



UNIVERSIDAD DE LAS PALMAS  
DE GRAN CANARIA

Instituto Universitario de Sistemas Inteligentes y Aplicaciones  
Numéricas en Ingeniería

## **Tesis Doctoral**

# **Análisis, Estimación y Corrección de la Distorsión de Lentes en Imágenes y Vídeos. Aplicaciones en Visión por Ordenador**

Daniel Elías Santana Cedrés

Noviembre 2015  
Las Palmas de Gran Canaria





UNIVERSIDAD DE LAS PALMAS  
DE GRAN CANARIA

Programa Oficial de Doctorado en Sistemas Inteligentes y  
Aplicaciones Numéricas en Ingenierías

Instituto Universitario de Sistemas Inteligentes y Aplicaciones  
Numéricas en Ingeniería

## **Tesis Doctoral**

# **Análisis, Estimación y Corrección de la Distorsión de Lentes en Imágenes y Vídeos. Aplicaciones en Visión por Ordenador**

Autor de la Tesis

Director de la Tesis

Daniel Elías Santana Cedrés

Luis Álvarez León

Las Palmas de Gran Canaria, Noviembre de 2015



*A mi padre, in memórium*

*A mi familia*

*A mis amigos, mi otra familia*



*”Las ciencias no tratan de explicar, incluso apenas tratan de interpretar, construyen modelos principalmente. Por modelo, se entiende una construcción matemática que, con la adición de ciertas interpretaciones verbales, describe los fenómenos observados. La justificación de tal construcción matemática es sólo y precisamente que se espera que funcione.”*

**John von Neumann**



## Agradecimientos

Después de varios años de trabajo, llega el momento de poner de relevancia el agradecimiento por la labor con la que, de un modo u otro, han contribuido diferentes personas a lo largo de la realización de la presente tesis doctoral.

Me gustaría empezar por mi director de tesis, Luis Álvarez León. Sin su apoyo y ayuda no hubiera sido posible la realización de este trabajo. Gracias por haberme dado la oportunidad de hacer esta tesis y por las constantes y pacientes explicaciones. Tampoco me gustaría olvidarme de otros miembros del Centro de I+D de Tecnologías de la Imagen, como Luis Gómez y Miguel Alemán. A ambos tengo que darles las gracias por su ayuda durante la realización de los diferentes artículos científicos. Tanto en la redacción como en los experimentos, su punto de vista y experiencia siempre han sido y serán de utilidad. A Agustín Salgado tengo que agradecerle su paciencia y ayuda en la realización de las demos de IPOL. Tampoco quiero olvidarme de Agustín Trujillo y Karl Krissian, con quienes hice mi Proyecto Fin de Carrera y me introdujeron en la investigación, permitiéndome colaborar en mis primeros artículos. Por otro lado también me gustaría hacer mención a Luis Mazorra, Julio Esclarín y a Javier Sánchez, Director del Centro de Tecnologías de la Imagen. Todos ellos forman un equipo que hace fácil lo difícil, y que transmiten pasión y esfuerzo por lo que hacen. Gracias a todos y cada uno de ustedes.

Las horas de trabajo y laboratorio nunca serían lo mismo sin las personas con las que compartes el día a día. Es por ello que me gustaría dar las gracias a esos compañeros que, por fortuna, también tengo la suerte de poder llamarlos amigos. Gracias a Sara Arencibia, Carlos y Sara Illera. Ellos fueron mis compañeros durante la época en la que estuve en el Grupo de Imagen, Tecnología Médica y Televisión. Tampoco quiero olvidarme de Kilian y Airam, así como de Miguel Ángel y José Manuel, con los cuales también he tenido la suerte de compartir esta etapa. Me gustaría hacer una mención especial a dos personas. La primera de ellas es Pedro, quien fuera el anterior doctorando en el grupo y mi compañero durante la primera etapa de mi doctorado. A él tengo que agradecerle muchas cosas y mentiría si dijera que a día de hoy no sigo usando las notas que él hizo sobre diversas configuraciones en el laboratorio. Gracias por haber estado y seguir ahí, por tu visión, por tu apoyo y consejo. La segunda persona es Nelson, con quien he compartido laboratorio todo este tiempo. Jamás habría pensado que cuando empezamos juntos la carrera, terminaríamos trabajando en el mismo sitio. Gracias por seguir ahí y vete preparando, porque tú eres el siguiente.

Quiero también dejar un espacio para el agradecimiento a mi familia. Coinci-

diendo con el desarrollo del doctorado, hemos pasado una época bastante difícil, hasta haber perdido a alguien muy importante para nosotros. A pesar de ello, las cosas siempre han ido saliendo, y estoy absolutamente convencido de que sin su apoyo y comprensión yo no habría sido capaz de seguir adelante. Sigue siendo complicado, pero junto a ustedes las cosas siempre terminan funcionando. Gracias familia.

Desde lo alto veo que aún me queda un trozo en esta página para recordar al resto de mis amigos: Adaya, Belén, Berto, Elena, Saray, Patricia, Tatiana, Natalia, ... y muchos otros más que siempre han estado ahí. Gracias por todo.

Gracias también a todas aquellas entidades sin cuya financiación no hubiera sido posible realizar este trabajo. En especial quería hacer referencia al Grupo de Investigación de Análisis Matemático de Imágenes.

Tengo por costumbre echar mano del álgebra de conjuntos para dar las gracias al conjunto complementario, evitando finalmente dejarme a nadie en los agradecimientos. No obstante, últimamente me encuentro planteando algunas restricciones adicionales, así que discúlpenme si me equivoco con los límites y condiciones de contorno.

¡Gracias!

# Índice general

<b>Introducción</b>	<b>1</b>
Contenido y organización de la tesis . . . . .	5
Principales aportaciones . . . . .	9
Publicaciones realizadas . . . . .	12
<b>1. Estado del arte sobre modelos de distorsión de lentes</b>	<b>17</b>
<b>2. Extracción de bordes en la imagen</b>	<b>33</b>
2.1. Introducción . . . . .	33
2.2. Método de detección de contornos de Canny . . . . .	34
2.3. Refinamiento de los bordes . . . . .	37
<b>3. Estimación automática de modelos de distorsión de un parámetro polinomiales y de división usando la transformada de Hough extendida</b>	<b>43</b>
3.1. Introducción . . . . .	43
3.2. Incluyendo la distorsión de la lente en el espacio de Hough . . . . .	45

3.3. Optimización del parámetro de distorsión . . . . .	55
3.4. Corrección de la distorsión . . . . .	61
3.5. Conclusiones . . . . .	62
<b>4. Estimación automática de modelos de distorsión de dos parámetros polinomiales y de división basada en la rectificación de líneas en la imagen</b>	<b>65</b>
4.1. Introducción . . . . .	65
4.2. Invertibilidad de los modelos de distorsión . . . . .	67
4.3. Normalización de parámetros . . . . .	73
4.3.1. Modelo polinomial . . . . .	74
4.3.2. Modelo división . . . . .	75
4.4. Estimación de un modelo de distorsión de dos parámetros para una colección de líneas dada . . . . .	76
4.5. Optimización iterativa de modelos de distorsión de dos parámetros y de la colección de líneas . . . . .	82
4.6. Corrección de la distorsión de la imagen . . . . .	83
4.7. Resultados experimentales . . . . .	85
4.8. Análisis de la complejidad computacional . . . . .	105
4.9. Conclusiones . . . . .	109
<b>5. Estimación automática de modelos de distorsión usando patrones de calibración</b>	<b>111</b>
5.1. Introducción . . . . .	111

---

5.2. Técnicas estándar de estimación del modelo de distorsión . . . . .	112
5.3. Método propuesto . . . . .	115
5.4. Resultados Experimentales . . . . .	118
5.5. Conclusiones . . . . .	119
<b>6. Aplicación a la detección de esquinas en el patrón de calibración de Zhang</b>	<b>123</b>
6.1. Introducción . . . . .	123
6.2. Correspondencia automática de las esquinas . . . . .	125
6.3. Resultados . . . . .	125
6.3.1. <i>Toolbox</i> de Matlab para calibración de cámaras . . . . .	126
6.3.2. Resultados usando el método propuesto . . . . .	128
6.4. Conclusión . . . . .	130
<b>7. Aplicación a la corrección de la distorsión de lentes en vídeos</b>	<b>131</b>
7.1. Introducción . . . . .	131
7.2. Contextualización del problema . . . . .	132
7.3. Herramienta desarrollada . . . . .	135
7.4. Resultados . . . . .	137
7.5. Conclusiones . . . . .	141
<b>8. Conclusiones y trabajo futuro</b>	<b>145</b>
8.1. Conclusiones . . . . .	145

8.2. Trabajo futuro . . . . . 149

# Índice de figuras

1.	Distorsión óptica tipo barril y cojín . . . . .	2
1.1.	Modelo proyectivo simple ( <i>pinhole model</i> ). . . . .	18
1.2.	Tipos de distorsión radial . . . . .	19
1.3.	Modelo de distorsión radial . . . . .	20
1.4.	Ejemplos de patrones de calibración . . . . .	20
1.5.	Ejemplo de aplicación del método LSD . . . . .	22
1.6.	Selección manual de primitivas . . . . .	25
2.1.	Diagrama de flujo del método de Canny . . . . .	36
2.2.	Ejemplo de refinado de contornos . . . . .	39
3.1.	Línea recta en el plano y en el espacio de parámetros . . . . .	46
3.2.	Representación de la recta en el espacio de Hough . . . . .	47
3.3.	Representación de la distorsión en función de un porcentaje . . . . .	48
3.4.	Matriz de votación 3D de Hough . . . . .	53
3.5.	Detalle de la concentración de las votaciones . . . . .	54

3.6. Evolución del número de votos con diferentes funciones de fiabilidad	57
3.7. Diagrama de flujo de Hough . . . . .	58
3.8. Comparativa Hough clásico frente a Hough mejorado . . . . .	59
3.9. Diagrama de flujo de corrección de la distorsión . . . . .	64
4.1. Caracterización matemática de los parámetros del modelo polinomial	70
4.2. Caracterización matemática de los parámetros del modelo división .	72
4.3. Diagrama de flujo de la optimización iterativa . . . . .	82
4.4. Imágenes originales utilizadas en los experimentos. . . . .	86
4.5. Colección inicial de líneas detectadas . . . . .	88
4.6. Líneas detectadas en la imagen de un patrón con el modelo polinomial	89
4.7. Líneas detectadas en la imagen de un patrón con el modelo división	90
4.8. Líneas detectadas en la imagen de un edificio con el modelo polinomial	91
4.9. Líneas detectadas en la imagen de un edificio con el modelo división	92
4.10. Líneas detectadas en la imagen de una capilla con el modelo polinomial	93
4.11. Líneas detectadas en la imagen de una capilla con el modelo división	94
4.12. Imágenes corregidas de un patrón con poca distorsión . . . . .	95
4.13. Imágenes corregidas de un patrón con fuerte distorsión . . . . .	96
4.14. Imágenes corregidas de la imagen de un edificio . . . . .	97
4.15. Imágenes corregidas de la imagen de una capilla con fuerte distorsión	98
4.16. Imágenes corregidas de la imagen de una capilla con poca distorsión	99

4.17. Evolución del número de puntos con optimización iterativa . . . . .	103
4.18. Múltiples ejemplos de corrección de la distorsión . . . . .	104
4.19. Imágenes usadas para el análisis de la complejidad computacional .	108
4.20. Gráfica de la complejidad computacional . . . . .	108
5.1. Modelo sintético del patrón de calibración e imagen real . . . . .	112
5.2. Perfil de distorsión de la lente Tokina DX 11-16mm . . . . .	113
5.3. Localización de los cuadrados de colores en la imagen . . . . .	118
5.4. Resultados lente Sigma DX 8-16 mm con cámara Nikon D90 . . . . .	120
5.5. Resultados lente Nikkor 14-24 mm con cámara Nikon D800 . . . . .	121
5.6. Resultados lente Nikkor 17-35mm y lente Nikkor 24-70mm con cámara Nikon D800 . . . . .	122
6.1. Esquinas obtenidas con Matlab, sin y con ayuda del usuario . . . . .	127
6.2. Detección de esquinas con el método propuesto . . . . .	129
7.1. Posicionamiento de las cámaras en el terreno de juego . . . . .	134
7.2. Diagrama de flujo de la herramienta desarrollada . . . . .	136
7.3. Detección de primitivas y corrección del primer <i>frame</i> . . . . .	139
7.4. Corrección de una serie de <i>frames</i> . . . . .	140
7.5. Aplicación de la técnica a cables de referencia horizontales en un fotograma . . . . .	142
7.6. Aplicación de la técnica a cables de referencia verticales en un foto- grama . . . . .	143

7.7. Referencia para el *tracking* de peatones . . . . . 143

# Introducción

La distorsión óptica es un fenómeno presente en las lentes de las cámaras, ya sean de vídeo o fotografía, especialmente en las provistas de lentes tipo gran angular o *zoom*. Este efecto se produce por la propia geometría de la lente que determina la forma en que la luz pasa a través de ella y se proyecta en el sensor del dispositivo. Con frecuencia, es necesario corregir este tipo de distorsión presente en las imágenes resultantes de la captura, bien por motivos puramente estéticos, bien porque en algunas aplicaciones importantes de Visión por Computador, como por ejemplo la calibración de cámaras, la medida de distancias en la escena, etc..., esta distorsión afecta negativamente a los resultados.

En esta tesis doctoral abordamos la problemática de la estimación del modelo de distorsión de la lente y sus aplicaciones. La mayoría de las lentes sufren en mayor o menor medida de distorsión óptica producida por su diseño. Dicha distorsión introduce en la imagen aberraciones ópticas que se manifiestan principalmente en la deformación de las líneas rectas que aparecen en la imagen resultante de la captura como líneas distorsionadas. Esta distorsión se denomina de tipo barril cuando las líneas se curvan hacia dentro y de tipo cojín cuando las líneas se curvan hacia fuera. En la figura 1 se ilustra el efecto de la distorsión óptica de la lente. Como puede apreciarse en dicha figura, visualmente existen diferencias significativas entre las distorsiones de tipo barril y cojín, sin embargo, en el modelado matemático y técnicas de corrección de la distorsión que se presentan en esta tesis no existen diferencias en el tratamiento de ambos tipos de distorsión.

La corrección de la distorsión en un punto cualquiera en la imagen  $(x, y)$  se modeliza habitualmente usando una transformación de tipo radial que se formaliza matemáticamente mediante la siguiente ecuación:

$$\begin{pmatrix} \hat{x} - x_d \\ \hat{y} - y_d \end{pmatrix} = L \left( \sqrt{(x - x_d)^2 + (y - y_d)^2} \right) \begin{pmatrix} x - x_d \\ y - y_d \end{pmatrix} \quad (1)$$

donde  $(\hat{x}, \hat{y})$  representa el resultado de la corrección de la distorsión en el punto

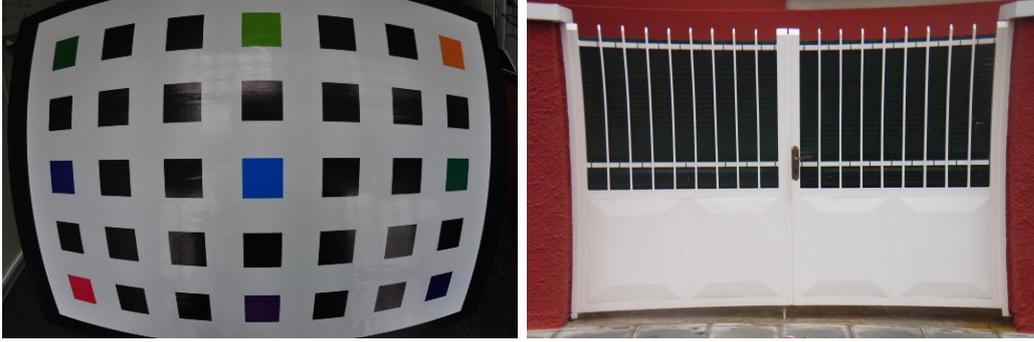


Figura 1: A la izquierda se muestra una imagen de un patrón de calibración plano que presenta una distorsión de tipo barril. A la derecha se muestra una imagen que presenta una distorsión de tipo cojín.

$(x, y)$ ,  $(x_d, y_d)$  representa el centro de la distorsión que habitualmente es un punto cercano al centro de la imagen y finalmente la función radial  $L(r)$  determina la magnitud de la corrección de la distorsión en función de la distancia del punto al centro de distorsión.

La función radial  $L(r)$  se modela habitualmente usando dos tipos de estrategias. La primera da lugar a los denominados modelos polinomiales y consiste en realizar una aproximación polinómica del tipo:

$$L(r) = 1 + k_1 r^2 + k_2 r^4 + \dots + k_n r^{2n} \quad (2)$$

la segunda estrategia, que da lugar a los denominados modelos de división, consiste en realizar una aproximación del tipo:

$$L(r) = \frac{1}{1 + k_1 r^2 + k_2 r^4 + \dots + k_n r^{2n}} \quad (3)$$

a los valores  $k_1, \dots, k_n$  se les denomina parámetros del modelo de distorsión. Estos valores, junto con las coordenadas del centro de distorsión  $(x_d, y_d)$ , determinan completamente el modelo de distorsión. En esta tesis se proponen nuevas técnicas de estimación automática de modelos de distorsión de uno y dos parámetros tanto polinomiales como de división.

En el trabajo desarrollado en esta tesis, la primera hipótesis fundamental que utilizamos, y que determina su ámbito de aplicación, es que en la escena 3D donde se adquiere la imagen hay visibles líneas rectas (al menos una) que se proyectan en la imagen en forma de líneas distorsionadas debido al efecto de la distorsión óptica de la lente. Esta hipótesis se cumple en una gran variedad de escenarios.

---

En particular en las fotografías de zonas urbanas suele haber con frecuencia líneas rectas visibles.

La segunda hipótesis de trabajo que utilizamos es que el modelo de distorsión genera una transformación invertible en el dominio de la imagen. Esto quiere decir que, por ejemplo, el trabajo desarrollado en esta tesis no es de aplicación en las lentes tipo “ojo de pez” que generan una imagen con una zona central circular visible rodeada de una zona negra. Es estos casos, la transformación que da lugar al modelo de distorsión de la lente no es una función invertible en la imagen pues en la zona negra no está correctamente definida. En esta tesis se presta especial atención al estudio de la invertibilidad de los modelos de distorsión y se muestra un resultado matemático que determina la invertibilidad en función de los parámetros del modelo de distorsión para el caso de modelos de uno y dos parámetros tanto polinomiales como de división.

Las diferentes estrategias que se proponen en esta tesis para la estimación automática de los modelos de distorsión se basan, como paso previo, en la extracción de las líneas distorsionadas presentes en la imagen. La técnica habitualmente utilizada para la extracción de líneas rectas en la imagen es la transformada de Hough. Sin embargo, cuando la imagen presenta una distorsión óptica significativa, la transformada de Hough estándar no funciona bien, debido a la distorsión que presentan las líneas. En esta tesis se propone una extensión de la transformada de Hough que tiene en cuenta la distorsión de las líneas.

La corrección de la distorsión de la lente en las imágenes tiene importantes aplicaciones prácticas que van desde la mejora estética de las imágenes hasta la resolución de problemas que se presentan en el ámbito de la Visión por Computador. En particular, en esta tesis se muestran algunas aplicaciones de las técnicas desarrolladas en el contexto de la conocida técnica de calibración de cámaras de Zhang, así como la corrección de la distorsión en vídeos deportivos para el estudio cinemático del movimiento de los jugadores.

El trabajo realizado en esta tesis doctoral se enmarca dentro de las líneas de investigación del grupo de Análisis Matemático de Imágenes (AMI), dirigido por el profesor D. Luis Álvarez León. Dicho grupo pertenece al Centro de I+D de Tecnologías de la Imagen (CTIM) de la Universidad de Las Palmas de Gran Canaria. Las actividades desarrolladas han sido financiadas por las siguientes entidades:

- La Universidad de Las Palmas de Gran Canaria mediante el programa de Personal Técnico de Apoyo a Grupos de Investigación de la ULPGC.

- 
- El Ministerio de Ciencia e Innovación del Gobierno de España a través del Proyecto de Investigación no Orientada **Modelización matemática de los procesos de calibración de cámaras de vídeo**, de Referencia: **MTM2010-17615** (subprograma MTM).
  - El Grupo de Investigación de Análisis Matemático de Imágenes (AMI) de la Universidad de Las Palmas de Gran Canaria a través de diversos contratos de apoyo a la investigación.
  - El Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería, a través de su programa de ayudas propias a la matrícula en estudios de máster y doctorado.
  - El Ministerio de Educación, Cultura y Deporte a través de la Secretaría de Estado de Educación, Formación Profesional y Universidades, mediante su programa de becas para estudios de postgrado.

---

## Contenido y organización de la tesis

Atendiendo a las aportaciones realizadas, así como a las diferentes aplicaciones de las técnicas propuestas, la presente tesis se estructura en un total de ocho capítulos. En el primero de ellos se hace una descripción del estado del arte sobre modelos de distorsión de lentes, mientras que en los siguientes se incluyen las principales aportaciones realizadas. Posteriormente, se explican diferentes aplicaciones de las técnicas propuestas, relacionadas con la calibración de cámaras y la corrección de secuencias de vídeo. Para terminar, se especifican las conclusiones que se obtienen de todo el trabajo realizado, así como las posibles líneas futuras a desarrollar. A continuación se describe en mayor detalle el contenido de cada uno de los capítulos que constituyen esta tesis doctoral:

- **Capítulo 1. Estado del arte sobre modelos de distorsión de lentes:** en este primer capítulo se describe en detalle la formalización de los modelos de distorsión de la lente. Se explica, en términos generales, cuáles son los pasos habituales seguidos para la estimación de la distorsión y su posterior corrección. Asimismo, se realiza un recorrido a través de la literatura disponible, describiendo diferentes enfoques tanto desde el punto de vista del análisis de la escena a través de la detección de primitivas, como respecto a distintos enfoques al abordar el problema de la distorsión de la lente.
- **Capítulo 2. Extracción de bordes en la imagen:** el primer paso, antes de proceder a la detección de primitivas en la imagen, es proporcionar un conjunto de puntos de interés que sirvan como partida para la búsqueda. En este capítulo presentamos el método de detección de bordes usado. Se trata de un método clásico, el método de Canny. Asimismo, también introducimos un paso opcional intermedio entre la extracción de los contornos y la detección de primitivas. En dicho paso se realiza un refinamiento de los contornos, consistente en eliminar *outliers* y preservar aquellos bordes con una orientación parecida dentro de una determinada vecindad.
- **Capítulo 3. Estimación automática de modelos de distorsión de un parámetro polinomiales y de división usando la transformada de Hough:** una vez se ha realizado la detección de bordes y se ha introducido un proceso de refinamiento, en el siguiente capítulo introducimos la estimación automática de modelos de un parámetro de tipo polinomial y división. Para dicho fin se hace uso de la transformada de Hough para detectar las primitivas presentes en la imagen de entrada. Sin embargo, el enfoque clásico proporciona resultados poco satisfactorios, puesto que la distorsión propicia

---

que las rectas sean detectadas como segmentos rectilíneos separados. Proponemos una ampliación del espacio de votación de Hough, de manera que se introduce la distorsión como una nueva dimensión. Eso permite detectar el conjunto de primitivas con el mejor ajuste a las líneas distorsionadas en la imagen de entrada, así como proporcionar un valor inicial del parámetro de distorsión de la lente. Tras ello se introduce un proceso de optimización no lineal, a través del cual se optimiza el modelo de distorsión mediante la minimización de la distancia entre las líneas y sus puntos de borde asociados una vez corregida la distorsión con el modelo.

- **Capítulo 4. Estimación automática de modelos de distorsión de dos parámetros polinomiales y de división basada en la rectificación de líneas en la imagen:** la técnica propuesta de una transformada mejorada de Hough mediante la introducción de la distorsión como una nueva dimensión en espacio de votación, permite detectar líneas distorsionadas fruto de la proyección de líneas rectas 3D reales. Sin embargo, cuando la distorsión es fuerte, la estimación de modelos de un parámetro se hace insuficiente, ya que no son capaces de abordar problemas con tal tipo de distorsión. Por ello, en este capítulo hablamos de la estimación automática de modelos de dos parámetros de tipo polinomial y división basándonos en la rectificación de las líneas detectadas. Para ello, introducimos en el proceso de optimización no lineal un segundo parámetro, así como el centro de distorsión. Esto permite un mejor modelado de la distorsión presente en imágenes afectadas por una fuerte distorsión de la lente. Además, caracterizamos matemáticamente las condiciones bajo las cuales los parámetros de distorsión definen una transformación biyectiva en el dominio de la imagen. Asimismo, introducimos un proceso de optimización iterativa, mediante el cual se mejoran tanto el modelo de distorsión obtenido como el conjunto de primitivas asociadas. De esta manera, la técnica propuesta permite la detección de líneas más largas y, por lo tanto, obtener una información más significativa de cara a la corrección de la distorsión. Todo ello añade robustez y fiabilidad a los modelos, ya que las condiciones de invertibilidad garantizan que es posible computar la transformación inversa y, por ende, calcular la imagen corregida. Para validar la técnica presentada, realizamos experimentos en una amplia variedad de imágenes, tanto de patrones de calibración sintéticos como de imágenes de escenarios reales.
- **Capítulo 5. Estimación automática de modelos de distorsión usando patrones de calibración:** el proceso de calibración de una cámara conlleva la obtención de aquellos parámetros que ponen en correspondencia la escena real que estamos analizando con un modelo de la misma. General-

---

mente esto se realiza sobre patrones de calibración y conlleva el cálculo de una homografía, que es una transformación en el plano que nos proporcionan esa correspondencia. Sin embargo, es importante introducir un buen modelo de la distorsión de la lente en el procedimiento, ya que de lo contrario el resultado obtenido podría no ser satisfactorio. En los capítulos anteriores se han descrito técnicas para obtener el modelo de distorsión sin tener ningún conocimiento de la escena, salvo la hipótesis de que existen líneas rectas visibles en ella pero sin conocer su localización. En este capítulo se hace uso de una escena conocida, la compuesta por un patrón de calibración, para introducir un criterio métrico en el proceso de estimación del modelo de distorsión de la lente. La técnica descrita consiste en la minimización del error de reproyección de los puntos detectados sobre las rectas conocidas del patrón de calibración. Inicialmente, se detecta la posición de tres cuadrados de colores conocidos que se encuentran en la región central del patrón, para posteriormente obtener las ocho rectas que pasan por ellos. La primera aproximación de la homografía se hallará poniendo en correspondencia esas líneas. Basándonos en dicha aproximación, se realiza la proyección de cada recta, estableciendo una correspondencia si el resultado de dicha proyección está lo suficientemente cercana a la recta real. La minimización del error de reproyección, se llevará a cabo haciendo uso del método de Levenberg–Marquard, para lo cual se parametriza la homografía mediante las coordenadas de las esquinas del cuadrado verde central. De esta manera, se tiene un mejor criterio, tanto para la estimación del modelo de distorsión de la lente, como para la obtención de la homografía entre el modelo del patrón de calibración y la fotografía del mismo.

- **Capítulo 6. Aplicación a la detección de esquinas en el patrón de calibración de Zhang:** dentro del proceso de calibración de una cámara, una de las etapas fundamentales es la detección de las esquinas del patrón de calibración que se esté usando para dicho procedimiento. En este capítulo, presentamos la aplicación de las técnicas expuestas para obtener la posición de las esquinas de los cuadrados sobre un patrón de calibración de Zhang. Para ello, inicialmente se obtienen el conjunto de primitivas presentes en la imagen, mediante los métodos anteriormente descritos y se ponen en correspondencia las líneas detectadas con las presentes en el modelo, mediante la técnica descrita en el capítulo anterior. Tras ello, se calcula una primera estimación de la posición de las esquinas de los cuadrados. Dicha estimación viene dada por el cálculo de la intersección de las líneas detectadas. Primeramente, se corrigen las primitivas y se calculan las intersecciones, para en un siguiente paso, obtener la primera aproximación de la posición de las esquinas mediante la transformación inversa del modelo de distorsión de la lente

---

obtenido. Por último, el conjunto final de las esquinas se obtiene poniendo en correspondencia las obtenidas mediante las intersecciones y las que se tienen como resultado del método de detección de esquinas de Harris. La aplicación de la técnica propuesta, proporciona una detección de las esquinas con una alta precisión, así como un número mucho mayor de esquinas detectadas, frente a otras técnicas que tienen problemas para abordar imágenes afectadas por una fuerte distorsión. Esto hace que el resultado obtenido a través de este método, proporcione información mucho más completa para el proceso de calibración de la cámara.

- **Capítulo 7. Aplicación a la corrección de la distorsión de lentes en vídeos:** los deportes de interior suponen un reto a la hora de colocar sistemas de cámaras, ya que se encuentran varios obstáculos que hacen difícil la colocación de las mismas. Por ello, realizar estudios sobre el comportamiento y el movimiento cinemático, conlleva abordar el problema con sistemas multicámara que graban simultáneamente la escena. En este capítulo se habla de la introducción de un sistema con una sola cámara provista de lente gran angular para capturar la secuencia. Describimos en detalle una técnica desarrollada para corregir la distorsión presente en el vídeo y proporcionar un vídeo corregido. Partiendo del primer fotograma de la secuencia, se estima un modelo de distorsión de la lente usando las técnicas descritas en capítulos anteriores. Tras ello, se corrige la secuencia al completo y se genera un vídeo de salida libre de distorsión. Este trabajo se ha hecho en colaboración con el Laboratory of Biomechanics and Motor Control de la Universidad de São Paulo, Brasil (LaBioCoM – USP).
- **Capítulo 8. Conclusiones y trabajo futuro:** en este capítulo presentamos las conclusiones obtenidas de todo el trabajo desarrollado y de las técnicas propuestas. Asimismo, se describen diversas líneas futuras a seguir, así como posibles aplicaciones de los métodos que se han descrito.

---

## Principales aportaciones

A continuación se relacionan brevemente las aportaciones principales de la tesis. Cabe destacar que para cada uno de los temas desarrollados se presentan una gran variedad de experimentos que justifican y validan el trabajo realizado. Por otro lado, a través de 2 publicaciones en la revista IPOL Journal of Image Processing Online, se ha hecho público, tanto el código de los principales algoritmos implementados como algunas demos online que permiten realizar experimentos vía web fácilmente de forma interactiva.

- ***Estimación de modelos de distorsión de 1 parámetro.***

Se propone una versión extendida de la transformada de Hough añadiendo un parámetro de distorsión al espacio de parámetros que permite automáticamente estimar modelos de distorsión de un parámetro tanto polinomiales como de división. Dicha estimación se basa en la extracción en la imagen de puntos de interés para lo cual usamos el detector de bordes de Canny. Proponemos también una nueva técnica de refinamiento de los contornos obtenidos por el método de Canny que se basa en la eliminación de puntos del contorno aislados o con curvatura alta. Este criterio se ajusta a los requerimientos del problema de estimación de modelos de distorsión que se basan en buscar estructuras alineadas. La transformada de Hough extendida suministra un valor inicial del parámetro de distorsión y una colección de rectas con sus correspondientes puntos del contorno asociados. Posteriormente se optimiza dicho parámetro de distorsión minimizando el error de rectificación dado por la distancia de los puntos a su recta asociada una vez corregida la distorsión usando el modelo.

- ***Estimación de modelos de distorsión de 2 parámetros.***

Se estudian en detalle los modelos de distorsión de lentes polinomiales y de división de 2 parámetros y se demuestran matemáticamente las condiciones que deben cumplir dichos parámetros para que sean transformaciones invertibles en la imagen. Se propone un nuevo algoritmo de optimización no-lineal para el cálculo de los parámetros de distorsión basado de nuevo en la rectificación de las líneas distorsionadas visibles en la imagen. Es de destacar aquí que este método no requiere de ningún conocimiento de la escena y que la única hipótesis que se se asume es que existen en la escena algunas líneas rectas visibles que son proyectadas en la imagen.

- ***Estimación de modelos de distorsión usando patrones de calibración.***

---

Se propone una nueva técnica para estimar los modelos de distorsión de 2 parámetros usando patrones de calibración. Se ha diseñado un patrón de calibración con unos cuadrados centrales de colores que permiten fijar una referencia para estimar automáticamente las correspondencias entre las rectas del patrón de calibración real y su imagen. Para explotar convenientemente el conocimiento métrico de las posiciones exactas de todas las rectas del patrón se propone usar como criterio de estimación del modelo de distorsión la minimización del error de reproyección entre los puntos de las rectas de la imagen y las rectas en el patrón de calibración. Dicho criterio da una medida más fiable y robusta sobre la calidad de un modelo de distorsión que las técnicas habituales usadas por las revistas especializadas. Utilizamos esta metodología para comparar los resultados de los modelos de distorsión obtenidos por las técnicas desarrolladas en este trabajo y las técnicas estándar en una variedad de lentes con un nivel de distorsión moderado.

- ***Aplicación a la detección de esquinas en el contexto del método de calibración de Zhang.***

El conocido método de calibración de Zhang utiliza varias vistas de un patrón de calibración plano para calibrar completamente la cámara calculando los parámetros intrínsecos y extrínsecos. El paso inicial del método de Zhang consiste en la detección y puesta en correspondencia de las esquinas de la proyecciones de los cuadrados en la imagen con las esquinas del patrón real. Cuando la cámara presenta una fuerte distorsión de lentes el problema de la detección automática de las esquinas y puesta en correspondencia resulta muy complejo. A partir de las técnicas desarrolladas en esta tesis se propone un nuevo método para localizar y poner en correspondencia las esquinas de los cuadrados del patrón de calibración basándose en la detección de las rectas distorsionadas y su intersección.

- ***Aplicación a la corrección de la distorsión en secuencias de vídeo.***

En colaboración con el Laboratory of Biomechanics and Motor Control de la Universidad de São Paulo, Brasil (LaBioCoM – USP), se ha desarrollado, usando las nuevas técnicas presentadas en esta tesis, una aplicación que permite corregir la distorsión en vídeos deportivos con objeto de poder hacer mediciones reales sobre los desplazamientos de los jugadores en el terreno de juego. De forma automática se detectan líneas en la escena (por ejemplo las líneas que delimitan el campo), y en base a ellas se construye el modelo de distorsión y se corrige el video original creando un nuevo vídeo con la distorsión corregida. Sobre este nuevo vídeo con la distorsión corregida los investigadores del LaBioCoM – USP realizan un estudio cinemático sobre el movimiento de los jugadores. Los resultados obtenidos muestran que con la

---

corrección de la distorsión, la calidad y precisión del estudio cinemático es equivalente al que se obtiene con sistemas de adquisición de vídeos mucho más complejos y caros que requieren un mayor volumen de almacenamiento y de trabajo en el postproceso de las secuencias.

---

## Publicaciones realizadas

En este apartado se describe el conjunto de publicaciones realizadas, así como las principales aportaciones de las mismas:

1. Miguel Alemán-Flores, Luis Alvarez, Luis Gomez y Daniel Santana-Cedrés: **Line detection in images showing significant lens distortion and application to distortion correction.** *Pattern Recognition Letters* 36 (2014), 261–271.

Las líneas rectas son unas de las primitivas básicas usadas por sistemas de percepción para analizar e interpretar una escena. Por tanto, su detección es muy importante desde el punto de vista de la robustez y flexibilidad de los sistemas de Visión por Computador. Sin embargo, si la imagen a procesar presenta una distorsión significativa, los métodos tradicionales de detección de líneas fallan. En este artículo se presenta una extensión de la transformada de Hough, introduciendo un nuevo parámetro que corresponde a la distorsión de la lente, de tal forma que el espacio de búsqueda pasa a ser tridimensional. En este nuevo espacio una línea vendría representada por tres parámetros: orientación, distancia al origen y distorsión. Usando la colección de líneas distorsionadas que se recuperan con esta nueva técnica, somos capaces de estimar la distorsión de la lente, corregirla y crear una nueva imagen sin distorsión.

2. Miguel Alemán-Flores, Luis Alvarez, Luis Gomez y Daniel Santana-Cedrés: **Wide-Angle Lens Distortion Correction Using Division Models.** In: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* 8258 (2013), 415–422. CIARP 2013 18th Iberoamerican Congress on Pattern Recognition.

En este artículo proponemos un nuevo método, basado en modelos de división, para corregir automáticamente la distorsión producida por lentes gran angular a partir de las líneas generadas por la proyección en la imagen de líneas rectas tridimensionales. Para ello han de afrontarse principalmente dos problemas: por un lado, las lentes gran angular producen una fuerte distorsión, lo que hace de la detección de las líneas distorsionadas una tarea particularmente difícil. Por otra parte, los modelos de distorsión de la lente polinomiales de un parámetro no son capaces de manejar distorsiones de este grado. Proponemos una extensión de la transformada de Hough mediante la adición un parámetro de distorsión para detectar las líneas distorsionadas, y el uso del modelo división de distorsión de la lente para manejar distorsiones

---

tipo gran angular. La ventaja de usar este tipo de modelos de distorsión radial se presenta comparando con modelos de distorsión polinomiales.

3. Miguel Alemán-Flores, Luis Alvarez, Luis Gomez y Daniel Santana-Cedrés: **Automatic Lens Distortion Correction Using One-Parameter Division Models.** *IPOJ Journal - Image Processing On Line* 4 (2014), 327–343.

Se presenta un método para corregir automáticamente la distorsión radial provocada por lentes gran angular usando las líneas distorsionadas producto de la proyección de líneas rectas reales en la imagen. La distorsión de la lente se estima mediante un modelo de distorsión de tipo división de un parámetro, lo que permite localizar el problema dentro del esquema de la transformada de Hough mediante la adición de un parámetro de distorsión para extraer mejor las líneas distorsionadas de la imagen de entrada. Este artículo describe un algoritmo que aplica esta técnica, proporcionando todos los detalles del diseño de la transformada de Hough modificada. Se realizan diversos experimentos usando patrones de calibración y también escenas reales que muestran una fuerte distorsión para ilustrar así la capacidad del método propuesto.

4. Miguel Alemán-Flores, Luis Alvarez, Luis Gomez y Daniel Santana-Cedrés: **Automatic Corner Matching in Highly Distorted Images of Zhang's Calibration Pattern.** In: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications* 8827 (2014), 754–761. CIARP 2014 19th Iberoamerican Congress on Pattern Recognition.

El método de Zhang es una técnica ampliamente extendida para la calibración de cámaras a partir de diferentes vistas de un patrón de calibración. Este patrón contiene un conjunto de cuadrados colocados de una cierta manera. Con el objetivo de calibrar la cámara, han de ponerse en correspondencia las esquinas de los cuadrados en la imagen y en el modelo de referencia. Cuando las imágenes muestran una fuerte distorsión de la lente, los métodos tradicionales para calcular la posición de las esquinas fallan, ya que las esquinas están desplazadas de las posiciones esperadas. En este artículo se propone un nuevo método que automáticamente estima la correspondencia de dichas esquinas considerando la distorsión de la lente. El método se basa en un algoritmo automático para la corrección de la distorsión de la lente que permite estimar las líneas distorsionadas que pasan por los bordes de los cuadrados.

5. Luiz H.P. Vieira, Emilio A. Pagnoca, Fabio Milioni, Ricardo A. Barbieri, Rafael P. Menezes, Luis Alvarez, Luis Gomez, Daniel Santana-Cedrés, Paulo

---

R.P. Santiago: **Tracking in futsal using a single wide-angle lens camera: accuracy analysis of the radial distortion correction based on an improved Hough transform algorithm.** *Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* (2015), 1–11.

En este artículo se propone el uso de una única cámara gran angular para el seguimiento de jugadores de fútbol sala y el estudio de propiedades cinemáticas del movimiento de los mismos. Para ello se compara con un sistema que usa tres cámaras provistas de lentes convencionales. Para poder realizar un *tracking* lo más preciso posible, es necesario corregir la distorsión presente en el vídeo provisto por la cámara gran angular. Se propone un sistema que, a partir del conjunto de *frames* del vídeo de entrada, estima un modelo de distorsión de la lente a partir del primer fotograma y posteriormente se realiza la corrección del conjunto completo, generando un vídeo de salida libre de distorsión. Esta nueva secuencia constituye la entrada para el estudio del seguimiento de los jugadores. Los resultados mostrados presentan que el sistema de una sola cámara proporciona resultados semejantes a sistemas más complejos y además presenta otras ventajas, como la reducción del espacio del almacenamiento y del tiempo de postproceso de la secuencia.

6. Daniel Santana-Cedrés, Luis Gomez, Miguel Alemán-Flores, Agustín Salgado, Julio Esclarín, Luis Mazorra, Luis Alvarez: **Invertibility and Estimation of Two-Parameter Polynomial and Division Lens Distortion Models.** *SIAM Journal on Imaging Sciences* 8(3) (2015), 1574–1606.

En este artículo estudiamos los modelos de distorsión de lentes de dos parámetros de tipo polinomial y división. En concreto, analizamos la invertibilidad de dichos modelos, caracterizando matemáticamente las condiciones bajo las cuales los parámetros de distorsión de ambos modelos definen una transformación biyectiva. Esto asegura que la transformación inversa está bien definida y la imagen corregida puede ser calculada de forma adecuada, lo que proporciona robustez a los modelos de distorsión. Se propone un nuevo método automático para corregir la distorsión radial, y se hace un análisis comparativo de este método usando tanto el modelo polinomial como el división. Con el objetivo de obtener una estimación precisa del modelo, proponemos un esquema de optimización que iterativamente mejora los parámetros para alcanzar una correspondencia mejor entre las líneas distorsionadas y los puntos de borde. El método propuesto estima modelos de distorsión radial de dos parámetros mediante la detección de las líneas distorsionadas más largas en la imagen. Esto se realiza a través de la aplicación de una transformada de Hough extendida con un parámetro de distorsión radial. Tras ello, se estima un modelo de dos parámetros usando un esquema iterativo de optimización

---

no lineal. Este esquema pretende minimizar la distancia de los puntos borde a las líneas asociadas mediante el ajuste de los dos parámetros de distorsión, así como de las coordenadas del centro de distorsión.

7. Daniel Santana-Cedrés, Luis Gomez, Miguel Alemán-Flores, Agustín Salgado, Julio Esclarín, Luis Mazorra, Luis Alvarez: **An Iterative Optimization Algorithm for Lens Distortion Correction Using Two-Parameter Models** *IPOL Journal - Image Processing On Line* (2015).

Este artículo se presentó conjuntamente al anterior y en él se describe en detalle la implementación de la técnica propuesta desde distintos niveles de abstracción. Una vez detectados los bordes presentes en la imagen de entrada y de que estos sean refinados, el método en primer lugar detecta el conjunto de líneas candidatas que mejor se ajusta a las líneas distorsionadas presentes en la imagen, además de obtener un valor inicial para el modelo de distorsión. Posteriormente, se lleva a cabo un proceso de optimización iterativo. A través de dicho proceso, por un lado se minimiza la distancia de las líneas detectadas a los bordes, mediante el ajuste de un modelo de distorsión de dos parámetros y del centro de distorsión. Por otra parte, también se intenta mejorar el conjunto de primitivas, aplicando tras la minimización una nueva detección de las primitivas proveyendo dicho modelo como entrada. Este proceso se lleva a cabo hasta que no se obtiene mejora en el resultado. Este artículo se encuentra en fase de publicación y ha sido aceptado como prepublicación, siendo accesible a través del enlace <http://www.ipol.im/pub/pre/130/>.

---

# Capítulo 1

## Estado del arte sobre modelos de distorsión de lentes

Las lentes, de la misma forma que otros sistemas ópticos, poseen defectos que pueden causar que la imagen resultante de una captura aparezca borrosa, con cambios en el color, distorsiones geométricas, etc. A dichos defectos en óptica normalmente se les llama aberraciones.

En el modelo proyectivo simple (*pinhole model*), un punto en el espacio 3D  $(x, y, z)$ , la escena, se proyecta en la imagen en un punto 2D  $(x, y)$ , el sensor, en coordenadas proyectivas (fig. 1.1). Las características de una cámara se establecen a través de una serie de parámetros, divididos principalmente en dos tipos: intrínsecos y extrínsecos. Los parámetros intrínsecos vienen determinados por el alto y el ancho de los píxeles de la imagen que se captura (*pixel aspect ratio*), la distancia focal (que puede ser variada mediante el mecanismo de *zoom*) y el punto focal. El foco de la cámara suele estar centrado respecto al negativo o, en el caso de cámaras digitales, respecto al sensor CCD en el que se proyecta la captura. Los parámetros extrínsecos de una cámara vienen caracterizados generalmente por una matriz de rotación y el foco. Estos parámetros determinan la posición y orientación de la cámara en el espacio. A diferencia de los parámetros extrínsecos, los intrínsecos son internos a la cámara y no dependen de su posición en el espacio.

En un sistema óptico ideal, cada uno de los puntos del objeto se proyectaría en un único punto de tamaño cero en la imagen. Sin embargo, un punto en el objeto real no es un punto en el plano de la imagen. El resultado es una distribución de intensidad en el volumen, el cual no tiene forma simétrica. En la práctica, la

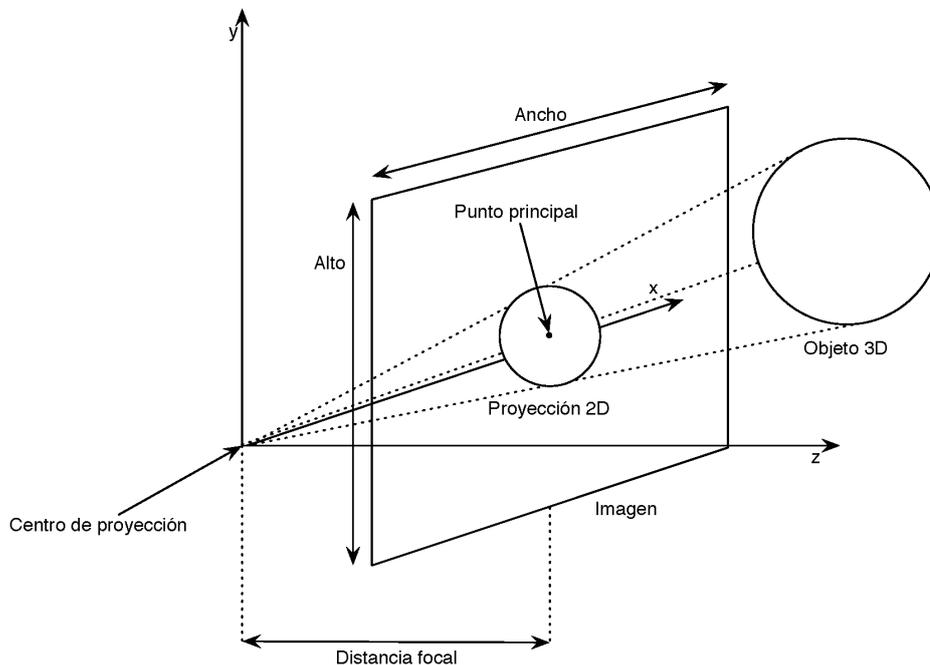


Figura 1.1: Modelo proyectivo simple (*pinhole model*).

distorsión es tanto mayor cuanto menor es la distancia focal, es decir, cuanto mayor es el ángulo de visión.

En aplicaciones de Visión por Computador, la distorsión radial es una aberración que tiene que ser estimada y corregida. Una lente típica tiene una distorsión de algunos píxeles, y generalmente no son visibles cuando se observa una escena natural. No obstante, cuando la imagen se usa para tomar medidas (fotogrametría), se hace necesaria la compensación de esa distorsión.

El modelo práctico de una aberración geométrica en una lente incluye principalmente dos componentes de distorsión. La primera componente es la **distorsión radial**, la cual es causada porque la lente proyecta el rayo más o menos que un caso ideal. La segunda, es el **desplazamiento del centro de distorsión** con respecto al punto central de la imagen.

La distorsión de cojín o de barril están centradas y son simétricas (fig. 1.2). Para aproximar la corrección necesaria, se suele usar un modelo matemático simple. Se trata de un polinomio de orden inferior dependiente de la distancia  $r$  del punto observado  $(x, y)$  respecto al centro de distorsión  $(x_d, y_d)$  del modelo. Las variables involucradas aparecen en la figura 1.3. Las coordenadas de un píxel después de la

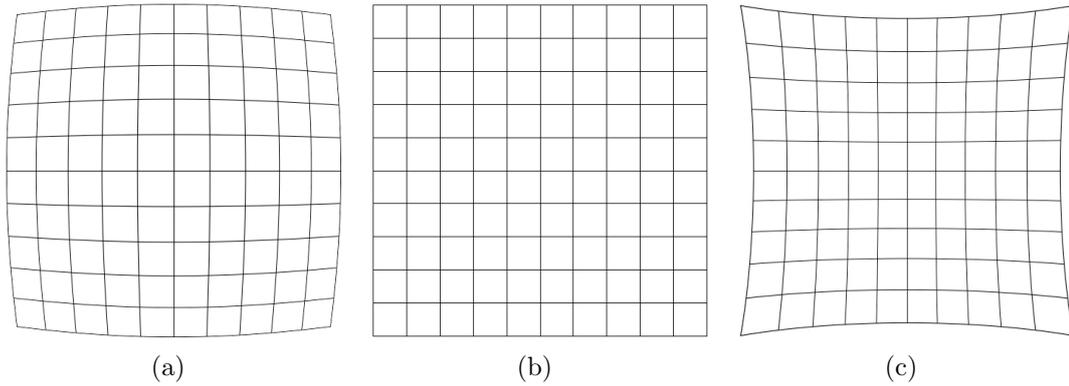


Figura 1.2: Distorsión radial sobre un grid de  $10 \times 10$  celdas: (a)distorsión tipo barril, (b)imagen original y (c)distorsión tipo cogín.

corrección son:

$$\begin{cases} \hat{x} = x + \Delta_x \\ \hat{y} = y + \Delta_y \end{cases}$$

La distorsión radial es aproximada como una función simétrica con respecto al centro de distorsión y dependiente de la distancia  $r$ :

$$r = \sqrt{(x - x_d)^2 + (y - y_d)^2},$$

la cual es medida desde el píxel  $(x, y)$  al centro de distorsión  $(x_d, y_d)$ .

De forma general, se usan polinomios de orden inferior que dependen de  $r$  para asegurar la simetría. Dicha aproximación es:

$$\begin{cases} \Delta_x = (x - x_d)(k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \\ \Delta_y = (y - y_d)(k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots) \end{cases} \quad (1.1)$$

La distorsión de la lente viene representada por los coeficientes  $k_i$ , generalmente denominados como parámetros de distorsión. Estos se obtienen experimentalmente para una lente particular, usando una imagen de calibración conocida, como son por ejemplo patrones regulares constituidos por diferentes estructuras. En la figura

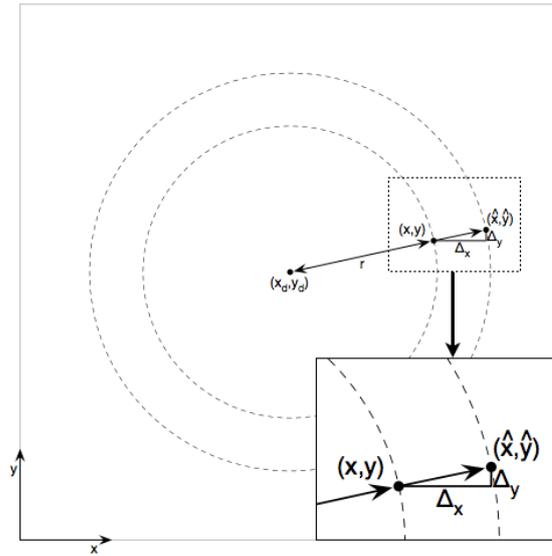


Figura 1.3: Modelo de distorsión radial:  $(x, y)$  coordenadas originales del punto.  $(\hat{x}, \hat{y})$  coordenadas del punto tras realizar la corrección.  $(x_d, y_d)$  coordenadas del centro de distorsión de la lente.  $(\Delta_x, \Delta_y)$  componentes de la corrección necesaria y  $r$  distancia entre el centro de distorsión y el punto sin corregir.

1.4, pueden verse algunos ejemplos de patrones de calibración usados [61]. Una descripción más exhaustiva acerca de diferentes patrones de calibración, así como sobre distintas técnicas usadas para el proceso de calibración de cámaras, puede encontrarse en [70].

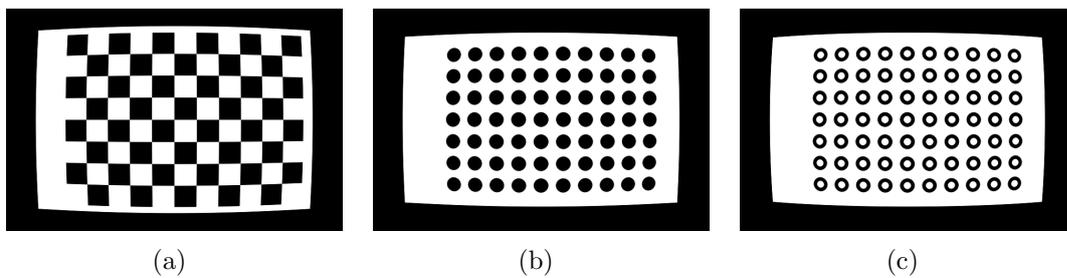


Figura 1.4: Diferentes ejemplos de patrones de calibración: (a) usando cuadrados, (b) círculos y (c) anillos [40].

De forma genérica, los algoritmos para la estimación del modelo de distorsión de la lente, se podrían dividir en los siguientes pasos:

1. Se calculan los bordes de la imagen usando un algoritmo detector de bordes.

- 
2. Se seleccionan algunos puntos de borde correspondientes a diferentes segmentos rectos 3D, que serán usados para ajustar los parámetros de distorsión del modelo.
  3. Si inicializa el vector de parámetros  $\mathbf{k} = (1, 0, \dots, 0)^T$ .
  4. Se elige cualquier par  $p, q \in \mathbb{Z}(1 \leq p, q \leq N_k)$  y se optimizan  $k_p, k_q$ .
  5. Se actualiza  $\mathbf{k}$  usando un factor de *zoom* tal que los puntos distorsionados y corregidos estén lo más cerca posible.

La detección de contornos sobre la imagen de entrada puede realizarse mediante cualquier detector de bordes. Uno de los que han sido usados de forma extensa, es el propuesto por John Canny en [19]. Se trata de un detector que cumple tres principios básicos: buena detección, buena localización y respuesta única. En el capítulo 2 entraremos en más detalle acerca de este detector.

Sin embargo, existen multitud de técnicas para detectar los bordes en una imagen. Recientemente, han tomado especial protagonismo métodos como el propuesto en [35]. En ese trabajo se propone un método para la detección de segmentos con precisión subpíxel que no precisa de parámetros. Para ello tiene en consideración el control preciso de los falsos positivos y falsos negativos resultantes de la detección. Para ello, generalizan el algoritmo propuesto por Desolneux et al. en [23], manteniendo los principios del control de falsos positivos/negativos, pero añadiendo precisión a la detección. Se realiza una adaptación de la transformada de Hough, en términos de realizar una búsqueda en un espacio de parámetros que representan segmentos, en lugar de detectar líneas como en la transformada tradicional. Se trata de un método adaptativo que usa la propia secuencia para automáticamente asignar los umbrales. En [36] se describe la implementación de dicho método para la detección de segmentos con precisión. En la figura 1.5 se muestra un ejemplo del resultado de detectar los bordes usando esta técnica en la imagen de un patrón de calibración.

Una vez se han detectado los puntos de interés presentes en la imagen, y con el objetivo de poder estimar la distorsión de la lente presente en ella, se hace necesario interpretar de alguna forma la información existente. Uno de los métodos clásicos usados para detectar formas a partir de los contornos es la transformada de Hough propuesta en [42] y patentada en [43]. En ellos se propone una técnica mediante la cual se detectan líneas rectas en la imagen, haciendo que cada punto borde vote por las líneas que vienen representadas en un espacio de votación en términos de orientación y distancia al origen. Sin embargo, este método no fue originalmente

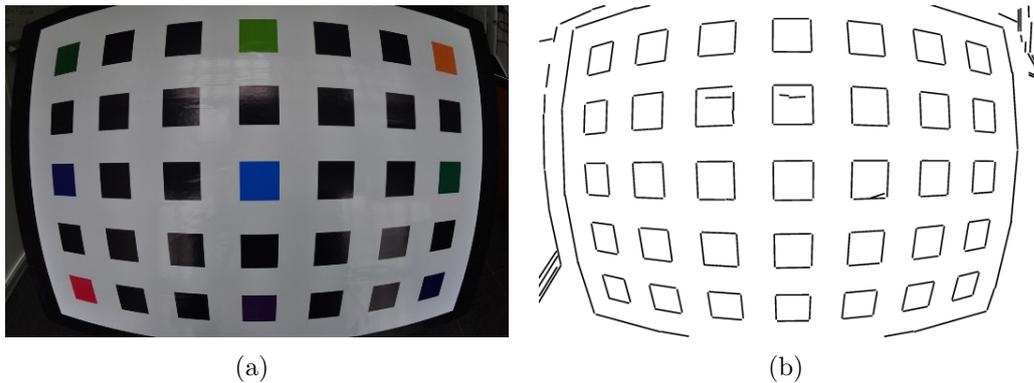


Figura 1.5: Ejemplo de aplicación del método *Line Segment Detector* a la imagen de un patrón de calibración: (a) imagen de entrada y (b) resultado de la detección de bordes.

concebido para abordar imágenes influenciadas por distorsión de la lente. En [25, 9], se introducen modificaciones al enfoque clásico, de forma que la técnica sea capaz de detectar otros tipos de estructuras en la imagen.

Se han hecho otras múltiples propuestas en relación a la detección de primitivas en la imagen. En [37], se propone un método para detectar segmentos rectos en la imagen, mediante el análisis de pequeños autovalores, explotando la relación entre dichos valores de la matriz de covarianza y el conjunto de los píxeles borde sobre una región conectada. Para ello, se parte de una imagen en la que ya se han detectado los contornos usando el método de Canny. Sin embargo, en el trabajo descrito en [49] se aborda el problema desde otro punto de vista, el análisis PCA. El análisis de componentes principales es un método estadístico que permite reducir el número de variables de un problema. En concreto, en este trabajo se propone el uso de esta técnica para la detección de líneas rectas y sus orientaciones, partiendo de un conjunto de bordes etiquetados. Para ello, se extraen inicialmente los contornos usando el método de Canny, para posteriormente etiquetarlos como filas y columnas. Tras ello, se aplica PCA para obtener las líneas rectas en la imagen.

En el trabajo descrito en [26], se presenta una transformada de Hough con un esquema de votación mejorado para la detección de líneas en tiempo real. El método propuesto evita que un punto vote a todas las rectas posibles, identificando grupos de segmentos colineales. Para cada uno de dichos puntos, se limita la región de votación mediante el uso de una pequeña región gaussiana de base elíptica. Esto permite restringir la participación sólo a unas pocas líneas del espacio de parámetros, dependiendo de la calidad del ajuste de sus píxeles a una línea recta.

---

Con ello se consigue una aceleración dentro del proceso de votación de todos los puntos al conjunto de posibles rectas que pueden pasar por ellos. Sin embargo, el conjunto de imágenes usadas en los experimentos, no presentan ningún tipo de distorsión de la lente. De la misma forma, en [31] se proponen algoritmos rápidos para la detección de líneas y círculos, basándose en mapas *hash* del gradiente inverso.

La detección de primitivas es una parte fundamental del proceso, ya que de ella depende poder estimar la distorsión de la imagen. Sin embargo, la mayor parte de las técnicas descritas no están específicamente orientadas a abordar el problema de la distorsión de la lente en imágenes, si bien algunas de ellas se han ido adaptando de alguna manera para dicho fin. Antes de pasar a describir algunos trabajos en el campo de la distorsión de lentes, existen otro tipo de factores que influyen en la distorsión y que deben ser también tomados en consideración. Tanto la profundidad de un punto de la escena respecto al plano de proyección, como la variación de la distancia focal, son características que influyen en la distorsión de la lente.

En [7], se propone un nuevo modelo de distorsión de la lente que depende de la profundidad, proporcionándose una técnica que permite estimar un nuevo modelo computacional de la lente de la cámara, dando un mayor nivel de precisión en la calibración. En el trabajo descrito en [28], se propone un modelo de corrección empírico, capaz de resolver ciertos problemas que no son capaces de resolver las aproximaciones de corrección geométricas tradicionales. Esta técnica también toma en consideración la estimación de la profundidad de campo.

En relación a la variación de la distancia focal mediante el mecanismo de *zoom*, en [4] los autores proponen nuevos modelos matemáticos, mediante los cuales se pretende incluir la modificación de la distancia focal en la estimación de la distorsión producida por la lente. Para dicho fin, usan una aproximación del modelo de distorsión radial de tipo polinomial de dos parámetros, en la cual se intenta minimizar una función objetivo que mide el error cuadrático medio que se produce al medir la distancia de los puntos de las primitivas a las líneas asociadas. La variación del modelo de distorsión de la lente se modela incluyendo la distancia focal efectiva  $f$ . Para verificar el correcto funcionamiento de su propuesta, realizan experimentos tanto sobre secuencias de imágenes sintéticas de un patrón de calibración, como en secuencias reales. Con ello, demuestran que la aproximación global de la secuencia es tan buena como la estimación del modelo de distorsión por cada fotograma, además de establecer que sólo usando un pequeño número de parámetros, se consigue estimar la distorsión para cualquier distancia focal efectiva del *zoom* de la lente.

Existen también otra serie de trabajos que toman en cuenta la distancia focal. En [47], se aborda el problema de los errores provocados por la metrología de patrones de calibración en la estimación precisa de los parámetros intrínsecos y extrínsecos para modelar un sistema de vídeo. Se propone una técnica para el cálculo de los parámetros tradicionales de calibración, al mismo tiempo que se extrae la geometría 3D del patrón usando múltiples imágenes en un algoritmo de calibración. Se consigue obtener de forma rápida y bien definida los parámetros de calibración tanto para distancias focales cortas como largas, haciendo uso de un material foto reflector para garantizar una detección precisa.

En el trabajo descrito en [68], se propone la estimación de la distorsión radial a partir de puntos de fuga detectados sobre una sola imagen. Para ello, estiman un único punto de fuga a partir de tres líneas distorsionadas en la imagen. Usan el modelo de distorsión de tipo división de un parámetro, con el centro de distorsión situado en el centro de la imagen. La estimación de la distorsión se realiza a partir de arcos circulares, detectados a partir de los contornos resultantes de aplicar el método de Canny. Esta detección se lleva a cabo a través de un método RANSAC, que usa tres tangentes para calcular el parámetro de distorsión y un punto de fuga, y dos tangentes más para la estimación de la distancia focal y dos puntos corregidos de fuga más. Los arcos se evalúan de forma que cumplan un umbral mínimo de distancia para cualquiera de los tres puntos de fuga estimados. Posteriormente, se lleva a cabo un procedimiento de refinamiento mediante optimización no lineal de la distorsión radial, la distancia focal y las posiciones de los puntos de fuga. Como medida se usa la consistencia de los arcos, definida en función de su longitud y el seno del ángulo que forman la normal del arco y la normal de la recta ideal que pasa por el punto de fuga.

Algunas técnicas para la estimación de la distorsión de la lente y su posterior corrección pasan por la selección, directamente sobre la imagen de entrada, de segmentos rectos 3D reales. En [5], los autores proponen una aproximación algebraica de la distorsión de la lente a través de la rectificación de líneas. Para dicho fin, se usa una aproximación general para determinar el polinomio de un modelo de distorsión radial mediante la imposición de que la proyección de líneas de la escena tridimensional en el plano bidimensional sean líneas rectas. A partir de ello, se propone una nueva técnica rápida, mediante la cual se obtienen los parámetros del modelo de distorsión de forma directa usando una nueva medida del error de la distorsión de la lente, basada en la desigualdad de Cauchy–Schwarz. Dichos autores en [6] proponen una implementación de esta metodología. La estimación del conjunto de parámetros del modelo de distorsión, se realiza a partir de la selección en la imagen de entrada de segmentos rectos, que aparecen distorsionados en la imagen resultante de la captura. En la figura 1.6, se muestra una secuencia en la

---

que se ha realizado la selección de diferentes primitivas mediante el marcado de puntos sobre un patrón de calibración.

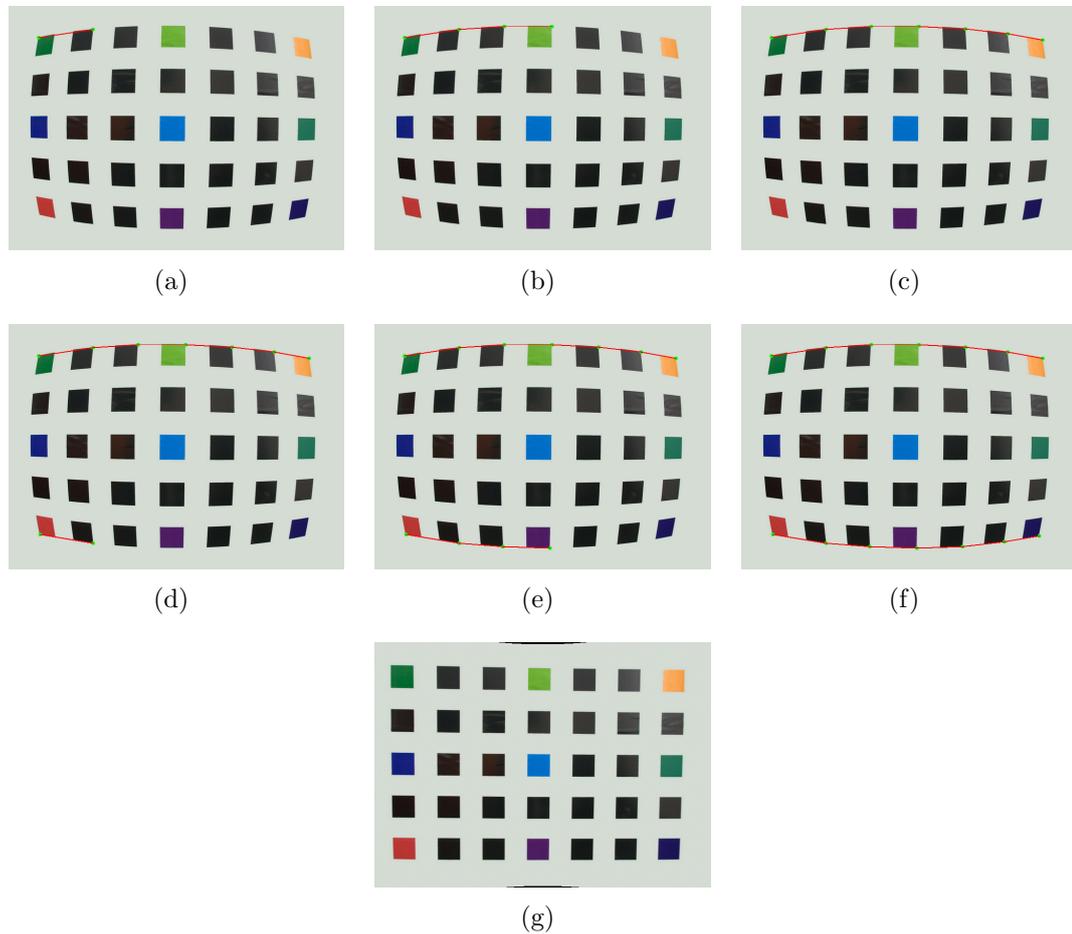


Figura 1.6: Selección manual de primitivas sobre la imagen de un patrón de calibración: (a)–(c) primera línea horizontal, (d)–(f) última línea horizontal y (g) imagen corregida.

En el trabajo descrito en [67], se presenta un método para la corrección de imágenes capturadas con lentes tipo ojo de pez, mediante proyección esférica. Usan una cámara de vigilancia con un ángulo de visión de hasta  $160^\circ$ . Sin embargo, se deben conocer de antemano parámetros como el centro óptico de la imagen y el radio de la esfera de proyección. La técnica propuesta se basa en el ajuste de elipses, pero requiere la extracción de los puntos distorsionados de las líneas rectas en la escena real, los cuales son seleccionados de forma manual, siendo necesarios al menos más de 5 puntos para extraer la elipse.

Atendiendo también al proceso de detección de las primitivas con intervención del usuario, en [22], se propone un método semi-automático para la corrección de la distorsión radial en imágenes y vídeos. La técnica propuesta se basa principalmente en dos pasos: tras obtener los contornos usando el método de Canny, se lleva a cabo la identificación automática o semi-automática de las líneas candidatas. En segundo lugar, se realiza una variación iterativa del parámetro de distorsión, hasta maximizar la rectitud de las líneas candidatas. Para ello se propone una nueva transformada de Hough basada en la corrección de la distorsión radial. Inicialmente debe conocerse un intervalo para el parámetro del modelo de distorsión, el centro y el número de muestras del parámetro de distorsión por iteración. La rectitud se evalúa mediante la obtención del máximo número de votaciones en el espacio de Hough. Sin embargo, su aplicación depende de la selección previa de una región de interés que incluya líneas rectas.

Generalmente, el modelado de la distorsión de la lente, cuanto mayor es su grado, requiere de un mayor número de parámetros en el modelo de distorsión. Sin embargo, cuando la imagen de entrada está afectada por una fuerte distorsión, los modelos de distorsión de tipo división parecen tener un mejor comportamiento, requiriéndose un número inferior de parámetros. En el trabajo descrito en [27], se puso de relevancia la utilidad del uso de los modelos de distorsión de tipo división. Sin embargo, este modelo fue originalmente propuesto en [50], pero fue este artículo el que atrajo una mayor atención sobre este modelo de distorsión de la lente. Una de las principales ventajas del uso de modelos de tipo división, es que precisan de un número inferior de parámetros para modelar distorsiones más severas. En el artículo, se propone una estimación lineal de la matriz fundamental. Para ello se hace uso del modelo división y se estima dicha matriz planteándola en función de un problema de autovalores cuadráticos. No se precisa de información previa, aunque han de proporcionarse al menos dos o más imágenes de la escena.

En relación a distintos modelos de distorsión, existen en la bibliografía trabajos acerca de diferentes estudios comparativos. En [44], se hace un estudio de diferentes métodos para la corrección de la distorsión radial y de otro tipo de distorsiones en cámaras con lentes tipo ojo de pez. Para ello, se aplican las diferentes técnicas en un conjunto de imágenes de test con diferentes grados y tipos de distorsión. Por otra parte, en el trabajo descrito en [63], se introducen los criterios de consistencia y universalidad. El criterio de consistencia se verifica mediante la medida del error residual obtenido al corregir una imagen que ha sido distorsionada con ese mismo modelo. De esta forma intenta demostrarse como de fiable es un modelo respecto a sí mismo. Por otro lado, la universalidad del modelo viene definida mediante el error residual resultante al corregir una imagen distorsionada con otro conjunto de modelos, revelando así su capacidad de afrontar diferentes fuentes de distorsión de

---

la lente.

En el trabajo descrito en [29], se realiza un estudio acerca del impacto de diferentes métodos usados para corregir la distorsión radial de la lente de tipo barril, clasificando su precisión en un contexto de diagnóstico médico por imagen. En concreto, la comparativa llevada a cabo se realiza sobre imágenes endoscópicas para la clasificación automática de patologías celíacas. Asimismo, se comparan una serie de métodos de interpolación usados para la reconstrucción de la imagen corregida. En el contexto descrito, los autores determinan que resulta difícil escoger algún método de estimación y corrección de la distorsión, puesto que los beneficios obtenidos en la clasificación dependen mucho del método usado para la extracción de características.

En [55], se analiza en profundidad la distorsión radial y los métodos para su identificación y eliminación. Especialmente se centra en el análisis de programas desarrollados para dicho fin, centrándose en la efectividad, precisión y aplicabilidad para la eliminación de la distorsión. Una de las aportaciones realizadas es el desarrollo de un nuevo algoritmo para medir dichos parámetros.

Respecto al uso de imágenes sintéticas para la evaluación de los modelos de distorsión de lentes o para procesos de calibración de cámaras, hemos visto algunos ejemplos de patrones de calibración, como los que aparecen en la figura 1.4. Una propuesta diferente es la que podemos encontrar en [34]. En este artículo se introduce un método *plumb line* para corregir y evaluar la distorsión de la lente de la cámara con alta precisión. El criterio de evaluación se define como la desviación estándar promedio de la rectitud de un conjunto de primitivas rectas, aproximadamente equiespaciadas, que son fotografiadas por la cámara uniformemente en todas las direcciones. De esta manera, su imagen cruza todo el campo de visión de la cámara. Para ello se usa un *arpa de calibración*, que es un marco en el cual se han estirado una serie de cuerdas para asegurar una rectitud muy precisa físicamente. El método comienza por estimar un modelo de distorsión polinomial de tercer orden, cuyo resultado se usa para inicializar uno de cuarto orden, continuando el proceso hasta orden 11. Tras ello, la estimación lineal se lleva a cabo de forma iterativa para refinar la precisión. La orientación de las líneas se inicia en primer lugar con la orientación de las líneas de regresión obtenidas de la minimización usando Levenberg–Marquardt. Éste método es usado para estimar los valores del polinomio del modelo de distorsión de forma creciente. En las siguientes iteraciones, las orientaciones toman los valores obtenidos en la iteración previa. La condición de parada es que se estabilice el resultado o que se alcance la precisión requerida. La función objetivo de la energía que se minimiza consiste en la suma de las distancias al cuadrado de los puntos a sus líneas de regresión,

considerando grupos de líneas e incluyendo un modelo de distorsión de la lente de tipo polinomial.

El mismo tipo de patrón de calibración descrito anteriormente se utiliza en el trabajo incluido en [62]. En este caso se usa el *arpa de calibración*, con el objetivo de obtener medidas más fiables, ya que no se puede asegurar que un patrón tradicional sea completamente plano. La distorsión presente en la imagen se evalúa mediante dos medidas diferentes y complementarias: el error cuadrático medio y el error máximo. El uso de este tipo de patrón de calibración proporciona varias ventajas. Por un lado, permite medir la precisión de la distorsión residual resultante de una corrección. Por otra parte, puede ser usado para la corrección de la distorsión, ya que la precisa detección de los puntos del arpa puede proporcionarse como entrada a cualquier método tradicional, dando resultados más precisos.

La amplia mayoría de los métodos se basan en la suposición de que la proyección en el plano imagen 2D de líneas rectas 3D reales, deben ser consecuentemente, rectas. En [24] se parte de la hipótesis de que en la escena capturada, independientemente del tipo que sea, hay líneas y segmentos rectos 3D. Una vez se extraen los contornos, se realiza una aproximación poligonal para detectar posibles bordes 3D en la escena. Tras ello, se buscan los parámetros de distorsión que minimizan la curvatura de dichos segmentos al ser proyectados en la imagen. Posteriormente, se vuelve a realizar una segunda aproximación poligonal, de forma que los segmentos anteriormente divididos a causa de la distorsión, deben aparecer como líneas rectas. Este procedimiento continua haciéndose iterativamente mientras no se alcance un mínimo del error de distorsión, que viene determinado por la suma de los cuadrados de las distancias de los puntos a las líneas.

Respecto a la selección automática de las primitivas en la imagen de entrada, existen diversos enfoques, pero generalmente se usa alguna modificación de una técnica estándar. En [66] hacen uso del modelo división descrito en [27], reformulándolo en términos de la ecuación de un círculo. Para obtener las líneas en la imagen, se determinan una serie de parámetros mediante el ajuste de las líneas distorsionadas a un círculo, usando un método directo de mínimos cuadrados. Con ello se obtiene una primera aproximación, que posteriormente se mejora realizando una minimización a través del método de Levenberg–Marquardt. Tras ello y partiendo de los arcos detectados, se realiza el cálculo del centro de distorsión de la lente. Con ello se calcula el valor de un parámetro de distorsión del modelo de tipo división, usando los parámetros obtenidos en los pasos anteriores. La principal ventaja de esta propuesta es que no precisa de información de los parámetros intrínsecos de la cámara, así como tampoco acerca de correspondencias con puntos 3D de la escena real.

---

En [32], se presenta un método numérico iterativo para la estimación automática y la corrección de la distorsión radial a partir de una sola imagen. Para ello, se consideran algunas restricciones como la localización de elementos rectilíneos y puntos de fuga. En ambos casos se parte de un modelo de distorsión de dos parámetros de tipo polinomial, para posteriormente, construir dos modelos relativos a las condiciones de colinealidad y a las restricciones geométricas de los puntos de fuga, respectivamente.

En [15, 16], los autores describen un método que estima la distorsión de la lente sobre una sola imagen sin necesidad de intervención humana ni el uso de patrones de calibración. La distorsión de la lente es modelada a través del modelo división de un parámetro. Usando un método RANSAC buscan automáticamente en la imagen arcos que no estén superpuestos. El centro de distorsión se calcula a partir de la resolución de un sistema lineal compuesto por tres de los arcos circulares detectados. Sin embargo, la principal limitación de este método es que cada arco circular debe ser una colección de puntos contiguos en la imagen y, por lo tanto, el método falla si no se encuentran ese tipo de arcos. Un ejemplo de ello puede obtenerse si se intenta aplicar esta técnica a un patrón de calibración compuesto por cuadrados.

En el trabajo descrito en [18], se propone la obtención automática de segmentos significativos mediante el uso de una transformada de Hough de orden superior. Para dicho fin, se usa una energía que mide la rectitud de las curvas en la imagen. Se asegura que se alcanza una mejor configuración determinando que los segmentos más significativos son aquellos que se obtienen en una configuración tal que la función de fiabilidad respecto del número de votos tiene valor máximo. Dicha función viene determinada por una potencia del número de votos en el espacio de Hough. La variación del exponente es usada para la estimación de un único parámetro del modelo de distorsión, seleccionando un valor para él de  $p \geq 3$ . De esta manera, se obtiene una primera estimación, que posteriormente es refinada mediante la búsqueda de picos en el espacio de Hough buscando un valor para  $k$  en un intervalo menor. Este intervalo viene determinado por el valor anterior obtenido  $\pm \delta k$ . Dicho incremento resulta de dividir el valor máximo del parámetro de distorsión entre el cuadrado de  $N \in [10 \dots 20]$ , siendo  $N$  el número de líneas a buscar.

Existen ciertos ámbitos en los que se hace necesario una corrección de la distorsión de la lente con el objetivo de facilitar labores de percepción. Un ejemplo de ello son las técnicas de cirugía no invasiva, que tienen principalmente dos inconvenientes: la ausencia de percepción táctil y de la profundidad. Además de ello, se tiene como añadido la distorsión introducida por el tipo de cámaras usadas en

el instrumental. En [48], se propone corregir la distorsión de este tipo de lentes, realizando una calibración y una deformación de la imagen en tiempo real. Se estima el parámetro de distorsión de un modelo división mediante el uso del cálculo de la entropía en la transformada de Hough. Tras detectar los contornos usando el método de Canny, se consideran una serie de parámetros candidatos representados por una tripleta formada por las coordenadas del centro y el parámetro de distorsión. Estos parámetros se usan para corregir los puntos borde. Se seleccionará aquel parámetro de distorsión que proporcione una entropía mínima. Para simplificar el cálculo de la transformada, se introduce el gradiente de la imagen en el modelo división, puesto que la entropía puede verse fácilmente influenciada por la presencia de múltiples líneas. Esta técnica incluye además la consideración del *zoom* de la lente. La corrección en tiempo real se lleva a cabo usando GPU.

Algunos enfoques no dan por hecho que existen líneas rectas en la escena real e intentan averiguar qué líneas curvadas en la imagen distorsionada son líneas rectas en la imagen corregida. En [17], se propone una técnica mediante la cual se obtienen automáticamente segmentos de líneas fiables para la estimación de los parámetros del modelo de distorsión de la lente. Para cada línea se calculan un conjunto de parámetros de distorsión, asignando un valor de fiabilidad a cada una de ellas. Partiendo de los bordes detectados, se obtienen las esquinas, que son usadas para generar segmentos curvos de líneas, eliminándose los más pequeños. Los parámetros del modelo de distorsión se obtienen a partir de la línea con la mayor fiabilidad pesada. Dicha fiabilidad depende de la inversa de la distancia promedio, a la que se le añade una constante y se multiplica por un peso que es el mínimo entre 1 y el doble de la distancia de la recta al centro de distorsión dividido por el radio. Sin embargo, este enfoque trata de determinar la distorsión de una forma local, centrándose en una única línea, cuando en realidad la distorsión de la lente es un efecto que se encuentra distribuido a lo largo de todo el dominio de la imagen de entrada.

La mayoría de las técnicas descritas usan la transformada de Hough o una modificación de la misma para la búsqueda de primitivas en la imagen, bien buscando líneas rectas o bien arcos de círculos. En el trabajo descrito en [46], se propone la obtención de los parámetros de distorsión de la lente a partir de una sola imagen utilizando *snakes*. El uso de *snakes* en Visión por Computador es común en técnicas de contornos activos, que generalmente son usados en áreas como el seguimiento de objetos, el reconocimiento de formas, la segmentación, la detección de bordes o las correspondencias entre imágenes obtenidas con un par estéreo. Sin embargo, el método propuesto plantea el uso de un tipo especial de *snakes*, las *snakes* de distorsión radial. Para ello, se precisa que el usuario dibuje en la imagen líneas que corresponden a líneas rectas 3D. Esta inicialización se toma en cuenta para

---

posteriormente usar *snakes* para obtener el mejor ajuste a las líneas, consiguiendo así los parámetros de distorsión. Para obtener dichos parámetros, se minimiza una función objetivo que tiene en cuenta los puntos de las *snakes* y un peso que es la fuerza del borde en cada punto. Los parámetros de distorsión se obtienen de forma sucesiva mediante la resolución de sistemas lineales de ecuaciones.

También existen métodos en los cuales se aborda la distorsión de la lente sin usar modelos de distorsión. En [33], se presenta un método de corrección de la distorsión de la lente que no usa ningún modelo. Este método considera que la distorsión puede ser cualquier difeomorfismo en el dominio de la imagen. La técnica propuesta parte de dos imágenes ligeramente diferentes tomadas sobre un patrón plano con parámetros constantes en la cámara. Posteriormente, se aplica un método SIFT de alta precisión para detectar los puntos de interés entre las imágenes tomadas y en el patrón original. Tras un proceso de eliminación de *outliers* mediante un bucle de validación, se interpolan las correspondencias restantes para obtener un campo denso de distorsión inverso. Aplicando dicho campo obtenido a todas las imágenes producidas por la cámara, hace que la misma se convierta en una cámara que virtualmente sigue el modelo proyectivo simple. Asimismo, se propone una mejora a esta técnica mediante la aplicación de un proceso iterativo. En dicha propuesta, inicialmente se aplicaría el método descrito y posteriormente se calcularía la forma física del patrón usando el método de Lavest et al. [47] sin usar modelo de distorsión.



# Capítulo 2

## Extracción de bordes en la imagen

### 2.1. Introducción

La primera de las fases a la hora de detectar primitivas en una imagen, consiste en encontrar una serie de puntos de interés que proporcionen algún tipo de información acerca de ellas. Este primer paso suele realizarse mediante la aplicación de detectores de contornos que permitan extraer los bordes presentes en la imagen. De esta manera, posteriormente se pueden aplicar técnicas que permitan agrupar estos puntos en estructuras más descriptivas, como rectas o elipses. Por tanto, es importante en este paso encontrar información descriptiva suficiente como para que sea significativa en el proceso de extracción posterior.

En este capítulo se describirá la técnica usada para la extracción de bordes. A través de este proceso se obtiene, por un lado, la imagen binaria que indica qué píxeles son borde. Por otro, también se tiene información acerca de la orientación, que viene proporcionada por las componentes del gradiente. De esta manera, no sólo se conoce la posición del punto que se pretende vincular a una primitiva, sino también de su orientación. Esto último permite que el proceso de búsqueda de líneas pueda restringirse a un cierto grado alrededor de la orientación del punto, puesto que la primitiva tendrá una cierta orientación predominante.

En primer lugar se empezará por describir el método de detección de contornos

usado. Se trata de un método clásico en la literatura del Procesado de Imágenes, el método de Canny. Esta técnica permite la extracción de bordes atendiendo a los criterios de buena detección, localización y respuesta única.

Tras ello, se hablará de una de las aportaciones de esta tesis, la introducción de un postproceso que elimina puntos que no proporcionan información tan buena como el resto debido a, por ejemplo, tener una alineación diferente.

## 2.2. Método de detección de contornos de Canny

El método de detección de bordes de Canny, comúnmente conocido como método de Canny, fue introducido por John Canny en [19]. Con él se pretenden detectar la colección de puntos borde que serán usados para estimar las líneas distorsionadas presentes en la imagen de entrada. Este método trata de cumplir tres condiciones: buena detección, buena localización y obtener una única respuesta para un solo borde. El operador de Canny está basado en un desarrollo analítico de optimización, en el que se parte de un modelo continuo de un escalón unidimensional.

Considérese una función escalón con amplitud “ $h_E$ ” afectada por un ruido blanco gaussiano de desviación típica  $\sigma_n$ . Supóngase que la detección de bordes se va a llevar a cabo mediante la convolución de una función continua unidimensional  $f(x)$  con una función respuesta impulsional antisimétrica  $h(x)$ , que tiene amplitud cero fuera del intervalo  $[-W, W]$ . El borde buscado de la función  $f(x)$  se marcará en donde aparezca el máximo local del gradiente, obtenido tras la convolución de  $f(x)$  con  $h(x)$ . Para determinar la función  $h(x)$  buscada se exige que ésta satisfaga los tres criterios citados anteriormente:

- **Buena detección.** Se maximiza la amplitud de la relación señal-ruido (SNR) del gradiente para obtener una mayor probabilidad de marcar el borde donde lo hay. Dicha relación para el modelo considerado es:

$$snr = \frac{h_E}{\sigma_n} S(h), \quad (2.1)$$

con

$$S(h) = \frac{\int_{-W}^0 h(x) dx}{\int_{-W}^W [h(x)]^2 dx} \quad (2.2)$$

- **Buena localización.** Los puntos del borde marcados por el operador deben estar tan cerca del centro del borde como sea posible. El factor de localización se define como:

$$LOC = \frac{h_E}{\sigma_n} L(h), \quad (2.3)$$

con

$$L(h) = \frac{h'(0)}{\int_{-W}^W [h(x)]^2 dx}, \quad (2.4)$$

donde  $h'(0)$  es la derivada de  $h(x)$ .

De acuerdo con estos dos criterios, el detector óptimo de bordes de tipo escalón es un escalón truncado (diferencia de escalones). Este operador, sin embargo, tiende a generar muchos máximos locales en su respuesta a bordes ruidosos de tipo escalón. Aunque estos bordes se deberían considerar erróneos de acuerdo con el primer criterio, sin embargo, no se ha considerado la interacción entre las respuestas en varios puntos próximos. Si se examina la salida del operador de diferencia de escalones, se ve que la respuesta a un escalón con ruido es un pico triangular con varios máximos en la vecindad del borde. Por ello es necesario incorporar el siguiente criterio que corrija este problema.

- **Respuesta única.** Debe haber una única respuesta para cada borde. La distancia entre picos del gradiente cuando sólo el ruido está presente, denotada por  $x_m$ , se establece como una fracción de  $k$  del ancho del operador, es decir:

$$x_m = kW. \quad (2.5)$$

Canny combina los tres criterios minimizando el producto  $S(h)L(h)$  sujeto a la restricción dada por la ecuación (2.5). Debido a la complejidad de esta formulación, no existe solución analítica de este problema.

Lo anteriormente expuesto se refiere a una señal unidimensional continua. en el caso de imágenes digitales (bidimensionales y discretas), el operador propuesto por Canny se aproxima mediante la derivada de la gaussiana en la dirección perpendicular al borde. En la práctica, aunque existen diferentes implementaciones, los pasos a seguir serían:

1. Calcular el módulo y la dirección del gradiente de la imagen suavizada aplicando el operador DroG (Derivada de la Gaussiana).

2. En la dirección del gradiente, eliminar puntos que no sean máximos locales del módulo (equivalente a encontrar el paso por cero en el operador LoG). Desechando los píxeles que no son máximos locales, se mejora la localización y se evitan falsas detecciones.

El proceso de eliminación de no-máximos locales se suele implementar siguiendo el borde en la dirección perpendicular al gradiente, considerando los 8-vecinos.

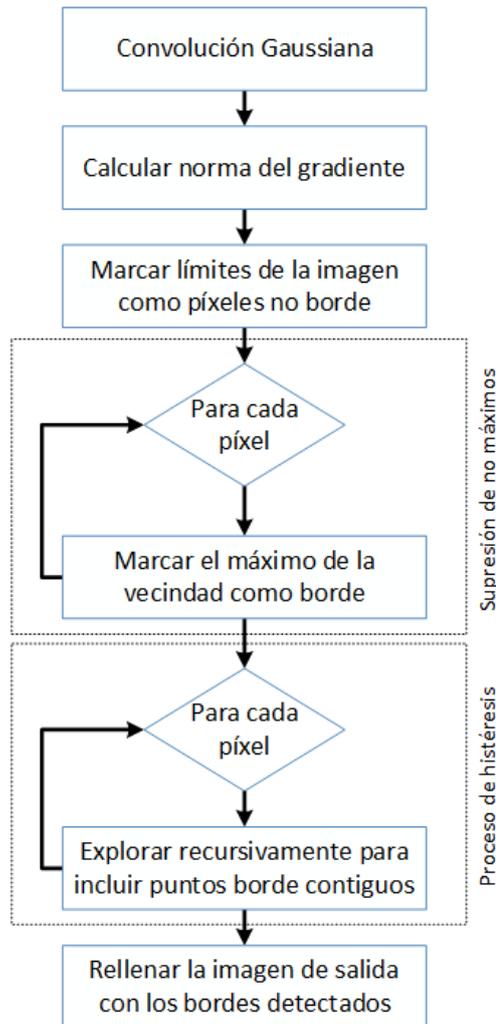


Figura 2.1: Diagrama de flujo del método de detección de contornos de Canny

Normalmente se utiliza una función de histéresis tal que el primer píxel de un segmento del borde debe tener un módulo del gradiente que supere un valor *gradiente\_alto*. Entonces se comienza a añadir píxeles vecinos en la dirección del

borde (perpendicular al gradiente) mientras que el valor del módulo de éste no caiga por debajo de un valor *gradiente\_bajo* (menor que *gradiente\_alto*) [45]. La figura 2.1 describe el diagrama de flujo que sigue este algoritmo, que a su vez se encuentra en el algoritmo 1.

## 2.3. Refinamiento de los bordes

Una de las aportaciones de esta tesis doctoral es la inclusión de un proceso de refinamiento de los bordes detectados. Antes de realizar la detección de primitivas y tras obtener los contornos usando el método de Canny, el proceso de refinamiento se usa para eliminar aquellos puntos que no aportan información significativa para el paso posterior. La idea de este procedimiento opcional radica en eliminar aquellos puntos que cuentan con una alta curvatura, así como preservar aquellos puntos que tengan cierta estabilidad en su orientación dentro de una cierta vecindad. Esto permite simplificar la colección de puntos y eliminar posibles *outliers*.

A partir de la detección de contornos realizada con el método de Canny, se obtiene información no sólo de la posición de los píxeles borde, sino también de su orientación. Además de ello, el proceso de refinamiento requiere de algunos parámetros más, principalmente relacionados con la vecindad que se va a procesar: su radio, el número mínimo de puntos borde dentro de la misma, el valor mínimo promedio del producto escalar de la orientación de los puntos dentro de dicha región y la distancia mínima entre puntos pertenecientes al contorno.

Ya que estamos interesados en extraer líneas a partir de los puntos borde detectados, se eliminan aquellos puntos en los que la orientación varía significativamente dentro de una determinada vecindad de un punto. Para cada punto  $p$  usamos como medida de la estabilidad de la orientación,  $EOSM(p)$ , la siguiente expresión:

$$EOSM(p) = \sum_{q \in \text{Puntos borde} \cap \text{Vecindad}(p)} \cos(\alpha_p)\cos(\alpha_q) + \sin(\alpha_p)\sin(\alpha_q) \quad (2.6)$$

donde  $\alpha_p, \alpha_q$  son las orientaciones de los puntos borde. Destacamos que cuanto menor es el valor de  $EOSM(p)$ , existe una mayor variación de la orientación del contorno en la vecindad de  $p$ . En particular, una esquina en la imagen tendrá un valor inferior de  $EOSM(p)$  que un punto borde que pertenezca a una línea recta.

**Algoritmo 1:** Detección de bordes usando el método de Canny, descrito en la figura 2.1.

---

```

input   : input: imagen de entrada.
           per_low, per_high: porcentajes para el umbral alto y bajo en [0,1].
output  : output: imagen con los bordes.
           sine, cosine: orientación de los bordes.
           edges: coordenadas de los bordes.

gauss_conv(input, blurred_image) ; // Convolución Gaussiana
grad(blurred_image, x_grad, y_grad) ; // Se calcula la norma del gradiente
foreach  $p \in \text{Dominio de la imagen}$  do
|   gradient_norm(p)  $\leftarrow \sqrt{x\_grad(p)^2 + y\_grad(p)^2}$ ;
end

// Se calculan los umbrales
low_threshold  $\leftarrow$  gradient_norm( per_low  $\cdot$  Number_pixels );
high_threshold  $\leftarrow$  gradient_norm( per_high  $\cdot$  Number_pixels );

// Supresión de no máximos
foreach  $p \in \text{Dominio de la imagen}$  do
|   if p es un máximo local en la dirección de máxima variación de la norma del
|   gradiente then
|   |   if gradient_norm(p)  $\geq$  high_threshold then
|   |   |   p es clasificado como borde
|   |   else
|   |   |   p es clasificado como máximo local del gradiente para ser considerado como
|   |   |   punto potencial a ser borde durante el proceso de histéresis
|   |   end
|   end
end

// Proceso de histéresis del método de Canny
foreach  $p \in \text{Dominio de la imagen}$  do
|   if p es clasificado como borde then
|   |   // Se explora recursivamente el vecindario de p para considerar puntos contiguos.
|   |   Cualquier punto contiguo cuyo gradiente es un máximo local y el valor de su
|   |   norma es mayor que low_threshold es clasificado como borde
|   end
end

// Se rellenan las imágenes de salida considerando la clasificación de bordes
foreach  $p \in \text{Dominio de la Imagen}$  do
|   if p es clasificado como borde then
|   |   output(p)  $\leftarrow$  255;
|   |   cosine(p)  $\leftarrow \frac{x\_grad(p)}{gradient\_norm(p)}$ ; sine(p)  $\leftarrow \frac{y\_grad(p)}{gradient\_norm(p)}$  ; // Orientación del borde
|   |   edges.add(p) ; // Posición del borde
|   else
|   |   output(p)  $\leftarrow$  0;
|   end
end

```

---

Por tanto, eliminamos puntos que no pertenecen a bordes de un segmento de línea recta, mediante mantener solamente puntos borde que tienen un valor mínimo de la medida  $EOSM(p)$  (dado como un umbral). Además, para acelerar el algoritmo fijamos una distancia mínima entre los puntos borde y eliminamos cualquier punto borde  $q$  tal que  $EOSM(q) < EOSM(p)$ , con  $q$  perteneciente a una vecindad cuadrada de  $p$  con un radio dado por dicha distancia mínima. El algoritmo 2 describe los diferentes pasos para refinar los contornos, proporcionados por el método de Canny, usando los parámetros descritos en la tabla 2.1. Es preciso destacar que, haciendo uso de este procedimiento de refinado, podemos reducir significativamente el número de puntos borde y, por lo tanto, el coste computacional del algoritmo se ve también disminuido.

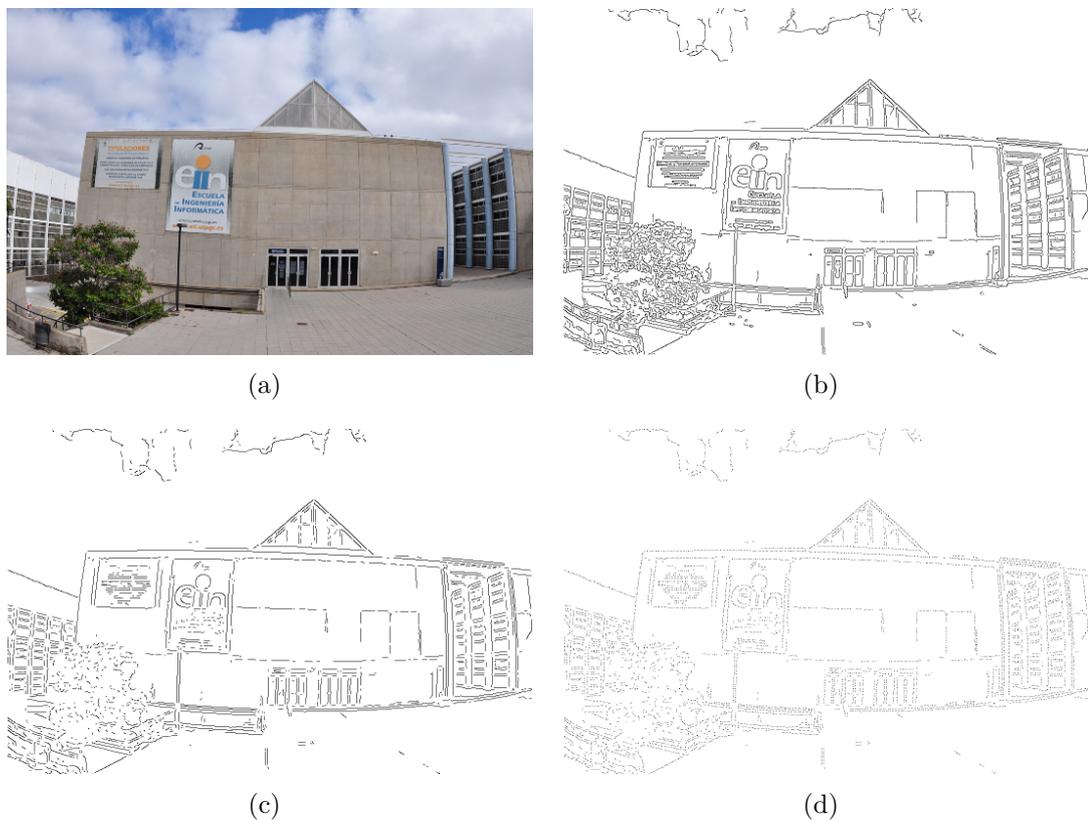


Figura 2.2: Ejemplo del proceso de refinado tras realizar la detección de contornos con el método de Canny: (a) imagen de entrada, (b) resultado de la detección de bordes, (c) bordes tras eliminar los puntos con una alta curvatura y (d) después de eliminar los puntos en una vecindad de  $3 \times 3$  píxeles.

En la figura 2.2 se muestra un ejemplo del procedimiento para el refinado de los contornos. A partir de la imagen original (fig. 2.2(a)), obtenemos los bordes

usando el método de Canny (fig. 2.2(b)) con los parámetros descritos en la tabla 2.1. Posteriormente, los puntos  $p$  con un valor alto de  $EOSM(p)$  son eliminados (fig. 2.2(c)) en base a los parámetros descritos en la tabla 2.2. Finalmente, solo aquellos puntos con el menor valor de  $EOSM(p)$  en el entorno de  $p$  son los que se preservan (fig. 2.2(d)), estableciendo la distancia mínima entre puntos borde a uno.

<b>Método de Canny</b>	
Parámetros	Valor por defecto
Desviación estándar de la Gaussiana	2.0
Umbral inferior para el método de Canny	0.7
Umbral superior para el método de Canny	0.8

Tabla 2.1: Parámetros del método de Canny.

<b>Proceso de refinado de los contornos</b>	
Parámetros	Valor por defecto
Radio de la vecindad	2.0
Número mínimo de puntos borde	2
Mínimo producto escalar promedio	0.95
Distancia mínima entre puntos borde	1

Tabla 2.2: Parámetros del proceso de refinado de los bordes.

Las consecuencias del uso de este proceso son principalmente dos. Por un lado, la información aportada a los pasos posteriores es más relevante. Esto se debe a que el proceso de votación de Hough se basa en la exploración de pares (distancia, orientación), los cuáles identifican a las rectas. Ya que durante el proceso de limpieza se eliminan aquellos puntos con altas curvaturas y con orientaciones dispares, el paso posterior se concentra en aquellos puntos que son candidatos más probables de formar parte de una recta. Por otro lado, el decremento de los puntos que van a ser procesados posteriormente, reduce de manera sensible el coste computacional del proceso de votación de la transformada de Hough.

---

**Algoritmo 2:** Proceso de refinamiento de los contornos.

---

**input** :  $p, \alpha_p$ : coordenadas y orientaciones de los puntos borde.  
*neighborhood\_radius*: radio de la vecindad a considerar.  
*min\_neighbor\_points*: número mínimo de puntos dentro de la vecindad.  
*min\_EOSM*: valor mínimo permitido de EOSM( $p$ ).  
*min\_distance\_point*: distancia mínima entre puntos borde.  
*width, height*: ancho y alto de la imagen.

**output** : Puntos borde refinados.

```

foreach  $p \in \text{Puntos borde}$  do
  // Calculamos una medida de la estabilidad (2.6) de la orientación dentro de una
  // vecindad cuadrada del punto  $p$ .
  EOSM( $p$ )  $\leftarrow$  0, number_neighbor_points( $p$ )  $\leftarrow$  0;
  for  $q \in \text{vecindad}(p)$  AND  $q \in \text{Puntos borde}$  do
    | EOSM( $p$ )  $\leftarrow$  EOSM( $p$ ) +  $\cos(\alpha_p) \cdot \cos(\alpha_q) + \sin(\alpha_p) \cdot \sin(\alpha_q)$ ;
    | number_neighbor_points( $p$ )  $\leftarrow$  number_neighbor_points( $p$ ) + 1;
  end
end

// Se eliminan puntos de la colección de bordes de acuerdo con el número mínimo de puntos
// vecinos y la estabilidad de la orientación del punto
foreach  $p \in \text{Puntos borde}$  do
  | if number_neighbor_points( $p$ ) < min_neighbor_points OR
  | EOSM( $p$ ) < number_neighbor_points( $p$ )  $\cdot$  min_EOSM then
  | | remove_point ( $p$ );
  | end
end

// Se eliminan iterativamente puntos borde cuyo número de bordes vecinos es inferior que el
// radio de la vecindad
while el número de puntos borde eliminados es mayor que 0 do
  | // Calculamos el número de puntos borde en cada vecindad del borde
  | // Eliminamos puntos borde con un número de bordes vecinos inferior que el radio de la
  | // vecindad
end

// Para acelerar el proceso sólo permitimos 4 iteraciones del while previo

// Posteriormente, para cada punto borde  $p$  eliminamos cualquier otro punto borde en una
// vecindad cuadrada centrada en  $p$  de radio min_distance_point con una mayor variación en
// la orientación del borde usando la medida dada por EOSM( $p$ ). Destacamos que este
// proceso depende del orden en que recorremos los puntos borde pero, en la práctica, dicho
// orden no tiene influencia significativa en la salida del algoritmo.

```

---



## Capítulo 3

# Estimación automática de modelos de distorsión de un parámetro polinomiales y de división usando la transformada de Hough extendida

### 3.1. Introducción

Como se ha descrito con anterioridad, la distorsión de la lente es un problema típico en los sistemas de capturas y que viene determinado por la propia geometría de la lente, así como por parámetros propios del sistema, como la variación de la distancia focal  $f$  mediante el mecanismo de *zoom*. En cualquier caso, esto determina en como una escena del mundo real se proyecta en un plano bidimensional que representa la imagen capturada.

Existen múltiples modelos de distorsión de la lente, pero de forma clásica, los modelos de distorsión radial han sido los más extensamente utilizados; en especial, aquellos basados en modelos de tipo polinomial. Estos representan la distorsión mediante un polinomio de cierto grado, determinando cómo se modifica la distancia de los puntos de la imagen respecto al centro de distorsión. Generalmente este centro se asume como el centro geométrico del plano de proyección. Sin embargo,

como se verá en capítulos posteriores, esta asunción no siempre es del todo correcta.

Para el modelado de la distorsión de la lente se hace necesario llevar a cabo, de alguna forma, una interpretación de la escena que se está capturando. Para ello, generalmente se intentan detectar formas conocidas en la escena, como primitivas. El uso extendido de patrones de calibración en Visión por Computador, hace que normalmente lo que se intente es detectar características de dichos patrones. Estos patrones consisten en una serie de formas, generalmente cuadrados o círculos, que se encuentran ordenados de una cierta manera. Lo que se intenta localizar son formas que den cierta información sobre la escena, como son los contornos de los cuadrados o puntos determinados en la imagen capturada. Existen muchas técnicas para dicho fin, pero la mayoría precisan de cierta intervención humana que, de alguna forma, proporcionen información sobre lo que se intenta localizar en la escena. Dicha intervención puede venir dada desde la selección de una cierta área donde buscar las características que se desean encontrar, hasta marcar de forma manual distintos puntos de la forma que se desea detectar en la imagen.

Una de las principales aportaciones de la presente tesis doctoral es el uso de una transformada de Hough extendida. Desde el punto de vista de la interpretación de la escena con el objetivo de conseguir información de la misma resulta de interés, en el caso que nos ocupa, la detección de líneas rectas. Cuando una imagen está distorsionada, ocurre que la proyección de dichas rectas en el plano tiene como consecuencia directa la curvatura de las mismas debido al efecto de la distorsión de la lente. La transformada clásica de Hough no es capaz de abordar este problema en el caso de presencia de distorsión. En ese caso, el algoritmo proporciona como salida líneas rectas, pero que han sido divididas en varios segmentos debido a la distorsión. Con la inclusión de la distorsión como una nueva dimensión del espacio de Hough, se pretende que dichas rectas sean obtenidas como tal, una estructura completa, en lugar de partidas en varios segmentos rectilíneos.

El presente capítulo se estructura de la siguiente manera: partiendo de la transformada de Hough tradicional para la detección de líneas rectas, se hablará de la inclusión de una nueva dimensión en dicho espacio, para así incluir la distorsión de la lente. Posteriormente, se describe el proceso de optimización del parámetro de distorsión, mediante la reducción de la distancia entre las primitivas detectadas y los puntos borde. Finalmente, se incluye el proceso de corrección de la distorsión, haciendo uso del modelo de distorsión con el parámetro optimizado a partir del conjunto de primitivas detectadas.

## 3.2. Incluyendo la distorsión de la lente en el espacio de Hough

La transformada de Hough es una técnica ampliamente usada en procesamiento de imágenes que permite extraer formas relevantes presentes en una imagen. Estas formas deben estar expresadas en forma paramétrica y, por lo tanto, esta técnica es principalmente usada para detectar formas geométricas como líneas, círculos o elipses. Inicialmente fue concebida para detectar líneas [42], pero más tarde se introdujo la transformada generalizada de Hough con el objetivo de detectar formas más complejas [9, 25].

Con el objetivo de encontrar formas que no sean del todo perfectas, la transformada de Hough usa un proceso de votación dentro de un espacio de parámetros. Esto se traduce en una estructura en la que se van acumulando las votaciones. Por último, los máximos locales dentro de dicha estructura proporcionan las formas detectadas. Una de las principales ventajas de la transformada de Hough es que no se ve afectada por huecos en los contornos de las formas. Sin embargo, una de las desventajas es el coste computacional, que principalmente depende del número de parámetros, como por ejemplo los grados de libertad de la forma y el número de valores que pueden tomar. Recibe el nombre de transformada porque la ecuación de una recta en el plano  $b = -ax + y$  se representa en un espacio de parámetros en el que en lugar de ser  $(x, y)$  las variables, lo son  $(a, b)$ . A esa transformación entre el plano imagen (coordenadas  $(x, y)$ ) y el plano o espacio de parámetros (coordenadas  $(a, b)$ ) es a lo que se denomina *Transformada de Hough* (fig. 3.1).

En dicho espacio de parámetros (también llamado espacio de Hough), cada punto  $(x, y)$  del plano imagen se convierte en una recta de pendiente  $-x$  y de ordenada en el origen  $y$ , que representa el lugar geométrico de todas las rectas que en el plano imagen pasan por el punto  $(x, y)$ . Por lo tanto, la intersección en el espacio de parámetros de dos rectas  $(x_i, y_i)$  y  $(x_j, y_j)$  da lugar al par  $(a', b')$  correspondiente a la recta que pasa por los puntos  $(x_i, y_i)$  y  $(x_j, y_j)$ . De esta manera, puntos alineados en el plano se transforman en el espacio de Hough en rectas que se cortan en un único punto  $(a', b')$ .

Para aplicar la transformada de Hough, es preciso discretizar el espacio de parámetros en una matriz de acumuladores. Esta discretización se realiza sobre los intervalos  $(a_{min}, a_{max})$  y  $(b_{min}, b_{max})$  que definen el conjunto de rectas buscadas. A esta estructura se le suele denominar acumulador y viene a estar representada por una matriz en cuyas posiciones se encuentra la cantidad de puntos que recibe cada

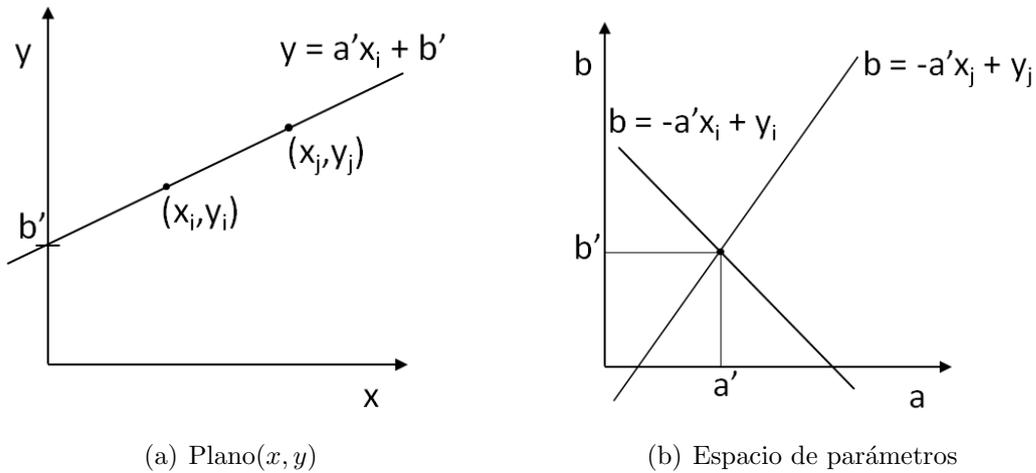


Figura 3.1: Representación de una línea recta: (a) en el plano  $(x, y)$  y (b) en el espacio de parámetros.

recta a la que está asociada dicha posición (la posición representa a una recta con un determinado ángulo y orientación).

El siguiente paso es evaluar la ecuación anterior para cada punto  $(x_k, y_k)$  de la imagen y cada uno de los valores de  $a$  en el intervalo  $(a_{min}, a_{max})$ . Los valores de  $b$  obtenidos se redondean al valor discreto más cercano. Así, si al tomar  $a_i$  se obtiene  $b_j$ , se incrementa en uno el valor de votos de la celda  $A(i, j)$ , es decir,  $A(i, j) \leftarrow A(i, j) + 1$ .

Cuando finaliza este proceso, el número de votos obtenidos en cada celda indica el número total de puntos que satisfacen la ecuación de la recta correspondiente. Por tanto, las celdas con mayor valor del acumulador de votos forman el conjunto de rectas detectadas.

Un problema de utilizar la ecuación explícita de la recta es que, tanto la pendiente como la ordenada en el origen tienden a infinito conforme la recta se acerca a posiciones verticales. Para evitar este problema se usa la representación normal de la recta dada por  $x \cos \theta + y \sin \theta = d$ , donde  $d$  es la distancia de la recta al origen y  $\theta$  es el ángulo que forma la normal con el eje  $x$  (fig. 3.2).

La forma de construir el acumulador en el plano  $\theta - d$  es similar al descrito anteriormente, aunque para cada punto que pertenece a la recta se obtiene en dicho plano una senoide en lugar de una recta como ocurría con la representación explícita [45].

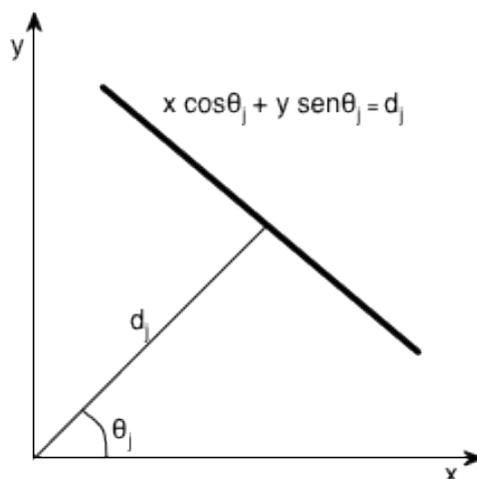


Figura 3.2: Representación de la recta en el espacio de Hough

La extracción de líneas a partir de una única imagen ha sido ampliamente estudiada [20, 37, 49]. Sin embargo, en nuestro caso en particular, tratamos de extraer líneas en imágenes que tienen un alto grado de distorsión, lo que supone un problema complejo en el campo del reconocimiento de formas. Si no se considera la distorsión, la transformada clásica de Hough descrita falla, ya que los puntos pertenecientes a una misma línea no votan por el mismo conjunto de parámetros, siendo por tanto considerados como segmentos de diferentes líneas. Por esta razón, el parámetro de distorsión debe ser incluido en el espacio de Hough. Por tanto, una línea en nuestro espacio vendrá descrita por tres parámetros: distancia al origen, orientación y distorsión.

Usamos un modelo de distorsión basado en una de las dos funciones siguientes:

$$L(r) = 1 + k_1 r^2 \quad o \quad L(r) = \frac{1}{1 + k_1 r^2}, \quad (3.1)$$

dependiendo de si se hace uso del modelo de distorsión de la lente polinomial o división, respectivamente.  $k_1$  es el parámetro de distorsión. Cuando la distorsión de la lente es corregida, la distancia  $r$  desde un punto al centro de distorsión es modificado de la siguiente forma:

$$\hat{r} = (1 + k_1 r^2)r \quad o \quad \hat{r} = \frac{r}{1 + k_1 r^2}, \quad (3.2)$$

para cada uno de los modelos.

Por razones prácticas, en lugar de considerar el valor del parámetro de distorsión dentro de este nuevo espacio de Hough, usamos un parámetro de distorsión normalizado  $p$ , que viene dado por el porcentaje de corrección de la distancia desde el centro de distorsión a la esquina más alejada,  $r_{max}$  (figura 3.3):

$$p = \frac{\hat{r}_{max} - r_{max}}{r_{max}}. \quad (3.3)$$

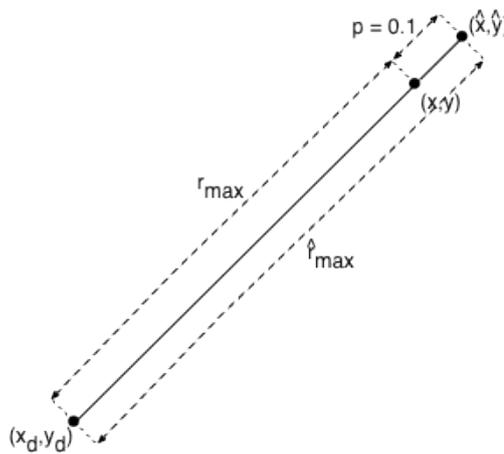


Figura 3.3: Representación de una variación de un diez por ciento de la distorsión mediante el parámetro normalizado  $p$ :  $(x_d, y_d)$  centro de distorsión,  $(x, y)$  punto original,  $(\hat{x}, \hat{y})$  punto corregido,  $r_{max}$  distancia del centro de distorsión al punto original,  $\hat{r}_{max}$  distancia desde el centro de distorsión al punto corregido y  $p$  parámetro de distorsión normalizado.

En particular, teniendo en cuenta los modelos presentados en la ecuación (3.1) y sustituyendo  $\hat{r}_{max}$  por la función correspondiente:

$$p = k_1 r_{max}^2, \quad (3.4)$$

para el caso polinomial, mientras que para el modelo división:

$$p = \frac{-k_1 r_{max}^2}{1 + k_1 r_{max}^2}. \quad (3.5)$$

---

### 3.2. Incluyendo la distorsión de la lente en el espacio de Hough

De manera equivalente se puede obtener la correspondencia con el parámetro de distorsión  $k_1$  de la forma:

$$k_1 = \frac{p}{r_{max}^2} \quad \text{o} \quad k_1 = \frac{-p}{(1+p)r_{max}^2}, \quad (3.6)$$

respectivamente. El uso de un parámetro de distorsión normalizado en lugar de  $k_1$  proporciona ciertas ventajas. Por un lado, usar el porcentaje  $p$  como parámetro de distorsión en el espacio de Hough nos permite independencia de la resolución de la imagen, puesto que está normalizado en términos de corrección de la distorsión. Por otra parte, la transformación dada por (3.2), debe ser una función creciente en el intervalo  $[0, r_{max}]$ . De otra manera, el modelo de distorsión de la lente no está bien definido. En concreto, su derivada debe ser positiva. Esto, en el caso del modelo polinomial nos lleva a:

$$1 + 3k_1r^2 > 0 \quad \forall r \in [0, r_{max}] \Rightarrow k_1 > -\frac{1}{3r_{max}^2}. \quad (3.7)$$

En términos de  $p$ , la condición anterior lleva a:

$$p > -1/3. \quad (3.8)$$

Por otro lado, en el caso del modelo división tendremos que:

$$\frac{(1 + k_1r^2) - 2k_1r^2}{(1 + k_1r^2)^2} > 0 \Rightarrow 1 - k_1r^2 > 0 \Rightarrow k_1 < \frac{1}{r_{max}^2}. \quad (3.9)$$

Usando la relación descrita en (3.6) para el modelo división, en términos de  $p$  se obtiene que:

$$p > -0,5. \quad (3.10)$$

En la práctica, para incluir el parámetro de distorsión en el espacio de Hough, seleccionamos un intervalo  $[p_{min}, p_{max}]$  (con  $p_{min} > -1/3$  o  $p_{min} > -0,5$  en función

del modelo) y tomamos:

$$p \in P = \left\{ p_{min} + n \cdot \delta_p : n = 0, 1, \dots, \frac{p_{max} - p_{min}}{\delta_p} \right\}, \quad (3.11)$$

donde  $\delta_p$  es el paso de discretización para el parámetro de distorsión. Así pues, una línea en el nuevo espacio de Hough viene determinada por la tupla  $(d, \alpha, p)$ , distancia, ángulo y parámetro de distorsión, respectivamente.

Para llevar a cabo el procedimiento de votación, es necesario primero extraer los puntos borde en la imagen, así como estimar la magnitud y orientación del borde, como se ha descrito en el Capítulo 2. Aquellos puntos donde la magnitud es significativa votan por un conjunto de líneas en nuestro espacio de Hough. Para cada valor  $p \in P$  (ec. (3.11)), en primer lugar se corrigen las coordenadas de los puntos borde, así como su orientación, usando el modelo de distorsión asociado a  $p$ . Tras ello, cada punto borde vota por una línea sólo si la orientación del punto de borde es coherente con la orientación de la línea. Para ello, se impone que la diferencia angular entre la orientación del borde y la de la línea debe ser menor que un cierto valor  $\delta_\alpha$ .

Podemos asumir que la distorsión que afecta a todas las líneas en la imagen es la misma para todas ellas. Esto significa que se estima un único valor para el parámetro de distorsión en toda la imagen. Es preciso destacar que, en este punto, se usa un modelo de distorsión de un solo parámetro. Además, cuanto más largas sean las líneas, más significativa será la información que proporcionan. Por estas razones, las líneas que son demasiado cortas son directamente rechazadas, y sólo seleccionamos las  $N$  líneas más largas para cada posible valor del parámetro de distorsión. El valor real hará que los puntos de borde se ajusten a las ecuaciones de las líneas, en lugar de votar aleatoriamente por líneas arbitrarias.

Una vez que las  $N$  líneas han sido seleccionadas para cada valor del parámetro de distorsión, debemos estimar cuál es el correcto. Para dicho fin calculamos una medida de fiabilidad para cada valor del parámetro de distorsión, favoreciendo aquellas líneas con una mayor puntuación y, por lo tanto, los valores para los cuales se han detectado las líneas más largas. La medida de fiabilidad  $v_i$  para el

i-ésimo valor del parámetro de distorsión es:

$$v_i = \sum_{j=1}^n s_j^i, \quad (3.12)$$

donde  $s_j^i$  es la puntuación total de la línea  $j$ -ésima del  $i$ -ésimo valor ( $p_i$ ). Dicha puntuación se calcula de acuerdo a la colección de puntos borde que votan por la  $j^{th}$  línea en el espacio de Hough. En otras palabras, si  $line_j$  es la ecuación de la línea asociada a la  $j^{th}$  línea, entonces  $s_j^i$  viene dada por:

$$s_j^i = \sum_{p \in Edge : distance(line_j, C_i(p)) < 1} \frac{1}{1 + distance(line_j, C_i(p))} \quad (3.13)$$

donde  $C_i(p)$  es la transformación del punto  $p$  dada por el modelo de distorsión de la lente asociado al  $i^{th}$  valor del parámetro de distorsión. En la ecuación (3.12), para cada valor  $i^{th}$ , se usan las  $N$  líneas en el espacio de Hough para calcular  $v_i$ .

De esta manera, para localizar las  $N$  líneas más votadas, buscamos un máximo local en la matriz de votaciones de Hough. Con la idea de evitar líneas duplicadas debido a una pequeña variación en los parámetros, se establece un entorno del extremo local en el espacio de Hough y se rechazan aquellas líneas dentro de dicho entorno de la línea seleccionada. Para ello se establecen dos parámetros que determinan la diferencia mínima en la orientación y en la distancia al origen entre dos líneas. Cuando la orientación y distancia de una línea difiere, con respecto a la línea seleccionada, una cantidad menor que la establecida en dichos parámetros, la línea es rechazada.

Denotamos como  $p_0$  al valor  $p \in P$  donde se alcanza el máximo usando el criterio de la ecuación (3.12). La figura 3.6(c) ilustra como varía el máximo de la votación dentro del espacio de Hough frente al porcentaje de corrección determinado por el parámetro normalizado de distorsión  $p$ .

Asociado al mejor valor de  $p$ , y de acuerdo al criterio anterior (ec.(3.12)), obtenemos una colección de  $N$  líneas. Asociamos a cada línea  $j$  los puntos borde  $\{\bar{x}_{ji} = (x_{ji}, y_{ji})\}_{i=1, \dots, N}$  que cumplen dos condiciones: que la diferencia entre la orientación de la línea y la orientación del punto borde corregido ( $\hat{x}_{ji}^p, \hat{y}_{ji}^p$ ) sea menor que un umbral  $\delta_\alpha$ . Además, la distancia del punto corregido a la línea debe ser menor a un umbral  $max_d$ . Destacamos que el punto ( $\hat{x}_{ji}^p, \hat{y}_{ji}^p$ ) se obtiene a partir de

la ecuación (3.2) (usando el modelo polinomial o división) con el modelo asociado al valor de  $p$ . En el algoritmo 3 se describe el funcionamiento de la transformada de Hough modificada.

Una vez se han obtenido las  $N$  líneas más votadas con la máxima votación asociada al valor del parámetro de distorsión  $p$ , se realizan una serie de procedimientos más. En primer lugar, se realiza una depuración de las primitivas detectadas. La depuración de las líneas consiste en calcular la orientación mayoritaria de los puntos que conforman la primitiva, borrando aquellos con una orientación distinta y que quedan en una situación minoritaria frente a los demás. Para obtener dicho cálculo, se evalúa cada punto con respecto a su orientación y a las componentes de la recta a la que pertenece mediante la expresión  $(b \cdot \cos_p - a \cdot \sin_p)$ . En ella,  $(a, b)$  son los coeficientes de la recta y  $(\cos_p, \sin_p)$  la orientación del punto. De cara a contabilizar la orientación mayoritaria, simplemente se tiene en cuenta el signo de dicha evaluación.

Debido a la eliminación de parte de los puntos pertenecientes a las primitivas, es necesario recalcular la ecuación de las primitivas respecto del conjunto de puntos que quedan.

El siguiente paso llevado a cabo consiste en eliminar las líneas demasiado pequeñas y unir aquellas primitivas que se encuentren muy cercanas. Desde el punto de vista de la información que proporcionan las líneas detectadas, aquellas líneas que tienen un número pequeño de puntos no aportan información significativa. Ya que la distorsión es un efecto global en la imagen, lo más adecuado es tener las líneas más largas posibles. En la práctica, hemos considerado que aquellas líneas que cuentan con un número inferior a 20 puntos, no aportan información valiosa.

A la hora de considerar si es preciso unir dos primitivas, en primer lugar se calcula la distancia a la que se encuentran. Si ambas se encuentran lo suficientemente cerca, se evalúa la orientación del primero de los puntos de cada una de ellas. En caso de cumplirse que tienen el mismo signo, se procede a unir ambas líneas. Como consecuencia, aquellas líneas que se han unido, precisan nuevamente recalcular su ecuación, puesto que los puntos que forman parte de ellas se han modificado.

De esta manera es como finalmente queda constituido el conjunto de primitivas con el mejor ajuste a las líneas distorsionadas en la imagen de entrada. Junto con el valor inicial obtenido para  $p$ , corresponden la salida de la transformada mejorada de Hough.

En la figura 3.4 se representan las matrices de votación para los diferentes

### 3.2. Incluyendo la distorsión de la lente en el espacio de Hough

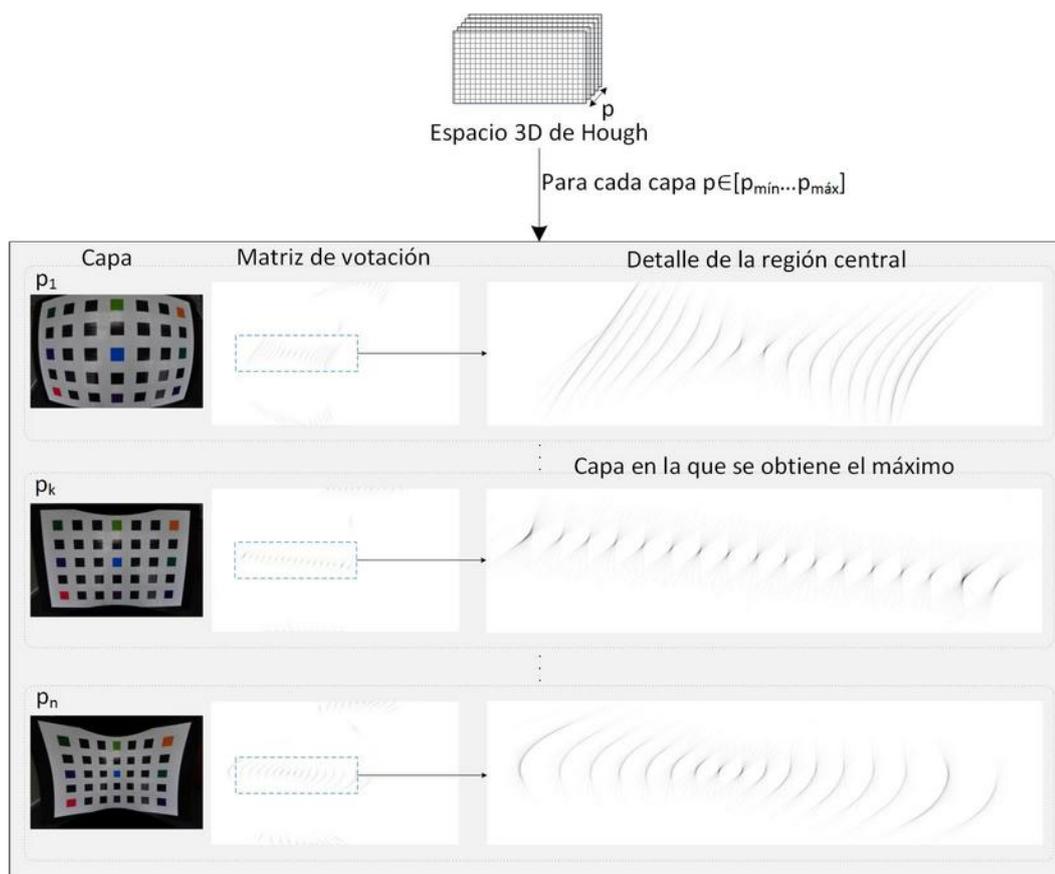


Figura 3.4: Representación de la matriz de votación dentro del espacio 3D de Hough para la imagen de un patrón de calibración.

valores de parámetro de distorsión discretizado  $p$  dentro del intervalo  $[p_{\min}, p_{\max}]$ . Para cada uno de ellos, la entrada es modificada, constituyendo el conjunto de puntos que van a votar por las líneas presentes en la imagen. En la parte derecha de la figura se destaca la región central de la matriz, que correspondería a las líneas verticales presentes en la imagen. Puede verse como en la capa en la que se alcanza el máximo, denotada como  $p_k$ , los votos se concentran en rectas específicas, en lugar de estar dispersados en el espacio de votación.

En la figura 3.5, se muestra en detalle el espacio de votación representado en la figura 3.4. En ambos casos se ha hecho un *zoom* sobre la misma región, a fin de poder comparar la concentración de votos en torno a varias posiciones en la matriz de votaciones. En la figura 3.5(a) se muestra cómo se ve la votación en el caso en que no se considera la distorsión. Puede observarse como, salvo en algún caso particular, los votos no parecen agruparse alrededor de una línea concreta,

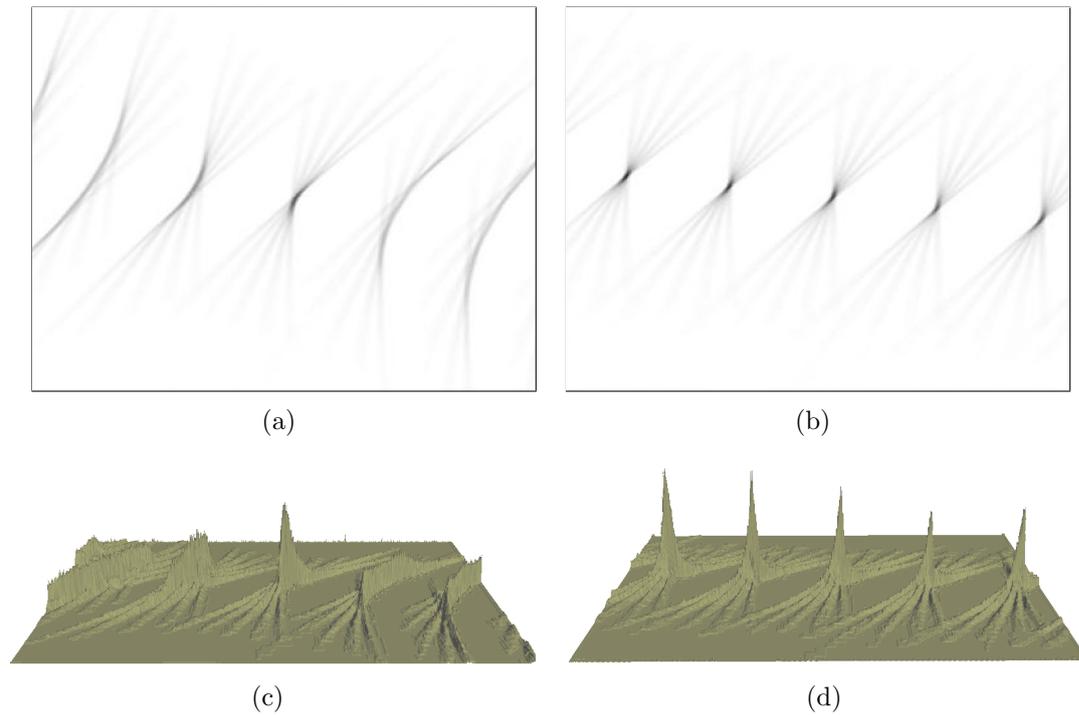


Figura 3.5: Detalle de la concentración de las votaciones mostradas en la figura 3.4: (a) votaciones sin tener en cuenta ningún parámetro de distorsión, (b) votaciones para el valor en el que se obtiene el máximo y (c)(d) sus respectivas representaciones en 3D.

sino que aparecen dispersos en el espacio de votación. Sin embargo, en la imagen representada en 3.5(b), que ilustra el caso en el que el valor del parámetro de distorsión proporciona el número máximo de votos, estos si están concentrados en un punto para cada una de las líneas que se muestran en la imagen del espacio de votación de Hough. Este efecto es más destacable si nos fijamos en la representación tridimensional de ambas regiones (figuras 3.5(c) y 3.5(d)), puesto que en el caso en el que se considera el parámetro de distorsión a la hora de realizar la votación (fig. 3.5(b) y su representación fig. 3.5(d)), la máximas cantidades de votaciones se encuentran concentradas en los puntos que representan a las primitivas.

Puede plantearse que, con el fin de garantizar que se seleccionan las líneas más largas, sería necesario incluir una función que así lo hiciese. En la práctica hemos observado que esto no es así. Sirva la figura 3.6 para argumentar esta afirmación. En ella se representa la evolución del total de votos dentro del espacio tridimensional de Hough al aplicar la técnica a la imagen de un patrón de calibración, usando diferentes funciones para la medida de fiabilidad. En 3.6(a) se represen-

ta la evolución usando la función  $f(x) = x^{3/2}$ , en 3.6(b) cuando se hace uso de  $f(x) = x^{0.9}$ , mientras que en 3.6(c) se usa la medida de fiabilidad presentada en la ecuación (3.12). Como puede observarse, independientemente de la función que se use, el perfil de la evolución de las votaciones es bastante similar. Y, lo que es más importante, el valor máximo se alcanza para el mismo valor  $p$  del parámetro de distorsión normalizado. Por tanto, a la hora de medir la fiabilidad para escoger el valor que proporcione el conjunto de primitivas con el mejor ajuste a las líneas distorsionadas en la imagen, no es necesario incluir ninguna de esas funciones.

En la figura 3.7 se describe el diagrama de flujo seguido por el algoritmo 3 que hace uso de la transformada mejorada de Hough a la que se le ha añadido la distorsión como una nueva dimensión.

Para ilustrar la mejora proporcionada por la inclusión de la distorsión como nueva dimensión en el espacio de votaciones de Hough, en la figura 3.8, se muestran las primitivas detectadas usando un modelo división de un parámetro, sobre la imagen de un patrón de calibración, usando la transformada de Hough clásica (fig. 3.8(a)) y con la mejora propuesta (fig. 3.8(b)).

### 3.3. Optimización del parámetro de distorsión

En la sección previa se ha descrito la técnica mediante la cual se introduce un nuevo parámetro en el espacio de Hough, la distorsión. Considerando este nuevo parámetro, el espacio de votaciones se convierte en un espacio tridimensional, en el que una recta viene denominada por la tupla  $(d, \alpha, p)$ , es decir, distancia al origen, ángulo y distorsión. Asimismo, se ha descrito como dicha introducción se hacía a través de la normalización del parámetro de distorsión en términos de la corrección de la distorsión, de forma que ésta variaba dentro del intervalo  $[p_{min}, p_{max}]$ . De esta anterior etapa se obtienen dos salidas. Por una lado, se obtiene el conjunto de primitivas que mejor se ajusta a las líneas distorsionadas en la imagen de entrada. Por otro, se tiene un valor para el parámetro de distorsión normalizado que ha permitido dicho ajuste.

El siguiente paso a considerar es la optimización del parámetro de distorsión. El objetivo es refinar el valor de  $p$  para obtener una aproximación más precisa del parámetro de distorsión. Previamente, se ha procedido a eliminar del conjunto de primitivas a aquellas que tienen un número pequeño de puntos asociados. Es necesario destacar que las líneas con un pequeño número de puntos no proporcionan

**Algoritmo 3:** Detección de líneas usando la transformada de Hough mejorada.

---

**input** :  $\{(x, y)\}$ : imagen con las coordenadas de los puntos borde.  
 $\{\alpha(x, y)\}$ : imagen con las orientaciones de los puntos borde.  
 $N$ : número máximo de líneas a detectar.

**output** : Estimación del parámetro  $p_0$  de distorsión normalizado y el conjunto de líneas más votadas

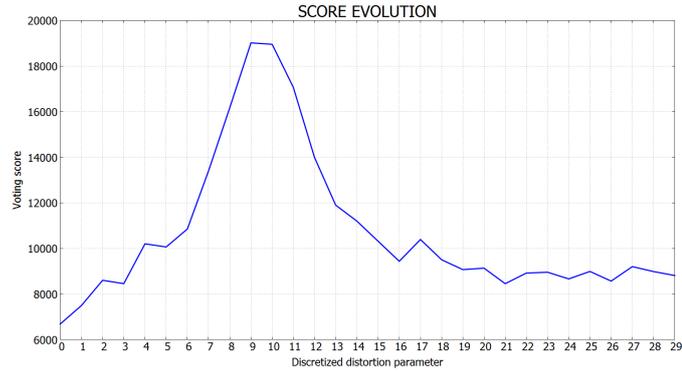
```

begin
     $max_{Hough} = -1$ ;
    for  $p = p_{min}$  to  $p_{max}$  do
         $k_1 = \frac{-p}{r_{max}^2(1+p)}$ ;
        for  $(x, y) \in Imagen\ de\ puntos\ borde$  do
             $r = \sqrt{(x - x_c)^2 + (y - y_c)^2}$ ;
             $\hat{x} = x_c + L(r)(x - x_c)$ ;
             $\hat{y} = y_c + L(r)(y - y_c)$ ;
            // Se actualiza la orientación del punto borde de acuerdo a la corrección de la
            // distorsión de la lente
             $\alpha_{new} = Update\_orientation(x, y, \alpha(x, y), k_1)$ ;
            for  $\beta \in [\alpha_{new} - \delta_\alpha, \alpha_{new} + \delta_\alpha]$  do
                 $d = Floor(-(\cos(\beta)\hat{x} + \sin(\beta)\hat{y}))$ ;
                for  $dist \in [d - 2, d + 2]$  do
                     $d_l = |\cos(\beta)\hat{x} + \sin(\beta)\hat{y} + d|$ ;
                     $votes(dist, \beta, p) = votes(dist, \beta, p) + \frac{1}{1+d_l}$ ;
                end
            end
        end
    end
    // Para las  $N$  líneas más votadas dadas por  $\{(d_{Hough,j}^p, \alpha_{Hough,j}^p)\}_{j=1}^N$  calculamos
     $sum = \sum_{j=1}^N votes(d_{Hough,j}^p, \alpha_{Hough,j}^p, p)$ ;
    if ( $sum > max_{Hough}$ ) then
         $p_0 = p$ ;
         $max_{Hough} = sum$ ;
        // Actualizamos las  $N$  líneas más votadas a  $\{(d_{Hough,j}^p, \alpha_{Hough,j}^p)\}_{j=1}^N$ 
    end
end
end
end

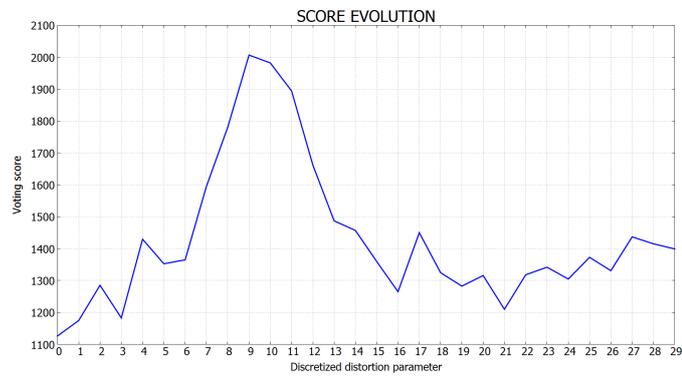
```

---

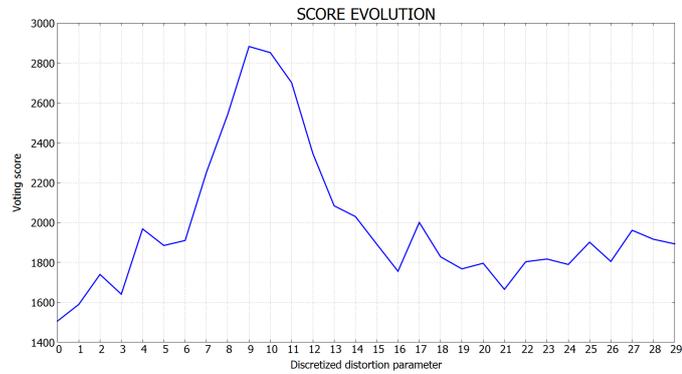
### 3.3. Optimización del parámetro de distorsión



(a)



(b)



(c)

Figura 3.6: Evolución del número de votos en el espacio tridimensional de Hough usando diferentes funciones para la fiabilidad: (a)  $f(x) = x^{3/2}$ , (b)  $f(x) = x^{0.9}$  y (c)  $f(x) = x$ .

información relevante (en comparación con las líneas más largas). Además, cuanto

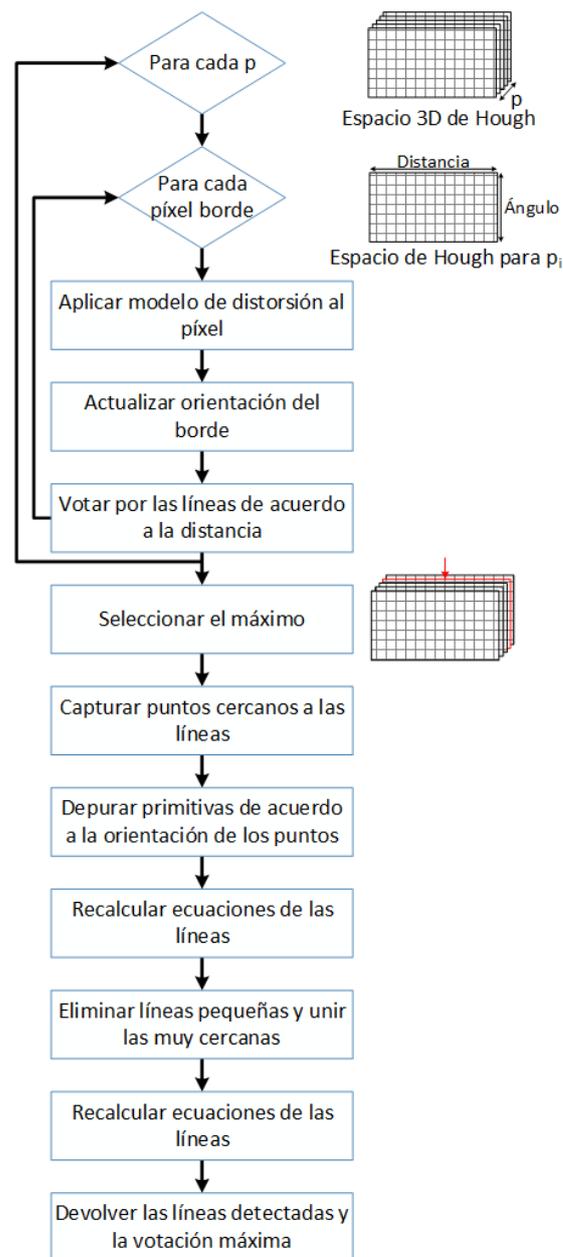


Figura 3.7: Diagrama de flujo de la transformada mejorada de Hough.

mayor sea el número de puntos de una línea, mayor es la probabilidad de que sea una línea real y no una línea espuria. Por otro lado, el parámetro inicial que representa el número máximo de líneas, se vuelve irrelevante puesto que es mayor que el número de líneas con más de un número mínimo de puntos asociados. Por tanto, finalmente obtenemos un número de líneas el cual es menor que el máximo

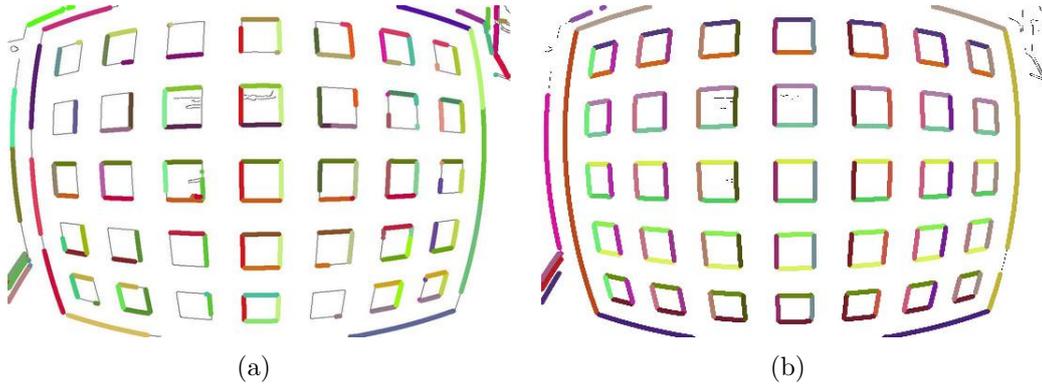


Figura 3.8: Comparativa de las primitivas detectadas sobre la imagen de un patrón de calibración usando el modelo de distorsión de la lente de tipo división de un parámetro: (a) con la transformada de Hough clásica y (b) con la mejora propuesta.

número anteriormente indicado. El valor del número de líneas obtenidas depende del número de líneas visibles en la imagen.

Para un dado  $p \in \mathfrak{R}$  y para una colección de puntos pertenecientes a líneas  $\{\bar{x}_{ji} = (x_{ji}, y_{ji})\}_{i=1, \dots, N(j)}$ , podemos optimizar la ecuación de la recta asociada minimizando la siguiente expresión:

$$D(\alpha, d) = \sum_{i=1}^{N(j)} (\cos(\alpha)\hat{x}_{ji}^p + \sin(\alpha)\hat{y}_{ji}^p + d)^2, \quad (3.14)$$

donde  $\{(\hat{x}_{ji}^p, \hat{y}_{ji}^p)\}_{i=1, \dots, N(j)}$  son los puntos corregidos usando el modelo asociado a  $p$ . Denotamos como  $(\alpha_j^p, d_j^p)$  al par donde la ecuación (3.14) alcanza su mínimo.

La optimización del modelo de distorsión de la lente se realiza mediante la minimización del promedio de la distancia al cuadrado de los puntos de las primitivas a su línea asociada. Por tanto, usando la notación introducida previamente, esta optimización consiste en minimizar la siguiente función de error:

$$E(p) = \frac{\sum_{j=1}^{Nl} \sum_{i=1}^{N(j)} (\cos(\alpha_j^p)\hat{x}_{ji}^p + \sin(\alpha_j^p)\hat{y}_{ji}^p + d_j^p)^2}{\sum_{j=1}^{Nl} N(j)}. \quad (3.15)$$

El valor inicial de  $p$  es  $p_0$ , obtenido mediante la transformada modificada de

Hough. En el algoritmo 4, se describe el proceso de minimización. Usamos una modificación del método de Newton–Raphson, donde incluimos un parámetro de convergencia  $\gamma$  para asegurar que el error decrece a lo largo de las iteraciones. Cuando  $\gamma$  está cerca de cero, el método se comporta como el método de Newton–Raphson y, cuando  $\gamma$  es grande, el método funciona como la técnica del gradiente descendente. Si el error para una solución candidata es menor que la anterior, aceptamos la nueva solución candidata y dividimos  $\gamma$  por 10 con el objetivo de acercarnos al método de Newton–Raphson. En otro caso, si rechazamos la solución candidata, multiplicamos  $\gamma$  por 10 para acercarnos al método del gradiente descendente y recalculamos la solución. Normalmente, el valor de  $\gamma$  es mayor cuanto más lejos se está de la solución y decrece en función de la cercanía a la misma. La idea de usar un parámetro de convergencia en algoritmos de optimización fue introducida por Levenberg en 1944 [51] en el contexto del conocido algoritmo de optimización de Levenberg–Marquardt.

---

**Algoritmo 4:** Algoritmo para optimizar  $p_0$  mediante la minimización de  $E(p)$ .

---

**input** :  $p_0$ : valor inicial del parámetro de distorsión normalizado.  
 $h$ : el paso discretizado para el cálculo de las derivadas.  
 $TOL$ : el valor de la tolerancia para la convergencia.

**output** : El modelo de distorsión de la lente optimizado

$p_{opt} = p_0$ ;  
 $\gamma = 1,0$ ;  
 $p = p_0 + TOL + 1$ ;  
**while** ( $|p - p_{opt}| > TOL$ ) **do**  
    //  $E$  está definida en la ecuación (3.15)  
     $E'(p_{opt}) = \frac{E(p_{opt}+h) - E(p_{opt}-h)}{2h}$ ;  
     $E''(p_{opt}) = \frac{E(p_{opt}+h) - 2E(p_{opt}) + E(p_{opt}-h)}{h^2}$ ;  
     $p_{new} = p_{opt} - \frac{E'(p_{opt})}{E''(p_{opt}) + \gamma}$ ;  
    **while** ( $E(p_{new}) > E(p_{opt})$ ) **do**  
         $\gamma = \gamma \times 10$ ;  
         $p_{new} = p_{opt} - \frac{E'(p_{opt})}{E''(p_{opt}) + \gamma}$ ;  
    **end**  
     $\gamma = \gamma / 10$ ;  
     $p = p_{opt}$ ;  
     $p_{opt} = p_{new}$ ;  
**end**  
**return**  $p_{opt}$ ;

---

### 3.4. Corrección de la distorsión

Una vez que se ha detectado el conjunto de primitivas con el mejor ajuste a las líneas distorsionadas de la imagen de entrada y se ha optimizado el valor del parámetro de distorsión normalizado, se pasa a corregir la distorsión.

Un uso interesante del método propuesto de detección automática de líneas es aplicarlo a la corrección de la distorsión en imágenes que muestran una distorsión radial significativa causada por la lente. Mediante el uso de la técnica presentada, las líneas en la imagen son detectadas. Nótese que, aunque el método propuesto funcionaría para el caso en el que se detectase una única línea, el error total RMS (suma de los cuadrados de las distancias de los puntos corregidos a las líneas para todas las líneas detectadas) se reduce mucho cuando se trata con más líneas. Por el momento, se asume que el centro de distorsión es el centro geométrico de la imagen.

Para obtener la imagen libre de distorsión  $I$ , es necesario calcular la posición correspondiente del píxel en la imagen distorsionada  $I'$  para cada píxel de  $I$ . Entonces, por medio de interpolación bicuadrática, se determina la intensidad del píxel y de esta manera se puede reconstruir adecuadamente  $I$ . Para llevar a cabo esta correspondencia entre imágenes, es necesario invertir el modelo de distorsión radial. Por tanto, buscamos una función radial  $G(\hat{r})$  que en el caso polinomial sería tal que:

$$\begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix} = G(\hat{r}) \begin{pmatrix} \hat{x} - x_c \\ \hat{y} - y_c \end{pmatrix}, \quad (3.16)$$

donde:

$$\hat{r} = \sqrt{(\hat{x} - x_c)^2 + (\hat{y} - y_c)^2}. \quad (3.17)$$

De la expresión anterior se obtiene que:

$$r = G(\hat{r})\hat{r}. \quad (3.18)$$

Por otro lado, tenemos que:

$$\begin{pmatrix} \hat{x} - x_c \\ \hat{y} - y_c \end{pmatrix} = L(r) \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix}, \quad (3.19)$$

y, por lo tanto:

$$\hat{r} = L(G(\hat{r})\hat{r})G(\hat{r})\hat{r}. \quad (3.20)$$

De esta manera, concluimos que  $G(\hat{r})$  es una raíz del polinomio:

$$P(\lambda) = 1 - L(\lambda\hat{r})\lambda = 1 - \sum_{j=0}^{N_k} k_j \hat{r}^j \lambda^{j+1}. \quad (3.21)$$

En el caso del modelo de tipo división, partiendo de la ecuación (3.2) obtenemos que:

$$\hat{r} + k_1 \cdot \hat{r} \cdot r^2 - r = 0, \quad (3.22)$$

el cual es un polinomio de segundo grado en  $r$ . Destacamos que sólo consideramos aquellos modelos de distorsión de la lente que satisfacen las condiciones (3.8) y (3.10), las cuáles definen una transformación biyectiva entre la imagen de entrada distorsionada y la corregida. En particular, para cualquier  $\hat{r} \in (0, \hat{r}_{max}]$  hay una única raíz  $r$  del polinomio anterior en el intervalo  $(0, r_{max}]$ . Dicha raíz debe ser la raíz positiva más pequeña de dicho polinomio, y viene dada por:

$$r = \frac{1 - \sqrt{1 - 4k_1(\hat{r})^2}}{2 \cdot k_1 \cdot \hat{r}}. \quad (3.23)$$

Nótese que si  $\hat{r} = 0$  entonces  $r = 0$ . En la práctica, en lugar de estimar dicha transformación inversa para cada píxel, primero se construye un vector donde guardamos el resultado de la expresión (3.23) en el intervalo  $[0, \hat{r}_{max}]$ . En el algoritmo 5 se describe el procedimiento completo para obtener la imagen corregida y el diagrama de flujo que sigue se encuentra en la figura 3.7.

## 3.5. Conclusiones

En el presente capítulo hemos visto una de las contribuciones de esta tesis: el uso de una transformada de Hough mejorada mediante la inclusión de la distorsión. Gracias a la incorporación de esta nueva dimensión, somos capaces de detectar de forma automática líneas que originalmente son rectas, pero que se encuentran distorsionadas por efecto de la distorsión de la lente. Hemos visto que, en caso de usar la transformada de Hough clásica, ésta no es capaz de afrontar la detección de líneas distorsionadas, ya que tiende a separarlas en distintos segmentos. Sin

---

**Algoritmo 5:** Cálculo de la imagen corregida.

---

```

input   : InputImage: imagen de entrada.
           : ldm: modelo de distorsión de la lente estimado.
output  : Imagen corregida
begin
  // Calculamos InverseVector en  $[0, \hat{r}_{max}]$  (como se ha explicado en el texto).
  for  $(\hat{x}, \hat{y}) \in \text{Píxeles de la imagen de salida}$  do
     $\hat{r} = \sqrt{(\hat{x} - x_c)^2 + (\hat{y} - y_c)^2}$ ;
     $index = \text{Floor}(\hat{r})$ ;
     $weight = \hat{r} - index$ ;
     $r = (1 - weight) \cdot \text{InverseVector}[index] + weight \cdot \text{InverseVector}[index + 1]$ ;
     $x = x_c + r \frac{\hat{x} - x_c}{\hat{r}}$ ;
     $y = y_c + r \frac{\hat{y} - y_c}{\hat{r}}$ ;
     $\text{OutputImage}(\hat{x}, \hat{y}) = \text{InputImage}(x, y)$ ;
    // En la expresión anterior usamos interpolación bilineal para estimar
     $\text{InputImage}(x, y)$ 
  end
end

```

---

embargo, usando la contribución expuesta, la transformada es capaz de detectar las líneas. La medida de fiabilidad propuesta en el nuevo espacio de Hough, se basa en el número de votos total obtenido. De manera que, para aquel valor del parámetro normalizado en el que se obtenga un mayor número de votos, se obtiene el conjunto de primitivas con mejor ajuste a las líneas distorsionadas en la imagen de entrada.

A partir de dicho conjunto de primitivas y del valor inicial del parámetro de distorsión, también hemos descrito una técnica para la optimización de dicho parámetro. Ésta se basa en la minimización de la distancia entre el conjunto de puntos que forman parte de las primitivas detectadas y sus líneas asociadas. De esta manera, se modifica el modelo de forma que proporcione el mejor ajuste posible a las líneas distorsionadas.

Por último, para poder construir una imagen libre de distorsión, se explica cómo se invierte el modelo de distorsión para poder realizar este proceso.

Sin embargo, la estimación automática de modelos de distorsión de un parámetro de tipo polinomial o división, puede encontrarse con ciertos problemas. El principal de ellos, es que la imagen que se proporciona como entrada posea un grado de distorsión tal que el modelo planteado no sea capaz de abordarla adecuadamente. Por tanto, parece lógica la introducción de un mayor número de parámetros, a fin

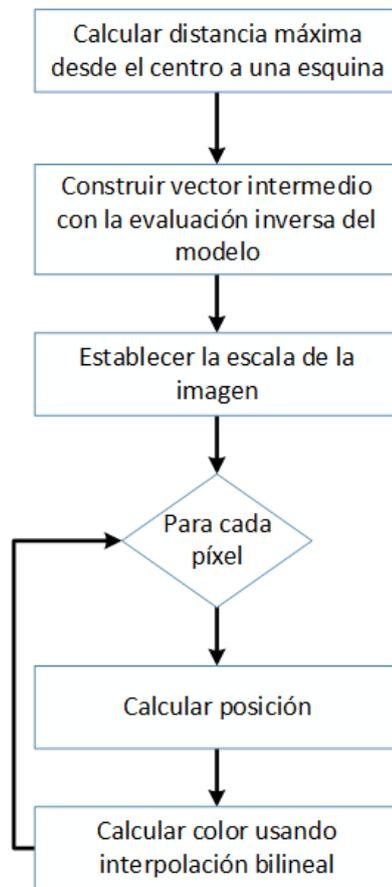


Figura 3.9: Diagrama de flujo de la corrección de la distorsión en la imagen.

de poder hacer frente a dicha problemática.

Por otro lado, como se ha descrito, la obtención de la imagen libre de distorsión, pasa por poder invertir el modelo de distorsión de la lente obtenido a través del proceso de estimación y optimización automático. Así pues, dotar al sistema de un mecanismo que permita verificar que el modelo se puede invertir, dotaría de robustez y fiabilidad al método propuesto.

A estas cuestiones trataremos de responder en los siguientes capítulos.

## Capítulo 4

# Estimación automática de modelos de distorsión de dos parámetros polinomiales y de división basada en la rectificación de líneas en la imagen

### 4.1. Introducción

En el capítulo anterior se describió la inclusión de un nuevo parámetro en el espacio de votación de Hough, la distorsión. Veíamos cómo la introducción de este parámetro permitía dos cosas. En primer lugar, obtener el conjunto de primitivas que mejor se ajustan a las líneas distorsionadas de la imagen de entrada. Por otro lado, se obtenía un primer valor para el parámetro de distorsión del modelo de la lente. También se explicó que el parámetro usado se normalizaba en términos de la corrección de la distorsión, lo cual aportaba dos ventajas: es más sencillo de interpretar (ya que representa un porcentaje de corrección) y da independencia respecto a la resolución de la imagen.

Una vez se ha obtenido el conjunto de primitivas con el mejor ajuste a las líneas distorsionadas y un primer valor para el parámetro de distorsión del modelo, éste era optimizado. Dicha optimización consistía en la minimización de la distancia

entre los puntos pertenecientes a las primitivas detectadas y sus líneas asociadas. De esta manera, se consigue un mejor ajuste de las primitivas detectadas respecto a los contornos de la imagen obtenidos en primera instancia. Finalmente, se corregía la distorsión presente en la imagen, mediante el cálculo de un vector intermedio que alberga el resultado de la evaluación de la inversa del modelo de distorsión obtenido. Así, finalmente se obtienen el color correspondiente a la nueva posición en la imagen resultante a través de una interpolación bilineal, y se puede construir la imagen corregida.

No obstante, cuando la distorsión por la que se ve afectada la imagen de entrada es demasiado fuerte, los modelos de distorsión de la lente de un parámetro no son capaces de abordar el problema, especialmente en el caso del modelo polinomial. Destacamos el caso del modelo polinomial puesto que en [27], se demuestra que los modelos de distorsión de tipo división o racionales, son capaces de hacer frente a distorsiones de un grado superior, haciendo uso de un menor número de parámetros en el modelo de distorsión de la lente. Sin embargo, también en el caso de este último de tipo de modelos, se observa una mejora cuando la distorsión con la que cuenta la imagen de entrada es fuerte.

En las siguientes secciones se va a presentar la estimación automática de modelos de distorsión de dos parámetros polinomiales y de división basada en la rectificación de líneas en la imagen. El primer paso usado en este capítulo se corresponde con el descrito en el capítulo anterior. Es decir, se obtiene una estimación inicial del valor del parámetro del modelo de distorsión de la lente y el conjunto de primitivas con el mejor ajuste. Ambos son producto de la propuesta del uso de una transformada de Hough mejorada mediante la inclusión de la distorsión como una nueva dimensión en el espacio de votación.

El presente capítulo se estructura en las siguientes secciones: en primer lugar se estudia la invertibilidad de los modelos de distorsión polinomiales y división de dos parámetros. Esto es, estudiar las condiciones bajo las cuales los parámetros de dichos modelos establecen una transformación biyectiva en el dominio de la imagen. Tras ello, y puesto que se incluye un nuevo parámetro a los modelos, se describe la normalización en términos de porcentaje de distancia al centro de distorsión para ambos tipos de modelo. Posteriormente, y de la misma forma que se hiciera en el capítulo 3, se optimiza el modelo de distorsión minimizando la distancia de las rectas a los bordes detectados. En esta ocasión, se incorpora además a dicha optimización el segundo parámetro y el centro de distorsión. Posteriormente, se describirá otra de las aportaciones de esta tesis doctoral, que consiste en la introducción de un proceso iterativo para mejorar el modelo de distorsión y el conjunto de primitivas asociadas. Una vez optimizado el modelo de distorsión

de la lente, incluyendo los dos parámetros y el centro de distorsión, se vuelven a buscar las líneas presentes en la imagen. A diferencia de la fase anterior, en este proceso de optimización iterativa se proporciona como entrada la imagen corregida con el modelo fruto de la optimización. Esto permite que se detecten líneas más largas, ya que el modelo optimizado hará que los puntos se ajusten más a las líneas rectas que se desean detectar. En la siguiente sección volverá hablarse de la corrección de la distorsión, aunque de forma breve, puesto que el procedimiento es similar al descrito en el capítulo previo. Tras hablar de la corrección, se mostrarán una serie de resultados experimentales, tanto sobre imágenes de patrones de calibración sintéticos, así como en escenas reales. Asimismo, también se describen resultados numéricos. Después se dedica un apartado al análisis de la complejidad computacional de la técnica propuesta, para finalmente describir las conclusiones de este capítulo.

## 4.2. Invertibilidad de los modelos de distorsión

Un modelo radial de distorsión de la lente de dos parámetros define una transformación en el dominio de la imagen. En esta sección, estudiamos la invertibilidad de dicha transformación de acuerdo a los valores de los parámetros de distorsión  $k_1, k_2$ . Ya que tratamos con transformaciones radiales, la invertibilidad de los modelos de distorsión es equivalente a la invertibilidad de la transformación radial asociada:

$$\hat{r}(r, k_1, k_2) = r \cdot L(r, k_1, k_2), \quad (4.1)$$

la cual determina la forma en la que se modifica la distancia  $r$  desde un punto al centro de distorsión cuando se aplica el modelo de distorsión. El dominio de definición de  $\hat{r}(r, k_1, k_2)$  con respecto a  $r$  es el intervalo  $[0, r_1]$ , donde  $r_1$  viene definido por:

$$r_1 = \max \{ \| (x, y) - (x_c, y_c) \| : (x, y) \in \text{Image Domain} \}. \quad (4.2)$$

En los siguientes teoremas presentamos algunos resultados relativos a la invertibilidad de  $\hat{r}(r, k_1, k_2)$  de acuerdo con los parámetros de distorsión  $k_1, k_2$  para los modelos polinomial y división.

**Teorema 1 (Invertibilidad de los modelos de distorsión de tipo polinomial)**

Para cualquier  $k_1, k_2 \in R$ , la función:

$$\hat{r}(r, k_1, k_2) = r(1 + k_1r^2 + k_2r^4) \quad (4.3)$$

es una función uno a uno con respecto a  $r$  en el intervalo  $[0, r_1]$  si y solo si  $k_1, k_2$  satisfacen:

$$\begin{cases} 9r_1^4k_1^2 - 20r_1^4k_2 < 0 & \text{si } r_1^2k_1 < -\frac{2}{3} \\ 5r_1^4k_2 + 3r_1^2k_1 + 1 > 0 & \text{si } r_1^2k_1 \geq -\frac{2}{3}. \end{cases} \quad (4.4)$$

Además, si  $k_1, k_2$  satisfacen (4.4), para cualquier  $r' \in (0, \hat{r}(r_1, k_1, k_2)]$ , existe un único  $r \in (0, r_1]$  tal que  $\hat{r}(r, k_1, k_2) = r'$ , dado por una raíz del polinomio:

$$P(r) = k_2r^5 + k_1r^3 + r - r' = 0. \quad (4.5)$$

**Demostración:** La derivada de  $\hat{r}(r, k_1, k_2)$  con respecto a  $r$  viene dada por:

$$\frac{\partial \hat{r}(r, k_1, k_2)}{\partial r} = 1 + 3k_1r^2 + 5k_2r^4. \quad (4.6)$$

Observamos que la derivada anterior es positiva en  $r = 0$ . Así, la función  $\hat{r}(r, k_1, k_2)$  será uno a uno en el intervalo  $[0, r_1]$  si y solo si  $\hat{r}(r, k_1, k_2)$  es una función creciente con respecto a  $r$  en el intervalo  $[0, r_1]$ . Esto es equivalente al hecho de que la derivada anterior es positiva en el intervalo  $[0, r_1]$  y, por tanto, sus raíces deben estar fuera del intervalo  $[0, r_1]$ . Formalmente, el potencial de raíces reales positivas del polinomio en (4.6) vienen dadas por:

$$\begin{cases} \sqrt{\frac{-3k_1 \pm \sqrt{9k_1^2 - 20k_2}}{10k_2}} & \text{si } k_2 \neq 0, 9k_1^2 - 20k_2 \geq 0 \text{ y } \frac{-3k_1 \pm \sqrt{9k_1^2 - 20k_2}}{10k_2} \geq 0 \\ \sqrt{-\frac{1}{3k_1}} & \text{si } k_2 = 0 \text{ y } k_1 < 0. \end{cases} \quad (4.7)$$

A continuación, consideramos el caso en que  $r_1 = 1$  y estudiamos las regiones en el espacio de los parámetros  $k_1, k_2$  donde todas las raíces reales de (4.6) están fuera del intervalo  $[0, 1]$ . Para encontrar tales regiones consideraremos las siguientes

curvas de nivel en el espacio  $k_1, k_2$ :

$$\begin{cases} k_2 = 0 \\ 9k_1^2 - 20k_2 = 0 \\ \frac{-3k_1 \pm \sqrt{9k_1^2 - 20k_2}}{10k_2} = 1. \end{cases} \quad (4.8)$$

Observamos que la última ecuación es equivalente a:

$$9k_1^2 - 20k_2 = (10k_2 + 3k_1)^2,$$

y un sencillo cálculo conduce a que esta ecuación es equivalente a:

$$5k_2 + 3k_1 + 1 = 0.$$

En la figura 4.1, ilustramos esas curvas de nivel. El área rosada representa las regiones donde todas las raíces reales del polinomio (4.6) están fuera del intervalo  $[0, 1]$  y, por lo tanto, la función  $r \rightarrow \hat{r}(r, k_1, k_2)$  es uno a uno en el intervalo  $[0, 1]$ . Observamos que, en caso de  $r_1 = 1$ , este área corresponde a aquella dada por la expresión (4.4), lo cual prueba la declaración del teorema para el caso en el que  $r_1 = 1$ . Para el caso general  $r_1 \neq 1$ , primeramente observamos que la función  $r \rightarrow \hat{r}(r, k_1, k_2)$  es uno a uno en el intervalo  $[0, r_1]$  si y solo si la función  $\tilde{r} \rightarrow \hat{r}(r_1\tilde{r}, k_1, k_2)$  es uno a uno en el intervalo  $[0, 1]$ . Por otro lado, tomando en consideración (4.3), obtenemos:

$$\hat{r}(r_1\tilde{r}, k_1, k_2) = r_1 \cdot \hat{r}(\tilde{r}, k_1r_1^2, k_2r_1^4),$$

y, por tanto, aplicando el resultado obtenido para el caso en que  $r_1 = 1$  a la función  $\tilde{r} \rightarrow \hat{r}(\tilde{r}, k_1r_1^2, k_2r_1^4)$ , obtenemos (4.4). Además, si la función  $r \rightarrow \hat{r}(r, k_1, k_2)$  es uno a uno, usando (4.3) obtenemos que la función inversa pueda ser calculada usando las raíces del polinomio  $P(r)$  dadas por (4.5), lo que concluye la prueba del teorema.

**Teorema 2 (Invertibilidad de los modelos de distorsión de tipo división)**  
*Para cualquier  $k_1, k_2 \in R$ , la función:*

$$\hat{r}(r, k_1, k_2) = \frac{r}{1 + k_1r^2 + k_2r^4} \quad (4.9)$$

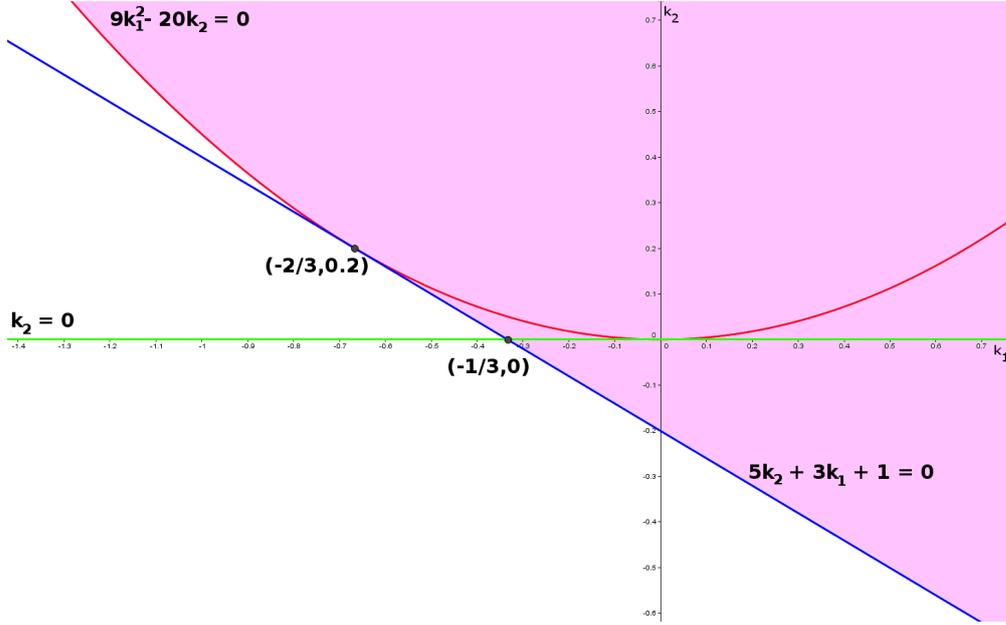


Figura 4.1: Modelo polinomial: curvas de nivel usadas para estudiar las regiones donde todas las raíces de la derivada de la función (4.6) están fuera del intervalo  $[0, 1]$ . El área rosa representa la región donde todas las raíces están fuera de  $[0, 1]$ .

es una función uno a uno con respecto a  $r$  en el intervalo  $[0, r_1]$  si y solo si  $k_1 r_1^2 > -2$  y:

$$\begin{cases} -1 - r_1^2 k_1 < r_1^4 k_2 < \frac{1 - r_1^2 k_1}{3} & \text{si } -2 < r_1^2 k_1 < 2 \\ -1 - r_1^2 k_1 < r_1^4 k_2 < -\frac{r_1^4 k_1^2}{12} & \text{si } r_1^2 k_1 \geq 2. \end{cases} \quad (4.10)$$

Además, si  $k_1, k_2$  satisfacen (4.10), para cualquier  $r' \in (0, \hat{r}(r_1, k_1, k_2)]$ , existe un único  $r \in (0, r_1]$  tal que  $\hat{r}(r, k_1, k_2) = r'$  y viene dado por una raíz del polinomio:

$$P(r) = r' k_2 r^4 + r' k_1 r^2 - r + r' = 0. \quad (4.11)$$

**Demostración:** Seguimos la misma técnica usada en el teorema 2 adaptada al modelo división. La derivada de  $\hat{r}(r, k_1, k_2)$  con respecto a  $r$  viene dada por:

$$\frac{\partial \hat{r}(r, k_1, k_2)}{\partial r} = \frac{1 - k_1 r^2 - 3k_2 r^4}{(1 + k_1 r^2 + k_2 r^4)^2}. \quad (4.12)$$

En este caso, tenemos que considerar las raíces del numerador y el denominador de la expresión anterior y obtenemos:

$$\left\{ \begin{array}{ll} \sqrt{\frac{-k_1 \pm \sqrt{k_1^2 - 4k_2}}{2k_2}} & \text{si } k_2 \neq 0, k_1^2 - 4k_2 \geq 0 \text{ y } \frac{-k_1 \pm \sqrt{k_1^2 - 4k_2}}{2k_2} \geq 0 \\ \sqrt{\frac{-k_1 \pm \sqrt{k_1^2 + 12k_2}}{6k_2}} & \text{si } k_2 \neq 0, k_1^2 + 12k_2 \geq 0 \text{ y } \frac{-k_1 \pm \sqrt{k_1^2 + 12k_2}}{6k_2} \geq 0 \\ \sqrt{\pm \frac{1}{k_1}} & \text{si } k_2 = 0 \text{ y } k_1 \neq 0. \end{array} \right. \quad (4.13)$$

A continuación, inicialmente asumimos que  $r_1 = 1$  y consideramos las siguientes curvas de nivel en el espacio  $k_1, k_2$  separando las diferentes regiones con respecto a las raíces:

$$\left\{ \begin{array}{l} k_2 = 0 \\ k_1^2 - 4k_2 = 0 \\ \frac{-k_1 \pm \sqrt{k_1^2 - 4k_2}}{2k_2} = 1 \rightarrow k_1 + k_2 + 1 = 0 \\ k_1^2 + 12k_2 = 0 \\ \frac{-k_1 \pm \sqrt{k_1^2 + 12k_2}}{6k_2} = 1 \rightarrow k_1 + 3k_2 - 1 = 0. \end{array} \right. \quad (4.14)$$

En la figura 4.2, mostramos dichas curvas de nivel. El área rosada representa las regiones donde todas las raíces reales de los polinomios en (4.12) están fuera del intervalo  $[0, 1]$  y, por tanto, la función  $r \rightarrow \hat{r}(r, k_1, k_2)$  es uno a uno en el intervalo  $[0, 1]$ . Para el caso general  $r_1 \neq 1$ , como en el teorema previo, probamos (4.10) mediante la reducción del problema al caso en el que  $r_1 = 1$  usando un cambio de variable. Usando (4.9), obtenemos que la función inversa puede calcularse usando las raíces del polinomio  $P(r)$  dado por (4.5), lo cual concluye la prueba del teorema.

Los resultados que se muestran en estos dos teoremas nos permiten comprobar fácilmente si para una configuración de parámetros  $k_1, k_2$  se define una transformación del modelo de distorsión uno a uno y si la transformación inversa está bien definida. En otro caso, la transformación de la distorsión de la lente genera singularidades y no se puede calcular apropiadamente la imagen corregida, porque se requiere la inversa de la transformación de la distorsión de la lente para hacerlo.

En el algoritmo 6, basado en los teoremas 1 y 2, se muestra el proceso seguido

---

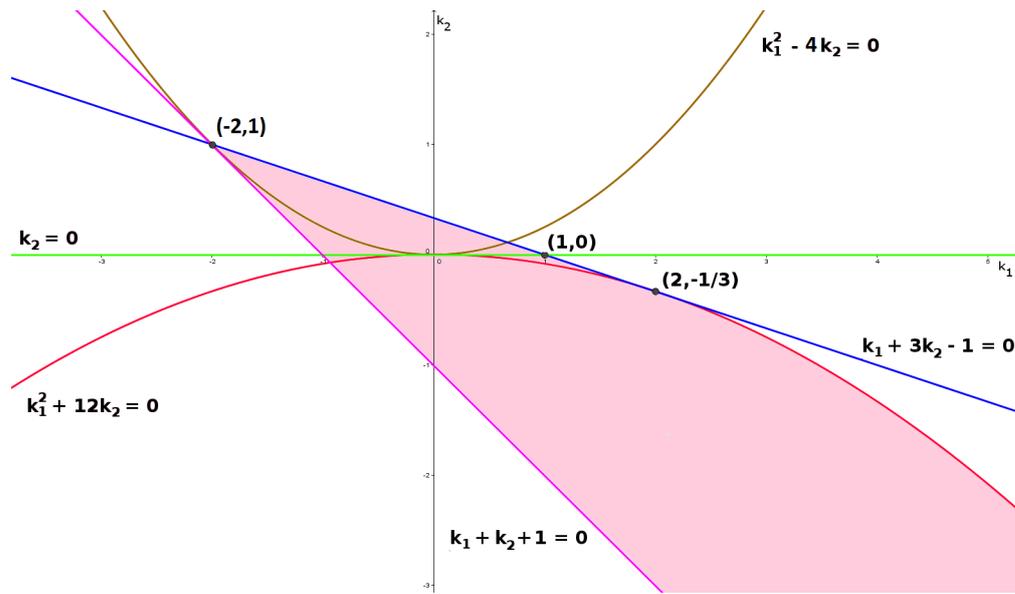


Figura 4.2: Modelo división: curvas de nivel usadas para estudiar las regiones donde todas las raíces de los polinomios en (4.12) están fuera del intervalo  $[0, 1]$ . El área rosa representa las regiones donde todas las raíces están fuera de  $[0, 1]$ .

---

**Algoritmo 6:** Comprobación de la invertibilidad del modelo de distorsión de la lente.

---

```

input   : ldm: modelo de distorsión de la lente a comprobar.
output : Verdadero si el modelo de distorsión es invertible y falso en caso contrario.

// Obtenemos la distancia desde el centro de distorsión a la esquina más alejada de la imagen
r ← get_radius(get_center(ldm), width, height);
hr ←  $\frac{r}{2}$ ;

if get_type(ldm) = POLYNOMIAL then
    k1 ←  $\frac{p_1 - 16p_2}{-12hr^2}$ ;
    k2 ←  $\frac{4p_2 - p_1}{-12hr^4}$ ;
    if (( $r^2k_1 < \frac{-2}{3}$ ) AND ( $9r^4k_1^2 - 20r^4k_2 < 0$ )) OR
        (( $r^2k_1 \geq \frac{-2}{3}$ ) AND ( $5r^4k_2 + 3r^2k_1 + 1 > 0$ )) then return true ;
else
    k1 ←  $\frac{-p_1 + 16p_2}{1+p_1 + 1+p_2}$ ;
    k2 ←  $\frac{-4p_2 + 1+p_1}{1+p_2 + 1+p_1}$ ;
    if ( $-2 < r^2k_1 < 2$ ) AND ( $-1 - r^2k_1 < r^4k_2 < \frac{1-r^2k_1}{3}$ ) then return true ;
    if ( $r^2k_1 \geq 2$ ) AND ( $-1 - r^2k_1 < r^4k_2 < -\frac{r^4k_1^2}{12}$ ) then return true ;
end
return false;

```

---

para comprobar la invertibilidad de ambos tipos de modelo de distorsión de la lente.

### 4.3. Normalización de parámetros

La ecuación general de un modelo de distorsión de la lente viene dada por la expresión (3.19) donde  $(x_c, y_c)$  representa el centro de distorsión,  $(x, y)$  es un punto en el dominio de la imagen,  $(\hat{x}, \hat{y})$  es el punto transformado,  $r = \| (x, y) - (x_c, y_c) \|$ , y  $L(r)$  representa la forma del modelo de distorsión. Como se comentó en el capítulo anterior, existen dos tipos de modelos de distorsión de la lente que son los más usados frecuentemente en visión por computador, debido a que son una buena solución de compromiso entre precisión y fácil cálculo. Dichos modelos son el modelo polinomial y el modelo división. El modelo polinomial (o el modelo de distorsión radial simple [39]), usando dos parámetros, se formula como:

$$L(r) = 1 + k_1 r^2 + k_2 r^4. \quad (4.15)$$

Los modelos de dos parámetros han sido usados en la literatura debido a su simplicidad y precisión [5, 46]. En [5], se propone un método algebraico, el cual es adecuado para corregir una distorsión radial significativa, además de ser computacionalmente eficiente. En [6] se puede encontrar una implementación en línea de dicho método algebraico.

El modelo división fue inicialmente propuesto en [50], pero ha recibido cierta atención de manera reciente tras la investigación llevada a cabo por Fitzgibbon [27]. En el caso de dos parámetros, se formula como:

$$L(r) = \frac{1}{1 + k_1 r^2 + k_2 r^4}. \quad (4.16)$$

La principal ventaja del modelo división es que requiere menos términos que el polinomial para abordar imágenes que tienen una distorsión severa. Por tanto, el modelo división parece ser más adecuado para lentes gran angular (puede consultarse una revisión reciente acerca de modelos de distorsión para lentes gran angular en [44]).

Con el objetivo de normalizar los parámetros  $k_1$  y  $k_2$  de los modelos de distorsión, de manera que el método sea independiente de la resolución de la imagen, y evitar trabajar con valores muy pequeños, introducimos en su lugar los parámetros  $p_1$  y  $p_2$ . El valor de  $p_1$  representa el porcentaje de corrección del punto más alejado al centro de distorsión en la imagen, mientras que  $p_2$  representa el mismo porcentaje de corrección, pero para el punto medio entre el centro de distorsión y el punto más alejado. De esta manera, los parámetros son más sencillos de interpretar y no dependen de la resolución de la imagen. En adelante, denotaremos como  $r_1$  la distancia desde el centro de distorsión al punto más alejado en el dominio de la imagen dada por la ecuación (4.2), así como  $r_2$  para la mitad de esta distancia.

### 4.3.1. Modelo polinomial

De acuerdo a las ecuaciones del modelo de distorsión polinomial, los valores descritos anteriormente se relacionan de la siguiente manera:

$$\begin{aligned}(1 + p_1)r_1 &= r_1(1 + k_1r_1^2 + k_2r_1^4), \\ (1 + p_2)r_2 &= r_2(1 + k_1r_2^2 + k_2r_2^4),\end{aligned}\tag{4.17}$$

y simplificando:

$$\begin{aligned}p_1 &= k_1r_1^2 + k_2r_1^4, \\ p_2 &= k_1r_2^2 + k_2r_2^4.\end{aligned}\tag{4.18}$$

Usando la relación  $r_1 = 2r_2$ , podemos reescribir las ecuaciones como:

$$\begin{aligned}p_1 &= k_14r_2^2 + k_216r_2^4, \\ p_2 &= k_1r_2^2 + k_2r_2^4.\end{aligned}\tag{4.19}$$

Conforme a esto, tenemos el sistema:

$$\begin{bmatrix} 4r_2^2 & 16r_2^4 \\ r_2^2 & r_2^4 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}, \quad (4.20)$$

el cual proporciona los siguientes valores para  $k_1$  y  $k_2$ :

$$k_1 = \frac{p_1 - 16p_2}{-12r_2^2}, \quad k_2 = \frac{4p_2 - p_1}{-12r_2^4}. \quad (4.21)$$

Por lo tanto, usando las expresiones de anteriores, podemos reemplazar  $k_1, k_2$  por  $p_1, p_2$  como parámetros de distorsión.

### 4.3.2. Modelo división

En el caso del modelo de distorsión de dos parámetros de tipo división, la relación entre los parámetros cambia de acuerdo con el modelo división:

$$\begin{aligned} (1 + p_1)r_1 &= \frac{r_1}{1 + k_1r_1^2 + k_2r_1^4}, \\ (1 + p_2)r_2 &= \frac{r_2}{1 + k_1r_2^2 + k_2r_2^4}, \end{aligned} \quad (4.22)$$

y por tanto:

$$\begin{aligned} p_1 &= \frac{1}{1 + k_1r_1^2 + k_2r_1^4} - 1, \\ p_2 &= \frac{1}{1 + k_1r_2^2 + k_2r_2^4} - 1. \end{aligned} \quad (4.23)$$

Para calcular  $k_1$  y  $k_2$ , tenemos un sistema similar al anterior:

$$\begin{bmatrix} r_1^2 & r_1^4 \\ r_2^2 & r_2^4 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \frac{-p_1}{1+p_1} \\ \frac{-p_2}{1+p_2} \end{bmatrix}. \quad (4.24)$$

Resolviendo este sistema, y considerando la relación  $r_1 = 2r_2$ , obtenemos:

$$k_1 = \frac{\frac{-p_1}{1+p_1} + \frac{16p_2}{1+p_2}}{-12r_2^2}, \quad k_2 = \frac{\frac{-4p_2}{1+p_2} + \frac{p_1}{1+p_1}}{-12r_2^4}. \quad (4.25)$$

#### 4.4. Estimación de un modelo de distorsión de dos parámetros para una colección de líneas dada

Denotamos por  $\mathbf{u}$  a la tupla  $(p_1, p_2, x_c, y_c)$ , la cual define el modelo de distorsión, y por  $\hat{x}_{ji}^{\mathbf{u}} = (\hat{x}_{ji}^{\mathbf{u}}, \hat{y}_{ji}^{\mathbf{u}})$  los puntos corregidos que se obtienen usando la ecuación (3.19) con el modelo dado por  $\mathbf{u}$ . Para un  $\mathbf{u}$  dado y una línea  $j$  dada, la ecuación de la línea recta asociada dada por  $\cos(\alpha_j^{\mathbf{u}})x + \sin(\alpha_j^{\mathbf{u}})y + d_j^{\mathbf{u}} = 0$  es definida a través del par  $(\alpha_j^{\mathbf{u}}, d_j^{\mathbf{u}})$  obtenido minimizando:

$$D(\alpha, d) = \sum_{i=1}^{N(j)} (\cos(\alpha)\hat{x}_{ji}^{\mathbf{u}} + \sin(\alpha)\hat{y}_{ji}^{\mathbf{u}} + d)^2. \quad (4.26)$$

La optimización del modelo de distorsión de la lente se lleva a cabo mediante la minimización del promedio de la distancia al cuadrado desde los puntos corregidos de las primitivas a sus líneas asociadas. Así, usando la notación introducida previamente, esta optimización consiste en minimizar la siguiente energía:

$$E(\mathbf{u}) = \frac{\sum_{j=1}^{Nl} \sum_{i=1}^{N(j)} (\cos(\alpha_j^{\mathbf{u}})\hat{x}_{ji}^{\mathbf{u}} + \sin(\alpha_j^{\mathbf{u}})\hat{y}_{ji}^{\mathbf{u}} + d_j^{\mathbf{u}})^2}{\sum_{j=1}^{Nl} N(j)}. \quad (4.27)$$

El valor inicial para  $p_1$  viene dado por el valor final obtenido de aplicar la transformada de Hough mejorada, mientras que el valor inicial para  $p_2$  es 0. El centro de distorsión es inicializado con el centro geométrico de la imagen. Para minimizar la energía (4.27), proponemos un esquema iterativo que es una combinación de los métodos del gradiente descendente y de Newton–Raphson. El método del

gradiente descendente está basado en el siguiente esquema iterativo:

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \lambda \nabla E(\mathbf{u}_n), \quad (4.28)$$

donde  $\lambda > 0$ . Es bien sabido que si  $\nabla E(\mathbf{u}_n) \neq 0$  y  $\lambda$  es lo suficientemente pequeño, entonces  $E(\mathbf{u}_n) < E(\mathbf{u}_{n+1})$ . El método de Newton–Raphson está basado en la expansión de Taylor de  $E(\mathbf{u})$  que viene dada por:

$$E(\mathbf{u}) = E(\mathbf{u}_n) + (\nabla E(\mathbf{u}_n))^T (\mathbf{u} - \mathbf{u}_n) + (\mathbf{u} - \mathbf{u}_n)^T \nabla^2 E(\mathbf{u}_n) (\mathbf{u} - \mathbf{u}_n) + \dots \quad (4.29)$$

Como queremos minimizar esta energía, derivamos la expresión previa y la igualamos a 0:

$$\nabla E(\mathbf{u}) \approx \nabla E(\mathbf{u}_n) + \nabla^2 E(\mathbf{u}_n) (\mathbf{u} - \mathbf{u}_n) = 0. \quad (4.30)$$

Por tanto,

$$\nabla^2 E(\mathbf{u}_n) (\mathbf{u} - \mathbf{u}_n) = -\nabla E(\mathbf{u}_n), \quad (4.31)$$

y el esquema de minimización de Newton-Raphson viene dado por:

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \nabla^2 E(\mathbf{u}_n)^{-1} \nabla E(\mathbf{u}_n), \quad (4.32)$$

donde:

$$\nabla E(\mathbf{u}_n) = \begin{pmatrix} \frac{\partial E(\mathbf{u}_n)}{\partial p_1} \\ \frac{\partial E(\mathbf{u}_n)}{\partial p_2} \\ \frac{\partial E(\mathbf{u}_n)}{\partial x_c} \\ \frac{\partial E(\mathbf{u}_n)}{\partial y_c} \end{pmatrix}, \quad (4.33)$$

$$\nabla^2 E(\mathbf{u}_n) = \begin{pmatrix} \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_1^2} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_1 \partial p_2} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_1 \partial x_c} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_1 \partial y_c} \\ \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_2 \partial p_1} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_2^2} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_2 \partial x_c} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial p_2 \partial y_c} \\ \frac{\partial^2 E(\mathbf{u}_n)}{\partial x_c \partial p_1} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial x_c \partial p_2} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial x_c^2} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial x_c \partial y_c} \\ \frac{\partial^2 E(\mathbf{u}_n)}{\partial y_c \partial p_1} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial y_c \partial p_2} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial y_c \partial x_c} & \frac{\partial^2 E(\mathbf{u}_n)}{\partial y_c^2} \end{pmatrix}. \quad (4.34)$$

Normalmente, el método de Newton–Raphson es más rápido que el método del gradiente descendente cuando la solución inicial no está lejos de un mínimo local de la energía, pero en el caso de que la solución inicial esté lejos de un mínimo local de la energía, el método puede divergir y, además, no hay forma de asegurar que la energía decrece a lo largo de las iteraciones. Para anular esta limitación del esquema de Newton–Raphson, en 1944 Levenberg [51] introdujo la idea de usar un parámetro de convergencia para combinar los esquemas del gradiente descendente y el de Newton–Raphson en el contexto del conocido algoritmo de minimización de Levenberg–Marquardt. Usando la aproximación de Levenberg, proponemos usar el siguiente esquema de minimización:

$$\mathbf{u}_{n+1} = \mathbf{u}_n - (\nabla^2 E(\mathbf{u}_n) + \gamma Id)^{-1} \nabla E(\mathbf{u}_n), \quad (4.35)$$

donde  $\gamma$  es el parámetro de convergencia usado para controlar la convergencia de la minimización de la siguiente manera:  $\gamma$  se actualiza en cada iteración para asegurar que  $E(\mathbf{u}_{n+1}) < E(\mathbf{u}_n)$ . Normalmente, su valor es mayor cuando estamos lejos de la solución y decrece cuando nos aproximamos a ella. Por tanto, el parámetro de convergencia ayuda a anular las inestabilidades de Newton–Raphson cuando la solución inicial está lejos de un mínimo local de la energía.

El parámetro de convergencia  $\gamma$  es inicializado a 10, y su evolución depende del valor obtenido para la energía optimizada en cada iteración. Con el objetivo de asegurar que la nueva energía calculada es menor que la anterior, el valor del parámetro es multiplicado/dividido por 10, en función del crecimiento/decrecimiento de la energía. Puesto que el proceso de optimización consiste en optimizar la tupla constituida por  $(p_1, p_2, x_c, y_c)$ , la condición de parada es que la diferencia de todos los parámetros optimizados y sus valores previos sea inferior a una tolerancia. Para evitar un alto número de iteraciones, se incluye una condición extra, estableciendo un límite de 100 iteraciones. En el peor caso, si ninguna iteración minimizase la energía, el modelo sería igual al de entrada, y el centro de distorsión correspondería al centro geométrico de la imagen.

Destacamos que buscamos un mínimo local de la energía  $E(\mathbf{u})$  porque, en general, el mínimo global de esta energía puede ser alcanzado en configuraciones singulares (especialmente cuando se usan muy pocos puntos de las primitivas). Para evitar estas soluciones singulares, verificamos si la transformación de la distorsión de la lente es biyectiva usando los resultados de los teoremas 1 y 2, y en otro caso, rechazamos la solución propuesta. Señalamos que, puesto que usamos un esquema iterativo, no podemos asegurar que el método converja siempre hacia la mejor solución. Sin embargo, el método se comporta bastante bien en la mayoría de las situaciones. En el algoritmo 7 se muestra el proceso seguido para la realización de este proceso de minimización.

---

**Algoritmo 7:** Optimización de  $p_1, p_2, x_c, y_c$  minimizando  $E(p_1, p_2, x_c, y_c)$ .

---

```

input   : ldm: el modelo de la lente dado por los parámetros  $p_1, p_2, x_c, y_c$ .
           : i_primitives: la colección de líneas.
           : TOL: valor para la convergencia.
           :  $h_1$ : variación para  $p_1$  y  $p_2$ .
           :  $h_2$ : variación para las coordenadas del centro de distorsión.
output  : ldm: el modelo de distorsión de la lente optimizado.
// Inicializamos el nuevo modelo de distorsión para entrar en el bucle while.
 $p_{1_n} = p_1 + TOL(|p_1| + 2)$ ;
// Inicializamos  $\gamma$ , usado para controlar la convergencia.
 $\gamma \leftarrow 10,0$ ;
 $Id(4, 4) \leftarrow$  Mariz identidad;
while ( $|p_{1_n} - p_1| > TOL(|p_1| + 1)$  OR  $|p_{2_n} - p_2| > TOL(|p_2| + 1)$ 
OR  $|x_{c_n} - x_c| > TOL(|x_c| + 1)$  OR  $|y_{c_n} - y_c| > TOL(|y_c| + 1)$ ) do
    // Calculamos el vector gradiente y la matriz Hessiana
     $grad_E \leftarrow$  Compute_gradient( $p_1, p_2, x_c, y_c, h_1, h_2$ ) ;           // Algoritmo 8
     $hess_E \leftarrow$  Compute_hessian( $p_1, p_2, x_c, y_c, h_1, h_2$ ) ;         // Algoritmo 9
    // Resolvemos el sistema de ecuaciones usando el método de Gauss
     $v_{res}(4) \leftarrow$  Gauss( $grad_E, hess_E + \gamma \cdot Id$ );
     $p_{1_{new}} \leftarrow p_1 + v_{res}(0)$ ;
     $p_{2_{new}} \leftarrow p_2 + v_{res}(1)$ ;
     $x_{c_{new}} \leftarrow x_c + v_{res}(2)$ ;
     $y_{c_{new}} \leftarrow y_c + v_{res}(3)$ ;
    // Iteramos hasta que la energía es reducida y el modelo es invertible
     $ldm \leftarrow$  lens_distortion_model ( $p_{1_{new}}, p_{2_{new}}, x_{c_{new}}, y_{c_{new}}$ );
    if  $E(p_{1_{new}}, p_{2_{new}}, x_{c_{new}}, y_{c_{new}}) < E(p_1, p_2, x_c, y_c)$  AND check_invertibility(ldm)
    then
        |  $p_{1_{new}} \leftarrow p_1 + v_{res}(0)$ ;
        |  $p_{2_{new}} \leftarrow p_2 + v_{res}(1)$ ;
        |  $x_{c_{new}} \leftarrow x_c + v_{res}(2)$ ;
        |  $y_{c_{new}} \leftarrow y_c + v_{res}(3)$ ;
        |  $\gamma \leftarrow \frac{\gamma}{10}$ ;
    else
        |  $\gamma \leftarrow \gamma \cdot 10$ ;
    end
end
// Actualizamos el porcentaje y centro de distorsión
 $(p_{1_n}, p_{2_n}, x_{c_n}, y_{c_n}) \leftarrow (p_1, p_2, x_c, y_c)$ ;
 $(p_1, p_2, x_c, y_c) \leftarrow (p_{1_{new}}, p_{2_{new}}, x_{c_{new}}, y_{c_{new}})$ ;

```

---

**Algoritmo 8:** Cálculo del gradiente de E.

**input** :  $p_1, p_2$ : valor normalizado para los parámetros de distorsión  $k_1$  y  $k_2$ .  
 $x_c, y_c$ : coordenadas  $(x, y)$  del centro de distorsión.  
 $h_1$ : paso de discretización para  $p_1$  y  $p_2$ .  
 $h_2$ : paso de discretización para las coordenadas del centro de distorsión.

**output** :  $grad_E$ : El vector gradiente

$$\begin{aligned} grad_E(0) &\leftarrow -\frac{(E(p_1+h_1, p_2, x_c, y_c) - E(p_1, p_2, x_c, y_c))}{h_1}; \\ grad_E(1) &\leftarrow -\frac{(E(p_1, p_2+h_1, x_c, y_c) - E(p_1, p_2, x_c, y_c))}{h_1}; \\ grad_E(2) &\leftarrow -\frac{(E(p_1, p_2, x_c+h_2, y_c) - E(p_1, p_2, x_c, y_c))}{h_2}; \\ grad_E(3) &\leftarrow -\frac{(E(p_1, p_2, x_c, y_c+h_2) - E(p_1, p_2, x_c, y_c))}{h_2}; \end{aligned}$$

**Algoritmo 9:** Cálculo del Hessiano de E.

**input** :  $p_1, p_2$ : valor normalizado para los parámetros de distorsión  $k_1$  y  $k_2$ .  
 $x_c, y_c$ : coordenadas  $(x, y)$  del centro de distorsión.  
 $h_1$ : paso de discretización para  $p_1$  y  $p_2$ .  
 $h_2$ : paso de discretización para las coordenadas del centro de distorsión.

**output** :  $hess_E$ : La matriz Hessiana.

$$\begin{aligned} hess_E(0, 0) &\leftarrow \left( \frac{(E(p_1+h_1, p_2, x_c, y_c) + E(p_1-h_1, p_2, x_c, y_c) - 2E(p_1, p_2, x_c, y_c))}{h_1^2} \right); \\ hess_E(1, 1) &\leftarrow \left( \frac{(E(p_1, p_2+h_1, x_c, y_c) + E(p_1, p_2-h_1, x_c, y_c) - 2E(p_1, p_2, x_c, y_c))}{h_1^2} \right); \\ hess_E(2, 2) &\leftarrow \left( \frac{(E(p_1, p_2, x_c+h_2, y_c) + E(p_1, p_2, x_c-h_2, y_c) - 2E(p_1, p_2, x_c, y_c))}{h_2^2} \right); \\ hess_E(3, 3) &\leftarrow \left( \frac{(E(p_1, p_2, x_c, y_c+h_2) + E(p_1, p_2, x_c, y_c-h_2) - 2E(p_1, p_2, x_c, y_c))}{h_2^2} \right); \\ \\ hess_E(0, 1) &\leftarrow hess_E(1, 0) \leftarrow \\ &\frac{(E(p_1+h_1, p_2+h_1, x_c, y_c) - E(p_1+h_1, p_2, x_c, y_c)) - (E(p_1, p_2+h_1, x_c, y_c) - E(p_1, p_2, x_c, y_c))}{h_1^2}; \\ hess_E(0, 2) &\leftarrow hess_E(2, 0) \leftarrow \\ &\frac{(E(p_1+h_1, p_2, x_c+h_2, y_c) - E(p_1+h_1, p_2, x_c, y_c)) - (E(p_1, p_2, x_c+h_2, y_c) - E(p_1, p_2, x_c, y_c))}{h_1 h_2}; \\ hess_E(0, 3) &\leftarrow hess_E(3, 0) \leftarrow \\ &\frac{(E(p_1+h_1, p_2, x_c, y_c+h_2) - E(p_1+h_1, p_2, x_c, y_c)) - (E(p_1, p_2, x_c, y_c+h_2) - E(p_1, p_2, x_c, y_c))}{h_1 h_2}; \\ hess_E(1, 2) &\leftarrow hess_E(2, 1) \leftarrow \\ &\frac{(E(p_1, p_2+h_1, x_c+h_2, y_c) - E(p_1, p_2+h_1, x_c, y_c)) - (E(p_1, p_2, x_c+h_2, y_c) - E(p_1, p_2, x_c, y_c))}{h_1 h_2}; \\ hess_E(1, 3) &\leftarrow hess_E(3, 1) \leftarrow \\ &\frac{(E(p_1, p_2+h_1, x_c, y_c+h_2) - E(p_1, p_2+h_1, x_c, y_c)) - (E(p_1, p_2, x_c, y_c+h_2) - E(p_1, p_2, x_c, y_c))}{h_1 h_2}; \\ hess_E(2, 3) &\leftarrow hess_E(3, 2) \leftarrow \\ &\frac{(E(p_1, p_2, x_c+h_2, y_c+h_2) - E(p_1, p_2, x_c+h_2, y_c)) - (E(p_1, p_2, x_c, y_c+h_2) - E(p_1, p_2, x_c, y_c))}{h_2^2}; \end{aligned}$$

## 4.5. Optimización iterativa de modelos de distorsión de dos parámetros y de la colección de líneas

Anteriormente, hemos descrito como los parámetros de ambos tipos de modelos de distorsión pueden ser optimizados para minimizar la distancia entre los puntos corregidos de las primitivas y sus rectas asociadas. Durante este proceso, consideramos los mismos puntos que fueron originalmente detectados como puntos pertenecientes a las líneas. Por esta razón, sólo los parámetros del modelo se varían durante la minimización, con el objetivo de obtener un mejor ajuste entre los puntos y las ecuaciones de las líneas.

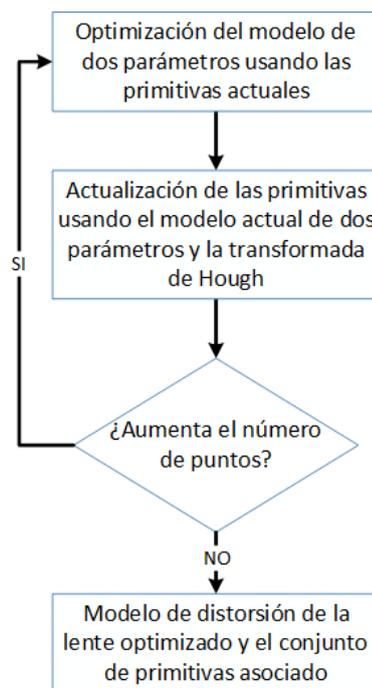


Figura 4.3: Diagrama de flujo del proceso de optimización iterativa del modelo de distorsión de la lente y del conjunto de primitivas asociadas.

Sin embargo, una vez que el modelo de dos parámetros y el centro de distorsión han sido optimizados usando dicha colección de líneas con sus puntos asociados, el nuevo modelo puede ser usado para mejorar las primitivas usando nuevamente la estrategia de votación de Hough. La idea subyacente consiste en que cuanto mayor sea la colección de puntos asociados a las primitivas, más largas serán las líneas y

mejor será la estimación del modelo de distorsión de la lente.

Por lo tanto, si integramos el modelo de distorsión optimizado dentro del proceso de Hough, los puntos borde volverán a votar por las líneas presentes en la imagen. La principal diferencia respecto a la aproximación inicial, es que en este caso los puntos han sido previamente corregidos usando el modelo optimizado. De esta manera, aquellos puntos que no fueron asociados inicialmente a las líneas, puesto que la correspondencia no era satisfactoria, pueden ser asociados ahora. Como resultado de este nuevo proceso de votación obtenemos una colección de primitivas mejorada, puesto que cuentan con un mayor número de puntos. Tras ello, con la colección de líneas actualizada, se puede optimizar de nuevo el modelo de dos parámetros usando este nuevo conjunto de primitivas. Este proceso de optimización se corresponde con el descrito en la sección previa, así que en cada iteración se minimiza la distancia desde los puntos corregidos a las nuevas líneas obtenidas. Este proceso iterativo finaliza si la cantidad global de puntos asociados a las líneas no aumenta al integrar el modelo optimizado en el proceso de votación de Hough. En la práctica, sólo hacen falta unas pocas iteraciones hasta alcanzar la convergencia.

En la figura 4.3 se muestra el diagrama de flujo del proceso de optimización iterativa del modelo de distorsión de la lente (algoritmo 10), así como del conjunto de primitivas asociadas.

## 4.6. Corrección de la distorsión de la imagen

En primer lugar, es necesario destacar que el modelo final de distorsión de la lente obtenido satisface las condiciones de los teoremas 1 y 2, porque verificamos dichas condiciones en las diferentes etapas del método, y cualquier modelo de distorsión de la lente que no satisfaga tales condiciones es rechazado. Por tanto, el modelo de distorsión obtenido es una transformación biyectiva y se puede calcular de forma fiable la transformación inversa (como se explicó en dichos teoremas). Usando dicha transformación inversa, obtenemos la imagen corregida mediante el cálculo de la distorsión inversa en los puntos de la imagen para obtener los canales de color RGB del punto correspondiente en la imagen original.

La corrección de la distorsión se lleva a cabo mediante el cálculo de la función inversa del modelo de distorsión de la lente obtenido tras los pasos anteriores. Para dicho fin, es necesario invertir la función radial:

---

**Algoritmo 10:** Optimización iterativa, descrita en el diagrama de flujo de la figura 4.3.

---

```

input   :  $np$ : cantidad previa de puntos.
            $ldm(p_1, p_2, x_c, y_c)$ : modelo de distorsión de la lente previo.
            $contours$ : bordes de la imagen de entrada.
            $i\_primitives$ : colección previa de líneas.
            $TOL$ : valor para la convergencia.
output :  $ldm$ : modelo de distorsión de la lente optimizado.
            $i\_primitives$ : colección de líneas optimizada.

current_model( $p_{1\_new}, p_{2\_new}, x_{c\_new}, y_{c\_new}$ )  $\leftarrow$   $ldm(p_1, p_2, x_c, y_c)$ ;
new_primitives  $\leftarrow$   $i\_primitives$ ;
new_points  $\leftarrow$   $np$ ; next_points  $\leftarrow$   $np$ ;
fails  $\leftarrow$  0;

while next_points  $\geq$  ( $new\_points \cdot (1 + TOL)$ ) OR fails  $<$  3 do
    new_points  $\leftarrow$  next_points;
    // Optimizamos current_model usando una minimización del error.
    Error_minimization( current_model, new_primitives, TOL );           // Algoritmo 7

    // Llamamos a Hough para actualizar new_primitives usando el modelo obtenido de la
    // minimización.
    next_points  $\leftarrow$  Hough( contours, new_primitives, current_model )

    // Actualizamos las primitivas y el modelo si Hough proporciona un mejor resultado
    if next_points  $>$   $np$  AND check_invertibility(current_model) then
        |  $i\_primitives \leftarrow$  new_primitives;
        |  $np \leftarrow$  next_points;
        |  $ldm \leftarrow$  current_model;
    else
        | fails  $\leftarrow$  fails+1;
    end
end

```

---

$$\hat{r}(r, k_1, k_2) = r \cdot L(r).$$

donde  $r$  es la distancia desde cualquier punto  $(x, y)$  del dominio de la imagen al centro de distorsión. Recordamos que, en este caso, dicho centro no coincide con el punto principal o con el centro geométrico del plano imagen, ya que su valor es dependiente del proceso de optimización que incluye ambos parámetros de distorsión y dichas coordenadas. Con el objetivo de acelerar el procedimiento, en lugar de calcular la inversa de  $\hat{r}(r, k_1, k_2)$  para cada punto  $(x, y)$  de forma independiente, en primer lugar calculamos la inversa en el conjunto discreto  $r = 1, 2, 3, \dots, \hat{r}(r_1, k_1, k_2)$  (donde  $r_1$  es la distancia máxima desde un punto de la imagen al centro de distorsión) y entonces, para cualquier punto  $(x, y)$ , calculamos la inversa de  $\hat{r}(r, k_1, k_2)$  usando un procedimiento de interpolación. El procedimiento llevado a cabo es similar al expuesto en el algoritmo 5, salvo que en esta ocasión, la construcción del vector con el resultado de la evaluación del modelo inverso, usaría los modelos de distorsión de dos parámetros descritos en este capítulo.

## 4.7. Resultados experimentales

Para comprobar el funcionamiento de la propuesta realizada en el presente capítulo, en esta sección se va a realizar una comparativa entre los resultados obtenidos con modelos de distorsión de un parámetro descritos en el capítulo previo, frente a los modelos polinomial y división con dos parámetros. En ese último caso, además, se incorpora el proceso de optimización iterativa para mejorar tanto el modelo de distorsión como el conjunto de primitivas asociadas.

En la figura 4.4, se muestran el conjunto de imágenes que se han usado para realizar los experimentos. Entre ellas se encuentran tanto imágenes con poca distorsión, como otras en las que el efecto es más fuerte. Además, se han usado imágenes consistentes en fotografías de patrones de calibración sintéticos e imágenes reales. De esta manera tratamos de reflejar cómo se comporta el algoritmo frente a diferentes situaciones.

Todo el conjunto de imágenes se fotografiaron usando una cámara Nikon D90. Por un lado, aquellas imágenes con una distorsión menos fuerte (figs. 4.4(a) y 4.4(e)) fueron captadas por una lente Nikkor DX con un *zoom* 17–200mm. Por otra parte, el resto de las imágenes (figs. 4.4(b), 4.4(c) y 4.4(d)) se capturaron usando una lente Tokina DX con un *zoom* 10–17mm. Las fotografías se hicieron usando

diferentes distancias focales (cambiando el *zoom* de la lente), lo que proporciona una cierta variedad de grados de distorsión.

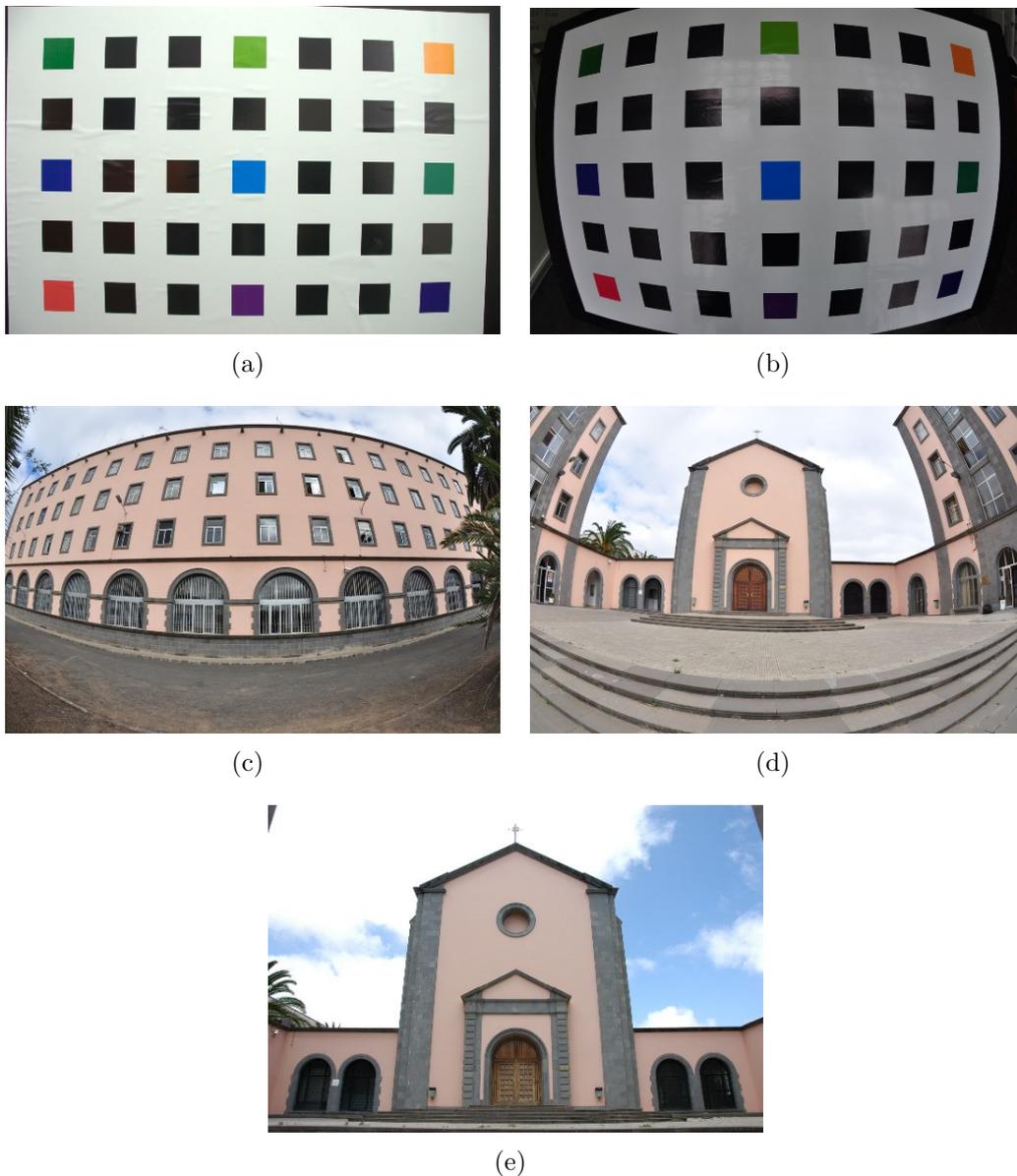


Figura 4.4: Imágenes originales utilizadas en los experimentos.

Las figuras 4.5 – 4.11, muestran los resultados de la detección de primitivas sobre imágenes con diferentes grados de distorsión. Todos aquellos puntos que tengan un mismo color, quiere decir que pertenecen a una misma línea recta. En el caso de la figura 4.5, solo se muestran el conjunto de primitivas resultantes

de aplicar el proceso con un único parámetro de distorsión, puesto que debido al bajo grado de distorsión, la mejora proporcionada por la introducción de la optimización iterativa no es apreciable. Las imágenes en las figuras 4.5(a), 4.6 y 4.7, muestran resultados de la detección de líneas en fotografías de patrones de calibración sintéticas. Por otro lado, las figuras 4.5(b), 4.8, 4.9, 4.10 y 4.11; contienen los resultados relativos a la misma detección pero sobre fotografías en escenarios reales.

Como puede verse, especialmente en aquellas imágenes afectadas por una fuerte distorsión, los resultados obtenidos con el método propuesto son mejores. La inclusión de un segundo parámetro de distorsión, la optimización del modelo y el centro de distorsión y además añadir el proceso de optimización iterativa, mejora el conjunto de primitivas detectadas. Esto puede verse en el hecho de que proporciona nuevas líneas, e incluso se obtienen líneas más largas, en las que el número de puntos es superior.

Por ejemplo, los cuatro cuadrados de las esquinas, que inicialmente no eran detectados usando el modelo polinomial (fig. 4.6(a)), son ahora detectados con el mismo modelo tras introducir la optimización iterativa (fig. 4.6(b)). En el caso del modelo división, las líneas son más largas, y por tanto se obtiene una información más fiable para la corrección de la distorsión (fig. 4.7(b)).

En los resultados que se muestran en las figuras 4.8, 4.9, 4.10 y 4.11 se observa algo similar. Como puede verse en las figuras 4.8 y 4.9, el modelo división es capaz de extraer líneas más largas, las cuáles se ven partidas cuando se hace uso del modelo polinomial. Esto puede observarse, por ejemplo, en las líneas horizontales de la parte superior. Además, en el caso de usar el proceso de optimización iterativa, permite detectar líneas paralelas muy cercanas. En las figuras 4.10 y 4.11, la introducción del método propuesto detecta más puntos pertenecientes a las líneas, y el uso del modelo división proporciona líneas más largas, que no se encuentran divididas en distintos segmentos. Un claro ejemplo de ello son las líneas que se encuentran en las escaleras de la parte inferior de la imagen.

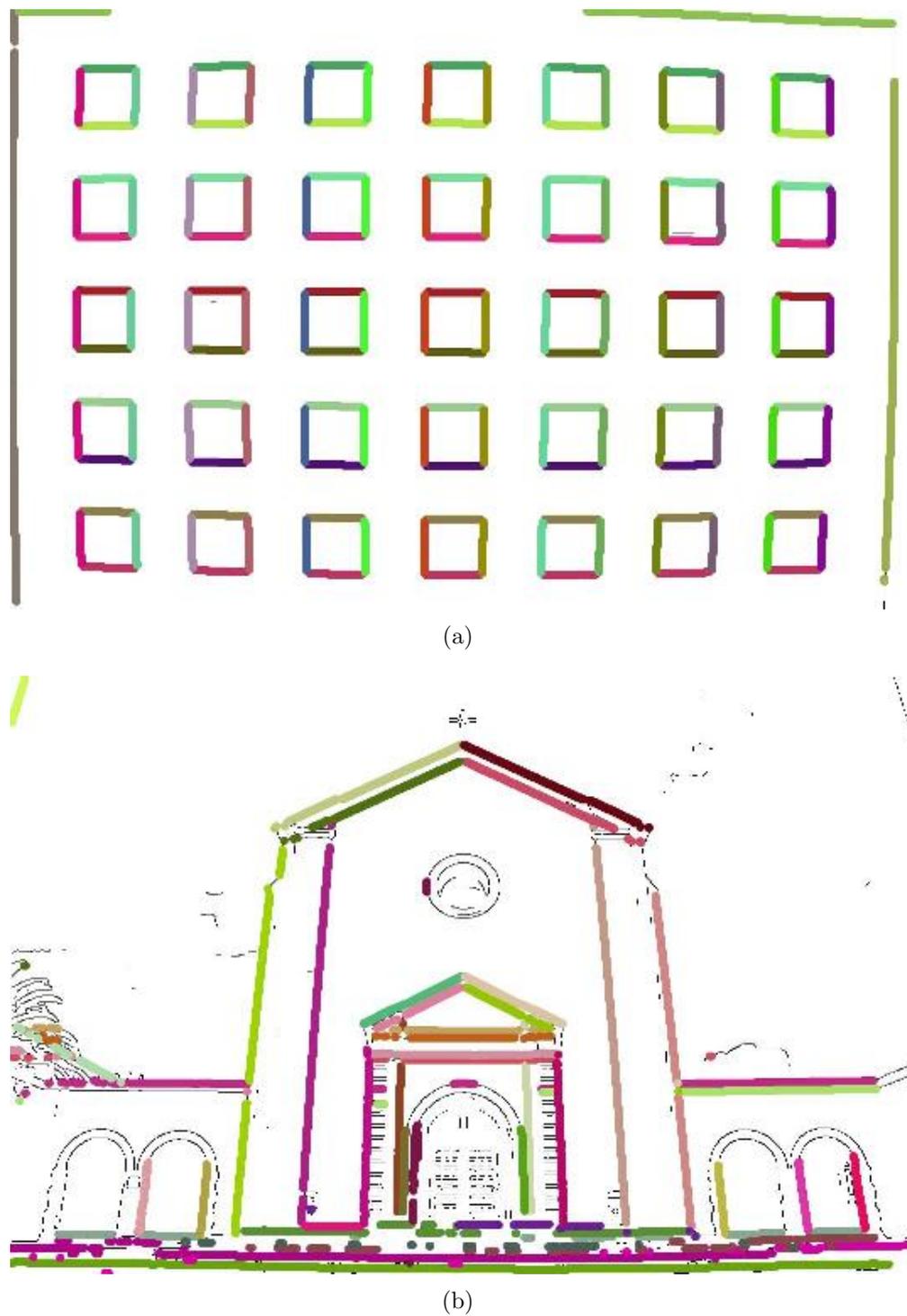


Figura 4.5: Colección inicial de líneas detectadas con la transformada mejorada de Hough usando el modelo polinomial para (a) un patrón de calibración y (b) un edificio.

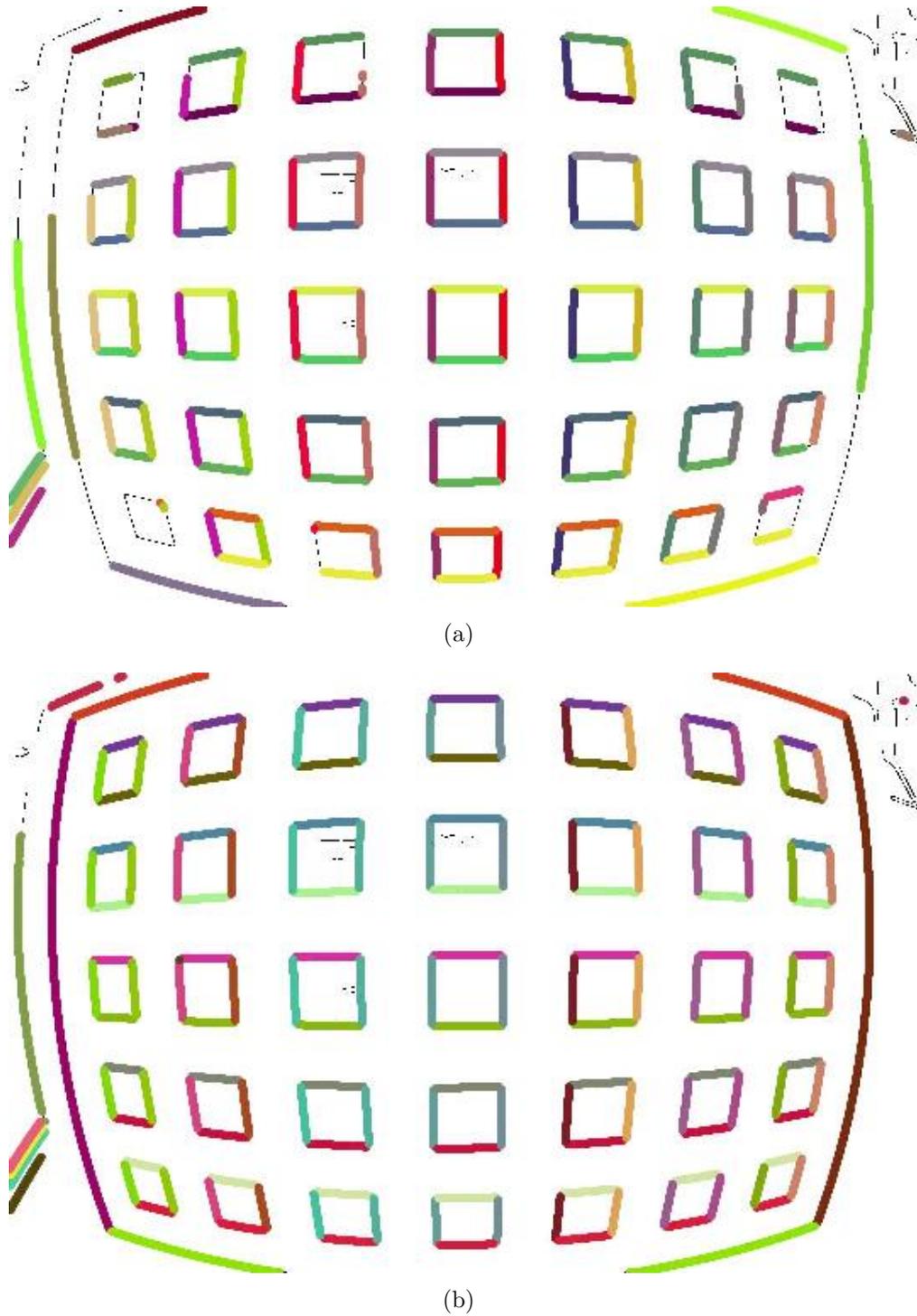


Figura 4.6: Colección de líneas detectadas con la transformada de Hough usando (a) el modelo polinomial y (b) el modelo polinomial con optimización iterativa.

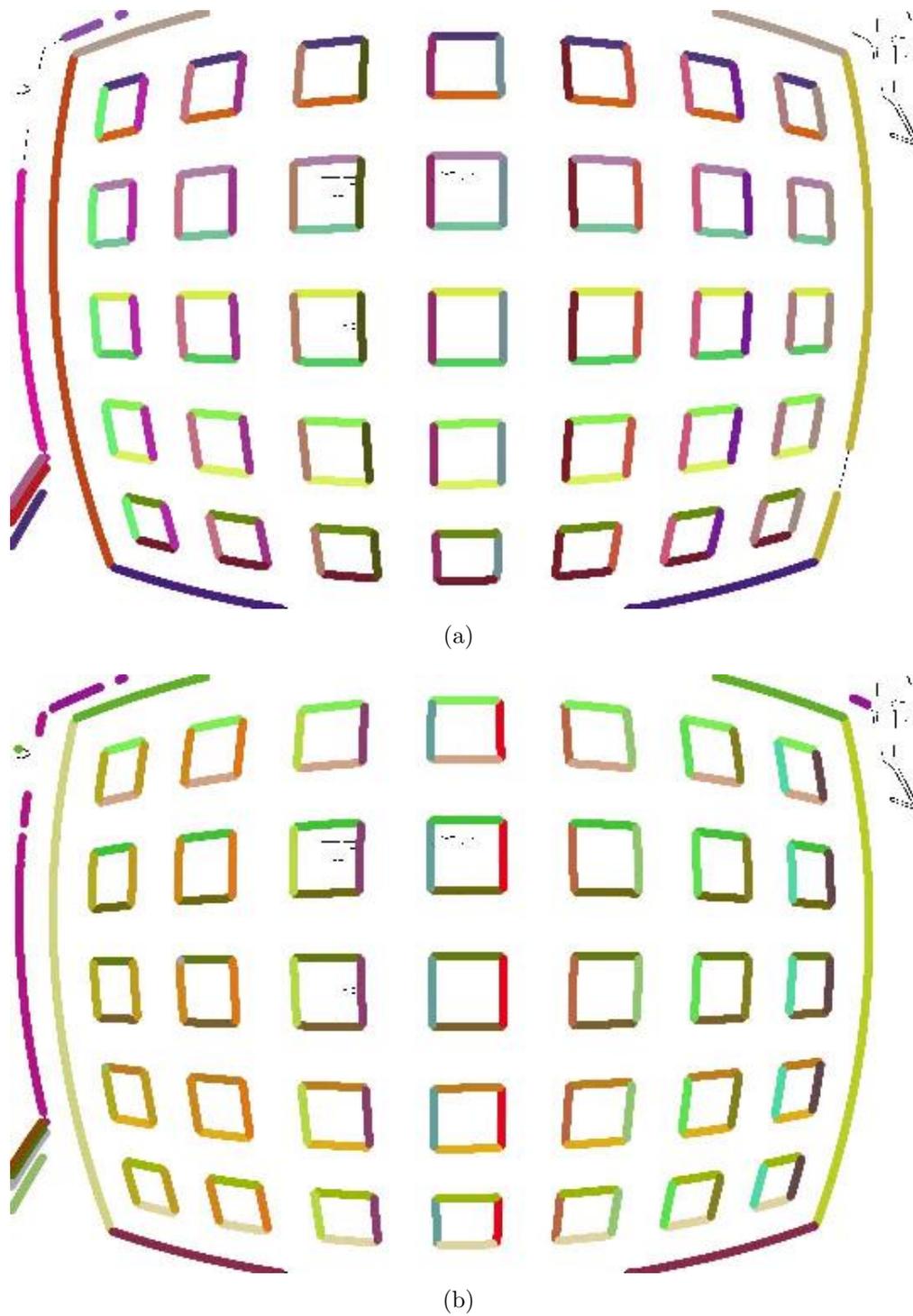
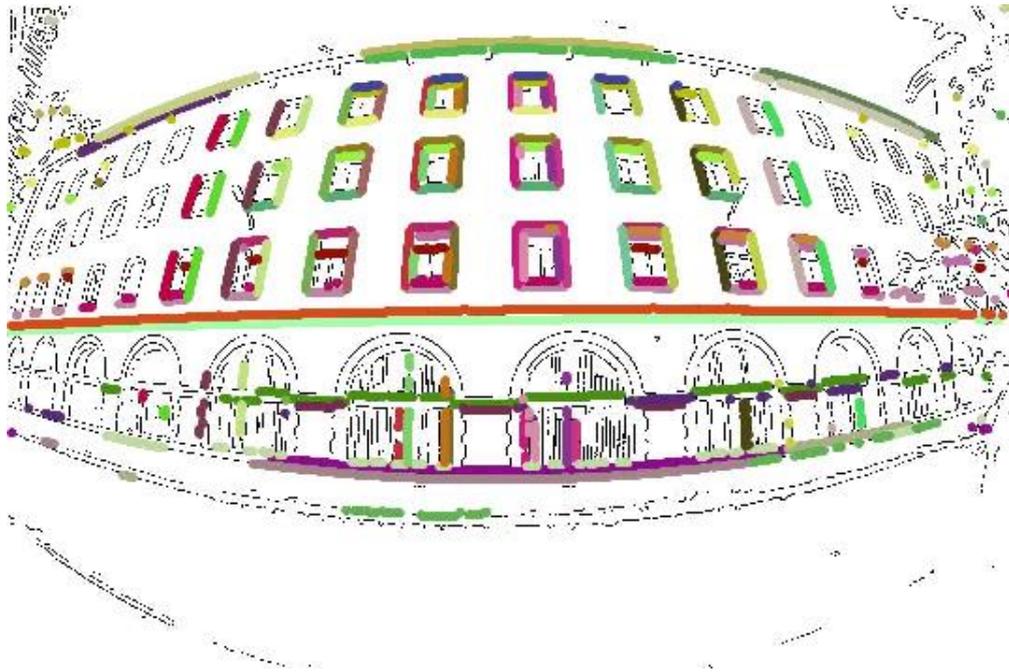
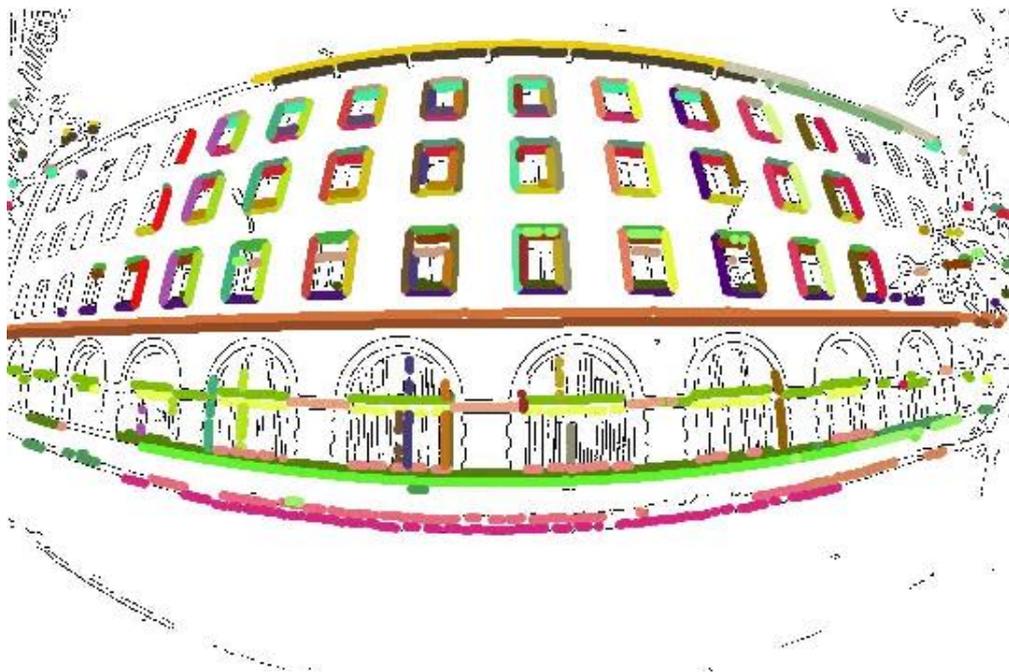


Figura 4.7: Colección de líneas detectadas con la transformada de Hough usando (a) el modelo división y (b) el modelo división con optimización iterativa.

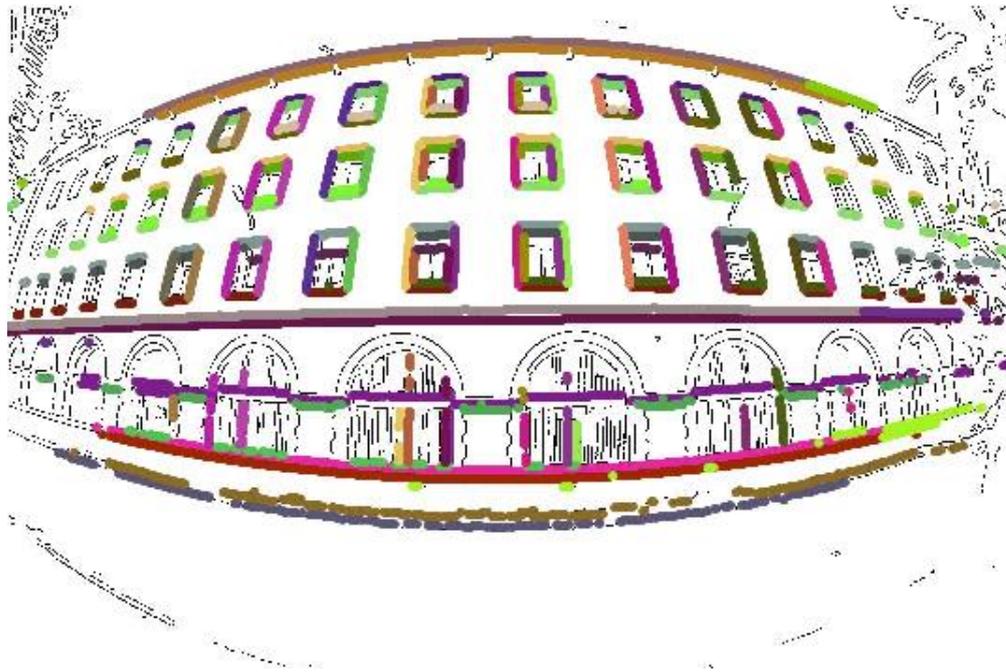


(a)

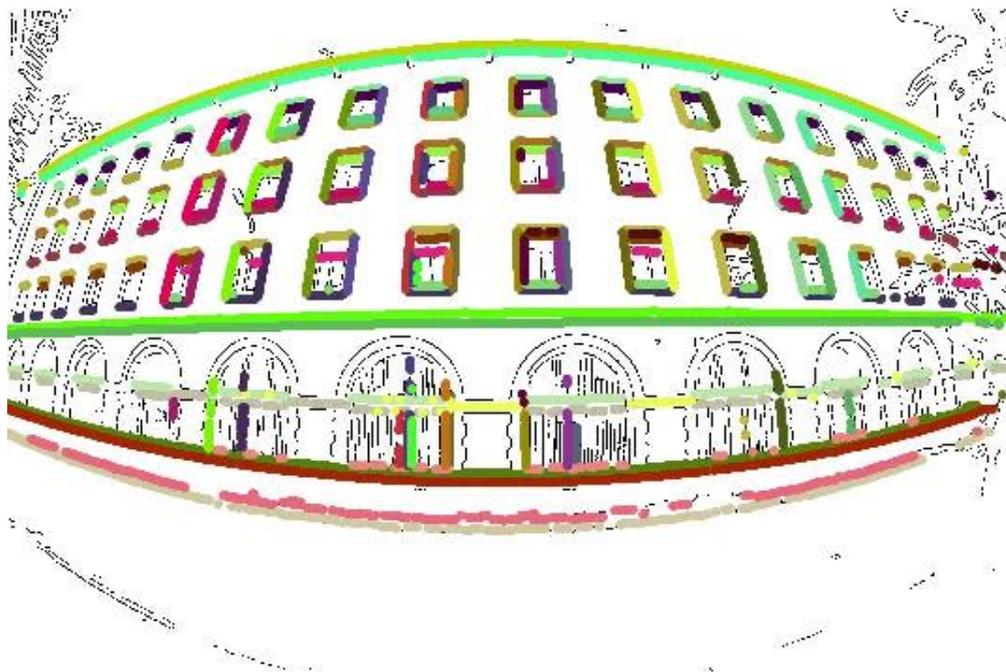


(b)

Figura 4.8: Colección de líneas detectadas con la transformada de Hough usando (a) el modelo polinomial y (b) el modelo polinomial con optimización iterativa.



(a)



(b)

Figura 4.9: Colección de líneas detectadas con la transformada de Hough usando (a) el modelo división y (b) el modelo división con optimización iterativa.



(a)



(b)

Figura 4.10: Colección de líneas detectadas con la transformada de Hough usando (a) el modelo polinomial y (b) el modelo polinomial con optimización iterativa.



(a)



(b)

Figura 4.11: Colección de líneas detectadas con la transformada de Hough usando (a) el modelo división y (b) el modelo división con optimización iterativa.

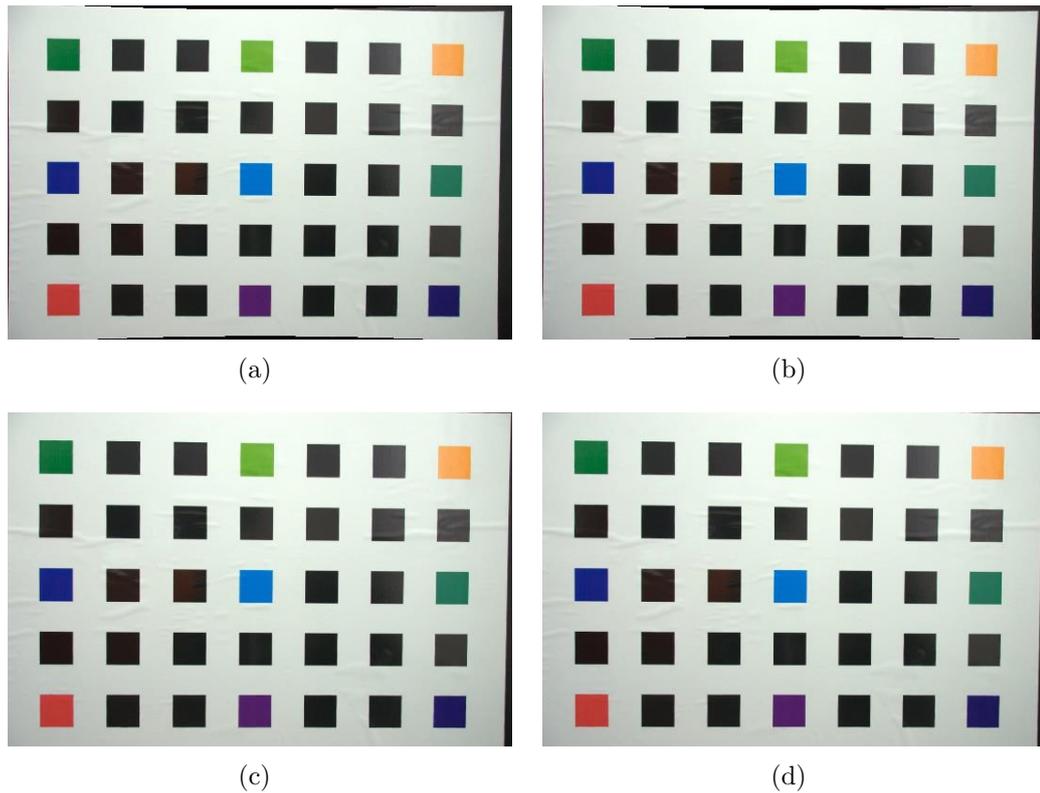


Figura 4.12: Imágenes corregidas usando las siguientes configuraciones: (a) modelo polinomial de un parámetro, (b) modelo división de un parámetro, (c) modelo polinomial de dos parámetros con optimización iterativa y (d) modelo división de dos parámetros con optimización iterativa.

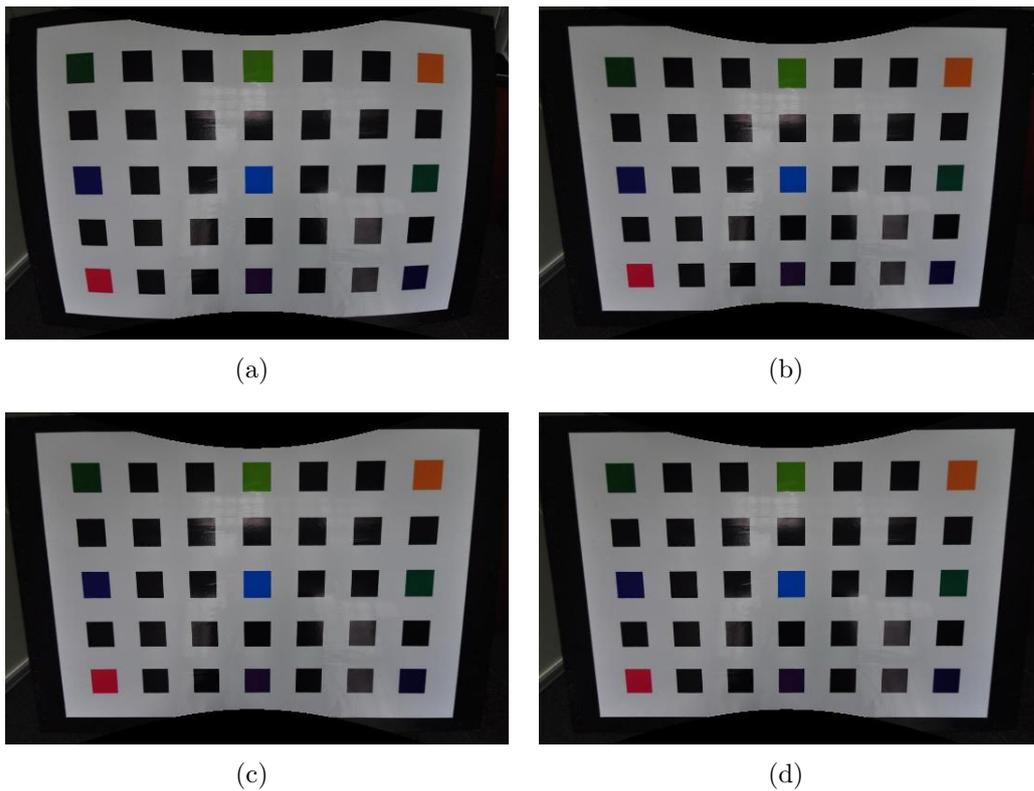


Figura 4.13: Imágenes corregidas usando las siguientes configuraciones: (a) modelo polinomial de un parámetro, (b) modelo división de un parámetro, (c) modelo polinomial de dos parámetros con optimización iterativa y (d) modelo división de dos parámetros con optimización iterativa.

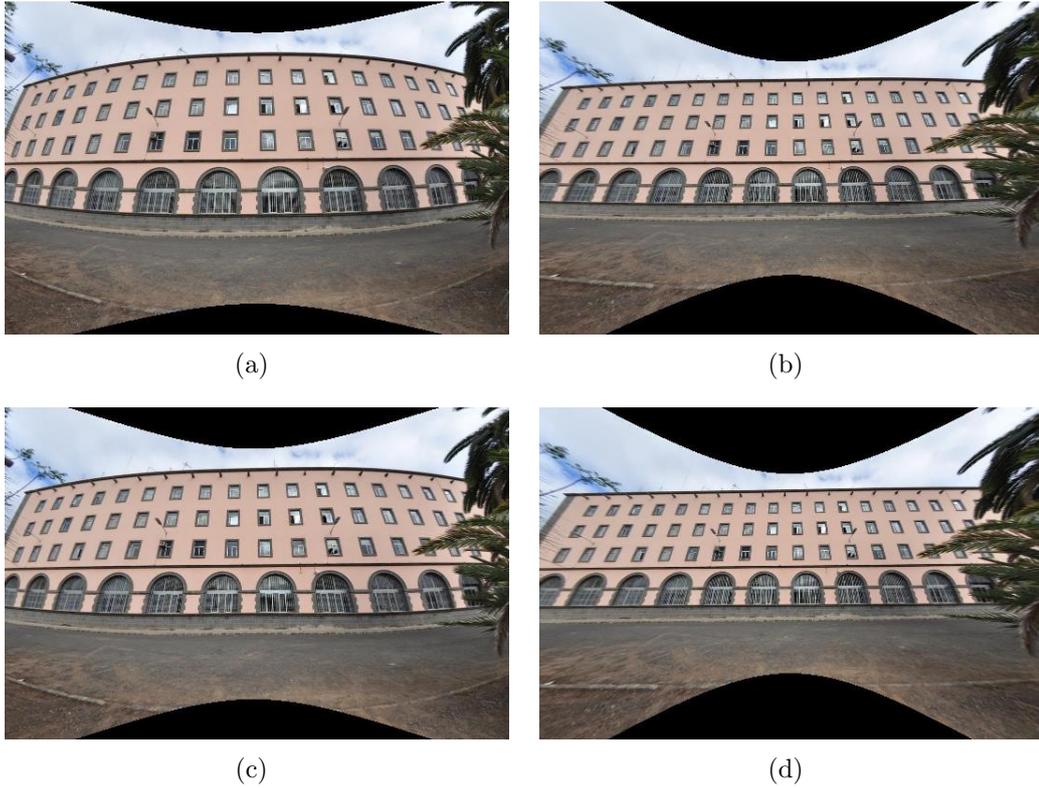


Figura 4.14: Imágenes corregidas usando las siguientes configuraciones: (a) modelo polinomial de un parámetro, (b) modelo división de un parámetro, (c) modelo polinomial de dos parámetros con optimización iterativa y (d) modelo división de dos parámetros con optimización iterativa.

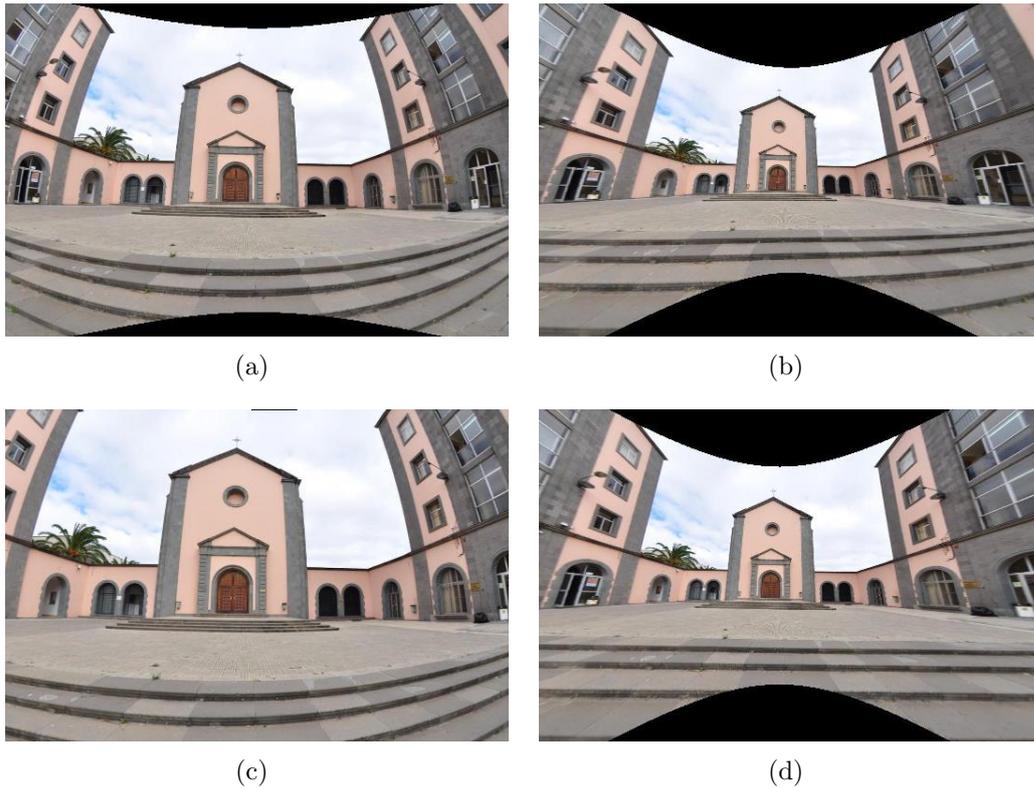


Figura 4.15: Imágenes corregidas usando las siguientes configuraciones: (a) modelo polinomial de un parámetro, (b) modelo división de un parámetro, (c) modelo polinomial de dos parámetros con optimización iterativa y (d) modelo división de dos parámetros con optimización iterativa.

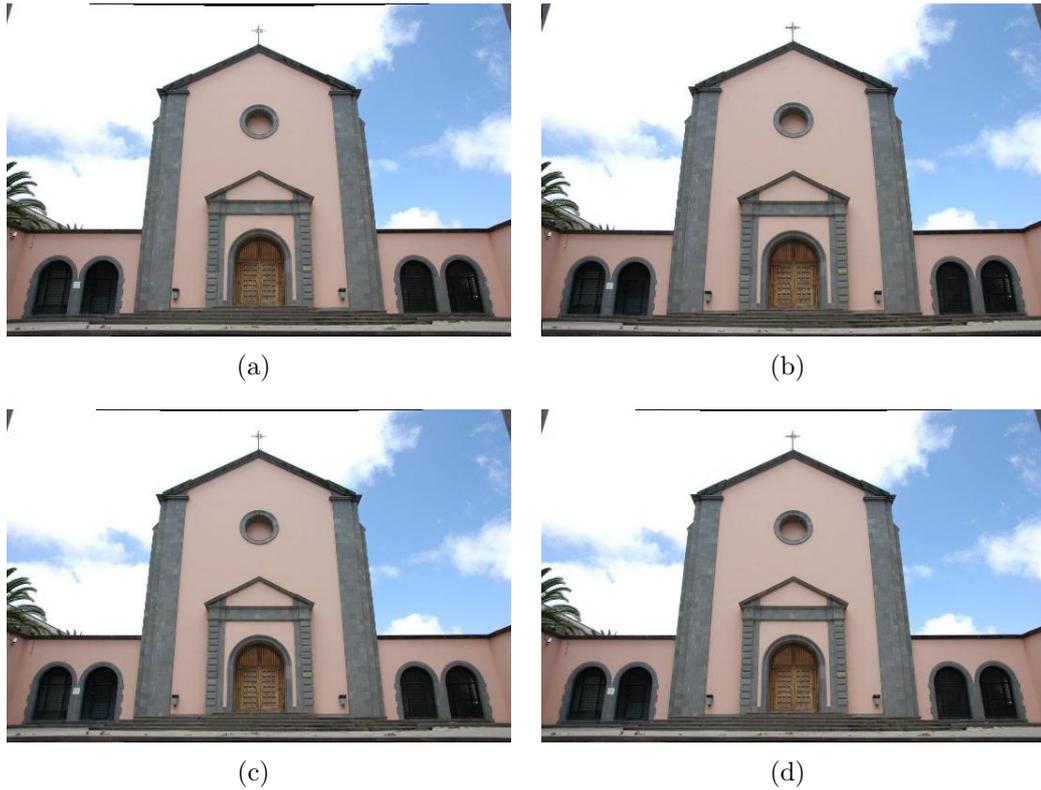


Figura 4.16: Imágenes corregidas usando las siguientes configuraciones: (a) modelo polinomial de un parámetro, (b) modelo división de un parámetro, (c) modelo polinomial de dos parámetros con optimización iterativa y (d) modelo división de dos parámetros con optimización iterativa.

Las figuras 4.12 – 4.16 muestran las imágenes corregidas usando los modelos de un parámetro polinomial y división (en la primera fila) así como el resultado de aplicar la técnica de la optimización iterativa para corregir la distorsión con los modelos de dos parámetros (segunda fila). Las figuras 4.12 y 4.13, ilustran los resultados de corregir la distorsión en el patrón de calibración con diferentes modelos. Aunque los resultados puedan parecer similares para 4.13(b) y 4.13(d), si analizamos las líneas detectadas en 4.6(b) y 4.7(b), podemos ver que la información proporcionada por el modelo división es más completa. Para la línea en la parte superior, aunque las dos partes de la línea están separadas y se encuentran en lugares completamente opuestos en la imagen, son detectadas como una sola línea. De manera similar, el proceso ha sido aplicado a las imágenes descritas anteriormente. Los resultados se muestran en las figuras 4.14 y 4.15.

En la figura 4.14, podemos observar que la introducción de la optimización iterativa en el modelo división 4.14(d) permite corregir la línea inferior en el suelo. El modelo polinomial no es capaz de corregir dicha distorsión, incluso si se incorpora el proceso de optimización iterativa. Esto se debe a las primitivas detectadas. Por lo tanto, si miramos la figura 4.9(b), las líneas detectadas son significativamente más largas que en el resto de los casos, especialmente en la parte superior e inferior del edificio.

En la figura 4.15, los beneficios de introducir la optimización iterativa son especialmente notables si observamos la región de las escaleras. Los modelos polinomiales de uno y dos parámetros no son capaces de corregir la distorsión presente en la imagen. Concretamente, en este caso, el uso del modelo división proporciona resultados mucho mejores cuando se aplica la optimización iterativa (fig. 4.15(d)). Considerando las líneas dibujadas en la figura 4.11(b), podemos observar que las líneas proporcionadas por el modelo de división de dos parámetros con la optimización iterativa son más largas que aquellas que dan como resultado las otras configuraciones. Este hecho es especialmente remarcable en las líneas de las escaleras en la parte inferior de la imagen, así como aquellas que pertenecen a los edificios a ambos lados de la capilla.

Modelo	N. Líneas	N. Puntos	$p_1$	$p_2$	centro	E
Pol1p	27	10504	6.33 %		(536,357.5)	0.494157
Pol2pIO	27	10632	6.96 %	2.79 %	(544.6,337.5)	0.101998
Div1p	30	10455	7.00 %		(536,357.5)	0.916431
Div2pIO	27	10634	6.86 %	2.72 %	(544,337.6)	0.098361

Tabla 4.1: Número de líneas y puntos, parámetros de distorsión, centro de distorsión y energía para el patrón de calibración en las figuras 4.5 y 4.12 usando los modelos polinomial y división de un parámetro (Pol1p y Div1p) y los modelos de dos parámetros con optimización iterativa (Pol2pIO y Div2pIO).

Modelo	N. Líneas	N. Puntos	$p_1$	$p_2$	centro	E
Pol1p	41	9822	4.09 %		(536,356.5)	2.48706
Pol2pIO	44	10293	3.12 %	0.82 %	(536,356.5)	2.03084
Div1p	42	9619	-0.73 %		(536,356.5)	3.97142
Div2pIO	42	10344	3.17 %	1.17 %	(536,356.5)	2.42632

Tabla 4.2: Número de líneas y puntos, parámetros de distorsión, centro de distorsión y energía para la imagen en las figuras 4.5 y 4.16 usando los modelos polinomial y división de un parámetro (Pol1p y Div1p) y los modelos de dos parámetros con optimización iterativa (Pol2pIO y Div2pIO).

Modelo	N. Líneas	N. Puntos	$p_1$	$p_2$	centro	E
Pol1p	38	9536	65.01 %		(536,356)	1.4673
Pol2pIO	34	11051	95.45 %	11.38 %	(525,361.9)	0.532611
Div1p	34	10846	98.72 %		(536,356)	1.53463
Div2pIO	34	<b>11137</b>	118.04 %	13.22 %	(525.1,362.3)	<b>0.261851</b>

Tabla 4.3: Número de líneas y puntos, parámetros de distorsión, centro de distorsión y energía para el patrón de calibración en las figuras 4.6, 4.7 y 4.13 usando los modelos polinomial y división de un parámetro (Pol1p y Div1p) y los modelos de dos parámetros con optimización iterativa (Pol2pIO y Div2pIO).

Modelo	N. Líneas	N. Puntos	$p_1$	$p_2$	centro	E
Pol1p	48	10405	73.24 %		(536,356)	2.63275
Pol2pIO	54	12577	121.02 %	24.41 %	(528.1,361.9)	1.64061
Div1p	47	12990	262.80 %		(536,356)	2.99372
Div2pIO	48	<b>14632</b>	555.61 %	21.41 %	(533,362.02)	<b>1.00159</b>

Tabla 4.4: Número de líneas y puntos, parámetros de distorsión, centro de distorsión y energía para la imagen en las figuras 4.8, 4.9 y 4.14 usando los modelos polinomial y división de un parámetro (Pol1p y Div1p) y los modelos de dos parámetros con optimización iterativa (Pol2pIO y Div2pIO).

Modelo	N. Líneas	N. Puntos	$p_1$	$p_2$	centro	E
Pol1p	47	9550	52.62 %		(536,356)	2.53074
Pol2pIO	47	10212	64.28 %	10.98 %	(571.6,363.2)	2.67954
Div1p	53	14209	321.66 %		(536,356)	3.98304
Div2pIO	45	<b>15874</b>	1732.70 %	22.25 %	(510.2,353.6)	<b>2.58055</b>

Tabla 4.5: Número de líneas y puntos, parámetros de distorsión, centro de distorsión y energía para la imagen en las figuras 4.10, 4.11 y 4.15 usando los modelos polinomial y división de un parámetro (Pol1p y Div1p) y los modelos de dos parámetros con optimización iterativa (Pol2pIO y Div2pIO).

Las tablas 4.1 – 4.5 muestran el número de líneas y puntos, los parámetros de los modelos, el centro de distorsión y el valor final de la energía usando las distintas configuraciones. Las filas denotadas con “Pol1p” y “Div1p” representan los resultados obtenidos usando [3] y [2] respectivamente. Por otro lado, las filas etiquetadas como “Pol2p” y “Div2p”, muestran los resultados usando el método propuesto en este capítulo. Como puede verse, cuando se introduce la optimización iterativa, el número de puntos detectados aumenta de forma significativa en ambos casos, tanto en el caso de usar modelos de tipo polinomial como en el tipo división. También es destacable que la introducción de la técnica propuesta permite minimizar la energía más que en las otras configuraciones. Como consecuencia, esta reducción del error medio significa que los puntos corregidos están más cerca de las líneas rectas que en los casos en los que usamos modelos de un sólo parámetro de distorsión.

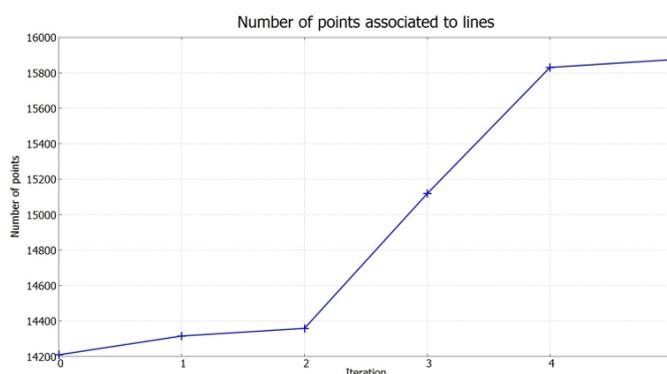


Figura 4.17: Variación del número de puntos de la imagen que se muestra en la figura 4.15(d) usando el esquema de optimización iterativa.

La gráfica de la figura 4.17, ilustra la variación del número de puntos a través del proceso de optimización iterativa usando el modelo división. El primer valor corresponde al número de puntos obtenidos mediante la transformada mejorada de Hough. Como puede observarse, desde la primera iteración de la optimización, obtenemos un número de puntos superior. Esto es debido a que la minimización de la función de energía, que proporciona un modelo optimizado de dos parámetros. Este modelo permite obtener puntos que no fueron considerados cuando usamos un modelo de un parámetro.

En la figura 4.18, mostramos la capacidad del método propuesto frente a una amplia variedad de imágenes. Vemos que el método es capaz de abordar con una fuerte distorsión en diferentes escenarios, obteniendo resultados satisfactorios.



Figura 4.18: Algunos ejemplos de corrección de la distorsión usando el método propuesto. Se puede acceder a algunos resultados adicionales, así como a las imágenes originales en <http://www.ctim.es/WideLensImageDatabase/>

## 4.8. Análisis de la complejidad computacional

En esta sección presentamos un análisis de la complejidad computacional de la técnica propuesta. Si bien este mismo análisis sería similar para el caso de un parámetro de distorsión, puede pensarse que la introducción de las modificaciones propuestas en este capítulo podría introducir cambios significativos. No obstante, la inclusión del proceso de optimización iterativa, en la práctica, sólo revierte en unas pocas iteraciones. Por lo tanto, la complejidad obtenida sería prácticamente la misma.

Antes de comenzar el análisis, empecemos por describir una serie de variables que van a ser usadas durante la descripción:

- $N_{pixels}$ : Número de píxeles en la imagen.
- $N_{edges}$ : Número de píxeles borde.
- $N_{lines}$ : Número de líneas a extraer en el espacio de Hough.
- $S_{distortion}$ : Tamaño del intervalo discretizado de los parámetros de distorsión que se usa en el espacio de Hough.
- $S_{Hough}^{matrix}$ : Tamaño de la matriz de votación de Hough.
- $S_{Hough}^{voting}$ : Tamaño de la vecindad usada por cada punto al votar en la matriz de Hough (esta vecindad incluye las líneas que pasan cerca del punto y con una orientación similar a la orientación del punto).

A continuación describiremos la complejidad atendiendo a cada uno de los pasos de la técnica propuesta.

**Detector de contornos de Canny:** En el algoritmo de detección de bordes de Canny, la complejidad viene determinada por el algoritmo que usamos para implementar la convolución gaussiana. Puesto que usamos el algoritmo rápido recursivo propuesto en [8] como una aproximación de la convolución gaussiana, la cual tiene una complejidad lineal respecto al tamaño de la imagen, podemos concluir que la complejidad del detector de bordes de Canny que usamos es:

$$O(N_{pixels}).$$

**Transformada mejorada de Hough:** La transformada tradicional de Hough tiene una complejidad relacionada con el número de puntos borde proporcionados y las dimensiones de la matriz de votación. En nuestro caso, dividiremos el análisis en las siguientes etapas:

1. Consideramos la complejidad del proceso de votación. El coste computacional de esta etapa está basado en el número de puntos detectados que se proporcionan mediante el detector de bordes de Canny a la transformada mejorada de Hough. Además, hemos de tener en cuenta la dimensión de la matriz de votación, que viene determinada por el tamaño del intervalo para el parámetro normalizado de distorsión de la lente y por el tamaño de los intervalos para la distancia discretizada y el ángulo de variación para cada píxel borde que va a votar. De esta manera, podemos escribir la complejidad de esta parte como:

$$O(N_{edges} \times S_{distortion} \times S_{Hough}^{voting}).$$

2. Así como en la descripción previa, en la selección del máximo consideramos también el tamaño de la matriz de votación de Hough. No obstante, como proporcionamos un número máximo de líneas a considerar, precisamos incluir este número en la estimación de la complejidad. Por tanto, el coste de la selección del máximo será:

$$O(N_{lines} \times S_{Hough}^{voting}).$$

3. La complejidad de llevar a cabo la asociación de los puntos borde detectados con las líneas depende del número de puntos bordes, así como del número de líneas a calcular. De esta manera:

$$O(N_{lines} \times N_{edges}).$$

**Optimización del parámetro de distorsión:** El método modificado de Newton–Raphson que hemos propuesto para la optimización del parámetro de distorsión, converge en unas pocas iteraciones para todos los experimentos que se han realizado. Por tanto, el tiempo de la complejidad viene determinado por la evaluación de la función a optimizar (4.27). En nuestro caso, dicha complejidad viene dada por:

$$O(N_{edges}).$$

**Corrección de la distorsión de la imagen:** Para la corrección de la distorsión, tenemos que recorrer los píxeles de la imagen y corregir la distorsión de la lente, así que la complejidad es:

$$O(N_{pixels}).$$

**Optimización iterativa:** La complejidad del proceso de optimización iterativa es similar al primer punto de la transformada mejorada de Hough descrito arriba. Sin embargo, en esta etapa el método se aplica sin un rango para el parámetro de distorsión normalizado. Así, podemos reescribir la complejidad como:

$$O(N_{edges} \times S_{Hough}^{voting}),$$

y esta complejidad se multiplicaría por el número de iteraciones. En la práctica, sólo hacen falta unas pocas iteraciones para alcanzar el modelo optimizado. Por lo tanto, la complejidad del método viene dada por:

$$O(N_{pixels}) + O(N_{pixels}) + O(N_{edges} \times S_{Hough}^{voting}).$$

Aunque proporcionar una estimación precisa para la relación entre las diferentes variables es difícil, puesto que dependen de la imagen, en general tenemos que  $N_{lines} \ll N_{edges} \ll N_{pixels} \approx S_{Hough}^{matrix}$ . Así pues, considerando las expresiones anteriores, la complejidad total estará alrededor de  $O(N_{pixels})$ .

Por lo tanto, es esperable que la complejidad temporal sea aproximadamente lineal con respecto al número de píxeles. Para respaldar e ilustrar este razonamiento, en la figura 4.20, mostramos el tiempo requerido para 8 imágenes de diferentes tamaños (cruces rojas), que corresponden a subimágenes de la imagen de un patrón de calibración. El tamaño de la imagen  $k^{th}$  (con  $k = 1, \dots, 8$ ) es  $\frac{k}{8}$  el tamaño de la imagen original (fig. 4.19). Usando esta colección de imágenes y el tiempo de CPU correspondiente a cada una, calculamos la regresión lineal (línea en verde). Obtenemos un coeficiente de correlación de 0.9884, lo que indica una complejidad computacional cercana a  $O(N_{pixels})$ .

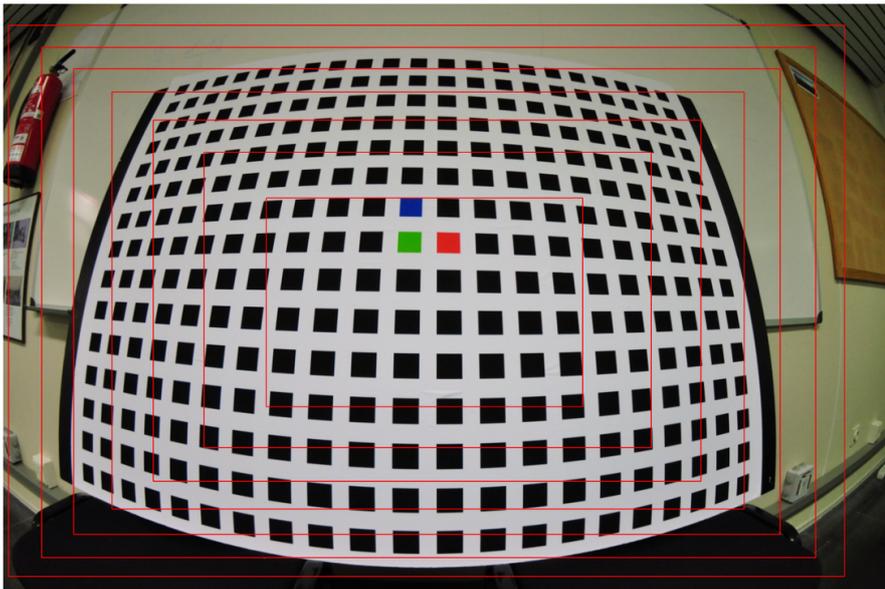


Figura 4.19: Conjunto de imágenes usadas para ilustrar la complejidad del algoritmo. En rojo aparecen las dimensiones de cada una de las imágenes. El tamaño de la imagen  $k^{th}$  (con  $k = 1, \dots, 8$ ) es  $\frac{k}{8}$  el tamaño de la imagen original.

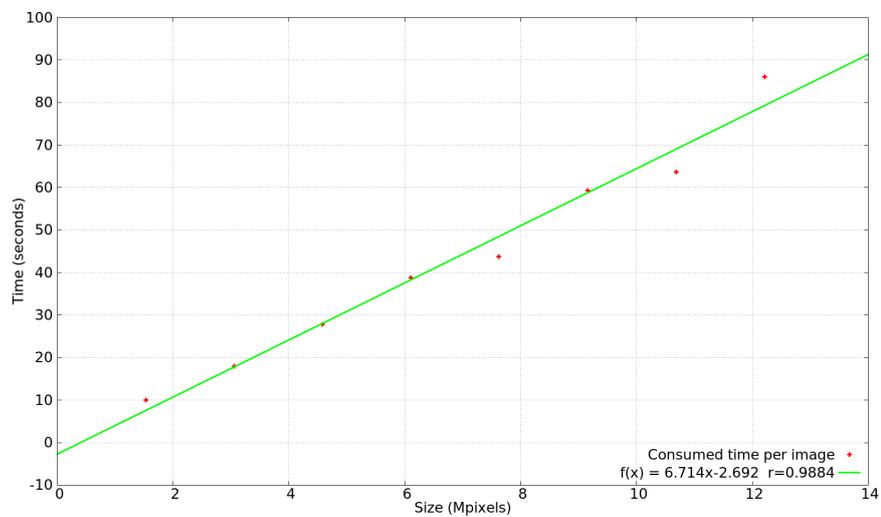


Figura 4.20: Tiempo de CPU para imágenes de diferentes tamaños obtenidas a partir de versiones escaladas del patrón de calibración (cruces rojas). Regresión lineal (línea en verde).

## 4.9. Conclusiones

En este capítulo hemos estudiado la invertibilidad de los modelos de distorsión de la lente de dos parámetros de tipo polinomial y división. Para ello, se han caracterizado matemáticamente las condiciones bajo las cuales los parámetros de distorsión de ambos tipos de modelos definen una aplicación biyectiva en el dominio de la imagen entre la imagen original y la corregida. Este estudio añade robustez a los modelos de distorsión, ya que la función inversa puede construirse de forma fiable y, de esta manera, calcular la imagen corregida. También hemos presentado como normalizar esos parámetros de distorsión en función de un porcentaje de variación del radio de la imagen. De la misma forma que vimos en el capítulo anterior, esto nos da independencia de las dimensiones de la imagen y nos evita trabajar con valores muy pequeños. La introducción del segundo parámetro ha hecho que asimismo se introduzca un segundo porcentaje, que en su caso en lugar de representar la variación desde el centro de distorsión a la esquina más alejada, representa dicha variación desde el punto medio de esa distancia.

Además, también se ha incluido una nueva minimización de la energía. Es decir, la reducción de la distancia entre los puntos corregidos de las primitivas y sus líneas asociadas. En este capítulo se ha incluido en dicha minimización el segundo parámetro de distorsión, además del centro. También hemos introducido un proceso de optimización iterativa, que permite mejorar tanto el modelo de distorsión obtenido como el conjunto de primitivas asociadas. Para ello se aplica sucesivamente una nueva detección de las primitivas en la imagen tras haber aplicado el modelo de distorsión obtenido tras la minimización. Esto nos aporta la capacidad de detectar líneas más largas y, por tanto, que la fiabilidad de cara a corregir la distorsión sea mayor. Todo ello proporciona un método automático para estimar modelos radiales de distorsión de dos parámetros.

Hemos presentado una validación del método sobre una amplia variedad de imágenes reales con diferentes grados de distorsión, obteniendo resultados bastante prometedores. Atendiendo a los experimentos, podemos concluir que el método propuesto es muy robusto y proporciona buenos resultados en una amplia variedad de imágenes. En aquellos casos en los que la distorsión es fuerte, se obtiene una mejora significativa con la introducción de los dos parámetros de distorsión (en lugar de uno) y, además, el modelo división parece proporcionar mejores resultados que el polinomial.

Las principales limitaciones del método propuesto son, primeramente, que la imagen debe contener la proyección de líneas rectas 3D visibles, y en segundo

lugar, que el modelo de distorsión debe ser invertible en la imagen completa. Por lo tanto, podría no funcionar con escenarios con un modelo de distorsión de la lente no invertible, como imágenes de tipo ojo de pez, donde la imagen se inscribe en un círculo rodeada de un fondo negro.

# Capítulo 5

## Estimación automática de modelos de distorsión usando patrones de calibración

### 5.1. Introducción

El uso de patrones de calibración planos formados por imágenes con una distribución de rectángulos, de los cuales conocemos su tamaño y localización, resulta de gran interés para la estimación de modelos de distorsión.

En los capítulos anteriores hemos desarrollado técnicas que permiten estimar el modelo de distorsión sin tener ningún conocimiento sobre la escena, suponiendo simplemente que hay un cierto número de líneas rectas visibles que son proyectadas en la imagen. En ese contexto, hemos usado como criterio de estimación de los modelos de distorsión la rectificación de las rectas distorsionadas visibles en la imagen. Este criterio tiene la ventaja de que no requiere ningún conocimiento de la escena, pero no es un criterio métrico, es decir, no se corresponde con un error de posición de objetos en la escena.

Cuando se manejan patrones de calibración planos tenemos un control absoluto sobre la escena que estamos observando y por tanto podemos diseñar técnicas que hagan uso de ese conocimiento. Concretamente en este capítulo proponemos una nueva técnica que permite simultáneamente estimar de forma automática la

homografía entre la imagen del patrón de calibración y el patrón real minimizando el error de reproyección de los bordes de los rectángulos de la imagen en la imagen del patrón real.

El error de reproyección es mucho más preciso que la mera rectificación de las rectas distorsionadas, porque es un error métrico y tiene en cuenta la posición real de los objetos en la escena. Por tanto, la minimización del error de reproyección es un criterio más sólido y robusto para estimar tanto el modelo de distorsión como la homografía entre la imagen del patrón y el patrón real. Además el error de reproyección nos sirve para comparar la calidad de los resultados obtenidos por diferentes modelos. En particular, nosotros lo usaremos aquí para comparar los resultados de los modelos polinomial y de división para algunas lentes existentes en el mercado. En la figura 5.1 se ilustra el tipo de patrón de calibración que nosotros utilizaremos en los experimentos de este capítulo. Para obtener las fotografías, el patrón real se ha imprimido en una plancha de poliuretano de dimensiones  $1.5m \times 1m$

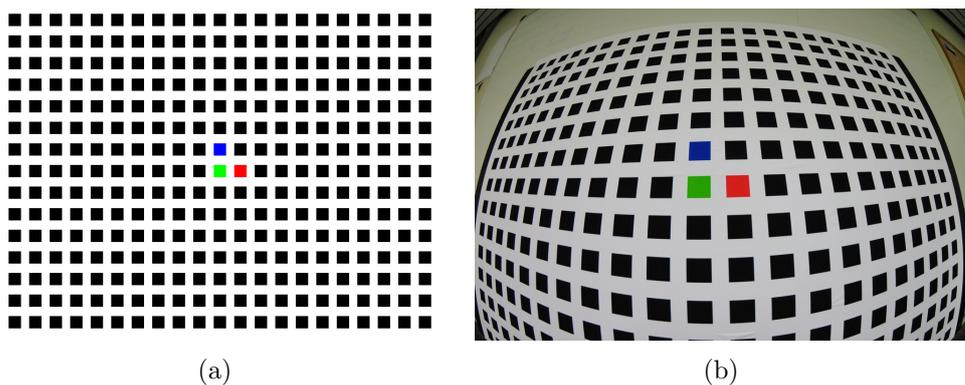


Figura 5.1: (a) Patrón de calibración usado en los experimentos de este capítulo y (b) fotografía de dicho patrón usando la lente Tokina DX 11-16mm.

## 5.2. Técnicas estándar de estimación de los modelos de distorsión de lentes usando patrones de calibración

Las revistas especializadas de fotografía como por ejemplo Photozone (<http://www.photozone.de/>), analizan en detalle las diferentes lentes que hay en el mercado y uno de los aspectos que valoran es la distorsión de la lente. Para ello usan

software profesional desarrollado por empresas especializadas. Uno de los software con mayor implantación en el mercado es el desarrollado por la empresa Imatest (<http://www.imatest.com/>). Dicha empresa ha desarrollado una aplicación específica para el análisis de la distorsión de la lente que usa principalmente modelos de distorsión polinomiales de 1 y 2 parámetros. Una vez estimado el modelo de distorsión, la manera habitual de ilustrar el resultado es usar una imagen como la que se muestra en la figura 5.2. En dicha figura se muestra el perfil de distorsión de la lente Tokina DX 11-16mm (para la distancia focal de 11mm) usando un modelo de distorsión de división con 2 parámetros obtenido a partir de la imagen de la figura 5.1 con las técnicas desarrolladas en este trabajo en capítulos anteriores. Se muestra una malla de líneas sin distorsionar (en color claro) que representan las dimensiones de la escena real que va a ser captada por la cámara y en la parte central (en color oscuro) la transformación de dichas líneas usando el modelo de distorsión obtenido que representaría las dimensiones de la imagen capturada. Como puede apreciarse, la distorsión que presenta esta lente es muy fuerte y de hecho el software suministrado por Imatest no es capaz de calcular un modelo de distorsión para esta lente.

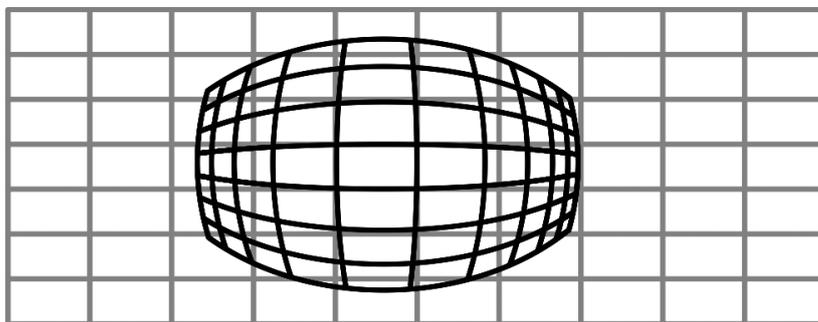


Figura 5.2: Ilustración del perfil de distorsión de la lente Tokina DX 11-16mm usando un modelo de distorsión de división con 2 parámetros. Se muestra una malla de líneas sin distorsionar (en color claro) y en la parte central (en color oscuro) la transformación de dichas líneas usando el modelo de distorsión obtenido.

La aplicación desarrollada por Imatest utiliza patrones de calibración rectangulares y el proceso de estimación de los modelos de distorsión, explicado en

<http://www.imatest.com/docs/distortion/>, pasa por las siguientes fases :

1. Adquisición con la cámara y la lente a analizar de una imagen de un patrón de calibración rectangular.
2. Estimación de las líneas distorsionadas verticales y horizontales.
3. Cálculo de los parámetros de los modelos de distorsión minimizando la distancia de los puntos a las líneas a las que pertenecen una vez corregida la distorsión con el modelo.

Esta metodología presenta las siguientes limitaciones:

1. Las rectas en la imagen del patrón deben ser aproximadamente verticales y horizontales. En caso contrario el método no es capaz de calcularlas correctamente. En particular, si la foto del patrón se aleja de la perspectiva fronto-paralela, el método falla.
2. Si la distorsión es muy fuerte la aplicación falla. Por ejemplo para la lente Tokina DX 11-16mm la aplicación no es capaz de estimar los modelos de distorsión.
3. Cuando se usan patrones de calibración tenemos una información métrica muy precisa sobre la posición relativa de cada una de las rectas que delimitan los rectángulos del patrón. Esta información métrica de distancias entre las diferentes rectas no se usa en absoluto cuando se aplica el criterio de estimar el modelo usando la distancia de los puntos a las líneas correspondientes. Es decir, dicho método no explota adecuadamente toda la información disponible en un patrón de calibración

La metodología propuesta en este trabajo para estimar los modelos de distorsión y evaluar una lente usando imágenes de un patrón de calibración que presentaremos a continuación resuelve estas limitaciones. Asimismo, creemos que representa una mejora respecto a las técnicas que se usan habitualmente en fotografía para evaluar la calidad de las lentes. Entre otras ventajas, cabe destacar, que el procedimiento propuesto es completamente automático, que la adquisición del patrón no tiene que ser fronto-paralela, funciona correctamente para lentes con fuerte distorsión (como la Tokina DX 11-16mm) y explota la información métrica del patrón de calibración relativa a la posición exacta de las rectas del patrón de calibración.

### 5.3. Método propuesto

En primer lugar, como el patrón de calibración es un objeto plano, en el caso en que la cámara se ajuste al modelo *pinhole* de proyección 3D simple, existe una homografía  $H$  dada por una matriz  $3 \times 3$  que determina la transformación, en coordenadas proyectivas, entre la imagen adquirida por la cámara y el patrón de calibración real. Es decir, si  $\bar{x} = (x, y)$  es un punto en la imagen, entonces su correspondiente  $(x_r, y_r)$  en el patrón de calibración real viene dado por el punto  $\bar{x}_r(\bar{x}, H)$  de coordenadas:

$$x_r(\bar{x}, H) = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + h_{22}} \quad (5.1)$$

$$y_r(\bar{x}, H) = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + h_{22}} \quad (5.2)$$

cuando la lente presenta distorsión, entonces es necesario corregirla previamente. Para manejar el modelo de distorsión utilizaremos la notación introducida en el capítulo anterior. Es decir, denotamos por  $\mathbf{u}$  al vector determinado por los parámetros del modelo, y dado un punto  $\bar{x} = (x, y)$  en la imagen denotamos por  $\hat{\bar{x}}^{\mathbf{u}} = (\hat{x}^{\mathbf{u}}, \hat{y}^{\mathbf{u}})$  al punto corregido usando el modelo  $\mathbf{u}$ . Introduciendo el modelo de distorsión en el modelo proyectivo anterior obtenemos las ecuaciones:

$$\bar{x}_r(\bar{x}, H, \mathbf{u}) = \begin{pmatrix} \frac{h_{00}\hat{x}^{\mathbf{u}} + h_{01}\hat{y}^{\mathbf{u}} + h_{02}}{h_{20}\hat{x}^{\mathbf{u}} + h_{21}\hat{y}^{\mathbf{u}} + h_{22}} \\ \frac{h_{10}\hat{x}^{\mathbf{u}} + h_{11}\hat{y}^{\mathbf{u}} + h_{12}}{h_{20}\hat{x}^{\mathbf{u}} + h_{21}\hat{y}^{\mathbf{u}} + h_{22}} \end{pmatrix} \quad (5.3)$$

El método que proponemos en este capítulo consiste en estimar  $H$  y  $\mathbf{u}$  para que la transformación anterior ajuste lo mejor posible la imagen del patrón con el patrón real. En primer lugar, el patrón real está formado por una colección de rectángulos, de los cuales conocemos su posición exacta. Por tanto, también conocemos la posición exacta de todas las rectas verticales y horizontales que delimitan estos rectángulos. Vamos a denotar por  $\{l_k^r\}_{k=1, \dots, N_{\text{lines}}}$  a dicha colección de rectas. Por otro lado, cuando aplicamos los métodos de estimación del modelo de distorsión desarrollados en capítulos anteriores a la imagen del patrón de calibración, y una vez corregida la distorsión con un modelo  $\mathbf{u}$ , obtenemos a su vez otra colección de rectas  $\{l_{k'}(\mathbf{u})\}_{k'=1, \dots, N'_{\text{lines}}}$ , (la  $\mathbf{u}$  indica que las rectas se han calculado después de corregir la distorsión con el modelo  $\mathbf{u}$ ).

Lo primero que vamos a desarrollar es un método que permite poner en correspondencia las rectas en el patrón de calibración  $l_k^r$  con las rectas obtenidas en la imagen  $l_{k'}(\mathbf{u})$ . Para ello, vamos a fijar un sistema de referencia que sea fácilmente localizable, tanto en la imagen como en el patrón real. Dicho sistema de referencia está formado por 3 cuadrados de colores distribuidos en forma de 'L', situados en el centro de la imagen (véase la figura 5.1 donde se ilustra el patrón de calibración usado con estos cuadrados de colores), así como por la fotografía de dicho patrón usando la lente Tokina DX 11-16mm. Usando este patrón de calibración, el método propuesto para encontrar automáticamente las correspondencias entre las rectas  $l_k^r$  y  $l_{k'}(\mathbf{u})$  se puede desglosar en los siguientes pasos:

1. Se detectan en la imagen las posiciones de los cuadros de colores.
2. Se detectan las 8 líneas que bordean estos cuadros de colores en la imagen, denotadas por  $\{l_{k'_1}(\mathbf{u}), \dots, l_{k'_8}(\mathbf{u})\}$ , y que identificamos con sus correspondientes líneas en el patrón de calibración  $\{l_{k_1}^r, \dots, l_{k_8}^r\}$ .
3. A partir de las correspondencias entre  $\{l_{k'_1}(\mathbf{u}), \dots, l_{k'_8}(\mathbf{u})\}$  y  $\{l_{k_1}^r, \dots, l_{k_8}^r\}$ , obtenemos una primera estimación de la homografía  $H$ .
4. Usamos la homografía  $H$  y el modelo de distorsión  $\mathbf{u}$  para proyectar cada una de las rectas  $l_{k'}(\mathbf{u})$  de la imagen, y verificamos si dicha proyección está suficientemente próxima a una recta  $l_k^r$ , en cuyo caso establecemos la correspondencia entre ambas rectas.

Al final de este procedimiento, tendremos una colección de rectas en correspondencia entre la imagen del patrón y el patrón real que denotaremos por  $\{l_{k'_n}(\mathbf{u})\}$  y  $\{l_{k_n}^r\}$ . Es decir, la recta  $l_{k'_n}(\mathbf{u})$  en la imagen se corresponde con la recta  $l_{k_n}^r$  en el patrón real. Además, cada una de esas rectas  $l_{k'_n}(\mathbf{u})$  tiene asociada una colección de puntos en la imagen que denotamos por  $\{\bar{x}_{k'_n,j}\}_{j=1,\dots,N_{k'_n}}$ . Para optimizar  $H$  y  $\mathbf{u}$ , vamos a minimizar el error de reproyección al proyectar los puntos  $\bar{x}_{k'_n,j}$  en el patrón de calibración y a calcular su distancia con la recta correspondiente  $l_{k_n}^r$ . Es decir, el error de reproyección se expresa como:

$$E(H, \mathbf{u}) = \sum_{n=1}^{N_l} \sum_{j=1}^{N_{k'_n}} (\text{distancia}(\bar{x}_r(\bar{x}_{k'_n,j}, H, \mathbf{u}), l_{k_n}^r))^2. \quad (5.4)$$

Es importante señalar que el error de reproyección se mide en términos de las medidas reales del patrón de calibración y, por tanto, es un error métrico

que explota totalmente la configuración del patrón de calibración. Asimismo, da una una información más precisa que el error asociado a la rectificación de las rectas distorsionadas, que es el usado cuando no se dispone de información sobre la posición en el espacio de las rectas.

Para minimizar el error de reproyección  $E(H, \mathbf{u})$  se utiliza el algoritmo de optimización de Levenberg–Marquardt [51]. Tomamos como aproximación inicial el vector  $\mathbf{u}$  que se obtiene de la estimación del modelo de distorsión mediante las técnicas explicadas en los capítulos anteriores. Como homografía  $H$ , utilizamos la que sale de poner en correspondencia las rectas en la imagen y en el patrón de calibración. El algoritmo de Levenberg–Marquardt requiere de una parametrización de la homografía  $H$ . Una homografía viene dada por una matriz  $3 \times 3$  módulo la multiplicación por una constante. Por ello, realmente viene determinada por 8 parámetros. Por otro lado, es bien conocido que  $H$  queda determinada por su valor en 4 puntos del plano que no estén alineados. Utilizando este resultado, en este trabajo parametrizamos la homografía a través de su valor en las 4 esquinas del cuadrado verde de referencia central. El valor de la homografía en cada esquina de dicho cuadrado da lugar a 2 coordenadas que determinan 2 grados de libertad. Por lo tanto, las 4 esquinas determinan 8 grados de libertad, que representan las variables que nosotros usamos para parametrizar la homografía  $H$ . Es decir, en  $E(H, \mathbf{u})$ ,  $H$  se parametriza a través de las 8 coordenadas de los puntos transformados de las esquinas del cuadrado verde de referencia. La implementación del algoritmo de optimización de Levenberg–Marquardt utilizada, así como la parametrización y cálculo básico de las homografías, ha sido desarrollada por investigadores del grupo AMI y no forma parte de las contribuciones realizadas en este trabajo.

Veamos ahora con algo más detalle cómo se localizan los cuadros de colores en la imagen del patrón dispuestos en forma de 'L'. En primer lugar, transformamos los canales RGB de la siguiente manera:

$$\begin{cases} R_{new} = \max\{0, R - \max\{G, B\}\} \\ G_{new} = \max\{0, G - \max\{R, B\}\} \\ B_{new} = \max\{0, B - \max\{R, G\}\} \end{cases}$$

Tras ello, convolucionamos los nuevos canales con una función gaussiana para eliminar el ruido. Finalmente, calculamos el máximo para cada canal, lo cual corresponde a puntos localizados en los cuadrados rojo, verde y azul. En la figura 5.3 ilustramos este procedimiento.

Una vez los cuadrados de colores han sido localizados, determinamos por proxi-

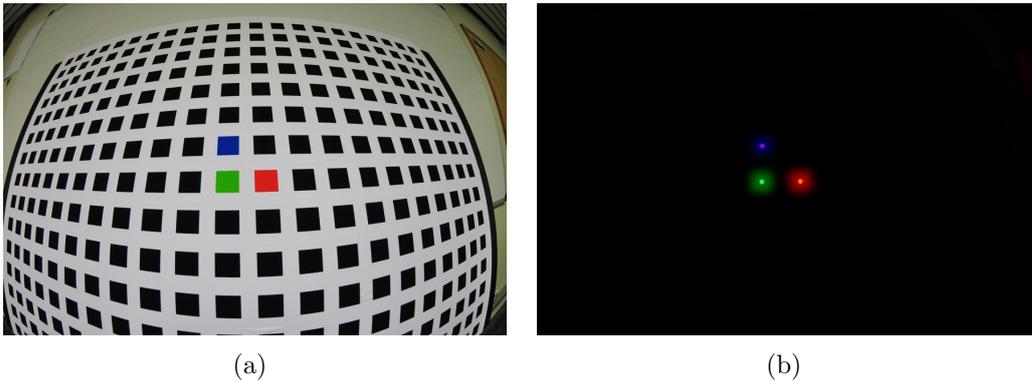


Figura 5.3: Localización de los cuadrados de colores en la imagen: (a) Patrón de calibración con los cuadrados de colores y (b) canales procesados  $R_{new}G_{new}B_{new}$ , donde se extrae el máximo de cada uno de los canales.

midad qué 8 líneas en la imagen corresponden a aquellas que pasan por los bordes de los cuadrados. Usando esas correspondencias de líneas, calculamos una homografía para proyectar todas las líneas en el modelo de referencia en la imagen, y asociarlas con la línea más cercana en la imagen, cuando es factible. Como resultados, obtendremos un conjunto de pares de líneas (referencia y líneas de la imagen) las cuales han sido satisfactoriamente asociadas.

## 5.4. Resultados Experimentales

En las figuras 5.4, 5.5 y 5.6, se ilustran los resultados obtenidos para una variedad de lentes con diferentes niveles de distorsión. Para cada experimento, se muestra la imagen utilizada y el perfil del modelo de distorsión obtenido. En la tabla 5.1 se muestra un análisis cuantitativo de los resultados obtenidos usando modelos polinomiales y de división de 2 parámetros con las técnicas introducidas en el capítulo 4. Asimismo, se presenta una comparación con el modelo polinomial de 2 parámetros, que para esas mismas lentes se puede obtener en la revista Photozone usando el software de Imatest. Estos resultados muestran que el error de rectificación, expresado en píxeles, es significativamente menor usando los modelos obtenidos por las técnicas introducidas en el capítulo 4, en comparación con el modelo obtenido a través de la revista Photozone. Por otro lado, los resultados también sugieren que para estas lentes, que presentan un nivel de distorsión moderado, el modelo polinomial parece dar resultados ligeramente mejores que el modelo de división.

Lente	$E_{rectif.}^{Imatest}$	$E_{rectif.}^{POL2p}$	$E_{rectif.}^{DIV2p}$	$E_{reproy.}^{POL2p}$	$E_{reproy.}^{DIV2p}$
Sigma 8-16 mm (8)	20.25	<b>2.66</b>	2.91	<b>0.36</b>	0.40
Sigma 8-16 mm (12)	7.11	<b>1.83</b>	1.85	<b>0.24</b>	0.25
Nikkor 14-24 (14)	3.87	<b>0.43</b>	0.43	<b>0.08</b>	0.08
Nikkor 14-24 (18)	2.82	<b>0.27</b>	0.27	<b>0.05</b>	0.05
Nikkor 17-35 (17)	3.66	<b>1.30</b>	1.34	<b>0.24</b>	0.25
Nikkor 24-70 (24)	2.97	<b>0.33</b>	0.33	<b>0.07</b>	0.07

Tabla 5.1: Errores obtenidos para diferentes tipos a de lentes. A la izquierda aparece la lente y, entre paréntesis, la distancia focal fijada. En las primeras 3 columnas aparecen los errores de rectificación en píxeles usando la expresión (4.27) para el modelo polinomial de 2 parámetros que se puede obtener en la revista Photozone, así como para los modelos polinomiales y de división obtenidos con las técnicas desarrolladas en el capítulo 4. En las columnas 4 y 5 se muestra el error de reproyección en milímetros para los modelos polinomial y de división usando la técnica desarrollada en este capítulo.

## 5.5. Conclusiones

En este capítulo se aborda la estimación de los modelos de distorsión de lentes usando patrones de calibración. Se ha diseñado un patrón de calibración con unos cuadrados centrales de colores, que permiten fijar una referencia para estimar automáticamente las correspondencias entre las rectas del patrón de calibración real y su imagen. Para explotar convenientemente el conocimiento métrico de las posiciones exactas de todas las rectas del patrón, se propone usar como criterio de estimación del modelo de distorsión la minimización del error de reproyección entre los puntos de las rectas de la imagen y las rectas en el patrón de calibración. Creemos que para valorar la distorsión de una lente, este criterio es más preciso y robusto que la mera rectificación de líneas distorsionadas en la imagen. Además, el método propuesto consigue estimar el modelo de distorsión para lentes con distorsión muy fuerte, donde los métodos habituales usados por las revistas de fotografía especializadas fallan. Los experimentos que se han realizado sobre una variedad de lentes comerciales, muestran que el método propuesto mejora los resultados obtenidos por los procedimientos habituales y que, además, en el caso de lentes con un nivel de distorsión moderado, los modelos polinomiales dan un resultado ligeramente mejor que los modelos de división.

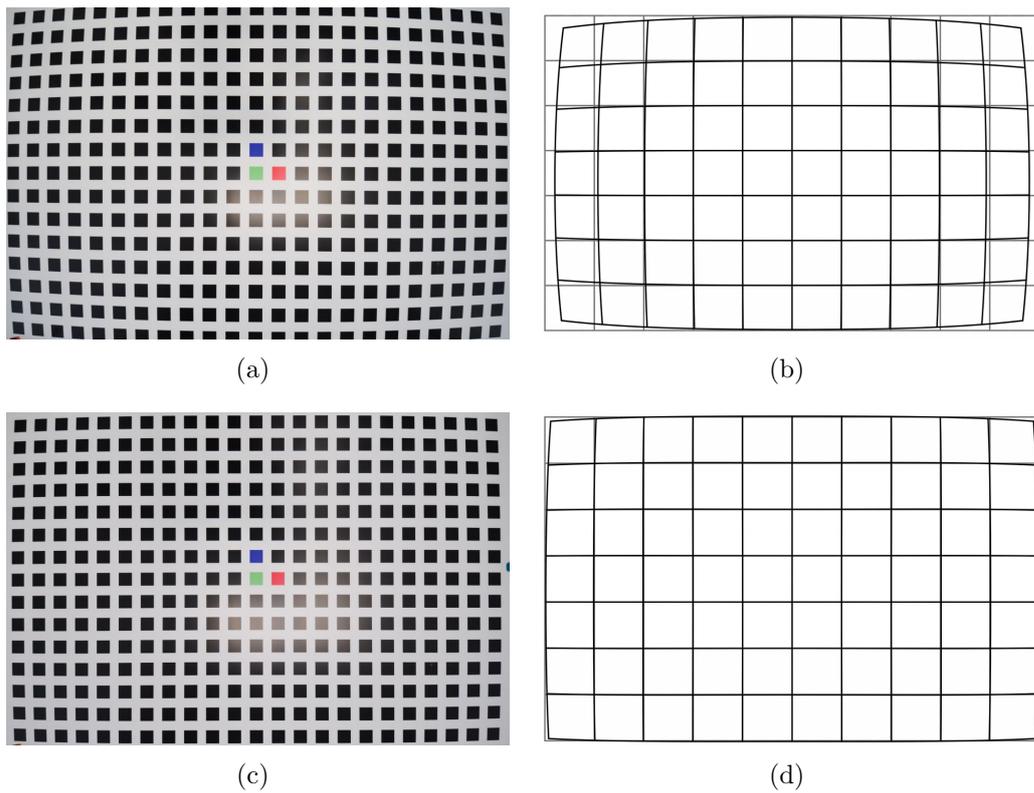


Figura 5.4: Resultados obtenidos para la lente Sigma DX 8-16 mm montada en una cámara Nikon D90. En (a)–(b) se usa la distancia focal 8 y en (c)–(d) la distancia focal 12.

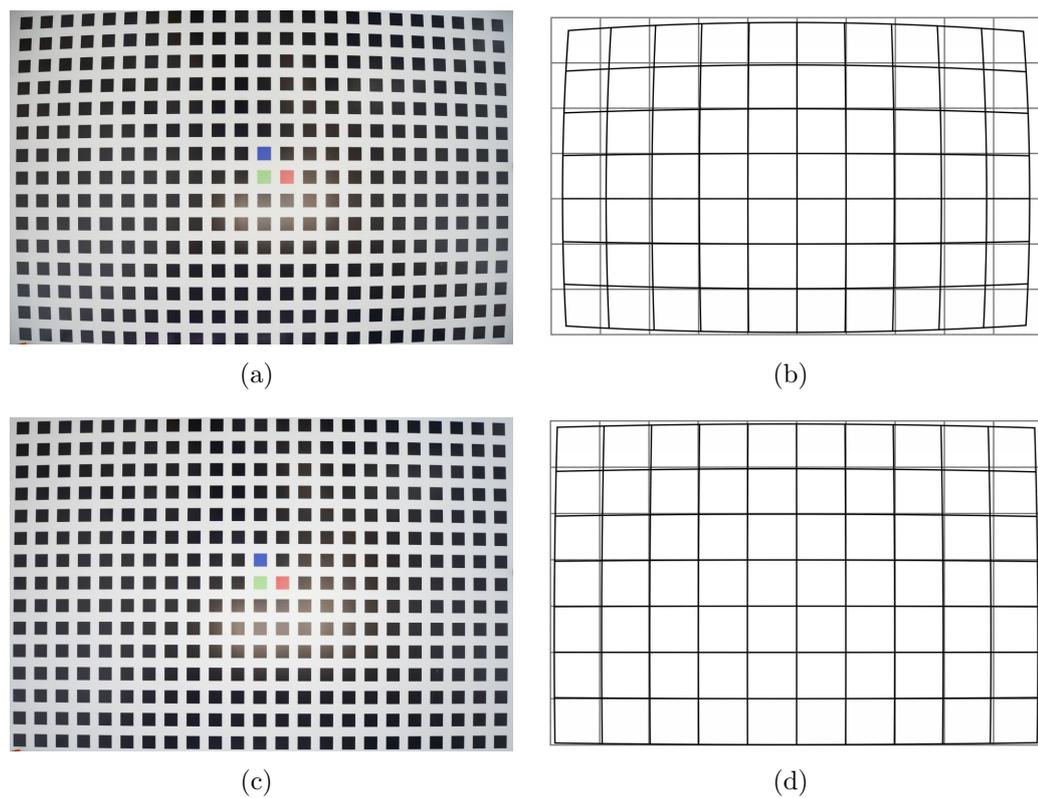


Figura 5.5: Resultados obtenidos para la lente Nikkor 14-24 mm montada en una cámara Nikon D800. En (a)–(b) se usa la distancia focal 14 y en (c)–(d) la distancia focal 18.

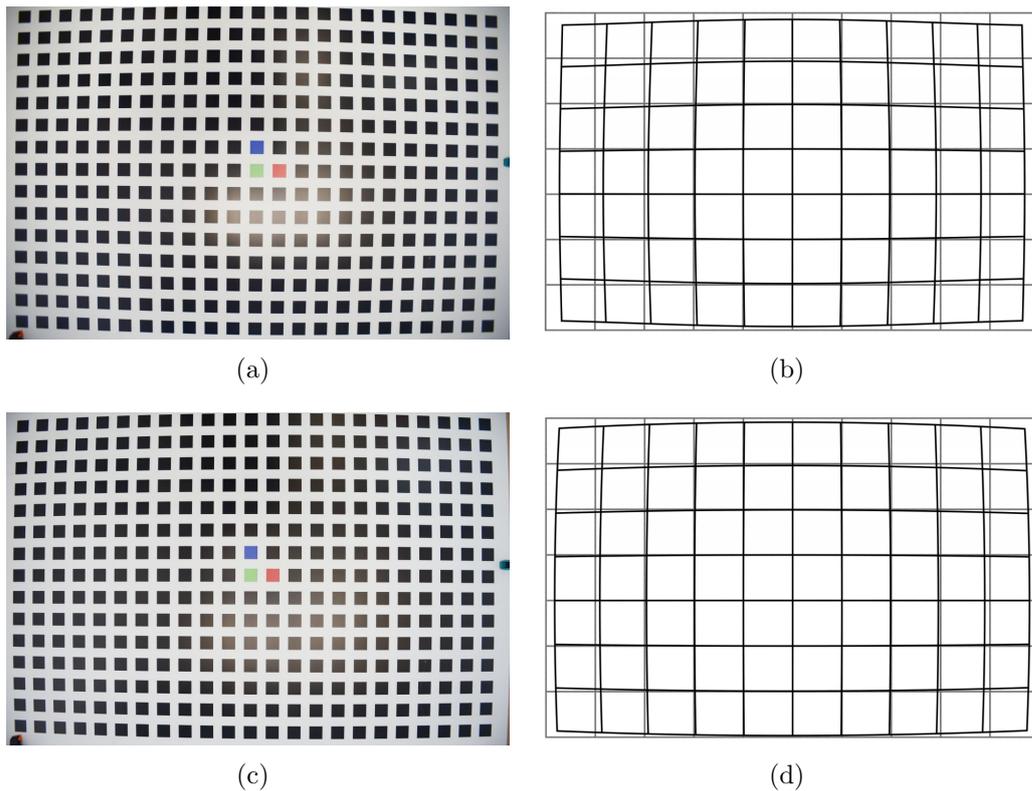


Figura 5.6: En (a)–(b) resultados obtenidos para la lente Nikkor 17-35mm montada en una cámara Nikon D800 fijando la distancia focal a 17. En (c)–(d) resultados obtenidos para la lente Nikkor 24-70mm montada en una cámara Nikon D800 fijando la distancia focal a 24.

# Capítulo 6

## Aplicación a la detección de esquinas en el patrón de calibración de Zhang

### 6.1. Introducción

La calibración de cámaras es uno de los problemas más relevantes en Visión por Computador y consiste en determinar el conjunto de parámetros que describen la correspondencia entre el mundo 3D y las coordenadas 2D de una imagen. Algunas técnicas de calibración no requieren del uso de un patrón de calibración, pero las aproximaciones estándar los requieren con unas medidas conocidas (calibración de fotogrametría). Para extraer los parámetros intrínsecos y extrínsecos de la cámara, se necesitan al menos dos vistas de un patrón plano o una sola vista de un patrón consistente en dos o tres planos ortogonales unos respecto a otros (pueden verse mas ejemplos de este tipo en [70]). Para cámaras de bajo coste o gran angular, la distorsión debida a la lente debe ser incorporada dentro del modelo de la cámara.

El método de Zhang, publicado en [69], es una técnica ampliamente usada para calibración de cámaras. Hace uso de un patrón de calibración plano que contiene una serie de cuadrados ordenados según una cierta configuración. En este capítulo, afrontamos el problema de la detección automática de las esquinas en un patrón de Zhang cuando las imágenes están afectadas significativamente por distorsión de la lente.

Para calibrar la cámara a partir de imágenes de un patrón de calibración de Zhang, el primer paso consiste en extraer un conjunto de características (las esquinas) que se harán corresponder con el patrón de calibración a través de una homografía. Este paso se lleva normalmente a cabo de forma automática para imágenes que no están afectadas por una distorsión significativa. Sin embargo, cuando la distorsión es alta, la técnica (automática) normal para estimar tal correspondencia, implementada en Matlab [1], falla. Esto se debe al hecho de que la distorsión de la lente produce un desplazamiento de la posición esperada de las esquinas de los cuadrados en la imagen y, por tanto, se precisa de intervención humana, lo que hace de la calibración una tarea tediosa y que consume mucho tiempo (una calibración típica usa alrededor de 20 imágenes, lo que implica cientos de esquinas).

Con la idea de afrontar este problema, proponemos usar los algoritmos automáticos para la estimación del modelo de distorsión de la lente expuestos en los capítulos anteriores, y para los cuales no se hace necesaria la intervención humana. La técnica propuesta puede resumirse en los siguientes pasos:

1. Extracción de las líneas distorsionadas en la imagen.
2. Detección de los cuadrados rojo, verde y azul, usados para establecer automáticamente un sistema de referencia en la imagen.
3. Estimación de la correspondencia entre las líneas distorsionadas en la imagen y las líneas del patrón de referencia.
4. Aproximación inicial de la correspondencia de las esquinas de los cuadrados usando la intersección de las líneas distorsionadas.
5. Cálculo de las esquinas con el detector de Harris y refinamiento de la localización de las esquinas usando la aproximación inicial obtenida en el paso previo.

Los pasos 1,2 y 3 utilizan la técnica introducida en el capítulo anterior. En la siguiente sección, presentamos nuestro método automático para extraer las esquinas en imágenes con una alta distorsión. Los resultados se comparan con la correspondencia de esquinas que proporciona Matlab, para ilustrar que nuestro método se comporta mejor en dichos casos. Adicionalmente, el método propuesto puede ser también aplicado para mejorar la calibración de la cámara en casos con una baja distorsión, puesto que la inclusión de la corrección de la distorsión proporciona un número mayor de esquinas y, por tanto, una mejor calibración.

## 6.2. Correspondencia automática de las esquinas

En esta sección presentamos la técnica propuesta para la correspondencia de esquinas. Básicamente, nuestro método consiste en extraer las líneas que delimitan los cuadrados en la imagen, asociarlas con aquellas líneas del patrón de referencia, estimar las esquinas mediante su intersección y refinar su posición mediante el detector de esquinas de Harris[38].

Utilizando la técnica introducida en el capítulo anterior, partimos de que, para cada imagen del patrón, hemos encontrado una colección de líneas en correspondencia entre el patrón real y su imagen. Puesto que necesitamos estimar la posición de las esquinas, debemos calcular la intersección de ciertos pares de líneas, pero previamente ha de corregirse su distorsión. Tras corregirla y calcular las intersecciones, usamos la inversa del modelo de distorsión de la lente para obtener una estimación inicial de la posición de las esquinas en la imagen. Ya que conocemos la correspondencia entre las líneas en la imagen y aquellas del modelo de referencia, podemos deducir la correspondencia de los puntos de las esquinas asociadas.

Para mejorar la estimación inicial de la localización de las esquinas que se han obtenido en el paso previo, usamos las esquinas obtenidas por el método de Harris. Para cada esquina extraída, buscamos la esquina más cercana proporcionada por el detector de esquinas de Harris. El resultado final de este paso es una colección de esquinas de Harris y su correspondencia en el modelo de referencia. Esta información es el punto de partida del método de calibración de Zhang.

## 6.3. Resultados

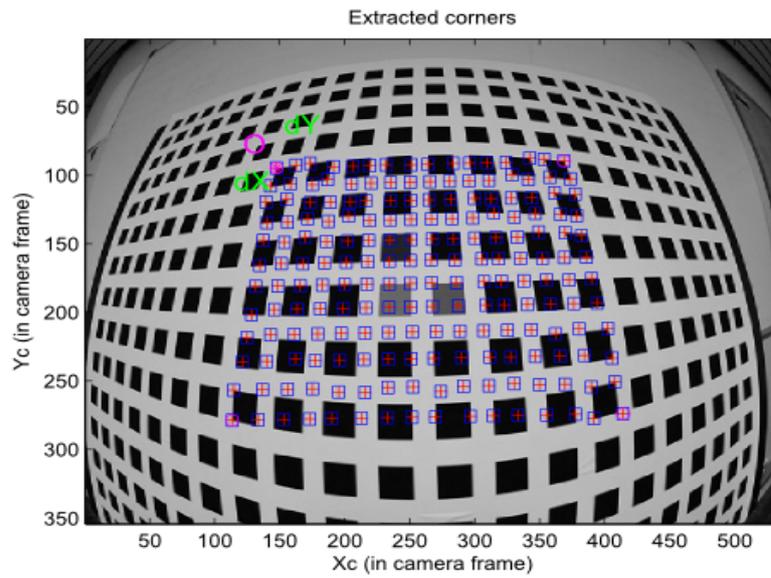
En esta sección comparamos el método propuesto con los resultados obtenidos usando la *toolbox* gratuita de Matlab para calibración de cámaras. Con ello tratamos de mostrar que, usando nuestro algoritmo de detección de líneas, la detección de esquinas mejora significativamente. Primeramente, daremos algunos detalles acerca de la *toolbox* de Matlab, la cual es comúnmente usada para calibración de cámaras.

### 6.3.1. *Toolbox* de Matlab para calibración de cámaras

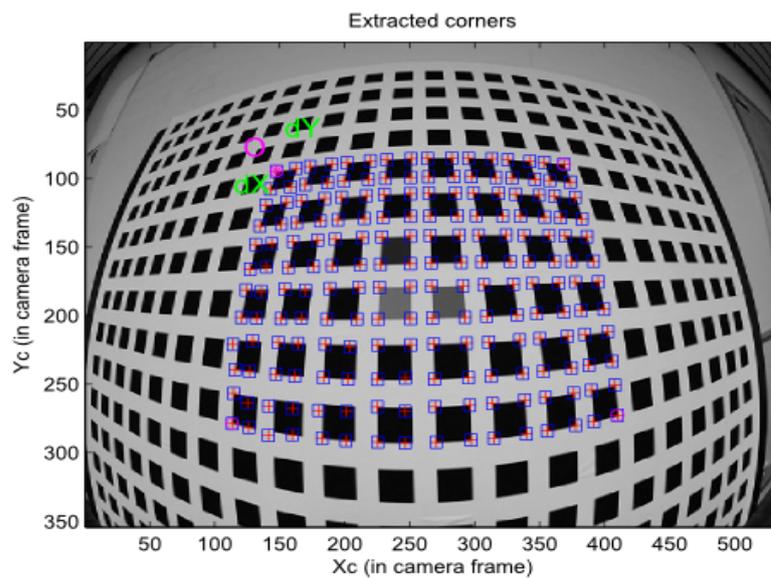
Un *toolbox* de Matlab es un programa específico que se puede descargar, escrito en Matlab y destinado a resolver un problema concreto. Generalmente suelen incluir una interfaz gráfica. En el caso que nos ocupa, la calibración de cámaras, usaremos la *toolbox* desarrollada por Jean-Yves Bouguet, y que puede descargarse de [14]. Esta *toolbox* permite calibrar una cámara usando una implementación particular, la cual es similar al método de Zhang y proporciona los mismos resultados.

El primer paso para usar esta *toolbox*, como en la mayoría de los métodos, consiste en extraer las esquinas para un patrón de calibración dado. Este patrón puede ser cualquier imagen con unas dimensiones conocidas, pero generalmente se usa un tablero de ajedrez plano (pueden verse ejemplos en [14]). La extracción de las esquinas con una alta precisión es una tarea crucial a la hora de obtener una calibración de cámara satisfactoria. Ésta es la razón por la cual esta etapa es extremadamente importante y todos nuestros esfuerzos en esta propuesta están centrados en ella. Adicionalmente, para obtener una calibración de cámara precisa, se requiere más de una imagen, siendo normalmente necesario manejar alrededor de 20 imágenes para cubrir varias perspectivas de cámara y afrontar la distorsión de la lente. Por ello, la extracción de esquinas debe ser llevada a cabo de forma rápida y automática.

Para el proceso de correspondencia, se requiere cierta información de las dimensiones del patrón. En este caso, el usuario debe proporcionar los tamaños de cada cuadrado a lo largo de las direcciones X–Y para facilitar la identificación de las esquinas, aunque se pueden usar algunos valores por defecto. Además, el usuario debe seleccionar una región en la que extraer la información deseada (las esquinas). Se recomienda que el tamaño de ese área sea lo más grande posible, con el objetivo de obtener una calibración precisa. El tamaño del área seleccionada está también limitada por la distorsión puesto que, para casos que muestran una distorsión severa, la extracción de esquinas no se lleva bien a cabo y el proceso debe reducirse a un área de menor tamaño. Un ejemplo para una imagen con una fuerte distorsión radial se muestra en la figura 6.1. Cuando el área que se selecciona es pequeña (como en el caso que se muestra), no se precisa intervención del usuario. Sin embargo, para imágenes con una fuerte distorsión radial, se requiere la intervención del usuario, incluso si el área es pequeña, como puede apreciarse en la figura 6.1(a). Por lo tanto, el usuario debe o bien seleccionar un área reducida, lo cual podría proporcionar una solución pobre de la calibración de la cámara, o ayudar al programa a mejorar la localización de las esquinas proporcionando una estimación de un solo parámetro de distorsión, lo cual no garantiza una solución



(a)



(b)

Figura 6.1: Esquinas proporcionadas por la *toolbox* de Matlab: (a) sin intervención del usuario y (b) con intervención del usuario para fijar un parámetro de distorsión ( $k = -0,4$ ).

6.1(b). Además, éste es un método tedioso, incluso para una única imagen, y bastante complejo de aplicar para un conjunto de imágenes. Corregir la distorsión mediante la estimación de un parámetro de distorsión adecuado, cuando la distorsión es fuerte o cuando el centro de distorsión no está situado en el centro de la imagen, puede ser una tarea extremadamente difícil.

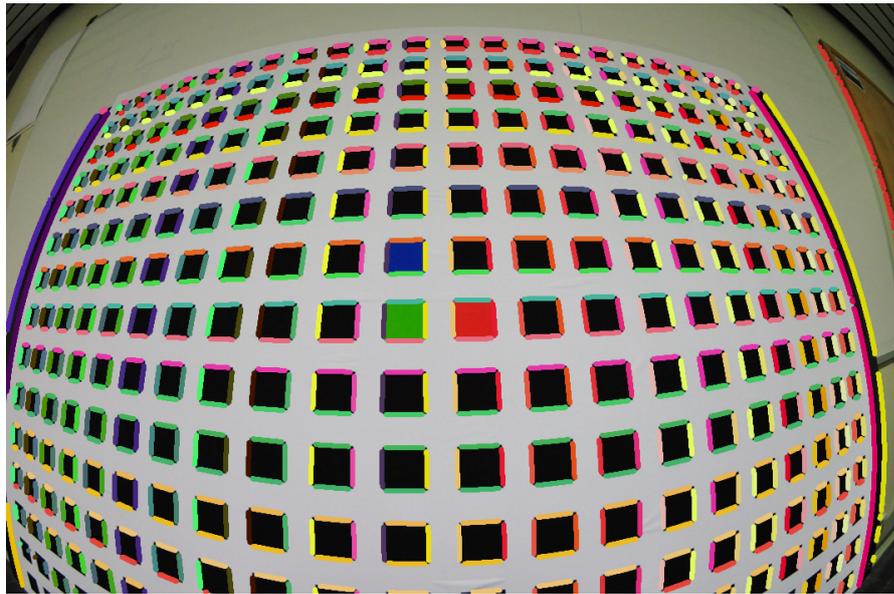
En la figura 6.1 se muestran dos ejemplos. Las cruces rojas corresponden a la posición final de las esquinas proporcionadas por este método. En el primer caso 6.1(a), usando la correspondencia totalmente automática, las esquinas no son identificadas de forma apropiada, puesto que las cruces rojas no encajan con las esquinas de los cuadrados. En el segundo caso 6.1(b), se aplica un valor definido por el usuario para el parámetro de distorsión radial, pero la colección de esquinas sigue siendo insatisfactoria. Aunque la correspondencia mejora notablemente, algunas esquinas no encajan correctamente. Es importante destacar que, usando estas correspondencias para la calibración de la cámara, el resultado esperado sería poco preciso.

### 6.3.2. Resultados usando el método propuesto

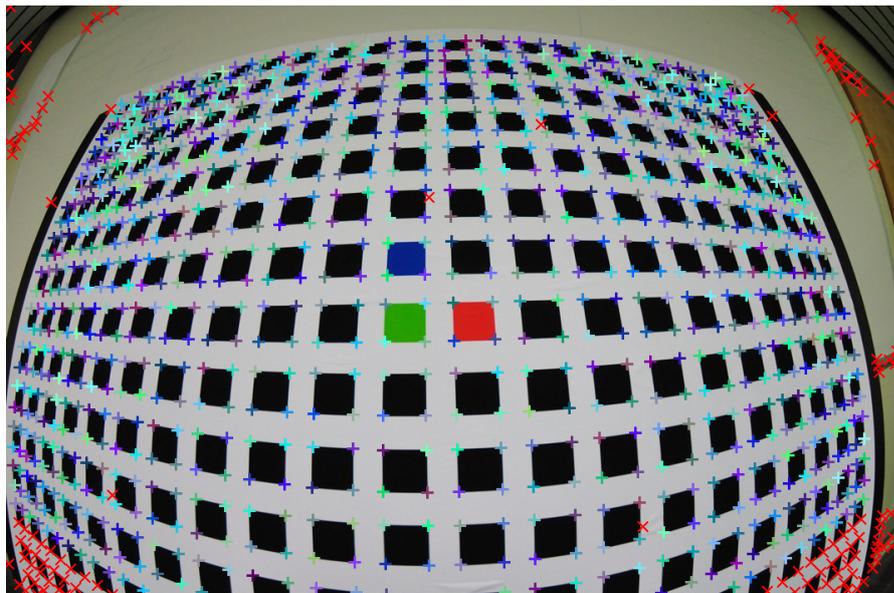
En la figura 6.2, mostramos los resultados obtenidos aplicando el método propuesto. Como se observa, estamos tratando con una imagen altamente distorsionada. Las líneas detectadas (fig. 6.2(a)) están dibujadas usando diferentes colores, donde cada uno de ellos refleja que esos puntos pertenecen a la misma línea. Nótese que todas las líneas son correctamente detectadas. En la figura 6.2(b), las esquinas estimadas mediante la proyección en el patrón usando la homografía y refinando sus posiciones con el detector de esquinas de Harris se muestran con cruces. Aquellas esquinas de Harris que no concuerdan con el modelo de referencia están dibujadas como X.

En vista de estos resultados, nuestra propuesta es capaz de detectar la mayoría de las esquinas en la imagen. De hecho, esto sucede incluso en las posiciones extremas, donde la distorsión radial es bastante severa. La principal ventaja es que este proceso se realiza sin demandar intervención humana.

Debido al hecho de que la imagen en la figura 6.2 muestra una distorsión severa, cuando se hace uso de la *toolbox* de Matlab, se hace necesario seleccionar un área pequeña para la extracción de las esquinas. En ese caso, el máximo número de esquinas a detectar es  $4 \times 6 \times 8 = 192$ . En áreas mayores, observamos que la *toolbox* de Matlab no funciona apropiadamente. Sin embargo, aplicando nuestra propuesta,



(a)



(b)

Figura 6.2: Detección de esquinas usando el método propuesto: (a) detección automática de las líneas distorsionadas y (b) detección final de las esquinas mediante el detector de Harris. Las esquinas que no han coincidido con el modelo de referencia se dibujan como X.

Método	Número de Esquinas	Coste Computacional
<i>Toolbox</i> de MatLab	169	Dependiente del usuario
Nuestra propuesta	1016	16.82 segs.

Tabla 6.1: Número de esquinas detectadas adecuadamente emparejadas y coste computacional usando la *toolbox* de Matalab (fig. 6.1) y nuestra aproximación (fig. 6.2).

se usa la imagen al completo para la extracción de las esquinas. Por lo tanto, el número de esquinas extraídas es mucho mayor (1016). La mejora proporcionada por el método propuesto, puede evaluarse a través del número de esquinas. En la tabla 6.1 se muestran algunos resultados comparativos.

## 6.4. Conclusión

La correspondencia de esquinas entre un patrón de calibración de Zhang y el modelo de referencia es una tarea complicada cuando las imágenes muestran una distorsión de la lente significativa. En particular, los métodos estándar para localizar las esquinas de los cuadrados fallan y se requiere una tediosa intervención por parte del usuario para mejorar dicha localización. En este capítulo, hemos descrito un procedimiento completamente automático para la correspondencia de esquinas que no precisa de intervención del usuario. El método se basa en un algoritmo automático para la corrección de la distorsión de la lente, el cual nos permite estimar las líneas distorsionadas que pasan a través de los bordes de los cuadrados. Tras ello, se obtiene una estimación inicial de las esquinas de los cuadrados a partir de la intersección de dichas líneas. Posteriormente, se mejora la localización de las esquinas usando el detector de esquinas de Harris. Cuando la imagen tiene una distorsión severa, el método propuesto detecta automáticamente un número bastante mayor de correspondencias de esquinas que el método propuesto en la *toolbox* de Matlab, incluso tras la intervención del usuario. Esto da como resultado una información mucho más completa y precisa para el posterior proceso de calibración de la cámara.

# Capítulo 7

## Aplicación a la corrección de la distorsión de lentes en vídeos

### 7.1. Introducción

Una extensión natural de las propuestas descritas en los capítulos anteriores, es la aplicación de estas técnicas a la corrección de la distorsión de lentes en vídeos. Un vídeo no es más que una sucesión de imágenes con ciertas características, como es el número de fotogramas por segundo (*frames per second* o *fps*), que determina el número de imágenes capturadas en un segundo. Existen multitud de formatos, algunos de ellos centrados en la calidad y otros en conseguir el menor tamaño posible de la grabación final mediante técnicas de compresión.

En relación a la distorsión de lentes, existen fundamentalmente dos aspectos a tener en cuenta acerca de la secuencia de vídeo. Por un lado, hay que considerar que la distorsión puede verse afectada por la profundidad de un punto de la escena respecto al plano de proyección de la cámara. En [7], los autores proponen un nuevo modelo de distorsión de la lente que depende de la profundidad, proporcionándose una técnica que permite estimar un nuevo modelo computacional de la lente de la cámara, dando como resultado un mayor nivel de precisión de la calibración de la misma.

Pero, por otro lado, existe otro factor aún más determinante que influye en la distorsión resultante en la captura, el *zoom* de la cámara. En algunos trabajos

como en [4], se propone el uso de nuevos modelos para estudiar la variación de la distorsión de la lente cuando se modifica la distancia focal mediante el *zoom*. Estos modelos se basan en una aproximación polinómica que permite tomar en consideración la variación de la distorsión radial cuando se modifica dicho parámetro. Para ello se usa la minimización de una función de energía del error global, midiendo la distancia entre secuencias de puntos distorsionados alineados y las líneas rectas resultantes después de la corrección de la distorsión.

En este capítulo se describirá el trabajo realizado en colaboración con el Laboratory of Biomechanics and Motor Control de la Universidad de São Paulo, Brasil (LaBioCoM – USP). En dicho laboratorio se desarrollan actividades de investigación en las áreas de deporte y salud. En relación al área de deporte, las investigaciones se centran en los aspectos biomecánicos relacionados con los disparos, acciones de defensa de porteros de fútbol, y la ejecución de saltos y aterrizajes entre otras acciones deportivas. Por otro lado, con respecto al área de salud, el objetivo principal que han tenido ha sido investigar los cambios de control de la marcha y la postura en personas mayores, con el fin de desarrollar propuestas de políticas que ayuden en la reducción de caídas en esta población. Por tanto, las actividades desarrolladas en LaBioCoM, pueden resumirse en el análisis cinemático del movimiento humano y el estudio de la postura y el equilibrio. Fruto de la colaboración llevada a cabo se publicó el trabajo descrito en [64].

El presente capítulo se estructura de la siguiente manera: en primer lugar, se describirá el escenario del problema, así como el estudio realizado a lo largo de la colaboración con el grupo LaBioCoM. Tras haber descrito el caso de estudio, se pasa a explicar la herramienta desarrollada, así como la aplicación de las propuestas de esta tesis a la corrección de la distorsión de la lente en vídeos. Finalmente, en este capítulo se describen los resultados obtenidos.

## 7.2. Contextualización del problema

En el trabajo presentado en [64] se estudia la aplicabilidad y fiabilidad de usar una única cámara gran angular GoPro<sup>®</sup> para realizar el *tracking* y el análisis cinemático de jugadores de fútbol sala. Para ello, se realiza una comparativa con un sistema más complejo, compuesto de tres cámaras equipadas con lentes convencionales.

Generalmente los entornos deportivos de interior presentan una serie de limi-

taciones, las cuales restringen la ubicación de cámaras, de forma que tengan un ángulo de visión lo más amplio posible. Un buen ángulo de visión es importante para realizar un seguimiento más completo de lo que sucede. Poder hacer un *tracking* de la mayor superficie posible del campo es interesante, sobre todo en deportes en los que la acción puede distribuirse por todo el terreno de juego. Debido a estas limitaciones, muchas veces no es posible ubicar una cámara convencional a la distancia suficiente como para cubrir toda la superficie necesaria. Esto lleva a la colocación de múltiples cámaras y al uso de perspectivas como la vista aérea o de pájaro. Por tanto, una posible solución planteada, es el uso de cámaras dotadas con lentes gran angular, que permiten un ángulo de visión de hasta  $180^\circ$  [57]. En el caso que nos ocupa, se propone el uso de una única cámara de estas características para realizar el seguimiento de los jugadores.

El uso de cámaras gran angular para el seguimiento de jugadores en eventos deportivos no es nuevo, y es una idea que se ha planteado en diversos trabajos de la literatura. Existen artículos sobre el uso de cámaras gran angular en escenarios deportivos para obtener variables de movimiento en tenis [52], baloncesto [58], balonmano [59] y squash [65]. La corrección de la distorsión radial se ha llevado a cabo mediante el uso de funciones exponenciales no paramétricas, en lugar de usar la aproximación tradicional con modelos polinomiales [56, 57].

El escenario usado para la grabación del vídeo procesado corresponde a un partido de un equipo profesional brasileño que disputaba los cuartos de final de la liga de São Paulo en el año 2013. Dicho partido fue grabado por cuatro cámaras de vídeo situadas en los puntos más altos del campo, capturando a una frecuencia de adquisición de 30Hz ( $720 \times 480$  y una resolución de 24-bits). Tres de las citadas cámaras estaban provistas de lentes convencionales (SONY<sup>TM</sup> DCD-SR21) y la otra una cámara GoPro<sup>®</sup> Hero3+, que capturaba el terreno de juego completo (fig. 7.1). Se trata de una cámara pequeña ( $55 \times 22 \times 41$ mm) con capacidad para grabar vídeo en alta definición. Puede ser activada por control remoto, pesa 0.74kgr, tiene WiFi y un ángulo de visión de hasta  $170^\circ$ .

Posteriormente, las secuencias de imágenes fueron transferidas a un ordenador con el objetivo de realizar la corrección de la distorsión de las imágenes tomadas con la cámara gran angular y calcular el *tracking* para ambos tipos de vídeo. Es importante destacar que, la distorsión que las lentes gran angular introducen en los vídeos, afecta de forma significativa a cualquier operación de postproceso. En concreto, realizar operaciones de seguimiento en una secuencia de vídeo requiere el uso de modelos lo más precisos posible. No obstante, el uso del modelo división no está tan extendido en el procesado de imágenes para labores de *tracking*.

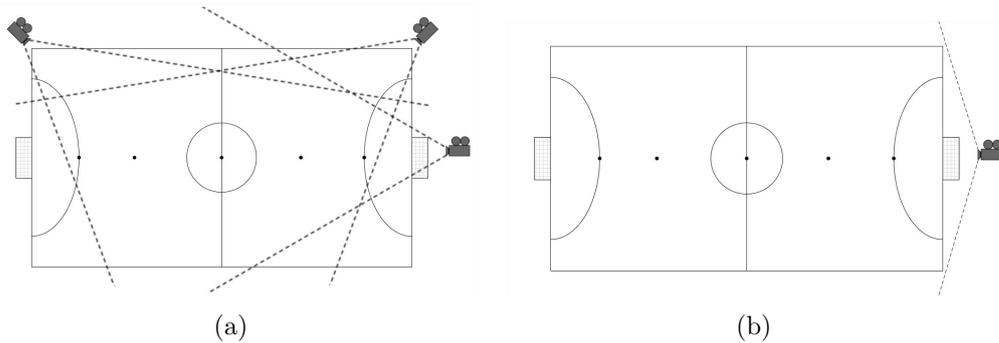


Figura 7.1: Disposición de las cámaras con respecto al campo de fútbol sala: (a) cámaras SONY™ DCD-SR21 provistas de lentes convencionales y (b) cámara GoPro® Hero3+ con lente gran angular

En investigaciones previas se hace uso de al menos dos cámaras [59, 58] y en algunos casos de hasta cuatro [54, 53]. El uso de un mayor número de cámaras tiene principalmente dos consecuencias. Por un lado, el volumen del contenido multimedia se ve aumentado, multiplicándose por el número de cámaras utilizadas durante la grabación. Por otro, y como consecuencia directa de lo anterior, el tiempo de postprocesado aumenta. Por ejemplo, en un partido completo de fútbol sala, una cámara grabando a 30Hz genera un total de 144.000 fotogramas. Por lo tanto, reducir el número de cámaras en este tipo de estudios podría representar una mejora significativa en volumen de almacenamiento y en tiempo de postprocesado/cómputo.

El uso de una única cámara gran angular para la grabación de deportes de equipo no está muy generalizado. Habitualmente este tipo de cámaras, especialmente la utilizada en este estudio, han sido usadas para grabaciones en primera persona de deportes individuales, como la bicicleta de montaña o más actualmente el *parkour*. Su uso ha sido recientemente introducido en evaluaciones de rendimiento en el deporte desde el punto de vista de la biomecánica, especialmente estudios cinemáticos del movimiento humano en *sprints* [30, 13], en un entorno acuático con el objetivo de identificar el patrón cinemático del pie mientras se corría sobre una cinta [41], o muy cerca del agua, como en la investigación sobre el rendimiento de los marineros [60]. En [21] se hace uso de una cámara GoPro® para el seguimiento de jugadores de fútbol. Sin embargo, sólo se describían parcialmente el error de los valores de *tracking* y de la corrección de la distorsión.

El principal objetivo en este capítulo es describir la herramienta desarrollada para corregir la distorsión de una secuencia de vídeo grabado por una cámara do-

tada con una lente gran angular. De esta manera, se podrá demostrar la fiabilidad y si es posible la aplicación de una única cámara para realizar el análisis cinemático del proceso computacional de *tracking* de jugadores de fútbol sala.

## 7.3. Herramienta desarrollada

Como se ha descrito con anterioridad, la cámara situada en la posición descrita en la figura 7.1(b), será aquella para la que se va a realizar una corrección de la distorsión. Dicha cámara se encontrará en una posición estática, enfocando al mismo punto y sin cambiar el *zoom*. Esto facilita el proceso de seguimiento del modelo a lo largo de la secuencia completa de *frames*, puesto que simplemente será necesario hacer su estimación en uno de ellos. Posteriormente, el proceso consistirá en aplicar la corrección a todos los fotogramas, para así generar un vídeo de salida corregido, sobre el que poder realizar el *tracking*.

Previamente a realizar la estimación del modelo de distorsión, es preciso extraer el conjunto de fotogramas presentes en la secuencia. Para ello se ha hecho uso del software FFmpeg [12], bajo licencia LGPL. Este programa es una solución multiplataforma que permite la grabación, conversión y transmisión de audio y vídeo. Permite de una forma sencilla múltiples operaciones sobre audio y vídeo, en concreto aquellas en las que estamos interesados en la aplicación a realizar: extracción de los fotogramas del vídeo y generación de uno nuevo a partir de las imágenes corregidas.

En la figura 7.2 se puede ver el diagrama de flujo de la herramienta desarrollada. El primer paso a realizar es la extracción de los *frames*. Para ello, como se ha mencionado anteriormente, se hace uso de la herramienta FFmpeg, que proporciona como salida el conjunto de fotogramas de la secuencia de vídeo sin corregir. Estos serán ubicados en un directorio denominado con el nombre del vídeo en cuestión seguido de “\_original.frames”. Este directorio proporciona el punto de partida para el método a aplicar. El siguiente paso, será realizar la estimación del modelo de distorsión de la lente de la cámara a partir del primer fotograma. Debido a que la cámara se encuentra en una posición estática, enfocando al mismo punto y no varía el *zoom*, simplemente bastará con realizar dicha estimación sobre el primer *frame* de la secuencia. Para dicho fin se aplica la técnica descrita en el capítulo 4, es decir:

1. Detección de bordes usando Canny y refinamiento de los mismos.

2. Búsqueda de un conjunto de líneas candidatas que mejor se ajustan a las líneas distorsionadas y un primer valor para el parámetro de distorsión de un modelo división de un parámetro.
3. Mejora iterativa del conjunto de primitivas así como del modelo de distorsión, pasándose a uno de dos parámetros donde además se optimiza el centro de distorsión de la lente.

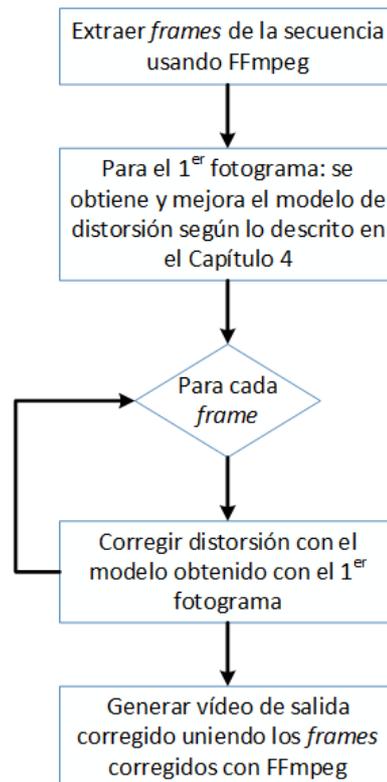


Figura 7.2: Diagrama de flujo del sistema desarrollado para realizar la corrección de la distorsión en una secuencia de vídeo

Una vez se ha obtenido la estimación del modelo de distorsión de la lente usando el primer fotograma de la secuencia de vídeo, se pasa al proceso de corrección. Dadas las características de la secuencia que se han descrito con anterioridad, no será necesario hacer una estimación posterior del modelo, ya que éste será el mismo para todo el vídeo. El proceso de corrección de los *frames* se hará de forma paralela. Es decir, mediante el uso de la librería OpenMP, se realiza la corrección de varios fotogramas de manera simultánea. En esta fase no es importante el orden en que se realice dicho proceso, ya que cada fotograma se encuentra numerado y

la construcción de la nueva secuencia no se llevará a cabo hasta haber terminado el proceso de corrección y todos se encuentren en nuevo directorio denominado “\_corrected\_frames”. Finalmente, y haciendo uso de nuevo de FFmpeg, se construye una nueva secuencia de vídeo en la que los fotogramas ya se encuentran corregidos y, por tanto, se puede realizar el proceso de *tracking* sobre los movimientos de los jugadores. De manera resumida, los pasos del algoritmo serían los que se muestran en la tabla 7.1.

Destacamos que el método propuesto es completamente automático y no requiere de intervención del usuario, más allá de proporcionar el vídeo de entrada.

Una vez se ha realizado la corrección de la distorsión de la lente de la cámara en el vídeo, el siguiente paso consiste en la realización del *tracking*. Para ello, los investigadores de LaBioCoM hicieron uso del software DVideo<sup>®</sup> [11, 10]. Para obtener la sincronización de las imágenes, se identifican eventos comunes en las regiones solapadas del sistema múltiples cámaras. Tras ello, se ha realiza una calibración basándose en 23 puntos de control en la superficie del campo, con medidas previas de distancias reales. El *tracking* se hace de forma manual *frame a frame* por un experto mediante clics de ratón. Se analizan ambas secuencias, tanto las obtenidas de las cámaras con lentes convencionales como la secuencia resultante de la corrección de la distorsión de la cámara con la lente gran angular. En total se procesaron los 8 primeros minutos de juego, lo que corresponde a 14.800 fotogramas.

Tras haberse realizado el marcado de los *frames*, los puntos de calibración y las trayectorias de los jugadores se construyen por medio del método de transformación lineal directa (*direct linear transformation* – DLT), lo que proporciona coordenadas 2D de los jugadores con respecto a las coordenadas del campo y comparadas con el *tracking* realizado para ambos tipos de cámaras. Los datos finales de la distancia recorrida, velocidad media y aceleración se calcularon usando Matlab.

## 7.4. Resultados

En esta sección se describirán los resultados obtenidos, tanto desde el punto de vista de la corrección de la distorsión de la lente de la cámara usada, como considerando la comparativa entre los dos sistemas de cámaras.

En la figura 7.3 se muestra el resultado de aplicar la estimación del modelo de

Etapa	Descripción
1.	Extracción de los <i>frames</i> de la secuencia de vídeo. in: Secuencia de vídeo. out: Conjunto de fotogramas extraídos.
2.	Para el primer fotograma: se extraen, automáticamente, el conjunto de líneas usando una transformada de Hough modificada en la que se incluye un modelo de distorsión de tipo división de un parámetro. El resultado de este paso es el conjunto de líneas distorsionadas y una estimación (valor inicial) del primer parámetro de distorsión. in: Primer <i>frame</i> distorsionado de la secuencia de vídeo. out: Estimación del parámetro de distorsión y el conjunto de líneas más votadas.
3.	Para el primer fotograma: se mejora el modelo de distorsión usando técnicas de optimización no lineal. El resultado de este paso es un modelo división de dos parámetros optimizado, así como la optimización del centro de distorsión de la lente y el conjunto de primitivas obtenido durante la optimización. in: Modelo de distorsión estimado y el conjunto de líneas más votadas. out: Modelo de distorsión de dos parámetros optimizado, incluyendo el centro y el conjunto de primitivas obtenidas.
4.	Corrección de la distorsión: se corrige la distorsión radial de cada fotograma del vídeo usando el modelo de distorsión de la lente estimado para el primer fotograma. in: Modelo de distorsión de la lente optimizado y <i>frames</i> de la secuencia de vídeo. out: Conjunto de todos los fotogramas corregidos.
5.	Generación del vídeo de salida. in: Conjunto de <i>frames</i> corregidos. out: Secuencia de vídeo corregida.

Tabla 7.1: Resumen de los pasos del algoritmo de corrección de la secuencia de vídeo.

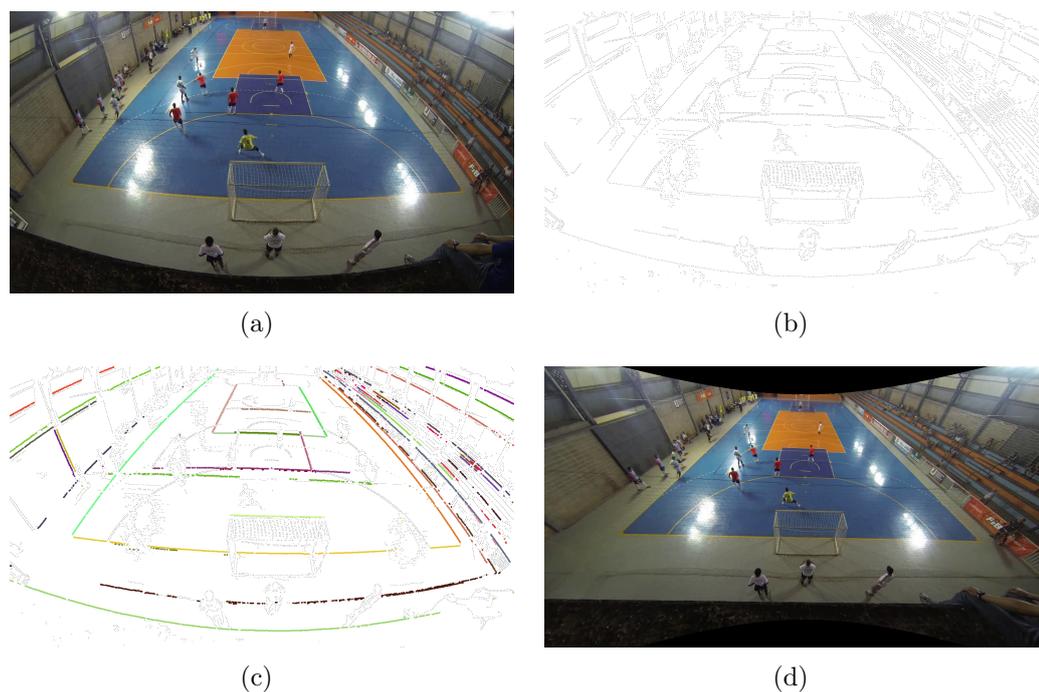


Figura 7.3: Ejemplo del proceso de estimación del modelo de distorsión de la lente para el primer *frame* de la secuencia: (a) fotograma original, (b) bordes detectados con el método de Canny y posteriormente refinados, (c) líneas finales detectadas tras el proceso iterativo y (d) imagen resultante de la corrección de la distorsión.

distorsión y posterior corrección en el primer *frame* de la secuencia de vídeo usada. En 7.3(a) se encuentra el fotograma original, en el cual puede apreciarse la notable distorsión introducida por la cámara GoPro<sup>®</sup>. En 7.3(b) aparece el resultado de la detección de bordes usando el método de Canny y el proceso de refinado de los contornos descrito en la sección 2.3 del capítulo 2, dedicado a la detección de contornos. En la imagen 7.3(c) puede observarse el resultado de detectar las primitivas a través del proceso iterativo descrito en el capítulo 4. A pesar de la distorsión introducida por la lente, se consiguen detectar prácticamente todas las líneas que delimitan el campo, las de las paredes laterales e incluso la recta perteneciente al borde del muro sobre el cual está colocada la cámara. Finalmente, en 7.3(d), se muestra el resultado tras realizar la corrección de la distorsión presente en la imagen de entrada correspondiente al primer *frame* de la secuencia de vídeo.

El siguiente paso llevado a cabo por la herramienta desarrollada, consiste en extrapolar la misma corrección realizada en el primer fotograma al resto de la secuencia. En la figura 7.4 se muestra un ejemplo de las entradas y salidas de

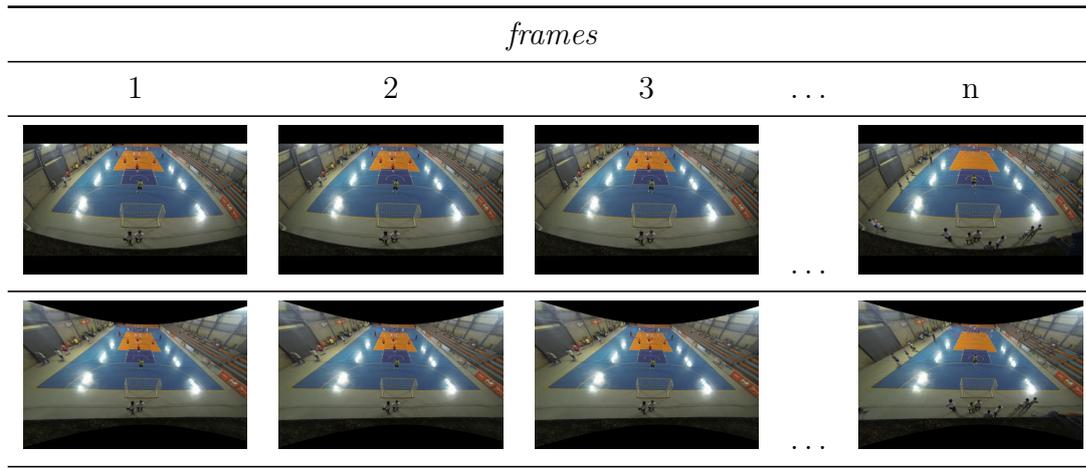


Figura 7.4: Ejemplo de la generalización de un mismo modelo de distorsión extraído del primer fotograma y aplicado a una sucesión de *frames* pertenecientes a una misma secuencia de vídeo.

este proceso. En la fila superior, se encuentra el conjunto de *frames* de entrada, mientras que en la fila inferior aparecen los corregidos. Puede verse que, aunque el modelo ha sido solamente estimado para el primero de los fotogramas de la secuencia de vídeo, proporciona una adecuada corrección también en el resto de ellos.

En relación al *tracking* de los jugadores de fútbol sala realizado a partir de la secuencia de vídeo corregido, se plantean varias medidas a considerar: distancia recorrida por los jugadores, velocidad y aceleración medias. Además del tratamiento de la secuencia de vídeo descrita en la sección anterior, también se realizó un experimento consistente en seleccionar a uno de los jugadores y hacerle recorrer ciertas distancias con una ruta predefinida un número de veces determinado. En el caso del cálculo de la distancia recorrida, se hizo una correlación cruzada demostrando que ambos tipos de secuencias (con cámaras dotadas de lentes convencionales y la cámara gran angular) proporcionaban resultados similares, con un coeficiente de correlación  $r = 0.99$ . En el caso de valores de variables de interés obtenidas mediante el *tracking* con la cámara gran angular, se calculó el error cuadrático medio, comparando el valor de referencia ( $M_{REF}$ ) con el valor obtenido experimentalmente ( $M_{EXP}$ ) en cada uno de los *frames* de la secuencia según la ecuación

(7.1).

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n |M_{EXP}^i - M_{REF}^i|^2} \quad (7.1)$$

## 7.5. Conclusiones

De los resultados presentados en [64] relativos a las distancias recorridas por los jugadores y a la velocidad y aceleración medias, se obtienen cifras competitivas comparadas con las existentes en la literatura y que usan sistemas con múltiples cámaras. Por tanto, es factible el uso de una sola cámara gran angular para la realización del seguimiento de parámetros cinemáticos de los jugadores de fútbol sala, siempre y cuando se aplique un sistema de corrección de la distorsión de la secuencia como el descrito. Además de conseguirse dichos resultados, cabe mencionar que también existen ventajas frente al uso de otros sistemas de grabación, como son la reducción del espacio de almacenamiento necesario y del tiempo de postprocesado.

En las figuras 7.5 y 7.6, se muestran dos ejemplos más de la detección de primitivas, estimación del modelo de distorsión y posterior corrección en dos *frames* de otra secuencia de vídeo. En la 7.5 se muestra un fotograma en el cual se han colocado cables en posición horizontal, mientras que en 7.6 se encuentran dispuestos verticalmente. En ambos casos las líneas son detectadas y se corrige la distorsión presente en ambos fotogramas, incluso cuando éstas no se encuentran totalmente visibles en la escena, como es el caso de la línea situada en la parte inferior en la figura 7.5. En ambos casos, se tratan de imágenes que son parte de una segunda colaboración que se está llevando a cabo con el Instituto de Tecnología de la Universidad Argentina de la Empresa<sup>1</sup>. El objetivo es, partir de la secuencia de vídeo corregida, realizar el seguimiento de ciertos peatones (que se distinguen del resto gracias a una referencia que puede verse en la figura 7.7), para hacer estimaciones acerca de sus trayectorias, cruces y colisiones.

---

<sup>1</sup>Colaboración realizada en el marco del proyecto de referencia PICT-2011-1238, dirigido por el Dr. Daniel Parisi y financiado por La Agencia Nacional de Promoción Científica y Tecnológica de la Argentina

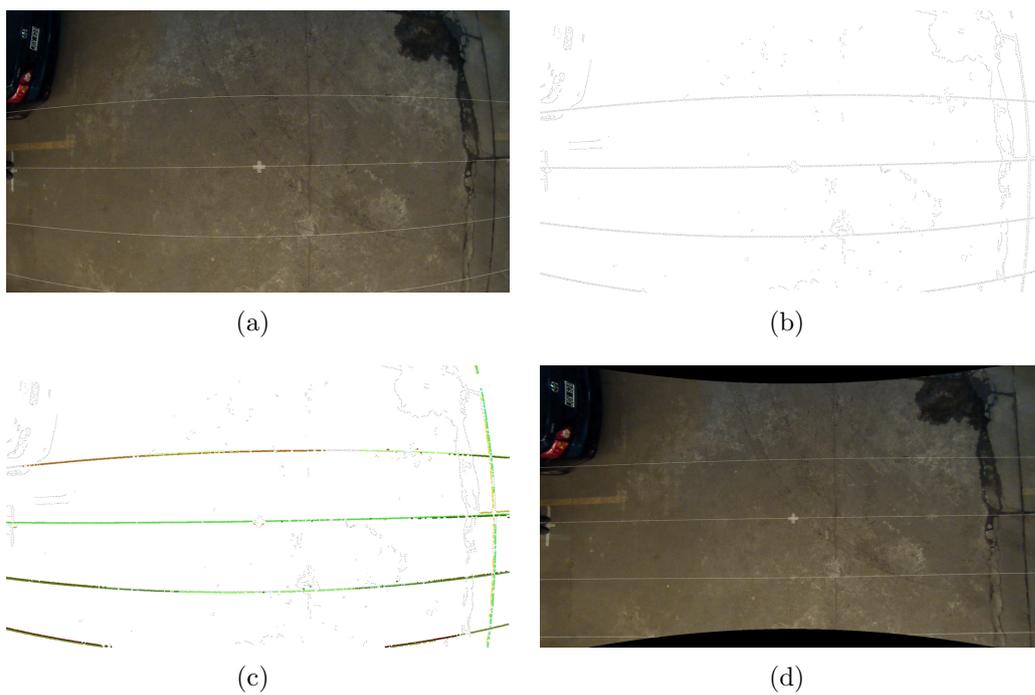


Figura 7.5: Resultados de la aplicación de los métodos descritos a la corrección de la distorsión de un *frame* de una secuencia en la que aparecen cables de referencia en posición horizontal: (a) imagen original, (b) bordes detectados, (c) primitivas detectadas y (d) imagen resultante de la corrección.

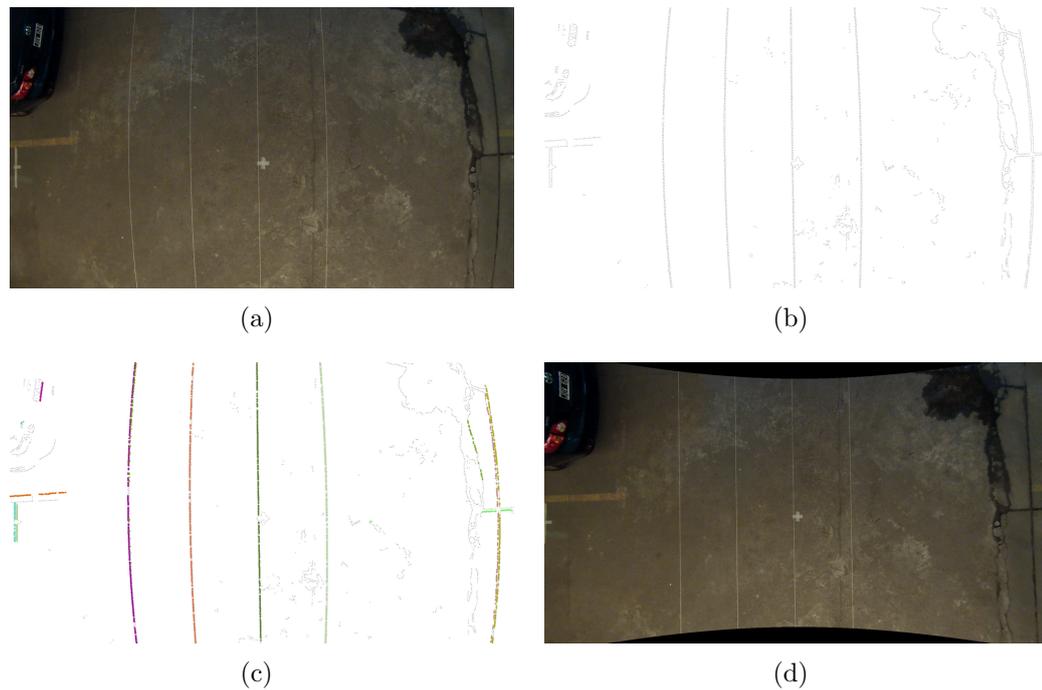


Figura 7.6: Resultados de la aplicación de los métodos descritos a la corrección de la distorsión de un *frame* de una secuencia en la que aparecen cables de referencia en posición vertical: (a) imagen original, (b) bordes detectados, (c) primitivas detectadas y (d) imagen resultante de la corrección.



Figura 7.7: Imagen de dos sujetos cruzándose mientras llevan los gorros usados como referencia para el proceso de seguimiento de peatones.



# Capítulo 8

## Conclusiones y trabajo futuro

### 8.1. Conclusiones

Al principio de esta tesis, hemos descrito como la propia geometría de la lente, así como la modificación de la distancia focal a través del mecanismo de *zoom*, produce una distorsión óptica que con frecuencia es necesario corregir. Como consecuencia de esta distorsión, las líneas rectas de la escena real 3D se proyectan en el plano imagen 2D de la captura en forma de líneas distorsionadas. Esto tiene un efecto adverso en multitud de aplicaciones que requieren una interpretación de la escena, como son por ejemplo: la realización de medidas en la imagen o la calibración de cámaras. Por lo tanto, se hace necesaria la detección, estimación y corrección de esta distorsión, como medio de proporcionar una entrada que se ajuste más a la escena real.

La interpretación de la escena capturada depende en gran medida de la capacidad de encontrar información significativa en la misma. En el caso concreto que nos ocupa, la distorsión de la lente, esta interpretación consiste en obtener primitivas que den cuenta del grado de distorsión presente, para de esta manera estimarla y poderla corregir. En ese sentido, técnicas como la transformada de Hough, que permite la detección de primitivas en una imagen, parecen una herramienta útil para este fin. Sin embargo, este método clásico no es capaz de abordar la detección de líneas rectas cuando éstas han sido distorsionadas. Debido a este efecto, tienden a ser detectadas como múltiples segmentos en lugar de como una línea completa.

En esta tesis hemos propuesto una transformada de Hough mejorada, mediante

la inclusión de la distorsión, haciendo referencia tanto a modelos de tipo polinomial como a los de tipo división. Con ello, se introduce una nueva dimensión en el espacio de votación de esta técnica, de forma que una línea viene representada por tres parámetros: ángulo, distancia al origen y distorsión. De esta manera, partiendo de los contornos detectados en la imagen de entrada, se consigue obtener el conjunto de primitivas que mejor se ajustan a las líneas distorsionadas en la imagen, así como un valor inicial del parámetro de distorsión. Con el objetivo de refinar dicho valor, posteriormente se introduce un proceso de optimización no lineal, mediante el cual se reduce la distancia cuadrática media de los puntos de las primitivas corregidas a sus rectas asociadas.

Sin embargo, cuando el grado de distorsión es alto, los modelos de distorsión de un parámetro no son capaces de abordar por completo el problema. Es por ello que proponemos un nuevo proceso de optimización no lineal. Tras aplicar la transformada mejorada descrita, se realiza una optimización en la que se incluye un segundo parámetro de distorsión, además del centro. Asimismo, hemos introducido también un proceso de optimización iterativa, mediante el cual se consigue mejorar el modelo de distorsión obtenido, así como el conjunto de primitivas asociadas al mismo. De esa forma, se consiguen detectar más puntos que dan lugar a líneas más largas, las cuales proporcionan información más significativa acerca de la distorsión presente en la imagen.

La corrección de la distorsión precisa de la obtención del modelo inverso, de forma que a través de él se obtenga la información necesaria para realizarla. Por esta razón, también hemos estudiado la invertibilidad de los modelos de distorsión polinomiales y de división de dos parámetros. Para ello, hemos caracterizado matemáticamente las condiciones bajo las cuales los parámetros de distorsión del modelo definen una transformación biyectiva o uno a uno. Con esto, se asegura que se puede calcular el modelo inverso de distorsión, añadiendo fiabilidad y robustez a las técnicas propuestas. Sin embargo, existe la limitación de que esta técnica no funcionaría en escenarios con modelos de distorsión no invertibles, como en el caso de lentes ojo de pez, donde la imagen se encuentra inscrita dentro de un círculo rodeado por un fondo negro.

Hasta ese momento, la bondad de la corrección de la distorsión realizada, se medía en términos del resultado obtenido, bien mediante la propia corrección definida como un porcentaje de variación del radio, a través del número de primitivas y puntos detectados, y más concretamente por la distancia promedio final entre los puntos corregidos de las primitivas y sus líneas asociadas. Con el objetivo de introducir un mejor criterio, se propone una nueva medida, a través de la estimación de modelos de distorsión mediante la minimización del error de reproyección

usando patrones de calibración. Para ello, se calcula una homografía, que permite obtener la correspondencia entre la imagen del patrón y su modelo sintético a través de una transformación en el plano. Con ello, se calculan las correspondencias entre las primitivas distorsionadas detectadas mediante las técnicas anteriormente descritas y las rectas del modelo de referencia. Posteriormente se realiza una minimización a través del método de Levenberg–Marquardt, lo cual permite mejorar el modelo de distorsión de la lente y minimizar el error de reproyección. De esta manera, se obtiene un criterio más preciso y robusto que la mera rectificación de líneas distorsionadas en la imagen. Asimismo, el método propuesto puede ser aplicado a imágenes afectadas por una fuerte distorsión, en las cuales los métodos tradicionales fallan.

Además de los métodos descritos, también hemos visto una serie de aplicaciones de los mismos. La primera de ellas es la detección de las esquinas de los cuadrados en un patrón de calibración de Zhang. Como parte previa al proceso de calibración de la cámara, se hace necesario la detección de esas características en la imagen de entrada. En el caso de usar métodos tradicionales sobre imágenes con una fuerte distorsión, estos fallan, ya que la posición real está desplazada respecto a la esperada. La técnica propuesta incluye la detección de las primitivas usando los métodos descritos y obteniendo las esquinas como intersecciones de las líneas distorsionadas detectadas. El conjunto final de las esquinas se obtiene mediante la intersección entre las obtenidas con nuestra metodología y las proporcionadas por el método de Harris. Como consecuencia, la aplicación descrita aporta información más significativa para el proceso de calibración de la cámara, puesto que proporciona un número mayor de esquinas y su localización es precisa.

La segunda de las aplicaciones descritas se vincula con la corrección de secuencias de vídeo. En el escenario descrito, se pretende comparar un sistema de tres cámaras dotadas con lentes convencionales, frente a una única cámara con lente gran angular que cubre todo el terreno de juego. Con el objetivo de medir parámetros cinemáticos de los jugadores, se hace necesaria una corrección de la distorsión, de forma que los resultados sean lo más precisos posible. Para ello se propone una herramienta que parte del primer fotograma, sobre el cual se estima el modelo de distorsión de la lente, aplicando las técnicas descritas a lo largo de la presente tesis. Debido a que no existe una modificación de la distancia focal, y que durante toda la secuencia la cámara se encuentra enfocando a una misma posición, no es necesario realizar más estimaciones del modelo a lo largo del vídeo. Por lo tanto, a partir del modelo estimado usando sólo un fotograma, es posible corregir la secuencia completa, proporcionando un vídeo libre de distorsión de la lente. Comparativamente, los resultados descritos presentan que los valores de los parámetros obtenidos a través de esta secuencia grabada con una lente gran angu-

lar, proporcionan una estimación similar al sistema más complejo. Además de ello, esto radica también en la simplificación de costes, desde el punto de vista del material usado para la captura y su posterior almacenamiento, así como una reducción significativa en el tiempo de postproceso. Asimismo, se introduce otra aplicación relativa a secuencias de vídeo, en la cual se pretende realizar el seguimiento de peatones en la calle.

En conclusión, el trabajo realizado en esta tesis aporta un cierto número de mejoras. La inclusión de la distorsión en el espacio de votación de la transformada de Hough, proporciona un medio por el cual es posible detectar líneas rectas que se encuentran distorsionadas en la imagen de entrada. Esto nos permite estimar la distorsión y con ello poder corregirla. En aquellos casos en los que se cuenta con un grado de distorsión alto, se ha abordado el problema mediante la introducción de un segundo parámetro de distorsión e incluyendo en el proceso de optimización el centro de distorsión de la lente. Asimismo, con el objetivo de obtener una mejor información de la distorsión presente en la imagen, se ha añadido un proceso de optimización iterativa, a través del cual se mejora tanto el modelo como el conjunto de primitivas asociado. La inclusión de la caracterización matemática de las condiciones bajo las cuales los parámetros de distorsión definen una aplicación biyectiva, ha dotado de robustez a las técnicas propuestas.

Los beneficios de los métodos descritos se hacen patentes en las aplicaciones descritas en los capítulos 6 y 7. Tanto en el caso de la detección de las esquinas en el patrón de calibración, como en el de la corrección de la distorsión en una secuencia de vídeo, las técnicas aplicadas proporcionan información significativa para fases posteriores. Desde el punto de vista de la calibración de cámaras, la información obtenida a través del método propuesto es más significativa, en número y en precisión de la detección. Asimismo, los procedimientos de *tracking* y el estudio cinemático de movimientos, se ven claramente beneficiados cuando se corrige la distorsión de la lente presente en la secuencia, dando resultados similares a sistemas multicámara más complejos.

Por fortuna, queda mucho trabajo por hacer. Por ello, en la sección siguiente se describirán algunas de las posibles líneas futuras a tener en cuenta.

## 8.2. Trabajo futuro

Considerando que el trabajo realizado se enmarca dentro de uno de los problemas clásicos en Visión por Computador, cuyo correcto modelado es necesario como medio de proporcionar información significativa en diversas aplicaciones, el conjunto de posibles líneas futuras a desarrollar es variado, tanto desde el punto de vista de las mejoras de las técnicas propuestas como de su aplicabilidad en diversos problemas.

- **Calibración de cámaras:** la calibración de cámaras consiste en obtener los parámetros que relacionan puntos 3D en la escena real con puntos 2D en una imagen. En este proceso se realiza previamente el cálculo de una homografía, que es una deformación en el plano que pone en correspondencia dos imágenes, la fotografía real del patrón y su modelo sintético. Generalmente el proceso de calibración se realiza usando imágenes de patrones usados para dicho fin como referencia, bien mediante diversas vistas de un mismo patrón plano o usando un patrón con dos o tres planos ortogonales. En esta tesis hemos descrito el uso de la minimización del error de reproyección para la estimación de modelos de distorsión de la lente en el capítulo 5. Para ello, se partía de las técnicas descritas en los capítulos 3 y 4. Asimismo, hemos presentado la aplicación de esas técnicas a la obtención precisa de las esquinas de los cuadrados en un patrón de calibración Zhang en el capítulo 6. Sin embargo, no se ha realizado por completo la calibración de la cámara, obteniendo sus parámetros característicos, que describen propiedades como la posición y orientación de la cámara en el espacio. Una de las posibles líneas futuras a llevar a cabo es aplicar los métodos descritos para realizar una calibración automática completa de la cámara. Para ello, sería necesario realizar este proceso partiendo de dos o más vistas distintas de un mismo patrón de calibración plano, que permitieran posteriormente obtener la matriz de proyección. La ventaja principal que ofrecerían las técnicas propuestas radican en que la detección, estimación y corrección de la distorsión se realizan de forma completamente automática, evitando cualquier tipo de interacción durante el proceso de localización de la información necesaria para el proceso de calibración. Asimismo, la introducción de la minimización del error de reproyección, mejoraría probablemente de forma significativa la calibración de la cámara respecto a otras técnicas usadas para dicho fin, especialmente cuando las diferentes imágenes de entrada del patrón de calibración se ven afectadas por una distorsión severa.
- **Inserción de objetos sintéticos en escenas reales. Realidad aumen-**

**tada:** la inserción de gráficos virtuales está cada vez más generalizada en multitud ámbitos, desde entornos deportivos, en los que se intenta analizar distintos parámetros de la escena o insertar publicidad, hasta aplicaciones industriales, donde se añade por ejemplo información extra relativa a piezas o procedimientos. Todo ello forma parte de lo que se denomina como realidad aumentada, que consiste en la visión de un entorno real en el que se han incorporado objetos virtuales. En aquellos casos en los que se precise una alta precisión en la inserción de los objetos, como por ejemplo que una determinada imagen o vídeo insertado se ajuste lo más posible a los bordes de un objeto de la escena real, es necesario considerar la distorsión de la lente. Por lo tanto, uno de los posibles escenarios en los que aplicar los métodos descritos, consistiría en hacer uso de las técnicas automáticas propuestas para modelar la distorsión de la lente del sistema que está visualizando la escena, previamente a la inserción de gráficos sintéticos en ella, proporcionando de esta manera una mayor precisión en este tipo de procedimientos.

- **Seguimiento de peatones:** durante la realización de la presente tesis, comenzó una colaboración con el Instituto de Tecnología de la Universidad Argentina de la Empresa. En dicha colaboración, se usaron las técnicas desarrolladas, con el objetivo de corregir secuencias de vídeo, sobre las que posteriormente se realizaría el seguimiento de peatones y se estudiarían características como sus trayectorias y colisiones o cruces. Dicha secuencia, al igual que en el caso expuesto a lo largo del capítulo 7, se realizó mediante el uso de una cámara provista de una lente gran angular. Dentro de las actividades futuras, se estima que esta colaboración seguirá su curso, dando lugar a una publicación conjunta entre el citado Instituto y el Grupo de Investigación de Análisis Matemático de Imágenes. En ella se llevarían a cabo la aplicación de las diferentes técnicas propuestas, a fin de proporcionar una fuente lo más precisa posible para el posterior estudio de la escena que se está grabando.
- **Corrección de secuencias de vídeo con variación de la distancia focal:** una de las aplicaciones descritas usando las técnicas propuestas en esta tesis, es la corrección de la distorsión en secuencias de vídeo. Sin embargo, estas secuencias son estáticas y carecen de variación en la distancia focal de la lente. Como se ha descrito, la distorsión se ve afectada por la profundidad del punto al que se enfoca, pero mucho más aún por la variación del *zoom*. Una posible extensión del trabajo realizado consistiría en aplicar las técnicas descritas a secuencias en las cuales existiera una variación de dichos parámetros. Con ello se proporcionaría una herramienta más genérica que en los casos que se han descrito. En este sentido, el seguimiento de primitivas

en la secuencia, así como del modelo y la estimación de la distancia focal, serían de utilidad.

- **Estimación de la distorsión de la lente en escenarios no invertibles:** una limitación descrita acerca de los métodos propuestos para detectar, estimar y corregir la distorsión, es que no funcionarían en escenarios en los que el modelo de distorsión de la lente no es invertible. La inclusión de los teoremas de invertibilidad, que añaden robustez, por contra limitan el funcionamiento en el caso de imágenes tomadas por ejemplo usando lentes tipo ojo de pez. Por ello, una posible línea futura es extender las técnicas a ámbitos más generales, en los cuales la formalización de los modelos y las condiciones bajo las cuales se define la transformación de la distorsión sean diferentes.
- **Corrección de la distorsión de la perspectiva:** además de la distorsión óptica, con frecuencia las imágenes presentan una distorsión de perspectiva generada por la orientación de la cámara respecto a la escena. Un ejemplo típico de ello son las fotografías de fachadas de edificios, donde la cámara no se puede situar de forma fronto-paralela, sino que es necesario posicionarla con una inclinación significativa para capturar la fachada completa. La corrección de la distorsión de la perspectiva se basa en la aplicación de una homografía en la imagen que corrige la posición de los puntos de fuga vertical y horizontal enviándolos al infinito para simular visualmente una vista fronto-paralela. Las técnicas desarrolladas en esta tesis permiten extraer en la imagen colecciones de líneas distorsionadas que pueden ser de gran utilidad para estimar de forma automática las posiciones de los puntos de fuga en la imagen y realizar la corrección de la distorsión de la perspectiva.



# Bibliografía

- [1] [www.mathworks.com/products/matlab](http://www.mathworks.com/products/matlab).
- [2] M. Alemán-Flores, L. Alvarez, L. Gomez, and D. Santana-Cedrés. Wide-angle lens distortion correction using division models. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 415–422. Springer Berlin Heidelberg, 2013.
- [3] M. Alemán-Flores, L. Alvarez, L. Gomez, and D. Santana-Cedrés. Line detection in images showing significant lens distortion and application to distortion correction. *Pattern Recognition Letters*, 36:261–271, 2014.
- [4] L. Alvarez, L. Gómez, and P. Henríquez. Zoom Dependent Lens Distortion Mathematical Models. *Journal of Mathematical Imaging and Vision*, 44(3):480–490, 2012.
- [5] L. Alvarez, L. Gomez, and J.R. Sendra. An algebraic approach to lens distortion by line rectification. *Journal of Mathematical Imaging and Vision*, 35(1):36–50, 2009.
- [6] L. Alvarez, L. Gomez, and J.R. Sendra. Algebraic Lens Distortion Model Estimation. *Image Processing On Line*, 1, 2010.
- [7] L. Alvarez, L. Gómez, and J.R. Sendra. Accurate Depth Dependent Lens Distortion Models: An Application to Planar View Scenarios. *Journal of Mathematical Imaging and Vision*, 39(1):75–85, 2011.
- [8] L. Alvarez and L. Mazorra. Signal and image restoration using shock filters and anisotropic diffusion. *SIAM Journal on Numerical Analysis*, 31(2):590–605, 1994.
- [9] D.H. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.

- [10] R.M.L. Barros, R.P. Menezes, T.G. Russomanno, M.S. Misuta, B.C. Brandão, P.J. Figueroa, N.J. Leite, and S.K. Goldenstein. Measuring handball players trajectories using an automatically trained boosting algorithm. *Computer methods in biomechanics and biomedical engineering*, 14(01):53–63, 2011.
- [11] R.M.L. Barros, M. S Misuta, R.P. Menezes, P.J. Figueroa, F.A. Moura, S.A. Cunha, R. Anido, and N.J. Leite. Analysis of the distances covered by first division brazilian soccer players obtained with an automatic tracking method. *Journal of sports science & medicine*, 6(2):233–242, 2007.
- [12] F. Bellard. FFmpeg. <https://www.ffmpeg.org>.
- [13] D.P. Born, H.C. Holmberg, F. Goernert, and B. Sperlich. A novel compression garment with adhesive silicone stripes improves repeated sprint performance - a multi-experimental approach on the underlying mechanisms. *BMC Sports Science, Medicine and Rehabilitation*, 6(1):21, 2014.
- [14] J.Y. Bouguet. [www.vision.caltech.edu/bouguetj/](http://www.vision.caltech.edu/bouguetj/).
- [15] F. Bukhari and M.N. Dailey. Robust radial distortion from a single image. In *Advances in Visual Computing*, pages 11–20. Springer, 2010.
- [16] F. Bukhari and M.N. Dailey. Automatic radial distortion estimation from a single image. *Journal of mathematical imaging and vision*, 45(1):31–45, 2013.
- [17] J. Cai. Automatic lens distortion correction using single images by finding reliable lines for rectification. In *Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on*, pages 1015–1020. IEEE, 2014.
- [18] J. Cai and S. Miklavcic. Automatic curve selection for lens distortion correction using hough transform energy. In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, pages 455–460. IEEE, 2013.
- [19] J. Canny. A Computational Approach to Edge Detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, Nov 1986.
- [20] H.L. Chung, Z.W. Lin, S.T. Huang, Y.H. Huang, and H.Y.M. Liao. New orientation-based elimination approach for accurate line-detection. *Pattern Recognition Letters*, 31(1):11–19, 2010.
- [21] F.M. Clemente, M.S. Couceiro, F.M. Martins, A.J. Figueiredo, and R.S. Mendes. Análise de jogo no Futebol: Métricas de avaliação do comportamento coletivo. *Motricidade*, 10(1):14–26, 2014.

- 
- [22] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani. A Hough transform-based method for radial lens distortion correction. In *Image Analysis and Processing, 2003. Proceedings. 12th International Conference on*, pages 182–187. IEEE, 2003.
- [23] A. Desolneux, L. Moisan, and J.M. Morel. Meaningful Alignments. *International Journal of Computer Vision*, 40(1):7–23, 2000.
- [24] F. Devernay and O. Faugeras. Straight lines have to be straight. *Machine vision and applications*, 13(1):14–24, 2001.
- [25] R.O. Duda and P.E. Hart. Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Commun. ACM*, 15(1):11–15, jan 1972.
- [26] L.A.F. Fernandes and M.M. Oliveira. Real-time line detection through an improved hough transform voting scheme. *Pattern Recognition*, 41(1):299 – 314, 2008.
- [27] A.W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In: *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 125–132, 2001.
- [28] C.S. Fraser and M.R. Shortis. Variation of distortion within the photographic field. *Photogrammetric Engineering and Remote Sensing*, 58(6):851–855, 1992.
- [29] M. Gadermayr, M. Liedlgruber, A. Uhl, and A. Vécsei. Evaluation of different distortion correction methods and interpolation techniques for an automated classification of celiac disease. *Computer methods and programs in biomedicine*, 112(3):694–712, 2013.
- [30] M. Gasser, B. Habegger, J. Goette, and M. Jacomet. Feature Extraction from Biological Motion with PARTwear An Application for Sprint Running. In *3rd International Conference on Ambulatory Monitoring of Physical Activity and Movement*, June 2013.
- [31] R. Gonzalez. Fast line and circle detection using inverted gradient hash maps. In *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, pages 1354–1358. IEEE, 2015.
- [32] D. González-Aguilera, J. Gómez-Lahoz, and P. Rodríguez-Gonzálvez. An automatic approach for radial lens distortion correction from a single image. *Sensors Journal, IEEE*, 11(4):956–965, 2011.
-

- [33] R. Grompone, P. Monasse, J.M. Morel, and Z. Tang. Towards high-precision lens distortion correction. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 4237–4240. IEEE, 2010.
- [34] R. Grompone, P. Monasse, J.M. Morel, and Z. Tang. Lens distortion correction with a calibration harp. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 617–620. IEEE, 2011.
- [35] R. Grompone von Gioi, J. Jakubowicz, J.M. Morel, and G. Randall. On Straight Line Segment Detection. *Journal of Mathematical Imaging and Vision*, 32(3):313–347, 2008.
- [36] R. Grompone von Gioi, J. Jakubowicz, J.M. Morel, and G. Randall. LSD: a Line Segment Detector. *Image Processing On Line*, 2:35–55, 2012.
- [37] D.S. Guru, B.H. Shekar, and P. Nagabhushan. A simple and robust line detection algorithm based on small eigenvalue analysis. *Pattern Recognition Letters*, 25(1):1–13, 2004.
- [38] C. Harris and M. Stephens. A combined corner and edge detector. In *The 4th Alvey Vision Conference*, volume 15, pages 147–151, 1988.
- [39] R.I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [40] M. Higuchi, A. Datta, and T. Kanade. Software Package for Precise Camera Calibration. [http://www.ri.cmu.edu/research\\_project\\_detail.html?project\\_id=617&menu\\_id=261](http://www.ri.cmu.edu/research_project_detail.html?project_id=617&menu_id=261).
- [41] J.P. Hoover Jr. *Differences in strike index between land treadmill and aquatic treadmill running in experienced distance runners*. PhD thesis, UTAH STATE UNIVERSITY, 2014.
- [42] P.V.C. Hough. Machine Analysis Of Bubble Chamber Pictures. In *Proceedings, 2nd International Conference on High-Energy Accelerators and Instrumentation, HEACC 1959*, volume C590914, pages 554–558, 1959.
- [43] P.V.C. Hough. Method and means for recognizing complex patterns, December 1962.
- [44] C. Hughes, M. Glavin, E. Jones, and P. Denny. Review of geometric distortion compensation in fish-eye cameras. In *Signals and Systems Conference, 208. (ISSC 2008). IET Irish*, pages 162–167, Galway, Ireland, 2008.

- 
- [45] J.G. Jiménez. *Visión por computador*. Ingeniería (Paraninfo). Paraninfo, 1999.
- [46] S.B. Kang. Radial distortion snakes. *IEICE Transactions on Information and Systems*, 84(12):1603–1611, 2001.
- [47] J.M. Lavest, M. Viala, and M. Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In Hans Burkhardt and Bernd Neumann, editors, *Computer Vision — ECCV’98*, volume 1406 of *Lecture Notes in Computer Science*, pages 158–174. Springer Berlin Heidelberg, 1998.
- [48] T.Y. Lee, T.S. Chang, C.H. Wei, S.H. Lai, K.C. Liu, and H.S. Wu. Automatic distortion correction of endoscopic images captured with wide-angle zoom lens. *Biomedical Engineering, IEEE Transactions on*, 60(9):2603–2613, 2013.
- [49] Y.S. Lee, H.S. Koo, and C.S. Jeong. A straight line detection using principal component analysis. *Pattern Recognition Letters*, 27(14):1744–1754, 2006.
- [50] R. Lenz. Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standardobjektiven für hochgenaue 3D - Messungen in Echtzeit. In Erwin Paulus, editor, *Mustererkennung 1987*, volume 149 of *Informatik-Fachberichte*, pages 212–216. Springer Berlin Heidelberg, 1987.
- [51] K. Levenberg. A method for the solution of certain non-linear problems in least-squares. *Quarterly of Applied Mathematics*, II(2):164–168, 1944.
- [52] R. Martínez-Gallego, J.F. Guzmán, N. James, J. Pers, J. Ramón-Llin, and G. Vuckovic. Movement characteristics of elite tennis players on hard courts with respect to the direction of ground strokes. *Journal of sports science & medicine*, 12(2):275, 2013.
- [53] E. Morais, A. Ferreira, S.A. Cunha, R.M.L. Barros, A. Rocha, and S. Goldenshtein. A multiple camera methodology for automatic localization and tracking of futsal players. *Pattern Recognition Letters*, 39:21–30, 2014.
- [54] F.A. Moura, J.E. Santana, A.L. Marche, T.H. Aguiar, A.C.M.A. Rodrigues, R.M. de Barros, and S.A. Cunha. Quantitative analysis of futsal players’ organization on the court. 2(11):105–108, 2011.
- [55] A.I. Papadaki and A. Georgopoulos. Development, comparison, and evaluation of software for radial distortion elimination. In *SPIE Optical Metrology*, pages 95280C–95280C. International Society for Optics and Photonics, 2015.
-

- [56] J. Perš and S. Kovacic. Nonparametric, model-based radial lens distortion correction using tilted camera assumption. In *Proceedings of the Computer Vision Winter Workshop*, volume 1, pages 286–295, 2002.
- [57] J. Pers, M. Kristan, M. Perse, and S. Kovacic. Analysis of Player Motion in Sport Matches. In Arnold Baca, Martin Lames, Keith Lyons, Bernhard Nebel, and Josef Wiemeyer, editors, *Computer Science in Sport - Mission and Methods*, number 08372 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2008. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany.
- [58] M. Perše, M. Kristan, S. Kovačič, G. Vučkovič, and J. Perš. A trajectory-based analysis of coordinated team activity in a basketball game. *Computer Vision and Image Understanding*, 113(5):612–621, 2009.
- [59] J. Perš, M. Bon, S. Kovačič, M. Šibila, and B. Dežman. Observation and analysis of large-scale human motion. *Human Movement Science*, 21(2):295–311, 2002.
- [60] J.P. Pluijms, R. Cañal-Bruland, M.J.M. Hoozemans, and G.J.P. Savelsbergh. Visual search, movement behaviour and boat control during the windward mark rounding in sailing. *Journal of sports sciences*, 33(4):398–410, 2015.
- [61] M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. Thomson-Engineering, 2007.
- [62] Z. Tang, R. Grompone von Gioi, P. Monasse, and J.M. Morel. High-precision camera distortion measurements with a “calibration harp”. *Journal of the Optical Society of America A*, 29(10):2134–2143, 2012.
- [63] Z. Tang, R. Grompone von Gioi, P. Monasse, and J.M. Morel. Self-consistency and universality of camera distortion models. 2012. <https://hal-enpc.archives-ouvertes.fr/hal-00739516>.
- [64] L.H.P. Vieira, E.A. Pagnoca, F. Milioni, R.A. Barbieri, R.P. Menezes, L. Alvarez, L. Gomez, D. Santana-Cedrés, and P.R.P. Santiago. Tracking futsal players with a wide-angle lens camera: accuracy analysis of the radial distortion correction based on an improved hough transform algorithm. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pages 1–11, 2015.
- [65] G. Vučković, J. Perš, N. James, and M. Hughes. Measurement error associated with the SAGIT/squash computer tracking software. *European Journal of Sport Science*, 10(2):129–140, 2010.

- [66] A. Wang, T. Qiu, and L. Shao. A simple method of radial distortion correction with centre of distortion estimation. *Journal of Mathematical Imaging and Vision*, 35(3):165–172, 2009.
- [67] Z. Wang, H. Liang, X. Wu, Y. Zhao, B. Cai, C. Tao, Z. Zhang, Y. Wang, S. Li, F. Huang, et al. A practical distortion correcting method from fisheye image to perspective projection image. In *Information and Automation, 2015 IEEE International Conference on*, pages 1178–1183. IEEE, 2015.
- [68] H. Wildenauer and B. Micusik. Closed form solution for radial distortion estimation from a single vanishing point. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2013.
- [69] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
- [70] Z. Zhang. *Camera Calibration*. Emerging Topics in Computer Vision. Prentice Hall Professional Technical Reference, 2004.