

Calibração geométrica de câmeras, aplicando o método dos mínimos quadrados

Airton Marco Polidório^{1*}, Flavio Bortolozzi² e Maurício Figueiredo¹

¹Departamento de Informática, Universidade Estadual de Maringá, Av. Colombo, 5790, 87020-900, Maringá-Paraná, Brazil. e-mail: ampolido@din.uem.br ²Pontifícia Universidade Católica do Paraná, Av. Imaculada Conceição, 115, Prado Velho, 80215-901, Curitiba-Paraná, Brazil. *Author for correspondence.

RESUMO. Para tornar possível a visão artificial, ou visão computacional, é necessário resolver uma série de problemas de razoável complexidade. O primeiro problema a ser resolvido é a calibração das câmeras que serão utilizadas pelo sistema de visão artificial. No processo de calibração, devem-se utilizar padrões conhecidos e com coordenadas $\mathbb{R}^3$ conhecidas, uma vez que o problema básico consiste em associar um ponto em $\mathbb{R}^3$, espaço tridimensional, com um outro ponto em $\mathbb{R}^2$, plano de imagem ou plano de retina da câmera, o que envolve fenômenos óticos, eletrônicos, eletrônico-numéricos e mecânicos, os quais contribuem com erros para o sistema. Para minimizá-los, é necessário que o sistema de visão artificial seja redundante, que se tenham mais equações do que incógnitas para resolvê-lo. Para resolver esse sistema, opta-se, neste trabalho, pelo método dos mínimos quadrados. Apresenta-se, também, uma metodologia para aquisição dos padrões necessários à calibração.

Palavras-chave: calibração de câmeras, visão por computador, mínimos quadrados.

ABSTRACT. Geometric calibration of cameras applying least-squares method. To make artificial vision or computer vision possible, it is necessary to solve a series of problems of reasonable complexity. The first problem is the calibration of the cameras which will be used by the artificial vision system. The calibration process requires the use of well-known patterns with known $\mathbb{R}^3$ coordinates, once the basic problem consists of associating a point in $\mathbb{R}^3$, tridimensional space, with another point in $\mathbb{R}^2$, image plane or plane of camera retina, involving optical, electronic, electronic-numerical and mechanical phenomena that contribute with mistakes to the system. To minimize them it is necessary that the artificial vision system be redundant having more equations than variables. The least-squares method was applied to solve this system. A methodology appropriate to the acquisition of the necessary calibration patterns is also presented.

Key words: camera calibration, computer vision, least squares.

A visão artificial tridimensional permite conferir atributos qualitativos aos sistemas que justificam o seu uso, por exemplo: robôs que se locomovem de forma independente (não monitorada), robôs fixos em uma linha de produção, os quais precisam determinar pontos específicos nas peças fabricadas, independente da translação e da rotação das mesmas com referência a um sistema de coordenadas fixo, conduzir, de forma autônoma, veículos automotores, processar imagens de satélites quando se necessita estabelecer coordenadas de forma automática de objetos ou de lugares (Fu *et al.*, 1987; Gonzalez e Thomason, 1978).

Para dotar um sistema com visão artificial tridimensional, é necessário utilizar mais de uma

câmera de aquisição de imagens, posicionadas em perspectivas diferentes, mas com um campo de “visão” em comum (interseção entre os planos imagem). Normalmente, são utilizadas duas câmeras para cada sistema, constituindo um sistema artificial de visão estéreo binocular. Entretanto, como existem muitos erros intrínsecos associados e, também, muitas incertezas no cruzamento de dados entre as duas câmeras (*matching*), muitos sistemas já aparecem dotados com visão estéreo trinocular, com três câmeras. Com isso, reduzem-se as incertezas pelo esquema de votação: um ponto, ou segmento de reta, para existir no sistema de visão artificial, tem que ser “visto” por, pelo menos, duas câmeras.

Ao considerar que o sistema de referência é fixo em relação às câmeras, por exemplo, o centro geométrico de um robô, têm-se que calibrar as câmeras do sistema de visão artificial em relação a esse sistema de referência. A esta operação dá-se o nome de calibração geométrica das câmeras e isto pode ser feito em condições controladas em um laboratório. (Mensbach, 1986; Tsai, 1987).

Calibrar geometricamente uma câmera consiste em estabelecer um conjunto de equações que promovam a transformação pontual de todo ponto $M(X, Y, Z)$ do espaço $\mathcal{R}^3$, que tem como referencial um sistema de coordenadas mundo (O_w, X_w, Y_w, Z_w), para o espaço $\mathcal{R}^2$, associado com um ponto imagem $M_i(x, y)$ no plano imagem ou plano retina da câmera que está definido por um referencial específico, (Figura 1).

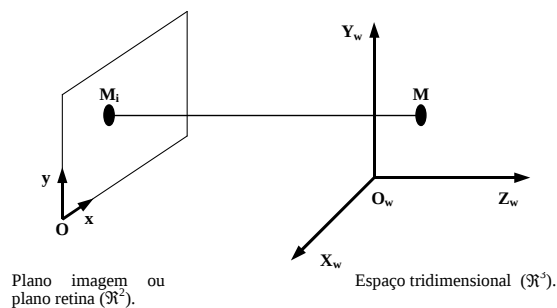


Figura 1. Ilustração dos aspectos geométricos gerais envolvidos na calibração de uma câmera

Este conjunto de equações deve representar uma realidade física que considera os seguintes fenômenos envolvidos: 1) ótico - pelo conjunto de lentes da câmera; 2) eletrônico - que transforma uma imagem tridimensional contínua em uma imagem analógica; 3) eletrônico-numérico - que consiste em transformar a imagem analógica em uma imagem digital (discreta) passível de quantificação pontual; 4) mecânicos - conversão do sistema de referência em $\mathcal{R}^3$ para $\mathcal{R}^2$.

Dessa forma, identificam-se dois tipos de parâmetros que são necessários para se calibrar uma câmera: os parâmetros extrínsecos (matrizes de rotação, translação) e os parâmetros intrínsecos (fatores de escala e projeção do centro ótico). Para melhor análise desses parâmetros, observe, a seguir, o modelo das transformações óticas de uma câmera.

Modelo de uma câmera

Na Figura 2, ilustra-se um modelo de transformação ótica que ocorre em uma câmera. Considere o ponto tridimensional P_k que é projetado no plano imagem da câmera, formando o ponto (*pixel*) m . Pode-se considerar que o plano imagem da câmera é um reticulado discreto de

pontos, o qual tem a capacidade de armazenar uma amostragem de pontos P_k , (Figura 3).

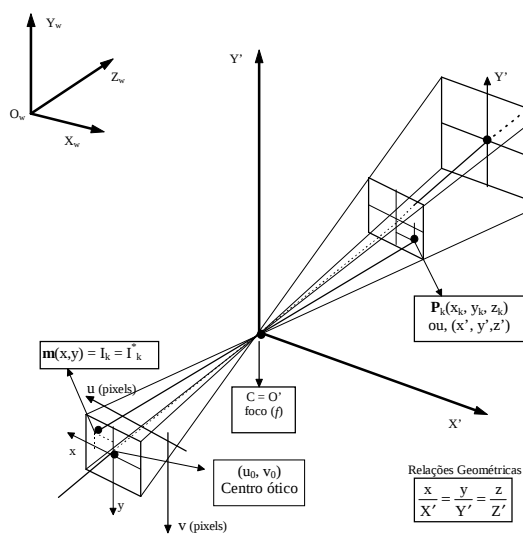


Figura 2. Modelo geométrico de uma câmera

Observe que o sistema reticulado do plano imagem da câmera é discreto e, portanto, pode existir um erro intrínseco associado ao sistema na conversão do ponto $P_k(x, y, z)$, espaço contínuo, para o ponto $m(x, y)$, espaço discreto. Por exemplo, supondo que o ponto $P_k(0, 2, 3)$ poderia ser convertido para $m(1, 2.5)$, indicando que o ponto no plano imagem da câmera ocuparia uma célula do reticulado correspondente à linha 1 e à coluna 2.5, o que é impossível, pois as células do reticulado são discretas e, portanto, só poderia ser (1, 2) ou (1, 3). Dessa forma, para que as relações geométricas mostradas na Figura 2 sejam verdadeiras, é necessário introduzir quatro coeficientes: os fatores de escala K_u e K_v e a projeção do centro ótico (u_0, v_0) sobre o plano imagem da câmera.

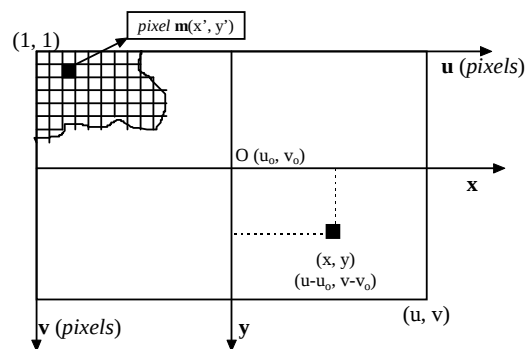


Figura 3. Reticulado do plano imagem da câmera

Os fatores de escala K_u e K_v são utilizados para se fazer a conversão de unidades entre os sistemas xy e uv que regem sobre o plano imagem da câmera. Um ponto no sistema xy é expresso em

unidades métricas, milímetros, por exemplo, e no sistema uv ele é expresso em *pixels*. O sistema de equações (1) relaciona os parâmetros envolvidos nesta conversão.

$$\begin{cases} x = \frac{u}{K_u} \\ y = \frac{v}{K_v} \end{cases} \quad (1)$$

Entretanto, como existem erros intrínsecos à construção da câmera, sabe-se que a projeção do centro ótico sobre o plano imagem não incide exatamente no centro do plano imagem. Necessita-se, então, transladar a origem do sistema xy na direção do ponto (u_o, v_o) que é a coordenada da projeção do centro ótico do plano imagem. Dessa forma, procedendo a translação (sistema de equações (2)), a coordenada (u, v) do *pixel* do ponto $\mathbf{m}$ fica relacionada com sua respectiva coordenada tridimensional (3D).

$$\begin{cases} x = \frac{u - u_o}{K_u} \\ y = \frac{v - v_o}{K_v} \end{cases} \quad \forall K_u, K_v > 0 \quad (2)$$

Na verdade, a determinação dos valores de k_u , k_v , u_o e v_o constitui o primeiro problema a ser resolvido para se calibrar uma câmera. Para algumas câmeras, como a CCD VÍDEO, pode-se obter o campo de tomada de vista e, a partir da distância focal (f), da profundidade, da largura e da altura do campo de vista, calcular os valores de k_u e k_v . Assim, ao dividir as duas relações do sistema de equações (2) pela distância focal f e fazendo:

$\alpha_u = K_u f$ e $\alpha_v = K_v f$ (2-a)
obtem-se o sistema de equações (3).

$$\begin{cases} u = u_o + \alpha_u \frac{x}{f} \\ v = v_o + \alpha_v \frac{y}{f} \end{cases} \quad (3)$$

Reescrevendo esse sistema como um sistema de coordenadas homogêneas (Noble, 1969) e considerando $Z \neq 0$ (reta suporte do foco da câmera - Figura 2) e fazendo $S = Z$ ser a constante de homogeneização e considerar, ainda, que $U = S.u$ e $V = S.v$, o sistema de equações ilustrado em (3) é transformado no sistema homogêneo de equações (4).

$$\begin{cases} U = \alpha_u.X + u_o.Z \\ V = \alpha_v.Y + v_o.Z \\ S = Z \end{cases} \quad (4)$$

O sistema de equações (4) pode ser escrito na forma matricial, sistema de equações (5), tendo, portanto, a forma geral: $\mathbf{U} = \mathbf{M.X}$, onde $\mathbf{M}$ é uma matriz 3×4 e é denominada: matriz de transformação perspectiva. Observe que, se $S = 0$, equivale dizer que $Z = 0$ e, portanto, o ponto no espaço tridimensional se encontra no plano focal da câmera e, dessa forma, os *pixels* de coordenada (u,v) contidos nesse plano não estão definidos no plano imagem da câmera.

Ainda, no sistema de equações (5), tem-se:

- se $\mathbf{X}$ e $\mathbf{M}$ são conhecidos, calculam-se as coordenadas (2D) dos pontos imagens $\mathbf{m}_i(x,y)$ associados com os pontos $\mathbf{P}_k(x, y, z)$;
- se $\mathbf{U}$ e $\mathbf{M}$ são conhecidos, podem-se reconstruir os pontos 3D, $\mathbf{P}_k(x, y, z)$, associados com os pontos imagens $\mathbf{m}_i(x,y)$;
- se $\mathbf{U}$ e $\mathbf{X}$ são conhecidos, calcula-se a matriz de transformação perspectiva $\mathbf{M}$.

$$\underbrace{\begin{pmatrix} U \\ V \\ S \end{pmatrix}}_{\mathbf{U}} = \underbrace{\begin{pmatrix} S.u \\ S.v \\ S \end{pmatrix}}_{\mathbf{M}} = \underbrace{\begin{pmatrix} \alpha_u & 0 & u_o & 0 \\ 0 & \alpha_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{\mathbf{M}} \underbrace{\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}}_{\mathbf{X}} \quad (5)$$

Sistemas de coordenadas com deslocamento rígido

Em algumas situações, para efeito de simplificações, pode-se fazer coincidir o sistema de coordenadas da câmera O' com o sistema de coordenadas mundo (O_w). Entretanto, as grandes aplicações exigem que esses sistemas de coordenadas sejam tratados em separado, como ilustra a Figura 4.

Os sistemas de coordenadas da câmera, O' e O'' , são fixos em relação à câmera e o referencial mundo (O_w) é fixo em relação ao mundo. Como os referenciais são independentes, os sistemas de coordenadas da câmera podem se mover, fixo ao movimento da câmera (O' é fixo em relação a O''), em relação ao referencial mundo. Nesse caso, o vetor $\mathbf{w}$ pode fornecer a posição da câmera e dos seus sistemas de coordenadas em qualquer instante em relação ao referencial mundo.

Ao se fazerem estas considerações, a câmera está livre para realizar movimentos de translação e de rotação em relação a O_w .

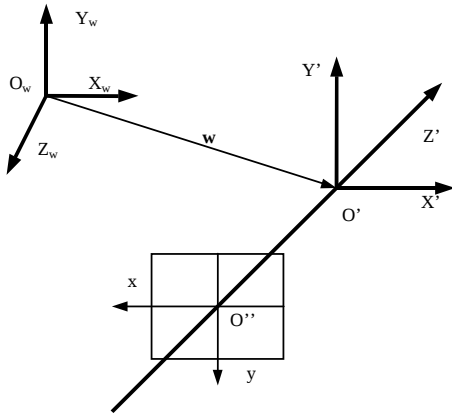


Figura 4. O Referencial mundo e os sistemas de coordenadas da câmera

Como todo referencial (O, X, Y, Z) pode ser calculado a partir do referencial mundo (O_w, X_w, Y_w, Z_w) pelo sistema de equações (6).

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \mathbf{R} \begin{pmatrix} X_w \\ Y_w \\ Z_w \end{pmatrix} + \mathbf{Tr} \quad (6)$$

onde:

$$\mathbf{R} = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \text{ é a matriz de rotação.}$$

$$\mathbf{Tr} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \text{ é a matriz de translação.}$$

Considerando a matriz de transformação perspectiva $\mathbf{M}$, sistema de equações (5), que pressupõe a coincidência dos referenciais mundo e os da câmera e o sistema de equações (6), que trata o referencial mundo e os referenciais da câmera em separado, deduz-se a Equação (7), que possibilita o cálculo da nova matriz de transformação perspectiva $\mathbf{T}$, exigida para que se possam tratar, em separado, os referenciais mundo e os da câmera.

$$\mathbf{T} = \mathbf{M} \begin{pmatrix} \mathbf{R} & \mathbf{Tr} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \quad (7)$$

onde:

$$\mathbf{0} = (0 \ 0 \ 0), \text{ e } \mathbf{1} = (1)$$

Desenvolvendo o algebrismo desse sistema, encontra-se o sistema (8).

$$\mathbf{T} = \begin{pmatrix} \alpha_u \mathbf{r}_1 + u_o \mathbf{r}_3 & \alpha_u t_x + u_o t_z \\ \alpha_v \mathbf{r}_2 + v_o \mathbf{r}_3 & \alpha_v t_y + v_o t_z \\ \mathbf{r}_3 & t_z \end{pmatrix} = \begin{pmatrix} t_1 & t_{14} \\ t_2 & t_{24} \\ t_3 & t_{34} \end{pmatrix} \quad (8)$$

Sabendo que os vetores linhas, $\mathbf{r}_i$, da matriz de rotação, $\mathbf{R}$, possuem as seguintes relações:

$$\mathbf{r}_i \cdot \mathbf{r}_j = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases} \quad (i, j = 1, 2, 3, \dots)$$

Através de algebrismo, pode-se, facilmente, constatar que os parâmetros intrínsecos (u_o, v_o), K_u (através de α_u) e K_v (através de α_v) são determinados pelas Equações (9-a), (9-b), (9-c) e (9-d), respectivamente.

$$\begin{aligned} u_o &= \mathbf{t}_1 \cdot \mathbf{t}_3^T & v_o &= \mathbf{t}_2 \cdot \mathbf{t}_3^T \\ (a) & & (b) \end{aligned} \quad (9)$$

$$\alpha_u = \sqrt{\mathbf{t}_1 \cdot \mathbf{t}_1^T - u_o^2} \quad \alpha_v = \sqrt{\mathbf{t}_2 \cdot \mathbf{t}_2^T - v_o^2} \quad (c) \quad (d)$$

De forma análoga, determinam-se os parâmetros extrínsecos (matriz de rotação $\mathbf{R} = (\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3)^T$ e matriz de translação $\mathbf{Tr} = (t_x \ t_y \ t_z)^T$), através das Equações 10-a, 10-b, 10-c, 10-d, 10-e e 10-f, respectivamente.

$$\mathbf{r}_1 = \frac{1}{\alpha_u} (\mathbf{t}_1 - u_o \mathbf{t}_3) \quad t_x = \frac{1}{\alpha_u} (t_{14} - u_o t_{34}) \quad (a) \quad (d)$$

$$\mathbf{r}_2 = \frac{1}{\alpha_v} (\mathbf{t}_2 - v_o \mathbf{t}_3) \quad t_y = \frac{1}{\alpha_v} (t_{24} - v_o t_{34}) \quad (b) \quad (e) \quad (10)$$

$$\mathbf{r}_3 = \mathbf{t}_3 \quad t_z = t_{34} \quad (c) \quad (f)$$

Cálculo da matriz de transformação perspectiva

Calibrar uma câmera, em relação a um referencial, significa determinar a matriz de transformação perspectiva existente entre a câmera, mais precisamente o plano imagem da câmera, e o referencial 3D associado a esta câmera.

A Figura 5 mostra o ponto $\mathbf{P}_k(X_k, Y_k, Z_k)$, pertencente a $\mathcal{R}^3$, sendo projetado no plano imagem de uma câmera, projeção esta que passa pelo foco C , centro ótico, e se define sobre o plano imagem no ponto $\mathbf{I}_k(u_k, v_k)$.

Considerando as coordenadas homogêneas da projeção de $\mathbf{P}_k(X_k, Y_k, Z_k)$ como sendo o ponto $\mathbf{I}_k^*(U_k, V_k, S_k)$, então, calibrar uma câmera

consiste em determinar a matriz de transformação perspectiva que permita transformar o ponto $\mathbf{P}_k$ no ponto $\mathbf{I}_k^*$ correspondente, o que é dado pela Equação 11.

$$\mathbf{I}_k^* = \mathbf{T} \cdot \begin{pmatrix} \mathbf{P}_k \\ 1 \end{pmatrix} \quad (11)$$

onde:

$$\mathbf{P}_k = (X_k \ Y_k \ Z_k)^T$$

$$\mathbf{I}_k^* = (U_k \ V_k \ S_k)^T$$

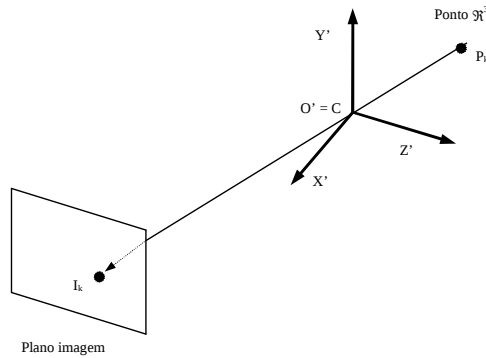


Figura 5. Projeção do ponto P_k (R_3) sobre o plano imagem (ponto I_k (R_2))

Supondo o caso geral $S_k \neq 0$, as coordenadas no plano imagem $\mathbf{I}_k$ (em *pixels*) são dadas pela Equação (12).

$$\mathbf{I}_k = \begin{pmatrix} u_k \\ v_k \end{pmatrix} = \begin{pmatrix} \frac{U_k}{S_k} \\ \frac{V_k}{S_k} \end{pmatrix} \quad (12)$$

Do sistema de equações (12) conclui-se que:

$$\begin{cases} U_k - u_k \cdot S_k = 0 \\ V_k - v_k \cdot S_k = 0 \end{cases} \quad (13)$$

Desenvolvendo-se a Equação (11), encontra-se:

$$\begin{cases} U_k = \mathbf{t}_1 \cdot \mathbf{P}_k + t_{14} \\ V_k = \mathbf{t}_2 \cdot \mathbf{P}_k + t_{24} \\ S_k = \mathbf{t}_3 \cdot \mathbf{P}_k + t_{34} \end{cases} \quad (14)$$

Substituindo (14) em (13), tem-se:

$$\begin{cases} \mathbf{t}_1 \cdot \mathbf{P}_k + t_{14} - u_k \cdot (\mathbf{t}_3 \cdot \mathbf{P}_k + t_{34}) = 0 \\ \mathbf{t}_2 \cdot \mathbf{P}_k + t_{24} - v_k \cdot (\mathbf{t}_3 \cdot \mathbf{P}_k + t_{34}) = 0 \end{cases} \quad (15)$$

No sistema de equações (15), no processo de calibração de câmeras, os valores $\mathbf{P}_k$, u_k e v_k são conhecidos. Logo, este sistema possui 11 incógnitas: $\mathbf{t}_1 = (t_{11} \ t_{12} \ t_{13})$, $\mathbf{t}_2 = (t_{21} \ t_{22} \ t_{23})$, $\mathbf{t}_3 = (t_{31} \ t_{32} \ t_{33})$, t_{14} e t_{24} . Então, para se resolver este sistema é necessário que se tenha $k=1, 2, 3, \dots, n$, onde n deve ser no mínimo 5, uma vez que, para cada valor de k , obtêm-se duas equações. Dessa forma, para se calcular a matriz de transformação perspectiva $\mathbf{T}$, são necessários, no mínimo, 6 pontos, não coplanares, $\mathbf{P}_k$ e, outros 6 pontos $\mathbf{I}_k$.

Na prática, para minimizar os erros intrínsecos ao sistema de visão artificial e de truncamento no processamento computacional, utiliza-se um número maior de pontos (da ordem de uma centena).

A aplicação do método dos mínimos quadrados

O problema em se aplicar o sistema de equações (15) é que nem sempre o parâmetro t_{34} é conhecido, ao supor que existe translação no sistema, tem-se que ter $t_z \neq 0$. Assim, substituindo o sistema de equações (8) no sistema (15) e dividindo o resultado por t_z , obtêm-se o sistema de equações (16).

$$\begin{cases} \mathbf{X}_1 \cdot \mathbf{P}_k - u_k \cdot \mathbf{X}_2 \cdot \mathbf{P}_k + C_1 = u_k \\ \mathbf{X}_3 \cdot \mathbf{P}_k - v_k \cdot \mathbf{X}_2 \cdot \mathbf{P}_k + C_2 = v_k \end{cases} \quad k = (1, \dots, n), n > 5 \quad (16)$$

O sistema de equações (16) pode ser representado pela forma matricial $\mathbf{A} \cdot \mathbf{X} = \mathbf{B}$, onde:

$$\mathbf{A} = \begin{pmatrix} \mathbf{P}_k^T & -u_k \cdot \mathbf{P}_k^T & \mathbf{0} & 1 & 0 \\ \mathbf{0} & -v_k \cdot \mathbf{P}_k^T & \mathbf{P}_k^T & 0 & 1 \end{pmatrix}_{(2n \times 11)} \quad k = (1, \dots, n), n > 5 \quad (16-a)$$

$$\mathbf{X} = (\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3 \ C_1 \ C_2)^T_{(11 \times 1)} \quad (16-b)$$

$$\mathbf{B} = \begin{pmatrix} u_k \\ v_k \end{pmatrix}_{(2n \times 1)} \quad k = (1, \dots, n), n > 5 \quad (16-c)$$

$$\mathbf{X}_1 = \frac{u_0}{t_z} \cdot \mathbf{r}_3 + \frac{\alpha_u}{t_z} \cdot \mathbf{r}_1 \quad (16-d)$$

$$\mathbf{X}_2 = \frac{1}{t_z} \cdot \mathbf{r}_3 \quad (16-e)$$

$$\mathbf{X}_3 = \frac{v_0}{t_z} \cdot \mathbf{r}_3 + \frac{\alpha_v}{t_z} \cdot \mathbf{r}_2 \quad (16-f)$$

$$C_1 = u_0 + \frac{\alpha_u}{t_z} \cdot t_x \quad (16-g)$$

$$C_2 = v_0 + \frac{\alpha_v}{t_z} \cdot t_y \quad (16-h)$$

A solução de $\mathbf{A}\mathbf{X} = \mathbf{B}$, pelo método dos mínimos quadrados é dada pela Equação (17).

$$\mathbf{X} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A} \cdot \mathbf{B} \quad (17)$$

e, partindo da Equação (16-e), sabendo que $\|\mathbf{r}_3\|^2 = 1$, pode-se mostrar que:

$$|t_z| = \frac{1}{\|\mathbf{X}_2\|} \quad (18)$$

Agora, basta determinar os parâmetros intrínsecos e extrínsecos da câmera para se determinar a matriz de transformação perspectiva, $\mathbf{T}$, Equação (8).

Os parâmetros intrínsecos a serem determinados são: (u_o, v_o) - coordenadas em *pixels* do centro ótico do plano imagem e (K_u, K_v) - fatores de escala (através de α_u e α_v pelas relações da Equação (2-a)). A nomenclatura das operações matemáticas usadas são as seguintes:

$\mathbf{X} \bullet \mathbf{Y}$ - produto interno entre os vetores $\mathbf{X}$ e $\mathbf{Y}$.

$\mathbf{X} \wedge \mathbf{Y}$ - produto vetorial ou externo entre os vetores $\mathbf{X}$ e $\mathbf{Y}$.

$\|\mathbf{X}\|$ - norma do vetor $\mathbf{X}$.

As Equações (19-a), (19-b), (19-c) e (19-d) fornecem os parâmetros necessários para se calcularem os parâmetros intrínsecos da câmera.

$$u_o = \frac{\mathbf{X}_1 \bullet \mathbf{X}_2^T}{\|\mathbf{X}_2\|^2} \quad v_o = \frac{\mathbf{X}_2 \bullet \mathbf{X}_3^T}{\|\mathbf{X}_2\|^2} \quad (19)$$

$$\alpha_u = \frac{\|\mathbf{X}_1 \wedge \mathbf{X}_2^T\|}{\|\mathbf{X}_2\|^2} \quad \alpha_v = \frac{\|\mathbf{X}_2 \wedge \mathbf{X}_3^T\|}{\|\mathbf{X}_2\|^2} \quad (19)$$

Os parâmetros extrínsecos: matriz de rotação, $\mathbf{R}$, e matriz de translação, $\mathbf{Tr}$, podem ser calculados pelas equações (20-a), (20-b), (20-c), (21-a), (21-b) e (21-c); lembrando que a matriz de rotação, de ordem 3×3 , é dada pelos seguintes vetores de ordem 1×3 : $\mathbf{r}_1$, $\mathbf{r}_2$ e $\mathbf{r}_3$, ou seja, $\mathbf{R} = (\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3)^T$ e a matriz de translação, ordem 3×1 , é dada pelos seguintes componentes: t_x, t_y e t_z ou, $\mathbf{Tr} = (t_x \ t_y \ t_z)^T$.

$$\mathbf{r}_1 = \frac{\|\mathbf{X}_2\|}{\|\mathbf{X}_1 \wedge \mathbf{X}_2^T\|} \left(\mathbf{X}_1 - \frac{\mathbf{X}_1 \bullet \mathbf{X}_2^T}{\|\mathbf{X}_2\|^2} \mathbf{X}_2 \right) \quad (20-a)$$

$$\mathbf{r}_2 = \frac{\|\mathbf{X}_2\|}{\|\mathbf{X}_2 \wedge \mathbf{X}_3^T\|} \left(\mathbf{X}_3 - \frac{\mathbf{X}_2 \bullet \mathbf{X}_3^T}{\|\mathbf{X}_2\|^2} \mathbf{X}_2 \right) \quad (20-b)$$

$$\mathbf{r}_3 = \frac{\mathbf{X}_2}{\|\mathbf{X}_2\|} \quad (20-c)$$

$$t_x = \frac{\|\mathbf{X}_2\|}{\|\mathbf{X}_1 \wedge \mathbf{X}_2^T\|} \left(C_1 - \frac{\mathbf{X}_1 \bullet \mathbf{X}_2^T}{\|\mathbf{X}_2\|^2} \right) \quad (21-a)$$

$$t_y = \frac{\|\mathbf{X}_2\|}{\|\mathbf{X}_2 \wedge \mathbf{X}_3^T\|} \left(C_2 - \frac{\mathbf{X}_2 \bullet \mathbf{X}_3^T}{\|\mathbf{X}_2\|^2} \right) \quad (21-b)$$

$$t_z = \frac{1}{\|\mathbf{X}_2\|} \quad (21-c)$$

A seguir, apresenta-se a proposta de um algoritmo para se obter a matriz de calibração do sistema ou a matriz de transformação perspectiva ($\mathbf{T}$).

Dados de entrada:

N : número de pontos I_k e P_k conhecidos;

Todos os N pontos $P_k(X_k, Y_k, Z_k)$: conhecidos sobre a mira de calibração;

Todos os N Pontos $I_k(u_k, v_k)$: conhecidos sobre o plano imagem em *pixels*;

Estruturas e valores iniciais:

para $k=1$ até N

definir a matriz $\mathbf{A}_{(2N \times 11)}$ (Equação (16-a))

definir a matriz $\mathbf{B}_{(11 \times 1)}$ (Equação (16-c))

Cálculos:

calcular $\mathbf{A}^T$

calcular $\mathbf{A}^T \cdot \mathbf{A}$

calcular $\mathbf{A}^T \cdot \mathbf{B}$

calcular $\mathbf{X} = (\mathbf{A}^T \cdot \mathbf{A})^{-1} \cdot \mathbf{A}^T \cdot \mathbf{B}$. Matriz solução pelo método dos mínimos quadrados. Equação (17).

decompor a matriz $\mathbf{X}$ em $(\mathbf{X}_1 \ \mathbf{X}_2 \ \mathbf{X}_3 \ C_1 \ C_2)^T$. Equação (16-b).

calcular u_o . Abscissa da coordenada do centro ótico. Equação (19-a).

calcular v_o . Ordenada da coordenada do centro ótico. Equação (19-b).

calcular α_u . Relação do fator de escala K_u . Equação (19-c).

calcular α_v . Relação do fator de escala K_v . Equação (19-d).

calcular $\mathbf{r}_1$. Primeiro vetor da matriz de rotação. Equação (20-a).

calcular $\mathbf{r}_2$. Segundo vetor da matriz de rotação. Equação (20-b).

calcular $\mathbf{r}_3$. Terceiro vetor da matriz de rotação. Equação (20-c).

calcular t_x . Primeiro componente da matriz de translação. Equação (21-a).

calcular t_y . Segundo componente da matriz de translação. Equação (21-b).

calcular t_z . Terceiro componente da matriz de translação. Equação (21-c).

calcular $\mathbf{T}$. Matriz de transformação perspectiva. Equação (8).

Metodologia para determinação dos pontos $\mathbf{P}_k$ e $\mathbf{I}_k$

Para calibrar uma câmera, pelo método aqui proposto, é necessário conhecer as coordenadas (X_k, Y_k, Z_k) dos pontos $\mathbf{P}_k$, $\mathcal{R}^3$ e as respectivas coordenadas (u_k, v_k) dos pontos imagem que são formados no plano imagem da câmera $\mathbf{I}_k$, $\mathcal{R}^2$.

Para isso, foram marcados, em preto, 60 quadrados, com 1cm de lado cada, em duas folhas de papel milimetrado, 30 em cada, de tal forma que era possível ter conhecimento da posição do centro de cada quadrado nas folhas de papel milimetrado. Essas folhas foram coladas, cada uma, em uma placa de papelão rígido e, em seguida, essas placas foram coladas uma na outra pelas suas bordas, de tal forma que o ângulo formado pelos planos das duas placas foi de 90° . Em seguida, foi posicionada uma câmera de vídeo para aquisição de uma imagem dos pontos contidos nas placas, como ilustra a Figura 6. Neste mesmo processo, pode-se estabelecer a origem do sistema de coordenada mundo, $\mathbf{O}_w$, como ilustra a própria Figura 6.

De posse de uma imagem digital do padrão de pontos $\mathbf{P}_k$, para se obterem as respectivas coordenadas (u_k, v_k) , em *pixels*, basta fazer um processamento em cada padrão de ponto na imagem, a fim de calcular os respectivos centróides, Figura 7, os quais correspondem diretamente às coordenadas procuradas.

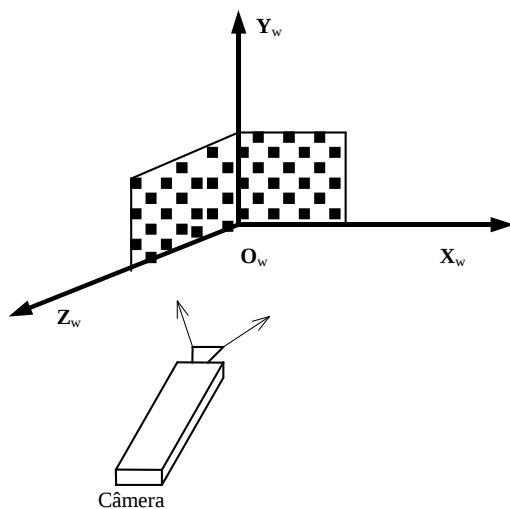


Figura 6. Ilustração da metodologia usada para aquisição dos pontos $\mathbf{P}_k$ e $\mathbf{I}_k$

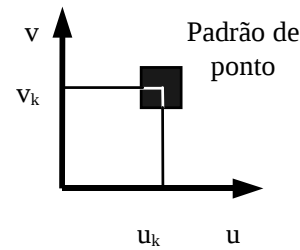


Figura 7. Ilustração de um padrão de ponto $\mathbf{I}_k$ e sua respectiva coordenada (u_k, v_k)

Dados e resultados obtidos

Os 60 pontos $\mathbf{P}_k$ pertencentes ao padrão, bem como os seus correspondentes pontos $\mathbf{I}_k$, são apresentados na Tabela 1.

Com esses dados, foram obtidos os seguintes valores para os parâmetros intrínsecos e extrínsecos da câmera:

$$u_o = 64.0958 \quad v_o = 59.1520 \quad \alpha_u = 1656.4103 \quad \alpha_v = 2433.9870$$

$$\mathbf{R} = \begin{pmatrix} 0.61670 & -0.02837 & -0.79759 \\ 0.04627 & -1.00005 & 0.05725 \\ -0.69894 & -0.14670 & -0.69997 \end{pmatrix}$$

$$\mathbf{Tr} = \begin{pmatrix} 0.5244 \\ 0.6025 \\ 3.8184 \end{pmatrix}$$

$$\mathbf{T} = \begin{pmatrix} 976.7094 & -56.3921 & -1365.9979 & 1113.2813 \\ 71.2828 & -2442.7879 & 97.9317 & 1692.3869 \\ -0.6989 & -0.1467 & -0.7000 & 3.8184 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \end{pmatrix}$$

Para se fazer uma análise qualitativa sobre esses resultados, foi adotado o seguinte procedimento:

1. Foi calculada a distância, d_{kw} , de cada ponto $\mathbf{P}_k$ com a origem de $\mathbf{O}_w = (0, 0, 0)$.
2. Os pontos $\mathbf{I}_k$ foram reconstruídos para $\mathcal{R}^3$, $\mathbf{P}_{ik}$, usando a matriz de transformação perspectiva obtida neste experimento e, também, foi calculada a distância, d_{ikw} , desses pontos com a origem de $\mathbf{O}_w$.
3. Procedeu-se, então, ao cálculo do desvio relativo entre estas distâncias: $dr = d_{kw} - d_{ikw}$.

Estes resultados podem ser observados na Tabela

1.

A calibração de uma câmera é apenas o estágio inicial de um sistema de visão artificial. Contudo, é nesse estágio que fica determinada a precisão com que o sistema irá operar.

Tabela 1. Dados e resultados produzidos no experimento

P_k			I_k		P_{ik}			d_{kw}	d_{ikw}	dr
X_k	Y_k	Z_k	u_k	V_k	X_{ik}	Y_{ik}	Z_{ik}			
0.0000	0.6000	0.4450	138	79	-0.0007	0.6003	0.4440	0.7470	0.7466	0.0004
0.0000	0.5000	0.4450	139	150	0.0022	0.5001	0.4471	0.6693	0.6708	-0.0015
0.0000	0.4000	0.4450	140	220	-0.0037	0.3996	0.4417	0.5984	0.5957	0.0027
0.0000	0.3000	0.4450	141	290	-0.0056	0.2985	0.4411	0.5367	0.5327	0.0040
0.0000	0.2000	0.4450	142	359	-0.0062	0.1984	0.4398	0.4879	0.4825	0.0054
0.0000	0.1000	0.4450	143	427	0.0027	0.1006	0.4474	0.4561	0.4586	-0.0025
0.0000	0.6000	0.3450	174	75	-0.0025	0.5996	0.3436	0.6921	0.6911	0.0010
0.0000	0.5000	0.3450	175	144	0.0011	0.5003	0.3460	0.6075	0.6083	-0.0008
0.0000	0.4000	0.3450	176	213	-0.0001	0.3994	0.3446	0.5282	0.5276	0.0007
0.0000	0.3000	0.3450	177	281	-0.0020	0.2995	0.3435	0.4572	0.4557	0.0015
0.0000	0.2000	0.3450	178	349	0.0041	0.2012	0.3488	0.3988	0.4027	-0.0040
0.0000	0.1000	0.3450	179	416	0.0004	0.1002	0.3445	0.3592	0.3588	0.0004
0.0000	0.6000	0.2450	209	70	0.0008	0.6003	0.2462	0.6481	0.6488	-0.0007
0.0000	0.5000	0.2450	210	139	-0.0012	0.4993	0.2445	0.5568	0.5560	0.0008
0.0000	0.4000	0.2450	211	206	0.0049	0.4010	0.2498	0.4691	0.4725	-0.0034
0.0000	0.3000	0.2450	211	273	-0.0013	0.2992	0.2435	0.3873	0.3858	0.0015
0.0000	0.2000	0.2450	212	339	-0.0009	0.2001	0.2439	0.3163	0.3155	0.0008
0.0000	0.1000	0.2450	213	405	0.0022	0.1009	0.2465	0.2646	0.2664	-0.0018
0.0000	0.6000	0.1450	243	66	-0.0015	0.5999	0.1432	0.6173	0.6167	0.0005
0.0000	0.5000	0.1450	244	133	0.0075	0.5013	0.1508	0.5206	0.5235	-0.0029
0.0000	0.4000	0.1450	244	199	0.0055	0.4014	0.1495	0.4255	0.4284	-0.0029
0.0000	0.3000	0.1450	245	265	0.0032	0.3009	0.1472	0.3332	0.3349	-0.0017
0.0000	0.2000	0.1450	245	330	0.0063	0.2019	0.1505	0.2470	0.2519	-0.0049
0.0000	0.1000	0.1450	246	395	0.0089	0.1028	0.1534	0.1761	0.1848	-0.0087
0.0000	0.6000	0.0450	275	63	-0.0112	0.5983	0.0346	0.6017	0.5994	0.0023
0.0000	0.5000	0.0450	276	128	0.0035	0.5006	0.0482	0.5020	0.5029	-0.0009
0.0000	0.4000	0.0450	276	193	0.0013	0.4003	0.0464	0.4025	0.4029	-0.0004
0.0000	0.3000	0.0450	276	258	-0.0097	0.2974	0.0365	0.3034	0.2998	0.0036
0.0000	0.2000	0.0450	277	322	0.0016	0.2000	0.0470	0.2050	0.2055	-0.0005
0.0000	0.1000	0.0450	277	385	0.0013	0.1001	0.0469	0.1097	0.1106	-0.0009
0.0450	0.6000	0.0000	304	62	0.0468	0.6002	0.0012	0.6017	0.6020	-0.0003
0.0450	0.5000	0.0000	304	128	0.0480	0.5003	0.0032	0.5020	0.5026	-0.0006
0.0450	0.4000	0.0000	304	193	0.0351	0.3983	-0.0091	0.4025	0.3999	0.0026
0.0450	0.3000	0.0000	305	257	0.0498	0.3012	0.0046	0.3034	0.3053	-0.0019
0.0450	0.2000	0.0000	305	321	0.0418	0.1996	-0.0031	0.2050	0.2039	0.0011
0.0450	0.1000	0.0000	305	385	0.0362	0.0973	-0.0086	0.1097	0.1042	0.0055
0.1450	0.6000	0.0000	336	65	0.1420	0.5999	-0.0030	0.6173	0.6165	0.0008
0.1450	0.5000	0.0000	337	132	0.1533	0.5015	0.0085	0.5206	0.5245	-0.0039
0.1450	0.4000	0.0000	337	198	0.1476	0.4005	0.0023	0.4255	0.4268	-0.0013
0.1450	0.3000	0.0000	337	264	0.1387	0.2984	-0.0060	0.3332	0.3292	0.0040
0.1450	0.2000	0.0000	337	329	0.1431	0.1996	-0.0023	0.2470	0.2456	0.0014
0.1450	0.1000	0.0000	337	394	0.1388	0.0982	-0.0060	0.1761	0.1701	0.0060
0.2450	0.6000	0.0000	370	69	0.2398	0.5993	-0.0042	0.6481	0.6455	0.0025
0.2450	0.5000	0.0000	371	137	0.2455	0.5002	-0.0005	0.5568	0.5572	-0.0004
0.2450	0.4000	0.0000	371	204	0.2508	0.4012	0.0049	0.4691	0.4732	-0.0041
0.2450	0.3000	0.0000	371	271	0.2452	0.3000	0.0001	0.3873	0.3875	-0.0001
0.2450	0.2000	0.0000	371	338	0.2453	0.2002	0.0003	0.3163	0.3166	-0.0003
0.2450	0.1000	0.0000	371	404	0.2478	0.1003	0.0029	0.2646	0.2674	-0.0027
0.3450	0.6000	0.0000	406	72	0.3514	0.6009	0.0071	0.6921	0.6962	-0.0040
0.3450	0.5000	0.0000	406	141	0.3459	0.5004	0.0007	0.6075	0.6083	-0.0009
0.3450	0.4000	0.0000	406	210	0.3534	0.4020	0.0091	0.5282	0.5353	-0.0071
0.3450	0.3000	0.0000	406	279	0.3480	0.3003	0.0028	0.4572	0.4596	-0.0024
0.3450	0.2000	0.0000	405	346	0.3421	0.1993	-0.0028	0.3988	0.3959	0.0028
0.3450	0.1000	0.0000	405	413	0.3424	0.0995	-0.0020	0.3592	0.3566	0.0026
0.4450	0.6000	0.0000	443	76	0.4457	0.5995	0.0005	0.7470	0.7470	0.0000
0.4450	0.5000	0.0000	442	146	0.4405	0.4994	-0.0043	0.6693	0.6659	0.0034
0.4450	0.4000	0.0000	442	217	0.4386	0.3986	-0.0074	0.5984	0.5927	0.0056
0.4450	0.3000	0.0000	442	286	0.4401	0.2989	-0.0048	0.5367	0.5320	0.0047
0.4450	0.2000	0.0000	442	355	0.4472	0.2009	0.0021	0.4879	0.4903	-0.0024
0.4450	0.1000	0.0000	442	424	0.4490	0.1009	0.0040	0.4561	0.4602	-0.0041

É fácil observar que, para calibrar uma câmera pelo método aqui proposto, são necessários 6 pontos P_k conhecidos e os seus respectivos pontos I_k , os quais geram 12 equações. Entretanto, o método dos mínimos quadrados permite que se criem sistemas redundantes, com mais equações do que incógnitas, e quanto maior for o número de equações no sistema, menores serão os erros extrínsecos e

intrínsecos do sistema. Neste trabalho, foram usados 60 pontos P_k , gerando, portanto, 120 equações.

A qualidade dos resultados obtidos pode ser observada na Tabela 1, onde se constata que o maior desvio percentual obtido é de 4.9%, mostrando a precisão do método e a qualidade da metodologia envolvida na parte experimental. Lembrando, ainda, que esse desvio é em relação à origem do sistema de coordenada mundo e não em relação à distância

entre os pontos $\mathbf{P}_k$ e $\mathbf{P}_{ik}$, se assim fosse, o desvio máximo seria de 1.3%. Esses desvios podem ser minimizados, simplesmente aumentando-se o número de padrões conhecidos (pontos $\mathbf{P}_k$).

Referências bibliográficas

- Mensbach, P. Calibration of a camera and light source by fitting to a physical model. *Comp. Vis., Graph. Im. Process.*, 35:200-219, 1986.
- Tsai, R.Y. A versatile camera calibration technique for high-accuracy 3D machine vision methodology using off-the-shelf TV cameras and lenses. *IEEE J. Rob. Autom.*, RA-3, 4:323-344, 1987.
- Fu, K.S.; Gonzalez, R.C.; Lee, C.S.G. *Robotics: control, sensing, vision and intelligence*. New York: McGraw-Hill, 1987.
- Gonzalez, R.C.; Safabakhsh, R. Computer vision techniques for industrial applications. *Computer*, 15(12):17-32, 1990.
- Noble, B. *Applied linear algebra*. New Jersey: Prentice-Hall, Englewood Cliffs, 1969.

Received on October 15, 1998.

Accepted on November 23, 1998.