Visão computacional

Lista 4 - Calibração de câmeras usando o Método linear

Isabelle Rodrigues Vaz de Melo

16 de setembro de 2023

1 Plotando os pontos

A questão fornece os pontos em 3D do vértice de um cubo, e de maneira análoga, os pontos respectivos na imagem em pixels detectados por uma câmera de um robô. A figura 1 mostra a imagem encontrada ao plotar os pontos.

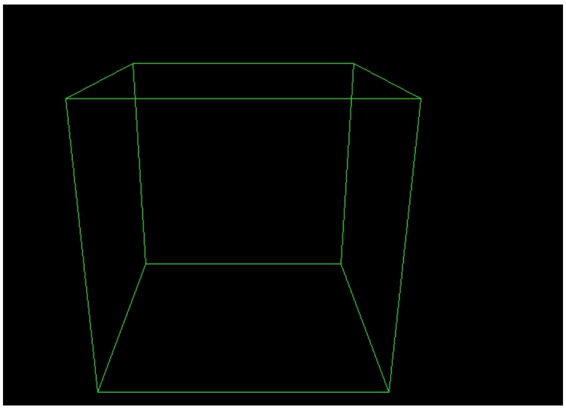

Figura 1: Pontos do cubo plotados formando a imagem

2 Matriz de projeção

As correspondências na imagem da câmera em 2D podem ser mapeadas para as coordenadas do mundo em 3D através de uma matriz de projeção $\mathbf{P}$. Não podemos esquecer que como estamos trabalhando em um espaço projetivo, é preciso homogenizar as coordenadas. A transformação é dada por:

$$\mathbf{x} = \mathbf{P}\mathbf{X} \tag{1}$$

Em que a equação pode ser reescrita de maneira a explicitar as coordenadas:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{2}$$

E assim como fizemos na lista 01, podemos reescrever este sistema de equações na forma $\mathbf{Ap} = \mathbf{0}$ de forma que o vetor $\mathbf{p}$ seja um vetor coluna das entradas da matriz $\mathbf{P}$ com 12 elementos.

Para N pontos:

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1X_1 & -x_1Y_1 & -x_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1X_1 & -y_1Y_1 & -y_1Z_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -x_2X_2 & -x_2Y_2 & -x_2Z_2 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -y_2X_2 & -y_2Y_2 & -y_2Z_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -x_NX_N & -x_NY_N & -x_NZ_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -y_NX_N & -y_NY_N & -y_NZ_N \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3}$$

Utilizamos a decomposição SVD da matriz **A** para resolução, e assim O vetor **p** é o autovetor correspondente ao menor valor singular de **A**.

A matriz de projeção encontrada com este procedimento foi:

$$P = \begin{bmatrix} 1.92485371e-01 & 2.82704820e-02 & 7.86051209e-02 & 7.34592038e-01 \\ 2.42367404e-07 & 2.04411170e-01 & 6.91856900e-05 & 6.12009747e-01 \\ 9.73363067e-10 & 9.42349769e-05 & 2.62017163e-04 & 2.44863818e-03 \end{bmatrix} \tag{4}$$

3 Parâmetros intrínscecos e extrínsecos

Para encontrar os parâmetros intrínscecos e extrínsecos da câmera ($\mathbf{K}$, $\mathbf{R}$, $\mathbf{t}$), devemos decompor a matriz de projeção. Para encontrar **K** e **R**, utilizamos uma fatorização RQ no bloco 3x3 esquerdo da matriz de projeção.

Como a fatorização não é única devido às possíveis formas de **R**, garantimos que a diagonal da matriz de calibração **K** seja positiva, ou seja, as distâncias focais estejam positivas.

$$\mathbf{K} = \begin{bmatrix} fx & s & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

Em que os f's representam as distâncias focais da câmera, s o parâmetro de *shear*, e c's as posições do centro ótico da câmera.

Pela decomposição e correção de valores, obtivemos os seguintes resultados:

$$\mathbf{K} = \begin{bmatrix} 6.91278652e+02 & -2.33581702e-06 & 3.00002308e+02 \\ 0.00000000e+00 & 6.90706665e+02 & 2.48678031e+02 \\ 0.00000000e+00 & 0.00000000e+00 & 1.00000000e+00 \end{bmatrix} \tag{6}$$

$$\mathbf{R} = \begin{bmatrix} -1.00000000e+00 & 1.18457256e-06 & 3.28884905e-06 \\ -1.62948648e-09 & -9.40991752e-01 & 3.38429494e-01 \\ -3.49567413e-06 & -3.38429494e-01 & -9.40991752e-01 \end{bmatrix} \tag{7}$$

Por fim, precisamos encontrar a matriz de translação. A forma utilizada foi escrevendo-a em função da matriz de calibração e o bloco 3x1 direito da matriz de projeção:

$$\mathbf{t} = \mathbf{K}^{-1} \begin{bmatrix} p_{14} \\ p_{24} \\ p_{34} \end{bmatrix} \tag{8}$$

Já com as matrizes com as correções de valores. O resultado computacional encontrado foi de:

$$\mathbf{t} = \begin{bmatrix} -7.33123607e-09 \\ 4.46965994e-06 \\ 2.44863818e-03 \end{bmatrix} \tag{9}$$

Podemos notar que o fator s é bem próximo de zero, como esperado, e que as coordenadas de translação também são valores bem pequenos tendendo a 0, o que mostra que não há uma translação considerável da câmera com relação ao objeto.