

Calibração de câmera e Tamanho de objetos

Avaliação de calibração de câmera e captura de imagem de objetos – Princípios de Visão Computacional – 05/10/2014

Vitor Quaresma Silveira de Hollanda Ramos
11/0143566

Departamento de Ciência da Computação-CIC
UNB – Universidade de Brasília
Brasília, Brasil
vitor.q.s.h.r@gmail.com

Mateus Mendelson Esteves da Silva
11/0017579

Departamento de Ciência da Computação-CIC
UNB – Universidade de Brasília
Brasília, Brasil
m.mendelson.unb@gmail.com

I. OBJETIVOS

Este projeto tem, como objetivo prático, a calibração de uma câmera e cálculo das medições reais de um objeto contido em uma imagem capturada pela câmera calibrada.

O objetivo acadêmico deste projeto é aprender na prática os conceitos teóricos aprendidos em sala de aula. Estes conceitos compreendem os conceitos de calibração de câmeras e de segmentação de objetos. O conceito de calibração de câmeras foi posto em prática por meio da calibração de uma câmera utilizando funções do OpenCV. O conceito de cálculo das dimensões reais de um objeto foi posto em prática, obtendo-se as dimensões do objeto em pixels na imagem, assim como a distância do objeto até a câmera e, então, realizando as transformações necessárias para realizar o cálculo das dimensões reais do objeto.

II. INTRODUÇÃO

A. Pressupostos

A calibração de câmeras define as relações entre as coordenadas 2D de uma imagem e suas referentes coordenadas em uma reconstrução 3D, que representa o mundo real. O sistema de referência da câmera pode ser localizado em relação ao sistema de referência do mundo.

B. Parâmetros intrínsecos

Os parâmetros intrínsecos, obtidos a partir da calibração da câmera, são utilizados para fazer a ligação entre os pontos da imagem com seus respectivos pontos no sistema de referência da câmera, caracterizando as propriedades óticas, geométricas e digitais da câmera visualizadora. Como a câmera utilizada é uma câmera do tipo *pinhole*, temos 3 conjuntos de parâmetros intrínsecos:

- a projeção perspectiva, cujo único parâmetro é o foco;
- a transformação entre os sistemas de referência de câmera e pixel;

- a distorção geométrica introduzida pelo sistema ótico.

Para realizar a transformação das coordenadas (X_{im} , Y_{im}) da imagem com as coordenadas (x , y) no mesmo ponto no sistema de referência da câmera em uma imagem sem distorção, devemos usar as fórmulas:

$$\begin{aligned}x &= -(x_{im} - o_x)s_x \\ y &= -(y_{im} - o_y)s_y\end{aligned}$$

Sendo (O_x , O_y) as coordenadas em pixels do centro da imagem e (S_x , S_y) o tamanho do pixel, em milímetros.

Para uma imagem com distorção, devemos utilizar as fórmulas:

$$\begin{aligned}x &= x_d(1 + k_1r^2 + k_2r^4) \\ y &= y_d(1 + k_1r^2 + k_2r^4)\end{aligned}$$

Sendo (X_d , Y_d) as coordenadas dos pontos distorcidos e k_1 e k_2 coeficientes de distorção.

C. Parâmetros extrínsecos

Os parâmetros extrínsecos são parâmetros de calibração que são identificados quando se conhece um sistema de referência do mundo e se desconhece um sistema de referência da câmera. Quando se obtém uma imagem qualquer, geralmente o sistema de referência da câmera é desconhecido, portanto devemos obter o sistema de referência por meio de informações da imagem.

Os parâmetros extrínsecos usam, em suas fórmulas de transformação, um vetor 3D de translação T que descreve as posições relativas das origens dos dois sistemas de referência e uma matriz de rotação R , que realiza a correspondência dos eixos dos dois sistemas de referência. Sendo P_w o sistema de referência do mundo e P_c o sistema de referência da câmera, temos a fórmula abaixo:

$$P_c = R(P_w - T)$$

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

D. Modelo de câmera perspectiva

Podemos definir a matriz de parâmetros intrínsecos (M_{int}) e a matriz de parâmetros extrínsecos (M_{ext}) da seguinte maneira:

$$M_{int} = \begin{pmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{pmatrix}$$

$$M_{ext} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & -R_1^t T \\ r_{21} & r_{22} & r_{23} & -R_2^t T \\ r_{31} & r_{32} & r_{33} & -R_3^t T \end{pmatrix}$$

Definimos, então, a equação matricial:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{int} M_{ext} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

Onde x_1/x_3 é a coordenada X_{im} da imagem e x_2/x_3 é a coordenada Y_{im} da imagem.

Podemos assumir, para simplificar, que $o_x = o_y$ e $s_x = s_y$, reescrevendo, assim, a matriz de projeção na forma:

$$M = \begin{pmatrix} -fr_{11} & -fr_{12} & -fr_{13} & fR_1^t T \\ -fr_{21} & -fr_{22} & -fr_{23} & fR_2^t T \\ r_{31} & r_{32} & r_{33} & R_3^t T \end{pmatrix}$$

Assumindo o modelo de projeção fraca, no qual a equação de projeção pode ser escrita como

$$\left| \frac{[R_3^t(P_i - P')]}{[R_3^t(P' - T)]} \right| \ll 1$$

, obtemos a matriz de projeção a seguir:

$$M_{wp} = \begin{pmatrix} -fr_{11} & -fr_{12} & -fr_{13} & fR_1^t T \\ -fr_{21} & -fr_{22} & -fr_{23} & fR_2^t T \\ 0 & 0 & 0 & R_3^t (P' - T) \end{pmatrix}$$

E. Resolvendo o problema de calibração

Para resolver o problema de calibração, utilizamos imagens com um certo número de pontos na cena de coordenadas conhecidas, determinando suas projeções na imagem e resolvendo as equações para obter os parâmetros de transformação. Os parâmetros de transformação são os de translação e rotação e, com eles, podemos obter os parâmetros intrínsecos e extrínsecos. Com as imagens, obtemos as coordenadas (x_i, y_i, z_i) do objeto conhecido e (x'_i, y'_i, z'_i) deste objeto após as operações de translação e rotação. Podemos, então, obter a matriz de rotação R deste objeto, como mostrado em (1).

$$(x'_i, y'_i, z'_i, 1) = \begin{bmatrix} R_{11} & R_{12} & R_{13} & D_1 \\ R_{21} & R_{22} & R_{23} & D_2 \\ R_{31} & R_{32} & R_{33} & D_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (1)$$

Obtendo as equações:

$$X_i = f \frac{R_{11}x_i + R_{12}y_i + R_{13}z_i + D_1}{R_{31}x_i + R_{32}y_i + R_{33}z_i + D_3}$$

$$Y_i = f \frac{R_{21}x_i + R_{22}y_i + R_{23}z_i + D_2}{R_{31}x_i + R_{32}y_i + R_{33}z_i + D_3}$$

Podemos assumir que os pontos são co-planares e, assim, eliminamos os parâmetros em Z desta matriz. Desta forma, podemos assumir um foco conhecido $f = 1$ e reescrever as equações da seguinte forma:

$$X_i = \frac{R_{11}x_i + R_{12}y_i + T_1}{R_{31}x_i + R_{32}y_i + T_3}$$

$$Y_i = \frac{R_{21}x_i + R_{22}y_i + T_2}{R_{31}x_i + R_{32}y_i + T_3}$$

F. OpenCV

A calibração do OpenCV obtém os coeficientes de distorção k_1, k_2, p_1, p_2, k_3 .

À partir destes coeficientes, é possível corrigir o fator radial. Para tal, utiliza-se a seguinte fórmula:

$$\begin{aligned}x_{\text{corrected}} &= x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \\y_{\text{corrected}} &= y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)\end{aligned}$$

E, para correção do fator tangencial, é utilizada a fórmula:

$$\begin{aligned}x_{\text{corrected}} &= x + [2p_1 xy + p_2(r^2 + 2x^2)] \\y_{\text{corrected}} &= y + [p_1(r^2 + 2y^2) + 2p_2 xy]\end{aligned}$$

Ainda, para a correção de unidades, utilizamos a fórmula:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

, onde

- $w = z$;
- f_x e f_y são as distâncias focais da câmera e
- c_x e c_y são as coordenadas do centro ótico em pixels.

III. MATERIAIS E MÉTODOS

Este projeto consiste de dois programas. O primeiro é o programa de calibração, que irá calibrar a câmera usando os métodos do OpenCV e gerar os arquivos de calibração; o segundo programa irá calcular a largura e a altura de um objeto em uma imagem capturada pela câmera calibrada. Para compilar os programas, foi utilizado o script `build_all.sh`, fornecido no Moodle da disciplina.

A. Calibração da câmera

A câmera foi calibrada utilizando o tutorial do OpenCV para calibração de câmeras. Primeiramente, foram capturadas diversas imagens contendo um padrão conhecido na cena, o padrão utilizado foi de um tabuleiro de xadrez (*cheesboard*). O número de imagens a serem capturadas, assim como o número de frames entre as imagens, é especificado pelo usuário. O programa, então, captura imagens da câmera e avisa ao usuário se conseguiu detectar ou não o tabuleiro de xadrez na cena, guardando as imagens onde foi detectado o tabuleiro de xadrez. Quando o programa conseguir detectar o número de imagens que foi escolhido pelo usuário, é iniciado o processo de calibração, alocando matrizes e chamando a função do OpenCV `CVClibrateCamera2`, guardando os coeficientes de calibração nas matrizes intrínsecas e de distorção. Esses coeficientes são, então, guardados em um arquivo XML, que será utilizado mais tarde, pelo segundo programa para calcular as dimensões do objeto.

B. Medição do objeto

O programa responsável por calcular as medições do objeto é o programa *mede_objeto*. A classe principal do programa é a classe *objectMeasure*, responsável por todo o processo de medição e ajuste. Primeiramente o programa pede a distância do objeto a ser medido (eixo Z), em metros, e guardando na variável *dist*. Depois, o programa cria os objetos referentes aos quatro pares de coordenadas do objeto, ou seja, o ponto mais acima, mais a direita, mais à esquerda e mais abaixo. Sendo que estes pontos são objetos do tipo *Pair*. Os objetos do tipo *Pair* realizam os cálculos, transformando os pontos da imagem distorcida em pontos reais do mundo, utilizando os parâmetros dos arquivos XML gerados pela calibração. Depois, o programa abre a câmera, mostrando para o usuário a imagem que está sendo capturada pela câmera. Enquanto isso, o programa abre os arquivos XML de calibração *Intrinsics.xml* e *Distortion.xml*, gravando em *CvMats*. Depois, percorre estas *CvMats* para obter os coeficientes intrínsecos e de distorção, guardando em suas variáveis correspondentes e mostrando estes coeficientes para o usuário, na tela. O programa, então, continua mostrando a imagem da câmera para o usuário, até que este aperte “Esc”. Quando o usuário aperta “Esc”, o programa guarda a imagem capturada em uma matriz e faz a detecção de contornos da imagem, primeiramente gerando contornos coloridos e depois transformando estes contornos em uma imagem em escala de cinza para melhor visualização. Na imagem em escalas de cinza, é criada uma barra para poder alterar o limiar de detecção de bordas, sendo que quanto menor for o limiar, mais bordas serão detectadas. Então, o usuário deve ajustar esse limiar até que sejam detectadas as bordas do objeto de interesse. O programa espera que o usuário clique nos pontos de coordenadas que serão utilizados na medição das dimensões, o usuário deve então, na imagem de contornos em escala de cinza, clicar primeiramente nos pontos que serão utilizados para calcular a largura, ou seja, pontos mais à direita e mais à esquerda, e, depois, nos pontos que serão utilizados para calcular a altura, ou seja, pontos mais abaixo e mais acima. Estes cliques são capturados pela função *click_callback*, após o usuário terminar estes cliques, o programa irá calcular as dimensões do objeto, primeiramente calculando o valor de r e depois realizando os cálculos matriciais para obter os valores de x_p e y_p a partir de x_d e y_d , como na equação abaixo:

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \begin{bmatrix} x_d \\ y_d \end{bmatrix} + \begin{bmatrix} 2p_1 x_d y_d + p_2(r^2 + 2x_d^2) \\ p_1(r^2 + 2y_d^2) + 2p_2 x_d y_d \end{bmatrix}$$

O programa, então, mostra ao usuário as dimensões do objeto, tanto em pixels quanto em metros.

IV. RESULTADOS

A seguir, serão mostrados os resultados obtidos pelo programa para 2 objetos distintos, em 3 posições diferentes.

A. Objeto 1 – Quadro

1) Medidas reais do objeto, obtidas por uma fita métrica

a) Largura: 0.945m

b) Altura: 0.649m

2) Dimensões do objeto, obtidas quando este estava próximo à câmera.

- a) Distância da câmera: 1.46 m
- b) Largura: 0.723 m
- c) Largura: 0.204 m



Fig. 1. Imagem próxima



Fig. 2. Imagem próxima após aplicar a detecção de bordas

3) Dimensões do objeto, obtidas quando este estava a uma distância média da câmera.

- a) Distância da câmera: 2.67m
- b) Largura: 0.452m
- c) Altura: 0.130m



Fig. 3. Imagem com distância média

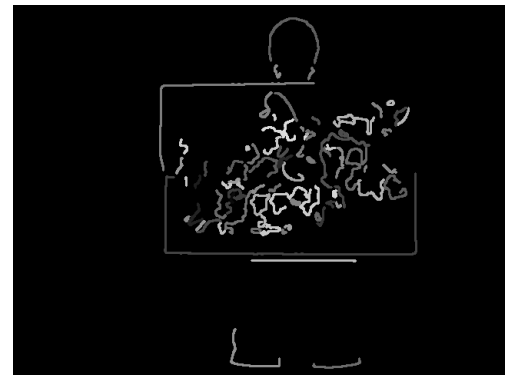


Fig. 4. Imagem com distância média após a detecção de bordas

4) Dimensões do objeto, obtidas quando este estava longe da câmera.

- a) Distância da câmera: 3.469m
- b) Largura: 0.367m
- c) Altura: 0.172 m



Fig. 5. Imagem com distância grande



Fig. 6. Imagem com distância grande após a detecção de bordas

B. Objeto 2 – Encosto de cadeira

1) Medidas reais do objeto, obtidas por uma fita métrica

- a) Largura: 0.398 m
- b) Altura: 0.251 m

2) Dimensões do objeto obtidas quando este estava próximo à câmera.

- a) Distância da câmera: 1.184m
- b) Largura: 0.200 m
- c) Altura: 0.135 m



Fig. 7. Imagem próxima



Fig. 8. Imagem próxima após aplicar a detecção de bordas

3) Dimensões do objeto obtidas quando este estava a uma distância média da câmera.

- a) Distância da câmera: 1.915m
- b) Largura: 0.185m
- c) Altura: 0.106m

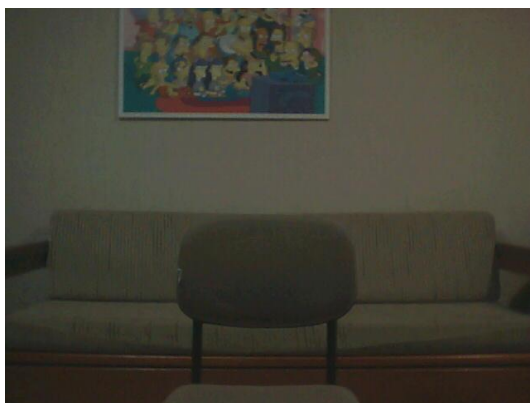


Figura 9 Imagem com distância média

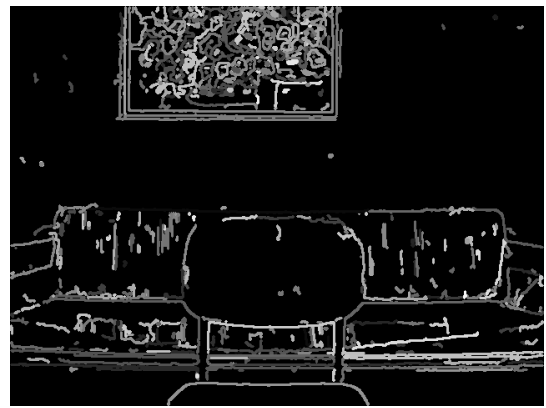


Fig. 10. Imagem com distância média após detecção de bordas

4) Dimensões do objeto obtidas quando este estava longe da câmera.

- a) Distância da câmera: 2.55m
- b) Largura: 0.203m
- c) Altura: 0.107m



Fig. 11. Imagem com distância grande

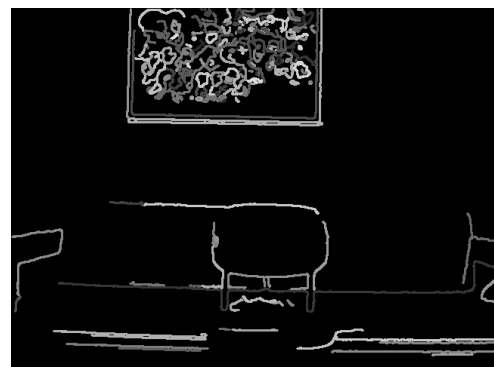


Fig. 12. Imagem com distância grande após a detecção de bordas

V. CONCLUSÃO

Podemos perceber, pelos resultados, que quanto mais próximo o objeto estava da distância de onde o padrão de calibração estava, maior a precisão da dimensão de largura e altura do objeto calculadas, ou seja, mais próximas as dimensões de largura e altura da realidade.

Podemos perceber no Objeto 1 que, quando este estava mais próximo, havia uma melhor precisão de suas dimensões de largura e altura. Isto ocorre porque o padrão de calibração estava próximo da câmera no momento da calibração.

No Objeto 2, as dimensões também foram mais precisas quando o objeto estava mais próximo da câmera, novamente porque estava mais próximo de onde o padrão de calibração estava no momento da calibração.

BIBLIOGRAFIA

- [1] Aula_13 – PVC , Prof. Dr. Flávio Vidal
- [2] Aula_14 – PVC , Prof. Dr. Flávio Vidal
- [3] Open-CV documentation
http://docs.opencv.org/doc/tutorials/calib3d/camera_calibration/camera_calibration.html