



UNICAMP

Universidade Estadual de Campinas

Faculdade de Engenharia Mecânica

IM563 – Processamento de Imagens
Aplicado à Automação e Robótica

19 de junho de 2024

Docentes: Paulo R. G. Kurka

Discente:

– Gabriel Toffanetto França da Rocha – 289320

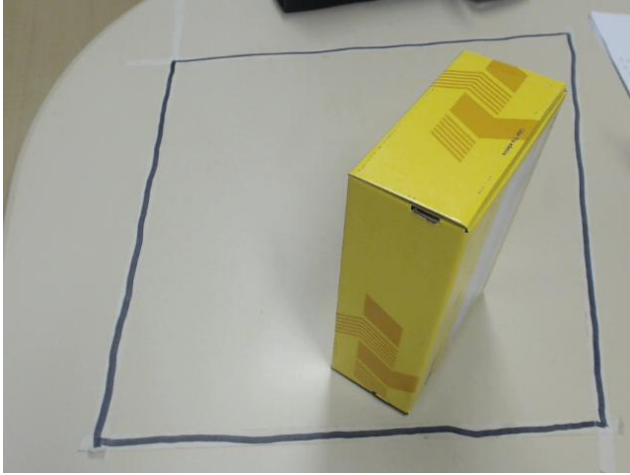
Trabalho 3

Sumário

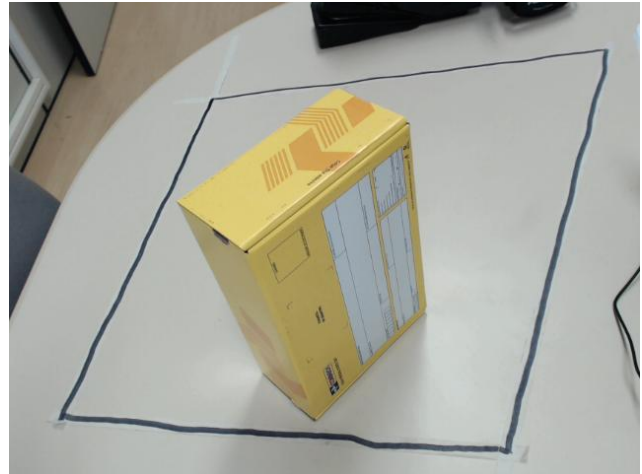
1	Calibração Intrínseca	2
1.1	Obtenção de pontos	2
1.2	Realização da calibração intrínseca	4
2	Transformação entre às câmeras	5
3	Conclusão	7
	Anexos	8
	Referências	8

1 Calibração Intrínseca

Para realização da calibração intrínseca da câmera, foram tomadas duas capturas de um objeto de calibração, sendo essa, uma caixa conhecida com 275 mm de altura, 87 mm de largura e 150 mm de profundidade. A Figura 1a apresenta a vista esquerda da caixa, enquanto na Figura 1b foi capturada a vista direita do objeto. Ambas as imagens foram capturadas pela mesma câmera, tendo tamanho de 640×480 pixels.



(a) Vista esquerda.



(b) Vista direita.

4

Figura 1: Imagens de calibração.

1.1 Obtenção de pontos

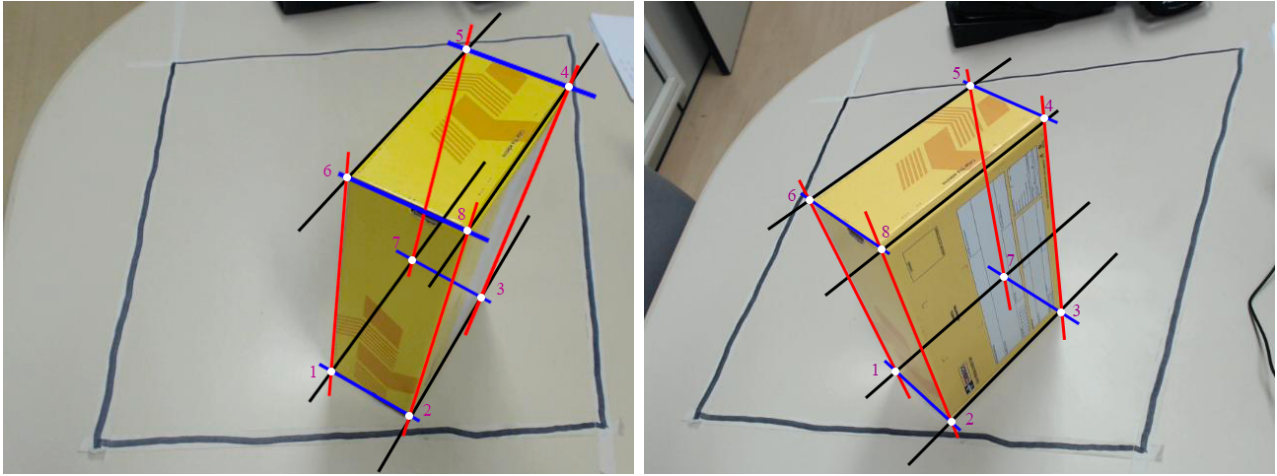
A definição dos pontos do sólido de calibração que serão utilizados para a calibração é crucial, tendo sido realizado de forma manual por meio do *software* de edição de imagens Adobe Photoshop. Os pontos são enumerados para corresponder ao vetor de pontos conhecidos utilizados na função de calibração, sendo o sistema global de coordenadas tendo origem no ponto 1 e os eixos definidos nas direções:

$$\text{Eixo } x \equiv \overrightarrow{1 \rightarrow 2} \quad (1)$$

$$\text{Eixo } y \equiv \overrightarrow{1 \rightarrow 7} \quad (2)$$

$$\text{Eixo } z \equiv \overrightarrow{1 \rightarrow 6} \quad (3)$$

O ponto oculto foi estimado por meio da projeção das arestas ocultas do sólido, por meio do paralelismo com as arestas visíveis, onde o ponto de interseção foi aproximado como o oitavo vértice do paralelepípedo.



(a) Vista esquerda.

(b) Vista direita.

Figura 2: Imagens de calibração com pontos conhecidos do mundo destacados.

Os pontos foram amostrados manualmente da imagem com os eixos, utilizando a função `getpts` do MATLAB, resultando em dois vetores, com as coordenadas x e y de cada ponto. Dado as coordenadas, os mesmos foram utilizados para montar a matriz de pontos de calibração, com o mesmo padrão mostrado em (4), sendo `ceil()` a função de truncamento para obtenção da coordenada de imagem do ponto, sendo obtida uma matriz de pontos para cada imagem. A notação com subscrito L será adotada para as equações relativas à imagem da vista esquerda, e o subscrito R para a vista direita.

$$\hat{\mathbf{p}} = \begin{bmatrix} \text{ceil}(\vec{x})_{[1 \times 8]} \\ \text{ceil}(\vec{y})_{[1 \times 8]} \\ \mathbf{1} \end{bmatrix}_{[3 \times 8]} \quad (4)$$

Dados os pontos em coordenadas de mundo do sólido de calibração, em metros, dado por (5), obteve-se as matrizes com os pontos relativos em coordenadas de imagem para a imagem da vista esquerda, (6), e para a vista direita, (7).

$$\mathbf{B} = \begin{bmatrix} 0 & 0.087 & 0.087 & 0.087 & 0 & 0 & 0 & 0.087 \\ 0 & 0 & 0.150 & 0.150 & 0.150 & 0 & 0.150 & 0 \\ 0 & 0 & 0 & 0.275 & 0.275 & 0.275 & 0 & 0.275 \end{bmatrix} \quad (5)$$

$$\hat{\mathbf{p}}_L = \begin{bmatrix} 334 & 412 & 486 & 575 & 470 & 350 & 416 & 472 \\ 377 & 423 & 302 & 89 & 50 & 180 & 265 & 235 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (6)$$

$$\hat{\mathbf{p}}_R = \begin{bmatrix} 256 & 312 & 424 & 407 & 331 & 169 & 366 & 242 \\ 378 & 428 & 317 & 121 & 86 & 203 & 282 & 253 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (7)$$

1.2 Realização da calibração intrínseca

Com a aplicação da função de calibração intrínseca $\text{RTAa}()$, obtém-se a matriz de parâmetros intrínsecos aproximada ($\hat{\mathbf{A}}$), a matriz de rotação aproximada ($\hat{\mathbf{R}}$) e o vetor de translação aproximado ($\hat{\mathbf{t}}$), sendo os dois últimos em relação à câmera e o sistema de coordenadas do mundo, ou seja, serão diferentes para cada uma das imagens.

Para a vista esquerda (Figura 2a), foram obtidas as seguintes matrizes:

$$\hat{\mathbf{A}}_L = \begin{bmatrix} 630.8066 & 0 & 426.9824 \\ 0 & 634.6096 & 216.9257 \\ 0 & 0 & 1 \end{bmatrix}; \quad (8)$$

$$\hat{\mathbf{R}}_L = \begin{bmatrix} 0.8682 & 0.4806 & 0.1239 \\ 0.4416 & -0.6343 & -0.6345 \\ -0.2263 & 0.6055 & -0.7630 \end{bmatrix}; \quad \hat{\mathbf{t}}_L = \begin{bmatrix} 0.1366 \\ -0.2138 \\ 0.5406 \end{bmatrix}; \quad (9)$$

Para a vista direita (Figura 2b), foram obtidas as seguintes matrizes:

$$\hat{\mathbf{A}}_R = \begin{bmatrix} 648.0318 & 0 & 330.0854 \\ 0 & 607.4200 & 200.7036 \\ 0 & 0 & 1 \end{bmatrix}; \quad (10)$$

$$\hat{\mathbf{R}}_R = \begin{bmatrix} 0.6871 & 0.7023 & -0.1861 \\ 0.4168 & -0.5909 & -0.6908 \\ -0.5951 & 0.3971 & -0.6987 \end{bmatrix}; \quad \hat{\mathbf{t}}_R = \begin{bmatrix} 0.3318 \\ -0.0902 \\ 0.5437 \end{bmatrix}; \quad (11)$$

Por fim, para uma melhor realização da matriz de parâmetros intrínsecos aproximada, tira-se a média das duas aproximações obtidas, tendo então:

$$\hat{\mathbf{A}} = \begin{bmatrix} 639.4192 & 0 & 378.5339 \\ 0 & 621.0148 & 208.8146 \\ 0 & 0 & 1 \end{bmatrix} \quad (12)$$

2 Transformação entre às câmeras

A matriz de rotação e o vetor de translação entre a câmera e o sistema de coordenadas do mundo podem ser representadas na forma de transformações homogêneas, que compactam a informação de rotação e translação, sendo construída com base em (13).

$$\mathbf{T} = \begin{bmatrix} \mathbf{R}_{[3 \times 3]} & \mathbf{t}_{[3 \times 1]} \\ 0 & 0 & 0 & 1 \end{bmatrix}_{[4 \times 4]} \quad (13)$$

Para a transformação entre dois *frames* $\{2\}$ e $\{1\}$ quaisquer, tem-se (14).

$$\mathbf{T}_{2/1} = \begin{bmatrix} \mathbf{R}_{2/1} & \mathbf{t}_{2/1} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (14)$$

De acordo com a propriedade algébrica das transformações homogêneas, a transformação do *frame* $\{0\}$ para o $\{2\}$ pode ser decomposta no produto da transformação do *frame* $\{0\}$ para o $\{1\}$, e do *frame* $\{1\}$ para o $\{2\}$.

$$\mathbf{T}_{0/2} = \mathbf{T}_{2/1} \mathbf{T}_{1/0} \therefore \mathbf{T}_{2/1} = \mathbf{T}_{0/2} \mathbf{T}_{1/0} = \mathbf{T}_{0/2} \mathbf{T}_{0/1}^{-1} \quad (15)$$

Dessa forma, uma vez que se conhece a transformação da câmera na vista esquerda para o sistema do mundo, e a transformação da câmera na vista direita para o sistema do mundo, a transformação entre as duas câmeras pode ser encontrada da forma (16).

$$\hat{\mathbf{T}}_{R/L} = \hat{\mathbf{T}}_{R/0} \hat{\mathbf{T}}_{0/L} = \hat{\mathbf{T}}_{R/0} \hat{\mathbf{T}}_{L/0}^{-1} \quad (16)$$

Com base em (9), pode-se escrever a transformação homogênea da câmera L para o sistema de coordenadas do mundo como (17).

$$\hat{\mathbf{T}}_{L/0} = \begin{bmatrix} \hat{\mathbf{R}}_L & \hat{\mathbf{t}}_L \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.8682 & 0.4806 & 0.1239 & 0.1366 \\ 0.4416 & -0.6343 & -0.6345 & -0.2138 \\ -0.2263 & 0.6055 & -0.7630 & 0.5406 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (17)$$

E por meio da (11), pode-se escrever a transformação homogênea da câmera R para o sistema de coordenadas do mundo como (18).

$$\hat{\mathbf{T}}_{R/0} = \begin{bmatrix} \hat{\mathbf{R}}_R & \hat{\mathbf{t}}_R \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0.6871 & 0.7023 & -0.1861 & 0.3318 \\ 0.4168 & -0.5909 & -0.6908 & -0.0902 \\ -0.5951 & 0.3971 & -0.6987 & 0.5437 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

Com isso, de acordo com (16), têm-se que, a transformação do referencial da câmera da esquerda para a câmera da direita é expressa por (19) e a transformação da câmera da direita para a câmera da esquerda é dada por (20).

$$\hat{\mathbf{T}}_{L/R} = \begin{bmatrix} 0.9153 & 0.2589 & -0.3085 & 0.2233 \\ -0.2945 & 0.9527 & -0.0743 & 0.0173 \\ 0.2747 & 0.1589 & 0.9483 & -0.0566 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (19)$$

$$\hat{\mathbf{T}}_{R/L} = \begin{bmatrix} 0.9153 & -0.2945 & 0.2747 & -0.1838 \\ 0.2589 & 0.9527 & 0.1589 & -0.0653 \\ -0.3085 & -0.0743 & 0.9483 & 0.1239 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (20)$$

Explicitando as matrizes de rotação e vetores de translação entre às câmeras:

$$\hat{\mathbf{R}}_{L/R} = \begin{bmatrix} 0.9153 & 0.2589 & -0.3085 \\ -0.2945 & 0.9527 & -0.0743 \\ 0.2747 & 0.1589 & 0.9483 \end{bmatrix}; \quad \hat{\mathbf{t}}_{L/R} = \begin{bmatrix} 0.2233 \\ 0.0173 \\ -0.0566 \end{bmatrix}; \quad (21)$$

$$\hat{\mathbf{R}}_{R/L} = \begin{bmatrix} 0.9153 & -0.2945 & 0.2747 \\ 0.2589 & 0.9527 & 0.1589 \\ -0.3085 & -0.0743 & 0.9483 \end{bmatrix}; \quad \hat{\mathbf{t}}_{R/L} = \begin{bmatrix} -0.1838 \\ -0.0653 \\ 0.1239 \end{bmatrix}; \quad (22)$$

3 Conclusão

Ao realizar a calibração intrínseca, observa-se que as realizações da matriz de parâmetros intrínsecos para as duas câmeras apresentam uma variação considerável, e que as coordenadas do centro optico em coordenadas de imagem, que pode ser obtido conhecendo o tamanho da imagem, se aproxima mais para a realização da câmera da direita, porém apresenta maior erro para a câmera da esquerda. Porém, ao realizar a média das duas realizações da matriz \hat{A} , obtém-se uma estimativa melhor para os parâmetros intrínsecos da câmera.

Dentre as causas da divergência da estimação obtida com o que se esperava basea-se primeiramente no método de obtenção de pontos. Dada as dimensões da caixa, e o traçado das arestas visíveis, o ponto oculto foi estimado por meio da estimação das arestas ocultas por paralelismo com as arestas visíveis, buscando manter também a noção de perspectiva das arestas. Além disso, os pontos foram tomados manualmente, e a caixa não apresentava arestas e vertices "afiados", permitindo uma incerteza ao traçar as linhas e tomar os pontos das interseções. Somado à isso, as coordenadas de câmera apresentam mais uma incerteza ao representar de forma discreta os pixels, onde no campo receptivo de cada pixel, todos os pontos presentes no mesmo serão representados pela mesma coordenada de imagem.

Dessa forma, popõem-se que para a melhor realização da calibração intrínseca, sejam capturadas inúmeras imagens do mesmo sólido, de ângulos diferentes, de forma que a média de todas as realizações da matriz de parâmetros intrínsecos representar com maior fidelidade a câmera, e com isso filtrar as incertezas que ocorrem durante o processo de calibração.

Para obtenção da transformação entre as câmeras, o uso de transformações homogêneas foi apenas um truque matemático para facilitar a concatenação de transformações de referencial e inversão das mesmas, onde translação e rotação são considerados simultaneamente. Dessa forma, no caso de um sistema de câmeras estéreo, tal transformação pode ser utilizada para a reconstrução 3D da cena, ou no caso de um sistema monoocular em movimento, a calibração intrínseca se faz útil para obter a rotação e a translação do sistema do instante de captura de uma imagem até o próximo, realizando assim o procedimento de odometria visual.

Anexos

Códigos fonte

Todos os códigos fonte e arquivos de dados utilizados para a elaboração deste documento podem ser encontrados no repositório do GitHub no link: github.com/toffanetto/im563.