CC222 – Visão Computacional

Câmeras

Instituto Tecnológico de Aeronáutica

Prof. Carlos Henrique Q. Forster – Sala 121 IEC

forster@ita.br

ramal 5981

Tópicos da aula

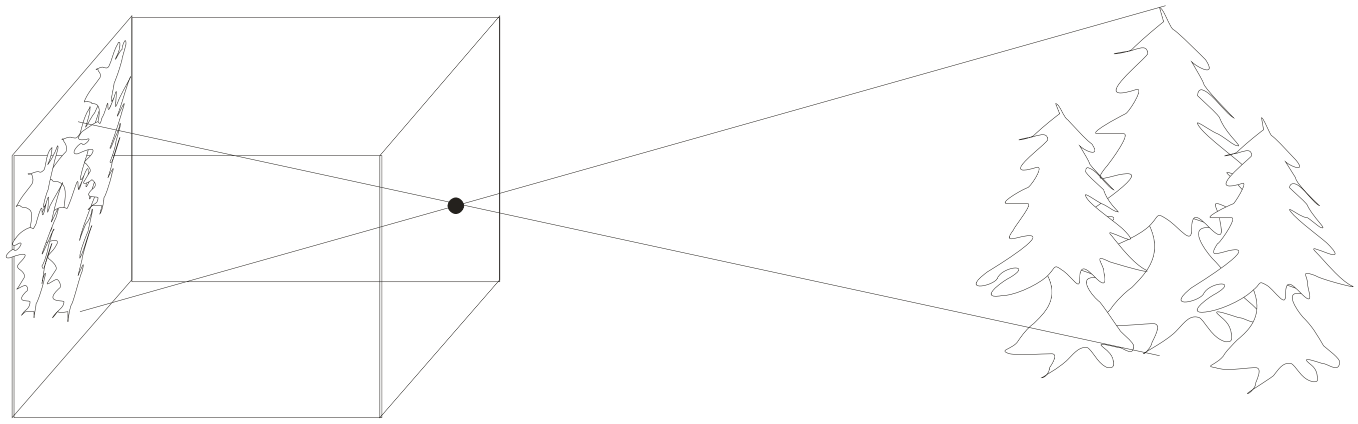
- Modelos de câmeras
- Aquisição de imagens
- Parâmetros da câmera
- Recuperação da matriz de projeção
- Calibração de Tsai

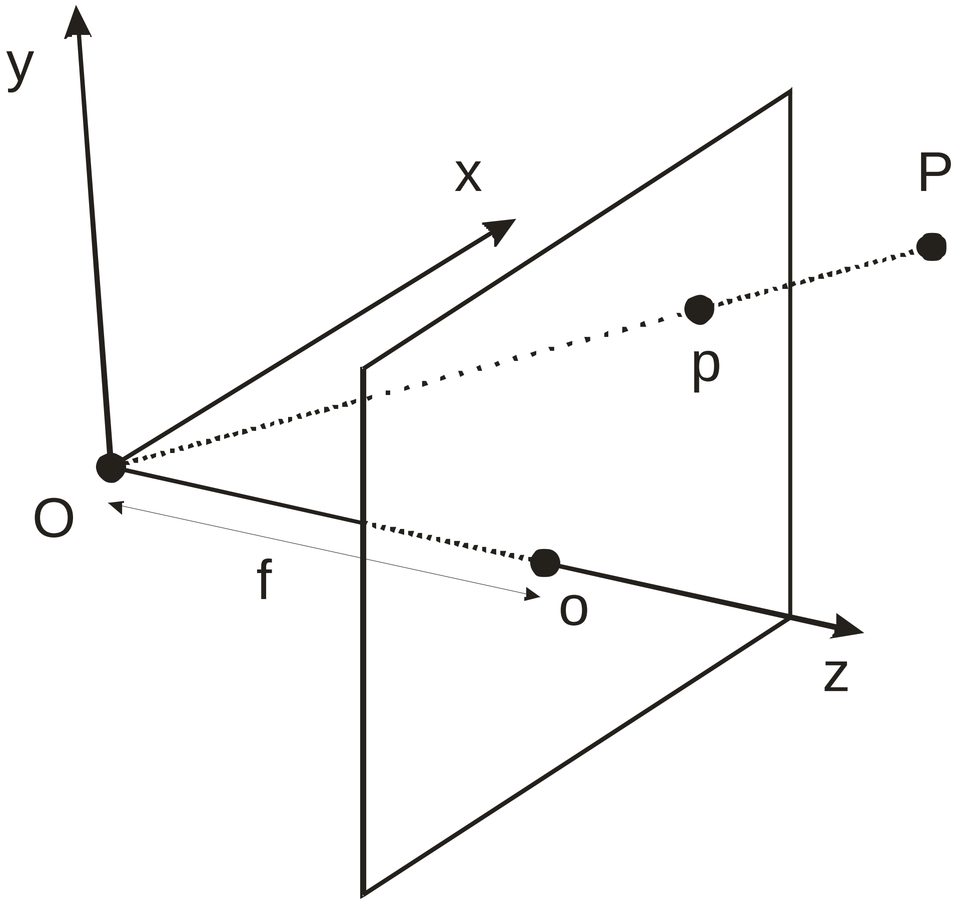
Livro para acompanhar essa aula

Trucco e Verri – Capítulo 6

Modelo de câmera de orifício

Câmera com projeção perspectiva





$$x = \frac{f}{Z} X \quad e \quad y = \frac{f}{Z} Y$$

Centro ou foco de projeção

Distância focal

Plano imagem

Eixo óptico

Ponto principal

Modelo de lente fina

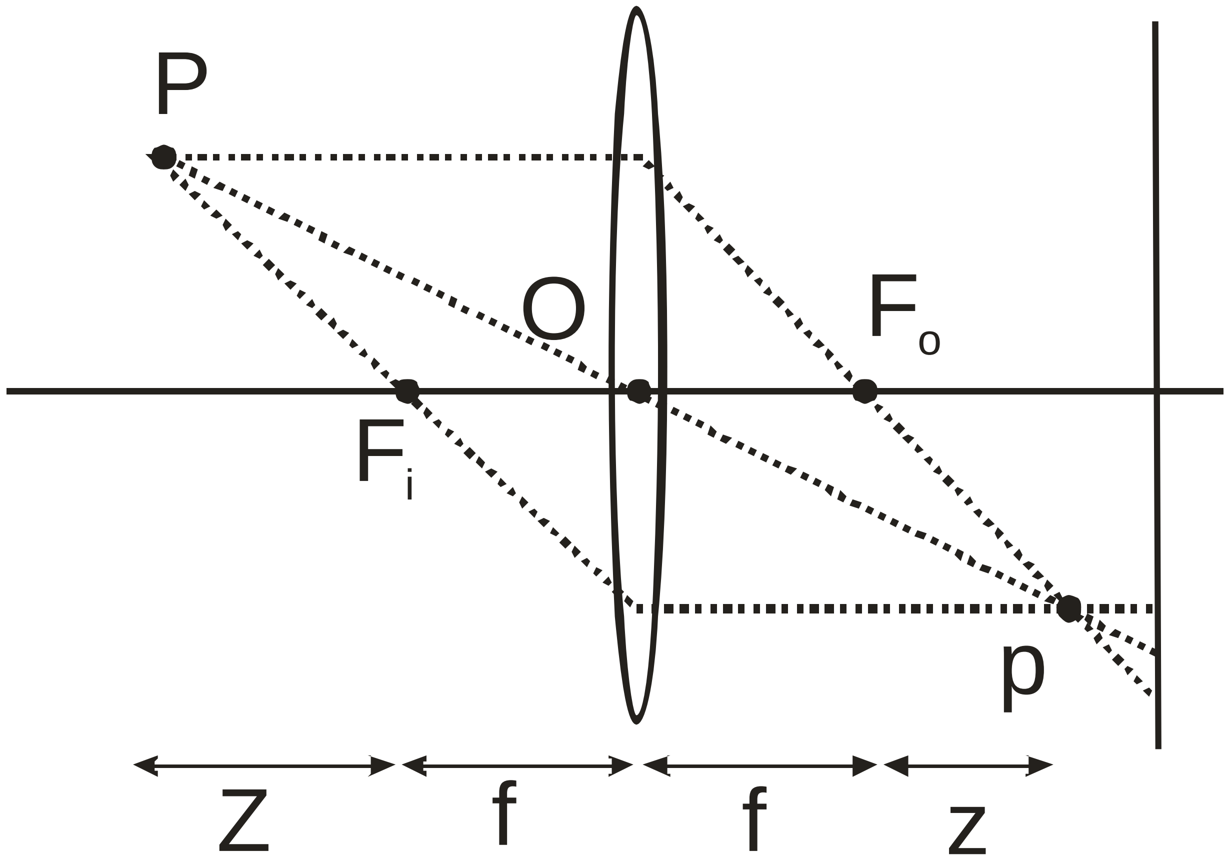
O modelo de câmera de orifício considera um furo mínimo sem dimensões.

Não modela adequadamente a radiometria da formação da imagem. Por exemplo, considerando o comprimento de onda da luz, para o qual é necessária uma abertura de tamanho finito.

Regras para raios em lentes finas:

- Todo raio que passa pelo foco, sai paralelo ao eixo óptico.
- Todo raio que incide paralelo ao eixo óptico, sai passando pelo foco.
- Todo raio que passa pelo centro da lente, sai não defletido.

A imagem de um ponto não focado corretamente resulta em uma região.



$$\frac{1}{Z+f}+\frac{1}{z+f}=\frac{1}{f}$$ equação fundamental de lentes finas

Perspectiva fraca

Aproximação da transformação projetiva por uma transformação afim.

Considera distâncias relativas pequenas em relação à distância ao centro de projeção.

Se $\bar{Z}$ for a distância média dos pontos do objeto ao centro de projeção, assume-se a coordenada Z constante e igual a $\bar{Z}$.

$$x = \frac{f}{\bar{Z}} X \quad e \quad y = \frac{f}{\bar{Z}} Y$$

Aquisição da imagem

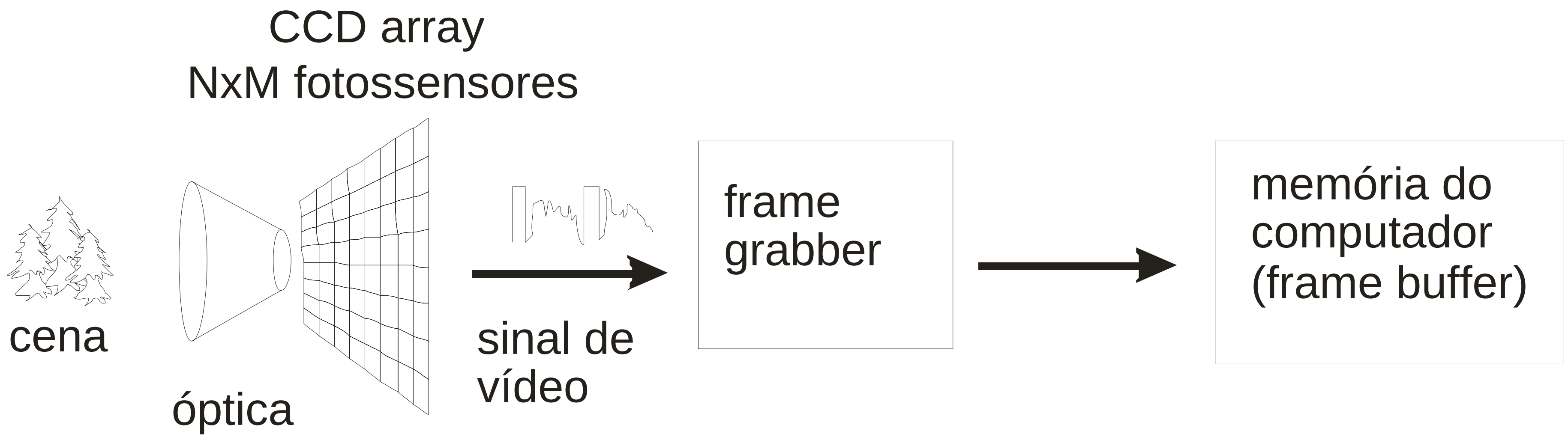


Imagem Digital

Pixels de coordenadas (i,j) com valores inteiros associados (tipicamente 0 a 255) ou triplas de valores inteiros (tipicamente RGB (vermelho, verde, azul)).

Imagem contínua: valores reais para coordenadas reais

Imagem discreta: valores reais para coordenadas inteiras (pixels são amostras de uma imagem contínua distribuidos espacialmente).

Amostragem Espacial

Se a distância entre CCDs vizinhos é d , então a freqüência espacial máxima que pode ser recuperada é

$$\nu_c = \frac{1}{2d}$$ (taxa de amostragem de Nyquist)

Considerando a óptica, não é possível capturar freqüências maiores que

$$\nu_c' = \frac{a}{\lambda\ f}$$, onde a é a abertura (diâmetro da lente), f é a distância focal e λ é o comprimento de onda.

Em geral, ν_c é muito menor que ν_c' .

Estimativa de ruído no CCD

N imagens da mesma cena sob mesmas condições: $E_0, E_1, \cdots, E_{N-1}$

Para cada pixel (i,j):

$$\overline{E}(i,j) = \frac{1}{N}\sum_{k=0}^{N-1} E_k(i,j)$$

$$\sigma(i,j) = \sqrt{\frac{1}{N-1}\sum_{k=0}^{N-1}\left(\overline{E}(i,j) - E_k(i,j)\right)^2}$$

Parâmetros de Câmeras

Equações relacionando coordenadas 3D com coordenadas da imagem.

Sistema de referência da câmera em relação a um sistema de referência global.

Distorções na óptica e posicionamento do CCD.

Parâmetros extrínsecos

Correspondem ao posicionamento e a orientação da câmera no espaço.

Movimento rígido: 6 DOF (graus de liberdade).

Corresponde a uma transformação de rotação em 3D e uma transformação de translação em 3D.

Parâmetros intrínsecos

Relacionam coordenadas de pontos da imagem com coordenadas do sistema de referência da câmera.

Correspondem a uma transformação em 2D.

Pode ser incluída uma componente de distorção não-linear na transformação.

Parâmetros intrínsecos

Distância focal f

Posição do ponto principal o_x, o_y

Tamanho efetivo do pixel (e aspecto ou aspect ratio) s_x, s_y

$$x = -(x_{im} - o_x)s_x$$

$$y = -(y_{im} - o_y)s_y$$

Parâmetros intrínsecos

Modelo de distorção radial

$$x = x_d(1 + k_1 r^2 + k_2 r^4)$$

$$y = y_d(1 + k_1 r^2 + k_2 r^4)$$

$$r^2 = x_d^{\ 2} + y_d^{\ 2}$$

Outros parâmetros:

Ângulo entre eixo x e eixo y

Rotação do CCD

Matrizes de câmera

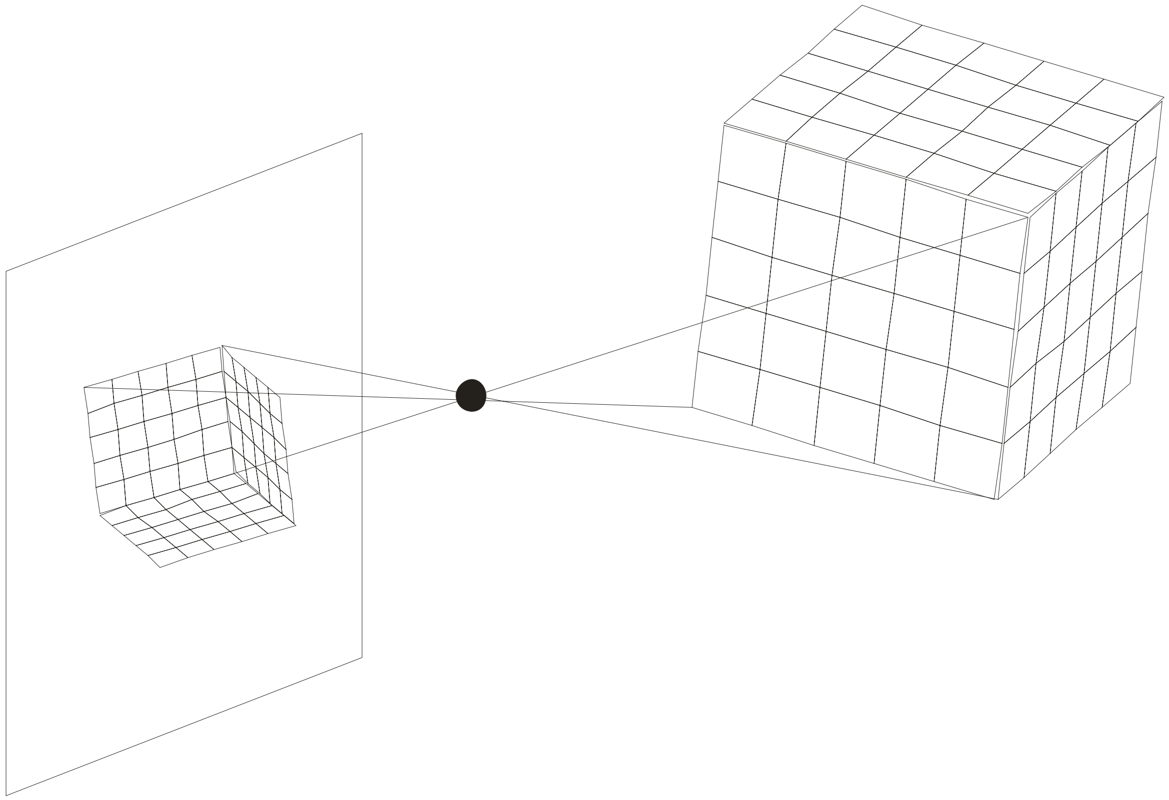
$$M_{\text{int}} = \begin{bmatrix} -f/s_x & & o_x \\ & -f/s_y & o_y \\ & & 1 \end{bmatrix}$$

$$M_{ext} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & -R_1^T\mathbf{t} \\ r_{21} & r_{22} & r_{23} & -R_2^T\mathbf{t} \\ r_{31} & r_{32} & r_{33} & -R_3^T\mathbf{t} \end{bmatrix}, \text{ onde } R_i = \begin{bmatrix} r_{i1} \\ r_{i2} \\ r_{i3} \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = M_{\text{int}} M_{ext} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \qquad \begin{aligned} x_{im} &= u/w \\ y_{im} &= v/w \end{aligned}$$

Calibração de câmera

Determinar os parâmetros da câmera, baseado em objetos com pontos de coordenadas conhecidas e associação desses pontos com pontos da imagem.



Estimação direta da matriz de projeção

$$\begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} = M \begin{bmatrix} X_i^w \\ Y_i^w \\ Z_i^w \\ 1 \end{bmatrix}$$

$$x_i = \frac{u_i}{w_i} = \frac{m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14}}{m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34}}$$

$$y_i = \frac{v_i}{w_i} = \frac{m_{21}X_i + m_{22}Y_i + m_{23}Z_i + m_{24}}{m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34}}$$

$$m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14} - m_{31}x_iX_i - m_{32}x_iY_i - m_{33}x_iZ_i - m_{34}x_i = 0$$

Constrói-se um sistema linear homogêneo $A\mathbf{m} = \mathbf{0}$, onde $m = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{21} & m_{22} & m_{23} & m_{24} & m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}^T$ representa os parâmetros a serem determinados da matriz de projeção e a matriz A é dada por

$$A = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -x_2 X_2 & -x_2 Y_2 & -x_2 Z_2 & -x_2 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -y_2 X_2 & -y_2 Y_2 & -y_2 Z_2 & -y_2 \\ & \vdots & & & & & & & & & & \\ & & & & & & & & & & & \end{bmatrix}$$

Para cada associação de um ponto da imagem com um ponto de coordenadas conhecidas no espaço, obtenho 2 equações. Há 12 incógnitas, mas como o sistema é homogêneo o rank máximo de A é 11.

Primeira forma de solução do sistema

Fazemos $m_{34} = 1$ porque qualquer múltiplo da solução não-trivial é também solução.

O sistema resultante é

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1X_1 & -x_1Y_1 & -x_1Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1X_1 & -y_1Y_1 & -y_1Z_1 \\ X_2 & Y_2 & Z_2 & 1 & 0 & 0 & 0 & 0 & -x_2X_2 & -x_2Y_2 & -x_2Z_2 \\ 0 & 0 & 0 & 0 & X_2 & Y_2 & Z_2 & 1 & -y_2X_2 & -y_2Y_2 & -y_2Z_2 \\ & \vdots & & & & & & & & & \\ & & & & & & & & & & \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ \vdots \\ m_{33} \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ \vdots \end{bmatrix}$$

Para 11 equações obtidas de 6 associações de pontos (para configurações não degeneradas) a solução é única.

Para mais equações, resolver utilizando a pseudo-inversa (mínimos quadrados).

Segunda forma de solução

Encontrar o espaço nulo da matriz A.

Qualquer matriz M resultante de vetor m pertencente ao espaço nulo de A satisfaz o sistema de equações.

Idealmente, se a configuração não for degenerada, o espaço nulo terá apenas uma dimensão livre que corresponde aos múltiplos de M.

Para encontrar o espaço nulo de A, decompô-la em valores singulares (SVD):

$$A = UDV^T$$

Obter m da coluna de V correspondente ao menor valor singular de A. (Muitos sistemas já ordenam os valores singulares).

Parâmetros da câmera a partir da matriz de projeção

$$M = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

$$M = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

Onde

$$f_x = f / s_x$$

$$f_y = f / s_y$$

$$q_1 = \begin{bmatrix} m_{11} \\ m_{12} \\ m_{13} \end{bmatrix}, q_2 = \begin{bmatrix} m_{21} \\ m_{22} \\ m_{23} \end{bmatrix}, q_3 = \begin{bmatrix} m_{31} \\ m_{32} \\ m_{33} \end{bmatrix}, q_4 = \begin{bmatrix} m_{14} \\ m_{24} \\ m_{34} \end{bmatrix}$$

Determinar fator de escala γ

$$\sqrt{m_{31}{}^2 + m_{32}{}^2 + m_{33}{}^2} = |\gamma|\sqrt{r_{31}{}^2 + r_{32}{}^2 + r_{33}{}^2} = |\gamma|$$

Dividir todos elementos m_{ij} por $|\gamma|$.

Obtendo da última linha da matriz M:

$$T_z = \pm m_{34}$$

$$r_{31} = \pm m_{31}, r_{32} = \pm m_{32}, r_{33} = \pm m_{33},$$

Obtenção do ponto principal da imagem

$$q_1 = \begin{bmatrix} -f_x r_{11} + o_x r_{31} \\ -f_x r_{12} + o_x r_{32} \\ -f_x r_{13} + o_x r_{33} \end{bmatrix}, q_3 = \begin{bmatrix} r_{31} \\ r_{32} \\ r_{33} \end{bmatrix}$$

$$q_1^T q_3 = -f_x r_{11} r_{31} + o_x r_{31}^2 - f_x r_{12} r_{32} + o_x r_{32}^2 - f_x r_{13} r_{33} + o_x r_{33}^2 =$$

$$= -f_x \left(r_{11} r_{31} + r_{12} r_{32} + r_{13} r_{33} \right) + o_x \left(r_{31}^2 + r_{32}^2 + r_{33}^2 \right) = o_x$$

$$o_y = q_2^T q_3$$

Obtenção da distância focal em pixels

$$q_1 = \begin{bmatrix} -f_x r_{11} + o_x r_{31} \\ -f_x r_{12} + o_x r_{32} \\ -f_x r_{13} + o_x r_{33} \end{bmatrix}$$

$$q_1{}^T q_1 = f_x{}^2\left(r_{11}{}^2 + r_{12}{}^2 + r_{13}{}^2\right) + o_x{}^2\left(r_{31}{}^2 + r_{32}{}^2 + r_{33}{}^2\right) - 2 f_x o_x \left(r_{11}r_{31} + r_{12}r_{32} + r_{13}r_{33}\right)$$

$$= f_x{}^2 + o_x{}^2$$

$$f_x = \sqrt{q_1{}^T q_1 - o_x{}^2}$$

$$f_y = \sqrt{q_2{}^T q_2 - o_y{}^2}$$

Obtenção do restante dos parâmetros.

$$r_{11} = \pm\left(o_x m_{31} - m_{11}\right)/ f_x \text{ , } r_{12} = \pm\left(o_x m_{32} - m_{12}\right)/ f_x \text{ , } r_{13} = \pm\left(o_x m_{33} - m_{13}\right)/ f_x \text{ ,}$$

$$r_{21} = \pm\left(o_y m_{31} - m_{21}\right)/ f_y \text{ , } r_{22} = \pm\left(o_y m_{32} - m_{22}\right)/ f_y \text{ , } r_{23} = \pm\left(o_y m_{33} - m_{23}\right)/ f_y \text{ ,}$$

$$T_x = \pm\left(o_x T_z - m_{14}\right)/ f_x \text{ ,}$$

$$T_y = \pm\left(o_y T_z - m_{24}\right)/ f_y$$

Calibração pelo método de Tsai

Assumir que as coordenadas da imagem partem do ponto principal.

$$x_i = -f_x \frac{r_{11}X_i + r_{12}Y_i + r_{13}Z_i + T_x}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + T_z}$$

$$y_i = -f_y \frac{r_{21}X_i + r_{22}Y_i + r_{23}Z_i + T_y}{r_{31}X_i + r_{32}Y_i + r_{33}Z_i + T_z}$$

Igualando os denominadores

$$x_i f_y \left(r_{21}X_i + r_{22}Y_i + r_{23}Z_i + T_y\right) = y_i f_x \left(r_{11}X_i + r_{12}Y_i + r_{13}Z_i + T_x\right)$$

Dividindo por f_y, e substituindo $\alpha = f_x / f_y$

$$x_i X_i v_1 + x_i Y_i v_2 + x_i Z_i v_3 + x_i v_4 - y_i X_i v_5 - y_i Y_i v_6 - y_i Z_i v_7 - y_i v_8 = 0$$

Onde

$$v_1 = r_{21}, v_2 = r_{22}, v_3 = r_{23}, v_4 = T_y, v_5 = \alpha r_{11}, v_6 = \alpha r_{12}, v_7 = \alpha r_{13}, v_8 = \alpha T_x$$

Sistema de N equações lineares para N pares de pontos

$$A\mathbf{v} = \mathbf{0}$$,

Com

$$A = \begin{bmatrix} x_1 X_1 & x_1 Y_1 & x_1 Z_1 & x_1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \\ \vdots & & & & & & & \end{bmatrix}$$

N>=7, A tem rank 7, a solução é o espaço nulo de A.

Recuperando o fator de escala γ

$$\sqrt{{v_1}^2 + {v_2}^2 + {v_3}^2} = |\gamma|\sqrt{{r_{21}}^2 + {r_{22}}^2 + {r_{23}}^2} = |\gamma|$$

Recuperando o aspecto α

$$\sqrt{{v_5}^2 + {v_6}^2 + {v_7}^2} = \alpha|\gamma|\sqrt{{r_{11}}^2 + {r_{12}}^2 + {r_{13}}^2} = \alpha|\gamma|$$

Encontrando T_z e f_x

$$x_i\left(r_{31}X_i + r_{32}Y_i + r_{33}Z_i + T_z\right) = -f_x\left(r_{11}X_i + r_{12}Y_i + r_{13}Z_i + T_x\right)$$

Resolver um sistema super-determinado da forma

$$A\begin{bmatrix} T_z \\ f_x \end{bmatrix} = \mathbf{b}$$

$$\begin{bmatrix} x_1 & r_{11}X_1 + r_{12}Y_1 + r_{13}Z_1 + T_x \\ \vdots & \end{bmatrix}\begin{bmatrix} T_z \\ f_x \end{bmatrix} = \begin{bmatrix} -x_1\left(r_{31}X_1 + r_{32}Y_1 + r_{33}Z_1\right) \\ \vdots \end{bmatrix}$$

Estimando o centro da imagem

Seja T o triângulo no plano-imagem definido pelos pontos de fuga de três conjuntos mutuamente ortogonais de retas paralelas no espaço. O centro da imagem é o ortocentro de T.