

## PROCEDIMIENTO COMPLETO PARA EL CALIBRADO DE CÁMARAS UTILIZANDO UNA PLANTILLA PLANA

**Carlos Ricolfé Viala, Antonio José Sánchez Salmerón**

*Departamento de Ingeniería de Sistemas y Automática  
Universidad Politécnica de Valencia  
Camino de Vera s/n, 46022 Valencia  
Email: cricolfe, asanchez@isa.upv.es*

**Resumen:** El proceso de calibración de una cámara es necesario para la obtención de información 3D a partir de imágenes 2D de la escena. Existen diferentes técnicas basadas en fotogrametría y auto-calibración. Como resultado se obtienen los parámetros intrínsecos y extrínsecos de la cámara. Mucho trabajo se ha realizado en el calibrado de cámaras y también en el pre y post procesado de los datos. Desde un punto de vista práctico, la bibliografía apunta al método basado en plantilla bidimensional como el más fácil de realizar y que mejores resultados obtiene. Sin embargo, no se han tenido en cuenta las mejoras que diferentes técnicas de pre y post procesado de datos pueden aportar a dicho método.

En este artículo se realiza un estudio de las diferentes técnicas de pre y post procesado de datos que mejoran el proceso de calibración de Zhang (Zhang, 1998, 2000). Estas técnicas condicionan de forma óptima los procesos de cálculo, eliminan ruidos en las coordenadas de los puntos y realizan una búsqueda no lineal adecuada al conjunto de parámetros de la cámara. Con este artículo se pretende definir el procedimiento completo para calibrar una cámara utilizando una plantilla plana y conseguir resultados óptimos con el proceso. *Copyright © 2008 CEA-IFAC*

**Palabras Clave:** calibración de cámaras, plantilla plana, rectificado de distorsión, normalización de los datos, calibración no lineal.

### 1. INTRODUCCIÓN

El proceso de calibración de una cámara es un paso necesario para obtener medidas de la escena a partir de imágenes de la misma. La exactitud de la calibración determinará posteriormente la precisión de las medidas que se realicen a partir de las imágenes. Es por este motivo que es imprescindible realizar la calibración de la cámara con plenas garantías de que los parámetros obtenidos son los más parecidos a los reales. Este compromiso implica tanto la elección del método de calibración así como la correcta utilización del mismo. Así pues, el proceso de calibrado se debe empezar realizando una exhaustiva revisión del estado del arte de los diferentes métodos de calibración para elegir el que mejores resultados obtiene y bajo que condiciones se consiguen estos resultados. Debido a la gran cantidad

de trabajos realizados en el campo de la calibración de cámaras, resulta una tarea ardua y poco sencilla la elección del método y las condiciones para desarrollarlo. Existen numerosas técnicas de pre y post procesado de los datos que pueden mejorar los resultados de la calibración pero puede suceder que no se apliquen porque no se encuentran en los artículos que describen un método de calibración. Por un lado existen trabajos relacionados con métodos de calibración, y por otro, estudios que tratan sobre el pre y post procesado de los datos. Por lo tanto, es necesario tener un conocimiento exhaustivo del estado del estado del arte para incluirlas dentro del procedimiento de calibración de forma correcta.

El presente artículo pretende definir un procedimiento completo de calibración que tiene en cuenta todas las técnicas que de forma genérica se pueden encontrar en el estado del arte de calibrado de

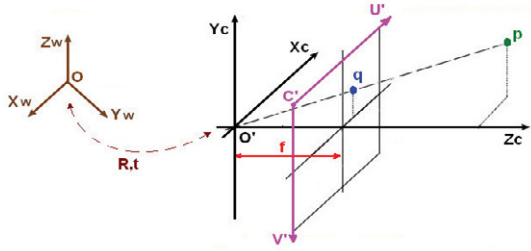


Figura 1. Modelo geométrico de la cámara

cámaras y que han sido revisadas por los autores. De esta forma, para un investigador novel en el mundo del calibrado de cámaras, resulta sencillo y rápido conocer el proceso completo de calibración de la cámara de acuerdo a un procedimiento que obtiene óptimos resultados.

En primer lugar se hace una breve descripción del modelo de la cámara. Después se tratan los diferentes métodos de calibración que existen y se elige el que mejores resultados obtiene. En tercer lugar se establecen las mejoras que se pueden introducir a este método realizando un pre procesamiento de los datos y planteando un índice correcto de búsqueda no lineal. En cuarto lugar, se establece el procedimiento con todas las etapas a seguir para calibrar la cámara, se realizan experimentos reales y simulados y se finaliza con unas conclusiones.

## 2. EL MODELO DE LA CÁMARA

El modelo básico de la cámara también llamado 'pin-hole' representa la transformación de las coordenadas de los puntos de la escena en las coordenadas de la imagen. Se muestra en la figura 1. El modelo lineal viene definido por la siguiente expresión.

$$q_i = \lambda \cdot K \cdot [R \quad t] \cdot p = \lambda \cdot K \cdot [r_1 \quad r_2 \quad r_3 \quad t_1] \cdot p_i \quad (1)$$

$q_i = (u_i, v_i, h_i)$  representa las coordenadas homogéneas en la imagen de un punto en la escena  $p_i = (x_i, y_i, z_i, 1)$ . El modelo contiene la orientación de la cámara en la escena y su posición expresadas con una matriz de rotación  $R$  y un vector de traslación  $t = (t_x, t_y, t_z)$  respectivamente, además de las características intrínsecas de la cámara contenidas en la matriz  $K$ .

### 2.1 Modelo de distorsión de la cámara.

Las imperfecciones en las formas de las lentes provocan desviaciones laterales del haz de luz, que las atraviesa. Como resultado se obtiene una posición del punto observado en la imagen diferente a su posición esperada como reflejo de un punto del espacio. La posición correcta en píxeles del punto  $q$  en la imagen  $q_p = (u_p, v_p)$  está relacionada con la posición observada del mismo  $q_d = (u_d, v_d)$ . Esta relación se expresa mediante las siguientes expresiones:

$$\begin{aligned} u_d &= u_p - \delta_u(u_p, v_p) \\ v_d &= v_p - \delta_v(u_p, v_p) \end{aligned} \quad (2)$$

Según (2) la cantidad de error geométrico en cada coordenada del punto  $q$  depende la posición del mismo dentro de la imagen. Según (Brown, 1966), (Faig, 1975), (Stama, *et al.*, 1980),  $p_1, p_2, p_3, \dots$  modela la distorsión tangencial, la radial se modela con los coeficientes  $k_1, k_2, k_3, \dots$ , y la de prisma con  $s_1, s_2, s_3, \dots$ . La distorsión efectiva puede ser modelada por la suma de los tres tipos de distorsiones definidas anteriormente. Si se descartan los términos de orden superior a 2 se obtiene la siguiente expresión.

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

Esta expresión representa el modelo no lineal de la cámara.  $r$  es la distancia de un píxel al punto principal de la imagen,  $c = (u_0, v_0)$ ,  $\Delta u = u_p - u_0$ ,  $\Delta v = v_p - v_0$ .  $r$  se calcula como  $r^2 = \Delta u^2 + \Delta v^2$ .

## 3. CALIBRACIÓN DE CÁMARAS

El proceso de calibración consiste en obtener los parámetros que definen el modelo de la cámara, tanto lineal como no lineal. En la actualidad existen numerosos métodos de calibración que utilizan diferentes tipos de plantillas y también sin utilizar ninguna de ellas. El método de Tsai (Tasi, 1997), representa un proceso clásico de calibración basado en las medidas de las coordenadas de los puntos de una plantilla 3D respecto a un punto de referencia fijo. Este método ha sido ampliamente utilizado en el siglo pasado (Kanade, *et al.*, 1997), (Raskar, *et al.*, 1998). En la comparación de métodos de calibración desarrollados entre 1982 y 1998 realizado por Salvi (Salvi, *et al.*, 2002), el método de Tsai muestra mejores resultados a pesar de que para conseguir buenos resultados es necesaria una precisión importante de los datos de entrada. Por el contrario, el método de Zhang (Zhang, 1998, 2000), el cual no está incluido en el estudio de Salvi (Salvi, *et al.*, 2002), representa una nueva era en el proceso de calibrado de la cámara. Este método utiliza las coordenadas de los puntos situados en una plantilla plana 2D. De esta forma se combinan las ventajas de los métodos de calibración basados en las medidas de las coordenadas de la plantilla con las ventajas de la auto calibración en la cual no es necesaria utilizar plantilla. Este modo de calibración resulta muy flexible desde el punto de vista de que tanto la cámara como la plantilla puede ser movida libremente y además se pueden tomar tantas imágenes como se quieran sin tener que volver a realizar medidas en la plantilla. Sun (Sun y Cooperstock, 2006) utiliza los resultados de Salvi (Salvi, *et al.*, 2002) y compara el método de Tsai con el método de Zhang. Por un lado el método de Tsai obtiene una estimación precisa de los parámetros de la cámara si los datos de entrada están poco contaminados con ruido. Teniendo en cuenta que son

necesarios al menos cien puntos en la plantilla y que las coordenadas han de estar referidas a un origen de coordenadas fijo, es imprescindible un adecuado diseño de la plantilla de calibración además de una medida exacta de las coordenadas de los puntos. A pesar de todo, la probabilidad de cometer errores en las medidas son altas tal y como confirman los experimentos realizados por Sun (Sun y Cooperstock, 2006). Por el contrario, el método de calibración de Zhang basado en plantilla 2D no requiere tal especial diseño de la plantilla, ni tampoco una medición tan exacta de los puntos de la misma. Sun realiza los experimentos con una plantilla hecha a mano y obtiene mejores resultados que con el método de Tsai. Además, la sensibilidad del algoritmo de calibración frente a errores en las medidas, puede ser mejorada aumentando el número de puntos en la plantilla, simplemente imprimiendo un tablero de ajedrez con más esquinas. Los resultados de la comparación muestran la flexibilidad y adaptabilidad del método de calibración de Zhang ya que se puede realizar en cualquier escenario.

Desde el punto de vista del modelo estimado, los experimentos con cámaras de calidad media realizados por Salvi (Salvi, *et al.*, 2002) y Sun (Sun y Cooperstock, 2006) indican que es razonable considerar distorsión tangencial ó desalineamiento cero y distorsión radial de segundo orden. Considerar una distorsión radial de orden cuatro puede ser factible si se disponen de datos con poco nivel de ruido. En caso de estimar distorsiones de orden seis o superiores el propio ruido de las medidas hace que los resultados obtenidos sean peores, además de complicar el algoritmo de calibración. Si se deja de considerar la distorsión radial, se pueden mejorar los resultados de la calibración añadiendo componentes de distorsión tangencial ó descentrado de la imagen.

### 3.1 Calibración lineal.

Así pues, basándose en las conclusiones obtenidas por Sun (Sun y Cooperstock, 2006) se establece como método base de calibración el que utiliza un plantilla plana presentado por Zhang (Zhang, 1998, 2000). Este método consiste en tomar varias imágenes de una plantilla plana desde diferentes posiciones y orientaciones de la cámara. Con cada una de estas imágenes es posible obtener una homografía  $H$  resolviendo la siguiente expresión:

$$\begin{bmatrix} p_i^T & 0 & -u_i q_i^T \\ 0 & p_i^T & -v_i q_i^T \end{bmatrix} h = 0 \quad (4)$$

$$A \cdot h = 0$$

Si se calculan  $n$  homografías a partir de  $n$  imágenes obtenidas de la plantilla se obtiene un sistema de ecuaciones de la forma  $V \cdot b = 0$ , donde  $V$  es una matriz de dimensiones  $2n \times 6$ , formada con los elementos de las homografías de la siguiente forma:

$$v_{ij}^T = [h_{11}h_{j1} \quad h_{11}h_{j2} + h_{12}h_{j1} \quad h_{12}h_{j2} \quad \dots \\ \dots \quad h_{i3}h_{j1} + h_{i1}h_{j3} \quad h_{i3}h_{j2} + h_{i2}h_{j3} \quad h_{i3}h_{j3}] \quad (5)$$

Si  $n \geq 3$  se tiene una solución general con solución única de  $b$  definida con un factor de escala. La solución del vector  $b$  es el vector propio de  $V^T \cdot V$  asociado al valor propio más pequeño. A partir del vector estimado  $b$  se obtienen los parámetros intrínsecos de la cámara. Una vez conocidos los parámetros intrínsecos se calculan los extrínsecos a partir de cada homografía. Ver (Zhang, 1998) para más detalles.

### 3.2 Calibración no lineal.

El algoritmo descrito en la subsección anterior obtiene los valores de los parámetros de la cámara de forma lineal. Si se realiza una estimación no lineal de los parámetros, se minimiza un error residual definido por la expresión  $e = b^{\#} - f(a)$ , donde  $a$  representa el vector de parámetros que deben satisfacer el vector de medidas  $b^{\#}$  con la función no lineal  $f$ . ( $^{\#}$  representa datos contaminados con ruido). El error se minimiza empezando con un valor de los parámetros y buscando por el espacio del vector de parámetros, siguiendo la dirección que minimiza el Laplaciano de  $f$  en el valor actual de los parámetros en cada iteración. El valor inicial para empezar la búsqueda de los parámetros se obtiene por métodos lineales.

En el caso del calibrado de la cámara, el error  $e$  representa la distancia geométrica entre los puntos medidos en la imagen  $q_i^{\#}$  y las coordenadas de los puntos proyectados con la plantilla de calibración y el valor actual de los parámetros  $q_i^* = M^{\#} p_i^{\#}$ . La función  $f$  se puede definir de diferentes formas y siempre expresa la relación entre coordenadas de los puntos en la imagen y la plantilla. Normalmente, los parámetros del modelo lineal de la cámara se utilizan para definir la función  $f$  y de esta forma se calculan directamente (también se pueden incluir los parámetros del modelo no lineal). Por lo tanto,  $a = (\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \phi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \phi_m, \psi_m)$ .  $m$  es el número de imágenes que participan en el proceso de calibración y  $n$  corresponde con el número de puntos. El espacio de búsqueda es  $4+6m$  donde 4 son los parámetros intrínsecos y 6 los extrínsecos. Los 6 elementos de la matriz de rotación se expresan con ángulos de Euler  $\theta, \phi, \psi$ . El índice a minimizar es:

$$I = \sum_{j=1}^m \sum_{i=1}^n \left( u_{ij} - \alpha_u \frac{r_{11j}x_i + r_{12j}y_i + r_{13j}z_i + t_{1j}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{3j}} - u_0 \right)^2 + \left( v_{ij} - \alpha_v \frac{r_{21j}x_i + r_{22j}y_i + r_{23j}z_i + t_{2j}}{r_{31j}x_i + r_{32j}y_i + r_{33j}z_i + t_{3j}} - v_0 \right)^2 \quad (6)$$

## 4. MEJORAS DEL MÉTODO DE CALIBRACIÓN

En el apartado anterior se describe una técnica lineal y no lineal para calibrar la cámara usando una

plantilla plana. Los parámetros óptimos  $\mathbf{p}^*$  satisfacen el modelo de la cámara para un conjunto de medidas  $\mathbf{q}_i^*$ ,  $\mathbf{p}_i^*$ . El conjunto de parámetros  $\mathbf{p}^*$  se obtienen partiendo de los valores iniciales de los parámetros  $\mathbf{p}_0$  cercanos a los definitivos siempre que el ruido que contaminaba las medidas  $\mathbf{q}_i^\#, \mathbf{p}_i^\#$  no sea excesivo. Los valores iniciales de los parámetros  $\mathbf{p}_0$  se obtienen en la etapa lineal. Con el objetivo de mejorar la estimación lineal de los parámetros de la cámara y obtener un conjunto de parámetros  $\mathbf{p}_0$  más cercanos a los óptimos  $\mathbf{p}^*$ , se va a realizar un preprocesado de los datos con técnicas desarrolladas por Hartley (Hartley, 1997) y por Ahmed (Ahmed y Farag, 2005) o Devernay (Devernay y Faugeras, 2001). Estas técnicas reducen notablemente el nivel de ruido de las medidas y se obtiene unos parámetros más exactos en la etapa lineal de cálculo.

Hartley realiza una normalización de las medidas de las coordenadas para obtener una matriz  $\mathbf{A}$  en (4) mejor condicionada y en consecuencia obtener unos valores más estables en el cálculo de los vectores propios de dicha matriz. Ahmed (Ahmed y Farag, 2005) o Devernay (Devernay y Faugeras, 2001) realizan una estimación de la distorsión que producen las lentes en la imagen y una posterior corrección de la misma. Esta técnica permite reducir el nivel de ruido en las medidas de las coordenadas de la imagen, ya que se corrigen las posiciones de los puntos en la misma y además, los parámetros estimados para la distorsión producida por la lente sirven de punto de partida para la búsqueda no lineal. El objetivo es introducir estos dos pasos previos antes de realizar el proceso de calibrado.

Después, en la etapa de búsqueda no lineal de los parámetros de la cámara, el índice de búsqueda se modifica para mejorar los resultados. En este caso, las coordenadas de los puntos en la plantilla se incluyen dentro del vector de parámetros para que la propia búsqueda no lineal tenga en cuenta su incertidumbre y corrija a su vez las posiciones de los mismos. Dado que las coordenadas de los puntos en la plantilla son corregidos para reducir el nivel de ruido, los errores de calibración sólo dependerán del nivel de ruido en las coordenadas de los puntos en la imagen. Esta técnica la presento (Lavest, *et al.*, 1998).

#### 4.1 Calibración de la distorsión de la cámara.

En este apartado se describe el método para calibrar los parámetros que modelan la distorsión de la cámara presentados en el apartado 2.1. Se considera la distorsión radial y tangencial principalmente, aunque la distorsión de descentrado también se puede considerar. Si se corrige la distorsión antes de realizar el calibrado de la cámara, el proceso de calibración mejora sus resultados. En este caso se ha utilizado el método descrito por Ahmed (Ahmed y Farag, 2005). Se basa en el principio de que una imagen de cualquier línea recta, siempre será recta. Ahmed expresa la distorsión en

función de la deformación de las líneas rectas. El método se describe en (Ahmed y Farag, 2005).

#### 4.2 Normalización de los datos.

La estabilidad en los resultados de la calibración de la cámara depende de las posiciones de los puntos en la imagen (Hartley y Zisserman, 2004). Formalmente se podría decir que si se tiene un conjunto de puntos  $\mathbf{p}_i$  en la escena, con sus correspondientes puntos  $\mathbf{q}_i$  en la imagen y las posiciones de estos puntos cambian de forma que  $\mathbf{p}_i^\circ = \mathbf{T}_p \cdot \mathbf{p}_i$  y  $\mathbf{q}_i^\circ = \mathbf{T}_q \cdot \mathbf{q}_i$ , donde  $\mathbf{T}_p$  y  $\mathbf{T}_q$  son sendas matrices de transformaciones de similitud, si los dos conjuntos de puntos iniciales están relacionados por una homografía  $\mathbf{H}$  de la forma  $\mathbf{q}_i = \mathbf{H} \cdot \mathbf{p}_i$ , esta expresión se transforma utilizando el nuevo conjunto de puntos en  $\mathbf{q}_i^\circ = \mathbf{T}_q \cdot \mathbf{H} \cdot \mathbf{T}_p^{-1} \cdot \mathbf{p}_i^\circ$ . Esta ecuación implica que la matriz de proyección  $\mathbf{H}^\circ$  que relaciona los conjuntos de puntos transformados  $\mathbf{p}_i^\circ$  y  $\mathbf{q}_i^\circ$  es  $\mathbf{H}^\circ = \mathbf{T}_q \cdot \mathbf{H} \cdot \mathbf{T}_p^{-1}$ . Teniendo esto en cuenta, si la situación de los puntos en el espacio no influye en la estimación de la homografía  $\mathbf{H}$ , se puede obtener la misma por un camino alternativo. Éste consistiría en transformar el conjunto de puntos iniciales  $\mathbf{p}_i$  y  $\mathbf{q}_i$  en  $\mathbf{p}_i^\circ = \mathbf{T}_p \cdot \mathbf{p}_i$  y  $\mathbf{q}_i^\circ = \mathbf{T}_q \cdot \mathbf{q}_i$  a partir de sendas transformaciones de similitud  $\mathbf{T}_p$  y  $\mathbf{T}_q$ . Después se calcula la homografía  $\mathbf{H}^\circ$  por los métodos explicados anteriormente basados en la minimización del error algebraico y a partir de ésta, obtener la homografía  $\mathbf{H}$  según  $\mathbf{H} = \mathbf{T}_q^{-1} \cdot \mathbf{H}^\circ \cdot \mathbf{T}_p$ , la cual, se hubiera obtenido directamente utilizando el conjunto de puntos iniciales  $\mathbf{p}_i$  y  $\mathbf{q}_i$ . Este hecho sin embargo, en la realidad no es cierto. Hartley (Hartley, 1997), demuestra que los resultados de los algoritmos basados en la minimización del error algebraico, son sensibles a la situación de los puntos en el espacio. Esto es debido a que en el momento de establecer las restricciones dentro del índice no es lo mismo fijar  $|\mathbf{h}|=1$  que  $|\mathbf{h}^\circ|=1$  y por lo tanto los resultados obtenidos por un camino son diferentes a los obtenidos por el camino inmediato.

La normalización de las coordenadas que Hartley propone en (Hartley, 1997), consiste en trasladar las mismas de forma que el centroide de todos los puntos esté situado en el origen de coordenadas del sistema de referencia. La traslación que se realiza es diferente en las coordenadas de los puntos de la plantilla, que en los puntos de la imagen. Además las coordenadas son escaladas de forma que la media de las distancias de todos los puntos al origen de coordenadas sea 1. El escalado en cada eje de coordenadas es igual y se realiza un escalado diferente para los puntos de la imagen y del escenario.

#### 4.3 Modificación del índice de búsqueda no lineal.

Lavest en (Lavest, *et al.*, 1998) calibra la cámara utilizando una plantilla tridimensional corrigiendo los puntos en la plantilla  $\mathbf{p}_i^\#$  ya que están contaminados con ruido. En cada iteración del proceso de búsqueda no lineal, un nuevo conjunto de puntos en la plantilla son estimados  $\mathbf{p}_i^*$ , los cuales

tienen un menor nivel de ruido que  $p_i^\#$  ya que generan un menor error geométrico. En este caso, las coordenadas de los puntos en la plantilla  $p_i^\#$ , se incluyen con los parámetros de la cámara que se van a estimar. Debido a que se está utilizando el método presentado por Zhang (Zhang 1998), la plantilla es plana y se tienen varias imágenes de la misma desde diferentes posiciones y orientaciones. Para utilizar toda la información, un índice con los parámetros intrínsecos de la cámara y los extrínsecos para cada una de las posiciones es minimizado. Además las coordenadas de los puntos en la plantilla se incluyen. El vector de parámetros queda según  $a=(\alpha_u, \alpha_v, u_0, v_0, t_{x1}, t_{y1}, t_{z1}, \theta_1, \varphi_1, \psi_1, \dots, t_{xm}, t_{ym}, t_{zm}, \theta_m, \varphi_m, \psi_m, x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n)$ .  $m$  es el número de imágenes de la plantilla y  $n$  corresponde con el número de puntos. La dimensión del espacio de búsqueda es  $4+6m+3n$  y el índice es el mismo que el que se definió en (6).

## 5. PROCEDIMIENTO DE CALIBRACIÓN

El procedimiento de calibración basado en plantilla plana debe seguir los siguientes pasos:

- Diseñar una plantilla tipo tablero de ajedrez con 70 esquinas.
- Tomar 10 imágenes de la plantilla desde varias posiciones y orientaciones. Las imágenes se tomarán de forma que la plantilla abarque la mayor superficie de la imagen.
- Detectar las esquinas de los cuadrados del tablero en las imágenes a nivel de subpixel.
- Corregir la distorsión de la imagen con el método descrito por Ahmed (Ahmed y Farag, 2005).
- Normalizar las coordenadas de los puntos en las imágenes y la plantilla según el método de Hartley (Hartley, 1997)
- Calcular una homografía para cada imagen.
- Con las homografías obtenidas, aplicar el método de calibrado de Zhang (Zhang, 1998, 2000).
- A partir de los resultados obtenidos realizar una búsqueda no lineal de los parámetros según el índice descrito en el apartado 4.3.

## 6. RESULTADOS EXPERIMENTALES

Para demostrar las mejoras que se producen sobre la calibración de la cámara, se van a realizar varios experimentos simulados y con cámaras reales. Las condiciones bajo las cuales se van a realizar los experimentos son parecidas a las realizadas Zhang en (Zhang, 1998) para comparar resultados. Según se muestra en la figura 3 de (Zhang, 1998), los resultados de calibración mejoran si los ángulos que forman el plano de la platilla con el los planos de las imágenes está comprendido entre 50 y 70 grados. Respecto al número de imágenes necesarias para realizar una calibración fiable, la figura 2 de (Zhang, 1998) muestra que a partir de 7 u 8 imágenes la

Figura 2. Efectos de la normalización y la corrección de la distorsión en la estimación lineal de los parámetros.

mejora en los resultados de la calibración ya no es significativa. Por lo tanto, para ver el efecto de la normalización de los datos y la corrección de la distorsión en el proceso de calibración de la cámara, se van a utilizar 8 imágenes tomadas de forma que el ángulo que forman el plano de la platilla con los respectivos planos de la imagen está comprendido entre 50 y 70 grados.

Las posiciones de la cámara son  $t_1=(150, 200, 580)$ ,  $t_2=(-50, 250, 880)$ ,  $t_3=(100, -20, 820)$ ,  $t_4=(-40, -150, 780)$ ,  $t_5=(-150, -150, 530)$ ,  $t_6=(-100, 125, 400)$ ,  $t_7=(140, -150, 500)$ ,  $t_8=(240, 50, 600)$ . La orientación de la cámara en cada una de las posiciones es tal que el eje óptico corta al plano de la plantilla en el centro de la misma.

Por otro lado, con el objetivo de poder comparar resultados simulados con los obtenidos por Zhang en (Zhang, 1998), las características de la cámara son las mismas que la simulada en el mencionado artículo. Éstas son  $\alpha_u=1250$ ,  $\alpha_v=900$ ,  $u_0=250$ ,  $v_0=250$  y  $\gamma=1.09083$  equivalente a  $89.95^\circ$ . La resolución de la imagen es  $512 \times 512$  píxeles. La plantilla, de la misma forma que hace Zhang es un tablero de ajedrez de  $10 \times 14 = 140$  esquinas. El tamaño de la plantilla es  $180 \times 250$  mm. En este caso, la plantilla está situada en el suelo de forma que todos los puntos de interés tienen coordenada  $z_w=0$ .

### 6.1 Efectos de la distorsión en la calibración.

Los efectos de la corrección de la distorsión de la cámara son evaluados en primer lugar. La cámara se calibra con puntos contaminados con distorsión extraídos directamente de las imágenes de la plantilla y con puntos cuya distorsión se ha corregido. Dado que se utilizan datos simulados se pueden variar fácilmente el nivel de ruido en las coordenadas de los puntos. Para realizar las simulaciones lo más parecidas a la realidad los valores de  $k_1$ ,  $k_2$ ,  $p_1$  y  $p_2$  son similares al de una cámara real. Estos corresponden a la cámara VCM50 de  $640 \times 480$  píxeles de resolución. Los valores de distorsión calculados son  $k_1=1.1834 \times 10^{-7}$ ,  $k_2=0.3872 \times 10^{-7}$ ,

(a) (b)

Figura 3. Efectos del ruido de la plantilla en la calibración no lineal. (a) los puntos de la plantilla están contaminados con ruido de  $\sigma=1\text{mm}$  (b) los puntos en la imagen están contaminados con ruido de  $\sigma=0.5$  píxeles.

(a) (b)

Figura 4. Errores cuando el número de puntos (a) y número de imágenes (b) cambia. Los puntos de la plantilla no se perturban y en la imagen tiene un nivel de ruido de  $\sigma=1$ .

$p_1=3.658 \times 10^{-7}$ ,  $p_2=0.74138 \times 10^{-7}$ ,  $c_u=312.5686$ ,  $c_v=248.0743$ . La figura 2 muestra los resultados de  $\alpha_u$  utilizando imágenes corregidas y no corregidas. Obviamente, se obtienen mejores resultados con imágenes corregidas. En caso de calibrar la cámara con métodos lineales corregir la distorsión resulta imprescindible. Si se realiza una calibración no lineal, la corrección de la imagen se debe realizar si no se tienen en cuenta los parámetros de distorsión en la búsqueda no lineal. Además se obtendrán mejores resultados si la imagen no está distorsionada. Los resultados obtenidos son parecidos para el resto de parámetros de la cámara  $\alpha_v$ ,  $u_0$ ,  $v_0$ .

## 6.2 Normalización de los datos.

Con la normalización de los datos se consigue mejorar la condición de la matriz  $A$ . Esto significa que el cálculo de las homografías es más estable frente al nivel de ruido en las medidas. En consecuencia se mejora la estimación de los parámetros de la cámara cuando la contaminación de los datos es importante. En este caso se ha simulado el efecto del ruido en las medidas, normalizando los datos y comparando los resultados con los obtenidos en el caso de la corrección de la distorsión de la

imagen. Las condiciones de simulación son las mismas que las anteriores. Los resultados mostrados para cada nivel de ruido son la media de 500 calibraciones. Estos se muestran en la figura 2. Se observa que efectivamente, si el nivel de ruido aumenta, la normalización de los datos realiza un filtrado del mismo y se consiguen mejores resultados. La mejoría es siempre relativa a si se ha realizado o no, la corrección de la distorsión de la imagen. Evidentemente, los mejores resultados se obtienen si se corrige la distorsión de la imagen y se normalizan los datos. Si el nivel de ruido no es muy alto, una solución aceptable para el cálculo lineal de los parámetros de la cámara sería solamente corrigiendo la distorsión de la imagen.

## 6.3 Búsqueda no lineal de parámetros.

Como último paso del proceso de calibración de la cámara se considera la búsqueda no lineal de los parámetros. Se parte de un punto de búsqueda inicial y se realiza una búsqueda de los parámetros de forma que se minimiza un índice planteado. En este caso, se van a minimizar dos de los índices planteados en el apartado 3.2 y 4.3. Los puntos iniciales de búsqueda son los calculados con la distorsión corregida en la

Tabla 1. Resultados de calibración para la cámara real VCM50

	Posiciones aleatorias			
	Calibración lineal	Preprocesado de datos + calibración lineal	Calibración NO lineal	Calibración NO lineal con puntos en la plantilla
$\alpha_u$	749.2±45.2	748.4±31.2	749.4±15.2	751.7±9.8
$\alpha_v$	752.2±52.4	749.1±42.5	740.1±22.5	743.1±14.7
$u_0$	318.5±5.9	328.5±5.1	319.4±5.1	321.7±3.6
$v_0$	237.9±8.7	245.9±8.1	245.9±7.1	246.2±3.2
Error estimación	0.345±0.059	0.232±0.093	0.153±0.045	0.053±0.012
Tiempo cómputo (s)	0.002	0.043	28.743	48.483

imagen y las coordenadas de los puntos normalizadas. Esta solución al estar más cercana a la definitiva, el algoritmo de búsqueda converge más rápidamente y la búsqueda no lineal es más corta.

La figura 3(a) muestra los resultados de la búsqueda no lineal de los parámetros de la cámara variando el nivel de ruido en las coordenadas de los puntos en la imagen. Las coordenadas de la plantilla se contaminan con ruido Gaussiano de  $\sigma=1$ . Se compara el caso de la búsqueda no lineal de sólo los parámetros de la cámara, y de los parámetros de la cámara junto con las coordenadas de los puntos en la plantilla. Se puede observar que el algoritmo es capaz de corregir las coordenadas de los puntos en la plantilla y obtener resultados casi exactos en los que el error de la calibración sólo depende del nivel de ruido en los puntos de la imagen. En el caso de aumentar el ruido en la imagen, el algoritmo que calcula posiciones de los puntos en la plantilla llega a conclusiones menos parecidas a las reales ya que se entiende que no existe referencia para corregir las coordenadas de los puntos en la plantilla.

La figura 3(b) presenta los efectos del error de medida cometidos en la plantilla. Si se realiza una búsqueda no lineal de los parámetros junto con el conjunto de puntos en la plantilla, se demuestra que el error cometido en la calibración sólo depende de las imprecisiones realizadas en el proceso de medida en los puntos en la imagen. En este caso, sólo dependerán del nivel de ruido en las medidas de la imagen. Esta mejora sólo se consigue si se realiza la etapa de búsqueda no lineal de los parámetros de la cámara con los puntos de la plantilla. Si se resuelve sólo la fase lineal del calibrado de la cámara, el ruido de los puntos en la plantilla sí afectará a la calibración.

#### 6.4 Efecto del número de puntos y de imágenes.

En este caso se simulan los efectos del número de datos que participan en el proceso de calibración

de la cámara. Se varían tanto el número de puntos en la plantilla así como las imágenes que se toman de las mismas. En el primer caso, se incrementa el número de puntos contaminando las coordenadas de los puntos en la imagen y en la plantilla con un ruido de  $\sigma=1$ . El número de imágenes que se toman son ocho al igual que en las experiencias simuladas anteriormente. La figura 4(a) muestra los resultados. Efectivamente contra más datos participan en la estimación, menores son los errores que se comenten en el cálculo de los parámetros. Como barrera significativa respecto del número de datos necesarios a ubicar en la plantilla, se podría decir que en torno a 60~70 puntos podría ser suficiente.

En segundo lugar se cambian el número de imágenes que se toman de la plantilla. En este caso los puntos que se utilizan son los 99 de la plantilla y el número de imágenes varía desde 3 que son las mínimas necesarias hasta 14. En la figura 4(b), sólo se muestra a partir de 5 imágenes para poder apreciar mejor los resultados. Con un número menor de imágenes, los resultados son extrapolables. Las posiciones desde donde se toman las imágenes son  $t_1=(150, 200, 580)$ ,  $t_2=(-50, 250, 880)$ ,  $t_3=(100, -20, 820)$ ,  $t_4=(-40, -150, 780)$ ,  $t_5=(-150, -150, 530)$ ,  $t_6=(-100, 125, 400)$ ,  $t_7=(140, -150, 500)$ ,  $t_8=(240, 50, 600)$ ,  $t_9=(50, -250, 880)$ ,  $t_{10}=(-240, -50, 600)$ ,  $t_{11}=(-140, 150, 600)$ ,  $t_{12}=(125, 20, 800)$ ,  $t_{13}=(25, -100, 700)$ ,  $t_{14}=(-25, 100, 600)$ .

En todos los casos se orienta la cámara para que el eje óptico coincida con el centro de la plantilla. Igual que antes, contra más imágenes se toman menos error se comete. Como barrera significativa del número de imágenes necesarias, se observa que con más de 8 imágenes la mejora de los resultados ya no resulta tan significativa. Estos resultados también se pueden observar en las experiencias realizadas por Zhang en (Zhang, 1998) y Sun en (Sun y Cooperstock, 2006).

### 6.5 Calibración de cámaras reales.

La cámara VCM50 de Vision Components GMBH se ha calibrado. Esta tiene una distancia focal fija y una resolución de 640x480 píxeles. La plantilla de calibración es un tablero de ajedrez de 15x11 esquinas de 280x200 mm impreso con una impresora láser. La rigidez se garantiza ya que se ha pegado a un tablero. Se han tomado 20 imágenes desde posiciones aleatorias para calibrar la cámara y 10 imágenes para testear los resultados. Los errores de calibración se han medido con la media de las distancias geométricas entre el conjunto de puntos medidos en las imágenes  $q_i^{\#}$  y los puntos de la plantilla  $p_i^{\#}$ , proyectados con los parámetros de la cámara estimados. Estos puntos se llaman  $q_i^*$ . El proceso de calibración se realiza con diferentes números de puntos e imágenes. También se realiza el preprocesado y la búsqueda no lineal de los parámetros incluyendo y sin incluir las coordenadas de los puntos de la plantilla. Para calibrar la cámara, se utilizan 10 imágenes y 100 puntos. Varias calibraciones se han realizado utilizando datos de diferentes posiciones aleatorias. La media y desviación típica de los parámetros de la cámara calibrados se muestran en la tabla 1.  $e$  representa el error geométrico entre los puntos  $q_i^{\#}$  y  $q_i^*$ , y  $t$  es el tiempo de computo empezando cuando ya se han detectado los puntos en las imágenes. Se ha calibrado utilizando un procesador AMD Athlon 64 a 3200 Mhz. Se observa que el preprocesado de los datos incrementa la robustez de la calibración. La calibración lineal tiene siempre un mayor error geométrico que la calibración no lineal. Además, si se incluyen el conjunto de puntos en la plantilla dentro de la búsqueda no lineal, los resultados que se obtienen son mucho más estables. No obstante, el tiempo de cálculo en caso de realizar una búsqueda no lineal crece considerablemente cuando se aumenta el número de parámetros buscados. Sin embargo, dado que la calibración es un proceso fuera de línea no resulta crítico que el tiempo de cómputo aumente, por lo que es totalmente válido incluir el conjunto de puntos en la plantilla.

## 7. CONCLUSIONES

El proceso de calibración consiste en estimar tanto la posición y orientación de la cámara en la escena, así como sus características intrínsecas. La robustez de la calibración depende de la calidad de las medidas, el modelo que se calcula, el método de calibración y las condiciones bajo las que se realiza. Varios tipos de plantillas se pueden utilizar incluso realizar una auto calibración en la cual no se utiliza la plantilla. Para un investigador ajeno al mundo de la calibración resulta complicado decidir que procedimiento utilizar para calibrar la cámara y bajo que condiciones. A partir de un estudio exhaustivo del estado del arte de calibrado de cámaras, se ha definido un procedimiento de calibración que obtiene buenos resultados. Éste está basado en el método de Zhang

que utiliza una plantilla de calibración plana. A este método se ha añadido un preprocesado de los datos y una mejora en la búsqueda no lineal de los parámetros. El preprocesado de los datos consiste en corregir los errores en las coordenadas de los puntos provocados por la distorsión de la cámara y la normalización de los mismos para mejorar la condición del proceso de calibración. Respecto al paso de búsqueda no lineal de los parámetros, es muy importante incluir las coordenadas de los puntos en la plantilla ya que de esta forma se disminuye el ruido que contamina los mismos y se mejora la calibración. De esta forma se establece un procedimiento de calibración fácil de resolver que garantiza unos óptimos resultados de la calibración.

## REFERENCIAS

- Ahmed M., Farag A. (2005) Non metric calibration of camera lens distortion: Differential methods and robust estimation. *IEEE Transactions on image processing*. Vol. 14, nº 8. pp. 1215-1230
- Brown D. C. (1966) Descentering distortion of lenses. *Photogrammetric Engineering and Remote Sensing*
- Criminisi A. (1999) *Accurate visual metrology from single and multiple uncalibrated images*. Dr. Ing. Tesis Department of engineering science. University of Oxford
- Devernay F., Faugeras O. (2001) Straight lines have to be straight. *Machine Vision Applications* 13 (1). p. 14-24
- Faig W., (1975) Calibration of close-range photogrammetry systems: Mathematical formulation. *Photogrammetric engineering and remote sensing*, Vol 41, p 1479-1486
- Hartley R. (1997) In defence of the eight point algorithm. *IEEE Transactions on pattern analysis and machine intelligence*. Vol 19 – 6
- Hartley R., Zisserman A. (2004) *Multiple View Geometry in computer Vision*. Cambridge University Press.
- Kanade T., Rander P., Narayanan P. (1997). Virtualized reality: constructing virtual worlds from real scenes. *IEEE Conference on Computer Vision and Pattern Recognition. IEEE Multimedia, Immersive Telepresence* pp.34-47
- Lavest J., Viala M. Dhome M. (1998) Do we really need accurate calibration pattern to achieve a reliable camera calibration. *European Conference on Computer Vision*, vol. I, pp 158-174
- Mikhail E., Bethel J., McGlone J. (2001) *Introduction to Modern Photogrammetry*. John Wiley & Sons Inc.
- Raskar R., Welch G., Cutts M., Lake A., Stesin L., Fuchs H. (1998). The office of the future: a unified approach to image-based modelling and spatial immersive displays. *SIGGRAPH 98*, pp 179-188
- Ricolfé-Viala C., Sánchez-Salmerón A. (2006). *Caracterización y optimización del proceso de calibrado de cámaras basado en plantilla*



- bidimensional* Ph.D Thesis, Universidad Politécnica de Valencia.
- Salvi J., Battle J., Mouraddib E. (1998). A robust-coded pattern projection for dynamic 3D scene measurement. *International Journal of Pattern Recognition Lett.* 19, pp 1055-1065
- Salvi J., Armangué X., Battle J. (2002) A Comparative review of camera calibrating methods with accuracy evaluation. *Pattern recognition* Vol. 35, pp 1617-1635
- Samtaney R., (1999) A method to solve an exterior camera calibration parameters for image resection. *Technical report NAS-99-003 MRJ Technology Solutions Inc*
- Stoer C., Bulirsch R. (1980). *Introduction to Numerical Analysis*. Springer, Berlin.
- Sturm P., Maybank S. (1999). On plane-based camera calibration: a general algorithm, singularities and applications. *IEEE Conference on Computer Vision and Pattern Recognition* pp 432-437.
- Sun W., Cooperstock J. (2006). An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques. *Machine Vision and Applications*. 17(1): 51-67
- Shih S., Hung Y., Lin W. (1995) When should we consider lens distortion in camera calibration. *Pattern recognition* Vol. 28, No.3 pp 447-461
- Taha H.A. (1987) *Operation research. An introduction*. Prentice Hall International Editions.
- Tarel J.P. (1994) Calibration de caméra fondée sur les ellipses. *Research report 2200*, INRIA
- Trucco E., Verri A. (1998) *Introductory techniques for 3-D computer vision*. Prentice Hall
- Tsai R., (1987) A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-self TV camera lenses. *IEEE Journal of Robotics and Automation*, Vol RA-3 No 4. pp. 323-344
- Weng J., Huang T.S., Ahuja N. (1989) Motion and structure from two perspective views: algorithms, error analysis and error estimation. *IEEE transactions on pattern analysis and machine intelligence*. Vol 11 p. 451-475
- Wilson R. (1994) *Modelling and calibration of automated zoom lenses*. PhD. dissertation. Department electronic computing engineering, Carnegie Mellon University.
- Zhang Z. (1998) A flexible new technique for camera calibration. *Technical Report MSR-TR-98-71*, Microsoft Research.
- Zhang Z. (2000) A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol.22, Iss.11, Nov 2000 pp 1330- 1334.