

JUAN CAMILO PÉREZ MUÑOZ

Maestría en Matemáticas

Directores:

Carlos Enrique Mejía Salazar

Andrés Fernando Osorio Arias

Escuela de Matemáticas

Facultad de Ciencias

Medellín

2009



OPTIMIZACIÓN NO LINEAL Y CALIBRACIÓN DE CÁMARAS FOTOGRAFICAS



OPTIMIZACIÓN NO LINEAL Y CALIBRACIÓN DE CÁMARAS FOTOGRÁFICAS

Autor: Juan Camilo Pérez Muñoz
Directores: Carlos Enrique Mejía Salazar
Andrés Fernando Osorio Arias

Trabajo de Grado Maestría en Matemáticas
Facultad de Ciencias
Universidad Nacional de Colombia
Sede Medellín
2009-I

Palabras Claves: Optimización, mínimos cuadrados, modelo Pinhole, Levenberg-Marquardt, estimación de parámetros.

Resumen

En este trabajo se estudian varios métodos de optimización no lineal con el ánimo de calibrar cámaras fotográficas usando el modelo Pinhole. Para tal fin se presentan las generalidades de los métodos de Gauss-Newton, del Gradiente y de Levenberg-Marquardt, escogiéndose este último como el método a usar en la resolución de los problemas asociados a la estimación de los parámetros del modelo Pinhole.

Para ayudar a entender al lector el modelo Pinhole, se presenta la construcción del mismo y se explica de manera detallada los efectos de distorsión en la imagen causados por los lentes. También se presentan las diversas estrategias encontradas en la literatura especializada para estimar todos o algunos de los parámetros de dicho modelo y se analiza la tolerancia a la incertidumbre y la precisión de los mismos, lo que permite concluir que los métodos no lineales basados en la solución de problemas de mínimos cuadrados son los métodos más apropiados cuando se exige un alto grado de precisión.

En este trabajo se comparan los resultados obtenidos con resultados reportados en la literatura y se presentan los resultados prácticos de calibrar cámaras instaladas en campo y se comparan los resultados al omitir e incluir los parámetros de distorsión radial en el proceso de calibración con puntos de control medidos en campañas topográficas.

Key Words: Optimization, least squares, Pinhole model, Levenberg-Marquardt, parameter estimation.

Abstract

Pointing to calibrate photographic cameras using Pinhole model, several non linear optimization methods are studied in this work. Generalities of Gauss-Newton, Gradient and Levenberg-Marquardt methods are presented, and the last one is chosen as the method used to solve the problems that arise when the Pinhole model parameter are estimated.

Whit the aim to help the reader to understand the Pinhole model, its construction process is presented and the distortion effects caused by the lens are explained in a detailed way. Also, several strategies used in the literature to estimate some or all the parameters of the model are presented and their precision and uncertainty tolerance are analyzed, then it's concluded that non linear methods based on solving least squares problems are the better choice when high precision is needed.

Results of this work are compared whit another results presented in literature and the practical results obtained when field installed cameras were calibrated are presented and compared when radial distortion parameters are omitted and included.

Optimización No Lineal y Calibración de Cámaras Fotográficas

por

Juan Camilo Pérez Muñoz



Trabajo presentado como requisito parcial
para optar al Título de

Magister en Matemáticas

Director: Carlos Enrique Mejía Salazar
Co-director: Andrés Fernando Osorio Arias

Universidad Nacional de Colombia
Sede Medellín

Facultad de Ciencias

Escuela de Matemáticas

Julio 2009

Este trabajo ha desarrollado dentro del proyecto HORUS y ha sido financiado por la Agencia Española de Cooperación Internacional y Desarrollo, AECID.

Resumen

Para realizar mediciones de distancias en imágenes fotográficas, los investigadores deben compensar los efectos causados por la perspectiva y por la distorsión asociada a los lentes de la cámara. Entre la comunidad científica goza de gran aceptación el uso del modelo Pinhole para modelar las cámaras fotográficas, el cual se basa en el principio de colinealidad para describir la relación entre un punto en el espacio (x, y, z) y su respectiva proyección en la imagen (u, v) , este modelo puede escribirse de la siguiente forma:

$$[wu, wv, w]^T = K [R | -RC] [x, y, z, 1]^T$$

Donde K es una matriz de Parámetros internos que dependen de la configuración de la cámara y los lentes, R es una matriz de rotación y C es un vector que indica la posición del centro óptico de la imagen, el cual es el punto donde se cortan todas las proyecciones bajo el principio de colinealidad.

Para encontrar los parámetros del modelo Pinhole correspondientes a una cámara con una configuración dada, se usa un conjunto de n puntos de control, es decir, puntos cuyas coordenadas en el espacio son conocidas al igual que sus coordenadas en la imagen. Con estos puntos de control es posible construir un vector de residuales

$$R\epsilon(\beta) = [u_1 - \hat{u}_1, v_1 - \hat{v}_1, \dots, u_n - \hat{u}_n, v_n - \hat{v}_n]^T$$

Donde (\hat{u}_n, \hat{v}_n) es la reproyección del punto de control con coordenadas (x_n, y_n, z_n) usando el modelo Pinhole con los parámetros dados por el vector β . El vector $R\epsilon(\beta)$ se conoce como vector de residuales y para obtener los estimados de los parámetros se propone minimizar el funcional $f(\beta) = R\epsilon(\beta)^T R\epsilon(\beta)$, el cual corresponde a la suma de los cuadrados de los residuales. Para minimizar dicho funcional se emplea un método iterativo y en éste trabajo se propone usar el método de Levenberg-Marquardt, el cual explota las ventajas del método de Gauss-Newton y el método del Gradiente.

Para ayudar a entender al lector el modelo Pinhole, se presenta la construcción del mismo y se explica de manera detallada los efectos de distorsión en la imagen causados por los lentes. También se presentan las diversas estrategias encontradas en la literatura

especializada para estimar todos o algunos de los parámetros de dicho modelo y se analiza la tolerancia a la incertidumbre y la precisión de los mismos, lo que permite concluir que los métodos no lineales basados en la solución de problemas de mínimos cuadrados son los métodos más apropiados cuando se exige un alto grado de precisión.

En este trabajo se presentan los resultados prácticos de calibrar cámaras instaladas en campo y se comparan los resultados al incluir los parámetros de distorsión radial en el proceso de calibración con puntos de control medidos en campañas topográficas. También se presentan los resultados obtenidos al estimar los parámetros del modelo Pinhole usando el método de Levenberg-Marquardt con datos sintéticos y se comparan estos resultados con otros resultados reportados en la literatura, mostrando mejoras importantes en dichos resultados al usar el método de Levenberg-Marquardt como mecanismo para resolver el problema de mínimos cuadrados asociado a la estimación de los parámetros.

Es importante resaltar que diversos autores han resaltado el uso del modelo Pinhole como herramienta básica para el monitoreo de variables ambientales en zonas costeras por medio de cámaras de video y que éste trabajo ha sido financiado por la Agencia Española de Cooperación Internacional y Desarrollo como parte del proyecto "Aplicación de Metodologías y técnicas basadas en sistemas de video para el seguimiento de los problemas costeros: Caso Cartagena, Colombia".

Contenido

Introducción	xiv
1 Problemas de Mínimos Cuadrados No lineales	1
1.1 Mínimos Cuadrados no lineales	1
1.2 Método de Gauss-Newton	4
1.3 Método del Gradiente	6
1.4 Método de Levenberg-Marquardt	7
1.5 Convergencia y velocidad de convergencia del método de Levenberg-Marquardt	16
2 Modelo Pinhole	22
2.1 Principios del Modelo Pinhole	23
2.2 Modelo Pinhole con distorsión	31
3 Calibración del modelo Pinhole en 2 etapas	37
3.1 Estimación de los parámetros internos y coeficientes de distorsión	38
3.1.1 Ecuaciones básicas	38
3.1.2 Estimación de los parámetros intrínsecos en forma cerrada	40
3.1.3 Estimación de los parámetros de distorsión radial en forma cerrada .	47
3.1.4 Estimación de los parámetros intrínsecos usando mínimos cuadrados	49
3.1.5 Matrices antisimétricas y de rotación, fórmula de Rodrigues	50

3.2	Corrección de la distorsión	52
3.2.1	Compensando la distorsión radial	53
3.2.2	Eliminación gráfica de la distorsión	55
3.3	Estimación de los parámetros extrínsecos	59
3.3.1	Solución geométrica: Problema PnP	60
3.3.2	Estimación de los parámetros extrínsecos con mínimos cuadrados . .	64
4	Calibración del modelo Pinhole completo	68
4.1	Transformada lineal Directa, DLT	68
4.1.1	Soluciones no homogéneas para H	71
4.1.2	Solución de la DLT con puntos en el espacio y su correspondencias en la imagen	72
4.1.3	Consideraciones sobre el rango de la matriz A en el caso 3D a 2D . .	73
4.1.4	Obteniendo los parámetros intrínsecos y extrínsecos a partir de la DLT	74
4.1.5	La transformación afín y su aplicación	77
4.2	Solución por mínimos cuadrados del modelo sin distorsión	79
4.3	Solución por mínimos cuadrados del modelo con distorsión	82
5	Experimentos y resultados	85
5.1	Cálculo de puntos de fuga	85
5.2	Estimación de los parámetros extrínsecos a partir de la DLT y la matriz de calibración interna	88
5.3	Estimación de los parámetros internos y de distorsión radial usando patrones de calibración	92
5.4	Estimación del modelo Pinhole sin distorsión	96
5.5	Estimación del modelo Pinhole con distorsión radial	109
5.6	Aplicación de la calibración del modelo Pinhole en la estación HORUS Carta- gena	113

6 Conclusiones	121
Bibliografía	126

Lista de Figuras

0-1	Foto aérea de la zona de Bocagrande, Cartagena de Indias. Foto tomada de GoogleEarth©.	xiv
0-2	Foto de la zona de Bocagrande, Cartagena, tomada con el sistema de captura HORUS.	xv
0-3	Imagen de una zona de interés en la región de Bocagrande, Cartagena, y su correspondiente rectificación usando el modelo Pinhole.	xvi
1-1	Función $F(x, y) = \cos(y - 2) (x - 1)^2 + (y - 2)^2$ y δ_t y δ_g en $(-1, 1)$	9
1-2	Función $F(x, y) = \cos(y - 2) (x - 1)^2 + (y - 2)^2$ y δ_t , δ_g y δ_{lm} en $(-1, 1)$	13
1-3	Función $F(x, y) = \cos^2(y - 2) (x - 1)^2 + (y - 2)^2$	17
1-4	Error cometido en cada iteración con los tres métodos de optimización, para la función de la ecuación (1.23).	18
1-5	Error de los métodos de optimización en la función de la ecuación (1.24). Con estimado inicial (1, 1) (Arriba) y con estimado inicial (3, 1) (Abajo).	21
2-1	Cámaras (Dentro de sus carcazas protectoras) usadas en la estación de captura HORUS en Cartagena y PC de control.	23
2-2	Cámaras STINGRAY y MARLIN, usadas en las estaciones de captura HORUS.	24
2-3	Esquema del Modelo Pinhole.	25
2-4	Proyección de la coordenada \widetilde{x} en el plano de la imagen, bajo el principio de colinealidad.	28

2-5	Plano de la imagen con sistema de coordenadas de rotación derecha.	28
2-6	Plano de la imagen con sistema de coordenadas de rotación izquierda.	29
2-7	Distintos tipos de lentes usados en las estaciones de captura HORUS.	32
2-8	Imagen con distorsión tomada con una cámara MARLIN y un lente de 8 mm.	33
2-9	Efecto de la distorsión radial. Línea continua: Sin distorsión. Líneas punteadas: Con distorsión radial.	34
3-1	Ejemplo de un punto de fuga en una imagen. Imagen del edificio Bavaria en Cartagena, donde se encuentra instalada una estación de captura HORUS.	43
3-2	Segmentos de línea usados para encontrar el punto de fuga (Líneas azules), punto de fuga encontrado y conjunto de líneas estimadas (Líneas rojas).	44
3-3	Puntos libres de distorsión en una imagen, puntos distorsionados y puntos con compensación de la distorsión obtenidos con la estrategia propuesta.	56
3-4	Proceso de creación de la imagen sin distorsión.	58
3-5	Píxeles alrededor de la coordenada (u, v) calculada.	59
3-6	Imagen sin distorsión obtenida a partir de la imagen de la figura 2-8.	60
3-7	Información disponible para resolver la matriz de orientación externa con 3 GCP.	61
3-8	Esquema del problema geométrico a resolver en el caso P3P.	62
3-9	Configuración de las regiones con solución única del problema P3P dadas en [Wang94] y [Xiao03].	65
4-1	Dos imágenes relacionadas por una transformación afín H , y su fusión.	79
4-2	Panorámica de Bocagrande, imagen fusionada de 3 cámaras con lente de 8 mm de la estación de captura HORUS instalada en Cartagena de Indias, Colombia.	80

4-3	Panorámica de Bocagrande, imagen fusionada de 3 cámaras con lente de 4.8 mm de la estación de captura HORUS instalada en Cartagena de Indias, Colombia.	80
5-1	Patrón de calibración y punto de fuga determinado usando el método de Levenberg-Marquardt.	87
5-2	Patrón de calibración y puntos usados para calcular la DLT de la imagen.	89
5-3	Interfaz gráfica de calibración de lentes HORUS.	94
5-4	Evolución del parámetro λ (Arriba) y del error de reproyección $F(\beta)^T F(\beta)$ usando el método de Levenberg-Marquardt.	101
5-5	Variación de los parámetros intrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 7 puntos de control.	102
5-6	Variación de los parámetros extrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 7 puntos de control.	103
5-7	Evolución de los errores de la calibración usando el método lineal.	105
5-8	Evolución de los errores de la calibración usando el método de Levenberg-Marquardt.	105
5-9	Variación de los parámetros intrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 43 puntos de control.	106
5-10	Variación de los parámetros extrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 43 puntos de control.	107
5-11	Evolución de los parámetros intrínsecos y de distorsión a medida que aumenta la incertidumbre en las mediciones en la imagen.	113
5-12	Evolución de los parámetros extrínsecos a medida que aumenta la incertidumbre en las mediciones en la imagen.	114
5-13	Evolución de los errores de la calibración del modelo Pinhole a medida que la incertidumbre aumenta.	115

5-14 Foto de 1024 x 768 píxeles, correspondiente a la cámara 1 de la estación HORUS Cartagena y los puntos de control usados para calibrarla.	116
5-15 Imagen rectificada de la cámara 1 de la estación HORUS en Cartagena usando los parámetros del modelo Pinhole obtenidos a partir de la topografía.	120

Agradecimientos

Quiero agradecer a todas las personas que de forma directa o indirecta contribuyeron a desarrollar este trabajo, fueron personas de diversas áreas y variados contextos, a todos muchas gracias. Entre ellos vale la pena agradecer explícitamente a algunos:

A los profesores Carlos Enrique Mejía Salazar y Andrés Fernando Osorio Arias por permitirme realizar este proyecto y acompañarme en todo momento dándome asesoría y discutiendo conmigo los detalles que acá aparecen explicados.

A mis compañeros en el proyecto HORUS en general: Hugo Escorcía, David Gutiérrez, Mauricio Sedano, Imen Turki, Sebastián Osorio y Juan David Osorio, sus comentarios y pruebas de los algoritmos desarrollados fueron de gran utilidad.

A Cristian Andrés Ortiz Alarcón por su ayuda incondicional, por todo el tiempo que invirtió discutiendo temas asociados a esta tesis, al proyecto HORUS y a miles de problemas presentados durante este tiempo.

Al grupo de Oceanografía e Ingeniería Costera de la Universidad Nacional, OCEANICOS, de la Facultad de Minas.

A Pedro Isaza, Jorge Mejía, y todos los profesores de la maestría en Matemáticas de la Universidad Nacional de Colombia Sede Medellín.

A mis profesores y compañeros de Judo, especialmente Juan Carlos Arboleda, Hernán Darío Arboleda y Juan David Ospina, el tiempo que invertí con ellos sirvió de válvula de escape cuando las cosas se ponían difíciles y su comprensión cuando me ausenté para trabajar en esta tesis y en mis proyectos académicos fue muy valiosa.

A mis compañeros de maestría, me alegra mucho poder decir que compartí este tiempo con ellos. Por las ayudas, por las veladas de estudio y por aquellas veladas menos académicas.

A Natalia Martínez Bilbao por diseñar la portada de ésta tesis.

A mi familia, su apoyo y buena energía durante todo mi pregrado y esta maestría han sido combustible para poner a andar mis proyectos.

A Laura Restrepo Mazo, su compañía me dió el empuje necesario para culminar con ánimo esta etapa y deseo de todo corazón que esa compañía perdure mucho tiempo.

Introducción

Es posible que el lector haya escuchado en algún momento la frase "Una imagen vale más que mil palabras" y en el caso de imágenes digitales dicha frase tiene mucha aplicación cuando éstas se usan para medir algún tipo de variable sobre ellas. Para poder medir distancias y/o áreas en una imagen es necesario tener una resolución constante, es decir, una distancia en píxeles medida sobre la imagen debe ser equivalente a la misma distancia en el espacio sin importar en que región de la imagen se mida. Éste hecho se garantiza cuando la cámara se encuentra perpendicular a la región sobre la que se desea medir, como es el caso de la fotografía aérea y satelital. En la figura 0-1 se puede apreciar un ejemplo de una foto satelital donde la distancia marcada con amarillo corresponde a 500 m de longitud.



Figura 0-1: Foto aérea de la zona de Bocagrande, Cartagena de Indias. Foto tomada de GoogleEarth©.

En el caso de los sistemas de monitoreo ambiental basados en cámaras de video (Ver [Osorio07] y [Ortiz08]) la idea es tomar fotos de una zona con períodos que van desde

minutos hasta horas, por lo que la fotografía aérea y la fotografía satelital deben descartarse como soluciones para dichos sistemas, dado que la primera resulta ser costosa a tan altas frecuencias y la segunda no permite períodos tan cortos entre una imagen y otra. Para solucionar este problema, los investigadores usan cámaras ubicadas en una parte elevada que permita abarcar la mayor extensión de área, obteniendo imágenes a la frecuencia deseada pero las cuales ya no presentan resolución constante como la que se presenta en la figura 0-2, en la cual puede observarse como los elementos que están más cerca de la cámara se ven más grandes que los que se encuentran más alejados, por lo que la proporción entre distancias en la imagen y en el espacio ya no es constante. Vale la pena resaltar que la imagen de la figura 0-2 fue tomada con el sistema de captura **HORUS** (<http://www.horusvideo.com>) instalado en la ciudad de Cartagena como parte del proyecto "Aplicación de Metodologías y técnicas basadas en sistemas de video para el seguimiento de los problemas costeros: Caso Cartagena, Colombia" el cual es co-financiado por la Agencia Española de Cooperación Internacional y Desarrollo, AECID.

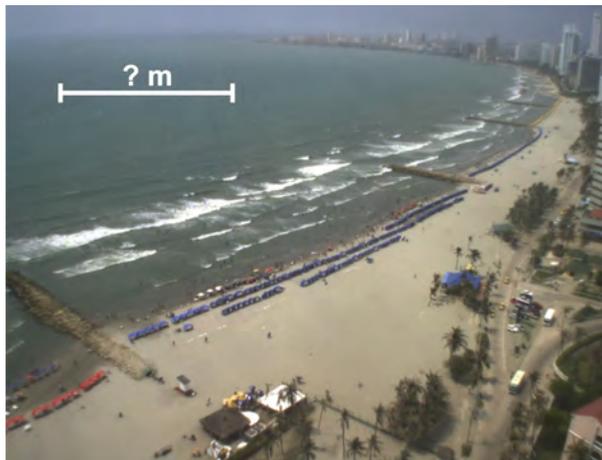


Figura 0-2: Foto de la zona de Bocagrande, Cartagena, tomada con el sistema de captura HORUS.

Como el lector ya se habrá dado cuenta, en los sistemas de monitoreo ambiental se sacrifica la resolución constante en las imágenes para poder tomar fotos con un período de tiempo corto, por lo que ya no es posible medir distancias o áreas directamente sobre la imagen. Para solucionar el problema de la ausencia de resolución constante, los investigadores recurren a usar un modelo de la cámara que relacione los puntos en un sistema coordinado en el espacio con su respectiva proyección en el plano de la imagen, entre los modelos usados el que más se destaca es el modelo Pinhole (Ver [Faugeras93], [Hartley03],

[Heikilla97], [Holland97], [Osorio07], [Ortiz08], entre otros). Dicho modelo permite realizar mediciones de distancias sobre la imagen e, incluso, construir fotos *rectificadas* de la misma, es decir, fotos que se aproximan a imágenes aéreas o a mapas de la región de estudio como la que se presenta en la figura 0-3.

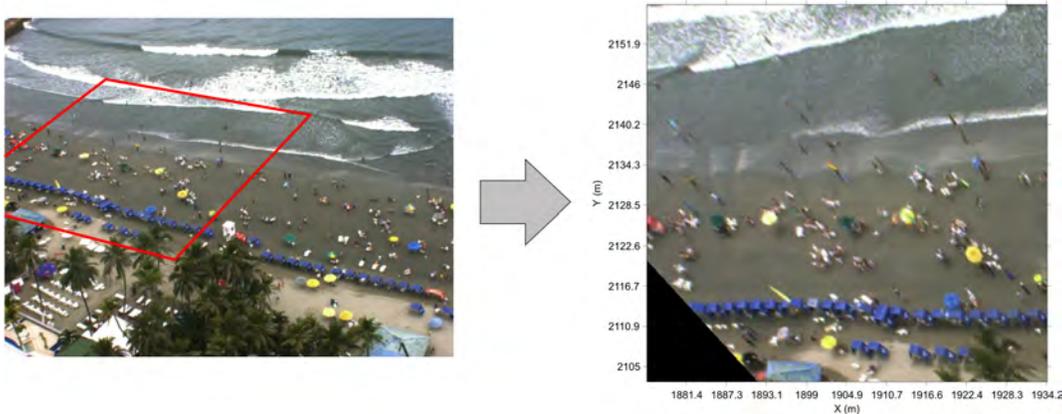


Figura 0-3: Imagen de una zona de interés en la región de Bocagrande, Cartagena, y su correspondiente rectificación usando el modelo Pinhole.

Para poder obtener la imagen rectificada de la figura 0-3 es necesario conocer todos los parámetros del modelo Pinhole, los cuales se pueden estimar usando métodos iterativos para minimizar algún funcional asociado al modelo (Ver [Hartley03], [Heikilla97], [Holland97] y [Salvi02], entre otros) o métodos basados en álgebra lineal (Ver [Faugeras93], [Osorio07] y [Zhang00]). Entre los métodos basados en álgebra lineal, se destaca el uso de la Transformada Lineal Directa (DLT por sus siglas en inglés), la cual es una aproximación matricial al modelo Pinhole que, además, permite obtener estimados de los parámetros del modelo bajo ciertas condiciones y son estos estimados el punto de partida para un método iterativo que permita obtener mejores aproximaciones de los parámetros (Ver [Salvi02], [Zhang00], [Zhang98]).

Varios autores han reportado el uso del modelo Pinhole como herramienta base para la medición de variables ambientales en zonas costeras, entre ellos [Holland97], [Osorio07] y [Ortiz08], lo que indica la importancia de estudiar dicho modelo y los métodos usados en la literatura para estimar sus parámetros, entre ellos los métodos de optimización basados en mínimos cuadrados. Debido a la importancia del modelo Pinhole como herramienta en el proceso de medición de variables ambientales usando imágenes fotográficas, el principal interés de este trabajo es estudiar las técnicas presentadas en la literatura para calibrar dicho modelo y, por tanto, los métodos de optimización necesarios para estimar los parámetros del modelo. Con base en dicho interés, en éste trabajo se plantean 3 objetivos principales:

1. Estudiar métodos de optimización no lineal basados en mínimos cuadrados, incluyendo en primer lugar el método de Gauss-Newton, con el ánimo de implementar un método que permita calibrar el modelo Pinhole.

2. Estudiar los métodos lineales presentados en la literatura para calibrar el modelo Pinhole y usar dichos métodos para obtener estimados iniciales para los métodos no lineales implementados en este trabajo.

3. Comparar el desempeño y los resultados obtenidos con los métodos implementados en este trabajo y otros métodos propuestos en la literatura.

Este trabajo se divide en 6 secciones, la primera es una presentación de los métodos de Gauss-Newton, del Gradiente y de Levenberg-Marquardt para la solución de problemas de mínimos cuadrados, dándole mayor trascendencia al último método debido a las ventajas que exhibe sobre los otros dos. En la segunda parte de este trabajo se presenta la construcción del modelo Pinhole con y sin distorsión, y se explica el significado físico de cada uno de sus parámetros. La tercera parte consiste en toda la teoría involucrada en la calibración del modelo Pinhole en dos etapas, haciendo énfasis en la solución de los parámetros de distorsión radial y la corrección de la misma. El cuarto capítulo presenta la teoría involucrada en la solución de la transformada lineal directa y el modelo Pinhole completo usando una sola imagen y métodos no lineales. La quinta parte presenta los experimentos y resultados que permiten validar la teoría sobre la solución del modelo Pinhole usando diversos métodos. Por último, el capítulo 6 presenta conclusiones importantes, resaltando los pequeños esfuerzos y aportes hechos en este trabajo.

Capítulo 1

Problemas de Mínimos Cuadrados No lineales

En este capítulo se estudiará la definición y clasificación de los problemas de mínimos cuadrados no lineales y algunos métodos de solución de dichos problemas, con los cuales se espera calibrar el modelo Pinhole. En primera instancia se presentará la construcción de la función a minimizar en dichos problemas, luego se introducen brevemente los métodos de Newton y del Gradiente, necesarios para entender las ideas sobre las que se basa el método de Levenberg-Marquardt y por último se presentan varios experimentos teóricos que permiten concluir sobre las ventajas que presenta el método de Levenberg-Marquardt.

1.1 Mínimos Cuadrados no lineales

Suponga que se desea ajustar un modelo f que tiene la forma:

$$y = f(x_1, x_2, \dots, x_m; \beta_1, \beta_2, \dots, \beta_k) = f(X, \beta) \quad (1.1)$$

Donde:

$X = (x_1, x_2, \dots, x_m)$ son variables independientes

$\beta = (\beta_1, \beta_2, \dots, \beta_k)$ son los parámetros del modelo

y es el valor observado, o medido, en el punto X

El ajuste del modelo f consiste en encontrar el conjunto de parámetros β tal que dado

un conjunto determinado de variables independientes, el valor entregado por el modelo sea lo más cercano posible al valor observado y . Para llevar este proceso a cabo, se supone que se tienen n muestras de la forma:

$$(y_i, X_i) = (y_i, x_{1i}, x_{2i}, \dots, x_{mi}) \text{ con } i = 1, 2, \dots, n$$

Entonces, el problema de calibrar el modelo de la ecuación (1.1) consiste en minimizar la función:

$$F(\beta) = \sum_{i=1}^n \|y_i - \hat{y}_i\|^2 \quad (1.2)$$

Donde \hat{y}_i es el valor de f usando el conjunto de parámetros β en el punto X_i . Otra forma de escribir este problema es definir el vector de residuales $\mathbf{Y} - \hat{\mathbf{Y}}(\beta)$ como $R_\epsilon(\beta)$:

$$R_\epsilon(\beta) = \mathbf{Y} - \hat{\mathbf{Y}}(\beta) = \begin{bmatrix} y_1 - \hat{y}_1 & y_2 - \hat{y}_2 & \dots & y_n - \hat{y}_n \end{bmatrix}^T \quad (1.3)$$

De este modo, la función $F(\beta)$ puede escribirse como

$$F(\beta) = \left[\mathbf{Y} - \hat{\mathbf{Y}}(\beta) \right]^T \left[\mathbf{Y} - \hat{\mathbf{Y}}(\beta) \right] = R_\epsilon(\beta)^T R_\epsilon(\beta) \quad (1.4)$$

En este caso, el Jacobiano de $R_\epsilon(\beta)$ se calcula como:

$$(R'_\epsilon(\beta))_{ij} = \left(\left(\mathbf{Y} - \hat{\mathbf{Y}}(\beta) \right)' \right)_{ij} = \frac{\partial (y_i - \hat{y}_i)}{\partial \beta_j} = -\frac{\partial f_i}{\partial \beta_j}, \quad 1 \leq i \leq n, \quad 1 \leq j \leq k \quad (1.5)$$

Mientras que el gradiente de $F(\beta)$ puede obtenerse con la expresión

$$\nabla F(\beta) = 2R'_\epsilon(\beta)^T R_\epsilon(\beta) \in \mathbb{R}^n \quad (1.6)$$

En [Kelley99] se presenta la clasificación para los problemas de mínimos cuadrados dependiendo del número de parámetros respecto al número de observaciones así:

- $n > k$: Problema sobredeterminado, este es el caso más común en la estimación de parámetros.
- $n = k$: Sistema de ecuaciones no lineal, en este caso la teoría y métodos presentados en [Kelley95] pueden ser usados.
- $n < k$: Problema subdeterminado, son los problemas menos comunes, pero pueden encontrarse en la solución de ecuaciones algebraicas de alto orden. También tienen el inconveniente que el límite de la iteración de Gauss-Newton no está determinado de forma única por la distancia del iterado inicial del punto donde $F(\beta^*) = 0$.

Como en los problemas de mínimos cuadrados lo que se desea es obtener un vector de residuales con norma cercana a 0, es posible clasificar dichos problemas según el valor de F en el punto donde se encuentra el mínimo local que se está buscando. En [Kelley99] se presenta dicha clasificación: Si β^* es el punto donde se encuentra el mínimo local de F , se tiene que si $F(\beta^*) = 0$ el problema de hallar el mínimo local de F recibe el nombre de problema de *residual cero*, y este no es el evento esperado cuando el problema es sobredeterminado o los datos usados están sometidos a la existencia de errores de medición. Por ejemplo, cuando se calibra una cámara, las mediciones sobre la imagen están sometidas al error de la persona que las realiza y a que la imagen sea de suficiente calidad para garantizar errores pequeños, por lo que en éste caso no se espera tener un problema de *residual cero*. Cuando $F(\beta^*)$ está cerca de 0 (recordar que $F(\beta) \geq 0$ para todo β), el problema es llamado un problema de *residual pequeño* y este es el tipo de problemas que se espera encontrar en la calibración de cámaras y en cualquier otro proceso de estimación de parámetros cuando el modelo representa de manera acertada la relación entre los datos medidos. En cualquier otro caso, se tiene un problema de *residual grande*.

Ahora, si se denota por $\mathbf{Y} - \hat{\mathbf{Y}}(\beta)$ al vector de residuales: La función $F(\beta)$ puede escribirse como: Donde el Jacobiano de $\mathbf{Y} - \hat{\mathbf{Y}}(\beta)$ corresponde exactamente a: Y así es posible definir el gradiente de F como:

En general, los métodos de optimización de mínimos cuadrados, siguen el siguiente algoritmo:

Algoritmo 1.1.1. Se tiene una aproximación inicial, $\beta^{(0)}$. Para $r \geq 0$, hacer:

1. Se usa ∇F y/o $R'_\epsilon(\beta^{(r)})$ para calcular un vector de corrección $\delta^{(r)}$ tal que $F(\beta^{(r)} + \delta^{(r)}) < F(\beta^{(r)})$.
2. Se hace $\beta^{(r+1)} = \beta^{(r)} + \delta^{(r)}$.
3. Hacer $r = r + 1$.
4. Si se cumple un criterio de parada, terminar, si no, ir al paso 1.

1.2 Método de Gauss-Newton

En el método de Gauss-Newton ([Kelley99]) se usa la aproximación lineal de Taylor para \mathbf{Y} alrededor de β_c ([Marquardt63]):

$$\langle \mathbf{Y}(\beta) \rangle = \hat{\mathbf{Y}}(\beta_c) - R'_\epsilon(\beta_c)(\beta - \beta_c) \quad (1.7)$$

Donde los corchetes $\langle \rangle$ se usan para distinguir las aproximaciones basadas en el modelo linealizado de las realizadas con el modelo no lineal, $\hat{\mathbf{Y}}$. Entonces, la aproximación de F calculada con (1.7) es

$$\langle F(\beta) \rangle = \sum_{i=1}^n \|Y_i - \langle Y_i \rangle\|^2 = [\mathbf{Y} - \langle \mathbf{Y}(\beta) \rangle]^T [\mathbf{Y} - \langle \mathbf{Y}(\beta) \rangle] = \|\mathbf{Y} - \langle \mathbf{Y}(\beta) \rangle\|^2 \quad (1.8)$$

En el método de Gauss-Newton se calcula el paso $\delta_t = \beta - \beta_c$, resolviendo la ecuación ([Kelley99], [Marquardt63]):

$$R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) \delta_t = -R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \quad (1.9)$$

El subíndice t es usado para indicar que el vector de corrección fue calculado a partir de la aproximación de Taylor de primer orden. La ecuación (1.9) tiene implícita la condición que $R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c)$ sea no singular, lo que obliga a tener la condición $n \geq k$. Otra interpretación que permite incluir los problemas subdeterminados ([Kelley99]) es que la iteración de Gauss-Newton busca la solución de *norma mínima* del modelo lineal del problema de mínimos cuadrados

$$\min \|\mathbf{Y} - \langle \mathbf{Y}(\beta) \rangle\|^2 = \min \left\| \mathbf{Y} - \hat{\mathbf{Y}}(\beta_c) + R'_\epsilon(\beta_c)(\beta - \beta_c) \right\|^2 \quad (1.10)$$

En este caso la suposición sobre la matriz $R'_\epsilon(\beta^*)^T R'_\epsilon(\beta^*)$, donde β^* es un minimizador local, es que sea de rango maximal ([Kelley99]), es decir:

- Si $n = k$, $R'(\beta^*)$ es no singular.
- Si $n > k$, $R'(\beta^*)$ es inyectiva (Tiene rango columna completo).
- Si $n < k$, $R'(\beta^*)$ es sobreyectiva (Tiene rango fila completo).

El método de Gauss-Newton puede resumirse con el siguiente algoritmo:

Algoritmo 1.2.1. Se tiene una aproximación inicial al minimizador local, $\beta^{(0)}$, y parámetros de parada ϵ y τ . Para $r \geq 0$, hacer:

1. Calcular $R_\epsilon(\beta^{(r)})$ y $R'_\epsilon(\beta^{(r)})$.
2. Calcular $\delta_t^{(r)}$ usando la ecuación (1.9) o la ecuación (1.10), con $\beta_c = \beta^{(r)}$.
3. Hacer $\beta^{(r+1)} = \beta^{(r)} + \delta_t^{(r)}$.
4. Hacer $r = r + 1$.
5. Si $F(\beta^{(r+1)}) < \epsilon$ y/o $\|\delta_t^{(r)}\| < \tau$, o si $R'(\beta^{(r)}) = 0$, terminar, si no, ir al paso 1.

1.3 Método del Gradiente

El método de Gauss-Newton presenta una desventaja al exigir que la aproximación inicial a la solución del problema de la ecuación (1.2) esté cerca de un minimizador local, lo que lo hace generalmente propenso a diverger cuando no se cumple dicha condición. Esto se debe principalmente a que β_c no está lo suficientemente cerca de la solución β^* , la dirección de actualización δ_t no sea una dirección de decrecimiento de la función F , es decir que no cumple la condición:

$$\nabla F(\beta_c)^T \delta_t < 0 \quad (1.11)$$

Esto debido a que es posible que $\nabla^2 F(\beta_c)$ no sea una matriz definida positiva. Por otra parte, el método del gradiente garantiza siempre que la dirección de actualización es una dirección de decrecimiento para F , al calcularla como:

$$\delta_g = -\lambda \nabla F(\beta_c) = -2\lambda R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \quad (1.12)$$

El paso δ_g así calculado cumple que

$$\begin{aligned} \nabla F(\beta_c)^T \delta_g &= -2\lambda \left[R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \right]^T R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \\ &= -2\lambda R_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) < 0 \end{aligned}$$

Puesto que $R'_\epsilon(\beta_c) R'_\epsilon(\beta_c)^T$ es una matriz definida positiva.

El método del gradiente tiene la ventaja de ir siempre en una dirección de decrecimiento de F , pero su inconveniente es que debe definirse el tamaño del paso, λ , puesto que si este es muy grande es posible que el nuevo estimado calculado a partir de $\beta = \beta_c + \delta_g$, no sea un punto donde $F(\beta) \leq F(\beta_c)$, aún cuando β_c no este cerca de un minimizador local. La razón para esto es que, a diferencia del método de Newton, $\nabla F(\beta)$ se escala con el valor de F . En [Kelley99] se presentan estrategias de búsqueda en línea (Line Search) y la regla de Armijo para resolver el problema del tamaño del paso de δ_g , sin embargo el método del

gradiente así planteado tiene el grave inconveniente de presentar convergencia lenta.

El método del gradiente puede resumirse con el siguiente algoritmo:

Algoritmo 1.3.1. Se tiene una aproximación inicial al minimizador local, $\beta^{(0)}$, y parámetros de parada ϵ y τ . Para $r \geq 0$, hacer:

1. Calcular $R_\epsilon(\beta^{(r)})$ y $R'_\epsilon(\beta^{(r)})$.
2. Calcular $\delta_g^{(r)}$ como $-2R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c)$.
3. Calcular $\lambda^{(r)}$ tal que $F(\beta^{(r)} + \lambda^{(r)}\delta_g^{(r)}) \leq F(\beta^{(r)})$.
4. Hacer $\beta^{(r+1)} = \beta^{(r)} + \lambda^{(r)}\delta_g^{(r)}$.
5. Hacer $r = r + 1$.
6. Si $F(\beta^{(r+1)}) < \epsilon$ y/o $\|\delta_g^{(r)}\| < \tau$, o si $R'_\epsilon(\beta^{(r)}) = 0$, terminar, si no, ir al paso 1.

1.4 Método de Levenberg-Marquardt

Tanto el método de Gauss-Newton como el método del gradiente exigen un control cuidadoso del tamaño del paso empleado una vez que se ha establecido una dirección para el vector de corrección, pero aún así el resultado obtenido es, generalmente, una convergencia lenta. Con la meta de solucionar este problema, Kenneth Levenberg publicó un algoritmo que luego fue redescubierto por Donald Marquardt, quien lo presentó en [Marquardt63]. Básicamente, la idea es obtener en cada iteración un vector de corrección que esté a lo más a 90° de la dirección negativa del gradiente de F , es decir de δ_g , puesto que con otras direcciones es de esperarse obtener valores de F mayores a lo largo del vector de corrección. Por otro lado, debido a la no linealidad de la función F y de la distancia del punto β_c de un minimizador local, la dirección de δ_t está usualmente a 90° de δ_g , con lo que el vector

de corrección deseado debería estar entre δ_t y δ_g . En la figura 1-1 se presenta un ejemplo de una función $F(x, y)$ y las direcciones δ_g y δ_t dadas en un punto cercano al mínimo local. La función F es

$$F(x, y) = R_\epsilon(x, y)^T R_\epsilon(x, y)$$

Con

$$R_\epsilon(x, y) = \begin{bmatrix} \sqrt{|\cos(y-2)|} (x-1) \\ (y-2) \end{bmatrix}$$

Luego, $F(x, y) = |\cos(y-2)| (x-1)^2 + (y-2)^2$ y, adicionalmente, tiene un mínimo local en $(1, 2)$. El punto donde se calcularon los vectores δ_t y δ_g es $(-1, 1)$.

El método de Levenberg-Marquardt puede resumirse con el siguiente algoritmo:

Algoritmo 1.4.1. Se tiene una aproximación inicial al minimizador local, $\beta^{(0)}$, un valor inicial para λ , $\lambda^{(0)}$, y parámetros de parada ϵ y τ . Para $r \geq 0$, hacer:

1. Calcular $R_\epsilon(\beta^{(r)})$ y $R'_\epsilon(\beta^{(r)})$.
2. Calcular $\delta_{lm}^{(r)}$ con $\left[R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) + \lambda^{(0)} I \right] \delta_{lm}^{(r)} = -R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c)$.
3. Si $F(\beta^{(r)} + \lambda^{(r)} \delta_{lm}^{(r)}) \leq F(\beta^{(r)})$, ir al paso 4. Si no, buscar $\lambda^{(r)}$ tal que se garantice que el $\delta_{lm}^{(r)}$ calculado como en el paso 2 cumple $F(\beta^{(r)} + \lambda^{(r)} \delta_{lm}^{(r)}) \leq F(\beta^{(r)})$.
4. Hacer $\beta^{(r+1)} = \beta^{(r)} + \lambda^{(r)} \delta_{lm}^{(r)}$.
5. Hacer $r = r + 1$.
6. Si $F(\beta^{(r+1)}) < \epsilon$ y/o $\left\| \delta_g^{(r)} \right\| < \tau$, o si $R'_\epsilon(\beta^{(r)}) = 0$, terminar, si no, ir al paso 1.

Las bases teóricas del algoritmo de Levenberg-Marquardt son presentadas en [Marquardt63] y se condensan en los dos teoremas siguientes y un tercer teorema que no se enuncia en

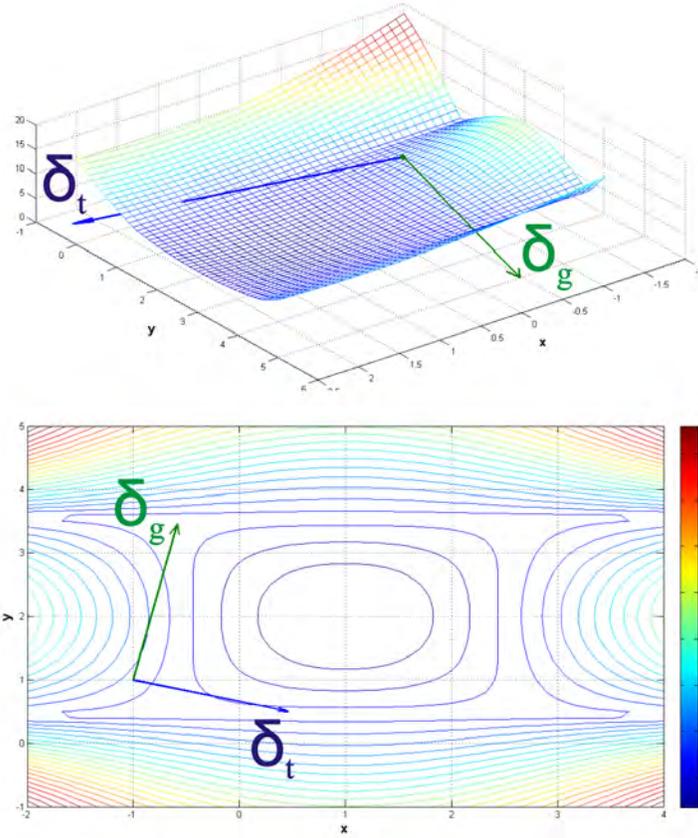


Figura 1-1: Función $F(x, y) = |\cos(y - 2)|(x - 1)^2 + (y - 2)^2$ y δ_t y δ_g en $(-1, 1)$.

este trabajo pero que aparece en [Marquardt63]. De ahora en adelante, por comodidad, se escribira $\hat{\mathbf{Y}} = \hat{\mathbf{Y}}(\beta_c)$ y $\delta = \beta - \beta_c$.

Teorema 1.4.1 *Sea $\lambda \geq 0$ arbitrario y sea δ_0 tal que satisface la ecuación*

$$\left[R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) + \lambda I \right] \delta_0 = -R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \quad (1.13)$$

Luego δ_0 minimiza $\langle F \rangle$ en la esfera cuyo radio $\|\delta\|$ satisface $\|\delta\|^2 = \|\delta_0\|^2$.

Prueba. Se busca δ tal que minimice la función

$$\langle F(\beta(\delta)) \rangle = \|\mathbf{Y} - \langle Y \rangle\|^2 = \left\| \mathbf{Y} - \hat{\mathbf{Y}}(\beta_c) + R'_\epsilon(\beta_c)(\beta - \beta_c) \right\|^2 \quad (1.14)$$

Bajo la restricción

$$\|\delta\|^2 = \|\delta_0\|^2$$

Usando el método de Lagrange, la condición necesaria para que un punto dado sea un punto estacionario es:

$$\frac{\partial u}{\partial \delta_1} = \frac{\partial u}{\partial \delta_2} = \dots = \frac{\partial u}{\partial \delta_k} = 0, \quad \frac{\partial u}{\partial \lambda} = 0 \quad (1.15)$$

donde $u(\delta, \lambda) = \left\| \mathbf{Y} - \hat{\mathbf{Y}} + R'_\epsilon(\beta_c)\delta \right\|^2 + \lambda \left(\|\delta\|^2 - \|\delta_0\|^2 \right)$ y λ es un multiplicador de Lagrange. Tomando las derivadas indicadas y simplificando se obtiene

$$\left[R'_\epsilon(\beta_c)^T \left(\mathbf{Y} - \hat{\mathbf{Y}} \right) + R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) \delta \right] + \lambda \delta = 0 \quad (1.16)$$

$$\|\delta\|^2 - \|\delta_0\|^2 = 0 \quad (1.17)$$

De la ecuación (1.16) se tiene que, para un λ dado, un punto estacionario δ de la función $\langle F(\delta) \rangle$ cumple

$$\left[R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) + \lambda I \right] \delta = -R'_\epsilon(\beta_c)^T \left(\mathbf{Y} - \hat{\mathbf{Y}} \right) = -R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \quad (1.18)$$

Con lo que queda demostrado que δ_0 es un punto estacionario. El hecho que δ_0 es un mínimo local viene del hecho que la matriz hessiana de $\langle F(\beta(\delta)) \rangle$, $\nabla^2 \langle F(\beta(\delta)) \rangle = [R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) + \lambda I]$ es definida positiva. ■

Teorema 1.4.2 Sea $\delta(\lambda)$ la solución de (1.13) para un valor dado de λ . Entonces $\|\delta(\lambda)\|^2$ es una función de λ continua y decreciente tal que si $\lambda \rightarrow \infty$, $\|\delta(\lambda)\|^2 \rightarrow 0$

Prueba. Como $R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c)$ es definida positiva, existen S ortogonal y D diagonal tales que $S^T R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) S = D$, y donde los elementos de D son positivos. Así, (1.13) se puede escribir como

$$\begin{aligned} S^T [R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) + \lambda I] \delta_0 &= -S^T R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \\ S^T [R'_\epsilon(\beta_c)^T R'_\epsilon(\beta_c) + \lambda I] S S^T \delta_0 &= -S^T R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \\ [D + \lambda I] S^T \delta_0 &= -S^T R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \end{aligned}$$

De donde:

$$\delta_0 = -S [D + \lambda I]^{-1} S^T R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c) \quad (1.19)$$

Definiendo $v = -S^T R'_\epsilon(\beta_c)^T R_\epsilon(\beta_c)$, se tiene que

$$\begin{aligned} \|\delta_0(\lambda)\|^2 &= [S [D + \lambda I]^{-1} v]^T [S [D + \lambda I]^{-1} v] \\ &= v^T [D + \lambda I]^{-1} S^T S [D + \lambda I]^{-1} v \\ &= v^T [D + \lambda I]^{-2} v \\ &= \sum_{j=1}^k \frac{v_j^2}{(D_j + \lambda)^2} \end{aligned}$$

Con lo que se tiene que $\|\delta_0(\lambda)\|^2$ es una función decreciente de λ tal que $\|\delta_0(\lambda)\|^2 \rightarrow 0$ cuando $\lambda \rightarrow \infty$. ■

Adicionalmente, Marquardt en [Marquardt63] prueba otro teorema que relaciona al ángulo entre δ_0 y δ_g , al que se denotará como γ , con el parámetro λ . El resultado principal

de dicho teorema es que $\gamma \rightarrow 0$ cuando $\lambda \rightarrow \infty$, lo que unido a la independencia de δ_g del parámetro λ permite concluir que la dirección δ_0 se acerca a la dirección δ_g a medida que $\lambda \rightarrow \infty$. En otras palabras, a medida que λ es más grande, el método de Levenberg-Marquardt se aproxima al método del gradiente.

Para ilustrar las ventajas que se pueden obtener al usar el método de Levenberg-Marquardt en el cálculo del vector de corrección, la figura 1-2 muestra el vector de actualización δ_{lm} calculado con $\lambda = 2$ en el punto $(-1, 1)$ para la función $F(x, y)$, también se graficaron los vectores de dirección de los métodos de Gauss-Newton y del Gradiente. Es notable que la dirección obtenida con el método de Levenberg-Marquardt es más deseable que las otras dos direcciones, adicionalmente el tamaño del paso fue determinado en el mismo momento que se encontró el vector δ_{lm} , mientras que en el método de Newton y del Gradiente es necesario calcular posteriormente dicho tamaño. γ

Claramente, durante todo el proceso del método de Levenberg-Marquardt, es posible cambiar el valor de λ de modo que pueda explotarse las ventajas de una dirección de corrección parecida a δ_g cuando se este lejos del mínimo local y también aprovechar las ventajas de una dirección de corrección cercana a δ_t cuando se está en una región lo suficientemente cercana al mínimo local de F , donde se pueda garantizar que se cumplen las condiciones para convergencia r-cuadrática del método de Gauss-Newton. Marquardt ([Marquardt63]) propuso una estrategia de actualización del $\lambda^{(r)}$ (λ correspondiente a la r-ésima iteración), buscando un $\lambda^{(r)}$ tal que

$$F^{(r+1)} < F^{(r)} \tag{1.20}$$

La estrategia de actualización de $\lambda^{(r)}$ de Marquardt, la cual es la correspondiente al paso **3** del algoritmo **1.4.1** se puede resumir así:

Algoritmo 1.4.2. Sean $v > 1$ y $\lambda^{(0)} = 10^{-2}$, y parámetros de parada ϵ y τ . Para $r \geq 0$, hacer:

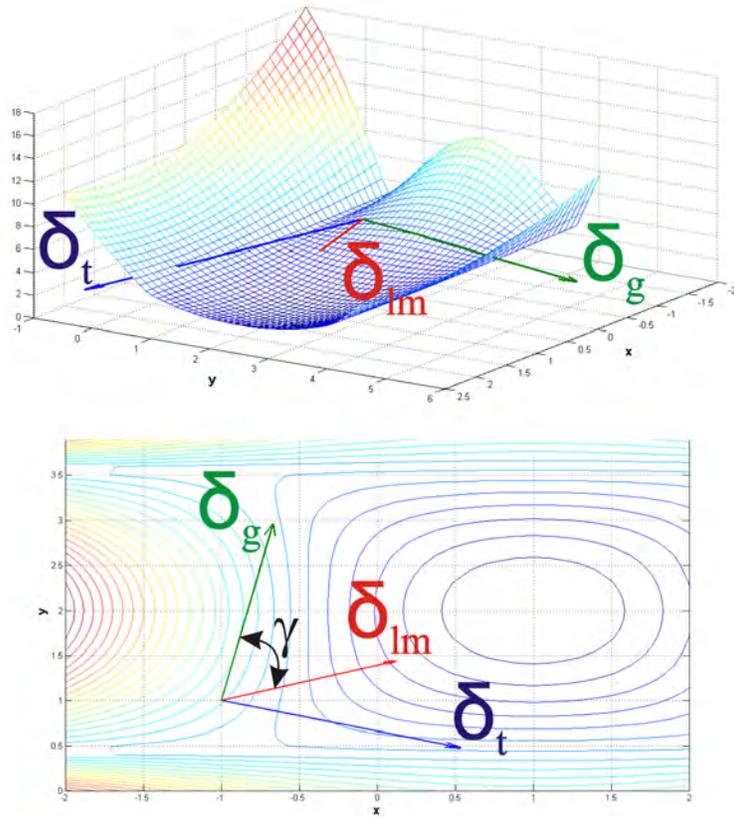


Figura 1-2: Función $F(x, y) = |\cos(y - 2)|(x - 1)^2 + (y - 2)^2$ y δ_t , δ_g y δ_{lm} en $(-1, 1)$.

1. Computar $F(\lambda^{(r-1)})$ y $F(\lambda^{(r-1)}/v)$.
2. Si $F(\lambda^{(r-1)}/v) \leq F^{(r)}$ entonces hacer $\lambda^{(r)} = \lambda^{(r-1)}/v$.
3. Si $F(\lambda^{(r-1)}/v) > F^{(r)}$ y $F(\lambda^{(r-1)}) \leq F^{(r)}$ entonces hacer $\lambda^{(r)} = \lambda^{(r-1)}$.
4. Si $F(\lambda^{(r-1)}/v) > F^{(r)}$ y $F(\lambda^{(r-1)}) > F^{(r)}$ entonces se incrementa λ al multiplicarla sucesivamente por v hasta obtener el más pequeño w tal que $F(\lambda^{(r-1)}v^w) \leq F^{(r)}$. En este caso, hacer $\lambda^{(r)} = \lambda^{(r-1)}v^w$.

El criterio de parada del algoritmo **1.4.1** se puede cambiar por el siguiente criterio:

$$\frac{|\delta_j^{(r)}|}{\epsilon + |\beta_j^{(r)}|} < \epsilon, \forall j \in \{1, \dots, k\} \quad (1.21)$$

En la ecuación (1.21) aparecen dos constantes, ϵ y ϵ , las cuales tienen funciones distintas: El valor ϵ es el valor mínimo para considerar que hay un cambio relativo importante en los parámetros, por lo que se debe parar el algoritmo cuando el cambio relativo de cada parámetro sea menor que ϵ , mientras que ϵ tiene como función evitar denominadores iguales a 0. Un valor usual para ϵ , propuesto en [Marquardt63], es 10^{-3} .

En [Marquardt63] se hace claridad que en el caso que se presente una alta correlación entre los parámetros β_j es posible que en el paso 4 del algoritmo **1.4.2** se incremente el valor de λ de forma desmesurada. En tal caso, el vector de actualización $\delta^{(r)}$ será de norma pequeña o despreciable, lo que se traduciría en un estancamiento del algoritmo. Para evitar este fenómeno, Marquardt ([Marquardt63]) propone no aumentar λ indiscriminadamente sino agregar el siguiente cambio al algoritmo **1.4.2**:

4. Si $F(\lambda^{(r-1)}/v) > F^{(r)}$ y $F(\lambda^{(r-1)}) > F^{(r)}$, hacer $K^{(r)} = 1$ y tomar, por ejemplo, $\gamma_0 = \pi/4$ (Valor recomendado en [Marquardt63]). Incrementar λ al multiplicarla sucesivamente por v hasta obtener $F(\lambda^{(r-1)}v^w) \leq F^{(r)}$ o hasta que $\gamma^{(r)}$, el ángulo entre la dirección de actualización dada por el algoritmo de Marquardt con $\lambda^{(r)} = \lambda^{(r-1)}v^w$ y δ_g , la dirección

del gradiente, sea menor que γ_0 . Si se llega primero a $\gamma^{(r)} < \gamma_0$, no aumentar más $\lambda^{(r)}$ y disminuir $K^{(r)}$ hasta que $F(\beta^{(r)}) = F(\beta^{(r-1)} + K^{(r)}\delta^{(r)}) \leq F(\beta^{(r-1)})$ con $\delta^{(r)}$ obtenido con el valor de $\lambda^{(r)}$.

El método de Levenberg-Marquardt es un caso especial de los métodos denominados de región de confianza (trust-region), [Kelley99]. Adicionalmente, se tiene el siguiente resultado de convergencia:

Teorema 1.4.3 *Sea $R_\epsilon \in \mathbb{R}^m \rightarrow \mathbb{R}^n$ continuamente diferenciable. Sean $\{\beta_k\}$ y $\{\lambda_k\}$ las secuencias de iterados y parámetros de Levenberg-Marquardt generados por dicho método. Suponga que $\{\lambda_k\}$ está acotada superiormente. Entonces $R'_\epsilon(\beta_k)^T R_\epsilon(\beta_k) = 0$ para algún k finito o*

$$\lim_{k \rightarrow \infty} R'_\epsilon(\beta_k)^T R_\epsilon(\beta_k) = 0$$

Más aún, si β^ es un punto límite de $\{\beta_k\}$ para el cual $R_\epsilon(\beta^*) = 0$ y $R'_\epsilon(\beta^*)$ tiene rango completo, entonces $\beta_k \rightarrow \beta^*$ q -cuadráticamente y $\lambda_k = 0$ para k suficientemente grande.*

Prueba. *La prueba de este teorema se presenta en [More78]. ■*

El resultado principal del teorema 1.4.3 puede resumirse así: Dado un vector de residuales R_ϵ tal que el método de Levenberg-Marquardt entrega valores de λ acotados y converge a un β^* que cumple $R_\epsilon(\beta^*) = 0$ y $R'_\epsilon(\beta^*)$ de rango completo, la velocidad de convergencia del método de Levenberg-Marquardt es comparable con la del método de Newton, el cual presenta convergencia r -cuadrática cuando el minimizador local cumple que $R'_\epsilon(\beta^*)$ sea de rango completo. Este resultado indica que el método de Levenberg-Marquardt es preferible al método del Gradiente en cuanto su velocidad de convergencia puede ser mayor y, debido a que evita trabajar con matrices singulares al agregarle λI a la matriz $R'_\epsilon(\beta_k)^T R_\epsilon(\beta_k)$, se hace más robusto que el método de Newton en tanto que puede aumentar la región de convergencia. Es importante tener en cuenta que estos resultados son en los casos que las condiciones del funcional a minimizar son las supuestas en el teorema y que, en muchos casos prácticos, los resultados obtenidos también dependen de usar

un método hecho a la medida del problema.

1.5 Convergencia y velocidad de convergencia del método de Levenberg-Marquardt

El teorema 1.4.3 presentado en la sección 1.4, invita a esperar que el algoritmo de Levenberg-Marquardt presente convergencia más rápida que el algoritmo del Gradiente y un tanto parecida a la del algoritmo de Newton cuando el estimado inicial está lo suficientemente cerca de la solución buscada. Para estudiar el comportamiento de dichos algoritmos, primero se considerará un vector de residuales, R_ϵ , que cumpla con la condición de ser continuamente diferenciable. Sea R_ϵ dado por

$$R_\epsilon(x, y) = \begin{bmatrix} \cos(y - 2)(x - 1) \\ y - 2 \end{bmatrix} \quad (1.22)$$

En este caso, el funcional a minimizar es

$$F(x, y) = \cos^2(y - 2)(x - 1)^2 + (y - 2)^2 \quad (1.23)$$

La gráfica de dicho funcional se presenta en la figura 1-3, donde puede apreciarse que $F(1, 2) = 0$ es el mínimo de la función.

Se usarán el algoritmo de Newton, del Gradiente y de Levenberg-Marquardt para aproximar el punto minimizador de la función, partiendo de la aproximación inicial $(4, 0)$. Debido a los inconvenientes que pueden presentarse en los métodos de Newton y Gradiente debido al tamaño de paso, en cada iteración se hace $\beta^{(r+1)} = \beta^{(r)} + K^{(r)}\delta^{(r)}$, donde $K^{(r)} \geq 1$ garantiza que $F(\beta^{(r+1)}) \leq F(\beta^{(r)})$. Esta pequeña modificación permite garantizar que ambos algoritmos no se alejen indiscriminadamente del minimizador cuando las condiciones no son las necesarias para garantizar convergencia. En cuanto al método de Levenberg-Marquardt, se usaron los parámetros $\lambda^{(0)} = 10$ y $v = 5$. En la figura 1-4 se presenta una

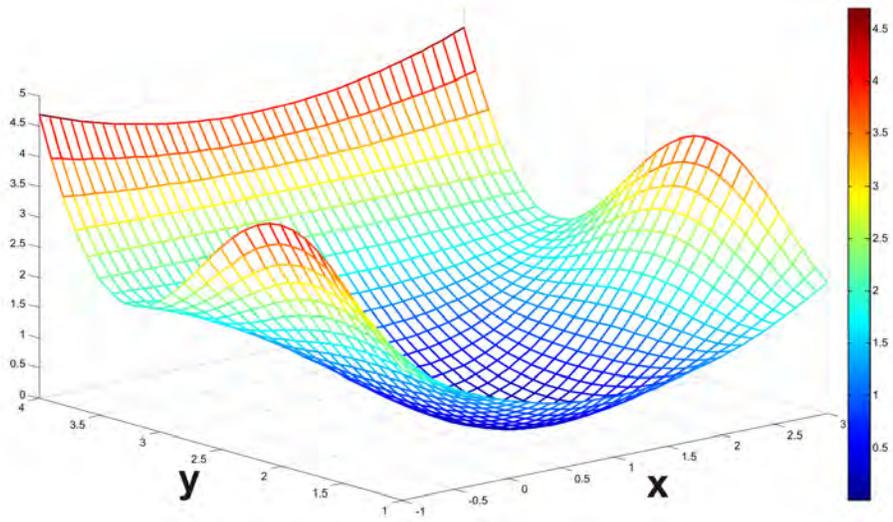


Figura 1-3: Función $F(x, y) = \cos^2(y - 2)(x - 1)^2 + (y - 2)^2$

gráfica de la magnitud de F con el vector de parámetros $\beta^{(r)}$ en cada iteración, que en este caso también representa el error cometido en cada paso, puede observarse como los tres métodos tardan en llegar a una región donde la convergencia se acelera y que el método de Newton tiene la velocidad de convergencia más alta, aunque el método de Levenberg-Marquardt también tiene una velocidad considerablemente mayor al método del Gradiente. El lector podría pensar en este momento que el método de Newton tiene ventajas sobre el método de Levenberg-Marquardt, pero dichas ventajas sólo son posibles cuando el estimado inicial está lo suficientemente cerca del minimizador local y el jacobiano del vector de residuales no se hace singular.

Para ilustrar con mayor claridad las ventajas del método de Levenberg-Marquardt sobre el método del Gradiente y el de Newton, se usa el vector de residuales presentado en uno de los ejemplos de [Madsen04], dicho vector está dado por

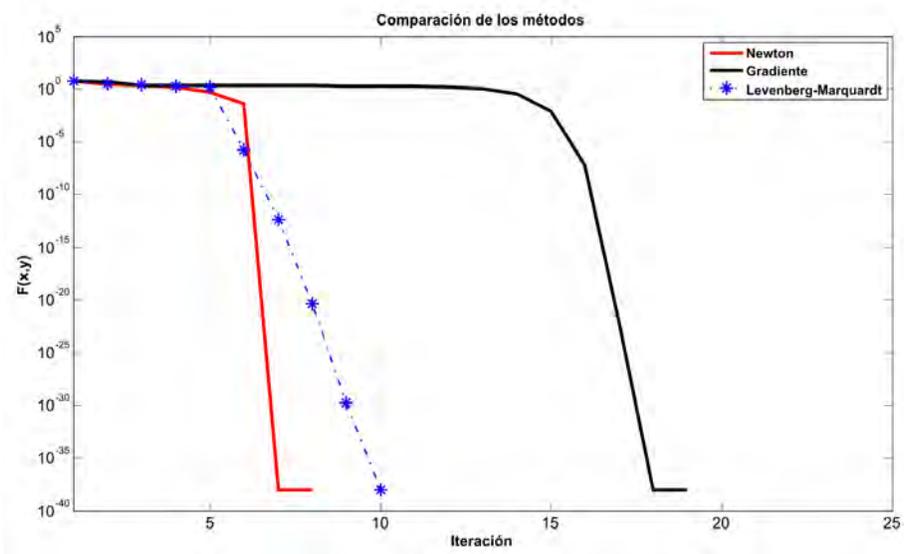


Figura 1-4: Error cometido en cada iteración con los tres métodos de optimización, para la función de la ecuación (1.23).

$$R_\epsilon(x, y) = \begin{bmatrix} x \\ \frac{10x}{x+0.1} + 2y^2 \end{bmatrix} \quad (1.24)$$

Para el vector de residuales dado en (1.24), el problema $\min R_\epsilon^T R_\epsilon$ tiene una única solución en $(0, 0)$, pero el Jacobiano de dicho vector de residuales es

$$R'_\epsilon(x, y) = \begin{bmatrix} 1 & 0 \\ (x+0.1)^2 & 4y \end{bmatrix} \quad (1.25)$$

Y dicho jacobiano es singular en la solución del problema, $(0, 0)$. En la figura 1-5 se presentan las gráficas de evolución del error cometido por los tres métodos de optimización al minimizar el funcional $F(x, y) = R_\epsilon^T(x, y) R_\epsilon(x, y)$, con R_ϵ dado en la ecuación (1.24), en la parte superior se presentan los resultados al usar como estimado inicial el punto $(1, 1)$ mientras que en la parte inferior se observan los resultados al usar como estimado inicial el punto $(3, 1)$. Nuevamente los parámetros iniciales para el método de Levenberg-Marquardt fueron $\lambda^{(0)} = 10$ y $v = 5$, también se usó el parámetro $K^{(r)}$ en los métodos del Gradiente y de Newton, el cual es el responsable del estancamiento observado en ambos casos.

En el caso donde la aproximación inicial es $(1, 1)$, el método de Levenberg-Marquardt y el método de Newton presentan la misma velocidad de convergencia, aunque el método de Levenberg-Marquardt tiene un comportamiento parecido al Gradiente en las primeras iteraciones. Queda claro que el método del Gradiente presenta rápidamente estancamiento lo que hace pensar que no es el método más apropiado para éste problema. El resultado de final obtenido con el método de Newton es el punto $(0, -0.8763 \times 10^{-7})$ que está más cerca de la solución exacta que los resultados obtenidos con el método del Gradiente y de Levenberg-Marquardt, $(-0.0044, -0.5079)$ y $(-4.2 \times 10^{-20}, 1.7 \times 10^{-7})$ respectivamente, aunque la diferencia de los estimados obtenidos con el método de Newton y Levenberg-Marquardt es del orden de 10^{-7} , lo que indica que ambos métodos son comparables en éste caso.

Cuando el estimado inicial fue $(3, 1)$, que está más alejado de la solución exacta (parte inferior de la figura 1-5), el método de Newton falló porque $\delta_t(3, 1)$ no resulta ser una dirección de decrecimiento:

$$\delta_t(3, 1) = - \left[R'_\epsilon(3, 1)^T R'_\epsilon(3, 1) \right]^{-1} R'_\epsilon(3, 1)^T R_\epsilon(3, 1) = \begin{bmatrix} -3 \\ 4.288 \end{bmatrix} \quad (1.26)$$

Con lo que $\nabla F(3, 1)^T \delta_t(3, 1) = \begin{bmatrix} -115.22 & -46.709 \end{bmatrix} \begin{bmatrix} -3 \\ 4.288 \end{bmatrix} = 145.36 > 0$ y por tanto, al no cumplirse la condición (1.11), el vector de actualización $K^{(r)}\delta_t^{(r)}$ no permitirá encontrar un $\beta^{(r+1)}$ tal que $F(\beta^{(r+1)}) \leq F(\beta^{(r)})$ sin importar cuan pequeño sea el valor $K^{(r)}$. En este segundo caso, el método de Levenberg-Marquardt entregó la mejor solución entre los tres métodos, $(-5.8848 \times 10^{-6}, 0.0172)$, tal y como podría esperarse teniendo en cuenta que el método de Levenberg-Marquardt tiene una región de convergencia casi tan grande como el método del Gradiente. En este caso el método del Gradiente y de Levenberg-Marquardt tienen un comportamiento parecido hasta la iteración 7, donde se encuentran mucho más cerca de la solución exacta y el método de Levenberg-marquardt empieza a incrementar la velocidad de convergencia.

Con base en los resultados presentados anteriormente, es posible concluir que el método de Levenberg-Marquardt presenta una región de convergencia más grande que el método de Newton y a la vez tiene una velocidad de convergencia mayor a la del método del Gradiente y casi tan alta como la de Newton y por ésta razón en los capítulos siguientes se usará exclusivamente el algoritmo de Levenberg-Marquardt para resolver los problemas de mínimos cuadrados asociados al proceso de calibración del modelo Pinhole, el cual se presenta en el capítulo 2.

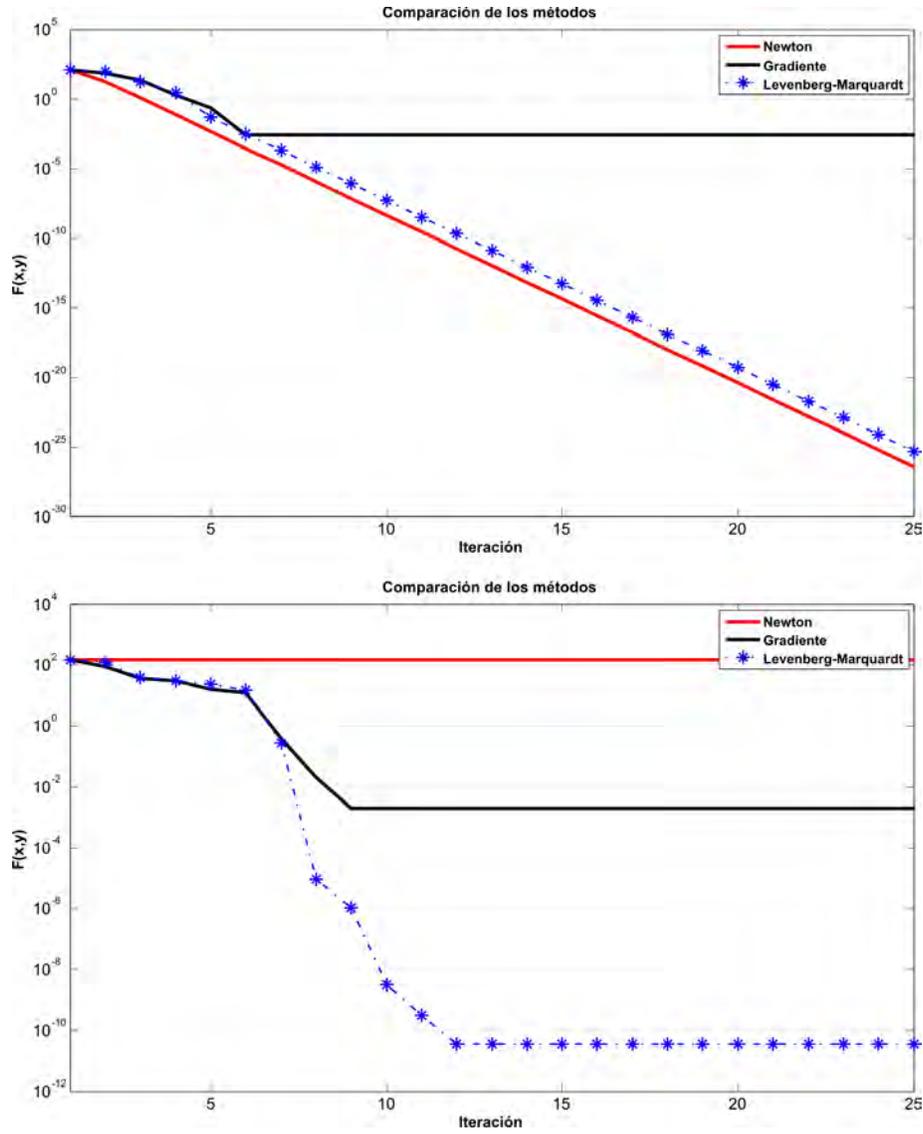


Figura 1-5: Error de los métodos de optimización en la función de la ecuación (1.24). Con estimado inicial (1, 1) (Arriba) y con estimado inicial (3, 1) (Abajo).

Capítulo 2

Modelo Pinhole

Como se mencionó en la introducción, este trabajo está enmarcado en un proyecto de monitoreo ambiental basado en el uso de cámaras de video desarrollado por el equipo HORUS de la Universidad Nacional de Colombia y la Universidad de Cantabria, España (Ver <http://www.horusvideo.com>). Para el desarrollo de dicho proyecto, se instalaron 4 cámaras en el edificio Bavaria, ubicado en Cartagena de Indias, las cuales están monitoreando las playas de Bocagrande y el Laguito, fotos correspondientes a algunas de estas cámaras se observan en la figura 2-1. Uno de los objetivos principales del proyecto es usar las estaciones de captura HORUS para cuantificar variables ambientales y de uso de las playas, como son la línea de costa, la densidad de usuarios, entre otras, (Para ampliar las ideas de las estaciones HORUS y sus usos, ver [Ortiz08]). Para llevar a cabo dichas cuantificaciones, es necesario poder medir en las imágenes como si fueran mapas o fotos satelitales, proceso que se dificulta en tanto que las cámaras no se encuentran perpendiculares a la zona de estudio.

En este caso, la estrategia usada es obtener un modelo de la cámara que permita relacionar las coordenadas en la imagen con coordenadas en un sistema de referencia en el espacio. Uno de los modelos más usados es el modelo Pinhole, el cual ha sido estudiado desde el punto de vista de la geometría proyectiva ([Faugeras93], [Hartley03]) y desde la fotogrametría ([Wolf00]). Varios autores han presentado estrategias para resolver dicho modelo ([Chartterjee93], [Heikilla97], [Salvi02]) y otros tantos han presentado diversas aplicaciones del mismo ([Holland97], [Ortiz08], [Osorio07]).

En la figura 2-2 se presentan las cámaras usadas en las estaciones de captura HORUS,



Figura 2-1: Cámaras (Dentro de sus carcazas protectoras) usadas en la estación de captura HORUS en Cartagena y PC de control.

instaladas en las playas de Cartagena de Indias, Colombia y en las playas de La Magdalena y El Sardinero, Santander, España.

2.1 Principios del Modelo Pinhole

El modelo Pinhole ([Faugeras93], [Hartley03], [Salvi02], [Wolf00]) se basa en el principio de colinealidad, es decir, supone que cada punto en el espacio y su correspondiente proyección en la imagen están sobre una línea recta que pasa por el centro óptico de la imagen. Para mayor claridad, la figura 2-3 presenta un esquema del modelo Pinhole, con cada uno de los sistemas de referencia usados durante el proceso de proyectar un punto del espacio sobre la imagen.

Donde:



Figura 2-2: Cámaras STINGRAY y MARLIN, usadas en las estaciones de captura HORUS.

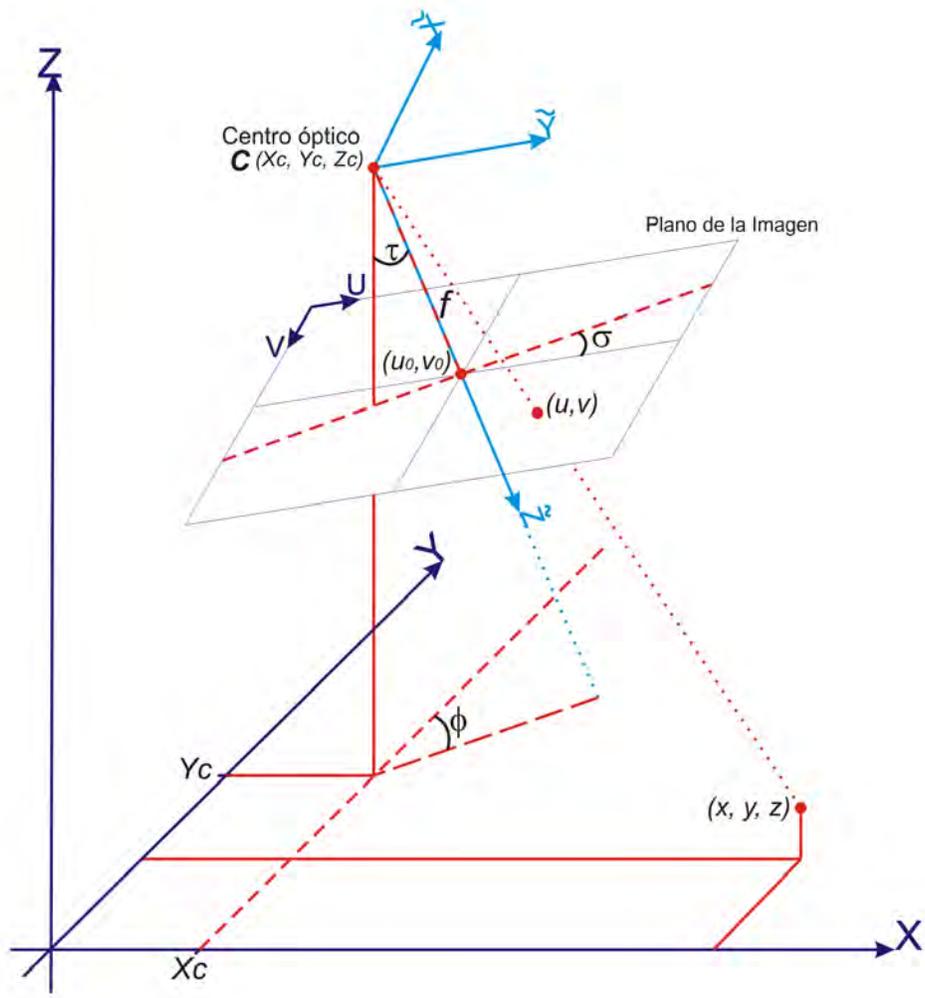


Figura 2-3: Esquema del Modelo Pinhole.

(x, y, z) son las coordenadas de un punto en el espacio,

medidas en el sistema de referencia XYZ

(u, v) son las coordenadas del punto (x, y, z) en la imagen,

medidas en el sistema de referencia UV en píxeles.

(τ, σ, ϕ) son ángulos de rotación, correspondientes al *tilt*,

roll y *azimuth*.

f es la distancia del centro óptico al plano de la imagen,

medida en unidades métricas.

(x_c, y_c, z_c) son las coordenadas del centro óptico C ,

medidas en el sistema de referencia XYZ .

(u_0, v_0) son las coordenadas del punto principal de la imagen,

medidas en el sistema de referencia UV en píxeles.

El punto principal de la imagen corresponde a la intersección entre el rayo principal (o línea focal) y el plano de la imagen ([Hartley03]). Para obtener la expresión del modelo Pinhole, primero es necesario obtener un sistema de referencia cuyo plano $\tilde{X}\tilde{Y}$ sea paralelo al plano de la imagen UV y su origen sea el centro óptico, en otras palabras, se necesita obtener un sistema coordenado cuyo eje \tilde{Z} positivo coincida con el rayo principal de la imagen. Para lograr esto, es necesario llevar a cabo una rotación y una traslación del sistema XYZ , de modo que un punto en el sistema coordenado del espacio, (x, y, z) , es transformado al punto $(\tilde{x}, \tilde{y}, \tilde{z})$ del sistema coordenado $\tilde{X}\tilde{Y}\tilde{Z}$. Dicha transformación se representa con la ecuación (2.1).

$$\begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x - x_c \\ y - y_c \\ z - z_c \end{bmatrix} \quad (2.1)$$

Los valores r_{11}, \dots, r_{33} dependen de los ángulos τ, σ, ϕ , tal como se muestra en la ecuación (2.2). En el caso de usar otro sistema de rotación, la matriz de rotación se mantiene constante a pesar que la expresión usada para calcularla cambie, tal y como se explica en [Wolf00].

$$\begin{aligned} r_{11} &= \cos \phi \cos \sigma + \sin \phi \cos \tau \sin \sigma \\ r_{12} &= -\sin \phi \cos \sigma + \cos \phi \cos \tau \sin \sigma \\ r_{13} &= \sin \tau \sin \sigma \\ r_{21} &= -\cos \phi \sin \sigma + \sin \phi \cos \tau \cos \sigma \\ r_{22} &= \sin \phi \sin \sigma + \cos \phi \cos \tau \cos \sigma \\ r_{23} &= \sin \tau \cos \sigma \\ r_{31} &= \sin \phi \sin \tau \\ r_{32} &= \cos \phi \sin \tau \\ r_{33} &= -\cos \tau \end{aligned} \quad (2.2)$$

Luego, se proyecta el punto $(\tilde{x}, \tilde{y}, \tilde{z})$ en el plano de la imagen, aprovechando que el principio de colinealidad permite aplicar semejanza de triángulos tal y como se muestra en la figura 2-4, para obtener el punto (\tilde{u}, \tilde{v}) , cuyas coordenadas están representadas en la ecuación 2.3.

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = \frac{f}{\tilde{z}} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} \quad (2.3)$$

Por último, es necesario realizar la conversión de unidades métricas del punto (\tilde{u}, \tilde{v}) a píxeles, con lo que se obtiene el punto (u, v) . Esta operación realiza un escalamiento que no es necesariamente igual en ambas direcciones, por lo que se usan parámetros de escalamiento D_u y D_v , y adicionalmente realiza una traslación del sistema $\tilde{U}\tilde{V}$. Pero es

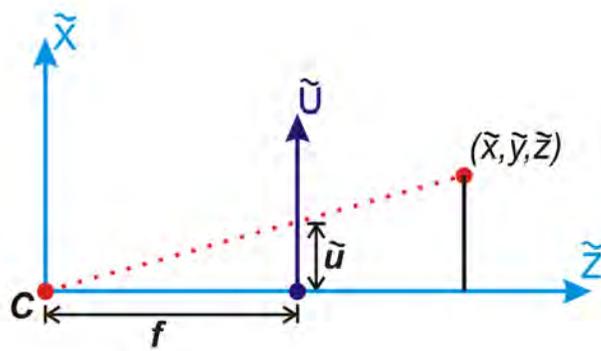


Figura 2-4: Proyección de la coordenada \tilde{x} en el plano de la imagen, bajo el principio de colinealidad.

importante tener cuidado con el sistema coordenado usado para medir en la imagen, puesto que si este es rotación derecha, como en la figura 2-5, la transformación a realizar está dada por la ecuación (2.4).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} D_u \tilde{u} \\ D_v \tilde{v} \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (2.4)$$

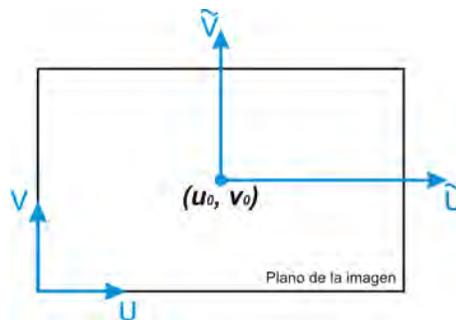


Figura 2-5: Plano de la imagen con sistema de coordenadas de rotación derecha.

Normalmente, la convención para medir la posición de un punto en una imagen es que

se use un sistema de rotación izquierda, como puede apreciarse en la figura 2-6. Este hecho obliga a tomar el opuesto en la dirección del eje \tilde{V} , para obtener la relación $(\tilde{u}, \tilde{v}) \leftrightarrow (u, v)$ dada por la ecuación (2.5).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} D_u \tilde{u} \\ -D_v \tilde{v} \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (2.5)$$

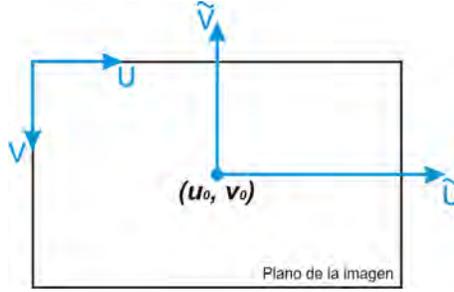


Figura 2-6: Plano de la imagen con sistema de coordenadas de rotación izquierda.

Luego, para cada punto en el espacio con coordenadas (x, y, z) su correspondiente representación en la imagen tiene coordenadas (u, v) dadas en las ecuaciones (2.6) y (2.7), las cuales son obtenidas a partir de las ecuaciones (2.1), (2.3) y (2.5).

$$u = f D_u \left(\frac{r_{11}(x - x_c) + r_{12}(y - y_c) + r_{13}(z - z_c)}{r_{31}(x - x_c) + r_{32}(y - y_c) + r_{33}(z - z_c)} \right) + u_0 \quad (2.6)$$

$$v = -f D_v \left(\frac{r_{21}(x - x_c) + r_{22}(y - y_c) + r_{23}(z - z_c)}{r_{31}(x - x_c) + r_{32}(y - y_c) + r_{33}(z - z_c)} \right) + v_0 \quad (2.7)$$

Haciendo uso de coordenadas homogéneas ([Hartley03]), el modelo Pinhole puede rep-

resentarse matricialmente de la forma:

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = KR [I_3] - C \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.8)$$

donde:

$$K = \begin{bmatrix} fD_u & 0 & u_0 \\ 0 & -fD_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \text{ es la matriz de calibración interna de la cámara.}$$

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \text{ es la matriz de rotación.}$$

$$C = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \text{ es el vector que representa la posición del centro óptico.}$$

$[I_3] - C$ es la matriz identidad de orden 3 ampliada con el vector $-C$

w es un factor de escala arbitrario.

La matriz K es una matriz no singular, puesto que tiene determinante distinto de 0, mientras que la matriz R es una matriz ortogonal.

En algunas ocasiones es más práctico no dejar explícita la posición del centro óptico de

la imagen, de modo que la representación dada en la ecuación (2.8) se convierte en:

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = K [R|t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2.9)$$

Donde $t = [t_x \ t_y \ t_z]^T = -RC$. La representación del modelo Pinhole dada por la ecuación (2.9) coincide con aquella dada en ([Salvi02]). La matriz $[R|t]$ se conoce como la matriz de orientación externa de la cámara, puesto que da la información referente a la orientación de la cámara respecto a un sistema coordenado.

Esta versión del modelo Pinhole no tiene en cuenta efectos de distorsión causados por los lentes de la cámara ni por las propiedades atmosféricas del aire, por lo cual es un modelo ideal. Para ser usado en situaciones reales, es necesario corregir los datos medidos o calibrar la distorsión presente en el sistema cámara-lente, la cual se explicará en la sección 2.2.

2.2 Modelo Pinhole con distorsión

En esta sección, se presentará una versión del modelo Pinhole ampliada para tener en cuenta efectos de la distorsión causada por los lentes usados en la cámara y sus imperfecciones. Es importante decir que la distorsión depende de los lentes y de la configuración interna de la cámara, lo que hace necesario recalibrar dicha distorsión cuando se cambian de lentes o simplemente se altera el enfoque de la cámara. En la figura 2-7 se presentan distintos tipos de lentes, todos usados en alguna estación de captura HORUS.

En la figura 2-8 se presenta un ejemplo de una imagen que evidentemente presenta distorsión. La fotografía corresponde a un patrón de calibración en forma de tablero de ajedrez, los cuales son ampliamente usados para calibrar los parámetros de distorsión de las cámaras, aunque también pueden usarse elementos tridimensionales para realizar la



Figura 2-7: Distintos tipos de lentes usados en las estaciones de captura HORUS.

calibración (Ver [Heikilla97], [Holland97], [Wei94], [Zhang98], [Zhang00]).

Como resultado de los diferentes tipos de imperfecciones en el diseño, construcción y ensamblaje de los lentes usados en la cámara, la expresión en la ecuación (2.3) no se cumple y debe ser reemplazada por expresiones que tengan en cuenta el error introducido en la posición:

$$\begin{bmatrix} \tilde{u}^* \\ \tilde{v}^* \end{bmatrix} = \begin{bmatrix} \tilde{u} + \delta_u(\tilde{u}, \tilde{v}) \\ \tilde{v} + \delta_v(\tilde{u}, \tilde{v}) \end{bmatrix} \quad (2.10)$$

donde (\tilde{u}, \tilde{v}) son las coordenadas en la imagen libres de distorsión, que no son observables, y $(\tilde{u}^*, \tilde{v}^*)$ son las correspondientes coordenadas con distorsión. Como puede observarse en la ecuación (2.10), la cantidad de distorsión a lo largo de cada eje coordenado depende de la posición. En general, se consideran 3 tipos de distorsión causadas por los lentes (Ver [Heikilla97] y [Weng92]):

- *Distorsión radial*: Este tipo de distorsión se debe principalmente a la curvatura de los lentes y posibles imperfecciones en la misma. La distorsión radial es estrictamente

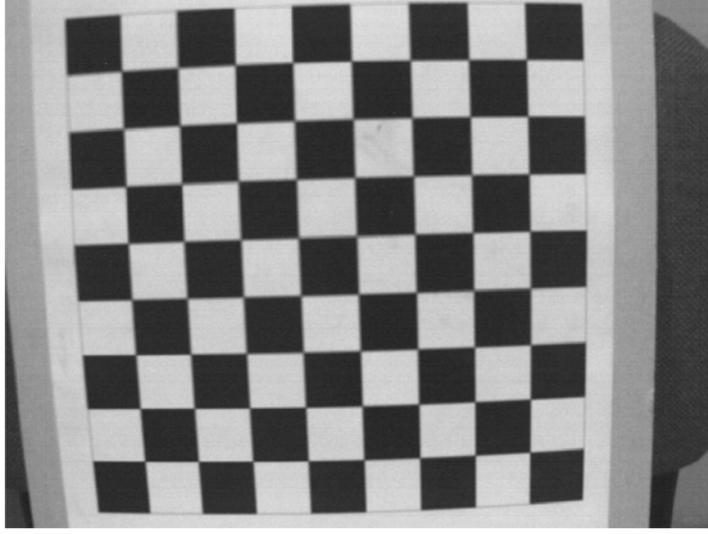


Figura 2-8: Imagen con distorsión tomada con una cámara MARLIN y un lente de 8 mm.

simétrica respecto al eje óptico, \tilde{Z} , por lo que ocasiona que un punto en la imagen se desplace radialmente sobre el plano de la imagen. Un desplazamiento radial negativo de los puntos de la imagen hace que los puntos más alejados del centro del plano de la imagen se acumulen y que la escala disminuya. Un desplazamiento radial positivo ocasiona que los puntos se alejen más a medida que aumenta el radio, aumentando la escala. En la figura 2-9 se ilustra el efecto de la distorsión radial.

La cantidad de distorsión radial en las direcciones \tilde{u} y \tilde{v} de unos lentes perfectamente centrados está gobernada por la expresión (Ver [Heikilla97], [Zhang00]):

$$\begin{bmatrix} \delta_{ur}(\tilde{u}, \tilde{v}) \\ \delta_{vr}(\tilde{u}, \tilde{v}) \end{bmatrix} = \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} [k_1 r^2 + k_2 r^4 + O[r^6]] \quad (2.11)$$

donde $r = \sqrt{\tilde{u}^2 + \tilde{v}^2}$. La expresión (2.11) indica que la cantidad de distorsión radial se modela como un polinomio y que los términos de grado igual o superior a 6 son

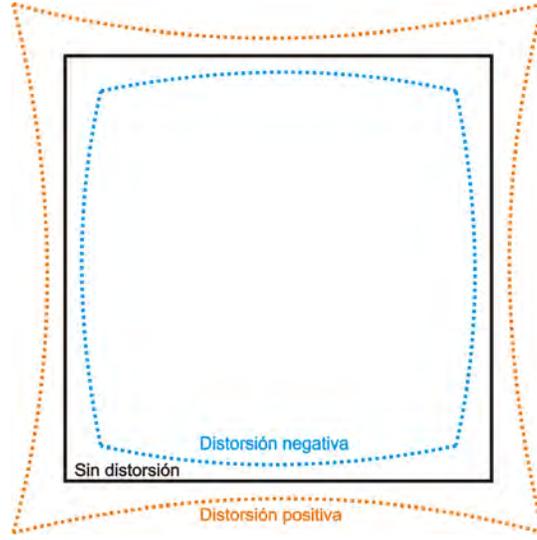


Figura 2-9: Efecto de la distorsión radial. Línea continua: Sin distorsión. Líneas punteadas: Con distorsión radial.

despreciables, incluso algunos autores (Ver [Weng92]) sólo tienen en cuenta el término cuadrático en dicha expresión.

- *Distorsión por descentramiento*: Este tipo de distorsión se debe a que los centros ópticos de los lentes del sistema no son estrictamente colineales. Este defecto produce distorsión en la componente radial y tangencial del plano de la imagen, con lo que la cantidad de distorsión tangencial en las direcciones u y v puede representarse por la expresión (Ver [Heikilla97], [Weng92]):

$$\begin{bmatrix} \delta_{ud}(\tilde{u}, \tilde{v}) \\ \delta_{vd}(\tilde{u}, \tilde{v}) \end{bmatrix} = \begin{bmatrix} p_1(3\tilde{u}^2 + \tilde{v}^2) + 2p_2\tilde{u}\tilde{v} + O[r^4] \\ 2p_1\tilde{u}\tilde{v} + p_2(\tilde{u}^2 + 3\tilde{v}^2) + O[r^4] \end{bmatrix} \quad (2.12)$$

- *Distorsión de prisma delgado*: La distorsión de prisma delgado se debe a la imperfección en el diseño y construcción de los lentes como también al ensamblaje de la cámara (Por ejemplo, una pequeña inclinación en los lentes o en el sensor de la

cámara). La expresión que modela la cantidad de distorsión a lo largo de los ejes u y v está dada por:

$$\begin{bmatrix} \delta_{up}(\tilde{u}, \tilde{v}) \\ \delta_{vp}(\tilde{u}, \tilde{v}) \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} [r^2 + O[r^4]] \quad (2.13)$$

Aunque cada uno de los tipos de distorsión presentes son independientes, las ecuaciones (2.11), (2.12) y (2.13) pueden combinarse en una sola ecuación de modo que ésta describa la cantidad total de distorsión en las direcciones u y v . La expresión que modela la distorsión total es:

$$\begin{bmatrix} \delta_u(\tilde{u}, \tilde{v}) \\ \delta_v(\tilde{u}, \tilde{v}) \end{bmatrix} = [k_1 r^2 + k_2 r^4] \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} + \begin{bmatrix} 2k_3 \tilde{u}\tilde{v} + k_4 (r^2 + 2\tilde{u}^2) + k_5 r^2 \\ k_3 (r^2 + 2\tilde{v}^2) + 2k_4 \tilde{u}\tilde{v} + k_6 r^2 \end{bmatrix} \quad (2.14)$$

Donde k_1, \dots, k_6 son los parámetros de distorsión de la cámara y, por ejemplo, k_1 y k_2 son los parámetros de distorsión radial. Usando la expresión dada en (2.14) en la expresión (2.10) permite completar el modelo Pinhole de modo que las coordenadas (u, v) observables en la imagen correspondientes a un punto con coordenadas (x, y, z) están dadas por:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} \tilde{u} + \delta_u(\tilde{u}, \tilde{v}) \\ \tilde{v} + \delta_v(\tilde{u}, \tilde{v}) \\ 1 \end{bmatrix} \quad (2.15)$$

donde K es la matriz de calibración interna de la cámara (Definida anteriormente en la ecuación (2.8)), y (\tilde{u}, \tilde{v}) están dados por la expresión (2.3). Una expresión análoga a aquella dada en (2.8) para el modelo Pinhole completo es:

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = K [R|t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} + K \begin{bmatrix} \tilde{w}\delta_u(\tilde{u}, \tilde{v}) \\ \tilde{w}\delta_v(\tilde{u}, \tilde{v}) \\ 0 \end{bmatrix} \quad (2.16)$$

donde $\begin{bmatrix} \tilde{w}\tilde{u} \\ \tilde{w}\tilde{v} \\ \tilde{w} \end{bmatrix} = [R|t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$ y $\delta_u(\tilde{u}, \tilde{v}), \delta_v(\tilde{u}, \tilde{v})$ están dados por la expresión (2.14).

Capítulo 3

Calibración del modelo Pinhole en 2 etapas

Para poder calibrar el modelo Pinhole, es decir, estimar sus parámetros, es necesario hacer uso de puntos de control terrestre (GCP, por sus siglas en inglés) que son puntos fácilmente identificables en la imagen y cuyas coordenadas en el espacio son conocidas, de modo que se conocen sus coordenadas (x, y, z) y (u, v) . Existen diversos métodos propuestos en la literatura para calibrar el modelo Pinhole dado que se cuente con suficientes puntos de control para hallar todos los parámetros, algunos de estos métodos resuelven directamente el modelo Pinhole con esquemas de optimización no lineal (ver [Chartterjee93] y [Salvi02]), otros recurren a métodos iterativos de dos pasos que resuelven una parte linealizada y luego realizan una optimización no lineal (ver [Salvi02] y [Weng92]) y algunos otros usan un modelo lineal obtenido a partir del modelo original llamado Transformada lineal Directa (DLT, por sus siglas en inglés, ver [Hartley03], [Holland97], [Osorio07] y [Salvi02]). En algunas otras ocasiones es posible medir los parámetros internos de la cámara en laboratorio al igual que los parámetros de distorsión (Ver [Holland97], [Zhang98] y [Zhang00]), y simplemente actualizar los parámetros de orientación externa de la cámara cada vez que sea necesario, siempre y cuando se mantenga constante la configuración de la cámara: Tamaño del lente, enfoque, tamaño de la imagen, etc. A este tipo de estrategia se le denominará *calibración en dos etapas*.

En este capítulo se abordará en primera instancia una estrategia que permite estimar los parámetros internos de la cámara y los coeficientes de distorsión, se presentarán las

estrategias para compensar o remover la distorsión de una imagen y por último se revisarán algunos métodos para estimar los parámetros de orientación externa de la cámara.

3.1 Estimación de los parámetros internos y coeficientes de distorsión

En esta sección se estudiarán estrategias para calcular los parámetros de la matriz de calibración interna, K , junto con los coeficientes de distorsión de la cámara, y a éste conjunto de parámetros se le conoce como *Parámetros Intrínsecos*. Principalmente se pretende estudiar una estrategia basada en el método de mínimos cuadrados no lineales, el cual se presenta en [Kelley99] y [Marquardt63]. Esta estrategia se presenta en [Hartley03] y [Zhang00]. La estimación de los coeficientes de distorsión tal y como aparecen en la ecuación (2.14) puede llevarse a cabo usando distintos patrones de calibración, los cuales pueden ser tridimensionales (Ver [Heikilla97]) o planares como los usados en la estrategia aquí presentada, un ejemplo del segundo tipo de patrón de calibración puede verse en la figura 2-8.

La estrategia presentada en [Zhang00] requiere la captura con la cámara de un patrón planar en distintas orientaciones. Este patrón consiste de un grupo de rectángulos blancos y negros en una configuración análoga a un tablero de ajedrez. Por simplicidad, es recomendable que todos los rectángulos tengan la misma dimensión e, incluso, sean cuadrados. En la figura 5 se presentó un ejemplo de dicho tipo de patrón.

3.1.1 Ecuaciones básicas

Sean (u, v) las coordenadas en la imagen correspondientes a un punto con coordenadas (x, y, z) en el espacio, se usará \tilde{g} para representar el vector $g = \begin{bmatrix} u & v \end{bmatrix}^T$ aumentado al agregar 1 como último elemento: $\tilde{g} = \begin{bmatrix} u & v & 1 \end{bmatrix}^T$ y de forma análoga $\tilde{G} = \begin{bmatrix} x & y & z & 1 \end{bmatrix}^T$ representará un vector de coordenadas en el espacio aumentado con un 1 como último ele-

mento. De acuerdo a esta notación y a la ecuación (2.9), la proyección g del punto G en el plano de la imagen está dada por:

$$w\tilde{g} = K [R|t] \tilde{G} \quad (3.1)$$

Sin pérdida de generalidad, se puede suponer que el plano modelo (el plano del patrón de calibración, como el que se presenta en la figura 2-8) se encuentra en el plano XY del sistema coordenado espacial, es decir, $Z = 0$ para todos los puntos sobre el patrón de calibración. Usando la notación $R = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}$ para representar la matriz de rotación R por medio de sus 3 columnas, es posible escribir la ecuación (3.1) como:

$$w \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} \vdots & \vdots & \vdots & \vdots \\ r_1 & r_2 & r_3 & t \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} x \\ y \\ 0 \\ 1 \end{bmatrix} = K \begin{bmatrix} \vdots & \vdots & \vdots \\ r_1 & r_2 & t \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3.2)$$

Haciendo un "inocente" abuso de notación al seguir usando G para denotar un punto en el plano modelo, se tiene:

$$w\tilde{g} = H\tilde{G} \quad (3.3)$$

con $H = K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}$ la *homografía* que relaciona los puntos en el patrón de calibración con sus correspondientes proyecciones sobre el plano de la imagen, $\tilde{g} = \begin{bmatrix} u & v & 1 \end{bmatrix}^T$ y $\tilde{G} = \begin{bmatrix} x & y & 1 \end{bmatrix}^T$. Como w es un factor de escala arbitrario, es claro que H está completamente definida excepto por un factor de escala. El nombre *homografía* es usado en geometría proyectiva y en este trabajo basta considerar una homografía como una matriz que permite realizar transformaciones entre imágenes.

En este trabajo, A^{-T} se usará como abreviatura para $(A^{-1})^T = (A^T)^{-1}$.

Denotando por $H = \begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix}$ la homografía entre un plano modelo y su imagen, donde h_i es la i -ésima columna de H , la ecuación (3.3) permite obtener la siguiente

expresión:

$$\begin{bmatrix} h_1 & h_2 & h_3 \end{bmatrix} = \lambda K \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \quad (3.4)$$

donde λ es un escalar arbitrario. Aprovechando el hecho que r_1 y r_2 son ortonormales, por ser columnas de la matriz de rotación R , se tiene:

$$h_1^T K^{-T} K^{-1} h_2 = 0 \quad (3.5)$$

$$h_1^T K^{-T} K^{-1} h_1 = h_2^T K^{-T} K^{-1} h_2 \quad (3.6)$$

Estas son las dos restricciones básicas para los parámetros internos dada una homografía (Ver [Zhang98]). Una homografía en este caso consta de 8 grados de libertad pero, debido a que se tienen 6 parámetros externos, solo es posible obtener 2 restricciones sobre los parámetros internos.

3.1.2 Estimación de los parámetros intrínsecos en forma cerrada

Sea

$$B = K^{-T} K^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{(fD_u)^2} & 0 & \frac{-u_0}{(fD_u)^2} \\ 0 & \frac{1}{(fD_v)^2} & \frac{-v_0}{(fD_v)^2} \\ \frac{-u_0}{(fD_u)^2} & \frac{-v_0}{(fD_v)^2} & \frac{u_0^2}{(fD_u)^2} + \frac{v_0^2}{(fD_v)^2} + 1 \end{bmatrix} \quad (3.7)$$

Claramente, B es una matriz simétrica, por lo que puede definirse por medio de un vector b de dimensión 6, de la forma:

$$b = \begin{bmatrix} B_{11} & B_{12} & B_{22} & B_{13} & B_{23} & B_{33} \end{bmatrix}^T \quad (3.8)$$

Si se escribe la i -ésima columna de H como $h_i = \begin{bmatrix} h_{1i} & h_{2i} & h_{3i} \end{bmatrix}^T$, se puede verificar que

$$h_i^T B h_j = v_{ij}^T b \quad (3.9)$$

donde

$$v_{ij} = \begin{bmatrix} h_{1i}h_{1j} & h_{1i}h_{j2} + h_{1i}h_{1j} & h_{2i}h_{2j} & h_{3i}h_{1j} + h_{1i}h_{3j} & h_{2i}h_{3j} + h_{3i}h_{2j} & h_{3i}h_{3j} \end{bmatrix}^T$$

Por consiguiente, las dos restricciones básicas (3.5) y (3.6) de una homografía dada pueden reescribirse como 2 ecuaciones homogéneas en b :

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (3.10)$$

Si se tienen n imágenes del plano modelo, apilando las n ecuaciones de la forma (3.10), se tiene el sistema:

$$Vb = 0 \quad (3.11)$$

Adicionalmente, con base en las ideas presentadas en [Hartley03] para la estimación de los parámetros internos a partir de la identificación de puntos de fuga y líneas de fuga (Vanishing points y vanishing lines, en inglés), se puede agregar al sistema dado en (3.11) una última ecuación que permita incluir la restricción $B_{12} = 0$, dicha ecuación se escribe como:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} b = 0 \quad (3.12)$$

Luego, se obtiene una matriz V de orden $(2n + 1) \times 6$. Si $n \geq 2$ y $\text{rango}(V) = 5$, se tiene una solución única b para (3.11) salvo por un factor de escala (Ver [Zhang98]). La solución de (3.11) es el vector propio de $V^T V$ asociado al menor valor propio (Equivalentemente, el vector singular derecho de V asociado con el menor valor singular). Es importante aclarar que en el caso que $\text{rango}(V) = 6$, la única solución para el sistema (3.11) es el vector $\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$, el cual no tiene ninguna utilidad para el proceso de calibración y, además, carece de interpretación física. En algunos casos es posible que agregar la ecuación (3.12) se traduzca en obtener $\text{rango}(V) = 6$, lo cual puede atribuirse a los errores en la toma

de los datos e , incluso, a que este paso no tiene en cuenta la distorsión. En [Hartley03] se propone solo usar el número mínimo de ecuaciones que garanticen $\text{rango}(V) = 5$, aunque esta estrategia no es práctica en el caso de contar con un número grande de imágenes, puesto que implicaría dejar de usar información de algunas imágenes. Bajo dicha situación, todavía es recomendable usar como solución para b , el vector singular derecho de V asociado a su menor valor singular aun cuando $\text{rango}(V) = 6$.

En algunos casos, es posible suponer que los píxeles son cuadrados, con lo que $fD_u = fD_v$, lo que se traduce en agregar al sistema dado en (3.11) la ecuación:

$$\begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \end{bmatrix} b = 0 \quad (3.13)$$

También en esta situación es recomendable usar la solución obtenida por medio de la descomposición de valores singulares de la matriz V .

La matriz B de la ecuación (3.7) coincide con la *imagen de la cónica absoluta*, un concepto de la geometría proyectiva bastante usado (Ver [Hartley03]), por lo cual es posible usar propiedades de la proyección en perspectiva para encontrar ecuaciones análogas a las usadas en el sistema (3.11) con el ánimo de estimar los parámetros de calibración interna. Una de estas propiedades es la existencia de puntos de fuga en la imagen, los cuales son los puntos donde se intersectan las imágenes de línea que son paralelas en el mundo real, como se muestra en la figura 3-1. Es claro que los puntos de fuga pueden caer dentro o fuera de la imagen, dependiendo de la perspectiva.

Para encontrar un punto de fuga es necesario contar con al menos dos línea en la imagen correspondientes a línea paralelas, pero es posible y recomendable hacer uso de un número mayor de línea, en especial cuando la imagen presenta algún tipo de distorsión. En dicho caso, es posible que las línea usadas no se crucen todas en el mismo punto debido a la distorsión y a los errores en las medidas y selección de las línea. Algunas estrategias para calcular el punto de fuga consisten en intersectar todas las línea por pares y usar el centroide de las intersecciones, o elegir el punto más cercano a las línea medidas (Ver



Figura 3-1: Ejemplo de un punto de fuga en una imagen. Imagen del edificio Bavaria en Cartagena, donde se encuentra instalada una estación de captura HORUS.

[Hartley03]). En cualquier caso, dichos procedimientos no son los más apropiados, por lo que en [Hartley03] se propone usar optimización de mínimos cuadrados para encontrar dicho punto.

El procedimiento para encontrar el punto de fuga por mínimos cuadrados consiste en encontrar un conjunto de líneas que se intersecten en un solo punto y que pasen cerca de los puntos extremos de los segmentos de línea medidos, de modo que se minimice la suma de los cuadrados de las distancias ortogonales de cada punto extremo a la línea correspondiente, un esquema de dicho conjunto se presenta en la figura 3-2, donde los segmentos medidos aparecen más oscuros y el punto de fuga encontrado aparece marcado como v .

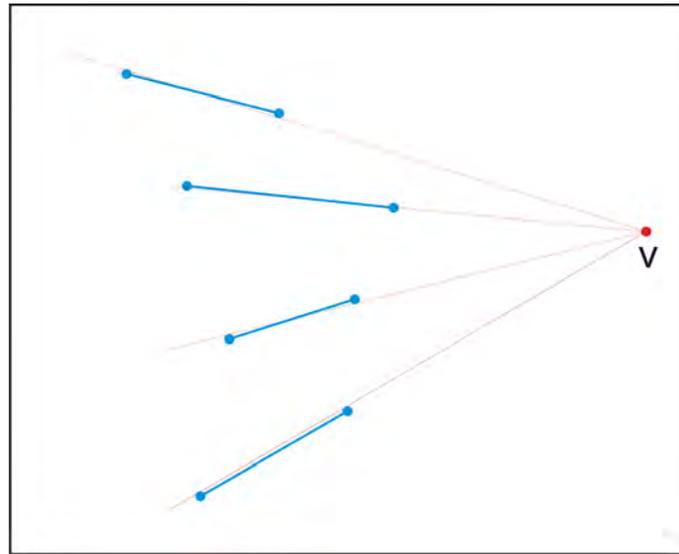


Figura 3-2: Segmentos de línea usados para encontrar el punto de fuga (Líneas azules), punto de fuga encontrado y conjunto de líneas estimadas (Líneas rojas).

Un par de puntos de fuga $v1 = [u_1 \ v_1 \ 1]^T$, $v2 = [u_2 \ v_2 \ 1]^T$ correspondientes a dos familias de línea paralelas que son ortogonales entre si, permiten construir la ecuación:

$$v1^T B v2 = \begin{bmatrix} u_1 u_2 & u_1 v_2 + u_2 v_1 & v_1 v_2 & u_1 + u_2 & v_1 + v_2 & 1 \end{bmatrix} b = 0 \quad (3.14)$$

Luego, si en cada una de las n imágenes tomadas al patrón de calibración se extraen 2 puntos de fuga correspondientes a líneas ortogonales, es posible apilar las ecuaciones para formar un sistema análogo al sistema (3.11) de n ecuaciones y 6 incógnitas. En el caso de ser necesario, pueden usarse las ecuaciones extras mencionadas anteriormente, con lo que se puede aumentar el número de ecuaciones del sistema. Nuevamente, la propuesta para resolver el sistema obtenido es usar descomposición de valores singulares.

Una vez que b ha sido estimado, es posible estimar los parámetros internos de la cámara así:

$$v_0 = \frac{-B_{23}}{B_{22}} \quad (3.15)$$

$$\lambda = B_{33} - \frac{[B_{13}^2 - v_0 B_{11} B_{23}]}{B_{11}} \quad (3.16)$$

$$fD_u = \sqrt{\frac{\lambda}{B_{11}}} \quad (3.17)$$

$$fD_v = \sqrt{\frac{\lambda}{B_{22}}} \quad (3.18)$$

$$u_0 = \frac{-B_{13} (fD_u)^2}{\lambda} \quad (3.19)$$

Con esto, la matriz K queda completamente determinada. Ahora, dada una imagen del plano modelo, en [Zhang00] se presenta una estrategia para estimar los parámetros externos de la cámara a partir de la ecuación (3.4), con lo que se obtiene:

$$r_1 = \lambda K^{-1} h_1 \quad (3.20)$$

$$r_2 = \lambda K^{-1} h_2 \quad (3.21)$$

$$r_3 = r_1 \times r_2 \quad (3.22)$$

$$t = \lambda K^{-1} h_3 \quad (3.23)$$

donde $\lambda = \frac{1}{\|K^{-1}h_1\|} = \frac{1}{\|K^{-1}h_2\|}$. Es claro que debido a los errores presentes en los datos usados (incertidumbres en la mediciones en tierra e incertidumbres en las marcaciones sobre las imágenes), la matriz R obtenida usando las ecuaciones (3.20) a (3.22) no satisface en general las propiedades de una matriz de rotación. En [Zhang00] se propone una estrategia para obtener la matriz de rotación R más "parecida" a una matriz general de orden 3×3 R^* . En dicha estrategia, "parecido" se usa para indicar que se desea minimizar la norma de Frobenius de la diferencia $R - R^*$.

La norma de Frobenius, $\|\cdot\|_F$ de una matriz A está definida como:

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{traza}(A^H A)} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2} \quad (3.24)$$

donde σ_i con $i = 1, \dots, \min\{m, n\}$ representa cada uno de los valores singulares de la matriz A . Teniendo esta definición en cuenta, el problema de obtener la matriz de rotación más "parecida" a una matriz R^* dada se puede escribir como:

$$\min_R \|R - R^*\|_F^2 \text{ sujeto a } R^T R = I \quad (3.25)$$

Se tiene que

$$\|R - R^*\|_F^2 = \text{traza} \left[(R - R^*)^T (R - R^*) \right] = 3 + \text{traza} \left[(R^*)^T R^* \right] - 2 \text{traza} (R^T R^*) \quad (3.26)$$

De modo que el problema dado en (3.25) es equivalente a maximizar la traza de $R^T R^*$.

Sea $USV^T = R^*$ la descomposición en valores singulares (SVD) de R^* , donde

$$S = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix}$$

Es claro que la matriz $Z = V^T R^T U$ es una matriz ortogonal. Usando la SVD de R^* se tiene que

$$\begin{aligned} \text{traza}(R^T R^*) &= \text{traza}(R^T U S V^T) = \text{traza}(V^T R^T U S) = \\ &= \text{traza}(Z S) = \sum_{i=1}^3 z_{ii} \sigma_i \leq \sum_{i=1}^3 \sigma_i \end{aligned} \quad (3.27)$$

Luego, $\text{traza}(R^T R^*)$ alcanza su valor máximo, $\sum_{i=1}^3 \sigma_i$, si $Z = I$, lo cual se cumple si $R = UV^T$, con lo que el problema dado en (3.25) queda resuelto.

3.1.3 Estimación de los parámetros de distorsión radial en forma cerrada

Algunos autores (Ver [Holland97], [Salvi02] y [Zhang00]) acostumbran despreciar los aportes de la distorsión por descentramiento y de prisma delgado en la ecuación (2.14), haciendo $k_3 = k_4 = k_5 = k_6 = 0$. Esta suposición es posible debido a que los sistemas de construcción y montaje de los lentes son cada vez más precisos, con lo que dichas distorsiones se eliminan desde fábrica. Bajo dicha condición, la ecuación (2.14) se convierte en

$$\begin{bmatrix} \delta_u(\tilde{u}, \tilde{v}) \\ \delta_v(\tilde{u}, \tilde{v}) \end{bmatrix} = [k_1 r^2 + k_2 r^4] \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} \quad (3.28)$$

Y esto permite escribir la relación (2.15), que describe la relación entre las coordenadas observables en la imagen (u, v) y las coordenadas teóricas (\tilde{u}, \tilde{v}) , como

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} [1 + k_1 r^2 + k_2 r^4] \tilde{u} \\ [1 + k_1 r^2 + k_2 r^4] \tilde{v} \\ 1 \end{bmatrix} \quad (3.29)$$

La ecuación (3.29) proporciona una expresión para las coordenadas observadas de un punto en la imagen, (u, v) . Por otro lado, la ecuación (2.5), que puede escribirse de la forma:

$$\begin{bmatrix} u^* \\ v^* \\ 1 \end{bmatrix} = K \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ 1 \end{bmatrix} \quad (3.30)$$

permite expresar las coordenadas ideales de un punto en la imagen, (u^*, v^*) , las cuales son no observables y libres de distorsión. Usando las ecuaciones (3.29) y (3.30), es posible obtener la siguiente expresión:

$$\begin{bmatrix} (u^* - u_0) r^2 & (u^* - u_0) r^4 \\ (v^* - v_0) r^2 & (v^* - v_0) r^4 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} u - u^* \\ v - v^* \end{bmatrix} \quad (3.31)$$

Cada punto observado en la imagen provee 2 ecuaciones para k_1 y k_2 , como muestra la ecuación (3.31). Ahora, si se tienen n puntos observados en p imágenes, es posible apilar np ecuaciones de la forma (3.31) para formar un sistema de la forma:

$$D \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = d \quad (3.32)$$

donde $D \in \mathbb{R}^{2np \times 2}$ y $d \in \mathbb{R}^{2np}$. La solución de mínimos cuadrados de (3.32) está dada por

$$\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = (D^T D)^{-1} D^T d \quad (3.33)$$

Con lo que quedan estimados los parámetros de distorsión de la cámara de forma lineal, normalmente estos estimados no son cercanos a los valores obtenidos usando métodos iterativos para su cálculo.

3.1.4 Estimación de los parámetros intrínsecos usando mínimos cuadrados

La solución obtenida en la sección anterior minimiza una distancia algebraica que no tiene sentido físico alguno. Es posible usar directamente toda la información de los puntos marcados en cada imagen del plano modelo para obtener una estimación de los parámetros internos de la cámara, incluso de los parámetros de distorsión, basada en mínimos cuadrados.

En este caso, se tienen p imágenes del plano modelo y hay n puntos marcados en el plano modelo. Es posible ([Zhang98]), suponer que los puntos en las imágenes están alterados por ruido independiente e idénticamente distribuido (Por ejemplo, ruido blanco). En el caso de contar con un conjunto de imágenes donde la distorsión es despreciable o ya ha sido corregida previamente, se puede estimar la matriz de calibración interna K al minimizar el funcional:

$$F(K) = \sum_{i=1}^p \sum_{j=1}^n \|g_{ij} - \hat{g}(K, R_i, t_i, G_j)\|^2 \quad (3.34)$$

donde $\hat{g}(K, R_i, t_i, G_j)$ es la proyección del punto $G_j = (x_j, y_j, z_j)$ en la imagen i usando la ecuación (2.9). El funcional $F(K)$ representa el error total, medido en píxeles, al estimar las proyecciones de los puntos G_i sobre cada imagen. Cada una de las matrices de rotación, R_i , se parametrizan con un vector de 3 parámetros, \mathbf{r}_i , el cual es paralelo al eje de rotación y cuya magnitud es igual al ángulo de rotación. Este vector \mathbf{r}_i se relaciona con R por medio de la fórmula de Rodrigues (Ver la sección 3.1.5 para detalles sobre la parametrización mencionada y la fórmula de Rodrigues).

También es posible usar la ecuación (3.34) alternadamente con la solución en forma cerrada de los parámetros de distorsión radial (Sección 3.1.3), para obtener mejores estimativos de todos los parámetros intrínsecos y de distorsión (Ver [Zhang00]). En este caso primero se resuelve (3.34), luego se estiman los parámetros de distorsión radial con la ecuación

(3.33), estos estimados se usan para volver a resolver (3.34) reemplazando $\hat{g}(K, R_i, t_i, G_j)$ por el \hat{g} obtenido con la ecuación (3.29), se vuelven a estimar los parámetros de distorsión radial con la ecuación (3.33) con los nuevos parámetros intrínsecos y así iterativamente.

El método de solución alternada propuesto puede presentar convergencia muy lenta (Ver [Zhang00]), por lo que es preferible estimar directamente todos los parámetros intrínsecos y de distorsión en cada paso del proceso iterativo, cambiando la expresión (3.34) por el funcional:

$$F(K, k_1, k_2, k_3, k_4, k_5, k_6) = \sum_{i=1}^n \sum_{j=1}^m \|g_{ij} - \hat{g}(K, k_1, k_2, k_3, k_4, k_5, k_6, R_i, t_i, G_j)\|^2 \quad (3.35)$$

Donde $\hat{g}(K, k_1, k_2, k_3, k_4, k_5, k_6, R_i, t_i, G_j)$ es la proyección del punto $G_j = (x_j, y_j, z_j)$ en la imagen i usando la ecuación (2.16). En cualquiera de los tres casos, el problema de minimizar (3.34) o (3.35) es un problema de minimización no lineal, que puede resolverse usando el algoritmo de Levenberg-Marquardt (Ver [Kelley99] y [Marquardt63]), el cual se presentó en el capítulo 1.4. En cada caso es necesario una aproximación inicial para K y $\{R_i, t_i | i = 1, \dots, n\}$ la cual puede ser obtenida usando la técnica descrita en la subsección anterior. En el caso de querer minimizar el funcional (3.35), también es necesario tener una estimación inicial de $(k_1, k_2, k_3, k_4, k_5, k_6)$ la cual puede tomarse como $(0, 0, 0, 0, 0, 0)$, en el caso de no contar con una más apropiada. En el caso descrito en la sección 3.1.3, es posible usar los estimados para k_1 y k_2 obtenidos al resolver la ecuación (3.33).

3.1.5 Matrices antisimétricas y de rotación, fórmula de Rodrigues

Sea H una matriz antisimétrica de orden 3×3 escrita como

$$H = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \quad (3.36)$$

Si se denota por \mathbf{r} al vector $\mathbf{r} = \begin{bmatrix} a & b & c \end{bmatrix}^T$, se tiene la siguiente proposición:

Proposición 3.1.1 *La matriz $R = e^H$ es una matriz de rotación. El vector \mathbf{r} es paralelo a su eje de rotación, y su magnitud $\|\mathbf{r}\|$ es igual al ángulo de rotación (mod 2π).*

Prueba. Para probar la proposición, se hace uso de los siguientes dos lemas:

Lema 3.1.2 *Sean A y B dos matrices tales que $AB = BA$. Entonces $e^A e^B = e^B e^A = e^{A+B}$.*

Lema 3.1.3 *Sea A una matriz cuadrada, entonces se cumple $\det(e^A) = e^{\text{traza}(A)}$.*

■

Prueba. Las pruebas de estos dos lemas se pueden encontrar en [Laub(05)]

Ahora es posible demostrar la proposición: Sea $R = e^H$. De la definición de matriz exponencial, se tiene que $(e^H)^T = e^{H^T}$. Adicionalmente se tiene que $H^T = -H$ (pues H es antisimétrica). Haciendo uso de estos hechos y del resultado del primer lema, como $-H = -IH$, se tiene: $RR^T = e^H (e^H)^{-1} = I$, con lo que R es ortogonal. Por otro lado, usando el segundo lema, $\det(R) = e^{\text{traza}(H)} = e^0 = 1$, con lo que queda probado que R es una matriz de rotación. Ahora, de la definición de H y \mathbf{r} se tiene que $H\mathbf{r} = 0$, y por tanto $R\mathbf{r} = \mathbf{r}$, con lo que queda probado que \mathbf{r} es paralelo al eje de rotación de R (Ver [Faugeras93]). Se sabe que los valores propios de una matriz de rotación son 1 , $e^{i\theta}$ y $e^{-i\theta}$, donde θ es el ángulo de rotación modulo 2π . Los valores propios de H son 0 , $i\|\mathbf{r}\|$ y $-i\|\mathbf{r}\|$. Se tiene que

$$e^{P_H^{-1}HP_H} = \begin{bmatrix} e^{\lambda_1} & q_{12} & q_{13} \\ & e^{\lambda_2} & q_{23} \\ & & e^{\lambda_3} \end{bmatrix} \quad (3.37)$$

Donde $\lambda_1, \lambda_2, \lambda_3$ son los valores propios de H y P_H es la matriz del teorema de **Schur** (Ver [Laub(05)]). Usando la ecuación (3.37) que permite diagonalizar H , se tiene que los valores propios de R cumplen $e^0 = 1$, $e^{i\|\mathbf{r}\|} = e^{i\theta}$ y $e^{-i\|\mathbf{r}\|} = e^{-i\theta}$ con lo que queda mostrado que $\theta = \|\mathbf{r}\| \pmod{2\pi}$. ■

Proposición 3.1.4 (Rodrigues) *La relación entre H y R puede ser escrita como*

$$R = e^H = I + \frac{\sin \theta}{\theta} H + \frac{1 - \cos \theta}{\theta^2} H^2 \quad (3.38)$$

Prueba. Se puede verificar que H satisface la siguiente relación:

$$H^3 = -\|\mathbf{r}\|^2 H = -\theta^2 H \quad (3.39)$$

De la relación (3.39) se tiene que

$$H^{2n-1} = (-1)^{n-1} \theta^{2(n-1)} H \quad \text{para } n = 1, 2, \dots$$

$$H^{2n} = (-1)^{n-1} \theta^{2(n-1)} H^2 \quad \text{para } n = 1, 2, \dots$$

Luego, usando la definición de matriz exponencial, se tiene:

$$\begin{aligned} R &= e^H = \lim_{k \rightarrow \infty} \sum_{n=0}^k \frac{1}{n!} (H)^n = \lim_{k \rightarrow \infty} \left[I + \sum_{n=1}^k \frac{1}{(2n-1)!} (H)^{2n-1} + \sum_{n=1}^k \frac{1}{(2n)!} (H)^{2n} \right] \\ &= \lim_{k \rightarrow \infty} \left[I + \sum_{n=1}^k \frac{(-1)^{n-1} \theta^{2(n-1)}}{(2n-1)!} H + \sum_{n=1}^k \frac{(-1)^{n-1} \theta^{2(n-1)}}{(2n)!} H^2 \right] \\ &= \lim_{k \rightarrow \infty} \left[I + \frac{H}{\theta} \sum_{n=0}^k \frac{(-1)^n \theta^{2n+1}}{(2n+1)!} + \frac{H^2}{\theta^2} \left(1 - \sum_{n=0}^k \frac{(-1)^n \theta^{2n}}{(2n)!} \right) \right] \\ &= I + \frac{\sin \theta}{\theta} H + \frac{1 - \cos \theta}{\theta^2} H^2 \end{aligned}$$

■

3.2 Corrección de la distorsión

Como se ha mencionado anteriormente, un modelo de distorsión de lentes puede ser suficientemente preciso si solamente se modela la distorsión radial (Ver [Holland97], [Salvi02])

y [Zhang00]), por lo que en esta sección se supondrá que solamente se presentan efectos debidos a este tipo de distorsión. Suponiendo que los parámetros internos de la cámara (Los parámetros de la matriz de calibración interna K) y los coeficientes de distorsión (Los cuales aparecen en la ecuación 2.14) han sido estimados, es posible compensar la distorsión radial presente en una imagen de dos formas: Calculando las coordenadas ideales de un punto, que están libres de distorsión, dadas las coordenadas medidas del punto, las cuales presentan distorsión. Este proceso exige encontrar ceros de polinomios, mientras que también puede corregirse la distorsión obteniendo una imagen sin distorsión a partir de la imagen distorsionada, para lo que es necesario usar interpolación en dos dimensiones.

3.2.1 Compensando la distorsión radial

En algunas aplicaciones, se cuenta con coordenadas (u, v) medidas en una imagen con distorsión y se desea usar el modelo de distorsión para encontrar las coordenadas ideales (u^*, v^*) libres de distorsión de la imagen, para luego calcular, por medio del modelo Pinhole sin distorsión, las coordenadas correspondientes en el espacio. Para llevar a cabo dicho proceso se usa la ecuación (3.29) para obtener la expresión:

$$K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ 1 \end{bmatrix} \quad (3.40)$$

Recordando que $r = \sqrt{\tilde{u}^2 + \tilde{v}^2}$ y llamando $\begin{bmatrix} \hat{u} & \hat{v} & 1 \end{bmatrix}^T = K^{-1} \begin{bmatrix} u & v & 1 \end{bmatrix}^T$, se tiene que la ecuación (3.40) puede reescribirse para obtener la expresión:

$$\sqrt{\hat{u}^2 + \hat{v}^2} = r + k_1 r^3 + k_2 r^5 \quad (3.41)$$

Claramente, la cantidad $\sqrt{\hat{u}^2 + \hat{v}^2}$ se puede calcular dado que se conocen las coordenadas (u, v) y la matriz de calibración interna K . De este modo, es posible hallar el valor de r

encontrando las raíces del polinomio $P(r) = k_2r^5 + k_1r^3 + r - \sqrt{\hat{u}^2 + \hat{v}^2}$. Dicho polinomio tiene grado 5, por lo que siempre tiene al menos una raíz real. En el caso que el polinomio $P(r)$ tenga una sola raíz real y esta sea positiva, dicha raíz es el valor de r y se pueden usar las ecuaciones (3.29) y (3.30) para obtener el valor de las coordenadas sin distorsión (u^*, v^*) :

$$\begin{bmatrix} u^* \\ v^* \\ 1 \end{bmatrix} = \left(\frac{1}{1 + k_1r^2 + k_2r^4} \right) \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3.42)$$

El problema de esta estrategia se puede presentar cuando el polinomio $P(r)$ tenga más de una raíz real positiva, con lo que debe plantearse una estrategia para seleccionar el valor adecuado para r . En principio, se podría pensar en usar siempre el valor positivo de r más cercano a cero, pero debido a la naturaleza de la distorsión radial es posible que dicho valor no sea el adecuado cuando se tienen coordenadas considerablemente alejadas del centro óptico de la imagen (Recordar que r es la distancia desde el centro óptico de la imagen y el punto sobre ésta en coordenadas métricas).

La estrategia propuesta en este trabajo consiste en no usar directamente la ecuación (3.42) para corregir la distorsión, sino realizar un paso intermedio que consiste en usar la ecuación (3.40) para encontrar las coordenadas $(\tilde{u}(r), \tilde{v}(r))$, las cuales son las coordenadas del punto sin distorsión en coordenadas métricas para cada uno de los valores reales positivos de r . Las coordenadas $(\tilde{u}(r), \tilde{v}(r))$ correctas deben estar suficientemente cerca de las coordenadas (\hat{u}, \hat{v}) que corresponden al punto con distorsión sobre el plano de la imagen en coordenadas métricas, puesto que la distorsión radial no debería, en general, aumentar la distancia entre el punto ideal y el punto distorsionado.

Para estudiar la validez de la estrategia propuesta se realizaron diversas pruebas de carácter cualitativo para estudiar los errores cometidos al compensar la distorsión. Primero se tomaron coordenadas en una imagen sin distorsión y se distorsionaron usando las ecuaciones (3.29) y (3.30) suponiendo distintos valores para la matriz de calibración interna y

los parámetros de distorsión. En la figura 3-3 se presenta el resultado gráfico de una de dichas pruebas usando los parámetros:

$$K = \begin{bmatrix} 1179 & 0 & 512 \\ & -1167 & 384 \\ & & 1 \end{bmatrix} \quad (3.43)$$

$$k_1 = -0.7 \text{ y } k_2 = 0.5 \quad (3.44)$$

Los resultados obtenidos con los distintos valores para los parámetros internos y los coeficientes de distorsión fueron similares y el error medio cuadrático de las correcciones fue del orden de 10^{-13} píxeles. En la figura 3-3 es posible notar que los puntos obtenidos al corregir la distorsión coinciden con los puntos originales, lo que sumado a un error de $5.38e-014$ píxeles en las correcciones permite concluir que la estrategia propuesta, a pesar de ser un poco más demandante computacionalmente, permite compensar de manera correcta la distorsión radial.

3.2.2 Eliminación gráfica de la distorsión

En el proceso de eliminación gráfica de la distorsión, lo que se desea obtener es una imagen que no presente distorsión, para tal fin se usa la ecuación (2.15) en la cual se tienen las coordenadas distorsionadas en la imagen, (u, v) , y la ecuación (3.32) que da una expresión para las coordenadas ideales libres de distorsión, (u^*, v^*) . En este proceso es posible eliminar tanto la distorsión radial como la distorsión por descentramiento y de prisma delgado, en el caso que estas últimas también se tengan modeladas y calibradas, debido a que en este caso el proceso consiste en calcular las coordenadas distorsionadas correspondientes a una imagen libre de distorsión, para de ese modo decidir qué valor de intensidad se le asignará a cada pixel de la imagen libre de distorsión.

Suponga que se tiene una imagen en escala de grises distorsionada $I_d(u, v)$, donde

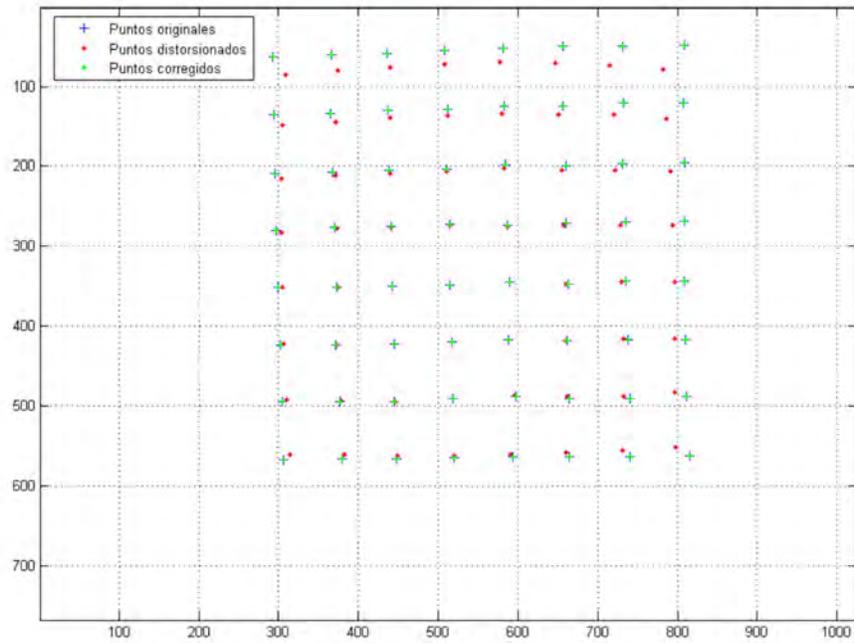


Figura 3-3: Puntos libres de distorsión en una imagen, puntos distorsionados y puntos con compensación de la distorsión obtenidos con la estrategia propuesta.

$I_d : \{1, 2, \dots, n\} \times \{1, 2, \dots, m\} \rightarrow [0, 255]$ representa la intensidad del pixel de coordenadas (u, v) . Una imagen sin distorsión $I(u^*, v^*)$, también es una función definida de $\{1, 2, \dots, n\} \times \{1, 2, \dots, m\}$ a $[0, 255]$, ahora, usando la ecuación (3.32) se tiene que:

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \\ 1 \end{bmatrix} = K^{-1} \begin{bmatrix} u^* \\ v^* \\ 1 \end{bmatrix} \quad (3.45)$$

Luego, usando la ecuación (2.15), se tendría que:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} \tilde{u} + \delta_u(\tilde{u}, \tilde{v}) \\ \tilde{v} + \delta_v(\tilde{u}, \tilde{v}) \\ 1 \end{bmatrix} = \begin{bmatrix} u^* \\ v^* \\ 1 \end{bmatrix} + K \begin{bmatrix} \delta_u(\tilde{u}, \tilde{v}) \\ \delta_v(\tilde{u}, \tilde{v}) \\ 1 \end{bmatrix} \quad (3.46)$$

donde $\delta_u(\tilde{u}, \tilde{v})$ y $\delta_v(\tilde{u}, \tilde{v})$ se calculan con la ecuación (2.14), usando las coordenadas (\tilde{u}, \tilde{v}) obtenidas con (3.45). La ecuación (3.46) relaciona las coordenadas ideales con las coordenadas distorsionadas, ahora se puede definir cuánto vale $I(u^*, v^*)$ para $(u^*, v^*) \in \{0, 1, 2, \dots, n\} \times \{0, 1, 2, \dots, m\}$ haciendo $I(u^*, v^*) = I_d(u(u^*, v^*), v(u^*, v^*))$. Este proceso se puede ilustrar en la figura 3-4.

Claramente, bajo este proceso es posible que las coordenadas u y v calculadas con (3.46) no sean números enteros, con lo que no corresponderían a la posición exacta de un pixel. Esto genera la necesidad de asignar un valor a $I_d(u(u^*, v^*), v(u^*, v^*))$ por medio de interpolación. Suponga que (r, s) es la coordenada entera más cercana a $(u(u^*, v^*), v(u^*, v^*))$, como se muestra en la figura 3-5, entonces se puede asignar el valor $I_d(u(u^*, v^*), v(u^*, v^*))$, por interpolación de vecino más próximo (*nearest neighbor* en inglés, ecuación (3.47)), bilineal (ecuación (3.48)) o bicúbica (Ver [Keys81]).

$$I_d(u, v) = I_d(r, s) \quad (3.47)$$

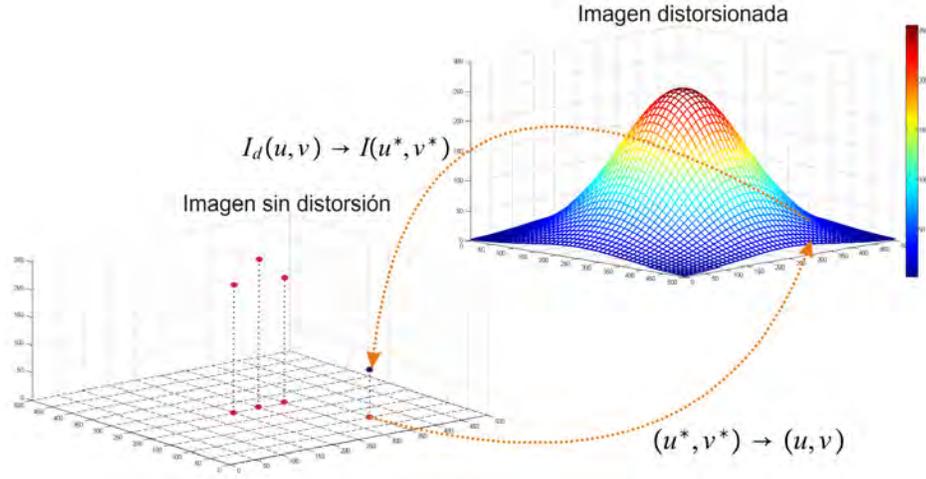


Figura 3-4: Proceso de creación de la imagen sin distorsión.

$$\begin{aligned}
 I_d(u, v) = & \frac{I_d(r-1, s-1)}{4} (r+1-u)(s+1-v) + \\
 & \frac{I_d(r+1, s-1)}{4} (u-r+1)(s+1-v) + \\
 & \frac{I_d(r-1, s+1)}{4} (r+1-u)(v-s+1) + \\
 & \frac{I_d(r+1, s+1)}{4} (u-r+1)(v-s+1)
 \end{aligned} \tag{3.48}$$

La figura 3-5 permite identificar las coordenadas relacionadas en las ecuaciones (3.47) y (3.48). En la figura 3-6 se presenta el resultado al corregir la distorsión de la imagen de la figura 2-8 usando el método aquí planteado con interpolación bilineal, puede observarse como las líneas que de antemano se sabe que son paralelas se observan efectivamente paralelas y que la escala de la imagen resultante es más pequeña (se perdió información de los extremos de la imagen original), debido a que la distorsión presente en dicha cámara es negativa lo que genera un estiramiento de la imagen en la corrección de la distorsión. En el caso de tener una imagen a color, por ejemplo en el sistema de color *RGB*, el proceso a realizar para eliminar la distorsión consiste en eliminar la distorsión de cada canal como si

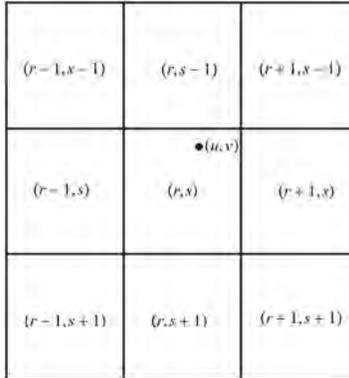


Figura 3-5: Píxeles alrededor de la coordenada (u, v) calculada.

fuera una imagen en escala de grises.

3.3 Estimación de los parámetros extrínsecos

La calibración del modelo Pinhole en dos etapas es muy útil cuando las cámaras están cambiando constantemente de orientación, con lo cual se hace necesario actualizar los parámetros extrínsecos continuamente, ejemplos de estas condiciones se evidencian en cámaras usadas en vigilancia o en las cámaras usadas para la navegación autónoma de robots por medio de sensores de video. En estos casos, se usan estrategias como las planteadas en la sección 3.1 para estimar los parámetros fD_u, fD_v, u_0, v_0 y, si es necesario, los parámetros de distorsión.

En otros casos, algunos de los parámetros de la orientación externa permanecen constantes mientras se mueve la cámara y, por lo tanto, solo es necesario recalcular los parámetros alterados. Un ejemplo de este hecho puede apreciarse cuando la cámara es rotada sobre si misma, dado que puede suponerse que el centro óptico permanece constante o que su desplazamiento es despreciable, mientras que los ángulos de orientación efectivamente deben cambiar ([Hartley03]).

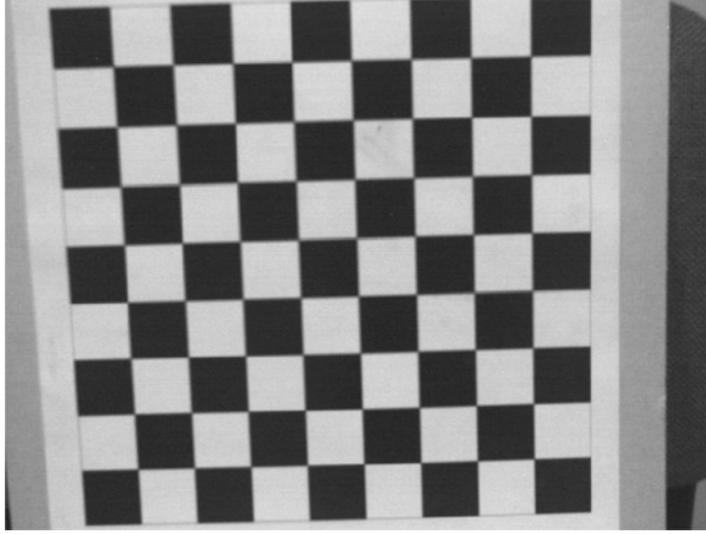


Figura 3-6: Imagen sin distorsión obtenida a partir de la imagen de la figura 2-8.

3.3.1 Solución geométrica: Problema PnP

En este caso, se suponen conocidos los parámetros intrínsecos de la cámara, D_u, D_v, f , y u_0, v_0 , y que la distorsión es despreciable o ha sido previamente corregida. Se desea encontrar la posición del centro óptico de la cámara, $C(x_c, y_c, z_c)$, y los ángulos de orientación τ, σ y ϕ , con los cuales es posible calcular la matriz de orientación externa, $[R|t] = R[I_3 | -C] \in \mathbb{R}^{3 \times 4}$. Estos parámetros aparecen explicados en la sección 2 y en la figura 2-3. Esto significa que se tienen 6 incógnitas y, como cada punto de control proporciona 2 ecuaciones, se necesitan como mínimo 3 puntos de control (GCP) para encontrar una solución. En [Hartley03] se propone como ejercicio mostrar que dados 3 GCP, es posible tener hasta 4 soluciones al problema, mientras que en [Fischler81] se propone una solución geométrica con 3 puntos de control que, efectivamente, puede conducir a 4 soluciones.

En [Fischler81] se da una definición del problema de Perspectiva con n Puntos (PnP):

Definición 3.3.1 *Dadas las ubicaciones espaciales relativas de n puntos de control, y da-*

dos los ángulos a cada par de puntos de control desde un punto adicional llamado el centro de perspectiva (C), encontrar las longitudes de los segmentos de línea que unen C con cada uno de los puntos de control.

Claramente, el problema PnP tiene solución si se cuenta con al menos 3 puntos de control. En la figura 3-7 se presenta un esquema de los datos disponibles en el caso del problema P3P, el GCP i tiene coordenadas (x_i, y_i, z_i) y su respectiva proyección en la imagen gcp i tiene coordenadas (u_i, v_i) , como ya se mencionó anteriormente, también se conoce la distancia focal, f , los parámetros de escalamiento, D_u, D_v , y la posición en píxeles de punto principal de la imagen, (u_0, v_0) . Este caso particular ha sido bastante estudiado debido a que, a pesar de poder presentar más de una solución, en muchas aplicaciones se desea usar el menor número de puntos de control posible, debido a que la cámara se está moviendo o la escena no cuenta con suficientes puntos claramente identificables en la imagen y/o fácilmente referenciables en el espacio.

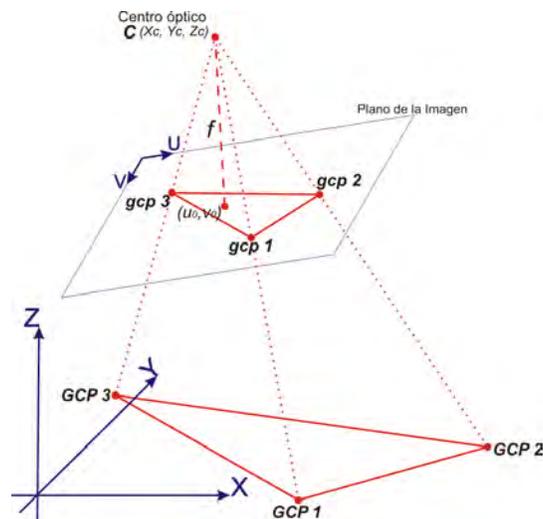


Figura 3-7: Información disponible para resolver la matriz de orientación externa con 3 GCP.

Las ideas presentadas [Fischler81] permiten encontrar una expresión algebraica para calcular las longitudes de los lados que unen cada punto de control con el centro óptico de la cámara, A, B, D . En la figura 3-8 se presentan dichas longitudes y los valores conocidos $\theta_{ab}, \theta_{ac}, \theta_{bc}, d_{12}, d_{13},$ y d_{23} , los cuales pueden calcularse haciendo uso de las coordenadas de los 3 puntos de control y de los parámetros internos de la cámara.

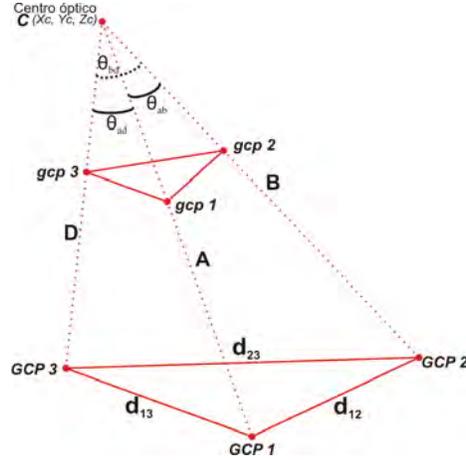


Figura 3-8: Esquema del problema geométrico a resolver en el caso P3P.

Definiendo las constantes $k_1 = \frac{d_{23}^2}{d_{13}^2}$ y $k_2 = \frac{d_{23}^2}{d_{12}^2}$, es posible obtener la ecuación polinómica

$$q_4x^4 + q_3x^3 + q_2x^2 + q_1x + q_0 = 0 \quad (3.49)$$

donde:

$$q_4 = k_2^2 + 2k_1k_2 - 2k_1k_2^2 + k_1^2 - 2k_1^2k_2 + k_1^2k_2^2 + 4k_1k_2 \cos^2(\theta_{bd}) \quad (3.50)$$

$$\begin{aligned}
q_3 &= -4k_1k_2 \cos(\theta_{ad}) \cos(\theta_{bd}) - \\
& 2k_1 \cos(\theta_{bd}) [4 \cos(\theta_{bd})k_2 \cos(\theta_{ab}) - 2k_1 \cos(\theta_{ad}) (1 - k_2)] + \\
& 4[(1 - k_1)(1 - k_2) - 1] (1 - k_1)k_2 \cos(\theta_{ab})
\end{aligned} \tag{3.51}$$

$$\begin{aligned}
q_2 &= 4k_1 \cos(\theta_{ad}) [2k_2 \cos(\theta_{bd}) \cos(\theta_{ab}) - k_1(1 - k_2) \cos(\theta_{ad})] - \\
& 2k_1 \cos(\theta_{bd}) [2(k_1 - k_2) \cos(\theta_{bd}) - 4k_1k_2 \cos(\theta_{ad}) \cos(\theta_{ab})] + \\
& [2(k_1 - 1)k_2 + 2k_1] [(1 - k_1)(1 - k_2) - 1] + 4(1 - k_1)^2 k_2^2 \cos^2(\theta_{ab})
\end{aligned} \tag{3.52}$$

$$\begin{aligned}
q_1 &= 4k_1 \cos(\theta_{ad}) [(k_1 - k_2) \cos(\theta_{bd}) - 2k_1k_2 \cos(\theta_{ad}) \cos(\theta_{ab})] - \\
& 4k_1^2 k_2 \cos(\theta_{bd}) \cos(\theta_{ad}) + [4(k_1 - 1)k_2 + 4k_1] (1 - k_1) \cos(\theta_{ab})
\end{aligned} \tag{3.53}$$

$$q_0 = [(k_1 - 1)k_2 + k_1]^2 + 4k_1^2 k_2 \cos^2(\theta_{ad}) \tag{3.54}$$

Las raíces de la ecuación (3.49) se pueden obtener en forma cerrada ([Griffiths47]) o por medio de un método iterativo. Para cada raíz positiva de (3.49), es posible determinar un valor positivo para cada uno de los lados A y B . En [Fischler81] se presenta un ejemplo de una configuración de 3 GCP y centro óptico que permite la existencia de 4 soluciones al problema P3P y se propone un método que use al menos 4 GCP para garantizar una solución única (Problema PnP con $n \geq 4$).

En [Haralick94] se estudia la estabilidad numérica de las soluciones directas del problema P3P y se presenta un método analítico para garantizar una solución numericamente estable.

En [Wang94] se presenta el estudio de ciertas regiones en el espacio donde la solución del problema P3P es única cuando los 3 GCP forman un triángulo isósceles en el espacio: Suponga que los 3 GCP usados están dados por los puntos A , B y C , los cuales forman un triángulo isósceles, como en la figura 3-9. Trace un plano v que pase por A y sea perpen-

dicular al segmento de línea AB . Luego, trace el plano u que también pasa por A pero que es perpendicular al segmento AC . A continuación se trazan el plano q que pasa por B y es perpendicular al segmento AB , el plano p que pasa por C y es perpendicular al segmento AC . Por último se trazan los planos n y m los cuales pasan por B y C , respectivamente, y son perpendiculares al segmento BC . En [Wang94] se demuestra por medios geométricos que si la cámara (Para ser más exactos el centro óptico) se ubica en las regiones III , IV , V , VI y VII entonces se tiene una única solución para el problema P3P, también en este mismo artículo se referencia un trabajo que, bajo ciertas condiciones, garantiza la existencia de una única solución al problema P3P cuando la cámara se encuentra en las regiones no cerradas I y II , dicho trabajo (Zhou Xin, Zhu Feng, “A note on unique solution conditions of the P3P problem,” Chinese Journal of Computers, vol. 26, no. 12, pp. 1696-1671, 2003) no se referencia aquí por encontrarse escrito en Mandarín.

En [Xiao03] se presenta una solución algebraica basada en la *descomposición en ceros de Wu-Ritt* para obtener una completa descomposición triangular del problema P3P, también se usa una estrategia geométrica para solucionar el problema y se da una completa clasificación de los casos en los que el problema presenta 1, 2, 3 o 4 soluciones, con lo que logran construir un método llamado *CASSC* que puede usarse para encontrar soluciones completas y numéricamente robustas del problema P3P.

3.3.2 Estimación de los parámetros extrínsecos con mínimos cuadrados

Teniendo presente que el interés principal de este trabajo es usar métodos de optimización no lineal para calibrar cámaras, la solución geométrica presentada en la sección anterior no es la de mayor interés, por lo que en esta sección se planteará una estrategia para convertir el problema PnP en un problema de mínimos cuadrados para ser resuelto usando el algoritmo de Levenberg-Marquardt. Nuevamente, se supondrá que se tienen n puntos de control (GCP por sus siglas en inglés) y que el GCP i tiene coordenadas (x_i, y_i, z_i) y su respectiva proyección en la imagen gcp i tiene coordenadas (u_i, v_i) , de igual forma, se

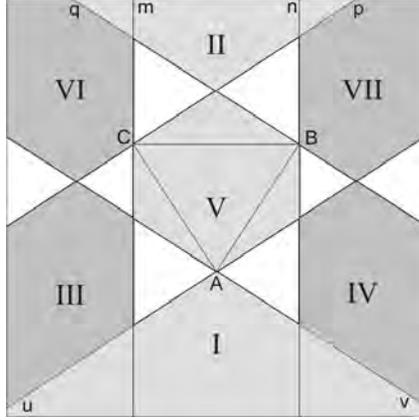


Figura 3-9: Configuración de las regiones con solución única del problema P3P dadas en [Wang94] y [Xiao03].

supondrán conocidos los parámetros de escalamiento, fD_u, fD_v , y la posición en píxeles de punto principal de la imagen, (u_0, v_0) . En este caso, se asumirá que la distorsión es despreciable o que fue previamente corregida.

Para escribir el problema PnP como un problema de mínimos cuadrados, primero se usará la fórmula de Rodrigues para parametrizar la matriz de rotación R con el vector $r = [a \ b \ c]^T$, donde se tiene que r es paralelo al eje de rotación y que $\|r\| = \theta(\text{mod } 2\pi)$, con θ el ángulo de rotación alrededor de r . Al construir la matriz antisimétrica H de la ecuación (3.36), se puede usar la fórmula (3.38) para obtener una expresión para R en términos de a, b y c y sustituirla en la ecuación (2.9), junto con $t = [t_x \ t_y \ t_z]^T$, con lo que se encuentra (\hat{u}_i, \hat{v}_i) , el estimado de (u_i, v_i) , la proyección en la imagen del punto de coordenadas (x_i, y_i, z_i) :

$$\begin{aligned}\hat{u}_i &= f_{(x_i, y_i, z_i)}(a, b, c, t_x, t_y, t_z) \\ \hat{v}_i &= f_{(x_i, y_i, z_i)}(a, b, c, t_x, t_y, t_z)\end{aligned}\tag{3.55}$$

Usando la expresión (3.55) para cada uno de los n puntos de control, es posible construir

una función $F : \mathbb{R}^6 \rightarrow \mathbb{R}$ de la siguiente forma:

$$R_\epsilon(a, b, c, t_x, t_y, t_z) = \begin{bmatrix} \hat{u}_1 - u_1 \\ \hat{v}_1 - v_1 \\ \vdots \\ \hat{u}_n - u_n \\ \hat{v}_n - v_n \end{bmatrix} \quad (3.56)$$

El vector $R_\epsilon(a, b, c, t_x, t_y, t_z)$ es el vector de errores de reproyección de los puntos de control sobre la imagen en cada coordenada. Luego, el problema PnP puede replantearse como el problema de minimizar la función:

$$F(a, b, c, t_x, t_y, t_z) = R_\epsilon(a, b, c, t_x, t_y, t_z)^T R_\epsilon(a, b, c, t_x, t_y, t_z) \quad (3.57)$$

En este caso, $\sqrt{\frac{F(a, b, c, t_x, t_y, t_z)}{2n}}$ representa el error medio cuadrático de las reproyecciones en cada coordenada de los puntos de control usando los parámetros (a, b, c, t_x, t_y, t_z) . Claramente, para poder resolver este problema de forma correcta se necesitan al menos 3 puntos de control ($n \geq 3$) pero se recomienda usar un número mayor cuando las mediciones sobre tierra y/o sobre la imagen estén sometidas a errores.

Al encontrar el conjunto de parámetros (a, b, c, t_x, t_y, t_z) que minimiza (3.57) se necesita luego obtener la matriz de rotación R , por medio de la fórmula de Rodrigues (3.38), para luego calcular los ángulos de rotación σ, ϕ, τ , a partir de la matriz de rotación R , usando la ecuación (2.2) y el vector C como $C = \begin{bmatrix} x_c & y_c & z_c \end{bmatrix}^T = -R^{-1}t$.

Para obtener una primera aproximación de los parámetros (a, b, c, t_x, t_y, t_z) se recomienda usar una transformación proyectiva, H , para dicha cámara y las ecuaciones (3.20) a (3.23) con los ajustes adecuados cuando H tenga 4 columnas en vez de 3 columnas. Dicha transformación se puede construir con al menos 6 puntos de control usando los métodos para estimar la Transformada lineal Directa que se presentan en el capítulo siguiente. Otra

estrategia para obtener la aproximación inicial es recurrir al método geométrico presentado en la sección anterior usando 3 o más puntos de control con la posibilidad, poco conveniente, de encontrarse con más de una posible solución.

Capítulo 4

Calibración del modelo Pinhole completo

En otros casos, las cámaras mantienen una posición y orientación constantes todo el tiempo, o al menos la orientación no cambia significativamente, ejemplo de esto son las cámaras usadas para el monitoreo de variables ambientales (Ver [Osorio07] para mayor información sobre este tipo de aplicaciones), en este caso puede ser de mayor utilidad calibrar todos los parámetros intrínsecos y extrínsecos de la cámara en el momento que esta ya ha sido ubicada y orientada. Para llevar a cabo este proceso, se acostumbra primero usar una aproximación lineal del modelo Pinhole, conocida como Transformada lineal Directa (DLT, por sus siglas en inglés), con la cual es posible obtener estimados iniciales para los parámetros del modelo Pinhole. En este capítulo se estudiará en primera instancia la DLT y otras homografías importantes, así como algunos de sus usos en la calibración de cámaras, para luego pasar a exponer las diferentes estrategias propuestas en la literatura para resolver el modelo Pinhole completo, teniendo en cuenta o no los efectos de la distorsión.

4.1 Transformada lineal Directa, DLT

En el caso de tener puntos en \mathbb{R}^3 con sus correspondencias en una imagen, el modelo Pinhole permite obtener la transformación que mapea puntos en el espacio con puntos en la imagen respetando las correspondencias dadas. En este caso, la transformación obtenida no es invertible pero en el caso del modelo Pinhole sin distorsión, se puede obtener una matriz H de $\mathbb{R}^{3 \times 4}$, la cual recibe el nombre de Transformada lineal Directa (DLT, Direct

linear Transform). La DLT se define a partir de (2.8) así:

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = KR[I_3 | -C] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = H \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.1)$$

Los métodos para encontrar H a partir de la correspondencia de puntos en el espacio (O en un plano) y puntos en una imagen, se conocen normalmente como métodos para solucionar la DLT. En [Hartley03] se presentan varios métodos para encontrar la DLT correspondiente a la transformación proyectiva entre dos planos (Una imagen respecto a otra o a un plano), con lo cual la matriz H pertenece a $\mathbb{R}^{3 \times 3}$ y puede ser llamada una homografía. La razón para estudiar esta transformación en particular es que los métodos para solucionarla son los mismos usados en el caso correspondiente a $H \in \mathbb{R}^{3 \times 4}$.

La primera estrategia para solucionar la DLT consiste en un método lineal el cual aparece construido en [Hartley03] tal y como se presenta en este trabajo. Suponga que se tienen n puntos en una imagen o un plano representados por los vectores $X_i = \begin{bmatrix} x_i & y_i & 1 \end{bmatrix}^T$ para $i = 1, 2, \dots, n$ y sus correspondencias en una imagen dadas por $Y_i = \begin{bmatrix} u_i & v_i & 1 \end{bmatrix}^T$. Entonces se quiere encontrar una transformación lineal, H , que satisfaga $Y_i = HX_i$ para todo $i = 1, 2, \dots, n$. Esta relación puede expresarse en términos del producto cruz como:

$$Y_i \times HX_i = 0 \quad (4.2)$$

Denotando la j -ésima fila de la matriz H como h^{jT} se tiene

$$HX_i = \begin{bmatrix} h^{1T} X_i \\ h^{2T} X_i \\ h^{3T} X_i \end{bmatrix} \quad (4.3)$$

Y se puede expresar el producto cruz dado en (4.2) como

$$Y_i \times HX_i = \begin{bmatrix} v_i h^{3T} X_i - h^{2T} X_i \\ h^{1T} X_i - u_i h^{3T} X_i \\ u_i h^{2T} X_i - v_i h^{1T} X_i \end{bmatrix} \quad (4.4)$$

Y de la ecuación (4.4) es posible encontrar un conjunto de 3 ecuaciones lineales en los elementos de H , de la forma

$$\begin{bmatrix} 0^T & -X_i^T & v_i X_i^T \\ X_i^T & 0^T & -u_i X_i^T \\ -v_i X_i^T & u_i X_i^T & 0^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (4.5)$$

Es importante tener presente que de estas 3 ecuaciones sólo 2 son linealmente independientes, por lo que cada correspondencia sólo aporta dos ecuaciones en los elementos de H , por lo que usualmente se acostumbra presentar el sistema (4.5) en la forma

$$\begin{bmatrix} 0^T & -X_i^T & v_i X_i^T \\ X_i^T & 0^T & -u_i X_i^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (4.6)$$

Y haciendo $h = [h^{1T} \quad h^{2T} \quad h^{3T}]^T \in \mathbb{R}^9$, esta ecuación puede escribirse como

$$A_i h = 0 \quad (4.7)$$

Ahora bien, si se tienen al menos 4 correspondencias entre un plano y la imagen, $n \geq 4$, es posible construir la matriz $A = [A_1 \quad A_2 \quad \dots \quad A_n]^T$ y con esta resolver para los elementos de H , encontrando una solución distinta de cero para $Ah = 0$. Es obvio que la solución $h = 0$ no es de interés práctico, por lo que es muy importante que la matriz A construida cumpla $\text{rango}(A) \leq 8$. El caso ideal es que $\text{rango}(A) = 8$, puesto que así las soluciones

diferentes de 0 para h son las pertenecientes al espacio nulo de A , el cual tiene en este caso dimensión 1. El vector generador del espacio nulo de A es, en este caso, el vector solución para los elementos de H , dado que H está unívocamente determinada por un factor de escala.

Debido a los errores en la medida de los puntos en la imagen y/o en el plano, es posible que $\text{rango}(A) < 8$ con lo que la solución para H no es exacta aparte de $h = 0$. En este caso, en vez de buscar una solución exacta es recomendable buscar una solución aproximada minimizando algún funcional. En [Hartley03] se propone minimizar la norma $\|Ah\|$ sujeta a la condición $\|h\| = 1$, de modo que la solución para h se calcula como el vector singular derecho de A asociado al valor singular más pequeño.

4.1.1 Soluciones no homogéneas para H

Otra estrategia para encontrar h es convertir el conjunto de ecuaciones lineales homogéneas de la forma (4.7) en un conjunto de ecuaciones no homogéneas forzando una condición de la forma $h_j = 1$, para alguno de los elementos del vector h . Forzar esta condición es justificable ya que H esta determinada exactamente excepto por un factor de escala. Algunos autores como [Holland97] y [Hartley03] proponen usar la condición $h_9 = H_{33} = 1$, de modo que las ecuaciones de la forma (4.6) se convierten en ecuaciones de la forma:

$$\begin{bmatrix} 0 & 0 & 0 & -x_i & -y_i & -1 & v_i x_i & v_i y_i \\ x_i & y_i & 1 & 0 & 0 & 0 & -u_i x_i & -u_i y_i \end{bmatrix} \tilde{h} = \begin{bmatrix} -v_i \\ u_i \end{bmatrix} \quad (4.8)$$

Donde \tilde{h} se obtiene con los primeros 8 elementos de h . Luego, se pueden concatenar n ecuaciones de la forma (4.8) para obtener un sistema con $2n$ ecuaciones y 8 incógnitas de la forma $M\tilde{h} = b$, el cual puede resolverse usando eliminación gaussiana en el caso que $M \in \mathbb{R}^{8 \times 8}$ y $\det(M) \neq 0$ o usando mínimos cuadrados en el caso de problemas con $n > 4$. En [Hartley03] se hace la aclaración que si la solución exacta es $h_j = 0$, entonces no existe un valor de escala k tal que $kh_j = 1$ y por lo tanto no podría encontrarse la solución

exacta usando el sistema no homogéneo. También en [Hartley03] se usan resultados de geometría proyectiva para mostrar que la normalización $h_9 = H_{33} = 1$ puede resultar en errores significativos en aplicaciones prácticas.

4.1.2 Solución de la DLT con puntos en el espacio y su correspondencias en la imagen

Cuando se cuenta con puntos en el espacio y no sobre un plano, es posible usar un método análogo al presentado en [Hartley03] para construir la DLT a partir de una correspondencia de al menos 6 puntos. La necesidad de por lo menos 6 correspondencias se debe a que en este caso la matriz H corresponde a la dada en la ecuación (4.1), la cual tiene 11 grados de libertad. En este caso, el punto i tiene coordenadas en el espacio, (x_i, y_i, z_i) , y este vector se representará en coordenadas homogéneas como $X_i^T = [x_i \ y_i \ z_i \ 1]^T$. Si los elementos de $H \in \mathbb{R}^{3 \times 4}$ son nuevamente representados por el vector $h = [h_1^T \ h_2^T \ h_3^T]^T = [h_{11} \ h_{12} \ h_{13} \ h_{14} \ h_{21} \ h_{22} \ h_{23} \ h_{24} \ h_{31} \ h_{32} \ h_{33} \ h_{34}]^T$, se tienen los siguientes equivalentes para las ecuaciones dadas en (4.6):

$$\begin{bmatrix} 0^T & -X_i^T & v_i X_i^T \\ X_i^T & 0^T & -u_i X_i^T \end{bmatrix} h = 0 \quad (4.9)$$

Nuevamente, las ecuaciones de (4.9) pueden representarse como $A_i h = 0$ para construir la matriz $A = [A_1 \ A_2 \ \dots \ A_n]^T$ y, si $n \geq 6$, se usa el vector singular derecho de A asociado al valor singular más pequeño como la solución del sistema $Ah = 0$. También en este caso es posible construir una sistema no homogéneo análogo al sistema dado con las ecuaciones de la forma (4.8), pero no es de interés en este trabajo debido a que no permite obtener resultados confiables en las aplicaciones prácticas.

Para este caso particular cuando $H \in \mathbb{R}^{3 \times 4}$ también es posible resolver para la DLT

minimizando la norma $\|Ah\|$ sujeta a la condición

$$\left\| \begin{bmatrix} h_{31} & h_{32} & h_{33} \end{bmatrix}^T \right\| = 1 \quad (4.10)$$

Dicha condición aparece al observar que $\begin{bmatrix} h_{31} & h_{32} & h_{33} \end{bmatrix}$ corresponde a la última fila de la matriz de rotación R asociada al modelo Pinhole, de donde se obtiene la DLT, por lo que dicho vector debe tener norma 1. En [Faugeras93] y [Hartley03] se presentan las estrategias para resolver el problema de minimización resultante, de las cuales es importante resaltar que el estimado lineal encontrado con el sistema de ecuaciones de la forma (4.9) es el punto de partida para minimizar el error de reproyección de la DLT teniendo como parámetros las componentes del vector h .

4.1.3 Consideraciones sobre el rango de la matriz A en el caso 3D a 2D

En el caso que se desee encontrar la DLT que represente una cámara proyectiva con puntos en el espacio, como en la sección anterior, se desea resolver un sistema homogéneo de $2n$ ecuaciones y 12 incógnitas, dado por

$$Ah = 0 \quad (4.11)$$

Donde $h \in \mathbb{R}^{12}$. Claramente, si $\text{rango}(A) = 12$ se tiene que la única solución para h es 0, lo cual no tiene significado físico alguno. En el caso que $\text{rango}(A) = 11$ se tiene que la solución para h es única salvo por un factor de escala, con lo que se tiene la solución para H .

En [Faugeras93] se presentan las siguientes dos proposiciones con su respectiva demostración:

Proposición 4.1.1 *Para n puntos ($n \geq 6$) en posición general, se tiene que $\text{rango}(A) = 11$.*

Proposición 4.1.2 *Para n puntos coplanares ($n \geq 4$) en posición general, se tiene que*

$\text{rango}(A) = 8$.

En las demostraciones de las proposiciones 4.1.1 y 4.1.2 en [Faugeras93], se define a qué se refiere el autor con el término *posición general*: Se refiere a que los puntos elegidos no caigan sobre una curva cúbica deformada (*twisted cubic*), la cual está definida por las dos ecuaciones

$$\begin{aligned}\lambda_1xz + \lambda_2yz + \lambda_3z^2 + \lambda_4z - \lambda_9x^2 - \lambda_{10}xy - \lambda_{11}x &= 0 \\ \lambda_5xz + \lambda_6yz + \lambda_7z^2 + \lambda_8z - \lambda_9xy - \lambda_{10}y^2 - \lambda_{11}y &= 0\end{aligned}\tag{4.12}$$

Donde $\lambda_i \in \mathbb{R}$. En [Faugeras93] también se hace la anotación que, normalmente, 7 puntos elegidos al azar no caen en una curva cúbica deformada, por lo que afirma que al tomar 6 puntos al azar y el centro óptico de la cámara se construye un conjunto de puntos que no caen en una cúbica deformada.

La proposición 4.1.2 permite realizar dos conclusiones importantes:

1. Si se está resolviendo la DLT para el caso general con puntos en el espacio, debe evitarse a toda costa usar más de 4 puntos coplanares, con el fin de evitar que $\text{rango}(A) < 11$.
1. En el caso de estar resolviendo la DLT para el caso que modelo la proyección de un plano sobre la imagen, al elegir 4 puntos al azar e incluir el centro óptico de la cámara se tiene que $\text{rango}(A) = 8$, por lo que la solución para la DLT será única.

4.1.4 Obteniendo los parámetros intrínsecos y extrínsecos a partir de la DLT

Recordando la ecuación (4.1), se tiene que

$$H = K [R|t]\tag{4.13}$$

Ahora bien, si se usa la notación r_i con $i = 1, 2, 3$ para denotar las filas de la matriz de rotación R se tiene que la Transformada lineal Directa puede escribirse como

$$H = K \begin{bmatrix} r_1 & t_x \\ r_2 & t_y \\ r_3 & t_z \end{bmatrix} = \begin{bmatrix} fD_u r_1 + u_0 r_3 & fD_u t_x + u_0 t_z \\ -fD_v r_2 + v_0 r_3 & -fD_v t_y + v_0 t_z \\ r_3 & t_z \end{bmatrix} = \begin{bmatrix} q_1^T & q_{14} \\ q_2^T & q_{24} \\ q_3^T & q_{34} \end{bmatrix} \quad (4.14)$$

Usando las ideas presentadas en [Faugeras93], es posible reconstruir todos los parámetros intrínsecos y extrínsecos dado que se cumpla la condición

$$\|q_3^T\| = 1 \quad (4.15)$$

Encontrar una matriz $H \in \mathbb{R}^{3 \times 4}$ que cumpla la condición (4.15) es un proceso que puede resolverse usando multiplicadores de Lagrange o mínimos cuadrados, como se muestra en [Faugeras93] y [Hartley03] respectivamente. Luego, si H cumple (4.15), la fila r_3 es igual a q_3^T salvo un factor de escala $\varepsilon = \pm 1$, de este modo se tiene que

$$r_3 = \varepsilon q_3^T \quad (4.16)$$

$$u_0 = q_1^T q_3 \quad (4.17)$$

$$v_0 = q_2^T q_3 \quad (4.18)$$

$$t_z = \varepsilon q_{34} \quad (4.19)$$

El hecho que $t_z = \varepsilon q_{34}$ se debe a que la matriz H calculada puede ser $\pm K [R|t]$. Ahora, si se calcula el cuadrado de la magnitud de q_1^T se tiene

$$\|q_1^T\|^2 = q_1^T q_1 = (fD_u)^2 r_1^T r_1 + 2u_0 fD_u r_3^T r_1 (u_0)^2 r_3^T r_3 = (fD_u)^2 + (u_0)^2 \quad (4.20)$$

De donde se obtiene

$$fD_u = \varepsilon_u \sqrt{q_1^T q_1 - u_0^2} \quad (4.21)$$

Donde $\varepsilon_u = \pm 1$. De forma análoga, se puede obtener la expresión

$$fD_v = \varepsilon_v \sqrt{q_2^T q_2 - v_0^2} \quad (4.22)$$

Donde igualmente se tiene $\varepsilon_v = \pm 1$. Ahora es posible obtener expresiones para r_1, r_2, t_x y t_y :

$$r_1 = \frac{\varepsilon (q_1^T - u_0 q_3^T)}{fD_u} \quad (4.23)$$

$$r_2 = \frac{-\varepsilon (q_2^T - v_0 q_3^T)}{fD_v} \quad (4.24)$$

$$t_x = \frac{\varepsilon (q_{14} - u_0 q_{34})}{fD_u} \quad (4.25)$$

$$t_y = \frac{\varepsilon (q_{24} - v_0 q_{34})}{-fD_v} \quad (4.26)$$

Se puede verificar que

$$\det(R) = \varepsilon \varepsilon_u \varepsilon_v \operatorname{signo} \left[\det \left(\begin{bmatrix} q_1^T \\ q_2^T \\ q_3^T \end{bmatrix} \right) \right] = 1 \quad (4.27)$$

Con la condición (4.27) se reduce el número de posibles conjuntos de valores para ε , ε_u y ε_v a 4 posibilidades. Es importante resaltar que las ecuaciones (4.16) a (4.27) solamente son válidas si se cumple la condición (4.15) y si se está trabajando con puntos en el espacio, es decir, $H \in \mathbb{R}^{3 \times 4}$.

También es posible que la matriz R obtenida con las ecuaciones (4.16), (4.23) y (4.24) no sea ortogonal debido a los errores en las mediciones en el espacio y en la imagen y también a que el método usado para calcular H no tiene en cuenta las propiedades de ortogonalidad de R , por lo que se debe usar la estrategia planteada en la sección 3.1.2 para encontrar la matriz ortogonal más "cercana" a la matriz R calculada.

4.1.5 La transformación afín y su aplicación

Los distintos métodos usados para calcular la DLT también pueden usarse para calcular otro tipo de transformaciones proyectivas como las transformaciones afines, las rotaciones entre imágenes o las traslaciones, por mencionar algunas transformaciones. En esta subsección se suponerá que se quieren realizar transformaciones entre dos imágenes, por lo que la matriz correspondiente a la transformación será $H \in \mathbb{R}^{3 \times 3}$. En el caso de tener dos o más imágenes de una escena que se solapen, es deseable *fusionar* dichas imágenes para obtener una sola imagen de la escena que contenga toda la información capturada por cada una de las imágenes originales. Este proceso de fusión consiste en encontrar una transformación entre cada par de imágenes y usar dicha transformación para hacer que ambas imágenes coincidan y puedan ser superpuestas sin que presenten diferencias importantes. La transformación usada en éste caso es la transformación afín, para la cual la matriz H tiene la estructura

$$H = \begin{bmatrix} A & t \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_u \\ a_{21} & a_{22} & t_v \\ 0 & 0 & 1 \end{bmatrix} \quad (4.28)$$

Donde $A \in \mathbb{R}^{2 \times 2}$ es una matriz no singular. Debido a que la transformación afín tiene una estructura bien definida, en la cual uno de sus componentes es igual a 1, es posible construir un sistema no homogéneo para encontrar los 6 parámetros que la determinan usando por lo menos 3 correspondencias entre puntos de la imagen. En este caso, se tiene que las coordenadas de uno de los puntos en la imagen 1 están dados por (u_i, v_i) , mientras que en la otra imagen se tiene que su correspondiente punto está dado por (u'_i, v'_i) . Así, de forma análoga a la que se construyeron las ecuaciones dadas en (4.8), es posible construir para el punto i las ecuaciones

$$A_i h = \begin{bmatrix} u'_i & v'_i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & u'_i & v'_i & 1 \end{bmatrix} h = \begin{bmatrix} -u_i \\ -v_i \end{bmatrix} \quad (4.29)$$

Donde $h = [a_{11} \ a_{12} \ t_u \ a_{21} \ a_{22} \ t_v]^T$. Construyendo la matriz $A = [A_1 \ A_2 \ \dots \ A_n]^T$, es posible encontrar h usando eliminación gaussiana o mínimos cuadrados lineales, según si se tiene $n = 3$ o $n > 3$.

Para llevar a cabo el proceso de *fusión* de dos imágenes, se encuentra la transformación afín, H , entre estas dos y se usa para transformar una imagen al sistema de referencia de la otra para luego solapar las imágenes obteniendo una nueva imagen, como se muestra en la figura 4-1.

En el caso de tener una tercera imagen a solapar, simplemente se fusiona esta tercera imagen con la imagen resultante de la primera fusión. En la figura 4-2 se presenta el resultado de usar transformaciones afines para fusionar 3 imágenes tomadas en Noviembre de 2008 con la estación de captura HORUS ubicada en Cartagena de Indias, usando cámaras con lentes de 8 mm y resolución de 1024 x 768 píxeles. La transformación afín permite fusionar imágenes sin importar el tamaño de las zonas que capturan, como puede corroborarse al ver la figura 4-3, la cual corresponde a la fusión de las mismas 3 cámaras usadas en el caso de la figura 4-2 pero con lentes de 4.8 mm, los cuales se instalaron en Mayo de 2009. En este caso, se logra capturar toda la zona de Bocagrande y aún cuando las

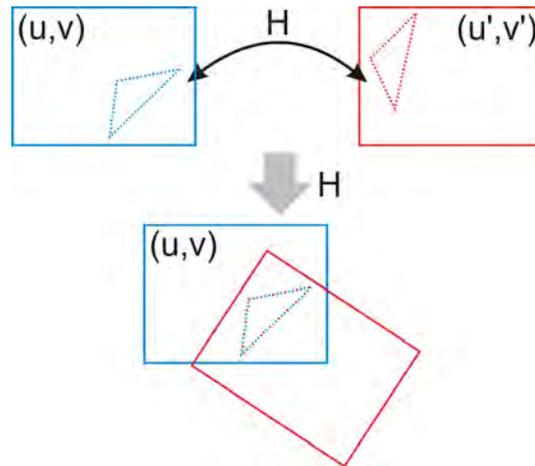


Figura 4-1: Dos imágenes relacionadas por una transformación afín H , y su fusión.

imágenes presentan distorsión radial los resultados de la fusión son satisfactorios.

En la figura 4-3 puede evidenciarse la distorsión radial causada por los lentes de 4.8 mm en la parte inferior de la imagen donde hay cercas que no aparecen rectas, ésta distorsión es un inconveniente para el proceso de fusión puesto que no hay garantía que las zonas de solape cambien de la misma forma en cada cámara, pero aún así los resultados son satisfactorios. Esto se debe a que las zonas de solape son tan amplias amplias que incluyen zonas con distorsión baja o nula, permitiendo ajustar las imágenes sin errores considerables. En base a ésta explicación se recomienda, en cualquier caso, intentar usar zonas amplias para el solape o eliminar la distorsión previamente a la estimación de las transformaciones afines.

4.2 Solución por mínimos cuadrados del modelo sin distorsión

Una condición importante para garantizar buenos resultados en un proceso de optimización no lineal es contar con una aproximación inicial que se encuentre lo suficientemente cerca del minimizador que se está buscando. Esta condición puede garantizarse al usar la DLT y



Figura 4-2: Panorámica de Bocagrande, imagen fusionada de 3 cámaras con lente de 8 mm de la estación de captura HORUS instalada en Cartagena de Indias, Colombia.

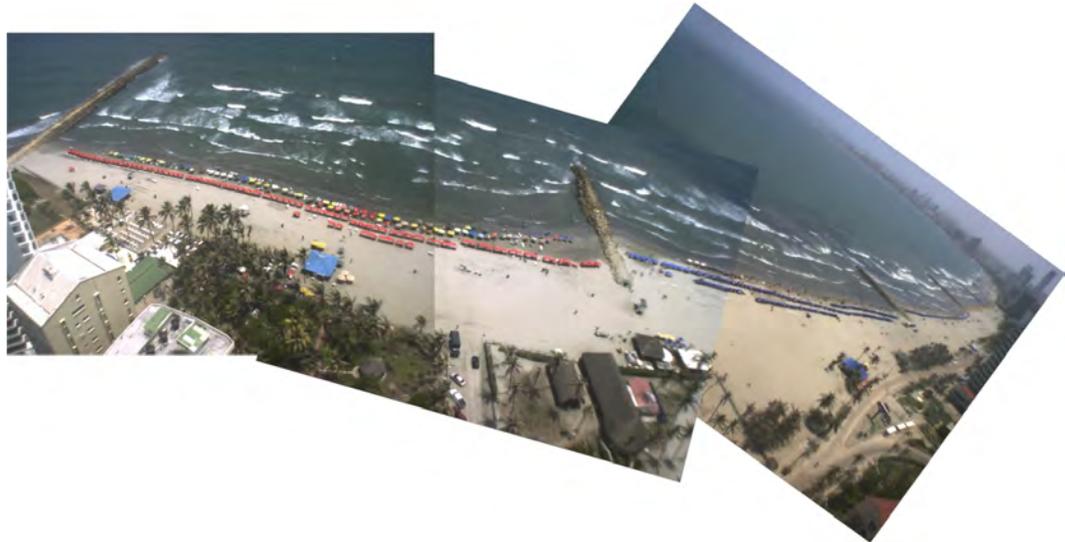


Figura 4-3: Panorámica de Bocagrande, imagen fusionada de 3 cámaras con lente de 4.8 mm de la estación de captura HORUS instalada en Cartagena de Indias, Colombia.

extraer los parámetros intrínsecos y extrínsecos como se propone en la sección 4.1.4, por lo que solo resta indicar en que forma se construye el funcional a minimizar usando el método de Levenberg-Marquardt presentado en la sección 1.4.

Nuevamente, se desea minimizar el error de reproyección de los puntos de control, por lo que se usarán las ecuaciones (2.6) y (2.7) luego de parametrizar la matriz de rotación R en términos de los elementos a, b, c del vector $r = \begin{bmatrix} a & b & c \end{bmatrix}^T$, por medio de la fórmula de Rodrigues presentada en la sección 3.1.5, con lo que se obtiene las expresiones para los errores de reproyección en las componentes u y v :

$$\epsilon_i^u = fD_u \left(\frac{r_{11}(x_i - x_c) + r_{12}(y_i - y_c) + r_{13}(z_i - z_c)}{r_{31}(x_i - x_c) + r_{32}(y_i - y_c) + r_{33}(z_i - z_c)} \right) + u_0 - u_i \quad (4.30)$$

$$\epsilon_i^v = -fD_v \left(\frac{r_{21}(x_i - x_c) + r_{22}(y_i - y_c) + r_{23}(z_i - z_c)}{r_{31}(x_i - x_c) + r_{32}(y_i - y_c) + r_{33}(z_i - z_c)} \right) + v_0 - v_i \quad (4.31)$$

Donde (u_i, v_i) son las coordenadas sobre la imagen del punto de control i cuyas coordenadas en el espacio son (x_i, y_i, z_i) y, claramente, $r_{ij} = g_{ij}(a, b, c)$ para $i, j \in \{1, 2, 3\}$. De este modo, haciendo uso de n puntos de control, se puede construir el vector de residuales R_ϵ como función de los parámetros $a, b, c, x_c, y_c, z_c, fD_u, fD_v, u_0, y v_0$:

$$R_\epsilon(a, b, c, z_c, y_c, z_c, fD_u, fD_v, u_0, v_0) = \begin{bmatrix} \epsilon_1^u \\ \epsilon_1^v \\ \epsilon_2^u \\ \vdots \\ \epsilon_n^v \end{bmatrix} \in \mathbb{R}^{2n} \quad (4.32)$$

Al observar el vector R_ϵ es claro que se necesitan al menos 5 puntos de control para garantizar la existencia de una solución al problema de mínimos cuadrados dado por

$$\min R_\epsilon^T R_\epsilon \quad (4.33)$$

Debido a que las estrategias para encontrar la DLT presentadas en la sección 4.1 exigen el uso de 6 puntos de control o más, en general se deberá contar con al menos 6 control y solucionar el problema de mínimos cuadrados sobredeterminado dado en la ecuación (4.33). En [Salvi02] se propone usar solamente la DLT y calcular los parámetros intrínsecos y extrínsecos usando la técnica presentada en la sección 4.1.4, mientras que en [Faugeras93] y [Hartley03] se propone minimizar el error de reproyección de la DLT directamente para luego encontrar los parámetros intrínsecos y extrínsecos como función de los 12 elementos de la DLT, aunque en [Faugeras93] también se plantea minimizar el error de reproyección directamente sobre los parámetros del modelo Pinhole.

4.3 Solución por mínimos cuadrados del modelo con distorsión

El modelo Pinhole sin distorsión es adecuado en aplicaciones que no necesitan un alto grado de precisión o donde dicha distorsión ya ha sido corregida de las imágenes (Por ejemplo, cuando la distorsión ha sido corregida graficamente usando métodos análogos al método presentado en la sección 3.2.2), pero en aplicaciones donde se necesita alta precisión y/o donde las cámaras no son de buena calidad es necesario tener en cuenta los efectos de la distorsión. En [Salvi02] se presenta el método cuyos autores denominan de *Faugeras con distorsión radial* para solucionar el modelo Pinhole suponiendo un sólo parámetro de distorsión radial, los autores presentan su solución usando el método de Newton obteniendo buenos resultados. En este caso, en vez de usar el error de reproyección sobre la imagen en píxeles, los autores usan el error de reproyección calculado en el plano de la imagen, bajo el sistema de referencia $\tilde{X}\tilde{Y}\tilde{Z}$ de la sección 2. Los autores de [Salvi02] no dan una justificación para usar dicho error de reproyección, aunque tampoco se encuentra en la literatura alguna razón para rechazar dicha opción. Otra diferencia significativa del método propuesto en [Salvi02] con el método que se presentará a continuación en esta sección es que los autores

no proponen usar la fórmula de Rodrigues para parametrizar la matriz de rotación, aunque este hecho es recomendado en [Faugeras93], [Zhang98] y [Zhang00].

La estrategia para construir el funcional a minimizar en el caso del modelo Pinhole con distorsión presentada en este trabajo se basará en la ecuación (2.16), teniendo presente que la matriz de rotación, R , nuevamente se parametrizará en términos de las componentes a, b, c del vector $r = [a \ b \ c]^T$ usando la fórmula de Rodrigues. En este caso, el vector de residuales R_ϵ es función de los parámetros $a, b, c, t_x, t_y, t_z, fD_u, fD_v, u_0, v_0, k_1, k_2, k_3, k_4, k_5$ y k_6 y es construido como

$$R_\epsilon = \begin{bmatrix} u_1 - \hat{u}_1 \\ v_1 - \hat{v}_1 \\ \vdots \\ u_n - \hat{u}_n \\ v_n - \hat{v}_n \end{bmatrix} \in \mathbb{R}^{2n} \quad (4.34)$$

Donde el punto de control i tiene coordenadas en el espacio (x_i, y_i, z_i) y coordenadas medidas en la imagen (u_i, v_i) y los estimados (\hat{u}_i, \hat{v}_i) cumplen

$$\begin{bmatrix} w\hat{u}_i \\ w\hat{v}_i \\ w \end{bmatrix} = K [R|t] \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} + K \begin{bmatrix} \tilde{w}\delta_u(\tilde{u}_i, \tilde{v}_i) \\ \tilde{w}\delta_v(\tilde{u}_i, \tilde{v}_i) \\ 0 \end{bmatrix} \quad (4.35)$$

Donde $\begin{bmatrix} \tilde{w}\tilde{u}_i \\ \tilde{w}\tilde{v}_i \\ \tilde{w} \end{bmatrix} = [R|t] \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix}$ y $\delta_u(\tilde{u}_i, \tilde{v}_i), \delta_v(\tilde{u}_i, \tilde{v}_i)$ están dados por la expresión (2.14).

Es importante aclarar que los valores w y \tilde{w} no son necesariamente los mismos para todos los puntos de control, aunque no aparezcan con el subíndice i en la ecuación (4.35)

En el caso de querer resolver el modelo Pinhole incluyendo los 6 parámetros de distorsión es necesario contar con al menos 8 puntos de control pero, debido a las consideraciones ya mencionadas en la sección 3.1.2, en este trabajo sólo se construirá el funcional $F = R_\epsilon^T R_\epsilon$ como función de los parámetros de distorsión radial, k_1 y k_2 , con lo que el número mínimo de puntos de control necesarios se reduce a 6, posibilitando usar el mínimo número de puntos de control para encontrar el estimado inicial de los parámetros intrínsecos y extrínsecos usando la DLT y las ecuaciones de la sección 4.1.2. En esta sección como en la anterior sección, no se presentan las versiones explícitas del vector R_ϵ ni su jacobiano R'_ϵ y tampoco se presentan las versiones explícitas del funcional F ni su gradiente, esto debido a que estas ecuaciones son bastante extensas y poco ilustrativas. Cabe aclarar que estas ecuaciones pueden obtenerse usando lenguajes para matemáticas simbólicas como *Maple*[®] (Ver www.maplesoft.com/).

Capítulo 5

Experimentos y resultados

Como ya se explicó al final del capítulo 1, para solucionar todos los problemas de mínimos cuadrados planteados en éste trabajo se hará uso del método de Levenberg-Marquardt debido a las ventajas que exhibe respecto a los métodos de Newton y del Gradiente. En primera instancia se presentaran resultados al usar el método de Levenberg-Marquardt para calcular puntos de fuga y para resolver el modelo Pinhole con o sin distorsión usando datos sintéticos, estos resultados permitirán validar el uso de dicho método y de los funcionales propuestos. Adicionalmente se presentarán resultados al agregar incertidumbre en los datos sintéticos con el ánimo de analizar la tolerancia al ruido de los algoritmos para luego usar los métodos en la solución del modelo Pinhole para el caso de las cámaras instaladas en la estación HORUS en Cartagena.

5.1 Cálculo de puntos de fuga

En la sección 3.1.2 se presentó el uso de los puntos de fuga en una imagen para estimar los parámetros intrínsecos de la cámara, en esta sección se mostrará la estrategia lineal y la no lineal para encontrar de manera acertada dichos puntos de fuga. Retomando las ideas de la sección 3.1.2, los parámetros que se desea encontrar para determinar el punto de fuga son sus coordenadas (u_f, v_f) y las pendientes de las rectas que pasan por los puntos que determinan cada uno de los segmentos de línea que determinan el punto de fuga como se muestra en la figura 3-2. En este caso lo que se desea minimizar es la distancia perpendicular de cada punto a la línea que pasa cerca de este. Para tal fin, se obtiene una expresión de dicha distancia en términos de la posición del punto de fuga y la

pendiente de la recta que pasa más cerca de dicho punto:

$$d_i^a = \sqrt{\left| \frac{m_i \left(\frac{u_i^a}{m_i} + v_i^a \right) - m_i (-m_i u_f + v_f)}{m_i^2 + 1} - u_i^a \right|^2 + \left| \frac{m_i^2 \left(\frac{u_i^a}{m_i} + v_i^a \right) - m_i u_f + v_f}{m_i^2 + 1} - v_i^a \right|^2} \quad (5.1)$$

$$d_i^b = \sqrt{\left| \frac{m_i \left(\frac{u_i^b}{m_i} + v_i^b \right) - m_i (-m_i u_f + v_f)}{m_i^2 + 1} - u_i^b \right|^2 + \left| \frac{m_i^2 \left(\frac{u_i^b}{m_i} + v_i^b \right) - m_i u_f + v_f}{m_i^2 + 1} - v_i^b \right|^2} \quad (5.2)$$

Donde (u_i^a, v_i^a) y (u_i^b, v_i^b) determinan el segmento de línea i , y la recta que aproxima dicho segmento tiene pendiente m_i y pasa por el punto de fuga (u_f, v_f) . El vector de residuales a construir tiene $2n$ ecuaciones y depende de $n + 2$ variables, así:

$$R_\epsilon(u_f, v_f, m_1, m_2, \dots, m_n) = \begin{bmatrix} d_1^a \\ d_1^b \\ \vdots \\ d_n^a \\ d_n^b \end{bmatrix} \in \mathbb{R}^{2n} \quad (5.3)$$

Para poder resolver el problema $\min R_\epsilon^T R_\epsilon$ usando el método de Levenberg-Marquardt primero es necesario contar con una aproximación inicial del vector $(u_f, v_f, m_1, m_2, \dots, m_n)$, para obtenerla se usa el método de mínimos cuadrados lineales, calculando cada pendiente como

$$m_i = \frac{v_i^b - v_i^a}{u_i^b - u_i^a} \quad (5.4)$$

Y el punto de fuga, (u_f, v_f) , se aproxima al resolver por mínimos cuadrados el sistema

$$\begin{bmatrix} m_1 & -1 \\ m_2 & -1 \\ \vdots & \vdots \\ m_n & -1 \end{bmatrix} \begin{bmatrix} u_f \\ v_f \end{bmatrix} = \begin{bmatrix} -m_1 u_1^a + v_1^a \\ -m_2 u_2^a + v_2^a \\ \vdots \\ -m_n u_n^a + v_n^a \end{bmatrix} \quad (5.5)$$

En la figura 5-1 se presenta una gráfica de un patrón de calibración y los 3 segmentos de línea usados para determinar un punto de fuga de la imagen.

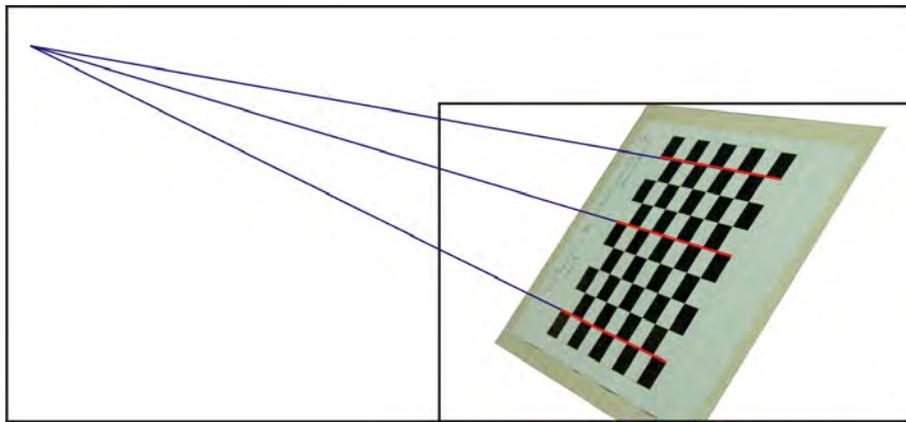


Figura 5-1: Patrón de calibración y punto de fuga determinado usando el método de Levenberg-Marquardt.

En la tabla **5.2.1** se presentan los resultados al aplicar el método de Levenberg-Marquardt para minimizar el funcional $R_\epsilon^T R_\epsilon$, puede observarse que los resultados obtenidos con el método lineal son suficientemente cercanos a los valores finales del método no lineal, aunque el error, $F(\beta^{(r)})$, en el primer paso de Levenberg-Marquardt tiene una disminución muy significativa.

**Tabla 5.2.1. Resultados del método de Levenberg-Marquardt
en la determinación de puntos de fuga.**

r	$\lambda^{(r)}$	$\beta^{(r)}$					$F(\beta^{(r)})$
		$u_f^{(r)}$	$v_f^{(r)}$	$m_1^{(r)}$	$m_2^{(r)}$	$m_3^{(r)}$	
0	–	–2589.43420	–410.13703	0.18276	0.30657	0.49004	625.860910
1	5	–2589.43280	–410.13874	0.18485	0.30283	0.49169	4.492460
2	2.5	–2589.43090	–410.14228	0.18485	0.30283	0.49170	4.491924
3	1.25	–2589.42730	–410.14886	0.18485	0.30284	0.49170	4.491852
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
20	19×10^{-8}	–2587.85220	–409.75388	0.18483	0.30286	0.49180	4.489924

En la tabla 5.2.1 puede observarse que el parámetro λ tiende a 0, lo cual es el comportamiento esperado cuando el método de Levenberg-Marquardt está convergiendo a un mimizador local, también es importante ver que en las primeras iteraciones el cambio en las coordenadas del punto de fuga fueron despreciables, lo que indica que el método lineal puede ser suficiente en aplicaciones que no necesiten alta precisión en las coordenadas del punto de fuga.

5.2 Estimación de los parámetros extrínsecos a partir de la DLT y la matriz de calibración interna

En esta sección se usarán las ecuaciones (3.20) a (3.23) para obtener un estimado lineal de la matriz de rotación R y el vector de traslación t , para luego usarlas como el estimado inicial en el proceso de minimización del funcional $F(a, b, c, t_x, t_y, t_z)$ dado en la ecuación (3.57). Los datos usados son los correspondientes a la imagen de la figura 5-2 en la cual se presenta una imagen del patrón de calibración a la cual se le ha removido la distorsión de manera gráfica. Se calculo su DLT a partir de los puntos marcados en la imagen, suponiendo que el extremo marcado como $(0, 0)$ corresponde al origen de coordenadas en

el espacio $(0, 0, 0)$, también puede observarse que todos los puntos son coplanares, por lo que la DLT calculada es una matriz 3×3 . La matriz de parámetros internos usada es:

$$K = \begin{bmatrix} 1229.665 & & 400 \\ & -1230.888 & 300 \\ & & 1 \end{bmatrix} \quad (5.6)$$

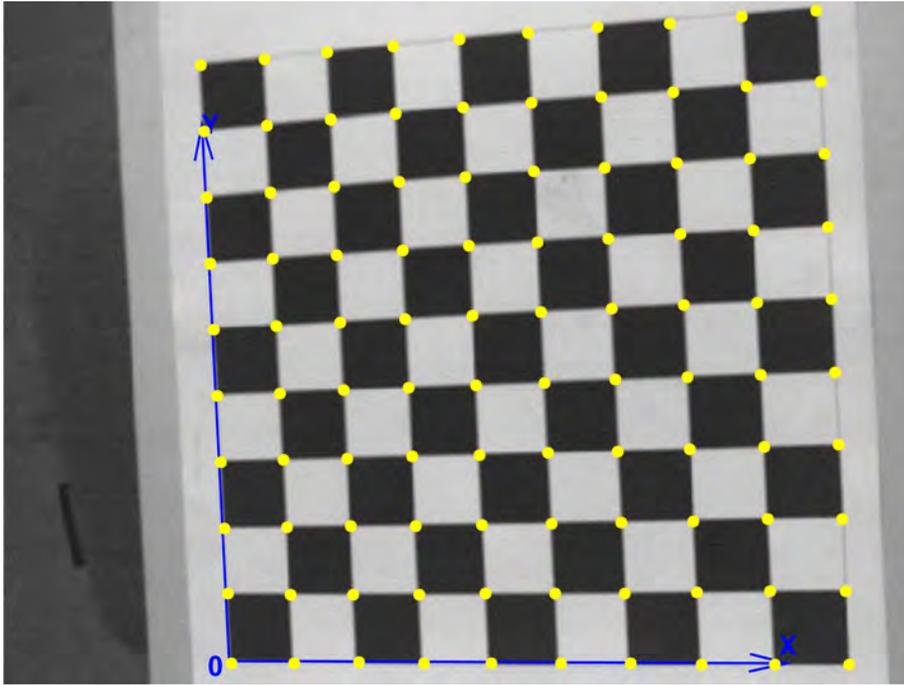


Figura 5-2: Patrón de calibración y puntos usados para calcular la DLT de la imagen.

La DLT obtenida usando el método de SVD propuesto en la sección 4.1 es:

$$H = \begin{bmatrix} 0.0028655435 & -0.0001524123 & 0.3253195545 \\ -0.0003124441 & -0.0031669341 & 0.9455930218 \\ -0.000000575 & 0.0000000116 & 0.0016318316 \end{bmatrix} \quad (5.7)$$

Al usar las ecuaciones (3.20) a (3.23) se obtiene la siguiente matriz R^* y el vector t :

$$R^* = \begin{bmatrix} 0.973926728381759 & -0.049421380346945 & 0.221974832899686 \\ 0.043955698793725 & 0.996460869490231 & 0.006595442184432 \\ -0.222563753309380 & 0.004521865547967 & 0.972652225891565 \end{bmatrix} \quad (5.8)$$

$$t = \begin{bmatrix} -103.0074255003998 \\ -143.3332364241051 \\ 631.2984680537493 \end{bmatrix} \quad (5.9)$$

Al hacer el producto $(R^*)^T (R^*)$ se puede ver que R^* no es una matriz ortogonal, por lo que se usa la estrategia dada en la sección **3.1.2** para obtener la matriz ortogonal más aproximada a la matriz R^* , obteniendo:

$$R = \begin{bmatrix} 0.973804787875467 & -0.046927056255536 & 0.222490643628011 \\ 0.046625525112510 & 0.998890563651087 & 0.006610768245465 \\ -0.222554028314024 & 0.003936145322807 & 0.974912412086952 \end{bmatrix} \quad (5.10)$$

La norma de Frobenius de la diferencia $R - R^{(*)}$ es $\|R - R^{(*)}\|_F = 0.004998580642437$. Al obtener la matriz R con el método lineal, se procede a parametrizarla en términos del vector $r = [a \ b \ c]^T$ por medio de la fórmula de Rodriguez, obteniendo:

$$r = \begin{bmatrix} 0.001349112669237 \\ -0.224486001449724 \\ -0.047189071660986 \end{bmatrix} \quad (5.11)$$

Ahora, con el vector $\beta^{(0)} = [r^T \ t^T]$ como estimado inicial, se usa el algoritmo de Levenberg-Marquardt para obtener mejores estimados, y los resultados se presentan en

la tabla **5.3.1**. En dicha tabla puede observarse que el vector de traslación no sufrió cambios significativos, mientras que el el vector r , que es paralelo al eje de rotación, si fue modificado fuertemente lo que causo cambios en los ángulos que determinan la matriz de rotación R , en este caso, se puede observar la fuerte disminución en el error de reproyección al usar el método de Levenberg-Marquardt. Las diferencias entre el método lineal y el método de Levenberg-Marquardt pueden explicarse al observar que el cálculo de la DLT no tiene en cuenta ninguna propiedad de la matriz H , lo que introduce error en el proceso de extracción de los parámetros del modelo Pinhole de dicha matriz. Adicionalmente, debido a esos errores, la matriz R estimada en principio no es ortogonal, lo que obliga a modificarla de modo que si lo sea, aumentando la desviación de los parámetros obtenidos a partir de la DLT. En el caso del método de Levenberg-Marquardt, los parámetros son estimados directamente usando el error de reproyección, por lo que se tiene muy en cuenta la estructura no lineal del modelo.

Tabla 5.3.1. Estimados de la matriz de rotación y del vector de traslación usando métodos lineales y no lineales			
Parámetro	Método lineal	Levenberg-Marquardt	Diferencia relativa (%)
r	$\begin{bmatrix} 0.00135 \\ -0.22449 \\ -0.04719 \end{bmatrix}$	$\begin{bmatrix} 0.00991 \\ -0.22036 \\ -0.047300 \end{bmatrix}$	$\begin{bmatrix} 86.38 \\ 1.872 \\ 0.234 \end{bmatrix}$
θ	13.143°	12.926°	1.684
τ	167.1389°	167.3627°	0.1337
σ	88.9868°	86.0638°	3.3962
ϕ	88.2981°	91.2151°	3.1979
t	$\begin{bmatrix} -103.00743 \\ -143.33324 \\ 631.29847 \end{bmatrix}$	$\begin{bmatrix} -103.21486 \\ -144.02935 \\ 632.88516 \end{bmatrix}$	$\begin{bmatrix} 0.2009 \\ 0.4833 \\ 0.2507 \end{bmatrix}$
$F(\beta)$	182639.691	106.7056	—

5.3 Estimación de los parámetros internos y de distorsión radial usando patrones de calibración

Para llevar a cabo el proceso de estimación de los parámetros internos y de distorsión radial usando patrones de calibración es necesario seguir los pasos dados en el algoritmo **5.4.1**.

Algoritmo 5.4.1. Para estimar los parámetros intrínsecos y de distorsión radial de una cámara usando un patrón de calibración planar como el presentado en la figura 5-2, es necesario:

1. Tomar p fotos del patrón de calibración en distintas orientaciones, p debe ser mayor o igual a 3.
2. Marcar o detectar automáticamente todos los n vértices del patrón de calibración en cada una de las p imágenes.
3. Suponiendo que el plano del patrón de calibración está sobre el plano $z = 0$, usar los vértices encontrados en el paso anterior para estimar las DLT (Homografías) de cada una de las imágenes.
4. Usar las DLT de las m imágenes y/o los puntos de fuga de cada imagen para estimar los parámetros internos con los métodos expuestos en la sección 3.1.2 y, si se desea, los parámetros de distorsión radial usando el método de la sección 3.1.3.
5. Usar los estimados del paso anterior para extraer las matrices de Rotación y los vectores de traslación de cada imagen, R_i, t_i con $i = 1, \dots, p$.
6. Usar los estimados de los parámetros intrínsecos, los parámetros de distorsión, las matrices R_i y los vectores t_i para optimizar los parámetros internos y de distorsión minimizando algún funcional, por ejemplo el presentado en la ecuación

(3.35). En este paso pueden usarse métodos no lineales como el Gradiente, Newton o Levenberg-Marquardt, también es posible usar métodos de dos pasos que minimizan en base a un subconjunto de parámetros y luego usan los nuevos estimados para minimizar respecto a los parámetros faltantes.

Para llevar a cabo cada uno de los pasos del algoritmo **5.4.1** se contruyó un toolbox de Matlab con su respectiva interfaz gráfica para facilitar la selección de imágenes, la extracción de los vértices y permitirle a un usuario final poder realizar calibraciones de distorsión sin necesidad de programar cada uno de los pasos. En la figura 5-3 se presenta la interfaz gráfica desarrollada, la cual se incorporó como una herramienta adicional del software que constituye el sistema HORUS.

El toolbox de calibración de HORUS utiliza un mecanismo de optimización en dos etapas: La primera etapa consiste en estimar los parámetros internos de la cámara para luego extraer los parámetros extrínsecos de cada imagen (R_i y t_i , para cada $i = 1, \dots, m$). La segunda etapa es el proceso de optimización del error de reproyección, donde se encuentran mejores aproximaciones de los parámetros intrínsecos y de distorsión radial. Terminada la segunda etapa, se repite el proceso de la primera, recalculando los parámetros extrínsecos de cada imagen, para luego volver a usar el método de optimización no lineal. En las primeras iteraciones del mecanismo de optimización se hace uso del método del Gradiente para aprovechar sus características de convergencia global, cambiando al método de Levenberg-Marquardt cuando los vectores de actualización calculados sean pequeños. Este método puede presentar estancamiento en mínimos locales que no sean soluciones aceptables del problema, hecho que obligaría a buscar estimados iniciales más adecuados, razón por la cual el toolbox de calibración cuenta con la posibilidad de definir los parámetros iniciales para el proceso de optimización.

En la actualidad es posible encontrar otros toolbox de calibración en internet, todos con sus fortalezas y debilidades frente al toolbox desarrollado en este trabajo, en particular se destaca el toolbox desarrollado por Jean-Yves Bouguet ([Bouguet]), el cual se encuentra



Figura 5-3: Interfaz gráfica de calibración de lentes HORUS.

disponible en la dirección http://www.vision.caltech.edu/bouguetj/calib_doc/index.html y funciona bajo matlab. Este toolbox tiene un funcionamiento similar al toolbox de este trabajo, con la diferencia de tener una interfaz menos gráfica y contar con un mecanismo de optimización que usa los parámetros intrínsecos de la cámara, los parámetros de distorsión y los parámetros extrínsecos asociados a cada imagen como el vector minimizador, lo que lo hace más robusto y reduce la posibilidad de estancamiento en mínimos locales. En la tabla 5.4.1 se presentan los resultados obtenidos al estimar los parámetros intrínsecos y de distorsión radial usando el toolbox de calibración de Bouguet y el toolbox HORUS, desarrollado en este trabajo. Puede observarse como los estimados de todos los parámetros son del mismo orden a excepción del estimado para k_2 , pero el hecho que ambos estimados sean cercanos a 0 indica que dicho parámetro no ejerce una influencia importante en ésta calibración. Es importante aclarar que al ser estos datos reales no es posible conocer la solución exacta del problema, si es que existe, y que por lo tanto la comparación de los resultados obtenidos con uno y otro método es la única herramienta que se tiene para validar dichos resultados.

Tabla 5.4.1. Resultados de la estimación de los parámetros internos y de distorsión radial usando dos diferentes toolbox		
Parámetro	Toolbox de Bouguet	Toolbox HORUS
fDu	1196.19607	1152.92556
fDv	1193.09726	1146.38759
u_0	420.14422	409.02738
v_0	298.21624	302.36306
k_1	-0.36662	-0.342904
k_2	0.05004	0.00403
$\sqrt{\frac{F(\beta)}{pn}}$	0.73519	0.85583
$\sigma(u_{ij})$	0.4818	0.6334
$\sigma(v_{ij})$	0.9225	1.0336

El valor $\sqrt{\frac{F(\beta)}{pn}}$ dado en la tabla 5.4.1 corresponde al error de reproyección medio cuadrático, mientras $\sigma(u_{ij})$ y $\sigma(v_{ij})$ corresponden a las desviaciones estandar de los errores de reproyección en las direcciones u y v , respectivamente. Puede observarse como el error obtenido con el toolbox de Bouguet es menor, sin ser muy alejado del error cometido con el toolbox HORUS. Para ver si el resultado entregado por Bouguet es el mejor estimado posible, se usaron sus resultados como estimados iniciales en el toolbox HORUS, obteniendo los resultados presentados en la tabla 5.4.2: El error de reproyección medio cuadrático disminuyo, al igual que las desviaciones en ambas direcciones, lo que indica que es posible obtener mejores resultados con el toolbox HORUS usando un estimado inicial cercano a una solución aceptable del modelo.

Tabla 5.4.2. Resultados de la estimación de los parámetros internos y de distorsión radial usando el toolbox HORUS	
Parámetro	Valor obtenido
fDu	1196.189386291161
fDv	1193.104593703543
$u0$	420.149688882363
$v0$	298.212730651266
k_1	-0.369663786425277
k_2	0.045542558896232
$\sqrt{\frac{F(\beta)}{pn}}$	0.735070689048368
$\sigma(u_{ij})$	0.4882
$\sigma(v_{ij})$	0.9189

5.4 Estimación del modelo Pinhole sin distorsión

En esta sección se usan los métodos presentados en la sección 4.1.4 para obtener estimados de los parámetros del modelo Pinhole a partir de la transformada lineal directa de la cámara y con estos estimados se inicializará el método de Levenberg-Marquardt para minimizar el

funcional de la ecuación (4.33). Para estudiar la validez y precisión de dicho procedimiento, primero se usarán datos sintéticos, contruidos usando los siguientes parámetros:

$$K = \begin{bmatrix} 1000 & 0 & 512 \\ & -1000 & 384 \\ & & 1 \end{bmatrix} \quad (5.12)$$

$$r = \begin{bmatrix} 1 \\ 1 \\ 0.4 \end{bmatrix} \quad (5.13)$$

$$C = \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix} \quad (5.14)$$

En este caso, la matriz de rotación R se calcula usando la fórmula de Rodrigues y el vector t se obtiene como $t = -RC$. Las coordenadas de los puntos de control usados se dan en la tabla **5.5.1**, los valores de u_i y v_i se obtuvieron usando la ecuación (2.9), de modo que estos valores no tienen error. Es importante resaltar que las coordenadas de dichos puntos sobre la imagen sólo abarcan una pequeña región sobre la imagen cuando se supone que ésta es de 1024 x 768 píxeles, éste hecho se debe a que se desea emular las condiciones que se encuentran en las aplicaciones en zonas costeras, donde sólo una pequeña porción de la imagen corresponde a playa y/o edificaciones que permitan marcar puntos de control.

Tabla 5.5.1. Datos de los puntos de control sintéticos

<i>GCP</i>	x_i	y_i	z_i	u_i	v_i
1	-280	670	1	212.5037	255.9812
2	-280	480	-1	56.903	410.1003
3	-47	450	3	209.9397	18.0808
4	-134	660	4	302.6338	75.0544
5	-159	560	4	218.4007	167.847
6	-188	570	3	203.9587	204.3205
7	-216	640	2	232.9359	197.1088

Al calcular la DLT usando el método de SVD propuesto en la sección 4.1.2 se obtiene la siguiente matriz H , la cual fue escalada para garantizar que $h_{31}^2 + h_{32}^2 + h_{33}^2 = 1$.

$$H = 10^5 \times \begin{bmatrix} -0.002558185 & -0.005772983 & -0.009292080 & 1.762324826 \\ 0.008830225 & 0.001932878 & -0.005747757 & -0.501534591 \\ 0.000005104 & -0.000008434 & -0.000001675 & 0.000500517 \end{bmatrix} \quad (5.15)$$

Usando el método de la sección 4.1.4 para estimar los parámetros del modelo Pinhole y la fórmula de Rodrigues para obtener el vector r se obtienen los valores presentados en la tabla 5.5.2, donde puede notarse que los errores son despreciables, lo cual era de esperarse pues los datos usados fueron generados sin considerar ningún tipo de incertidumbre.

Tabla 5.5.2 Parámetros del modelo Pinhole obtenidos a partir de la DLT			
Parámetro	Estimado	Exacto	Error relativo (%)
fDu	999.999999	1000	9×10^{-8}
fDv	1000	1000	0
$u0$	512	512	0
$v0$	384	384	0
τ	80.3559°	80.3559°	0
σ	58.8176°	58.8177°	1×10^{-4}
ϕ	148.8175°	148.8177°	2.5×10^{-4}
t_x	100	100	0
t_y	99.9999	100	1×10^{-5}
t_z	99.9999	100	1×10^{-5}

Para estudiar la precisión del método que minimiza el funcional de la ecuación (4.33) usando Levenberg-Marquardt, se usaron los mismos datos sin error, pero como estimado inicial se usó el vector $\beta^{(0)}$ dado en la tabla 5.5.3 y se obtiene el vector $\beta^{(217)}$ dado en la misma tabla. Puede observarse en dicha tabla que los errores son 0, resultado esperado al usar datos sin incertidumbre, de modo que el funcional de la ecuación (4.33) tiene como mínimo absoluto el valor 0.

Tabla 5.5.3 Parámetros del modelo Pinhole obtenidos con Levenberg-Marquardt				
Parámetro	$\beta^{(0)}$	$\beta^{(217)}$	Exacto	Error relativo (%)
fDu	76.5593	1000	1000	0
fDv	112.3980	1000	1000	0
$u0$	512	512	512	0
$v0$	384	384	384	0
τ	172.4839°	80.3559°	80.3559°	0
σ	61.4627°	58.8177°	58.8177°	0
ϕ	160.2610°	148.8177°	148.8177°	0
t_x	105.2634	100	100	0
t_y	278.1039	100	100	0
t_z	21.9885	100	100	0

En la figura 5-4 se presenta la evolución del parámetro λ y del error de reproyección $F(\beta)^T F(\beta)$, puede observarse como el valor $\lambda^{(r)} \rightarrow 0$ al igual que el error de reproyección, lo que era de esperarse teniendo en cuenta los resultados del 1.4.3.

Para estudiar la tolerancia al ruido del método propuesto se agregó a las coordenadas en la imagen de los puntos de control un número aleatorio distribuido normalmente con media 0 y desviación estandar entre 0 y 1.5, el cual puede interpretarse como la incertidumbre de las mediciones hechas sobre la imagen. En las figuras 5-5 y 5-6 puede observarse como cambiaron los estimados de cada parámetro con el método lineal y con el método de Levenberg-Marquardt a medida que aumentaba la incertidumbre de las mediciones en la imagen.

En la figura 5-5 se observa que el método no lineal es más robusto que el método lineal, incluso el método lineal presenta grandes variaciones en los estimados a medida que la incertidumbre de las mediciones en la imagen aumenta, hecho que puede explicarse al tener en cuenta que el método para calcular la DLT no tiene en cuenta las propiedades del

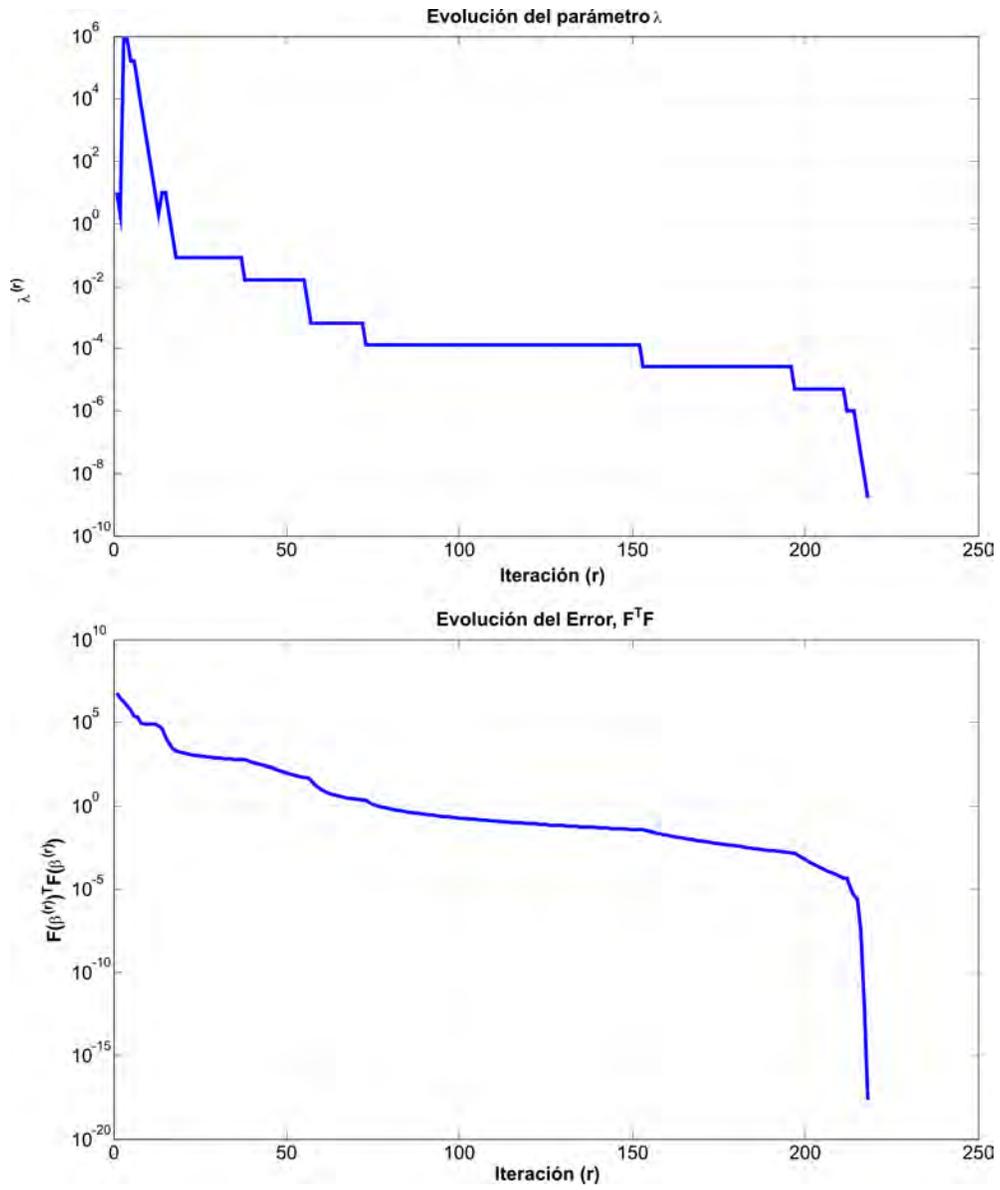


Figura 5-4: Evolución del parámetro λ (Arriba) y del error de reproyección $F(\beta)^T F(\beta)$ usando el método de Levenberg-Marquardt.

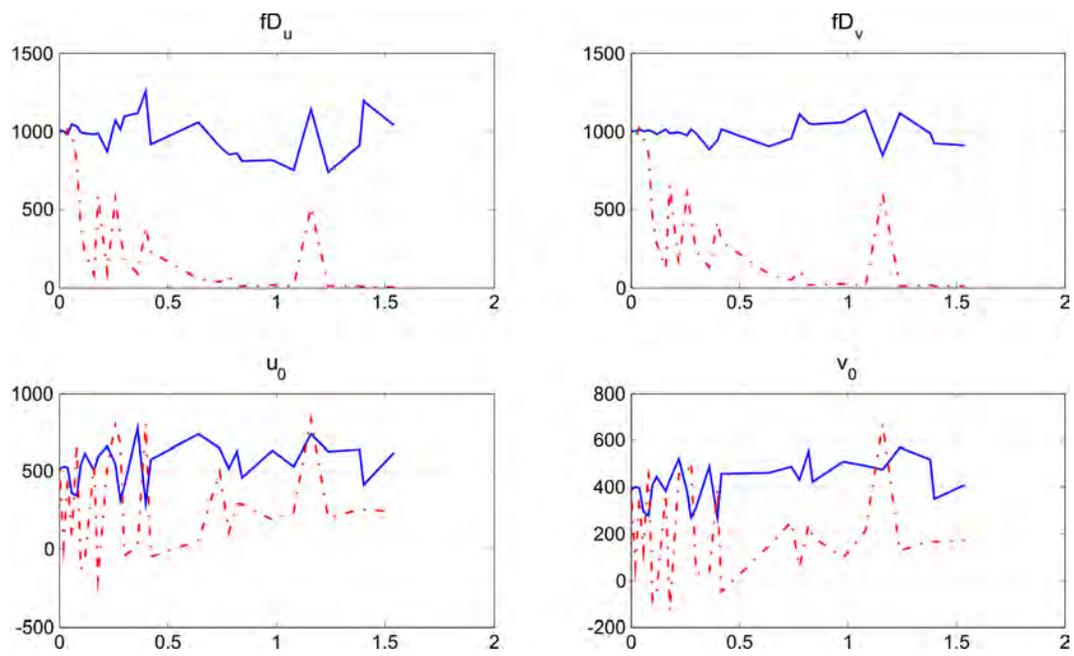


Figura 5-5: Variación de los parámetros intrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 7 puntos de control.

modelo Pinhole, solo busca encontrar la homografía que minimice el error de reproyección, sin prestarle importancia a la estructura de dicha matriz. También es notable que los parámetros que tienen mayor variación son los parámetros intrínsecos, lo que se puede explicar teniendo presente que un error en dichos parámetros puede ser compensado por los parámetros de traslación y rotación de la cámara.

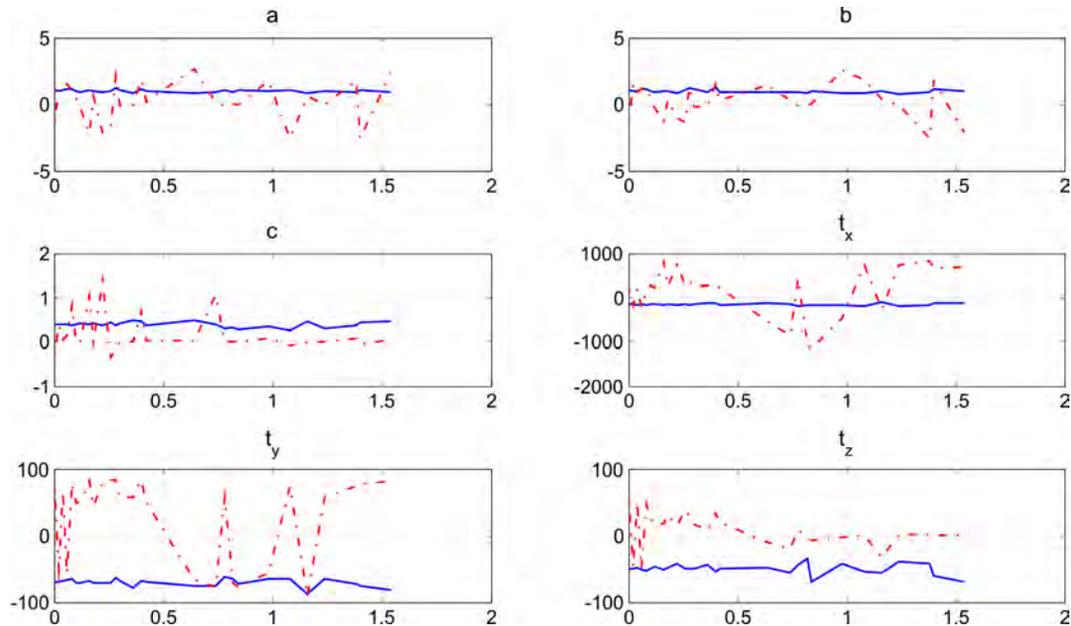


Figura 5-6: Variación de los parámetros extrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 7 puntos de control.

Para poder estudiar hasta donde los modelos obtenidos con el método lineal y con Levenberg-Marquardt son confiables, se calculó para cada valor de la incertidumbre el error de reproyección medio cuadrático, EMC , también las desviaciones estándar de las reproyecciones en las direcciones u y v , $\sigma(u)$ y $\sigma(v)$ respectivamente, y, por último, el *error de normalización calibrado* (NCE por sus siglas en inglés) que fue introducido en [Weng92] y fue propuesto como un método para comparar distintas calibraciones de una misma cámara en [Salvi02]. El error NCE se calcula usando la siguiente fórmula:

$$NCE = \frac{1}{n} \sum_{i=1}^n \left[\frac{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2}{z_i^2 (fD_u^2 + fD_v^2) / 12} \right]^{1/2} \quad (5.16)$$

Donde (x_i, y_i, z_i) es la posición exacta del punto de control i y (\hat{x}_i, \hat{y}_i) es la proyección de las coordenadas en la imagen de dicho punto de control, (u_i, v_i) , intersectada con el plano $z = z_i$. En [Salvi02] y [Weng92] se explica que si $NCE < 1$ entonces el error de la proyección es menor, en media, que el error en la digitalización, lo que es bastante deseable. Si $NCE \approx 1$ entonces se tiene una buena calibración. Valores de $NCE \gg 1$ revelan una calibración poco precisa. En el caso lineal, MCE , $\sigma(u)$ y $\sigma(v)$ se calcularon usando la matriz H correspondiente a la DLT y NCE se calculó con los parámetros del modelo Pinhole extraídos de dicha matriz, mientras que en el caso no lineal se usaron siempre los estimados de los parámetros obtenidos con el método de Levenberg-Marquardt. En la figura 5-7 se presenta la evolución de cada error en el caso lineal y en la figura 5-8 se presenta la evolución de cada error usando el método de Levenberg-Marquardt.

Puede observarse en la figura 5-7 que en la mayoría de los casos $NCE \gg 1$, lo que indica que los parámetros del modelo Pinhole obtenidos directamente de la DLT no permiten obtener una calibración aceptable, mientras que en la figura 5-8 se puede ver que $NCE \ll 1$ en todos los casos, por lo que puede decirse que los parámetros estimados con el método de Levenberg-Marquardt permiten obtener una calibración bastante confiable. Sin embargo, los errores obtenidos con la transformada lineal directa (Ver figura 5-7) son comparables con los errores obtenidos por el método no lineal, lo que indica que en el caso de aplicaciones donde se necesite calibrar la cámara pero no sea relevante tener explícitamente los parámetros del modelo Pinhole es aceptable usar solamente la transformada lineal directa.

Al observar las altas variaciones en los estimados de los parámetros en las figuras 5-5 y 5-6, es posible imaginar que el método propuesto en este trabajo para estimar los parámetros no es lo suficientemente confiable, pero una de las principales razones para

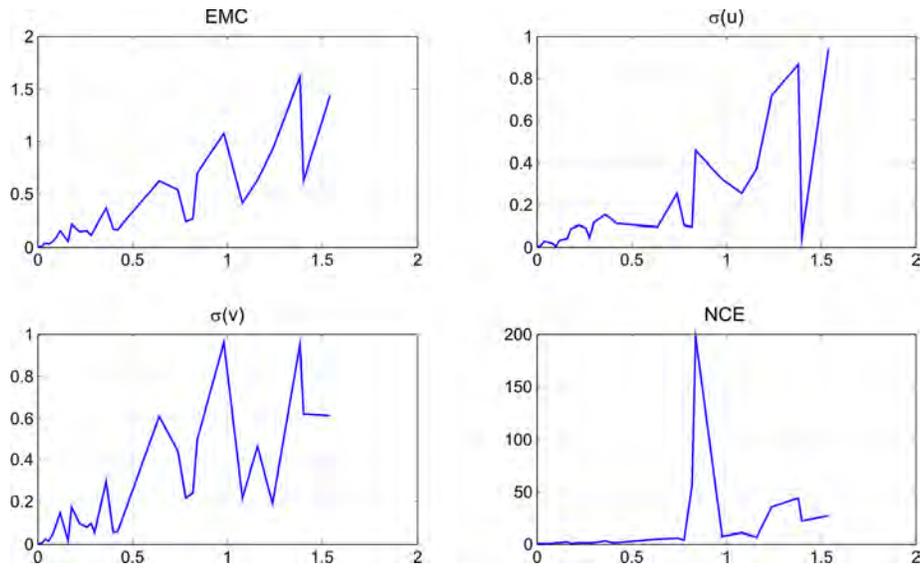


Figura 5-7: Evolución de los errores de la calibración usando el método lineal.

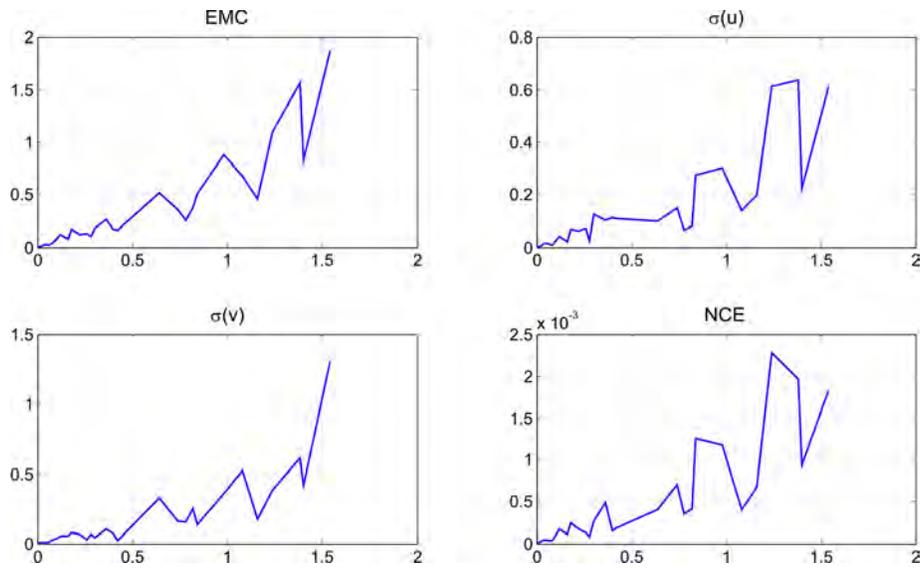


Figura 5-8: Evolución de los errores de la calibración usando el método de Levenberg-Marquardt..

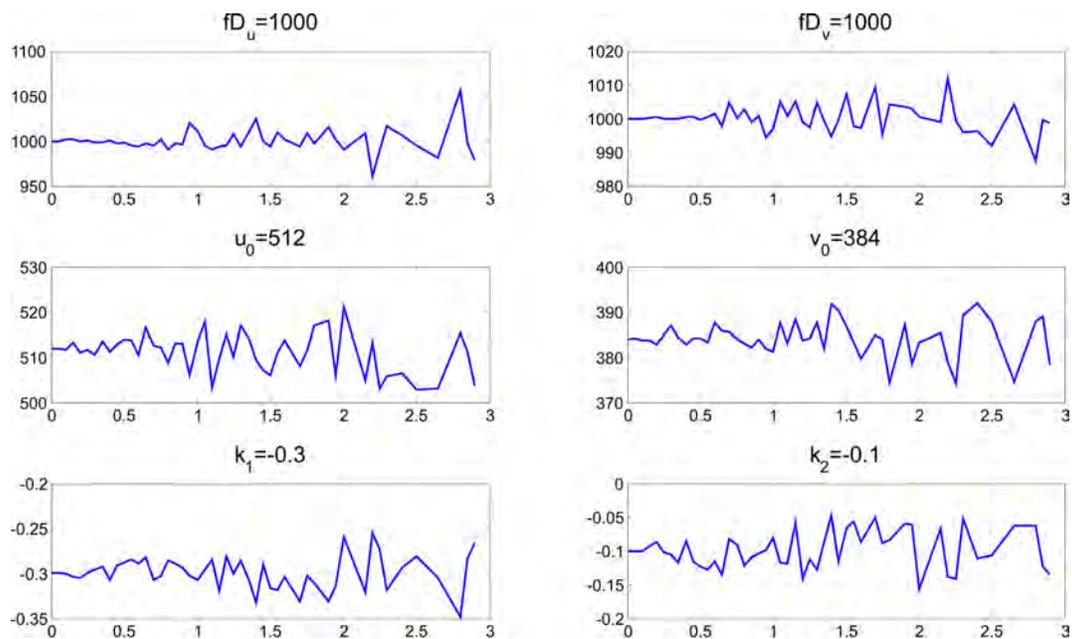


Figura 5-9: Variación de los parámetros intrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 43 puntos de control.

dichas variaciones es que solamente se usaron 7 puntos de control, cuando lo normal es hacer éste tipo de pruebas con un número de puntos de control mayor: Por ejemplo, en [Salvi02] y [Tsai87] se usan 300 puntos de control sintéticos para estimar los parámetros del modelo Pinhole. Debido a esto, se llevo a cabo una segunda prueba de tolerancia a la incertidumbre usando 43 puntos de control, obteniendo las gráficas de evolución de los parámetros presentadas en las figuras 5-9 y 5-10 en las cuales puede observarse que la variación en los estimados disminuye considerablemente y que, incluso, es posible tener una incertidumbre más alta. El hecho que la incertidumbre pueda llegar a tener una desviación estándar de 3 píxeles es un caso poco comun, ya que los puntos de control siempre se escogen de forma que sea posible identificarlos claramente en la imagen garantizando incertidumbres bajas.

Para terminar la comparación con datos sintéticos, se usaron los datos sintéticos usados

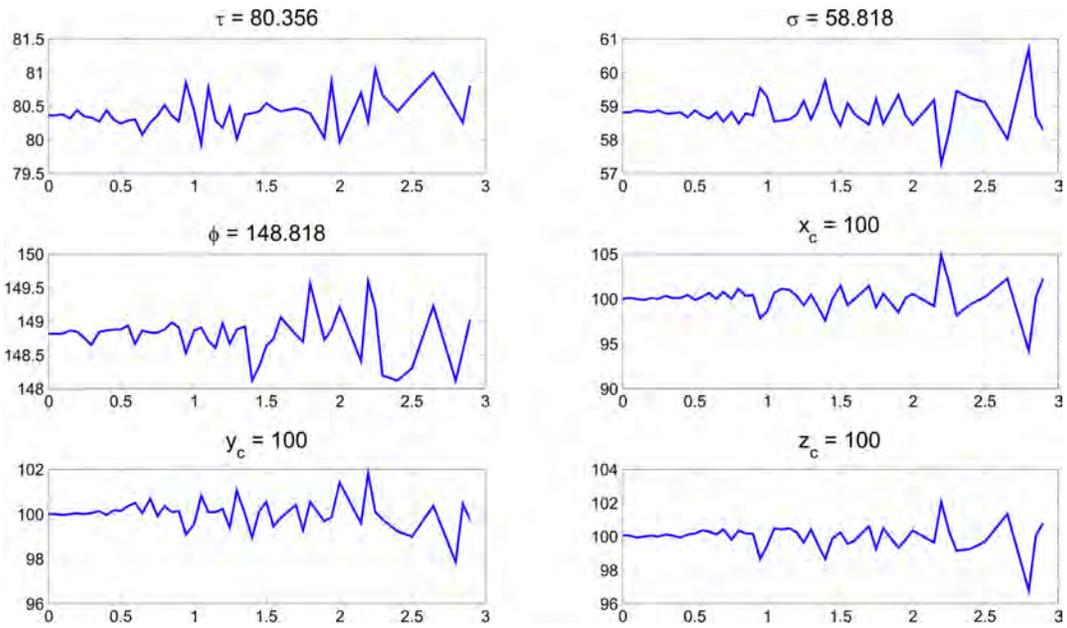


Figura 5-10: Variación de los parámetros extrínsecos al aumentar la incertidumbre en las coordenadas de la imagen usando 43 puntos de control.

en [Salvi02] y [Tsai87] para estimar los parámetros del modelo Pinhole usando los métodos propuestos en este trabajo. En la tabla 5.5.4 se presentan el error medio cuadrático, el máximo error de reproyección cometido, $\max(\epsilon)$ y el error de calibración normalizado para los distintos métodos reportados en [Salvi02] y para los métodos propuestos en éste trabajo, los cuales se identifican con el nombre HORUS. La abreviatura NR indica que el método fue optimizado usando Newton-Raphson, propuesta hecha en [Tsai87]. Puede observarse que los métodos que no tienen en cuenta efectos de distorsión, Faugeras, Faugeras NR sin distorsión y HORUS sin distorsión, presentan mayores errores debido a que los datos sintéticos fueron generados considerando distorsión radial, mientras que los métodos que tienen en cuenta los efectos de la distorsión radial presentan errores bastante menores. También puede observarse que los resultados del método HORUS son mejores en términos de máxima desviación, NCE, y error medio cuadrático, lo que permite concluir que usar el algoritmo de Levenberg-Marquardt permitió obtener resultados de mayor calidad que los reportados con los otros métodos.

Tabla 5.5.4. Errores cometidos al estimar los parámetros del modelo Pinhole usando los datos de [Tsai87]			
Método	<i>EMC</i>	$\max(\epsilon)$	<i>NCE</i>
Faugeras	0.2689	1.2701	0.6555
Faugeras NR sin distorsión	0.2770	1.3692	0.6784
Faugeras NR con distorsión	0.0840	0.2603	0.2042
Tsai optimizado con punto principal optimizado	0.0836	0.2500	0.2033
Weng	0.0845	0.2680	0.2064
HORUS sin distorsión	0.2109	0.7761	0.6575
HORUS con distorsión	0.0632	0.2424	0.2053

5.5 Estimación del modelo Pinhole con distorsión radial

En esta sección se usarán nuevamente los resultados presentados en la sección 4.1.4 para obtener estimados iniciales de los parámetros intrínsecos y extrínsecos de la cámara y para los estimados iniciales de los parámetros de distorsión se usará el valor 0. Dichos estimados de la distorsión se usarán teniendo presente las siguientes consideraciones:

1. Los sistemas cámara-lente de buena calidad presentan distorsiones nulas o despreciables.
2. Para poder estimar los parámetros de distorsión radial con métodos lineales, es necesario contar con más de una imagen de la misma escena desde distintas orientaciones y con el sistema cámara-lente en la misma configuración o al menos contar con imágenes de varias escenas con la misma configuración de la cámara.
3. El sistema HORUS usa cámaras fijas, por lo que sólo es posible obtener una imagen para realizar la calibración luego de que la cámara ha sido instalada.
4. Debido a que las cámaras HORUS están configuradas con distancias focales muy grandes, la distorsión radial disminuye considerablemente.
5. Es preferible para un usuario con poco tiempo y/o conocimiento, marcar tantos puntos de control como sea posible en la misma imagen, que tener que preocuparse por conseguir distintas imágenes para llevar a cabo la calibración.

Nuevamente, para poder estudiar la precisión del método de Levenberg-Marquardt se usaron datos sintéticos. En primera instancia se usaron los datos de la sección 5.4, para verificar que el algoritmo logra converger a un estimado donde los parámetros de distorsión sean 0. En la tabla **5.6.1** se presentan los estimados iniciales, los estimados finales con el método de Levenberg-Marquardt luego de 60 iteraciones y se comparan estos con los parámetros reales. Es claro que en el caso de contar con datos sin incertidumbre y con distorsión nula, el método es capaz de obtener estimados con errores despreciables.

Tabla 5.6.1. Estimados de los parámetros del Modelo Pinhole obtenidos con el método de Levenberg-Marquardt

Parámetro	Estimado inicial	Estimado final	Valor real
fDu	900	1000	1000
fDv	900	1000	1000
u_0	500	512	512
v_0	300	384	384
k_1	0.1	-1.49×10^{-13}	0
k_2	0.2	5.291×10^{-13}	0
a	1.1	1	1
b	0.8	1	1
c	0.35	1	1
t_x	-120	-150.606	-150.606
t_y	-80	-69.3733	-69.3733
t_z	-100	-50.0517	-50.0517

Ahora, usando los datos de la sección 5.4, se generan nuevos valores de las coordenadas u y v , que corresponden a un modelo con distorsión radial dada por los parámetros $k_1 = -0.3$ y $k_2 = -0.1$. En la tabla **5.6.2** se presentan los datos obtenidos.

<i>GCP</i>	x_i	y_i	z_i	u_i	v_i
1	-280	670	1	222.3726	260.1996
2	-280	480	-1	87.2381	408.3606
3	-47	450	3	231.8722	44.6501
4	-134	660	4	311.7882	88.5629
5	-159	560	4	230.6273	176.8484
6	-188	570	3	216.2094	211.4663
7	-216	640	2	242.735	203.6713

Estos datos fueron usados para estimar nuevamente los parámetros del modelo Pinhole, incluyendo los parámetros de distorsión, obteniendo los resultados presentados en la tabla **5.6.3**. El método de Levenberg-Marquardt tardó 62 iteraciones en encontrar el estimado final, usando $\lambda^{(0)} = 10$ y $v = 5$. Nuevamente se obtuvo un estimado con error 0, debido a que los datos usados no tienen incertidumbre alguna.

Tabla 5.6.3. Estimados de los parámetros del Modelo Pinhole obtenidos con el método de Levenberg-Marquardt

Parámetro	Estimado inicial	Estimado final	Valor real
fDu	900	1000	1000
fDv	900	1000	1000
u_0	500	512	512
v_0	300	384	384
k_1	0	-0.3	-0.3
k_2	0	-0.1	-0.1
a	1.1	1	1
b	0.8	1	1
c	0.35	1	1
t_x	-120	-150.606	-150.606
t_y	-80	-69.3733	-69.3733
t_z	-100	-50.0517	-50.0517

Para estudiar la robustez de este método, nuevamente se agregó incertidumbre a los datos u y v , la cual consiste en un ruido gaussiano con media 0 y desviación estándar con valores entre 0 y 3. En la figura 5-11 se presenta la evolución de los parámetros intrínsecos y de distorsión del modelo Pinhole a medida que aumenta la incertidumbre, mientras en la figura 5-12 se presenta la evolución de los parámetros extrínsecos. Puede observarse como los parámetros del modelo Pinhole varían fuertemente, esto puede explicarse debido a que se están usando sólo 7 puntos de control el cual es un número de puntos de control muy pequeño que no permite encontrar un valor adecuado para los coeficientes de distorsión. Una estrategia para evitar dichas variaciones puede consistir en el mayor número de puntos de control que sea posible. En este trabajo solo se están usando 7 debido a que algunas aplicaciones en grandes espacios abiertos con pocos elementos claramente identificables, como las playas estudiadas con el sistema HORUS, no permiten el uso de un gran número

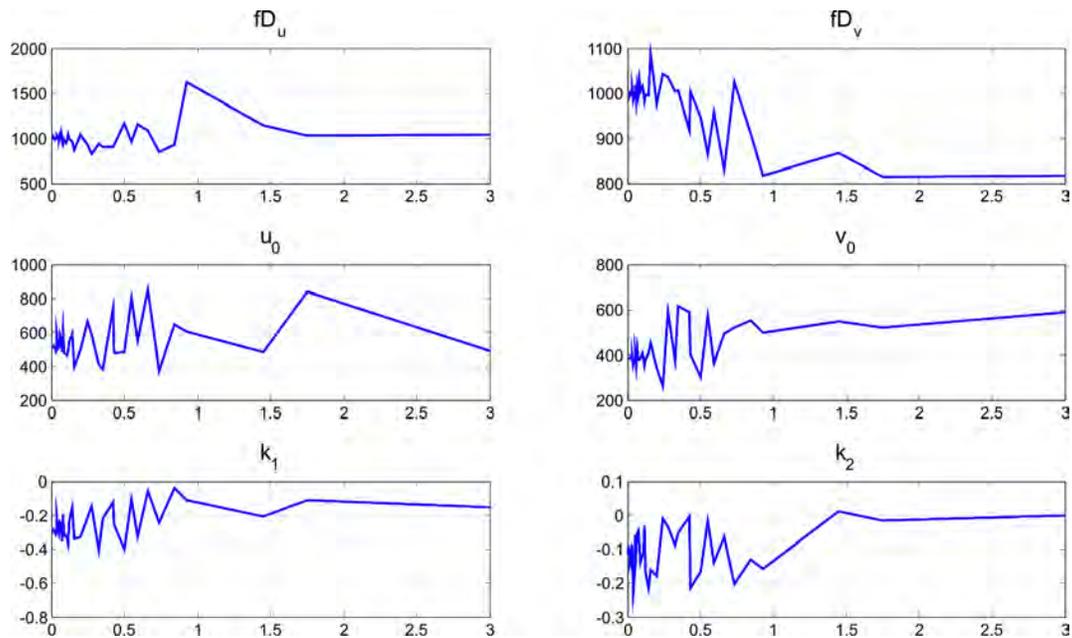


Figura 5-11: Evolución de los parámetros intrínsecos y de distorsión a medida que aumenta la incertidumbre en las mediciones en la imagen.

de puntos de control.

Para validar los resultados obtenidos, en la figura 5-13 se presenta la evolución de los errores MCE , $\sigma(u)$, $\sigma(v)$ y NCE a medida que aumenta la incertidumbre de las mediciones realizadas en la imagen. En este caso puede comprobarse que, a pesar de obtener errores crecientes a medida que aumenta la incertidumbre, las calibraciones obtenidas son de buena calidad según el criterio del NCE .

5.6 Aplicación de la calibración del modelo Pinhole en la estación HORUS Cartagena

En esta sección se presentarán los resultados obtenidos al calibrar una de las cámaras instaladas en la estación de captura HORUS Cartagena usando la Transformada lineal

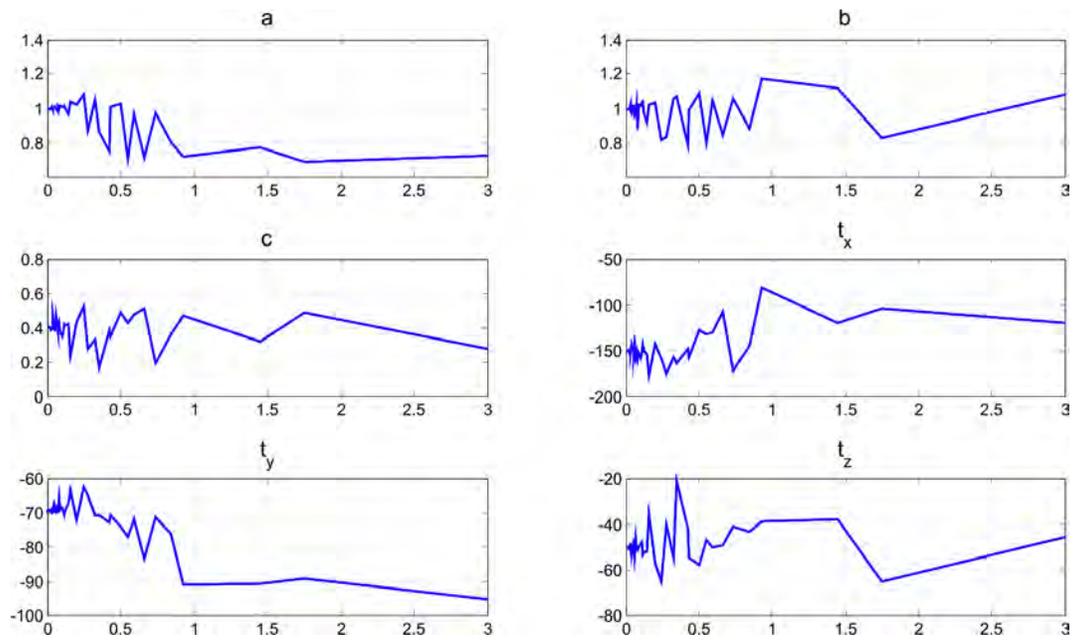


Figura 5-12: Evolución de los parámetros extrínsecos a medida que aumenta la incertidumbre en las mediciones en la imagen.

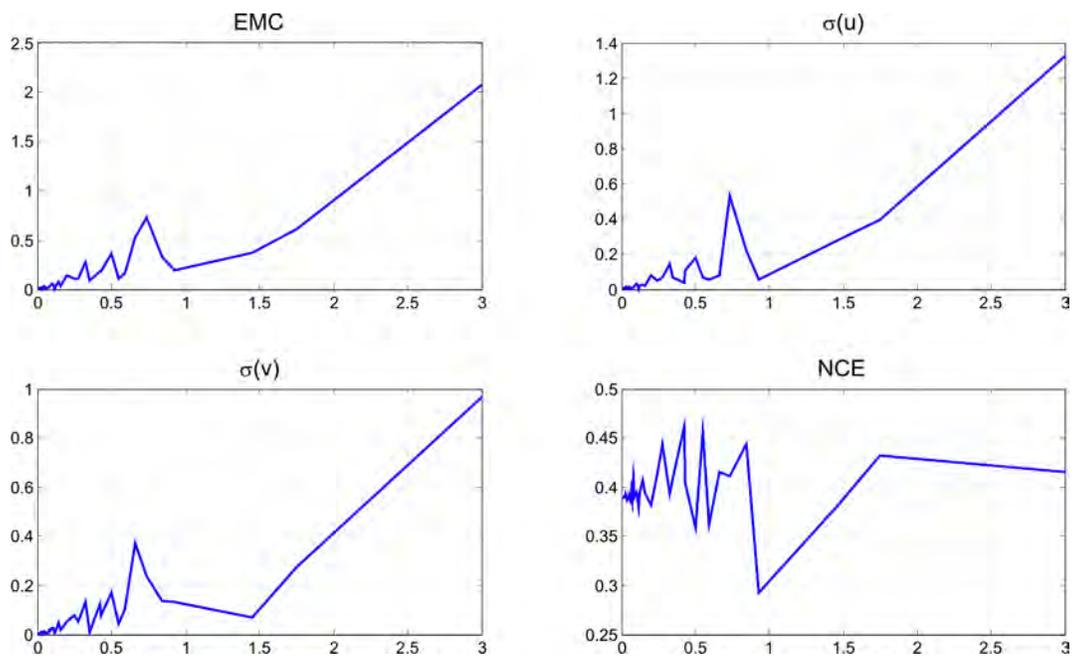


Figura 5-13: Evolución de los errores de la calibración del modelo Pinhole a medida que la incertidumbre aumenta.

Directa y el modelo Pinhole completo. En este caso, se usaron 7 puntos de control, los cuales se muestran en la figura 5-14, y cuyas coordenadas están dadas en la tabla 5.7.1. Las coordenadas XYZ de los puntos de control están tomadas en un sistema de referencia local.



Figura 5-14: Foto de 1024 x 768 píxeles, correspondiente a la cámara 1 de la estación HORUS Cartagena y los puntos de control usados para calibrarla.

Tabla 5.7.1 Puntos de control usados para calibrar la cámara 1 de la estación HORUS Cartagena					
gcp	x	y	z	u	v
1	1893.6827	2105.7929	0.819	37.6918	430.6744
2	1901.7136	2105.2867	0.926	98.3594	452.0852
3	1937.6282	2114.6094	1.437	496.3835	470.6713
4	1937.7197	2116.3825	1.124	507.7952	459.1173
5	1939.5385	2116.1316	1.434	525.1323	464.069
6	1939.2525	2114.4301	1.136	511.7029	477.2375
7	1927.2492	2092.5744	-0.995	250.0514	611.0356

En el caso de la calibración con la Transformada lineal Directa, se usó el método no homogéneo y el método homogéneo presentados en la sección 4.1 obteniendo los resultados presentados en la tabla **5.7.2**, es importante aclarar que el NCE no puede calcularse directamente de la DLT puesto que el NCE depende explícitamente de los parámetros del modelo Pinhole. Puede observarse que los errores asociados a la DLT obtenida con el método no homogéneo son mayores que los presentados por el método homogéneo, esto se debe a que se supone que en la matriz H de la DLT la componente $h_{34} = 1$, lo que no siempre es exacto en aplicaciones prácticas. También es importante resaltar que, aunque el error de reproyección medio cuadrático (EMC) cometido al usar el método homogéneo es reducido, el error de calibración normalizado (NCE) es bastante superior a 1, lo que indica que los parámetros obtenidos a partir de la DLT causan que el error de reproyección sea sustancialmente mayor que el error asociado con la digitalización, por lo que se hace relevante usar un método de optimización no lineal que permita encontrar mejores estimados de los parámetros del modelo Pinhole.

Tabla 5.7.2 Errores al calibrar la cámara con la DLT				
Método	<i>EMC</i>	$\sigma(u)$	$\sigma(v)$	<i>NCE</i>
No homogéneo	0.5018	0.2671	0.2598	12.2559
Homogéneo	0.3070	0.1386	0.1450	85.2971

En [Holland97] y en [Heikilla97], entre otros trabajos, se propone resolver la DLT con el método no homogéneo aunque en el trabajo de [Hartley03], por ejemplo, se explican las razones por las cuales dicha solución no es recomendable, entre ellas que no se tiene garantía que la componente de H que se supone 1 no tenga un valor real cercano a 0. Tomando como argumento ésta falencia del método no homogéneo es que se recomienda usar la SVD para resolver el sistema homogéneo que se presenta en la sección 4.1.2 y obtener la solución de la DLT de la cámara.

Al utilizar el método de Levenberg-Marquardt para estimar los parámetros del modelo Pinhole, se realizó la estimación de los parámetros con y sin distorsión radial, usando como estimados iniciales en primera instancia los parámetros extraídos de la DLT calculada con el método homogéneo y luego usando estimados iniciales basados en las mediciones de los puntos de control y de la posición de la cámara realizadas en campo con técnicas de topografía. Los errores asociados a cada uno de los casos se presentan en la tabla **5.7.3** donde es posible comparar la calidad de cada una de las calibraciones y observar que en términos del error de calibración normalizado y de los errores de reproyección los mejores resultados son aquellos obtenidos al incluir los parámetros de distorsión radial en el proceso de calibración, pero debe tenerse en cuenta que dichos errores se calculan con los puntos de control usados, por lo que no cuantifican los errores cometidos en las regiones de la imagen donde no existen puntos de control.

Tabla 5.7.3 Errores al calibrar el modelo Pinhole con el método de Levenberg-Marquardt						
Método	Parámetros		Errores			
	k_1	k_2	EMC	$\sigma(u)$	$\sigma(v)$	NCE
Con DLT	-	-	1.3925	1.1756	0.9382	0.0001
Distorsión con DLT	$-1.15E-9$	$-4.51E-17$	0.9045	0.8711	0.4425	0.0001
Con Topografía	-	-	1.2686	1.2179	0.6280	0.0001
Distorsión con topografía	0.252	-0.2202	0.3223	0.3397	0.0761	$2.32E-5$

También es posible usar el modelo Pinhole calibrado para una cámara para obtener imágenes rectificadas de zonas de interés, como la que se presenta en la figura 5-15. Estas imágenes se obtienen suponiendo que todos los puntos de la imagen son proyecciones de puntos sobre el plano $z = z^*$, de modo que es posible ubicar los puntos con coordenadas (x, y, z^*) y aplicar una técnica de interpolación análoga a la que se presentó en la sección 3.2.2 para eliminar gráficamente la distorsión. La figura 5-15, la cual corresponde a la misma imagen usada en el proceso de calibración de la cámara 1 de la estación HORUS en Cartagena, se obtuvo suponiendo que la imagen es del plano $z = 0$ y puede observarse que, a pesar que los puntos de control tienen valores de z distintos a 0, sus coordenadas x y y son aproximadas a las medidas en campo, lo que permite concluir que el modelo obtenido para dicha cámara es útil para realizar mediciones de longitud.

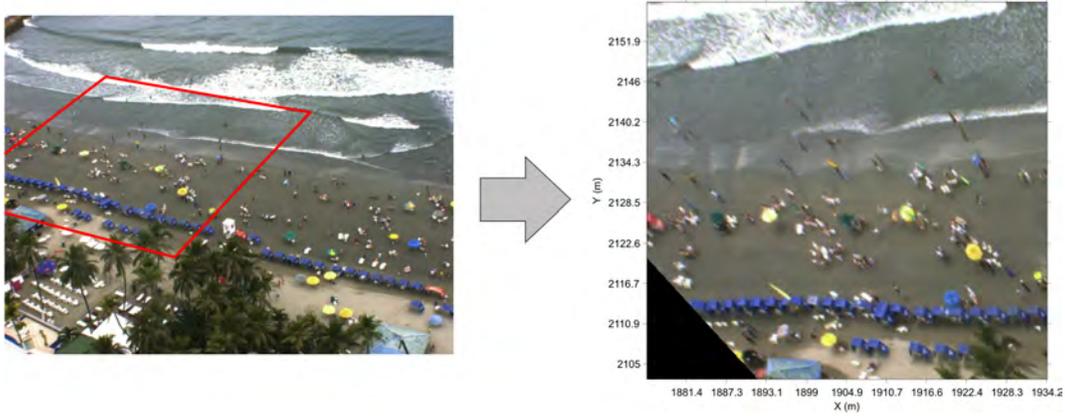


Figura 5-15: Imagen rectificada de la cámara 1 de la estación HORUS en Cartagena usando los parámetros del modelo Pinhole obtenidos a partir de la topografía.

Capítulo 6

Conclusiones

Antes de hablar de las conclusiones de éste trabajo, es necesario recordar que el interés principal era estudiar métodos que permitan estimar los parámetros del modelo Pinhole, obteniendo resultados coherentes y competitivos respecto a los resultados presentados en la literatura. Dicho hecho obligo a estudiar, en primera instancia, métodos para solucionar problemas de mínimos cuadrados con el animo de entender las propiedades más importantes de estos y la formulación de los mismos. Entre los métodos encontrados en la literatura, se eligieron los métodos de Gauss-Newton, el Gradiente y Levenberg-Marquardt, el cual goza de gran aceptación entre la comunidad científica. En el capítulo 1 se resaltaron las ideas presentadas por Marquardt en [Marquardt63] para justificar su método y se señalaron las propiedades que hacen a este método más ventajoso que los métodos de Newton y el Gradiente en la mayoría de los casos de aplicación. En este punto cabe resaltar el esfuerzo realizado en este trabajo para unificar la notación en la presentación de los tres métodos y de sus aplicaciones, lo que permitió estudiar cada uno teniendo presente las propiedades de los otros y adicionalmente permitió implementarlos en Matlab en forma de algoritmos genéricos que pueden usarse directamente para solucionar problemas prácticos o como ayudas didácticas para clases avanzadas de Métodos Numéricos. Por medio de algunos ejemplos, se mostraron las fortalezas del método de Levenberg-Marquardt entre las que se resalta que permite obtener convergencia en casos donde el método de Gauss-Newton falla debido a que el vector de actualización de Newton no está en una dirección de decrecimiento y que presenta velocidad de convergencia más alta que la del método del Gradiente. Las ventajas expuestas conducen a tomar la decisión de usar el método de

Levenberg-Marquardt para resolver todos los problemas asociados a la estimación de los parámetros del modelo Pinhole.

En el segundo capítulo se presentó el modelo Pinhole, usado en la calibración de cámaras fotográficas, se realizó una completa descripción del modelo incluyendo el caso más general que tiene en cuenta los efectos de la distorsión causada por los lentes. Se resalta en ésta parte del trabajo el uso de notación matricial para escribir el modelo Pinhole, lo que permite mostrar en el capítulo 4 el origen de la Transformada Lineal Directa y permite explicar el proceso por el cual se extraen los parámetros del modelo una vez ha sido estimada la DLT. Respecto al modelo Pinhole con distorsión, se presentan las tres fuentes de distorsión más comunes y se construye el modelo que tiene en cuenta los efectos de las tres fuentes, aunque más adelante se explican las razones por las cuales en éste trabajo sólo se estiman los parámetros de distorsión radial y se desprecian los efectos de la distorsión por descentramiento y de prisma delgado.

En el capítulo 3 se presentaron métodos lineales basados en Álgebra lineal para encontrar los parámetros intrínsecos del modelo Pinhole, usando una o varias fotografías, las cuales pueden ser escenas en el espacio (Locaciones interiores o exteriores) o imágenes de patrones de calibración. Entre los métodos presentados sobresale la inclusión de uno basado en geometría proyectiva: El uso de puntos de fuga para la estimación de la matriz de calibración interna. También es de gran importancia el uso de la descomposición en valores singulares para la solución de los sistemas homogéneos asociados a la estimación lineal de los parámetros intrínsecos y de distorsión del modelo Pinhole y para la obtención de matrices de rotación a partir de aproximaciones de las mismas. En este capítulo se explican las razones propuestas por varios autores para despreciar algunos de los parámetros de distorsión y se presentan estrategias para eliminar la distorsión de las imágenes, tanto de manera numérica como de manera gráfica. Se resalta la presentación y uso de la fórmula de Rodrigues, la cual permite parametrizar las matrices de rotación por medio de un vector de tres componentes y a su vez le agrega estabilidad numérica al proceso de estimación

de los parámetros usando mínimos cuadrados. Los métodos presentados en éste capítulo permiten plantear el algoritmo de calibración del modelo Pinhole en dos etapas que consiste en estimar los parámetros intrínsecos y de distorsión en laboratorio y luego estimar los parámetros extrínsecos en campo una vez las cámaras han sido orientadas. Entre los métodos estudiados para la estimación de los parámetros extrínsecos se presenta el método geométrico, denominado problema PnP, y el método basado en la minimización del error de reproyección usando mínimos cuadrados.

El cuarto capítulo está dedicado a los métodos que permiten estimar todos los parámetros del modelo Pinhole, entre los que es relevante resaltar el estudio de la Transformada lineal Directa, de los métodos usados para solucionarla y de los métodos que permiten extraer de ésta los parámetros del modelo Pinhole. Al estudiar la estructura de la LDT y las estrategias para solucionarla, apareció otra herramienta bastante útil para extraer información de una escena a partir de un conjunto de imágenes: La Transformación Afín, la cual permite construir imágenes fusionadas con gran detalle de toda una escena a partir de imágenes más pequeñas las cuales tienen alto detalle y áreas de solape entre una y otra. También se debe destacar la inclusión de las explicaciones basadas en geometría proyectiva para garantizar la solución única de la DLT y para evitar hacer uso de métodos no homogéneos durante la solución de la misma. La DLT se convierte en la herramienta esencial en el proceso de estimación de los parámetros del modelo Pinhole, porque es a partir de ella que se obtienen los estimados iniciales para resolver el modelo usando métodos de mínimos cuadrados para minimizar el error de reproyección. En esta sección se explica como se construye dicho error y como se construye el vector de residuales correspondiente a los métodos de minimización del capítulo 1, lo que permite usar el método de Levenberg-Marquardt para obtener los estimados de los parámetros del modelo Pinhole.

En el capítulo 5 se presenta el uso del método de Levenberg-Marquardt en los problemas de mínimos cuadrados asociados a la estimación de los parámetros del modelo Pinhole, donde cabe resaltar la calibración de los parámetros externos (Cuando se cuenta con los

parámetros internos y de distorsión calculados anteriormente), la estimación de puntos de fuga, y la calibración del modelo Pinhole completo con o sin distorsión radial usando una sola imagen. Los resultados de la calibración del modelo Pinhole mostraron que en el caso de no distorsión, la DLT entrega resultados competitivos con el método lineal, pero tiene la desventaja que sus parámetros no son interpretables y que en la mayoría de los casos, los estimados de los parámetros del modelo Pinhole a partir estos pueden generar calibraciones de poca precisión en términos del error de calibración normalizado, ENC, razón por la cual se hace más importante usar el método de Levenberg-Marquardt cuando es necesario conocer explícitamente dichos parámetros. Para llevar a cabo el proceso de calibración en 2 etapas se desarrolló un toolbox en Matlab que calcula estimados de los parámetros intrínsecos y de distorsión usando los métodos del Gradiente y de Levenberg-Marquardt, el cual fue incluido como una herramienta adicional del software que constituye el sistema HORUS. Para estudiar el funcionamiento del toolbox HORUS se usaron datos reales y se compararon los resultados obtenidos con otro toolbox, disponible en [Bouguet], el cual entrega mejores resultados en términos de los errores de reproyección. En los estimados obtenidos se observa una diferencia marcada en uno de los parámetros de distorsión radial, pero dicho parámetro es cercano a 0, por lo que no tiene un efecto considerable sobre las reproyecciones y, por lo tanto, las diferencias obtenidas al eliminar la distorsión usando los dos conjuntos de parámetros obtenidos no son identificables visualmente, por ser diferencias del orden de 0.5 píxeles. Cabe resaltar que al usar datos reales, se desconoce la solución exacta de los parámetros, si existe, por lo que los resultados obtenidos con ambos toolbox sólo sirven para concluir que la solución obtenida es aceptable. En éste capítulo se usaron distintas medidas del error de la calibración para comparar los métodos lineales y no lineales y hacer pruebas de la tolerancia a la incertidumbre de los mismos, encontrando que los métodos no lineales son en general más robustos y permiten encontrar calibraciones de muy buena calidad, razón por la cual estos son los métodos finalmente implementados en el sistema HORUS para la estimar los parámetros del modelo Pinhole. También pudo comprobarse

que en la presencia de incertidumbre los estimados obtenidos son muy sensibles a la incertidumbre cuando el número de puntos de control es bajo, mientras que dicha sensibilidad disminuye a medida que se aumenta el número de puntos de control usados. También se comparó los resultados al usar el método de Levenberg-Marquardt con los métodos presentados en la literatura, haciendo uso de datos sintéticos usados por varios autores y se concluyó que el método de Levenberg-Marquardt permite obtener estimados mejores en términos de los errores de reproyección. También se comprobó la utilidad práctica de la estimación de los parámetros del modelo Pinhole usando Levenberg-Marquardt para calibrar las cámaras instaladas en la estación HORUS Cartagena usando datos medidos en campo con una campaña topográfica. De los resultados prácticos se resalta que el método no homogéneo para la solución de la DLT no presenta errores de reproyección mayores a los que se obtiene cuando se usa el método homogéneo.

Como conclusión principal de éste trabajo se resalta que el método de Levenberg-Marquardt permite obtener resultados de buena calidad en la calibración de cámaras fotográficas y dichos resultados son competitivos respecto a los métodos reportados en la literatura, por lo que se recomienda el uso de dicho método en la solución de estos problemas. Aunque es importante hacer énfasis en la recomendación de usar el mayor número de puntos de control posible, para minimizar la sensibilidad de los estimados a la incertidumbre presente en las medidas en las imágenes y en el espacio.

Bibliografía

- [Bouguet] Bouguet, Jean-Yves (2008). "Camera Calibration Toolbox for Matlab". Computer Vision Research Group, Department of Electrical Engineering, California Institute of Technology, USA. Disponible en línea en http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.
- [Chatterjee93] Chatterjee, C., y Roychowdhury, V. (2000). "Algorithms for coplanar camera calibration". En *Machine Vision and Applications*, No 12, pp 84-97.
- [Faugeras93] Faugeras, O. (1993). *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, Cambridge, Massachusetts, USA, 663 p.
- [Fischler81] Fischler, M. y Bolles, R. (1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". En *Communications of the ACM*, Volume 24 , Issue 6, pp 381-395
- [Griffiths47] Griffiths, L. (1947). *Introduction to the theory of equations*. 2nd Edition, Chapman & Hall, USA, 278 p.
- [Hartley03] Hartley, R. y Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*, 2d Edition. Cambridge University Press, New York, NY, USA., 700 p.

- [Haralick94] Haralick, B., Lee, C., Ottenberg, K. y Nölle, M. (1994). "Review and analysis of solutions of the three point perspective pose estimation problem". En *International Journal of Computer Vision*, Vol 13, No 3, pp. 331-356.
- [Heikilla97] Heikkilä, J. y Silvén, O. (1997). "A Four-step Camera Calibration Procedure with Implicit Image Correction". En *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*. IEEE Computer Society Washington, DC, USA.
- [Holland97] Holland, K.T., Holman, R.A., Lippmann, T.C. (1997). "Practical Use of Video Imagery in Nearshore Oceanographic Field Studies". En *IEEE Journal of Oceanic Engineering*, Vol.22, No 1, p. 81-92.
- [Kelley95] Kelley, C. T. (1995). *Iterative Methods for linear and Nonlinear Equations*, 1st Edition. SIAM, USA, 166 p.
- [Kelley99] Kelley, C. T. (1999). *Iterative Methods for Optimization*, 1st Edition. SIAM, USA, 180 p.
- [Keys81] Keys, R. (1981) Cubic convolution interpolation for digital image processing. En *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 29, IssueNo 6, p. 1153 - 1160.
- [Laub(05)] Laub, A. J. (2005). *Matrix analysis for scientists and engineers*. SIAM, USA, 157 p.
- [Madsen04] Madsen, K., Nielsen, H.B. y Tingleff, O. (2004). *Methods for Non-linear Least Squares Problems*, 2nd Edition. Informatics and Mathematical Modelling, Technical University of Denmark, 58p.
- [Marquardt63] Marquardt, W. (1963). "An Algorithm for Least-Squares Estimation of

Nonlinear Parameters". En *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, No. 2, pp. 431-441

- [More78] Moré, J. (1978). "The Levenberg-Marquardt algorithm: Implementation and theory ". En *Numerical Analysis, G. A. Watson, ed., Lecture Notes in Mathematics*, Vol. 630, Springer-Verlag, Berlin, 1977, pp. 105-116.
- [Osorio07] Osorio, A., Pérez, J., Ortiz, C., y Medina, R. (2007) "Técnicas basadas en imágenes de video para cuantificar variables ambientales en zonas costeras". En *Avances en recursos Hidraulicos*, No 16, pp 51-64.
- [Ortiz08] Ortiz, C.A (2008). M.I Tesis, Metodología para la construcción de indicadores morfodinámicos a través del uso de cámaras de vídeo. Caso de aplicación: playa de la Magdalena (Cantabria, España). Facultad de Minas, Universidad Nacional de Colombia, Sede Medellín.
- [Salvi02] Salvi, J., Armangué, X. y Batlle, J. (2002) "A comparative review of camera calibrating methods with accuracy evaluation". En *The Journal of Pattern Recognition*, No 35, pp 1617-1635.
- [Tsai87] Tsai, R. (1987) "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses". En *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, p. 323-344.
- [Wang94] Wang, T., Wang, Y. y Yao, C. (2006). "Some Discussion on the Conditions of the Unique Solution of P3P Problem" En *Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, Luoyang, China.
- [Wei94] G. Wei y S., Ma. (1994). "Implicit and explicit camera calibration: The-

ory and experiments". En *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 16, No 5, pp 469-480.

[Weng92] Weng, J., Cohen, P. y Herniou, M. (1992). "Camera calibration with distortion models and accuracy evaluation". En *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 14, No 10, pp 965-980,

[Wolf00] Wolf, P. y Dewitt, B. (2000). *Elements of Photogrammetry with Applications in GIS*, 3d Edition. McGraw-Hill, Boston, Massachusetts, 608 p.

[Xiao03] Xiao-Shan, G., Xiao-Rong, H., Jianliang T., y Hang-Fei C. (2003). "Complete Solution Classification for the Perspective-Three-Point Problem" En *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, pp. 930-943.

[Zhang98] Zhang, Z. (1998). A Flexible New Technique for Camera Calibration, Technical Report MSR-TR-98-71, Microsoft Research, Dec. 1998. Se puede conseguir en línea en: <http://research.microsoft.com/~zhang/calib/>

[Zhang00] Zhang, Z. (2000). "A Flexible New Technique for Camera Calibration". En *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp 1330-1334.