



Departamento de Ingeniería de Sistemas y Automática
Universidad Politécnica de Valencia



TESIS DOCTORAL

**CARACTERIZACIÓN Y
OPTIMIZACIÓN DEL PROCESO
DE CALIBRADO DE CÁMARAS
BASADO EN PLANTILLA
BI DIMENSIONAL**

Realizada por:

Carlos Ricolfe Viala

Dirigida por:

Antonio José Sánchez Salmerón

25 de Octubre de 2006

Grupo de Robótica

INSTITUTO DE AUTOMÁTICA E INFORMÁTICA INDUSTRIAL

RESUMEN

El procedimiento de calibrado de una cámara acaba siendo un paso necesario para la obtención de información 3D del entorno a partir de imágenes 2D del mismo. Existen diferentes técnicas las cuales se basan en fotogrametría o auto calibración. Los métodos basados en fotogrametría capturan una imagen de una escena conocida compuesta por una plantilla tridimensional, bidimensional o unidimensional. Las técnicas de auto calibración se basan en la obtención de varias imágenes de una misma escena aprovechando la rigidez de la misma para establecer restricciones que permitan realizar la calibración de la cámara.

Como resultado de la calibración de la cámara se obtienen los parámetros intrínsecos y extrínsecos de la misma. La obtención de todos los parámetros de la cámara mediante calibración, no es exacta debido a imprecisiones que perturban el proceso. Estas imprecisiones surgen por imperfecciones constructivas de las lentes, desalineamientos mecánicos de las mismas o del sensor, y también por el hecho de procesar la imagen y obtener posiciones de los puntos dentro de ellas. Los resultados que se obtienen dependen tanto de la plantilla de calibración utilizada, como del algoritmo para resolverla, así como del tratamiento previo que se les pueda realizar a los datos. En la literatura existen multitud de técnicas de calibrado de cámaras y de preprocesado de datos.

Desde el punto de vista que es imposible obtener un valor exacto para cada uno de los parámetros de la cámara, resulta interesante obtener un intervalo en el cual esté contenido. Estas incertidumbres asociadas a los parámetros de la cámara permitirán mejorar los procedimientos de reconstrucción 3D y de medida que se realicen a partir de los mismos. También, a la hora de calibrar una cámara surgen preguntas acerca del algoritmo de calibración ó plantilla a utilizar, número de puntos a colocar en la plantilla de calibración, número de imágenes a tomar de la misma, así como las posiciones y orientaciones desde las que tomar las imágenes.

Esta tesis pretende dar respuesta a todas estas cuestiones. En primer lugar se adopta el método de calibración que mejor resultados obtiene basándose en los métodos existentes tanto de calibración como de preprocesado de datos. Se plantea un preprocesado de datos que permite disminuir el ruido de las medidas y asociar una incertidumbre a cada una de ellas de forma que después se pueda utilizar para el cálculo de los parámetros. Como resultado se obtienen los parámetros de la

cámara con su incertidumbre asociada. También se caracteriza desde un punto de vista estadístico el proceso de cálculo para determinar como afecta la incertidumbre de los parámetros calculados en función del número de imágenes o número de puntos en la plantilla. Esto permite posteriormente definir valores acerca del número de imágenes a tomar de la plantilla o el número de puntos a colocar en la misma para obtener unos valores razonables de incertidumbre. Finalmente, se definen unas posiciones y orientaciones para tomar las imágenes de la plantilla de forma que el procedimiento de calibración se realice de forma óptima. Estas posiciones y orientaciones no dependen en ningún momento de las características de la cámara.

RESUM

El procediment de calibrat d'una càmera acaba sent un pas necessari per a l'obtenció d'informació 3D de l'entorn a partir d'imatges 2D del mateix. Hi ha diferents tècniques les quals es basen en fotogrametria o auto calibratge. Els mètodes basats en fotogrametria capturen una imatge d'una escena coneguda composta per una plantilla tridimensional, bidimensional o unidimensional. Les tècniques d'auto calibratge es basen en l'obtenció de diverses imatges d'una mateixa escena aprofitant la rigidesa de la mateixa per a establir restriccions que permeten realitzar la calibratge de la càmera.

Com resultat del calibratge de la càmera s'obtenen els paràmetres intrínsecs i extrínsecs de la mateixa. L'obtenció de tots els paràmetres de la càmera per mitjà del calibratge, no és exacta a causa d'imprecisions que pertorben el procés. Estes imprecisions sorgixen per imperfeccions constructives de les lents, desalineaments mecànics de les mateixes o del sensor, i també pel fet de processar la imatge i obtenir posicions dels punts dins d'elles. Els resultats que s'obtenen depenen tant de la plantilla de calibratge utilitzada, com de l'algoritme per a resoldre-la, així com del tractament previ que se'ls puga realitzar a les dades. En la literatura hi ha multitud de tècniques de calibratge de càmeres i de preprocessat de dades.

Des del punt de vista que és impossible obtenir un valor exacte per a cada un dels paràmetres de la càmera, resulta interessant obtenir un interval en el qual estiga contingut. Estes incerteses associades als paràmetres de la càmera permetran millorar els procediments de reconstrucció 3D i de mesura que es realitzen a partir dels mateixos. També, a l'hora de calibrar una càmera sorgixen preguntes sobre l'algoritme de calibratge o plantilla a utilitzar, nombre de punts a col·locar en la plantilla de calibratge, nombre d'imatges a prendre de la mateixa, així com les posicions i orientacions des de les que prendre les imatges.

Esta tesi pretén donar resposta a totes estes qüestions. En primer lloc s'adopta el mètode de calibratge que millor resultats obté basant-se en els mètodes existents tant de calibratge com de preprocessat de dades. Es planteja un preprocessat de dades que permet disminuir el soroll de les mesures i associar una incertesa a cada una d'elles de manera que després es puga utilitzar per al càlcul dels paràmetres. Com resultat s'obtenen els paràmetres de la càmera amb la seua incertesa associada. També es caracteritza des d'un punt de vista estadístic el

procés de càlcul per a determinar com a afecta la incertesa dels paràmetres calculats en funció del nombre d'imatges o nombre de punts en la plantilla. Açò permet posteriorment definir valors sobre el nombre d'imatges a prendre de la plantilla o el nombre de punts a col·locar en la mateixa per a obtindre uns valors raonables d'incertesa. Finalment, es definixen unes posicions i orientacions per a prendre les imatges de la plantilla de manera que el procediment de calibratge es realitze de forma òptima. Estes posicions i orientacions no depenen en cap moment de les característiques de la càmera.

ABSTRACT

Camera calibration is necessary to obtain 3D information from 2D images of the scene. There are different techniques which are based on photogrammetry or self-calibration. Photogrammetric techniques take an image of a known scene which has a three, two or one-dimensional template. With the coordinates of the points in the template and its correspondences in the image the camera is calibrated. Self-calibration techniques take several images of the scene and using the rigidity of the scene they fix restrictions to resolve the camera calibration.

As a result of the calibration the camera intrinsic and extrinsic parameters are computed. Computing the parameters is not an exact process because of the uncertainties. This uncertainty came from constructive imperfections and mechanical misalignments of the lenses or sensor, and because of the image is processed to obtain points positions within them. Results depend so much of the calibration template as the calibration algorithm and also on processing of the data before resolving the camera calibration. A lot of techniques exist to resolve it.

Since it is impossible to compute the exact value of the camera parameter, it proves to be very interesting to compute an interval in which the parameter is. The camera parameter together with its uncertainty improves all procedures in which it participates such as 3D scene reconstruction. By the other hand, when a camera calibration is going to be carried out, a lot of questions arise about the number of points to put into the calibration template, number of images to take of the template and also from where and how to orient the camera to take the images.

This work gives the answer to all these questions. First, the best performance camera calibration algorithm is chosen. A previous data processing has been added to this calibration method to improve the results. Data processing allows to reduce the noise of the measurements and also to associate the uncertainty to each data. If each data has an individual uncertainty, it can be use afterwards in the calibration process according with its uncertainty. As a result, camera parameters are computed together with its uncertainty also. To resolve the questions about the number of points in the template, or number of images of it, the camera calibration process has been characterized from a statistical point of view. This allows defining how the results are affected by the uncertainty of the

input data and therefore, the number of images and points can be determined in order to obtain the best performance of the calibration algorithm. Finally, the optimal positions and orientations of the camera to take the images are defined. These conditions do not depend on the camera features and therefore they can be used with in any camera calibration process.

AGRADECI MI ENTOS

Dar las gracias a mi director de tesis por la complicidad demostrada en todo momento. Las conversaciones mantenidas han servido para encauzar las euforias en momentos de alegría y para levantar el ánimo en momentos menos jubilosos. En todo momento me he sentido apoyado y satisfecho de su labor. También dar las gracias a Martín por hacerme sentir parte del grupo de investigación de Robótica que él mismo dirige atendiendo de forma eficiente las peticiones y necesidades que me han surgido en todo momento.

Desde un punto de vista personal dar las gracias a mis padres por ser lo que soy y como soy sin los cuales este trabajo por supuesto que no hubiera sido posible. Cada una de las líneas de esta tesis es un fruto indirecto de sus esfuerzos realizados durante tantos años. A mí hermana y cuñado por el honor del primer ahijado. Agradecer a mi princesita querida todo lo demás. Risas, alegrías, emociones y cariño han provocado que sea una fuente de estímulo inagotable y necesario para acabar este trabajo.

Agradecer también la confianza de todos mis amigos y en especial, dar las gracias a Miguel Ángel alias Pele, por todos sus ánimos e interpretaciones y a Arturo por todos sus teoremas, sin los cuales el concepto general hubiera quedado desubicado del lado orientado a la jurisdicción que toda tesis tecnológica facultativa experimental y orientada a la ideosincracia social debe tener.

*... a mi familia
y en especial para mi princesita.*

... ante ciertos libros uno se pregunta: ¿quién los leerá?

... ante ciertas personas uno se pregunta. ¿qué leerá?

Al final las personas y los libros se encuentran.

(André Gide, 1869-1951)

... entre las almas ocurre lo mismo.

ÍNDICE

1 INTRODUCCIÓN	1
1.1. TRATAMIENTO DE INCERTIDUMBRES.....	2
1.2. CALIBRACIÓN DE LA CÁMARA.....	3
1.3. OBJETIVOS.....	6
2 EL MODELO DE LA CÁMARA	9
2.1. LA FORMACIÓN DE LA IMAGEN.....	10
2.2. MODELO PIN-HOLE DE LA CÁMARA.....	11
2.3. MODELO LINEAL DE LA CÁMARA.....	15
2.4. MODELO LINEAL APROXIMADO DE LA CÁMARA.....	19
2.5. RESTRICCIONES EN LA MATRIZ DE PROYECCIÓN.....	20
2.6. OBTENCIÓN DE LOS PARÁMETROS DEL MODELO LINEAL.....	23
2.7. HOMOGRAFÍAS.....	29
2.8. EL MODELO NO LINEAL DE LA CÁMARA.....	30
2.9. CONCLUSIONES.....	33
3 EL PROCESO DE CALIBRADO	35
3.1. MÉTODOS DE CALIBRACIÓN.....	36
3.2. CALIBRACIÓN BASADA PLANTILLA 3D.....	41
3.2.1. Solución no homogénea.....	45
3.2.2. Solución homogénea.....	47
3.3. CALIBRACIÓN BASADA EN PLANTILLA BIDIMENSIONAL.....	50
3.3.1. Cálculo de homografías.....	51
3.3.2. Proceso de calibrado con plantillas 2D.....	52

3.4. ESTIMACIÓN NO LINEAL DE LOS PARÁMETROS DE LA CÁMARA.....	54
3.4.1. Errores en la estimación de los parámetros.....	56
3.4.2. Técnicas de optimización	59
3.4.3. Funciones de coste	61
3.5. MEJORAS EN LA ESTIMACIÓN LINEAL DE LOS PARÁMETROS.....	64
3.5.1. Corrección de la distorsión en la imagen	65
3.5.2. Normalización de los datos	68
3.6. MÉTODO DE CALIBRACIÓN DE LA CÁMARA	70
3.7. RESULTADOS EXPERIMENTALES	70
3.8. CONCLUSIONES	84
4 ESTIMACIÓN BASADA EN INCERTIDUMBRES	87
4.1. ERRORES DE ESTIMACIÓN	88
4.2. ESTIMACIÓN ESTADÍSTICA	89
4.2.1. Distribuciones Gaussianas.....	89
4.2.2. Distribuciones χ^2	90
4.2.3. Estimación de máxima verosimilitud	91
4.2.4. Evaluación de algoritmos	93
4.3. ESTIMACIÓN DE PARÁMETROS BASADA EN INCERTIDUMBRES	96
4.3.1. Corrección de los datos.....	97
4.3.2. Cálculo del valor Δu_i	98
4.3.3. Consideraciones prácticas.	99
4.3.4. Matriz de covarianzas de los datos corregidos.....	100
4.3.5. Estimación del vector de parámetros	101
4.3.6. Matriz de covarianzas del vector de parámetros estimado	102
4.3.7. Testeo de hipótesis y estimación del nivel de ruido.....	102
4.4. CALIBRACIÓN DE LA CÁMARA CON LAS INCERTIDUMBRES EN LAS MEDIDAS.....	103
4.4.1. Corrección de la distorsión en la imagen	105
4.4.2. Corrección de errores en la plantilla.....	109

4.4.3. Estimación de las homografías	110
4.4.4. Estimación de los elementos de la matriz $A^{-T} \cdot A^{-1}$	112
4.4.5. Cálculo de los parámetros de la cámara a partir del vector b	113
4.4.6. Búsqueda directa de los parámetros de la cámara	114
4.5. EVALUACIÓN DE ALGORITMOS DE CALIBRACIÓN	115
4.6. RESULTADOS EXPERIMENTALES	118
4.7. CONCLUSIONES	135
5 CONDICIONES ÓPTIMAS.....	137
5.1. ¿QUÉ SE ENTIENDE COMO CONDICIONES ÓPTIMAS?.....	137
5.2. CONDICIONES ÓPTIMAS PARA LA CALIBRACIÓN	139
5.2.1. Condición del sistema.....	141
5.2.2. Posiciones óptimas para tomar las imágenes	143
5.3. CONSIDERACIONES PRÁCTICAS	149
5.4. RESULTADOS EXPERIMENTALES	151
5.5. CONCLUSIONES	153
6 CONCLUSIONES	157
6.1. APORTACIONES REALIZADAS.....	161
6.2. TRABAJO FUTURO.....	162
BIBLIOGRAFÍA.....	163

ÍNDICE DE FIGURAS

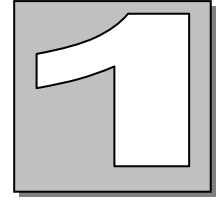
2.1	Formación de la imagen con el modelo Pin-Hole.....	10
2.2	Uso de la lente para obtener mayor iluminación en el plano de la imagen.....	11
2.3	Modelo geométrico de la cámara.....	12
2.4	Cambio del sistema de referencia dentro de la imagen.....	13
2.5	Situación geométrica de los vectores que definen la matriz de proyección.....	21
2.6	Interpretación geométrica del modelo lineal de la cámara.....	22
2.7	Efecto de la distorsión radial y tangencial en la imagen.....	31
2.8	Transformaciones lineales y no lineales contenidas en el modelo de la cámara.....	33
3.1	Plantilla 3D para la calibración de la cámara.....	38
3.2	Calibración de cámara sin conocimiento del escenario mediante auto calibración.	39
3.3	Interpretación geométrica del algoritmo de ajuste por mínimos cuadrados.....	47
3.4	Situación de la cámara en el escenario de simulación.....	71
3.5	Corrección de la distorsión en la imagen capturada.....	72
3.6	Efecto de la distorsión sobre la estimación de los parámetros de la cámara.....	73
3.7	Efecto de la corrección de la distorsión de la imagen en la estimación lineal de los parámetros de la cámara.....	74
3.8	Efecto de la normalización de los datos sobre la estimación lineal de los parámetros de la cámara.....	75
3.9	Efecto del ruido en la plantilla sobre la calibración lineal de la cámara. Los puntos en la imagen tienen un nivel de ruido de $\sigma=1.5$ píxeles.....	76
3.10	Efecto del ruido en la imagen sobre la calibración no lineal de la cámara. Los puntos en las plantillas tienen un nivel de ruido de $\sigma=1$ mm.....	77
3.11	Efecto del ruido en la imagen sobre la calibración no lineal de la cámara. Los puntos en las plantilla no están contaminados con ruido.....	78

3.12	Efecto del ruido en la plantilla sobre la calibración no lineal de la cámara. Los puntos en la imagen tienen un nivel de ruido de $\sigma=0.5$ píxeles.....	79
3.13	Errores cambiando el número de puntos. El ruido en las coordenadas de la imagen es $\sigma=0.5$ y en las coordenadas de la plantilla es $\sigma=0.5$	80
3.14	Errores cambiando el número de imágenes. El ruido en las coordenadas de la imagen es $\sigma=1$ y en las coordenadas de la plantilla es nulo.....	81
3.15	Imágenes reales tomadas para la calibración. Izq. Logitech Quick Cam Pro. Drcha. Canon Powershot A690.....	82
4.1	Interpretación geométrica de los errores en el espacio de medidas.....	95
4.2	Error residual y de estimación en el cálculo de los elementos de una homografía con y sin ruido en las coordenadas de la plantilla y de los parámetros de la cámara con ruido en la plantilla variando el número de imágenes y el número de puntos.....	117
4.3	Diferentes posiciones de las coordenadas de los puntos en la plantilla y en la imagen. El efecto de la corrección hace que los puntos cumplan las supuestas rectas.....	121
4.4	Estimación no lineal basada en incertidumbres de los parámetros de la cámara. Se muestran las incertidumbres de los parámetros calculados. El ruido en la plantilla varía mientras en la imagen permanece constante.....	122
4.5	Estimación no lineal basada en incertidumbres de los parámetros de la cámara. Se muestran las incertidumbres de los parámetros calculados. El ruido en la plantilla es nulo mientras en la imagen varía.....	126
4.6	Efectos del número de puntos y número de imágenes en el cálculo de los parámetros. Si se varía el número de puntos el ruido en la plantilla y en las imágenes es $\sigma=0.5$. Variando en número de imágenes no hay ruido en la plantilla y mientras en la imagen es $\sigma=1$	128
4.7	Estimación no lineal basada en incertidumbres por etapas de los parámetros de la cámara. Se muestran las incertidumbres de los parámetros calculados. El ruido en la plantilla es nulo mientras en la imagen varía.....	130
4.8	Estimación de los parámetros extrínsecos para la primera posición de la cámara. Se muestran las incertidumbres. El ruido en la plantilla es nulo mientras en la imagen varía.....	131
4.9	Errores de estimación de los parámetros intrínsecos con los diferentes algoritmos. El ruido en la plantilla es nulo mientras en la imagen vale $\sigma=1$. Se cambia el	

	número de puntos.....	132
4.10	Escenario de calibración.....	133
5.1	Escenario simulado de calibración para definir las condiciones óptimas.....	139
5.2	Ángulos de rotación de la cámara expresados a partir de la posición de la misma, de forma que el eje óptico pasa por el origen de coordenadas del sistema de referencia del escenario de calibración y eje de coordenadas de la cámara xc permacene paralelo al plano formado por los ejes de coordenadas $xw-yw$	140
5.3	Posiciones óptimas para obtener las imágenes de la plantilla. Una coordenada en el suelo se mantiene nula mientras la otra varía. La altura se calcula partir de las expresiones (5.23) ó (5.24).....	148
5.4	Altura de la cámara en función de lo alejada que se encuentra del origen de medidas. La altura se calcula partir de las expresiones (5.23) ó (5.24).....	149
5.5	Comparación de errores de estimación de los parámetros intrínsecos en diferentes condiciones de calibración. El ruido en la plantilla es nulo mientras en la imagen varía.....	152
5.6	Comparación de errores de estimación de los parámetros intrínsecos en diferentes condiciones de calibración. El ruido en la plantilla es nulo mientras en la imagen vale $\sigma=I$. Se cambia el número de puntos.....	153

ÍNDICE DE TABLAS

3.1 Comparación de resultados de calibración de dos cámaras reales con diferentes métodos.....	83
4.1 Error residual y de estimación para cada una de las etapas del proceso de calibrado de la cámara.....	116
4.2 Comparación de resultados de calibración de dos cámaras reales con diferentes métodos.....	134
5.1 Comparación de resultados de calibración de dos cámaras reales utilizando 3 imágenes tomadas desde posiciones predefinidas y posiciones aleatorias.....	154



INTRODUCCIÓN

Para operaciones autónomas de robots inteligentes en entornos desconocidos, es necesario tener un modelo 3D del entorno en que se encuentra. La fuente fundamental de información 3D para un robot es la visión. Hoy en día, extraer información 3D de imágenes y construir un modelo 3D del entorno es una de las áreas de investigación que atrae más la atención en todo el mundo [5], [6], [24], [30], [39], [46], [68], [99]. La dificultad intrínseca en la obtención del modelo 3D radica en que tanto los datos recogidos como los parámetros del modelo de la cámara no son del todo ciertos. Las imprecisiones del modelo de la cámara surgen ya que los datos obtenidos para realizar la calibración de la misma son erróneos. Los errores en los datos extraídos de las imágenes se originan porque las imágenes que se forman en la cámara sufren distorsiones ópticas, llamadas aberraciones, debidas a imperfecciones constructivas o desalineamiento de las lentes o del sensor. Causas electrónicas que producen distorsiones son las posibles diferencias que puedan existir en las frecuencias de muestreo vertical y horizontal del sensor. Teóricamente, estas distorsiones sistemáticas se pueden corregir a priori si se calibran previamente. Sin embargo, en la realidad es muy difícil eliminar absolutamente estas distorsiones, aunque se tuviera una cámara perfecta desde un punto de vista mecánico y electrónico. En este caso, se producirían imprecisiones debido a falta de iluminación, enfoque defectuoso, poca resolución y factores dependientes de los elementos de la escena tales como sombras, reflejos, brillos, etc.... En consecuencia, los puntos y las líneas que se detectan en la imagen mediante un procesado de la misma no están localizados donde tendrían que estar. En general todo este tipo de errores y desviaciones de los puntos y líneas de sus posiciones ideales es lo que se llama ruido en la imagen. Este ruido en la imagen afecta tanto al proceso de calibrado de la cámara como en la reconstrucción 3D del entorno.

El objetivo del presente trabajo es caracterizar y optimizar el proceso de calibración. Se parte de la incertidumbre de los datos de entrada para mejorar la estimación de los parámetros

de la cámara. El método de calibración elegido es el que mejor resultados obtiene teniendo en cuenta el estado del arte actual. A dicho método se le añade un paso de preprocesado de datos y también se caracteriza para dar respuestas a preguntas del tipo ¿Cuántos puntos son necesarios en la plantilla de calibración? y ¿Cuántas imágenes tomo de la plantilla? Como resultado se obtienen los parámetros de la cámara más parecidos a los reales junto con su incertidumbre asociada. Finalmente se definen las condiciones óptimas bajo las cuales se debe realizar el proceso de calibrado para sacarle el máximo rendimiento de los datos de entrada. En este capítulo de introducción se realiza una breve descripción de los pasos seguidos para conseguir estos objetivos.

1.1. Tratamiento de incertidumbres

Obtener la incertidumbre de los datos, indistintamente de cual sea su fuente, es uno de los grandes retos para los investigadores en el campo de la visión artificial y la robótica. Diferentes técnicas se han utilizado para contrarrestar los efectos del ruido en la imagen, las cuales se pueden clasificar en las que utilizan métodos de cálculo intervalar o las que usan procedimientos estadísticos. Las técnicas intervalares tratan la incertidumbre como un intervalo dentro del cual el dato puede tomar cualquier valor con la misma probabilidad. Asumiendo que el valor del dato puede tomar ciertos valores con más probabilidad que otros, se utilizan las técnicas estadísticas, las cuales tratan la incertidumbre como una probabilidad expresada en una matriz de varianzas/covarianzas. En el caso del usar técnicas estadísticas, existen dos formas de gestionar la incertidumbre. Una de ellas disminuye la incertidumbre a partir del tratamiento de múltiples datos o medidas repetidas. La incertidumbre de los resultados se obtiene a partir de la propagación de la varianza de los datos de partida por el proceso de cálculo [6], [68]. Este planteamiento resulta muy aproximado ya que no resulta fácil establecer la incertidumbre de los datos iniciales. Se establece una estimación de la incertidumbre para todos los datos de partida basada en el error residual que se genera de la estimación de los parámetros. De esta forma, todos los datos participan de igual forma en el proceso de cálculo indistintamente de la incertidumbre que presenten, ya que a todos ellos se les supone la misma. Como resultado se obtiene un resultado aproximado con una incertidumbre sobredimensionada.

Para mejorar el procedimiento de estimación desde un punto de vista estadístico, la incertidumbre se contrarresta partiendo del conocimiento a priori de la estructura de los datos, es decir, si estos cumplen algunos lugares geométricos determinados. Esta forma de eliminar la incertidumbre aprovecha las restricciones geométricas que las medidas deben cumplir en ausencia de ruido para corregirlas. Para un punto o una línea, estas restricciones son posibles colinealidades o coplanaridades entre ellas. Esto permite corregir los errores de medida además de establecer una incertidumbre con cierta confianza para cada uno de los datos de forma

individual. Posteriormente, en la fase de estimación de parámetros, es posible utilizar la incertidumbre de los datos de entrada de forma particular. Esto permite ponderar cada uno de ellos dentro del índice de forma que los datos con menos incertidumbre se tengan más en cuenta que los que sean menos precisos. Esta técnica de estimación es la llamada de máxima verosimilitud, la cual minimiza la distancia de Mahalanobis dentro del espacio de parámetros [46].

La técnica de estimación de parámetros llamada de máxima verosimilitud dispone de herramientas que permiten caracterizar el proceso de estimación de parámetros. En primer lugar se habla de “*precisión del resultado*”. Esta precisión se establece mediante una matriz de covarianzas derivada inicialmente asumiendo que el ruido tiene una distribución Gaussiana. El límite teórico se calcula asumiendo que el número de datos es infinito y como resultado se tiene una relación entre el ruido de los datos de partida y la incertidumbre de los resultados. Esta expresión permite evaluar que condiciones afectan en mayor o menor medida a la estimación de los parámetros. Además, se dispone de un límite el cual sirve para comparar los resultados que se obtienen con diferentes métodos para así decidir cual de todos es más exacto. De esta forma es fácil demostrar que la optimización basada en mínimos cuadrados no resulta del todo óptima.

A partir de la varianza de los parámetros estimados se puede obtener un “*intervalo de confianza*”. Este intervalo de confianza define los valores entre los cuales estará el valor del parámetro con un error establecido en %. Para ello se utiliza la función χ^2 . Esta permite calcular los intervalos de confianza en los que se encuentra un parámetro asumiendo un porcentaje de error para el mismo. Contra menos error se desee obtener, más ancho será el intervalo. Este intervalo definirá la calidad del resultado en términos cuantitativos. Así por ejemplo, si un robot desconoce la calidad de un modelo 3D, no se puede saber si el robot alcanzará una posición determinada dentro de un determinado intervalo de confianza. El saber que la distancia a un objeto estimada de forma óptima es 5 m tiene poca utilidad sino se sabe si la incertidumbre es de ± 1 cm. ó ± 1 m.

1.2. Calibración de la cámara

El proceso de calibración de una cámara es el primer paso para resolver aplicaciones en las que sea necesario obtener datos cuantitativos de la imagen. Aunque es posible obtener información de la escena a partir de imágenes tomadas con cámaras sin calibrar [31], el proceso de calibración resulta esencial cuando se trata de obtener medidas de las mismas. La calibración precisa de la cámara permite obtener distancias en el mundo real a partir de las imágenes tomadas del mismo [23], [30]. Con esta información es posible resolver aplicaciones industriales de ensamblado de piezas [15], evitar obstáculos en la navegación de un robot [9], [16], controlar

un brazo robot o realizar una planificación de trayectorias [48], [53], [54]. Si por el contrario nos centramos en la reconstrucción 3D de objetos, cada punto en la imagen determina un rayo óptico que pasa a través del centro óptico de cámara hacia la escena. El manejo de varias imágenes de una misma escena en la cual no exista movimiento, permite relacionar ambos rayos ópticos para obtener la posición de los puntos 3D en la escena [2], [8], [43]. En este caso es necesario resolver el paso previo de las correspondencias de un objeto en las diferentes imágenes [102]. Una vez se ha podido realizar la reconstrucción 3D del objeto, éste se puede comparar con un modelo almacenado para determinar las imperfecciones del mismo fruto del proceso de fabricación, lo que implica una mejoría importante respecto a la inspección humana [68].

La calibración de la cámara consiste en la estimación de los parámetros intrínsecos de la misma los cuales modelan la geometría interna de la cámara y las características ópticas del sensor. Los parámetros extrínsecos miden la posición y la orientación de la cámara respecto al sistema de coordenadas establecido para la escena. Estos dan la relación respecto al sistema de coordenadas del usuario en lugar del sistema de coordenadas de la cámara.

En la actualidad existen varios métodos para la calibración de una cámara. Atendiendo al método de resolución se pueden clasificar en lineales frente a no lineal o iterativos. Los métodos lineales obtienen un modelo aproximado, sin embargo son fáciles de implementar y muy rápidos para su ejecución [1], [29], [22], [90]. Si por el contrario se requiere un modelo de la cámara más complejo, en el cual se incluyan las distorsiones que produce la cámara es necesario minimizar índices no lineales de forma iterativa [36], [94], [98], [100]. Basándose en el resultado de la calibración, ésta puede ser explícita cuando se obtienen directamente los parámetros del modelo de la cámara [7], [93], [103], mientras que en una implícita se obtienen matrices de transformación que contienen el conjunto de todos los parámetros. Aunque no se conoce el valor exacto de alguno de los parámetros, los resultados pueden ser utilizados para realizar medidas 3D y la generación de coordenadas en la imagen [29], [97], [98]. Atendiendo a los parámetros que forman el modelo de la cámara, los métodos de calibración también se pueden clasificar en intrínsecos y extrínsecos. Los métodos de calibración intrínsecos sólo obtienen los parámetros físicos y ópticos de la cámara [52], [70]. Por el contrario, los extrínsecos calculan la posición y orientación de la cámara en la escena [57], [96]. Finalmente, atendiendo a las características de la plantilla que se utiliza para la calibración, existen métodos basados en fotogrametría o procedimientos de auto calibración. Los métodos que utilizan plantillas de referencia, éstas pueden ser 3D, 2D, 1D y basan la calibración de la cámara en establecer una relación entre las coordenadas conocidas de los puntos en la plantilla y sus correspondientes coordenadas en la imagen. En el caso de plantillas 3D con una sólo imagen de la misma es posible realizar la calibración [22], [37], [44], [94]. En el caso de utilizar plantillas 2D es necesario tomar varias imágenes de la misma desde varias posiciones o cambiar la posición y orientación de la plantilla. No es necesario conocer las posiciones desde donde se toman las imágenes [58], [64], [81], [104].

Este método resulta más versátil ya que la elaboración de la plantilla se puede realizar fácilmente. Los métodos de calibración basados en plantillas 1D son muy útiles en el caso de calibrar sistemas con varias cámaras. En el caso de utilizar métodos basados en plantillas 3D ó 2D, dado que es necesario que todas ellas vean varios puntos de la plantilla de calibración a la vez, resulta complicado establecer una posición para la misma a no ser que la plantilla sea transparente. Es por este motivo que el método de calibración basado en una plantilla 1D resulta atractivo a la hora de calibrar un sistema con varias cámaras [105]. Por otro lado, las técnicas que no utilizan ningún objeto de calibración pueden ser consideradas como de plantilla 0D ya que sólo es necesario establecer correspondencias de un punto en diferentes imágenes. Solamente moviendo la cámara en una escena estática, la rigidez de la escena provoca en general dos restricciones dentro de los parámetros intrínsecos de la cámara [59], [62]. Por lo tanto con varias imágenes de la misma escena tomadas con los mismos parámetros intrínsecos, las correspondencias entre tres imágenes son suficientes para calcular tanto los parámetros intrínsecos como extrínsecos. Esta regla permite a su vez realizar la reconstrucción 3D de una estructura a partir de varias imágenes de la misma.

Otro aspecto a tener en cuenta es la familia de algoritmos que calculan los parámetros que modelan la distorsión producida por las lentes en la imagen sin el uso de objetos de calibración y por lo tanto sin conocer ninguna estructura 3D. Estos métodos se basan en que dentro de una proyección perspectiva ideal, la cámara transforma líneas rectas en el espacio 3D en líneas rectas dentro del espacio 2D correspondiente a la imagen. Por lo tanto reforzando la linealidad de las partes de la imagen que se ven curvas debido a la distorsión de la lente de la cámara, es posible estimar la deformación que ésta produce [3], [13], [19], [71], [83]. Hay métodos que utilizan restricciones epipolares y trilineales entre pares y tripletes de imágenes respectivamente para estimar la distorsión radial [86].

Desde el punto de vista de los errores cometidos en el proceso de calibrado de la cámara existen una serie de factores a tener en cuenta. Claramente, con datos más exactos es posible conseguir una calibración mejor. Sin embargo, resulta más efectivo saber que precisión se puede conseguir en los resultados de calibración con los escenarios y métodos que se disponen. Existen estudios sobre los errores cometidos en la calibración debidos a imprecisiones en las medidas del plano de la imagen o en la plantilla de calibración [49], [50], [51], [55], [103]. Además al proceso de calibración, no sólo afecta el ruido de las medidas, sino también la asociación errónea de un punto en el mundo con un punto en la imagen. Algunos autores utilizan herramientas estadísticas para detectar esta situación [25], [40]. También se pueden tener en cuenta errores de cálculo debido a inestabilidades de la herramienta matemática [87]. Una normalización de los datos permite que el algoritmo sea menos sensible al ruido de las medidas y obtenga unos resultados más exactos [32]. Otro factor importante es el modelo de cámara que se pretende estimar. Un modelo muy complejo puede complicar mucho el algoritmo

de búsqueda y los resultados obtenidos no producir mejoras significativas frente a otro modelo más simple. Un modelo excesivamente complicado incluso puede producir inestabilidades en el proceso de búsqueda y llegar a resultados absurdos [100].

A partir de todo este conocimiento, desde un punto de vista práctico, se puede decir que a la hora de calibrar una cámara lo interesante es saber que algoritmo de calibración ó plantilla utilizar, número de puntos a colocar en la plantilla de calibración, número de imágenes tomadas de la misma, así como las posiciones y orientaciones desde las que tomar las imágenes.

1.3. Objetivos

Esta tesis pretende dar respuesta a todas las cuestiones que se plantean a la hora de calibrar una cámara. En primer lugar establece un estado del arte de los métodos de calibración de cámaras así como de las técnicas existentes para la mejora del proceso. A partir de aquí, se adopta el método de calibración que mejor resultados obtiene basándose en los estudios existentes y propios realizados por el propio autor, tanto de calibración como de preprocesado de datos. Se plantea un preprocesado de datos que permite disminuir el ruido de las medidas y asociar un error a cada una de ellas de forma que el cálculo de los parámetros se realice basándose en la incertidumbre de los datos. Como resultado se obtienen los parámetros de la cámara con un intervalo asociado. También se caracteriza desde un punto de vista estadístico el proceso de cálculo obteniéndose valores acerca del número de imágenes a tomar de la plantilla o el número de puntos a colocar en la misma y también se determina como se ve afectada la incertidumbre de los parámetros calculados en función del número de imágenes o número de puntos en la plantilla. Finalmente, se pretenden definir unas posiciones y orientaciones para tomar las imágenes de la plantilla de forma que el procedimiento de calibración sea óptimo. Estas posiciones y orientaciones deben ser generales y no depender en ningún momento de las características de la cámara.

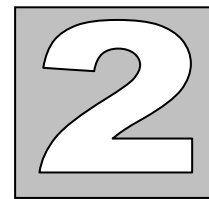
En el capítulo 2 de la tesis se realiza una revisión de todos los modelos geométricos de la cámara que se utilizan en las diferentes técnicas de calibración. Esto permite definir la nomenclatura de los diferentes parámetros que definen la cámara y las restricciones que existen entre ellos las cuales se utilizarán en capítulos posteriores.

En el capítulo 3 se hace un estudio del estado del arte referente a todos los métodos de calibración de cámaras basados en técnicas de fotogrametría. Se tienen en cuenta todos los métodos basados en plantilla tridimensional, bidimensional y unidimensional. También se evalúan todos los métodos que existen en la actualidad para realizar un preprocesado de los datos antes de utilizarlos para calibrar la cámara. Con este preprocesamiento se trata de corregir las distorsiones que se producen en la captura de la imagen, o la normalización de las

coordenadas para condicionar bien el algoritmo de cálculo. El objetivo es definir el procedimiento de calibración que mejores resultados obtenga. Para ello se han utilizado experiencias realizadas por otros autores así como simulaciones realizadas por el propio autor con el objetivo de decidir que metodología seguir a la hora de calibrar una cámara.

En el capítulo 4, una vez definido el método de calibración que mejor resultados obtiene se trata de mejorarlo y caracterizarlo. Esta caracterización y mejora se realiza con el uso de herramientas estadísticas. Se establece una metodología para corregir la medida de las coordenadas en la imagen de forma que se consiga contrarrestar los errores de las medidas y a su vez se obtenga las incertidumbres de los datos que van a participar en la calibración. Posteriormente, con los datos corregidos junto con sus incertidumbres se realiza el procedimiento de calibración de la cámara basado en plantilla bidimensional con técnicas de máxima verosimilitud. Esto permite obtener los parámetros de la cámara con incertidumbres además de utilizar cada dato de forma inversamente proporcional al error que presenta. Además el uso de técnicas de estimación estadística permite caracterizar el método de calibración y determinar el número de imágenes a tomar de la plantilla o el número de puntos a colocar en la plantilla. También permite definir la incertidumbre mínima que se puede conseguir a partir de una incertidumbre de los datos de entrada o en que medida se ve afectada la incertidumbre por parámetros tales como número de puntos en la plantilla o número de imágenes tomadas de la plantilla.

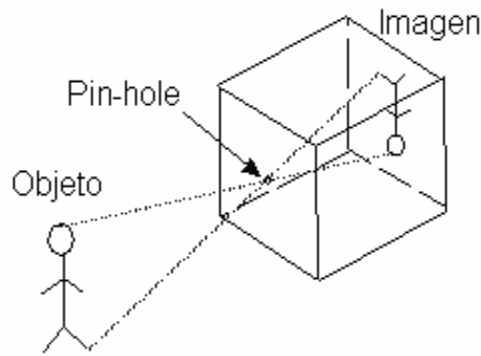
En el capítulo 5 y último se responde a la cuestión referente a la posición y orientación desde las que tomar las imágenes de la plantilla. En este capítulo se plantean las circunstancias que condicionan de forma óptima el proceso de cálculo que involucra el calibrado de la cámara. El algoritmo de calibración resuelve diferentes expresiones cuyas soluciones son vectores propios asociados a los valores propios más pequeños de una matriz. En este caso se plantea optimizar la condición de dichas matrices en la medida de lo posible para que los vectores y valores propios resultantes sean poco sensibles al ruido de los elementos de la matriz. Para ello se define un escenario de calibración genérico junto con una pequeña modificación del modelo de forma que permita determinar las posiciones y orientaciones desde las que tomar las imágenes para que la matriz resultante esté bien condicionada. Estas posiciones aseguran un cálculo de los parámetros de la cámara más aproximados a los reales en el caso de resolver el algoritmo de calibración en su fase lineal solamente. Además estas posiciones no dependen de las características de la cámara lo que permite que se puedan utilizar para cualquier cámara. En la fase de búsqueda no lineal de los parámetros, si se asegura que físicamente ciertos parámetros ya tienen un determinado valor, esto reduce el vector de parámetros a evaluar y en consecuencia el algoritmo converge de forma más rápida y además los resultados son más cercanos a los reales.



EL MODELO DE LA CÁMARA

En este capítulo se describe el modelo geométrico asociado al proceso de formación de una imagen con la cámara. El modelo está caracterizado por un cierto número de parámetros los cuales caracterizan las propiedades intrínsecas de la cámara y su situación en el mundo. Desde un punto de vista geométrico la situación de un punto en la imagen es resultado de transformaciones lineales y perspectivas aplicadas a las coordenadas del punto en el mundo. El modelo geométrico consiste en una matriz de proyección, la cual contiene todas estas operaciones. Si las propiedades intrínsecas no cambian y la situación de la cámara es la misma, es posible saber, utilizando esta matriz de proyección, la posición de un punto en la imagen conociendo la posición del mismo en el mundo. Asimismo, partiendo del modelo de la cámara es posible localizar la cámara en un entorno a partir de las imágenes obtenidas del mismo, así como obtener las características intrínsecas de la cámara.

El capítulo esta estructurado de la siguiente forma. En la primera parte se obtiene el modelo lineal de la cámara a partir del modelo básico “pin-hole” y las transformaciones lineales que experimentan las coordenadas de los puntos en la imagen. También se describe la formación del modelo a partir de los parámetros que definen las características de la cámara y su posición en el mundo. En la segunda parte se realiza una interpretación geométrica del modelo de la cámara para realizar el paso inverso. Este paso consiste en la obtención de los parámetros intrínsecos de la cámara y su situación en el mundo a partir de una matriz de proyección dada. Para finalizar se describen las transformaciones no lineales que sufren las coordenadas de los puntos en la imagen, las cuales son consecuencia de las distorsiones que introduce la cámara real en el proceso de formación de la imagen.



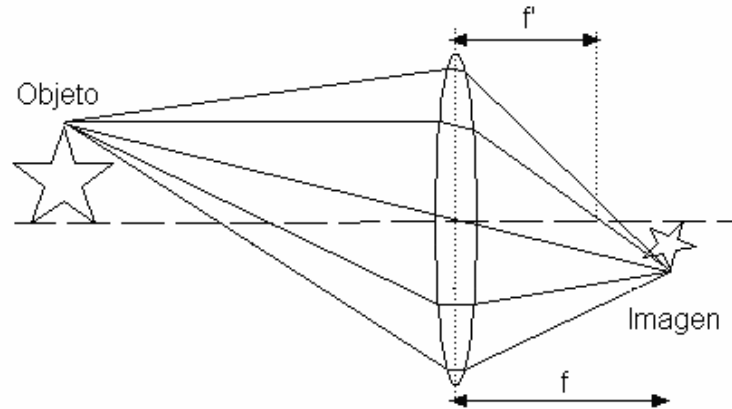
2.1 Formación de la imagen con el modelo Pin-Hole

2.1. La formación de la imagen

El proceso de formación de una imagen consta de dos partes a tener en cuenta. La primera parte, principalmente geométrica, determina la posición en el plano de la imagen, a partir de la proyección de un punto de la escena. La segunda consiste en la naturaleza de la luz la cual determina el brillo de un punto en el plano de la imagen en función de las propiedades de la superficie y de la iluminación.

El sistema más elemental para formar una imagen de una escena 3D en una superficie 2D consiste en una caja totalmente cerrada. De esta caja se utilizan dos lados enfrentados los cuales simulan dos pantallas situadas de forma paralela. En la primera pantalla se realiza un pequeño orificio muy pequeño a través del cual solamente tendría que pasar un fotón de luz. Este agujero permite que los rayos de luz emitidos o reflejados por un objeto de la escena traspasen la primera pantalla y formen una imagen invertida en la segunda tal como muestra la figura 2.1. Para recoger la imagen simplemente es necesario colocar un elemento fotosensible en la segunda pantalla. Dado que en la práctica el pequeño orificio no deja pasar suficiente luz para excitar el elemento fotosensible, se colocan lentes en la apertura para enfocar el paquete de rayos que llega de un punto de la escena al correspondiente punto del plano de la imagen. El esquema que representa este proceso se ve reflejado en la figura 2.2.

El efecto de las lentes se basa en el principio de refracción de la luz. Un haz de luz se refracta cuando encuentra un obstáculo transparente, y como resultado, éste sufre un cambio en su trayectoria. Este cambio de trayectoria viene determinado por el ángulo de refracción, el cual depende del ángulo de incidencia y de la longitud de onda del rayo de luz. Utilizando el ángulo de refracción es posible conseguir con una lente que todos los rayos de luz procedentes de un mismo punto de la escena se corten en un solo punto detrás de la lente. Este punto de corte depende del ángulo de incidencia sobre la lente de los rayos de luz. Para conseguir una imagen



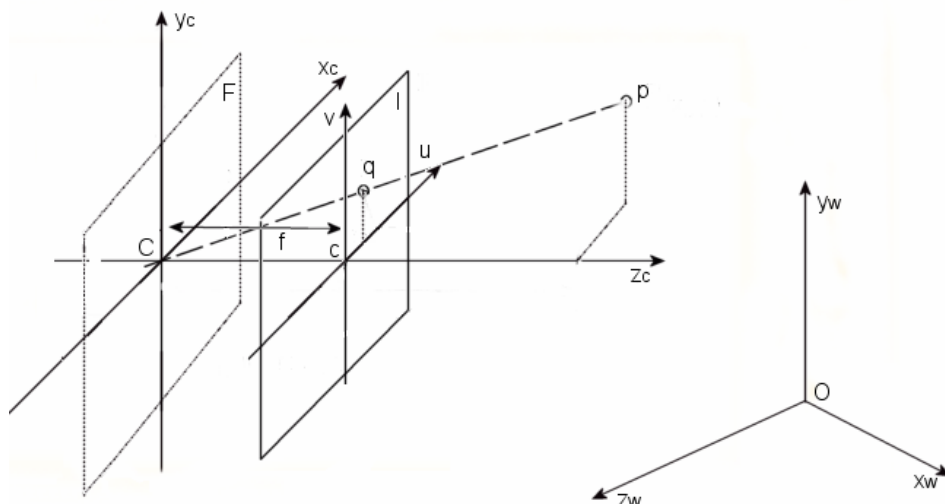
2.2 Uso de la lente para obtener mayor iluminación en el plano de la imagen

nítida de la escena, es necesario que el plano de la imagen esté situado justo en esa distancia de corte. Ésta se llama distancia focal del sistema de visión. La distancia focal de una lente es la separación entre la lente y el punto de corte de rayos de luz procedentes del infinito. Ésta no tiene porqué coincidir con la distancia focal del sistema de visión ya que ésta última se ajusta en función de la distancia a que se encuentra la escena a retratar.

El elemento sensor es normalmente un CCD del inglés Charge-Coupled Device, “dispositivo de cargas (eléctricas) interconectadas”. El CCD es un circuito integrado que contiene un número determinado de condensadores enlazados o acoplados. Bajo el control de un circuito interno, cada condensador puede transferir su carga eléctrica a uno o a varios de los condensadores que estén a su lado en el circuito impreso. La alternativa digital a los CCD son los dispositivos CMOS (Complementary Metal Oxide Semiconductor) utilizados en algunas cámaras digitales y en numerosas Webcam. En la actualidad los CCD son mucho más populares en aplicaciones profesionales y en cámaras digitales. Los primeros dispositivos CCD fueron inventados por Willard Boyle y George Smith en 1969 en los Laboratorios Bell. Los condensadores del CCD acaban funcionando como diminutas células fotoeléctricas que registran la imagen. Desde allí la imagen es procesada por la cámara y registrada en la tarjeta de memoria.

2.2. Modelo pin-hole de la cámara

Éste es el modelo más simple el cual se utiliza como base para todos los demás modelos. Representa una cámara ideal libre de distorsiones. Se puede obtener el modelo geométrico de la cámara partiendo de la figura 2.3. I es el plano de la imagen. F es el plano focal en el cual todos los puntos tienen $z_c=0$. El punto o' es el centro óptico o centro de proyección, el cual se sitúa a una distancia focal f del origen del sistema de coordenadas de la cámara c (distancia focal del



2.3 Modelo geométrico de la cámara

sistema de visión). El centro óptico se utiliza para formar la imagen de un punto p en el plano de la imagen I . La imagen del punto p llamada punto q , se obtiene por la intersección de la recta $o'p$ con el plano de la imagen I . El eje óptico es la recta que pasa por el centro óptico o' y es perpendicular al plano I . El punto c es la intersección del eje óptico con el plano de la imagen, también llamado punto principal. Para el modelo pin-hole de la cámara se va a tomar este punto como origen de medidas en la imagen. El plano focal F es paralelo al plano de la imagen y pasa por el centro óptico o' . Los puntos situados en el plano focal no tienen imagen en el plano de la imagen I ya que la recta $o'p$ es paralela al plano de la imagen I y nunca la corta.

Desde el punto de vista de la geometría proyectiva, esta recta corta al plano de la imagen en el infinito. Se fija un sistema de coordenadas 3D para el escenario ($o, \{x_w, y_w, z_w\}$), otro sistema de coordenadas 3D para la cámara ($o', \{x_c, y_c, z_c\}$) y otro 2D en el plano de la imagen ($c, \{u, v\}$). El eje óptico está alineado con el eje Z del sistema de referencia de la cámara tal como se observa en figura 2.3. Las coordenadas de los puntos de interés en el espacio están referidas al sistema de coordenadas del escenario ($o, \{x_w, y_w, z_w\}$) y sus correspondientes posiciones en la imagen están referidas respecto al sistema de coordenadas de la misma ($c, \{u, v\}$).

Para caracterizar el modelo ideal de la cámara hay que tener en cuenta dos transformaciones que realiza la misma. Con estas transformaciones se obtienen las coordenadas de los puntos en la imagen partiendo de sus posiciones respecto al sistema de referencia del escenario. En primer lugar existe una transformación del sistema de referencia métrico del escenario, al sistema de referencia métrico asociado a la cámara. En segundo lugar, existe una proyección perspectiva que transforma un punto de interés situado en el escenario en un punto 2D de la imagen. Dadas las coordenadas de un punto p respecto al sistema de referencia del

escenario $\mathbf{p}_w=(x_w, y_w, z_w)$, las coordenadas de este mismo punto \mathbf{p} respecto al sistema de referencia de la cámara $\mathbf{p}_c=(x_c, y_c, z_c)$ son:

$$\mathbf{p}_c = \mathbf{R} \cdot \mathbf{p}_w - \mathbf{t} \quad (2.1)$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11}x_w & r_{12}y_w & r_{13}z_w \\ r_{21}x_w & r_{22}y_w & r_{23}z_w \\ r_{31}x_w & r_{32}y_w & r_{33}z_w \end{bmatrix} - \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

\mathbf{R} es la matriz de rotación 3x3 ortonormal que relaciona ambos sistemas de referencia y $\mathbf{t}=(t_x, t_y, t_z)$ son las coordenadas del centro del proyección \mathbf{o}' respecto al origen de coordenadas del escenario. La expresión (2.1) representa la relación que existe en las coordenadas de un mismo punto expresadas según el sistema de referencia del escenario o el sistema de referencia de la cámara. La matriz \mathbf{R} y el vector \mathbf{t} expresan la relación entre ambos sistemas de referencia.

Para la obtención de las coordenadas del punto \mathbf{q} respecto del sistema de coordenadas de la imagen $\mathbf{q}_c=(u_c, v_c)$ se realiza la proyección perspectiva de las coordenadas $\mathbf{p}_c=(x_c, y_c, z_c)$ respecto al sistema de coordenadas de la cámara.

$$\frac{f}{z_c} = \frac{u_c}{x_c} = \frac{v_c}{y_c} \quad (2.2)$$

$$\mathbf{q}_c = \begin{bmatrix} u_c \\ v_c \end{bmatrix} = f \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \end{bmatrix} \quad (2.3)$$

f es la distancia focal del sistema de visión. En el espacio proyectivo, si $z_c=0$ indica que el punto está en el plano focal de la cámara. En este caso las coordenadas del punto en la imagen $\mathbf{q}_c=(u_c, v_c)$ no están definidas y corresponden a un punto en el infinito.

Los parámetros que relacionan el sistema de coordenadas del escenario y el sistema de coordenadas de la cámara determinan la posición y la orientación de la cámara en el escenario. Estos son llamados parámetros extrínsecos. La distancia focal caracteriza una cámara determinada y es independiente de la posición de la misma en el escenario. La distancia focal forma parte de los parámetros intrínsecos de la cámara. Si se utilizan coordenadas homogéneas el modelo pin-hole de la cámara queda definido según la siguiente expresión:

$$\mathbf{q}_c = F \mathbf{R}_t \mathbf{p}_w \quad (2.4)$$

$$\begin{bmatrix} w \cdot u_c \\ w \cdot v_c \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & -t_x \\ r_{21} & r_{22} & r_{23} & -t_y \\ r_{31} & r_{32} & r_{33} & -t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

La matriz \mathbf{R}_t representa la traslación y la rotación del sistema de referencia de la cámara en el sistema de referencia del escenario y \mathbf{F} representa la matriz de proyección perspectiva. La ventaja que presenta el uso de coordenadas homogéneas es que la proyección perspectiva se convierte en una relación lineal entre las coordenadas. Las coordenadas cartesianas pueden ser obtenidas realizando la división por el factor de escala representado por la última componente.

La relación entre la distancia focal f y la matriz \mathbf{F} es directa. Lo mismo ocurre con la traslación $\mathbf{t}=(t_x, t_y, t_z)$ y la matriz de \mathbf{T} . Sin embargo, la relación entre los ángulos de rotación y la matriz \mathbf{R} es más complicada. \mathbf{R} es ortonormal y sus columnas expresan coordenadas de los ejes en el sistema de coordenadas rotado. \mathbf{R} tiene nueve parámetros pero sólo tres grados de libertad. Normalmente una rotación se expresa en ángulos de Euler como una secuencia de tres rotaciones elementales $\omega_x, \phi_y, \varphi_z$ a lo largo de los ejes del sistema de coordenadas cartesiano. El problema de esta representación es que tiene una singularidad cuando $\phi_y=\pi/2$ ó $\varphi_z=\pi/2$. Esta representación puede ser aceptable cuando las posiciones de la cámara están restringidas en el espacio. Una representación alternativa la cual no presenta singularidades es mediante el uso de cuaterniones. Una explicación detallada se puede encontrar en [69].

Hasta este momento se ha realizado la transformación del sistema de coordenadas del escenario al que tiene la cámara en dos pasos. Primero se realiza una rotación del sistema de coordenadas con la matriz \mathbf{R} y después se traslada desde el origen de medidas del escenario al de la cámara una traslación \mathbf{t} , tal como se refleja en la expresión (2.1). Desde otro punto de vista, también es posible realizar la transformación entre los dos sistemas de coordenadas realizando primero una traslación \mathbf{t} del sistema de referencia y después realizar la rotación del mismo con la matriz \mathbf{R} . Algunos autores utilizan esta segunda interpretación. En este caso las expresiones que se obtienen son las siguientes:

$$\mathbf{p}_c = \mathbf{R} \cdot (\mathbf{p}_w - \mathbf{t}) \quad (2.5)$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11}(x_w - t_x) & r_{12}(y_w - t_y) & r_{13}(z_w - t_z) \\ r_{21}(x_w - t_x) & r_{22}(y_w - t_y) & r_{23}(z_w - t_z) \\ r_{31}(x_w - t_x) & r_{32}(y_w - t_y) & r_{33}(z_w - t_z) \end{bmatrix}$$

Si se añaden la proyección perspectiva de \mathbf{p}_c el modelo de la cámara según esta segunda interpretación es el siguiente:

$$\mathbf{q}_c = \mathbf{F} \mathbf{R} \mathbf{T} \mathbf{p}_w \quad (2.6)$$

$$\begin{bmatrix} w \cdot u_c \\ w \cdot v_c \\ w \end{bmatrix} = \begin{bmatrix} -f & 0 & 0 \\ 0 & -f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & -t_x \\ 0 & 1 & 0 & -t_y \\ 0 & 0 & 1 & -t_z \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

En este caso T representa la matriz de traslación del origen de medidas del escenario al origen de medidas de la cámara y R representa la matriz de rotación que existe entre ambos sistemas de coordenadas, la cual se realiza en segundo lugar. Ambos modelos son igualmente válidos.

2.3. Modelo lineal de la cámara

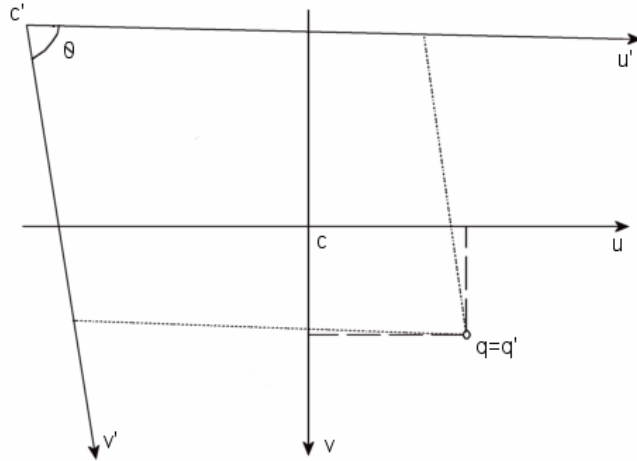
El modelo lineal de la cámara es una matriz de transformación proyectiva de 3x4 elementos, la cual relaciona las coordenadas de un punto $p_w=(x_w, y_w, z_w)$ en el sistema de coordenadas del escenario con su posición en el sistema de coordenadas del plano de la imagen $q_c=(u_c, v_c)$.

$$q_c = M p_w \quad (2.7)$$

$$\begin{bmatrix} w \cdot u_c \\ w \cdot v_c \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

Utilizando la matriz de transformación proyectiva M se puede tratar una cámara, como una caja negra que realiza la transformación del espacio proyectivo P^3 al espacio proyectivo P^2 . Este modelo es una extensión del modelo pin-hole. En el modelo pin-hole, el origen de coordenadas del sistema de referencia de la imagen es el punto c , el cual según se observa en la figura 2.3, es la intersección del eje óptico con el plano de la imagen. Además, las unidades de medida del sistema de referencia de la imagen son milímetros. Es necesario adaptar este modelo para que se represente el comportamiento de las cámaras reales. En este caso, los puntos de la imagen se miden en píxeles respecto al sistema de referencia $(c', \{u', v'\})$ asociado a una esquina de la imagen según se ve en la figura 2.4. La transformación de milímetros a píxeles que se realiza en la cámara en cada uno de los ejes es diferente y ésta depende de la tecnología de la cámara. Por otro lado, desde el punto de vista del montaje mecánico de la cámara, es posible que el sensor CCD no esté montado perfectamente perpendicular al eje óptico. Un pequeño giro del plano de la imagen respecto a una de sus diagonales puede producir una pérdida de ortogonalidad de los ejes del sistema de coordenadas de la imagen. Esta pérdida de ortogonalidad se ve reflejada posteriormente en un diferente escalado en cada uno de los ejes.

Así pues, el modelo pin-hole se ve modificado ligeramente debido al comportamiento que presentan las cámaras reales. Estas variaciones son el desplazamiento del origen de medidas a una esquina de la imagen, la no ortogonalidad en los ejes de medida y las diferentes escalas en cada uno de los ejes que realizan el cambio de unidades de milímetros a píxeles. La obtención de las coordenadas en milímetros $q_c=(u_c, v_c)$ obtenidas a partir de coordenadas en píxeles extraídas



2.4 Cambio del sistema de referencia dentro de la imagen

de la imagen $\mathbf{q}_p=(u_p, v_p)$ se puede realizar de varias formas. A continuación se muestran dos de ellas.

Modelo 1:

En este caso la relación entre las coordenadas del punto \mathbf{q} expresadas en píxeles $\mathbf{q}_p=(u_p, v_p)$ y las coordenadas expresadas en milímetros $\mathbf{q}_c=(u_c, v_c)$ es:

$$\mathbf{q}_p = \mathbf{S}' \cdot \mathbf{q}_c - \mathbf{s} \quad (2.8)$$

$\mathbf{s}=(u_0, v_0)$ son las coordenadas en píxeles del punto principal o centro de la imagen \mathbf{c} respecto del sistema de referencia $(\mathbf{c}', \{\mathbf{u}', \mathbf{v}'\})$, según se observa en la figura 2.4. La matriz \mathbf{S}' contiene la diferencia de escalado en cada uno de los ejes y su pérdida de ortogonalidad.

Suponiendo que los ejes de coordenadas de la imagen son ortogonales, las coordenadas del punto \mathbf{q} en píxeles $\mathbf{q}_p=(u_p, v_p)$ a partir de sus coordenadas en milímetros $\mathbf{q}_c=(u_c, v_c)$ son:

$$\begin{aligned} u_p &= k_u \cdot u_c - u_0 \\ v_p &= k_v \cdot v_c - v_0 \end{aligned} \quad (2.9)$$

Esta relación expresada de forma matricial utilizando coordenadas homogéneas es la siguiente:

$$\mathbf{q}_p = \mathbf{S} \mathbf{q}_c \quad (2.10)$$

$$\begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = \begin{bmatrix} k_u & 0 & -u_0 \\ 0 & k_v & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix}$$

Si se considera la pérdida de ortogonalidad de los ejes de coordenadas en la imagen debidas a imperfecciones en el proceso de construcción de la cámara, y suponiendo que el eje v está alineado con el eje v' , la relación entre las coordenadas del sistema de referencia ortogonal $(c, \{u, v\})$ con el nuevo sistema de referencia $(c', \{u', v'\})$ es la siguiente [38]:

$$\begin{aligned} u' &= u \sin \theta + v \cos \theta \\ v' &= v \end{aligned} \quad (2.11)$$

Teniendo en cuenta la pérdida de ortogonalidad, la expresión (2.10) debe ser redefinida de la siguiente forma:

$$\begin{aligned} q_p &= S_\theta q_c \quad (2.12) \\ \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} &= \begin{bmatrix} k_u & k_v \cot \theta & u'_0 \\ 0 & k_v / \sin \theta & v'_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} \\ u'_0 &= u_0 + v_0 \cot \theta \\ v'_0 &= v_0 / \sin \theta \end{aligned}$$

Modelo 2:

Según Melen [63], la relación entre las coordenadas reales y las distorsionadas es,

$$\begin{aligned} q_c &= B \cdot (q_p - v) \quad (2.13) \\ \begin{bmatrix} u_c \\ v_c \end{bmatrix} &= \begin{bmatrix} 1+b_1 & b_2 \\ b_2 & 1-b_1 \end{bmatrix} \left(\begin{bmatrix} u_p \\ v_p \end{bmatrix} - \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \right) \end{aligned}$$

Igual que antes, $s=(u_0, v_0)$ representa las coordenadas en milímetros del punto principal o centro de la imagen c respecto del sistema de referencia $(c', \{u', v'\})$, según se observa en la figura 2.3. La matriz B contiene los coeficientes b_1 y b_2 los cuales compensan la diferencia de escalado y la pérdida de ortogonalidad entre los ejes de la imagen. Utilizando coordenadas homogéneas se tiene la expresión equivalente:

$$\begin{aligned} q_c &= B \cdot V \cdot q_p \quad (2.14) \\ \begin{bmatrix} u_c \\ v_c \\ 1 \end{bmatrix} &= \begin{bmatrix} 1+b_1 & b_2 & 0 \\ b_2 & 1-b_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} \end{aligned}$$

Para obtener las coordenadas en píxeles $q_p=(u_p, v_p)$ a partir de las coordenadas en milímetros $q_c=(u_c, v_c)$ es necesario realizar las inversas de las matrices B y V ,

$$q_p = V^{-1} \cdot B^{-1} \cdot q_c \quad (2.15)$$

$$B^{-1} = \begin{bmatrix} \frac{1-b_1}{1-b_1^2-b_2^2} & -\frac{b_2}{1-b_1^2-b_2^2} & 0 \\ -\frac{b_2}{1-b_1^2-b_2^2} & \frac{1+b_1}{1-b_1^2-b_2^2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$V^{-1} = \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

B será invertible siempre que $b_1^2 + b_2^2 < 1$. Esta restricción implica que $|b_1| < 1$ y que $|b_2| < 1$, la cual cosa es razonable ya que si $|b_1| \gg 1$ éste cambiaría la dirección de uno de los ejes y si $|b_2| \gg 1$ cambiaría un eje por el otro.

Modelo lineal de la cámara

Una vez obtenidas las transformaciones que sufren las coordenadas de los puntos dentro de una cámara real, es posible obtener el modelo lineal de la cámara. Este modelo obtiene las coordenadas de un punto en la imagen, tomado con una determinada cámara y sabiendo que se encuentra situado en una posición en la escena. El modelo matemático representado por la expresión (2.7) se puede obtener a partir de la expresión (2.4) o (2.6) junto con el modelo del comportamiento de la cámara que se desee utilizar representado por (2.12) o (2.15). En este caso se asume que la cámara es localizada en la escena a partir de una traslación del origen de coordenadas y una posterior rotación de los ejes de coordenadas. También se supone que la construcción de la cámara no ha sido perfecta y por lo tanto existe una cierta pérdida de ortogonalidad en los ejes de coordenadas de la imagen. Esto implica que el modelo matemático de la cámara (2.7) se puede obtener a partir de la expresión (2.4) y (2.12):

$$M = \lambda \cdot S_\theta \cdot F \cdot R_t \quad (2.16)$$

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} = \lambda \begin{bmatrix} k_u & k_v \cot \theta & u'_0 \\ 0 & k_v / \sin \theta & v'_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & -t_x \\ r_{21} & r_{22} & r_{23} & -t_y \\ r_{31} & r_{32} & r_{33} & -t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.17)$$

Utilizando el segundo modelo (2.15) junto con la segunda interpretación en el cambio de sistema de referencia (2.6), se tiene:

$$M = \lambda \cdot V^{-1} \cdot B^{-1} \cdot F \cdot R \cdot T \quad (2.18)$$

$$\begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} = \lambda \begin{bmatrix} \frac{1-b_1}{1-b_1^2-b_2^2} & -\frac{b_2}{1-b_1^2-b_2^2} & 0 \\ -\frac{b_2}{1-b_1^2-b_2^2} & \frac{1+b_1}{1-b_1^2-b_2^2} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.19)$$

$$\begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -f & 0 & 0 \\ 0 & -f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \end{bmatrix}$$

El factor de escala λ que multiplica a toda la expresión es un escalar arbitrario que nunca puede valer cero. Su valor únicamente afecta a la coordenada homogénea del punto de la imagen w según se ve en la expresión (2.7). Aunque hay dieciocho parámetros en el modelo de la cámara, sólo existen doce grados de libertad debido a que hay seis restricciones de ortogonalidad en la matriz de rotación \mathbf{R} . De estos doce grados de libertad once corresponden a los parámetros de la cámara ya que la igualdad está expresada en coordenadas homogéneas.

Los parámetros f , k_u , k_v , u_0' , v_0' y θ definen las características de la cámara. k_u , k_v , f representan el escalado de milímetros a píxeles, u_0' , v_0' definen indirectamente el tamaño en píxeles de la imagen y θ indica el ángulo que forman los ejes de la imagen. En ninguno de los casos dependen de la posición ni de la orientación de la cámara en la escena. Es por este motivo que dentro del modelo de la cámara son llamados *parámetros intrínsecos*. Las matrices \mathbf{T} y \mathbf{R} constituyen la posición y la orientación de la cámara respecto al sistema de referencia de la escena. Dado que no dependen para nada de la característica de la misma, estas matrices contienen los llamados *parámetros extrínsecos* del modelo de la cámara.

2.4. Modelo lineal aproximado de la cámara

En este apartado se presenta un modelo aproximado al modelo general de la cámara. Se realiza una aproximación de primer orden la cual es válida cuando la distancia de la escena a la cámara es relativamente grande respecto a la separación de los puntos del eje óptico.

Asumiendo la representación de la figura 2.3, partiendo de las coordenadas en milímetros de un punto \mathbf{p} en la escena $\mathbf{p}_w=(x_w, y_w, z_w)$, se pueden obtener las coordenadas de su correspondiente punto \mathbf{q} de la imagen $\mathbf{q}_c=(u_c, v_c)$, según la siguientes expresiones;

$$\begin{aligned} u_c &= f \frac{r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} = f \frac{r_1 \mathbf{p}_w + t_x}{r_3 \mathbf{p}_w + t_z} \\ v_c &= f \frac{r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} = f \frac{r_2 \mathbf{p}_w + t_y}{r_3 \mathbf{p}_w + t_z} \end{aligned} \quad (2.20)$$

En este caso u_c , v_c son las coordenadas en milímetros respecto al sistema de referencia de la cámara. r_1 , r_2 , r_3 son los vectores fila de la matriz de rotación \mathbf{R} . Si se dividen estas expresiones por la traslación de la cámara en el eje Z definido por t_z , se obtiene,

$$\begin{aligned} u_c &= \frac{r_1' p_w + t_x / t_z}{e + 1} \\ v_c &= \frac{r_2' p_w + t_y / t_z}{e + 1} \end{aligned} \quad (2.21)$$

r_1 , r_2 y e valen:

$$r_1' = \frac{f r_1}{t_z} \quad r_2' = \frac{f r_2}{t_z} \quad e = \frac{r_3 p_w}{t_z} \quad (2.22)$$

A partir del valor de e se deduce que si la distancia de la cámara a la escena t_z es grande o los puntos de interés de la escena no están muy separados del origen de coordenadas de la misma el modelo de la cámara se puede aproximar a

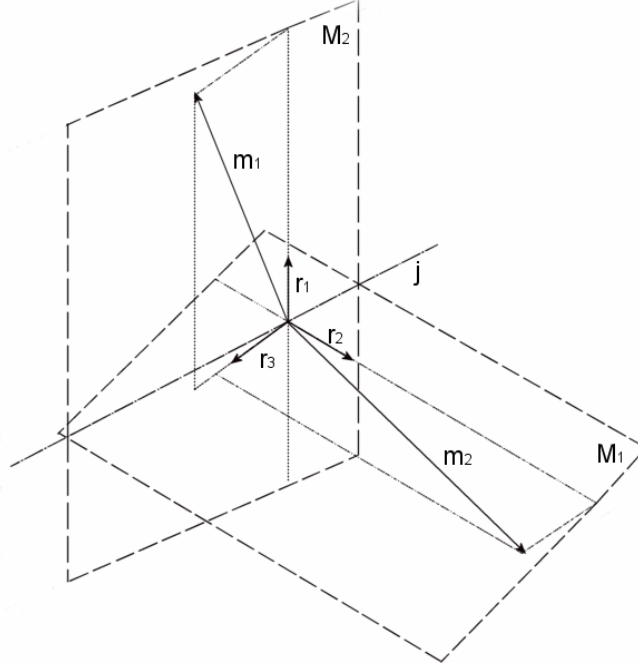
$$\begin{bmatrix} w \cdot u \\ w \cdot v \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (2.23)$$

Esta situación se produce cuando la distancia focal es corta y la cámara está situada físicamente muy alejada de la misma en la dirección del eje óptico. En estos casos se puede utilizar el modelo lineal aproximado de la cámara.

2.5. Restricciones en la matriz de proyección

Evidentemente, cualquier matriz de dimensión 3×4 no puede ser escrita en la forma de la matriz (2.16) ó (2.18) ya que ésta depende de los once parámetros intrínsecos y extrínsecos que modelan la cámara. En este apartado se realiza un estudio de la matriz de proyección \mathbf{M} desde un punto de vista geométrico. Esta interpretación geométrica será útil para demostrar las fuentes de errores que se producen en el proceso de calibración de la cámara y para obtener los parámetros del modelo de la cámara a partir de la misma.

Con el desarrollo de la expresión (2.16) se obtiene la forma general de la matriz de proyección \mathbf{M} , la cual depende de cuatro parámetros intrínsecos y seis extrínsecos de la siguiente forma,



2.5 Situación geométrica de los vectores que definen la matriz de proyección

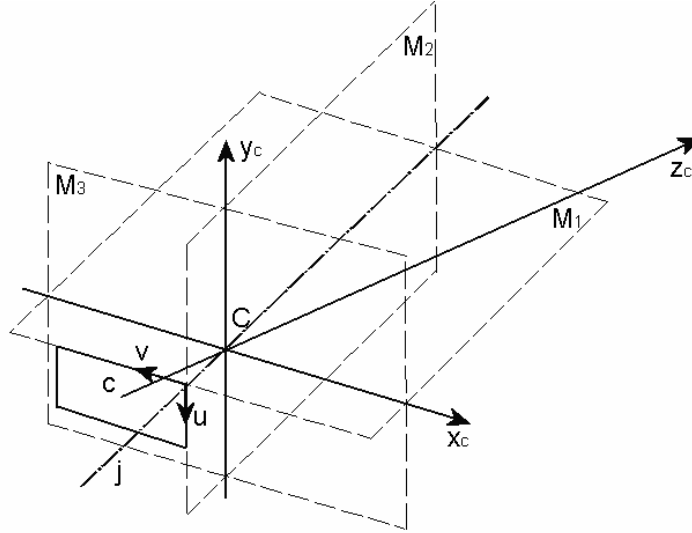
$$M = \begin{bmatrix} \alpha_u r_1 + u_0 r_3 & \alpha_u t_x + u_0 t_z \\ \alpha_v r_2 + v_0 r_3 & \alpha_v t_y + v_0 t_z \\ r_3 & t_z \end{bmatrix} \quad (2.24)$$

$\alpha_u = f \cdot k_u$ y $\alpha_v = f \cdot k_v$ representan el escalado de milímetros a píxeles, u_0 , v_0 son las coordenadas del centro de la imagen. Los vectores \mathbf{r}_1 , \mathbf{r}_2 , \mathbf{r}_3 son las filas de la matriz de rotación \mathbf{R} y $\mathbf{t} = (t_x, t_y, t_z)$ representa la traslación desde el origen de coordenadas del escenario al origen de coordenadas de la cámara. Para no complicar la interpretación geométrica se ha supuesto que los ejes de coordenadas del sistema de referencia de la imagen son ortogonales, es decir $\theta = \pi/2$.

La expresión (2.24) representa una matriz de proyección \mathbf{M} que se puede descomponer de la siguiente forma:

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \end{bmatrix} = \begin{bmatrix} \vec{m}_1 & m_{14} \\ \vec{m}_2 & m_{24} \\ \vec{m}_3 & m_{34} \end{bmatrix} = \begin{bmatrix} \alpha_u \vec{r}_1 + u_0 \vec{r}_3 & \alpha_u x t_x + u_0 t_z \\ \alpha_v \vec{r}_2 + v_0 \vec{r}_3 & \alpha_v t_y + v_0 t_z \\ \vec{r}_3 & t_z \end{bmatrix} \quad (2.25)$$

M_i son cada una de las filas de la matriz \mathbf{M} de forma que $M_i = (\mathbf{m}_i, m_{i4})$. Cada uno estos vectores forma el plano proyectivo $M_i: \mathbf{p} = 0$ siendo \mathbf{p} las coordenadas homogéneas de un punto respecto al sistema de referencia del escenario. \mathbf{m}_i es el vector ortogonal al plano M_i y m_{i4} marca la distancia del plano al origen de coordenadas del sistema de referencia del escenario \mathbf{o} [69]. Las componentes de los vectores \mathbf{m}_1 y \mathbf{m}_2 son una combinación lineal de los vectores \mathbf{r}_1 , \mathbf{r}_2 , \mathbf{r}_3 . Esta combinación lineal viene determinada por los parámetros intrínsecos α_u , α_v , u_0 y v_0 . El vector \mathbf{m}_3 corresponde directamente con el vector \mathbf{r}_3 . La figura 2.5 muestra la disposición geométrica de



2.6 Interpretación geométrica del modelo lineal de la cámara

cada uno de estos vectores junto con los planos que representan. Hay que tener en cuenta que esta representación no está a escala de lo que sería una cámara real. En una cámara real los vectores $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ tienen modulo unitario, los parámetros α_u, α_v son del orden de 1000 y u_0, v_0 son del orden de 250.

Por otro lado, el plano proyectivo $M_3 \cdot \mathbf{p} = 0$ corresponde con todos los puntos del plano de la imagen cuya coordenada homogénea vale cero, es decir $w = 0$ (2.7). Estos puntos son los llamados puntos en el infinito que corresponden con aquellos puntos del escenario que tienen $z_c = 0$ respecto al sistema de coordenadas de la cámara. Por lo tanto el plano $M_3 \cdot \mathbf{p} = 0$ corresponde el plano focal de la cámara F indicado en la figura 2.3. El plano con ecuación $M_1 \cdot \mathbf{p} = 0$ corresponde con todos aquellos puntos que en el plano de la retina tiene la coordenada $u = 0$ y el plano $M_2 \cdot \mathbf{p} = 0$ con aquellos que tienen la coordenada $v = 0$. La intersección de estos dos planos es la línea j , indicada en las figuras 2.5 y 2.6. Esta línea es la que une el origen de coordenadas del sistema de referencia de la imagen con el origen de coordenadas del sistema de medidas de la cámara \mathbf{o}' . En general, no coincide con el eje óptico de la cámara. De esta interpretación geométrica de la matriz de proyección \mathbf{M} se pueden obtener algunas conclusiones. De acuerdo con la figura 2.6 el centro óptico u origen de coordenadas del sistema de referencia de la cámara \mathbf{o}' es el punto de intersección de los tres planos $\mathbf{M} \cdot \mathbf{o}' = 0$. La posición de este punto respecto al sistema de referencia del escenario coincide con la posición de la cámara en el mismo. Si se expresa $\mathbf{M} = [\mathbf{M}' \ \mathbf{m}]$, \mathbf{o}' se puede obtener como,

$$\mathbf{o}' = -\mathbf{M}'^{-1} \mathbf{m} \quad (2.26)$$

Otra conclusión interesante es la ecuación de la línea que forman el centro óptico \mathbf{o}' con un punto \mathbf{q} de la imagen, figura 2.3. Esta línea es el rayo óptico definido por el punto \mathbf{q} y el

centro óptico \mathbf{o}' . De la conclusión anterior se obtiene la posición del punto \mathbf{o}' . Otro punto de esta línea es un punto en el infinito \mathbf{d} cuyas coordenadas homogéneas son $\mathbf{d}=(d_1, d_2, d_3, 0)$. Este punto satisface a la ecuación:

$$\mathbf{d} = M^{-1} \cdot \mathbf{q} \quad (2.27)$$

El punto \mathbf{q} se expresa en coordenadas homogéneas. \mathbf{d} es paralelo al vector \mathbf{m}_i que pasa por el centro de proyección \mathbf{o}' . Un punto \mathbf{p} en esta línea se puede expresar como $\mathbf{p}=\mathbf{M}^T \lambda \mathbf{p}$ donde λ varía entre $\pm\infty$.

Finalmente, de la interpretación geométrica se pueden deducir las restricciones que deben de cumplir los elementos de la matriz de proyección \mathbf{M} . Dado que el vector \mathbf{m}_3 corresponde con el vector \mathbf{r}_3 y éste forma parte de la matriz de rotación \mathbf{R} en la cual todos los vectores tienen módulo unitario, se deduce que el módulo de \mathbf{m}_3 también será unitario.

$$|\mathbf{m}_3| = 1 \quad (2.28)$$

Además, los vectores $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ son ortogonales entre sí, tal y como se puede ver en la figura 2.5. Los vectores \mathbf{m}_1 y \mathbf{m}_3 son ortogonales con el vector \mathbf{r}_2 , y también a su vez, \mathbf{m}_2 y \mathbf{m}_3 son ortogonales con el vector \mathbf{r}_1 . Por lo tanto, el vector resultado del producto vectorial de \mathbf{m}_1 y \mathbf{m}_3 será ortogonal al producto vectorial de \mathbf{m}_2 y \mathbf{m}_3 . Formalmente se puede expresar como,

$$(\mathbf{m}_1 \wedge \mathbf{m}_3) \cdot (\mathbf{m}_2 \wedge \mathbf{m}_3) = 0 \quad (2.29)$$

Desarrollando esta expresión se obtiene la restricción en función de los parámetros del modelo de la cámara,

$$(\mathbf{m}_1 \wedge \mathbf{m}_3) \cdot (\mathbf{m}_2 \wedge \mathbf{m}_3) = (\alpha_u r_1 + u_0 r_3) \wedge r_3 \cdot (\alpha_v r_2 + v_0 r_3) \wedge r_3 = (\alpha_u r_1 \wedge r_3) \cdot (\alpha_v r_2 \wedge r_3) = 0 \quad (2.30)$$

2.6. Obtención de los parámetros del modelo lineal

La matriz de transformación proyectiva \mathbf{M} (2.7) que conforma el modelo lineal de la cámara, puede ser obtenida de diferentes formas según (2.16). El problema inverso de calcular los once parámetros físicos de la cámara junto con el factor de escala λ a partir de una matriz de transformación proyectiva \mathbf{M} dada fue resuelto completamente en la década de los 90 [24]. En este apartado se describen varios métodos para obtener dichos parámetros a partir de una matriz de proyección \mathbf{M} dada. Algunos de ellos se basan en la interpretación geométrica realizada en el apartado anterior.

El cálculo del factor de escala λ es común a los tres métodos y se basa en la restricción $|\mathbf{m}_3|=1$. Teniendo esto en cuenta

$$\lambda = \pm \sqrt{\mathbf{m}_3 \cdot \mathbf{m}_3^T} \quad (2.31)$$

Ecuaciones explícitas

Si la matriz de proyección cumple con las dos restricciones anteriores, es posible obtener los parámetros que modelan la cámara con las expresiones dadas por Faugeras [24]. Dado que M está definida con un factor de escala λ el cual puede ser positivo o negativo $\lambda=\pm 1$ por la condición $|\mathbf{m}_3|=1$, existen diferentes soluciones para los parámetros de la cámara que satisfacen la matriz de proyección M dada. Se supone que el ángulo θ que forman los ejes de coordenadas de la imagen es ortogonal $\theta=\pi/2$. La matriz de proyección M se puede descomponer de la siguiente forma, igual que se hizo para la interpretación geométrica.

$$M = \begin{bmatrix} \alpha_u r_1 + u_0 r_3 & \alpha_u t_x + u_0 t_z \\ \alpha_v r_2 + v_0 r_3 & \alpha_v t_y + v_0 t_z \\ r_3 & t_z \end{bmatrix} = \begin{bmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{bmatrix} \quad (2.32)$$

Por lo tanto

$$t_z = \pm m_{34} r_3 = \pm m_3 \quad (2.33)$$

A partir de la descomposición realizada para la matriz de proyección se observa que \mathbf{m}_1 es la suma de dos vectores ortonormales \mathbf{r}_1 y \mathbf{r}_3 según

$$\mathbf{m}_1 = \alpha_u \mathbf{r}_1 + u_0 \mathbf{r}_3 \quad (2.34)$$

Utilizando esta expresión se tiene que,

$$|u_0 r_3| = u_0 = |\mathbf{m}_1| \text{sen} \alpha \quad (2.35)$$

α es el ángulo que forman los vectores \mathbf{m}_1 y \mathbf{m}_3 . El seno del ángulo que forman dos vectores se puede obtener mediante el producto escalar de ambos. Si se tiene en cuenta que $\mathbf{m}_3 = \pm \mathbf{r}_3$ y que el módulo de \mathbf{r}_3 vale 1 ya que es uno de los vectores que forman la matriz de rotación se obtiene la siguiente expresión para el producto escalar de \mathbf{m}_1 y \mathbf{m}_3 .

$$\mathbf{m}_1 \cdot \mathbf{m}_3 = |\mathbf{m}_1| \cdot |\mathbf{m}_3| \text{sen} \alpha = |\mathbf{m}_1| \text{sen} \alpha \quad (2.36)$$

Con las dos expresiones anteriores se deduce el valor de u_0 .

$$u_0 = (\mathbf{m}_1 \cdot \mathbf{m}_3) \quad (2.37)$$

De la misma forma es posible deducir el valor de v_0 .

$$v_0 = (\mathbf{m}_2 \cdot \mathbf{m}_3) \quad (2.38)$$

Utilizando la propiedad del producto escalar se puede obtener los valores de α_u y α_v con el mismo procedimiento realizado para u_0 y v_0 .

$$\alpha_u = \pm |\mathbf{m}_1 \wedge \mathbf{m}_3| \quad \alpha_v = \pm |\mathbf{m}_2 \wedge \mathbf{m}_3| \quad (2.39)$$

O también

$$\alpha_u = \pm\sqrt{\mathbf{m}_1 \cdot \mathbf{m}_1^T - u_0} \quad \alpha_v = \pm\sqrt{\mathbf{m}_2 \cdot \mathbf{m}_2^T - v_0} \quad (2.40)$$

Habiendo obtenido los parámetros intrínsecos α_u , α_v , u_0 y v_0 es posible calcular los parámetros extrínsecos \mathbf{r}_1 , \mathbf{r}_2 , t_x y t_y de la siguiente forma:

$$\begin{aligned} r_1 &= \pm \frac{1}{\alpha_u} (\mathbf{m}_1 - U_0 \mathbf{m}_3) & r_2 &= \pm \frac{1}{\alpha_v} (\mathbf{m}_2 - V_0 \mathbf{m}_3) \\ t_x &= \pm \frac{1}{\alpha_u} (\mathbf{m}_{14} - U_0 \mathbf{m}_{34}) & t_y &= \pm \frac{1}{\alpha_v} (\mathbf{m}_{24} - V_0 \mathbf{m}_{34}) \end{aligned} \quad (2.41)$$

Los parámetros α_u y α_v son siempre diferentes de cero ya que el rango de la matriz \mathbf{M} vale 3. Es fácil de verificar que $|\mathbf{r}_i|=1$ para $i=1,2,3$ y que $\mathbf{r}_i^T \cdot \mathbf{r}_3=0$ para $i=1,2$, dado que son las condiciones que cumple cualquier matriz de rotación. Sin embargo, en general se cumplirá que $\mathbf{r}_1^T \cdot \mathbf{r}_2 \neq 0$. El valor que toma la multiplicación escalar de $\mathbf{r}_1^T \cdot \mathbf{r}_2$ es:

$$r_1^T r_2 = \frac{(\mathbf{m}_3 \wedge (\mathbf{m}_1 \wedge \mathbf{m}_3)) \cdot (\mathbf{m}_3 \wedge (\mathbf{m}_2 \wedge \mathbf{m}_3))}{\|\mathbf{m}_1 \wedge \mathbf{m}_3\| \cdot \|\mathbf{m}_2 \wedge \mathbf{m}_3\|} \quad (2.42)$$

Usando propiedades de determinantes y teniendo en cuenta que $|\mathbf{m}_3|=1$ se puede simplificar esta expresión como:

$$r_1^T r_2 = \pm \frac{(\mathbf{m}_1 \wedge \mathbf{m}_3) \cdot (\mathbf{m}_2 \wedge \mathbf{m}_3)}{|\mathbf{m}_1 \wedge \mathbf{m}_3| \cdot |\mathbf{m}_2 \wedge \mathbf{m}_3|} \quad (2.43)$$

El término $\mathbf{r}_1^T \cdot \mathbf{r}_2$ es el coseno del ángulo entre los ejes de coordenadas de la imagen \mathbf{u}' , \mathbf{v}' . Si la restricción $(\mathbf{m}_1 \wedge \mathbf{m}_3) \cdot (\mathbf{m}_2 \wedge \mathbf{m}_3) = 0$ es satisfecha, existen cuatro series posibles de soluciones, las cuales corresponden a la posición del origen de coordenadas del mundo 3D respecto del de la cámara. Éstas son si está delante de la cámara $t_z > 0$, o detrás $t_z < 0$ y la elección de los ejes de coordenadas u' y v' en la imagen [24].

En los parámetros obtenidos hasta ahora se asume que los ejes de coordenadas en el plano de la imagen son ortogonales, $\theta = \pi/2$. En el caso de que esto no sea así, el ángulo θ que forman estos dos ejes del plano de la imagen es un quinto parámetro intrínseco a añadir a los cuatro que ya tenía el modelo anterior. Si se pretende obtener los parámetros de la cámara, incluyendo este ángulo θ se parte de la matriz de proyección \mathbf{M} general extendida:

$$M = \begin{bmatrix} \alpha_u r_1 + (\alpha_v \cot \theta) r_2 + u_0' r_3 & \alpha_u t_x + (\alpha_v \cot \theta) t_y + u_0' t_z \\ (\alpha_v / \sin \theta) r_2 + v_0' r_3 & (\alpha_v / \sin \theta) t_y + v_0' t_z \\ r_3 & t_z \end{bmatrix} \quad (2.44)$$

En este caso también hay que tener en cuenta que \mathbf{M} está definida con un factor de escala λ el cual puede ser positivo o negativo $\lambda = \pm 1$ por la condición $|\mathbf{m}_3|=1$. Igual que antes, se nombran diferentes vectores dentro de la matriz de proyección de la siguiente forma:

$$\begin{bmatrix} \alpha_u r_1 + (\alpha_v \cot \theta) r_2 + U'_0 r_3 & \alpha_u t_x + (\alpha_v \cot \theta) t_y + U'_0 t_z \\ (\alpha_v / \sin \theta) r_2 + V'_0 r_3 & (\alpha_v / \sin \theta) t_y + V'_0 t_z \\ r_3 & t_z \end{bmatrix} = \begin{bmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{bmatrix} \quad (2.45)$$

Con el producto escalar de \mathbf{m}_3 con \mathbf{m}_1 y \mathbf{m}_2 se puede obtener u_0 y v_0 respectivamente.

$$\begin{aligned} u_0' &= \mathbf{m}_1 \cdot \mathbf{m}_3 \\ v_0' &= \mathbf{m}_2 \cdot \mathbf{m}_3 \end{aligned} \quad (2.46)$$

Efectuando el producto vectorial de \mathbf{m}_2 y \mathbf{m}_3 se obtiene:

$$\frac{\alpha_v}{\sin \theta} = |\mathbf{m}_2 \wedge \mathbf{m}_3| \quad (2.47)$$

Por tanto, r_2 será:

$$r_2 = \frac{\mathbf{m}_2 - (\mathbf{m}_2 \cdot \mathbf{m}_3) \mathbf{m}_3}{|\mathbf{m}_2 \wedge \mathbf{m}_3|} \quad (2.48)$$

El producto escalar $\mathbf{m}_1 \cdot \mathbf{m}_2$ permite obtener el coseno entre los ángulos de los ejes

$$\cos \theta = \frac{(\mathbf{m}_1 \cdot \mathbf{m}_2) - (\mathbf{m}_1 \cdot \mathbf{m}_3)(\mathbf{m}_2 \cdot \mathbf{m}_3)}{|\mathbf{m}_2 \wedge \mathbf{m}_3|^2} \quad (2.49)$$

Con los productos vectoriales de $\mathbf{m}_1 \wedge \mathbf{m}_3$ y $\mathbf{m}_1 \wedge \mathbf{m}_2$ se obtiene

$$\begin{aligned} \alpha_u &= \frac{\|\cos \theta (\mathbf{m}_1 \wedge \mathbf{m}_2) - (\cos \theta \cdot V'_0 - U'_0) (\mathbf{m}_1 \wedge \mathbf{m}_3)\|}{\frac{\alpha_v}{\sin \theta} \cos \theta p_3 - U'_0 r_2} \\ r_1 &= \frac{1}{\alpha_u} \left(\mathbf{m}_1 - U'_0 \mathbf{m}_3 - \frac{\alpha_v}{\sin \theta} \cos \theta r_2 \right) \end{aligned} \quad (2.50)$$

El vector de traslación se obtiene según las siguientes expresiones:

$$\begin{aligned} t_z &= m_{34} \\ t_y &= \frac{\alpha_v}{\sin \theta} (m_{24} - V'_0 m_{34}) \\ t_x &= \frac{1}{\alpha_u} \left(m_{14} - U'_0 m_{34} - \frac{\alpha_v}{\sin \theta} \cos \theta t_y \right) \end{aligned} \quad (2.51)$$

Descomposición de Cholesky

Seedahmed [84] plantea un método de obtención de los parámetros del modelo de la cámara a partir de la matriz de proyección basado en la factorización de Cholesky. Mediante la descomposición de Cholesky es posible obtener los parámetros intrínsecos de la cámara. Estos incluyen tanto las coordenadas del punto principal de la imagen u_0 y v_0 , el factor de escala en cada uno de los ejes de coordenadas α_u y α_v , así como el ángulo θ que existe entre los dos ejes de la imagen. Una vez obtenida la matriz \mathbf{K} , siendo ésta $\mathbf{K} = \lambda \mathbf{F}$ se obtienen los parámetros

extrínsecos de la cámara. Se parte del modelo de la cámara que tiene en cuenta la pérdida de ortogonalidad que existe en los ejes de coordenadas de la imagen. Éste cumple la restricción $|\mathbf{m}_3|=1$. A partir de la expresión (2.16) se obtiene el siguiente modelo de la cámara:

$$M = \lambda \cdot F \cdot R \cdot T = K \cdot R \cdot T$$

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} = \begin{bmatrix} \alpha_u & \alpha_v \cot \theta & u_0' \\ 0 & \alpha_v / \sin \theta & v_0' \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{bmatrix} \quad (2.52)$$

De la expresión anterior se puede extraer la matriz D de la siguiente forma

$$D = K \cdot R \quad (2.53)$$

Donde se cumple que:

$$D = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \quad -K \cdot R \mathbf{t} = d \quad (2.54)$$

d es igual a:

$$d = \begin{bmatrix} m_{14} \\ m_{24} \\ m_{34} \end{bmatrix} \quad (2.55)$$

\mathbf{t} es el vector que define la traslación de la cámara $\mathbf{t}=(t_x, t_y, t_z)$. De la expresión (2.53) se puede escribir

$$DD^T = (K \cdot R)(K \cdot R)^T \quad (2.56)$$

Dado que cualquier matriz de rotación es ortonormal se tiene que $\mathbf{R} \cdot \mathbf{R}^T = \mathbf{I}$ y por lo tanto:

$$DD^T = KK^T \quad (2.57)$$

Se trata de una forma cuadrática se obtiene como resultado una matriz simétrica definida positiva. Esto provoca que con la descomposición directa en factores Cholesky no se obtenga la verdadera estructura de la matriz en términos ordenados de mayor a menor. Para obtener la verdadera estructura de la matriz es necesario realizar la inversa de la misma y aplicar el método de factorización de Cholesky. Después, para obtener la matriz \mathbf{K} con los parámetros intrínsecos de la cámara, es necesario volver a invertir el resultado de la factorización. Con estos tres pasos (invertir DD^T , realizar la descomposición de Cholesky y volver a realizar la inversa del resultado) se obtiene una matriz \mathbf{K} , la cual su elemento k_{33} es diferente de 1. Por lo tanto, será necesario dividir la matriz obtenida por k_{33} .

$$K = \frac{1}{k_{33}} \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & k_{33} \end{bmatrix} \quad (2.58)$$

Una vez obtenida la matriz \mathbf{K} que contiene los parámetros intrínsecos de la cámara, es posible obtener la matriz \mathbf{R} y el vector \mathbf{t} que contienen los parámetros extrínsecos de la cámara a partir de las expresiones (2.53) y (2.54).

$$\mathbf{R} = \mathbf{K}^{-1} \mathbf{D} \quad \mathbf{t} = -(\mathbf{K} \cdot \mathbf{R})^{-1} \mathbf{d} = -\mathbf{D}^{-1} \mathbf{d} \quad (2.59)$$

Descomposición RQ

Melen [63] propone un método para la extracción de los parámetros de la cámara de la matriz de proyección basado en la descomposición RQ. En este caso se buscan los elementos que forman el modelo 2 (2.15). Se parte del modelo de la cámara (2.18) el cual cumple la restricción $|\mathbf{m}_3|=1$

$$\mathbf{M} = \mathbf{V}^{-1} \cdot \mathbf{B}^{-1} \cdot \mathbf{F} \cdot \mathbf{R} \cdot \mathbf{T} \quad (2.60)$$

Si se define

$$\mathbf{D} = \mathbf{V}^{-1} \mathbf{B}^{-1} \mathbf{K} \cdot \mathbf{R} = \mathbf{G} \cdot \mathbf{R} \quad (2.61)$$

En este caso se tiene:

$$\mathbf{D} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (2.62)$$

$$\mathbf{G} = \begin{bmatrix} \frac{-(1-b_1)f}{1-b_1^2-b_2^2} & -\frac{b_2}{1-b_1^2-b_2^2} & -u_0 \\ -\frac{b_2}{1-b_1^2-b_2^2} & \frac{-(1+b_1)f}{1-b_1^2-b_2^2} & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{23} & g_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

Las matrices \mathbf{D} y \mathbf{G} cumplen las siguientes restricciones dado que $b_1^2 + b_2^2 < 1$ y $f > 0$.

$$g_{12} = g_{21} \quad g_{11} > 0 \quad g_{21} > 0 \quad (2.63)$$

Realizando la descomposición RQ de la matriz \mathbf{D} se tiene como resultado una matriz triangular superior y una matriz ortogonal. La matriz ortogonal \mathbf{R} corresponde con la matriz de rotación.

$$\mathbf{D} = \mathbf{G} \cdot \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & \pm 1 \end{bmatrix} \cdot \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \quad (2.64)$$

El siguiente paso es encontrar una matriz S que haga que:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & \pm 1 \end{bmatrix} \cdot \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{23} & g_{23} \\ 0 & 0 & 1 \end{bmatrix} \quad (2.65)$$

Claramente se observa que $s_{31}=0$ y $s_{32}=0$. Para que la matriz S sea ortogonal, ésta debe ser una matriz de rotación elemental

$$S = \begin{bmatrix} \cos \mathcal{G} & \sin \mathcal{G} & 0 \\ -\sin \mathcal{G} & \cos \mathcal{G} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.66)$$

Con la restricción $g_{12}=g_{21}$ se puede determinar el ángulo de rotación como:

$$r_{11} \sin \mathcal{G} + r_{12} \cos \mathcal{G} = -r_{22} \sin \mathcal{G} \quad \tan \mathcal{G} = -\frac{r_{12}}{r_{11} + r_{22}} \quad (2.67)$$

En este caso se tiene que:

$$\sin \mathcal{G} = -\frac{\tan \mathcal{G}}{\sqrt{1 + \tan^2 \mathcal{G}}} \quad \cos \mathcal{G} = -\frac{1}{\sqrt{1 + \tan^2 \mathcal{G}}} \quad (2.68)$$

Una vez obtenida la matriz G es posible obtener los parámetros intrínsecos del modelo 2 de la cámara según:

$$b_1 = -\frac{g_{11} - g_{22}}{g_{11} + g_{22}} \quad b_2 = -2\frac{g_{12}}{g_{11} + g_{22}} \quad f = 2\frac{g_{11}g_{22} - g_{12}^2}{g_{11} + g_{22}} \quad (2.69)$$

La matriz R que define los parámetros extrínsecos de rotación de la cámara se obtuvieron en el paso de la descomposición RQ de la matriz D . El cálculo de vector t de traslación de la cámara se realiza con la expresión (2.59) utilizada en la descomposición de Cholesky.

2.7. Homografías

Un caso particular de matriz de proyección es aquel que los puntos en la escena están distribuidos en un plano, es decir, las coordenadas de los puntos son bidimensionales. En este caso esta matriz de transformación que establece la relación entre el plano de la imagen y el plano de la plantilla se llama homografía. Formalmente se puede decir que dado un conjunto de puntos p_i en el espacio proyectivo P^2 y sus correspondientes q_i en P^2 , existe una transformación proyectiva llamada homografía H que relaciona cada punto de un plano con su correspondiente del otro plano. Esto representa el caso práctico de tomar una imagen de una plantilla plana en las cuales tanto los puntos en la imagen como los de la plantilla están contenidos dentro de un

espacio proyectivo P^2 . También podría representar el caso de la relación existente entre dos puntos correspondientes de dos imágenes distintas, considerando cada imagen como un plano proyectivo P^2 . La relación entre los dos puntos está representada por la homografía \mathbf{H} . La transformación está expresada mediante la ecuación

$$\mathbf{q}_i = \mathbf{H} \cdot \mathbf{p}_i \quad (2.70)$$

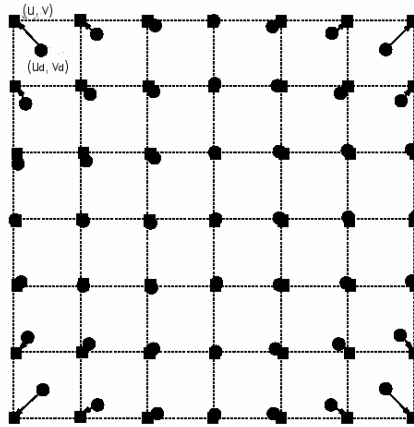
La matriz \mathbf{H} es de 3x3 elementos que constituyen la homografía. Partiendo de la matriz de proyección \mathbf{M} (2.7) y asumiendo sin pérdida de generalidad que la coordenada $z_w=0$ en el caso de tener una plantilla bidimensional, se puede obtener la siguiente expresión

$$\begin{bmatrix} w \cdot u_p \\ w \cdot v_p \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \quad (2.71)$$

Esta expresión relaciona las coordenadas homogéneas bidimensionales de los puntos en la escena $\mathbf{p}_w=(x_w, y_w, 1)$, con sus correspondientes coordenadas homogéneas bidimensionales del punto en la imagen $\mathbf{q}_p=(u_p, v_p, 1)$. Los elementos de la homografía \mathbf{H} son idénticos a las columnas primera, segunda y cuarta de la matriz de proyección \mathbf{M} . La homografía \mathbf{H} al igual que la matriz de proyección \mathbf{M} contiene los parámetros intrínsecos y extrínsecos de la cámara. A partir de un conjunto de parámetros se puede generar una matriz de proyección \mathbf{M} y eliminando la tercera columna, se obtiene la correspondiente homografía \mathbf{H} . Sin embargo, dada una homografía \mathbf{H} no es posible extraer los parámetros intrínsecos y extrínsecos que definen las características de la cámara y su posición en el mundo.

2.8. El modelo no lineal de la cámara

Todo lo visto hasta ahora presenta un punto de vista ideal del proceso de formación de la imagen. La obtención de una imagen real se ve afectada por perturbaciones producidas tanto por imprecisiones mecánicas de la cámara así como del proceso de captura y muestreo de la imagen. Las imperfecciones mecánicas tienen que ver con el proceso de fabricación y ensamblaje de las lentes. Estos defectos provocan que el camino que recorre la luz hasta que llega al plano de la imagen no sea una línea recta. Por lo tanto el modelo descrito hasta ahora no es del todo válido ya que se había supuesto una perfecta proyección perspectiva en la cámara. Estas distorsiones son necesarias tenerlas en cuenta en el modelo ya que pueden afectar significativamente a los resultados. Diferentes autores han estudiado las fuentes y el efecto de estas distorsiones en la imagen [89], [100]. Respecto a las distorsiones producidas en el proceso



2.7 Efecto de la distorsión radial y tangencial en la imagen

de captura y muestreo de la imagen se pueden considerar despreciables debido a la evolución de los sistemas de visión [85].

Las imperfecciones en las formas de las lentes provocan desviaciones laterales del haz de luz, que las atraviesa. Como resultado se obtiene una posición del punto observado en la imagen diferente a su posición real como reflejo de un punto del espacio. La desviación de la posición es radial respecto a un centro de distorsión. En el caso de las imperfecciones debidas al ensamblaje de las lentes, se generan tanto distorsiones radiales como tangenciales de las posiciones de los puntos en la imagen. Las consecuencias son desplazamientos geométricos de los puntos en la imagen. El efecto visual se muestra en la figura 2.7.

La posición correcta en píxeles del punto \mathbf{q} en la imagen $\mathbf{q}_p = (u_p, v_p)$ está relacionada con la posición observada del mismo $\mathbf{q}_d = (u_d, v_d)$. Esta relación se expresa mediante las siguientes expresiones:

$$\begin{aligned} u_d &= u_p - \delta_u(u_p, v_p) \\ v_d &= v_p - \delta_v(u_p, v_p) \end{aligned} \quad (2.72)$$

Según (2.72) la cantidad de error geométrico en cada coordenada del punto \mathbf{q} depende la posición del mismo dentro de la imagen.

Distorsión radial:

La distorsión radial produce un desplazamiento de la posición del punto a lo largo de la línea que une el punto con el centro de distorsión radial de la imagen. Normalmente, éste no coincide con el centro de la imagen. Esta distorsión está causada por los defectos que existen en la curvatura de la lente. Un desplazamiento radial negativo de los puntos de la imagen es llamado también distorsión de barril. El efecto que se produce es que los puntos situados en los bordes de la imagen se acercan y la escala disminuya. Si el desplazamiento es positivo, se

conoce también como distorsión de cojín. En este caso los puntos del borde de la imagen se alejan y la escala aumenta. Esta distorsión es simétrica respecto al eje óptico de la cámara. Si se asume que el centro de distorsión se encuentra en el centro de la imagen, la expresión matemática que representa esta distorsión es la siguiente [21], [61]:

$$\begin{aligned}\delta_{ur}(u_p, v_p) &= \Delta u \cdot (k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6 + \dots) \\ \delta_{vr}(u_p, v_p) &= \Delta v \cdot (k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6 + \dots)\end{aligned}\quad (2.73)$$

r es la distancia del píxel al punto principal de la imagen, $\mathbf{c}=(u_0, v_0)$, $\Delta u=u_p-u_0$, $\Delta v=v_p-u_0$. Los coeficientes de la distorsión radial son k_1, k_2, k_3, \dots . La distancia r se calcula como:

$$r^2 = \Delta u^2 + \Delta v^2 \quad (2.74)$$

Distorsión tangencial o descentrado de la imagen:

Los sistemas ópticos están sujetos a varios grados de descentrado debido a que los centros ópticos de las diferentes lentes no están situados en la misma recta. Este defecto provoca una distorsión llamada de descentrado de la imagen. Esta distorsión tiene tanto componente radial como tangencial, la cual se pueden describir mediante la siguiente expresión [12], [61]:

$$\begin{aligned}\delta_{ud}(u_p, v_p) &= p_1(r^2 + 2 \cdot \Delta u^2) + 2p_2 \cdot \Delta u \cdot \Delta v + \dots \\ \delta_{vd}(u_p, v_p) &= 2p_1 \cdot \Delta u \cdot \Delta v + p_2(r^2 + 2 \cdot \Delta v^2) + \dots\end{aligned}\quad (2.75)$$

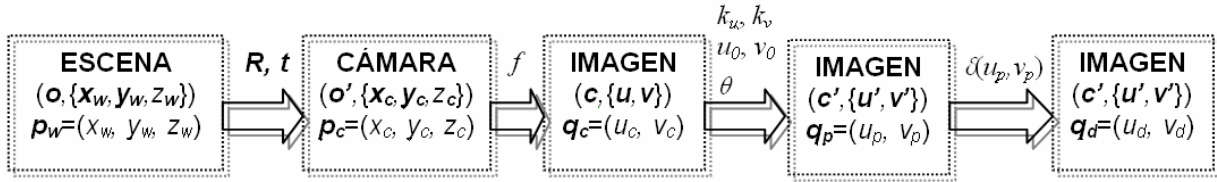
p_1, p_2, p_3, \dots son los coeficientes que modelan la distorsión tangencial y r es la distancia del píxel al punto principal de la imagen

Distorsión de Prisma:

La distorsión de prisma es debida a las imperfecciones en el diseño de las lentes, a su fabricación y ensamblado. Principalmente se trata de ligeros desplazamientos de algunas lentes que resultan en una falta de perpendicularidad con el eje óptico de la cámara. Esta distorsión se puede modelar añadiendo un pequeño prisma al sistema óptico de la cámara el cual causa una mayor distorsión radial y tangencial [12], [21]. En este caso la distorsión se modela según:

$$\begin{aligned}\delta_{up}(u_p, v_p) &= s_1 \cdot r^2 + s_2 \cdot r^4 + s_3 \cdot r^6 + \dots \\ \delta_{vp}(u_p, v_p) &= s_1 \cdot r^2 + s_2 \cdot r^4 + s_3 \cdot r^6 + \dots\end{aligned}\quad (2.76)$$

r es la distancia del píxel al punto principal de la imagen definida por (2.74) y s_1, s_2, s_3, \dots son los coeficientes que modelan este tipo de distorsión.



2.8 Transformaciones lineales y no lineales contenidas en el modelo de la cámara

Distorsión total:

La distorsión efectiva puede ser modelada por la suma de los tres tipos de distorsiones definidas anteriormente [12], [21]. La cantidad de distorsión total en un punto de la imagen lo dará la suma de las expresiones (2.73), (2.75), (2.76). Si se descartan los términos de orden superior a 3 se obtiene la siguiente expresión.

$$\begin{aligned} \delta_u(u_p, v_p) &= u_p \cdot k_1 \cdot r^2 + p_1(3u_p^2 + v_p^2) + 2p_2 \cdot u_p \cdot v_p + s_1 \cdot r^2 \\ \delta_v(u_p, v_p) &= v_p \cdot k_1 \cdot r^2 + 2p_1 \cdot u_p \cdot v_p + 2p_2(u_p^2 + 3v_p^2) + s_2 \cdot r^2 \end{aligned} \quad (2.77)$$

La figura 2.8 muestra las diferentes transformaciones que sufren las coordenadas de un punto respecto a los diferentes sistemas de coordenadas. La flecha indica hacia que sentido se produce el cambio de coordenadas asociadas a un sistema de referencia distinto.

2.9. Conclusiones

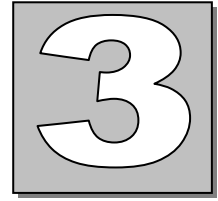
Para estudiar el comportamiento de la cámara y poder tratar la misma como una caja negra que realiza una transformación de las coordenadas del espacio proyectivo P^3 al espacio P^2 existen diferentes modelos los cuales se ajustan más o menos al comportamiento real de la misma. El modelo básico es el llamado pin-hole a partir del cual se construyen los demás modelos. Este modelo básico solamente transforma las coordenadas de los puntos en la escena. Esta transformación da la relación entre el sistema de referencia de la cámara y el del escenario. No tiene en cuenta las transformaciones que se producen en la formación de la imagen. Un modelo más ajustado al comportamiento real de la cámara añade al modelo pin-hole las transformaciones lineales que sufren las coordenadas de los puntos en la imagen. Éstas son el escalado que permite el cambio de unidades de píxeles a milímetros, la localización del origen de medidas en la esquina superior izquierda de la imagen y la no ortogonalidad de los ejes de medidas debida a imperfecciones en el proceso de construcción. El hecho de que los ejes de la imagen no formen 90 grados puede ser tratado de forma lineal según la expresión (2.12). El modelo más ajustado al comportamiento de la cámara es el que tiene en cuenta las distorsiones que se producen en el sistema de visión por las imperfecciones constructivas de las lentes. Estas imperfecciones producen desviaciones en la trayectoria del haz de luz que a priori se supone

recta. El efecto de estas desviaciones son distorsiones radiales y de descentrado en el CCD que recoge la imagen.

Las diferentes transformaciones que sufren las coordenadas de los puntos se ven reflejadas en la figura 2.8. Cada cuadro representa las coordenadas del punto dentro de cada sistema de referencia. Las transformaciones entre cuadro y cuadro están parametrizadas por los parámetros indicados en las flechas. Estos parámetros se clasifican en intrínsecos o extrínsecos según describan características de la cámara o su posición y orientación en la escena. Con un conjunto de parámetros dado es posible construir la matriz de transformación que los representa. Del mismo modo, a partir de una matriz de transformación es posible extraer los parámetros intrínsecos y extrínsecos que la generan.

En principio una matriz de proyección relaciona puntos en el espacio tridimensional situados en el escenario, con sus correspondientes puntos colocados en un espacio bidimensional que es la imagen. Un caso particular de matriz de proyección es la llamada homografía, la cual relaciona dos conjuntos de puntos situados en sendos espacios bidimensionales. Estos espacios bidimensionales pueden ser una plantilla plana relacionada con su imagen, o dos imágenes diferentes de la misma plantilla. En este caso, al contrario que ocurre con la matriz de transformación, de una homografía no se pueden extraer los parámetros intrínsecos y extrínsecos de la cámara. Sin embargo, con un conjunto de parámetros si que es posible generar una homografía.

Por otro lado, también es interesante tener en cuenta que el modelo de la cámara formado por una matriz de proyección homogénea permite tratar la proyección perspectiva como una transformación lineal.



EL PROCESO DE CALIBRADO

El proceso de calibrado de una cámara es un paso necesario para extraer información 3D de imágenes 2D. Mucho trabajo se ha realizado en este campo el cual consiste en estimar los parámetros intrínsecos y extrínsecos de la misma a partir de una o varias imágenes obtenidas de una plantilla. En este capítulo se presenta el estado del arte tanto de las diferentes técnicas de calibración, como de los diferentes aspectos que afectan al proceso de calibrado de una cámara. En primer lugar se nombran los métodos existentes en la actualidad para calibrar una cámara, estableciendo diferentes clasificaciones según técnicas, resultados y elementos necesarios para la calibración. También se presentan las diferentes peculiaridades que afectan al proceso de calibrado. Éstas pretenden decidir hasta que punto pueden afectar al resultado los errores en las medidas, las deficiencias en la elección del modelo o los errores en la asignación de puntos en el mundo con los de la imagen. Seguidamente, basándose en todos los métodos existentes se han elegido un proceso de calibración que a juicio del autor según diferentes estudios realizados por él mismo y por otros autores [76], [82], [42], resulta más interesante para realizar la calibración de una cámara bajo ciertas condiciones. Este proceso incluye desde la corrección de las medidas utilizando técnicas de normalización de datos o técnicas para la corrección de distorsiones hasta la definición del proceso de cálculo de los parámetros de la cámara. El objetivo es definir el método de calibración más completo que permita resolver el mayor número de situaciones posibles basándose en las técnicas de calibración existentes hasta la fecha. Una vez definido el método de calibración, se procederá a calcular intervalos en los resultados obtenidos y optimizarlo utilizando técnicas estadísticas. Esta parte se desarrolla en el siguiente capítulo.

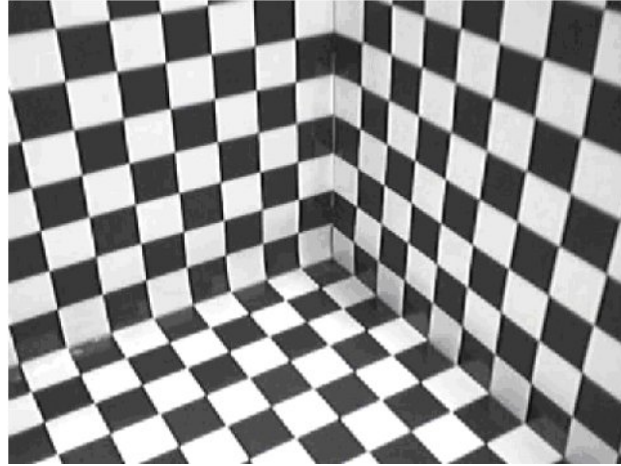
3.1. Métodos de calibración

El proceso de calibración de una cámara es el primer paso para resolver aplicaciones en las que sea necesario obtener datos cuantitativos de la imagen. Aunque es posible obtener información de la escena a partir de imágenes tomadas con cámaras sin calibrar [31], el proceso de calibración resulta esencial cuando se trata de obtener medidas de las mismas. La calibración precisa de la cámara permite obtener distancias en el mundo real a partir de las imágenes tomadas del mismo [23], [30]. Por ejemplo, desde un punto de vista de localización de objetos, es posible situarlos en el mundo real cuando se tiene una imagen de los mismos. Esta localización puede ser absoluta respecto a un origen de coordenadas del mundo o relativa respecto a cualquier otro objeto. Con esto es posible resolver aplicaciones industriales de ensamblado de piezas [15] o evitar obstáculos en la navegación de un robot [9], [16]. También, si la cámara está colocada en un robot, ya bien sea en un robot móvil o en un brazo mecánico, la posición y orientación de la cámara puede ser calculada localizando marcas conocidas en el mundo dentro de la imagen. Incluso, si estas medidas son almacenadas, es posible determinar la trayectoria del robot o del brazo mecánico. Esta información resulta muy útil a la hora de controlar el robot o realizar una planificación de trayectorias [48], [53], [54]. Si por el contrario nos centramos en la reconstrucción 3D de objetos, cada punto en la imagen determina un rayo óptico que pasa a través del centro óptico de cámara hacia la escena. El manejo de varias imágenes de una misma escena en la cual no exista movimiento, permite relacionar ambos rayos ópticos para obtener la posición de los puntos 3D en la escena [2], [8], [43]. En este caso es necesario resolver el paso previo de las correspondencias de un objeto en las diferentes imágenes [102]. Para tomar varias imágenes de la escena se pueden utilizar una única cámara en movimiento, varias cámaras montadas sobre un sistema estereoscópico o una fuente de luz estructurada. Una vez se ha podido realizar la reconstrucción 3D del objeto, éste se puede comparar con un modelo almacenado para determinar las imperfecciones del mismo fruto del proceso de fabricación. La inspección visual es una herramienta muy útil para el control de calidad, la cual permite examinar todos los productos de forma automática y exhaustiva, lo que implica una mejoría importante respecto a la inspección humana la cual requiere de herramientas estadísticas para su realización [68].

Una parte de la calibración de la cámara es la estimación de los parámetros intrínsecos de la misma los cuales modelan la geometría interna de la cámara y las características ópticas del sensor. Estos parámetros determinan las coordenadas de un punto en la imagen a partir de la posición del punto en la escena respecto del sistema de coordenadas de la cámara. También es necesario tener en cuenta los parámetros que miden la distorsión de la imagen producida por las imperfecciones constructivas de la cámara. Estos parámetros permiten corregir la posición de los

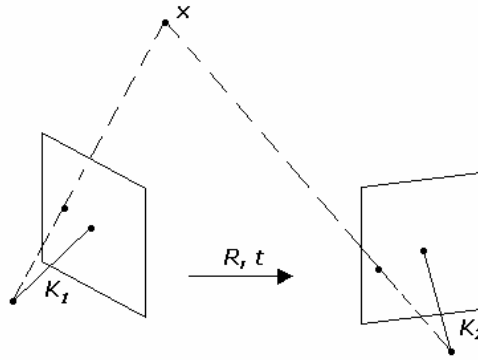
puntos en la imagen consiguiendo un comportamiento ideal de la cámara correspondiente al modelo pin-hole. La estimación de la relación geométrica entre la cámara y la escena o entre diferentes cámaras, es también importante dentro del proceso de calibración. Esta relación geométrica se caracteriza con los parámetros extrínsecos. Los parámetros extrínsecos miden la posición y la orientación de la cámara respecto al sistema de coordenadas establecido para el mundo. Estos dan la relación respecto al sistema de coordenadas del usuario en lugar del sistema de coordenadas de la cámara.

En la actualidad existen varios métodos para la calibración de una cámara. Estos métodos pueden clasificarse según diferentes criterios. Por ejemplo, atendiendo al método de resolución se pueden clasificar en lineales frente a no lineales o iterativos. Los métodos lineales utilizan métodos de resolución de sistemas de ecuaciones basados en mínimos cuadrados. Estos métodos obtienen una matriz de transformación que relaciona los puntos 3D en el mundo con sus proyecciones 2D en la imagen. En este caso no se calculan los parámetros que modelan la distorsión de la cámara por lo que los resultados obtenidos son bastante aproximados, sin embargo son fáciles de implementar y muy rápidos para su ejecución [1], [22], [29], [41], [90]. Si por el contrario se requiere un modelo de la cámara más complejo, en el cual se incluyan las distorsiones que produce la cámara es necesario minimizar índices no lineales de forma iterativa. El índice a minimizar suele incluir la distancia entre los puntos medidos en la imagen y los puntos proyectados obtenidos con el modelo de la cámara. La ventaja de estos métodos iterativos es que cualquier modelo puede ser calculado y además la exactitud del mismo aumenta con el número de iteraciones hasta que converge. Sin embargo, son mucho más lentos y necesitan partir de una buena aproximación de los parámetros para garantizar esta convergencia [79], [74]. Es por este motivo que se utilizan los resultados obtenidos por métodos lineales para iniciar la búsqueda no lineal de los parámetros. Con métodos lineales se calculan una parte del conjunto de parámetros y posteriormente utilizando métodos iterativos se mejoran estos parámetros y se estima el resto. Esta calibración realizada en dos pasos permite reducir considerablemente el número de iteraciones garantizando además la convergencia de la búsqueda iterativa de los parámetros [36], [94], [98], [100]. Otra clasificación de los métodos de calibración puede realizarse basándose en el resultado de la misma. Una calibración explícita obtiene directamente los parámetros del modelo de la cámara [7], [93], [103], mientras que en una implícita se obtienen matrices de transformación que contienen el conjunto de todos los parámetros. Aunque no se conoce el valor exacto de alguno de los parámetros, los resultados pueden ser utilizados para realizar medidas 3D y la generación de coordenadas en la imagen. Los métodos implícitos no son aptos para modelar la cámara ya que los parámetros que se obtienen no corresponden con los reales de la cámara [29], [97], [98]. Atendiendo a los parámetros que forman el modelo de la cámara, los métodos de calibración también se pueden clasificar en intrínsecos y extrínsecos. Los métodos de calibración intrínsecos sólo obtienen los parámetros físicos y ópticos de la cámara [52], [70]. Por el contrario, los extrínsecos calculan la posición y



3.1 Plantilla 3D para la calibración de la cámara

orientación de la cámara en la escena [57], [96]. Finalmente, atendiendo a las características de la plantilla que se utiliza para la calibración, existen métodos que utilizan plantillas 3D, 2D, 1D o no utilizan plantilla. Los métodos que utilizan plantillas de referencia basan la calibración de la cámara en establecer una relación entre las coordenadas conocidas de los puntos en la plantilla y las coordenadas de estos puntos en la imagen. En el caso de plantillas 3D con una sola imagen de la misma es posible realizar la calibración. En este caso la plantilla consiste en dos o tres planos ortogonales entre ellos. Como se asumen los puntos en planos se evitan errores de medidas de las coordenadas de puntos en la plantilla ya que se asume la misma para todos los puntos del mismo plano. Por el contrario este tipo de calibración requiere de una elaboración costosa de la plantilla para realizarse [22], [37], [44], [94]. En el caso de utilizar plantillas 2D es necesario tomar varias imágenes de la misma desde varias posiciones o cambiar la posición y orientación de la plantilla. No es necesario conocer las posiciones desde donde se toman las imágenes [58], [64], [81], [104]. Este método resulta más versátil ya que la elaboración de la plantilla se puede realizar fácilmente. Los métodos de calibración basados en plantillas 1D son muy útiles en el caso de calibrar sistemas con varias cámaras. En el caso de utilizar métodos basados en plantillas 3D ó 2D, dado que es necesario que todas ellas vean varios puntos de la plantilla de calibración a la vez, resulta complicado establecer una posición para la misma a no ser que la plantilla sea transparente. Es por este motivo que el método de calibración basado en una plantilla 1D resulta atractivo a la hora de calibrar un sistema con varias cámaras [105]. En algunos casos, los métodos basados en plantilla necesitan conocer la relación entre los planos para establecer más restricciones en el proceso de calibración y poder conseguir unos resultados más exactos. También el uso de propiedades geométricas de la escena del tipo líneas al infinito [95], o elementos como líneas rectas [17], [20] o circunferencias dentro de la misma [36], [64], permiten realizar la calibración sin necesidad de tener medidas de las coordenadas de los puntos en la plantilla. Por otro lado, las técnicas que no utilizan ningún objeto de calibración pueden ser consideradas como de plantilla 0D ya que sólo es necesario establecer correspondencias de un



3.2 Calibración de cámara sin conocimiento del escenario mediante auto calibración

punto en diferentes imágenes. Solamente moviendo la cámara en una escena estática, la rigidez de la escena provoca en general dos restricciones dentro de los parámetros intrínsecos de la cámara [59], [62]. Por lo tanto con varias imágenes de la misma escena tomadas con los mismos parámetros intrínsecos, las correspondencias entre tres imágenes son suficientes para calcular tanto los parámetros intrínsecos como extrínsecos. Esta regla permite a su vez realizar la reconstrucción 3D de una estructura a partir de varias imágenes de la misma. En estos casos aunque no es necesaria una plantilla, es necesario calcular un gran número de parámetros, resultando en un problema matemático bastante complejo. Debido a la dificultad para iniciar la búsqueda los métodos de auto calibración tienden a ser inestables [10], [26]. Para terminar, con el sumario de los diferentes métodos que existen para la calibración de cámaras, también es necesario tener en cuenta la familia de algoritmos que calculan los parámetros que modelan la distorsión producida por las lentes en la imagen sin el uso de objetos de calibración y por lo tanto sin conocer ninguna estructura 3D. Estos métodos se basan en que dentro de una proyección perspectiva ideal, la cámara transforma líneas rectas en el espacio 3D en líneas rectas dentro del espacio 2D correspondiente a la imagen. Por lo tanto reforzando la linealidad de las partes de la imagen que se ven curvas debido a la distorsión de la lente de la cámara, es posible estimar la deformación que ésta produce [3], [13], [19], [71], [83]. Hay métodos que utilizan restricciones epipolares y trilineales entre pares y tripletes de imágenes respectivamente para estimar la distorsión radial [86].

Desde el punto de vista de los errores cometidos en el proceso de calibrado de la cámara existen una serie de factores a tener en cuenta. Claramente, con datos más exactos es posible conseguir una calibración mejor. Sin embargo, resulta más efectivo saber que precisión se puede conseguir en los resultados de calibración con los escenarios de calibración y métodos que se disponen. Existen estudios sobre los errores cometidos en la calibración debidos a imprecisiones en las medidas del plano de la imagen o en la plantilla de calibración [49], [50], [51], [55], [103]. Además al proceso de calibración, no sólo afecta el ruido de las medidas, sino también la asociación errónea de un punto en el mundo con un punto en la imagen. Algunos

autores utilizan herramientas estadísticas para detectar esta situación [25], [40]. También se pueden tener en cuenta errores de cálculo debido a inestabilidades de la herramienta matemática [87]. Una normalización de los datos permite que el algoritmo sea menos sensible al ruido de las medidas y obtenga unos resultados más exactos [32]. Otro factor importante es el modelo de cámara que se pretende estimar. Un modelo muy complejo puede complicar mucho el algoritmo de búsqueda y los resultados obtenidos no mejorar significativamente frente a otro modelo más simple. Un modelo excesivamente complicado incluso puede producir inestabilidades en el proceso de búsqueda y llegar a resultados absurdos [100].

Basándose en el estado actual de los procesos de calibración descritos hasta ahora, resulta complicado elegir un método eficiente que permita calibrar la cámara en cualquier situación. El método de Tsai [94], representa un proceso clásico de calibración basado en las medidas de las coordenadas de los puntos de una plantilla 3D respecto a un punto de referencia fijo. Este método ha sido ampliamente utilizado en el siglo pasado [45], [73]. En la comparación de métodos de calibración desarrollados entre 1982 y 1998 realizado por Salvi [76], el método de Tsai muestra mejores resultados a pesar de que para conseguir buenos resultados es necesaria una precisión importante de los datos de entrada. Por el contrario, el método de Zhang [103], el cual no está incluido en la comparación de Salvi [76], representa una nueva era en el proceso de calibrado de la cámara. Este método utiliza las coordenadas de los puntos situados una plantilla plana 2D tomando diferentes imágenes de la misma desde diferentes posiciones y orientaciones. De esta forma se combinan las ventajas de los métodos de calibración basados en las medidas de las coordenadas de la plantilla con las ventajas de la auto calibración en la cual no es necesaria utilizar plantilla. Este modo de calibración resulta muy flexible desde el punto de vista de que tanto la cámara como la plantilla puede ser movida libremente y además se pueden tomar tantas imágenes como se quieran sin tener que volver a realizar medidas en la plantilla. Sun [82] compara el método de Tsai con el método de Zhang. Por un lado el método de Tsai obtiene una estimación precisa de los parámetros de la cámara si los datos de entrada están poco contaminados con ruido. Teniendo en cuenta que son necesarios al menos cien puntos en la plantilla y que las coordenadas han de estar referidas a un origen de coordenadas fijo, es imprescindible un adecuado diseño de la plantilla de calibración además de una medida exacta de las coordenadas de los puntos. A pesar de todo, la posibilidades de cometer errores en las medidas son altas tal y como confirman los experimentos realizados por Sun [82]. Por el contrario, el método de calibración de Zhang basado en plantilla 2D no requiere tal especial diseño de la plantilla, ni tampoco una medición tan exacta de los puntos de la misma. Sun realiza los experimentos con una plantilla hecha a mano y obtiene mejores resultados que con el método de Tsai. Además, la sensibilidad del algoritmo de calibración frente a errores en las medidas, puede ser mejorada aumentando el número de puntos en la plantilla, simplemente imprimiendo un tablero de ajedrez que con más esquinas. Los resultados de la comparación

muestran la flexibilidad y adaptabilidad del método de calibración de Zhang ya que se puede realizar en cualquier escenario.

Desde el punto de vista del modelo estimado, los experimentos con cámaras de calidad media realizados por Salvi [76] y Sun [82] indican que es razonable considerar distorsión tangencial ó desalineamiento cero y distorsión radial de segundo orden. Considerar una distorsión radial de orden cuatro puede ser factible si se disponen de datos con poco nivel de ruido. En caso de estimar distorsiones de orden seis o superiores el propio ruido de las medidas hace que los resultados obtenidos sean peores, además de complicar el algoritmo de calibración. Si se deja de considerar la distorsión radial, se pueden mejorar los resultados de la calibración añadiendo componentes de distorsión tangencial ó descentrado de la imagen.

A partir de los resultados obtenidos por estos dos autores se toma el método de Zhang como referencia hasta el momento para el calibrado de la cámara. Diferentes variantes aparecen de este método. Meng [64] utiliza círculos en la plantilla para evitar la medida de las coordenadas en la plantilla. Además, al proceso de calibrado de Zhang se pueden incluir las mejoras definidas por Hartley [32] con la normalización de los datos, Lavest [50] con la corrección de los puntos en la plantilla y Ahmed [3],[4] o Devernay [19] estimando la distorsión radial de la cámara antes de realizar el proceso de calibrado de la misma. Por otro lado surge la duda si se obtendrán resultados más exactos utilizando el método de calibración basado en plantilla 1D presentado por Zhang [105]. Con el objetivo de incluir mejoras respecto al método de calibración de Zhang utilizando plantilla 2D se van a tomar estas aportaciones para definir un método que permita calibrar la cámara de un modo exacto y sirva como base para calcular incertidumbres en el proceso de calibrado y estimar las condiciones óptimas de calibración.

3.2. Calibración basada plantilla 3D

Se parte de una escena conocida para estimar los parámetros de la cámara. Como escena conocida se refiere a que una plantilla 3D está situada en el escenario cuyos puntos \mathbf{p}_{wi} , $i=1\dots n$, tienen coordenadas conocidas. En este caso se realiza una estimación de los elementos de la matriz de proyección \mathbf{M} , a partir de la posición conocida de los puntos $\mathbf{p}_{wi}=(x_{wi}, y_{wi}, z_{wi})$, $i=1\dots n$, en el mundo respecto al sistema de referencia del escenario $(\mathbf{o}, \{\mathbf{x}_w, \mathbf{y}_w, \mathbf{z}_w\})$ y de la posición de los puntos correspondientes en la imagen $\mathbf{q}_{di}=(u_{di}, v_{di})$, $i=1\dots n$, respecto al sistema de coordenadas de la misma $(\mathbf{c}', \{\mathbf{u}', \mathbf{v}'\})$. Al indicar las coordenadas del punto en la imagen como \mathbf{q}_d se quiere indicar que la información que se dispone son los puntos medidos de la imagen y por lo tanto contaminados con ruido y por la distorsión que introduce la cámara. Teniendo en cuenta que la matriz de proyección \mathbf{M} que se pretende estimar representa la relación lineal que existe entre los conjuntos de puntos \mathbf{p}_{wi} y \mathbf{q}_{pi} , se va a suponer para la demostración que la

distorsión de las coordenadas de los puntos es nula, es decir $\mathbf{q}_{di}=\mathbf{q}_{pi}$. De esta forma, para simplificar la notación se va a utilizar $\mathbf{p}_i=(x_i, y_i, z_i)$ para referirse al conjunto de puntos de la plantilla y $\mathbf{q}_i=(u_i, v_i)$ para referirse al conjunto de puntos de la imagen, asumiendo que la distorsión de la imagen es nula. Por lo tanto, la relación lineal entre el conjunto de puntos \mathbf{p}_i y \mathbf{q}_i esta dada por la matriz de proyección \mathbf{M} la cual se puede obtener según (2.16) o (2.18).

$$\mathbf{q}_i = \mathbf{M} \mathbf{p}_i \quad (3.1)$$

$$\begin{bmatrix} w_i \cdot u_i \\ w_i \cdot v_i \\ w_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (3.2)$$

Para tener una notación más compacta se va a separar la matriz de proyección \mathbf{M} en varias submatrices, teniéndose:

$$\mathbf{M} = \begin{bmatrix} m_1 & m_{14} \\ m_2 & m_{24} \\ m_3 & m_{34} \end{bmatrix} = \begin{bmatrix} m^1 \\ m^2 \\ m^3 \end{bmatrix} \quad (3.3)$$

m^1, m^2, m^3 son los vectores fila de la matriz \mathbf{M} y m_1, m_2, m_3 son los 1x3 vectores fila de la matriz \mathbf{M} sin la última coordenada.

Existen dos planteamientos iniciales equivalentes los cuales son igual de validos bajo las mismas condiciones. En los dos se obtiene la relación que existe entre las coordenadas de los puntos utilizados para la calibración y los elementos de la matriz de proyección.

Planteamiento 1

En este planteamiento se tiene un punto en la plantilla 3D \mathbf{p}_i junto con su equivalente en la imagen 2D \mathbf{q}_i , junto con la matriz de proyección que los relaciona $\mathbf{q}_i = \mathbf{M} \cdot \mathbf{p}_i$. Por lo tanto, el vector \mathbf{q}_i y el vector resultado de $\mathbf{M} \cdot \mathbf{p}_i$ tienen que ser iguales. Dado que están expresados en coordenadas homogéneas, esta condición de igualdad se traduce en que ambos dos tengan la misma dirección pero pueden tener distinto módulo. Por lo tanto se tiene que cumplir que el producto vectorial de ambos sea nulo. Esto garantiza que tienen la misma dirección.

$$\mathbf{q}_i \times \mathbf{M} \cdot \mathbf{p}_i = \mathbf{0} \quad (3.4)$$

Para un punto en la platilla 3D \mathbf{p}_i , el vector $\mathbf{M} \cdot \mathbf{p}_i$ está definido de la siguiente forma:

$$\mathbf{M} \cdot \mathbf{p}_i = \begin{bmatrix} m^1 p_i \\ m^2 p_i \\ m^3 p_i \end{bmatrix} \quad (3.5)$$

Por lo tanto el producto vectorial de ambos vectores será:

$$q_i \times M \cdot p_i = \begin{bmatrix} v_i \cdot m^3 \cdot p_i - m^2 \cdot p_i \\ m^1 \cdot p_i - u_i m^3 \cdot p_i \\ u_i \cdot m^2 \cdot p_i - v_i \cdot m^1 \cdot p_i \end{bmatrix} = 0 \quad (3.6)$$

Organizado en función de los elementos de la matriz de proyección queda:

$$\begin{bmatrix} 0 & -p_i^T & v_i \cdot p_i^T \\ p_i^T & 0 & -u_i \cdot p_i^T \\ -v_i \cdot p_i^T & u_i \cdot p_i^T & 0 \end{bmatrix} \cdot \begin{bmatrix} m^{1T} \\ m^{2T} \\ m^{3T} \end{bmatrix} = 0 \quad (3.7)$$

Dado que la tercera fila es una combinación lineal de las dos primeras, ya que se trabaja en coordenadas homogéneas, sólo se tienen dos ecuaciones linealmente independientes con cada pareja de puntos

$$\begin{bmatrix} 0 & -p_i^T & v_i \cdot p_i^T \\ p_i^T & 0 & -u_i \cdot p_i^T \end{bmatrix} \cdot \begin{bmatrix} m^{1T} \\ m^{2T} \\ m^{3T} \end{bmatrix} = 0 \quad (3.8)$$

Si se extiende esta expresión a todo el conjunto N de puntos disponibles, se puede expresar de forma genérica como:

$$A \cdot m = 0 \quad (3.9)$$

El vector m son los elementos de la matriz de proyección M colocados en un vector columna, los cuales son desconocidos y A es la matriz formada por las coordenadas de los puntos de interés tanto en la imagen como en el mundo 3D, la cuales son conocidas.

$$m = [m^1 \quad m^2 \quad m^3]^T \quad (3.10)$$

Planteamiento 2

Este método se basa en la propiedad de que las coordenadas en la imagen están expresadas en coordenadas homogéneas. Para cada posición conocida de un punto en la plantilla $p_i=(x_i, y_i, z_i)$ y su correspondiente en la imagen $q_i=(u_i, v_i)$, se pueden obtener dos ecuaciones con 12 incógnitas.

$$\begin{aligned} u_i(m_{31}x_i + m_{32}y_i + m_{33}z_i + m_{34}) &= m_{11}x_i + m_{12}y_i + m_{13}z_i + m_{14} \\ v_i(m_{31}x_i + m_{32}y_i + m_{33}z_i + m_{34}) &= m_{21}x_i + m_{22}y_i + m_{23}z_i + m_{24} \end{aligned} \quad (3.11)$$

Si se disponen de n puntos, se puede obtener un sistema lineal de $2 \cdot n$ ecuaciones con 12 incógnitas de la siguiente forma:

$$\begin{bmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1 \cdot x_1 & -u_1 \cdot y_1 & -u_1 \cdot z_1 & -u_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1 \cdot x_1 & -v_1 \cdot y_1 & -v_1 \cdot z_1 & -v_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_n & y_n & z_n & 1 & 0 & 0 & 0 & 0 & -u_n \cdot x_n & -u_n \cdot y_n & -u_n \cdot z_n & -u_n \\ 0 & 0 & 0 & 0 & x_n & y_n & z_n & 1 & -v_n \cdot x_n & -v_n \cdot y_n & -v_n \cdot z_n & -v_n \end{bmatrix} \cdot \begin{bmatrix} m_{11} \\ m_{12} \\ \dots \\ m_{33} \\ m_{34} \end{bmatrix} = 0 \quad (3.12)$$

El sistema de ecuaciones lineales se puede expresar en forma matricial como:

$$\mathbf{A} \cdot \mathbf{m} = 0 \quad (3.13)$$

\mathbf{A} es la matriz de dimensión $(2 \cdot N) \times 12$ que contiene las coordenadas de los puntos en la plantilla 3D y sus correspondientes posiciones de las proyecciones en la imagen, todas ellas conocidas. \mathbf{m} es un vector que contiene todos los elementos de la matriz de proyección \mathbf{M} colocados en vector columna los cuales se desconocen.

$$\mathbf{m} = [m_1 \quad m_{14} \quad m_2 \quad m_{24} \quad m_3 \quad m_{34}]^T \quad (3.14)$$

Resolución de la expresión para obtener \mathbf{m}

Con los dos planteamientos anteriores se llega a una expresión $\mathbf{A} \cdot \mathbf{m} = 0$ en la cual \mathbf{A} se conoce y se busca el vector \mathbf{m} que satisface la expresión. Con cada pareja de puntos se obtiene un par de ecuaciones lineales independientes. Dado un conjunto de 6 parejas de puntos se obtiene una matriz \mathbf{A} con las coordenadas de cada pareja. Con esta matriz se busca la solución a la expresión $\mathbf{A} \cdot \mathbf{m} = 0$ para obtener los elementos de la matriz de proyección \mathbf{M} . El rango de la matriz \mathbf{A} da una idea de cómo se deben situar los puntos para realizar la calibración de la cámara. Se sabe de álgebra lineal que el rango de una matriz de dimensiones $n \times m$ depende de la dimensión de su espacio nulo según la relación:

$$\text{rango}(\mathbf{A}) + \text{null}(\mathbf{A}) = h \quad (3.15)$$

En este caso $h=12$. Por lo tanto existen varias soluciones:

- $\text{rango}(\mathbf{A})=12$: La dimensión del espacio nulo es 0 y por lo tanto existe sólo una única solución para el sistema que es $\mathbf{m}=\mathbf{0}$, la cual no es muy significativa.
- $\text{rango}(\mathbf{A})=11$: La dimensión del espacio nulo es 1. Esto significa que la solución es única. En este caso los elementos de la matriz \mathbf{M} son únicos a excepción de que están definidos con un factor de escala dado que la matriz de proyección es homogénea y sólo los ratios son significativos.
- $\text{rango}(\mathbf{A}) < 11$: En este caso existen infinito número de soluciones para el vector \mathbf{m} . Una forma de que esto ocurra es si todos los puntos utilizados para la calibración están situados en el mismo plano.

Por lo tanto para encontrar una solución única para el vector \mathbf{m} de los elementos de \mathbf{M} es necesario asegurar que los puntos utilizados para la calibración se encuentran en una posición general que no sea un plano. Si seis o más puntos son seleccionados aleatoriamente la probabilidad de que estén situados en el mismo plano es muy pequeña. En el caso de que esto ocurra no se podrá obtener un valor de \mathbf{m} válido.

El proceso de calibración consiste en encontrar el vector \mathbf{m} que mejor satisface las expresiones (3.9) o (3.13). Dado que las medidas de las coordenadas de los puntos utilizados en la calibración están contaminadas con ruido esto provocará que el rango de la matriz \mathbf{A} sea 12 y por lo tanto la solución sea la trivial $\mathbf{m}=\mathbf{0}$. Por lo tanto, para evitar esta situación y no obtener la solución trivial $\mathbf{m}=\mathbf{0}$, la cual no aporta ninguna información, es necesario introducir algún tipo de restricción. Éstas pueden ser las condiciones bajo las cuales dicha matriz de proyección \mathbf{M} puede ser definida las cuales se describen en el apartado 2.5. Las condiciones son las restricciones que debe cumplir la matriz de proyección \mathbf{M} debidas a la naturaleza geométrica del modelo de la cámara. También se pueden introducir restricciones debidas a la condición de matriz homogénea. Por lo tanto, el cálculo del vector \mathbf{m} se debe realizar con algoritmos de optimización de funciones sujetas a dichas restricciones. Existen diferentes técnicas las cuales se describen en los siguientes apartados.

3.2.1. Solución no homogénea

Una alternativa para obtener los elementos de la matriz de proyección contenidos en el vector \mathbf{m} es expresar el conjunto de ecuaciones (3.9) ó (3.13) en forma no homogénea imponiendo que $m_{34}=1$. Esta decisión está justificada desde el punto de vista que la solución está sujeta a un factor de escala ya que la matriz de proyección es homogénea. m_{34} es la componente t_z del vector de traslación \mathbf{t} que sitúa a la cámara en el sistema de referencia del escenario, tal y como se puede ver en la descomposición por parámetros de la matriz \mathbf{M} (2.24). Por lo tanto, este elemento nunca ó sólo en casos muy concretos será nulo. En estos casos no será válido el método de estimación de los elementos de la matriz de proyección descrito en este apartado. Teniendo esto en cuenta, la expresión (3.9) ó (3.13) para cada uno de los casos se transforma en:

$$\mathbf{A}'\cdot\mathbf{m}'=\mathbf{b} \quad (3.16)$$

\mathbf{A}' es la matriz formada por la once primeras columnas de la matriz \mathbf{A} y \mathbf{b} el vector columna con la doceava columna de \mathbf{A} . En este caso el vector \mathbf{m}' son los elementos de la matriz de proyección menos m_{34} , el cual no es necesario estimar su valor ya que éste se ha forzado a 1. Para obtener el valor del vector \mathbf{m}' que satisface la expresión (3.16) se recurre a un algoritmo de mínimos cuadrados. Se tiene un sistema de ecuaciones lineales en el que hay más ecuaciones que incógnitas. La ecuación (3.16) se puede escribir de la forma:

$$\mathbf{b} - \mathbf{A}' \cdot \mathbf{m}' = \mathbf{e} \quad (3.17)$$

\mathbf{e} representa un vector de error para cualquier valor estimado de \mathbf{m}' . La mejor solución de \mathbf{m}' será aquella que minimiza el módulo del vector de error \mathbf{e} . Por lo tanto \mathbf{m}' es el vector tal que $|\mathbf{e}|^2 = \mathbf{e}^T \cdot \mathbf{e}$ es mínimo. Además el valor de este error también depende del número de parejas de puntos que se utiliza en el proceso de estimación. Por lo tanto se plantea el siguiente índice el cual depende de \mathbf{m}' y n , siendo n el número de puntos utilizados.

$$J(\mathbf{m}', n) = \frac{1}{2} \mathbf{e}^T \mathbf{e} = \frac{1}{2} (\mathbf{b} - \mathbf{A}' \cdot \mathbf{m}')^T (\mathbf{b} - \mathbf{A}' \cdot \mathbf{m}') = \mathbf{m}'^T \mathbf{A}'^T \mathbf{A}' \mathbf{m}' - 2 \mathbf{m}'^T \mathbf{A}'^T \mathbf{b} + \mathbf{b}^T \mathbf{b} \quad (3.18)$$

La existencia de un valor de \mathbf{m}' que hace mínima esta expresión será aquel que haga que la primera derivada sea nula.

$$\frac{\delta J}{\delta \mathbf{m}} = \mathbf{A}'^T \mathbf{A}' \cdot \mathbf{m}' - \mathbf{A}'^T \mathbf{b} = \mathbf{0} \quad (3.19)$$

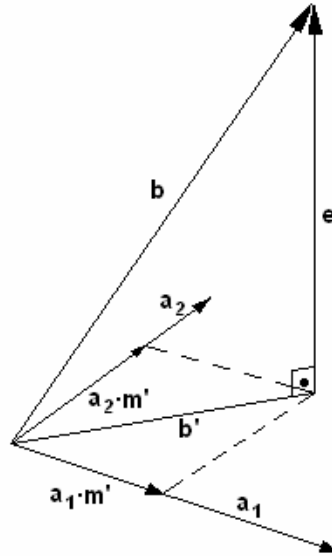
Para garantizar que este valor es mínimo, la segunda derivada del índice respecto a \mathbf{m}' debe ser positiva.

$$\frac{\delta^2 J}{\delta \mathbf{m}^2} = \mathbf{A}'^T \mathbf{A}' > \mathbf{0} \quad (3.20)$$

El valor de la segunda derivada siempre positivo garantiza la existencia de un valor de \mathbf{m}' que hace mínimo el error, es decir, existen unos valores para la matriz de proyección que satisfacen la relación entre la posición de los puntos en el escenario $\mathbf{p}_i = (x_i, y_i, z_i)$ y su posición en la imagen $\mathbf{q}_i = (u_i, v_i)$. Estos valores son los once elementos de la matriz de proyección y se calculan despejando de la ecuación (3.19) de la siguiente forma:

$$\mathbf{m}' = (\mathbf{A}'^T \mathbf{A}')^{-1} \mathbf{A}'^T \mathbf{b} \quad (3.21)$$

Con esto se consigue estimar la matriz de proyección \mathbf{M} siendo el elemento $m_{34} = 1$. Desde un punto de vista geométrico, el ajuste por los mínimos cuadrados puede ser interpretado como un problema de encontrar las constantes que forman el vector \mathbf{m}' de manera que el vector \mathbf{b} se aproxime lo máximo posible a la combinación lineal de los vectores formados por las filas de \mathbf{A}' . Si \mathbf{b} es el vector buscado y \mathbf{b}' es el vector generado por la combinación lineal de los vectores fila de \mathbf{A}' , los cuales son $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{2n}$, tomando $\mathbf{e} = \mathbf{b} - \mathbf{b}'$, se puede observar en la figura 3.3 que el vector \mathbf{e} será mínimo cuando es ortogonal a cada uno de los vectores. Luego en el caso de la figura 3.3 en la cual se representa un caso bidimensional se tiene que dar que el vector \mathbf{e} sea perpendicular al plano formado por los vectores \mathbf{a}_1 y \mathbf{a}_2 . Por lo tanto se tiene que $\mathbf{a}_1 \cdot \mathbf{e} = \mathbf{0}$, el cual expresado de forma matricial para todas las filas de \mathbf{A}' queda como $\mathbf{A}' \cdot \mathbf{e} = \mathbf{0}$. Esto indica que el vector de parámetros estimados será único si los vectores $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{2n}$ son linealmente independientes. Por lo tanto configuraciones de los puntos los cuales estén situados en planos o



3.3 Interpretación geométrica del algoritmo de ajuste por mínimos cuadrados

rectas darán problemas de estimación. Es por este motivo que se necesita una plantilla 3D para estimar los parámetros de la matriz de proyección M .

La matriz de proyección M definitiva se obtendrá multiplicando cada uno de los elementos calculados por la componente t_z del vector de traslación de la cámara en el sistema de coordenadas de la escena. Éste puede ser uno de los inconvenientes que presenta este método. Es necesario conocer la traslación de la cámara a lo largo del eje Z del sistema de coordenadas del mundo 3D para saber cual es el factor de escalado de la matriz obtenida. Una forma alternativa de conocer el factor de escalado de la matriz de proyección es mediante la restricción $|\mathbf{m}_g|=1$. El factor de escalado que hace que el modulo de vector \mathbf{m}_g de la matriz obtenida sea unitario, coincide con la traslación de la cámara a lo largo del eje Z y por lo tanto se puede utilizar para calcular la matriz de proyección M definitiva.

Una vez estimada la matriz de proyección M es posible obtener los parámetros intrínsecos y los extrínsecos de la cámara. Para ello es necesario tener en cuenta la propiedad de ortonormalidad que presenta la matriz de la rotación R . Las técnicas que se pueden utilizar son las presentadas en el apartado 2.6. En el mismo apartado se explica como poder obtener el modelo extendido de la misma.

3.2.2. Solución homogénea

Con restricción $|\mathbf{m}|=1$

Otra alternativa para obtener los elementos de la matriz de proyección contenidos en el vector \mathbf{m} es utilizar las ecuaciones (3.9) ó (3.13) en forma homogénea. Dada que la solución obvia $\mathbf{m}=\mathbf{0}$ no tiene interés, es necesario imponer algún tipo de restricción. Un vector \mathbf{m} que sea

solución de la expresión (3.9) ó (3.13), y cualquier otro vector $k \cdot \mathbf{m}$ son igual de válidos, ya que la norma del vector \mathbf{m} no es importante debido a que éste, está definido con un factor de escala. Por lo tanto, ya que no hay una solución exacta para solución para $\mathbf{A} \cdot \mathbf{m} = \mathbf{0}$, se puede minimizar la norma de $|\mathbf{A} \cdot \mathbf{m}|$ sujeta a la restricción $|\mathbf{m}|=1$. Esto es lo mismo que encontrar el mínimo del cociente $|\mathbf{A} \cdot \mathbf{m}|/|\mathbf{m}|$.

Para la obtención de \mathbf{m} se trata de resolver un problema de minimización de un índice sujeto a restricciones. El índice en este caso es:

$$J = \mathbf{m}^T \mathbf{A}^T \mathbf{A} \mathbf{m} \quad (3.22)$$

Está sujeto a la restricción $|\mathbf{m}|=1$. La correspondiente función de Lagrange es:

$$L = \mathbf{m}^T \mathbf{A}^T \mathbf{A} \mathbf{m} - \lambda(\mathbf{m}^T \mathbf{m} - 1) \quad (3.23)$$

Obteniendo el mínimo de esta función realizando la derivada respecto de \mathbf{m} e igualando a cero se tiene:

$$\frac{\delta L}{\delta \mathbf{m}} = 2 \mathbf{A}^T \mathbf{A} \mathbf{m} - 2\lambda \mathbf{m} = \mathbf{0} \quad (3.24)$$

Esto se puede expresar como:

$$\mathbf{A}^T \mathbf{A} \mathbf{m} = \lambda \mathbf{m} \quad (3.25)$$

Por lo tanto, la solución \mathbf{m} es el vector propio unitario de la matriz $\mathbf{A}^T \cdot \mathbf{A}$ con valor propio más pequeño. Este vector se puede obtener a partir de la descomposición en valores propios de la matriz simétrica definida positiva $\mathbf{A}^T \cdot \mathbf{A}$.

Con restricción $|\mathbf{m}_3|=1$

Faugeras [24] continúa su análisis examinando la matriz de transformación \mathbf{M} . Su análisis muestra que la solución no homogénea en la cual $m_{34}=1$ (3.21), hace que los valores obtenidos para los parámetros intrínsecos dependan del sistema de coordenadas del escenario. Debido a este inconveniente, se impone una nueva restricción $|\mathbf{m}_3|=1$. Para la obtención del vector \mathbf{m} sujeto a esta nueva restricción es necesario evaluar una nueva función de Lagrange.

Calculando explícitamente los coeficientes de \mathbf{M} en función de los parámetros que componen la matriz de proyección se encuentra la particularidad de que $\mathbf{m}_3 = \mathbf{r}_3$, en donde \mathbf{r}_3 corresponde con la tercera fila de la matriz de rotación. Es fácilmente demostrable que

$$r_{31}^2 + r_{32}^2 + r_{33}^2 = 1 \quad (3.26)$$

En consecuencia se tiene que:

$$|\mathbf{m}_3| = m_{31}^2 + m_{32}^2 + m_{33}^2 = 1 \quad (3.27)$$

Por lo tanto se puede obtener el vector \mathbf{m} con los parámetros de la matriz de proyección \mathbf{M} minimizando la ecuación $\mathbf{A} \cdot \mathbf{m} = \mathbf{0}$ teniendo en cuenta esta última restricción $|\mathbf{m}_3| = 1$. La ecuación (3.9) ó (3.13) se puede escribir de la forma:

$$D \cdot \mathbf{m}_9 + F \cdot \mathbf{m}_3 = 0 \quad (3.28)$$

siendo

$$D = \begin{bmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -u_1 \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -v_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_n & y_n & z_n & 1 & 0 & 0 & 0 & 0 & -u_n \\ 0 & 0 & 0 & 0 & x_n & y_n & z_n & 1 & -v_n \end{bmatrix} \quad (3.29)$$

$$F = \begin{bmatrix} -u_1 \cdot x_1 & -u_1 \cdot y_1 & -u_1 \cdot z_1 \\ -u_1 \cdot x_1 & -u_1 \cdot y_1 & -u_1 \cdot z_1 \\ \dots & \dots & \dots \\ -u_n \cdot x_n & -u_n \cdot y_n & -u_n \cdot z_n \\ -v_n \cdot x_n & -v_n \cdot y_n & -v_n \cdot z_n \end{bmatrix} \quad (3.30)$$

El vector \mathbf{m} que contiene los parámetros de la matriz de proyección a estimar también se divide en dos según:

$$\begin{aligned} \mathbf{m}_9 &= [m_1 \quad m_{14} \quad m_2 \quad m_{24} \quad m_{34}]^T \\ \mathbf{m}_3 &= [m_3]^T \end{aligned} \quad (3.31)$$

La función a minimizar es

$$Q = |D \cdot \mathbf{m}_9 + F \cdot \mathbf{m}_3| \quad \text{sujeta a la restricción } |\mathbf{m}_3| = 1 \quad (3.32)$$

Si se introduce la restricción dentro de la función a minimizar aplicando multiplicadores de Lagrange se obtiene:

$$Q = |D \cdot \mathbf{m}_9 + F \cdot \mathbf{m}_3| + \lambda(1 - |\mathbf{m}_3|) \quad (3.33)$$

que desarrollado se tiene

$$Q = \mathbf{m}_9^T \cdot D^T D \cdot \mathbf{m}_9 + \mathbf{m}_3^T \cdot F^T F \cdot \mathbf{m}_3 + \mathbf{m}_9^T \cdot D^T F \cdot \mathbf{m}_3 + \mathbf{m}_3^T \cdot F^T D \cdot \mathbf{m}_9 + \lambda(1 - \mathbf{m}_3^T \mathbf{m}_3) \quad (3.34)$$

Realizando las derivadas parciales respecto a \mathbf{m}_9 y \mathbf{m}_3 .

$$\begin{aligned} \frac{\delta Q}{\delta \mathbf{m}_9} &= D^T D \cdot \mathbf{m}_9 + D^T F \cdot \mathbf{m}_3 = 0 \\ \frac{\delta Q}{\delta \mathbf{m}_3} &= F^T F \cdot \mathbf{m}_3 + F^T D \cdot \mathbf{m}_9 - \lambda \mathbf{m}_3 = 0 \end{aligned} \quad (3.35)$$

De estas expresiones se obtiene que:

$$\begin{aligned} \mathbf{m}_g &= -(D^T D)^{-1} D^T F \cdot \mathbf{m}_3 \\ (F^T F - F^T D (D^T D)^{-1} D^T F) \mathbf{m}_3 &= \lambda \mathbf{m}_3 \end{aligned} \quad (3.36)$$

La última expresión indica que \mathbf{m}_3 es un vector propio de la matriz.

$$G = (F^T F - F^T D (D^T D)^{-1} D^T F) \quad (3.37)$$

G es una matriz simétrica positiva de 3x3 y por lo tanto, tiene sus valores propios positivos. \mathbf{m}_3 es un vector propio de G asociado a un valor propio. El valor de \mathbf{m}_3 es el vector propio asociado al valor propio más pequeño a la matriz G . Para obtener su valor es necesario calcular los valores propios de G , escoger el más pequeño el cual tendrá asociado el vector propio que minimizará la expresión (3.33). Cuando se ha obtenido \mathbf{m}_3 se puede obtener \mathbf{m}_g con la ecuación (3.36). Los elementos de la matriz de proyección M se obtienen a partir de los vectores \mathbf{m}_g y \mathbf{m}_3 . Dado que el signo de \mathbf{m}_3 no está definido, existen dos posibles soluciones de la matriz de proyección, M y $-M$. Se puede elegir una de las dos soluciones sabiendo si el objeto utilizado para la calibración se encuentra en una situación, respecto al eje Z del sistema de coordenadas, delante o detrás de la cámara. Si es positivo $t_z = m_{34} > 0$.

Una vez obtenido la matriz de proyección M es posible extraer los parámetros intrínsecos y extrínsecos que modelan la misma. Las expresiones para obtener cada uno de los parámetros son las expuestas en el apartado 2.6. De la misma forma es posible obtener el modelo extendido de la cámara el cual contiene un quinto parámetro intrínseco, el ángulo que forman los ejes de coordenadas de la imagen.

3.3. Calibración basada en plantilla bidimensional

Zhang [105] propone una técnica de calibración basada en la observación de una plantilla plana desde varias posiciones. Este método resulta más versátil ya que la elaboración de la plantilla se puede realizar fácilmente. La cámara se puede desplazar manualmente ya que no es necesario conocer las posiciones de la cámara desde donde se han tomado las imágenes de la plantilla. Esto hace que sea una técnica muy flexible ya que el método anterior necesita un diseño de plantilla más complicado además de conocer las posiciones exactas de los puntos dentro de la misma. Para calibrar la cámara con este método es necesario estimar las homografías de cada una de las imágenes tomadas de la plantilla. Con las homografías estimadas se calculan los parámetros de la cámara. Previamente a la descripción del método de calibración de la cámara se detalla el proceso de estimación de las homografías.

3.3.1. Cálculo de homografías

En el apartado anterior se ha tratado la estimación de la matriz de proyección a partir de una plantilla formada por varios planos de forma que contiene puntos con coordenadas en las tres dimensiones. En el caso de querer estimar una homografía a partir de una plantilla bidimensional, el razonamiento es parecido al realizado hasta el momento. Las coordenadas de un punto $\mathbf{q}_i=(u_i, v_i)$ en la imagen se calculan a partir de la posición de su correspondiente en la escena $\mathbf{p}_i=(x_i, y_i)$ y de los elementos que forman la homografía según las siguientes expresiones:

$$\begin{aligned} u_i &= \frac{h_{11}x_i + h_{12}y_i + h_{13}}{h_{31}x_i + h_{32}y_i + h_{33}} \\ v_i &= \frac{h_{21}x_i + h_{22}y_i + h_{23}}{h_{31}x_i + h_{32}y_i + h_{33}} \end{aligned} \quad (3.38)$$

Reordenando estas dos ecuaciones, para cada posición conocida de un punto en la plantilla $\mathbf{p}_i=(x_i, y_i)$ y su correspondiente en la imagen $\mathbf{q}_i=(u_i, v_i)$, se pueden obtener dos ecuaciones con 9 incógnitas, que corresponden a los elementos de la homografía.

$$\begin{aligned} u_i(h_{31}x_i + h_{32}y_i + h_{33}) &= h_{11}x_i + h_{12}y_i + h_{13} \\ v_i(h_{31}x_i + h_{32}y_i + h_{33}) &= h_{21}x_i + h_{22}y_i + h_{23} \end{aligned} \quad (3.39)$$

Si se tiene disponibles n puntos, se puede obtener un sistema lineal de $2 \cdot n$ ecuaciones con 9 incógnitas de la siguiente forma:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1 \cdot x_1 & -u_1 \cdot y_1 & -u_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1 \cdot x_1 & -v_1 \cdot y_1 & -v_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_n & y_n & 1 & 0 & 0 & 0 & -u_n \cdot x_n & -u_n \cdot y_n & -u_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -v_n \cdot x_n & -v_n \cdot y_n & -v_n \end{bmatrix} \cdot \begin{bmatrix} h_{11} \\ h_{12} \\ \dots \\ h_{32} \\ h_{33} \end{bmatrix} = \mathbf{0} \quad (3.40)$$

El sistema de ecuaciones lineales se puede expresar en forma matricial como:

$$\mathbf{A} \cdot \mathbf{h} = \mathbf{0} \quad (3.41)$$

\mathbf{A} en este caso, la matriz de dimensión $(2 \cdot n) \times 9$ que contiene las coordenadas de los puntos en la plantilla 3D y sus correspondientes posiciones de las proyecciones en la imagen, todas ellas conocidas. \mathbf{h} es un vector que contiene todos los elementos de la homografía \mathbf{H} colocados en vector.

$$\mathbf{h} = [h_{11} \ h_{12} \ h_{13} \ h_{21} \ h_{22} \ h_{23} \ h_{31} \ h_{32} \ h_{33}]^T \quad (3.42)$$

Para obtener los elementos de la homografía \mathbf{h} , dado que la solución $\mathbf{h}=\mathbf{0}$ no tiene interés, se impone la restricción $|\mathbf{h}|=1$. Un vector \mathbf{h} que sea solución de la expresión (3.41), y cualquier otro vector $k \cdot \mathbf{h}$ son igual de válidos, ya que \mathbf{h} , está definido con un factor de escala. Por lo tanto, ya que no hay una solución exacta para solución para $\mathbf{A} \cdot \mathbf{h}=\mathbf{0}$, se puede minimizar

la norma de $|\mathbf{A} \cdot \mathbf{h}|$ sujeta a la restricción $|\mathbf{h}|=1$. Es la misma operación realizada para el cálculo de la matriz de proyección \mathbf{M} en el apartado 3.3.2. La solución \mathbf{h} es el vector propio unitario de la matriz $\mathbf{A}^T \cdot \mathbf{A}$ asociado al valor propio más pequeño. Este vector se puede obtener a partir de la descomposición en valores propios de la matriz simétrica definida positiva $\mathbf{A}^T \cdot \mathbf{A}$.

3.3.2. Proceso de calibrado con plantillas 2D

El modelo lineal de la cámara que se va a estimar es el descrito en el capítulo anterior, mediante el cual la obtención de las coordenadas del punto en la imagen se pueden obtener según:

$$q_i = \lambda \cdot K \cdot [R \quad t] p = \lambda \cdot K \cdot [r_1 \quad r_2 \quad r_3 \quad t] p_i$$

$$\begin{bmatrix} w_i \cdot u_i \\ w_i \cdot v_i \\ w_i \end{bmatrix} = \lambda \cdot \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (3.43)$$

λ es un factor de escala ya que los puntos están expresados en coordenadas homogéneas. \mathbf{R} y \mathbf{t} representan los parámetros extrínsecos de la cámara siendo \mathbf{r}_i las columnas de la matriz de rotación \mathbf{R} . \mathbf{K} contiene los parámetros intrínsecos siendo α_u y α_v los factores de escala en cada uno de los ejes de la imagen, u_0 y v_0 son las coordenadas del punto principal de la imagen y γ es el parámetro que representa la pérdida de ortogonalidad de los ejes de coordenadas en la imagen.

Si se parte de que la plantilla de calibración es plana, se puede asumir que los puntos de dicha plantilla están colocados de forma que su coordenada $z_w=0$. En este caso el modelo se reduce según la siguiente expresión:

$$q_i = \lambda \cdot K \cdot [R \quad t] p = \lambda \cdot K \cdot [r_1 \quad r_2 \quad r_3 \quad t] p_i = \lambda \cdot K \cdot [r_1 \quad r_2 \quad t] p_i \quad (3.44)$$

Ahora el modelo inicial se transforma en una homografía \mathbf{H} que relaciona las coordenadas de los puntos de la plantilla plana del escenario con sus correspondientes en la imagen.

$$q_i = \lambda \cdot K \cdot [r_1 \quad r_2 \quad t] p_i = H \cdot p_i \quad (3.45)$$

Esta homografía puede ser calculada según el método descrito en el apartado anterior. Si se separan las columnas de la homografía se obtiene que

$$[h_1 \quad h_2 \quad h_3] = \lambda \cdot K \cdot [r_1 \quad r_2 \quad t] \quad (3.46)$$

Dado que \mathbf{R} es una matriz de rotación, los vectores que la componen cumplen las restricciones de ortonormalidad, es decir, $\mathbf{r}_1^T \cdot \mathbf{r}_1 = \mathbf{r}_2^T \cdot \mathbf{r}_2$ y que $\mathbf{r}_1^T \cdot \mathbf{r}_2 = 0$. Por lo tanto, si se extraen los vectores de rotación a partir de la última expresión se tiene que

$$\begin{aligned} h_1^T K^{-T} K^{-1} h_1 &= h_2^T K^{-T} K^{-1} h_2 \\ h_1^T K^{-T} K^{-1} h_2 &= 0 \end{aligned} \quad (3.47)$$

Estas son las dos restricciones básicas de los parámetros intrínsecos dada una homografía. Para resolver este problema se propone una solución analítica seguida de una optimización no lineal. La matriz $\mathbf{K}^T \cdot \mathbf{K}^{-1}$ está compuesta por los parámetros intrínsecos de la cámara de la siguiente forma

$$\begin{aligned} K^{-T} K^{-1} = B &= \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \\ &= \begin{bmatrix} 1/\alpha^2 & -\gamma/\alpha^2\beta & V_0\gamma - U_0\beta/\alpha^2\beta \\ -\gamma/\alpha^2\beta & \gamma^2/\alpha^2\beta + 1/\beta^2 & -\gamma(V_0\gamma - U_0\beta)/\alpha^2\beta - U_0/\beta^2 \\ V_0\gamma - U_0\beta/\alpha^2\beta & -\gamma(V_0\gamma - U_0\beta)/\alpha^2\beta - U_0/\beta^2 & (V_0\gamma - U_0\beta)^2/\alpha^2\beta + V_0/\beta^2 + 1 \end{bmatrix} \end{aligned} \quad (3.48)$$

Teniendo en cuenta que se trata de una matriz simétrica se puede definir con un vector de 6 elementos de la forma:

$$\mathbf{b} = [b_{11} \quad b_{12} \quad b_{22} \quad b_{13} \quad b_{23} \quad b_{33}]^T \quad (3.49)$$

Por lo tanto si la columna i -ésima de la matriz \mathbf{H}_i es $\mathbf{h}_i = (h_{i1}, h_{i2}, h_{i3})$, se tiene que

$$h_i^T B h_i = v_{ij}^T \mathbf{b} \quad (3.50)$$

siendo

$$v_{ij}^T = [h_{i1}h_{j1} \quad h_{i1}h_{j2} + h_{i2}h_{j1} \quad h_{i2}h_{j2} \quad h_{i3}h_{j1} + h_{i1}h_{j3} \quad h_{i3}h_{j2} + h_{i2}h_{j3} \quad h_{i3}h_{j3}] \quad (3.51)$$

Con estas expresiones se pueden escribir las restricciones de la matriz \mathbf{R} en dos ecuaciones homogéneas en función de \mathbf{b} de la siguiente forma

$$\begin{bmatrix} v_{11}^T & -v_{22}^T \\ v_{12}^T \end{bmatrix} \mathbf{b} = \mathbf{0} \quad (3.52)$$

Si se calculan n homografías a partir de n imágenes obtenidas de la plantilla se obtiene un sistema de ecuaciones de la forma

$$V \cdot \mathbf{b} = \mathbf{0} \quad (3.53)$$

V es una matriz de dimensiones $2n \times 6$. Si $n \geq 3$ se tiene una solución general con solución única de \mathbf{b} definida con un factor de escala. Si $n=2$ se puede imponer la restricción $\gamma=0$ la cual se añade a las ecuaciones anteriores. Siguiendo el mismo razonamiento realizado para la resolución de la expresión (3.9), en la cual se adoptaba la solución homogénea restringiendo el módulo $|\mathbf{m}|=1$, la solución del vector \mathbf{b} es el vector propio de $V^T \cdot V$ asociado al valor propio más pequeño. En este caso se tiene también que $|\mathbf{b}|=1$. Cuando se ha estimado \mathbf{b} se pueden obtener los parámetros intrínsecos de la cámara que forman la matriz S según:

$$\begin{aligned}
v_0 &= \frac{(b_{12}b_{13} - b_{11}b_{23})}{(b_{11}b_{22} - b_{12}^2)} \\
\lambda &= b_{33} - \frac{(b_{13}^2 + v_0(b_{12}b_{13} - b_{11}b_{23}))}{b_{11}} \\
\alpha_u &= \sqrt{\frac{\lambda}{b_{11}}} \\
\alpha_v &= \sqrt{\frac{\lambda b_{11}}{(b_{11}b_{22} - b_{12}^2)}} \\
\gamma &= \frac{-b_{12}\alpha_u^2\alpha_v^2}{\lambda} \\
u_0 &= \frac{\gamma \cdot v_0}{\alpha_u} - \frac{b_{13}\alpha_v^2}{\lambda}
\end{aligned} \tag{3.54}$$

λ representa un factor de escala entre la matriz B y la matriz $K^T \cdot K^{-1}$. Utilizando los parámetros intrínsecos se forma la matriz S la cual permite estimar los parámetros extrínsecos según las siguientes expresiones:

$$\begin{aligned}
r_1 &= \lambda \cdot S^{-1} \cdot h_1 \\
r_2 &= \lambda \cdot S^{-1} \cdot h_2 \\
r_3 &= r_1 \times r_2 \\
t &= \lambda \cdot S^{-1} \cdot h_3
\end{aligned} \tag{3.55}$$

3.4. Estimación no lineal de los parámetros de la cámara

Hasta el momento las técnicas presentadas para la estimación de los parámetros de la cámara realizaban la calibración de la misma basándose únicamente en técnicas lineales de optimización. Debido a las técnicas utilizadas los parámetros estimados de la cámara han sido aquellos que representaban una relación lineal entre las coordenadas del punto \mathbf{p}_w en el escenario

y las coordenadas de su correspondiente punto en la imagen \mathbf{q}_i . En un caso se toman puntos distribuidos en un plantilla tridimensional y en el otro caso, los puntos formaban un plano bidimensional. Los parámetros que establecen esta relación lineal son \mathbf{R} y \mathbf{t} que representan los parámetros extrínsecos de la cámara y \mathbf{K} que contiene los parámetros intrínsecos siendo $\alpha_u, \alpha_v, u_0, v_0$ y γ . En ninguno de los casos se ha contemplado el cálculo de los parámetros que establecen relaciones no lineales. Estos parámetros son aquellos que modelaban la distorsión de la imagen debida a imperfecciones en el proceso de construcción y montaje de las lentes.

Por este motivo, dado un conjunto de puntos establecidos dentro de un tipo de plantilla $\mathbf{p}_i, i=1..n$ y otro conjunto de puntos $\mathbf{q}_i, i=1..n$, que corresponden a la imagen de los anteriores tomada con una cámara de ciertas características desde un punto del escenario, en ninguno de los casos, el modelo de la cámara calculado cumplirá la expresión $\mathbf{q}_i = \mathbf{M} \cdot \mathbf{p}_i$ que establece la relación lineal entre un punto \mathbf{p}_i y su imagen \mathbf{q}_i dada una matriz de proyección \mathbf{M} .

Si a pesar de todo, se tomara la imagen de la plantilla con una cámara ideal en la cual no existiera distorsión en la imagen debida a imperfecciones constructivas de las lentes, la expresión $\mathbf{q}_i = \mathbf{M} \cdot \mathbf{p}_i$ seguiría sin ser cierta para los parámetros de la cámara calculados. En este caso, la imperfección de los parámetros calculados vendría dada por el ruido que contamina las medidas de las coordenadas de los puntos de interés tanto en el plantilla, como en la imagen.

Llegados a este punto es necesario separar entre los valores reales y valores medidos ó estimados. Las coordenadas reales de los conjuntos de puntos en la imagen y en la plantilla se denotan mediante \mathbf{q}_i y \mathbf{p}_i respectivamente, y representan los valores reales de las coordenadas. Las coordenadas de estos conjuntos de puntos que son medidas y utilizadas posteriormente para la estimación de los parámetros de la cámara se denotan mediante $\mathbf{q}_i^\#$ y $\mathbf{p}_i^\#$. La matriz de proyección real se nombra como \mathbf{M} . Ésta satisface la expresión $\mathbf{q}_i = \mathbf{M} \cdot \mathbf{p}_i$ y contiene los valores reales de los parámetros de la cámara $\mathbf{K}, \mathbf{R}, \mathbf{t}$. La matriz de proyección estimada de forma lineal a partir de las coordenadas $\mathbf{q}_i^\#$ y $\mathbf{p}_i^\#$ se denota mediante $\mathbf{M}^\#$ y contiene los valores estimados de los parámetros de la cámara $\mathbf{K}^\#, \mathbf{R}^\#, \mathbf{t}^\#$. En este caso, no se satisface la expresión $\mathbf{q}_i^\# = \mathbf{M}^\# \cdot \mathbf{p}_i^\#$. El error que se produce es lo que se llamará posteriormente distancia geométrica o error de estimación.

Así pues, dos fuentes de errores existen en la estimación de los parámetros de la cámara. Una de ella es debida a la imperfección del modelo, ya que se pretende estimar un conjunto de parámetros que no representan el comportamiento real de la cámara. La otra fuente de imperfección es el ruido en las coordenadas de los puntos de interés $\mathbf{q}_i, \mathbf{p}_i$ debidas a las imprecisiones en las medidas. Estas dos fuentes de errores, hacen que los métodos lineales presentados anteriormente estimen los parámetros de forma más o menos aproximada. Esta aproximación está en función del nivel de ruido en las medidas de las coordenadas de los puntos de interés y de la distorsión que introduce la cámara en la imagen. Unos métodos son más

sensibles que otros al nivel de ruido. Además con estos métodos de estimación, las restricciones (2.28) ó (2.29) que debe satisfacer la matriz de proyección, sólo se garantiza la primera de ella debido a que tiene carácter lineal. La segunda restricción no está garantizada que se cumpla en ningún momento.

Para obtener el conjunto de parámetros que obtengan un modelo más aproximado al real es necesario realizar una búsqueda no lineal de los mismos. Ésta consiste en plantear un índice de coste a minimizar. Según el índice de coste planteado se consigue minimizar un error y obtener unos resultados diferentes. A continuación se describen algunos de estos errores.

3.4.1. Errores en la estimación de los parámetros

Asumiendo la estimación del modelo lineal de cámara, se definen diferentes errores que miden la calidad de los parámetros estimados. A continuación se definen varios de ellos y se realiza una comparativa. Estos errores servirán a posteriori para definir funciones de coste que permitan determinar los parámetros de la cámara de forma no lineal.

Distancia algebraica

En este caso se minimiza la norma $|\mathbf{A}^\# \cdot \mathbf{m}^\#|$ donde la matriz $\mathbf{A}^\#$ está formada por datos conocidos, los cuales están contaminados con ruido y $\mathbf{m}^\#$ es el vector de parámetros a calcular de forma lineal. Coincide con la estimación basada en plantilla tridimensional pero se podría aplicar igualmente al caso bidimensional de la homografía. En este caso, primero se trataría la estimación de las homografías mediante la expresión (3.41) y después la estimación de los parámetros auxiliares del vector \mathbf{v} (3.53). Siguiendo en el caso de la calibración de cámara mediante plantilla 3D, el vector de error \mathbf{e} es el vector residuo $|\mathbf{A}^\# \cdot \mathbf{m}^\#| = \mathbf{e}$ y es la norma de este vector lo que se pretende minimizar. Las componentes de este vector surgen de las filas de la matriz $\mathbf{A}^\#$, formadas con los datos medidos. Cada fila contribuye a \mathbf{e} con un error parcial e_i . En el caso del cálculo de la homografía el error e_i es el error algebraico asociado a la pareja de puntos $\mathbf{q}_i^\#$ y $\mathbf{p}_i^\#$ junto con la matriz de proyección $\mathbf{M}_i^\#$ estimada. La norma del vector \mathbf{e} es un escalar llamado distancia algebraica. La suma de todos los módulos del error algebraico asociado a cada pareja de puntos constituye la distancia algebraica total.

$$\sum_i d_{alg}(\mathbf{q}_i^\#, \mathbf{M}_i^\# \cdot \mathbf{p}_i^\#)^2 = \sum_i |e_i|^2 = |\mathbf{A}^\# \cdot \mathbf{m}^\#|^2 = |\mathbf{e}|^2 \quad (3.56)$$

La distancia algebraica no tiene ningún significado geométrico. Por este motivo, la reducción de esta distancia puede llevar a situaciones absurdas aunque algebraicamente la distancia sea pequeña. Si se normaliza la información de partida los resultados obtenidos con la optimización de esta distancia son buenos[32]. Además, la minimización de esta distancia es

poco costosa desde el punto de vista computacional. Dado que el planteamiento del índice es lineal, solamente es posible estimar parámetros que representen relaciones lineales entre los datos de partida. Representa la estimación de los parámetros realizada hasta el momento.

Distancia geométrica

Ésta es la distancia euclídea entre dos conjuntos de puntos. Estos dos conjuntos de puntos son el conjunto de puntos de la imagen con coordenadas sin ruido \mathbf{q}_i y el conjunto de puntos resultado de transformar los puntos de la plantilla con la matriz de proyección estimada $\mathbf{M}_2^\#$. Es decir, en este caso existen un conjunto de puntos en la imagen \mathbf{q}_{i2}^* , resultado de la transformación de los puntos del escenario con una matriz de proyección estimada $\mathbf{M}_2^\#$ y también está el conjunto de puntos de la imagen iniciales $\mathbf{q}_i^\#$ utilizados para estimar la matriz de proyección $\mathbf{M}_2^\#$. Con el conjunto de puntos con coordenadas sin ruido de la escena \mathbf{p}_i y el conjunto de la imagen $\mathbf{q}_i^\#$ con ruido se estima una matriz de proyección $\mathbf{M}_2^\#$. En este caso sólo se considera que están contaminadas con ruido las coordenadas de los puntos en la imagen. Con esta matriz de proyección estimada $\mathbf{M}_2^\#$ se obtiene un nuevo conjunto de puntos en la imagen \mathbf{q}_i^* a partir de los puntos de la escena \mathbf{p}_i según $\mathbf{q}_{i2}^* = \mathbf{M}_2^\# \cdot \mathbf{p}_i$. Existe un punto \mathbf{q}_{i2}^* en el nuevo conjunto correspondiente al punto inicial utilizado para la estimación $\mathbf{q}_i^\#$. La distancia geométrica de estos dos puntos \mathbf{q}_{i2}^* y $\mathbf{q}_i^\#$ es la distancia euclídea que existe entre ellos. Para todo el conjunto de puntos, la distancia geométrica será la suma de las distancias euclídeas de todos los puntos.

$$d_{geo} = \sum_i d_{eucli}(\mathbf{q}_i^\#, \mathbf{M}_2^\# \cdot \mathbf{p}_i)^2 \quad (3.57)$$

En el caso de tener ruido en las medidas de las coordenadas de los puntos de la plantilla $\mathbf{p}_i^\#$ la matriz de proyección $\mathbf{M}_3^\#$ se estima en este caso con los puntos de la imagen sin ruido \mathbf{q}_i y el conjunto de puntos de la plantilla con ruido $\mathbf{p}_i^\#$. Con esta matriz de proyección estimada se obtiene un nuevo conjunto de puntos en la imagen \mathbf{q}_{i3}^* , los cuales son el resultado de transformar las coordenadas con ruido de los puntos del espacio $\mathbf{p}_i^\#$ con la matriz de proyección estimada $\mathbf{q}_{i3}^* = \mathbf{M}_3^\# \cdot \mathbf{p}_i^\#$. En este caso la distancia geométrica es la distancia euclídea que existe entre los puntos \mathbf{q}_{i3}^* y \mathbf{q}_i . La distancia geométrica total será la suma de todas.

$$d_{geo} = \sum_i d_{eucli}(\mathbf{q}_i, \mathbf{M}_3^\# \cdot \mathbf{p}_i^\#)^2 \quad (3.58)$$

En el caso de tener ruido tanto en los puntos de la imagen $\mathbf{q}_i^\#$ como en los del escenario $\mathbf{p}_i^\#$, se estima una matriz de proyección a partir de los coordenadas contaminadas y $\mathbf{q}_i^\#$ obteniéndose una matriz $\mathbf{M}_4^\#$. Con esta matriz de proyección estimada $\mathbf{M}_4^\#$ y el conjunto de puntos con ruido de la plantilla $\mathbf{p}_i^\#$ se obtiene un nuevo conjunto de puntos en la imagen

$\mathbf{q}_i^* = \mathbf{M}_4^\# \cdot \mathbf{p}_i^\#$. La distancia geométrica en este caso es la distancia euclídea entre los puntos \mathbf{q}_i^* y $\mathbf{q}_i^\#$. Éste correspondería a un caso más real.

$$d_{geo} = \sum_i d_{eucli}(\mathbf{q}_i^\#, \mathbf{M}_4^\# \cdot \mathbf{p}_i^\#)^2 \quad (3.59)$$

Error de reproyección

El error de reproyección es la distancia euclídea que existe entre los puntos con ruido de la imagen utilizados para la estimación $\mathbf{q}_i^\#$ y $\mathbf{p}_i^\#$, y un conjunto de puntos que satisface plenamente dicha matriz de proyección estimada \mathbf{q}_i^* y \mathbf{p}_i^* . Cuando se estima una matriz de proyección con las coordenadas de los puntos con ruido $\mathbf{q}_i^\#$ y $\mathbf{p}_i^\#$ se obtiene una matriz $\mathbf{M}_4^\#$ la cual no satisface los conjuntos de los puntos, es decir $\mathbf{q}_i^\# \neq \mathbf{M}_4^\# \cdot \mathbf{p}_i^\#$. Para el cálculo del error de reproyección es necesario encontrar dos conjuntos de puntos, uno en el escenario \mathbf{p}_i^* y otro en la imagen \mathbf{q}_i^* que satisface plenamente la transformación estimada $\mathbf{M}_4^\#$ con las coordenadas de los puntos con ruido, es decir $\mathbf{q}_i^* = \mathbf{M}_4^\# \cdot \mathbf{p}_i^*$. Por lo tanto, el error de reproyección es la suma de todas las distancias euclídeas entre los puntos \mathbf{q}_i^* y $\mathbf{q}_i^\#$.

$$d_{geo} = \sum_i d_{eucli}(\mathbf{q}_i^\#, \mathbf{M}_4^\# \cdot \mathbf{p}_i^*)^2 \quad (3.60)$$

Para la obtención de los puntos \mathbf{q}_i^* y \mathbf{p}_i^* que satisfacen plenamente la matriz de proyección estimada $\mathbf{M}_4^\#$ se realiza por métodos iterativos, según explica Sampson [77]. La minimización de este error implica tanto obtener una matriz de proyección $\mathbf{M}_4^\#$ como los conjuntos de puntos que la satisfacen \mathbf{q}_i^* y \mathbf{p}_i^* .

Comparación del error algebraico con el error geométrico

Se tienen las coordenadas de la imagen con ruido $\mathbf{q}_i^\# = (w_i^\# \cdot u_i^\#, w_i^\# \cdot v_i^\#, w_i^\#)$ y las coordenadas de los puntos en la plantilla sin ruido $\mathbf{p}_i = (x_i, y_i, z_i, 1)$. Con ambos conjuntos de puntos se estima la matriz de proyección $\mathbf{M}^\#$ con la cual se obtienen un conjunto de puntos en la imagen $\mathbf{q}_i^* = (w_i^* \cdot u_i^*, w_i^* \cdot v_i^*, w_i^*)$ resultado de, $\mathbf{q}_i^* = \mathbf{M}^\# \cdot \mathbf{p}_i$, se va a estudiar la relación que existe entre el error algebraico y el error geométrico.

Por un lado, atendiendo al planteamiento 1 se tiene que el error algebraico según (3.9) en este caso vale:

$$A \cdot m = \begin{bmatrix} 0 & -w_i^\# \cdot p_i^T & w_i^\# \cdot v_i^\# \cdot p_i^T \\ w_i^\# \cdot p_i^T & 0 & -w_i^\# \cdot u_i^\# \cdot p_i^T \end{bmatrix} \cdot \begin{bmatrix} m^{1T} \\ m^{2T} \\ m^{3T} \end{bmatrix} = e_i \quad (3.61)$$

Desarrollando la expresión se tiene el vector e_i de residuos vale:

$$e_i = \begin{bmatrix} -w_i^\# \cdot p_i^T \cdot m^{2T} + w_i^\# \cdot v_i^\# \cdot p_i^T \cdot m^{3T} \\ w_i^\# \cdot p_i^T \cdot m^{1T} - w_i^\# \cdot u_i^\# \cdot p_i^T \cdot m^{3T} \end{bmatrix} = \begin{bmatrix} -w_i^\# \cdot v_i^\# \cdot w_i^* + w_i^\# \cdot v_i^\# \cdot w_i^* \\ w_i^\# \cdot u_i^\# \cdot w_i^* - w_i^\# \cdot u_i^\# \cdot w_i^* \end{bmatrix} \quad (3.62)$$

Por lo tanto el error algebraico será el módulo de este vector:

$$d_{alg}(q_i^\#, q_i^*) = \left(\left(-w_i^\# \cdot v_i^\# \cdot w_i^* + w_i^\# \cdot v_i^\# \cdot w_i^* \right)^2 + \left(w_i^\# \cdot u_i^\# \cdot w_i^* - w_i^\# \cdot u_i^\# \cdot w_i^* \right)^2 \right)^{\frac{1}{2}} \quad (3.63)$$

Por otro lado, el error geométrico entre los puntos $q_i^\# = (w_i^\# \cdot u_i^\#, w_i^\# \cdot v_i^\#, w_i^\#)$ y $q_i^* = (w_i^* \cdot u_i^*, w_i^* \cdot v_i^*, w_i^*)$ se puede calcular como:

$$\begin{aligned} d_{geo}(q_i^\#, q_i^*) &= \left(\left(\frac{w_i^\# \cdot u_i^\#}{w_i^\#} - \frac{w_i^* \cdot u_i^*}{w_i^*} \right)^2 + \left(\frac{w_i^\# \cdot v_i^\#}{w_i^\#} - \frac{w_i^* \cdot v_i^*}{w_i^*} \right)^2 \right)^{\frac{1}{2}} \\ &= \left(\left(\frac{w_i^\# \cdot u_i^\# \cdot w_i^* - w_i^\# \cdot u_i^\# \cdot w_i^*}{w_i^\# \cdot w_i^*} \right)^2 + \left(\frac{-w_i^\# \cdot v_i^\# \cdot w_i^* + w_i^\# \cdot v_i^\# \cdot w_i^*}{w_i^\# \cdot w_i^*} \right)^2 \right)^{\frac{1}{2}} \\ &= \frac{d_{alg}(q_i^\#, q_i^*)}{w_i^\# \cdot w_i^*} \end{aligned} \quad (3.64)$$

Por lo tanto, la distancia algebraica y la geométrica no son iguales. Sólo serán iguales en el caso de que $w_i^\# = w_i^* = 1$, es decir, cuando se expresan las coordenadas de los puntos de la forma $q_i^\# = (u_i^\#, v_i^\#, 1)$ y $q_i^* = (u_i^*, v_i^*, 1)$ y además, la matriz de proyección tenga la tercera fila formada por el vector $m^3 = (0, 0, 0, 1)$, cosa que sólo se puede conseguir en configuraciones muy concretas de los puntos elegidos para la estimación de la matriz.

3.4.2. Técnicas de optimización

Hasta el momento las técnicas lineales utilizadas para la obtención de los elementos de la matriz de proyección tratan de minimizar el error algebraico. Dado que este error no tiene ningún significado físico, se pueden llegar a soluciones que son correctas desde un punto de vista algebraico pero que resultan absurdas desde un punto de vista geométrico. Por otro lado, estas técnicas lineales no garantizan el cumplimiento de las restricciones que debe satisfacer la matriz de proyección (2.28) ó (2.29) dado que son relaciones no lineales. También, es imposible de estimar por técnicas lineales los parámetros del modelo no lineal de la cámara, el cual tiene en cuenta las distorsiones que produce la cámara en la imagen (2.77). El no cumplimiento de las restricciones y la imposibilidad de estimar el modelo no lineal de la cámara, es debido a que se plantean funciones no lineales imposibles de resolver con las técnicas lineales presentadas. Por este motivo, con el objetivo de poder encontrar los parámetros que mejor satisfacen estas funciones no lineales, se plantea la estimación de los parámetros de la cámara mediante técnicas de optimización no lineales. En este caso se opera como en los casos anteriores, se plantea una

función de coste no lineal, la cual se pretende minimizar obteniendo los valores de los parámetros que mejor la satisfacen.

Para la solución del problema se utilizan técnicas de optimización iterativas tales como el método de Newton o el método de Levenberg-Marquardt. La idea general consiste en buscar ceros en una función que relaciona el espacio de medidas con el espacio de parámetros. Formalmente, se tiene el vector \mathbf{b} de medidas en el espacio euclideo R^N y se tiene un vector de parámetros \mathbf{a} en el espacio euclideo R^M relacionados por una función $\mathbf{b}=f(\mathbf{a})$. Se tiene un vector de medias $\mathbf{b}^\#$ que se aproxima al vector de valores reales \mathbf{b} y se desea obtener un vector de parámetros \mathbf{a}^* que mejor satisface la relación. En concreto, se busca un vector \mathbf{a}^* que satisface $\mathbf{b}^\#=f(\mathbf{a}^*)+\mathbf{e}$ para el cual $|\mathbf{e}|$ es mínimo. En el caso de que f sea una función no lineal, se parte de un valor inicial de los parámetros \mathbf{a}_0 , a partir del cual se inicia la búsqueda. El valor de los parámetros buscados va cambiando suponiendo que la función es localmente lineal en ese valor. Se tiene $\mathbf{e}_0=\mathbf{b}^\#-f(\mathbf{a}_0)$. En la siguiente iteración se busca un vector \mathbf{a}_1 el cual viene dado por $\mathbf{a}_1=\mathbf{a}_0+\Delta\mathbf{a}_0$ que hace que \mathbf{e}_1 sea menor que \mathbf{e}_0 . Haciendo una aproximación de Taylor de la función f en el valor del vector \mathbf{a}_0 se tiene el valor de la función en \mathbf{a}_0 a partir del valor en \mathbf{a}_0 :

$$f(\mathbf{a}_1) = f(\mathbf{a}_0 + \Delta\mathbf{a}_0) = f(\mathbf{a}_0) + J \cdot \Delta\mathbf{a}_0 \quad \text{donde } J = \frac{\partial f(\mathbf{a}_0)}{\partial \mathbf{a}} \quad (3.65)$$

De esta última expresión y dado que el error cometido con el nuevo vector \mathbf{a}_1 debe ser menor, se obtiene que:

$$\mathbf{b}^\# - f(\mathbf{a}_1) = \mathbf{b}^\# - f(\mathbf{a}_0) - J \cdot \Delta\mathbf{a}_0 = \mathbf{e}_0 - J \cdot \Delta\mathbf{a}_0 \quad (3.66)$$

Por lo tanto, se calcula un $\Delta\mathbf{a}_0$ que minimice el módulo de $|\mathbf{e}_0 - J \cdot \Delta\mathbf{a}_0|$. El vector $\Delta\mathbf{a}_0$ se calcula resolviendo la ecuación normal $\mathbf{J}^T \cdot \mathbf{J} \cdot \Delta\mathbf{a}_0 = \mathbf{J} \cdot \mathbf{e}_0$. $\Delta\mathbf{a}_0$ es la pseudo inversa de \mathbf{J} multiplicada por el vector de errores \mathbf{e}_0 , $\Delta\mathbf{a}_0 = \mathbf{J}^+ \cdot \mathbf{e}_0$. Partiendo de un vector de parámetros iniciales \mathbf{a}_0 y realizando sucesivas iteraciones, se calculan sucesivos incrementos $\Delta\mathbf{a}_i$ del vector de parámetros que hacen que finalmente se obtenga un nuevo vector de parámetros que satisface mejor la función f :

$$\mathbf{a}_{i+1} = \mathbf{a}_i + \Delta\mathbf{a}_i \quad (3.67)$$

$\Delta\mathbf{a}_i$ es la solución de la ecuación normal:

$$\mathbf{J} \cdot \Delta\mathbf{a}_i = \mathbf{e}_i \quad (3.68)$$

\mathbf{J} es la matriz Jacobiana de f evaluada en el vector \mathbf{a}_i y $\mathbf{e}_i = \mathbf{b}^\# - f(\mathbf{a}_i)$. En el caso de funciones multidimensionales en las cuales se buscan soluciones basadas en ajuste por mínimos cuadrados, la aplicación de esta técnica es relativamente inmediata. En teoría, esta búsqueda de los parámetros que mejor satisfacen la función no lineal debe converger a un mínimo absoluto. Sin embargo, en la práctica ocurre que este mínimo que se pretende alcanzar es local o algunas veces, ni siquiera se llega a alcanzar.

Por lo tanto, para la búsqueda no lineal de parámetros es necesario un conjunto de medidas $\mathbf{b}^\#$, un valor inicial \mathbf{a}_0 para el vector de parámetros que se desea estimar, y una función f cuyos parámetros deben satisfacer el conjunto de medidas dado $\mathbf{b} = f(\mathbf{a})$. El vector de parámetros y las funciones de coste a plantear en el caso de calibración de la cámara se tratan a continuación. Estas funciones tratan de minimizar diferentes errores vistos en el apartado anterior. Las condiciones para tener éxito en la búsqueda no lineal de los parámetros son partir de unos valores iniciales de los parámetros \mathbf{a}_0 bastante aproximados a los valores reales, que la función f que relaciona los parámetros con el conjunto de medidas se ajuste lo mejor al comportamiento de los mismos, y que además, el conjunto de medidas estén poco o nada contaminadas con ruido.

3.4.3. Funciones de coste

El algoritmo de estimación de parámetros descrito en el apartado anterior se utiliza para la estimación no lineal de los elementos de la matriz de proyección. Se trata de minimizar el error definido por $\mathbf{e} = \mathbf{b}^\# - f(\mathbf{a})$, donde \mathbf{a} representa el vector de parámetros que satisfacen el vector de medidas $\mathbf{b}^\#$ según la función f . Básicamente todas las funciones de coste que se van a presentar tratan de minimizar el error geométrico entre las coordenadas de los puntos de la imagen $\mathbf{q}_i^\#$ y las coordenadas resultado de proyectar los puntos de la plantilla con los parámetros de la cámara estimados $\mathbf{q}_i^* = \mathbf{M}^\# \cdot \mathbf{p}_i^\#$. La definición de la función f se puede realizar de diferentes formas. El objetivo de este apartado es presentar diferentes funciones en las que cada una de ellas representa la relación entre los puntos de la plantilla y los puntos de la imagen de diferente forma.

Matriz de proyección

En este caso el vector de parámetros está formado por los elementos de la matriz de proyección \mathbf{M} , $\mathbf{a} = (m_{11}, m_{12}, \dots, m_{34})$. El vector de medidas corresponde a las coordenadas de los puntos en la plantilla \mathbf{p}_i . La función $f(\mathbf{a})$ viene definida por la siguiente expresión:

$$f(m_{11}, \dots, m_{34}) = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \frac{m_{11}x_i + m_{12}y_i + m_{13}z_i + m_{14}}{m_{31}x_i + m_{32}y_i + m_{33}z_i + m_{34}} \\ \frac{m_{21}x_i + m_{22}y_i + m_{23}z_i + m_{24}}{m_{31}x_i + m_{32}y_i + m_{33}z_i + m_{34}} \end{bmatrix} \quad (3.69)$$

El índice a minimizar se representa según:

$$I = \sum_{i=1}^n d_{eucli}(\mathbf{q}_i^\#, \mathbf{M}^\# \cdot \mathbf{p}_i^\#)^2 \quad (3.70)$$

Matriz de proyección y restricciones de la matriz de proyección

En este caso el vector de parámetros \mathbf{a} y la función f son las mismas que en el caso anterior. La diferencia está en que además, se intenta también que la matriz de proyección calculada cumpla con las restricciones (2.28) y (2.29) obtenidas en el apartado 2.5. El vector de errores está compuesto por los errores geométricos de cada uno de los puntos \mathbf{e}_{geo} y el valor de la restricción \mathbf{e}_{res} .

$$\mathbf{e} = \begin{pmatrix} \mathbf{e}_{geo} & \mathbf{e}_{res} \end{pmatrix} \quad (3.71)$$

El error geométrico se calcula de la misma forma que se hizo en el apartado anterior. El error en el valor de la restricción será el valor de la segunda restricción que debe cumplir la matriz de proyección $(\mathbf{m}_1 \wedge \mathbf{m}_3) \cdot (\mathbf{m}_2 \wedge \mathbf{m}_3)$ siendo \mathbf{m}_1 , \mathbf{m}_2 , \mathbf{m}_3 los vectores formados por las filas de la matriz de proyección (ver apartado 2.5). Teniendo en cuenta que la matriz de proyección es homogénea y está definida con un factor de escala, es necesario escalar el valor de la segunda restricción con este valor de escala. Este factor de escala viene determinado por el módulo del vector \mathbf{m}_3 de forma que se tiene que conseguir que éste sea unitario $|\mathbf{m}_3|=1$. Ésta es la primera restricción que tiene que cumplir la matriz de proyección. Por este motivo, el valor de la segunda restricción se ve afectada por el valor de la primera restricción. De esta forma para garantizar que ambas restricciones se cumplen, la función no lineal que expresa el valor de la restricción en función de los elementos de la matriz de proyección es la siguiente:

$$R = \frac{(\vec{m}_1 \wedge \vec{m}_3) \cdot (\vec{m}_2 \wedge \vec{m}_3)}{(\vec{m}_3 \cdot \vec{m}_3^T)^2} = 0 \quad (3.72)$$

Formalmente, el índice de coste que se plantea para obtener los valores de los elementos de la matriz de proyección, en este caso el índice es:

$$I = \sum_{i=1}^n d_{eucli} \left(q_i^{\#}, M_{p^{\#}q^{\#}}^* \cdot p_i^{\#} \right)^2 + \alpha \cdot R \quad (3.73)$$

Para ponderar el valor de la restricción dentro del índice y garantizar el cumplimiento de la misma se introduce el factor α .

Parámetros de la cámara

Hasta este momento se ha planteado la calibración de la cámara en dos pasos. En el primer paso se estiman los elementos de la matriz de proyección, y seguidamente se extraen los parámetros del modelo de la cámara a partir de la misma. Mediante la técnica de búsqueda no lineal de los elementos de la matriz de proyección, es posible estimar directamente los parámetros del modelo de la cámara, si se plantea la función no lineal que relaciona las medidas de las coordenadas de los puntos de interés con los parámetros de la cámara. En este caso los parámetros corresponden a los del modelo de la cámara $\mathbf{a} = (\alpha_u, \alpha_v, u_0, v_0, t_x, t_y, t_z, \theta, \phi, \psi)$. Los

parámetros θ , ϕ , ψ , representan los ángulos de rotación del sistema de referencia de la cámara respecto de la del escenario. Estos ángulos dan como resultado la matriz de rotación \mathbf{R} . De esta forma la función no lineal que relaciona las medidas de los puntos con los parámetros del modelo de la cámara es:

$$f(\alpha_u, \alpha_v, u_0, v_0, \mathbf{R}, t_x, t_y, t_z) = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \alpha_u \cdot \frac{r_{11}x_i + r_{12}y_i + r_{13}z_i + t_x}{r_{31}x_i + r_{32}y_i + r_{33}z_i + t_z} + u_0 \\ \alpha_v \cdot \frac{r_{21}x_i + r_{22}y_i + r_{23}z_i + t_y}{r_{31}x_i + r_{32}y_i + r_{33}z_i + t_z} + v_0 \end{bmatrix} \quad (3.74)$$

El índice a minimizar en este caso es el siguiente:

$$I = \sum_{i=1}^n \left(u_i - \alpha_u \frac{r_{11}x_i + r_{12}y_i + r_{13}z_i + t_x}{r_{31}x_i + r_{32}y_i + r_{33}z_i + t_z} - u_0 \right)^2 + \left(v_i - \alpha_v \frac{r_{21}x_i + r_{22}y_i + r_{23}z_i + t_y}{r_{31}x_i + r_{32}y_i + r_{33}z_i + t_z} - v_0 \right)^2 \quad (3.75)$$

En un principio se tiene que minimizar un índice que depende de 4+9 parámetros, de los cuales 4 corresponden a los parámetros intrínsecos y 9 a los extrínsecos, 6 a los elementos de la matriz de rotación y 3 al vector de traslación. Sin embargo es posible reducir las dimensiones del espacio de búsqueda si se expresa la matriz de rotación en función de los ángulos de Euler θ , ϕ , ψ , los cuales representan los giros que se realizan respecto a tres ejes. De esta forma la dimensión del espacio de búsqueda es 4+6. Los elementos de la matriz de rotación obtenidos a partir de los ángulos de Euler son los siguientes:

$$R = \begin{bmatrix} \cos\psi \cos\phi - \cos\theta \sin\phi \sin\psi & \cos\psi \sin\phi + \cos\theta \cos\phi \sin\psi & \sin\theta \sin\psi \\ -\sin\psi \cos\phi - \cos\theta \sin\phi \cos\psi & -\sin\psi \sin\phi + \cos\theta \cos\phi \cos\psi & \sin\theta \cos\psi \\ \sin\theta \sin\phi & -\sin\theta \cos\phi & \cos\theta \end{bmatrix} \quad (3.76)$$

Parámetros de la cámara y el conjunto de puntos en la plantilla 3D

Desde el punto de vista que los puntos que se miden en la plantilla $\mathbf{p}_i^\#$ están contaminados con ruido se realiza una búsqueda de los parámetros de la cámara junto con el conjunto de puntos en la plantilla $\mathbf{p}_i^\#$ que mejore el error geométrico. En teoría, este nuevo conjunto de puntos estimados \mathbf{p}_i^* , tendrán un menor nivel de ruido que los iniciales $\mathbf{p}_i^\#$ ya que consiguen minimizar el error geométrico. De esta forma, para mejorar las coordenadas de los puntos en la plantilla $\mathbf{p}_i^\#$, éstas se incluyen junto con los parámetros de la cámara en el vector de elementos a estimar. La búsqueda en este caso se hará sobre los parámetros de la cámara y las coordenadas de los puntos $\mathbf{a} = (\alpha_u, \alpha_v, u_0, v_0, t_x, t_y, t_z, \theta, \phi, \psi, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$. La dimensión de búsqueda es 4+6+3n, siendo n el número de puntos.

Parámetros de la cámara y el conjunto de puntos en la plantilla 2D desde diferentes posiciones

En este caso se trata de utilizar toda la información disponible si se está utilizando el método de Zhang [104] para calibrar la cámara. Con el método de Zhang [104] la plantilla es plana y se toman varias imágenes de la misma desde diferentes posiciones y orientaciones. Para utilizar toda la información se plantea un índice que contiene los parámetros intrínsecos de la cámara, los extrínsecos correspondientes a todas las imágenes que se han tomado de la cámara y las coordenadas de los puntos en la plantilla, las cuales también se modifican. El vector de parámetros es $\mathbf{a}=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$. m es el número de imágenes tomadas de la plantilla y n corresponde al número de puntos de interés en la plantilla. La dimensión de búsqueda es $4+6m+3n$. El índice queda como:

$$I = \sum_{j=1}^m \sum_{i=1}^n \left(u_{ij} - \alpha_u \frac{r_{11j}x_i + r_{12j}y_i + r_{13j}z_i + t_{xj}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{zj}} - u_0 \right)^2 + \left(v_{ij} - \alpha_v \frac{r_{21j}x_i + r_{22j}y_i + r_{23j}z_i + t_{yj}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{zj}} - v_0 \right)^2 \quad (3.77)$$

3.5. Mejoras en la estimación lineal de los parámetros

En el apartado 3.4.2 se describen técnicas de optimización no lineal y de las condiciones que se tienen que cumplir para encontrar los parámetros óptimos \mathbf{p}^* que satisfacen la función f para un conjunto de medidas \mathbf{x}' . La conclusión obtenida en este apartado es que se necesitan unos valores iniciales de los parámetros \mathbf{p}_0 bastante cercanos a los definitivos y que además el ruido que contaminaba las medidas $\mathbf{x}^\#$ no debe ser excesivo. En el caso de la estimación de los parámetros de la cámara utilizando una plantilla bidimensional, los datos utilizados para la estimación son el conjunto de puntos en la plantilla $\mathbf{p}_i^\#$, junto con sus correspondientes en las imágenes $\mathbf{q}_i^\#$. Los valores iniciales de los parámetros de la cámara son los obtenidos de forma lineal según el método desarrollado por Zhang [104] explicado en el apartado 3.3. Con el objetivo de mejorar la estimación lineal de los parámetros de la cámara, que sirven como punto inicial de búsqueda no lineal de los mismos, se van a utilizar dos técnicas desarrolladas por Hartley [32] y por Ahmed [3],[4] o Devernay [19] que permiten un preprocesado de los datos antes de utilizarlos en el proceso de estimación.

Hartley realiza una normalización de las medidas de las coordenadas. Esta técnica permite obtener una matriz \mathbf{A} (3.41) mejor condicionada y en consecuencia obtener unos valores más estables en el cálculo de los vectores propios de dicha matriz. Según se explica en 3.3.1 la matriz \mathbf{A} (3.41) contiene las medidas de las coordenadas de los puntos.

Ahmed [3],[4] o Devernay [19] realizan una estimación de la distorsión que producen las lentes en la imagen y una posterior corrección de la misma. Esta técnica permite reducir el nivel

de ruido en las medidas de las coordenadas de la imagen, ya que se corrigen las posiciones de los puntos en la misma y además, los parámetros estimados para la distorsión producida por la lente sirven de punto de partida para la búsqueda no lineal. En el caso de establecer únicamente una búsqueda lineal de los parámetros, los valores iniciales de partida para la búsqueda no lineal de los parámetros que modelan la distorsión de la imagen es 0. Con esta técnica es posible partir de unos valores más aproximados para estos parámetros.

El objetivo es introducir estos dos pasos previos antes de realizar el proceso de calibrado. En el apartado de resultados experimentales se realizan simulaciones para ver el efecto de estas modificaciones en el proceso de calibración de la cámara. A continuación se explican estas dos técnicas.

3.5.1. Corrección de la distorsión en la imagen

En este caso, la mejora que se plantea en el proceso de calibración de la cámara consiste en realizar una estimación previa de la distorsión que introducen las lentes en el proceso de formación de la imagen. Esto permite tanto corregir los puntos de la imagen para tener un menor nivel de ruido en las medidas, así como disponer de una estimación inicial de los parámetros que modelan la distorsión para iniciar la búsqueda no lineal.

En el apartado 2.8 se establecían diferentes tipos de distorsiones, cada una de las cuales se modelaban con infinitos términos. Sin embargo, desde un punto de vista práctico, sólo dos o tres parámetros deben ser tenidos en cuenta a la hora de determinar la distorsión que producen las lentes en la imagen. Los términos de orden superior son comparativamente insignificantes y sólo sirven para complicar el modelo [94], [98], [100]. Además, para la mayoría de aplicaciones de visión artificial, no es necesario considerar el término de descentrado o distorsión tangencial ya que éste se puede considerar despreciable si se tienen en cuenta los demás [13], [94]. Si los parámetros que definen la distorsión de la lente de la cámara son conocidos a priori, la corrección de la distorsión es fácilmente aplicable. Desafortunadamente, esta información no está disponible. Además, en una producción en serie de lentes, cada una de ellas varía sus características con respecto a las otras. También es necesario tener en cuenta que los parámetros que modelan la distorsión varían con el enfoque de la cámara. Por lo tanto, es necesario definir un método para estimar los parámetros que modelan la distorsión.

Normalmente los parámetros que modelan la distorsión son estimados en la fase no lineal del proceso de calibrado, al mismo tiempo que se calibran los demás parámetros intrínsecos y extrínsecos. El problema de este método es que debido al acoplamiento que existe entre los parámetros intrínsecos y los extrínsecos de cámara, esta búsqueda general de todos los parámetros desemboca en errores de la estimación de los parámetros intrínsecos de la cámara.

Para realizar la estimación de la distorsión de las lentes previa al proceso de calibrado existen diferentes técnicas las cuales se basan generalmente en características geométricas invariantes de los puntos en la imagen. Los métodos de Ahmed [3],[4] o Devernay [19] basan la estimación de la distorsión de la lente en el análisis de las líneas rectas en la imagen. De esta forma se puede medir la distorsión en la imagen para poder ser calculada utilizando técnicas de búsqueda no lineal. En este caso se expresa la distorsión de la imagen en función de la deformación de las líneas rectas en la misma. Dado que la plantilla de calibración está basada en un tablero de ajedrez, la detección de las líneas en la imagen resulta obvia.

El modelo para la distorsión radial y tangencial transforma las coordenadas medidas de la imagen $\mathbf{q}_d=(u_d, v_d)$, las cuales han sufrido la distorsión de las lentes, en las coordenadas ideales libres de distorsión $\mathbf{q}_p=(u_p, v_p)$, las cuales son desconocidas. Las expresiones que modelan estas distorsiones se presentan en el apartado 2.8. En este caso, se van a modelar la distorsión radial y tangencial o de descentrado de la imagen. Por lo tanto, utilizando (2.73), (2.75) la relación entre las coordenadas será:

$$\begin{aligned} u_p &= u_d + \Delta u \cdot (k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6 + \dots) + p_1 (r^2 + 2 \cdot \Delta u^2) + 2 p_2 \cdot \Delta u \cdot \Delta v + \dots \\ v_p &= v_d + \Delta v \cdot (k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6 + \dots) + 2 p_1 \cdot \Delta u \cdot \Delta v + p_2 (r^2 + 2 \cdot \Delta v^2) + \dots \end{aligned} \quad (3.78)$$

r es la distancia del píxel, al punto principal de la imagen $\mathbf{c}=(u_0, v_0)$, $\Delta u = u_d - u_0$, $\Delta v = v_d - v_0$.

$$r = \sqrt{\Delta u^2 + \Delta v^2} \quad (3.79)$$

Los coeficientes que modelan la distorsión radial son k_1, k_2, k_3, \dots y p_1, p_2, p_3, \dots son los coeficientes que modelan la distorsión tangencial.

Si se tiene una línea recta en el plano de la imagen que pasa por un píxel de la misma, en cada punto de esa misma recta se satisface la expresión $a \cdot u_p + b \cdot v_p + c = 0$ donde a, b, c son constantes para una determinada línea, siendo $-a/b$ la pendiente de la misma. Cada punto $\mathbf{q}_p=(u_p, v_p)$ en la recta tiene relación con un punto $\mathbf{q}_d=(u_d, v_d)$ en el plano de la imagen distorsionado dado por (3.78). Esto significa que las coordenadas del punto en la recta son función de $\mathbf{q}_d=(u_d, v_d)$. Por lo tanto, la ecuación de la recta se puede escribir como:

$$f(u_d, v_d) = a \cdot u_p(u_d, v_d) + b \cdot v_p(u_d, v_d) + c = 0 \quad (3.80)$$

$f(u_d, v_d)$ describe la ecuación de la curva correspondiente en el plano distorsionado de la imagen. La derivada de f en cualquier punto distorsionado de la imagen $\mathbf{q}_d=(u_d, v_d)$ dará la pendiente de la curva en ese punto. Esta pendiente en ese punto de la curva corresponde a la pendiente teórica de la recta sin distorsión. Esta pendiente se expresa como:

$$\delta f = a \cdot \left(\frac{\delta u_p}{\delta u_d} \delta u_d + \frac{\delta u_p}{\delta v_d} \delta v_d \right) + b \cdot \left(\frac{\delta v_p}{\delta u_d} \delta u_d + \frac{\delta v_p}{\delta v_d} \delta v_d \right) = 0 \quad (3.81)$$

Las derivadas parciales se calculan de la expresión (3.78). Arreglando esta última expresión, si la pendiente de la recta que en el plano sin distorsión era $-a/b$, en este caso está relacionada con la pendiente de la recta tangente $\delta v_d/\delta u_d$ de la curva en el punto $\mathbf{q}_d=(u_d, v_d)$ mediante:

$$s(u_d, v_d) = \frac{\frac{\delta v_p}{\delta u_d} + \frac{\delta v_p}{\delta v_d} \frac{\delta v_d}{\delta u_d}}{\frac{\delta u_p}{\delta u_d} + \frac{\delta u_p}{\delta v_d} \frac{\delta v_d}{\delta u_d}}; \quad s = -\frac{a}{b} \quad (3.82)$$

Si se tienen los valores exactos de los parámetros que modelan la distorsión, las pendientes calculadas con la expresión (3.82) en todos los puntos de la imagen distorsionada, tienen que ser la misma si los puntos pertenecen a la misma recta. Este coeficiente permite definir una medida de distorsión. Dado un conjunto N de puntos $\mathbf{q}_{di}=(u_{di}, v_{di})$, $i=1\dots N$ en el plano de la imagen distorsionada que pertenecen a la misma recta, se calculan las pendientes de las tangentes en cada punto y entonces se pueden obtener los parámetros que modelan la distorsión minimizando el siguiente error:

$$e_s = \sum_{i=2}^N (s(u_{di}, v_{di}) - s(u_{di-1}, v_{di-1})) \quad (3.83)$$

Es obvio que este error será cero si todos los puntos están situados en la misma línea. Si varias curvas de la imagen son utilizadas, el error será la suma de todas las rectas utilizadas. Para calibrar la distorsión, la medida de la distorsión de la última expresión, puede ser minimizada por métodos no lineales, tales como Levenberg-Marquardt. En el punto inicial de búsqueda, el centro de distorsión será el centro de la imagen (u_0, v_0) . Los demás parámetros que modelan la distorsión se inicializarán a cero. En el caso de asumir el centro de la imagen como centro de distorsión, se puede plantear una solución lineal al problema de estimación de los parámetros k_1, k_2, k_3, \dots y p_1, p_2, p_3, \dots . Para obtener esta solución lineal es necesario que el número de parámetros a estimar sea ≤ 2 . En este caso, las derivadas parciales usadas en la expresión (3.82) asumiendo solamente distorsión radial y tangencial de primer orden, son las siguientes expresiones:

$$\begin{aligned} \frac{\delta v_p}{\delta u_d} &= 2 \cdot k_1 \cdot \Delta u \cdot \Delta v + 2 \cdot p_1 \cdot \Delta v & \frac{\delta v_p}{\delta v_d} &= 1 + 3 \cdot k_1 \cdot \Delta v^2 + k_1 \cdot \Delta u^2 + 2 \cdot p_1 \cdot \Delta u \\ \frac{\delta u_p}{\delta u_d} &= 1 + 3 \cdot k_1 \cdot \Delta u^2 + k_1 \cdot \Delta v^2 + 6 \cdot p_1 \cdot \Delta u & \frac{\delta u_p}{\delta v_d} &= 2 \cdot k_1 \cdot \Delta u \cdot \Delta v + 2 \cdot p_1 \cdot \Delta v \end{aligned} \quad (3.84)$$

Todas estas expresiones son lineales para los parámetros que modelan la distorsión k_i y p_i en cualquier punto. Utilizando (3.82) y (3.84) se calcula el valor teórico de la pendiente de la recta tangente a la curva en cada punto con distorsión $\mathbf{q}_{di}=(u_{di}, v_{di})$. Esta pendiente $s_i(u_{di}, v_{di})$ dada por cada punto con distorsión, corresponde a una pendiente s teórica desconocida de la recta que formarían los puntos sin distorsión. Esta pendiente s puede ser estimada con los

puntos de la curva por mínimos cuadrados. Siendo conocida la pendiente de la recta tangente a la curva en cada punto, para cada punto se puede obtener una ecuación lineal en función de los coeficientes de la distorsión. Todas las ecuaciones pueden ser ordenadas de la forma $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ donde \mathbf{A} es una matriz de dimensiones $2 \times N$, \mathbf{x} es un vector desconocido de los parámetros que modelan la distorsión y \mathbf{b} es un vector conocido de longitud N . Este sistema sobredeterminado puede ser resuelto mediante descomposición en valores singulares. Si por ejemplo se desea estimar los valores de k_1 y p_1 las ecuaciones son las siguientes:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ \dots & \dots \\ \dots & \dots \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} k_1 \\ p_1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} s_1 - s \\ \dots \\ \dots \end{bmatrix} \quad (3.85)$$

Los elementos de estas matrices son:

$$\begin{aligned} a_{11} &= 2 \cdot \Delta x_1 \cdot \Delta v_1 + s(3 \cdot \Delta v_1^2 + \Delta u_1^2) - s_1(3 \cdot \Delta u_1^2 + \Delta v_1^2 + 2 \cdot s \cdot \Delta u_1 \cdot \Delta v_1) \\ a_{12} &= 2 \cdot \Delta v_1 + 2 \cdot s \cdot \Delta x_1 - s_1(6 \cdot \Delta u_1 + 2 \cdot s \cdot \Delta v_1) \end{aligned} \quad (3.86)$$

s_i representa la pendiente de la recta tangente a la curva en el punto i , y s la pendiente de dicha recta sin distorsión. La matriz \mathbf{A} se forma con cada uno de los puntos que forma la recta distorsionada. Si se tienen más rectas en la imagen distorsionada, éstas se deben añadir a la matriz \mathbf{A} en filas sucesivas ya que los parámetros que modelan la distorsión son los mismos para todas las rectas. En cualquier caso, se puede hacer una aproximación lineal de los parámetros y posteriormente utilizar los resultados para minimizar el error de distorsión utilizando Levenberg-Marquardt tal y como se explica en el apartado de búsqueda no lineal.

3.5.2. Normalización de los datos

Desde el punto de vista de la robustez de los algoritmos de estimación de la matriz de proyección \mathbf{M} , es deseable que los resultados que se obtienen sean invariables, independientemente de la situación de los puntos en el espacio. Es decir, se desea que el resultado obtenido de la matriz de proyección con un conjunto de puntos sea el mismo que el que se obtiene con otro conjunto de puntos sabiendo la relación que existe entre los mismos. Formalmente se podría decir que se tiene un conjunto de puntos \mathbf{p}_i en la escena, con sus correspondientes puntos \mathbf{q}_i en la imagen, las posiciones de estos puntos cambian de forma que $\mathbf{p}_i^o = \mathbf{T}_p \cdot \mathbf{p}_i$ y $\mathbf{q}_i^o = \mathbf{T}_q \cdot \mathbf{q}_i$, donde \mathbf{T}_p y \mathbf{T}_q son sendas matrices que representan transformaciones de similitud. Si los dos conjuntos de puntos iniciales están relacionados por la matriz de proyección de la forma $\mathbf{q}_i = \mathbf{M} \cdot \mathbf{p}_i$, esta expresión se transforma utilizando el nuevo conjunto de puntos en $\mathbf{q}_i^o = \mathbf{T}_q \cdot \mathbf{M} \cdot \mathbf{T}_p^{-1} \cdot \mathbf{p}_i^o$. Esta ecuación implica que la matriz de proyección \mathbf{M}^o que relaciona los conjuntos de puntos transformados \mathbf{p}_i^o y \mathbf{q}_i^o es $\mathbf{M}^o = \mathbf{T}_q \cdot \mathbf{M} \cdot \mathbf{T}_p^{-1}$.

Teniendo esto en cuenta, si la situación de los puntos en el espacio no influye en la estimación de la matriz de proyección \mathbf{M} se puede obtener la misma por un camino alternativo. Éste consistiría en transformar el conjunto de puntos iniciales \mathbf{p}_i y \mathbf{q}_i en $\mathbf{p}_i^o = \mathbf{T}_p \cdot \mathbf{p}_i$ y $\mathbf{q}_i^o = \mathbf{T}_q \cdot \mathbf{q}_i$ a partir de sendas transformaciones de similitud \mathbf{T}_p y \mathbf{T}_q . Después se calcula la matriz de proyección \mathbf{M}^o por los métodos explicados anteriormente basados en la minimización del error algebraico y a partir de ésta, obtener la matriz de proyección \mathbf{M} según $\mathbf{M} = \mathbf{T}_q^{-1} \cdot \mathbf{M}^o \cdot \mathbf{T}_p$, la cual, se hubiera obtenido directamente utilizando el conjunto de puntos iniciales \mathbf{p}_i y \mathbf{q}_i .

Este hecho sin embargo, en la realidad no es cierto. Hartley [32], [33] demuestra que los resultados de los algoritmos basados en la minimización del error algebraico, son sensibles a la situación de los puntos en el espacio. Esto es debido a que en el momento de establecer las restricciones dentro del índice no es lo mismo fijar $|\mathbf{m}|=1$ que $|\mathbf{m}^o|=1$ y por lo tanto los resultados obtenidos por un camino son diferentes a los obtenidos por el camino inmediato. En cambio si se utilizan métodos basados en la optimización del error geométrico, esto no ocurre así, ya que las transformaciones de similitud que se realizan en el conjunto de puntos, no afectan a las distancias relativas de los mismos en el espacio.

Partiendo de este resultado y aprovechando la particularidad de que los resultados de los algoritmos basados en la minimización del error algebraico son sensibles a la posición de los puntos en el espacio, es interesante obtener la situación óptima de los puntos \mathbf{p}_i y \mathbf{q}_i que haga que el error algebraico obtenido en la estimación de la matriz de proyección sea el menor posible. Una posibilidad es realizar una transformación de similitud \mathbf{T}_p y \mathbf{T}_q del conjunto de puntos que se dispone para estimar la matriz de proyección la cual consiste en una traslación y un escalado. Con el nuevo conjunto de puntos \mathbf{p}_i^o y \mathbf{q}_i^o , se realiza la estimación de la matriz \mathbf{M}^o y finalmente se obtiene la matriz \mathbf{M} definitiva mediante la expresión $\mathbf{M} = \mathbf{T}_q^{-1} \cdot \mathbf{M}^o \cdot \mathbf{T}_p$. Esta normalización de las coordenadas a parte de mejorar los resultados de la estimación, consigue que los resultados no sean sensibles a la situación de los puntos en el escenario. Las coordenadas de los puntos de partida para la calibración cumplirán unas características que siempre serán las mismas independientemente del sistema de referencia elegido.

La normalización de las coordenadas que Hartley propone en [32] y [33], consiste en trasladar las mismas de forma que el centroide de todos los puntos esté situado en el origen de coordenadas del sistema de referencia. La traslación que se realiza es diferente en las coordenadas de los puntos de la plantilla, que en los puntos de la imagen. Además las coordenadas son escaladas de forma que la media de las distancias de todos los puntos al origen de coordenadas sea 1. El escalado en cada eje de coordenadas es igual y se realiza un escalado diferente para los puntos de la imagen y del escenario.

Esta normalización de los datos es imprescindible utilizarla en la estimación de la matriz de proyección con técnicas basadas en la minimización de un error algebraico. A parte de

proporcionar un método de calibración independiente del sistema de coordenadas, consigue mejorar la condición de la matriz \mathbf{A} . En el caso de tener sistemas mal condicionados, el ruido que contamina la medida de las coordenadas afecta bastante a los resultados. Con la normalización se consiguen sistemas bien condicionados cuya resolución es más estable y menos sensible al ruido.

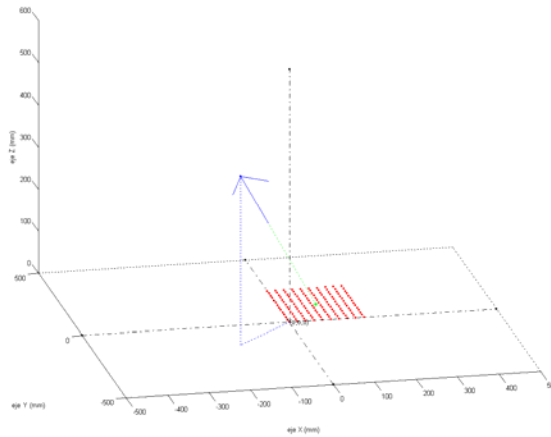
3.6. Método de calibración de la cámara

En este apartado del capítulo de calibración se va a definir un método de calibración con el que se obtienen los mejores resultados. La definición de este método de calibración se basa tanto en los estudios realizados por otros autores Salvi [76] y Sun [82] así como en los realizados por el propio autor. Teniendo en cuenta los estudios realizados por Salvi [76] y Sun [82] el método de calibración que mejores resultados obtiene es el que se utiliza una plantilla bidimensional presentado por Zhang [104]. Respecto al modelo que mejor se ajusta al comportamiento de la cámara, estos autores defienden el modelo lineal presentado en el capítulo 2, y la parte del modelo no lineal que tiene en cuenta una distorsión radial de segundo orden considerando la distorsión tangencial ó desalineamiento nulo.

Partiendo de esta premisa, se va a definir un método de calibración basado en una plantilla bidimensional añadiendo diferentes mejoras. Antes de aplicar el método de calibración, se va a realizar un tratamiento de los datos disponibles para la calibración. Primero se va a calibrar la distorsión de la imagen producida por las lentes. Esto obtendrá una reducción significativa del nivel de ruido en las medidas de las coordenadas de la imagen. En segundo lugar, una vez corregida la distorsión en las coordenadas de los puntos, se procederá a normalizar los mismos para que la solución lineal de los parámetros sea más cercana a la exacta. Con las coordenadas de los puntos en la imagen corregidas y normalizadas, se realizará el proceso de calibración descrito en el apartado 3.3, para obtener la solución lineal de los parámetros de la cámara. Posteriormente, se realizará una búsqueda no lineal de los mismos, partiendo de la solución obtenida en el paso anterior. En este caso el error geométrico a minimizar contemplará la estimación de las posiciones de los puntos en la plantilla. Esto garantiza que el ruido que se tiene en las medidas de las coordenadas de los puntos de la platilla no afecta al proceso de calibrado tal y como demuestra Lavest en [50].

3.7. Resultados experimentales

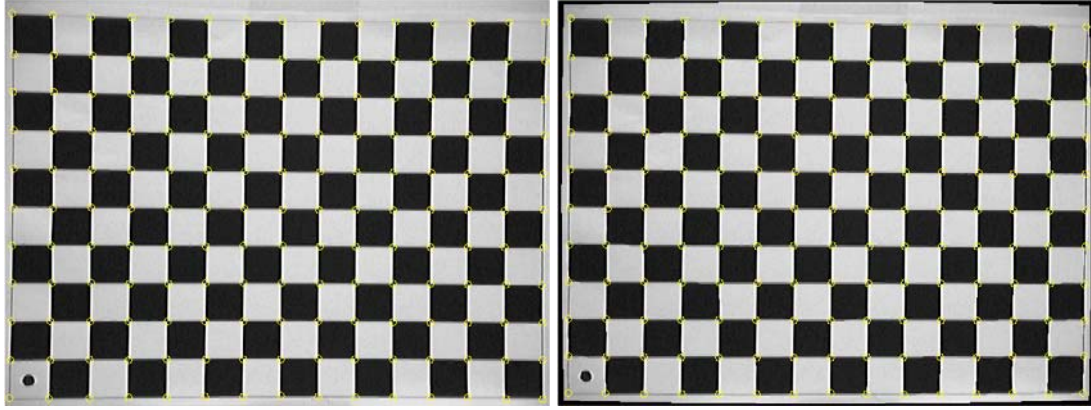
A partir del método de calibración definido en el apartado anterior, se procede a realizar diferentes experimentos para determinar que efectivamente, el método elegido presenta



3.4 Situación de la cámara en el escenario de simulación

mejoras respecto al procedimiento existente basado en plantilla bidimensional. Los experimentos se realizan tanto con datos simulados como con imágenes reales para cada una de las etapas. Para examinar el efecto del tratamiento previo de los datos se realiza la calibración de la cámara utilizando datos simulados en diferentes condiciones, es decir, datos a los que se le ha corregido la distorsión de la imagen, datos solamente normalizados y datos sin distorsión y con normalización. De esta forma se puede definir en que casos es necesario el preproceso de los datos en el caso de realizar una estimación lineal de los parámetros de la cámara. Desde el punto de vista de la estimación no lineal de los parámetros se implantarán los diferentes índices descritos en el apartado 3.4.3, y se tomará como datos de partida para la búsqueda no lineal, las estimaciones lineales realizadas con los diferentes preprocesados de los datos.

Para fijar las condiciones bajo las cuales se van a realizar los experimentos con datos simulados se parte de los resultados obtenidos por Zhang en [103]. Según se muestra en la figura 3 de [103], los resultados de calibración mejoran si los ángulos que forman el plano de la platilla con el los planos de las imágenes está comprendido entre 50 y 70 grados. Respecto al número de imágenes necesarias para realizar una calibración fiable la figura 2 de [103] muestra que a partir de 7 u 8 imágenes la mejora en los resultados de la calibración ya no es significativa. Por lo tanto, para ver el efecto de la normalización de los datos y la corrección de la distorsión en el proceso de calibración de la cámara, se van a utilizar 8 imágenes tomadas de forma que el ángulo que forman el plano de la platilla con los respectivos planos de la imagen está comprendido entre 50 y 70 grados. Las posiciones de la cámara son $\mathbf{t}_1=(150, 200, 580)$, $\mathbf{t}_2=(-50, 250, 880)$, $\mathbf{t}_3=(100, -20, 820)$, $\mathbf{t}_4=(-40, -150, 780)$, $\mathbf{t}_5=(-150, -150, 530)$, $\mathbf{t}_6=(-100, 125, 400)$, $\mathbf{t}_7=(140, -150, 500)$, $\mathbf{t}_8=(240, 50, 600)$. La orientación de la cámara en cada una de las posiciones es tal que el eje óptico corta al plano de la plantilla en el centro de la misma. Además, la orientación también está calculada para que el eje de coordenadas sea paralelo al plano en el que se encuentra la platilla. Esto intenta simular que la plantilla está situada en el suelo, y las



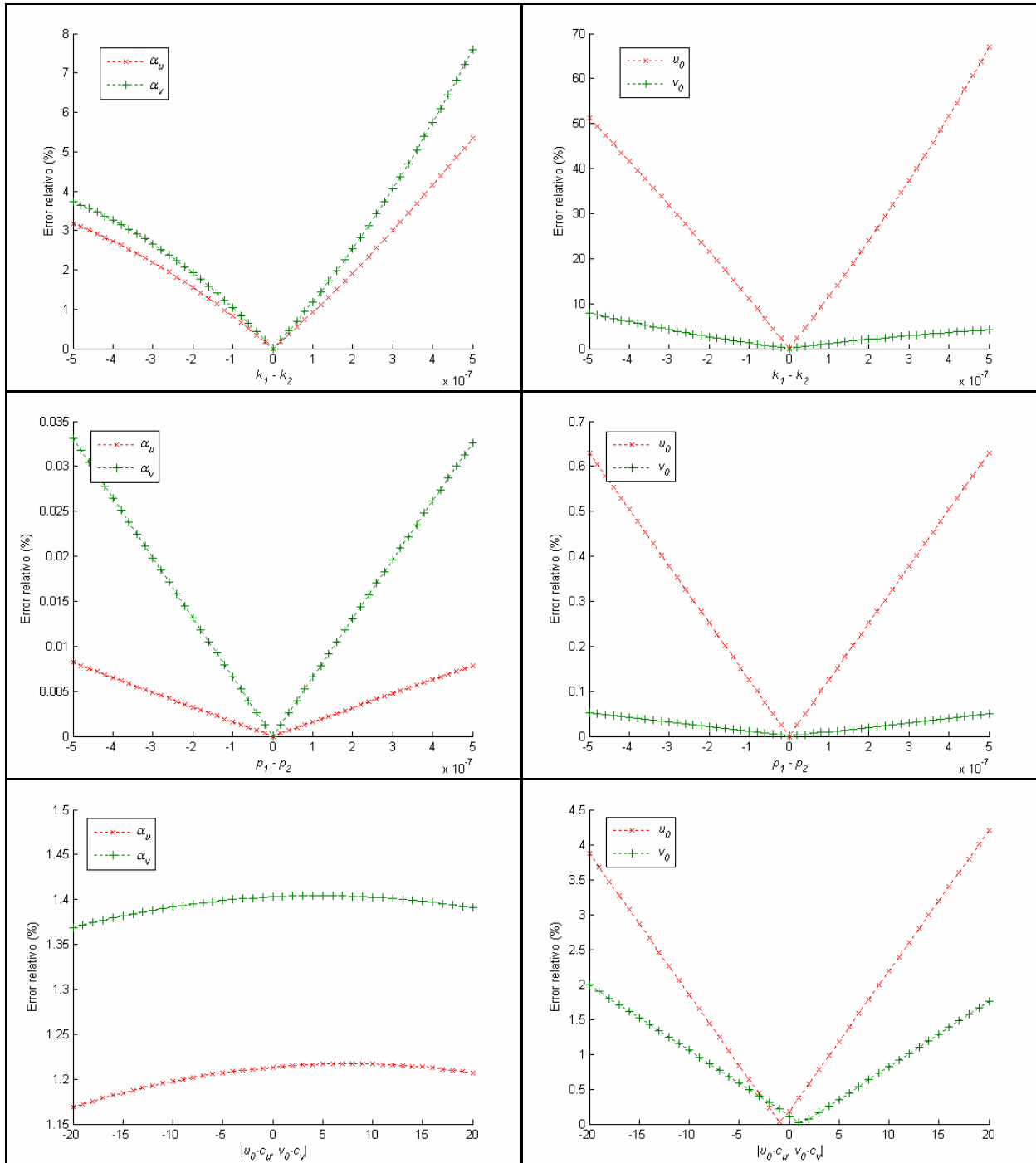
3.5 Corrección de la distorsión en la imagen capturada

imágenes son tomadas con la cámara utilizando un trípode en las posiciones indicadas. En concreto, la situación de la cámara en el escenario de simulación propuesta para la posición t_5 es la que se muestra en la figura 3.4. Por otro lado, con el objetivo de poder comparar resultados simulados con los obtenidos por Zhang en [103], las características de la cámara son las mismas que la simulada en el mencionado artículo. Éstas son $\alpha_u=1250$, $\alpha_v=900$, $u_0=250$, $v_0=250$ y $\gamma=1.09083$ equivalente a 89.95° . La resolución de la imagen es 512×512 píxeles. La plantilla, de la misma forma que hace Zhang es un tablero de ajedrez de $10 \times 14 = 140$ esquinas. El tamaño de la plantilla es 180×250 mm. En este caso, la plantilla está situada de forma que todos los puntos de interés tienen coordenada $z_w=0$, tal y como se muestra en la figura 3.4.

Estimación de la distorsión de la cámara

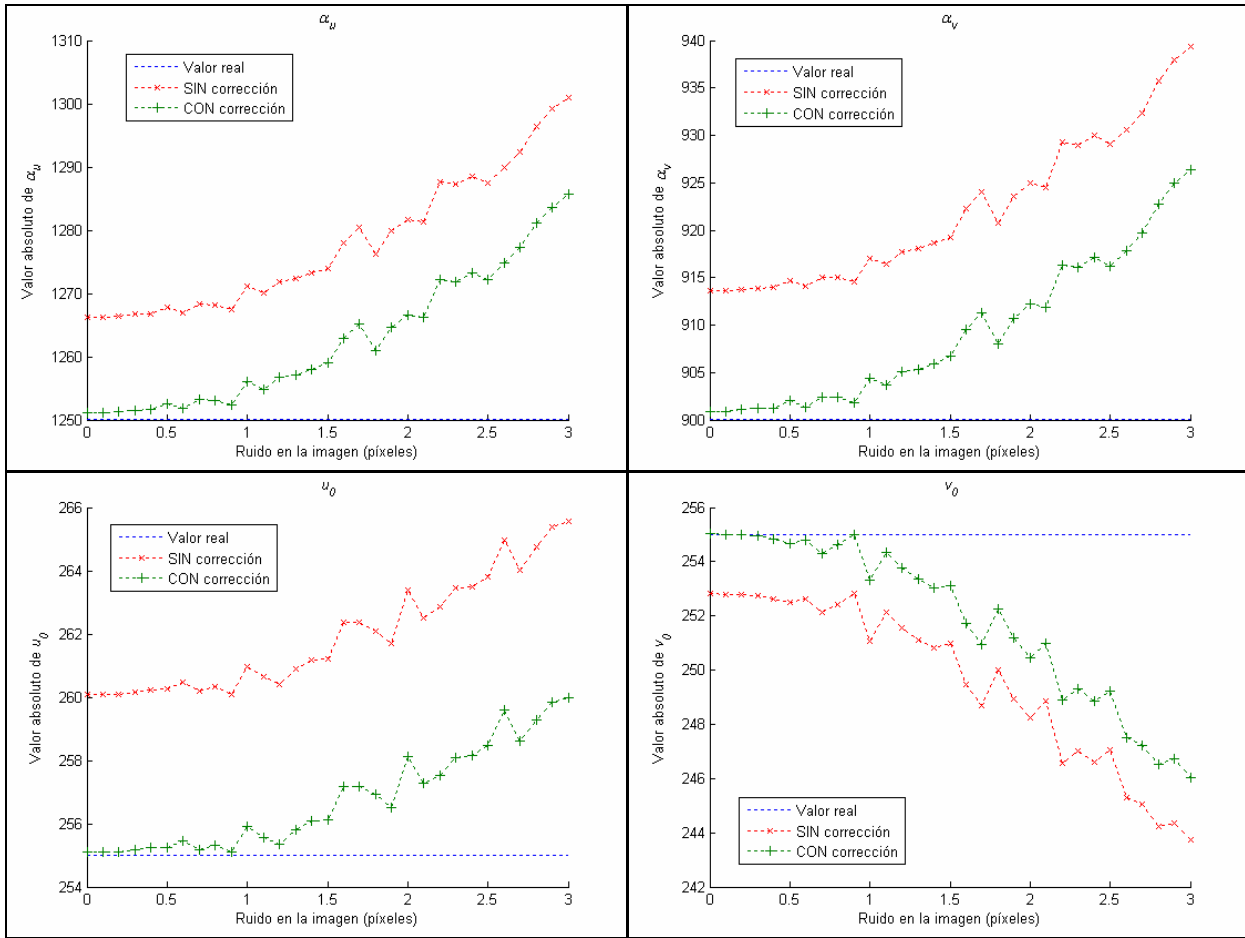
En primer lugar se procede a testear el efecto sobre el resultado de la calibración de la distorsión en los puntos de la imagen. En este caso no se trata de caracterizar el error de estimación de los parámetros que modelan la distorsión de la cámara. Este procedimiento ya se realiza bajo diferentes condiciones en el trabajo presentado por Ahmed [3],[4]. En este caso se va a utilizar este método para estimar la distorsión que permita corregir la imagen y a partir de los datos corregidos estimar los parámetros de la cámara utilizando el método basado en plantilla bidimensional. Se realiza una comparación de los resultados de calibración obtenidos haciendo una corrección de la distorsión previa y tomando directamente las coordenadas medidas de la imagen. Dado que se va a experimentar con datos simulados, se testeará la eficacia del método teniendo en cuenta diferentes condiciones en el proceso de calibración.

En primer lugar se va a proceder a determinar como afecta la distorsión en la imagen al proceso de calibración. En este caso se toman las imágenes de la plantilla desde las posiciones indicadas anteriormente y se simula la distorsión producida en la imagen según el modelo (3.78). Con las coordenadas sometidas a distorsión de los puntos en la imagen se realiza en proceso de calibración lineal explicado en el apartado 3.3.2.



3.6 Efecto de la distorsión sobre la estimación de los parámetros de la cámara

Para detectar el efecto de la distorsión sobre el proceso de calibración lineal, se realizan simulaciones cambiando los valores de cada uno de los parámetros que modelan la distorsión. Primero se varía k_1 dentro de un rango de valores dejando los demás igual a cero, después k_2 y así sucesivamente con p_1 y p_2 . Finalmente se realiza la misma operación cambiando el centro de distorsión que se había supuesto el centro de la imagen. En este caso se cambia el centro de distorsión manteniendo los parámetros de la distorsión diferentes de cero. Para realizar una simulación lo más parecida a la realidad, se deberían tomar unos valores de k_1 , k_2 , p_1 y p_2 ,

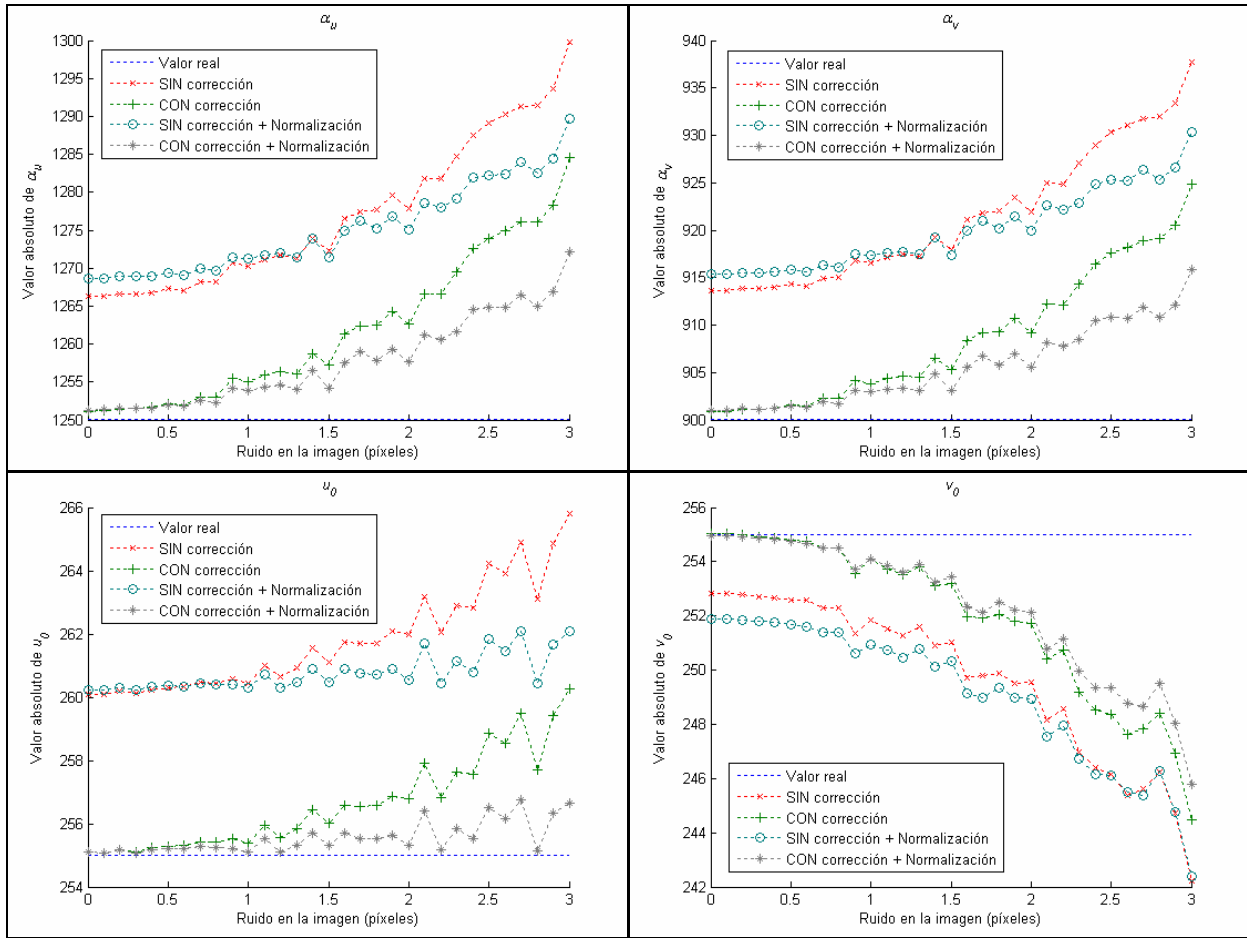


3.7 Efecto de la corrección de la distorsión de la imagen en la estimación lineal de los parámetros de la cámara

similares a los de una cámara real. Para ello se han tomado imágenes de una plantilla real con una webcam de Logitech modelo QuickCam Pro 4000, entendiéndose que este tipo de cámaras de gama baja añaden todo tipo de distorsiones a la imagen capturada.

El objetivo es estimar los parámetros de distorsión de esta cámara para establecer un rango de valores bajo los cuales es razonable encontrarlos. El método utilizado es el definido en el apartado 3.5.1. Los puntos utilizados para la estimación de los parámetros k_1 , k_2 , p_1 y p_2 y el centro de distorsión $\mathbf{c}=(c_u, c_v)$ son las esquinas de los cuadrados del tablero de ajedrez. Una vez calibrada la distorsión, se ha corregido la imagen. La imagen original y la corregida se muestran en la figura 3.5. En ellas se observa como las líneas que deben seguir todos los puntos de la plantilla se corrigen para que sean rectas. Los resultados de la calibración de la distorsión son $k_1=1.1834 \times 10^{-7}$, $k_2=0.3872 \times 10^{-7}$, $p_1=3.658 \times 10^{-7}$, $p_2=0.74138 \times 10^{-7}$, $c_u=312.5686$, $p_2=248.0743$.

Una vez obtenidos los datos de una cámara real, se va a simular el efecto que tiene la distorsión de la imagen en la estimación lineal de los parámetros de la cámara. Para ello se varían los valores de k_1 , k_2 , p_1 y p_2 así como los del centro de distorsión $\mathbf{c}=(c_u, c_v)$, para ver el efecto que tienen en la estimación lineal de los parámetros intrínsecos de la cámara α_u , α_v , u_0 , y

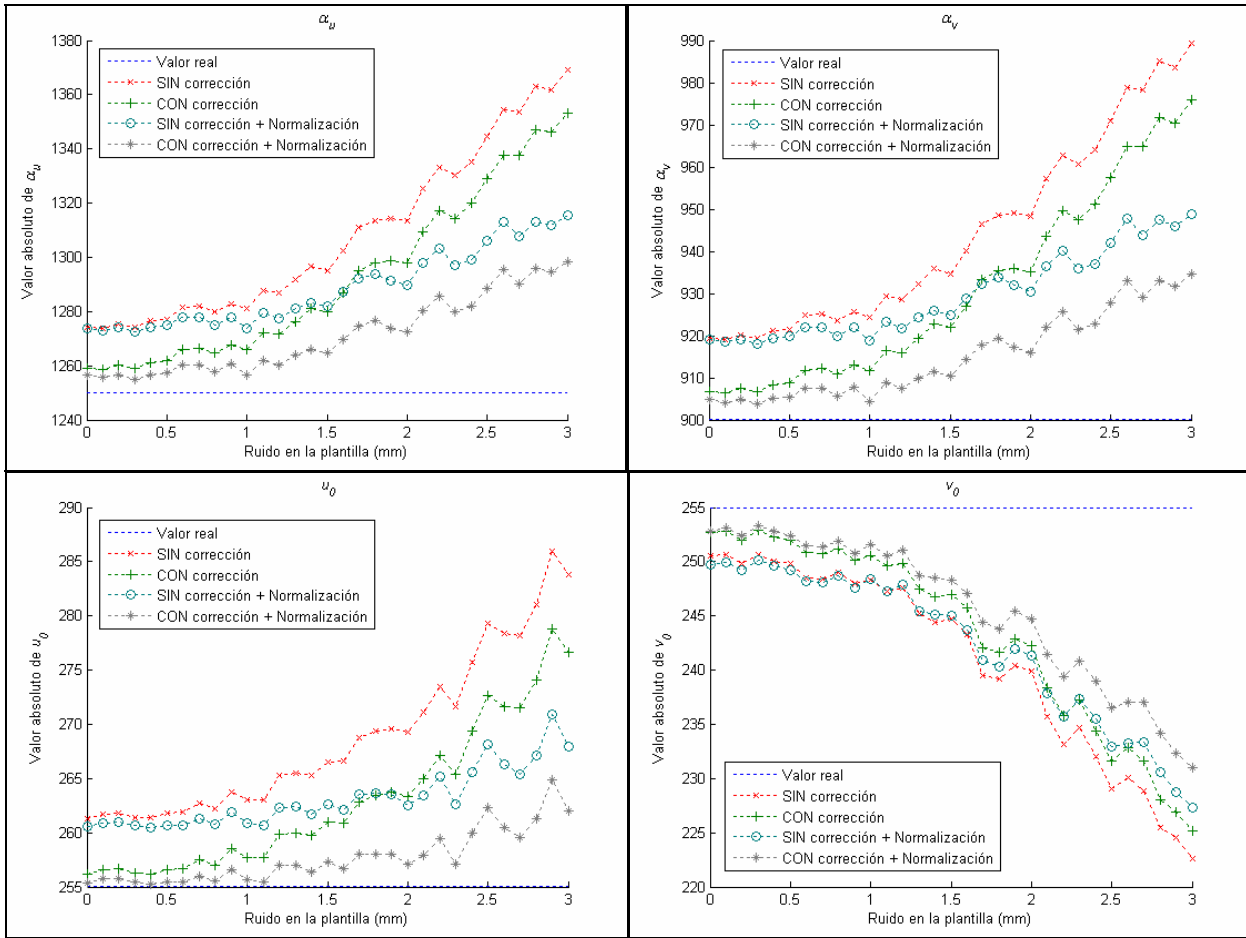


3.8 Efecto de la normalización de los datos sobre la estimación lineal de los parámetros de la cámara

v_0 . En este caso se realiza la calibración de la cámara presentada en el apartado 3.7 en las condiciones allí indicadas. Los resultados obtenidos se muestran en la figura 3.6.

Se han emparejado los resultados de la distorsión radial y tangencial, ya que con el parámetro k_1 se obtiene el mismo resultado que con el parámetro k_2 . Lo mismo ocurre con p_1 y p_2 . En el caso del centro de distorsión se presentan los resultados del error que se comete frente a la distancia del centro del CCD (u_0, v_0) con el centro de distorsión (c_u, c_v). Se observa que la distorsión radial es la que más afecta a la estimación lineal de α_u, α_v, u_0 , y v_0 . También es importante considerar el centro de distorsión de la imagen como fuente de errores.

En segundo lugar, dado que se tiene el comportamiento de la estimación lineal de los parámetros frente a los diferentes elementos que determinan la distorsión en la imagen, se procede a simular el efecto de la corrección de la imagen como paso previo antes de proceder a la estimación lineal de los parámetros. Partiendo de los valores de k_1, k_2, p_1 y p_2 obtenidos de la cámara real, se simula el efecto de la estimación lineal de α_u, α_v, u_0 , y v_0 , utilizando la imagen corregida y sin corregir. La corrección de la imagen se va a realizar teniendo en cuenta que los niveles de ruido pueden variar en las coordenadas de los puntos extraídos de la imagen. En



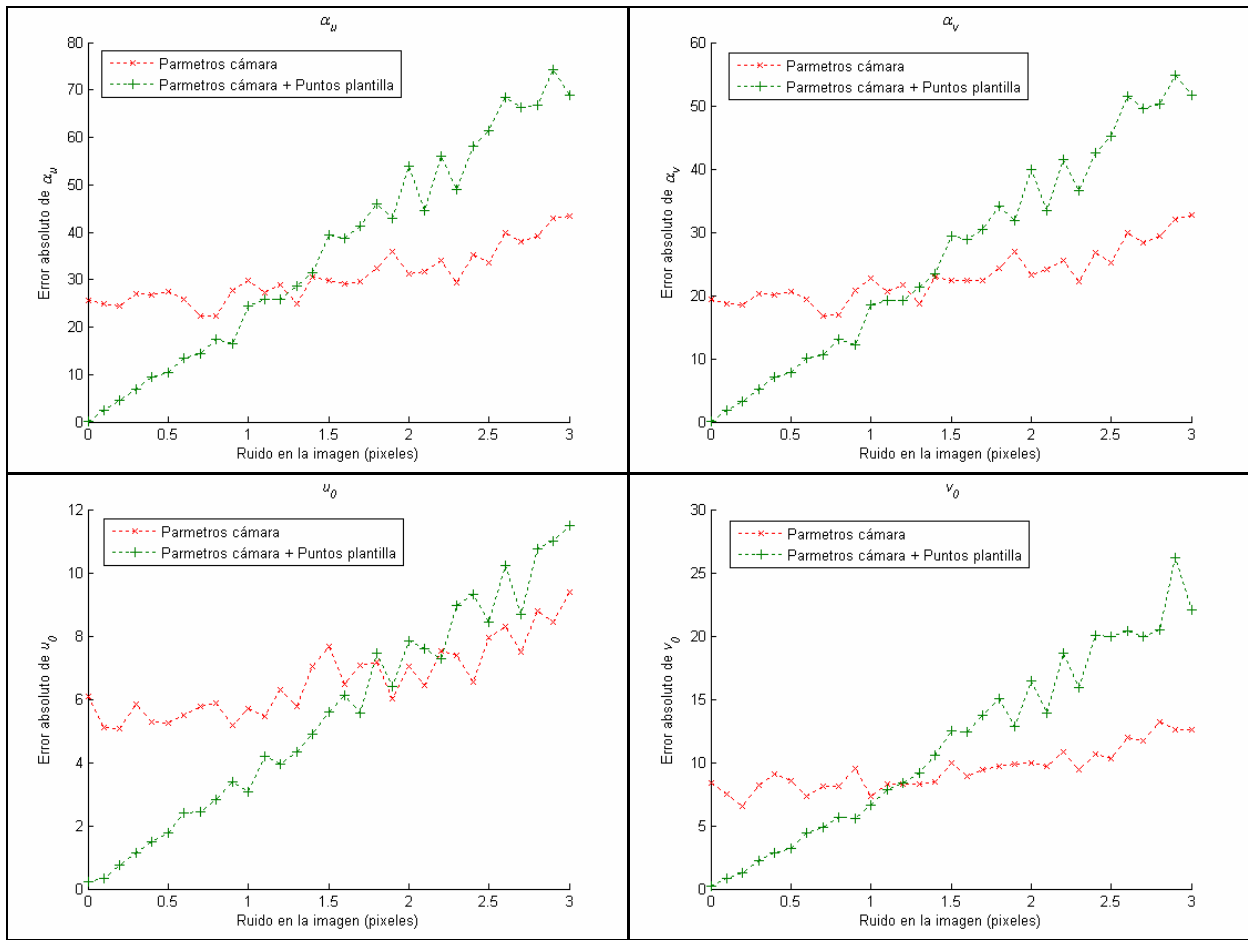
3.9 Efecto del ruido en la plantilla sobre la calibración lineal de la cámara. Los puntos en la imagen tienen un nivel de ruido de $\sigma=1.5$ píxeles

todos los casos se supone que las coordenadas de los puntos de la plantilla no están contaminadas con ruido.

La figura 3.7 muestra los resultados de la estimación lineal de los parámetros α_u , α_v , u_0 , y v_0 utilizando las imágenes corregidas y sin corregir. Obviamente, los resultados obtenidos con la imagen corregida son mejores. En el caso de realizar simplemente una estimación lineal de los parámetros, sería necesario este paso previo, para alcanzar mejores resultados. En el apartado de estimación no lineal de los parámetros α_u , α_v , u_0 , y v_0 se verá si es necesaria o no esta corrección previa de la imagen, ya que el punto inicial de búsqueda es importante para que el algoritmo de búsqueda converja al valor exacto.

Normalización de los datos

Según lo explicado en el apartado 3.5.2 con la normalización de los datos se consigue mejorar la condición de la matriz \mathbf{A} . Esto significa que el cálculo de las homografías es más estable frente al nivel de ruido en las medidas. En consecuencia se mejora la estimación de los parámetros de la cámara cuando la contaminación de los datos es importante. En este caso se ha

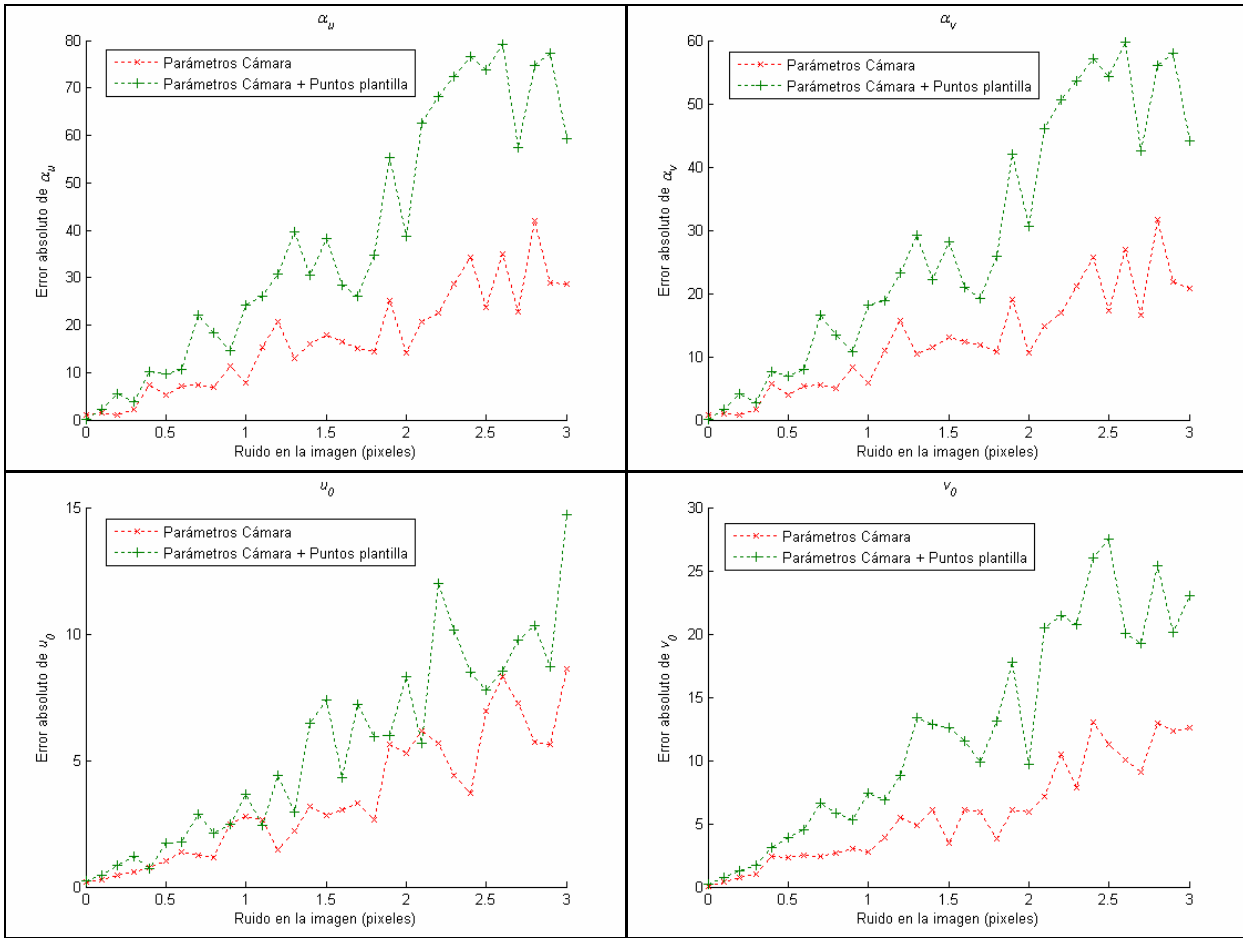


3.10 Efecto del ruido en la imagen sobre la calibración no lineal de la cámara. Los puntos en las plantillas tienen un nivel de ruido de $\sigma=1$ mm

simulado el efecto del ruido en las medidas, normalizando los datos y comparando los resultados con los obtenidos en el caso de la corrección de la distorsión de la imagen. Las condiciones de simulación son las mismas que las anteriores. Los resultados mostrados para cada nivel de ruido son la media de 500 calibraciones. Estos se muestran en la figura 3.8. Se observa que efectivamente, si el nivel de ruido aumenta, los resultados calculados son más cercanos a los valores reales si se realiza una normalización de los datos. La mejoría es siempre relativa a si se ha realizado o no, la corrección de la distorsión de la imagen. Evidentemente, los mejores resultados se obtienen si se corrige la distorsión de la imagen y se normalizan los datos. Si el nivel de ruido no es muy alto, una solución aceptable para el cálculo lineal de los parámetros de la cámara sería solamente corrigiendo la distorsión de la imagen. En estos casos incluso se obtienen resultados peores si se realiza una normalización de los datos.

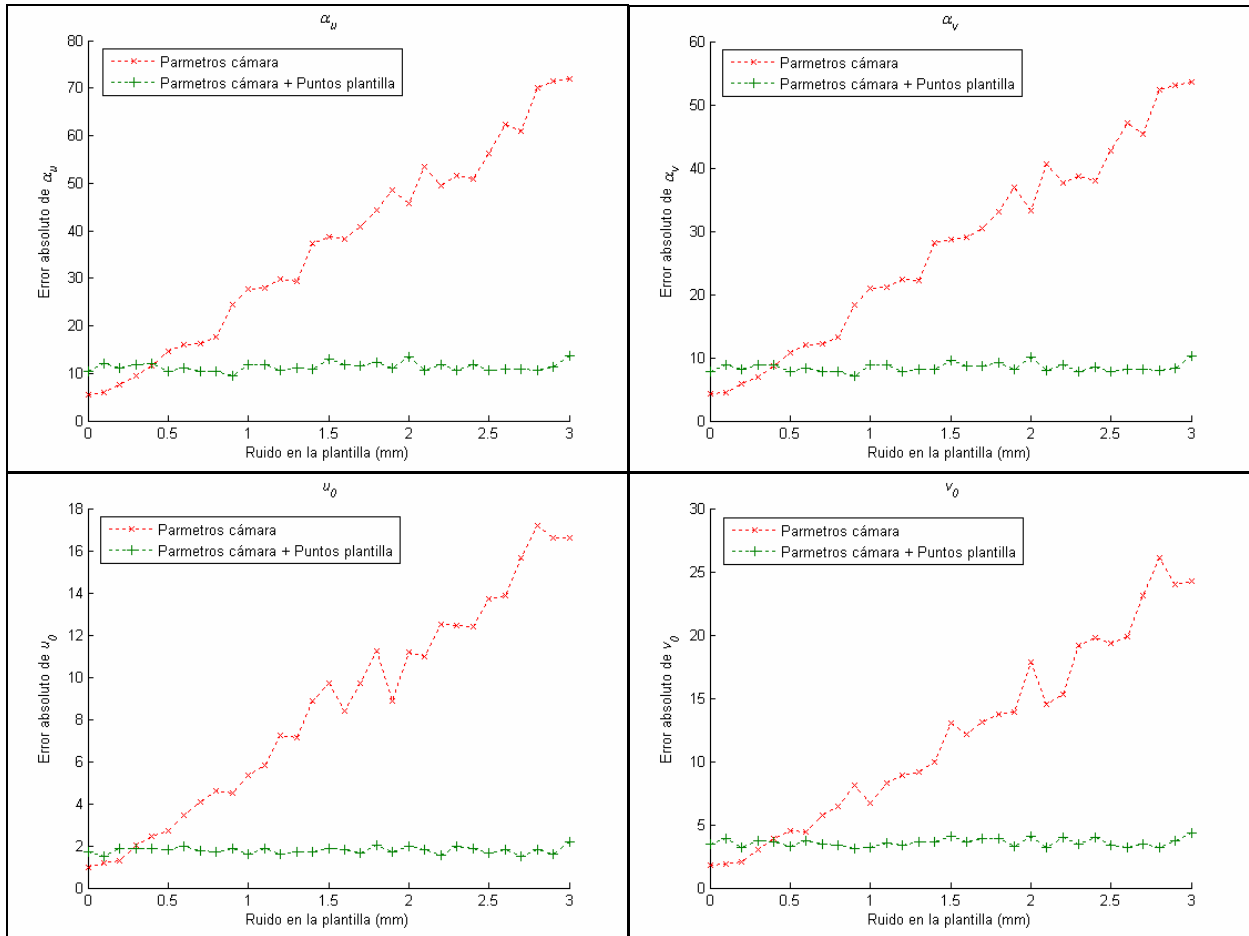
Calibración lineal de la cámara

En un proceso de calibración real, los datos de entrada están contaminados con ruido. El efecto del ruido en la imagen sobre el resultado de la calibración se ha visto en los apartados



3.11 Efecto del ruido en la imagen sobre la calibración no lineal de la cámara. Los puntos en las plantilla no están contaminados con ruido

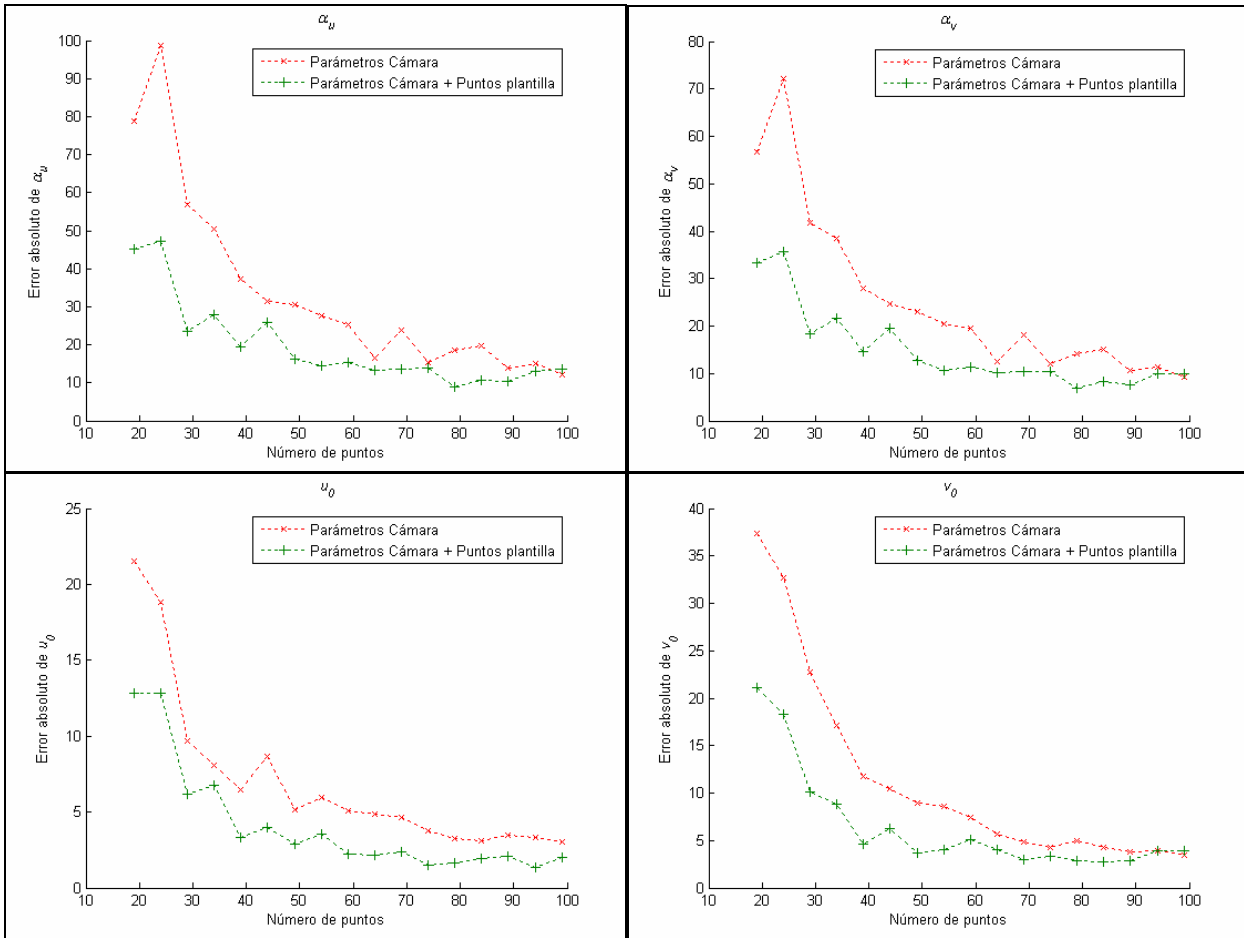
anteriores, e incluso se ha visto la forma de disminuir sus efectos. En el apartado de calibración lineal de la cámara se va a simular el efecto del ruido en las coordenadas de los puntos de interés en la plantilla sobre el resultado de calibración. Las condiciones de calibración y la cámara son las mismas que en apartados anteriores. También los resultados mostrados son la media de 500 calibraciones. En este caso, el nivel de ruido para las coordenadas de la imagen es de 1.5 píxeles. Se han estimado los parámetros de la cámara realizando la corrección de la distorsión de la cámara y la normalización de los datos. Los resultados obtenidos se muestran en la figura 3.9. Se observa como efectivamente la mejor estimación de los parámetros se realiza con la corrección de la distorsión y normalización de los datos. En el caso de tener que elegir entre corrección de la distorsión o normalización de los datos, se aprecia que cuando aumenta el nivel de ruido en la plantilla, la estimación de los parámetros es más fiable si se realiza simplemente el normalizando de los datos.



3.12 Efecto del ruido en la plantilla sobre la calibración no lineal de la cámara. Los puntos en la imagen tienen un nivel de ruido de $\sigma=0.5$ píxeles

Estimación no lineal de los parámetros de la cámara

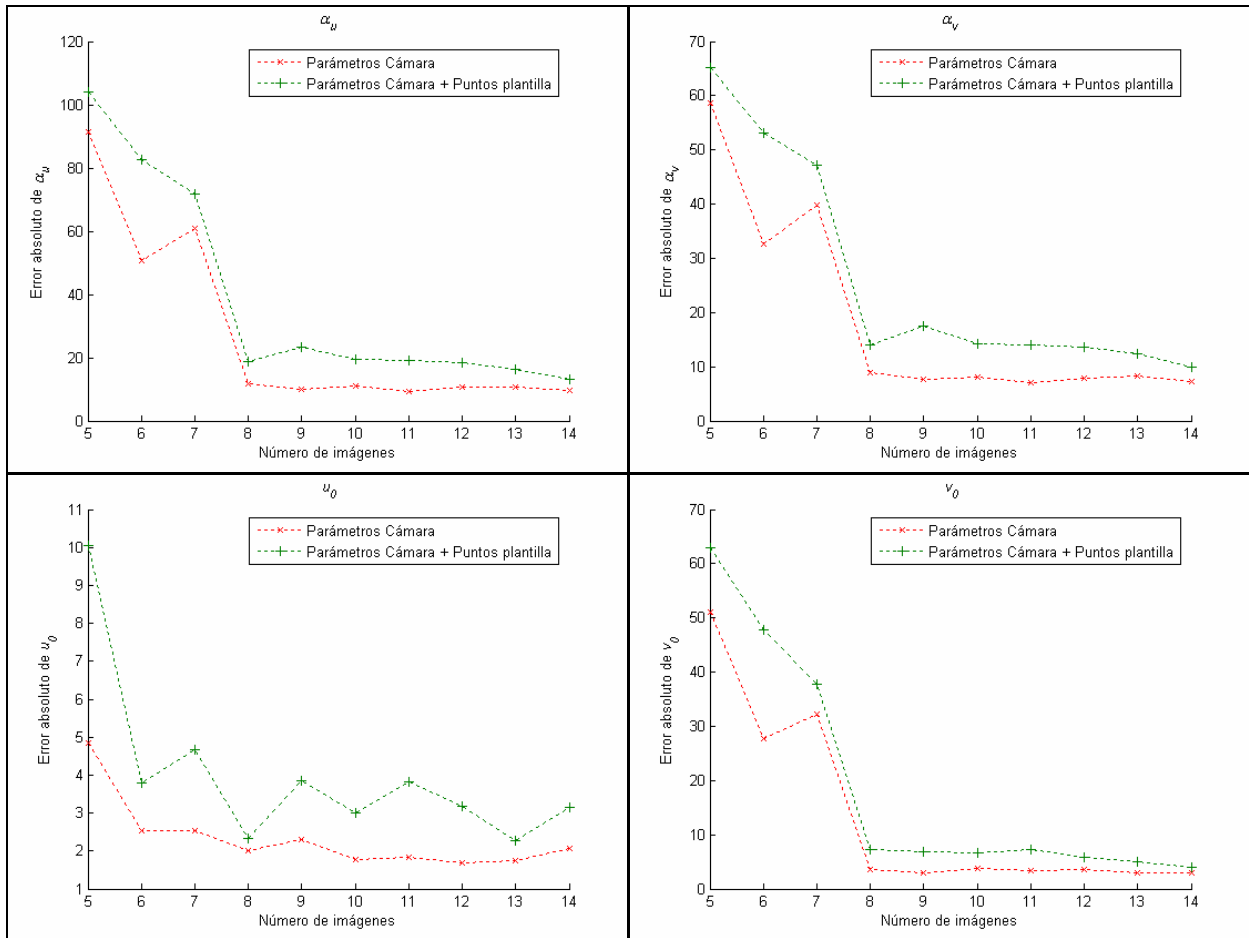
Como último paso del proceso de calibración de la cámara se considera la búsqueda no lineal de los parámetros. Se parte de un punto de búsqueda inicial y se realiza una búsqueda de los parámetros de forma que se minimiza un índice planteado. En este caso, se van a minimizar dos de los índices planteados en el apartado 3.4. Los puntos iniciales de búsqueda son aquellos calculados en los apartados anteriores de forma lineal. Dado que se parte del mismo conjunto de puntos aunque los valores iniciales de los parámetros sean diferentes, el algoritmo converge a los mismos valores de los parámetros. Es por este motivo que no importa los valores de los parámetros tomados de partida para las situaciones simuladas en este apartado. En todos los casos se ha llegado a los mismos valores con la búsqueda iterativa, si bien es cierto que con unos valores de los parámetros más cercanos a los reales la búsqueda no lineal acaba antes. Por lo tanto, como punto inicial de búsqueda se ha tomado el calculado con la distorsión corregida en la imagen y las coordenadas de los puntos normalizadas. Esta solución al estar más cercana a la definitiva, el algoritmo de búsqueda converge más rápidamente y la búsqueda no lineal es más corta.



3.13 Errores cambiando el número de puntos. El ruido en las coordenadas de la imagen es $\sigma=0.5$ y en las coordenadas de la plantilla es $\sigma=0.5$.

Desde el punto de vista del índice de coste a minimizar, se han simulado los resultados para la búsqueda de los elementos de la matriz de proyección y el cálculo directo de los parámetros de la cámara. Notables mejoras se obtienen el caso de realizar una búsqueda directa de los parámetros de la cámara. En este caso la función de coste que se plantea presenta unos mayores cambios de pendiente. Esto permite que si el punto inicial de búsqueda está cercano el definitivo, el algoritmo de búsqueda converja rápidamente al valor real.

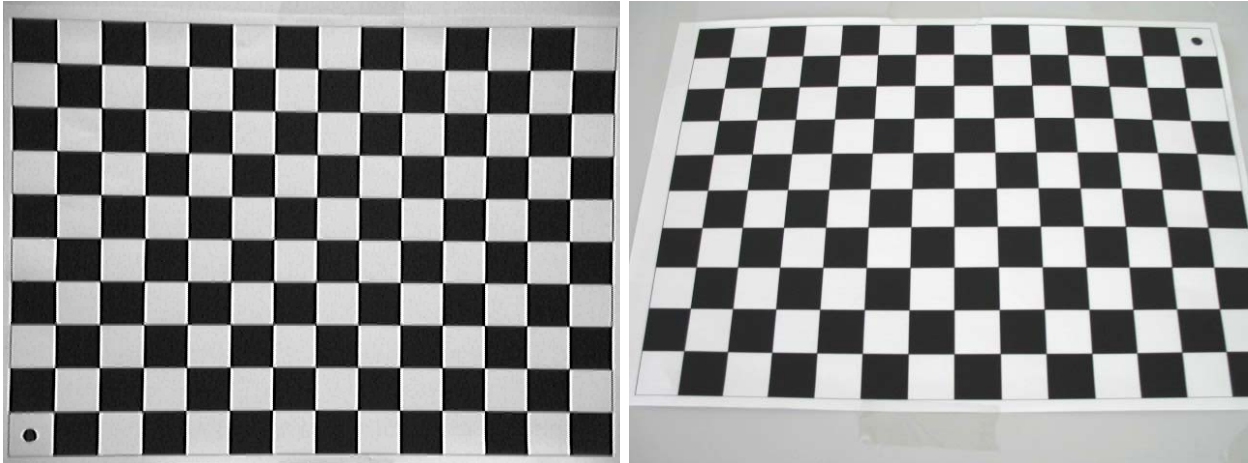
Se tiene el riesgo que sea imposible de encontrar el óptimo si el punto inicial se encuentra muy alejado del verdadero ya que la búsqueda puede evolucionar por otro camino. En el caso de la matriz de proyección, el índice de coste es “más lineal” y las mejoras con la búsqueda no lineal apenas se aprecian. Por el contrario esta búsqueda no lineal tiene un comportamiento más estable en el caso de tener un punto de partida muy alejado. Los resultados obtenidos se muestran en las figuras 3.10, 3.11 y 3.12. La figura 3.10 muestra los resultados de la búsqueda no lineal de los parámetros de la cámara variando el nivel de ruido en las coordenadas de los puntos en la imagen. Las coordenadas de la plantilla se contaminan con ruido Gaussiano de $\sigma=0.5$. Se compara el caso de la búsqueda no lineal de sólo los parámetros



3.14 Errores cambiando el número de imágenes. El ruido en las coordenadas de la imagen es $\sigma=1$ y en las coordenadas de la plantilla es nulo.

de la cámara, y de los parámetros de la cámara junto con las coordenadas de los puntos en la plantilla. Se puede observar que el algoritmo es capaz de corregir las coordenadas de los puntos en la plantilla y obtener resultados casi exactos en los que el error de la calibración sólo depende del nivel de ruido en los puntos de la imagen. En el caso de aumentar el ruido en la imagen, el algoritmo que calcula posiciones de los puntos en la plantilla llega a conclusiones menos parecidas a las reales ya que se entiende que no existe referencia para corregir las coordenadas de los puntos en la plantilla.

La figura 3.11 muestra la misma simulación pero en este caso, las coordenadas de los puntos en la plantilla no están contaminadas con ruido. El algoritmo que no busca los puntos de la plantilla consigue mejores resultados ya que estas medidas son exactas. La figura 3.12 presenta los efectos del error de medida cometidos en la plantilla. En este caso se demuestra que el error cometido en la calibración sólo depende de las imprecisiones realizadas en el proceso de medida en los puntos en la imagen. Realizando una búsqueda no lineal de los parámetros de la cámara junto con las coordenadas de los puntos en la plantilla, los errores cometidos en la calibración sólo dependerán del nivel de ruido en las medidas de la imagen. Esto puede indicar



3.15 Imágenes reales tomadas para la calibración. Izqd. Logitech Quick Cam Pro. Drcha. Canon Powershot A690

que la búsqueda de parámetros junto con el conjunto de puntos en la plantilla sea la más interesante a la hora de calibrar una cámara. Sin embargo se observa que en el caso de tener un nivel de ruido pequeño en la imagen, se obtienen mejores resultados buscando solamente los parámetros de la cámara. Además, basándose en las simulaciones anteriores presentadas en la figura 3.10, el algoritmo de búsqueda será capaz de corregir las coordenadas de los puntos de la plantilla si los errores en las medidas de los puntos en la imagen no son excesivos. En caso contrario los resultados son más parecidos a los reales si sólo se calculan los parámetros de la cámara.

Por lo tanto existe una ventana en la cual es mejor un método u otro. Esta ventana la dará en todos los casos el nivel de ruido de las medidas de la imagen.

Efectos del número de puntos y número de imágenes que participan en la calibración

En este caso se simulan los efectos del número de datos que participan en el proceso de calibración de la cámara. Se varían tanto el número de puntos en la plantilla así como las imágenes que se toman de las mismas. En el primer caso, se incrementa el número de puntos contaminando las coordenadas de los puntos en la imagen y en la plantilla con un ruido de $\sigma=1$. El número de imágenes que se toman son ocho al igual que en las experiencias simuladas anteriormente. La figura 3.13 muestra los resultados. Efectivamente contra más datos participan en la estimación, menores son los errores que se comenten en el cálculo de los parámetros. Como barrera significativa respecto del número de datos necesarios a ubicar en la plantilla, se podría decir que en torno a 50-60 puntos podría ser suficiente. En segundo lugar se cambian el número de imágenes que se toman de la plantilla. En este caso los puntos que se utilizan son los 99 de la plantilla y el número de imágenes varía desde 3 que son las mínimas necesarias hasta 14. En la figura 3.14, sólo se muestra a partir de 5 imágenes para poder apreciar mejor los resultados. Con

Logitech Quick Cam Pro

	Lineal				No Lineal	
	SIN corrección	CON corrección	SIN corrección + Normalización	CON corrección + Normalización	Parámetros cámara	Parámetros cámara + Puntos plantilla
α_u	931.2	932.2	933.5	929.3	932.1	930.4
α_v	947.2	942.2	945.4	941.5	942.5	940.1
u_0	312.5	318.5	326.8	322.4	323.5	317.5
v_0	238.7	237.9	242.5	242.9	242.8	241.9
φ	1.12	1.45	1.12	1.45	1.11	0.98
k_1	0	2.26	0	2.12	6.2	3.9
k_2	0	0.25	0	0.38	0.61	0.31
e	0.454	0.345	0.546	0.321	0.245	0.356
t	0.001	0.002	0.001	0.002	1.453	15.72

Canon PowerShot A640

	Lineal				No Lineal	
	SIN corrección	CON corrección	SIN corrección + Normalización	CON corrección + Normalización	Parámetros cámara	Parámetros cámara + Puntos plantilla
α_u	845.6	838.5	815.8	833.5	831.9	829.5
α_v	862.3	855.2	875.2	845.2	837.8	839.9
u_0	302.5	312.5	308.5	315.2	317.2	318.4
v_0	255.5	246.3	257.8	246.2	239.8	238.5
φ	0.54	0.98	0.79	0.96	0.15	0.65
k_1	0	0.551	0	0.554	0.646	0.889
k_2	0	0.05	0	0.05	0.06	0.04
e	0.585	0.58	0.556	0.473	0.456	0.398
t	0.001	0.002	0.001	0.002	1.657	15.65

3.1 Comparación de resultados de calibración de dos cámaras reales con diferentes métodos

un número menor de imágenes, los resultados son extrapolables. Las posiciones desde donde se toman las imágenes son $\mathbf{t}_1=(150, 200, 580)$, $\mathbf{t}_2=(-50, 250, 880)$, $\mathbf{t}_3=(100, -20, 820)$, $\mathbf{t}_4=(-40, -150, 780)$, $\mathbf{t}_5=(-150, -150, 530)$, $\mathbf{t}_6=(-100, 125, 400)$, $\mathbf{t}_7=(140, -150, 500)$, $\mathbf{t}_8=(240, 50, 600)$, $\mathbf{t}_9=(50, -250, 880)$, $\mathbf{t}_{10}=(-240, -50, 600)$, $\mathbf{t}_{11}=(-140, 150, 600)$, $\mathbf{t}_{12}=(125, 20, 800)$, $\mathbf{t}_{13}=(25, -100, 700)$, $\mathbf{t}_{14}=(-25, 100, 600)$. En todos los casos se orienta la cámara para que se el eje óptico coincida con el centro de la imagen. Igual que antes, contra más imágenes se toman menos error se comete. Como barrera significativa del número de imágenes necesarias, se observa que con más de 8 imágenes la mejora de los resultados ya no resulta tan significativa.

Resultados obtenidos para cámaras reales

Sendas calibraciones se han realizado de dos cámaras reales. Una de ellas de gama baja, la logitech Quick Cam Pro, y la otra la Canon Powershot A640. En ambos casos 8 imágenes se han tomado de la plantilla. Las resoluciones han sido en ambos casos de 640x480 píxeles. La plantilla está compuesta de un tablero de ajedrez de 15x11 esquinas de dimensiones 280x200 mm. Se asume que está colocado de forma que el plano de la plantilla corresponde con el plano $z=0$ de la escena. En la figura 3.15 se muestra una de ellas correspondiente a cada cámara. Las posiciones desde donde se han tomado las imágenes se ha tratado que sean parecidas a la de la cámara simulada. La suma de todos los errores geométricos entre los puntos $\mathbf{q}_i^\#$ y \mathbf{q}_i^* partido por el número de puntos se muestra en los resultados. Las calibraciones de las cámaras se han realizado en los mismos casos que la cámara simulada, es decir haciendo un preprocesado de los datos y sin hacerlo, y también realizando una búsqueda no lineal con los dos índices de coste

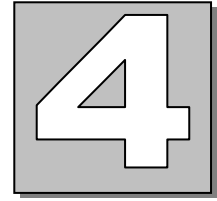
utilizados en simulación. Los resultados se muestran en la tabla 3.1 para la cámara Logitech y Canon. e representa el error geométrico cometido en cada caso. t representa el tiempo en segundos que cuesta hacer la calibración. Se computa sólo el tiempo de cálculo del algoritmo a partir de los puntos detectados de la imagen. El calibrado se ha realizado con un procesador AMD Athlon 64 con una velocidad de reloj de 3200 Mhz. Se observa que en el caso de realizar una búsqueda no lineal de los parámetros, el error geométrico resultado es menor que en el caso de realizar una estimación lineal de los mismos. Si se ajusta la búsqueda no lineal junto con el conjunto de puntos en el mundo se consiguen mejores resultados. Por el contrario, el tiempo de cómputo crece de forma inversamente proporcional a la calidad de los resultados. La búsqueda no lineal junto con el conjunto de puntos en el mundo multiplica por 10^5 el tiempo de cómputo. Para aplicaciones en las que el tiempo de cómputo sea crítico es necesario pensar en una calibración lineal junto con un preprocesado de los datos.

3.8. Conclusiones

El proceso de calibrado de una cámara es un paso necesario si se desea utilizar la información de la imagen para realizar medidas de la escena o para efectuar reconstrucciones 3D de la misma. Este proceso consiste en estimar los parámetros que caracterizan la cámara así como su posición en el escenario. Para llevarlo a cabo se pueden utilizar tanto plantillas tridimensionales, bidimensionales o unidimensionales, así como también el procedimiento llamado auto calibración que no necesita plantillas de calibración. Básicamente, dado que se trata de una estimación de parámetros o modelado de un sistema, los métodos de calibración constan de dos etapas. En la primera etapa se calcula un valor aproximado de una parte de los parámetros utilizando métodos de optimización lineal. Esta aproximación se toma como punto de partida para una búsqueda no lineal de todo el conjunto de parámetros que modelan la cámara y la sitúan en el escenario. En esta segunda etapa se utilizan métodos iterativos de búsqueda. La efectividad del proceso de calibración depende tanto de la calidad en las medidas que intervienen, así como del modelo a estimar, sin olvidar el método utilizado para la estimación. A lo largo del capítulo se ha realizado un exhaustivo estado del arte sobre las diferentes técnicas de calibración. No sólo se han tratado técnicas de calibración en si mismas, sino que también se han tenido en cuenta posibles técnicas que mejoren la estimación de los parámetros. El objetivo ha sido definir un método de calibración que resulte efectivo y válido para la mayoría de los problemas de calibración que se puedan presentar en la actualidad. Para ello se han utilizado estudios previos además de experiencias realizadas por el autor. Como conclusión de estudios previos realizados se puede decir que el método presentado por Zhang en [103] resulta el más eficiente a la hora de calibrar una cámara. Partiendo de éste método, se ha definido un preprocesado de la información y una mejora en el índice de coste que perfecciona la

búsqueda no lineal. El preprocesado de los datos de la imagen consiste en corregir errores en las medidas debidos a la distorsión que introducen las lentes, y también formalizar una normalización de los datos para condicionar mejor el proceso de cálculo del algoritmo. En la parte de cálculo no lineal de los parámetros se ha planteado también la búsqueda de las coordenadas de los puntos de la plantilla. Se puede decir que el caso de no realizar una búsqueda no lineal en los parámetros, el preprocesado de los datos permite obtener un conjunto de parámetros muy cercanos a los reales. Esto permite calibraciones ‘on line’ para utilizarlas por ejemplo en la localización de un vehículo mediante una cámara. Si se pretende ajustar más el resultado y se utilizan técnicas lineales de optimización, se concluye que el ruido en las medidas de las coordenadas en la plantilla no afecta al proceso de calibrado. Solamente el ruido en las coordenadas de la imagen perturba al proceso de calibración. Dado que se realiza un preprocesado de los datos para minimizar este ruido en la imagen, se puede decir que se consiguen resultados casi exactos en el proceso de calibración que sólo dependen del nivel de ruido de la imagen.

A partir de este método de calibración basado en plantilla bidimensional al que se le han introducido significantes mejoras, se pretende obtener los intervalos en los cuales se están contenidos los parámetros estimados. A pesar de obtener resultados casi exactos, estos siempre serán imperfectos. Para calcular esta incertidumbre se van a utilizar técnicas de estimación estadísticas que tienen en cuenta los errores de los datos de partida en el proceso de cálculo. Este hecho permite descartar medidas con incertidumbre elevada. Con las técnicas utilizadas hasta ahora, estos datos provocan perturbaciones importantes en el procedimiento de cálculo.



ESTIMACIÓN BASADA EN INCERTIDUMBRES

En el proceso de calibrado de los parámetros de la cámara se introducen errores los cuales se reflejan en la calidad de los resultados. No es suficiente con tener los valores de los parámetros de la cámara, sino que también es necesario disponer de una medida de incertidumbre de los mismos. Además para contrarrestar los efectos de las incertidumbres en las medidas, surgen cuestiones sobre el número de puntos necesarios en la plantilla de calibración o el número de imágenes a tomar para que el proceso tenga resultados estables. También, dado que las medidas son inciertas, es interesante utilizar cada una de ellas de acuerdo a la incertidumbre que presentan o corregir las mismas para minimizar su error de acuerdo a algún criterio.

En este apartado se desarrolla una metodología que permite caracterizar el proceso de calibración de la cámara. En primer lugar se realiza una estimación de la incertidumbre de los datos de partida procedentes de la imagen. Después basándose en la incertidumbre de los datos de partida se realiza la estimación de los parámetros de la cámara. Esto permite pesar la información a la hora de utilizarla, además de obtener unos parámetros junto con su incertidumbre, la cual viene expresada en una matriz de covarianzas. Esta metodología también permite definir valores para el número puntos necesarios o número de imágenes en función de las incertidumbres de las medidas, además de corregir los datos para que cumplan ciertas restricciones de acuerdo a la incertidumbre que los define. Basándose en los resultados obtenidos en el apartado anterior, esta metodología se utiliza para caracterizar el proceso de calibración utilizando una plantilla bidimensional. Al final del capítulo se pone en práctica el algoritmo calculando los parámetros de la cámara junto con sus incertidumbres y definiendo valores óptimos para realizar la calibración de la cámara con plenas garantías.

4.1. Errores de estimación

Los errores en la estimación de los parámetros de la cámara proceden de diferentes fuentes. Éstas pueden ser tanto errores en las medidas en la imagen, como del número de puntos utilizados, así como de la configuración espacial de los mismos. Por un lado, los errores en las medidas de la imagen vienen provocados tanto por las distorsiones producidas por defectos constructivos de la propia cámara, o también por el algoritmo encargado de la búsqueda de características de interés dentro de la imagen. Además, también dependen del ángulo desde el que se toman las imágenes, así como de imperfecciones constructivas de la plantilla. Esta detección de características de interés en la imagen acarrea tanto errores de cuantización y como de detección.

Por otro lado, otra fuente de error muy importante en el proceso de calibrado es la situación espacial de los puntos en el espacio tridimensional. Los algoritmos utilizados para la calibración son relativamente sensibles a los errores en las coordenadas de los puntos en función de las situaciones de los mismos. Recordar que hay situaciones prohibidas en las cuales el algoritmo no permite estimar la matriz de proyección. Esto provoca posiciones de los puntos que llevan a situaciones degeneradas matemáticamente, que debido al ruido en las coordenadas de los mismos se puedan resolver. Asimismo, si partiendo de una situación degenerada matemáticamente, si se mueve un punto para obtener una configuración con solución matemática, dicha configuración resultará muy sensible al ruido. Por lo tanto, en estos casos la solución obtenida será inservible.

Existen diversos métodos para medir la incertidumbre tanto en las medidas de las coordenadas como en los parámetros estimados a partir de las mismas. La incertidumbre basada en la determinación del peor de los casos obtiene un intervalo definido por el límite o valor máximo del error. Este método es muy útil para aplicaciones con errores de redondeo en las cuales se desea estimar el error a partir de una determinada longitud de palabra. Debido a que la longitud de palabra utilizada en la implementación de los algoritmos de calibración es más que suficiente, los errores cometidos por el propio cálculo son despreciables frente a los errores que se cometen por el uso de medidas de coordenadas de puntos de interés contaminadas con ruido. Con estos métodos, la estimación de la incertidumbre basada en el peor de los casos resulta en un valor muy grande, el cual aporta poca información. Se puede pensar por ejemplo en el límite superior de una variable aleatoria con distribución Gaussiana. Debido a que este límite casi nunca es alcanzado, la utilidad de esta estimación de la incertidumbre es muy limitada.

Para combatir el efecto de las medidas inciertas, se utilizan métodos de cálculo basados en observaciones redundantes. Estos consisten en la minimización de algún índice compuesto por un mayor número de medidas de las estrictamente necesarias. Con los métodos de cálculo

basados en observaciones redundantes, el nivel de error que se obtiene en el resultado es relativamente estable para un nivel de ruido fijo en los datos de entrada. Esta estabilidad se debe a la redundancia de las observaciones. Dado un conjunto de medidas inciertas es posible determinar su incertidumbre a partir de la aproximación de sus parámetros estadísticos o teniendo en cuenta un conocimiento a priori de las condiciones que cumplen dichas medidas. Si se tiene conocimiento de las relaciones geométricas que cumplen estas medidas en ausencia de ruido, es posible corregirlas de forma que se adapten a las restricciones que ellas mismas cumplen, y de esta forma calcular sus incertidumbres las cuales siempre serán menores que las que se obtendrían en el caso de aproximar sus parámetros estadísticos directamente. Esta técnica es la llamada corrección geométrica mediante la cual el conjunto de medidas son desplazadas de forma óptima para que cumplan las restricciones geométricas de forma exacta desde un punto de vista de máxima verosimilitud. Por otro lado, dado un conjunto de medidas inciertas, es posible estimar los parámetros que de forma óptima, representan las restricciones geométricas que cumplen el conjunto de medidas sin ruido. Dado que el conjunto de medidas son inciertas los parámetros estimados también lo son y su incertidumbre puede ser caracterizada.

Si se realiza una estimación de parámetros a partir de medidas inciertas, será necesario tener en cuenta en todo momento conceptos tales como el límite de máximo de optimalidad que se puede alcanzar en la estimación a partir de una incertidumbre en los datos de entrada o también, la incertidumbre de los parámetros calculados. Además, también habrá que tener en cuenta la eficiencia del algoritmo desde el punto de vista que sea estable en los cálculos frente a los mismos errores en los datos de entrada o errores de redondeo, así como en que medida es posible adaptar el conjunto de datos al modelo determinado por los parámetros. El planteamiento estadístico del problema de estimación de parámetros de la cámara permite tener en cuenta todos estos aspectos del proceso de cálculo. A continuación se describen diferentes herramientas estadísticas que posteriormente se aplicarán al proceso de calibrado de la cámara.

4.2. Estimación estadística

4.2.1. Distribuciones Gaussianas

La distribución de probabilidad fundamental de una variable aleatoria N -dimensional es la distribución Gaussiana multidimensional o también llamada distribución normal. Esta distribución se caracteriza mediante su media y su matriz de covarianzas. Dada una variable aleatoria N -dimensional representada vector $\mathbf{x}^\#$, su distribución Gaussiana multidimensional está

representada por su media \mathbf{x} y su matriz de covarianzas Σ_x según la siguiente densidad de probabilidad:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_x|}} e^{-\left(\mathbf{x}^\# - \mathbf{x}\right)^T \cdot \Sigma_x^{-1} \cdot (\mathbf{x}^\# - \mathbf{x}) / 2} \quad (4.1)$$

Esta densidad de probabilidad define una distribución a lo largo de todo el espacio N -dimensional R^n . La densidad de probabilidad en la superficie en R^n es constante en las direcciones definidas por la expresión $\mathbf{c} = (\mathbf{x}^\# - \mathbf{x})^T \cdot \Sigma_x \cdot (\mathbf{x}^\# - \mathbf{x})$. Si $\mathbf{x} = \mathbf{0}$ y $\Sigma_x = \mathbf{I}$ la distribución se llama normal. En el caso de que los elementos del vector aleatorio no estén correlados la matriz de covarianzas Σ_x es una matriz diagonal en la que cada elemento de la diagonal es una varianza σ_i^2 asociada al elemento i del vector aleatorio $\mathbf{x}^\#$. El valor σ_i se llama desviación típica. Para cada elemento del vector x_i , el intervalo definido por $[x_i - \sigma_i, x_i + \sigma_i]$ se llama intervalo estándar y la probabilidad de que el valor esté dentro del intervalo estándar es del 68.27%. Si la matriz de covarianzas Σ_x no es diagonal, los elementos de fuera de la diagonal indican las relaciones de varianza que existen entre cada uno de los elementos del vector. Dada una distribución Gaussiana en la cual el rango r de de la matriz de covarianzas es menor que n , las varianzas de los elementos $i > r$ serán nulas. Esto indica que la distribución está limitada el subespacio r -dimensional. La base de este espacio R -dimensional vendrá dada por los vectores propios \mathbf{u}_i asociados a la matriz de covarianzas, cada uno de los cuales tendrá asociado un valor propio que corresponderá con la varianza en dicha dirección.

$$\Sigma_x = \sum_{i=1}^n \sigma_i^2 \mathbf{u}_i \mathbf{u}_i^T \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n (=0) \quad (4.2)$$

Si la matriz de covarianzas tiene una descomposición espectral dada por la expresión anterior, \mathbf{u}_i indica la dirección en la cual es más probable que la desviación ocurra y σ_i es la desviación estándar en esa dirección. El intervalo $[\mathbf{x} - \sigma_i \cdot \mathbf{u}_i, \mathbf{x} + \sigma_i \cdot \mathbf{u}_i]$ define la orientación en la cual la desviación es más probable que ocurra.

Si $\mathbf{x}^\#$ es un vector n -dimensional, aleatorio de media \mathbf{x} y matriz de covarianzas Σ_x , un vector $\mathbf{y}^\#$ m -dimensional definido por una matriz arbitraria \mathbf{A} de dimensiones $n \times m$ según $\mathbf{y}^\# = \mathbf{A} \cdot \mathbf{x}^\#$, $\mathbf{y}^\#$ es una variable aleatoria de distribución Gaussiana de media $\mathbf{A} \cdot \mathbf{x}$ y matriz de covarianzas $\mathbf{A} \cdot \Sigma_x \cdot \mathbf{A}^T$.

4.2.2. Distribuciones χ^2

Si x_1, x_2, \dots, x_r son muestras independientes de una distribución Gaussiana unidimensional, de media x y varianza σ^2 , se define la distribución R de la siguiente forma:

$$R = \frac{(x_1 - x)^2}{\sigma^2} + \frac{(x_2 - x)^2}{\sigma^2} + \dots + \frac{(x_r - x)^2}{\sigma^2} \quad (4.3)$$

Esta función llamada distribución χ^2 con r grados de libertad tiene como valor medio r y la desviación típica de R es $2 \cdot r$. Las propiedades más importantes de una distribución χ^2 son:

- Si $R_1, R_2, R_3, \dots, R_n$ son variables independientes χ^2 , con $r_1, r_2, r_3, \dots, r_n$ grados de libertad respectivamente, la suma $R = R_1 + R_2 + R_3 + \dots + R_n$ es una variable χ^2 con $r = r_1 + r_2 + r_3 + \dots + r_n$ grados de libertad.
- Si \mathbf{x} es una variable aleatoria con distribución Gaussiana de media $\mathbf{0}$ y matriz de covarianzas Σ_x de rango r , la forma cuadrática $R = \mathbf{x}^T \cdot \Sigma_x^{-1} \cdot \mathbf{x}$ es una variable χ^2 con r grados de libertad.
- Si $\mathbf{x}_i, i=1 \dots n$, son variables aleatorias Gaussianas independientes, en la que cada una de ellas tiene media representada por el vector $\mathbf{0}$ y matriz de covarianzas Σ_{x_i} de rango r_i , la suma $R = \sum_{i=1 \dots n} (\mathbf{x}_i^T \cdot \Sigma_{x_i}^{-1} \cdot \mathbf{x}_i)$ es una variable χ^2 con $\sum_{i=1 \dots n} r_i$ grados de libertad.
- Se tiene el vector N -dimensional \mathbf{x} y el vector M -dimensional \mathbf{y} siendo ambos variables aleatorias de media $\mathbf{0}$ y Σ_x, Σ_y sus respectivas matrices de covarianzas. n y r ($n \geq r$) son los rangos de Σ_x, Σ_y respectivamente. Si existe una matriz \mathbf{A} de dimensiones $m \times n$ que relaciona ambos vectores de la forma $\mathbf{y} = \mathbf{A} \cdot \mathbf{x}$, según el teorema de Cochran, la diferencia $R = (\mathbf{x}^T \cdot \Sigma_x^{-1} \cdot \mathbf{x}) - (\mathbf{y}^T \cdot \Sigma_y^{-1} \cdot \mathbf{y})$ es una variable χ^2 con $n - r$ grados de libertad.

La distribución χ^2 es una herramienta estadística para testear hipótesis. En muchas situaciones se puede definir una variable aleatoria R de la forma (4.3) en la cual cada término x_i tiene media nula bajo ciertas condiciones. Estas condiciones se plantean como hipótesis, y pueden ser testeadas observando el valor de R para un conjunto de muestras. La hipótesis es rechazada si este valor es relativamente grande.

4.2.3. Estimación de máxima verosimilitud

Dada una variable aleatoria Gaussiana representada por el vector N -dimensional \mathbf{x} de media $\mathbf{0}$ y matriz de covarianzas Σ_x , si es de rango completo se define la norma $|\mathbf{x}|_\Sigma$ de \mathbf{x} como

$$|\mathbf{x}|_\Sigma = \sqrt{\mathbf{x}^T \cdot \Sigma_x^{-1} \cdot \mathbf{x}} \quad (4.4)$$

Esta norma $|\mathbf{x}|_\Sigma$ es llamada distancia de Mahalanobis del vector \mathbf{x} . Si \mathbf{x} es elegida de forma aleatoria, $|\mathbf{x}|_\Sigma^2$ es una variable χ^2 con n grados de libertad.

Se tiene un vector n -dimensional \mathbf{x} , una matriz \mathbf{A} de dimensiones $m \times n$ y también se tiene una variable aleatoria Gaussiana representada por el vector n -dimensional \mathbf{e} de media $\mathbf{0}$ y matriz de covarianzas Σ_e . El vector M -dimensional $\mathbf{y}^\# = \mathbf{A} \cdot \mathbf{x} + \mathbf{e}$ es una variable aleatoria de media $\mathbf{y} = \mathbf{A} \cdot \mathbf{x}$ y matriz de covarianzas Σ_e . La densidad de probabilidad de $\mathbf{y}^\#$ está definida por la siguiente expresión.

$$p(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_e|}} e^{-\left(\mathbf{y}^\# - \mathbf{A} \cdot \mathbf{x}\right)^T \cdot \Sigma_e^{-1} \cdot \left(\mathbf{y}^\# - \mathbf{A} \cdot \mathbf{x}\right) / 2} \quad (4.5)$$

Si se considera el problema de estimar el parámetro \mathbf{x} a partir de una muestra de $\mathbf{y}^\#$, la estimación de máxima verosimilitud busca el valor de \mathbf{x} que maximiza la densidad de probabilidad $p(\mathbf{y})$ descrita en la expresión anterior. El problema se reduce a minimizar la distancia de Mahalanobis $|\mathbf{y}^\# - \mathbf{A} \cdot \mathbf{x}|_{\Sigma_e}$ planteado en un índice de la siguiente forma:

$$J(\mathbf{x}) = \left(\mathbf{y}^\# - \mathbf{A} \cdot \mathbf{x}\right)^T \cdot \Sigma_e^{-1} \cdot \left(\mathbf{y}^\# - \mathbf{A} \cdot \mathbf{x}\right) \quad (4.6)$$

La solución al problema de estimación lo da la siguiente expresión

$$\mathbf{x}^* = \left(\mathbf{A}^T \cdot \Sigma_e^{-1} \cdot \mathbf{A}\right)^{-1} \mathbf{A}^T \cdot \Sigma_e^{-1} \cdot \mathbf{y}^\# \quad (4.7)$$

La media del valor estimado \mathbf{x}^* coincide con su valor real \mathbf{x} y la varianza está definida por la expresión:

$$V(\mathbf{x}^*) = \left(\mathbf{A}^T \cdot \Sigma_e^{-1} \cdot \mathbf{A}\right)^{-1} \quad (4.8)$$

El residuo de de la estimación $J(\mathbf{x}^*)$ está definido por la siguiente expresión

$$J(\mathbf{x}^*) = \left(\mathbf{y}^{\#T} \cdot \Sigma_e^{-1} \cdot \mathbf{y}^\#\right) - \left(\mathbf{x}^* - \mathbf{A}^T \cdot \Sigma_e^{-1} \cdot \mathbf{A} \cdot \mathbf{x}^*\right) \quad (4.9)$$

Este residuo representa una variable χ^2 con $n-m$ grados de libertad, donde m representa el rango de la matriz \mathbf{A} .

Si cada componente de \mathbf{e} se distribuye independientemente y de forma isotópica con la misma σ^2 , la matriz de covarianzas de \mathbf{e} es vale $V(\mathbf{e}) = \sigma^2 \cdot \mathbf{I}$. En este caso el problema es equivalente a la resolución de mínimos cuadrados.

En el caso de que existan restricciones que tengan que cumplirse, la optimización se plantea de la siguiente manera. Se tiene un vector N -dimensional $\mathbf{x}^\#$ que representa la variable aleatoria con distribución Gaussiana de media \mathbf{x} y matriz de covarianzas Σ_x de rango completo. Se supone que la media \mathbf{x} satisface la restricción lineal $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$. Se considera el problema de estimar el valor de \mathbf{x} a partir del valor medido $\mathbf{x}^\#$. La densidad de probabilidad de \mathbf{x} está definida según la siguiente expresión:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_x|}} e^{-\left(\mathbf{x}^\# - \mathbf{x}\right)^T \cdot \Sigma_x^{-1} \cdot \left(\mathbf{x}^\# - \mathbf{x}\right) / 2} \quad (4.10)$$

La estimación de máxima verosimilitud para \mathbf{x} se obtiene minimizando el siguiente índice bajo la restricción lineal $\mathbf{A}\cdot\mathbf{x}=\mathbf{b}$.

$$J(\mathbf{x}) = (\mathbf{x}^\# - \mathbf{x})^T \cdot \Sigma_x^{-1} \cdot (\mathbf{x}^\# - \mathbf{x}) \quad (4.11)$$

Si se define $\Delta\mathbf{x}=\mathbf{x}^\#-\mathbf{x}$, el problema se reduce a minimizar $(\Delta\mathbf{x}^T \cdot \mathbf{A} \cdot \Delta\mathbf{x})$ bajo la restricción $\mathbf{A} \cdot \Delta\mathbf{x} = \mathbf{A} \cdot \mathbf{x} - \mathbf{b}$. Si es posible satisfacer la restricción, la estimación de máxima verosimilitud $\mathbf{x}^* = \mathbf{x}^\# - \Delta\mathbf{x}$ se obtiene según:

$$\mathbf{x}^* = \mathbf{x}^\# - \Sigma_x \cdot \mathbf{A}^T (\mathbf{A} \cdot \Sigma_x \cdot \mathbf{A}^T)^{-1} (\mathbf{A} \cdot \mathbf{x}^\# - \mathbf{b}) \quad (4.12)$$

La media de la estimación \mathbf{x}^* tiende al valor real \mathbf{x} y su matriz de covarianzas vale

$$\Sigma_{x^*} = \Sigma_x - \Sigma_x \cdot \mathbf{A}^T (\mathbf{A} \cdot \Sigma_x \cdot \mathbf{A}^T)^{-1} \mathbf{A} \cdot \Sigma_x \quad (4.13)$$

El residuo de la $\mathbf{J}(\mathbf{x}^*)$ de la función $\mathbf{J}(\mathbf{x})$ vale:

$$J(\mathbf{x}^*) = (\mathbf{A} \cdot \mathbf{x}^\# - \mathbf{b})^T (\mathbf{A} \cdot \Sigma_x \cdot \mathbf{A}^T)^{-1} (\mathbf{A} \cdot \mathbf{x}^\# - \mathbf{b}) \quad (4.14)$$

Este residuo es una variable χ^2 con m grados de libertad, donde m representa el rango de la matriz \mathbf{A} .

4.2.4. Evaluación de algoritmos

Una vez un algoritmo ha sido desarrollado, es necesario evaluar su comportamiento ante diferentes condiciones de cálculo. Este paso se puede realizar tanto con datos simulados o datos reales. En el capítulo anterior se testearon diferentes algoritmos de calibración de cámaras utilizando datos simulados y los resultados obtenidos con cada uno de ellos eran comparados con los restantes e incluso, también con los valores reales de los parámetros de la cámara ya que estos eran conocidos. En ningún momento se plantearon cuestiones relacionadas con la calidad de los parámetros obtenidos en relación con los datos de partida, o que aproximación de los parámetros se podría conseguir para una incertidumbre dada de los datos iniciales. Teniendo en cuenta estos argumentos, si se dispone de herramientas para medir la calidad de los resultados dada una calidad en las medidas iniciales, es posible mejorar la estimación de los parámetros de forma que la calidad de la estimación sea máxima para una incertidumbre de los datos de partida. En este apartado se describe una herramienta estadística que permite determinar el límite máximo al cual se puede llegar en la estimación, dada una incertidumbre en los datos de partida. Una vez definido este límite, será posible determinar que algoritmos utilizan mejor la información incierta contenida en los datos de partida.

Una posible vía para determinar la calidad de la estimación de los parámetros está basada en el cálculo del error residual que existe entre el vector de medidas iniciales $\mathbf{y}^\#$ y el vector de medidas \mathbf{y}^* , resultado de los parámetros estimados \mathbf{x}^* dado por la expresión $\mathbf{y}^* = \mathbf{A} \cdot \mathbf{x}^*$.

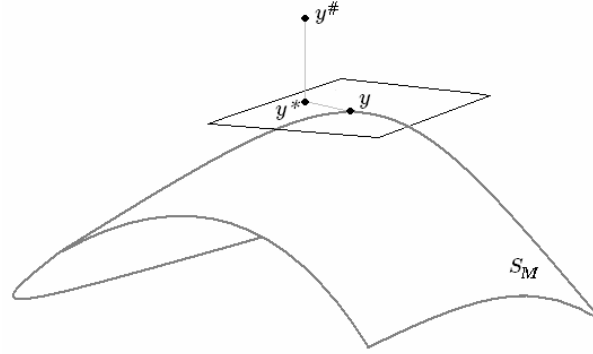
Para obtener un error residual esperado en cada uno de los datos de entrada se divide el error residual total por el número N de datos quedando la siguiente expresión:

$$e_r = \frac{1}{N} \sum_{i=1}^N d(y_i^{\#T} \cdot y_i^*)^2 \quad (4.15)$$

Suponiendo que el número de datos N coincide con el rango m de la matriz \mathbf{A} , la solución óptima calculada para los parámetros \mathbf{x}^* será tal que los datos medidos $\mathbf{y}^{\#}$, coincidirán exactamente con los datos resultado de los parámetros calculados \mathbf{y}^* . En este caso, el error residual que se obtendrá será cero, que bajo todos los puntos de vista, sería el mejor que se podría obtener. El vector de parámetros calculados \mathbf{x}^* satisface plenamente el conjunto de medidas $\mathbf{y}^{\#}$ contaminadas con ruido. Sin embargo, es necesario tener en cuenta que los valores exactos de los parámetros serán aquellos que cumplan con las medidas libres de ruido \mathbf{y} . Así pues en el caso de tener un número de datos igual al rango de la matriz \mathbf{A} , los parámetros del modelo calculado encajan perfectamente con el conjunto de medidas contaminadas con ruido, pero no dan una buena aproximación de los valores reales los cuales tendrían que satisfacer las medidas sin ruido. Por lo tanto, este error residual e_r no es una medida significativa de la calidad de la solución obtenida.

Si el número de datos n se incrementa, el valor del error residual se verá incrementado al mismo tiempo que el modelo calculado se aproxima mejor a los valores reales. Asintóticamente, la incertidumbre del modelo calculado debe disminuir de forma inversamente proporcional al número de datos de entrada, sin embargo, el error residual aumenta. Es por este motivo que el error residual en si mismo no es un buen índice para decidir la calidad de la estimación de los parámetros del modelo. Para que sea útil, es necesario comparar el valor obtenido con algún otro que sirva de referencia. Si se es capaz de definir un error residual esperado e_r que se supone que se debe obtener de forma teórica para un número y un nivel de incertidumbre en los datos de partida, se dispone de una herramienta para definir la calidad de los parámetros estimados ya que el error residual obtenido se podrá comparar con este valor teórico que se obtendría en el mejor de los casos.

Para definir el valor límite del error residual esperado e_r que debe obtenerse en cada caso, se va a utilizar la estimación de máxima verosimilitud. La estimación de máxima verosimilitud tiene en cuenta la incertidumbre de los datos de partida y el objetivo es proyectar esta incertidumbre en los parámetros estimados. Esta proyección dará el límite teórico que es posible alcanzar para el error residual en el caso de tener un conjunto N de datos con una determinada incertidumbre expresada mediante una varianza σ^2 . Dado otro error residual obtenido con otro algoritmo de estimación, es posible compararlo con este valor teórico que se puede alcanzar y tener así un grado de calidad en los parámetros del modelo calculados con ese otro algoritmo.



4.1. Interpretación geométrica de los errores en el espacio de medidas

Para obtener el valor teórico del error residual se realiza una interpretación geométrica de la estimación de máxima verosimilitud. Se tiene $\mathbf{y}^\#$ como un vector de medidas elegidas de acuerdo a la distribución Gaussiana isotópica con valor medio \mathbf{y} , y varianza $n \cdot \sigma^2$. La varianza total de distribución Gaussiana N -dimensional es la traza de la matriz de covarianzas, es decir, la suma de las varianzas en cada una de las direcciones axiales. Por lo tanto, para la distribución Gaussiana isotópica con varianzas independientes σ^2 , la varianza total será $N \cdot \sigma^2$. Como el vector de parámetros estimados \mathbf{x}^* , varía en torno al vector de parámetros reales \mathbf{x} , el valor de la función $f(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$, traza una superficie S_M en R^N en torno al punto definido por $\mathbf{y} = \mathbf{A} \cdot \mathbf{x}$. Esta superficie S_M se ve definida en la figura 4.1 y viene dada por el rango de la matriz \mathbf{A} . La dimensión de la superficie S_M , es un subespacio de R^N y es igual a d , donde d determina el número de parámetros necesarios para caracterizar el modelo. Dado un vector de medidas, la estimación de máxima verosimilitud \mathbf{y}^* es el punto en S_M más cercano a \mathbf{y} . El estimador de máxima verosimilitud es aquel que obtiene un vector de parámetros \mathbf{x}^* tal que el vector de medidas estimadas $\mathbf{y}^* = \mathbf{A} \cdot \mathbf{x}^*$ son las que están dentro de la superficie S_M tal como muestra la figura 4.1.

Se asume que en el entorno de \mathbf{y} , la superficie S_M es esencialmente plana y que se puede aproximar por la tangente en ese punto, por lo menos para un entorno del orden de magnitud del ruido en las medidas $\mathbf{y}^\#$. Dentro de esta aproximación lineal, las medidas resultado dadas por la estimación de máxima verosimilitud según $\mathbf{y}^* = \mathbf{A} \cdot \mathbf{x}^*$, son aquellas perpendiculares al plano tangente que pasan por el vector de medidas $\mathbf{y}^\#$. El error residual es la distancia desde el punto $\mathbf{y}^\#$ al valor estimado \mathbf{y}^* . La distancia desde el punto \mathbf{y}^* al punto siempre desconocido \mathbf{y} es la distancia desde el valor estimado de forma óptima al valor real, tal y como se observa en la figura 4.1. Se trata de obtener los valores esperados de estos errores y estos serán los límites teóricos que son posibles alcanzar para un valor de incertidumbre en las medidas y el número de datos determinado.

El error residual esperado con una estimación de máxima verosimilitud se reduce a obtener la distancia esperada de la variable aleatoria $\mathbf{y}^\#$ al hiperplano de dimensión d que pasa

por \mathbf{y} . $\mathbf{y}^\#$ está definida en R^N con media \mathbf{y} , y varianza $N \cdot \sigma^2$. La variable aleatoria Gaussiana definida en R^N proyectada sobre el plano tangente de dimensión d , da una distribución del error de estimación, es decir la diferencia entre el valor estimado \mathbf{y}^* y el valor real \mathbf{y} . La proyección en la superficie de dimensión $N-d$, normal a la tangente, da la distribución del error residual. Si se tiene en cuenta que de una distribución Gaussiana isotópica definida en R^N con varianza total igual $N \cdot \sigma^2$ proyectada en un subespacio de dimensión s , da como resultado una distribución isotópica Gaussiana con varianza total $s \cdot \sigma^2$, se puede obtener los valores esperados de las distancias expresadas anteriormente. En el caso de la estimación de los parámetros, se tiene que N medidas son modeladas como una función dependiente de un conjunto de d parámetros. Se supone que las medidas están sujetas a ruido Gaussiano con desviación típica σ en cada medida.

Por lo tanto, para el error residual la proyección se hace sobre la superficie de dimensión $N-d$. El error residual tendrá una varianza total de $(N-d) \cdot \sigma^2$. Para definir un valor del error residual esperado para cada dato será necesario calcular la desviación típica y dividir por el número de datos. El límite teórico de error residual esperado para cada dato es:

$$e_r^* = \sqrt{\frac{1}{N}(N-d)\sigma^2} = \sqrt{\left(1 - \frac{d}{N}\right)} \cdot \sigma \quad (4.16)$$

Este valor da un límite inferior para el error residual el cual sirve de referencia para compararlo con cualquier otra estimación realizada con otro algoritmo. Por otro lado, en el caso del error de estimación la proyección se realiza sobre el plano tangente de dimensión d . La varianza del error de estimación será $d \cdot \sigma^2$. Igual que antes, definiendo un error esperado para cada medida, el error de estimación se define como:

$$e_e^* = \sqrt{\frac{d}{N} \cdot \sigma^2} = \sqrt{\frac{d}{N}} \cdot \sigma \quad (4.17)$$

4.3. Estimación de parámetros basada en incertidumbres

Utilizando las herramientas estadísticas descritas en el apartado anterior, se va a detallar el proceso de estimación de parámetros teniendo en cuenta que los datos de partida son inciertos. Se considera un número de datos N los cuales se representan mediante un vector \mathbf{u}_i , $i=1 \dots N$. El conjunto de N datos se asume que son estadísticamente independientes ya que se obtienen de forma separada, pero los componentes de \mathbf{u}_i pueden estar correlados. Esta correlación se representa en la matriz de covarianzas asociada a cada dato $\Sigma \mathbf{u}_i$. Si \mathbf{u}_i es el valor real del dato obtenido en ausencia de ruido, el valor medido de dicho dato vendrá representado por $\mathbf{u}_i^\# = \mathbf{u}_i + \Delta \mathbf{u}_i$. Además, se considera un vector \mathbf{a} de dimensión m que representa los parámetros de una determinada función $R^M \times R^N$. El vector \mathbf{a} representa el lugar geométrico S_M

que satisfacen el conjunto de datos $\{\mathbf{u}_i\}$. Se supone que los valores reales $\mathbf{u}_1 \dots \mathbf{u}_N$ satisfacen l ecuaciones representadas por $f^{(k)}(\dots)$, que son funciones del vector \mathbf{u}_i , cuyos parámetros están contenidos en \mathbf{a} .

$$f^{(k)}(\mathbf{u}_i, \mathbf{a}) = 0 \quad (4.18)$$

Las l ecuaciones representan las funciones $R^M x R^N$ que son en definitiva las restricciones que tiene que cumplir los datos \mathbf{u}_i medidos bajo unos ciertos parámetros \mathbf{a} . \mathbf{u}_i acaban siendo valores siempre desconocidos. Estas restricciones no tienen porqué ser independientes. En caso de no serlo, r se denomina el número de ecuaciones independientes, es decir, el rango de la restricción. El problema a resolver en un primer lugar es corregir de forma óptima $\mathbf{u}_1, \dots, \mathbf{u}_N$ para que estas l restricciones parametrizadas con el vector \mathbf{a} dado, sean satisfechas. En segundo lugar, utilizando la expresión de esta corrección se pretende resolver el problema de estimar de forma óptima el valor del vector de parámetros \mathbf{a} , a partir del conjunto de datos $\mathbf{u}_1, \dots, \mathbf{u}_N$ y sus matrices de covarianzas $\Sigma \mathbf{u}_1, \dots, \Sigma \mathbf{u}_N$.

4.3.1. Corrección de los datos

Para un vector \mathbf{a} de parámetros dados se pretende obtener una corrección sobre las medidas realizadas para que cumplan plenamente las restricciones (4.18). Sustituyendo $\mathbf{u}_i^\# = \mathbf{u}_i + \Delta \mathbf{u}_i$ en $f^{(k)}(\mathbf{u}_i, \mathbf{a})$ se obtiene $f^{(k)}(\mathbf{u}_i^\#, \mathbf{a}) = f^{(k)}(\mathbf{u}_i, \mathbf{a}) + \nabla_{\mathbf{u}_i} f^{(k)} \cdot \Delta \mathbf{u}_i$. Realizando una aproximación lineal de primer orden y teniendo en cuenta que $f^{(k)}(\mathbf{u}_i, \mathbf{a}) = 0$, se tiene que

$$f^{(k)\#} = \nabla_{\mathbf{u}_i} f^{(k)T} \cdot \Delta \mathbf{u}_i \quad (4.19)$$

$f^{(k)\#}$ es la abreviatura de $f^{(k)}(\mathbf{u}_i^\#, \mathbf{a})$ y $\nabla_{\mathbf{u}_i} f^{(k)}$ representa la derivada parcial de la restricción $f^{(k)}(\mathbf{u}_i, \mathbf{a}) = 0$ respecto al vector \mathbf{u}_i evaluada en el valor real \mathbf{u}_i . Esta restricción linealizada es satisfecha si existe al menos una solución $\Delta \mathbf{u}_i$, que la cumple. En este caso la corrección toma la forma $\hat{\mathbf{u}}_i = \mathbf{u}_i^\# + \Delta \mathbf{u}_i$, para una aproximación de primer orden. Infinitas soluciones existen para la corrección $\Delta \mathbf{u}_i, i=1 \dots N$. De todas ellas, se elige aquella que minimiza el cuadrado de las distancias de Mahalanobis $|\Delta \mathbf{u}_i|_{\Sigma \mathbf{u}_i}$ bajo la restricción (4.19). Este índice a minimizar viene representado por:

$$J = \sum_{i=1}^N \left(\Delta \mathbf{u}_i^T \cdot \Sigma \mathbf{u}_i^{-1} \cdot \Delta \mathbf{u}_i \right) \quad (4.20)$$

Desde un punto de vista geométrico, esta corrección se interpreta como la proyección de $\mathbf{u}_i^\#$ en el punto más cercano al lugar geométrico S_M parametrizado por el vector \mathbf{a} determinado por (4.18). Gráficamente se observa en la figura 4.1 en la que se utiliza \mathbf{y} , en lugar de \mathbf{u} . La cercanía de este punto se mide según la distancia de Mahalanobis con respecto de la matriz de covarianza total. Este criterio está justificado como la estimación más probable si el

ruido que contamina las medidas tiene una distribución Gaussiana. Esto quiere decir que los errores $\Delta \mathbf{u}_1 \dots \Delta \mathbf{u}_N$ tienen una densidad de probabilidad de la forma:

$$p(\Delta \mathbf{u}_1, \dots, \Delta \mathbf{u}_N) = K \cdot e^{-\sum_{i=1}^N (\Delta \mathbf{u}_i^T \cdot \Sigma \mathbf{u}_i^{-1} \cdot \Delta \mathbf{u}_i) / 2} \quad (4.21)$$

Maximizar esta probabilidad es equivalente a minimizar el índice (4.20) sujeto a la restricción (4.19). Esta corrección es exacta si solamente existe un objeto que tenga que satisfacer el lugar geométrico S_M , es decir, si sólo se tiene un objeto, la posición más probable del mismo será el valor real ya que la campana de Gauss se centra sobre este valor. Teniendo en cuenta que existen N objetos las correcciones de cada uno son tales que minimizan el índice (4.20) planteado. Por otro lado la dirección en la cual se va a realizar la corrección $\Delta \mathbf{u}_i$ de cada punto vendrá determinada por los vectores propios contenidos en la matriz de covarianzas asociada a cada objeto geométrico $\Sigma \mathbf{u}_i$. Estos vectores dan los sentidos hacia los cuales es más probable que se encuentre el punto que satisface el lugar geométrico. El cálculo de $\Delta \mathbf{u}_i$ se realiza mediante una estimación de máxima verosimilitud del índice (4.20) sujeta a las restricciones (4.19), tal como se describió en el apartado 4.2.3. El desarrollo se hace en el siguiente apartado.

4.3.2. Cálculo del valor $\Delta \mathbf{u}_i$

En el cálculo del valor de $\Delta \mathbf{u}_i$ se utiliza la técnica de los multiplicadores de Lagrange para minimizar el índice (4.20) sujeto a la restricción (4.19).

$$J - 2 \sum_{k=1}^L \lambda_k \cdot \sum_{i=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)T} \cdot \Delta \mathbf{u}_i \right) \rightarrow \min \quad (4.22)$$

Derivando respecto a cada $\Delta \mathbf{u}_i$ e igualando a cero se obtiene el resultado

$$\Sigma \mathbf{u}_\alpha^{-1} \cdot \Delta \mathbf{u}_\alpha = \sum_{k=1}^L \lambda_k \cdot \nabla_{\mathbf{u}_i} f^{(k)} \quad (4.23)$$

Despejando $\Delta \mathbf{u}_i$ se tiene que

$$\Delta \mathbf{u}_i = \sum_{k=1}^L \lambda_k \cdot \Sigma \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(k)*} \quad (4.24)$$

Sustituyendo este valor en la ecuación de la restricción (4.19) se tiene:

$$f^{(k)} = \sum_{l=1}^L \left(\sum_{i=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)T} \cdot \Sigma \mathbf{u}_\alpha \cdot \nabla_{\mathbf{u}_i} f^{(l)} \right) \right) \lambda_l \quad (4.25)$$

Por lo tanto el valor de λ_l vendrá dado por:

$$\lambda_l = \sum_{k=1}^L W^{(kl)} f^{(k)} \quad (4.26)$$

$W^{(kl)}$ es el elemento (k,l) de la matriz $\mathbf{W}=\mathbf{V}^{-1}$, donde \mathbf{V} es la matriz de dimensiones $L \times L$ definida como

$$V = \sum_{i=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)T} \cdot \sum \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(l)} \right) \quad (4.27)$$

Finalmente, la corrección óptima $\Delta \mathbf{u}_i$ del objeto geométrico \mathbf{u}_i vendrá dada por la siguiente expresión:

$$\Delta \mathbf{u}_i = \sum \mathbf{u}_i \cdot \sum_{k,l=1}^L W^{(kl)} \cdot f^{(l)} \cdot \nabla_{\mathbf{u}_i} f^{(k)} \quad (4.28)$$

4.3.3. Consideraciones prácticas.

La corrección óptima $\Delta \mathbf{u}_i$ del dato \mathbf{u}_i dada por (4.28) obtenida en el apartado anterior es meramente una expresión teórica, ya que las matrices de covarianzas $\sum \mathbf{u}_i$ se tienen en cuenta evaluadas en los valores reales de los datos $\mathbf{u}_1, \dots, \mathbf{u}_N$, los cuales se quieren calcular. Estos datos se desconocen y solamente se sabe de ellos una aproximación dada por $\mathbf{u}_1^\#, \dots, \mathbf{u}_N^\#$. Por lo tanto, una posible aproximación es utilizar los datos conocidos $\mathbf{u}_1^\#, \dots, \mathbf{u}_N^\#$ en lugar de los desconocidos $\mathbf{u}_1, \dots, \mathbf{u}_N$, junto con sus matrices de covarianzas $\sum \mathbf{u}_i$ evaluadas en los datos contaminados con ruido $\mathbf{u}_1^\#, \dots, \mathbf{u}_N^\#$. El uso de datos aproximados a la hora de evaluar las derivadas parciales de las funciones $\nabla_{\mathbf{u}_i} f^{(k)}$ hace que la matriz \mathbf{V} tome la siguiente forma,

$$V^\# = \sum_{i=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)\#T} \cdot \sum \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(l)\#} \right) \quad (4.29)$$

El término $\#$ representa las funciones evaluadas con los datos contaminados. Si la matriz \mathbf{V} es aproximada con la matriz $\mathbf{V}^\#$, puede ocurrir que el rango de la matriz $\mathbf{V}^\#$ sea diferente al rango de matriz \mathbf{V} . Por lo tanto la matriz $W^{(kl)} = \mathbf{V}^{-1}$ será diferente de $W^{(kl)\#} = \mathbf{V}^{\#-1}$. Es por este motivo que una solución práctica a este problema es realizar la inversa de la matriz de covarianzas $\mathbf{V}^\#$ teniendo en cuenta esta restricción de rango, es decir, realizar la inversa de la matriz $\mathbf{V}^\#$ asumiendo que el rango esta matriz corresponde al de la matriz que representa \mathbf{V} . De esta forma, la corrección óptima $\Delta \mathbf{u}_i$ del objeto geométrico \mathbf{u}_i tendrán un valor aproximado dado $\Delta \hat{\mathbf{u}}_i$, el cual vale:

$$\Delta \hat{\mathbf{u}}_i = \sum \mathbf{u}_i^\# \cdot \sum_{k,l=1}^L W^{(kl)\#} \cdot f^{(l)} \cdot \nabla_{\mathbf{u}_i} f^{(k)\#} \quad (4.30)$$

El termino $W^{(kl)}$ se define teniendo en cuenta el rango de \mathbf{V} para los datos sin ruido.

$$W^{(kl)\#} = \left(\sum_{i=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)\#T} \cdot \sum \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(l)\#} \right) \right)_r^{-1} \quad (4.31)$$

Dado que esta corrección está basada en una aproximación lineal, los valores $\hat{\mathbf{u}}_i = \mathbf{u}_i^\# - \Delta \hat{\mathbf{u}}_i$ obtenidos con (4.30) no cumplirán exactamente las restricciones (4.19). Los valores $\hat{\mathbf{u}}_i$ que cumplen plenamente las restricciones (4.19) pueden ser calculados usando un método iterativo de forma que en cada iteración se utilizaran como \mathbf{u}_i los valores $\hat{\mathbf{u}}_i$ obtenidos en la iteración anterior. En este caso la convergencia es cuadrática y con dos o tres iteraciones es suficiente. En estas iteraciones la matriz de varianzas-covarianzas $\sum \mathbf{u}_i$ debe ser actualizada en cada iteración ya que la varianza de los nuevos valores estimados $\hat{\mathbf{u}}_i$ es diferente de la varianza de los datos de partida \mathbf{u}_i . Esta nueva varianza se calcula en el siguiente apartado.

4.3.4. Matriz de covarianzas de los datos corregidos

Aunque las restricciones (4.19) están impuestas en el método de optimización, los valores que se obtienen de los datos corregidos $\{\hat{\mathbf{u}}_i\}$ son variables aleatorias ya que están calculados a partir de determinadas aproximaciones y con datos contaminados con cierto ruido $\{\mathbf{u}_i^\#\}$. Llamando $\mathbf{u}_i^\# = \mathbf{u}_i + \Delta \mathbf{u}_i$ a los valores iniciales observados, $\Delta \mathbf{u}_i$ representa el error inicial para el dato $\mathbf{u}_i^\#$. Después, la corrección aproximada $\Delta \hat{\mathbf{u}}_i$ se calcula según la expresión (4.30) la cual modifica $\mathbf{u}_i^\#$ en una primera aproximación resultando en $\hat{\mathbf{u}}_i$ según $\hat{\mathbf{u}}_i = \mathbf{u}_i^\# - \Delta \hat{\mathbf{u}}_i$.

$$\begin{aligned} \hat{\mathbf{u}}_i &= (\mathbf{u}_i + \Delta \mathbf{u}_i) - \sum \mathbf{u}_i \cdot \sum_{k,l=1}^L W^{(kl)\#} \cdot f^{(k)}(\mathbf{u}_1^\#, \dots, \mathbf{u}_N^\#) \nabla_{\mathbf{u}_i} f^{(l)} = \\ &= \mathbf{u}_i + \left(\Delta \mathbf{u}_i - \sum \mathbf{u}_i \cdot \sum_{k,l=1}^L W^{(kl)\#} \cdot \sum_{j=1}^N \left(\nabla_{\mathbf{u}_j} f^{(k)\#T} \cdot \Delta \mathbf{u}_j \right) \nabla_{\mathbf{u}_i} f^{(l)} \right) \end{aligned} \quad (4.32)$$

Si se llama $\Delta \mathbf{u}_i^* = \hat{\mathbf{u}}_i - \mathbf{u}_i$, éste es el error para el valor corregido del dato $\hat{\mathbf{u}}_i$. La matriz de covarianzas para los datos corregidos $\{\hat{\mathbf{u}}_i\}$ se pueden obtener según $\Sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) = E(\Delta \mathbf{u}_i^* \cdot \Delta \mathbf{u}_j^{*\top})$. Sustituyendo $\Delta \mathbf{u}_i^*$ por sus correspondientes valores y operando, el resultado que se obtiene para la matriz de covarianzas de los objetos geométricos corregidos es:

$$\Sigma \hat{\mathbf{u}}_i = \Sigma \mathbf{u}_i - \sum_{k,l=1}^L W^{(kl)\wedge} \left(\sum \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(k)\wedge} \right) \left(\sum \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(l)\wedge} \right)^T \quad (4.33)$$

$$\Sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) = - \sum_{k,l=1}^L W^{(kl)\wedge} \left(\sum \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(k)\wedge} \right) \left(\sum \mathbf{u}_j \cdot \nabla_{\mathbf{u}_j} f^{(l)\wedge} \right)^T \quad (4.34)$$

$\nabla_{\mathbf{u}_i} f^{(k)\wedge}$ representa $\nabla_{\mathbf{u}_i} f^{(k)}(\hat{\mathbf{u}}_1 \dots \hat{\mathbf{u}}_N)$. Dado que $\Sigma(\hat{\mathbf{u}}_i, \hat{\mathbf{u}}_j) \neq \mathbf{0}$, esto indica que existe correlación entre los valores corregidos $\{\hat{\mathbf{u}}_i\}$ aunque los valores originales $\{\mathbf{u}_i^\#\}$ sean independientes.

4.3.5. Estimación del vector de parámetros

Hasta este momento es posible corregir los datos de entrada para que cumplan un lugar geométrico parametrizado con un vector \mathbf{a} . El objetivo ahora es determinar el valor del vector de parámetros \mathbf{a} , a partir de un conjunto de datos $\{\mathbf{u}_i^\#\}$. Se parte de N datos que satisfacen la misma relación geométrica $f^{(k)}(\mathbf{u}_i, \mathbf{a}) = \mathbf{0}$ parametrizada con el vector \mathbf{a} . Si $\Delta \mathbf{u}_i$ es una variable independiente aleatoria Gaussiana de media $\mathbf{0}$ y matriz de covarianzas $\Sigma \mathbf{u}_i$, la densidad de probabilidad conjunta de todos los $\Delta \mathbf{u}_i$ tiene la siguiente forma:

$$p(\Delta \mathbf{u}_1, \dots, \Delta \mathbf{u}_N) = K \cdot e^{-\sum_{i=1}^N \left(\Delta \mathbf{u}_i^T \cdot \Sigma \mathbf{u}_i^{-1} \cdot \Delta \mathbf{u}_i \right) / 2} \quad (4.35)$$

Esta densidad de probabilidad puede ser vista como la posibilidad de que los valores que se han medido sean los reales. Si J_i^* es el residuo de J_i para cada objeto geométrico obtenido cuando se sustituye la ecuación (4.28) en (4.20), la probabilidad de que los valores medidos sean los reales se transforma en la siguiente expresión:

$$p(\Delta \mathbf{u}_1, \dots, \Delta \mathbf{u}_N) = K \cdot e^{-\sum_{i=1}^N \left(J_i^{*T} \cdot \Sigma \mathbf{u}_i^{-1} \cdot J_i^* \right) / 2} \quad (4.36)$$

Esta probabilidad es solamente función del vector de parámetros \mathbf{a} . La estimación del vector de parámetros \mathbf{a} consiste en calcular aquel que maximiza esta probabilidad, es decir, el valor de máxima verosimilitud. Esto es equivalente a minimizar la suma de todos los residuos J_i^* los cuales están definidos en función del vector de parámetros \mathbf{a} . El índice de coste queda en función de este vector de parámetros $J(\mathbf{a})$.

$$J(\mathbf{a}) = \sum_{i=1}^N J_i^* \rightarrow \min \quad (4.37)$$

Sustituyendo la ecuación (4.28) en la ecuación (4.20) y operando, la función a minimizar tiene la siguiente forma:

$$J(\mathbf{a}) = \sum_{i=1}^N \sum_{k,l=1}^L W_i^{(kl)}(\mathbf{a}) \cdot f^{(k)}(\mathbf{u}_i^\#, \mathbf{a}) \cdot f^{(l)}(\mathbf{u}_i^\#, \mathbf{a}) \rightarrow \min \quad (4.38)$$

En la práctica, la función $J(\mathbf{a})$ no puede ser calculada con los datos disponibles ya que $W^{(kl)}(\mathbf{a})$ necesita los valores reales de los datos \mathbf{u}_i para ser calculada. Como en el apartado anterior, una matriz $W^{(kl)\#}(\mathbf{a})$ puede ser obtenida de forma aproximada realizando la inversa de la matriz $\mathbf{V}(\mathbf{a})$ teniendo en cuenta el rango de la misma.

$$W_i^{(kl)\#}(\mathbf{a}) = \left(\nabla_{\mathbf{u}_i} f^{(k)\#T} \cdot \Sigma \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(l)\#} \right)_r^{-1} \quad (4.39)$$

Utilizando el valor aproximado $W^{(kl)\#}(\mathbf{a})$ para la matriz $W^{(kl)}(\mathbf{a})$, la función a minimizar será $J(\mathbf{a})$. En principio una solución $\hat{\mathbf{a}}$ es obtenida por búsqueda numérica. Métodos iterativos del tipo Newton-Rapson o Levenberg-Marquardt deben ser utilizados para calcularla.

4.3.6. Matriz de covarianzas del vector de parámetros estimado

Se asume que se ha calculado una solución óptima $\hat{\mathbf{a}}$, la cual es una variable aleatoria ya que está calculada a partir del conjunto de datos $\{\mathbf{u}_i^\#\}$. Se asume que \mathbf{a} es el valor verdadero del vector de parámetros que satisface plenamente las hipótesis (4.18). Si se sustituye $\mathbf{u}_i^\# = \mathbf{u}_i + \Delta \mathbf{u}_i$ y $\mathbf{a}^\# = \mathbf{a} + \Delta \mathbf{a}$ en $f^{(k)}(\mathbf{u}_i, \mathbf{a})$ y se expande en \mathbf{u}_i y \mathbf{a} se obtiene:

$$f^{(k)}(\mathbf{u}_i, \mathbf{a}) = \nabla_{\mathbf{u}_i} f^{(k)\#T} \cdot \Delta \mathbf{u}_i + \nabla_{\mathbf{a}} f^{(k)\#T} \cdot \Delta \mathbf{a} + \mathcal{O}(\Delta \mathbf{u}_i, \Delta \mathbf{a})^2 \quad (4.40)$$

El término $\mathcal{O}(\Delta \mathbf{u}_i, \Delta \mathbf{a})^2$ incluye todos los términos de orden 2 y superiores. Teniendo en cuenta esta expresión, la ecuación (4.38) se puede expresar en función de las incertidumbres $\Delta \mathbf{u}_i$ y $\Delta \mathbf{a}$. Minimizando este índice se puede expresar la incertidumbre $\Delta \mathbf{a}$ en función de $\Delta \mathbf{u}_i$. Finalmente, teniendo en cuenta que $\Sigma(\hat{\mathbf{a}}) = E(\Delta \mathbf{a} \cdot \Delta \mathbf{a}^T)$ la varianza del vector de parámetros estimados se calcula con:

$$\Sigma \hat{\mathbf{a}} = \left(\sum_{i=1}^N \sum_{k,l=1}^L W_i^{(kl)}(\hat{\mathbf{a}}) \cdot \nabla_{\mathbf{a}} f^{(k)T}(\hat{\mathbf{u}}_i, \hat{\mathbf{a}}) \nabla_{\mathbf{a}} f^{(l)}(\hat{\mathbf{u}}_i, \hat{\mathbf{a}}) \right)^{-1} \quad (4.41)$$

$\{\hat{\mathbf{u}}_i\}$ representa el conjunto de datos corregidos de forma que sean los más probables para el vector de parámetros estimados. Sin embargo el conjunto inicial de datos $\{\mathbf{u}_i^\#\}$ puede ser utilizado para calcular la varianza de la estimación en una primera aproximación.

4.3.7. Testeo de hipótesis y estimación del nivel de ruido

Hasta el momento las restricciones (4.18) se asumen dadas. Sin embargo, éstas pueden ser hipotéticas y el procedimiento anterior de corrección del conjunto de datos que las satisfacen puede ayudar a discernir si estas hipótesis que cumplen los datos son ciertas o no. Se supone que hipotéticamente los valores verdaderos \mathbf{u}_i satisfacen las restricciones (4.18) y que $\mathbf{u}_i^\# = \mathbf{u}_i + \Delta \mathbf{u}_i$, $i=1 \dots N$, son los valores observados de forma que $\Delta \mathbf{u}_i$ es una variable aleatoria independiente con distribución Gaussiana de media $\mathbf{0}$ y matriz de covarianza $\Sigma \mathbf{u}_i$. Dado que $\nabla_{\mathbf{u}_i} f^{(k)}$ son valores determinísticos, la ecuación (4.18) implica que para una primera aproximación, $f^{(k)}$ es una variable aleatoria con distribución Gaussiana de media $\mathbf{0}$. Se puede calcular la covarianza de $f^{(k)}$ y de $f^{(l)}$ según $\Sigma(f^{(k)}, f^{(l)}) = E(\Delta f^{(k)} \cdot \Delta f^{(l)T})$, de la siguiente forma:

$$\begin{aligned} \Sigma(f^{(k)}, f^{(l)}) &= E \left(\sum_{i=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)T} \cdot \Delta \mathbf{u}_i \right) \sum_{j=1}^N \left(\nabla_{\mathbf{u}_j} f^{(l)T} \cdot \Delta \mathbf{u}_j \right) \right) = \sum_{i,j=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)T} \cdot E(\Delta \mathbf{u}_i \cdot \Delta \mathbf{u}_j) \cdot \nabla_{\mathbf{u}_j} f^{(l)} \right) = \\ &= \sum_{i,j=1}^N \left(\nabla_{\mathbf{u}_i} f^{(k)T} \cdot \Sigma \Delta \mathbf{u}_i \cdot \nabla_{\mathbf{u}_j} f^{(l)} \right) = V^{(kl)} \end{aligned} \quad (4.42)$$

La matriz $W = W^{(kl)}$ definida según la expresión (4.31) tiene rango r y por lo tanto la forma cuadrática,

$$J^* = \sum_{k,l=1}^L \left(W^{(kl)} \cdot f^{(k)T} \cdot f^{(l)} \right) \quad (4.43)$$

es una variable χ^2 chi cuadrado con r grados de libertad. Se puede demostrar que el valor de J^* definido según la ecuación (4.43) coincide con el valor obtenido cuando se sustituye la expresión (4.28) en el índice a minimizar (4.20). Esto significa que J^* es el residuo de la optimización (4.20). Si este valor es mucho mayor que el supuesto para el nivel de ruido, las hipótesis iniciales (4.18) deben ser cuestionadas. Por lo tanto, las hipótesis (4.18) pueden ser testeadas mediante χ^2 . Las hipótesis son rechazadas con un nivel $d\%$ si $J^* > \chi_{r,d}^2$, donde $\chi_{r,d}^2$ es el valor de semejanza $d\%$ de un valor χ^2 con r grados de libertad. De forma intuitiva se puede decir que las hipótesis representadas por (4.18) son rechazadas si el desplazamiento del conjunto de datos $\{\mathbf{u}_i\}$ para que cumplan las hipótesis (4.18) es muy grande medido en distancia de Mahalanobis.

Por otro lado, en muchos casos las características geométricas del ruido que perturba las medidas pueden ser predichas. Sin embargo, el valor absoluto de la magnitud del ruido es muy difícil de estimar a priori. En estos casos, se puede escribir la matriz de covarianzas $\sum \mathbf{u}_i$ de la forma que $\sum \mathbf{u}_i = \varepsilon^2 \sum_o \mathbf{u}_i$, $i=1 \dots N$, donde $\sum_o \mathbf{u}_i$ representa las características geométricas del ruido conocidas a priori representadas en una matriz de varianzas-covarianzas normalizada y ε^2 representa la magnitud del nivel de ruido desconocida a priori. De las expresiones (4.28) y (4.31) se ve que la corrección óptima no depende de la multiplicación de $\sum \mathbf{u}_i$ por una constante positiva arbitraria. Esto significa que en vez de utilizar $\sum \mathbf{u}_i$ se puede utilizar la matriz de varianzas covarianzas normalizada $\sum_o \mathbf{u}_i$ para calcular la corrección. Por lo tanto, no es necesario conocer el valor absoluto del ruido para obtener el valor óptimo de la corrección.

Una vez la solución óptima está calculada, es posible calcular el nivel del ruido ε^2 a posteriori. Si se usa la matriz normalizada de varianzas-covarianzas $\sum_o \mathbf{u}_i$, el residuo resultante J^* es igual a $\varepsilon^2 \cdot J_0^*$. Dado que J^* es una variable χ^2 con r grados de libertad, su media y su varianza son r y $2 \cdot r$ respectivamente. Por lo tanto una estimación $\varepsilon^{2\#}$ de ε^2 es:

$$\varepsilon^{2\#} = \frac{J_0^*}{r} \quad (4.44)$$

La media de esta estimación es ε^2 y su varianza viene dada por $\sum \varepsilon^2 = (2 \cdot \varepsilon^2) / r$.

4.4. Calibración de la cámara con las incertidumbres en las medidas

En este apartado se va a optimizar el proceso de calibrado de un cámara descrito en el capítulo 3.6 utilizando la técnica de estimación de parámetros basada en incertidumbres descrita en el apartado 4.3 de este capítulo.

Por un lado, el método de calibración de la cámara está basado en el proceso presentado por Zhang [103], [104], al cual se ha añadido un preprocesado de la información de partida y una mejora en el índice de coste que perfecciona la búsqueda no lineal. El preprocesado consiste en corregir errores en las medidas debidos a la distorsión que introducen las lentes, y también normalizar las mismas para condicionar mejor el proceso de cálculo del algoritmo. En el cálculo no lineal de los parámetros se incluye la búsqueda de las coordenadas de los puntos de la plantilla para que el ruido de éstas no afecte al proceso de calibración.

Por otro lado se tiene una técnica de búsqueda de parámetros basada en incertidumbres descrita en el apartado anterior, la cual puede resumirse de la siguiente manera. Se parte de un conjunto de N datos $\{\mathbf{u}_i^\#\}$ los cuales se suponen están contaminados con un ruido de media nula y varianza expresada según la matriz de covarianzas $\Sigma \mathbf{u}_i$. Dicho conjunto de datos cumplen un lugar geométrico o modelo representado mediante un vector de parámetros \mathbf{a} que lo definen. Para cada dato $\mathbf{u}_i^\#$ se cumple que $f^{(k)}(\mathbf{u}_i, \mathbf{a}) = \mathbf{0}$, $i=1\dots N$, $k=1\dots L$, donde \mathbf{u}_i es el valor real del dato, N es el número de datos y L es el rango de las restricciones que definen el modelo. El objetivo es estimar el vector de parámetros \mathbf{a}^* más probable a partir del conjunto de datos contaminados con ruido $\{\mathbf{u}_i^\#\}$. Para ello el índice que se minimiza es el siguiente:

$$J(\mathbf{a}) = \sum_{i=1}^N \sum_{k,l=1}^L W_i^{(kl)\#}(\mathbf{a}) \cdot f^{(k)}(\mathbf{u}_i^\#, \mathbf{a}) \cdot f^{(l)}(\mathbf{u}_i^\#, \mathbf{a}) \rightarrow \min \quad (4.45)$$

$W^{(kl)}(\mathbf{a})$ corresponde a:

$$W_i^{(kl)\#}(\mathbf{a}) = \left(\nabla_{\mathbf{u}_i} f^{(k)\#T} \cdot \Sigma \mathbf{u}_i \cdot \nabla_{\mathbf{u}_i} f^{(l)\#} \right)^{-1} \quad (4.46)$$

El vector de parámetros estimado \mathbf{a}^* tiene una incertidumbre ya que está calculado a partir de datos con ruido. La incertidumbre de la estimación viene expresada mediante la matriz de varianzas $\Sigma \mathbf{a}^*$ según:

$$\Sigma \mathbf{a}^* = \left(\sum_{i=1}^N \sum_{k,l=1}^L W_i^{(kl)\#}(\mathbf{a}^*) \cdot \nabla_{\mathbf{a}} f^{(k)T}(\hat{\mathbf{u}}_i, \mathbf{a}^*) \nabla_{\mathbf{a}} f^{(l)}(\hat{\mathbf{u}}_i, \mathbf{a}^*) \right)^{-1} \quad (4.47)$$

$\hat{\mathbf{u}}_i$ son las medidas corregidas que satisfacen los parámetros \mathbf{a}^* . Particularizando esta técnica de estimación basada en incertidumbres para el proceso de calibrado descrito anteriormente, es necesario determinar en cada una de las etapas del calibrado, que elementos constituyen los datos de entrada $\{\mathbf{u}_i^\#\}$ junto con las incertidumbres que los caracterizan $\Sigma \mathbf{u}_i$, y que elementos son los parámetros a estimar que forman el vector \mathbf{a} . Una vez definidos los datos de partida y los parámetros a estimar se establecen las funciones $f^{(k)}(\mathbf{u}_i, \mathbf{a}) = \mathbf{0}$ que relacionan ambos y a partir de aquí sólo queda determinar las derivadas parciales $\nabla_{\mathbf{u}_i} f^{(k)}$ para poder aplicar las expresiones anteriores (4.45), (4.46), (4.47). Como resultado de cada etapa debe obtenerse un conjunto de parámetros óptimos \mathbf{a}^* junto con la incertidumbre asociada a cada uno de ellos $\Sigma \mathbf{a}^*$. Para evaluar la incertidumbre asociada a cada parámetro estimado es necesario obtener en

conjunto de medidas corregidas \hat{u}_i que satisfacen los parámetros \mathbf{a}^* . Este proceso aplicado a cada una de las etapas de calibración, se describen en los subapartados siguientes.

Para la calibración de la cámara basada en datos inciertos, no se realiza normalización de los datos ya que las mejoras introducidas no son muy significativas. Este hecho se observó en el apartado de resultados experimentales del apartado 3.7. Por lo tanto, con el objetivo de simplificar el método de calibración, se utilizarán directamente los datos corregidos de la distorsión para estimar las homografías sin realizar una normalización de los mismos.

4.4.1. Corrección de la distorsión en la imagen

En esta primera etapa del proceso de calibración se trata de corregir las distorsiones producidas por defecto de fabricación en las lentes. Esta corrección se realiza partiendo del hecho que las líneas rectas en la imagen deben permanecer rectas a pesar de la posición y orientación de la cámara. En este caso los datos de partida son las coordenadas de los puntos medidos en la imagen $\{\mathbf{q}_{di}^\#\}$ y como resultado se obtienen un conjunto de puntos $\{\mathbf{q}_{pi}^*\}$ que satisfacen lo mejor posible las rectas ideales que deberían cumplir en el caso de no existir distorsión ni ruido en las medidas. En este momento es necesario diferenciar las diferentes formas de expresar las coordenadas de un punto i de la imagen. Sus coordenadas distorsionadas contaminadas con ruido se expresan mediante $\mathbf{q}_{di}^\#$, sus coordenadas libres de distorsión contaminadas con ruido son $\mathbf{q}_{pi}^\#$ y las coordenadas \mathbf{q}_{pi}^* son aquellas libres de ruido que mejor satisfacen la recta r y que por lo tanto el nivel de ruido que las contamina es menor que el de $\mathbf{q}_{di}^\#$. Las coordenadas libres de ruido y distorsión se nombran como \mathbf{q}_{pi} .

El proceso para estimar los parámetros que modelan la distorsión y corregir las medidas de los puntos $\{\mathbf{q}_{di}^\#\}$ para obtener $\{\mathbf{q}_{pi}^\#\}$ de describió en el apartado 3.5.1 y se pueden resumir en los siguientes pasos:

1. Detectar los puntos que forman rectas r en la imagen.
2. Calcular la pendiente teórica s_r de cada recta $\mathbf{r}=(a,b,c)$. Este proceso se realiza por mínimos cuadrados utilizando la expresión de la recta $a \cdot u_p + b \cdot v_p + c = 0$ y el conjunto de puntos medidos de la imagen que supuestamente la forman $\{\mathbf{q}_{di}^\#\}$.
3. Calcular las pendientes $s_i(u_{di}, v_{di})$ de las rectas tangentes a la curva que forman los puntos distorsionados en cada uno de ellos mediante (3.82) y (3.84).
4. Calcular los elementos a_{ij} que forman la matriz \mathbf{A} según (3.86).
5. Resolver los parámetros de la distorsión radial y tangencial k_l y p_l con mínimos cuadrados según (3.85).

6. Corregir las medidas de las coordenadas distorsionadas $\{q_{di}^{\#}\}$ para obtener las coordenadas libres de distorsión $\{q_{pi}^{\#}\}$ utilizando la expresión (3.78).

Básicamente se tienen dos procesos de estimación de parámetros en los pasos 2 y 5, y un proceso de corrección de medidas en el apartado 6. Desde un punto de vista de máxima verosimilitud se debería establecer la búsqueda de estos parámetros teniendo en cuenta la incertidumbre de las medidas en ambos casos y calculando los parámetros que definen la recta \mathbf{r} , así como los que miden la distorsión k_l y p_l con sus matrices de covarianzas. Estas matrices de covarianzas definirían la incertidumbre de los parámetros estimados. Además, en el paso 6, la corrección de la distorsión se debería realizar teniendo en cuenta que los parámetros que modelan la distorsión tienen una incertidumbre asociada. Aquí se tendría que propagar la incertidumbre de las medidas $\{q_{di}^{\#}\}$ expresadas en una matriz de covarianzas, a las medidas $\{q_{pi}^{\#}\}$. Éstas serían las que cumplirán la condición de máxima verosimilitud y por lo tanto podrían ser nombradas como $\{q_{pi}^{*}\}$.

A partir de este razonamiento, para corregir la distorsión de la imagen se dispone de todos los datos de partida excepto de las incertidumbres de las medidas de coordenadas de la imagen $\{q_{di}^{\#}\}$. Un método para estimar esta incertidumbre puede ser el definido en el apartado 4.3.7 mediante el cual se calculan las incertidumbres de las medidas que intervienen en un proceso de estimación de parámetros, a partir del residuo que genera dicha estimación. Dado que el índice que se está minimizando es una variable χ^2 , el residuo generado por la estimación J_0^* sirve para obtener un valor medio del error de las medidas expresado mediante ε^2 . Por lo tanto, teniendo en cuenta esta técnica se podría definir el nivel de ruido de las medidas $\{q_{di}^{\#}\}$ a partir del residuo generado con la estimación de los parámetros de la recta \mathbf{r} . De esta forma, se calcula un valor de error ε^2 a partir del residuo generado por la estimación J_0^* , y se asume una incertidumbre para cada medida de la forma $\sum q_{di}^{\#} = \varepsilon^2 \mathbf{I}$, donde \mathbf{I} es una matriz identidad de dimensiones dadas por el vector de medidas expresado en coordenadas homogéneas.

Si se tiene en cuenta que los puntos $\{q_{di}^{\#}\}$ utilizados para la estimación de la pendiente de la recta están sometidos a la distorsión de la lente y ellos mismos forman una curva que corresponde a la línea recta de la plantilla distorsionada, la estimación de los parámetros de la recta no serán del todo fiables con estos puntos, al mismo tiempo que el residuo que se genera de la estimación dará un valor de error inadecuado para los datos $\{q_{di}^{\#}\}$. Si se utiliza un valor inadecuado para la incertidumbre de las medidas, se cometerán errores en la estimación de los parámetros que modelan la distorsión así como en su incertidumbre. Por lo tanto, teniendo en cuenta que la estimación con este método no resultará del todo fiable, es necesario plantear un método alternativo para obtener un conjunto de puntos $\{q_{pi}^{*}\}$ que satisfacen desde el punto de vista de máxima verosimilitud las rectas que forman la plantilla de calibración. Este método alternativo toma como datos de entrada el conjunto de medidas corregidas $\{q_{pi}^{\#}\}$ obtenidas

como resultado del proceso de calibración de la distorsión realizado tal y como se describe en el apartado 3.5.1. Estas medidas $\{q_{pi}^{\#}\}$ han sido corregidas para que cumplan las rectas estimadas a partir de los datos $\{q_{di}^{\#}\}$. Por lo tanto se puede realizar una segunda corrección de $\{q_{pi}^{\#}\}$ para obtener $\{q_{pi}^{*}\}$ desde un punto de vista de máxima verosimilitud con el objetivo de obtener los valores $\{q_{pi}^{*}\}$ con un menor nivel de ruido y una incertidumbre que lo defina.

De este modo, lo que se va a realizar es una corrección del conjunto de puntos $\{q_{pi}^{\#}\}$ tal y como se describe en el apartado 4.3.1. Para cada punto $q_{pi}^{\#}$ se obtendrá una corrección optima $\Delta q_{pi}^{\#}$ desde el punto de vista de máxima verosimilitud, de forma que se cumple que $q_{pi}^{*} = q_{pi}^{\#} + \Delta q_{pi}^{\#}$. Este $\Delta q_{pi}^{\#}$ se calcula según el apartado 4.3.2. El nuevo conjunto de puntos corregidos $\{q_{pi}^{*}\}$ forman la supuesta recta ideal libre de ruido. Dado que la estimación de la corrección $\Delta q_{pi}^{\#}$ se realiza a partir de los datos inciertos $\{q_{pi}^{\#}\}$, el conjunto de puntos obtenidos como resultado $\{q_{pi}^{*}\}$ tendrá una incertidumbre expresada con una matriz de covarianzas la cual se puede calcular con el método descrito en apartado 4.3.4.

Por otro lado para realizar la corrección de los datos $\{q_{pi}^{\#}\}$ es necesario calcular la recta que mejor se adapta a estos. Hasta el momento se dispone de una recta la cual fue obtenida a partir de coordenadas sometidas a distorsión $\{q_{di}^{\#}\}$. Con el objetivo de obtener la recta que mejor se adapta al conjunto de datos corregidos $\{q_{pi}^{\#}\}$ y obtener el nivel de ruido de las medidas sin distorsión $\{\Delta q_{pi}^{\#}\}$, se va a determinar un proceso de estimación de máxima verosimilitud para la recta definida por $a \cdot u_{pi} + b \cdot v_{pi} + c = 0$. La incertidumbre de los datos corregidos vendrá dada, ya que el índice a minimizar que se plantea es una función χ^2 con $N-2$ grados de libertad. Este proceso se describió en el apartado 4.3.7.

Por lo tanto, se define el proceso de corrección de la distorsión del siguiente modo. Primero se calculan los parámetros que modelan la distorsión k_l y p_l tal y como se describe en el apartado 3.5.1. Esto permite corregir la distorsión de las medidas $\{q_{di}^{\#}\}$ y obtener el conjunto de puntos libres de distorsión $\{q_{pi}^{\#}\}$. En segundo lugar se estima la recta $a \cdot u_{pi} + b \cdot v_{pi} + c = 0$ que mejor satisface el conjunto de puntos $\{q_{pi}^{\#}\}$ desde el punto de vista de máxima verosimilitud. Esto permite conocer los parámetros de la recta libre de distorsión y las incertidumbres de las medidas $\{q_{pi}^{\#}\}$. En tercer lugar, se corrigen las medidas $\{q_{pi}^{\#}\}$ de forma que cumplan la recta ideal calculada anteriormente. El resultado es un conjunto de puntos $\{q_{pi}^{*}\}$ libre de distorsión y con un nivel de ruido menor que $\{q_{pi}^{\#}\}$ dado que satisfacen la recta que mejor se adapta a la imagen de la plantilla libre de distorsión. Los pasos de estimación de la recta y corrección de las medidas se describen a continuación.

Estimación de la recta libre de distorsión en la imagen

En este caso se utiliza el proceso de estimación de parámetros de máxima verosimilitud definido por las expresiones (4.45), (4.46), (4.47). Se utilizan como datos de entrada las coordenadas de los puntos corregidas libres de distorsión $\{q_{pi}^{\#}\}$ y se pretende estimar los parámetros de la recta $r=(a,b,c)$ definida por la función $f(r, q_{pi}^{\#})=r^T \cdot q_{pi}^{\#}$ para que se cumpla que $a \cdot u_{pi}^{\#} + b \cdot v_{pi}^{\#} + c = 0$. Dado que no se conocen a priori la matriz de covarianzas que determinan la incertidumbre de los datos de entrada $\{q_{pi}^{\#}\}$, se asume esta matriz de covarianzas como la matriz identidad I , cuyo elemento (3,3) tiene valor cero ya que no se considera incertidumbre en la coordenada homogénea. El rango de la función $f(r, q_{pi}^{\#})$ vale 1. El índice a minimizar está definido por la expresión (4.45). Si se particulariza para la estimación de los parámetros de la recta, teniendo en cuenta que $|r|=1$, éste queda como sigue:

$$J_0(r) = \sum_{i=1}^N r^T \cdot q_{pi}^{\#} \rightarrow \min \quad (4.48)$$

El residuo $J_0^*(r)$ es una variable χ^2 con $N-2$ grados de libertad. Obviamente, el número N de puntos en la imagen debe ser tal que $N \geq 2$. El valor de las incertidumbres de las medidas $\{q_{pi}^{\#}\}$ vendrá dado por la siguiente expresión:

$$\varepsilon^2 = \frac{J_0^*(r)}{N-2} \quad (4.49)$$

Si se tienen sólo 2 puntos, el residuo es nulo y por lo tanto se entiende que no existe incertidumbre en las medidas ya que la recta calculada satisface plenamente los puntos utilizados. Dado que la estimación de la incertidumbre se basa en el residuo del índice minimizado, el cual mide la aproximación de los datos a los parámetros estimados, es necesario tener un número de datos que permita una estimación fiable.

Corrección de los puntos en la imagen

Atendiendo a la expresión (4.30) es posible corregir el conjunto de puntos $\{q_{pi}^{\#}\}$ de forma que mejor se adapte a la recta estimada en el paso anterior. Particularizando la expresión (4.30) se corrige cada punto según:

$$\Delta q_{pi}^{\#} = \sum q_{pi}^{\#} \cdot \frac{r^T \cdot q_{pi}^{\#}}{r^T \cdot \sum q_{pi}^{\#} \cdot r} \cdot r \quad (4.50)$$

El resultado es el conjunto de puntos $\{q_{pi}^{\#}\}$ el cual se obtendrá con $q_{pi}^* = q_{pi}^{\#} + \Delta q_{pi}^{\#}$. La varianza del nuevo conjunto de puntos se calcula utilizando la expresión (4.33). En este caso se particulariza según la siguiente expresión:

$$\sum q_{pi}^* = \sum q_{pi}^{\#} - \frac{\sum q_{pi}^{\#} \cdot r \cdot r^T \cdot \sum q_{pi}^{\#}}{r^T \cdot \sum q_{pi}^{\#} \cdot r} \quad (4.51)$$

De esta forma se consigue corregir el conjunto de puntos medidos de la imagen de forma que la distorsión queda eliminada y el nivel de ruido se reduce bastante ya que todos ellos pertenecen a las rectas que más se aproximan a las definidas por la imagen de la plantilla libre de distorsión y ruido. Dado que el proceso parte de un conjunto de puntos inciertos, el resultado también presenta cierta incertidumbre expresada por $\sum q_{pi}^*$. Los datos q_{pi}^* , junto con sus incertidumbres $\sum q_{pi}^*$ se utilizan en el siguiente paso para estimar las homografías que relacionan la plantilla con las diferentes imágenes que se toman de estas desde diferentes posiciones.

4.4.2. Corrección de errores en la plantilla

Del mismo modo que se producen errores de medida en las coordenadas de los puntos en la imagen, también se producen errores de medida en las coordenadas de los puntos en la plantilla. En este caso, los errores son procedentes del proceso de medida de las coordenadas en la plantilla y no están sujetos a ningún tipo de distorsión, a parte de la propia que pueda tener la plantilla por ser un elemento físico el cual se supone plano y puede estar deformado en algún aspecto. Con el objetivo de definir el orden de magnitud del error en las coordenadas de los puntos en la plantilla y establecer una corrección que disminuya este error, se utiliza un procedimiento parecido a la corrección de los puntos en la imagen. Dado que la plantilla está formada por puntos que satisfacen una línea recta, se pueden estimar los parámetros de estas rectas para ver en que medida las coordenadas de los puntos las cumplen. Este índice da el error de las coordenadas tal y como se explica en el apartado 4.3.7 y se ha utilizado posteriormente para la estimación del error en las coordenadas en la imagen. Posteriormente, una vez se conoce la recta que deben satisfacer los puntos, es posible corregir los mismos para que la cumplan. En este caso el ruido de las coordenadas corregidas es inferior al de las coordenadas iniciales.

Se tienen el conjunto de coordenadas de los puntos en la plantilla $\{p_i^\#\}$ y se estiman los parámetros de la recta $r=(a,b,c)$ definida por la función $f(r,p_i^\#)=r^T \cdot p_i^\#$ para que se cumpla que $a \cdot x_i^\# + b \cdot y_i^\# + c = 0$. Se asume una matriz identidad I como matriz de covarianzas inicial para $\{p_i^\#\}$. El índice a minimizar para obtener los parámetros de la recta, teniendo en cuenta que $|r|=1$ es:

$$J_0(r) = \sum_{i=1}^N r^T \cdot p_i^\# \rightarrow \min \quad (4.52)$$

El residuo $J_0^*(r)$ da la incertidumbre de las medidas $\{p_i^\#\}$ según la expresión

$$\varepsilon^2 = \frac{J_0^*(r)}{N-2} \quad (4.53)$$

Una vez conocida la recta se corrigen los puntos $\{p_i^\#\}$ para que se adapte a ésta. La corrección para cada punto viene dada por la siguiente expresión:

$$\Delta p_i^\# = \sum p_i^\# \cdot \frac{r^T \cdot p_i^\#}{r^T \cdot \sum p_i^\# \cdot r} \cdot r \quad (4.54)$$

El nuevo conjunto de puntos se obtiene con $p_i^* = p_i^\# + \Delta p_i^\#$. La varianza del nuevo conjunto de puntos viene dada por la siguiente expresión:

$$\sum p_i^* = \sum p_i^\# - \frac{\sum p_i^\# \cdot r \cdot r^T \cdot \sum p_i^\#}{r^T \cdot \sum p_i^\# \cdot r} \quad (4.55)$$

De esta forma se consigue corregir el conjunto de coordenadas de los puntos de la plantilla y obtener una incertidumbre que define el intervalo en el que están contenidos.

4.4.3. Estimación de las homografías

En este caso los datos de partida son las coordenadas óptimas de los puntos en la plantilla $\{p_i^*\}$ y sus correspondientes en la imagen $\{q_{pi}^*\}$. A las coordenadas de los puntos en cada imagen se les ha corregido la distorsión producida por las lentes y el nivel de ruido ya que todos los puntos cumplen una línea recta. Los puntos de la plantilla están corregidos para que cumplan la recta a la que pertenecen. Los datos de partida se expresan en coordenadas homogéneas y ordenados en un vector de datos d_i , $d_i = (x_i, y_i, 1, u_i, v_i, 1)$, donde x_i, y_i representan las coordenadas del punto en la plantilla y u_i, v_i son las coordenadas de su correspondiente en la imagen. Estos valores se obtienen de $\{p_i^*\}$ y $\{q_{pi}^*\}$. $\sum u_i$ es una matriz que representa la incertidumbre de los puntos en la imagen y en la plantilla. Ésta se forma a partir de las matrices $\sum q_{pi}^*$, $\sum p_i^*$ obtenidas en el paso de corrección de datos. $\sum u_i$ incluye también las correlaciones que existen entre las coordenadas de los puntos.

El objetivo es estimar los elementos de las homografías H_β , $\beta=1..M$, donde β representa el número de imagen tomada de la plantilla. Para cada homografía se dispone de un conjunto de puntos en la imagen $\{^{\beta}q_{pi}^*\}$ pero el conjunto de las coordenadas en la plantilla son las mismas $\{p_i^*\}$. Los elementos de la matriz que forma la homografía se ordenan en un vector h_β . En este caso las relaciones que existen entre los elementos de la matriz que representa la homografía h_β y las coordenadas de los puntos en la imagen y la plantilla son las siguientes:

$$\begin{aligned} g^{(1)}(h_\beta, d_i) &= x_i \cdot h_{\beta 11} + y_i \cdot h_{\beta 12} + h_{\beta 13} - u_i \cdot x_i \cdot h_{\beta 31} - u_i \cdot y_i \cdot h_{\beta 32} - u_i \cdot h_{\beta 33} = 0 \\ g^{(2)}(h_\beta, d_i) &= x_i \cdot h_{\beta 21} + y_i \cdot h_{\beta 22} + h_{\beta 23} - v_i \cdot x_i \cdot h_{\beta 31} - v_i \cdot y_i \cdot h_{\beta 32} - v_i \cdot h_{\beta 33} = 0 \end{aligned} \quad (4.56)$$

Para obtener los elementos de la homografía que mejor satisfacen el conjunto de puntos en la plantilla y en la imagen desde el punto de vista de máxima verosimilitud, el índice a minimizar es el dado por las expresiones (4.45) y (4.46). Particularizado para este caso queda:

$$\mathcal{J}[h_\beta] = \sum_{i=1}^N \sum_{k,l=1}^2 W_i^{(kl)}(h_\beta) \cdot g^{(k)}(h_\beta, d_i) \cdot g^{(l)}(h_\beta, d_i) \rightarrow \min \quad (4.57)$$

$$W_i^{(kl)}(h_\beta) = \left(\nabla_{d_i} g^{(k)}(h_\beta, d_i) \cdot \sum d_i \cdot \nabla_{d_i} g^{(l)}(h_\beta, d_i) \right)_r^{-1}$$

Siguiendo con el procedimiento de máxima verosimilitud la expresión (4.47) da la varianza de los elementos estimados de la homografía. En este caso queda la siguiente expresión:

$$\sum h_\beta^* = \left(\sum_{i=1}^N \sum_{k,l=1}^2 W_i^{(kl)}(h_\beta^*) \cdot \nabla_h g^{(k)}(h_\beta^*, d_i^*) \cdot \nabla_h g^{(l)}(h_\beta^*, d_i^*) \right)_r^{-1} \quad (4.58)$$

En este momento se dispone de las relaciones que existen entre los elementos de la homografía y las coordenadas de los puntos tanto en la plantilla como en la imagen. Utilizando las expresiones anteriores, es posible estimar los parámetros de la homografía y su varianza. Sólo es necesario definir los términos $\nabla_h g^{(1)}$, $\nabla_h g^{(2)}$, $\nabla_{d_i} g^{(1)}$ y $\nabla_{d_i} g^{(2)}$ los cuales son:

$$\nabla_h g^{(1)}(h_\beta, d_i) = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -u_i \cdot x_i & -u_i \cdot y_i & -u_i \end{bmatrix} \quad (4.59)$$

$$\nabla_h g^{(2)}(h_\beta, d_i) = \begin{bmatrix} 0 & 0 & 0 & x_i & y_i & 1 & -v_i \cdot x_i & -v_i \cdot y_i & -v_i \end{bmatrix} \quad (4.60)$$

$$\nabla_{d_i} g^{(1)}(h_\beta, d_i) = \begin{bmatrix} h_{\beta 11} - u_i \cdot h_{\beta 31} & h_{\beta 12} - u_i \cdot h_{\beta 32} & 0 & -x_i \cdot h_{\beta 31} - y_i \cdot h_{\beta 32} - h_{\beta 33} & 0 & 0 \end{bmatrix} \quad (4.61)$$

$$\nabla_{d_i} g^{(2)}(h_\beta, d_i) = \begin{bmatrix} h_{\beta 21} - v_i \cdot h_{\beta 31} & h_{\beta 22} - v_i \cdot h_{\beta 32} & 0 & 0 & -x_i \cdot h_{\beta 31} - y_i \cdot h_{\beta 32} - h_{\beta 33} & 0 \end{bmatrix} \quad (4.62)$$

Por lo tanto, utilizando estas expresiones es posible realizar una estimación de los elementos que componen la homografía \mathbf{H}_β entre la plantilla y la imagen β , desde un punto de vista de máxima verosimilitud. Será necesario realizar una estimación para cada imagen β tomada desde diferente posición y orientación. Además, también es posible asociar a cada homografía la incertidumbre correspondiente representada en la matriz de covarianzas $\sum h_\beta^*$. Esta incertidumbre se calcula a partir de la expresión (4.58). En esta expresión existe un término llamado d_i^* . Este término corresponde a las coordenadas de los puntos en la imagen y la plantilla que satisfacen la homografía estimada. Será necesario calcular una nueva corrección para cada vector de coordenadas las cuales cumplan la homografía estimada desde el punto de vista de máxima verosimilitud. En teoría esta matriz de covarianzas se tendría que calcular en el valor exacto de las coordenadas libres de ruido pero dado que éstas son desconocidas, se calculan en las coordenadas que satisfacen la homografía, haciendo una aproximación tal y como se explica en el apartado 4.3.3. En este caso se podría realizar una segunda aproximación utilizando el conjunto de coordenadas $\{d_i\}$. Sin embargo, para obtener un valor más cercano al real es necesario corregir las mismas para obtener $\{d_i^*\}$. Atendiendo a la expresión (4.30) es posible corregir el conjunto de puntos $\{d_i\}$ de forma que mejor se adapte a la homografía estimada en el paso anterior. Particularizando la expresión (4.30) se corrige cada punto según:

$$\Delta d_i = \sum d_i \cdot \sum_{k,l=1}^2 W_i^{(kl)} \cdot g^{(l)} \cdot \nabla_{d_i} g^{(k)} \quad W_i^{(kl)} = \left(\sum_{i=1}^N \left(\nabla_{d_i} g^{(k)T} \cdot \sum d_i \cdot \nabla_{d_i} g^{(l)} \right) \right)_r^{-1} \quad (4.63)$$

El vector de coordenadas que satisface la homografía se calcula según $\mathbf{d}_i^* = \mathbf{d}_i + \Delta \mathbf{d}_i$.

4.4.4. Estimación de los elementos de la matriz $\mathbf{A}^{-T} \cdot \mathbf{A}^{-1}$

En esta etapa del proceso de calibrado de la cámara se estiman los elementos de la matriz $\mathbf{A}^{-T} \cdot \mathbf{A}^{-1}$. Dado que la matriz $\mathbf{A}^{-T} \cdot \mathbf{A}^{-1}$ es diagonal, dichos elementos se ordenan en un vector \mathbf{b} . En este caso, la estimación del vector \mathbf{b} se realiza teniendo en cuenta la incertidumbre que tienen los elementos de las homografías. Los datos utilizados para la estimación de este vector son las homografías \mathbf{h}_β , así como la matriz de covarianzas $\Sigma \mathbf{h}_\beta$ correspondiente a cada una de ellas calculadas en el paso anterior. El procedimiento de estimación de los elementos del vector \mathbf{b} es similar a los anteriores. Se tiene unos datos de partida con ciertas incertidumbres y unas relaciones que se tienen que cumplir según unos parámetros a estimar. En este caso la relación que existe entre los elementos del vector \mathbf{b} y las diferentes homografías de cada imagen de la plantilla es la siguiente:

$$f^{(1)}(\mathbf{h}_\beta, \mathbf{b}) = (h_{\beta 11}^2 - h_{\beta 12}^2) \mathbf{b}_{11} + 2(h_{\beta 21} \cdot h_{\beta 11} - h_{\beta 22} \cdot h_{\beta 12}) \mathbf{b}_{12} + (h_{\beta 21}^2 - h_{\beta 22}^2) \mathbf{b}_{22} + \\ + 2(h_{\beta 31} \cdot h_{\beta 11} - h_{\beta 32} \cdot h_{\beta 12}) \mathbf{b}_{13} + 2(h_{\beta 31} \cdot h_{\beta 21} - h_{\beta 22} \cdot h_{\beta 32}) \mathbf{b}_{23} + (h_{\beta 31}^2 - h_{\beta 32}^2) \mathbf{b}_{33} \quad (4.64)$$

$$f^{(2)}(\mathbf{h}_\beta, \mathbf{b}) = h_{\beta 11} \cdot h_{\beta 12} \cdot \mathbf{b}_{11} + (h_{\beta 11} \cdot h_{\beta 22} + h_{\beta 21} \cdot h_{\beta 12}) \mathbf{b}_{12} + h_{\beta 21} \cdot h_{\beta 22} \cdot \mathbf{b}_{22} + \\ + (h_{\beta 31} \cdot h_{\beta 12} + h_{\beta 11} \cdot h_{\beta 32}) \mathbf{b}_{13} + (h_{\beta 31} \cdot h_{\beta 22} + h_{\beta 21} \cdot h_{\beta 32}) \mathbf{b}_{23} + h_{\beta 31} \cdot h_{\beta 32} \cdot \mathbf{b}_{33} \quad (4.65)$$

El rango de esta restricción es 2. En este caso el índice a minimizar será:

$$J[\mathbf{b}] = \sum_{\beta=1}^M \sum_{k,l=1}^2 W_\beta^{(kl)}(\mathbf{b}) \cdot f^{(k)}(\mathbf{h}_\beta, \mathbf{b}) \cdot f^{(l)}(\mathbf{h}_\beta, \mathbf{b}) \rightarrow \min W_\beta^{(kl)}(\mathbf{b}) = \left(\nabla_{\mathbf{h}_\beta} f^{(k)}(\mathbf{h}_\beta, \mathbf{b}) \cdot \Sigma \mathbf{h}_\beta \cdot \nabla_{\mathbf{h}_\beta} f^{(l)}(\mathbf{h}_\beta, \mathbf{b}) \right)^{-1} \quad (4.66)$$

La varianza del vector \mathbf{b}^* estimado se calcula según la siguiente expresión

$$\Sigma \mathbf{b}^* = \left(\sum_{\beta=1}^M \sum_{k,l=1}^2 W_\beta^{(kl)}(\mathbf{b}^*) \cdot \nabla_{\mathbf{b}} f^{(k)}(\mathbf{b}^*, \mathbf{h}_\beta^*) \cdot \nabla_{\mathbf{b}} f^{(l)}(\mathbf{b}^*, \mathbf{h}_\beta^*) \right)^{-1} \quad (4.67)$$

Los términos $\nabla_{\mathbf{h}_\beta} f^{(1)}$, $\nabla_{\mathbf{h}_\beta} f^{(2)}$, $\nabla_{\mathbf{b}} f^{(1)}$ y $\nabla_{\mathbf{b}} f^{(2)}$ se definen de la siguiente forma:

$$\nabla_{\mathbf{h}_\beta} f^{(1)}(\mathbf{h}_\beta, \mathbf{b}) = \begin{bmatrix} 2 \cdot h_{\beta 11} \cdot \mathbf{b}_{11} + 2 \cdot h_{\beta 21} \cdot \mathbf{b}_{12} + 2 \cdot h_{\beta 31} \cdot \mathbf{b}_{13} \\ 2 \cdot h_{\beta 12} \cdot \mathbf{b}_{11} - 2 \cdot h_{\beta 22} \cdot \mathbf{b}_{12} - 2 \cdot h_{\beta 32} \cdot \mathbf{b}_{13} \\ 0 \\ 2 \cdot h_{\beta 11} \cdot \mathbf{b}_{12} + 2 \cdot h_{\beta 21} \cdot \mathbf{b}_{22} + 2 \cdot h_{\beta 31} \cdot \mathbf{b}_{23} \\ - 2 \cdot h_{\beta 12} \cdot \mathbf{b}_{12} - h_{\beta 22} \cdot \mathbf{b}_{22} - h_{\beta 32} \cdot \mathbf{b}_{23} \\ 0 \\ 2 \cdot h_{\beta 11} \cdot \mathbf{b}_{13} + 2 \cdot h_{\beta 21} \cdot \mathbf{b}_{23} + 2 \cdot h_{\beta 31} \cdot \mathbf{b}_{33} \\ - 2 \cdot h_{\beta 12} \cdot \mathbf{b}_{13} - 2 \cdot h_{\beta 22} \cdot \mathbf{b}_{23} - 2 \cdot h_{\beta 32} \cdot \mathbf{b}_{33} \\ 0 \end{bmatrix}; \quad \nabla_{\mathbf{h}_\beta} f^{(2)}(\mathbf{h}_\beta, \mathbf{b}) = \begin{bmatrix} h_{\beta 12} \cdot \mathbf{b}_{11} + h_{\beta 22} \cdot \mathbf{b}_{12} + h_{\beta 32} \cdot \mathbf{b}_{13} \\ h_{\beta 11} \cdot \mathbf{b}_{11} + h_{\beta 21} \cdot \mathbf{b}_{12} + h_{\beta 31} \cdot \mathbf{b}_{13} \\ 0 \\ h_{\beta 12} \cdot \mathbf{b}_{12} + h_{\beta 22} \cdot \mathbf{b}_{22} + h_{\beta 32} \cdot \mathbf{b}_{23} \\ h_{\beta 11} \cdot \mathbf{b}_{12} + h_{\beta 21} \cdot \mathbf{b}_{22} + h_{\beta 31} \cdot \mathbf{b}_{23} \\ 0 \\ h_{\beta 12} \cdot \mathbf{b}_{13} + h_{\beta 22} \cdot \mathbf{b}_{23} + h_{\beta 32} \cdot \mathbf{b}_{33} \\ h_{\beta 11} \cdot \mathbf{b}_{13} + h_{\beta 21} \cdot \mathbf{b}_{23} + h_{\beta 31} \cdot \mathbf{b}_{33} \\ 0 \end{bmatrix} \quad (4.68)$$

$$\nabla_b f^{(1)}(h_\beta, b) = \begin{bmatrix} h_{\beta_{11}}^2 - h_{\beta_{12}}^2 \\ 2(h_{\beta_{21}} \cdot h_{\beta_{11}} - h_{\beta_{22}} \cdot h_{\beta_{12}}) \\ h_{\beta_{21}}^2 - h_{\beta_{22}}^2 \\ 2(h_{\beta_{31}} \cdot h_{\beta_{11}} - h_{\beta_{32}} \cdot h_{\beta_{12}}) \\ 2(h_{\beta_{31}} \cdot h_{\beta_{21}} - h_{\beta_{22}} \cdot h_{\beta_{32}}) \\ h_{\beta_{31}}^2 - h_{\beta_{32}}^2 \end{bmatrix}; \quad \nabla_b f^{(2)}(h_\beta, b) = \begin{bmatrix} h_{\beta_{11}} \cdot h_{\beta_{12}} \\ h_{\beta_{11}} \cdot h_{\beta_{22}} + h_{\beta_{21}} \cdot h_{\beta_{12}} \\ h_{\beta_{21}} \cdot h_{\beta_{22}} \\ h_{\beta_{31}} \cdot h_{\beta_{12}} + h_{\beta_{11}} \cdot h_{\beta_{32}} \\ h_{\beta_{31}} \cdot h_{\beta_{22}} + h_{\beta_{21}} \cdot h_{\beta_{32}} \\ h_{\beta_{31}} \cdot h_{\beta_{32}} \end{bmatrix} \quad (4.69)$$

Igual que ha pasado en la homografía, para la estimación de la incertidumbre de los elementos del vector \mathbf{b}^* , aparecen los términos \mathbf{h}_β^* . Estos términos corresponden a aquellos que mejor satisfacen la homografía estimada y no a aquellos utilizados para la estimación que eran el resultado del paso anterior. Por este motivo es necesario corregir los elementos de las homografías para calcular las incertidumbres del vector \mathbf{b}^* estimado. Los vectores \mathbf{h}_β^* se calculan con $\mathbf{h}_\beta^* = \mathbf{h}_\beta + \Delta \mathbf{h}_\beta$, en donde $\Delta \mathbf{h}_\beta$ es una corrección dada por:

$$\Delta \mathbf{h}_\beta^* = \sum h_\beta \cdot \sum_{k,l=1}^2 W^{(kl)} \cdot f^{(l)} \cdot \nabla_{h_\beta} f^{(k)} \quad W^{(kl)} = \left(\sum_{\beta=1}^M \left(\nabla_{h_\beta} f^{(k)T} \cdot \sum h_\beta \cdot \nabla_{h_\beta} f^{(l)} \right) \right)^{-1} \quad (4.70)$$

4.4.5. Cálculo de los parámetros de la cámara a partir del vector \mathbf{b}

Para la obtención de los parámetros de la cámara, se utiliza la propagación de la varianza la cual se define en el apartado 4.2.1. Si $\mathbf{x}^\#$ es un vector N -dimensional, aleatorio de media \mathbf{x} y matriz de covarianzas Σ_x , un vector $\mathbf{y}^\#$ M -dimensional definido por una matriz arbitraria \mathbf{A} de dimensiones $N \times M$ según $\mathbf{y} = \mathbf{A} \cdot \mathbf{x}$, $\mathbf{y}^\#$ es una variable aleatoria de distribución Gaussiana de media $\mathbf{A} \cdot \mathbf{x}$ y matriz de covarianzas $\mathbf{A} \cdot \Sigma_x \cdot \mathbf{A}^T$. La matriz \mathbf{A} representa la relación lineal entre los elementos del vector \mathbf{x} e \mathbf{y} . En el caso de que exista una relación no lineal entre ambos vectores $\mathbf{y} = f(\mathbf{x})$, la media de \mathbf{y} vale $f(\mathbf{x})$ y su varianza Σ_y se calcula a partir de linealizar la relación f no lineal para los valores del vector \mathbf{x}_0 .

$$\Sigma_y = \nabla_x f(\mathbf{x}_0) \cdot \Sigma_x \cdot \nabla_x f(\mathbf{x}_0)^T \quad (4.71)$$

De esta forma, para la obtención de los parámetros de la cámara, se tiene un vector \mathbf{b}^* aleatorio con matriz de covarianzas $\Sigma \mathbf{b}^*$ obtenida en el apartado anterior y se quiere calcular los parámetros de la cámara junto con sus incertidumbres. La relación entre los parámetros de la cámara $\alpha_u, \alpha_v, u_0, v_0, \mathbf{t}, \mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ y el vector \mathbf{b} viene dado por las expresiones (3.54) y (3.55). Los valores medios de los parámetros de la cámara se calculan directamente a partir de las expresiones (3.54) y (3.55). Para propagar la varianza de los elementos del vector \mathbf{b} a los parámetros de la cámara es necesario calcular las derivadas de las expresiones (3.54) y (3.55) con respecto a cada uno de los elementos que intervienen.

4.4.6. Búsqueda directa de los parámetros de la cámara

Con el proceso de calibración descrito a lo largo de los apartados 4.4.3 al 4.4.5 se realiza una búsqueda no lineal de los parámetros que intervienen en cada etapa desde un punto de vista de máxima verosimilitud. El objetivo es obtener los parámetros que mejor se adaptan a los datos de entrada de cada fase junto con sus incertidumbres para utilizarlos como datos de partida en la siguiente etapa del proceso de calibrado.

Un procedimiento alternativo consiste en utilizar las coordenadas de los puntos en la imagen y la plantilla para obtener directamente los parámetros de la cámara. El objetivo es minimizar una función de coste basada en el modelo de la cámara, teniendo en cuenta que se toman m imágenes de la plantilla y se tienen n puntos de interés en la plantilla. Como datos de partida para la búsqueda de los parámetros se toman los conjuntos de puntos corregidos tanto de la imagen como de la plantilla $\{\mathbf{p}_i^*\}$ y $\{\mathbf{q}_{pi}^*\}$ junto con sus matrices de covarianzas $\Sigma \mathbf{q}_{pi}^*$, $\Sigma \mathbf{p}_i^*$ obtenidas en el paso de corrección de datos descrito en los apartados 4.4.1 y 4.4.2 respectivamente. El vector a estimar se compone de los parámetros intrínsecos de la cámara junto con los extrínsecos para cada una de las posiciones desde las que se han tomado imágenes $\mathbf{a}=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$. Tal y como se vio en la parte experimental del capítulo 3 dedicado al método de calibración de la cámara, se observa que si se incluyen las coordenadas de los puntos en la plantilla, el error en la estimación sólo depende de los errores cometidos en las medidas de las coordenadas en las imágenes. En este caso el vector de elementos estimados es $\mathbf{a}=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$. Las coordenadas de los puntos en la plantilla pueden ser tanto las originales $\{\mathbf{p}_i^\#\}$ como las corregidas $\{\mathbf{p}_i^*\}$. La búsqueda no lineal de los parámetros corregirá las mismas de forma que mejor se adapten a la estimación. Como datos de partida se toman las coordenadas corregidas de los puntos en las imágenes $\{\mathbf{q}_{pi}^*\}$ junto con sus errores expresados en la matriz de covarianzas $\Sigma \mathbf{q}_{pi}^*$.

Las funciones $f^{(k)}(\mathbf{q}_{ij}, \mathbf{a})=0$ que relacionan los datos $\{\mathbf{q}_{ij}^*\}$ con el vector de parámetros \mathbf{a} son en este caso:

$$f^{(1)}(\mathbf{q}_{ij}, \mathbf{a}) = u_{ij} - \alpha_u \frac{r_{11j}x_i + r_{12j}y_i + r_{13j}z_i + t_{xj}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{zj}} - u_0$$

$$f^{(2)}(\mathbf{q}_{ij}, \mathbf{a}) = v_{ij} - \alpha_v \frac{r_{21j}x_i + r_{22j}y_i + r_{23j}z_i + t_{yj}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{zj}} - v_0$$
(4.72)

El subíndice i corresponde con el número de punto en la imagen nombrada con el subíndice j . Hay que tener en cuenta que en estas funciones aparecen los elementos de la matriz de rotación $r_{11j}, r_{12j}, \dots, r_{33j}$ y que se tienen como parámetros a estimar los ángulos θ_j, ϕ_j, ψ_j , que

representan la rotación en cada uno de los ejes de coordenadas. Los elementos $r_{11}, r_{12}, \dots, r_{33}$ se calculan a partir de los ángulos de rotación θ_j, ϕ_j, ψ_j según (3.76).

A partir de aquí sólo queda determinar las derivadas parciales $\nabla_{u_i} f^{(k)}$. Debido a la complejidad de la expresión dada por la relación que existe entre los ángulos θ_j, ϕ_j, ψ_j , y la matriz de rotación que forman $r_{11j}, r_{12j}, \dots, r_{33j}$, se opta por aproximar numéricamente el valor de la derivada de la función en cada punto.

4.5. Evaluación de algoritmos de calibración

En el apartado 4.2.4 se definió una regla que permitía saber cual era el error residual mínimo que se podía obtener de una estimación sabiendo la incertidumbre de los datos de partida utilizados. En concreto la expresión (4.16) definía ese límite teórico de error residual e_r^* que cualquier algoritmo debería alcanzar en el caso de sacar el máximo rendimiento a los datos inciertos de entrada que se utilizaban para la estimación. También se definía un error de estimación óptimo e_e^* .

$$e_r^* = \sqrt{\frac{1}{N}(N-d)\sigma^2} = \sqrt{\left(1 - \frac{d}{N}\right)} \cdot \sigma \qquad e_e^* = \sqrt{\frac{d}{N}\sigma^2} = \sqrt{\frac{d}{N}} \cdot \sigma \qquad (4.73)$$

Estos errores dependen tanto del número N de datos utilizados para la estimación como del número de parámetros d estimados. En el caso de disponer de un número infinito de datos, es posible saber la tendencia del error. Desde un punto de vista práctico, dado que es imposible tener un conjunto infinito de elementos de entrada para una estimación, se puede establecer un número mínimo de elementos con los que conseguir un error de estimación determinado, o que a partir del cual, la mejora significativa de este error no sea importante. Por lo tanto, por un lado se tiene un error residual teórico que permite comparar los resultados de diferentes algoritmos, y por el otro se establece un error de estimación que permite calcular ese número mínimo de datos necesarios para conseguir una estimación de parámetros dentro de unos límites de error de estimación. Dado que el proceso de calibrado de la cámara acaba realizándose en diferentes etapas en cada una de las cuales se estiman diferentes parámetros, el cálculo del error residual y el error de estimación permitirá comparar algoritmos de calibración y definir el número mínimo de elementos necesarios para realizar una estimación de forma fiable. La tabla 4.1 muestra los resultados en cada caso. n indica el número de puntos que se dispone en la plantilla de calibración y j indica el número de imágenes tomadas de la plantilla. Así pues, en el caso de calcular los parámetros de la recta u homografías, n es el número de puntos que intervienen en la estimación. Si se trata de calcular los elementos de la matriz $\mathbf{A}^{-T} \cdot \mathbf{A}^{-1}$, j indica las homografías que participan en la estimación, cada una de las cuales corresponden a una de las imágenes tomadas. El caso de obtener los parámetros de la cámara a partir del modelo de la misma se

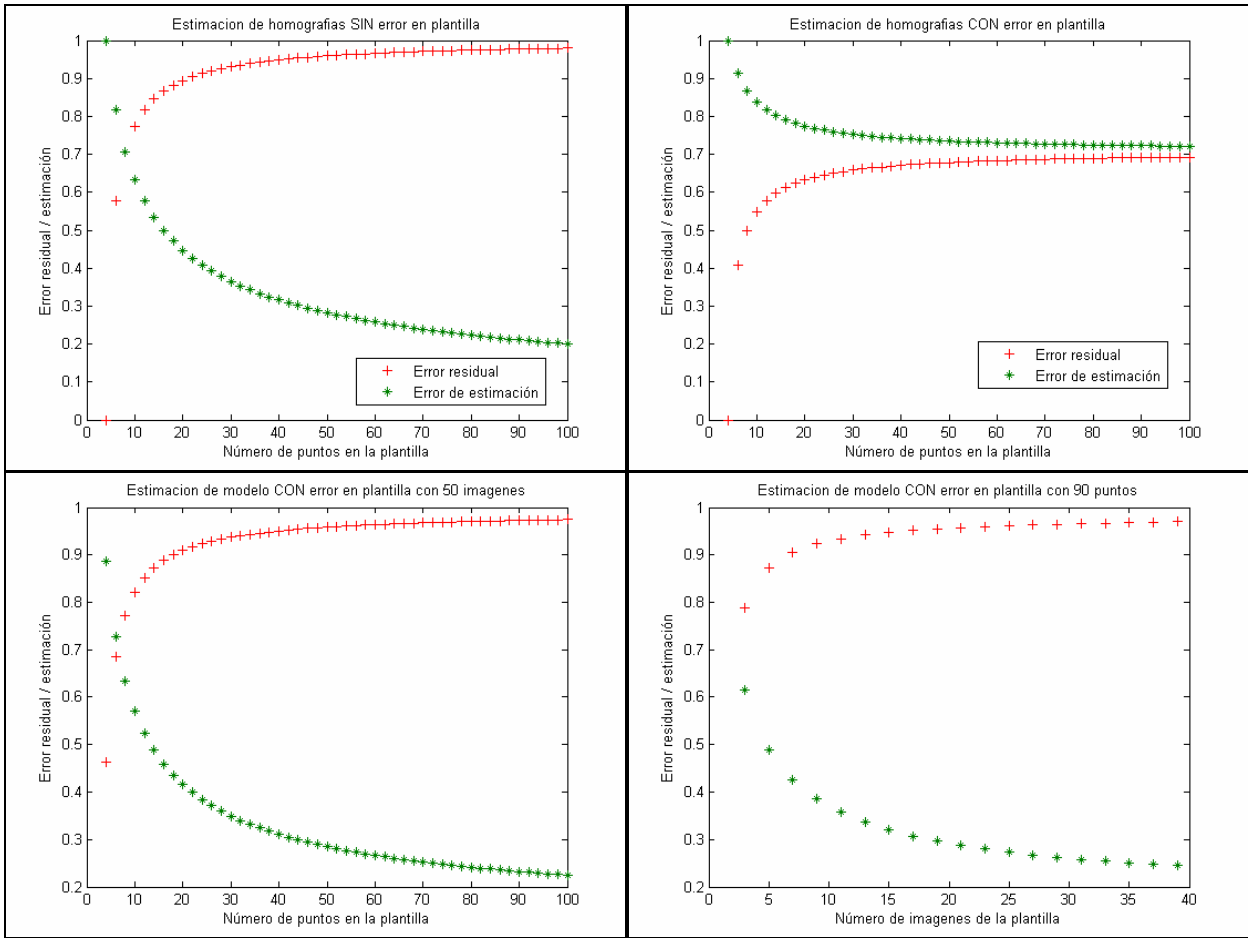
	N	d	e_r^*	e_e^*
Parámetros de la recta	$2n$	2	$e_r^* = \sqrt{\left(1 - \frac{1}{n}\right)} \cdot \sigma$	$e_e^* = \sqrt{\left(\frac{1}{n}\right)} \cdot \sigma$
Homografía SIN ruido en la plantilla	$4n$	8	$e_r^* = \sqrt{\left(1 - \frac{4}{n}\right)} \cdot \sigma$	$e_e^* = \sqrt{\left(\frac{4}{n}\right)} \cdot \sigma$
Homografía CON ruido en la plantilla	$4n$	$8+2n$	$e_r^* = \sqrt{\left(\frac{n-4}{2n}\right)} \cdot \sigma$	$e_e^* = \sqrt{\left(\frac{4+n}{2n}\right)} \cdot \sigma$
Matriz $\mathbf{A}^T \cdot \mathbf{A}^{-1}$	$8j$	6	$e_r^* = \sqrt{\left(1 - \frac{3}{4j}\right)} \cdot \sigma$	$e_e^* = \sqrt{\left(\frac{3}{4j}\right)} \cdot \sigma$
Modelo de la cámara SIN ruido en la plantilla	$2n+2n \cdot j$	$6+6j$	$e_r^* = \sqrt{\left(1 - \frac{3}{n}\right)} \cdot \sigma$	$e_e^* = \sqrt{\left(\frac{3}{n}\right)} \cdot \sigma$
Modelo de la cámara CON ruido en la plantilla	$2n \cdot j$	$6+6j+2n$	$e_r^* = \sqrt{\left(\frac{(2n-6) \cdot j - (2n+6)}{2n \cdot j}\right)} \cdot \sigma$	$e_e^* = \sqrt{\left(\frac{6+6j+2n}{2n \cdot j}\right)} \cdot \sigma$

4.1 Error residual y de estimación para cada una de las etapas del proceso de calibrado de la cámara

estiman los parámetros intrínsecos de la cámara junto con los parámetros extrínsecos que definen la posición y orientación de la cámara en cada una de las imágenes. Como datos de entrada se tienen las coordenadas de los puntos en la plantilla y las coordenadas de los mismos en cada una de las imágenes.

Realizando un estudio detallado de las expresiones obtenidas en la tabla 4.1, se pueden obtener las siguientes conclusiones. En el caso de la estimación de las homografías, el error de estimación decrece de forma asintótica con el número de puntos utilizados. En el caso de disponer del conjunto de puntos en la plantilla sin ruido el límite de esta asíntota es cero cuando el número de puntos es infinito. Sin embargo, en el caso de tener ruido en las coordenadas de la plantilla, el límite de esta asíntota es $\sigma/\sqrt{2}$. En el caso de no tener ruido en la plantilla, dado que se tiene una de las medidas del punto perfecta es posible estimar la homografía de forma asintóticamente exacta aumentando el número de puntos. En el caso de tener contaminadas las dos medidas de un punto, la estimación nunca puede ser perfecta ya que no se tiene ningún valor de referencia exacto para realizar la estimación. Se tienen dos medidas de la posición de cada punto, una en la plantilla y otra en la imagen, relacionadas con la matriz de transformación llamada homografía, ninguna de las cuales es correcta.

Otra conclusión interesante está relacionada con la búsqueda no lineal de los parámetros de la cámara basada en el modelo de la misma. En este caso, j representa el número de imágenes tomadas de la plantilla desde diferentes posiciones y orientaciones. Cabe destacar



4.2 Error residual y de estimación en el cálculo de los elementos de una homografía con y sin ruido en las coordenadas de la plantilla y de los parámetros de la cámara con ruido en la plantilla variando el número de imágenes y el número de puntos

como resultado que en el caso de disponer de medidas en las coordenadas de la plantilla exactas, el número de imágenes que se tomen de la misma, no influye en el proceso de mejora de la estimación. En este caso sólo depende del número de puntos contenidos en la plantilla y la tendencia del error de estimación es muy parecida a la estimación de la homografía, la cual sólo depende del número de puntos que intervienen en la estimación. Este dato es muy interesante desde el punto de vista del número de imágenes a tomar en el proceso de calibración basado en plantilla plana. En el caso de tener ruido en las coordenadas de los puntos en la plantilla, la tendencia del error de estimación es $\sigma/\sqrt{2n}$. Igual que en el caso de la estimación de la homografía, dado que no se tiene ningún dato exacto es imposible calcular los valores exactos de los parámetros y el error de estimación dependerá del número de puntos utilizados.

Gráficamente se han representado los errores residuales y de estimación para el caso de la homografía y la búsqueda no lineal de los parámetros de la cámara basada en el modelo de la misma. Para la representación se toma $\sigma=1$. Se comparan los errores que se consiguen en el caso de tener coordenadas de los puntos de la plantilla con y sin ruido. En el caso de la homografía se representan los errores frente al número de puntos que intervienen en el proceso de cálculo. En

el caso de la búsqueda no lineal de los parámetros de la cámara basada en el propio modelo, se representan frente al número de puntos utilizados y número de imágenes solamente para el caso de tener ruido en las coordenadas de los puntos de la plantilla y las imágenes. El caso de no tener ruido en las coordenadas de la plantilla no se representa, dado que frente al número de puntos su comportamiento es parecido al de las homografías, y frente al número de imágenes tomadas, los errores no experimentan ninguna variación. Estas gráficas son muy útiles para comparar los errores residuales obtenidos con los diferentes algoritmos de calibración. Además ayudan a definir el número de puntos o imágenes necesarias para conseguir un error de estimación dentro de unos límites.

Por ejemplo, si se pretende definir el número de puntos necesarios para diseñar una plantilla que garantice unos errores de estimación mínimos, la representación del error de estimación indica que con 40 puntos son suficientes en el caso de tener ruido en las coordenadas de la plantilla. La mejora que se consigue en la estimación es poco significativa a partir de ese número de puntos. En el caso de no tener coordenadas con ruido en la plantilla 80~90 puntos pueden ser suficientes. En teoría llegando a infinitos puntos se pueden conseguir errores de estimación nulos. En el caso de determinar el número de imágenes a tomar, si se consiguen coordenadas en la plantilla sin ruido, 3 imágenes son suficientes para poder resolver linealmente los parámetros de la cámara y después utilizarlos como punto de partida para la búsqueda no lineal. El número de puntos en la plantilla se define a partir de la homografía. El número de imágenes a capturar en el caso de tener coordenadas en la plantilla contaminadas con ruido dependerá del número de puntos utilizados en la estimación. En las gráficas correspondientes a los errores geométricos y de estimación de la búsqueda no lineal de los parámetros de la figura 4.2, se utilizan 50 imágenes y 90 puntos. Teniendo en cuenta estos resultados con 90 puntos en la plantilla, en torno a 20 imágenes serían suficientes para obtener un error de estimación aceptable. También se podría decir a con 50 imágenes hacen falta en torno a 40~50 puntos.

Estos valores se obtienen suponiendo como aceptable un error de estimación en los parámetros en torno al 30% del error de los datos de entrada. Obviamente todas estas expresiones y valores aquí obtenidos están basadas en valores teóricos los cuales pocos algoritmos pueden alcanzar. Sirva pues estos valores para utilizarlos de referencia para definir las condiciones de calibración de la cámara.

4.6. Resultados experimentales

En este apartado se obtienen los resultados correspondientes a los algoritmos de calibración descritos en el capítulo. Dado que las posibilidades son múltiples se van a testear diferentes algoritmos basándose en los resultados del apartado 3.7 del capítulo anterior. El

objetivo en este caso es ver la mejora en la estimación de los parámetros dado que se tiene en cuenta las incertidumbres en las medidas. También se pretende ver como evoluciona la incertidumbre de los parámetros estimados cuando cambian las condiciones de calibración. Al mismo tiempo que se mejora la estimación de los parámetros y se calcula su incertidumbre se van a evaluar los diferentes algoritmos para entender hasta que punto obtienen el máximo rendimiento de los datos de entrada. En este momento se disponen de herramientas para determinar la eficacia del algoritmo e incluso para obtener los valores mínimos del error de estimación. Las expresiones se reflejan en la tabla 4.1.

Se tiene un algoritmo de calibración tolerante a los errores de las coordenadas de los puntos en la plantilla, el cual se describe en la segunda parte del apartado 4.4.6. Desde el punto de vista de tiempo de cómputo, este método es el más costoso. Sin embargo, es la técnica más robusta de calibración ya que los resultados obtenidos no se ven afectados por los errores de medida en los puntos de la plantilla. En segundo lugar ordenado según el tiempo de cómputo, se tiene el algoritmo que busca directamente los parámetros de la cámara tomando como datos de partida las coordenadas de los puntos en la plantilla y en la imagen. En este caso se obtiene una mejora significativa de los resultados frente al cálculo lineal de los parámetros y el tiempo de cómputo se reduce bastante frente al método anterior.

Otro procedimiento de calibración plantea el caso de realizar el proceso de cálculo de los parámetros de la cámara a partir del método de calibración de Zhang realizando una búsqueda no lineal de los parámetros en cada una de las etapas, para obtener unos parámetros óptimos junto con sus incertidumbres. La ventaja a priori de este método puede ser el desacoplamiento de los parámetros. Al realizar la búsqueda no lineal de todos los parámetros de la cámara, los parámetros que están acoplados dentro del modelo se compensan los errores y esto hace que se minimice el índice buscado sin obtener los parámetros correctos de la cámara. Si se realiza una búsqueda no lineal en cada una de las etapas de proceso de calibración de Zhang, este acoplamiento desaparece y se puede llegar a unos resultados menos inciertos que los obtenidos con la búsqueda directa de los parámetros.

El apartado se divide en seis partes. En cada una de ella se tiene en cuenta uno de los factores que afectan al proceso de calibrado utilizando una plantilla plana y se caracterizan sus efectos. En primer lugar se corrigen las coordenadas de los puntos medidos en la imagen y la plantilla para que cumplan las rectas teóricas que deberían cumplir en el caso de ausencia de ruido. En segundo lugar se tiene en cuenta el ruido en la plantilla. Seguidamente se cambia el valor del ruido en las coordenadas de los puntos de la imagen. En cuarto y quinto lugar se ven los efectos del número de puntos en la plantilla y el número de imágenes tomadas de la misma. Para terminar el apartado se presenta una comparativa de los diferentes algoritmos utilizados teniendo en cuenta el error residual que generan, el error de estimación y el tiempo de cómputo utilizado para llegar a la solución.

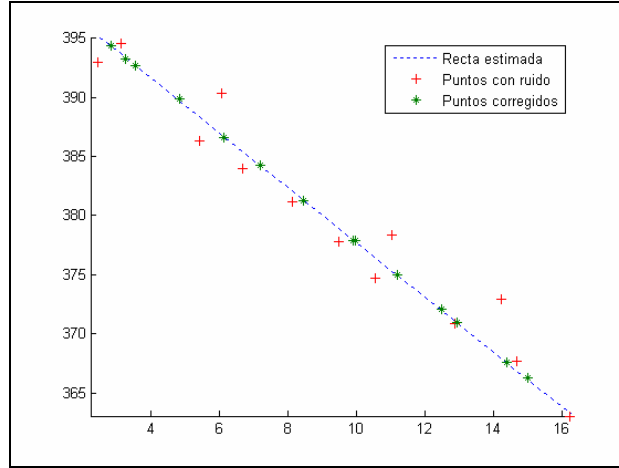
Todos los datos para evaluar los algoritmos son generados mediante simulación. Las condiciones de simulación son idénticas a las utilizadas en el apartado 3.7. Las características de la cámara son $\alpha_u=1250$, $\alpha_v=900$, $u_0=250$, $v_0=250$ y $\gamma=1.09083$ equivalente a 89.95° . La resolución de la imagen es 512×512 píxeles. La plantilla, de la misma forma que hace Zhang es un tablero de ajedrez de $10 \times 14 = 140$ esquinas. El tamaño de la plantilla es 180×250 mm. En este caso, la plantilla está situada de forma que todos los puntos de interés tienen coordenada $z_w=0$, según se muestra en la figura 3.4. Las posiciones de la cámara son $\mathbf{t}_1=(150, 200, 580)$, $\mathbf{t}_2=(-50, 250, 880)$, $\mathbf{t}_3=(100, -20, 820)$, $\mathbf{t}_4=(-40, -150, 780)$, $\mathbf{t}_5=(-150, -150, 530)$, $\mathbf{t}_6=(-100, 125, 400)$, $\mathbf{t}_7=(140, -150, 500)$, $\mathbf{t}_8=(240, 50, 600)$. La orientación de la cámara en cada una de las posiciones es tal que el eje óptico corta al plano de la plantilla en el centro de la misma. Además, la orientación también está calculada para que el eje de coordenadas sea paralelo al plano en el que se encuentra la plantilla.

Corrección de las coordenadas de la plantilla y de las imágenes

Una vez corregidos los errores de las coordenadas de los puntos producidos por las distorsiones que introducen los defectos de fabricación en la cámara, se procede a cambiar las coordenadas de los puntos de forma que cumplan las rectas teóricas que los definen. Dado que la plantilla esta formada por un tablero de ajedrez los puntos de interés en la plantilla cumplen sendas rectas. Por otro lado, dado que la imagen de cualquier recta acaba siendo otra recta, se corrigen los puntos de las imágenes para que se cumplan las imágenes de las rectas del tablero de ajedrez. La corrección se realiza desde el punto de vista de máxima verosimilitud con el objetivo de obtener una varianza que defina la incertidumbre de las medidas de las coordenadas. El procedimiento se describió en el apartado 4.4.1 y 4.4.2. Los resultados se muestran en la figura 4.3. Se observa como se tiene un valor corregido procedente de la distorsión y éste se cambia para que cumpla la recta teórica. Además se obtiene una incertidumbre de las coordenadas del nuevo punto que se utiliza en el proceso de calibrado de la cámara.

Ruido en las coordenadas de los puntos de la plantilla

Se toman como datos de partida para la búsqueda de los parámetros los conjuntos de puntos corregidos tanto de la imagen como de la plantilla $\{\mathbf{p}_i^*\}$ y $\{\mathbf{q}_{pi}^*\}$ junto con sus matrices de covarianzas $\Sigma \mathbf{q}_{pi}^*$, $\Sigma \mathbf{p}_i^*$ obtenidas en el paso anterior. Para obtener unos resultados simulados lo más parecidos a los reales, las medidas de las imágenes $\{\mathbf{q}_{pi}^*\}$ están contaminadas con ruido dado que las medidas originales $\{\mathbf{q}_{pi}^\#\}$ se contaminan con un ruido con propiedades Gaussianas de amplitud $\sigma=1$. El vector a estimar se compone de los parámetros intrínsecos de la cámara junto con los extrínsecos para cada una de las posiciones desde las que se han tomado imágenes $\mathbf{a}_I=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$. También se puede



4.3 Diferentes posiciones de las coordenadas de los puntos en la plantilla y en la imagen. El efecto de la corrección hace que los puntos cumplan las supuestas rectas

incluir las coordenadas de los puntos en la plantilla para que el error en la estimación sólo dependa de los errores cometidos en las medidas de las coordenadas en las imágenes. En este caso el vector de elementos estimados es $\mathbf{a}_2 = (\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$.

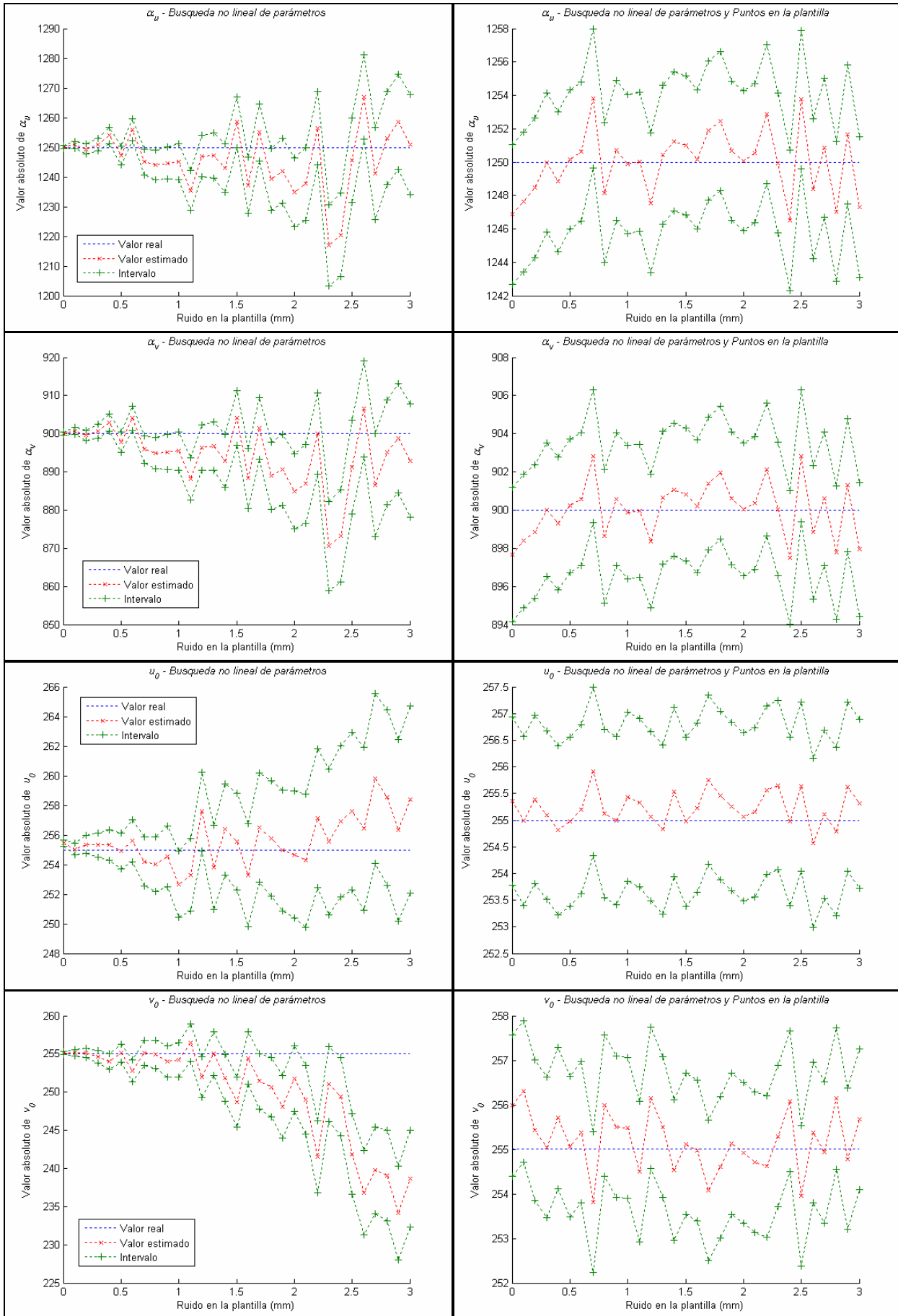
Las funciones $f^{(k)}(\mathbf{q}_{ij}, \mathbf{a}) = \mathbf{0}$ que relacionan los datos $\{\mathbf{q}_{ij}^*\}$ con el vector de parámetros \mathbf{a} son en este caso:

$$f^{(1)}(\mathbf{q}_{ij}, \mathbf{a}) = u_{ij} - \alpha_u \frac{r_{11j}x_i + r_{12j}y_i + r_{13j}z_i + t_{xj}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{zj}} - u_0 \quad (4.74)$$

$$f^{(2)}(\mathbf{q}_{ij}, \mathbf{a}) = v_{ij} - \alpha_v \frac{r_{21j}x_i + r_{22j}y_i + r_{23j}z_i + t_{yj}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{zj}} - v_0$$

El subíndice i corresponde con el número de punto en la imagen nombrada con subíndice j . Hay que tener en cuenta que en estas funciones aparecen los elementos de la matriz de rotación $r_{11j}, r_{12j}, \dots, r_{33j}$ y que se tienen como parámetros a estimar los ángulos θ_j, ϕ_j, ψ_j , que representan la rotación en cada uno de los ejes de coordenadas. Los elementos $r_{11}, r_{12}, \dots, r_{33}$ se calculan a partir de los ángulos de rotación θ_j, ϕ_j, ψ_j según (3.76).

Con estas relaciones se plantea minimizar el índice (4.45). Las derivadas parciales $\nabla_{\mathbf{a}} f^{(k)}$ se aproxima numéricamente su valor en cada punto, debido a la complejidad de la expresión dada por la relación que existe entre los ángulos θ_j, ϕ_j, ψ_j , y la matriz de rotación que forman $r_{11j}, r_{12j}, \dots, r_{33j}$. Se realiza una búsqueda no lineal de los parámetros utilizando el algoritmo de Levenberg-Madquart tomando como datos iniciales de búsqueda los obtenidos linealmente según el método descrito en el apartado 3.6. Una vez se ha obtenido el valor óptimo, se calcula la varianza de los parámetros de la cámara según la expresión (4.47).



4.4 Estimación no lineal basada en incertidumbres de los parámetros de la cámara. Se muestran las incertidumbres de los parámetros calculados. El ruido en la plantilla varía mientras en la imagen permanece constante

La figura 4.4 muestra los resultados de la experiencia. En una columna se observan los resultados de la búsqueda no lineal de los parámetros $\mathbf{a}_1=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$ y en la otra los correspondientes al vector de parámetros $\mathbf{a}_2=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$. En todos los casos se muestra el valor calculado para los elementos del vector parámetros correspondiente junto con su incertidumbre. La incertidumbre está definida por un intervalo el cual se calcula a partir del doble de la desviación típica obtenida para el parámetro. Dado que se asume una distribución Gaussiana del error tomando $2 \cdot \sigma$, se garantiza que el parámetro estará dentro del intervalo el 95% de las veces.

La lectura de las gráficas obtenidas es inmediata a partir de las expresiones que daban los errores residuales y de estimación según el algoritmo utilizado mostrados en la tabla 4.1. En primer lugar se observa que la búsqueda de los parámetros del vector $\mathbf{a}_2=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$, es más estable que el vector $\mathbf{a}_1=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$. Dado que en este caso están contaminadas con ruido tanto las coordenadas de los puntos en la imagen como las coordenadas de los puntos en la plantilla, se trata por lo tanto del último caso de la tabla 4.1. Como ambos conjuntos de coordenadas están contaminados con ruido, el vector de parámetros a estimar es el \mathbf{a}_2 y no el \mathbf{a}_1 .

En el caso de plantear el vector de parámetros \mathbf{a}_1 , el algoritmo de estimación asume que las coordenadas de los puntos en la plantilla son correctas y trata de ajustar la estimación apoyándose en estos datos inciertos. Aunque las coordenadas de los puntos de la plantilla se utilizan en el algoritmo junto con sus matrices de covarianzas, estas pesan más en el proceso de estimación y hacen que la estimación sea menos robusta. Por el contrario, si se estima el vector de parámetros \mathbf{a}_2 la estimación es más robusta ya que el algoritmo tiene en cuenta las correcciones de las coordenadas de la plantilla para realizar la estimación de los parámetros de forma más exacta.

Desde el punto de vista de los errores de estimación que se comenten, es posible predecirlos a partir de la tabla 4.1. Si se asume un número de datos de entrada n y un número de imágenes j iguales en ambos casos, según la tabla 4.1, el error de estimación dependerá en ambos casos de la incertidumbre de los datos de entrada, la cual se mide con el parámetro σ . En el caso de la estimación de \mathbf{a}_1 , esta incertidumbre es tanto la de las coordenadas de los datos de la plantilla, como la de las imágenes. En las imágenes el ruido permanece constante con $\sigma=1$ y se cambia el ruido en las coordenadas de la plantilla desde $\sigma=0$ hasta $\sigma=3$. Por lo tanto, es de esperar que el error de estimación incremente a medida que crece el ruido en las coordenadas de la plantilla. De todas formas, es necesario concretar que su valor no será predecible a partir de los datos de entrada ya que el planteamiento del índice a minimizar no es el correcto dado que

se tienen contaminadas las coordenadas de la plantilla. Es por este motivo que la estimación realizada del vector de parámetros \mathbf{a}_2 es totalmente predecible e incluso se pueden estimar los errores de estimación cometidos. En este caso, el error de estimación de los datos también depende solamente de la incertidumbre de los datos de entrada. En este caso, estos son las coordenadas de los puntos en las imágenes y las coordenadas de los puntos en la plantilla participan en la estimación como un valor inicial de búsqueda de las coordenadas reales libres de ruido. El algoritmo tiene en cuenta esta estimación y su comportamiento es acorde a lo esperado. Dado que las incertidumbres de los datos de entrada son constantes, es decir, el ruido en las coordenadas de los puntos en las imágenes es el mismo a lo largo de todo el experimento, el error de estimación permanece constante independientemente del ruido con el que se contaminan las coordenadas de la plantilla. En este caso el posible predecir el error de estimación dado que el algoritmo asume que los datos de la plantilla son inciertos. Gráficamente los errores de estimación se observan mejor en la figura 3.12 del capítulo anterior. En esta figura se representan los errores de estimación en ambos casos.

Extrapolando los resultados obtenidos de la tabla 4.1 se puede decir en el caso de tener contaminados tanto las coordenadas de la plantilla como las de las imágenes, que nunca se obtendrán los valores exactos y el error de estimación tenderá al límite dado por:

$$e_e^* = \sqrt{\left(\frac{6+6j+2n}{2nj}\right)} \cdot \sigma \quad (4.75)$$

Desde el punto de vista de las incertidumbres calculadas es necesario tener en cuenta la expresión (4.47) junto con la (4.46) para interpretar los resultados.

$$\begin{aligned} \sum \mathbf{a}^* &= \left(\sum_{i=1}^N \sum_{k,l=1}^L W_i^{(kl)\#}(\mathbf{a}^*) \cdot \nabla_{\mathbf{a}} f^{(k)T}(\hat{u}_i, \mathbf{a}^*) \nabla_{\mathbf{a}} f^{(l)}(\hat{u}_i, \mathbf{a}^*) \right)^{-1} \\ W_i^{(kl)\#}(\mathbf{a}) &= \left(\nabla_{u_i} f^{(k)\#T} \cdot \sum u_i \cdot \nabla_{u_i} f^{(l)\#} \right)_r^{-1} \end{aligned} \quad (4.76)$$

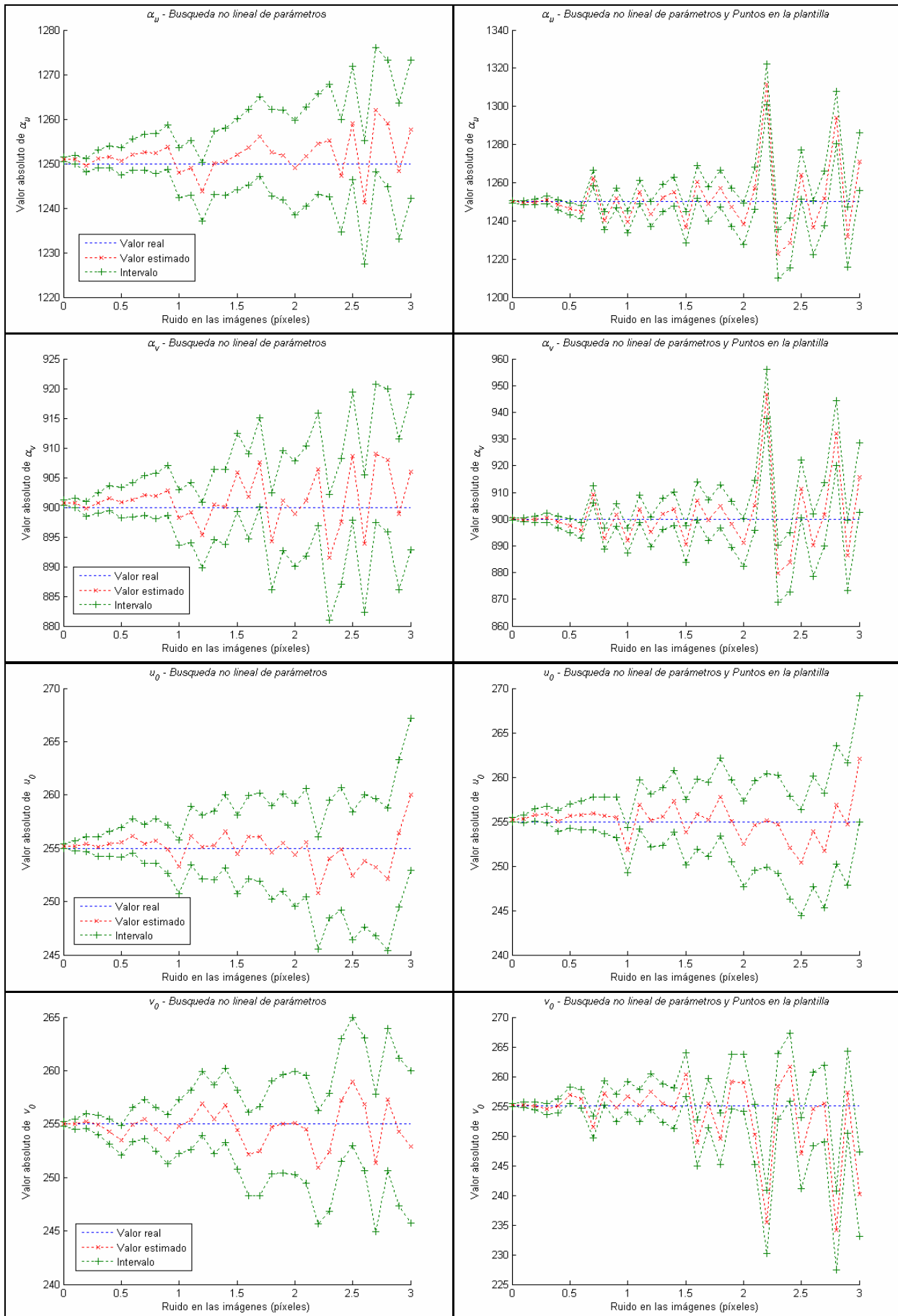
Por un lado se tiene el valor del gradiente de las funciones respecto de los parámetros buscados $\nabla_{\mathbf{a}} f^{(k)}$ para los parámetros estimados. Dado que se realiza una búsqueda no lineal, el valor de $\nabla_{\mathbf{a}} f^{(k)}$ para los parámetros estimados será siempre el mismo independientemente de la incertidumbre de los datos de partida. Decir que será el mismo significa que el algoritmo de búsqueda no lineal hará que los parámetros estimados sean tales que el valor de $\nabla_{\mathbf{a}} f^{(k)}$ en ese punto será nulo o casi nulo. Se ha comprobado este hecho y es cierto que independientemente de los errores de los datos de entrada el algoritmo consigue ajustar $\nabla_{\mathbf{a}} f^{(k)}$ a un valor cercano a cero dado que ésta es la condición para que el algoritmo finalice la búsqueda. Por otro lado se tiene el gradiente de las funciones respecto a los datos de entrada $\nabla_{\mathbf{a}} f^{(k)}$. Este término junto con la matriz de covarianzas de los datos de entrada pondera la información que aporta cada dato al algoritmo de búsqueda.

Una matriz de covarianzas asociada a un dato $\sum \mathbf{u}_i$ con una gran incertidumbre hace que éste, tenga menos peso en la búsqueda de los parámetros según se observa en la expresión (4.45). $\mathbf{W}_i^{(kl)}$ interviene como denominador en cada término del sumatorio del índice dado por (4.45), lo que hace que se pondere el peso del dato dentro del índice en función de su incertidumbre representada por $\sum \mathbf{u}_i$. Al mismo tiempo que en la estimación de los parámetros $\mathbf{W}_i^{(kl)}$ pondera con poco peso los datos con mucha incertidumbre, en el caso del cálculo de la varianza de los datos estimados aparece en el numerador y esto hace que datos con mucha incertidumbre generen incertidumbre en los datos estimados. Utilizando esta interpretación se observa que la incertidumbre de los datos estimados del vector $\mathbf{a}_2=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$ es siempre la misma dentro de unas tolerancias ya que los puntos de las imágenes que se utilizan como datos de estimación, siempre están contaminados con el mismo nivel de ruido. Por el contrario, en la estimación del vector de parámetros $\mathbf{a}_1=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$ dado que se utilizan las coordenadas de los puntos en la plantilla como datos de entrada, su incertidumbre varía y en consecuencia, la incertidumbre de los datos estimados también. Además, dado que el algoritmo planteado no representa la situación correcta del estado de los parámetros, la estimación de la incertidumbre es también muy poco robusta ya que cuando el nivel de ruido aumenta, las incertidumbres calculadas para los parámetros estimados, no corresponden a las reales ya que no incluyen el valor real. Se debe descartar pues la búsqueda del vector de parámetros $\mathbf{a}_1=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$ ya que no representa las condiciones de calibración en este caso.

Ruido en las coordenadas de los puntos de la imagen

En este caso, se realiza la misma experiencia que en el apartado anterior pero variando el nivel de ruido de las coordenadas en las imágenes. Dado que si se contaminan las coordenadas de la plantilla se trata del mismo caso del apartado anterior pero cambiando el ruido en la imagen, en este caso se opta por no contaminar las coordenadas de los puntos en la plantilla. De esta forma se espera que la estimación del vector de parámetros $\mathbf{a}_1=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$ sea más ajustada que la estimación de los parámetros junto con las coordenadas de los puntos en la plantilla representadas por el vector de parámetros $\mathbf{a}_2=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$.

Efectivamente en la figura 4.5 se muestran los resultados. En este caso la columna de la izquierda que muestra las estimaciones de los parámetros basados en el vector \mathbf{a}_1 , muestra unos resultados más estables que en el caso del vector \mathbf{a}_2 representados en la columna de la derecha.

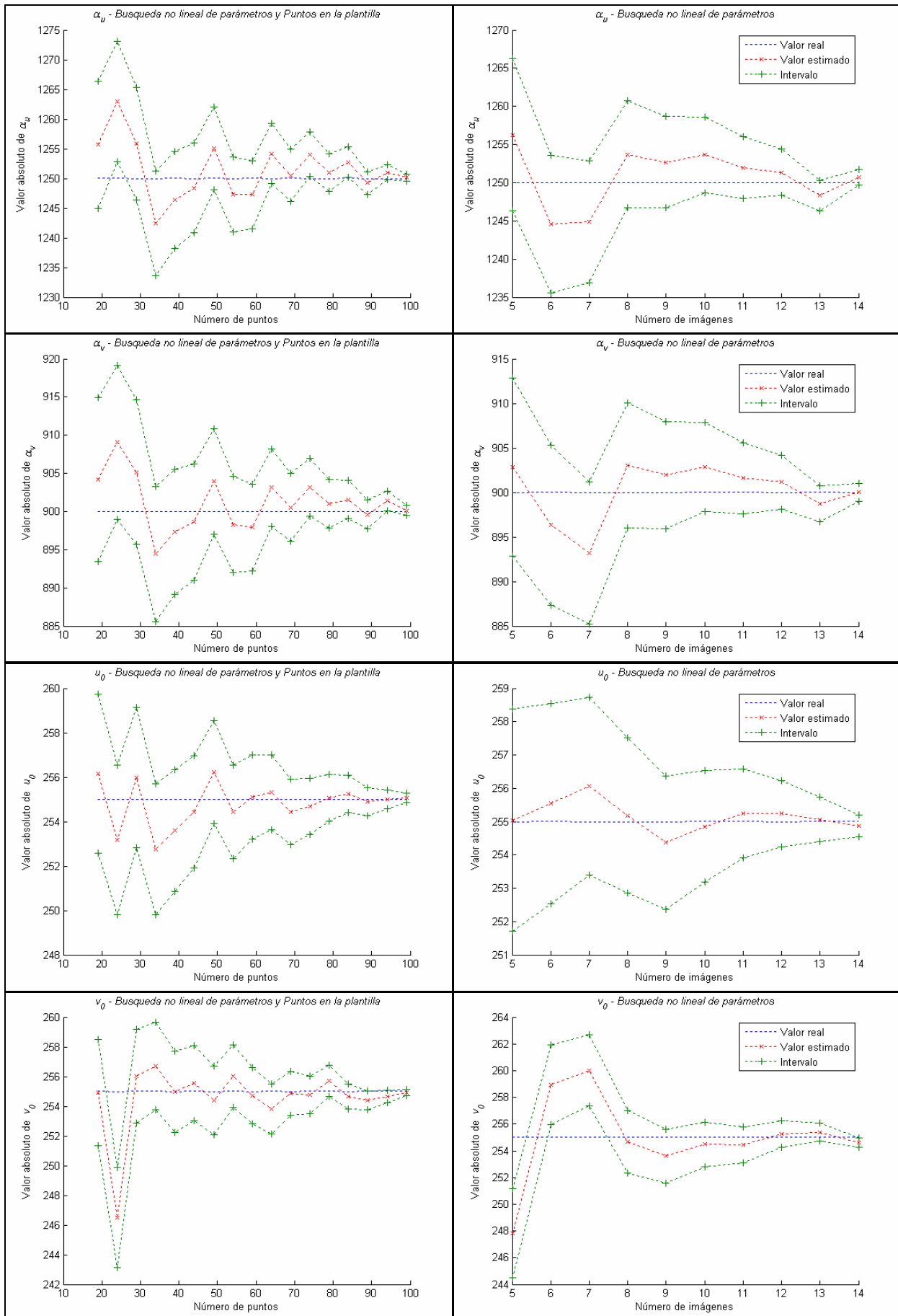


4.5 Estimación no lineal basada en incertidumbres de los parámetros de la cámara. Se muestran las incertidumbres de los parámetros calculados. El ruido en la plantilla es nulo mientras en la imagen varía

Dado que las coordenadas de los puntos en la plantilla son una referencia exacta para la búsqueda de los parámetros, el algoritmo de búsqueda consigue resultados más estables y más aproximados a los valores reales. Dentro del algoritmo de búsqueda las coordenadas de los puntos de la plantilla participan con mucho peso ya que su covarianza se supone nula. Esto hace que la información que aportan tenga mucho peso en la búsqueda de los parámetros. Por el contrario, si se incluyen las coordenadas de los puntos en la plantilla a pesar de ser exactas, el algoritmo de búsqueda desvirtúa esta información y obtiene resultados menos fiables. Incluso en el cálculo de las incertidumbres de los parámetros estimados, se cometen errores ya que en muchos de los casos el valor calculado no corresponde con el verdadero. Desde el punto de vista del error de estimación que se comete, según la tabla 4.1, en ambos casos el error aumenta con la incertidumbre de los datos de entrada. Dado que en este caso se varía el error de las coordenadas de las imágenes, el error de estimación aumentará según se aprecia en la figura 3.11. En esta figura se muestran los errores de estimación que se cometen cuando el ruido en la imagen aumenta y no se tiene ruido en las coordenadas de la plantilla. Como es de esperar, el error de estimación es menor si se calcula el vector \mathbf{a}_1 , ya que representa mejor las condiciones de calibrado. En este caso, plantear la estimación del vector \mathbf{a}_2 resulta absurda ya que al no representar las condiciones de calibrado, se comete un error superior en la estimación de los parámetros al obtenido en el caso del vector \mathbf{a}_1 .

Efecto del número de puntos y del número de imágenes

En este caso se comparan los efectos que producen el número de puntos y el número de imágenes en la estimación de los parámetros. Se realiza una simulación cambiando el número de puntos que tiene la plantilla. Se contaminan las coordenadas de los puntos en la imagen y en la plantilla con un ruido de $\sigma=0.5$. El número de imágenes que se toman son ocho al igual que en las experiencias simuladas anteriormente. La columna izquierda de la figura 4.6 muestra los resultados. Se realiza la estimación del vector de parámetros $\mathbf{a}_2=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$ dado que las coordenadas de los puntos en la plantilla están contaminadas con ruido. Efectivamente contra más datos participan en la estimación, menores son los errores que se cometen en el cálculo de los parámetros. En el caso de tener pocos puntos en la estimación, los resultados son menos estables y esto provoca que el cálculo de la incertidumbre no incluya el valor real. Desde un punto de vista estadístico, se asume que el error que contamina las coordenadas de los puntos tiene propiedades Gaussianas. Esto se garantiza si el número de puntos es grande. En el caso de tener pocos puntos ocurre que las propiedades Gaussianas del ruido se desvirtúan y en consecuencia la estimación no acaba siendo todo lo robusta que se espera.

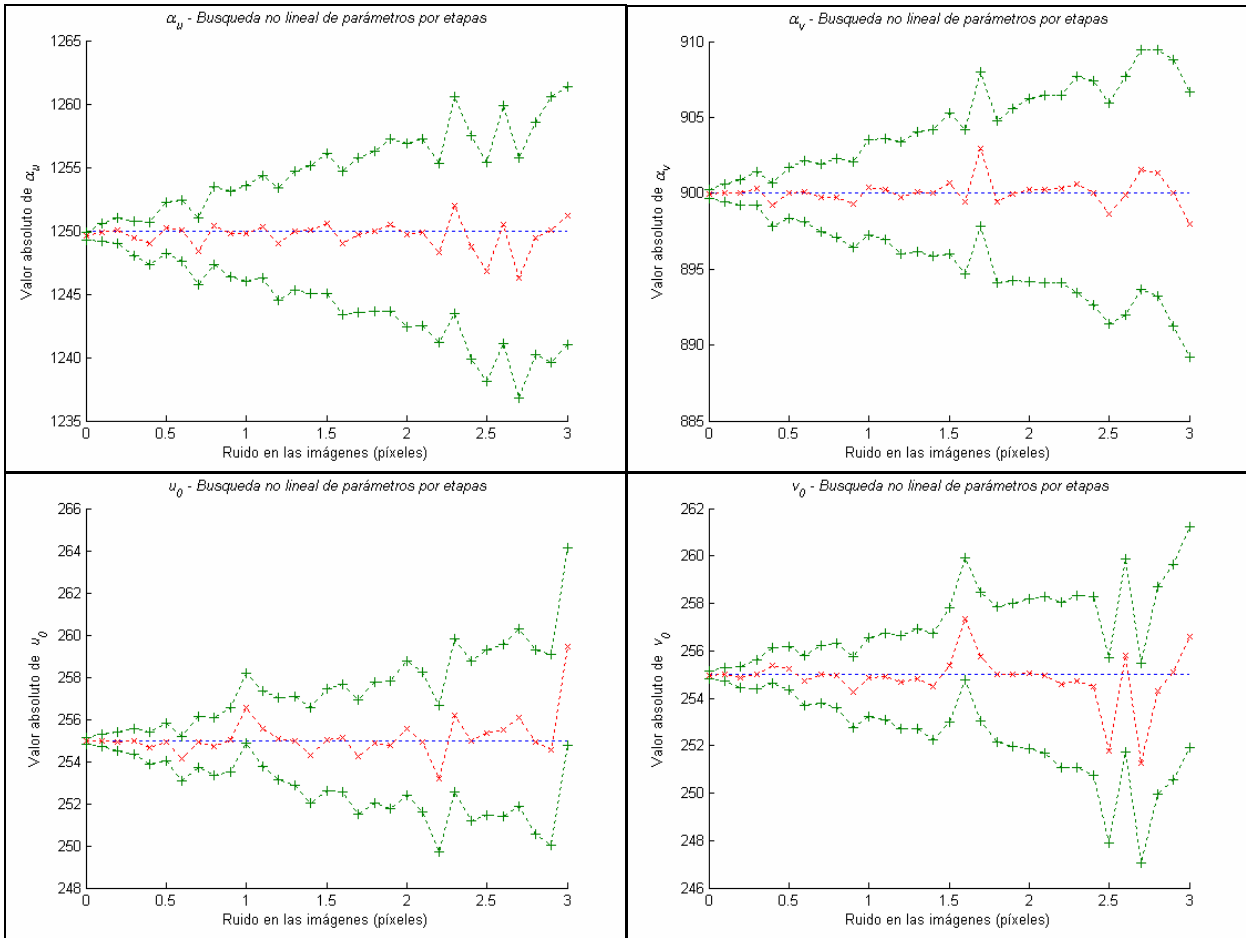


4.6 Efectos del número de puntos y número de imágenes en el cálculo de los parámetros. Si se varía el número de puntos el ruido en la plantilla y en las imágenes es $\sigma=0.5$. Variando en número de imágenes no hay ruido en la plantilla y mientras en la imagen es $\sigma=1$

En segundo lugar se cambian el número de imágenes que se toman de la plantilla. En este caso los puntos que se utilizan son los 99 de la plantilla y el número de imágenes varía desde 3 que son las mínimas necesarias hasta 14. En la columna derecha de la figura 4.6, sólo se muestra a partir de 5 imágenes para poder apreciar mejor los resultados. Con un número menor de imágenes, los resultados son extrapolables. Las posiciones desde donde se toman las imágenes son $\mathbf{t}_1=(150, 200, 580)$, $\mathbf{t}_2=(-50, 250, 880)$, $\mathbf{t}_3=(100, -20, 820)$, $\mathbf{t}_4=(-40, -150, 780)$, $\mathbf{t}_5=(-150, -150, 530)$, $\mathbf{t}_6=(-100, 125, 400)$, $\mathbf{t}_7=(140, -150, 500)$, $\mathbf{t}_8=(240, 50, 600)$, $\mathbf{t}_9=(50, -250, 880)$, $\mathbf{t}_{10}=(-240, -50, 600)$, $\mathbf{t}_{11}=(-140, 150, 600)$, $\mathbf{t}_{12}=(125, 20, 800)$, $\mathbf{t}_{13}=(25, -100, 700)$, $\mathbf{t}_{14}=(-25, 100, 600)$. En todos los casos se orienta la cámara para que se el eje óptico coincida con el centro de la imagen. Igual que antes, contra más imágenes se toman menores errores se cometen. En este caso las coordenadas de los puntos en la plantilla no se contaminan con ruido. En consecuencia se plantea la estimación del vector de parámetros $\mathbf{a}_I=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$. El hecho de no contaminar las coordenadas de los puntos de la plantilla con ruido es para comprobar que el número de imágenes que se toman de la plantilla no afecta a la estimación de los parámetros. Esta regla se definió en la tabla 4.1 en la que se ve que cuando se estiman los parámetros de la cámara basándose en el modelo y no se tienen contaminados los puntos de la plantilla, el número de imágenes utilizadas no afecta ni en el error de estimación ni tampoco en el residual. Este efecto se aprecia mejor en la figura 3.14. En esta figura se observa que a partir de 8 imágenes, el error de estimación que se comete es constante en el caso de estimar el vector de parámetros \mathbf{a}_I . Esto indica que efectivamente el número de imágenes que participan en el proceso de calibración no afectan al éxito de los resultados. Se puede obtener como conclusión que al menos es necesario 8 imágenes para garantizar un ruido con características Gaussianas. A partir de las simulaciones se desprende que con menos de 8 imágenes los resultados no son muy robustos, incluso fuera de los esperados. Se asume que el ruido que contamina las coordenadas en el caso de disponer de pocas imágenes no tiene características Gaussianas o también que las posiciones desde las que se toman las imágenes no garantizan que los resultados sean los esperados. En el próximo capítulo se tratan los efectos de las posiciones desde las que se toman las imágenes y como afectan éstas al proceso de calibración.

Cálculo no lineal de parámetros en cada una de las etapas del proceso de calibrado

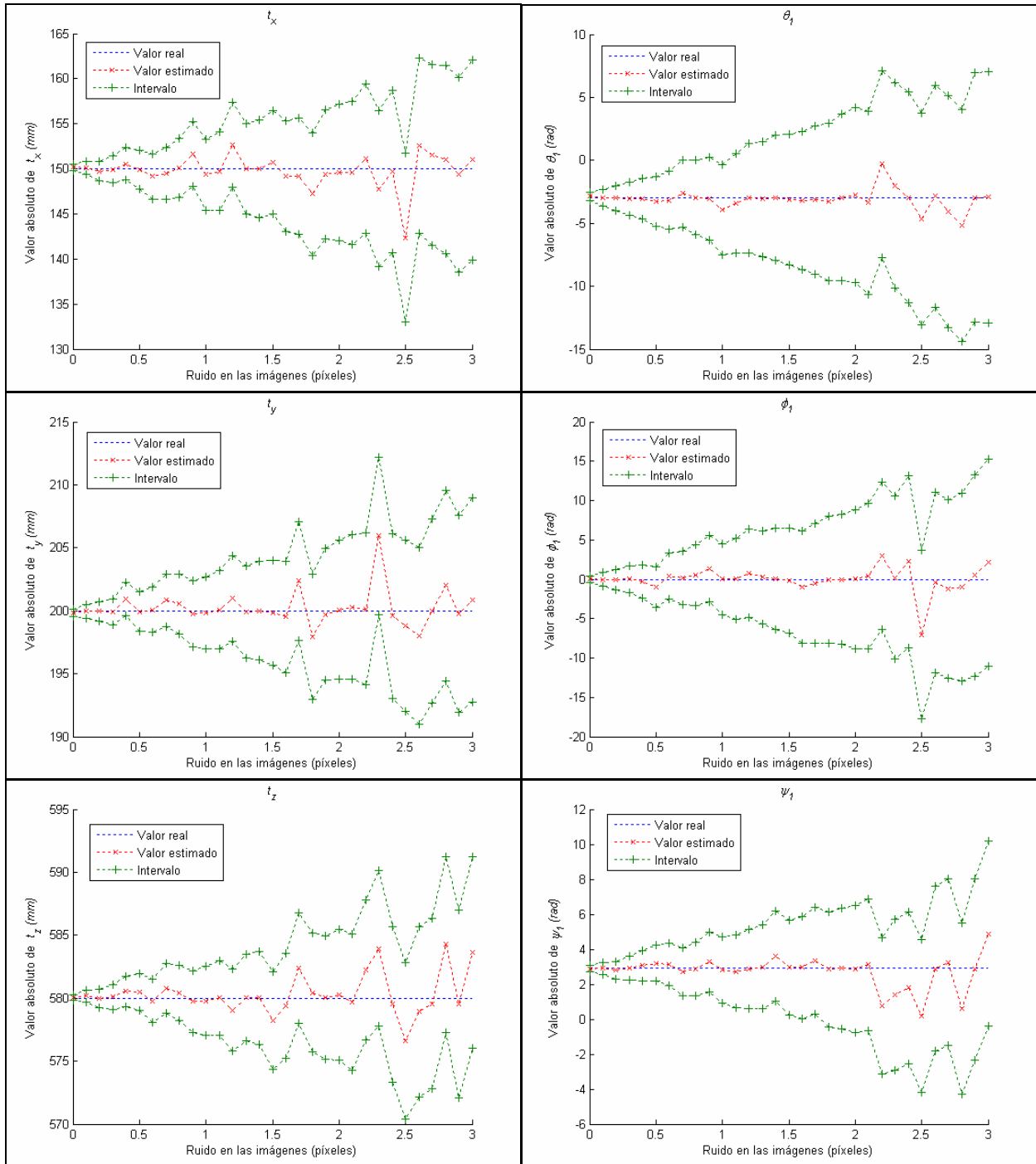
En este caso se plantea la estimación de los parámetros de la cámara por etapas. El procedimiento es el mismo que la denominada búsqueda lineal de parámetros pero en este caso se ajustan los parámetros estimados en cada etapa con una búsqueda no lineal teniendo en cuenta la incertidumbre de los datos de entrada para esa etapa. Los parámetros estimados en cada etapa se calculan junto con una matriz de covarianzas que determina la incertidumbre de los mismos. Estos mismos parámetros junto con sus incertidumbres se utilizan como datos de



4.7 Estimación no lineal basada en incertidumbres por etapas de los parámetros de la cámara. Se muestran las incertidumbres de los parámetros calculados. El ruido en la plantilla es nulo mientras en la imagen varía

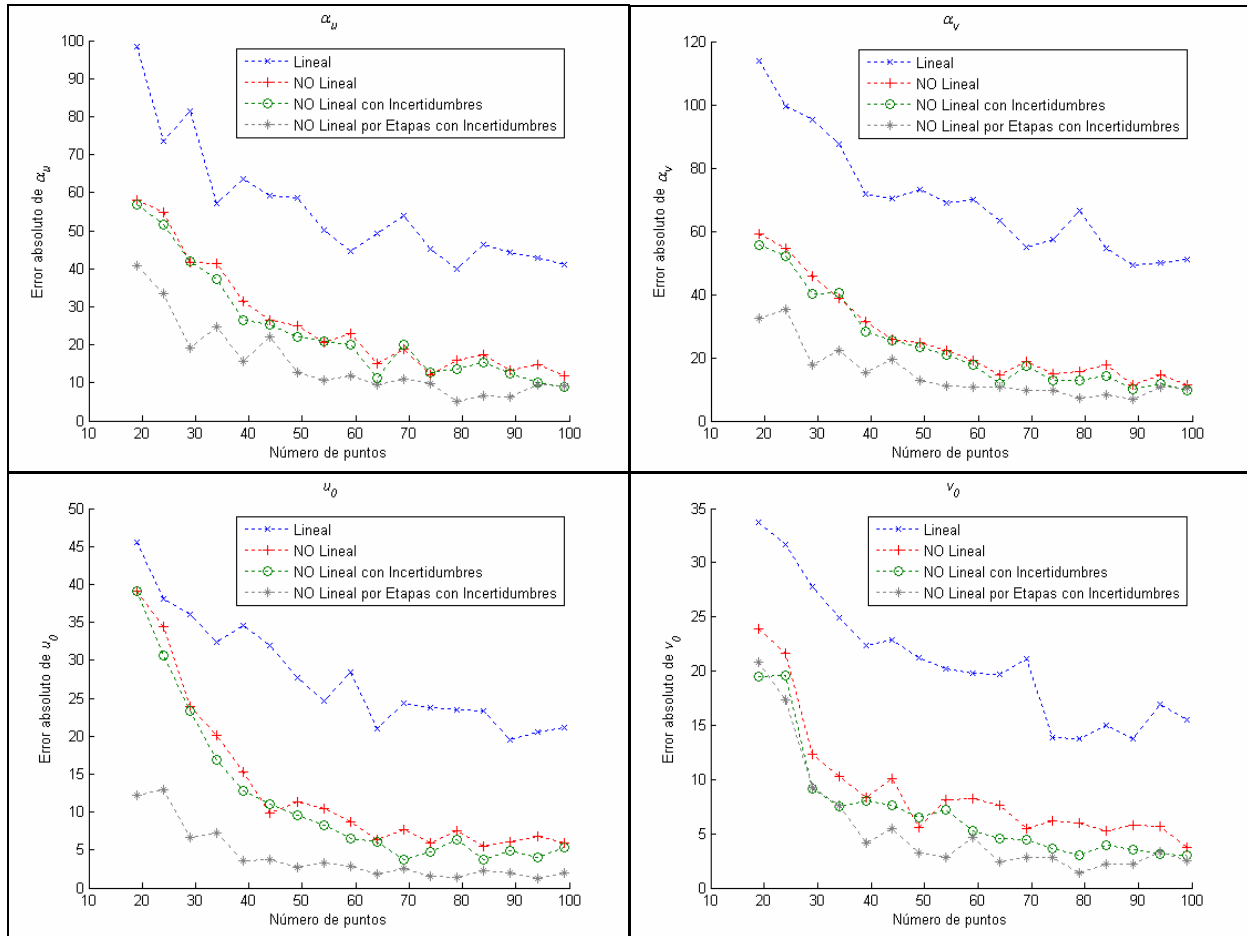
entrada para la siguiente etapa. La forma de llevar a cabo cada etapa se describe en los apartados 4.4.1, 4.4.2, 4.4.3, 4.4.4, 4.4.5. Los resultados obtenidos se muestran en la figura 4.7. En este caso se varía el nivel de ruido de las coordenadas de las imágenes desde un valor nulo a $\sigma=3$. Las coordenadas de los puntos en la plantilla no se contaminan con ruido para poder comparar los resultados con los procedimientos anteriores contaminando con ruido los puntos de las imágenes. A partir de los resultados obtenidos anteriormente, en este caso se plantea el índice en la etapa de estimación de homografías teniendo en cuenta que los puntos en la plantilla no están contaminados con ruido.

En este caso y de igual forma se que se hace en todos los experimentos anteriores las gráficas muestran la media de los resultados obtenidos para 100 simulaciones del proceso de calibración. Se observa que los resultados se acercan más a los valores reales de los parámetros que las mostradas en la figura 4.5. Debido a que en cada etapa se utilizan unos valores de los parámetros más cercanos a los reales junto con sus incertidumbres, el algoritmo de búsqueda utiliza la información que cada dato aporta de forma correcta dándole el peso que se merece según la incertidumbre del mismo. Además, debido a que se plantea un índice con significado puramente algebraico los posibles acoplamientos que puedan presentarse entre los parámetros de



4.8 Estimación de los parámetros extrínsecos para la primera posición de la cámara. Se muestran las incertidumbres. El ruido en la plantilla es nulo mientras en la imagen varía

la cámara desaparecen y al final del proceso se obtienen resultados más fiables. Se observa en la figura 4.5 que los errores que se cometen en la estimación de α_u y α_v están correlados en cierta forma. Lo mismo ocurre con los parámetros u_0 y v_0 . En los resultados de la figura 4.7 esta correlación no se aprecia. Por el contrario, se corre el riesgo que en algunos casos debido al significado algebraico del índice planteado, el algoritmo de búsqueda se desvíe hacia una



4.9 Errores de estimación de los parámetros intrínsecos con los diferentes algoritmos. El ruido en la plantilla es nulo mientras en la imagen vale $\sigma=1$. Se cambia el número de puntos

solución con poco sentido geométrico pero que minimiza el índice algebraico. Este hecho no ha ocurrido en las simulaciones planteadas.

Cálculo basado en incertidumbres de los parámetros extrínsecos de la cámara

Hasta el momento sólo se han calculado los parámetros intrínsecos de la cámara. El método de calibración basado en plantilla plana calcula los parámetros intrínsecos de la cámara, y a partir de los cuales se obtienen los parámetros extrínsecos. Dado que se tiene múltiples posiciones de la cámara desde donde se toman todas las imágenes, existe un conjunto de parámetros para cada una de ellas. Dado que los parámetros intrínsecos se obtienen junto con una incertidumbre, los parámetros extrínsecos también se calculan teniendo en cuenta esa incertidumbre. Como resultado se puede obtener los intervalos en los cuales se localizaba y orientaba la cámara cuando se toma la imagen de la plantilla. En la figura 4.8 se muestran los valores de obtenidos junto con sus incertidumbres.



4.10 Escenario de calibración

Comparación de algoritmos por errores de estimación

Una vez puestos a prueba los diferentes procedimientos de cálculo de los parámetros de la cámara es hora de realizar comparaciones basadas en algún tipo de referencia para decidir cual de ellos es más eficiente. En la figura 4.9 se muestran los errores de estimación que se obtienen con los diferentes algoritmos. Dado que se conocen los valores reales de los parámetros de la cámara se obtiene un error de estimación con los parámetros calculados con cada uno de los métodos. Se comparan los tres métodos definidos por búsqueda lineal de parámetros, búsqueda no lineal de parámetros y cálculo de parámetros realizando una búsqueda no lineal por etapas.

La búsqueda lineal de parámetros se realiza para tener un punto de partida para la búsqueda no lineal. La búsqueda no lineal por etapas presenta la mejor estimación ya que el proceso de cálculo está más cuidado. Sin embargo, desde el punto de vista del tiempo de cómputo, éste es el algoritmo que necesita mayor tiempo para llegar a una solución. La búsqueda no lineal de parámetros es muy costosa desde el punto de vista computacional y con este procedimiento se realiza varias veces.

Resultados de calibración con cámaras reales

Al igual que se hizo en el capítulo anterior se han calibrado las mismas cámaras utilizando el proceso de calibración basado en incertidumbres descrito en éste capítulo. Una de ellas es de gama baja, la logitech Quick Cam Pro, y la otra la Canon Powershot A640. En ambos casos se toman 8 imágenes de la plantilla entendiendo que son suficientes a partir de las simulaciones realizadas. Las resoluciones han sido en ambos casos de 640x480 píxeles. Igual que en el capítulo anterior, la plantilla está compuesta de un tablero de ajedrez de 15x11 esquinas de dimensiones 280x200 mm. Se asume que está colocado de forma que el plano de la plantilla corresponde con el plano $z=0$ de la escena. En la figura 3.15 se muestra una de las imágenes de

Logitech Quick Cam Pro

	Lineal	No Lineal	No Lineal basado en Incertidumbres	No Lineal por etapas basado en Incertidumbres
α_u	932.2	932.1	930.4±35.2	928.6±35.2
α_v	942.2	942.5	940.1±42.5	943.5±27.5
u_0	318.5	323.5	318.5±5.7	321.5±4.8
v_0	237.9	242.8	245.9±9.1	241.8±7.8
φ	1.45	1.11	-	-
k_1	2.26	6.2	-	-
k_2	0.25	0.61	-	-
e	0.345	0.245	0.253	0.298
t	0.002	1.453	2.743	12.565

Canon PowerShot A640

	Lineal	No Lineal	No Lineal basado en Incertidumbres	No Lineal por etapas basado en Incertidumbres
α_u	838.5	831.9	829.5±22.5	932.4±18.9
α_v	855.2	837.8	839.9±19.8	840.2±17.8
u_0	312.5	317.2	318.4±4.5	320.4±4.8
v_0	246.3	239.8	238.5±5.2	240.8±5.7
φ	0.98	0.15	-	-
k_1	0.551	0.646	-	-
k_2	0.05	0.06	-	-
e	0.58	0.456	0.489	0.504
t	0.002	1.657	5.635	25.451

4.2 Comparación de resultados de calibración de dos cámaras reales con diferentes métodos

la plantilla correspondientes a cada cámara. Las posiciones desde donde se han tomado las imágenes se ha tratado que sean parecidas a la de la cámara simulada.

En este caso, antes de realizar la estimación de los parámetros de la cámara, se corrigen las coordenadas de los puntos de forma que cumplan supuestamente las imágenes reales de las rectas que forman los puntos de la plantilla. Con este procedimiento se obtienen unas medidas más exactas de los puntos en las imágenes, además de conseguir sus incertidumbres expresadas en una matriz de covarianzas. En el caso de las coordenadas de los puntos en la plantilla se asume que no están contaminadas con ruido. Para conseguir unas coordenadas libres de ruido en la plantilla, ésta se diseña con un programa de dibujo en el que se colocan las esquinas de los puntos de interés a una distancia de 20mm. Posteriormente se utiliza una impresora láser para obtener una máxima resolución en la generación de la plantilla. Además, para garantizar que efectivamente la plantilla es plana, ésta se ha pegado en un espejo. Normalmente, los espejos suelen garantizar una superficie plana ya que sino, los objetos que se reflejan en el mismo se verían deformados. De esta forma se puede tener una plantilla plana en la cual se asumen que las medidas de las coordenadas no están contaminadas con ruido. El origen de coordenadas del mundo se sitúa en una de las esquinas de la plantilla. La figura 4.10 muestra el escenario de calibración.

A partir de los puntos corregidos de las imágenes y las coordenadas de los puntos de la plantilla se calculan los parámetros de la cámara basándose en las incertidumbres de los datos de entrada. Los parámetros que se obtienen están sujetos a incertidumbres las cuales se expresan mediante una matriz de covarianzas. En la tabla 4.2 se muestran los resultados. Las

incertidumbres de los parámetros se expresan con un intervalo obtenido a partir de 2σ . Desde un punto de vista estadístico el valor real del parámetro estará contenido dentro del intervalo aproximadamente el 95% de las veces. Se calculan los parámetros con dos procedimientos. El primero consiste en obtener unos valores aproximados de los parámetros de forma lineal y después utilizarlos como punto de partida en la búsqueda no lineal. El segundo consiste en realizar todas las etapas del proceso de calibración utilizando plantilla plana realizando una búsqueda no lineal de los parámetros implicados en cada una de ellas. Los resultados obtenidos en el segundo caso son mejores a cambio de un mayor coste computacional. En los dos casos se asume que las coordenadas de los puntos en la plantilla están libres de ruido. También se muestran los errores residuales obtenidos en cada caso, entendiendo el error geométrico que existe entre el conjunto de puntos medidos en cada imagen $\{q_{ii}^{\#}\}$ y los que se obtienen de forma óptima a partir de los parámetros estimados $\{q_{pi}^0\}$. También se muestra el tiempo de cómputo en segundos a partir de la corrección de las coordenadas de los puntos en la imagen. El calibrado se ha realizado con un procesador AMD Athlon 64 con una velocidad de reloj de 3200 Mhz.

Se observa que en el caso de realizar una búsqueda no lineal de los parámetros en cada etapa, el error residual resultado es mayor que en el caso de realizar la búsqueda no lineal directa de los parámetros de la cámara. Esto es debido a que en el caso de la búsqueda no lineal de los parámetros a partir del modelo de la misma, el índice que se plantea sirve para minimizar directamente este error residual. En el caso de realizar una búsqueda no lineal de parámetros en cada etapa del proceso de calibración, el índice minimiza un error entre parámetros los cuales tienen un significado meramente algebraico. Al final se obtiene un error residual mayor entre los puntos $\{q_{ii}^{\#}\}$ y $\{q_{pi}^0\}$ sin embargo, esto no significa que la estimación sea menos parecida a los valores reales. Como se explica en el apartado 4.5 un error geométrico menor no significa unos parámetros estimados más parecidos a los reales. Desde el punto de vista del tiempo de cómputo, el procedimiento que mejor resultados obtiene acaba siendo el más costoso de tiempo.

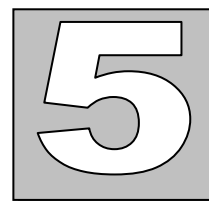
4.7. Conclusiones

El uso de herramientas estadísticas para resolver el procedimiento de calibración de una cámara permite obtener mejores resultados además de caracterizar el método. Por un lado se obtienen mejores resultados ya que el algoritmo tiene en cuenta la información de acuerdo a la incertidumbre que ella misma presenta y además previamente se corrige la información de partida para eliminar dentro de lo posible los errores de medida. Las medidas de las coordenadas se corrigen de forma que cumplan la imagen de la recta y esto permite definir una incertidumbre de las coordenadas de los puntos si no lo hacen. Después, las medidas de las coordenadas se utilizan en la estimación de los parámetros de la cámara de acuerdo a su incertidumbre. Como

consecuencia se obtienen un conjunto de parámetros más cercanos a los reales además de una incertidumbre de los mismos. Se han testeado dos formas de obtener los parámetros teniendo en cuenta la incertidumbre de los datos. Uno de ellos consiste en la búsqueda no lineal basada en incertidumbres de los parámetros de la cámara basada en el modelo de la misma. Como punto inicial de partida se toman los resultados obtenidos de forma lineal. El segundo método consiste en resolver el procedimiento de calibración basado en plantilla plana de Zhang [103], realizando una búsqueda no lineal basada en incertidumbres en cada una de las etapas. Como resultado se obtiene la incertidumbre de los parámetros para que sirva de dato de entrada en la siguiente etapa. Las gráficas obtenidas en los resultados experimentales demuestran que el segundo método obtiene mejores resultados así como menos incertidumbre en los parámetros. Dado que se plantea un índice algebraico sin significado físico, los acoplamientos que puedan existir en los parámetros de la cámara desaparecen y la estimación es más cercana a la real. Por el contrario, el tiempo de cómputo necesario para resolver este proceso de calibración es muchísimo más largo. Esta incertidumbre permite utilizar posteriormente los resultados de la calibración de forma adecuada en caso de realizar reconstrucción 3D o localización de robots con visión.

Por otro lado es posible caracterizar el procedimiento de calibración desde el punto de vista de conceptos tales como el límite de máximo de optimalidad que se puede alcanzar en la estimación a partir de una incertidumbre en las medidas de las coordenadas de los puntos ó número de puntos o imágenes a utilizar en el proceso. Este límite ha sido definido para cada una de las etapas del proceso de calibración. A partir de ahí es posible comparar algoritmos y decidir cual funciona de forma más eficiente. Se llega a la conclusión de que con ruido en las medidas de las coordenadas de la plantilla e imagen es imposible llegar a la solución exacta de los parámetros, sin embargo, si sólo se tiene ruido en las coordenadas de las imágenes, con un número de puntos considerable en la plantilla si que sería posible. Del mismo modo, dado que el límite de optimalidad depende tanto del número de puntos en la plantilla como del número de imágenes tomadas de la misma, se han definido valores en torno a los cuales los resultados están garantizados dentro de un nivel de incertidumbre. Como conclusión teórica interesante decir que el número de imágenes tomadas de la plantilla no influye en la mejora de los resultados de calibración si las coordenadas de los puntos de la plantilla no están contaminadas con ruido. Esta conclusión a nivel práctico se traduce en que por lo menos 8 imágenes son necesarias para que esto sea cierto.

En definitiva, se descarta la idea de que una mejor estimación lleva asociada un menor error residual, ya que como se comprueba a lo largo del capítulo lo que parecía inicialmente una ventaja que era incluir los puntos en la plantilla dentro del índice de calibración resulta interesante si el nivel de ruido de éstas es considerable. En caso contrario es mejor apostar por un diseño exacto de la plantilla que permita tener las coordenadas con un bajo nivel de ruido y realizar el proceso de calibración sin incluir el conjunto de puntos de la plantilla en el índice.



CONDICIONES ÓPTIMAS

En este capítulo se pretende definir unas condiciones bajo las cuales el proceso de calibrado de la cámara se realiza desde algún punto de vista de forma óptima. Resulta complicado definir un procedimiento que asegure la optimalidad de la calibración en todas las condiciones, si embargo, si que se puede definir una metodología a partir de la cual realizar la calibración de cámara de forma óptima bajo unas condiciones muy concretas. Algunos autores definen diferentes criterios que permiten calibrar una parte del modelo de la cámara de forma óptima. Un ejemplo de estos trabajos son los de Kanatani [47] que define unas condiciones óptimas para calcular la distancia focal o Luxen [60] que obtiene la rotación de la cámara. También se pueden plantear otros puntos de vista a tener en cuenta los cuales sirvan para condicionar el proceso de calibración. Estos pueden ser por ejemplo la relación señal/ruido de los datos de entrada planteada por Hartley [32], o desacoplamientos entre parámetros a estimar. Igualmente se pueden tratar condiciones numéricas que permitan la resolución matemática del algoritmo de calibración de forma estable, o del mismo modo condiciones bajo las cuales la incertidumbre de los parámetros estimados sea mínima. En este capítulo, se optimiza el proceso de calibración de la cámara a partir de las posiciones en las que se toman las imágenes. Las posiciones se calculan a partir de restricciones que permiten obtener una buena condición en el proceso de cálculo.

5.1. ¿Qué se entiende como condiciones óptimas?

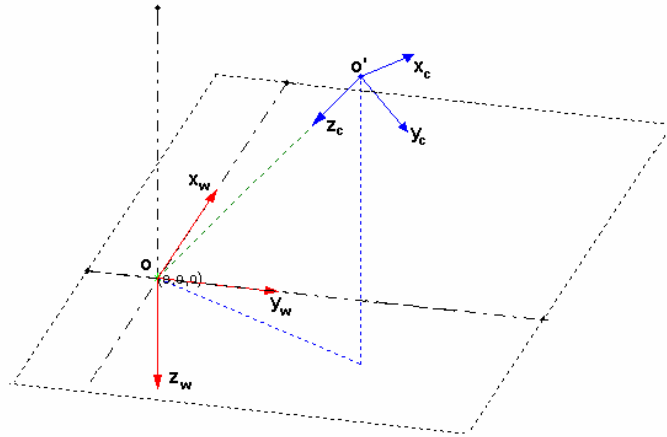
Hasta el momento se ha estudiado y caracterizado desde diferentes puntos de vista el método de calibración de cámaras basado en plantilla plana. Este método permite obtener los parámetros de la cámara a partir de varias imágenes de la plantilla tomadas desde diferentes posiciones y orientaciones. A priori parece fácil calibrar una cámara con este método. Sin

embargo, en el momento de plantearse el procedimiento de calibración siempre surgen cuestiones del tipo, cual es el tamaño de la plantilla de calibración a utilizar o cuantas y desde donde se tienen que tomar las imágenes. Estas preguntas se han resuelto de una forma cualitativa a partir de la caracterización del procedimiento tanto teórico como práctico realizado en el capítulo anterior. Las pautas obtenidas hasta el momento definen unos valores aproximados basados en unas condiciones teóricas de calibración que casi nunca se pueden alcanzar pero que sirven de alguna forma para calibrar la cámara bajo ciertas condiciones de optimalidad. Éstas deciden el número de imágenes a tomar o el número de puntos a utilizar en la plantilla pero en ningún momento definen desde que posiciones y orientaciones es más aconsejable tomar las imágenes o si los puntos de interés definidos en la plantilla colocada de alguna forma en concreto mejoran el proceso de calibración. Hasta ahora se ha hablado de un tablero de ajedrez pero quizás existan otras formas de definir los puntos en la plantilla que permitan detectar las posiciones de los mismos en las imágenes con mayor exactitud, o si estos se sitúan de alguna forma determinada, condicionen mejor el cálculo del algoritmo de calibración.

Los resultados del capítulo anterior hacen pensar sobre la posibilidad de condicionar mejor el proceso de calibrado de la cámara a partir de las posiciones desde las que se toman las imágenes. En el caso de calibrar la cámara teniendo las coordenadas de los puntos en la plantilla libres de ruido, se espera que el número de imágenes no afecte a la calidad de los resultados. Sin embargo, este hecho se produce a partir de un número de imágenes determinadas. Quizás es que las posiciones desde las que se toman las imágenes no condicionan bien el sistema y esto provoca que las propiedades estadísticas de ruido Gaussiano se desvirtúen.

Para obtener las posiciones óptimas de la cámara se pueden tener en cuenta dos criterios. Uno de ellos es meramente estadístico. En este caso se buscan unas posiciones que permitan minimizar los errores que se producen cuando se calcula la posición de un punto en la imagen a partir de sus coordenadas en la plantilla, la posición y orientación de la cámara y sus características. De esta forma, dado que los errores en las coordenadas de la plantilla afectan poco al cálculo de las coordenadas en la imagen, el proceso de búsqueda no lineal acabará con unos parámetros con una varianza mínima que se ajusten al conjunto de puntos utilizados. Para ello se tiene en cuenta la expresión que calcula la varianza de los parámetros estimados (4.47).

Otra forma de poder obtener las posiciones óptimas desde las que tomar las imágenes para realizar la calibración es condicionando bien el proceso de calibración. Para ello se utiliza el número de condición de la matriz que sirve para calcular los parámetros intrínsecos de la cámara usando una plantilla plana (3.53). Esta matriz se compone a partir de las homografías que relacionan las imágenes con la plantilla plana que a su vez se componen de las posiciones desde las que se toman las imágenes. Por lo tanto se pueden obtener las posiciones desde las que tomar las imágenes que permitan condicionar de forma óptima esta matriz. A continuación se detalla este segundo planteamiento.

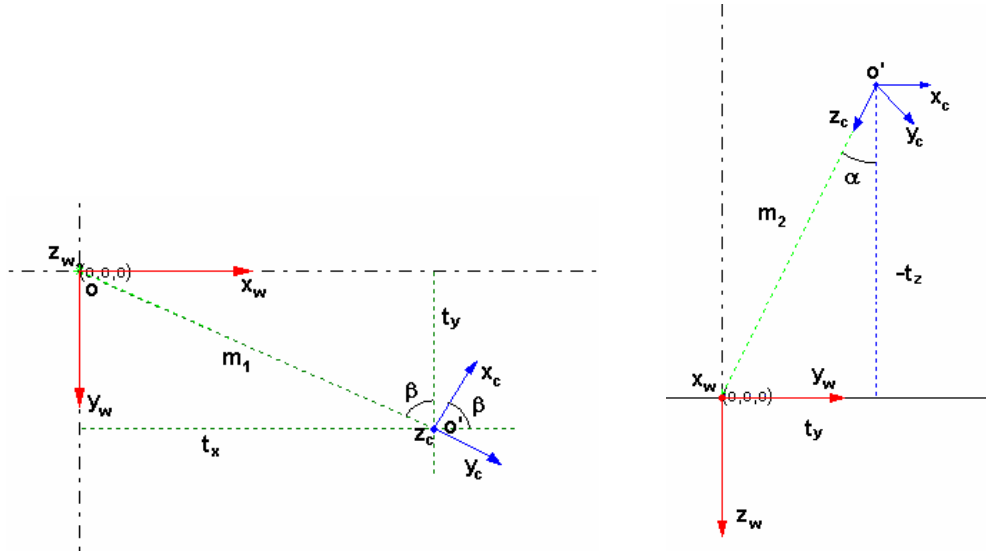


5.1 Escenario simulado de calibración para definir las condiciones óptimas

5.2. Condiciones óptimas para la calibración

Para obtener las posiciones óptimas desde las que se toman las imágenes de la plantilla plana se va a definir primero el escenario de calibración. Sin pérdida de generalidad se establece el origen de coordenadas 3D para el escenario $(o, \{x_w, y_w, z_w\})$ y su orientación tal y como se muestra en la figura 5.1. De forma genérica se establece el origen de coordenadas del escenario en el centro de la plantilla. También se fija el origen de coordenadas 3D para la cámara $(o', \{x_c, y_c, z_c\})$ con traslación negativa en el eje Z respecto al sistema de coordenadas del mundo. La cámara se supone siempre orientada de forma que el eje óptico atraviesa el centro de la plantilla, es decir, el origen de coordenadas del escenario. La figura 5.1 muestra esta situación. Esta forma de orientar la cámara permite definir la rotación de la misma en función de su posición en el escenario. Además, permite realizar la deducción sin llegar a situaciones absurdas en las que la cámara quedara orientada de una forma que fuera imposible visualizar la plantilla. A partir de aquí, se asume que el eje X_c de coordenadas de la cámara permanece paralelo al plano X_w-Y_w de la escena. Desde un punto de vista práctico, este escenario de calibración representa la plantilla plana colocada en el plano del suelo y la cámara se coloca sobre un soporte en el cual se mantiene el borde superior de la imagen paralelo al plano del suelo. El origen de medidas del escenario se establece en el centro de la plantilla.

Inicialmente se supone el sistema de coordenadas de la cámara en la misma orientación que el sistema de referencia de la escena de calibración. A partir de aquí, se rota la cámara primero un ángulo β respecto al eje Z_c . Esta situación se representa en la parte derecha de la figura 5.2. Seguidamente, respecto al eje X_c que ha quedado girado, se rota la cámara un ángulo α . Este giro está representado en la parte izquierda de la figura 5.2. Estos dos giros hacen que el eje X_c permanezca paralelo al plano $X-Y$ de escenario de calibración. Los ángulos de rotación



5.2 Ángulos de rotación de la cámara expresados a partir de la posición de la misma, de forma que el eje óptico pasa por el origen de coordenadas del sistema de referencia del escenario de calibración y eje de coordenadas de la cámara x_c permanece paralelo al plano formado por los ejes de coordenadas x_w - y_w

respecto a cada uno de los ejes se determinan a partir de la posición de la cámara. Basándose en la figura 5.2 dichos ángulos expresados en función de la posición de la cámara son los siguientes:

$$\cos \beta = \frac{t_y}{m_1} \quad \sin \beta = \frac{t_x}{m_1} \quad \cos \alpha = \frac{|t_z|}{m_2} \quad \sin \alpha = \frac{m_1}{m_2} \quad (5.1)$$

$$m_1 = \sqrt{t_x^2 + t_y^2} \quad m_2 = \sqrt{t_x^2 + t_y^2 + t_z^2}$$

m_1 representa el módulo de la traslación de la cámara en el plano X_w - Y_w que simboliza el suelo y m_2 es la distancia desde el origen de coordenadas de la cámara hasta el origen de coordenadas del escenario. La matriz de rotación de la cámara obtenida a partir de los ángulos de Euler, queda definida según la siguiente expresión.

$$R = \begin{bmatrix} \cos \beta & \sin \beta & 0 \\ -\sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix} = \begin{bmatrix} \cos \beta & \cos \alpha \sin \beta & \sin \alpha \sin \beta \\ -\sin \beta & \cos \beta \cos \alpha & \sin \alpha \cos \beta \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix} \quad (5.2)$$

Utilizando las expresiones de los ángulos obtenidas anteriormente a partir de la traslación de la cámara, se tiene la matriz de rotación expresada en función de la posición de la cámara. En este momento el modelo de la cámara sólo depende de los parámetros intrínsecos de la misma y de su posición en el escenario de calibración. De esta forma es posible definir las posiciones que definen un proceso de calibración óptimo. A partir de las posiciones que se obtengan se orientará la cámara de forma que el eje óptico atraviese el centro de la plantilla que se encuentra situado en el origen de coordenadas.

Para reducir las expresiones que aparecen a la hora de definir unas condiciones óptimas de calibración se van a expresar las medidas de las coordenadas de los puntos en la imagen tomando como origen de coordenadas el centro de la misma, es decir que si las especificaciones

técnicas de la cámara dicen que la cámara monta un CCD con dimensiones 640x480 píxeles se asume que el origen de coordenadas de las medidas en la imagen es el punto central establecido el píxel (320,240). Este cambio en el origen de medidas dentro de las imágenes no reduce la generalidad de los resultados. Los parámetros intrínsecos u_0 , v_0 que definen las coordenadas de la intersección del eje óptico con el plano de la imagen, se expresan siempre con respecto al origen de medidas en la imagen. Si las coordenadas de los puntos en la imagen se refieren la esquina superior izquierda, estos parámetros toman valores cercanos a la mitad de las dimensiones del CCD. Si se traslada el origen de medidas dentro en la imagen al centro de la misma se puede asumir que estos parámetros son nulos y en consecuencia reducir significativamente las expresiones que se obtienen. En el conjunto de parámetros que se obtiene del proceso de calibración habrá que tener en cuenta el hecho de que los parámetros u_0 , v_0 están referidos a este nuevo origen de coordenadas.

Otra suposición que se realiza en el modelo de la cámara es referente a los factores de escalado α_u y α_v en cada uno de los ejes de la imagen. Sin pérdida de generalidad y para obtener unas condiciones óptimas de calibración se asume que es el mismo para los dos ejes. Con esto se supone que los píxeles son cuadrados y el factor de escalado se llamará k . Posteriormente se discutirá sobre la validez o no de esta suposición.

Por lo tanto, el modelo de la cámara definido de esta forma sólo depende de la posición de la misma dentro del escenario y el factor de escala α que expresa la relación de píxeles a milímetros en la imagen.

5.2.1. Condición del sistema

El cálculo de los parámetros de la cámara en su fase lineal acaba reduciéndose al cálculo del vector propio de una matriz. El número de condición de una matriz muestra como su nombre indica la condición de la misma, es decir, si la información que contiene la matriz es redundante o cubre todos los ejes dentro del espacio R -dimensional en el que está definida. Toda matriz tiene asociados unos vectores y valores propios los cuales se utilizan para calcular este número de condición. Los vectores propios definen una base ortogonal dentro del espacio R -dimensional en el que se define la matriz. Los vectores propios son ortogonales y de alguna forma, sus longitudes son una medida del número de vectores dentro de la matriz que están orientados hacia esa dirección. Cada uno de los vectores propios tiene asociado un valor propio que indica el módulo del mismo. El cociente entre el módulo del vector más corto con el módulo del vector más largo es lo que se llama el número de condición de la matriz.

De esta forma, una matriz mal condicionada tiene todos los vectores orientados hacia unos determinados ejes. En el caso de estimación de parámetros cada uno de los ejes dentro del

espacio R -dimensional, corresponde a uno de ellos. Por lo tanto, este hecho indica que la información que contiene la matriz solamente está referida a la parte del vector de parámetros que representan estos ejes, pero sin embargo el resto del vector de parámetros está poco afectado por la información contenida en la matriz. En consecuencia, la parte del vector de parámetros a los que se refiere dicha matriz será calculado a partir de datos que no le afectan directamente y por lo tanto será muy sensible al ruido de los mismos. Para obtener unos resultados robustos y fiables de los parámetros, es necesario partir de matriz bien condicionada. Esto significa montar la matriz a partir de vectores que cubran todos los ejes del espacio R -dimensional y además con medidas similares. Dado que los vectores de la matriz se forman a partir de los datos de entrada será necesario tener unos datos que garanticen este hecho. En definitiva las condiciones que tienen que cumplir los vectores de la matriz es que todos ellos sean ortogonales entre ellos y con idéntica longitud. El número de vectores mínimo y necesario será aquel que coincida con la dimensión del espacio R para que la matriz sea de rango completo. Normalmente, dado que los datos se obtienen sin tener en cuenta ninguna restricción se acumulan datos con la esperanza de que todos ellos cubran todo el espacio R -dimensional. Lo que aquí se va a hacer es obtener las restricciones que hacen posible que se obtengan los datos para formar el mínimo número de vectores y que todos ellos sean ortogonales y de módulo idéntico para garantizar una matriz bien condicionada y que los resultados de calibración en la fase lineal sean robustos y fiables. Por supuesto que se pueden añadir más vectores dentro de la matriz, pero dado que se tiene los mínimos necesarios para cubrir el espacio R -dimensional, estos aportarán información redundante.

Particularizando para el caso del cálculo de parámetros de la cámara basado en plantilla bidimensional, se tienen dos etapas intermedias en las que se que se calculan de forma lineal diferentes parámetros. En una de ellas se toman las coordenadas de los puntos en la plantilla y sus correspondientes en las imágenes y se calculan sendas homografías para cada imagen capturada. En una segunda etapa se toman los datos de las homografías y se utilizan para calcular los elementos de la matriz que contiene los parámetros intrínsecos de la cámara. Así pues, teniendo en cuenta el escenario de calibración definido anteriormente se va a definir un procedimiento para que se garantice que las matrices que se forman a partir de los datos tengan un número de condición óptimo. Para ello, sus vectores tendrán que tener módulos idénticos y todos ellos serán ortogonales entre sí. En primer lugar se va a optimizar el proceso de cálculo de los elementos de la matriz que contiene los parámetros intrínsecos. De esta forma se obtendrán unas posiciones óptimas para tomar las imágenes de la plantilla. Posteriormente, a partir de las posiciones calculadas en este caso, se puede establecer alguna regla para que el cálculo de las homografías sea óptimo. En este caso sólo se puede tratar las posiciones de los puntos en la plantilla. Esto permite definir una plantilla de calibración óptima.

5.2.2. Posiciones óptimas para tomar las imágenes

En el proceso de calibración basado en plantilla plana hay una etapa en la que se calculan los elementos que forman la matriz que contiene los parámetros intrínsecos. En concreto en el apartado 3.3.2 se obtienen los elementos de la matriz $\mathbf{K}^{-T} \cdot \mathbf{K}^{-1}$ que contiene los parámetros intrínsecos de la cámara. Tal y como se explica en el apartado 3.3.2, teniendo en cuenta que se trata de una matriz simétrica se obtiene un vector de 6 elementos de la forma:

$$\mathbf{b} = [b_{11} \quad b_{12} \quad b_{22} \quad b_{13} \quad b_{23} \quad b_{33}]^T \quad (5.3)$$

A partir de las distintas homografías calculadas con cada una de las imágenes tomadas de la plantilla bidimensional, si la columna i -ésima de la homografía l , nombrada como ${}^l\mathbf{H}$, es ${}^l\mathbf{h}_i = ({}^l h_{i1}, {}^l h_{i2}, {}^l h_{i3})$, se tiene que:

$${}^l\mathbf{h}_i^T \mathbf{B} {}^l\mathbf{h}_i = {}^l\mathbf{v}_{ij}^T \mathbf{b} \quad (5.4)$$

$${}^l\mathbf{v}_{ij}^T = \begin{bmatrix} {}^l h_{i1} & {}^l h_{j1} & {}^l h_{i1} {}^l h_{j2} + {}^l h_{i2} {}^l h_{j1} & {}^l h_{i2} & {}^l h_{j2} & {}^l h_{i3} {}^l h_{j1} + {}^l h_{i1} {}^l h_{j3} & {}^l h_{i3} {}^l h_{j2} + {}^l h_{i2} {}^l h_{j3} & {}^l h_{i3} & {}^l h_{j3} \end{bmatrix}$$

Con estas expresiones se pueden escribir las restricciones de de la matriz de rotación \mathbf{R} en dos ecuaciones homogéneas en función de \mathbf{b} de la siguiente forma:

$$\begin{bmatrix} {}^l\mathbf{v}_{11}^T & -{}^l\mathbf{v}_{22}^T \\ {}^l\mathbf{v}_{12}^T \end{bmatrix} \mathbf{b} = 0 \quad (5.5)$$

Si se calculan n homografías a partir de n imágenes obtenidas de la plantilla se obtiene un sistema de ecuaciones de la forma

$$\mathbf{V} \cdot \mathbf{b} = 0 \quad (5.6)$$

\mathbf{V} es una matriz de dimensiones $2n \times 6$. El número de homografías necesarias son por lo tanto $n \geq 3$ para obtener una solución general con solución única de \mathbf{b} definida con un factor de escala. Siguiendo este planteamiento, si se quiere obtener una matriz \mathbf{V} bien condicionada, será necesario formarla a partir de vectores ${}^l\mathbf{v}_{ij}$ los cuales sean ortogonales entre sí y a su vez tengan el mismo módulo. El subíndice ij representa el vector \mathbf{v} formado a partir de las columnas i -ésima y j -ésima de la homografía nombrada con el supraíndice l . Es necesario pues tener unas homografías cuyos elementos formen un conjunto de vectores ${}^l\mathbf{v}_{ij}$ los cuales sean todos ortogonales entre sí y con un módulo idéntico. Para ello será necesario fijarse en la formación de los elementos de la homografía.

La formación de una homografía está definida en el apartado 3.3.1 y se basa en el modelo de la cámara teniendo en cuenta que la coordenada z de todos los puntos de la plantilla es nula. Si además se tienen en cuenta las definidas en el escenario de calibración en las que el origen de coordenadas en la imagen es el centro de la misma, $u_0=0$, $v_0=0$, y el factor de escalado

en ambas direcciones de la imagen es el mismo, los elementos de una homografía se obtienen según la siguiente expresión:

$${}^lH = \begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^l r_{11} & {}^l r_{12} & {}^l t_x \\ {}^l r_{21} & {}^l r_{22} & {}^l t_y \\ {}^l r_{31} & {}^l r_{32} & {}^l t_z \end{bmatrix} = \begin{bmatrix} {}^l h_{11} & {}^l h_{12} & {}^l h_{13} \\ {}^l h_{21} & {}^l h_{22} & {}^l h_{23} \\ {}^l h_{31} & {}^l h_{32} & {}^l h_{33} \end{bmatrix} \quad (5.7)$$

${}^l r_{ab}$ representa cada uno de los elementos de la matriz de rotación correspondiente a la homografía l . La posición desde donde se toma cada imagen está representada en el modelo mediante ${}^l t_x, {}^l t_y, {}^l t_z$. Si se tienen en cuenta las condiciones bajo las cuales se ha definido el escenario de calibración, se tiene que la matriz de rotación \mathbf{R} que determina la orientación de la cámara está expresada en función de la traslación de la misma. Formalmente queda expresado con las ecuaciones (5.1) y (5.2). Por lo tanto, las homografías resultantes de cada imagen sólo dependen de la posición en la que se coloca la cámara y el factor de escalado en la imagen. El resto de parámetros se pueden calcular como consecuencia de estos. Extrapolando esta conclusión a los vectores ${}^l \mathbf{v}_{ij}$ que forman la matriz \mathbf{V} , dado que estos se forman a partir de los elementos de la homografía ${}^l \mathbf{H}$, es posible definir sus elementos a partir de la posición en la que se coloca la cámara. El procedimiento es el que sigue. Substituyendo en la matriz de rotación \mathbf{R} definida por (5.2) los valores de las expresiones (5.1), se tiene la orientación de la cámara en función de la traslación de la misma. Utilizando (5.7) se obtienen los elementos de la homografía los cuales sólo dependen de la posición de la cámara y el factor de escalado en la imagen. Estos elementos de la homografía se utilizan para formar los vectores ${}^l \mathbf{v}_{ij}$ según la expresión (5.4). Finalmente por cada homografía se obtiene una pareja de vectores según (5.5) que forman parte de la matriz \mathbf{V} .

Para obtener una matriz \mathbf{V} bien condicionada, las expresiones que tienen que cumplir los vectores que la forman son las siguientes, teniendo en cuenta que se tienen que cumplir para todo i, j, l .

$$\left({}^l \mathbf{v}_{11}^T - {}^l \mathbf{v}_{22}^T \right) {}^l \mathbf{v}_{12} = 0 \quad (5.8)$$

$$\left({}^l \mathbf{v}_{11}^T - {}^l \mathbf{v}_{22}^T \right) \left({}^l \mathbf{v}_{11} - {}^l \mathbf{v}_{22} \right) - {}^l \mathbf{v}_{12}^T \cdot {}^l \mathbf{v}_{12} = 0 \quad (5.9)$$

La primera representa la condición de ortogonalidad entre los dos vectores y la segunda representa la igualdad de módulos. i, j , toman valores 1 y 2 ya que corresponden a los dos primeras columnas de la homografía l . El número l de homografías son como mínimo 3. Si se establecen 3 homografías cuyos vectores sean ortogonales y tengan el mismo módulo se tendrá toda la información necesaria para calcular los elementos del vector \mathbf{b} ya que se tendrá una matriz \mathbf{V} bien condicionada. Se van a establecer diferentes reglas cada una de las cuales garantiza una condición de ortogonalidad o igualdad de módulo entre los vectores. Éstas surgen a partir de plantear las expresiones para las condiciones de ortogonalidad e igualdad de módulo y encontrar las condiciones bajo las cuales son ciertas. El objetivo será establecer un conjunto de

reglas que hagan que se cumplan el máximo número de condiciones posibles. Las dos primeras reglas obtienen las condiciones bajo las cuales los dos vectores generados con una misma homografía tienen el mismo módulo y son ortogonales. A partir de la tercera se establecen reglas que garantizan las restricciones para el mayor número de vectores posibles. Si se tiene en cuenta que dos homografías generan 4 vectores, es necesario establecer las posiciones de las mismas para que los cuatro vectores tengan el mismo módulo y además sean ortogonales entre sí. Eso acaba representando doce expresiones las cuales tienen que ser ciertas bajo las mismas condiciones. Basándose en la expresión (5.5), en el caso de dos homografías las ecuaciones que tienen que ser ciertas son:

$$\left({}^1v_{11}^T - {}^1v_{22}^T\right){}^1v_{12} = 0 \quad (5.10)$$

$$\left({}^1v_{11}^T - {}^1v_{22}^T\right)\left({}^1v_{11} - {}^1v_{22}\right) - \left({}^1v_{12}^T \cdot {}^1v_{12}\right) = 0 \quad (5.11)$$

$$\left({}^2v_{11}^T - {}^2v_{22}^T\right){}^2v_{12} = 0 \quad (5.12)$$

$$\left({}^2v_{11}^T - {}^2v_{22}^T\right)\left({}^2v_{11} - {}^2v_{22}\right) - \left({}^2v_{12}^T \cdot {}^2v_{12}\right) = 0 \quad (5.13)$$

$$\left({}^1v_{11}^T - {}^1v_{22}^T\right){}^2v_{12} = 0 \quad (5.14)$$

$$\left({}^1v_{11}^T - {}^1v_{22}^T\right)\left({}^1v_{11} - {}^1v_{22}\right) - \left({}^2v_{12}^T \cdot {}^2v_{12}\right) = 0 \quad (5.15)$$

$$\left({}^1v_{11}^T - {}^1v_{22}^T\right)\left({}^2v_{11} - {}^2v_{22}\right) = 0 \quad (5.16)$$

$$\left({}^1v_{11}^T - {}^1v_{22}^T\right)\left({}^1v_{11} - {}^1v_{22}\right) - \left({}^2v_{11}^T - {}^2v_{22}^T\right)\left({}^2v_{11} - {}^2v_{22}\right) = 0 \quad (5.17)$$

$${}^1v_{12}^T \cdot \left({}^2v_{11}^T - {}^2v_{22}^T\right) = 0 \quad (5.18)$$

$${}^1v_{12}^T \cdot {}^1v_{12} - \left({}^2v_{11}^T - {}^2v_{22}^T\right)\left({}^2v_{11} - {}^2v_{22}\right) = 0 \quad (5.19)$$

$${}^1v_{12}^T \cdot {}^2v_{12} = 0 \quad (5.20)$$

$$\left({}^1v_{12}^T \cdot {}^1v_{12}\right) - \left({}^2v_{12}^T \cdot {}^2v_{12}\right) = 0 \quad (5.21)$$

Para establecer las condiciones de ortogonalidad e igualdad de módulo entre los vectores generados a partir de una homografía se generan los vectores ${}^l v_{11}$, ${}^l v_{22}$, y ${}^l v_{12}$ utilizando la expresión (5.4). Se utilizan las expresiones (5.1), (5.2) y (5.7) para expresarlos en función de la posición de la cámara ${}^l t_x$, ${}^l t_y$, ${}^l t_z$ y su factor de escalado k . Cada una de las siguientes reglas garantiza una restricción.

Regla 1

En este caso se establece la condición de ortogonalidad entre los dos vectores generados a partir de la misma homografía l . El objetivo es que se cumplan las expresiones (5.10) y (5.12). Simplificando la expresión para encontrar una posición de la cámara que garantice la ortogonalidad entre los dos vectores generados con la homografía l se obtienen tres reglas:

1a) Si ${}^l t_x = 0$, ${}^l t_y$, ${}^l t_z$ pueden tomar cualquier valor

1b) Si ${}^l t_y=0$, ${}^l t_x, {}^l t_z$ pueden tomar cualquier valor

1c) Si ${}^l t_x \neq 0$ y ${}^l t_y \neq 0$ entonces ${}^l t_z$ puede tomar 4 valores diferentes dados por:

$${}^l t_z = \pm \sqrt{\frac{\left(1 + 2 \cdot {}^l m_1^2 - 2 \cdot {}^l m_1^4\right) {}^l t_x^2 + {}^l t_y^2 + 2 \cdot {}^l m_1^2 \cdot {}^l t_y^2 - {}^l m_1^4 \cdot {}^l t_y^2 \pm \sqrt{\left(3 - 4 \cdot {}^l m_1^2 + 3 \cdot {}^l m_1^4\right) {}^l m_1^4}}{-2 \cdot {}^l m_1^4}} \quad (5.22)$$

Regla 2

Ahora se trata que los vectores generados a partir de la misma homografía tengan el mismo módulo. De todas las posiciones posibles sólo se toman las definidas por las reglas 1a, 1b y 1c. En este caso, se tiene que cumplir las expresiones (5.11) y (5.13). La regla 2a surge de la 1a, la 2b sale a partir de 1b, y 1c se utiliza para determinar la regla 2c. De alguna forma estas reglas son las mismas que las reglas 1 pero más restrictivas ya que a la condición de ortogonalidad se ha añadido la condición de igualdad de módulo. Las expresiones obtenidas en este caso son:

2a) Si ${}^l t_x=0$, ${}^l t_y$ puede tomar cualquier valor y ${}^l t_z$ toma 4 valores diferentes:

$${}^l t_z = \pm \sqrt{-\frac{2 \cdot {}^l t_y^2 \left(1 + 2 \cdot {}^l t_y^2 + 2 \cdot {}^l t_y^4 + \sqrt{\left(-3 - 4 \cdot {}^l t_y^2 + 12 \cdot {}^l t_y^4\right)}\right)}{1 + 3 \cdot {}^l t_y^2 + {}^l t_y^4}} \quad (5.23)$$

$${}^l t_z = \pm \sqrt{-\frac{2 \cdot {}^l t_y^2 \left(-1 - 2 \cdot {}^l t_y^2 - 2 \cdot {}^l t_y^4 + \sqrt{\left(-3 - 4 \cdot {}^l t_y^2 + 12 \cdot {}^l t_y^4\right)}\right)}{1 + 3 \cdot {}^l t_y^2 + {}^l t_y^4}}$$

2b) Si ${}^l t_y=0$, ${}^l t_x$ puede tomar cualquier valor y ${}^l t_z$ toma 4 valores diferentes:

$${}^l t_z = \pm \sqrt{-\frac{2 \cdot {}^l t_x^2 \left(1 + 2 \cdot {}^l t_x^2 + 2 \cdot {}^l t_x^4 + \sqrt{\left(-3 - 4 \cdot {}^l t_x^2 + 12 \cdot {}^l t_x^4\right)}\right)}{1 + 3 \cdot {}^l t_x^2 + {}^l t_x^4}} \quad (5.24)$$

$${}^l t_z = \pm \sqrt{-\frac{2 \cdot {}^l t_x^2 \left(-1 - 2 \cdot {}^l t_x^2 - 2 \cdot {}^l t_x^4 + \sqrt{\left(-3 - 4 \cdot {}^l t_x^2 + 12 \cdot {}^l t_x^4\right)}\right)}{1 + 3 \cdot {}^l t_x^2 + {}^l t_x^4}}$$

2c) Si ${}^l t_x \neq 0$ y ${}^l t_y \neq 0$ entonces ${}^l t_z$ puede tomar 4 valores diferentes determinados por la regla 1c. Sin embargo, para que se cumpla la condición de igualdad de módulo, los valores de ${}^l t_x$ y ${}^l t_y$ están relacionados con la siguiente expresión:

$${}^l t_x = \pm \frac{\sqrt{\pm \sqrt{2} - 2 \cdot {}^l t_y^2}}{\sqrt{2}} \quad (5.25)$$

La regla 2a y 2b establecen situar la cámara de modo que una de las coordenadas de la traslación en el suelo sea nula y la otra pueda tomar cualquier valor. A partir de esta otra coordenada, la altura del suelo con que se sitúe la cámara se calcula en función de la coordenada no nula en el plano X-Y. De las cuatro alturas que se calculan, dos de ellas son imaginarias y las otras dos son cambiadas de signo, es decir, que una se refiere a una altura positiva y la otra a una negativa. Por lo tanto, si se establece el plano del suelo como plano X-Y difícilmente se

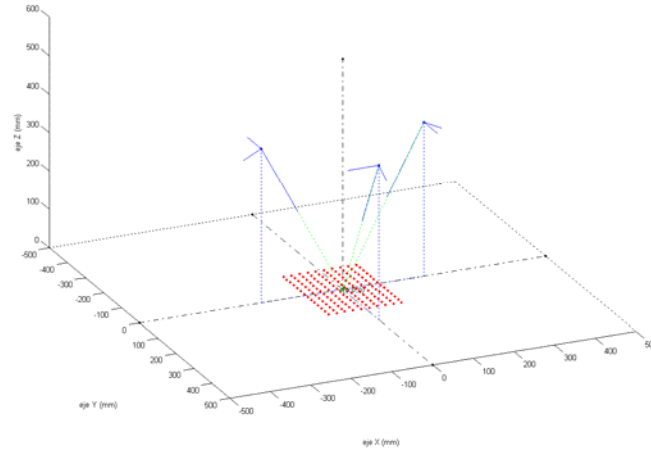
podrán obtener traslaciones positivas en el eje Z respecto a esta referencia ya que eso significa colocar la cámara debajo del suelo. Así pues, de las cuatro posiciones que se calculan, sólo es válida aquella que tiene la traslación en el eje Z negativa, ya que es así como se ha definido el escenario de calibración. La conclusión establecida por la regla 2c restringe solamente la posición de la cámara a valores de 1t_x y 1t_y muy cercanos a cero. Estos valores de 1t_y serán aquellos que hagan que el radical de la expresión (5.25) sea positivo. En caso contrario el resultado obtenido será imaginario y carecerá de utilidad práctica.

Hasta el momento se tienen las posiciones para tomar la primera imagen y conseguir que los vectores generados con la homografía correspondiente sean ortogonales y del mismo módulo. Ahora se tiene que conseguir que estos dos vectores sean ortogonales con los dos siguientes, obtenidos con la segunda homografía. Igual que antes se generan los vectores ${}^1\mathbf{v}_{11}$, ${}^1\mathbf{v}_{22}$, y ${}^1\mathbf{v}_{12}$ correspondientes a primera homografía y ${}^2\mathbf{v}_{11}$, ${}^2\mathbf{v}_{22}$, y ${}^2\mathbf{v}_{12}$ utilizando la segunda homografía. Se usan las expresiones (5.4) y (5.1), (5.2) y (5.7) para expresarlas en función de las posiciones de las cámaras 1t_x , 1t_y , 1t_z y 2t_x , 2t_y , 2t_z respectivamente y el factor de escalado α de la misma. Por supuesto, que las condiciones obtenidas para 1t_x , 1t_y , 1t_z son completamente extrapolables a 2t_x , 2t_y , 2t_z ya que con la homografía correspondiente a la segunda imagen se genera dos vectores para componer la matriz \mathbf{V} . Es por este motivo que se van a utilizar las obtenidas anteriormente para definir las nuevas posiciones.

Regla 3

En este caso se considera hacer nula la restricción expresada por (5.14) que garantiza la ortogonalidad entre dos vectores generados por dos homografías diferentes. Como resultado se obtiene que esta expresión es cierta siempre que la traslación a lo largo del eje X ó Y del sistema de referencia del escenario sea nula para las dos imágenes a la vez. Cualquier combinación es posible, es decir, ${}^1t_x = {}^2t_x = 0$, ${}^1t_x = {}^2t_y = 0$, ${}^1t_y = {}^2t_x = 0$, ${}^1t_y = {}^2t_y = 0$. Esto implica que situando la cámara en estas condiciones se consigue satisfacer la ortogonalidad de los vectores correspondientes a las propias homografías y el caso concreto de los dos vectores planteados en esta regla. Asimismo, las condiciones definidas por esta regla 3 son también válidas en el caso de intentar hacer cierta la expresión (5.18).

En el caso de tomar las conclusiones de las reglas 1c y 2c, se establece una complicada relación entre las coordenadas 1t_y y 2t_y para que se cumpla la expresión (5.14). Dada la complejidad, esta solución se desecha. Llegados a este punto las posiciones óptimas para capturar las imágenes de la plantilla están situadas a lo largo de los ejes X e Y del plano del suelo. Bajo estas condiciones sería ideal que se cumplieran el resto de las restricciones que quedan pendientes. Dado que no se cumplen se va a establecer una nueva regla que haga que sean ciertas algunas de las restantes.

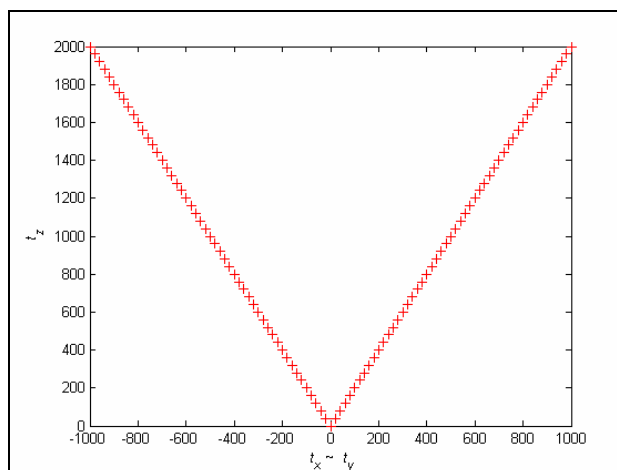


5.3 Posiciones óptimas para obtener las imágenes de la plantilla. Una coordenada en el suelo se mantiene nula mientras la otra varía. La altura se calcula partir de las expresiones (5.23) ó (5.24)

Regla 4

Partiendo de los resultados obtenidos de la regla 3 se coloca la cámara en sendas posiciones tales que la coordenada X ó Y respecto del sistema de referencia del escenario de calibración sea nula. Con esto se garantizan todas las reglas definidas hasta el momento que hacen ciertas las expresiones. Para garantizar el cumplimiento de las expresiones (5.15), (5.17) y (5.21), se establece que la coordenada restante en el plano del suelo sea complementaria en ambas imágenes. Por ejemplo, si ${}^1t_x = {}^2t_x = 0$ se coloca la cámara de forma que ${}^1t_y = -{}^2t_y$. De la misma forma, si se elige la opción ${}^1t_y = {}^2t_y = 0$ la coordenada restante para cada una de las posiciones será ${}^1t_x = -{}^2t_x$. De este modo, utilizando la expresión de la regla 2a ó 2b la traslación de la cámara en el eje Z para las dos imágenes será la misma, es decir ${}^1t_z = {}^2t_z$. Así se consigue que la restricciones (5.15), (5.17) y (5.21) sean ciertas y además las (5.10), (5.11), (5.12), (5.13), (5.14) y (5.18). Sólo quedan que sean ciertas las restricciones (5.16), (5.19) y (5.20).

Para que las restricciones (5.16), (5.19) y (5.20) sean ciertas se debe de acotar el único grado de libertad que queda libre. Éste es el de las coordenadas que se definen complementarias de la forma ${}^1t_x = -{}^2t_x$. Hasta el momento esta traslación se elegía aleatoriamente. Para cada una de ellas se cumple que una coordenada determinada hace la restricción nula. Es por este motivo que no se pueden fijar unas condiciones generales para que todas las restricciones sean ciertas, pero si que se puede conseguir que la mayoría de ellas lo sean. Éstas son, que una de las traslaciones a lo largo de los ejes del suelo sea nula y la otra sea complementaria. Una vez la cámara se tiene situada en el suelo se elige su altura a partir de las expresiones (5.23), (5.24) dependiendo de la coordenada que se haya elegido nula. La figura 5.4 muestra como evoluciona la altura de la cámara en función de la traslación que se elige en el suelo. Teniendo en cuenta que al menos tres imágenes son necesarias, se pueden capturar dos de ellas haciendo la traslación a lo largo del eje



5.4 Altura de la cámara en función de lo alejada que se encuentra del origen de medidas. La altura se calcula partir de las expresiones (5.23) ó (5.24)

X de plano de suelo nula, y otras dos tomando una traslación a lo largo del eje Y nula. La figura 5.3 muestra estas posibles posiciones.

5.3. Consideraciones prácticas

En el apartado anterior se han definido las posiciones desde las que se deben de tomar las imágenes para conseguir una estimación lineal del vector \mathbf{b} a partir del cual se extraen los parámetros intrínsecos de la cámara. Con esto se tiene desde un punto de vista teórico una estimación de los parámetros en su fase lineal con cierta robustez. Las posiciones óptimas de la cámara serán aquellas que tengan una de las coordenadas ${}^l t_x$ ó ${}^l t_y$ nula, y dependiendo de lo alejada que esté la cámara en el otro eje que representan el plano del suelo, se calculará la altura de la cámara expresada en una traslación ${}^l t_z$ a lo largo del eje Z , dada por la expresión (5.23) ó (5.24). Es necesario establecer una distancia fija de la cámara al origen de medidas del escenario para todas las posiciones. Esto hará que se tenga la cámara colocada a la misma altura en todas ellas y de esta forma se condicione de forma óptima el proceso de calibración. Haciendo una representación gráfica de las expresiones (5.23) ó (5.24) se puede obtener la evolución de la altura de la cámara en función de lo alejada que se coloque del origen de coordenadas, la cual se muestra en la figura 5.4. Se observa que a pesar de lo aparatosa que resulta la expresión, la relación que existe entre el alejamiento y la altura de la cámara es bastante lineal, resultando además con una pendiente de valor dos, es decir la altura de la cámara será el doble del desplazamiento que se tome. Esto indica que el ángulo que se establece entre el plano de la imagen y el plano de la plantilla son siempre 63.5° , teniendo en cuenta como dato importante que no depende de las características de la cámara. De esta forma, la posición de la cámara se determinará de forma manual más o menos cercana al origen, dependiendo de las características de la cámara y basándose en el encuadre de la imagen sobre la plantilla. Se deberá conseguir que

la plantilla ocupe la mayor parte de imagen acercando o alejando la cámara manualmente, teniendo en cuenta que la altura de la misma sea el doble del alejamiento del origen de medidas en el suelo. Finalmente se tendrá cuidado de orientar la cámara para que el eje óptico atraviese el origen de coordenadas del escenario y además garantizar que el borde superior de la imagen sea paralelo al plano del suelo en el que se encuentra situada la plantilla.

Todas estas condiciones se han definido de forma teórica y por supuesto que rara vez se cumplirán de forma exacta. Por un lado se pueden calcular unas posiciones para tomar cada una de las imágenes de la plantilla pero se puede entender que físicamente será complicado situar la cámara de forma exacta en la posición calculada. Además, una vez situada la cámara es necesario orientarla de forma que el eje óptico atraviese exactamente el centro de la plantilla que coincide con la coordenada $(0, 0, 0)$ del sistema de coordenadas del escenario de calibración, y además alinear el borde superior de la imagen con el plano del suelo. Asimismo, una vez obtenidas las imágenes desde posiciones óptimas, los errores de medidas en las coordenadas de los puntos hacen que se cometan errores en las estimaciones de las homografías y en consecuencia afecten a la condición de la matriz V . Por lo tanto se asume que estas condiciones son teóricas y que ayudan a definir unas condiciones óptimas de calibración que se cumplirán de forma aproximada.

Por otro lado, en el proceso de estimación de las homografías, las distorsiones no lineales que se producen debidas a imperfecciones de construcción de la cámara no se han tenido en cuenta en la definición de las posiciones óptimas. Este hecho tampoco desvirtúa notablemente el resultado, ya que aunque se hubieran tenido en cuenta, el propio ruido en las medidas de las coordenadas de los puntos hubiera hecho que los vectores resultantes formados a partir de las coordenadas no hubieran sido perfectamente ortogonales ni con el mismo módulo. Además, teniendo en cuenta que se ha determinado un procedimiento para corregir la distorsión de la imagen antes de proceder con la calibración de la cámara, se entiende que con coordenadas libres de ruido se podrían conseguir las condiciones óptimas.

Así pues, está claro que las expresiones obtenidas en el apartado anterior son puramente teóricas, pero sin embargo, sí que ayudan a determinar unas condiciones óptimas para calibrar. De forma orientativa se tienen unas condiciones entorno a las cuales la estimación lineal de los parámetros resulta estable y robusta. Además, en la fase de búsqueda no lineal de parámetros estas condiciones precalculadas, ayudan notablemente en la consecución de mejores resultados. Si se es capaz de ajustar físicamente la cámara a algunas de las condiciones precalculadas, es posible conocer con cierto nivel de confianza el parámetro conseguido físicamente. Esto hace que el parámetro en cuestión no sea necesario añadirlo al conjunto de parámetros que se buscan en la fase no lineal. Por lo tanto la dimensión del espacio de búsqueda disminuye y las probabilidades de obtener resultados más cercanos a los reales aumentan. Por supuesto que físicamente no se podrá colocar la cámara exactamente en una determinada

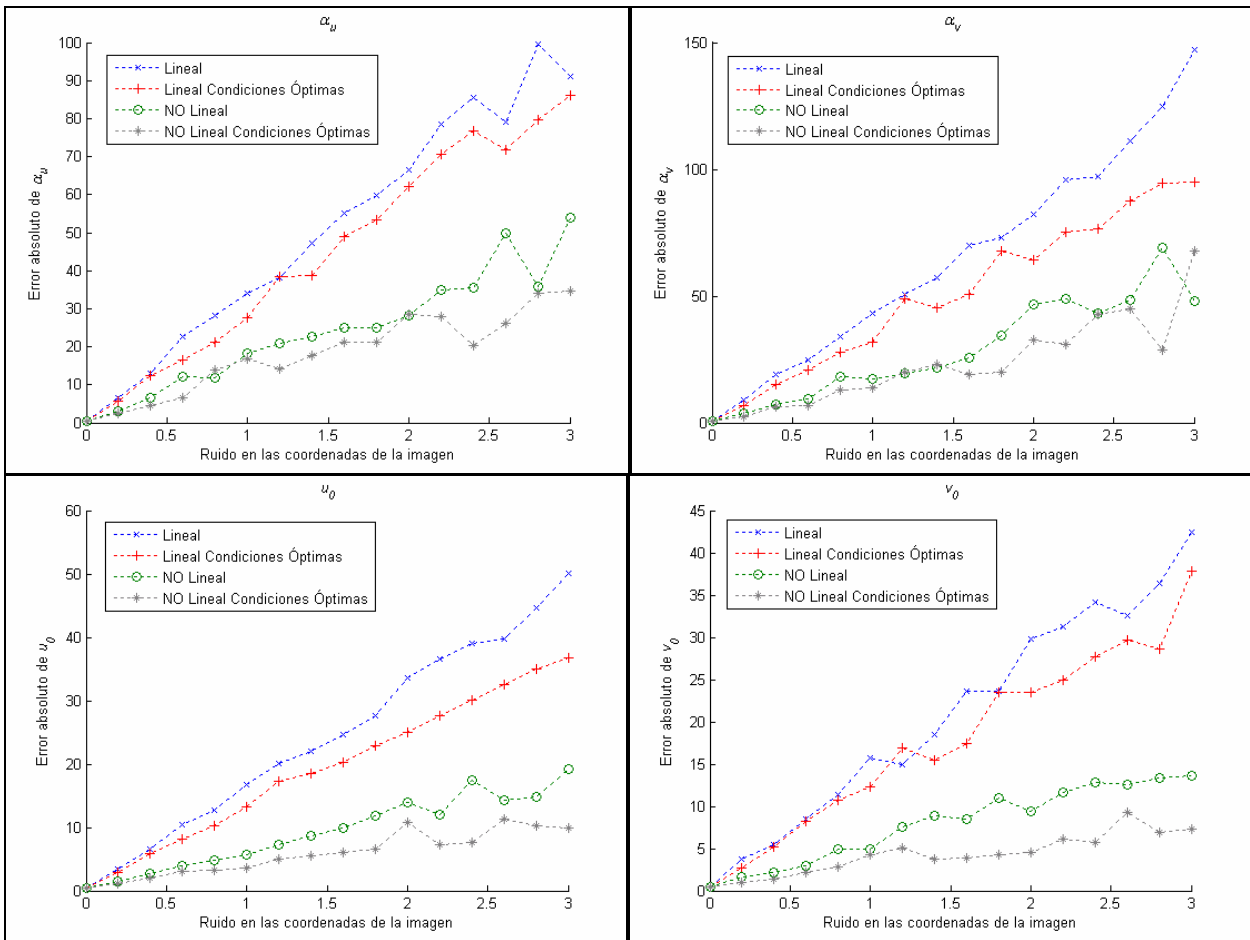
posición, pero si se consigue aproximar bastante, este valor será igual de válido que el que se pueda obtener como resultado de una búsqueda no lineal del parámetro. La reducción de la dimensión del espacio de búsqueda dependerá del número de parámetros ajustados físicamente.

En el caso de utilizar por ejemplo un brazo robot para posicionar la cámara en el escenario, las traslaciones a lo largo de los tres ejes de coordenadas del escenario pueden considerarse conocidas. También es conocido uno de los ángulos de rotación de la cámara, ya que si se considera el borde superior de la imagen paralelo al plano de la cámara, uno de los ángulos de rotación es nulo. Si se utilizan técnicas más rudimentarias como podría ser un trípode, al menos un desplazamiento respecto a uno de los ejes si que se puede ajustar y considerarlo cero al mismo tiempo que se considera un ángulo de rotación nulo manteniendo el borde superior de la imagen paralelo al plano de la imagen. El mantener el borde superior de la imagen paralela al plano del suelo es relativamente fácil ya que no existe el grado de libertad que permita cambiar este ángulo si se utiliza un trípode convencional.

5.4. Resultados experimentales

En este apartado se comparan los resultados obtenidos en el capítulo anterior calibrando la cámara desde posiciones aleatorias con los que se obtienen en el caso de situar la cámara de forma óptima para tomar las imágenes. Se corrigen las coordenadas de los puntos medidos en la imagen para que cumplan las rectas teóricas que deberían cumplir en el caso de ausencia de ruido y así obtener unas homografías más cercanas a las reales. De esta forma se garantizan en mejor medida las restricciones planteadas para que el procedimiento de cálculo sea óptimo.

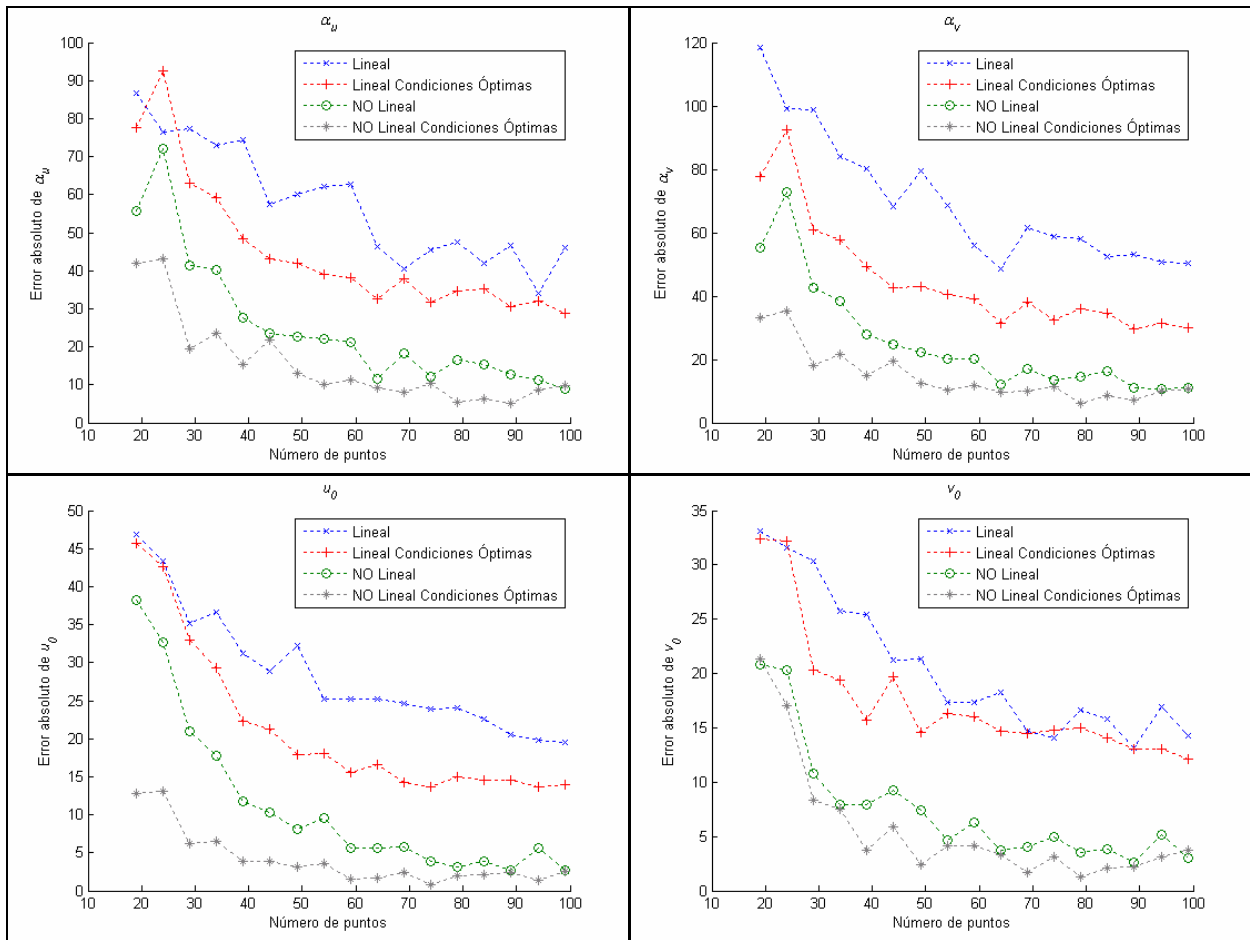
Todos los datos para evaluar los algoritmos son generados mediante simulación. Las características de la cámara son $\alpha_u=1250$, $\alpha_v=900$, $u_0=250$, $v_0=250$. La resolución de la imagen es 512×512 píxeles. La plantilla, de la misma forma que hace Zhang es un tablero de ajedrez de $10 \times 14 = 140$ esquinas. El tamaño de la plantilla es 180×250 mm. En este caso, la plantilla está situada tal como se define el escenario de calibración. Ésta coincide con el de simulaciones anteriores. La plantilla está situada en el suelo de forma que todos los puntos de interés tienen coordenada $z_w=0$, según se muestra en la figura 5.3. Las posiciones utilizadas para obtener unas condiciones de calibración óptimas son $\mathbf{t}_1=(200, 0, 401.8)$, $\mathbf{t}_2=(-200, 0, 401.8)$, $\mathbf{t}_3=(0, 200, 401.8)$. Los resultados de calibración en este caso son comparados con los resultados obtenidos en el caso de colocar la cámara en cualquier punto. Los puntos elegidos son $\mathbf{t}_1=(150, 200, 580)$, $\mathbf{t}_2=(-50, 250, 880)$, $\mathbf{t}_3=(100, -20, 820)$. Se toman tres imágenes en ambos casos para poder comparar resultados con igual número de imágenes. La orientación de la cámara en todos los casos es tal que el eje óptico corta al plano de la plantilla en el centro de la misma.



5.5 Comparación de errores de estimación de los parámetros intrínsecos en diferentes condiciones de calibración. El ruido en la plantilla es nulo mientras en la imagen varía.

En el caso de situar la cámara en posiciones óptimas se asume como conocido el parámetro de la posición de la cámara que es igual a cero. Con esto se reduce el vector de parámetros en la fase de búsqueda no lineal. También es necesario tener en cuenta que las coordenadas de los puntos en las imágenes se expresan respecto al origen de coordenadas situado en el centro de la misma. En este caso los parámetros u_0 , v_0 que se calculan están expresados respecto a este origen de coordenadas en el centro de la imagen.

En primer lugar se evalúa como afecta el ruido de las medidas de la imagen en las condiciones óptimas de calibrado. La figura 5.5 muestra los resultados. El resultado de los parámetros obtenidos en la fase lineal de calibración se mejora en el caso de colocar la cámara en posiciones preestablecidas. En el caso de la búsqueda no lineal de los parámetros, debido a que se reduce el espacio de búsqueda los resultados que se obtienen en este caso también son ligeramente mejores. En segundo lugar se ven los efectos del número de puntos en la plantilla los cuales se muestran en la figura 5.6. La calibración de la cámara tomando imágenes desde posiciones precalculadas obtiene menores errores en los parámetros.



5.6 Comparación de errores de estimación de los parámetros intrínsecos en diferentes condiciones de calibración. El ruido en la plantilla es nulo mientras en la imagen vale $\sigma=1$. Se cambia el número de puntos

Las cámaras reales también se han calibrado a partir de tres imágenes tomadas en posiciones precalculadas y se han comparado con los resultados en el caso de tomarlas desde tres posiciones aleatorias. Se ajustan al máximo las posiciones desde las que se toman las imágenes con el objetivo de conseguir un procedimiento de calibración bien condicionado. Los resultados obtenidos se muestran en la tabla 5.1. En condiciones óptimas en la fase lineal se obtienen resultados parecidos a los obtenidos con la búsqueda no lineal en posiciones aleatorias.

5.5. Conclusiones

Partiendo de un criterio de optimización se han definido unas condiciones bajo las cuales el proceso de calibración de una cámara resulta robusto en la fase lineal de cálculo y elimina variables de búsqueda en su fase no lineal o iterativa. El criterio utilizado es el número de condición de la matriz formada por los elementos de las homografías llamada V . Las condiciones óptimas de calibración obtenidas han sido las posiciones desde las cuales se capturan las imágenes ya que de esta forma se consigue un proceso de cálculo más estable y robusto

Logitech Quick Cam Pro

	Lineal	Lineal Condiciones Óptimas	No Lineal	No Lineal Condiciones Óptimas
α_u	942.4	931.1	939.5±55.2	937.5±22.3
α_v	937.5	932.8	947.3±52.5	934.5±21.5
u_0	302.7	326.7	314.3±8.9	322.6±5.7
v_0	257.1	245.9	245.8±8.1	242.5±6.5
φ	1.57	1.01	-	-
k_1	2.32	5.12	-	-
k_2	0.26	0.72	-	-
e	0.545	0.245	0.248	0.198
t	0.001	0.001	12.743	12.145

Canon PowerShot A640

	Lineal	Lineal Condiciones Óptimas	No Lineal	No Lineal Condiciones Óptimas
α_u	845.3	834.2	833.8±25.5	932.4±14.7
α_v	862.4	838.8	835.2±17.6	840.2±16.9
u_0	308.7	319.4	316.5±3.2	322.4±3.4
v_0	251.9	245.5	248.6±7.1	245.8±3.5
φ	0.96	0.91	-	-
k_1	0.491	0.416	-	-
k_2	0.12	0.16	-	-
e	0.585	0.246	0.219	0.178
t	0.001	0.001	12.739	12.542

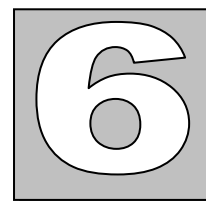
5.1 Comparación de resultados de calibración de dos cámaras reales utilizando 3 imágenes tomadas desde posiciones predefinidas y posiciones aleatorias

cuando se resuelve de forma lineal. Teniendo en cuenta que se establecen unas posiciones en las que se debe colocar la cámara para capturar las imágenes, si se tiene cuidado en alcanzar físicamente estas posiciones antes de tomar las imágenes, se pueden considerar como válidas para el calibrado de la cámara. De esta forma, los parámetros considerados como válidos no se calculan en el proceso de búsqueda no lineal, se reduce el espacio de búsqueda y se consigue más acierto en los parámetros buscados. Hay que tener en cuenta que cualquier proceso de cálculo de los parámetros de la cámara va a cometer errores, y por lo tanto es totalmente factible asumir ciertos parámetros como válidos si se ha tenido cuidado en ajustarlos previamente.

Las condiciones óptimas de calibración han quedado definidas respecto a un escenario. En este escenario se coloca la plantilla de calibración en el plano X - Y considerado como suelo. La traslación negativa del eje Z del escenario es sobre el suelo. La cámara se coloca sobre el suelo y se orienta de forma que el eje óptico atraviesa el centro de la plantilla. De esta forma las posiciones óptimas para realizar la calibración es colocar la cámara de forma que una de las traslaciones respecto a los ejes del suelo sea nula en todas las posiciones. La otra traslación respecto al plano que simboliza el suelo ha de ser siempre la misma en valor absoluto. A partir de ésta, se calcula la altura de la cámara sobre el suelo que de forma aproximada es el doble del alejamiento de la cámara del origen de coordenadas. En consecuencia el ángulo que forma el plano de la imagen con el plano de la plantilla son 63.5° . Estas posiciones no garantizan plenamente que la matriz V que se forma tiene un número de condición unitario. Existen varias de las condiciones de ortogonalidad e igualdad de módulo que no se satisfacen. Sin embargo si que dan una idea de cómo situar la cámara para realizar la calibración. Cualquier otra posición

seguro que hace que se cumplan un menor número de restricciones nombradas en las expresiones de la (5.10) a la (5.21).

Es interesante resaltar que las posiciones óptimas no dependen en ningún momento de las características de la cámara. Esto permite poder utilizarlas con cualquier cámara sin tener que conocer previamente ninguno de sus parámetros. Las condiciones particulares de la cámara servirán para realizar el encuadre de la plantilla de forma que ocupe la mayor parte posible de la imagen.



CONCLUSIONES

Para estudiar el comportamiento de la cámara y poder obtener un modelo de la misma que realice una transformación del espacio proyectivo P^3 al espacio P^2 existen diferentes formas las cuales se ajustan más o menos al comportamiento real de la misma. El modelo básico es el llamado pin-hole a partir del cual se construyen los demás modelos. Este modelo básico solamente transforma las coordenadas de los puntos en la escena sin expresar la relación entre el sistema de referencia de la cámara y el del escenario, ni tampoco tiene en cuenta las transformaciones que se producen en la formación de la imagen. Un modelo más ajustado al comportamiento real de la cámara añade al modelo pin-hole las transformaciones lineales que sufren las coordenadas de los puntos en la imagen. Éstas son el escalado que permite el cambio de unidades de píxeles a milímetros, la localización del origen de medidas en la esquina superior izquierda de la imagen y la no ortogonalidad de los ejes de medidas debida a imperfecciones en el proceso de construcción. El hecho de que los ejes de la imagen no formen 90 grados puede ser tratado de forma lineal según la expresión (2.12). El modelo más ajustado al comportamiento de la cámara es el que tiene en cuenta las distorsiones que se producen en el sistema de visión por las imperfecciones constructivas de las lentes. Estas imperfecciones producen desviaciones en la trayectoria del haz de luz que a priori se supone recta. El efecto de estas desviaciones se traduce en distorsiones radiales y de descentrado de la imagen en el CCD. El modelo de la cámara contiene todas las transformaciones que sufren las coordenadas de un punto en la escena para poder expresarlas respecto al sistema de coordenada de la imagen. Las transformaciones están definidas por los parámetros intrínsecos o extrínsecos según describan las características propias de la cámara o su posición y orientación respecto al sistema de referencia de la escena. Con un conjunto de parámetros dado es posible construir la matriz de proyección que los representa. Del

mismo modo, a partir de una matriz de proyección es posible extraer los parámetros intrínsecos y extrínsecos que la generan.

En el caso de tener un modelo implícito, existe una matriz de proyección que relaciona puntos en espacio tridimensional situados en el escenario, con sus correspondientes puntos colocados en un espacio bidimensional que es la imagen. Un caso particular de matriz de proyección es la llamada homografía, la cual relaciona dos conjuntos de puntos situados en sendos espacios bidimensionales. Estos espacios bidimensionales pueden ser una plantilla plana relacionada con su imagen, o dos imágenes diferentes de la misma plantilla. En este caso, al contrario que ocurre con la matriz de proyección, de una homografía no se pueden extraer los parámetros intrínsecos y extrínsecos de la cámara. Sin embargo, con un conjunto de parámetros si que es posible generar una homografía. Finalmente cabe destacar sobre el modelo de la cámara que se trabaja en coordenadas homogéneas para poder tratar la proyección perspectiva como una transformación lineal.

El modelo de la cámara es un elemento necesario si se desea utilizar la información de las imágenes para realizar medidas de la escena o para efectuar reconstrucciones 3D de la misma. Para obtener un modelo de la cámara es necesario realizar una calibración de la misma. Este proceso consiste en estimar los parámetros intrínsecos y extrínsecos de la cámara o su correspondiente matriz de proyección. Para llevarlo a cabo se pueden utilizar tanto plantillas tridimensionales, bidimensionales o unidimensionales, así como también el procedimiento llamado auto calibración que no necesita plantillas de calibración. Básicamente, en función de la exactitud que se desee para los parámetros o el número de parámetros a estimar se realiza en dos etapas. En la primera etapa se calcula un valor aproximado de una parte de los parámetros utilizando métodos de optimización lineal. Esta aproximación se toma como punto de partida para una búsqueda no lineal de todo el conjunto de parámetros que modelan la cámara y la sitúan en el escenario. En esta segunda etapa se utilizan métodos iterativos de búsqueda. La efectividad del proceso de calibración depende tanto de la calidad en las medidas que intervienen, así como del modelo a estimar, sin olvidar el método utilizado para la estimación. A lo largo del capítulo 3 se ha realizado un exhaustivo estado del arte sobre las diferentes técnicas de calibración. No sólo se han tratado técnicas de calibración en si mismas, sino que también se han tenido en cuenta posibles técnicas que mejoren la estimación de los parámetros. Como conclusión se ha definido un método de calibración que resulta efectivo y válido para la mayoría de los problemas de calibración que se puedan presentar en la actualidad. Para ello se han utilizado estudios previos presentes en la literatura, además de experiencias realizadas por el autor. Como conclusión de estudios previos realizados se puede decir que el método presentado por Zhang en [103] resulta el más eficiente a la hora de calibrar una cámara. Partiendo de éste método, se ha definido un preprocesado de la información y una mejora en el índice de coste que perfecciona la búsqueda no lineal. El preprocesado de los datos de la imagen consiste en corregir

errores en las medidas debidos a la distorsión que introducen las lentes, y también formalizar una normalización de los datos para condicionar mejor el proceso de cálculo del algoritmo. En la parte de cálculo no lineal de los parámetros se ha planteado también la búsqueda de las coordenadas de los puntos de la plantilla. Se puede decir que el caso de no realizar una búsqueda no lineal en los parámetros, el preprocesado de los datos permite obtener un conjunto de parámetros muy cercanos a los reales. Esto permite calibraciones ‘on line’ para utilizarlas por ejemplo en la localización de un vehículo mediante una cámara. Si se pretende ajustar más el resultado y se utilizan técnicas lineales de optimización, se concluye que el ruido en las medidas de las coordenadas en la plantilla no afecta al proceso de calibrado. Solamente el ruido en las coordenadas de la imagen perturba al proceso de calibración. Dado que se realiza un preprocesado de los datos para minimizar este ruido en la imagen, se puede decir que se consiguen resultados casi exactos en el proceso de calibración que sólo dependen del nivel de ruido de la imagen.

A partir de este método de calibración basado en plantilla bidimensional al que se le han introducido significantes mejoras, se pretende obtener los intervalos en los cuales están contenidos los parámetros estimados. A pesar de obtener resultados casi exactos, estos siempre serán imperfectos. Para calcular esta incertidumbre se han a utilizar técnicas de estimación estadísticas que tienen en cuenta los errores de los datos de partida en el proceso de cálculo. El uso de herramientas estadísticas para resolver el procedimiento de calibración de una cámara permite obtener mejores resultados además de caracterizar el procedimiento. Por un lado se obtienen mejores resultados ya que el algoritmo tiene en cuenta la información de acuerdo a la incertidumbre que ella misma presenta y además previamente se corrige la información de partida para eliminar dentro de lo posible los errores de medida. Las medidas de las coordenadas se corrigen de forma que cumplan la imagen de la recta y esto permite definir una incertidumbre de las coordenadas de los puntos si no lo hacen. Después, las medidas de las coordenadas se utilizan en la estimación de los parámetros de la cámara de acuerdo a su incertidumbre. Como consecuencia se obtienen un conjunto de parámetros más cercanos a los reales además de una incertidumbre de los mismos. Se han testeado dos formas de obtener los parámetros teniendo en cuenta la incertidumbre de los datos. Uno de ellos consiste en la búsqueda no lineal basada en incertidumbres de los parámetros de la cámara basada en el modelo de la misma. Como punto inicial de partida se toman los resultados obtenidos de forma lineal. El segundo método consiste en resolver el procedimiento de calibración basado en plantilla plana de Zhang [103], realizando una búsqueda no lineal basada en incertidumbres en cada una de las etapas. Como resultado se obtiene la incertidumbre de los parámetros para que sirva de dato de entrada en la siguiente etapa. Las gráficas obtenidas en el apartado de resultados experimentales del capítulo 4 demuestran que el segundo método obtiene mejores resultados así como menos incertidumbre en los parámetros. Dado que se plantea un índice algebraico sin significado físico, los acoplamientos que puedan existir en los parámetros de la cámara desaparecen y la estimación es más cercana a

la real. Por el contrario, el tiempo de cómputo necesario para resolver este proceso de calibración es muchísimo más largo.

La incertidumbre obtenida para los parámetros de la cámara permite utilizar posteriormente los resultados de la calibración de forma adecuada en caso de realizar reconstrucción 3D o utilizar una cámara para localizar un robot. Por otro lado, es posible caracterizar el procedimiento de calibración desde el punto de vista de conceptos tales como el límite de máximo de optimalidad que se puede alcanzar en la estimación a partir de una incertidumbre en las medidas de las coordenadas de los puntos. Este límite ha sido definido para cada una de las etapas del proceso de calibración. A partir de ahí es posible comparar algoritmos y decidir cual funciona de forma más eficiente. Se llega a la conclusión de que con ruido en las medidas de las coordenadas de la plantilla e imagen es imposible llegar a la solución exacta de los parámetros, sin embargo, si sólo se tiene ruido en las coordenadas de las imágenes, con un número de puntos considerable en la plantilla si que sería posible. Del mismo modo, dado que el límite de optimalidad depende tanto del número de puntos en la plantilla como del número de imágenes tomadas de la misma, se han definido valores en torno a los cuales los resultados están garantizados dentro de un nivel de incertidumbre. Como conclusión teórica interesante decir que el número de imágenes tomadas de la plantilla no influye en la mejora de los resultados de calibración si las coordenadas de los puntos de la plantilla no están contaminadas con ruido. Esta conclusión a nivel práctico se traduce en que por lo menos 8 imágenes son necesarias para que esto sea cierto.

En definitiva, se han descartado conceptos tales como que una mejor estimación lleva asociada un menor error residual, ya que como se comprueba a lo largo del capítulo 4 lo que parecía inicialmente una ventaja que era incluir los puntos en la plantilla dentro del índice de calibración resulta interesante si el nivel de ruido de éstas es considerable. En caso contrario, es mejor apostar por una diseño de la plantilla exacto que permita tener las coordenadas con un bajo nivel de ruido y a partir de ahí, realizar el proceso de calibración sin incluir el conjunto de puntos de la plantilla en el índice.

Para finalizar sólo queda definir las posiciones óptimas que mejor condicionan el proceso de calibración de la cámara. Partiendo de un criterio de optimización se han definido unas condiciones bajo las cuales el proceso de calibración de una cámara resulta robusto en la fase lineal de cálculo y elimina variables de búsqueda en su fase no lineal o iterativa. El criterio utilizado es el número de condición de la matriz formada por los elementos de las homografías llamada **V**. Las condiciones óptimas de calibración obtenidas han sido las posiciones desde las cuales se capturan las imágenes, ya que de ésta forma se consigue un proceso de cálculo más estable y robusto cuando se resuelve de forma lineal. Teniendo en cuenta que se establecen unas posiciones en las que se debe colocar la cámara para capturar las imágenes, si se tiene cuidado en alcanzar físicamente estas posiciones antes de tomar las imágenes, se pueden considerar como

válidas para el calibrado de la cámara. De esta forma, los parámetros considerados como válidos no se calculan en el proceso de búsqueda no lineal, se reduce el espacio de búsqueda y se consigue más acierto en los parámetros restantes buscados. Hay que tener en cuenta que cualquier proceso de cálculo de los parámetros de la cámara va a cometer errores, y por lo tanto es totalmente factible asumir ciertos parámetros como válidos si se ha tenido cuidado en ajustarlos previamente.

Las condiciones óptimas de calibración han quedado definidas respecto a un escenario de calibración general. En este escenario se coloca la plantilla de calibración en el plano X - Y considerado como suelo. La traslación negativa del eje Z del escenario es sobre el suelo. La cámara se coloca sobre el suelo y se orienta de forma que el eje óptico atraviesa el centro de la plantilla. De esta forma las posiciones óptimas para realizar la calibración es colocar la cámara de forma que una de las traslaciones respecto a los ejes del suelo sea nula en todas las posiciones. La otra traslación respecto al plano que simboliza el suelo ha de ser siempre la misma en valor absoluto. A partir de ésta, se calcula la altura de la cámara sobre el suelo que de forma aproximada es el doble del alejamiento de la cámara del origen de coordenadas. El ángulo que de forma óptima debe formar el plano de la imagen con el plano de la plantilla son 63.5° . Estas posiciones no garantizan plenamente que la matriz \mathbf{V} que se forma tiene un número de condición unitario. Existen varias de las condiciones de ortogonalidad e igualdad de módulo que no se satisfacen. Sin embargo si que dan unas pautas a seguir de cómo situar la cámara para realizar la calibración ya que mejoran considerablemente la condición de la matriz \mathbf{V} . Cualquier otra posición seguro que hace que se cumplan un menor número de restricciones nombradas en las expresiones de la (5.10) a la (5.21).

Es interesante resaltar que las posiciones óptimas no dependen en ningún momento de las características de la cámara. Esto permite poder utilizarlas con cualquier cámara sin tener que conocer previamente ninguno de sus parámetros. Las condiciones particulares de la cámara servirán para realizar el encuadre de la plantilla de forma que ocupe la mayor parte de la imagen.

6.1. Aportaciones realizadas

Las aportaciones realizadas en esta tesis se pueden considerar las siguientes:

- Definición de un método de calibrado de cámara que mejores resultados obtiene basado en el estado del arte, tanto de los métodos de calibración como de las técnicas de preprocesado de datos.
- Mejora de la técnica de preprocesado de los datos, obteniendo una incertidumbre asociada de forma particular a cada uno de ellos.

- Mejora del proceso de calibración ya que éste utiliza la incertidumbre de los datos de partida para optimizar la estimación de los parámetros.
- Cálculo de una incertidumbre asociada a los parámetros calculados la cual no esté sobredimensionada.
- Caracterización del proceso de calibración para definir que parámetros tales como número de imágenes o número de puntos intervienen en la incertidumbre de los resultados.
- Definición de unas posiciones y orientaciones óptimas que permiten condicionar el proceso de cálculo y mejorar los resultados.

6.2. Trabajo futuro

Las posibles líneas sobre las cuales seguir esta investigación son las siguientes

- Realizar la corrección de los puntos en la imagen teniendo en cuenta que el punto pertenece a dos rectas en el caso de tomar como plantilla un tablero de ajedrez.
- Basar la corrección y estimación de las rectas teniendo en cuenta que las rectas paralelas se juntan en el infinito en el llamado “vanishing points”.
- Estudiar posiciones óptimas para tomar las imágenes basándose en los acoplamientos que puedan existir en los parámetros. Tomar como base la matriz de varianzas / covarianzas de los parámetros estimados.
- Dado que existen vectores de la matriz V que no acaban siendo ortogonales, es porque no existe información que cubra esas dimensiones del espacio. Obtener la información referente a las direcciones espaciales que quedan por cubrir para conseguir un número de condición unitario de la matriz V .
- Determinar a que parámetros de la cámara afecta la ausencia de información en las direcciones espaciales que quedan por cubrir.
- Ver como afectan las perturbaciones a las condiciones óptimas de calibración
- Si se definen unas condiciones de calibración que se cumplen físicamente, se tiene unos parámetros que son conocidos. Por ejemplo si $t_y=0$, $h_{23}=0$. ¿Cómo afecta este dato en el caso del cálculo de las homografías?
- Definir una plantilla de calibración para condicionar de forma óptima el cálculo de las homografías.

BIBLIOGRAFÍA

- [1] Abdel-Aziz Y.I., Karara H. M. (1971) Direct linear transformation into object space coordinates in close-range photogrammetry. Proceedings symposium on close-range photogrammetry, Urbana, Illinois, p. 1-18
- [2] Ahlers R., Lu J. (1989) Stereoscopic vision – an application oriented overview, SPIE-Opt. Illumination, Image sensing Mach. Vision IV, p. 298-307
- [3] Ahmed M., Farag A. (2001) Non metric calibration of camera lens distortion. Proceedings of the International Conference on Image Processing. Grecia. pp. 157-160
- [4] Ahmed M., Farag A. (2005) Non metric calibration of camera lens distortion: Differential methods and robust estimation. IEEE Transactions on image processing. Vol. 14, n° 8. pp. 1215-1230
- [5] Aloimonos J., Shulman D. (1993) Active perception. Lawrence Erlbaum Associates, Hillsdale, N.J., USA
- [6] Ballard H., Brown C. (1982) Computer Vision. Prentice-Hall. Englewood Cliff, N.J., USA
- [7] Batista J., Araujo J., de Almeida A. (1999). Iterative multistep explicit camera calibration. IEEE International Journal on Robotics and Automation 15, p. 897-916
- [8] Batlle J., Mouaddib E., Salvi J. (1998). A survey: recent progress in coded structured light as a technique to solve the correspondence problem. Int. J. Pattern Recognition 31, p. 963-982

- [9] Broggi A. (1998). Vision-based driving assistance in vehicles of the future. *IEEE Intelligent Systems* 13 (6) pp.22-23.
- [10] Bougnoux S. (1998). From projective to Euclidean space under any practical situation, a criticism of self-calibration. *Proceedings of the 6th International Conference on Computer Vision*, pp 790-796.
- [11] Bowker A. H. (1981) *Estadística para ingenieros*. Prentice-Hall International
- [12] Brown D. C. (1966) Descentering distortion of lenses. *Photogrammetric Engineering and Remote Sensing*
- [13] Brown D. C. (1971) Close-Range camera calibration. *Photogrammetric Engineering and Remote Sensing* 37 (8), pp 855-866
- [14] Caprile A., Torre V. (1990). Using vanishing points for camera calibration. *International Journal of Computer Vision* pp 127-140
- [15] Casals A. (1989). *Sensor devices and systems for robotics*. Vol. 52, NATO ASI Series, Springer, Berlin, Heidelberg.
- [16] Charvonnier L., Fournier A. (1995) Heading guidance and obstacles localization for an indoor mobile robot. *IEEE International Conference on Advanced Robotics*, pp. 507-513
- [17] Chen L., Tsai W. (1990) A systematic approach to analytic determination of camera parameters by line features. *Pattern Recognition* 23 pp. 859-877
- [18] Criminisi A. (1999) *Accurate visual metrology from single and multiple uncalibrated images*. Dr. Ing. Tesis Department of engineering science. University of Oxford
- [19] Devernay F., Faugeras O. (2001) Straight lines have to be straight. *Machine Vision Applications* 13 (1). p. 14-24
- [20] Echigo T. (1975). A camera calibration technique using three sets of parallel lines. *Machine Vision Applications* 3, pp 159-167
- [21] Faig W., (1975) Calibration of close-range photogrammetry systems: Mathematical formulation. *Photogrammetric engineering and remote sensing*, Vol 41, p 1479-1486
- [22] Faugeras O., Toscani G. (1986) The calibration problem for stereo. *Proceedings Conference on Computer Vision and Pattern Recognition*, 1986, p. 15-20
- [23] Faugeras O., Luong T., Maybank S. (1992). Camera self-calibration: theory and experiments. *European Conference on Computer Vision*, p. 321-334

- [24] Faugeras O. (1993) Three dimensional computer vision: A geometric viewpoint. The MIT press, Cambridge.
- [25] Förstner W. (1987) Reliability analysis of parameter estimation in linear models with applications to measurement problems in computer vision. *Computer Vision, Graphics and image processing* 40, pp 273-310
- [26] Fusiello A. (2000). Uncalibrated euclidean reconstruction: a review. *Image and Vision Computing* 18, p. 555-263
- [27] García J., López M.. (1992) *Álgebra lineal y geometría*. Marfil.
- [28] Golub G.H., Van Loan C.F., (1996) *Matrix computation*, third edition. The Jonh Hopkins University Press
- [29] Hall E., Tio J. (1982). Measuring curved surfaces for robot vision. *Computing Journal* 15. pp 42-54
- [30] Haralick R., Shapiro L. (1993). *Computer and Robot Vision*. Vol. 2, Addison-Wesley Publishing Company, Reading.
- [31] Hartley R. (1993), Euclidean reconstruction from uncalibrated views. Second European Workshop on Applications of Invariance in Computer Vision. Pp 237-257
- [32] Hartley R. (1997) In defence of the eight point algorithm. *IEEE Transactions on pattern analysis and machine intelligence*. Vol 19 – 6
- [33] Hartley R., Zisserman A. (2000). *Multiple views geometry in computer vision*. Cambridge University Press
- [34] Heikkilä J., Silven O. (1996) Calibration procedure for short focal length off-the-shelf CCD cameras. *Proceedings of 13th International Conference on Pattern Recognition*. pp 166-170
- [35] Heikkilä J., Silven O. (1997) A four step camera calibration procedure with implicit image correction. *Proceedings of 14th International Conference on Pattern Recognition*
- [36] Heikkilä J. (2000). Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, pp 1066-1077
- [37] Hong Z., Yang J. (1993) An algorithm for camera calibration using a three-dimensional reference point. *Pattern Recognition* 26 pp 1655-1660.
- [38] Horaud R., Monga O. (1995) *Vision par ordinateur*. Hermes.

- [39] Horn B. (1986) Robot Vision. MIT Press. Cambridge.
- [40] Huynh D., Hartley R., Heyden A. (2003) Outlier correction in image sequences for the affine camera. Proceedings of the 9th IEEE International Conference on Computer Vision.
- [41] Ito J. (1991). Robot vision modelling – camera modelling and camera calibration. Advances in Robotics 5. pp 321-335
- [42] Isern J. (2003) Estudio experimental de métodos de calibración y autocalibración de cámaras. Dr. Ing. Tesis. Departamento de Informática y Sistemas. Universidad de las Palmas de Gran Canaria.
- [43] Jarvis R. (1983) A perspective on range finding techniques for computer vision. IEEE Transactions on Pattern Analysis and Machine Intelligence. Pp 122-139.
- [44] Kamata S., Eason M., Tsuji M., Kawaguchi E. (1992). A camera calibration using 4 points tarjets. International Conference on Pettern Recognition 1 pp.550-553
- [45] Kanade T., Rander P., Narayanan P. (1997). Virtualized reality: cosntructing virtual worlds from real scenes. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Multimedia, Immersive Telepresence pp.34-47
- [46] Kanatani K. (1995). Statistical Optimization for Geometric Computation. Theory and Practice. Dover
- [47] Kanatani K., Maruyama T. (1996). Optimal pattern for focal lenght calibration. Advanced Robotics, vol 10, No 1, pp.81-103
- [48] Khadraoui D., Motyl G., Martinet P., Gallice J., Chaumette F. (1996). Visual servoing in robotics scheme using a camera/laser-stripe sensor. IEEE International Journal Robotics and Atoimation 12, pp.743-750
- [49] Kopparapu S., Corke P. (2001). The effect of noise on camera calibration parameters. Graphical Models 63, p 277-303
- [50] Lavest J., Viala M. Dhome M. (1998) Do we really need accurate calibration pattern to achieve a reliable camera calibration. European Conference on Computer Vision, vol. I, pp 158-174
- [51] Lai J. (1993). On the sensitivity of camera calibration. Journal on Image Vision and Computing. pp 656-664
- [52] Lenz R., Tsai R. (1988) Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology. IEEE Transaction on Pattern Analysis and Machine Intelligence 10. pp 713-720

- [53] Lenz R., Tsai R. (1989) Calibrating a Cartesian robot with eye-on-hand configuration independent of eye-to-hand relationship. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 11. pp 916-928
- [54] Li M. (1994). Camera calibration of a head-eye system for active vision. *European Conference on Computer Vision*. pp 543-554
- [55] Li M., Lavest J. (1996). Some Aspects of Zoom-Lens camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (11).
- [56] Liebowitz D., Zisserman A. (1998). Metric rectification for perspective images of planes. *IEEE Conference on Computer Vision and Pattern Recognition*, vol. I. pp.482-488
- [57] Liu Y., Huang T., Faugeras O. (1990). Determination of camera location from 2D to 3D line and point correspondences. *IEEE International Journal on Robotics and Automation* 8. pp 42-54
- [58] Lucchese L. (2005) Geometric camera calibration of digital cameras through multi-view rectification. *Image and Vision Computing*, Vol. 23, No. 5, pp. 517-539.
- [59] Luong Q., Faugeras O., Maybank S. (1992) Camera self-calibration: theory and experiments. *Proceeding of the European Conference on Computer Vision*, pp 321-334.
- [60] Luxen M., Förstner W. (2001). Optimal camera orientation from points and straight lines. *Lecture Notes in Computer Science*, vol 2191, pp.84-92
- [61] *Manual of Photogrammetry*. American Society of Photogrammetry (1980) 4th edition.
- [62] Maybank S., Faugeras O. (1992) A theory of self-calibration of a moving camera. *International Journal of Computer Vision* 8(2), pp 123-151.
- [63] Melen T. (1994) Geometrical modelling and calibration of video cameras for underwater navigation. Dr. Ing. Tesis. Institutt for teknisk kybernetikk trondheim.
- [64] Meng X., Li H., Hu Z. (2000) A new easy camera calibration technique based on circular points. *The British Machine Vision Conference*.
- [65] Mohr R., Triggs B. (1996) *Projective geometry for image analysis*. ISPRS.
- [66] Naftel A., Boot J. (1991) An iterative linear transformation algorithm for solution of the collinearity equations. *American society for photogrammetry and remote sensing* Vol. 57 No 7 pp 913-919

- [67] Newman T., (1995) A survey of automated visual inspection. *Image Understanding* 61, pp 231-262
- [68] Nevatia R. (1982) *Machine perception*. Prentice Hall, Englewood Cliffs, N.J., USA
- [69] Noble B, Daniel J. (1989) *Algebra Lineal Aplicada*. Prentice-Hall.
- [70] Penna M. (1991) Camera calibration: a quick and easy way to detection of scale factor. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12. pp 1240-1245
- [71] Prescott B., McLean G. (1997) Line-based correction of radial lens distortion. *Graphical Models and Image Processing* 59 (1) p 39-47
- [72] Puget P., Skordas T., (1990) An Optimal solution for mobile camera calibration. *ECCV European conference computer vision*
- [73] Raskar R., Welch G., Cutts M., Lake A., Stesin L., Fuchs H. (1998). The office of the future: a unified approach to image-based modelling and spatial immersive displays. *SIGGRAPH 98*, pp 179-188
- [74] Salvi J. (1997) An approach to coded light to obtain three dimensional information. Ph.D. Thesis Universitat de Girona, Departament d'Electrònica, Informàtica I Automàtica.
- [75] Salvi J., Battle J., Mouraddib E. (1998). A robust-coded pattern projection for dynamic 3D scene measurement. *International Journal of Pattern Recognition Lett.* 19, pp 1055-1065
- [76] Salvi J., Armangué X., Batlle J. (2002) A Comparative review of camera calibrating methods with accuracy evaluation. *Pattern recognition* Vol. 35, pp 1617-1635
- [77] Sampson P.D. (1982) Fitting conic sections to 'very scattered' data: An iterative refinement of the Bookstein algorithm. *Computer Vision, graphics and image processing*.
- [78] Samtaney R., (1999) A method to solve an exterior camera calibration parameters for image resection. Technical report NAS-99-003 MRJ Technology Solutions Inc
- [79] Stama C., Theurer C. Henriksen (1980). *Manual of Photogrammetry*, 4th Edition, American Society of Photogrammetry. Falls church.
- [80] Stoer C., Bulirsch R. (1980). *Introduction to Numerical Analysis*. Springer, Berlin.

- [81] Sturm P., Maybank S. (1999). On plane-based camera calibration: a general algorithm, singularities and applications. *IEEE Conference on Computer Vision and Pattern Recognition* pp 432-437.
- [82] Sun W., Cooperstock J. (2004). Requiriments for camera calibration: Must accuracy come with a high price?. *IEEE Workshop on Applications of Computer Vision*.
- [83] Swaminathan R., Nayar S. (2002). Non metric calibration of wide-angle lenses and polycameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (10) p 1172-178.
- [84] Seedahmed G.H., Schenk T. (2002) Retrieval of the calibration matrix from 3D projective camera model. 15th International conference on vision interface
- [85] Shih S., Hung Y., Lin W. (1995) When should we consider lens distortion in camera calibration. *Pattern recognition* Vol. 28, No.3 pp 447-461
- [86] Stein G. (1997) Lens distortion calibration using point correspondences. *Proceedings of the International Conference on Computer Vision and Pattern Recognition* p 602-608
- [87] Stewart G. (1990) Perturbation theory for the singular value. Tech. Report TR90 -124, University of Maryland.
- [88] Taha H.A. (1987) Operation research. An introduction. Prentice Hall International Editions.
- [89] Tarel J.P. (1994) Calibration de caméra fondée sur les ellipses. Research report 2200, INRIA
- [90] Toscani G. (1987) Systemès de calibration et perception du movement en vision artificielle. Ph. D. Thesis Université Paris Sud.
- [91] Triggs B. (1998). Autocalibration from planar scenes. *European Conferrence on Computer Vision*, pp 89-105
- [92] Trucco E., Verri A. (1998) Introductory techniques for 3-D computer vision. Prentice Hall
- [93] Tsai R. (1986) An efficient an accurate camera calibration technique for 3D machine vision. *IEEE computer vision and pattern recognition*, p. 364-374
- [94] Tsai R., (1997) A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-self TV camera lenses. *IEEE Journal of Robotics and Automation*, Vol RA-3 No 4. pp. 323-344
- [95] Wang C., Tsai W. (1991) Camera calibration by vanishing lines for 3D computer

- vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 13. pp 370-376
- [96] Wang C. (1992) Extrinsic calibration of a vision sensor mounted on a robot. IEEE International Journal on Robotics and Automation 8. pp 161-175
- [97] Wei G., de Ma S. (1993). A complete two-plane camera calibration method and experimental comparison. Proceedings of 4th International Conference on Computer Vision pp. 439-446
- [98] Wei J., de Ma S. (1994). Implicit and explicit camera calibration: Theory and experiments. IEEE Transactions on Pattern Analysis and machine intelligence. Vol 16 p. 15-20
- [99] Weng J., Huang T.S., Ahuja N. (1989) Motion and structure from two perspective views: algorithms, error analysis and error estimation. IEEE transactions on pattern analysis and machine intelligence. Vol 11 p. 451-475
- [100] Weng J., Cohen P., Herniou M. (1992) Camera calibration with distortion models and accuracy evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (10) pp 965-980
- [101] Wilson R. (1994) Modelling and calibration of automated zoom lenses. PhD. dissertation. Department electronic computing engineering, Carnegie Mellon University.
- [102] Zhang Z. (1993) The matching problem: the state of the art. Technical report No. 2146, Institute National de Recherche en Informatique et en Automatique.
- [103] Zhang Z. (1998) A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research.
- [104] Zhang Z. (2000) A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, pp 1330-1334.
- [105] Zhang Z. (2002) Camera calibration with one-dimensional objects. Technical Report MSR-TR-2001-120 Microsoft research.