

# Calibración de parámetros para cámaras CCD: Una guía práctica.

Arturo Espinosa Romero  
 Institute of Perception Action and Behaviour, Division of Informatics,  
 University of Edinburgh  
 E-mail: <arturoe@dal.ed.ac.uk>

Recibido: Mayo de 2001. Aceptado marzo de 2002.

## RESUMEN

La utilización de cámaras electrónicas para la captura y procesamiento de imágenes es cada vez más importante en el estudio de la robótica y la visión por computadora. Con el fin de obtener la relación métrica entre la cámara y la escena a observar, es necesario conocer las características físicas de la cámara. En este artículo se presenta una guía para la estimación de los parámetros internos de cámaras CCD. El método descrito en este trabajo -basado en el trabajo Kanatani- toma en cuenta posibles distorsiones de la imagen producidas por la óptica de la cámara y presenta una caracterización del ruido inherente al dispositivo

*Palabras clave:* Robótica y visión, cámara CCD, calibración, instrumentación.

## ABSTRACT

The importance of the use of electronic cameras for the capture and processing of images in the fields of robotics and computational vision is continuously increasing. With the purpose of estimating the metric relationship between the camera and the scene being observed, it is necessary to know the camera's physical characteristics. In this article a guide to the estimation of the internal parameters of a CCD cameras is presented. The described method -based on the work of Kanatani- takes in account image distortions produced by the camera's optics and presents a characterization of the device's inherent noise.

*Key words:* Robotics & vision, CCD camera, calibration, instrumentation.

## 1 INTRODUCCIÓN

Llamamos calibración de cámaras al proceso mediante el cual encontramos los parámetros internos y/o externos de una cámara; los parámetros internos determinan las características geométricas y ópticas de la cámara, los parámetros externos determinan la orientación y posición de la misma con respecto a un marco de coordenadas. Los parámetros internos y externos son independientes entre sí.

Este artículo tiene como objetivo hacer una presentación básica de la calibración de cámaras, en particular en el cómputo de los parámetros internos de una cámara CCD (*Charged-Coupled Device*). Existen varios métodos para la calibración de cámaras que han sido propuestos en los últimos años: un ejemplo clásico

es el propuesto por Tsai, que propone un método automático en dos etapas para la estimación de los parámetros internos y externos de una cámara.

El método descrito en este artículo está basado en el trabajo de Kanatani [3]; aunque el método descrito no es automático es lo suficientemente simple para poder calibrar la cámara en poco tiempo (una vez que se ha creado el código necesario). También tiene la ventaja de que, puesto que se toma en cuenta la incertidumbre inherente del proceso de captura, es posible obtener una estimación de la incertidumbre de los parámetros estimados; el cálculo de la incertidumbre puede ser de extrema utilidad para ciertas aplicaciones, como la robótica, en donde suele ser necesario definir intervalos de confianza en el movimiento del robot.

El artículo está organizado como sigue: En la primera sección se describe el modelo de la cámara, y cuáles son los parámetros internos de la misma. En la segunda sección se describe el método utilizado para eliminar las distorsiones producidas por el lente de la cámara. En la cuarta sección se introducen algunos conceptos preliminares de geometría proyectiva que son necesarios para definir el método de calibración. En la quinta sección se describe el método utilizado para hallar el centro óptico de la cámara, mientras el método utilizado para averiguar la longitud focal es descrito en la sexta sección.

## 2 Modelo de la cámara.

Los parámetros que definen las características geométricas de la cámara son básicamente la longitud focal  $f$  las coordenadas del centro óptico de la cámara  $(X_0, Y_0)$ , y la razón entre las escalas verticales y horizontales de la imagen<sup>1</sup>  $(k_x, k_y)$ . La relación que describe el modelo de la cámara es descrita mediante la siguiente ecuación lineal:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

(Ec. 1)

en donde  $(x, y)$ , es la proyección del punto en la escena  $P = (X, Y, Z)$  en el plano de la imagen  $I$  (véase la Fig. 1), y los factores escala  $(\alpha_x, \alpha_y)$ , que combinan la longitud focal y la razón de aspecto de la imagen  $(\alpha_x = k_x f, \alpha_y = k_y f)$ . Se asume que el lente de la cámara está enfocado al infinito, por lo que la distancia focal representa la distancia entre el plano de la imagen y el origen (véase la Fig. 1). Las unidades de la distancia focal son las mismas utilizadas para representar las distancias en el espacio polinomiales (e.g. metros) por lo que las unidades de las coordenadas en el plano de la imagen  $(x, y)$  son también representadas en tales unidades. La imagen obtenida con la cámara es una ver-

sión discreta de la imagen proyectada en el sensor de la cámara. Como la información acerca de la densidad de píxeles en el sensor no siempre está disponible, es más conveniente concebir al parámetro  $f$  como un factor de conversión, cuyas unidades son píxeles en lugar de metros. Esta convención se utilizará en el resto de este artículo.

Las características ópticas del modelo de la cámara son descritas como la distorsión producida por el lente de la cámara (distorsión radial y tangencial). Como fue indicado por Tsai [5] en la mayoría de las polinomiales industriales de visión automatizada, únicamente se considera la distorsión radial.

## 3 Eliminación de la distorsión.

El primer paso en el proceso de calibración es la eliminación de las distorsiones en la imagen producidas por el lente de la cámara. La distorsión es modelada con un par de funciones que mapean el píxel  $(x, y)$  en el plano de la imagen capturado por la cámara (imagen distorsionada) a una nueva posición  $(x', y')$ .

En nuestro caso, las funciones se definen como las ecuaciones polinomiales:

$$x' = \sum_{i=0}^m \sum_{j=0}^{m-i} a_{ij} x^i y^j \quad y' = \sum_{i=0}^m \sum_{j=0}^{m-i} b_{ij} x^i y^j$$

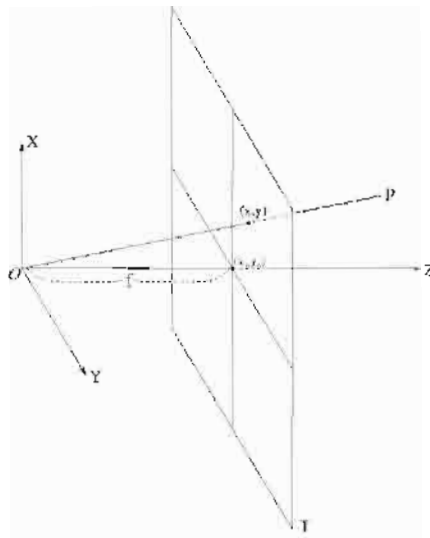
(Ec. 2)

en donde  $m = 2$   $(x, y)$  y  $(x', y')$  y corresponden a las coordenadas de los puntos en la imagen original y la imagen deseada.

Un número  $N$  de coordenadas son obtenidas de ambas imágenes, y al sustituirlas en la Ec. 2 obtenemos un conjunto de ecuaciones lineares con respecto a los coeficientes  $a_{ij}$  y  $b_{ij}$ ; resolviendo dichas ecuaciones bajo un criterio de mínimos cuadrados [1] se halla una aproximación de las funciones que describen el mapeo.

La forma más común de obtener la serie de coordenadas, necesarias en el proceso de eliminación de la distorsión, es utilizar la imagen de una malla regular de cuadrados tomada con la cámara perfectamente perpendicular a ésta (imagen a) en la Fig. 2). Las esquinas de los

<sup>1</sup> Razón de aspecto de la imagen, o "image aspect ratio".



$l$	Plano de la imagen
$(X_0, Y_0)$	Centro óptico
$P$	Punto en la escena
$(x, y)$	Proyección de $P$ en el plano de la imagen $l$
$O$	Origen del sistema de coordenadas de la cámara

Fig.1: Proyección perspectiva de la escena en el plano de la imagen.

cuadrados son usadas como puntos de referencia; para extraerlos se utiliza un detector de bordes Canny [2] (imagen b) en la figura 2), y líneas pueden entonces ser ajustadas a los bordes de cada cuadrado. Las coordenadas de las intersecciones de líneas de cada cuadro se utilizan como una estimación de las coordenadas de las esquinas de los cuadrados, o puntos de referencia. En las imágenes utilizadas en la Fig. 2, cada cuadrado mide 1 cm. por lado, y la distancia entre la cámara y la malla es de 13.5 cm. La imagen ideal fue modelada con una malla regular compuesta por 280 puntos, misma que fue centrada en la malla obtenida de la imagen original. Las coordenadas de los puntos de la malla original y de la malla ideal fueron substituidas en la Ec.2, y el conjunto de parámetros  $a_{ij}$ , y  $b_{ij}$  fueron estimados.

El paso a seguir es usar las ecuaciones para construir un par de mapas de transformación ( $M_x, M_y$ ) que registra la correspondencia entre cada pixel en la imagen original ( $I(x, y)$ ) y la imagen corregida ( $I'(x, y)$ ).

$$I'(x, y) = I(M_x(x, y), M_y(x, y))$$

En el ejemplo mostrado en la Fig. 2 el tamaño original de la imagen obtenida por la cámara es 640 x 480 pixeles; después del mapeo el tamaño de la imagen fue 672 x 487.

Es conveniente, al momento de calcular las funciones que definen los mapas, normalizar las mallas real e ideal de puntos, de tal manera que sus valores varíen entre -1 y 1, para así evitar problemas de precisión. También se puede experimentar con valores pares mayores a 2 para el orden de los polinomios utilizados ( $m=4$  ó  $m=6$ ) en la Ec. 2.

Es también necesario tomar en cuenta que al momento de hacer el mapeo entre la imagen original y la imagen corregida, cierto tipo de interpolación de los valores de la imagen en cada pixel es necesaria. En las imágenes utilizadas en este artículo se utilizó una interpolación bilinear; una muy buena descripción de los posibles tipos de interpolación que pueden utilizarse puede encontrarse en Milan Sonka [5].

#### 4 Conceptos Preliminares.

Esta sección contiene una breve introducción a conceptos básicos de geometría proyectiva y está basada principalmente en el trabajo seminal de Kanatani [3], que resalta de otros trabajos sobre el tema en los que se toma en cuenta la incertidumbre del proceso de captura, lo cual es muy importante para el trabajo que se hace en visión computacional, en donde no siempre se puede asumir una *descripse sensa*. Por motivos de espacio, no se incluyen aquí las demostraciones de lo que en esta sección se afirma, pero

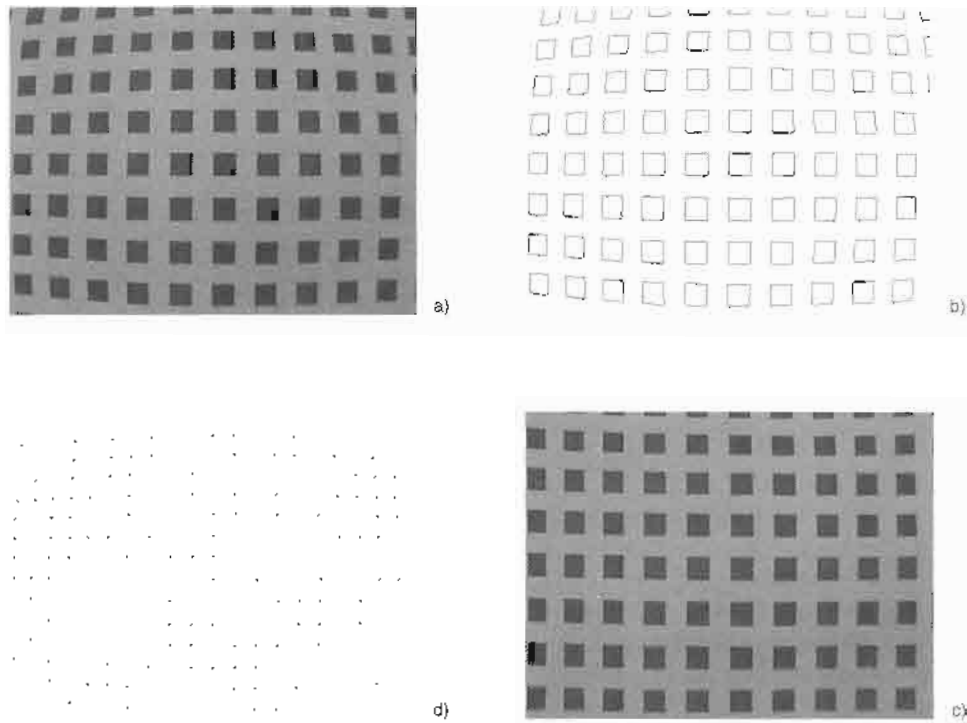


Fig. 2: Proceso de eliminación de la distorsión: la imagen a) muestra la imagen original, b) muestra los bordes obtenidos con el detector de bordes Canny, c) muestra la malla original de puntos que se ajusta a las esquinas de los cuadrados en la imagen b), y d) muestra la imagen corregida.

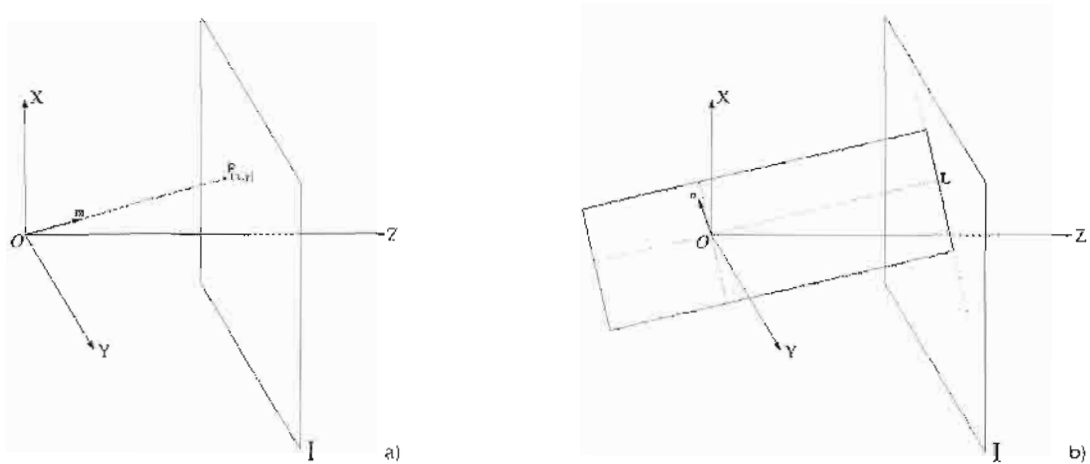


Fig. 3: Coordenadas homogéneas para un punto y una línea en el plano de la imagen.

se pueden encontrar a detalle en el texto de Kanatani.

#### 4.1 Representación de puntos y líneas en el plano de la imagen.

Los puntos y líneas en el plano de la imagen  $I$  son representados por medio de coordenadas homogéneas, que son definidas como una triplete de números reales  $m = (m_1, m_2, m_3)^T$ , en donde no todos ellos son iguales a 0. La relación entre las coordenadas homogéneas  $m$ , de un punto  $P$ , en el plano de la imagen y sus coordenadas cartesianas  $(x, y)$  está dado por

$$x = f \frac{m_1}{m_3}, \quad y = f \frac{m_2}{m_3}$$

(Ec. 3)

Una línea en el plano de la imagen puede definirse por una triplete de números reales  $n = (n_1, n_2, n_3)^T$ , en donde no todos ellos son iguales a 0; la línea representada por  $n$  aparecerá en el plano de la imagen, en las coordenadas cartesianas definidas por

$$n_1 x + n_2 y + n_3 = 0$$

(Ec. 4)

Las coordenadas homogéneas del punto  $P$ , en el plano de la imagen, pueden ser concebidas como el vector que partiendo del origen  $O$  y que apunta hacia el punto  $P$  (véase la imagen a) en la Fig. 4). Las coordenadas homogéneas de la línea  $L$ , en el plano de la imagen pueden, a su vez, ser concebido como el vector normal al plano que pasa por el origen  $O$ , e intersecta la línea  $L$  (véase la imagen b) en la Fig. 4).

Como la relación entre coordenadas homogéneas y su equivalente representación Cartesiana depende únicamente de la orientación del vector, las coordenadas homogéneas pueden ser multiplicadas arbitrariamente por un número diferente a cero, sin perder sus propiedades; con el fin de mantener la magnitud del vector en un rango razonable, la dividimos por su norma. En el resto de este artículo nos referiremos a los puntos y a las líneas en el plano de la imagen, usando su vector normalizado. De acuerdo a esto, un punto en el plano de la imagen con

coordenadas  $(x, y)$  será representado por el vector normalizado

$$n = \pm N \left[ \begin{pmatrix} x \\ y \\ f \end{pmatrix} \right]$$

(Ec. 5)

mientras la línea que yace en el plano de la imagen, descrita por  $Ax + By + C = 0$ , es representada por el vector normalizado

$$m = \pm N \left[ \begin{pmatrix} A \\ B \\ C/f \end{pmatrix} \right]$$

(Ec. 6)

en donde  $f$  representa la longitud focal y  $N$  es el operador de normalización definido como sigue:

$$N(\mathbf{u}) = \frac{\mathbf{u}}{\|\mathbf{u}\|}$$

(Ec. 7)

#### 4.2 Modelo estadístico del ruido.

Cuando la imagen es capturada por la cámara, el ruido inherente del proceso tiene que ser tomado en cuenta, para así poder estimar la confiabilidad de las mediciones utilizadas para estimar los parámetros internos de la cámara (en nuestro caso, puntos y líneas en el plano de la imagen).

El ruido del punto de la imagen  $P$  cuyas coordenadas homogéneas están definidas por el vector  $m$  se define como la matriz de covarianza  $V[m]$  con las siguientes propiedades:

- Es simétrica y positiva semi-definida.
- Es singular con respecto al vector  $m$ :  $V[m]m = 0$
- Su descomposición espectral está definida por

$$V[m] = \sigma_1^2 \mathbf{u}\mathbf{u}^T + \sigma_2^2 \mathbf{v}\mathbf{v}^T + \sigma_3^2 \mathbf{m}\mathbf{m}^T$$

(Ec. 8)

donde  $\sigma_1^2$ ,  $\sigma_2^2$  y 0, ( $\sigma_1 > \sigma_2 > 0$ ) son los tres eigenvalores de  $V[m]$  y  $\{u, v, m\}$  son los eigenvectores correspondientes.

- El RMS de la proyección ortogonal del ruido toma su máximo en la orientación del vector y el mínimo en la orientación del vector
- La magnitud RMS del ruido es

$$\sqrt{\text{tr}(V[m])} = \sqrt{\sigma_1^2 + \sigma_2^2}$$

El modelo del ruido asume que éste ocurre en cada punto de la imagen, y es igualmente probable en cualquier orientación con la misma magnitud RMS (medida en píxeles). Data esta condición la matriz de covarianza  $V[m]$  del vector  $m$  del punto en el plano de la imagen  $P$  a una distancia  $r$  del origen de la imagen está dado por

$$V[m] = \frac{\varepsilon / f^2}{2(1 + r^2 / f^2)} \left( uu^T + \frac{vv^T}{1 + r^2 / f^2} \right) \quad (\text{Ec. 9})$$

donde

$$\mathbf{k} = (0, 0, 1)^T \quad \mathbf{u} = \pm \sqrt{1 + \frac{f^2}{r^2}} \mathbf{m} \times \mathbf{k} \\ \mathbf{v} = \pm \mathbf{u} \times \mathbf{m} \quad (\text{Ec. 10})$$

en donde  $\varepsilon$  representa la magnitud RMS del error, que en nuestro caso definimos como 1.

#### 4.3 Colinealidad de puntos y concurrencia de líneas.

El vector  $m$  del punto de intersección  $P$  de dos líneas en el plano de la imagen  $l$  y  $l'$  representados por los vectores  $n$  y  $n'$  respectivamente está dado por

$$m = N[n \times n'] \quad (\text{Ec. 11})$$

De la misma manera el vector  $n$  de la línea  $L$  que une los puntos  $P$  y  $P'$  representados por los vectores  $m$  y  $m'$  está dado por

$$n = N[m \times m'] \quad (\text{Ec. 12})$$

Cuando tenemos un conjunto de tres o más puntos, decimos que son colineales si hay una línea pasando a través de todos ellos; cuando tenemos un conjunto de tres o más líneas, decimos que son concurrentes si todas ellas comparten el mismo punto de intersección. Formalmente decimos que un conjunto de puntos son colineales si el rango de su conjunto de vectores es menor a tres, y un conjunto de líneas es concurrente si el rango de su conjunto de vectores es menor a tres. El problema con esta definición es que es sumamente sensitiva al ruido; unas definiciones más robustas de concurrencia y de colinearidad son como sigue:

Un conjunto de  $N$  puntos en el plano de la imagen  $P_\alpha$  con coordenadas homogéneas  $m_\alpha$ ,  $\alpha=1 \dots N$  son colineales si el eigenvalor mas pequeño de su matriz de momentos

$$M = \sum_{\alpha=1}^N W_{M\alpha} m_\alpha m_\alpha^T \quad (\text{Ec. 13})$$

donde  $W_{M\alpha}$  son constantes positivas, es igual a cero; el eigenvector asociado  $n$  es el vector de la línea  $L_n$ , que pasa a través de  $P_n$ . De la misma manera, un conjunto de líneas en el plano de la imagen  $L_\alpha$  con coordenadas homogéneas  $n_\alpha$ ,  $\alpha=1 \dots N$ , son concurrentes si el eigenvector mas pequeño de la matriz de momentos

$$N = \sum_{\alpha=1}^N W_{N\alpha} n_\alpha n_\alpha^T \quad (\text{Ec. 14})$$

donde  $W_{N\alpha}$  son constantes positivas, es igual a cero; el eigenvector asociado  $m$  es el vector del punto  $P$  que es la intersección común de las líneas  $L_\alpha$ .

Los conjunto de constantes  $W_{M\alpha}$  y  $W_{N\alpha}$  son pesos que deben ser elegidos de tal manera

que a datos confiables les correspondan pesos grandes mientras a datos no confiables se les asignen pesos pequeños. Los pesos óptimos en el sentido de máxima verosimilitud para las Ec. 13 y 14 están definidos respectivamente por:

$$W_{M\alpha} = \frac{1}{\mathbf{n} \cdot (V[\mathbf{m}_\alpha] \mathbf{n})}$$

(Ec. 15a)

$$W_{N\alpha} = \frac{1}{\mathbf{m} \cdot (V[\mathbf{n}_\alpha] \mathbf{m})}$$

(Ec. 15b)

Como el cómputo de los pesos óptimos involucra el uso del vector  $m$  o  $n$  que queremos averiguar, se utilizan las siguientes aproximaciones

$$W_{M\alpha} = \frac{1}{\text{trace}(V[\mathbf{m}_\alpha])}$$

(Ec. 16a)

$$W_{N\alpha} = \frac{1}{\text{trace}(V[\mathbf{n}_\alpha])}$$

(Ec. 16b)

Una vez que hemos definido la estimación óptima de la línea formada de un conjunto de puntos colineares, o un punto como la intersección de líneas concurrentes, es necesario considerar que el ruido de cada uno de los puntos que forman la línea, o de cada una de las líneas que intersectan un punto se propagará en la nueva línea o en el nuevo punto. Por eso, definimos la matriz de covarianza de la estimación óptima del vector  $m$ , que representa el punto de intersección de  $N$  líneas concurrentes como

$$V[\mathbf{m}] = \frac{\mathbf{u}\mathbf{u}^T}{\lambda_u} + \frac{\mathbf{v}\mathbf{v}^T}{\lambda_v}$$

(Ec. 17)

donde  $\{m, u, v\}$  es el sistema de eigenvectores de la matriz de momentos óptima, con los correspondientes eigenvalores  $\{0, \lambda_u, \lambda_v\}$ . De mane-

ra análoga, la matriz de covarianza de la estimación óptima del vector  $n$ , de la línea común entre  $N$  puntos colineares, esta definida por

$$V[\mathbf{n}] = \frac{\mathbf{u}\mathbf{u}^T}{\lambda_u} + \frac{\mathbf{v}\mathbf{v}^T}{\lambda_v}$$

(Ec. 18)

donde  $\{n, u, v\}$  es el sistema de eigenvectores de la matriz óptima de momentos, con los correspondientes eigenvalores  $\{0, \lambda_u, \lambda_v\}$ .

#### 4.4 Puntos en el infinito.

El punto en el infinito de una línea en el espacio es definido como el límite de la proyección de un punto que se mueve a lo largo de dicha línea; si situamos a una línea en el origen de la cámara  $O$ , el punto en el infinito de tal línea está dado por la intersección de dicha línea con el plano de la imagen. De esta interpretación, se desprende que, las líneas paralelas en el espacio, cuando se proyectan en el plano de la imagen, se intersectarán en un punto común en el infinito.

#### 5 Estimación del centro óptico.

El centro óptico de la cámara, o centro de la imagen, es calculado encontrando el foco de expansión (FDE) de un conjunto de puntos espaciales de referencia, mientras la cámara es trasladada a lo largo de su eje óptico. El FDE se define como el punto en el infinito de las trayectorias de los puntos que se mueven en la escena; si movemos la cámara a lo largo del eje óptico, las trayectorias de los puntos de referencia serán paralelas al eje, y por lo tanto intersectarán al plano de la imagen en el centro óptico.

El procedimiento para estimar el centro óptico será descrito a continuación:

- La cámara se monta en un "photographic copy stand", el cual nos permite controlar la distancia ente la cámara y el objeto que se va a capturar. La cámara se monta apuntando hacia abajo, con su eje principal paralelo al eje de movimiento. Una malla de cuadrados similar a la utilizada para el proceso de eliminación de la distorsión óptica se usa como referencia. Se captura una secuencia de  $N$  imá-

genes, cada imagen se toma con la cámara situada a diferentes distancias de la malla de calibración, y el proceso de eliminación de la distorsión descrito anteriormente, se aplica a cada una de ellas.

- Un número  $M$  de esquinas de la malla de calibración se eligen como puntos de referencia. Las coordenadas de cada punto en el plano de la imagen se obtienen de una manera similar al utilizado en el proceso de eliminación de la distorsión; esto es, ajustando líneas a los bordes de los cuadros en la malla y, tomando la intersección de las líneas, como los puntos de referencia.
- $M$  Líneas se ajustan a los  $M$  conjuntos de  $N$  puntos colineales, que fueron obtenidos de las coordenadas de cada esquina registrada en la secuencia de imágenes. La correspondiente matriz de covarianza de cada línea se calcula usando el método descrito en la sección 4.3 (Ec. 13 y 14).
- Se calcula el punto de concurrencia de las líneas y su matriz de covarianza utilizando el método descrito en la sección 4.3 (Ec. 13 y 14).

#### 6 Estimación de la longitud focal.

Utilizamos el concepto de "conjugacy" para estimar la longitud focal. Decimos que dos puntos  $P$  y  $P'$  con coordenadas homogéneas  $m$  y  $m'$  se conjugan entre ellos si y sólo si

$$m \cdot m' = 0 \quad (\text{Ec. 19})$$

En otras palabras, dos puntos en el plano de la imagen se conjugan entre sí, si sus vectores correspondientes son mutuamente ortogonales. Podemos definir conjugación en términos de coordenadas cartesianas substituyendo la Ec. 3 en la Ec. 19; con ello podemos afirmar lo siguiente: dos puntos en el plano de la imagen, con coordenadas cartesianas  $(a,b)$  y  $(a', b')$  se conjugan entre sí, si y sólo si

$$aa' + bb' + f^2 = 0 \quad (\text{Ec. 20})$$

De la definición de punto en el infinito y el concepto de conjugación podemos deducir lo siguiente: dos líneas espaciales son ortogonales

entre sí, si y sólo si sus puntos en el infinito son conjugados entre si en el plano de la imagen.

El procedimiento utilizado para estimar la longitud focal involucra la estimación de los puntos en el infinito con coordenadas homogéneas  $m$  y  $m'$  de dos líneas espaciales ortogonales proyectadas en el plano de la imagen. Como los puntos en el infinito no necesariamente caen en la imagen capturada por el sensor de la cámara, la mejor manera de representarlos es utilizando sus vectores normalizados. Como coordenadas homogéneas normalizadas no proveen una expresión explícita para su equivalente en el plano de la imagen (si no conocemos el valor de la longitud focal), no podemos utilizar directamente la Ec. 20. En lugar de eso expresamos la conjugación en términos de la razón de cambio del vector normalizado  $m$  al variar la longitud focal del sistema. De la definición de vector normalizado (Ec. 5) obtenemos la siguiente expresión: si  $m$  es el vector que representa un punto en la imagen con respecto a la longitud focal  $f$ , el punto  $m'$  representando el mismo punto con respecto a la distancia focal  $f'$  está dado por

$$m' = \pm N \left[ \begin{array}{c} m_1 \\ m_2 \\ (f'/f)m_3 \end{array} \right] \quad (\text{Ec. 21})$$

Sustituyendo la Ec. 21 en la Ec. 20 obtenemos la siguiente expresión

$$m_1 m_1' + m_2 m_2' + \frac{\hat{f}^2}{f^2} m_3 m_3' = 0 \quad (\text{Ec. 22})$$

donde  $m$  y  $m'$  son los vectores de los puntos en el infinito de dos líneas en el espacio mutuamente ortogonales, definidas con respecto a una longitud focal tentativa  $f$ , mientras  $\hat{f}$  es la verdadera longitud focal. De aquí se obtiene que la verdadera longitud focal se puede determinar usando la siguiente ecuación

$$\hat{f} = f \sqrt{\frac{m_1 m_1' + m_2 m_2'}{m_3 m_3'}} \quad (\text{Ec. 23})$$



El par de puntos en el infinito conjugados puede ser encontrado por medio de la imagen de una superficie rectangular situada en la escena (véase la Fig. 4). El punto en el infinito VP es determinado como la intersección de las líneas que pasan a través de los segmentos AB y DC, y el punto en el infinito VP' es determinado como la intersección de las líneas que pasan a través de los segmentos DC y BC.

Como los puntos en el infinito VP y VP' con coordenadas homogéneas  $m$  y  $m'$  son calculados a partir de una imagen real, una cierta cantidad de ruido, descrita por las respectivas matrices de covarianza  $V[m]$  y  $V[m']$  deben ser consideradas. La varianza de la estimación de la distancia focal está dada por

$$V[f] = \frac{f^2}{4} \frac{m' \cdot (V[m]m') + m \cdot (V[m']m)}{(m_3 m'_3)^2} \quad \dots 24)$$

La confiabilidad de la estimación de la longitud focal también depende de la configuración espacial de las líneas espaciales usadas para determinar los puntos en el infinito; la precisión de la estimación es degradada si los puntos en el infinito encontrados están lejos del centro de la imagen, pero como por definición ellos son conjugados entre sí, hay un número limitado de configuraciones que les permitiera a ambos estar relativamente cerca del centro de la imagen. En lugar de hallar la configuración ideal para poder determinar la longitud focal, la podemos estimar como el promedio ponderado sobre  $N$  mediciones obtenidas de imágenes diferentes, en donde se prueban distintas configuraciones. La longitud focal óptima estimada  $\bar{f}$  se define como

$$\bar{f} = \sum_{\alpha=1}^N W_{\alpha} f_{\alpha} \quad \sum_{\alpha=1}^N W_{\alpha} = 1 \quad (\text{Ec. 25})$$

donde  $f_{\alpha}$ ,  $\alpha = 1, \dots, N$  son los estimados de la longitud focal sobre  $N$  imágenes diferentes.

Los pesos se definen en términos de la varianza de las estimaciones independientes de la longitud focal  $V[f_{\alpha}]$  como

$$W_{\alpha} = \frac{1}{V[f_{\alpha}]} / \sum_{\beta=1}^N \frac{1}{V[f_{\beta}]} \quad (\text{Ec. 26})$$

y la varianza del estimado óptimo de la longitud focal está dada por

$$V[\bar{f}] = 1 / \sum_{\alpha=1}^N \frac{1}{V[f_{\alpha}]} \quad (\text{Ec. 27})$$

De la varianza calculada de la estimación óptima de la longitud focal se puede calcular a un intervalo de confianza. La estadística  $(f - \bar{f}) / \sqrt{V[\bar{f}]}$  donde  $f$  es la verdadera longitud focal, obedece a la distribución normal estándar. Un intervalo de confianza  $(100 - \alpha)\%$  puede ser definido como

$$\left[ \bar{f} - \lambda_{\alpha} \sqrt{V[\bar{f}]}, \bar{f} + \lambda_{\alpha} \sqrt{V[\bar{f}]} \right] \quad (\text{Ec. 28})$$

donde  $\lambda_{\alpha}$  es el punto  $\alpha\%$  en la distribución normal estándar.

## 7 Procedimiento para la estimación de la longitud focal.

El procedimiento para estimar la longitud focal se resume en las siguientes líneas:

- Se capturan  $N$  imágenes del marco de calibración con configuraciones espaciales diferentes, y la distorsión óptica producida por el lente se elimina.
- Los bordes del marco de calibración se extraen utilizando un detector de bordes Canny (véanse imágenes d)-f) en la Fig. 5).
- Las líneas se ajustan a los bordes del marco de calibración y sus matrices de covarianza se estiman utilizando el método descrito en la sección (ecuaciones 13, 15a, 15b y 18).
- Los puntos en el infinito y sus respectivas matrices de covarianza se estiman utilizando las ecuaciones (véanse imágenes g)-i) en la Fig. 5).
- La longitud focal se estima utilizando el método descrito en la sección 6.

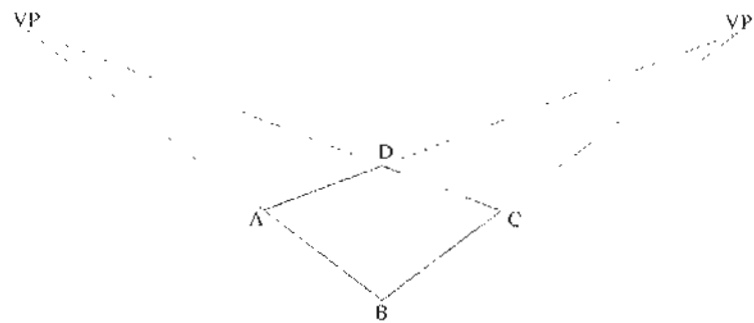


Fig. 4: Proyección de un rectángulo en la escena en el plano de la imagen. La intersección de las líneas espaciales  $AB, DC$  y  $AD, BC$  definen el par de puntos en infinito mutuamente conjugados  $VP$  y  $VP'$ .

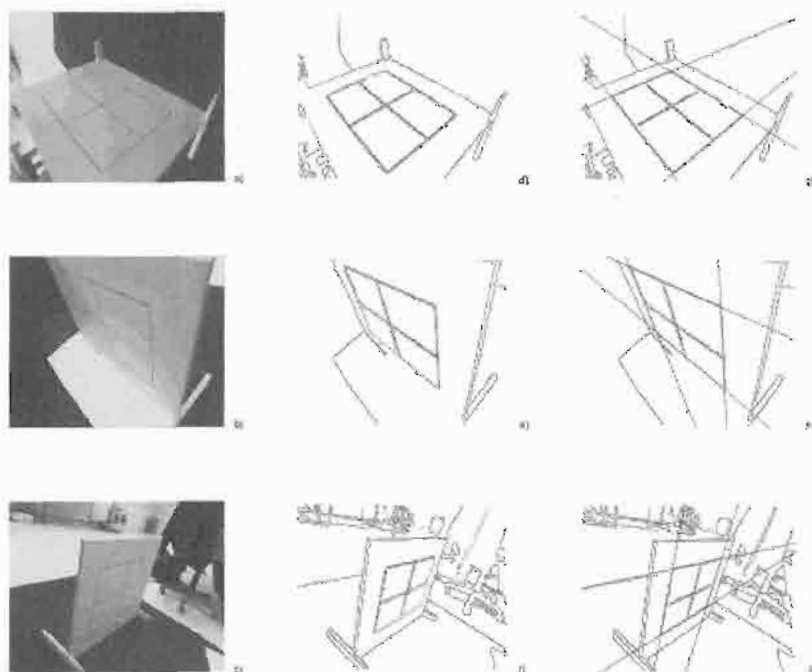


Fig. 5: Imágenes a)-c) muestran el mismo marco de calibración visto desde diferentes posiciones y orientaciones, imágenes d)-f) muestran las imágenes resultantes de aplicar el detector de bordes Canny a las figuras a)-c), imágenes g)-i) muestran el rectángulo ajustado al marco de calibración en rojo, y en azul los dos pares de líneas cuyos puntos en el infinito son usados para estimar la distancia focal.

#### REFERENCIAS

1. Richard L. Burden y J Douglas Faires *Análisis Numérico*, Grupo Editorial Iberoamericana, 1985.
2. J. F. Canny. A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679-698, Noviembre 1986.
3. Kenichi Kanatani. Geometric Computation for Machine Vision, Núm. 37, en *Oxford Engineering Sciences Series*. Oxford Science Publications, 1993.
4. Milan Sonka, Vaclav Hlavac, y Roger Boyle. *Image Processing, Analysis and Machine Vision*, Chapman & Hall, 1993.
5. Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses, *IEEE Journal of Robotics and Automation*, RA-3(4):323-344, August 1987.