



Revista Facultad de Ingeniería

ISSN: 0717-1072

facing@uta.cl

Universidad de Tarapacá

Chile

Aracena Pizarro, Diego; Campos, Pedro; Tozzi, Clésio Luis  
Comparación de técnicas de calibración de cámaras digitales  
Revista Facultad de Ingeniería, vol. 13, núm. 1, 2005, pp. 57-67  
Universidad de Tarapacá  
Arica, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=11413106>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica  
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## COMPARACIÓN DE TÉCNICAS DE CALIBRACIÓN DE CÁMARAS DIGITALES

Diego Aracena Pizarro<sup>1</sup> Pedro Campos<sup>2</sup> Clésio Luis Tozzi<sup>3</sup>

Recibido el 28 de marzo de 2005, aceptado el 25 de abril de 2005

### RESUMEN

Este trabajo presenta una comparación de las técnicas de calibración propuestas por Tuceryan et al. [15], Trucco et al. [13] y Zhang [20], los cuales proponen maneras diferentes para determinar los parámetros a partir de puntos 3D conocidos y sus correspondientes en el plano imagen 2D. La correspondencia es extraída en forma manual, centrando la atención en la exactitud de los parámetros de cámara entregados (intrínsecos y extrínsecos). Los patrones pasivos utilizados consideran las técnicas de calibración con diferentes homografías, dependiendo de los planos existentes en el proceso. Zhang [20] propone una aproximación basada en patrones acomodados en una superficie planar 2D, donde la cámara o el patrón se desplaza sin conocer el movimiento. En cambio Tuceryan et al. [15] y Trucco et al. [13] se basan en un modelo conocido 3D y su correspondencia en el plano imagen 2D, diferenciándose en la consideración de los parámetros intrínsecos que deben ser determinados o de influencia en la matriz de cámara final.

Los algoritmos fueron implementados en Matlab 6.0, de acuerdo a la interpretación de los trabajos y técnicas de optimización extraídas desde la literatura citada.

Palabras clave: Visión computacional, calibración de cámara, homografía 3D-2D, matrices de proyección.

### ABSTRACT

*This work presents a comparison of the calibration techniques proposed by Tuceryan et al. [15], Trucco et al. [13] and Zhang [20], suggesting different ways to determine the parameters, starting from known 3D points and their corresponding points in the 2D image plane. The correspondence is manually extracted, being the obtained parameter accuracies (extrinsic and intrinsic) the most important. The passive patterns used consider the calibration techniques with different homographs, depending on the existing process planes. Zhang [20] proposes an approximation based on patterns located on a 2D planar surface, where the camera or the pattern are displaced without knowledge of the motion. Tuceryan et al. [15] and Trucco et al. [13] base their procedure in a known 3D model and its correspondence on a 2D image plane, the difference being in the consideration of the intrinsic parameters that must be determined or considered in the final camera matrix. The algorithms were implemented using Matlab 6.0, according to the interpretation of the works and optimization techniques extracted from the cited literature.*

*Keywords: Computer vision, camera calibration, 3D-2D homograph, projection matrix.*

### INTRODUCCIÓN

En aplicaciones con visión computacional un aspecto importante a considerar es el proceso denominado calibración de cámara, el cual consiste en determinar los parámetros internos de la cámara, tales como distancia focal, factores de distorsión y puntos centrales del plano imagen. En la literatura se presentan diversos métodos, los cuales difieren en la forma de capturar desde las imágenes los parámetros intrínsecos, como también si el proceso es automático 2D-2D o manual 3D-2D. Las

comparaciones que considera este trabajo son para aplicaciones en las cuales se deben determinar los parámetros, por calibración, al comienzo de los procesos (Ej. Realidad aumentada, inspección, monitoreo, máquinas de visión), y estos permanecen fijos hasta que cambien las condiciones internas de la cámara, sólo se considera la determinación de manera dinámica los parámetros externos, los cuales varían porque se producen movimientos, ya sea por movimiento de la cámara o de los objetos en la escena, estos normalmente se determinan mediante técnicas de estimación de movimiento, a través

<sup>1</sup> Universidad de Tarapacá, Arica Chile, Académico Departamento de Computación e Informática, Facultad de Ingeniería, daracena@uta.cl

<sup>2</sup> Universidad de Tarapacá, Arica Chile, Estudiante Magíster Ingeniería de Software, Facultad de Ingeniería, pcampos@uta.cl

<sup>3</sup> Universidad Estatal de Campinas, São Paulo Brasil, Profesor Titular del Departamento de Ingeniería en Computación y Automatización, Facultad de Ingeniería Eléctrica y Computación, clesio@dca.fee.unicamp.br

dos vistas [19], [9], [16] o filtros de predicción Extendido de Kalman [4], [3], [2] o de Dimensión Variables [10].

El proceso de calibración de cámaras, es necesario para poder extraer información métrica a partir de imágenes 2D del mundo 3D. Aunque existen técnicas para inferir esta información desde imágenes captadas por cámaras no calibradas, a través de procesos iterativos de autocalibración o el antiguo proceso de calibración de los 3 puntos [7], el proceso de calibración de cámara abre la posibilidad de realizar aplicaciones efectivas en visión, tal como realidad aumentada, reconocimiento, seguimiento y reconstrucción 3D, los cuales se basan en el conocimiento de calibración y pose de la cámara [13].

La idea principal tras los procesos de calibración de cámara es describir el modelo de proyección que relaciona los sistemas de coordenadas que permiten obtener los parámetros de la cámara. En esencia el proceso de calibración de cámara consiste en determinar la geometría y características internas de la cámara, parámetros intrínsecos (tamaño horizontal ( $1/s_u$ ) y vertical ( $1/s_v$ ) de los píxeles o aspecto proporcional, coordenadas de proyección  $O'(r_0, c_0)$  del centro óptico, largo focal), y los parámetros extrínsecos (rotación y translación), que representan la localización y orientación (pose) de la cámara relativa a una imagen en un sistema de coordenadas. Estos parámetros normalmente son calculados desde un patrón de calibración (figura 1) que contiene rasgos fácilmente detectables de manera precisa en la imagen capturada [15].

Para ubicar objetos en el mundo real, establecemos un marco de referencia, que en el contexto de visión computacional es llamado *marco de referencia del mundo*. Un objeto en una imagen es medido en términos de coordenadas de píxeles, los cuales están en el *marco de referencia de la imagen*. El sólo conocer la distancia en píxeles entre puntos en una imagen, no nos permite determinar la distancia correspondiente a los mismos puntos en el mundo real. Por lo tanto, es necesario establecer las ecuaciones que unan el marco de referencia del mundo con el marco de referencia de la imagen, de manera de establecer la relación entre los puntos en coordenadas en el espacio 3D y los puntos en coordenadas de imagen 2D. Desafortunadamente, no se puede establecer esta relación directamente, haciéndose necesario establecer un marco de referencia intermedia, llamado *marco de referencia de la cámara*. Por lo tanto, se deben encontrar las ecuaciones que unan el marco de referencia de la cámara con el marco de referencia de la imagen, y las ecuaciones que unan el marco de referencia del mundo con el marco de referencia de la cámara. Al

resolver el sistema generado se obtiene la relación buscada, lo cual es equivalente a encontrar las características de la cámara [12].

En este trabajo se presenta una comparación de las técnicas o métodos de calibración de cámara directo propuestos por Tuceryan et al [15] y Trucco et al. [13] y el método de calibración de cámara por homografía o matriz de proyección propuesto por Zhang [20]. De esta manera contribuye a encontrar las fortalezas y exactitud en el proceso de calibración y pose, porque nos permite medir la precisión en realizar la reproyección y con esto se puede inicialmente garantizar un primer nivel de estimación del comportamiento de la aplicación a seguir (en nuestro caso aplicaciones de Realidad Aumentada). Para tal efecto se realiza una comparación de los métodos fotogramétricos recientes, descartándose el algoritmo de los 3 puntos por el conocimiento que se tiene de su escasa precisión y su poca exactitud reportada en aplicaciones de AR e industriales [18], considerando las técnicas recientes basada en métodos de calibración directa y homográfico propuestos.

## CALIBRACIÓN DE CÁMARA

Existen diversas técnicas de calibración de cámara. De acuerdo a la literatura estas pueden ser clasificadas en dos grandes categorías: Calibración fotogramétrica y autocalibración [20]:

- Calibración fotogramétrica. Se realiza mediante la observación de patrones cuya geometría en el espacio 3D es conocida con un buen nivel de precisión. Los patrones de calibración normalmente están posicionados en dos o tres planos ortogonales entre ellos. En algunos casos, basta con un único plano, cuya traslación es perfectamente conocida [14] y [20]. Este tipo de calibración requiere una configuración elaborada, pero sus resultados son eficientes.
- Autocalibración. Este método se basa en el movimiento de la cámara observando una escena estática, a partir de sus desplazamientos y usando únicamente la información de la imagen. La rigidez de la escena impone en general restricciones sobre los parámetros de cámara. Tres imágenes tomadas por una misma cámara con parámetros intrínsecos fijos son suficientes para obtener tanto los parámetros extrínsecos como intrínsecos. Aunque esta técnica es muy flexible, aun no se encuentra madura [5].

Dos métodos comunes de calibración, en la categoría de calibración fotogramétrica, son “Calibración directa de parámetros” y “Recuperación desde la homografía o matriz de proyección”. A través de la “Calibración directa de parámetros” se obtienen los parámetros extrínsecos e intrínsecos directamente, y a través de la “Recuperación desde la matriz de proyección” se obtienen los parámetros como una solución de “manera cerrada”, es decir, no se obtienen directamente los valores de los parámetros, sino que se calculan a partir de los valores de la homografía o matriz de proyección. Los métodos de calibración directa que estamos interesados requieren de patrones de calibración consistentes, en al menos dos planos ortogonales. Sin embargo, la calibración mediante homografía o matriz de proyección, una variante recientemente propuesta, consiste en utilizar patrones que descansan en un único plano. Este trabajo presenta una comparación entre estos tipos de algoritmos, a fin de conocer cual resulta más eficiente en el cálculo de los parámetros.

#### Métodos de calibración de cámara directo

- Coordenadas del centro imagen conocidas. Trucco et al. [13] presenta un método de calibración directo, el cual considera determinar los parámetros intrínsecos tales como distancia focal, aspecto proporcional y los parámetros extrínsecos. Para lograr esto, se basa en la ecuación de perspectiva y recuperación de pose de los puntos de interés en patrones, como muestra la figura 1.

La pose es representada por la matriz de rotación, escritas como  $r_{i,j}$ , y el vector de translación o desplazamiento  $T = [T_x, T_y, T_z]$ . Por otro lado, el largo focal efectivo  $f_x$  (largo focal dividido por el tamaño de píxel horizontal) y el aspecto proporcional  $\alpha$  del lente deben ser conocidos, en orden a calcular en forma precisa la pose y en consecuencia los parámetros intrínsecos y extrínsecos de la cámara. En forma resumida, los puntos correspondientes encontrados, están relacionados por

$$x = -f_x \frac{r_{11}X^w + r_{12}Y^w + r_{13}Z^w + T_x}{r_{31}X^w + r_{32}Y^w + r_{33}Z^w + T_z} \quad (1)$$

$$y = -f_y \frac{r_{21}X^w + r_{22}Y^w + r_{23}Z^w + T_y}{r_{31}X^w + r_{32}Y^w + r_{33}Z^w + T_z} \quad (2)$$

En las ecuaciones 1 y 2,  $x = c - c_0$  e  $y = r - r_0$  son conocidas, se tiene el mismo denominador y se introduce la tasa de proporcionalidad definida por

$\alpha = f_x / f_y$ , para cada par de correspondencia; simplificando las ecuaciones se llega a un sistema lineal con 8 incógnitas  $V = [v_1, v_2, \dots, v_8]$  de la forma:

$$\begin{aligned} x_i f_y (r_{21}X^w + r_{22}Y^w + r_{23}Z^w + T_y) = \\ y_i f_x (r_{11}X^w + r_{12}Y^w + r_{13}Z^w + T_x) \end{aligned} \quad (3)$$

$$\begin{aligned} x_i X_i^w v_1 + x_i Y_i^w v_2 + x_i Z_i^w v_3 + x_i v_4 + \\ -y_i X_i^w v_5 - y_i Y_i^w v_6 - y_i Z_i^w v_7 - y_i v_8 = 0 \end{aligned} \quad (4)$$

donde,

$$\begin{aligned} v_1 &= r_{21} & v_5 &= \alpha r_{11} \\ v_2 &= r_{22} & v_6 &= \alpha r_{12} \\ v_3 &= r_{23} & v_7 &= \alpha r_{13} \\ v_4 &= T_y & v_8 &= \alpha T_x \end{aligned} \quad (5)$$

La ecuación (4) permite armar un sistema de ecuaciones lineales de la forma  $AV = 0$ , para N correspondencias entre los puntos imagen  $(x_i, y_i)$  y el modelo de puntos 3D  $(X^w, Y^w, Z^w)$ . Para  $N \geq 7$  y los N puntos son no coplanares, el sistema tiene una solución no trivial (única hasta un factor de escala desconocido), que puede ser determinado desde el SVD de A, con

$$A = UDV^T \quad (6)$$

La matriz de rotación (R), los primeros dos elementos  $(T_x, T_y)$  del vector de translación y tasa proporcional  $\alpha$  son calculados desde V, bajo restricciones de ortogonalidad de la matriz de rotación, con  $\tilde{R} \cdot \tilde{R}^T = I$ , la suma de los elementos al cuadrado de las filas es la unidad y donde tercera fila de R es  $r_3 = r_1 \times r_2$ .

Finalmente, obtener  $T_z$  y  $f_x$ , a partir de la ecuación (1) la cual puede ser re-escrita como

$$\begin{aligned} x_i (r_{31}X_i^w + r_{32}Y_i^w + r_{33}Z_i^w + T_z) = \\ -f_x (r_{11}X_i^w + r_{12}Y_i^w + r_{13}Z_i^w + T_x) \end{aligned} \quad (7)$$

resolviendo el sistema sobre dimensionado de N ecuaciones lineales

$$A \begin{pmatrix} T_z \\ f_x \end{pmatrix} = b \quad (8)$$

Esta aproximación difiere del problema tradicional de calibración, en que el largo focal efectivo y tasa proporcional son calculadas para minimizar el error, en este caso se optimiza el cálculo de la rotación  $R$ , a través del método indicado en Weng et al. [17]. Una vez calculada la calibración, estos parámetros permanecen fijos, en el caso de estar interesado en seguimiento (tracking), mejora la robustez del proceso y reduce el número de parámetros a estimar en cada secuencia. La matriz de calibración de cámara resultante es

$$K = \begin{pmatrix} -f_x & 0 & c_0 \\ 0 & -f_y & r_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (9)$$

### Coordenadas del centro de imagen desconocidas

Tuceryan et al. [15], considera el sistema de referencia que muestra la figura patrón de calibración (fig. 1), de la cual, se pueden definir las siguientes relaciones de perspectiva:

$$u = f \frac{x}{z} \quad (10)$$

$$v = f \frac{y}{z} \quad (11)$$

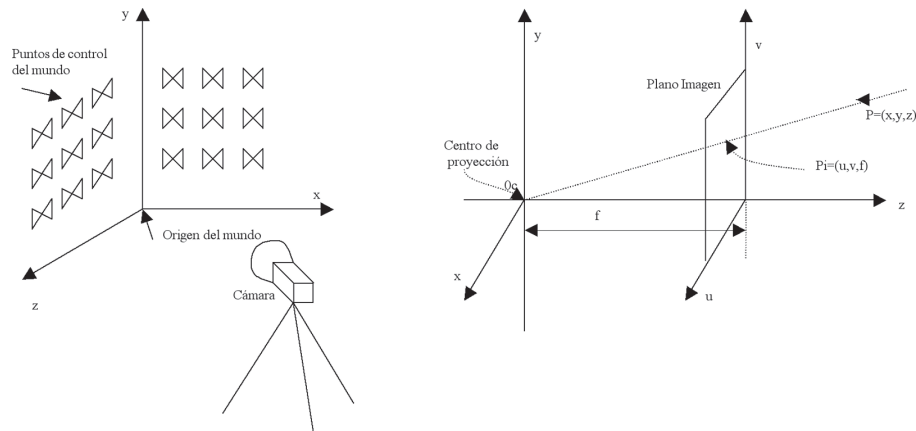


Fig. 1 Grilla patrón de Calibración y sistema de referencia (figuras tomadas de Tuceryan et al. [15]).

Las coordenadas de los píxeles son relativas a las coordenadas del plano imagen por las siguientes ecuaciones ( con  $r$  filas y  $c$  columnas)

$$-(r - r_0) = s_v v \quad (12)$$

$$(c - c_0) = s_u u \quad (13)$$

Donde  $((r_0, c_0))$  son el centro del sistema de coordenadas del plano imagen  $O'$ ,  $s_u$  y  $s_v$  reflejan la naturaleza no isotrópica de las cámaras (ejemplo, cámaras con píxeles no regulares) y el aspecto proporcional de la cámara. Sean  $f_u = s_u f$  y  $f_v = s_v f$ , reduciendo los parámetros intrínsecos de la cámara a  $f_u$ ,  $f_v$ ,  $r_0$ ,  $c_0$ , queda una matriz de cámara de la forma

$$A = \begin{pmatrix} f_u & 0 & c_0 \\ 0 & f_v & r_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (14)$$

y los parámetros extrínsecos o del movimiento ( $R$  y  $T$ ), logrando relacionar un punto 3D  $P=(x,y,z)$  y su proyección  $(r, c)$  de la forma

$$\begin{aligned} \frac{u}{f} &= \frac{r - r_0}{s_u f} = \frac{r_{11}x + r_{12}y + r_{13}z + t_1}{r_{31}x + r_{32}y + r_{33}z + t_3}, \\ \frac{v}{f} &= \frac{c - c_0}{s_v f} = \frac{r_{21}x + r_{22}y + r_{23}z + t_2}{r_{31}x + r_{32}y + r_{33}z + t_3} \end{aligned} \quad (15)$$

Acomodando las ecuaciones y dejando todo en función de datos conocidos se llega a

$$\begin{aligned} (r_i - r_0)(r_{31}x_i + r_{32}y_i + r_{33}z_i + t_3) + \\ -f_u(r_{11}x_i + r_{12}y_i + r_{13}z_i + t_1) = 0 \end{aligned} \quad (16)$$

$$\begin{aligned} (c_i - c_0)(r_{31}x_i + r_{32}y_i + r_{33}z_i + t_3) + \\ -f_v(r_{21}x_i + r_{22}y_i + r_{23}z_i + t_2) = 0 \end{aligned} \quad (17)$$

Arreglando y redefiniendo algunos términos (acomodados para resolver un sistema lineal de ecuaciones) para  $n$  puntos de calibración, obtenemos un sistema lineal de la forma  $AW = 0$ , donde  $W$  es un cambio estándar de variables definidas en la literatura de visión la cual linealiza las ecuaciones:

$$\begin{aligned}
 \bar{r}_0 &= W_1^T W_3, \\
 \bar{c}_0 &= W_2^T W_3, \\
 \bar{f}_u &= -\|W_1 - \bar{r}_0 W_3\|, \\
 \bar{f}_v &= \|W_2 - \bar{c}_0 W_3\|, \\
 W_1 &= f_u R_1 + r_0 R_3, \\
 W_2 &= f_v R_2 + c_0 R_3, \\
 W_3 &= R_3, \\
 w_4 &= f_u t_1 + r_0 t_3, \\
 w_5 &= f_v t_2 + c_0 t_3, \\
 w_6 &= t_3, \\
 \Rightarrow \bar{t}_1 &= (w_4 - \bar{r}_0 w_6) / \bar{f}_u, \\
 \bar{t}_2 &= (w_5 - \bar{c}_0 w_6) / \bar{f}_v, \\
 \bar{t}_3 &= w_6, \\
 \bar{R}_1 &= (W_1 - \bar{r}_0 W_3) / \bar{f}_u, \\
 \bar{R}_2 &= (W_2 - \bar{c}_0 W_3) / \bar{f}_v, \\
 \bar{R}_3 &= W_3
 \end{aligned} \quad (18)$$

Considerando la solución en (18), como estimación inicial de la matriz de rotación  $R$ , su optimización se realiza mediante técnicas de minimización propuestas en Weng et al. [17].

Uno de los factores que afecta la precisión del proceso de calibración de cámara, y consecuentemente los pasos siguientes, es la distorsión de los lentes, que es no lineal en la cámara física. La distorsión no lineal puede causar problemas debido a la inconsistencia con el modelo pinhole. Aunque los parámetros lineales estimados no son válidos, sirven como aproximación. Una solución para este problema es modelar las distorsiones de cámara no lineales y estimarlos. Trabajos recientes apuntan en ese sentido [20] y [8]. Esto es atractivo si el interés está centrado en tareas donde el modelo de cámara juega un papel importante en el análisis del contenido de la imagen. Por otro lado, si se desea renderizar en ambientes de computación gráfica, el hecho de extraer la distorsión pierde su importancia.

#### Método basado en Homografía

Zhang [20] presenta un algoritmo para estimación de pose basado en homografía o matriz de proyección. Este algoritmo requiere a lo menos 4 puntos coplanares correspondientes. Brevemente se describe el algoritmo.

El modelo de cámara pinhole describe las relaciones entre un punto 3D, coordenadas globales  $M = [X, Y, Z, 1]^T$ , y su proyección en el plano imagen  $m = [u, v, 1]^T$  como

$$sm = A[RT]M \quad (19)$$

donde  $s$  es el factor de escala desconocido,  $R$  la matriz de rotación,  $T$  es el vector de translación y  $A$  es la matriz de cámara o de parámetros intrínsecos, los cuales están en función de las coordenadas del centro de imagen  $(u_0, v_0)$ ,  $\alpha$  y  $\beta$  los cuales se derivan del largo focal y tamaño del píxel y  $\gamma$  la asimetría. Así la matriz de cámara es de la forma

$$A = \begin{pmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (20)$$

Cuando todos los puntos 3D son coplanares en el modelo, se selecciona el sistema de coordenadas globales tal que en el modelo el plano  $Z$  es cero. Por tanto la ecuación (19) puede ser re-escrita como

$$\begin{aligned}
 sm &= A[r_1 \ r_2 \ r_3 \ T][X \ Y \ 0 \ 1]^T \\
 sm &= A[r_1 \ r_2 \ T][X \ Y \ 1]^T
 \end{aligned} \quad (21)$$

Se considera  $M' = [X \ Y \ 1]^T$  como un punto en el plano imagen, luego  $m$  y  $M'$  pueden relacionarse por una matriz de homografía  $H$ :

$$sm = HM', \text{ con } H = A[r_1 \ r_2 \ T]. \quad (22)$$

Si se definen a lo menos 4 puntos no-colineales en el modelo, la homografía  $H$  puede ser determinada hasta un factor de escala. Se puede escribir

$$H = [h_1 \ h_2 \ h_3] = \lambda A[r_1 \ r_2 \ T], \quad (23)$$

Teóricamente se puede descomponer la homografía  $H$  para obtener la matriz intrínseca  $A$ , en este caso aplicamos el método descrito en Zhang [20] sección 3.1.

Luego se puede obtener la matriz de rotación  $R$  y el vector de translación  $T$ , de la siguiente manera:

$$\begin{aligned}
 r_1 &= \lambda A^{-1} h_1, \quad r_2 = \lambda A^{-1} h_2, \\
 r_3 &= r_1 \times r_2, \quad T = \lambda A^{-1} h_3,
 \end{aligned} \quad (24)$$

$$\text{donde} \quad \lambda = 1/|A^{-1} h_1| = 1/|A^{-1} h_2| \quad (25)$$

Debido a que el resultado obtenido en la ecuación (24) no fuerza la ortogonalidad en  $R$ , se necesita forzar esta ortogonalidad en  $R$ , en nuestro caso utilizamos el método descrito en Weng et al. [16]. Por minimización se refina

la estimación final de la matriz de calibración y los parámetros R y T.

De los procesos de calibración presentados se requiere conocer puntos 3D y su proyección correspondiente en puntos de imagen 2D, estableciéndose un sistema de ecuaciones que permiten determinar los parámetros intrínsecos y extrínsecos por el empleo del método de minimización por mínimos cuadrados. Para mas detalles de estos procesos ver [6], [14], [8], [20], [11], [13], [15] o [1].

## RESULTADOS DE LA COMPARACIÓN

Se utilizaron datos simulados y datos reales para la comparación de las técnicas de calibración propuestas en Tuceryan et al. [15], Trucco et al. [13] y Zhang [20]. Para los datos simulados, se realizó lo siguiente:

- Aprovechando los trabajos originales, se consideraron datos de los parámetros intrínsecos y extrínsecos, para una cámara virtual y se generaron puntos de imagen a partir de coordenadas de un mundo ficticio, utilizando estos parámetros.
- Posteriormente, se utilizó los algoritmos implementados con los pares de puntos correspondientes, a fin de verificar los resultados obtenidos.
- Finalmente, se verifica la exactitud de los parámetros de cámara obtenidos con cada algoritmo.

Para las pruebas con datos reales, se construyeron dos esquemas con patrones rectangulares negros y fondo claro, como muestran las figuras 2 y 3. Cabe hacer notar que los patrones deben ser figuras geométricas, donde sea fácil detectar rasgos, tales como puntos, líneas, esquinas, etc. En Tuceryan et al. [15] considera los centros de las mariposas, como punto de calibración, en este trabajo fueron considerados los puntos esquinas de los cuadrados existentes en los plano xy, xz e/o yz. Las imágenes fueron capturadas con una WebCam Cmos, Microtek conectable a la entrada USB. Estas pruebas permitieron verificar el nivel de error entre los puntos reprojectados y los extraídos manualmente desde las imágenes capturadas. El error se calcula por la distancia absoluta en píxeles.

### Prueba con datos simulados

Para el caso de los algoritmos no planares, se definieron los siguientes valores de parámetros extrínsecos a considerar en las pruebas:

$$R = \begin{bmatrix} 0.5985 & -0.0224 & -0.8008 \\ -0.2762 & 0.9326 & -0.2325 \\ 0.7520 & 0.3603 & 0.5520 \end{bmatrix} \quad T = \begin{bmatrix} 0.1 \\ -0.3 \\ -50 \end{bmatrix} \quad (26)$$

En primer lugar se muestra los puntos imagen que se logran dados los parámetros (tablas a.1, b.1, c.1, c.2 y c.3) y posteriormente se muestran la determinación de éstos, a partir de los puntos obtenidos aplicados a los algoritmos implementados (tablas a.2, b.2, c.4 y c.5).

- a) Para el método de calibración directa propuesto por Trucco et al. [13], considerando como parámetros intrínsecos  $\alpha = 0.9$  y  $f_x = -500$ , a partir de las coordenadas ficticias (X, Y, Z) se obtienen los puntos de imagen (r, c), los resultados se muestran en la tabla a.1.

Tabla a.1 Cálculo manual de los puntos.

Nº	X	Y	Z	C	R
1	18	10	0	157.9690	171.4554
2	0	9	6	376.4664	154.3444
3	6	0	6	333.2042	284.1562
4	6	11	0	278.5235	128.9363
5	0	15	9	413.9146	77.4235
6	3	0	6	352.7388	271.5538
7	12	3	0	229.5779	251.3715
8	0	12	9	410.5956	119.9224
9	12	0	6	287.1145	313.8902

Tabla a.2 Parámetros obtenidos desde la aplicación de los algoritmos.

Parámetro	Valor Definido	Valor Obtenido	Error Relativo
$\alpha$	0.9	0.9000	0.00000
$f_x$	-500	-488.6944	0.02261
R[1,1]	0.5985	0.5985	0.00000
R[1,2]	-0.0224	-0.0224	0.00000
R[1,3]	-0.8008	-0.8008	0.00000
R[2,1]	-0.2762	-0.2762	0.00000
R[2,2]	0.9326	0.9326	0.00000
R[2,3]	-0.2325	-0.2325	0.00000
R[3,1]	0.7520	0.7520	0.00000
R[3,2]	0.3603	0.3603	0.00000
R[3,3]	0.5520	0.5520	0.00000
T[1]	0.1	0.0724	0.27600
T[2]	-0.3	-0.2172	0.27600
T[3]	-50	-49.1520	0.01696



En este método de las pruebas realizadas no se detectan grandes errores en los parámetros intrínsecos (tabla a.2), pero si en lo que se refiere a la translación (parámetro extrínseco), presenta un error relativo a considerar, la experiencia indica que un error menor a 20% es aceptable [1], en este caso es cerca del 30%.

- b) Para el método de calibración directa propuesto por Tuceryan et al. [15], considerando como parámetros intrínsecos  $f_u = -555.5556$ ,  $f_v = -500$ ,  $r_0 = 240$  y  $c_0 = 320$ , se obtienen los siguientes puntos r, c de la imagen (tabla b.1)

Tabla b.1 Cálculo manual de los puntos.

Nº	X	Y	Z	r	c
1	18	10	0	420.0344	381.6901
2	0	9	6	177.2595	397.0900
3	6	0	6	225.3287	280.2594
4	6	11	0	286.0850	419.9574
5	0	15	9	135.6505	466.3188
6	3	0	6	203.6236	291.6015
7	12	3	0	340.4690	309.7657
8	0	12	9	139.3383	428.0698
9	12	0	6	276.5395	253.4988

Tabla b.2 Parámetros obtenidos desde la aplicación de los algoritmos.

Parámetro	Valor Definido	Valor Obtenido	Error Relativo
$f_u$	-555.5556	-555.5451	0.00002
$f_v$	-500	500.0161	0.00003
$r_0$	240	240.0220	0.00009
$c_0$	320	320.0128	0.00004
R[1,1]	0.5985	-0.5985	(*)0.00000
R[1,2]	-0.0224	0.0224	0.00000
R[1,3]	-0.8008	0.8008	0.00000
R[2,1]	-0.2762	-0.2762	0.00000
R[2,2]	0.9326	0.9326	0.00000
R[2,3]	-0.2325	-0.2325	0.00000
R[3,1]	0.7520	-0.7250	0.00000
R[3,2]	0.3603	-0.3603	0.00000
R[3,3]	0.5520	-0.5520	0.00000
T[1]	0.1	-0.0980	0.02000
T[2]	-0.3	-0.3013	0.00433
T[3]	-50	49.9995	0.00030

Cabe hacer notar que existen diferencias en el signo de los valores de R. Esto se debe a la definición inicial de la propuesta, que es distinta a la comúnmente usada, en este

caso consideramos el valor absoluto de estos parámetros (tabla b.2).

- c) Para el caso del método de calibración por homografía, propuesto por Zhang [20], se definieron los siguientes valores para los parámetros intrínsecos de la cámara:

$$u_0 = 320, v_0 = 240, \alpha = 1000, \beta = 900 \text{ y } \gamma = 1.1$$

Puesto que en este método se deben capturar varias imágenes del mismo plano, con diferentes ángulos y posiciones de la cámara, los valores de R y T varían, por tanto existen diferentes valores de R y T, para cada imagen.

Imagen ficticia 1, la tabla c.1 muestra los puntos logrados con parámetros extrínsecos

$$R1 = \begin{bmatrix} 0.9910 & -0.0192 & 0.1323 \\ 0.0081 & 0.9964 & 0.0840 \\ -0.1334 & -0.0822 & 0.9876 \end{bmatrix} \quad T1 = \begin{bmatrix} -3.8191 \\ 3.1214 \\ 14.0933 \end{bmatrix} \quad (27)$$

Tabla c.1 Cálculo manual de los puntos.

Nro.	X	Y	U	v
1	0	-0.5	50.6844	407.0308
2	0.5	-0.5	84.6303	408.0818
3	0.5	0	83.2953	440.5408
4	0	0	49.2567	439.3330
5	6.2	-6.7	497.3100	11.7418
6	6.7	-6.7	534.2061	10.8997
7	6.7	-6.2	534.1880	41.9194
8	6.2	-6.2	497.1819	43.6091

Imagen ficticia 2, la tabla c.2 muestra los puntos logrados con parámetros extrínsecos

$$R2 = \begin{bmatrix} 0.9965 & 0.0036 & 0.0839 \\ 0.0136 & 0.9791 & -0.2031 \\ -0.0829 & 0.2036 & 0.9755 \end{bmatrix} \quad T2 = \begin{bmatrix} -3.6792 \\ 3.2103 \\ 14.4592 \end{bmatrix} \quad (28)$$

Tabla c.2 Cálculo manual de los puntos.

Nro.	X	Y	U	V
1	0	-0.5	63.8250	410.5514
2	0.5	-0.5	97.8876	411.4727
3	0.5	0	99.6181	440.8212
4	0	0	65.7903	439.8223
5	6.2	-6.7	516.4366	6.4103
6	6.7	-6.7	556.8205	6.1262
7	6.7	-6.2	555.0984	42.8628
8	6.2	-6.2	515.0442	43.0245



Imagen ficticia 3, la tabla c.3 muestra los puntos logrados con parámetros extrínsecos

$$R3 = \begin{bmatrix} 0.9120 & -0.0317 & 0.4090 \\ -0.0082 & 0.9954 & 0.0953 \\ -0.4101 & -0.0902 & 0.9076 \end{bmatrix} T3 = \begin{bmatrix} -2.9223 \\ 3.1931 \\ 15.5574 \end{bmatrix} \quad (29)$$

Tabla c.3 Cálculo manual de los puntos.

Nº	X	Y	u	v
1	0	-0.5	133.9090	395.4789
2	0.5	-0.5	161.0458	397.3098
3	0.5	0	159.5821	426.9486
4	0	0	132.3859	424.7217
5	6.2	-6.7	535.9178	6.9286
6	6.7	-6.7	573.2122	3.0907
7	6.7	-6.2	572.9218	35.7967
8	6.2	-6.2	535.5078	39.1532

En este caso, en la tabla c.4 y c.5, se muestra la recuperación de los parámetros intrínsecos y los valores de R y T, obtenidos desde la primera imagen solamente.

Tabla c.4 Parámetros obtenidos desde la aplicación de nuestro algoritmo.

Parámetro	Valor Definido	Valor Obtenido	Error Relativo
$\alpha$	1000	1026	0.02600
$\beta$	900	919	0.02111
$\gamma$	1.1	3.6	2.27273
$u_0$	320	322	0.00625
$v_0$	240	256	0.06667
R[1,1]	0.9910	0.9889	0.00212
R[1,2]	-0.0192	-0.0196	0.02083
R[1,3]	0.1323	0.1471	0.11187
R[2,1]	0.0081	0.0041	0.49383
R[2,2]	0.9964	0.9945	0.00191
R[2,3]	0.0840	0.1049	0.24881
R[3,1]	-0.1334	-0.1484	0.11244
R[3,2]	-0.0822	-0.1031	0.25426
R[3,3]	0.9876	0.9835	0.00415
T[1]	-3.8191	-3.8350	0.00416
T[2]	3.1214	2.8717	0.07999
T[3]	14.0933	14.3950	0.02141

En este método se observa un nivel de error relativo alto en el parámetro de asimetría de los lentes de la cámara, su cambio de valor incide en el nivel de error que se produce en los parámetros de R en los elementos R[2,1], R[2,3], R[3,2].

Considerando la implementación (código ejecutable) facilitada por Z. Zhang [20] en su página <http://research.microsoft.com/~zhang/Calib>, también produce el mismo efecto detectado en nuestra implementación, la tabla c.5, a continuación muestra los resultados obtenidos al introducir nuestros datos al algoritmo original.

Tabla c.5 Parámetros obtenidos desde la aplicación del algoritmo implementado por Zhang[20].

Parámetro	Valor Definido	Valor Obtenido	Error Relativo
$\alpha$	1000	1000.34	0.00034
$\beta$	900	896.25	0.00417
$\gamma$	1.1	0.78	0.29091
$u_0$	320	329.32	0.02913
$v_0$	240	247.775	0.03240
R[1,1]	0.9910	0.9896	0.00141
R[1,2]	-0.0192	-0.0184	0.04167
R[1,3]	0.1323	0.1430	0.08088
R[2,1]	0.0081	0.0035	0.56790
R[2,2]	0.9964	0.9946	0.00181
R[2,3]	0.0840	0.1034	0.23095
R[3,1]	-0.1334	-0.1441	0.08021
R[3,2]	-0.0822	-0.1019	0.23966
R[3,3]	0.9876	0.9843	0.00334
T[1]	-3.8191	-3.9339	0.03006
T[2]	3.1214	3.0017	0.03835
T[3]	14.0933	14.0393	0.00383

Nuevamente, se puede observar el error relativo de la asimetría y de los elementos de R[2,1], R[2,3], R[3,2], los cuales son significativos al igual que lo producido con nuestra versión.

### Pruebas con imágenes reales

Para los algoritmos de Trucco et al. [13], Tuceryan et al. [15] y Zhang [20], se consideraron los patrones de calibración que muestran las figuras 2 y 3. De cada imagen se extrajeron los puntos de corner de los patrones cuadrados, en el sistema de referencia global o del mundo 3D. Posteriormente se encontraron manualmente los puntos correspondientes en las imágenes capturadas. Con este conjunto de puntos se aplicaron los algoritmos refinados implementados y se lograron los parámetros intrínsecos y extrínsecos, para finalmente utilizarlos en la reproyección de los puntos en cada imagen.



Fig. 2 Patrón 3D Calibración directa y puntos reproyectados a) Trucco y b) Tuceryan.

Para poder medir la distancia promedio de separación, entre el píxel medido y el reproyectado en la imagen, como también la desviación estándar, estos puntos se generaron de la siguiente manera:

- (1) 21 puntos distribuidos equitativamente entre los planos XY, YZ y XZ;
- (2) 14 puntos distribuidos equitativamente entre los planos XY e YZ.

El resultado se muestra en la tabla d.1.

Tabla d.1 Disparidad promedio y desviación, encontrada en los métodos de calibración directa.

Algoritmo	Promedio distancias (píxeles)	$\sigma$ desviación estándar
Trucco(1)	2.9405	3.2586
Trucco(2)	2.5641	2.9607
Tuceryan(1)	0.7992	0.9359
Tuceryan(2)	0.7693	0.9326

Para la reproyección con el método de Zhang [20], se utilizó un esquema análogo, de la forma:

- (1) 60 puntos distribuidos equitativamente en las 3 imágenes.
- (2) 100 puntos en las 3 imágenes, según muestra la figura 3.
- (3) Similar a (1), pero utilizando la implementación del algoritmo original de Zhang.

El resultado se muestra en la tabla d.2.

Tabla d.2 Disparidad promedio y desviación, encontrada en los métodos de calibración por homografía.

Algoritmo	Promedio distancias (píxeles)	$\sigma$ desviación estándar
Zhang(1)	0.9324	1.0496
Zhang(2)	1.3278	2.5779
Zhang(3)	0.9308	1.0435

Cabe hacer notar que los mejores resultados en esta fase, los muestran los algoritmos implementados de Tuceryan y Zhang, mostrando una disparidad promedio y desviación estándar aproximadas. El algoritmo de Trucco muestra un corrimiento mayor de los puntos, en la mayoría de las pruebas realizadas.

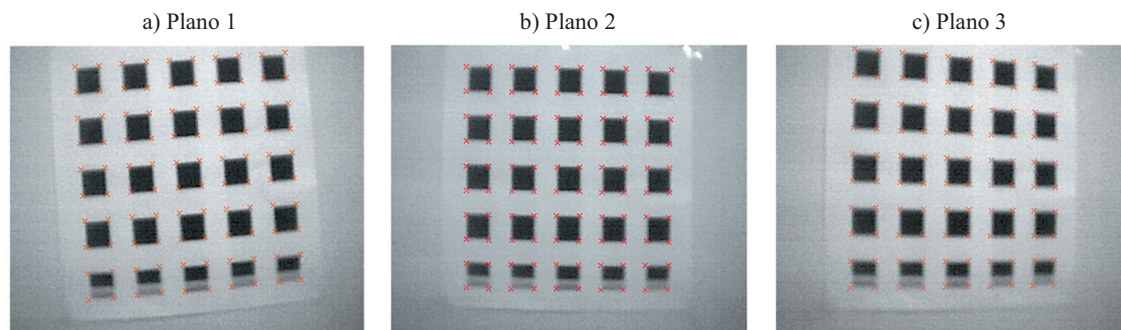


Fig. 3 Patrón 2D. Calibración por homografía y puntos reproyectados en los 3 planos imagen.

### Análisis de los valores obtenidos

Luego de comparar los valores obtenidos mediante los distintos algoritmos, se puede observar que los algoritmos que ofrecen el mejor desempeño en general, de acuerdo a los resultados, son: Tuceryan, a pesar que en la mayoría de las pruebas muestra diferencias, en cuanto a los signos en la matriz de rotación y vector de desplazamiento, lo interesante es que en conjunto entrega mayor precisión en la recuperación o reproyección de los puntos. Zhang es más regular en comportamiento, es más, tiene mayor precisión en la segunda y tercera imagen, en cuanto a reproyección.

Al observar los datos simulados, comparación de valores de parámetros de cámara obtenidos, se puede observar que los errores en los valores de los parámetros tanto extrínsecos como intrínsecos son bastante bajos en los algoritmos que utilizan patrones 3D, en comparación con el algoritmo de Zhang. Los errores en los primeros (calibración directa) son del orden de  $10^{-5}$ , en tanto que por homografía presenta errores del orden de  $10^{-1}$  en nuestra implementación, mientras que con la implementación original de Zhang los errores son del orden de  $10^{-2}$ . Se debe notar que los errores en los parámetros de T son mayores en los algoritmos de Trucco y Tuceryan, en tanto que en el algoritmo de Zhang los errores en este parámetro son bastante menor en magnitud.

Al verificar la reproyección de los puntos logrados con nuestros algoritmos implementados, en este caso, el error que se produce entre la distancia del punto medido y el punto generado con los parámetros de cámara obtenidos, se observa que los mejores valores son logrados por el algoritmo de Tuceryan, seguido por el algoritmo de Zhang. Las distancias son del orden de 0.9 pixel. A la luz de los resultados logrados en las diversas pruebas, se puede deducir que la disparidad que presenta Trucco, es producida por el error que introduce el vector de traslación.

### CONCLUSIONES

Los algoritmos de calibración de cámaras analizados son estables y tienen buen resultado. En particular, los algoritmos que utilizan patrones más complejos (3D), los cuales presentan un nivel más bajo de errores en el cálculo de los parámetros de cámara, no así en los parámetros extrínsecos, aunque su desempeño en conjunto es bastante satisfactorio. El algoritmo de Zhang, si bien tiene un nivel de error mayor, no muestra grandes diferencias al momento de reproyectar la ubicación de los puntos en la imagen, además de presentar la ventaja de utilizar un patrón sencillo, lo que lo hace muy atractivo, especialmente en el entorno de la computación personal,

computación móvil y aplicaciones de realidad aumentada y de visión computacional.

El desarrollo futuro del área seguramente estará dado por el refinamiento de algoritmos como el propuesto por Zhang, que utilizan patrones más sencillos para la calibración, lo que abre la puerta a una expansión del área de la visión computacional, puesto que pone al alcance de personas no iniciadas o con menos recursos, la utilización de complejos algoritmos de visión que requieren el paso previo de la calibración de cámara.

### TRABAJO FUTURO

Como trabajo futuro se considera realizar calibración de manera automática, sólo es necesario conocer un patrón colocado en un sistema de referencia conocido, captar vistas, insertar una fase de reconocimiento del patrón utilizado y detectar los puntos de control, para finalmente recuperar los parámetros de cámara y pose. Parámetros que son claves en aplicaciones de Realidad Aumentada, actualmente en desarrollo.

### REFERENCIAS

- [1] D. Aracena, C. Tosí. "Algoritmo de Seguimiento para aplicaciones en Realidad Aumentada", X Encuentro Chileno de la Computación, Copiapó Chile, 2002.
- [2] A. Azarbayejani and A. Pentland. "Recursive estimation of motion, structure, and focal length", IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(6):562-575, 1995.
- [3] Y. Bar-Shalom, X. Li. "Estimation and Tracking: principles, techniques and software", Artech House, Boston, 1993.
- [4] T. Broida, S. Chandrashekar and R. Chellappa. "Recursive 3D Motion Estimation from a Monocular Image Sequence", IEEE Transactions on Aerospace and Electronic Systems Vo. 26, 4, July. Pp. 639-656, 1990.
- [5] F. Dornaika and R. Chung ; "An Algebraic Approach to Camera Self Calibration", Computer Vision and Image Understanding, Vol. 83, pp. 195-215, 2001.
- [6] O. Faugeras. "Tridimensional Computer Vision: A Geometric Viewpoint" Artificial Intelligence. MIT Press, 1993.

- [7] R.M. Haralick, C. Lee, K. Ottenberg and N. Nolle. "Re-view and analysis of solutions of the three point perspective pose estimation problem" *International Journal of Computer Vision*, Vol. 13, 3, pp. 331-356, 1994.
- [8] R.I. Hartley. "In Defence of the 8-point Algorithm", in *Proceeding of the IEEE International Conference on Computer Vision*, 1995.
- [9] R.I. Hartley. "Kruppa's Equations Derived from Fundamental Matrix", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, (2), pp. 133-135, Feb, 1997.
- [10] P. McLauchlan and D. Murray. "A Unifying Framework for Structure and Motion recovery From Image Sequence", In *Proc. 5th Int'l Conf. On Computer Vision*, Boston, pp. 314-320, June, 1995.
- [11] L. Shapiro and G. Stockman. "Computer Vision" updated in March, 2000, <http://george.ee.washington.edu/courses.html>
- [12] Z. Sun y C. Rayburn. "Camera Calibration"; UNR & Ford Motor Company, 2001.
- [13] E. Trucco and A. Verri. "Introductory to techniques for 3-D computer vision"; Prentice Hall. 1998.
- [14] R. 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*, Vol. RA-3,4, pp. 323-344. August, 1987.
- [15] M. Tuceryan, D. Greer, R. Whitaker, D. Breen, C. Crapton and K. Ahlers. "Calibration Requirements and Procedures for a Monitor-Based Augmented Reality System", *IEEE Trans. On Visualization and Computer Graphics*, Vol. 1, 3, Sep. Pag 255-273, 1995.
- [16] J. Weng, T. Huang and N. Ahuja. "Motion and Structure from two Perspective Views: Algorithms, Error Analysis and Error Estimation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, (5), pp. 451-476, 1989.
- [17] J. Weng, P. Cohen and Herniou. "Camera Calibration with Distortion Models Accuracy Evaluation". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, (10), pp. 965-980, 1992.
- [18] X. Zhang and N. Navab. "Tracking and Pose Estimation for Computer Assisted Localization in Industrial Environments", *Fifth IEEE Workshop on Applications of Computer Vision*, pp. 214-221, December 04-06. 2000.
- [19] Z. Zhang, R. Deriche, O. Faugeras, Q. Luong. "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry", *INRIA Research Report 2273*, May. 1996.
- [20] A. Zhang. "A Flexible New Technique for Camera Calibration"; 1999  
<http://research.microsoft.com/~zhang>.