



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

FACULTAD DE BELLAS ARTES DE SAN CARLOS
DEPARTAMENTO DE PINTURA

Programa de Doctorado:
ARTES VISUALES E INTERMEDIA

Tesis doctoral:

**ENTORNOS MULTIMEDIA DE REALIDAD AUMENTADA
EN EL CAMPO DEL ARTE**

Presentada por:

CRISTINA PORTALÉS RICART

Dirigida por:

Dra. MARÍA JOSÉ MARTÍNEZ DE PISÓN

Dr. JOSÉ LUIS LERMA GARCÍA

MAYO 2008

RESUMEN DE LA TESIS

La relación entre Ciencia y Arte ha mantenido a lo largo de la historia momentos de proximidad o distanciamiento, llegando a entenderse como dos culturas diferentes¹, pero también se han producido situaciones interdisciplinarias de colaboración e intercambio que en nuestros días mantienen como nexo común la cultura digital y el uso del ordenador. Según (Berenguer, 2002) desde la aparición del ordenador, científicos y artistas están encontrando un espacio común de trabajo y entendimiento. Mediante el empleo de las nuevas tecnologías, la distancia que separa ambas disciplinas es cada vez más corta. En esta tesis, cuyo título es "entornos multimedia de realidad aumentada en el campo del arte", se presenta una investigación teórico-práctica de la tecnología de realidad aumentada aplicada al arte y campos afines, como el edutainment. La investigación se ha realizado en dos bloques: en el primer bloque se trata la tecnología desde distintos factores que se han considerado relevantes para su entendimiento y funcionamiento; en el segundo se presentan un total de seis ensayos que constituyen la parte práctica de esta tesis.

El primer capítulo del primer bloque empieza por una definición de la tecnología de realidad aumentada (RA) y tecnologías afines, a partir de las aportadas por varios autores en periodos distintos; se puede ver como el término ha evolucionado para abarcar un campo mucho más extenso, en el que se incluyen dispositivos

¹ En 1956 C. P. Snow ya señaló el problema en su artículo "The Two Cultures" publicado en la revista inglesa *New Statesman*.

multimodales; se adoptan una serie de características que definen esta tecnología, a partir de las cuales se sientan las bases del concepto considerado en esta tesis. También se presenta una evolución histórica desde principios del siglo XX de tecnologías cercanas que han llevado a la actualidad de la RA, así como diversos ejemplos de aplicaciones y desarrollos llevados a cabo por investigadores en distintos campos, a parte del arte, como en educación, entretenimiento, ingeniería o medicina, con la finalidad de describir la utilidad y aportar una visión general de esta tecnología novedosa. En siguientes capítulos se habla de las relaciones espaciales entre el entorno físico 3D y la imagen capturada por la cámara en tiempo real, así como de las transformaciones que relacionan dicho entorno con su homólogo virtual, para construir una escena aumentada que sea consistente, es decir, verosímil. También se ha hecho una revisión bibliográfica de los dispositivos fundamentales en esta tecnología: los sensores y los displays. Para ello se ha mostrado una clasificación de los mismos, aportando diversos ejemplos tanto de desarrollos como de aplicaciones de RA que se valen de ellos. Se ha considerado también relevante tratar los distintos programas que existen en la actualidad para la generación de escenas de realidad aumentada, centrando la atención en aquellos basados en el reconocimiento de patrones de marcas planas, por sus ventajas y su uso extendido. Finalmente, en este bloque se ha tratado la interacción y experiencia del usuario final en entornos de RA.

Los ensayos aportados en el segundo de los bloques tecnológicamente se enmarcan dentro de los capítulos aportados en el primer bloque, y conceptualmente tratan diversos aspectos bajo un punto de vista cercano al campo del arte, edutainment o entretenimiento. Así pues, en el primero de los ensayos, bajo el título de *Usuario Aumentado*, se invierte la disposición espacial típica de los sistemas de RA y se utiliza la metáfora del espejo mágico para provocar una transformación de la imagen percibida de uno mismo, que cuestiona la propia identidad; técnicamente se utilizan las librerías de ARToolKit para superponer (en la imagen) una serie de modelos 3D animados a unas marcas que el usuario lleva puestas a modo de vestimenta. En el segundo de los ensayos, *UrbanMix*, se realiza una aplicación propia de RA para espacios abiertos en la que se entremezclan dos espacios urbanos distintos; como caso de estudio, se presenta la integración de dos monumentos históricos del casco de Valencia en la Universidad Politécnica de Valencia; tecnológicamente, se presenta un registro del usuario mediante la combinación de un sensor inercial y técnicas de fotogrametría o visión por computador; además, también se resuelve el problema de la oclusión. En el tercer ensayo, *Hot Images*, se propone una nueva cartografía de un

entorno urbano, mediante la adición de connotaciones humanas a los edificios inertes; se propone un sistema de navegación para viandantes novedoso, mediante una gradación de colores que abarcan desde el rojo (significando cercanía, caliente) hasta el azul (lejanía, frío), idea aportada a partir del popular juego infantil de caliente/frío; técnicamente la localización del usuario se realiza mediante la combinación de un receptor GPS y un sensor inercial. En el cuarto de los ensayos, *Live LEGO House*, se propone un juego para niños mediante una casita "aumentada" hecha con los bloques de LEGO, en la que conviven dos muñecos; se utiliza la metáfora del espejo mágico y se establece un juego de interacciones con los distintos elementos de la casa, haciendo que se ejecuten diversos sonidos y elementos gráficos; mediante una serie de condicionales establecidos en el entorno de programación, se premia la "buena conducta" de los muñecos, fomentando así pues la convivencia entre ellos; además, se introducen diversos conceptos multiculturales mediante los elementos generados por el ordenador. En el quinto ensayo, *CatedrAR*, se hace uso de un programa de detección de rasgos naturales para aumentar dos elementos de la Catedral de Valencia; de este modo, el interior de la bóveda renacentista situada sobre el altar mayor, se aumenta con la imagen de su homóloga barroca, hoy en día desmantelada; de otro modo, el retablo del altar mayor se aumenta con un retablo de plata desaparecido, que data del siglo XV. En el último de los ensayos, *AR-Jazz*, se crea un entorno de RA para visualizar el sonido y el movimiento de diversos músicos interpretando piezas de Jazz sobre un escenario, en una *Jam session*, precedido por una fase de aprendizaje mediante un taller.

Finalmente, se presentan unas conclusiones, tanto del primer bloque como del segundo, en la que no sólo se establecen criterios a partir de cada uno de los temas tratados, sino que además, se comparan los distintos ensayos realizados para evaluar de forma crítica los resultados obtenidos, tanto de la adecuación de los dispositivos utilizados como del éxito alcanzado a nivel conceptual.

RESUM DE LA TESI

La relació entre la ciència i l'art ha tingut al llarg de la història moments de proximitat o distanciament, i s'ha arribant a entendre-les com dues cultures diferents², però també s'han produït situacions interdisciplinàries de col·laboració i intercanvi que en els nostres dies mantenen com a nexes comuns la cultura digital i l'ús de l'ordinador. Segons (Berenguer, 2002), des de l'aparició de l'ordinador, científics i artistes han trobat un espai comú de treball i d'entesa. Mitjançant l'ús de les noves tecnologies, la distància que separa ambdues disciplines és cada vegada més curta. En aquesta tesi, el títol de la qual és *Entorns multimèdia de realitat augmentada en el camp de l'art*, es presenta una investigació teoricopràctica de la tecnologia de realitat augmentada aplicada a l'art i a camps afins a aquest, com ara l'entreteniment educatiu (*edutainment*). La investigació s'ha realitzat en dos blocs: en el primer es tracta de la tecnologia des de diversos punts de vista que s'han considerat rellevants per a entendre-la i per al mateix funcionament d'aquesta, i en el segon es presenta un total de sis assajos que constitueixen la part pràctica d'aquesta tesi.

El primer capítol del primer bloc comença amb una definició de la tecnologia de realitat augmentada (RA) i les tecnologies afins, a partir de les aportades per diversos autors en períodes diferents; es pot veure com el terme ha evolucionat per a abraçar un camp molt més extens, en el qual s'inclouen dispositius multimodals; s'enumera un seguit

² En 1956 C. P. Snow ja va assenyalar el problema en el seu article "The Two Cultures" publicat en la revista anglesa *New Statesman*.

de característiques que defineixen aquesta tecnologia, a partir de les quals es posen les bases del concepte considerat en aquesta tesi. També es presenta una evolució històrica des de principis del segle XX de les tecnologies pròximes que han portat a l'actualitat de la RA, així com diversos exemples d'aplicacions i desenvolupaments portats a terme per investigadors en diferents camps, a banda de l'art, com són l'educació, l'entreteniment, l'enginyeria o la medicina, amb la finalitat de descriure la utilitat d'aquesta tecnologia nova i aportar-ne una visió general. En els capítols següents es parla de les relacions espacials entre l'entorn físic 3D i la imatge capturada per la càmera en temps real, així com de les transformacions que relacionen aquest entorn amb l'homòleg virtual, elements que serveixen per a construir una escena augmentada que siga consistent, és a dir, versemblant. També s'ha fet una revisió bibliogràfica dels dispositius fonamentals en aquesta tecnologia: els sensors i els *displays*. Amb aquesta finalitat es mostra una classificació dels dispositius esmentats, aportant diversos exemples tant de desenvolupaments com d'aplicacions de RA que en fan ús. S'ha considerat també rellevant referir-se als diversos programes que hi ha en l'actualitat per a la generació d'escenes de realitat augmentada, centrant l'atenció en aquells basats en el reconeixement de marques fiducials, pels avantatges i l'ús tan estès que presenten. Finalment, en aquest bloc s'ha tractat de la interacció i l'experiència de l'usuari final en entorns de RA.

Els assajos aportats en el segon dels blocs s'emmarquen, quant a la tecnologia, dins dels capítols del primer bloc, i conceptualment tracten de diversos aspectes sota un punt de vista pròxim al camp de l'art, l'*edutainment* o l'entreteniment. Així, doncs, en el primer dels assajos, sota el títol d'"*Usuario Aumentado*", s'inverteix la disposició espacial típica dels sistemes de RA i s'utilitza la metàfora de l'espill màgic per a provocar una transformació de la imatge percebuda d'un mateix, que qüestiona la pròpia identitat; tècnicament s'utilitzen les llibreries d'ARToolKit per a superposar (en la imatge) una sèrie de models 3D animats a unes marques que l'usuari porta posades a manera de vestimenta. En el segon dels assajos, "*UrbanMix*", es realitza una aplicació pròpia de RA per a espais oberts, en la qual s'entremesclen dos espais urbans distints; com a cas d'estudi, es presenta la integració de dos monuments històrics del nucli urbà de València a la Universitat Politècnica; tecnològicament, es presenta un registre de l'usuari mitjançant la combinació d'un sensor inercial i tècniques de fotogrametria o visió per computador; a més, també es resol el problema de l'oclusió. En el tercer assaig, "*Hot Images*", es proposa una nova cartografia d'un entorn urbà, mitjançant l'addició de connotacions humanes als edificis inerts; es proposa un nou sistema de

navegació per a vianants, mitjançant una gradació de colors que comprenen des del roig (significant proximitat, calor) fins al blau (llunyania, fred), idea derivada del popular joc infantil de calent/fred; tècnicament la localització de l'usuari es realitza mitjançant la combinació d'un receptor GPS i un sensor inercial. En el quart dels assajos, "Live LEGO House", es proposa un joc per a xiquets mitjançant una caseta *augmentada* feta amb els blocs de LEGO, en la qual conviuen dos ninotets; s'utilitza la metàfora de l'espill màgic i s'estableix un joc d'interaccions amb els diferents elements de la casa, fent que s'executen diversos sons i elements gràfics; mitjançant una sèrie de condicions establides en l'entorn de programació, es premia la *bona conducta* dels ninotets, i així es fomenta la convivència entre ells; a més, s'introdueixen diversos conceptes multiculturals mitjançant els elements generats per l'ordinador. En el cinquè assaig, "CatedrAR", es fa ús d'un programa de detecció de trets naturals per a augmentar dos elements de la Seu de València; d'aquesta manera, l'interior de la volta renaixentista situada sobre l'altar major s'augmenta amb la imatge de l'homòloga barroca d'aquesta, avui dia desmantellada; a més, el retaule de l'altar major s'augmenta amb un retaule d'argent desaparegut, que datava del segle XV. En l'últim dels assajos, "AR-Jazz", es crea un entorn de RA per a visualitzar el so i el moviment de diversos músics que interpreten peces de Jazz sobre un escenari, en una *Jam session* precedida per una fase d'aprenentatge mitjançant un taller.

Finalment, es presenten unes conclusions, tant del primer bloc com del segon, en les quals no sols s'estableixen criteris a partir de cadascun dels temes tractats, sinó que, a més, es comparen els diversos assajos realitzats per a avaluar-ne de manera crítica els resultats, pel que fa a l'adequació dels dispositius utilitzats i, també, a l'èxit aconseguit a nivell conceptual.

SUMMARY OF THE DOCTORAL THESIS

The relationship between Science and Arts has kept during the course of the history moments of proximity or distancing, arriving to be understood as two different cultures³, but also interdisciplinary situations of collaboration and interchange have occurred that nowadays keep as common nexus the digital culture and the use of computers. According to (Berenguer, 2002), since the appearance of computers, scientifics and artists have found a common space of work and understanding. Within the use of new technologies, the distance that separates both disciplines is each time shorter. In this doctoral thesis, whose title is "augmented reality multimedia environments in the field of arts", a theoretical-practical of research of augmented reality technology is presented, applied to the field of arts and related areas, such as edutainment. The research has been carried out in two blocks: in the first block the technology is reviewed according to different factors that have been considered to be relevant for their understanding and functioning; in the second block a total of six rehearsals are presented, which form the practical part of this doctoral thesis.

The first chapter of first block begins within a definition of augmented reality (AR) technology and related technologies, from the ones given by different authors in different periods; it can be seen how the term has evolved to cover a much wider field, in which multimodal devices are included; several characteristics defining this

³ In 1956 C. P. Snow already pointed the problem in his article "The Two Cultures", published in the English magazine *New Statesman*.

technology are taken, from which the concept considered in this doctoral thesis is derived. A historical evolution is also presented from the beginnings of XX century of close technologies that have lead to the actuality of AR, as well as several examples of applications and developments fulfilled by researchers in different areas, apart from arts, as education, entertainment, engineering or medicine, with the aim to describe the utility and give a general vision of this novel technology. In subsequent chapters the spatial relation between the physical 3D environment and the image captured by the camera in real time is described, as well as the transformations that relate that environment with its homologous virtual, to build an augmented scene that is consistent, i.e., plausible. A bibliographic review has also been carried out regarding to fundamental devices of this technology: sensors and displays. To that purpose a classification has been done, bringing diverse examples of developments as well as AR applications that have them integrated. It has been considered also to be relevant to review the different software that exists nowadays for the generation of augmented reality scenes, centring the attention in those based on the recognition of fiducial markers, due to their advantages and extended use. Finally, in this block the final user interaction and experience in AR environments has been reviewed.

The rehearsals at the second block, technologically are within the chapters in the first block, and conceptually treat different aspects upon a point of view close to arts, edutainment or entertainment. Therefore, in the first of the rehearsals, under the title *Usuario Aumentado*, the typical spatial relationship of AR systems is inverted and the magic mirror metaphor is used to cause a transformation of the perceived own image, that questions the own identity; technically the ARToolkit libraries are used to overlap (at the image) a set of animated 3D models to a markers that the user wears. In the second of the rehearsals, *UrbanMix*, an own outdoors AR application is achieved in which different urban spaces are mixed; as case study, the integration of two historical monuments of the centre of Valencia into the University of Valencia is presented; technologically, a tracking of the user is presented via the combination of an inertial sensor and photogrammetric or computer vision techniques; furthermore, the occlusion problem is also solved. In the third rehearsal, *Hot Images*, a new cartography of an urban environment is proposed, by means of addition of human connotations to the inert buildings; a novel pedestrian navigation system is proposed, by means of colour gradations that go from red (meaning close, hot) to blue (remoteness, cold), idea taken from the popular childish game of hot/cold; technically the location of the user is achieved by the combination of a GPS receiver and an inertial sensor. In the forth of the

rehearsals, *Live LEGO House*, a game for children is proposed by means of an "augmented" house made with the LEGO blocks, in which two dolls live together; the magic mirror metaphor is used and a game of interactions is established within the different elements of the house, executing diverse sound and graphic elements; by means of a set of conditions established within the programming environment, "good behaviour" of the dolls is rewarded, promoting therefore the coexistence between them; furthermore, several multicultural concepts are introduced by computer generated objects. In the fifth rehearsal, *CatedrAR*, a natural feature tracking based program is used to augment two elements of the Cathedral of Valencia; in this way, the interior of the renaissance vault located over the major altar, is augmented with the image of its homologous baroque, nowadays dismantled; on the other hand, the reredos of the major altar is augmented with a disappeared silver reredos, dating from sigle XV. In the last of the rehearsals, *AR-Jazz*, an AR environment is recreated to visualize sound and movement of several musicians interpreting Jazz fragments on a stage, in a *Jam session*, preceded by a learning phase by means of a workshop.

Finally, some conclusions are presented, so much of the first block as of the second one, in which not only criteria are established from each of the items considered, but also the different presented rehearsals are compared in order to critically evaluate the obtained results, so much as the adequacy of the used devices as the success reached regarding to the conceptual level.

AGRADECIMIENTOS

En estas primeras líneas quisiera transmitir mi gratitud a todas aquellas personas que de algún modo han contribuido directa o indirectamente a la realización de esta tesis doctoral. El trabajo descrito en esta memoria es fruto de una estrecha relación entre distintos campos del conocimiento (arte, ciencia, ingeniería, informática, etc.), por lo que en algunos casos se ha trabajado en colaboración con distintos investigadores y profesionales pertenecientes a ellos.

En primer lugar quisiera mostrar mis agradecimientos hacia mis directores de tesis, los doctores María José Martínez de Pisón y José Luis Lerma. A María José por su paciencia y dedicación acortando las distancias que separan los mundos científico y artístico. También por proporcionarme la libertad de dirigir mi investigación hacia aquellos aspectos que he considerado más relevantes y por introducirme en algunos otros relacionados con el mundo del Arte. A José Luis por sus buenos consejos, apoyo constante y entusiasmo desde mis comienzos en los estudios universitarios. Y a ambos por aportar su conocimiento y amplia experiencia. A continuación quisiera agradecer la inestimable ayuda del doctor Francisco Giner, junto con el que he realizado dos de los ensayos de esta tesis y difundido los resultados en diversos congresos. También al doctor Francisco Sanmartín, que de igual modo comparte la autoría de uno de los ensayos.

De igual modo quisiera dar las gracias a todos los integrantes de Laboratorio de Luz, los cuales me acogieron muy gratamente desde un principio. Los cuatro años que he pasado junto a ellos han sido personalmente muy enriquecedores.

A todos y cada uno de los miembros e investigadores del Mixed Reality Laboratory de la Universidad de Nottingham con los que coincidí durante mi visita, especialmente a Macarena Gómez, Bartosz Wietrzyk, Leif Oppermann y Mauricio Capra, por prestarme su ayuda desinteresada en algunos momentos. También a Christopher Hide del Institute of Engineering Surveying and Space Geodesy, por su ayuda en lo referente a sensores inerciales. De igual modo quisiera mostrar mis agradecimientos a todos los integrantes e investigadores del Mixed Reality Lab, actualmente perteneciente a la National University of Singapore, que han coincidido durante mi estancia allí. En especial a Liu Wei y a Johannes Wagner por prestarme su ayuda y consejos. También a su director Adrian David Cheok por permitirme formar parte activa de su laboratorio y recibirme con entusiasmo.

De igual modo, quisiera agradecer la ayuda prestada en cuestiones de programación a Vincent Lepetit, uno de los desarrolladores de BazAR. También a Oliver Bimber, por difundir gratuitamente los resultados de su valiosa investigación en el campo de la realidad aumentada espacial. Y en general, a todos aquellos investigadores y profesionales que comparten, a través de la red y de forma desinteresada, sus conocimientos y desarrollos.

Muy especialmente quisiera agradecer la ayuda, paciencia y constante apoyo de la persona con la que comparto mi vida, Carlos David Perales, compositor y director con el que también comparto la autoría de uno de los ensayos.

Esta tesis ha sido realizada gracias al programa doctoral de Formación de Personal Investigador del Ministerio de Ciencia y Tecnología, bajo el proyecto "Percepción ampliada: un sistema dual de realidad aumentada" (BHA2003-02720).

ÍNDICE

RESUMEN DE LA TESIS.....	3
RESUM DE LA TESI.....	7
SUMMARY OF THE DOCTORAL THESIS	11
AGRADECIMIENTOS	15
ÍNDICE	17
LISTA DE FIGURAS, TABLAS Y ECUACIONES	25
INTRODUCCIÓN.....	39
BLOQUE I: ASPECTOS TEÓRICOS.....	55
1. MARCO GENERAL DE REALIDAD AUMENTADA	57
1.1. DEFINICIONES	58
1.1.1. REALIDAD MEZCLADA Y EL CONTÍNUO DE MILGRAM	58
1.1.2. REALIDAD VIRTUAL	59
1.1.3. VIRTUALIDAD AUMENTADA.....	60
1.1.4. REALIDAD AUMENTADA	61
1.1.5. REALIDAD MEDIATIZADA	63
1.2. ASPECTOS GENERALES.....	64
1.2.1. CARACTERÍSTICAS DE LA REALIDAD AUMENTADA.....	64
1.2.2. CLASIFICACIÓN DE LOS SISTEMAS DE RA	67
1.3. RECORRIDO HISTÓRICO.....	70
1.4. DESARROLLOS Y APLICACIONES	78
1.4.1. ARTE.....	78
1.4.2. ENTRETENIMIENTO.....	82
1.4.3. EDUCACIÓN	84

1.4.4.	EDUTAINMENT	86
1.4.5.	MEDICINA.....	88
1.4.6.	PSICOLOGÍA	89
1.4.7.	INGENIERÍA Y ARQUITECTURA.....	90
1.4.8.	ROBÓTICA Y TELERROBÓTICA.....	91
1.4.9.	INDUSTRIA	93
1.4.10.	MILITAR	94
2.	PRINCIPIOS MATEMÁTICOS.....	97
2.1.	RELACIONES ESPACIALES ENTRE EL ENTORNO 3D Y LA IMAGEN 2D	98
2.1.1.	SISTEMAS DE COORDENADAS	98
2.1.2.	PARÁMETROS DE LA ORIENTACIÓN EXTERNA	101
2.1.3.	ECUACIONES DE LA PROYECCIÓN CENTRAL	102
2.1.4.	ESTÉREO-VISIÓN: GEOMETRÍA EPIPOLAR.....	104
2.1.5.	ORIENTACIÓN ESPACIAL	106
2.1.5.1.	MÉTODOS DIRECTOS	106
2.1.5.1.1.	Posición a partir de un receptor GPS	107
2.1.5.1.2.	Rotaciones a partir de un sensor INS	109
2.1.5.2.	MÉTODOS INDIRECTOS	110
2.1.5.2.1.	Calibración de la cámara	111
2.1.5.3.	MÉTODOS MIXTOS	113
2.1.5.3.1.	INS y cámara	113
2.2.	RELACIONES ESPACIALES ENTRE LOS ENTORNOS VIRTUAL Y REAL	113
2.3.	RENDERIZACIÓN EN SISTEMAS DE RA	115
2.3.1.	OCCLUSIÓN	116
2.3.2.	SOMBREADO.....	118
2.4.	CARACTERÍSTICAS DE SENSORES Y DISPLAYS.....	120
3.	ENTRADAS: SENSORES	121
3.1.	SENSORES MECÁNICOS	122
3.1.1.	GONIÓMETROS.....	122
3.1.2.	ENLACES DE SUELO	125
3.2.	SENSORES MAGNÉTICOS	126
3.3.	SENSORES ÓPTICOS	128
3.3.1.	SISTEMAS DE ESTÉREO-VISIÓN PASIVOS	129
3.3.2.	SISTEMAS DE MARCAS ACTIVAS	130
3.3.3.	SISTEMAS DE PATRONES DE MARCAS PLANAS	132
3.3.4.	SISTEMAS DE REGISTRO DE RASGOS NATURALES	134
3.3.5.	RADAR LÁSER	136
3.3.6.	SENSORES OCULARES	137
3.4.	SENSORES ACÚSTICOS	139

3.5.	SENSORES INERCIALES	141
3.6.	SENSORES DE RADIOFRECUENCIA.....	143
3.6.1.	GPS.....	143
3.6.2.	RFID.....	144
4.	SALIDAS: DISPLAYS	147
4.1.	DISPLAYS VISUALES.....	148
4.1.1.	DISPLAYS PARA LA CABEZA	149
4.1.1.1.	DISPLAYS RETINIANOS	149
4.1.1.2.	DISPLAYS SUJETOS A LA CABEZA.....	151
4.1.1.3.	PROYECTORES SUJETOS A LA CABEZA	154
4.1.2.	DISPLAYS DE MANO	158
4.1.3.	DISPLAYS ESPACIALES	161
4.1.3.1.	DISPLAYS ESPACIALES DE VÍDEO	161
4.1.3.2.	DISPLAYS ESPACIALES ÓPTICOS	162
4.1.3.3.	DISPLAYS ESPACIALES DE PROYECTORES.....	165
4.2.	DISPLAYS SONOROS	167
4.3.	DISPLAYS HÁPTICOS.....	170
4.4.	OTROS DISPLAYS	171
4.5.	SISTEMAS MULTIMODALES.....	173
5.	SOFTWARE	177
5.1.	SW BASADO EN PATRONES DE MARCAS PLANAS.....	178
5.1.1.	ARTOOLKIT.....	178
5.1.1.1.	PRINCIPIOS BÁSICOS.....	178
5.1.1.2.	CALIBRACIÓN DE LA CÁMARA	179
5.1.1.3.	DEFINICIÓN DE LAS MARCAS Y REGISTRO	180
5.1.1.4.	LIMITACIONES.....	182
5.1.2.	MXRTOOLKIT	183
5.1.2.1.	PRINCIPIOS BÁSICOS.....	184
5.1.2.2.	CALIBRACIÓN DE LA CÁMARA	185
5.1.2.3.	DEFINICIÓN DE LAS MARCAS Y REGISTRO	186
5.1.2.4.	LIMITACIONES.....	187
5.1.3.	STUDIERSTUBE.....	188
5.1.4.	ARTAG.....	188
5.1.5.	OSGAR.....	189
5.1.6.	OSGART	190
5.1.7.	AMIRE.....	191
5.2.	SW DE RECONOCIMIENTO DE RASGOS NATURALES	192
5.2.1.	BAZAR.....	192
5.2.1.1.	PRINCIPIOS BÁSICOS.....	193

5.2.1.2.	CALIBRACIÓN DE LA CÁMARA	193
5.2.1.3.	PROCESO DE REGISTRO.....	195
5.2.1.4.	LIMITACIONES.....	196
5.3.	SW MULTITAREA	196
5.3.1.	MAX/MSP JITTER.....	197
5.3.1.1.	REGISTRO DE ELEMENTOS EN LA IMAGEN.....	197
5.3.1.2.	OBJETOS 3D Y CÁMARAS VIRTUALES.....	198
5.3.1.3.	INTEGRACIÓN DE SENSORES.....	201
5.3.1.4.	LIMITACIONES.....	202
6.	EL USUARIO FINAL.....	203
6.1.	INTERACTUANDO CON EL ENTORNO AUMENTADO	205
6.1.1.	FORMAS DE INTERACTUAR	206
6.1.1.1.	MANIPULACIÓN DE OBJETOS VIRTUALES / REALES	206
6.1.1.2.	NAVEGACIÓN	209
6.1.1.3.	INTERACCIÓN CON OTROS USUARIOS.....	210
6.1.1.3.1.	Colaboración cara a cara/presencial	211
6.1.1.3.2.	Colaboración remota.....	212
6.1.1.3.3.	Entornos aumentados colaborativos	213
6.1.2.	INTERACCIÓN SIN COSTURAS	215
6.1.3.	VESTIDOS COMPUTACIONALES.....	217
6.2.	LA EXPERIENCIA EN EL ENTORNO AUMENTADO.....	221
6.2.1.	PERCEPCIÓN HUMANA – PERCEPCIÓN AMPLIADA.....	221
6.2.2.	INMERSIÓN	222
6.2.3.	COMPATIBILIDAD	224
6.2.4.	ENTRETENIMIENTO EN ENTORNOS DE RA.....	226
6.2.5.	APRENDIZAJE EN ENTORNOS DE RA	228
6.2.5.1.	APRENDER HACIENDO.....	230
6.2.5.2.	APRENDER JUGANDO: EDUTAINMENT.....	231
6.2.6.	EL COMPROMISO ARTÍSTICO EN ENTORNOS DE RA.....	233
	BLOQUE II: ENSAYOS	235
7.	USUARIO AUMENTADO.....	237
7.1.	PLANTEAMIENTO	239
7.1.1.	MARCO EN EL QUE SE INSCRIBE EL ENSAYO	241
7.1.2.	REFERENTES.....	243
7.2.	DESCRIPCIÓN TÉCNICA	246
7.2.1.	COMPONENTES DE LA APLICACIÓN	246
7.2.1.1.	DISPOSITIVOS FÍSICOS	246
7.2.1.2.	ELEMENTOS VIRTUALES.....	248
7.2.2.	DESARROLLO, PROCESO	249

7.2.3.	CONSIDERACIONES / RESTRICCIONES TÉCNICAS	251
7.3.	MEMORIA DE LAS DISTINTAS FASES.....	252
7.3.1.	PRIMERA FASE: PRUEBAS DE LABORATORIO	252
7.3.1.1.	DESCRIPCIÓN TÉCNICA.....	253
7.3.1.2.	MODELOS VIRTUALES	254
7.3.1.3.	INTERACCIÓN	255
7.3.2.	SEGUNDA FASE: CAMBIO DE PERSONALIDAD	255
7.3.2.1.	DESCRIPCIÓN TÉCNICA.....	255
7.3.2.2.	MODELOS VIRTUALES	257
7.3.2.3.	INTERACCIÓN	258
7.3.3.	TERCERA FASE: CARRUSEL DE IMÁGENES	258
7.3.3.1.	DESCRIPCIÓN TÉCNICA.....	259
7.3.3.2.	MODELOS VIRTUALES	259
7.3.3.3.	INTERACCIÓN	260
7.3.4.	CUARTA FASE: TÚNEL MÁGICO	260
7.3.4.1.	DESCRIPCIÓN TÉCNICA.....	261
7.3.4.2.	MODELOS VIRTUALES	263
7.3.4.3.	INTERACCIÓN	264
7.3.5.	QUINTA FASE: UN [∞]	264
7.3.5.1.	DESCRIPCIÓN TÉCNICA.....	267
7.3.5.2.	MODELOS VIRTUALES	269
7.3.5.3.	INTERACCIÓN	272
7.3.6.	SEXTA FASE: CINEATRO (CINE + TEATRO)	273
7.3.6.1.	DESCRIPCIÓN TÉCNICA.....	273
7.3.6.2.	MODELOS VIRTUALES	275
7.3.6.3.	INTERACCIÓN	276
7.4.	SISTEMAS DE VERIFICACIÓN	277
8.	URBANMIX.....	283
8.1.	PLANTEAMIENTO	285
8.1.1.	MARCO EN EL QUE SE INSCRIBE EL ENSAYO	286
8.1.2.	REFERENTES.....	287
8.2.	DESCRIPCIÓN TÉCNICA	288
8.2.1.	COMPONENTES DE LA APLICACIÓN	288
8.2.1.1.	DISPOSITIVOS FÍSICOS	288
8.2.1.2.	ELEMENTOS VIRTUALES	290
8.2.2.	DESARROLLO, PROCESO.....	292
8.2.2.1.	SISTEMA DE REFERENCIA.....	292
8.2.2.2.	CONSIDERACIÓN DE LA OCLUSIÓN.....	294
8.2.2.3.	ENTORNO DE PROGRAMACIÓN	295

8.2.3.	CONSIDERACIONES / RESTRICCIONES TÉCNICAS	299
9.	HOT IMAGES.....	301
9.1.	PLANTEAMIENTO	303
9.1.1.	MARCO EN EL QUE SE INSCRIBE EL ENSAYO	305
9.1.2.	REFERENTES.....	307
9.2.	DESCRIPCIÓN TÉCNICA	308
9.2.1.	COMPONENTES DE LA APLICACIÓN.....	308
9.2.1.1.	DISPOSITIVOS FÍSICOS	308
9.2.1.2.	ELEMENTOS VIRTUALES.....	309
9.2.2.	DESARROLLO, PROCESO	310
9.2.2.1.	INTERFAZ DE NAVEGACIÓN.....	310
9.2.2.2.	SISTEMA DE REFERENCIA.....	311
9.2.2.3.	ENTORNO DE PROGRAMACIÓN	312
9.2.2.4.	STORYBOARD	318
9.2.3.	CONSIDERACIONES / RESTRICCIONES TÉCNICAS	321
9.3.	SISTEMAS DE VERIFICACIÓN.....	322
10.	LIVE LEGO HOUSE	325
10.1.	PLANTEAMIENTO	327
10.1.1.	MARCO EN EL QUE SE INSCRIBE EL ENSAYO	328
10.1.2.	REFERENTES	330
10.2.	DESCRIPCIÓN TÉCNICA	332
10.2.1.	COMPONENTES DE LA APLICACIÓN.....	332
10.2.1.1.	DISPOSITIVOS FÍSICOS	332
10.2.1.2.	ELEMENTOS VIRTUALES.....	333
10.2.2.	DESARROLLO, PROCESO	334
10.2.2.1.	PROCESO DE REDUCCIÓN DE MARCAS.....	334
10.2.2.2.	INTERFAZ, INTERACCIÓN Y MÁQUINA DE ESTADOS FINITOS.....	336
10.2.2.3.	DESARROLLO EN EL SOFTWARE UTILIZADO	340
10.2.3.	FASES DEL JUEGO	345
10.2.3.1.	CONSTRUIR LA CASA CON LOS BLOQUES DE LEGO	345
10.2.3.2.	PROCESO DE CALIBRACIÓN DE LAS MARCAS.....	346
10.2.3.3.	EXPLORAR LOS ASPECTOS DE CONVIVENCIA Y MULTICULTURALISMO	347
10.2.4.	CONSIDERACIONES / RESTRICCIONES TÉCNICAS	349
10.3.	MEMORIA DE LAS DISTINTAS FASES.....	351
10.3.1.	PRIMERA FASE: LLH V.1	351
10.3.1.1.	DESCRIPCIÓN TÉCNICA.....	351
10.3.2.	SEGUNDA FASE: LLH V.2.....	353
10.3.2.1.	DESCRIPCIÓN TÉCNICA.....	353
10.4.	SISTEMAS DE VERIFICACIÓN.....	355

11. CATEDRAR.....	359
11.1. PLANTEAMIENTO	361
11.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO	363
11.1.2. REFERENTES	364
11.2. DESCRIPCIÓN TÉCNICA	366
11.2.1. COMPONENTES DE LA APLICACIÓN.....	366
11.2.1.1. DISPOSITIVOS FÍSICOS	366
11.2.1.2. ELEMENTOS VIRTUALES	367
11.2.2. DESARROLLO, PROCESO	368
11.2.2.1. IMPLEMENTACIÓN DE IMÁGENES CON CANAL ALFA	368
11.2.2.2. CASOS DE ESTUDIO	371
11.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS	374
12. AR-JAZZ.....	377
12.1. PLANTEAMIENTO	379
12.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO	380
12.1.2. REFERENTES	381
12.2. DESCRIPCIÓN TÉCNICA	383
12.2.1. COMPONENTES DE LA APLICACIÓN.....	383
12.2.1.1. DISPOSITIVOS FÍSICOS	383
12.2.1.2. ELEMENTOS VIRTUALES	384
12.2.2. DESARROLLO, PROCESO	385
12.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS	388
12.3. SISTEMAS DE VERIFICACIÓN	389
12.3.1. TALLER.....	389
12.3.2. REPRESENTACIÓN EN VIVO	390
CONCLUSIONES.....	397
BIBLIOGRAFÍA	407
ANEXOS.....	429

LISTA DE FIGURAS, TABLAS Y ECUACIONES

Figuras

Figura 1. Pirámide de las etapas de desarrollo de la tecnología de RA. Elaborada a partir de (Bimber et al., 2005d, 6).	46
Figura 2. Continuo de Milgram. A partir de (Milgram et al., 1994).....	58
Figura 3. Continuo de Milgram ampliado con otros conceptos.	59
Figura 4. Imágenes del entorno virtual The 21st Century Virtual Reality Color Organ (2000) de Jack Ox, en (Ox et al., 2000).....	60
Figura 5. Ejemplo de un entorno de VA, en (Simsarian et al., 1997).....	61
Figura 6. Ejemplo de entorno de RA mediante las librerías BazAR, en (CVLAB - Computer Vision Laboratory, 2007).	63
Figura 7. Aplicación de realidad mediatizada. En ambas imágenes la cafetera es el objeto virtual. Los filtros estilo cartoon se han aplicado a la imagen de la derecha. En (Fischer et al., 2005).....	64
Figura 8. Diversas escenas cinematográficas: a) The Lord of the Rings: The Return of the King (2003); b) Fantastic Four: Rise of the Silver Surfer (2007); c) Spierman 3 (2007).....	65
Figura 9. Ejemplo de un sistema RA de videoconferencia en el que los datos están espacializados. En (Kato et al., 2001).	67
Figura 10. Sistema de periscopio, en (Sherman et al., 2003, 24).	70
Figura 11. Cinerama: a) Dibujo esquemático, en (Norwood, 1997); b) Cinerama Dome en Los Angeles, en (Coate, 2007); c) Sistema con cinco proyectores, en (Hart, 1995).	71
Figura 12. a) y b) Sensorama; c) Sensorama 3-D Motion Picture Camera. En (Heilig, 2007; Kunsthochschule für Medien Köln, 2007).....	72

Figura 13. Sistema HMD binocular, en (Ellis, 2007).	72
Figura 14. Sutherland interactuando con Sketchpad. En (Mischitz, 2001; Chavarro García, 2007).	73
Figura 15. HMD estereoscópico de Sutherland, en (Update Centre, 2007).....	73
Figura 16. Dispositivos del sistema CAVE, en (Cruz-Neira et al., 2002).	74
Figura 17. The Virtual Showcase, en (Bimber et al., 2001b).....	76
Figura 18. Ejemplos de utilización y GUI del AMIRE.	77
Figura 19. Mando de la consola Wii.....	78
Figura 20. Imágenes de un usuario interactuando con la pieza Can You See Me Now? (2001), en (Blast Theory, 2007).....	79
Figura 21. Usuarios interactuando con The Hidden Worlds of Noise and Voice. En (Levin et al., 2008).....	80
Figura 22. Imágenes de la pieza Red Libre Red Visible (2004) de Diego Díaz y Clara Boj, en (Díaz, 2007).....	80
Figura 23. Imágenes de usuarios interactuando con la pieza lifeClipper (2004), en (Torpus, 2005).	81
Figura 24. U-Tsu-Shi-O-Mi: The Virtual Humanoid You Can Reach (2006), de Michihiko Shoji, en (SIGGRAPH06, 2006).	81
Figura 25. Imágenes de Touch-Space (2002), en (Cheok et al., 2002b).	82
Figura 26. Imágenes de Human Pacman, en (Cheok et al., 2003).	83
Figura 27. Usuarios interactuando con The Invisible Train (2004), en (Wagner et al., 2004a).....	83
Figura 28. Usuarios interactuando con una mesa de tenis aumentada, en (Park et al., 2006).....	84
Figura 29. Usuarios jugando con PAC-LAN y visualizaciones en el móvil, en (Rashid et al., 2006). ..	84
Figura 30. Usuarios interactuando con Construct3D, en (Kaufmann, 2004).	85
Figura 31. Distintas visualizaciones moleculares. En (Gillet et al., 2004).	86
Figura 32. Usuarios interactuando con MagicBook, en (Technology Research News, 2000; HITLab, 2006).....	87
Figura 33. Aplicaciones de RA en el campo del edutainment: a) Solar System; b) y c) Plant Story. En (MXR, 2006).	87
Figura 34. Sistema de cámara-proyector y ejemplos de funcionamiento del sistema sobre un cráneo de plástico, en (Glossop et al., 2003).	88
Figura 35. Sistema de RA adaptado a un escáner médico, en (Wacker et al., 2005).	89
Figura 36. Tratamiento de fobias: a) Primer sistema de RA para tratar fobias a insectos; b) Sistema de RA sin marcas; c) Sistema de RA mediante HMD óptico.....	90
Figura 37. Sistemas para el tratamiento de vértigo: a) Usuario inmerso en un sistema CAVE; b) Sistema de RA.....	90
Figura 38. Usuarios interactuando con el sistema Luminous Planning Table, en (Ben-Joseph et al., 2001).....	91
Figura 39. Ejemplos del sistema ARVino, en (King et al., 2005).	91

Figura 40. Ejemplos de funcionamiento del sistema para robots, en (Stilman et al., 2005).	92
Figura 41. Robot y modelos virtual y real, en (Shaheen et al., 2001).....	92
Figura 42. Sistema de RA en procesos de manufactura, en (Dangelmaier et al., 2005).....	93
Figura 43. Dispositivo de RA, en (Behringer et al., 1999).	93
Figura 44. Visualizaciones de Eyekon, en (Hicks et al., 2002; Hicks et al., 2003).....	94
Figura 45. Visualizaciones del sistema BARS y usuario equipado, (Livingston et al., 2002).	95
Figura 46. Sistema de coordenadas imagen y sistema espacial de coordenadas imagen.	98
Figura 47. Sistemas de coordenadas: a) Terrestre global; b) Terrestre local; c) Centrado en objetos.	99
Figura 48. Esquema de la perspectiva central. Relación espacial entre un punto físico, su representación bidimensional en la imagen, y el centro de proyección de la cámara. Rotaciones de la cámara con respecto al sistema terrestre.	102
Figura 49. Rotaciones de un elemento virtual con respecto al sistema de referencia de la marca, y su correspondiente matriz de rotación; Rotaciones angulares (0, 0, 48).	104
Figura 50. Visión binocular.	104
Figura 51. Geometría epipolar. A partir de (Gruen et al., 2001, 201).	105
Figura 52. Datos NMEA de un receptor BTGPS II Trine de EMTAC.	107
Figura 53. Significado de la línea \$GPGGA.....	107
Figura 54. Superficies de referencia.....	108
Figura 55. Dos posiciones de la cámara para el cálculo de la distancia focal.....	112
Figura 56. Sistemas de coordenadas real/virtual en aplicaciones basadas en registro por sensores para un sistema de referencia terrestre y un registro directo. A partir de (Azuma, 2004).....	114
Figura 57. Sistemas de coordenadas real/virtual en aplicaciones basadas en registro por cámaras para un sistema de referencia terrestre y un registro indirecto. A partir de (Azuma, 2004).	115
Figura 58. Generación semi-automática de modelos 3D. En (Lepetit et al., 2000b; Lepetit et al., 2000a).	117
Figura 59. Detección de oclusiones en frente de fondos estáticos: a) Fondo estático oculto parcialmente por un objeto; b) Mapa de oclusiones (en blanco). En (Fischer et al., 2003).	118
Figura 60. Cambio del sombreado de un objeto virtual y de la sombra proyectada en el entorno real a medida que se éste se eleva, en (Gibson et al., 2003).	118
Figura 61. Sombreado: a) Imagen original; b) Imagen generada; c) Imagen generada con la inclusión de un objeto virtual. En (Wang et al., 2003).	119
Figura 62. Guantes: a) CyberGlove II de Immersion; b) software para CATIA; c) VirtualHand para MotionBuilder. En (Immersion Corporation, 2007b).	123
Figura 63. Sistema CyberGrasp de Inmersion. En (Immersion Corporation, 2007a).	124
Figura 64. Aplicación de RA con el uso del CyberGlove. En (Malkawi et al., 2005).....	124
Figura 65. Usuario con guantes de RA. En (Teichrieb et al., 2007).	125

Figura 66. CyberForce junto al CyberGrasp Exoskeleton de Inmersion. En (Immersion Corporation, 2007a).	125
Figura 67. Sistemas de realidad mezclada que emplean enlaces de suelo: a) Dispositivo para aplicaciones médicas; b) Display háptico. En (Inami et al., 2000).	126
Figura 68. FASTRAK de Polhemus. En (Polhemus, 2007a).	127
Figura 69. MiniBIRD de Ascension. En (Ascension Technology Corporation, 2007b).	128
Figura 70. Esquema y usuarios interactuando en un sistema que utiliza diversos sensores magnéticos. En (Kiyokawa et al., 2000b).	128
Figura 71. Cámaras estereoscópicas: a) Bumblebee; b) Digiclops. En (Point Grey, 2007).	129
Figura 72. Usuario con marcas retroreflectivas y sistema HMD binocular; imágenes izquierda y derecha e identificación de las marcas. En (Dorfmüller, 1999).	130
Figura 73. ReActor 2 de Ascension. En (Ascension Technology Corporation, 2007c).	131
Figura 74. Registro óptico-electrónico: a) HMD con dos videocámaras en miniatura y tres LEDs de infrarrojos; b) Mujer realizando una simulación de una biopsia con el sistema descrito. En (Rosenthal et al., 2002).	132
Figura 75. Ejemplo de realidad aumentada con sensor óptico y marcas de ARToolKit.	133
Figura 76. Imágenes del escenario Zygotes en el sistema Living Room. En (Torpus, 2004).	134
Figura 77. Reconocimiento de marcas proyectadas imperceptibles para el ojo humano. En (Bimber, 2007a).	134
Figura 78. Detección de puntos homólogos en imágenes contiguas y vectores de desplazamiento calculados. En (Prince et al., 2002).	135
Figura 79. Reconocimiento de puntos en superficies deformables. En (Salzmann et al., 2006).	136
Figura 80. LaserBIRD 2 de Ascension y esquema de funcionamiento. En (Ascension Technology Corporation, 2007a).	136
Figura 81. Sistema de láser para el registro del entorno 3D y posterior utilización en un sistema de realidad aumentada espacial. En (Bimber, 2007).	137
Figura 82. Sensor ocular: a) VisionTrak; b) VisionTrack Binocular Desktop 300; c) VisionTrak Raw Eye Movement Data Acquisition Software. En (Polhemus, 2007b).	138
Figura 83. Sistema RA de videoconferencia con registro de la mirada. En (Barakonyi et al., 2007).	139
Figura 84. 3D HeadTracker de Logitech. En (Logitech, 2007).	140
Figura 85. Usuarios interactuando con El Ball del Fanalet. En (Hoberman et al., 1998).	140
Figura 86. Sensor inercial MT9 de Xsens y su representación virtual. En (XSens, 2007).	142
Figura 87. Aplicación de RA en la que el usuario lleva GPS y un sensor inercial para el registro de su posición y orientación, y marcas para el reconocimiento de gestos de las manos. En (Piekarski, 2006).	142
Figura 88. Receptor BTGPS II Trine y antena externa activa, de EMTAC. En (EMTAC, 2007).	144
Figura 89. RD5000 Mobile RFID Reader de Motorota. En (Motorola, 2008).	145

Figura 90. Juego aumentado del Aprendiz de Brujo, con la utilización de RFID y fotorresistencias. En (Peitz et al., 2006).	146
Figura 91. Dispositivo móvil con RFID adherido. En (Reilly et al., 2006).	146
Figura 92. Clasificación de los displays visuales en función de la distancia al usuario. A partir de (Bimber et al., 2005d, 72).	148
Figura 93. Esquema de un display retiniano. A partir de (Bimber et al., 2005d, 73).	150
Figura 94. Display retinal de Microvision y aplicación de RA para ingeniería. En (Microvision, 2008).	150
Figura 95. Esquemas de los HMD: a) HMD de vídeo; b) HMD óptico.	151
Figura 96. HMD óptico que proporciona oclusión mutua gracias a unos paneles LCD. En (Kiyokawa et al., 2002).	152
Figura 97. a) Varioscope AR e instrumental quirúrgico equipados con un sistema óptico para el registro; b) Varioscope AR. En (Wanschitz et al., 2002).	153
Figura 98. HMD de vídeo utilizado en los ensayos de esta tesis: i-glasses PC/SVGA.	153
Figura 99. HMD de video con cámaras integradas: a) ARvision-3D Goggles; b) ARvision-3D HMD. En (Trivisio, 2008).	154
Figura 100. Esquema del funcionamiento de los proyectores sujetos a la cabeza. A partir de (Kiyokawa et al., 2000a).	154
Figura 101. Esquema del funcionamiento de un PHMD. A partir de (Kijima et al., 1997).	155
Figura 102. a) HMDP, en (Inami et al., 2000); b) visualización de la estructura ósea, en (Kawakami et al., 1998).	156
Figura 103. Esquema de camuflaje óptico mediante el empleo de HMPD y ejemplos. En (Inami et al., 2000; Bimber et al., 2007).	157
Figura 104. Sistema de HMDP óptico y tele-transportación de la cara del usuario a un espacio remoto. En (Rolland et al., 2002).	157
Figura 105. Esquema de los distintos dispositivos de mano empleados en RA. A partir de (Bimber et al., 2007, 121).	158
Figura 106. Diversas escenas del juego CitiTag: a) Aviso de que se puede “pillar” a un jugador; b) Jugador pillado; c) Aviso de rescate de un jugador del mismo equipo. En (Vogiazou et al., 2007).	159
Figura 107. Dispositivo AR-Pad y usuarios interactuando. En (Mogilev et al., 2002).	160
Figura 108. Esquema de funcionamiento e imágenes de un display de mano óptico con un transductor ultrasónico. En (Stetten et al., 2001).	160
Figura 109. Dispositivo iLamp: a) Componentes; b) Prototipo de proyector aumentado; c) Registro por marcas fiduciales y superposición de proyecciones. En (Raskar et al., 2003).	161
Figura 110. The Periscope, visualización a través del display y niños interactuando. En (Wilde et al., 2003).	162
Figura 111. Geometría de la proyección en los displays espaciales ópticos. A partir de (Bimber et al., 2007, 126).	163

Figura 112. Visuo-Haptic Display. En (Kawakami et al., 1998).....	164
Figura 113. Simulador de visión de Rayos-X: a) Display LCD, sistema de registro (Optotrak 3020) y espejo semitransparente; b) Punto de vista del cirujano. En (Nikou et al., 2000).	164
Figura 114. Espejo semitransparente a modo de pizarra. En (Bimber et al., 2001c).....	165
Figura 115. Esquema de los displays espaciales de proyectores. A partir de (Bimber et al., 2007, 131).....	166
Figura 116. Objeto tridimensional de color neutro y el mismo objeto texturizado mediante Shader Lamps, para lo que se han utilizado un total de tres proyectores. En (Bimber et al., 2007, 162). .	167
Figura 117. Paredes y paredes aumentadas mediante proyecciones. En (Low et al., 2001).....	167
Figura 118. Esquema del funcionamiento de ec(h)o. En (Hatala et al., 2004).....	168
Figura 119. Usuarios interactuando con Notebook en la sesión demo del congreso ACE'05, UPV.	169
Figura 120. LISTEN LISBOA. Compositora grabando sonidos y cartografía del sonido virtual. En (Le Prado et al., 2007).....	170
Figura 121. Componentes de SmartFinger. En (Ando et al., 2002).....	171
Figura 122. Forehead Retina System. En (SIGGRAPH06, 2006).....	171
Figura 123. Usuario interactuando con Straw-like User Interface, imagen del dispositivo y esquema. En (Hashimoto et al., 2005).....	172
Figura 124. Wearable olfactory display. En (Yamada et al., 2006).....	172
Figura 125. SpotScents: a) Registro de la posición de la nariz; b) Esquema del funcionamiento del sistema; c) Cuarto prototipo del cañón de aire. En (Yanagida et al., 2004; Bimber et al., 2007).	173
Figura 126. Secuencia de los procesos de registro de una marca y superposición de elementos virtuales en ARToolKit. En (Kato et al., 2000a, 8).	179
Figura 127. Patrón de calib_distortion y dos ejemplos de calibración de la cámara. En (Kato et al., 2000a, 27).....	180
Figura 128. Patrón de calib_cparam y dos ejemplos de calibración de la cámara. En (Kato et al., 2000a, 28).	180
Figura 129. Marcas de ARToolKit: a) PattHiro; b) PattKanji; c) PattSample1; d) PattSample2. En (Human Interface Technology Laboratory, 2007).	181
Figura 130. Proceso de registro: a) Imagen original; b) Imagen umbralizada; c) Componentes conectados; d) Contorno; e) Extracción de lados y esquinas; f) Cubo ajustado a la marca. En (Human Interface Technology Laboratory, 2007).	182
Figura 131. Secuencia de los procesos de registro de una marca y superposición de elementos virtuales en MXRToolKit. En (Mixed Reality Lab, 2003c).....	184
Figura 132. Proceso de calibración: a) Patrón de calibración con el borde verde en forma de U; b) Cuadro de diálogo mostrando los parámetros de calibración por defecto; c) Proceso de captura de la calibración, mostrando tres barras coloreadas que indican el estado del proceso; d) Proceso de revisión, en el que el usuario puede visualizar la bondad del proceso. En (Díaz, 2007).....	186

Figura 133. Ejemplos de marcas: a) Frame que consta de una sola marca; b) Frame que consta de un conjunto de cuatro marcas.	187
Figura 134. Ejemplo de aplicación en Studierstube: Cows vs. Aliens. En (Fiala, 2004, 11; Fisher et al., 2004).	188
Figura 135. Marcas de ARTag y ejemplo de registro con oclusión por un dedo.	189
Figura 136. Ejemplos de escenas de RA disponibles en AMIRE y componentes conectados para su elaboración: a) Un objeto 3D se asigna a una marca (Marker1); b) Se ejecuta un sonido cuando dos marcas (Marker1 y Marker2) son visibles. En (CVLAB - Computer Vision Laboratory, 2007).	192
Figura 137. Proceso de calibración geométrica, superposición de una figura 3D y calibración fotométrica de BazAR.	195
Figura 138. Ejemplo de registro de color mediante el objeto tap.jit.colortrack de Tap Tools.	198
Figura 139. Ejemplo de cambio de posición y dirección del eje óptico de la cámara virtual.	199
Figura 140. Objeto jit.gl.model.	200
Figura 141. Modelo obj modificado mediante jit.gl.model: a) Mapeado de vídeo; b) Modo de alambres; c) Transparencia (blend activado); d) Salida matricial.	200
Figura 142. Objeto mt9 para Jitter.	201
Figura 143. Conexión del receptor GPS con Jitter: a) Mediante el objeto serial y el patch programado por (Bleecker, 2005); b) Mediante MaxLink y el código escrito en Processing por (Wang et al., 2006)	202
Figura 144. Esquemas dibujados por Ishii: GUI vs. TUI. En (Wang et al., 2006).	208
Figura 145. The Legible City. En (Cheok et al., 2003; Benford et al., 2006; Rashid et al., 2006; Le Prado et al., 2007).	210
Figura 146. Usuarios interactuando con la pieza Vídeo Balancín. En (Billinghurst et al., 2002b).	212
Figura 147. Poultry Internet. En (Szalavári et al., 1998).	213
Figura 148. Usuarios interactuando en un juego colaborativo y esquema de la red del juego. En (Sanchez-Segura, 2004, 236).	214
Figura 149. Diseño e interacción con ClearBoard-2. En (Mann, 1997).	217
Figura 150. Vestidos computacionales: a) Teclado que controla la chaqueta musical de Levi's, estando completamente hecho de tela; b) Tela elaborada por el MIT Media Lab, completamente lavable; c) Uno de los primeros prototipos de Smart Shirt desarrollada en Georgia Tech. En (Barfield et al., 2001b, 490-495).	219
Figura 151. Elementos gráficos en Kyoto Garden. En (Barfield, 1995).	222
Figura 152. Usuario interactuando con Magic Land. En (Charles et al., 2004; Dow et al., 2006).	227
Figura 153. Secuencia del juego interactivo James Bond. En (Dow et al., 2006).	227
Figura 154. Escena del juego de realidad aumentada Façade y usuario interactuando. En (Stanney, 2002, 894; Gardenfors, 2005).	228
Figura 155. Ejemplos de instalaciones de RA en museos: a) The S.O.L.A.R. System, en (Bimber et al., 2005a); b) Integración de proyecciones sobre cuadros, en (Roussou, 2004).	231

Figura 156. Continuo de realidad mezclada. En (Courchesne, xxxx).	234
Figura 157. Entornos de RA: a) Sistema tradicional de RA, b) Sistema de Usuario Aumentado....	240
Figura 158. Exposición "Michelangelo Pistoletto: A Refected World" (1966) en Walker Art Center, Minneapolis. La imagen de la izquierda corresponde a la pieza Girls on a Balcony (1962-1964). En (Paul, 2003, 103).	243
Figura 159. Imágenes de Hallucination (1988-90), de Jim Campbell.....	243
Figura 160. Diversos videojuegos para EyeToy Play 2.	244
Figura 161. Usuarios interactuando con un sistema de RA basado en la metáfora del espejo mágico. En (Portalés Ricart, 2006).....	245
Figura 162. Intercambio de la imagen de la cara: a) Attack and Swap (2006) de Benedikt Gross; b) AR Magic System (2007) de Diego Díaz y Clara Boj.....	245
Figura 163. Esquema básico de la distribución espacial del sistema de Usuario Aumentado.....	246
Figura 164. Componentes físicos de Usuario Aumentado.	247
Figura 165. Imágenes de las marcas libres.	253
Figura 166. Imágenes del gorro con prisma ajustable.....	254
Figura 167. Modelos virtuales utilizados en las pruebas de laboratorio.	254
Figura 168. Ejemplos de interacción en las primeras pruebas de laboratorio.....	255
Figura 169. Imágenes del prisma con diadema.....	256
Figura 170. Imágenes de las marcas ajustables para las manos.....	257
Figura 171. Modelos virtuales utilizados en Cambio de Personalidad.	257
Figura 172. Los autores de Usuario Aumentado interactuando con la versión de "Cambios de Personalidad", en unos ensayos realizados en uno de los estudios de la facultad de BBAA.	258
Figura 173. Esquemas de las bandas giratorias de Carrusel de Imágenes: a) Una sola banda; b) Tres bandas que se entrecruzan.	259
Figura 174. Ejemplos de interacción con el Carrusel de Imágenes.	260
Figura 175. Esquema del sistema de Túnel Mágico.	262
Figura 176. Montaje de Túnel Mágico en uno de los estudios de la facultad de BBAA.....	262
Figura 177. Imágenes del montaje y funcionamiento del prisma con diadema retroiluminado.	263
Figura 178. Modelos virtuales utilizados en Túnel Mágico.	263
Figura 179. Interactuando con Túnel Mágico.....	264
Figura 180. Esquema de la distribución espacial de los dispositivos de Un [∞]	267
Figura 181. Los integrantes del Laboratorio de Luz montando la pieza Un [∞] en el Centro Párraga de Murcia.....	268
Figura 182. Imágenes del prisma con diadema mejorada.	268
Figura 183. Imágenes de las marcas a modo de guantes.	269
Figura 184. Modelos virtuales utilizados en el espejo de las emociones: a) Máscara griega femenina; b) Máscara griega masculina; c) Frases vinculadas a las manos.....	270

Figura 185. Modelos virtuales utilizados en el espejo de los estereotipos: a) Modelo para la cabeza masculino; b) Modelo para la cabeza femenino; c) Esferas de palabras para las manos.	271
Figura 186. Modelos virtuales utilizados en el espejo de género-multicultural: a) y b) Modelos de cajas para la cabeza; c) Palabras en forma de espiral y figuras radiales para las manos.	272
Figura 187. Secuencia de imágenes de varios usuarios interactuando con la pieza de Un [∞] . Centro Párraga (Murcia).	272
Figura 188. Esquema de las relaciones espaciales en el sistema de Cineatro.	274
Figura 189. Imágenes de las marcas a modo de peto.	275
Figura 190. Modelos virtuales utilizadas en Cineatro.	275
Figura 191. Diversos usuarios (niños y adultos) interactuando con la pieza de Cineatro.	276
Figura 192. Puntuaciones otorgadas por los usuarios; distinción entre hombres y mujeres.	278
Figura 193. Puntuaciones otorgadas por los usuarios; distinción entre mayores y menores de 30 años.	279
Figura 194. Usuarios interactuando con el sistema de Usuario Aumentado en la facultad de BBAA; fase de Cambio de Personalidad y Túnel Mágico.	280
Figura 195. Usuarios interactuando con una demo de Usuario Aumentado en el congreso VSMM 2005, en Gante, Bélgica. Fase de cambios de personalidad.	281
Figura 196. Ubicación del Campus de Vera respecto al centro histórico de Valencia y los monumentos Les Torres de Serrans y El Micalet.	285
Figura 197. Imágenes de Emperadores Desplazados, Arquitectura Relacional 2. En (Höllner et al., 1999).	287
Figura 198. Usuario con los dispositivos de RA y visualización de un edificio derruido. En (Wilde et al., 2003).	288
Figura 199. Dispositivos físicos utilizados en UrbanMix.	289
Figura 200. Campus de Vera: a) Imagen aérea, en (Lynch, 1960); b) Modelo tridimensional realizado con el programa 3D Studio Max.	290
Figura 201. Imágenes de Les Torres de Serrans y modelo 3D.	291
Figura 202. Imágenes de El Micalet y modelo 3D.	291
Figura 203. Sistema de referencia terrestre local.	292
Figura 204. Labores topográficas.	293
Figura 205. Puntos de apoyo utilizados para el registro del usuario.	293
Figura 206. Ubicación espacial de los edificios históricos en el campus de Vera.	294
Figura 207. Superposición de un elemento virtual a la imagen vídeo en tiempo real sin consideración del problema de la oclusión, donde: a) Edificio virtual, b) Entorno aumentado.	294
Figura 208. Superposición de un elemento virtual a la imagen vídeo en tiempo real resolviendo el problema de la oclusión, donde: a) Modelo virtual del campus de Vera, b) Modelo virtual del campus de vera junto al modelo virtual de Les Torres de Serrans integrado espacialmente, c)	

Objetos virtuales superpuestos a la imagen vídeo, d) Entorno aumentado con resolución de la oclusión mediante el canal alfa sobre el modelo del campus de Vera.	295
Figura 209. Arquitectura de programación.....	296
Figura 210. Patch principal de UrbanMix.	297
Figura 211. Sub-patches para el cálculo de la orientación de la cámara: a) p camori; b) p leastsquares.	298
Figura 212. Sub-patches: a) p trackpoint; b) p model.....	299
Figura 213. Sub-patch p light.	299
Figura 214. Imágenes de Olton Av., Beeston.	303
Figura 215. Field-Work@Alsace, en (Höllerer et al., 1999; Reitmayr et al., 2003; Kolbe, 2004).	307
Figura 216. Sistema aumentado de LBS en interiores. En (Kolbe, 2004).	307
Figura 217. Sistema aumentado de LBS en una zona urbana. En (Bleecker, 2005).....	308
Figura 218. Dispositivos de Hot Images.	309
Figura 219. Modelo Virtual simplificado del Jubilee Campus.....	310
Figura 220. Interfaz de navegación y relaciones conceptuales.	311
Figura 221. Mapa de cinco puntos calientes en el campus.	311
Figura 222. Sistema de referencia terrestre local, donde O es el origen de coordenadas (0,0) y A representa un punto cualquiera con coordenadas (X_A, Y_A)	312
Figura 223. Esquema de la arquitectura de programación.	313
Figura 224. a) Patch principal de Hot Images; b) Ventana de OpenGL en vista 2D y 3D.	314
Figura 225. Sub-patch p IMU.	315
Figura 226. Cálculo de la posición del usuario: a) Sub-patch p GPS; b) Código insertado en el objeto programado por (Wilde et al., 2003) en Processing.....	315
Figura 227. Sub-patch p dist.	316
Figura 228. Sub-patch p catchimg.	316
Figura 229. Sub-patch p plane.	316
Figura 230. Sub-patch p campus.	317
Figura 231. Sub-patch p images.....	318
Figura 232. Sub-patch p textures.....	318
Figura 233. Storyboard.	321
Figura 234. La instalación KidsRoom y usuarios interactuando, en (Wikipedia, 2008c).	330
Figura 235. Diversas escenas: a) En The Sims (2000); b) En The Sims Online (2002).	331
Figura 236. Componentes de la aplicación; Imágenes de la segunda fase del ensayo (apartado 10.3.2).....	332
Figura 237. Objetos 3D utilizados en LLH.	334
Figura 238. Asignación de marcas a los frames.	335
Figura 239. Código de programación que describe la interacción entre LEGOMan y la tele.	338
Figura 240. Distintas escenas en la tele: a) Los dos muñecos están sentados en sus respectivos asientos y aparece un vídeo que muestra a varias personas en una ceremonia dentro de un	

templo Hindú; b) Sólo está sentada LEGOWoman, y aparece un vídeo que muestra la zona comercial de Chinatown; c) Sólo LEGOMan está sentado y se ve un vídeo que muestra un hawker centre.	339
Figura 241. Ejemplos de distintos escenarios físicos.	346
Figura 242. Distintas escenas de LLH a modo de storyboard.	348
Figura 243. Imagen mostrando el problema de la oclusión: a) Con una cámara en posición baja; b) Con la cámara en una posición elevada.	350
Figura 244. Primera versión de la casa y muñecos.	351
Figura 245. Algunas mejoras añadidas: TV sobre fondo de pared; posición de la cámara más elevada para aumentar la zona visualizada y evitar oclusiones.	352
Figura 246. Salvapantallas: a) Vídeo de la calle desde el patio de una casa; b) Imagen de la pared de la casa con cromas en las ventanas; c) Vídeo resultado de la composición del primer vídeo con la pared de la casa.	353
Figura 247. Elementos de LLH en la segunda fase.	354
Figura 248. Usuarios interactuando con LLH V.1 en el Mixed Reality Lab.	355
Figura 249. Niña de dos años jugando con LLH.	357
Figura 250. Bóveda del altar mayor: a) Bóveda gótica (antes de ser restaurada); b) Bóveda barroca (antes de ser desmantelada).	361
Figura 251. Modelo 3D de la bóveda barroca obtenido con un láser escáner terrestre. En (Las Provincias, 2006).	362
Figura 252. Retablo del altar mayor: a) Retablo actual con las puertas cerradas; b) Retablo actual con las puertas abiertas; c) Retablo de plata desaparecido.	362
Figura 253. Reconstrucciones virtuales 3D: a) Modelo 3D de la tumba del emperador Maximiliano I, en (Gutierrez et al., 2004); b) Sistema CAVE para el antiguo suburbio musulmán Sinhaya, en (Koutsoudis et al., 2007); c) Reconstrucción 3D fotorrealística de la ciudad de Xanthi, en (Aguiló et al., 2001).	365
Figura 254. ARCHEOGUIDE: a) Usuario equipado con el sistema; b) Imágenes de ruinas griegas; c) Imagen aumentada con reconstrucciones 3D. En (Hall et al., 2002; Bimber et al., 2003; Brown et al., 2003; Hatala et al., 2004; Wojciechowski et al., 2004; Woods et al., 2004).	365
Figura 255. Escenas de RA. En (Lepetit et al., 2005).	366
Figura 256. Dispositivos de CatedrAR.	367
Figura 257. Imágenes utilizadas para el mapeado de planos, donde la zona en sombra corresponde a la transparencia del canal alfa: a) Imagen de la bóveda; b) Imagen del retablo con las puertas abiertas.	367
Figura 258. Usuario con los dispositivos de CatedrAR: a) Entorno real; b) Simulación de realidad aumentada con una bóveda virtual tridimensional.	368
Figura 259. Imagen de la bóveda tomada desde debajo de esta: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.	372

Figura 260. Imagen del retablo: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.....	373
Figura 261. Imagen tomada desde el altar mayor: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.....	374
Figura 262. Imagen tomada desde una esquina del altar mayor: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.	374
Figura 263. a) Piano – as image media; b) Telesymphony. En (Levin, 2006).	382
Figura 264. Messa di Voice. En (Cycling '74, 2008).	382
Figura 265. Esquema de los componentes y distribución espacial de AR-Jazz.	383
Figura 266. NURBS cilíndrico a partir de js jitaudio2nurbs.js, donde: a) Con textura; b) Representación alámbrica; c) Con textura y deformaciones a partir de un audio; d) Representación alámbrica y deformaciones a partir de un audio.	385
Figura 267. Patch principal de AR-Jazz.	386
Figura 268. Sub-patch p sonido.	387
Figura 269. Sub-patch p model plano.....	387
Figura 270. Sub-patch p mood_machine.	388
Figura 271. Sub-patch p IMU.	388
Figura 272. Imágenes tomadas durante el taller de AR-Jazz.	389
Figura 273. Visualizaciones del sonido de acuerdo al instrumento utilizado y a la nota emitida.	393
Figura 274. Imágenes tomadas durante la Jam session.....	394

Tablas

Tabla 1. Evolución cronológica de los ensayos y distintas actuaciones dentro del periodo de duración de la beca FPI.	51
Tabla 2. Comparación de dispositivos de registro. En (Neumann, 2001, 333).....	101
Tabla 3. Datos que registra un INS; donde, de izquierda a derecha: Tiempo medido en segundos; Aceleraciones en los tres ejes medido en m/s^2 ; Velocidad de los giros en los tres ejes medidos en rad/s ; Valor del campo magnético en los tres ejes medidos en a.u. (intensidad del campo magnético: 1 a.u. = $235052 \text{ Wb}/m^2$); Temperatura en el interior del sensor medida en $^{\circ}C$	109
Tabla 4. Usos de la RA en función de los estímulos virtualmente reproducidos. En (Barfield, 1995, 568-570).	176
Tabla 5. Secuencia del código en ARToolKit. En (Human Interface Technology Laboratory, 2007).	179
Tabla 6. Elementos con interacción, acciones derivadas y elementos multimedia que se muestran.....	340
Tabla 7. Letras correspondientes a los elementos de la casa en el proceso de calibración.....	347
Tabla 8. Puntuaciones individuales y medias (MED).....	356

Tabla 9. Comparación entre los distintos programas analizados en esta tesis. Ventajas e inconvenientes encontrados.	404
Tabla 10. Comparación entre los distintos métodos de registro utilizados en los ensayos de esta tesis. Ventajas e inconvenientes encontrados.	404

Ecuaciones

Ecuación 1. Ecuaciones de colinealidad.	102
Ecuación 2. Matriz de orientación.	103
Ecuación 3. Términos de la matriz de orientación.	103
Ecuación 4. Longitud de un arco de meridiano.	108
Ecuación 5. Longitud de un arco de paralelo.	108
Ecuación 6. Ecuaciones para el cálculo de la distancia focal.	111
Ecuación 7. Sistema de ecuaciones para la obtención del centro de la imagen y coeficientes de distorsión.	112

INTRODUCCIÓN

*"Aquél que posee arte y ciencia tiene religión; quien no las posee, necesita religión"*⁴ (Goethe, 1774).

Tradicionalmente, el Arte y la Ciencia se perciben como dos disciplinas contrarias: o se es artista o se es científico. Existe la creencia popular de que la Ciencia es racional, objetiva y fría, en tanto que el Arte es subjetivo, irracional y emotivo.

Según (Díaz, 2008) *"todo arte consiste, en esencia, en la creación de formas, en una transformación que se manifiesta, finalmente, en la producción de una estructura. A su vez, toda forma, natural o creada por el ser humano tiene, potencialmente, información, es decir, puede transmitirse en el proceso que llamamos comunicación. La obra de arte es así un vínculo entre quien la produce y quien la observa y experimenta"*. Por otra parte, las teorías y verdades científicas surgen a partir del contraste de datos observados de la Naturaleza, siendo por ello limitadas, condicionadas y relativas. Así pues, la grandeza de la Ciencia radica en que no puede ofrecer un espacio de certeza absoluta, aunque su aporte al conocimiento es esencial

⁴ Texto original:

*Wer Wissenschaft und Kunst besitzt,
Hat auch Religion;
Wer jene beiden nicht besitzt,
Der habe Religion.*

y trascendente. Según (García Canal, 2008): "la Ciencia está basada en la objetividad, no en la creencia, permitiendo que los fenómenos adquieran una definición racional, una descripción coherente y universal al independizarla del hombre particular que hace la observación o la teoría". De este modo, mientras al científico le mueve la realidad y la lógica (mundo objetivo), el artista se basa en la imaginación y la emoción (mundo subjetivo).

Sin embargo, siglos atrás esta distinción no era tan evidente. El artista era un conocedor de la técnica y la ciencia, haciendo uso de éstas en su proceso creativo; mientras que el científico era un ser imaginativo y creativo que se aventuraba a defender nuevas teorías, antes incluso de poder ser demostradas empíricamente.

Un claro ejemplo lo podemos encontrar en la figura del célebre Leonardo da Vinci, artista, pensador e investigador italiano (Vinci, Toscana, 1452 - Amboise, Turena, 1519) interesado por todas las ramas del saber y por todos los aspectos de la vida. La obra pictórica de Leonardo es la que le ha hecho destacar como un personaje cumbre en la historia del arte. Sin embargo, según su criterio no debía existir separación entre el arte y la ciencia, como no la hubo en sus investigaciones, dirigidas de forma preferente hacia temas como la anatomía humana, la zoología, la geología, la astronomía, la física o la ingeniería. En este último terreno fue donde quedó más patente su talento precursor, ya que Leonardo concibió multitud de máquinas que no dio a conocer entre sus contemporáneos y que la técnica ha acabado por convertir en realidad siglos más tarde: aparatos de navegación (como un submarino, una campana de buceo y un salvavidas), máquinas voladoras (como el paracaídas, una especie de helicóptero y unas alas inspiradas en las de las aves para hacer volar a un hombre), obras de ingeniería civil (como canalizaciones de agua o casas prefabricadas), etc.

Otro ejemplo lo podemos encontrar en Johannes Kepler (Weil der Stadt, Alemania, 1571 – Ratisbona, Alemania, 1630) astrónomo alemán conocido por enunciar las tres leyes que describen el movimiento de los planetas en sus órbitas alrededor del Sol. Kepler también tuvo su incursión en el arte, al estudiar las órbitas de los planetas para establecer una relación entre el movimiento de estos cuerpos celestes con una teoría musical. En su libro *Harmonices Mundi*, postuló que las velocidades angulares de cada planeta producían sonidos consonantes y escribió seis melodías, una para cada uno de los planetas del Sistema Solar conocidos hasta entonces. Al combinarse, estas melodías podían producir cuatro acordes distintos, siendo uno de ellos el acorde producido en el momento de la creación y otro el que marcaría el momento del fin del universo.

Aunque encontremos hechos comunes en ambos campos, es indudable que deben existir ciertas diferencias entre la Ciencia y el Arte. Para (Díaz, 2008), estas diferencias se basan esencialmente en el método y el objetivo:

- El método del científico se basa en emplear técnicas muy elaboradas para realizar sus observaciones. Necesita instrumentos cada vez más complejos y precisos. Una vez obtenidos los datos, el científico realiza la última etapa del método: la escritura del artículo científico, que es la obra propiamente dicha. El método artístico también parte de la observación, sin embargo esta observación se basa en el refinamiento de factores perceptivos, cognitivos y emocionales propios: el artista depura su sensibilidad. En este caso, y a diferencia de la ciencia, no se generan datos (generalmente numéricos) a los que es necesario dar una interpretación, sino que se genera una representación más directa; la técnica en el arte se emplea, fundamentalmente, en la producción de la obra.
- El arte es representación y el objeto artístico es la expresión de dicha representación. De igual modo, la ciencia es una representación del mundo y la producción de teorías a partir de ella. Sin embargo, hay diferencias en los objetivos: el propósito de la ciencia es producir conocimiento certero y general sobre aspectos restringidos del mundo; el del arte es producir una emoción estética.

Según (Berenguer, 2002) el divorcio entre artistas y científicos se inició con Newton y su modelo mecanicista del Universo, y se consolidó con las consecuencias de su método, especialmente durante la Revolución Industrial: a partir del siglo XIX las máquinas sustituyeron al hombre, ahorrándole esfuerzo pero, a la vez, restándole identidad y humanidad. Los artistas, como reacción, se refugiaron en sí mismos. Sin embargo, algunos artistas se plantearon una visión alternativa. Gino Severino, un artista futurista, decía: *"Yo preveo el fin del cuadro y de la estatua. Estas formas de arte, incluso empleadas con el espíritu más genuinamente innovador, limitan la libertad creativa del artista. Ellas mismas contienen sus destinos: museos y galerías de coleccionistas, en otras palabras, cementerios"*. Los futuristas se planteaban la búsqueda de nuevas formas artísticas, en la que la máquina no era un enemigo, sino un aliado y hasta una obra susceptible de valoración estética. Por otra parte, en los años 60 Robert Rauschenberg y Billy Kluver fundaban EAT (*Experiments in Art and Technology*), una organización dedicada a potenciar creativamente la síntesis entre arte y tecnología. Entre las

actuaciones de EAT destacó la exposición "Nine evenings", celebrada en Nueva York en 1966; por primera vez la autoría de las obras exhibidas se hallaba repartida entre el artista y el técnico. De igual modo, algunos científicos se acercaban al mundo del arte. Cabe destacar la figura de Isaac Asimov, bioquímico y escritor estadounidense de origen ruso, reconocido como uno de los maestros del género de ciencia ficción.

Además, existen evidencias que demuestran como esta separación entre Arte y Ciencia se desmorona. En primer lugar, la propia Ciencia no puede prescindir de las consideraciones morales o sociales de sus descubrimientos, mientras que el Arte en muchas ocasiones se basa en hechos científicos o hace uso de tecnologías desarrolladas por éstos. Según (García Canal, 2008): *"los científicos para hacer Ciencia usan de su imaginación y de su inspiración en muchas oportunidades. Por el contrario, muchas veces el Arte surge como resultado del agregado, o del desagregado, de partes sobre bases racionales. Entonces, Arte y Ciencia son nada más que diferentes formas de la creatividad. Por lo tanto podemos decir que Arte y Ciencia son realmente complementarias"*. Además, hay elementos gozosos en la ciencia así como también hay elementos cognitivos en el arte. Según (Díaz, 2008): *"el científico goza el placer estético que le produce un experimento bien diseñado, al que califica de "elegante", y el artista o el crítico bien saben que la reflexión y la contrastación no están excluidas del arte; de hecho, le son consustanciales"*.

Tanto el Arte como la Ciencia tienen una influencia decisiva en la generación de puntos de vista; un creador inventa una nueva manera de ver y de expresar, se inspira o basa en lo existente y su trabajo afecta a quienes lo siguen. Ciertamente, hacer Ciencia implica desconfiar severamente de los argumentos de autoridad: el científico es libre de formular cualquier pregunta, de dudar de cualquier aseveración, de buscar nuevas evidencias y de corregir cualquier error. Por ello la Ciencia es peligrosa para doctrinas con pretensión de certidumbre. De igual modo, el Arte, contribuye al proceso de comprensión y aprecio de nuestra propia existencia. Enseña a no aceptar ni a auto-imponer límites a la imaginación.

A finales de los años setenta surge un elemento revolucionario que hace que Arte y Ciencia encuentren un marco común evidente: el ordenador. En la actualidad, la aplicación del ordenador a la comunicación y expresión audiovisual ha progresado extraordinariamente; su práctica constituye un claro ejemplo de fusión entre Arte y Ciencia a través de la tecnología. Según (Berenguer, 2002): *"La potencia simuladora del ordenador permite reproducir, hasta un límite imprevisible, las actividades humanas"*

de concebir, diseñar, imaginar, comunicar... El ordenador es pues una prótesis de la mente y, como tal, se halla a disposición del artista; en particular, del artista para el cual creatividad y técnica constituyen las dos caras de la misma moneda".

Uno de los productos que ha generado el mundo digital es la denominada realidad virtual: el ordenador es capaz de simular un entorno alternativo a nuestro entorno físico, y de generar artificialmente estímulos sensoriales (gráficos, audio, olores, etc.). En este nuevo entorno las leyes físicas y matemáticas que rigen nuestro mundo pueden emularse, o de otro modo pueden ser alteradas para producir nuevas dimensiones alternativas. Los usuarios de esta tecnología pueden interactuar con la máquina o con otros usuarios dentro de este nuevo mundo, alcanzando un grado de aceptación que les permite adentrarse física y psicológicamente en estas realidades artificiales, concepto conocido como "inmersión". A partir de la realidad virtual nacen nuevos conceptos y tecnologías, como el ciberespacio, telepresencia, vida artificial, realidad mezclada, realidad aumentada, etc. Todos ellos modelos actuales del avance científico y la creación artística.

En esta tesis se busca una sinergia entre Arte y Ciencia en el campo concreto de la realidad aumentada, tecnología que ha emergido en los últimos años a partir de la realidad virtual. El método científico y las leyes físicas y matemáticas se emplean para dar lugar a unos resultados alternativos que poco tienen que ver con el primero. Así pues, el método utilizado es el científico, pero el objetivo perseguido es el artístico. Todo ello se realiza a través de una serie de ensayos, en los que la imaginación y el proceso creativo son comunes, como hemos argumentado, a ambas disciplinas.

Motivación personal

Desde muy joven he sentido un interés especial por la Ciencia. El método científico es apasionante y la Ciencia me produce admiración y belleza. Para mí, la Ciencia es un Arte, y el científico un creador. Pongamos como ejemplo a Leonardo da Vinci, Kepler, Galileo, Einstein o Isaac Asimov. Quizá por ello me cuesta entender la barrera que durante tanto tiempo ha separado estas dos disciplinas, a mi juicio, complementarias: el científico busca nuevas "verdades", mientras que el artista las pone en duda o examina bajo otros puntos de vista. Por lo tanto, necesariamente el artista debe ser conocedor de dichas verdades; mientras que los interrogantes y la crítica de éstas animan al científico a avanzar en su búsqueda. Además, es un hecho que el científico

se vale del Arte como vehículo para la difusión de sus conocimientos; mientras que los avances en la Ciencia se ven reflejados en el Arte. Así pues, Arte y Ciencia caminan juntas de la mano.

Es por ello que, aunque mi formación es científica, no dudé cuando tuve la oportunidad de embarcarme en una tesis doctoral dentro del mundo de las Artes. Para mí ha supuesto una oportunidad excelente para completar mi formación académica y personal, abriendo la mente hacia nuevas perspectivas y maneras de afrontar los retos, nuevos modos de ver, sentir, escuchar, actuar... La temática de la tesis – entornos de realidad aumentada – me facilita dicho nexo pretendido entre ambas disciplinas, puesto que las leyes físicas y matemáticas que rigen esta nueva tecnología están estrechamente relacionadas con mi titulación, por ejemplo: las ecuaciones matemáticas que describen la transformación entre distintos sistemas de coordenadas, los sensores de registro tipo GPS o INS, el cálculo de diversos parámetros de orientación a partir del análisis digital de la imagen, etc. Sin embargo, gran parte de estos conceptos suelen ser desconocidos en el mundo del Arte. Es por ello que esta tesis se ha planteado en dos bloques: en el primero de ellos se introduce la tecnología, metodologías y dispositivos, abordando los temas relevantes pero sin entrar en explicaciones matemáticas exhaustivas; mientras que en el segundo de los bloques se hace uso de dicha tecnología para obtener resultados específicos, aplicados al mundo del Arte o relacionados con éste.

Contexto académico

La titulación con la que accedí a estudios de doctorado es la de Ingeniera en Geodesia y Cartografía. En el año 2003 inicié dichos estudios en el Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría de la UPV, donde realicé los cursos de doctorado y un trabajo de investigación. A mediados del año 2004 me concedieron una beca de Formación de Personal Investigador (FPI) del Ministerio de Ciencia y Tecnología, que aproveché para unirme al grupo de Investigación Laboratorio de Luz en la Facultad de Bellas Artes, en el proyecto “Percepción ampliada: un sistema dual de realidad aumentada” y en el programa de doctorado “Artes visuales e intermedia”. Dentro del periodo que abarca esta beca he realizado dos estancias en centros de investigación extranjeros (ver Tabla 1): en el año 2005, una estancia de 6 meses en el Mixed Reality Laboratory (MRL) de la University of Nottingham (UK), bajo la tutoría de Tom Rodden; en el año 2006, una estancia de 6 meses en el Mixed Reality Lab (MXR) de

la National University of Singapore, bajo la tutoría de Adrian David Cheok. También he realizado dos Diplomas de Estudios Avanzados (DEA), uno en el Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría y otro en el Departamento de Pintura. Cumplidos dos años de la beca FPI, se me contrató como técnico superior de investigación de la UPV, figura contractual que mantengo en estos momentos. También he colaborado como docente en la asignatura "Videoarte" de cuarto año de carrera para la titulación de BBAA durante el curso académico 2006/2007, asignatura de la que es titular la profesora María José Martínez de Pisón; durante este curso 2007/2008 colaboraré como docente en la asignatura "Entornos de realidad aumentada/virtual en el Arte" del Máster en Producción Artística, asignatura de la que es titular el profesor Francisco Giner.

Contexto científico

Como hemos introducido anteriormente, la realidad aumentada es un paradigma que surge a partir de la realidad virtual. Sin embargo, así como la realidad virtual es una tecnología bastante desarrollada y asentada, en el caso de la realidad aumentada todavía es necesario que pasen unos años para poder afirmar lo mismo. Tanto para la realidad virtual como para la realidad aumentada podemos decir que existen una serie de fases, representadas en la pirámide de la Figura 1 elaborada a partir de (Bimber et al., 2005d, 6), y que se lee de abajo a arriba. En el primer nivel de esta figura nos encontramos los elementos básicos de la tecnología, como los sistemas de tracking y registro, dispositivos de visualización o displays y renderización; en un segundo nivel están los elementos más complejos, como son la interacción, presentación y autoría; el tercer nivel lo forman las aplicaciones que constituyen el interfaz con el usuario; y en el último nivel están los usuarios o consumidores finales de dicha tecnología. Según (Bimber et al., 2005d, 6), mientras que la comunidad científica se encuentra en el segundo nivel de desarrollo para el caso de la realidad virtual, la realidad aumentada se encuentra todavía en el primer nivel, por lo que no es de extrañar que las aplicaciones actualmente desarrolladas no sean lo suficientemente robustas, que los interfaces no tengan un aspecto depurado (en muchos casos los interfaces son de

propia construcción y no están comercializados; ej.: HMD con cámaras acopladas), o que no se haya profundizado lo suficiente en temas referentes al usuario final⁵.

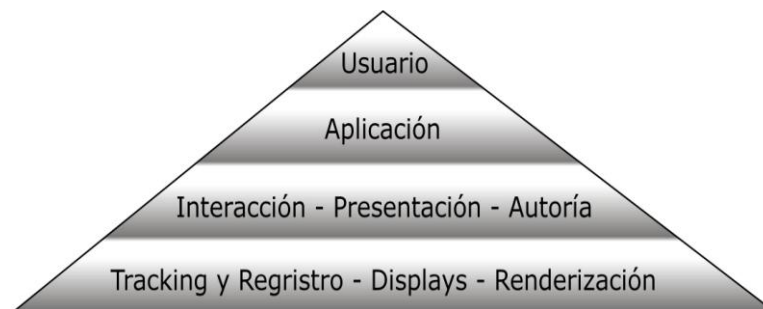


Figura 1. Pirámide de las etapas de desarrollo de la tecnología de RA. Elaborada a partir de (Bimber et al., 2005d, 6).

Sin embargo, en este esquema la parte fundamental es el usuario final, la punta de la pirámide, puesto que dependiendo de la utilidad real o beneficios que la tecnología de realidad aumentada pueda aportar a la sociedad o a un grupo de usuarios, se seguirá invirtiendo en ella y se mejorarán sus aspectos técnicos, para convertirla en una tecnología robusta, eficaz y asequible. Pero hay que tener en cuenta que su actual fragilidad impide que se elaboren aplicaciones lo suficientemente estables, complejas y/o atractivas, por lo que, según (Livingston, 2005), muchas de ellas quedan como simples ensayos de laboratorio, sin posibilidad de comercializarse. Debido a ello, los estudios realizados en cuanto al usuario final son también escasos (ver capítulo 6). Sin embargo, si no se demuestran los beneficios que esta tecnología puede aportar con respecto a otras, difícilmente se fomentará su desarrollo.

Estas reflexiones datan del año 2005, poco después de empezar esta tesis. Si bien son todavía ciertas, desde entonces se aprecia un interés creciente por el estudio de aspectos relacionados con el usuario final, como en (Cheok et al., 2006; Dünser et al., 2006; Ha et al., 2007). De igual modo en esta tesis, en la medida de lo posible, se han evaluado los ensayos mediante pruebas de usuarios, obteniendo resultados y conclusiones concretas.

⁵ Usuario o usuario final: persona que usa un sistema informático y sus programas de aplicación para realizar tareas y producir resultados. En (Giannetti, 2002, 197).

Límites de la investigación

- **Naturaleza de la investigación.** Esta tesis tiene carácter teórico-práctico. Aunque se ha centrado en la elaboración de una serie de ensayos, también se ha hecho una revisión bibliográfica, extrayendo aquellos conceptos, metodologías, conclusiones, etc. que se han considerado relevantes para el entendimiento de la tecnología, estado actual de desarrollo y corrientes futuras. En la parte práctica se han elaborado un total de seis ensayos para la resolución de diversas cuestiones planteadas a partir de la tecnología de realidad aumentada.
- **Acotación del área de estudio.** El área de estudio en la que se basa esta tesis es la tecnología de realidad aumentada, otorgando especial atención a los distintos entornos que se pueden generar a partir de un estudio general, dispositivos y software que se utilizan, consideraciones referentes al usuario final, y las distintas alternativas tecnológicas y metodológicas que han aportado diversos investigadores para establecer dichos entornos.
- **Ensayos realizados.** En esta tesis se han elaborado un total de seis ensayos, algunos de ellos consistentes en varias fases. El número y temática no se había planteado en un principio, sino que han ido surgiendo durante la investigación llevada a cabo y se han visto condicionados por distintas circunstancias, explicadas en cada uno de los ensayos.
- **Acotación metodológica.** Para la parte teórica de esta tesis (bloque I) se ha empleado una metodología de búsqueda bibliográfica y puesta en común de investigaciones y desarrollos aportados por distintos autores. Por otra parte, la metodología seguida en la parte práctica (bloque II) es más extensa, de acuerdo a los siguientes pasos: en primer lugar se ha planteado un problema; a continuación se ha realizado una revisión bibliográfica para ver cómo los distintos autores han aportado soluciones a problemas similares; después se ha elaborado una solución propia, desarrollado y obtenido resultados; en la medida de lo posible, se han validado los sistemas obtenidos a partir de diversas pruebas; por último, se han redactado escritos y presentado en congresos especializados de ámbito nacional e internacional. Este último punto se ha visto limitado por cuestiones económicas. Todos estos pasos se explican con detalle en el apartado "metodología" de esta introducción.
- **Acotación bibliográfica.** La bibliografía consultada en esta tesis ha sido extensa y variada, aunque se han consultado en mayor medida artículos publicados en

congresos. Ello es debido en primer lugar a que la novedad relativa de la tecnología de realidad aumentada no ha dado lugar todavía a una extensa bibliografía editada de otro modo; por otra parte, a que éstos son el medio normal de difusión de los últimos avances e investigaciones, que han sido especialmente numerosos en los últimos 7-8 años. Es por ello que la mayoría de bibliografía consultada está en inglés. En la medida de lo posible, estos documentos bibliográficos se han recopilado en formato *pdf*, para insertarlos a continuación en una base de datos bibliográfica elaborada con el programa EndNote. El tipo de citación utilizado es *author-date*, en el que las citas aparecen de la forma: (autor, año); en algunos casos (principalmente, cuando la cita se refiere a una página específica de un libro o documento extenso) se ha añadido el número de página: (autor, año, página). En el apartado de bibliografía, los datos presentados consta únicamente de los autores citados en el texto; sin embargo, la base de datos elaborada con EndNote consta de un mayor número de autores, y se pone a disposición en el DVD de la tesis para poder ser consultada y utilizada por todos aquellos que lo deseen.

- **Actualidad de la investigación.** Esta tesis se basa en una tecnología que está actualmente en continuo desarrollo y expansión. Este hecho tiene la ventaja de que la temática tratada es actual y novedosa; sin embargo, tiene la desventaja de que quizá, en el momento de lectura de esta memoria, hayan acontecido nuevos desarrollos que no estén aquí contemplados.

Objetivos

El principal objetivo del primer bloque es establecer un marco de la tecnología de realidad aumentada de acuerdo a su estado actual y de desarrollo de los últimos años, haciendo un análisis de los dispositivos que intervienen. Se han introducido en el texto una gran diversidad de ejemplos aportados por diversos autores, para ofrecer una visión más cercana de los temas tratados. De otro modo, el principal objetivo del segundo bloque es la resolución de cuestiones concretas (planteadas al comienzo de cada ensayo) mediante el empleo de esta tecnología. Se pueden enumerar una serie de objetivos generales:

- Estudiar la incursión de la realidad aumentada en el mundo del arte y disciplinas afines, así como elaborar nuevas aplicaciones en estos campos.

- Investigación sobre las posibilidades de esta tecnología a partir de un estudio de las partes que la hacen posible: sistemas integrados, entornos, sensores, displays, software, usuario final, campos de aplicación, etc.
- Mediante los ensayos se persigue estudiar la viabilidad y adecuación de sistemas económicos para la generación de aplicaciones de realidad aumentada en diversos campos, mediante el empleo y combinación, en la medida de lo posible, de distintos sensores, displays, software, etc.

Metodología

El contenido del bloque I se ha elaborado durante la fase de redacción de la memoria de la tesis (ver Tabla 1), por lo que la metodología empleada ha consistido en una búsqueda bibliográfica y posterior contraste con distintos autores; la elaboración de una base de datos bibliográfica ha facilitado el proceso en gran medida. De otro modo, en la fase práctica – los ensayos – se ha seguido la siguiente metodología:

- Se acotó el marco técnico y temático que se podía abordar con la tecnología de RA.
- Revisión bibliográfica para ver cómo distintos autores han aportado soluciones a problemas similares, y evaluar el estado actual de desarrollo.
- Realización de pruebas de laboratorio para, finalmente, elaborar una solución propia a partir de ciertos desarrollos. En total, se han obtenido seis ensayos, cada uno respondiendo a unas necesidades concretas, dentro de un área específica.
- Los resultados obtenidos se han analizado y discutido, aceptando las limitaciones o condicionantes en cada caso. Cuando ha sido posible, se han realizado pruebas con usuarios para validar los sistemas obtenidos (a partir de cuestiones referentes al usuario final) y establecer un punto de referencia para futuros desarrollos.
- Algunos de estos ensayos se han presentado a congresos y revistas científicas de alto impacto. Con ellos, no sólo se difunde entre la comunidad investigadora el trabajo realizado, sino que además existe un feedback que permite plantearse nuevas formas de seguir la investigación. El feedback es mutuo, por lo que también se conoce de primera mano el trabajo realizado por otros investigadores a nivel nacional e internacional, con la oportunidad de entablar debates sobre cuestiones concretas, y adquiriendo una visión generalizada sobre los intereses actuales de investigación y próximos desarrollos dentro del

campo en cuestión. En los anexos de esta tesis se presentan las ponencias presentadas; también se presenta una ponencia aceptada para el congreso ENGAGE 2006 como “full paper”, que finalmente no fue publicado por no poder inscribirse debido a cuestiones económicas (principalmente, porque el lugar de celebración era Sydney, Australia). Dos artículos enviados a revistas de impacto no se incluyen por encontrarse actualmente en fase de revisión.

La siguiente tabla refleja la duración de cada uno de los ensayos junto con las distintas actuaciones realizadas dentro del periodo de la beca FPI, destacando la parte académica (obtención de los DEA y redacción de la memoria de la tesis), las estancias realizadas en centros internacionales de investigación, las asignaturas impartidas y las publicaciones en congresos y revistas.

PERIODO		ACADÉMICO	ESTANCIAS/ ASIGNATURAS	PUBLICACIONES/ EXPOSICIONES	ENSAYOS						
2004	JUN				USUARIO AUMENTADO	URBANMIX					
	JUL										
	AGO										
	SEP										
	OCT	DEA DICGF									
	NOV										
	DIC										
2005	ENE						HOT IMAGES				
	FEB		Estancia: Mixed Reality Laboratory (MRL) de Nottingham								
	MAR										
	ABR										
	MAY										
	JUN			Conf. ACE							
	JUL										
	AGO										
	SEP										
	OCT			Conf. VSMM							
	NOV			Conf. NMTA (2)							
	DIC										
2006	ENE										
	FEB										
	MAR	DEA Pintura		Catálogo Centro Párraga							

	ABR	Estancia: Mixed Reality Lab (MXR) de Singapur	Exp. Centro Párraga							
	MAY									
	JUN		Catálogo UPV							
	JUL		Exposición UPV							
	AGO									
	SEP									
	OCT									
	NOV									
	DIC									
2007	ENE	Redacción de la memoria de la Tesis	Asignatura: Videoarte							
	FEB									
	MAR									Conf. INTED (2)
	ABR									
	MAY									
	JUN									Revista SOJEL, Conf. ACE (2)
	JUL									Evento Jazz & Arts Meeting
	AGO									
	SEP									
	OCT									
	NOV									
	DIC									
2008	ENE		Asignatura: Entornos RA/RV							
	FEB									
	MAR									
	ABR									
	MAY									

Tabla 1. Evolución cronológica de los ensayos y distintas actuaciones dentro del periodo de duración de la beca FPI.

Estructura de la memoria

Como ya se ha dicho anteriormente, la memoria de esta tesis se estructura en dos grandes bloques. En el primero de ellos (capítulos 1-6) se hace una revisión de la

tecnología de realidad aumentada; en el segundo de ellos (capítulos 7-12) se presentan los ensayos elaborados. A continuación se explican los contenidos abarcados en los capítulos pertenecientes a cada bloque.

El primer capítulo recoge algunos conceptos generales de la realidad aumentada. Hay que tener en cuenta que esta tecnología es relativamente nueva, por lo que en sus comienzos su propia definición es distinta a la utilizada hoy en día; es por ello que se aportan diversas definiciones, conceptos generales y evolución histórica de esta tecnología. Finalmente se estudian, de manera generalizada, distintos campos de aplicación a partir de aplicaciones concretas aportadas por distintos autores.

En el segundo capítulo se explican las bases matemáticas de esta tecnología. Para ello se analizan los distintos sistemas de referencia que intervienen en el proceso de generación de la escena aumentada (imagen, objeto, terreno, etc.; del entorno real y del virtual) y las transformaciones espaciales entre ellos. También se trata el problema de la oclusión y las condiciones de sombreado para que la escena generada sea creíble, aportando algunas consideraciones que deben tenerse en cuenta en la elección de los dispositivos de registro y displays.

En el tercer capítulo se enumeran y describen algunos dispositivos de registro o sensores que se utilizan en aplicaciones de realidad aumentada, los cuales envían datos de entrada (*inputs*) al ordenador para el cálculo (generalmente) de los parámetros de orientación externa de la cámara. Se realiza una clasificación de los sensores a partir de la aportada en (National Research Council Staff, 1994). Además, para cada tipo de sensor analizado se muestran uno o dos ejemplos de dispositivos explicando su funcionamiento y datos técnicos, así como algunos usos aportados por distintos autores en aplicaciones de realidad aumentada.

Siguiendo un esquema similar, en el capítulo cuatro se analizan tipos de displays utilizados en aplicaciones de realidad aumentada. Se ofrece una clasificación de los mismos a partir de la aportada por (Bimber et al., 2005d) para el caso de los de tipo visual, en los que se enumeran sus ventajas e inconvenientes; también se tratan los displays de tipo sonoro, háptico, olfativo, gustativo y multimodales. Para todos los casos se introducen distintos desarrollos y/o aplicaciones aportadas por diversos autores.

En el capítulo cinco se analiza el software disponible (generalmente de libre distribución a través de Internet) para el desarrollo de aplicaciones de RA. Estos

programas se han clasificado en tres grupos: aquellos que realizan el registro con cámaras mediante el reconocimiento de patrones de marcas planas, los que realizan el registro con cámaras mediante el reconocimiento de una serie de rasgos naturales, y aquellos que, aun no siendo específicos para aplicaciones de realidad aumentada, se pueden elaborar tales aplicaciones ya que permiten la entrada de datos de sensores y la gestión de imágenes, objetos 3D, sonido, etc., en tiempo real. Se abordan con más detalle aquellos que se han utilizado en el desarrollo de los ensayos de esta tesis: ARToolKit, MXRToolKit, BazAR y Max/MSP Jitter.

En el capítulo seis se analiza el usuario final, bajo el punto de vista de la interacción y la experiencia en entornos de realidad aumentada. Se distinguen distintas formas de interactuar (con el ordenador y/o con otros usuarios), se introduce el concepto de interacción sin costuras, y se aborda el tema de los vestidos computacionales. En cuanto a la experiencia, se abordan algunos aspectos referentes a la percepción humana, inmersión y compatibilidad; finalmente se introducen algunos estudios o conceptos referentes al usuario final que aportan los campos del entretenimiento, aprendizaje y arte.

Ya en el bloque II de la tesis, en el capítulo siete se presenta uno de los ensayos realizados, cuyo título es *Usuario Aumentado*, junto con sus diversas fases (un total de seis). Éste es una instalación de realidad aumentada en el que se introduce el concepto del espejo mágico para proponer un cambio de imagen, cuestionando la identidad del usuario final y provocando una modificación de su conducta.

El capítulo ocho contiene el ensayo titulado *UrbanMix*, una aplicación de realidad aumentada para espacios abiertos en la que se genera un nuevo entorno con la mezcla de distintos espacios urbanos. La localización del usuario dentro de dicho entorno se consigue mediante la combinación de distintos dispositivos de registro.

En el capítulo nueve se describe el ensayo denominado *Hot Images*, aplicación para espacios abiertos elaborada durante mi estancia en MRL (ver Tabla 1). En este ensayo se propone una nueva cartografía al dotar a los edificios del entorno urbano de connotaciones humanas, elaborando al mismo tiempo una interfaz de navegación novedosa a partir de un juego infantil.

En el capítulo diez se explica el ensayo *Live LEGO House*, elaborado durante mi estancia en MXR (ver Tabla 1). En éste se presenta un juego para niños en el que se

aumenta un juguete tradicional (una casa elaborada con bloques de LEGO) con diversos elementos generados por ordenador, activados a partir de la interactividad entre dos muñecos y los elementos de la casa. En este juego se introducen diversos factores multiculturales y se premia la buena convivencia entre los muñecos.

En el capítulo once se explica el describe *CatedrAR*, una aplicación elaborada para aumentar diversos elementos del interior de la Catedral de Valencia. La condición de que el entorno físico de la Catedral no se vea alterado, influye en la elección de los dispositivos y el software utilizado.

En el capítulo doce se explica el ensayo *AR-Jazz*, una instalación elaborada para representaciones en vivo en la que se ofrece una visualización del sonido y del movimiento a partir de las interpretaciones de un combo⁶. Este ensayo ha sido elaborado para el evento *Jazz & Arts Meeting*, en el que también se ha ofrecido un taller dirigido a los músicos que intervienen a continuación en una *Jam session*⁷.

Se ha seguido una estructura similar para presentar cada uno de los ensayos: en primer lugar se realiza una introducción donde se resume el trabajo realizado, la motivación y el marco temporal en el que se inscribe el ensayo. A continuación se explica el concepto o idea general, enmarcando el ensayo dentro de los distintos dispositivos y conceptos abordados en el bloque I, introduciendo también algunos referentes (principalmente) artísticos. El siguiente apartado contiene la descripción técnica a partir de los componentes de la aplicación (físicos y virtuales), desarrollo, proceso, consideraciones y restricciones técnicas encontradas; este apartado puede variar dependiendo de las necesidades de cada ensayo. Después se introduce un apartado de las fases de desarrollo, en caso de haberlas. Por último, en algunos ensayos se introduce un apartado de verificación a partir de pruebas de usuarios.

Una vez abordados los dos grandes bloques de esta tesis, se aporta un apartado de conclusiones, en el que se resumen, discuten y valoran cada uno de los capítulos, aportando además un análisis comparativo entre los distintos dispositivos y programas utilizados en los ensayos. Finalmente, en los anexos se aportan las publicaciones que se han elaborado durante la realización de esta tesis.

⁶ Combo: agrupación reducida de músicos de Jazz.

⁷ Jam session: improvisación sobre melodía, sin mucha preparación y arreglos definidos.

BLOQUE I: ASPECTOS TEÓRICOS

1. MARCO GENERAL DE REALIDAD AUMENTADA

La realidad aumentada es una tecnología novedosa que surge a partir de la tecnología de realidad virtual. Aunque su aparición es bastante reciente, hoy en día su utilización es cada vez más extendida, y podemos encontrar diversos ejemplos y aplicaciones en áreas muy distintas, tales como el arte, educación, medicina o ingeniería. Debido a su reciente aparición, se pueden encontrar en la bibliografía distintas definiciones, variando las características que la definen. Es por ello que se ha procurado unificar criterios y clarificar la definición que se utiliza en esta tesis doctoral. Además, existen otras tecnologías muy próximas, como la virtualidad aumentada, que deben ser revisadas y diferenciadas.

1.1. DEFINICIONES

1.1.1. REALIDAD MEZCLADA Y EL CONTÍNUO DE MILGRAM

El concepto de **realidad mezclada (RM, *mixed reality*)** fue definido por primera vez en (Milgram et al., 1994) a partir del continuo realidad-virtualidad, comúnmente conocido como **continuo de Milgram** (Figura 2) que aparece referenciado por múltiples autores, como en (Milgram et al., 1994; Milgram et al., 1997; Billinghurst et al., 2001a; Florins et al., 2005; Jacobs et al., 2006; Lam et al., 2006). Según este continuo, dependiendo de la cantidad de entorno generado por ordenador, se puede establecer una clasificación que va desde el entorno completamente real al completamente virtual, pasando por los estados intermedios de realidad aumentada (RA) y virtualidad aumentada (VA). La tecnología de RM engloba a la RA y la VA. La diferencia entre estas dos últimas tecnologías radica en la cantidad de entorno sintético (o contenido generado por ordenador) frente a la cantidad de entorno real (estos conceptos se explicarán con más detalle en los siguientes sub-apartados), aunque en muchos casos la línea que separa ambas tecnologías es difícil de trazar y algunas aplicaciones no se pueden catalogar claramente en una de estas categorías. Cabe señalar que Milgram diferencia entre: Entorno Real – Realidad Aumentada – Virtualidad Aumentada – Entorno Virtual. Sin embargo, a mi parecer en esta clasificación se entremezclan distintos conceptos ("entorno" frente a la tecnología de RA o VA). Es por ello que en la Figura 3 se muestra una modificación (o ampliación) del continuo de Milgram, ya que contempla la diversidad de conceptos paralelos aportados por otros autores, comúnmente referenciados en la bibliografía consultada.

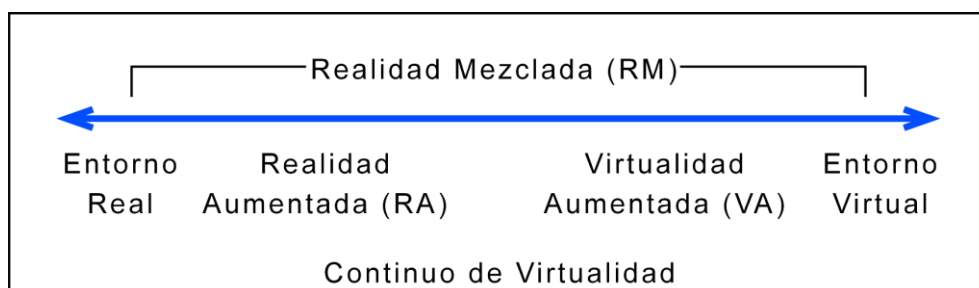


Figura 2. Continuo de Milgram. A partir de (Milgram et al., 1994).

Realidad	Realidad Mezclada / Realidad Híbrida		Virtualidad
	Realidad Aumentada	Virtualidad Aumentada	
Entorno Real / Mundo Real	Entorno Mezclado / Entorno Híbrido / Entorno Aumentado		Entorno Virtual / Entorno Sintético / Mundo Virtual
	Entorno Real Aumentado	Entorno Virtual Aumentado	
Realidad Real	Realidad Videográfica		Realidad Virtual / Realidad Sintética
0% contenidos generados por ordenador]0%, 50%[contenidos generados por ordenador]50%, 100%[contenidos generados por ordenador	100% contenidos generados por ordenador

Figura 3. Continuo de Milgram ampliado con otros conceptos.

1.1.2. REALIDAD VIRTUAL

La **realidad virtual (RV, virtual reality)**, es según (TechTarget, 2008) la “*simulación de un entorno real o imaginado que se puede experimentar visualmente en las tres dimensiones de anchura, altura y profundidad, y que puede proveer adicionalmente una experiencia interactiva visual en un movimiento completamente en tiempo real con sonido y posiblemente con retroalimentación háptica o de otro tipo. La forma más simple de realidad virtual es una imagen 3D que se puede explorar de manera interactiva en un ordenador personal, normalmente por medio de la manipulación de teclado o ratón, de tal forma que el contenido de la imagen se mueve en la misma dirección, se acerca o se aleja*”.

Esta definición no es unánime entre los diferentes autores. Por ejemplo, en (Sherman et al., 2003, 13) se da una definición más compacta de RV, en la que sin embargo no se incluye el término “3D” ni se indica que es lo que se “simula”: “*la realidad virtual es un medio compuesto de simulaciones por ordenador interactivas que registran las posiciones y acciones de los participantes y reemplazan o aumentan el feedback a uno o más sentidos, dando la sensación de estar mentalmente inmerso o presente en la simulación (un mundo virtual)*”.

Se pueden distinguir dos tipos de RV: 1) Simulación de un entorno real, como el interior de un edificio; se suelen emplear para aplicaciones museísticas, de entrenamiento o de aprendizaje; 2) Simulación un entorno puramente ficticio; generalmente utilizado dentro del campo del arte, los videojuegos y el entretenimiento. En este segundo caso nos podríamos cuestionar el término "realidad" dentro del término RV, y quizá sería más conveniente hablar de **entorno virtual (VE, Virtual Environment)**, que abarca ambos casos y no da lugar a conflictos conceptuales.

Un buen ejemplo de entornos de realidad virtual en el campo del arte es la pieza *The 21st Century Virtual Reality Color Organ* (2000) de Jack Ox (Figura 4), descrita por su autora como "un sistema computacional para trasladar composiciones musicales a representaciones visuales" (Ox et al., 2000). Esta pieza está diseñada para sistemas interactivos e inmersivos, como es el caso de la CAVE (ver apartado 1.3).

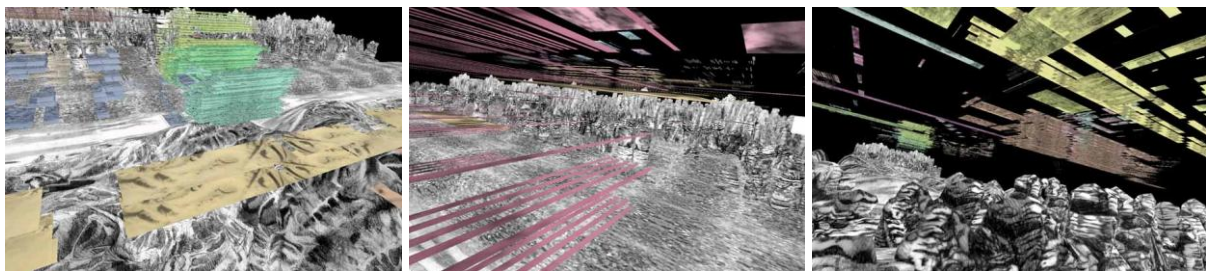


Figura 4. Imágenes del entorno virtual *The 21st Century Virtual Reality Color Organ* (2000) de Jack Ox, en (Ox et al., 2000).

1.1.3. VIRTUALIDAD AUMENTADA

Según (Zlatanova, 2002), la **virtualidad aumentada (VA, augmented virtuality)** define un entorno principalmente virtual, el cual puede ser aumentado mediante la inclusión de vídeos o texturas del mundo real. En el primer caso, el vídeo puede ser pregrabado o en tiempo real. En el segundo caso, las texturas pueden corresponderse con las que los objetos virtuales (si son modelos de objetos reales) tienen en la realidad, lo que generalmente se conoce como modelos foto-realistas. En (Simsarian et al., 1997) se presenta una aplicación de VA, en la que diversos objetos virtuales son mapeados con texturas y vídeos fotorrealistas (Figura 5).



Figura 5. Ejemplo de un entorno de VA, en (Simsarian et al., 1997).

En todo caso, hay que tener en cuenta el continuo de Milgram, explicado en el apartado 1.1.1. La transición entre un mundo puramente virtual y la virtualidad aumentada, o entre la virtualidad aumentada y la realidad aumentada puede ser muy sutil, dándose el caso de que algunas aplicaciones rocen el límite entre una categoría u otra. Es decir, ¿qué cantidad de virtualidad y/o realidad debe aparecer en una aplicación para ser considerada como virtualidad aumentada? En términos generales podemos decir que en un entorno virtual en el que no existe ningún enlace (generalmente visual) con el mundo real, es un entorno de realidad virtual; si el entorno creado es esencialmente virtual pero existe alguna representación (como una imagen o vídeo) del mundo real, hablaríamos de virtualidad aumentada; si el entorno es principalmente real y existen algunos elementos virtuales, nos referiremos a la realidad aumentada, la cual se define en el siguiente apartado.

1.1.4. REALIDAD AUMENTADA

La **realidad aumentada (RA, augmented reality)** es una disciplina relativamente nueva. Las primeras definiciones que podemos encontrar restringen la RA a aquellos sistemas que utilizan displays tipo Head Mounted Displays (HMD, ver apartado 4.1.1.2), apuntando además a un carácter exclusivamente visual. Definiciones posteriores extienden el concepto de RA, abriendo el campo a otros dispositivos y a contenidos multimedia no necesariamente visuales, por lo que surgen nuevas posibilidades. A continuación se apuntan algunas de estas definiciones:

- En (Janin et al., 1993) se utiliza el término de RA para definir un HMD que permite al usuario ver su entorno como una superposición de gráficos generados por ordenador.

- Según (Beaudouin-Lafon, 1994, 14), *“la RA reconoce que la gente está acostumbrada al mundo real, el cual no puede ser reproducido de forma fidedigna en un ordenador. Por lo que la RA lo que hace es construir en el mundo real aumentándolo con capacidades computacionales”*. A continuación hace una comparativa con la RV: *“en lugar de insertar al usuario en un mundo generado por ordenador, la RA recubre el mundo real con el mundo virtual, o incrusta/fusiona ambos mundos”*.
- De acuerdo con (Barfield, 1995, 543) el concepto de RA *“se refiere a la ampliación del mundo real con imágenes sintéticas, por lo que no se requiere que la escena esté completamente generada por ordenador, sino que la imagen sintética se utiliza como complemento de la escena del mundo real”*. Sin embargo a continuación se añade que *“la RA no debe entenderse como exclusivamente visual, sino que debe incluir también información táctil/tangible y auditiva”*.
- En (Bimber et al., 2005d, 2) no se da una definición formal de lo que es RA, sino que se deja a disposición del lector quien, después de haber leído el libro, deberá reflexionar sobre la cuestión. Sin embargo, sí se hace una comparativa con los sistemas de RV, diciendo que, en contraposición a estos, *“en los sistemas de RA el entorno real no se suprime por completo, sino que juega un papel fundamental ya que, en lugar de someter a los usuarios en una inmersión dentro de un mundo puramente sintético, la RA intenta adherir suplementos sintéticos al mundo real”*. Subraya además, que existe un fuerte vínculo entre ambos mundos (real y virtual) establecido principalmente mediante una relación espacial.

Esta última reflexión se aproxima más al concepto de realidad aumentada actual, pues considera un terreno más amplio, sin restringir la tecnología de realidad aumentada a los sistemas que utilizan dispositivos HMD o son puramente visuales. También aporta un dato fundamental que no encontramos en las definiciones anteriores: la relación espacial que se establece entre el entorno real y el mundo virtual (o elementos virtuales) generados por ordenador. Este hecho implica directamente la necesidad de que exista algún tipo de sensor(es) que establezca(n) dicha relación. Considerando el entorno real como un espacio de tres dimensiones (3D), el problema queda resuelto mediante el cálculo de seis parámetros, que son tres traslaciones y tres rotaciones (ver apartado 2.1.2).

Según (Sherman et al., 2003, 517), los sistemas de RA tienen constreñimientos de diseño que difieren de los sistemas de RV. Mientras que en la RV el objetivo es la visualización

de suficientes polígonos para que el entorno creado sea creíble, en la RA el mundo ya existe, y sólo es necesario añadir una pequeña cantidad de información. Así pues, para la RA el reto consiste en producir sistemas independientes, portables, y que sean capaces de registrar y/o ubicar de forma precisa el entorno virtual en el entorno real. Además, los entornos visualizados en sistemas de RV normalmente permiten al usuario moverse de forma virtual por el espacio. Sin embargo, en RA el desplazamiento físico del usuario o de un robot (por ejemplo, en telepresencia) es generalmente el único método aceptado para viajar. En la siguiente figura se muestra un ejemplo de un entorno de RA mediante el empleo de las librerías BazAR (descritas en el apartado 5.2.1).

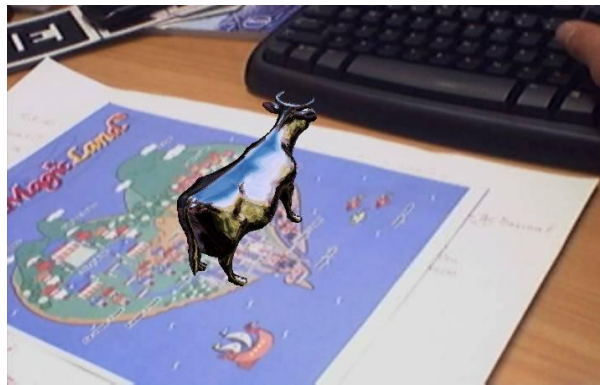


Figura 6. Ejemplo de entorno de RA mediante las librerías BazAR, en (CVLAB - Computer Vision Laboratory, 2007).

1.1.5. REALIDAD MEDIATIZADA

Aunque este término aparece menos referenciado, tal como señala (Mann, 2001, 304) la **realidad mediatizada (RMe, mediated reality)**, es un término más general que engloba a la realidad aumentada. La RMe incluye la adición de objetos virtuales a la realidad visual, pero también la posibilidad de eliminar, alterar, reducir y cambiar de manera significativa la percepción de la realidad visual. La RMe trata pues de “mediatizar” visualmente los objetos reales, para lo que se suelen emplear operaciones de filtrado que alteran la visión de la realidad, para después superponer los objetos virtuales.

Este tipo de efecto visual sólo se podrá conseguir mediante el empleo de displays de realidad aumentada basados en imagen-vídeo (ver capítulo 3), puesto que presupone una alteración de la imagen de la realidad antes de ser mostrada al usuario. Aunque bien es cierto que la imagen que el usuario recibe de la realidad ya es en sí una

alteración, puesto que depende tanto del sistema óptico que captura la imagen como del tipo de display, y por tanto habrá una serie de alteraciones inevitables en cuanto a color, contraste, brillo, resolución, distorsiones, etc. Es por ello que esta definición puede resultar un poco contradictoria en sí misma y convendría distinguir entre dos tipos de alteraciones: aquellas que responden a las limitaciones de los sistemas físicos empleados, y por lo tanto son inevitables o difíciles de corregir; y aquellas que se introducen a propósito con la finalidad de alterar la imagen adquirida, como por ejemplo los filtros de detección de bordes, granulado, etc. A éstas segundas nos referimos al hablar de realidad mediatizada.

Un claro ejemplo lo podemos encontrar en (Fischer et al., 2005), donde se aplican diversas alteraciones (en este caso, a la imagen resultante de la fusión real-virtual) como el efecto de dibujos animados (resalte de bordes) o un efecto de pincel (granulado y alteración del color de la imagen). En la siguiente figura podemos ver unas imágenes de este trabajo.



Figura 7. Aplicación de realidad mediatizada. En ambas imágenes la cafetera es el objeto virtual. Los filtros estilo *cartoon* se han aplicado a la imagen de la derecha. En (Fischer et al., 2005).

1.2. ASPECTOS GENERALES

1.2.1. CARACTERÍSTICAS DE LA REALIDAD AUMENTADA

Como hemos visto, la definición de RA puede resultar en algunos casos ambigua. Es por ello que se hace necesario el adoptar unas características generales que la definan y distinguan de otras tecnologías próximas. En este sentido, en esta tesis se sigue

el trabajo presentado en (Azuma, 1997), donde se define la realidad aumentada como los sistemas que contienen simultáneamente las tres características:

1. Mezcla de lo real y lo virtual.
2. Interactividad en tiempo real.
3. Registro tridimensional.

Según estas características, películas como *El Señor de los Anillos* (2001-03), *Los Cuatro Fantásticos* (2005-07) o *Spiderman* (2002-07) en la que se mezclan elementos filmicos con elementos tridimensionales generados por ordenador, no pueden considerarse sistemas de RA puesto que no son interactivos (Figura 8).

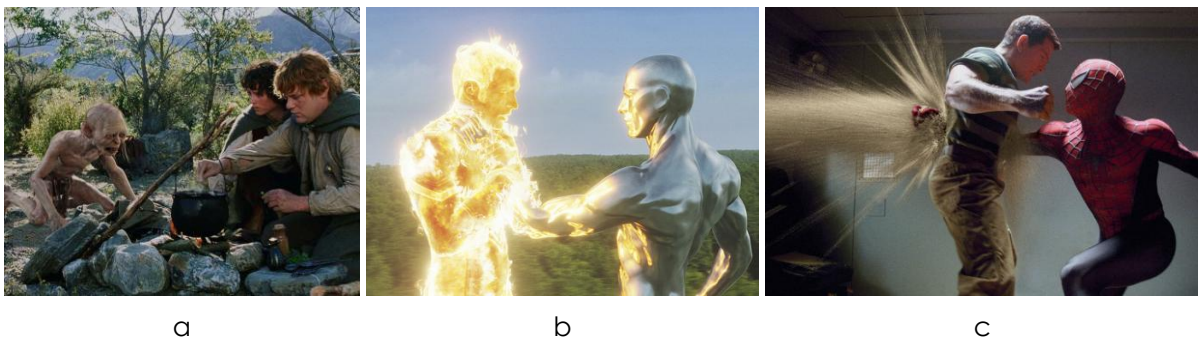


Figura 8. Diversas escenas cinematográficas: a) *The Lord of the Rings: The Return of the King* (2003); b) *Fantastic Four: Rise of the Silver Surfer* (2007); c) *Spiderman 3* (2007).

Sin embargo, en la aportación de Azuma cabe discutir el primer punto – “mezcla de lo virtual y lo real” – en cuanto a cómo se produce dicha mezcla. Basándonos en sistemas visuales (que son los predominantes), en las aplicaciones de RA elaboradas hasta finales de los años 90 esta mezcla se produce a nivel de imágenes (del entorno real y el virtual) que se visualizan de forma conjunta en el dispositivo display (generalmente HMD). Sin embargo, a partir de finales de los 90 surgen otros sistemas en donde la mezcla del entorno real y virtual ya no implica que las imágenes de ambos estén fusionadas en un mismo display, sino que la representación del entorno real en la imagen se sustituye por el propio entorno real, visualizándose únicamente la parte virtual, que sigue guardando una relación espacial con el entorno real y es, al mismo tiempo, interactiva. Estas aplicaciones, sin embargo, suelen catalogarse genéricamente como de realidad mezclada. Un claro ejemplo son las aplicaciones *Can you see me now?* (2001) o *PAC-LAN* (2006), descritas en los apartados 1.4.1 y 1.4.2, respectivamente, que utilizan dispositivos móviles. Otro tipo de trabajos que surgen son las proyecciones estereoscópicas sobre objetos físicos tridimensionales, estableciendo

las bases de la RA espacial (apartado 1.2.2), siendo Bimber (Bimber et al., 2005c) pionero y uno de los máximos referentes actuales.

En otro orden de ideas, también se puede destacar (Billinghurst et al., 2002b), donde se añaden diversos aspectos que proporciona la tecnología de realidad aumentada, como habilidades “potenciales”, contemplando los siguientes aspectos:

- **Interacción sin costuras** entre los entornos virtual y real. La costura de una interfaz fue definida en 1994 por Ishii, y se explica con detalle en el apartado 6.1.2. En la realidad aumentada la comunicación entre usuarios se produce de manera natural, ya que permite un contacto visual directo, permitiendo además que los usuarios puedan ver los elementos virtuales sin excluir la imagen del entorno real. Como ejemplo, en (Kiyokawa et al., 2000b) se realiza un estudio sobre la colaboración entre usuarios a partir de un sistema de RA que permite el diseño y la manipulación de elementos 3D, y que está creado para minimizar las costuras en la interacción.
- La habilidad de **ampliar la realidad**. Mediante los sistemas de RA, a la realidad física se le pueden añadir elementos sensoriales, tales como imágenes, modelos virtuales, sonidos, etc. También se pueden alterar algunas partes de la realidad (mediante filtros), que sería el caso de la realidad mediatizada. A parte de añadir o alterar, existe una tercera posibilidad, la de eliminar elementos del entorno real (Azuma, 1997), por ejemplo, mediante la utilización de máscaras que oculten objetos físicos.
- La presencia de **datos espacializados** (*spatial cues*) para la colaboración remota y presencial. Es decir, permite que los elementos generados por ordenador se puedan distribuir espacialmente conforme al entorno físico de manera interactiva. Por ejemplo, en (Kato et al., 2001) se presenta un trabajo de videoconferencia (Figura 9) en el que los usuarios pueden distribuir libremente en el espacio físico monitores virtuales, pudiendo ubicar a los oyentes alrededor de una mesa, visualizándolos a través de un único display (es decir, sin la necesidad de poner varios monitores reales en el entorno físico).
- Soporte de una **interfaz tangible**. En la RA existe una relación íntima entre los objetos virtuales y los del mundo físico. Los objetos físicos se pueden aumentar mediante diversos elementos generados por ordenador, permitiendo la superposición de dichos elementos de forma dinámica, pudiendo ser accesibles a todos los usuarios (público) o de modo individualizado (privado), y establecer

interacciones físicas. Los objetos físicos se pueden utilizar para manipular los elementos virtuales de una manera intuitiva, sin necesidad de tener conocimientos previos sobre un determinado programa o aplicación. De este modo, las personas que no tienen un mínimo conocimiento del funcionamiento de los ordenadores, pueden sin embargo tener experiencias enriquecedoras dentro de entornos de RA.

- La habilidad de **transitar** de forma suave entre la realidad y la virtualidad. Como se ha visto anteriormente (apartado 1.1.1), según (Milgram et al., 1994) los interfaces pueden situarse en un continuo dependiendo de la cantidad de entorno generada por ordenador. En los extremos del continuo se sitúan el entorno real y el puramente virtual; entre medias, existe una amplia variedad de interfaces que permiten que ambos entornos se entremezclen, haciendo que predomine más un entorno que otro. En las aplicaciones de RA esta propiedad se puede utilizar para establecer transiciones graduales entre el entorno real y el virtual.

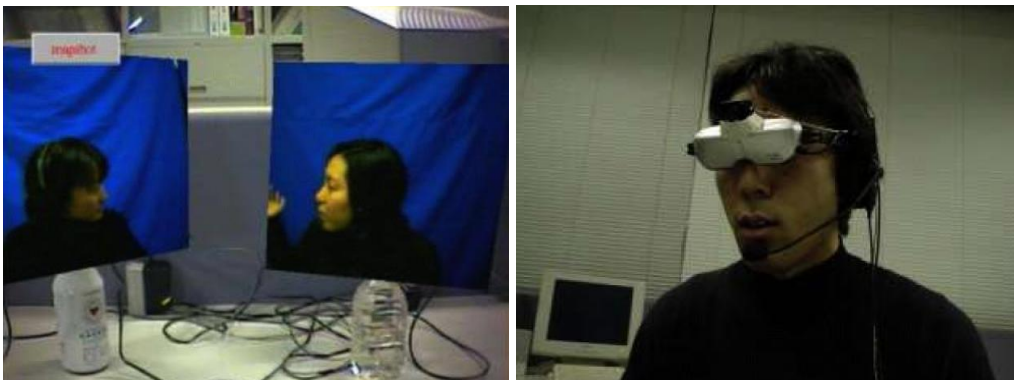


Figura 9. Ejemplo de un sistema RA de videoconferencia en el que los datos están espacializados. En (Kato et al., 2001).

1.2.2. CLASIFICACIÓN DE LOS SISTEMAS DE RA

Actualmente, no existe una única o clara clasificación de los sistemas de realidad aumentada, sin embargo existen algunos términos que establecen cierta catalogación, atendiendo a distintos criterios:

- Según el entorno físico en el que se desarrolla la aplicación, podemos diferenciar entre sistemas dentro de recintos o **cerrados** (*indoors*) vs. sistemas **al aire libre** o **abiertos** (*outdoors*). La diferencia entre aplicaciones dentro de recintos y las aplicaciones al aire libre condiciona muchos aspectos de los

sistemas de RA establecidos, principalmente el tipo de dispositivos de registro y displays utilizados. De hecho, en aplicaciones al aire libre, se suelen emplear receptores GPS y sensores inerciales para el registro, y PDAs o HMDs para los displays; mientras que en los recintos, los dispositivos de registro y displays son mucho más variados. Ejemplos de aplicaciones en espacios cerrados las encontramos en (Bimber et al., 2000; Broll et al., 2004); en espacios abiertos en (Piekarski et al., 2003; Kameda et al., 2004; Avery et al., 2005).

- Según la extensión que abarquen, podemos distinguir entre sistemas **locales** (*local*) vs. **ubicuos** (*ubiquitous*). Los sistemas locales se desarrollan en un ámbito acotado, bien en espacios abiertos o dentro de recintos. Ejemplos de aplicaciones locales las podemos encontrar en (Schnädelbach et al., 2002; Park et al., 2006). Las aplicaciones ubicuas (u omnipresentes) son aquellas en las que la extensión del entorno aumentado es tan amplio, que el usuario tiene la sensación de que vaya donde vaya seguirá estando inmerso en la aplicación. En este tipo de RA se suelen utilizar diversas tecnologías de registro que se complementen, para permitir que el usuario pueda entrar en recintos o salir a espacios abiertos. Además, los sistemas suelen ser móviles, ya que el usuario debe desplazarse libremente. En (Kangas et al., 2002; Newman et al., 2003; Wagner et al., 2004b; Sandor et al., 2005) se presentan técnicas de registro y aplicaciones en el ámbito de la RA ubicua.
- Según la movilidad de los dispositivos de registro y/o displays podemos distinguir entre sistemas **móviles** (*mobile*) vs. sistemas **espaciales** (*spatial*). En los sistemas móviles generalmente el usuario lleva consigo los dispositivos de registro, el display e, incluso, el ordenador o PDA que gestiona la aplicación. Estas aplicaciones son usuales, aunque no restrictivas, de espacios abiertos. Algunos ejemplos los encontramos en (Höllner et al., 1999; Benford et al., 2006; Lee et al., 2006b). Contrariamente, en aplicaciones espaciales, los dispositivos de registro y/o displays están fijos en el entorno, utilizándose en muchos casos sistemas de proyección sobre superficies tridimensionales. En (Bimber et al., 2005c) se hace un estudio exhaustivo sobre la RA espacial.
- Según el número de usuarios que simultáneamente pueden interactuar con el sistema, podemos distinguir entre sistemas **individuales** (*individual*) vs. **colaborativos** (*collaborative*). Así pues, el sistema es individual cuando existe la posibilidad de interacción de un solo usuario, y colaborativa si existe, como mínimo, la posibilidad de participación de dos usuarios de forma simultánea. Ejemplos de aplicaciones colaborativas las encontramos en (Szalavári et al.,

1998; Broll et al., 2004; Cooper et al., 2004); de aplicaciones individuales en (Schnädelbach et al., 2002).

- Según el tipo de colaboración establecida (para sistemas colaborativos), podemos distinguir entre sistemas **presenciales** o **cara a cara** (*face-to-face*) vs. **remotos** (*remote*). En las aplicaciones colaborativas, se distingue si los usuarios colaboran cara a cara, estando físicamente presentes en el mismo entorno, o si la colaboración es remota, estableciéndose generalmente la comunicación a través de la red. Ejemplos de sistemas presenciales los podemos encontrar en (Rolland et al., 2002; Kaufmann et al., 2003); de sistemas remotos en (Billinghurst et al., 2002a; Flintham et al., 2003).

A parte de estos conceptos, en (Mackay, 1998) se describe otra clasificación de los sistemas de realidad aumentada introducida anteriormente en el año 1993, atendiendo principalmente a la localización de los dispositivos de registro. Según su autora, se pueden diferenciar varias formas de aumentar la realidad:

1. **Aumentar al usuario:** El usuario lleva consigo algún dispositivo, generalmente sujeto a la cabeza o en las manos, para obtener información sobre objetos físicos.
2. **Aumentar el objeto físico:** El objeto físico se transforma mediante la inclusión de entradas, salidas o dispositivos computacionales en él.
3. **Aumentar el entorno, el usuario y el objeto:** Ni el usuario ni el objeto son afectados directamente. En su lugar, dispositivos independientes recogen información del entorno, visualizando dicha información sobre objetos y capturando información sobre la interacción de los usuarios con los objetos.

Debemos matizar que esta clasificación aportada por Mackay no es de utilización extendida, y puede resultar ambigua si no se contextualiza exclusivamente en la localización de los dispositivos de registro y los displays, puesto que generalmente el objeto aumentado es aquel que está enriquecido con elementos generados por ordenador, independientemente de que lleve consigo o no diversos dispositivos. Según veremos en el ensayo de *Usuario Aumentado* (capítulo 7), la definición empleada es completamente diferente a la apuntada en el punto 1 de Mackay, ya que en este ensayo lo que se hace es “aumentar al usuario” mediante elementos virtuales que afectan directamente a la imagen que el usuario obtiene de sí mismo y que, por lo tanto, formarán parte de él en el entorno de RA así establecido. Siguiendo la

clasificación de Mackay, *Usuario Aumentado* es, sin embargo, una aplicación en la que se aumenta el objeto físico.

1.3. RECORRIDO HISTÓRICO

La idea de proporcionar datos adicionales a un entorno real ni es novedosa, ni surge con los sistemas actuales de realidad aumentada. Este concepto se ha utilizado en diversos campos; por ejemplo, existen muchos libros en los que se ha utilizado la superposición, sobre fotografías, de elementos no existentes en el entorno físico, dibujados sobre una transparencia de tal forma que se puedan ver fusionados los elementos existentes con los imaginados. En este sentido, se podría argumentar que la realidad aumentada es una disciplina que, aunque tecnológicamente es novedosa, conceptualmente no lo es. Sin embargo, en esta tesis se considera el hecho de que el término "realidad aumentada" surgió para dar nombre a una tecnología y, aunque es cierto que podríamos establecer un juego de palabras y cuestionar que la "realidad" se puede "aumentar" de muchas formas (no exclusivamente mediante la tecnología), lo cierto es que estaríamos hablando de conceptos distintos. Es por ello que en esta tesis al hablar de RA nos referimos a la "tecnología de realidad aumentada", basándose en las características aportadas en (Azuma, 1997) y descritas en el apartado 1.2.1. En este sentido, los comienzos de la RA van unidos a los de la realidad virtual, ya que comparten muchas características. A continuación se muestra un recorrido desde el año 1916 hasta nuestros días, seleccionando los aspectos que se consideran más destacables de la tecnología de realidad aumentada o de otras tecnologías cercanas relacionadas con su creación y desarrollo.

1916

En este año Albert B. Pratt patenta en los EE.UU. un sistema de periscopio para la cabeza (Figura 10). Esta patente con el número 1183492 es considerada el primer HMD de la historia (Sherman et al., 2003, 24).

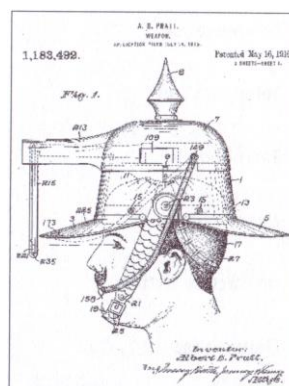


Figura 10. Sistema de periscopio, en (Sherman et al., 2003, 24).

1954

El *Cinerama* (Figura 11) fue inventado por el ingeniero Fred Waller. Su trabajo en los estudios Paramount en los años 30 le llevaron a entender que se podía lograr un efecto de profundidad y realismo en las proyecciones de cine por medio de pantallas curvadas de 180° que incluyeran la visión periférica del público. Sus primeros experimentos fueron con once cámaras de 16mm y proyectores que emitiesen una imagen hemisférica. El sistema se redujo hasta cinco cámaras y proyectores (Hart, 1995; Grau, 2003, 155-157).

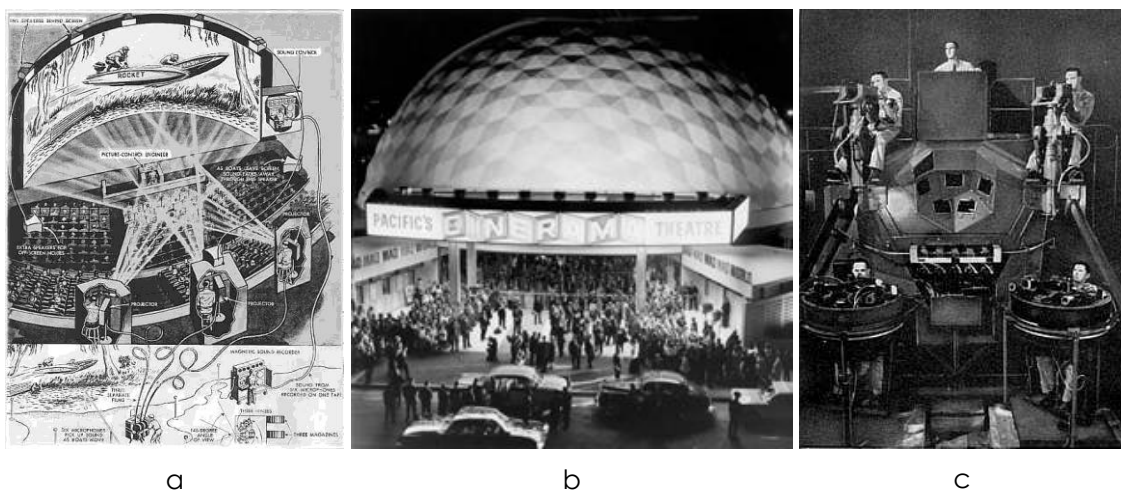


Figura 11. Cinerama: a) Dibujo esquemático, en (Norwood, 1997); b) *Cinerama Dome* en Los Angeles, en (Coate, 2007); c) Sistema con cinco proyectores, en (Hart, 1995).

1956

Morton Heilig desarrolla *Sensorama* (Figura 12), inspirándose en Cinerama. *Sensorama* es la razón por la que algunos autores definen a Morton como el padre de la realidad virtual. Este invento consiste en un sistema de experiencia multimodal, en el que una única persona puede percibir una experiencia pre-grabada mediante vistas estereoscópicas, sonido estereofónico, olor, vibración de la silla y viento. Sin embargo, todavía no se le proporciona al usuario la posibilidad de interactuar con el sistema. Este sistema se patentó en 1962 con el número 3050870. Otros sistemas son el *Sensorama Motion Picture Projector* y el *Sensorama 3-D Motion Picture Camera* (Grau, 2003, 158-159; Sherman et al., 2003, 24; Heilig, 2007; Kunsthochschule für Medien Köln, 2007).

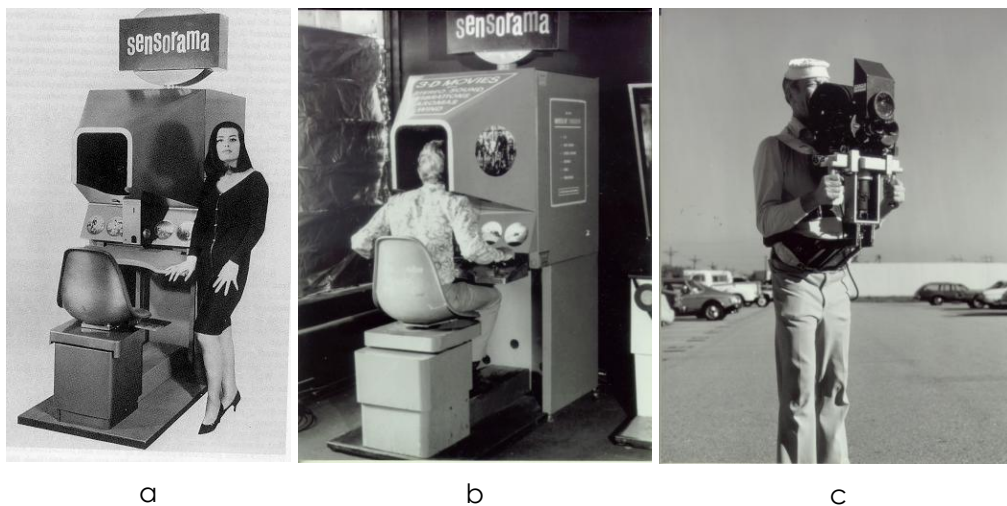


Figura 12. a) y b) Sensorama; c) Sensorama 3-D Motion Picture Camera. En (Heilig, 2007; Kunsthochschule für Medien Köln, 2007).

1961

Los ingenieros Comeau y Bryan de Philco crean un sistema HMD binocular (Figura 13) para utilizarlo como un sistema de visión de una video cámara remota, siguiendo los movimientos de la cabeza mediante un sensor de orientación electromagnético (Sherman et al., 2003; Ellis, 2007).

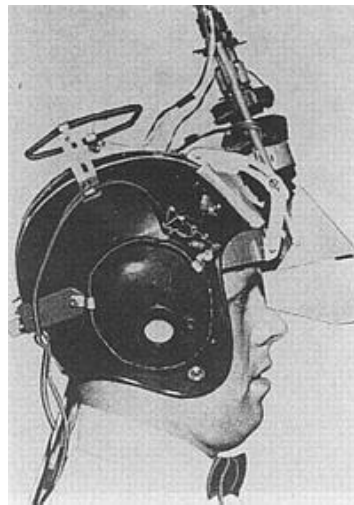


Figura 13. Sistema HMD binocular, en (Ellis, 2007).

1963

Ivan Sutherland, estudiante de Doctorado del MIT (Massachusetts Institute of Technology), desarrolló en su tesis doctoral un programa denominado "Sketchpad" que permitía la manipulación directa de objetos gráficos en una pantalla CRT usando un lápiz óptico (Figura 14). El usuario podía dibujar con precisión los gráficos e interactuar con ellos acercando o alejando su imagen en la pantalla. Este sistema constituye la primera interfaz gráfica de usuario (Grau, 2003, 162; Sherman et al., 2003, 26).

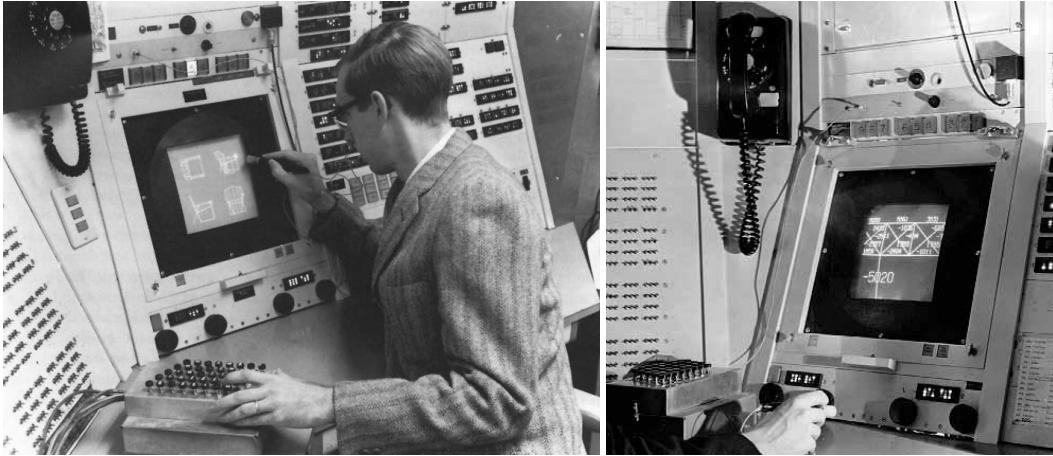


Figura 14. Sutherland interactuando con Sketchpad. En (Mischitz, 2001; Chavarro García, 2007).

1965

Ivan Sutherland explica el concepto del *ultimate display* en su presentación en el congreso International Federation for Information Processing (IFIP). Sutherland sentó las bases del concepto de realidad virtual, definiendo un display con el que el usuario puede interactuar con objetos en un entorno que no necesariamente sigue las leyes del mundo físico: “es un espejo dentro de un país de maravillas de las matemáticas”. La descripción de su display incluye estímulos visuales y hápticos (Sutherland, 1965; Sherman et al., 2003, 26).

1968

Ivan Sutherland construye en la Universidad de Harvard el primer HMD estereoscópico con imágenes sintéticas generadas por ordenador, para formas de geometría muy sencilla. Este HMD era tan pesado, que se ideó un sistema de brazos anclados al techo para sujetarlo, de tal manera que el usuario no tuviera que soportar su peso (Figura 15). El display utilizaba mini tubos de rayos catódicos, con la óptica adaptada para presentar imágenes separadas a cada uno de los ojos y una interfaz de sensores mecánicos y ultrasónicos (Sherman et al., 2003; Kunsthochschule für Medien Köln, 2007).

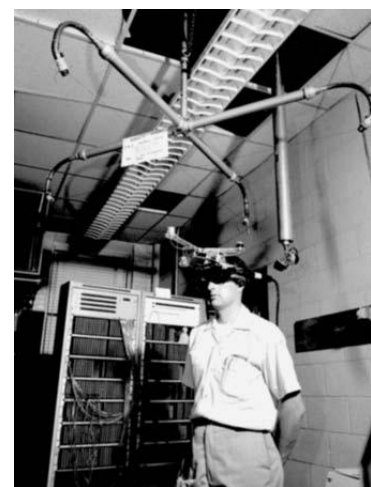


Figura 15. HMD estereoscópico de Sutherland, en (Update Centre, 2007).

1981

Se diseña *Super Cockpit* bajo la dirección de Tom Furness en la base militar aérea de Wright Patterson. Este sistema incluye un display de visión transparente (*see-through*) montado en el casco del piloto. Cuando el piloto mira en varias direcciones, la visión de su entorno se aumenta con diferente información. Por ejemplo, al mirar el ala se visualiza el número de misiles que quedan (Sherman et al., 2003, 29).

En el MIT, un grupo de investigadores integrado por Chris Schmandt, Eric Hulteen, Jim Zamiska y Scott Fisher, entre otros, empieza a trabajar con un display de realidad aumentada que permite a los usuarios explorar materias como dibujos 3D, visualización arquitectónica, y trazados 3D de chips de computador. El dispositivo utiliza espejos semitransparentes para superponer una imagen generada por ordenador sobre las manos u otras partes del cuerpo del usuario (Sherman et al., 2003, 29).

1992

La realidad virtual proyectada se introduce como una alternativa a los sistemas basados en HMD, en la conferencia de SIGGRAPH'92, en Chicago. El sistema más importante es CAVE (Cave Automatic Virtual Environment, Figura 16), desarrollada por el Electronic Visualization Lab de la Universidad de Illinois en Chicago (UIC), con una variedad de aplicaciones científicas y artísticas demostrando la tecnología (Sherman et al., 2003, 33).

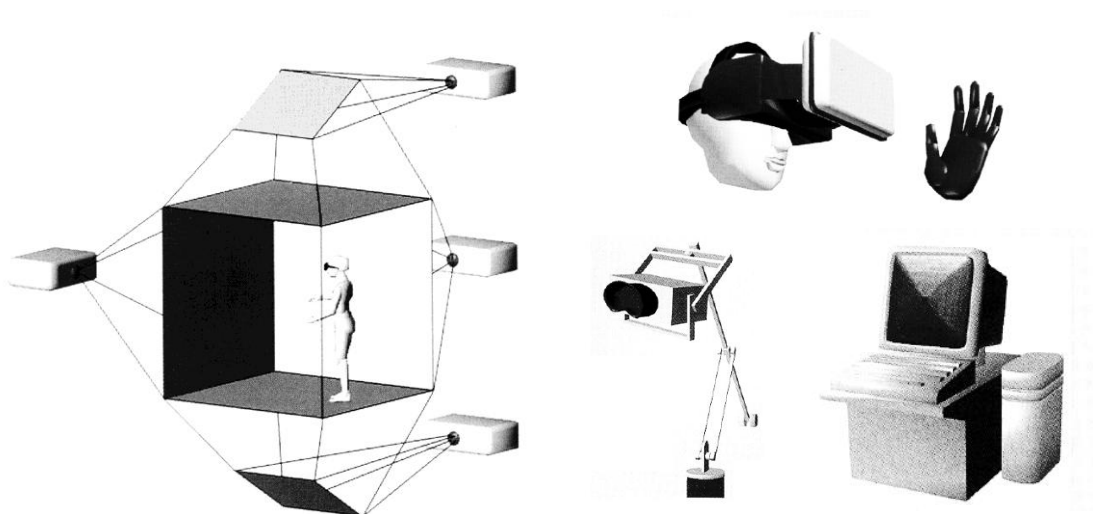


Figura 16. Dispositivos del sistema CAVE, en (Cruz-Neira et al., 2002).

En este mismo año Tom Caudell acuña el término de “realidad aumentada” en un artículo para el congreso HICCS (*IEEE Hawaii International Conference on System Science*), presentado conjuntamente con David Mizell, bajo el título de “Augmented Reality: an Application of Heads-Up Display Technology to Manual Manufacturing Processes”. El trabajo realizado está descrito en (Mizell, 2001; Sherman et al., 2003, 513-520), y a continuación se muestra un resumen por su relevancia en el establecimiento de la tecnología de RA: En la era de la manufactura, el grupo *Virtual Systems Research and Technology* del *Boeing Computer Services* elabora una aplicación para ayudar en la tarea del ensamblaje del cableado para aviones. El director del grupo es David Mizell y el proyecto fue concebido por Tom Caudell. La investigación se lleva a cabo en colaboración con el CMU, Honeywell, y Virtual Visión, Inc., bajo un programa de tres años de duración llamado *Wearable Computer Systems with Transparent, Head-Mounted Displays*, iniciado en agosto de 1994 y que finalizaría en el mismo mes del año 1997. El sistema de RA desarrollado provee al usuario con líneas (virtuales) superpuestas sobre el tablero de ensamblajes (real) mediante un HMD. En sus inicios, antes de tener financiación para el proyecto, se utilizó un HMD de bajo coste, denominado *Private Eye*; éste es un sistema monocular que abarca sólo una pequeña porción del campo de visión. En 1993 se elabora un prototipo con dos CRT (Cathode Ray Tube) estéreo y un sistema de registro magnético. En 1997, se presenta un prototipo que consiste en un sistema ligero de HMD y un ordenador puesto alrededor de la cadera, elaborado por TriSen Systems, utilizando visión por computador y marcas fiduciales para el registro del usuario.

1998

En este año se desarrolla ARToolKit (Human Interface Technology Laboratory, 2007), unas librerías de libre distribución diseñadas para aplicaciones de RA, entre la colaboración de la Universidad de Washington, Seattle, Human Interfaces Technology Laboratory AR y ATR Media Integration & Communication en Kyoto, Japón (ver apartado 5.1.1).

A partir de este año nacen los primeros congresos dedicados exclusivamente a realidad aumentada. A continuación se indican la fecha, lugar de celebración y nombre de estos:

- 1 Noviembre, 1998, San Francisco, CA (USA). IEEE International Workshop on Augmented Reality (**IWAR**).
- 9-11 Marzo, 1999, Yokohama (Japón). International Symposium on Mixed Reality (**ISMR**).
- 5-6 Octubre, 2000, Munich (Alemania). IEEE, ACM, and Eurographics International Symposium on Augmented Reality (**ISAR**).
- 30 Septiembre – 1 Octubre, 2002, Darmstadt (Alemania). International Symposium on Mixed and Augmented Reality (**ISMAR**). Este congreso unió los dos congresos de ISMR e ISAR.

2001

Oliver Bimber, junto con otros investigadores, presenta *The Virtual Showcase* (Bimber et al., 2001b; Bimber et al., 2003), un nuevo display óptico de realidad aumentada que permite que varios usuarios observen e interactúen con el contenido físico aumentado situado dentro de la caja del display (Figura 17). El efecto se consigue mediante la combinación de espejos semitransparentes divisores del haz (*half-silvered mirror beam splitters*), que permiten presentaciones estereoscópicas y superposiciones 3D junto a los elementos físicos.

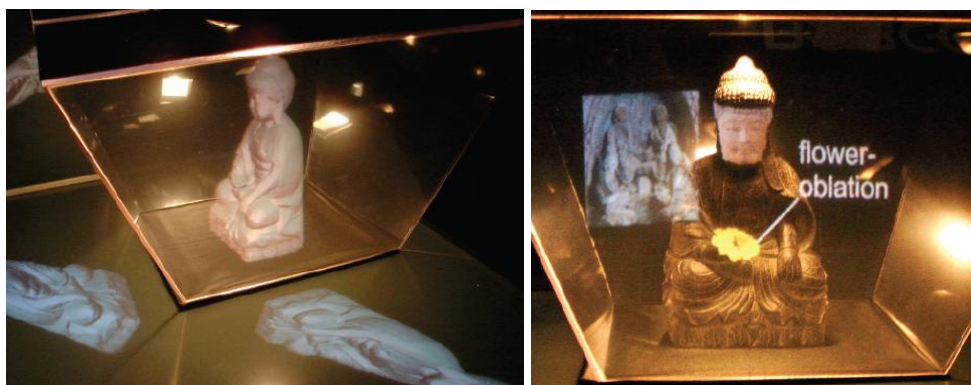


Figura 17. *The Virtual Showcase*, en (Bimber et al., 2001b).

2002

Desde este año y hasta el 2004 se pone en marcha el proyecto AMIRE patrocinado por el programa EU IST. AMIRE (Figura 18) es un software de autor basado en las librerías ARToolKit, y que permite la creación y modificación de aplicaciones de realidad aumentada de manera sencilla, sin la necesidad de programación (AMIRE, 2004a). Este software se retoma en el apartado 5.1.7.

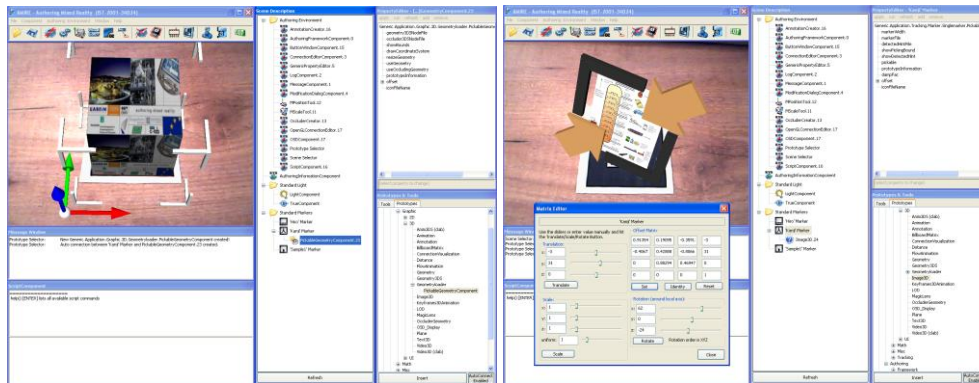


Figura 18. Ejemplos de utilización y GUI del AMIRE.

2003

Se empieza a implantar en Europa la tercera generación de la telefonía móvil (3G), permitiendo disponer en el móvil de aplicaciones y servicios nuevos, caracterizados por una alta calidad de imagen y una elevada velocidad de transmisión de datos. El incremento de capacidad de almacenamiento de los móviles, la integración de cámaras digitales y la posibilidad de localización con RFID, hacen de éste un dispositivo idóneo para ser utilizado en aplicaciones de RA móviles. Un claro ejemplo es el trabajo realizado en el Instituto Computer Graphics and Vision, de la Graz University of Technology al desarrollar ARToolKitPlus (2003-2006), unas librerías hechas a partir de ARToolKit para introducir aplicaciones de RA en los teléfonos móviles y PDAs (Wagner et al., 2007).

2006

Nace Wii (Figura 19), la quinta videoconsola casera de Nintendo. El dispositivo más atrayente es su controlador inalámbrico, *Wii Remote*, que se puede utilizar como puntero y puede detectar aceleración y orientación en 3D. Este control remoto es un dispositivo relativamente económico que puede utilizarse en aplicaciones de realidad aumentada para capturar las rotaciones de la cámara.



Figura 19. Mando de la consola Wii.

1.4. DESARROLLOS Y APLICACIONES

La realidad aumentada se ha introducido en diversos campos, abarcando desde los ámbitos de la medicina y la ingeniería, hasta el arte y la educación. Dado que esta investigación se dirige principalmente a las áreas de aplicación del arte, entretenimiento, educación y edutainment, se abordarán diversos casos de estudio dirigidos a estas áreas en concreto, y a continuación se mencionan someramente algunos ejemplos dirigidos a otras áreas, para comprender la amplitud de la tecnología de realidad aumentada.

1.4.1. ARTE

Podemos encontrar diversos ejemplos de piezas de realidad aumentada y realidad mezclada en el ámbito del arte, como por ejemplo en (Shaw, 1989; Hoberman et al., 1998; Koleva et al., 2001; Paul, 2003; Galantay et al., 2004; Wagner et al., 2004a; Paulos et al., 2005; Torpus, 2005; Levin, 2006; Díaz, 2007). En muchos de estos trabajos se entremezclan habilidades artísticas y técnicas, surgiendo interesantes colaboraciones entre artistas, informáticos, tecnólogos, ingenieros, etc. A continuación se muestran algunas piezas con más detalle.

Can You See Me Now? (2001) es un proyecto que surge de la colaboración del grupo Blast Theory y Mixed Reality Laboratory de la Universidad de Nottingham (UK). Consiste en un juego que se desarrolla en las calles de una ciudad, en el que diversos jugadores están conectados en-línea y son perseguidos a través de un modelo virtual de la ciudad por los corredores (actores profesionales equipados con tecnología GPS y WiFi) que tienen que correr físicamente por las calles de la ciudad para dar caza a los

jugadores (Benford et al., 2006). Conceptualmente, el grupo Blast Theory pretendía indagar en el ámbito de penetración que tiene la tecnología móvil, incluso en las zonas más pobres, entornos rurales, adolescentes y, en general, en demografías que habitualmente no tienen a su alcance las nuevas tecnologías (Blast Theory, 2007). La pieza se exhibió por primera vez en el festival b.tv de Sheffield. En el 2002 fue nominada para un premio del Interactive Arts BAFTA en UK, y en el 2003 obtuvo el Golden Nica en Ars Electrónica, Linz (Austria). En las siguientes imágenes puede verse un usuario interactuando con el sistema.



Figura 20. Imágenes de un usuario interactuando con la pieza *Can You See Me Now?* (2001), en (Blast Theory, 2007).

The Hidden Worlds of Noise and Voice (2002) de Golan Levin y Zachary Lieberman es una instalación audiovisual interactiva de realidad aumentada, instalada durante dos años en la exposición de Ars Electronica Museum of the Future en Linz, Austria. El tema central es la relación mágica entre el habla y el medio etéreo que lo transporta. Los usuarios son capaces de ver una representación de sus voces, mediante figuraciones gráficas animadas que emergen de sus bocas al hablar, a través de un HMD. Cuando uno de los usuarios habla o canta, aparecen diversos gráficos de colores que emergen de su boca, representando dichos sonidos y adquiriendo una gran variedad de formas y comportamientos que están relacionados con las cualidades únicas de su voz (volumen, altura y timbre). Esta instalación permite la participación simultánea de seis usuarios, con un amplio rango de inmersión audiovisual y juego conversacional. Aquellos participantes que no estén equipados con los HMD, pueden ver una proyección en la mesa que representa la sombra de los gráficos generados (Figura 21).

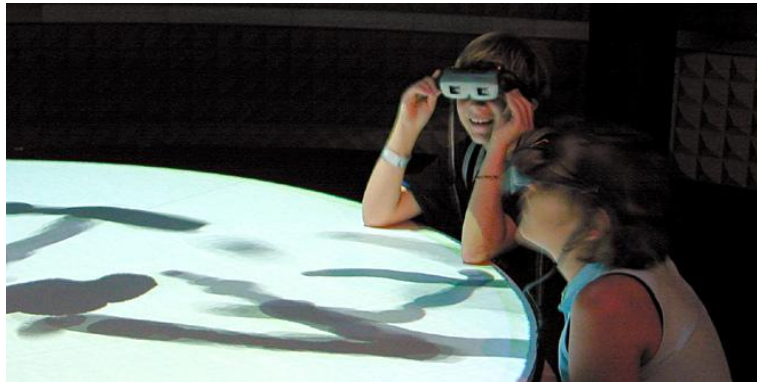


Figura 21. Usuarios interactuando con *The Hidden Worlds of Noise and Voice*. En (Levin et al., 2008).

Red Libre Red Visible (2004) de Diego Díaz y Clara Boj, es un proyecto que nace en el Interaction and Entertainment Research Centre, de la Nanyang Technological University de Singapur, y en el que colaboran los investigadores Liu Wei, Xu Ke y Dui Ta. En esta pieza se construye un escenario de realidad aumentada en el entorno urbano mediante el empleo de las librerías MXRToolKit, que usa las posibilidades de las nuevas tecnologías para crear nuevos paisajes en el espacio público, a través de la visualización de los datos que fluyen por las redes digitales (Díaz, 2007, 273). Se buscan nuevas configuraciones virtuales, personales y subjetivas, y se fomenta el acceso e intercambio de información a través de una red libre y gratuita (Figura 22).



Figura 22. Imágenes de la pieza *Red Libre Red Visible* (2004) de Diego Díaz y Clara Boj, en (Díaz, 2007).

LifeClipper (2004), desarrollada por un grupo de investigadores y artistas de University of Applied Sciences Northwestern Switzerland y coordinados por Jan Torpus, es una pieza de realidad aumentada móvil que consiste en una experiencia audiovisual enriquecida con elementos virtuales (Figura 23). Técnicamente se basa en un ordenador portátil que el usuario debe llevar recorriendo un área específica de interés cultural. Un sensor GPS registra en todo momento su posición y dirección, y mediante un HMD el usuario ve y escucha una serie de imágenes y sonidos superpuestas al entorno real, ambos

relacionados con el lugar hacia donde está mirando. Según (Torpus, 2005), los usuarios se sienten como si estuvieran viendo una película en la que ellos participan como observadores activos, y en la que los diversos personajes virtuales se dirigen a ellos.



Figura 23. Imágenes de usuarios interactuando con la pieza *lifeClipper* (2004), en (Torpus, 2005).

Desde los inicios de las conferencias y exhibiciones del ACM SIGGRAPH (Association for Computing Machinery's Special Interest Group on Graphics and Interactive Techniques), una de sus partes esenciales ha estado integrada por artistas e investigadores que han utilizado la computación como parte de su proceso creativo. En SIGGRAPH'06 se presentó *U-Tsu-Shi-O-Mi: The Virtual Humanoid You Can Reach* (2006), de Michihiko Shoji (Shoji et al., 2006), un sistema de realidad aumentada que sincroniza un robot humanoide y un avatar virtual (Figura 24). Los usuarios pueden experimentar las sensaciones que se producen al darle la mano a un avatar generado por ordenador. Para lograr el registro, se ha desarrollado un sistema de control integrado (mediante sensores ópticos y de fuerza), de tal forma que los movimientos del robot y del avatar virtual permanezcan sincronizados durante el contacto físico con los usuarios. La mezcla de los entornos real y virtual en la imagen video vista por el usuario, se consigue mediante un croma.



Figura 24. *U-Tsu-Shi-O-Mi: The Virtual Humanoid You Can Reach* (2006), de Michihiko Shoji, en (SIGGRAPH06, 2006).

1.4.2. ENTRETENIMIENTO

Dentro del campo del entretenimiento existe una gran diversidad de sistemas de realidad aumentada. Algunos ejemplos los podemos encontrar en (Bobick et al., 1999; Badiqué et al., 2002; Cheek et al., 2002b; Andersen et al., 2004; Cooper et al., 2004; Charles et al., 2004; Matyszczok et al., 2004; Romero et al., 2004; Avery et al., 2005; Barakonyi et al., 2005; Bimber et al., 2005b; Benford et al., 2006; Brandt, 2006; Lam et al., 2006; Park et al., 2006; Peitz et al., 2006; Rashid et al., 2006). No es de extrañar el creciente interés que despiertan estas aplicaciones en este campo; basta con tener en cuenta la gran industria del videojuego que existe en la actualidad con los sistemas de realidad virtual, y las grandes posibilidades que puede ofrecer la realidad aumentada como tecnología novedosa.

Una aplicación dentro de este campo es *Touch-Space* (2002). Según (Cheek et al., 2002b) este juego está basado en la tecnología de realidad mezclada que captura la interacción de los usuarios en el entorno real como elementos principales del juego, mientras que mantiene al mismo tiempo las características de los mundos ficticios propios de aplicaciones de entretenimiento por ordenador (Figura 25). El juego se lleva a cabo en un entorno físico del tamaño de una habitación, que permite la interacción háptica usuario-usuario y usuario-entorno. Su narrativa sigue una historia concreta: una princesa es capturada por una bruja; los jugadores tienen que encontrar dos piezas de un mapa y una serie de tesoros; a continuación tienen que sobrevolar la Tierra y buscar el castillo; una vez encontrado, tienen que luchar y vencer a la bruja; finalmente, deben entrar al castillo para rescatar a la princesa y cumplir la misión.

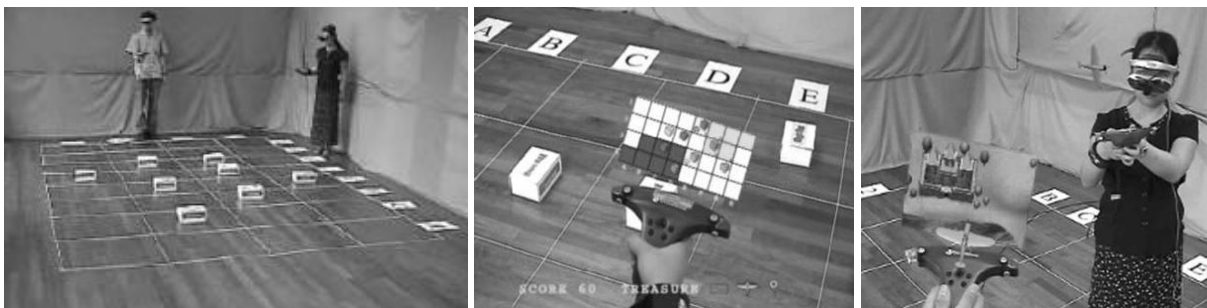


Figura 25. Imágenes de *Touch-Space* (2002), en (Cheek et al., 2002b).

Human Pacman (2003) es una aplicación de RA móvil que nació en el Mixed Reality Lab de la National University of Singapore (NUS). Está basada en el conocido videojuego *Pacman* de Arcade de los años 80. Según sus autores, es un sistema pionero en las aplicaciones dentro del campo del entretenimiento, basándose en la materialización

del juego, movilidad, interacción social y computación ubicua (Figura 26). Los jugadores se mueven físicamente por un área urbana, realizando tareas para lograr sus objetivos (comer galletas), llevando consigo ordenadores portátiles, HMD y GPS. El juego se desarrolla tanto fuera como dentro de un recinto, y los jugadores pueden interactuar de manera presencial con otros jugadores cuando físicamente están próximos, y de forma remota mediante una red inalámbrica (LAN) (Cheek et al., 2003; Cheek et al., 2004).

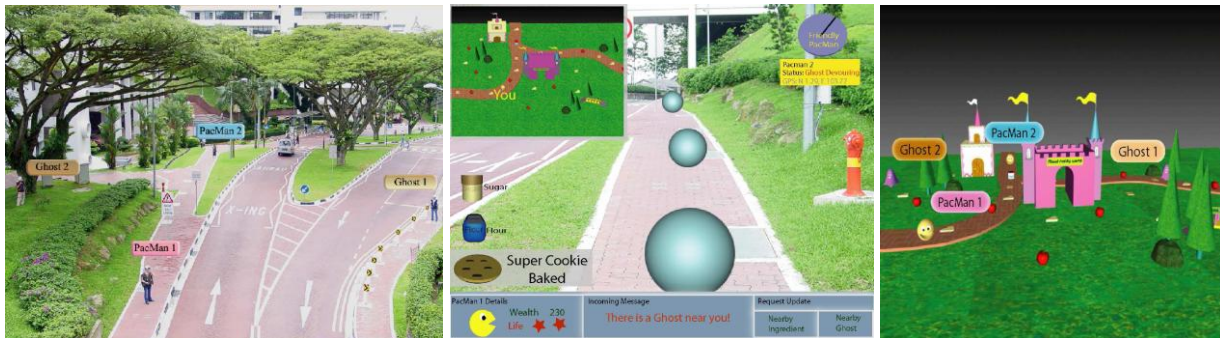


Figura 26. Imágenes de *Human Pacman*, en (Cheek et al., 2003).

Otra aplicación que se puede catalogar dentro del campo del entretenimiento es *The Invisible Train* (2004). Según sus autores, *The Invisible Train* es la primera aplicación de RA multi-usuario para PDAs (Wagner et al., 2004a). La aplicación consiste en un juego en el que los usuarios controlan una serie de trenecitos virtuales que circulan sobre una maqueta de madera que contiene los raíles (Figura 27). Los usuarios interactúan con el juego controlando los cambios de vías y la velocidad de los trenes virtuales, y están sincronizados mediante una red inalámbrica. El objetivo del juego es evitar que los trenes colisionen. Este proyecto ha ganado diversos premios, entre los que destaca el del Japan Media Arts Festival en el año 2004; también se ha expuesto en numerosas ocasiones, destacando su participación en el festival de Ars Electronica 2004 en Linz (Austria) y en el Wired NextFest 2005, en Chicago (EEUU).



Figura 27. Usuarios interactuando con *The Invisible Train* (2004), en (Wagner et al., 2004a).

En (Park et al., 2006) se presenta un juego de tenis de mesa aumentado mediante el empleo de las librerías ARToolkit para el registro de las acciones de los usuarios (Figura 28). En este juego, sus autores implementan propiedades físicas para establecer la trayectoria de la pelota virtual de acuerdo al lanzamiento de cada jugador y los rebotes en la mesa física.

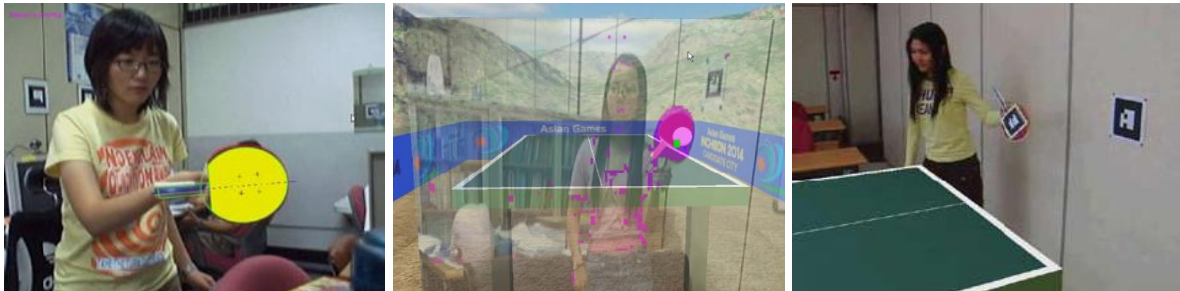


Figura 28. Usuarios interactuando con una mesa de tenis aumentada, en (Park et al., 2006).

PAC-LAN (2006) es una aplicación de RM móvil basada, al igual que en el caso de *Human Pacman*, en el tradicional videojuego *Pacman* de Arcade. En este caso, los jugadores tienen un teléfono móvil que les sirve como display, dispositivo de registro (mediante RFID), y elemento de comunicación entre usuarios (Figura 29), permitiendo mayor libertad de movimiento (Rashid et al., 2006).



Figura 29. Usuarios jugando con PAC-LAN y visualizaciones en el móvil, en (Rashid et al., 2006).

1.4.3. EDUCACIÓN

Vinculadas al campo de la educación podemos encontrar también un creciente número de aplicaciones, como las presentadas en (Hall et al., 2002; Kaufmann, 2002; Kaufmann et al., 2003; Gillet et al., 2004; Kaufmann, 2004; Lu et al., 2004; Roussou, 2004; Klopfer et al., 2005; Garzotto et al., 2006; Pan et al., 2006). En algunas ocasiones también se utiliza el término inglés de *e-learning* para referirse a sistemas de aprendizaje basados en tecnología multimedia.

Un claro ejemplo en este ámbito es *Construct3D* (2002) descrita en (Kaufmann, 2002; Kaufmann et al., 2003; Kaufmann, 2004), una aplicación de realidad aumentada específicamente diseñada para la educación en el campo de las matemáticas y la geometría, combinando cuatro áreas de investigación: geometría, pedagogía, psicología y realidad aumentada. Es un sistema que trata de maximizar la transferencia de conocimientos y que permite varios escenarios de interacción entre alumnos y profesores, implementando métodos flexibles para el contexto y visualizaciones individualizadas de algunas partes. Se han implementado tres modos diferentes: modo independiente, en el que cada estudiante sólo puede ver los objetos construidos por él mismo; modo colaborativo, visible para todos; y modo de profesor, en el que el profesor puede seleccionar la visibilidad mediante un panel de interacción personal. De este modo, sus autores proponen una clase aumentada, en la que los alumnos llevan puestos unos HMD y pueden interactuar con diversos objetos virtuales, de acuerdo con las indicaciones del profesor que lleva el control de la actividad propuesta (Figura 30).

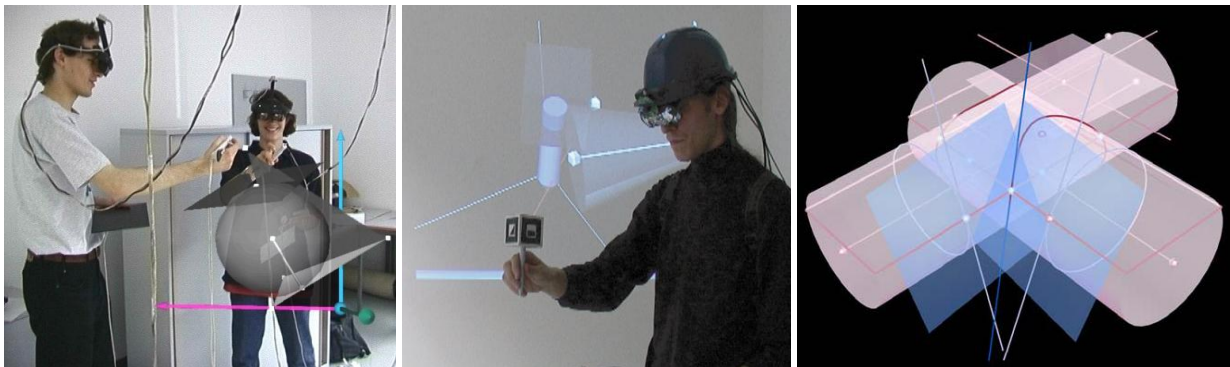


Figura 30. Usuarios interactuando con *Construct3D*, en (Kaufmann, 2004).

En (Gillet et al., 2004) se describe un sistema de RA para visualizar modelos físicos de estructuras biológicas complejas (Figura 31). La aplicación utiliza las marcas de ARToolKit para la localización espacial de la cámara y el establecimiento del entorno aumentado, adaptado para visualizar los modelos tridimensionales a partir del programa Python Molecular Viewer. Los usuarios pueden cambiar de forma sencilla el contenido visualizado mediante distintas representaciones de una molécula, visualización de sus propiedades, o de información dinámica. Según sus autores, el modelo físico representado proporciona una interfaz intuitiva para manipular los modelos computacionales, haciendo eficiente la interacción entre las decisiones humanas, el modelo físico, y la actividad computacional.

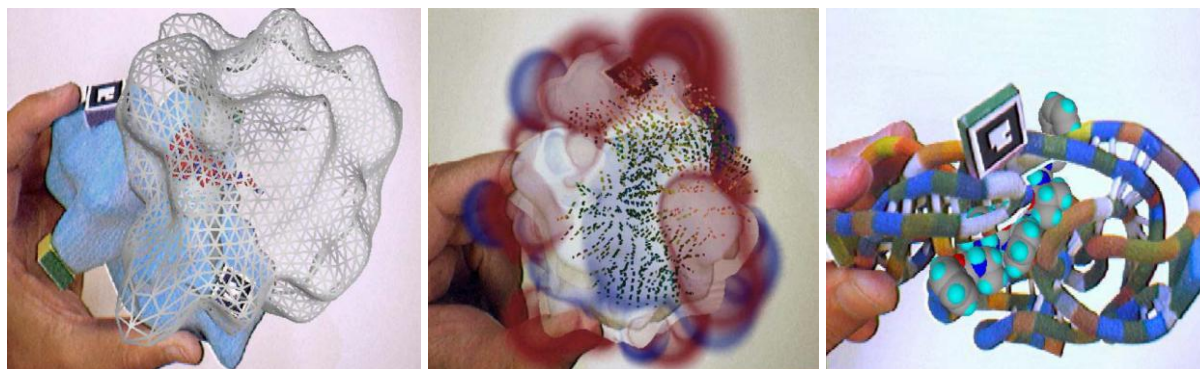


Figura 31. Distintas visualizaciones moleculares. En (Gillet et al., 2004).

1.4.4. EDUTAINMENT

El término *edutainment* nace de la fusión de educación y entretenimiento (**education + entertainment**). Estos sistemas se basan en el aprendizaje mediante el juego, con el objetivo de ser más eficaces y atractivos. Algunos trabajos de RA dentro de este campo los podemos encontrar descritos en (Bobick et al., 1997; Bobick et al., 1999; Billinghamurst et al., 2001a; Cheok et al., 2002a; Song et al., 2004; Wojciechowski et al., 2004; Woods et al., 2004; Cheok et al., 2006; De Souza e Silva et al., 2006; Garzotto et al., 2006; MXR, 2006).

Una de las aplicaciones más conocidas de realidad aumentada en este campo es *Magic Book* (2000), que se mostró por primera vez en una exhibición de tecnologías emergentes en el congreso SIGGRAPH del año 2000. *Magic Book* tiene imágenes y texto al igual que un libro tradicional, pero está aumentado con diversos elementos generados por ordenador mediante la tecnología de ARToolKit (Billinghurst et al., 2001a). Los usuarios pueden leer el libro conjuntamente, interactuando mientras lo leen y viendo al mismo tiempo las imágenes virtuales desde distintos ángulos a través de un HMD. Los usuarios también pueden adentrarse en el entorno virtual al girar un pequeño interruptor que activa un sensor inercial (*InterTrax* de InterSense) que registra sus movimientos, llegando a formar parte de la acción. Si más de un usuario se adentra en la escena, los participantes se ven unos a otros como avatares (Figura 32). Según Billinghamurst (ScienceDaily, 2000), la aplicación crea un continuo por el que puede moverse la gente y pasar de un entorno puramente real a uno aumentado, o sentirse inmersos dentro de un entorno completamente virtual. *Magic Book* puede contener desde un cuento para niños a un manual técnico, un libro de texto o cualquier otra narración. Uno de los contenidos más conocidos trata de un cuento infantil que narra la historia de una princesa japonesa, un guerrero samurai y un ninja demoníaco que

roba el espejo de la princesa. Al tiempo que los distintos personajes son presentados, los usuarios pueden verlos mediante una representación tridimensional con animación y sonido: la princesa llora porque no tiene su espejo; el guerrero samurai trata de buscar su caballo; y el ninja practica ejercicios de combate.



Figura 32. Usuarios interactuando con *MagicBook*, en (Technology Research News, 2000; HITLab, 2006).

En (MXR, 2006) se describen varios sistemas que pretenden educar a través del juego, permitiendo a los usuarios la interacción colaborativa en un entorno aumentado para explorar diversos aspectos relacionados con la ciencia, matemáticas, idioma inglés y otras materias (Figura 33). Dos ejemplos los podemos encontrar en *Solar System* (2005) y *Plant Story* (2006). En el primero se explican varios conceptos relacionados con el sistema solar, como la composición del Sol y de los diversos planetas, la geometría espacial para que se produzca un eclipse lunar, la disposición de los distintos planetas orbitando alrededor del Sol, etc. En el segundo se invita al usuario a que plante una planta en una maceta, la riegue y la ponga al Sol para que crezca y se produzca la fotosíntesis.

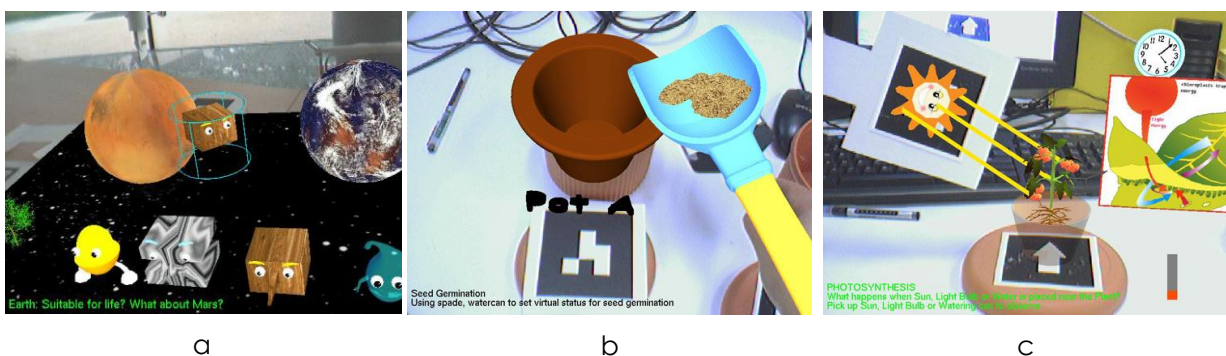


Figura 33. Aplicaciones de RA en el campo del edutainment: a) *Solar System*; b) y c) *Plant Story*.

En (MXR, 2006).

1.4.5. MEDICINA

En el campo de la medicina es importante que los sistemas de realidad aumentada utilizados tengan un alto grado de precisión y fiabilidad en el registro, sobre todo si estos sistemas se van a emplear en operaciones que puedan resultar críticas para el paciente. A pesar de la poca madurez de la tecnología de realidad aumentada, como se ha explicado en la introducción de esta tesis, existen muy diversas actuaciones dentro del campo de la medicina, como por ejemplo las explicadas en (Nikou et al., 2000; Weidenbach et al., 2000; Rosenthal et al., 2002; Aschke et al., 2003; Balasingham et al., 2003; Fischer et al., 2004; Gelenbe et al., 2005; Konishi et al., 2005; Wacker et al., 2005; Mischkowski et al., 2006; Fischer et al., 2007).

En (Glossop et al., 2003) se presenta un dispositivo de realidad aumentada que utiliza un láser escáner para mostrar información directamente sobre los pacientes. El sistema también es capaz de medir la localización espacial de la información proyectada mediante una cámara estereoscópica para asegurar la precisión, aplicar las correcciones necesarias a la proyección y ayudar en el registro. El sistema de proyección tiene su propio interfaz de programación de aplicaciones (API). En la Figura 34 se puede ver como el sistema es probado sobre un cráneo de plástico. Tras realizar varios ensayos, se estimó que el error en la proyección era de 1 mm. Según sus autores, este sistema se puede emplear para la identificación de estructuras anatómicamente críticas, visualización de trayectorias preestablecidas, y aplicaciones de tele-cirugía y tele-consulta.

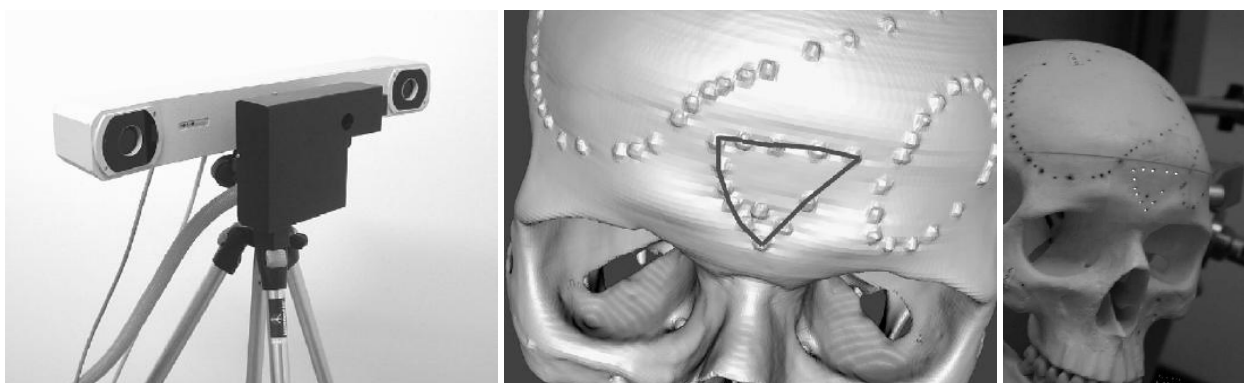


Figura 34. Sistema de cámara-proyector y ejemplos de funcionamiento del sistema sobre un cráneo de plástico, en (Glossop et al., 2003).

En (Wacker et al., 2005) se hace un estudio para adaptar un sistema de RA a un escáner médico, y se prueba su viabilidad para utilizarlo en un tipo determinado de

biopsias (Figura 35). El sistema de RA empleado consiste en un HMD de tipo vídeo con dos cámaras a color, lo que proporciona una imagen estereoscópica, y una tercera cámara que se utiliza para el registro. Ésta cámara calcula la posición del usuario en relación con una serie de marcas ópticas situadas sobre la mesa del escáner. Las pruebas se realizaron sobre muñecos y se obtuvo un resultado óptimo en 11 biopsias, y una pequeña desviación de 1.5 mm en 9 biopsias. Finalmente se establece que el método desarrollado puede ser de ayuda en este tipo de operaciones.

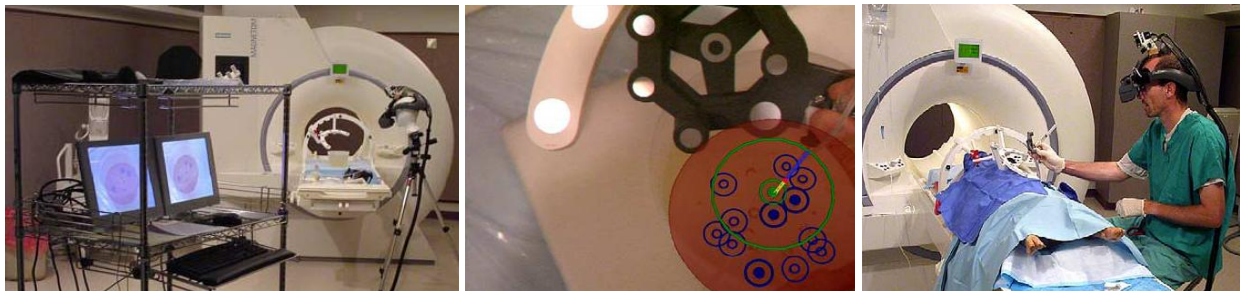


Figura 35. Sistema de RA adaptado a un escáner médico, en (Wacker et al., 2005).

1.4.6. PSICOLOGÍA

En el campo de la psicología cabe destacar el trabajo desarrollado en la Universidad Politécnica de Valencia por (Juan et al., 2005; Juan et al., 2006a; Juan et al., 2006b; Juan et al., 2007; Pérez et al., 2007), que es pionero en el desarrollo de sistemas de RA para el tratamiento de distintos trastornos psicológicos. En (Juan et al., 2005) se presenta un sistema de RA para el tratamiento de fobias a animales pequeños como cucarachas y arañas (Figura 36), iniciado con un proyecto final de carrera en la facultad de informática de la UPV, realizado por el alumno César Carrión Villaplana en el año 2004. Esta aplicación se probó con pacientes reales que, después de alrededor de una hora de tratamiento con este sistema, fueron capaces de interactuar y matar animales reales. Se realizaron otras dos versiones de este sistema, uno con marcas invisibles (Juan et al., 2006b) y otro mediante un HMD óptico (Juan et al., 2007). Ambos sistemas se evaluaron para presencia, ansiedad y juicio de la realidad con usuarios sin fobia.

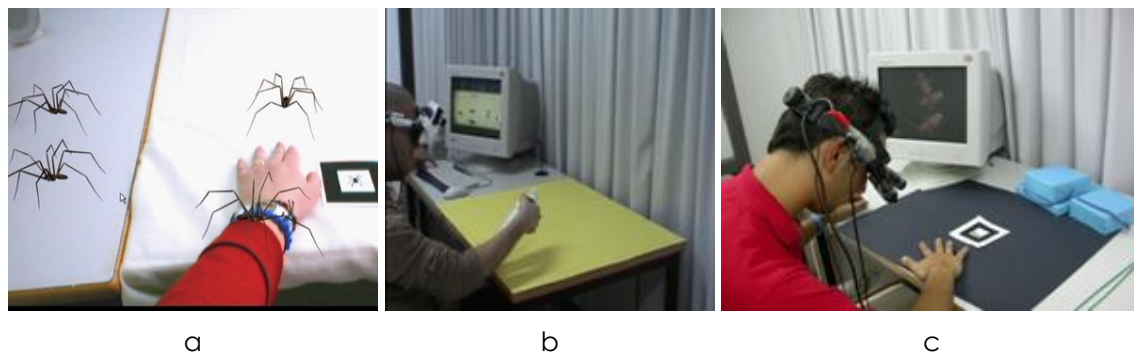


Figura 36. Tratamiento de fobias: a) Primer sistema de RA para tratar fobias a insectos; b) Sistema de RA sin marcas; c) Sistema de RA mediante HMD óptico.

También se desarrollaron dos sistemas de RA para el tratamiento del miedo a las alturas (Figura 37), evaluados de igual forma que los trabajos anteriores; el primero de ellos incluye, como elementos virtuales, fotografías inmersivas; en el segundo los elementos virtuales están mezclados con el entorno real de una habitación, simulando que el suelo se cae y las paredes se elevan. El sistema de fotos navegables se comparó con un entorno real (Juan et al., 2006a), mientras que el segundo se comparó con un sistema de realidad virtual con las mismas características (habitación real modelada), (Pérez et al., 2007).

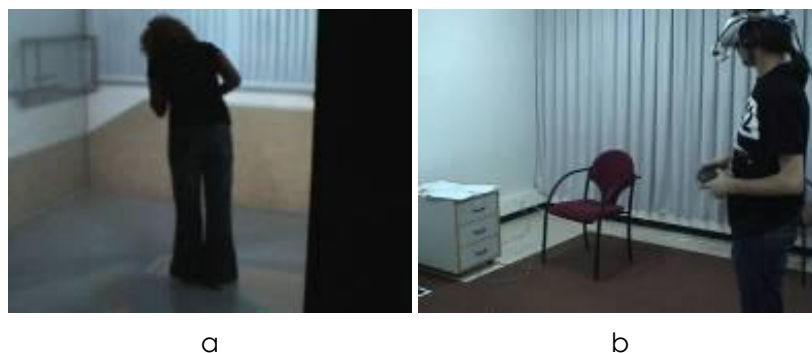


Figura 37. Sistemas para el tratamiento de vértigo: a) Usuario inmerso en un sistema CAVE; b) Sistema de RA.

1.4.7. INGENIERÍA Y ARQUITECTURA

Dentro del campo de la ingeniería y la arquitectura se pueden encontrar diversos ejemplos, como los descritos en (Ben-Joseph et al., 2001; Zlatanova, 2001; Bimber et al., 2003; Kaufmann et al., 2003; Broll et al., 2004; Gillet et al., 2004; Kaufmann, 2004; King et al., 2005; Kealy et al., 2006).

En (Ben-Joseph et al., 2001) se presenta *Luminous Planning Table* (Figura 38), un sistema de realidad aumentada que utiliza una interfaz computerizada tangible, aplicado a la planificación de zonas urbanas. En su día constituyó un sistema novedoso, tanto por su diseño como por la presentación de los procesos, ya que proporciona un medio integrado que muestra conjuntamente los modelos físicos y las simulaciones y animaciones generadas por ordenador, mediante un sistema de proyectores.

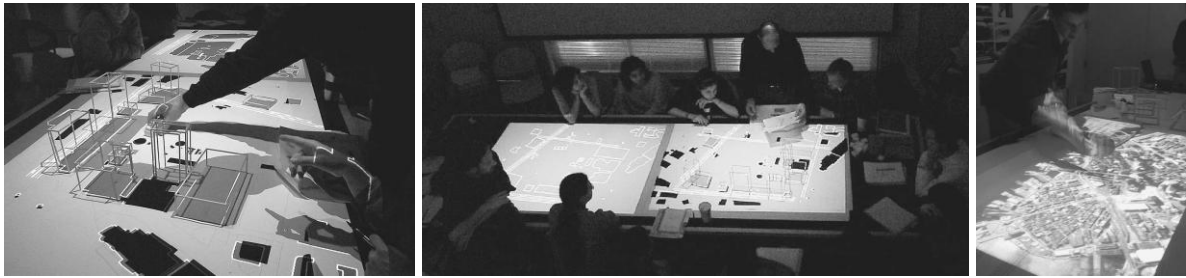


Figura 38. Usuarios interactuando con el sistema *Luminous Planning Table*, en (Ben-Joseph et al., 2001).

Otro ejemplo lo encontramos descrito en (King et al., 2005), donde se presenta *ARVino* (Figura 39), un sistema móvil de RA diseñado para asistir a los viticultores en la visualización in situ de diversos parámetros que afectan directamente a sus plantaciones y la calidad de las viñas. El sistema es una combinación de realidad aumentada y GIS (Geographic Information System).



Figura 39. Ejemplos del sistema *ARVino*, en (King et al., 2005).

1.4.8. ROBÓTICA Y TELERROBÓTICA

Dentro del área de la robótica y telerrobótica podemos encontrar algunos sistemas de RA descritos en (Halme et al., 1999; Shaheen et al., 2001; Marchand et al., 2002; Aschke et al., 2003; Stilman et al., 2005; Nawab et al., 2007).

En (Stilman et al., 2005) se presenta una aplicación para robots experimentales que permite realizar pruebas unificadas de subsistemas individuales al actuar como parte de un sistema integrado por componentes reales y virtuales (Figura 40). Se integran los últimos avances en velocidad y precisión de captura del movimiento mediante sistemas ópticos para determinar la localización del robot, registrar los objetos del entorno, y extraer los parámetros de orientación externa en tiempo real de diversas cámaras en movimiento. De este modo se construyen correspondencias espaciales y temporales entre elementos virtuales y reales. Los resultados del sistema se muestran mediante una aplicación on-line para el desarrollo de robots humanoides autónomos.



Figura 40. Ejemplos de funcionamiento del sistema para robots, en (Stilman et al., 2005).

En (Shaheen et al., 2001) se presenta un sistema de realidad aumentada desarrollado para reconstruir escenas tridimensionales a partir de la visión de una sola cámara (Figura 41). La cámara está previamente calibrada con respecto al sistema de referencia considerado. El modelo tridimensional se conoce y está correctamente ubicado con respecto a su imagen bidimensional. Los diversos algoritmos desarrollados se aplican en el campo de la telerrobótica para manipular un objeto (que el usuario señala en la imagen) por medio de un robot con 4 grados de libertad (*DOF, degree of freedom*).

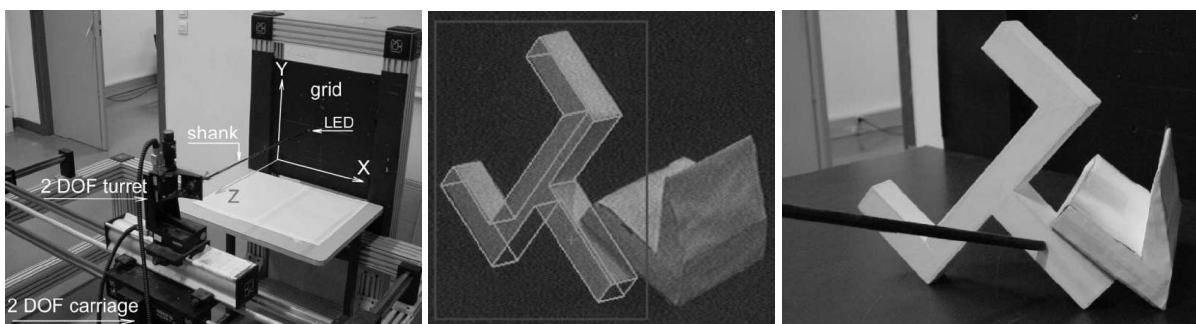


Figura 41. Robot y modelos virtual y real, en (Shaheen et al., 2001).

1.4.9. INDUSTRIA

En el campo de la industria podemos encontrar diversos ejemplos descritos en (Ciulli et al., 1998; Behringer et al., 1999; Barfield et al., 2001a; Bass et al., 2001; Anastassova et al., 2005; Dangelmaier et al., 2005; Wang et al., 2006).

En (Dangelmaier et al., 2005) se describe la arquitectura de un sistema de realidad aumentada para dar soporte al proceso de planificación de sistemas de manufactura complejos (Figura 42). El sistema asiste al usuario en el modelado, la validación del modelo de simulación, y la consiguiente optimización de la producción del sistema.



Figura 42. Sistema de RA en procesos de manufactura, en (Dangelmaier et al., 2005).

En (Behringer et al., 1999) se muestra un sistema de realidad aumentada diseñado por el Rockwell Science Center (RSC), que simula visión de rayos X sobre objetos reales, mediante la superposición de modelos 3D, animaciones y anotaciones de texto sobre una imagen vídeo del objeto (Figura 43). Un sistema de reconocimiento del habla permite al usuario preguntar sobre el estado de los componentes del dispositivo, posibilitando a su vez que las anotaciones verbales queden registradas en forma de texto. La posición del usuario con respecto al objeto se registra mediante un sistema de visión por computador basado en marcas fiduciales.

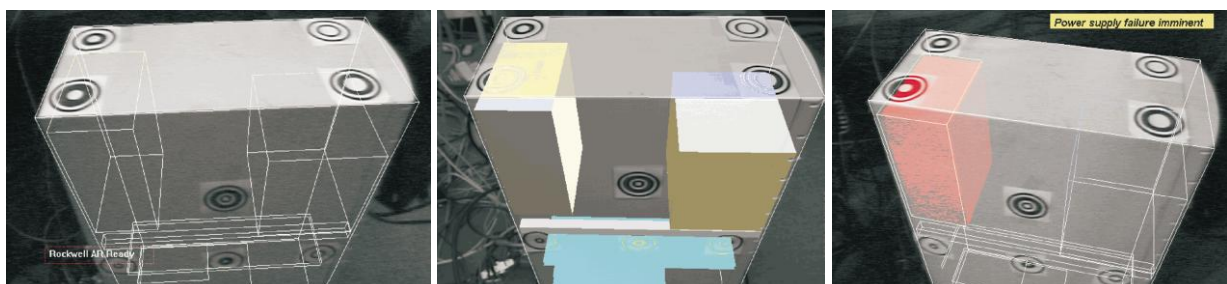


Figura 43. Dispositivo de RA, en (Behringer et al., 1999).

1.4.10. MILITAR

La industria militar es quizá uno de los campos que más ha impulsado diversas tecnologías relacionadas con la RA. Por ejemplo, tanto los vestidos computacionales como los dispositivos móviles para llevarlos en las manos tienen una procedencia de uso militar. De igual modo, el GPS se desarrolló para ser utilizado por el ejército de los Estados Unidos, y gran parte del software de simulación con tecnología de realidad virtual y aumentada se ha desarrollado para la industria militar. Podemos encontrar algunos ejemplos de aplicaciones en este campo en (Tappert et al., 2001; Hicks et al., 2002; Livingston et al., 2002; Hicks et al., 2003).

En (Hicks et al., 2002; Hicks et al., 2003) se presenta un proyecto en el que diversos iconos gráficos se superponen a la señal de vídeo de la mira del arma de un soldado (Figura 44). El proyecto, denominado *EyeKon*, es un sistema inteligente que soporta decisiones basadas en agentes, ubicado en el ordenador que lleva el soldado. *EyeKon* utiliza la posición del soldado para proporcionar una vista en perspectiva en la mira del arma.



Figura 44. Visualizaciones de EyeKon, en (Hicks et al., 2002; Hicks et al., 2003).

En (Livingston et al., 2002) se desarrolla *BARS* (Battlefield Augmented Reality System), un sistema que consiste en un ordenador para llevar, un sistema en red inalámbrico, y un HMD, que está pensado para operaciones militares en un entorno urbano (Figura 45). El ordenador genera gráficos que, desde la perspectiva del usuario, aparecen alineados con el entorno real. Por ejemplo, un edificio puede ser aumentado mostrando su nombre, un plano de su interior, iconos que representen la localización de francotiradores, y los nombres de las calles adyacentes.



Figura 45. Visualizaciones del sistema BARS y usuario equipado, (Livingston et al., 2002).

2. PRINCIPIOS MATEMÁTICOS

Como hemos visto en el apartado 1.2.1, una de las características que definen la RA es el registro tridimensional, de tal forma que los mundos reales y virtuales queden espacialmente alineados. Para que ello sea posible, es necesario considerar las transformaciones espaciales que relacionan los distintos sistemas de referencia que intervienen que son: el sistema de coordenadas bidimensional de la imagen capturada por la cámara; el sistema de coordenadas tridimensional del mundo físico; y el sistema de coordenadas tridimensional de mundo virtual. Una vez establecidas dichas relaciones espaciales, la escena aumentada es geoméricamente plausible. Además, existen otros factores que pueden considerarse para que la escena aumentada sea más verosímil, como son la oclusión y la luminosidad o sombreado.

2.1. RELACIONES ESPACIALES ENTRE EL ENTORNO 3D Y LA IMAGEN 2D

2.1.1. SISTEMAS DE COORDENADAS

Un **sistema de coordenadas** es un conjunto de valores que permiten definir inequívocamente la posición de cualquier punto de un espacio euclídeo. Normalmente los sistemas de coordenadas son ortogonales y se distinguen entre (Lerma García, 2002, 53-55):

1. Sistema de **coordenadas imagen** y sistema **espacial de coordenadas imagen** (Figura 46): El sistema de coordenadas imagen (x_{img} , y_{img}) es ortogonal, está centrado en el punto principal de la imagen y teóricamente está exento de errores sistemáticos, como por ejemplo los producidos por la distorsión de la lente. La posición de un punto en la imagen viene definido por sus coordenadas bidimensionales. El sistema espacial de coordenadas imagen se utiliza para definir la posición de un punto de la imagen en el espacio. Éste es un sistema tridimensional que toma como origen el centro de proyección de la cámara; el eje z_e coincide con la dirección del eje óptico; los ejes x_e e y_e pasan por el centro de proyección y son paralelos a los ejes x e y del sistema de coordenadas imagen.

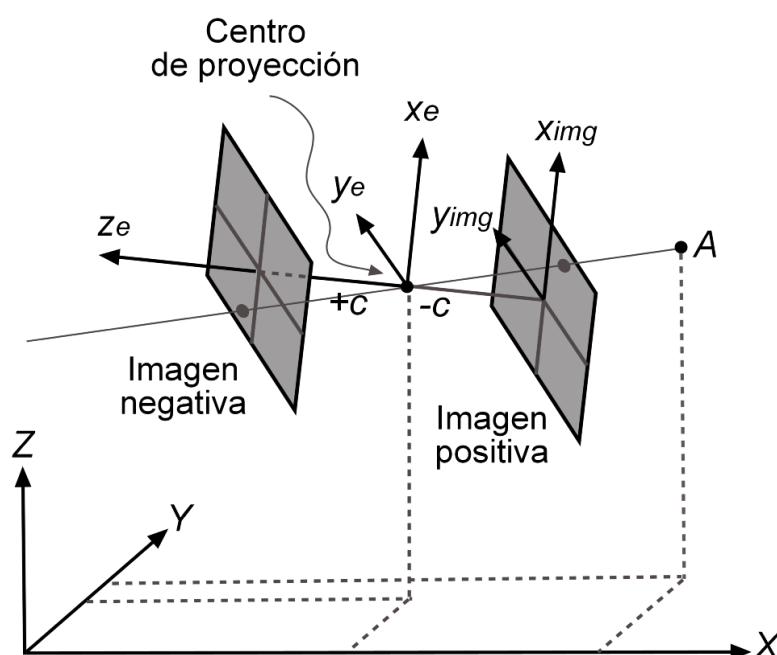


Figura 46. Sistema de coordenadas imagen y sistema espacial de coordenadas imagen.

2. Sistema de coordenadas **terreno/objeto**: El sistema de coordenadas terreno u objeto permite definir la posición espacial de puntos pertenecientes a la superficie terrestre o a cualquier otra superficie. Éste es un sistema cartesiano tridimensional definido por las coordenadas (X,Y,Z) y puede ser:
- a. **Terrestre**: Son sistemas de coordenadas fijos o anclados al terreno, y se puede distinguir entre:
 - i. Sistema global (Figura 47-a), cuando se toma en consideración la Tierra y por lo tanto el sistema de coordenadas viene dado de antemano (ej. sistema elipsoidal, sistema astronómico, etc.).
 - ii. Sistema local (Figura 47-b), cuando se toma en consideración una pequeña proporción terrestre, como por ejemplo una habitación, y el sistema de coordenadas se puede establecer arbitrariamente.
 - b. **Centrado en objetos** (Figura 47-c): Son sistemas de coordenadas situados en objetos, y por lo tanto no dependen del entorno en que esté situado dicho objeto. El objeto puede permanecer estático o estar en movimiento con respecto a la cámara, y se pueden dar distintos casos en aplicaciones de RA:
 - i. El objeto y la cámara permanecen estáticos (su posición relativa será constante en el tiempo),
 - ii. La cámara está fija y el objeto en movimiento,
 - iii. La cámara está en movimiento y el objeto fijo,
 - iv. Tanto el objeto como la cámara están en movimiento.

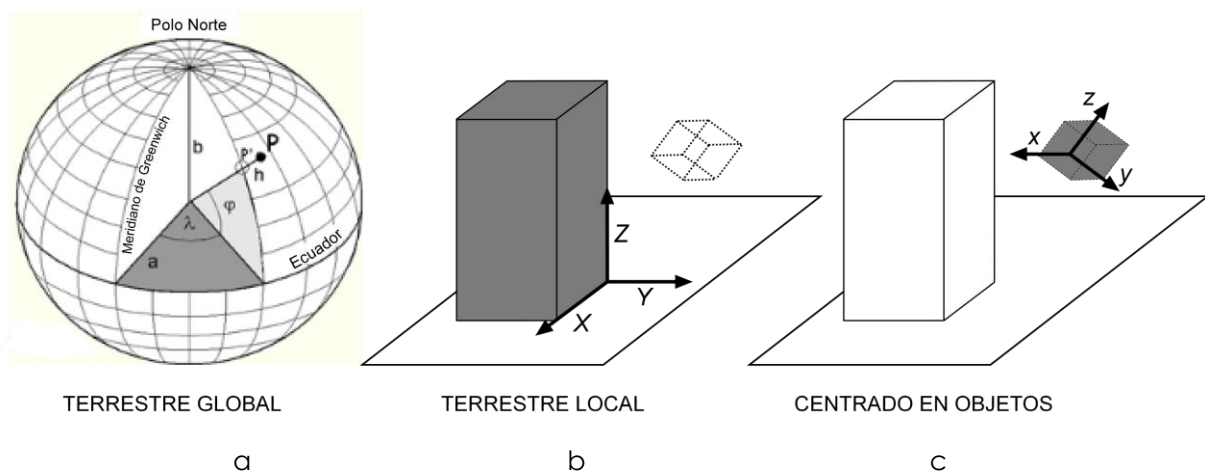


Figura 47. Sistemas de coordenadas: a) Terrestre global; b) Terrestre local; c) Centrado en objetos.

Los casos b.ii, b.iii y b.iv se pueden considerar como uno sólo desde el punto de vista matemático, puesto que los tres implican que la posición relativa entre la cámara y el objeto no permanece constante, y por lo tanto es necesario el registro de la orientación externa en todo momento. De otro modo, en el caso b.i sólo será necesaria la previa calibración del sistema cámara-objeto.

En los sistemas centrados en objetos, la no dependencia del entorno físico puede suponer una gran ventaja, si lo comparamos con los sistemas terrestres que implican una transformación adicional, es decir, la orientación externa de la cámara y del objeto deben conocerse en el sistema terrestre para posteriormente poder obtener la orientación externa relativa cámara-objeto, necesaria para la construcción del entorno aumentado. El problema se reduce considerablemente cuando la cámara o el objeto están fijos, puesto que sólo será necesaria la obtención de la orientación externa en todo momento del objeto móvil, mientras que el otro bastará con calibrarlo de antemano. Es por ello que el registro en sistemas de referencia terrestres puede presentar limitaciones, puesto que implica que los objetos del entorno estén calibrados con respecto del sistema de referencia del sensor y que, después de la calibración, estos no se muevan (Neumann, 2001, 329-333) – o de moverse, tanto el objeto como la cámara deberían llevar sensores, con el consiguiente encarecimiento del sistema, mayor complejidad de cálculo y disminución de precisiones (ya que los sensores deben estar calibrados/alineados entre sí).

Por otra parte, existen una serie de características propias de los sistemas terrestres globales que los hacen apropiados para muchas aplicaciones. La principal característica o ventaja es la posibilidad de abarcar grandes extensiones de terreno, por lo que son idóneos para aplicaciones abiertas (ver apartado 1.2.2). En este sentido, el registro en los sistemas de referencia terrestres locales o centrados en objetos suelen estar limitados al entorno de los objetos y/o al entorno visto por la cámara, siendo más utilizados en aplicaciones dentro de recintos.

A continuación se muestra una tabla comparativa entre los sistemas de registro terrestres y los centrados en objetos:

Sistemas de registro terrestres	Sistemas de registro centrados en objetos
Asumen que los objetos están calibrados con respecto a la infraestructura	Permite que los objetos registrados se muevan libremente en el entorno
El registro puede abarcar grandes extensiones	El registro es local al entorno de los objetos
Pueden haber grandes infraestructuras de registro que sean partes permanentes del entorno	Los objetos llevan consigo una infraestructura de registro mínima, que tiene pequeño impacto en el entorno
La resolución y precisión están en unidades terrestres fijas relativas a los componentes del sistema de registro (ej. centímetros o pulgadas)	La resolución y precisión son relativas a la visión que la cámara tiene de los objetos (los errores en unidades de píxel varían en un rango de dimensiones terrestres)
El registro de las oclusiones es dificultoso	El registro de las oclusiones se puede llevar a cabo de manera semiautomática

Tabla 2. Comparación de dispositivos de registro. En (Neumann, 2001, 333).

Además de las consideraciones generales de sistemas de referencia terrestres o centrados en objetos (dentro del entorno real), hay que señalar que el entorno virtual también tendrá su propio sistema de referencia, el cual deberá coincidir con el real o, de otro modo, deberá conocerse la matriz de transformación que los relaciona.

2.1.2. PARÁMETROS DE LA ORIENTACIÓN EXTERNA

Los parámetros de orientación externa de la cámara establecen la posición del centro de proyección en el sistema de coordenadas terreno/objeto (X_o , Y_o , Z_o) y la orientación angular del sistema espacial de coordenadas imagen en el sistema de coordenadas terreno/objeto (Lerma García, 2002, 55). La orientación angular puede definirse mediante tres ángulos independientes, siendo varias las posibles combinaciones de orientación según el sentido de giro (dextrógiro, levógiro) y el orden de sucesión de los giros, aunque también puede expresarse directamente con la matriz de rotación. Existen además diversos sistemas de rotación: *inclinación-giro-acimut* (*tilt-swing-azimuth*), *omega-phi-kappa* (ω, ϕ, χ) y el sistema *balanceo-cabeceo-guiñada* (*roll-pitch-yaw*). Este último es el sistema empleado por el sensor inercial. Estos parámetros se pueden obtener a partir de una sola imagen mediante un proceso de resección espacial (utilizando la condición de colinealidad o el método de Church).

También se pueden obtener a partir de múltiples imágenes mediante triangulación espacial (por el método de los haces de rayos).

2.1.3. ECUACIONES DE LA PROYECCIÓN CENTRAL

Las ecuaciones de la proyección central o ecuaciones de colinealidad establecen una relación matemática entre un punto u objeto físico (en el entorno real) y su representación en la imagen bidimensional. Se basan en el hecho de que, una vez tomada la imagen, un punto del terreno/objeto, el centro de proyección y la imagen del punto están alineados (Kraus, 1997, 19; Lerma García, 2002, 56). La relación espacial de la cámara respecto al sistema de referencia terrestre/objeto puede describirse en la Figura 48 y explicarse mediante la Ecuación 1:

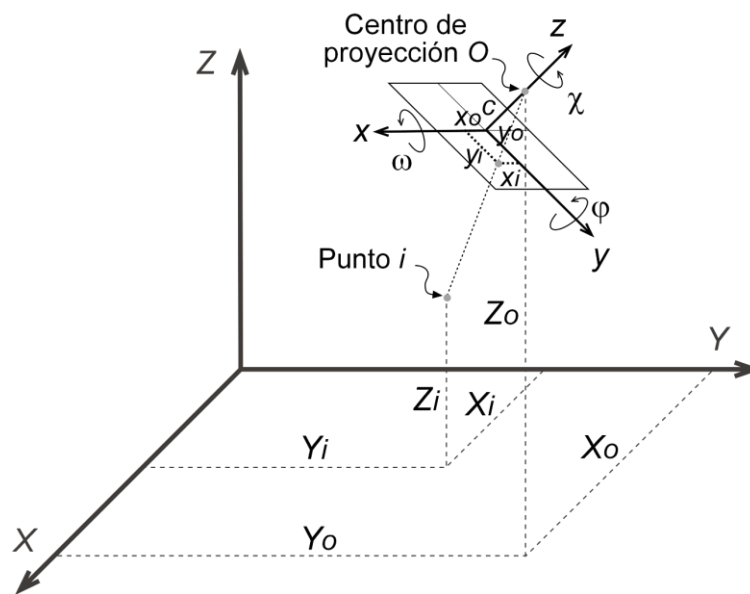


Figura 48. Esquema de la perspectiva central. Relación espacial entre un punto físico, su representación bidimensional en la imagen, y el centro de proyección de la cámara.

Rotaciones de la cámara con respecto al sistema terrestre.

$$x_i = x_o - c \frac{r_{11}(X_i - X_o) + r_{12}(Y_i - Y_o) + r_{13}(Z_i - Z_o)}{r_{31}(X_i - X_o) + r_{32}(Y_i - Y_o) + r_{33}(Z_i - Z_o)}$$

$$y_i = y_o - c \frac{r_{21}(X_i - X_o) + r_{22}(Y_i - Y_o) + r_{23}(Z_i - Z_o)}{r_{31}(X_i - X_o) + r_{32}(Y_i - Y_o) + r_{33}(Z_i - Z_o)}$$

Ecuación 1. Ecuaciones de colinealidad.

Donde (x_i, y_i) son las coordenadas imagen de un punto i , (X_i, Y_i, Z_i) son las coordenadas tridimensionales (en el sistema de referencia terrestre) de un punto i , (x_0, y_0) son las coordenadas del punto principal, c es la distancia focal de la cámara (debe permanecer constante) y (X_0, Y_0, Z_0) son las coordenadas tridimensionales del centro de proyección O de la cámara. Los términos r_{ij} son los elementos de la matriz de orientación tridimensional que describe las rotaciones de la cámara en el sistema de referencia terrestre.

La matriz de orientación es una matriz cuadrada y ortogonal compuesta de nueve elementos, de los que sólo tres son independientes. La matriz de orientación que se suele emplear en los sistemas de RA está basada en la rotación de tres giros (por ej. ω, φ, χ) alrededor de los ejes (X, Y, Z) , y por eso también se suele llamar matriz de rotación (Kraus, 1997, 14-16; Lerma García, 2002, 62-63). Ésta tiene la forma:

$$\mathbf{R}_{\omega\varphi\chi} = \mathbf{R}_{\chi}\mathbf{R}_{\omega}\mathbf{R}_{\varphi} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}$$

Ecuación 2. Matriz de orientación.

Cuyos términos r_{ij} son:

$$\begin{aligned} r_{11} &= \cos \chi \cos \varphi \\ r_{12} &= \sin \varphi \sin \omega \cos \chi - \sin \chi \cos \omega \\ r_{13} &= \sin \varphi \cos \omega \cos \chi + \sin \omega \sin \chi \\ \\ r_{21} &= \sin \chi \cos \varphi \\ r_{22} &= \sin \chi \sin \omega \sin \varphi + \cos \chi \cos \omega \\ r_{23} &= \sin \varphi \cos \omega \sin \chi - \sin \omega \cos \chi \\ \\ r_{31} &= -\sin \varphi \\ r_{32} &= \cos \varphi \sin \omega \\ r_{33} &= \cos \varphi \cos \omega \end{aligned}$$

Ecuación 3. Términos de la matriz de orientación.

A continuación se muestra un ejemplo realizado con el programa AMIRE, en el que se observan las rotaciones de un objeto 3D y su correspondiente matriz de rotación:

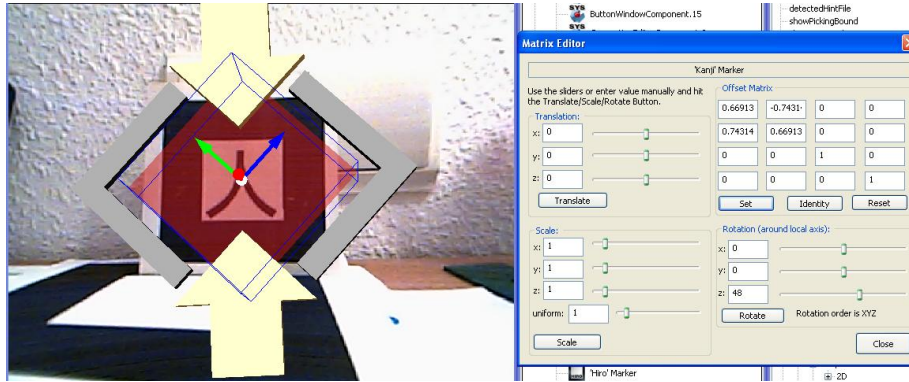


Figura 49. Rotaciones de un elemento virtual con respecto al sistema de referencia de la marca, y su correspondiente matriz de rotación; Rotaciones angulares (0, 0, 48).

2.1.4. ESTÉREO-VISIÓN: GEOMETRÍA EPIPOLAR

Algunos displays de realidad aumentada, generalmente ciertos HMD, son estereoscópicos. Aunque en esta tesis no se han empleado dichos dispositivos, debido a su relevancia, a continuación pasamos a describir brevemente los principios de la visión estereoscópica y la geometría epipolar.

El mecanismo de visión humano es estéreo, es decir, somos capaces de apreciar, a través de la visión binocular (Figura 50), las diferentes distancias y volúmenes en el entorno que nos rodea. Nuestros ojos, debido a su separación (distancia interpupilar, d), obtienen dos imágenes del entorno con pequeñas diferencias entre ellas, a lo que denominamos disparidad o paralaje. Nuestro cerebro procesa las diferencias entre ambas imágenes y las interpreta de forma que percibimos la sensación de profundidad, lejanía o cercanía de los objetos que nos rodean, en un proceso denominado estereopsis. La distancia interpupilar más habitual es de 65 mm, pero en algunas personas puede variar desde los 45 a los 75 mm.

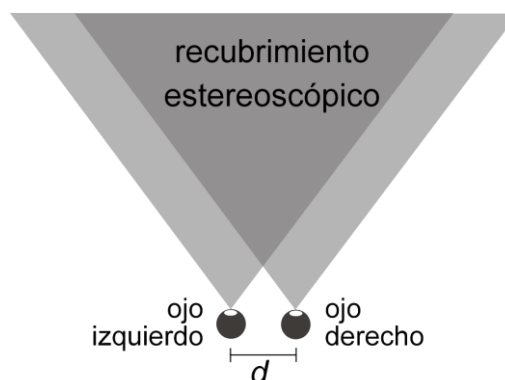


Figura 50. Visión binocular.

En la estereopsis intervienen diversos mecanismos. Cuando observamos objetos muy lejanos, los ejes ópticos de nuestros ojos son paralelos. Cuando observamos un objeto cercano, nuestros ojos giran para que los ejes ópticos estén alineados sobre él, es decir, convergen. A su vez se produce la acomodación o enfoque para ver nítidamente el objeto. Este proceso conjunto se llama fusión. No todo el mundo tiene la misma capacidad de fusionar un par de imágenes en una sola tridimensional. Alrededor de un 5% de la población tiene problemas de fusión. La agudeza estereoscópica es la capacidad de discernir, mediante la estereopsis, detalles situados en planos diferentes y a una distancia mínima. Hay una distancia límite a partir de la cual no somos capaces de apreciar la separación de planos, y que varía de unas persona a otras. Así, la distancia límite a la que dejamos de percibir la sensación estereoscópica puede variar desde unos 60 metros hasta cientos de metros.

La geometría epipolar es el constreñimiento básico que deriva de la existencia de dos puntos de vista. Supongamos que dos cámaras toman una imagen de una misma escena, bajo dos puntos de vista distintos (Figura 51).

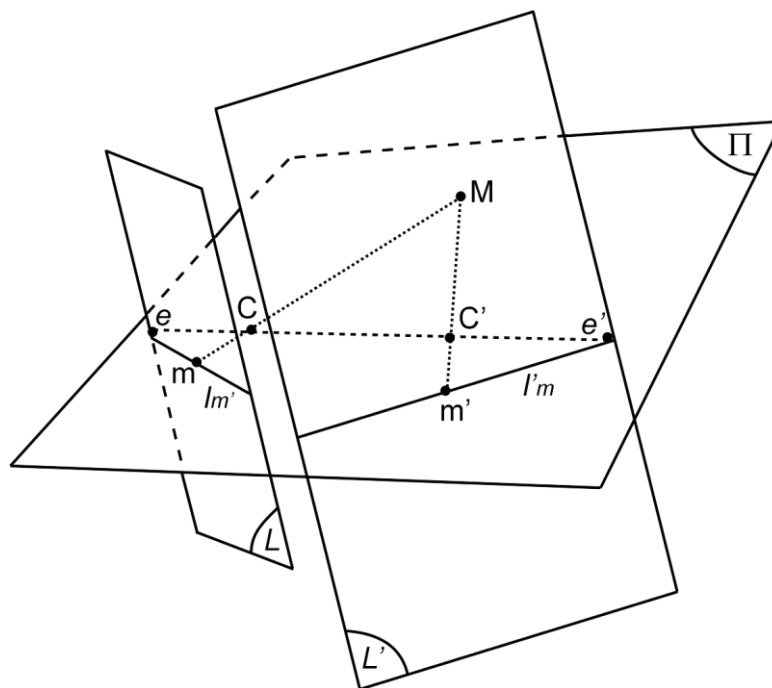


Figura 51. Geometría epipolar. A partir de (Gruen et al., 2001, 201).

Siendo C el centro óptico de la primera cámara y C' el centro óptico de la segunda cámara. La línea (C, C') se proyecta en un punto e en la primera imagen L , y en un punto e' en la segunda imagen L' . Los puntos e y e' son los epípolos. La línea a través de e en la primera imagen y la línea a través de e' en la segunda imagen son las líneas epipolares. Los centros ópticos de las cámaras, un punto cualquiera del terreno M y sus

puntos homólogos correspondientes en las imágenes, m y m' , se encuentran en el mismo plano epipolar \square . La constante epipolar es bien conocida en la visión estereoscópica: para cada punto m en la primera retina, su correspondiente punto m' está sobre la línea l'_m (Gruen et al., 2001, 201).

2.1.5. ORIENTACIÓN ESPACIAL

Según el tipo de sensores y la metodología empleada para el cálculo de los parámetros de la orientación externa de la cámara, podemos distinguir varios tipos de sistemas:

1. Métodos **directos**: Obtención de parámetros de la orientación externa directamente a través de los sensores, como un receptor GPS (registro de la posición espacial) o un INS (registro de las rotaciones).
2. Métodos **indirectos**: Cálculo de la orientación externa, mediante técnicas y algoritmos matemáticos, a partir de los valores capturados por un sensor. Este apartado hace especialmente referencia a las cámaras (*vision-based tracking*), y a la necesidad de utilizar técnicas de visión por computador o de fotogrametría para el cálculo de la orientación externa a partir del análisis de la imagen.
3. Métodos **mixtos**: Combinación de los anteriores. Como por ejemplo, un sistema combinado de INS y cámara que registra marcas fiduciales.

Cada tipo de sensor registra unos datos distintos o de distinta manera, con lo que el cálculo posterior para obtener la orientación externa varía de unos a otros. A continuación se explican únicamente los casos que se han empleado en esta tesis.

2.1.5.1. MÉTODOS DIRECTOS

La orientación externa consta de un total de seis parámetros: tres de traslación respecto al origen de coordenadas establecido y tres de rotación alrededor de los ejes de dicho sistema. A continuación se explican la obtención de los parámetros de posición a partir de un GPS y de orientación a partir de un sensor inercial.

2.1.5.1.1. Posición a partir de un receptor GPS

Los datos que el usuario obtiene directamente a partir de un receptor están en formato NMEA (National Marine Electronics Association) y se muestran como texto de la siguiente manera:

```
$GPVTG,,T,,M,0.112121,N,0.207643,K*4A
$GPGGA,183613.270,3924.1171,N,00023.9433,W,1,04,3.8,83.4,M,51.8,M,0.0,0000*51
$GPGSA,A,3,04,25,10,13,,,,,,,,,12.1,3.8,11.5*3E
$GPGSV,2,1,08,27,71,115,00,25,59,065,48,02,53,264,32,08,50,176,00*79
$GPGSV,2,2,08,04,47,202,38,13,44,047,34,10,26,306,38,23,08,065,00*7A
$GPRMC,183615.270,A,3924.1173,N,00023.9434,W,0.127173,,160208,,*0D
```

Figura 52. Datos NMEA de un receptor BTGPS II Trine de EMTAC.

De todos estos datos lo que realmente interesa es la longitud y latitud, que nos indicarán en qué lugar del globo terrestre se encuentra nuestro receptor y podremos transformar en coordenadas lineales X, Y. Por otro lado, la altitud nos proporcionará la coordenada Z. Estos datos se pueden obtener por ejemplo a partir de la línea \$GPGGA (señalado en negrita), que a continuación se analiza con más detalle:

```
$GPGGA,183613.270,3924.1171,N,00023.9433,W,1,04,3.8,83.4,M,51.8,M,0.0,0000*51

Donde:
GGA          Global Positioning System Fix Data
183613.270   Fix taken at 18:36:13.270 UTC
3924.1171,N  Latitude 39 deg 24.1171' N
00023.9433,W Longitude 0 deg 23.9433' W
1            Fix quality: 0 = invalid
                    1 = GPS fix (SPS)
                    2 = DGPS fix
                    3 = PPS fix
                    4 = Real Time Kinematic
                    5 = Float RTK
                    6 = estimated (dead reckoning) (2.3 feature)
                    7 = Manual input mode
                    8 = Simulation mode
04           Number of satellites being tracked
3.8          Horizontal dilution of position
83.4,M      Altitude, Meters, above mean sea level
51.8,M       Height of geoid (mean sea level) above WGS84 ellipsoid
0.0          time in seconds since last DGPS update
0000         DGPS station ID number
*51          the checksum data, always begins with *
```

Figura 53. Significado de la línea \$GPGGA.

Para entender los datos de altura dados por el sensor, hay que tener en cuenta tres superficies distintas: la superficie terrestre, la del geode y la del elipsoide de referencia (Figura 54). La altura que el GPS considera es en realidad la que existe entre el receptor y el elipsoide de referencia (altura elipsoidal, h). Sin embargo, con un buen modelo de ondulación del geode (distancia N en el dibujo) se puede calcular fácilmente la altura ortométrica ($H=h+N$) o altitud sobre el nivel medio del mar. A partir de la línea \$GPGGGA se obtienen tanto la altura elipsoidal como la ondulación del geode. La altura que nos interesa es la ortométrica, aunque debido a las pobres precisiones que ofrece esta medida en el caso del GPS, en aplicaciones de RA se emplea una altura media constante, que suele ser la altura de la cámara sobre el mismo terreno donde se desarrolla la aplicación.

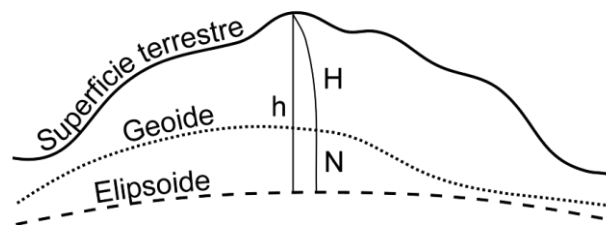


Figura 54. Superficies de referencia.

De toda esta información, lo que nos interesa realmente son las coordenadas espaciales de longitud y latitud. Hay que tener en cuenta que estas coordenadas están referidas al elipsoide de referencia, que en el caso del GPS es el GRS80 (sistema geodésico WGS84), definidos como valores angulares. Normalmente, este sistema de coordenadas (denominado geodésico: latitud, longitud y altura elipsoidal) no se utiliza directamente debido a la inconveniencia de tener que expresar los valores en unidades no lineales (Lerma García, 2002, 55). Para el paso de estos valores angulares a valores lineales (coordenadas X , Y) se procede del siguiente modo (Guillem Picó et al., 2001):

$$s = a(1 - e^2) \left[A(\varphi_2 - \varphi_1) - \frac{B}{2}(\sin 2\varphi_2 - \sin 2\varphi_1) + \frac{C}{4}(\sin 4\varphi_2 - \sin 4\varphi_1) \right]$$

Ecuación 4. Longitud de un arco de meridiano.

$$s' = N \cos \varphi (\lambda_2 - \lambda_1)$$

Ecuación 5. Longitud de un arco de paralelo.

Donde:

$$A = 1 + \frac{3}{4}e^2 + \frac{45}{64}e^4 + \dots$$

$$B = \frac{3}{4}e^2 + \frac{15}{16}e^4 + \dots$$

$$C = \frac{15}{64}e^4 + \dots$$

$$N = \frac{a}{(1 - e^2 \sin^2 \varphi)^{1/2}}, \text{ radio transverso de curvatura.}$$

$a = 6378137$, valor de semieje mayor del elipsoide GRS80.

$e = 0,081819$, valor de excentricidad del elipsoide GRS80.

2.1.5.1.2. Rotaciones a partir de un sensor INS

El tipo de datos que registra el sensor son los siguientes:

[s]	[m/s ²]			[rad/s]			[a.u.]			[°C]
time	Xacc	Yacc	Zacc	Xgyr	Ygyr	Zgyr	Xmag	Ymag	Zmag	temp
0	-2,787496	-0,476608	9,423267	-0,033508	0,009019	0,00896	0,217609	-0,377115	-0,740355	24,6875
0,01	-2,783835	-0,42474	9,453347	-0,018159	0,012469	0,002843	0,221124	-0,377775	-0,745848	24,6875
0,02	-2,782214	-0,429893	9,440937	-0,012043	0,025411	-0,006204	0,220242	-0,377805	-0,738753	24,6875
0,03	-2,785613	-0,442034	9,444966	-0,005883	0,00941	0,002022	0,211589	-0,375691	-0,743848	24,6875
0,04	-2,794121	-0,43157	9,439033	0,004916	0,011328	-0,004901	0,219906	-0,376536	-0,741448	24,6875
0,05	-2,780413	-0,438665	9,462478	0,003299	0,009465	0,001257	0,221038	-0,376903	-0,741009	24,6875
0,06	-2,768454	-0,44745	9,477598	-0,004406	0,018961	-0,001671	0,217742	-0,375721	-0,748213	24,6875
0,07	-2,773601	-0,416131	9,457727	-0,012593	-0,000312	0,041324	0,228294	-0,374853	-0,754547	24,6875
0,08	-2,794032	-0,416014	9,449859	-0,008531	0,000828	-0,014397	0,212478	-0,374974	-0,740433	24,6875

Tabla 3. Datos que registra un INS; donde, de izquierda a derecha: Tiempo medido en segundos; Aceleraciones en los tres ejes medido en m/s²; Velocidad de los giros en los tres ejes medidos en rad/s; Valor del campo magnético en los tres ejes medidos en a.u. (intensidad del campo magnético: 1 a.u. = 235052 Wb/m²); Temperatura en el interior del sensor medida en °C.

A partir de estos datos, el propio sensor calcula las rotaciones, y muestra los resultados de forma matricial (nueve componentes r_{ii}), en cuaterniones (q_0, q_1, q_2, q_3), o como ángulos de Euler (*roll, pitch, yaw*).

2.1.5.2. MÉTODOS INDIRECTOS

En los métodos de registro basados en la imagen (*vision-based*), la obtención de la orientación externa de la cámara se realiza a partir del análisis de una o más imágenes. Se han desarrollado diversos sistemas de registro los cuales se pueden clasificar, atendiendo a los parámetros de entrada, en (Neumann, 2001, 333):

- Tres o más puntos con coordenadas 3D conocidas (que deben ser visibles en una imagen).
- Una secuencia de imágenes con correspondencias a partir de una cámara móvil, donde las coordenadas 3D de los puntos pueden ser conocidas o desconocidas.
- Un modelo 3D de la escena o plantillas/patrones de imagen (estando disponibles para relacionarlos a una imagen en concreto).

Los puntos necesarios para el registro pueden extraerse a partir de elementos naturales (esquinas, agujeros, etc.) o especialmente diseñados y ubicados en el entorno (puntos fiduciales, puntos luminosos, marcas de RA, etc.).

En el proceso de cálculo de los sistemas indirectos, se suelen emplear diversos conceptos que cabe distinguir:

- Parámetros de orientación interna de la cámara: O parámetros intrínsecos de la cámara, incluyen la distancia focal, las coordenadas del centro de la imagen y distorsiones radiales y tangenciales de la lente.
- Calibración de la cámara: Proceso mediante el cual se determinan los parámetros de orientación interna de la cámara. Algunos autores también incluyen en este término los parámetros de orientación externa (De la Escalera Hueso, 2001).
- Parámetros de orientación externa de la cámara: O parámetros extrínsecos de la cámara, incluyen la matriz de rotación (3x3) y el vector 3D de translación. También se utiliza como sinónimo “la matriz de transformación”.
- Orientación de la cámara: Proceso mediante el cual se calculan los parámetros de orientación externa de la cámara (Jähne et al., 1999, 442). Cabe reseñar que en algunos métodos de orientación de la cámara (por ej., DLT), se obtienen simultáneamente los parámetros de orientación interna y los de orientación externa.

La orientación externa varía en aplicaciones en que la cámara se mueve y por tanto su cálculo deberá ser en tiempo real, mientras que la orientación interna suele permanecer constante y su cálculo será normalmente necesario una única vez, antes del cálculo de la orientación externa. También existen otros métodos por los que se calcula de manera simultánea tanto la orientación externa como la interna (por ej. una transformación lineal directa o DLT); éstos métodos no suelen ser utilizados en aplicaciones de RA puesto que, al permanecer normalmente la orientación interna constante, sería un cálculo excesivo/redundante.

2.1.5.2.1. Calibración de la cámara

El cálculo de la orientación interna de la cámara es un paso imprescindible y previo al cálculo de la orientación externa. En él se determinan las condiciones geométricas intrínsecas de la cámara, como son la distancia focal, las coordenadas del centro de la imagen (x_0 , y_0) y la distorsión radial y tangencial de la lente. En (Kraus, 1997; Jähne et al., 1999; De la Escalera Hueso, 2001; Neumann, 2001; Lerma García, 2002) podemos encontrar distintas metodologías de calibración: con banco de calibración (*test-field calibration*), con líneas rectas, con colimadores, etc. A continuación se expondrá la explicada en (Neumann, 2001, 334-335) por ser similar a la que utilizan las librerías ARToolKit. Este método se divide en dos partes: en primer lugar se procede al cálculo de la distancia focal; en segundo lugar al cálculo de la distorsión de la lente y coordenadas del centro de la imagen.

Para la determinación de la distancia focal utilizan un plano que contiene un patrón de rejillas en el que se conoce la separación entre las líneas. Se toman diversas imágenes de dicho patrón, cada una de ellas separada de la anterior por una distancia D a lo largo el eje óptico. Para varios pares de imágenes se calcula la distancia focal f a partir de la Ecuación 6 y la Figura 55. La distancia focal aceptada es la media de las anteriores.

$$f = \frac{K_1 K_2 D}{K_1 - K_2}, \text{ donde: } K_1 = \frac{x_2 - x_1}{d} \text{ y } K_2 = \frac{x_4 - x_3}{d}$$

Ecuación 6. Ecuaciones para el cálculo de la distancia focal.

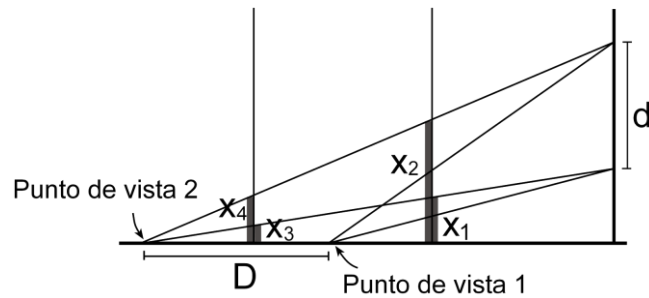


Figura 55. Dos posiciones de la cámara para el cálculo de la distancia focal.

Para el cálculo de los parámetros restantes también se utiliza un sistema de rejillas como el explicado anteriormente y una exposición de la cámara perpendicular a éste. El sistema de ecuaciones es un poco más complejo, por ello no se expone aquí el desarrollo, sino simplemente las ecuaciones resultantes:

$$U + \Delta U = U + f_x (k_1 U_r D^2 + t_1 (D^2 + 2U_r^2) + 2t_2 U_r V_r) = \frac{f_x X_0 + Z_0 U_0 - f_x X}{Z_0}$$

$$V + \Delta V = V + f_y (k_1 V_r D^2 + t_2 (D^2 + 2V_r^2) + 2t_1 U_r V_r) = \frac{f_y Y_0 + Z_0 V_0 - f_y Y}{Z_0}$$

Ecuación 7. Sistema de ecuaciones para la obtención del centro de la imagen y coeficientes de distorsión.

Donde:

$$U_r = \frac{U - U_0}{f_x}, \quad V_r = \frac{V - V_0}{f_y}, \quad D = \sqrt{U_r^2 + V_r^2}$$

U, V : Coordenadas imagen de un punto tridimensional (x, y, z)

U_0, V_0 : Coordenadas imagen del centro de la distorsión de la lente

f_x, f_y : Distancias focales efectivas a lo largo de los ejes u y v

X_0, Y_0 : Centro de la imagen

k_1 : Coeficiente de distorsión radial

t_1, t_2 : Coeficientes de distorsión tangencial

Y los parámetros a calcular mediante un proceso iterativo mínimo cuadrático son: X_0, Y_0, k_1, t_1 y t_2 .

2.1.5.3. MÉTODOS MIXTOS

Como ya se ha apuntado anteriormente, dependiendo de los sensores utilizados el sistema de ecuaciones será diferente. Como ejemplo se explica la combinación de INS y cámara, tal y como se utiliza en la aplicación *UrbanMix* (capítulo 8).

2.1.5.3.1. INS y cámara

En la aplicación *UrbanMix* se ha utilizado un sistema mixto INS-cámara para el cálculo de la orientación externa de la cámara. En el cómputo se han empleado las ecuaciones de colinealidad (Ecuación 1). El procedimiento es el siguiente: los términos r_{ij} (Ecuación 2) se obtienen directamente a partir del INS, siempre que éste esté alineado con los ejes de la cámara; en caso de no estarlo, debería conocerse la matriz de rotación que relaciona ambos sistemas (cámara-INS). Por lo tanto, en el sistema de ecuaciones planteado en Ecuación 1, tendremos que las incógnitas son únicamente las coordenadas tridimensionales del centro de proyección de la cámara (X_o , Y_o , Z_o). Para el cálculo de tres incógnitas, necesitaremos un sistema de como mínimo tres ecuaciones; puesto que se pueden plantear dos ecuaciones por cada punto medido en la imagen, necesitaremos al menos identificar dos puntos en la imagen de coordenadas 3D conocidas (2 puntos x 2 ecuaciones = 4 ecuaciones). A continuación el sistema planteado se puede resolver por procedimientos mínimo-cuadráticos.

2.2. RELACIONES ESPACIALES ENTRE LOS ENTORNOS VIRTUAL Y REAL

En la mayoría de aplicaciones de VR, el requerimiento visual del entorno virtual es simplemente que la escena generada aparezca consistente para el observador. Sin embargo, como ya se ha dicho en varias ocasiones, para la mayoría de las aplicaciones de RA un requerimiento adicional es que las imágenes generadas por ordenador concuerden de manera precisa con el entorno del mundo real. En estas aplicaciones debe obtenerse información precisa del mundo real para poder llevar a cabo la calibración de los entornos real y virtual. Sin el conocimiento preciso tanto de la geometría del mundo real como de la escena generada por ordenador, el registro exacto no es posible. Para alinear adecuadamente entre ellas la imagen video y las

imágenes generadas por ordenador, se deben tener en consideración varios sistemas de referencia (Barfield et al., 1995, 550).

Como se puede ver en la Figura 56, un punto situado sobre el área de trabajo se proyecta sobre la pantalla virtual. Los objetos del entorno real se pueden definir a partir de un sistema de coordenadas terrestre. El sistema de coordenadas del entorno virtual se define relacionando su localización con respecto a la óptica del display físico. Los parámetros del entorno virtual incluyen tres rotaciones y tres traslaciones que determinan su localización respecto al sistema terrestre.

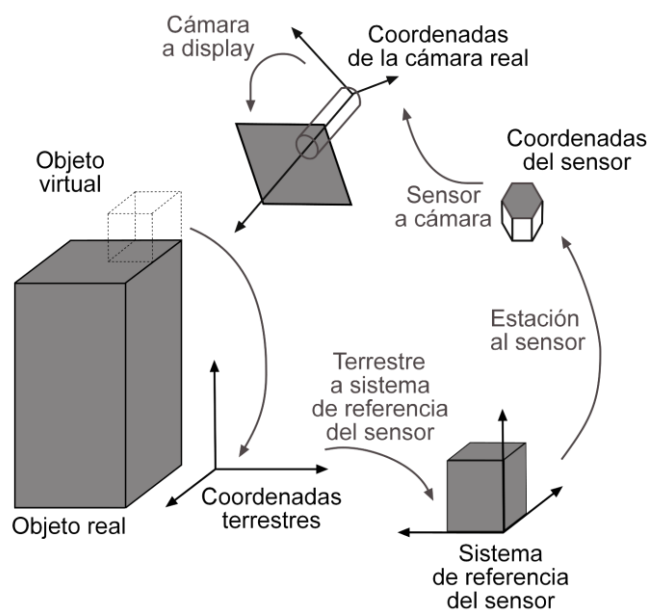


Figura 56. Sistemas de coordenadas real/virtual en aplicaciones basadas en registro por sensores para un sistema de referencia terrestre y un registro directo. A partir de (Azuma, 2004).

Hay que tener en cuenta que en sistemas estereoscópicos hay dos sistemas de coordenadas virtuales (uno para cada ojo), que definen el punto de vista. Además, puede haber otro par de sistemas de coordenadas que deben tenerse en consideración: el sistema de coordenadas propio del sensor que detecta la posición y rotación de la cabeza (ej. un sensor inercial combinado con un receptor GPS); y el sistema de coordenadas propio de los transmisores (ej. los satélites GPS) o dispositivos a partir de los cuales los sensores calculan su posición y rotación. El transmisor es un objeto físico que está anclado al sistema de coordenadas terrestres (o del que se conocen sus coordenadas en todo momento). Al mover el usuario su cabeza, la posición y la orientación del receptor cambia con respecto a los transmisores. Además, hay otros parámetros que cambian dependiendo de los usuarios, como la distancia interocular o la distancia vertical desde la posición del sensor (normalmente situado en

la parte superior de la cabeza) hasta el plano horizontal de los ojos. Así pues, para un registro adecuado de las imágenes, todos estos parámetros deben tenerse en cuenta en la calibración del sistema.

Los sistemas de coordenadas que emplean cámaras y marcas fiduciales difieren del esquema anterior. La diferencia esencial es que desaparece un sistema de coordenadas, el propio del sensor, por lo que el esquema se simplifica (Figura 57).

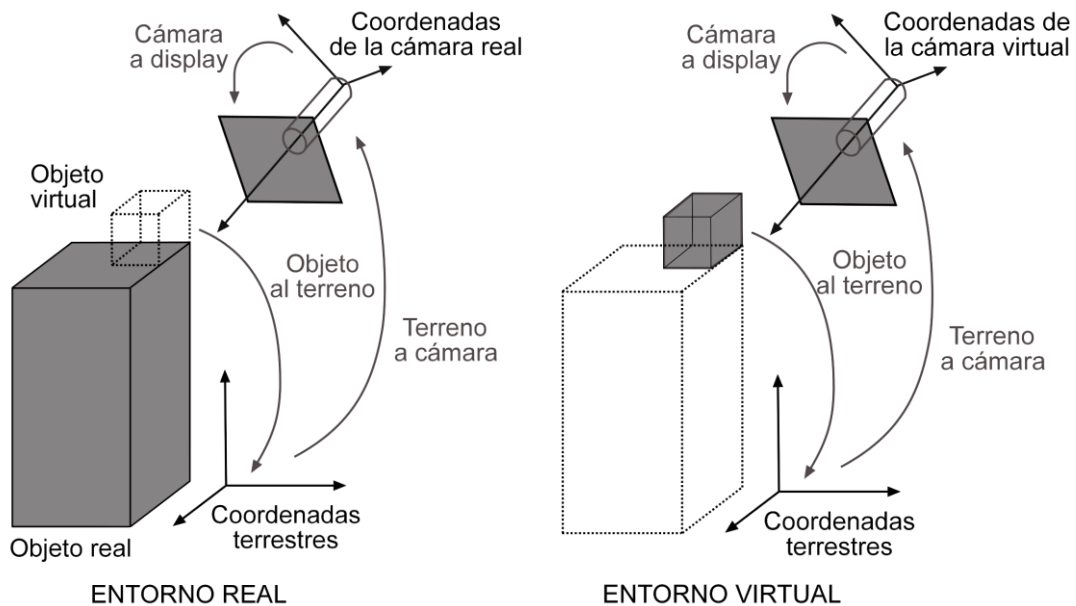


Figura 57. Sistemas de coordenadas real/virtual en aplicaciones basadas en registro por cámaras para un sistema de referencia terrestre y un registro indirecto. A partir de (Azuma, 2004).

2.3. RENDERIZACIÓN EN SISTEMAS DE RA

Tras un registro correcto de los entornos reales y virtuales, ambos mundos deben mostrarse al usuario de forma conjunta (en las aplicaciones más tradicionales de RA; ver apartado 1.2.1). En este apartado se abordarán los aspectos técnicos más relevantes para lograr dicha unión.

Para que la mezcla de los entornos real y virtual aparezca consistente al ojo humano, debe asegurarse que los dos niveles de la imagen vídeo resultante (real y virtual) se fusionen estableciendo una lógica visual creíble utilizando para ello el efecto *keyer*, de modo que el entorno 3D actúa como una figura sin fondo al sustituir esos datos de la imagen por los de la entrada de vídeo por circuito cerrado (Barfield et al., 1995, 548).

Además de esta composición de las dos entradas de vídeo, también hay que procurar que el entorno virtual generado guarde una **perspectiva, oclusión y sombreado** de acuerdo con el entorno real (Bimber, 2003). De estos tres puntos, quizá el más evidente y/o importante es el primero. Una correcta perspectiva se consigue haciendo que la imagen de la escena del mundo virtual coincida geoméricamente con la de la escena del mundo real, para lo que el correcto registro de la cámara (cálculo de la orientación externa) en tiempo real es crucial. Este hecho se ha tratado ya en los apartados 2.1 y 2.2, por lo que a continuación haremos una revisión de la oclusión y el sombreado, y de cómo distintos autores han tratado de implementar estos factores.

2.3.1. OCLUSIÓN

Las aplicaciones de RA en las que la escena mostrada en el display es el resultado de la fusión de los entornos real y virtual no están exentas de oclusiones entre los objetos reales y los virtuales. Dependiendo del punto de vista de la cámara, puede darse el caso de que los objetos virtuales oculten total o parcialmente objetos del entorno real, o el caso contrario, que objetos reales oculten total o parcialmente objetos virtuales. El primero de los casos no supone un problema, ya que la imagen del mundo virtual aparece siempre en un primer plano. Sin embargo, el segundo caso debe ser tratado de forma especial ya que deben determinarse las partes de los objetos virtuales que no deben ser renderizados, constituyendo el denominado "problema de la oclusión", que muchos autores han considerado en sus trabajos (Breen et al., 1996; Fuhrmann et al., 1998; Tatham, 1999; Lepetit et al., 2000b; Lepetit et al., 2000a; Fischer et al., 2003; Gibson et al., 2003; Fischer et al., 2004).

En (Breen et al., 1996) se distinguen dos metodologías diferentes: el método basado en modelos y el método basado en profundidad, que se describen a continuación.

Método basado en modelos. Consiste en realizar modelos tridimensionales de los objetos físicos que aparecen en la escena, y situarlos espacialmente en el mundo virtual de acuerdo con la perspectiva de la cámara de tal forma que, en la imagen obtenida por la cámara en tiempo real, los objetos físicos y sus homólogos virtuales coincidan. Una vez que los modelos virtuales de los objetos reales están correctamente posicionados en el entorno virtual, éstos pueden ser utilizados a modo de máscara. Por ejemplo, dependiendo del sistema utilizado, a la textura del modelo del objeto real se le puede asignar el color negro, por lo que no aparecerá en la imagen resultante

(fusión de los entornos real y virtual), sino que a través de ellos se verá la imagen vídeo, haciendo el efecto de ocultar aquellas partes del resto de elementos (puramente) virtuales que deban ser ocultadas.

En (Lepetit et al., 2000b; Lepetit et al., 2000a) se muestra un método semiautomático que puede catalogarse como método basado en modelos, ya que se generan modelos 3D de los objetos reales para construir máscaras de vídeo. La novedad reside en la manera en que se obtienen dichos modelos, generados de forma semiautomática a partir de un silueteado de la representación de los objetos físicos (que producen la oclusión) en varias imágenes (de referencia) capturadas por la cámara.

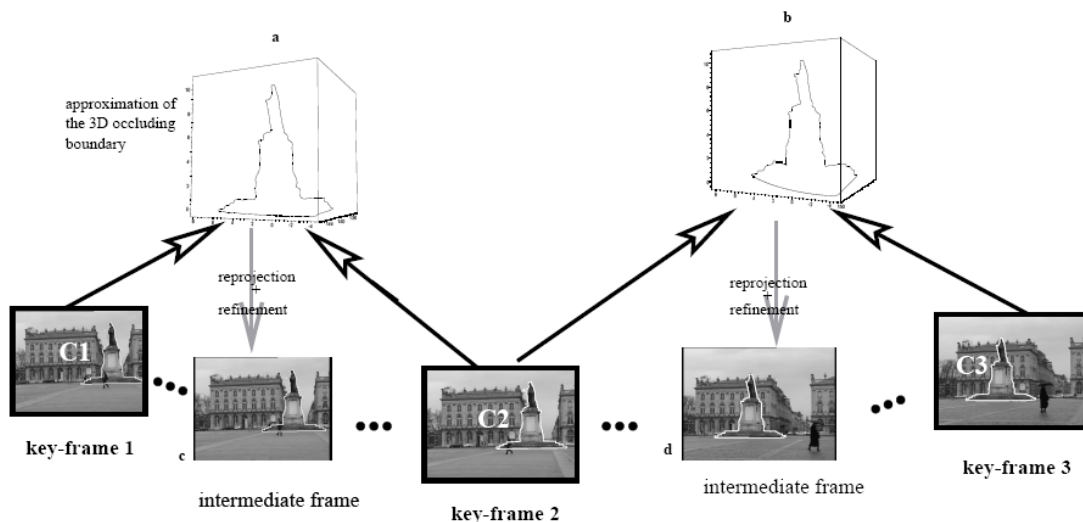


Figura 58. Generación semi-automática de modelos 3D. En (Lepetit et al., 2000b; Lepetit et al., 2000a).

Método basado en profundidad. Consiste en utilizar un mapa de profundidad del entorno real para producir oclusiones. El mapa de profundidad se puede teselar (hacer un mosaico) y a continuación subdividirse para crear una superficie poligonal. El modelo poligonal representa las superficies combinadas de los objetos físicos vistos desde la cámara. El resultado es un modelo geométrico de la escena actual, y puede utilizarse de igual forma que en el método anterior.

Existen sin embargo otras metodologías, como la descrita en (Fischer et al., 2003), que explica un **método basado en la detección de oclusiones en frente de fondos estáticos**. El algoritmo se basa en la comparación gráfica de la imagen del fondo (invariable) con la de la escena actual recogida por la cámara. Si la imagen de la cámara difiere

de la del fondo, se asume que existe oclusión, si no detecta diferencias se asume que el fondo no tiene oclusiones.

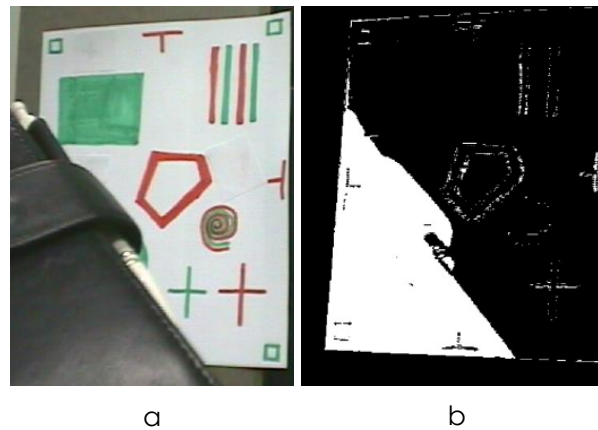


Figura 59. Detección de oclusiones en frente de fondos estáticos: a) Fondo estático ocluido parcialmente por un objeto; b) Mapa de oclusiones (en blanco). En (Fischer et al., 2003).

2.3.2. SOMBREADO

La determinación de condiciones de luminosidad y/o sombreado han sido tomadas en cuenta por múltiples autores, tales como (Stauder, 1999; Tatham, 1999; Gibson et al., 2003; Wang et al., 2003; Jacobs et al., 2006). De acuerdo con (Gibson et al., 2003), para que la escena generada sea consistente, un factor importante es la iluminación. En el entorno virtual deben utilizarse las mismas condiciones de luminosidad que las del entorno real pero, además, deben determinarse en que medida los objetos virtuales afectan la iluminación presente en la escena y considerarse los reflejos y sombras que los objetos virtuales causan sobre los reales. Además, según (Wang et al., 2003), para manipular de forma realista imágenes generadas del entorno real, es necesario el conocimiento de las direcciones de luminosidad. En la Figura 60 y Figura 61 podemos ver ejemplos de los trabajos desarrollados por ambos autores.

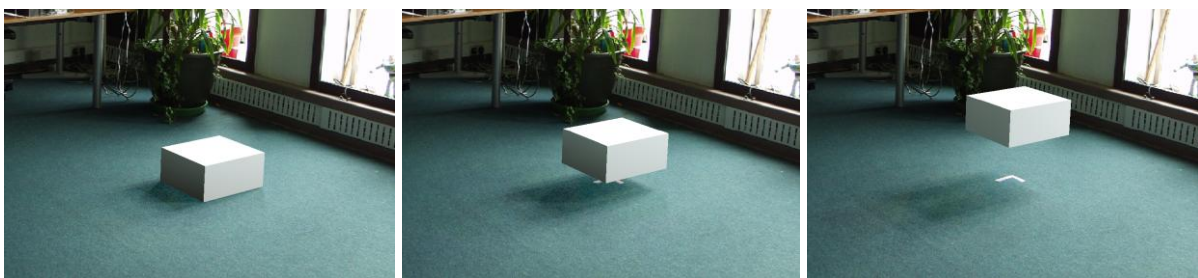


Figura 60. Cambio del sombreado de un objeto virtual y de la sombra proyectada en el entorno real a medida que se éste se eleva, en (Gibson et al., 2003).

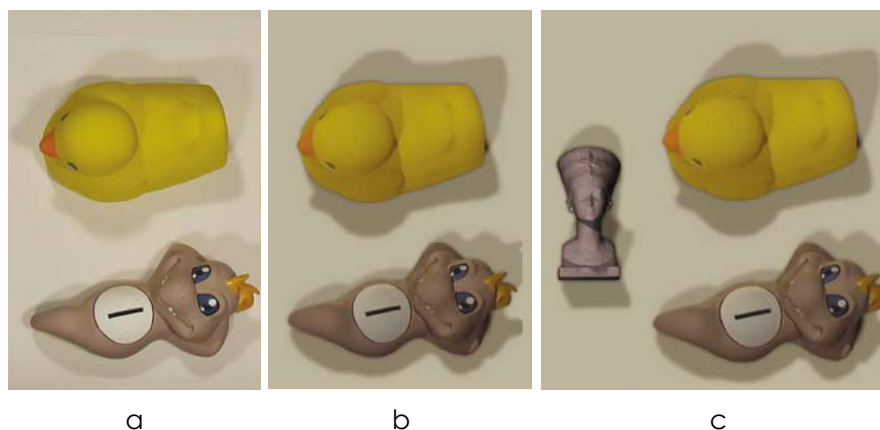


Figura 61. Sombreado: a) Imagen original; b) Imagen generada; c) Imagen generada con la inclusión de un objeto virtual. En (Wang et al., 2003).

En (Jacobs et al., 2006) se identifican tres clasificaciones de las metodologías de iluminación: iluminación común, re-iluminación y métodos basados en la iluminación inversa. Éstas se describen a continuación.

Iluminación común. A esta categoría pertenecen aquellos métodos que proporcionan un cierto nivel de iluminación compuesta (*blending*), como es la adición de sombras proyectadas desde los objetos reales sobre los virtuales, o viceversa. Estas técnicas no permiten ninguna modificación de la iluminación general de la escena. Se pueden considerar dos tipos de iluminación: local y global. Para la iluminación local común, generalmente no se requiere de ninguna información sobre la BRDF (Bidirectional Reflectance Distribution Function). Para la iluminación global común es recomendable tener información sobre el tipo de material de los objetos físicos.

Re-iluminación después de eliminación de la luz. Las técnicas de re-iluminación permiten cambiar las condiciones de luminosidad de la escena en dos pasos. En primer lugar, los efectos lumínicos existentes de la escena real se analizan y, si es posible, se eliminan. En segundo lugar, se generan nuevos efectos de iluminación (sombras, cambios de intensidad, etc.) basándose en un nuevo patrón de iluminación. Estos métodos generalmente no requieren un conocimiento exacto de la BRDF, pero sí de un modelo geométrico detallado de la escena real.

Iluminación inversa. Estos métodos tratan de recobrar propiedades fotométricas de todos los objetos de la escena. Se estiman los valores de la BRDF de manera tan precisa como sea posible, así como las propiedades y localizaciones de focos de luz en la escena real.

2.4. CARACTERÍSTICAS DE SENSORES Y DISPLAYS

Los sensores y displays de RA se tratarán con más detalle en el capítulo 3 y 4. En este apartado únicamente vamos a mencionar unas características generales que deben cumplir para que la escena generada sea creíble, logrando una mayor sensación de inmersión. Estos incluyen la frecuencia de imagen, frecuencia de actualización, retardos del sistema, y el rango y sensibilidad de los sensores de registro (Barfield et al., 1995, 554).

La frecuencia de la imagen es una variable dependiente del hardware y que determina el número de imágenes renderizadas por segundo. La frecuencia de refresco del display es el ratio en el cual se presentan nuevas imágenes. Según (Barfield et al., 1995, 554-555) las imágenes en entornos virtuales se mapean al mundo real dado el límite del ratio de actualización del sistema de registro (los sensores) y del display, y no con el límite del ratio de adquisición de las imágenes. Un retardo en la generación de la imagen y registro se aprecia en todos los dispositivos de display, pero se acentúa dramáticamente en los HMD ópticos (apartado 4.1.1.2). En este caso la visión del mundo real se mueve de forma consistente con los movimientos de la cabeza, pero las imágenes generadas por ordenador sufren del retardo. Este es un problema crítico si se requiere un registro de la imagen exacto. Para entornos virtuales, en muchos casos es suficiente conocer de forma aproximada la posición de la cabeza del usuario. Sin embargo, en aplicaciones de RA, pequeños errores de registro se detectan fácilmente por el sistema visual.

Además, la elección del tipo de sensor es crucial en las aplicaciones de realidad aumentada, ya que los distintos sistemas ofrecen precisiones muy variadas (ver capítulo 3). Un sensor tiene que ser preciso en una pequeña fracción de grado para la orientación y pocos milímetros en la posición. Los errores en la orientación de la cabeza afectan al registro de la imagen más que los errores en la posición, llevando así pues a requerimientos más restrictivos para el registro de la orientación de la cabeza. En (Barfield et al., 1995, 555) se menciona que los errores en el registro de la posición no deben exceder los 1-2 mm para aplicaciones de realidad aumentada, especialmente en los HMD ópticos.

3. ENTRADAS: SENSORES

Los sensores realizan el registro del usuario y/o la cámara en tiempo real y proporcionan los datos de entrada (*inputs*) al ordenador en sistemas de realidad aumentada. Según (Bimber et al., 2005d, 4-5), el problema del registro es uno de los temas más desafiantes de la RA en nuestros días. El registro preciso, rápido y robusto de los usuarios así como de los objetos reales y virtuales dentro de un entorno, es crítico para aplicaciones de RA. Algunos autores distinguen entre registro de *fuera-a-dentro* y de *dentro-a-fuera*. El primer tipo, se refiere a los casos en que se aplican sensores fijos en el entorno que registran a una serie de emisores sobre objetos móviles. El segundo tipo, utiliza sensores que están sujetos al objeto móvil. Estos sensores son capaces de determinar su posición relativa a emisores fijos en el entorno. El cálculo de la orientación externa y el registro continuo a lo largo de una amplia área no siempre es posible con el empleo de una sola tecnología, ya que cada uno de los dispositivos tiene sus limitaciones. Por ejemplo, los sensores magnéticos y de radiofrecuencia están sujetos a interferencias metálicas; los sensores ópticos son muy precisos pero se ven afectados por la visibilidad; el GPS diferencial no tiene suficiente precisión para muchas de las aplicaciones de RA y sufre del problema del trayecto múltiple, con lo que difícilmente puede ser empleado dentro de edificios; etc. Es por ello que en algunas aplicaciones de RA se emplean conjuntamente diversas tecnologías, aprovechando los beneficios que cada una de ellas presenta. Éstas se conocen como tecnologías híbridas, y muchas empresas ofrecen productos comerciales ya integrados.

3.1. SENSORES MECÁNICOS

Los sensores mecánicos pueden resultar un sistema económico y relativamente preciso para registrar la posición de la cabeza o de una parte del cuerpo (National Research Council Staff, 1994, 190). Este tipo de sensores pueden medir incluso el movimiento del cuerpo por completo y no tienen latencias intrínsecas. La fuerza de reflexión se incorpora montando actuadores en las juntas de los enlaces. Se pueden distinguir dos tipos dependiendo de si los enlaces mecánicos los lleva enteramente el usuario sujetos a su cuerpo (*body-based*) o si están parcialmente anclados/enlazados al suelo (*ground-based*).

3.1.1. GONIÓMETROS

Los enlaces sujetos al cuerpo – denominados exoesqueletos o goniómetros – se han utilizado frecuentemente en biomecánica para la medición angular de las juntas (National Research Council Staff, 1994, 191). Para interfaces hápticos, estos sensores forman la base de los guantes de RA (*master gloves*) y exoesqueletos de fuerza refleja (*force-reflecting exoskeletons*). Debido a que estos dispositivos están sujetos al cuerpo (se llevan puestos a modo de vestimenta), resultan ergonómicos y facilitan la movilidad. Sin embargo, si hay un desplazamiento del cuerpo, se requiere además de un sistema adicional de registro.

Se pueden distinguir dos maneras de utilizar datos del goniómetro: deducir los ángulos de las juntas o deducir posiciones entre puntos. Existe el problema de que las uniones en los seres humanos no representan uniones perfectas de bisagra o uniones esféricas: los ejes de la rotación se mueven con los ángulos de las juntas.

Un ejemplo de este tipo de sensores es el guante de captura de movimiento inalámbrico *CyberGlove II* de Immersion (Immersion Corporation, 2007b). Está equipado con 18 ó 22 (según modelo) medidores de alta precisión de juntas de ángulos (Figura 62). Utiliza tecnología de registro de la resistencia de plegado (*resistive bend-sensing*) para transformar los movimientos de la mano y los dedos en datos digitales en tiempo real. El modelo con 22 sensores tiene tres sensores de flexión por dedo, cuatro sensores

de abducción⁸, un sensor del arco de la palma de la mano, y sensores que miden la flexión y abducción. Los sensores son muy delgados y flexibles para reducir su presencia dentro del guante elástico. Tiene una resolución 0.5° con una desviación estándar de 1° , y un alcance de unos 9 m desde el puerto USB.

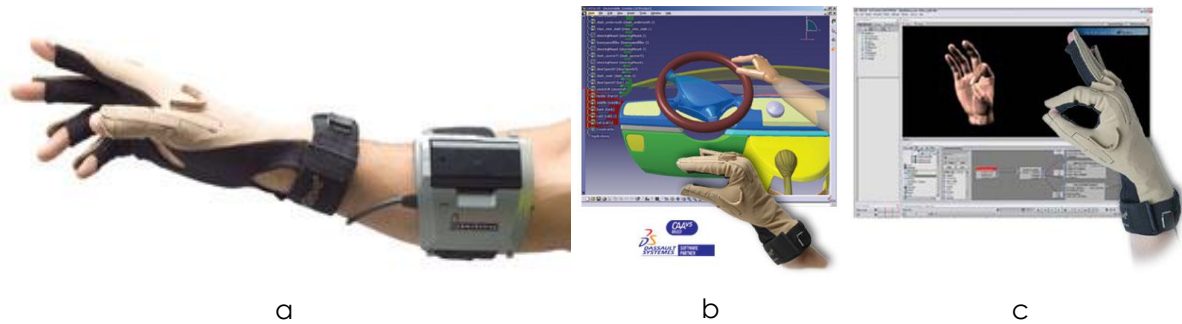


Figura 62. Guantes: a) CyberGlove II de Immersion; b) software para CATIA; c) VirtualHand para MotionBuilder. En (Immersion Corporation, 2007b).

Otro ejemplo de la misma compañía Immersion, es el *CyberGrasp Exoskeleton* (Immersion Corporation, 2007a). Es un sistema de fuerza retroalimentada para los dedos y manos, que permite tanto la manipulación de objetos generados por ordenador como de objetos físicos de forma remota (Figura 63). Al ser un sistema que devuelve la fuerza, los usuarios son capaces de sentir el tamaño y forma de los objetos dentro de un entorno virtual. Las fuerzas de agarre se producen por medio de una red de tendones llevados a las yemas de los dedos por el exoesqueleto. Hay un total de cinco actuadores, uno para cada dedo, que pueden ser individualmente programados para evitar que penetren o aplasten objetos virtuales sólidos. Los actuadores están colocados en un módulo pequeño que se puede dejar sobre la mesa. Además, como el sistema no está anclado al suelo, se puede llevar en una mochila para aplicaciones móviles. El dispositivo ejerce fuerzas de agarre que son aproximadamente perpendiculares a las yemas de los dedos a través de su rango de movimiento. Entre sus especificaciones técnicas se incluyen: fuerza continua máxima de 12 N por dedo; 12 bits de resolución de la fuerza; 350 g de peso; alcance de 1 m; interfaz vía RS-232 y Ethernet.

⁸ Con el término abducción se entiende el movimiento de salida del dedo del eje del brazo. El movimiento de extensión/abducción es la capacidad de extensión del pulgar hacia la parte exterior y flexión hacia el interior de la palma.



Figura 63. Sistema CyberGrasp de Immersion. En (Immersion Corporation, 2007a).

Podemos encontrar diversos sistemas de RA que utilizan estos u otros tipos de guantes para registrar los movimientos de las manos de los usuarios. Algunos ejemplos están descritos en (Matysczok et al., 2004; Malkawi et al., 2005; Teichrieb et al., 2007). En (Malkawi et al., 2005) se describe un sistema de RA interactivo basado en el reconocimiento del habla y de los gestos de las manos. El sistema se utiliza para visualizar e interactuar con edificios y su entorno térmico (Figura 64). Contiene una librería de reconocimiento de diversos gestos de las manos; gracias a ellos el usuario puede realizar tareas e interactuar con la visualización del espectro térmico.

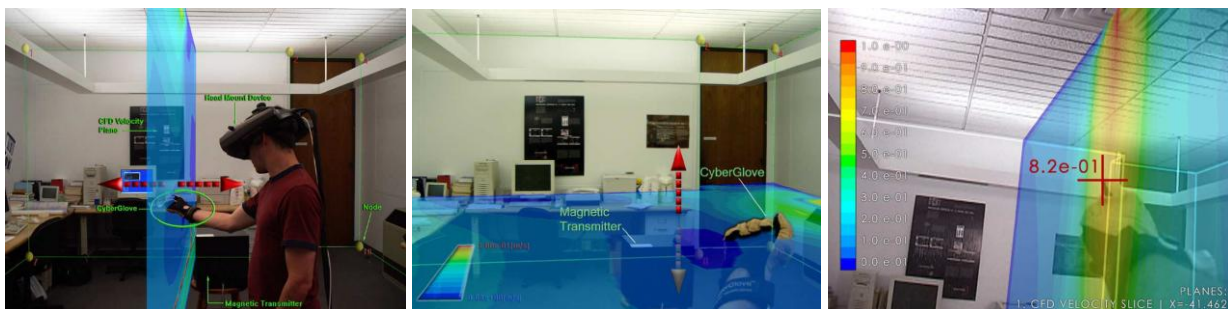


Figura 64. Aplicación de RA con el uso del CyberGlove. En (Malkawi et al., 2005).

En (Teichrieb et al., 2007) se presenta un proyecto de ambientes aumentados, mediante la construcción de un escenario altamente interactivo, utilizando para ello la integración de diversos sistemas de registro (marcas de ARToolKit y guante) con dispositivos multimedia (Figura 65). El sistema se ha probado en un museo virtual, donde los usuarios se unen a una red de servicios que incluye visualizaciones de piezas de arte, entrevistas, chat y subastas remotas en vivo.



Figura 65. Usuario con guantes de RA. En (Teichrieb et al., 2007).

3.1.2. ENLACES DE SUELO

Los enlaces de suelo (*ground-based linkages*) se utilizan principalmente para registrar los seis parámetros de orientación externa, tales como la posición y orientación de la cabeza o de la mano. Se asume que el usuario puede agarrar rígidamente el manipulador, o que el display está sujeto a la cabeza (National Research Council Staff, 1994, 192).

Un ejemplo es el dispositivo *CyberForce Tactile Feedback System* de Immersion (Figura 66). Este es un sistema que registra los movimientos de la mano y del brazo y devuelve la fuerza. Al estar anclado al suelo (o a una mesa) registra los 6 parámetros de orientación externa – posición y rotaciones en 3D. Se puede utilizar conjuntamente con *CyberGrasp Exoskeleton* para poder poner la mano sobre un volante, sentir la fuerza e inercia mientras se coge un objeto virtual pesado, o sentir la resistencia impenetrable de una pared generada por ordenador. Entre sus especificaciones técnicas se incluye: fuerzas de posición de entre 6.6 N y 8.8 N; peso de 8.6 Kg (sin el *CyberGrasp Exoskeleton*); alcance de un cuadrado de barrido de 30x30 cm a través de 133° con un radio de 49.5 cm; resolución en la posición de entre 0.06 mm y 0.07 mm; resolución en orientación de 0.09°.



Figura 66. CyberForce junto al CyberGrasp Exoskeleton de Immersion. En (Immersion Corporation, 2007a).

En (Inami et al., 2000) se muestran varios sistemas de realidad mezclada que utilizan enlaces de suelo. En la Figura 67-a podemos ver un dispositivo para su uso en aplicaciones médicas; el enlace de suelo es de un material retro-reflectante, con lo que aparece camuflado gracias a un sistema de proyección sobre un espejo semitransparente (ver apartado 4.1.1.3). En la Figura 67-b se muestra otro dispositivo que emplea HMD ópticos (ver apartado 4.1.1.2) y volúmenes del mismo material retro-reflectante sobre el que se visualizan objetos virtuales, haciendo que el usuario pueda coger los objetos virtuales (display háptico). La localización se realiza en este caso mediante un enlace de suelo sujeto al HMD.

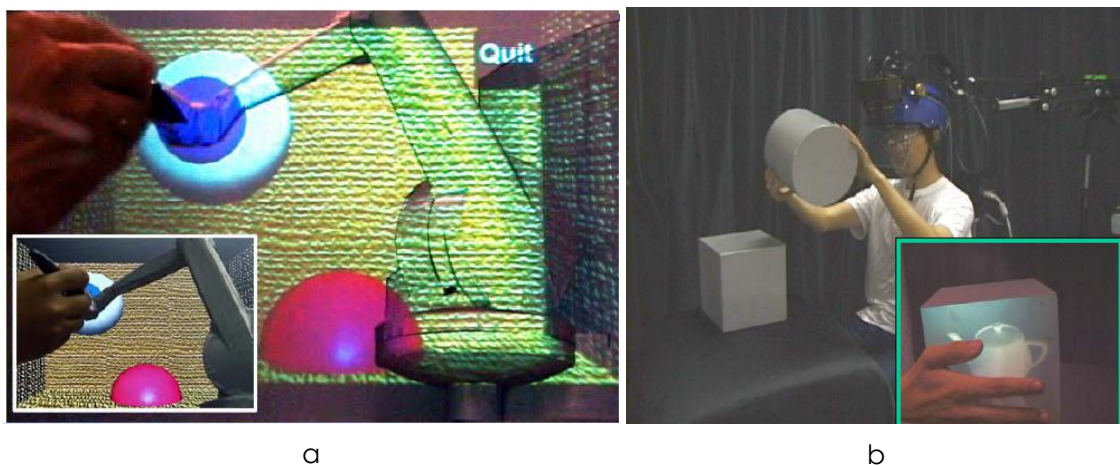


Figura 67. Sistemas de realidad mezclada que emplean enlaces de suelo: a) Dispositivo para aplicaciones médicas; b) Display háptico. En (Inami et al., 2000).

3.2. SENSORES MAGNÉTICOS

Uno de los sensores de posicionamiento más populares son los de tipo magnético debido a su bajo coste, precisión modesta pero razonable y comodidad de manejo. Los sensores magnéticos no sufren de problemas de visibilidad, aunque son sensibles a campos magnéticos y materiales ferromagnéticos. Se suelen emplear varios sensores para registrar el movimiento del cuerpo entero.

Un ejemplo de este tipo de sensores es *FASTRAK* (Figura 68) de Polhemus (Polhemus, 2007a). Este sensor ofrece un registro de los seis parámetros de orientación externa en tiempo real con una latencia de 4 ms. El sistema ofrece la posición y orientación de un pequeño sensor mientras se mueve por el espacio. *FASTRAK* incluye una unidad de

sistema electrónico (SEU), suministro eléctrico, un sensor y un transmisor. La SEU contiene el hardware y software necesarios para generar y registrar los campos magnéticos, calcular la posición y orientación, y comunicarse con el ordenador por medio de un puerto RS-232 y USB. El transmisor contiene bobinas electromagnéticas encerradas en una caja de plástico que emite los campos magnéticos y constituye el sistema de referencia para las mediciones del receptor. El receptor contiene bobinas electromagnéticas encerradas en una caja de plástico que detectan los campos magnéticos generados por el transmisor. Su posición y orientación se registran a medida que se mueve. El receptor es completamente pasivo, por lo que no requiere de ningún voltaje.

Para proyectos que requieran del registro de varios elementos, se pueden conectar hasta cuatro sensores a la SEU. También se pueden sincronizar hasta cuatro sistemas *FASTRAK* para permitir el registro simultáneo de 16 sensores. Ofrece una frecuencia de actualización de 120Hz, multiplexada⁹ entre los sensores conectados. Ofrece unas precisiones estáticas de 0.04 mm y 0.15°, siendo el alcance con el transmisor estándar TX2 de 1.52 m, y con el transmisor extendido de 4.6 m.



Figura 68. *FASTRAK* de Polhemus. En (Polhemus, 2007a).

Otro ejemplo similar es *MiniBIRD* (Figura 69) de Ascension (Ascension Technology Corporation, 2007b). En este caso las precisiones estáticas que ofrece son de 1.8 mm en la posición y 0.5° en orientación, y están garantizadas para un rango de 20 cm a 45.7 cm para el modelo 500, y un rango de 20.3 cm a 76.2 cm para el modelo 800.

⁹ Dividida.



Figura 69. MiniBIRD de Ascension. En (Ascension Technology Corporation, 2007b).

Algunos trabajos de RA que utilizan sensores magnéticos los podemos encontrar descritos en (Kiyokawa et al., 2000b; Matysczok et al., 2004; Hirose et al., 2005; Konishi et al., 2005; Malkawi et al., 2005). En la mayoría de ellos, estos sensores se utilizan de forma conjunta con algún otro tipo de sensor. Así pues, en (Matysczok et al., 2004; Malkawi et al., 2005) se utiliza el sensor *FASTRAK* adherido a los guantes para proporcionar la posición y orientación de la mano; con la combinación de ambos, se puede obtener la posición absoluta de todos los dedos. En (Kiyokawa et al., 2000b) se presenta una aplicación colaborativa para el diseño y manipulación de objetos 3D (Figura 70). Cada usuario tiene dos dispositivos de entrada de datos, y cada uno de estos dispositivos tiene un *FASTRAK* con botones integrados.

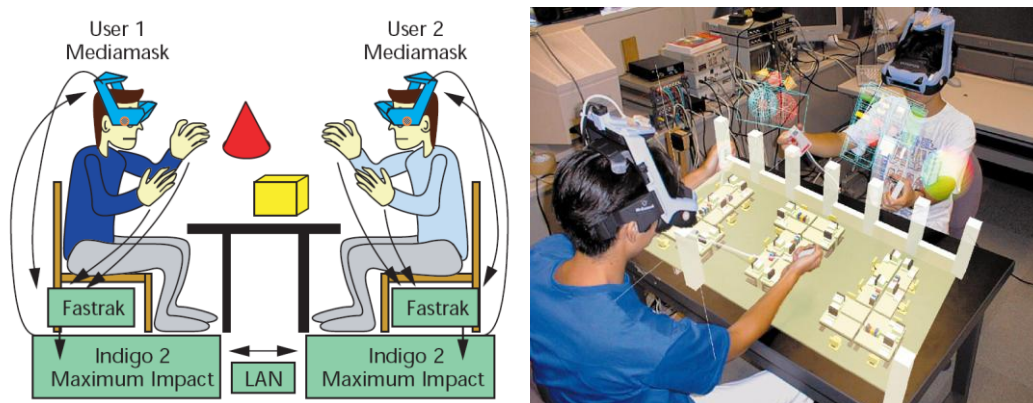


Figura 70. Esquema y usuarios interactuando en un sistema que utiliza diversos sensores magnéticos. En (Kiyokawa et al., 2000b).

3.3. SENSORES ÓPTICOS

Dentro de los sensores ópticos, hay distintas metodologías para el registro de la posición y/o orientación. Según (National Research Council Staff, 1994, 194) la distancia se

puede medir por triangulación (por ej. estéreo-visión), por tiempos (radar láser), o por interferometría. Se puede utilizar la luz pasiva del entorno (sistemas de estéreo-visión), la luz puede ser pulsada (radar láser), o se pueden poner sobre un cuerpo móvil una serie de marcas activas (diodos emisores de luz infrarroja – IREDs) o pasivas. Las cámaras o captadoras pueden emplear series de CCDs (Charge Coupled Device) lineales o planos, PSDs (Position Sensing Detectors), o fotodiodos. En los siguientes subapartados se presentan únicamente los sensores ópticos comúnmente utilizados en aplicaciones de RA.

3.3.1. SISTEMAS DE ESTÉREO-VISIÓN PASIVOS

Los sistemas de estéreo-visión pasivos utilizan un conjunto de dos o más cámaras para resolver el problema de correspondencia mediante la relación de una serie de puntos testigo (*sample points*) en dos imágenes de un mismo entorno tomadas bajo distintos puntos de vista.

En (Point Grey, 2007) podemos encontrar varios tipos de videocámaras estereoscópicas mediante las que se puede, por ejemplo, registrar el movimiento de un objeto en tiempo real. El sistema *Bumblebee* (Figura 71-a) consta de dos CCD con resoluciones de 640x480 (a 30 fps) o 1024x768 (a 15 fps). La focal puede ser de 2, 4 o 6 mm. El sistema *Digiclops* (Figura 71-b) es un sistema estereoscópico que utiliza tres cámaras CCD para proporcionar imágenes digitales 3D en tiempo real. Ofrece resoluciones de 640x480 (a 24 fps) o 1024x768 (a 12 fps). La focal puede ser de 2, 4 o 6 mm.

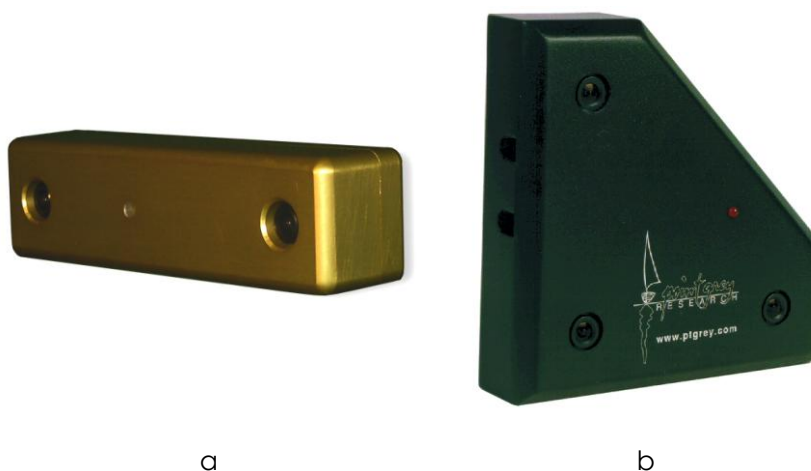


Figura 71. Cámaras estereoscópicas: a) *Bumblebee*; b) *Digiclops*. En (Point Grey, 2007).

Algunos ejemplos de aplicaciones de RA que utilizan este tipo de sensores los podemos encontrar en (Auer et al., 1999; Dorf Müller, 1999; Ribo et al., 2001). En (Auer et al., 1999) se utiliza un sistema combinado de registro óptico y magnético. El sensor magnético utilizado es el *Flock of Birds* de Ascension; El registro óptico se realiza a partir del reconocimiento de las esquinas de rectángulos negros de forma estereoscópica. Por otro lado, en (Dorf Müller, 1999) se utiliza un sistema óptico de marcas retro-reflectante para el cálculo de la orientación externa en aplicaciones de realidad aumentada (Figura 72). Las marcas son iluminadas con luz infrarroja (940 nm) mediante dispositivos sujetos a cada una de las cámaras de un sistema HMD binocular. El registro de las marcas se hace de forma estereoscópica, es decir, mediante la correspondencia entre la imagen izquierda y derecha. Por último, en (Ribo et al., 2001) se presenta un sistema de registro estereoscópico basado en la extracción de manchas (*blobs*), predicción bidimensional, constreñimiento epipolar y reconstrucción 3D.

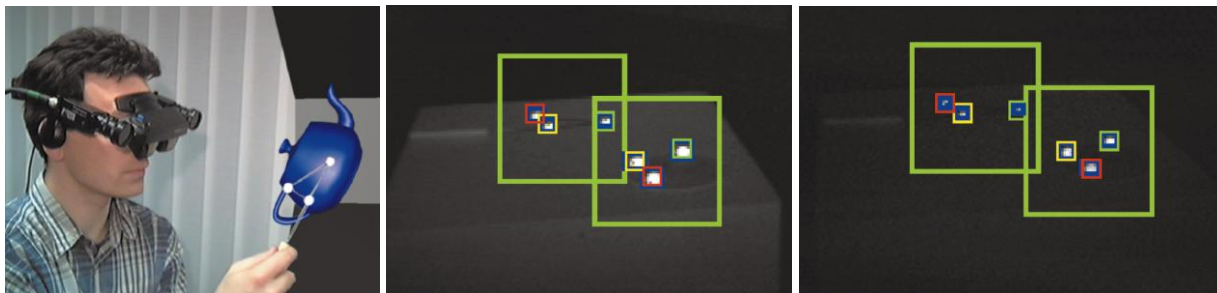


Figura 72. Usuario con marcas retro-reflectantes y sistema HMD binocular; imágenes izquierda y derecha e identificación de las marcas. En (Dorf Müller, 1999).

3.3.2. SISTEMAS DE MARCAS ACTIVAS

El problema de correspondencia se resuelve en los sistemas de marcas activas gracias a la identificación de una serie de puntos fiduciales que son registrados sobre objetos móviles. El procedimiento más sencillo y preciso es utilizar una serie de IRED (Infra-Red Emitting Diode), que crean unas marcas muy brillantes en las imágenes. Los IRED pueden ser pulsados en secuencia con la adquisición de la cámara para detectar inequívocamente cada marca. Para la detección, una técnica frecuentemente empleada es mediante PSDs (también llamados fotodiodos de efecto lateral). La luz incidente induce una corriente que se mide en cada esquina del cuadrado para el cálculo de las coordenadas X e Y del centroide. Un problema fundamental en el uso de PSDs son las reflexiones de la luz IRED de las superficies del entorno que mueven el

centroide aparente de la luz registrada; la cantidad de la luz reflejada es alta, alrededor del 25% del total (National Research Council Staff, 1994, 195).

Un ejemplo es el sistema *ReActor 2* (Figura 73) de Ascension (Ascension Technology Corporation, 2007c), mediante el que se captura el movimiento de un usuario, que puede moverse libremente dentro del área de captura delimitada por barras modulares unidas entre ellas. Los detectores digitales incluidos en las barras proporcionan una cobertura total del usuario al mismo tiempo que minimizan las marcas ocultas. Un total de 544 detectores digitales fijos a lo largo de las barras de aluminio (un total de 12) envuelven al usuario y reciben las señales de las marcas, con lo que se producen menos oclusiones que con sistemas ópticos pasivos. Hay un total de 30 marcas sobre el cuerpo del usuario, estando cada marca de IR compuesta por un conjunto de 42 LEDs ($30 \times 42 = 1260$ LEDs) a 900 Hz, con una alta precisión de 3 mm. La tecnología de reconocimiento instantáneo de marcas es posible gracias a que cada marca está etiquetada con una identidad individual y secuencia de destello, así que los detectores conocen siempre la posición de las marcas. Cada una de las marcas se asigna a una parte del cuerpo mediante un software llamado FusionCore. El usuario puede salir y volver a entrar al área de registro, sin necesidad de recalibrar el sistema. Las marcas se pueden utilizar sobre diversos usuarios al mismo tiempo.

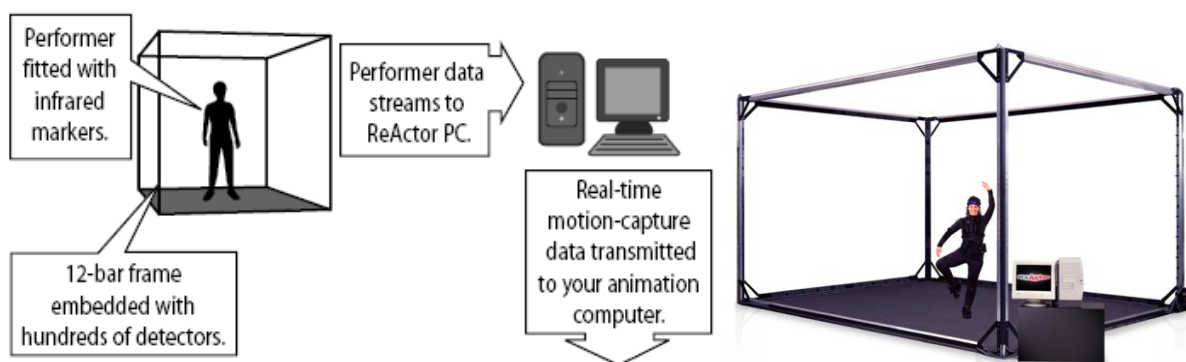


Figura 73. *ReActor 2* de Ascension. En (Ascension Technology Corporation, 2007c).

Las aplicaciones de RA basadas en la detección de marcas activas no son muy comunes. Sin embargo podemos encontrar algunos ejemplos (Madritsch et al., 1996; Rolland et al., 2002; Rosenthal et al., 2002). En (Madritsch et al., 1996) se explica un sistema de registro que consta de dos cámaras CCD y tres LEDs para la obtención de los seis parámetros de orientación externa. El algoritmo de detección de los LEDs trabaja con precisiones sub-píxel, por lo que el método se puede utilizar en aplicaciones de RA que requieran elevadas precisiones. Por otro lado, en (Rosenthal et

al., 2002) utilizan un registro óptico-electrónico para elaborar una sistema de realidad aumentada y utilizarlo en biopsias con agujas guiadas por ultrasonidos (Figura 74). El HMD, la prueba de ultrasonido y la aguja de biopsia están equipadas con LEDs de infrarrojos. El sensor óptico-electrónico registra la posición de estos LEDs. En (Rolland et al., 2002) se utiliza también un sistema óptico-electrónico (el *OptoTrack 3020*), para localizar la posición de un HMD desarrollado por ellos.



Figura 74. Registro óptico-electrónico: a) HMD con dos videocámaras en miniatura y tres LEDs de infrarrojos; b) Mujer realizando una simulación de una biopsia con el sistema descrito. En (Rosenthal et al., 2002).

3.3.3. SISTEMAS DE PATRONES DE MARCAS PLANAS

Los sistemas de patrones de marcas planas (*planar pattern marker systems*), también denominados de marcas fiduciales, resuelven el problema de la obtención de los seis parámetros de orientación externa mediante la identificación de una sola de las marcas por una sola cámara, aplicando para ello técnicas de visión por computador. Previamente al registro la cámara debe ser calibrada, proceso en el cual se obtienen la distancia principal, las coordenadas planas del centro de proyección, y diversos parámetros de distorsión.

Existen varios sistemas que permiten la detección de marcas, entre los que cabe destacar ARToolKit (Figura 75), ARTag, StudierStube, OSGART, etc. Estos sistemas están compuestos por una serie de librerías que recogen las técnicas de visión por computador necesarias para realizar el posicionamiento a la vez que el desarrollo de aplicaciones de RA. Generalmente, estas librerías tienen una licencia de tipo GPL (GNU General Public License), lo que permite su uso libre con fines no comerciales. Este hecho, unido a los mínimos requerimientos de elementos físicos (una cámara Web y las

marcas), a parte del ordenador, hace que estos sistemas sean unos de los más económicos, inmediatos y flexibles, y por ello cobran gran importancia dentro del mundo de la investigación. De estos sistemas se hablará con más detalle en el capítulo 5.

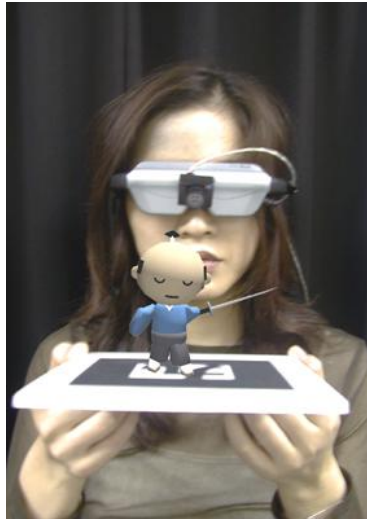


Figura 75. Ejemplo de realidad aumentada con sensor óptico y marcas de ARToolkit.

En (Klinker et al., 1999; Kato et al., 2001; Hedley et al., 2002; Cooper et al., 2004; Fischer et al., 2004; Galantay et al., 2004; Matyszczok et al., 2004; Walczak et al., 2005; Calado Lopes et al., 2006; Freeman et al., 2006; Teichrieb et al., 2007) podemos encontrar aplicaciones basadas en sistemas ópticos de marcas pasivas para el registro de la cámara. En la mayoría de ellos se utilizan las librerías ARToolkit o similares, como ARToolkitPlus.

Zygotes (2004) de Jan Torpus es una pieza narrativa interactiva de realidad aumentada (Galantay et al., 2004). Esta pieza es uno de los escenarios del sistema *Living Room* del mismo autor, en la que se utilizan las librerías de ARToolkit para realizar el registro del usuario dentro del entorno creado; para la generación del escenario aumentado, utiliza el software Max/MSP Jitter. El usuario se mueve dentro de una habitación y, dependiendo de su posición, algunos humanoides/zygotes entran en la escena formando una constelación. Una vez que el usuario encuentra todos los zygotes añorando al sol, la escena sigue una sucesión lineal autónoma que conduce hacia un final, después del cual empieza una nueva sesión (Figura 76).



Figura 76. Imágenes del escenario *Zygotes* en el sistema *Living Room*. En (Torpus, 2004).

En (Bimber, 2007a) se describe una técnica novel de proyecciones de marcas imperceptibles, considerando los parámetros de percepción visual del ojo humano. Una imagen codificada que es invisible a los usuarios se integra de forma temporal en la imagen proyectada, siendo ésta capturada por una cámara sincronizada con la proyección. Además, la representación de las marcas se adapta a los movimientos del usuario, para verse más grandes o más pequeñas dependiendo de las partes de la pantalla que queden ocultas por éste. El sistema se muestra como ejemplo en una aplicación de realidad aumentada para la predicción del tiempo televisada (Figura 77).

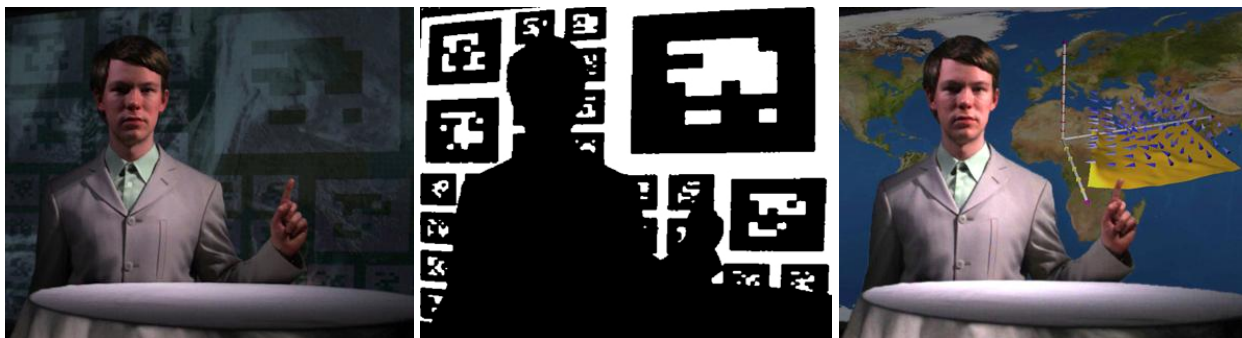


Figura 77. Reconocimiento de marcas proyectadas imperceptibles para el ojo humano. En (Bimber, 2007a).

3.3.4. SISTEMAS DE REGISTRO DE RASGOS NATURALES

Los sistemas de registro de rasgos naturales (*natural feature tracking*) funcionan de manera similar a los de los patrones de marcas, mediante el análisis de la imagen capturada por técnicas de visión por computador. En primer lugar se procede la identificación de una serie de puntos o entidades homólogas entre una imagen testigo y la imagen capturada en tiempo real, o entre dos imágenes contiguas capturadas por la cámara. La correspondencia entre ambas imágenes se describe con la matriz de

transformación, que incluye rotaciones y translaciones, o mediante vectores de desplazamiento. Con estos elementos se establecen los parámetros de orientación externa de la cámara.

Por ejemplo, BazAR (CVLAB - Computer Vision Laboratory, 2007) son unas librerías de libre distribución bajo licencia GPL escritas con lenguaje C++. Estas librerías realizan el análisis de la imagen para el cálculo de la orientación externa de la cámara a partir de la identificación de rasgos naturales puntuales (*natural feature point tracking*), permitiendo también el desarrollo de aplicaciones de realidad aumentada. Estas librerías se explican con más detalle en el apartado 5.2.1.

Algunas aplicaciones de RA que realizan este tipo de registro quedan descritas en (Koller et al., 1997; Neumann et al., 1999; Cornelis et al., 2001; Prince et al., 2002), donde se presentan distintos sistemas de la estimación de los movimientos de la cámara, a partir de los vectores de desplazamiento entre puntos naturales homólogos en imágenes contiguas de una secuencia de vídeo (Figura 78).

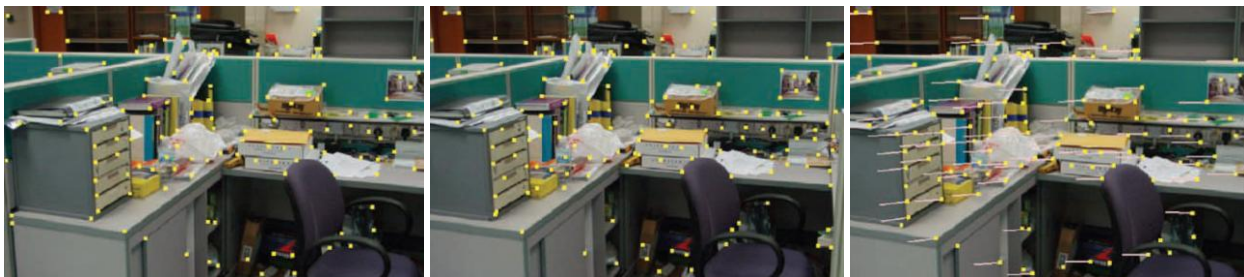


Figura 78. Detección de puntos homólogos en imágenes contiguas y vectores de desplazamiento calculados. En (Prince et al., 2002).

De otro modo, en (Lepetit et al., 2004a; Lepetit et al., 2004b; Lepetit et al., 2005; Pilet et al., 2006) se describe el funcionamiento de las librerías BazAR, y en (Salzmann et al., 2006) se presenta un desarrollo para dichas librerías mediante el cual el sistema óptico monocular es capaz de registrar la matriz de transformación a partir de imágenes sobre superficies deformables (Figura 79).

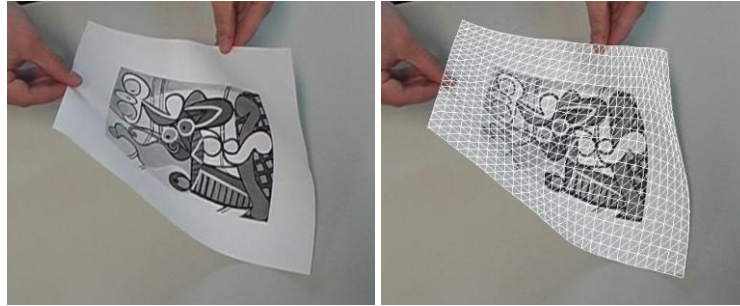


Figura 79. Reconocimiento de puntos en superficies deformables. En (Salzmann et al., 2006).

3.3.5. RADAR LÁSER

La tecnología de radar láser se basa en el cálculo del tiempo necesario que tarda un pulso de luz en viajar desde un emisor, reflejarse en un objeto, y volver a un detector. A partir de la medida del tiempo del recorrido, se calcula la distancia que separa el detector del objeto. En aplicaciones de realidad aumentada, estos sistemas son comúnmente utilizados para un registro de la geometría tridimensional del entorno.

Un ejemplo es *LaserBIRD 2* (Figura 80) de la firma Ascension (Ascension Technology Corporation, 2007a). El sistema escanea constantemente el entorno, permitiendo que su sensor registre de forma instantánea los pulsos de láser emitidos, los procesa, y envía datos de orientación y posición a un ordenador vía conexión RS-232. Algunos datos técnicos son: campo de visión horizontal de $\pm 50^\circ$ y vertical de $\pm 52^\circ$; frecuencia de medición de 240 Hz; alcance de 0.15 m a 1.83 m; precisión en la posición de 0.7 mm RMS¹⁰; resolución estática a 1 m de 0.1 mm; rango angular del sensor de $\pm 85^\circ$ en azimut y elevación, $\pm 180^\circ$ en balanceo (o $\pm 180^\circ$ en azimut, $\pm 85^\circ$ en elevación y balanceo); precisión angular de 0.5° RMS; resolución angular estática a 1 m de 0.05°.

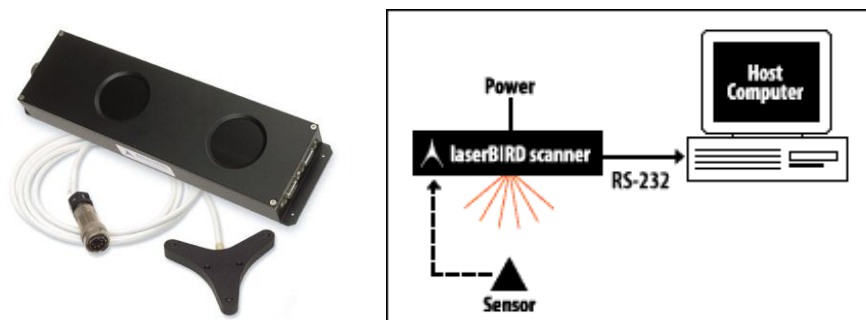


Figura 80. *LaserBIRD 2* de Ascension y esquema de funcionamiento. En (Ascension Technology Corporation, 2007a).

¹⁰ RMS: *Root Mean Square* (raíz media cuadrática).

En (Bimber, 2007b) se presenta un sistema de RA espacial para ser utilizado en entornos físicos arbitrarios (Figura 81). En primer lugar, mediante un láser se realiza un registro de la geometría del entorno; a continuación, la superficie tridimensional capturada se utiliza para realizar una serie de correcciones geométricas a diversas proyecciones sobre las paredes del entorno. El sistema también realiza un registro de punteros láser sobre la superficie y marcas de LEDs situadas en el espacio 3D. La aplicación se utiliza para la manipulación (desplazamientos, rotaciones y ampliaciones) de una ventana virtual proyectada sobre las paredes del entorno, con la finalidad de poder ser utilizado en los campos de la arquitectura o la ingeniería.

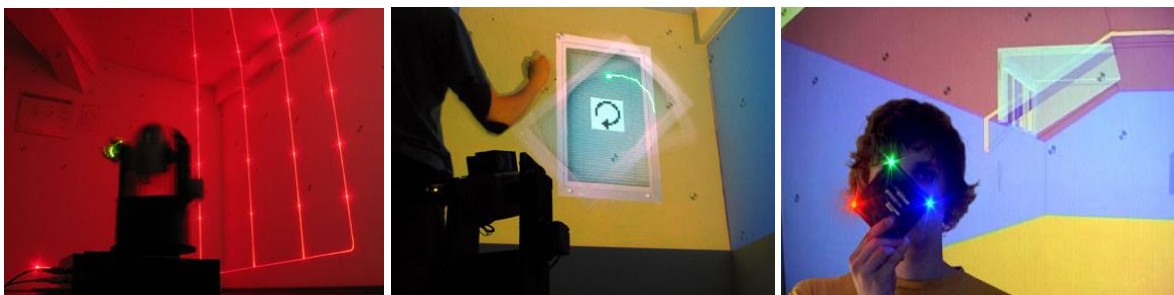


Figura 81. Sistema de láser para el registro del entorno 3D y posterior utilización en un sistema de realidad aumentada espacial. En (Bimber, 2007b).

3.3.6. SENSORES OCULARES

Los sensores oculares son generalmente electro-oculares (medición del potencial corneo-retiniano con electrodos), electromagnéticos (medición del voltaje electromagnético inducido sobre una bobina adherida a una lente ocular), y ópticos (reflexiones de la superficie de los ojos). Únicamente los sistemas ópticos parecen adecuados para un uso generalizado, porque son menos invasivos y pueden ser razonablemente precisos (National Research Council Staff, 1994, 201).

Un ejemplo de este tipo de sensores es *VisionTrak* (Figura 82-a) de Polhemus (Polhemus, 2007b). Este sensor está disponible en versión para sujetarlo a la cabeza y para escritorio. Registra el tamaño de la pupila, el movimiento del ojo y hacia donde mira, todo en tiempo real sin restringir el movimiento de la cabeza. Unas cruces filares se centran automáticamente sobre las pupilas y córneas para indicar un registro adecuado. El sistema *VisionTrack Binocular Desktop 300* – versión de escritorio (Figura 82-b) – permite al usuario estar sentado cómodamente sin apoyar la barbilla. Ambos ojos se localizan utilizando un sistema remoto de video cámara de infrarrojos (IR). La versión para la cabeza, consiste en una serie de componentes de captura muy

pequeños y ligeros montados en una especie de gorra o turbante, permitiendo al usuario libertad de movimiento. Estos componentes se pueden también montar fácilmente sobre un dispositivo HMD. El registro de la escena libre de paralaje con calibración del punto de vista del usuario es suficientemente preciso sobre una profundidad de campo completa. El *VisionTrak Raw Eye Movement Data Acquisition Software* (Figura 82-c), permite que la captura de los ojos y el registro de los datos sean ajustables para cualquier usuario. Los datos capturados pueden verse gráficamente en tiempo real y a su vez ser analizados o exportados en tiempo real a otros dispositivos. El software guarda los datos en formato nativo y en ASCII, y permite una revisión completa de los datos y determinación automática de las velocidades y aceleraciones calibradas.

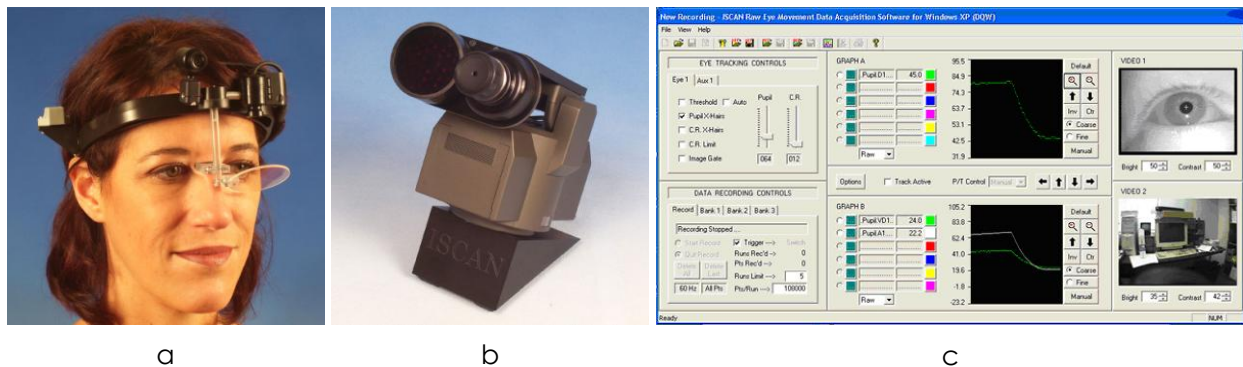


Figura 82. Sensor ocular: a) *VisionTrak*; b) *VisionTrack Binocular Desktop 300*; c) *VisionTrak Raw Eye Movement Data Acquisition Software*. En (Polhemus, 2007b).

En (Barakonyi et al., 2007) se describe un sistema RA de videoconferencia (Figura 83) en el que diversos objetos virtuales se pueden manipular mediante el movimiento de unas marcas y mediante la mirada (gracias a un sensor ocular). El sistema consta de los siguientes elementos: cada uno de los usuarios está equipado con dos monitores para visualizar el espacio local y el remoto; una cámara Web situada encima de la pantalla del ordenador apunta hacia el usuario para registrar el entorno local; también se utilizan monitores de audio para permitir comunicación sonora; finalmente, uno de los usuarios está equipado con un sensor ocular para monitorizar su mirada.

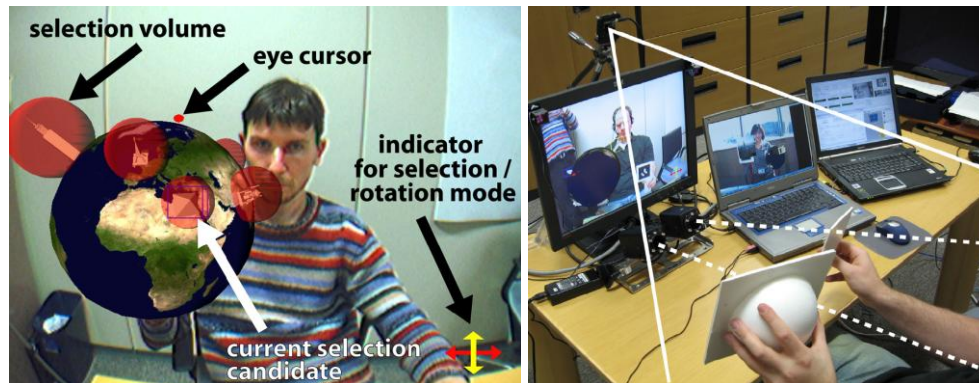


Figura 83. Sistema RA de videoconferencia con registro de la mirada. En (Barakonyi et al., 2007).

3.4. SENSORES ACÚSTICOS

Los sensores acústicos necesitan de al menos tres micrófonos para triangular la señal proveniente de un emisor sobre un cuerpo móvil. Para aplicaciones que requieran de precisiones y velocidades modestas, estos sensores son una alternativa económica a los sensores magnéticos, ya que los rangos son mayores y no se producen interferencias magnéticas. Sin embargo, se debe mantener una clara línea de visión entre el transmisor y el receptor, y la latencia es proporcional a la distancia más larga medida. Otro inconveniente se debe a las posibles interferencias acústicas causadas por diversos aparatos y a los ecos de las superficies duras, que pueden tener hasta un 90% de reflectividad a las ondas ultrasónicas (National Research Council Staff, 1994, 199).

Un ejemplo es *3D HeadTracker* (Figura 84) de Logitech (Logitech, 2007), que normalmente se monta sobre un HMD. El sistema registra los seis parámetros de orientación externa, y se pueden conectar hasta cuatro *HeadTrackers* para registrar datos de forma simultánea. El sistema consiste en un transmisor, un receptor que puede montarse en el HMD y una caja de control. El transmisor consiste en tres altavoces que emiten señales ultrasónicas, y el receptor contiene tres micrófonos en miniatura que reciben dichas señales. El transmisor y el receptor se comunican a través de una línea de visión, teniendo que estar ambos encarados. Una caja de control se conecta a ambos dispositivos y determina la posición del receptor basándose en el tiempo que tardan los micrófonos en registrar las señales. Algunas especificaciones técnicas incluyen: resolución de 250 dpi en modo 3D, o de 400 dpi en modo 2D; velocidad de transmisión de 19200 baudios en modo 3D o 1200 baudios en modo 2D; frecuencia de

actualización de 50 Hz; velocidad de registro de 76 cm/s; espacio de registro de 1.5 m a lo largo en un cono de 100°.

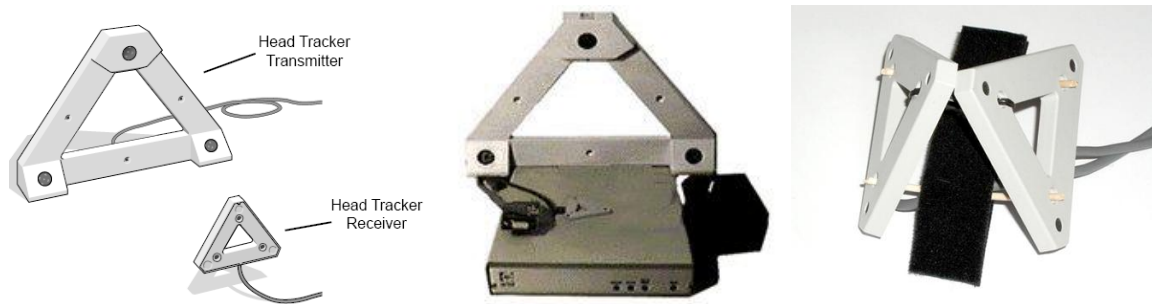


Figura 84. 3D HeadTracker de Logitech. En (Logitech, 2007).

Algunos ejemplos de la utilización de sensores acústicos en aplicaciones de RA las podemos encontrar en (Hoberman et al., 1998; Lee et al., 2006b). En (Lee et al., 2006b) se presenta una aplicación de realidad aumentada denominada *Poultry Internet* en la que un usuario interactúa con un ave de corral virtual mediante sensores de ultrasonidos. Para el registro de la cámara se utilizan marcas de ARToolKit. En (Hoberman et al., 1998) se describe una instalación de realidad mezclada destacada en el ámbito del arte, *El Ball del Fanalet* (1998) de Perry Hoberman y Galeria Virtual (Figura 85), donde la posición de diversos faroles se calcula mediante sensores de ultrasonidos basados en cuatro altavoces situados en las esquinas de una sala, los cuales emiten pulsos de ultrasonidos que son captados por pequeños micrófonos situados en los faroles. Esta información es enviada de forma inalámbrica a un ordenador central que calcula la posición espacial de cada farol. A continuación se generan unas imágenes y sonidos en tiempo real que se corresponden con las acciones de los participantes. Dichas imágenes se proyectan sobre la pista de baile, generando un sistema inmersivo.



Figura 85. Usuarios interactuando con El Ball del Fanalet. En (Hoberman et al., 1998).

3.5. SENSORES INERCIALES

La inercia se define como la propiedad de los cuerpos a mantener constante la velocidad lineal o rotacional, a menos que otras fuerzas actúen sobre éstos, de acuerdo a la primera ley de Newton. Por lo tanto, un sistema de referencia inercial es un sistema de coordenadas en el que se cumplen las leyes de Newton. Estos sistemas no se encuentran en rotación ni aceleración (Grewal et al., 2001, 10). Los sensores inerciales (*INS, Inertial Navigation System*, o *IMU, Inertial Motion Unit*) se basan en fenómenos físicos para medir la aceleración y rotación relativas al sistema de referencia terrestre. Estos sensores son sistemas integrados que constan de giroscopios, acelerómetros y, en algunas ocasiones, de magnetómetros:

- **Giroscopios:** Un giroscopio mecánico es un sistema basado en el principio de conservación del momento angular que dice que un objeto rotado con una velocidad angular elevada, en ausencia de momentos externos, conserva su velocidad angular.
- **Acelerómetros:** Un acelerómetro mide la aceleración lineal del objeto al que está sujeto.
- **Magnetómetros:** Los magnetómetros miden la orientación de un objeto con respecto al campo magnético terrestre.

A partir de los datos registrados por el INS se obtienen de manera inmediata rotaciones 3D. Para el cálculo de la posición, los datos registrados por éste deben ser doblemente integrados, asumiendo que sus condiciones iniciales de posición y velocidad son conocidas. Sin embargo, el resultado es sensible a la deriva y al sesgo del sensor, y los errores en la posición calculada se acumulan con el tiempo (Rolland et al., 2001, 87).

Un ejemplo es el sensor inercial MT9 (Figura 86) de Xsens (Xsens, 2007), que se ha utilizado en algunos ensayos de esta tesis (capítulos 8, 9 y 12). El MT9 es un dispositivo de pequeñas dimensiones que integra un giroscopio de referencia para la medida de rotaciones, utilizando sensores inerciales de tecnología MEMS (Micro-Electro-Mechanical Systems). El sensor inercial consta de tres tecnologías integradas: un acelerómetro, un giroscopio y un magnetómetro (consistente en una película fina magneto-resistiva). Su procesador interno, de bajo consumo, provee una orientación 3D sin deriva (en forma de cuaterniones, medidas angulares o matriz de rotación), junto con las aceleraciones triaxiales, las velocidades angulares (*rate of turn*) 3D, la orientación en el campo magnético 3D y la temperatura.

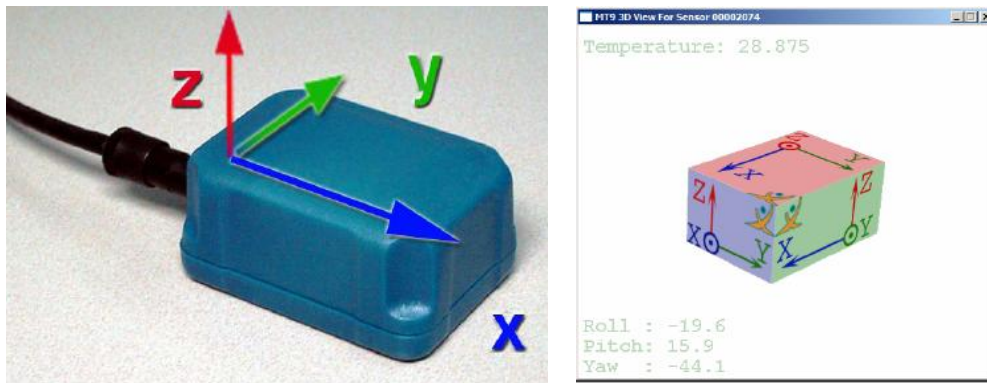


Figura 86. Sensor inercial MT9 de Xsens y su representación virtual. En (XSens, 2007).

Como se ha explicado, mediante los sensores inerciales se obtienen únicamente las rotaciones; es por ello que muchos autores los utilizan en combinación con otros dispositivos de registro para el cálculo adicional de la posición espacial. En (Piekarski, 2006) desarrollan un sistema de RA en espacios abiertos (Figura 87) en el que se utiliza la combinación de un GPS y un sensor inercial (el *Inertia Cube 2* de Intersense) para calcular los seis parámetros de orientación externa del usuario. También se utilizan una serie de marcas para registrar movimientos de la mano del usuario mediante un sistema óptico. Otros autores que describen sistemas combinados de GPS y sensor inercial son (Azuma et al., 1999; Reitmayr et al., 2004; Avery et al., 2005; Demiris et al., 2005; Dow et al., 2005a).



Figura 87. Aplicación de RA en la que el usuario lleva GPS y un sensor inercial para el registro de su posición y orientación, y marcas para el reconocimiento de gestos de las manos. En (Piekarski, 2006).

3.6. SENSORES DE RADIOFRECUENCIA

3.6.1. GPS

El sistema de posicionamiento global (GPS, Global Positioning System) constituye un sistema de radionavegación a escala global, basado en el registro de tiempos y frecuencias (Grewal et al., 2001; Rolland et al., 2001). La constelación GPS está integrada por un total de 24 satélites ordenados en órbitas, de tal manera que en todo momento existen 4 satélites visibles desde cualquier punto del globo terrestre. Además, hay 6 estaciones de monitorización, siendo cuatro de ellas antenas terrestres, una estación central de control y una estación de seguridad. Ofrecen dos niveles de servicio: un servicio de posicionamiento estándar (SPS), y un servicio de posicionamiento preciso (PPS); éste último sólo está disponible para usuarios autorizados por el gobierno de EEUU. Cada uno de los satélites tiene un reloj atómico con una precisión de 340 ns para el SPS. La precisión de los relojes es crucial, puesto que un error de 1 ms se puede transformar en un error horizontal de 300 Km. La estación central controla la órbita de los satélites y aplica correcciones a los relojes si es necesario.

Teóricamente, el sistema puede determinar la posición de un usuario dotado de un receptor GPS con la señal recibida por tres de los satélites, mediante el cómputo del tiempo que tarda el receptor en recibir las respectivas señales. En la práctica, sin embargo, la señal recibida por el reloj del GPS tiene un error (*bias*) adicional, por lo que hacen falta un mínimo de cuatro satélites para el cálculo conjunto de la posición y el error del reloj. Con el sistema SPS, el GPS ofrece precisiones en torno a los 100 m; con el PPS las precisiones son diez veces mejores. Sin embargo, gracias al posicionamiento diferencial, mediante estaciones emisoras terrestres que mejoran la resolución del sistema, se consiguen precisiones centimétricas. Existen varios errores de posición que pueden aparecer limitando la precisión. Los errores principales incluyen errores de órbita de satélites, una mala geometría de los satélites, señales de trayecto múltiple, retardo atmosférico y la sincronización del receptor.

Un ejemplo es el receptor *BTGPS II Trine* (Figura 88), de EMTAC (EMTAC, 2007), utilizado en uno de los ensayos de esta tesis (capítulo 9). Durante el intento de conexión con un punto de posición, el receptor necesita localizar al menos tres satélites y utiliza la señal que pueda recibir al igual que los datos de la última posición almacenada en la memoria digital del receptor. La posición puede fijarse rápidamente en sólo 10

segundos a partir de un estado de inicio “en caliente”¹¹, pasando este tiempo a ser de 80 segundos “en frío”¹², o de 3 a 5 minutos en un estado de reinicio absoluto, por ejemplo, si se encuentra a más de 500 Km de distancia del punto origen. Trine utiliza señales de satélite para calcular una ubicación geodésica mediante un método de triangulación, con un margen de precisión de 10 m en disponibilidad selectiva (SA). Los datos de posición se convierten en coordenadas de latitud y longitud. Una característica de este modelo de sensor es que dispone de cuatro puertos Bluetooth, para poder conectarse a cuatro dispositivos de manera simultánea. También se le puede acoplar una antena externa activa que va provista de un imán para su adhesión a superficies metálicas.



Figura 88. Receptor BTGPS II Trine y antena externa activa, de EMTAC. En (EMTAC, 2007).

En el apartado 3.5 se han descrito varios sistemas que utilizan GPS en combinación con sensores inerciales. También podemos encontrar otras aplicaciones descritas en (Höllner et al., 2001; Flintham et al., 2003; Reimann et al., 2005; Bartie et al., 2006; Dasgupta et al., 2006; Kealy et al., 2006; Narzt et al., 2006), la mayoría de ellas describiendo sistemas de navegación de realidad aumentada o mezclada.

3.6.2. RFID

El sistema RFID (*Radio Frequency IDentification*, en español identificación por radiofrecuencia) es un sistema de almacenamiento y recuperación de datos remoto

¹¹ Inicio “en caliente”: Modo de inicio del receptor GPS en el que la posición actual, la compensación de la sincronización, tiempo GPS aproximado y los datos actuales de efemérides están todos disponibles (EMTAC, 2004).

¹² Inicio “en frío”: Modo de inicio en la que el receptor GPS puede llegar a una solución de navegación sin disponer de una posición, hora, efemérides y datos de almanaque iniciales (EMTAC, 2004).

que usa dispositivos denominados etiquetas (*tags*). El propósito fundamental de esta tecnología es transmitir la identidad de un objeto mediante ondas de radio. Una etiqueta RFID es un dispositivo pequeño, similar a una pegatina, que puede ser adherida o incorporada a cualquier objeto. Contienen antenas para poder recibir y responder a peticiones por radiofrecuencia desde un emisor-receptor RFID. Las etiquetas pueden ser pasivas o activas; las primeras no necesitan alimentación eléctrica interna, mientras que las segundas sí lo requieren. Una de las ventajas del uso de radiofrecuencia es que no se requiere visión directa entre emisor y receptor (Wikipedia, 2008b). También se distingue la tecnología M-RFID (Mobile RFID), definida como los "servicios que proporcionan información en objetos equipados con una etiqueta RFID a través de una red de telecomunicación" (Wikipedia, 2008a).

El dispositivo RD5000 Mobile RFID Reader (Figura 89) de Motorola (Motorola, 2008) es un lector RFID que se puede instalar en cualquier parte, estando especialmente diseñado para su uso en dispositivos móviles. El dispositivo proporciona conexión inalámbrica LAN, permitiendo un registro continuo.

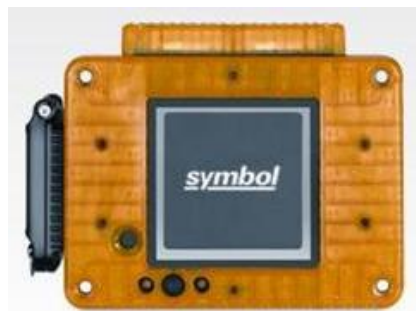


Figura 89. RD5000 Mobile RFID Reader de Motorola. En (Motorola, 2008).

Existen diversas aplicaciones de RA que utilizan este tipo de dispositivo. En (Peitz et al., 2006) se describe el desarrollo y funcionamiento un juego de mesa de realidad aumentada basado en "El aprendiz de brujo" (Figura 90). La tecnología RFID se utiliza para detectar los movimientos de unas cartas y figuritas sobre el tablero, y el resultado al lanzar los dados. Para ello se elaboró un sistema en el que un lector RFID pudiera soportar varias antenas RFID. También se utilizaron varias fotorresistencias para detectar la localización de las fichas.

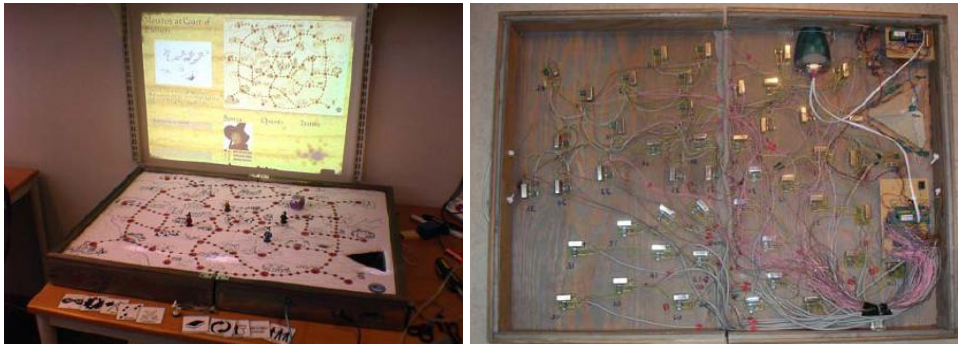


Figura 90. Juego aumentado del Aprendiz de Brujo, con la utilización de RFID y fotorresistencias.

En (Peitz et al., 2006).

En (Reilly et al., 2006) se describe un sistema para interactuar con mapas, pósteres y quioscos utilizando sistemas móviles (Figura 91). La localización del dispositivo móvil se realiza mediante la combinación de RFID y una cámara integrada en el móvil. Se muestra un ejemplo para interactuar con mapas cartográficos, en el que el usuario puede seleccionar diversos elementos del mapa al centrarlos en el campo de visión de la cámara. El sistema dispone de una serie de algoritmos de reconocimiento de imagen, para reconocer diversos elementos del mapa. Al dispositivo móvil se le adhiere un lector RFID, y una malla de RFID se utiliza para proporcionar puntos de anclaje y constreñir el espacio de búsqueda.



Figura 91. Dispositivo móvil con RFID adherido. En (Reilly et al., 2006).

En (Rashid et al., 2006) se describe *PAC-LAN*, un juego de realidad mezclada en espacios urbanos, en donde la localización de los usuarios se realiza mediante M-RFID. Este trabajo se ha introducido anteriormente en el apartado 1.4.2.

4. SALIDAS: DISPLAYS

Los displays son aquellos dispositivos que muestran al usuario la información generada por ordenador. Así como los sensores proporcionan los datos de entrada al ordenador, en este caso, podemos afirmar que los displays proporcionan los datos de salida (*outputs*) desde el ordenador. En aplicaciones de realidad aumentada, este tipo de información es predominantemente de tipo gráfica o visual, aunque existen otros tipos de displays capaces de ofrecer información de tipo sonoro, háptico, olfativo o gustativo (estos dos últimos casos son menos comunes). En los primeros años de la tecnología, únicamente se tenían en consideración los de tipo visual, aunque hoy en día existe un interés por la generación de sistemas de RA que aúnen varios tipos de displays; éstos son los llamados sistemas multimodales.

4.1. DISPLAYS VISUALES

Los displays visuales o gráficos son dispositivos que permiten visualizar la imagen y están configurados por una serie de componentes ópticos, electrónicos y mecánicos para generar imágenes en algún lugar entre los ojos del usuario y el objeto físico que se aumenta (Bimber et al., 2005d, 71). Dependiendo de la óptica que se utilice, la imagen se puede formar en un plano o en una superficie no-plana más compleja.

En la Figura 92 se puede ver una clasificación de los displays según la distancia al usuario. Los displays para la cabeza (*head-attached displays*), como los displays retinianos, displays sujetos a la cabeza (*head mounted displays – HMD*) o proyectores sujetos a la cabeza (*head-mounted projectors – HMP*) tienen que ser llevados (o “vestidos”, *worn*) por los usuarios. Mientras que algunos displays son de mano (*hand-held*), otros están situados en el entorno y completamente alejados del usuario. Los displays retinianos y otros basados en proyecciones forman imágenes curvas, bien sobre la retina del usuario o directamente sobre el objeto físico. Sin embargo, la mayoría de los displays forman imágenes sobre planos, y pueden estar sujetos a la cabeza o situados en el entorno.

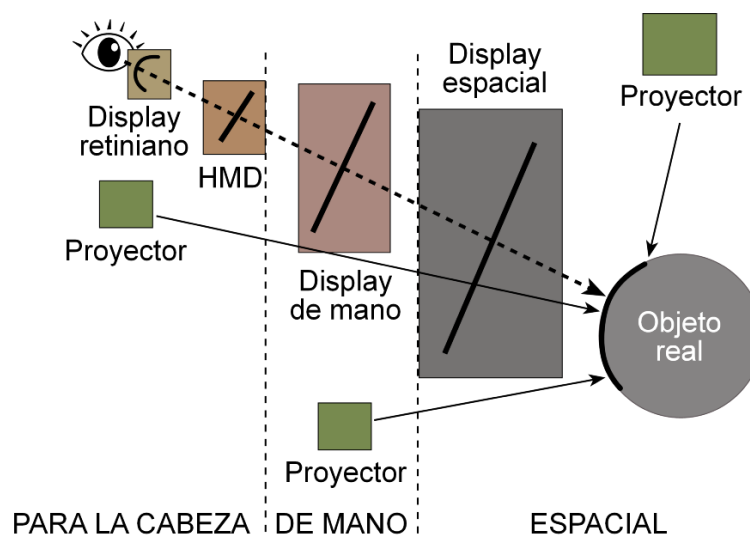


Figura 92. Clasificación de los displays visuales en función de la distancia al usuario. A partir de (Bimber et al., 2005d, 72).

Además, también existe la posibilidad de clasificar los displays para la cabeza según presenten visión monoscópica (ambos ojos ven la misma imagen) o estereoscópica (cada ojo ve una imagen distinta). Para el caso monoscópico, únicamente es posible un HMD basado en imágenes de vídeo; Sin embargo el caso de la estereoscopía es

posible tanto con sistemas puramente basados en vídeo (*video see-through HMD*) como aquellos combinados con un sistema óptico (*optical see-through HMD*). El controlador del vídeo, mezcla la imagen generada en tiempo real con los gráficos generados por ordenador, mostrando el resultado en la pantalla; el sistema óptico muestra la imagen generada por el ordenador junto con la imagen del entorno real tal y como lo ve el usuario a través de un combinador óptico.

4.1.1. DISPLAYS PARA LA CABEZA

Los displays para la cabeza (*head-attached displays, HAD*) requieren que el usuario lleve puesto el display en su cabeza, a la altura de sus ojos. Dependiendo de la tecnología de generación de la imagen, existen tres tipos fundamentales (Bimber et al., 2005d, 72): displays retinianos, que emplean tecnología láser a baja potencia para proyectar imágenes directamente sobre la retina del ojo; HMD que utilizan pequeños displays a modo de gafas delante de los ojos; y HMP que utilizan mini-proyectores o mini-paneles LCD con luz negra (*blacklighting*) y proyectan imágenes sobre superficies del entorno real.

Conviene resaltar que varios autores utilizan términos distintos para referirse a los displays sujetos a la cabeza (HMD), que son una variedad de los HAD. Por ejemplo, en (Barfield, 1995, 544) se distingue entre *opaque HMD* y *transparent* o *see-through HMD*. Sin embargo, en (Bimber et al., 2005d, 74) se habla de *video see-through HMD* y *optical see-through HMD*. Los términos *opaque HMD* y *video see-through HMD* son equivalentes, mientras que *see-through HMD* y *optical see-through HMD* también lo son. En esta tesis se utiliza el término de “HMD de vídeo” para referirse a los primeros y “HMD ópticos” para referirse a los segundos. En aplicaciones de RA ambos sistemas son transparentes (*see-through*), ya que a través de ellos se puede ver el entorno físico – también conocido como la metáfora de “ver a través de”.

4.1.1.1. DISPLAYS RETINIANOS

Los displays retinianos o para la retina (*retinal displays*) utilizan láseres semiconductores de baja potencia para escanear luz modulada directamente sobre la retina del ojo (Figura 93), en lugar de proveer de pantallas delante de los ojos. Esto produce una imagen mucho más brillante y de mayor resolución, y con un mayor campo de visión que un display basado en pantallas.

Las principales ventajas de estos sistemas son el elevado brillo y contraste, a la vez que su bajo consumo, haciéndolos muy útiles para aplicaciones en espacios abiertos. Se espera que en un futuro los sistemas soporten imágenes estereoscópicas policromáticas, altas resoluciones y abarquen grandes ángulos de visión. Los principales inconvenientes de estos dispositivos son (Bimber et al., 2005d, 73):

- Sólo se utilizan imágenes monocromáticas (rojo), ya que el láser azul y verde a baja potencia resulta un sistema demasiado caro.
- El sentido de la acomodación ocular no se soporta, siendo la distancia focal fija.
- No existe la versión estereoscópica.

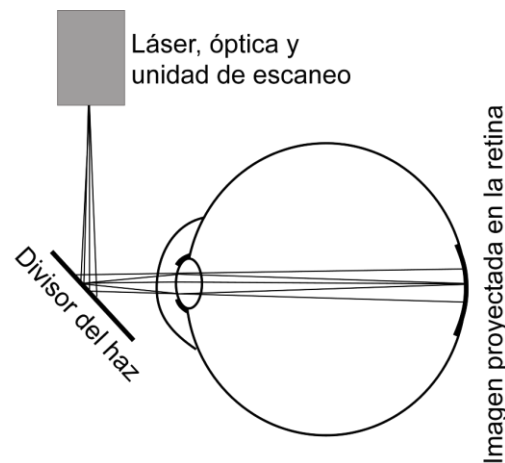


Figura 93. Esquema de un display retiniano. A partir de (Bimber et al., 2005d, 73).

Estos sistemas se han utilizado tradicionalmente en aplicaciones de tipo militar, industrial o simulaciones. En (Microvision, 2008) se han desarrollado distintos displays que utilizan la tecnología láser, incluyendo pequeños proyectores, displays para llevar puestos (*wearable*) y displays para vehículos (Figura 94).



Figura 94. Display retiniano de Microvision y aplicación de RA para ingeniería. En (Microvision, 2008).

4.1.1.2. DISPLAYS SUJETOS A LA CABEZA

Los displays sujetos a la cabeza (*head-mounted displays, HMD*) son los displays que tradicionalmente más se han utilizado en aplicaciones de realidad aumentada. Existen dos tipos de tecnologías para superponer los gráficos generados por ordenador a la imagen del mundo real vista por el usuario: HMD de vídeo (Figura 95-a), que utiliza la tecnología de mezcla de vídeos y muestra el resultado sobre los pequeños monitores ubicados delante de los ojos; HMD óptico (Figura 95-b), que utiliza combinadores ópticos (principalmente espejos semitransparentes o LCD transparentes).

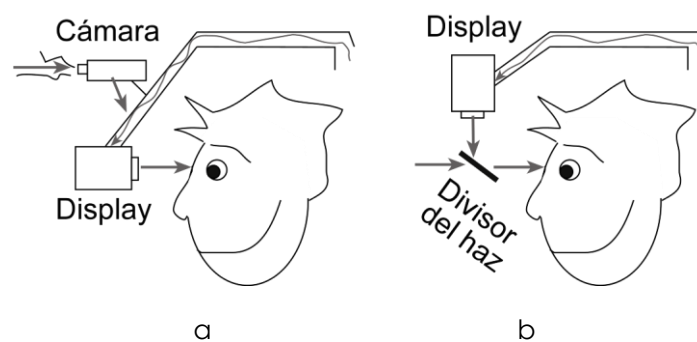


Figura 95. Esquemas de los HMD: a) HMD de vídeo; b) HMD óptico.

Existen algunas ventajas asociadas al utilizar vídeo para entornos de realidad aumentada (Barfield, 1995, 547): con una imagen vídeo es posible procesar la información visual capturada por la cámara para realzar o resaltar una porción particular de la imagen. Además, es posible montar las cámaras en un entorno que no es local al usuario. Finalmente, también es posible eliminar partes de la escena de vídeo y reemplazarla por imágenes sintéticas. Sin embargo conviene tener en cuenta una serie de desventajas (Bimber et al., 2005d, 75):

- Bajas resoluciones debido a las limitaciones de los pequeños displays. En el caso del HMD óptico, este hecho afecta solamente los gráficos generados por ordenador.
- Campo de visión limitado, debido a las limitaciones de la óptica aplicada.
- Pérdidas de percepción visual causadas por la profundidad de imagen constante. Para el caso de HMD ópticos: debido a que los objetos del entorno real y la imagen plana mostrada se registran en diferentes profundidades, los ojos son forzados a cambiar constantemente el foco entre los dos niveles de profundidades, o a percibir uno de esos niveles de forma borrosa. Este hecho se conoce como el problema de la focal fija.

- Los HMD ópticos requieren de una dificultosa calibración (para cada sesión y usuario) y precisan de un traqueo de la cabeza para asegurar una correcta superposición de los gráficos.
- Incremento de la sensación de mareo debido a la adhesión de la imagen plana en la cabeza.
- Los de tipo óptico son incapaces de proporcionar efectos de oclusión adecuados entre los objetos reales y virtuales debido a los combinadores ópticos, ya que reflejan la luz de los mini displays interfiriendo con la proveniente del entorno físico.

Para solucionar este último problema, en (Kiyokawa et al., 2000a; Kiyokawa et al., 2002) se introduce un HMD óptico que utiliza paneles LCD adicionales para bloquear la luz de forma selectiva. En la Figura 96 se muestra el dispositivo y dos imágenes, la primera de ellas sin la resolución de la oclusión.

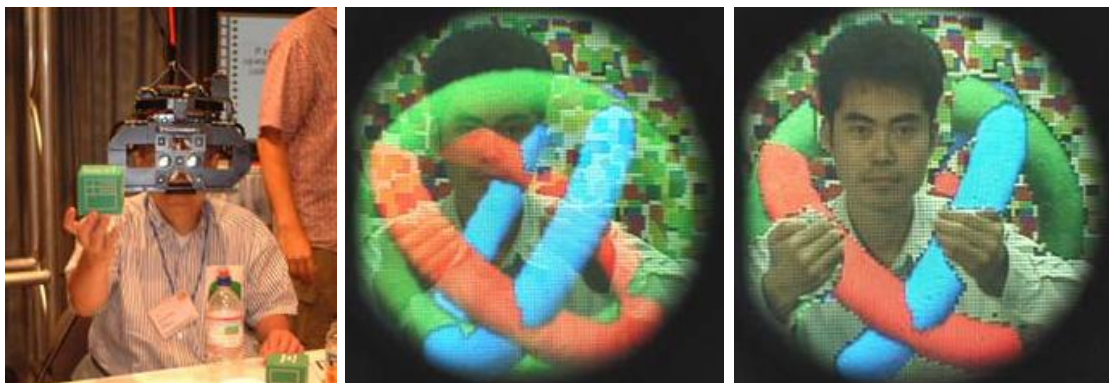


Figura 96. HMD óptico que proporciona oclusión mutua gracias a unos paneles LCD. En (Kiyokawa et al., 2002).

Los dispositivos HMD son muy comunes en aplicaciones de RA, bien de tipo óptico (Auer et al., 1999; Kato et al., 1999; Kiyokawa et al., 2000b; Wanschitz et al., 2002; Avery et al., 2005; Malkawi et al., 2005) o de vídeo (Szalavári et al., 1997; Fuhrmann et al., 1998; Höllerer et al., 1999; Höllerer et al., 2001; Cheek et al., 2003; Cheek et al., 2004; Bekel et al., 2005; Teichrieb et al., 2007); algunas de estas aplicaciones ya se han introducido en anteriores capítulos. En (Wanschitz et al., 2002) desarrollan un HMD óptico denominado *Varioscope AR* para ser utilizado en procesos de implantaciones por cirugía (Figura 97). La peculiaridad de este display es que tiene autofocus y zoom. La distancia a la que se puede trabajar varía entre 300 y 600 mm, y los aumentos entre 3.6 y 7.2. Se le añadió un sistema óptico para el registro con el que se lograron precisiones mejores de 2 mm en la superposición de los gráficos sobre el instrumental quirúrgico.

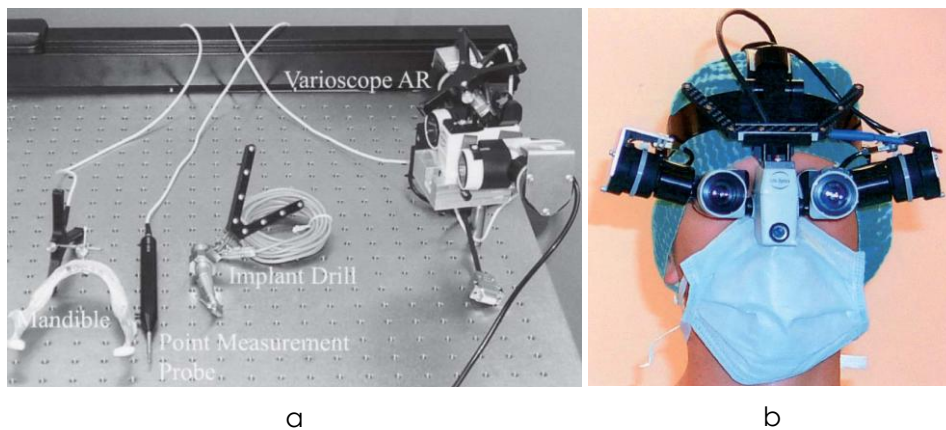


Figura 97. a) *Varioscope AR* e instrumental quirúrgico equipados con un sistema óptico para el registro; b) *Varioscope AR*. En (Wanschitz et al., 2002).

En varios ensayos de esta tesis (capítulos 8, 9 y 12) se ha utilizado el HMD de vídeo monoscópico *i-glasses PC/SVGA* (Figura 98). Algunas de las características técnicas de este dispositivo son: resolución 800x600 píxeles; 100% de superposición de imágenes; campo de visión diagonal de 26°; refresco de pantalla 100 Hz; peso de 220 g; auriculares estéreo; elementos de control: encendido, volumen del audio, color, tono, contraste, luminosidad.



Figura 98. HMD de vídeo utilizado en los ensayos de esta tesis: *i-glasses PC/SVGA*.

Respecto a otros sistemas, la firma Trivisio (Trivisio, 2008) es una de las pocas en el mercado que ofrece HMD (de vídeo) con cámaras integradas – ya que los HMD comercializados más comunes son los utilizados en sistemas de realidad virtual, los cuales no necesitan de la visión del entorno real. Tiene tanto la versión monoscópica (una sola cámara) como estereoscópica (dos cámaras), y los modelos se presentan a modo de gafas (Figura 99-a) o sujetos a la cabeza con una especie de turbante (Figura 99-b). Algunas características técnicas del modelo *ARvision-3D Goggles* son: resolución de las cámaras PAL, 1/3" CMOS; la convergencia y focal se ajustan de forma manual; iris fijo; lente de 6 mm; dos displays SVGA; refresco de 120 Hz; 100% de superposición de

imágenes; campo de visión diagonal de 40°; distancia interpupilar ajustable entre 60-70 mm; peso de 230 g; dimensiones de 155x50x90 mm. El modelo *ARvision-3D HMD* es de características similares, variando en el peso y en la distancia interpupilar (60-75 mm).

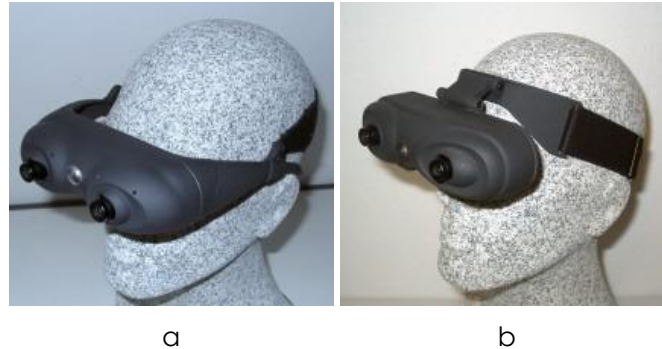


Figura 99. HMD de video con cámaras integradas: a) *ARvision-3D Goggles*; b) *ARvision-3D HMD*.
En (Trivisio, 2008).

4.1.1.3. PROYECTORES SUJETOS A LA CABEZA

Los displays de proyectores sujetos a la cabeza (*Head-Mounted Projective Displays, HMPD*) redirigen el tronco de la proyección con un espejo divisor del haz de tal forma que las imágenes son transmitidas sobre superficies retro-reflectantes colocadas en frente del visor (Figura 100). Una superficie retro-reflectante está cubierta con muchos miles de micro cubos angulares. Debido a que cada micro cubo angular tiene la propiedad óptica de reflejar la luz de vuelta a través de su dirección de incidencia, tales superficies reflejan imágenes más brillantes que las superficies normales que difuminan la luz.

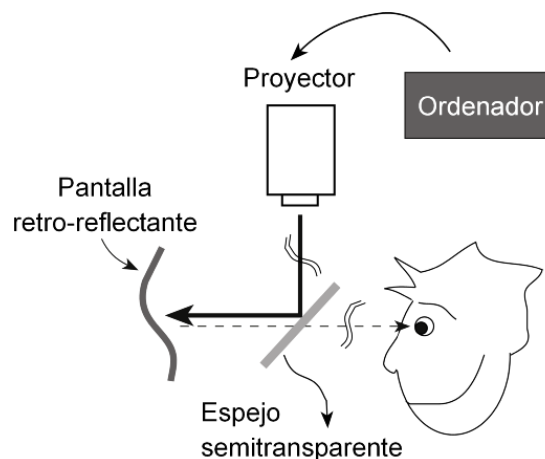


Figura 100. Esquema del funcionamiento de los proyectores sujetos a la cabeza. A partir de (Kiyokawa et al., 2000a).

En (Kijima et al., 1997) se introduce el concepto de displays sujetos a la cabeza proyectivos (*Projective Head-Mounted Displays, PHMD*), que transmiten las imágenes generadas sobre techos regulares, en lugar de sobre superficies especiales que apuntan al observador (Figura 101). Se utilizan dos espejos semitransparentes para integrar la imagen proyectada sobre el campo de visión del observador de tal forma que los parámetros de los proyectores concuerdan con los parámetros del observador. Al cambiar la disposición espacial del proyector para que se encuentre alineado con el punto de vista del observador, se evitan algunas distorsiones de la imagen que se producen al llevar incorrectamente puestos los HMD.

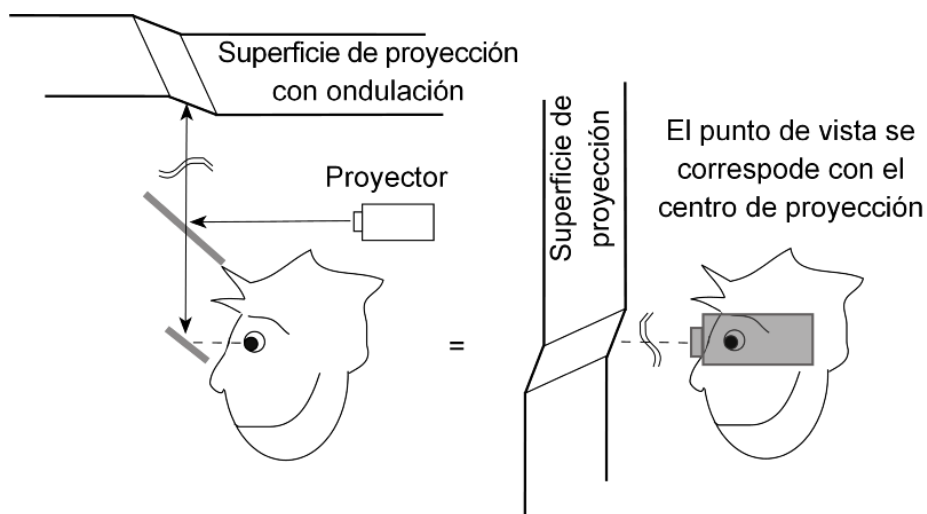


Figura 101. Esquema del funcionamiento de un PHMD. A partir de (Kijima et al., 1997).

Estos dispositivos reducen el efecto de inconsistencia de acomodación y convergencia que se relaciona con los HMD, proporcionando un mayor campo de visión sin necesidad de lentes adicionales que puedan introducir distorsiones. También previenen distorsiones incorrectas de paralaje causadas por la no correspondencia de la distancia interpupilar (si los HMD se llevan de forma incorrecta). Sin embargo, también introducen algunas desventajas (Bimber et al., 2005d, 78):

- Ambos, HMPD y PHMD, actualmente son engorrosos y pesados. Sin embargo, los nuevos prototipos tienden a ser más pequeños y ergonómicos.
- Los mini-proyectores/LCD integrados ofrecen resoluciones y luminosidad limitados.
- Los HMPD puede que requieran superficies de proyección especiales (ej. superficies retro-reflectantes) para proveer imágenes luminosas.
- Para los PHMD, la luminosidad de las imágenes depende de las condiciones de luz ambiental.

- Los PHMD sólo se pueden utilizar en espacios cerrados, ya que requieren de la presencia de un techo.

En la bibliografía se pueden encontrar diversas aplicaciones de realidad aumentada que utilizan este tipo de displays, como en (Kawakami et al., 1998; Inami et al., 2000; Rolland et al., 2002), (Figura 102-a). En (Kawakami et al., 1998) se describe una aplicación de realidad aumentada en la que se visualiza la estructura ósea de una persona a través de un dispositivo HMPD (Figura 102-b), gracias a un material recubierto por una superficie retro-reflectante.

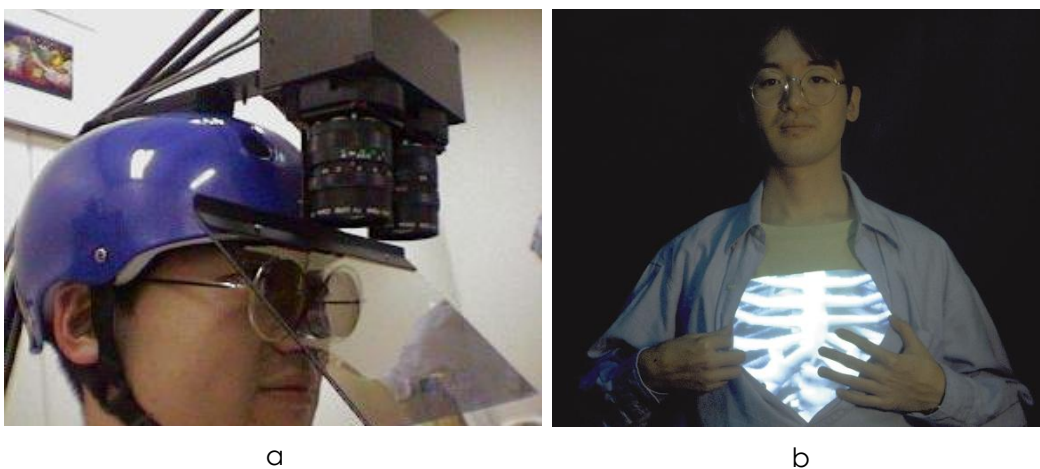


Figura 102. a) HMDP, en (Inami et al., 2000); b) visualización de la estructura ósea, en (Kawakami et al., 1998).

El mismo sistema es utilizado en (Inami et al., 2000) para presentar una aplicación en donde se simula que un objeto físico opaco se vuelve transparente al verlo a través del HMPD, produciéndose un efecto de "camuflaje óptico" (*optical camouflage*). Para lograr este efecto, el objeto físico se recubre del material retro-reflectante, y en el dispositivo HMPD lo que se proyecta es una imagen del fondo de la sala. En la Figura 103 se muestra un esquema del funcionamiento de este sistema de camuflaje y ejemplos sobre una caja y una camisa de materiales retro-reflectantes.

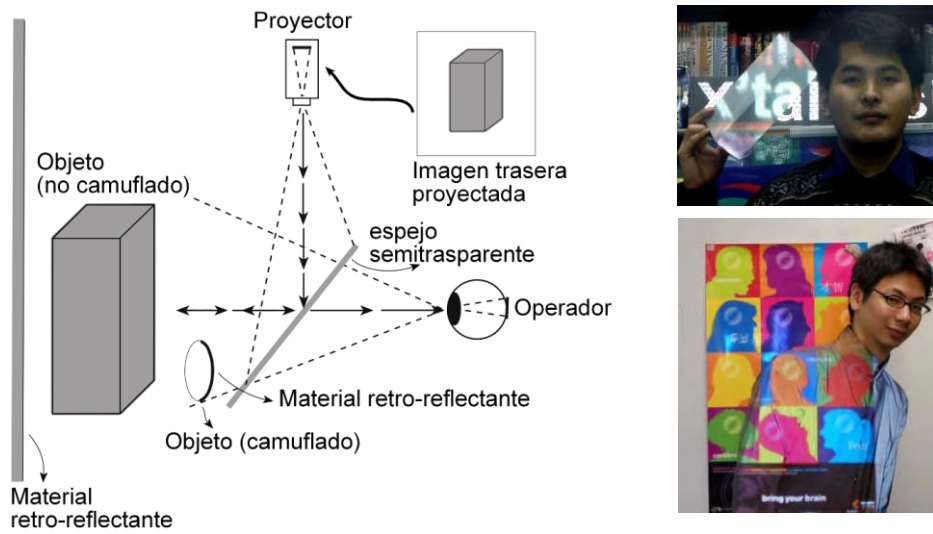


Figura 103. Esquema de camuflaje óptico mediante el empleo de HMPD y ejemplos. En (Inami et al., 2000; Bimber et al., 2007).

En (Rolland et al., 2002) se presenta un display denominado *Teleportal Head-Mounted Projective Display (THMPD)*, (Figura 104). Consiste en una óptica de mini proyecciones sujetas a la cabeza y materiales retro-reflectantes dispuestos estratégicamente en el entorno, al igual que los HMPD explicados anteriormente, aunque con dimensiones más reducidas. Como novedad, el dispositivo permite capturar imágenes estereoscópicas de la cara del usuario que lleva puesto el THMPD gracias a unos mini espejos frontales y unas mini cámaras sujetas a los lados del casco, de tal forma que la imagen de la cara del usuario se puede tele-transportar a un lugar remoto vía Internet, para formar parte de entornos virtuales o aumentados.

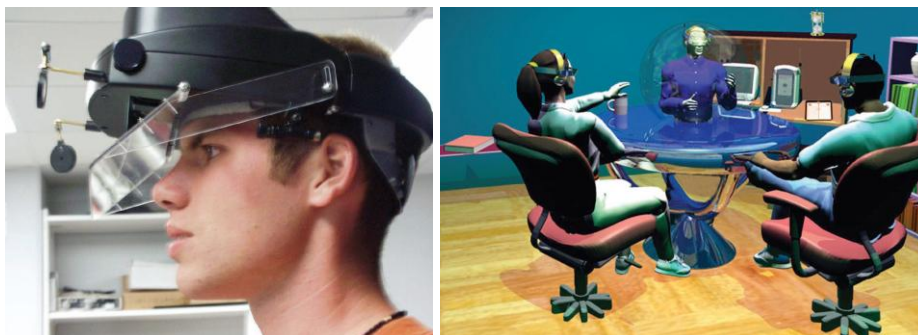


Figura 104. Sistema de HMDP óptico y tele-transportación de la cara del usuario a un espacio remoto. En (Rolland et al., 2002).

4.1.2. DISPLAYS DE MANO

Los displays de mano (*hand-held displays, HHD*), tales como tablet-PCs, PDAs o teléfonos móviles (Figura 105), generan imágenes “al alcance de las manos”. Todos estos ejemplos combinan procesador, memoria, display y tecnología de interacción en un solo dispositivo. Las cámaras integradas capturan vídeo en tiempo real del entorno, al que se le superponen elementos gráficos generados por ordenador antes de ser mostradas. Aunque la mayoría de los displays de mano están basados en imagen vídeo, al igual que en el caso de los HMD también existe la tecnología óptica y la proyectada.

Las desventajas que nos podemos encontrar con estos tipos de displays son (Bimber et al., 2005d, 80):

- El análisis y renderizado de los componentes de la imagen es intensa a nivel de procesador y memoria. Este hecho es crítico para los dispositivos de bajo coste como las PDAs y teléfonos móviles.
- El tamaño de pantalla limitado de la mayoría de los dispositivos de mano, restringe el ángulo de visión que pueden abarcar; mover el dispositivo para navegar a través de un espacio de información que es esencialmente mayor que el éste, ocasiona un fenómeno de percepción visual que se conoce como el *efecto Parks*.
- La óptica y los chips de los sensores de imagen en las cámaras integradas en dispositivos de mano comerciales no soportan auto-foco, con lo que sólo pueden ser efectivas en un cierto rango de profundidad.
- No permiten trabajar con las manos libres.

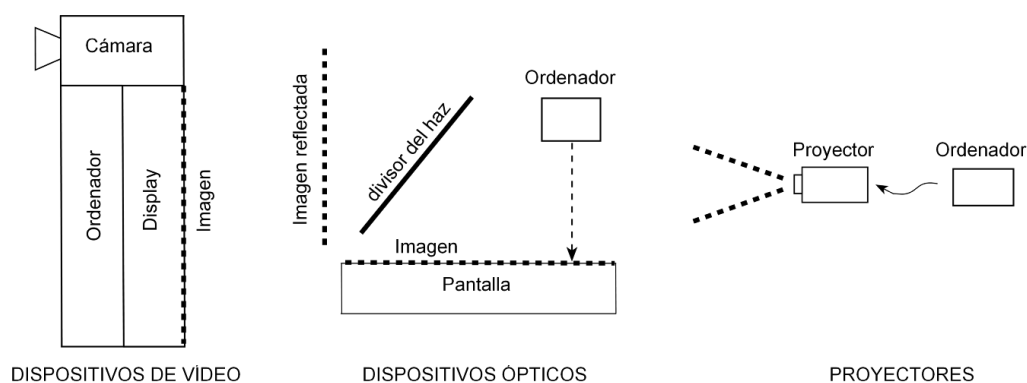


Figura 105. Esquema de los distintos dispositivos de mano empleados en RA. A partir de (Bimber et al., 2007, 121).

Los dispositivos de mano representan una alternativa real a los HMD para aplicaciones móviles. Especialmente dispositivos como las PDAs y los teléfonos móviles tienen el potencial de acercar la realidad aumentada a un público más amplio, habiéndose incrementado en los últimos años de forma considerable las aplicaciones de RA que utilizan estos dispositivos, como en (Mogilev et al., 2002; Flintham et al., 2003; Dow et al., 2005a; Dow et al., 2005b; Calado Lopes et al., 2006; Rashid et al., 2006; Reilly et al., 2006; Vogiazou et al., 2007). La mayoría de estas aplicaciones se desarrollan en espacios abiertos y utilizan GPS para su localización. Por ejemplo, en (Vogiazou et al., 2007) se describe *CitiTag*, un juego de realidad mezclada multiusuario, utilizando PDAs conectadas a receptores GPS y una red inalámbrica (Figura 106). Este juego tiene lugar en espacios urbanos y se basa en los juegos de un patio de recreo, en cuanto a simplicidad y espontaneidad. En este juego existen dos equipos (verdes y rojos), y cada usuario pertenece a uno u otro. Los jugadores deben acercarse a los usuarios del equipo contrario para “pillarlos”; si pillan a un usuario, tiene que rescatarlo uno de su mismo equipo; los jugadores son avisados por el sistema si alguien necesita ser rescatado.

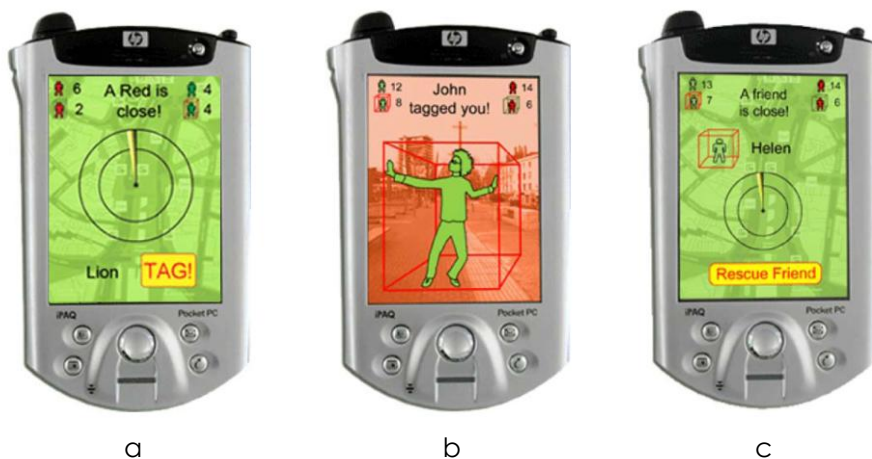


Figura 106. Diversas escenas del juego *CitiTag*: a) Aviso de que se puede “pillar” a un jugador; b) Jugador pillado; c) Aviso de rescate de un jugador del mismo equipo. En (Vogiazou et al., 2007).

En (Mogilev et al., 2002) se presenta *AR Pad*, un dispositivo de mano que consiste en una pantalla LCD, un mando de videojuegos y una cámara Web, estando todo ello conectado a un PC (Figura 107). Este dispositivo es a la vez display e interfaz, ya que permite que el usuario pueda interactuar con los objetos virtuales mostrados en el display mediante el mando; para el registro y localización de dichos objetos se utilizan las librerías ARToolkit.

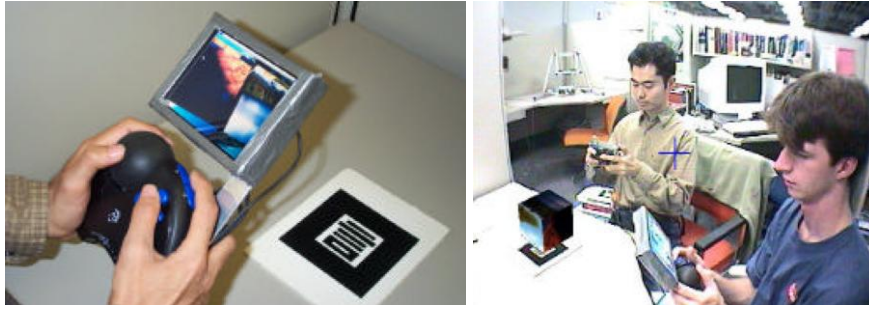


Figura 107. Dispositivo AR-Pad y usuarios interactuando. En (Mogilev et al., 2002).

También existen algunas aplicaciones de RA que utilizan dispositivos de mano de tipo óptico. Por ejemplo, en (Stetten et al., 2001) utilizan un dispositivo de mano óptico para superponer datos de tomografías en tiempo real sobre el paciente (Figura 108). Consiste en un transductor de ultrasonido que escanea láminas de objetos situados en frente de él. Estas láminas se muestran secuencialmente sobre un monitor y luego se reflejan por un espejo semitransparente de tal forma que la imagen virtual está exactamente alineada con el área de la lámina escaneada.

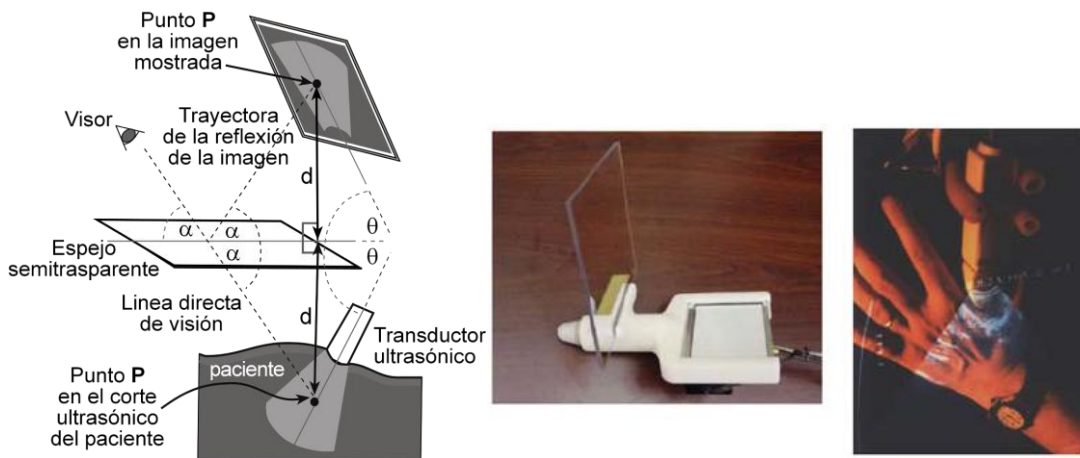


Figura 108. Esquema de funcionamiento e imágenes de un display de mano óptico con un transductor ultrasónico. En (Stetten et al., 2001).

En (Raskar et al., 2003) se introduce el concepto de *iLamp* mediante la adaptación de proyectores de vídeo para ser utilizados como displays de mano (Figura 109), actuando como verdaderas linternas que superponen diversa información sobre objetos físicos (planos o curvos). El sistema móvil consta de un proyector, una cámara Web, un sensor de inclinación y un ordenador portátil con interfaz y comunicación inalámbrica. El sistema se puede comunicar con otros dispositivos y objetos para el conocimiento de diversas relaciones geométricas en tiempo real. Se describe una aplicación para

interactuar con las proyecciones, en donde el registro de la orientación externa del sensor se hace a partir de unas marcas fiduciales situadas en el entorno físico.

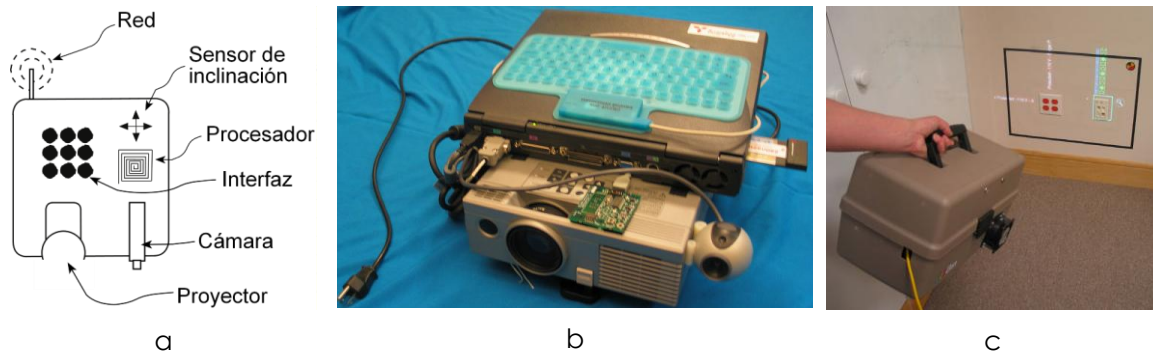


Figura 109. Dispositivo iLamp: a) Componentes; b) Prototipo de proyector aumentado; c) Registro por marcas fiduciales y superposición de proyecciones. En (Raskar et al., 2003).

4.1.3. DISPLAYS ESPACIALES

En contraste con los displays sujetos al cuerpo (sujetos a la cabeza o de mano), los displays espaciales separan la mayoría de la tecnología del usuario y la integran en el entorno físico. Al igual que en los casos anteriores, existen tres técnicas que principalmente difieren en la manera en que aumentan el entorno: de vídeo, ópticos o proyectores.

4.1.3.1. DISPLAYS ESPACIALES DE VÍDEO

Los displays espaciales de vídeo (*screen-based video see-through displays, SBVD*) hacen uso de mezcla de vídeos y muestran las imágenes fusionadas en un monitor corriente. La técnica de aumentación basada en monitores de vídeo es bastante frecuente en aplicaciones de RA no móviles. Esto es debido a que representa probablemente el sistema más económico de RA, ya que sólo se requieren componentes de hardware comerciales y un PC estándar. Sin embargo, existen una serie de desventajas (Bimber et al., 2005c, 8):

- Al igual que sucede con algunos sistemas de RV basados en monitores, estos sistemas proveen de un bajo grado de inmersión, ya que el campo de visión se limita y restringe al tamaño del monitor, su alineación espacial con respecto al observador, y su distancia al observador.

- Campos de visión reducidos debido al tamaño relativamente pequeño de los monitores.
- La mayoría proveen una vista remota, en lugar de soportar una metáfora de "ver a través de".
- Normalmente no soportan interacción directa con el entorno real y la aumentación gráfica. Solamente soportan técnicas de interacción indirecta y/o remota.

En (Wilde et al., 2003) se describe *The Periscope*, un dispositivo de RA que utiliza como display un monitor camuflado, con un sistema que permite elevarlo y rotarlo (Figura 110). *The Periscope* está diseñado para el ámbito de la educación, aumentando una "excursión al monte" mediante diversa información adicional sobre el ciclo de la vida en ese entorno. El dispositivo contiene una serie de cajas Petri dotadas con etiquetas RFID, permitiendo a los niños seleccionar diversos experimentos que serán mostrados a continuación en el monitor.



Figura 110. *The Periscope*, visualización a través del display y niños interactuando. En (Wilde et al., 2003).

4.1.3.2. DISPLAYS ESPACIALES ÓPTICOS

En contraste con los displays ópticos sujetos a la cabeza o de mano, los displays espaciales ópticos (*spatial optical see-through displays, SOD*) generan imágenes que están alineadas con el entorno físico (Figura 111). Los combinadores espaciales ópticos, tales como espejos divisores del haz, pantallas transparentes u hologramas ópticos, son componentes característicos de estos displays. Sus principales ventajas son una mayor facilidad de acomodación visual y convergencia, una resolución mayor y escalable, un mayor campo de visión, factores ergonómicos mejorados, una calibración más

sencilla y más estable, y un entorno más controlable. Sin embargo, también existen una serie de desventajas (Bimber et al., 2005c, 9):

- No soportan aplicaciones móviles.
- En la mayoría de los casos, la óptica aplicada impide una interacción directa con los objetos virtuales y reales que están ubicados detrás de la óptica.
- El número de observadores que pueden soportar simultáneamente se encuentra restringido por la óptica aplicada.
- Por las mismas razones que para los HMD ópticos, no soportan una oclusión mutua entre los objetos virtuales y los reales, aunque existen técnicas de iluminación basadas en proyectores que resuelven este problema.
- Debido al tamaño limitado de los monitores y combinadores ópticos, los objetos virtuales que se encuentran fuera del área de display se cortan de manera artificial. Este efecto se conoce como violación de ventana.

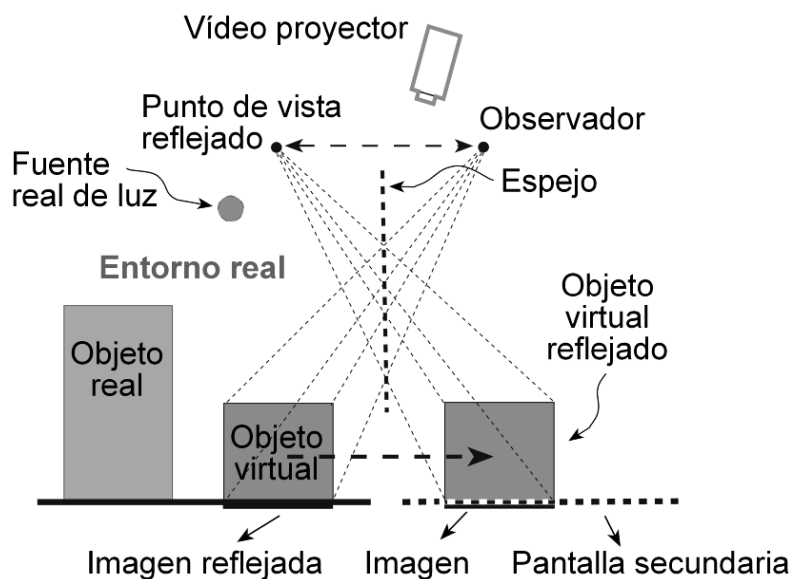


Figura 111. Geometría de la proyección en los displays espaciales ópticos. A partir de (Bimber et al., 2007, 126).

Algunas aplicaciones de RA con el empleo de estos displays se describen por ejemplo en (Kawakami et al., 1998; Inami et al., 2000; Nikou et al., 2000; Bimber et al., 2001a; Bimber et al., 2001b; Bimber et al., 2001c). En (Kawakami et al., 1998; Inami et al., 2000) se introduce el concepto de *Visuo-Haptic Display*, un dispositivo mediante el cual se pueden tocar elementos virtuales al aparecer mapeados sobre superficies físicas (Figura 112). Para lograr este efecto se utilizan proyecciones sobre espejos semitransparentes y materiales retro-reflectantes.

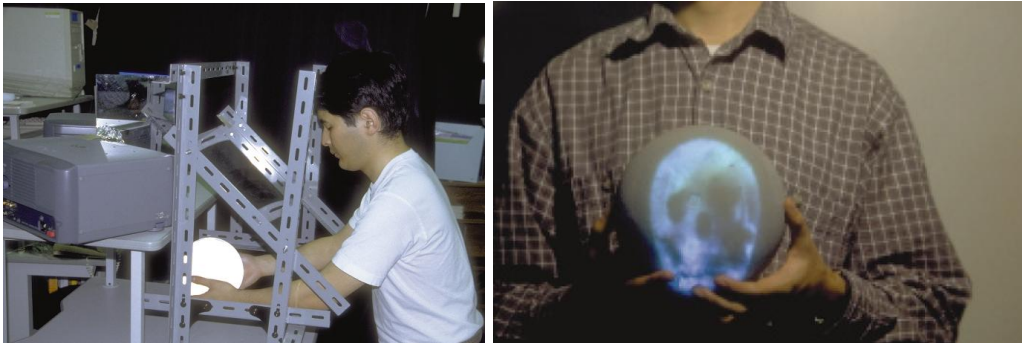


Figura 112. *Visuo-Haptic Display*. En (Kawakami et al., 1998).

En (Nikou et al., 2000) se hace un análisis de distintos dispositivos de realidad aumentada que se habían utilizado hasta ese momento en el mundo de la medicina para simular visión de Rayos-X sobre el paciente. Uno de los trabajos que describe es el realizado en el campo de la cirugía ortopédica por el Carnegie Mellon University y el University del Pittsburgh Medical Center. El dispositivo que emplean está compuesto por un display LCD y un espejo semitransparente sobre el que se visualizan los gráficos 3D. Tanto el paciente, como el espejo y la cabeza del cirujano se registran para proporcionar de forma adecuada las proyecciones de las estructuras tridimensionales (Figura 113).

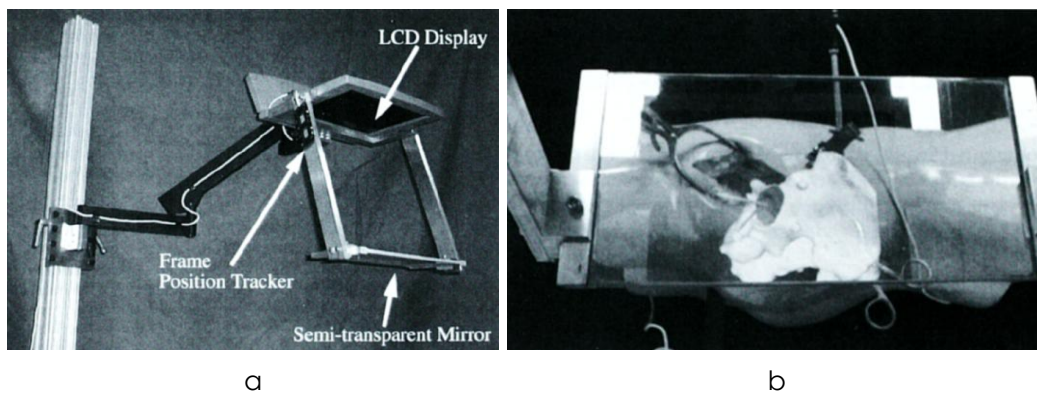


Figura 113. Simulador de visión de Rayos-X: a) Display LCD, sistema de registro (Optotrak 3020) y espejo semitransparente; b) Punto de vista del cirujano. En (Nikou et al., 2000).

En (Bimber et al., 2001a; Bimber et al., 2001b; Bimber et al., 2001c) se describen varios dispositivos. Uno de los más conocidos es el *Virtual ShowCase* (Bimber et al., 2001b), introducido anteriormente en el apartado 1.3. En (Bimber et al., 2001c) se propone un sistema de proyecciones sobre pantalla para una aplicación de RA en el ámbito de la ingeniería. La pantalla utilizada es un espejo semitransparente de dimensiones 1x1.5 m con un marco de madera alrededor, simulando una pizarra o tablón (Figura 114). Esta

pantalla se utiliza como un combinador óptico que refleja gráficos 3D integrados con el espacio físico. Al mismo tiempo, la pantalla puede ser utilizada a modo de pizarra, permitiendo a los usuarios realizar dibujos y anotaciones. Para el registro se utilizaron sensores electromagnéticos (Flock of Birds de Ascension) y un sensor híbrido (Mimio, ultrasónico e infrarrojos) para el registro 2D sobre superficies planas.



Figura 114. Espejo semitransparente a modo de pizarra. En (Bimber et al., 2001c).

4.1.3.3. DISPLAYS ESPACIALES DE PROYECTORES

Los displays espaciales de proyectores (*projector-based spatial displays, PSD*) aplican proyección frontal para proyectar imágenes directamente sobre superficies de objetos físicos (Figura 115). Se pueden utilizar tanto proyectores orientables como estáticos, o incluso varios proyectores para incrementar el área de display. Según (Bimber et al., 2005c, 9-10), si se muestran gráficos tridimensionales sobre la superficie de los objetos, se requiere una proyección estereoscópica, que a su vez depende del punto de vista del observador. Algunas ventajas de estos displays son que proporcionan un campo de visión teóricamente ilimitado, resolución escalable, y una mayor facilidad de acomodación visual. Sin embargo, también introducen nuevos problemas:

- Aparición de sombras sobre objetos físicos y sobre los usuarios que interactúan, debido a la proyección frontal.
- Restricciones del área de display que se limita al tamaño, forma y color de la superficie de los objetos físicos.
- Condicionado a un solo usuario en el caso de sistemas estereoscópicos.
- Los proyectores convencionales solo pueden focalizar en un único plano focal situado a una distancia constante. Por lo tanto, al proyectar sobre superficies no planas se producen imágenes borrosas. Como excepción están los proyectores-láser.

- La complejidad de alineamiento geométrico consistente y calibración del color crece con el número de proyectores empleados.

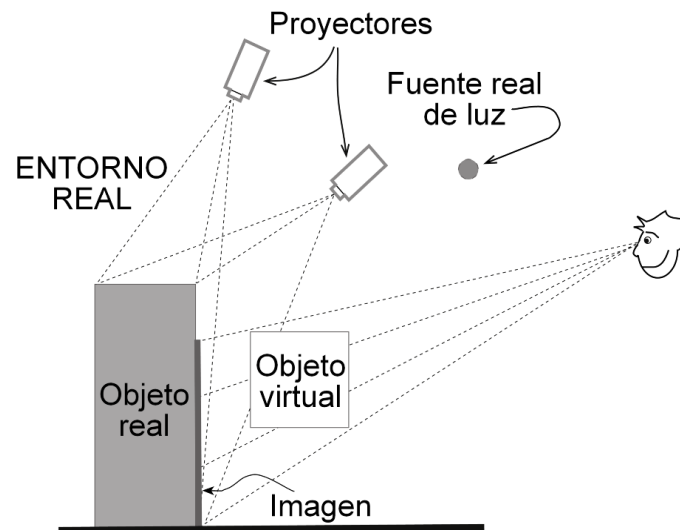


Figura 115. Esquema de los displays espaciales de proyectores. A partir de (Bimber et al., 2007, 131).

En (Ben-Joseph et al., 2001; Low et al., 2001; Ciger et al., 2003; Bimber et al., 2005a; Bimber et al., 2005b; Bimber et al., 2005d; Bimber et al., 2007) se describen varias aplicaciones con el uso de estos displays. En (Ben-Joseph et al., 2001) se utilizan proyecciones sobre una mesa de trabajo (ver apartado 1.4.7). En (Bimber et al., 2005d, 280) se describe el término de *Shader Lamps*, que define un tipo de proyección sobre objetos físicos tridimensionales, mediante la utilización de diversos proyectores dispuestos estratégicamente en el entorno, de manera que el objeto físico quede perfectamente mapeado y tenga una textura geoméricamente correcta (Figura 116). Estas proyecciones se pueden utilizar además para alternar texturas, iluminación o añadir animaciones. Para lograr este efecto, en una primera fase (pre-proceso) el objeto físico debe ser modelado, para adquirir una figura 3D virtual del mismo. A continuación los proyectores se posicionan; para ello en primer lugar se disponen en el espacio de forma aproximada y se ajusta su relación de geometría con respecto al objeto físico mediante el cálculo de los parámetros de orientación interna del proyector y la matriz de transformación entre el objeto y el proyector. Para finalizar esta fase se calcula una corrección de intensidad. En una segunda fase (proceso) las imágenes se proyectan sobre el objeto de acuerdo a la matriz de transformación calculada y se le aplica la corrección de intensidad dependiente de la geometría del modelo. Por último, deben aplicarse diversas atenuaciones en los solapamientos de las proyecciones.

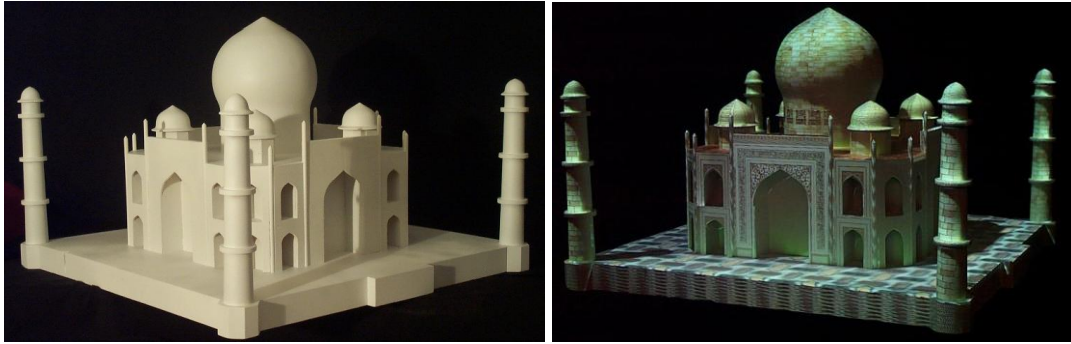


Figura 116. Objeto tridimensional de color neutro y el mismo objeto texturizado mediante *Shader Lamps*, para lo que se han utilizado un total de tres proyectores. En (Bimber et al., 2007, 162).

En (Low et al., 2001) se utiliza el principio de *Shader Lamps* en un sistema de proyecciones para aumentar algunas paredes dentro de un espacio cerrado, simulando una geometría tridimensional sobre cada uno de sus planos (Figura 117). Los autores señalan que las principales ventajas de este sistema son: campo de visión extenso (en contraposición a los HMD), exploración física del lugar (en contraposición a entornos virtuales), reducción a la sensibilidad de los errores de registro, reducción en la latencia del sistema, visión auto-estereoscópica (sin necesidad de tecnología de estereovisión), adición natural de virtualidad aumentada y provisión de interfaz háptica.



Figura 117. Paredes y paredes aumentadas mediante proyecciones. En (Low et al., 2001).

4.2. DISPLAYS SONOROS

El display que reproduce la señal digital sonora hacia el mundo analógico es el altavoz o auricular. Este display es un transductor, en concreto, un transductor electroacústico que sigue un doble procedimiento: eléctrico-mecánico-acústico. En la primera etapa convierte las ondas eléctricas en energía mecánica, y en la segunda convierte la

energía mecánica en energía acústica. Es por tanto la puerta por donde sale el sonido al exterior desde los aparatos que posibilitaron su amplificación, su transmisión por medios telefónicos o radioeléctricos, o su tratamiento.

Según los estudios aportados en (Zhou et al., 2004), la adición de sonido 3D puede ser un modo muy efectivo de complementar la experiencia del usuario en entornos de RA. Ayuda a la percepción de profundidad, realización de tareas y facilita la colaboración entre usuarios. Además, permite un entorno más realista y una sensación de inmersión mayor.

Existen diversas aplicaciones de RA que utilizan sonido (Behringer et al., 1999; Hatala et al., 2004; Dow et al., 2005a; Dow et al., 2005b; Mañas et al., 2005; Le Prado et al., 2007). En (Behringer et al., 1999) – explicado en el apartado 1.4.9 – se utiliza un sistema de audio 3D para indicar la localización de los objetos que no están dentro del campo de visión del usuario. En (Hatala et al., 2004) se presenta *ec(h)o*, una interfaz de realidad aumentada que utiliza paisajes sonoros espacializados y una aplicación Web para gestionar la información (Figura 118). El sistema se desarrolla para un museo de historia y ciencias naturales, creando una experiencia museística que consiste en aumentar las piezas del museo mediante la vinculación de paisajes sonoros. Las librerías de sonidos están conectadas a otros museos mediante una aplicación Web. La interfaz permite a los usuarios interactuar con el sistema mediante el movimiento y manipulación de ciertos objetos, sin la necesidad de utilizar directamente un ordenador.

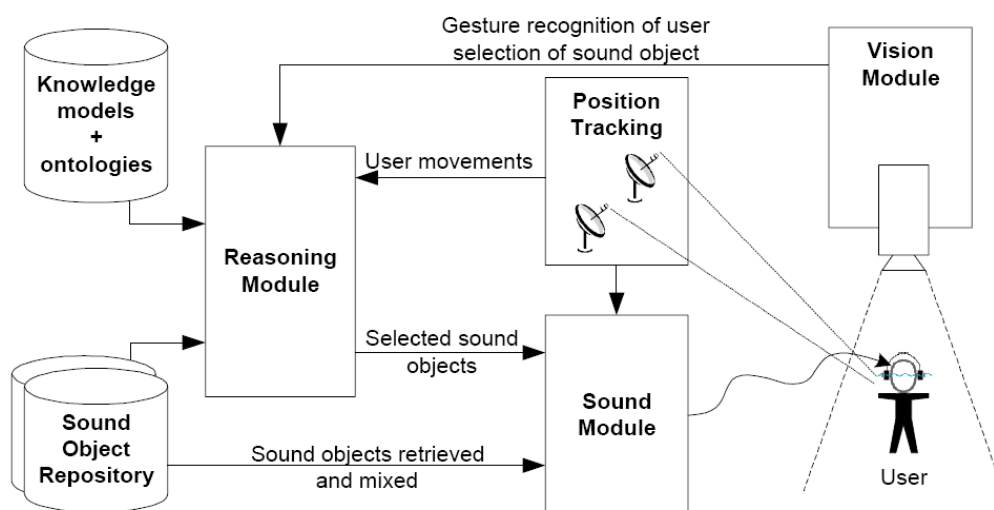


Figura 118. Esquema del funcionamiento de *ec(h)o*. En (Hatala et al., 2004).

En (Dow et al., 2005b) se presenta *WOz*, una aplicación de realidad mezclada basándose en una simulación del Mago de Oz. Se establece una acción narrativa contada por actores que interpretan los personajes del cuento, dando lugar a una experiencia auditiva que se desarrolla en un lugar histórico. En (Mañas et al., 2005) se describe *Notbook AR* (2005), una instalación artística multiusuario de narración multimedia (Figura 119), elaborada por Laboratorio de Luz y MedicLab de la UPV. El sistema propone al espectador explorar los caminos que forman distintas lecturas interrelacionadas, jugando con las diversas pistas que le ofrecen las imágenes 3D y el sonido. La instalación se desarrolla en un espacio cuadrangular vacío, donde dos usuarios se sitúan llevando en sus manos dos tablet-PC equipadas, cada una de ellas, con una cámara Web, un sensor inercial y auriculares. Cada usuario ve en su display objetos 3D superpuestos a la imagen del espacio físico. Esos objetos recrean el despliegue de fragmentos de páginas sueltas y flotantes que se mueven sutilmente en el espacio, desligadas del orden que le impondría un libro. El mapeado de los fragmentos de página hace referencia gráficamente a unos textos que el usuario escucha a través de los auriculares. El audio es la lectura de tres textos relacionados con el barrio del Cabanyal de Valencia. En el espacio de la sala se articulan tres zonas de audio que están relacionadas con los objetos 3D de la simulación. Uno de los audios nos cuenta la repercusión sobre el entorno natural de los cambios ocurridos en el barrio desde su asentamiento. El segundo narra los cambios económicos de la zona tomando el puerto como motor principal. Por último la tercera zona de audio cuenta la relación de las personas que habitan el barrio incidiendo en la memoria individual y colectiva de los vecinos.



Figura 119. Usuarios interactuando con *Notbook AR* en una demo del congreso ACE'05, UPV.

En (Le Prado et al., 2007) se presenta *LISTEN LISBOA*, una pieza en la que los usuarios experimentan la sensación de adentrarse en un mundo sonoro virtual coexistente con el entorno real (Figura 120). Los sonidos han sido grabados en Lisboa, bien de forma

monofónica, estereofónica o ambisónica (sonido 3D), y posteriormente transformados y editados con Protools para ser utilizados como composiciones interactivas escuchadas por los auriculares de los usuarios. Se utilizan un total de cinco fuentes de sonido: tres voces, el puerto y una zona musical. La navegación del usuario está planteada como si se tratase de un videojuego en el que hay varios niveles; en el primer nivel solo se escuchan las voces para que el usuario tome conciencia de la relación entre el espacio virtual sonoro y el entorno físico; en el segundo nivel se pueden explorar dos zonas virtuales mapeadas en el suelo para entender la cartografía invisible del sonido; en el tercer y último nivel se añaden algunos constreñimientos dependiendo del tiempo que el usuario pase en un área determinada y de la habilidad que tenga de atender a un personaje virtual, apareciendo nuevos sonidos.

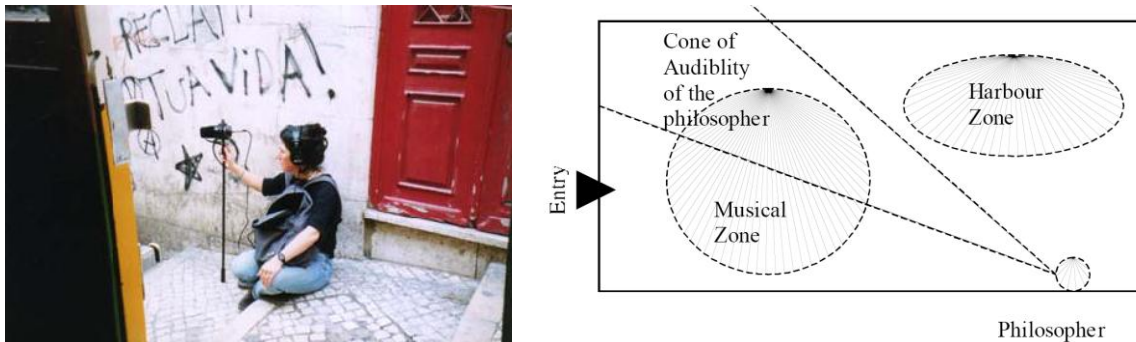


Figura 120. *LISTEN LISBOA*. Compositora grabando sonidos y cartografía del sonido virtual. En (Le Prado et al., 2007).

4.3. DISPLAYS HÁPTICOS

En contraste con la naturaleza sensorial de la visión y audición, únicamente los sistemas hápticos permiten a la vez sentir y actuar sobre el entorno, por eso son controladores y displays simultáneamente (ver apartado 6.1), conociéndose con el nombre genérico de interfaces hápticas. Estos dispositivos se emplean para tareas que normalmente se realizan con las manos en el entorno real, tales como exploraciones y manipulaciones de objetos. Las interfaces hápticas más comunes utilizadas en entornos de realidad aumentada son aquellos guantes o exoesqueletos que pueden devolver fuerzas al usuario, explicadas ya en el apartado 3.1.

Existen sin embargo otros desarrollos de sistemas hápticos que funcionan únicamente como displays. En la Figura 121 se puede ver *SmartFinger* (Ando et al., 2002), un

dispositivo que se sujeta en las uñas de los dedos, proporcionando al usuario sensaciones táctiles adicionales. Según su autor, un usuario que lleve puesto el *SmartFinger* y pase su dedo por un dibujo, puede sentir sus trazos en forma de bordes y rugosidades virtuales a partir de su textura. Esto se consigue generando pequeñas vibraciones con una bobina de voz (*voice coil*). Otros elementos del sensor son un foto detector (*photo detector*) y un sensor para las uñas (*fingernail sensor*).

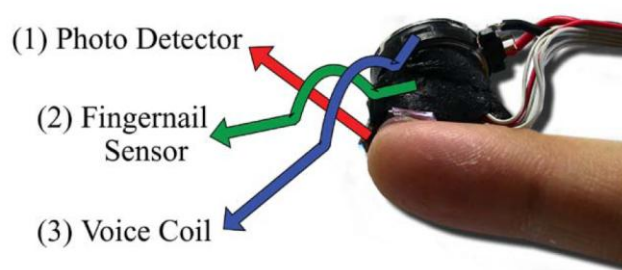


Figura 121. Componentes de *SmartFinger*. En (Ando et al., 2002).

En SIGGRAPH 2006 se presentó *Forehead Retina System* (Figura 122), desarrollado por Hiroyuki Kajimoto en la Universidad de Tokio. El sistema se pone en la frente como un turbante y proporciona al usuario una representación háptica de lo que está delante de él. El usuario también lleva puestas unas gafas de sol con una mini cámara integrada que captura la imagen del entorno en tiempo real, para procesarla en un pequeño ordenador que envía (de forma esquemática) los datos analizados a un total de 512 electrodos situados en la frente del usuario, de tal forma que el usuario puede “sentir lo que ve”.



Figura 122. *Forehead Retina System*. En (SIGGRAPH06, 2006).

4.4. OTROS DISPLAYS

La mayoría de los displays existentes en la actualidad son de tipo visual, sonoro o háptico. Hace una década se afirmaba que estos eran los únicos estímulos reproducibles en un ordenador (Barfield, 1995, 568-570). Sin embargo, existen algunos

desarrollos en el campo olfativo y gustativo que tratan de reproducir sintéticamente olores y sabores.

Por ejemplo, en (Hashimoto et al., 2005) se describe *Straw-like User Interface*, un sistema novel que reproduce de forma virtual las sensaciones que se experimentan al beber (Figura 123). Estas sensaciones se crean a partir de datos reales de presión, vibración y sonido producidos al beber de forma natural a través de una pajita. El sistema está dispuesto de tal forma que se puede experimentar la sensación de beber tanto líquido como alimentos sólidos.

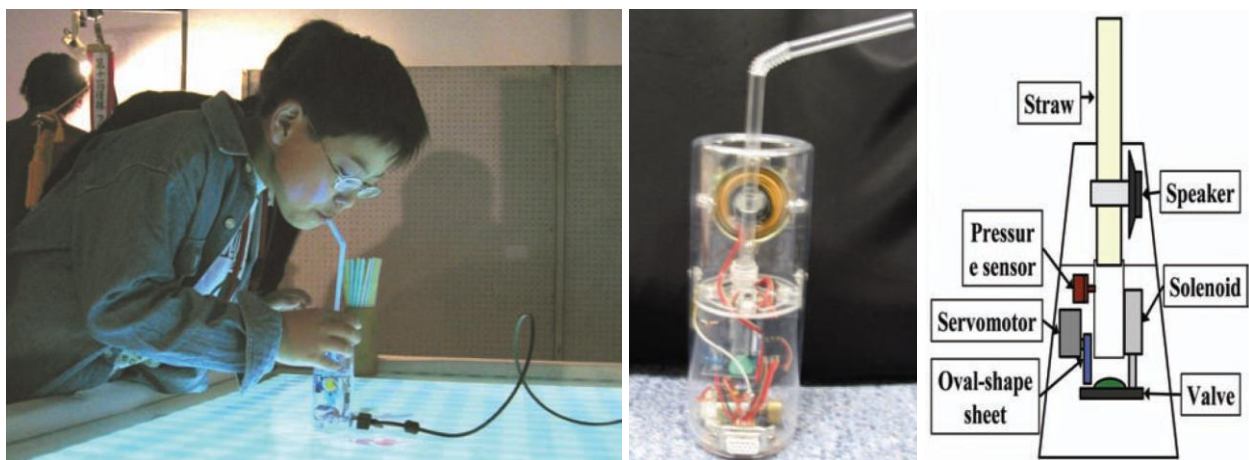


Figura 123. Usuario interactuando con *Straw-like User Interface*, imagen del dispositivo y esquema. En (Hashimoto et al., 2005).

En (Yamada et al., 2006) se presenta un display olfativo para llevar puesto (*wearable olfactory display*) que reproduce la espacialización del olor en un entorno abierto (Figura 124). El sistema trata los olores en estado gaseoso, y el aire perfumado se proyecta a la nariz del usuario por medio de unos tubos. Además, dependiendo de la posición del usuario en el espacio, estos olores son más o menos fuertes, según se aleje o acerque a aquello que teóricamente es lo que huele, siendo el usuario quien elige la posición de la fuente del olor. Versiones posteriores de este sistema utilizan un dispositivo de chorro de tinta que contiene los olores de forma líquida.

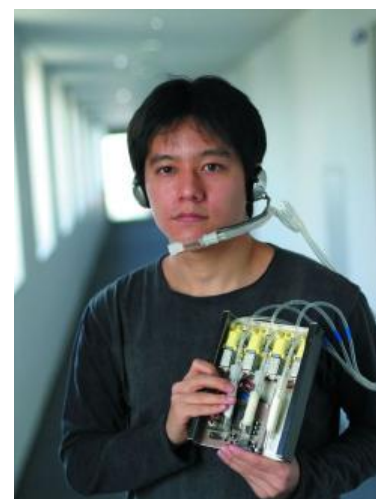


Figura 124. *Wearable olfactory display*. En (Yamada et al., 2006).

En (Yanagida et al., 2004) se describe un sistema que proyecta un grupo de olores cerca de la nariz del usuario (Figura 125). Para ello se utilizan cañones de aire que generan remolinos toroidales de las fragancias. La nariz del usuario se registra mediante una cámara de tal forma que los cañones se puedan dirigir a su localización.

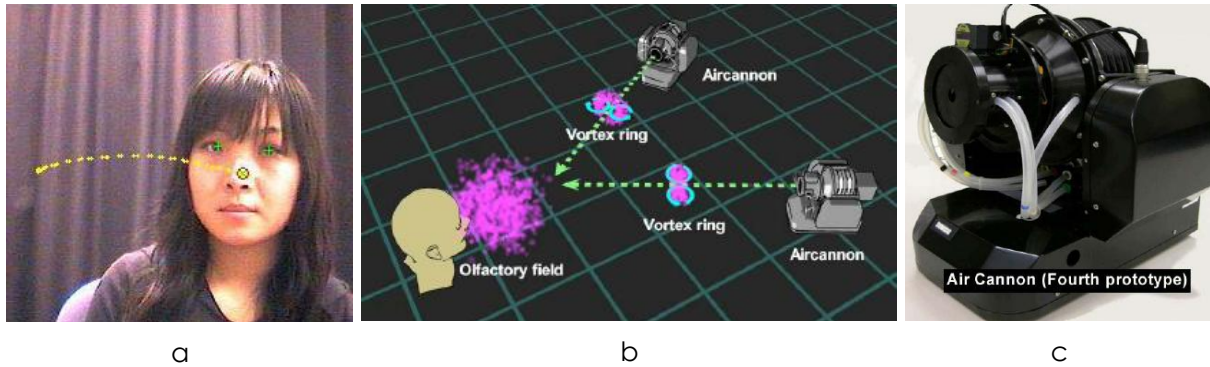


Figura 125. SpotScents: a) Registro de la posición de la nariz; b) Esquema del funcionamiento del sistema; c) Cuarto prototipo del cañón de aire. En (Yanagida et al., 2004; Bimber et al., 2007).

4.5. SISTEMAS MULTIMODALES

Los sistemas multimodales son aquellos que integran distintos estímulos reproducidos de forma virtual. Con este tipo de sistemas se enriquecen de manera considerable los entornos aumentados, proporcionando un mayor grado de inmersión mental (o sensación de presencia) del usuario. Este hecho ha sido corroborado en (Ha et al., 2007), en el que se realizan varias pruebas para determinar el grado de inmersión de los usuarios dependiendo de los estímulos reproducidos (solamente visual, sonoro y háptico). El trabajo concluye que los sistemas que mayor grado de inmersión proporcionan son los integrados por “visual + sonoro + háptico”, seguido por “visual + háptico” y “visual + sonoro”.

En (Barfield, 1995, 568-570) se muestra una tabla que relaciona los distintos usos que se le pueden dar a la tecnología RA en función de estímulos sensoriales reproducidos virtualmente (sólo para las más comunes: visual, sonora y háptica), estableciendo usos tanto para sistemas mono como multi-sensoriales. Esta tabla se muestra a continuación:

Estímulos	Método de presentación	Posibles usos/tareas
Visual	<i>Comentarios generales</i>	<i>Todas las técnicas se pueden emplear en mono o estero-visión si se requiere</i>
	HMD de vídeo	Tareas que requieran visión de 360° y elevada fidelidad visual; tareas que incluyan registro visual con el propio cuerpo del usuario/operador
	HMD óptico	Tareas que requieran visión de 360°; tareas que requieran grandes IFOV (campo de visión instantáneo); tareas que requieran grandes resoluciones del mundo real; particularmente apropiadas para entornos locales
	Monitor o pantalla de proyección	Minimiza los instrumentos que el usuario lleva consigo; fácil implementación para múltiples usuarios; tareas que requieran elevada fidelidad visual
Sonoro	<i>Comentarios generales</i>	<i>RA mezclada se puede obtener mediante la supresión de sonidos del mundo real con reducción activa del sonido</i>
	Altavoces de campo libre (<i>free-field</i>)	Minimiza los instrumentos que el usuario lleva consigo; requiere tantos altavoces como localizaciones de fuentes de sonido; no se restringe el sonido del mundo real; tareas que requieran elevada precisión de localización del sonido; no requiere calibración individual
	Sonido convolucionado presentado mediante altavoces	Minimiza los instrumentos que el usuario lleva consigo; adecuado para fuentes de sonido múltiples y móviles; solo produce alta definición de la fuente del sonido con auriculares individuales; no se restringe el sonido del mundo real; adecuado para múltiples usuarios (con restricción de la localización de los usuarios)
	Sonido convolucionado presentado mediante auriculares	Para tareas multiusuario; permite determinar la reducción de sonidos no deseados del mundo real; adecuado para tareas que requieran alta definición de la localización de la fuente del sonido

Háptico	Comentarios generales	<i>Para tareas tales como telepresencia, manipulación remota de objetos; los interfaces táctiles interfieren con la percepción de la información táctil del mundo real; los aparatos táctiles están actualmente limitados al área superficial de la piel</i>
	Aparatos de esfuerzo (<i>force-exerting</i>) corporales (<i>body-mounted</i>) utilizando neumáticos o fuerzas electromagnéticas	Tareas que necesiten fuerzas resistivas que contrarresten la contracción de los músculos; tienen la capacidad de moderar los niveles de libertad del movimiento (los ejemplos incluyen el entrenamiento de los usuarios para interactuar en entornos con alta o baja gravedad y con respuesta de la fuerza ocasionada por los movimientos)
	Aparatos de esfuerzo (<i>force-exerting</i>) no corporales (<i>world-mounted</i>): Normalmente incluyen equipamientos hidráulicos o electromecánicos	Tareas que requieran escalar las fuerzas ejercidas por el usuario o expuestas a él (ejemplo: la dirección asistida de un automóvil)
	Aparatos hápticos corporales (<i>body-mounted</i>)	Tareas que incluyan detección de contacto físico tales como detección de la localización de la superficie y características de la propiedad de la superficie; amplias posibilidades de aplicación para personas con deficiencias visuales o auditivas
Visual y sonoro	Comentarios generales	<i>Se pueden hacer muchas combinaciones de presentación, pero no todas ellas resultan en buenas simbiosis; esta RA bi-modal puede resultar más conveniente que cualquier otra RA uni-modal</i>
	HMD combinado con auriculares	Tareas que requieran movilidad de los usuarios; tareas que requieran grados elevados de orientación espacial o tareas que requieran la búsqueda de objetivos
	Monitor de ordenador o sistema de proyección combinado con sonido convolucionado presentado con dos altavoces	Alta fidelidad de sonido e imagen; tareas multiusuario; tareas con limitada movilidad de usuario; minimiza el instrumental llevado por los usuarios

	<i>Comentarios generales</i>	<i>Estas modalidades son particularmente adecuadas para la investigación y manipulación de objetos virtuales no sonoros en el entorno real</i>
Visual y háptico	Monitor de ordenador o sistema de proyección combinado con aparatos de fuerza-retroalimentada (<i>force-feedback</i>) anclados al suelo (<i>world-mounted</i>)	Manipulación de objetos muy pesados o altas precisiones en el posicionamiento de los objetos en el mundo real (los objetos pueden ser virtuales)
	HMD combinado con dispositivos táctiles	Tareas en las que el usuario puede tocar objetos reales y virtuales en un entorno enteramente integrado; tareas que incluyen la exploración de las características o límites de superficies virtuales
Sonoro y háptico	<i>Comentarios generales</i>	<i>Estas modalidades se refuerzan en lugar de complementarse y se pueden utilizar para objetos de advertencia auditivos y táctiles</i>
	Sonido espacializado de múltiples altavoces combinados con dispositivos táctiles	Refuerzo de la detección de colisión de objetos donde el control directo visual carece de suficiente precisión; exploración auditiva y táctil de entornos virtuales
Visual, sonoro y háptico	<i>Comentarios generales</i>	<i>Ésta representa la multi-modalidad de RA más avanzada; se puede aplicar para tareas que requieran una gran sensación de presencia</i>

Tabla 4. Usos de la RA en función de los estímulos virtualmente reproducidos. En (Barfield, 1995, 568-570).

5. SOFTWARE

Desde el desarrollo de las librerías ARToolKit (introducidas en el apartado 1.3), se han derivado diversas librerías, generalmente de tipo freeware, diseñadas para realizar las tareas necesarias de registro así como la composición de la escena aumentada en tiempo real, tales como MXRToolKit, ARTag o Studierstube. Este tipo de librerías utilizan técnicas de visión por computador basadas en el registro de una serie de marcas (patrones de marcas planas) para el cálculo de la matriz de transformación, siendo su uso muy extendido entre la comunidad científica. Sin embargo, existen otras librerías de visión por computador en las que el posicionamiento se realiza mediante el registro de rasgos naturales, como es el caso de las librerías BazAR.

Las librerías mencionadas anteriormente generalmente se basan en técnicas de visión por computador, sin considerar otro tipo de sensores (a excepción de OSGAR, apartado 5.1.5). Es por ello que su uso debe realizarse en otro tipo de software existente o integrarse mediante programación por código. Existen sin embargo otro software, como Max/MSP Jitter, que permiten la generación de sistemas interactivos en tiempo real, permitiendo la incorporación de sensores para la entrada de datos y de displays para la generación de la escena aumentada, así como la posibilidad de generar datos multimedia (gráficos, sonoros, etc.).

5.1. SW BASADO EN PATRONES DE MARCAS PLANAS

Los programas que se van a explicar a continuación están específicamente diseñados para la elaboración de aplicaciones de RA. No sólo realizan el registro en tiempo real, sino que además, realizan las tareas necesarias para combinar los mundos reales y virtuales.

5.1.1. ARTOOLKIT

ARToolKit (Human Interface Technology Laboratory, 2007) es un software que consta de un conjunto de librerías escritas en C y C++ que permite a los usuarios el desarrollo de aplicaciones de RA. Utiliza técnicas de visión por computador para el cálculo de la orientación externa de la cámara relativa a una serie de marcas con unas características específicas, permitiendo que se superpongan a ellas objetos virtuales. Funciona con los sistemas operativos de SGI IRIS, Linux, Mac OS X, y Windows (95/98/NT/2000/XP). Soporta aplicaciones de RA basadas tanto en displays de video como ópticos. ARToolKit se distribuye gratuitamente para su uso no comercial bajo licencia GPL.

Estas librerías son muy populares, habiéndose utilizado por muchos autores y en muy distintas aplicaciones, como en (Kato et al., 1999; Kato et al., 2000b; Billinghurst et al., 2001a; Kato et al., 2001; Slay et al., 2001; Hedley et al., 2002; Sinclair et al., 2002; MacIntyre et al., 2003; Andersen et al., 2004; Asai et al., 2004; Cooper et al., 2004; Fischer et al., 2004; Galantay et al., 2004; Lee et al., 2004; Matyszczok et al., 2004; Barakonyi et al., 2005; Walczak et al., 2005; Calado Lopes et al., 2006; Park et al., 2006; Barakonyi et al., 2007), alguna de las cuales se han explicado brevemente en apartados anteriores.

5.1.1.1. PRINCIPIOS BÁSICOS

Las tareas de registro e identificación de las marcas, cálculo de la matriz de transformación y renderizado de modelos virtuales, las realiza ARToolKit mediante la secuencia de procesos esquematizada en la siguiente figura:

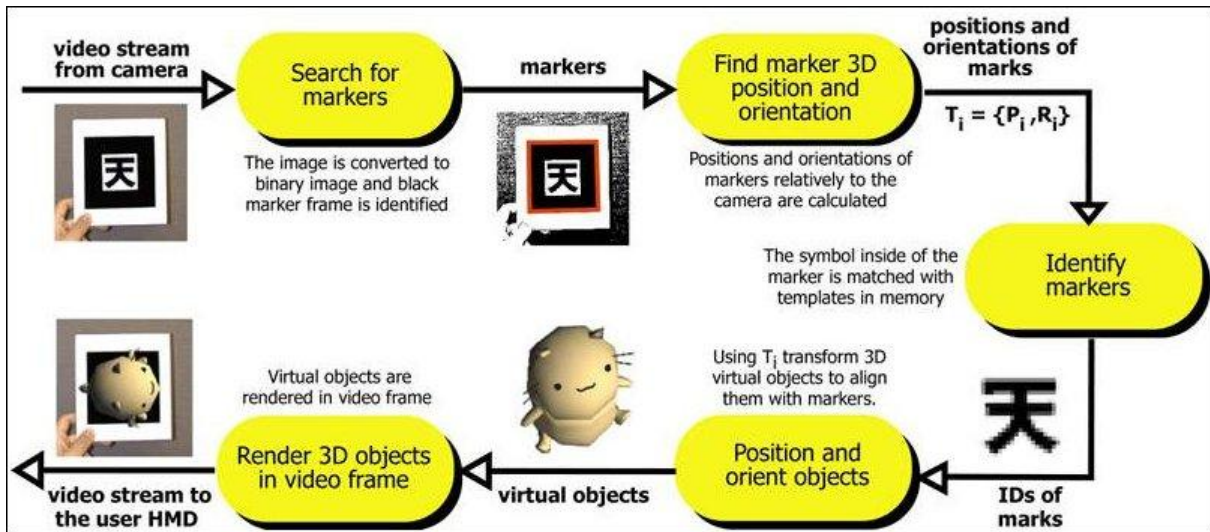


Figura 126. Secuencia de los procesos de registro de una marca y superposición de elementos virtuales en ARToolkit. En (Kato et al., 2000a, 8).

Estos procesos deben formar parte del código de cualquier aplicación programada con las librerías de ARToolkit siguiendo un orden concreto que se detalla a continuación:


Inicialización	1. Inicializar la captura de video y leer los ficheros de los patrones de las marcas y los parámetros de la cámara.
Bucle principal 	2. Seleccionar un fotograma de la entrada de vídeo.
	3. Detectar las marcas y reconocer los patrones en el fotograma de entrada del vídeo.
	4. Calcular la transformación de la cámara relativa a los patrones detectados.
	5. Dibujar los objetos virtuales sobre los patrones detectados.
Cierre	6. Cerrar la captura del vídeo.

Tabla 5. Secuencia del código en ARToolkit. En (Human Interface Technology Laboratory, 2007).

Los pasos 2 a 5 se repiten continuamente hasta que la aplicación se cierra, mientras que los pasos 1 y 6 sólo se ejecutan al empezar y terminar la aplicación, respectivamente.

5.1.1.2. CALIBRACIÓN DE LA CÁMARA

La calibración de la cámara es un paso previo a la ejecución de las aplicaciones de ARToolkit. Con la calibración, el programa es capaz de corregir las distorsiones de la imagen capturada por la cámara. Los parámetros por defecto de la cámara se

guardan en un archivo llamado *camera_para.dat*, que se lee cada vez que se inicializa el programa. Aunque estos parámetros pueden ser útiles para una gran variedad de cámaras, es recomendable utilizar un archivo separado de calibración para cada una de las cámaras que se van a utilizar. La calibración de la cámara se puede hacer mediante ARToolKit con una serie de ejecutables denominados *calib_distortion* y *calib_cparam*. El primero de ellos se utiliza para medir el punto central de la imagen y la distorsión de la lente, mientras que el segundo mide parámetros adicionales. La calibración se hace también por técnicas de visión por computador mediante el empleo de patrones que consisten en una malla de nodos circulares en el primer caso (Figura 127), y una rejilla de cuadrados en el segundo (Figura 128). El usuario debe seguir en todo momento las instrucciones dadas por los programas, como voltear la cámara o alejar la cámara de los patrones.

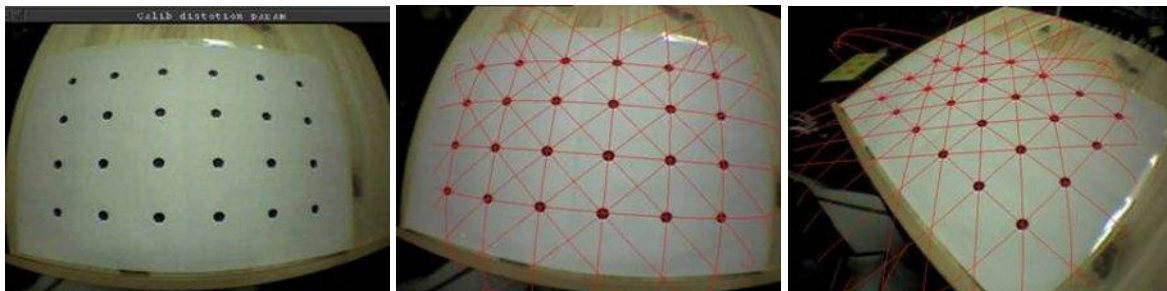


Figura 127. Patrón de *calib_distortion* y dos ejemplos de calibración de la cámara. En (Kato et al., 2000a, 27).



Figura 128. Patrón de *calib_cparam* y dos ejemplos de calibración de la cámara. En (Kato et al., 2000a, 28).

5.1.1.3. DEFINICIÓN DE LAS MARCAS Y REGISTRO

Las marcas de ARToolKit constan de un borde negro de un ancho determinado. En su interior, un dibujo negro o patrón hace que se diferencien unas de otras. En la Figura 129 se muestran los cuatro ejemplos de marcas distribuidas con el software por tratarse de unos clásicos (son las mismas marcas distribuidas desde las primeras versiones de

ARToolKit). Cabe remarcar que la simetría respecto al eje vertical no es un problema en las marcas b y c, puesto que el software no considera la posibilidad de que éstas sean vistas de espaldas, es decir, la inclinación va de -90° a $+90^\circ$, siendo 0° la posición horizontal, tal y como se muestran.

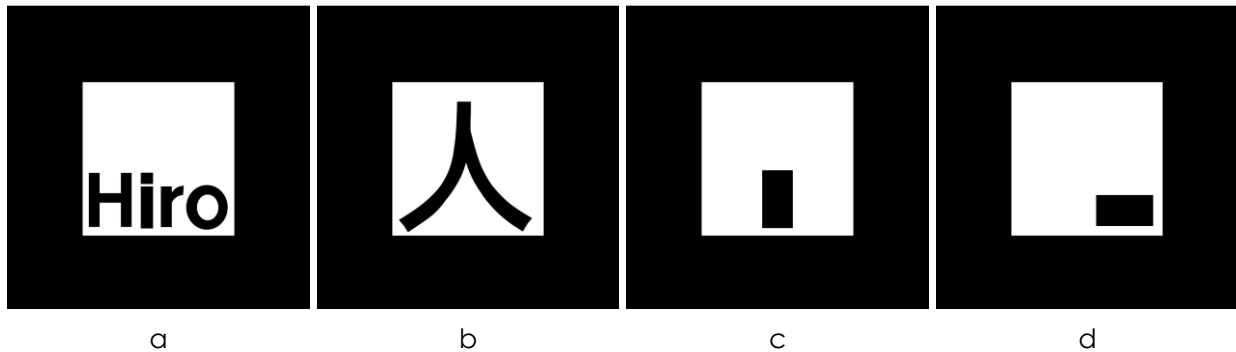


Figura 129. Marcas de ARToolKit: a) PattHiro; b) PattKanji; c) PattSample1; d) PattSample2. En (Human Interface Technology Laboratory, 2007).

El registro se basa en un procedimiento básico de detección de esquinas con un algoritmo de estimación de los parámetros de orientación externa de la cámara o matriz de transformación. En primer lugar la imagen original se transforma en una imagen binaria (blanca y negra) basada en un valor umbral. A continuación se buscan las regiones cuadradas. Para cada cuadrado, se registra el patrón definido en su interior, y se compara con los patrones que se han definido previamente y que el programa tiene guardados en memoria. Si se encuentra una correspondencia (*match*), significa que el programa ha encontrado una de las marcas. Entonces el programa utiliza el cuadrado, cuyas dimensiones (tamaño de los lados) y orientación (a partir de los patrones) son conocidas, para calcular la matriz de transformación de la cámara con respecto a la marca. Esta matriz tiene dimensiones de 3×4 , cuyos parámetros son los nueve elementos que definen la matriz de rotación y las tres traslaciones respecto a los ejes de coordenadas. Esta matriz se utiliza para definir la cámara virtual en el sistema de coordenadas. De este modo, ambos sistemas de coordenadas (real y virtual) se hacen coincidir para que los objetos virtuales se dibujen en su posición correcta respecto a los objetos del entorno real (Kato et al., 2000a). En la siguiente figura, se resume el proceso:

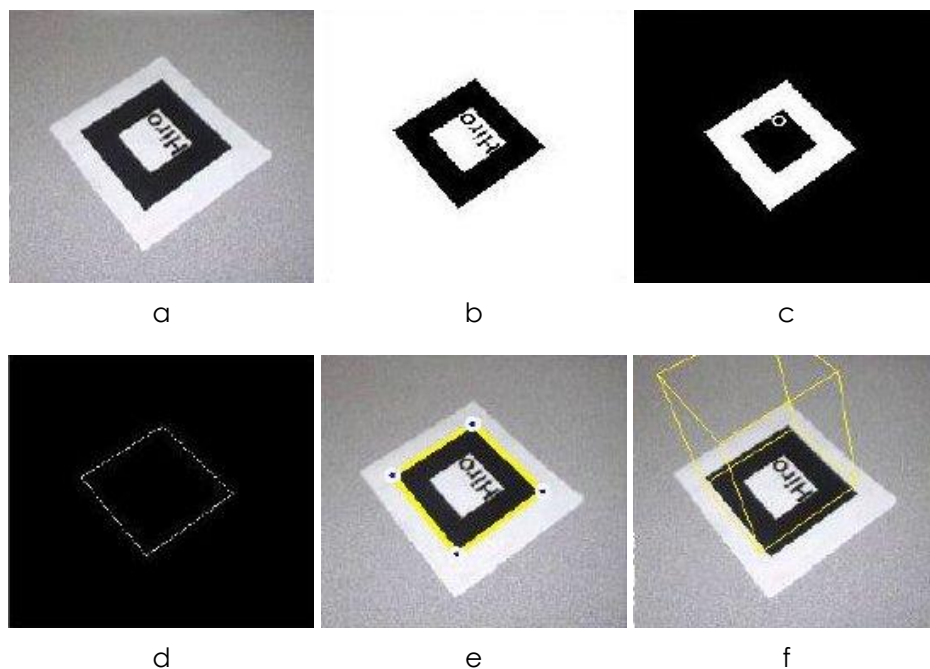


Figura 130. Proceso de registro: a) Imagen original; b) Imagen umbralizada; c) Componentes conectados; d) Contorno; e) Extracción de lados y esquinas; f) Cubo ajustado a la marca. En (Human Interface Technology Laboratory, 2007).

En (Dorfmueller-Ulhaas, 2003) se explican con más detalle los procesos antes descritos.

5.1.1.4. LIMITACIONES

Hay una serie de limitaciones propias de los sistemas de RA basados en visión por computador:

- Los objetos virtuales solamente aparecen cuando las marcas físicas son **visibles** a la cámara. Este hecho limita el movimiento del usuario en el entorno y, además, implica que las marcas no deben ocultarse total ni parcialmente (por ejemplo, por las manos del usuario o por otros objetos físicos).
- Existen también unas **distancias** críticas que deben ser consideradas para que la cámara reconozca correctamente las marcas. A mayor tamaño de las marcas físicas, más se podrá alejar la cámara de ésta.
- Esta distancia también se ve afectada por la complejidad de los **patrones** dentro de la marca. A mayor simplicidad de los patrones, el reconocimiento será mejor. Los patrones con grandes regiones negras y blancas (patrones de baja frecuencia) son los más efectivos.

- El registro también se ve condicionado por la **orientación** relativa entre la marca y la cámara. A mayor grado de inclinación, menos fiable será el reconocimiento.
- Por último, el registro también depende de las condiciones de **iluminación** ambiente. Las luces directas pueden producir reflejos y puntos brillantes sobre las marcas. Para evitar este problema se pueden utilizar superficies mate.

En las aplicaciones desarrolladas en esta tesis (capítulo 7), la limitación más importante se ha encontrado en las condiciones de iluminación debido a la utilización de retroproyecciones, ya que ha sido especialmente dificultosa la tarea de establecer entornos de baja iluminación pero suficientes para que se vean las marcas, evitando en todo momento la luz directa para que no se produjesen reflexiones sobre éstas.

5.1.2. MXRTOOLKIT

Los principios básicos y el funcionamiento de MXRToolKit se asemejan mucho los de ARToolKit, aunque existen algunas diferencias. Una de las ventajas que presenta MXRToolKit es la manera en que está estructurado el software, con una interfaz más sencilla orientada al usuario, ya que permite que el aprendizaje sea mucho más rápido y que, incluso los usuarios no expertos en programación, puedan crear sus propias aplicaciones de RA. Por ejemplo, la herramienta *Object Aligment Tool* (Mixed Reality Lab, 2003b) permite situar visualmente y de forma sencilla los objetos virtuales con respecto a las marcas, mientras que la herramienta *Camera Calibration Tool* (Mixed Reality Lab, 2003a) está dotada de una interfaz amigable que permite realizar la calibración de la cámara. Otro ejemplo atractivo de este software son los distintos tipos de objetos generados por ordenador que se pueden relacionar directamente con las marcas definidos en el objeto *mrxMedia*, como por ejemplo vídeo (AVI, MPG), audio (WAV, MIDI) o imágenes (JPG, BMP, TGA), a parte de modelos 3D (VRML, LWO, MD2, OBJ, ASE, MS3D, 3DS). Además, existen tres manuales (Mixed Reality Lab, 2003b; Mixed Reality Lab, 2003a; Mixed Reality Lab, 2003c) que explican detalladamente los pasos necesarios para la instalación, funcionamiento y desarrollo de aplicaciones propias de RA. Sin embargo la última versión de este software data del año 2003, y actualmente el software no se sigue desarrollando.

El uso de estas librerías no es tan extendido como en el caso de ARToolKit, aunque podemos encontrar varios autores que las utilizan como, por ejemplo, en la pieza *Red*

Libre Red Visible, anteriormente explicada en el apartado 1.4.1. En esta tesis se han utilizado en el ensayo descrito en el capítulo 10.

5.1.2.1. PRINCIPIOS BÁSICOS

La Figura 131 muestra el esquema empleado en el manual (Mixed Reality Lab, 2003c) para explicar los principios básicos del software mediante distintos objetos utilizados en la programación. Como se puede apreciar, éstos son similares a los de ARToolKit (recordar Figura 126): la imagen se captura (*mxrCaptureStream*) y se visualiza sobre el fondo (*mxrFrustum*); al mismo tiempo, la imagen (*mxrImage*) se analiza, ya que el ordenador busca las marcas bidimensionales (*mxrFrame*) mostradas en la escena; finalmente, visualiza los objetos virtuales (*mxrMedia*) sobre la imagen vídeo, de acuerdo con los parámetros de la matriz de transformación (*mxrTransform*) que contiene la orientación y posición de la marca con respecto a la cámara.

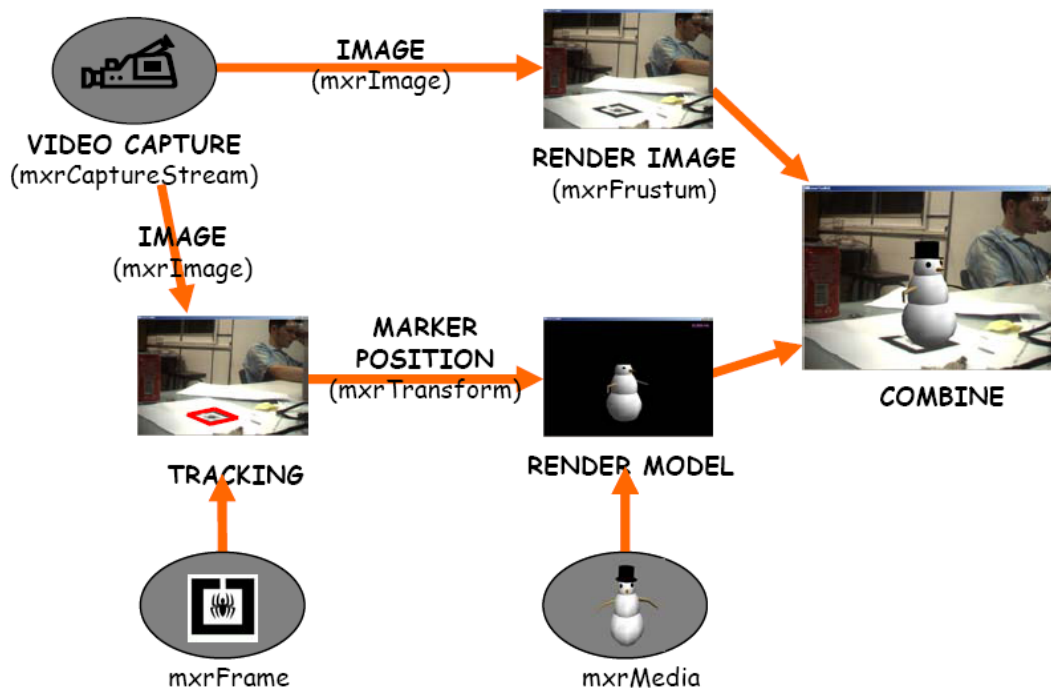


Figura 131. Secuencia de los procesos de registro de una marca y superposición de elementos virtuales en MXRToolKit. En (Mixed Reality Lab, 2003c).

El software MXRToolKit depende de las librerías DirectX 9.0 SDK y del controlador PGRFlyCapture de Point Grey Research Inc para cámaras Firewire y Dragonfly. Hay un total de cinco proyectos en área de trabajo (Mixed Reality Lab, 2003c):

- **mxrSDKDLL**, **mxrSDKMediaDLL** y **simpleTrack**: Los dos primeros se utilizan para construir los ficheros *dll* de MXRToolKit; el tercero es un pequeño ejemplo que utiliza estas dos *dlls*.
- **mxrCalibration**: Contiene el código fuente de la herramienta *Camera Calibration Tool*.
- **winOGL**: Contiene el código fuente de la herramienta *Object Alignment Tool*.

5.1.2.2. CALIBRACIÓN DE LA CÁMARA

Como se ha explicado anteriormente, la calibración de la cámara se realiza mediante la herramienta *Camera Calibration Tool*. Al igual que en el caso de ARToolKit, la calibración se realiza a partir de un patrón, que en este caso consta de un dibujo de rejillas cuadradas blancas y negras, a modo de tablero de ajedrez (Figura 132).

En (Mixed Reality Lab, 2003a) se explican los pasos a seguir para lograr la calibración de una cámara de manera sencilla, los cuales se resumen a continuación:

1. Imprimir el patrón de calibración (facilitado en formato *.pdf*) y pegarlo sobre una superficie rígida.
2. Abrir la herramienta de calibración de la cámara.
3. Aceptar los parámetros por defecto que aparecen en el cuadro de diálogo.
4. Seleccionar el origen de la cámara. En este momento, se visualiza la entrada de vídeo de la cámara en un recuadro.
5. Mover el patrón hacia atrás y hacia adelante, variando el ángulo y posición con respecto a la cámara. Si este proceso es satisfactorio, aparecen tres barras de colores (amarilla, verde y roja) moviéndose de izquierda a derecha en la ventana principal, indicando la calidad del conjunto de datos capturados. Cuando las tres barras estén a más de 2/3 del ancho de la ventana, presionar el icono *calibrate* que se halla en la barra de herramientas para abrir el menú de calibración. Presionar el botón *calibrate* y esperar. En pocos instantes, se muestran los parámetros calculados.
6. Guardar los parámetros en un fichero.

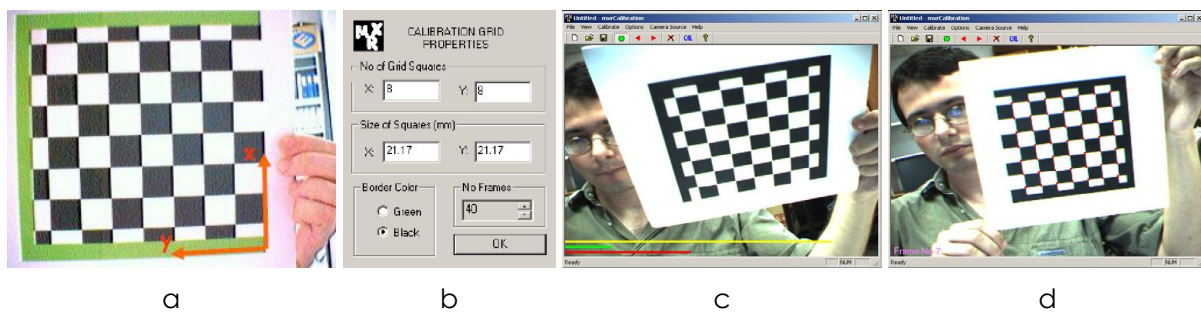


Figura 132. Proceso de calibración: a) Patrón de calibración con el borde verde en forma de U; b) Cuadro de diálogo mostrando los parámetros de calibración por defecto; c) Proceso de captura de la calibración, mostrando tres barras coloreadas que indican el estado del proceso; d) Proceso de revisión, en el que el usuario puede visualizar la bondad del proceso. En (Mixed Reality Lab, 2003a).

En este proceso se calculan un total de 9 parámetros, que se describen brevemente a continuación:

- Parámetros lineales:
 - f_x : Distancia focal en la dirección x expresada en píxeles;
 - f_y : Distancia focal en la dirección y expresada en píxeles;
 - ox : Coordenada x del centro de la imagen, tomando como origen la esquina superior izquierda;
 - oy : Coordenada y del centro de la imagen, tomando como origen la esquina superior izquierda;
 - $skew$: Describe la oblicuidad de la matriz de CCD con respecto al eje de la cámara.

- Parámetros no lineales o componentes de la distorsión radial:
 - r_2 : Distorsión radial de segundo orden;
 - r_4 : Distorsión radial de cuarto orden;
 - xy_1 : Distorsión tangencial primera;
 - xy_2 : Distorsión tangencial segunda.

5.1.2.3. DEFINICIÓN DE LAS MARCAS Y REGISTRO

Las marcas de MXRToolKit constan de un borde negro de un ancho determinado, con una pequeña abertura blanca en el lado superior y un dibujo en su interior. A diferencia

de ARToolKit, en estas marcas el dibujo o patrón no puede tocar el borde negro. Además, se pueden registrar de forma individual o colectiva; el segundo caso, implica que un conjunto de marcas están conectadas entre sí, formando en realidad una sola, por lo que la matriz de transformación se calcula solamente para dicho conjunto, no individualmente. Para ello, MXRToolKit define el objeto *mxrFrame*, que puede contener una marca individual o un conjunto, denominándose en ambos casos como *frames*; de este modo la matriz de transformación se calcula sólo para los distintos *frames* establecidos. En la siguiente figura se muestran dos ejemplos.

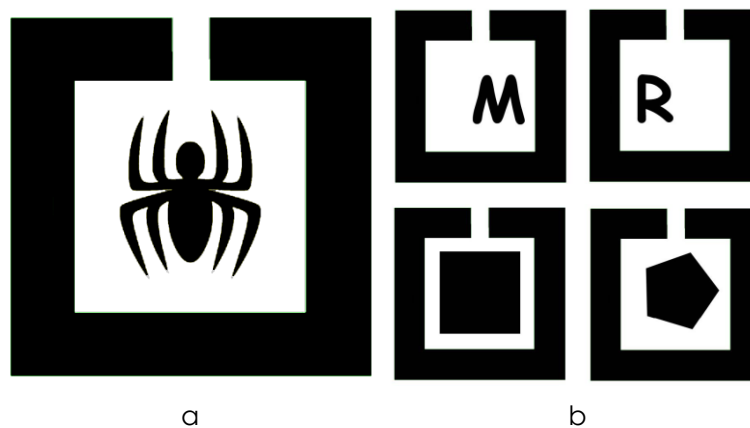


Figura 133. Ejemplos de marcas: a) *Frame* que consta de una sola marca; b) *Frame* que consta de un conjunto de cuatro marcas.

En la documentación no se encuentra de forma detallada la explicación del algoritmo empleado para el reconocimiento de las marcas, pero su funcionamiento es similar al de ARToolKit. La principal diferencia reside en el hecho de que las marcas de MXRToolKit tienen una pequeña abertura en uno de los lados del borde negro que las define, y por ello en lugar de 4 puntos, el algoritmo utiliza un total de 8 para el cálculo de la matriz de transformación (Díaz, 2007). Dicha matriz se guarda en el objeto *mxrTransform* que consta de un total de 12 parámetros: 9 que definen la matriz de rotación (r_{11} , r_{12} , r_{13} , r_{21} , r_{22} , r_{23} , r_{31} , r_{32} , r_{33}), y tres que definen la translación (t_x , t_y , t_z), en ambos casos respecto a los tres ejes de coordenadas.

5.1.2.4. LIMITACIONES

Las mismas explicadas para ARToolKit (apartado 5.1.1.4).

5.1.3. STUDIERSTUBE

El desarrollo de la plataforma Studierstube empezó en la Universidad Tecnológica de Viena (TU-Wien) en el año 1996. Sin embargo, desde el 2004 las principales tareas de investigación y desarrollo se llevan a cabo en la Universidad Tecnológica de Graz. El grupo de investigación de Graz desarrolló el software ARToolKitPlus (Wagner, 2007a), basado en ARToolKit. Tras varios años de desarrollo obtuvieron un software que no compartía ninguna de las librerías de ARToolKit, y lo denominaron Studierstube, vocablo alemán que significa "aula de estudio". Consta de dos versiones, una es Studierstube Tracker (Wagner, 2007c) y la segunda Studierstube ES (Wagner, 2007b). La primera comprende las librerías para la detección y estimación de la orientación externa a partir de las marcas. La segunda es la adaptación de las librerías para que funcionen en pequeños dispositivos móviles, tales como PDAs o teléfonos móviles.

Una de las aplicaciones más conocidas desarrollada con estas librerías es *Invisible Train* (introducida en el apartado 1.4.2), aunque también se han desarrollado otros trabajos, como *Cows vs. Aliens* (Figura 134) descrito en (Mulloni, 2007).



Figura 134. Ejemplo de aplicación en Studierstube: Cows vs. Aliens. En (Mulloni, 2007).

5.1.4. ARTAG

ARTag (Fiala, 2004, 11) es un sistema para la realización de aplicaciones de RA basado en marcas de patrones planos. El sistema tiene un total de 2002 marcas que consisten en un borde exterior, al igual que ARToolKit, pero con la diferencia de que la detección del borde se basa en los ejes, y el procesado del patrón interno se hace con un procedimiento digital en lugar de mediante correlaciones. El borde cuadrado en ambos sistemas (ARTag y ARToolKit) permite el cálculo de los seis parámetros de orientación externa de la marca con respecto a la cámara a partir del registro de los cuatro puntos. ARTag localiza las marcas potenciales en una imagen mediante la detección de contornos con cuatro esquinas, utiliza los puntos de las esquinas para

definir la homografía, crea una rejilla de muestra y analiza su interior para determinar si se trata o no de un patrón. El interior se muestrea para determinar un '0' o un '1' digital para cada elemento de la rejilla, y el procesado se realiza en el dominio digital. Las marcas de ARTag se componen de un total de 10x10 unidades, con un grosor de borde de 2 unidades, dejando 36 celdas en su interior para contener la información del patrón. Cada celda es únicamente blanca o negra y contiene un bit de datos digitales. Los bordes pueden ser tanto negros como blancos (Figura 135).



Figura 135. Marcas de ARTag y ejemplo de registro con oclusión por un dedo.

Una de las ventajas de ARTag respecto a ARToolKit es la menor dependencia de las condiciones de luminosidad. Además, las marcas pueden ocultarse parcialmente y tener un tamaño más reducido.

Aunque estas librerías no se han utilizado en ninguno de los ensayos de esta tesis (principalmente por no disponer de licencia GPL), se ha probado su uso derivándose algunas características comparativas con otras librerías, que se describen en el apartado de conclusiones de esta memoria.

5.1.5. OSGAR

OSGAR (Augmented Environments Laboratory, 2007) extiende las librerías OpenSceneGraph para que puedan ser utilizadas en aplicaciones de RA, soportando transformaciones que tienen asociadas ciertas incertidumbres o errores. OSGAR proporciona a los programadores un amplio rango de facilidades para consultar y actuar dependiendo del error estimado en el proceso de registro. A continuación se enumeran una serie de capacidades que se han añadido a las librerías de OpenSceneGraph:

- Visualizar una entrada de vídeo en el fondo (background).
- Detectar marcas artificiales.

- Conectar diferentes sensores mediante el cliente VRPN (*Virtual Reality Peripheral Network*).
- Soporte de múltiples sensores.
- Estimación en tiempo real del error en el registro.

5.1.6. OSGART

OSGART (Grasset et al., 2005; HITLab NZ, 2006) son unas librerías escritas en C++ para la realización de aplicaciones de RA, combinando las librerías de registro de ARToolKit con las librerías de renderizado OpenSceneGraph, y permitiendo una gran variedad de integración de vídeo. Al cambiar de GLUT¹³ (utilizado en ARToolKit) a OpenSceneGraph, se incrementa notablemente la calidad de los gráficos generados por ordenador, siendo visualmente equiparables a los que se utilizan en los videojuegos estándar. La versión básica de estas librerías también es libre (licencia GPL), pudiéndose utilizar con fines académicos y de investigación. Los usuarios pueden desarrollar nuevas aplicaciones/tareas en C++, pero también los usuarios no-programadores tienen acceso a través de las interfaces de Python, Lua o Ruby (mediante los módulos de *osgBindings* u *osgIntrospection*). Algunos aspectos destacados que definen estas librerías son:

- Diseño orientado a objetos.
- Soporte de múltiples entradas de vídeo (en tiempo real, mediante ficheros o a través de la red).
- Integración de objetos de vídeo de alto nivel (fondos de vídeo, planos, carteleras, etc.)
- Concepto de sombreado de vídeos mediante GLSL¹⁴.
- Concepto de marcas genéricas.
- Soporte de múltiples formas de marcas y de tecnologías de reconocimiento de rasgos naturales (ARToolKit, ARToolKit++, ARTag, BazAR, etc.).
- Soporte de técnicas de renderizado de sombras (actualmente profundidad de sombras).

¹³ GLUT: *The OpenGL Utility Toolkit*.

¹⁴ GLSL: *OpenGL Shading Language*. Una tecnología parte del API estándar de OpenGL, que permite especificar segmentos de programas gráficos que serán ejecutados sobre el GPU (Wikipedia, 2007b).

- Interfaz de programación en C++, pero también en Python, Lua, Ruby, C# y Java (a través de osgBindings u osgIntrospection).
- Soporte para todas las posibilidades de OpenSceneGraph (renderizado 3D de alta calidad, estadísticas, exportar desde 3DS Max/MAYA, múltiples plataformas, etc.).

5.1.7. AMIRE

AMIRE (AMIRE, 2004a) es un software de libre distribución (SourceForge, 2007) para Windows 2000/XP, basado en las librerías ARToolKit, que proporciona una sencilla GUI para la elaboración de aplicaciones de RA sin la necesidad de programar (Figura 136). Este programa consta de una serie de subsistemas que se resumen a continuación:

- Marco de trabajo (*framework*): Basado en el entorno de programación C++ para la generación de aplicaciones de RA mediante la conexión de diversos componentes.
- Componentes: Bloques de construcción para las funciones necesarias en aplicaciones de RA, que se pueden conectar unos con otros.
- Autoría: Interfaz visual de usuario que permite seleccionar y conectar los diversos componentes para construir una escena de RA, sin la necesidad de programación ni compilación de código.
- Formato XML de la escena: Las escenas en AMIRE se guardan en formato XML. Tiene la ventaja que este tipo de formato se puede escribir mediante otros programas o editores de texto.
- Ejecución (*runtime*): Programa de inicio para ejecutar aplicaciones de RA una vez guardadas en formato XML, sin la necesidad de abrir el programa AMIRE.
- GEMs: Base de datos del software y técnicas de programación que permite una reutilización del código.

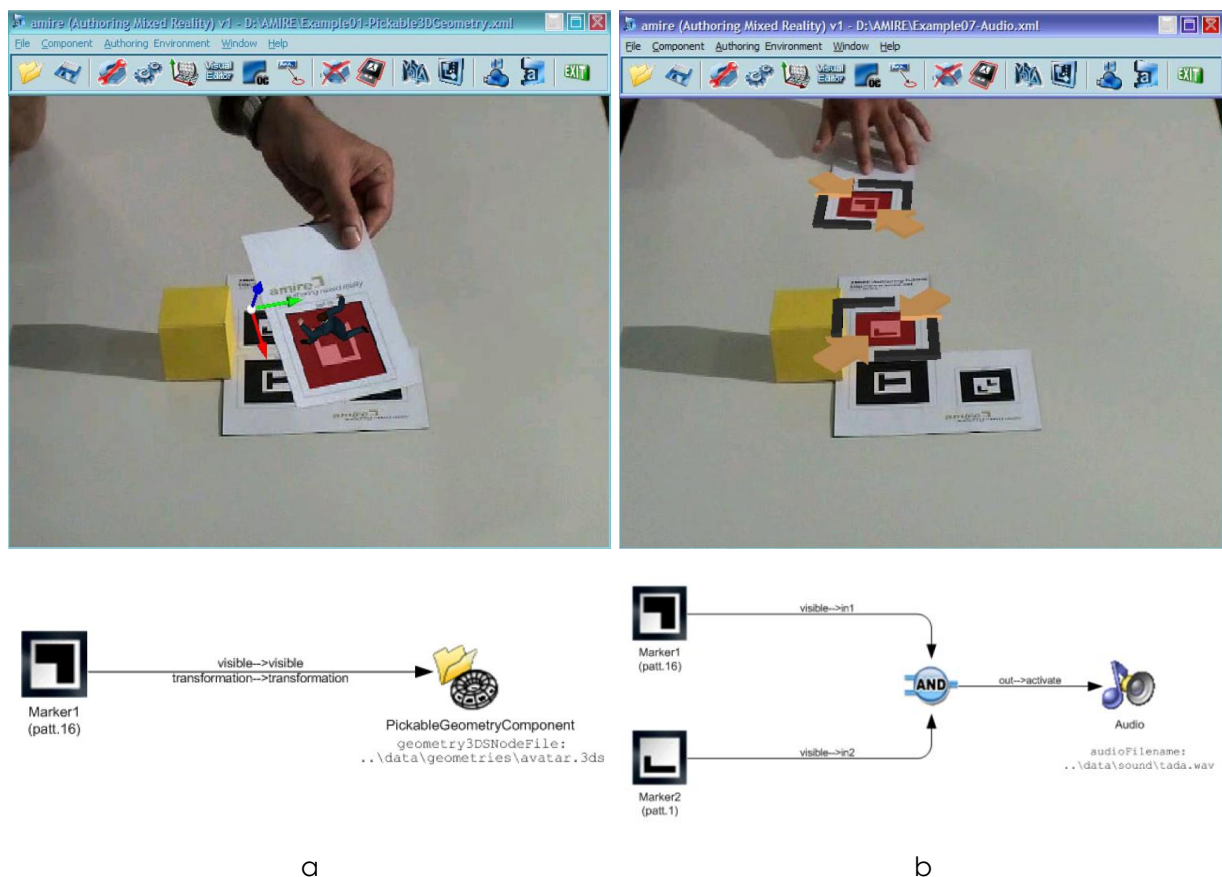


Figura 136. Ejemplos de escenas de RA disponibles en AMIRE y componentes conectados para su elaboración: a) Un objeto 3D se asigna a una marca (Marker1); b) Se ejecuta un sonido cuando dos marcas (Marker1 y Marker2) son visibles. En (AMIRE, 2004b).

5.2. SW DE RECONOCIMIENTO DE RASGOS NATURALES

5.2.1. BAZAR

BazAR son unas librerías de visión por computador basadas en la detección y correspondencia de rasgos naturales puntuales. En particular, es capaz de detectar y registrar rápidamente objetos conocidos en imágenes planas. También contiene un sistema de calibración geométrica y fotométrica de la cámara bastante potente. Este software se ha desarrollado y se distribuye por el laboratorio CVLab de la EPFL (Ecole Polytechnique Fédérale Lausanne) (CVLAB - Computer Vision Laboratory, 2007). En esta tesis, estas librerías se han utilizado en el ensayo descrito en el capítulo 11.

5.2.1.1. PRINCIPIOS BÁSICOS

Las librerías de BazAR dependen de OpenCV y de OpenGL (GLUT y GLEW¹⁵) para su funcionamiento. BazAR está organizado en dos módulos:

1. **Garfeild:** Este módulo depende del segundo (*starter*) y proporciona las herramientas para detectar puntos clave o rasgos puntuales, y establecer su correspondencia respecto de una serie de puntos almacenados (puntos modelo). Garfeild está diseñado para detectar rápidamente objetos planos en una imagen mediante dos fases de proceso: 1) Entrenamiento, o fase off-line, 2) Reconocimiento y correspondencias, o fase on-line. Durante la fase de entrenamiento se detectan los rasgos puntuales de una imagen modelo. Se extraen sus elementos (*patches*), se corrigen sus orientaciones y se aprende su apariencia. Durante la segunda fase, un algoritmo de reconocimiento de objetos planos detecta los rasgos puntuales de la imagen de entrada para establecer las correspondencias con los puntos modelo. A continuación, se activa un procedimiento para discernir correspondencias buenas y malas, y encontrar la homografía que relaciona los puntos modelo con los de entrada.
2. **Starter:** Contiene las estructuras básicas y las herramientas matemáticas.

5.2.1.2. CALIBRACIÓN DE LA CÁMARA

En (Pilet et al., 2006) se explica el proceso de calibración de BazAR, un sistema sencillo con el que se pueden calibrar varias cámaras simultáneamente. En el proceso se calcula la geometría de las cámaras, su respuesta fotométrica y un mapa de iluminación ambiental. La única intervención manual requerida es el movimiento arbitrario de la imagen plana dentro del campo de visión de la cámara. En los siguientes puntos (y figuras) se explican los pasos para la calibración y puesta en marcha de una aplicación:

- a. **Adquisición del modelo:** En primer lugar hay que presentar la imagen plana a la cámara, a continuación presionar la barra espaciadora para crear una imagen fija, y finalmente arrastrar un marco verde que aparece sobre la imagen para seleccionar el modelo (imagen de referencia o patrón de calibración) deseado dentro de la imagen.

¹⁵ GLEW: *The OpenGL Extension Wrangler Library*.

- b. **Adquisición de datos** para la calibración de la cámara: Hay que mover el patrón de calibración dentro del campo de visión de la cámara. Unos puntos verdes muestran los rasgos puntuales correctamente detectados.
- c. Cómputo de la **calibración geométrica**: El sistema calcula de forma automática la calibración de la cámara o cámaras. En primer lugar el sistema calcula las homografías entre el plano geométrico de la imagen capturada y la imagen modelo. A continuación se retienen las homografías más fiables y sus *frames* correspondientes para calcular los parámetros intrínsecos de la cámara y la matriz de transformación respecto al objeto plano en cada uno de los *frames* correspondientes. Finalmente se calcula la matriz de transformación entre las diversas cámaras. Durante el proceso aparece un histograma que muestra los errores de la re-proyección en unidades de píxel. La parte negra muestra el límite de un píxel. Este paso es completamente automático.
- d. **Calibración fotométrica**: Después de la calibración geométrica, es posible aumentar los objetos mediante elementos virtuales tridimensionales. Sin embargo todavía no existe ningún mapa de iluminación para proporcionar una visualización más idónea de los objetos virtuales. Es por ello que al inicio de la fase de aumentación el sistema recoge datos de iluminación para calcular el mapa de irradiación. Para la calibración fotométrica se utiliza la misma imagen modelo que en la calibración geométrica. Para cada una de las cámaras, siempre y cuando la iluminación y los parámetros (*settings*) de las cámaras no cambien, las intensidades de los píxeles dentro del patrón de calibración dependen únicamente de su normal. Es decir, cada imagen de entrada en la que se detecta el patrón proporciona un número de ejemplos correspondientes a la orientación de una superficie individual. Debido a que se registran automáticamente muchos ejemplos, se puede simultáneamente recuperar la ganancia (*gain*) y el sesgo/desviación (*bias*) de cada una de las cámaras y crear un modelo de iluminación ambiental que se puede utilizar con fines de re-iluminación.
- e. **Aumentación**: Finalmente todo está a punto para mostrar elementos virtuales (con iluminación) superpuestos a la imagen de entrada. El mapa de irradiación calculado previamente aparece en la parte superior derecha de la imagen vista por el usuario.

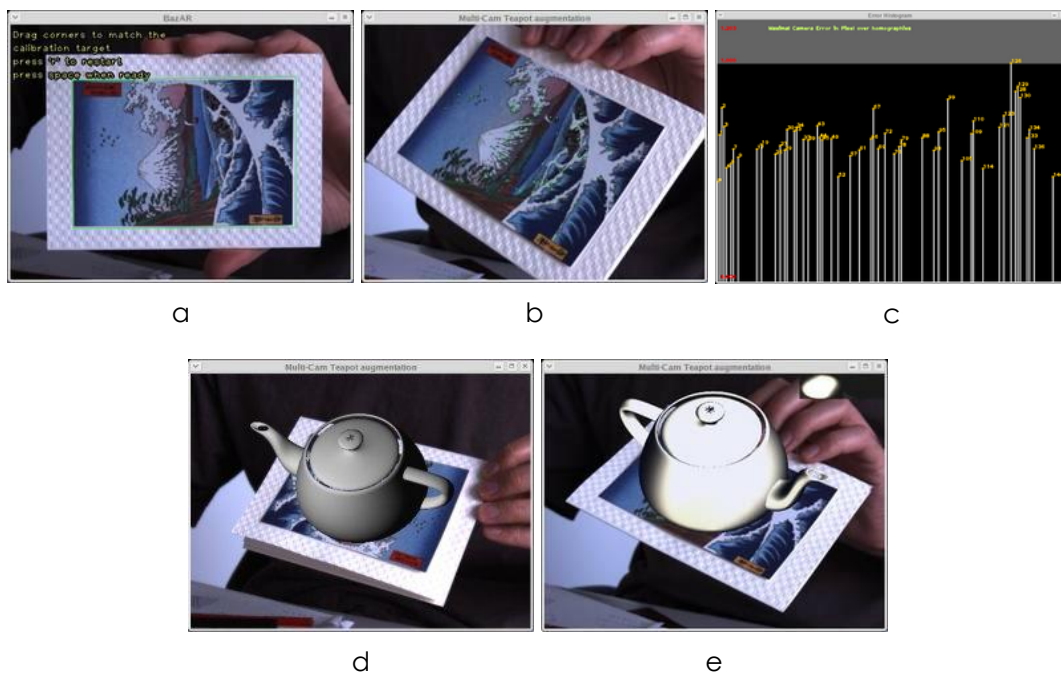


Figura 137. Proceso de calibración geométrica, superposición de una figura 3D y calibración fotométrica de Bazar.

5.2.1.3. PROCESO DE REGISTRO

Según sus autores, en Bazar se incluye un procedimiento novedoso para la correspondencia de rasgos puntuales bajo numerosos puntos de vista y cambios de iluminación, que es adecuado para la estimación precisa de la matriz de transformación con un coste computacional menor que en otros métodos actuales. Este método también se aplica para el registro sobre superficies deformables (aunque este desarrollo no se incluye en la versión gratuita de las librerías de Bazar).

En (Lepetit et al., 2004b) se explica el procedimiento, que se basa en tratar las correspondencias de rasgos puntuales como un problema de clasificación, en el que cada clase se corresponde con un conjunto de todas las posibles vistas de dicho rasgo puntual. Durante la fase de entrenamiento, proporcionando como mínimo una imagen del objeto en cuestión (o imagen patrón), se sintetizan un gran número de vistas de puntos clave individuales. Si se asume que el objeto es localmente plano, bastará con envolver los elementos (*patches*) de la imagen alrededor de los puntos bajo deformaciones afines u homografías. Por otra parte, dado un modelo 3D, se utilizan técnicas estándar de mapeado de texturas. Este segundo procedimiento es más complejo pero relaja los supuestos de superficie plana. En tiempo de ejecución, se utilizan técnicas potentes y rápidas de clasificación para decidir a qué conjunto, si lo

hay, pertenece el elemento observado. Una vez que las correspondencias potenciales se han establecido entre los puntos de interés de la imagen de entrada y los del objeto (imagen modelo), se aplica un método basado en el procedimiento estándar denominado RANSAC para estimar la matriz de transformación.

5.2.1.4. LIMITACIONES

En este software se reducen algunas de las limitaciones encontradas en sistemas anteriormente descritos, por ejemplo:

- La imagen se puede ocultar parcialmente sin impedir el registro, por lo que el usuario puede seguir viendo los objetos virtuales.
- Trabaja mejor en condiciones de luz menos favorables.
- No existe la necesidad de ubicar elementos artificiales (marcas) en el entorno físico.

Sin embargo existen una serie de limitaciones:

- El software que se distribuye sólo contempla la posibilidad del reconocimiento de una sola imagen, en contraposición a los sistemas anteriormente vistos.
- La complejidad del algoritmo de registro hace que el software requiera de ordenadores más potentes.

5.3. SW MULTITAREA

Existe una gran diversidad de software multitarea con los que se pueden programar aplicaciones de RA. Entre ellos cabe destacar *Processing* (Processing, 2008), *vvvv* (vvvv group, 2008), *Pure Data* (PD, 2008) y *Max/MSP Jitter* (Cycling '74, 2008). Los tres primeros son de libre distribución, mientras que el último es software multitarea. Aunque es recomendable la utilización de software de libre distribución, sobre todo en el ámbito de la investigación, el programa *Max/MSP Jitter* es, a día de hoy, el más completo en cuanto a integración, gestión y procesado de datos multimedia, permitiendo al mismo tiempo múltiples posibilidades de interacción. Es por ello que se ha preferido su uso en algunos de los ensayos de esta tesis (los descritos en los capítulos 8, 9 y 12).

5.3.1. MAX/MSP JITTER

Jitter es un conjunto de más de 130 objetos de vídeo, matrices y gráficos 3D para el programa Max/MSP. Estos objetos extienden la funcionalidad del programa para generar, modificar y manipular datos matriciales – cualquier dato que se pueda expresar en forma de filas y columnas, como vídeos o imágenes estáticas, geometría 3D, texto, hojas de cálculo, sistemas de partículas, vóxeles¹⁶, o audio. Max/MSP es un entorno de programación desarrollado en un principio para su uso en aplicaciones interactivas de música, pero ha ido ampliándose con nuevos objetos que hacen de él una herramienta multimedia muy estable.

La arquitectura de Jitter está optimizada para realizar tareas con video (pregrabado o en tiempo real), existiendo para ello una serie de operadores matemáticos, análisis, tratamiento del color, procesamiento del canal alfa, deformaciones espaciales, filtros de convolución y otros efectos. Además, la compatibilidad de Jitter con la arquitectura de QuickTime (QT) de Apple, le permite soportar todos los formatos admitidos por QT, así como la creación de ficheros, operaciones de edición, importar/exportar, efectos de QT, digitalización de vídeo, etc. Para aplicaciones en tiempo real, Jitter permite el control de cámaras de vídeo digital (DV), entrada y salida vía FireWire, y soporte para múltiples monitores.

El soporte integrado de gráficos 2D/3D proporciona herramientas para utilizar gráficos OpenGL conjuntamente con video, incluyendo la capacidad de texturizar geometría 3D con entradas de vídeo en tiempo real, convertir entradas de audio y vídeo en datos geométricos, y renderizar modelos, NURBS, texto 2D/3D, y otras geometrías.

Debido a las grandes posibilidades que ofrece este programa y a la gran cantidad de objetos que contiene, en los siguientes subapartados se introducen únicamente los utilizados en los ensayos de esta tesis.

5.3.1.1. REGISTRO DE ELEMENTOS EN LA IMAGEN

Mediante el objeto *jit.findbounds* de Jitter se realiza el registro de un color en la imagen de vídeo, devolviendo las coordenadas imagen de la esquina inferior izquierda y superior derecha de una pequeña venta que encuadra el color. Sin embargo existen otros autores que han desarrollado herramientas externas para realizar el registro de

¹⁶ Vóxel: píxel tridimensional.

color o de elementos puntuales a partir de un análisis de la imagen. Por ejemplo, el objeto *cv.jit.track* de la herramienta *cv.jit* (Pelletier, 2006) realiza un registro de la posición de un determinado número de rasgos puntuales en una imagen vídeo. Este objeto utiliza una implementación piramidal de la técnica de Lucas-Kanade. Otro ejemplo es el objeto *tap.jit.colortrack* de la herramienta *Tap Tools 1.5* de Timothy A. Place (Electrotap, 2008). Este objeto registra de forma simultánea hasta un máximo de cuatro colores diferentes, basándose en el análisis del tono, saturación y luminosidad de los píxeles de la imagen para encuadrar aquella zona en la que se encuentre dicho color, dentro de unas tolerancias especificadas por el usuario. El programa devuelve las coordenadas imagen (x,y) de la esquina inferior izquierda y superior derecha de la ventana. Ambos objetos se han utilizado en la tesis (capítulo 8). En la Figura 138 se puede ver *tap.jit.colortrack* en funcionamiento.

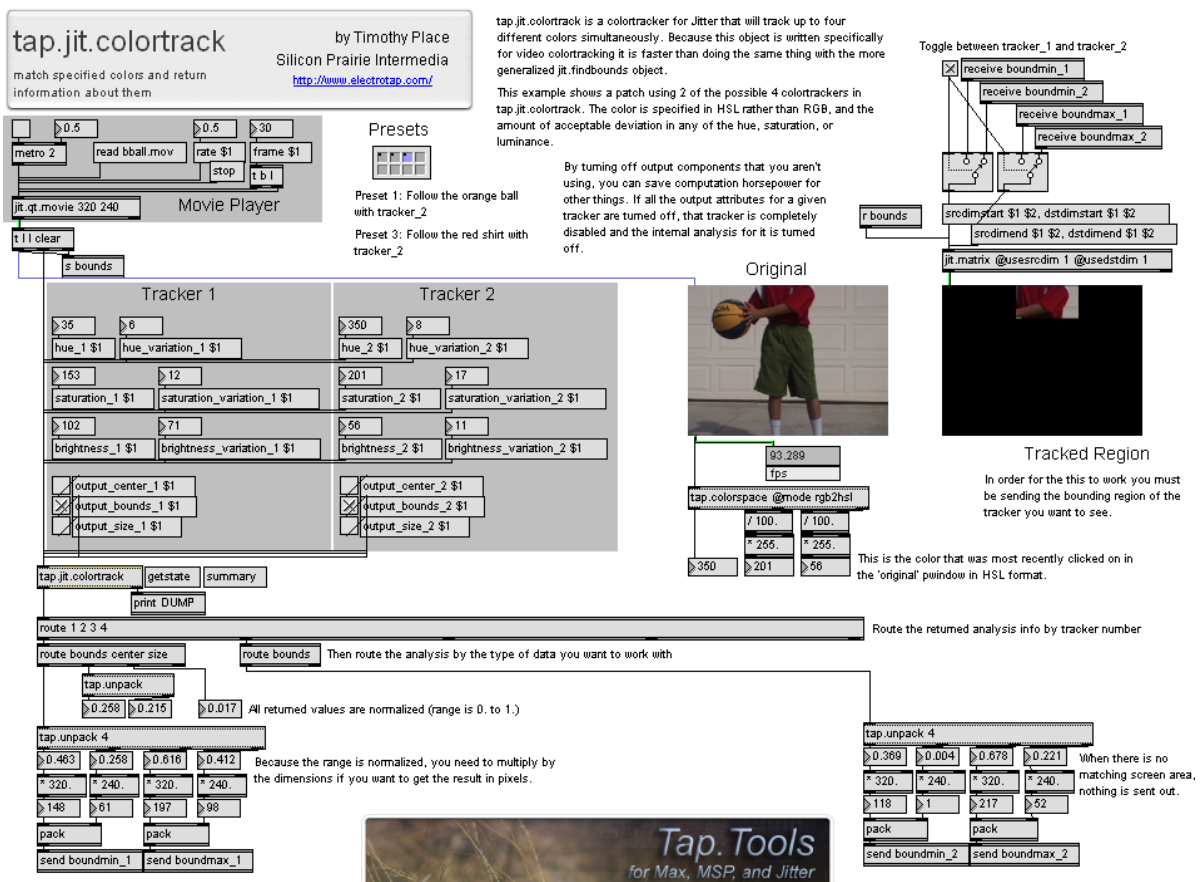


Figura 138. Ejemplo de registro de color mediante el objeto tap.jit.colortrack de Tap Tools.

5.3.1.2. OBJETOS 3D Y CÁMARAS VIRTUALES

En este apartado se explican los objetos de control de la cámara virtual, los de los gráficos 3D y su posterior renderización en una ventana de OpenGL. Los objetos de

control de la cámara virtual son *pak camera* para la posición 3D del centro de proyección y *pak lookat* para indicar la dirección del eje óptico de la cámara (hacia dónde apunta el eje óptico de la cámara). También podemos emplear *pak up* para establecer el ángulo de rotación a través de la dirección del eje óptico y *lens_angle* (a través de *jit.gl.render*) para el ángulo de apertura. En la Figura 139 se muestra un ejemplo en el que primero se mueve la cámara (de la posición 0, 0, 2 a la posición 2, 2, 6) y a continuación se cambia la dirección del eje óptico (de 0, 0, 0 a 2, 0, 0).

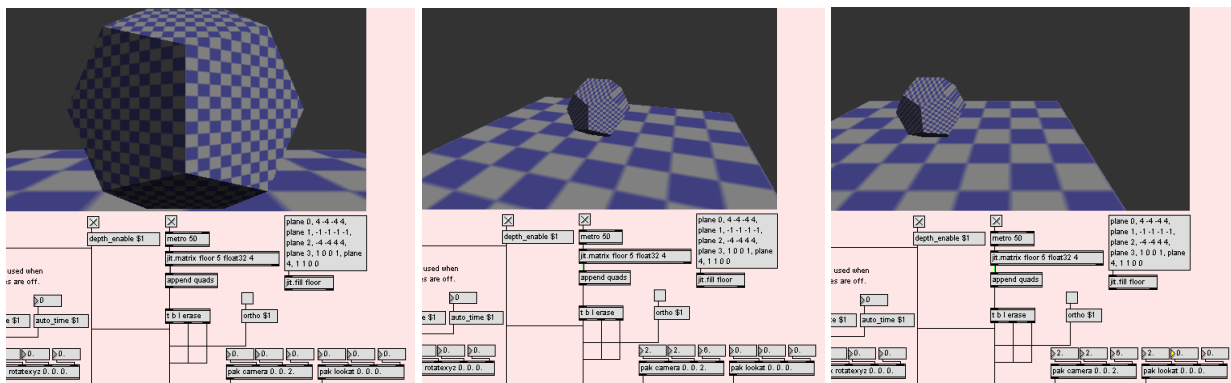


Figura 139. Ejemplo de cambio de posición y dirección del eje óptico de la cámara virtual.

Mediante el objeto *jit.gl.render* se pueden dibujar y manipular gráficos 2D/3D en una ventana *p.window* de OpenGL. El objeto *jit.gl.model* (Figura 140) lee las descripciones de objetos 3D como colecciones de polígonos. Estas descripciones están guardadas en ficheros con formato *obj*. Así pues, los modelos 3D deben tener esta extensión para ser cargados en Jitter.

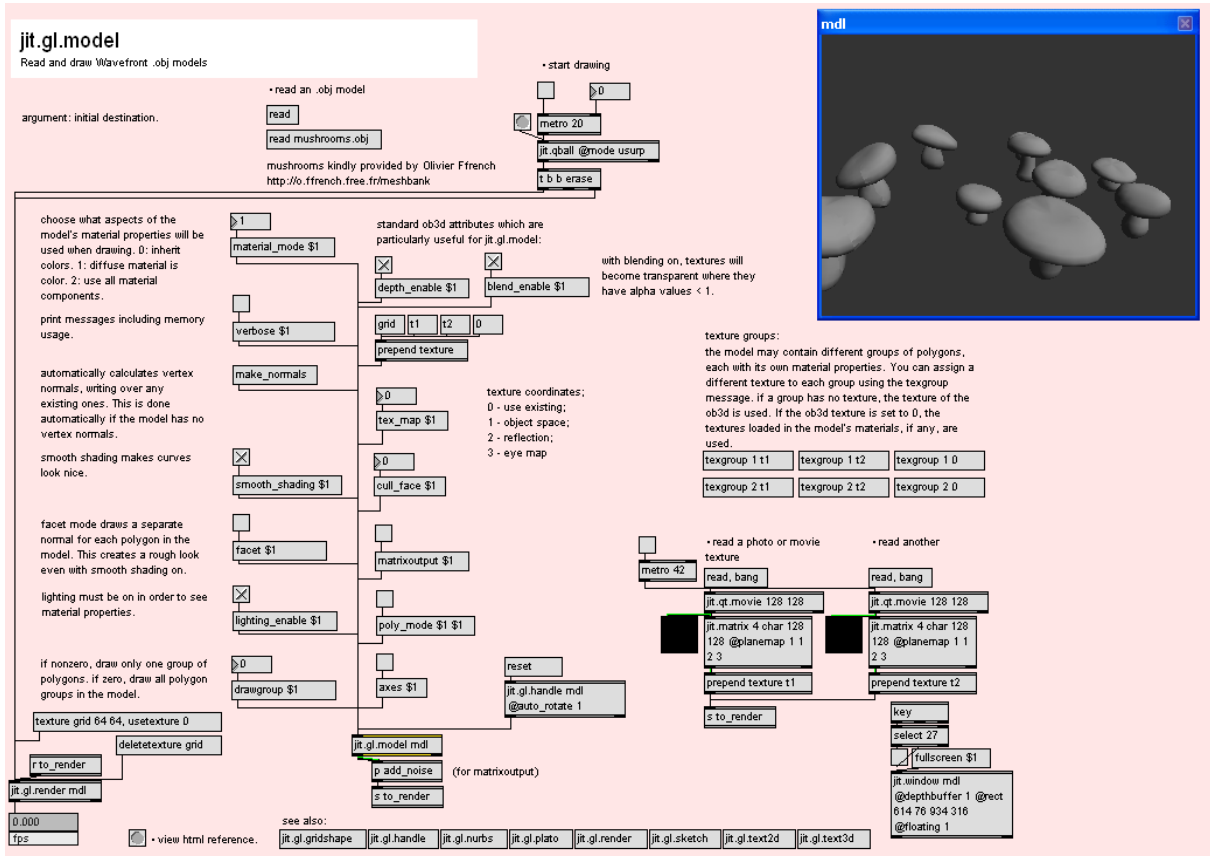


Figura 140. Objeto *jit.gl.model*.

Mediante *jit.gl.model* se pueden modificar diversos atributos de los modelos 3D: iluminación y sombreado, textura (imagen estática, vídeo pregrabado, vídeo en tiempo real, o color neutro), grupo de objetos renderizados, tipo de material, etc. En la Figura 141 podemos ver algunos ejemplos de visualizaciones para un mismo modelo.

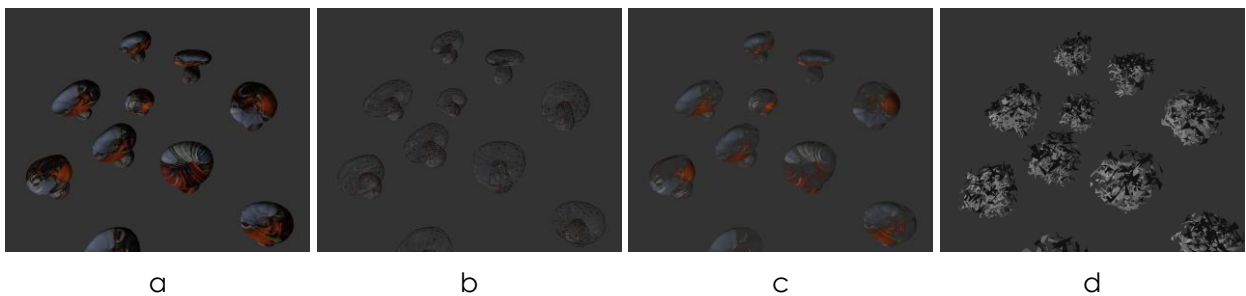


Figura 141. Modelo *obj* modificado mediante *jit.gl.model*: a) Mapeado de vídeo; b) Modo de alambres; c) Transparencia (blend activado); d) Salida matricial.

Finalmente, mediante el objeto *jit.window* establecemos las propiedades (dimensiones, modo flotante, etc.) de la ventana OpenGL, y con el objeto *jit.gl.render*, designamos los gráficos 2D/3D que se van a dibujar en dicha ventana.

5.3.1.3. INTEGRACIÓN DE SENSORES

En este apartado se explican los objetos empleados en los ensayos de esta tesis para integrar distintos sensores en Jitter: cámara Web, micrófono, sensor inercial y receptor GPS.

La entrada de la cámara web en Jitter se puede hacer mediante el objeto *jitter.qt.grab* o el objeto *jitter.dx.grab*. La ventaja del segundo en el caso de utilizar un PC, es que no requiere del programa QuickTime VDIG, ya que trabaja con DirectX. Estos objetos deben unirse con *jitter.gl.render* para que la imagen capturada se dibuje en la ventana de OpenGL.

De otro modo, la entrada de la señal del micrófono se gestiona con el objeto *adc~* (*analog to digital converter*), mediante el cual se registra, a un muestreo determinado, el voltaje recogido por el micro.

La conexión del sensor inercial MT9 de Xsens con el programa Jitter, se puede hacer mediante el objeto *mt9* programado por Laboratorio de Luz (Figura 142). Este objeto devuelve las medidas en bruto del sensor (tiempo, aceleración, giros, campo magnético y temperatura) junto con los 9 parámetros de la matriz de rotación calculada.

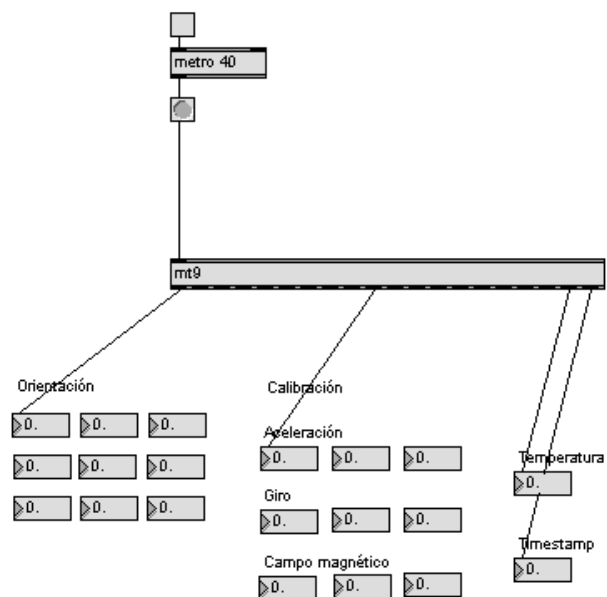


Figura 142. Objeto *mt9* para Jitter.

La conexión del GPS con Jitter se puede hacer de varias formas. Si el receptor GPS tiene conexión a puerto serie, se puede emplear el objeto *serial* y, mediante un programa escrito por José Manuel Berenguer (Berenguer, 2008), se puede desglosar el código recibido para extraer los datos de latitud y longitud (Figura 143). De otro modo, si la conexión es mediante bluetooth (puerto COM), se puede hacer uso de un objeto programado para el software Processing (Processing, 2008) por Julian Bleecker (Bleecker, 2005), y mediante el empleo de *MaxLink* (Kriss, 2008) para conectar

posteriormente Processing con Jitter. Para que los datos del GPS los reciba Jitter, el programa Processing debe estar en funcionamiento.

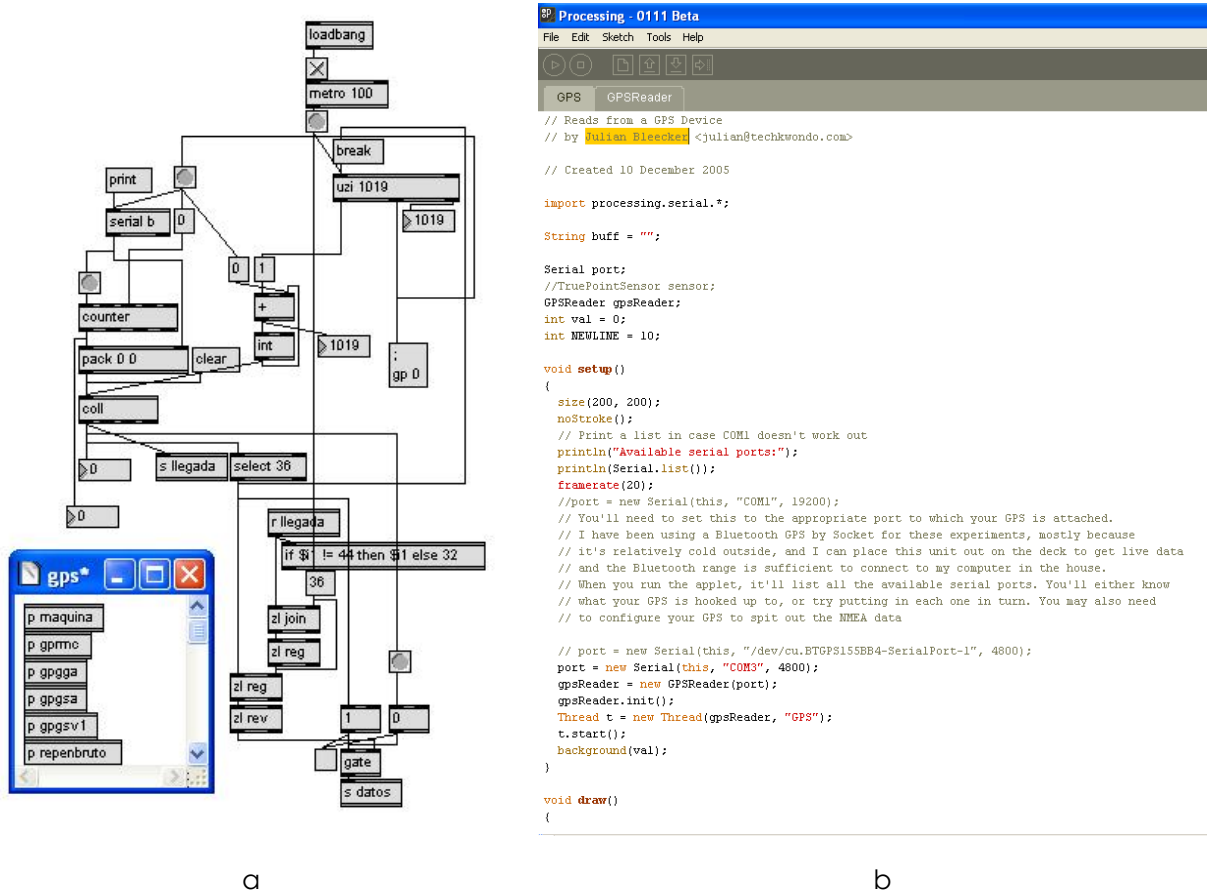


Figura 143. Conexión del receptor GPS con Jitter: a) Mediante el objeto serial y el patch programado por (Berenguer, 2008); b) Mediante MaxLink y el código escrito en Processing por (Bleeker, 2005).

5.3.1.4. LIMITACIONES

La principal limitación de este software es que no está expresamente elaborado para la realización de aplicaciones de realidad aumentada. Los procesos de cálculo que la tecnología RA implica (registro y combinación del entorno real y el virtual en tiempo real) hacen que se necesite de diversos objetos que realicen dichas tareas de forma conjunta. La utilización de imágenes de vídeo, gráficos 3D y, tal vez, registro de marcas fiduciales en las imágenes, hace necesaria la utilización de un hardware suficientemente potente.

6. EL USUARIO FINAL

Este capítulo constituye una parte esencial en el estudio de aplicaciones de realidad aumentada puesto que, como se ha visto en la introducción de esta tesis, el usuario constituye el último eslabón de la pirámide, y por lo tanto los factores referentes a éste deberían ser considerados en el diseño y evaluación de los sistemas o aplicaciones de realidad aumentada. Sin embargo, debido a la juventud de la tecnología de RA, los actuales esfuerzos realizados por la comunidad científica se centran en mejorar la tecnología, dejando en muchas ocasiones de lado al usuario final. Por ejemplo, en (Wang et al., 2006) se advierte que se han realizado muchos estudios referentes al usuario en entornos de realidad virtual, pero muy pocos en el campo específico de la realidad aumentada. Además, según (Livingston, 2005), muchos de los sistemas de realidad aumentada actualmente en desarrollo no son más que prototipos de laboratorio debido, en primer lugar, a que los investigadores necesitan hardware más avanzado que el que existe actualmente en el mercado para implementar dichos sistemas y, en segundo lugar, a que la comunidad investigadora necesita todavía resolver diversas cuestiones sobre el usuario. Sin embargo ambos factores están relacionados, puesto que los problemas ocasionados por las limitaciones en el hardware dificultan el estudio de cuestiones relacionadas con el usuario final. Por ejemplo, algunos displays tienen deficiencias en la resolución, campo de visión, brillo y contraste, visión estereoscópica, ergonomía, oclusión entre objetos reales y virtuales, o enfoque. Además los dispositivos de registro no ofrecen suficiente alcance, precisión, fiabilidad, latencia, ergonomía o facilidad en la calibración. Es por ello que los propios

dispositivos pueden imponer una barrera que aleja al usuario, sintiendo la experiencia como algo extraño, excesivamente tecnológico o poco intuitivo.

Salvando estos inconvenientes, surge la necesidad de ahondar en la temática de los entornos de realidad aumentada bajo el punto de vista del usuario; por ejemplo, es interesante comprender de qué manera se puede establecer una interacción más *natural* entre usuarios y ordenadores, estudiar la comunicación que se establece entre diferentes usuarios que colaboran en un entorno aumentado, la ergonomía de los dispositivos portables, la percepción del sistema creado, el grado de inmersión alcanzado, etc. Según (Livingston, 2005) si intentamos comparar el factores referentes al usuario en entornos de realidad aumentada con el de otros sistemas más tradicionales (como los entornos de realidad virtual), existe el riesgo de probar/validar una interfaz que todavía no ha sufrido pruebas rigurosas de diseño perceptivo y utilidad en la aplicación, frente a otra que sí lo ha alcanzado. Sin embargo, si no se demuestran los beneficios que puede aportar la realidad aumentada en este sentido, no habrá un interés por mejorar sus deficiencias. Es por ello que en los siguientes apartados se hace un recorrido de factores referentes al usuario basándose, en algunos casos, en estudios realizados en entornos de realidad virtual, abalados por la disciplina denominada Interacción Persona-Computador (*HCI – Human Computer Interaction*). En verdad, debido a la parte virtual/sintética propia de la realidad aumentada y a las muchas características que comparte con la tecnología de realidad virtual (por ejemplo, algunos displays, controladores, concepto de inmersión, etc.), podemos afirmar que muchos de los estudios realizados en el campo de la realidad virtual serán válidos también para la realidad aumentada, al menos en un cierto porcentaje.

6.1. INTERACTUANDO CON EL ENTORNO AUMENTADO

En muchas de las aplicaciones de RA el usuario (o grupo de usuarios) juegan un rol activo, es decir, pueden interactuar con el entorno aumentado. Estas interacciones están gestionadas por el ordenador, y la disciplina que las estudia recibe el nombre de Interacción Persona-Computador (*HCI, Human-Computer Interaction*), también conocida como Interacción Hombre-Máquina (*MMI, Man-Machine Interaction*) o Interacción Computador-Persona (*CHI, Computer-Human Interaction*). Según (Wikipedia, 2007c), la finalidad de la HCI es mejorar las interacciones entre usuarios y ordenadores, haciendo los segundos más útiles y receptivos a las necesidades de los primeros. Uno de los objetivos a largo plazo de la HCI, es diseñar sistemas que minimicen la barrera entre el modelo cognitivo de las personas de lo que quieren hacer y el entendimiento de dichas tareas por parte de los ordenadores. Específicamente, la HCI se ocupa de:

- Las metodologías y procesos para diseñar las interfaces;
- Los métodos para implementar interfaces;
- Las técnicas para evaluar y comparar interfaces;
- El desarrollo de nuevas interfaces y técnicas de interacción;
- El desarrollo de modelos descriptivos y predecibles de teorías de interacción.

Según (Roussou, 2004), en los entornos integrados por seres humanos y ordenadores, el término interacción se ha definido de diversas formas, dependiendo de si el contexto es operacional, mecánico, práctico, educacional, social, artístico o recreativo. Por ejemplo, en un nivel operacional se ha definido la interacción como la función de entrada requerida por el usuario (al ordenador) y la naturaleza de la respuesta del sistema debida a dicha entrada – aunque en muchas ocasiones se ha confundido el término interacción con la habilidad de simplemente mover un joystick o hacer clic con el ratón. En el contexto de exhibiciones públicas, donde existe una perspectiva más social y afectiva, el término interacción se ha definido como una experiencia que puede involucrar (*engage*) activamente al usuario de forma física, intelectual, emocional y/o social. En contextos educacionales, se ha definido la interacción como un mecanismo necesario y fundamental para la adquisición de conocimientos y el desarrollo de habilidades cognitivas y físicas. Generalmente, el término interacción se define como una característica intrínseca de la práctica educacional en el sentido de comunicación social, pero también es una propiedad inherente de cualquier entorno

interactivo multimedia o de RV que incluye actividades y respuestas físicas, sensoriales y mentales.

La interacción entre personas y computadores se produce a través de la interfaz, que puede ser un controlador (*inputs* o entradas de datos al ordenador), un display (*outputs* o salidas de datos del ordenador) o ambos (*inputs + outputs*). También cabe distinguir entre distintos tipos de interfaces (por ejemplo, gráficas, hápticas, sonoras o multimodales); y el hecho de que no todas las interfaces son de usuario.

6.1.1. FORMAS DE INTERACTUAR

6.1.1.1. MANIPULACIÓN DE OBJETOS VIRTUALES / REALES

En (Sherman et al., 2003) se distinguen cuatro métodos de manipulación de los objetos en entornos de realidad virtual: 1) Control directo del usuario: es un método de manipulación en el que el usuario interactúa con los objetos en el mundo virtual de igual forma que lo haría en el real; 2) Control físico: son aquellos que controlan el mundo virtual a partir de dispositivos reales, es decir, mediante interfaces físicas, las cuales proporcionan un feedback al usuario; 3) Control virtual: suelen ser representaciones de controles reales generadas por ordenador; 4) Control mediante agentes: son aquellos que permiten al usuario especificar comandos a través de un intermediario, un agente inteligente, que realizará la tarea especificada.

En un entorno de realidad aumentada los puntos 1 y 2 se entremezclan, siendo éstos la forma más común de interacción. Recordemos que una de las características que definen a los entornos de RA es el soporte de una metáfora de **interfaz tangible** (ver apartado 1.2.1). Es decir, los usuarios pueden manipular físicamente objetos habituales de su entorno, que pueden actuar como interfaz y estar directamente relacionados con objetos virtuales, por lo que el usuario puede literalmente “manipular” objetos virtuales. Es por ello que en este apartado se centra la atención en las interfaces hápticas.

Las manos son capaces de presionar, agarrar, apretar, y golpear objetos; pueden explorar las propiedades de los objetos tales como textura, forma y suavidad; pueden manipular herramientas para reparar otros dispositivos. El hecho de ser capaces de tocar, sentir y manipular objetos en un entorno, a parte de verlos y oírlos, proporciona

una sensación de inmersión en el entorno que de otra manera no sería posible. La exploración y manipulación de objetos sólidos, cubre un amplio rango de funciones hápticas, proporcionando un marco de tareas con el cual los roles de los subsistemas de la biomecánica, sensorio-motriz, y cognitivo se pueden entender. La información sensorial háptica de las manos en contacto con cualquier objeto, puede dividirse en dos clases: 1) Información táctil, refiriéndose al sentido del contacto con el objeto; 2) Información cinestésica, refiriéndose al sentido de la posición y movimiento de los cuerpos con las fuerzas asociadas (National Research Council Staff, 1994, 161-162).

Uno de los grupos de investigación destacado en el desarrollo de interfaces tangibles es el Tangible Media Group del MIT, dirigido por Hiroshi Ishii. Las piezas y trabajos realizados por este grupo responden al concepto del TUI (*Tangible User Interface*) en contraposición al GUI (*Graphical User Interface*). En la Figura 144 se aprecia un esquema comparativo dibujado por Ishii, que se explica a continuación según descrito en (Moggridge, 2007, 527):

- **GUI:** La interfaz gráfica de un ordenador provee de múltiples componentes, como ventanas, menús o iconos, pero los representa como píxeles intangibles, siendo el medio básico de representación de estos píxeles la pantalla del ordenador. Para controlar estas representaciones, la GUI provee de controladores genéricos, como ratones o teclados.
- **TUI:** La idea de la interfaz tangible elaborada por el grupo de Ishii, es darle forma física (o representación tangible) a la información y la computación. La representación está estrechamente asociada con los procesos internos del ordenador pero, al ser una representación física, también sirve como mecanismo de control, permitiendo a los usuarios coger y manipular elementos, controlando de este modo los procesos internos o la información digital. Sin embargo, uno de los problemas de la representación física es que no se pueden cambiar de forma sencilla parámetros como el color, la forma, etc. Es por ello que el grupo de Ishii asocia representaciones tangibles con intangibles, como una proyección de vídeo o un sonido, creando la ilusión de que ambas representaciones son perceptivamente una.

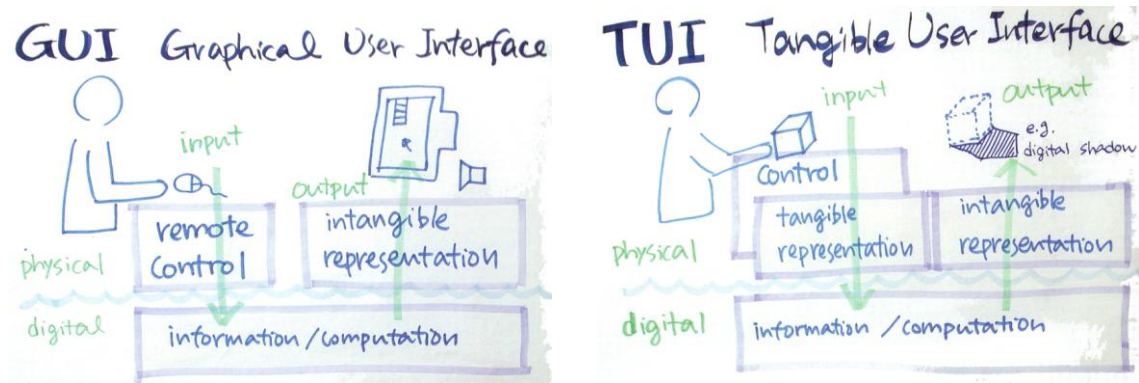


Figura 144. Esquemas dibujados por Ishii: GUI vs. TUI. En (Moggridge, 2007, 526).

En muchas de sus piezas, Ishii utiliza la metáfora del “ordenador invisible”, ya que el ordenador permanece oculto: aunque sea éste el que realice las tareas computacionales, el usuario no tiene por qué ser consciente de ello, incluso es irrelevante el hecho de que sepa o no manejar un ordenador, puesto que las interfaces son sencillas al estar asociadas a objetos cotidianos y al uso que de ellos se hace.

En otro orden de ideas, en (Wang et al., 2006) podemos encontrar un estudio de las aplicaciones de RA dentro del contexto de la arquitectura, ingeniería y construcción, distinguiendo entre una serie de interacciones cuando el usuario debe manipular objetos del mundo físico. Elaboran una clasificación que aborda factores a considerar en el diseño de entornos de RA. Dicha clasificación es la siguiente:

- **Ocupación de las manos:** distinguen cuatro tipos de ocupación de las manos: 1) Las manos no manipulan objetos y el usuario no interactúa con la aplicación de RA: estos sistemas se pueden emplear para aumentar actividades cotidianas, proporcionando la información de manera continua; 2) Las manos están ocupadas únicamente para interactuar con la aplicación de RA: las manos interactúan con la información digital mediante dispositivos tangibles que no tienen de manera necesaria una entrada física a una tarea, como en procesos de mantenimiento o inspección; 3) Las manos están ocupadas únicamente para realizar tareas físicas: las manos se ocupan de realizar las tareas manuales, únicamente con la visualización de información digital; 4) Las manos se utilizan tanto para realizar tareas como para interactuar.
- **Tareas físicas concurrentes:** el cuarto caso del apartado anterior involucra procesos multitarea (*dual-task*), tareas con las manos y el cambio asiduo de atención entre tareas. Si el tiempo requerido para cambiar de tareas es excesivo, el rendimiento será menor. Esta carga excesiva se puede evitar

integrando información aumentada sobre las propias tareas (por ejemplo, mediante un dispositivo tangible que actúe tanto como interfaz y como el objeto de la tarea). Esta mezcla de entidades reales y virtuales corresponde al problema de combinar dos modalidades de interacción, una unida a la entidad virtual y otra a la entidad real.

- **Tareas mentales concurrentes:** Muchas labores requieren que los usuarios realicen varias tareas mentales al mismo tiempo. En estos casos, los entornos de realidad aumentada deben estar diseñados de tal modo que se facilite la monitorización e integración de tareas concurrentes. El empleo de dispositivos tangibles puede ser una alternativa para solucionar estas cuestiones.

6.1.1.2. NAVEGACIÓN

Otra de las formas de actuación que es característica de los entornos virtuales y/o aumentados es la navegación. Según dice (Murray, 1999, 142), *“la habilidad para moverse a través de los espacios virtuales es un placer en sí misma, independientemente del contenido de tales espacios”*; y añade *“construir el espacio y moverse a través de él explorándolo (...) es una actividad placentera independientemente de que el espacio sea virtual o real”*. Así pues, muchas aplicaciones de RV y RA – sobre todo en el campo del patrimonio cultural, arqueología o actividades museísticas – se basan en el simple hecho de la navegación.

Según (Roussou, 2004) los interfaces y aplicaciones actuales de RV están diseñados con la conciencia de la interacción, dotando al usuario de medios para que literalmente se sienta ubicado en la escena y comprometido con su entorno. En muchas aplicaciones, la interactividad se identifica generalmente con la habilidad de elegir el recorrido dentro del entorno generado y navegar libremente por éste. Tradicionalmente, la navegación se suele producir con un joystick (sobre todo en entornos de RV), aunque existen otro tipo de aplicaciones, sobre todo artísticas, en las que la interfaz resulta más novedosa e interesante. Un ejemplo destacable es la pieza de *The Legible City* (1989) de Jeffrey Shaw (Figura 145), en la que el usuario se monta en una bicicleta estática para navegar por una peculiar representación virtual de la ciudad que consta de una serie de letras 3D generadas por ordenador, formando palabras y frases a lo largo de los lados de las calles (Shaw, 1989). La ubicación de estas letras responde a los planos de ciudades reales (Manhattan, Ámsterdam y Karlsruhe). El viaje a través de estas ciudades constituye una jornada de lectura, donde el usuario puede elegir el camino que toma. El manillar y los pedales de la bicicleta constituyen la interfaz usuario-

máquina, otorgando al usuario el control interactivo de la dirección y velocidad del viaje. Un vídeo proyector se utiliza para la visualización de las imágenes generadas por ordenador. Otro monitor pequeño en frente de la bicicleta muestra al usuario una planta simplificada de cada ciudad, indicando en todo momento la posición del ciclista.

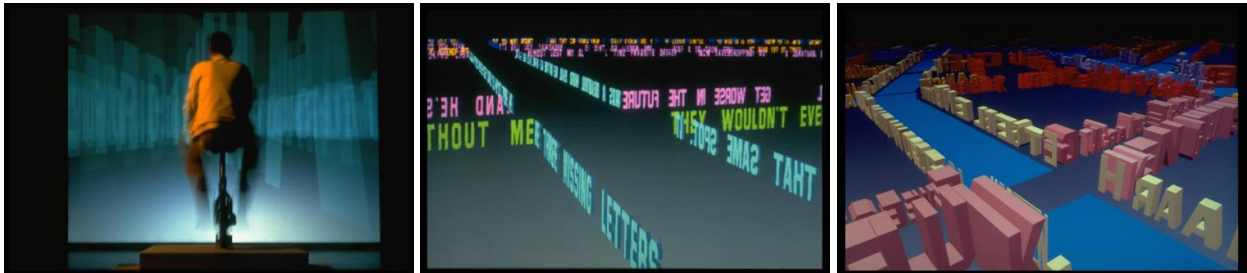


Figura 145. *The Legible City*. En (Shaw, 1989).

En la tecnología de realidad aumentada la navegación es el tipo de interacción predominante en las aplicaciones móviles, en donde los usuarios interactúan con el mundo aumentado principalmente a partir de su localización, como se ha visto anteriormente en (Cheek et al., 2003; Benford et al., 2006; Rashid et al., 2006; Le Prado et al., 2007).

6.1.1.3. INTERACCIÓN CON OTROS USUARIOS

Otra característica importante a tener en cuenta en cuanto a la forma en que se produce la interacción, es si existe la posibilidad de que varios usuarios interactúen simultáneamente con el ordenador, o incluso entre ellos. Hay sistemas que se basan en una perspectiva **centrada en un solo usuario**, excluyendo de este modo la participación simultánea de un grupo de gente (la mayoría de los mundos virtuales en el área de arquitectura o patrimonio cultural se basan en este principio); otros sistemas ofrecen la posibilidad de establecer **entornos colaborativos** y por lo tanto permiten la participación activa de diversos usuarios (Roussou, 2004). Aunque los sistemas de RA para un solo usuario ofrecen muchas posibilidades, el mayor potencial de la RA reside en el desarrollo de nuevas formas de establecer comunicación para un grupo de usuarios, mientras se realizan simultáneamente diversas tareas. Así pues, en este apartado se va a indagar en la interacción entre humanos dentro de un entorno aumentado, con la mediación de los ordenadores.

Después de un estudio sobre interfaces utilizadas en varios sistemas de RA de tipo colaborativo, en (Billinghurst et al., 2002b) se concluye que existen una serie de características propias de la tecnología de la RA (ver apartado 1.2.1) que permiten el diseño de interfaces que mejoran la comunicación, por lo que la tecnología de RA se puede utilizar para mejorar la comunicación cara a cara/presencial y la colaboración remota de un modo que sería difícil conseguir con tecnología tradicional.

6.1.1.3.1. Colaboración cara a cara/presencial

Según (Billinghurst et al., 2002b), en la colaboración cara a cara las personas utilizan toda una serie de elementos, como el habla, gestos, miradas o indicaciones no verbales para poderse comunicar de la manera más clara posible. Los objetos físicos también pueden formar un papel importante en la comunicación, – por ejemplo, cuando se está explicando como se realiza una tarea – soportando la colaboración tanto en sus apariencias como en sus alcances físicos, sus usos como representaciones semánticas, sus relaciones espaciales, y sus habilidades para ayudar a mantener la atención; son los constituyentes de la actividad colaborativa, crean marcos de referencia para la comunicación y alteran la dinámica de la interacción, especialmente en entornos multiusuario. En contraste, muchas interfaces computacionales en la colaboración presencial crean una barrera artificial (espacial, temporal o funcional) entre el mundo real y el mundo digital compartido.

También, según (Billinghurst et al., 2001b, 564), en entornos de RA donde se establece una colaboración presencial (*collocated collaboration*), los usuarios pueden referir a notas, diagramas, libros y otros objetos reales mientras siguen viendo los objetos virtuales; también pueden utilizar herramientas reales que les resulten familiares para manipular los objetos virtuales, incrementando el carácter intuitivo de la interfaz; además, pueden ver la expresión facial del resto de colaboradores, sus gestos y lenguaje corporal, incrementando así el grado de comunicación. Todos estos factores permiten que el proceso de comunicación mediatizado por ordenador sea más natural.

Un ejemplo es la pieza *Zona de Recreo* (2003) de Diego Díaz y Clara Boj (Díaz, 2007), que consiste en una plataforma interactiva multiusuario que surge de la adaptación de un balancín tradicional mediante la incorporación de nuevas tecnologías (Figura 146). En una pantalla se visualizan distintas aplicaciones gráficas a través de las cuales los

usuarios pueden navegar mediante la inclinación de la plataforma. Para una correcta navegación, los usuarios deberán mover sus cuerpos de forma acorde unos con otros.

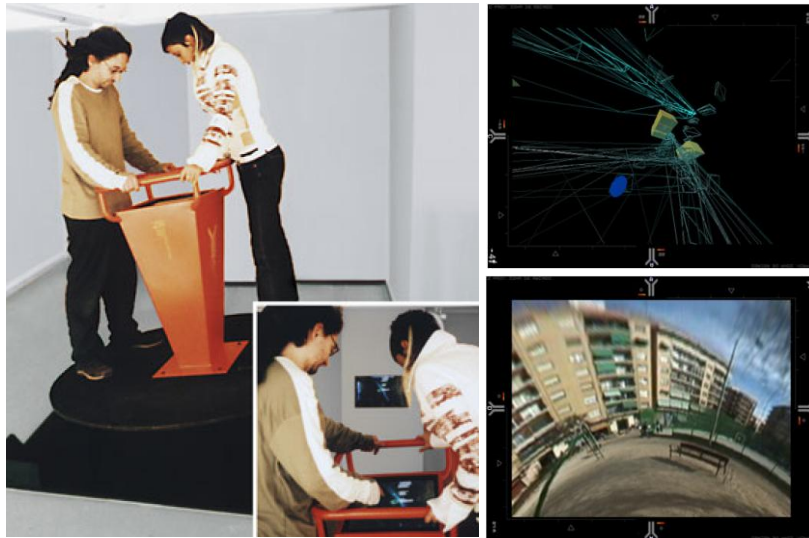


Figura 146. Usuarios interactuando con la pieza *Vídeo Balancín*. En (Díaz, 2007).

6.1.1.3.2. Colaboración remota

En (Billinghurst et al., 2002b) se argumenta que gran parte de la tecnología actual para la colaboración remota afecta directamente al proceso de comunicación. Por ejemplo, la utilización de interfaces únicamente sonoras eliminan las entradas visuales vitales para el establecimiento de turnos en conversaciones, incrementando las interrupciones y solapamientos, propiciando la ambigüedad entre interlocutores y dificultando la determinación del deseo de interacción de los otros. Con las interfaces convencionales de videoconferencia los movimientos sutiles o gestos no pueden ser capturados, no hay datos espacializados (*spatial cues*) entre los participantes, el número de éstos se limita a la resolución del monitor y no pueden establecer fácilmente contacto visual. Los oradores no pueden saber cuando la gente les está prestando atención o si pueden mantener conversaciones paralelas. Por todo ello, en el proceso de comunicación remota mediada por ordenador puede que sea imposible obtener la misma experiencia que en la colaboración presencial, por la misma naturaleza del medio, pero sí es posible compensar estos defectos y ofrecer un entorno de interacción incluso más rico. Según (Billinghurst et al., 2001b, 546), en lugar de utilizar los nuevos medios para tratar de emular la comunicación presencial, muchos de los investigadores se preocupan por estudiar los nuevos atributos que los medios pueden ofrecer para que se establezca una comunicación eficaz a pesar de la lejanía

física de los colaboradores. También se advierte que la espacialización permite una comunicación remota más natural.

En muchos sistemas de RA para comunicación remota, la finalidad es sencillamente establecer un ámbito de conversación apropiado y enriquecido, en el que los usuarios se sientan inmersos. Podemos citar diversas aplicaciones en el campo de la telecomunicación, como (Billinghamurst et al., 2002a; Lee et al., 2006a; Barakonyi et al., 2007). Dentro del ámbito de la comunicación remota, existe una aplicación muy peculiar de realidad mezclada en la que se establece una comunicación remota entre una persona y un animal de corral: *Poultry Internet* (Figura 147), descrita en (Cheek et al., 2006; Lee et al., 2006b; MXR, 2006). El sistema consta de dos espacios, uno en la oficina y otro en el campo, donde reside el animal. En la oficina, está instalada una representación física del animal montada sobre un brazo mecánico que permite desplazamientos horizontales y giros, acordes con los movimientos reales del animal. En el otro espacio se encuentra el animal equipado con un vestido computacional que simula la sensación del tacto de tal forma que, cuando el usuario acaricia al muñeco, el animal lo percibe.

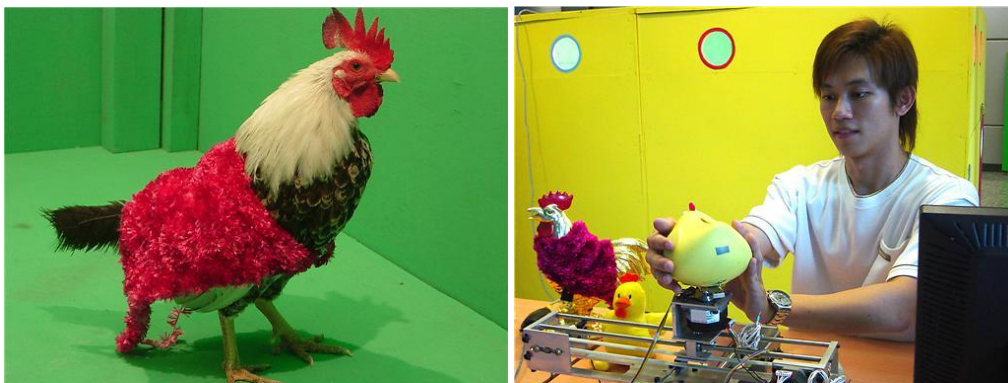


Figura 147. Poultry Internet. En (MXR, 2006).

6.1.1.3.3. Entornos aumentados colaborativos

En 1998 se publica un artículo (Szalavári et al., 1998) sobre una aplicación de RA basada en un juego tradicional chino denominado Mah-Jongg, en el que la comunicación social es tan importante como el mantener cierta privacidad para establecer decisiones estratégicas (Figura 148). Los autores advierten que, hasta el momento, los juegos de ordenador no son capaces de mantener la comunicación social al mismo nivel que en los juegos tradicionales, en los que es más natural. Tras la puesta en práctica de su aplicación, identifican una serie de factores cruciales para los

juegos colaborativos de RA: la habilidad para mostrar información diferente a cada participante, comunicación social sin obstáculos, e interacción precisa y rápida. Este juego constituye uno de los primeros entornos colaborativos de realidad aumentada.

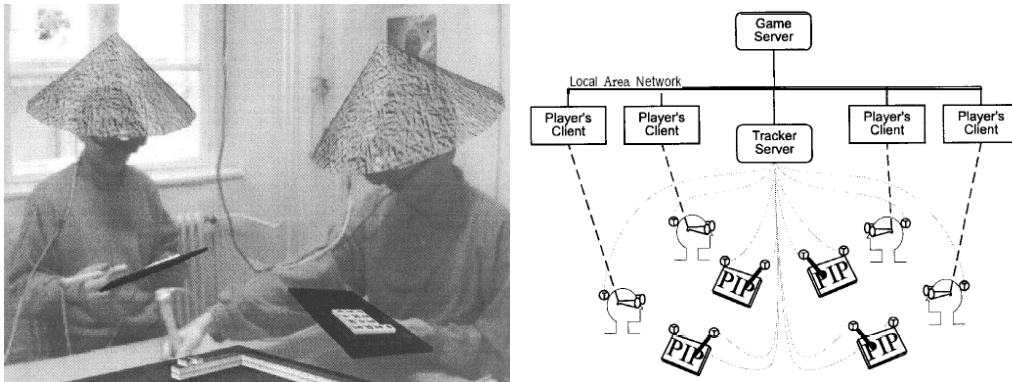


Figura 148. Usuarios interactuando en un juego colaborativo y esquema de la red del juego. En (Szalavári et al., 1998).

En este apartado se introduce el concepto de entorno aumentado colaborativo a partir de su equivalente virtual. Según (Sanchez-Segura, 2004, 236), *“un entorno virtual colaborativo (collaborative virtual environment, CVE) es un espacio generado por ordenador, tridimensional en el cual una serie de usuarios pueden interactuar en tiempo real”*. A partir de esta definición, podemos decir que un entorno aumentado colaborativo (*collaborative augmented environment, CAE*) es un espacio híbrido, generado en parte por ordenador, tridimensional en el cual una serie de usuarios pueden interactuar en tiempo real y que permite establecer contacto visual entre ellos. Las diferencias entre los CVE y los CAE pueden ser significativas: en los procesos generales de comunicación colaborativa se produce una interacción natural humano-humano; por otro lado, los sistemas de CVE y CAE tratan de emular dicha interacción para facilitar la comunicación mediatizada por ordenadores; sin embargo, el hecho de que los sistemas de CAE permitan establecer contacto visual, hace que la comunicación mediatizada por ordenador se acerque más a una comunicación intuitiva, rompiendo una de las barreras de los CVE.

Según (Sanchez-Segura, 2004, 236), un CVE (también válido para un CAE) tiene que soportar audio, video, y datos potencialmente hápticos a los participantes, así como capacidades interactivas y reactivas, de tal forma que pueda responder al participante y presentar un mundo interesante. Además, mientras que tanto las entradas como las salidas se producen en tiempo real, el sistema también necesita distribuir todos los datos a múltiples usuarios situados en lugares diferentes, de tal modo

que puedan colaborar. Estos sistemas se centran en distribuir correctamente en el entorno local de cada participante el volumen de datos generados. Además, en su diseño es importante tener en cuenta aspectos característicos de la comunicación colaborativa que se establece independientemente de los ordenadores. De este modo, si se estudian diversos aspectos de la conducta humana en dicho proceso, se podrán mejorar ambos sistemas. Algunos aspectos que se pueden explorar sobre la interacción y comportamiento de los participantes en un CVE son, según (Sanchez-Segura, 2004):

- **Apariencia:** representación de los usuarios mediante avatares;
- **Conciencia:** lo que el avatar puede percibir sobre el entorno virtual y la situación;
- **Interacción:** con objetos del entorno y con el mismo entorno;
- **Toma de turno:** para comunicar, interactuar con los objetos u otros usuarios;
- **Contenido de la comunicación:** los intercambios comunicativos reales;
- **Modos de comunicación:** las formas en que la comunicación de los avatares se puede presentar;
- **Pedagogía:** los roles y tácticas que los usuarios adoptan para exponer varios tópicos.

Estos aspectos se pueden extrapolar a los sistemas de CAE. Finalmente, en (Billinghurst et al., 2001b, 564-565) se describen cinco características que definen los CAE:

- **Virtualidad:** los objetos generados por ordenador se pueden ver y examinar;
- **Aumentación:** los objetos reales se pueden aumentar con anotaciones virtuales;
- **Cooperación:** múltiples usuarios se pueden ver entre sí y cooperar de manera natural;
- **Independencia:** cada usuario controla su propio punto de vista independiente;
- **Individualidad:** los datos mostrados pueden ser diferentes para cada usuario para permitir diferentes roles.

6.1.2. INTERACCIÓN SIN COSTURAS

Según se vio en el apartado 1.2.1, una de los aspectos que proporciona la tecnología de RA es la “interacción sin costuras entre los entornos virtual y real”. En este apartado

se profundizará en este tema. En primer lugar, cabe analizar el término "costura en una interfaz" definido en 1994 por Ishii (Ishii et al., 1994) como "*un constreñimiento espacial, temporal o funcional que fuerza al usuario a cambiar entre una variedad de espacios o modos de operación*". Las costuras pueden ser de dos tipos:

- Costuras **funcionales**: Discontinuidades entre dos espacios de trabajo diferentes, obligando al usuario a cambiar los modos de operación. Una de las costuras funcionales más importantes es la que se establece entre los espacios de trabajo compartidos y los interpersonales. Los primeros constituyen el espacio de trabajo compartido entre los colaboradores, mientras que los segundos son el espacio de comunicación común. En la comunicación cara a cara el espacio de trabajo compartido suele ser un subgrupo del interpersonal, y el cambio de atención de uno a otro se realiza de forma sencilla y dinámica mediante comunicación no verbal. Sin embargo, en la comunicación mediada por ordenador se suelen establecer costuras arbitrarias al centrar la atención en la pantalla de un ordenador, ya que se obstaculiza la comunicación no verbal (pérdida de contacto visual entre colaboradores).
- Costuras **cognitivas**: Discontinuidades entre prácticas de trabajo existentes y nuevas, obligando al usuario a que aprenda nuevas maneras de trabajar. Una costura cognitiva común es la que se da entre las herramientas ofrecidas por un computador y las herramientas tradicionales. Normalmente las herramientas ofrecidas por un ordenador son rechazadas si obligan al usuario a cambiar su modo de trabajo, si llevan asociadas un aprendizaje adicional o si no se pueden integrar de manera eficaz con las herramientas existentes.

Sin embargo, es posible diseñar interfaces sin costuras que mejoren la comunicación. Para que una interfaz minimice tanto las costuras funcionales como las cognitivas, debe tener las siguientes características (Billinghurst et al., 2001b, 544):

- Debe soportar las herramientas y técnicas de trabajo existentes.
- Los usuarios deben ser capaces de aplicar objetos del mundo real a la interfaz.
- El espacio de trabajo compartido tiene que ser un subgrupo del espacio interpersonal.
- Tiene que haber comunicación sonora y visual entre los participantes.
- Los colaboradores deben ser capaces de mantener contacto visual y mirada consciente.

En (Ishii et al., 1994) se presenta *ClearBoard-2*, una interfaz diseñada para asistir en entornos colaborativos de trabajo (Figura 149). El sistema consta de dos módulos, cada uno de ellos consistente en una mesa de trabajo con una pantalla especial (compuesta por un espejo semitransparente, película polarizada, pantalla de proyección y una lámina transparente digitalizadora). Por medio de la combinación de proyecciones, espejos, cámaras CCD y pantallas especiales, se logra el efecto de mantener contacto visual directo con el compañero de trabajo situado en la otra mesa, al mismo tiempo que los dibujos de ambos se entremezclan.

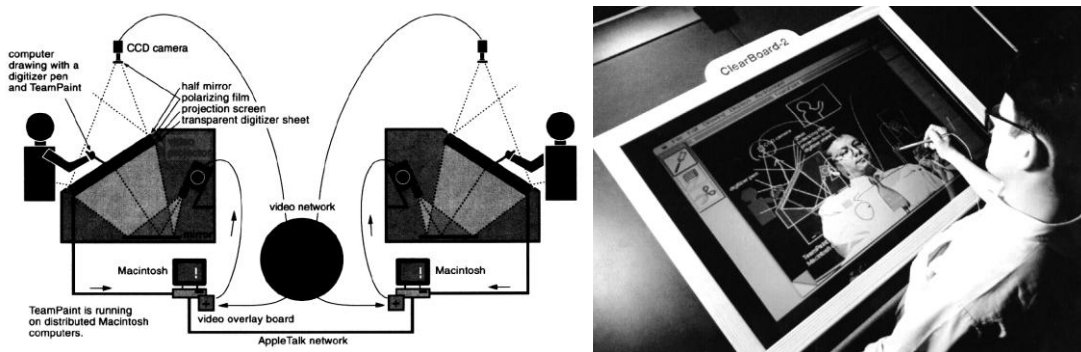


Figura 149. Diseño e interacción con *ClearBoard-2*. En (Ishii et al., 1994).

6.1.3. VESTIDOS COMPUTACIONALES

En el estudio de la interacción persona-computador cabe mencionar el papel de los vestidos computacionales, por ser verdaderos ordenadores integrados en el cuerpo de la persona. Según (Mann, 1997), *“los vestidos con capacidades computacionales consisten en un ordenador que se subsuma al espacio personal del usuario, controlado por el usuario, y que tiene constancia tanto operacional como interactiva”*.

Cabe distinguir entre los términos “computadores para llevar puestos” (*wearable computers*) y “vestidos computacionales” (*computational clothing*). En el primer caso los ordenadores pueden ser llevados por las personas por ejemplo, dentro de una mochila, pero no forman parte de su cuerpo. Sin embargo, el vestido computacional tiene toda la capacidad computacional de un ordenador pero, además, está integrado con la persona que lo lleva a modo de vestimenta. Según (Barfield et al., 2001b, 474) hay una serie de características o atributos fundamentales que tendrán los vestidos computacionales (del futuro):

- **Memoria fotográfica:** Si el sistema utiliza una cámara, el usuario podrá en todo momento recobrar la información registrada.

- **Memoria compartida:** En un entorno colaborativo, los usuarios podrán compartir su información.
- **Inteligencia colectiva conectada:** En un entorno colaborativo, unos usuarios podrán colaborar, mientras que otros podrán estar realizando otra tarea.
- **Seguridad personal:** En contraste con las redes de vigilancia centralizada construidas en la arquitectura de la propia red, en los vestidos computacionales, la vigilancia estará construida en la arquitectura (vestido) del usuario.
- **Operaciones sin cadenas:** Permitirán la movilidad, evitando la conexión por cables a una salida eléctrica o a una línea de comunicación.
- **Sinergia:** La finalidad es conseguir una sinergia entre usuario y ordenador, en la que cada uno realiza la tarea que mejor sabe hacer. Existe la expectativa de que, en un futuro, los vestidos computacionales serán una prolongación del cuerpo y mente del ser humano, y no una parte adicional.

Además, se pueden distinguir tres modos operacionales que representan la interacción humano-computador asociados a los vestidos computacionales (Barfield et al., 2001b, 476):

- **Constancia:** La parte computadora de estos vestidos funciona de manera continua y está siempre preparada para interactuar con el usuario. Esto no sucede con dispositivos tales como PDA, ordenadores portátiles, etc.
- **Aumentación:** En estos sistemas la parte principal no es la computación, ya que se supone que el usuario estará haciendo otras tareas al mismo tiempo que interactuando con la parte computacional del vestido. Por lo tanto, la tarea principal de estos sistemas es aumentar el intelecto del usuario, o sus sentidos.
- **Mediación:** Los vestidos computacionales encapsulan o envuelven al usuario, aunque no necesariamente completamente. Podemos distinguir dos aspectos de este encapsulamiento:
 - **Soledad:** El vestido computacional puede servir como una barrera que filtra la información que el usuario no desea recibir.
 - **Privacidad:** Puede también funcionar como un intermediario para interactuar con sistemas de poca confianza.

En (Bonsor, 2008) se presentan algunos ejemplos de vestidos computacionales (Figura 150). La chaqueta musical de Levi's está hecha de seda organza y se controla con un teclado de tela totalmente flexible, duradero y sensible al tacto, hecho a partir de hilo

conductor. El teclado está conectado a un mini sintetizador MIDI que reproduce música. Otro ejemplo es un teclado de tela desarrollado en el MIT Media Lab, hecho a partir de materiales conductivos y no conductivos entrelazados en filas y columnas. El producto tiene el aspecto de un edredón confeccionado a partir de retales. Las columnas conductivas acolchadas están aisladas y forman las filas conductoras mediante una tela suave. Diversos agujeros en la tela aislada permiten a los conductores de filas y columnas hacer contacto cuando el usuario los presiona. Este material se puede lavar en la lavadora al igual que una pieza normal de ropa. Finalmente, *Smart Shirt* del Georgia Institute of Technology puede monitorizar la respiración y el ritmo cardíaco utilizando fibras conductoras ópticas y eléctricas, que están tejidas en la tela de la camiseta.

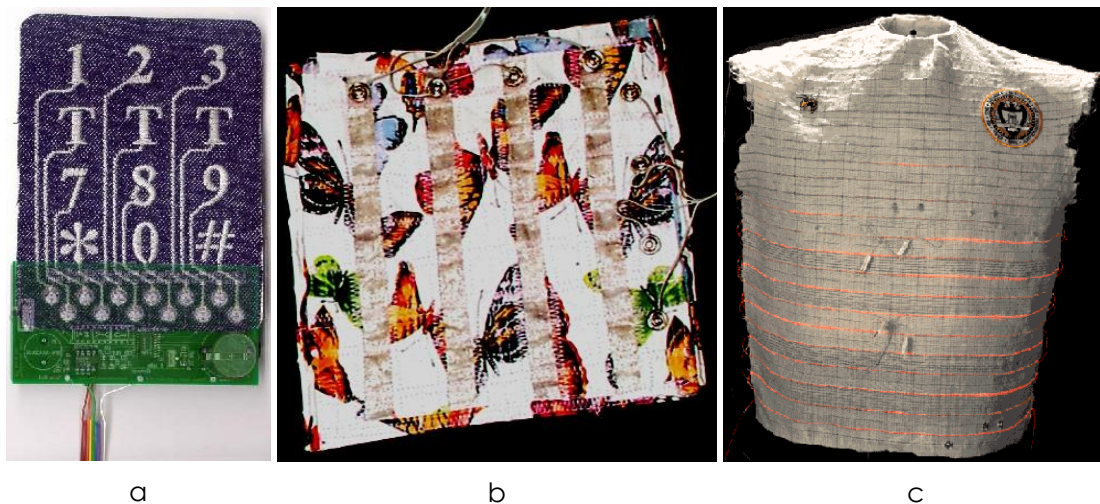


Figura 150. Vestidos computacionales: a) Teclado que controla la chaqueta musical de Levi's, estando completamente hecho de tela; b) Tela elaborada por el MIT Media Lab, completamente lavable; c) Uno de los primeros prototipos de *Smart Shirt* desarrollada en Georgia Tech. En (Bonsor, 2008).

Los diseñadores de interacción de la Carnegie Mellon University, han elaborado un total de trece pautas para el diseño de vestidos computacionales, basándose en la hipótesis de que a pesar de la visión de textiles y prendas computacionales para el futuro, será también necesario integrar de manera ergonómica las formas sólidas de los actuales ordenadores o dispositivos en el cuerpo humano. Estas pautas son (Barfield et al., 2001b, 490-495):

- Pauta 1: El diseño requiere una localización confortable sobre el cuerpo humano, para permitir una portabilidad dinámica.

- Pauta 2: Adaptación a la forma dinámica del cuerpo para asegurar un ajuste cómodo y estable.
- Pauta 3: El movimiento del cuerpo humano ofrece tanto un constreñimiento como un recurso para el diseño de formas dinámicas. La libertad de movimiento se puede conseguir mediante el diseño alrededor de las zonas más activas de las uniones/juntas o creando espacios dentro del vestido en los que el cuerpo se pueda mover.
- Pauta 4: Diseñar para determinadas distancias de relación social. El ser humano reacciona de forma diferente a las relaciones que se producen en su espacio personal, social o íntimo, como si de un campo activo en torno al cuerpo se tratara y cuyas distancias varían según las culturas. Los vestidos computacionales deben tener en cuenta estas distancias según sea su finalidad. El estudio de estas capas de percepción alrededor del cuerpo se denomina proxémica.
- Pauta 5: Deben diseñarse para que puedan acoplarse al mayor número de usuarios, permitiendo la variación en cuanto a tamaño.
- Pauta 6: Las formas creadas se pueden adherir de manera confortable envolviéndolos alrededor del cuerpo, en lugar de usar adhesiones puntuales tales como clips o ganchos.
- Pauta 7: En el diseño de vestidos computacionales, en muchos casos se requiere que el vestido contengan diversos objetos.
- Pauta 8: El peso del vestido no debe dificultar el movimiento del cuerpo o balancearse.
- Pauta 9: Tener en cuenta las normas de accesibilidad y uso para que el producto sea más útil.
- Pauta 10: La interacción sensorial, tanto activa como pasiva.
- Pauta 11: Hay tres aspectos térmicos en el diseño de objetos para el cuerpo: funcional, biológico, y perceptivo. El cuerpo necesita respirar y es muy sensible a los focos de calor.
- Pauta 12: Un aspecto importante de la forma y función de cualquier objeto para el cuerpo es la estética. La cultura y el contexto dictaminan la forma, materiales, texturas y colores que perceptivamente encajen con los usuarios y con el entorno.
- Pauta 13: Probar los efectos que puedan producirse en el cuerpo humano por el uso de los vestidos a largo plazo.

6.2. LA EXPERIENCIA EN EL ENTORNO AUMENTADO

En este apartado se trata la experiencia del usuario dentro de un entorno aumentado. Para ello se analizan los elementos básicos que inciden en dicha experiencia, atendiendo a factores de percepción, inmersión y compatibilidad. A continuación se centra la atención sobre diversos factores relacionados con el usuario en entornos de RA dentro del ámbito del entretenimiento, educación y arte, por ser los que se han tratado en los ensayos de esta tesis (capítulos 7-12).

6.2.1. PERCEPCIÓN HUMANA – PERCEPCIÓN AMPLIADA

Las representaciones en el entorno aumentado están filtradas por el sistema de percepción humano, y es por ello que en el diseño de aplicaciones de RA deben tenerse en cuenta diversas características humanas, a nivel fisiológico, físico y/o emocional. Los seres humanos percibimos el entorno que nos rodea a través de nuestros sentidos, siendo la visión el sentido de percepción más poderoso y, a la vez, más complejo. Sin embargo no somos conscientes de ello: con un simple abrir de ojos automáticamente vemos el espacio que nos rodea, y con un solo vistazo creemos que somos capaces de comprenderlo. Las imágenes están llenas de mensajes que nuestro cerebro puede procesar e interpretar a gran velocidad, velocidad que se incrementa cada vez más dado que nuestra sociedad está inmersa en un mundo predominantemente visual. Artilugios como la televisión, cámaras digitales, ordenadores, etc. se han convertido en elementos domésticos que generan y/o emiten imágenes en nuestro entorno cotidiano. Además, hay una tendencia a visualizarlo todo; incluso intentamos visualizar todo tipo de datos por medio de gráficos o esquemas para poder comprenderlos mejor y, sobre todo, más rápido. Es por ello que las aplicaciones de realidad aumentada y/o mezclada tienen un fuerte componente visual, en algunos casos, el único estímulo empleado.

Según se describe en (Sherman et al., 2003), una de las características interesantes de la cognición y percepción humana es la habilidad de generalizar. Por ejemplo, al ver una mesa sabemos que es una mesa aunque no la hayamos visto antes, sean cuales sean sus características (tamaño, color, material, etc.), ya que somos capaces de extraer unos conceptos comunes a todas las mesas (una tabla horizontal apoyada sobre tres o más patas). Esta habilidad es importante en el mundo de la informática, ya que permite la creación de objetos simplificados o símbolos que sean entendibles por

todo el mundo, con el consiguiente ahorro computacional que ello supone. La generalización nos permite agrupar objetos y conceptos con características similares. En este sentido, las aplicaciones de realidad aumentada están llenas de generalizaciones y simbolismos, simplificando en gran medida los elementos generados por ordenador, que son una abstracción de los elementos referenciales que representan. Por ejemplo, en *Kyoto Garden* (Cheek et al., 2006) los elementos de un jardín japonés se simplifican tanto en su representación virtual 3D como en su representación gráfica 2D (en los patrones de las marcas físicas); además, las instrucciones para interactuar con dichos elementos se muestran mediante símbolos virtuales convencionales (Figura 151). Esto contribuye tanto a un menor consumo computacional como a un mayor entendimiento del juego, mediante simbolismos que resultan familiares y por lo tanto sencillos de entender.

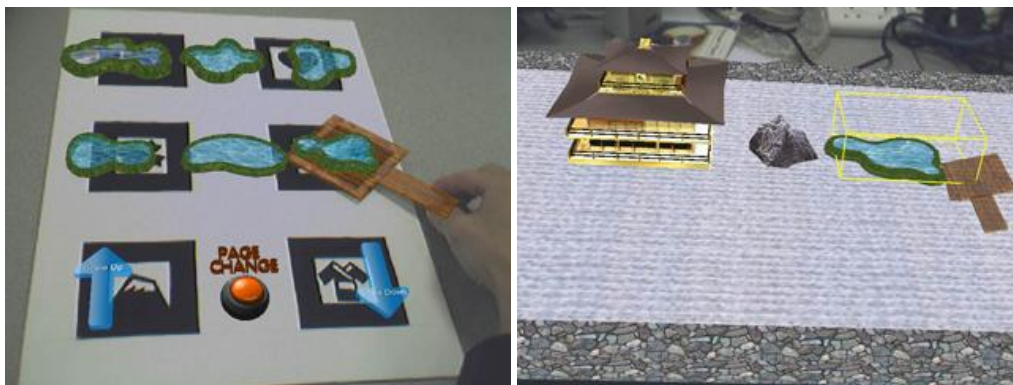


Figura 151. Elementos gráficos en *Kyoto Garden*. En (Cheek et al., 2006).

En un entorno de realidad aumentada, el reto consiste en integrar técnica y perceptivamente ambos mundos (el real y el virtual) de tal forma que el usuario pueda formar un único modelo mental de la escena aumentada. Según (Barfield, 1995), a mayor realismo de los dos mundos, a mayor correspondencia entre ambas imágenes en cuanto a brillo y color, y a registros de imagen más precisos, mayor facilidad tendrá el usuario para ver la escena aumentada como un único entorno. Sin embargo, otros autores mantienen que el realismo de ambos mundos no es crucial para que el conjunto sea creíble, sino que lo que realmente importa es la compatibilidad (ver apartado 6.2.3).

6.2.2. INMERSIÓN

Según (Sherman et al., 2003, 7, 381) existen dos tipos de inmersión: la inmersión física (o sensorial) y la inmersión mental. La expresión “estar inmerso” generalmente se utiliza

refiriéndose a un estado mental o emocional (sentimiento de estar participando de una experiencia). En los entornos de realidad aumentada y realidad virtual se utiliza el término inmersión física para referirse a la propiedad de estos sistemas de reemplazar o aumentar los estímulos hacia los sentidos de los usuarios. A continuación se definen estos conceptos de acuerdo con el citado autor:

- **Inmersión:** Sensación de estar dentro de un entorno; puede ser un estado puramente mental o puede estar completado a través de medios físicos.
- **Inmersión mental:** Estado de encontrarse profundamente involucrado; suspensión de incredulidad; participación. Se suele emplear como sinónimo de "sensación de presencia".
- **Inmersión física:** Entrar en un medio de forma física; estímulos sintéticos de los sentidos del cuerpo a través del uso de la tecnología; esto no implica que todos los sentidos o el cuerpo entero estén inmersos.

En un entorno de realidad aumentada la inmersión física se lleva a cabo mediante la presentación al usuario del entorno creado basado en su localización y orientación – el cual puede estar o no representado por un avatar –, y proporcionando estímulos sintéticos a uno o más de sus sentidos en respuesta a sus movimientos y acciones. Es decir, el ordenador sabe dónde se encuentra el usuario o qué acciones realiza mediante su registro (ver capítulo 3). Cuando el usuario se mueve, las cualidades gráficas, sonoras o hápticas que establecen la inmersión física dentro del entorno aumentado cambiarán respondiendo a dicho movimiento o acción.

El grado de inmersión mental requerido dependerá del tipo de aplicación. Por ejemplo, en aplicaciones dentro de la industria del entretenimiento o el campo artístico, la inmersión mental es muy importante, para que el usuario sienta que forma parte del juego/rol propuesto e incitar a su participación activa o crítica. En otras aplicaciones, como por ejemplo aquellas cuyo objetivo es meramente explorar diversos elementos aumentados (piezas documentales), el grado de inmersión mental no es tan relevante.

Como ya se ha visto anteriormente (apartado 4.5), las aplicaciones de RA multimodales ayudan a incrementar la inmersión mental del usuario. Existen otros factores que pueden incrementar el grado de inmersión, como por ejemplo, la representación tridimensional de la parte virtual (en contraposición a una representación bidimensional, plana) o las características de los dispositivos utilizados (frecuencia de imagen y de actualización, retardos del sistema, rango y sensibilidad de

los sensores de registro, etc.). De ello se desprende que el grado de compatibilidad entre el mundo real y virtual es fundamental para que la escena creada parezca creíble o verosímil al usuario, incrementando la sensación de presencia.

6.2.3. COMPATIBILIDAD

Para mezclar adecuadamente las entidades reales y virtuales en un entorno de realidad aumentada, debe estudiarse la **compatibilidad** entre ambas entidades. La compatibilidad se consigue cuando múltiples representaciones de diferentes entidades (contenido virtual y objetos reales) aseguran que el usuario interpreta correctamente la información percibida. La compatibilidad se puede aplicar a un nivel perceptivo y cognitivo. Por lo tanto, se puede evaluar mediante el estudio de las relaciones de los entornos perceptivos y procesos cognitivos correspondientes. El usuario de un sistema de RA debe ser capaz de percibir e interpretar diferentes entidades, y combinar la información recibida desde entornos perceptivamente separados. Se pueden establecer una serie de directrices para mejorar dicha compatibilidad. En (Wang et al., 2006) se explican tres directrices, y (Barfield, 1995, 556) aporta una más; estas se pueden resumir:

- Selección del **display** basándose en la **compatibilidad perceptiva**: En el nivel perceptivo, la incompatibilidad se define como el desequilibrio entre los entornos perceptivos involucrados, principalmente debido a localizaciones diferentes de ambos entornos o grados de separación entre ellos. En las aplicaciones de RA, esta incompatibilidad está directamente relacionada con el display utilizado, siendo ideal aquél que minimice el tiempo de acceso a la información o el esfuerzo de cambiar la atención de un lugar a otro.
- Selección de los **medios de representación** basándose en la **compatibilidad cognitiva**: En el nivel cognitivo, la incompatibilidad se define como la diferencia entre modelos cognitivos que están asociados a representaciones distintas concurrentes (real, simbólica, etc.) de una entidad percibida. Debido a que el usuario puede percibir formatos diferentes de información de las representaciones del mundo real y virtual, la incompatibilidad se debe principalmente a la diferencia entre ambas representaciones. En el diseño de RA, la incompatibilidad en el nivel cognitivo está directamente relacionada con la elección de los medios de representación utilizados para transmitir entidades reales, debiéndose considerar las características de las tareas reales o las

herramientas que hay que utilizar. Los medios utilizados de representación de los objetos virtuales deben estar en concordancia con sus partes reales, es decir, deben tener rasgos visuales y funciones similares.

- **Adecuar el display a la tarea** mediante **requerimientos de centricidad**: Adecuar escenarios y tareas en tipos de displays y características apropiados es esencial para el desarrollo de sistemas aumentados o mezclados. El punto de vista del usuario es importante, y debe tener una visión no obstruida, detallada y clara de la información virtual, así como de los objetos reales. Los puntos de vista en el entorno virtual pueden ser egocéntricos (punto de vista coincidente con el del usuario) o excéntricos (punto de vista desde fuera del usuario, mirando hacia él). El primero de ellos se emplea cuando los usuarios tienen que experimentar una fuerte sensación de presencia (conciencia de percepción local); el segundo de ellos se utiliza cuando es importante la localización relativa al usuario de otros objetos o de múltiples usuarios (conciencia de percepción global). De este modo, dependiendo de los requerimientos de centricidad, los usuarios pueden tener una visión egocéntrica, excéntrica, o múltiple. Dado un entorno aumentado, las tareas representativas del usuario pueden implícitamente sugerir la mezcla de centricidad, distorsión y dirección de la vista requerida. Cada una de las tareas debe tener su correspondiente dispositivo de display.
- Proceso de **selección** para **evitar el abarrotamiento** de información visualizada: Sin una clara consideración de cómo ambos tipos de información (real y virtual) deben combinarse, puede producirse un exceso de información visualizada e interferir en la tarea del usuario. Además, la información virtual no debe ocultar partes importantes de la escena real.

Para lograr la consistencia, podemos encontrar dos vertientes: hacer que la escena sea lo más real posible (Jacobs et al., 2006), o aplicar filtros tanto a la parte real como a la virtual (Haller, 2004). Así pues, la segunda directriz no implica que los elementos virtuales deban ser realistas, sino simplemente que el conjunto realidad-virtualidad tenga concordancia entre sí. Por ejemplo, si la parte virtual tiene aspecto *cartoon*, se pueden aplicar filtros a la entrada de vídeo de la parte real para que tenga el mismo carácter visual.

6.2.4. ENTRETENIMIENTO EN ENTORNOS DE RA

La industria del entretenimiento mueve miles de millones de euros al año a nivel mundial: videojuegos, cine, espectáculos, eventos deportivos, etc. Existe una gran variedad de opciones para todo tipo de público. El entretenimiento no es un campo reservado para niños; jóvenes y adultos son también grandes consumidores. Quizá por ello la finalidad de muchas aplicaciones de realidad mezclada es el entretenimiento o que, en otras, se utilice el juego como recurso de comunicación o interactividad con el usuario, haciendo que permanezca de alguna forma “enganchado” (*engaged*). En este segundo grupo podemos encontrar un buen ejemplo en la pieza *Can you see me now?*, introducida en el apartado 1.4.1, que junto a la estrategia comunicativa del juego de persecuciones por la ciudad, cuestiona también a nivel conceptual aspectos sociales y políticos de la tecnología.

Así pues, la tecnología de realidad mezclada surge como un nuevo medio capaz de ofrecer entretenimiento, y algunas investigaciones en los últimos años se han interesado en estudiar las nuevas relaciones de tipo humano que surgen en estos sistemas. Por ejemplo en (Cheek et al., 2006) se hace un estudio de los nuevos paradigmas de la interacción física y social que surgen en entornos de realidad mezclada a través del estudio de diversas piezas. Según se argumenta, la experiencia social no se basa únicamente en proveer experiencias multiusuario, donde los usuarios puedan competir y colaborar unos con otros, sino que además hay que aumentar la implicación emocional de los jugadores mediante la introducción de interacciones cara a cara y nuevas interfaces entre ellos y el entorno virtual. Por otro lado, la interacción física proporciona la seguridad psicológica que hará que los jugadores se sientan efectivamente inmersos en el mundo aumentado. La naturaleza altamente dinámica de las simulaciones por ordenador y el uso sensato de capacidades audiovisuales pueden potencialmente proveer situaciones de juego más inmersivas y satisfactorias que mediante las tecnologías tradicionales. Uno de los trabajos que se analizan es *Magic Land*, una pieza de realidad mezclada en la que los jugadores mueven físicamente diversos recipientes que contienen personajes virtuales e imágenes 3D de personas reales – digitalizadas en una habitación-croma contigua – para interactuar con ellos (Figura 152). En cuanto a la interacción social, los usuarios pueden jugar el papel del personaje principal en una historia social interactiva, estar sumergido totalmente en un mundo virtual y explorarlo mientras que, al mismo tiempo, otros jugadores pueden ver y construir el escenario de realidad mezclada con nuevos personajes para que interactúen con el personaje principal.



Figura 152. Usuario interactuando con Magic Land. En (Cheek et al., 2006).

Recientemente, y gracias a los avances en la inteligencia artificial, han surgido algunas aplicaciones de realidad mezclada en el ámbito del entretenimiento en las que se ha incluido la interacción con personajes virtuales y narrativas no lineales, como en (Charles et al., 2004; Dow et al., 2006). La narrativa interactiva consiste en la generación en tiempo real de acciones narrativas que toman en consideración las consecuencias de las intervenciones del usuario, mediante la regeneración de la historia a medida que el entorno se modifica por las acciones del usuario, basándose para ello en técnicas de inteligencia artificial. En (Charles et al., 2004) el escenario narrativo está basado en una aventura de James Bond, en la que el usuario juega el papel activo del Profesor (Figura 153). Este sistema soporta tres componentes de actuación: la actitud del protagonista mediante el reconocimiento de sus gestos; palabras habladas al conversar con los personajes virtuales, mediante el uso de sistemas de reconocimiento del habla; interacción física del usuario sobre objetos o personajes virtuales.

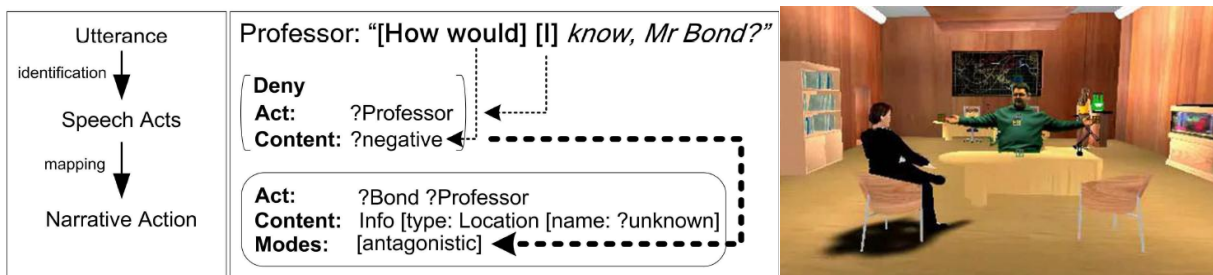


Figura 153. Secuencia del juego interactivo James Bond. En (Charles et al., 2004).

En (Dow et al., 2006) se describe un juego de realidad aumentada, basado en el videojuego Façade, en el que los jugadores se mueven físicamente por un apartamento y utilizan gestos y el habla para interactuar con dos personajes virtuales autónomos, Trip y Grace, desarrollándose una acción dramática basada computacionalmente en inteligencia artificial (Figura 154). El jugador, que ha sido invitado por sus amigos Grace y Trip para tomar unas copas en su apartamento en un

momento en el que su matrimonio tiene problemas, sin darse cuenta entrará en un juego psicológico con ellos, haciendo de conciliador o de abogado del diablo.



Figura 154. Escena del juego de realidad aumentada Façade y usuario interactuando. En (Dow et al., 2006).

6.2.5. APRENDIZAJE EN ENTORNOS DE RA

Según sugiere (Stanney, 2002, 894; Gardenfors, 2005), desde la aparición de los primeros ordenadores, las simulaciones bien diseñadas pueden permitir el acceso a experiencias de aprendizaje que no era posible realizar con los medios tradicionales. Por ejemplo, los viajes virtuales no consumen gasolina, los laboratorios simulados no explotan, las bisecciones virtuales no matan animales. Hace unos años había una tendencia a pensar que si se instalaban ordenadores suficientes en una clase, muchos de los problemas educacionales desaparecerían. Sin embargo, hoy se cuestiona el valor real de la tecnología de la informática en la educación y si la construcción de complejos sistemas virtuales o aumentados ayuda realmente en el proceso de aprendizaje.

Según (Gardenfors, 2005), no existe una teoría unificada sobre el aprendizaje que se pueda utilizar como base para la generación de entornos aumentados, y por ello surge la necesidad de revisar su compatibilidad con distintas metodologías, así como evaluar los posibles tipos de implementación para asegurar los objetivos educativos. En (Stanney, 2002, 895) se proponen varias metodologías, que tienen su fundamento en teorías establecidas de modos de aprendizaje:

- Exploración de mundos preconstruidos como actividades de aprendizaje constructivistas.
- Creación o modificación de mundos como actividad de aprendizaje construccionista.

- Role-play por múltiples participantes como una actividad de aprendizaje situado.

En (Stanney, 2002, 896-897) se explican brevemente dichas teorías:

- **Constructivismo:** Se ha establecido que los individuos construyen activamente modelos internos del mundo antes que aceptar datos de forma pasiva, construyéndose activamente nuevos conocimientos a medida que se interactúa con el entorno. En la teoría constructivista, todo el conocimiento útil es un conocimiento procedural. Todo lo que un individuo lee, ve, oye, siente o toca se contrasta con su conocimiento anterior y si encaja dentro del mundo que hay en su mente, puede formar un nuevo conocimiento. Este conocimiento se refuerza si puede aplicarse en el entorno que le rodea. El individuo no es sólo un banco de memoria que absorbe información pasivamente, ni se le puede 'transmitir' conocimiento sólo leyendo algo o escuchando a alguien. Esto no significa que no pueda aprender nada leyendo una página Web o asistiendo a una lección, sólo indica que se trata más de un proceso de interpretación que de una transferencia de información de un cerebro a otro.
- **Construccionismo:** El construccionismo explica que el aprendizaje es particularmente efectivo cuando se construye algo que debe llegar a otros. Esto puede ir desde una frase hablada o enviar un mensaje en Internet, a artefactos más complejos como una pintura, una casa o un paquete de software, se aprende construyendo activamente los significados. El construccionismo es una extensión del constructivismo, en el cual los individuos tienen que crear artefactos – preferiblemente interactivos – para integrar completamente su modelo del mundo.
- **Aprendizaje situado:** El aprendizaje situado, se basa en la idea de que el aprendizaje se estimula o impulsa cuando se le vincula a circunstancias basadas en historias personales. Por ejemplo, en lugar de calcular el área de un triángulo, a los estudiantes se les pide que ayuden a los personajes de una historia a resolver un problema complejo, como la construcción de una casa. Evidentemente, en dicha historia aparecerán muchos problemas individuales y su solución estará motivada por las necesidades naturales del personaje de la historia. El **role-play** lleva el aprendizaje situado un paso más allá involucrando múltiples estudiantes que se identifican con personajes específicos de la historia. Las situaciones simuladas pueden ser competitivas, cooperativas o ambas. Al

introducir varios estudiantes surgen diversos problemas, tales como la sincronización, el reparto de tareas, la disciplina, la capacidad de resolución etc., pero ofrece como potencial el desarrollo de la interacción social.

El paso de la transferencia del conocimiento desde un sistema altamente guiado a un proceso de aprendizaje social abierto, basado en actividades, es también apropiado en el desarrollo de tecnología educacional, en el modo en que los nuevos recursos mediáticos se forman para vehicular y apoyar nuevos métodos de aprendizaje. El constructivismo ha emergido en la última década como una alternativa pedagógica íntimamente relacionada con los avances en la educación tecnológica, proporcionando a los tecnólogos los fundamentos teóricos para el desarrollo de entornos virtuales de aprendizaje (*VLE, virtual learning environments*), (Roussou, 2004).

6.2.5.1. APRENDER HACIENDO

Se pueden crear aplicaciones de RA en las que los participantes puedan de manera activa construir sus propios conocimientos, probando ideas y conceptos basados en experiencias y conocimientos personales preexistentes, a la vez que involucrando a los usuarios en el diseño y construcción de artefactos. El constructivismo se adopta como base para el desarrollo de entornos altamente interactivos y participativos, donde el usuario es capaz de modificar, construir, probar ideas, e involucrarse activamente en la resolución de un problema (Roussou, 2004).

Los museos son un buen ejemplo de esta práctica, mediante el desarrollo de exhibiciones en las que algunas piezas se pueden tocar y manipular. Conforme los museos se vuelven más abiertos y vinculados con tecnologías interactivas, también emerge una nueva concepción de la audiencia como participantes activos o incluso como creadores (Roussou, 2004). Ejemplos de instalaciones museísticas basadas en entornos aumentados los podemos encontrar en (Hall et al., 2002; Bimber et al., 2003; Brown et al., 2003; Hatala et al., 2004; Sparacino, 2004; Wojciechowski et al., 2004; Woods et al., 2004; Bimber et al., 2005a). En (Woods et al., 2004) se describe *The S.O.L.A.R System* (Figura 155-a), una pieza educacional de realidad aumentada creada para el TheManawa Science Centre (Nueva Zelanda). Mediante esta pieza se pretende enseñar los visitantes la posición que ocupa cada planeta en el sistema solar. El sistema contiene un total de nueve marcas de RA que el visitante puede distribuir sobre una mesa en la que están dibujadas las órbitas de los planetas. Cuando las

marcas se encuentran en la posición correcta, se ve una animación en 3D en la que los planetas empiezan a orbitar alrededor del sol. En (Bimber et al., 2005a) se utiliza un sistema para integrar diversas proyecciones sobre elementos pictóricos (Figura 155-b). Los visitantes disponen de una representación tangible del cuadro, sobre la que pueden realizar varias operaciones mediante el tacto, como ampliar, rayos-X, desenfocar, desaturar, resaltar partes, etc. El resultado se proyecta sobre el cuadro original colgado en la pared, gracias a la utilización de una lámina sobre éste que tiene dos propiedades: es completamente transparente y difunde una fracción de la luz proyectada sobre él.



Figura 155. Ejemplos de instalaciones de RA en museos: a) *The S.O.L.A.R. System*, en (Woods et al., 2004); b) Integración de proyecciones sobre cuadros, en (Bimber et al., 2005a).

La tecnología de RA es capaz de proporcionar un amplio rango de posibilidades para este tipo de interacción y permite la participación activa en la formación de contenido. De este modo, se convierte en un medio adecuado para su uso en escuelas, museos y centros de entretenimiento.

6.2.5.2. APRENDER JUGANDO: EDUTAINMENT

El vocablo *edutainment* nace de la combinación de educación y entretenimiento (*education + entertainment*). El juego puede unir imaginación e intelecto en más de una manera, y ayudar a los usuarios a descubrir cosas en su propio espacio y a su modo. Es indudable el hecho de que la actividad favorita de los niños es el juego, y por ello se supone que el aprendizaje se dará mejor en entornos de diversión, retos y variedad. Sin embargo, según (Roussou, 2004) si el entorno establecido tiene

“demasiada diversión” el aprendizaje no está garantizado, por lo que además de diversión, éste debe incluir tareas significativas para que el usuario tome el aprendizaje de manera consciente y aprenda a resolver tareas difíciles.

La creación de experiencias interactivas empieza a converger con el impulso de la industria del entretenimiento para el desarrollo de exposiciones museísticas dentro del campo del *edutainment*. Los museos científicos y para niños ofrecen programas basados en el juego, descubrimiento y compromiso como herramienta principal para el aprendizaje. Existen muchos ejemplos de museos en los que las piezas se pueden tocar, manipular y experimentar, incentivando a los espectadores a descubrir y aprender. Incluso en algunos museos de arte se incluyen zonas de juego físicas o basadas en páginas Web para sus audiencias más jóvenes (Roussou, 2004).

Aunque todavía existe una división entre software para el aprendizaje y software específicamente dirigido al entretenimiento (los videojuegos), el diseño de software educacional intenta incluir muchos de los recursos que caracterizan el diseño de videojuegos, tales como la finalidad de alcanzar un objetivo, la habilidad de personalizar la experiencia, el incremento de dificultad al avanzar el juego, etc. Existen diversas teorías que unen las posibilidades de aprendizaje aportadas por los videojuegos a la motivación y el compromiso. Los primeros cuatro tipos de motivación intrínsecas (desafío, curiosidad, control y fantasía) pueden estar presentes en cualquier situación de aprendizaje, incluso en aquellas que involucran a una sola persona. Los otros elementos de motivación intrínseca (competición, cooperación y reconocimiento) se categorizan como motivaciones interpersonales, ya que se basan en la existencia de otros jugadores. La interacción también se puede considerar como una propiedad intrínseca del diseño de un juego. Los juegos mantienen la atención de los usuarios mediante características interactivas, tanto si su intención es avanzar en la historia del juego como si permiten el desarrollo y exploración de interacciones y relaciones sociales.

Por lo tanto, existe una conexión entre interactividad, compromiso y aprendizaje. Juntos pueden formar la base para el desarrollo de espacios de juego virtuales o aumentados, que permitan a los niños involucrarse en un juego constructivo y creativo, consiguiendo la mezcla ideal entre los valores educativos y recreativos (Roussou, 2004). Un claro ejemplo es el sistema KidsRoom (Bobick et al., 1997; Bobick et al., 1999), explicado en el apartado 1.4.4, donde la espontaneidad y naturaleza colaborativa de un lugar físico de juego se unen a una historia fantástica.

6.2.6. EL COMPROMISO ARTÍSTICO EN ENTORNOS DE RA

El arte, en general, analiza y cuestiona muchos aspectos de la vida cotidiana, las relaciones sociales, actitudes o comportamientos del momento presente (la época) en que se desarrolla; las artes visuales dirigen este cuestionamiento principalmente a través de la imagen y en el momento actual las tecnologías de la imagen están incidiendo de forma evidente en los modos de visualización de nuestra época, por ello las prácticas artísticas vinculadas a los nuevos medios tecnológicos tienen una presencia cada vez más destacada en el contexto cultural y social. La forma en que el arte incorpora estos medios es creativa y crítica, indagando en las fronteras o límites de lo establecido por los recursos comerciales para romper convenciones, buscando siempre otras formas de uso que hagan visible aquello que física o socialmente limita nuestra percepción.

En el ámbito de la realidad aumentada también existen voces críticas, y algunos autores argumentan que cualquier tipo de medio capaz de transportarnos a otra realidad puede considerarse realidad mezclada. Por ejemplo, el cine es uno de los medios tradicionales más inmersivos que existen en la actualidad, utilizando para ello las convenciones del lenguaje cinematográfico elaborado durante décadas: tipos de planos, secuencias, diálogos, narrativa, audio, lugares, vestuario, etc. De igual forma, la lectura de un libro de aventuras o ciencia ficción también puede transportar a su lector a un mundo ficticio, del que él mismo puede ser protagonista. Incluso se puede argumentar también que cualquier tipo de medio convencional puede ser considerado como realidad mezclada, ya que todas las experiencias son físicas e interactúan con el mundo real de alguna manera. En (Stapleton et al., 2002) se recogen algunas de estas ideas. Sus autores introducen el “continuo de los medios”, estableciendo una transición gradual entre el mundo real y el mundo virtual. A continuación añaden que, para que la realidad mezclada llegue a ser un “medio” convincente, su descripción no puede basarse únicamente en la tecnología, sino que debe considerar el continuo de los medios, elaborando para ello un nuevo espectro realidad-virtualidad en el que se incluye la imaginación como componente fundamental (Figura 156). En este espectro tienen cabida tanto las formas nuevas como las antiguas de realidad mezclada. Las formas más tradicionales de realidad mezclada, como un libro o un antiguo parque temático, no utilizaban la tecnología de hoy en día, sino que con el mundo físico y la imaginación era suficiente. Otros medios, como el cine, se basan en el compendio de tecnología y convenciones artísticas para despertar la imaginación y transformar el mundo real. De este modo, los medios

tradicionales pasivos, como una novela, están situados en el continuo más cerca de la imaginación, mientras que los medios tradicionales activos, como una película, están más cerca de la tecnología. En la parte superior del espectro están situadas las tecnologías que tratan de emular la realidad con diversos grados de realidad virtual (el continuo de Milgram). Según sus autores, el arte de la realidad mezclada, como una forma de medios compilados y expresivos, estará en la habilidad de sacar ventaja del espectro de realidades, por lo que aparece centrado en el diagrama. Así pues, la realidad mezclada se desarrollará mediante la convergencia de todas las formas de arte y tecnología de los medios.

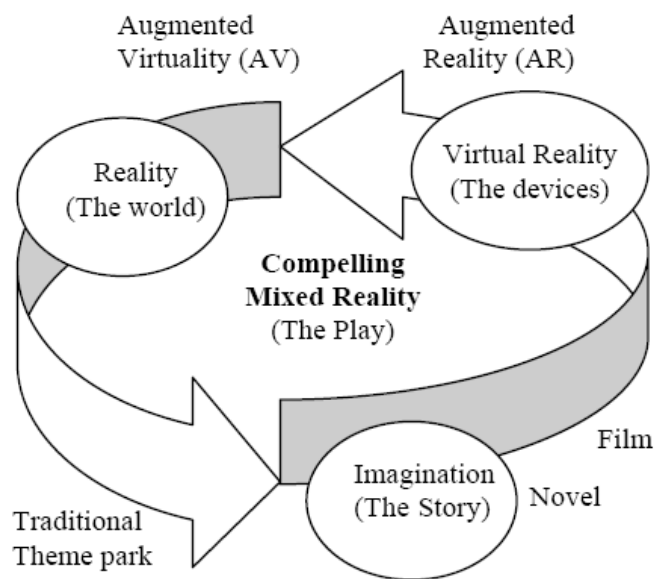


Figura 156. Continuo de realidad mezclada. En (Stapleton et al., 2002).

Algunos ejemplos de RM en el campo del arte se han descrito en el apartado 1.4.1. Podemos destacar *Can you see me now?*, aplicación de RM móvil en la que no se produce una mezcla de vídeo de los mundos reales y virtuales; se puede decir que la mezcla de ambos mundos se realiza a nivel cognitivo.

BLOQUE II: ENSAYOS

7. USUARIO AUMENTADO

Usuario Aumentado es un sistema novel de RA en el que se utiliza la metáfora del espejo mágico para alterar la apariencia física que el usuario percibe de sí mismo, proponiendo un cambio de identidad que cuestiona su personalidad y provoca en él una alteración de su comportamiento, todo ello a partir de un juego de máscaras y disfraces virtuales. Este cambio de imagen se percibe a través de una o diversas pantallas que funcionan a modo de espejos mágicos, los cuales transforman el espacio de representación típico de los sistemas tradicionales de RA, surgiendo una serie de dualidades (sujeto/objeto y dentro/fuera). Este ensayo, incluyendo sus diferentes fases, se ha realizado con la colaboración del doctor Francisco Giner.

Motivación

Las aplicaciones de realidad aumentada generalmente requieren la utilización de un material bastante caro y frágil, como por ejemplo los HMD o los ordenadores portátiles a modo de vestimenta. Además, para la adquisición de la orientación externa de la cámara en tiempo real, también se emplean distintos sensores (como GPS, sensores magnéticos, inerciales, etc.) con su coste adicional. Todos estos factores hacen que muchas aplicaciones de RA sean solamente accesibles para un reducido número de usuarios. Es por ello que en este ensayo se pretende elaborar un sistema novedoso de realidad aumentada que sea económico, ergonómico y fácilmente utilizable por cualquier persona. Esto se ha logrado, en primer lugar, mediante la utilización de

librerías de libre distribución (ARToolKit) para el cálculo de la orientación externa de la cámara; en segundo lugar, con el propio diseño de marcas de RA para vestir (ligeras y adaptables para cualquier usuario); y finalmente, con un sistema que permite una interacción persona-computador natural, ya que no necesita de un aprendizaje previo.

Marco temporal

Este ensayo tiene una duración dilatada en el tiempo, comenzando en el año 2004 y finalizando a principios del 2007; consta de múltiples fases, fruto de la evolución de la pieza a partir de diversas consideraciones por parte de los usuarios, y de un interés propio por investigar nuevas posibilidades. Así pues, surgen diferentes temáticas que abarcan distintas áreas, como el entretenimiento, el edutainment o el arte.

7.1. PLANTEAMIENTO

“El arte último de la representación es interactivo, inmersivo, viviente, transporta al observador a otro mundo, lo introduce en su lógica; mezcla las referencias identitarias ya que, al otro lado del espejo, un visitante no tiene necesariamente la misma razón de ser” (Courchesne, 2004).

En este ensayo se introduce el concepto de **metáfora del espejo mágico**, en contraposición a la **metáfora de la lente mágica**, muy utilizada en sistemas de realidad aumentada. Basándose en la distinción entre lente y espejo, se puede decir que, dependiendo del tipo de display utilizado y de la disposición espacial de la cámara, la realidad aumentada se puede visualizar a través de una ventana a un mundo ficticio o como un reflejo de éste. La metáfora de la lente mágica se hace más evidente en los dispositivos para llevar en la mano, como PDAs o teléfonos móviles con cámaras integradas, que se asemejan a una lente a través de la cual se descubre una nueva realidad (la aumentada). En *Usuario Aumentado* se muestra esa nueva realidad como un reflejo alterado del entorno real, invirtiéndose para ello la disposición espacial de la cámara que muestra al usuario como parte fundamental del entorno aumentado, incrementando también la sensación de presencia/inmersión.

La utilización de la metáfora del espejo mágico no es banal. Como bien dice (Domènech, 2006, 23): *“El espejo es símbolo de la imaginación, como reflejo de las formas del mundo y símbolo del pensamiento, como elemento de análisis del universo, de auto contemplación. Mucho se ha escrito sobre el espejo y su reflejo, lo que vemos y sentimos ante él, una imagen idealizada, imaginada, una máscara de nosotros mismos, nuestro doble”*. Por otro lado, al referirse a los espejos mágicos en (Revilla, 2003, 167): *“... los espejos mágicos, constantes en muchas culturas: se trata de los que calan más allá de las apariencias y desechando éstas reflejan precisamente lo que ellas ocultan... Ciertos espejos mágicos han poseído capacidades adivinatorias, suscitando por ejemplo la imagen de individuos que existirán en el futuro, para prever su actuación”*. *Usuario Aumentado* es el reflejo de un mundo ficticio, en el que la realidad física del usuario que se ve a sí mismo es alterada por medio del espejo mágico/aumentado. El juego – el mirarse a uno mismo alterado o disfrazado – constituye el elemento atrayente de la pieza. Mediante el disfraz somos capaces de inhibirnos y expresar emociones y sentimientos que de otra manera permanecerían ocultos. El yo interno se revela y se muestra a través del espejo aumentado.

Al introducir la metáfora del espejo mágico, se ha alterado la relación espacial de los dispositivos con respecto a los sistemas más tradicionales de RA (Figura 157). En el nuevo sistema establecido surgen una serie de dualidades a nivel conceptual (sujeto/objeto) y físico (dentro/fuera):

- **Sujeto/objeto:** El usuario es parte del objeto. En las aplicaciones de RA basadas en el reconocimiento de marcas, las marcas se sitúan en el espacio adheridas a los objetos que van a ser aumentados, siendo el usuario externo a ellos. En nuestro caso, el usuario lleva puestas las marcas, por lo que es él mismo el elemento aumentado, formado parte del objeto.
- **Dentro/fuera:** Incorporación del usuario dentro del espacio de representación. Normalmente, en las aplicaciones de RA el usuario lleva un HMD con una cámara adherida a éste, haciendo coincidir (en lo posible) el punto de vista de la cámara con la del usuario. En nuestra aplicación, el usuario ya no lleva el HMD, y la cámara está fija en el entorno físico, dirigida hacia el lugar que ocupa el usuario, por lo cual el espacio de observación se transforma en espacio de representación. En las aplicaciones de RA basadas en la metáfora de la lente mágica, el usuario puede llevar o no un HMD, pero sigue estando fuera del espacio de representación.

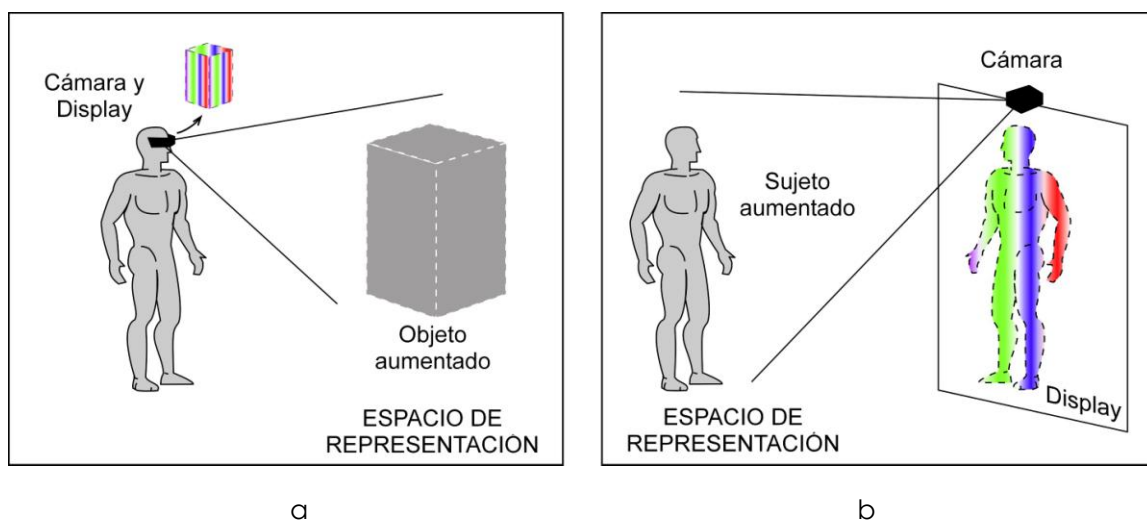


Figura 157. Entornos de RA: a) Sistema tradicional de RA, b) Sistema de *Usuario Aumentado*.

El espectador/usuario pasa a ser la parte central de la pieza, al ser él mismo el objeto aumentado. De hecho, no hay pieza sin usuario, ya que nada ocurre si nadie se sitúa en el espacio de representación. Las pantallas reflejan el blanco de la pared de enfrente como escena vacía hasta que el espectador entra en ese espacio de

representación y todo cambia: el usuario se convierte en una especie de humanoide, un híbrido entre lo real y lo virtual. La interacción se produce con sus propios movimientos, siendo éste a la vez espectador, objeto e interfaz. El usuario es la parte fundamental de la pieza por ser él mismo el objeto aumentado, el cual está en continua transformación.

Este ensayo se aproxima así a la contraposición: *“ser es percibir (esse est percipere) / ser es ser percibido (esse est percipi)”*¹⁷, afectando también en cuestiones de identidad, avatar y juego. Es decir, la imagen proyectada actúa como un espejo cuyo reflejo es visto tanto por el usuario como por otros espectadores (nos observamos y somos observados), propiciando un entorno colaborativo y permitiendo una experiencia compartida, multiusuario. Además, en algunas fases de este ensayo se utilizan varias pantallas para permitir que el usuario adopte distintas apariencias dependiendo de en cual de ellas se mire. Así pues, el usuario adquiere tantas apariencias o identidades como espejos diferentes lo observen. Yo sigo siendo yo, pero me reconozco a través del espejo como un ser diferente. Esto puede condicionar temporalmente el comportamiento, al tratar de moverse y actuar de acuerdo a esa nueva identidad adquirida. En este momento, la identidad constituye una dualidad. En tanto uno se va descubriendo a través de los diferentes espejos, la dualidad se convierte en pluralidad. Estos espejos nos recuerdan las reflexiones de Ascott sobre el individuo mediatizado: *“... cada uno está hecho de muchos individuos, de una colección completa de yos... Todos somos interfaz. Todos somos seres mediatizados y ampliados por los ordenadores. Estas nuevas formas de conceptualizar y percibir la realidad implican más que cierta clase de cambio cualitativo en nuestro modo de ver, pensar y actuar en el mundo. Constituyen un cambio cualitativo en nuestro ser, una facultad completamente nueva, la facultad postbiológica de la cibercepción”* (Ascott, 1998, 95).

7.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO

Siguiendo las aportaciones del bloque I, se puede elaborar un marco general de este ensayo:

- Ubicación dentro del continuo de Milgram (apartado 1.1.1):
 - Realidad aumentada.

¹⁷ Máxima del obispo y filósofo irlandés George Berkeley (1685-1753).

- Clasificación (apartado 1.2.2):
 - Según el entorno físico en el que se desarrolla la aplicación:
 - Cerrada.
 - Según la extensión que abarca:
 - Local.
 - Según la movilidad de los dispositivos de registro y/o displays:
 - Espacial.
 - Según los usuarios que simultáneamente pueden interactuar con el sistema:
 - Colaborativa.
 - Según el tipo de colaboración establecida:
 - Presencial.
 - Clasificación según Mackay (apartado 1.2.2):
 - Aumentar el objeto físico.

- Campo de aplicación (apartado 1.4):
 - Arte.
 - Entretenimiento.
 - Edutainment.

- Sistema de coordenadas terreno/objeto (apartado 2.1.1):
 - Centrado en objetos.

- Orientación espacial (apartado 2.1.5):
 - Método indirecto.

- Dispositivos de registro / controladores (capítulo 3):
 - Sensores ópticos: Marcas pasivas.

- Displays (capítulo 4):
 - Displays espaciales: Proyector.

- Software (capítulo 5):
 - Basado en patrones de marcas planas: ARToolKit.

- Forma de interactuar (apartado 6.1.1):
 - Manipulación de objetos virtuales/reales.
 - Interacción con otros usuarios:
 - Colaboración presencial/directa.

7.1.2. REFERENTES

Existen muchos ejemplos de la utilización de la metáfora del espejo en los campos del arte, entretenimiento o edutainment. En el campo del arte los espejos han sido utilizados por muchos autores como parte principal de sus piezas, como en las obras de Christian Megert, Robert Morris, Isa Genzken o John Armleder. Cabe sin embargo destacar diversas obras de Michelangelo Pistoletto (Figura 158), ya que inserta en sus espejos elementos ficticios, las imágenes de personas que no están físicamente presentes.



Figura 158. Exposición "Michelangelo Pistoletto: A Reflected World" (1966) en Walker Art Center, Minneapolis. La imagen de la izquierda corresponde a la pieza *Girls on a Balcony* (1962-1964). En (Walker Art Center, 2001; Pistoletto, 2008).

Dentro del arte digital, cabe destacar la pieza *Hallucination* (1988-90) de Jim Campbell, en la que una imagen de vídeo actúa como un espejo en el cual los espectadores se ven a sí mismos envueltos en llamas, acompañados también de un audio de llamas (Figura 159). El trabajo de Campbell juega con la alucinación de la realidad y la interactividad en la imagen de la televisión, diluyendo las fronteras entre una imagen viva y una construida y cuestionando el nivel de control del usuario (Paul, 2003, 103).



Figura 159. Imágenes de *Hallucination* (1988-90), de Jim Campbell.

En el mundo del entretenimiento hay que mencionar *la casa de los espejos*, típica atracción de ferias, en los que se establece un juego laberíntico a partir de espejos que deforman la imagen de los visitantes. En el campo de los videojuegos nos encontramos con la tecnología de *EyeToy Play 2* (2003), un dispositivo para la videoconsola PlayStation 2 que incluye una cámara Web que captura la imagen del usuario, por lo que éste ve la imagen de sí mismo dentro del entorno de los videojuegos (Figura 160). La interacción con los elementos gráficos generados por ordenador se produce mediante el simple movimiento del usuario. Hay que resaltar que dichos elementos gráficos son de geometría plana, y el registro del usuario se produce de igual forma en un espacio bidimensional, por lo que no se puede considerar como una aplicación de RA ortodoxa (ver apartado 1.2.1). Sin embargo, posteriormente surgen nuevas versiones con esta tecnología en las que sí se consideran objetos virtuales tridimensionales, aunque el registro de los movimientos del usuario sigue siendo bidimensional.

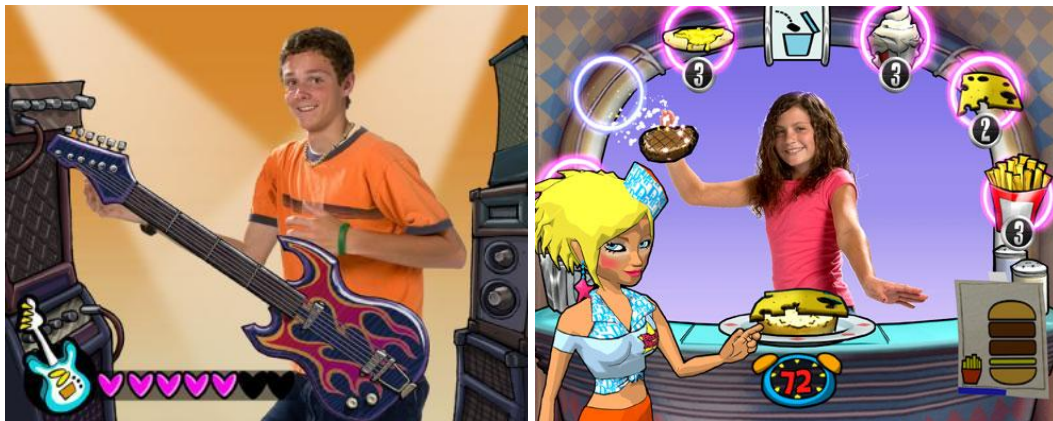


Figura 160. Diversos videojuegos para EyeToy Play 2.

Dentro de las aplicaciones de realidad aumentada, parece ser que la utilización de la metáfora del espejo mágico para alterar la apariencia del usuario es una característica novel de nuestro sistema. Recordemos que la RA nació con los HMD, por lo que necesariamente los objetos aumentados eran externos al observador. Sin embargo, posteriormente a la primera versión de *Usuario Aumentado* publicada en (Giner Martínez et al., 2005), se han encontrado otros referentes que utilizan sistemas similares, como en (ARTag, 2007), (Figura 161).



Figura 161. Usuarios interactuando con un sistema de RA basado en la metáfora del espejo mágico. En (ARTag, 2007).

También hay que destacar las piezas *Attack and Swap* (2006) de Benedikt Gross (Figura 162-a) y *AR Magic System* (2007) de Diego Díaz y Clara Boj (Figura 162-b), en las que se establece un juego visual de intercambio de caras de los usuarios a partir del reconocimiento de éstas en tiempo real. Aunque estas piezas también constituyen técnicamente sistemas planos, guardan relación con *Usuario Aumentado* en tanto que se establece un juego de cambio de identidades a partir de una transformación del rostro de los usuarios.

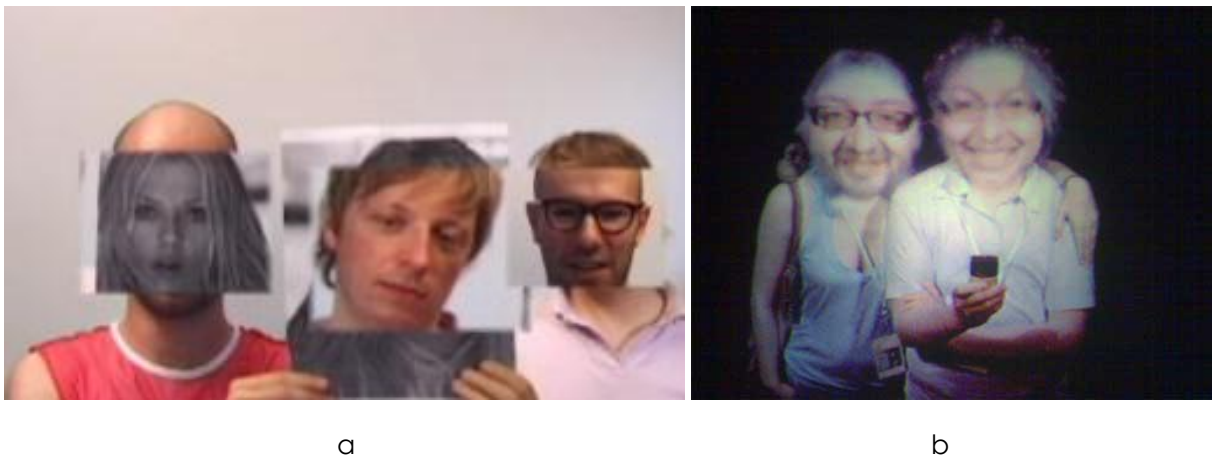


Figura 162. Intercambio de la imagen de la cara: a) *Attack and Swap* (2006) de Benedikt Gross; b) *AR Magic System* (2007) de Diego Díaz y Clara Boj.

7.2. DESCRIPCIÓN TÉCNICA

7.2.1. COMPONENTES DE LA APLICACIÓN

7.2.1.1. DISPOSITIVOS FÍSICOS

El sistema físico de *Usuario Aumentado* consiste en una pantalla de retroproyección a modo de espejo. Sobre ésta se sitúa una cámara Web de 640x480 de resolución, que apunta hacia el espacio de representación (Figura 163). Al entrar en el campo de visión de la cámara, la imagen de los usuarios/audiencia se visualiza en dicha pantalla gracias a un proyector multimedia que se encuentra situado detrás de ésta (en el *backstage*).

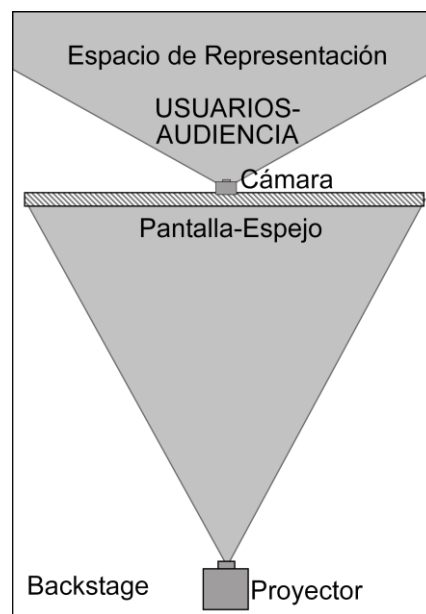


Figura 163. Esquema básico de la distribución espacial del sistema de *Usuario Aumentado*.

Para el registro de la orientación externa de la cámara se han empleado las librerías de ARToolkit; las marcas empleadas en este caso las llevan puestas los usuarios a modo de vestimenta. De este modo, a la imagen de los usuarios en tiempo real se superponen diversos elementos virtuales que van asociados a cada una de las marcas, dando lugar a una representación del usuario en la pantalla-espejo a modo de híbrido entre lo real y lo virtual. En algunas versiones de *Usuario Aumentado*, este sistema de pantallas se triplica.

También se requiere la disposición de unos focos para controlar las condiciones de luminosidad del entorno, de tal forma que la parte posterior de la pantalla quede en penumbra, mientras que el espacio de representación quede iluminado, evitando luces directas que puedan ocasionar reflejos o zonas de sombras. El fondo del espacio de representación debe ser uniforme (por ejemplo, de color blanco), para evitar que la imagen capturada por la cámara tenga información innecesaria, lo que puede ocasionar que el sistema reconozca falsas marcas. En la Figura 164 se muestran de forma esquemática todos estos elementos.

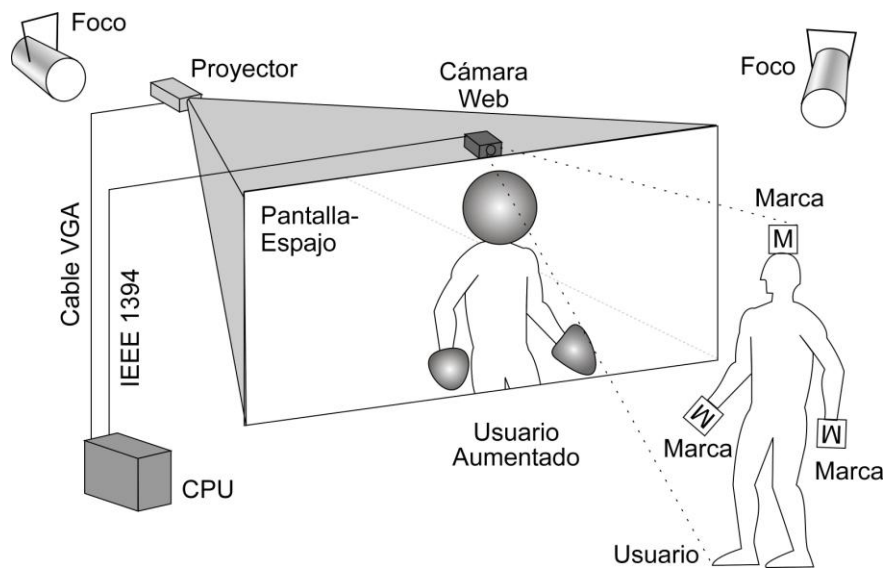


Figura 164. Componentes físicos de *Usuario Aumentado*.

Una de las ventajas que presenta el empleo de ARToolKit con respecto a otros sistemas de registro es el factor económico así como su perdurabilidad, ya que no necesita apenas mantenimiento y las marcas no se rompen fácilmente. Además, su bajo peso y fácil manejabilidad hacen que este sistema tan sencillo sea utilizable por todo tipo de usuarios independientemente de la edad o condición física. Si bien es cierto que una de las desventajas de utilizar sistemas de marcas para el registro es la misma presencia de las marcas en el entorno – que tiende a ser innatural – en *Usuario Aumentado* esta desventaja se hace menos evidente al estar las marcas integradas en el usuario a modo de vestimenta y no ser visibles en la imagen de la pantalla-espejo por estar ocultas bajo los objetos virtuales, por lo que el usuario no repara en ellas en el momento de la interacción, ya que su atención recae sobre la imagen de sí mismo transformado.

Como ya se ha apuntado en apartados anteriores, las librerías ARToolKit necesitan visualizar una o varias marcas para realizar las operaciones de posicionamiento y superposición de elementos virtuales solo que, en nuestro sistema, las marcas forman

parte del usuario. Es por ello que se nos planteó la necesidad de crear una serie de marcas que el usuario pudiera ponerse en diferentes partes de su cuerpo (a partir de ahora las llamaremos marcas para llevar puestas), a modo de vestimenta, que fuesen cómodas de llevar para que no interfirieran en su libertad de movimiento y, a su vez, que fuesen constantemente visibles por la cámara. En este sentido, se diseñaron y probaron diferentes prototipos, pudiéndose utilizar cada uno de ellos en condiciones diferentes.

Nuestro sistema de marcas para vestir constituye un sistema económico y ligero. Ambos hechos permiten que sea fácilmente llevado por todo tipo de usuarios, incluyendo niños, sin el temor de romper las usualmente caras interfaces de visión o dispositivos de registro (sensores inerciales, GPS, etc.). A modo cronológico, los prototipos diseñados son:

- Marca libre;
- Gorro con prisma ajustable;
- Prisma con diadema;
- Marcas ajustables para las manos;
- Prisma con diadema iluminado;
- Prisma con diadema mejorada;
- Marcas a modo de guantes;
- Marcas a modo de peto;

En todos los casos, la parte de color negro de las marcas se realizó con una tela aterciopelada adhesiva, para conseguir una superficie totalmente mate y evitar así los no deseados reflejos lumínicos, que pueden llevar a una identificación errónea de las marcas. Según la evolución de *Usuario Aumentado* se han utilizado unas marcas u otras; es por ello que éstas se explican con detalle en el apartado 7.3 y subapartados, junto con las ventajas e inconvenientes encontrados.

7.2.1.2. ELEMENTOS VIRTUALES

Los elementos utilizados para aumentar a los usuarios son modelos tridimensionales. En ARToolKit existe la posibilidad de modelar objetos 3D a partir de OpenGL. Éste es, sin embargo, un lenguaje de programación por código, por lo que no suele utilizarse por diseñadores gráficos. Sin embargo, las librerías ARToolKit disponen de un módulo en el que se integran las librerías OpenVRML, permitiendo la utilización de objetos 3D con el

formato VRML. Este módulo es el que se ha utilizado en *Usuario Aumentado*. Para ello, en primer lugar se han modelado los objetos 3D con el programa 3D Studio Max, añadiendo en algunos casos texturas fotorrealistas y/o pequeñas animaciones; a continuación se han exportado a formato VRML. Cada elemento 3D va asociado a una marca o a un conjunto de marcas, y visualmente coinciden con alguna parte del cuerpo del usuario (cabeza, manos, torso o cuerpo entero). Debido a la gran variedad de modelos utilizados, éstos se mostrarán en las diferentes fases del ensayo (apartado 7.3 y subapartados).

7.2.2. DESARROLLO, PROCESO

En las primeras fases de *Usuario Aumentado* se ha trabajado con la versión 2.65 de ARToolkit y con el ejemplo *simpleVRML.exe*. Este archivo se compiló con Microsoft Visual C++ 6.0. En primer lugar la cámara debe calibrarse mediante el ejecutable *calib_camera2.exe*; el resultado de la calibración se guarda en el fichero *camera_para.dat*. La información sobre las marcas que intervienen en la aplicación, así como el archivo que contiene diversos parámetros del objeto VRML vinculado a la marca, se encuentran en el fichero *vrmldata*. Los archivos que contienen la ruta de los VRML y diversa información, como las condiciones de luminosidad, se encuentran en ficheros con la extensión *dat*. De este modo, para una única aplicación habrá únicamente un fichero *vrmldata*, y tantos archivos *dat* como modelos virtuales vinculados a marcas. A continuación se muestra el fichero *vrmldata* y uno de los *dat* utilizados en la primera versión del *Usuario Aumentado*.

Contenido del fichero *vrmldata*:

```
7 # número de marcas para reconocer

# marca 1
VRML Wrl/cabeza0.dat # ruta y nombre del archivo que contiene los parámetros del VRML
Data/patt.ter05b # ruta y nombre de la marca a la que se asigna el modelo VRML
83.0 # ancho de la marca (en mm)
0.0 0.0 # desplazamientos en x e y del modelo VRML con respecto al centro de la marca

# marca 2
VRML Wrl/cabeza1.dat
Data/patt.ter06b
83.0
0.0 0.0
```

```

...

# marca 7
VRML  Wrl/manoDb.dat
Data/patt.ter02n
64.0
0.0 0.0

```

Contenido del fichero *cabeza0.dat*:

```

Wrl/cabeza.WRL      # ruta y nombre del archivo VRML
0.0 -200.0 -30.0    # Translación (en mm) en x,y,z desde el centro de la marca
0.0 0.0 0.0 0.0    # Ángulo de rotación y eje, por ejemplo 90.0 1.0 0.0 0.0
0.9 0.9 0.9        # Escala en x,y,z

LIGHT POINT        # Luz puntual
0.3                # Ambiente
1.0 0.0 0.0        # Atenuación
1.0 1.0 1.0        # RGB
1.0                # Intensidad
0.0 9.0 9.0        # Localización
2.0                # Radio

```

En la última versión de *Usuario Aumentado* se ha trabajado con la extensión 2.71.3, con el ejemplo *simpleVRMLd.exe*. Este ejemplo se ha compilado con Microsoft Visual C++ .NET. Aunque el funcionamiento es bastante similar a la versión anteriormente descrita, en este caso el fichero donde se guarda la información de las marcas se denomina *object_data_vrml*. Además, hay otro fichero denominado *VDM_camera.xml* con el que se pueden controlar algunos parámetros de la cámara, que en este caso se han utilizado para voltear la imagen capturada por ésta a través de su eje vertical, con el fin de conseguir el efecto de espejo. En anteriores versiones, el efecto de espejo lo hacíamos directamente volteando la imagen desde el proyector, en aquellos casos en los que no se puede hacer desde el comando de control de la propia cámara (como es en la cámara Fire-i).

Contenido del archivo *VDM_camera.xml*:

```

<?xml version="1.0" encoding="UTF-8"?>
<!--Sample XML file generated by XMLSpy v2005 rel. 3 U (http://www.altova.com)-->
<dsvl_input xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="C:\Documents and Settings\Thomas\My Documents\
projects\ARToolKit & DSVideoLib\ARVideoLib\DsVideoLib\DsVideoLib.xsd">

```

```
<camera show_format_dialog="true" friendly_name="PGR">
  <pixel_format>
    <RGB32 flip_h="false" flip_v="true"/>
  </pixel_format>
</camera>
</dsvl_input>
```

7.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS

En todas las versiones de *Usuario Aumentado* se ha trabajado con cámaras Web con 640x480 de resolución, bien mediante conexión FireWire o USB 2.0. En las primeras versiones, debido a la capacidad limitada de los ordenadores utilizados, el renderizado se hacía a 15 fps, mientras que en las últimas versiones utilizamos renderizados de 30 fps.

La iluminación es uno de los factores que más ha condicionado la pieza de *Usuario Aumentado*. En las primeras versiones se utilizaron focos de luz blanca ubicados de tal forma que no incidiese la luz directamente sobre la parte frontal de la pantalla, al mismo tiempo que proporcionando la suficiente luminosidad en el espacio de representación para que las marcas fuesen visibles a la cámara, evitando, en la medida de lo posible, la aparición de sombras; en una de las versiones (ver apartado 7.5.3) se utilizó luz negra. Para evitar reflejos en las marcas, éstas se han elaborado con material mate; la parte de color blanca es de cartón pluma (excepto en una de las versiones, ver apartado 7.5.3), mientras que la parte negra está hecha con tela adhesiva de terciopelo.

Debido a limitaciones propias de las librerías ARToolKit, se han utilizado un máximo de 14 marcas en una sola aplicación, puesto que al incrementar el número de éstas, el programa puede confundirlas y aumentar el consumo de procesador al tener que reconocerlas e identificarlas en tiempo real. Además, las marcas deben tener un tamaño mínimo. En las distintas versiones de *Usuario Aumentado*, el tamaño mínimo de marcas que se ha utilizado es de 64 mm, trabajando a una distancia usuario-cámara de 2.0 m aproximadamente. Además, el fondo del espacio de representación debe ser homogéneo (preferiblemente blanco), para que no se reconozcan falsas marcas.

A pesar de todas estas consideraciones, las limitaciones en el registro de ARToolKit (ver apartado 5.1.1.4) unida a las condiciones especiales necesarias en sistemas de pantallas de retroproyección (penumbra detrás de la pantalla), hacen que el resultado

no sea muy estable; este hecho se hace más evidente con movimientos rápidos de las marcas, por lo que los usuarios están limitados a movimientos pausados.

7.3. MEMORIA DE LAS DISTINTAS FASES

Cabe destacar que este proyecto se ha ido retro-alimentando a sí mismo, y han surgido diferentes variantes gracias a un feed-back directo con los usuarios, espectadores (gracias a pruebas de usuarios y a diversas exposiciones) e investigadores ajenos al ensayo (gracias a congresos). Así pues, hablaremos de un total de seis versiones o fases que se describen a continuación.

7.3.1. PRIMERA FASE: PRUEBAS DE LABORATORIO

Para describir los inicios de *Usuario Aumentado*, retomaremos unas notas, publicadas anteriormente, que reflejan de forma directa ese momento (Portalés Ricart, 2006): *“Aunque las librerías de ARToolKit se han utilizado en numerosos proyectos de realidad aumentada, nunca antes se había propuesto un proyecto en el que el objeto aumentado fuese el mismo usuario, sustituyendo partes de su cuerpo por los objetos 3D generados por computador. Esta ingeniosa idea surgió de forma un tanto casual, que yo recuerdo además muy divertida, y por ello me gustaría narrarla a continuación: Un buen día en el Laboratorio de Luz estaba probando con mi compañero – Paco Giner – las posibilidades de juego del programa. Yo estaba en el ordenador, sujetando la cámara, mientras Paco sujetaba una de las marcas. El objeto visualizado era uno de los que vienen como ejemplo, una pequeña nave espacial de color magenta que gira sobre sí misma. Entonces empezamos a cambiar alguno de los parámetros que determinan la posición y giros del objeto – la nave, en este caso – con respecto a la marca... ahora más a la derecha, ahora más arriba, ahora se gira un poco... en fin, que al final terminó la nave sobrevolando la cabeza de Paco, a modo de sombrero... y nos reímos un poco. Esto nos pareció un tanto divertido, así que probamos lo mismo con el muñeco de nieve, que también viene de ejemplo. Bueno, poner un muñeco de nieve sobre la cabeza a modo de sombrero ya no era tan gracioso... entonces – creo que fue Paco quién lo sugirió – cambiamos la escala del muñeco para hacerlo más grande, y lo centramos con respecto a la marca, para que el muñeco coincidiese exactamente con la altura de Paco, que sujetaba la marca a la altura del pecho. Esto*

sí que era divertido: ¡Paco se había convertido en muñeco de nieve!... entonces empezó a actuar de forma graciosa, dando saltitos y moviéndose de lado a lado... ¡que risa!... O por lo menos, así es como lo recuerdo. Lo que sí es seguro, es que fue un hallazgo muy divertido".

7.3.1.1. DESCRIPCIÓN TÉCNICA

Las pruebas se realizaron utilizando la pantalla del PC como display. La iluminación era la propia del laboratorio, dos tubos fluorescentes en el techo. Se utilizó la cámara Fire-i de Unibrain, con conexión FireWire y 640x480 píxeles de resolución. La distancia marca-cámara varía entre 0.5 a 1.5 m. En esta primera fase de *Usuario Aumentado* se utilizaron la marca libre y el gorro con prisma ajustable, que se describen a continuación:

Marca libre: Este constituye nuestro primer prototipo de marcas y el más sencillo de todos (Figura 165). Consiste en una única marca con aristas de 18 cm sobre superficie rígida. Las (relativamente) grandes dimensiones de la marca la hacen fácilmente reconocible por el sistema, haciendo factible los movimientos (un tanto) rápidos del usuario, el cual sujeta la marca con sus propias manos. En principio, puede parecer que este hecho disminuya la libertad de movimientos del propio usuario, al tener que sujetar la marca; sin embargo, esto también posibilita que mueva la marca a su antojo pudiendo, por ejemplo, convertirse él mismo en un personaje diferente o tener un amigo virtual, a modo de marioneta.

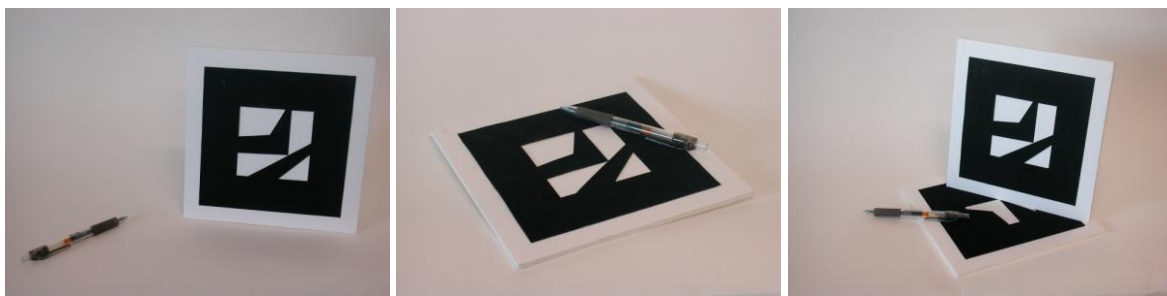


Figura 165. Imágenes de las marcas libres.

Gorro con prisma ajustable: La cabeza es un sitio idóneo para situar las marcas, puesto que pocas veces quedarán ocultas por los brazos de los usuarios, proporcionando así una mayor libertad de movimiento. En un primer intento de situar las marcas en la cabeza, se diseñó una especie de prisma con base triangular variable (Figura 166); la base triangular significa que el prisma está dotado de tres marcas, para que el usuario pueda girar sobre sí mismo, permitiendo que la cámara en todo momento visualice

una de las marcas, las cuales están vinculadas a un mismo objeto/personaje virtual; la base variable significa que las marcas no están estrictamente rígidas entre ellas, sino que estarán más o menos distanciadas (dentro de un rango pequeño) dependiendo del usuario en cuestión, es decir, el prisma se ajusta a la cabeza de cada usuario. Para permitir este ajuste, cada una de las marcas sobre superficie rígida, se adosó a un gorro de lana. Las dimensiones de las marcas siguen siendo relativamente grandes para que el sistema las reconozca fácilmente (aristas de 18 cm). Este prototipo, aunque realmente atractivo, resulta un tanto aparatoso, sobre todo por las dimensiones de las marcas; otro problema encontrado se debe a la base irregular ya que, al estar las tres marcas vinculadas a un mismo modelo virtual, se producen pequeños desajustes de dichos modelos cuando el usuario gira la cabeza, es decir, cuando el sistema pasa de reconocer una marca a otra o reconoce dos marcas al mismo tiempo.



Figura 166. Imágenes del gorro con prisma ajustable.

7.3.1.2. MODELOS VIRTUALES

Los modelos virtuales que se utilizaron para las primeras pruebas de laboratorio fueron los VRML que se distribuyen como ejemplo en el paquete de ARToolKit: un muñeco de nieve y una nave espacial. En ensayos posteriores, se modeló la cara de una mujer (Figura 167).

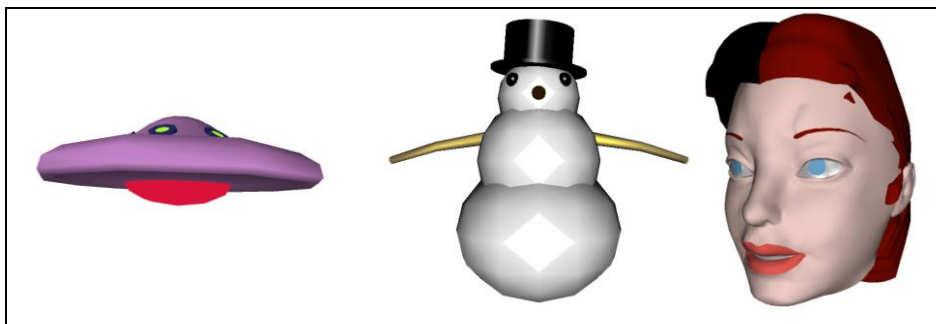


Figura 167. Modelos virtuales utilizados en las pruebas de laboratorio.

7.3.1.3. INTERACCIÓN

En la siguiente figura se muestran algunos ejemplos de interacción con las marcas y modelos virtuales descritos anteriormente.



Figura 168. Ejemplos de interacción en las primeras pruebas de laboratorio.

7.3.2. SEGUNDA FASE: CAMBIO DE PERSONALIDAD

El primer sistema que se diseñó en este ensayo se denominó *Cambio de Personalidad*, ya que los usuarios se ven transformados en diversos personajes, alguno de ellos conocidos, convirtiendo la aplicación en un juego de disfraces mediante el cambio de identidades. No hay nada más allá de la diversión al mirarse a uno mismo o a los otros disfrazados. Así pues, mediante *Cambio de Personalidad* se potencia el juego a través de la metáfora del espejo mágico. En esta primera fase únicamente pretendíamos observar la reacción de los usuarios al interactuar con nuestra pieza. Constituye una prueba de fuego, pues no sabíamos si el sistema de *Usuario Aumentado* resultaría o no atractivo. Tras las primeras pruebas con usuarios (apartado 7.4), los resultados fueron más que satisfactorios.

7.3.2.1. DESCRIPCIÓN TÉCNICA

Las primeras pruebas se realizaron en uno de los estudios de la facultad de BBAA, donde se montó una pantalla de 1.5x2.5 m de plástico blanco sobre marco de madera. En el estudio las condiciones de luz fueron fácilmente controlables, utilizándose un par de focos de recorte para iluminar el espacio de representación. La

cámara utilizada en este caso es la Fire-i de Unibrain, y la distancia cámara-marcas está dentro del rango 1.5-2.5 m. Se emplearon un máximo de siete marcas por usuario, con la posibilidad de que interactuasen al mismo tiempo dos usuarios, de tal forma que se estableciese un entorno colaborativo. Las marcas que se utilizaron fueron el prisma con diadema y las marcas ajustables para las manos, que se describen a continuación:

Prisma con diadema: El prisma con diadema resuelve los problemas encontrados con el prototipo del gorro. Este modelo es un prisma de base triangular rígida (evitando así la no coincidencia de objetos virtuales), el cual se apoya sobre una diadema (Figura 169). Las diademas siguen siendo ajustables a todo tipo de usuarios pero, al estar el prisma completamente sobre la cabeza, las dimensiones de éste deben ser más reducidas. Al reducir las dimensiones de las marcas, la identificación de éstas por el sistema es más complicada. Es por ello que se hicieron varias pruebas para ver cual sería la dimensión apropiada. Finalmente, se utilizaron marcas con aristas de 8 cm. Un problema que surge es la poca estabilidad de las marcas, debido a que están sujetas a la diadema mediante velcro, siendo la zona de contacto bastante reducida. Es por ello que se añadió una segunda diadema, resultando el conjunto un poco más estable. Otro problema encontrado es el hecho de que a algunos usuarios les resultaba un poco incómodo el llevar diadema. Este hecho se observó sobre todo en niños, para los que este sistema resultaba bastante incómodo.



Figura 169. Imágenes del prisma con diadema.

Marcas ajustables para las manos: Se realizaron unas marcas para las manos fácilmente ajustables para cada usuario mediante el empleo de velcros a modo de pasador. Este mismo modelo también se puede emplear para las piernas o pies. De este modo, un usuario que vista un prisma con diadema junto con las marcas para las manos, verá su cabeza y sus manos aumentadas, es decir, transformadas o sustituidas por un objeto virtual. La posibilidad de juego aumenta a medida que se van añadiendo marcas. En la siguiente figura vemos una mano con las marcas ajustables.

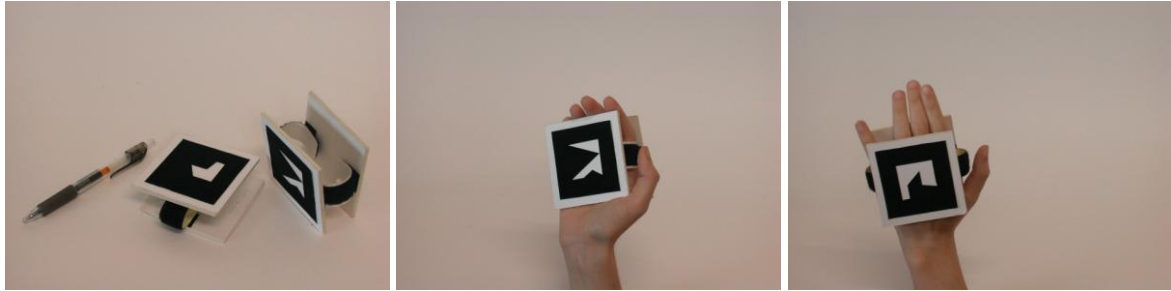


Figura 170. Imágenes de las marcas ajustables para las manos.

7.3.2.2. MODELOS VIRTUALES

Los personajes que se modelaron en esta aplicación son (Figura 171):

- El Sr. y la Sra. Patata.
- Las manos de Mickey.
- Las cabezas de Jack Lemmon, Walther Matthau y Spencer Tracy.

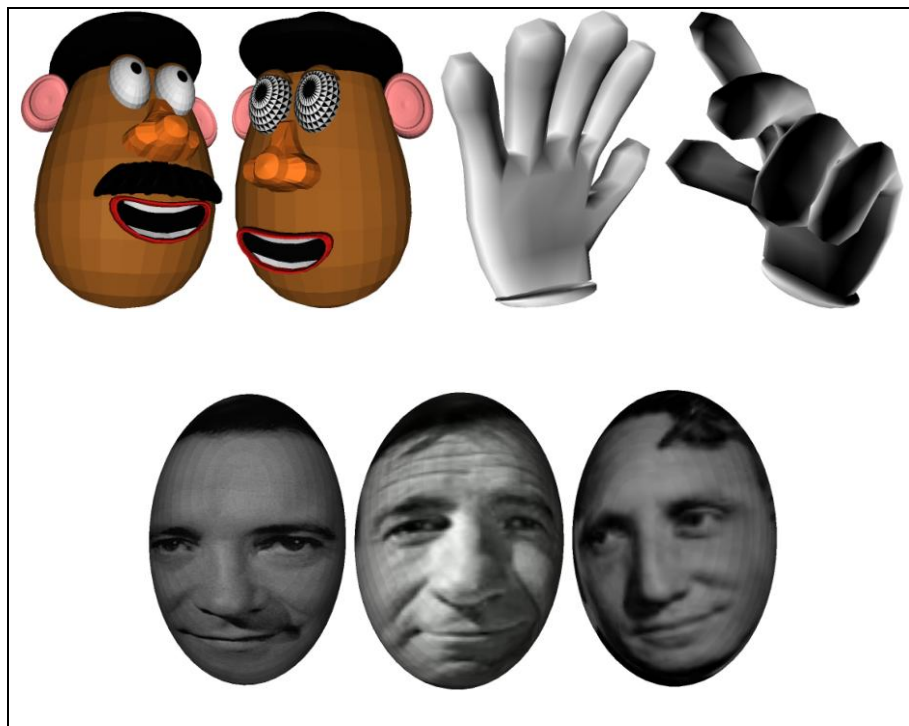


Figura 171. Modelos virtuales utilizados en *Cambio de Personalidad*.

Hay que destacar que las cabezas de Jack Lemmon, Walther Matthau y Spencer Tracy se han modelado mediante simples figuras ovaladas, mapeadas con una foto de la cara de los citados actores. Estos modelos remiten a algunas de las expresiones artísticas de Tony Oursler, que proyecta vídeos de caras humanas sobre una especie

de cabezas que actúan como pantallas de muñecos inanimados en posiciones y escalas no habituales. Todos los modelos están presentados con una escala un tanto exagerada con respecto al usuario, para que sean más visibles y cobren mayor importancia. En esta fase del ensayo, los modelos son estáticos.

7.3.2.3. INTERACCIÓN

En las siguientes imágenes podemos ver a los autores interactuando con el sistema de *Cambio de Personalidad*.

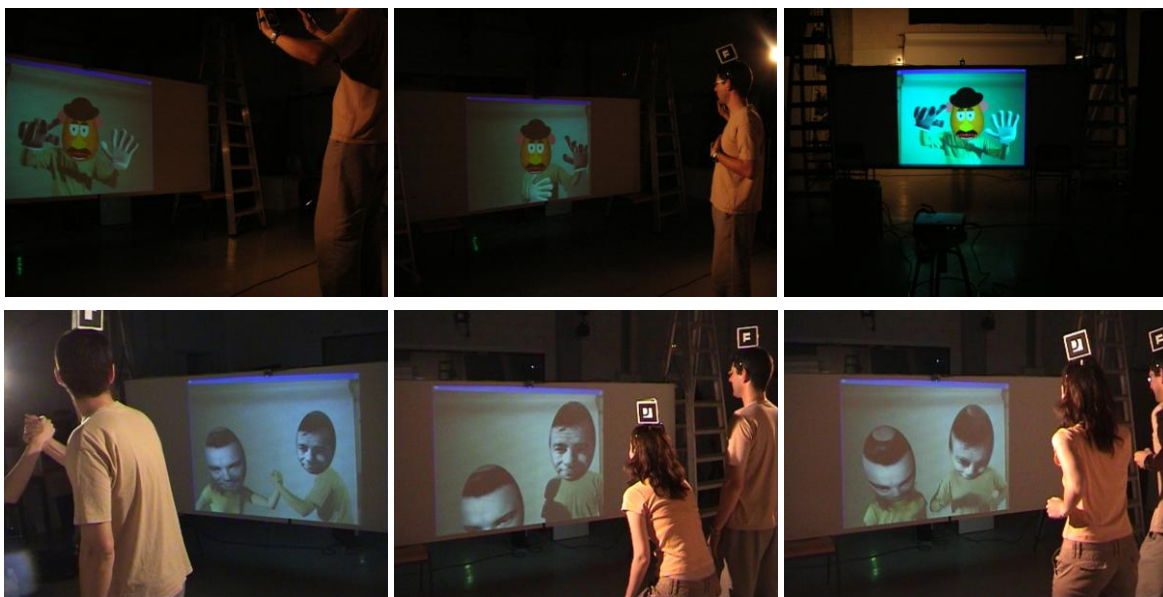


Figura 172. Los autores de *Usuario Aumentado* interactuando con la versión de “Cambios de Personalidad”, en unos ensayos realizados en uno de los estudios de la facultad de BBAA.

7.3.3. TERCERA FASE: CARRUSEL DE IMÁGENES

En esta fase se pretende aludir a la memoria de los usuarios mediante la visualización de una serie de imágenes, a modo de álbum de fotos aumentado que se distribuye geoméricamente por el espacio; es por ello que se ha denominado *Carrusel de Imágenes*. El carrusel virtual consiste en una serie de imágenes ordenadas de forma circular en distintas bandas giratorias que se cruzan, estando el usuario situado en el centro (Figura 173).

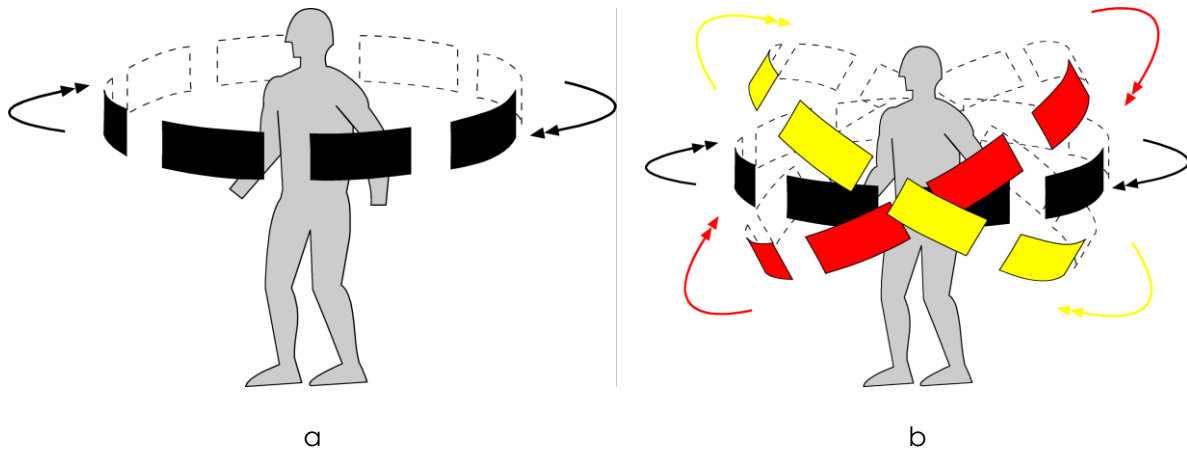


Figura 173. Esquemas de las bandas giratorias de *Carrusel de Imágenes*: a) Una sola banda; b) Tres bandas que se entrecruzan.

Las imágenes pueden aludir a distintos eventos, lugares, escenas familiares o personales, mostrándose alrededor del cuerpo o de la cabeza, aludiendo a los recuerdos del usuario. Mediante las distintas bandas se pueden establecer relaciones temporales o conceptuales: eventos pasados versus recientes, lugares lejanos versus cercanos, elementos conocidos versus desconocidos, etc., que se entremezclan en algún momento. Así pues, el contenido de las imágenes puede tener carácter temático, expositivo, narrativo, etc.

7.3.3.1. DESCRIPCIÓN TÉCNICA

Los elementos técnicos de esta fase son los mismos que los de *Cambio de Personalidad*, con la única salvedad que como marcas se ha utilizado únicamente el prisma con diadema, ya que las manos no se han aumentado.

7.3.3.2. MODELOS VIRTUALES

El modelo virtual consta de una serie de superficies planas, ordenadas de forma circular, sobre las que se mapean las imágenes. Se pueden establecer distintas bandas, cruzándose entre ellas en un punto determinado, formando un efecto visual similar al del kaleidoscopio. Las características del modelo virtual se pueden modificar, cambiando la velocidad de rotación, número de las superficies planas, número e inclinación de las bandas, intersecciones, etc. A un mismo modelo virtual se le pueden asignar una serie de imágenes distintas.

En esta aplicación nos encontramos con el problema de la oclusión, ya que el modelo virtual gira alrededor del usuario, por lo que el usuario debe ocultar la parte del modelo que queda a sus espaldas. Las librerías de ARToolKit no contemplan la posibilidad de solucionar este problema, que de por sí es complejo, por lo que se optó por una solución alternativa en la fase de modelado: la información de mapeado (en 3D Studio Max) se dio únicamente para las caras externas de las bandas, quedando las internas sin textura. Esto provoca que la parte interna de las bandas aparezca como invisible, no dejando ver a través de ella la parte externa.

7.3.3.3. INTERACCIÓN

En las siguientes imágenes podemos ver a uno de los autores interactuando con el espejo de *Carrusel de Imágenes*.



Figura 174. Ejemplos de interacción con el *Carrusel de Imágenes*.

7.3.4. CUARTA FASE: TÚNEL MÁGICO

Teniendo en cuenta que la misma marca puede ser asignada a diferentes modelos, pero cada sistema solo puede asignar una marca cada vez a un modelo específico, en la segunda fase del ensayo (*Cambio de Personalidad*) no podríamos disfrutar de cambios de modelo dentro de un mismo sistema. Se puede considerar un sistema cerrado a una visualización determinada. Existen a priori diferentes posibilidades para subsanar este problema, desde programar un cambio de modelo asignado a una marca, lo que conllevaría un cambio asignado a una variable aleatoria o a un contador temporal y por lo tanto un cambio ajeno a la voluntad del usuario, hasta el relanzamiento del sistema con un modelo diferente cada vez, proceso que conlleva una espera mientras un sistema se desconecta y otro se carga. Decidimos que lo interesante sería buscar un procedimiento por el cual el sistema ofreciera de manera inmediata al usuario la posibilidad de adoptar diferentes identidades alteradas en función de su voluntad.

En este punto surgió la tercera fase del ensayo, denominada *Túnel Mágico*. Utilizando diferentes sistemas en paralelo, se otorga al usuario la posibilidad para cambiar de identidad cuando quiera o recrearse en la actual todo el tiempo que desee, y también de hacerlo de una forma natural e inmediata, simplemente saliendo del campo de visión de una cámara y entrando en la de otra. Al estar los sistemas cargados de antemano, la proyección es inmediata en cuanto se reconocen las marcas por cada cámara. Este sistema múltiple de visualización de los usuarios y los modelos permite una gran interactividad, puesto que no sólo puede un usuario ser aumentado, sino que también es capaz de ver las transformaciones de sus acompañantes en el sistema autónomo paralelo y comprobar sus evoluciones con cada una de estas identidades. La sensación de cambiar visualmente de aspecto y actuar en consecuencia si el usuario reconoce la nueva identidad que ha adoptado es inmediata, produciéndose una contaminación del comportamiento muy interesante.

El túnel hace referencia a un transcurrir de un lugar a otro, atravesando varios espejos, siendo el entorno a su vez más íntimo. En la primera fase del ensayo se observó que, algunos usuarios, se sentían un poco cuartados bajo la mirada del desconocido. Con este sistema se pretendía transformar el espacio de representación en un lugar más recogido, con luces tenues, donde los usuarios no se sintiesen continuamente observados (aunque hubiera otros espectadores), y pudieran actuar libremente de acuerdo con su nueva personalidad.

7.3.4.1. DESCRIPCIÓN TÉCNICA

El sistema de *Túnel Mágico* consta de tres sistemas básicos de *Usuario Aumentado* dispuestos de forma paralela, quedando las tres pantallas en línea (Figura 175). Las pantallas utilizadas en este caso son también de plástico blanco sobre marco de madera, con unas dimensiones cuadradas de 1.5x1.5 m. La iluminación del espacio escénico se realizó mediante focos de recorte, aunque con mayores dificultades para obtener una correcta iluminación en el sistema central. Al estar el sistema de *Usuario Aumentado* triplicado, se utilizaron un total de tres cámaras Fire-i de Unibrain, dispuestas sobre cada una de las pantallas. En la Figura 176 podemos ver imágenes de su montaje.

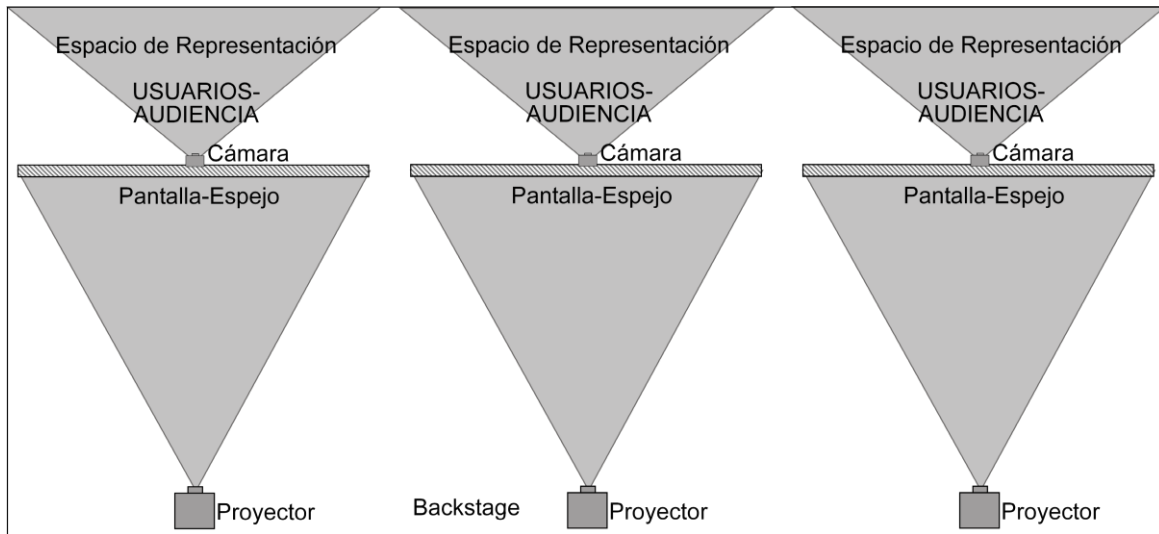


Figura 175. Esquema del sistema de *Túnel Mágico*.

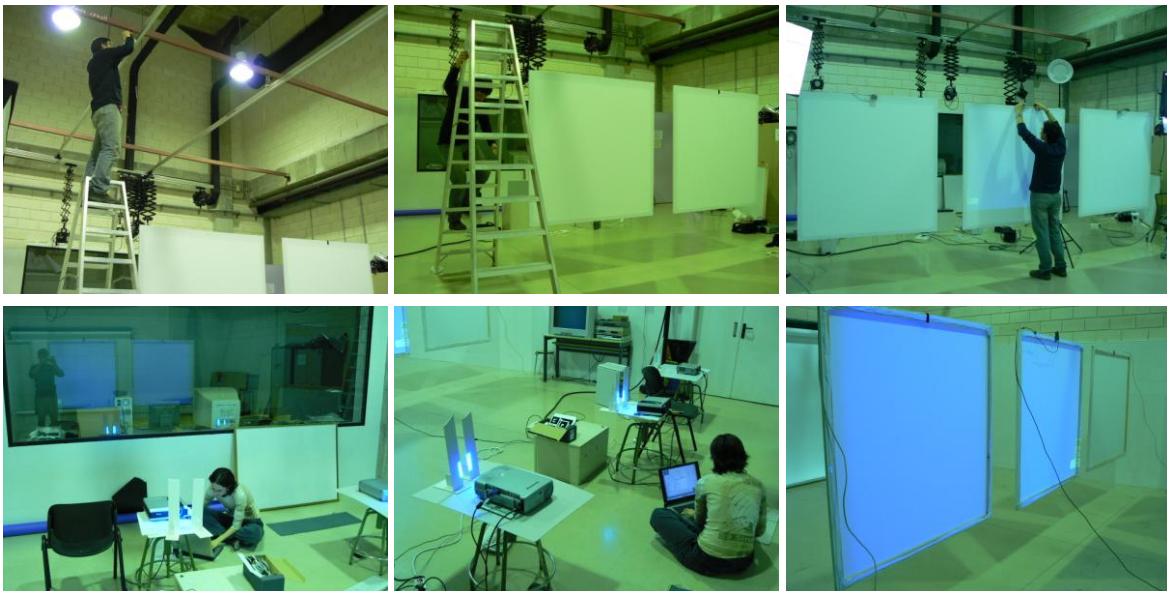


Figura 176. Montaje de *Túnel Mágico* en uno de los estudios de la facultad de BBAA.

En esta fase se utilizaron el prisma con diadema y las marcas ajustables para las manos, y también se probó una nueva marca, el prisma con diadema retroiluminado, que se describe a continuación:

Prisma con diadema retroiluminado: Una versión diferente del prisma con diadema, es el prisma con diadema retroiluminado. Este prisma está hecho con material translúcido (la parte blanca de las marcas) para poder dejar pasar la luz de un total de 6 LEDs insertados en su interior, y alimentadas por una batería en su base (Figura 177). De este modo, bastará que el escenario tenga una luz tenue para que el sistema pueda reconocer las marcas sin dificultad. Las marcas, al tener luz propia, siempre serán visibles por la cámara.

En la siguiente figura se aprecia el proceso de fabricación y funcionamiento de este prisma.

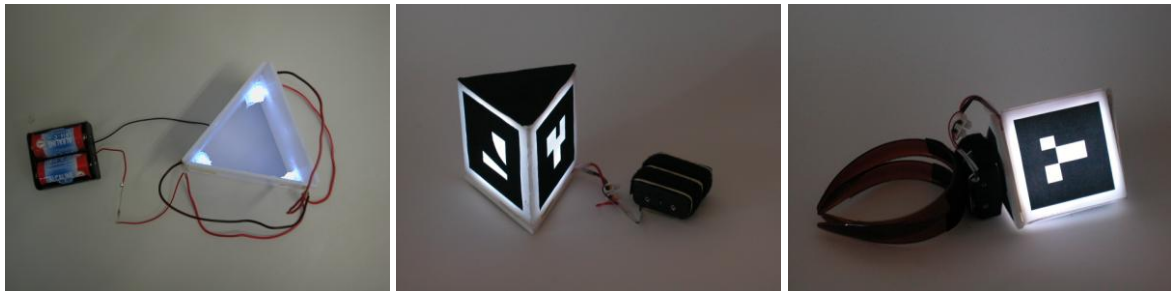


Figura 177. Imágenes del montaje y funcionamiento del prisma con diadema retroiluminado.

7.3.4.2. MODELOS VIRTUALES

Los modelos virtuales que se modelaron en esta fase son (Figura 178):

- Tres de las señoritas de Avignon de P. Picasso.
- Tres figuras poliédricas con mapeados de las caras de los autores.
- Tres figuras míticas del cine: Charles Chaplin, Stan Laurel y Oliver Hardy.



Figura 178. Modelos virtuales utilizados en *Túnel Mágico*.

En el caso de las figuras poliédricas se utilizaron imágenes de la cabeza de frente, de perfil, y desde arriba para mapear los modelos VRML. Las figuras del cine se mapearon con una sola imagen de la cara del correspondiente actor vista de frente.

7.3.4.3. INTERACCIÓN

En las siguientes imágenes podemos mostramos el sistema de *Túnel Mágico* montado al completo junto con diversos usuarios interactuando con los distintos espejos.



Figura 179. Interactuando con *Túnel Mágico*.

7.3.5. QUINTA FASE: UN[∞]

Una de las mayores críticas recibidas en cuanto a las primeras fases de *Usuario Aumentado* es la "falta de contenidos". Es por ello que en esta cuarta fase se planteó la posibilidad de incidir directamente sobre la naturaleza del individuo que interactúa con la pieza, mediante la elaboración de tres pantallas-espejos que cuestionan la identidad del usuario bajo diferentes puntos de vista. El nombre de *Un[∞]* (leído "un infinito") hace referencia a uno mismo con infinitas personalidades.

En los espejos de *Un[∞]* hablaremos de tres tipos de identidades: la identidad psicológica, la física y la social. Como apunta Villanueva “...*las personas se ven amenazadas en su existencia... en parte por la socialización o lo que es lo mismo por la pérdida de su individualidad a manos de la sociedad que las masifica y disuelve*” (Villanueva, 1995, 151). Sin embargo, “*Las personas son gregarias, no pueden sustraerse de la vida en común con los demás*” (Villanueva, 1995, 153). Así pues, parece inevitable que exista una pérdida o modificación de identidad individual a favor de una masificación o globalización, en tanto que se persigue una aceptación o integración social. Villanueva habla también de una disparidad entre identidad aparente e identidad real, de cómo aparecen las personas a sí mismas y a los demás. En este momento se nos pueden presentar dos cuestiones. En primer lugar: ¿hasta qué punto somos capaces de modificar nuestra conducta a favor de un reconocimiento social?, ¿somos conscientes de nuestra “otra” identidad, la que tenemos oculta?, ¿nos vemos a nosotros mismos como nos mostramos a los demás?... De otro modo: si la noción de identidad implica la noción de colectividad (pluralidad de miembros), deberemos analizar el individuo dentro de la sociedad en la que se haya, y habrá que tener en cuenta que su identidad (o parte de ella) se forma dentro de un conjunto socio-cultural y quizá no tenga sentido hablar de identidad real o verdadera: ¿cómo despojar al individuo de su identidad no real si para descubrirla deberíamos apartarlo de la sociedad?, ¿no perdería en ese momento su identidad?, ¿cómo definirse a uno mismo si no es comparándose con el resto? De todo esto hablan los espejos de *Un[∞]*: de nuestra identidad dentro de una sociedad plural (identidad social) y de nuestra adaptabilidad dentro de esa sociedad (identidad física y psicológica), en la que nos identificamos como individuos. Se invita al usuario a reflexionar sobre la siguiente cuestión: ¿hasta qué punto somos dueños de nuestra identidad?

Así pues, se proponen tres cambios de identidad, tres espejos que tratan de indagar en los siguientes aspectos:

Espejo de las emociones: Versa sobre la identidad psicológica, el yo interno verdadero frente al aparente, el que se pretende mostrar o somos capaces de exteriorizar. A través de la cara nos reconocemos y comunicamos nuestros sentimientos más internos. “*La cara...; en ella se organiza una falsa individualidad, una pretendida originalidad personal... Las personas se disfrazan, esconden y metamorfosean, se sirven de artificios, ocultamientos y máscaras para cuestionar su identidad, reinventarla y dar una imagen plural de su existencia*” (García Cortés, 1997, 107). En este sentido, podemos afirmar que: *La cara “no siempre” es el reflejo del alma.*

Esta dualidad de estado emocional interno y estado emocional exteriorizado, de ser y aparentar, se muestra mediante un juego de máscaras que revelan diferentes estados emocionales. Como bien dijo el antropólogo francés Lévi-Strauss (Lévi-Strauss, 1982, 44), *"la máscara niega tanto como afirma"*. Las máscaras ocultan a la vez que muestran. Así pues, las máscaras se utilizan como alegorías del yo interno frente al exteriorizado; la capacidad de las personas para adaptarse a diversas situaciones.

Espejo de los estereotipos: Versa sobre la identidad física, el yo físico aparente o inventado, diseñado a gusto personal y para los demás. Se basa en el concepto de posthumano: *"el arte posthumano... un arte en el cual, ..., el cuerpo natural se diluye, en un primer momento, en beneficio de la representación del cuerpo ajeno a sí mismo, para, con posterioridad, hacerlo en beneficio de la representación de un cuerpo monstruoso"* (Cruz Sánchez et al., 2004, 268); entendiendo la monstruosidad como *"la evidencia de una diferencia, con el añadido de una deficiencia"* (Cruz Sánchez et al., 2004, 274). El cuerpo posthumano constituye por una parte un *cuerpo reconstruido* "a la carta", gracias a los avances científicos, y por otra, un *cuerpo proteico*, mediante la incorporación a nuestro cuerpo de diversos artilugios (Cruz Sánchez et al., 2004, 23-24). Pensemos en el cuerpo reconstruido como la búsqueda de la belleza a partir de una (falsa) perfección facial y corporal mediante la cirugía estética, o, incluso, ante la posibilidad de una manipulación genética. De otro modo, pensando en el cuerpo proteico, nos vienen a la cabeza imágenes de piercings, uñas postizas, pelucas, extensiones, e incluso los cyborgs (mitad hombre mitad máquina).

Espejo de género-multicultural: Versa sobre la identidad social, el yo dentro de un contexto social. En este espejo hablamos de la sociedad frente al individuo. Aquí es donde se hace más evidente la búsqueda de la identidad personal frente a la pluralidad social, algo que nos identifique y nos distinga de los demás. La sociedad en la que vivimos es cada vez más cosmopolita. En (Nash et al., 2001, 21) se apunta a que *"... el multiculturalismo contempla la existencia de la diversidad cultural en el seno de la sociedad. Pretende asimismo elaborar políticas de reconocimiento de sus diversas expresiones y establecer bases para la igualdad de oportunidades"*. Cuanto antes comprendamos que la naturaleza humana es plural, antes lograremos una sociedad con igualdad de oportunidades.

7.3.5.1. DESCRIPCIÓN TÉCNICA

En esta fase se utilizó el sistema básico de *Usuario Aumentado* por triplicado, al igual que en *Túnel Mágico*, con la salvedad que los tres displays se unificaron en una única pantalla de plástico blanco sobre marco de madera con dimensiones de 6.0x1.5 m.

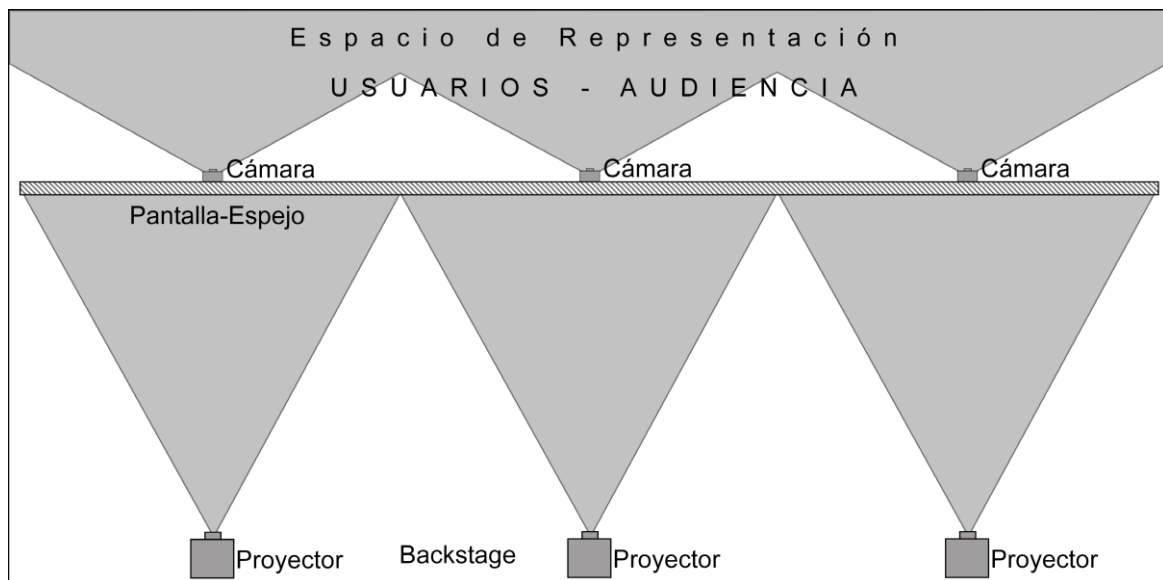


Figura 180. Esquema de la distribución espacial de los dispositivos de Un^∞ .

La pieza se montó en la sala del Centro Párraga de Murcia, junto con otras piezas de Laboratorio de Luz en la exposición denominada "Especulaciones a un Tiempo" (2006). En la sala, se establecieron condiciones de luminosidad uniformes para todas las piezas. Para poder resaltar las marcas, se añadieron cuatro tubos de luz negra en la parte inferior de la pantalla, dispuestos uno al lado del otro, de forma longitudinal (Figura 181).





Figura 181. Los integrantes del Laboratorio de Luz montando la pieza *Un[∞]* en el Centro Párraga de Murcia.

Se utilizaron un total de 14 marcas simultáneas (7 para un solo usuario) y se diseñaron nuevas marcas, las marcas del prisma con diadema mejorada y las marcas a modo de guantes, que se describen a continuación:

Prisma con diadema mejorada: Estos prismas se diseñaron con la finalidad de mejorar, funcional y estéticamente, los prismas con diadema. Hay que señalar que son un diseño de Ricardo P. Bochons (Figura 182). Resultan mucho más estables que los prismas con diadema, no se clavan a la cabeza, las marcas son un poco mayores y están unos centímetros separadas de la cabeza para evitar oclusiones (ocasionadas por el pelo). En lugar de cartón-pluma, se ha utilizado un material un poco más rígido, al que se le aplicó una capa de pintura blanca especial, para que resaltase aún más con la acción de la luz negra. Sin embargo tienen un peso ligeramente mayor que las anteriores, y siguen siendo incómodas para los niños.

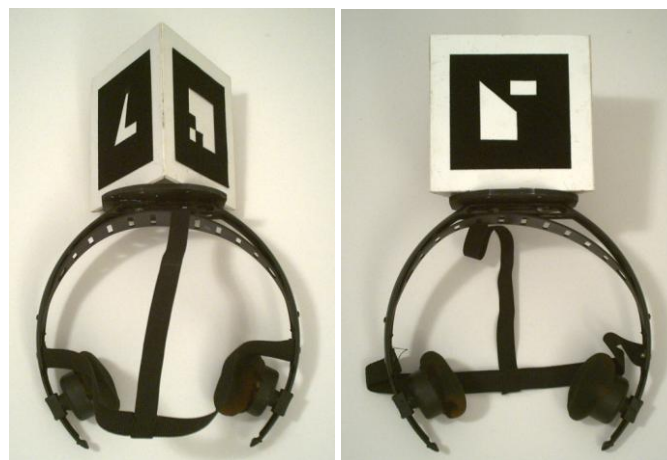


Figura 182. Imágenes del prisma con diadema mejorada.

Marcas a modo de guantes: Los guantes están hechos con tela blanca de algodón un poco elástica, por lo que son ajustables. Cosidas sobre ellas están las marcas, con dimensiones de 64 mm. El material sobre el que se apoya la marca es el mismo que en el

caso del prisma con diadema mejorada, y también se le aplicó una capa de pintura blanca especial (Figura 183). Una vez elaboradas, surgió el problema de que resultaba dificultoso para un único usuario ponerse el segundo guante una vez puesto el primero.

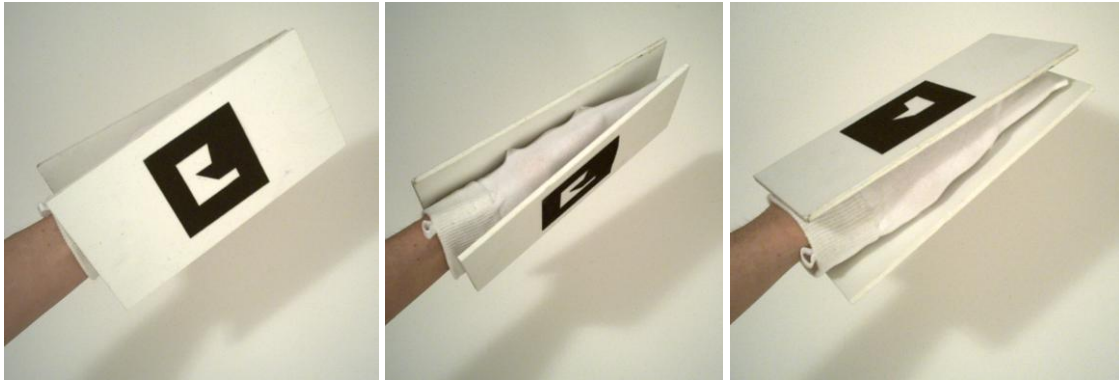


Figura 183. Imágenes de las marcas a modo de guantes.

7.3.5.2. MODELOS VIRTUALES

Como ya se ha dicho, en esta pieza se proponen tres cambios de identidad, tres espejos que tratan de indagar en los siguientes aspectos. A continuación se describen y muestran los modelos virtuales utilizados en cada espejo.

Espejo de las emociones: La cabeza de los usuarios se aumenta mediante máscaras griegas tradicionales, con sus facciones cambiando lentamente, pasando de la felicidad a la tristeza y de la sorpresa a la indiferencia (Figura 184-a y -b). Las manos se aumentan mediante textos relacionados con las máscaras (Figura 184-c): "la máscara niega tanto como afirma", "la máscara hace visible lo invisible, negando la evidencia existente", "en la máscara coinciden realidad y deseo, el ser y el querer ser", "se clausura el espacio de la inhibición y se inaugura el gran teatro del mundo", "todas las transformaciones tienen algo de profundamente misterioso y vergonzoso a la vez", "cuando la figura, la cara, es sólo transformación, la máscara es el espejo del alma". Estos textos aparecen divididos de tal forma que cada usuario tiene la mitad de cada uno de ellos. Así pues se propone un juego, en el que los dos usuarios deben colaborar para formar las frases completas, simplemente juntando sus manos de forma adecuada.

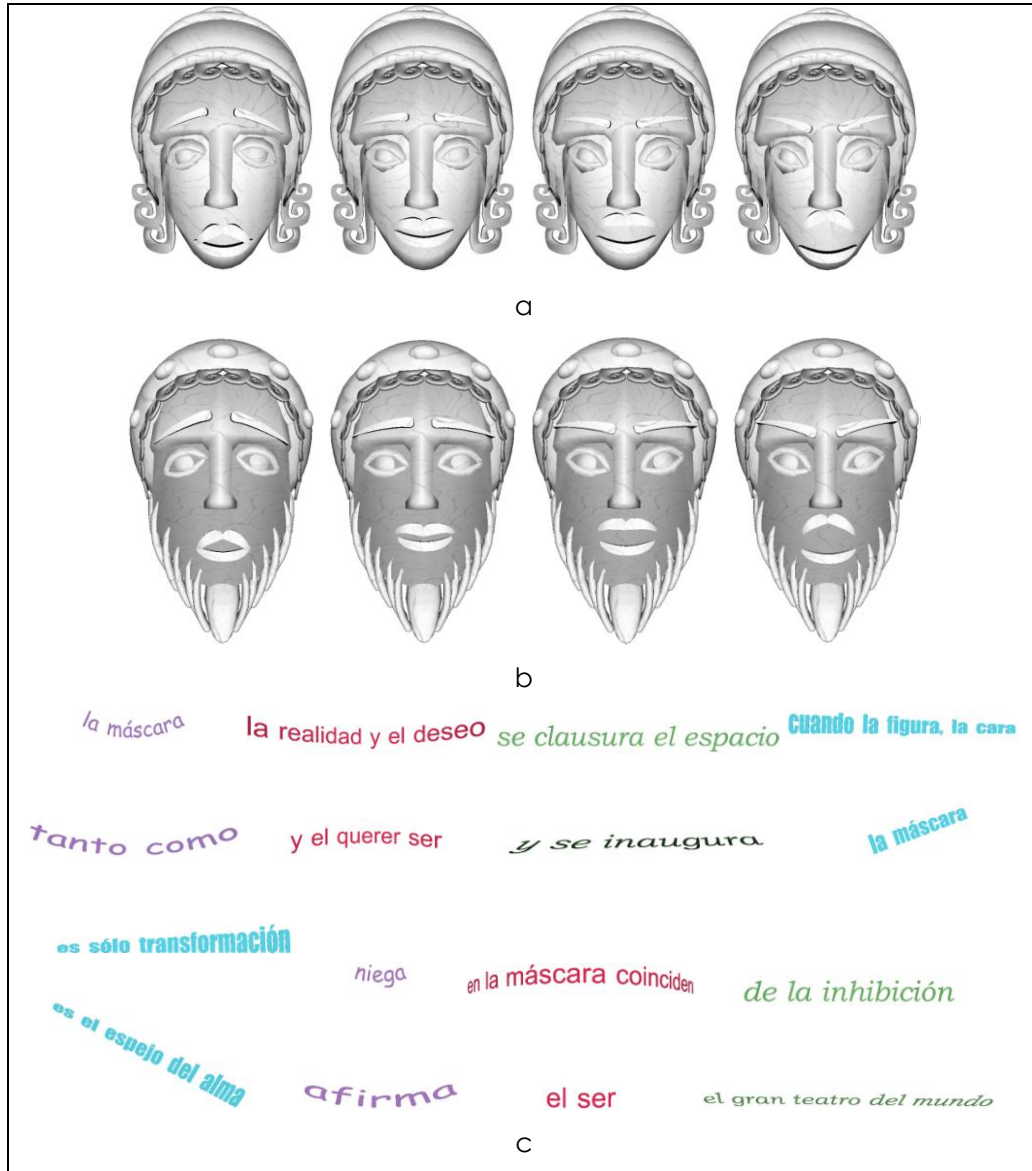


Figura 184. Modelos virtuales utilizados en el espejo de las emociones: a) Máscara griega femenina; b) Máscara griega masculina; c) Frases vinculadas a las manos.

Espejo de los estereotipos: En este espejo los usuarios son aumentados con unas cabezas que recuerdan a las Suprernas o a las Tanga Girls, siendo una de ellas masculina (Figura 185-a) y la otra femenina (Figura 185-b). Sus facciones aparecen y desaparecen, de acuerdo a la teoría de la reconstrucción del cuerpo posthumano. Las manos de los usuarios son aumentadas con una serie de palabras relacionadas con este espejo, como (Figura 185-c): cirugía, dieta, moda, body-pump, piercing, tattoo, etc. Las palabras están distribuidas conformando una especie de esferas giratorias que dificultan su lectura, con lo que hay que poner atención para poder leerlas; metafóricamente aludiendo a que no siempre nos damos cuenta de la imposición de estas reconstrucciones.

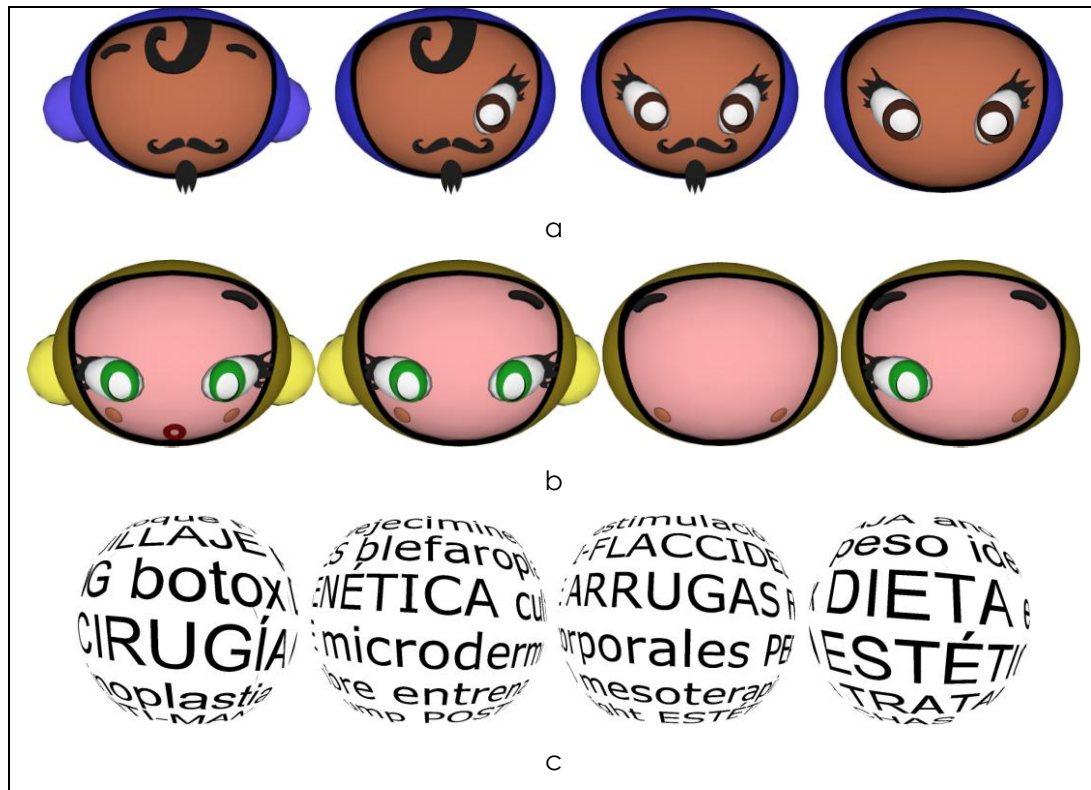
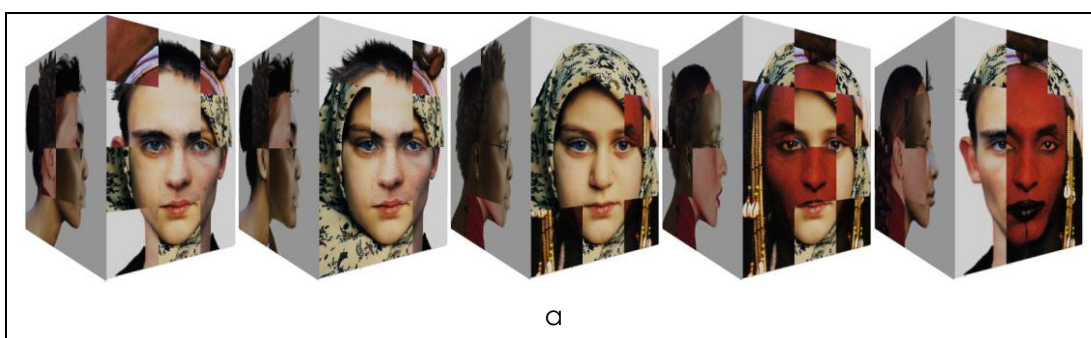


Figura 185. Modelos virtuales utilizados en el espejo de los estereotipos: a) Modelo para la cabeza masculino; b) Modelo para la cabeza femenino; c) Esferas de palabras para las manos.

Espejo de género-multicultural: En este tercer espejo los usuarios son aumentados con unas cabezas en forma de cajas, las cuales aparecen mapeadas formando mosaicos mediante la composición de varios retratos, abarcando personas de distintas edades, género, culturas, etc. (Figura 186-a y -b). Estos mosaicos tienen forma de rejilla, estando cada cara de la caja compuesta por una rejilla de 3x4 (frente) o 4x4 (laterales) cuadrados, que recuerdan al famoso cubo de Rubik, típico de los años 80. Cada cierto tiempo, los cuadrados de las rejillas van cambiando de mapeado, mostrando en su conjunto una cara cambiante compuesta por diferentes rostros. Las manos de los usuarios se aumentan mediante una serie de palabras relacionadas con este espejo, como por ejemplo (Figura 186-c): cosmopolita, plural, racial, mujer, género, etc. Las palabras están ordenadas formando unas espirales y figuras radiales giratorias.



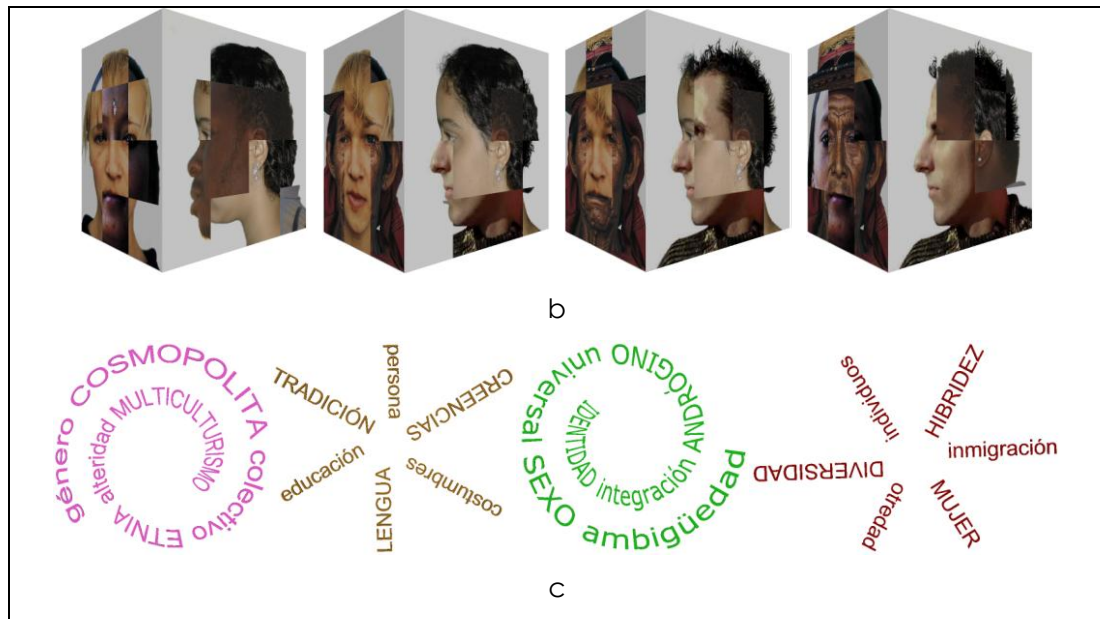


Figura 186. Modelos virtuales utilizados en el espejo de género-multicultural: a) y b) Modelos de cajas para la cabeza; c) Palabras en forma de espiral y figuras radiales para las manos.

7.3.5.3. INTERACCIÓN

En las siguientes imágenes podemos ver el sistema de Un^∞ montado al completo junto con diversos usuarios interactuando con los distintos espejos.



Figura 187. Secuencia de imágenes de varios usuarios interactuando con la pieza de Un^∞ . Centro Párraga (Murcia).

7.3.6. SEXTA FASE: CINEATRO (CINE + TEATRO)

En esta fase se aborda el tema del edutainment. Esta era una idea que nos llamaba la atención desde un primer momento, pero que por la evolución propia de la pieza no se había abordado hasta ahora. Se desarrolla una pieza que va dirigida a un público infantil, para edades de 5 a 12 años.

En este ensayo se cambia el sistema de *Usuario Aumentado* en cuanto al espacio ocupado por los usuarios y la audiencia, de tal forma que la audiencia ve a los actores aumentados desde la parte trasera de la pantalla-espejo, como si estuvieran en una sala de cine; sin embargo la acción tiene lugar en tiempo real, estando los actores en el escenario, por lo que también se asemeja a una sala de teatro; de esta combinación surge el nombre de *Cineatro*.

7.3.6.1. DESCRIPCIÓN TÉCNICA

En esta fase se planteó un sistema a modo de teatro, en la que una serie de niños pudiesen representar un cuento, cada uno de ellos transformado en un personaje distinto por el sistema del *Usuario Aumentado*. En esta fase incluimos algunas diferencias con respecto a la distribución espacial del sistema establecido en *Usuario Aumentado* (Figura 188):

- En primer lugar, la parte de detrás del espejo (el backstage) se convierte en la sala de la audiencia; de este modo se invade la parte trasera del espejo, que permanecía reservada únicamente a los dispositivos (ordenadores y proyectores) en las anteriores versiones.
- En segundo lugar, se desglosa la unión usuario-audiencia; el usuario es esta vez actor, mientras que la audiencia únicamente ve la imagen del entorno aumentado, sin ver físicamente a los usuarios-actores que interactúan con la pieza. De igual forma, los usuarios-actores no pueden ver a la audiencia, sino únicamente su reflejo aumentado en el espejo mágico; al verse a sí mismos representando su personaje, pueden controlar su actuación.
- Los dos puntos anteriores conceptualmente conllevan que: para los usuarios-actores la pantalla-espejo sigue la metáfora del espejo mágico; para la audiencia, la pantalla-espejo sigue la metáfora de la lente mágica.

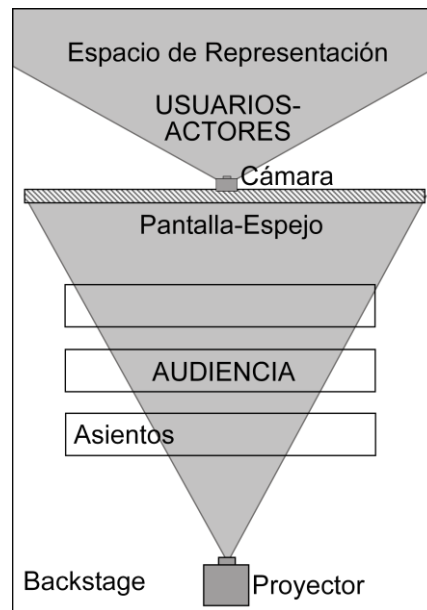


Figura 188. Esquema de las relaciones espaciales en el sistema de *Cineatro*.

En este ensayo, a diferencia de los anteriores, se le exige al usuario que represente un papel determinado, el de su personaje, en lugar de dejar que el usuario asuma libremente la identidad aportada por su representación en la pantalla-espejo. Además, en el sistema de *Cineatro* se incrementa el número de usuarios que interactúan con la pieza, ya que en las versiones anteriores nunca se consideraron más de dos usuarios por cada pantalla-espejo.

El sistema de *Cineatro* se montó en una habitación para realizar pruebas con distintos usuarios. En este caso se empleó una única pantalla de plástico blanco sobre marco de madera, con dimensiones de 2.5x1.5 m. La luminosidad de la sala se controló con un foco de luz regulable. La cámara web utilizada es la Live! Cam Voice de Creative, que tiene conexión USB 2.0 y un FOV de 85°. Aunque esta cámara ofrece la posibilidad de trabajar con una resolución de 1280x960, se utilizó la misma resolución que en el resto de fases (640x480) para evitar problemas en la velocidad de procesamiento. La distancia cámara-maracas está dentro del rango de 1.0-2.0 m. Para esta fase se diseñaron unas nuevas marcas a modo de peto; cada usuario porta un peto que consta de dos marcas, por lo que en total el sistema reconoce 10 marcas. Éstas se describen a continuación:

Marcas a modo de peto: Estas marcas se diseñaron para ser llevadas por niños, en sustitución de los prismas de la cabeza. Constan de dos marcas unidas por cintas, que se ponen en el torso a modo de peto, quedando una marca en el pecho y la otra en la espalda. Las cintas tienen un sencillo sistema para poder ser ajustadas dependiendo

de la altura y contorno de cada usuario (Figura 189). Tienen la desventaja que se pueden producir oclusiones ocasionales por los brazos, pudiendo coartar la libertad de movimiento.

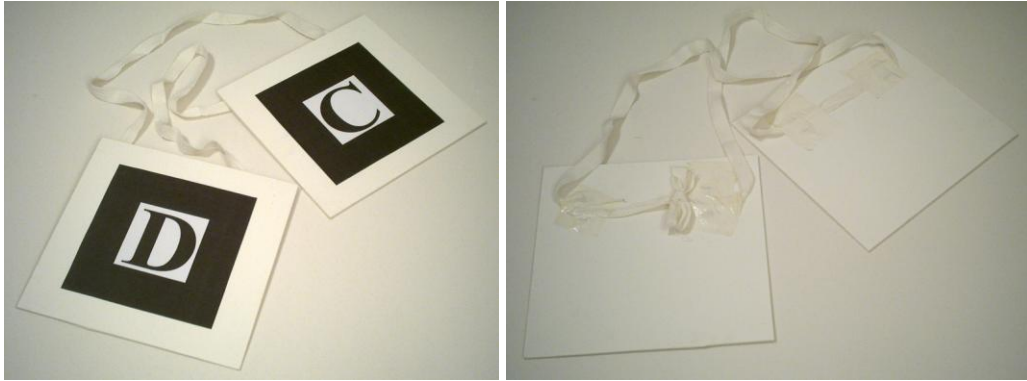


Figura 189. Imágenes de las marcas a modo de peto.

7.3.6.2. MODELOS VIRTUALES

Se realizaron unas pruebas para las que se modelaron una serie de personajes del cuento de las aventuras de Alicia en el País de las Maravillas (Figura 190): la taza, la tetera, el reloj, la reina de corazones y un naipe de la guardia real. Este cuento es una obra de literatura infantil creada por el matemático, sacerdote anglicano y escritor británico Charles Lutwidge Dodgson, bajo el más conocido pseudónimo de Lewis Carrol. Cuenta la historia de una niña, llamada Alicia que tras caer a través de una madriguera de un conejo, entra en un mundo de absurdos y paradojas lógicas.



Figura 190. Modelos virtuales utilizadas en *Cineatro*.

7.3.6.3. INTERACCIÓN

El sistema fue probado por un grupo de niños entre 7-12 años. No todos los niños conocían el cuento propuesto, por lo que la representación fue improvisada, dando lugar a una nueva historia a partir de los personajes. En la Figura 191 pueden verse estos niños (junto con dos de sus madres) interpretando distintos papeles: un niño echando azúcar dentro de la taza y removiéndola; los naipes aparecen erguidos y con actitud seria, ya que van a arrestar a Alicia; el reloj advierte que es la hora del té, y la tetera echa un poco de té en la taza; etc. Finalmente, todos juntos hacen un baile para despedirse del público.

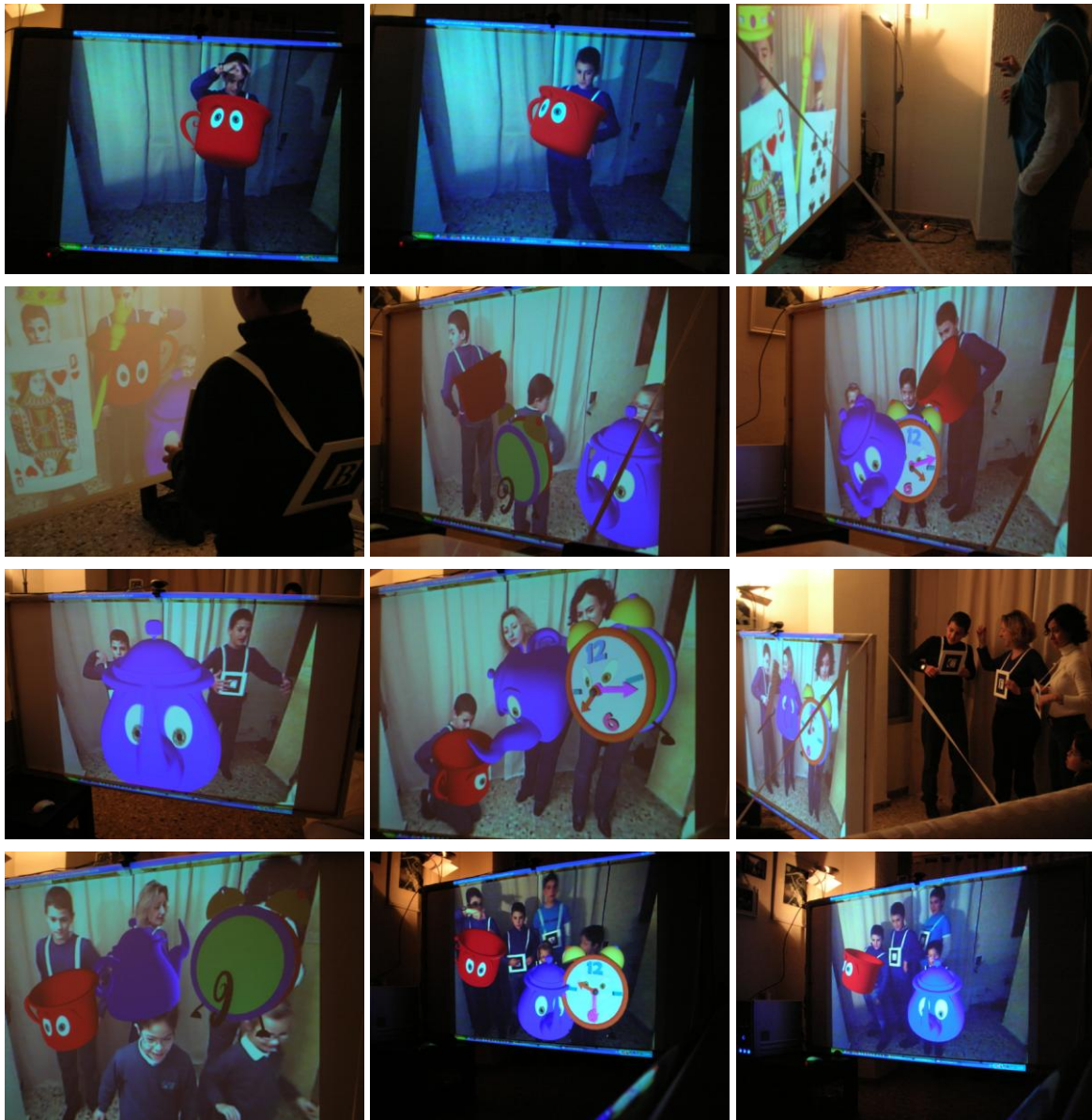


Figura 191. Diversos usuarios (niños y adultos) interactuando con la pieza de *Cineatro*.

7.4. SISTEMAS DE VERIFICACIÓN

La pieza de *Usuario Aumentado* se puso en marcha a finales de Julio del 2005 (entre la segunda y la tercera fase) en la sala de exposiciones de la facultad de BBAA de la UPV durante dos días. El primer día se hizo una filmación exhaustiva del sistema, probando todo tipo de marcas diseñadas. Este primer ensayo se realizó por los propios autores con el fin de probar el sistema y corregir o ajustar factores tales como la luminosidad, resolución de la cámara, distancia usuario/cámara y movilidad del usuario dentro del escenario establecido. El segundo día se abrió la exposición al público en general, que tuvieron la posibilidad de probar nuestro sistema de modo individual y colectivo. La respuesta de los usuarios fue mejor de la esperada y totalmente enriquecedora. La primera impresión que produce es de “asombro”, al verse el usuario a sí mismo transformado en otro personaje. A continuación, generalmente se establecía un ambiente de juego y diversión, especialmente si la aplicación era colaborativa, es decir, varios usuarios interactuando al mismo tiempo con el sistema. Los usuarios permanecían un tiempo interactuando en una sola pantalla; posteriormente pedían más cambios de personalidad, por lo que simplemente tenían que interactuar con otra de las pantallas para descubrir el nuevo personaje en el que se transformaban y crear nuevos juegos. De esta experiencia se derivan las siguientes conclusiones:

- A mayor cantidad y variabilidad de personalidades adquiridas, mayor interés;
- A mayor número de partes del cuerpo aumentadas, mayor posibilidad de juego;
- A mayor número de usuarios interactuando con el sistema, mayor implicación.

Observamos que las personas reconocían la imagen de la pantalla-espejo como si fueran ellos mismos, cambiando su comportamiento. Por ejemplo, si se transformaban en Charles Chaplin, empezaban a andar como él; si se transformaban en el Sr. Patata, empezaban a hacer movimientos cómicos. Les gustaba observar su propia imagen mientras se movían, girando, saltando, bailando o simplemente caminando, siempre mirando a su nueva imagen. A través de esta pieza, los usuarios son capaces de expresarse y desinhibirse; y con los cambios de personalidad propuestos, pueden ser cualquier cosa, moverse libremente, adoptando nuevos roles en cada cambio. La interacción con diferentes usuarios es también una parte crucial. Observamos que las experiencias colaborativas permiten un juego más divertido y un diálogo entre los espejos y los usuarios en el entorno compartido, incrementando la sensación de inmersión. También observamos que para los usuarios resulta muy sencillo interactuar con los espejos, sin necesidad de explicaciones previas.

Pedimos a los usuarios que rellenasen unas encuestas para, según su criterio, evaluar el sistema; en total se rellenaron 34 encuestas. La edad de los participantes oscila entre 7 y 60 años, la mitad de ellos hombres y la otra mitad mujeres. En esta encuesta, les pedíamos que evaluaran del 1 (malo) al 10 (excelente) un total de 9 aspectos:

1. Facilidad de interacción con el sistema.
2. Libertad de movimientos del usuario.
3. Diseño y confort de las marcas.
4. La idea de espejo como interfaz.
5. Personalidades adquiridas.
6. Posibilidad de juego/diversión.
7. Posibilidad en el área del aprendizaje.
8. Posibilidades terapéuticas.
9. Posibilidad de incluir el sistema en museos.

A continuación se muestran unas figuras con las puntuaciones medias para cada uno de los 9 aspectos, desglosados en primer lugar entre hombres (50%) y mujeres (50%), y en segundo lugar entre usuarios menores de 30 años (32%) y mayores de 30 años (68%).

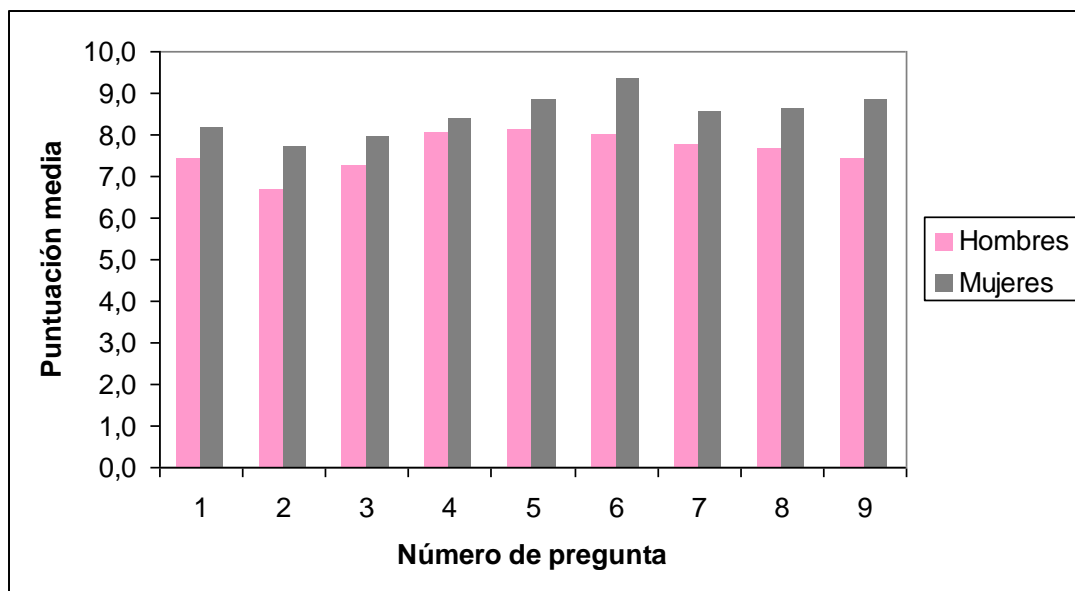


Figura 192. Puntuaciones otorgadas por los usuarios; distinción entre hombres y mujeres.

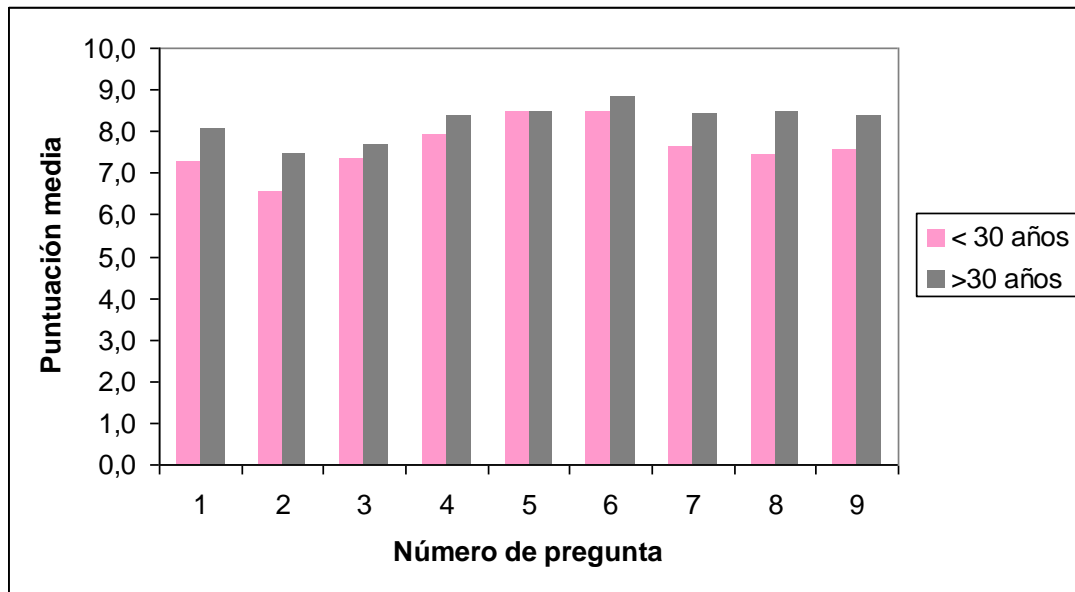


Figura 193. Puntuaciones otorgadas por los usuarios; distinción entre mayores y menores de 30 años.

Como se puede observar, todas las puntuaciones están por encima de 6.5 y, en general, las mujeres y los usuarios mayores de 30 años son los que disfrutaron (un poco) más de la experiencia. El aspecto peor puntuado es el 2, que corresponde a la libertad de movimientos del usuario; no es de extrañar la valoración que los usuarios han aportado en esta cuestión, pues, como se ha apuntado anteriormente (apartado 7.2.3), si las marcas se mueven de forma rápida el sistema no es capaz de detectarlas, perdiendo en ese momento la visión aumentada y restando inmersión.

Finalmente, también pedimos que nos diesen su opinión personal y que anotasen si cambiarían o añadirían algo. Algunas de las aportaciones que nos parecieron más interesantes son: *“Puede tener grandes aplicaciones didácticas, al igual que terapéuticas, especialmente en tratamientos de fobias”*; *“Su uso en psicoterapia: juegos de rol, terapia familiar, psicodrama. Sería interesante incorporar la voz en paralelo o integrada para que fuera más completo”*; *“Me he divertido mucho, falta igual ambientación musical de acuerdo con el personaje creado”*; *“Se podría crear un entorno. Por ejemplo el usuario convertido en Tarzán y el espacio en el que se mueva, ambientado como si fuera la selva...”*; *“Incluir más elementos. Profundizar en lo didáctico”*. En la siguiente figura podemos ver diversos usuarios interactuando con la pieza.



Figura 194. Usuarios interactuando con el sistema de *Usuario Aumentado* en la facultad de BBAA; fase de *Cambio de Personalidad* y *Túnel Mágico*.

También pudimos recoger las impresiones de algunos investigadores y artistas en el congreso VSMM 2005, celebrado del 3 al 7 de Octubre en la Flanders Expo en Gante (Bélgica). La pieza de *Usuario Aumentado* se presentó en la sección de arte mediante una comunicación y un póster. El tercer día del congreso, presentamos en la zona de exposiciones una versión reducida del sistema, con un ordenador portátil, mostrando a Charles Chaplin y al Sr. y la Sra. Patata. A la gente le llamó mucho la atención la pieza, pero al principio se sentían reacios a ponerse las marcas; a continuación apareció un reconocido investigador, Bernard Fisher, al que le pareció muy interesante nuestra propuesta, haciendo muchas fotos de la pieza y probándose las marcas; en ese momento se armó un gran revuelo y se hizo un gran corro alrededor de la pieza;

finalmente, muchos de los presentes se sintieron seducidos por la pieza de *Usuario Aumentado* y se vistieron con las marcas. En la siguiente figura se muestran algunas imágenes del evento.



Figura 195. Usuarios interactuando con una demo de *Usuario Aumentado* en el congreso VSMM 2005, en Gante, Bélgica. Fase de cambios de personalidad.

8. URBANMIX

UrbanMix es una aplicación de realidad aumentada que propone una mezcla de entornos urbanos mediante la incorporación de modelos virtuales en el espacio físico donde se desarrolla la aplicación. Los edificios modelados también existen, pero en un lugar diferente, de tal forma que se acercan y se entremezclan dos ámbitos de distintas características. El usuario se siente inmerso en un espacio urbano híbrido, pudiendo hallarse, conceptualmente, en dos lugares al mismo tiempo. Técnicamente, *UrbanMix* se construye con un sistema propio de realidad aumentada gestionado por un software que no está hecho para este fin. Además, se aborda el problema de la oclusión. Este ensayo se ha realizado con la colaboración de Francisco Giner y Francisco Sanmartín.

Motivación

Las aplicaciones de RA móviles en entornos abiertos suelen basar la localización del usuario en el registro mediante GPS, principalmente porque permite un registro continuo en un sistema de coordenadas global – donde el usuario puede desplazarse libremente sin límites espaciales – y tiene bajos costes. El registro por técnicas de computador se evita por varios motivos: las condiciones de luminosidad son difíciles de controlar; a menudo no se pueden poner marcas físicas en el lugar; y la complejidad del entorno real suele complicar el registro por el reconocimiento de elementos naturales. Sin embargo, los receptores GPS de bajo coste presentan limitaciones en

cuanto a precisión que impiden establecer una correcta alineación entre el entorno virtual y el real, por lo que en muchas de estas aplicaciones la visualización a través del display es exclusivamente del entorno virtual. Así, en este ensayo se prescinde del GPS, gracias a un método de registro mixto que combina una cámara y un sensor inercial, para tratar de aprovechar las ventajas de ambos dispositivos de registro: la precisión del sistema visual y la independencia de las condiciones ambientales/visuales del sensor inercial unido a su carácter global. También se reducen los problemas del registro visual, puesto que únicamente necesitamos dos puntos de apoyo con coordenadas 3D conocidas. En otro orden de ideas, el tratamiento del problema de la oclusión no es sencillo, y muchas de las aplicaciones de RA no lo abordan. En este ensayo si se toma en cuenta, y se resuelve mediante la incorporación de modelos 3D – del entorno físico donde se desarrolla la aplicación – que actúan como máscaras.

Marco temporal

Este ensayo, realizado en el Laboratorio de Luz durante 2004-2005, fue la primera aplicación de RA que realizábamos con un software multitarea. *UrbanMix* fue presentado en el congreso ACE (Advanced Computer Entertainment) 2005 celebrado en el campus de Vera de la UPV, y organizado por Laboratorio de Luz en colaboración con Mixed Reality Lab de la National University of Singapore. En este congreso se presentó una ponencia titulada *Back to the 70's*, que proponía una vuelta nostálgica a la universidad de los años 70, cuando sus edificios estaban todavía ubicados en el centro de la ciudad.

8.1. PLANTEAMIENTO

La Universidad Politécnica de Valencia tiene cuatro campus: el de Vera, el de Gandía, el de Játiva y el de Alcoy, siendo el primero el central. El campus de Vera está situado en el límite norte de Valencia. Pero antaño, la universidad estuvo integrada en el centro de la ciudad. La historia de esta integración data de hace más de 500 años, cuando el núcleo de la población era muy pequeño y compacto, con una estructura de desarrollo urbano de herencia medieval. Durante siglos, la presencia de estas instituciones influyó en la vida de los ciudadanos, pero con el tiempo, el desarrollo urbano y el crecimiento demográfico hicieron necesaria la ampliación de las instalaciones, adquiriendo en principio espacios en los límites del centro urbano que fueron absorbidos por el rápido crecimiento de la urbe, con lo que las escuelas y facultades pronto se vieron dispersas y desconectadas. Además, el encarecimiento del terreno en las zonas limítrofes de la ciudad unido a la creciente demanda de plazas para estudiantes, provocaron que los nuevos edificios de las universidades se ubicaran en grandes extensiones alejadas del centro de la ciudad, en campus autosuficientes, con los servicios y recursos necesarios (comedores, papelerías, bancos, gimnasios, etc.).

En este ensayo se propone una vuelta nostálgica al reencuentro entre la Universidad y el centro de la ciudad, posible gracias a la tecnología de realidad aumentada: diversos edificios emblemáticos del centro de la ciudad de Valencia – Les Torres de Serrans y El Micalet (Figura 196) – se transportan e integran en el campus de Vera, formando una atractiva composición visual en la que se unen realidad e ilusión, modernidad e historia.



Figura 196. Ubicación del Campus de Vera respecto al centro histórico de Valencia y los monumentos Les Torres de Serrans y El Micalet.

8.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO

Siguiendo las aportaciones del bloque I, se puede elaborar un marco general de este ensayo:

- Ubicación dentro del continuo de Milgram (apartado 1.1.1):
 - Realidad aumentada.
- Clasificación (apartado 1.2.2):
 - Según el entorno físico en el que se desarrolla la aplicación:
 - Abierta.
 - Según la extensión que abarca:
 - Local.
 - Según la movilidad de los dispositivos de registro y/o displays:
 - Móvil.
 - Según los usuarios que simultáneamente pueden interactuar con el sistema:
 - Individual.
 - Según el tipo de colaboración establecida:
 - (No existe colaboración).
 - Clasificación según Mackay (apartado 1.2.2):
 - Aumentar el usuario.
- Campo de aplicación (apartado 1.4):
 - Arte.
- Sistema de coordenadas terreno/objeto (apartado 2.1.1):
 - Terrestre local.
- Orientación espacial (apartado 2.1.5):
 - Método mixto.
- Dispositivos de registro / controladores (capítulo 3):
 - Sensores ópticos: Rasgos naturales.
- Displays (capítulo 4):
 - Display sujeto a la cabeza: De vídeo.
- Software (capítulo 5):
 - Multimedia: Max/MSP Jitter.
- Forma de interactuar (apartado 6.1.1):
 - Navegación.

8.1.2. REFERENTES

Existen varios autores que relacionan espacios urbanísticos y/o arquitectónicos distantes espacialmente y/o temporalmente. Por ejemplo, la pieza *Emperadores Desplazados, Arquitectura Relacional 2* (1997) de Rafael Lozano-Hemmer (Figura 197), presentada en el Festival Ars Electronica (Linz) en colaboración con Hill Bauer y Susie Ramsay, es una instalación que relaciona dos espacios arquitectónicos separados espacial y temporalmente mediante una serie de proyecciones sobre uno de los edificios. Los participantes utilizan el interfaz "arquitecto" para transformar el Castillo Habsburgo de Linz (Austria) en el Castillo de Chapultepec, residencia de los Emperadores Habsburgo de Méjico. Por medio de sensores, se calcula el lugar hacia donde los participantes apuntan con el brazo, para controlar una gran mano proyectada que, en su desplazamiento por la fachada del castillo, va descubriendo las imágenes del otro castillo. Además, por diez chelines, el espectador podía apretar el botón de Moctezuma y activar una animación post-colonial del Penacho Azteca que permanece en el Museo etnológico de Viena, acompañada de música étnica mejicana.

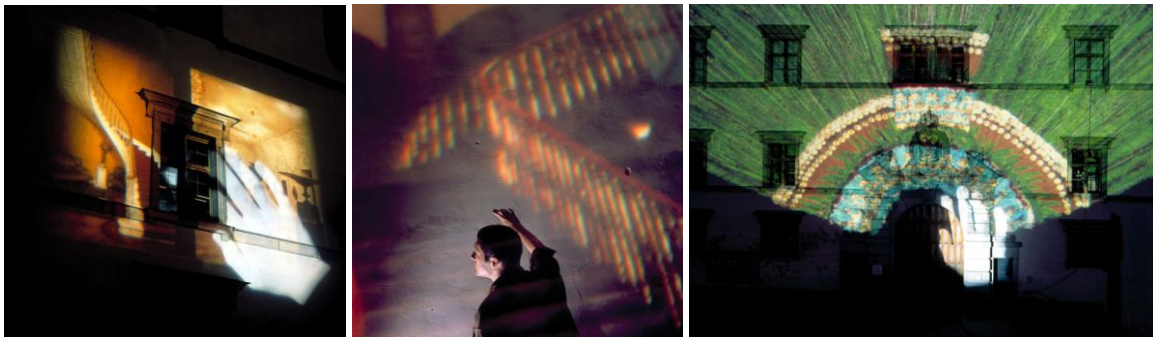


Figura 197. Imágenes de *Emperadores Desplazados, Arquitectura Relacional 2*. En (Lozano-Hemmer, 2007).

De otro modo, en la pieza *lifeClipper* (2004) de Jan Torpus (ver apartado 1.4.1) se mezclan distintas etapas históricas de un mismo espacio urbano. En el proyecto MARS (Figura 198), descrito en (Höllner et al., 1999), se entremezclan espacios urbanos separados temporalmente: se trata de un conjunto de dispositivos (GPS, ordenador portátil, etc.) que permiten la localización y movilidad del usuario dentro de los límites de la Columbia University. Mediante un HMD óptico se pueden visualizar elementos virtuales integrados en el espacio urbano, como por ejemplo edificios que han sido demolidos. Se establece también un sistema paralelo de realidad aumentada en un espacio cerrado, mediante el cual el usuario ve una representación aumentada de un

plano del campus universitario, y puede mover diversos elementos del plano que afectarán la visualización del usuario que está en el entorno abierto, estableciéndose una comunicación entre ambos entornos.



Figura 198. Usuario con los dispositivos de RA y visualización de un edificio derruido. En (Höllerer et al., 1999).

8.2. DESCRIPCIÓN TÉCNICA

8.2.1. COMPONENTES DE LA APLICACIÓN

8.2.1.1. DISPOSITIVOS FÍSICOS

Los dispositivos físicos utilizados en *UrbanMix* son (Figura 199):

- Un ordenador portátil con una tarjeta gráfica de 512 MB y con el software Max/MSP Jitter instalado.
- Un HMD: I-glasses SVGA video (ver apartado 4.1.1.2).
- Baterías para el HMD.
- Una cámara Web Logitech, con una resolución de 640x480 píxeles y conexión USB 2.0.
- Un sensor inercial: MT9 de Xsens.
- Una mochila que transporta el ordenador portátil y las baterías.



Figura 199. Dispositivos físicos utilizados en *UrbanMix*.

El usuario lleva consigo un ordenador portátil en una mochila, el cual procesará toda la información recibida en tiempo real y realizará las operaciones necesarias para crear el entorno urbano aumentado a través de una aplicación realizada con *Max/MSP Jitter*. Los displays utilizados en aplicaciones de RA móviles al aire libre suelen ser HMD o dispositivos pequeños de mano (teléfonos móviles o PDAs), aunque en algunos casos se utilizan también pantallas (de ordenadores portátiles o de tablet-PC). Las pantallas tienen el inconveniente de que, a plena luz del día, su luminosidad no es suficientemente potente, por lo que se requiere un sistema adicional que proteja la pantalla de la luz directa, como se hace en (Wilde et al., 2003), introducido en el apartado 4.1.3.1 La desventaja de los dispositivos pequeños de mano es que su capacidad de almacenaje y proceso para aplicaciones de RA es limitada. Es por ello que en *UrbanMix* optamos por la utilización de un HMD de vídeo conectado al ordenador portátil.

Otro dispositivo indispensable en esta aplicación es una cámara, cuyo centro de proyección debe aproximarse al punto de vista del usuario, por lo que deberá situarse próximo a sus ojos. La resolución de la cámara no deberá ser ni demasiado alta (para permitir el proceso en tiempo real) ni demasiado baja (para no ofrecer una imagen precaria); en este caso se utiliza una cámara Web de Logitech, con una resolución de 640x480 píxeles. Por último, se utiliza un sensor inercial MT9 de Xsens, situado sobre la cámara y alineado con ésta, que registra las rotaciones de la cabeza del usuario. El sensor va conectado al ordenador portátil y a las baterías mediante cables.

8.2.1.2. ELEMENTOS VIRTUALES

Al tratar el problema de la oclusión, en este ensayo necesitamos dos tipos de modelos virtuales: los modelos virtuales de los edificios que se van a insertar en el entorno físico mediante la aplicación de RA; y modelos virtuales de los edificios del entorno físico donde se desarrolla la aplicación, para ser utilizados como máscaras en el proceso de la oclusión.

El espacio urbano en el que se desarrolla la aplicación (el campus de Vera) se modela tridimensionalmente (Figura 200). Teniendo en cuenta que el modelado no ha de tener demasiado detalle, podemos simplificar la geometría para que la cantidad de polígonos resultante no sea excesiva. La planta de los edificios se puede obtener de varias formas. La más sencilla es por medio de mapas urbanos o imágenes aéreas a partir de las cuales digitalizar el contorno de los edificios; por ejemplo, en Google Earth se pueden adquirir imágenes aéreas con suficiente resolución, dependiendo de la zona. A continuación se asignan alturas a los edificios calculadas mediante técnicas fotogramétricas o topográficas (Kraus, 1997). El proceso de modelado se ha realizado con el programa 3DStudio Max. Finalmente, los modelos se han exportado a formato *obj* para poder ser cargados en Max/MSP Jitter.

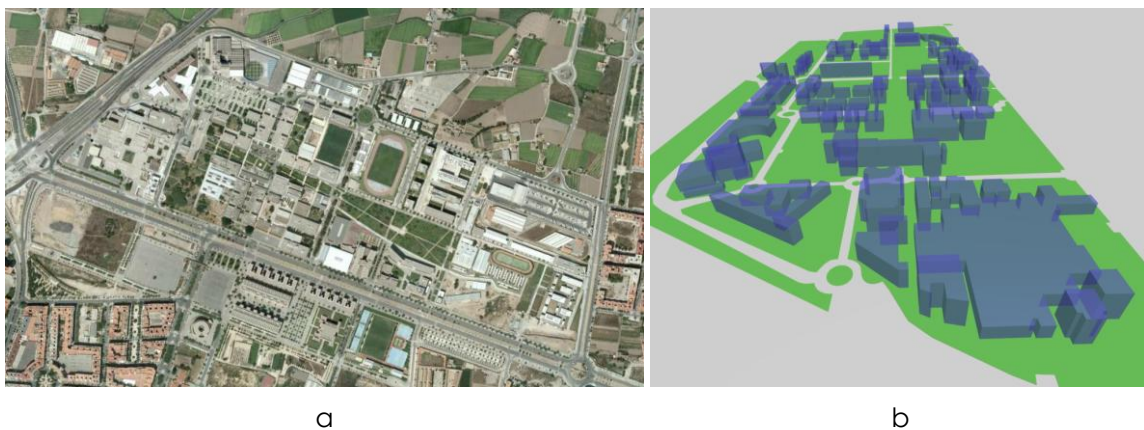


Figura 200. Campus de Vera: a) Imagen aérea, en (Google, 2007); b) Modelo tridimensional realizado con el programa 3D Studio Max.

Respecto a los edificios que introducimos en la aplicación de forma virtual, hemos de recurrir a un nivel de detalle mayor en su geometría y a un mapeado posterior (a partir de imágenes fotográficas o de texturas similares). La geometría de los modelos virtuales es más compleja y consta de una cantidad de polígonos más elevada. Se sigue el

mismo proceso de modelado que en el caso anterior, pero de forma más precisa y detallada. Finalmente, los modelos obtenidos también se exportan a formato *obj*.

Los edificios modelados en *UrbanMix* fueron dos del casco histórico de Valencia, Les Torres de Serrans (Figura 201) y El Micalet (Figura 202). Se tomaron múltiples imágenes desde distintos puntos de vista y distancias, para capturar la volumetría total y los detalles.

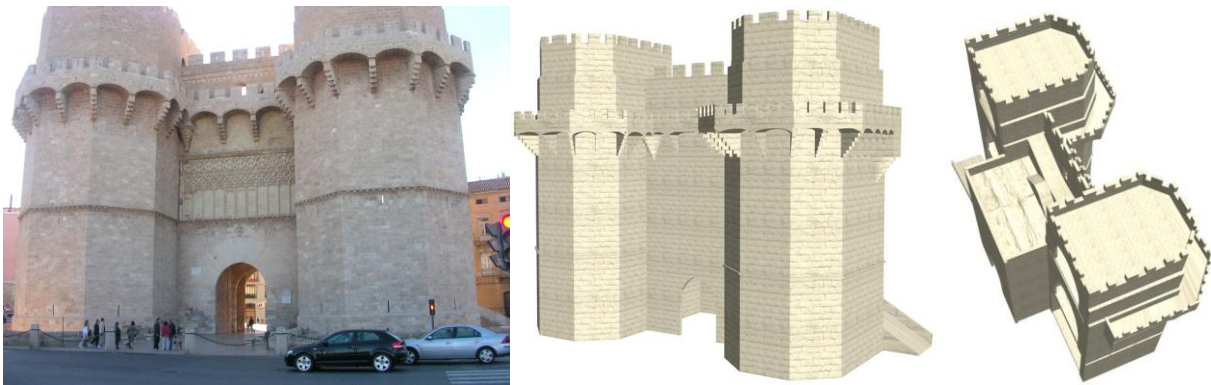


Figura 201. Imágenes de Les Torres de Serrans y modelo 3D.



Figura 202. Imágenes de El Micalet y modelo 3D.

Los programas de modelado 3D poseen una herramienta de texturización, comúnmente llamada mapeado, que introduce en la imagen distintos efectos (rugosidades, brillos, abombamientos, etc.); sin embargo, al exportar el objeto 3D al formato *obj* y cargarlos en Max/MSP Jitter, estos efectos programados se pierden, ya que se suele exportar simplemente la información volumétrica y el enlace a la imagen original que ha de servir de textura. Por ello, se ha realizado la integración y mapeado de estas imágenes directamente en el programa Max/MSP Jitter.

8.2.2. DESARROLLO, PROCESO

8.2.2.1. SISTEMA DE REFERENCIA

Para que la escena generada por ordenador esté en concordancia con la del entorno físico, debemos establecer un marco de referencia común. En *UrbanMix* se ha establecido un sistema de referencia terrestre local (ver apartado 2.1.1) en el jardín de la facultad de Bellas Artes, definido a partir de la dirección de sus tres ejes ortogonales y el origen de coordenadas (Figura 203). El eje Y sigue la dirección de la fachada Este de uno de los edificios de la zona, alineada con la dirección Norte; el eje X apunta hacia el Este; el eje Z es ortogonal al plano definido por los ejes X e Y, y apunta hacia el cenit. El centro de coordenadas coincide con la esquina de uno de los edificios. A partir de este sistema de referencia se realiza el registro del usuario (posición y rotaciones) en tiempo real.



Figura 203. Sistema de referencia terrestre local.

A continuación se hicieron una serie de mediciones topográficas (Figura 204) para calcular la altura del edificio semicircular de la facultad de Bellas Artes y coordenadas tridimensionales a los centros de dos elementos circulares, uno de color amarillo y otro de color magenta (Figura 205), que serán utilizados para el registro del usuario en tiempo real mediante un tracking de color (ver apartado 8.2.2.3); a partir de ahora llamaremos a estos elementos circulares "puntos de apoyo". También se hicieron mediciones adicionales como reserva. El instrumental topográfico utilizado fue una estación total Leica con distanciómetro y prisma para la medida de distancias.



Figura 204. Labores topográficas.



Figura 205. Puntos de apoyo utilizados para el registro del usuario.

Una vez establecido el sistema de coordenadas y realizados los modelos virtuales de los edificios de la zona, los otros modelos virtuales de los dos edificios del casco histórico de Valencia se ubican en el espacio de tal forma que el usuario, situado en el centro del jardín, al mirar hacia ellos pueda ver simultáneamente los dos puntos de apoyo necesarios para el registro. Además, los edificios del casco histórico deben situarse en un lugar donde no haya otros edificios, para que la composición sea plausible. Finalmente, los dos edificios del casco histórico se ubicaron en la zona ocupada por un aparcamiento al aire libre (Figura 206).

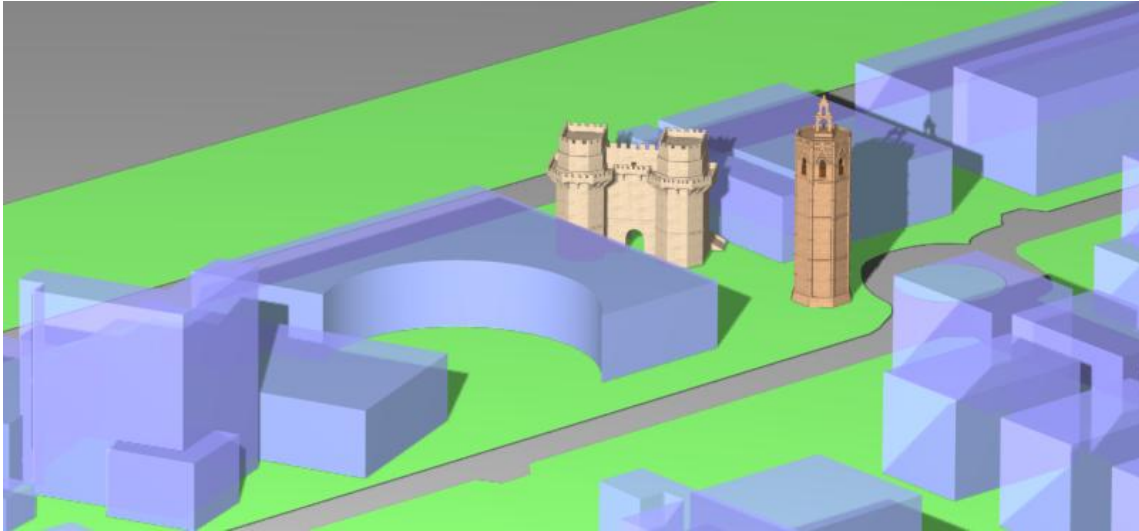


Figura 206. Ubicación espacial de los edificios históricos en el campus de Vera.

8.2.2.2. CONSIDERACIÓN DE LA OCLUSIÓN

Para integrar adecuadamente los edificios virtuales con los reales, en este ensayo consideramos el problema de la oclusión. Como se ha explicado anteriormente (apartado 2.3.1) la oclusión que se debe tratar es la que se produce cuando un objeto real está situado espacialmente entre la cámara y el objeto virtual; el caso contrario (el virtual entre la cámara y el objeto real) no necesita de un tratamiento especial, puesto que la capa de vídeo que contiene a los virtuales aparece siempre superpuesta a la que contiene la imagen del entorno físico. En la Figura 207 se muestra un ejemplo de composición de vídeo sin el tratamiento de la oclusión: el objeto virtual, en este caso Les Torres de Serrans, aparece superpuesto a la imagen de fondo, cuando espacialmente está ubicado detrás del edificio semicircular de la facultad de Bellas Artes (es por ello que las torres aparecen como flotando).

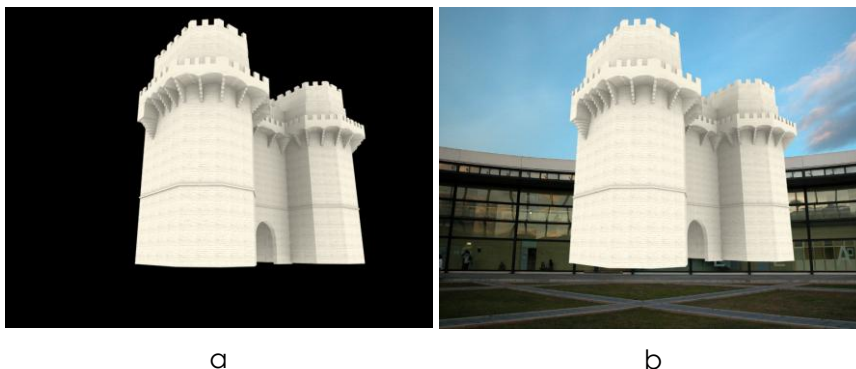


Figura 207. Superposición de un elemento virtual a la imagen vídeo en tiempo real sin consideración del problema de la oclusión, donde: a) Edificio virtual, b) Entorno aumentado.

Diversos autores han propuesto metodologías para resolver la oclusión (ver apartado 2.3.1). En *UrbanMix*, este problema se ha solucionado mediante un modelo virtual de los edificios del campus de Vera para aplicarles un canal alfa, de forma que actúen como una máscara de los edificios virtuales del casco histórico, a través de la cual se ve la capa de vídeo del entorno real. Este proceso se ha realizado en el programa Max/MSP Jitter (ver apartado 8.2.2.3). En la siguiente figura podemos ver unas imágenes que describen el proceso del método empleado para resolver la oclusión.

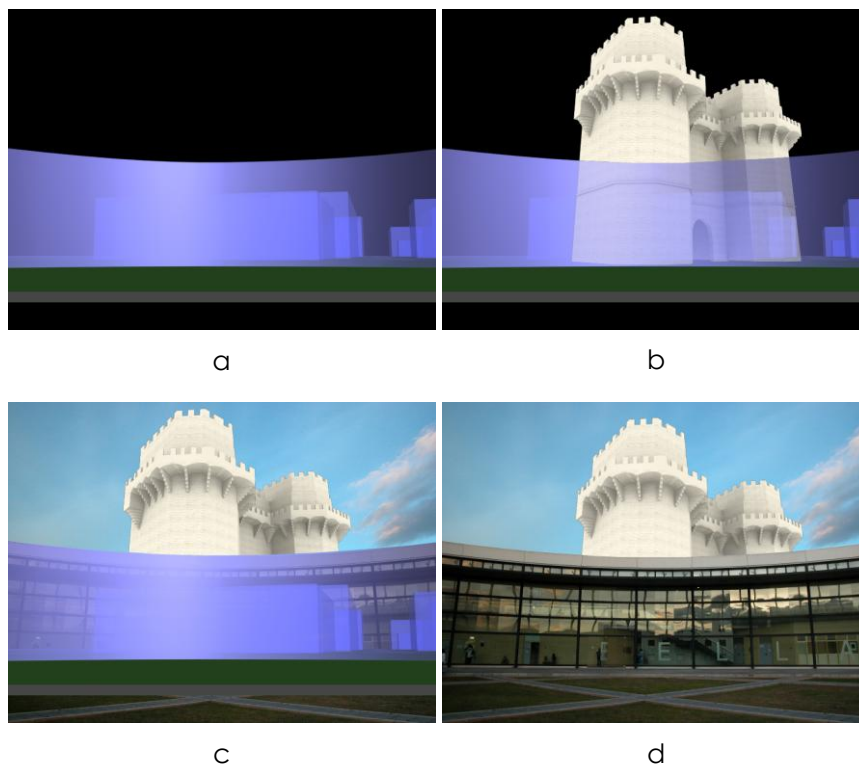


Figura 208. Superposición de un elemento virtual a la imagen vídeo en tiempo real resolviendo el problema de la oclusión, donde: a) Modelo virtual del campus de Vera, b) Modelo virtual del campus de vera junto al modelo virtual de Les Torres de Serrans integrado espacialmente, c) Objetos virtuales superpuestos a la imagen vídeo, d) Entorno aumentado con resolución de la oclusión mediante el canal alfa sobre el modelo del campus de Vera.

8.2.2.3. ENTORNO DE PROGRAMACIÓN

La Figura 209 muestra el esquema de funcionamiento de *UrbanMix*: la captura de la cámara en tiempo real tiene su entrada en el programa a través del objeto *jitter.qt.grab*, para ser posteriormente analizada mediante un tracking de color de los puntos de apoyo; esta información junto con los datos aportados por el sensor inercial entra a formar parte de las ecuaciones de la proyección central (ver Ecuación 1, apartado

2.1.3). El resultado es la matriz de transformación (posiciones y rotaciones) de la cámara en el sistema de referencia establecido, los cuales se trasladan al mundo virtual para construir la escena mediante el objeto *jit.gl.render*. Por otro lado, los modelos virtuales de los edificios se gestionan mediante el objeto *jit.gl.model*, asignándole transparencia al modelo del campus que hace la función de máscara. Todos estos procesos se organizan mediante una serie patches, habiendo un patch principal y diversos sub-patches que se detallan a continuación.

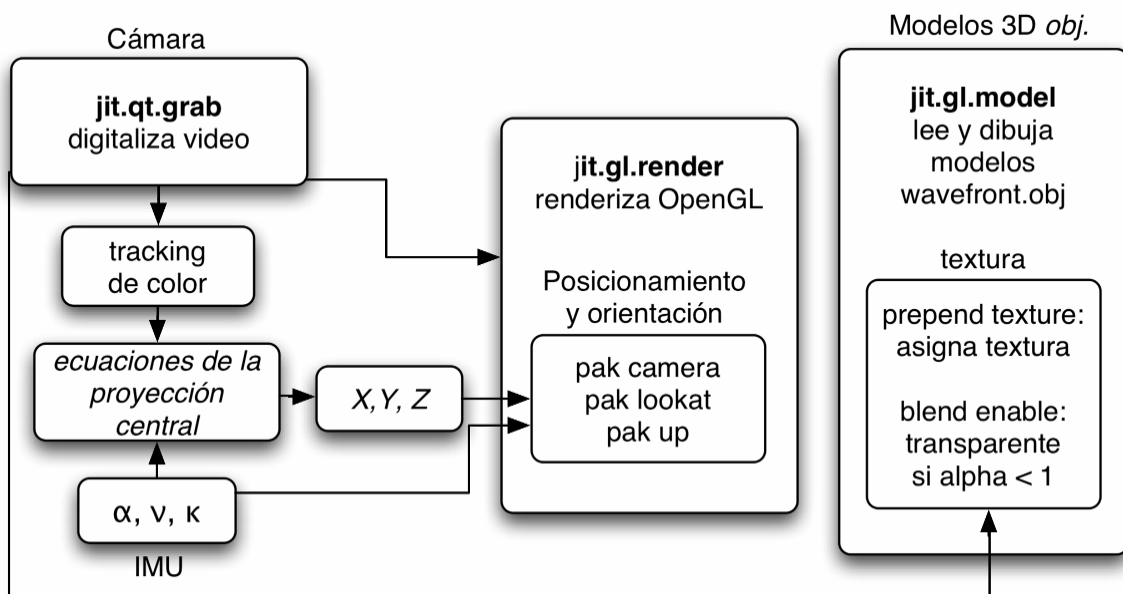


Figura 209. Arquitectura de programación.

En el patch principal de la aplicación (Figura 210) se activa la entrada de la cámara mediante la orden *open* al objeto *jit.qt.grab*; la renderización en la ventana de OpenGL se activa mediante el botón Start Rendering. También podemos visualizar la posición y orientación de la cámara mediante *pak camera*, *pak lookat* y *pak up*, descritos en el apartado 5.3.1.2; estos parámetros se actualizan en tiempo real a partir de los datos recibidos a través de *p camori*.

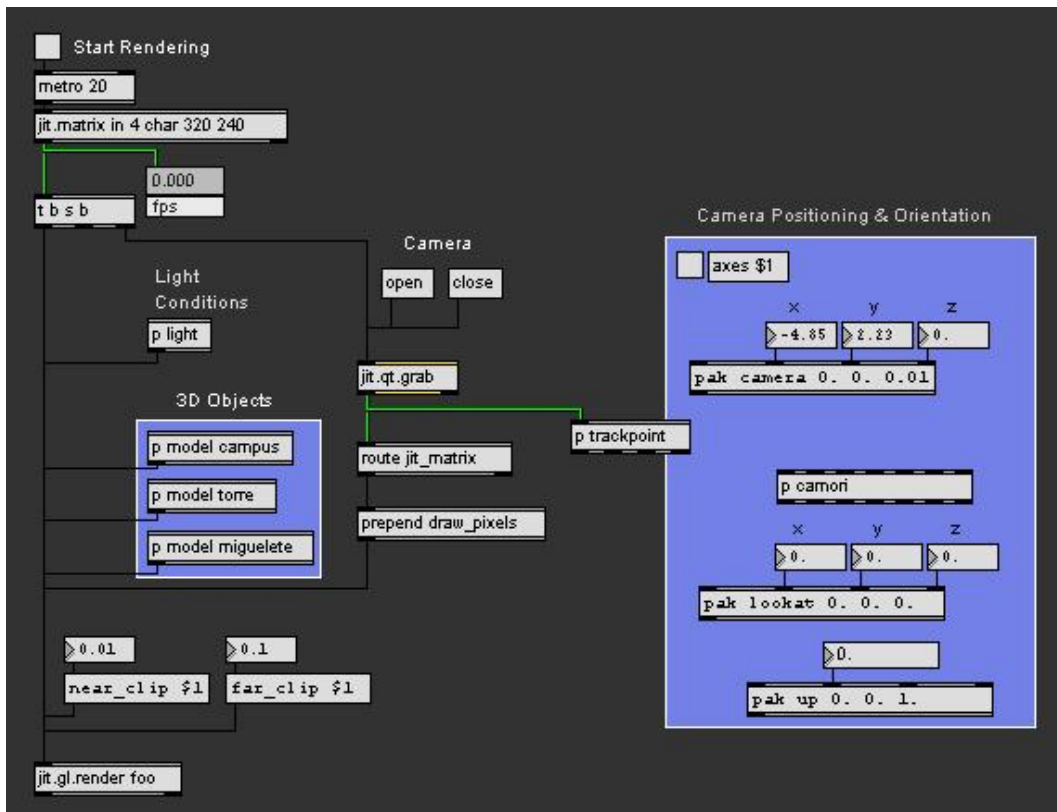


Figura 210. Patch principal de *UrbanMix*.

Mediante los sub-patches *p camori* (Figura 211-a) y *p leastsquares* (Figura 211-b) se realizan las operaciones matemáticas necesarias para el cálculo de la orientación de la cámara, tal y como se describe en el apartado 2.1.5.3. En *p camori* están las rotaciones del sensor inercial (*roll*, *pitch*, *yaw*) calculadas a partir de los datos del objeto *mt9* (descrito en el apartado 5.3.1.3), la distancia focal de la cámara, las coordenadas imagen (x_1, y_1, x_2, y_2 ; calculadas en *p trackpoint*) y terreno ($X_1, Y_1, Z_1, X_2, Y_2, Z_2$) de los dos puntos de apoyo del entorno, y las ecuaciones de la proyección central (Ecuación 1); por otro lado, *p leastsquares* contiene el procedimiento mínimo cuadrático para la resolución de las ecuaciones de la proyección central, cuyo resultado son las coordenadas tridimensionales de la cámara (X_0, Y_0, Z_0), que se envían al objeto *pak camera* del patch principal. Además, si tomamos en consideración que la coordenada Z_0 de la cámara va a permanecer estable dentro de unas tolerancias (la altura del usuario), podemos añadir este constreñimiento al sistema de ecuaciones para que el resultado sea más estable.

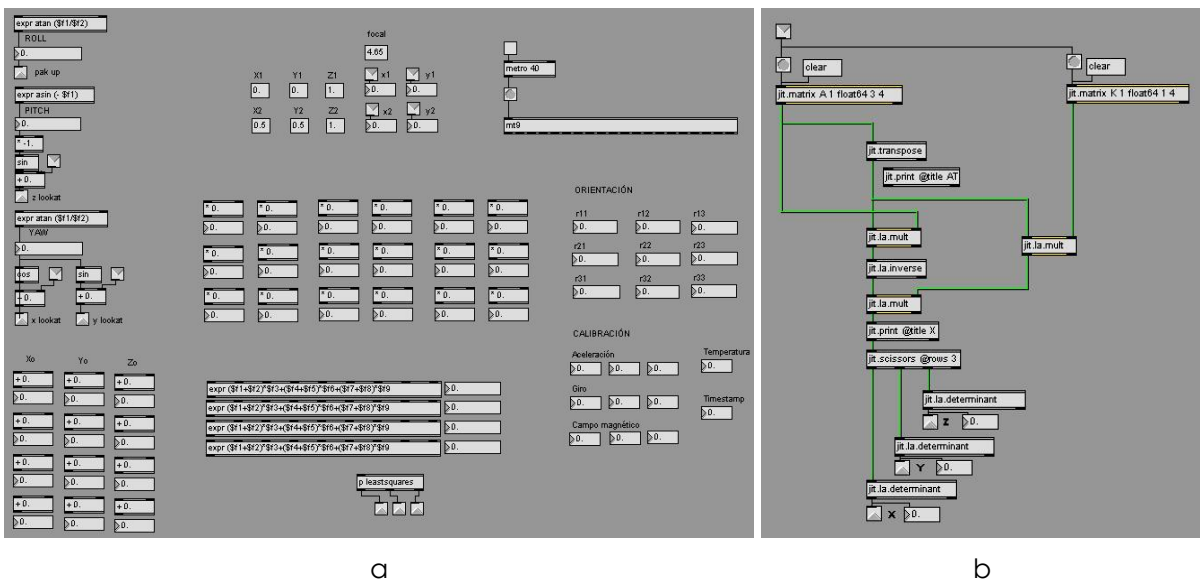


Figura 211. Sub-patches para el cálculo de la orientación de la cámara: a) *p camori*; b) *p leastsquares*.

En el sub-patch *p trackpoint* (Figura 212-a) se realiza el registro de dos colores (amarillo y magenta) a partir del objeto *tap.jit.colortrack* explicado anteriormente en el apartado 5.3.1.1; los valores de salida son las coordenadas planas de los puntos de apoyo en la imagen bidimensional (x_1, y_1, x_2, y_2), que son enviados al sub-patch *p camori* para que entren a formar parte del proceso de resolución de las ecuaciones planteadas.

Cada uno de los sub-patch *p model* (Figura 212-b) gestiona las propiedades de los modelos virtuales *obj*, tanto los que van a ser insertados en el entorno real como el modelo del espacio físico que actúa como máscara, mediante el objeto *jit.gl.model*, descrito en el apartado 5.3.1.2.

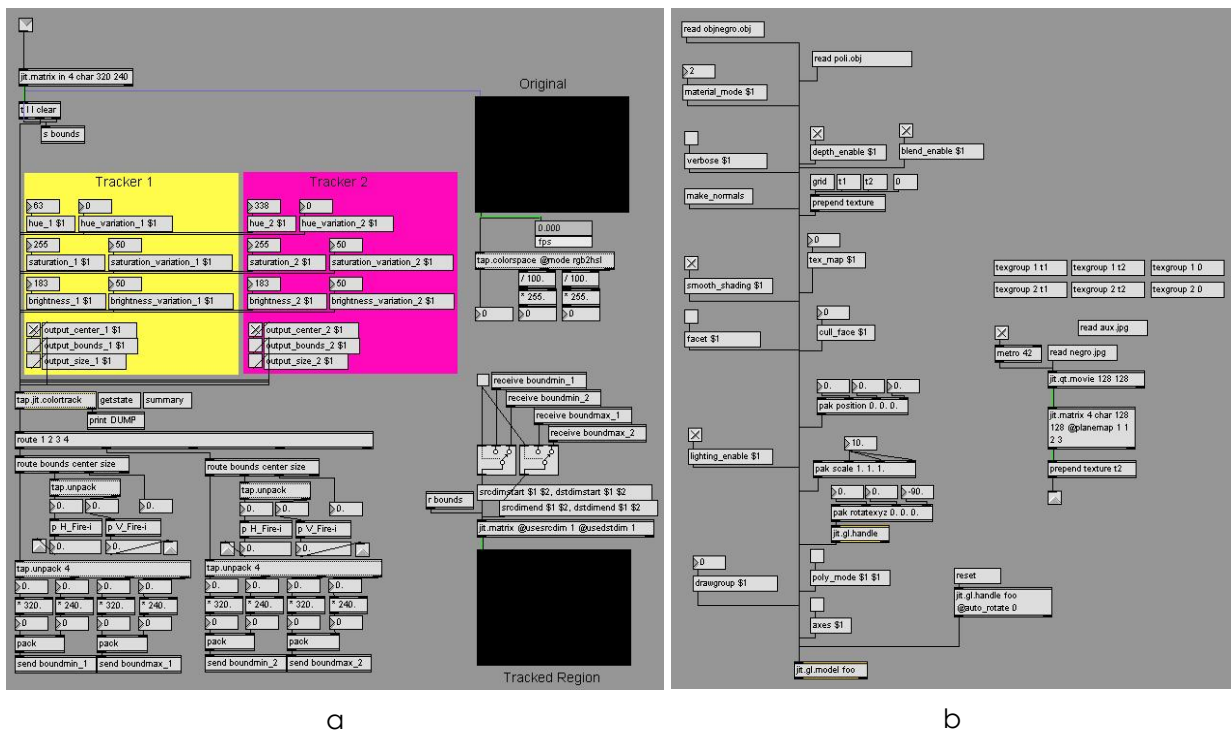


Figura 212. Sub-patches: a) *p trackpoint*; b) *p model*.

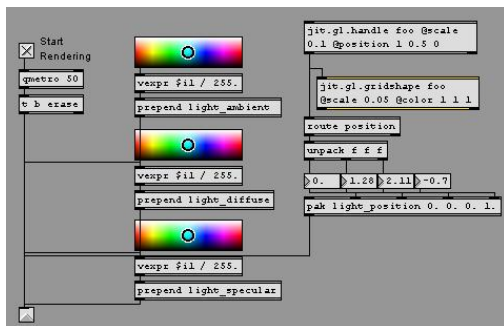


Figura 213. Sub-patch *p light*.

En el sub-patch *p light* (Figura 213) se controlan las condiciones de luminosidad en la ventana de OpenGL. Estas condiciones son: luz ambiente (*light_ambient*), difusa (*light_diffuse*) y especular (*light_specular*); posición espacial del punto de luz (*light_position*). Estas propiedades se envían al objeto *jit.gl.render* del patch principal.

8.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS

El trabajo presentado en este ensayo es técnicamente complejo, pues requiere la integración de distintos dispositivos y metodologías para el registro. Además, los edificios modelados son de complejidad media y se ha resuelto el problema de la oclusión. Todo esto hace que la gestión de las tareas necesarias para la construcción del escenario aumentado sea complicada mediante el empleo de un ordenador portátil estándar. Además, el programa Max/MSP Jitter es nativo para Macintosh, y menos estable para la versión PC. Sin embargo, el plug-in utilizado para el sensor

inercial es compatible únicamente con PC, por lo que necesariamente *UrbanMix* debe utilizarse con este sistema.

También encontramos una serie de problemas vinculados a la movilidad del usuario. Por ejemplo, el dispositivo HMD y el sensor inercial necesitan baterías para su funcionamiento, por lo que ha sido necesario diseñar un pequeño kit que debe llevar el usuario además del ordenador portátil; el peso de estos elementos, aunque no excesivo, puede restar confort y movilidad. Por otra parte, para que el entorno virtual aparezca correctamente alineado con el real, el campo de visión de la cámara debe contener los puntos de apoyo; de lo contrario, los dos espacios se desajustan.

Otro inconveniente encontrado es la utilización de un registro de color, por la necesidad de ubicar puntos de apoyo artificiales en el entorno. Es por ello que también se realizaron pruebas mediante el objeto *cv.jit.track*, descrito en 5.3.1.1, con el que no es necesario la utilización de puntos artificiales. Sin embargo el registro mediante este objeto era inestable bajo las condiciones ambientales propias de los espacios abiertos; además, el edificio de la facultad de Bellas Artes, sobre el que se realizaron las pruebas, tiene una fachada acristalada, por lo que los problemas con los reflejos de la luz solar se incrementan de forma notable.

9. HOT IMAGES

En este capítulo se presenta *Hot Images*, una pieza de virtualidad aumentada que indaga la relación entre las personas y los entornos urbanos, al mismo tiempo que propone una nueva cartografía de un entorno urbano (mediante el mapeado de imágenes personales) e interfaz de navegación (mediante un juego de colores). En el entorno virtualmente recreado en *Hot Images*, las construcciones inertes de la ciudad se acercan al campo humano, transformando su aspecto impersonal mediante el reflejo de relaciones subjetivas y afectivas.

Motivación

En el campo de la cartografía multimedia surgen diversos sistemas de navegación guiados con GPS para viandantes y/o para vehículos. La finalidad de estos nuevos mapas interactivos es guiar a sus usuarios a través de espacios urbanos, redes de carreteras, etc. Las aplicaciones de RA para espacios urbanos pueden superponer información gráfica a la imagen capturada por la cámara. Estos gráficos son principalmente indicaciones para guiar a los viandantes hacia una localización concreta a través de una ruta calculada por el sistema, por ejemplo, el camino más corto de la posición actual del usuario a su destino. Mediante la utilización de un sistema de navegación con GPS, este ensayo propone una nueva cartografía de la ciudad, mediante la generación de modelos virtuales de los edificios dotados de connotaciones humanas. Las herramientas de navegación se establecen a partir de un

juego de colores, de tal forma que no se da una indicación clara del camino a seguir. El usuario es guiado a unas localizaciones concretas sin saber de antemano su destino. La finalidad objetiva del mapa se transforma en signos subjetivos que, a modo de juego, pueden ayudar a explorar un entorno desconocido.

Marco temporal

Hot Images se ha desarrollado en el año 2005, durante una estancia en el Mixed Reality Laboratory (MRL) de la University of Nottingham, en Reino Unido.

9.1. PLANTEAMIENTO

Esta aplicación la desarrolle durante mi estancia en el Mixed Reality Laboratory de la Universidad Técnica de Nottingham (UK). Allí estuve viviendo en una típica casa inglesa de protección oficial del barrio Beeston, cuyas calles y casas mantienen un aspecto bastante uniforme, por lo que resulta sencillo desorientarse y perderse. Se puede decir que el barrio tiene poca *imageability*, término introducido en (Lynch, 1960) que lo describe como la cualidad de un objeto físico, que puede evocar en cualquier observador una imagen viva, fuerte. De este modo, una ciudad con alta "imagenibilidad" sería aquella bien formada, con partes distintivas, e instantáneamente reconocible por cualquier habitante. Beeston representa todo lo contrario: topónimos de calles mal indicados, autopistas que cruzan la ciudad, caminos que no llevan a ninguna parte, gran similitud de las calles (las casas fueron construidas de acuerdo a una normativa rigurosa del ayuntamiento, por lo que todas se parecen sustancialmente), poca iluminación nocturna, autobuses (con el mismo número) que dependiendo del horario tienen rutas distintas, etc. Los hallazgos de Lynch se utilizan actualmente en la planificación urbanística de muchas ciudades, aunque todavía se sigue investigando en temas relacionados con la cognición espacial.



Figura 214. Imágenes de Olton Av., Beeston.

Las bajas temperaturas en invierno propician que las calles permanezcan casi desiertas, especialmente por la tarde/noche, desarrollándose las principales actividades diarias en espacios cerrados. La ciudad refleja un ambiente frío e impersonal, casi inerte, tal vez por ello surgió la idea de aproximar la ciudad al ámbito humano, debido a la necesidad personal de dotar a los alrededores de un toque distinto que cambiase la/mi percepción de la ciudad.

La necesidad de incrementar la imagenibilidad de un entorno visualmente frío, aproximándolo al entorno humano se aborda con la tecnología de la realidad

aumentada mediante: el establecimiento de una nueva cartografía aumentada personalizada y la navegación por el entorno aumentado mediante el juego de caliente-frío.

Cartografía aumentada

“El mundo es complejo, dinámico, multidimensional; el papel es estático, plano. ¿Cómo vamos a representar el rico mundo visual de experiencia y dimensión sobre un simple plano?” (Tufte, 1990).

En los últimos años el soporte de los mapas ha evolucionado desde el papel a la pantalla de ordenador, pasando de la representación bidimensional a la tridimensional, de la cartografía tradicional a la denominada cartografía multimedia (Cartwright et al., 1999). Hoy en día podemos representar datos geoespaciales a partir de modelos 3D e incluso visualizar fenómenos dinámicos mediante simulaciones. Las técnicas emergentes de realidad virtual y aumentada han enriquecido los mapas, proporcionando al mismo tiempo nuevas herramientas para la navegación de los usuarios.

Una de las funciones primordiales de un mapa urbano es la de orientar a los usuarios dentro de la ciudad, mostrando la disposición urbanística del lugar, con topónimos de sus calles o avenidas principales y resaltando generalmente los edificios o zonas singulares. En este sentido, la cartografía ha evolucionado con la incorporación de receptores GPS de bajo coste a diversos dispositivos (que contienen un mapa base en formato digital), dando lugar a los denominados LBS (*Location Based Services*) y los sistemas de navegación para viandantes (*pedestrian navigation systems*). Algunos trabajos en estos campos se describen en (Gartner, 2004; Retscher et al., 2004).

Existen múltiples tipos de mapas con distintas características y finalidades, por ejemplo, los mapas geológicos representan los estratos del terreno con sus características de interés geológico; los mapas físicos representan factores geográficos, como las cadenas montañosas, ríos, lagunas, etc.; los mapas topográficos representan la forma del terreno mediante curvas de nivel. A la hora de confeccionar un mapa (que es siempre una representación de una parte de la realidad física) es necesario saber a quién va dirigido y para qué finalidad; dependiendo de estos factores, la representación y simbolización empleadas serán diferentes. En este sentido surge la

cuestión: ¿sería posible realizar un mapa personalizado de la ciudad, cuya finalidad sea acercar el entorno urbano al ámbito personal del usuario?

La aplicación *Hot Images* se sirve de los sistemas de navegación para viandantes y la tecnología de realidad aumentada para establecer una nueva cartografía de la ciudad, proponiendo un sistema de navegación guiado a través del juego de caliente-frío.

Juego caliente-frío

Como ya se ha dicho, *Hot Images* es una aplicación de virtualidad aumentada en la que se confrontan dos espacios diferentes: el espacio exterior/público, impersonal; y el espacio privado, personal. Las experiencias personales se mezclan con el entorno urbano mediante una serie de imágenes de eventos pasados relacionados con el usuario. Estas imágenes están ocultas (virtualmente) en ciertas localizaciones (puntos calientes) del entorno considerado. El usuario tiene que explorar dicho entorno para poder encontrar las imágenes y recuperarlas, guiándose por las indicaciones del sistema de navegación basado en colores hacia los puntos calientes, de acuerdo al juego infantil de caliente-frío. Este juego consiste en encontrar ciertos objetos que están escondidos; la persona que ha escondido los objetos guía al usuario a través de las palabras “caliente” o “frío”; la segunda significa que el buscador está lejos de encontrar el objeto, mientras que la primera significa que se está aproximando a él. De igual forma, en *Hot Images* se establece una relación de cercanía-lejanía a las imágenes escondidas, pero en lugar de emplear las palabras clave, éstas se corresponden con los colores azul (frío, lejanía) y rojo (caliente, cercanía), estableciendo así un juego visual. Cuando el usuario llega a un punto caliente, la imagen escondida en esa localización pasa a formar parte de las fachadas de los modelos virtuales correspondientes a edificios de su entorno. La última finalidad del juego es convertir el espacio urbano frío e impersonal en un lugar cálido, que resulte familiar y agradable para el usuario.

9.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO

Siguiendo las aportaciones del bloque I, se puede elaborar un marco general de este ensayo:

- Ubicación dentro del continuo de Milgram (apartado 1.1.1):
 - Virtualidad aumentada.
- Clasificación (apartado 1.2.2):
 - Según el entorno físico en el que se desarrolla la aplicación:
 - Abierta.
 - Según la extensión que abarca:
 - Local.
 - Según la movilidad de los dispositivos de registro y/o displays:
 - Móvil.
 - Según los usuarios que simultáneamente pueden interactuar con el sistema:
 - Individual.
 - Según el tipo de colaboración establecida:
 - (No existe colaboración).
 - Clasificación según Mackay (apartado 1.2.2):
 - Aumentar el usuario.
- Campo de aplicación (apartado 1.4):
 - Arte.
 - Entretenimiento.
- Sistema de coordenadas terreno/objeto (apartado 2.1.1):
 - Terrestre global.
- Orientación espacial (apartado 2.1.5):
 - Método directo.
- Dispositivos de registro / controladores (capítulo 3):
 - Receptor GPS.
 - Sensor inercial.
- Displays (capítulo 4):
 - Display de mano.
- Software (capítulo 5):
 - Multimedia: Max/MSP Jitter.
- Forma de interactuar (apartado 6.1.1):
 - Navegación.

9.1.2. REFERENTES

En la instalación interactiva *Field-Work@Alsace* (2002) de Masaki Fujihata (Figura 215) se establece una representación de las dimensiones espaciales y temporales mediante una serie de videoimágenes móviles conectadas entre sí dentro de un entorno virtual 3D. Las imágenes corresponden a distintos lugares de Alsace y están dotadas de una georreferenciación mediante el empleo de un receptor GPS y una brújula electrónica. A partir de estos datos Fujihata construye el entorno virtual al ubicar las imágenes en la línea trazada por el recorrido generado con los sensores.



Figura 215. Field-Work@Alsace, en (Fujihata, 2002).

En cuanto a trabajos de realidad aumentada en el ámbito de los LBS podemos citar varios autores, como (Höllerer et al., 1999; Reitmayr et al., 2003; Kolbe, 2004). En (Höllerer et al., 1999) se describe uno de los primeros (sino el primer) LBS de realidad aumentada (ver apartado 8.1.2). En (Reitmayr et al., 2003) se elabora un sistema para ser utilizado en espacios cerrados mediante un registro con patrones de marcas planas. El usuario selecciona la habitación hacia donde quiere ir mediante la selección de ésta en una representación virtual de la planta del edificio o seleccionándola de entre una lista de habitaciones. La aplicación calcula el camino más corto desde su posición actual a la habitación para, a continuación, guiarlo a través de él. Como podemos ver en las imágenes (Figura 216), las indicaciones que guían al usuario a través de los pasillos del edificio son unos símbolos en forma de flecha.



Figura 216. Sistema aumentado de LBS en interiores. En (Reitmayr et al., 2003).

De otro modo, en (Kolbe, 2004) se presenta una LBS de realidad aumentada para guiar a los usuarios en espacios abiertos. El dispositivo utilizado como display es en este caso una PDA o un teléfono móvil. La novedad reside en que las imágenes vistas en el display son vídeos pregrabados aumentados con información relevante para la navegación, ejecutándose de acuerdo con la localización espacial y orientación del usuario, por lo que los vídeos están grabados siguiendo los dos sentidos de cada calle. Además, en las intersecciones de las calles, se han realizado imágenes panorámicas de 360° (también con indicaciones para la navegación) que son visualizadas de acuerdo con la orientación del usuario. Como podemos ver en la Figura 217, las indicaciones del camino que debe seguir el usuario se hacen a través de flechas. Sin lugar a dudas este símbolo es el más empleado en los LBS para indicar el camino a seguir, puesto que es sencillo y fácilmente reconocible.



Figura 217. Sistema aumentado de LBS en una zona urbana. En (Kolbe, 2004).

9.2. DESCRIPCIÓN TÉCNICA

9.2.1. COMPONENTES DE LA APLICACIÓN

9.2.1.1. DISPOSITIVOS FÍSICOS

El escenario de *Hot Images* se compone de los siguientes dispositivos y software (Figura 218):

- Un ordenador portátil o tablet-PC que el usuario lleva consigo en las manos, con el software Max/MSP Jitter instalado.

- Una cámara Web mediante la cual el sistema reconoce una serie de elementos del entorno físico que están vinculados a las imágenes. Se ha utilizado una cámara Logitech con una resolución de 640x480 píxeles y conexión USB 2.0.
- Un receptor GPS para establecer la localización planimétrica (X, Y) del usuario en el entorno físico, que se corresponde a la vez con el entorno virtual. Se ha utilizado el receptor Trine II, descrito en el apartado 3.6.1.
- Un sensor inercial para establecer la orientación del usuario. También es posible establecer las inclinaciones, pero debido a problemas de estabilidad esta opción se ha omitido. El sensor empleado es el MT9 de Xsens, descrito en el apartado 3.5.



Figura 218. Dispositivos de *Hot Images*.

9.2.1.2. ELEMENTOS VIRTUALES

La aplicación *Hot Images* se desarrolló en el Jubilee Campus de la Universidad Técnica de Nottingham, y por ello el modelo virtual utilizado corresponde a dicho campus (Figura 219). Este modelo fue realizado por personal del Mixed Reality Laboratory mediante el programa 3D Studio Max. Para su utilización en la aplicación *Hot Images* se hicieron algunos cambios: para simplificar el modelo se eliminó información no necesaria (como elementos de detalle y árboles), se agruparon los elementos y se asignó en 3D Studio Max un mapeado diferente para cada grupo y se exportó el resultado a formato *obj*. Finalmente, las texturas fueron cargadas en Max/MSP Jitter, desde donde se mapea el modelo a medida que el usuario va encontrando las imágenes.

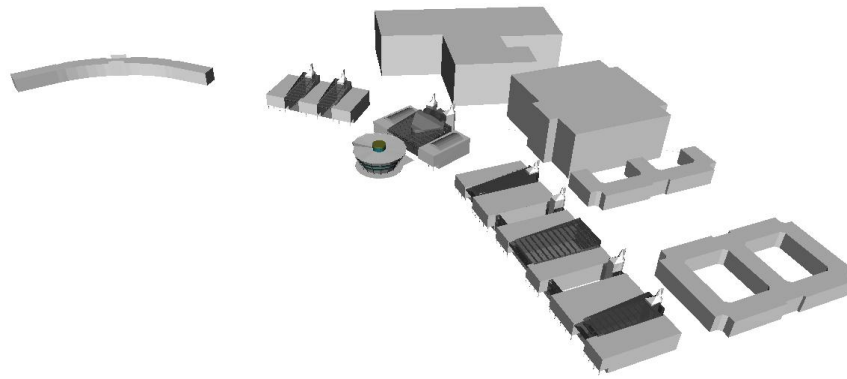


Figura 219. Modelo Virtual simplificado del Jubilee Campus.

La aplicación puede albergar distintos contenidos o temáticas dependiendo del tipo de audiencia y finalidad buscada, como por ejemplo: imágenes sobre la historia de la ciudad o un evento pasado relacionado con el entorno establecido, si la aplicación tiene la finalidad de memoria histórica y va dirigida en general a los ciudadanos o transeúntes de una ciudad, estableciendo a la vez relaciones espacio-temporales; imágenes relacionadas con la naturaleza si la finalidad es denunciar un abuso urbanístico no sostenible, sobre todo en zonas donde el ecosistema es frágil; etc.

9.2.2. DESARROLLO, PROCESO

9.2.2.1. INTERFAZ DE NAVEGACIÓN

La interfaz de navegación propuesta en *Hot Images* es a través de una graduación de colores que abarcan desde el rojo al azul, pasando por las tonalidades intermedias que surgen. Escondidas en el entorno, se encuentran algunas imágenes relacionadas con el usuario, en localizaciones concretas, a las que hemos denominado "puntos calientes". Dependiendo de la proximidad del usuario a uno de estos puntos calientes los edificios virtuales se verán con el correspondiente color: 100% rojo cuando esté a menos de 1 m; 100% azul cuando esté alejado más de 50 m. Cuando el usuario llega a un punto caliente, deberá localizar el lugar exacto con la cámara, la cual realiza un registro de la imagen del entorno en busca de un patrón concreto. Al conseguir una correspondencia, la imagen guardada en esa localización pasa a formar parte de entorno, mapeándose en la fachada de algún edificio, con este cambio de aspecto el entorno pasa a tener connotaciones humanas. Las relaciones espaciales y conceptuales que surgen a través de la interfaz propuesta se resumen en la siguiente figura:

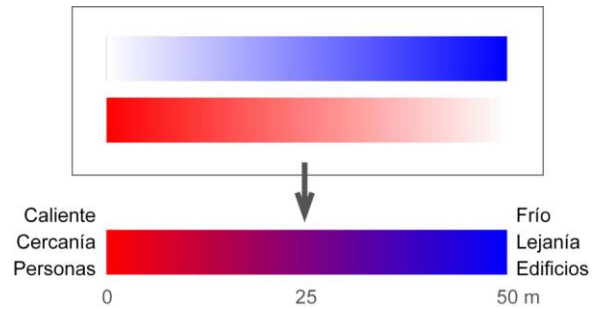


Figura 220. Interfaz de navegación y relaciones conceptuales.

Visto en planta, estos colores forman un mapa predominantemente azul con zonas circulares cuyo centro es rojo, y a medida que se aleja de éste se degrada hacia el azul. A continuación se muestra un ejemplo del mapa que forman cinco puntos calientes sobre una vista aérea del campus (Figura 221).



Figura 221. Mapa de cinco puntos calientes en el campus.

9.2.2.2. SISTEMA DE REFERENCIA

El sistema de referencia terrestre global en el que se basa el GPS, con sus correspondientes coordenadas polares (*longitud, latitud*), se ha transformado en un sistema de referencia terrestre local cartesiano, con coordenadas (X, Y). A continuación se muestra un ejemplo de cálculo:

- En primer lugar, fijamos el origen de coordenadas en un punto situado a la parte izquierda inferior de la zona (Figura 222), como por ejemplo en:

$$\varphi_0 = 52^{\circ}57'00.19''N = 52.95005278^{\circ}N \rightarrow X_0 = 0$$

$$\lambda_0 = 1^{\circ}11'32.53''W = 1.19236944^{\circ}W \rightarrow Y_0 = 0$$

- El punto para el cual queremos calcular las coordenadas planas es A:

$$\varphi_A = 52^\circ 57' 13.30'' N = 52.95369444^\circ N$$

$$\lambda_A = 1^\circ 11' 18.39'' W = 1.188441667^\circ W$$

- Y la latitud media de la zona

$$\varphi = 52^\circ 57' 08.96''$$

- El cálculo de las coordenadas planas de A se realiza a partir de la Ecuación 4 y la Ecuación 5, resultando en:

$$X_A = 263.989 \text{ m}$$

$$Y_A = 405.263 \text{ m}$$



Figura 222. Sistema de referencia terrestre local, donde O es el origen de coordenadas (0, 0) y A representa un punto cualquiera con coordenadas (X_A , Y_A).

En estos cálculos se han considerado únicamente las coordenadas planas (X , Y), ya que el valor de la coordenada Z permanece constante: aproximadamente a 1.6 m del suelo en la vista 3D, y a una altura de unos 10 m sobre el suelo en la vista 2D.

9.2.2.3. ENTORNO DE PROGRAMACIÓN

El entorno de programación de *Hot Images* se ha hecho con el software multitarea Max/MSP Jitter, introducido en el apartado 5.3.1. En la Figura 223 se puede ver un esquema de la arquitectura de programación. Como se ha explicado anteriormente,

los dispositivos externos son un receptor GPS, un sensor inercial y una cámara Web. Los dos primeros realizan el registro del usuario dentro del entorno real mientras que la cámara Web se utiliza para un registro visual de una escena en concreto. Los datos proporcionados por estos tres dispositivos se gestionan a través del programa, el cual consta de un patch o ventana principal y una serie de sub-patches.

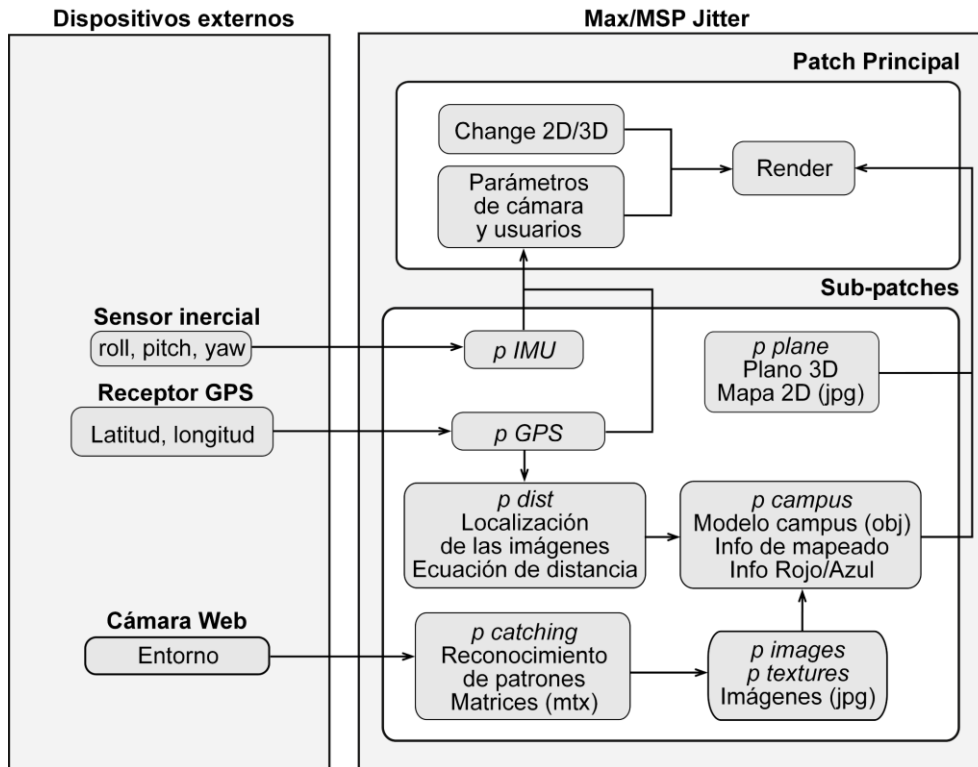


Figura 223. Esquema de la arquitectura de programación.

El patch principal de la aplicación consiste en una GUI que contiene los siguientes objetos/elementos (Figura 224-a):

- **Change 2D/3D:** Permite cambiar en cualquier momento la forma de visualización del entorno, que puede ser en 2D o en 3D. En la vista 2D, la cámara virtual se sitúa en lo alto, justo encima de la localización del usuario y apuntando hacia éste (hacia el nadir). De esta forma el entorno virtual se ve a modo de mapa centrado en la localización del usuario, siendo su avatar visible en el centro de la porción del mapa visualizado. La ventaja reside en que el usuario tiene una vista más amplia del entorno en el que se mueve, ya que es capaz de ver los edificios que están fuera de su campo de visión. La información referente a la orientación puede ser activada o desactivada; en el segundo caso, el mapa permanecerá siempre orientado hacia el norte; en el primer caso la orientación del mapa coincidirá con la del usuario, resultando en un mapa

(conocido como) egocéntrico. Por otro lado, en la vista 3D la cámara virtual coincide (aproximadamente) con el punto de vista del usuario, por lo que la información de la orientación permanece siempre activada. En este caso el usuario no ve su avatar, sino lo que éste ve.

- *Parámetros de la cámara y del usuario:* La localización del usuario se establece con las coordenadas planas (X , Y) y el ángulo de orientación (azimut). Adicionalmente, también puede considerarse (en la vista 3D) la información de las inclinaciones dadas por el sensor inercial, aunque esta opción puede resultar un poco desconcertante debido a los cambios constantes de inclinación que se producen en el movimiento del usuario al andar, resultando en una visualización inestable.
- *Render:* Aquí se define la ventana de OpenGL en la que se visualiza el entorno virtual aumentado (Figura 224-b).

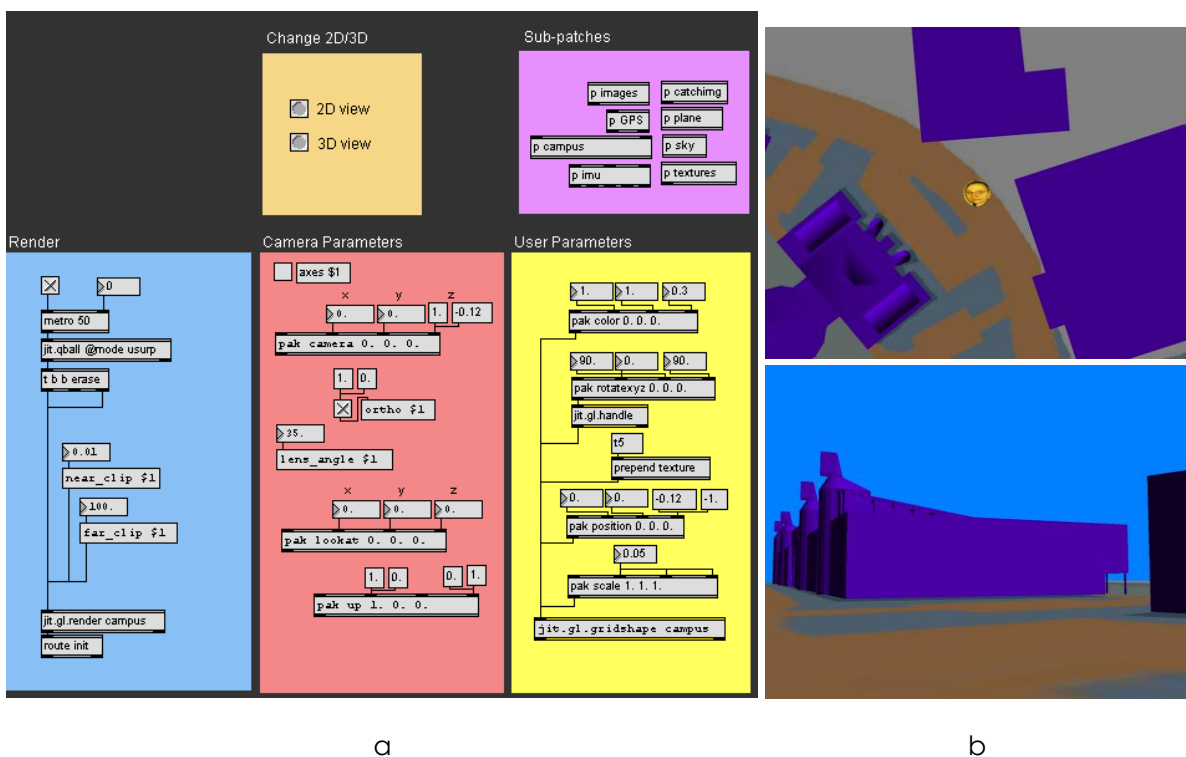


Figura 224. a) Patch principal de *Hot Images*; b) Ventana de OpenGL en vista 2D y 3D.

Los sub-patches son una serie de ventanas secundarias en las que se encuentran diversos objetos necesarios para el funcionamiento de la aplicación. Estas ventanas no son visibles al usuario en el entorno de *Hot Images*. Estos sub-patches son:

p IMU (Figura 225): Recibe los nueve parámetros de la matriz de rotación del sensor inercial. A partir de estos valores, se calculan los ángulos de rotación (roll, pitch, yaw),

que son enviados a los objetos de control de la cámara (*pak lookat* y *pak up*) en el patch principal.

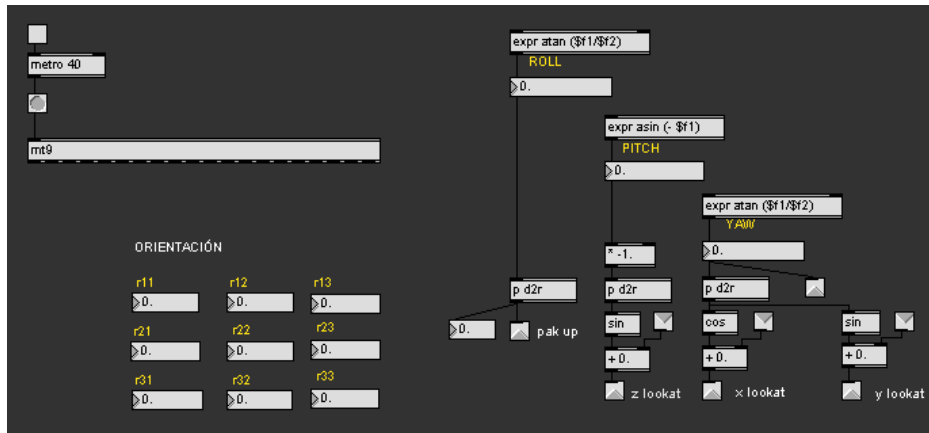


Figura 225. Sub-patch *p IMU*.

p GPS (Figura 226-a): Para la lectura de los valores registrados por el receptor GPS se ha utilizado en este caso el objeto de (Bleecker, 2005) programado con Processing, introducido en el apartado 5.3.1.3. Este sub-patch recibe la información del receptor GPS a partir del objeto *mxj jk.link GPS2max*, que sirve de nexo entre el programa Processing y Max/MSP Jitter. A partir de las coordenadas de latitud y longitud, calcula las correspondientes coordenadas (X, Y) del usuario – de acuerdo a los cálculos detallados en el apartado 9.2.2.2 – y las envía al objeto *pak position* de la cámara en el patch principal. En el programa Processing debe llamarse desde el código a las librerías *MaxLink*, y detallar los valores de salida hacia Max/MSP mediante *link.output* (Figura 226-b).

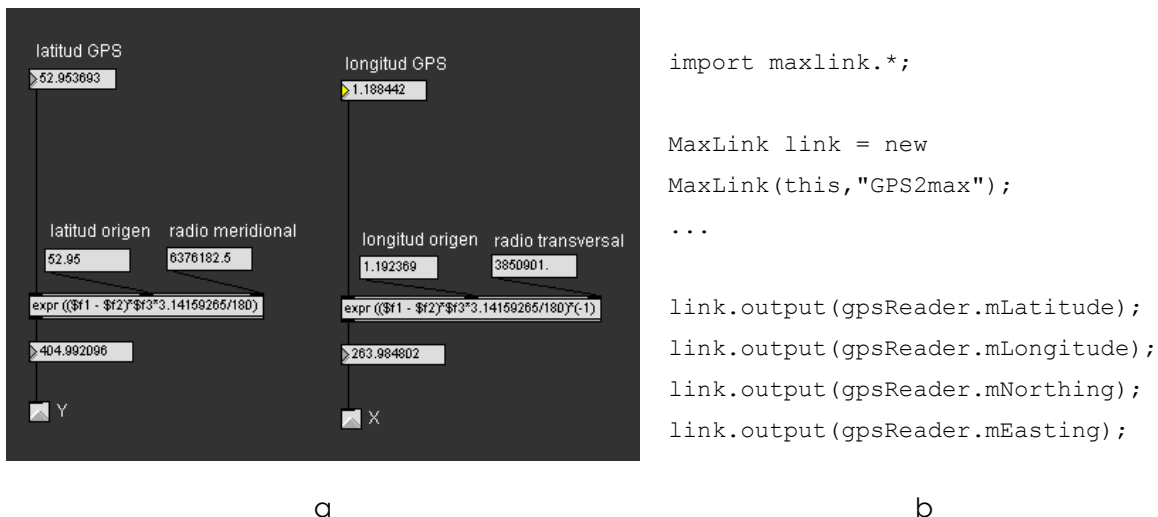
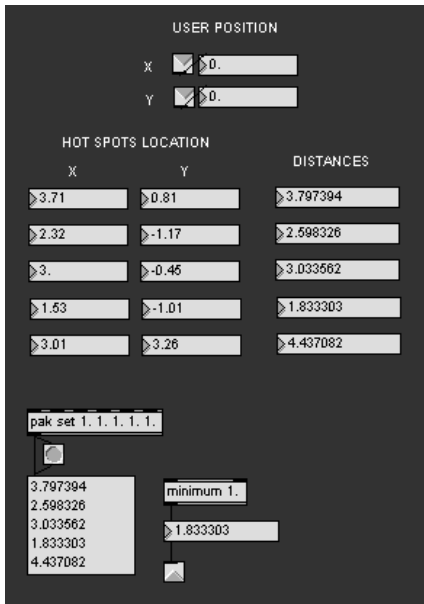
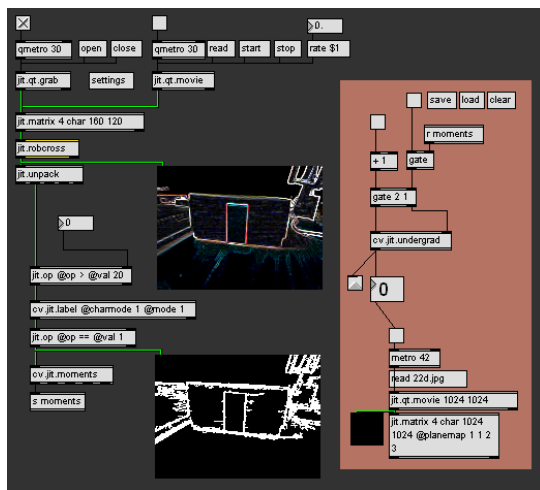


Figura 226. Cálculo de la posición del usuario: a) Sub-patch *p GPS*; b) Código insertado en el objeto programado por (Bleecker, 2005) en Processing.



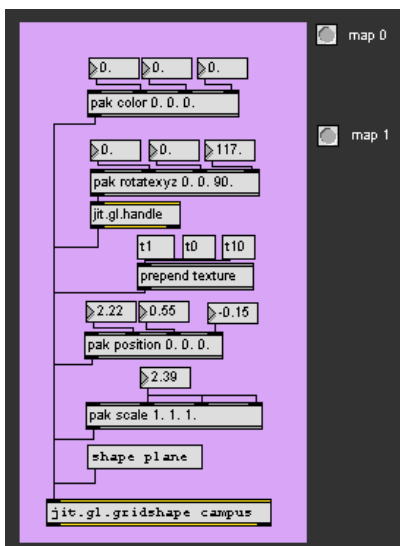
p dist (Figura 227): Realiza el cálculo de la distancia entre el usuario y los puntos calientes. La menor distancia se transforma en el color (del azul al rojo) que tendrán los edificios en ese momento. Cuando el usuario está a más de 50 m de cualquier punto caliente, los edificios se ven completamente azules; a menores distancias, se va disminuyendo el valor del azul y aumentando el valor del rojo. El resultado funciona como un mapa de buffers circulares (Figura 221).

Figura 227. Sub-patch *p dist*.



p catching (Figura 228): Se aplica el algoritmo de reconocimiento de patrones a la imagen capturada por la cámara Web, en el momento en que el usuario llega a un punto caliente, mediante el objeto `cv.jit.undergrad`. Esta imagen se compara con otra imagen guardada previamente en formato matricial *mtx*; cuando se halla una correspondencia, la imagen escondida en dicha localización pasa a formar parte del entorno aumentado.

Figura 228. Sub-patch *p catching*.



p plane (Figura 229): Dibuja un plano en la ventana de OpenGL a partir del objeto `shape plane`. Este plano constituye la figura llana del terreno, que puede ser mapeada según dos texturas distintas (`map 0`, `map 1`). La orientación y posición del plano quedan definidas mediante los objetos `pak rotatexyz` y `pak position`.

Figura 229. Sub-patch *p plane*.

p campus (Figura 230): Contiene el objeto *jit.gl.render* que carga el modelo 3D del entorno en formato *obj* mediante la orden *read*. En esta ventana se controlan diversos parámetros del modelo, como la orientación, posición, escala o texturas. Las texturas (*t101*, *t102*, etc.) se mapean en diversas partes del modelo a partir de unos grupos establecidos (*texgroup*). Al comienzo del juego, los edificios carecen de textura; al finalizarlo, todos ellos aparecen mapeados con las distintas imágenes encontradas por el usuario. Además, los modelos de los edificios aparecen con tonalidades que van desde el azul al rojo, dependiendo de la distancia que separa al usuario de algún punto caliente.

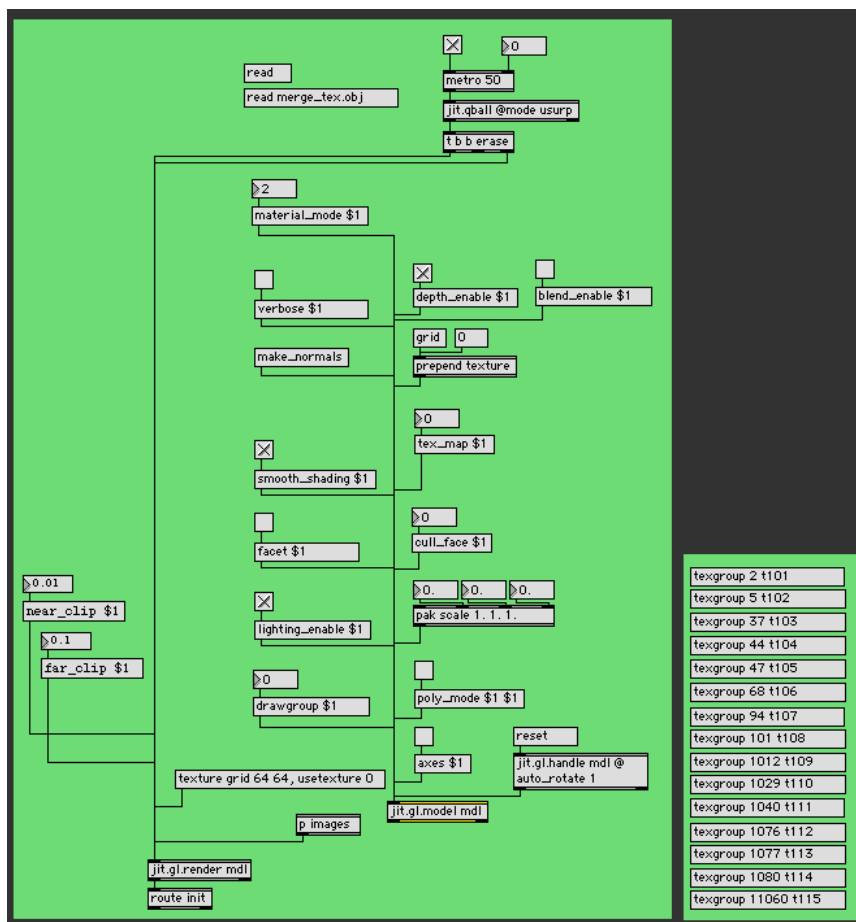


Figura 230. Sub-patch *p campus*.

p images (Figura 231): En este sub-patch se cargan las imágenes a medida que el usuario las va encontrando mediante el objeto *read*, y se asignan a una textura (*t101*, *t102*, etc.) mediante el objeto *prepend texture*. Estas imágenes están guardadas en una carpeta.

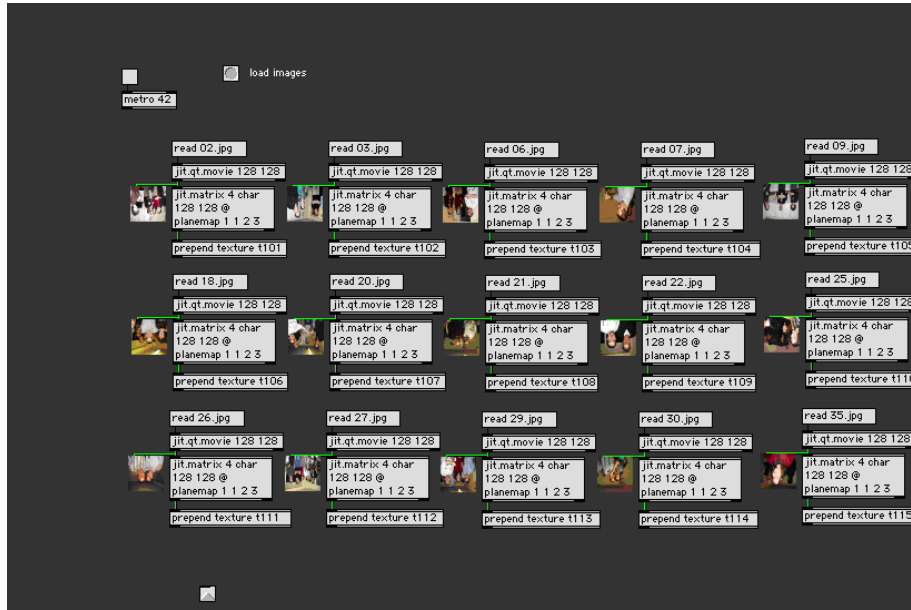
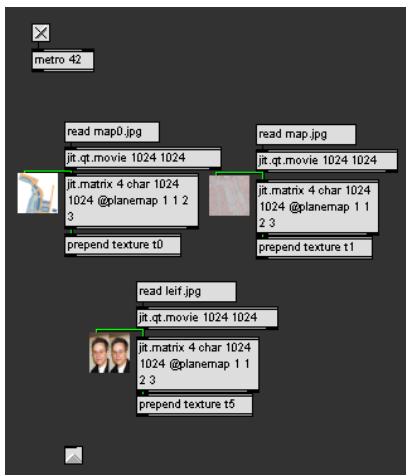


Figura 231. Sub-patch *p images*.



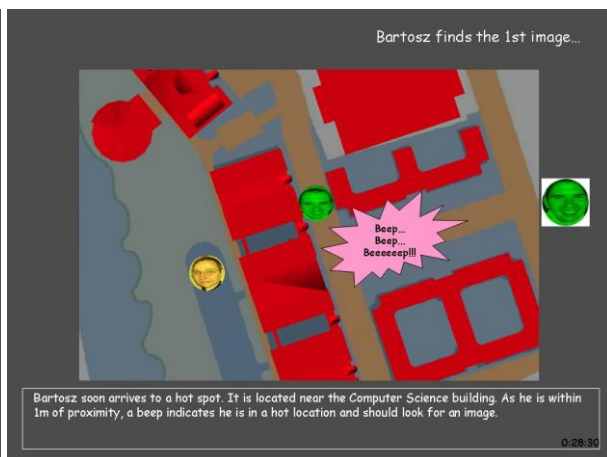
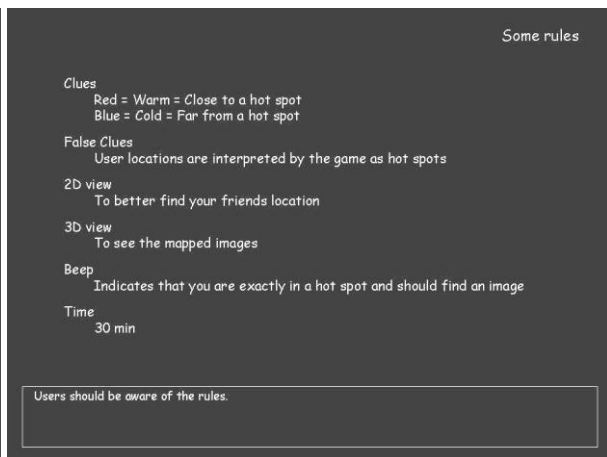
p textures (Figura 232): En esta ventana se cargan todas aquellas texturas empleadas en el juego, a excepción de las imágenes de los puntos calientes (cargadas en *p images*). Estas texturas son: diversos mapas de la distribución urbanística del terreno y las imágenes de los usuarios.

Figura 232. Sub-patch *p textures*.

9.2.2.4. STORYBOARD

A continuación se presenta una simulación del juego a modo de storyboard (Figura 233). Esta simulación se realizó en el MRL antes de tener el proyecto finalizado, con la intención de presentarlo al resto de investigadores y a investigadores invitados, en una de las reuniones que semanalmente tenían lugar cada viernes con la finalidad de difundir entre los miembros del grupo el trabajo realizado individual y colectivamente. Como podemos ver, en un principio *Hot Images* estaba pensado como un juego colectivo, en el que tres participantes debían encontrar las imágenes de una fiesta antes de que se agotara el tiempo (30 min). El color rojo significa cercanía a una imagen o a uno de los jugadores, dando lugar en el segundo caso a una pista falsa. Además, la cercanía a un punto caliente se refuerza mediante un *beep*, que indica

que el usuario ha llegado a un punto caliente y debe buscar la imagen escondida. En el storyboard se recogen distintos casos que se pueden producir durante el juego, como que un jugador pierda la señal del GPS, que el sensor inercial se vea afectado por un campo magnético, que las condiciones de luminosidad ambientales no sean las apropiadas para el reconocimiento de patrones mediante la cámara Web, e incluso que uno de los jugadores es atacado por una de las muchas ocas que allí había (hecho real). A pesar de los imprevistos, los jugadores consiguen finalmente su premio: recuperar las imágenes de la fiesta. En la elaboración de este storyboard colaboraron tres investigadores: Macarena Gómez, Leif Oppermann y Bartosz Wietrzyk.



... gets it...

From his location, he makes a round, hoping the WebCam detects the hidden image. When the WebCam points to the right location, a pattern recognition algorithm detects the place (shows a 7), the image is shown...

0:28:15

... and is mapped

... and is mapped in some of the surrounding virtual buildings. Bartosz changes to the 3D visualisation, to see the image. Macarena and Leif do the same, as they can also see the mapped image.

0:28:00

False clues...

Macarena continues navigating within the 3D view. She is getting closer to a hot spot, as the virtual environment is turning to red...

0:25:00

Be aware at your friends location!

Unfortunately, she meets Leif just when turning a corner. This means they were following a false clue... they change to the 2D visualisation.

0:24:40

Lost GPS signal

Macarena walks in between two buildings. There are also some trees. These makes the GPS signal to be lost, and the positions given by the single IMU device highly drift in few seconds, giving her a completely wrong location. She goes away to recover the signal... Hope there is no a hot spot at that location.

0:20:15

Wrong INS signal

Bartosz is getting closer to the 3rd image, which seems to be somewhere near the parking. When passing next to the cars, the IMU Heading values seems to become crazy. This is due to the iron content of cars that distort the magnetic field, thus making the heading values drift. Bartosz decides to change to 2D view.

0:15:05

Hurry up!

Macarena is really enjoying the game. She finds the images very nice and wants to look at them. Suddenly she aware the time is fast running, and she only has found two images... hurry up!

0:10:00

Bad weather

Leif has arrived to his 4th hot spot. He tries to find the image, but the pattern recognition algorithm seems not to find it, maybe because of poor light conditions... He keeps trying...

0:09:00



Figura 233. Storyboard.

9.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS

La precisión planimétrica del receptor GPS utilizado está en torno a los 3 m, mientras que la precisión del sensor inercial es de 1°. Estas precisiones, que a priori son insuficientes para establecer una correcta alineación entre el entorno real y el virtual, son sin embargo adecuadas para la aplicación propuesta al establecer un entorno principalmente virtual, en el que el mundo real aparece como imagen fija dentro de

los límites de las fachadas de los edificios, por lo que el posicionamiento del usuario es necesario únicamente de modo aproximado.

Podemos afirmar que el display utilizado no parece el más óptimo para aplicaciones en espacios abiertos, puesto que la luminosidad que suelen ofrecer los ordenadores portátiles estándar no contrarresta la luz ambiental, surgiendo la necesidad de adaptar un dispositivo a modo de visera, como en (Wilde et al., 2003). De otro modo, como la localización del usuario no es a través de un dispositivo óptico, éste podría ser utilizado en horarios de poca luminosidad, incluso de noche, facilitando la visibilidad del display. En este último caso, los puntos calientes deberían encontrarse en zonas iluminadas (por ejemplo, cerca de farolas) para que la cámara reconociese correctamente (mediante reconocimiento de patrones) el lugar donde se esconden las imágenes.

Otro factor a tener en cuenta es la calidad de las imágenes y el entorno virtual modelado. Para que el ordenador sea capaz de gestionar toda esta información gráfica en tiempo real, ha sido necesario aplicar una simplificación de polígonos en el modelo virtual; en el caso de las imágenes utilizadas, tienen un tamaño de 600x450 píxeles y una resolución de 72 ppp.

Otra dificultad que ha surgido es la de establecer un juego colaborativo, tal y como se propone en el apartado 9.2.2.4. El principal inconveniente reside en disponer de diversos receptores GPS y sensores inerciales, uno por cada participante. Otro factor a tener en cuenta es la necesidad de que los usuarios estén conectados en línea, cuestión que no se ha abordado en esta tesis.

9.3. SISTEMAS DE VERIFICACIÓN

En este ensayo no se ha elaborado ningún sistema de verificación con usuarios de la aplicación propuesta, sin embargo sí se ha elaborado un pequeño formulario de dos preguntas para verificar la relación entre los conceptos planteados en el sistema de navegación (rojo/azul, caliente/frío) de acuerdo a diferentes culturas. Este formulario y los resultados se detallan a continuación:

1. ¿Conoces el juego de caliente-frío?
 - a. Sí.

- b. No.
2. ¿Qué opción tiene más sentido para ti?
- a. Caliente puede relacionarse con el rojo, y frío con el azul.
 - b. Caliente puede relacionarse con el azul, y frío con el rojo.
 - c. Caliente-frío no se puede relacionar con ningún color.
 - d. Caliente-frío se pueden relacionar con otros colores (indica cuales).

Los participantes que contestaron a este pequeño formulario son de las siguientes nacionalidades (sólo 1-3 participantes por nacionalidad): alemana, austriaca, australiana, brasileña, china, estadounidense, española, finlandesa, italiana y singaporense. Con respecto a la primera pregunta, todos contestaron "a", a excepción de un participante de origen chino. A la segunda pregunta todos contestaron "a", pero el participante de origen chino especificó que relacionaría el frío con el azul oscuro. Aunque no todas las culturas entienden los colores de la misma forma (Heller, 2004), parece ser que en cuanto a la relación establecida en *Hot Images* hay unanimidad, por lo que se puede afirmar que el sistema de navegación resulta bastante intuitivo, independientemente de la nacionalidad a la que pertenezca el usuario.

10. LIVE LEGO HOUSE

Live LEGO House (LLH) es un espacio interactivo de realidad aumentada para experimentar algunos aspectos de la convivencia a través del juego. El sistema consiste básicamente en una casa física construida con los bloques de LEGO (LEGO, 2008) que está aumentada con diversos contenidos multimedia generados por ordenador, incluyendo sonidos, vídeos, objetos 3D y pequeñas animaciones. A través de los contenidos de los vídeos y sonidos se introducen algunos elementos multirraciales. Los aspectos relacionados con la convivencia se exploran a partir de una serie de acontecimientos relacionados con la interactividad entre unos muñecos y los elementos de la casa, haciendo que los usuarios tengan que colaborar y ponerse de acuerdo al realizar ciertas acciones. No hay normas explícitas para jugar; los usuarios interactúan libremente con la casa mediante dos muñecos, descubriendo un entorno lleno de sonidos y gráficos que responden a sus acciones. Sin embargo, cada una de las acciones (individuales y en común) condiciona diferentes aspectos del juego, por lo que los usuarios deberán decidir qué acciones realizar para beneficiar a ambos.

Motivación

Según se define en (Wikipedia, 2007a) edutainment es una forma de entretenimiento diseñado para educar al mismo tiempo que divertir. El edutainment generalmente busca instruir o socializar a su audiencia insertando objetivos educativos dentro de alguna forma familiar de entretenimiento, como por ejemplo, programas de televisión,

vídeo juegos, películas, música, páginas Web, etc. Normalmente, esta disciplina busca tutelar en uno o varios temas específicos, o cambiar ciertos comportamientos mediante la enseñanza de actitudes socio-culturales específicas. Se considera un recurso complementario para la educación, cuyo aspecto más destacado reside en aportar un método divertido, lúdico y atractivo para el aprendizaje. Por otro lado, retoma también el potencial que las tecnologías multimedia han aportado en el campo de la educación, debido a que ofrecen una base tecnológica excelente para el desarrollo de habilidades y la transmisión de conocimientos, generando metodologías más eficientes para la educación y el aprendizaje. De igual forma, dentro del área del entretenimiento y juego, estas tecnologías abren nuevas perspectivas que ofrecen grandes posibilidades en esta faceta humana. En *LLH* se indaga en el campo del edutainment mediante la elaboración de un juego de RA con fines educativos, potenciando el juego colaborativo y la interactividad.

Marco temporal

Este ensayo lo realicé durante una estancia de seis meses en el Mixed Reality Lab (MXR, 2006) de la National University of Singapore, laboratorio de prestigio mundial, especializado en la construcción, evaluación y difusión de piezas de realidad mezclada, dirigido por Adrian David Cheok e integrado por una serie de investigadores que provienen principalmente del campo de la Ingeniería Informática, aunque también participan artistas vinculados a los media, matemáticos y físicos.

10.1. PLANTEAMIENTO

La idea de desarrollar un juego basado en la convivencia nació en Singapur, ciudad-estado que se caracteriza por su condición cosmopolita, siendo su población mayoritariamente una mezcla de Chinos, Indios y Malayos, naciones con fuertes arraigos religiosos (entre los que destacan el Budismo, Hinduismo, Islamismo y Catolicismo) y con diferentes lenguas maternas (mandarín, tamil y malayo), con el inglés como nexo común. Aunque existen algunos barrios característicos (como Chinatown), la gente de diferentes culturas conviven día a día (estudian en los mismos centros, trabajan juntos, van de compras a los mismos centros comerciales, comen en los mismos restaurantes, se divierten juntos, etc.), respetando las diferentes costumbres, religiones y tradiciones del resto de culturas. Por ejemplo, cabe resaltar el caso de los *hawker center*, que son típicos establecimientos singapurenses donde se sirve comida a todas horas: en ellos se encuentran las cocinas de todas estas culturas representadas (comida vegetariana china, comida picante india, carnes de cordero para los musulmanes, etc.); un grupo de amigos, por ejemplo, integrado por diferentes culturas, elegirá su menú del estante que desee, pero luego se sentarán todos en la misma mesa para comer; además, también existen platos y cubiertos de diferente color o forma, para distinguir aquellos que nunca van a estar en contacto con la carne de cerdo (para los musulmanes). Otro ejemplo es el respeto por las distintas indumentarias tradicionales: las mujeres indias (aunque no todas) visten con el típico sari mientras que los hombres visten con camisas y pantalones; las musulmanas llevan un pañuelo que cubre su cabello, y pueden llevar o no la típica chillaba, mientras que los hombres también pueden llevarla, aunque suelen ir con pantalones y camisa; la cultura china tiene un modo de vestir más occidental y a las mujeres jóvenes les gusta mucho vestir con faldas y pantalones extra-mini, aunque algunas personas mayores pueden verse con vestidos más tradicionales.

Para una persona occidental puede resultar asombroso ver tanta mezcla cultural (sobre todo en términos religiosos) y lo bien que aparentemente se lleva todo el mundo. Sin embargo, hay que tener en cuenta el riguroso sistema político de Singapur, que en algunos aspectos se asemeja a una dictadura (falta de libertad de expresión en algunos temas; re-elección disimuladamente impuesta del mismo presidente; pena de muerte; etc.), aunque con un aire moderno y comercial que copia mucho de EEUU. Desde la creación del ASEAN (Asociación de Naciones del Sudeste Asiático) en 1967, tanto Indonesia, Malasia como Singapur iniciaron un acelerado avance hacia el

desarrollo económico. A comienzos de los años 90, se caracterizaban por alcanzar sus logros a través de políticas de apertura y libre mercado, guiadas principalmente por el desarrollo tecnológico, siendo los ingresos *per cápita* del país los segundos de Asia, después de Japón. El desarrollo de esta riqueza económica ha sido parejo al régimen liberal-conservador del partido que ha dirigido la política de la ciudad-estado desde su independencia en 1965.

De otro modo, en esa misma época me planteaba la posibilidad de investigar aplicaciones en el área del edutainment. Había realizado ya aplicaciones que se basaban en el juego y en el entretenimiento, pero siempre pensé en explorar el ámbito de la educación, aunque de manera implícita y con el juego como reclamo para captar la atención de los usuarios. Pensé entonces en elaborar un juego interactivo de RA, preferentemente para niños (o con un aspecto visual infantil), en el que se mostrasen diversos aspectos multiculturales característicos de Singapur (a modo de álbum de fotos, pero con vídeos y audio) y de alguna manera se incentivase la buena convivencia dentro de un espacio determinado (una casa).

La elección de los bloques de LEGO fue motivada por la facilidad que este juguete ofrece para modificar de forma sencilla del espacio físico de la casa. Los bloques de LEGO han traspasado culturas y generaciones, ya que son mundialmente conocidos y muchos niños y adultos juegan o han jugado alguna vez con ellos. El hecho de que “resulten familiares” a un gran número de personas de todas las edades, ayuda a que los usuarios sientan una cierta nostalgia o cercanía, involucrándose por ello más en el juego. De hecho, normalmente “el juego gusta antes de jugar”.

A nivel técnico, era el momento idóneo para realizar una aplicación con el software MXRToolKit, ya que fue desarrollado por los investigadores del laboratorio en donde me encontraba.

10.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO

Siguiendo las aportaciones del bloque I, se puede elaborar un marco general de este ensayo:

- Ubicación dentro del continuo de Milgram (apartado 1.1.1):
 - Realidad aumentada.
- Clasificación (apartado 1.2.2):

- Según el entorno físico en el que se desarrolla la aplicación:
 - Cerrada.
- Según la extensión que abarca:
 - Local.
- Según la movilidad de los dispositivos de registro y/o displays:
 - Espacial.
- Según los usuarios que simultáneamente pueden interactuar con el sistema:
 - Colaborativa.
- Según el tipo de colaboración establecida:
 - Presencial.
- Clasificación según Mackay (apartado 1.2.2):
 - Aumentar el objeto físico
- Campo de aplicación (apartado 1.4):
 - Entretenimiento
 - Edutainment
- Sistema de coordenadas terreno/objeto (apartado 2.1.1):
 - Centrado en objetos
- Orientación espacial (apartado 2.1.5):
 - Método indirecto.
- Dispositivos de registro / controladores (capítulo 3):
 - Sensores ópticos: Marcas pasivas.
- Displays (capítulo 4):
 - Displays espaciales: De vídeo.
 - Displays sonoros: Altavoces.
- Software (capítulo 5):
 - Basado en patrones de marcas planas: MXRToolKit.
- Forma de interactuar (apartado 6.1.1):
 - Manipulación de objetos virtuales/reales.
 - Interacción con otros usuarios:
 - Colaboración presencial/directa.

10.1.2. REFERENTES

Los bloques de LEGO son popularmente conocidos, y diversos autores los han utilizado para desarrollar proyectos muy variados. Por ejemplo, en (Rieth, 2005) se presenta un kit de laboratorio hecho a partir de estos bloques, a los que se le han añadido diversos componentes electrónicos, como resistencias, cableado, bombillas, pequeños motores, etc. Otros autores han utilizado los bloques de LEGO en aplicaciones de realidad aumentada, como en (Tavares et al., 2003; Andersen et al., 2004; Barakonyi et al., 2004). Existen otros trabajos de RA donde se explora el espacio de una casa, como en (Dow et al., 2006), introducido en el apartado 6.2.4. Pero uno de los referentes más importantes de realidad mezclada en el campo del edutainment lo podemos encontrar en *KidsRoom* (1997), un proyecto que, aunque inicialmente no fue diseñado con propósitos educativos concretos, según (Roussou, 2004) su combinación de narrativa, juego, y actividades individuales y en grupo, abren un gran abanico de posibilidades pedagógicas. Sus autores (Bobick et al., 1997; Bobick et al., 1999) lo definen como un espacio de juego para niños basado en la percepción, interactividad y narrativa, en el que se utilizan imágenes, música, narración, luces y efectos de sonido para transformar la habitación de un niño en un lugar fantástico, en donde los niños son guiados a través de una historia de aventuras. El sistema se diseñó atendiendo a varios objetivos: mantener la atención de la acción del usuario e interacción en el espacio físico, no en el virtual; permitir que un grupo colaborativo de personas participen en una experiencia interactiva, combinando objetos reales y virtuales; utilizar algoritmos de visión por computador sin la necesidad de que los participantes lleven consigo ningún dispositivo adicional o que se vistan de una forma determinada; la utilización de la narrativa para constreñir el reconocimiento perceptivo permitiendo a los participantes que conduzcan la narrativa; y por último, crear una habitación que constituya un entorno inmersivo e interactivo (Figura 234).



Figura 234. La instalación *KidsRoom* y usuarios interactuando, en (Bobick et al., 1999).

En otro orden de ideas, la pieza de *LLH* recuerda en algunos aspectos al videojuego *The Sims* (Figura 235), y la gran mayoría de las personas que lo conocen los comparan rápidamente, dando a entender que *LLH* es una mera versión del videojuego. Aunque es cierto que *LLH* mantiene ciertas referencias con *The Sims*, como la convivencia en una casa y la interacción con algunos de sus elementos, también existen diferencias sustanciales, que se enumeran a continuación:

1. El objetivo del juego. En *The Sims* el objetivo del juego es organizar el tiempo de los Sims para lograr que alcancen metas personales (Wikipedia, 2008c). En *LLH*, aunque aparentemente no muestre una finalidad determinada, se premia la buena convivencia y tiene un fondo multicultural (a través de los vídeos y audio).
2. Las características del entorno. En *The Sims* el entorno reproducido es puramente virtual. En *LLH* existe un entorno real (físico) y un entorno virtual; a través de la pantalla del ordenador se muestra el entorno real aumentado con los elementos virtuales.
3. Interacción persona-ordenador. En *The Sims* las entradas al ordenador se producen a través del teclado o control de mandos de la videoconsola; en *LLH* la interacción se produce de manera natural a través de los muñecos físicos (interfaz tangible).
4. Interacción persona-persona. *The Sims* es un juego inicialmente creado para un sólo usuario; en su versión online se puede jugar con otros usuarios de forma remota; sin embargo esta versión no ha tenido mucho éxito (Wikipedia, 2008d). *LLH* es un espacio colaborativo para dos usuarios en el que la comunicación entre ellos es presencial, con contacto visual directo.



a

b

Figura 235. Diversas escenas: a) En *The Sims* (2000); b) En *The Sims Online* (2002).

10.2. DESCRIPCIÓN TÉCNICA

10.2.1. COMPONENTES DE LA APLICACIÓN

10.2.1.1. DISPOSITIVOS FÍSICOS

Los componentes de la aplicación de LLH son (Figura 236):

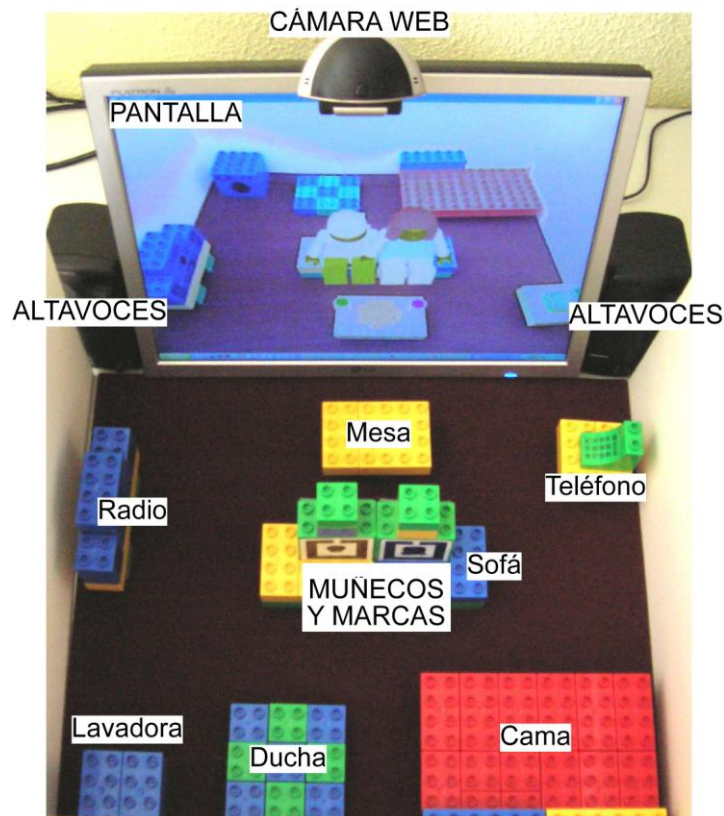


Figura 236. Componentes de la aplicación; Imágenes de la segunda fase del ensayo.

1. Un ordenador de mesa, con las librerías de MXRToolKit.
2. Una cámara Web. En la primera fase del ensayo (apartado 10.3.1) se utilizó la cámara Fire-i de Unibrain con conexión Firewire, resolución de 640x480 píxeles y un campo de visión horizontal 42°. En la segunda fase del ensayo (apartado 10.3.2) se ha utilizado la cámara *Live! Cam Voice* de Creative, con una conexión USB 2.0, una resolución de 640x480 ppp y un campo de visión de 85°.
3. Como display se ha utilizado la pantalla del ordenador, siguiendo la metáfora del espejo mágico.

4. Dos muñecos, LEGOMan y LEGOWoman, contruidos con los bloques de LEGO y que llevan dos marcas respectivamente (una en el dorso y otra en el anverso). Constituyen una interfaz tangible.
5. Bloques de LEGO que forman la casa y los elementos que hay en ella.
6. Un tablero sobre el que se apoyan los bloques.
7. Dos altavoces donde se reproducen diversos sonidos.

10.2.1.2. ELEMENTOS VIRTUALES

LLH es una aplicación de RA multimedia, basada en elementos gráficos, sonoros y tangibles (los muñecos físicos). En este apartado se explicarán los distintos elementos multimedia que son reproducidos por el ordenador durante el desarrollo de la aplicación: sonidos, vídeos y objetos 3D.

Sonidos: Los sonidos utilizados en *LLH* proceden del banco de sonidos del Ministerio de Educación y Ciencia (Ministerio de Educación y Ciencia, 2006). En esta dirección Web hay disponibles diversos elementos multimedia para su libre utilización en ámbitos educacionales. Algunos de los sonidos utilizados son: agua corriente, mujer cantando en la ducha, sintonizar la radio, músicas étnicas, llamada de teléfono, ronquidos, risas, beber, etc.

Vídeos: Se han utilizado un total de seis vídeos que muestran diferentes espacios de Singapur, teniendo cada uno de ellos una duración de entre 1-5 min y una resolución de 640x480 píxeles (condicionada por las limitaciones del sistema).

Objetos 3D: Se han utilizado un total de 15 elementos 3D, que son principalmente los muñecos con distintas posturas o vestimenta (Figura 237) – los objetos correspondientes a la comida y bebida se han introducido en la segunda fase del ensayo (apartado 10.3.2). Estos objetos se han modelado con el software 3D Studio Max, incluyendo algunas animaciones, y posteriormente se han exportado al formato VRML para poder ser ejecutados desde el programa MXRToolKit. Las animaciones son sencillas, y se han utilizado únicamente para la acción de bailar y para dar movimiento a los sueños mientras están durmiendo. Los muñecos se han modelado de acuerdo a los muñecos originales de LEGO. Hay que resaltar que, el único hecho diferenciador de géneros que muestran los muñecos tradicionales de LEGO es en la cabeza: las mujeres llevan pelo largo mientras que los hombres lo llevan corto o llevan una gorra; además, la forma que representa el pelo es desmontable y se pueden intercambiar, dando lugar a ambigüedad de sexos.

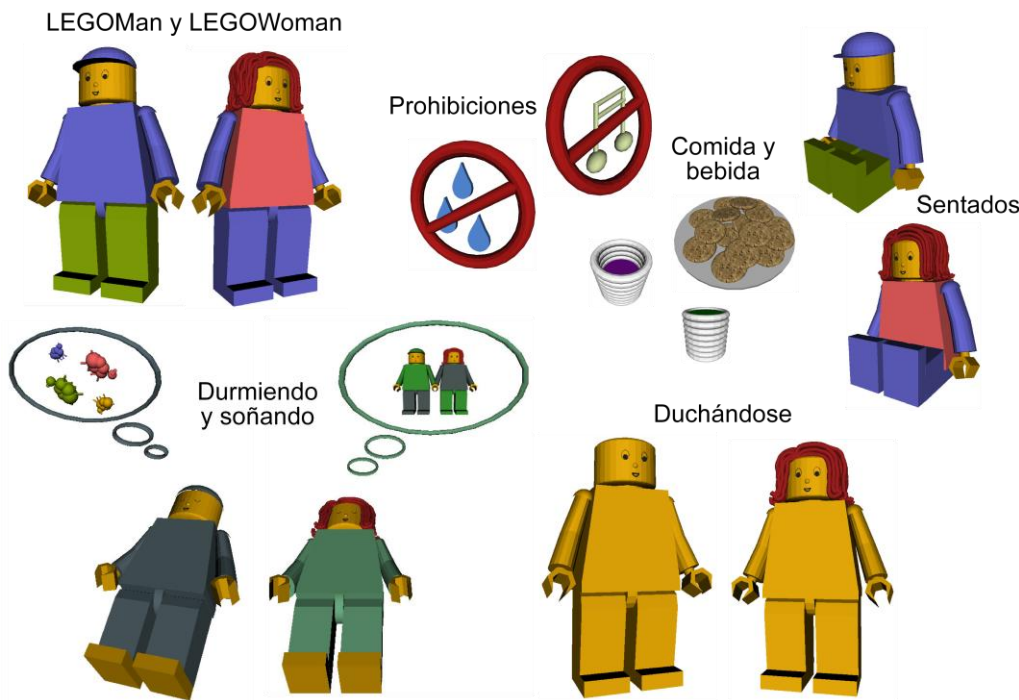


Figura 237. Objetos 3D utilizados en LLH.

10.2.2. DESARROLLO, PROCESO

10.2.2.1. PROCESO DE REDUCCIÓN DE MARCAS

Un aspecto importante de *LLH* es la cantidad de diferentes elementos multimedia que se muestran en el juego comparado con el número de marcas físicas: se utilizan 97 elementos multimedia, de los cuales 45 corresponden a objetos 3D (*wrl*), 46 a sonidos (*wav*) y 6 a vídeos (*avi*), mientras que únicamente hay cuatro marcas físicas, dos por muñeco. Ese hecho es inusual en este tipo de aplicaciones, en las que generalmente se asigna un objeto por marca o grupo de marcas. Esta es sin duda una desventaja de estos sistemas; si incrementásemos el número de marcas en *LLH* de tal forma que cada elemento multimedia tuviese su correspondiente marca, el espacio físico se llenaría de éstas, y no cabría a penas lugar para otros elementos físicos, existiendo además el problema de la oclusión causada por el usuario mientras interactúa con el sistema. Además, el incremento de número de marcas ralentizaría el proceso de registro de las mismas, dando lugar también a que el programa confundiese unas con otras (esto está directamente relacionado con la resolución de la cámara y el tamaño de las marcas). Es por ello que en *LLH* se han reducido el número de marcas considerablemente, para lo cual se han seguido tres pasos:

1. El primer paso consiste en asignar varios elementos multimedia a una misma marca. Por ejemplo, la marca que correspondería a la radio se le han asignado un total de 30 archivos de audio. Esto no significa que se vayan a ejecutar todos a la vez, sino que se ejecutará uno u otro dependiendo de las interacciones que se produzcan en el juego.
2. El segundo paso consiste en eliminar todas aquellas marcas físicas que permanecen fijas con respecto a la cámara durante el desarrollo del juego, aunque manteniendo su definición dentro del juego. Ya que estas marcas no están físicamente, las denominaremos a partir de ahora marcas virtuales. Este paso es posible porque la cámara permanece estática, por lo que algunas marcas, las relativas a los elementos de la casa (cama, ducha, etc.), tendrán una matriz de transformación constante durante todo el juego. Así pues, únicamente será necesario conocer dicha matriz a principio del juego, y para lograrlo se ha ideado un sistema de “calibración de las marcas” mediante teclado, que se explicará a continuación.
3. El tercer y último paso para la reducción del número de marcas consiste en asignar a todos los *frames* definidos, una única marca (*marker*) virtual. Esto se puede hacer porque la cámara nunca va a encontrar en el entorno dicha marca. En la Figura 238 se muestra la programación de asignación de marcas a los *frames*.

```
// asignamos al frame 0 las marcas hombre1.jpg y hombre2.jpg:
mrxFrameReadDoll(&frame[0], 20, 72, "hombre1.jpg", "hombre2.jpg");
// asignamos al frame 0 las marcas mujer1.jpg y mujer2.jpg:
mrxFrameReadDoll(&frame[1], 20, 72, "mujer1.jpg", "mujer2.jpg");

// asignamos al resto de frames la misma marca marker12.frame:
mrxFrameRead(&frame[2], "marker12.frame");
mrxFrameRead(&frame[3], "marker12.frame");
mrxFrameRead(&frame[4], "marker12.frame");
mrxFrameRead(&frame[5], "marker12.frame");
...
mrxFrameRead(&frame[15], "marker12.frame");
```

Figura 238. Asignación de marcas a los frames.

Con estas consideraciones, el número total de marcas y *frames* queda como sigue:

- Cantidad de *frames* definidos en el programa: 16
 - *Frames* en los muñecos: 2 (uno por cada muñeco)

- *Frames* en los elementos de la casa: 14 (hay 7 elementos con 2 marcos cada uno, puesto que hay 2 muñecos)
- Cantidad de marcas físicas: 4 (2 por cada muñeco)
- Cantidad e marcas virtuales: 1 (que engloba los 14 elementos de la casa)

10.2.2.2. INTERFAZ, INTERACCIÓN Y MÁQUINA DE ESTADOS FINITOS

La cámara Web apunta hacia la casa registrando las acciones de los muñecos. La casa es como un elemento vivo ya que responde dependiendo de dichas acciones, mostrando (a través de la pantalla o altavoces) diferentes elementos dependiendo del comportamiento de ambos muñecos. Por ejemplo, la música que un muñeco escucha cuando se acerca a la radio depende de sus acciones anteriores (si se ha duchado, ha visto la tele, etc.), pero también depende de si el otro muñeco también se acerca o no a la radio y de las acciones anteriores de este segundo muñeco. De este modo, da la impresión de que la casa está viva, y de que guía a los muñecos a través del juego. También da cierta sensación de aleatoriedad, evitando el conocimiento en todo momento de qué es lo que va a pasar.

Técnicamente, este hecho se consigue mediante el establecimiento de una máquina de estados finitos sencilla que, en términos computacionales, se asemeja a la inteligencia artificial ya que produce la sensación de que el ordenador piensa por sí mismo. La arquitectura de estas máquinas se divide en: lógica combinacional de entrada; bloque secuencial de registros de estados; lógica combinacional de salida. En *LLH* se han programado una serie de condicionales y variables para crear este efecto: cada vez que el muñeco se acerque a un objeto de la casa, dependiendo del estado actual se producirá un tipo de interacción distinto, y se asignarán estados de salida que condicionarán la siguiente interacción.

La interfaz usuario-ordenador

El desarrollo de interfaces amigables es la clave para que los usuarios, incluso aquellos no familiarizados con los ordenadores, puedan beneficiarse de las posibilidades que éstos ofrecen, implicando la eliminación del denominado "digital gap". En *LLH* la interfaz la constituyen los muñecos de LEGO, por lo que la comunicación con el ordenador se convierte en un proceso muy sencillo e intuitivo, simplemente con las inclinaciones y desplazamientos del muñeco por el entorno creado, es decir, mediante

las acciones propias de un juego infantil, sin necesidad de advertir que existe un ordenador (siguiendo la metáfora del ordenador invisible).

Interacción

En las siguientes líneas se explica la manera en que se produce la interacción:

- La interacción se activa de dos maneras diferentes:
 - Proximidad: Se han establecido diferentes radios alrededor de los elementos de la casa, de tal forma que cuando uno de los muñecos entra en la zona afectada por este radio se activa algún proceso.
 - Rotaciones: Cuando un muñeco esté dentro de una zona de interacción, la orientación del muñeco también activa ciertos procesos.
- La secuencia de las acciones de los muñecos condiciona los elementos multimedia que se van a mostrar. Técnicamente esto se ha logrado mediante la definición de una serie de variables (aux, aux2, aux3, ..., aux15) que van cambiando de valor a medida que los muñecos interactúan con los elementos de la casa, y que, según su valor actual, condicionan los elementos multimedia que se van a mostrar. En la Figura 239 se muestra un ejemplo de programación comentado para el momento en que se produce la interacción entre LEGOMan y la tele.
- Las acciones de un muñeco condicionan al otro. Por ejemplo:
 - Si ambos muñecos están viendo la tele, se muestra un programa específico (el que les gusta a los dos).
 - Si uno de los muñecos está durmiendo, el otro no puede encender la radio ni llamar por teléfono, ya que no se permite hacer ruido.
 - Para que los muñecos puedan comer y beber, deberán estar sentados los dos en el sofá; si sólo se sienta uno de ellos podrá realizar sólo una de las acciones (comer o beber).

```

check8 = check_dist(&dollInfoh, &fInfo08); // distancia LEGOMan-sofá
if (check8 <= distmin){ // si LEGOMan está sentado en su parte del sofá
  aux10 = 2; // aux10, que controla la radio, se le asigna un valor igual a 2
  aux11 = 1; // aux11, que controla el sonido de comer/beber de LEGOMan, se le asigna
un valor igual a 1
  if (check9 > distmin){ // si LEGOWoman no está sentada en su parte del sofá
    if (aux2 == 0){ // si la variable aux2 es igual a 0
      mxrMediaRender(&mediaObj2, &frame[8].T); // la tele muestra el vídeo del
mediaObj2
    }
  }
}

```

```

    }
    if (aux2 == 1){ // si la variable aux2 es igual a 1
        mxrMediaRender(&mediaObj3,&frame[8].T); // la tele muestra el vídeo del mediaObj3
    }
    if (aux2 == 2){ // si la variable aux2 es igual a 2
        mxrMediaRender(&mediaObj26,&frame[8].T); // la tele muestra el vídeo del
mediaObj26
    }
    if (aux2 == 3){ // si la variable aux2 es igual a 3
        mxrMediaRender(&mediaObj27,&frame[8].T); // la tele muestra el vídeo del
mediaObj27
    }
}
if (check9 <= distmin){ // si LEGOWoman está sentada en su parte del sofá
    mxrMediaRender(&mediaObj2,&frame[8].T); // la tele muestra el vídeo del mediaObj2
}
}
}

```

Figura 239. Código de programación que describe la interacción entre LEGOMan y la tele.

La televisión es uno de los elementos centrales del juego, ya que muestra distintos vídeos de la cultura singaporense. Metafóricamente, la pantalla del ordenador es la tele que, cuando está apagada, funciona a modo de espejo reflejando el espacio físico (aumentado), y cuando está encendida deja de verse dicho espacio para ver los vídeos. La televisión se enciende cuando los muñecos se sientan en su lado respectivo del sofá (el amarillo para el muñeco amarillo y el azul para el muñeco azul), siendo necesario que se siente al menos uno de ellos. Si los muñecos se sientan en el lado opuesto, la tele permanece apagada, y es cuando se activan las acciones de comer y beber. Es por ello que el sofá tiene doble interacción: ver la tele y comer/beber.

Cuando varias personas ven la tele, deben llegar a un acuerdo de qué quieren ver, generalmente, aquel programa, documental o película que sea del agrado de todos. Se han asignado diferentes vídeos a cada uno de los muñecos, teniendo sólo un vídeo en común (que les gusta a los dos). De este modo, cuando los dos muñecos estén sentados en el sofá, únicamente se verá dicho vídeo. Cuando sólo un muñeco esté sentado, se verá alguno de los vídeos que tiene asignado, dependiendo siempre de las acciones anteriores. En la Figura 240 se muestra un ejemplo.

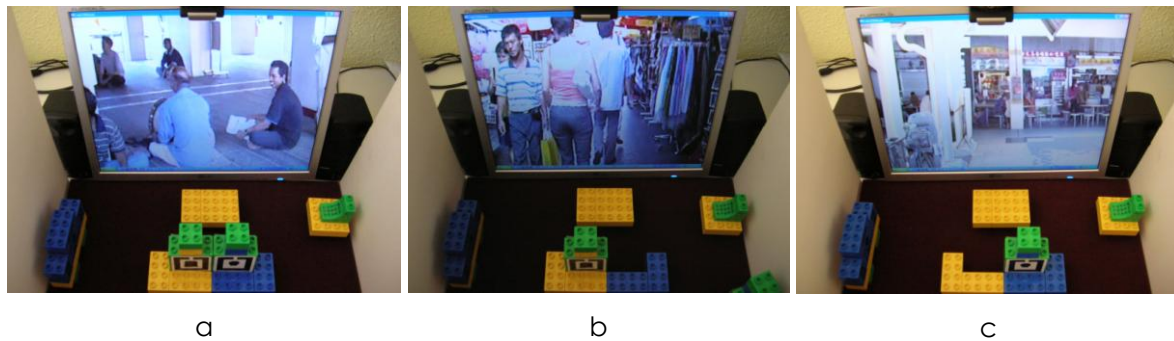


Figura 240. Distintas escenas en la tele: a) Los dos muñecos están sentados en sus respectivos asientos y aparece un vídeo que muestra a varias personas en una ceremonia dentro de un templo Hindú; b) Sólo está sentada LEGOWoman, y aparece un vídeo que muestra la zona comercial de Chinatown; c) Sólo LEGOMan está sentado y se ve un vídeo que muestra un *hawker centre*.

También existen una serie de restricciones respecto al agua y al sonido. Así pues, la cantidad diaria de agua que se puede utilizar está delimitada a una ducha por muñeco y a una puesta de lavadora. En cuanto al sonido, como se ha mencionado antes, ningún muñeco podrá realizar alguna acción que active elementos sonoros si el otro muñeco está durmiendo.

Elemento	Interacción	Acciones	Multimedia
Lavadora	Proximidad	Poner la lavadora / Prohibición	Sonido: Lavadora en funcionamiento 3D: Señal de prohibición de agua
Radio	Proximidad / Roll	Encender la radio / Cambiar de sintonía / Bailar / Prohibición	Sonido: Canciones y programas de radio 3D: Muñecos bailando / Señal de prohibición de sonido
Sofá (mesa)	Proximidad / Pitch	Sentarse / Comer / Beber / Degustar / Bostezar	3D: Muñecos sentados / Vasos con zumo / Un plato con galletas Sonido: Degustar / Beber / Bostezar
Sofá (TV)	Proximidad	Encender la TV / Cambiar el canal de la TV	Videos: 6 vídeos mostrando escenas de Singapur
Cama	Proximidad / Pitch	Ponerse el pijama / Cerrar los ojos / Roncar / Soñar	3D: Muñecos con pijamas / Sueños animados Sonido: Roncar
Ducha	Proximidad	Desvestirse / Darse una ducha / Cantar	3D: Muñecos desnudos / Señal de prohibición de agua Sonido: Cantar / Agua corriente

Teléfono	Proximidad / Roll / Pitch	Llamar por teléfono / Recibir una llamada / Contestar al teléfono / Prohibición	Sonido: Marcar / Ring / Descolgar el teléfono / Decir hola 3D: Señal de prohibición de sonido
----------	---------------------------------	---	---

Tabla 6. Elementos con interacción, acciones derivadas y elementos multimedia que se muestran.

10.2.2.3. DESARROLLO EN EL SOFTWARE UTILIZADO

Para desarrollar la aplicación de *LLH* se ha utilizado el software *MXRToolkit* (ver apartado 5.1.2). Como se ha visto, la filosofía de este programa es proporcionar una interfaz de programación lo suficientemente sencilla para que pueda ser utilizada por usuarios no expertos en programación. Es por ello que la integración de distintos elementos multimedia (sonido, vídeo, texto, 3D, etc.) resulta mucho más sencillo que en otros programas similares, tales como *ARToolKit*. La utilización de *MXRToolkit* ha facilitado la realización de una aplicación colaborativa con interfaces tangibles, para la que además de usar objetos característicos de este software (como *mxrMediaRead*, *mxrMediaTranslate*, *mxrGLDrawVideo*, etc.), también se han definido algunos nuevos objetos en C++ para su utilización en la aplicación de *LLH*. A continuación se muestran los códigos comentados de estos nuevos objetos:

1. Distancia espacial entre un muñeco y una marca virtual: **check_dist**.

```
double check_dist(DollInfo *dInfo, FurnitureInfo *fInfo){

    double    dist2;
    double    x, y, z;

    mxrTransform TFurnitureInv,T;
    mxrTransformInvert(&TFurnitureInv, fInfo->T);
    mxrTransformMult(&T, &TFurnitureInv, dInfo->T);

    x = T.tx; // distancia entre las 2 marcas en el eje x
    y = T.ty; // distancia entre las 2 marcas en el eje y
    z = T.tz; // distancia entre las 2 marcas en el eje z
    dist2= sqrt(x*x+y*y+z*z); // distancia espacial entre las 2 marcas

    return dist2;
}
```

2. Distancia espacial entre los dos muñecos: **check_proxi**.

```
double check_proxi(DollInfo *d1Info, DollInfo *d2Info){

    double    proxi;
    double    x, y, z;

    mxrTransform TDollInv,T;
    mxrTransformInvert(&TDollInv, d2Info->T);
    mxrTransformMult(&T, &TDollInv, d1Info->T);

    x = T.tx; // distancia entre los 2 muñecos en el eje x
    y = T.ty; // distancia entre los 2 muñecos en el eje y
    z = T.tz; // distancia entre los 2 muñecos en el eje z
    proxi= sqrt(x*x+y*y+z*z); // distancia espacial entre los 2 muñecos

    return proxi;
}
```

3. Rotaciones entre dos marcas, alrededor de los tres ejes coordenados: **check_roll**, **check_pitch**, **check_yaw**.

```
double check_roll(DollInfo *dInfo, FurnitureInfo *fInfo){

    double    roll;

    mxrTransform TFurnitureInv,T;
    mxrTransformInvert(&TFurnitureInv, fInfo->T);
    mxrTransformMult(&T, &TFurnitureInv, dInfo->T);

    roll = atan (T.r23/T.r33)*180/3.1416; // rotación a través del eje x
    return roll;
}

double check_pitch(DollInfo *dInfo, FurnitureInfo *fInfo){

    double    pitch;

    mxrTransform TFurnitureInv,T;
    mxrTransformInvert(&TFurnitureInv, fInfo->T);
    mxrTransformMult(&T, &TFurnitureInv, dInfo->T);

    pitch = asin (-T.r13)*180/3.1416; // rotación a través del eje y
    return pitch;
}
```

```
double check_yaw(DollInfo *dInfo, FurnitureInfo *fInfo){

    double    yaw;

    mxrTransform TFurnitureInv,T;
    mxrTransformInvert(&TFurnitureInv, fInfo->T);
    mxrTransformMult(&T, &TFurnitureInv, dInfo->T);

    yaw = atan (T.r12/T.r11)*180/3.1416; // rotación a través del eje z
    return yaw;
}
```

4. Dos marcas opuestas (las de los muñecos) funcionando como una única marca (esta función es una modificación de la función *mxrFrameReadCube*): **mxrFrameReadDoll**.

```
bool mxrFrameReadDoll(mxrFrame *frame, float dollWidth, float dpi, char *front,
char *back){

    mxrFrameCreate(frame);
    mxrTransform T;

    // Añadir una marca al frente
    mxrTransformLoadIdentity(&T);
    mxrTransformCWRotateZ(&T,3.1413);
    mxrFrameAddOneMarkerFromImage (frame,front,dpi,&T);

    // Añadir una marca al dorso
    mxrTransformLoadIdentity(&T);
    mxrTransformCWRotateZ(&T,3.1413);
    mxrTransformCWRotateY(&T,3.1413);
    mxrTransformTranslate(&T,0,0,-dollWidth);
    mxrFrameAddOneMarkerFromImage (frame,back,dpi,&T);

    return(1);
}
```

5. Conjunto de funciones necesarias para realizar el efecto de espejo (voltear la imagen horizontalmente): **getRed, getGreen, getBlue, setColor, mxrImageFlipX**.

```
char getRed(mxrImage *img, int x, int y) {
    return img->ubData[y*img->x*3+x*3];
}

char getGreen(mxrImage *img, int x, int y) {
    return img->ubData[y*img->x*3+x*3+1];
}
```

```

}

char getBlue(mxrImage *img, int x, int y) {
    return img->ubData[y*img->x*3+x*3+2];
}

void setColor(mxrImage *img, int x, int y, char r, char g, char b) {
    img->ubData[y*img->x*3+x*3] = b;
    img->ubData[y*img->x*3+x*3+1] = g;
    img->ubData[y*img->x*3+x*3+2] = r;
}

void mxrImageFlipX(mxrImage *outImg, mxrImage *inImg){

    int xSize, ySize;
    xSize = inImg->x;
    ySize = inImg->y;

    for (int y = 0; y<ySize; y++)
        for (int x = 0; x<xSize; x++)
            setColor(outImg, x, y, getRed(inImg,xSize-x-1, y),
                    getGreen(inImg,xSize-x-1, y),
                    getBlue(inImg,xSize-x-1, y));

    return;
}

```

6. Conjunto de funciones necesarias para realizar un croma (solamente se utilizó en la primera versión de LLH): **findColor**, **setColor**, **extractBackground**, **changeBackground**, **getColor**.

```

int x = 0;
int width;
int y = 0;
int height;

bool findColor(int x, int y, mxrImage * im, int threshold, int lThreshold, char
color) { //encontrar un color a partir de ciertos umbrales establecidos

    px Px;
    getColor(x,y,im,&Px);
    bool isColor = false;
    switch(color) { //dependiendo de la tecla presionada se elige un color
        case 'B':
            isColor = (Px.B>threshold && Px.R<lThreshold && Px.G<lThreshold); //Blue
            break;

```

```

    case 'R':
        isColor = (Px.R>threshold && Px.G<lThreshold && Px.B<lThreshold); //Red
        break;
    case 'G':
        isColor = (Px.G>threshold && Px.R<lThreshold && Px.B<lThreshold); //Green
        break;
    case 'D':
        isColor = (Px.R==0 && Px.G == 0 && Px.B==0); // Black
        break;
    case 'W':
        isColor = (Px.R==255 && Px.G == 255 && Px.B==255); // White
        break;
    case 'O':
        isColor = (Px.R>200 && Px.G >150 && Px.B < 150); // Orange
        break;

}
return isColor;
}

void setColor(int x, int y, mxrImage *im, int R, int G, int B, int A){ // función
sobrecargada; en este caso, se aplica para imágenes con valor alpha: RGBA

    im->ubData[y*4*im->x + x*4]=R;
    im->ubData[y*4*im->x + x*4+1]=G;
    im->ubData[y*4*im->x + x*4+2]=B;
    im->ubData[y*4*im->x + x*4+3]=A;
}

void setColor(int x, int y, mxrImage *im, int R, int G, int B){ // función
sobrecargada; en este caso, se aplica para imágenes sin valor alpha: RGB

    im->ubData[y*3*im->x + x*3]=R;
    im->ubData[y*3*im->x + x*3+1]=G;
    im->ubData[y*3*im->x + x*3+2]=B;
}

void extractBackground(mxrImage *im, int threshold, int lThreshold, char color,
mxrImage * outIm) {

    bool isColor;
    for (int i=0; i<im->x; i++) {
        for (int j=0; j<im->y; j++){
            isColor = (findColor(i,j,im,threshold,lThreshold,color));
            if(isColor) setColor(i,j,outIm, 0, 0, 0, 100); // set Background transparent
        }
    }
}

```



```

}

void changeBackground(mxrImage *im, mxrImage *imBack, int threshold, int
lThreshold, char color,mxrImage * outIm) {

    px Px;
    bool isColor;
    for (int i=0; i<im->x; i++) {
        for (int j=0; j<im->y; j++){
            isColor = (findColor(i,j,im,threshold,lThreshold,color));
            if(isColor) {
                getColor(i,j,imBack,&Px);
                setColor(i,j,outIm, Px.R,Px.G,Px.B);// set Background transparent
            }
        }
    }
}

void getColor(int x, int y, mxrImage *im, px *Pixel){
    Pixel->B = im->ubData[y*3*im->x + x*3+2];
    Pixel->G = im->ubData[y*3*im->x + x*3+1];
    Pixel->R = im->ubData[y*3*im->x + x*3];
}
}

```

10.2.3. FASES DEL JUEGO

El juego de *LLH* consta de tres fases:

- Construir la casa con los bloques de LEGO.
- Proceso de calibración de las marcas.
- Explorar los aspectos de convivencia y multiculturalismo.

Para niños pequeños el proceso de calibración resulta complejo, por lo que debe ser ayudado por un adulto. En los siguientes subapartados se explican con más detalle.

10.2.3.1. CONSTRUIR LA CASA CON LOS BLOQUES DE LEGO

El primer paso consiste en construir una casa con los bloques de LEGO, dando lugar a que los usuarios se valgan de su imaginación para construir el entorno físico de la aplicación. En la siguiente figura se muestran algunos ejemplos:

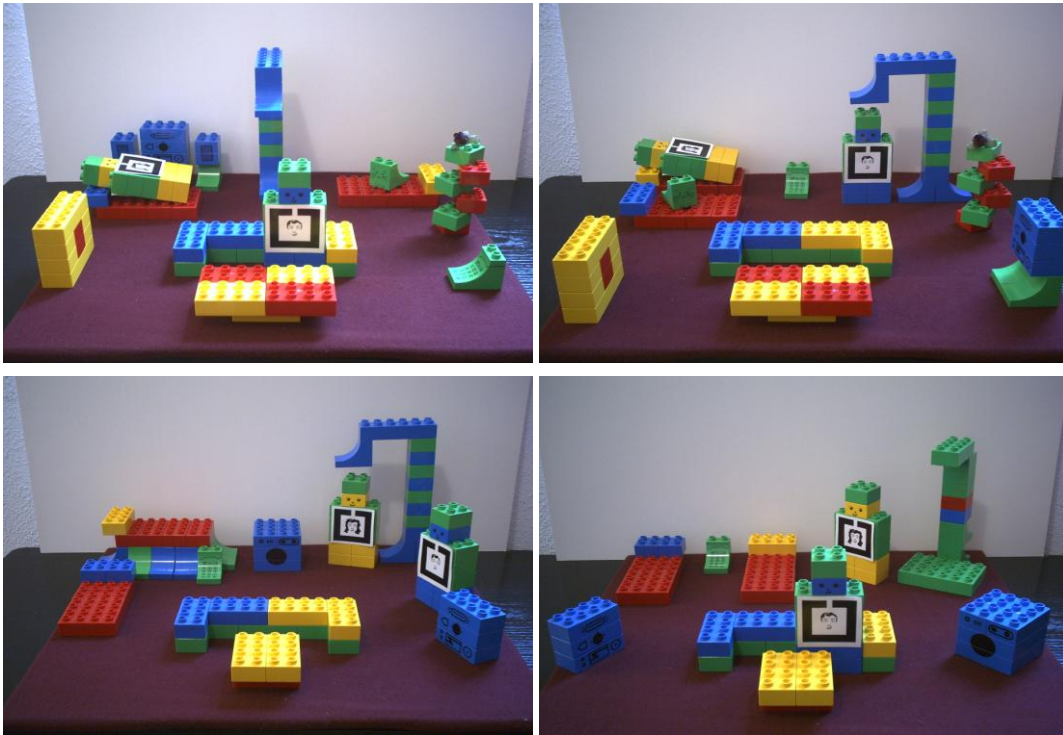


Figura 241. Ejemplos de distintos escenarios físicos.

Existen sin embargo una serie de pautas o restricciones que han de tenerse en cuenta. En primer lugar, la casa se construye sobre un tablero de dimensiones determinadas, que está situado de una manera específica en relación a la pantalla del ordenador y a la cámara Web (ver apartado 10.2.1.1). En segundo lugar, los elementos de la casa que tienen asignados algún tipo de interacción son determinados (lavadora, sofá, mesa, cama, ducha, teléfono, radio), por lo que preferiblemente deberán estar presentes, y su orientación debe ser tal que los muñecos no se queden nunca de perfil a la cámara Web (ya que no se verían las marcas). Por último, la ubicación del sofá debe ser preferentemente en primera fila, delante de la pantalla, para que las dimensiones de los vídeos mostrados sean adecuadas.

10.2.3.2. PROCESO DE CALIBRACIÓN DE LAS MARCAS

Aunque es cierto que los elementos de la casa permanecen fijos con respecto a la cámara a lo largo del juego, también hay que tener en cuenta que cada usuario puede construir la casa como quiera, por lo que las matrices de transformación permanecerán constantes durante un mismo juego pero variarán en juegos distintos. De este hecho se deriva la necesidad de un proceso de calibración semi-automático que guarde dichas matrices al inicio de cada juego. Este proceso, sin embargo, está programado de manera que se puede rehacer en cualquier momento, permitiendo al

usuario que pueda mover los elementos de la casa durante el juego, siempre que repita el proceso de calibración (sólo para los elementos que se muevan).

El proceso de calibración se hace de una forma sencilla mediante los muñecos y a través del teclado. Cada elemento de la casa tiene asignada una letra del teclado, que es mayúscula para LEGOWoman y minúscula para LEGOMan; estas correspondencias se muestran en la Tabla 7. Los pasos a seguir son:

1. Activar el modo Mayúsculas en el teclado.
2. Coger el muñeco LEGOWoman y situarlo al lado de uno de los elementos de la casa; a continuación presionar la letra del teclado que corresponde. Repetir para cada elemento de la casa. De este modo la marca virtual correspondiente a cada uno de los elementos pasa tener la matriz de transformación que LEGOWoman tiene en el instante que se presiona la tecla.
3. Desactivar el modo Mayúsculas en el teclado.
4. Repetir el paso 2 con el muñeco LEGOMan.

Letra	Elementos
A/a	Lavadora
B/b	Radio
C/c	Sofá (TV)
D/d	Sofá (mesa)
E/e	Cama
F/f	Ducha
G/g	Teléfono

Tabla 7. Letras correspondientes a los elementos de la casa en el proceso de calibración.

10.2.3.3. EXPLORAR LOS ASPECTOS DE CONVIVENCIA Y MULTICULTURALISMO

En este apartado se muestra a modo de *storyboard* un ejemplo de ejecución del juego, mediante una serie de imágenes comentadas que se muestran a continuación.

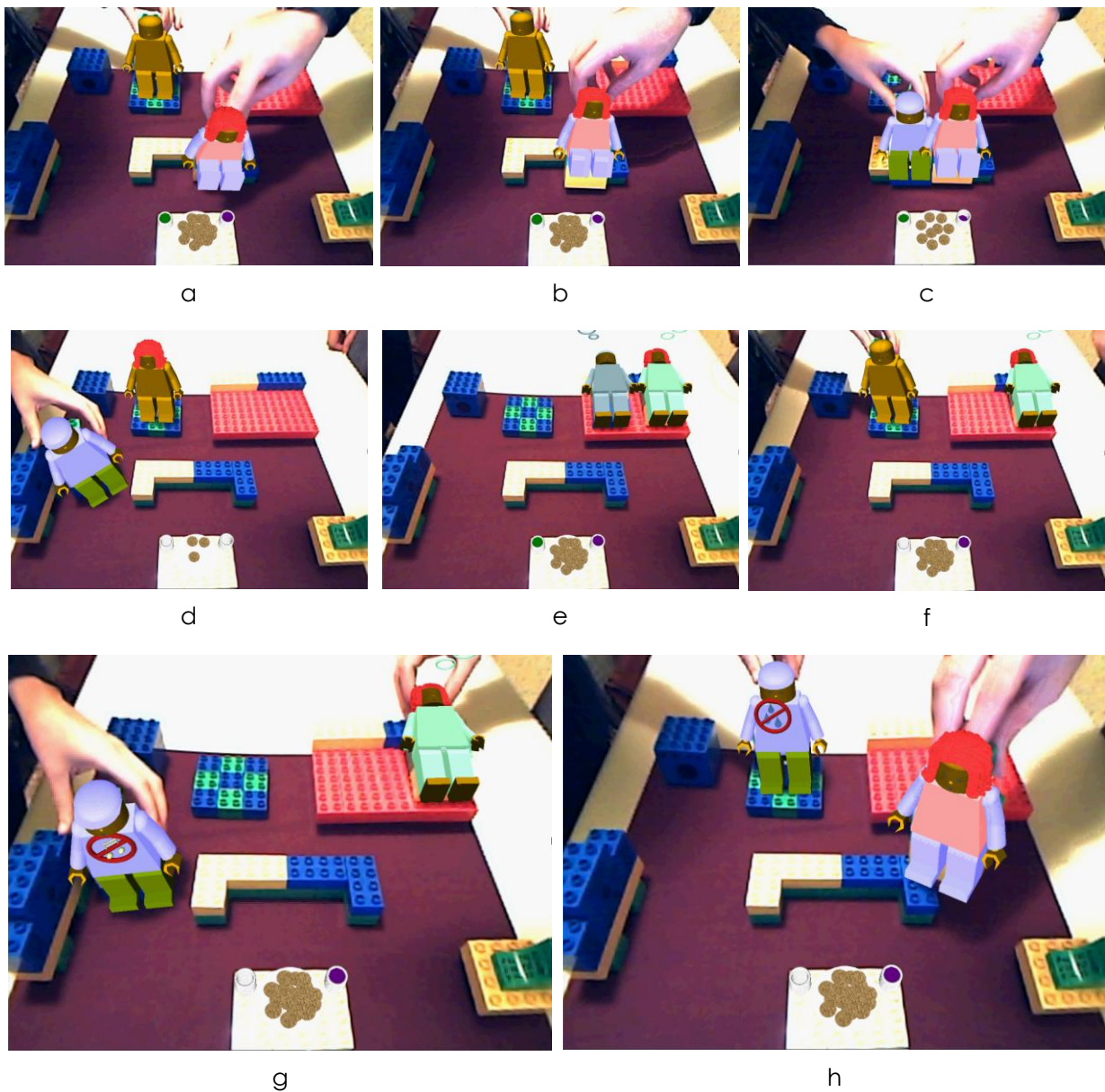


Figura 242. Distintas escenas de LLH a modo de storyboard.

- a) Empieza un nuevo día. LEGOMan toma su ducha diaria mientras LEGOWoman come (movimiento hacia delante y hacia atrás).
- b) Ídem.
- c) Los dos muñecos comen y beben. Disminuye la cantidad de galletas en el plato y el líquido del vaso.
- d) LEGOMan escucha la radio y baila mientras LEGOWoman toma su ducha diaria. En la mesa casi no queda comida ni bebida.
- e) Los dos se acuestan, con lo que a partir de aquí empieza un nuevo día. El plato y los vasos vuelven a estar llenos y no hay restricciones de "agua".

- f) LEGOMan se levanta y toma su ducha diaria mientras LEGOWoman sigue durmiendo. Se observa como anteriormente ha desayunado, puesto que no le queda líquido en el vaso. Sin embargo no ha podido comer porque LEGOWoman no se ha sentado con él a desayunar, por lo que el plato de galletas sigue intacto.
- g) LEGOMan intenta poner música, pero prevalece la restricción de “sonido” porque LEGOWoman sigue durmiendo.
- h) LEGOWoman se levanta. LEGOMan intenta ducharse, pero prevalece la restricción de “agua” porque ya se ha duchado una vez en el mismo día.

10.2.4. CONSIDERACIONES / RESTRICCIONES TÉCNICAS

Para construir el escenario de *LLH* se han tomado en cuenta diversos aspectos condicionantes:

1. **Superficie satinada de los bloques de LEGO:** La luz reflejada es un problema en este tipo de aplicaciones, en las que el registro se basa en el reconocimiento de marcas a partir de técnicas de visión por computador. Los bloques de LEGO son bastante reflectantes, especialmente aquellos con colores claros o blancos. En la primera versión de *LLH*, se utilizó una base para la casa de color gris claro, y los muñecos estaban hechos con bloques blancos. Estas superficies producían bastantes reflejos, dando lugar a que el programa se colapsara bastante a menudo. Para mitigar este problema, en la segunda versión de *LLH* se ha utilizado una superficie recubierta de una tela aterciopelada de color oscuro; además, los muñecos ya no son blancos.
2. **Campo de visión de la cámara:** Debido a que las marcas de los muñecos son pequeñas (3.5x3.5 cm), la cámara debe estar lo bastante cerca para detectarlas y reconocerlas adecuadamente. Sin embargo, si la cámara se acerca demasiado, puede que su campo de visión no abarque toda la casa, y por lo tanto se queden partes sin visualizar. En la primera versión de *LLH* se utilizó una cámara con un FOV de 50°, teniendo que situarse ésta a una altura aproximada de 45 cm sobre la casa. En la segunda versión de *LLH*, se utilizó una cámara con un FOV de 85°, permitiendo situarla más cerca de la casa a pesar de que sus dimensiones son un poco mayores que en el caso anterior. La cámara empleada tiene un sistema automático de corrección de distorsión de la lente debido a la abertura angular, por lo que las distorsiones en la imagen son mínimas y no afectan al registro de las marcas.

3. **Oclusión:** *LLH* es una aplicación de mesa en la que existen elementos reales sobre ella. Los elementos físicos pueden ocasionar problemas de oclusión, haciendo que los elementos virtuales aparezcan visualmente desplazados. Para evitar el problema de la oclusión en *LLH* se han tomado en consideración dos aspectos. En primer lugar, la posición de la cámara debe ser un poco elevada con respecto a la casa (justo encima de la pantalla); en segundo lugar los elementos de la casa no deben ser muy altos (max. 2-3 bloques). Este problema se aprecia en la Figura 243, donde el muñeco está situado físicamente entre una mesa (de color rojo, detrás del muñeco) y una silla (de color azul, delante del muñeco). La imagen en la figura 10.10-a está vista con una cámara a poca altura, por lo que el muñeco virtual aparece en primera línea, delante de la silla; la imagen correspondiente a la figura 10.10-b está tomada con una cámara a mayor altura, apareciendo el muñeco virtual en su posición correcta. Así pues, se puede decir que el problema de la oclusión restringe la ubicación de la cámara.

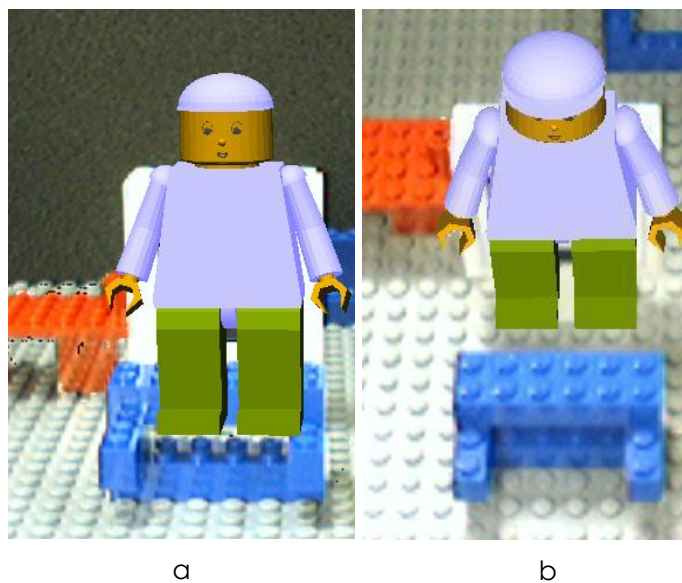


Figura 243. Imagen mostrando el problema de la oclusión: a) Con una cámara en posición baja; b) Con la cámara en una posición elevada.

10.3. MEMORIA DE LAS DISTINTAS FASES

10.3.1. PRIMERA FASE: LLH V.1

En una primera fase de laboratorio se elaboraron diversos prototipos de casas y diversos elementos gráficos. Se añadieron gradualmente al MXRToolKit los elementos generados por ordenador descritos en el apartado 10.2.1.2, se añadieron de forma paralela al programa otros elementos gráficos (ver apartado 10.3.1.2) y se programaron nuevas rutinas, todo ello con la finalidad de encontrar los límites de la tecnología y del equipo utilizado. A partir de su análisis se aceptaron unos y se rechazaron otros. Finalmente se elaboró una prueba de usuarios con la finalidad de que evaluaran el sistema (ver apartado 10.4).

10.3.1.1. DESCRIPCIÓN TÉCNICA

En las Figura 244 y Figura 245 podemos ver los primeros prototipos de la casa, los muñecos y la posición de la cámara en el entorno físico. Como se puede ver, en un primer momento la cámara estaba situada más cerca de la casa; la escena de captura no abarcaba todo el espacio físico de ésta debido a su posición y campo de visión. En pruebas posteriores (Figura 245), la cámara se elevó unos centímetros, para poder recoger toda la escena. También se modificaron los gráficos mostrados en el display, al introducir el dibujo de una pared y un televisor como fondo de la ventana de realidad aumentada.



Figura 244. Primera versión de la casa y muñecos.



Figura 245. Algunas mejoras añadidas: TV sobre fondo de pared; posición de la cámara más elevada para aumentar la zona visualizada y evitar oclusiones.

En esta fase de laboratorio, también se probaron distintas posibilidades que finalmente se desecharon por diversos motivos. Estas "pruebas fallidas" son:

1. Se pensó en la utilización de un elemento plano de forma rectangular para colgarlo en las paredes de la casa, el cual se pudiera quitar, poner, o cambiar de lugar a elección de los usuarios. Dependiendo de en qué lugar de la pared estuviese, su representación en la imagen virtual pudiese ser una ventana (abierta o cerrada), un cuadro, un cartel, etc. Para evitar el problema de la oclusión, se pensó en programar un croma key, de tal forma que el objeto rectangular debería ser de un color determinado. Este croma se programó (ver apartado 10.2.2.3), pero finalmente se desechó su uso por consumir demasiados recursos de la CPU, ralentizando la aplicación.
2. También se realizó un vídeo de gente pasando por la calle desde el patio de una casa; este vídeo se fusionó con las imágenes de las ventanas situadas a los lados del televisor mediante un croma (con el programa Premiere), de tal forma que a través de dichas ventanas se viesan las personas transcurriendo de un lugar a otro (Figura 246). Finalmente, este vídeo se puso como protector de pantalla, para que permaneciese de fondo durante la aplicación de LLH. Al igual que en el caso anterior, esta prueba se desechó por consumir demasiados recursos de la CPU.

3. La ventana de realidad aumentada en esta primera fase corresponde al centro del televisor dibujado (figura 10.12), con unas dimensiones aproximadas de 15x11 cm. Esta ventana resultaba demasiado pequeña, ya que restaba importancia a la parte gráfica de la aplicación, por lo que finalmente se decidió que las dimensiones del televisor fueran coincidentes con las del display empleado (pantalla del PC).

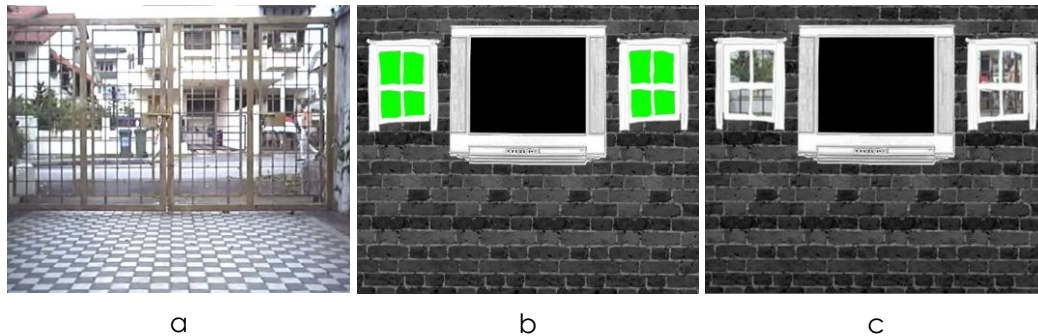


Figura 246. Salvapantallas: a) Vídeo de la calle desde el patio de una casa; b) Imagen de la pared de la casa con croma en las ventanas; c) Vídeo resultado de la composición del primer vídeo con la pared de la casa.

10.3.2. SEGUNDA FASE: LLH V.2

Tras las primeras pruebas de laboratorio y las pruebas de usuarios realizadas (apartado 10.4), en esta segunda fase el sistema de *LLH* se modifica, introduciendo algunas novedades que tratan de fomentar los aspectos positivos y eliminar los negativos.

10.3.2.1. DESCRIPCIÓN TÉCNICA

Las nuevas aportaciones técnicas con respecto a la primera fase de laboratorio son (Figura 247):

1. Se incrementa el tamaño de la ventana de RA, haciéndola coincidir con la pantalla de PC utilizada (de 20"). De este modo la parte gráfica de la aplicación cobra mayor importancia, ya que se aprecian mejor los objetos 3D así como las pequeñas animaciones realizadas.
2. Se utilizan muñecos y piezas de LEGO más grandes, mediante los bloques de Duplo (bloques de LEGO para edades entre 2 y 5 años). Las dimensiones generales de la casa aumentan al igual que la de los muñecos, haciendo más

evidente la parte tangible de la aplicación, valorada muy positivamente por los usuarios.

3. Al incrementar el tamaño de los muñecos, las marcas de RA son un poco más grandes, ayudando a un mejor reconocimiento e identificación de éstas, permitiendo al mismo tiempo que puedan alejarse más de la cámara.
4. La superficie del suelo de la casa se recubre con una tela aterciopelada de tonalidad oscura, de tal modo que se reducen de forma considerable las reflexiones ocasionadas por las superficies satinadas de los bloques.
5. Se sustituye la cámara Web por otra con un mayor ángulo de apertura, de tal forma que al situarla más cerca de la casa abarque por completo el espacio físico de esta.
6. Se añaden elementos virtuales de comida y bebida sobre la mesa, y los sonidos correspondientes de comer y beber, programándose nuevas interacciones entre ambos muñecos y la casa: para terminarse la bebida o comida los muñecos deben sentarse juntos en la mesa y realizar la acción de comer o beber (activados mediante la inclinación); cuando se acaba la comida, hasta el día siguiente no se rellenan los vasos y el plato.
7. El proceso de calibración se simplifica mediante la incorporación al teclado de una serie de símbolos referentes a los muñecos y a los elementos de la casa.



Figura 247. Elementos de LLH en la segunda fase.

10.4. SISTEMAS DE VERIFICACIÓN

En las siguientes líneas se detallan los resultados de una prueba realizada en el Mixed Reality Lab de la National University of Singapore con la primera versión de *LLH* (Figura 248). Los usuarios que probaron el sistema fueron un total de seis de entre los investigadores del laboratorio e investigadores invitados, los cuales estaban en ese momento trabajando en el desarrollo de diversas aplicaciones interactivas dentro del ámbito de la realidad mezclada. Las áreas de especialización de los participantes abarca un diverso rango: informática, HCI, RA/RM, juego, arte neomedial, artes plásticas, ingeniería, multimedia y música. Sus edades están comprendidas entre los 24 y los 30 años. También cabe mencionar sus nacionalidades, puesto que pueden influir en la manera de valorar algunos aspectos; estos son: Alemán, Austriaco, Chino, Español, Singaporense de origen Chino-Malayo y Singaporense de origen Chino.



Figura 248. Usuarios interactuando con *LLH* V.1 en el Mixed Reality Lab.

Se valoraron un total de 16 aspectos de la aplicación, puntuándolos de 1 a 5: 1) Muy malo; 2) Malo; 3) Mejorable; 4) Bueno; 5) Muy bueno). A continuación se muestra una tabla con las características valoradas, las puntuaciones individuales de cada participante (A-F) y la media (MED):

CARACTERÍSTICAS	A	B	C	D	E	F	MED
El rol del juego: Explorar la coexistencia	4	5	4	4	4	5	4,33
El escenario colaborativo: 2 usuarios	5	4	4	4	4	5	4,33
El procedimiento de calibración mediante teclado	3	4	3	4	3	-	3,40
Los muñecos físicos como interfaz	5	5	5	5	5	5	5,00
La idea de que la casa esté "viva"	5	5	4	5	4	5	4,67
Datos generados por ordenador: Sonido	5	5	4	4	5	4	4,50
Modelos 3D	5	5	3	3	3	4	3,83
Videos	4	4	4	4	4	4	4,00
Ambigüedad en el género de los muñecos	4	3	3	5	4	5	4,00
Restricciones: Agua / sonido	4	5	2	4	4	4	3,83

La aparentemente aleatoriedad de los datos generados por ordenador	4	4	3	3	4	5	3,83
La TV como display para mostrar la RA	5	5	5	5	4	3	4,50
La manera en que se produce la interacción: Distancias y rotaciones	4	5	2	4	5	4	4,00
La coexistencia con tu pareja	5	5	3	4	3	5	4,17
Las posibilidades de educación mediante el juego propuesto	3	5	4	3	4	5	4,00
Puntuación general	4	5	4	4	4	4	4,17

Tabla 8. Puntuaciones individuales y medias (MED).

En esta tabla aparecen resaltadas las puntuaciones (para una misma característica) en las que todos los participantes individualmente han otorgado un 4 (bueno) o un 5 (muy bueno), con al menos tres puntuaciones de 5. Sin lugar a dudas cabe destacar la unanimidad en otorgar la puntuación máxima a la característica de “utilización de los muñecos físicos como interfaz”. También existe unanimidad en cuanto a los “vídeos” utilizados (refiriéndose a la temática de éstos), con una puntuación de 4 (bueno). La nota media más baja corresponde al “procedimiento de calibración mediante teclado”. La nota media individual más baja es un 2 (malo) y corresponde a las “restricciones de agua y sonido” y a “la manera en que la interacción se produce mediante distancias y rotaciones”; ambas puntuaciones fueron dadas por el mismo participante y contrastan con la de los otros participantes para estas mismas características (entre 4 y 5).

En la hoja de la prueba también se facilitó un espacio para añadir comentarios; a continuación se muestran y/o comentan algunos de ellos:

1. Uno de los participantes encontró que la utilización de la TV como ventana al mundo aumentado resultaba confusa al cambiar entre los vídeos (TV encendida) y el entorno aumentado (TV apagada). Proponía cambiar el tipo de display por un HMD.
2. Mayor interacción entre los muñecos, es decir, que puedan comer o bailar juntos, en lugar de separadamente.
3. Incrementar las actividades de los muñecos y elaborar más la interactividad.
4. Añadir más efectos para hacer que la coexistencia sea más evidente en el juego; por ejemplo, poner comida encima de la mesa dependiendo de si se sienta un muñeco o los dos.

5. El procedimiento de calibración resulta complicado para que pueda ser utilizado por los niños.

A parte de estas pruebas escritas, el programa también lo han probado otros usuarios, incluyendo niños. En la Figura 249 pueden verse imágenes de una niña de dos años interactuando con la segunda versión de LLH.



Figura 249. Niña de dos años jugando con LLH.

Tras estas pruebas se puede concluir que en niños menores de 4/5 años el display carece de importancia, puesto que no son capaces de entender la aplicación de RA (no relacionan los espacios reales y virtuales) y se fijan más en su propia imagen representada en la pantalla. Sin embargo, sienten un gran interés por la parte física del juego (los bloques de LEGO) ya que les resultan altamente familiares; La parte mediática/virtual pasa casi desapercibida.

11. CATEDRAR

En este ensayo se presenta *CatedrAR*, un trabajo de realidad aumentada aplicado al campo del patrimonio cultural, para mostrar de forma conjunta elementos presentes y pasados (parcial o totalmente desaparecidos). El trabajo se desarrolla para aumentar el interior de la bóveda de la Catedral de Valencia en su antiguo estado barroco, y para aumentar el actual retablo del altar con una imagen de su homólogo renacentista. La finalidad es recrear ambos escenarios dentro de la Catedral.

Motivación

El modelado virtual y la reconstrucción 3D son herramientas comúnmente utilizadas en el campo del patrimonio cultural para recrear, analizar y visualizar objetos de pequeña escala como excavaciones arqueológicas y monumentos arquitectónicos, y objetos de gran escala, como esculturas, piezas de cerámica, etc. En la actualidad es posible aplicar diferentes tecnologías para realizar modelos 3D fotorrealísticos, sin embargo, el modelado preciso de estos elementos es normalmente costoso y caro, puesto que la realidad física es por sí compleja; además, los modelos representados en este campo suelen pertenecer a monumentos/elementos derruidos, desaparecidos, o que en la actualidad no se conservan en un estado óptimo. En este sentido la tecnología de la realidad aumentada puede jugar un papel importante, ya que permite la visualización de forma simultánea de elementos reales y virtuales, sin la necesidad de modelar el objeto entero.

Marco temporal

Este ensayo comienza en el año 2007 y surge a partir de una colaboración entre los Departamentos de Ingeniería Cartográfica, Geodesia y Fotogrametría y de Conservación y Restauración del Patrimonio Cultural de la UPV, en un proyecto de recuperación y restauración de la bóveda renacentista del altar mayor de la Catedral de Valencia. La finalidad de la aplicación creada es la visualización de elementos arquitectónicos desaparecidos en el entorno físico donde estaban antaño, con la condición de que el entorno físico debe permanecer inalterado, es decir, no se ha situado en el entorno físico ningún tipo de dispositivo (ej. sensores, displays o marcas artificiales).

11.1. PLANTEAMIENTO

En *CatedrAR* se aumentan dos elementos arquitectónicos de la Catedral de Valencia: el interior de la bóveda y el retablo del altar mayor. A continuación se presenta una breve historia de la Catedral, introduciendo ambos elementos arquitectónicos tal y como están en la actualidad y como estaban antaño, sirviendo este último estado como elemento virtual en el entorno aumentado establecido.

La Catedral de Valencia data del año 1262, y presenta una gran variedad de estilos, incluyendo el románico, gótico y barroco; principalmente refleja un estilo gótico temprano, aunque el altar mayor es de estilo barroco a excepción del retablo central, que es renacentista. A partir de unos trabajos de restauración que empezaron en el año 2004 en la zona del altar mayor, se encontró que la bóveda barroca (Figura 250-b) escondía una bóveda más antigua de estilo gótico (Figura 250-a), que había permanecido oculta durante más de cinco siglos. Sin embargo, el hecho más relevante es que la bóveda gótica es el soporte físico de unos frescos únicos renacentistas pintados por los maestros italianos Franco Pagano y Paolo de San Leocadio (Pérez García (ed.), 2006). Después de una serie de discusiones socio-políticas, la bóveda barroca fue finalmente desmantelada para que los frescos renacentistas fueran visibles y pudieran ser también restaurados.

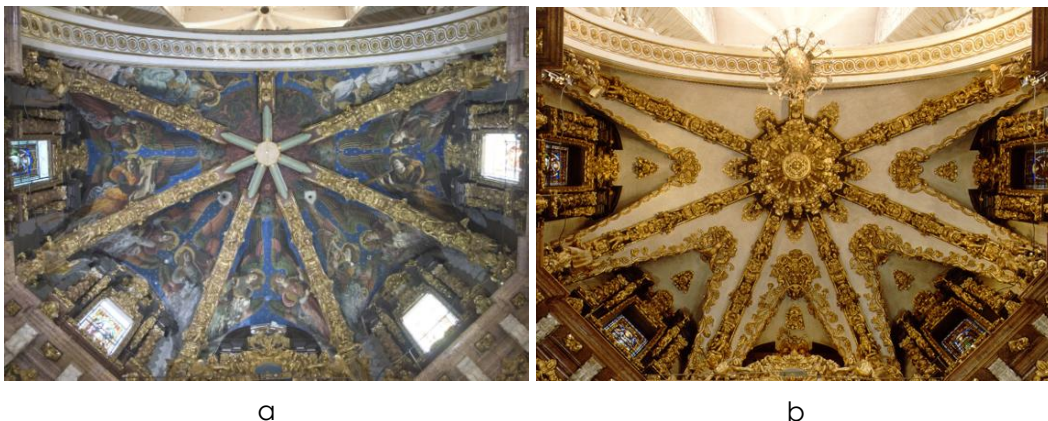


Figura 250. Bóveda del altar mayor: a) Bóveda gótica (antes de ser restaurada); b) Bóveda barroca (antes de ser desmantelada).

Previamente al desmantelamiento de la bóveda barroca, se hicieron varios trabajos de fotogrametría con tal de adquirir los modelos tridimensionales, uno de cada bóveda, mediante tecnología láser escáner terrestre. El proceso de documentación gráfica y morfológica aparece descrito en (Lerma et al., 2005; Biosca Taronger et al., 2007).

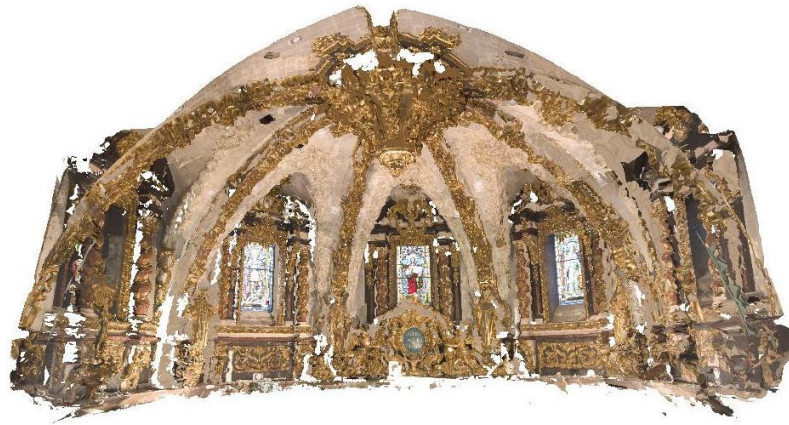


Figura 251. Modelo 3D de la bóveda barroca obtenido con un láser escáner terrestre. En (Biosca Taronger et al., 2007).

En cuanto al retablo de plata desaparecido, podemos encontrar algunos datos en (Las Provincias, 2006). El retablo de plata fue elaborado entre 1489 y 1506 por Piero da Ponce, Augustine Nicos y los plateros valencianos Francisco Cetina y Bernat J. Cetina. El retablo representaba los Siete Gozos de la Virgen, con una imagen central representando la Asunción, y una coronación de la Virgen en la parte superior. El conjunto se completa con sus puertas monumentales que se abrían únicamente en ocasiones especiales, decoradas por las dos caras con motivos marianos. En el siglo XIX el retablo de plata fue derretido en Mallorca, durante la guerra de la Independencia, para convertirlo en moneda y pagar a las tropas extranjeras que cooperaban con España en su lucha contra Napoleón. En la actualidad, únicamente prevalece un dibujo del siglo XVIII guardado en los archivos de la Catedral; una imagen de este dibujo es la que se utiliza en este ensayo para aumentar el actual retablo.

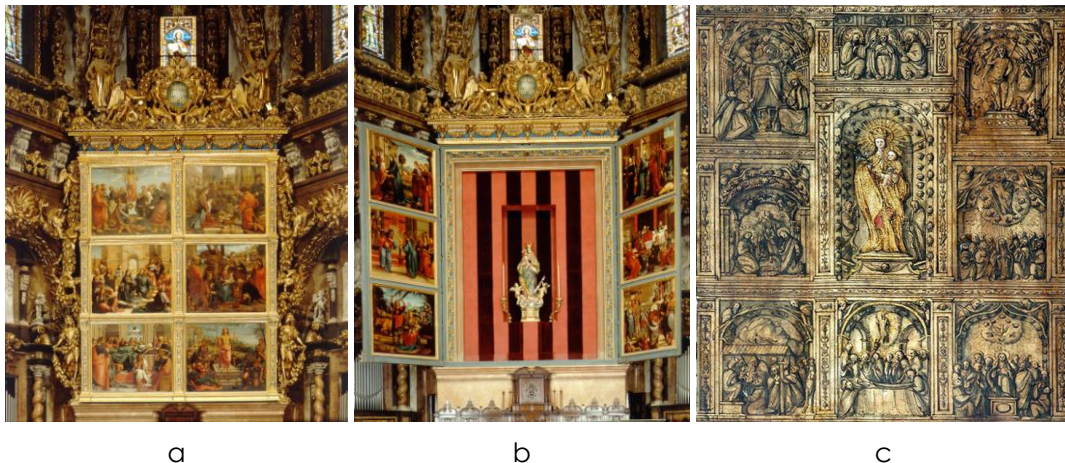


Figura 252. Retablo del altar mayor: a) Retablo actual con las puertas cerradas; b) Retablo actual con las puertas abiertas; c) Retablo de plata desaparecido.

Hay que señalar que la finalidad de la aplicación creada en este ensayo es la visualización de elementos arquitectónicos desaparecidos en el entorno físico donde estaban antaño, y por ello no tiene otra lectura conceptual. Sin embargo, desde el punto de vista técnico resulta muy interesante, puesto que una de las restricciones establecidas – la no utilización de ningún sensor, display o marcas artificiales en el entorno físico – condiciona el tipo de registro y software utilizado (las librerías BazAR), que resulta novedoso con respecto al resto de ensayos presentados en esta tesis, dando lugar a un estudio comparativo más completo (aportado en las conclusiones de la tesis). Las cuestiones que se resuelven en *CatedrAR*, desde el punto de vista técnico, son esencialmente la implementación de imágenes con canal alfa (apartado 11.2.2.1) y la elección de la imagen de referencia (apartado 11.2.2.2).

11.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO

Siguiendo las aportaciones del bloque I, se puede elaborar un marco general de este ensayo:

- Ubicación dentro del continuo de Milgram (apartado 1.1.1):
 - Realidad aumentada.
- Clasificación (apartado 1.2.2):
 - Según el entorno físico en el que se desarrolla la aplicación:
 - Cerrada.
 - Según la extensión que abarca:
 - Local.
 - Según la movilidad de los dispositivos de registro y/o displays:
 - Móvil.
 - Según los usuarios que simultáneamente pueden interactuar con el sistema:
 - Individual.
 - Según el tipo de colaboración establecida:
 - (No existe colaboración).
 - Clasificación según Mackay (apartado 1.2.2):
 - Aumentar el usuario.
- Campo de aplicación (apartado 1.4):
 - Arte (patrimonio cultural).

- Educación.
- Sistema de coordenadas terreno/objeto (apartado 2.1.1):
 - Centrado en objetos.
- Orientación espacial (apartado 2.1.5):
 - Método indirecto.
- Dispositivos de registro / controladores (capítulo 3):
 - Sensores ópticos: Rasgos naturales.
- Displays (capítulo 4):
 - Display sujeto a la cabeza: De vídeo.
- Software (capítulo 5):
 - Reconocimiento de rasgos naturales: BazAR.
- Forma de interactuar (apartado 6.1.1):
 - Navegación.

11.1.2. REFERENTES

El rápido incremento de la capacidad de procesamiento y almacenamiento en los ordenadores durante la última década ha posibilitado que técnicamente sea asequible la adquisición, almacenamiento y manipulación de datos virtuales 3D, incluso mediante el uso de ordenadores personales de gama media. Según (Pieraccini et al., 2001) el creciente interés de la comunidad científica respecto al modelado 3D del patrimonio cultural se debe a las necesidades de conservación, incluyendo tareas como la compilación de archivos digitales de modelos 3D, la adquisición de réplicas minuciosas de piezas artísticas, la difusión del patrimonio cultural, o la posibilidad de restauración y monitorización digital. Por ejemplo, en (Boehler et al., 2003) se consigue un modelo digital de la tumba del emperador alemán Maximiliano I a partir de técnicas de fotogrametría y escáner 3D (Figura 253-a). En (Gutierrez et al., 2004) se reconstruye un antiguo suburbio musulmán de Sinhaya (Zaragoza) siguiendo técnicas topográficas, a partir de las cuales se adquieren modelos precisos y texturas, aplicándoles a continuación condiciones de luminosidad virtuales para incrementar el realismo; este modelo 3D se presenta al usuario mediante un sistema de proyecciones tipo CAVE (Figura 253-b). En (Koutsoudis et al., 2007) se presenta una reconstrucción 3D fotorrealística para una parte de la vieja ciudad de Xanthi (Tracia, Grecia) mediante la utilización de software libre basado en gráficos 3D (Figura 253-c).

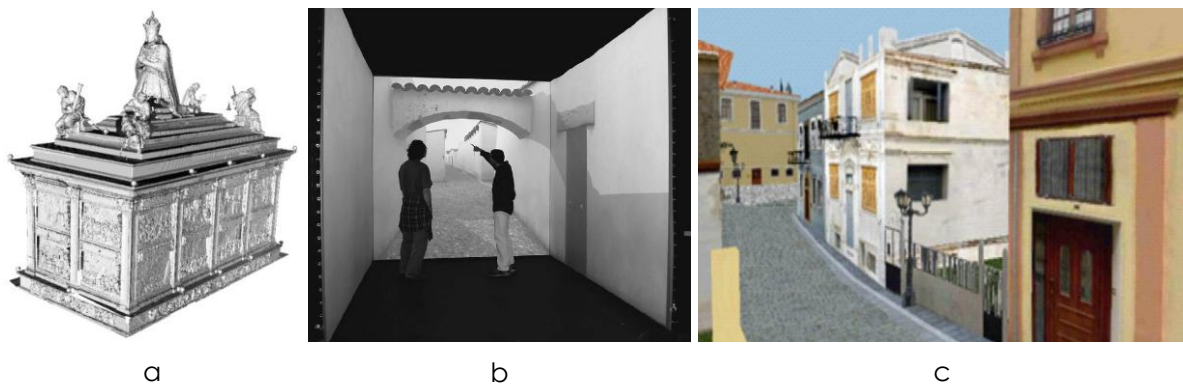


Figura 253. Reconstrucciones virtuales 3D: a) Modelo 3D de la tumba del emperador Maximiliano I, en (Boehler et al., 2003); b) Sistema CAVE para el antiguo suburbio musulmán Sinhaya, en (Gutierrez et al., 2004); c) Reconstrucción 3D fotorrealística de la ciudad de Xanthi, en (Koutsoudis et al., 2007).

Con el mayor incremento de capacidades computacionales en los últimos años, surgen algunos trabajos en el campo del patrimonio cultural en el que se hace uso de la tecnología de realidad aumentada. Por ejemplo, en (Aguiló et al., 2001) se presenta un sistema de realidad aumentada para visitas a excavaciones arqueológicas. De igual modo, en (Gleue et al., 2001; Vlahakis et al., 2001) se presenta el proyecto ARCHEOGUIDE, un sistema de realidad aumentada que visualiza reconstrucciones de antiguas ruinas basándose en la posición y orientación del usuario a partir de datos GPS y una brújula magnética (Figura 254).

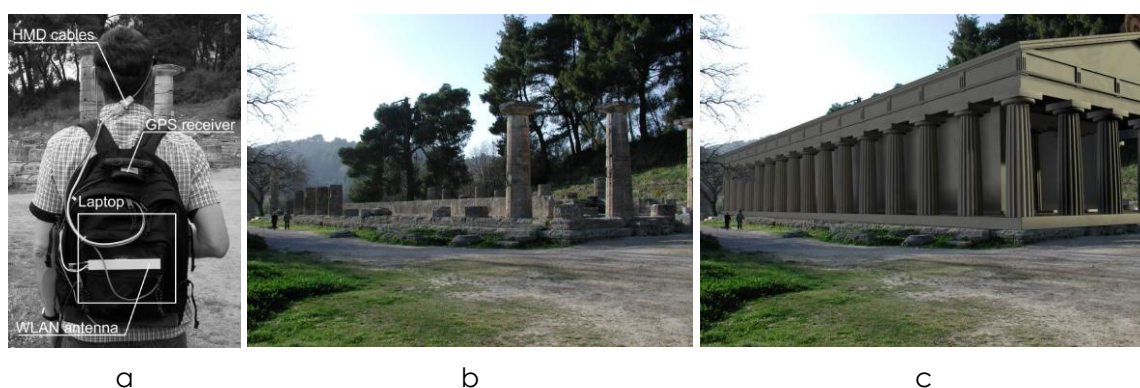


Figura 254. ARCHEOGUIDE: a) Usuario equipado con el sistema; b) Imágenes de ruinas griegas; c) Imagen aumentada con reconstrucciones 3D. En (Gleue et al., 2001; Vlahakis et al., 2001).

Existen otros autores que han realizado sistemas de RA para museos, como en (Hall et al., 2002; Bimber et al., 2003; Brown et al., 2003; Hatala et al., 2004; Wojciechowski et al., 2004; Woods et al., 2004). Por ejemplo, en (Wojciechowski et al., 2004) se presenta un sistema de RA para que los visitantes interactúen con diversas piezas modeladas en 3D,

las cuales también son accesibles a través de Internet. El trabajo se incluye dentro del programa ARCO (*Augmented Representation of Cultural Objects*), que impulsa el desarrollo de las tecnologías que permiten la creación, manipulación, gestión y presentación de objetos culturales digitalizados en exhibiciones virtuales accesibles dentro y fuera de museos (Figura 255).

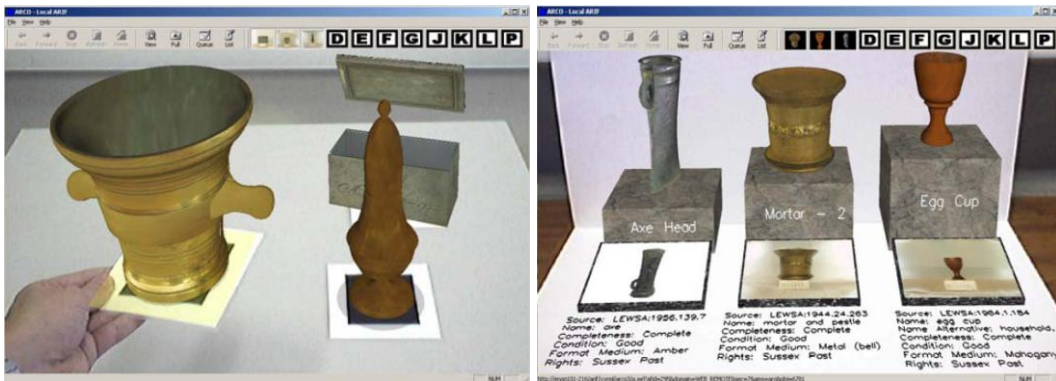


Figura 255. Escenas de RA. En (Wojciechowski et al., 2004).

11.2. DESCRIPCIÓN TÉCNICA

11.2.1. COMPONENTES DE LA APLICACIÓN

11.2.1.1. DISPOSITIVOS FÍSICOS

El sistema de *CatedrAR* consta de los siguientes dispositivos (Figura 256):

- Un ordenador estándar dentro de una mochila;
- Un HMD: I-glasses SVGA video (ver apartado 4.1.1.2, Figura 98);
- Una cámara Web: *CREATIVE Live! Cam Voice* con una resolución de 640x480 píxeles, conexión USB 2.0 y FOV de 85°, situada encima del HMD;
- Baterías para el HMD (opcional, ya que también existe toma de luz directa).

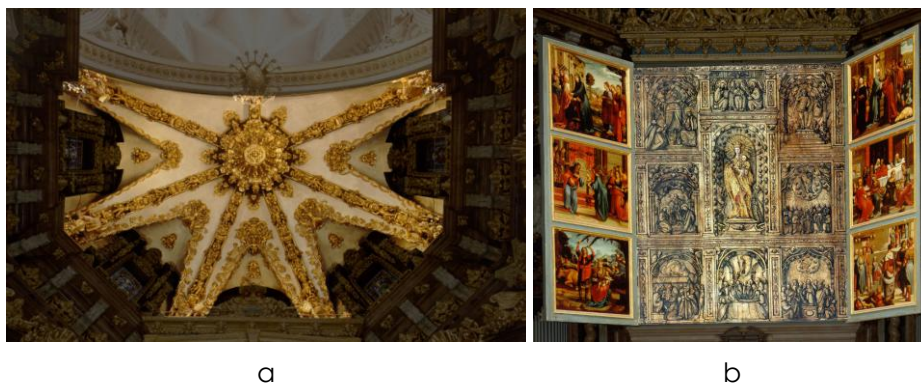


Figura 256. Dispositivos de *CatedrAR*.

Para el registro del usuario en tiempo real y la construcción del escenario aumentado se han empleado las librerías BazAR. Esta elección (en lugar de ARToolKit o MXRToolKit) está condicionada por la imposibilidad de poner ningún elemento artificial en el entorno físico. Así pues, en *CatedrAR* el usuario lleva consigo todos los dispositivos necesarios para el registro, generación y presentación de la escena de RA, de tal forma que el entorno permanece inalterado. El usuario puede moverse libremente por la zona del altar mayor.

11.2.1.2. ELEMENTOS VIRTUALES

Para aumentar tanto la bóveda como el retablo, se han modelado dos planos en OpenGL y a continuación se han mapeado con texturas en formato targa, con su correspondiente canal alfa en ambos casos (Figura 257). Las librerías de BazAR no tienen implementadas la posibilidad de cargar imágenes con canal alfa, por lo que ha sido necesario su desarrollo (ver apartado 11.2.2.1)



a

b

Figura 257. Imágenes utilizadas para el mapeado de planos, donde la zona en sombra corresponde a la transparencia del canal alfa: a) Imagen de la bóveda; b) Imagen del retablo con las puertas abiertas.

También se ha elaborado un modelo 3D simplificado a partir del modelo 3D obtenido con el escáner (Figura 258-b). Este modelo se prevé pueda implementarse en futuros desarrollos de las librerías BazAR, por ejemplo, mediante la integración de las librerías OpenVRML, al igual que se hace en ARToolKit.

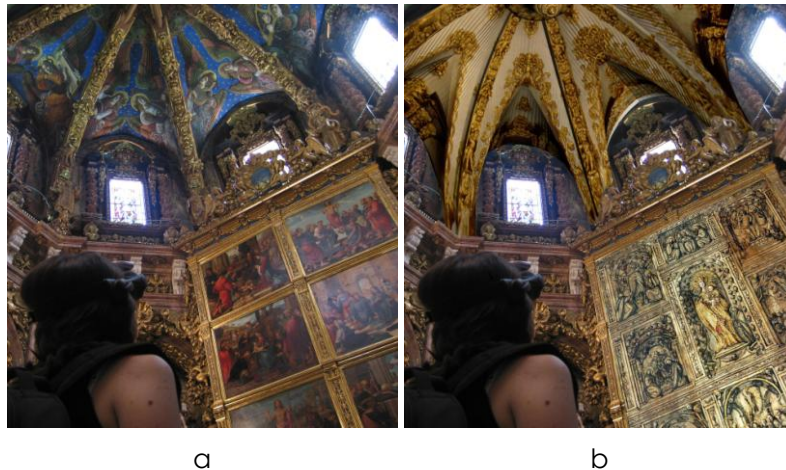


Figura 258. Usuario con los dispositivos de *CatedrAR*: a) Entorno real; b) Simulación de realidad aumentada con una bóveda virtual tridimensional.

11.2.2. DESARROLLO, PROCESO

La primera vez que se ejecuta el programa bazAR, el sistema realiza una auto-calibración de la cámara. Para ello, se selecciona manualmente una zona de la imagen de entrada (imagen de referencia), como base para realizar una correspondencia entre puntos de interés de la imagen y los del objeto, basándose en el registro de rasgos naturales puntuales (Lepetit et al., 2005). A continuación, se calcula la matriz de transformación de la cámara con respecto al objeto, y se dibujan los objetos virtuales, aplicándoles también iluminación a partir de una calibración fotométrica. En los siguientes subapartados se explica la implementación del objeto virtual empleado y se muestran distintas pruebas (o casos de estudio) para la elección de la imagen de referencia para calcular las correspondencias.

11.2.2.1. IMPLEMENTACIÓN DE IMÁGENES CON CANAL ALFA

Como se ha explicado en el apartado 11.2.1.2, en este ensayo se ha utilizado como elemento virtual un plano dibujado mediante código OpenGL y mapeado con una imagen que contiene canal alfa. A continuación se detalla el código que se ha

insertado en el entorno de programación de BazAR para cargar una imagen de tipo targa y mapearla sobre el plano del entorno 3D.

Cargar una imagen targa, a partir de (Hawkins, 2001):

```
typedef struct tagTGAFILEHEADER
{
    unsigned char imageIDLenght;
    unsigned char colorMapType;
    unsigned char imageTypeCode;
    short int colorMapOrigin;
    short int colorMapLenght;
    short int colorMapEntrySize;
    short int imageXOrigin;
    short int imageYOrigin;
    short int imageWidth;
    short int imageHeight;
    unsigned char bitCount;
    unsigned char imageDescriptor;
} TARGAFILEHEADER;

typedef struct
{
    unsigned char imageTypeCode;
    short int imageWidth;
    short int imageHeight;
    unsigned char bitCount;
    unsigned char *imageData;
} TGAFILE;

int LoadTGAFfile(char *filename, TGAFILE *tgaFile)
{
    FILE *filePtr;
    unsigned char ucharBad;
    short int sintBad;
    long imageSize;
    int colorMode;
    long imageIdx;
    unsigned char colorSwap;

    filePtr = fopen(filename, "rb");
    if (!filePtr)
        return 0;

    fread(&ucharBad, sizeof(unsigned char), 1, filePtr);
    fread(&ucharBad, sizeof(unsigned char), 1, filePtr);
```

```
fread(&tgaFile->imageTypeCode, sizeof(unsigned char), 1, filePtr);

if((tgaFile->imageTypeCode != 2) && (tgaFile->imageTypeCode != 3))
{
    fclose(filePtr);
    return 0;
}

fread(&sintBad, sizeof(short int), 1, filePtr);
fread(&sintBad, sizeof(short int), 1, filePtr);
fread(&ucharBad, sizeof(unsigned char), 1, filePtr);
fread(&sintBad, sizeof(short int), 1, filePtr);
fread(&sintBad, sizeof(short int), 1, filePtr);

fread(&tgaFile->imageWidth, sizeof(short int), 1, filePtr);
fread(&tgaFile->imageHeight, sizeof(short int), 1, filePtr);

fread(&tgaFile->bitCount, sizeof(unsigned char), 1, filePtr);

fread(&ucharBad, sizeof(unsigned char), 1, filePtr);

colorMode = tgaFile->bitCount/8;
imageSize = tgaFile->imageWidth * tgaFile->imageHeight * colorMode;

tgaFile->imageData = (unsigned char*)malloc(sizeof(unsigned char)*imageSize);

fread(tgaFile->imageData, sizeof(unsigned char), imageSize, filePtr);

for (imageIdx = 0; imageIdx < imageSize; imageIdx += colorMode)
{
    colorSwap = tgaFile->imageData[imageIdx];
    tgaFile->imageData[imageIdx] = tgaFile->imageData[imageIdx + 2];
    tgaFile->imageData[imageIdx + 2] = colorSwap;
}

fclose(filePtr);

return 1;
}
```

Dibujar un plano y mapearlo con la imagen:

```
glDisable(GL_LIGHTING);

// Cargamos la imagen TGA:
```

```
    TGAFILE *myTGA;
    myTGA = (TGAFILE*)malloc(sizeof(TGAFILE));
    LoadTGAFile("Cenital-barroca-alpha-72.tga", myTGA);

    glTexEnvf(GL_TEXTURE_ENV, GL_TEXTURE_ENV_MODE, GL_REPLACE); // No combinamos la
    textura con el color original del plano.

    glTexImage2D(GL_TEXTURE_2D, 0, 4, myTGA->imageWidth, myTGA->imageHeight, 0,
    GL_RGBA, GL_UNSIGNED_BYTE, myTGA->imageData); // definimos la textura 2D.

    glEnable (GL_TEXTURE_2D); // Habilitamos el mapeado de la textura.

    // Definimos el ancho y alto del plano:
    int w = 260;
    int h = 160;

    glTranslatef(10,0,0); // Aplicamos una traslación.

    // Dibujamos el plano:
    glBegin(GL_QUADS);
    glTexCoord2f(0, 0);
    glVertex2f(-w, h);
    glTexCoord2f(1, 0);
    glVertex2f(w, h);
    glTexCoord2f(1, 1);
    glVertex2f(w, -h);
    glTexCoord2f(0, 1);
    glVertex2f(-w, -h);
    glEnd();

    free(myTGA->imageData);

    glDisable (GL_TEXTURE_2D); // Deshabilitamos el mapeado de la textura.
```

11.2.2.2. CASOS DE ESTUDIO

En este apartado se presentan diversos casos para la elección de la imagen de referencia. En el caso de la aumentación de la bóveda, un gran inconveniente encontrado es que el objeto que se va a aumentar no constituye una geometría plana

– normalmente, el objeto a aumentar constituye la imagen de referencia. De otro modo, en el caso del retablo, éste sí se ajusta a dicha geometría, por lo que se espera que los resultados sean más robustos. A continuación se presentan tres casos estudiados: aumentar sólo la bóveda, aumentar sólo el retablo, aumentar el retablo y la bóveda.

Primer caso: Recreación de la bóveda

En primer lugar se realizaron varias pruebas para aumentar únicamente la bóveda. Para ello, se tomó como imagen de referencia la porción interna de la bóveda (Figura 259), que se aproxima más a una figura plana debido a la distancia que la separa de la cámara (unos 20 m).

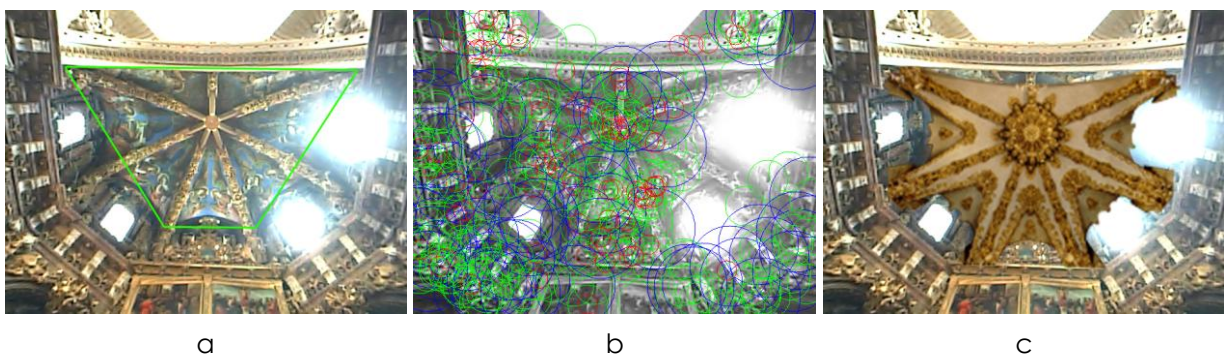


Figura 259. Imagen de la bóveda tomada desde debajo de esta: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.

La aplicación funciona de forma correcta en un área delimitada de unos 2x2 m, justo debajo de la bóveda; esta es el área que geoméricamente se corresponde con la imagen de referencia, puesto que las deformaciones ocasionadas por la concavidad de la bóveda al cambiar la perspectiva están dentro de un rango de tolerancia que permite que se establezca la correspondencia. Sin embargo, podemos añadir que, aunque la versión libre de bazAR no permite el registro sobre imágenes no planas, existen algunos desarrollos de sus autores, descritos en (Salzmann et al., 2006), mediante los que se establece correspondencia incluso sobre superficies deformables.

Las condiciones de luz implican un problema adicional en la imagen de referencia de la bóveda, ya que durante gran parte del día la luz solar incide directamente sobre una o varias de las cuatro ventanas situadas alrededor del fresco (dos orientadas al

Este y las otras dos al Oeste). Es por ello que se hicieron algunas pruebas de noche, con iluminación artificial; está solución ofrecía resultados más estables.

Segundo caso: Recreación del retablo

En este caso se aumentó el retablo del altar mayor. Al estar el retablo cerrado, se tomó como imagen de referencia los dibujos que éste contiene; dicha imagen encaja en una superficie plana.

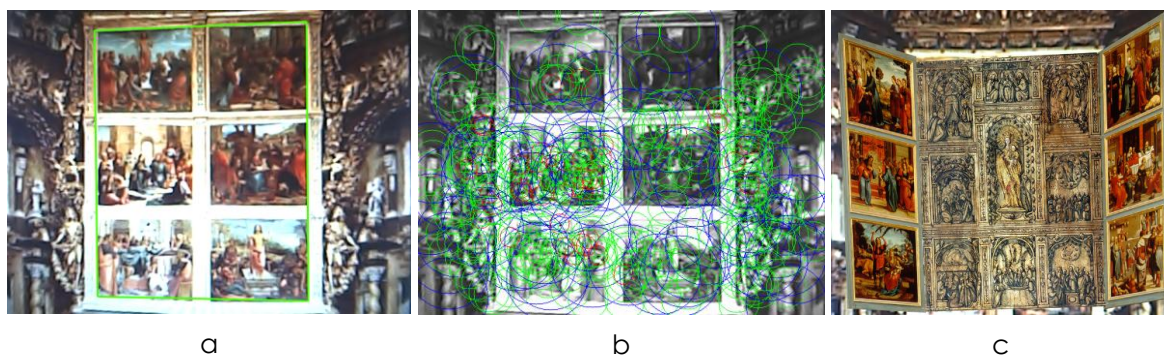


Figura 260. Imagen del retablo: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.

En este caso, la aplicación funciona mejor en un área de mayores dimensiones debido a que se evitan los problemas encontrados en el primer caso, ya que la imagen de referencia es una figura plana y las condiciones de luminosidad son adecuadas. Es por ello que este caso es más estable que el primero, tanto de día como de noche.

Tercer caso: Recreación de la bóveda y del retablo

Finalmente, se aumentaron conjuntamente la bóveda y el retablo. Se realizaron varias pruebas desde diferentes puntos de vista, tomando como imagen de referencia distintas superficies. En un primer momento se pensó en la imagen del retablo como imagen de referencia por ser de geometría plana, sin embargo, para que el usuario vea de forma conjunta la bóveda y el retablo, debe alejarse bastante, por lo que la superficie que ocupa el retablo en la imagen es demasiado pequeña para establecerla como referencia en el cálculo de las correspondencias. Es por ello que la imagen de referencia debe abarcar un área mucho mayor, cubriendo gran parte de la imagen registrada por la cámara, por lo que vuelve a surgir el problema de la

tridimensionalidad del objeto registrado. En Figura 261 y Figura 262 podemos ver dos ejemplos de imágenes de referencia tomadas desde distintos puntos de vista. Al igual que en el primer caso de estudio, la aplicación funciona correctamente dentro de un área pequeña, restando movilidad al usuario.

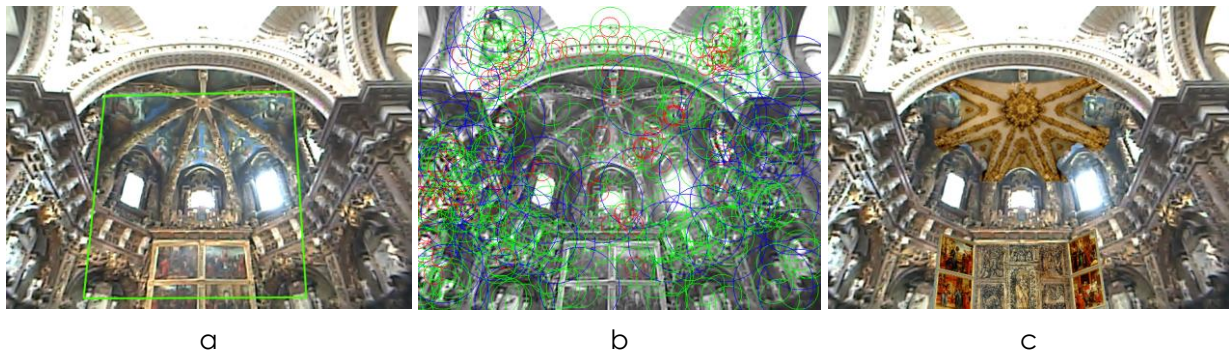


Figura 261. Imagen tomada desde el altar mayor: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.

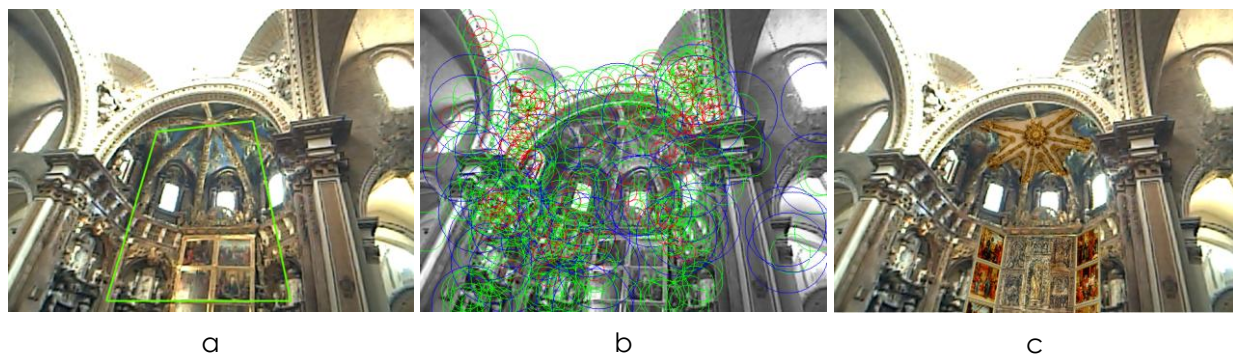


Figura 262. Imagen tomada desde una esquina del altar mayor: a) Imagen de referencia; b) Selección de puntos para la correspondencia; c) Imagen aumentada.

11.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS

Una de las principales restricciones encontradas ha sido la exigencia de no alterar de ninguna manera el entorno físico. Ello ha condicionado tanto el tipo de display y hardware utilizado como el sistema de registro: el usuario lleva consigo todos los dispositivos (display, cámara y ordenador); y para el sistema de registro se han utilizado las librerías de BazAR, que calculan la matriz de transformación a partir de una imagen del entorno físico (y no de una marca artificial, como es el caso de ARToolkit). Sin embargo, la utilización de estas librerías requiere del registro a partir de una imagen sobre una geometría plana. Como se ha visto en el apartado 11.2.2.2, la geometría del

entorno aumentado y las condiciones de luz específicas de la parte de la bóveda, han hecho que el establecimiento de una solución estable no sea fácil, por lo que se ha restringido el espacio físico por donde se puede mover el usuario.

12. AR-JAZZ

AR-Jazz es una aplicación de realidad aumentada pensada especialmente para representaciones instrumentales en vivo. En este ensayo se explora el sonido y el movimiento como entrada de datos que gestionan la representación proyectada de una figura geométrica, siendo un grupo de instrumentistas y/o intérpretes quienes de forma conjunta controlan la visualización del objeto aumentado.

Motivación

El campo de la realidad aumentada ha tenido hasta el momento poca investigación en lo referente al sonido. Los parámetros de altura, intensidad, timbre y espacialización son los aspectos que determinan el lenguaje del sonido. Así pues, estos cuatro factores pueden servirnos como punto de partida en la indagación de realidades sonoras aumentadas. El ensayo que se muestra a continuación desarrolla algunos de estos parámetros aumentándolos en una proyección visual. El contacto físico del músico con su instrumento permite establecer otro parámetro, como es el de la sinestesia. Esta percepción gestual en la técnica interpretativa se transmite a la forma y características del objeto aumentado, creándose una interacción entre el músico y el computador a través del sonido emitido, que puede condicionar la interpretación del primero.

Marco temporal

Este ensayo se ha realizado para un evento específico, Jazz & Arts Meeting dentro del festival Jazz Panorama 2007, organizado por la Big Band Sedajazz y celebrado en el auditorio de la localidad de Torrent los días 18, 19 y 20 de julio. En este evento se proponen diferentes actividades, entre las que se incluyen: ponencias, experiencias, instalaciones, exposiciones, talleres, combos, videoconferencias, *Jam sessions*, clases de instrumentos, big band, conciertos y representaciones (*performances*).

12.1. PLANTEAMIENTO

El festival Jazz Panorama es un evento anual organizado por Sedajazz desde el año 1997. Inicialmente se ha celebrado en el Hort de Trénor de Torrent, pero desde hace unos años tiene lugar en el Auditori de Torrent. En este festival se organizan una serie de conciertos en vivo al que son invitados distintos artistas del mundo del Jazz; para este año cabe destacar la participación de Latino Blanco Band, Eliane Elias, Pierrick Pedron Quartet, Kurt Elling y Sopra Le Nuvole.

Jazz & Arts Meeting es un evento celebrado por primera vez en el año 2007, dentro del festival Jazz Panorama, para propiciar la interacción del Jazz con otras artes como el cine, teatro, pintura, escultura, poesía, electrónica, multimedia, etc., mediante diferentes experiencias de carácter formativo y con el fin de crear sinergias creativas. El evento se organiza mediante una serie de clases magistrales y prácticas, dirigidas tanto a músicos como a artistas de otras disciplinas. Con todo ello se pretende acortar distancias entre artistas de diferentes ámbitos, para seguir con la difusión del Jazz como lenguaje universal.

Para este evento, se nos propuso preparar un taller y realizar una representación en vivo de 15 minutos de duración con los alumnos asistentes al taller, que tendría lugar en el escenario del auditorio junto con las representaciones de otros talleres. Para se realizó una aplicación de realidad aumentada en la que se visualizaran los sonidos generados por los instrumentistas y movimientos de uno de ellos dentro de un entorno aumentado. En la fase del taller, los músicos adquieren conocimientos sobre la realidad aumentada, cómo funciona la aplicación y cómo interactuar con ella; también se explica el sistema a la audiencia antes de la representación. Así pues, podemos catalogar este ensayo dentro del ámbito del arte y del edutainment. Hay que mencionar que durante la fase de taller, el feedback que se estableció con los músicos permitió modificar algunos parámetros de la aplicación para que se visualizase de una forma más didáctica y precisa.

En este ensayo conté con la colaboración del músico y compositor Carlos David Perales, al que debo agradecer su implicación tanto en la fase de laboratorio como en la preparación del taller y posterior actuación en la performance como pianista. También debo agradecer la participación del Laboratorio de Luz durante la fase de montaje y a Toni Belenguer de Sedajazz por su implicación en la fase del ensayo y representación posterior.

12.1.1. MARCO EN EL QUE SE INSCRIBE EL ENSAYO

Siguiendo las aportaciones del bloque I, se puede elaborar un marco general de este ensayo:

- Ubicación dentro del continuo de Milgram (apartado 1.1.1):
 - Realidad aumentada.
- Clasificación (apartado 1.2.2):
 - Según el entorno físico en el que se desarrolla la aplicación:
 - Cerrada.
 - Según la extensión que abarca:
 - Local.
 - Según la movilidad de los dispositivos de registro y/o displays:
 - Espacial.
 - Según los usuarios que simultáneamente pueden interactuar con el sistema:
 - Colaborativa.
 - Según el tipo de colaboración establecida:
 - Presencial.
 - Clasificación según Mackay (apartado 1.2.2):
 - Aumentar el entorno, el usuario y el objeto.
- Campo de aplicación (apartado 1.4):
 - Arte.
 - Entretenimiento.
 - Educación.
- Sistema de coordenadas terreno/objeto (apartado 2.1.1):
 - Terrestre local.
- Orientación espacial (apartado 2.1.5):
 - Método directo.
- Dispositivos de registro / controladores (capítulo 3):
 - Sensor inercial.
 - Sensores acústicos: Micrófono.
- Displays (capítulo 4):
 - Displays espaciales: Proyector.
- Software (capítulo 5):
 - Multitarea: Max/MSP Jitter.

- Forma de interactuar (apartado 6.1.1):
 - Manipulación de objetos virtuales/reales.
 - Interacción con otros usuarios:
 - Colaboración presencial/directa.

12.1.2. REFERENTES

Diversos autores han tratado de darle forma al sonido mediante la generación de gráficos por ordenador. Por ejemplo, en (Paul, 2003, 137) se describe la pieza del artista japonés Toshio Iwai, *Piano – as image media* (1995), en la que se utiliza una partitura virtual para activar las teclas de un piano, induciendo la proyección de imágenes generadas por ordenador sobre una pantalla (Figura 263-a). La partitura la escriben los usuarios, mediante la asignación de puntos sobre una rejilla proyectada delante del piano. Tanto la melodía como las imágenes gráficas se generan de acuerdo a la distribución de los puntos en la rejilla.

Telesymphony (2001) de Golan Levin es una performance en la que los sonidos están generados por los teléfonos móviles de la audiencia (Paul, 2003). Esta pieza se presentó en un concierto que tuvo lugar en el festival de Ars Electronica de Linz, en el año 2001 (Figura 263-b). A la entrada, se les pidió a los participantes registraran sus números de teléfono móvil; a continuación se les asignó un asiento en la sala y se les enviaron una serie de nuevos tonos. Como cada uno de los asistentes tenía un asiento y tono asignado, la performance se podía dirigir perfectamente desde un equipo central mediante llamadas a los móviles. Al sonar el móvil de algún asistente, éste era iluminado por un foco de luz. El resultado se visualizaba en una serie de pantallas a los lados del escenario.

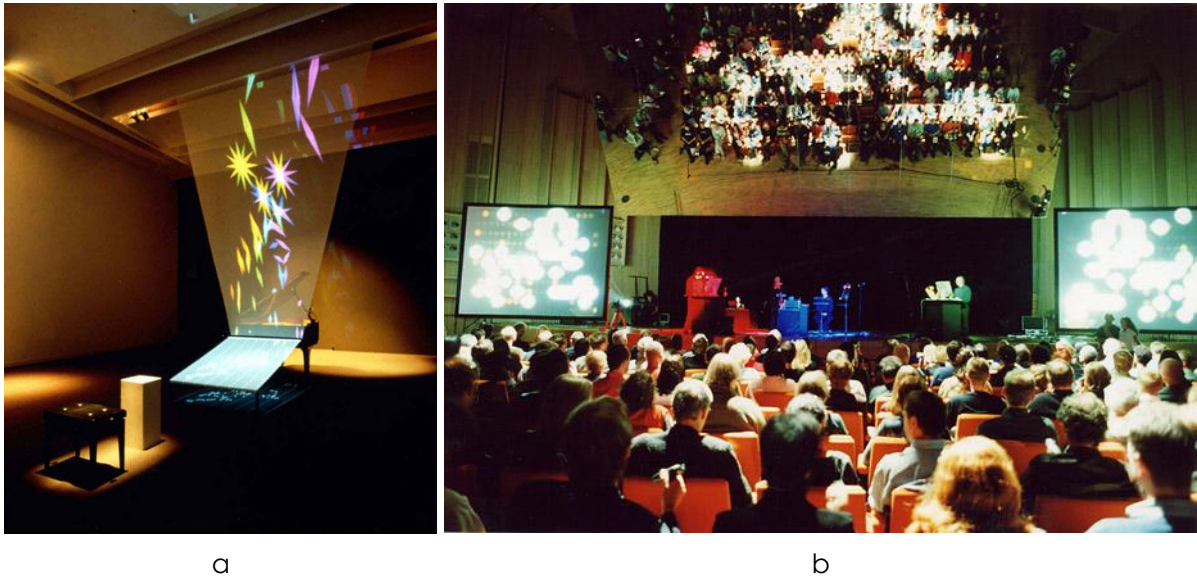


Figura 263. a) Piano – as image media; b) Telesymphony. En (Paul, 2003).

Messa di Voice (2003) es una performance audiovisual de realidad aumentada (Figura 264), creada por Golan Levin en colaboración con Zachary Lieberman, en la que se utiliza el registro de los movimientos del cuerpo de dos actores en combinación con el análisis de los sonidos que emiten (habla, gritos, canciones, etc.), para generar unos elementos gráficos que se proyectan sobre una pantalla (Levin, 2006). Otro referente de estos mismos autores es la pieza *Hidden Worlds of Noise and Voice* (2002), introducida anteriormente en el apartado 1.4.1.

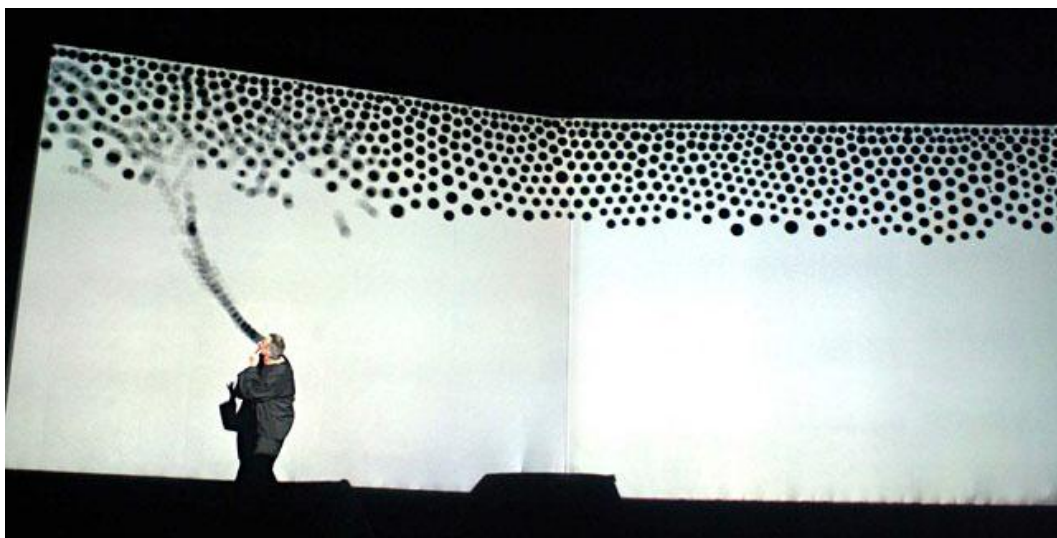


Figura 264. *Messa di Voice*. En (Levin, 2006).

12.2. DESCRIPCIÓN TÉCNICA

12.2.1. COMPONENTES DE LA APLICACIÓN

12.2.1.1. DISPOSITIVOS FÍSICOS

Los elementos que conforman la aplicación de *AR-Jazz* son (Figura 265):

- Un PC de mesa estándar.
- El software Max/MSP (versión 4.5.5) con el módulo de Jitter (versión 1.5.1).
- Un sensor inercial. Se ha utilizado el MT9 de Xsens.
- Una cámara Web. Se ha utilizado la cámara *Live Cam! Voice* de Creative, con un FOV de 85° y una resolución de 640x480.
- Una pantalla que funciona a modo de espejo. Las dimensiones de la pantalla del Auditori de Torrent son de 15x11 metros, aproximadamente.
- Un proyector multimedia.
- Un micrófono.
- Instrumentistas.

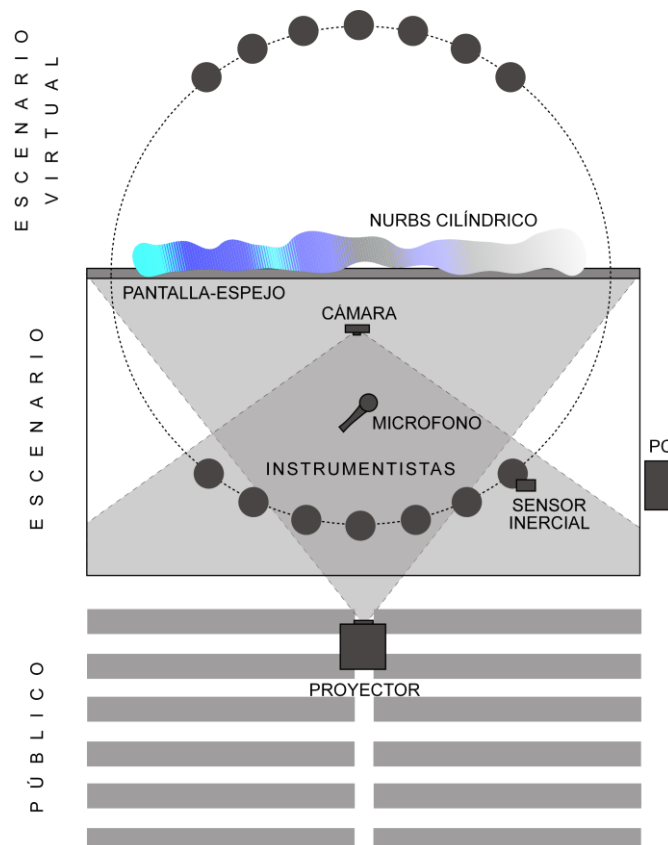
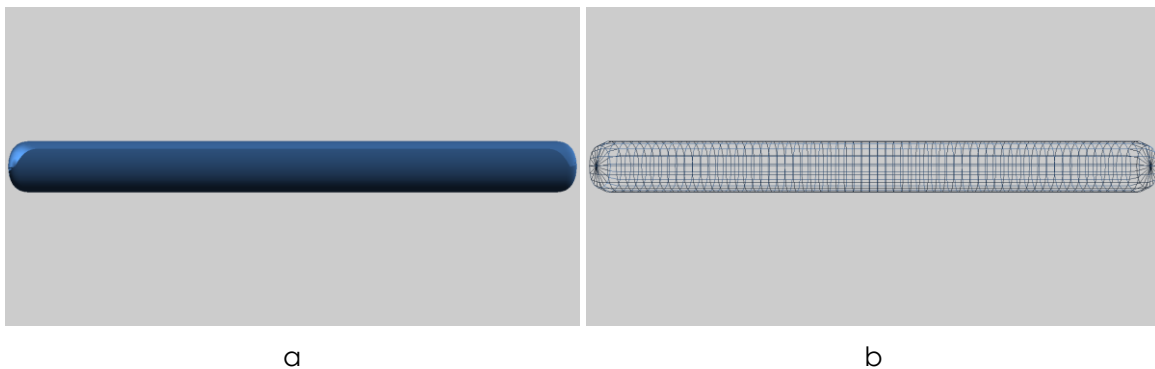


Figura 265. Esquema de los componentes y distribución espacial de *AR-Jazz*.

La cámara permanece fija en el espacio y apunta hacia los instrumentistas, de tal forma que la pantalla siga la metáfora del espejo mágico. El micrófono se sitúa entre la cámara y los instrumentistas. La distribución espacial de los instrumentistas (o músicos) en el escenario es tal que, junto con la representación de su imagen vista en la pantalla, forman un círculo en medio del cual se encuentra la representación visual (NURBS cilíndrico; ver apartado 12.2.1.2) del sonido que emiten en conjunto. Los músicos dan la espalda al público, con la finalidad de visualizar en todo momento el elemento virtual. El público puede ver el rostro de los músicos a partir de la imagen representada en la pantalla a modo de espejo; una vez más se utiliza la metáfora del espejo mágico para elaborar el entorno de realidad aumentada.

12.2.1.2. ELEMENTOS VIRTUALES

Para la visualización del sonido se ha utilizado un elemento virtual que consiste en un NURBS (Non Uniform Rational B-Splines) cilíndrico moldeable a partir del objeto *js jitaudio2nurbs.js* (ver apartado 12.2.2), cuya forma, color, luminosidad y texturizado varía de acuerdo a diversas características sonoras. De este modo, el elemento virtual variará dependiendo del timbre de cada instrumento, las notas que emita (altura) y la intensidad (amplitud). También se han establecido otros parámetros que varían de forma aleatoria o siguiendo un ritmo temporal (ver apartado 12.2.2), así como un cambio en la orientación a partir de las rotaciones del sensor inercial. El NURBS tiene la posición y un ángulo fijos; solamente varían dos de sus rotaciones, las cuales se registran a partir del sensor inercial, que lo puede llevar consigo uno de los instrumentistas, un bailarín o un usuario cualquiera.



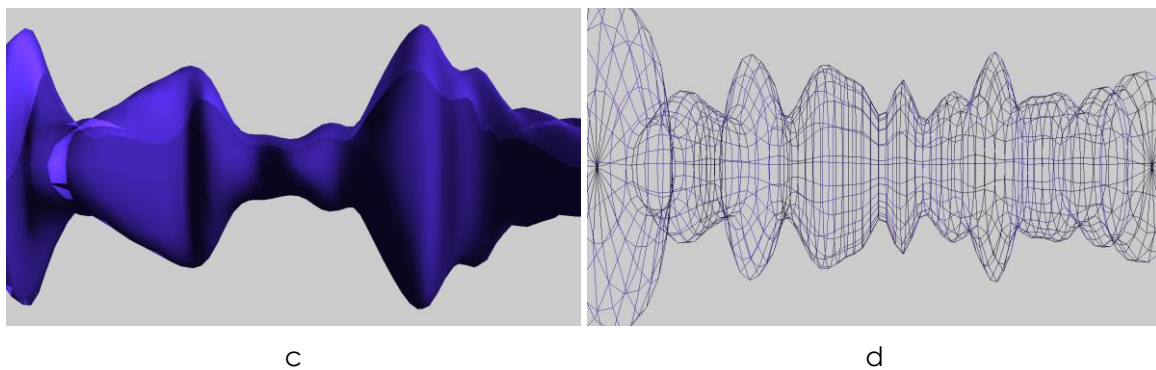


Figura 266. NURBS cilíndrico a partir de *js jitaudio2nurbs.js*, donde: a) Con textura; b) Representación alámbrica; c) Con textura y deformaciones a partir de un audio; d) Representación alámbrica y deformaciones a partir de un audio.

12.2.2. DESARROLLO, PROCESO

Para la elaboración de *AR-Jazz* se ha utilizado el programa *Max/MSP Jitter* (Cycling '74, 2008), en el que se ha desarrollado una aplicación propia de realidad aumentada mediante la combinación de varios objetos. La interacción persona-ordenador se realiza a dos niveles: sonoro y háptico. En el primer caso, la interfaz es un micrófono que recoge los sonidos emitidos por los músicos; en el segundo caso, la interfaz es un sensor inercial que lo puede llevar uno de los instrumentistas u otra persona que desee participar en la visualización del NURBS. A continuación mostramos el patch principal de programación y los sub-patches, explicando cada uno de ellos con detalle.

El patch principal (Figura 267) está basado en *jit.gl.nurbs-audiorender.pat* que se distribuye con el módulo de *Jitter*. El objeto principal es el *js jitaudio2nurbs.js*, que define la forma del NURBS cilíndrico, cuya superficie se deforma de acuerdo al audio mediante programación javascript. La entrada del micrófono se abre mediante *startwindow* o mediante el objeto *IO*. La ventana de OpenGL queda definida con el objeto *jit.window*. Las dimensiones de la superficie cilíndrica varía de 1 a 4 cada 100 segundos (mediante el objeto *line*), mostrándose en el primer caso con formas angulosas, que se van suavizando a medida que aumentamos el valor numérico. También se han establecido unos tiempos para definir la forma en que se dibuja la superficie (mediante el objeto *metro*), con lo que una parte del cilindro cambia de texturizado a alámbrico cada 40 segundos, y la otra parte lo hace cada 70 segundos, coincidiendo en el mismo durante ciertos intervalos de tiempo. Todos los elementos que definen al cilindro se envían al objeto *jit.gl.nurbs*.

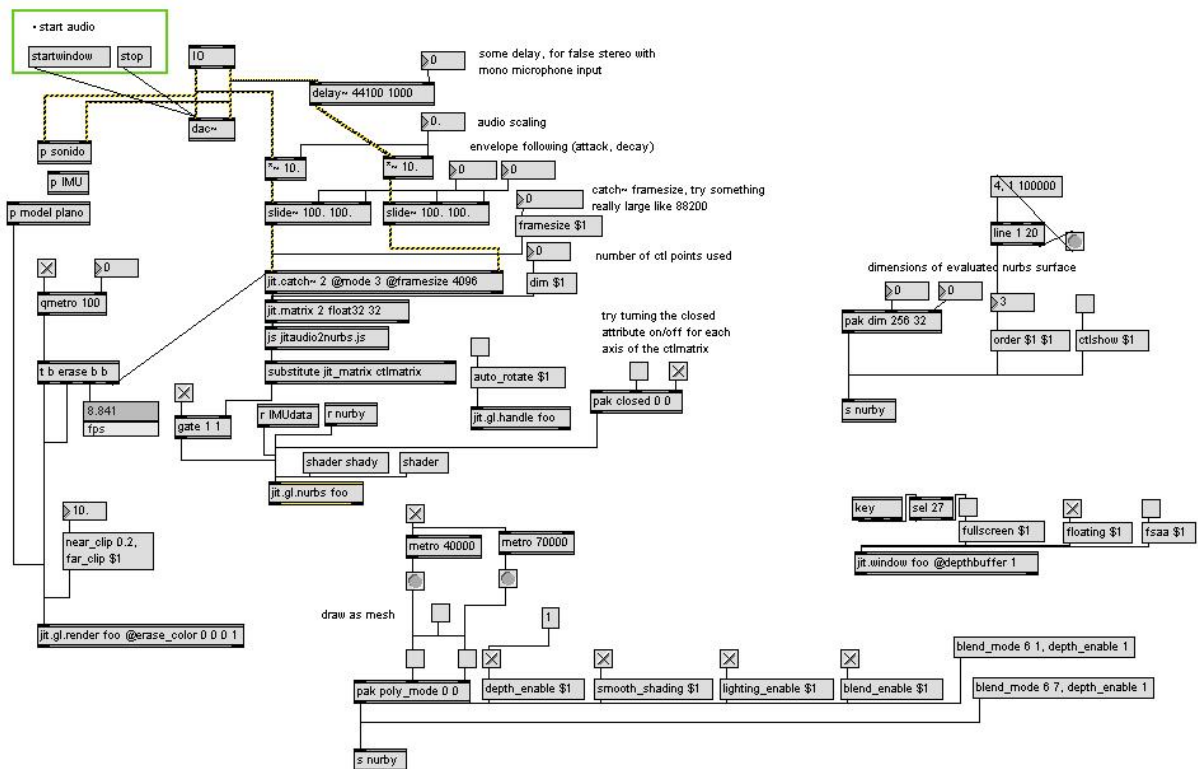


Figura 267. Patch principal de AR-Jazz.

p sonido (Figura 268): en este sub-patch se hace un análisis del sonido que recoge el micrófono mediante el objeto *fiddle~*, obteniendo la altura y amplitud del sonido. Debido a que estos valores cambian muy rápidamente, se calcula la media de los últimos 10 o 20 valores, mediante *zl group* y *mean*. A continuación estos valores se escalan mediante división por 100 y son asignados al valor del rojo y a la luminosidad del cilindro virtual. La ganancia de la entrada del micro está ajustada de tal forma que los valores de la amplitud registrada varían entre 70 y 90 (dB), por lo que los valores del rojo serán bastante altos (entre 0.7 y 0.9), potenciándose así los colores cálidos en el cilindro. El valor de luminosidad tiene establecido un rango, de tal forma que puede variar entre 0.3 y 1, ya que valores más próximos al 0 hacen que el cilindro se visualice como una mancha negra, y valores por encima del 1 lo emblaquece demasiado. Los valores del verde y del azul son determinados aleatoriamente por el subpatch *p mood machine*.

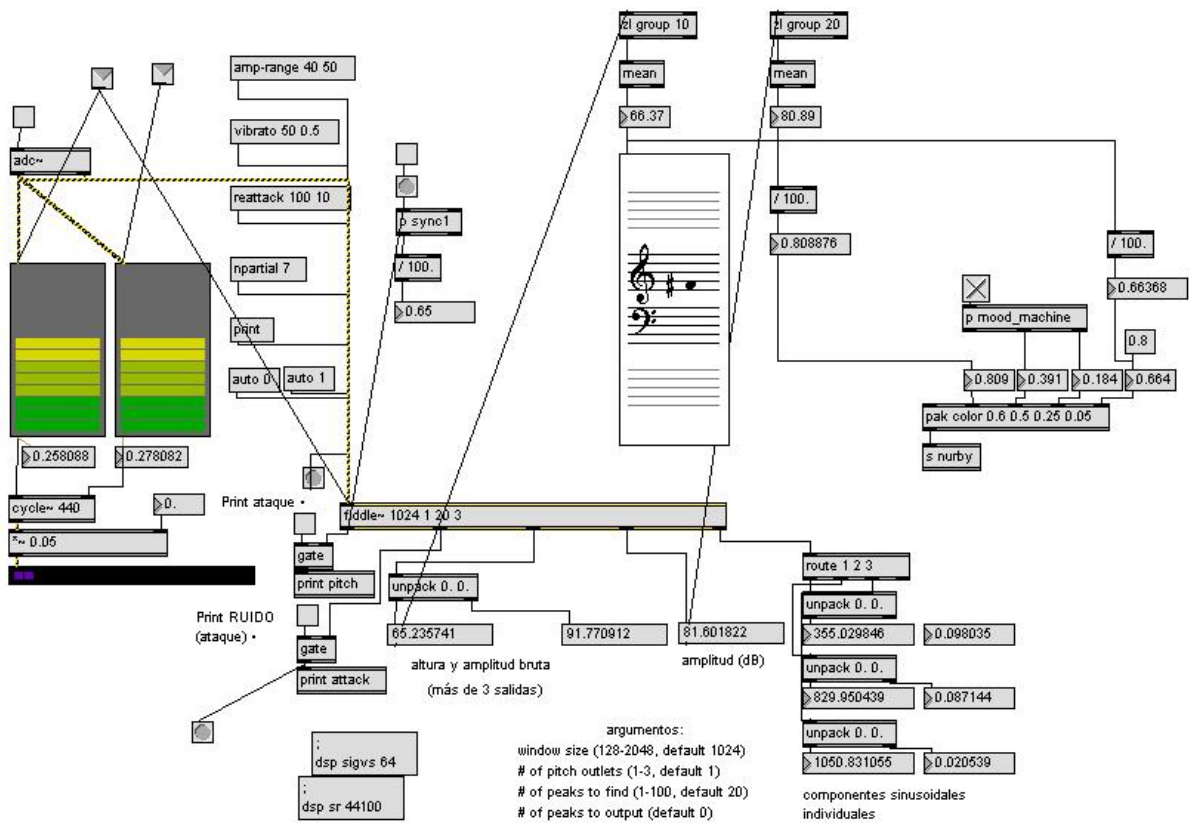


Figura 268. Sub-patch *p sonido*.

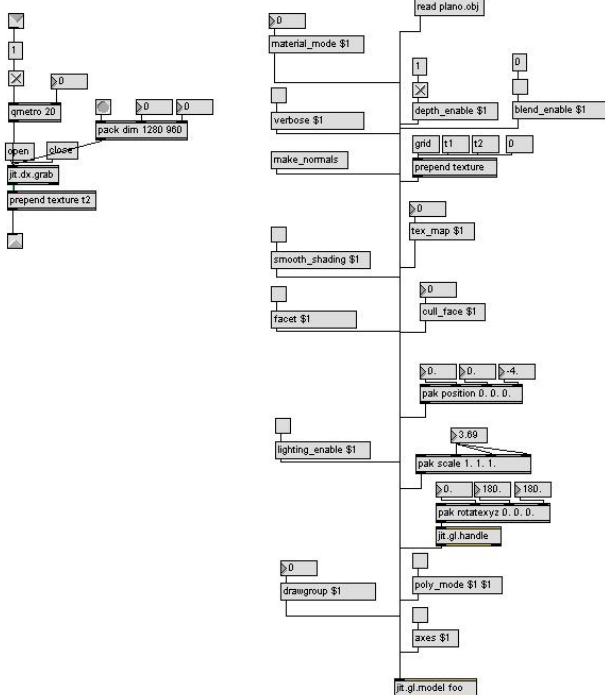


Figura 269. Sub-patch *p model plano*.

p model plano (Figura 269): Carga un plano en formato *obj* sobre el que se mapea la imagen capturada en tiempo real por la cámara mediante el objeto *prepend texture*. Este plano se envía a la ventana de OpenGL (con el nombre de *foo*, en este caso) mediante el objeto *jit.gl.model*. La posición del plano dentro de la ventana de OpenGL queda establecida con el objeto *pak position*, la orientación mediante *pak rotatexyz*, y la escala mediante *pak scale*.

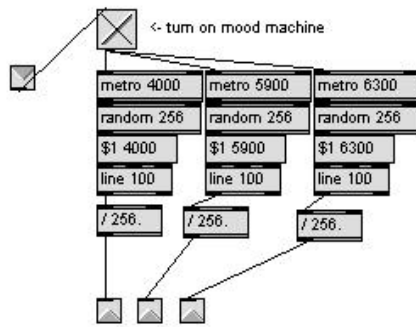


Figura 270. Sub-patch *p mood_machine*.

p mood_machine (Figura 270): En este patch se eligen valores entre 0 y 256 de forma aleatoria mediante el objeto *random*. Estos valores se envían a continuación al sub-patch *p sonido*, para asignarlos a los parámetros del verde y del azul del cilindro virtual, dentro de la gama RGB.

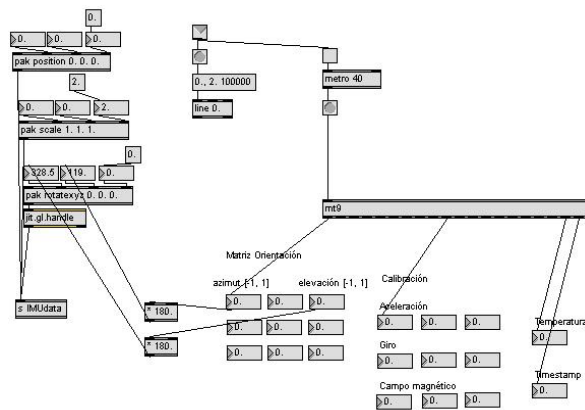


Figura 271. Sub-patch *p IMU*.

p IMU (Figura 271): Recoge los valores registrados por el sensor inercial mediante el objeto *mt9* y asigna dos de sus rotaciones a las rotaciones en X e Y del cilindro virtual, mediante el objeto *pak rotatexyz*. La posición y escalado del cilindro también se determinan en este patch mediante los objetos *pak position* y *pak scale*, respectivamente.

12.2.3. CONSIDERACIONES / RESTRICCIONES TÉCNICAS

Es importante mencionar que algunas partes de este ensayo son en cierta manera improvisadas, ya que las características del lugar y las necesidades de los músicos condicionaron algunos aspectos del ensayo. En cuanto al escenario, sus dimensiones eran excesivas para el número de músicos que formaron el combo integrado por los asistentes al taller. La utilización de dispositivos que necesariamente van conectados al ordenador a través de cables (sensor inercial y cámara Web), obligó a que la agrupación tuviese que situarse cerca del PC, por lo que los músicos no pudieron ubicarse completamente en el centro del escenario. Finalmente, las condiciones lumínicas de la sala condicionaron algunos parámetros de la aplicación, viéndose la imagen capturada por la cámara demasiado oscura durante la representación en vivo, por lo que el efecto de la pantalla a modo de espejo era poco luminoso.

12.3. SISTEMAS DE VERIFICACIÓN

12.3.1. TALLER

El taller se realizó en el escenario del auditorio durante un par de horas a media mañana y otro par de horas por la tarde, al cual asistieron un total de nueve músicos: un batería, una flauta, una guitarra eléctrica, un piano de cola, un saxo alto, dos trombones, una tuba y una cantante. Tras montar la pieza en el escenario, se explicó brevemente el funcionamiento de la aplicación creada con Max/MSP Jitter y el modo en que se producía la interacción con NURBS generado. A continuación los músicos fueron probando uno a uno su instrumento, para ver el resultado de su interpretación, tras lo que se ajustaron algunos de los parámetros de la aplicación mediante un feedback directo con los ellos. Finalmente, los músicos que participaban en el taller realizaron un ensayo en conjunto de diversas piezas musicales, visualizando en todo momento el objeto virtual fruto de su interpretación. En la siguiente figura se muestran algunas imágenes del transcurso del taller.



Figura 272. Imágenes tomadas durante el taller de AR-Jazz.

12.3.2. REPRESENTACIÓN EN VIVO

La representación en vivo tuvo lugar a las 20:00 en el auditorio con la asistencia de diverso público. En primer lugar se explicó a la audiencia en que consiste la realidad aumentada y el funcionamiento de la pieza *AR-Jazz*. Para que entendieran la relación directa entre sonido emitido por los instrumentistas y la visualización gráfica en la pantalla, cada uno de los músicos hizo un solo con su instrumento, mostrando la variación entre notas graves y agudas, dinámicas, articulaciones, y recursos tímbricos. Para que la explicación resultase más clara, en esta fase se mostró el NURBS sobre un fondo grisáceo y se restringieron algunos parámetros, como las rotaciones. En las siguientes figuras podemos ver algunos ejemplos de la forma que adquiere el NURBS cilíndrico dependiendo del instrumento y la nota emitida (de acuerdo al índice acústico franco-belga). Sin embargo, no resulta fácil establecer una relación directa entre las distintas notas musicales, especialmente porque los sonidos son emitidos en directo, produciéndose ruidos en la señal recogida.

BATERÍA



caja clara



plato hit hat



plato suspendido



redoble tom tom



tom tom grave



tom tom medio

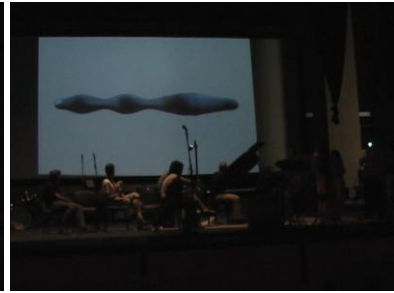
FLAUTA



do 3



fa 4



fa 5



re 4



si_b 3



si_b 4

GUITARRA ELÉCTRICA



fa 2



fa 4



si mayor séptima



re 1



re 2

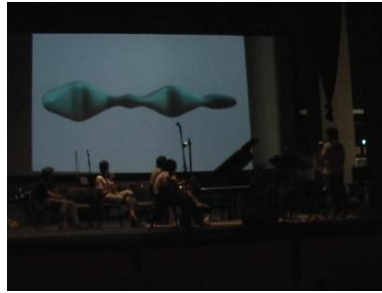


mi_b mayor sustained

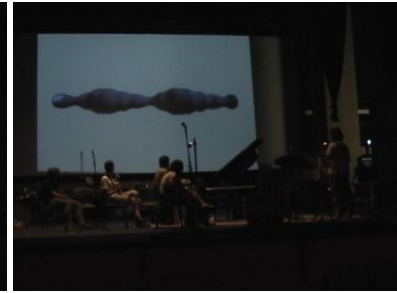
SAXO ALTO



fa 3



fa 4



do 4



si, 1



la, 2



la, 3

TROMBÓN



do 4



re 3



mi 2



la, 1

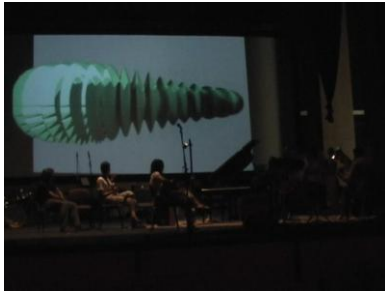


la, 2



la, 3

TUBA



fa 2



fa 2 (forte)



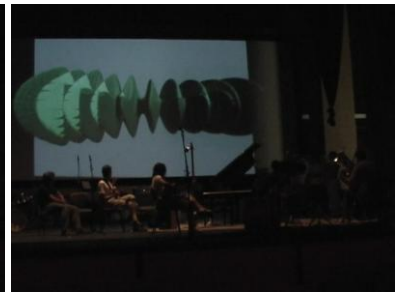
fa# 2



do 2



sol 1



si b 1

VOZ



re 3



mi 3



fa# 3



sol 3



la 3



si 3

Figura 273. Visualizaciones del sonido de acuerdo al instrumento utilizado y a la nota emitida.

Una vez terminada la explicación, se procedió a la representación de diversas piezas de Jazz, en una *Jam session*. Respecto a las rotaciones de la visualización del sonido, finalmente se decidió que las controlara un voluntario del público, mujer en este caso,

la cual subió al escenario y se le apuntaron algunas indicaciones para el correcto manejo del sensor, permaneciendo al lado de los músicos durante toda la representación. Las piezas representadas fueron dos temas que se habían ensayado en el taller de la mañana; la primera de ellas fue una improvisación sobre la escala frigia (o escala española); la segunda fue una improvisación a partir de un ritmo de blues. En la siguiente figura podemos ver una serie de imágenes del evento.

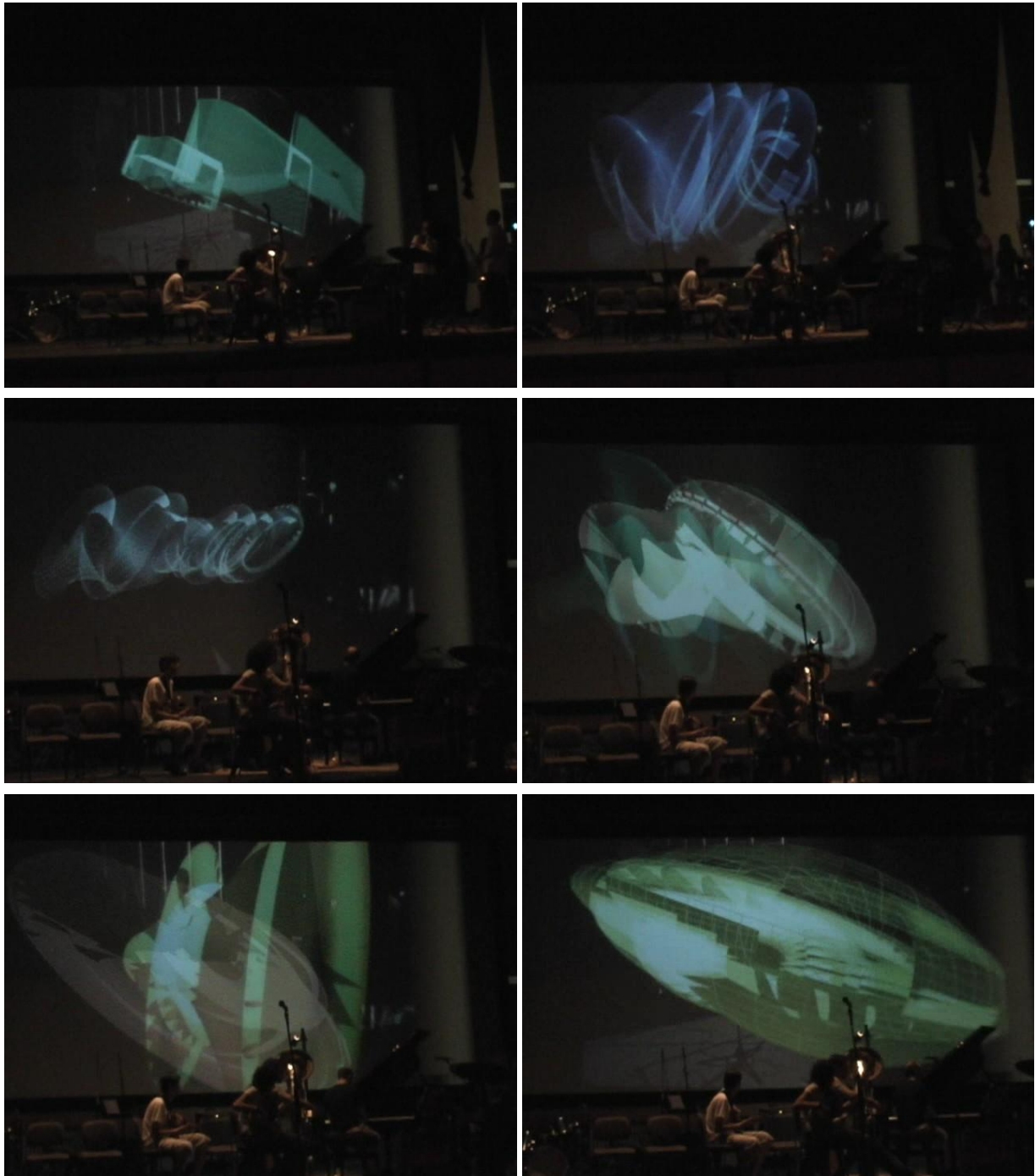


Figura 274. Imágenes tomadas durante la *Jam session*.

En la representación se ha observado la dificultad que experimentan los músicos a situarse en el escenario dando la espalda al público, puesto que están acostumbrados a dirigirse siempre hacia ellos para poder proyectar el sonido de su instrumento hacia la audiencia. Es por ello que algunos prefirieron sentarse de medio lado, para poder ver simultáneamente al público y a la pantalla.

CONCLUSIONES

Para exponer las conclusiones se va a hacer un estudio transversal entre los capítulos expuestos de teoría y práctica, discutiendo algunos aspectos que se consideran más relevantes. También se aporta una visión personal tanto de las carencias como del potencial que ofrece la tecnología de RA.

La realidad aumentada como tecnología emergente en estos momentos, tiene un **gran potencial en campos del conocimiento muy variados** (arte, entretenimiento, educación, medicina, psicología, arquitectura, etc.), y distintos autores han creado sistemas que así lo evidencian. En esta tesis, esto se ha mostrado a través de la referencia a diversas aplicaciones, desarrollos y estudios realizados mediante esta tecnología. A partir de estos trabajos, se puede vislumbrar como actualmente la tecnología de RA avanza hacia sistemas colaborativos, multimodales, altamente interactivos, y con la integración de dispositivos híbridos que realizan un registro 3D en tiempo real. Los ordenadores, cada vez más potentes, permiten que los objetos virtuales sean más fotorrealistas (por ejemplo, en OSGART se integran las librerías de OpenSceneGraph en ARToolkit). Todos estos dispositivos son también cada vez más ligeros y ergonómicos, por lo que en las aplicaciones móviles de RA, el ordenador portátil se sustituye por PDAs o teléfonos móviles con cámaras integradas y/o GPS, con capacidad suficiente para soportar sistemas de RA. Las aplicaciones no se limitan a dispositivos de visualización tipo HMD, sino que surge una gran variedad de displays, incluyendo proyecciones 3D.

En los últimos años existe un interés creciente, por parte de la comunidad investigadora, en estudiar la reacción y el comportamiento del usuario en aplicaciones de RA. Como se ha visto en la introducción de esta tesis, el usuario está situado en lo alto de la pirámide de la tecnología de RA, por lo que surge la necesidad de estudiar los aspectos relacionados con el componente humano. A partir de los ensayos realizados en esta tesis, se pueden derivar algunas conclusiones referentes al **usuario final**:

1. La utilización de marcas de patrones planos permite una **interacción tangible**, ya que los objetos virtuales, al estar relacionados con marcas físicas, pueden ser literalmente cogidos con las manos para manipularlos de una forma natural (se pueden girar, mover, acercar o alejar), emulando incluso la manera en que sus correspondientes objetos reales son utilizados;
2. El diseño de **interfaces sencillas** hace más natural la comunicación persona-computador, permitiendo incluso que personas no familiarizadas con los ordenadores puedan interactuar con el sistema;
3. Los displays gráficos basados en proyecciones facilitan una **visualización compartida**, posibilitando la participación colaborativa con el uso de un único dispositivo (con lo que se reducen los costes) al tiempo que permiten una escalabilidad de los datos mostrados;
4. De otro modo, los dispositivos HMD permiten la distribución **individual** de información, ya que la escena aumentada sólo es visible para aquel usuario que lleve puesto el HMD y, en caso de haber varios dispositivos, éstos pueden mostrar información distinta a cada usuario;
5. Gracias a la **metáfora del espejo mágico** utilizada en algunos de los ensayos, los elementos virtuales se pueden visualizar conjuntamente con la imagen de los usuarios, proporcionando una mayor sensación de presencia. Además, la inversión del sistema tradicional de RA posibilita el hecho de que el único dispositivo que el usuario debe llevar consigo son las marcas, evitando en gran medida el impacto que ciertos dispositivos típicos de RA puedan ocasionar sobre algunos usuarios (por ejemplo, la incomodidad de llevar sujeto a la cabeza un HMD unido a los mareos que éste pueda ocasionar) y haciéndolo más universal (ya que también es adecuado para niños);
6. El **juego** es un recurso que se ha utilizado en algunos de los ensayos con la finalidad de captar la atención del usuario, incrementando de igual modo la sensación de presencia. En el caso de *LLH*, el juego colaborativo con contacto

visual directo permite una mayor complicidad entre los usuarios, haciendo menos evidente la separación entre tecnología y el juego más tradicional.

Se puede afirmar que **la realidad aumentada es un nuevo “sintetizador” que está impulsando con fuerza nuevas relaciones entre Arte y Ciencia**. Este hecho se ha demostrado a través de los ensayos de esta tesis en los que también se ha dado respuesta a través de esta tecnología a distintas cuestiones conceptuales: en *Usuario Aumentado* se cuestiona la identidad de los usuarios a través de la metáfora del espejo mágico, transformando la percepción visual que los usuarios tienen de su cuerpo; en *UrbanMix* se confrontan visualmente distintos espacios urbanos mezclando pasado y presente; en *AR-Jazz* se realiza la visualización del sonido y movimiento de los intérpretes en una *Jam session*, para introducirlo visualmente en el centro de un círculo formado por éstos y su representación gráfica; etc. Todo ello se ha logrado gracias a la sinergia entre el mundo real y el virtual en tiempo real, por lo que se deduce que todas estas cuestiones son difícilmente abordables mediante otras tecnologías.

Sin embargo, en la realización de dichos ensayos se ha visto como, en muchos aspectos, la tecnología que existía (o la tecnología a la que se podía acceder) tenía diversas carencias o no ofrecía suficientes recursos para su resolución. Es por ello que ha sido necesario integrar distintas metodologías (como en *UrbanMix* o *Hot Images*), haciendo que el trabajo resultante tenga un fuerte componente técnico. Ciertamente, en muchas ocasiones se ha tenido el presentimiento de que la tecnología que se estaba empleando estaba en un “estado inicial de desarrollo” por lo que, de no haber incidido en la parte técnica, muchas de las cuestiones planteadas no se hubiesen podido abordar mediante la RA.

Como se ha visto en el capítulo 1 de esta tesis, la realidad aumentada es un paradigma que surge a partir de la realidad virtual, aportando un nuevo potencial: mientras que la realidad virtual nos aísla del mundo en que vivimos, la realidad aumentada lo enriquece, añadiendo, alterando o incluso, eliminando información. Sin embargo, hay que ser conscientes de que actualmente, esta tecnología tiene **grandes carencias en cuanto a desarrollo técnico**, que dificultan en gran medida la realización de aplicaciones atractivas y estables. Comparando la tecnología de RA con la de RV, estas carencias se hacen aún más evidentes. No hay más que ver la calidad de los gráficos generados por ordenador que se consiguen con técnicas de realidad virtual – incluso, en tiempo real – y su gran difusión y consumo por un público muy variado. En este sentido, la RA es una tecnología débil y una gran desconocida.

Como se ha argumentado a lo largo de esta tesis, la principal ventaja de la RA frente a la RV – la mezcla de un mundo real y uno virtual que guardan una relación espacial – constituye a la vez el principal inconveniente para ser abordado por una tecnología robusta. Esto es, la **dificultad de proporcionar un registro tridimensional en tiempo real** que sea suficientemente preciso y fácilmente gestionado por el hardware para, al mismo tiempo, generar una escena aumentada que sea (como mínimo) **geoméricamente consistente**. Por ello es necesario conocer los elementos que intervienen en este proceso – sensores, métodos de registro y software – para utilizar, en cada caso concreto, aquel o aquellos que por sus características resuelvan mejor la finalidad buscada. Por este motivo, en los ensayos realizados se han utilizado distintos sistemas de registro, abarcando métodos directos, indirectos y mixtos; así como distintos dispositivos y software. En general, se ha evidenciado la dificultad de integración de todos estos elementos en sistemas robustos de RA, condicionando en muchos casos los resultados obtenidos: para resolver una cuestión determinada se han buscado los elementos más idóneos, pero estos elementos a su vez, han condicionado muchos aspectos de la escena aumentada.

Por ejemplo, la **falta de precisión aportada por receptores GPS de bajo coste** (del orden de pocos metros) no se puede resolver técnicamente, puesto que es una característica propia de este tipo de sensores, que además, también trae consigo el problema del trayecto múltiple. Es por ello que debemos ser conscientes de estas limitaciones y utilizar estos receptores sólo en aquellas aplicaciones de RA en que éstas no constituyan un inconveniente: por ejemplo, en aplicaciones al aire libre en las que el mundo generado por ordenador sea principalmente virtual, como el mostrado en el ensayo *Hot Images*, que constituye una aplicación de virtualidad aumentada en la que el mundo real no se registra de forma precisa al virtual, sino que simplemente se aporta una localización aproximada del usuario dentro de éste.

De otro modo, en el ensayo *UrbanMix* se ha utilizado un **método mixto** para la localización espacial (sensor inercial y técnicas de visión por computador). Aunque los inconvenientes encontrados en este caso han sido múltiples (como la alteración del entorno físico, dependencia fuerte de condiciones de luminosidad o escala local de actuación), se puede afirmar que la consideración de un método mixto puede aportar considerables ventajas si se potencian las cualidades positivas de cada uno de los elementos que intervienen. Para mejorar el método utilizado en *UrbanMix*, debería modificarse el registro de color para establecer otro, a partir de elementos naturales del

entorno y considerando un mayor número de entidades; también se podría utilizar una cámara de IR para reducir la dependencia con las condiciones de luminosidad.

La **integración de sensores** en aplicaciones de RA generalmente se suele hacer con **programas multitarea**, como es el caso de Max/MSP Jitter. Este programa dispone de objetos que posibilitan de manera sencilla la integración de distintos periféricos, tanto para el registro 3D en tiempo real como para posibilitar una interacción más variada. Además, la inclusión de datos multimedia se realiza de forma sencilla. Sin embargo, el principal inconveniente que presenta este software es que no está diseñado específicamente para aplicaciones de RA, por lo que su adaptación requiere un conocimiento de las relaciones matemáticas entre la geometría del mundo físico y la imagen capturada por la cámara, así como de la concordancia entre el mundo real y el virtual (explicadas en el capítulo 2). En este sentido, la realización de aplicaciones de RA sólo es posible a partir del establecimiento de una serie de sistemas de ecuaciones que, en algunos casos, pueden resultar complejas.

La solución más extendida para la realización de aplicaciones de RA es la utilización de **software basado en el reconocimiento de patrones planos**, tales como ARToolKit o MXRToolKit, que están diseñados específicamente para la generación de escenas aumentadas. Entre sus ventajas cabe destacar: la precisión alcanzada con el registro 3D (del orden del cm); la peculiaridad de que los objetos generados por ordenador se pueden relacionar fácilmente con las marcas (permitiendo interfaces tangibles); la posibilidad de implementar elementos de interacción (a partir de distancias y rotaciones entre marcas); y la economía de los sistemas (ya que suelen disponer de licencia GPL y para el registro únicamente es necesaria una cámara Web). Sin embargo, también hay que tener en cuenta sus limitaciones, por ejemplo: la fuerte dependencia de las condiciones de luminosidad; la escala local de trabajo (distancia cámara-marca); y el hecho de que el entorno físico se ve alterado por las marcas. Además, un gran inconveniente que se ha encontrado en la **utilización de software libre** es la falta de garantía de que éste funcione de forma correcta, la ausencia en muchos casos de un manual detallado o de asistencia técnica ante cualquier problema. Si a estos factores le añadimos la necesidad de programación en un lenguaje de alto nivel, como es el C++, hacen que su utilización masiva no sea posible, estando generalmente reducida a desarrolladores e investigadores.

Estableciendo una **comparativa entre los software** de este tipo, empleados en los ensayos de esta tesis (ARToolKit y MXRToolKit) y otros que, aunque no se han utilizado

por diversos motivos, sí se han analizado (ARTag y AMIRE), se pueden establecer las siguientes conclusiones: los más estables son ARToolKit y ARTag, mientras que los que más recursos consumen son MXRToolKit y AMIRE, haciendo que en algunas ocasiones el programa finalice de forma inesperada; en cuanto a la integración de elementos multimedia, tanto MXRToolKit como AMIRE proporcionan mayores posibilidades (distintos formatos de objetos 3D, vídeos, imágenes, sonido, etc.); el registro más robusto lo proporciona ARTag, ya que tiene una menor dependencia con las condiciones de luminosidad, las marcas se pueden ocultar de manera parcial sin interrumpir el registro, y el sistema es capaz de identificar un total de 2002 marcas distintas; sin embargo, el inconveniente de ARTag es que no posee licencia GPL; en cuanto a programación, en ARToolKit y ARTag se requiere un nivel de experto y en MXRToolKit un nivel medio, mientras que en AMIRE no son necesarios conocimientos de programación.

Otro software similar que se ha utilizado es BazAR que realiza el **registro a partir de la identificación de rasgos naturales puntuales**. A diferencia de los anteriores, mediante la utilización de este software no se altera el entorno físico, al no requerir marcas artificiales, por lo que puede ser idóneo para ciertas aplicaciones, como ha sido el caso de *CatedrAR*. Sin embargo hay que tener en cuenta la gran cantidad de recursos que consume – necesita un ordenador más potente que en los casos anteriores –, la consideración de una única imagen de referencia (en contraposición a los anteriores, que pueden identificar varias marcas) y la poca variedad en cuando a elementos multimedia (únicamente OpenGL). El nivel de programación requerido es también de experto.

Estos conceptos referentes a software se resumen en la Tabla 9 a modo comparativo. Se resaltan aquellos aspectos que se consideran ventajosos frente a los que se consideran inconvenientes. De otro modo, en la Tabla 10 se resumen de igual forma los métodos de registro empleados en los ensayos.

Software	Ventajas (+)	Inconvenientes (-)
ARToolKit	Muy estable. Software libre.	Nivel de programación: experto. Datos multimedia: OpenGL y VRML. Marcas visibles para registro. Fuerte dependencia de la luz ambiental. Proceso semiautomático de calibración de la cámara. Registro basado en el coeficiente de

		correlación no permite el uso de muchas marcas (del orden de 10-20)
MXRToolKit	Datos multimedia: OpenGL, VRML (y otros formatos 3D), sonido, imagen, vídeos. Nivel de programación: medio. Software libre.	Poco estable. Marcas visibles para registro. Su desarrollo no se sigue actualmente. Fuerte dependencia de la luz ambiental. Proceso semiautomático de calibración de la cámara. Registro basado en el coeficiente de correlación no permite el uso de muchas marcas (del orden de 10-20)
ARTag	Datos multimedia: OpenGL y otros formatos 3D. Marcas parcialmente ocultas permiten registro. Menor dependencia de la luz ambiental. Registro robusto permite consideración de muchas marcas (un total de 2002).	Nivel de programación: experto. No es software libre. Proceso semiautomático de calibración de la cámara.
AMIRE	Datos multimedia: OpenGL, 3DS, sonido, imagen, vídeos. Nivel de programación: no necesario. Software libre.	Consume muchos recursos de la CPU. Su desarrollo no se sigue actualmente. Proceso semiautomático de calibración de la cámara.
BazAR	Imagen parcialmente oculta permite registro. Software libre. Menor dependencia de la luz ambiental. Proceso automático de calibración de la cámara. Proceso automático de calibración fotométrica e iluminación de los objetos virtuales.	Nivel de programación: experto. Datos multimedia: OpenGL. Una sola imagen como referencia.
Max/MSP Jitter	Nivel de programación: no necesario. Datos multimedia: OpenGL, VRML (y otros formatos 3D), sonido, imagen, vídeos. Fácil integración de periféricos	No es software libre. No es software específico de RA. Registro por técnicas de visión por computador insuficiente (alto consumo de recursos).

	permitiendo mayor interacción.	
--	--------------------------------	--

Tabla 9. Comparación entre los distintos programas analizados en esta tesis. Ventajas e inconvenientes encontrados.

Método	Ventajas (+)	Inconvenientes (-)
Visión por computador: ARToolKit (capítulo 7). MXRToolKit (capítulo 10).	Bajo coste. Ligero (marcas). Muy preciso (del orden del mm/cm).	Entorno físico alterado (marcas). Dependiente de la luz. Procesamiento lento (requiere de una buena tarjeta gráfica, etc.). A escala local (distancia cámara-marca).
Visión por computador: BazAR (capítulo 11).	Entorno físico no alterado. Bajo coste. Muy preciso (del orden del mm/cm).	Dependiente de la luz. Procesamiento lento (requiere de una buena tarjeta gráfica, etc.). A escala local (distancia cámara-marca). Más pesado (ordenador portátil).
Sensores: GPS + INS + Max/MSP Jitter (capítulo 9). INS + Max/MSP Jitter (capítulo 12).	Entorno físico no alterado. No dependiente de la luz (funciona de noche, en la oscuridad). Procesamiento rápido. A escala global.	Mayores costes (sensores + sw). Más pesados (ordenador portátil + GPS + INS + baterías). Poco preciso (del orden de pocos metros en el GPS y de 1° en el INS).
Sensores + visión por computador: Max/MSP Jitter + INS (capítulo 8).	Precisión media (del orden del cm).	Entorno físico alterado (elementos circulares de colores). Mayores costes (sensores + sw). Más pesados (ordenador portátil + INS + baterías). Dependiente de la luz. Procesamiento lento (debido al registro de color). A escala local (los elementos circulares deben ser visibles).

Tabla 10. Comparación entre los distintos métodos de registro utilizados en los ensayos de esta tesis. Ventajas e inconvenientes encontrados.

Cabe mencionar que, durante el periodo de realización de esta tesis, se ha experimentado una **rápida evolución de los programas basados en el registro de patrones planos**, con la consiguiente necesidad de actualizar continuamente las librerías utilizadas. Por ejemplo, en el caso de ARToolKit la primera versión con la que se ha trabajado es la 2.65, mientras que los últimos ensayos se llevaron a cabo con la versión 2.73; en el primer caso las librerías se compilaban con Microsoft Visual C++ 6.0, mientras que en el segundo caso con Microsoft Visual C++ .NET. Además, actualmente se ha desarrollado OSGART, que incluye las librerías OpenSceneGraph en el programa ARToolKit. También se han experimentado **incompatibilidades entre algunos programas y distintos dispositivos**. Por ejemplo, el plug-in del sensor inercial MT9 está programado para la versión Max/MSP 4.5, siendo incompatible con la versión actual. Además, este plug-in únicamente es compatible para PC. El **rápido desarrollo de software y hardware** crean una incertidumbre respecto a los próximos años: ¿se podrán ejecutar estas aplicaciones en la próxima generación de ordenadores?

Líneas futuras

A partir de la experiencia adquirida en estos ensayos, sería importante realizar nuevas investigaciones y desarrollar algunas de las cuestiones más relevantes del panorama investigador actual. Se pueden citar varias líneas de actuación:

1. **Mejorar el registro en tiempo real.** Siendo uno de los mayores retos de esta tecnología, se considera relevante el estudio de la integración entre diversos sensores, potenciando las ventajas de cada uno de ellos e intentando ofrecer un registro ubicuo.
2. **Integración de sistemas multimodales.** Como se ha argumentado en el apartado 4.5, los sistemas multimodales están surgiendo como nuevas alternativas de RA que ofrecen mayores grados de inmersión. En los ensayos estudiados, las escenas aumentadas son básicamente visuales, aunque también se exploran en algunos casos los estímulos de tipo sonoro y háptico. Respecto a los sonoros, únicamente se utilizan en dos de ellos: en *LLH*, donde el sonido constituye una salida del ordenador; y en *AR-Jazz*, donde el sonido es una entrada para la generación del objeto virtual. En cuanto al háptico, únicamente se utiliza en *LLH*. Aunque la exploración de estos estímulos haya sido somera, se aprecia su gran potencial para la generación de escenas aumentadas que ofrezcan una experiencia más enriquecedora.

3. **Explorar otros dispositivos o metodologías de RA no abarcadas durante la tesis.**

Un aspecto que está creciendo en importancia es la realización de aplicaciones de RA a partir de pequeños dispositivos móviles, tales como PDAs o teléfonos móviles, principalmente en el campo del entretenimiento. En este sentido, se considera relevante la exploración de las librerías Studierstube (introducidas en el apartado 5.1.3) así como de otras plataformas que soporten la entrada y gestión de datos GPS. Otro aspecto que no se ha abordado en esta tesis son los displays basados en proyecciones sobre superficies no planas; la exploración de estas técnicas se considera relevante ya que estas proyecciones permiten romper los límites rectangulares impuestos por las pantallas y ofrecer una integración más versátil para la visualización con el entorno físico.

4. **Aspectos referentes al usuario final.** En algunos de los ensayos se han podido realizar pruebas de usuarios, con la finalidad de evaluar los sistemas elaborados y establecer líneas futuras de actuación, obteniendo en todos los casos unos resultados satisfactorios. Es por ello que en futuros desarrollos se prevé de igual forma realizar dichos análisis. En concreto, existen una serie de factores aportados por (Wang et al., 2006) que se deberían considerar: aspectos perceptivos (mezcla mental del mundo real y el virtual); atención; procesos de información humana; punto de vista; factores humanos en colaboración.

5. **Campos de aplicación.** Se considera importante el estudio de aspectos sociales que la tecnología de RA pueda aportar, principalmente en los campos de la educación y del entretenimiento.

BIBLIOGRAFÍA

- Aguiló, C., J. Lorés y E. Junyent (2001). Enhanced Cultural Heritage Environments by Augmented Reality Systems. International Conference on Virtual Systems and Multimedia (VSMM'01), IEEE Computer Society, Washington, DC, USA.
- AMIRE. (2004a). "AMIRE." Visto el 31 Enero, 2008, en <http://www.amire.net/index.html>.
- AMIRE (2004b). AMIRE Authoring Tutorial, Version 2004-04-16: 68.
- Anastassova, M., J.-M. Burkhardt, C. Megard y P. Ehanno (2005). "Results from a user-centred critical incidents study for guiding future implementation of augmented reality in automotive maintenance." International Journal of Industrial Ergonomics **35**(1): 67-77.
- Andersen, T. L., S. Kristensen, B. W. Nielsen y K. Grønbaek (2004). Designing an Augmented Reality Board Game with Children: The BattleBoard 3D Experience. Proceeding of the 2004 conference on Interaction design and children: building a community IDC '04, ACM Press, New York, NY, USA.
- Ando, H., T. Miki, M. Inami y T. Maeda (2002). The Nail-Mounted Tactile Display for the behavior modeling. Computer Graphics and Interactive Techniques San Antonio, Texas, ACM New York, NY, USA
- ARTag. (2007). "ARTag." Visto el 30 September, 2007, en <http://www.artag.net/>.
- Asai, K., Y. Y. Sugimoto, N. Hemmakorn, N. Osawa y K. Kondo (2004). Language-support system using character recognition. Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry VRCAI '04, ACM Press, New York, NY, USA.
- Ascension Technology Corporation. (2007a). "laser BIRD 2™." Visto el 5 Noviembre, 2007, en <http://www.ascension-tech.com/products/laserbird.php>.

- Ascension Technology Corporation. (2007b). "miniBIRD@ 500 & 800." Visto el 5 Noviembre, 2007, en <http://www.ascension-tech.com/products/minibird.php>.
- Ascension Technology Corporation. (2007c). "ReActor@2: DIGITAL ACTIVE-OPTICAL MOCAP SYSTEM." Visto el 5 Noviembre, 2007, en <http://www.ascension-tech.com/products/reactor.php>.
- Ascott, R. (1998). La arquitectura de la cibercepción. Ars Telemática. Telecomunicación, Internet y Ciberespacio. C. G. Ed. Barcelona, L'àngelot.
- Aschke, M., C. R. Wirtz, J. Raczowsky, H. Worn y S. Kunze (2003). "Stereoscopic augmented reality for operating microscopes." International Congress Series **1256**: 408-413.
- Auer, T. y A. Pinz (1999). "The integration of optical and magnetic tracking for multi-user augmented reality." Computers & Graphics **23**(6): 805-808.
- Augmented Environments Laboratory. (2007). "AR Scene Graph: OSGAR." Visto el 12 Noviembre, 2007, en <http://www.cc.gatech.edu/projects/acl/projects/ARSceneGraph.html>.
- Avery, B., B. H. Thomas, J. Velikovsky y W. Piekarski (2005). Outdoor augmented reality gaming on five dollars a day. Proceedings of the Sixth Australasian conference on User interface - Volume 40, Australian Computer Society, Inc. Darlinghurst, Australia, Australia.
- Azuma, R. (2004). Overview of Augmented Reality. SIGGRAPH'04.
- Azuma, R., J. Weon Lee, B. Jiang, J. Park, S. You y U. Neumann (1999). "Tracking in unprepared environments for augmented reality systems." Computers & Graphics **23**(6): 787-793.
- Azuma, R. T. (1997). "A Survey of Augmented Reality." Presence: Teleoperators and Virtual Environments **6**(4): 355-385.
- Badiqué, E., M. Cavazza, G. Klinker, G. Mair, T. Sweeney, D. Thalmann y N. M. Thalmann (2002). Entertainment Applications of Virtual Environments. Handbook of Virtual Environments: Design, Implementation, and Applications. K. M. Stanney, Lawrence Erlbaum Associates: 1143 - 1166.
- Balasingham, I., E. Samset, A. Hansen y L. Aurdal (2003). "An interactive augmented reality 3D visualization system for destroying liver tumor using cryoablation." International Congress Series **1256**: 690-695.
- Barakonyi, I., H. Prendinger, D. Schmalstieg y M. Ishizuka (2007). Cascading Hand and Eye Movement for Augmented Reality Videoconferencing. IEEE Symposium on 3D User Interfaces (3DUI'07), Charlotte NC, USA.
- Barakonyi, I., T. Psik y D. Schmalstieg (2004). Agents That Talk And Hit Back: Animated Agents in Augmented Reality. International Symposium on Mixed and Augmented Reality (ISMAR'04), IEEE Computer Society, Washington, DC, USA.
- Barakonyi, I., M. Weilguny, T. Psik y D. Schmalstieg (2005). MonkeyBridge: autonomous agents in augmented reality games. Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology, ACM Press New York, NY, USA.

- Barfield, W. (1995). Virtual Environments and Advanced Interface Design. Cary, NC, USA, Oxford University Press.
- Barfield, W., K. Baird, J. Skewchuk y G. Ioannou (2001a). Applications of Wearable Computers and Augmented Reality to Manufacturing. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 695-714.
- Barfield, W., S. Mann, K. Baird, F. Gemperle, C. Kasabach, J. Stivoric, M. Bauer, R. Martin y G. Cho (2001b). Computational Clothing and Accessories. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 471-510.
- Barfield, W., C. Rosenberg y W. A. Lotens (1995). Augmented-Reality Displays. Virtual Environments and Advanced Interface Design. W. Barfield y T. A. Furness, Oxford University Press, USA: 557 - 590.
- Bartie, P. J. y W. A. Mackaness (2006). "Development of a Speech-Based Augmented Reality System to Support Exploration of Cityscape." Transactions in GIS **10**(1).
- Bass, L., D. Siewiorek, M. Bauer, R. Casciola, C. Kasabach, R. Martin, J. Siegel, A. Smailagic y J. Stivoric (2001). Constructing Wearable Computers for Maintenance Applications. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 663-694.
- Beaudouin-Lafon, M. (1994). Beyond the Workstation: Mediaspaces and Augmented Reality. People and Computers. G. Cockton, S. W. Drapery y G. R. S. Weir. Glasgow, Cambridge University Press. **IX**.
- Behringer, R., S. Chen, V. Sundareswaran, W. Kenneth y M. Vassiliou (1999). "A distributed device diagnostics system utilizing augmented reality and 3D audio." Computers & Graphics **23**(6): 821-825.
- Bekel, H., G. Heidemann y H. Ritter (2005). "Interactive image data labeling using self-organizing maps in an augmented reality scenario." Neural Networks **18**(5-6): 566-574.
- Ben-Joseph, E., H. Ishii, J. Underkoffler, B. Piper y L. Yeung (2001). "Urban Simulation and the Luminous Planning Table: Bridging the Gap between the Digital and the Tangible." Journal of Planning Education and Research **21**(2): 196-203.
- Benford, S., A. Crabtree, M. Flintham, A. Drozd, R. Anastasi, M. Paxton, N. Tandavanitj, M. Adams y J. Row-Farr (2006). Can You See Me Now? ACM Transactions on Computer-Human Interaction (TOCHI), ACM Press New York, NY, USA.
- Berenguer, J. M. (2008). "Sonoscop, Archivo de Arte Sonoro." Visto el 20 Febrero, 2008, en <http://www.sonoscop.net/jmb/pub/gps.pat.zip>.
- Berenguer, X. (2002). "Arte y tecnología: una frontera que se desmorona." Visto el 22 Febrero, 2008, en <http://www.uoc.edu/artnodes/esp/art/xberenguer0902/xberenguer0902.html>.
- Billingshurst, M., A. Cheek, S. Prince y H. Kato (2002a). "Real World Teleconferencing." IEEE Computer Graphics and Applications **22**(6): 11-13.

- Billinghurst, M. y H. Kato (2002b). Collaborative Augmented Reality. Communications of the ACM.
- Billinghurst, M., H. Kato y I. Poupyrev (2001a). "The MagicBook - Moving Seamlessly between Reality and Virtuality." Computers and Graphics: 2 - 4.
- Billinghurst, M., E. Miller y S. Weghorst (2001b). Collaboration With Wearable Computers. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 539-578.
- Bimber, O. (2003). Augmented Reality, Part 3, Basic Rendering, Faculty of Media, Bauhaus University of Weimar: 34.
- Bimber, O. (2007a). "Dynamic Adaptation of Projected Imperceptible Codes." Visto el 4 Febrero, 2008, en <http://www.uni-weimar.de/medien/ar/research.php>.
- Bimber, O. (2007b). "Laser Pointer Tracking in Projector-Augmented Architectural Environments." Visto el 4 Febrero, 2008, en <http://www.uni-weimar.de/medien/ar/research.php>.
- Bimber, O., F. Coriand, A. Kleppe, E. Bruns, S. Zollmann y T. Langlotz (2005a). "Superimposing Pictorial Artwork with Projected Imagery." IEEE MultiMedia **12**(1): 16 - 26.
- Bimber, O., A. Emmerling y T. Klemmer (2005b). "Embedded Entertainment with Smart Projectors." Computer **38**(1): 48 - 55.
- Bimber, O., L. M. Encarnação y P. Branco (2001a). "The Extended Virtual Table: An Optical Extension for Table-Like Projection Systems." Presence: Teleoperators and Virtual Environments **10**(6): 613-631.
- Bimber, O., L. M. Encarnação y D. Schmalstieg (2003). The virtual showcase as a new platform for augmented reality digital storytelling. Proceedings of the workshop on Virtual environments 2003, ACM Press. New York, NY, USA.
- Bimber, O., L. M. Encarnao y D. Schmalstieg (2000). "Augmented Reality with Back-Projection Systems using Transflective Surfaces." Computer Graphics Forum **19**(3).
- Bimber, O., B. Fröhlich, D. Schmalstieg y L. M. Encarnação (2001b). "The Virtual Showcase." IEEE Computer Graphics and Applications **21**(6): 48-55.
- Bimber, O. y R. Raskar (2005c). Spatial Augmented Reality. SIGGRAPH 2005 Course 30 Notes.
- Bimber, O. y R. Raskar (2005d). Spatial Augmented Reality: Merging Real and Virtual Worlds.
- Bimber, O., R. Raskar y M. Inami (2007). Spatial Augmented Reality. SIGGRAPH 2007 Course 17 Notes.
- Bimber, O., A. Stork y P. Branco (2001c). Projection-Based Augmented Engineering. Human Computer Interaction International (HCI'01).
- Biosca Taronger, J. M., S. Navarro Tarín y J. L. Lerma García (2007). Modelado tridimensional de una bóveda barroca mediante la combinación de láser escáner y fotogrametría. 7 Setmana Geomàtica, Barcelona (España).
- Blast Theory. (2007). "Can you see me now?" Visto el 10 diciembre, 2007, en http://www.blasttheory.co.uk/bt/work_cysmn.html.
- Bleecker, J. (2005). "Processing + GPS." Visto el 8 Febrero, 2008, en <http://interactive.usc.edu/members/jbleecker/archives/005865.html>.

- Bobick, A., J. Davis y S. Intille (1997). "The KidsRoom: An Example Application Using a Deep Perceptual Interface." Workshop on Perceptual User Interfaces: 1 - 4.
- Bobick, A. F., S. S. Intille, J. W. Davis, F. Baird, C. S. Pinhanez, L. W. Campbell, Y. A. Ivanov, A. Schütte y A. Wilson (1999). "The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment." Presence: Teleoperators and Virtual Environments **8**(4): 367-391.
- Boehler, W., M. B. Vicent, K. Hanke y A. Marbs (2003). Documentation of German Emperor Maximilian I's Tomb. XIX International Symposium of The ICCMOSS & ISPRS Committee for Documentation of Cultural Heritage (CIPA'03), Antalya, Turkey.
- Bonsor, K. (2008). "How Smart Clothes Work." Visto el 10 Febrero, 2008, en <http://computer.howstuffworks.com/computer-clothing2.htm>.
- Brandt, E. (2006). Designing exploratory design games: a framework for participation in Participatory Design? Proceedings of the ninth conference on Participatory design: Expanding boundaries in design, ACM Press New York, NY, USA.
- Breen, D. E., R. T. Whitaker, E. Rose y M. Tuceryan (1996). "Interactive Occlusion and Automatic Object Placement for Augmented Reality." Computer Graphics Forum **15**(3).
- Broll, W., I. Lindt, J. Ohlenburg, M. Wittkämper, C. Yuan, T. Novotny, A. F. g. Schiecky, C. Mottramy y A. Strothmann (2004). "ARTHUR: A Collaborative Augmented Environment for Architectural Design and Urban Planning." Journal of Virtual Reality and Broadcasting **1**(1): 1 - 10.
- Brown, B., I. MacColl, M. Chalmers, A. Galani, C. Randell y A. Steed (2003). Lessons From The Lighthouse: Collaboration In A Shared Mixed Reality System. Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press New York, NY, USA.
- Calado Lopes, A. y J. M. Dias Sales (2006). "Integration of Geo-Referenced Data for Visual Simulation in Location-Based Mobile Computing." Computer-Aided Civil and Infrastructure Engineering **21**(7).
- Cartwright, W., M. Peterson y G. Gartner (1999). Multimedia Cartography. Berlin, Springer Verlag.
- Ciger, J., M. Gutierrez, F. Vexo y D. Thalmann (2003). The Magic Wand.
- Ciulli, N., S. Giordano y D. Sparano (1998). "A cooperative environment based on augmented reality: from telepresence to performance issues." Computer Networks and ISDN Systems **30**(16-18): 1447-1455.
- Coate, M. (2007). "This is Cinerama in Los Angeles." Visto el 10 diciembre, 2007, en http://www.fromscripttodvd.com/cinerama_la.htm.
- Cooper, N., A. Keatley, M. Dahlquist, S. Mann, H. Slay, J. Zucco, R. Smith y B. H. Thomas (2004). Augmented Reality Chinese Checkers.
- Cornelis, K., M. Pollefeys, M. Vergauwen, F. Verbiest y L. V. Gool (2001). Tracking Based Structure and Motion Recovery for Augmented Video Productions. ACM symposium on virtual reality software and technology (VRST'01), Banff, Canada, ACM.

- Courchesne, L. (2004). *De l'autre côté du miroir. ÊJouable. Art, jeu et interactivité*, Haute école d'arts appliqués. HES. Genève. École nationale supérieure des arts décoratifs. Université Paris.
- Cruz-Neira, C., D. J. Sandin, T. A. DeFanti, R. V. Kenyon y J. C. Hart (2002). "The CAVE: audio visual experience automatic virtual environment." *Communications of the ACM* **35**(6): 64 - 72.
- Cruz Sánchez, P. A. y M. Á. Hernández-Navarro (2004). *Cartografías del cuerpo. La dimensión corporal en el arte contemporáneo*, Cendeac.
- CVLAB - Computer Vision Laboratory. (2007). "bazAR Documentation." Visto el 29 September, 2007, en <http://cvlab.epfl.ch/software/bazar/html/index.html>.
- Cycling '74. (2008). "Jitter, a Brilliant Collection of Video, Matrix, and 3D Graphics Objects for Max." Visto, en <http://www.cycling74.com/products/jitter>.
- Charles, F., M. Cavazza, S. J. Mead, O. Martin, A. Nandi y X. Marichal (2004). *Compelling Experiences in Mixed Reality Interactive Storytelling*. Advanced Computer Entertainment (ACE'04).
- Chavarro García, O. X. (2007). "Computer Graphics Timeline." Visto el 10 diciembre, 2007, en http://sophia.javeriana.edu.co/~ochavarr/computer_graphics_history/historia/.
- Cheok, A. D., S. W. Fong, K. H. Goh, X. Yang, W. Liu y F. Farbiz (2003). *Human Pacman: a sensing-based mobile entertainment system with ubiquitous computing and tangible interaction*. Network and System Support for Games, Redwood City, California, ACM New York, NY, USA
- Cheok, A. D., K. H. Goh, W. Liu, F. Farbiz, S. W. Fong, S. L. Teo, Y. Li y X. Yang (2004). "Human Pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing." *Personal and Ubiquitous Computing* **8**(2): 71 - 81.
- Cheok, A. D., K. S. Teh, T. H. D. Nguyen, T. C. T. Qui, S. P. Lee, L. Wei, C. C. Li, D. Diaz y C. Boj (2006). "Social and physical interactive paradigms for mixed-reality entertainment." *ACM Computers in Entertainment* **4**(2).
- Cheok, A. D., W. Weihua, X. Yang, S. Prince, F. S. Wan, M. Billinghurst y H. Kato (2002a). *Interactive Theatre Experience in Embodied + Wearable Mixed Reality Space*. ISMAR.
- Cheok, A. D., X. Yang, Z. Z. Ying, M. Billinghurst y H. Kato (2002b). "Touch-Space: Mixed Reality Game Space Based on Ubiquitous, Tangible, and Social Computing." *Personal and Ubiquitous Computing* **6**: 430 - 442.
- Dangelmaier, W., M. Fischer, J. Gausemeier, M. Grafe, C. Matysczok y B. Mueck (2005). "Virtual and augmented reality support for discrete manufacturing system simulation." *Computers in Industry* **56**(4): 371-383.
- Dasgupta, S. y A. Banerjee (2006). "An Augmented-Reality-Based Real-Time Panoramic Vision System for Autonomous Navigation." *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS*, **36**(1): 154 - 161.
- De la Escalera Hueso, A. (2001). *Visión por computador. Fundamentos y métodos*. Madrid.

- De Souza e Silva, A. y G. C. Delacruz (2006). "Hybrid Reality Games Reframed: Potential Uses in Educational Contexts." Games and Culture **1**(3): 231 - 251.
- Demiris, A. M., V. Vlahakis, A. Makri, M. Papaioannou y N. Ioannidis (2005). "intGuide: A platform for context-aware services featuring augmented-reality, based on the outcome of European Research Projects." Signal Processing: Image Communication **20**(9-10): 927-946.
- Díaz, D. J. (2007). De la plaza al chat: análisis de las transformaciones del espacio público desde la práctica artística neomedial. Facultad de Bellas Artes. Departamento de Escultura. Valencia, Universidad Politécnica de Valencia. **Tesis doctoral**: 459.
- Díaz, J. L. (2008). "VI. Ciencia y Arte: la inteligencia de las musas." El ábaco, la lira y la rosa. Las regiones del conocimiento Visto el 22 Febrero, 2008, en http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen3/ciencia3/152/htm/sec_11.htm.
- Domènech, M. (2006). Algunas reflexiones sobre identidad, juego, interactividad. Laboratorio de Luz. Especulaciones a un tiempo, Consejería de Educación y Cultura. Centro Párraga: 97 - 110.
- Dorfmueller-Ulhaas, K. (2003). Augmented Reality, Part 8, Markerdetektion, Faculty of Media, Bauhaus University of Weimar: 50.
- Dorfmueller, K. (1999). "Robust tracking for augmented reality using retroreflective markers." Computers & Graphics **23**(6): 795-800.
- Dow, S., J. Lee, C. Oezbek, B. MacIntyre, J. D. Bolter y M. Gandy (2005a). Exploring spatial narratives and mixed reality experiences in Oakland cemetery. Advanced Computer Entertainment (ACE'05), Valencia.
- Dow, S., J. Lee, C. Oezbek, B. MacIntyre, J. D. Bolter y M. Gandy (2005b). Wizard of Oz Interfaces for Mixed Reality Applications. (CHI'05).
- Dow, S., M. Mehta, A. Lausier, B. MacIntyre y M. Mateas (2006). Initial Lessons From AR Façade, An Interactive Augmented Reality Drama. Advanced Computer Entertainment (ACE'06).
- Dünser, A., K. Steinbügl, H. Kaufmann y J. Glück (2006). Virtual and Augmented Reality as Spatial Ability Training Tools. (Chinz'06).
- Electrotap. (2008). "Software and hardware for innovative music, media, and art." Visto el 8 Febrero, 2008, en <http://www.electrotap.com/taptools/>.
- Ellis, S. R. (2007). "Virtual Environments and Environmental Instruments." Visto el 12 Noviembre, 2007, en http://human-factors.arc.nasa.gov/ihh/spatial/papers/pdfs_se/Ellis_1996_Virtual_Environments.html.
- EMTAC (2004). BTGPS II Trine Receptor GPS Guía de usuario.
- EMTAC. (2007). "EMTAC Wireless Solutions." Visto el 5 Febrero, 2008, en <http://www.emtac.com/>.
- Fiala, M. (2004). ARTag Revision 1, a Fiducial Marker System Using Digital Techniques. Canada, Computational Video Group, Institute for Information Technology, National Research Council Canada: 47.

- Fischer, J., D. Bartz y W. Straßer (2004). Occlusion Handling for Medical Augmented Reality using a Volumetric Phantom Model. (VRST'04).
- Fischer, J., D. Bartz y W. Straßer (2005). Artistic Reality: Fast Brush Stroke Stylization for Augmented Reality. (VRST'05).
- Fischer, J., M. Eichler, D. Bartz y W. Straer (2007). "A hybrid tracking method for surgical augmented reality." Computers & Graphics **31**(1): 39-52.
- Fischer, J., H. Regenbrecht y G. Baratoff (2003). "Detecting Dynamic Occlusion in front of Static Backgrounds for AR Scenes." ACM Transactions on Computer-Human Interaction: 153 - 161.
- Flintham, M., R. Anastasi, S. Benford, T. Hemmings, A. Crabtree, C. Greenhalgh, T. Rodden, N. Tandavanitj, M. Adams y J. Row-Farr (2003). Where On-Line Meets On-The-Streets: Experiences With Mobile Mixed Reality Games. CHI.
- Florins, M., D. G. Trevisan y J. Vanderdonckt (2005). The continuity property in mixed reality and multiplatform systems: a comparative study. Computer-Aided Design of User Interfaces IV. R. Jacob. Netherlands, Kluwer Academic Publishers. **26**: 323 - 334.
- Freeman, R. y A. Steed (2006). Interactive Modelling and Tracking for Mixed and Augmented Reality. VRST'06.
- Fuhrmann, A., H. Löffelmann, D. Schmalstieg y M. Gervautz (1998). "Collaborative Visualization in Augmented Reality." IEEE Computer Graphics and Applications: 54-59.
- Fujihata, M. (2002). "«Field-Work@Alsace»." Visto el 14 Febrero, 2008, en <http://www.medienkunstnetz.de/works/field-work/>.
- Galantay, R., J. Torpus y M. Engeli (2004). "living-room" Interactive, Space-Oriented Augmented Reality. ACM Multimedia (MM'04).
- García Canal, C. (2008). "Ciencia = Cultura." rei, la revista electrónica del IFIC Visto el 24 Febrero, 2008, en http://rei.ific.uv.es/rei/index.php/rei/arte_y_ciencia/ciencia_cultura.
- García Cortés, J. M. (1997). El rostro velado (Travestismo e Identidad en el arte). Donostia, Koldomixelena Kulturunea.
- Gardenfors, P. (2005). Cognition, Education, and Communication Technology, Mahwah, NJ, USA: Lawrence Erlbaum Associates.
- Gartner, G. (2004). Location-based mobile pedestrian navigation services – the role of multimedia cartography. ICA UPIMap, Tokyo.
- Garzotto, F. y M. Forfori (2006). Hyperstories and Social Interaction in 2D and 3D Edutainment Spaces for Children. HT'06.
- Gelenbe, E., K. Hussain y V. Kaptan (2005). "Simulating autonomous agents in augmented reality." Journal of Systems and Software **74**(3): 255-268.
- Giannetti, C. (2002). Estética Digital. Sintopía del arte, la ciencia y la tecnología. Barcelona, ACC L'Angelot.
- Gibson, S., J. Cook, T. Howard y R. Hubbard (2003). Rapid Shadow Generation in Real-World Lighting Environments. Eurographics Symposium on Rendering.

- Gillet, A., M. Sanner, D. Stoffler, D. Goodsell y A. Olson (2004). Augmented Reality with Tangible Auto-Fabricated Models for Molecular Biology Applications. IEEE Visualization.
- Giner Martínez, F. y C. Portalés Ricart (2005). The Augmented User: A Wearable Augmented Reality Interface. International Conference on Virtual Systems and Multimedia (VSMM'05) Hal Thwaites.
- Gleue, T. y P. Dähne (2001). Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the ARCHEOGUIDE Project. Virtual Reality, Archaeology, and Cultural Heritage International Symposium (VAST'01), Glyfada, Nr Atenas, Grecia.
- Glossop, N. D. y Z. Wang (2003). "Laser projection augmented reality system for computer-assisted surgery." International Congress Series **1256**: 65-71.
- Goethe, J. W. v. (1774). Goethe, Zahme Xenien IX. (Anfang).
- Google. (2007). "Google Earth." Visto el 12 Febrero, 2008, en <http://earth.google.es/>.
- Grasset, R., J. Looser y M. Billinghurst (2005). OSGARToolKit: Tangible + Transitional 3D Collaborative Mixed Reality Framework. ICAT.
- Grau, O. (2003). Virtual Art. From Illusion to Immersion. London, The MIT Press.
- Grewal, M. S., L. R. Weill y A. P. Andrews (2001). Global Positioning Systems, Inertial Navigation, and Integration, John Wiley & Sons.
- Gruen, A. y T. S. Huang (2001). Calibration and Orientation of Cameras in Computer Vision. Berlin, Springer.
- Guillem Picó, S. y P. Acosta Matarredona (2001). Geodesia clásica. Conceptos generales. Valencia, Universidad Politécnica de Valencia.
- Gutierrez, D., F. J. Seron, J. A. Magallon, E. J. Sobreviela y P. Latorre (2004). "Archaeological and cultural heritage: bringing life to an unearthed Muslim suburb in an immersive environment." Journal of Cultural Heritage **5**: 63 - 74.
- Ha, T., Y. Chang y W. Woo (2007). Usability Test of Immersion for Augmented Reality Based Product Design. Edutainment.
- Halme, A., J. Suomela y M. Savela (1999). "Applying telepresence and augmented reality to teleoperate field robots." Robotics and Autonomous Systems **26**(2-3): 117-125.
- Hall, T., L. Ciolfi, L. Bannon, M. Fraser, S. Benford, J. Bowers, C. Greenhalgh, S.-O. Hellström, S. Izadi, H. Schnädelbach y M. Flintham (2002). The Visitor as Virtual Archaeologist: Explorations in Mixed Reality Technology to Enhance Educational and Social Interaction in the Museum. ACM
- Haller, M. (2004). Photorealism or/and Non-Photorealism in Augmented Reality. ACM
- Hart, M. B. (1995). "The American WideScreen Museum. Cinerama." Visto el 10 diciembre, 2007, en <http://www.widescreenmuseum.com/widescreen/wingcr1.htm>.
- Hashimoto, Y., N. Nagaya, M. Kojima, J. Ohtaki, T. Mitani, A. Yamamoto, S. Miyajima y M. Inami (2005). Straw-like user interface. Computer Graphics and Interactive Techniques, Los Angeles, California, ACM New York, NY, USA.

- Hatala, M., L. Kalantari, R. Wakkary y K. Newby (2004). Ontology and Rule based Retrieval of Sound Objects in Augmented Audio Reality System for Museum Visitors. ACM Symposium on Applied Computing (SAC'04).
- Hawkins, K. (2001). OpenGL Game Programming, Roseville : Prima Tech, cop.
- Hedley, N. R., M. Billinghamurst, L. Postner, R. May y H. Kato (2002). "Explorations in the use of augmented reality for geographic visualization." Presence: Teleoperators and Virtual Environments **11**(2): 119 - 133.
- Heilig, M. L. (2007). "Inventor in the Field of Virtual Reality." Visto el 12 Noviembre, 2007, en <http://www.mortonheilig.com/InventorVR.html>.
- Heller, E. (2004). Psicología del color. Barcelona, Ed. Gustavo Gili, S.A.
- Hicks, J., R. Flanagan, P. Petrov y A. Stoyen (2003). EyeKon: Distributed Augmented Reality for Soldier Teams. International Command and Control Research and Technology Symposium (ICCRTS'03), Washington, DC.
- Hicks, J. D., R. A. Flanagan, P. V. Petrov y A. D. Stoyen (2002). EyeKon: augmented reality for battlefield soldiers. Software Engineering Workshop, NASA Goddard/IEEE.
- Hirose, R. y H. Saito (2005). A Vision-Based AR Registration Method Utilizing Edges and Vertices of 3D Model. ICAT.
- HITLab. (2006). "Human Interface Technology Laboratory." Visto el 12 Noviembre, 2007, en <http://www.hitl.washington.edu/home/>.
- HITLab NZ. (2006). "OSGART: ARToolKit for OpenSceneGraph." Visto el 12 Noviembre, 2007, en <http://www.artoolworks.com/community/osgart/>.
- Hoberman, P., N. Parés y R. Parés. (1998). "El ball del fanalet." Visto el 4 Febrero, 2008, en http://www.iaa.upf.es/~gvirtual/lghtmls/lghtml_c.htm.
- Höllerer, T., S. Feiner, D. Hallaway, B. Bell, M. Lanzagorta, D. Brown, S. Julier, Y. Baillot y L. Rosenblum (2001). "User interface management techniques for collaborative mobile augmented reality." Computers & Graphics **25**(5): 799-810.
- Höllerer, T., S. Feiner, T. Terauchi, G. Rashid y D. Hallaway (1999). "Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system." Computers & Graphics **23**(6): 779-785.
- Human Interface Technology Laboratory. (2007). "ARToolKit." Visto el 30 September, 2007, en <http://www.hitl.washington.edu/artoolkit/>.
- Immersion Corporation. (2007a). "CyberGrasp™ Exoskeleton." Visto el 5 Noviembre, 2007, en http://www.immersion.com/3d/products/cyber_grasp.php.
- Immersion Corporation. (2007b). "Wireless Data Glove: The CyberGlove® II System." Visto el 5 Noviembre, 2007, en http://www.immersion.com/3d/products/cyber_glove.php.
- Inami, M., N. Kawakami, D. Sekiguchi, Y. Yanagida, T. Maeda y S. Tachi (2000). Visuo-Haptic Display Using Head-Mounted Projector. Virtual Reality, IEEE Computer Society Washington, DC, USA.

- Ishii, H., M. Kobayashi y K. Arita (1994). "Iterative design of seamless collaboration media." Communications of the ACM **37**(8): 83 - 97.
- Jacobs, K. y C. Loscos (2006). "Classification of Illumination Methods for Mixed Reality." Computer Graphics Forum **25**(1).
- Jähne, B., H. Haussecker y P. Geissler, Eds. (1999). Handbook of Computer Vision and Applications. Sensors and Imaging.
- Janin, A. L., D. W. Mizell y T. P. Caudell (1993). Calibration of head-mounted displays for augmented reality applications. IEEE Virtual Reality Annual Int. Symp., Seattle, WA.
- Juan, M. C., M. Alcañiz, J. Calatrava, I. Zaragoza, R. Baños y C. Botella (2007). An Optical see-through AR system for the treatment of phobias to small animals. Lecture Notes in Computer Science, ICVR'07.
- Juan, M. C., M. Alcañiz, C. Monserrat, C. Botella, R. M. Baños y B. Guerrero (2005). "Using augmented reality to treat phobias." IEEE Computer Graphics and Applications **25**(6): 31-37.
- Juan, M. C., R. Baños, C. Botella, D. Pérez, M. Alcañiz y C. Monserrat (2006a). "An Augmented Reality System for the Treatment of Acrophobia: The Sense of Presence Using Immersive Photography." Presence: Teleoperators & Virtual Environments **15**(4): 393-402.
- Juan, M. C., D. Joele, R. Baños, C. Botella, M. Alcañiz y C. van der Mast (2006b). A Markerless Augmented Reality System for the treatment of phobia to small animals. Presence Conference.
- Kameda, Y., T. Takemasa y Y. Ohta (2004). Outdoor see-through vision utilizing surveillance cameras. Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04).
- Kangas, K. J. y J. Röning (2002). "Using Mobile Code to Create Ubiquitous Augmented Reality." Wireless Networks **8**: 199 – 211.
- Kato, H. y M. Billinghurst (1999). Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System. 2nd International Workshop on Augmented Reality (IWAR'99), San Francisco, USA.
- Kato, H., M. Billinghurst, K. Morinaga y K. Tachibana (2001). The Effect of Spatial Cues in Augmented Reality Video Conferencing. 9th International Conference on Human-Computer Interaction (HCI International 2001), New Orleans, LA, USA.
- Kato, H., M. Billinghurst y I. Poupyrev (2000a). ARToolKit, version 2.33.
- Kato, H., M. Billinghurst, I. Poupyrev, K. Imamoto y K. Tachibana (2000b). Virtual object manipulation on a table-top AR environment. International Symposium on Augmented Reality (ISAR'00).
- Kaufmann, H. (2002). Construct3D: An Augmented Reality Application for Mathematics and Geometry Education. ACE'02.
- Kaufmann, H. (2004). Geometry Education with Augmented Reality. Institut für Softwaretechnik und Interaktive Systeme. Viena, Technischen Universität Wien. **Tesis doctoral**: 179.

- Kaufmann, H. y D. Schmalstieg (2003). "Mathematics and geometry education with collaborative augmented reality." Computers & Graphics **27**(3): 339-345.
- Kawakami, N., M. Inami, Y. Yanagida y S. Tachi (1998). Object-oriented displays. Computer Graphics and Interactive Techniques, Orlando, Florida, United States, ACM New York, NY, USA
- Kealy, A. y S. Scott-Young (2006). "A Technology Fusion Approach for Augmented Reality Applications." Transactions in GIS **10**(2).
- Kijima, R. y T. Ojika (1997). Transition Between Virtual Environment and Workstation Environment with Projective Head Mounted Display. Virtual Reality Annual International Symposium, IEEE.
- King, G. R., W. Piekarski y B. H. Thomas (2005). ARVino — Outdoor Augmented Reality Visualisation of Viticulture GIS Data. International Symposium on Mixed and Augmented Reality (ISMAR'05), IEEE Computer Society.
- Kiyokawa, K., Y. Kurata y H. Ohno (2000a). An optical see-through display for mutual occlusion of real and virtual environments. Augmented Reality (ISAR'00), Munich, Germany.
- Kiyokawa, K., H. Ohno y Y. Kurata (2002). Occlusive Optical See-through Displays in a Collaborative Setup. Computer Graphics and Interactive Techniques, San Antonio, Texas, ACM New York, NY, USA.
- Kiyokawa, K., H. Takemura y N. Yokoya (2000b). "SeamlessDesign for 3D Object Creation." Multimedia Computing and Systems: 22 - 33.
- Klinker, G., D. Stricker y D. Reinert (1999). "Optically based direct manipulation for augmented reality." Computers & Graphics **23**(6): 827-830.
- Klopfer, E., J. Perry, K. Squire y M.-F. Jan (2005). Collaborative Learning through Augmented Reality Role Playing. Computer Support for Collaborative Learning, Taipei, Taiwan, International Society of the Learning Sciences.
- Kolbe, T. H. (2004). Augmented Videos and Panoramas for Pedestrian Navigation. Location Based Services & TeleCartography, Viena, G. Gartner (ed.).
- Koleva, B., I. Taylor, S. Benford, M. Fraser, C. Greenhalgh, H. Schnädelbach, D. v. Lehn, C. Heath, J. Row-Farr y M. Adams (2001). Orchestrating a Mixed Reality Performance. CHI'01, ACM.
- Koller, D., G. Klinker, E. Rose, D. Breen, R. Whitaker y M. Tuceryan (1997). Real-time Vision-Based Camera Tracking for Augmented Reality Applications. ACM symposium on Virtual reality software and technology (VRST'97), Lausanne, Switzerland, ACM New York, NY, USA
- Konishi, K., M. Hashizume, M. Nakamoto, Y. Kakeji, I. Yoshino, A. Taketomi, Y. Sato, S. Tamura y Y. Maehara (2005). "Augmented reality navigation system for endoscopic surgery based on three-dimensional ultrasound and computed tomography: Application to 20 clinical cases." International Congress Series **1281**: 537-542.
- Koutsoudis, A., F. Arnaoutoglou y C. Chamzas (2007). "On 3D reconstruction of the old city of Xanthi. A minimum budget approach to virtual touring based on photogrammetry." Journal of Cultural Heritage **8**: 26 - 31.

- Kraus, K. (1997). Photogrammetry. Advanced Methods and Applications. Bonn.
- Kriss, J. (2008). "jklabs :: MaxLink." Visto el 8 Febrero, 2008, en <http://jklabs.net/maxlink/>.
- Kunsthochschule für Medien Köln. (2007). "Virtual Reality - Historie." Visto el 12 Noviembre, 2007, en <http://www.khm.de/~vrmkurs/historie/historie.html>.
- Lam, A. H. T., K. C. H. Chow, E. H. H. Yau y M. R. Lyu (2006). ART: Augmented Reality Table for Interactive Trading Card Game. VRCIA'06, ACM.
- Las Provincias. (2006). "El Retablo de Plata de la Catedral de Valencia." Visto el 27 Octubre, 2007, en http://www.lasprovincias.es/valencia/prensa/20061229/cultura/retablo-plata-catedral-valencia_20061229.html.
- Le Prado, C. y S. Natkin (2007). LISTEN LISBOA: Scripting Languages for Interactive Musical Installations. Sound and Music Computing Conference, Lefkada, Greece.
- Lee, G. A., M. Billinghurst y G. J. Kim (2004). Occlusion based Interaction Methods for Tangible Augmented Reality Environments, ACM.
- Lee, S.-Y., S. C. Ahn, H.-G. Kim y M. Lim (2006a). "Real-time 3D video avatar in mixed reality: An implementation for immersive telecommunication." Simulation & Gaming **37**(4): 491 - 506.
- Lee, S. P., A. D. Cheok, T. K. S. James, G. P. L. Debra, C. W. Jie, W. Chuang y F. Farbiz (2006b). "A mobile pet wearable computer and mixed reality system for human-poultry interaction through the internet." Pers Ubiquit Comput **10**: 301 - 317.
- LEGO. (2008). "LEGO." Visto el 18 Diciembre, 2007, en <http://www.lego.com>.
- Lepetit, V. y M.-O. Berger (2000a). Handling Occlusion in Augmented Reality Systems: A Semi-Automatic Method. IEEE and ACM International Symposium on Augmented Reality, Munich, Germany.
- Lepetit, V. y M.-O. Berger (2000b). A Semi-Automatic Method for Resolving Occlusion in Augmented Reality. IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina (USA).
- Lepetit, V. y P. Fua (2004a). Towards Recognizing Feature Points using Classification Trees. Lausanne, Switzerland, Swiss Federal Institute of Technology (EPFL): 1 - 13.
- Lepetit, V., P. Lagger y P. Fua (2005). Randomized Trees for Real-Time Keypoint Recognition. Computer Vision and Pattern Recognition, San Diego, CA.
- Lepetit, V., J. Pilet y P. Fua (2004b). Point Matching as a Classification Problem for Fast and Robust Object Pose Estimation. Computer Vision and Pattern Recognition, Washington, DC.
- Lerma García, J. L. (2002). Fotogrametría Moderna: Analítica y Digital.
- Lerma, J. L., A. García y C. Pérez (2005). 3D Reconstruction and Modelling of Architectural Domes. CIPA XX Internacional Symposium, Torino (Italy).
- Lévi-Strauss, C. (1982). The Way of the Masks, Washington Press.
- Levin, G. (2006). "Computer Vision for Artists and Designers: Pedagogic Tools and Techniques for Novice Programmers." Journal of Artificial Intelligence and Society **20**(4): 462 - 482.
- Levin, G. y Collaborators. (2008). "Hidden Worlds of Noice and Voice." Visto el 19 Febrero, 2008, en <http://www.flong.com/projects/hwnv/>.

- Livingston, M. A. (2005). "Evaluating Human Factors in Augmented Reality Systems." IEEE Computer Graphics and Applications. Projects in VR: 6 - 9.
- Livingston, M. A., L. J. Rosenblum, S. J. Julier, D. Brown, Y. Baillot, J. E. Swan, J. L. Gabbard y D. Hix (2002). An augmented reality system for military operations in urban terrain. Interservice / Industry Training, Simulation, & Education Conference (I/ITSEC '02), Orlando, FL.
- Logitech. (2007). "VRRealities." Visto el 20 Diciembre, 2007, en <http://www.vrealities.com/logitech.html>.
- Low, K.-L., G. Welch, A. Lastra y H. Fuchs (2001). Life-sized projector-based dioramas. Virtual Reality Software and Technology, ACM New York, NY, USA.
- Lozano-Hemmer, R. (2007). "Rafael Lozano-Hemmer." Visto el 15 Febrero, 2008, en <http://www.lozano-hemmer.com>.
- Lu, B., K. Lim, J. Zheng y Y. Cai (2004). Learning Molecular Biology By VR Playing, ACM.
- Lynch, K. (1960). The Image of the City. Cambridge, Massachusetts, The MIT Press.
- MacIntyre, B. y J. D. Bolter (2003). "Single-narrative, multiple point-of-view dramatic experiences in augmented reality." Virtual Reality 7: 10–16.
- Mackay, W. E. (1998). Augmented Reality: linking real and virtual worlds. A new paradigm for interacting with computers. Advanced visual interfaces (AVI'98), L'Aquila, Italy, ACM Press.
- Madritsch, F. y M. Gervautz (1996). "CCD-Camera Based Optical Beacon Tracking for Virtual and Augmented Reality." Computer Graphics Forum 15(3).
- Malkawi, A. M. y R. S. Srinivasan (2005). "A new paradigm for Human-Building Interaction: the use of CFD and Augmented Reality." Automation in Construction 14(1): 71-84.
- Mann, S. (1997). "Smart Clothing: The Wearable Computer and WearCam." Personal Technologies 1(1): 21-27.
- Mann, S. (2001). Fundamental Issues in Mediated Reality, WearComp, and Camera-Based Augmented Reality. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 295-328.
- Mañas, M., F. Giner, J. Montesa, C. Portalés, J. Soler, M. Alcañiz, M. J. Martínez De Pisón y D. Díaz (2005). Notbook AR. International Conference on Advances in Computer Entertainment Technology (ACE'05), Valencia (España), ACM SIGCHI.
- Marchand, r. y F. Chaumette (2002). "Virtual Visual Servoing: a framework for real-time augmented reality." Computer Graphics Forum 21(3).
- Matyszczok, C., R. Radkowski y J. Berssenbruegge (2004). AR-Bowling: Immersive and Realistic Game Play in Real Environments Using Augmented Reality. ACE'04, Singapore, ACM.
- Microvision. (2008). "A World of Display and Imaging Opportunities." Visto el 7 Febrero, 2008, en <http://www.microvision.com/>.
- Milgram, P. y D. Drascic (1997). Perceptual effects in aligning virtual and real objects in augmented reality displays. 41 st Annual Meeting of Human Factors and Ergonomics Society.

- Milgram, P. y F. Kishino (1994). "A taxonomy of mixed reality visual displays." IEICE Transactions on Networked Reality **E77-D**(12): 1321 - 1329.
- Ministerio de Educación y Ciencia. (2006). "Banco de sonidos." Visto el 6 Mayo, 2006, en <http://recursos.cnice.mec.es/bancoimagenes/sonidos/index.php>.
- Mischitz, G. E. (2001). "The History of Human Computer Interaction." Visto el 10 diciembre, 2007, en <http://www2.iicm.tugraz.at/cguetl/education/projects/mischitz/Seminar.htm>.
- Mischkowski, R. A., M. J. Zinser, A. C. Kubler, B. Krug, U. Seifert y J. E. Zoller (2006). "Application of an augmented reality tool for maxillary positioning in orthognathic surgery - A feasibility study." Journal of Cranio-Maxillofacial Surgery **34**(8): 478-483.
- Mixed Reality Lab (2003a). MXRToolKit: Mixed Reality Camera Calibration Tool, National University Singapore.
- Mixed Reality Lab (2003b). MXRToolKit: Mixed Reality Model Aligment Tool, National University Singapore: 11.
- Mixed Reality Lab (2003c). MXRToolKit: Mixed Reality Software Developement Kit, National University Singapore.
- Mizell, D. (2001). Boeing's Wire Bundle Assembly Project. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 447-470.
- Moggridge, B. (2007). Designing interactions. Cambridge, Massachusetts; London, England, MIT Press.
- Mogilev, D., K. Kiyokawa, M. Billinghurst y J. Pair (2002). AR Pad: An Interface for Face-to-Face AR Collaboration. Human Factors in Computing Systems, Minneapolis, Minnesota, USA.
- Motorola. (2008). "RD5000 Mobile RFID Reader." Visto el 4 Febrero, 2008, en <http://www.motorola.com/business/v/index.jsp?vgnextoid=95ebe90e3ae95110VgnVCM1000008406b00aRCRD>.
- Mulloni, A. (2007). "Handheld Augmented Reality. Augmented Reality on Truly Mobile Devices: Cows vs. Aliens." Visto el 12 Noviembre, 2007, en http://studierstube.icg.tugraz.ac.at/handheld_ar/cows_vs_alien.php.
- Murray, J. H. (1999). Hamlet en la Holocubierta : El futuro de la narrativa en el ciberespacio. Barcelona, Paidós.
- MXR. (2006). "Mixed Reality Lab." Visto el 1 Febrero, 2008, en <http://www.mixedrealitylab.org/>.
- Narzt, W., G. Pomberger, A. Ferscha, D. Kolb, R. Müller, J. Wieghardt, H. Hörtner y C. Lindinger (2006). "Augmented reality navigation systems." Universal Access in the Information Society **4**(3): 177-187.
- Nash, M. y D. Marre (2001). Multiculturalismos y género. Un estudio interdisciplinar. Barcelona, Edicions bellaterra.
- National Research Council Staff (1994). Virtual Reality : Scientific and Technological Challenges, Washington, DC, USA: National Academies Press.

- Nawab, A., K. Chintamani, D. Ellis, G. Auner y A. Pandya (2007). Joystick mapped Augmented Reality Cues for End-Effector controlled Tele-operated Robots. Virtual Reality (VR'07).
- Neumann, U. (2001). STAR: Tracking for Object-Centric Augmented Reality. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 329-352.
- Neumann, U. y S. You (1999). "Natural Feature Tracking for Augmented Reality." IEEE TRANSACTIONS ON MULTIMEDIA **1**(1): 53 - 64.
- Newman, J., M. Wagner, T. Pintaric, A. MacWilliams, M. Bauer, G. Klinker y D. Schmalstieg (2003). Fundamentals of Ubiquitous Tracking for Augmented Reality. Viena, TU-Wien: 18.
- Nikou, C., A. M. Digioia lii, M. Blackwell, B. Jaramaz y T. Kanade (2000). "Augmented reality imaging technology for orthopaedic surgery." Operative Techniques in Orthopaedics **10**(1): 82-86.
- Norwood, S. E. (1997). "Cinerama." Visto el 10 diciembre, 2007, en <http://www.redballoon.net/~snorwood/book/>.
- Ox, J. y D. Britton (2000). "The 21st Century Virtual Reality Color Organ." Artful Media: 2-5.
- Pan, Z., A. D. Cheok, H. Yang, J. Zhu y J. Shi (2006). "Virtual reality and mixed reality for virtual learning environments." Computers & Graphics **30**: 20 - 28.
- Park, J.-S., T. Kim y J.-H. Yoon (2006). AR Table Tennis: A Video-Based Augmented Reality Sports Game. ICAT'06, Springer-Verlag Berlin Heidelberg.
- Paul, C. (2003). Digital Art. 224, Thames & Hudson world of art.
- Paulos, E. y T. Jenkins (2005). Urban Probes: Encountering Our Emerging Urban Atmospheres. CHI'05, ACM.
- PD. (2008). "Pure Data." Visto el 8 Febrero, 2008, en <http://puredata.info/>.
- Peitz, J., S. Björk y A. Jäppinen (2006). Wizard's Apprentice gameplay-oriented design of a computer-augmented board game. ACE'06, California, USA, ACM.
- Pelletier, J.-M. (2006). "cv.jit computer vision for jitter." Visto el 8 Febrero, 2008, en <http://www.iamas.ac.jp/~jovan02/cv/>.
- Pérez, D., M. C. Juan, M. Alcañiz, R. Baños y C. Botella (2007). Two similar systems for the treatment of acrophobia. An AR system and a Virtual Reality system. Lecture Notes in Computer Science, ICVR'07.
- Pérez García (ed.), M. d. C. (2006). Los ángeles músicos de la catedral de Valencia. Valencia, Generalitat Valenciana.
- Piekarski, W. (2006). "3D Modeling with the Tinmith Mobile Outdoor Augmented Reality System." IEEE Computer Graphics and Applications. Projects in VR: 14 - 17.
- Piekarski, W. y B. H. Thomas (2003). ThumbsUp: Integrated Command and Pointer Interactions for Mobile Outdoor Augmented Reality Systems. Human-computer Interaction: Theory and Practice. C. Stephanidis, Lawrence Erlbaum Associates. **1**: 1223 - 1227.
- Pieraccini, M., G. Guidi y C. Atzeni (2001). "3D digitizing of cultural heritage." Journal of Cultural Heritage **2**: 63 - 70.

- Pilet, J., A. Geiger, P. Laguerre, V. Lepetit y P. Fua (2006). An All-In-One Solution to Geometric and Photometric Calibration. International Symposium on Mixed and Augmented Reality (ISMAR'06).
- Pistoletto, M. (2008). "Michelangelo Pistoletto, Works." Visto el 7 Marzo, 2008, en <http://www.pistoletto.it/eng/crono04.htm#>.
- Point Grey. (2007). "Point Grey Research Products, Stereo Vision." Visto el 4 Febrero, 2008, en <http://www.ptgrey.com/products/legacy.asp>.
- Polhemus. (2007a). "FASTRAK." Visto el 5 Noviembre, 2007, en http://www.polhemus.com/?page=Motion_Fastrak.
- Polhemus. (2007b). "VisionTrak." Visto el 5 Noviembre, 2007, en http://www.polhemus.com/?page=Eye_VisionTrak.
- Portalés Ricart, C. (2006). Existe un Universo reflejo si y sólo si el Universo es igual a la intersección entre las personas con la unión de las dimensiones espacio y tiempo. Laboratorio de Luz. Especulaciones a un tiempo, Consejería de Educación y Cultura. Centro Párraga: 97 - 110.
- Prince, S. J. D., K. Xu y A. D. Cheek (2002). "Augmented Reality Camera Tracking with Homographies." IEEE Computer Graphics and Applications: 39 - 45.
- Processing. (2008). "Processing." Visto el 8 Febrero, 2008, en <http://processing.org/>.
- Rashid, O., W. Bamford, P. Coulton, R. Edwards y J. Sheible (2006). "PAC-LAN: Mixed-Reality Gaming with RFID Enabled Mobile Phones." ACM Computers in Entertainment **4**(4).
- Raskar, R., J. v. Baar, P. Beardsley, T. Willwacher, S. Rao y C. Forlines (2003). "iLamps: Geometrically Aware and Self-Configuring Projectors." ACM Transactions on Graphics (TOG) **22**(3): 809 - 818.
- Reilly, D. y H. Chen (2006). Mobile Lenses: a Hybrid Approach to Direct Interaction with Maps and Kiosks. Pervasive Mobile Interaction Devices (PERMID'06), Dublin, Ireland.
- Reimann, C. y V. Paelke (2005). Adaptive Mixed Reality Games. ACE'05, Valencia, ACM.
- Reitmayr, G. y D. Schmalstieg (2003). Location based applications for mobile augmented reality. Australasian User Interface Conference, Adelaide, Australia.
- Reitmayr, G. y D. Schmalstieg (2004). Scalable Techniques for Collaborative Outdoor Augmented Reality. Viena, TU-WIEN.
- Retscher, G. y M. Thienelt (2004). "NAVIO – A Navigation and Guidance Service for Pedestrians." Journal of Global Positioning Systems **3**(1-2): 208-217.
- Revilla, F. (2003). Diccionario de iconografía y simbología. Madrid, Ed. Cátedra: 167.
- Ribo, M., A. Pinz y A. L. Fuhrmann (2001). A new Optical Tracking System for Virtual and Augmented Reality Applications. IEEE Instrumentation and Measurement Technology Conference, Budapest, Hungary.
- Rieth, C. (2005). "Lego Electronic Lab Kit." Visto el 25 Febrero, 2008, en <http://computer.team.lu/legoelectronic/index.html>.

- Rolland, J., L. Davis, Y. Ha, C. Meyer, V. Shaoulov, A. Akcay, H. Zheng, R. Banks y B. D. Vento (2002). "3D Visualization and Imaging in Distributed Collaborative Environments." IEEE Computer Graphics and Applications. Projects in VR: 11 - 13.
- Rolland, J. P., L. D. Davis y Y. Baillot (2001). A Survey of Tracking Technologies for Virtual Environments. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 67-112.
- Romero, L., J. Santiago y N. Correia (2004). "Contextual Information Access and Storytelling in Mixed Reality Using Hypermedia." ACM Computers in Entertainment **2**(3).
- Rosenthal, M., A. State, J. Lee, G. Hirota, J. Ackerman, K. Keller, E. D. Pisano, M. Jiroutek, K. Muller y H. Fuchs (2002). "Augmented reality guidance for needle biopsies: An initial randomized, controlled trial in phantoms." Medical Image Analysis **6**(3): 313-320.
- Roussou, M. (2004). "Learning by Doing and Learning Through Play: An Exploration of Interactivity in Virtual Environments for Children." ACM Computers in Entertainment **2**(4): 1 - 23.
- Salzmann, M., S. Ilic y P. Fua (2006). Physically Valid Shape Parameterization for Monocular 3-D Deformable Surface Tracking.
- Sanchez-Segura, M. I. (2004). Developing Future Interactive Systems, Hershey, PA, USA: Idea Group Publishing.
- Sandor, C. y G. Klinker (2005). "A rapid prototyping software infrastructure for user interfaces in ubiquitous augmented reality." Pers Ubiquit Comput **9**: 169 – 185.
- ScienceDaily. (2000). "Readers Become Part Of The Action Through High-Tech Mixture Of Traditional Storytelling And Virtual Reality." Visto el 10 diciembre, 2007, en <http://www.sciencedaily.com/releases/2000/09/000921073201.htm>.
- Schnädelbach, H., B. Koleva, M. Fintham, M. Fraser, S. Izadi, P. Chandler, M. Foster, S. Benford, C. Greenhalgh y T. Rodden (2002). The Auguroscope: a mixed reality interface for outdoors. CHI'02.
- Shaheen, M., M. Mallem y F. Chavand (2001). "Visual command of a robot using 3D-scene reconstruction in an augmented reality system." Control Engineering Practice **9**(4): 375-385.
- Shaw, J. (1989). "The legible city. Manhattan version (1989), Amsterdam version (1990), Karlsruhe version (1991)." Visto el 5 Octubre, 2007, en http://www.jeffrey-shaw.net/html_main/show_work.php3?record_id=83.
- Sherman, W. R. y A. B. Craig (2003). Understanding virtual reality : interface, application, and design. Amsterdam, Morgan Kaufmann.
- Shoji, M., K. Miura y A. Konno (2006). U-Tsu-Shi-O-Mi: the virtual humanoid you can reach. International Conference on Computer Graphics and Interactive Techniques, Boston, Massachusetts, ACM New York, NY, USA
- SIGGRAPH06. (2006). "The 33rd International Conference and Exhibitions on Computer Graphics and Interactive Techniques." Visto, en <http://www.siggraph.org/s2006/>.

- Simsarian, K. T. y K.-P. Åkesson (1997). Windows on the World: An example of Augmented Virtuality. Interfaces 97: Man-Machine Interaction. Montpellier, Francia.
- Sinclair, P. A. S., K. Martinez, D. E. Millard y M. J. Weal (2002). Links in the Palm of your Hand: Tangible Hypermedia using Augmented Reality. ACM transactions on computer-human interaction.
- Slay, H., M. Phillips, R. Vernik y B. Thomas (2001). Interaction Modes for Augmented Reality Visualization. Australian Symposium on Information Visualisation. Conferences in Research and Practice in Information Technology, Sydney.
- Song, M., T. Elias, I. Martinovic, W. Müller-Wittig y T. K. Y. Chan (2004). Digital Heritage Application as an Edutainment Tool.
- SourceForge. (2007). "AMIRE." Visto el 20 Febrero, 2008, en <http://amire.sourceforge.net>.
- Sparacino, F. (2004). Scenographies of the Past and Museums of the Future: From the Wunderkammer to Body-Driven Interactive Narrative Spaces. (MM'04).
- Stanney, K. M. (2002). Handbook of Virtual Environments : Design, Implementation, and Applications, Mahwah, NJ, USA: Lawrence Erlbaum Associates.
- Stapleton, C., C. Hughes y J. M. Moshell (2002). Mixed Reality and the Interactive Imagination: Adding the art to the science and technology of Mixed Reality for training, education and entertainment. Swedish-American Workshop on Modeling and Simulation (SAWMAS'02), Orlando, FL.
- Stauder, J. (1999). "Augmented Reality with Automatic Illumination Control Incorporating Ellipsoidal Models." IEEE TRANSACTIONS ON MULTIMEDIA **1**(2): 136 - 143.
- Stetten, G., V. Chib, D. Hildebrand y J. Bursee (2001). Real Time Tomographic Reflection: Phantoms for Calibration and Biopsy. Augmented Reality, IEEE and ACM.
- Stilman, M., P. Michel, J. Chestnutt, K. Nishiwaki, S. Kagami y J. Kuffner (2005). Augmented Reality for Robot Development and Experimentation, Robotics Institute, Carnegie Mellon University.
- Sutherland, I. E. (1965). "The Ultimate Display." Visto el 12 Noviembre, 2007, en http://www.cise.ufl.edu/~lok/teaching/dcvef05/papers/ultimate_display.html.
- Szalavári, Z., E. Eckstein y M. Gervautz (1998). Collaborative Gaming in Augmented Reality. VRST'98, Taipei (Taiwan), ACM.
- Szalavári, Z. y M. Gervautz (1997). "The Personal Interaction Panel a Two-Handed Interface for Augmented Reality." Computer Graphics Forum **16**(3).
- Tappert, C. C., A. S. Ruocco, K. A. Langdorf, F. J. Mabry, K. J. Heineman, T. A. Brick, D. M. Cross, S. V. Pellissier y R. C. Kaste (2001). Military Applications of Wearable Computers and Augmented Reality. Fundamentals of wearable computers and augmented reality. W. Barfield y T. Caudell. Mahwah, New Jersey, Lawrence Erlbaum Associates: 625-648.
- Tatham, E. W. (1999). "Optical Occlusion and Shadows in a 'See-through' Augmented Reality Display." Information Visualization.

- Tavares, D., A. Burlamaqui, A. Dias, M. Monteiro, V. Antunes, G. Thó, T. Tavares, C. Lima, L. Gonçalves, G. Lemos, P. Alsina y A. Medeiros (2003). Hyperpresence - an application environment for control of multi-user agents in mixed reality spaces. Annual Simulation Symposium (ANSS'03), IEEE Computer Society.
- Technology Research News. (2000). "Pop-up books meld real with virtual." Visto el 1 Febrero, 2008, en http://www.trnmag.com/Stories/100400/Magic_Book_100400.htm.
- TechTarget. (2008). "Whatis?com, Computer Dictionary." Visto el 31 Enero, 2008, en <http://whatis.techtarget.com/>.
- Teichrieb, V., S. G. Neto, T. Farias, J. M. Teixeir, J. P. Lima, G. Almeida y J. Kelner (2007). Augmented Ambient: An Interactive Mobility Scenario. Human-Computer Interaction (CHI'07), Beijing, China.
- Torpus, J. (2004). "Living Room." Visto el 5 Octubre, 2007, en <http://projekte.idk.ch/livingroom/>.
- Torpus, J. (2005). "lifeClipper." Visto el 5 Octubre, 2007, en <http://www.torpus.com/lifeclipper/>.
- Trivisio. (2008). "Trivisio Prototyping GmbH." Visto el 7 Febrero, 2008, en <http://www.trivisio.com/>.
- Tufte, E. R. (1990). Envisioning Information. Cheshire, Connecticut, Graphics Press.
- Update Centre. (2007). "Virtual Reality. Encyclopædia Britannica Article." Visto el 10 diciembre, 2007, en <http://updatecenter.britannica.com/eb/article?articleId=1382>.
- Villanueva, E. (1995). Las Personas. Méjico, Centro de Neurobiología – UNAM.
- Vlahakis, V., J. Karigiannis, M. Tsotros, M. Gounaris, L. Almeida, D. Stricker, T. Gleue, I. T. Christou, R. Carlucci y N. Ioannidis (2001). ARCHEOGUIDE: First Results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites.
- Vogiazou, Y., B. Raijmakers, E. Geelhoed, J. Reid y M. Eisenstadt (2007). "Design for emergence: experiments with a mixed reality urban playground game." Pers Ubiquit Comput **11**: 45 - 58.
- vvvv group. (2008). "vvvv: a multipurpose toolkit." Visto el 8 Febrero, 2008, en <http://vvvv.org/tiki-index.php>.
- Wacker, F. K., S. Vogt, A. Khamene, F. Sauer, M. Wendt, J. L. Duerk, J. S. Lewin y K. J. Wolf (2005). "MR image-guided needle biopsies with a combination of augmented reality and MRI: A pilot study in phantoms and animals." International Congress Series **1281**: 424-428.
- Wagner, D. (2007a). "Handheld Augmented Reality. Augmented Reality on Truly Mobile Devices: ARToolKitPlus 2.1.1." Visto el 12 Noviembre, 2007, en http://studierstube.icg.tu-graz.ac.at/handheld_ar/artoolkitplus.php.
- Wagner, D. (2007b). "Handheld Augmented Reality. Augmented Reality on Truly Mobile Devices: Studierstube ES (StB ES)." Visto el 12 Noviembre, 2007, en http://studierstube.icg.tu-graz.ac.at/handheld_ar/stbes.php.
- Wagner, D. (2007c). "Handheld Augmented Reality. Augmented Reality on Truly Mobile Devices: Studierstube Tracker." Visto el 12 Noviembre, 2007, en http://studierstube.icg.tu-graz.ac.at/handheld_ar/stbtracker.php.

- Wagner, D., T. Pintaric y D. Schmalstieg. (2004a). "The Invisible Train: A Multi-Player Handheld Augmented Reality Game." Visto el 12 Noviembre, 2007, en http://studierstube.icg.tu-graz.ac.at/invisible_train/.
- Wagner, D. y D. Schmalstieg (2007). ARToolKitPlus for Pose Tracking on Mobile Devices. Computer Vision Winter Workshop, St. Lambrecht, Austria.
- Wagner, M., A. MacWilliams, M. Bauer, G. Klinker, J. Newman, T. Pintaric y D. Schmalstieg (2004b). Fundamentals of Ubiquitous Tracking. Viena, TU-Wien: 6.
- Walczak, K. y R. Wojciechowski (2005). Dynamic Creation of Interactive Mixed Reality Presentations. (VRST'05). Monterey, California, USA, ACM: 167 - 176.
- Walker Art Center. (2001). "Collections and resources." Visto el 7 Marzo, 2008, en http://collections.walkerart.org/item/enlarge_fs.html?id=5&type=event&image_num=3.
- Wang, X. y P. S. Dunston (2006). "Compatibility issues in Augmented Reality systems for AEC: An experimental prototype study." Automation in Construction **15**(3): 314-326.
- Wang, Y. y D. Samaras (2003). "Estimation of multiple directional light sources for synthesis of augmented reality images." Graphical Models **65**(4): 185-205.
- Wanschitz, F., W. Birkfellner, M. Figl, S. Patruta, A. Wagner, F. Watzinger, K. Yerit, K. Schicho, R. Hanel, F. Kainberger, H. Imhof, H. Bergmann y R. Ewers (2002). "Computer-enhanced stereoscopic vision in a head-mounted display for oral implant surgery." Clinical Oral Implants Research **13**(6).
- Weidenbach, M., C. Wick, S. Pieper, K. J. Quast, T. Fox, G. Grunst y D. A. Redel (2000). "Augmented Reality Simulator for Training in Two-Dimensional Echocardiography." Computers and Biomedical Research **33**(1): 11-22.
- Wikipedia. (2007a). "Edutainment." Visto el 20 Enero, 2007, en <http://en.wikipedia.org/wiki/Edutainment>.
- Wikipedia. (2007b). "GLSL." Visto el 02 Noviembre, 2007, en <http://es.wikipedia.org/wiki/GLSL>.
- Wikipedia. (2007c). "Human Computer-Interaction." Visto el 02 Noviembre, 2007, en http://en.wikipedia.org/wiki/Human-computer_interaction.
- Wikipedia. (2008a). "Mobile RFID." Visto el 20 Enero, 2008, en http://en.wikipedia.org/wiki/Mobile_RFID.
- Wikipedia. (2008b). "RFID." Visto el 20 Enero, 2008, en <http://es.wikipedia.org/wiki/RFID>.
- Wikipedia. (2008c). "The Sims." Visto el 12 Febrero, 2008, en http://en.wikipedia.org/wiki/The_Sims.
- Wikipedia. (2008d). "The Sims Online." Visto el 12 Febrero, 2008, en http://en.wikipedia.org/wiki/The_Sims_Online.
- Wilde, D., E. Harris, Y. Rogers y C. Randell (2003). "The Periscope: Supporting a Computer Enhanced Field Trip for Children." Pers Ubiquit Comput **7**: 227-233.
- Wojciechowski, R., K. Walczak, M. White y W. Cellary (2004). Building Virtual and Augmented Reality Museum Exhibitions, ACM: 135 - 144.

- Woods, E., M. Billingham, G. Aldridge y B. Garrie (2004). Augmenting the Science Centre and Museum Experience, ACM: 230 - 236.
- Xsens. (2007). "Xsens Motion Technologies." Visto el 2 Febrero, 2007, en <http://www.xsens.com/>.
- Yamada, T., S. Yokoyama, T. Tanikawa, K. Hirota y M. Hirose (2006). Wearable Olfactory Display: Using Odor in Outdoor Environment. Virtual Reality, IEEE Computer Society Washington, DC, USA.
- Yanagida, Y., S. Kawato, H. Noma, A. Tomono y N. Tetsutani (2004). Projection-Based Olfactory Display with Nose Tracking. Virtual Reality 2004 (VR'04), IEEE Computer Society Washington, DC, USA.
- Zhou, Z., A. D. Cheok, X. Yang y Y. Qiu (2004). "An experimental study on the role of 3D sound in augmented reality environment." Interacting with Computers **16**(6): 1043-1068.
- Zlatanova, S. (2001). 3D modelling for augmented reality. Dynamic and Multi-dimensional GIS. Bangkok (Thailand).
- Zlatanova, S. (2002). Augmented Reality Technology. Delft, Section GIS technology, Faculty of Civil Engineering and Geosciences, TU Delft: 77.

ANEXOS

Publicaciones

Giner Martínez, F. y C. Portalés Ricart (2005). The Augmented User: A Wearable Augmented Reality Interface. International Conference on Virtual Systems and Multimedia (VSMM'05) Hal Thwaites.

Giner Martínez, F. y C. Portalés Ricart (2005). El Túnel Mágico. Nuevos materiales y tecnologías para el arte, Madrid (España), Universidad Complutense de Madrid.

Giner Martínez, F. y C. Portalés Ricart (2007). Augmented teaching. International Technology, Education and Development Conference (INTED'07), Valencia (España), L. Gómez Chova, D. Martí Belenguer, I. Candel Torres. International Association of Technology, Education and Development. IATED, Valencia.

Portalés Ricart, C. (2007). "Live LEGO House: a Mixed Reality Game for the Edutainment." Schottish Online Journal of E-Learning **1**(1): 19 - 28.

Portalés Ricart, C. (2007). Live LEGO House: an interactive space to explore coexistence through gaming. International Technology, Education and Development Conference (INTED'07), Valencia (España), L. Gómez Chova, D. Martí Belenguer, I. Candel Torres. International Association of Technology, Education and Development. IATED, Valencia.

Portalés Ricart, C. (2007). The private city through the Hot Images. International Conference on Advances in Computer Entertainment Technology (ACE'07), Salzburg (Austria), ACM SIGCHI.

Portalés Ricart, C., F. Giner Martínez y F. Sanmartín Piquer (2005). Back to the 70's. International Conference on Advances in Computer Entertainment Technology (ACE'05), Valencia (España), ACM SIGCHI.

Portalés Ricart, C., F. Giner Martínez y F. Sanmartín Piquer (2005). Urbanmix. Nuevos materiales y tecnologías para el arte, Madrid (España), Universidad Complutense de Madrid.

Portalés Ricart, C., C. D. Perales Cejudo y A. Cheok (2007). Exploring Social, Cultural and Pedagogical Issues in AR-Gaming Through The Live LEGO House. International Conference on Advances in Computer Entertainment Technology (ACE'07), Salzburg (Austria), ACM SIGCHI.

Aceptados para publicar

Portalés Ricart, C. (2006). The Private City Through the Hot Images. ENGAGE, Sydney (Australia).

The Augmented User: a Wearable Augmented Reality Interface

Francisco GINER, Cristina PORTALÉS

Camino de Vera s/n, Laboratorio de Luz, BB.AA., Univ. Politécnica de Valencia, Spain

fginerm@pin.upv.es, criporri@doctor.upv.es

Abstract. In this paper, 'The Augmented User' project is presented, a new Augmented Reality (AR) system based on the ARToolKit and MRXToolKit technologies. We have developed an innovative system in which the user is visible in the AR scenario and indispensable part of the final result: he/she is in fact the carrier of the markers that are recognized by the system allowing his/her own transformation; this opens the possibilities of the users that are immersed in an AR system, being at the same time the augmented matter and the spectator. We also introduce two interesting applications for our system: the 'changing personalities' game and the 'image carousel' performance, as simple examples of the multiple applications that our system actually have. The collaborative part is also proposed and presented, which makes the applications even more interesting. Finally, some conclusions are presented and further work is proposed.

1. Introduction

The Augmented User is an AR application based on the well known ARToolKit [2] and MXRToolKit [9] libraries; these libraries are indeed the base of a wide range of AR applications [1, 3, 4, 5, 6, 7, 8, 10, 11, 12]. Most of these applications are applied to visualize elements virtually inserted in a pre-existing physical context. In our system we try to relativize some of the dualities or binary structures that arise in these kinds of applications regarding to the perception, for example:

- Relationship user/object: the user is part of the object¹.
- Inside/out: user incorporation inside the representation space.
- Refers to the 'mirror stage'² by changing user identity.

To deal with these dualities, the spatial distribution of some basic elements is inverted:

- Normally the user establishes identification between the display and his/her vision, identification increased by the physical location of the HMD (Head Mounted Display) which is placed at the height of his/her eyes. In our application the mechanical eye, i.e., the camera is far away from the user.
- It is also common that the user equipped with the HMD observes markers (to achieve user pose) as an exterior object. In our case, the markers form part of the user. The user acts, sees, thinks and feels at the same time the AR experience.
- The projected image works as a mirror whose reflex is viewed by the user and others, thus allowing a collaborative experience. Furthermore, within a single AR scenario, several markers wear by different users can be tracked, thus allowing

¹ 'Be-like-object for the Other'. Jean Paul Sartre

² 'The stage of the training mirror like of the function of I (je)' [1949] Jacques Lacan (Written 1)

several augmented users that may or may not be together with other non-augmented users.

Our application is generically approximated to the contraposition: ‘to be is to perceive (esse is percipere) / to be is being perceived (esse est percipi)’³, affecting also in questions of identity, avatar and playing.

In the following sections, the camera pose, our AR system and the developed marker prototypes are explained. Some examples are shown, including a collaborative AR scenario with multiple users, whose experiences are also collected. Finally, some conclusions are given and further work is proposed.

2. Camera pose tracking

Normally, in AR applications the camera pose (position and orientation) corresponds to user’s head pose, so the camera point of view is considered to coincide with that of the user. In our case, the camera is still and far away from the user.

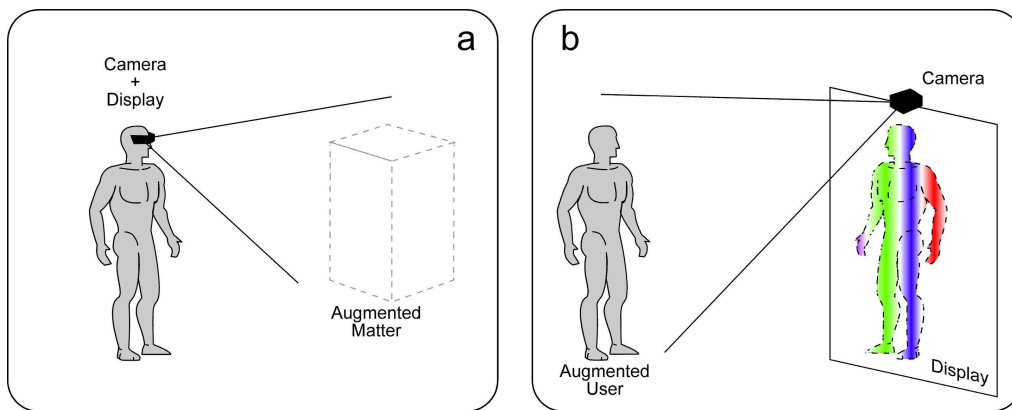


Figure 1. Relation between user and camera:
a) in traditional AR system; b) in the Augmented User system.

Our interest yields in acquiring the relative pose between the markers and the camera. Camera pose tracking is one of the most crucial parts in AR applications [5]. This consists on determining the 6 outer orientation parameters of the camera, i.e., three translations and three rotations. In our application, this task was achieved with the ARToolkit library, which was developed at the HIT Lab of the University of Washington. This technology provides the necessary computer vision techniques required to track some kind of markers, and accurately calculate the camera pose relative to them. Virtual objects are then accurately posed – relative to the markers – in the ‘real’ world. We also used the MXRToolkit developed at the MXR Lab in Singapore, which consists of a library of routines to help with all aspects of building mixed reality applications.

3. System

Our system consists of:

- The physical wearable markers,
- an overhead screen,
- a DV-Camera, connected by an IEEE 1394 cable to the CPU,
- a computer graphics workstation,

³ By G. Berkeley.

- two (or more) potent and controlled focuses,
- ARToolKit and/or MXRToolkit libraries.

The markers are carried by the user and are linked to some 3D virtual models. The virtual models are in the VRML format and, for some applications, are mapped with images. The overhead screen has a dimension of 2.5x1.5 m, and is made of a translucent white plastic. The dimension of the screen and the distance to the projector should be taken into account in order to achieve a representation scale similar to the real one, thus incrementing the mirror effect. The user is situated within a distance of 1-3 m to the screen and at the same height. The camera is situated at the very top of the screen pointing to the user. Two focuses are needed to properly illuminate the user for the camera to track the markers; the light should be directed in such a way that the screen itself remains in half-light. In the following figure, the spatial situation of all devices and the user can be seen.

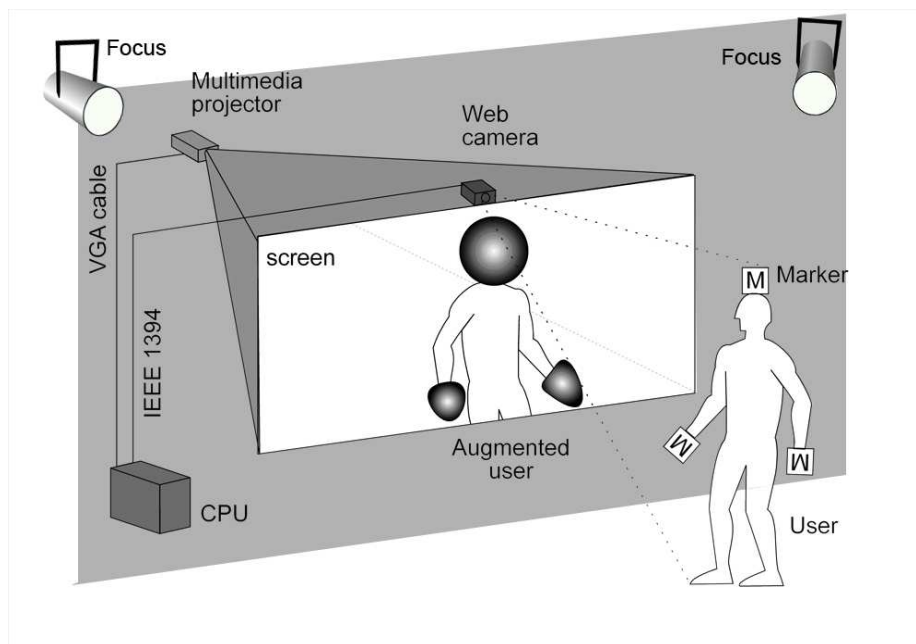


Figure 2. System scheme.

4. Marker Prototypes

Several prototypes were made. At the beginning, basic virtual shapes were tested with the user carrying a simple marker on his/her hands (figure 3).

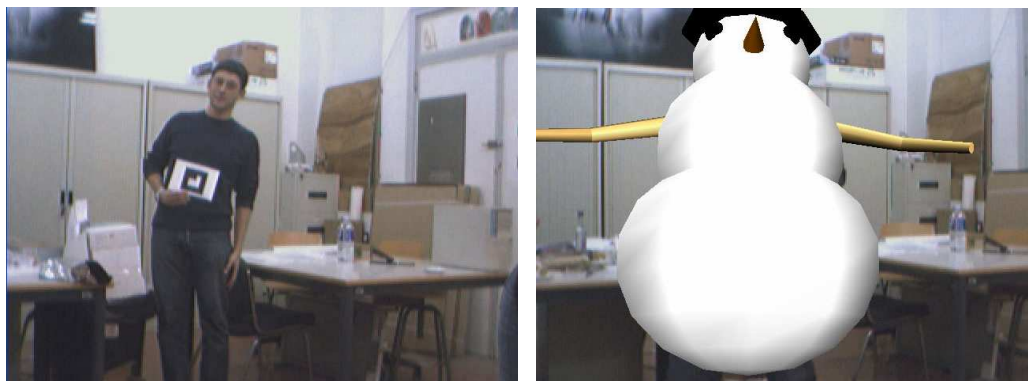


Figure 3. Simple marker and user as a snowman.

When the user moves with the marker, the virtual object acquires the same movement, thus allowing a playful experience. The first impression acquired by the users was very positive, as they really enjoyed the fact of being transformed into a new character that followed their movements. Nevertheless, the users found that carrying a marker was not comfortable, as they wanted to have their hands free. We found another problem with this system, that the users need to handle the marker at a certain position, neither lower nor higher, for the virtual object to appear in the correct place. We thought about the possibility of doing a vest-marker for the marker to be fixed at a certain location, but somehow we needed several vest-sizes to allow all kind of users to wear it. Another problem found was the fact that, within the natural user's motion sometimes his/her arms partially hid the markers, thus losing the virtual object at that time. Therefore, the idea of making a vest-marker was abandoned.

After thinking about the best suited place for the user to wear a marker, we concluded that the best one was on his/her head, which is the highest point of the human being, solving the occasional problem of the arms hiding the marker. We decided to make a hat in such a way that it could be wearable by all users. So the hat itself is mainly made of wool, which makes it to be elastic. This hat is composed by a set of three markers, forming a prism with a triangular base in such a way that the virtual object will never be loosed although the user rotates, allowing a more motion freedom.



Figure 4. User wearing the prismatic hat; Man-to-Lady.

A posterior and natural evolution of our hat marker is the hair band marker, with both reduced dimension and weight, and more comfortable for the user to wear. Additionally, we also modelled markers for the user hands, thus allowing the incorporation of new virtual objects for different parts of the user body. Each one of the virtual models is independently animated by the user, thus incrementing gaming possibilities. It should be pointed that reducing the marker dimensions implies the need of both a better camera resolution and illumination, thus marker dimension is always dependent on our system specifications. In the next figures the last prototypes can be seen.

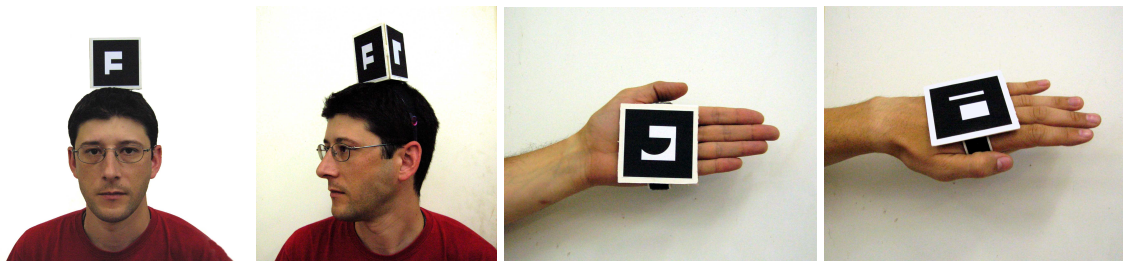


Figure 5. User wearing the hair band marker and the hand markers.

The wearable markers are indeed an inexpensive system for the AR applications. Furthermore, the markers are extremely light to wear. These two facts allow this system to be used by everybody, including children, with no fear to break normally expensive vision interfaces (as HMD, hand-held screens, laptops, etc.) or pose sensors (as inertial sensors, GPS, etc.). In [4] the user wears a glove, with a marker attached; nevertheless, in this case

the information acquired from the marker is used to navigate within the virtual world. Therefore, we can say our system is innovative.

5. Changing personalities

For this application we modelled different characters (a lady, a young child, Mr. Potato, etc.) in the VRML format. The users see themselves with the new personality projected into the overhead screen that acts as a mirror. The projected image is rendered in real time, allowing a visual dialog in which the user can be introduced very easily. In the following figures we can see the user transformed in Mr. Potato. He is wearing the hair band marker that is linked to the Mr. Potato's head, and the right and left hand markers that are linked to Mr. Potato's right and left hands, respectively. This makes the user to experience a personality change, which is not only seen by other users, but principally by him/herself from the exterior of his/her own augmented body; the user is an interactive spectator and an indispensable part of the application for being the matter of itself, which is in continuous transformation.

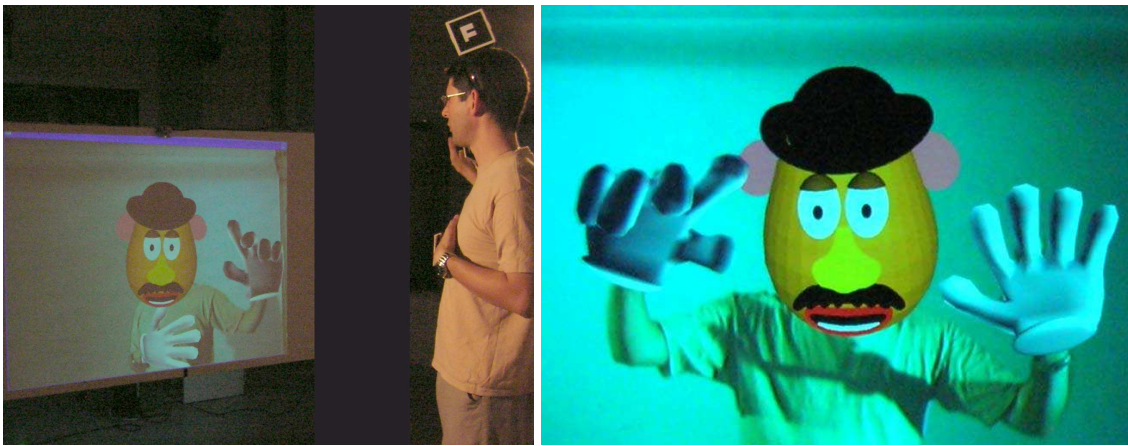


Figure 6. User transformed in Mr. Potato.

It should be pointed that one of the well-known drawbacks of using marker-board tracking is the presence of the marker itself, which tends to look unnatural. In our application this problem disappears, as the markers are always hidden by the virtual objects, thus avoiding them to appear in the projected image. The model scale with respect to the user is exaggerated; therefore the playful and/or cartoon effect is increased.

6. Image carrousel

This application consists of a set of images disposed in several bands that are animated, thus rotating around the user. This virtual model is also in the VRML format and the application runs under the MXRToolKit libraries. The mapped images can show past events, far places, personal and/or private spaces, user's relatives, etc. Therefore, the image content can have a thematic, expositive, narrative and/or random character. We can also establish a relationship between past, present and future, close and far places, known and unknown, etc. Images are mapped on a VRML animated model consisting of a set of planar surfaces disposed in different bands. These bands can intercept each other at a certain point, showing an interesting visual effect similar to that of a kaleidoscope (figure 7).

The virtual model can be modified by changing its rotation velocity, number of planar surfaces, number and tilting of the bands, intersections, etc. For the same virtual model,

different images can be linked. Very interesting is the sum of both model and user movements.

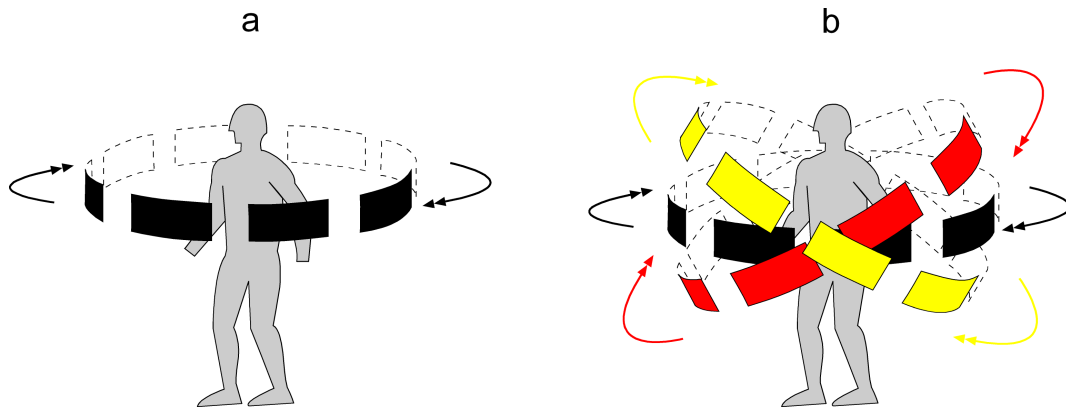


Figure 7. Carrousel model scheme:
a) with only one band; b) with three bands intersecting each other.

In this application we faced the occlusion problem, i.e., the non-desirable effect when a real object (in our case, the user) partially should hide the virtual one. We solved this problem in a very easy way, just giving mapping information to the faces of the planar surfaces that point outwards the user, thus the faces pointing to the user remains transparent, and are not seen in the projected image when being backward to the user. In the next figure, we can see the user with a thematic Image Carrousel application, showing some images of Ghent.



Figure 8. Image Carrousel with images of Ghent (Belgium)

7. Collaborative Scenario

Multiple users can immerse in each one of the applications at the same time. Thus, a collaborative space is created. Within this new space, each one of the users is wearing a set of markers, related to different virtual objects, but with some kind of relation between them. For example, in the ‘changing personalities’ application, the heads of couples of famous characters were modeled; these models are in fact simple quasi-spherical shapes mapped with images, which are very easily interpreted by the user. This specific application remits to some artistic expressions of Tony Oursler⁴, who projects videos of human faces on some kind of heads acting like screens of inanimate dolls in a non habitual positions and scales, adopting finally these pieces a similar aesthetics to ours, and provoking the reflection of the spectator; in our case, the user gives ‘live’ with his/her own movement to the augmented hybrid in which he/she is converted. We do not discard to map the heads with some videos instead of images, which will allow a major interaction.

To map the models, firstly we focused our attention on legendary Hollywood couples, as Jack Lemmon & Walther Matthau, Spencer Tracy & Katharine Hepburn, C3PO & R2D2 or

⁴ <http://www.tonyoursler.com/>

Fred Aster & Ginger Rogers. Therefore, user 1 could be Fred Aster and user 2 Ginger Rogers. They both could dance as in the legendary *Shall we Dance* movie. We also thought about couples than can never be together or never ever met, but have lot in common, as Isaac Newton & Albert Einstein, Wolfgang Amadeus Mozart & Ludwig van Beethoven, Hergé & Tintin or Dr. Jekyll & Mr. Hyde. Could you imagine Mozart and Beethoven talking about a Symphony?, or Dr. Jekyll facing with Mr. Hyde? What would ask Titin to his creator, Hergé?. In the following figures, two users transformed into Jack Lemmon & Walther Matthau are playing together.



Figure 9. Different users playing with their new identities: Jack Lemmon & Walther Matthau; Laurel & Hardy; Mr. & Mrs. Potato

For the ‘image carousel’, an augmented collaborative application was created in which images from three different spaces were shown. User 1 could see the images from a trip to London (city), user 2 from a trip to Ordesa (field), and user 3 some personal/private (home) images. In the next figure the three different models can be seen.

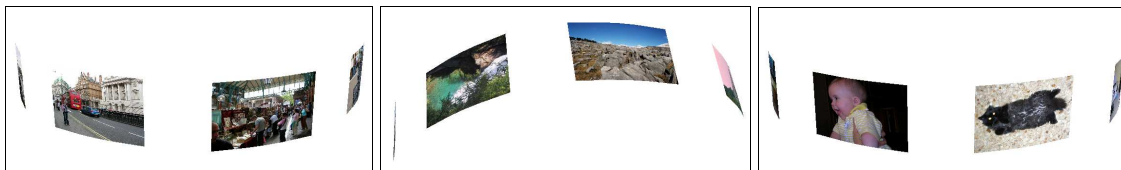


Figure 10. Different images for a collaborative image carousel application showing three different spaces: city, field and home

8. User experience

User response with both ‘changing personalities’ and ‘image carousel’ applications has been very positive. A total of 17 men and 17 women ranging from 7 to 60 years old tested our system; after the experience all the users fulfilled a short test. We asked them to score between 1 (bad) and 10 (excellent) a total of 9 different facts of our system. The facts/questions to be scored were:

1. Interaction
2. Fastness of the system (regarding to the user movements)
3. Design and comfort of the wearable marks
4. The idea of the mirror as an interface
5. 3D virtual models
6. Gaming possibilities
7. Learning possibilities
8. Therapeutic possibilities
9. Gallery/Museum possibilities

In figure 11 the mean scores are shown. As it can be seen all the scores are above 6.5 and, in general, women gave better punctuations; the fact that was better scored by women was the question 6 (9.4 points) and by men the questions 4 and 5 (8.1 points).

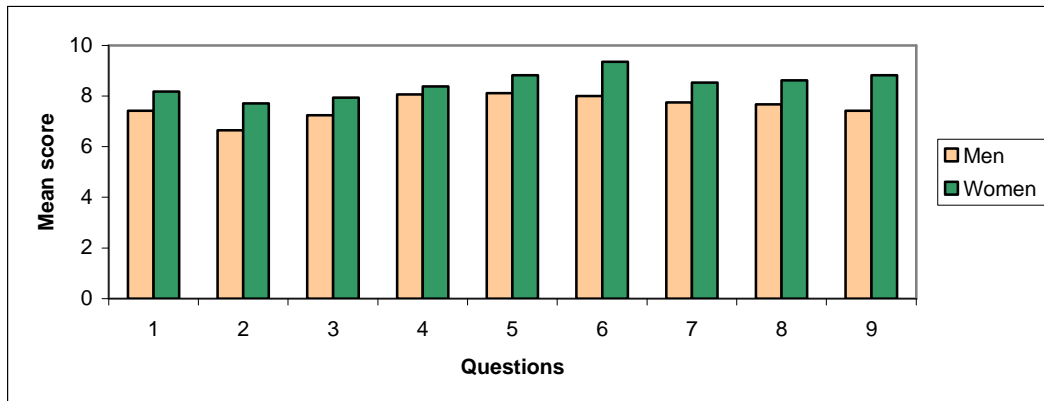


Figure 11. Scores given by the users to 9 different questions about our system

We also asked the users to write a short conclusion about the experience. In general, they empathised the fact of the gaming possibilities, and gave us some new ideas to improve our system or use it in a different scenario. Some of the comments were: “I enjoyed myself very much. It could be interesting to add some kind of music related to the new virtual personality”; “A new virtual scenario could be created”; “Improve the system velocity and use bigger mirrors”; “Add more virtual objects and pay more attention in the learning possibilities”. In general, the applications that people enjoyed the most were the collaborative ones.

9. Conclusion

In this paper we have presented the Augmented User, an innovative AR application where the whole system is inverted: the camera is far from the user who is augmented by some kind of wearable markers. The user interacts with the virtual world with his/her own movements, being at the same time the ‘user’ and/or ‘matter’ of the application and the ‘spectator’, being transformed and acquiring a new identity inside the already-known real space.

The identification of the user with the interface is immediate since he/she recognizes him/herself and realizes about the gaming possibilities, those that permits his/her own body by the movement, since the system does not have physical limitations beyond the space in which is developed, i.e., the prepared AR scenario. The movements of every user are different and particular; the way of reacting when turning into an augmented user is also different depending on the connotations of every transformation and what they mean for every user, making the application to be different, unique and unrepeatable regarding to personal feelings and sensorial perception. The collaborative experiences allow a more playful game and interesting interaction between different augmented users within a shared space.

Regarding to the whole system, it should be pointed that the facility of programming allows changes and variations with very little cost of time and material resources, which makes possible a rapid adequacy and update of the application to new spaces and needs. Moreover, the rapid and easy evolution that we are experimenting makes us think about a very stable system of high interaction in real time, of easy use and with multiple possible ways of development, including different fields of application, from playful to documentary or therapeutic.

10. Further work

We are working in new marker prototypes that are auto illuminated, in such a way that the dependency of external luminous devices disappears. Furthermore, the incorporation of a new camera with more resolution and a greater angular will allow us to decrease the markers dimensions thus allowing the user to experiment a more freedom sensation. A further incorporation of video images to map the models, or even, some kind of in-real-time video, could increase the conceptual potential of our applications.

Finally, in the presented collaborative applications, it could be interesting that users could really face each other, thus allowing an AR face-to-face collaboration [7]. Within our currently interface this is not possible, because both users have to look to the same screen to see the virtual objects. Therefore, we are working in order to enrich our interface. One of our ideas is an interface using three screens placed in orthogonal positions (figure 12), in such a way that the users could always see each other, although they look in different directions.

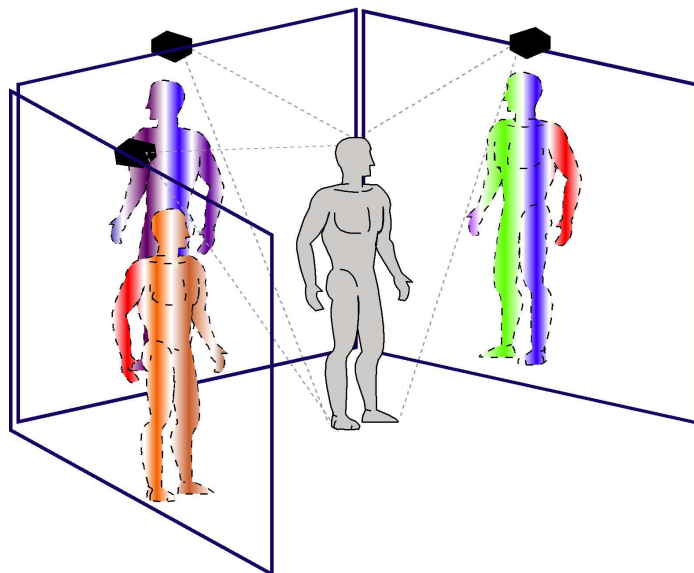


Figure 12. Interface prototype

References

- [1] A. D. Cheok, W. Weihua, X. Yang, S. Prince, F. S. Wan, M. Billinghurst and H. Kato, Interactive Theatre Experience in Embodied + Wearable Mixed Reality Space., *International Symposium on Mixed and Augmented Reality (ISMAR 2002)*, ISBN: 0-7695-1781-1, 2002, pp. 59-317.
- [2] ARToolKit: <http://www.hitl.washington.edu/artoolkit/>
- [3] D. Díaz and C. Boj, Red Libre Red Visible, Workshop: Taller de realidad aumentada y redes inalámbricas, 15-18 February 2005.
- [4] E. Woods, P. Mason and M. Billinghurst, MagicMouse: an Inexpensive 6-Degree-of-Freedom Mouse, *Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia*. ISBN:1-58113-578-5, ACM Press, NY, USA, 2003, pp. 285-286.
- [5] G. Gordon, M. Billinghurst, M. Bell, J. Woodfill, B. Kowalik, A. Erendi and J. Tilander, The Use of Dense Stereo Range Data in Augmented Reality, *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR'02)*, ISBN: 0-7695-1781-1, IEEE Computer Society, Washington, DC, USA, 2002, pp. 14-23.
- [6] H. Kato, M. Billinghurst, I. Pourpyrev, K. Imamoto and K. Tachibana, Virtual Object Manipulation on a Table-Top AR Environment, *IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*. ISBN: 0-7695-0846-4, Munich, Germany, October 05-06, 2000, pp. 111-119.

- [7] M. Billinghurst and H. Kato, Collaborative Augmented Reality. *Communications of the ACM, How the virtual inspires the real*, Volume 45 (7), ISSN:0001-0782, ACM Press, New York, NY, USA, pp. 64-70.
- [8] M. Billinghurst, K. Hirakazu and I. Poupyrev, The MagicBook– Moving Seamlessly between Reality and Virtuality, *IEEE Computer Graphics and Applications*, Volume 21 (3), ISSN:0272-1716. IEEE Computer Society Press, Los Alamitos, CA, USA, 2001, pp. 6-8.
- [9] MXRToolKit: <http://mxrtoolkit.sourceforge.net/>
- [10] N. R. Hedley, M. Billinghurst, L. Postner and R. May, Explorations in the use of Augmented Reality for Geographic Visualization, *Presence: Teleoperators and Virtual Environments*, Volume 11 (2), ISSN: 1054-7460, MIT Press, Cambridge, MA, USA, 2002, pp. 119-133
- [11] P. A. S. Sinclair and K. Martinez, Tangible Hypermedia using ARToolKit, *Augmented Reality Toolkit, The First IEEE International Workshop*, ISBN: 0-7803-7680-3, 2002, 2 pp.-.
- [12] P. A. S. Sinclair, K. Martinez, D. E. Millard and M. J. Weal, Links in the Palm of your Hand: Tangible Hypermedia using Augmented Reality, *Proceedings of the thirteenth ACM conference on Hypertext and hypermedia*. ISBN:1-58113-477-0, ACM Press, New York , NY, USA, 2002, pp. 127-136.

EL TÚNEL MÁGICO

GINER MARTÍNEZ, Francisco. Laboratorio de Luz, BB. AA., Universidad Politécnica de Valencia, camino de Vera s/n 46022, Valencia. fginerm@pin.upv.es

PORTALÉS RICART, Cristina. Laboratorio de Luz, BB. AA., Universidad Politécnica de Valencia, camino de Vera s/n 46022, Valencia. criporri@doctor.upv.es

Palabras clave: realidad aumentada, sistema de marcas, interactividad.

1. INTRODUCCIÓN

“Realidad aumentada”, se puede definir como la amplificación de la capacidad sensorial de la percepción del mundo que nos rodea mediante el uso de dispositivos que superponen elementos virtuales a la imagen real.

El Túnel Mágico es una aplicación de realidad aumentada basada en las librerías de ARToolKit [1]; estas librerías son de libre distribución y se han utilizado en numerosas aplicaciones. La mayoría de estas aplicaciones tratan de mezclar mundos y/o escenarios nuevos, insertando elementos virtuales dentro de la realidad que nos rodea. En [3] se puede encontrar una descripción y clasificación de diferentes sistemas de realidad aumentada y su evolución en el tiempo.

El Túnel Mágico consta de una serie de pantallas a modo de espejo con las que diversos usuarios pueden interactuar al verse transformados en un personaje pseudo-virtual. Estas pantallas están dispuestas linealmente, a modo de túnel, y el usuario se puede mover por el escenario y así verse reflejado en las distintas pantallas, cada una de ellas aportándole una nueva identidad.

En nuestro sistema, tratamos de relativizar alguna de las dualidades o estructuras binarias que conllevan este tipo de aplicaciones en términos de percepción, como por ejemplo:

- Relación usuario/objeto: el usuario es parte del objeto.
- Dentro/fuera: incorporación del usuario en el espacio de representación.
- Referencia al ‘estadio del espejo’¹, cambiando la identidad del usuario.

Para ello se invierte la disposición espacial de algunos elementos básicos:

- Normalmente el usuario de aplicaciones de realidad Aumentada establece una identificación entre el display y su visión, identificación incrementada por la localización física de las gafas o HMD (Head Mounted Display) a la altura de sus ojos. En este ensayo el ojo maquinal, la cámara DV está alejado del sujeto.
- También normalmente el usuario equipado con las gafas observa la(s) marca(s) de realidad aumentada como objeto exterior. En el ensayo la marca forma parte del sujeto.
- La imagen proyectada funciona a modo de espejo, cuyo "reflejo" es visible a este usuario y otros, la experiencia más que multiusuario es compartida.

¹ Lacan

Este ensayo se aproxima de forma genérica a la contraposición: "ser es percibir (esse es percipere) / ser es ser percibido (esse est percipi)"², incidiendo también en cuestiones de identidad, avatar y juego.

El Túnel Mágico es un proyecto en continuo proceso, en el cual pretendemos crear un sistema propio de realidad aumentada que, principalmente, se puede encuadrar dentro de las áreas de diversión, narración y exposición, no descartando en un futuro desarrollarlo para su uso en la educación o, incluso, para fines terapéuticos.

En los siguientes apartados, se describe con detalle los componentes y el funcionamiento de nuestro sistema de realidad aumentada. Así mismo, se dedica una sección para narrar las experiencias de los usuarios en el Túnel Mágico y se apuntan las pautas a seguir para seguir enriqueciendo nuestro sistema.

2. ARTOOLKIT: ANTECEDENTES

Como ya hemos apuntado en la introducción, el ARToolKit es un software de libre distribución, originalmente creado por el HIT Lab de la Universidad de Washington, y cuyas librerías se encargan principalmente de calcular la posición y orientación del usuario con respecto a una marca determinada, que es reconocible por el sistema a partir de técnicas de visión por computador; posteriormente, los objetos virtuales se posicionan correctamente con respecto a estas marcas. Este hecho lo hace especialmente idóneo para aplicaciones de realidad aumentada, puesto que, en este tipo de aplicaciones, lo que se pretende es mezclar la realidad con la ficción, es decir, añadir objetos virtuales (creados por ordenador) en nuestro entorno real; es por ello que la realidad debe casar con la ficción en términos de "posición/orientación" (el objeto virtual ocupará un espacio real dentro del escenario), y para ello, el sistema debe conocer la posición y orientación espacial del sujeto o cámara dentro del entorno real [7; 9]. Muchas aplicaciones se han venido desarrollando en los últimos años. Por ejemplo, en [4] se muestra un sistema de teatro interactivo, en el que el usuario carga con ordenador a modo de vestimenta, y las marcas se utilizan para posicionar dentro del escenario a diversos actores; los autores de [8] proponen un sistema de marcas para visualizar e interactuar con datos geográficos; en [10] y [11] se muestra un sistema hipermedia tangible de realidad aumentada; en [2] se propone un libro interactivo de realidad aumentada; el sistema desarrollado en [5], propone insertar las marcas por la ciudad, a modo de paneles o pancartas, con el fin de crear nuevos paisajes en el espacio público a través de la visualización de los datos que fluyen por las redes digitales.

A diferencia de las anteriormente citadas de aplicaciones de realidad aumentada, en las que la cámara se corresponde con el ojo del espectador (o punto de vista), en el Túnel Mágico el sistema se invierte, y la cámara permanece estática y alejada del usuario, siendo las marcas las que se mueven. Esta idea de invertir el sistema surgió poco tiempo atrás, concepto que denominamos como "Usuario Aumentado" [6]. En la figura 1 se puede ver ilustrada esta dualidad planteada.

Este concepto, aunque realmente novedoso, necesitaba mejoras para crear aplicaciones creíbles, en las que el usuario o usuarios se sientan realmente inmersos en una realidad aumentada y formen parte de ella. En este sentido, el Túnel Mágico es una evolución natural de nuestro anterior sistema, en el que se propone una aplicación real en términos de juego, narración y exposición, incluyendo la posibilidad de un entorno colaborativo (hasta 6 usuarios).

² Por G. Berkeley.

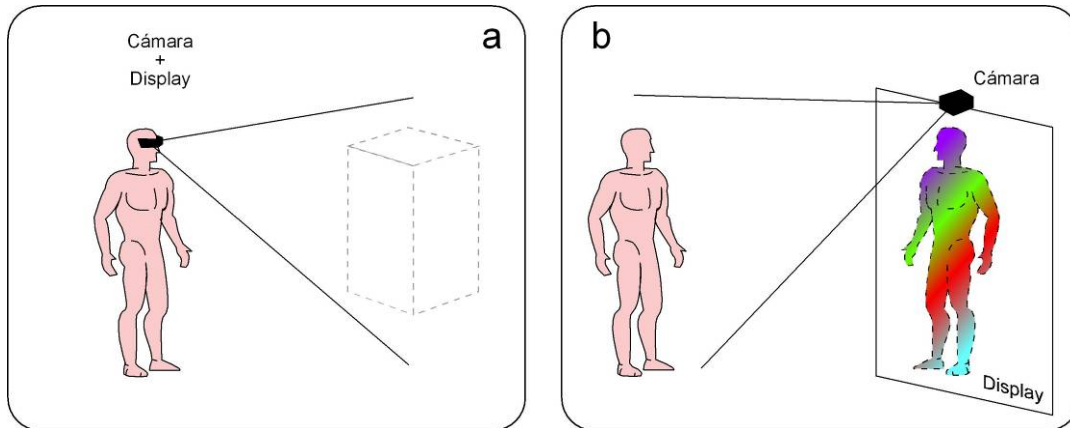


Figura 1. Relación cámara/usuario en aplicaciones de realidad aumentada:
a) En sistemas tradicionales; b) En nuestro sistema.

3. SISTEMA DE REALIDAD AUMENTADA

Nuestro sistema consiste en:

- Las marcas físicas
- 3 Pantallas de retroproyección
- 3 Cámaras Web, conectadas por firewire (IEEE 1394)
- 3 Estaciones gráficas
- 3 Proyectores multimedia
- Focos de luz que permitan ser controlados en intensidad y dirección
- Librerías ARToolKit

El usuario se viste con las marcas y estas son vinculadas a algunos modelos 3D. Estos modelos están en formato VRML, que es susceptible de acoger imágenes digitales como textura, proceso comúnmente llamado mapeado.

Las pantallas de retroproyección tienen una dimensión de 1,40 x 1,40 m y está realizado con un material plástico traslucido. El tamaño de la pantalla y la distancia de ésta al proyector multimedia se han de tener en cuenta en el montaje para asemejar el tamaño de la imagen proyectada a la escala real del usuario, esto ayuda a que la sensación de encontrarse ante espejo sea más real y se refuerce con la proyección ajustada en volteo horizontal para que la imagen sea totalmente simétrica al usuario.

El usuario se sitúa dentro del campo de visión de la cámara y en un rango de distancia que permita al sistema reconocer las marcas con suficiente definición, de 1 a 3 metros. Cada cámara se sitúa en la parte superior de cada pantalla, centrada, y se conecta a una estación gráfica distinta, que proyecta autónomamente en su respectiva pantalla.

Es necesario controlar la luz para que el sistema pueda reconocer las marcas con garantías, y al mismo tiempo las pantallas de retroproyección queden en penumbra y se iluminen únicamente por el proyector multimedia y la imagen del usuario alterado. En la siguiente figura, la situación espacial de todos los componentes del sistema puede ser observada.

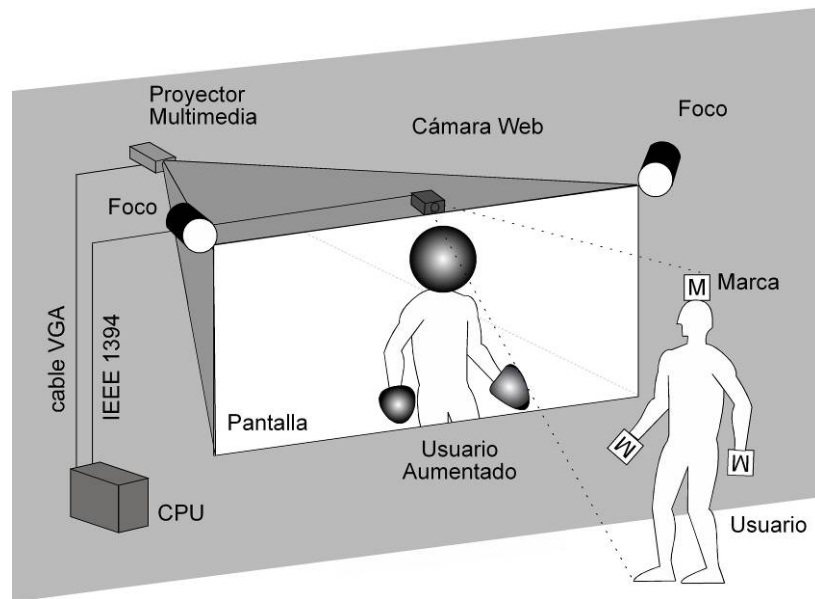


Figura 2. Esquema básico de nuestro sistema.

4. SISTEMA DE MARCAS PARA VESTIR

Como ya se ha apuntado en apartados anteriores, las librerías del ARToolKit necesitan visualizar una o varias marcas para realizar las operaciones de posicionamiento y superposición de elementos virtuales solo que, en nuestro sistema, las marcas forman parte del usuario. Es por ello que se nos planteó la necesidad de crear una serie de marcas que el usuario pudiera ponerse en diferentes partes de su cuerpo (a partir de ahora las llamaremos marcas ponibles), a modo de vestimenta, que fuesen cómodas de llevar para que no interfieran en la libertad de movimiento de los usuarios y, a su vez, que fuesen constantemente visibles por la cámara. En este sentido, se diseñaron y probaron diferentes prototipos, pudiéndose utilizar cada uno de ellos en condiciones diferentes. Estos son:

- marca libre;
- gorro con prisma ajustable;
- prisma con diadema;
- prisma con diadema iluminado;
- marcas ajustables para las manos.

En todos los casos, la parte de color negro de las marcas se realizó con terciopelo, para conseguir una superficie totalmente mate y evitar así los no deseados reflejos lumínicos, que supondrían un reconocimiento erróneo de las marcas por el sistema. En los siguientes sub-apartados se explican todos nuestros prototipos con detalle, junto con sus ventajas e inconvenientes:

4.1. Marca libre

Este constituye nuestro primer prototipo y el más sencillo de todos. Consiste en una única marca con aristas de 18 cm sobre superficie rígida. Las (relativamente) grandes dimensiones de la marca la hacen fácilmente reconocible por el sistema, haciendo

factible los movimientos rápidos del usuario, el cual sujeta la marca con sus propias manos. En principio, puede parecer que este hecho disminuya la libertad de movimientos del propio usuario, al tener que sujetar la marca; sin embargo, esto también posibilita que éste mueva la marca a su antojo pudiendo, por ejemplo, convertirse él mismo en un personaje diferente o tener un amigo virtual, a modo de marioneta. En la figura 3 se aprecia este doble juego.



Figura 3. Marca libre; usuario transformado y usuario con amigo virtual.

4.2. Gorro con prisma ajustable

La cabeza es un sitio idóneo para situar las marcas, puesto que pocas veces quedarán ocultas por los brazos de los usuarios, proporcionándoles así una mayor libertad de movimiento. En un primer intento de situar las marcas en la cabeza, diseñamos una especie de prisma con base triangular variable; la base triangular significa que el prisma está dotado de tres marcas, para que el usuario pueda girar sobre sí mismo permitiendo que la cámara en todo momento visualice una de las marcas, las cuales están vinculadas a un mismo objeto/personaje virtual; la base variable significa que las marcas no están estrictamente rígidas entre ellas, sino que estarán más o menos distanciadas (dentro de un rango pequeño) dependiendo del usuario en cuestión, es decir, el prisma se ajusta a la cabeza de cada usuario. Para permitir este ajuste, cada una de las marcas sobre superficie rígida, se adosó a un gorro de lana. Las dimensiones de las marcas siguen siendo relativamente grandes para que el sistema las reconozca fácilmente (aristas de 18 cm). Este prototipo, aunque realmente atractivo, es un poco aparatoso, sobre todo por las dimensiones de las marcas; otro problema encontrado se debe a la base irregular ya que, al estar las tres marcas vinculadas a un mismo modelo virtual, se producen pequeños desajustes de dichos modelos cuando el usuario gira la cabeza, es decir, cuando el sistema pasa de reconocer una marca a otra o reconoce dos marcas al mismo tiempo. En la figura 4 se ve esto ilustrado.



Figura 4. Usuario con gorro; problemática de la no coincidencia de objetos virtuales.

4.3. Prisma con diadema

El prisma con diadema resuelve los problemas encontrados con el prototipo del gorro. Este modelo es un prisma de base triangular rígida (evitando así la no coincidencia de objetos virtuales), el cual se apoya sobre una diadema. Las diademas siguen siendo ajustables a todo tipo de usuarios pero, al estar el prisma completamente sobre la cabeza, las dimensiones de éste deben ser más reducidas. Al reducir las dimensiones de las marcas, la identificación de éstas por el sistema es más complicada. Es por ello que se hicieron varias pruebas para ver cual sería la dimensión apropiada. Finalmente, se utilizaron marcas con aristas de 8 cm., aunque se debe tener gran cuidado con las condiciones lumínicas del escenario de realidad aumentada, situando varios focos y evitando, en lo posible, que se produzcan sombras. En este caso, la resolución de la cámara también juega un papel importante, y deberemos elegir una resolución mínima de 640x480. En la figura 5 se aprecia el prisma con diadema.



Figura 5. Prisma con diadema.

4.4. Prisma con diadema retroiluminado

Una versión diferente del prisma con diadema, es el prisma con diadema retroiluminado. Este prisma está hecho con material translúcido (la parte blanca de las marcas) para poder dejar pasar la luz de un total de 6 leds insertados en su interior, y alimentadas por una batería en su base. Para utilizar esta nueva variante, es necesario cambiar las condiciones lumínicas del escenario. Así pues, los focos apuntados en el apartado 3 se ajustarán para que aporten una luz tenue, suficiente para que la cámara pueda mostrar al usuario en la pantalla a modo de espejo. Las marcas, al tener luz propia, siempre serán visibles por la cámara. En la figura 6 se aprecia el proceso de fabricación y funcionamiento de este prisma.

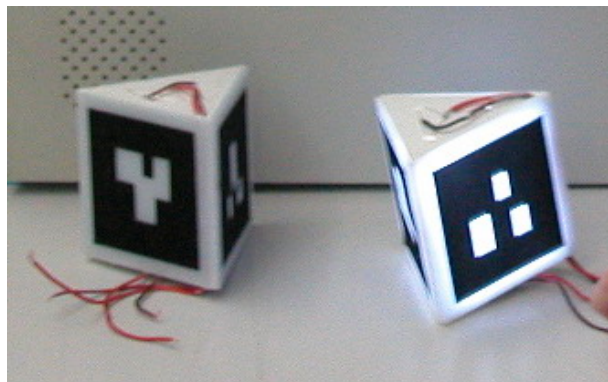


Figura 6. Prisma retroiluminado.

4.5. Marcas ajustables para las manos

Anteriormente se han descrito una serie de marcas que pueden aumentar al usuario (marca libre) o parte de éste (cabeza, en gorro y prisma con diadema) por un objeto virtual; sin embargo, también es interesante el planteamiento de que el usuario sea aumentado por varios objetos virtuales. Así pues, se pueden diseñar otras marcas que aumenten diferentes partes del cuerpo. Es por ello que se realizaron unas marcas para las manos fácilmente ajustables para cada usuario mediante el empleo de velcros a modo de pasador. Este mismo modelo también se puede emplear para las piernas o pies. De este modo, un usuario que vista un prisma con diadema junto con las marcas para las manos, verá su cabeza y sus manos aumentadas, es decir, transformadas o sustituidas por un objeto virtual. La posibilidad de juego aumenta a medida que se van añadiendo marcas. En la siguiente figura vemos una mano con las marcas ajustables.

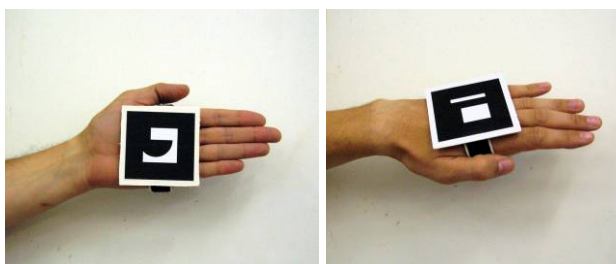


Figura 7. Mano con marcas ajustables mediante velcro.

Nuestro sistema de marcas ponibles es sin lugar a duda un sistema económico y, además extremadamente ligero de llevar; Ambos hechos permiten que este sistema sea fácilmente llevado por todo tipo de usuarios, incluyendo niños, sin el temor de romper los usualmente caros interfaces de visión (HMD, pantallas de ordenador, etc.) o sistemas de posicionamiento (sensores inerciales, GPS, etc.). La única referencia que encontramos de un sistema similar es la apuntada en [12], en la que el usuario lleva un guante con una marca; sin embargo ésta se utiliza simplemente para navegar dentro de un mundo puramente virtual (no aumentado) creado íntegramente por ordenador, y en ningún caso se mezcla realidad con ficción ni se aumenta al usuario.

5. APLICACIÓN FINAL: SISTEMA BÁSICO

Como se ha explicado anteriormente en el apartado 3, el esquema básico de nuestro sistema esta formado por un panel a modo de espejo, una cámara que captura al usuario y las marcas vinculadas a objetos virtuales. Como veremos a continuación, para crear el escenario del Túnel Mágico, este esquema básico se repetirá tantas veces como sea necesario o se considere oportuno. En este apartado vamos a explicar con detalle que sucede en un único sistema (un panel) cuando un único usuario interactúa con éste.

En principio, en el panel puede verse el entorno capturado por la cámara; la imagen se invierte horizontalmente, para dar la sensación de espejo. El escenario visualizado por la cámara estará libre de objetos y será preferiblemente blanco, para evitar que cualquier elemento del entorno sea falsamente reconocido por el sistema como una de las marcas. El usuario entonces se viste con las marcas diseñadas, por ejemplo el prisma con diadema y las marcas en las manos. En el momento que el usuario se hace visible a la cámara, el sistema reconoce las marcas y las reemplaza por los objetos virtuales a los que cada una de ellas está vinculada. Estos objetos virtuales son en realidad modelos 3D, realizados con un programa de diseño por ordenador y exportados a formato VRML para poder ser utilizado

por las librerías del ARToolKit, mediante un modulo especial para este tipo de formato. La elección del formato VRML está justificado por la posibilidad de mapear los modelos con imágenes reales; esto nos ha posibilitado el poder modelar, por ejemplo, míticos actores de cine e incluir su imagen real en el modelo, de tal forma que sea más creíble. También hemos utilizado modelos tipo cartoon, como por ejemplo la pareja del Sr. y la Sra. Patata. Con respecto a los modelos vinculados a las manos, realizamos unas manos también de tipo cartoon, con un tamaño exagerado para incrementar la sensación de cambio. Sucesivamente hemos mejorado la calidad de los modelados que asignamos a cada conjunto de marcas, con lo que el numero de polígonos de cada volumen aumenta y por lo tanto la necesidad de procesar en tiempo real el redibujado de las formas, pero hemos obtenido la misma respuesta de inmediata eficacia por parte del sistema. El siguiente paso fue asignar imágenes a esos modelados de mayor resolución, cosa que tampoco alteró significativamente la rapidez de respuesta. En la siguiente figura se puede ver a un usuario convertido en el Sr. Patata.



Figura 8. Usuarios vistiéndose con las marcas y adquiriendo el aspecto del Sr. y Sra. Patata.

Hay que resaltar que uno de los inconvenientes de la utilización del sistema de marcas reconocibles por el ARToolKit, es precisamente la presencia física de las propias marcas, que otorgan al entorno un aspecto un tanto innatural. En nuestro sistema, aunque las marcas están presentes en el entorno real, estas no se ven en el espejo, puesto que siempre quedan ocultas por los propios objetos virtuales.

6. EL TÚNEL MÁGICO

Teniendo en cuenta que la misma marca puede ser asignada a diferentes modelos, pero cada sistema solo puede asignar una marca cada vez a un modelo específico, no podríamos disfrutar de cambios de modelo dentro de un mismo sistema. Lo podríamos considerar un sistema cerrado a una visualización determinada.

Existen a priori diferentes posibilidades para subsanar este problema, desde programar un cambio de modelo asignado a una marca, lo que conllevaría un cambio asignado a una variable aleatoria o a un contador temporal y por lo tanto un cambio ajeno a la voluntad del usuario/espectador, hasta el relanzamiento del sistema con un modelo diferente cada vez, proceso que conlleva una espera mientras un sistema se desconecta y otro se carga. Decidimos que lo interesante sería buscar un procedimiento por el cual el sistema ofreciera de manera inmediata al usuario la posibilidad de adoptar diferentes identidades alteradas en función de su voluntad.

En este punto surgió el proyecto del Túnel Mágico. Utilizando diferentes sistemas en paralelo, otorgamos al usuario la potestad para cambiar de identidad cuando quiera o recrearse en la actual todo el tiempo que lo desee, y también de hacerlo de una forma natural e inmediata, simplemente andando del campo de visión de una cámara a la de otra. Al estar los sistemas cargados de antemano, la proyección es inmediata en cuanto se reconocen las marcas por cada cámara. Este sistema múltiple de visualización de los usuarios y los modelos permite una gran interactividad, puesto que no sólo puede un usuario verse modificado, ampliado o menguado, sino que también es capaz de ver las transformaciones de sus acompañantes en el sistema autónomo paralelo y comprobar sus evoluciones con cada una de estas identidades.

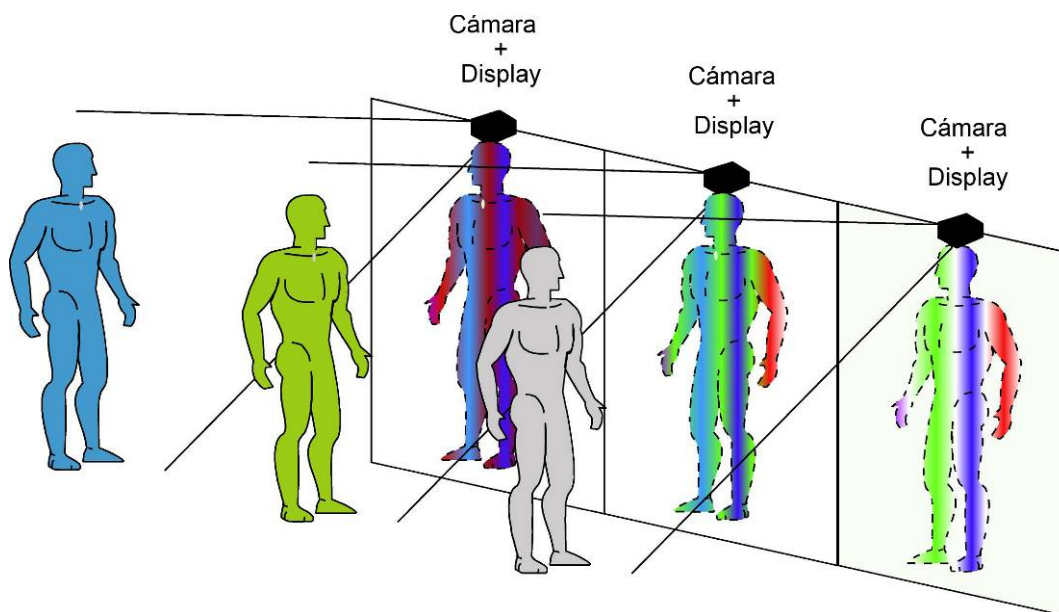


Figura 9. Esquema de la disposición y efecto del túnel mágico

La secuenciación de los sistemas y las condiciones de luz necesarias para su correcto funcionamiento han propiciado el título del proyecto, pero también nos han sugerido posibles usos didácticos, expositivos, pedagógicos y comerciales. Por ejemplo, si el sistema fuese utilizado en un museo de ciencias naturales, podríamos conseguir que un usuario fuese capaz de observarse detenidamente, al transformarse en oruga, crisálida y finalmente mariposa, y observar con detalle las diferencias morfológicas de estos tres estados de una manera totalmente subjetiva utilizando simplemente su movimiento corporal a voluntad. Si el sistema se instalase en una exposición histórica el usuario podría convertirse en Napoleón o Wellington y sentir por sí mismo como se podría desenvolver en un determinado espacio de época recreado que sirviese al mismo tiempo de plató al sistema.

En nuestro caso, y ante la ausencia de un objetivo determinado, nos hemos dedicado a comprobar la fiabilidad del sistema con parejas de identidades asignadas a cada pantalla. Así, hemos utilizado modelados de las cabezas de Walter Mathau, Jack Lemmon, Oliver y Hardy y Charles Chaplin, que en algunos casos hemos complementado con marcas ajustables para las manos.

Si bien los primeros modelados, rápidos y sencillos, recuerdan a trabajos de Tony Oursler³, la sensación de cambiar de cabeza y actuar en consecuencia si el usuario reconoce la nueva identidad que ha adoptado es inmediata y se produce una contaminación

³ <http://www.tonyoursler.com/>

del comportamiento muy interesante. Si en los trabajos de Tony Oursler la reflexión del espectador deriva de la observación de los modelos estáticos objeto de las proyecciones, en este caso, el espectador es también usuario y objeto observado al mismo tiempo, lo que conlleva una participación activa y por lo tanto un resultado diferente siempre de la interacción personal que acomete cada uno ante esa situación, el movimiento natural de cada persona, su lenguaje corporal, pasan a formar parte de la obra, contaminado como antes decíamos, por la identidad que el sistema le asigna.



Figura 10. Usuario a través del Túnel Mágico.
Transformándose en: Oliver, Chaplin y Hardy.

7. RESPUESTA DE LOS USUARIOS

El Túnel Mágico se puso en marcha en la sala de exposiciones de la facultad de BB. AA. de la Universidad Politécnica de Valencia durante dos días. El primer día se hizo una filmación exhaustiva del sistema, probando todo tipo de marcas diseñadas. Este primer ensayo se realizó por los propios autores con el fin de probar el sistema y corregir o ajustar factores tales como la luminosidad, resolución de la cámara, distancia usuario/cámara y movilidad del usuario dentro del escenario establecido. El segundo día se abrió la exposición al público en general, que tuvieron la posibilidad de probar nuestro sistema a modo individual y colectivo. La respuesta de los usuarios fue mejor de la esperada y totalmente enriquecedora. La primera impresión que produce nuestro sistema es de "asombro", al verse el usuario a sí mismo transformado en otro personaje. A continuación, generalmente se establecía un ambiente de juego y diversión, especialmente si la aplicación era colaborativa, es decir, varios usuarios interactuando al mismo tiempo con el sistema. Los usuarios permanecían un tiempo interactuando en una sola pantalla; posteriormente pedían más cambios de personalidad, por lo que simplemente tenían que interactuar con otra de las pantallas para descubrir el nuevo personaje en el que se transformaban y crear nuevos juegos. De esta experiencia se derivan las siguientes conclusiones:

- A mayor cantidad y variabilidad de personalidades adquiridas, mayor interés;
- A mayor número de partes del cuerpo aumentadas, mayor posibilidad de juego;
- A mayor número de usuarios interactuando con el sistema, mayor diversión.

En la siguiente figura, se muestran imágenes de los usuarios interactuando con el sistema:



Figura 11. Distintos usuarios interactuando con el sistema

También pedimos a los usuarios que rellenasen unas encuestas para, según su criterio, evaluar el sistema; en total se rellenaron 34 encuestas. La edad de los participantes oscila entre 7 y 60 años, la mitad de ellos hombres y la otra mitad mujeres. En esta encuesta, les pedíamos que evaluarasen del 1 (malo) al 10 (excelente) un total de 9 aspectos. Estos son, junto con las puntuaciones medias dadas por los usuarios, diferenciando hombres de mujeres:

- Facilidad de interacción con el sistema: H (7,4), M (8,2)
- Libertad de movimientos del usuario: H (6,6), M (7,7)
- Diseño y confort de las marcas posibles: H (7,2), M (7,9)
- La idea de espejo como interface: H (8,1), M (8,4)
- Personalidades adquiridas: H (8,1), M (8,8)
- Posibilidad de juego/diversión: H (8,0), M (9,4)
- Posibilidad en el área del aprendizaje: H (7,8), M (8,5)
- Posibilidades terapéuticas: H (7,7), M (8,6)
- Posibilidad de incluir el sistema en museos: H (7,4), M (8,8)

Como se puede observar, todas las puntuaciones están por encima de 6,5 y, en general, las mujeres disfrutaron más de la experiencia.

Finalmente, también pedimos que nos diesen su opinión personal y que anotasen si cambiarían o añadirían algo al Túnel Mágico. Algunas de las aportaciones que nos parecieron más interesantes son: "Puede tener grandes aplicaciones didácticas, al igual que terapéuticas, especialmente en tratamientos de fobias"; "Su uso en psicoterapia: juegos de rol, terapia familiar, psicodrama. Sería interesante incorporar la voz en paralelo o integrada para que fuera más completo"; "Me he divertido mucho, falta igual ambientación musical de acuerdo con el personaje creado"; "Se podría crear un entorno. Por ejemplo el usuario convertido en Tarzán y el espacio en el que se mueva, ambientado como si fuera la selva..."; "Incluir más elementos. Profundizar en lo didáctico". Estas consideraciones nos parecen muy enriquecedoras, puesto que a partir de ellas podemos mejorar nuestro sistema, y por ello serán consideradas en futuros desarrollos.

8. CONCLUSIÓN

En esta comunicación hemos presentado el Túnel Mágico, una aplicación de la tecnología de Realidad Aumentada usando las librerías ARToolKit. En nuestro sistema, la presentación habitual se invierte, la cámara no pretende reemplazar el punto de vista del

usuario, sino que facilita su propia contemplación dentro de un sistema al que siempre había accedido como mero espectador.

El usuario se viste con marcas que son reconocidas por las cámaras que incorpora el sistema y que éste sustituye por objetos virtuales precargados. Una vez realizada esta sustitución, el sistema proyecta la imagen aumentada del usuario con los objetos virtuales superpuestos.

Como las marcas se sitúan en partes de la anatomía del usuario, para contemplarlas mejor o moverlas, el usuario, ha de mover su cuerpo, algo que se intuye muy pronto y que se realiza de forma natural, con lo que el resultado, la experiencia final, no es adquirida a través de un proceso o interface extraño o complicado para el usuario, la herramienta es su cuerpo y lo que hace con él.

En este sistema, el usuario puede moverse libremente, pudiendo realizar rotaciones, inclinaciones y acercamiento o alejamiento a las pantallas que funcionan a modo de espejo; sin embargo, es aconsejable acotar el espacio físico por donde se puede mover, para permanecer siempre dentro del campo de visión de la cámara.

El sistema puede ser precargado con cualquier tipo de modelado virtual en formato vrmf y con imágenes para texturizar al mismo en formatos *.bmp, *.jpg o *.gif, de suficiente resolución para su proyección con garantías de calidad dentro del sistema. Este proceso, muy fácil de programar, se realiza con un muy pequeño coste de tiempo o material.

La duplicación de los recursos permite que la misma marca sea reconocida en todos los sistemas, pero que el modelado asignado a cada marca varíe.

La disposición de los sistemas uno al lado del otro permite disponer un hilo temporal y/o temático conductor del sistema de un extremo al del otro, con diferentes intereses u objetivos, narrativo, expositivo, lúdico, comercial, didáctico...

La utilización de sistemas paralelos con campos de interacción autónomos y separados permite que el usuario cambie de identidad a voluntad simplemente abandonando el campo de visión de una cámara y entrando en el de otra.

La rápida evolución de nuestra aplicación nos hace prever un sistema muy estable de alta interacción en tiempo real, de fácil uso y con diversos campos de aplicación.

REFERENCIAS

- [1] ARToolKit: <http://www.hitl.washington.edu/artoolkit/>
- [2] BILLINGHURST, M., HIRKAZU, K., POUPYREV, I. (2001). The MagicBook– Moving Seamlessly between Reality and Virtuality, *IEEE Computer Graphics and Applications*, Volume 21 (3), ISSN:0272-1716. IEEE Computer Society Press, Los Alamitos, CA, USA, pp. 6-8.
- [3] BILLINGHURST, M., KATO, H. (2002). Collaborative Augmented Reality. *Communications of the ACM, How the virtual inspires the real*, Volume 45 (7), ISSN:0001-0782, ACM Press, New York, NY, USA, pp. 64-70.
- [4] CHEOK, A. D., WEIHUA, W., YANG, X., PRINCE, S., WAN, F. S., BILLINGHURST, M., KATO, H. (2002). Interactive Theatre Experience in Embodied + Wearable Mixed Reality Space., *International Symposium on Mixed and Augmented Reality (ISMAR'02)*, ISBN: 0-7695-1781-1, pp. 59-317.
- [5] DÍAZ, D., BOJ, C. (2005). Red Libre Red Visible, Taller de realidad aumentada y redes inalámbricas, 15-18 Febrero 2005.
- [6] GINER, P., PORTALÉS, C. (2005). The Augmented User. *Proceedings of the 11th International Conference on Virtual Systems and Multimedia (VSMM'05)*, Gante, Bélgica, Octubre 03-07.
- [7] GORDON, G., BILLINGHURST, M., BELL, M., WOODFILL, J., KOWALIK, B., ERENDI, A., TILANDER, J. (2002). The Use of Dense Stereo Range Data in Augmented Reality, *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR'02)*, ISBN: 0-7695-1781-1, IEEE Computer Society, Washington, DC, USA, pp. 14-23.
- [8] HEDLEY, N. R., BILLINGHURST, M., POSTNER, L., MAY, R. (2002). Explorations in the use of Augmented Reality for Geographic Visualization, *Presence: Teleoperators and Virtual Environments*, Volumen 11 (2), ISSN: 1054-7460, MIT Press, Cambridge, MA, USA, pp. 119-133.

- [9] KATO, H., BILLINGHURST, M., POURPYREV, I., IMAMOTO, K., TACHIBANA, K. (2000). Virtual Object Manipulation on a Table-Top AR Environment, *IEEE and ACM International Symposium on Augmented Reality (ISAR'00)*. ISBN: 0-7695-0846-4, Munich, Alemania, Octubre 05-06, pp. 111-119.
- [10] SINCLAIR, P. A. S., MARTINEZ, K. (2002). Tangible Hypermedia using ARToolKit, *Augmented Reality Toolkit, The First IEEE International Workshop*, ISBN: 0-7803-7680-3, 2 pp.-.
- [11] SINCLAIR, P. A. S., MARTINEZ, K., MILLARD, D. E., WEAL, M. J. (2002). Links in the Palm of your Hand: Tangible Hypermedia using Augmented Reality, *Proceedings of the thirteenth ACM conference on Hypertext and hypermedia*. ISBN:1-58113-477-0, ACM Press, New York , NY, USA, pp. 127-136.
- [12] WOODS, E., MASON, P., BILLINGHURST, M. (2003). MagicMouse: an Inexpensive 6-Degree-of-Freedom Mouse, *Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia*. ISBN:1-58113-578-5, ACM Press, NY, USA, pp. 285-286.

AUGMENTED TEACHING

Francisco Giner, Cristina Portalés

Laboratorio de Luz, Universidad Politécnica de Valencia
Valencia (Spain)

fginerm@pin.upv.es, criporri@doctor.upv.es

Abstract

In this paper we present a new proposal for the E-learning through the Augmented Reality (AR) technology. We have developed two possible applications based on the same idea. The first one is mainly built for youth and children, where a theatre environment is established to live performing a tale, within an entertainment environment. In this so called AR-theatre, children are the actors and the AR is used for the customs. The second one is thought as a complementary tool for history learning for all ages, where the teacher can 'manipulate' 3D-virtual ancient tools and show them to the pupils, with the additional advantage that virtual objects are shown with a real scale and related to the human being. Finally, conclusions are given and further work is proposed.

Keywords

Augmented reality, technology, e-learning, theatre, collaborative, real-time.

1. INTRODUCTION

Electronic learning or E-learning is an all-encompassing term generally used to refer to computer-enhanced learning, although it is often extended to include the use of mobile technologies such as PDA and MP3 players. It may include the use of web-based teaching materials and hypermedia in general, multimedia CD-ROMs or web sites, discussion boards, collaborative software, e-mail, blogs, text chat, computer aided assessment, educational animation, simulations, games, learning management software, electronic voting systems and more, with possibly a combination of different methods being used. On the other hand, AR is a field of computer research which deals with the combination of real world and computer generated data. At present, most AR research is concerned with the use of live video imagery which is digitally processed and "augmented" by the addition of computer generated graphics. Advanced research includes the use of motion tracking data, marker recognition using machine vision, and the construction of controlled environments containing any number of sensors and actuators [12].

With the term Augmented Teaching we are extending the term E-learning to the AR technology. We say "teaching" instead of "learning" because our system is thought as a complementary tool to aid teachers or other people that need to teach something, although self-learning is also possible. Anyway, the term teaching implies also learning, if real bidirectional communication exists.

In the following sections, we will explain some technical aspects of the developed system and the two proposals for the E-learning. In both cases, some characteristics and/or advantages of our teaching system will be shown.

2. THE AUGMENTED TEACHING SYSTEM

Augmented Teaching is an AR-system based on the well-known ARToolKit libraries [1]. These libraries are marked based and open source, and have been used to develop several applications for different purposes, as for example, for geographic visualization [6; 7], video conferencing [2; 8], tangible hypermedia [10; 11], gaming [3; 9] or arts [4; 5]. In comparison to these applications, in our system a relativization of the spatial distribution is introduced (Fig. 1), thus some kind of dualities regarding to the perception stand out:

- **Inside/Out:** User incorporation inside the representation space. Normally, in the AR applications the user is wearing a head mounted display (HMD), while the camera is attached to it, representing the user's point of view. In our application the user is not longer wearing a HMD, while the camera is fixed in the space and pointing to him, thus being the user inside the representation space.
- **Subject/Object:** The user is part of the object. In AR marker based applications, the markers are located in the space, and the user looks at them in order to see the augmentation. In our case, the user wears the markers, thus being him the augmented thing.

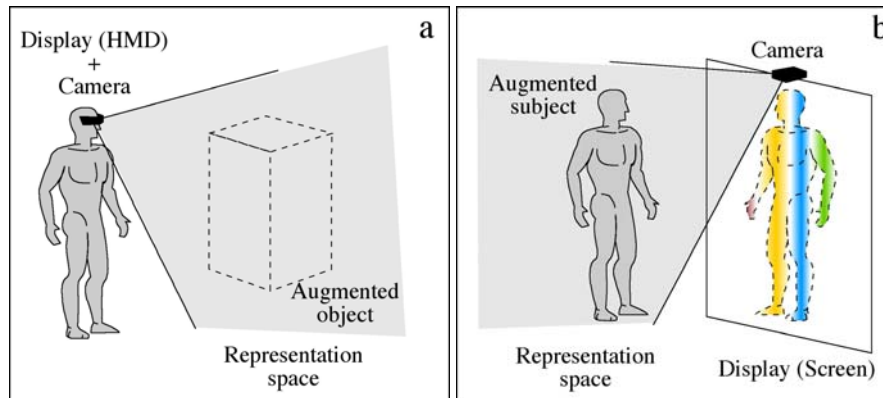


Fig. 1. AR environment: a) Traditional system; b) Our system.

The system of Augmented Teaching consists on (Fig. 2):

- the ARToolKit libraries,
- some physical wearable markers,
- an overhead screen,
- a multimedia projector,
- a web cam with USB or IE1394 connection,
- a computer graphics workstation,
- controlled light conditions.

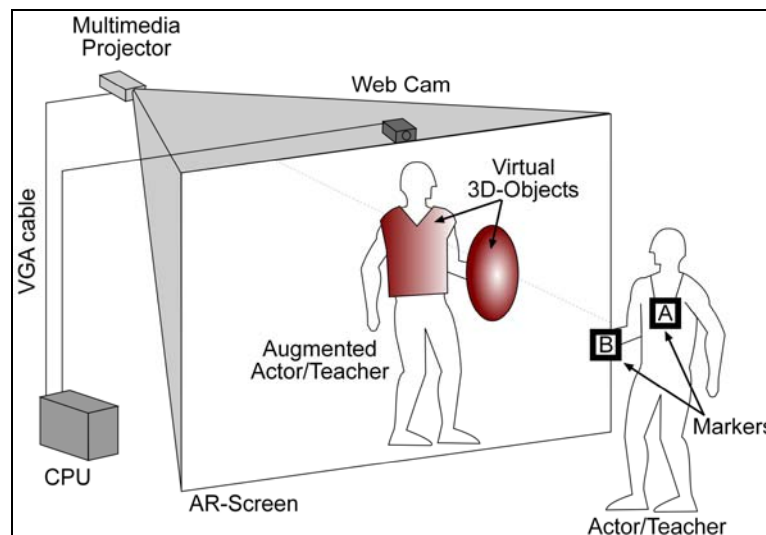


Fig 2. Schema of the Augmented Teaching system.

The ARToolKit libraries provide the necessary computer vision techniques required to track some kind of markers, and accurately calculate the camera pose relative to them. Virtual objects are then posed – relative to the markers – in the 'real' world. In our system, users are carrying a set of markers, in some cases attached to their bodies (wearable markers). These markers are related to some 3D virtual models in the VRML format.

The overhead screen has a dimension of 2.0 x 1.5 m, and is made of a translucent white plastic. The web cam is situated at the very top of the screen pointing to the actors/teachers. The dimension of the screen and the distance to the projector should be taken into account in order to achieve a representation scale similar to the real one, thus allowing what we call “human dimension” (see section 4). Also it is important to consider that the markers should not be too far away from camera, in order that the software can properly identify them. Thus the representation space is restricted within a distance of 3 m to the screen. The used web cam in this case (Live! Cam Voice from Creative) is a USB 2.0 with a resolution of 640 x 480 ppp and a FOV of 85°. Also a web cam with IE1394 connection can be used.

3. THEATRE LEARNING FOR CHILDREN

Tales can be represented by children, each one acquiring different personalities according to the interpreted character. One child can play different characters, with the only need to change markers. Markers can be carried with their hands or hanged with a necklace. The AR-theatre environment is built as shown in Fig. 3, where audience see the performance in real time through the AR-screen, and actors are at the back of the screen, where the web cam is pointing to them. Some of the characteristics that define the so established AR-theatre include:

1. **Collaborative reality:** AR allows users to interact with graphical objects while still allowing users to see the real world. This allows several users to naturally collaborate with a computer generated objects as if they were real world items.
2. **Natural interaction:** actors interact with the computer and within themselves in a natural way, as they only have to move their bodies or the markers in order to activate interaction.
3. **Real-time:** the system responds almost immediately to the movements of children, thus allowing a real-time performance.
4. **Wearable markers:** comfortable wearable markers were designed in order to avoid interferences with the children’s natural movements. For this reason, necklace-form-markers were designed. Each virtual object is attached to a pair of markers, one to show the front of the object and the other to show the back.
5. **AR-mirror:** the interface between user-computer is an overhead screen. This allows actors to see themselves while acting, as there was a mirror in front of them, thus they can control their performance and if the augmentation is success. Audience will see the augmented-actors projected on the screen, as if they were on a cinema, with the characteristic that the AR-theatre is in real time. Actors do not see the audience directly.
6. **3D virtual models:** through the 3D virtual models the actors experience a transformation of their body. By changing the markers, they change the model and thus their role.
7. **Fast user adaptation.** The curve of learning to use the application is very short and pronounced, since the possibilities of use are very intuitive. Moreover, the perceived results of the changes in the manipulation of the markers are immediate.

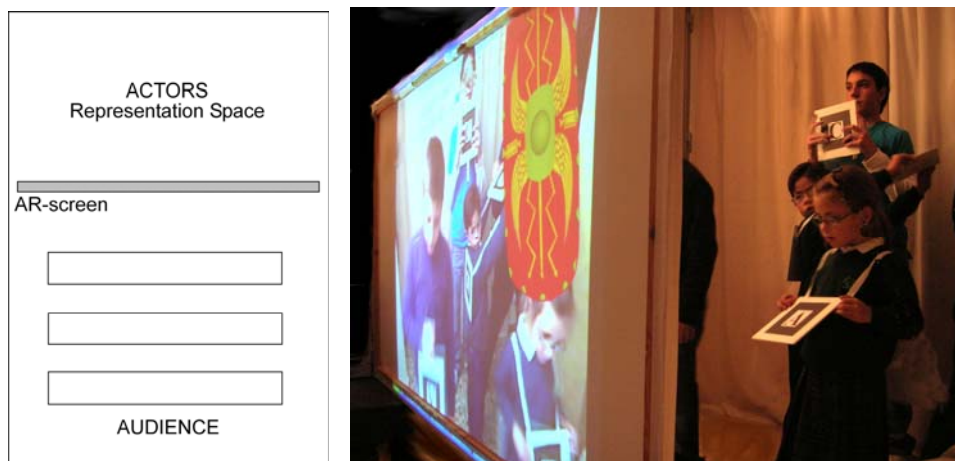


Fig. 3. Scheme of the AR-theatre environment.

We proposed a group of 5 children between 7-12 years old to do a performance of the well-known tale of *Alice's Adventures in Wonderland*. This is a work of children's literature by the English mathematician and author, Reverend Charles Lutwidge Dodgson, written under the pseudonym Lewis Carroll. It tells the story of a girl named Alice who falls down a rabbit-hole into a fantasy realm populated by talking playing cards and anthropomorphic creatures. For our AR-theatre, some of tale's characters were modelled: the clock, the teapot, the cup, the Queen of hearts and the 10th of clubs.



Fig. 4. Children performing the tale "Alice's Adventures in Wonderland".

Most of them knew about the tale, but were not completely sure about their role: what they had to say, how, in which order, etc. So, instead of doing a reliable performance, they invented most of their roles. They acted accordingly to their characters and played with quite a lot of imagination.

In Fig. 4 we can see: the teapot trying to pour some tea into the cup; the child performing the cup is adding some sugar inside; the cards are steady and hanging their sticks; the clock is talking with the teapot and becomes angry because he also wants some tea; etc. Throughout the time, different possibilities of game arose and children never got bored. They asked for more roles, thus they wanted to change markers with their friends in order to play new characters. This enriched the interaction within the group, thus increasing the learning possibilities.

4. HISTORY LEARNING

Virtual objects can be 'carried' by real users. In this sense, teachers can perform augmented history learning, by carrying different tools/cloths used in the past, with a real representation scale. A set of markers can be linked to different virtual objects, thus being related to an specific theme. Teachers can carry these markers and show them to the pupils that will be able to see the objects with a human dimension. Objects can be moved, rotated and carried according to their real use in the past, i.e., they can be 'manipulated'. They can be put together and compared between them. Therefore, comparing this AR-system with traditional book-learning or even e-learning through a single computer screen, some advantages stand out:

1. **Manipulation.** Imagine that we are interested in ancient shields, as the one shown in Fig. 5. Traditional learning through books can give us a detailed description of these objects (shape, materials, colours, dimension, etc.), but it would be hard to build in our mind a reliable image of the object if the description is not accompanied with a drawing or even a photograph. But images are static, so maybe we would need at least two or three images of a single object to have a complete description (front, back and side). Depending of the complexity of an object, even more images can be needed (showing details, corners, holes, etc.), but the more images, the more unconnected can appear that object to us, leading to misunderstandings. On the other hand, virtual objects do not present this problem, as they can be manipulated. This means that we can rotate, translate or zoom a 3D object, i.e., we can 'navigate' through every corner, twists and turns of the object, approximate our vision to better see the details or zoom away to see the whole. This characteristic is not only present in AR-systems, but in almost every software capable to work with 3D-virtual objects.



Fig 5. Different views of a Greek shield:
Front, side, back, detail at the back, detail at the front.

2. **Human dimension.** When we hear that an object has, for instance, a dimension of 1.75 x 0.65 x 0.27 m, we have to imagine what it really means, thus opening our arms long, wide and high to compare it with ourselves. Therefore, by "human dimension" we mean that real physical dimension of an object that we try to represent in our minds. Books can show 1:1 objects if they are not too big (smaller than the book itself). Virtual objects rendered in a computer screen can be zoomed, but they do not solve human dimension problem, as a 1:1 dimension can vary depending on our screen size. On the other hand, our AR-system is able to solve this problem, as the object is both carried by a human being and having a 1:1 dimension (Fig. 6). The AR-mirror is indeed greater than a book or a computer screen, allowing bigger objects to be represented. Furthermore, they can bring the experience of teaching to the dimension of apprehension or perception of their physical surrounding

environment, which is something that we are used to do in a natural way across our senses and across the space that our body occupies.



Fig 6. A teacher carrying ancient weapons. Left: a Romanic and a Greek shield; Middle: Romanic shield and sword; Right: back of a Romanic shield.

3. **Tangible interaction.** In AR-technology, computer virtual objects can be attached to real-world physical objects by means of markers. By changing the position of the real object with our own hands, we can correspondingly change the position of the virtual object. This is known as “tangible interaction” and provides an extremely natural interface for manipulating computer generated objects which does not need to be learnt by the user. To sum up, we can really “carry” the objects to manipulate them, instead of simply using a mouse or a joystick within a computer. Also, objects can be attached to our body by means of wearable markers, becoming part of ourselves (Fig. 7).

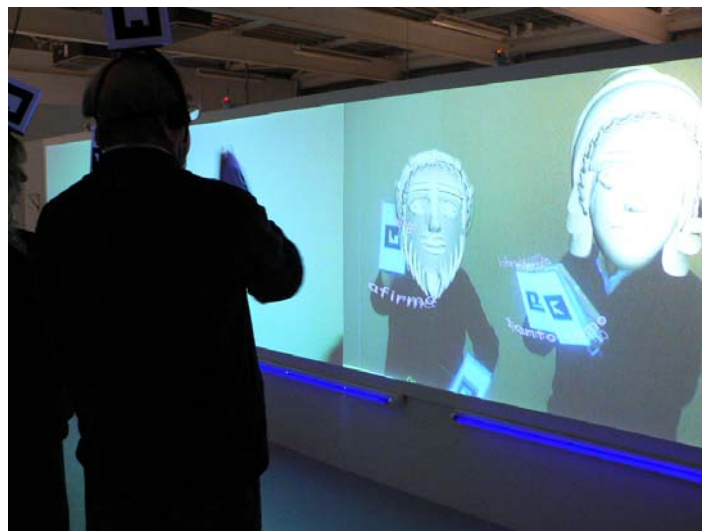


Fig. 7. Carrying some wearable markers. Ancient Greek masks are shown.

5. CONCLUSION

Through this paper we have shown the great potential of E-learning by means of the AR-technology. Two practical examples have been shown: the first one called “theatre learning for children” and the second one “history learning”. Through these proposals, we have explored some of the advantages of our system regarding to other traditional learning tools, as for example, collaborative reality, natural interaction, manipulation, human dimension and tangible interaction, while still allowing face-to-face learning.

Although Augmented Teaching is not a panacea, we think it has a great potential for the teaching and/or learning process, because it allows exploring virtual objects and relate them to our body in a very intuitive and natural way. Furthermore, it is a non-expensive system because it is based on open-source software and tracking is achieved via markers, thus still having a great potential to add, change or personalize new contents. This means, that with a low initial inversion, the system can be fast promoted.

Another great advantage is that there is no limitation in age, and everybody can easily interact with the system. This is mostly thanks to the wearable markers, which are non-expensive, comfortable and light-weighted interfaces that require no previous knowledge for the interaction.

6. FURTHER WORK

Further work will be focused on the real teaching experience through the Augmented Teaching system in order to evaluate the potential of the learning process. Based on these experiences, also an improvement of our system will be done. Moreover, we think that the system can also be used to aid in some kind of therapies, as for example sociological or psychological, or new educational methods, so they disinhibit themselves and establish communication with the roles of other users. In this sense, we can say that users and audience perceive the system as a funny game in a fast and natural way.

References

- [1] ARToolKit: <http://www.hitl.washington.edu/artoolkit/>
- [2] Billinghurst, M., Cheok, A. D., Prince, S., and Kato, H. (2002) Real World Teleconferencing, IEEE Computer Graphics and Applications. Volume 22, Issue 6. IEEE Computer Society Press, Los Alamitos, pp 11-13.
- [3] Cheok, A. D., Teh, K. S., Nguyen, T. H. D., Qui, T. C. T., Lee, S. P., Liu, W., Li, Ch. C., Diaz, D., and Boj, C. (2006) Social and physical interactive paradigms for mixed-reality entertainment, Computers in Entertainment (CIE) archive. Volume 4, Issue 2. ACM Press, New York.
- [4] Cheok, A. D., Weihua, W., Yang, X., Prince, S., Wan, F. S., Billinghurst, M., and Kato, H. (2002) Interactive Theatre Experience in Embodied + Wearable Mixed Reality Space, International Symposium on Mixed and Augmented Reality (ISMAR'02), pp 59-68.
- [5] Galantay, R., Torpus, J., and Engeli, M. (2004) "living-room": interactive, space-oriented augmented reality. Proceedings of the 12th annual ACM international conference on Multimedia. ACM Press, New York, pp 64-71.
- [6] Gordon, G., Billinghurst, M., Bell, M., Woodfill, J., Kowalik, B., Erendi, A., and Tilander, J. (2002) The Use of Dense Stereo Range Data in Augmented Reality, Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR'02), IEEE Computer Society, Washington, DC, pp 14-23.
- [7] Hedley, N. R., Billinghurst, M., Postner, L., and May, R. (2002) Explorations in the use of Augmented Reality for Geographic Visualization, Presence: Teleoperators and Virtual Environments, Volume 11 (2). MIT Press, Cambridge, MA, pp 119-133.
- [8] Kato, H., Billinghurst, M., Morinaga, K., and Tachibana, K. (2001) The effect of spatial cues in augmented reality video conferencing, Proceedings of the 9th International Conference on Human-Computer Interaction. Lawrence Erlbaum Associates, Mahwah, N.J., pp 478-481.
- [9] Kato, H., Billinghurst, M., Pourpyrev, I., Imamoto, K., and Tachibana, K. (2000) Virtual Object Manipulation on a Table-Top AR Environment, IEEE and ACM International Symposium on Augmented Reality (ISAR'00). Munich, Germany, pp 111-119.
- [10] Sinclair, P. A. S., and Martinez, K. (2002) Tangible Hypermedia using ARToolKit, Augmented Reality Toolkit, The First IEEE International Workshop, 2 pp.-.
- [11] Sinclair, P. A. S., Martinez, K., Millard, D. E., and Weal, M. J. (2002) Links in the Palm of your Hand: Tangible Hypermedia using Augmented Reality, Proceedings of the thirteenth ACM conference on Hypertext and hypermedia. ACM Press, New York, pp 127-136.
- [12] Wikipedia: <http://en.wikipedia.org/>

LIVE LEGO HOUSE, A MIXED REALITY SPACE FOR THE EDUTAINMENT

Cristina Portalés
Laboratorio de Luz
Universidad Politécnica de Valencia
Camino de Vera s/n
46022 Valencia (Spain)
Telephone number: (0034) 963877007
E-mail: criporri@doctor.upv.es

Abstract

Gaming is a natural way of learning and multimedia technologies bring new possibilities in the field of edutainment. Live LEGO House in an interactive space to explore coexistence and multicultural factors through gaming based on the Mixed Reality (MXR) technology. The system consists basically on a physical/real house built with the LEGO blocks, which is enriched with different non-physical/virtual multimedia files, including sounds, videos and 3D animations with multicultural content. Children interact with the house via two dolls, each of them having different characteristics and needs. They (children) will need to agree in some aspects in order to play the game properly, to coexist. Tactile, visual and auditory stimuli are present in the game.

Keywords: mixed reality, interaction, edutainment, technology.

1. Introduction

Edutainment (also educational entertainment or entertainment-education) is a form of entertainment designed to educate as well as to amuse. Edutainment typically seeks to instruct or socialize its audience by embedding lessons in some familiar form of entertainment: television programs, computer and video games, films, music, websites, multimedia software, etc. Most often, edutainment seeks either to tutor in one or more specific subjects, or to change behaviour by engendering specific socio-cultural attitudes. Successful edutainment is discernible by the fact that learning becomes fun and teachers or speakers educate an audience in a manner which is both engaging and amusing (Wikipedia, 2007).

Multimedia technologies have demonstrated a great potential in the field of education, since they offer an excellent technological base for the training and transmission of knowledge, which generates more efficient methodologies of education and learning. Equally in the area of entertainment and leisure, these technologies open new perspectives that offer new chances in this important human facet.

Live LEGO House (LLH) is an interactive game for children based on the MXR technology, which takes great advantage of multimedia systems (sounds, images, videos, etc.). It is thought to entertain as the same time as to educate in the area of coexistence and multicultural factors. Nevertheless, education is not explicit, as no rules are necessary to play. Two dolls – led by two children – will have to interact with the house, discovering a place where plenty of sounds, videos and 3D virtual objects will respond to their acts. But actions do have consequences, so they have to decide what to do: what they want or what seems to be better for both.

LLH has been specially developed for, but not restricted to, children between 3-7 years old. To keep children's attention at these ages, a large variety of funny sounds, graphics and videos are introduced. The thematic of all these media is 'multicultural', so children can indirectly learn about other cultures. Furthermore, in order to reinforce good environmental and social concepts, some restrictions regarding to water and sound are introduced, in such a way that positive coexistence between players is fomented.

This application was developed by the author in Singapore, while a research intern at the Mixed Reality Lab. The idea of developing a game based on coexistence and multiculturalism was born there. One of the characteristics of Singapore is the actual coexistence between people from different races (Chinese, Malaysian and Indian) and religions (Buddhism, Islamism, Catholicism, Hinduism, etc.), and even different mother languages (Mandarin, Malay and Tamil), all of them having English as a unifier. Although there are some characteristic ghettos (as Chinatown), cultures are quite mixed. Some examples of this coexistence are: people from different countries live and work together, respecting their culture and traditions; a Hindu Temple can be found in Chinatown; all kind of traditional food is offered at the stalls of a single hawker centre; different traditions are celebrated during the whole year; different languages are in use with the English as link; etc. Nevertheless, an "equal" coexistence is still not a fact, but this seems to be more a problem of social class and/or poverty than of cultural roots. Some of these coexistence facts are shown in LLH by means of some videos that have been recorded there.

2. Edutainment through MXR

MXR refers to the combination of computer graphics and real-world objects. This encompasses both augmented reality, which involves placing computer graphics objects into the real world and augmented virtuality, which involves placing real-world objects into virtual environments. MXRToolKit (2007) consists of a library of routines to help with all aspects of building mixed reality applications. The philosophy of the SDK is to keep the interface extremely simple. Therefore, one of the advantages of the MXRToolKit compared with other marker based programs is the speed with which you can write your own applications – even for non-programmer experts – and integrate different kind of multimedia files (sound, videos, 3D objects, animations, etc) in a single project.

Comparing MXR applications with Virtual Reality, two major advantages stand out:

1. **Collaborative reality.** Mixed reality allows users to interact with graphical objects while still letting them to see the real world. Several users can collaborate with a computer generated object as if they were real world items, thus leading to a natural human-human interaction.
2. **Tangible interaction.** Computer graphical objects can be attached to real-world physical objects by means of markers. By manipulating the position of the real object, we can correspondingly manipulate the position of the virtual object. This is known as "tangible interaction" and provides an extremely natural interface for manipulating computer generated objects which does not need to be learnt by the user.

The game can be divided in the following steps:

1. Construct your own house: use your imagination and creativity to build a house within the LEGO blocks.

2. Tell the program where the objects are: this is called the calibration process. Simply by pressing some buttons on a prepared keyboard you tell the pc how your house is arranged. Nevertheless, this step is perhaps too difficult for small children, so they will need some help.
3. Explore the matters of coexistence and try to solve them wisely: you will discover how to play taking into account your partner.

3. Game design

According to (Liarokapis, 2006), important issues of game design in MXR include:

1. Technical characteristics: advanced visualization and interaction; adaptability and usability issues.
2. Interactivity: advanced using computer sensors and natural methods.
3. Social, cultural, and pedagogical issues: can be direct and indirect.
4. Collaboration: can be direct and/or indirect.
5. Game scenarios: more pragmatic; need to merge scenario with real world.

In the following sub-sections, these issues are seen in more detail for the LLH.

3.1 Technical characteristics

LLH is built within the MXRToolKit, which consists of a library of routines to help with all aspects of building mixed reality applications. The whole system of the LLH consists on (figure 1):

- A computer work station, with the MXRToolKit libraries;
- Display system: a computer screen;
- Tracking system: a web cam and MXRToolKit square markers;
- Tangible part: The LEGO blocks that form the house/furniture and the dolls;
- Two speakers where sounds are reproduced.

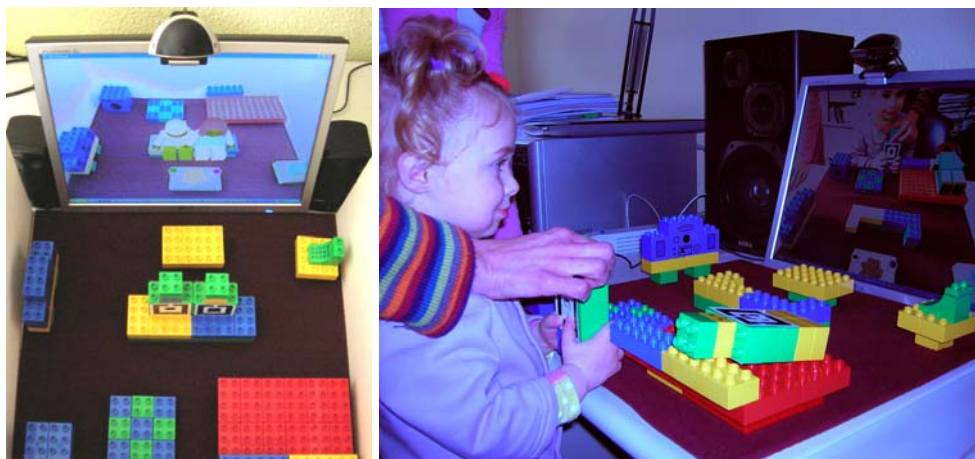


Figure 1: Left: LLH environment; Right: small child playing with her mother

MXR markers are necessary in order to run the application. Here are the total amount of multimedia files, frames and markers used in the LLH:

- Amount of multimedia files: 65
 - 3D Objects: 14
 - Sounds: 45
 - Videos: 6
- Amount of defined frames: 16
 - Dolls: 2

- Furniture: 14
- Amount of markers: 5
 - 4 real (dolls)
 - 1 virtual (rest of frames)

As can be seen, a huge variety of multimedia files are present in the game, with a view to increase the playful possibilities. The total amount (65) compared with the 4 single markers that the user can see is quite unusual in this kind of application as normally one multimedia file is related to a single marker (or even to a single set of markers), but this has the disadvantage that only a reduced amount of multimedia files can be used, as increasing the number of markers would slow down the system, or would even cause the system not to distinguish properly between markers (this is related to the camera resolution and the markers size). It would be quite hard to build an application with the MXRToolKit by using a total of 65 markers. Furthermore, for the LEGO Live House it was thought that using so many markers on the LEGO blocks would be unattractive. Also the problem of a user hiding markers while playing would be a great problem. Three steps were thought out in order to reduce the amount of markers:

1. The first step was to assign different multimedia files to the same marker. These multimedia files are rendered depending on the actions of the user. For example, when a user comes close to the radio, the radio is on and some music is played; the user can change the music by a rotation (pitch value: rotation about Y axis).
2. The second step was to 'remove' all the physical markers that remained still with respect to the camera. This is possible because the camera remains still (HMD is used) and is based on the mathematical principle that this kind of marker would have a constant transformation (rotation & translation) with respect to the camera. The objects that remain are also still: the furniture of the house, so the physical markers for them are not needed.
3. The last step was to join all the furniture markers into a single marker. This can be done because the program is never going to 'physically' find the marker. This single marker is referred as a 'virtual' marker.

Users can build the house as they wish. That means that furniture do not remain in the same position with respect to the camera depending on the users. But, as was explained before, the transformation matrix of furniture should be known beforehand, as markers are not placed on the objects. That means that whenever the application is started, the user should 'tell the program' where the real objects are placed. This step is called 'calibration procedure', where the user should tell the program where the objects are by means of the keyboard. When the doll is placed next to each of the objects, the user has to press the corresponding key; at that moment, the transformation matrix of the doll is passed to the furniture.

Some further aspects should be considered in order to build the MXR scenario properly for the LLH:

1. **Reflective surfaces of the LEGO blocks.** Reflected light is a problem in these kinds of applications, where tracking is based on marker recognition. LEGO blocks are quite reflective, especially light colours. To solve this problem, in the last version of LLH the surface of the ground was covered with deep-red velvet.

2. **Field of view of the camera:** As the markers on dolls are small (3.5 x 3.5 cm), the camera should be close to the house. The closer the camera is, the less part of the house is seen. Therefore, a web cam of a wide field of view is preferred. The Live! Cam Voice of Creative is in use. This is a USB 2.0 web cam with a video resolution of 1280 x 960 at 15fps and a FOV of 85°. Radial distortion is automatically corrected.
3. **Occlusion.** In the LLH the occlusion problem comes out. This is a typical problem in MXR technology, and arises when real objects are placed in between virtual objects and the camera, because the virtual objects are always rendered at the top, hiding any object that is in the video-image. To solve occlusion, sophisticated techniques are required (Breen et al., 1996; Fischer et al., 2004; Lepetit and Berger, 2000), which are not available within the MXRToolKit. There is a simple way to avoid occlusion in LLH: first of all, the camera should be placed at a higher position than the house (around 20 cm above), pointing down; secondly, the furniture of the house should not be that high (max. 3 blocks). With these two simple considerations, the occlusion can be avoided. This fact is illustrated in figure 2, where the real doll is placed in between a table (in red, at the back) and an armchair (in blue, at the foreground). In the first case, placing the camera too low makes evident the occlusion problem; in the second case, with the camera at a higher position, the occlusion is avoided.

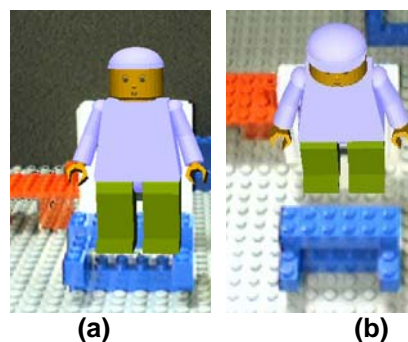


Figure 2: Image showing the occlusion problem: (a) with the camera at a lower position; (b) with the camera at a high position enough to avoid occlusion.

3.2 Interactivity

3.2.1 The house

The house is the interactive space where the MXR happens. The camera pointing to the house is like an eye that registers the actions of the dolls, and acts accordingly. Thus, it can be said that the house is 'alive', as different multimedia files will be shown depending on doll's behaviour, e.g. good actions (positive coexistence) will be rewarded, whereas bad actions (negative coexistence) will lead to some kind of punishment (e.g. prohibition of listening to the music while your partner is sleeping). Therefore, it can be said that rather than the users, the house is the one leading the game.

3.2.2 The dolls

Human-computer interaction is achieved through markers attached to real dolls. The development of amicable interfaces is a key goal to achieve, so that even users not familiar with computers can enjoy the advantages that they offer. This implies the elimination of the 'digital gap'. If someone is asked to communicate with a computer, they may think of a mouse, a keyboard or a joystick as the 'normal' interfaces. Most people probably do not ever think about a 'doll' as an interface. But this is exactly

what happens with the LLH: the mouse is changed by a doll, making the interaction more tangible. This is in fact a natural interface for children which they do not need to “learn” beforehand.

Interaction is triggered by two ways (table 1): by dolls’ proximity to the elements of the house or by dolls’ tilting (roll and pitch).

Furniture	Interaction	Actions	Multimedia Files
Washing machine	Proximity /	Put the washing machine / Prohibition	Sound: Washing machine on 3D: Water prohibition signal
Radio	Proximity / Roll	Turn on the radio / Tuning the radio / Dance / Prohibition	Sound: Different songs & radio programs 3D: Animated dolls dancing / Sound prohibition signal
Armchair	Proximity / Pitch	Sit down / Eat / Drink / Yawn	3D: Dolls sit down, drink & eat / Glasses with juice / A plate with cookies Sound: Eating / Drinking / Yawning
TV	Proximity /	Turn on the TV / Change TV program	Videos: Different videos
Bed	Proximity / Pitch	Put the pyjama on / Close eyes / Snore / Dream	3D: Dolls wear pyjamas / Animated dreams Sound: Snoring
Shower	Proximity	Undress / Have a shower / Sing	3D: Nude dolls / Water prohibition signal Sound: Singing / Falling water
Telephone	Proximity / Roll / Pitch	Make a call / Telephone ring / Prohibition	Sound: Making a call / Telephone ringing / Picking up the telephone / Say hello 3D: Sound prohibition signal

Table 1: Furniture related to interaction, possible actions and multimedia files.

The sequence of actions by a single doll also conditions multimedia files. This is technically managed with some kind of variables associated to the real objects and gives a ‘surprise’ element to the game. Therefore, users have the feeling that the multimedia files do happen randomly, but in fact they are not random.

Also, restricted actions are introduced to water and sound. Daily amounts of water are restricted, so dolls cannot have too many showers or put the washing machine on too many times per day; to have more water, they have to wait till next day (go to sleep!). But when a doll is sleeping, the other doll cannot switch the radio on or make a call.

3.3 Social, cultural and pedagogical issues

Social, cultural and pedagogical factors are implicit in the LLH via the multimedia files, which include sounds, videos and 3D-objects. Further pedagogical factors are implicit within some restrictions included in the game that somehow reward positive coexistence.

3.3.1 Sounds

Sound is a very important part of the game. Some of the sounds have been taken from the web page of the Spanish Ministry of Education and Science (2007), for example: falling water, singing in the shower, tuning the radio, telephone ring,

snoring, laughing, eating, etc. Songs related to the LLH radio are multicultural: samba, paso doble, rumba, waltz, sirtaki, tribal, traditional Chinese, etc. Afterwards, these sounds have been edited and others designed by electronic synthesis in order to enrich sound texture and timbre. This gives an interesting and 'funny' touch to the game that will catch children's attention. Moreover, each time a doll interacts with an item, the sound changes; for example, a total of 10 different rings are related to the telephone, and more than 30 songs are linked to the radio. The large variety of different sounds included in the game, makes the experience more playful and increases children curiosity.

3.3.2 Videos

The TV is one of the central elements in the game. Users will pay attention to what is happening inside. This element is also interactive, and it will show the augmented dolls (when TV is off) or some video files (when TV is on), depending on users' actions:

- If any doll is sitting down on their corresponding armchairs (yellow for LEGOWoman and blue for LEGOMan): the TV is off and the AR scenario is shown.
- If at least one doll is sitting down on its armchair: the TV is on and a video file is rendered.
- If both dolls are sitting down on their armchairs, only a common program that delights both dolls is seen.

The videos are also multicultural, which were recorded in Singapore.

3.3.3 3D-objects

The 3D objects (figure 3) were modelled with the 3DStudioMax software. Within this, some objects were animated (for example, when the dolls are sleeping). Afterwards, the objects were exported into the VRML format in order to be rendered by the MXRToolKit.

As is evident, the dolls have been modelled according to the original LEGO dolls. This means that no extra 'bending' has been introduced: for example, when the dolls are sitting down, the entire leg is a single block. Also no gender has been added: when the dolls are having a shower, both bodies look the same, and man/woman can only be distinguished because of the hair.

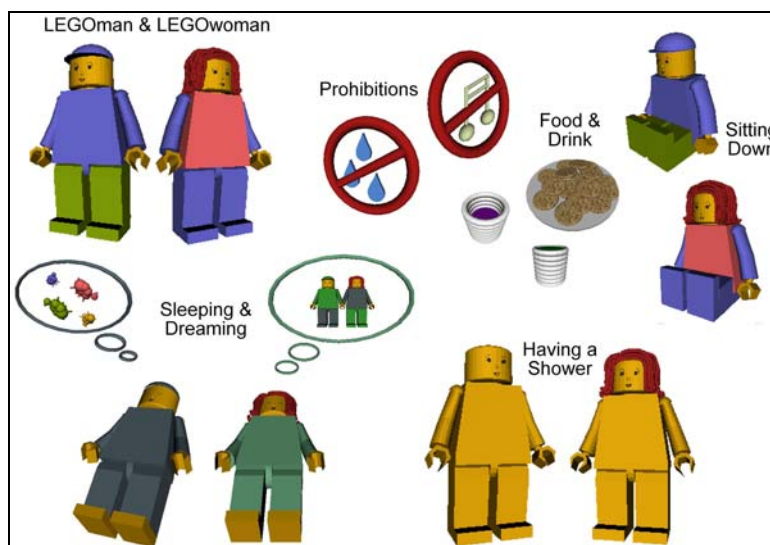


Figure 3: 3D virtual objects.

3.4 Collaboration

In LLH children collaboration is achieved face-to-face. Children will have to agree in some aspects, as the sequence of their actions condition the multimedia files that are rendered, and actions of a doll conditions the other doll. These dependencies are made in order to foment positive collaboration. Negative coexistence is avoided via some kind of restrictions and/or prohibitions related to sound and water. For example:

- If both dolls are watching TV, a program that delights both will be displayed. If children want to change a TV channel, they will have to agree (one doll should get up from the armchair).
- If one of the dolls is sleeping, the sound is not allowed, which means that the other doll cannot listen to the radio or make a telephone call.
- If one doll has too many showers in a single day, no more water is left, and the other doll cannot have a shower or put on the washing machine.
- If a doll sits in front of the table to have lunch, it only will be able to eat or drink, but not both at the same time. To have lunch properly, both dolls need to do the action at the same time.

Therefore some funny situations can arise: when the amount of water is finished, they have to wait till next day to have more, so both dolls have to go to sleep; but when a doll is sleeping, the other cannot switch the radio on or make a call. Therefore, children would need to agree in order to control the game properly. Mutual agreement will lead to positive coexistence, thus social factors are once more fomented.

3.5 Game scenarios

Game takes place in a mixture of real and virtual environments, in the real house and in the house shown on the screen. Commonly, AR is thought to be mainly visual. But it has to be noticed that, although video-based approaches are the primary interest in AR environments, sounds or other kind of media can be incorporated in order to enrich the experience. According to the virtual stimulus, the LLH scenario integrates the three possibilities, which are tactile, visual and auditory, as olfactory and gustatory do not have their virtual equivalent. According to Barfield et al. (1995), this

represents the most advanced multi-modal AR. That means, children can touch, see and hear in the LLH game. Figure 4 shows some samples of mutual coexistence.

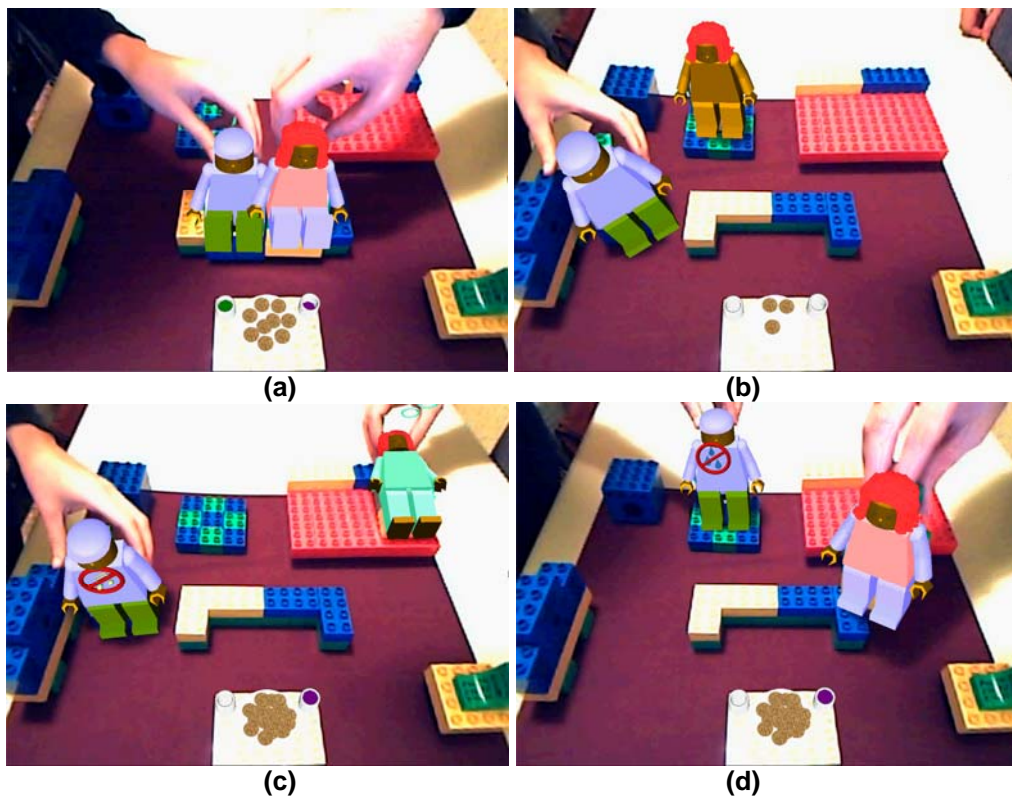


Figure 4: Mutual coexistence at the LLH scenario. (a): both dolls are having lunch together; (b): one doll is having a shower and the other is dancing next to the radio; (c): one doll is sleeping, so the other cannot make sound (radio off); (d): one doll tries to have a shower, but the amount of water per day has been exceeded.

4. Conclusion

In this paper, the Live LEGO House interactive game has been shown. This game is based on the MXR technology and has been thought to entertain children as well as to teach them in the area of coexistence. The learning process is not explicit, and positive and/or negative coexistence arise through exploration. Some advantageous characteristics of MXR that stand out in this game are collaborative reality and tangible interaction. These imply that children can play together and also interact with a computer in a natural way with no previous knowledge. LLH retains the basic LEGO model, allowing children to collaborate and use their imagination and creativity while playing, but adds some kind of multimedia files that enrich the game experience and foment other kinds of mutual collaboration and interaction.

Although some works have been developed within the MXR technology in the area of entertainment and/or education (Cheek et al., 2002; Andersen et al., 2004; Galantay et al., 2004; Brederode et al., 2005; Metaxas et al., 2005; Cheek et al., 2006), demonstrating a great potential in these fields, MXR for edutainment can still be considered a novel tool, and more research would be necessary, placing special emphasis on user experiences and learning processes.

References

- Andersen, T. L., Kristensen, S., Nielsen, B. W. and Grønbæk, K. (2004). 'Designing an Augmented Reality Board Game with Children: The BattleBoard 3D Experience'. In Proceeding of the 2004 conference on Interaction design and children: building a community. ACM Press, New York, NY, pp. 137-138.
- Barfield, W., Rosenberg, C. and Lotens, W. A. (1995). 'Augmented-reality displays. Virtual environments and advanced interface design'. Oxford University Press, Inc., New York, NY, pp. 542-575.
- Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A. and de Ridder, H. (2005). Powerball: The design of a novel mixed-reality game for children with mixed abilities. In Proceeding of the 2005 conference on Interaction design and children (IDC '05). ACM Press, New York, NY, pp. 32-39.
- Breen, D.E., Whitaker, R.T., Rose, E. and Tuceryan, M. (1996). 'Interactive Occlusion and Automatic Object Placement for Augmented Reality', Eurographics '96 Proceedings. Blackwell Publishers, Poitiers, France, pp. 11-22.
- Cheok, A. D., Teh, K. S., Nguyen, T. H. D., Qui, T. C. T., Lee, S. P., Liu, W., Li, Ch. C., Diaz, D., and Boj, C. (2006). 'Social and physical interactive paradigms for mixed-reality entertainment', Computers in Entertainment (CIE) archive. Volume 4, Issue 2. ACM Press, New York.
- Cheok, A. D., Weihua, W., Yang, X., Prince, S., Wan, F. S., Billingham, M., and Kato, H. (2002). 'Interactive Theatre Experience in Embodied + Wearable Mixed Reality Space', International Symposium on Mixed and Augmented Reality (ISMAR'02), pp 59-68.
- Fischer, J., Bartz, D., and Straer, W. (2004). 'Occlusion Handling for Medical Augmented Reality using a Volumetric Phantom Model'. In Proceedings of ACM Symposium on Virtual Reality Software and Technology (VRST), pp. 174 – 177.
- Galantay, R., Torpus, J., and Engeli, M. (2004). 'Living-room: interactive, space-oriented augmented reality'. Proceedings of the 12th annual ACM international conference on Multimedia. ACM Press, New York, pp 64-71.
- Lepetit, V. and Berger, M.-O. (2000). 'Handling Occlusions in Augmented Reality Systems: A Semi-Automatic Method'. In *Proceedings of IEEE and ACM International Symposium on Augmented Reality*, Munich, Germany. October, pp. 34-47.
- Liarokapis, F. (2006). An exploration from virtual to augmented reality gaming. *Simulation & Gaming*, Vol. 37, No. 4. SAGE Publications, pp. 507-533.
- Metaxas, G., Metin, B., Schneider, J., Shapiro, G., Zhou, W. and Markopoulos, P. (2005). 'SCORPIODROME: An Exploration in Mixed Reality. Social Gaming for Children'. In Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology (ACE'05). ACM Press, New York, NY, pp. 229-232.
- MXRToolKit documentation (2007). Available at: <http://mxrtoolkit.sourceforge.net/>. [Viewed 27th April, 2007].
- Spanish Ministry of Education and Science (2007). Available at: <http://recursos.cnice.mec.es/bancoimagenes/sonidos/index.php>. [Viewed 27th April, 2007].
- Wikipedia (2007). Available at: <http://www.wikipedia.org/>. [Viewed 27th April, 2007].

LIVE LEGO HOUSE, AN INTERACTIVE SPACE TO EXPLORE COEXISTENCE THROUGH GAMING

Cristina Portalés Ricart

Laboratorio de Luz, Universidad Politécnica de Valencia
Valencia (Spain)
criporri@doctor.upv.es

Abstract

Gaming is a natural way of learning. Live LEGO House is an interactive space to explore coexistence through gaming based on the Mixed Reality (MXR) technology. The system consists basically on a physical/real house built with the LEGO blocks [7], which is enriched with different non-physical/virtual multimedia files, including sounds, videos and 3D animations. Children interact with the house via two dolls, each of them having different characteristics and needs. They will need to agree in some aspects in order to rule the game properly, to coexist.

Keywords

Mixed Reality, interaction, edutainment, technology.

1. INTRODUCTION

Edutainment (also educational entertainment or entertainment-education) is a form of entertainment designed to educate as well as to amuse. Edutainment typically seeks to instruct or socialize its audience by embedding lessons in some familiar form of entertainment: television programs, computer and video games, films, music, websites, multimedia software, etc. There are also lessons that use edutainment as a basis for teaching in a more efficient and faster way. Most often, edutainment seeks either to tutor in one or more specific subjects, or to change behaviour by engendering specific socio-cultural attitudes. Successful edutainment is discernible by the fact that learning becomes fun and teachers or speakers educate an audience in a manner which is both engaging and amusing [12].

Multimedia technologies have demonstrated a great potential in the field of education, since they offer an excellent technological base for the training and transmission of knowledge, which generates more efficient methodologies of education and learning. Equally in the area of entertainment and leisure, these technologies open new perspectives that offer new chances in this important human facet.

Live LEGO House (LLH) is an interactive game for children based on the MXR technology, which takes great advantage of the multimedia systems (sounds, images, videos, etc.). It is thought to entertain as the same time as to educate in the area of coexistence. Nevertheless, education is not explicit, as no rules are necessary to play. Two dolls – leaded by two children – will have to interact with the house, discovering a place plenty of sounds, videos and 3D virtual objects that will respond to their acts. But actions do have consequences, so they have to decide what to do: what they want or what seems to be better for both.

I developed this application while I was in Singapore, where I stayed as a research intern at the Mixed Reality Lab [10]. The idea of developing a game based on coexistence was born there. One of the characteristics of Singapore is the actual coexistence between people from different races (Chinese, Malaysian and Indian) and religions (Buddhism, Islamism, Catholicism, Hinduism, etc.), and even different mother languages (Mandarin, Malay and Tamil), having all of them English as a unifier. Although there are some characteristic ghettos (as Chinatown), cultures are quite mixed. Some examples of this coexistence are: people from different countries live and work together, respecting their culture and traditions; a Hindu Temple can be found in Chinatown; all kind of traditional food is offered at the stalls of a single hawker centre; different traditions are celebrated during the whole year; different languages are in use with the English as link; etc. Nevertheless, an “equal” coexistence is still not a fact, but this seems to be more a problem of social class and/or poverty than of cultural

roots. Some of these coexistence facts are shown in LLH by means of some videos that have been recorded there.

2. LLH ENVIRONMENT

2.1 Edutainment through MXR

MXR refers to the combination of computer graphics and real-world objects. This encompasses both augmented reality, which involves placing computer graphics objects into the real world and augmented virtuality, which involves placing real-world objects into virtual environments.

MXRToolKit [11] consists of a library of routines to help with all aspects of building mixed reality applications. The philosophy of the SDK is to keep the interface extremely simple. Therefore, one of the advantages of the MXRToolKit compared with other marker based programs is the fastness to write your own applications – even for non-programmer experts – and integrate different kind of multimedia files (sound, videos, 3D objects, animations, etc) in a single project.

Comparing MXR applications with Virtual Reality, two major advantages stand out:

1. **Collaborative reality:** mixed reality allows users to interact with graphical objects while still letting them to see the real world. Several users can collaborate with a computer generated objects as if they were real world items, thus leading to a natural human-human interaction.
2. **Tangible interaction:** Computer graphical objects can be attached to real-world physical objects by means of markers. By manipulating the position of the real object, we can correspondingly manipulate the position of the virtual object. This is known as “tangible interaction” and provides an extremely natural interface for manipulating computer generated objects which does not need to be learnt by the user.

The game can be divided in the following steps:

1. Construct your own house: use your imagination and creativity to build a house within the LEGO blocks.
2. Tell the program where the objects are: this is called the calibration process. Simply by pressing some buttons on a prepared keyboard you tell the pc how your house is arranged. Nevertheless, this step is perhaps too difficult for small children that will need some help.
3. Explore the matters of coexistence and try to solve them wisely: you will discover how to play having into account your partner.

2.2 The interactive space

The house is the interactive space where the MXR happens (Fig. 1). The whole system consists on:

- A PC screen, where a TV is displayed showing the Augmented Reality;
- A web cam pointing to the house;
- The LEGO blocks that form the house/furniture and represent the tangible part of the application;
- The dolls that are also built with LEGO blocks and having the markers for the MXR, that represent the interface between user and computer and are also tangible;
- Two speakers where sounds are reproduced.

The camera pointing to the house is like an eye that registers the actions of the dolls, and acts accordingly. Thus, it can be said that the house is ‘alive’, as different multimedia files will be shown depending on doll’s behaviour, e.g. good actions (positive coexistence) will be rewarded, whereas bad actions (negative coexistence) will lead to some kind of punishment (e.g. prohibition of listening to the music while your partner is sleeping). Therefore, it can be said that rather than the users, the house is the one leading the game.



Fig. 1: The interactive scenario.

2.3 The interface user-computer

The development of amicable interfaces is the key to achieve that even users not familiar with computers can enjoy the advantages that they offer them. This implies the elimination of the "digital gap". If you are asked to communicate with a computer, it suddenly comes to your mind a mouse, a keyboard or a joystick as the 'normal' interfaces. You probably do not ever think about a 'doll' as an interface. But this is exactly what happens with the LLH (Fig. 2): the mouse is changed by a doll, making the interaction really easy as one only has to play around with the dolls to make everything happens. Dolls are made with the LEGO blocks as well, and two MXR markers are placed on each doll, one at the front and one at the back. This is in fact a natural interface for children that don't need to be learnt beforehand, i.e., interaction will happen with the only need to play in the 'real' house, they even don't need to think about computers.

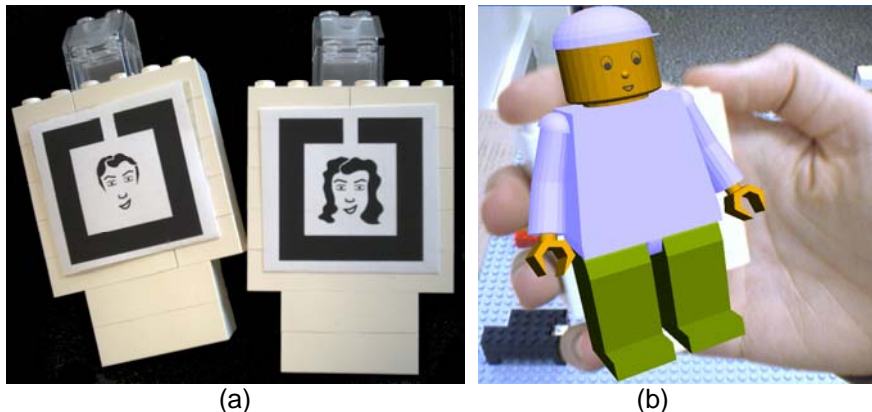


Fig. 2: The dolls: (a) LEGOman and LEGOwoman with attached markers; (b) User carrying the augmented LEGOman.

2.4 Markers and calibration

MXR markers are necessary in order to run the application. Here are the total amount of multimedia files, frames and markers used in the current version of the LLH:

- Amount of multimedia files: 65
 - 3D Objects: 14
 - Sounds: 45
 - Videos: 6

- Amount of defined frames: 16
 - Dolls: 2
 - Furniture: 14
- Amount of markers: 5
 - 4 real (dolls)
 - 1 virtual (rest of frames)

The most important part of this application is the amount and different kind of multimedia files that are shown during the game. The total amount (65) compared with the 4 single markers that the user can see is quite unusual in this kind of applications. Normally, one multimedia file is related to a single marker (or even to a single set of markers), but this has the disadvantage that only a reduced amount of multimedia files can be used, as increasing the number of markers would slow down the system, or even would cause the system not to distinguish properly between markers (this is related to the camera resolution and the markers size). It would be quite hard to build an application with the MXRToolKit by using a total of 65 markers. Furthermore, for the LEGO Live House it was thought that using so many markers on the LEGO blocks would be unattractive. Also the problem of user hiding markers while playing would be a great problem.

Three steps were thought in order to reduce the amount of markers:

1. The first step was to assign different multimedia files to a same marker. These multimedia files are rendered depending on the actions of the user. For example, when a user comes close to the radio, the radio is on and some music is played; user can change the music by a rotation (pitch value: rotation about Y axis).
2. The second step was to 'remove' all the physical markers that remained still with respect to the camera. This is possible because the camera remains still (not HMD is used) and is based on the mathematical principle that this kind of markers would have a constant transformation (rotation & translation) with respect to the camera. The objects that remain also still are the furniture of the house (all out of the dolls), so the physical markers for them are not needed.
3. The last step was to join all the furniture markers into a single marker. This can be done because the program is never going to 'physically' find the marker. This single marker is referred as a 'virtual' marker.

Users can build the house as they wish. That means that furniture do not remain in the same position with respect to the camera depending on the users. But, at it was explained before, the transformation matrix of furniture should be known beforehand, as markers are not placed on the objects. That means that whenever the application is started, the user should 'tell the program' where the real objects are placed. This step is called 'calibration procedure', where the user should tell the program where the objects are by means of the keyboard (table 1). In the following table the keyboard buttons are shown for each doll. Note that the doll should be placed next to each of the objects and at that position the user should press the indicated button. In this way, the transformation matrix of the doll is passed to the furniture.

LEGOWoman	LEGOman	Furniture
A	a	Washing machine
B	b	Radio
C	c	Armchair
D	d	TV
E	e	Bed
F	f	Shower
G	g	Telephone

Table 1: Keyboards related to furniture.

3. MULTIMEDIA FILES

3.1 Interaction

In the following lines, the way the interaction is produced is explained:

- Interaction is triggered by two different ways (table 2 and Fig. 3):
 - Proximity to the elements of the house: when dolls are close to different real elements of the house, some multimedia files related to them are activated.
 - Dolls tilting (roll and pitch) also activate different multimedia files.
- The sequence of actions by a single doll also conditions multimedia files. This is technically managed with some kind of variables associated to the real objects and gives a ‘surprise’ element to the game. Note that users have the feeling that the multimedia files do happen randomly, but in fact they are not.
- Actions of a doll condition other doll:
 - If both dolls are watching TV, a common program will be displayed.
 - If one of the dolls is sleeping, the sound is not allowed, that means that the other doll cannot listen to the radio or make a telephone call.

Also restricted actions are introduced regarding to water and sound: daily amount of water is restricted, so dolls cannot have too many showers or put the washing machine too many times per day; Then, to have more water, they have to wait till next day (go to sleep!). But when a doll is sleeping, the other cannot switch the radio on or make a call.

Furniture	Interaction	Actions	Multimedia Files
Washing machine	Proximity /	Put the washing machine / Prohibition	Sound: Washing machine on 3D: Water prohibition signal
Radio	Proximity / Roll	Turn on the radio / Tuning the radio / Dance / Prohibition	Sound: Different songs & radio programs 3D: Animated dolls dancing / Sound prohibition signal
Armchair	Proximity / Pitch	Sit down / Eat / Drink / Yawn	3D: Dolls sit down, drink & eat / Glasses with juice / A plate with cookies Sound: Eating / Drinking / Yawning
TV	Proximity /	Turn on the TV / Change TV program	Videos: Different videos
Bed	Proximity / Pitch	Put the pyjama on / Close eyes / Snore / Dream	3D: Dolls wear pyjamas / Animated dreams Sound: Snoring
Shower	Proximity	Undress / Have a shower / Sing	3D: Nude dolls / Water prohibition signal Sound: Singing / Falling water
Telephone	Proximity / Roll / Pitch	Make a call / Telephone ring / Prohibition	Sound: Making a call / Telephone ringing / Picking up the telephone / Say hello 3D: Sound prohibition signal

Table 2: Furniture related to interaction, possible actions and multimedia files.

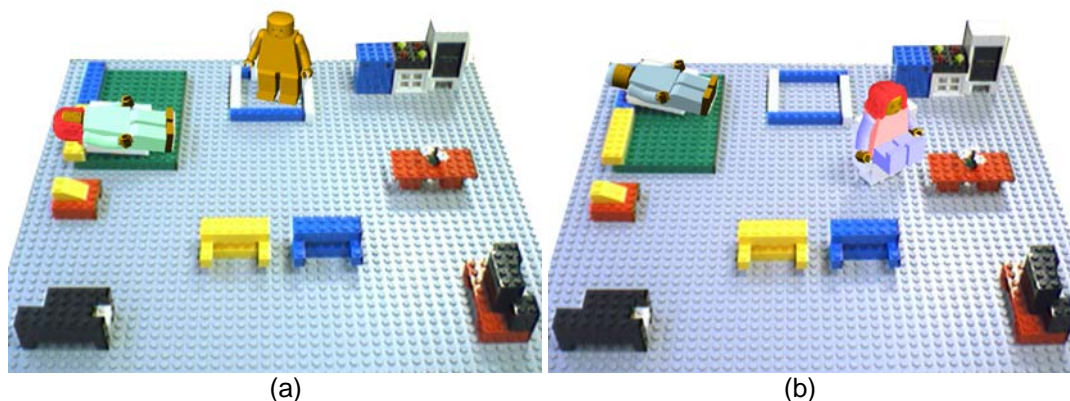


Fig. 3: (a) LEGOman is having a shower while LEGOwoman is sleeping;
(b) LEGOwoman is sitting quiet while LEGOman is sleeping.

3.2 Sounds

Sound multimedia files are mainly taken from the web page at [9]. In this web page, a set of free multimedia files for education purposes can be found, including sounds, videos and images. Between others, the following sounds are included in the game: falling water, singing in the shower, tuning the radio, telephone ring, snoring, laughing, eating, etc.

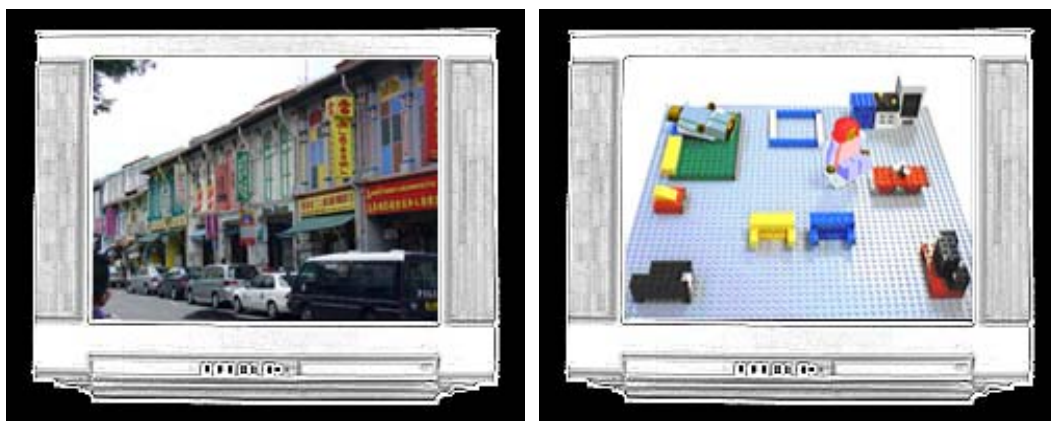
3.3 Videos

The TV is maybe the central element in the game (Fig. 4). Users will pay attention to what is happening inside. This element is also interactive, and it will show the augmented dolls (when TV is off) or some video files (when TV is on), depending on users' actions:

- If at least one doll is sitting down in front of the TV:
 - The TV is on
 - A video file is rendered
- Otherwise:
 - The TV is off
 - The AR scenario is shown, as the house was reflected on a real TV-screen
- If both dolls are sitting, only the common program that delights both dolls is seen:

Doll 1	Doll 1&2	Doll 2
Program 1	Program 1	Program 1
Program 2		Program 4
Program 3		Program 5

Table 2. Programs associated to dolls.



(a)

(b)

Fig. 4: TV as displayed in the PC screen: (a) the TV is on; (b) the TV is off.

3.4 Three-dimensional objects

The 3D objects (Fig. 5) were modelled with the 3DStudioMax software. Within this, also some objects were animated (for example, when the dolls are sleeping). Afterwards, the objects were exported into the VRML format in order to be rendered by the MXRToolKit.

Dolls have been modelled according to the original LEGO dolls. That means that not extra 'bending' has been introduced: for example, when the dolls are sitting down, the entire leg is a single block. Also no gender has been added: when the dolls are having a shower, both bodies look like the same, and man/woman can only be distinguished because of the hair.

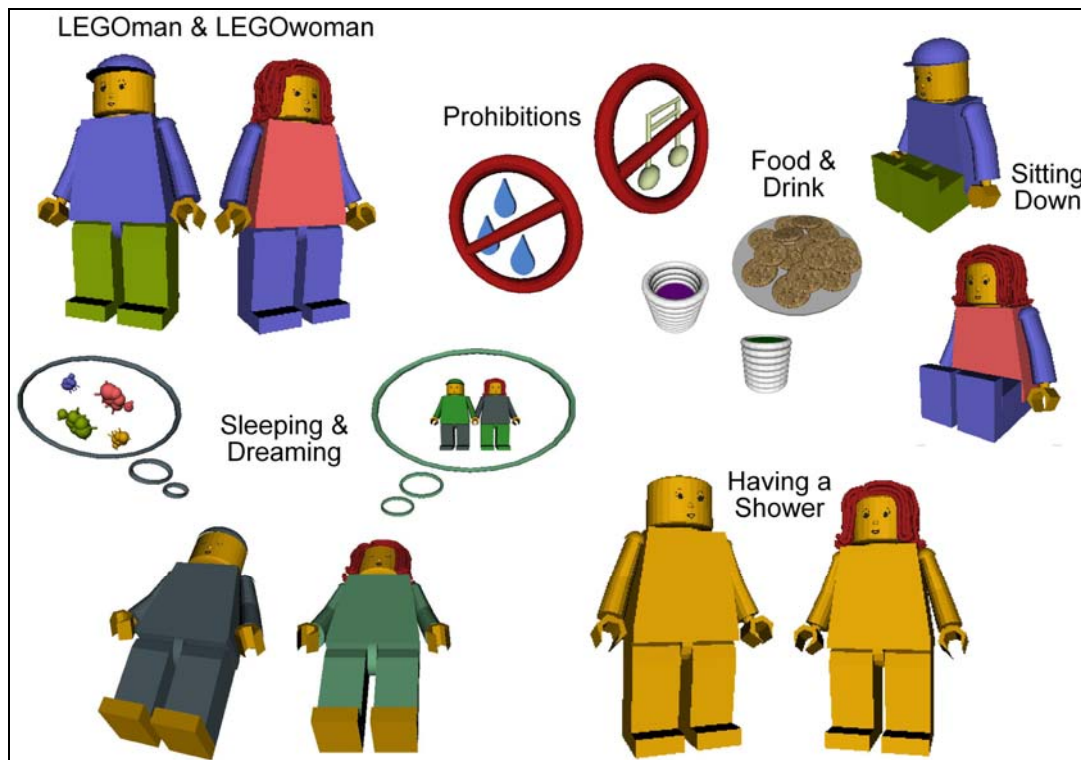


Fig. 5: 3D virtual models.

4. GENERAL ASPECTS

Some general aspects should be considered in order to properly build the MXR scenario for the LLH:

1. **Reflective surfaces of the LEGO blocks:** Reflected light is a problem in these kind of applications, where pose tracking is based on marker recognition. LEGO blocks are quite reflective, especially light colours. To solve this problem, in the last version of LLH the surface of the ground was covered with deep-red velvet.
2. **Field of view of the camera:** As the markers on dolls are small (3.5 x 3.5 cm), the camera should be close to the house. The closer the camera is, the less part of the house is seen. Therefore, a web cam of a wide field of view is preferred. The Live! Cam Voice of Creative [4] is in use. This is a USB 2.0 web cam with a video resolution of 1280 x 960 at 15fps and a FOV of 85°. Radial distortion is automatically corrected.
3. **Occlusion:** in the LLH the occlusion problem comes out. This is a typical problem in MXR technology, and arises when real objects are placed in between virtual objects and the camera, because the virtual objects are always rendered at the top, hiding any object that is in the video-image. To solve occlusion sophisticated techniques are required [1; 5; 8], which are not available within the MXRToolKit. Anyway, there is a simple way to avoid occlusion in LLH: first of all, the camera should be placed at a higher position than the house (around 20 cm above), pointing down; secondly, the furniture of the house should be not that high (max. 3 blocks); within this two simple considerations, the occlusion can be avoided. This fact is illustrated in Fig. 6, where the real doll is placed in between a table (in red, at the back) and an armchair (in blue, at the foreground). In the first case, placing the camera too low makes evident the occlusion problem; in the second case, with the camera at a higher position, the occlusion is avoided.

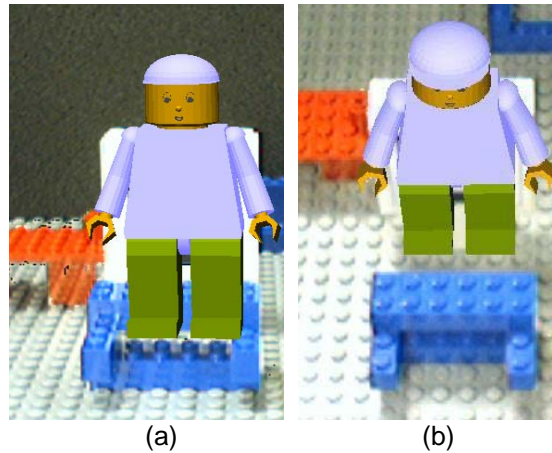


Fig. 6: Image showing the occlusion problem: (a) with the camera at a lower position; (b) with the camera at a high position enough to avoid occlusion.

5. CONCLUSION

In this paper, the Live LEGO House interactive game has been shown. This game is based on the MXR technology and has been thought to entertain children as well as to teach them in the area of coexistence. Learning process is not explicit, and positive and/or negative coexistence arise thought exploration. Some advantageous characteristics of MXR that stand out in this game are collaborative reality and tangible interaction. These imply that children can play together and also interact with a computer in a natural way with no previous knowledge.

Although some works have been developed within the MXR technology in the area of entertainment and/or education [2; 3; 6], demonstrating a great potential in these fields, MXR for edutainment can still be considered a novel tool, and more research would be necessary putting special interest in user experiences and learning processes.

References

- [1] Breen, D.E., Whitaker, R.T., Rose, E. and Tuceryan, M. "Interactive Occlusion and Automatic Object Placement for Augmented Reality", Eurographics '96 Proceedings. Blackwell Publishers, Poitiers, France, August 1996. pp. 11-22.
- [2] Cheok, A. D., Teh, K. S., Nguyen, T. H. D., Qui, T. C. T., Lee, S. P., Liu, W., Li, Ch. C., Diaz, D., and Boj, C. "Social and physical interactive paradigms for mixed-reality entertainment", Computers in Entertainment (CIE) archive. Volume 4, Issue 2. ACM Press, New York, 2006.
- [3] Cheok, A. D., Weihua, W., Yang, X., Prince, S., Wan, F. S., Billingham, M., and Kato, H. "Interactive Theatre Experience in Embodied + Wearable Mixed Reality Space", International Symposium on Mixed and Augmented Reality (ISMAR'02), 2002, pp 59-68.
- [4] Creative official web site: <http://www.creative.com/>
- [5] Fischer, J., Bartz, D., and Straer, W. "Occlusion Handling for Medical Augmented Reality using a Volumetric Phantom Model". In Proceedings of ACM Symposium on Virtual Reality Software and Technology (VRST), 2004, pp. 174 – 177.
- [6] Galantay, R., Torpus, J., and Engeli, M. "Living-room: interactive, space-oriented augmented reality". Proceedings of the 12th annual ACM international conference on Multimedia. ACM Press, New York, 2004, pp 64-71.
- [7] LEGO official web site: <http://www.lego.com>
- [8] Lepetit, V. and Berger, M.-O. "Handling Occlusions in Augmented Reality Systems: A Semi-Automatic Method". In *Proceedings of IEEE and ACM International Symposium on Augmented Reality*, Munich, Germany. October 2000. 34-47.
- [9] Ministerio de Educación y Ciencia (España). Banco de sonidos: <http://recursos.cnice.mec.es/bancoimagenes/sonidos/index.php>
- [10] Mixed Reality Lab: <http://www.mixedreality.nus.edu.sg/>
- [11] MXRToolKit documentation: <http://mxrtoolkit.sourceforge.net/>
- [12] <http://www.wikipedia.org/>

The Private City through the Hot Images

ABSTRACT

Hot Images is an artistic mixed reality application that deals with the relation between human beings and city environments, thus proposing a novel cartography and navigation tool for the city. Within the virtual recreated environment of the Hot Images, the inert city is closed to the human field, transforming the cold aspect of the city into the hotness of the close friends and relatives. The user is carrying a laptop where a 3D virtual model of the city is presented within a blue color. Some images of past events related to the user are hidden in the city, and s/he has to find them. As the user gets closer to the images locations, the environment changes to red color. Once the user arrives to a location, s/he can catch some images that will appear mapped on some of the buildings' facades, thus changing the aspect of the city. The location of the user is achieved with a combination of a GPS and an inertial sensor. The interface is done with the Max/MSP Jitter software.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation (HCI)]: Multimedia Information Systems – *Artificial, augmented, and virtual realities.*

General Terms

Algorithms, Performance, Design, Experimentation, Human Factors, Languages.

Keywords

Mixed Reality, Location Based Services, Color Navigation, Urban Environments, Human Space.

1. INTRODUCTION

The world is complex, dynamic, multidimensional; the paper is static, flat. How are we going to represent the rich visual world of experience and measurement on mere flat? [10]. This is one of the questions that most worried the cartographers since the cartography exists. In the last years, the support for maps has

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference'04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

evolved from a plain paper to the screen of a computer and traditional cartography has evolved to multimedia cartography [2]. Nowadays we can represent geospatial data through 3D models, and even visualize dynamic phenomena with some kind of simulators. The emergent techniques of virtual and augmented reality [5, 7] have enriched the maps, thus giving new tools for the user orientation.

The main purpose of a map is not other than to orient the users. In this sense, cartography has also evolved with the incorporation of the GPS and inertial sensors, which provide a navigation system for the so called location based services (LBS) and pedestrian navigation systems [1, 3, 4, 9]. But there are maps with other different purposes, as those intended to visualize some amount of data related to the physical space (as for example, population, temperature, precipitation, etc.). Others are intended to classify the soil according to different criteria (natural resources, human activities, etc.). Topographic maps show the relief, whereas geologic maps show the underlying materials. At the time of making a map – which is a simplification of the reality – it is very important to know who are going to use it (general users, expertise users, children, etc.); depending on that, the cartographic encoding/representation would be different. Then, could it be possible to do a customized map of the city, whose main purpose is to approach the city to the user personal field?

The Hot Images application tries to fulfill this task. I developed this application while I was living in a residential area called Beeston, characterized for not having anything special. It can be said that this area has low 'imageability'. This term was introduced by Lynch [8] and can be described as the quality of a physical object, which can evoke in any observer a strong, vivid image. Then, a highly imageable city would be well formed, would contain very distinct parts, and would be instantly recognizable to the common inhabitant. The low temperatures at winter time made that the streets remained almost desert, especially at evening/night, so the daily activity of people was mainly indoors. Therefore, the city seemed to be cold and inert, like deprived of life. It was in this cold area where I met the idea of approaching the city to the personal scope, due to the necessity to give a distinguishing touch to the singsong surroundings thus changing the perception of the city.

2. THE APPLICATION

Hot Images is a mixed reality application in which two different spaces are confronted: the exterior space, impersonal; and the private space, personal. The personal experiences are mixed with the urban surroundings by means of some images related to

certain past events. These images are hidden in certain locations of the city. The user will have to explore his surroundings in order to find and recover them. In this way, the user is forced to travel around the entire area in order for him to meet new places. It is a way of exploring the city. Users discover the city and, at the same time, build their own private map. The way of finding the images can be compared to the popular child game of 'hot-cold', which consists of finding certain objects hidden in the surroundings. The one that has hidden the object is the one in charge to guide the player with these two words: 'cold' means that he is far away from the hidden object; 'hot' means that he is near to the objective. Hot Images uses this role to guide the user in his task of finding the images (the hidden object). The game relates the distance (cold) to the blue color and the proximity (hot) with the red color, thus given the user the clues to walk along the city streets (figure 1).

When the user arrives to a location where a set of images are hidden he has to look to the surroundings to find them, which are then mapped into some of the buildings' facades. Thus, the final mission of the user will be to turn the blue/cold city into a red/hot place. Then, the inert surroundings acquire human connotations.

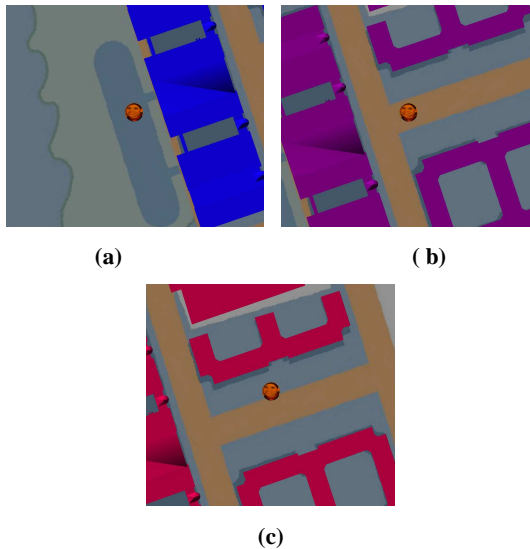


Figure 1 - Buildings' color depending on the distance to a hot location, with: a) > 100 m: cold; b) 30 m; c) < 10 m: hot.

3. TECHNICAL ASPECTS

3.1 External Devices

The mixed reality scenario is built with the following devices:

- Laptop: the user carries a laptop in which he sees the 3D model of the city.
- WebCam: A webcam is needed in order to find the images hidden in the surroundings, which is achieved with a pattern recognition algorithm. A camera resolution of 640 x 480 ppp is enough.
- GPS receiver: To achieve user location in real time, the Trine II GPS receiver was used. This receiver has the advantage that it can be connected to four different devices at the same

time via Bluetooth. Then, a collaborative scenario could be built in which different users share the experience.

- Inertial sensor: for the orientation and tilting information, the MT9 Xsens inertial sensor was used. Data can be received in form of Euclidean angles, rotation matrix or quaternion.



Figure 2. A user running the application.

3.2 Software Environment

The application runs under the Max/MSP Jitter software. Inside the program, the different routines are organized into patches and sub-patches (figure 3).

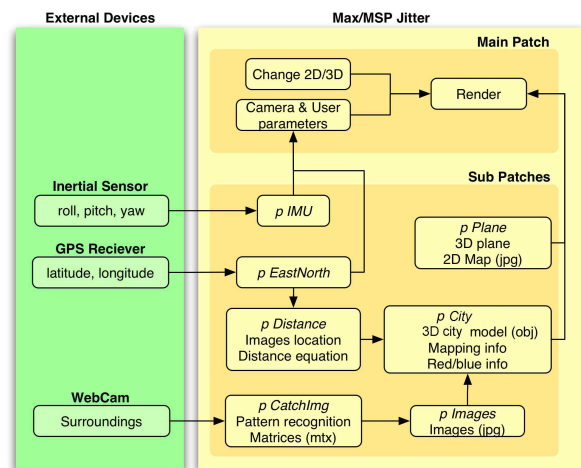


Figure 3. Application architecture.

The main patch consists of a window where the main tasks are performed. This is the actual window that the user can see when the application starts, and contains the following objects:

- *Change 2D/3D*: The user can choose at any time whether to visualize the application in the 2D or 3D view. In the 2D view, he sees the surroundings as in a map, thus he is able to see the buildings that remain out of his actual sight; at this view, the tilting information is not used, and the centre of the map is his actual location. The heading information can be activated or deactivated; in the first case, it leads to an egocentric map. In the 3D view, the model is shown according to the user real point of view. Tilting and heading information are used in this case.
- *Camera and user parameters*: the location is set with the positioning (Easting, Northing and height), the tilting (pitch and roll) and the heading information. In the 3D view, the

user parameters are set to coincide with the camera ones, whereas in the 2D view the height is set different, thus the camera is placed at the top looking to the nadir and pointing to the user. Therefore, in the 2D view, the centre of the map is the location of the user. Users are represented as simple spheres mapped with a picture of their own face, so they are the ones exploring the virtual world.

- *Render*: This is the OpenGL window where the models and images are rendered.

The sub-patches are some windows that the user is not able to see, but that are necessary in order to run the application. These are:

- *p IMU*: In this patch the data from the inertial sensor is received.
- *p EastNorth*: This patch transforms the received latitude and longitude values from the GPS into local X and Y coordinates.
- *p Distance*: Here the distance from the user to all the images is calculated. The shortest one is transformed into a color (from red to blue). When the user is more than 100 m far away from a target, the scene is seen completely blue; as the user get closer to an image, the 3D world turns more and more red. This patch will act as if it were circular buffers defined around images (figure 4).

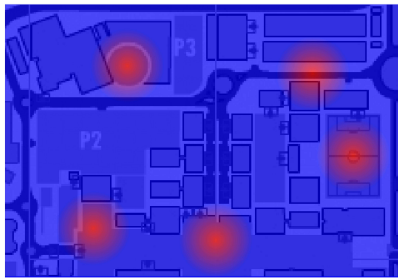


Figure 4. Hot Locations at Vera Campus.

- *p Catching*: In this patch, a pattern recognition procedure is applied to the incoming image, in order to find the correspondence with a stored matrix. When the matrix and the webcam image match, then the user recovers one of the hidden images and is automatically mapped to some of the buildings' facades.
- *p Plane*: This patch contains an OpenGL plane with its texture being the map of the area. For the example shown (figures 1 and 3), a map vectorized in a cartoon style was used.
- *p City*: a 3D model of the surroundings is needed in the *obj* format in order to be read by the program. The model consists on a simple representation of the buildings but without texture. In its place they appear colored in tones that vary from the red to the blue, indicating the distance or proximity to the images. The mapping information (i.e., how to map) is given to them beforehand (for example, with the 3DStudioMax before being exported to *obj*), although the texture (i.e., what to map) is applied afterwards in the Max/MSP Jitter.

- *p Images*: In this patch the images are called. Once the user arrives to a hot location and finds correspondence with the incoming webcam image, the hidden image is called and automatically mapped.

4. TESTING THE SYSTEM

The application was first tested at the campus of the Polytechnic University of Valencia (Spain). People working at the Laboratorio de Luz were asked to be involved in testing the system, so they were already familiar with the campus. The content of the images was selected in order to give the audience a different perception of this environment. All images showed people inside the buildings of the University in different situations (studying, working, eating, etc.). They tell us diverse personal stories that happen in the interior of the buildings, thus making 'visible' from the outside the human relationship with the inert constructions. In the following figure, some of the selected images can be seen.



Figure 5. Images at the inside of the buildings.

As the user is approaching a hot location, the environment turns into red (figure 6 - top). Images related to that hot location are then mapped into some of the buildings (figure 6 - middle). Once found, images will remain visible even if the user goes further away (figure 6 - bottom).



Figure 6. Shortcuts.

Some of the audience noticed that the proposed game of 'hot-cold' is an interesting navigation tool for exploring the campus, and the relationship between proximity and color resulted to be quite intuitive. The game of hot-cold was familiar to all users

participating in the performance, therefore only a brief introduction was needed to explain the role of the proposed game.

It has to be pointed that color changes in a gradual way, there are not jumps between colors. As a navigation tool, it can be said that when using color in this way is not fast to arrive to a location, as there are not arrows or other clear indications that guides the users through a specific direction. These means, the user has to really explore the environment, and destination is not known until he arrives. At the beginning, audience was a little bit confused about this fact, because they expected a straight forward navigation, and also were somehow anxious to know the destination. Some users had also problems finding hot locations due to color, as they thought to have arrived to a hot location before the environment had entirely turned into red.

5. BRIEF DISCUSSION

Image content can be discussed. Depending on the expected audience and the purpose of the exhibition, different contents can be chosen. Very different topics can be thought, for example: images about the history of a city or an specific past event related to the city, if the application is to be shown for a general audience living or arriving there; images of fauna/flora, if the purpose is to denunciate the abuse of constructions to the detriment of wild/green areas; images of a private party or common past event, if the audience is a specific group; etc. Using 2D images to construct a virtual 3D world is not new, and a good example can be found in the projects 'Field-Works' of Masaki Fuhata.

Whether the hot-cold game is known in different cultures or if red can be easily related to hot and blue to cold is a fact that would need some further discussion. In order to explore into the universality of these facts, I prepared a short test and asked some friends from different countries. The questions were:

1. Are you familiar with the hot-cold game?
 - a. Yes
 - b. No
2. Which option has more sense to you?
 - a. Hot can be related to red, and cold to blue
 - b. Hot can be related to blue, and cold to red
 - c. I cannot relate hot-cold with any color
 - d. I would relate hot-cold with other colors (say which)

People participating in the test come from: Germany, Austria, Italy, Brazil, Finland, USA, Singapore, China, and of course, Spain. To the first question, all of them answered 'a' out of the one coming from China. To the second question, the answer was 'a' in all cases, but the Chinese guy specified 'dark blue'. Although this short test can't be seen as a rigorous surveying, it gives an idea about the possibility to extend the proposed navigation tool into other cultures.

Anyway, whether audience is familiar with the hot-cold game or not, shouldn't be a problem when understanding the navigation tool, as the game is very simple and intuitive, thus needing short explanation. Whether the chosen red-blue colors can be understood as hot-cold is a question of visual culture [6], and maybe people from other countries would use other colors to describe it. If this were the case, it would very simple to change the color information in the Hot Images application in order to be better understood by the audience.

6. CONCLUSION

Although the main purpose of a map is to orient the users, a huge variety of facts are visualized within cartography, and lots of purposes can be defined. In the Hot Images application, a new relationship between users and cartography has been proposed, thus altering the perception of the environment: the impersonal city is closed to the personal scope by virtually changing the look of its buildings. A popular childish game is proposed as a navigation tool: users have to find some images hidden at specific locations (hot locations) with the only clues of 'hot-cold', which are visualized with the colors 'red-blue'.

Therefore, users are guided through the city according to these color clues. The aim is to turn the inert environment into a hot place, providing the buildings with human connotations. At hot locations images related to the audience are hidden. Then, when a hot location is reached, some images are discovered and mapped on some of the buildings' facades, thus acquiring human connotations.

The application is build into the Max/MSP Jitter environment. For the location of the user, a GPS receiver and an inertial sensor are in use. A webcam is also needed in order to find the hidden images that are recognized via a pattern recognition algorithm.

7. REFERENCES

- [1] Aslan, I., and Krüger, A. The Bum Bag Navigator (BBN): An advanced pedestrian navigation system. In AIMS, Nottingham, UK, 2004.
- [2] Cartwright, W., Peterson, M., and Gartner, G.. *Multimedia Cartography*. Berlin: Springer Verlag, 1999.
- [3] Fang, L., Antsaklis, P. J., Montestrucque, L., McMickell, M. B., Lemmon, M., Sun, Y., Fang, H., Koutroulis, I., Haenggi, M., Xie, M., and Xie, X. Design of a Wireless Assisted Pedestrian Dead Reckoning System – The NavMote Experience. *IEEE Trans. Instrumentation and Measurement*, Vol. 54, 2005, 2342-2358.
- [4] Gartner, G. Location-based mobile pedestrian navigation services – the role of multimedia cartography. ICA UPIMap, Tokyo, 2004.
- [5] Hedley, N. R., Billingham, M., Postner, L., and May, R. Explorations in the use of Augmented Reality for Geographic Visualization, *Presence: Teleoperators and Virtual Environments*, Volume 11 (2). The MIT Press, Cambridge, 2002, 119-133.
- [6] Heller, E. *Psicología del color*. Ed. Gustavo Gili, SA, Barcelona, 2004.
- [7] Kolbe, T. Augmented Videos and Panoramas for Pedestrian Navigation. In *LBS & TeleCartography*, G. Gartner, Ed. *Geowissenschaftliche Mitteilungen*, vol. 66, 2003, 45-52.
- [8] Lynch, K. *The Image of the City*. The MIT Press. Cambridge, Massachusetts, 1960.
- [9] Retscher, G., and Thienelt, M. NAVIO – A Navigation and Guidance Service for Pedestrians. *Journal of Global Positioning Systems*, Vol. 3, No. 1-2, 2004, 208-217.
- [10] Tufte, E. R. *Envisioning Information*. Graphics Press, Cheshire, Connecticut, 1990.

Back to the 70's

Cristina Portalés
Dipl.-Ing. Geodesy
Light Laboratory

Francisco Giner
Univ.Ass.Dr. Fine Arts
Light Laboratory

Francisco Sanmartín
Univ.Ass.Dr. Fine Arts
Light Laboratory

Camino de Vera s/n , 46022 Valencia
+34 963877690 Ext.76923

Camino de Vera s/n, 46022 Valencia
+34 963877690 Ext.76923

Camino de Vera s/n, 46022 Valencia
+34 963877690 Ext.76923

criporri@doctor.upv.es

fginerm@pin.upv.es

frasanpi@pin.upv.es

ABSTRACT

Back to the 70's is an augmented reality application, designed to deliver the Vera University campus back to the city centre as it was 30 years ago. Within this, the user is immersed in a new hypothetical environment where the campus buildings are seen together with some historical ones belonging to Valencia City. This paper describes the methodologies followed to deal with the occlusion problem via a conjunction of 3D objects acting as a subtracting models, and the acquisition of the user orientation inside the application, which is achieved via a combination of a visual and an inertial tracker.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation (HCI)]:
Multimedia Information Systems – *Artificial, augmented, and virtual realities.*

General Terms

Algorithms, Performance, Design, Experimentation, Human Factors, Languages.

Keywords

Augmented reality, camera orientation, 3D-models, occlusion problem.

1. INTRODUCTION

The Polytechnic University of Valencia has its headquarters in four campuses: the Vera campus, the Gandia campus, the Xativa campus and the Alcoi campus, being the former the central/main one. This campus is placed at the outskirts of the city of Valencia, at its northern limit, near the highway and railroad exits of the city towards Barcelona. The surrounding rural environment and its distance to the city centre, gives to the campus a very slightly urban identity.

Nevertheless, this situation not always was like that. The history of the city of Valencia in relation to the University dates from

more than 500 years ago, when the population nucleus was very small and compact, with an urban development structure of clearly medieval inheritance. The building of the first Valencian University, nailed in the historical and geographical centre of the city, anchors the university spirit to the historical identity of Valencia. Throughout the centuries the presence and influence of these educational institutions in the life of the citizens had always been something tangible, but as the time went by, the urban development and thus demographic increment favored a change.

The Universities had to extend their teaching institutions and services, making it first under a disintegration politics, acquiring real estates at the limits of the urban centre where to construct their amplifications, which soon were absorbed by the unstoppable development of the city; schools and faculties turned out to be dispersed and unconnected for the whole city. Finally, the price to the rise of the considered bordering areas to the city and the everlasting need of adequacy of the Universities to the increasing demand of place for students, were the trigger that campuses of the ancient and new Universities were established as urbanized cores of new construction, in big terrain extensions placed to a relative distance of the urban centre and, in certain form, isolated; thought as communities that were self-sufficient in the major possible measure and independent from the city.

In the back to the 70's application a hypothetical and nostalgic return to the communion between University and city is proposed, which is made possible thanks to the Augmented Reality approach. Several historical buildings of the city centre are placed in the Vera campus landscape, forming a visual and conceptual extremely attractive and playful game that tries to join in the same scenario reality and illusion, outskirts and city centre, modernity and history.

2. GENERAL DESCRIPTION

2.1 Historical Buildings

Several historical buildings of the city of Valencia were selected for being considered to be emblematic, monumental and popular. Photographs were taken with a digital camera and from multiple points of view in order to model the buildings following photogrammetric techniques [10] to ensure the fidelity with the reality. Models can be achieved in very different formats; in our case, we used the *.obj* for being compatible with OpenGL. Our systems permits the use of complex models, thus the huge number of polygons used seems not to affect the in real-time rendering (see section 3.2). In figure 1 the ancient Serrano's Towers and the obtained 3D model are shown. This monument will be used as an example in the following sections.



Figure 1. Picture and model of the Serrano's Towers.

2.2 Augmented Reality Scenario

The user carries a laptop in a backpack and is immersed in the application via a Head Mounted Display (HMD). Rigidly mounted to the HMD, there is a camera and an Inertial Motion Unit (IMU). The former captures a video image of the real environment and serves as a visual sensor to achieve the user position and orientation inside the application in real-time; the second is exclusively used to sense the user orientation (see section 3.3). He can move freely inside a delimited area, the Fine Arts rounded garden (figure 2), and looking at the landscape discovers a new environment; in the typical architectural space of the University, already known, he finds himself immersed with the virtual buildings of the city centre perfectly integrated with the real ones of the campus, in a virtual space of monumental character.

3. METHODOLOGY

The integration between the real buildings of the campus and the virtual ones of the city centre is carried out in several steps:

- Firstly, a 3D-Cartesian coordinate system has to be established in the campus, and models of the city centre buildings should be 'placed' into it.
- Secondly, the occlusion problem should be taken into account.
- Finally, the user viewing point (i.e., the camera projection centre) should be positioned and oriented in real-time according to the campus coordinate system.

In the following sections, all these steps are explained in more detail.

3.1 3D-Campus Reference Frame

A 3D coordinate system was established for the campus following geodetic techniques. This is a local Cartesian 3D system, defined with three orthogonal axes (X , Y , Z) and with its origin at (X_0 , Y_0 , Z_0). This system will be called campus reference frame or simply campus frame (figure 5). Measurements of available 3D control points were achieved in order to align the Y -axes to the North direction, thus the X -axes is pointing to the East.

Heights of the surrounding buildings were measured with a digital theodolite. Within these measurements and an available – enough accurate – ground map, a 3D campus model was achieved, that was used to deal with the occlusion problem, as will be explained in the following section.

Our major goal was to achieve a proper and reliable integration of the campus and building models. For this reason, some tests were done in order to check the effect produced by mixing both campus

and buildings models. This was realized by using 3D visualization techniques, looking for the best possible integration between them (figure 2).

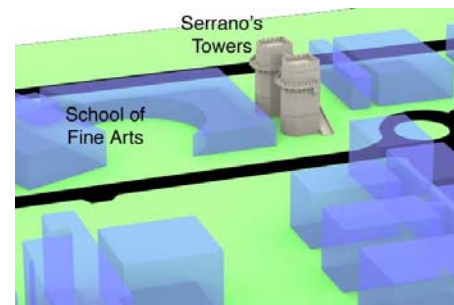


Figure 2. Virtual 3D model of the Vera campus together with the Serrano's Towers model.

Additionally, in order to make the virtual models visually 'fit' better with the reality, some lights conditions were applied in such a way that the intensity and the direction of the virtual luminous focus be as similar as possible to the actual sunlight conditions.

3.2 Occlusion Problem

As it was explained in the previous section, a set of real and virtual objects are integrated together inside the campus reference frame. To achieve a proper integration, the occlusion problem should be taken into account. This problem arises when the virtual objects should be hidden by the real ones. In figure (3a) an example is shown. As it can be seen, the model of Serrano's Towers is at the image foreground – the virtual model hides the real one –, when it should be behind of the Fine Arts' real building according to figure 2. Several authors have tried to solve this problem in different ways [4, 5, 6, 11]. In the Back to the 70's application, this problem was solved with the program MSP/Max Jitter [2], where a virtual model of the campus works as a 3D mask occluding the parts of the city models hidden by the campus ones. In figure (3b), the Serrano's Towers and the Fine Arts' virtual models are seen together. The occlusion is solved by giving to the 3D campus model some special properties, in such a way that it acts as a subtracting model, thus making not visible those parts of the historical buildings that are behind it. In figure (3c) the resulting combination is seen.



(3a)

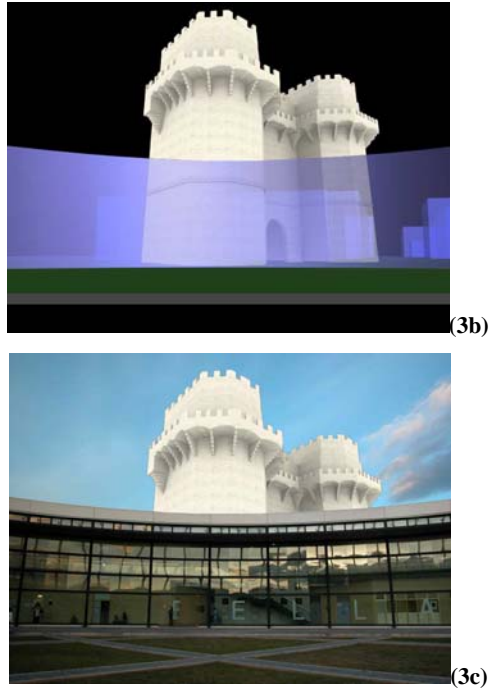


Figure 3. Solving the occlusion problem: (a) model of the Serrano's Towers at the foreground of the video image; (b) Serrano's Towers and mask models; (c) final video image seen by the user

In figure 4, a schema of the patch we created in Jitter is shown.

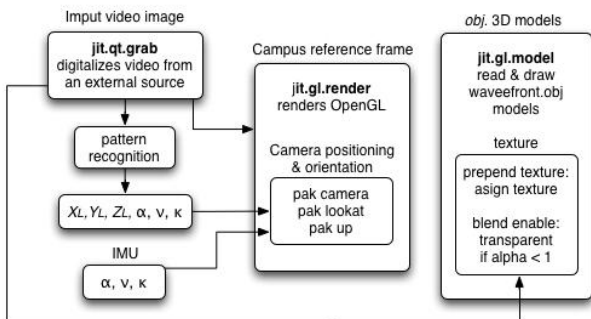


Figure 4. Diagram of the patch created in Max/MSP Jitter to solve the occlusion problem.

This patch consists of a combination of three objects:

- *jit.qt.grab*, that digitalizes the video signal send by the camera (real environment)
- *jit.gl.model*, that reads and draws the *.obj* models (buildings and campus models)
- *jit.gl.render*, that renders via OpenGL the 3D objects overlapping the video image send by the camera

The image acquired with the camera is digitalized to appear as the background image. To achieve the occlusion effect, the 3D campus model is texturized and an alpha value less than one is applied, thus appearing as to be transparent when *blend enable* is

activated. Therefore, the video image appears where the 3D campus model should be, remaining visible only those parts of the virtual historical buildings that are not occluded by the real ones. Finally, the 3D model and the video image are rendered with OpenGL *jit.gl.render* and shown in the output window called *jit.window*. The camera is moved and oriented inside the application within the *pak camera*, *pak lookat* and *pak up* objects.

3.3 User Positioning and Orientation

User positioning and orientation inside the back to the 70's application is achieved via the combination of a camera and an IMU.

The camera is used to act as a visual tracker [1, 9, 14, 15, 16] by sensing some landmarks placed in the surrounding buildings. Each single landmark consists of a set of control points, whose 3D-spatial coordinates are known. These points are basically made of small mirrors, in order to appear in the sensed image as very luminous and therefore to be later on easily recognized by a patten recognition algorithm. Once these points are identified, some photogrammetric rules are applied [10, 12, 17] in order to achieve the six orientation parameters, i.e., three coordinates (X_L , Y_L , Z_L) and three angles (ω , φ , κ or α , v , κ). This process is automatically achieved in the program Max/MSP Jitter [2], with a patch that we named *camori.jit*.

The second one is a device that continuously measures 3D acceleration, 3D rate of turn (rate gyro) and 3D earth-magnetic field data, and outputs three rotations (*roll*, *pitch*, *yaw*) in real-time. In our application, the Xsens MT9 was in use, which is a low cost IMU but accurate enough for augmented reality applications. Furthermore, we also used a magnetic field model in order to avoid possible magnetic disturbances caused by the electronic devices carried by the user. To ensure a maximum accuracy, the inertial sensor was mounted as close as possible to the camera projection centre [13]. Figure 5 shows a schema of the spatial relationship between different frames.

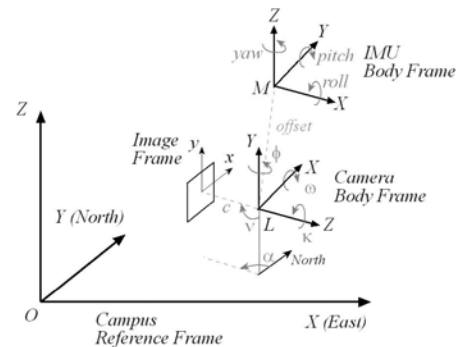


Figure 5. Spatial relationship between campus reference frame, image frame and camera and IMU body frames.

4. CONCLUSIONS

We have presented Back to the 70's, an Augmented Reality application that shows a virtual environment of an alternative/hypothetical present, so answering the questions of how the Vera campus could be nowadays if their origins had not been moved, and how would be its integration at the core of the city of Valencia.

From a conceptual point of view, we can highlight that the utilization of popular architectural elements of the city centre, that possess a historical and emblematic identity of the first order, and their incorporation in an environment of relative new construction, gives course for a visual confrontation, which propitiates diverse conceptual equally valid and attractive interpretations. Specially interesting is the join of the past and the present, across a curious metaphor: virtually annulling the physical distances to show in conjunction diverse elements that cannot be never together in the real sphere, due to their monumental dimensions.

The technological contributions in this paper are the great effort in integrating virtual elements thought the real ones, and to preserve the last ones from a not wished visual occlusion. The use of a campus 3D model that acts as a subtracting model, thus making not visible those parts of the virtual buildings that are occluded by the real ones. Important is also the successful combination of two different sensors, a camera and an inertial device, to track the user's orientation inside the application. And last, but not least, the capacity of the system to render the building models, which are quite complex thus consisting of a great amount of polygons, in real-time whilst running the necessary algorithms to achieve the user's orientation and to solve the occlusion problem.

5. FURTHER WORK

We are thinking in the possibility of combining GPS with the inertial sensor to achieve the camera orientation [3, 8, 13]. This technique is currently been used in the direct orientation process of vehicles, in which an optimum integration of both systems is achieved via a Kalman filter [3, 7, 10]. This will give directly the absolute user's position and orientation making possible the free motion of the user among the entire campus.

6. ACKNOWLEDGMENTS

We would like to thank the staff of the Light Laboratory for their unconditional support.

7. REFERENCES

- [1] ARToolKit. <http://www.hitl.washington.edu/artoolkit>.
- [2] Bernstein, J., Clayton, J. K., Dobrian, Ch., DuBois, R. L., Jones, R., and Taylor, G. *Jitter Tutorials and Topics*. Cycling '74, 2002-2004.
- [3] Cramer, M. GPS/INS Integration. In *Photogrammetric Week'97*. Fritsch D., Spiller R. (Eds.), 1997, 1-10.
- [4] Fischer, J., Bartz, D., and Straßer, W. Occlusion Handling for Medical Augmented Reality using a Volumetric Phantom Model. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST'2004)* (Hong Kong, November 10-12, 2004).
- [5] Fischer, J., Regenbrecht, H., and Baratoff, G. Detecting Dynamic Occlusion in front of Static Backgrounds for AR Scenes. In *Proceedings of the workshop on Virtual environments (EGVE '03)* (Zurich, Switzerland, May 22-23, 2003). ACM Press, New York, NY, 2003, 153-161.
- [6] Fuhrmann, A., Löffelmann, H., Schmalstieg, D., and Gervautz, M. Collaborative Visualization in Augmented Reality. In *IEEE Computer Graphics and Applications*. Vol. 18, no. 4, July/August 1998, 54-59.
- [7] Jekeli, Ch. *Inertial Navigation Systems with Geodetic Applications*. Walter de Gruyter, Berlin, New York, 2001.
- [8] Kameda, Y., Takemasa, T., and Ohta, Y. Outdoor See-Through Vision Utilizing Surveillance Cameras. In *Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'04)* (Arlington, VA, USA, November 02-05, 2004). IEEE Computer Society, 2004, 151-160.
- [9] Koller, D., Klinker, G., Rose, E., Breen, D., Whitaker, R., and Tuceryan, M. Real-time vision-based camera tracking for augmented reality applications. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST'97)* (Lausanne, Switzerland, September 15-17, 1997). ACM Press, New York, NY, 1997, 87-94
- [10] Kraus, K. *Photogrammetry Volume 2, Advanced Methods and Applications*. Ferdinand Duemmler (4th edition). Vienna, 1997.
- [11] Lee, G. A., Billingham, M., and Kim, G. J. Occlusion based Interaction Methods for Tangible Augmented Reality Environments. In *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry (VRCAI'04)* (Singapore, June 16-18, 2004). ACM Press, New York, NY, 2004, 419-426.
- [12] Lerma, J. L. *Fotogrametría Moderna: Analítica y Digital*. Universidad Politécnica de Valencia, 2002.
- [13] Mostafa, M., Hutton, J. and Reid, B. GPS/IMU products – the Applanix approach. In *Photogrammetric Week'01*. Fritsch D., Spiller R. (Eds.), 2001, 63-83.
- [14] Neumann, U., and You, S. Natural Feature Tracking for Augmented Reality. In *IEEE Transactions on Multimedia*. Vol. 1, no. 1, March 1999, 53-64.
- [15] Piekarski, W., and Thomas, B. H. Using ARToolKit for 3D Hand Position Tracking in Mobile Outdoor Environments. In *Proceedings CD of 1st International Augmented Reality Toolkit Workshop*. (Darmstadt, Germany, September 29, 2002).
- [16] Prince, S. J. D., Xu, K., and Cheok, A. D. Augmented Reality Camera Tracking with Homographies. In *IEEE Computer Graphics and Applications*. Vol. 22, no. 6, November/December 2002, 39-45.
- [17] Rönnholm, P., Hyypä, H., Pöntinen, P., Haggrén, H. Interactive relative orientation between terrestrial images and airborne laser scanning data. In *Proceedings of the ISPRS working group III/3 workshop, Vol. XXXIV-3/W13* (Dresden, Germany, October 8-10, 2003). Editors: H.-G. Maas, G. Vosselman, A. Streilein.

URBANMIX

PORTALÉS RICART, Cristina. Laboratorio de Luz, BB. AA., Universidad Politécnica de Valencia, camino de Vera s/n 46022, Valencia. criporri@doctor.upv.es

GINER MARTÍNEZ, Francisco. Laboratorio de Luz, BB. AA., Universidad Politécnica de Valencia, camino de Vera s/n 46022, Valencia. fginerm@pin.upv.es

SANMARTÍN PIQUER, Francisco Javier. Laboratorio de Luz, BB. AA., Universidad Politécnica de Valencia, camino de Vera s/n 46022, Valencia. frasanpi@pin.upv.es

Palabras clave: realidad aumentada, tiempo real, espacio abierto.

1. INTRODUCCIÓN

“Realidad aumentada”, se puede definir como la amplificación de la capacidad sensorial de la percepción del mundo que nos rodea mediante el uso de dispositivos que superponen elementos virtuales a la imagen real.

La aplicación Urbanmix, nos posibilita este concepto a escala urbana y en exteriores usando un sistema propio de realidad aumentada, que se basa en modelos 3D y el programa Max/MSP Jitter [2] de edición de video en tiempo real.

A través de la aplicación Urbanmix, podemos acceder a una visión de la ciudad o espacio en el que nos encontremos, ampliada con algunos edificios o monumentos provenientes de otras. Con ello, y utilizando escalas 1:1 en el sistema, podemos realizar fácilmente comparaciones reales y así tomar verdadera conciencia de la magnitud de los edificios que visualizamos.

Ante la gran dimensión de algunos de los edificios que podemos introducir en el sistema, es necesario ubicarlos a una distancia lo suficientemente grande como para que el sistema pueda tomar las referencias necesarias para ubicarlo correctamente y su visualización continúe siendo creíble; por la falta de espacio en las ciudades lo normal es que algunos edificios reales se superpongan generalmente a los virtuales, por una superposición de canales de video, con lo que nos encontramos con un problema de oclusión de formas reales a virtuales que no se corresponde con lo habitual en estos sistemas, oclusión de formas virtuales a reales. El problema se resuelve con una máscara tridimensional de video que actualiza sus datos en tiempo real, que permite al usuario moverse a su antojo dentro del área acotada en donde se desarrolle la aplicación, y continuar viendo la inmersión del edificio o monumento escogido en el espacio precargado desde cualquier punto de vista.

2. ANTECEDENTES

Urbanmix es un proyecto en continuo desarrollo. Nuestro primer prototipo [11] estaba pensado para su ejecución dentro de un único espacio, el jardín de la escuela de BB.AA. de la Universidad Politécnica de Valencia. Actualmente, pensamos que el proyecto puede ser fácilmente trasladable a cualquier espacio abierto que sea susceptible de ser medido y modelado. Esta tarea se facilita enormemente gracias a la herramienta de libre distribución Google Earth (ver apartado 3.3.), con la que es posible adquirir la planta y ubicación relativa de los necesarios modelos 3D de casi cualquier ciudad del mundo; pero también, por el empleo de una serie de pegatinas/marcas

(fácilmente transportables) para realizar cálculos de la posición del usuario dentro de la aplicación mediante técnicas de visión por computador (ver apartado 3.2.).

3. METODOLOGÍA

Para una correcta integración de los edificios reales con los virtuales, es necesario realizar los siguientes pasos:

- En primer lugar, se debe definir el sistema de realidad aumentada, es decir, los dispositivos que intervendrán en la aplicación.
- Después, se definirá y modelará el escenario virtual, que se superpondrá al real, y los objetos 3D que se incluirán en el sistema. Para resolver el problema de la oclusión, el sistema aplica una máscara sustractiva de video entre ambos modelados; la oclusión se da cuando un objeto real se interpone entre el objeto virtual y el punto de vista del usuario.
- Finalmente, el sistema debe ser capaz de calcular en tiempo real la posición y orientación del usuario que, en este caso, coincide con el centro de proyección de la cámara.

A continuación se explican todos estos pasos con detalle.

3.1. Sistema de realidad aumentada

En este tipo de aplicaciones, se necesita de un dispositivo que permita al usuario percibir la nueva realidad creada. Estos dispositivos se denominan *interfaces*, y a través de ellos se muestran superpuestas la realidad y la ficción. Generalmente, se suelen utilizar unas gafas especiales, conocidas como HMD (Head Mounted Display), aunque también podemos encontrar diversas aplicaciones en las que se hace uso de pantallas (bien de ordenadores portátiles o de tablet pc). Ambos dispositivos tienen sus ventajas e inconvenientes. Una de las grandes ventajas que ofrece una pantalla, es que puede ser visualizada al mismo tiempo por varios usuarios, permitiendo un entorno colaborativo. Sin embargo, al tener lugar nuestra aplicación en un espacio abierto, optamos por la elección de un HMD, principalmente porque una pantalla de ordenador se vería afectada por los reflejos de la luz solar, impidiendo una correcta percepción de la nueva realidad. Además, las HMD son extremadamente ligeras en comparación con las pantallas, y ofrecen también una inmersión más creíble del usuario dentro de la aplicación.

Otro dispositivo indispensable en nuestra aplicación es una cámara, cuyo centro de proyección, en lo posible, se debe hacer coincidir con el punto de vista del espectador; es por ello que la cámara deberá situarse lo más cerca posible de los ojos del usuario. Esta cámara se encarga de registrar nuestro entorno en tiempo real, para posteriormente la imagen-vídeo ser procesada por el sistema, el cual nos ofrecerá (a través del HMD) dicha imagen-vídeo superpuesta con los objetos virtuales creados por ordenador. La resolución de la cámara deberá ser óptima, ni demasiado alta (para permitir el proceso en tiempo real) ni demasiado baja (para no ofrecer una imagen demasiado pobre). En nuestro caso utilizamos una Fire-i de Unibrain [15] con una resolución de 640 x 480 píxeles.

Otro factor a tener en cuenta es la posición y orientación del usuario dentro de nuestra aplicación, que se llevará a cabo en tiempo real mediante sensores. En nuestro caso, hemos utilizado la combinación de un sensor visual y uno inercial (ver apartado

3.4.). El sensor visual es la cámara anteriormente descrita; como sensor inercial hemos utilizado el MT9 de Xsens [14].

Finalmente, el usuario deberá cargar consigo un ordenador portátil en una cómoda mochila. Este se encargará de procesar toda la información recibida en tiempo real y realizar las operaciones necesarias para crear el nuevo entorno que verá el usuario.

En la siguiente figura pueden apreciarse todos los dispositivos anteriormente descritos.

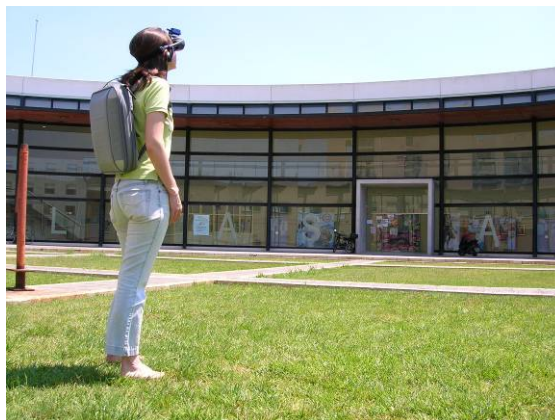


Figura 1. Dispositivos de realidad aumentada utilizados en Urbanmix.

3.2. Resolución de la oclusión

Para alcanzar una integración apropiada, debe ser resuelto el problema de la oclusión. Este problema se presenta cuando el canal en el que aparecen los objetos virtuales se superpone al canal de video de la cámara, entonces los objetos virtuales aparecen superpuestos a la realidad.

Si pretendemos que los objetos virtuales se integren en una trama urbana debemos poder escoger qué parte de los objetos virtuales ha de quedar visible, y cual ocultar para que dejen ver a través la realidad que se supone que ha de presentarse en un plano más cercano a ellos.

Diversas metodologías se han probado para resolver este sistema [3; 4; 9]. En Urbanmix, este problema ha sido solucionado con una aplicación realizada con el programa de edición de video en tiempo real Max/MSP/Jitter [2], donde un modelo virtual de la ciudad funciona como una máscara 3D que oculta partes del modelo virtual del edificio integrado según la posición del usuario. La oclusión es solucionada dando al modelo 3D de la ciudad algunas características especiales, de manera tal que aparece transparente pero oculta aquellas partes del modelado posterior a las que se superpone desde el punto de vista del usuario.

Por lo tanto, la imagen video aparece donde el modelo 3D de la ciudad debería estar visible, ocultando solamente esas partes del edificio virtual que son ocluidos por los verdaderos. Finalmente, el modelo 3D y la imagen video se renderizan con OpenGL mediante el objeto *jit.gl.render* y se muestran en la ventana de la salida llamada *jit.window*.

En la siguiente figura podemos ver el funcionamiento de la oclusión. Como ejemplo se ha modelado el entorno de la Puerta del Sol de Madrid, que es donde se desarrollará la aplicación; como objeto virtual perteneciente a otra ciudad, se ha modelado el edificio Chrysler, de Nueva York. Para entender mejor el proceso hemos superpuesto estos modelados a una imagen que no se corresponde con los modelados cargados, esta

imagen la recogería la cámara web. En segundo lugar, vemos el resultado al aplicar el efecto transparencia y asignar la máscara sustractiva, quedando visible la parte del Chrysler que no quedaría oculta por el modelado.

En último lugar, vemos el efecto que se produce cuando hacemos coincidir el modelado virtual de la ciudad (transparente también) con la imagen recogida por la cámara web.



Figura 2. Funcionamiento de la oclusión.

3.3. Modelado 3D

Como ya hemos apuntado en el anterior apartado de la oclusión, para que el sistema funcione, es necesario contar con un par de modelados que han de funcionar en canales distintos, el del entorno en el que nos movamos, generalmente urbano, y que ha de funcionar como máscara, y el del edificio o monumento que queramos introducir en él.

El espacio urbano en el que vayamos a hacer funcionar el sistema, necesario para poder contemplar a escala real los modelos introducidos, se ha modelar tridimensionalmente. Para ello partimos de vistas aéreas de esa parte de la ciudad, que podemos encontrar, por ejemplo, en la aplicación de libre distribución Google Earth.

Teniendo en cuenta que este modelado no ha de tener demasiado detalle, sino que más bien, su naturaleza ha de ser sencilla y su acabado limpio para que el efecto de

sustracción sea efectivo y rotundo, la cantidad de polígonos con la que lo construimos no es excesiva.



Figura 3. Creación del modelado urbano desde vista aérea

El proceso es sencillo, desde una imagen aérea con suficiente calidad, tomamos las dimensiones bidimensionales y los datos de ubicación relativa en planta de cada volumen, y con técnicas fotogramétricas y/o topográficas [6; 8], le asignamos a cada uno de ellos, su altura y las peculiaridades específicas que el sistema necesita para funcionar; se ha de cuidar en específico el modelado del contorno superior, ya que es generalmente el que limita con la visión del edificio o monumento que aumenta la realidad al introducirse en el sistema.

El proceso se puede realizar en principio con cualquier programa de creación 3d, siempre que el resultado final pueda exportarse al formato nativo de opengl, con extensión *.obj.

En cuanto al edificio que introducimos en el sistema, hemos de recurrir en principio a un nivel de acabado mucho mayor y a una texturización del mismo con fotografías reales del edificio para conseguir un efecto mayor de realidad. Este modelado generalmente es mucho más detallado, y contempla por tanto una cantidad mayor de polígonos pues ha de soportar con garantías la visualización del usuario-espectador y la comparación que se realiza con la realidad al ubicarlo en el mismo plano de visión en el que están los edificios reales. A partir de la información que se pueda reunir sobre el edificio, se realiza un modelado exhaustivo, en el que son muy importantes las proporciones y su altura.

Por otro lado, su textura no puede ser de origen sintético, pues se evidencia inmediatamente y se percibe como una construcción digital, la solución más efectiva es cubrir el modelado entero con fotografías del edificio original, consiguiendo así un efecto mucho más creíble, las fuentes de obtención de éstas, pueden ir desde la toma directa si es posible, hasta la utilización de imágenes tomadas de Internet, ya que la resolución final de visualización es la de pantalla, 72 ppp.

Es preciso recalcar que todos los programas de creación 3d, poseen un apartado de texturización, comúnmente llamado mapeado, que introduce sobre la imagen que vayamos a utilizar, los efectos necesarios para



integrarla en la imagen final; al exportar el objeto al formato *.obj, estos efectos programados se pierden, ya que generalmente, se exporta simplemente la información volumétrica y el enlace a la imagen original que ha de servir de textura. Por ello la integración y adecuación de estas imágenes se ha de realizar en un programa externo de edición de las mismas, como el Adobe Photoshop o el Corel Photopaint, de esta forma, el modelo *.obj, se encarga de ubicar estas imágenes, ya preparadas, en el modelo tridimensional.

Una vez obtenidos los dos modelados, se ubican en el espacio, el primero, correspondiente al entorno real, y del cual se utiliza únicamente su geometría, se sitúa con respecto a la ubicación del usuario-espectador, exactamente en las coordenadas relativas que coinciden con las de los volúmenes reales, buscando así una superposición de formas lo más exacta posible.

El segundo se puede ubicar en donde queramos, pero siendo conscientes, como decíamos anteriormente, de que lo interesante es que se contemple todo su volumen dentro del mismo campo visual que nos posibilita ver los edificios reales cercanos. De lo contrario, no existiría esa comparación que hace al sistema novedoso y útil.

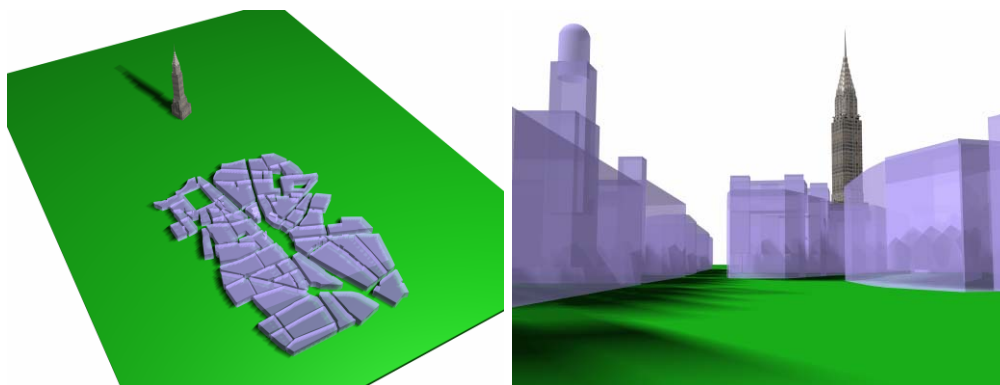


Figura 3. Modelos 3D de los edificios pertenecientes al entorno real y modelo del edificio Chrysler.

3.4. Orientación y posicionamiento del usuario

Para una correcta ubicación del usuario dentro de una aplicación de realidad aumentada, es necesario calcular un total de seis parámetros, generalmente denominados como “parámetros de orientación externa”, en contraposición a los “parámetros de orientación interna” (distancia focal y punto principal de la cámara). Estos seis parámetros son: tres rotaciones u “orientaciones”, que son el acimut, cabeceo, y balanceo (α , ν , κ); y tres coordenadas o “posiciones”, que son la (X, Y, Z) del usuario dentro del sistema 3D establecido. Como ya se ha apuntado en el apartado 3.1., la orientación y posición del usuario dentro de nuestra aplicación se realiza mediante la combinación de dos sensores, uno visual y otro inercial (comúnmente denominado IMU: Inertial Motion Unit). Ambos sensores son de características diferentes, pero si se combinan adecuadamente, pueden ofrecernos la precisión necesaria para nuestra aplicación. Dichas características o particularidades son:

- El sensor inercial es un tipo de sensor “directo”, es decir, registra automáticamente y de forma continua una serie de parámetros sin necesidad de realizar cálculos. Los parámetros que registra el MT9 son: la aceleración, variación del giro y campo magnético local en tres direcciones ortogonales, y la

- temperatura ambiente. A partir de estos datos, directamente se obtiene la orientación del usuario, que puede venir expresada como una matriz de rotación.
- Por el contrario, la cámara es un sensor de tipo “indirecto”, es decir, se necesitan de una serie de operaciones matemáticas para poder calcular los parámetros que necesitamos a partir de un previo análisis de la imagen recogida. Específicamente, la cámara deberá de reconocer una serie de puntos repartidos por el espacio; estos puntos se denominan “puntos de apoyo” y tienen la peculiaridad que sus coordenadas tridimensionales son conocidas. Estos puntos pueden hacerse visibles mediante pegatinas o elementos reflectantes, de tal modo que la cámara los pueda registrar y reconocer mediante técnicas de visión por computador (en nuestro caso, tracking de color). A partir de ellos, y mediante una serie de operaciones matemáticas, obtendremos finalmente la posición del usuario.

En la siguiente figura se puede ver esquematizado el funcionamiento de nuestro sistema, tal y como hemos explicado:

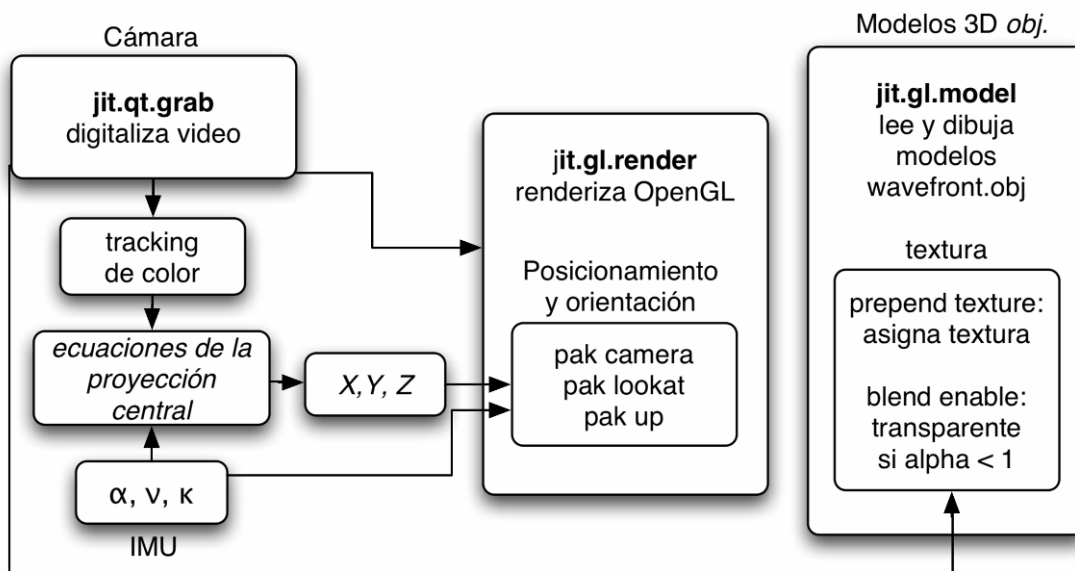


Figura 4. Esquema del funcionamiento de nuestro sistema dentro del programa Max/MSP Jitter.

Realmente, tanto la posición como orientación del usuario podría realizarse con el único empleo de un tracking visual [1; 5; 9; 10; 12]; el problema que surge en este caso, es que el número mínimo de puntos de apoyo que debería registrar la cámara simultáneamente son seis [8], hecho que ralentizaría considerablemente el proceso de adquisición y procesado de los datos y, por lo tanto, dificultaría considerablemente la ejecución de la aplicación en tiempo real. En Urbanmix, al encargarse el IMU de adquirir de forma directa la matriz de rotación, la cámara únicamente deberá registrar simultáneamente un mínimo de dos puntos, de acuerdo con las ecuaciones de la proyección central [6; 8; 13], para calcular la posición.

Respecto al tracking o rastreo de color, hemos utilizado el objeto *tap.jit.colortrack*, que detecta colores específicos de una imagen-vídeo y devuelve información numérica de ellos. De esta manera obtendremos en cada instante información que proporciona las coordenadas de los centros de los objetos rastreados que, como ya hemos dicho, son

una especie de pegatinas circulares de colores diferentes y que contrasten con el entorno.



Figura 5. Imagen de la Puerta del Sol con las pegatinas;
Objetos registrados por el tracking de color.

4. CONCLUSIONES

En esta comunicación se ha presentado Urbanmix, un sistema de realidad aumentada de creación propia.

En este, el usuario se ve inmerso, mediante un dispositivo HMD, en un entorno aumentado; en particular, el espacio donde se desarrolla la aplicación puede ser cualquier ciudad o entorno susceptible de ser modelado; como ejemplo, hemos tratado de recrear nuestra aplicación en la plaza de la Puerta del Sol de Madrid.

El usuario percibe su entorno real mezclado con una serie de objetos virtuales creados por ordenador; estos objetos son edificios de otra ciudad diferente, siguiendo nuestro ejemplo, el edificio Chrysler de Nueva York, y que están dotados de textura foto-realística para aumentar la sensación de credibilidad del sistema en su conjunto. De este modo, el usuario percibe una fusión de ciudades (Madrid-NY), alterando su percepción del entorno y viéndose inmerso en una nueva realidad.

Este proyecto se ha realizado con el programa Max/MSP Jitter que, en principio, está pensado para gestionar en tiempo real imágenes-video y objetos 3D, pero que nosotros hemos adaptado para lograr una combinación y/o mezcla de la realidad con la ficción.

Para ello, hemos creado un *patch* que resuelve el problema de la oclusión mediante el empleo de modelos 3D de los objetos pertenecientes al entorno real (Puerta del Sol de Madrid) y máscaras de vídeo; además, la posición y orientación del usuario dentro de Urbanmix, se ha realizado también mediante un *patch* que combina un tracking visual con uno inercial, todo ello en tiempo real.

5. REFERENCIAS

- [1] ARToolKit, <http://www.hitl.washington.edu/artoolkit>
- [2] BERNSTEIN, J., CLAYTON, J. K., DOBRIAN, CH., DUBOIS, R. L., JONES, R., TAYLOR, G. (2002-2004) *Jitter Tutorials and Topics*. Cycling '74.
- [3] FISCHER, J., BARTZ, D., STRAßER, W. (2004). Occlusion Handling for Medical Augmented Reality using a Volumetric Phantom Model. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST'04)* (Hong Kong, 10-12 Noviembre).

- [4] FISCHER, J., REGENBRECHT, H., BARATOFF, G. (2003). Detecting Dynamic Occlusion in front of Static Backgrounds for AR Scenes. *Proceedings of the workshop on Virtual environments (EGVE '03)* (Zurich, Switzerland, 22-23 Mayo, 2003). ACM Press, New York, NY, pp. 153-161.
- [5] KOLLER, D., KLINKER, G., ROSE, E., BREEN, D., WHITAKER, R., AND TUCERYAN, M. (1997). Real-time vision-based camera tracking for augmented reality applications. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST'97)* (Lausanne, Switzerland, 15-17 Septiembre, 1997). ACM Press, New York, NY, pp. 87-94
- [6] KRAUS, K. (1997). *Photogrammetry Volume 2, Advanced Methods and Applications*. Ferdinand Duenmiller (4th edition). Vienna.
- [7] LEE, G. A., BILLINGHURST, M., KIM, G. J. (2004). Occlusion based Interaction Methods for Tangible Augmented Reality Environments. *Proceedings of the 2004 ACM SIGGRAPH international Conference on Virtual Reality continuum and its applications in industry (VRCAI'04)* (Singapore, 16-18 Junio, 2004). ACM Press, New York, NY, pp. 419-426.
- [8] LERMA, J. L. (2002). *Fotogrametría Moderna: Analítica y Digital*. Universidad Politécnica de Valencia.
- [9] NEUMANN, U., YOU, S. (1999). Natural Feature Tracking for Augmented Reality. *IEEE Transactions on Multimedia*. Vol. 1, no. 1, Marzo, pp. 53-64.
- [10] PIEKARSKI, W., THOMAS, B. H. (2002). Using ARToolKit for 3D Hand Position Tracking in Mobile Outdoor Environments. *Proceedings CD of 1st International Augmented Reality Toolkit Workshop*. (Darmstadt, Germany, 29 Septiembre, 2002).
- [11] PORTALÉS, C., GINER, F., SANMARTÍN, F. J. (2005). Back to the 70's. *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE'05)*. (Valencia, España, 15-17 Junio, 2005). ACM Press, New York, NY, pp. 209-212.
- [12] PRINCE, S. J. D., XU, K., CHEOK, A. D. (2002). Augmented Reality Camera Tracking with Homographies. In *IEEE Computer Graphics and Applications*. Vol. 22, no. 6, Noviembre/Diciembre, pp. 39-45.
- [13] RÖNNHOLM, P., HYYPPÄ, H., PÖNTINEN, P., HAGGRÉN, H. (2003). Interactive relative orientation between terrestrial images and airborne laser scanning data. In *Proceedings of the ISPRS working group III/3 workshop, Vol. XXXIV-3/W13* (Dresden, Alemania, 8-10 Octubre). Editors: H.-G. Maas, G. Vosselman, A. Streilein.
- [14] Unibrain, <http://www.unibrain.com>
- [15] Xsens, <http://www.xsens.com>

Exploring Social, Cultural and Pedagogical Issues in AR-Gaming through the Live LEGO House

ABSTRACT

Live LEGO House (LLH) is an interactive Augmented Reality (AR) environment to explore coexistence and multicultural factors through gaming. The system consists basically of a physical/real house built with the LEGO blocks [11], which is enriched with different non-physical/virtual multimedia files (sounds, videos and 3D animations), insisting in social, cultural and pedagogical issues that the AR technology can bring for children. Tactile, visual and auditory stimuli are present at the game.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation (HCI)]: Multimedia Information Systems – *Artificial, augmented, and virtual realities.*

General Terms

Design, Experimentation, Human Factors, Languages.

Keywords

Augmented Reality, Interaction, Collaborative, Gaming.

1. INTRODUCTION

AR-gaming has increased in the last years [1; 3; 6; 7; 8], demonstrating its great potential for the entertainment. Other works exploit the potential of physical and social interactions that AR can offer [4]. Following these works, in LLH more than a game is proposed, as it pays special attention in social, cultural and pedagogical issues for children, demonstrating the great potential that the AR technology can bring in these areas.

LLH has been specially developed, but not restricted, for children between 3-7 years old. To keep children's attention at these ages, a large variety of funny sounds, graphics and videos are introduced. The thematic of all these media is 'multicultural', so children can indirectly learn about other cultures. Furthermore, in

order to reinforce good environmental and social concepts, some restrictions regarding to water and sound are introduced (see section 2.4), in such a way that positive coexistence between players is fomented, which is the base to achieve a better society: if two persons cannot coexist and save energy at their own home, how are we going to lead to a multicultural society and/or preserve world natural resources?

The game can be divided in the following steps:

1. Build your own house within the LEGO blocks.
2. Tell the program where the objects are (calibration procedure).
3. Explore the AR environment within your partner.

It is very important to remark that in the LLH the goal is not to build a computer video game within the AR technology, but to 'augment' a traditional (non-computer) childish game to explore the matters of coexistence and multiculturalism. There are no rules: children can explore the LLH as they wish; they do not need to arrive to a target or to find some objects that give them extra points to keep playing. This means, in LLH there is no 'game over'. Children can play until they wish, like in any traditional non-computer game. Within the so established AR game, children are highly encouraged to use their imagination and to learn from not good events/actions (negative coexistence) and good events/actions (positive coexistence). They have to decide when is time to go to bed or to have lunch, for instance.

Although a computer is necessary to play, children do not need to care about that, as they are playing in the real world; computer can be considered as the tool that renders different multimedia files (on the screen or speakers) depending on dolls actions. Decisions on how to rule the game are made by children, not by the computer.

2. GAME DESIGN

According to [5], important issues of game design in AR include:

1. Technical characteristics: advanced visualization and interaction; adaptability and usability issues.
2. Interactivity: advanced using computer sensors and natural methods.
3. Social, cultural, and pedagogical issues: can be direct and indirect.
4. Collaboration: can be direct and/or indirect.

- Game scenarios: more pragmatic; need to merge scenario with real world.

2.1 Technical Characteristics

The LLH is built within the MXRToolKit [9], which consists of a library of routines to help with all aspects of building mixed reality applications.

The whole system of the LLH consists on (Figure 1):

- A computer work station, with the MXRToolKit libraries;
- Display system: a computer screen;
- Tracking system: a web cam and MXRToolKit square markers;
- Tangible part: The LEGO blocks that form the house/furniture and the dolls;
- Two speakers where sounds are reproduced.

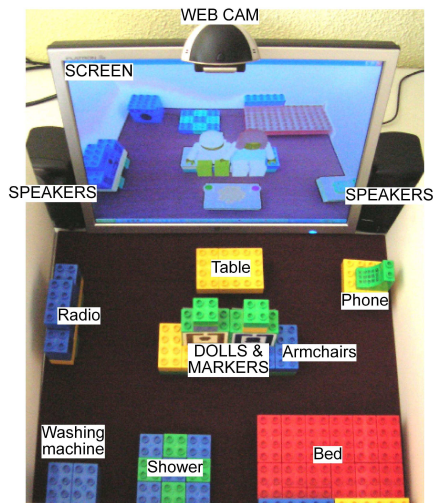


Figure 1. Schema of the LLH system.

2.2 Interactivity

Human-computer interaction is achieved through markers attached to real dolls (Figure 2). The development of amicable interfaces is a key goal to achieve, so that even users not familiar with computers can enjoy the advantages that they offer them. This implies the elimination of the ‘digital gap’. If someone is asked to communicate with a computer, they may think of a mouse, a keyboard or a joystick as the ‘normal’ interfaces. Most people probably do not ever think about a ‘doll’ as an interface. But this is exactly what happens with the LLH: the mouse is changed by a doll, making the interaction more tangible. This is in fact a natural interface for children that don’t need to learn beforehand.

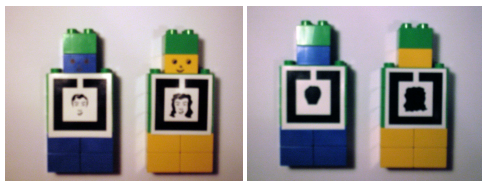


Figure 2. Front and back of the dolls.

Users can build the house as they wish, so furniture are somehow non-static objects. That means that whenever the application is started or furniture is moved, the user should ‘tell the program’ where the real objects are placed. This step is called ‘calibration procedure’, and is easily achieved via the keyboard.

Interaction is triggered by two ways (Table 1): by dolls’ proximity to the elements of the house or by dolls’ tilting (roll and pitch). Produced actions and rendered multimedia files depend upon previous or common interactions.

Furniture	Interaction	Actions	Multimedia Files
Washing machine	Proximity	Put the washing machine / Prohibit	Sound: Washing machine on 3D: Water prohibition signal
Radio	Proximity / Roll	Turn on the radio / Tuning the radio / Dance / Prohibit	Sound: Different songs & radio programs 3D: Animated dolls dancing / Sound prohibition signal
Armchair	Proximity / Pitch	Sit down / Eat / Drink / Yawn	3D: Dolls sit down, drink & eat / Glasses with juice / A plate with cookies Sound: Eating / Drinking / Yawning
TV	Proximity	Turn on the TV / Change TV program	Videos: Different videos
Bed	Proximity / Pitch	Put the pyjama on / Close eyes / Snore / Dream	3D: Dolls wear pyjamas / Animated dreams Sound: Snoring
Shower	Proximity	Undress / Have a shower / Sing	3D: Nude dolls / Water prohibition signal Sound: Singing / Falling water
Telephone	Proximity / Roll / Pitch	Make a call / Telephone ring / Prohibit	Sound: Making a call / Telephone ringing / Picking up the telephone / Say hello 3D: Sound prohibition signal

Table 1: Furniture related to interaction, possible actions and multimedia files.

2.3 Social, Cultural, and Pedagogical Issues

Social and/or cultural issues can be established in direct communication between players, if they belong to different social or cultural environments. Also, social, cultural and/or pedagogical factors are implicit in the LLH via the multimedia files, which include sounds, videos and 3D-objects.

2.3.1 Sounds

Sound is a very important part of the game. Some of the sounds have been taken from the web page at [10], like: falling water, singing in the shower, tuning the radio, telephone ring, snoring, laughing, eating, etc. Songs related to the LLH radio are multicultural: samba, paso doble, rumba, waltz, sirtaki, tribal, traditional Chinese, etc. Afterwards, these sounds have been edited and others designed by electronic synthesis in order to enrich sound texture and timbre. This gives an interesting and 'funny' touch to the game that will catch children's attention. Moreover, each time a doll interacts with an item, the sound changes; for example, a total of 10 different rings are related to the telephone, and more than 30 songs are linked to the radio. The large variety of different sounds included in the game, makes the experience more playful and increases children curiosity.

2.3.2 Videos

The TV is one of the central elements in the game. This element is also interactive, and it will show the augmented dolls (when TV is off) or some video files (when TV is on), depending on users' actions (see section 2.4):

- If any doll is sitting down on their corresponding armchairs (yellow for LEGOWoman and blue for LEGOMan): the TV is off and the AR scenario is shown.
- If at least one doll is sitting down on its armchair: the TV is on and a video file is rendered.
- If both dolls are sitting down on their armchairs, only a common program that delights both dolls is seen.

Thematic of the videos are also multicultural (Figure 3).

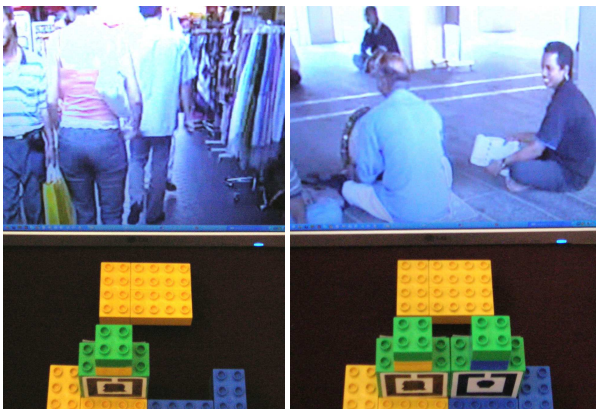


Figure 3. Left: people walking at China Town in Singapore. Right: people at a Hindu Temple in Singapore.

Most of the videos were recorded in Singapore, which is indeed a good example of multiculturalism. One of the characteristics of Singapore is the actual coexistence between people from different races (Chinese, Malaysian and Indian) and religions (Buddhism,

Islamism, Catholicism, Hinduism, etc.). Although there are some characteristic ghettos (as Chinatown), cultures are quite mixed. Some examples of this multiculturalism are: people from different countries live and work together respecting their culture and traditions; a Hindu Temple can be found in Chinatown; all kind of traditional food is offered at the stalls of a single hawker centre; different traditions are celebrated during the whole year; different languages are in use with the English as unifier; etc. Also, videos from European countries have been included. Through these videos, children will be able to see people from different cultures and countries working, praying, having lunch, walking on the streets, etc., showing an example of positive coexistence.

2.3.3 3D-Objects

3D objects, including animations, are within the VRML format. Dolls have been modelled according to the original LEGO dolls: for example, when the dolls are having a shower, both bodies look like the same (no gender), and dolls can only be distinguished because of the hair.

2.4 Collaboration

In LLH game, children collaboration is achieved face-to-face. Children will have to agree in some aspects, as the sequence of their actions conditions the multimedia files that are rendered, and actions of a doll condition other doll. These dependencies are made in order to foment positive collaboration. Negative coexistence is avoided via some kind of restrictions and/or prohibitions regarding to sound and water. For example:

- If both dolls are watching TV, a program that delights both will be displayed. If children want to change TV channel, they will have to agree (one doll should get up from the armchair).
- If one of the dolls is sleeping, the sound is not allowed, that means that the other doll cannot listen to the radio or make a telephone call.
- If one doll has too many showers in a single day, no more water is left, and the other doll cannot have a shower or put the washing machine.
- If a doll sits in front of the table to have lunch, it only will be able to eat or drink, but not both at the same time. To properly have lunch, both dolls need to do the action at the same time.

Therefore some funny situations can arise: when the amount of water is finished, they have to wait till next day to have more, so both dolls have to go to sleep; But when a doll is sleeping, the other cannot switch the radio on or make a call. Therefore, children would need to agree in order to rule the game properly. Mutual agreement will lead to positive coexistence, thus social factors are once more fomented.

2.5 Game Scenarios

Game takes place in the mixture of real and virtual environments, in the real house and in the house shown at the screen. Commonly, AR is thought to be mainly visual. But it has to be noticed that, although video-based approaches are the primary interest in AR environments, also sounds or other kind of media can be incorporated in order to enrich the experience. According to the virtual stimulus, the LLH scenario integrates the three

possibilities, which are tactile, visual and auditory, as olfactory and gustatory do not have their virtual equivalent. According to [2], this represents the most advanced multi-modal AR. That means, children can touch, see and hear in the LLH game. In (Figure 4), some samples of mutual coexistence are shown.

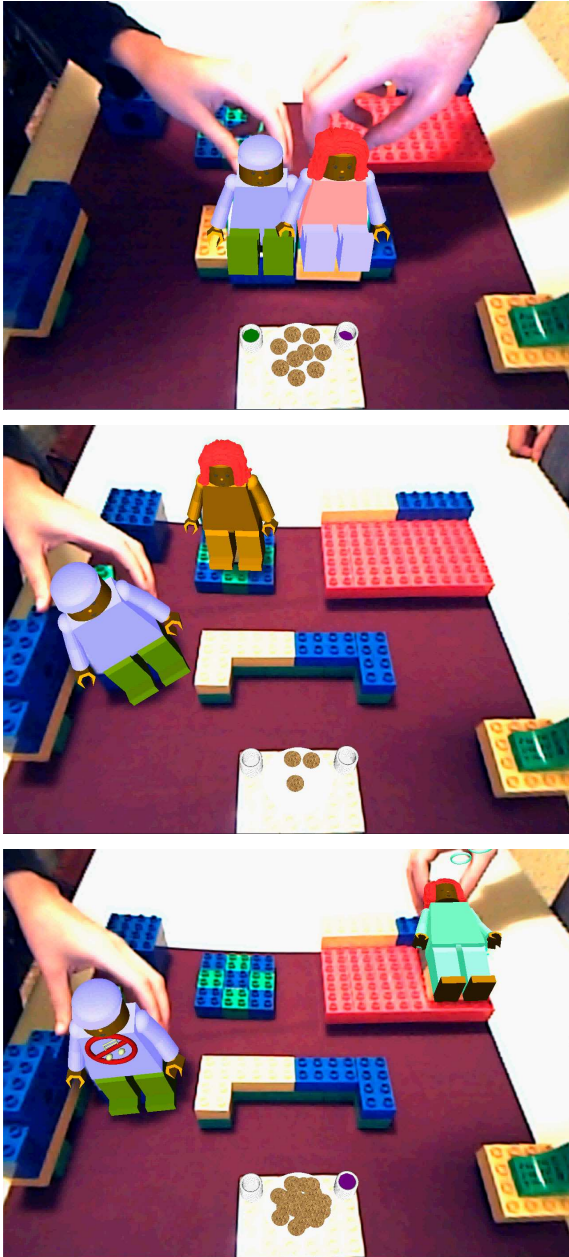


Figure 4: Mutual coexistence at the LLH scenario. Top: both dolls are having lunch together; Middle: one doll is having a shower and the other is dancing next to the radio. Bottom: one doll is sleeping, so the other cannot make sound (radio off).

3. CONCLUSION

In this paper, the LLH interactive AR game has been shown. While children explore the house, several multicultural sounds, videos and 3D models are rendered, thus incising in the great

potential that AR-systems can bring for the social, cultural and pedagogical issues for children. LEGO blocks are very popular between children (and adults), so they will find the physic part of the game very familiar and easy to play with. LLH still keeps the bases of the LEGO, allowing children to collaborate and use their imagination and creativity while playing, but adds some kind of multimedia files that enrich the game experience and foment other kind of mutual collaboration and interaction.

4. REFERENCES

- [1] Andersen, T. L., Kristensen, S., Nielsen, B. W. and Grønbaek, K. Designing an Augmented Reality Board Game with Children: The BattleBoard 3D Experience. In *Proceeding of the 2004 conference on Interaction design and building a community*. ACM Press, New York, NY, 2004, 137-138.
- [2] Barfield, W., Rosenberg, C. and Lotens, W. A. *Augmented-reality displays*. Virtual environments and advanced interface design. Oxford University Press, Inc., New York, NY, 1995, 542-575.
- [3] Brederode, B., Markopoulos, P., Gielen, M., Vermeeren, A. and de Ridder, H. pOwerball: The design of a novel mixed-reality game for children with mixed abilities. In *Proceeding of the 2005 conference on Interaction design and children (IDC '05)*. ACM Press, New York, NY, 2005, 32-39.
- [4] Cheok, A. D., Teh, K. S., Nguyen, T. H. D., Qui, T. C. T., Lee, S. P., Liu, W., Li, C. C., Díaz, D. and Boj, C. *Social and physical interactive paradigms for mixed-reality entertainment*. Computers in Entertainment (CIE), Volume 4, Issue 2. ACM Press, New York, NY, 2006.
- [5] Liarakapis, F. An exploration from virtual to augmented reality gaming. *Simulation & Gaming*, Vol. 37, No. 4. SAGE Publications, 2006, 507-533.
- [6] Metaxas, G., Metin, B., Schneider, J., Shapiro, G., Zhou, W. and Markopoulos, P. SCORPIODROME: An Exploration in Mixed Reality. Social Gaming for Children. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology (ACE'05)*. ACM Press, New York, NY, 2005, 229-232.
- [7] Qui, T. C. T., Nguyen, T. H. D., Mallawaarachchi, A., Xu, K., Liu, W., Lee, S. P., Zhou, Z. Y., Teo, S. L., Teo, H. S., Thang, L. N., Li, Y., Cheok, A. D. and Kato, H. Magic Land: Live 3d Human Capture Mixed Reality Interactive System. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA, April 02-07, 2005). ACM Press, New York, NY, 2005, 1142-1143.
- [8] Zhou, Z., Cheok, A. D., Pan, J. H. and Li, Y. Magic Story Cube: an Interactive Tangible Interface for Storytelling. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology (ACE '04)*. ACM Press, New York, NY, USA, 2004, 364-365.
- [9] <http://mxrtoolkit.sourceforge.net/>
- [10] <http://recursos.cnice.mec.es/bancoimagenes/sonidos/>
- [11] <http://www.lego.com>

The Private City through the Hot Images

Cristina Portalés Ricart

Hot Images is an artistic mixed reality application that deals with the relation between human beings and city environments, thus proposing a novel cartography and navigation tool for the city. Within the virtual recreated environment of the Hot Images, the inert city is closed to the human field, transforming the cold aspect of the city into the hotness of the close friends and relatives. The user is carrying a laptop where a 3D virtual model of the city is presented within a blue color. Some images of past events related to the user are hidden in the city, and s/he has to find them. As the user gets closer to the images locations, the environment is changing to red color. Once the user arrives to a location, s/he can catch some images that will appear mapped on some of the buildings' facades, thus changing the aspect of the city. The location of the user is achieved with a combination of a GPS and an inertial sensor. The interface is done with the Max/MSP Jitter software.

T*he world is complex, dynamic, multidimensional; the paper is static, flat. How are we going to represent the rich visual world of experience and measurement on mere flat?' (Tufte 1990).* This is one of the questions that most worried the cartographers since the cartography exists. In the last years, the support for maps has evolved from a plain paper to the screen of a computer and traditional cartography has evolved to multimedia cartography (Cartwright 1999). Nowadays we can represent geospatial data through 3D models, and even visualize dynamic phenomena with some kind of simulators. The emergent techniques of virtual and augmented reality (Hedley et. al. 2002, Kolbe 2003) have enriched the maps, thus giving new tools for the user orientation.

The main purpose of a map is not other that to orient the users. In this sense, cartography has also evolved with the incorporation of the GPS and inertial sensors, which provide a navigation system for the so called location based services (LBS) and pedestrian navigation systems (Aslan et. al. 2004, Fang et. al. 2005, Gartner 2004, Retscher et. al. 2004). But there are maps with other different purposes, as those intended to visualize some amount of data related to the physical space (as for example, population, temperature, precipitation, etc.). Others are intended to classify the soil according to different criteria (natural resources, human activities, etc.). Topographic maps show the relief, whereas geologic maps show the underlying materials. At the time of making a map – which is a simplification of the reality – it is very important to know who are going to use it (general users, expertise users, children, etc.); depending on that, the cartographic

encoding/representation would be different. Then, could it be possible to do a customized map of the city, whose main purpose is to approach the city to the user personal field?

The Hot Images application tries to fulfill this task. I developed this application while I was living in a residential area called Beeston, characterized for not having anything special. It can be said that this area has low 'imageability'. This term was introduced by Lynch (1960) and can be described as the quality of a physical object, which can evoke in any observer a strong, vivid image. Then, a highly imageable city would be well formed, would contain very distinct parts, and would be instantly recognizable to the common inhabitant. The low temperatures at winter time made that the streets remained almost desert, especially at evening/night, so the daily activity of people was mainly indoors. Therefore, the city seemed to be cold and inert, like deprived of life. It was in this cold area where I met the idea of approaching the city to the personal scope, due to the necessity to give a distinguishing touch to the singsong surroundings thus changing the perception of the city.

1. The Hot Images application

Hot Images is a mixed reality application in which two different spaces are confronted: the exterior space, impersonal; and the private space, personal. The personal experiences are mixed with the urban surroundings by means of some images related to certain past events. These images are hidden in certain locations of the city. The user will have to explore his surroundings in order to find and recover them. In this way, the user is forced to travel around the entire area in order for him to meet new places. It is a way of exploring the city. Users discover the city and, at the same time, build their own private map. The way of finding the images can be compared to the popular child game of 'hot-cold', which consists of finding certain objects hidden in the surroundings. The one that has hidden the object is the one in charge to guide the player with these two words: 'cold' means that he is far away from the hidden object; 'hot' means that he is near to the objective. Hot Images uses this role to guide the user in his task of finding the images (the hidden object). The game relates the distance (cold) to the blue color and the proximity (hot) with the red color, thus given the user the clues to walk along the city streets (figure 1).

When the user arrives to a location where a set of images are hidden he has to look to the surroundings to find them, which are then mapped into some of the buildings' facades. Thus, the final mission of the user will be to turn the blue/cold city into a red/hot place. Then, the inert surroundings acquire human connotations.

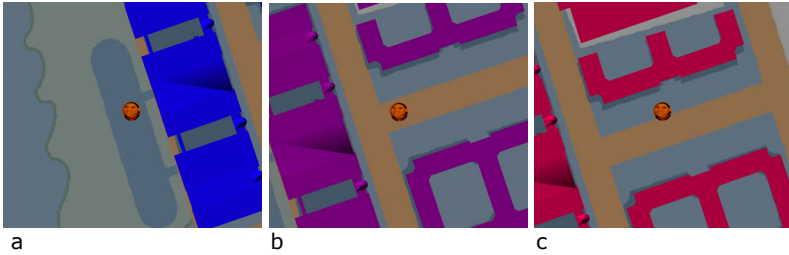


Figure 1 - Buildings' color depending on the distance to a hot location, with: a) > 100 m: cold; b) 30 m; c) < 10 m: hot.

2. Technical aspects

2.1. External devices

The mixed reality scenario is built with the following devices:

- Laptop: the user carries a laptop in which he sees the 3D model of the city.
- WebCam: A webcam is needed in order to find the images hidden in the surroundings, which is achieved with a pattern recognition algorithm. A camera resolution of 640 x 480 ppp is enough.
- GPS receiver: To achieve user location in real time, the Trine II GPS receiver was used. This receiver has the advantage that it can be connected to four different devices at the same time via Bluetooth. Then, a collaborative scenario could be built in which different users share the experience.
- Inertial sensor: for the orientation and tilting information, the MT9 Xsens inertial sensor was used. Data can be received in form of Euclidean angles, rotation matrix or quaternion.



Figure 2

2.2. Software environment

The application runs under the Max/MSP Jitter software. Inside the program, the different routines are organized into patches and sub-patches (figure 3). In the following sections, these are briefly explained.

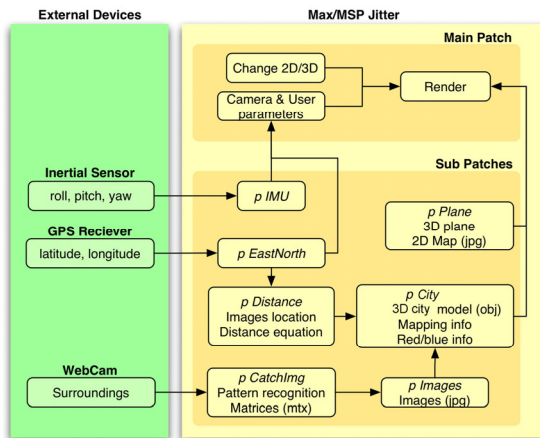


Figure 3 - Application architecture.

2.2.1. Main patch

The main patch consists of a window where the main tasks are performed. This is the actual window that the user can see when the application starts, and contains the following objects:

- *Change 2D/3D:* The user can choose at any time whether to visualize the application in the 2D or 3D view. In the 2D view, he sees the surroundings as in a map, thus he is able to see the buildings that remain out of his actual sight; at this view, the tilting information is not used, and the centre of the map is his actual location. The heading information can be activated or deactivated; in the first case, it leads to an egocentric map. In the 3D view, the model is shown according to the user real point of view. Tilting and heading information are used in this case.
- *Camera and user parameters:* the location is set with the positioning (Easting, Northing and height), the tilting (pitch and roll) and the heading information. In the 3D view, the user parameters are set to coincide with the camera ones, whereas in the 2D view the height is set different, thus the camera is placed at the top looking to the nadir and pointing to the user. Therefore, in the 2D view, the centre of the map is the location of the user. Users are represented as simple spheres mapped with a picture of their own face, so they are the ones exploring the virtual world.
- *Render:* This is the OpenGL window where the models and images are rendered.

2.2.2. Sub-patches

The sub-patches are some windows that the user is not able to see, but that are necessary in order to run the application. These are:

- *p IMU*: In this patch the data from the inertial sensor is received.
- *p EastNorth*: This patch transforms the received latitude and longitude values from the GPS into local X and Y coordinates.
- *p Distance*: Here the distance from the user to all the images is

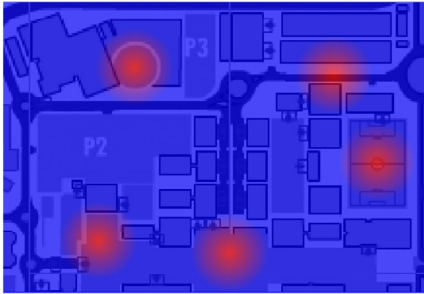


Figure 4

calculated. The shortest one is calculated. The shortest one is transformed into a color (from red to blue). When the user is more than 100 m far away from a target, the scene is seen completely blue; as the user get closer to an image, the 3D world turns more and more red. This patch will act as if it were circular buffers defined around images (figure 4).

- *p Catching*: In this patch, a pattern recognition procedure is applied to the incoming image, in order to find the correspondence with a stored matrix. When the matrix and the webcam image match, then the user recovers one of the hidden images and is automatically mapped to some of the buildings' facades.
- *p Plane*: This patch contains an OpenGL plane with its texture being the map of the area. For the example shown (figures 1 and 3), a map vectorized in a cartoon style was used.
- *p City*: a 3D model of the surroundings is needed in the *obj* format in order to be read by the program. The model consists on a simple representation of the buildings but without texture. In its place they appear colored in tones that vary from the red to the blue, indicating the distance or proximity to the images. The mapping information (i.e., how to map) is given to them beforehand (for example, with the 3DStudioMax before being exported to *obj*), although the texture (i.e., what to map) is applied afterwards in the Max/MSP Jitter.
- *p Images*: In this patch the images are called. Once the user arrives to a hot location and finds correspondence with the incoming webcam image, the hidden image is called and automatically mapped.

3. Testing the system

The application was first tested at the campus of the Polytechnic University of Valencia (Spain). People working at the Laboratorio de Luz

were asked to be involved in testing the system, so they were already familiar with the campus. The content of the images was selected in order to give the audience a different perception of this environment. All images showed people inside the buildings of the University in different situations (studying, working, eating, etc.). They tell us diverse personal stories that happen in the interior of the buildings, thus making 'visible' from the outside the human relationship with the inert constructions. In the following figure, some of the selected images can be seen.



Figure 5 – Some images used in testing the system.



Figure 6

As the user is approaching a hot location, the environment turns into red (figure 6 - top). Images related to that hot location are then mapped into some of the buildings (figure 6 - middle). Once found, images will remain visible even if the user goes further away (figure 6 - bottom).

Some of the audience noticed that the proposed game of 'hot-cold' is an interesting navigation tool for exploring the campus, and the relationship between proximity and color resulted to be quite intuitive. The game of hot-cold was familiar to all users participating in the performance, therefore only a brief introduction was needed to explain the role of the proposed game.

It has to be pointed that color changes in a gradual way, there are not jumps between colors. As a navigation tool, it can be said that when using color in this way is not fast to arrive to a location, as there are not

arrows or other clear indications that guides the users through a specific direction. These means, the user has to really explore the environment, and destination is not known until he arrives. At the beginning, audience was a little bit confused about this fact, because they expected a straight forward navigation, and also were somehow anxious to know the destination. Some users had also problems finding hot locations due to color, as they thought to have arrived to a hot location before the environment had entirely turned into red.

4. Discussion

Image content can be discussed. Depending on the expected audience and the purpose of the exhibition, different contents can be chosen. Very different topics can be thought, for example: images about the history of a city or an specific past event related to the city, if the application is to be shown for a general audience living or arriving there; images of fauna/flora, if the purpose is to denunciate the abuse of constructions to the detriment of wild/green areas; images of a private party or common past event, if the audience is a specific group; etc. Using 2D images to construct a virtual 3D world is not new, and a good example can be found in the projects 'Field-Works' of Masaki Fujiata.

Whether the hot-cold game is known in different cultures or if red can be easily related to hot and blue to cold is a fact that would need some further discussion. In order to explore into the universality of these facts, I prepared a short test and asked some friends from different countries. The questions were:

1. Are you familiar with the hot-cold game?
 - a. Yes
 - b. No
2. Which option has more sense to you?
 - a. Hot can be related to red, and cold to blue
 - b. Hot can be related to blue, and cold to red
 - c. I cannot relate hot-cold with any color
 - d. I would relate hot-cold with other colors (say which)

People participating in the test come from: Germany, Austria, Italy, Brazil, Finland, USA, Singapore, China, and of course, Spain. To the first question, all of them answered 'a' out of the one coming from China. To the second question, the answer was 'a' in all cases, but the Chinese guy specified 'dark blue'. Although this short test can't be seen as a rigorous surveying, it gives an idea about the possibility to extend the proposed navigation tool into other cultures.

Anyway, whether audience is familiar with the hot-cold game or not, shouldn't be a problem when understanding the navigation tool, as the game is very simple and intuitive, thus needing short explanation.

Whether the chosen red-blue colors can be understood as hot-cold is a question of visual culture (Heller 2004), and maybe people from other countries would use other colors to describe it. If this were the case, it would very simple to change the color information in the Hot Images application in order to be better understood by the audience.

5. Conclusion

Although the main purpose of a map is to orient the users, a huge variety of facts are visualized within cartography, and lots of purposes can be defined. In the Hot Images application, a new relationship between users and cartography has been proposed, thus altering the perception of the environment: the impersonal city is closed to the personal scope by virtually changing the look of its buildings. A popular childish game is proposed as a navigation tool: users have to find some images hidden at specific locations (hot locations) with the only clues of 'hot-cold', which are visualized with the colors 'red-blue'.

Therefore, users are guided through the city according to these color clues. The aim is to turn the inert environment into a hot place, providing the buildings with human connotations. At hot locations images related to the audience are hidden. Then, when a hot location is reached, some images are discovered and mapped on some of the buildings' facades, thus acquiring human connotations.

The application is build into the Max/MSP Jitter environment. For the location of the user, a GPS receiver and an inertial sensor are in use. A webcam is also needed in order to find the hidden images that are recognized via a pattern recognition algorithm.

References

- Aslan, I., and Krüger, A. (2004) The Bum Bag Navigator (BBN): An advanced pedestrian navigation system. In AIMS 2004, Nottingham, UK.
- Cartwright, W., Peterson, M., and Gartner, G. (1999) Multimedia Cartography. Berlin: Springer Verlag.
- Fang, L., Antsaklis, P. J., Montestruque, L., McMickell, M. B., Lemmon, M., Sun, Y., Fang, H., Koutroulis, I., Haenggi, M., Xie, M., and Xie, X. (2005) Design of a Wireless Assisted Pedestrian Dead Reckoning System - The NavMote Experience. IEEE Trans. Instrumentation and Measurement, Vol. 54, pp 2342-2358.
- Gartner, G. (2004) Location-based mobile pedestrian navigation services - the role of multimedia cartography. ICA UPIMap, Tokyo.
- Hedley, N. R., Billinghurst, M., Postner, L., and May, R. (2002) Explorations in the use of Augmented Reality for Geographic Visualization, Presence: Teleoperators and Virtual Environments, Volume 11 (2). The MIT Press, Cambridge, pp 119-133.
- Heller, E. (2004) Psicología del color. Ed. Gustavo Gili, SA, Barcelona.

Kolbe, T. (2003) Augmented Videos and Panoramas for Pedestrian Navigation. In LBS & TeleCartography, G. Gartner, Ed. Geowissenschaftliche Mitteilungen, vol. 66, pp 45-52.

Lynch, K. (1960) The Image of the City. The MIT Press. Cambridge, Massachusetts.

Retscher, G., and Thienelt, M. (2004) NAVIO – A Navigation and Guidance Service for Pedestrians. Journal of Global Positioning Systems, Vol. 3, No. 1-2, pp 208-217.

Tufte, E. R. (1990) Envisioning Information. Graphics Press, Cheshire, Connecticut.