Proceso de calibración de cámaras digitales basado en el modelo Pin-Hole

A.I. Barranco Gutiérrez[2] y J.J. Medel Juárez[1,2]

[1]*Centro de Investigación en Computación del Instituto Politécnico Nacional*. A.A. 1055, Cúcuta, Colombia. [Times 9 pt]
[2]*Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada del Instituto Politécnico Nacional, Legaria 694. Colonia Irrigación, 11500 México D. F.*

Resumen
Mostramos el análisis de una nueva técnica flexible para facilitar la calibración de cámaras digitales. Solo se requiere una cámara para observar un patrón de muestra en diferentes orientaciones (al menos dos). Cada cámara o patrón plano pueden ser movidos libremente. El movimiento no necesita ser conocido. La Simulación por computadora y datos reales se han utilizado para examinar la técnica propuesta, y se han obtenido muy buenos resultados. Comparado con las clásicas técnicas que usan equipos costosos como los planos ortogonales calibrados, la técnica propuesta es fácil y flexible de usar.

Introducción
Actualmente en sistemas automáticos es necesario observar el sentido paramétrico de los objetos [3], que se encuentran en la imagen para que sistemas automáticos como robots o computadoras tomen decisiones en sustitución del ser humano [1]. Para poder medir las dimensiones reales de un objeto dentro de una imagen, es necesario conocer los parámetros de la cámara con la que se tomó la fotografía basados en un modelo y en base a esta información obtener dimensiones. Un modelo muy estudiado para resolver este problema es el llamado Pin-Hole [2].

Transformación lineal de coordenadas del espacio a coordenadas de una fotografía
Cuando vemos una fotografía, nosotros como seres humanos, podemos interpretar algunas distancias en 3 dimensiones de la foto. Por ejemplo en una fotografía de una persona podemos estimar el tamaño de su pie debido a información que poseemos con anticipación. También debido a que nuestro cerebro puede calcular algunas distancias debido a que conoce los parámetros de nuestros ojos y sabemos por que la foto fue tomada de un mundo en 3 dimensiones (3D), pero ahora se encuentra en 2 dimensiones (2D).La transformación lineal de modelo Pin-Hole esta expresada como sigue:

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad (1)$$

Calibración de una cámara
La calibración de una cámara consiste en conocer los parámetros $m_{i,j}$ de la transformación lineal. Si partimos de la ecuación (1) podemos observar que al dar un punto $[X,Y,Z,1]^T$ del espacio podemos mapearlo en el plano de la cámara con coordenadas (λu, λv, λ); si consideramos a λ como la relación de escala de la fotografía, entonces obtenemos el siguiente conjunto de ecuaciones:

$$u = \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \qquad (2)$$

$$v = \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \qquad (3)$$

Estas expresiones se utilizan en el caso de que se conoce un punto en el espacio y su correspondencia en la imagen. Por cada punto expresado en las ecuaciones (2) y (3) contamos con 2 ecuaciones; por lo tanto para obtener el valor de los 12 mij es necesario tener por lo menos 6 correspondencias de puntos y resolver el sistema de ecuaciones. Las ecuaciones (2) y (3) quedan expresadas en (4) y (5) para el i-ésimo conjunto de correspondencias de puntos.

$$m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14} - (m_{31}u_iX_i + m_{32}u_iY_i + m_{33}u_iZ_i) = u_im_{34} \qquad (4)$$

$$m_{21}X_i + m_{22}Y_i + m_{23}Z_i + m_{24} - (m_{31}v_iX_i + m_{32}v_iY_i + m_{33}v_iZ_i) = v_im_{34} \qquad (5)$$

podemos re-escribirlo también en forma matricial:

$$\begin{bmatrix} & & & & \vdots & \vdots & \vdots & & & & \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_iX_i & -u_iY_i & -u_iZ \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_iX_i & -v_iY_i & -v_iZ \\ & & & & \vdots & \vdots & \vdots & & & & \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \end{bmatrix} = \begin{bmatrix} \vdots \\ u_im_{34} \\ v_im_{34} \\ \vdots \end{bmatrix} \qquad (6)$$

[Solución por mínimos cuadrados]
Entonces obtenemos un sistema de ecuaciones de la forma $Ax = b$, que puede ser resuelto si primero multiplicamos ambos lados de la ecuación por AT:

$$A^T Ax = A^T b$$

Después despejamos el vector solución x:

$$(A^TA)^{-1}(A^TA)x = (A^TA)^{-1}A^Tb$$

$$x = (A^TA)^{-1}A^Tb$$

Agradecimientos
Agradecemos al CONACyT y a la Secretaria de Investigación y Posgrado (SIP) del IPN por su valioso apoyo a este trabajo.

Referencias
[1] Barranco, G.A.I (2008): Vision estereoscópica por computadora, México D.F.: CIC-IPN.
[2] Hartley R., Zisserman A. (2003): Multiple View Geometry in Computer Vision., Canberra, Australia. Cambridge University Press.
[3] Sossa H., (2006), Rasgos Descriptores para el reconocimiento de Objetos, México D.F., Insituto Politécnico Nacional.