

CC222 – Visão Computacional

Reconstrução por Estéreo

Instituto Tecnológico de Aeronáutica

Prof. Carlos Henrique Q. Forster – Sala 121 IEC

forster@ita.br

ramal 5981

Tópicos da aula

- Auto-calibração de um par estéreo
- Análise da matriz fundamental
- Reconstrução 3D no caso calibrado
- Reconstrução 3D no caso não calibrado
- Reconstrução densa (shape-from-stereo)

Livro para acompanhar essa aula

Trucco e Verri – cap 7

Calibrando um sistema estéreo

Se não for possível calibrar cada câmera individualmente (por exemplo: o sistema estéreo pode se tratar de uma câmera em movimento), é ainda possível calibrar o par estéreo se for conhecida a associação de 8 pontos da imagem esquerda com os 8 pontos correspondentes (homólogos) da imagem direita.

A auto-calibração do sistema estéreo consiste da estimação da matriz fundamental a partir da associação de 8 pontos. (O algoritmo é conhecido algoritmo dos 8 pontos).

Se forem conhecidos os parâmetros intrínsecos das câmeras, então a matriz essencial também pode ser obtida.

O algoritmo dos 8 pontos

Uma associação de um ponto pE da imagem esquerda com um ponto pD da imagem direita cria uma restrição sobre os elementos da matriz fundamental dada pela equação de Longuet-Higgins

$$\tilde{p}_D^T \cdot F \cdot \tilde{p}_E = 0$$

Expandindo
$$\begin{bmatrix} x_D \\ y_D \\ 1 \end{bmatrix}^T \cdot \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \cdot \begin{bmatrix} x_E \\ y_E \\ 1 \end{bmatrix} = 0$$

$$x_D x_E f_{11} + y_D x_E f_{21} + x_E f_{31} + x_D y_E f_{12} + y_D y_E f_{22} + y_E f_{32} + x_D f_{13} + y_D f_{23} + f_{33} = 0$$

Sistema de 8 equações e 9 incógnitas → Buscar o espaço nulo

Esse método da forma apresentada é numericamente instável, sendo necessárias normalizações (não definir como 1 as coordenadas homogêneas).

Um outro problema é a possível obtenção da matriz F com rank completo, quando sabemos que deve ter rank 2. O truque que se utiliza é decompor em SVD e fixar o menor valor singular em zero, então recompor a matriz F .

Localização dos epipólos dada matriz fundamental

Como o epipólo \tilde{e}_E da imagem esquerda pertence a qualquer linha epipolar para qualquer ponto \tilde{p}_D da imagem direita, vale a relação

$$\tilde{p}_D^T \cdot F \cdot \tilde{e}_E = 0$$

Como F não é identicamente nula, então isto só é possível no caso em que

$$F \cdot \tilde{e}_E = 0$$

Assim, decompondo F em SVD, a coluna de V correspondente ao valor singular nulo representa o epipólo \tilde{e}_E . A coluna de U correspondente ao valor singular nulo representa o epipólo \tilde{e}_D .

Obtenção da linha de base pela matriz essencial

Lembrando que $E = RS$

$$E^T E = S^T R^T RS$$

Mas, como R é matriz de rotação, sendo ortonormal $R^T R = I$

$$E^T E = S^T S = \begin{bmatrix} T_y^2 + T_z^2 & -T_x T_y & -T_x T_z \\ -T_y T_x & T_z^2 + T_x^2 & -T_y T_z \\ -T_z T_x & -T_z T_y & T_x^2 + T_y^2 \end{bmatrix}$$

O tamanho da linha de base é obtido da soma dos elementos da diagonal (traço)

$$\|T\| = \sqrt{\frac{1}{2} \text{Traço}(E^T E)}$$

O vetor T pode ser obtido de qualquer coluna de $E^T E$ por substituição (mas, perde-se o sinal).

Obtenção da orientação relativa pela matriz essencial

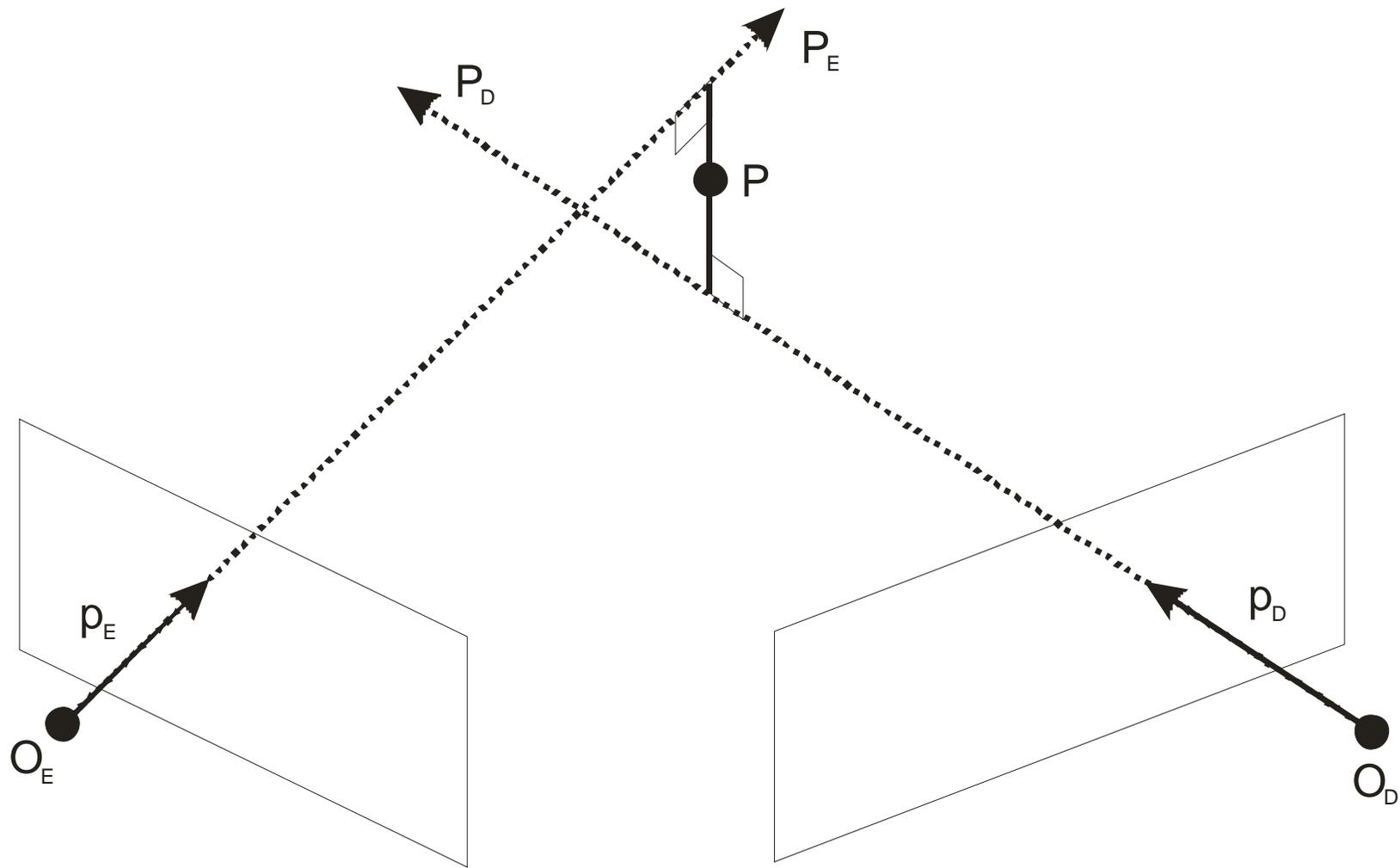
Sejam $\hat{E} = E / \|T\|$ e $\hat{T} = T / \|T\|$

Definimos $w_i = \hat{E}_i \times \hat{T}$, $i = 1, 2, 3$, onde \hat{E}_i são linhas de \hat{E} .

A matriz de rotação é dada por $R_i = w_i + w_j \times w_k$ para as triplas (i,j,k) iguais a (1,2,3), (2,3,1) e (3,1,2).

Problemas na determinação do sinal do vetor T e da matriz R.

Reconstrução – Caso calibrado



Como nossas medidas são imprecisas, as retas podem não ser concorrentes, mas reversas. Então é necessário encontrar um ponto de distância mínima às duas retas.

$a \cdot p_E$ corresponde ao raio P_E e $T + b \cdot R^T p_D$ ao raio P_D .

w é um vetor ortogonal às duas retas, portanto, uma escolha seria $p_E \times R^T p_D$.

A soma dos 3 vetores deve ser igual ao vetor T

$$a \cdot p_E - b \cdot R^T p_D + c \cdot (p_E \times R^T p_D) = T$$

Resolvemos o sistema linear para (a, b, c) e encontramos o ponto médio do segmento entre as duas retas. $\frac{1}{2}(a \cdot p_E + T + b \cdot R^T p_D)$

Reconstrução – Caso não calibrado

Primeiro caso: Conhecemos os parâmetros intrínsecos das câmeras.

- Calibramos com o algoritmo dos 8 pontos.
- Estimamos a matriz essencial
- Normalizamos a matriz essencial, porque não conhecemos o tamanho da linha de base
- Estimamos T e R a partir da matriz essencial
- Reconstruir a profundidade dos pontos (testar as 4 possibilidades de sinal)
- O resultado obtido é dependente de um fator de escala

Reconstrução – Caso não calibrado

Segundo caso: não conhecemos os parâmetros intrínsecos.

- Calibramos com o algoritmo dos 8 pontos
- Estimamos as posições dos epipólos a partir da matriz fundamental
- Definimos uma base projetiva padrão (5 pontos) considerando os 5 primeiros pontos a serem reconstruídos (e que não formam uma configuração degenerada) como pontos da base
- Encontrar as transformações projetivas planares T_E e T_D que levam 4 pontos da imagem esquerda e 4 pontos da direita na base projetiva padrão em 2D. Aplicar as transformações a todos pontos da imagem e aos epipólos.
- Estimar as coordenadas dos centros de projeção na base projetiva 3D definida e as matrizes de projeção.
- Efetuar a triangulação, o resultado é dependente do conhecimento da posição de 5 pontos no espaço 3D.

Reconstrução Densa (Horn)

Reconstrução a partir de pares de pontos pode requerer uma interpolação de pontos intermediários para se obter uma imagem completa da disparidade. Vejamos um método que computa diretamente o mapa de disparidade baseado nas diferenças de intensidade pela imagem.

Considerando uma linha do sistema estéreo com coordenadas x_E e x_D correspondentes.

$$x_E = f \frac{x + b/2}{z} \quad \text{e} \quad x_D = f \frac{x - b/2}{z}$$

Queremos encontrar $z(x)$ tal que

$$E_E \left(f \frac{x + b/2}{z} \right) = E_D \left(f \frac{x - b/2}{z} \right) \quad (\text{intensidades dos pixels})$$

Substituindo $x' = fx/z$ e $d(x') = bf/z$ (mapa de disparidade)

$$E_E\left(x + \frac{1}{2}d(x)\right) = E_D\left(x - \frac{1}{2}d(x)\right)$$

Restrição de suavidade do mapa de disparidade (a ser minimizada)

$$e_s = \iint (\nabla^2 d(x, y))^2 dx dy$$

Restrição a minimizar para igualar as intensidades das imagens

$$e_i = \iint (E_E(x, y) - E_D(x, y))^2 dx dy$$

Função objetivo total a ser minimizada

$$F = e_s + \lambda e_i, \lambda \text{ é o parâmetro de regularização de Tikonov}$$

Solução da minimização é solução da equação de Euler associada:

$$F_d - \frac{\partial}{\partial x} F_{d_x} - \frac{\partial}{\partial y} F_{d_y} = 0$$

Expandindo:

$$\nabla^4 d = \lambda(E_E - E_D) \frac{1}{2} \left(\frac{\partial E_E}{\partial x} + \frac{\partial E_D}{\partial x} \right)$$

O operador bi-harmônico é definido por:

$$\nabla^4 = \frac{\partial^4}{\partial x^4} + 2 \frac{\partial^4}{\partial x^2 y^2} + \frac{\partial^4}{\partial y^4}$$

Substituindo por diferenças finitas, construímos o esquema iterativo:

$$d_{ij}^{m+1} = \bar{d}_{ij}^m - \frac{\lambda}{K} (E_E - E_D) \frac{1}{2} \left(\frac{\partial E_E}{\partial x} + \frac{\partial E_D}{\partial x} \right)$$

As derivadas de E e de d são obtidas por convolução com um padrão apropriado.