

DESARROLLO DE HERRAMIENTAS PARA LA CALIBRACIÓN DE CÁMARAS CON DISTORSIÓN

TESIS

Que para obtener el grado de
MAESTRÍA EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA

presenta

Cuauhtémoc Gómez Suárez

Dr. Leonardo Romero Muñoz

Director de Tesis

Universidad Michoacana de San Nicolás de Hidalgo

Noviembre 2007

Resumen

Un problema que se presenta en las imágenes capturadas por las cámaras de amplia apertura, comúnmente utilizadas en robots móviles, es conocido como distorsión radial.

En esta tesis, la corrección de la distorsión radial se plantea como un problema de optimización no lineal y para resolverlo se usa el método de Levenberg-Marquardt. Para ello se utiliza un patrón formado por cuadros blancos con fondo negro y el problema se plantea como encontrar un conjunto de parámetros que minimice la distorsión de la imagen capturada por la cámara.

Esta tesis analiza varias formulaciones para resolver el problema y presenta una técnica adecuada para remover la distorsión radial aún en cámaras con lentes de muy amplia apertura, donde precisamente otros trabajos anteriores de tesis no lo resuelven.

Removida la distorsión radial, también se aborda en esta tesis el problema de la calibración global de una cámara; es decir, la determinación de una matriz que relaciona los puntos en el entorno tridimensional y el correspondiente punto en la imagen de la cámara.

Para la calibración de una cámara se utilizan las correspondencias entre las características de un patrón de calibración formado por dos planos ortogonales y su imagen. Se formula el problema como un sistema de ecuaciones homogéneo y se resuelve por medio de la descomposición en valores singulares.

Como resultado del presente trabajo, se desarrolló una herramienta para la calibración de una cámara con el fin de facilitar posteriores trabajos en el área de visión.

Abstract

A problem that appears in images captured by wide opening cameras, commonly used in mobile robots, is known as radial distortion.

In this thesis, radial distortion correction is considered as a nonlinear optimization problem, and it is solved using the Levenberg-Marquardt method. A black and white checkboard pattern is used to enable the set of parameters to be calculated such that the radial distortion is minimized.

This thesis analyzes several ways to solve the problem and it presents a technique to remove the radial distortion even with very wide opening cameras, where other fail to solve the radial distortion problem.

Once the radial distortion is removed, this thesis focuses on the global calibration problem. It consists on finding a matrix that relates the tridimensional features in the environment to the bidimensional points in the images.

For the camera calibration we use the features correspondences between a pattern built with two orthogonal planes and the image captured by the camera. The problem is formulated as an homogenous system of equations and it is solved by the singular values decomposition method.

As a result of this work, a camera calibration tool was developed, its purpose is to facilitate later works in the vision area.

Contenido

Resumen	III
Abstract	V
Contenido	VI
Lista de Figuras	XI
Lista de Algoritmos	XII
Lista de Símbolos	XV
Lista de Publicaciones	XVII
1. Introducción	1
1.1. Motivación	1
1.2. Antecedentes	3
1.3. Objetivo de la tesis	5
1.3.1. Alcance	5
1.3.2. Contribuciones	6
1.4. Organización del documento	6
2. Calibración de cámaras	9
2.1. Introducción	9
2.2. Proyección de perspectiva ideal	11
2.3. Formación de la imagen.	14
2.4. Calibración	17
2.5. Conclusiones	18
3. Corrección de la distorsión radial	19
3.1. Errores ópticos	19
3.2. Distorsión	21
3.2.1. Distorsión radial	23
3.3. Transformación proyectiva	25
3.3.1. Interpolación bilineal	28
3.4. Ajuste de un modelo de distorsión	30
3.4.1. Modelo D-U	33
3.5. Método de correspondencias de puntos	35

3.5.1. Selección de puntos característicos	37
3.6. Optimización no lineal	40
3.7. Corrección de la distorsión radial	40
3.8. Conclusiones	43
4. Calibración Global	45
4.1. Algoritmo de Transformación Lineal Directa (DLT)	46
4.1.1. Número mínimo de correspondencias	47
4.1.2. Sistema de ecuaciones	48
4.2. Conclusiones	50
5. Resultados experimentales	53
5.1. Introducción	53
5.1.1. Material utilizado	53
5.2. Corrección de distorsión radial	54
5.3. Calibración global	57
5.4. Conclusiones	58
6. Conclusiones y trabajos futuros	65
6.1. Conclusiones	65
6.2. Trabajos futuros	66
A. Optimización de la función de error total	67
A.1. Optimización no lineal	67
B. Manual de Usuario	75
B.1. Descripción	75
B.2. Requerimientos generales	75
B.3. Instalación General	76
B.4. Corrección radial	76
B.4.1. Requerimientos	76
B.5. Como ejecutar el programa	77
B.6. Calibración Global	78
B.6.1. Requerimientos	78
B.6.2. Como ejecutar el programa	78
B.7. Simulador del modelo de distorsión	79
B.7.1. Como ejecutar los programas	80
C. Normalización de puntos	83

<i>Contenido</i>	IX
D. Códigos	89
D.1. Corrección de la distorsión radial	89
D.2. Diseño de simulador del modelo de distorsión radial	90
D.3. Calibración Global	90
Referencias	93
Glosario	97

Lista de Figuras

2.1. Modelo de la cámara ideal (pinhole) [Forsyth y Ponce(2002)]	11
2.2. Modelo de cámara de perspectiva	13
2.3. Geometría de la cámara de pinhole	14
2.4. Marco de referencia global	16
3.1. Error de curvatura	20
3.2. Corrección de una imagen con distorsión radial	21
3.3. Distorsión radial y tangencial	22
3.4. Tipo de error tangencial que ocasiona la distorsión	23
3.5. Efectos de la distorsión radial	23
3.6. Modelado de la distorsión	24
3.7. Transformación proyectiva	25
3.8. Escalamiento de una imagen	27
3.9. El problema de interpolar el valor para $\mathbf{x}$	28
3.10. Interpolación bilineal	29
3.11. Efecto de la distorsión radial	31
3.12. Dos funciones diferentes para modelar la distorsión	34
3.13. Experimento del modelo de distorsión U-D	35
3.14. Imágenes utilizadas en la corrección de la distorsión	36
3.15. Objetivo en la corrección de imágenes	37
3.16. Correspondencias de puntos	39
4.1. Patrón de calibración	46
4.2. Proyección de un punto 3D sobre un punto 2D	47
5.1. Cámaras Videre Design	54
5.2. Patrones de calibración	55
5.3. Imagen tomada con la cámara	56
5.4. Corrección de la distorsión en imágenes	56
5.5. Error en la corrección de la distorsión radial	57
5.6. Experimentos de corrección de la distorsión radial	61
5.7. Experimentos de corrección de la distorsión radial	62

5.8. Corrección de la distorsión radial utilizando k_1 y k_2	63
5.9. Patrón de calibración 3D	63
B.1. Patrón de calibración	77
B.2. Interfaz de usuario que muestra el efecto de la corrección de la distorsión radial	82

Lista de Algoritmos

1.	Algoritmo de Optimización no lineal - GNLM	41
2.	Algoritmo DLT, para la estimación de una matriz $\mathbf{P}$ a partir de correspondencias de puntos de 3D a 2D [Hartley y Zisserman(2004)]	51

Lista de Símbolos

f	Distancia focal
O_c	Origen en el marco de referencia de la cámara
π	El plano de la imagen
O_i	El centro de la imagen
M_c	Marco de referencia de la cámara
M_g	Marco de referencia global
$\mathbf{R}_{3 \times 3}$	Matriz de rotación
$\mathbf{T}_{3 \times 1}$	Vector de traslación
$\mathbf{P}$	Matriz de calibración global
$\mathbf{X}_g^h$	Un punto $\mathbf{P}$ en el marco de referencia global, con coordenadas homogéneas
$\mathbf{x}_c^h$	Un punto $\mathbf{p}$ en el marco de referencia de la cámara, con coordenadas homogéneas
P^u	Un punto P sin distorsión
c_d	Centro de distorsión
P^d	Un punto P con distorsión
f_1	Función de distorsión radial que representa el modelo D-U
f_2	Función de distorsión radial que representa el modelo U-D
θ_1^d	Conjunto de parámetros que representan la distorsión radial en el modelo D-U
θ_2^d	Conjunto de parámetros que representan la distorsión radial en el modelo U-D
$(x^r, y^r)^T$	Coordenadas reales de un plano
$\mathbf{H}$	Homografía, matriz de transformación proyectiva
(x_e, y_e)	Parte entera de las coordenadas de un punto de la imagen transformada
(x_f, y_f)	Parte fraccional de las coordenadas de un punto de la imagen transformada
$(x_u', y_u')^T$	Coordenadas de un punto en la representación de una imagen sin distorsión
$(x_d', y_d')^T$	Coordenadas de un punto en la representación de una imagen

	con distorsión
c_d	Centro de distorsión radial
r_u	Distancia radial entre un punto P^u y el centro de distorsión radial c_d
$(c_x, c_y)^T$	Coordenadas del centro de distorsión radial
k_1, k_2 y k_3	Parámetros que representan la distorsión radial
$\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$	Conjunto de parámetros que denotan la distorsión radial y proyección de la imagen
θ^d	Conjunto de parámetros que denotan la distorsión radial
θ^p	Conjunto de parámetros que denotan la proyección
I_r	Imagen de referencia
I_d	Imagen tomada por la cámara
r_j	Número de píxeles en blanco, que componen el j-ésimo cuadro blanco en una imagen
$x(P_{ij})$	es la coordenada x del punto P_i de un pixel blanco correspondiente al cuadro blanco j
$\mathbf{X}_i$	Es un punto de la escena con el marco de referencia global
$\mathbf{x}_i$	Es un punto de la imagen con el marco de referencia de la cámara
$\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$	Correspondencia de puntos
$\mathbf{p}^{jT}$	j-ésima fila de la matriz $\mathbf{Q}$
$\mathbf{A}_i$	Matriz de 3×12 que representa las ecuaciones de una correspondencia de puntos entre la escena y una imagen
$\mathbf{P}_v$	Es un vector columna que tiene 12 elementos, representado en una matriz $\mathbf{P}$
$\mathbf{T}$	Matriz de transformación de similaridad para normalizar los puntos de la imagen
$\mathbf{U}$	Matriz de transformación de similaridad para normalizar los puntos de la escena

Lista de Publicaciones

Leonardo Romero and Cuauhtemoc Gomez. Correcting Radial Distortion of Cameras With Wide Angle Lens Using Point Correspondences. A. Gelbukh, S. Torres, I. Lopez (Eds.) Advances in Computer Science and Engineering. Research in Computing Science 19, 2006, pp. 185-196, IPN, Mexico. ISSN: 1665-9899.

Leonardo Romero and Cuauhtémoc Gómez G. Book Chapter Obinata and A. Dutta (Eds.) Vision Systems. The book is scheduled to be published in April of 2007 by Advanced Robotic Systems. Vienna, Austria, Europe. ISBN 979-902613-06-6.

Capítulo 1

Introducción

Dentro de la robótica móvil el uso de cámaras es muy importante. Para ello es necesario llevar a cabo una calibración de las cámaras como un paso esencial antes de su uso. La calibración consiste en obtener una matriz, llamada matriz de calibración, por medio de la cual se obtienen correspondencias de puntos del espacio tridimensional (3D) con puntos asociados en las imágenes. Sin embargo, generalmente las cámaras utilizadas en robots móviles son cámaras que tienen como característica una apertura de vista amplia, requisito necesario para que el robot detecte los objetos a su alrededor. No obstante, dichas cámaras presentan imágenes con cierto grado de distorsión o deformación.

Esta tesis se enfoca en la calibración de una cámara con lente de vista amplia, abordando también la corrección de la distorsión debida al lente.

1.1. Motivación

Los humanos reciben información a través de su vista y con dicha información son capaces de adaptarse al ambiente. De manera análoga los robots por medio de las cámaras como su sistema de visión, procesan imágenes y toman decisiones de acuerdo a éstas.

Una imagen es una manera de representar datos en un formato donde hay una correspondencia directa entre la escena y su representación [Murphy(2000)].

Las cámaras que se usan en robótica son dispositivos que trabajan con imágenes no comprimidas a diferencia de las cámaras de video. Este tipo de cámaras son utilizadas en aplicaciones científicas, industrial, microscopía, cámaras web y aplicaciones fotográficas, entre otras.

Podemos pensar que un elemento sensible a la luz, opera como un pequeño capacitor que se carga inicialmente. Al incidir la luz en el elemento sensible, se liberan electrones que son retenidos por los campos eléctricos dentro del elemento sensible. De este modo la cantidad de carga retenida está relacionada con la cantidad de luz incidente. El elemento sensible a la luz en una foto tomada por la cámara se le conoce como *pixel* (Picture Element), para abreviar utilizaremos pixel al referirnos a dicho elemento sensible.

Enseguida se enuncian las diferencias entre la tecnología *CCD* y *CMOS*, las dos tecnologías más comunes empleadas en la fabricación de cámaras, tomando en cuenta la referencia [Siegwart y Nourbakhsh(2000)].

- En la tecnología CCD las cargas de cada pixel se miden “en serie”. Esto es, el proceso de medición se lleva en una sola parte del chip y a ese lugar se transportan las cargas de cada pixel en secuencia.
- La tecnología CMOS es más nueva y su operación es similar a la tecnología CCD. La diferencia central es la forma de medición de la carga de cada pixel. A diferencia de la anterior, cada pixel tiene un pequeño circuito para medir su carga, habilitando una lectura en paralelo. Las cámaras CMOS son más económicas y consumen menor energía que las cámaras CCD. Sin embargo las cámaras CCD están más desarrolladas y tienen mejores desempeños. Se espera que las cámaras CMOS, alcancen rápidamente el nivel de las cámaras CCD, con la ventaja de ser más económicas [Siegwart y Nourbakhsh(2000)].

Las cámaras son deseables en el uso de robots móviles, por ser sensores pasivos (no emiten energía) y por que es posible tener una visión amplia del ambiente. En [Kortenkamp et al.(1998)] citan el uso de varios robots en ambientes reales, entre ellos “POLLY”, un robot que se utilizó como guía de personas en un piso de un edificio.

Generalmente un robot móvil sensa, planea y actúa [Kortenkamp et al.(1998)]. Para sensar el ambiente, el robot “POLLY” utiliza una cámara y con ella es capaz de deambular por el piso gracias a un mapa previamente grabado en la memoria del robot, así como las restricciones del ambiente. En particular “POLLY” asume que todos los obstáculos descansan sobre el piso y la textura del piso es única.

En el proyecto “POLLY” se omite la distorsión en cámaras, dado que se utilizan imágenes muy pequeñas, lo suficiente como para obtener la información que se ocupa. En la actualidad el procesar imágenes con una mayor resolución y con un mayor ángulo de visión ayuda a tener una mejor percepción del ambiente.

Uno de los objetivos que se persigue en esta tesis es el desarrollo de un software que sirva como una herramienta para que estudiantes de nivel maestría puedan utilizar cámaras en sus estudios de visión computacional.

1.2. Antecedentes

En la mayoría de la literatura en materia de visión computacional, la corrección de la distorsión se omite, debido a que las cámaras que se utilizan tienen poca o nula distorsión.

La comparación entre 3 modelos de distorsión descrita en [Ma et al.(2003)], analiza cual modelo es más simple y por ende más rápido y eficiente. Sin embargo presenta experimentos con cámaras que generan poca distorsión.

Anteriormente en la tesis de Maestría de Adrián Núñez [Núñez(2004)] se elaboró un software en el cual se corrige la distorsión radial en cámaras con distorsión media. En el capítulo 5 se prueba dicho software con cámaras equipadas con lentes de muy amplia apertura sin conseguir corregir la distorsión.

En el artículo [Romero y Calderón(2003)] se utilizan dos métodos para la corrección de la distorsión: correspondencia de imágenes y correspondencia de puntos característicos. En las pruebas realizadas a ambos métodos se observa que la correspondencia de puntos muestra mejores resultados. Este método se basa en la minimización de una función de error entre dos conjuntos de puntos; por un lado los puntos que pertenecen a un patrón de calibración y por otro los correspondientes puntos en la

imagen obtenida por la cámara. El objetivo es obtener un conjunto de parámetros de proyección conjuntamente con los parámetros de distorsión, con los cuales se logre empatar las dos imágenes: la distorsionada y una sin distorsión. Sin embargo en [Romero y Calderón(2003)] tampoco muestra pruebas que demuestren la eficiencia en imágenes con distorsión elevada.

Existe otro método presentado en [Romero y Calderón(2004)] utilizando una forma diferente de corrección. En él se describe la diferencia entre corrección de la distorsión paramétrica y no paramétrica. Los métodos paramétricos se basan en obtener un modelo con un conjunto pequeño de parámetros para corregir la distorsión y dicho modelo se utiliza para corregir toda la imagen. En los métodos no paramétricos, no se tiene un pequeño conjunto de parámetros sino que se tiene uno o varios parámetros por cada pixel de la imagen. El método de corrección no paramétrica presenta problemas en donde no es clara la asociación entre imágenes. Sin embargo puede ser útil para modelar otro tipo de distorsiones de los lentes, sobre todo de los de bajo costo.

En el artículo presentado en [Romero et al.(2004)] se corrigen las imágenes con el método utilizado en [Romero y Calderón(2003)]. Se utilizan los puntos detectados por un telémetro láser en forma de valores iniciales de disparidad para un sistema de visión estéreo trinocular.

Otro método más para la corrección de la distorsión radial en las cámaras es el que presenta en [Benhimane y Malis(2004)] el cual se basa en dos imágenes: una tomada con una cámara con parámetros de distorsión desconocidos, mientras la segunda es tomada con una cámara sin distorsión. Los puntos son utilizados para calcular las variaciones que existen entre los parámetros que se obtienen en un modelo de una cámara ideal y los parámetros obtenidos con un modelo de distorsión radial estándar. Los puntos escogidos son esquinas.

Un ejemplo más de calibración de cámaras con la técnica de registro de imágenes es el que utiliza Toru Tamaki en su artículo [Tamaki et al.(2000)], el método está basado en la técnica de registro de imágenes el cual utiliza todos los puntos de una imagen y sus correspondencias en el espacio. El método es dividido en tres etapas: *transformación afin*, *transformación proyectiva planar* y eliminación de la distorsión en los

lentes. Para esto se utiliza una imagen patrón libre de distorsión y su correspondiente imagen tomada por la cámara.

1.3. Objetivo de la tesis

Esta tesis aborda el desarrollo de una herramienta de software que calibre cámaras digitales equipadas con lentes de amplia apertura. Se contemplan los casos de poca, mediana o mucha distorsión radial, ocasionada por los lentes, enfatizando el hecho de que la calibración sea robusta, precisa y requiera poca intervención humana.

En particular se plantearon los siguientes objetivos:

1. Comparar el desempeño de algoritmos de corrección de la distorsión eligiendo aquel que tenga mejores desempeños con poca, mediana y alta distorsión radial.
2. Desarrollar una herramienta robusta y precisa, para la determinación de los parámetros del modelo de distorsión elegido en el objetivo anterior.
3. Desarrollar una biblioteca de funciones para llevar a cabo rápidamente la corrección de imágenes adquiridas por una o múltiples cámaras.
4. Desarrollar un simulador didáctico para apreciar el desempeño del modelo de distorsión radial elegido, permitiendo variar fácilmente los parámetros del modelo y observando su efecto inmediato en las imágenes.
5. Desarrollar una herramienta para la determinación de la matriz de calibración, utilizando el procedimiento basado en la descomposición de valores singulares descrito en [Hartley y Zisserman(2004)].

1.3.1. Alcance

El software desarrollado será en el sistema operativo Linux, usando “C” como lenguaje de desarrollo.

La adquisición de imágenes digitales se basa en el proyecto IEEE Linux1394 para LINUX, el cual es un conjunto de controladores para utilizar el *bus serial firewire* para

la transferencia de imágenes de la cámara a la computadora [Bombe y Collins(2000)]. Las bibliotecas que se utilizan para la comunicación entre las cámaras y los controladores son las que se encuentran en la biblioteca libdc1394 [Dennedy et al.(2000)].

Para el manejo de vectores, matrices y operaciones del álgebra lineal que se requieren se utiliza GNU ¹ Scientific Library (GSL) [Galassi et al.(2004)]. GSL es una biblioteca de dominio público bastante madura que se utiliza en Linux y que normalmente se incluye dentro del Sistema Operativo.

1.3.2. Contribuciones

- Se logró el desarrollo de un software para la calibración de cámaras.
- Se desarrolló una librería para la corrección de imágenes con distorsión. La corrección se basa en el uso de un patrón de cuadros blancos y fondo negro, y la imagen distorsionada de dicho patrón que adquiere la cámara.
- Se desarrolló un simulador gráfico del modelo de distorsión radial, con fines didácticos. Apoyándose en la herramienta de desarrollo *GTK+*, herramienta que sirve para crear interfaces gráficas para usuarios.
- Se compararon dos modelos de distorsión, encontrándose ventajas y desventajas en cada uno de ellos.

1.4. Organización del documento

Esta tesis se divide en 4 partes:

- El capítulo 2 explica la calibración de las cámaras.
- El capítulo 3 presenta los métodos de corrección de distorsión radial.
- El capítulo 4 aborda el tema de la calibración, partiendo de que tenemos una cámara que obtiene imágenes sin distorsión radial. La calibración consiste en

¹GNU.- Es un acrónimo recursivo que significa “GNU no es Unix”. Se refiere a aquel software de uso libre. Software que todos los usuarios pueden “ejecutarlo, copiarlo, modificarlo y distribuirlo”

obtener una matriz que mapea los puntos en el espacio 3D con puntos asociados en la imagen de la cámara.

- Finalmente se presentan apéndices donde se muestra el manual de usuario y la explicación de las partes más relevantes de código.

Capítulo 2

Calibración de cámaras

En general, el proceso de calibración de una cámara, puede verse como encontrar la función o correspondencia entre puntos del espacio tridimensional y puntos de la imagen capturada por la cámara.

En este capítulo se presentan las herramientas matemáticas y un modelo simple de una cámara para formular adecuadamente el problema de calibración de una cámara. También se revisan las diferentes técnicas o métodos de calibración descritos en la literatura especializada sobre este tema.

Por sencillez en la presentación, la corrección de la distorsión, que puede verse como una parte de la calibración de una cámara, se reserva para el capítulo 3.

2.1. Introducción

Las imágenes digitales son muy importantes en el área de visión computacional. Se pueden dividir en dos clases [Trucco y Verri(1998)]:

Las imágenes por intensidad. Codifican intensidades de luz adquiridas por las cámaras.

Las imágenes por rango. Codifican la distancia adquirida por sensores especiales como son los sonares y telémetros láser.

Las imágenes por intensidad, que son las que nos ocupan en esta tesis miden la cantidad de luz que afecta un dispositivo fotosensible. Además, la manera en que se guardan es una matriz bidimensional de números que representan el nivel de intensidad de luz adquirida [Ma et al.(2004)].

En la siguiente sección se inicia con un análisis de un modelo simplificado de cámaras. Las cámaras son un medio por el cual es posible adquirir imágenes. Sin embargo, existen otros tipos de dispositivos utilizados en la adquisición de imágenes: dispositivos de adquisición de imágenes por medio de rayos “X”, escáneres utilizados para la adquisición de imágenes a partir de negativos, etc.

Para entender el proceso de formación de imágenes es conveniente revisar los principales parámetros físicos que afectan la adquisición de imágenes por intensidad de acuerdo con la categorización que se menciona en [Trucco y Verri(1998)]:

- **Parámetros ópticos de los lentes.** Se refieren a parámetros de la óptica sensorial del sistema de visión. Entre ellos se encuentran: tipos de lentes, apertura angular y longitud focal.
- **Parámetros fotométricos.** Los parámetros fotométricos son aquellos relacionados con las condiciones de la luz y reflectancia del entorno. Así al obtener la imagen en la cámara, ésta se verá influenciada por factores como la dirección de la iluminación o por las propiedades de la reflectancia de las diferentes superficies que participan. Los parámetros fotométricos son externos a la cámara y por lo tanto no son objeto de la calibración. En adelante no se tomarán en cuenta en este documento.
- **Parámetros Geométricos.** Con ayuda de los parámetros geométricos es posible determinar la posición en la imagen en la cual un punto 3D será proyectado. Estos parámetros, además determinan:
 - El tipo de proyección.
 - La posición y orientación de la cámara en el espacio.
 - La distorsión involucrada en la formación de la imagen.

2.2. Proyección de perspectiva ideal

El modelo ideal para la adquisición de imágenes en cámaras, conocido como *modelo de "pinhole"* ¹[Hartley y Zisserman(2004)] el cual se muestra en la figura 2.1. Este modelo consiste en una caja hermética con un orificio por donde ingresan los rayos de luz, los cuales se proyectan en la parte posterior de la caja. Idealmente, el orificio debería ser tan estrecho para permitir únicamente el paso de un haz luminoso de diámetro infinitesimal.

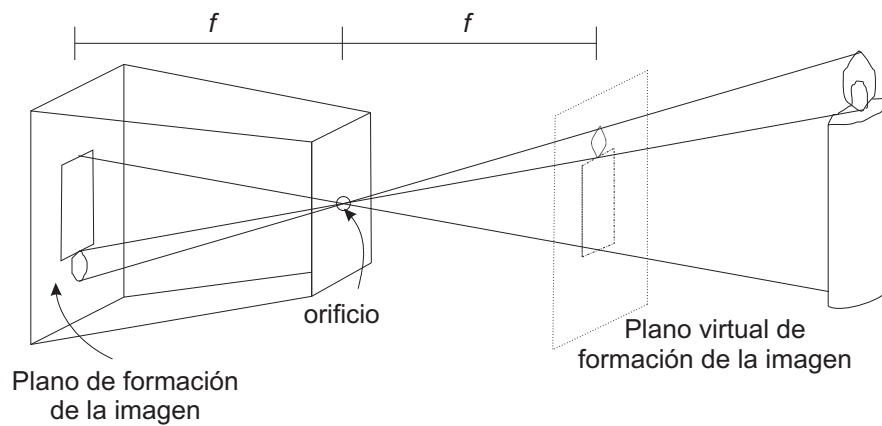


Figura 2.1: Modelo de la cámara ideal (pinhole) [Forsyth y Ponce(2002)]

En la figura 2.1 se hace referencia a una distancia f que va desde el orificio de la cámara hasta el plano donde se proyecta la imagen y se conoce como *distancia focal*. Para simplificar el modelo, en adelante se supone que la imagen se forma al frente del orificio, justo en un plano paralelo al de la imagen situado a una distancia f , como se ilustra en la figura 2.1. Visto de esta manera, podemos decir que la imagen se forma cuando el rayo que une un punto del ambiente con el orificio de la cámara, intersecta el plano virtual de formación de la imagen. A diferencia de la imagen real, la *imagen virtual* no está invertida.

En adelante cuando nos referimos al plano de la imagen se deberá entender el

¹Pinhole es una palabra de origen inglés que se puede traducir como *agujero hecho o como si fuera hecho por un alfiler*.

plano virtual de la imagen. También, el orificio de la cámara se le denotará como *centro de proyección* de la cámara o simplemente como centro de la cámara.

El objetivo de utilizar este modelo es ilustrar el principio de funcionamiento de la adquisición de imágenes en una cámara. A partir de él, es posible modelar matemáticamente la correspondencia entre puntos de la escena y puntos en la imagen. La figura 2.2 introduce el marco de referencia de la cámara. Sea el centro de la cámara O_c , el origen del marco de referencia de la cámara. El plano de la imagen π está situado a una distancia focal f , ortogonal al eje óptico (el eje Z) de la cámara. El centro de la imagen O_i , es la intersección del plano de la imagen con el eje óptico.

En la figura 2.2 se puede ver como un punto $\mathbf{X}$ con coordenadas de la escena $(X_c, Y_c, Z_c)^T$ con respecto del marco de referencia de la cámara se proyecta en el punto $\mathbf{x}$ de la imagen con coordenadas $(x_i, y_i)^T$ con respecto al marco de referencia de la imagen. De esta manera todos los rayos que parten de puntos de la escena y pasan por el centro de la cámara, se intersectan con el plano de la imagen formando la imagen virtual.

La figura 2.3 ilustra la manera de obtener la coordenada y_i de un punto en la imagen virtual a partir de un punto en la escena. Tenemos en este caso un punto $\mathbf{X}$ en la escena y su correspondiente punto $\mathbf{x}$ proyectado en la imagen virtual. De aquí es posible ver que el triángulo rectángulo que forman $\langle p, O_c, O_i \rangle$, es similar al triángulo rectángulo formado por $\langle P, O_c, Z_c \rangle$. La tangente de α es la misma en ambos triángulos. Esto es:

$$\frac{y_i}{f} = \frac{Y_c}{Z_c}$$

Despejando y_i , tenemos la coordenada deseada:

$$y_i = f \frac{Y_c}{Z_c} \quad (2.1)$$

De forma similar es posible encontrar la coordenada x_i del punto p en el plano π . Así tenemos que la coordenada x_i del punto p_i está dado por:

$$x_i = f \frac{X_c}{Z_c} \quad (2.2)$$

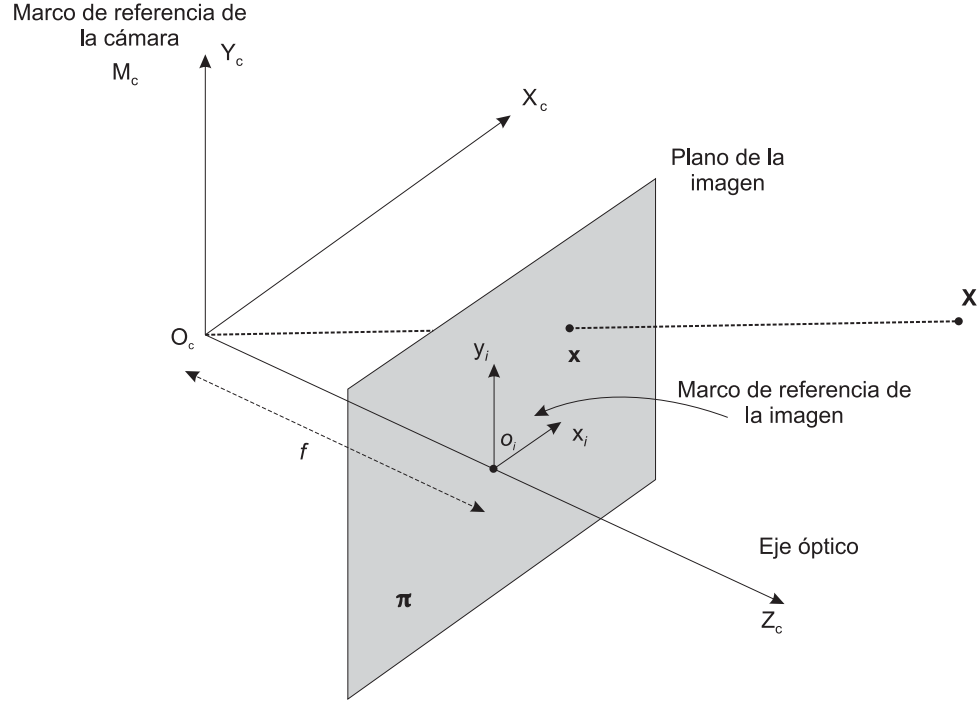


Figura 2.2: Modelo de cámara de perspectiva. El punto $\mathbf{X}$ de la escena se proyecta en el punto $\mathbf{x}$ de la imagen [Trucco y Verri(1998)].

En resumen, si un punto $\mathbf{X}$ tiene coordenadas $\mathbf{X} = (X_c, Y_c, Z_c)^T$ relativas al marco de referencia de la cámara, y su correspondiente punto $\mathbf{x}$ en la imagen, con coordenadas $\mathbf{x} = (x_i, y_i)^T$ se dice que éstas están relacionadas por la llamada proyección de perspectiva ideal de la siguiente forma:

$$x_i = f \frac{X_c}{Z_c}, \quad y_i = f \frac{Y_c}{Z_c} \quad (2.3)$$

Enseguida se aborda una formulación matricial conveniente para el proceso de formación de imágenes en las cámaras.

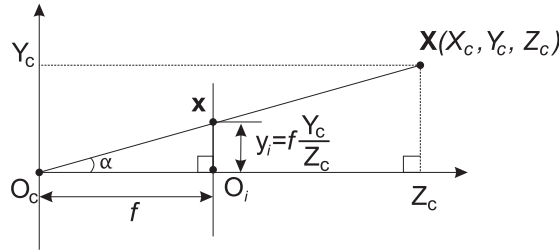


Figura 2.3: Geometría de la cámara de pinhole. Proyección de un punto en la escena con el plano de la imagen.

2.3. Formación de la imagen.

En la sección anterior se mencionó el modelo matemático con las ecuaciones (2.3), las cuales corresponden al modelo de una cámara de perspectiva ideal conocida como pinhole. Note que las ecuaciones (2.3) son no lineales debido al factor $\frac{1}{Z_c}$. Tampoco preservan la distancia entre puntos ni los ángulos entre líneas, sin embargo, mapean líneas en líneas [Trucco y Verri(1998)].

En el modelo de la cámara de pinhole que se explicó en la sección anterior (Figura 2.2) vemos que el punto $\mathbf{X}$ es proyectado en el plano de la imagen en el punto $\mathbf{x}$.

$$\mathbf{x} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \frac{f}{Z_c} \begin{bmatrix} X_c \\ Y_c \end{bmatrix} \quad (2.4)$$

Observemos que la ecuación 2.4 no es lineal, por el factor $\frac{1}{Z_c}$. Es importante resaltar que el cambio del sistema de coordenadas a un sistema de coordenadas homogéneo tiene la gran ventaja de linealizar la ecuación 2.4 [Hartley y Zisserman(2004)]. Enseguida se explica brevemente el uso de las coordenadas homogéneas para simplificar las operaciones del modelo de perspectiva ideal. Un punto del plano $(x_i, y_i)^T$ (en coordenadas reales) equivale a otro punto $(x^h, y^h, w^h)^T$ en coordenadas homogéneas. Donde las coordenadas reales se relacionan con las *coordenadas homogéneas* de la siguiente manera:

$$x_i = \frac{x^h}{w^h}, \quad y_i = \frac{y^h}{w^h} \quad (2.5)$$

De igual manera un punto del espacio con coordenadas $(X, Y, Z)^T$ equivale a un punto $(X^h, Y^h, Z^h, W^h)^T$ en coordenadas homogéneas, y ambas se relacionan por:

$$X = \frac{X^h}{W^h}, \quad Y = \frac{Y^h}{W^h}, \quad Z = \frac{Z^h}{W^h} \quad (2.6)$$

El lector interesado en conocer más sobre geometría proyectiva y en particular sobre el uso de coordenadas homogéneas puede consultar [Hartley y Zisserman(2004)].

En coordenadas homogéneas la ecuación 2.4 puede ser escrita como:

$$\begin{bmatrix} x^h \\ y^h \\ w^h \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c^h \\ Y_c^h \\ Z_c^h \\ 1 \end{bmatrix} \quad (2.7)$$

En forma compacta tenemos:

$$\mathbf{x}^h = \mathbf{C}_i \mathbf{X}_c^h \quad (2.8)$$

donde $\mathbf{x}^h = (x^h, y^h, w^h)^T$, $\mathbf{X}_c^h = (X_c^h, Y_c^h, Z_c^h, 1)^T$ y $\mathbf{C}_i$ es una matriz de 3×4 . Sin embargo, en muchas ocasiones es más conveniente referir los puntos de la escena a otro marco de referencia diferente del marco de referencia de la cámara. En adelante llamaremos a este marco nuevo como *marco de referencia global*.

El marco de referencia global debe ser muy fácil de identificar o localizar y típicamente tiene como origen al centro geométrico del robot, un punto externo visible de la cámara, (p.ej. el tornillo de sujeción de la cámara) una esquina del ambiente, etc.

La figura 2.4 ilustra la relación entre el marco de referencia de la cámara M_c y el marco de referencia global M_g . En coordenadas homogéneas, un punto $\mathbf{X}_g^h$ del ambiente con respecto a M_g , tiene sus correspondientes coordenadas $\mathbf{X}_c^h$ con respecto a M_c , expresadas de la siguiente forma [Trucco y Verri(1998)]:

$$\mathbf{X}_c^h = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ 0 & 1 \end{bmatrix} \mathbf{X}_g^h \quad (2.9)$$

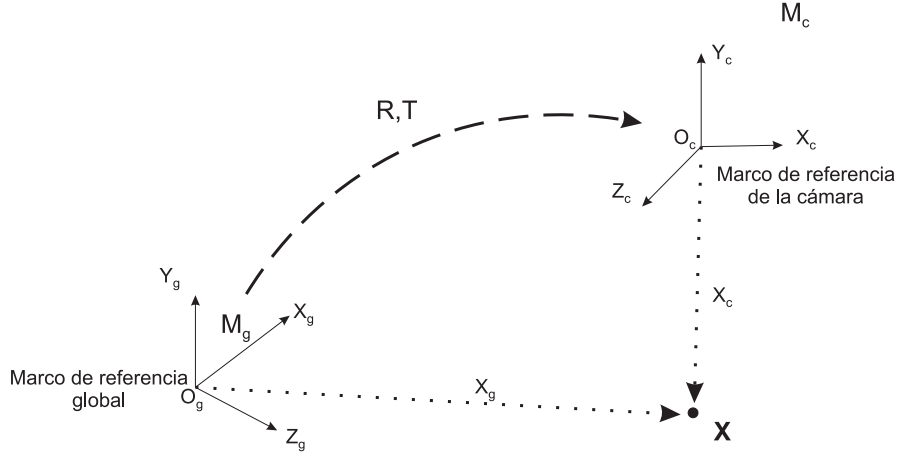


Figura 2.4: Relación entre el marco de referencia global y el marco de referencia de la cámara

donde $\mathbf{R}_{3 \times 3}$ es una matriz de rotación de 3×3 en cada uno de los ejes del marco de referencia y $\mathbf{T}_{3 \times 1}$ es una traslación en 3D. De este modo, dicha matriz representa la traslación y rotaciones involucradas entre ambos marcos de referencia. El lector interesado en el detalle de las rotaciones y traslaciones, conocidos como movimientos de cuerpos rígidos, puede consultar [Ma et al.(2004)].

En forma compacta la ecuación 2.9 se puede reescribir como:

$$\mathbf{X}_c^h = \mathbf{C}_e \mathbf{X}_g^h \quad (2.10)$$

donde $\mathbf{C}_e = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ 0 & 1 \end{bmatrix}$ es una matriz de 4×4 .

Sustituyendo la ecuación 2.10 en 2.8 tenemos:

$$\mathbf{x}^h = \mathbf{C}_i \mathbf{C}_e \mathbf{X}_g^h \quad (2.11)$$

Si llamamos a $\mathbf{C}_i \mathbf{C}_e = \mathbf{P}$, como la matriz de calibración global, tenemos:

$$\mathbf{x}^h = \mathbf{P} \mathbf{X}_g^h \quad (2.12)$$

donde $\mathbf{P}$ es una matriz de 3×4 .

2.4. Calibración

La calibración es el proceso de calcular los *parámetros intrínsecos* y *extrínsecos* de la cámara, directamente de las imágenes tomadas por la misma. Cabe mencionar que el proceso de calibración de un sistema óptico no es totalmente automático y se requiere en mayor o menor grado la intervención humana.

Existe una relación entre los parámetros de las cámaras con el modelo ideal de pinhole. Es decir, es posible determinar las coordenadas de los píxeles de un punto de la imagen a partir de las correspondientes coordenadas en el marco de referencia de la cámara (*calibración interna*) y determinar la posición relativa y orientación de las cámaras (*calibración externa*). Esto es, la calibración interna (determinación de parámetros intrínsecos, internos a la cámara) intenta determinar los parámetros que forman la matriz $\mathbf{C}_i$; la calibración externa (determinación de parámetros extrínsecos, externos a la cámara) intenta determinar los parámetros que forman la matriz $\mathbf{C}_e$. También se puede hablar de una calibración global que intenta determinar la matriz $\mathbf{P}$. Con esta calibración global nos volveremos a encontrar en el capítulo 4, cuando con ella se efectúen las calibraciones de cámaras.

A continuación se mencionan algunos métodos de calibración existentes.

Un primer tipo de calibración usa un patrón de calibración con ciertas características, cuyas coordenadas son conocidas. Estas características pueden ser puntos, esquinas o cualquier otra característica que pueda ser extraída de la imagen adquirida [Tsai y Lenz(1988)]. Este método calcula los parámetros intrínsecos y extrínsecos al mismo tiempo, lo cual conlleva a ciertos errores producidos por el fuerte acoplamiento entre los parámetros externos e internos, siendo éstos últimos los que se ven mayormente afectados [Weng et al.(1992)].

Otra clase de calibración incluye aquellos que no necesitan el conocimiento acerca de la posición de los puntos en el mundo. En este caso es posible separar calibración interna de la calibración externa y el paso de la optimización puede ser llevado a cabo únicamente sobre los parámetros internos. Un ejemplo de este método se presenta en [Devernay y Faugeras(1995)].

La calibración con un *par estéreo* (a partir de dos cámaras, se obtiene una repre-

sentación de objetos 3D en un plano) puede calibrar los parámetros de las cámaras por separado, calculando primeramente los parámetros internos, sin tomar en cuenta para ello los parámetros externos de la misma, es decir, sin importar la posición que tenga la cámara con respecto a la escena. El proceso de calibración interna minimiza una función de distorsión radial de una imagen que contenga únicamente líneas rectas y la calibración externa se basa en algunas propiedades de los *mapas de disparidad* [Iocchi y Konolige(1998)].

En [Trucco y Verri(1998)] se lleva a cabo el proceso de la calibración sin considerar la distorsión radial, suponiendo que las cámaras generan imágenes con poca distorsión, y por lo tanto puede ser ignorada cuando no se requiera mucha exactitud en todas las regiones de la imagen, o bien, cuando los píxeles de la periferia puedan ser descartados.

Otro tipo de calibración es el que no necesita conocer las coordenadas de las marcas en el mundo tridimensional. Este enfoque, conocido como autocalibración, asume que todos los parámetros son desconocidos y se busca encontrar el que mejor ajuste a través de métodos iterativos [Iocchi y Konolige(1998)].

Resumiendo la calibración interna y externa de las cámaras puede realizarse en forma conjunta o por separado. Además debe tomarse en cuenta la corrección de la distorsión en imágenes generadas por cámaras con visión amplia.

2.5. Conclusiones

Con lo expuesto en el presente capítulo se introduce un modelo matemático para la correspondencia de puntos de una escena, en el marco de referencia global, con puntos en el plano de la imagen. Particularmente, la calibración de la cámara del presente trabajo, se refiere a una calibración paramétrica, basada en la correspondencia de puntos. En el siguiente capítulo se aborda el tema de la distorsión radial en cámaras de visión amplia.

Además se revisaron algunos métodos que existen en la literatura para la calibración de cámaras. Asimismo, el proceso de calibración global de cámaras, una vez corregida la distorsión, se presenta en el capítulo 4.

Capítulo 3

Corrección de la distorsión radial

Las cámaras digitales comúnmente utilizan lentes para adquirir imágenes. Si dichos lentes fueran totalmente equivalentes al modelo de pinhole producirían imágenes perfectas. Sin embargo, esto generalmente no es así y los lentes ocasionan que las imágenes tengan ciertos defectos o errores ópticos en su formación.

En particular, este trabajo aborda en detalle un tipo de error llamado *distorsión radial* que ocurre frecuentemente al utilizar lentes de amplia apertura, de uso común en aplicaciones de robótica.

3.1. Errores ópticos

El término error óptico se refiere a la imperfección de un sistema óptico. Los errores ópticos se presentan en las imágenes resultantes tomadas por una cámara y dependen entre otros factores, del índice de refracción de los lentes, de la posición del haz de luz que entra en el objetivo o debido a que diferentes áreas de los lentes tienen ligeramente diferentes longitudes focales [Forsyth y Ponce(2002)].

Enseguida se explica de forma breve los tipos de errores ópticos, el lector que quiera ahondar más en el tema puede consultar [Hecht(1987)]:

Error esférico: es un defecto de las lentes en el que los rayos de luz que inciden paralelamente al eje óptico, son llevados a diferentes focos. Los rayos que inciden

en los lentes y que se encuentran más alejados del eje óptico tienden a interceptar el eje óptico a una distancia más corta que los rayos incidentes al eje óptico.

Error cromático: en este tipo de error el nivel de refracción de un lente depende de la longitud de onda (color) de la luz incidente. Es decir, como la longitud de onda de distintos colores es diferente, la refracción de los rayos no logra que todos los rayos incidentes de luz coincidan en un mismo punto focal. Este error se manifiesta en imágenes borrosas al igual que el error esférico.

Error de curvatura: el error existe cuando la imagen que se está representando se plasma en una superficie curvada en lugar de una plana. Este error ocurre por que los rayos de luz de la periferia de los lentes tienen que cruzar una distancia más grande para alcanzar la superficie donde se plasma la imagen. A diferencia de los rayos que se encuentran más cerca del eje óptico. El resultado de este tipo de error es que las imágenes tienen un enfoque mayor en el centro que en la periferia o viceversa. En la figura 3.1 se ilustra este tipo de error.

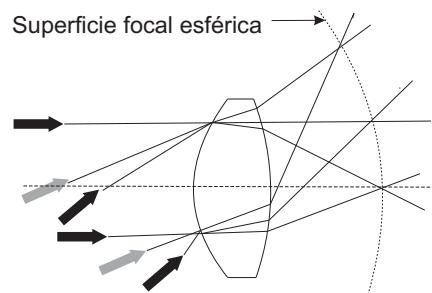


Figura 3.1: Error de curvatura

Error de distorsión: este tipo de error difiere de los demás por que modifica la imagen en su conjunto. Por ejemplo, las partes centrales de una imagen pueden ser más grandes que las de la periferia o viceversa.

Existen otros tipos de errores como el *astigmatismo* que es ocasionado por defectos en algunas zonas de los lentes y que produce imágenes con menor nitidez. A diferencia

del error esférico, que afecta a la imagen completa, el astigmatismo afecta únicamente algunas zonas de la imagen.

En la figura 3.2 (a) se muestra una imagen real capturada por una cámara Videre Design, Modelo MDCS2. Con un rango de adquisición de imágenes de 15 *fps*, con una resolución máxima de 1280×960 píxeles, una interfaz firewire IEEE 1394 y con lentes de 3.5 *mm*, donde se puede apreciar la distorsión ocasionada por los lentes. En la figura 3.2(b) se muestra la imagen una vez que se corrigió el error de distorsión. El resultado es una imagen bastante aceptable. El error de distorsión en las cámaras de amplia apertura es muy significativo, por lo tanto, a continuación se abordará este tipo de error.



(a) Imagen tomada por la cámara.

(b) Imagen corregida

Figura 3.2: Corrección de una imagen con distorsión radial.

3.2. Distorsión

Dependiendo del tipo de lentes, existen dos tipos de distorsión: *radial* y *tangencial* [Weng et al.(1992)]. La distorsión radial se refiere al desplazamiento del punto P^u (sin distorsión o posición ideal) de un objeto (punto) en la imagen, a una nueva posición que se acerca o se aleja del centro de distorsión (c_d) en forma radial. Es decir, se mueve sobre la línea que conecta el punto en cuestión con el centro de distorsión.

En la figura 3.3 la distorsión radial de P^u es de alejamiento. En la misma figura 3.3 se encuentra representada la distorsión tangencial, que es el desplazamiento de un punto en dirección perpendicular a la línea que forma el centro de distorsión con el punto en cuestión. Nótese que el centro de distorsión es invariante a ambos errores o transformaciones en las imágenes.

En la figura 3.2 (a) se ilustra el efecto de una distorsión radial de acercamiento, donde el centro de distorsión se ubica cerca del centro geométrico de la imagen.

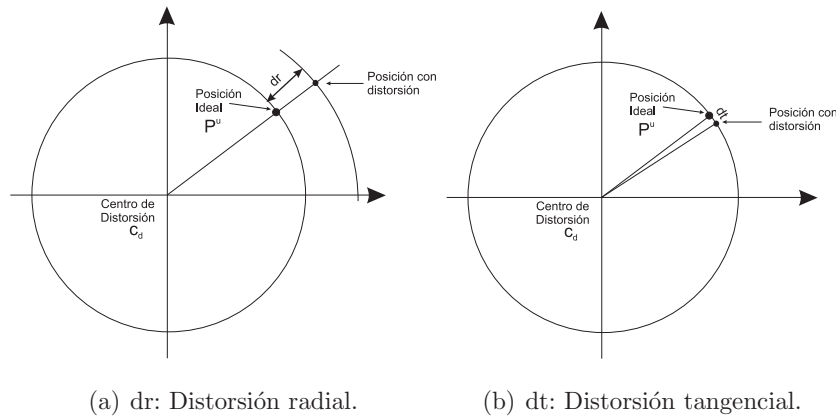


Figura 3.3: Tipos de distorsión: distorsión radial y distorsión tangencial [Weng et al.(1992)].

En la figura 3.4 observamos la distorsión tangencial y su efecto en una imagen. Vemos como las líneas punteadas están desfasadas de su posición ideal cierta distancia que se incrementa conforme los puntos de la imagen se encuentran más alejados del centro de distorsión de la imagen. También vemos que en ciertas regiones de la imagen la distorsión tangencial llegar a ser mínima.

El presente trabajo se enfoca en la corrección de imágenes que presentan distorsión radial. De hecho en la figura 3.2 se corrigió solamente la distorsión radial, debido a que la tangencial no es significativa. Vamos a describir entonces un poco más sobre este tipo de distorsión en la siguiente sección.

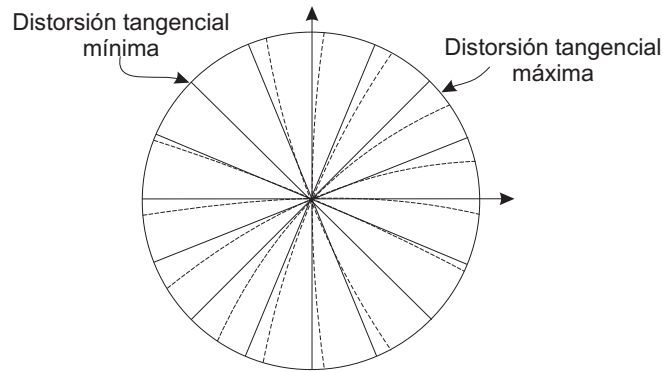


Figura 3.4: Efecto de la distorsión tangencial. Líneas sólidas: sin distorsión; líneas punteadas: con distorsión tangencial [Weng et al.(1992)].

3.2.1. Distorsión radial

La figura 3.5 muestra dos tipos de distorsión que comúnmente se presentan al utilizar lentes. Si la cámara ve un rectángulo en un plano ortogonal al eje óptico, la imagen sin distorsión contiene un rectángulo. El rectángulo puede aparecer como una figura geométrica curvada hacia adentro (distorsión radial tipo barril o de acercamiento) o curvada hacia fuera (distorsión radial tipo almohadilla o de alejamiento) [Weng et al.(1992)].

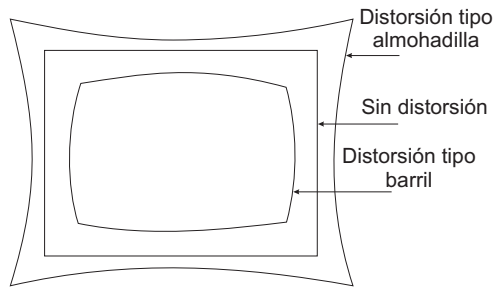
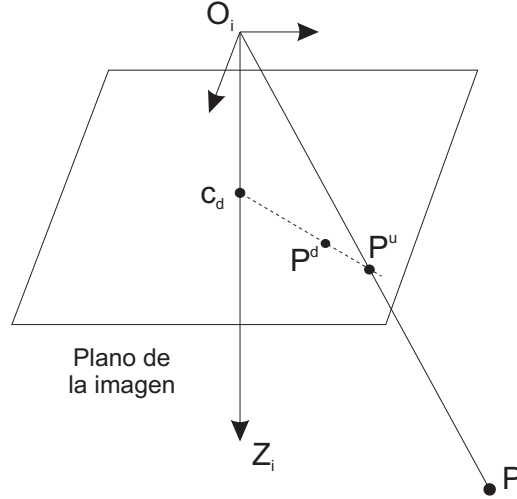


Figura 3.5: Efectos de la distorsión radial. Imagen tomada de [INFAIMON(2003)]

La distorsión radial hacia adentro ocurre frecuentemente en lentes de amplia apertura, mientras que el otro caso ocurre con lentes de muy pequeña apertura (*telefoto*).

Usualmente la distorsión radial de una imagen es modelada de la siguiente manera.

Un punto P en el espacio 3D es proyectado en el plano de la imagen en un punto P^u , como se observa en la figura 3.6. A continuación, debido al efecto de la distorsión el punto P^u es movido a un punto P^d como se muestra en la figura 3.6.



(a) Distorsión de un punto

Figura 3.6: Modelado de la distorsión [Tamaki(2005)].

Para atenuar la distorsión generalmente se realizan tres pasos: selección de un modelo de distorsión, estimación de los coeficientes de distorsión y por último corregir la distorsión de la imagen [Ma et al.(2003)].

Para encontrar los parámetros dado un modelo de distorsión tenemos dos maneras de encontrarlos [Tamaki(2005)]: una es la formulación D-U (del inglés *Distortion-Undistortion*) y otra la formulación U-D (del inglés *Undistortion-Distortion*). En la formulación D-U un punto sin distorsión P^u es expresado en función de las coordenadas distorsionadas P^d y un conjunto de parámetros θ_1^d .

Esto es:

$$P^u = f_1(P^d, \theta_1^d) \quad (3.1)$$

En la formulación U-D las coordenadas con distorsión son expresadas en función de las coordenadas sin distorsión. Es decir:

$$P^d = f_2(P^u, \theta_2^d) \quad (3.2)$$

Hasta ahora hemos definido dos diferentes modelos para representar la distorsión radial. Falta por modelar la *transformación proyectiva*, la cual forma parte del proceso de formación de la imagen.

3.3. Transformación proyectiva

Consideremos que se coloca un objeto planar al frente de la cámara. Dicho plano, denominado π , contiene una cierta imagen o figura, un rectángulo por ejemplo. Se puede demostrar que la figura del plano π se transforma en otra figura en el plano π' (el plano donde se forma la imagen de la cámara) mediante una transformación proyectiva modelada por una matriz H de dimensión 3×3 llamada homografía [Hartley y Zisserman(2004)]. La transformación proyectiva conserva las líneas, pero no necesariamente el paralelismo entre líneas. Es decir, dependiendo de como se presente el plano π al frente de la cámara, las líneas en la imagen pueden no aparecer paralelas. La figura 3.7 (b) ilustra esta situación, donde el plano π no es ortogonal al eje óptico de la cámara.

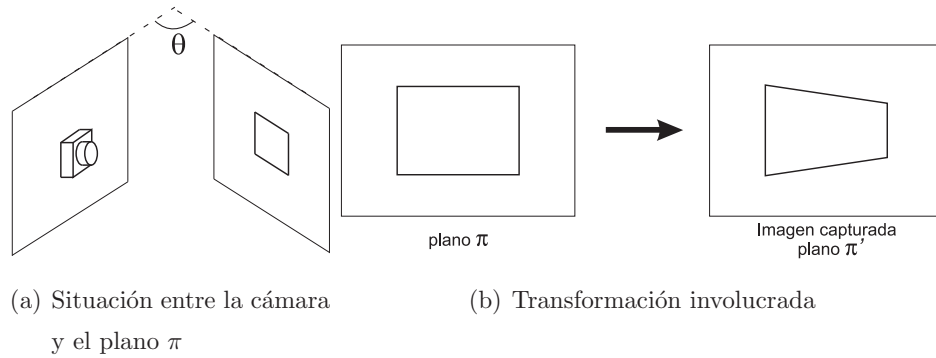


Figura 3.7: Transformación proyectiva.

Un punto $(x_1^r, x_2^r)^T$ en el plano π es representado en coordenadas homogéneas por un vector de 3 coordenadas $(x_1, x_2, x_3)^T$ y ambas coordenadas están relacionadas

por:

$$x_1^r = \frac{x_1}{x_3} \quad x_2^r = \frac{x_2}{x_3} \quad (3.3)$$

La transformación proyectiva transforma un punto $(x_1, x_2, x_3)^T$ del plano π en otro punto $(x_1', x_2', x_3')^T$ en el plano π' mediante la siguiente relación:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (3.4)$$

o en forma breve:

$$\mathbf{x}' = \mathbf{H}\mathbf{x} \quad (3.5)$$

donde $\mathbf{x}$ es un punto en el plano π , $\mathbf{x}'$ denota el correspondiente punto en el plano π' y $\mathbf{H}$ es la homografía [Hartley y Zisserman(2004)].

Tomando en cuenta que un punto $\mathbf{x}$ y otro $k\mathbf{x}$, donde k es una constante real diferente de cero representan el mismo vector en coordenadas homogéneas, la matriz $\mathbf{H}$ tiene 8 grados de libertad [Hartley y Zisserman(2004)]. Es decir, sólo 8 parámetros independientes, de entre los 9 parámetros que forman la matriz.

Si el plano π se presenta cerca de una posición ortogonal al eje óptico de la cámara, h_{33} tiene un valor diferente de cero y la ecuación 3.4, puede representarse de la siguiente forma:

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{pmatrix} x_1^r \\ x_2^r \\ 1 \end{pmatrix} \quad (3.6)$$

Donde $x_1 = x_1^r$, $x_2 = x_2^r$ y $x_3 = 1$.

De esta manera, para calcular los valores $x_1^{r'}$ y $x_2^{r'}$, se tiene:

$$\begin{aligned}
x_1^{r'} &= \frac{x_1'}{x_3'} = \frac{h_{11}x_1^r + h_{12}x_2^r + h_{13}}{h_{31}x_1^r + h_{32}x_2^r + 1} \\
x_2^{r'} &= \frac{x_2'}{x_3'} = \frac{h_{21}x_1^r + h_{22}x_2^r + h_{23}}{h_{31}x_1^r + h_{32}x_2^r + 1}
\end{aligned}
\tag{3.7}$$

En este caso particular, para calcular la transformación proyectiva se requieren conocer h_{11} , h_{12} , h_{13} , h_{21} , h_{22} , h_{23} , h_{31} y h_{32} , un total de ocho parámetros. A esta lista de parámetros se agregarán los parámetros correspondientes a la corrección de la distorsión, como veremos más adelante.

Con lo anterior vimos que al aplicar una homografía a una imagen resulta otra imagen transformada. Para calcular una imagen transformada, dada una homografía $\mathbf{H}$, se puede pensar en utilizar la ecuación 3.5. Sin embargo, por este método se tiene el siguiente problema de interpolación. Consideremos, por ejemplo, que la matriz $\mathbf{H}$ amplifica 3 veces las coordenadas de un punto, como lo ilustra la figura 3.8. De esta forma, no existe información en los píxeles con coordenadas $(1, 1)$, $(1, 2)$, $(0, 1)$, etc. de la imagen transformada y su valor tendría que interpolarse de los valores vecinos.

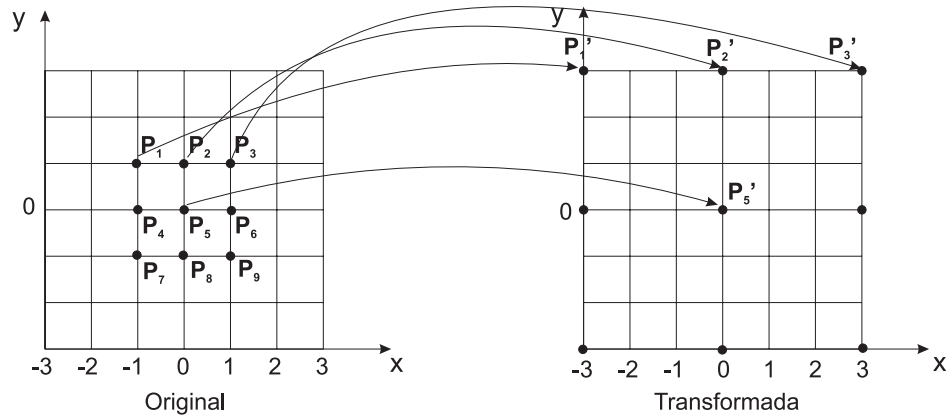


Figura 3.8: Escalamiento de una imagen.

El problema de interpolación se complica aún más si las coordenadas de los puntos

transformados no son coordenadas enteras.

Afortunadamente, el problema de cálculo de la imagen transformada puede plantearse a la inversa, y existe una interpolación llamada *bilineal* [Press et al.(1992)] que resuelve satisfactoriamente el problema.

3.3.1. Interpolación bilineal

A partir de la ecuación 3.5 se puede calcular $\mathbf{x}$ en función de $\mathbf{x}'$, de la siguiente forma:

$$\mathbf{x} = \mathbf{H}^{-1}\mathbf{x}' \quad (3.8)$$

Visto de esta forma el problema de interpolación es más sencillo. Para calcular el tono de gris de un pixel con coordenadas enteras $\mathbf{x}'$, $\mathbf{x}$ daría las coordenadas (reales en el caso general) correspondientes en la imagen original.

Ahora el problema consiste en conocer el nivel de gris correspondiente a dichas coordenadas (reales) en función de los cuatro vecinos más cercanos que tienen coordenadas enteras en la imagen original. La figura 3.9 ilustra esta situación.

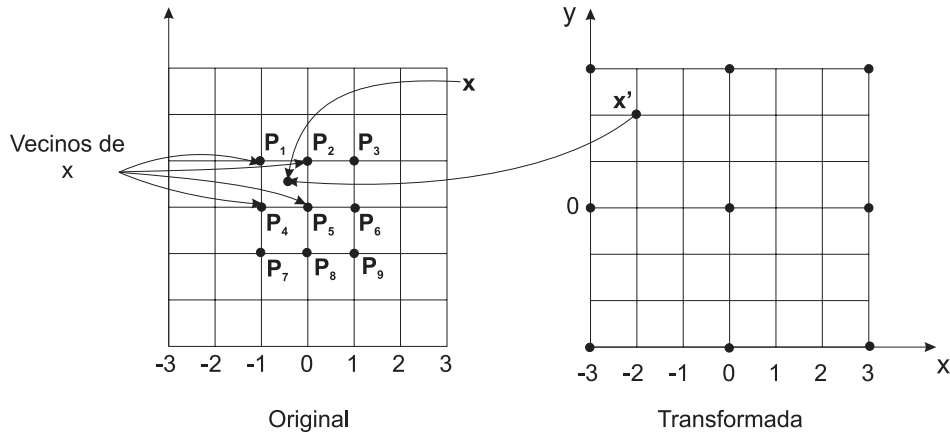


Figura 3.9: El problema de interpolar el valor para $\mathbf{x}$ a partir de los vecinos.

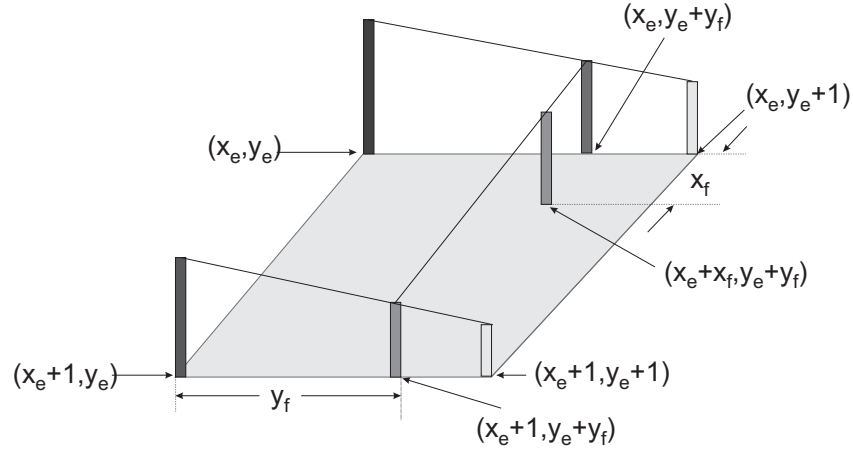


Figura 3.10: Interpolación bilineal.- Los niveles de gris de los píxeles se encuentran representados por barras verticales.

Sea x_e y x_f la parte entera y la parte fraccional respectivamente de x . Sea y_e y y_f la parte entera y la parte fraccional de y . En la figura 3.10 se ilustra el método de la interpolación bilineal para encontrar el valor de intensidad del píxel $I(x_e + x_f, y_e + y_f)$, basándose en la intensidad de los 4 píxeles adyacentes [Press et al.(1992)].

En la figura 3.10 primero se interpola $I(x_e, y_e + y_f)$ en base a $I(x_e, y_e)$ y $I(x_e, y_e + 1)$, asumiendo una relación lineal entre ellos, simbolizada por la línea recta que conecta las partes superiores de las barras verticales (representando el nivel de gris) de los píxeles involucrados. Posteriormente se interpolan $I(x_e + 1, y_e + y_f)$ en base a los valores de intensidad de los píxeles $I(x_e + 1, y_e)$ y $I(x_e + 1, y_e + 1)$. Por último, de los dos valores calculados previamente se obtiene finalmente el valor deseado, en este caso $I(x_e + x_f, y_e + y_f)$.

La notación matemática de la interpolación bilineal quedaría entonces de la siguiente manera [Press et al.(1992)]:

$$\begin{aligned}
 I(x_e, y_e + y_f) &= (1 - y_f)I(x_e, y_e) + y_f I(x_e, y_e + 1) \\
 I(x_e + 1, y_e + y_f) &= (1 - y_f)I(x_e + 1, y_e) + y_f I(x_e + 1, y_e + 1) \\
 I(x_e + x_f, y_e + y_f) &= (1 - x_f)I(x_e, y_e + y_f) + x_f I(x_e + 1, y_e + y_f)
 \end{aligned} \tag{3.9}$$

Hasta aquí hemos visto el modelado de la transformación proyectiva y la forma

de calcular una imagen transformada. Enseguida se aborda el modelado de la distorsión radial y más tarde la forma de calcular los parámetros de los modelos de transformación proyectiva y distorsión radial.

3.4. Ajuste de un modelo de distorsión

Una forma de corregir la distorsión radial generada por las cámaras digitales, consiste en tomar como referencia una imagen que no contiene error, es decir una imagen ideal. La solución al problema consiste en encontrar una transformación que al aplicarse a una imagen tomada por la cámara nos de una imagen transformada. Entre la imagen de referencia y la imagen transformada deberá haber una diferencia mínima. Para esto será necesario encontrar un método que nos ayude a encontrar los parámetros de la transformación.

La ecuación 3.2, página 25, que se repite a continuación, representa la transformación U-D de las coordenadas de los puntos sin distorsión P^u a coordenadas de los puntos con distorsión P^d en base a un conjunto de parámetros θ_2^d [Hartley y Zisserman(2004)].

$$P^d = f_2(P^u, \theta_2^d)$$

La ecuación anterior, considerando una distorsión radial, puede reescribirse como sigue:

$$\begin{pmatrix} x_d' \\ y_d' \end{pmatrix} = f_2(r_u) \begin{pmatrix} x_u' \\ y_u' \end{pmatrix} \quad (3.10)$$

Donde:

- $(x_u', y_u')^T$ Es la posición de un punto en la representación ideal de la imagen (sin distorsión) con respecto al centro de distorsión radial c_d .
- $(x_d', y_d')^T$ Es la posición de un punto en la imagen distorsionada, con respecto al centro de distorsión radial c_d .

- r_u Es la distancia radial $\sqrt{x_u'^2 + y_u'^2}$ entre P^u y el centro de la distorsión radial c_d .
- $f_2(r_u)$ Es una función de distorsión radial la cual depende de r_u .

En la figura 3.11 se muestra el efecto que produce la distorsión radial en una imagen. En la figura 3.11(a) se muestra una imagen sin distorsión. En la figura 3.11(b) se muestra como sería la imagen con distorsión, y finalmente en la figura 3.11(c) se muestra la sobreposición de las dos imágenes. En la figura 3.11(c) se ve como el punto P^u se movió hasta llegar a la posición que tiene el punto P^d . Podemos concluir, que la imagen con distorsión está encogida con respecto a la imagen sin distorsión. En otras palabras, un punto P^u con coordenadas $(x_u', y_u')^T$ está más lejos del centro que el punto P^d con coordenadas $(x_d', y_d')^T$.

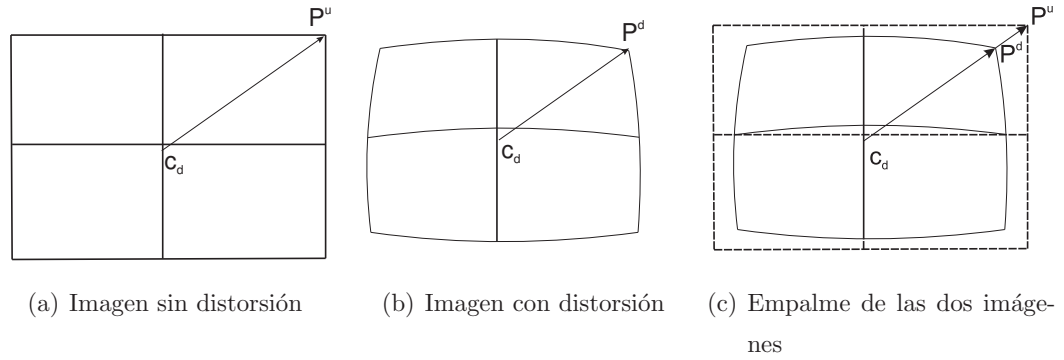


Figura 3.11: Efecto de la distorsión radial.

La diferencia $x_u' - x_d'$ y $y_u' - y_d'$ determinan la distorsión (compresión) de la imagen en el punto (x_d', y_d') . El efecto de la distorsión es distinto para cada punto de la imagen y dependerá que tan cerca o alejado esté del centro de distorsión c_d , con coordenadas (c_x, c_y) , el cual no necesariamente coincide con el centro geométrico de la imagen. Por esta razón a esta transformación de la imagen se le conoce como distorsión radial, ya que la distorsión está en función de la distancia que existe desde el centro de distorsión hasta el punto en cuestión. Es decir, entre más alejado esté el punto, mayor será la diferencia entre P^u y P^d .

Si se tratara de una cámara ideal los dos puntos tanto P^u y P^d tendrían las mismas coordenadas. Es decir:

$$\begin{aligned} x_u &= x_d \\ y_u &= y_d \end{aligned} \quad (3.11)$$

Sin embargo, esto sucede solo en el modelo ideal de una cámara.

De la ecuación 3.10 se desprenden las siguientes ecuaciones del modelo U-D; considerando que $x'_d = x_d - c_x$, $y'_d = y_d - c_y$, $x'_u = x_u - c_x$, $y'_u = y_u - c_y$, tenemos.

$$\begin{aligned} x_d - c_x &= (x_u - c_x)f_2(r_u) \\ y_d - c_y &= (y_u - c_y)f_2(r_u) \end{aligned} \quad (3.12)$$

Donde r_u se puede calcular como sigue:

$$r_u = \sqrt{(x_u - c_x)^2 + (y_u - c_y)^2} \quad (3.13)$$

De manera análoga las correspondientes ecuaciones para el modelo D-U son las siguientes:

$$\begin{aligned} x_u - c_x &= (x_d - c_x)f_1(r_d) \\ y_u - c_y &= (y_d - c_y)f_1(r_d) \end{aligned} \quad (3.14)$$

Donde:

$$r_d = \sqrt{(x_d - c_x)^2 + (y_d - c_y)^2} \quad (3.15)$$

Tanto f_1 como f_2 pueden ser de cualquiera de las siguientes dos formas [Ma et al.(2003)]:

$$\begin{aligned} f_a(r) &= 1 + k_1 r^2 + k_2 r^4 + k_3 r^6 + \dots \\ f_b(r) &= 1 + k_1 r + k_2 r^2 + k_3 r^3 + \dots \end{aligned} \quad (3.16)$$

Donde r puede ser r_d ó r_u dependiendo de la formulación (U-D o D-U). La función f_b es una expansión de serie de Taylor de r , en tanto que la otra función f_a es una expansión en r^2 . La segunda forma es la expresión clásica utilizada en la literatura, mientras que la primera es descrita en [Ma et al.(2003)] y resulta adecuada para

modelar distorsiones pequeñas con mejor precisión. Un modelo de distorsión queda completo seleccionando el tipo de formulación (U-D o D-U), el tipo de función (f_a o f_b) y determinando experimentalmente los valores de cada parámetro. Normalmente k_1 es suficiente para modelar lentes de poca apertura y se requiere de k_1 , k_2 y k_3 para lentes de amplia apertura. En adelante consideramos 3 parámetros de manera que los parámetros del modelo son k_1, k_2, k_3, c_x y c_y . Para encontrar estos parámetros es necesario encontrar una función que minimice los errores existentes entre una imagen ideal o patrón, y una imagen capturada por la cámara.

En resumen, para modelar la distorsión utilizaremos un total de 5 parámetros (c_x, c_y, k_1, k_2, k_3), que en conjunto con los 8 parámetros que modelan la proyección, tendremos un total de 13 parámetros.

Denotaremos dichos parámetros como el conjunto de parámetros $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, donde $n = 13$ (5 de distorsión y 8 de proyección).

No obstante, que existen dos formulaciones para encontrar el conjunto de parámetros Θ , únicamente el modelo D-U resulta adecuado para modelar altas distorsiones radiales de tipo barril, como se describe a continuación.

3.4.1. Modelo D-U

El modelo D-U como se explicó en la sección 3.4, utiliza las coordenadas de la imagen con distorsión radial, para obtener las coordenadas sin distorsión. La otra formulación, el modelo U-D se basa en las coordenadas sin distorsión para obtener las coordenadas con distorsión. La diferencia que existe entre los modelos D-U y U-D puede ser vista más claramente si graficamos las funciones f_1 y f_2 . Sea r_d y r_u la distancia euclidiana entre un punto con distorsión P^d y el centro de distorsión c_d y entre un punto P^u y c_d , respectivamente.

La relación de las distancias radiales r_d y r_u pueden ser modeladas de la siguiente manera:

$$r_u = r_d f_1(r_d^2) \quad (\text{Modelo D-U}) \quad (3.17)$$

$$r_d = r_u f_2(r_u^2) \quad (\text{Modelo U-D}) \quad (3.18)$$

Donde las funciones f_1 y f_2 , pueden ser aproximadas con ayuda de la serie de Taylor [Ma et al.(2003)] (ver ecuaciones 3.16). La figura 3.12 muestra el comportamiento de las funciones f_1 y f_2 para modelar una distorsión radial de acercamiento o barril. La función f_1 modela una compresión, donde el punto P^u se comprime hasta ubicarse en el punto P^d como se muestra en la figura 3.11. En contraste, la función f_2 modela una expansión, donde el punto P^d se expande hasta alcanzar el punto P^u . Para graficar el comportamiento de estos modelos utilizamos la función $f_a(r) = 1 + k_1 r^2$, tanto en la función f_1 como en f_2 .

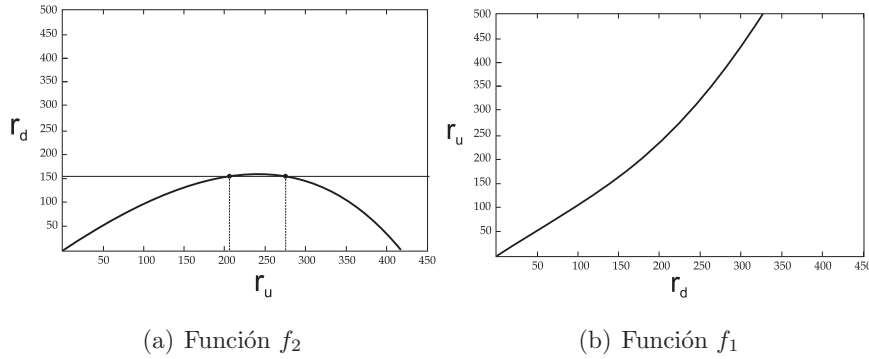
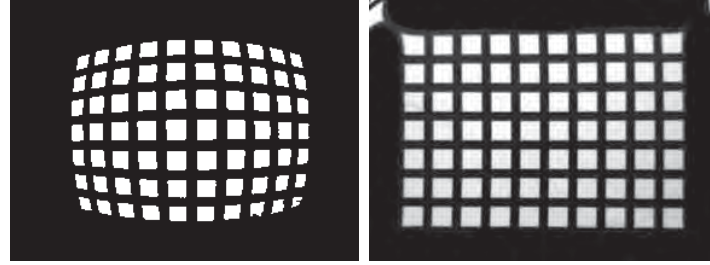


Figura 3.12: Dos funciones diferentes para modelar la distorsión.

La figura 3.13 muestra el efecto que se obtiene en la corrección de imágenes utilizando la función f_2 . En la figura 3.13 (a) tenemos una imagen tomada por la cámara y en la figura 3.13 (b) se muestra la imagen corregida. Podemos observar que en las esquinas de la parte superior de la imagen corregida (figura 3.13 (b)) aparecen zonas blancas, donde no debería de haber. La línea horizontal de la figura 3.12 (a) ayuda a entender la aparición de estas zonas blancas erróneas. Vemos que dos diferentes r_u se relacionan con un mismo r_d . Es decir, las esquinas superiores de la imagen toman su nivel de gris de la misma posición donde aparecen los cuadros blancos en la imagen distorsionada.

Este problema no ocurre con f_1 (modelo D-U) pues la relación entre r_u y r_d es única para la distorsión radial tipo barril. Por ello, en adelante solo consideramos el

modelo D-U.



(a) Imagen de la cámara.

(b) Imagen corregida.

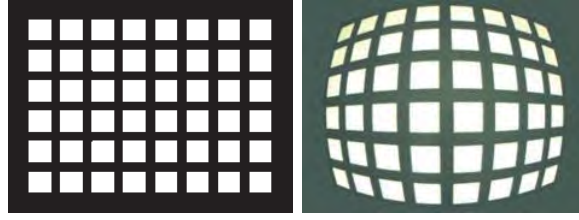
Figura 3.13: Efecto de la corrección de la distorsión radial utilizando el modelo U-D.

Una vez que se define el modelo a utilizar para la corrección en imágenes con distorsión, toca el turno de escoger una forma de comparar las imágenes de referencia y la imagen transformada obtenida a partir de la imagen de la cámara. En la siguiente sección se aborda un método basado en la selección de puntos característicos que se tomarán en cuenta para obtener un conjunto de correspondencias de puntos entre las dos imágenes: la imagen de referencia I_r y la imagen tomada por la cámara I_d . Enseguida analizaremos el método de correspondencia de puntos, así como el criterio de selección de correspondencias de puntos entre dichas imágenes.

3.5. Método de correspondencias de puntos

El objetivo es encontrar un conjunto de parámetros $\Theta = [c_x, c_y, k_1, k_2, k_3, h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}]$ a partir de dos imágenes: una imagen de referencia que llamaremos I_r y la imagen capturada por la cámara I_d , cuando la imagen I_r (imagen impresa) se coloca al frente de la cámara. La figura 3.14 muestra un ejemplo de I_r e I_d .

La idea para encontrar los parámetros del modelo es encontrar un conjunto de n puntos correspondientes entre ambas imágenes. Por simplicidad y facilidad de cálculo, se utilizan los centros o las esquinas de cada uno de los cuadros blancos de las imágenes.



(a) Imagen de referencia I_r . (b) Imagen I_d tomada por la cámara.

Figura 3.14: Imágenes utilizadas en la corrección de la distorsión.

Sea P'_i un punto de I_r (imagen sin distorsión), y P_i el correspondiente punto de la imagen I_d . El error apropiado a minimizar es la distancia euclidiana que existe entre los puntos P'_i y $f(\Theta, P_i)$, donde $f(\Theta, P_i)$ denota la transformación de corrección de la distorsión y la transformación proyectiva. La figura 3.15 muestra gráficamente como opera $f(\Theta, P_i)$. Primero el punto P_i de I_d se transforma en un punto sin distorsión radial logrando con ello que los cuadros blancos no tengan curvas; y después se transforma dicho punto en otro que lo transforma proyectivamente en un punto similar a P'_i .

Usaremos la notación $\|P' - P\|^2$, para denotar el cuadrado de la distancia euclidiana entre puntos no homogéneos representados por P' y P . Entonces, el error a minimizar por el conjunto de n correspondencias es:

$$E_t = \sum_i^n \|P'_i - f(\Theta, P_i)\|^2 \quad (3.19)$$

donde Θ es el argumento buscado que minimice la sumatoria de distancias. Es decir: $\Theta = \operatorname{argmin} E_t$.

El siguiente problema es elegir los puntos. Dadas dos imágenes I_r e I_d tenemos que encontrar el conjunto de correspondencias de puntos entre las dos imágenes. A continuación se muestran los criterios que se tomaron en cuenta para obtener las coordenadas de los puntos en las imágenes, puntos que llamaremos *puntos característicos*.

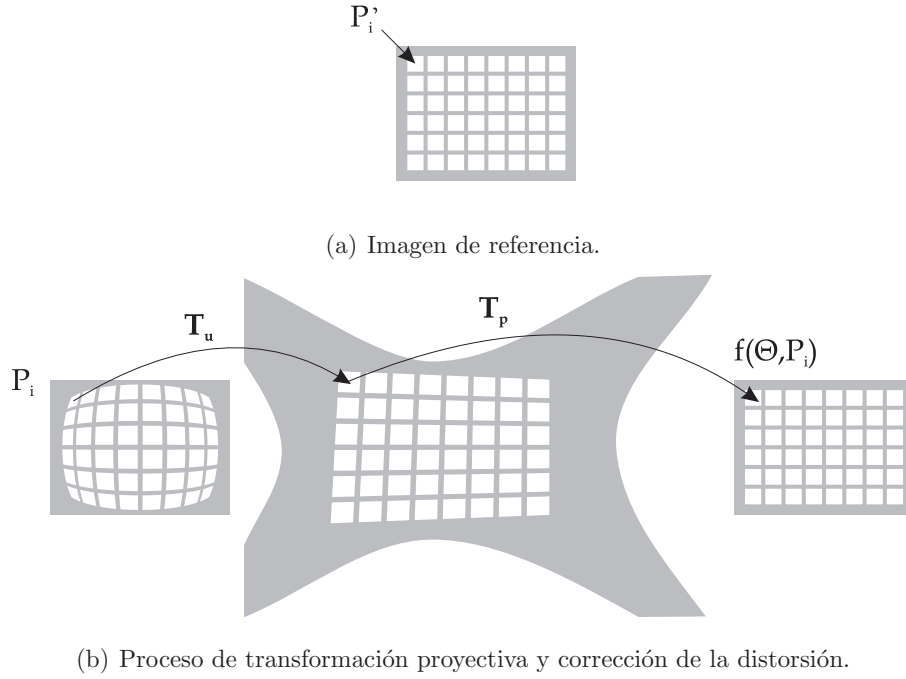


Figura 3.15: El objetivo es que un punto P'_1 en la imagen de referencia I_r , coincida con el punto $f(\Theta, P_i)$ de la imagen transformada de la cámara. La transformación global considera primero la transformación para quitar la distorsión radial (T_u) y después la transformación proyectiva (T_p).

3.5.1. Selección de puntos característicos

Para facilidad de uso, sería altamente deseable disponer de un método automático de selección de puntos característicos de las imágenes I_d e I_r . Como se observa en la figura 3.14 (b), no es fácil determinar las esquinas de los cuadros blancos cercanos a las orillas de la imagen. Por ello se prefirió, en una primera aproximación para calcular los parámetros Θ , utilizar los centros geométricos de los cuadros blancos.

Para el cálculo automático de dichos puntos, primero se calculan los centros de todos los cuadros blancos de la imagen de referencia I_r . Sea k el número de cuadros detectados.

Enseguida se procede a detectar todos los cuadros blancos de la imagen I_d , que pueden ser más de k cuadros en el caso de que la cámara capture zonas adicionales al

marco de referencia. Un ejemplo de ello se muestra en la figura 3.14 (b). Para desechar los cuadros espurios, se consideran las dos siguientes suposiciones:

- En la imagen I_d aparece aproximadamente centrada la imagen del marco de referencia.
- Los cuadros blancos “buenos” no tocan los límites de la imagen I_d .
- El marco de referencia no está muy girado, de manera que se conserva la relación entre los cuadros blancos y las esquinas de la imagen de referencia I_r con respecto a los cuadros correspondientes y las esquinas de la imagen I_d . Por ejemplo, el cuadro superior izquierdo es el más cercano a la esquina superior izquierda.

Aprovechando la primera y segunda suposiciones se desechan los cuadros espurios, procediendo primero a considerar los cuadros más cercanos al centro de la imagen.

La segunda suposición es clave para establecer la asociación automática de los puntos característicos de I_d e I_r .

Una vez encontrada una primera aproximación para los parámetros del modelo, se puede utilizar la imagen corregida para mejorar la estimación de Θ .

Para ello se aprovecha que es mucho más sencillo determinar los puntos característicos de la imagen corregida (p. ej. esquinas o centros geométricos de los cuadros). A partir de los puntos característicos determinados en la imagen corregida, se pueden calcular los puntos correspondientes en I_d considerando el conjunto de parámetros Θ recién obtenidos. La figura 3.16 ilustra esta situación. Con ello se mejora la estimación de los puntos característicos en I_d y el proceso se puede repetir para obtener un conjunto de parámetros Θ' . Finalmente el proceso puede iterarse hasta que el conjunto de parámetros no cambie o el error a minimizar no disminuya.

Para obtener el centro geométrico de cada uno de los cuadros blancos de las imágenes; primero la imagen pasa por un proceso de *binarización*, el cual permite mediante un umbral separar los píxeles blancos de los negros. Es decir, a los píxeles negros se les asigna un valor de 0 binario y a los píxeles blancos se les asigna un valor de 1. Una vez hecho esto se tratan de manera independiente las áreas blancas y con las siguientes formulas es posible obtener las coordenadas de los centros de dichas áreas.

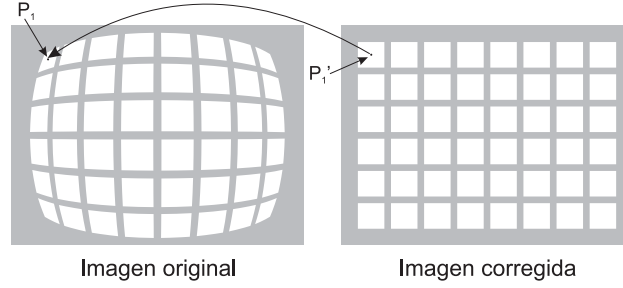


Figura 3.16: Correspondencias de puntos. Determinando el punto característico P_1' en la imagen corregida se puede calcular el correspondiente punto P_1 en la imagen original I_d .

Sea k el número de cuadros blancos. Las coordenadas (x_j, y_j) del centro geométrico del cuadro j -ésimo ($j = 1, 2, \dots, k$) se pueden calcular como sigue:

$$\begin{aligned} x_j &= \frac{1}{r} \sum_{i=1}^{r_j} x(P_{ij}) , \\ y_j &= \frac{1}{r} \sum_{i=1}^{r_j} y(P_{ij}) \end{aligned} \quad (3.20)$$

Donde:

- r_j representa el número de píxeles en blanco, que componen el j -ésimo cuadro blanco;
- $x(P_{ij})$ es la coordenada x del punto P_i de un píxel blanco correspondiente al cuadro blanco j ;

Una vez que se obtienen el conjunto de correspondencias de puntos es necesario minimizar el error geométrico, para ello, existen métodos de optimización no lineal que resuelven el problema. En la siguiente sección explicamos el método sobre el que nos basamos, el cuál ofrece algunas ventajas sobre otros métodos. Este método resulta adecuado en la resolución de los sistemas no lineales involucrados.

3.6. Optimización no lineal

Existen diferentes métodos para encontrar una solución a ecuaciones similares a la ecuación 3.19 del página 36, dentro los que destacan, por su utilidad, los métodos: mínimos cuadrados, Newton, Gauss-Newton y Gauss-Newton-Levenberg-Marquardt [Nocedal y Wright(1999)]. Cada uno ha representado una mejora con respecto al que se le antecede, principalmente en la optimización con que se realizan los cálculos y la convergencia en las iteraciones. Para minimizar el error geométrico (ecuación 3.19), se requiere del uso de técnicas iterativas, como es el caso de los métodos mencionados. En particular, el método que presentaremos y que fue utilizado en nuestra investigación para calcular el argumento mínimo de la ecuación 3.19, es el método Gauss-Newton-Levenberg-Marquardt (GNLM), el cual es un método iterativo de optimización no lineal, y que se presenta a lo largo de la presente sección.

En cada iteración, el incremento de los parámetros $\delta\theta$, es calculado resolviendo la siguiente ecuación matricial:

$$[\mathbf{A} + \lambda]\delta\Theta = \mathbf{B} \quad (3.21)$$

En el apéndice A se muestra el desarrollo de dicha ecuación y se explican las matrices $\mathbf{A}$ y $\mathbf{B}$.

En nuestro problema tenemos n correspondencias de puntos y m parámetros en Θ , $\mathbf{A}$ es una matriz de $m \times m$ y $\mathbf{B}$ es una matriz de $m \times 1$. En el algoritmo 1 se muestra el método iterativo GNLM.

En el apéndice A se muestra el desarrollo del procedimiento utilizado para la optimización de la función de error dada por la ecuación 3.19.

Una vez que se obtuvieron los parámetros Θ , se puede corregir la imagen con distorsión radial. La siguiente sección muestra este proceso.

3.7. Corrección de la distorsión radial

Dado que se cuenta con el conjunto de parámetros θ_1^d , así como el modelo D-U, se puede mapear un punto con coordenada $(x_p^r, y_p^r)^T$ a un punto con coordenada

Algoritmo 1 Algoritmo de Optimización no lineal - GNLM

Objetivo.

Encontrar el conjunto de parámetros Θ los cuales minimicen la función de error.

Algoritmo.

Dada una función diferenciable $f(\Theta, P_i)$ (que representa la transformación de la imagen) y un valor inicial los parámetros Θ .

Inicializar $\lambda = 0.001$, $inc = 10$, $lim_sup_ \lambda = 1e^{30}$

Hacer**Hacer**

Calcular $E_t(\Theta)$ el error total, ecuación (3.19).

Resolver $A\delta\Theta = B$ ecuación (3.21).

Calcular el error $E_t(\Theta + \delta\Theta)$.

Si $E_t(\Theta + \delta\Theta) > E_t(\Theta)$ **Entonces**

$$\lambda = \lambda * inc$$

Fin si

Mientras ($E_t(\Theta + \delta\Theta) > E_t(\Theta)$ y $\lambda < lim_sup_ \lambda$).

Si $E_t(\Theta + \delta\Theta) < E_t(\Theta)$ **entonces**

$$\lambda = \lambda / inc$$

$$\Theta = \Theta + \delta\Theta.$$

Fin si

Mientras ($\lambda < lim_sup_ \lambda$)

$(x_d^r, y_d^r)^T$. Para esto, es necesario invertir la matriz de transformación $\mathbf{H}$ para obtener las coordenadas homogéneas $(x_u, y_u, w_u)^T$.

$$(x_u, y_u, w_u)^T = \mathbf{H}^{-1}(x_p, y_p, w_p)^T \quad (3.22)$$

Con lo anterior, es posible obtener la coordenada $(x_u^r, y_u^r)^T$. Sin embargo, para obtener la coordenada $(x_d^r, y_d^r)^T$ no existe una manera directa de calcularlo. Por lo que se utilizará el método de la bisección [Chapra y Canale(1999)], donde el objetivo

es encontrar $f_1(r_d^2)$ dado $(x_u^r, y_u^r)^T$ y θ_1^d . Una vez que $f_1(r_d^2)$ ha sido encontrado, la coordenada $(x_d^r, y_d^r)^T$ puede ser calculada a partir de la ecuación 3.14, de la página 32, con las ecuaciones:

$$\begin{aligned} x_d^r &= \frac{(x_u^r - c_x)}{f_1(r_d^2)} + c_x \\ y_d^r &= \frac{(y_u^r - c_y)}{f_1(r_d^2)} + c_y \end{aligned}$$

A partir de la ecuación 3.17 formulamos una nueva ecuación:

$$f(r_d) = r_u - r_d f_1(r_d^2) = r_u - r_d(1 + k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) \quad (3.23)$$

En la ecuación 3.23, se debe obtener la raíz de $f(r_d)$, es decir el valor de r_d tal que $f(r_d) = 0$, dado que se conoce k_1, k_2 y k_3 , así como (x_u, y_u) . Si $x_u^r = c_x$ y $y_u^r = c_y$, de la ecuación 3.14, página 32, se tiene que $x_d^r = x_u^r$, $y_d^r = y_u^r$ y el problema está resuelto. Para el caso cuando $x_u^r \neq c_x$ y $y_u^r \neq c_y$, se necesita encontrar un número, r_{d_0} , tal que $f(r_{d_0}) > 0$ y un número, r_{d_1} , tal que $f(r_{d_1}) < 0$. Con estos límites es posible utilizar el método de la bisección, y encontrar la raíz, y con algunas iteraciones encontrar el resultado con la precisión necesaria.

El límite r_{d_0} se encuentra fácilmente notando que $f(0) = r_u - (0)f_1(r_d^2) = r_u$, siendo $r_u > 0$ en el caso de que (x_u, y_u) no sea el centro de distorsión. Es decir, $r_{d_1} = 0$.

El límite r_{d_1} se estará incrementando el valor de r_d hasta que $f(r_d) < 0$. En la implementación se incrementa r_d en un valor de una unidad en cada iteración.

Una vez que se calibra la cámara, es posible crear un mapeo entre las coordenadas de la imagen tomada por la cámara y la imagen corregida. El mapeo será entonces de coordenadas reales en la imagen tomada por la cámara a coordenadas enteras en la imagen corregida. Debido a ésto, será necesario aplicar una interpolación bilineal para obtener el tono de gris en las coordenadas de la imagen corregida. Además, se desarrolló una interfaz de usuario para observar el efecto que produce la modificación de los parámetros de distorsión k_1, k_2, k_3, c_x y c_y en una imagen. El efecto de transformación de la imagen es inmediato, ya sea por medio de una barra de desplazamiento o por la captura dentro de un cuadro de texto de los parámetros de distorsión θ_1^d .

Con lo anterior damos por concluida la corrección radial en imágenes, enseguida algunas conclusiones al respecto.

3.8. Conclusiones

El presente capítulo muestra como el modelo D-U ayuda a encontrar el conjunto de parámetros que corrigen la distorsión radial en cámaras de muy amplia apertura. También se ha mostrado que utilizando el método U-D no es posible obtener una corrección adecuada en imágenes con distorsión radial generadas por cámaras de amplia apertura.

Debido a que el método D-U requiere de un método iterativo para poder realizar el mapeo entre la imagen tomada por la cámara y la imagen patrón, puede decirse que el método no tiene una solución directa. Sin embargo, esos valores son precargados en memoria para la corrección de las subsiguientes imágenes. Con esto se logra un cálculo más rápido como se mostró en el capítulo.

Se ha dicho que el robot requiere de información del ambiente, es decir necesita obtener un mapeo entre el mundo tridimensional y las imágenes que se obtienen por las cámaras. El primer paso fue corregir las imágenes. Es decir, minimizar la distorsión radial. Ahora se está en la posibilidad de encontrar un mapeo entre el mundo tridimensional y las imágenes tomadas por la cámara. Es en el siguiente capítulo donde hablaremos de la calibración global como una aplicación de la corrección en imágenes.

Capítulo 4

Calibración Global

Una aplicación típica de la corrección de la distorsión radial en imágenes, la encontramos en la *calibración global* de una cámara. La calibración global consiste en encontrar las ecuaciones que relacionen las coordenadas de un punto en el espacio tridimensional, con el correspondiente punto en la imagen tomada por la cámara [Trucco y Verri(1998)]. Si se asume que no existe distorsión en la imagen que toma o captura la cámara, entonces se involucra una transformación proyectiva que se puede modelar de manera lineal al utilizar coordenadas homogéneas. Este capítulo se va a dedicar a encontrar la matriz u homografía que define la transformación proyectiva particular de una cámara.

En la calibración global se trata de encontrar una matriz de proyección. Es decir, habrá una transformación entre puntos de una escena tridimensional y puntos con una imagen obtenida de dicha escena. Los puntos de la escena se obtiene en base a un marco de referencia conocido.

La calibración global se basa en una imagen de un patrón de calibración, como el mostrado en la figura 4.1. Este patrón de calibración es un objeto tridimensional, el cual está formado de dos planos ortogonales. Cada plano contiene varios cuadros blancos con fondo negro y facilita localizar sus puntos esquina más fácilmente.

Para obtener la matriz de transformación que representa la proyección entre las coordenadas del mundo y las coordenadas de la cámara utilizaremos el algoritmo de *Transformación Lineal Directa* [Hartley y Zisserman(2004)] que explicaremos a

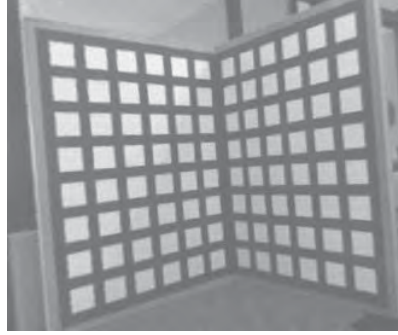


Figura 4.1: Patrón de calibración.

continuación.

4.1. Algoritmo de Transformación Lineal Directa (DLT)

Comenzaremos por modelar el problema de manera geométrica. Primero, considérese un punto en el espacio tridimensional $\mathbf{X}_i$, con coordenadas homogéneas $(X, Y, Z, W)^T$ y un punto en el espacio bidimensional $\mathbf{x}_i$ con coordenadas $(x, y, w)^T$. $\mathbf{X}_i$ es el punto en el marco de referencia global (M_g), mientras que $\mathbf{x}_i$ es el punto en el marco de referencia de la cámara (M_c). En la figura 4.2 se muestra la proyección de un punto $\mathbf{X}_i$ sobre un punto $\mathbf{x}_i$ en el plano de la imagen. Asimismo, el origen del marco de referencia de la cámara, es el centro de proyección y el eje z se encuentra sobre el eje óptico.

La idea del algoritmo es encontrar un conjunto de n correspondencias $\{\mathbf{X}_1 \leftrightarrow \mathbf{x}_1, \mathbf{X}_2 \leftrightarrow \mathbf{x}_2, \dots, \mathbf{X}_n \leftrightarrow \mathbf{x}_n\}$ entre puntos $\mathbf{X}_i$ y puntos $\mathbf{x}_i$. Para esto, requerimos encontrar una matriz de proyección $\mathbf{P}$ o matriz de transformación de dimensión 3×4 . Esta matriz nos definirá las correspondencias entre los puntos mediante la ecuación:

$$\mathbf{x}_i = \mathbf{P}\mathbf{X}_i \quad (4.1)$$

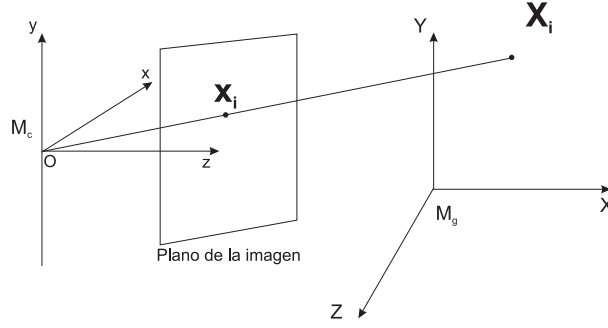


Figura 4.2: Proyección de un punto 3D $\mathbf{X}_i$ a un punto 2D $\mathbf{x}_i$.

Nótese que esta es una ecuación que involucra vectores homogéneos; quiere decir que los vectores $\mathbf{x}_i$ y $\mathbf{P}\mathbf{X}_i$ no son totalmente iguales, ellos difieren en magnitud por un factor de escala diferente de 0 [Trucco y Verri(1998)]. En base a lo anterior, la ecuación 4.1 puede ser expresada con el producto cruz de dos vectores de la siguiente manera:

$$\mathbf{x}_i \times \mathbf{P}\mathbf{X}_i = 0 \quad (4.2)$$

El problema de la calibración global consiste en calcular la matriz $\mathbf{P}$ definida por la ecuación 4.1 en base a las n correspondencias de puntos. De esta manera, la solución del sistema de ecuaciones es lineal, y requerirá un cierto mínimo de ecuaciones para encontrar la matriz $\mathbf{P}$, como se discute a continuación.

4.1.1. Número mínimo de correspondencias

Dado que la matriz $\mathbf{P}$ contiene 12 valores, en total se consideran 11 grados de libertad (ignorando el factor de escala). Cada correspondencia $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$, se refiere a 2 restricciones. Entonces, requeriremos un mínimo de 6 ecuaciones para encontrar la solución a este sistema de ecuaciones. Enseguida se muestra la obtención del sistema de ecuaciones dado un conjunto de correspondencias de puntos.

4.1.2. Sistema de ecuaciones

Primero fijaremos una notación para los elementos de la matriz $\mathbf{P}$. Sea $\mathbf{p}^{jT}$ la j -ésima fila de la matriz $\mathbf{P}$, como se muestra enseguida:

$$\mathbf{P} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} = \begin{bmatrix} \mathbf{p}^{1T} \\ \mathbf{p}^{2T} \\ \mathbf{p}^{3T} \end{bmatrix} \quad (4.3)$$

Podemos escribir $\mathbf{P}\mathbf{X}_i$ como:

$$\mathbf{P}\mathbf{X}_i = \begin{pmatrix} \mathbf{p}^{1T}\mathbf{X}_i \\ \mathbf{p}^{2T}\mathbf{X}_i \\ \mathbf{p}^{3T}\mathbf{X}_i \end{pmatrix}$$

Considerando que $\mathbf{x}_i = (x_i, y_i, w_i)^T$ en la ecuación 4.2 y realizando el producto cruz obtenemos:

$$\mathbf{x}_i \times \mathbf{P}\mathbf{X}_i = \begin{pmatrix} y_i\mathbf{p}^{3T}\mathbf{X}_i - w_i\mathbf{p}^{2T}\mathbf{X}_i \\ w_i\mathbf{p}^{1T}\mathbf{X}_i - x_i\mathbf{p}^{3T}\mathbf{X}_i \\ x_i\mathbf{p}^{2T}\mathbf{X}_i - y_i\mathbf{p}^{1T}\mathbf{X}_i \end{pmatrix} = 0$$

Tomando en cuenta que $\mathbf{p}^{jT}\mathbf{X}_i = \mathbf{X}_i^T\mathbf{p}^j$ el sistema anterior puede ser representado como un conjunto de 3 ecuaciones, las cuales pueden ser escritas de la siguiente forma:

$$\begin{bmatrix} \mathbf{0}^T & -w_i\mathbf{X}_i^T & y_i\mathbf{X}_i^T \\ w_i\mathbf{X}_i^T & \mathbf{0}^T & -x_i\mathbf{X}_i^T \\ -y_i\mathbf{X}_i^T & x_i\mathbf{X}_i^T & \mathbf{0}^T \end{bmatrix} \begin{pmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{pmatrix} = 0 \quad (4.4)$$

Esta ecuación tiene la forma de $\mathbf{A}_i\mathbf{P}_v = 0$ donde $\mathbf{A}_i$ es una matriz de 3×12 y $\mathbf{P}_v$ es un vector de 12 elementos, representado en una matriz $\mathbf{P}$.

$$\mathbf{P}_v = \begin{pmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{pmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{12} \end{bmatrix} \quad (4.5)$$

Aunque existen 3 ecuaciones en la ecuación 4.4, únicamente dos de ellas son linealmente independientes (la tercera fila es obtenida por una combinación lineal de la 1ª y 2ª filas). Así, la tercera ecuación se omite, y de cada correspondencia de puntos obtenemos dos ecuaciones:

$$\begin{bmatrix} \mathbf{0}^T & -w_i \mathbf{X}_i^T & y_i \mathbf{X}_i^T \\ w_i \mathbf{X}_i^T & \mathbf{0}^T & -x_i \mathbf{X}_i^T \end{bmatrix} \begin{pmatrix} \mathbf{p}^1 \\ \mathbf{p}^2 \\ \mathbf{p}^3 \end{pmatrix} = 0 \quad (4.6)$$

Así de un conjunto de n correspondencias de puntos obtenemos una matriz $\mathbf{A}$ de dimensión $2n \times 12$. La matriz de proyección $\mathbf{P}$ es calculada resolviendo un sistema de ecuaciones del tipo $\mathbf{A}\mathbf{P}_v = 0$ donde $\mathbf{P}_v$ es el vector columna que contiene los elementos de la matriz $\mathbf{P}$ como se encuentra representado en la ecuación 4.3.

Así, para n puntos, considerando las correspondencias $\{\mathbf{X}_1 \leftrightarrow \mathbf{x}_1, \mathbf{X}_2 \leftrightarrow \mathbf{x}_2, \dots, \mathbf{X}_n \leftrightarrow \mathbf{x}_n\}$ tenemos $\mathbf{X}_i^T = [X_i, Y_i, Z_i, W_i]$ y el sistema de ecuaciones adquiere la forma:

$$\mathbf{A}\mathbf{P}_v = 0 \quad (4.7)$$

donde la matriz $\mathbf{A}$ está representada por:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -w_1 X_1 & -w_1 Y_1 & -w_1 Z_1 & -w_1 W_1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 W_1 \\ w_1 X_1 & w_1 Y_1 & w_1 Z_1 & w_1 W_1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 & -x_1 W_1 \\ 0 & 0 & 0 & 0 & -w_2 X_2 & -w_2 Y_2 & -w_2 Z_2 & -w_2 W_2 & -y_2 X_2 & -y_2 Y_2 & -y_2 Z_2 & -y_2 W_2 \\ w_2 X_2 & w_2 Y_2 & w_2 Z_2 & w_2 W_2 & 0 & 0 & 0 & 0 & -x_2 X_2 & -x_2 Y_2 & -x_2 Z_2 & -x_2 W_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & -w_n X_n & -w_n Y_n & -w_n Z_n & -w_n W_n & -y_n X_n & -y_n Y_n & -y_n Z_n & -y_n W_n \\ w_n X_n & w_n Y_n & w_n Z_n & w_n W_n & 0 & 0 & 0 & 0 & -x_n X_n & -x_n Y_n & -x_n Z_n & -x_n W_n \end{bmatrix} \quad (4.8)$$

Una vez que se tiene el sistema de ecuaciones se puede calcular la matriz $\mathbf{P}_v$ por medio de descomposición en valores singulares [Press et al.(1992)]. Sin embargo, para un resultado óptimo se recomienda la normalización de las coordenadas de los puntos. El

apéndice C muestra el procedimiento requerido para la normalización de los puntos. En el algoritmo 2 se muestra el procedimiento para obtener la matriz $\mathbf{P}_v$ dadas n correspondencias de puntos.

Con el cálculo de la matriz $\mathbf{P}$, se logra el objetivo de la calibración global que es el mapear puntos de la escena con puntos en la imagen. El mapeo se logra aplicando la ecuación 4.1. Es decir, dado un conjunto de puntos de la escena es posible obtener sus correspondientes puntos en la imagen.

4.2. Conclusiones

El proceso de calibración global mostrado considera que las imágenes no tienen distorsión radial y es un método sencillo que nos permite encontrar la matriz de calibración $\mathbf{P}$ deseada. Sin embargo, es posible obtener la matriz $\mathbf{P}$ ya sea normalizando previamente los puntos o sin normalización de puntos. En el capítulo 5 podremos ver una comparación de la calibración obtenida con o sin normalización de puntos.

Algoritmo 2 Algoritmo DLT, para la estimación de una matriz $\mathbf{P}$ a partir de correspondencias de puntos de 3D a 2D [Hartley y Zisserman(2004)]

Objetivo.

Dadas n correspondencias de puntos del marco de referencia 3D al marco de referencia 2D $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$, $i = 1, 2, \dots, n$ determinar el vector $\mathbf{P}_v$ de calibración de la cámara.

Algoritmo.

1. Normalizar. En el apéndice C se describe como encontrar una transformación de similaridad $\mathbf{T}$, para normalizar los puntos de la imagen y otra transformación de similaridad $\mathbf{U}$ para normalizar los puntos 3D.
2. Solución lineal.
 - a) Para cada correspondencia $\{\mathbf{X}_i \leftrightarrow \mathbf{x}_i\}$ calcular un par de renglones de la matriz $\mathbf{A}$ (ecuación 4.7).
 - b) Unir las $2n \times 12$ ecuaciones en una matriz $\mathbf{A}$.
 - c) Obtener la descomposición de valores singulares de $\mathbf{A}$ (SVD). Donde la solución de $\mathbf{P}_v$ es la última columna de $\mathbf{V}$.
 - d) La matriz $\mathbf{P}'$ es determinada a partir de $\mathbf{P}_v$ (ecuación 4.3).
3. Eliminar normalización. Para eliminar la normalización es necesario aplicar las matrices de similaridad obtenidas en el primer paso a la matriz $\mathbf{P}'$ obtenida en el paso anterior

$$\mathbf{P} = \mathbf{T}^{-1}\mathbf{P}'\mathbf{U}$$

$\mathbf{P}$ es la matriz deseada.

Capítulo 5

Resultados experimentales

El objetivo del capítulo es presentar los resultados más relevantes de esta investigación. Básicamente, el capítulo presenta los resultados de la etapa de experimentación de los capítulos 3 y 4: experimentos de la corrección de la distorsión radial en imágenes tomadas por cámaras, así como la calibración global.

5.1. Introducción

El conjunto de experimentos que se realizaron incluyeron: 1) la construcción de un par de patrones de calibración, 2) La adquisición de imágenes, 3) La obtención de un modelo de corrección de distorsión radial en imágenes tomadas por la cámara y 4) la calibración global de una cámara.

Tomando en cuenta lo anterior el capítulo se divide en tres partes: el material utilizado, la experimentación en la corrección de la distorsión radial en imágenes tomadas por la cámara, así como la experimentación de la calibración global.

5.1.1. Material utilizado

Los experimentos se realizaron en el laboratorio de Sistemas Computacionales del edificio de la División de Estudios de Posgrado de la Facultad de Ingeniería Eléctrica de la Universidad Michoacana de San Nicolás de Hidalgo. Para los experimentos del

presente trabajo utilizamos cámaras Videre Design, Modelo MDCS2. Con un rango de adquisición de imágenes de 15 *fps*, con una resolución máxima de 1280×960 pixeles, una interfaz *firewire IEEE 1394*, con lentes de 3.5 *mm*. En la figura 5.1 se muestra el modelo de cámara que se utilizó.



Figura 5.1: Cámara Videre design, Modelo MDCS2.

Además se utilizaron 2 patrones de calibración que se muestran en la figura 5.2. En la parte inferior de la figura se muestra el patrón plano utilizado para la corrección de la distorsión radial en imágenes tomadas por la cámara. El patrón de calibración de la parte superior (dos planos ortogonales), fue el utilizado para la calibración global de cámaras.

Con estos patrones se realizaron los experimentos que enseguida se describen.

5.2. Corrección de distorsión radial

En esta parte empezamos por tomar una imagen del patrón de calibración plano. En la figura 5.3 se muestra la imagen tomada por la cámara al patrón de calibración. Una prueba que se realizó fue el de comparar los resultados de ambos modelos para la corrección de la distorsión radial, tanto el modelo U-D, así como el modelo D-U. La figura 5.7 presenta los resultados usando ambos modelos para la imagen de entrada que presenta la figura 5.3. Puede observarse que el modelo U-D, descrito en la página 33, no logra corregir la distorsión radial, mientras que el modelo D-U tiene éxito en la corrección.

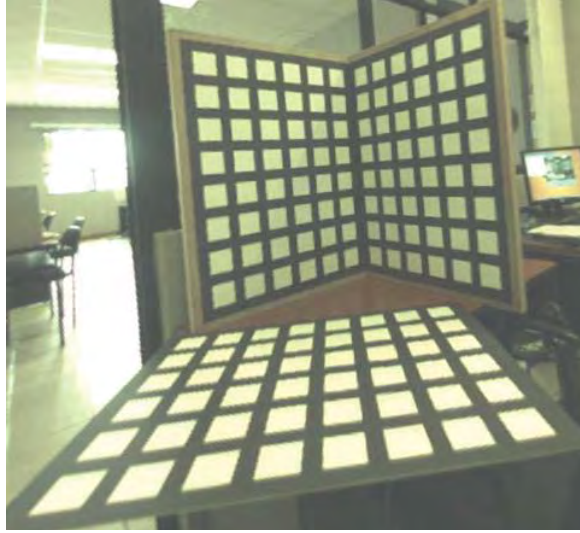


Figura 5.2: Patrones de calibración

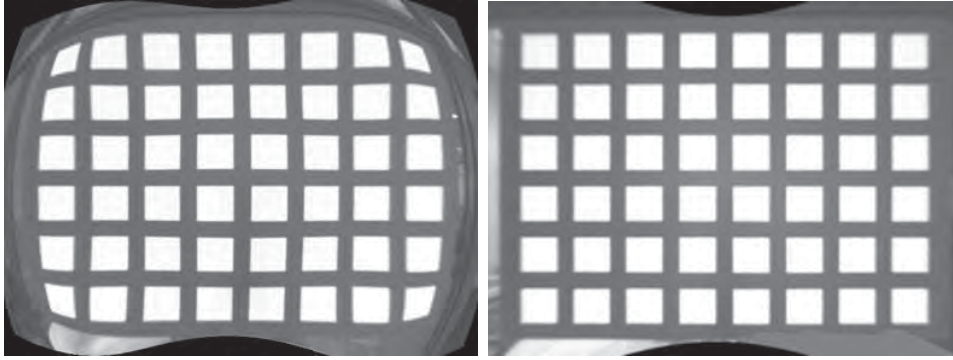
En la figura 5.5 se muestra una imagen de la diferencia que existe entre el patrón de referencia y la imagen corregida con el modelo D-U. Si las imágenes fueran totalmente iguales esta imagen no tendría zonas claras. Las zonas claras indican diferencias entre las imágenes. En adelante todos los experimentos consideran solamente el modelo D-U.

En la tabla 5.1 se presentan los errores calculados para imágenes de diferentes dimensiones. Se presenta una columna “*Error máximo*” que indica la máxima distancia euclidiana entre correspondencias de puntos. La columna “*Error acumulado*” indica la suma de diferencias de distancias al cuadrado de todas las correspondencias. Las “*etapas*” se refieren al número de veces que el método GNLM se aplicó. Lo que determina el número de etapas que se aplicarán a la corrección de las imágenes es la diferencia que existe en el error acumulado. Si el siguiente error es más grande se opta por detener el proceso.

Por último, en la tabla 5.2 se muestran el conjunto de parámetros Θ que se obtuvo para una imagen de dimensión 1280×960 así como para una imagen de dimensión 640×480 . La tabla muestra los 9 parámetros de la matriz de proyección y los 5



Figura 5.3: Imagen tomada con la cámara



(a) Imagen corregida con el modelo U-D (b) Imagen corregida con el modelo D-U

Figura 5.4: Imágenes corregidas utilizando dos modelos de distorsión.

parámetros que determinan la distorsión radial.

Además de probar con los tres parámetros de distorsión radial k_1 , k_2 y k_3 se probó únicamente con dos de ellos: k_1 y k_2 . La figura 5.8 muestra la corrección con solo dos parámetros de distorsión radial, k_1 y k_2 . Asimismo, en la tabla 5.3 se muestra los errores obtenidos en este experimento, tomando en cuenta las diferentes resoluciones de imágenes tomadas por la cámara. Como se puede observar en los resultados de las tablas 5.1 y 5.3 el uso de los tres parámetros produce una mejor corrección de la distorsión radial. Por esta razón se utilizan 3 parámetros en el modelo distorsión

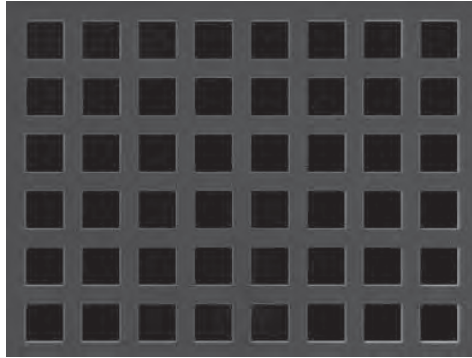


Figura 5.5: Error en la corrección de la distorsión radial.

radial.

5.3. Calibración global

En el capítulo 4 se presentó un algoritmo para calcular la matriz de transformación $\mathbf{P}$, la cual determina la calibración global. En este apartado se muestran los resultados obtenidos en el cálculo de dicha matriz de transformación. En este experimento se considera que las imágenes de la cámara no presentan distorsión radial. Es decir, se ha corregido previamente.

El experimento consiste en tomar un mínimo de 6 correspondencias de puntos entre el patrón de calibración (consistente en dos planos ortogonales) y la imagen tomada por la cámara. Deberán hacerse ciertas consideraciones o configuraciones para el mejor cálculo de la matriz de transformación $\mathbf{P}$:

1. La elección de los puntos deberá incluir puntos en los dos planos ortogonales.
2. Los puntos a considerar deberán ser esparcidos, de tal manera que los puntos no sean colineales.
3. La cámara deberá estar cerca del patrón. Si la cámara se encuentra alejada del patrón es un factor de riesgo para que el sistema de ecuaciones resultante pueda estar degenerado [Hartley y Zisserman(2004)].

Tabla 5.1: Cálculo de errores en la corrección de la distorsión radial.

Imagen de 1280×960			
Etapas	No. Iteraciones	Error máximo	Error acumulado
1 ^a	8	2.039811	1.600898E+05
2 ^a	940	1.953687	1.476844E+05
3 ^a	930	2.322362	1.475795E+05
4 ^a	927	1.704078	1.475528E+05
5 ^a	943	2.043169	1.473519E+05
Imagen de 640×480			
Etapas	No. Iteraciones	Error máximo	Error acumulado
1 ^a	10	1.184061	4.130012E+04
2 ^a	629	1.414867	3.951668E+04
3 ^a	785	1.914177	3.943574E+04
4 ^a	971	1.582132	3.937112E+04

En la tabla 5.4 se muestran las 6 correspondencias de puntos 3D-2D utilizadas para la calibración. En la figura 5.9 se muestra el patrón de calibración indicando con $\mathbf{X}_i$ ($i = 1, \dots, 6$) los puntos 3D utilizados. En la imagen capturada por la cámara se obtienen los correspondientes puntos $\mathbf{x}_i$ ($i = 1, \dots, 6$) (determinados visualmente por el autor de esta tesis). Finalmente la tabla 5.5 muestra tanto los puntos 2D como los puntos 2D calculados usando la matriz $\mathbf{P}$ recién obtenida.

En la tabla 5.6 se muestra el cálculo de los errores obtenidos al utilizar 6 correspondencias de puntos. El “error máximo” representa la mayor distancia entre un punto calculado y su correspondiente punto en la imagen tomada por la cámara. El “error acumulado” indica la sumatoria de distancias al cuadrado. Puede observarse que los errores son pequeños y que la mejora al utilizar la normalización de puntos no es muy relevante.

5.4. Conclusiones

El modelo D-U con parámetros k_1 , k_2 y k_3 permite remover la distorsión radial generada por los lentes de las cámara utilizadas.

Los experimentos fueron realizados en LINUX y el código fuente se encuentra

Tabla 5.2: Resultados del cálculo de los parámetros de calibración de imágenes con distorsión radial.

θ	Dimensión de la imagen	
	1280×960	640×480
m_0	7.43606133E-01	7.25215668E-01
m_1	-4.18113070E-03	-3.54220012E-02
m_2	1.34558534E+02	9.05559980E+01
m_3	-1.47227928E-02	4.37710696E-02
m_4	7.47765117E-01	6.69664211E-01
m_5	1.28616198E+02	9.43844262E+01
m_6	-2.11016884E-05	1.48459733E-04
m_7	-1.05491710E-05	-1.39183006E-04
m_8	1.0	1.0
k_1	1.22245369E-06	4.50886473E-06
k_2	-5.09827015E-13	5.13396034E-12
k_3	6.80909747E-18	3.26886144E-16
c_x	5.11850604E+02	2.43304660E+02
c_y	6.36174397E+02	3.13426139E+02

disponible¹. Se utilizó el lenguaje *C* standard, las bibliotecas GSL para las operaciones con matrices, así como GTK para el diseño de la interfaz de usuario.

¹http://faraday.fie.umich.mx/~m_cgomez/calibrate.html

Tabla 5.3: Cálculo de errores en la corrección de la distorsión radial únicamente tomando en cuenta k_1 y k_2 .

Imagen de 1280×960			
Etapas	No. Iteraciones	Error máximo	Error acumulado
1 ^a	5	3.145792	1.603125E+05
2 ^a	626	2.723954	1.481887E+05
Imagen de 640×480			
Etapas	No. Iteraciones	Error máximo	Error acumulado
1 ^a	5	1.480182	4.138167E+04
2 ^a	647	1.488102	3.956966E+04
3 ^a	647	1.838706	3.952795E+04
4 ^a	651	1.675386	3.952416E+04

Tabla 5.4: Correspondencias de puntos 3D-2D utilizadas en el cálculo de la matriz $\mathbf{P}$.

	Coordenadas de los puntos	
i	3D	2D
1	(5, 0, 5)	(582,685)
2	(90, 0 5)	(136,913)
3	(90, 0,120)	(97, 61)
4	(5, 0,120)	(578,336)
5	(0,90,120)	(1076, 49)
6	(0,90, 5)	(1049,912)

Tabla 5.5: Comparación entre los puntos 2D determinados visualmente y los puntos calculados utilizando la matriz $\mathbf{P}$, los puntos se encuentran normalizados.

	Coordenadas de los puntos	
i	2D determinados visualmente	2D calculados como $\mathbf{x}_i = \mathbf{P}\mathbf{X}_i$
1	(582,685)	(581.606, 685.004)
2	(136,913)	(136.479, 912.977)
3	(97, 61)	(96.4849, 61.0233)
4	(578,336)	(578.405, 335.996)
5	(1076, 49)	(1075.5, 48.9841)
6	(1049,912)	(1049.47, 912.015)



(a) Imagen tomada por la cámara



(b) Imagen corregida

Figura 5.6: Experimento de la corrección radial en diferentes ambientes.



(a) Imagen tomada por la cámara



(b) Imagen corregida

Figura 5.7: Experimento de la corrección radial en diferentes ambientes. Continuación

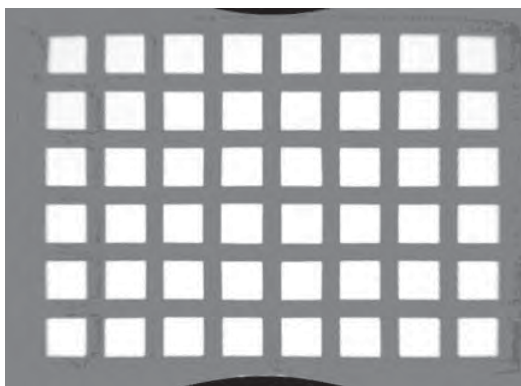


Figura 5.8: Corrección de la distorsión radial utilizando k_1 y k_2

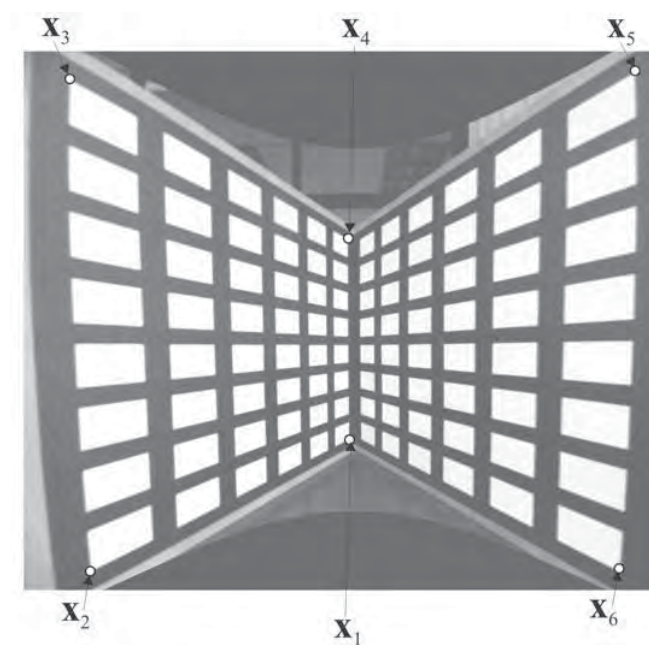


Figura 5.9: Patrón de calibración tridimensional que muestra los puntos que se utilizaron en el cálculo de la matriz $\mathbf{P}$

Tabla 5.6: Resultados del cálculo de la matriz $\mathbf{P}$ en la calibración global de una cámara.

$\mathbf{p}^i$ i-ésima fila	Valores de matriz $\mathbf{P}$			
$\mathbf{p}^1$	-3.388	-1.0875	-9.2910e-02	3.3676e+02
$\mathbf{p}^2$	-1.9555	-1.9857	-1.7115	3.9447e+02
$\mathbf{p}^3$	-3.7550e-03	-3.7854e-03	-1.3421e-04	5.6855e-01
			Error máximo	Error acumulado
Con normalización			0.515631	1.28398
Sin normalización			0.515631	1.28398

Capítulo 6

Conclusiones y trabajos futuros

En esta tesis se abordó el problema de corrección de la distorsión radial y la calibración global de una cámara.

6.1. Conclusiones

En particular se realizó lo siguiente:

1. Se compararon los desempeños de modelos de distorsión, el modelo U-D y el modelo D-U. El modelo D-U resultó adecuado para modelar distorsiones radiales ocasionadas por lentes de amplia apertura en cámaras digitales.
2. Se desarrolló una herramienta de software para la determinación automática del modelo de distorsión a partir de una imagen de referencia y la correspondiente imagen capturada por la cámara. La imagen o patrón de referencia consiste de cuadros blancos sobre un fondo negro, de manera que se facilita el cálculo automático de las correspondencias de puntos, en particular utilizando el centro geométrico de los cuadros blancos. En el apéndice B se describe la herramienta desarrollada.
3. Se desarrolló una biblioteca de funciones para corregir imágenes adquiridas por una o múltiples cámaras. En el apéndice D se presenta el detalle de las funciones desarrolladas.

4. Se desarrolló un simulador didáctico para visualizar el efecto de los parámetros del modelo de distorsión en la corrección de las imágenes. En el apéndice B.7 se muestra el simulador desarrollado, así como algunos ejemplos de uso.
5. Se desarrolló una biblioteca de funciones para calcular la matriz $\mathbf{P}$ de calibración global en base a las correspondencias de puntos 3D-2D. En el apéndice B.6 se muestra el uso de este software.

6.2. Trabajos futuros

Como trabajos futuros se sugiere abordar los siguientes:

1. Utilizar algoritmos paralelos para la corrección rápida de la distorsión radial de una imagen, aprovechando las capacidades de multiprocesamiento de los CPUs más avanzados.
2. Continuar el desarrollo de una herramienta de calibración global donde se automatize la determinación de los puntos característicos de la imagen capturada por la cámara.

Apéndice A

Optimización de la función de error total

A.1. Optimización no lineal

El objetivo de este apéndice es encontrar un modelo que minimice una función objetivo $f(x)$ y encontrar una solución óptima x^* [Nocedal y Wright(1999)], [Ma et al.(2004)]. Esto es:

$$x^* = \underset{x}{\operatorname{argmin}} f(x) \quad (\text{A.1})$$

Enseguida se presenta el caso de búsqueda de un modelo que permita corregir la distorsión radial en imágenes.

En nuestro caso, el objetivo es encontrar un conjunto de parámetros de proyección θ^p y un conjunto de parámetros de distorsión θ^d . Dichos parámetros transformarán la imagen tomada por la cámara I_d , en una imagen I_p . Se trata de obtener un modelo paramétrico que minimice la diferencia entre la imagen I_p y una imagen de referencia sin distorsión I_r . Para hacer esto, utilizamos un conjunto de correspondencias de puntos entre las imágenes I_p e I_r (como es expuesto en la sección 3.5).

Sea n el número de correspondencias. Sea (x_{rk}, y_{rk}) las coordenadas de la k -ésima correspondencia en la imagen sin distorsión. Además sea (x_{dk}, y_{dk}) la k -ésima

correspondencia en la imagen tomada por la cámara. Formulamos la siguiente función objetivo a minimizar:

$$f(x) = E_t(\Theta) = \sum_{k=1}^n (e_{xk}^2 + e_{yk}^2) \quad (\text{A.2})$$

$$\Theta = \underset{\Theta}{\operatorname{argmin}} E_t(\Theta) \quad (\text{A.3})$$

$$e_{xk} = x_{pk}(\theta, x_{dk}, y_{dk}) - x_{rk} \quad (\text{A.4})$$

$$e_{yk} = y_{pk}(\theta, x_{dk}, y_{dk}) - y_{rk} \quad (\text{A.5})$$

Donde $\Theta = \theta^p \cup \theta^d$ y x_{pk} y y_{pk} son las nuevas coordenadas del punto (x_{dk}, y_{dk}) utilizando el modelo Θ .

Para que $E_t(\Theta)$ sea un mínimo se debe de cumplir que $\frac{\partial E_t}{\partial \theta_i} = 0$ para $i = 1, 2, \dots, m$, donde m significa el número de parámetros y n el número de correspondencias de puntos que se tienen para encontrar el argumento mínimo de la función de error.

Es decir para el primer parámetro tendríamos:

$$\frac{\partial E_t}{\partial \theta_1} = 2 \sum_{k=1}^n \left[(x_{pk} - x_{rk}) \left(\frac{\partial x_{pk}}{\partial \theta_1} \right) + (y_{pk} - y_{rk}) \left(\frac{\partial y_{pk}}{\partial \theta_1} \right) \right] = 0 \quad (\text{A.6})$$

Tomando las ecuaciones de A.4 y A.5 y sustituyéndolas en la ecuación anterior obtenemos:

$$f_{\theta_1} = \sum_{k=1}^n \left(e_{xk} \frac{\partial x_{pk}}{\partial \theta_1} + e_{yk} \frac{\partial y_{pk}}{\partial \theta_1} \right) = 0 \quad (\text{A.7})$$

De esta manera y aplicando las derivadas para cada uno de los parámetros en Θ podemos minimizar $E_t(\Theta)$, lo cual implica resolver un sistema de ecuaciones simultáneas para todas las variables.

En este caso tendríamos tantas funciones como variables existan, es decir:

$$\begin{aligned}
f_{\theta_1}(\Theta) &= \frac{\partial E(\Theta)}{\partial \theta_1} = 0 \\
f_{\theta_2}(\Theta) &= \frac{\partial E(\Theta)}{\partial \theta_2} = 0 \\
&\vdots \\
f_{\theta_m}(\Theta) &= \frac{\partial E(\Theta)}{\partial \theta_m} = 0
\end{aligned} \tag{A.8}$$

Para resolver este sistema de ecuaciones se puede utilizar una aproximación con la serie de Taylor, que sería una linealización local [McCallum et al.(1999)]. Considerando únicamente los términos lineales de la serie de Taylor tenemos:

$$\begin{aligned}
f_{\theta_1}(\theta_1 + \delta\theta_1, \theta_2 + \delta\theta_2, \dots, \theta_m + \delta\theta_m) &\approx f_{\theta_1}(\Theta) + \frac{\partial f_{\theta_1}}{\partial \theta_1} \delta\theta_1 + \frac{\partial f_{\theta_1}}{\partial \theta_2} \delta\theta_2 + \dots + \frac{\partial f_{\theta_1}}{\partial \theta_m} \delta\theta_m \\
f_{\theta_2}(\theta_1 + \delta\theta_1, \theta_2 + \delta\theta_2, \dots, \theta_m + \delta\theta_m) &\approx f_{\theta_2}(\Theta) + \frac{\partial f_{\theta_2}}{\partial \theta_1} \delta\theta_1 + \frac{\partial f_{\theta_2}}{\partial \theta_2} \delta\theta_2 + \dots + \frac{\partial f_{\theta_2}}{\partial \theta_m} \delta\theta_m \\
&\vdots \\
f_{\theta_m}(\theta_1 + \delta\theta_1, \theta_2 + \delta\theta_2, \dots, \theta_m + \delta\theta_m) &\approx f_{\theta_m}(\Theta) + \frac{\partial f_{\theta_m}}{\partial \theta_1} \delta\theta_1 + \frac{\partial f_{\theta_m}}{\partial \theta_2} \delta\theta_2 + \dots + \frac{\partial f_{\theta_m}}{\partial \theta_m} \delta\theta_m
\end{aligned} \tag{A.9}$$

Donde m representa el número de parámetros.

Dado que las funciones deben ser igualadas a 0, es posible agrupar términos y representar las ecuaciones de A.9 en forma matricial. Considerando la representación matricial se puede obtener la siguiente expresión:

$$-\begin{bmatrix} f_{\theta_1}(\Theta) \\ f_{\theta_2}(\Theta) \\ \vdots \\ f_{\theta_m}(\Theta) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_{\theta_1}}{\partial \theta_1} & \frac{\partial f_{\theta_1}}{\partial \theta_2} & \dots & \frac{\partial f_{\theta_1}}{\partial \theta_m} \\ \frac{\partial f_{\theta_2}}{\partial \theta_1} & \frac{\partial f_{\theta_2}}{\partial \theta_2} & \dots & \frac{\partial f_{\theta_2}}{\partial \theta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{\theta_m}}{\partial \theta_1} & \frac{\partial f_{\theta_m}}{\partial \theta_2} & \dots & \frac{\partial f_{\theta_m}}{\partial \theta_m} \end{bmatrix} \begin{bmatrix} \delta\theta_1 \\ \delta\theta_2 \\ \vdots \\ \delta\theta_m \end{bmatrix} \tag{A.10}$$

De hecho, se logra obtener una ecuación del tipo $-\mathbf{B} = \mathbf{A}\delta\Theta$. Los elementos de $\mathbf{B}$ se pueden calcular fácilmente:

$$f_{\theta_i} = \sum_{k=1}^n \left(e_{xk} \frac{\partial x_{pk}}{\partial \theta_i} + e_{yk} \frac{\partial y_{pk}}{\partial \theta_i} \right) \tag{A.11}$$

Para encontrar la matriz $\mathbf{A}$ es necesario derivar las ecuaciones del tipo A.11. La primera derivada $\frac{\partial f_{\theta_1}}{\partial \theta_1}$, descartando las segundas derivadas (Método de Gauss-Newton) quedaría como sigue:

$$\begin{aligned}
 \frac{\partial f_{\theta_1}}{\partial \theta_1} &= \frac{\partial}{\partial \theta_1} \left[\sum_{k=1}^n \left(e_{xk} \frac{\partial x_{pk}}{\partial \theta_1} + e_{yk} \frac{\partial y_{pk}}{\partial \theta_1} \right) \right] \\
 &= \sum_{k=1}^n \left[\left(e_{xk} \frac{\partial}{\partial \theta_1} \left(\frac{\partial x_{pk}}{\partial \theta_1} \right) + \frac{\partial x_{pk}}{\partial \theta_1} \frac{\partial e_{xk}}{\partial \theta_1} \right) + \left(e_{yk} \frac{\partial}{\partial \theta_1} \left(\frac{\partial y_{pk}}{\partial \theta_1} \right) + \frac{\partial y_{pk}}{\partial \theta_1} \frac{\partial e_{yk}}{\partial \theta_1} \right) \right] \\
 &= \sum_{k=1}^n \left[\frac{\partial x_{pk}}{\partial \theta_1} \frac{\partial e_{xk}}{\partial \theta_1} + \frac{\partial y_{pk}}{\partial \theta_1} \frac{\partial e_{yk}}{\partial \theta_1} \right] \tag{A.12}
 \end{aligned}$$

De manera similar, para $\frac{\partial f_{\theta_1}}{\partial \theta_2}$ se tendría:

$$\frac{\partial f_{\theta_1}}{\partial \theta_2} = \sum_{k=1}^n \frac{\partial x_{pk}}{\partial \theta_1} \frac{\partial e_{xk}}{\partial \theta_2} + \frac{\partial y_{pk}}{\partial \theta_1} \frac{\partial e_{yk}}{\partial \theta_2}$$

Representando el resultado de las derivadas en la suma de dos matrices, tenemos:

$$\mathbf{A} = \begin{bmatrix} \sum_{k=1}^n \frac{\partial x_{pk}}{\partial \theta_1} \frac{\partial e_{xk}}{\partial \theta_1} & \cdots & \sum_{k=1}^n \frac{\partial x_{pk}}{\partial \theta_1} \frac{\partial e_{xk}}{\partial \theta_m} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n \frac{\partial x_{pk}}{\partial \theta_m} \frac{\partial e_{xk}}{\partial \theta_1} & \cdots & \sum_{k=1}^n \frac{\partial x_{pk}}{\partial \theta_m} \frac{\partial e_{xk}}{\partial \theta_m} \end{bmatrix} + \begin{bmatrix} \sum_{k=1}^n \frac{\partial y_{pk}}{\partial \theta_1} \frac{\partial e_{yk}}{\partial \theta_1} & \cdots & \sum_{k=1}^n \frac{\partial y_{pk}}{\partial \theta_1} \frac{\partial e_{yk}}{\partial \theta_m} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n \frac{\partial y_{pk}}{\partial \theta_m} \frac{\partial e_{yk}}{\partial \theta_1} & \cdots & \sum_{k=1}^n \frac{\partial y_{pk}}{\partial \theta_m} \frac{\partial e_{yk}}{\partial \theta_m} \end{bmatrix} \tag{A.13}$$

De este modo $\mathbf{A}$ se puede agrupar como una suma de dos matrices, como se ve en la siguiente ecuación:

$$\mathbf{A} = \mathbf{A}_x + \mathbf{A}_y \tag{A.14}$$

Ahora, la matriz $\mathbf{A}_x$ puede verse como el producto de dos matrices, como se observa en la siguiente ecuación:

$$\mathbf{A}_x = \begin{bmatrix} \frac{\partial x_{p_1}}{\partial \theta_1} & \frac{\partial x_{p_2}}{\partial \theta_1} & \cdots & \frac{\partial x_{p_n}}{\partial \theta_1} \\ \frac{\partial x_{p_1}}{\partial \theta_2} & \frac{\partial x_{p_2}}{\partial \theta_2} & \cdots & \frac{\partial x_{p_n}}{\partial \theta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_{p_1}}{\partial \theta_m} & \frac{\partial x_{p_2}}{\partial \theta_m} & \cdots & \frac{\partial x_{p_n}}{\partial \theta_m} \end{bmatrix} \begin{bmatrix} \frac{\partial x_{p_1}}{\partial \theta_1} & \frac{\partial x_{p_1}}{\partial \theta_2} & \cdots & \frac{\partial x_{p_1}}{\partial \theta_m} \\ \frac{\partial x_{p_2}}{\partial \theta_1} & \frac{\partial x_{p_2}}{\partial \theta_2} & \cdots & \frac{\partial x_{p_2}}{\partial \theta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_{p_n}}{\partial \theta_1} & \frac{\partial x_{p_n}}{\partial \theta_2} & \cdots & \frac{\partial x_{p_n}}{\partial \theta_m} \end{bmatrix} \quad (\text{A.15})$$

Donde $p_1, p_2, \dots, p_n$ representan las n correspondencias de puntos. De la misma manera podemos agrupar la matriz $\mathbf{A}_y$. Además, con la ayuda de la representación matricial anterior, es posible visualizar la representación del *jacobiano* de $f(x_p)$, el cual es una matriz de derivadas parciales de primer orden representada por una matriz de $n \times m$ definida por [Nocedal y Wright(1999)]:

$$J_x \stackrel{def}{=} \begin{bmatrix} \nabla f_1^T(\mathbf{x}_{p_1}) \\ \vdots \\ \nabla f_m^T(\mathbf{x}_{p_n}) \end{bmatrix} = \begin{bmatrix} \frac{\partial x_{p_1}}{\partial \theta_1} & \cdots & \frac{\partial x_{p_1}}{\partial \theta_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_{p_n}}{\partial \theta_1} & \cdots & \frac{\partial x_{p_n}}{\partial \theta_m} \end{bmatrix} \quad (\text{A.16})$$

donde $\nabla f_1(\mathbf{x}_{p_1})$ representa el gradiente de f_1 en el punto $\mathbf{x}_{p_1}$.

Utilizando la representación de J_x , la ecuación A.14 se reduce a la siguiente:

$$\mathbf{A} = J_x^T J_x + J_y^T J_y \quad (\text{A.17})$$

Resumiendo, de la ecuación A.10 hemos obtenido la matriz $\mathbf{A}$, representada por la ecuación A.17, y la matriz $\mathbf{B}$ representada en la ecuación A.11. Además, si existen n puntos de correspondencias y m parámetros en Θ , $\mathbf{A}$ es una matriz de dimensión $m \times m$ y $\mathbf{B}$ es una matriz de dimensión $m \times 1$. En resumen, los elementos de $\mathbf{A}$ y $\mathbf{B}$ se pueden calcular como sigue:

$$a_{i,j} = \sum_{k=1}^n \left(\frac{\partial x_p}{\partial \theta_i} \frac{\partial x_{pk}}{\partial \theta_j} + \frac{\partial y_p}{\partial \theta_i} \frac{\partial y_{pk}}{\partial \theta_j} \right), \quad b_i = - \sum_{k=1}^m \left(\frac{\partial x_{pk}}{\partial \theta_i} e_{xk} + \frac{\partial y_{pk}}{\partial \theta_i} e_{yk} \right) \quad (\text{A.18})$$

Para simplificar la notación, utilizaremos x_p en lugar de x_{pk} y y_p en lugar de y_{pk} . Las primeras 8 derivadas $\frac{\partial x_p}{\partial \theta_i}$ y $\frac{\partial y_p}{\partial \theta_i}$ (correspondientes al modelo de proyección) se obtienen al derivar las ecuaciones 3.7 de la página 27 con respecto de cada uno de los

parámetros. En seguida se muestran las ecuaciones x_p y y_p , así como cada una de sus derivadas.

$$\begin{aligned}
 x_p &= \frac{h_{11}x_u + h_{12}y_u + h_{13}}{h_{31}x_u + h_{32}y_u + 1} \\
 y_p &= \frac{h_{21}x_u + h_{22}y_u + h_{23}}{h_{31}x_u + h_{32}y_u + 1}
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial x_p}{\partial h_{11}} &= \frac{x_u}{D} & \frac{\partial y_p}{\partial h_{11}} &= 0 \\
 \frac{\partial x_p}{\partial h_{12}} &= \frac{y_u}{D} & \frac{\partial y_p}{\partial h_{12}} &= 0 \\
 \frac{\partial x_p}{\partial h_{13}} &= \frac{1}{D} & \frac{\partial y_p}{\partial h_{13}} &= 0 \\
 \frac{\partial x_p}{\partial h_{21}} &= 0 & \frac{\partial y_p}{\partial h_{21}} &= \frac{x_u}{D} \\
 \frac{\partial x_p}{\partial h_{22}} &= 0 & \frac{\partial y_p}{\partial h_{22}} &= \frac{y_u}{D} \\
 \frac{\partial x_p}{\partial h_{23}} &= 0 & \frac{\partial y_p}{\partial h_{23}} &= \frac{1}{D} \\
 \frac{\partial x_p}{\partial h_{31}} &= \frac{-x_u x_p}{D} & \frac{\partial y_p}{\partial h_{31}} &= \frac{-x_u y_p}{D} \\
 \frac{\partial x_p}{\partial h_{32}} &= \frac{-y_u x_p}{D} & \frac{\partial y_p}{\partial h_{32}} &= \frac{-y_u y_p}{D}
 \end{aligned} \tag{A.19}$$

Donde $D = h_{31}x_u + h_{32}y_u + 1$.

Las derivadas parciales de los 5 parámetros de distorsión se obtienen en base a la ecuación 3.12 de la página 32. Aplicando la regla de la cadena se obtiene:

$$\frac{\partial x_p}{\partial \theta_i} = \frac{\partial x_p}{\partial x_u} \frac{\partial x_u}{\partial \theta_i} + \frac{\partial x_p}{\partial y_u} \frac{\partial y_u}{\partial \theta_i}, \quad \frac{\partial y_p}{\partial \theta_i} = \frac{\partial y_p}{\partial x_u} \frac{\partial x_u}{\partial \theta_i} + \frac{\partial y_p}{\partial y_u} \frac{\partial y_u}{\partial \theta_i} \tag{A.20}$$

$$\begin{aligned}
 \frac{\partial x_p}{\partial x_u} &= \frac{Dh_{11} - (h_{11}x_u + h_{12}y_u + h_{13})h_{31}}{D^2}, \\
 \frac{\partial x_p}{\partial y_u} &= \frac{Dh_{12} - (h_{11}x_u + h_{12}y_u + h_{13})h_{32}}{D^2}, \\
 \frac{\partial y_p}{\partial x_u} &= \frac{Dh_{21} - (h_{21}x_u + h_{22}y_u + h_{23})h_{31}}{D^2}, \\
 \frac{\partial y_p}{\partial y_u} &= \frac{Dh_{22} - (h_{21}x_u + h_{22}y_u + h_{23})h_{32}}{D^2}
 \end{aligned} \tag{A.21}$$

Donde x_u y y_u se definieron en la ecuación 3.14 de la página 32 como:

$$\begin{aligned}x_u &= (x_d - c_x)(1 + k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) + c_x \\y_u &= (y_d - c_y)(1 + k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) + c_y\end{aligned}$$

Donde r_d está representado por:

$$r_d = \sqrt{(x_d - c_x)^2 + (y_d - c_y)^2} \quad (\text{A.22})$$

Finalmente como segundo factor en la regla de la cadena se obtienen las siguientes derivadas:

$$\begin{aligned}\frac{\partial x_u}{\partial k_1} &= r_d^2 (x_d - c_x) \\ \frac{\partial y_u}{\partial k_1} &= r_d^2 (y_d - c_y) \\ \frac{\partial x_u}{\partial k_2} &= r_d^4 (x_d - c_x) \\ \frac{\partial y_u}{\partial k_2} &= r_d^4 (y_d - c_y) \\ \frac{\partial x_u}{\partial k_3} &= r_d^6 (x_d - c_x) \\ \frac{\partial y_u}{\partial k_3} &= r_d^6 (y_d - c_y) \\ \frac{\partial x_u}{\partial c_x} &= -(k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) - 2(k_1 + 2k_2 r_d^2 + 3k_3 r_d^4)(x_d - c_x)^2 \\ \frac{\partial y_u}{\partial c_x} &= -2(k_1 + 2k_2 r_d^2 + 3k_3 r_d^4)(x_d - c_x)(y_d - c_y) \\ \frac{\partial x_u}{\partial c_y} &= -2(k_1 + 2k_2 r_d^2 + 3k_3 r_d^4)(x_d - c_x)(y_d - c_y) \\ \frac{\partial y_u}{\partial c_y} &= -(k_1 r_d^2 + k_2 r_d^4 + k_3 r_d^6) - 2(y_d - c_y)^2(k_1 + 2k_2 r_d^2 + 3k_3 r_d^4)\end{aligned} \quad (\text{A.23})$$

Ahora el método GNLM puede ser aplicado para minimizar la función de error A.2, y resolver $\Theta = (\theta_0, \dots, \theta_{12})$.

Apéndice B

Manual de Usuario

B.1. Descripción

El programa se compone de tres módulos: el módulo de corrección radial en imágenes, la calibración global en cámaras de visión amplia y el módulo del simulador del modelo de distorsión radial. El software se puede bajar de: http://faraday.fie.umich.mx/~m_cgomez/calibrate.html. Cada una de estas partes tiene sus requerimientos propios, así como también tienen requerimientos generales. A continuación los requerimientos generales de las aplicaciones.

B.2. Requerimientos generales

El software presentado es de uso libre basado en el sistema operativo LINUX. Se requiere instalar las siguientes bibliotecas:

1. Bibliotecas libdc1394.- son las bibliotecas propias del uso de cámaras tipo *firewire*. Se utilizan en la captura de imágenes propias de este tipo de interfaz.
2. Bibliotecas GSL.- *GNU Scientific Library (GSL)*. Contiene el conjunto de funciones necesarias para realizar los cálculos con matrices.
3. Bibliotecas GTK.- en el diseño de la interfaz de usuario se utilizaron estas bibliotecas para diseñar una interfaz gráfica.

Además de este conjunto de bibliotecas se utilizaron dos patrones de calibración que describiremos en su correspondiente módulo.

B.3. Instalación General

El software viene en un archivo comprimido de extensión *tgz*. Para su instalación se descomprime ejecutando el comando:

```
$tar -xzf calibrate.tgz
```

El comando creará la organización de las carpetas y descomprimirá los archivos.

Para que la instalación quede completa, únicamente se ejecuta el siguiente comando desde la consola:

```
$make clean all
```

Las siguientes secciones están divididas en los manuales de usuario para la corrección radial, calibración global de las cámaras y el simulador.

B.4. Corrección radial

B.4.1. Requerimientos

La corrección radial en las imágenes tomadas por las cámaras se basa en el uso de un patrón de calibración de un plano. Por un lado se tiene una imagen ideal y por otro la imagen tomada con la cámara al patrón de calibración. Antes de utilizar las herramientas es necesario tener a la mano las imágenes necesarias para su uso. A continuación la descripción de como obtener cada una de las imágenes.

Imagen de referencia

1. Imprimir un patrón de calibración (ver figura B.1) con cuadros blancos y fondo negro usando una impresora láser.

Esta imagen de referencia se guarda en un archivo. En este trabajo nos referimos a esta imagen como la imagen patrón *pattern.pgm*.

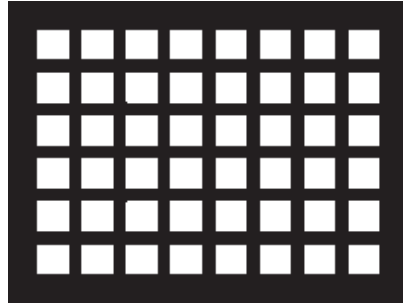


Figura B.1: Patrón de calibración

2. Poner el patrón de calibración en frente de la cámara y obtener una imagen de la cámara.

La imagen puede ser tomada con el programa ***capture***, esta imagen es llamada de aquí en adelante *image_from_camera.pgm*. La imagen requerida deberá estar en formato *PGM*.

B.5. Como ejecutar el programa

Para ejecutar el programa ***calibrate***, el cual calibra la cámara y calcula el conjunto de parámetros (Archivo *final_parameters.dat*) para empatar la imagen de la cámara (*image_from_camera.pgm*) con la imagen patrón (*pattern.pgm*).

Sintaxis:

```
$/calibrate image_from_camera.pgm pattern.pgm initial_parameters.dat final_parameters.dat
```

Donde:

1. *image_from_camera.pgm*.- La imagen adquirida por la cámara.
2. *pattern.pgm*.- La imagen patrón.
3. *initial_parameters*.- Los parámetros iniciales.
4. *final_parameters*.- El nombre del archivo donde los parámetros serán guardados.

El programa **correct** aplica un conjunto de parámetros a una imagen y obtiene una imagen transformada.

Sintaxis:

```
$/correct final_parameters.dat image_in.pgm image_out.pgm
```

Donde:

1. *final_parameters.dat*.- El archivo generado por el programa **calibrate**.
2. *image_in.pgm*.- Una imagen tomada por la cámara.
3. *image_out.pgm*.- La imagen corregida por el programa.

B.6. Calibración Global

B.6.1. Requerimientos

En la calibración global se utiliza un patrón de calibración de dos planos ortogonales como el que se muestra en la figura 4.1. Se toma una imagen al patrón de calibración y con las herramientas de la corrección de la distorsión radial se mejora la imagen. Y por último se localizan el conjunto de correspondencias de puntos que se utilizaran en la calibración global de la cámara. En nuestro caso utilizamos una herramienta de diseño gráfico como es el *GIMP* y localizamos las coordenadas de los puntos correspondientes a los que escojamos en el patrón de dos planos. En las siguientes secciones se muestra un ejemplo de las correspondencias entre una imagen bidimensional y un patrón tridimensional.

B.6.2. Como ejecutar el programa

El programa **getP** se encarga de obtener la matriz de calibración **P**.

Sintaxis:

```
$/getP matrix3D matrix2D
```

Donde:

1. *matrix3D*.- Se refiere al conjunto de puntos en el patrón tridimensional.

2. *matrix2D*.- Conjunto de puntos en la imagen.

A continuación un ejemplo de cada uno de los archivos que lleva por parámetros el programa *getP*.

```
$cat matrix3D
5 0 5
75 0 5
75 0 55
5 0 55
0 75 5
0 75 55
```

```
$cat matrix2D
582 685
136 913
97 61
578 336
1049 912
1076 49
```

Una vez generada la matriz de calibración $\mathbf{P}$, es posible comprobar los resultados. El programa *get_from_PX* obtiene los correspondientes puntos en la imagen dado un conjunto de puntos en el espacio tridimensional. Cabe mencionar que la matriz de calibración global $\mathbf{P}$ se genera en el paso anterior y se guarda en un archivo intermedio. El programa que se describe a continuación toma dicho archivo para el cálculo de puntos en la imagen. Además, utiliza los mismos archivos que usa el programa *getP* para su funcionamiento. Para obtener el conjunto de correspondencias en la imagen utiliza el archivo *matrix3D*. El archivo *matrix2D* sirve para calcular el error entre los puntos calculados y los puntos de referencia.

Sintaxis:

```
$/get_from_PX matrix3D matrix2D
```

B.7. Simulador del modelo de distorsión

El simulador del modelo de distorsión incluye dos programas gráficos que muestran como se afecta una imagen al variar el conjunto de parámetros de distorsión radial.

1. *imagetool*.- herramienta que se utiliza para mostrar el efecto de la modificación de los parámetros de distorsión en imágenes previamente tomadas por la cámara.

2. ***capture***.- herramienta que muestra las imágenes tomadas por la cámara y si se desea puede congelarse la imagen con el fin de modificar los parámetros para observar su efecto en la imagen.

En la figura B.2 se muestra la interfaz que presentan ambos programas que a continuación se describen la manera en que operan.

B.7.1. Como ejecutar los programas

El programa ***capture*** puede usar un conjunto de parámetros de distorsión y proyección como valores iniciales para ser aplicados a las imágenes (con el objetivo de reducir la distorsión).

Sintaxis:

`./capture initial_parameters.dat`

Un ejemplo del archivo *initial_parameters.dat* es:

```
$cat initial_parameters.dat
k1 0.00E+00 1.00E-07
k2 0.00E+00 1.00E-12
k3 0.00E+00 1.00E-17
Cx 6.40E+02 1
Cy 4.80E+02 1
h11 1
h12 0
h13 0
h21 0
h22 1
h23 0
h31 0
h32 0
h33 1
```

El conjunto de parámetros puede ser ajustado mientras el programa está en funcionamiento:

1. El primer paso es congelar la imagen presionando el botón “Freeze”.
2. Puedes modificar los parámetros de distorsión utilizando las barras de desplazamiento y ver los efectos que producen en la imagen.
3. Los nuevos parámetros serán guardados en el archivo *final_parameters.dat* cuando el programa termine.

Nota 1: El archivo de los parámetros de distorsión sigue la siguiente sintaxis: `<name_of_parameter><space><value><space><increment>`. Al principio los parámetros asociados son fijos en `<value>` y el `<increment>` es usado para definir el efecto relacionado a la barra de desplazamiento.

Nota 2: También existe el programa ***imagetool*** que puede ejecutar los mismos tipos de ajustes como ***capture***, pero sobre una imagen precargada desde un archivo, no de la cámara. Esto debería ser útil para encontrar un conjunto de parámetros iniciales, sin tener una cámara conectada a la computadora.

Sintaxis:

`$/imagetool initial_parameters.dat final_parameters.dat`

El archivo *final_parameters.dat* tendrá la forma:

```
$cat final_parameters.dat
k1 4.64E-06
k2 -1.71E-13
k3 3.48E-16
Cx 2.43E+02
Cy 3.13E+02
h11 1
h12 0
h13 0
h21 0
h22 1
h23 0
h31 0
h32 0
h33 1
```

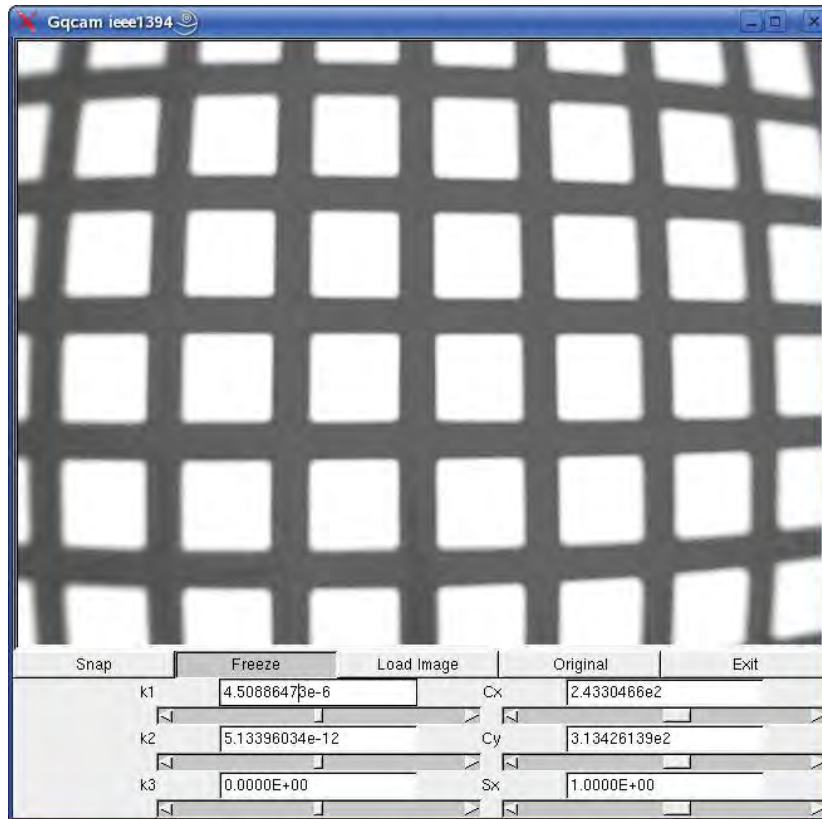


Figura B.2: Interfaz de usuario que muestra el efecto de la corrección de la distorsión radial. Es posible modificar los parámetros de distorsión: k_1 , k_2 y k_3 , el centro de distorsión c_x y c_y , así como un factor de escala s_x . La barra de desplazamiento permite aumentar o disminuir el parámetro correspondiente y en la caja de texto es posible capturar directamente el valor.

Apéndice C

Normalización de puntos

Un proceso típico que mejora las condiciones de las ecuaciones utilizadas en transformaciones de imágenes es la *normalización de los datos* [Hartley y Zisserman(2004)]. Esto, debido a que se utilizan dos marcos de referencia diferentes en los datos que se utilizan para la transformación. Entonces, es conveniente uniformar estos marcos de referencia, de tal forma que el resultado sea invariante a transformaciones de escalamiento.

El algoritmo DLT visto en el capítulo 4 página 51 hace uso del proceso de normalización. Una de las razones para no omitir el proceso de normalización es que permite obtener un resultado invariante a transformaciones de similaridad.

En este apéndice veremos como normalizar los datos de las coordenadas 2D así como las coordenadas 3D.

En el caso de la normalización de los puntos 2D, se traslada y escala un conjunto de puntos homogéneos de tal manera que su *centroide* es el origen y su distancia media desde el origen es $\sqrt{2}$.

Por otro lado en la normalización de los puntos 3D, de manera similar se traslada y escala los correspondientes puntos homogéneos, de tal forma que su centroide es el origen y su distancia media desde el origen es $\sqrt{3}$.

En resumen el proceso de normalización requiere los siguientes pasos:

- Todos los puntos son trasladados, de tal manera que su centroide es el origen.

- Los puntos son escalados de tal manera que la distancia promedio del origen a cualquier punto es de $\sqrt{2}$, en el caso de puntos en imágenes, y $\sqrt{3}$ en el caso de puntos en $3D$.
- Los pasos anteriores son aplicados a los dos conjuntos de datos.

En cada caso se encuentra una matriz de transformación que traslada y escala los puntos [Hearn y Baker(1996)]. Como se muestra en la siguiente ecuación donde un punto $\mathbf{x}_i$ es transformado en un punto $\tilde{\mathbf{x}}_i$.

$$\tilde{\mathbf{x}}_i = \mathbf{T}\mathbf{x}_i \quad (\text{C.1})$$

A continuación explicaremos como encontrar dicha matriz de similaridad en imágenes. Primero, necesitamos una matriz que nos desplace los puntos, de tal manera que su centroide sea el origen. Para esto, tenemos la matriz que nos traslada un punto en coordenadas homogéneas, en las dos direcciones [Hearn y Baker(1996)].

$$\mathbf{x}' = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x} \quad (\text{C.2})$$

donde t_x y t_y representan el desplazamiento del punto en las coordenadas x y y .

Todos los puntos serán desplazados cierta distancia. Sin embargo, debe cumplir que su centroide sea el origen. Entonces, es necesario encontrar una matriz de transformación la cual desplace el centroide al origen.

Sea (x_c, y_c) las coordenadas del centroide de la imagen. La siguiente ecuación desplaza dada una matriz de desplazamiento $\mathbf{T}_d$ un punto x a un punto x' .

$$\mathbf{x}' = \mathbf{T}_d \mathbf{x} \quad (\text{C.3})$$

Por lo tanto, la matriz que desplaza el centroide al origen es:

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} \quad (\text{C.4})$$

El siguiente paso es escalar los puntos de tal manera que la distancia promedio D_p del origen a cualquier punto sea $\sqrt{2}$. La distancia del origen a cualquier punto escalado es:

$$d_i = \sqrt{(ex_i)^2 + (ey_i)^2} \quad (\text{C.5})$$

La matriz que escala un punto es la siguiente [Hearn y Baker(1996)]:

$$\mathbf{x}'' = \begin{bmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}' \quad (\text{C.6})$$

donde e representa el factor de escala.

Sea n el número de puntos que serán escalados, entonces la distancia promedio D_p es:

$$D_p = \frac{1}{n} \sum_{i=1}^n d_i = \sqrt{2} \quad (\text{C.7})$$

Sustituyendo la ecuación C.5 en la ecuación C.7 y despejando e , se tiene:

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n \sqrt{(e^2 x_i^2 + e^2 y_i^2)} &= \sqrt{2} \\
e^2 \left(\frac{1}{n} \sum_{i=1}^n x_i^2 + y_i^2 \right) &= 2 \\
e^2 &= \frac{2}{\left(\frac{1}{n} \sum_{i=1}^n x_i^2 + y_i^2 \right)} \\
e &= \sqrt{\frac{2}{\frac{1}{n} \sum_{i=1}^n x_i^2 + y_i^2}} \quad (C.8)
\end{aligned}$$

Por último, al multiplicar las dos matrices: escalamiento y desplazamiento, obtenemos la matriz de transformación de similaridad $\mathbf{T}$, como se muestra a continuación:

$$\mathbf{T} = \begin{bmatrix} e & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_c \\ 0 & 1 & -y_c \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} e & 0 & -ex_c \\ 0 & e & -ey_c \\ 0 & 0 & 1 \end{bmatrix} \quad (C.9)$$

Y sustituyendo en la ecuación C.1 obtendremos los puntos normalizados.

$$\begin{pmatrix} \tilde{x}_i \\ \tilde{y}_i \\ 1 \end{pmatrix} = \begin{bmatrix} e & 0 & -ex_c \\ 0 & e & -ey_c \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \quad (C.10)$$

De manera similar se encuentra una matriz de transformación $\mathbf{U}$ en los puntos 3D.

$$\widetilde{\mathbf{X}}_i = \mathbf{U} \mathbf{X}_i \quad (C.11)$$

De manera matricial se tiene:

$$\begin{pmatrix} \tilde{x}_i \\ \tilde{y}_i \\ 1 \end{pmatrix} = \begin{bmatrix} e & 0 & 0 & -ex_c \\ 0 & e & 0 & -ey_c \\ 0 & 0 & e & -ez_c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix}$$

En resumen, primero los puntos son trasladados de tal manera que su centroide es el origen. Los puntos son escalados, a fin de que la distancia promedio del origen es $\sqrt{2}$ o $\sqrt{3}$ en el caso 2D o 3D respectivamente. Esto significa que las coordenadas homogéneas del punto promedio serán $(1, 1, 1)^T$ o $(1, 1, 1, 1)^T$.

Esta transformación, que es una transformación de similaridad o escalamiento isotrópico, es aplicada a cada conjunto de coordenadas.

Apéndice D

Códigos

El código desarrollado para la realización de los experimentos es presentado a continuación. No obstante, es posible obtener el código completo, además de algunos ejemplos para probar, en la página web http://faraday.fie.umich.mx/~m_cgomez/calibrate.html. El código fuente se divide en tres partes:

1. La parte de corrección de la distorsión radial.
2. Diseño de una interfaz gráfica para su uso práctico en la corrección de la distorsión radial y
3. La calibración global.

D.1. Corrección de la distorsión radial

El archivo Makefile incluye la secuencia con la cual se compilan las librerías necesarias y el conjunto de programas que se desarrollaron para la parte de experimentación. A continuación describimos la parte que corresponde a la corrección de la distorsión radial en imágenes tomadas por la cámara.

1. `calibrate.c.-` es el programa que se encarga de encontrar el conjunto de parámetros necesarios para la corrección de la distorsión radial.

2. `correct.c`.- programa que corrige la distorsión radial dado un conjunto de parámetros.
3. `libcorrige.a`.- es una librería que incluye las funciones necesarias para la corrección de la distorsión radial.
4. `calibra.c`.- contiene todas las funciones necesarias para la corrección de la distorsión radial.
5. `pointcorresp.c`.- en este archivo se tienen las funciones para encontrar las correspondencias de puntos. Asimismo, las funciones para leer los parámetros iniciales para encontrar el conjunto de parámetros más óptimo. La función principal para aplicar el modelo D-U.

D.2. Diseño de simulador del modelo de distorsión radial

El diseño de la interfaz utiliza como principal archivo *frontend.c*. el cual se basa en el uso de bibliotecas GTK, propias para el desarrollo de interfaz en el sistema operativo *LINUX*.

D.3. Calibración Global

En esta sección encontramos el código fuente de la calibración global de las cámaras. Los archivos que contiene este módulo son:

1. `Makefile`.- archivo que contiene el orden en que se crearán los archivos ejecutables para obtener la matriz $\mathbf{P}$.
2. `obten.c`.- contiene las funciones de normalización de puntos tanto tridimensionales como bidimensionales. Así como el cálculo de la matriz $\mathbf{P}$. Hace uso de las librerías GSL.

3. obten.h.- tiene el conjunto de funciones cabeceras del módulo de calibración global.

Referencias

- [Benhimane y Malis(2004)] Benhimane, Selim y Malis, Ezio. Self-calibration of distortion of a zooming camera by matching points at different resolutions. En *IEEE/RSJ International Conference on Intelligent Robots and Systems*, tomo 2567. Sendai, Japan, 2004.
- [Bombe y Collins(2000)] Bombe, Andreas y Collins, Ben. Proyecto linux1394. 2000.
URL <http://www.linux1394.org>
- [Chapra y Canale(1999)] Chapra, Steven C. y Canale, Raymond P. *Métodos numéricos para ingenieros*. McGraw-Hill, 1999.
- [Dennedy et al.(2000)] Dennedy, Dan, et al. Proyecto libdc1394. 2000.
URL <http://sourceforge.net/projects/libdc1394/>
- [Devernay y Faugeras(1995)] Devernay, F. y Faugeras, O. Automatic calibration and removal of distortion from scenes of structured environments. En *SPIE*, tomo 2567. San Diego, CA, 1995.
URL citeseer.ist.psu.edu/devernay95automatic.html
- [Forsyth y Ponce(2002)] Forsyth, David y Ponce, Jean. *Computer Vision: A Modern Approach*. Prentice Hall, 2002. ISBN 0130851981.
- [Galassi et al.(2004)] Galassi, Mark, et al. GNU scientific library reference manual. 2004.
URL <http://www.gnu.org/software/gsl/>

- [Hartley y Zisserman(2004)] Hartley, R. I. y Zisserman, A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edición, 2004.
- [Hearn y Baker(1996)] Hearn, Donald y Baker, M. Pauline. *Computer Graphics C Version*. Prentice Hall, 1996.
- [Hecht(1987)] Hecht, E. *Optics*. Addison-Wesley, 1987.
- [INFAIMON(2003)] INFAIMON, S.L. Fundamentos teóricos - Ópticas, 2003.
URL <http://www.infaimon.com/catalogo/industria/optica/teoricos/criterios.htm>
- [Iocchi y Konolige(1998)] Iocchi, L. y Konolige, K. A multiresolution stereo vision system for mobile robots. En *AIIA (Italian AI Association) Workshop*. 1998.
- [Jones et al.(1998)] Jones, Joseph L., et al. *Mobile Robots Inspiration to Implementation*. Natick, Massachusetts, 1998.
- [Kortenkamp et al.(1998)] Kortenkamp, David, et al. *Artificial Intelligence and Mobile Robots*. The MIT Press, 1998.
- [Ma et al.(2003)] Ma, Lili, et al. A new analytical radial distortion model for camera calibration. 2003.
URL <http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:cs/0307046>
- [Ma et al.(2004)] Ma, Yi, et al. *An invitation 3D Vision*. Springer, 2004.
- [McCallum et al.(1999)] McCallum, William G., et al. *Cálculo de Varias Variables*. CECSA, 1999.
- [Murphy(2000)] Murphy, Robin. *Introduction to AI robotics*. The MIT Press, 2000.
- [Núñez(2004)] Núñez, Adrián. *Fusión de lecturas de un telémetro láser con un sistema de visión estereo trinocular para detectar obstáculos en tres dimensiones*. Pro-

- yecto Fin de Carrera, Facultad de Ingeniería Eléctrica, Universidad Michoacana de San Nicolás de Hidalgo, 2004.
- [Nocedal y Wright(1999)] Nocedal, Jorge y Wright, Stephen J. *Numerical Optimization*. Springer-Verlag New York, Inc, 1999.
- [Press et al.(1992)] Press, William H., et al. *Numerical Recipes in C - The Art of Scientific Computing*. Cambridge University Press, 1992.
- [Romero y Calderón(2003)] Romero, Leonardo y Calderón, Félix. Correcting radial lens distortion using image and point correspondences. En *8th Iberoamerican Congress on Pattern Recognition, CIARP2003*, páginas 204–211. Universidad Michoacana de San Nicolás de Hidalgo, 2003.
- [Romero y Calderón(2004)] Romero, Leonardo y Calderón, Félix. Non-parametric register as way to obtain an accurate camera calibration. En *MICAI 2004: Advances in Artificial Intelligence*, páginas 584–591. Universidad Michoacana de San Nicolás de Hidalgo, 2004.
- [Romero et al.(2004)] Romero, Leonardo, et al. Fusing a laser range finder and a stereo vision system to detect obstacles in 3D. En *IBERAMIA 2004*, páginas 555–561. Universidad Michoacana de San Nicolás de Hidalgo, 2004.
- [Siegwart y Nourbakhsh(2000)] Siegwart, Rolan y Nourbakhsh, Illah R. *Introduction to Autonomous Mobile Robots*. The MIT Press, 2000.
- [Tamaki(2005)] Tamaki, Toru. Unified Approach to Image Distortion: D-U and U-D Models. *IEICE Trans Inf Syst*, E88-D(5):1086–1090, 2005.
URL <http://ietisy.oxfordjournals.org/cgi/content/abstract/E88-D/5/1086>
- [Tamaki et al.(2000)] Tamaki, Toru, et al. An automatic camera calibration method with image registration. En *IJCV Proc. of SCI2000*, páginas 317–322. 2000.
- [Trucco y Verri(1998)] Trucco, Emanuele y Verri, Alessandro. *Introductory techniques for 3D Computer Vision*. Prentice Hall, 1998.

- [Tsai y Lenz(1988)] Tsai, Roger Y. y Lenz, Reimar K. Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology. *IEEE Trans. Pattern Anal. Mach. Intell.*, 10(5), 1988.
- [Weng et al.(1992)] Weng, Juyang, et al. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(10):965–980, 1992. ISSN 0162-8828.

Glosario

Cuádrica Lugar geométrico de los puntos del espacio cuyas coordenadas cartesianas satisfacen una ecuación de segundo grado; p. ej. Una elipse.

Diafragma Control de la cámara que regula la cantidad de luz que puede entrar para la formación de la imagen.

Difracción Fenómeno característico de las ondas que consiste en la dispersión aparente de las ondas cuando encuentran un obstáculo.

Estereografía Método para representar objetos tridimensionales en un plano por medio de sus proyecciones.

IEEE 1394 Es un bus similar al que utiliza el bus serial USB, pero funciona a velocidades hasta 800 Mbps. Tiene un modo de transmisión que garantiza el ancho de banda, lo cual lo hace ideal para cámaras de video digital y dispositivos similares.

Obturador Determina el tiempo el cual va a recibir la luz una cámara para la formación de la imagen.

Píxel Acrónimo formado con las palabras inglesas “Picture” “Element”. Es el elemento más simple de una imagen digital.

Reflectancia La reflectancia es la proporción de luz reflejada por una superficie y se determina comparando los lúmenes que inciden en ella (iluminancia) con los que refleja (luminancia).

Refracción Es el cambio de dirección que experimenta una onda al pasar de un medio a otro.

Sónar El sónar (acrónimo de Sound Navigation And Ranging) es, básicamente, un sistema de navegación y localización similar al radar pero que, en lugar de emitir señales de radiofrecuencia, emite impulsos ultrasónicos [Jones et al.(1998)].

Telefotografía Arte de tomar y transmitir fotografías a distancia mediante sistemas electromagnéticos.

Translúcido Dicho de un cuerpo que deja pasar la luz, pero que no deja ver nítidamente los objetos.

Índice

- astigmatismo, 20
- Binarización, 38
- Cámara ideal, 11
- Calibración
 - externa, 17
 - Global, 45
 - interna, 17
- Centro de proyección, 12
- Centroide, 83
- Coordenadas
 - homogéneas, 14
 - reales, 14
- Coordenadas homogéneas, 15
- distancia focal, 11
- Distorsión
 - radial, 19, 21
 - tangencial, 21, 22
- Distorsión radial
 - tipo almohadilla, 23
 - tipo barril, 23
- Error
 - óptico, 19
 - cromático, 20
 - de curvatura, 20
 - de distorsión, 20
 - esférico, 19
- Error acumulado, 55
- Error máximo, 55
- escalamiento
 - isotrópico, 87
- Firewire, 75
- Firewire IEEE 1394, 5, 54
- gradiente, 71
- Homografía, 25, 45
- Imágenes por intensidad, 9
- Imágenes por rango, 9
- Imagen virtual, 11
- Interpolación
 - bilineal, 28
- Método Gauss Newton Levenberg Marquardt, 40
- Marco
 - de referencia de la cámara, 15
 - de referencia global, 15
- Marco de referencia
 - de la cámara, 12

global, 15

Matriz

de proyección, 45

de rotación, 16

Modelo

D-U, 33

de perspectiva ideal, 14

U-D, 33

Modelo de distorsión

D-U Distorsión - Sin distorsión, 24

U-D Sin Distorsión - Distorsión, 24

Modelo de pinhole, 11

Normalización, 83

Parámetros

ópticos, 10

extrínsecos, 17

fotométricos, 10

geométricos, 10

intrínsecos, 17

Puntos característicos, 36

reflectancia, 10

Telefoto, 23

Transformación

proyectiva, 25

Vector

de traslación, 16