CC222 – Visão Computacional

Visão Estéreo

Instituto Tecnológico de Aeronáutica

Prof. Carlos Henrique Q. Forster – Sala 121 IEC

forster@ita.br

ramal 5981

Tópicos da aula

- Caso simples de visão estéreo (triangulação)
- Problema de matching
- Geometria epipolar
- Matrizes essencial e fundamental
- Retificação
- Calibração do sistema estéreo
- Reconstrução por triangulação
- Reconstrução dependente de um fator de escala
- Reconstrução dependente de uma base projetiva

Livro para acompanhar essa aula

Trucco e Verri –cap 7.

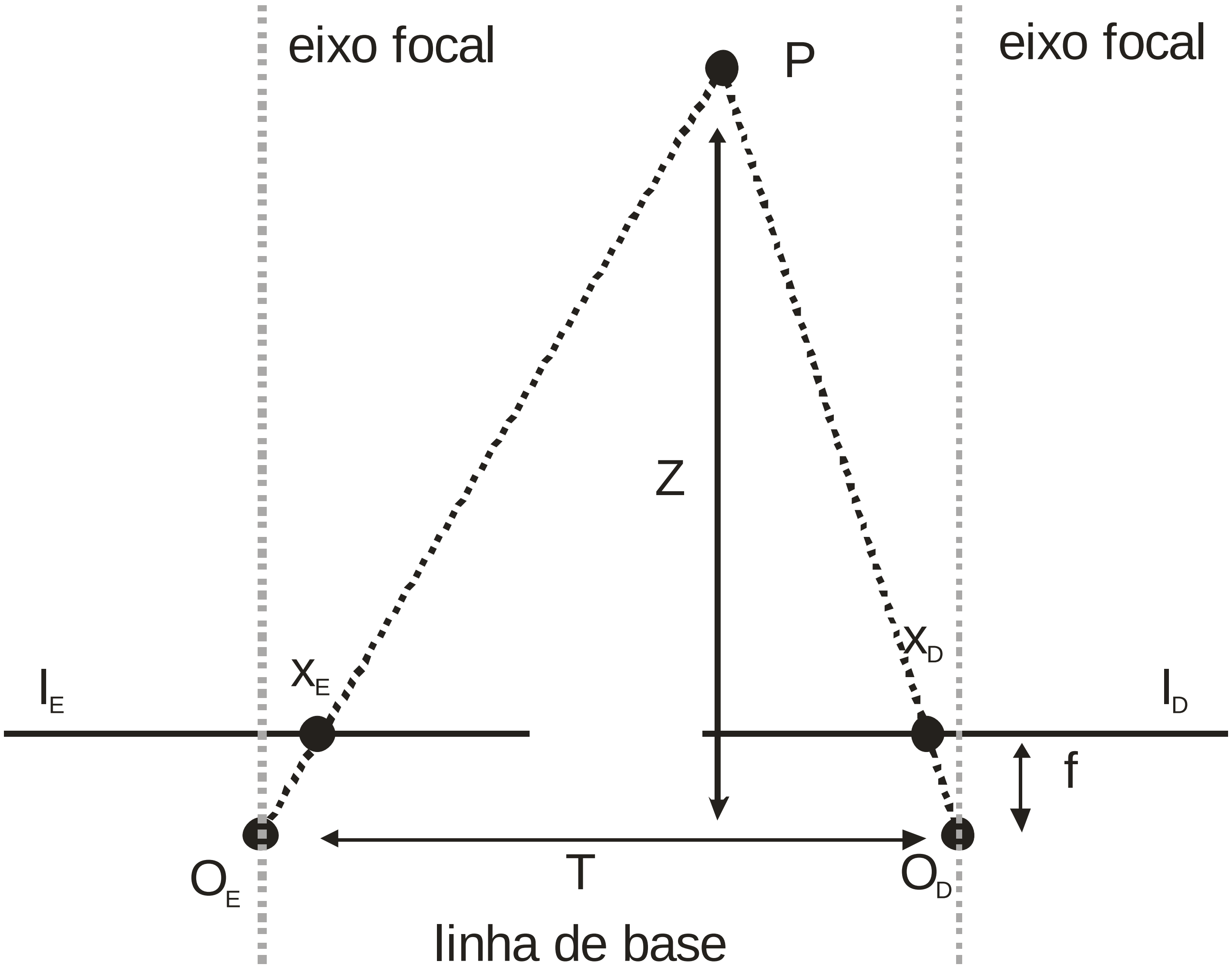
Caso simples de visão estéreo

Duas câmeras colocadas lado a lado com eixos ópticos paralelos.

Os planos-imagem coincidem (mas possuem seus próprios sistemas de coordenadas).

Os eixos horizontais coincidem e a escala coincide.

As distâncias focais são iguais.


eixo focal
P
eixo focal
Z
x_D
x_E
I_E
I_D
f
O_E
T
O_D
linha de base

$$\frac{T}{Z} = \frac{T + x_E - x_D}{Z - f}$$

Resolvendo para Z,

$$Z = f \frac{T}{x_D - x_E}$$

Exemplo, par de imagens estéreo neste caso

Os pontos correspondentes estão sobre as linhas horizontais de mesma coordenada y.

Exemplo: uma linha correspondente das duas imagens

Problema: determinar pontos correspondentes de uma imagem na outra.

A diferença de coordenadas (disparidade) corresponde ao inverso da profundidade.

Um método para determinar a correspondência: utilizar correlação cruzada de partes de uma imagem sobre a outra.

Exemplo de solução

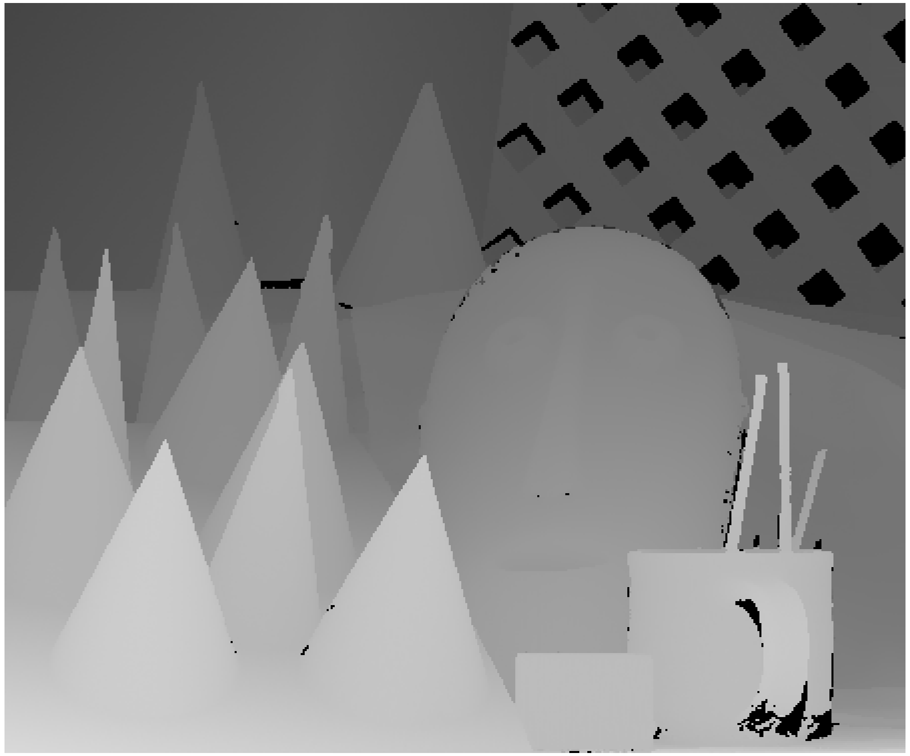
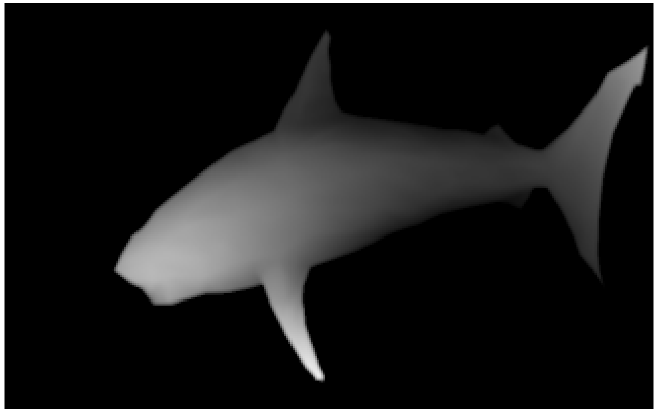


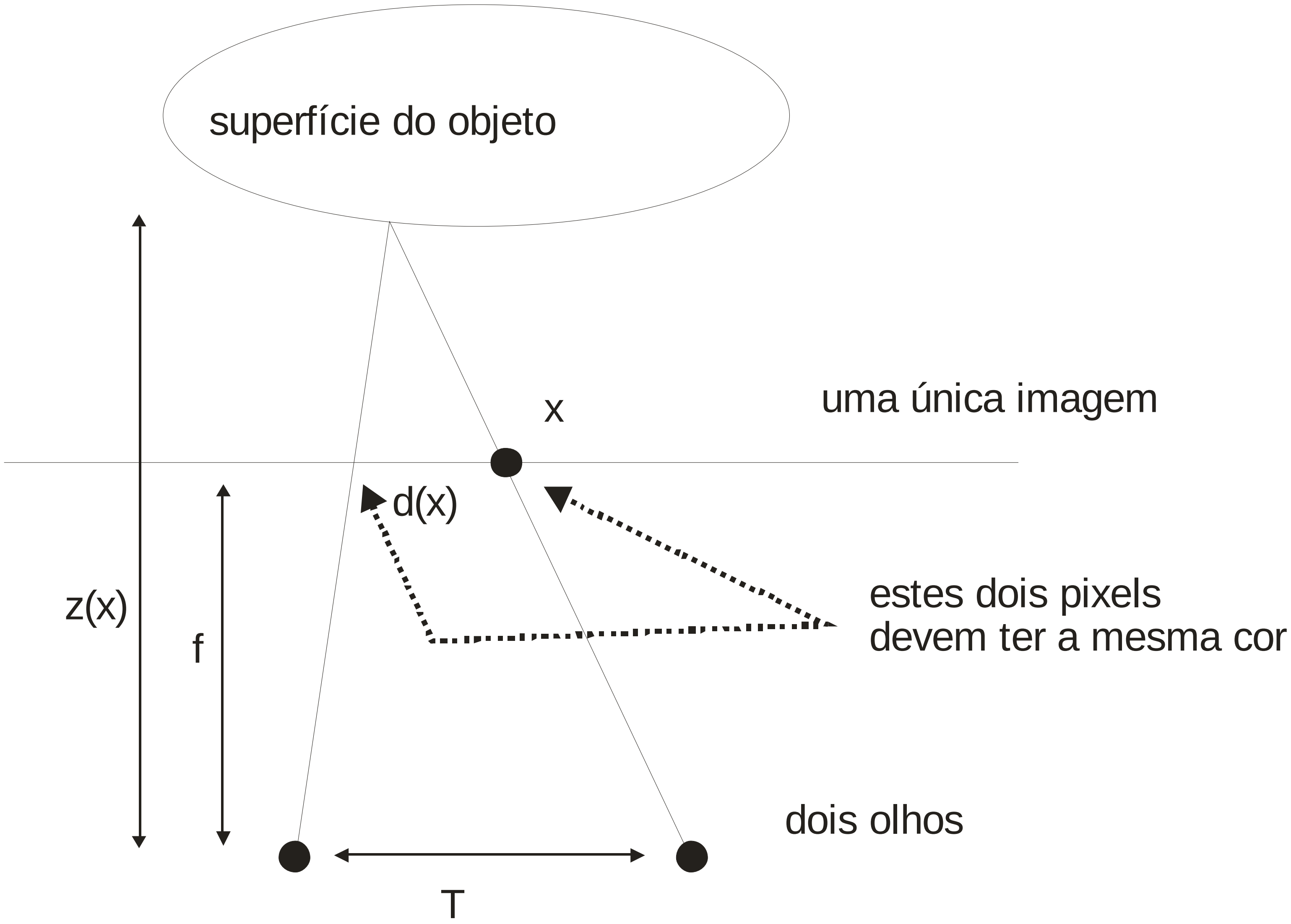
Imagem de profundidade z(x,y)

Exercício

Modelar a construção de um autoestereograma


(imagem de profundidade)


(auto estereograma)

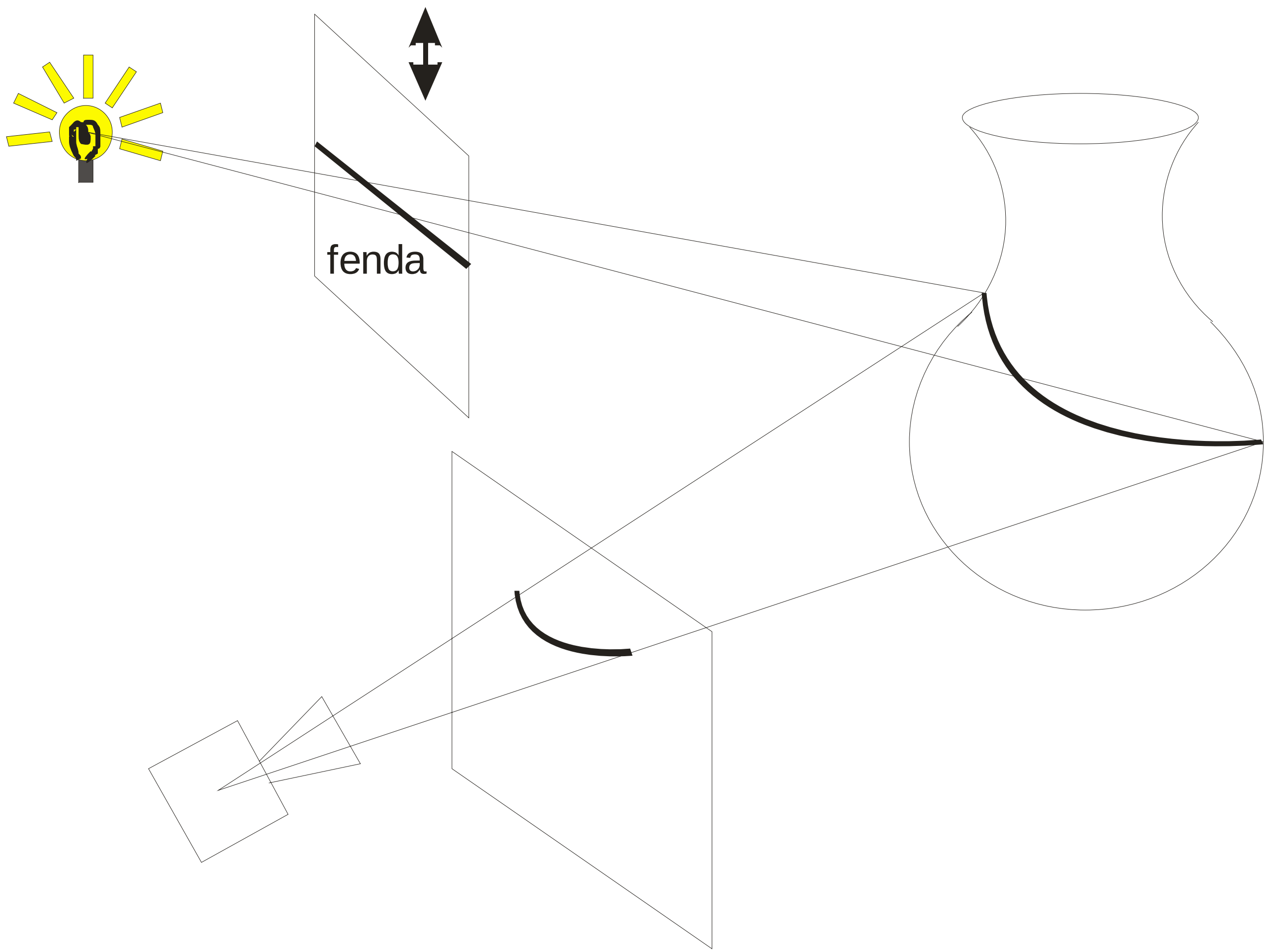

superfície do objeto
x
uma única imagem
d(x)
estes dois pixels
devem ter a mesma cor
z(x)
f
dois olhos
T

$$\frac{d(x)}{T} = \frac{z(x) - f}{z(x)} = 1 - \frac{f}{z(x)}$$

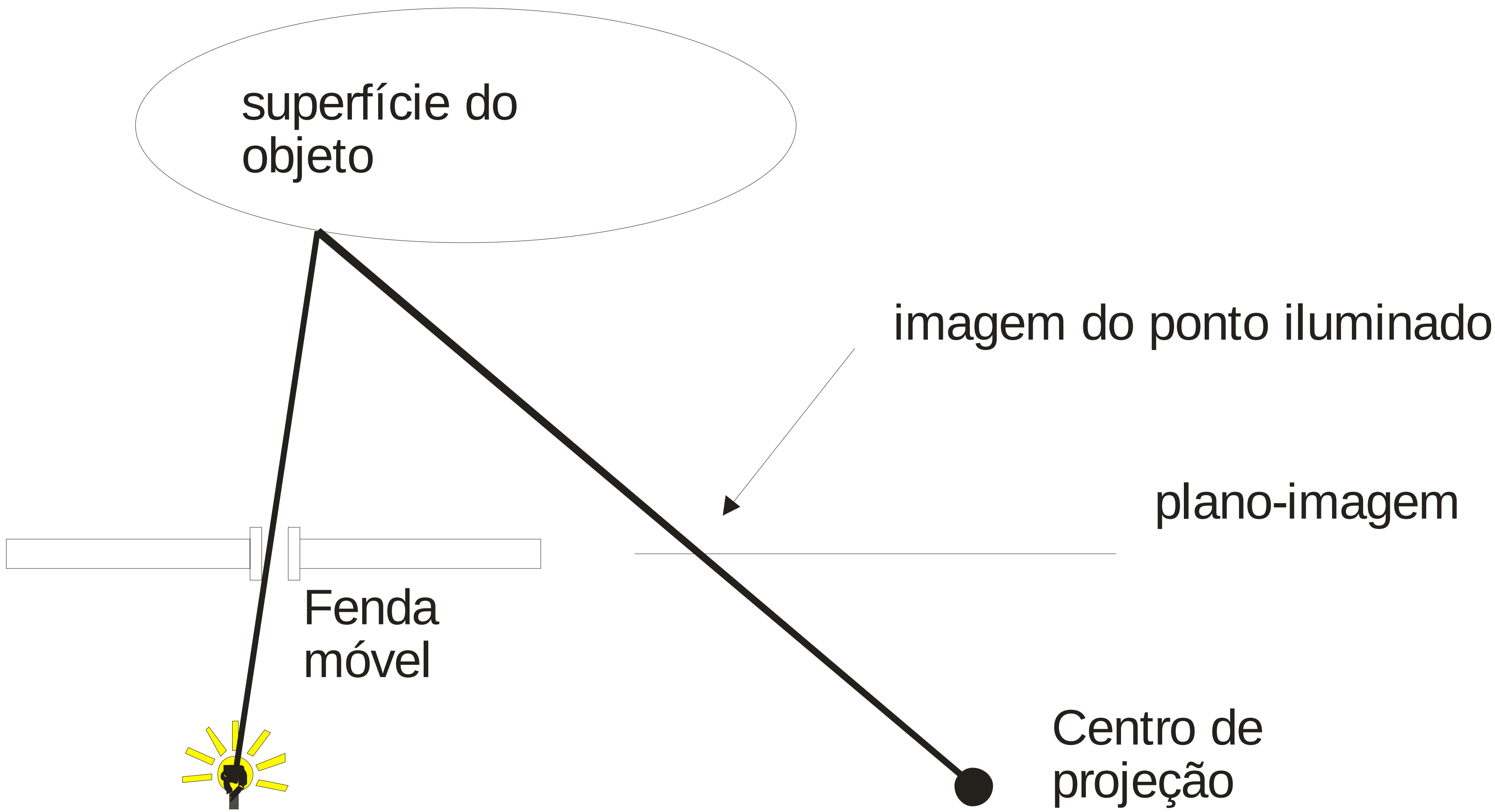
$$d(x) = T - \frac{Tf}{z(x)}$$

$$E(x) = E(x - d(x))$$

Exercício – Light Stripe



Modelar algoritmo de reconstrução 3d


superfície do
objeto
imagem do ponto iluminado
plano-imagem
Fenda
móvel
Centro de
projeção

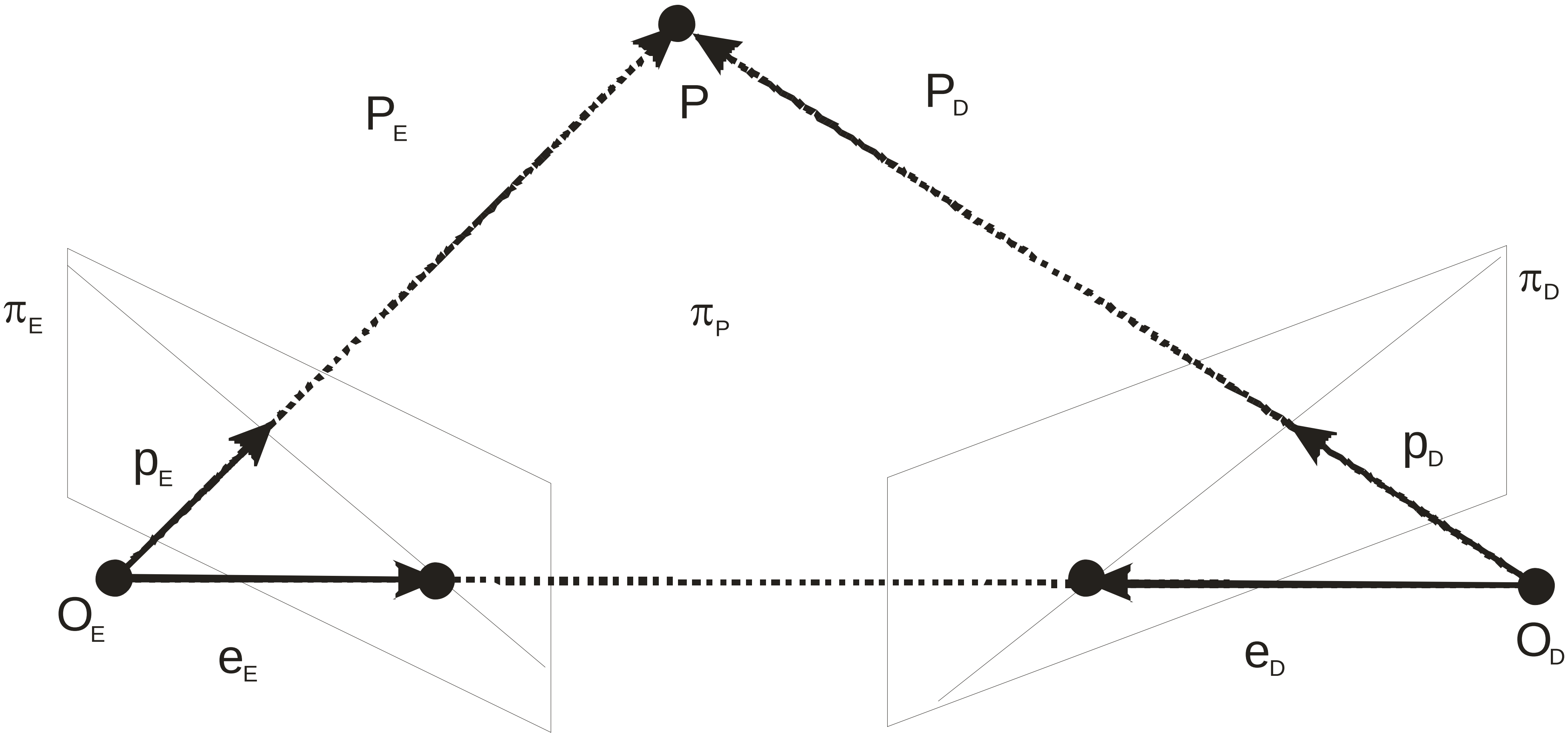
Geometria Epipolar

Como proceder se as câmeras não forem paralelas ou então se tratar de uma só câmera em movimento?

Estudar a geometria de um par de câmeras.

Propriedade da geometria epipolar: Existe um ponto em cada imagem chamado epipólo que consiste da projeção do centro de projeção da outra câmera.

Dado um ponto na imagem E, o ponto correspondente na imagem D está sobre uma determinada reta do plano-imagem D que passa pelo epipólo.



Centro de projeção, epipólo, reta epipolar, plano epipolar, ponto-objeto, ponto-imagem

Orientação relativa entre câmeras é dada por uma matriz de parâmetros extrínsecos (movimento rígido: translação e rotação apenas).

Translação: $T = O_D - O_E$

Rotação: $P_D = \mathbf{R}(P_E - T)$

Formação das imagens

$$p_E = \frac{f_E}{z_E} P_E \quad \text{e} \quad p_D = \frac{f_D}{z_D} P_D$$

Matriz essencial

Coplanaridade de $P_E, T, P_E - T$

Volume do prisma = 0

$\left(P_E - T\right)^T T \times P_E = 0$, substituindo P_D

$\left(R^T P_D\right)^T T \times P_E = 0$, escrevendo produto vetorial na forma matricial

$$T \times P_E = SP_E, onde\ S = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$ (rank 2)

$P_D{}^T \cdot R \cdot S \cdot P_E = 0$, substituindo $E = RS$ matriz essencial.

$P_D{}^T \cdot E \cdot P_E = 0$

$$P_D{}^T \cdot E \cdot P_E = 0$$

$$p_D{}^T \cdot E \cdot p_E = 0$$

Esta última configura uma equação da reta em função dos pontos de π_D (veja que o plano epipolar passa pela origem que é o centro de projeção da câmera)

Os coeficientes da reta epipolar gerada por p_E são dados por $u_D = E \cdot p_E$

Assim $p_D{}^T \cdot u_D = 0$

Matriz Fundamental

Sejam as matrizes de parâmetros intrínsecos M_E e M_D.

$$p_E = M_E^{-1}\tilde{p}_E$$

$$p_D = M_D^{-1}\tilde{p}_D$$

Seja a matriz fundamental F dada por

$$F = M_D^{-T} \cdot E \cdot M_E^{-1}$$

Temos a relação sobre as coordenadas homogêneas de pontos das imagens (Longuett-Higgins)

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

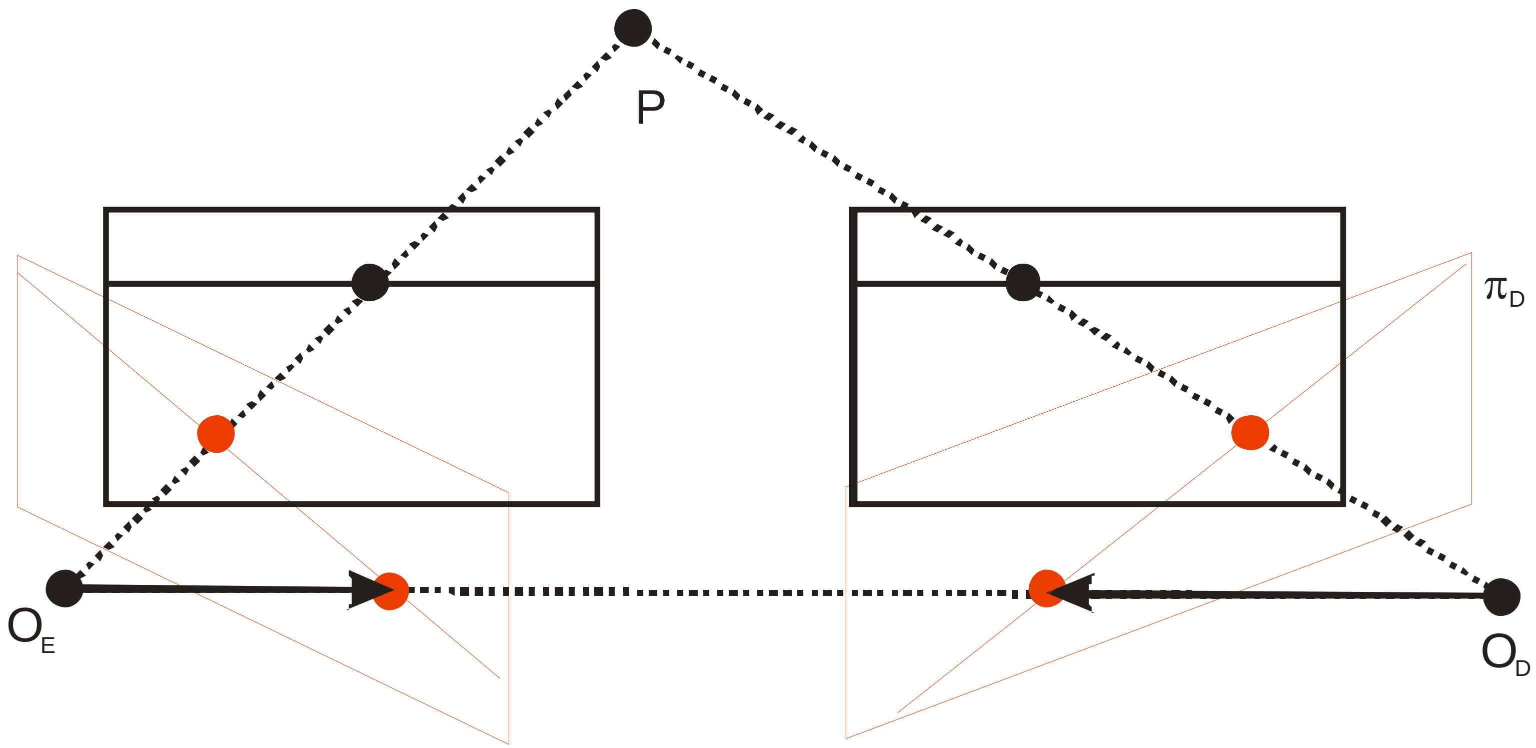
A reta epipolar sobre o plano imagem π_D para o ponto $\tilde{p}_E$ é dada pelos coeficientes:

$$\tilde{u}_D = F \cdot \tilde{p}_E$$

Assim, pode-se, por exemplo, saber em que reta sobre a imagem da direita deve-se procurar o ponto correspondente a um ponto dado da imagem da esquerda.

Retificação

É possível transformar um caso genérico de visão estéreo para o caso mais simples através de warping das imagens. É necessário encontrar as duas projetividades para as quais, aplicadas às imagens, as restrições do caso mais simples (retificado) são cumpridas. Esse processo é chamado retificação do par de imagens estéreo.



No caso retificado:

- Um par de retas epipolares conjugadas se torna colinear e paralelas a um dos eixos da imagem (eixo x).
- Retas epipolares conjugadas horizontais terão a mesma coordenada y nas imagens.
- Os epipólos estão no infinito (porque os planos-imagem são paralelos à ret que une os centros de projeção (linha de base)).

Assumindo:

- A origem do sistema de referência das imagens é o ponto principal.
- A distância focal é f.

Algoritmo:

- Rotacionar a câmera esquerda para que o epipólo vá ao infinito na direção horizontal (só depende da posição dos centros de projeção).
- Aplicar a mesma rotação à câmera direita para recobrar a geometria original.
- Rotacionar a câmera direita de R (parâmetro extrínseco).
- Ajustar a escala dos sistemas de coordenadas das câmeras.

Definir base de 3 vetores ortonormais (ortogonais entre si e unitários).

$$e_1 = \frac{T}{\|T\|}$$, onde T é o vetor do centro de projeção esquerdo para o direito.

$$e_2 = \frac{1}{\sqrt{T_x^{\,2} + T_y^{\,2}}} \begin{bmatrix} -T_y \\ T_x \\ 0 \end{bmatrix}$$, um vetor perpendicular a T.

$$e_3 = e_1 \times e_2$$, vetor perpendicular a T e a e2.

A rotação definida por $R_{rect} = \begin{bmatrix} e_1^{\,T} \\ e_2^{\,T} \\ e_3^{\,T} \end{bmatrix}$ rotaciona a câmera esquerda levando o epipólo para o infinito no eixo horizontal. (verificar a multiplicação por T).

Algoritmo (Trucco-Fusiello):

1. Construir a matriz R_{rect}.
2. Fazer $R_E = R_{rect}$ e $R_D = R \cdot R_{rect}$
3. Para cada ponto para a câmera esquerda $p_E = [x, y, f]^T$
   a. $[x', y', z']^T = R_E \cdot p_E$
   b. $p_E' = \frac{f}{z'}[x', y', z']^T$ ponto retificado
4. Repetir para a câmera direita com R_D e p_D.


200
400
600
800
1000
1200
100
200
300
400
500
600
700
800
900