

# Aplicacions per a un diccionari de gestos base d'un sistema de comunicacions

*Autor:* M<sup>a</sup>Ángeles Calabuig Navarro

*Director1:* Valery Naranjo Ornedo

*Resum* — L'objectiu d'aquesta tesina és la implementació de diferents aplicacions necessàries per a l'estudi prospectiu d'un diccionari de gestos genèric que controla els elements de la interfaç gràfica de terminals que proporcionen major sensació d'immersió així com una major accessibilitat als usuaris.

La interfaç del terminal d'usuari d'un sistema de comunicacions de vídeo de nova generació és gestual i es pretén que els gestos que controlen aquesta interfaç gràfica 3D siguin el més naturals possible. D'ací va sorgir la idea d'aquesta tesina.

En aquest document s'explica l'origen i en què consisteix breument (secció 1), així com el sistema general per al que s'implementa (secció 2). A continuació es descriu el plantejament inicial de l'activitat "Interfaç d'usuari" dins d'aquest context (secció 3) i les diferents parts que la componen, començant amb la interfaç gestual a la secció 4 i seguidament amb la interfaç gràfica a la secció 5. Finalment, es presenten les diferents aplicacions gràfiques que s'han implementat per a aquest diccionari gestual.

Autor: Calabuig Navarro, M<sup>a</sup>Ángeles, email: [mcalabuig@labhuman.i3bh.es](mailto:mcalabuig@labhuman.i3bh.es)

Director 1: Naranjo Ornedo, Valery, email: [vnaranjo@dcom.upv.es](mailto:vnaranjo@dcom.upv.es)

Data d'entrega: 4-12-09

**ÍNDEX**

I.	INTRODUCCIÓ .....	5
II.	SISTEMA GLOBAL .....	6
III.	PLANTEJAMENT INICIAL DE L'ACTIVITAT .....	7
IV.	INTERFAÇ GESTUAL.....	7
IV.1	Gestos Reconeguts .....	8
IV.2	Consideracions del nou diccionari .....	8
IV.3	Experiment accions-gestos .....	9
IV.4	Interacció .....	12
V.	INTERFAÇ GRÀFICA .....	13
V.1	Aplicacions proposades .....	13
V.2	Brainstorm eStudio .....	16
V.3	FUNCIONAMENT.....	17
V.4	EVOLUCIÓ .....	21
V.5	INDEPENDÈNCIA .....	22
V.6	MICRODEMOSTRADORS .....	22
V.6.1	<i>Estela</i> .....	24
V.6.2	<i>Pissarra</i> .....	25
V.6.3	<i>Focs artificials</i> .....	26
V.6.4	<i>Cotxe</i> .....	27
V.6.5	<i>Vídeos</i> .....	29
V.6.6	<i>Joc</i> .....	32
V.6.7	<i>Slider horitzontal i vertical</i> .....	34
V.7	POSSIBLES MILLORES .....	35
VI.	CONCLUSIÓ.....	36
	AGRADEIXEMENTS .....	37
	BIBLIOGRAFIA.....	39
	Annexes.....	41

**ÍNDIX DE FIGURES**

Fig. 1 – Sistema global .....	6
Fig. 2 – Acceptar/Seleccionar opció 1 .....	9
Fig. 3 – Obrir diferents menús.....	10
Fig. 4 – Obrir/Tancar menú.....	10
Fig. 5 – Interacció horitzontal i vertical .....	12
Fig. 6 – Interacció gest-interfaç.....	12
Fig. 7 – Zoom de mapes .....	13
Fig. 8 – Canves amb ferramentes .....	14
Fig. 9 – Explorador de fotos .....	14
Fig. 10 – Moviment de partícules.....	14
Fig. 11 – Salva pantalles.....	15
Fig. 12 – Llista de llums per defecte .....	16
Fig. 13 – Pantalla inicial.....	17
Fig. 14 – Nombre de botons configurable.....	18
Fig. 15 – Selecció d'un microdemostrador.....	18
Fig. 16 – Eixida d'un microdemostrador.....	19
Fig. 17 – Gest per canviar al panell de l'esquerra .....	19
Fig. 18 – Distinció pel color del panell .....	20
Fig. 19 – Pantalla completa per 1 panell .....	21
Fig. 20 – Proves desenvolupament interfaç.....	21
Fig. 21 – Microdemostrador descartat.....	22
Fig. 22 – Estela.....	24
Fig. 23 – Usuari amb l'estela.....	24
Fig. 24 – Pissarra virtual.....	25
Fig. 25 – Diferent color pissarra.....	25
Fig. 26 – Usuari amb la pissarra.....	26
Fig. 27 – Fons negre .....	26
Fig. 28 – Fons blau .....	27
Fig. 29 – Gest de moure el cotxe / Gest de selecció de color.....	27
Fig. 30 – Usuari movent cotxe cap avall .....	28
Fig. 31 – Major nombre de colors de la paleta .....	28
Fig. 32 – Aspecte final .....	29
Fig. 33 – Vídeos .....	29
Fig. 34 – Moure's a esquerra (ambdues opcions).....	30
Fig. 35 – Zoom in .....	30

Fig. 36 – Gest Zoom in/out .....	31
Fig. 37 – Alternativa gest Zoom in/out .....	31
Fig. 38 – Joc .....	32
Fig. 39 – Puntuació màxima.....	33
Fig. 40 – Possibilitats Gest-Joc .....	33
Fig. 41 – Slider vertical i horitzontal.....	34
Fig. 42 – Seqüència de gestos sliders .....	35

## **ÍNDIX DE TAULES**

Taula 1 – Accions i gestos amb major percentatge d'ús .....	11
Taula 2 – Accions i gestos per al menú.....	20
Taula 3 - Relació microdemostrador-gestos.....	23

## I. INTRODUCCIÓ

Aquesta memòria forma part de l'entorn d'un projecte CENIT<sup>1</sup> anomenat VISION del Ministeri d'Indústria, Turisme i Comerç gestionat pel CDTI (Centre per al Desenvolupament Tecnològic Industrial) emmarcat dins del programa Ingeni 2010 del Govern espanyol per augmentar la inversió en investigació i desenvolupament (I+D).

Actualment la tecnologia està present a la societat que ens rodeja, i encara que ja existeixen mitjans prou sofisticats per a comunicar-se, es pretén aconseguir un sistema de comunicació de vídeo que transmeti una sensació de presència real amb alta qualitat.

Ja han sigut desenvolupats molts sistemes de videoconferència i gràcies a ells es poden evitar desplaçaments innecessaris amb la implementació d'un entorn en el qual no existisquen diferències d'interlocució amb una persona situada físicament al teu costat d'una altra que està en qualsevol altre lloc remot. No obstant, molts usuaris d'aquests sistemes encara senten que interaccionen amb una màquina en lloc de sentir-se immersos en el entorn reproduït i parlant amb la gent que està als altres terminals enfront seu. Per aquesta raó eviten l'ús d'aquests sistemes i prefereixen traslladar-se presencialment a on es realitzi la reunió. En canvi, el sistema global de videoconferència del què forma part aquesta interfaç pretén que els usuaris estiguen immersos el major possible sentint-se còmodes en qualsevol entorn.

Al principi, es pensà quines situacions podien representar millor aquest sistema. D'aquest mode, van ser diversos els escenaris plantejats per a testear-lo, des de els englobats en un entorn domèstic com pot ser qualsevol familiar en el qual els seus membres es troben separats per algun motiu i necessiten sentir-se prop, fins els d'un nivell més professional. Per exemple, un professor de ioga que impartisca les seues lliçons des de la seua pròpia casa o una orquestra on cada músic s'ubica en una ciutat diferent i el director en una altra. En qualsevol d'aquests entorns i amb aquest sistema, els usuaris de la videoconferència no deuen sentir que interactuen amb màquines sinó que viuen una experiència equiparable a una tradicional on totes les persones estan realment a la mateixa habitació o lloc físic, realitzant-se estudis d'usabilitat per aconseguir el prototip desitjat.

Aquest sistema ha d'estar ben dissenyat i reflectir la major part de característiques que es donen en les situacions comuns independentment de que hi haja o no tecnologia. Algunes d'aquestes són el contacte visual, la interactivitat real mitjançant gestos naturals, l'absència de retards, la profunditat i la comunicació no verbal.

---

<sup>1</sup> CENIT: Consorcis Estratègics Nacionals en Investigació Tècnica, on es financien grans projectes integrats en investigació industrial de caràcter estratègic, gran dimensió i llarg alcanç científic-tècnic orientats a una investigació planificada en àrees tecnològiques de futur i amb projecció internacional.

Finalment, aquest sistema es divideix en diversos mòduls. Per una banda està la part que es troba en contacte directe amb els usuaris als extrems de la comunicació, l'anomenada terminal d'usuari, on està inclosa la interfàç desenvolupada. Altres mòduls s'encarreguen de permetre la interacció amb l'usuari per a la provisió dels serveis oferits, així com de capturar i sintetitzar les escenes. Per una altra banda com es lògic existeix una codificació i descodificació dels fluxos de vídeo generats que s'envien a través d'una òptima xarxa de comunicacions.

## II. SISTEMA GLOBAL

En el sistema global de videoconferència, cadascun dels mòduls que el formen tenen unes determinades funcions bàsiques enumerades a continuació.

- Mòdul de captura: Encarregat de l'adquisició de les imatges i el so, així com de la generació del model tridimensional de l'escena i de les tasques de seguiment i reconeixement gestual dels interlocutors [1][2].
- Mòdul de reconstrucció: Realitzador de la síntesi del model 3D mitjançant algoritmes de renderització de l'escena a partir del model generat en captura [3][4][5][6].
- Mòdul de presentació: Responsable de visualitzar l'escena 3D reconstruïda a l'usuari remot mitjançant una interfàç gràfica que apareix a les pantalles del que es disposa [7][8][9][10].
- Mòduls de codificació i descodificació: Processadors dels fluxos de vídeo generats pel mòdul de captura abans de la transmissió a la xarxa (codificació) i després d'ella, abans de donar-los en l'altre extrem al mòdul de reconstrucció (descodificació), optimitzant d'aquesta forma l'ús de la xarxa.
- Mòdul de comunicacions: Encarregat de l'establiment de les comunicacions entre els diferents sistemes d'usuari.

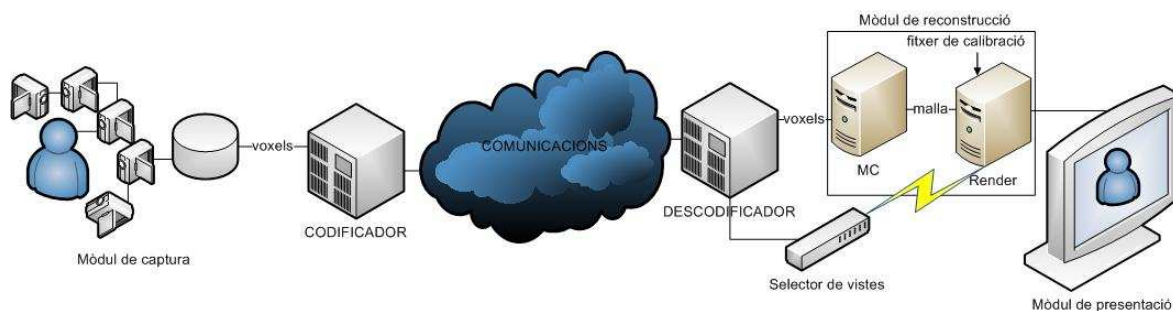


Fig. 1 – Sistema global

Inicialment s'implementà un sistema punt a punt en aquest context, passant després a multipunt, al qual se li han aplicat les tècniques més innovadores en cadascun dels mòduls, el que ha permès

publicar articles sobre el sistema. A més a més s'han definit les interfícies òptimes per a la interoperabilitat entre els diferents mòduls, així com la interfaç d'usuari considerada com la més apropiada.

Després d'una implementació inicial de la interfaç d'usuari d'aquest sistema de videoconferència, es demostrà que l'entrada més natural aconseguida era mitjançant una interfaç gestual que permet el control de la interfaç gràfica tridimensional. No obstant, encara que s'aconseguia accedir a les aplicacions que incloïa, en certs punts no s'intuïa fàcilment el que es tenia que fer i alguns usuaris acabaven desesperant-se. És a dir, aquesta aplicació de videoconferència era manejable mitjançant gestos que no eren fàcils de recordar tal com es veu amb les proves realitzades als usuaris i tal com explica la teoria d'aprenentatge [11][12][13][14], ni intuïtius per a tot tipus d'usuaris, incloent qualsevol edat i professió. Per aquesta raó començà l'activitat que s'explica a aquesta tesina com a millora de diversos aspectes de la interfaç d'usuari .

### **III. PLANTEJAMENT INICIAL DE L'ACTIVITAT**

En primer lloc, es plantejà un estudi d'un diccionari de gestos genèric per aconseguir distingir entre aquells gestos que foren naturals i intuïtius per a qualsevol usuari i per tant, millorar el sistema de videoconferència. Com a conseqüència, va sorgir la necessitat d'implementar diferents aplicacions gràfiques en les que els usuaris pogueren interaccionar i utilitzar els diferents gestos per a realitzar determinades accions com pot ser l'acció d'obrir o tancar un menú i d'aquest mode aconseguir major sensació d'immersió.

Quan es va començar en aquesta activitat coneguda com "Interfaç d'usuari", sortiren diverses qüestions com quins gestos es tenien que reconèixer i com es tenien que realitzar, quines accions s'havien de dir als usuaris que realitzaren per estudiar els gestos que realitzaven. També quines interfícies gràfiques eren les més adequades, així com el seu aspecte de visualització i la manipulació dels elements gràfics, per això es proposaren inicialment una sèrie de idees tant per a la interfaç gestual com per a la interfaç gràfica.

En esta activitat es diferenciaven principalment tres parts: la de reconeixement de gestos, la de interfaç gràfica i per últim, la de factors humans, on s'estudià el seu disseny i s'avaluà el sistema resultant.

### **IV. INTERFAÇ GESTUAL**

Respecte a la interfaç gestual, es a dir, als gestos emprats per a interactuar amb la interfaç gràfica que posseeixen els terminals [15], es parteix dels gestos utilitzats inicialment no considerats manejables i per això, es proposa millorar-los i afegir uns altres nous si fóra necessari.

### ***IV.1 Gestos Reconeguts***

El diccionari de gestos que ja es reconeixia abans de començar amb aquest estudi, estava format per gestos que deuriem realitzar-se amb una única mà i la palma orientada a la càmera. Entre aquestos gestos inicials estaven els següents:

- Selecció: usat per a escollir qualsevol objecte.
- Desplaçament vertical cap a dalt i cap a baix, així com en direcció horitzontal a esquerra i dreta de qualsevol objecte.
- Cercle a dretes i a esquerres: Traçar un cercle en una mà en una de les dos possibles direccions.
- Trajectòria recta diagonal de la mà tant des de dalt com des de baix cap a la dreta i esquerra.

Aquest sistema de reconeixement gestual inicial permet obtindre la posició aproximada del centre de masses de la mà detectada en tres dimensions (x,y,z), així com té la possibilitat d'introduir gestos que utilitzen la distància entre la càmera i la mà. No obstant, un inconvenient d'aquest sistema és que no permet utilitzar les dos mans [16].

### ***IV.2 Consideracions del nou diccionari***

Per a millorar el diccionari de gestos inicial s'inclogueren gestos usant tant una mà com les dos mans, ja que el resultat d'usar ambdues és més aparent i en moltes aplicacions necessari. Encara que no s'ha d'oblidar que els gestos amb les dos mans a nivell d'usabilitat requereixen una major complexitat per als usuaris i pot reduir el nombre d'usuaris que pot utilitzar dita aplicació, per la qual cosa si existeixen casos que es poden resoldre amb una única mà, pot resultar una aplicació més usable.

Entre els gestos que es consideren útils per al sistema de videoconferència inicial que conté aplicacions com el joc del tres en ratlla o el “*photolog*” foren els d'apuntar o seleccionar, el de seguiment d'algun objecte, el d'agafar i el de rotació en totes les direccions tant horitzontal com verticalment (dalt, baix, esquerra o dreta).

A més a més, es realitzà una classificació bàsica [17] amb dos tipus de gestos: els “atòmics” (sense informació associada darrere d'ells) com per exemple prémer qualsevol tecla tal com “Play”, “Stop” o indicar una posició determinada del cursor (x, y, z); i els “analògics”, que tenen una informació associada ja que per exemple impliquen direcció, com pot ser un moviment del cursor, arrossegar-lo, polsar un click o doble click.

Per una altra part, es consideraren gestos amb Z, es a dir, on es té en compte també l'altura o profunditat. Per exemple alçar o baixar una mà ja que si la pantalla està situada frontalment en



vertical i es desitja seleccionar algun element, el gest més natural és el d'apropar la mà sense doblar la monyica, es a dir, com si es cridara al timbre d'una porta, pel que s'usa l'eix Z.

En quant a la distància de l'usuari a la pantalla, es plantejà una interfaç de gestos desvinculada de les dimensions i distància de la pantalla, es a dir, per un costat els gestos "atòmics" deuen de comportar-se com símbols desvinculats i els gestos "analògics" deuen anar lligats a l'espai d'interacció de l'usuari, per exemple, un pam de recorregut de la mà correspondria amb 100 píxels en pantalla independentment de la distància a la que se es trobe d'ella.

Partint de totes aquestes especificacions considerades necessàries, es va fer una recopilació d'accions i es realitzà un experiment amb un cert grup d'usuaris per a agrupar els gestos associats a cadascuna de les accions definides estimant el percentatge corresponent d'usabilitat per a cadascun.

### IV.3 Experiment accions-gestos

El procés de confecció d'aquest diccionari començà amb la realització d'un experiment amb un grup variat d'usuaris sense cap aplicació davant, utilitzant diferents objectes (un bidimensional, un altre tridimensional i un slider) de forma que se'ls demani als usuaris unes determinades accions per obtindre una base més sòlida de gestos reals utilitzats per un sector de la població representatiu. Aquestos resultats reflecteixen quins gestos s'utilitzen amb major freqüència per a realitzar cadascuna de les accions.

Una vegada fet aquest estudi, s'acordà limitar l'ampli nombre de gestos analitzats que s'arreglaven en un document anomenat "*Usabilitat al llenguatge gestual*" [18] a una sèrie de gestos que es consideraren més útils per al nostre sistema.

Por exemple es va veure com els gestos més emprats quan als usuaris se les deia que acceptaren alguna opció o que seleccionaren algun objecte eren els que apareixen en la figura següent per a cada cas. Per aquesta raó com a primera opció s'implementaren els següents gestos.

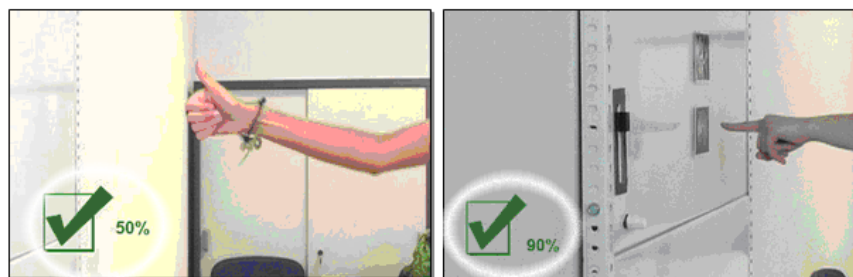


Fig. 2 – Acceptar/Seleccionar opció 1

Un altre exemple important és el gest d'obrir i tancar un menú, ja que per a accedir a totes les aplicacions dels sistema de videoconferència (la videoconferència tal qual, el photolog i el joc) s'utilitzen aquestos gestos, pel que la freqüència d'ús és elevada.

A la "Fig.3" es veuen diferents situacions on és necessari usar aquest gest. A la part superior esquerra es tracta de l'accés inicial a qualsevol de les tres aplicacions; a la superior dreta es

visualitza el moment de la videoconferència on s'ha d'escollir amb quin usuari es vol mantindre la comunicació (A, B o A3D); a la part inferior dreta reflexa el moment que es desitja canviar a una altra aplicació diferent de la del photolog que és la que s'està usant en eixe moment. Per últim, en el cantó inferior esquerra es mostra quan s'està jugant i es necessita iniciar una videoconferència amb un altre usuari pel que és necessari obrir el menú d'elecció de l'usuari.

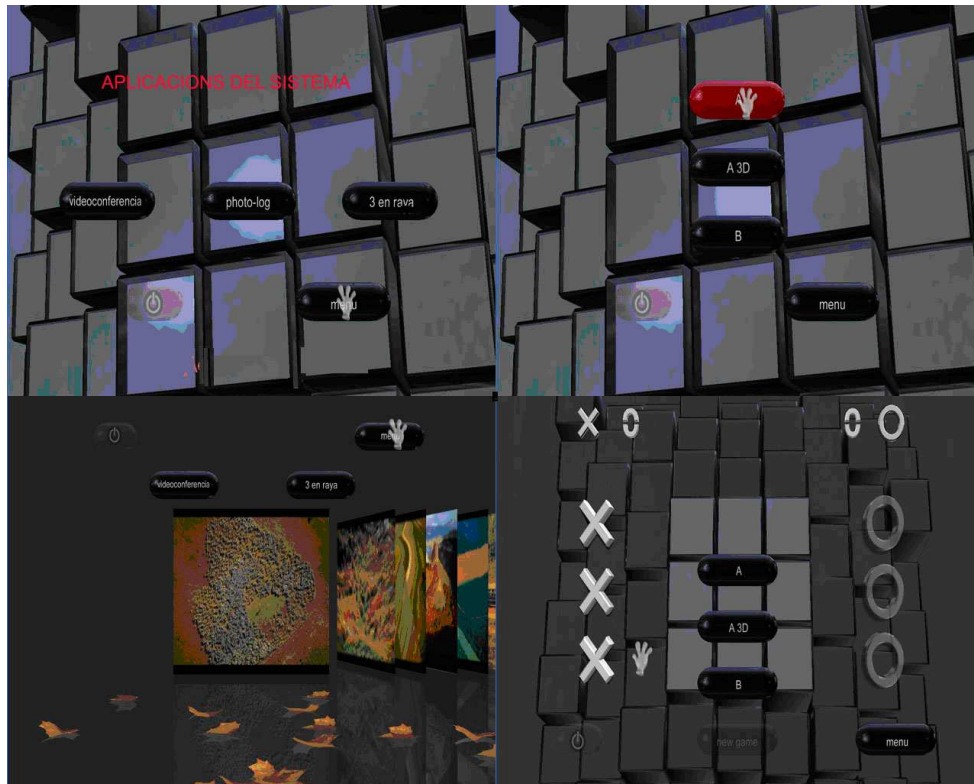


Fig. 3 – Obrir diferents menús

El percentatge que es va obtenir d'usabilitat per aquest gest d'obrir i tancar el menú és el que es mostra en la figura de baix i aquest gest deurà capturar-se amb independència del lloc de la pantalla en el que siga efectuat.

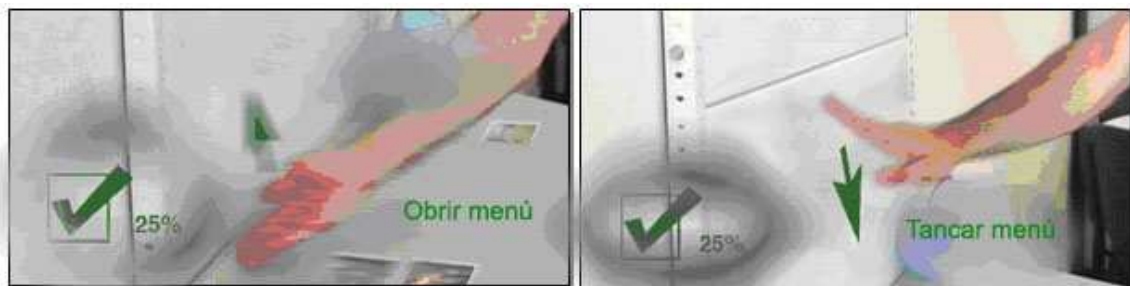




Fig. 4 – Obrir/Tancar menú

Aquest mateix procediment es segueix per a totes les accions que es necessitaven, obtenint els gestos que componen el diccionari. D'aquesta forma es recolliren tots els gestos amb major percentatge d'ús associats a cada acció i amb aquesta llista que es mostra en la taula següent, es va passar a dissenyar la part d'interfaç gràfica que es descriu al següent apartat.

Accions	Gestos amb major percentatge d'ús
<i>Seleccionar o apuntar alguna cosa</i>	- Selecció (dit índex estirat i apuntant on es vol, resta de dits tancats) 
<i>Moure's a dreta o a esquerra (per exemple passar de panell o de vídeo)</i>	- Puny tancat amb el dit polze estirat senyalant a dreta o a esquerra per passar al panell de la dreta o esquerra respectivament  - Dit índex i el del mig estirats per moure's a esquerra i dreta 
<i>Agafar un objecte i girar o moure'l en qualsevol direcció</i>	- Palma de la mà oberta amb dits estirats i movent-la en direcció corresponent  - Mà oberta amb dits estirats per agafar l'objecte, puny tancat per moure'l 
<i>Soltar un objecte</i>	- Obrir la mà i apartar-la cap a darrere 
<i>Traçar una trajectòria</i>	- Mateix gest que el de selecció però continu marcant la trajectòria
<i>Fer zoom d'un vídeo</i>	- Puny tancant amb el dit polze cap a dalt 
<i>Acabar amb el zoom (parar-lo)</i>	- Palma de la mà oberta i plantada cap a dalt 

Taula 1 – Accions i gestos amb major percentatge d'ús

Per últim, es va implementar un sistema de reconeixement format pels gestos amb major percentatge d'ús o considerats necessaris per al sistema en qüestió. Aquest reconeixement es basa

en la captura de la silueta de la mà o mans utilitzades per a efectuar el gest, obtenint la seua màscara, a continuació es busca el seu vector de característiques que modela la mà o mans i es reconeix el gest dinàmic [19]. Finalment, aquesta informació se li envia a la interfaç gràfica mitjançant el protocol UDP<sup>2</sup> [20].

#### IV.4 Interacció

Un altre aspecte important de ser considerat, va ser la interacció gest-interfaç, determinant que la pantalla d'interacció per a aquest estudi fora vertical en lloc d'horitzontal, que era com es situà al terminal d'usuari del sistema de videoconferència inicialment, per a comparar-la posteriorment amb una interacció vertical.

Donat que existeixen diferències significatives entre els moviments realitzats en una pantalla vertical d'una horitzontal, s'analitzen tant els moviments en vertical com en horitzontal.



Fig. 5 – Interacció horitzontal i vertical

Respecte a la càmera de captura, es deu situar a la part inferior de la pantalla, tal i com es mostra a la "Fig.6", de manera que els gestos siguin capturats des d'un angle inferior.

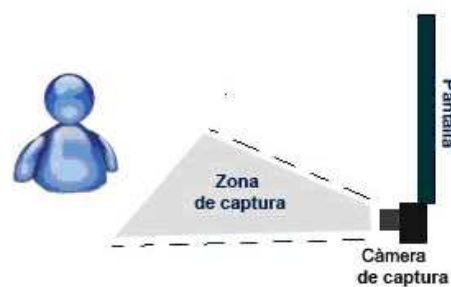


Fig. 6 – Interacció gest-interfaç

---

<sup>2</sup> UDP ("User Datagram Protocol"): Protocol del nivell de transport basat en l'intercanvi de datagrames no orientat a la connexió. No té confirmació ni control de fluxos, per la qual cosa els paquets poden avançar-se els uns als altres i tampoc es sap si han arribat correctament.

## V. INTERFAÇ GRÀFICA

Pel que respecta a la part de la interfaç gràfica, es plantejaren diverses propostes d'aplicacions útils per als gestos que es pretenien estudiar i amb totes elles, es començà amb la seua implementació.

### V.1 *Aplicacions proposades*

En un principi, algunes aplicacions es consideraren que podien ser útils a l'hora de dissenyar la biblioteca de gestos. Aquestes aplicacions devien englobar totes les accions i gestos que formaren el diccionari ja que són els que s'usen per a realitzar l'estudi.

Per exemple, es necessiten aplicacions on s'esculla algun element, es gire un cub o qualsevol altre objecte de forma vertical (eix Y dalt-baix-dalt) o horitzontal (eix X dreta-esquerra-dreta), es mogui un dels elements a una altra part de la pantalla, es faça zoom in/out en una fotografia o vídeo, s'oculten els elements (com si foren els continguts d'un menú i es volguera arrastrar cap avall aquesta pantalla per a ocultar tots eixos elements i veure'n uns altres diferents) o bé, el cas contrari, mostrar uns elements que no estaven a la pantalla.

A continuació es mostren aplicacions on s'usen les accions acabades de mencionar i que serveixen per a reflectir el que s'ha de fer a la interfaç gràfica implementada.

- Sistema de menús HTC: Sistema de menús en tres dimensions aconseguint un aspecte tridimensional amb forma de T on es poden trobar les successives aplicacions, a les quals es pot accedir desplaçant el dit a esquerra, dreta, dalt i/o baix en una pantalla tàctil. Amb això es pretén no pensar únicament en menús plans sinó també en gestos per a menús 3D. Existeix un exemple de vídeo [21] amb la interfaç gràfica del “*HTC Touch Diamond: TouchFLO 3D*”, així com de la interfaç gràfica inicial de HTC [22].
- Sistema de “panning”<sup>3</sup>, zoom i gir de mapes: Com el seu nom indica, l'aplicació consisteix bàsicament en moure mapes, ja siga apropant-los o separant-los de la vista de l'usuari (Fig. 7).

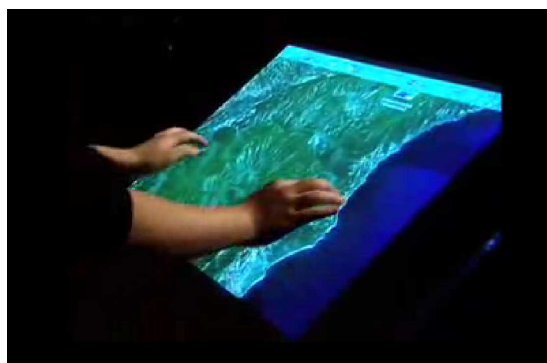


Fig. 7 – Zoom de mapes

---

<sup>3</sup> “Panning”: Moviment perpendicular a la vista.

- Canves amb ferramentes: Lloc on es poden usar ferramentes com la selecció de colors per a dibuixar i esborrar-los usant una mà, tal i com s'observa en la figura següent.

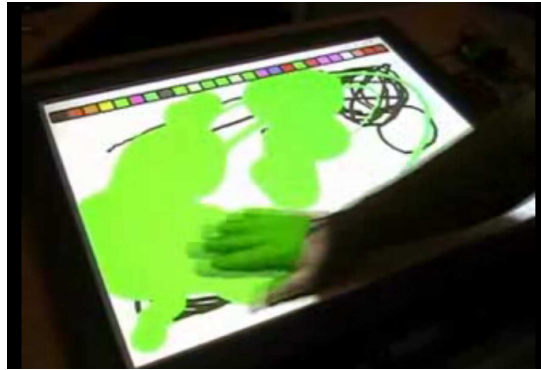


Fig. 8 – Canves amb ferramentes

- Explorador d'una col·lecció de carpetes d'imatges: Col·lecció de fotos col·locades per la pantalla que es poden moure a qualsevol lloc d'aquesta, així com modificar la seua mida ampliant-les i reduint-les, canviar-les el color i altres característiques de les imatges.



Fig. 9 – Explorador de fotos

- Moviment de partícules: Conjunt de partícules que l'usuari pot moure, canviar-li el tamany, experimentar amb elles sense tindre realment cap objectiu final, tal com il·lustra la imatge de baix.

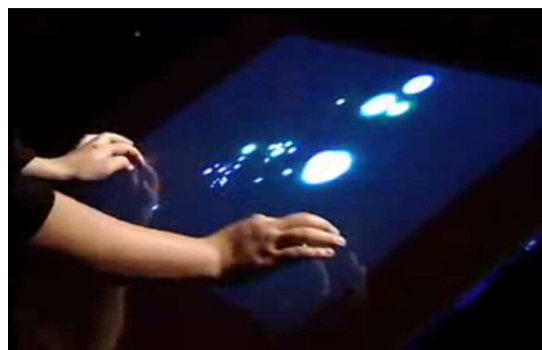


Fig. 10 – Moviment de partícules

- “Salva pantalles” canviant de forma i vistosos, que captin l’atenció, tal com apareix a la figura següent.



Fig. 11 – Salva pantalles

- “*Virtual Wall*”: Mur virtual infinit per a navegar per fotos, vídeos i altres aplicacions. Un exemple es el realitzat per “*Piclens*” [23].

En qualsevol d’aquestes possibles aplicacions es poden incloure la reproducció de sons per a simular millor la realitat. A partir de tots aquests exemples d’aplicacions es va implementar la interfàç gràfica amb els microdemostradors que s’expliquen més avant.

Considerant les idees sobre aplicacions comentades en l’apartat anterior, es començà amb el desenvolupament d’una interfàç basada en un menú HTC amb el funcionament base que aquest porta associat ja que és la solució més coherent per al que es pretenia aconseguir. Cadascuna de les aplicacions anteriors podia pensar-se tant com un interfàç com aplicació en concret accessible des de la interfàç.

El motiu de l’elecció d’aquest tipus d’interfàç HTC és l’estructura organitzada que presenta. Respecte a les interfícies de tipus "sistema de panning, zoom i gir de mapes" estan molt orientades al tipus d’aplicació en les que s’usen però no tenen un ús general, i les del “explorador d’una col·lecció de carpetes de fotos”, són interfícies sorprenents, no contrastant això, la navegació d’elements que mostra no és funcional, és a dir, amb un nombre xicotet d’elements, la llibertat que aquest tipus d’interfàç proporciona a l’usuari és fantàstica però si el nombre és elevat i es necessita ordenar-los o localitzar un en concret, no és una interfàç pràctica. En canvi, no es va descartar la idea de contemplar-lo com una opció d’entreteniment o microdemostrador dels gestos per als usuaris i no com una forma d’interacció funcional i real.

Una altra raó per al seu desenvolupament és la desaparició dels elements que no són necessaris per a l’usuari, és a dir, tant el menú com els elements d’interacció innecessaris durant una aplicació

queden ocults, el que evita errors d'interacció i dóna seguretat als usuaris, així com no interfereix en el sentit de presència a l'afegir elements estranys a la tasca.

Una vegada concretat el tipus de menú, es va començar amb la definició de les característiques específiques d'aquest menú HTC. Es considerà que el menú estiguera format per tres panells o cares ja que aquest nombre era suficient per a simular una aparença 3D, encara que aquest nombre de panells fóra variable, és a dir, el sistema devia fer-se de forma que es poguera canviar fàcilment per disposar de més o menys panells en cas de necessitar-se.

En cada una d'aquestes cares s'ubicarien els botons d'accés a les diferents aplicacions, puguent-se realitzar moviments cap a dalt, baix, dreta i esquerra.. A més a més, abans de començar a descriure més a fons aquest sistema, s'ha de dir que ha sigut desenvolupat gràcies a la eina “*Brainstorm eStudio*” [24][25][26].

## V.2 *Brainstorm eStudio*

“*Brainstorm eStudio*” és un potent programa per desenvolupar entorns gràfics 3D d'alta qualitat renderitzant i controlant la simulació en temps real. Tots els seus fitxers de configuració (excepte alguns de textures i objectes 3D en format binari o de sons) són fitxers ASCII en sintaxi del llenguatge de programació Python [27]. La seua interfàç s'organitza en llistes d'elements i editors. Cadascuna de les llistes representa un concepte i cada element de la llista és una instància al concepte. Per exemple, existeix una llista de llums i cada element d'aquesta llista és una de les llums de la simulació (Fig. 12). A cada llista d'elements li corresponen uns determinats editors, on cadascun executa una acció diferent sobre els objectes de la llista.



Fig. 12 – Llista de llums per defecte

Combinant els elements de les diferents llistes (objectes, materials, textures, accions, binds, maps, dlls, temporitzadors, partícules, geometries, corbes...) és possible obtenir resultats com la desenvoludapa a la interfàç gràfica d'aquest sistema. On una vegada detectats els gestos i enviats al terminal d'usuari que conté la interfàç gràfica, aquestos s'han d'interpretar i associar a diferents elements de les llistes per poder moure els diferents objectes que s'han creat en aquest entorn. Per



exemple per a crear un botó d'un dels panells s'ha tingut que crear un objecte associat a una primitiva amb un material que li dona el color corresponent i a una textura que afegeix el dibuix que té cadascun dels botons. A més a més, s'ha tingut que crear un element de la llista d'unions o binds que associa una determinada acció al gest de selecció d'un botó o click del ratolí per a que aquest responga cada vegada que s'efectue aquest gest. Amb altres combinacions d'elements i ajudant-se del llenguatge de programació Python s'ha implementat el sistema que s'explica a continuació.

### V.3 *FUNCIONAMENT*

En primer lloc, s'ha dissenyat per a que a l'arrancar aquest sistema, aparegui directament la pantalla inicial formada per una sèrie de botons d'accés a les diverses aplicacions o microdemostradors diferents.



Fig. 13 – Pantalla inicial

El número de botons es configurable. A la figura següent es mostren dos exemples de la interfaz amb diferent nombre de botons. A la part esquerra es mostra un panell amb dotze botons i a la part dreta amb sis botons, ja que el disseny d'aquesta permet modificar el nombre de botons a col·locar en cada panell, encara que tal i com s'ha il·lustrat a la imatge anterior, l'aplicació final té sis botons a cadascun dels panells.



Fig. 14 – Nombre de botons configurable

Una vegada l'usuari ja ha iniciat el sistema i té davant la pantalla inicial, pot seleccionar qualsevol microdemostrador mitjançant el gest elegit per a realitzar aquesta acció (seleccionar un botó), com s'observa a la imatge de baix. En esta figura també es pot veure la posició de la càmera i la pantalla que s'ha descrit al subapartat d'interacció gest-interfaç.

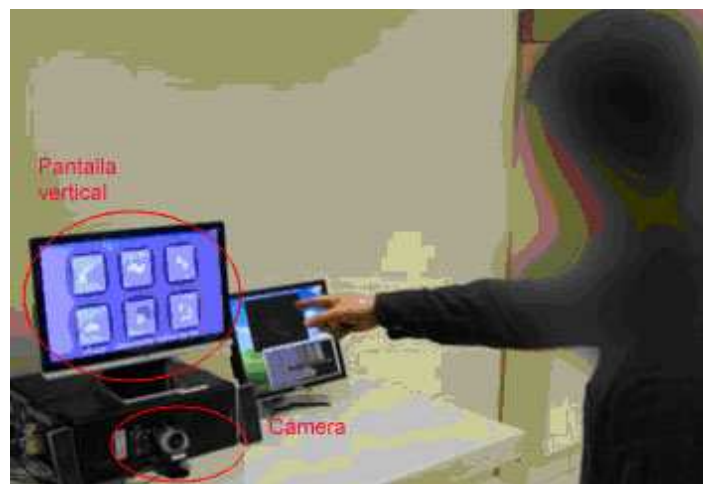


Fig. 15 – Selecció d'un microdemostrador

Si es vol passar a un altre microdemostrador o simplement eixir del sistema, es tornarà a la pantalla inicial dels menús al realitzar-se el gest de desplaçament cap a avall (gest associat a l'acció "Tancar menú" o "Eixida d'un microdemostrador"), visualitzant-se la pantalla principal dels menús per a accedir a les aplicacions. Per tant, les aplicacions no poden ser sensibles a aquest moviment cap a avall de la mà, ja que aquest gest està associat a eixir d'elles.



Fig. 16 – Eixida d'un microdemostrador




D'aquest mode, quan s'inicia el sistema apareix un menú amb una sèrie de botons que permeten accedir als diversos microdemostradors, en cadascun dels quals es pretén provar una sèrie de gestos.

A més a més, des d'aquest menú es pot accedir a un altre panell de dalt com a un de baix, així com a esquerra i dreta, és a dir a qualsevol dels seus quatre panells veïns, on en cadascun hi ha uns altres botons amb altres funcionalitats. No obstant, al considerar que per a l'estructura del diccionari de gestos que es vol, no es necessita anar al panell de dalt i de baix, es va decidir que hi hauria sols tres panells que era suficient per a una aparença 3D, de forma que no existeixen els panells veïns de dalt i baix encara que l'aplicació si que està configurada per poder-se afegir quan es considere oportú. Per tant, quan s'està a un panell sols es pot realitzar el moviment a esquerra i dreta, on es visualitzen altres panells amb altres botons que al prémer-los, mostren microdemostradors diferents.



Fig. 17 – Gest per canviar al panell de l'esquerra

Un resum dels gestos emprats als panells es recull en la següent taula.

Acció en el menú	Gestos utilitzat
<i>Selecció d'un microdemostrador</i>	Selecció d'un botó 
<i>Canvi de panell</i>	Moure's a esquerra i dreta 
<i>Eixida d'un microdemostrador</i>	Moviment de la mà cap avall 

Taula 2 – Accions i gestos per al menú

Per a saber en quina pantalla o panell es troba l'usuari, s'han dibuixat cadascun en diferents colors, criteri que s'ha considerat com a suficient per distingir-los. Els colors elegits han sigut el violeta per al panell inicial que es el que apareix a l'arrancar el sistema, el color verd per al panell de la dreta respecte al panell inicial i el color blau turquesa per al de l'esquerra. També s'ha de destacar que sols podrà hi haure una aplicació oberta en cada moment. A continuació es mostren tots els panells i està indicat amb les fletxes com l'usuari es pot moure tant a dreta com a esquerra des de qualsevol panell.



Fig. 18 – Distinció pel color del panell

En quant a l'aspecte i dimensions de la interfàç, aquesta deurà ocupar tota la pantalla, tal i com s'observa a la part dreta de la figura de baix, però podran també ser configurables les seues dimensions i relació d'aspecte.

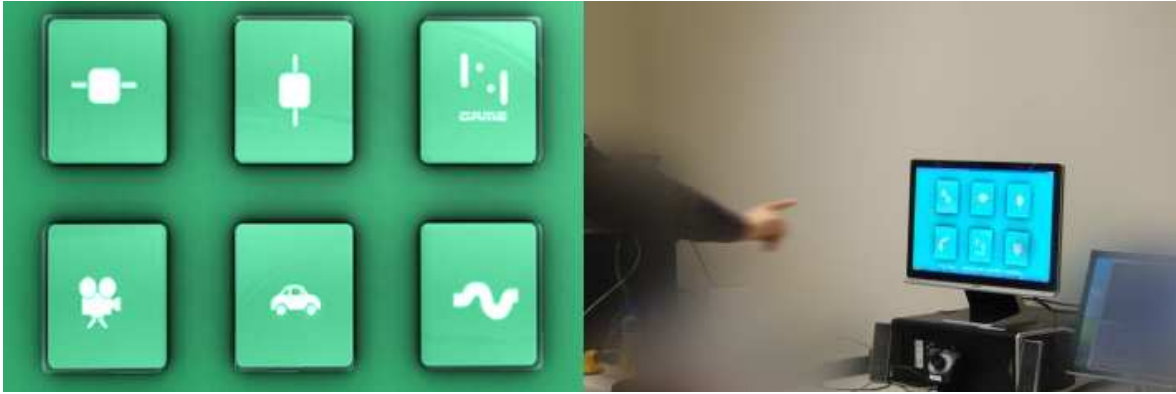


Fig. 19 – Pantalla completa per 1 panell

#### V.4 EVOLUCIÓ

Durant la seua implementació, es féren diversos canvis per a intentar aconseguir una interfàç gràfica tridimensional que s'ajustés al màxim possible als gestos elegits i aquestos es pogueren avaluar adequadament. A més a més amb aquesta interfàç gràfica es pretén donar idees per a futurs serveis que poden ser desenvolupats al sistema global de forma que a part d'incorporar el sistema de videoconferència, un photolog i el joc del tres en ratlla, puga contenir altres aplicacions com pot ser qualsevol de la desenvolupada en els microdemostradors.

Entre els diferents canvis que s'han tingut que realitzar durant el seu desenvolupament, un d'ells va ser l'aspecte final dels botons del menú inicial. Tot seguit es mostra, una d'aquestes versions que finalment no ha sigut la definitiva. Com també s'observa el dibuix de cadascun dels botons, que pretenen representar el contingut de cada microdemostrador.

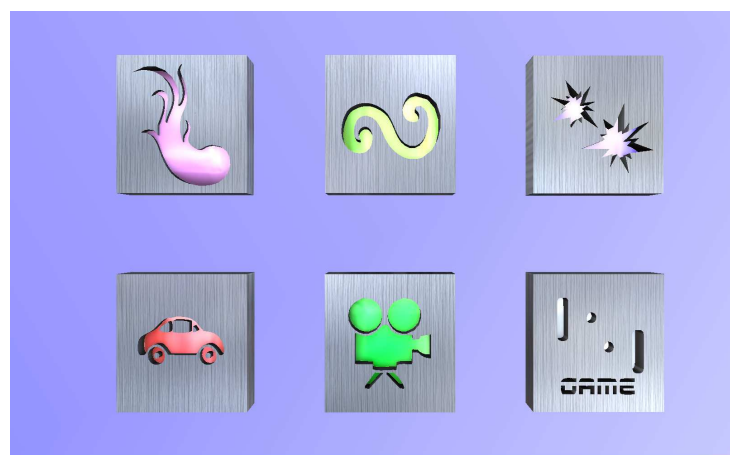


Fig. 20 – Proves desenvolupament interfàç

## V.5 INDEPENDÈNCIA

Es tracta d'una aplicació independent ja que funciona sense el sistema de reconeixement gestual. D'aquesta manera, el teclat i el ratolí simulen els esdeveniments que posteriorment van a ser associats als gestos reconeguts.

Per exemple, per a canviar de panell s'utilitzen les tecles de les fletxes a esquerra i dreta del teclat que reproduïxen els gestos de moviment a esquerra i dreta o bé el gest associat a seleccionar un ítem per accedir a qualsevol dels microdemostradors es realitza amb un click de ratolí. Per últim per a eixir dels microdemostradors s'utilitza la fletxa cap a avall del teclat associada al gest de "Tancar menú" que consisteix en un moviment de la mà cap a avall.

## V.6 MICRODEMOSTRADORS

Una vegada es seleccionat algun botó d'un dels panells, s'accedeix al microdemostrador en qüestió. Després de totes les idees proposades, les aplicacions finalment desenvolupades són les que s'expliquen a continuació.

Alguns microdemostradors es van desestimar per ser redundants com ocorria amb el denominat "Visualitzador d'objectes" on apareixen una sèrie d'objectes que anaven variant cada cert temps i l'usuari podia agafar-los i girar-los. No obstant, aquesta mateixa idea queda arrellegada al microdemostrador del cotxe que s'ha implementat. Avall, es mostra un llibre que era un dels diferents objectes que es mostraven a aquest microdemostrador.





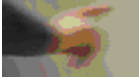





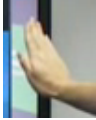
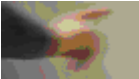





Fig. 21 – Microdemostrador descartat

Cadascun dels microdemostradors següents s'utilitza per a avaluar els diferents gestos que s'havien obtingut després de l'estudi de la interfaç gestual que s'ha explicat anteriorment.

En la "taula 3" es mostren les diferents aplicacions desenvolupades junt amb els gestos principals que s'utilitzen en cadascuna i a les que s'accedeix seleccionant el botó corresponent al

microdemostrador en qüestió i com es veu, s'intenta reduir el nombre de gestos el màxim possible per a que l'aplicació siga més fàcil d'utilitzar pels usuaris.

Aplicació	Gestos utilitzables i reconeguts
<i>Estela</i>	Selecció i seguiment de la mà o dibuix d'una trajectòria 
<i>Pissarra</i>	Selecció del punt inicial d'escriptura i seguiment de la línia a traçar 
<i>Focs artificials</i>	Selecció de la situació de la carcassa 
<i>Audi</i>	<p>Agafar i girar un objecte en qualsevol direcció </p> <p>Seleccionar per a canvi de color </p>
<i>Vídeos</i>	<p>Girar a esquerra o dreta per elegir un vídeo  ó </p> <p>Selecció d'un vídeo o fer zoom d'un vídeo  ó </p> <p>Finalitzar el zoom d'un vídeo  ó </p>
<i>Joc</i>	<p>Selecció </p> <p>Pillar paleta i seguiment per a colpejar la pilota </p>
<i>Slider horitzontal i vertical</i>	<p>Selecció i arrossegament fi </p> <p>Agafar un objecte, moure'l i soltar-lo </p>

Taula 3 - Relació microdemostrador-gestos

Seguidament s'expliquen cadascun dels microdemostradors de la taula anterior amb més detall.

### V.6.1 Estela

El primer microdemostrador implementat consisteix en una estela, tal i com apareix en la imatge de baix, que va seguint la mà de l'usuari en qüestió, la qual es pot deformar i jugar amb ella sense ningun altre fi, mentre va deixant rastre de la seua trajectòria. Aquest microdemostrador es va realitzar basant-se en la idea proposada de moure les partícules com les dels salva pantalles.



Fig. 22 – Estela

El gest emprat per l'usuari tal com es veu a la figura següent, és el de selecció d'un punt i amb eixe mateix gest anant traçant la trajectòria que es vullga. L'aplicació també treballa seguint la trajectòria marcada pel ratolí.



Fig. 23 – Usuari amb l'estela



### V.6.2 *Pissarra*

Aquest altre microdemostrador permet a l'usuari dibuixar el que es vulgui mentre estiga seleccionat, que és el gest avaluat. Es reconeix el punt inicial de l'escriptura quan es fa un click amb el ratolí o es realitza el gest associat de selecció d'un punt que és d'on es comença a dibuixar, després es traça la línia seguint la trajectòria de la mà o bé l'arrossegament del ratolí. Sorgeix a partir de l'aplicació proposada anomenada "Canves amb ferramentes".

Si transcorri cert temps sense seleccionar res, desapareix el que s'havia dibuixat fins eixe moment, netejant així la pissarra, deixant-la buida tal i com estava inicialment.

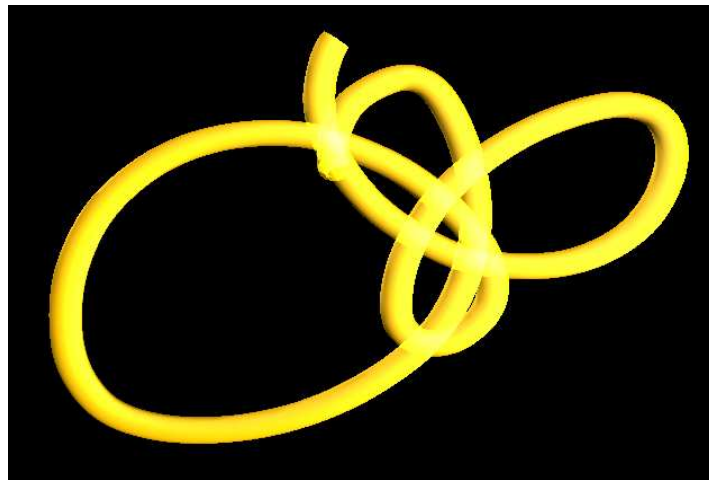


Fig. 24 – Pissarra virtual

Un dels canvis possibles a realitzar en aquest microdemostrador és el color del traç de la línia a dibuixar. Tal com es veu en la imatge anterior, en un principi era de color adorat i finalment, el aspecte és verd, com apareix en la figura següent.



Fig. 25 – Diferent color pissarra

Finalment es pot veure un dels usuaris intentant dibuixar alguna cosa al sistema. Com es pot apreciar a la imatge, és més senzill dibuixar el que es vol amb el ratolí que per mitjà de gestos, ja que s'ha de moure la mà lentament evitant els moviments bruscos però també s'aconsegueix el que es vol si es té paciència i delicadesa.



Fig. 26 – Usuari amb la pissarra

### V.6.3 Focs artificials

Tal i com el seu nom indica, es tracta d'un microdemostrador on s'il·lustren les carcasses de focs artificials amb so associat simulant l'esclafit de cada carcassa. Per a fer-ho, l'usuari deu seleccionar durant un instant el/els punt/s on vulgui situar eixa/es carcassa/es i ja el programa s'encarrega de traçar la trajectòria d'aquestos a través de corbes associades. Aquesta selecció de la/es posició/ns puntual/s de la/es carcassa/es es pot fer amb el gest o mitjançant el click del ratolí.

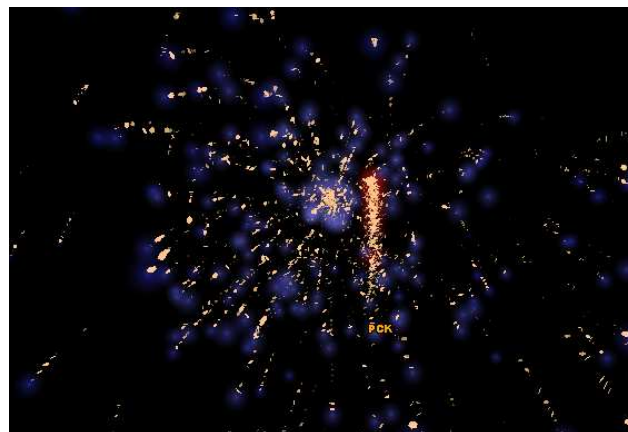


Fig. 27 – Fons negre

Finalment, per a representar una escena més real, el color del fons elegit ha sigut el blau, intentant reflectir el cel en lloc de negre com s'observa a la imatge anterior. A més a més, en la part dreta de la següent figura es pot veure un usuari seleccionant el punt d'una de les carcasses mitjançant el gest corresponent.



Fig. 28 – Fons blau

#### V.6.4 Cotxe

Per a poder avaluar els gestos de pillar i girar un objecte, s'ha implementat aquest microdemostrador. En ell es pot agafar l'objecte, en aquest cas el cotxe en qüestió on se li ha intentat ressaltar la seua geometria 3D amb una textura real. L'usuari pot pillar-lo, girar-lo i canviar-li el color.

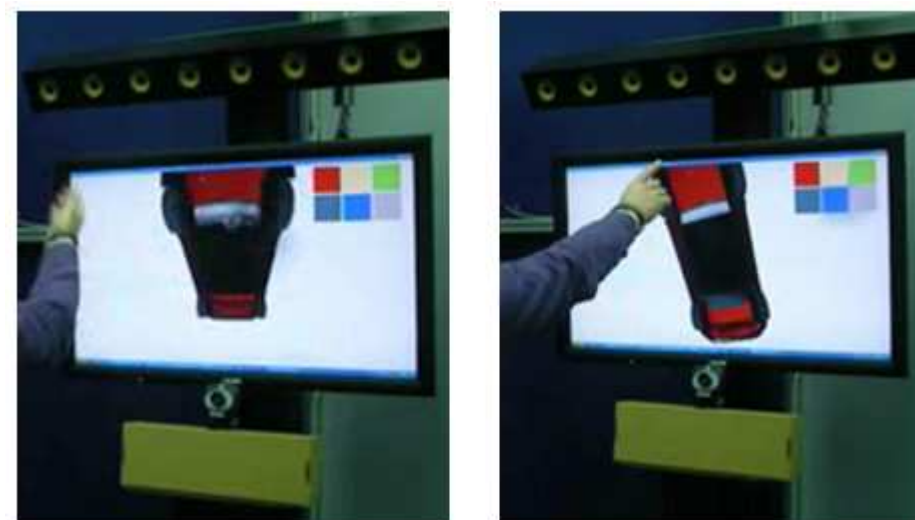


Fig. 29 – Gest de moure el cotxe / Gest de selecció de color

Com es pot observar a la figura anterior, a la part esquerra de la imatge si es fixes en la posició de la mà de l'usuari, es pot apreciar com està movent el cotxe cap a dalt. En canvi, en la part dreta

està seleccionant un color per a canviar-li'l. En aquest cas, elegeix el color groc i en la següent figura es veu com una vegada ha variat el color, l'usuari continua movent el cotxe, esta vegada cap avall com s'aprecia amb el gest.



Fig. 30 – Usuari movent cotxe cap avall

Al començament, es va col·locar una paleta amb un major nombre de colors, però analitzant-lo es va arribar a la conclusió que era millor que cada quadre de cada color tinguera una major àrea per a disposar d'una major àrea d'interacció del gest de l'usuari amb la pantalla encara que així es disposara de menys colors.



Fig. 31 – Major nombre de colors de la paleta

Per tant, per qüestions d'usabilitat, l'aspecte final del microdemostrador és el següent, on sols hi ha sis colors possibles, el que fa que l'usuari tingui més àrea per a interaccionar amb el gest que desitja triar un determinat color per al seu cotxe.



Fig. 32 – Aspecte final

Aquest microdemostrador a l'igual que la resta, apart de ser emprat amb gestos es pot usar amb el teclat i el ratolí, on s'elegeix el color que es vol amb el click del ratolí i es mou el cotxe amb les tecles de les fletxes cap a dalt, baix, dreta i esquerra.

### V.6.5 Vídeos

Després de plantejar-se com a idea inicial, la utilitat de què l'usuari interaccionés amb una fotografia o vídeo, s'ha implementat el microdemostrador de vídeos, on apareixen diversos vídeos en una espècie de cercle com s'il·lustra en la imatge de baix.



Fig. 33 – Vídeos

L'usuari pot arribar al vídeo que li interessi girant amb un moviment a esquerra o dreta amb els gestos corresponents tal com es veu en la imatge de baix. Hi ha dos possibilitats per a realitzar l'acció de moure's a esquerra i dreta. En la figura següent l'usuari està realitzant el moviment a esquerra de dues formes diferents.

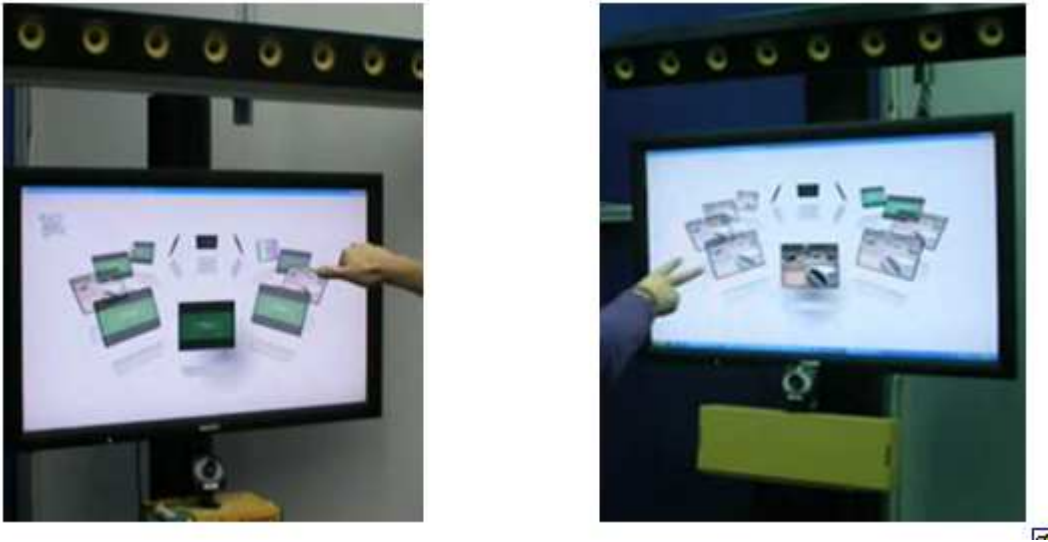


Fig. 34 – Moure's a esquerra (ambdues opcions)

També es pot fer un zoom in/out del vídeo que es volga quan estiga davant d'ell. Aquest zoom és el que es mostra seguidament, on un sols vídeo ocupa tota la pantalla.



Fig. 35 – Zoom in

A continuació es mostren els gestos dels usuaris tant al moment de selecció del vídeo per fer-li un zoom com de tornar al cercle amb tots els vídeos.

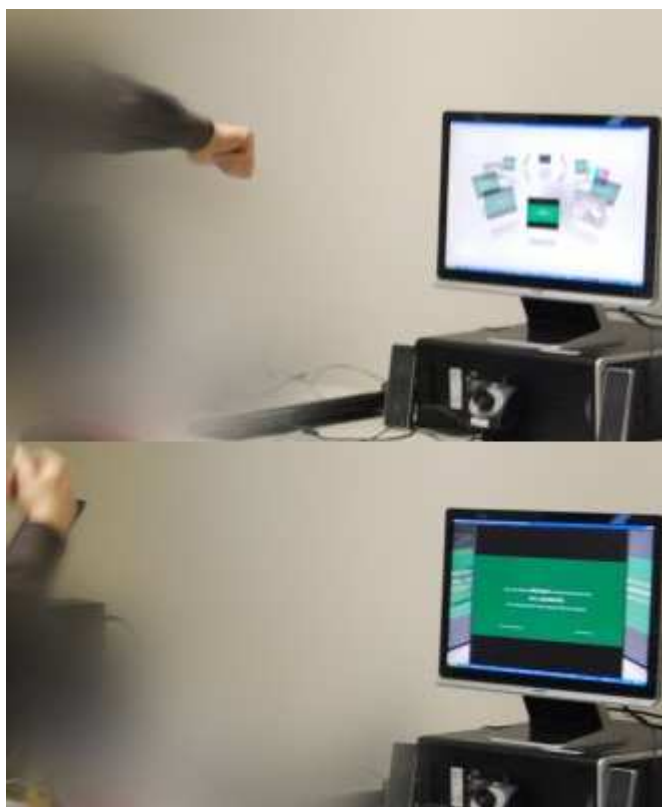


Fig. 36 – Gest Zoom in/out

Una alternativa per als gestos de selecció del vídeo (zoom in) i per al de tornar a la roda de tots els vídeos (zoom out) són els que es mostren a continuació.



Fig. 37 – Alternativa gest Zoom in/out

Tant com ocorre als altres microdemostradors, es pot utilitzar mitjançant la interfàç gestual o amb les tecles de la fletxa a dreta i esquerra per als moviments a dreta i esquerra respectivament com és lògic i amb un click de ratolí per a fer zoom in i zoom out.

### V.6.6 *Joc*

S'ha considerat interessant incloure també un microdemostrador que consisteix en un joc ja que el prototip inicial del sistema global ja en tenia un, el del tres en ratlla. En aquest cas, és el típic joc de "ping-pong", on es tracta de seleccionar la raqueta, agafar-la i seguir la pilota movent la raqueta per a pegar-li a la pilota, tal qual es veu en la imatge següent.

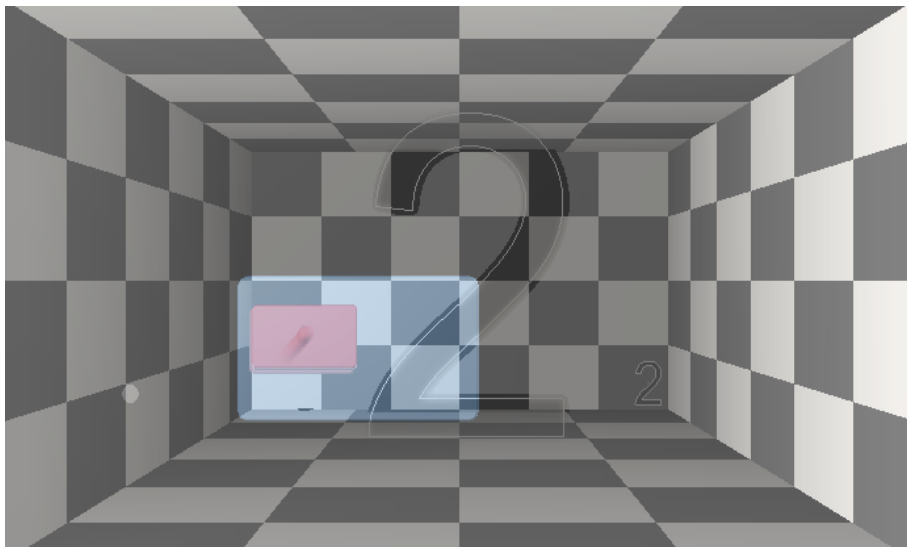


Fig. 38 – Joc

Al fons de l'escenari es marca la puntuació obtinguda en cada partida, reiniciant-se cada vegada que l'usuari no li pega a la pilota, tal i com s'observa en la figura següent. A més a més, en el cantó inferior dret del fons es manté la màxima puntuació obtesa de totes les partides que s'han jugat, de forma que l'usuari quan juga una nova partida sap quina és la puntuació que ha de superar per a fixar un nou record. A la figura de baix, es veu un exemple, on l'usuari té com a puntuació màxima 5 i en la partida actual que està jugant ha obtés per ara una puntuació de 3.



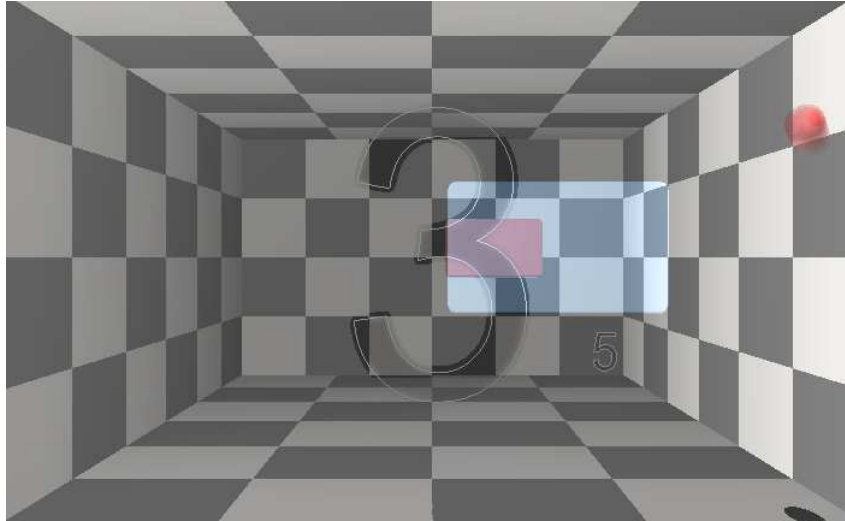


Fig. 39 – Puntuació màxima

Per últim, es veu com un usuari executa aquest microdemostrador amb les dos possibilitats que existeixen: el gest seleccionar (el de la part esquerra de la imatge), sent l'únic que s'usa a tot el microdemostrador, ja que es selecciona la raqueta i es mou cap a la direcció que es vullga per donar-li a la pilota o bé, el gest d'agafar un objecte (en aquest cas la raqueta) i moure-la o traçar una trajectòria en qualsevol direcció (part dreta de la figura) sent aquest gest el mateix. Per tant, en tot el joc sols s'usa un únic gest.

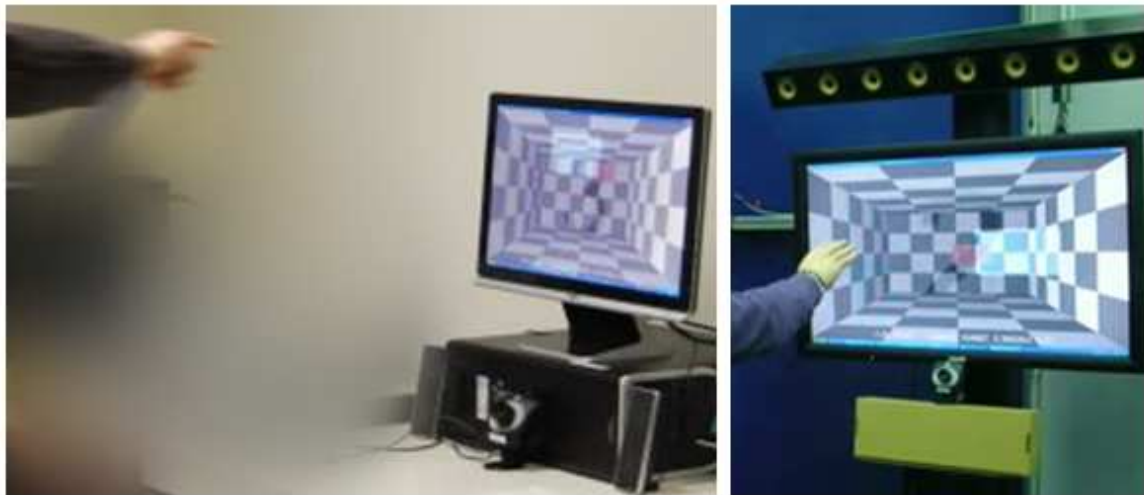


Fig. 40 – Possibilitats Gest-Joc

Si es vol usar sense gestos, simplement es tracta de moure el ratolí que simula el moviment de la raqueta amb la que s'ha de colpejar a la pilota.

### V.6.7 *Slider horitzontal i vertical*

Per acabar amb els microdemostradors, s'han implementat dos microdemostradors on es poden avaluar els gestos de seleccionar i agafar així com el d'un arrossegament fi tant horitzontal com verticalment a l'utilitzar els sliders implementats que van indicant el percentatge de moviment numèricament. Com també ocorre als microdemostradors anteriors, aquests apart d'usar-se mitjançant la interfàcia gestual, també es poden moure amb el ratolí.

A la figura següent es mostren el slider horitzontal i vertical respectivament en diferents posicions, així com es veuen els seus límits en la pantalla.

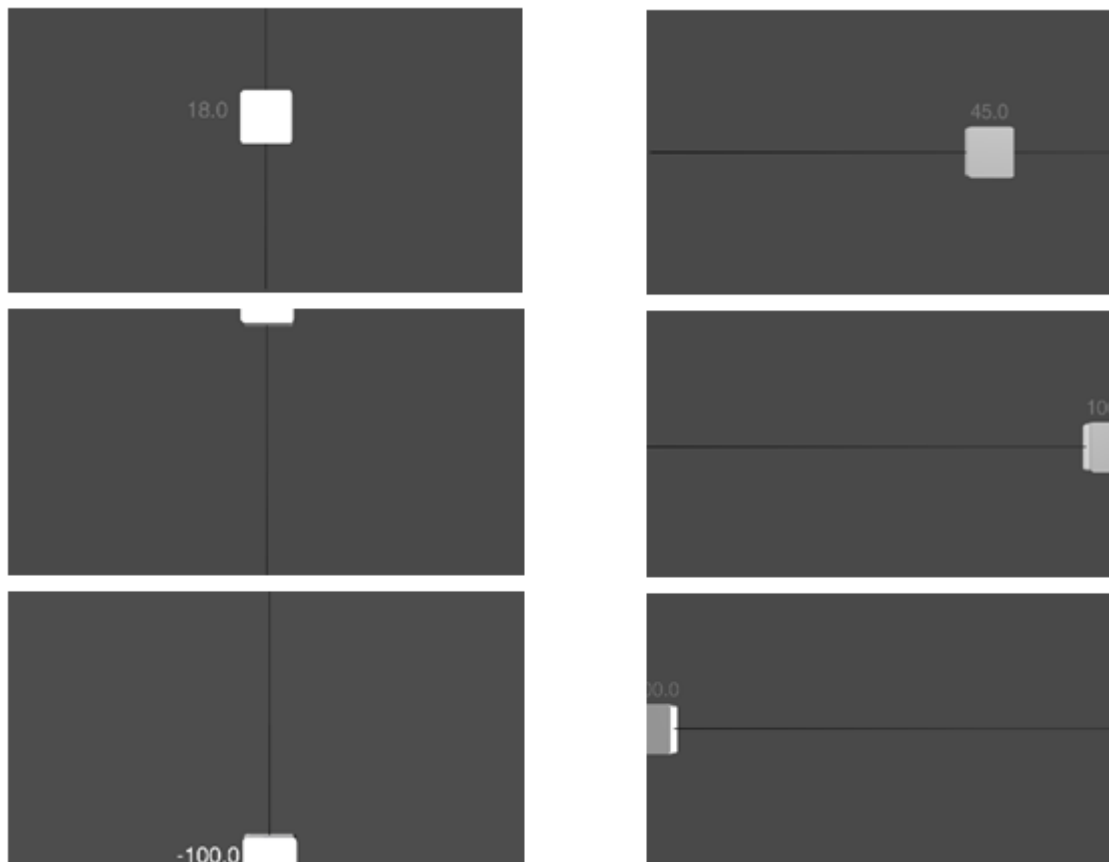


Fig. 41 – Slider vertical i horitzontal

Finalment, la figura de baix reflexa els gestos emprats a aquests dos microdemostradors. A la primera fila està la primera possibilitat que es utilitzant el gest de selecció i amb aquest mateix gest traçar la línia vertical o horitzontal depenent del slider en el que estiga l'usuari i per acabar obrir la mà cap avall. La segona possibilitat de combinació de gestos a emprar en aquests microdemostradors s'observa a la segona i tercera fila per al microdemostrador del slider horitzontal i vertical respectivament. En aquesta segona possibilitat, els gestos són el d'agafar un

objecte (en aquest cas, el slider) i soltar-lo que és amb la mà oberta i els dits estirats i el de moure-lo amb el puny tancat.

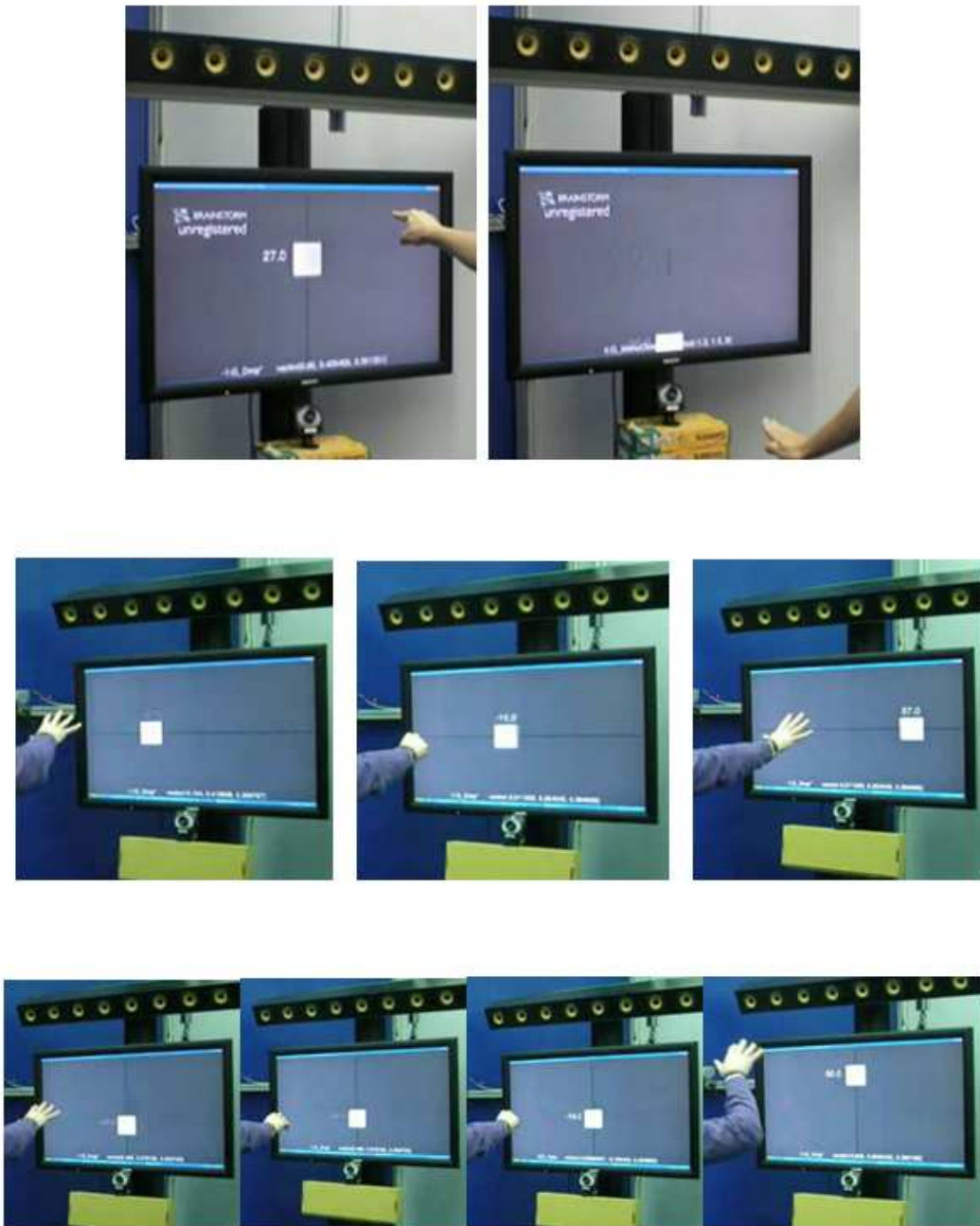


Fig. 42 – Seqüència de gestos sliders

### V.7 POSSIBLES MILLORES

Algunes millores possibles sobre la interfaç gràfica tipus menú HTC realitzada serien les citades a continuació.

La primera consisteix en col·locar en la part inferior de la pantalla una ruta de situació on es puga decidir si l'usuari pot interaccionar o no amb ella ja que si els usuaris posseeixen nocions informàtiques molt baixes, inicialment es senten perduts al canviar de dreta a esquerra malgrat l'existència del panells de diferents colors. No saben exactament on estan ni què hi ha a la dreta i a l'esquerra encara que se les canvie el color del panell. L'aprenentatge en alguns casos és ràpid. No obstant, memoritzar el maneig depèn de la freqüència d'ús i passats uns dies sense haver-lo usat de nou, es repeteix el mateix i tornen a desorientar-se els usuaris, per això una possible millora seria situar una barra on s'indica en cada moment en quin panell està l'usuari i quins són els panells veïns.

Una altra millora possible a realitzar és afegir una sèrie de indicacions que apareguen a la pantalla com poden ser "Per favor selecciona l'objecte" o "Per favor gira l'objecte a la dreta", per a què a nivell d'usabilitat l'usuari no estiga obligat a aprendre un llenguatge de signes així com si hi ha gran quantitat de gestos, amb aquestes indicacions, l'usuari entén què significa cada gest a priori.

També és aconsellable que hi haja un menú de configuració, permetent que l'usuari cree un perfil personalitzat, que es guardi i ajusti a les seues necessitats, permetent també establir una relació gest-tipus d'interacció predeterminada basada en les proves d'usabilitat inicials. A més a més, en aquesta configuració inicial es pot considerar si l'usuari és esquerrà o dextre i segons l'opció escollida per cada usuari, la configuració de gestos treballa en mode "Dextre" (entenen aquest com el mode predeterminat ja que la majoria de la població ho és) o "Esquerrà". (que seria un mode tipus mirall del mode anterior).

## VI. CONCLUSIÓ

Una vegada efectuat tot l'estudi previ sobre els gestos amb persones de perfil variat, s'ha tractat d'esbrinar quins són els gestos més naturals que s'usen per a moure o pillar un objecte, apuntar o seleccionar un punt, etc. Sense tindre davant cap terminal, és a dir, evitant la tecnologia. De forma que únicament dient-los que feren certes accions com per exemple com seguirien el recorregut d'una formiga o com passarien una fulla d'un llibre, es pot conèixer la seua resposta gestual.

Aquest experiment permet fer una estimació dels gestos amb els quals la gent es maneja habitualment i a partir d'ells s'implementà un sistema de reconeixement de gestos el més senzill possible i es desenvoluparen les aplicacions on es podien aplicar simplificant-les al màxim.

El sistema de reconeixement de gestos consisteix en la captura de la silueta de la mà o les mans, tenint així la màscara d'ella o d'elles. A continuació s'extreia el vector de característiques per a obtenir les dades que modelen la o les mans. Seguidament, es reconeix el gest dinàmic i per últim, una vegada ja reconegut el gest es comunica amb la part de la interfàç gràfica per UDP.

La part de la interfàç gràfica es basa en un sistema, que tal i com s'ha dit, està format per tres panells, encara que aquest nombre és configurable i l'usuari reconeix en quin panell es troba pel color que té cadascun. Des de cadascun d'ells, l'usuari pot moure's a dreta o esquerra, ja que el moviment cap a avall queda reservat per a eixir de cadascun dels microdemostradors.

D'aquesta forma, es sotmet posteriorment als usuaris a un altre experiment, aquesta vegada davant de dites aplicacions per a esbrinar si s'han aconseguit gestos més eficients i per tant si s'han millorat respecte al que es tenia inicialment. És a dir, si els usuaris es senten més familiars amb les aplicacions i és més accessible a tot el món.

Com a conclusió, dir que els diversos microdemostradors han aconseguit obtenir gestos més propers a la població en general. Un exemple clar està reflectit al microdemostrador dels vídeos, on l'usuari passa els diferents vídeos tal qual ho feia amb un llibre real i selecciona el que més l'interessa, és en eixe moment quan es realitza un zoom.

A més a més, alguns dels microdemostradors implementats per aquesta avaluació de gestos com és el dels focs artificials o el del cotxe, han sigut posteriorment incorporats com a serveis al sistema de videoconferència global, el qual inicialment sols tenia el joc del tres en ratlla, el photolog i el propi servei de videoconferència. Per tant, els resultants han sigut favorables i s'ha complert l'objectiu marcat d'aconseguir millorar aquest sistema global nou de videoconferència.

Per últim, afegir que d'aquesta tesina han sorgit dos publicacions que estan en procés de revisió. Una s'anomena "A 3D real-time immersive rendering system for videoconferencing", on es descriu aquest sistema de videoconferència 3D de forma general, explicant en cadascun dels apartats de l'article cadascun dels mòduls que componen el sistema, des de la captura de l'escena, la seua codificació i transmissió per la xarxa de comunicacions fins el seu sistema de render i reconstrucció presentada al terminal d'usuari remot, emfatitzant amb la interfàç natural que proporciona major presència.

La segona publicació anomenada "A comparative study of Marching Cubes techniques for 3D real-time rendering of scenes in videoconferencing systems" realitza un estudi comparatiu entre les diferents tècniques i variants de l'algoritme usat al mòdul de reconstrucció de models 3D en temps real obtenint la màxima eficiència i flexibilitat, així com el procés de render influeix a l'algoritme de reconstrucció que es presentat a la interfàç d'usuari.

## **AGRADEIXEMENTS**

En primer lloc, desitge reconèixer la gran dedicació de la meua Directora de Tesina, Valery, pel seu constant treball que ha fet, qui inicialment va ser professora meua a la carrera universitària i ara apart de directora és companya a l' institut d'investigació on treballem (I3BH). També agraeix al personal del departament LabHuman ("*Human centered Technology*") en concret dins d'aquest

institut que és on estic treballant pel seu recolzament diari en aquesta investigació i implementació, així com per tots els coneixements que m'han anant ensenyant.

Finalment, citar el projecte d'investigació anomenat VISION liderat per Telefònica I+D que està estretament relacionat amb aquesta memòria.

**BIBLIOGRAFIA**

- [1] Semantic video analysis for adaptive content delivery and automatic description, IEEE Transactions on Circuits and Systems for Video Technology, vol. 15(10) of 1200-1209, A. Cavallaro, O. Steiger, and T. Ebrahimi, 2005.
- [2] Detecting moving objects, ghosts and shadows in video streams, IEEE Trans. on Patt. Anal. and Machine Intell., vol. 25, pp. 1337\_1342, R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, 2003
- [3] Video-Based Rendering, Editorial, Sales, and Customer Service Office, A.K. Peters, Marcus A. Magnor, Max-Planck-Institut für Informatik Saarbrücken, Germany, Wellesley, Massachusetts, 2005
- [4] Scientific Visualization from Linköping University, Sweden. Focus on research in 3D video coding technology and GPU based image processing/rendering [www.mpi-inf.mpg.de/~gziegler/](http://www.mpi-inf.mpg.de/~gziegler/)
- [5] GPU-Friendly Marching Cubes <http://appsrv.cse.cuhk.edu.hk/~gywang/papers/mc/mc.html>
- [6] NVision-Never underestimate your graphics hardware, <http://nvision.sourceforge.net/>
- [7] Designing the user interface for multimodal speech and pen-based gesture applications: state-of-the-art systems and future research directions, Sharon Oviatt, Phil Cohen, Lizhong Wu, John Vergo, Lisbeth Duncan, Bernhard Suhm, Josh Bers, Thomas Holzman, Terry Winograd, James Landay, Jim Larson, David Ferro, Human-Computer Interaction, Volume 15 , Issue 4, Pages 263-322, 2000
- [8] Put-That-There: Voice and Gesture at the Graphics Interface, Richard A. Bolt, Architecture Machine Group, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, International Conference on Computer Graphics and Interactive Techniques, Proceedings of the 7th annual conference on Computer graphics and interactive techniques , Seattle, Washington, United State, pages: 262 - 270, 1980
- [9] Speech/gesture interface to a visual-computing environment, Sharma, R.; Zeller, M.; Pavlovic, V.I.; Huang, T.S.; Lo, Z.; Chu, S.; Zhao, Y.; Phillips, J.C.; Schulten, K.; Computer Graphics and Applications, IEEE, Volume 20, Issue 2, March-April 2000 Page(s):29 - 37
- [10] Toward Multimodal Human-Computer Interface, Rajeev Sharme, Vladimir I. Pavlovic, Thomas S. Huang, IEEE
- [11] Ámbitos aplicados de Psicología de la Memoria, Madrid: Alianza, cap 2, pp 47-65, Sánchez Cábaco A. et al , 2001
- [12] Recognizing: the judgment of previous occurrence. Psychological Review, 87 (3), 252-271, Mandler, G. 1980
- [13] Your face looks familiar but I can't re-member your name: A review of dual process theory. Em E. William, E. Hockley & E. S. Lewandowsky (Eds.). Relating theory and data: essays on human memory in honour of Bennet B. Murdock (pp. 207-225). Hillsdale: Erlbaum, Mandler, G. , 1991
- [14] Activation makes words more accessible, but not necessarily more retrievable. Journal of Verbal Learning and Verbal Behavior, 23 (5), 553-568, Graf, P. & Mandler, G., 1984.
- [15] La construcción de un interfaz multimodal para interacciones hombre/maquina, Departamento de I+D, NLU de ISS, Ricard Pérez, Procesamiento del lenguaje natural, boletín nº10, 1991
- [16] Two-handed gesture in multi-modal natural dialog, Richard A. Bolt, Edward Herranz, Symposium on User Interface Software and Technology, Proceedings of the 5th annual ACM symposium on User interface software and technology, Monterey, California, United States, Pages: 7 - 14, 1992

- [17] Finger-Pointer: Pointing interface by image processing, Masaaki Fukimoto, Yasuhito Suenaga, Kenji Mase, NTT Human Interface Laboratories, Yokosuka, Japan, *Comput. & Graphics*, Vol. 18, nº 5, pp. 633-642, 1994
- [18] E.8.2.7. Evaluación de la interfaz de usuario. Usabilidad en el lenguaje gestual: línea de base mediante el test de usuarios, CENIT VISION, Diana Castilla López, Ignacio Miralles Tena, Juana M. Bretón López, Yolanda Vizcaíno Dragón, M.Ángeles Pérez Ara, M. Vicenta Navarro Haro, UJI, 2007
- [19] E6.5.2b. Sistema de reconocimiento de objetos y gestos, CENIT VISION: Comunicaciones de Vídeo de Nueva Generación, Tomasz Adamek, Paulo Villegas, Christian Ferran, Montse Pardàs, Jose M. Martínez, Jesús Bescós, J.C. San Miguel, Javier Molina, Marcos Escudero, TID, UAM, UPC, 2007
- [20] Wikipedia, La enciclopedia libre, <http://es.wikipedia.org>
- [21] Menú HTC, Youtube, [youtube.com/watch?v=CoaXA9un7YE&feature=related](http://youtube.com/watch?v=CoaXA9un7YE&feature=related), [youtube.com/watch?v=Gj8PMcvYTZo](http://youtube.com/watch?v=Gj8PMcvYTZo)
- [22] Interfaz gráfica HTC cero, <http://alt1040.com/2009/06/htc-hero-con-la-nueva-interfaz-grafica-sense>
- [23] Piclens, [www.piclens.com](http://www.piclens.com)
- [24] Introducción a la programación con Python, Departamento de Lenguajes y Sistemas Informáticos, Universitat Jaume I, Andrés Marzal, Isabel Gracia, 2003
- [25] Brainstorm eStudio Manual v11.2, The definitive broadcast 3D graphics suite, Brainstorm team, January 2008
- [26] Brainstorm eStudio Tutorials v11.2, The definitive broadcast 3D graphics suite, Brainstorm team, January 2008
- [27] Python, Lenguaje de programación orientado a objetos interpretado, <http://www.python.org/>



### **Annexes**

Es deuen incloure les publicacions que el autor tinga relacionades amb la tesina i els justificants d'acceptació en el caso que no hagen sigut publicats encara els articles/congressos