Reconstrução 3D

Thiago Souto Maior
mouse@cin.ufpe.br
Judith Kelner
jk@cin.ufpe.br

Grupo de Pesquisa em Realidade Virtual e Multimídia
Universidade Federal de Pernambuco, Centro de Informática

12/04/2010

Agenda

- Introdução
- Modelo de Câmera
- Geometria Projetiva
- Geometria Epipolar
- Pipeline de Reconstrução
- Conclusão

Introdução

Introdução

- Visão Computacional
  - Capturar informações de um objeto a fim de recuperar seu modelo 3D
  - Imagens foto-realístico
- Aplicações
  - Modelagem 3D

Objetivos

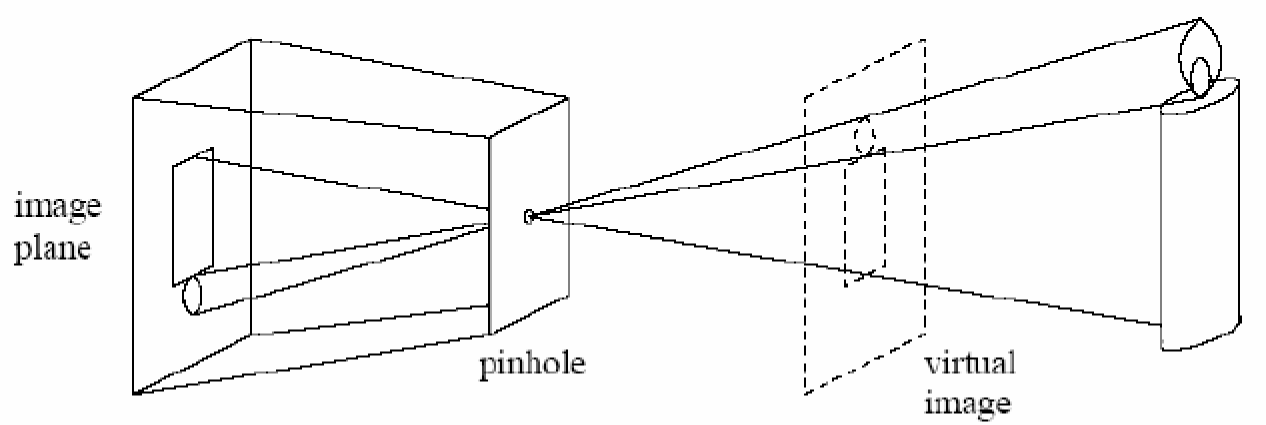
- Entender a relação entre várias imagens de um objeto
- Entender os princípios gerais para estimar os parâmetros necessários para a reconstrução de um objeto 3D
- Entender como tratar uma cena e as propriedades da câmera a partir de imagens do mundo real

Modelo de Câmera

Modelo de Câmera

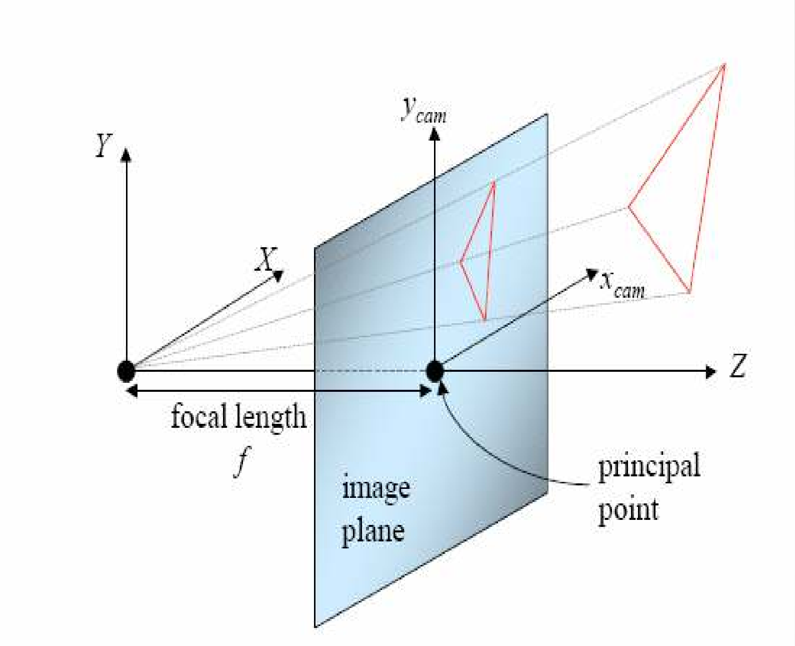
- Representação 2D de objetos 3D
- Definir a pose de um objeto de acordo com um frame de referência:
  - Posição
  - Orientação

Câmera Pin-hole



Modelo de Câmera

- Modelo de Câmera
  - Perspectiva de Projeção
  - Plano da Retina



Modelo de Câmera

- O ponto 3D (X, Y, Z) é projetado no plano de imagem como (x_{cam}, y_{cam})

$$\lambda \begin{bmatrix} x_{cam} \\ y_{cam} \\ f \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} onde \quad \lambda = \frac{z}{f}$$

Modelo de Câmera

- Pode-se utilizar as coordenadas homogêneas para representar a projeção dos pontos 3D no plano da imagem
- Utilizando as coordenadas homogêneas, a projeção é representada por uma transformação linear

$$\mathbf{x} = \mathbf{P}\,\mathbf{X}$$

Matriz de projeção da câmera

Projeção da Câmera

- Matriz de Calibração:
  - Distância focal
  - Posição da Câmera
  - Pixel (tamanho e formato)

- Movimento da Câmera

$$M' = \begin{bmatrix} R & \mathrm{t} \\ 0_3 & 1 \end{bmatrix} M$$

Projeção de Câmera

- P = K [R|t]
  - K -> matriz de calibração da câmera
  - R -> Rotação => matriz que descreve a orientação da câmera
  - t -> matriz que descreve a translação da origem da câmera em relação à um sistema de coordenada

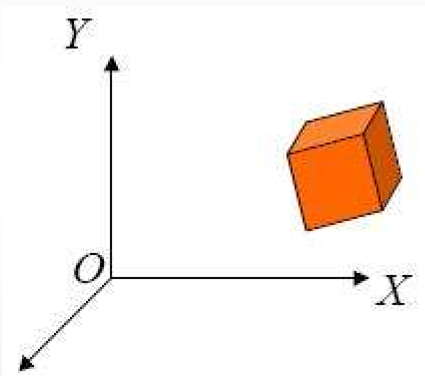
Geometria Projetiva

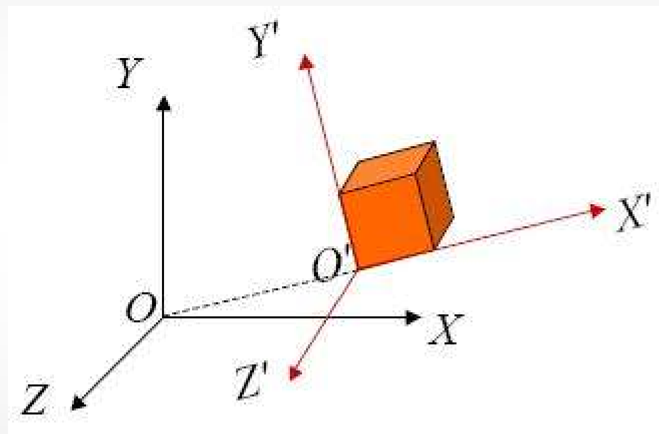
Coordenadas Homogêneas

- Principal sistema de coordenadas utilizadas em:
  - Visão Computacional
  - Computação Gráfica
  - Modelagem Geométrica
- Representar os objetos 3D e descrever a pose do mesmo:
  - Orientação e posição

Coordenadas Euclidianas

- Qual a orientação do cubo?





- Translação
- Rotação

Transformação entre Frames

- Representação da transformação entre dois frames usando coordenadas euclidianas:

$$p = t + Rp'$$

- Usando coordenadas homogêneas, a transformação é representada por uma simples multiplicação:

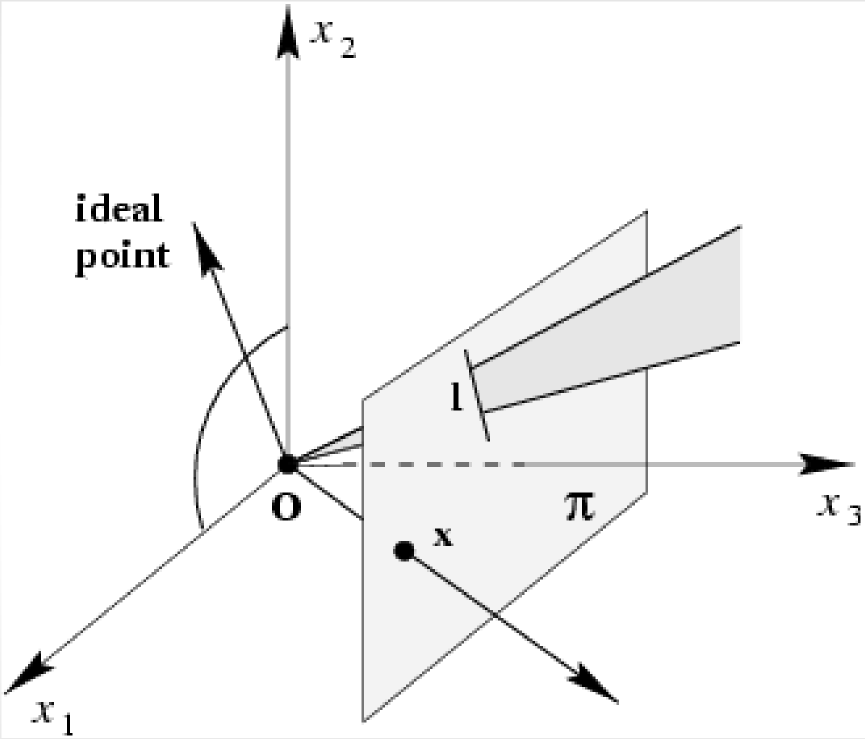
$$\mathbf{p = Tp'}$$

Coordenadas Homogêneas

- Para representar pontos e vetores em um espaço d dimensional, utiliza-se vetores de comprimento d+1
  - pontos tem a última coordenada = 1
  - vetores tem a última coordenada = 0

Pontos no infinito

- Considera que a imagem real encontra-se no infinito
- Representa os pontos no finito
- Definir a distância de um ponto



Coordenadas homogêneas

This allows us to write the transformation

$$\mathbf{p} = \mathbf{R}\mathbf{p}' + \mathbf{t}$$

$$\begin{bmatrix} p_X \\ p_Y \\ p_Z \end{bmatrix} = \mathbf{R} \begin{bmatrix} p_X' \\ p_Y' \\ p_Z' \end{bmatrix} + \mathbf{t}$$

Non-homogenous representation

as

$$\tilde{\mathbf{p}} = \mathbf{T}\tilde{\mathbf{p}}'$$

$$\tilde{\mathbf{p}} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \tilde{\mathbf{p}}'$$

$$\begin{bmatrix} p_X \\ p_Y \\ p_Z \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} p'_X \\ p'_Y \\ p'_Z \\ 1 \end{bmatrix}$$

Homogeneous representation

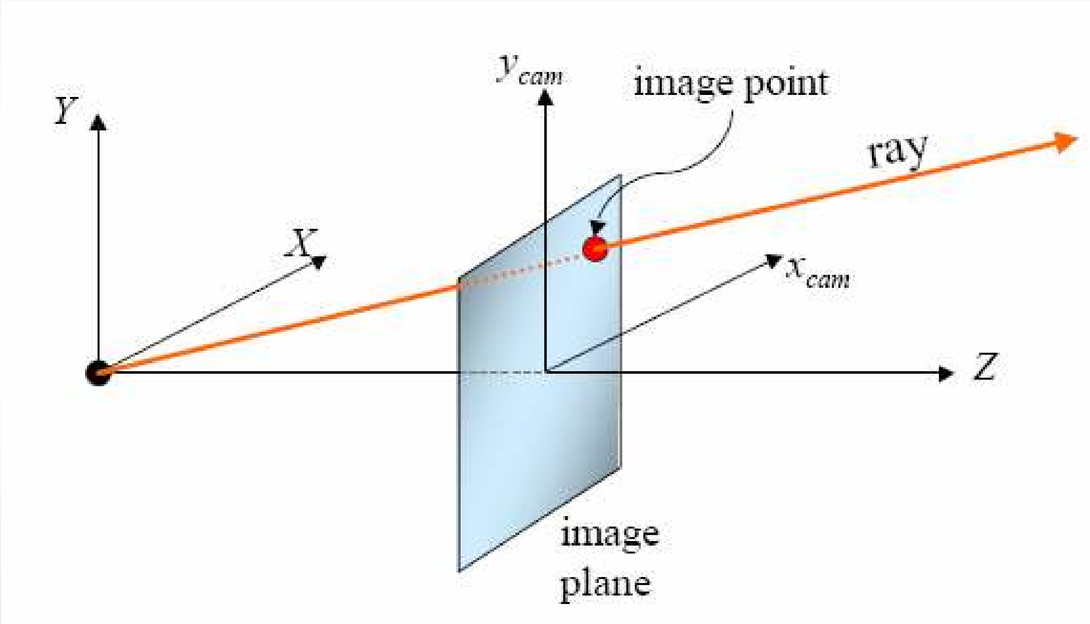
Geometria Epipolar

Geometria Epipolar

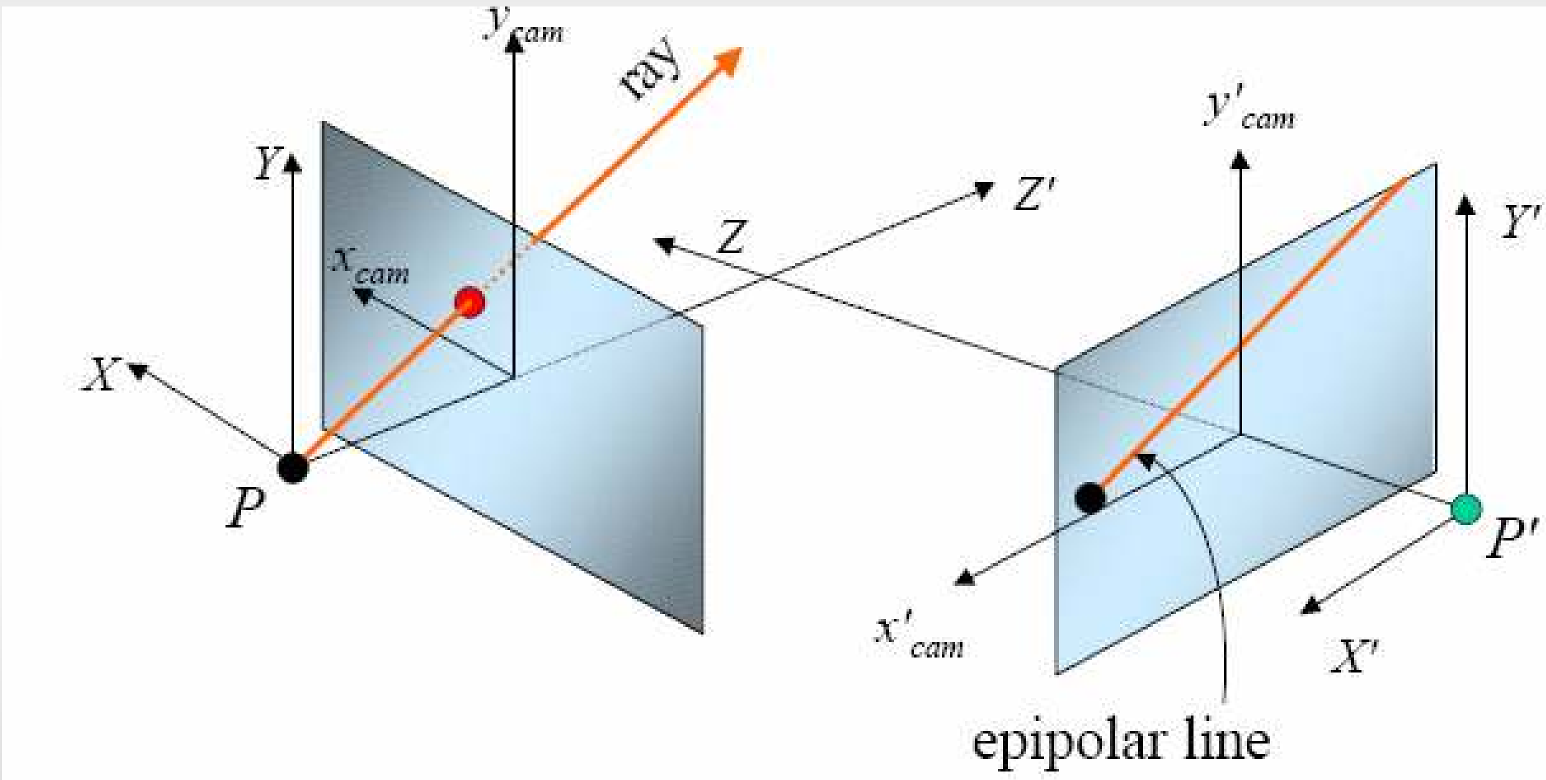
- Determinar a estrutura a partir de várias imagens

Geometria Epipolar

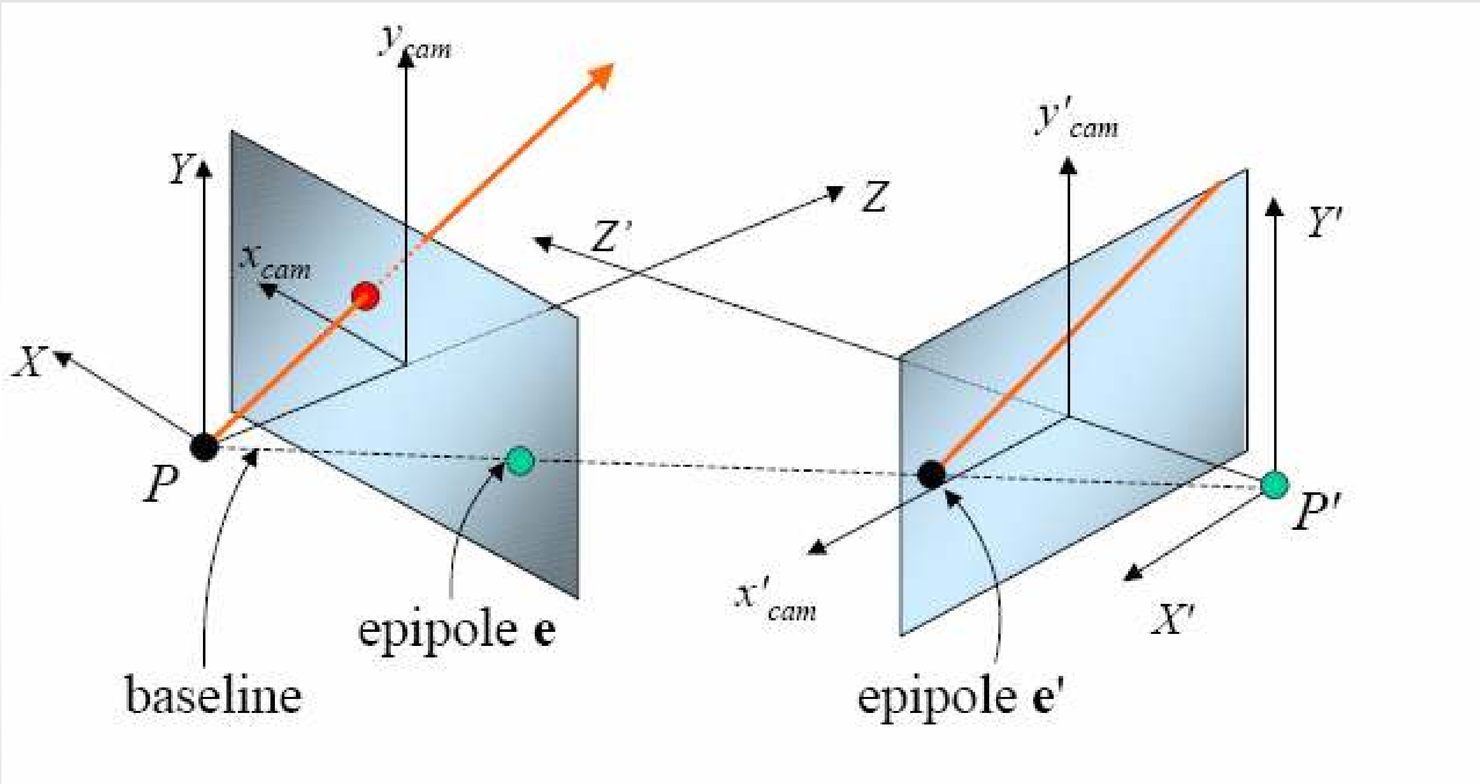
- Um ponto em uma imagem representa um raio no espaço 3D



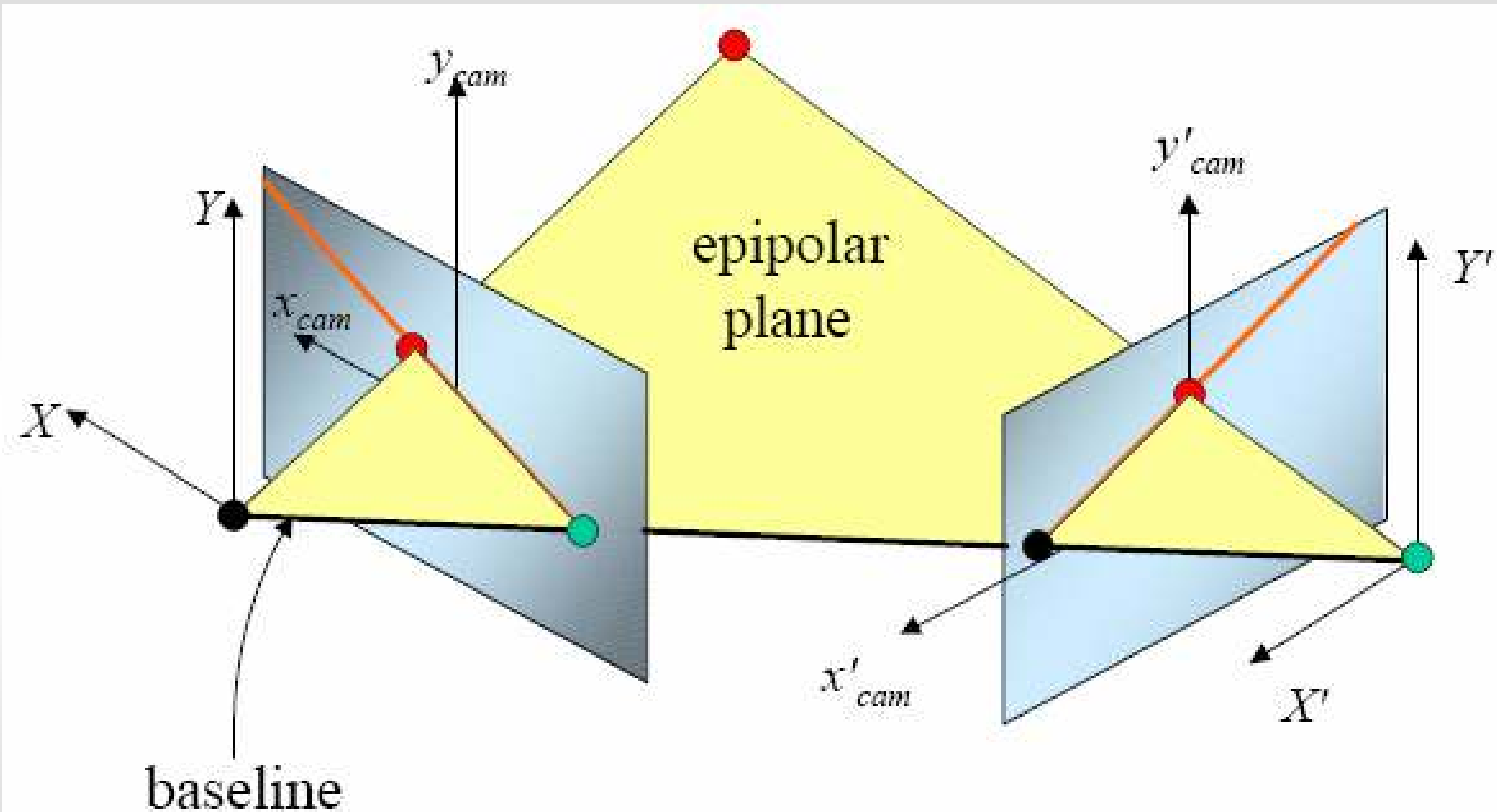
Geometria Epipolar



Geometria Epipolar

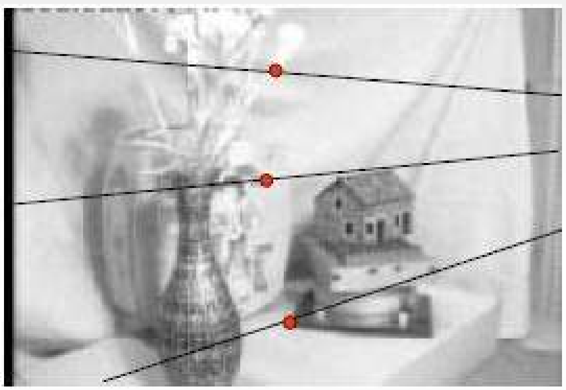


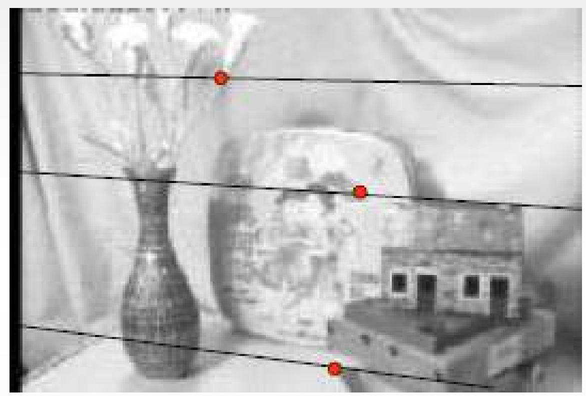
Geometria Epipolar



Geometria Epipolar

Geometria Epipolar





Geometria Epipolar

- Depende apenas dos parâmetros da câmera (intrínsecos e extrínsecos)
- Não depende da estrutura da cena

Matriz fundamental

- É a representação algébrica da geometria epipolar
- Para cada par de pontos correspondentes $(x \longleftrightarrow x')$

$$x'^{T} F x = 0$$

- Onde F é uma matriz 3x3

Matriz Essencial

- É uma especialização da matriz fundamental
  - Coordenadas normalizadas
  - Câmeras com a mesma calibração (ou a utilização de uma mesma câmera)

$$x'^{T} E x = 0$$

- Relação entre a matriz essencial e a fundamental

$$E = F'^{T} E F$$

Pipeline de Reconstrução

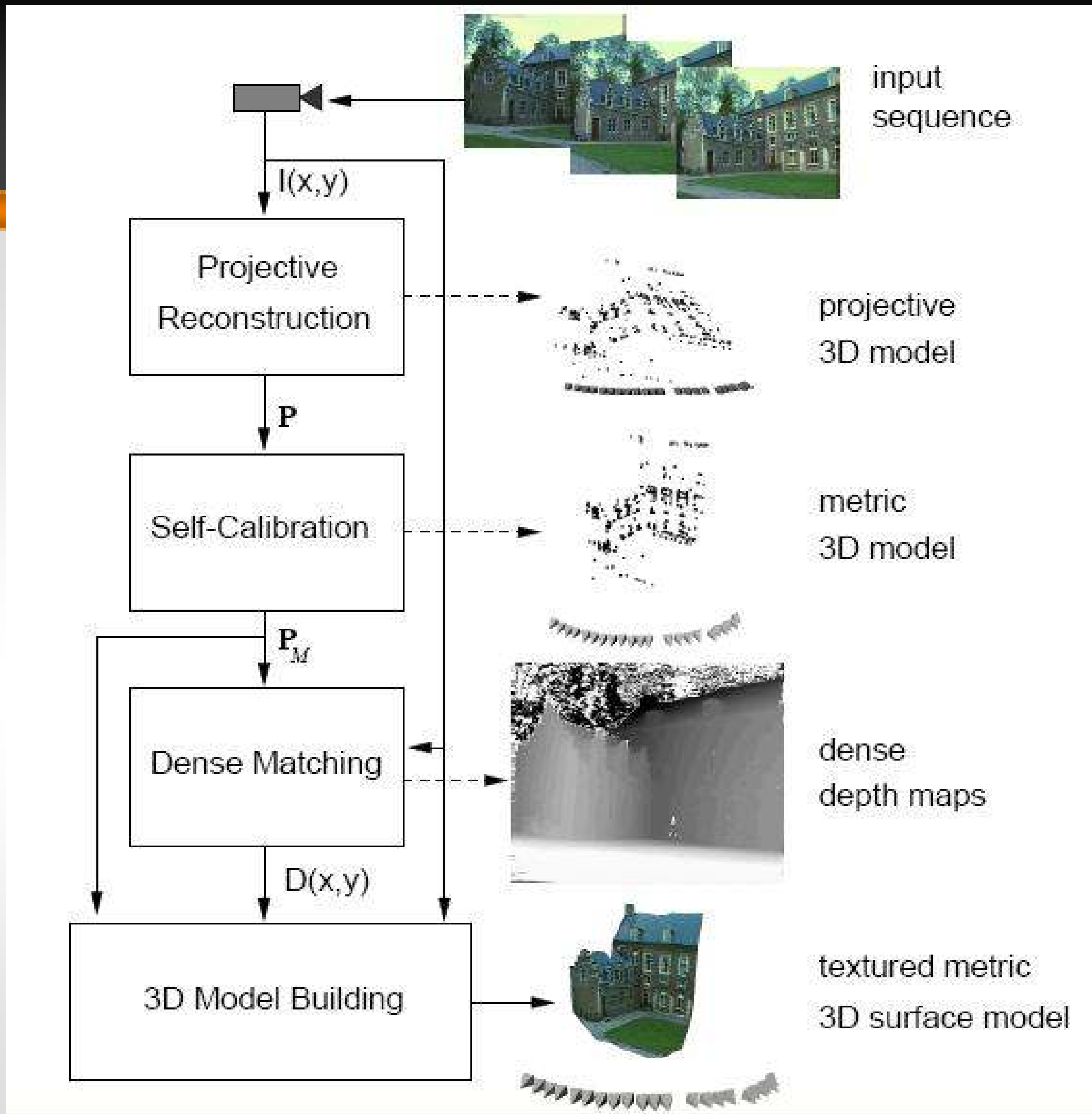
Por Pollefeys

Pipeline de Reconstrução

- Recuperar uma superfície 3D a partir de uma seqüência de imagens
- Imagens adquiridas pela livre movimentação da câmera sobre o objeto a ser tratado

Pipeline de Reconstrução

- Movimento da câmera e a configuração da mesma não precisam ser conhecidos
  - Obtidos gradualmente no decorrer do processo de reconstrução
- Versão reduzida do modelo
- Textura obtida a partir das imagens


input
sequence
I(x,y)
Projective
Reconstruction
projective
3D model
P
Self-Calibration
metric
3D model
P_M
Dense Matching
dense
depth maps
D(x,y)
3D Model Building
textured metric
3D surface model

Elementos Básicos da Reconstrução

- Rastreamento 2D
- Geração de keyframes
- Struct from Motion
- Mapeamento 2D-3D
- Refinamento (Bundle Adjustment)
- Reconstrução densa
- Geração de malha e Texturização

Rastreamento 2D

- Técnicas de rastreamento
  - Fluxo óptico (KLT)
  - Análise de similaridade
  - SIFT (Scale Invariant Feature Transform)

Geração de Keyframes

- A geometria epipolar é medida entre frames que devem ter uma distância mínima (baseline) e uma rotação mínima (aproximadamente 15º)
- A primeira estimativa de poses e pontos 3D só pode ser realizada após a seleção dos 2 primeiros keyframes

Heurísticas para Geração de Keyframes

- Inspeção visual
- Decaimento de features durante o rastreamento
- Análise de borramento (baixas freqüências)

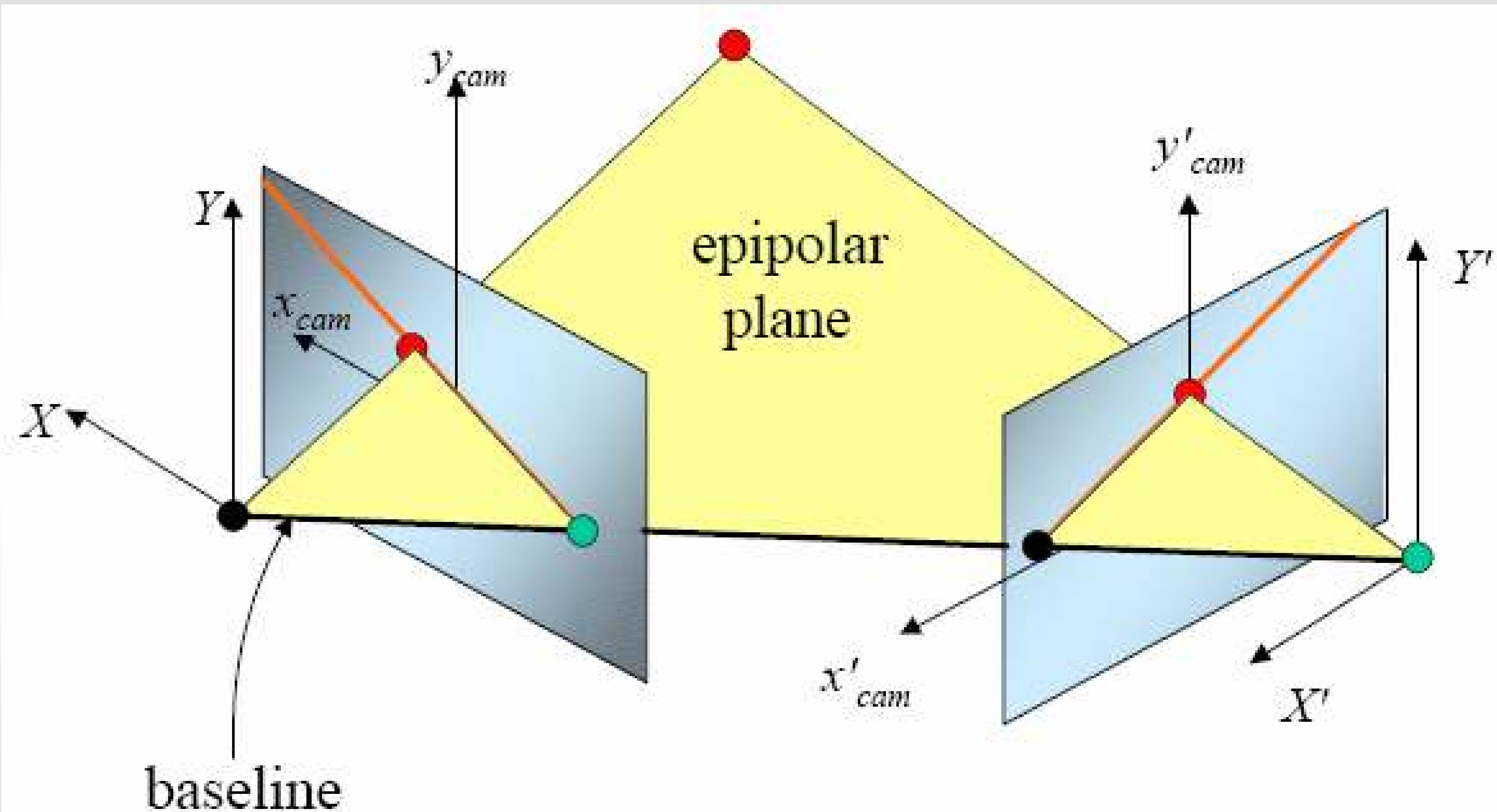
Struct from Motion

- É a técnica na qual é calculada a matriz fundamental ou essencial de um par de imagens.
- Após a obtenção da matriz, são extraídas as poses das câmeras e os pontos 3D são triangulados a partir do rastreamento dos pontos 2D
- Podem ser utilizadas 5, 7, 8 ou n > 8 correspondências

Struct from Motion

- Falsos casamentos
  - Inliers e outliers
- Cálculo de hipótese robusto
  - RANSAC
  - Contagem de inliers
  - Cálculo do erro de reprojeção

Triangulação



Mapeamento 2D-3D

- As novas poses são adquiridas através da correlação entre o rastreamento contínuo das features (tracks) e os pontos 3D já reconstruídos por elas

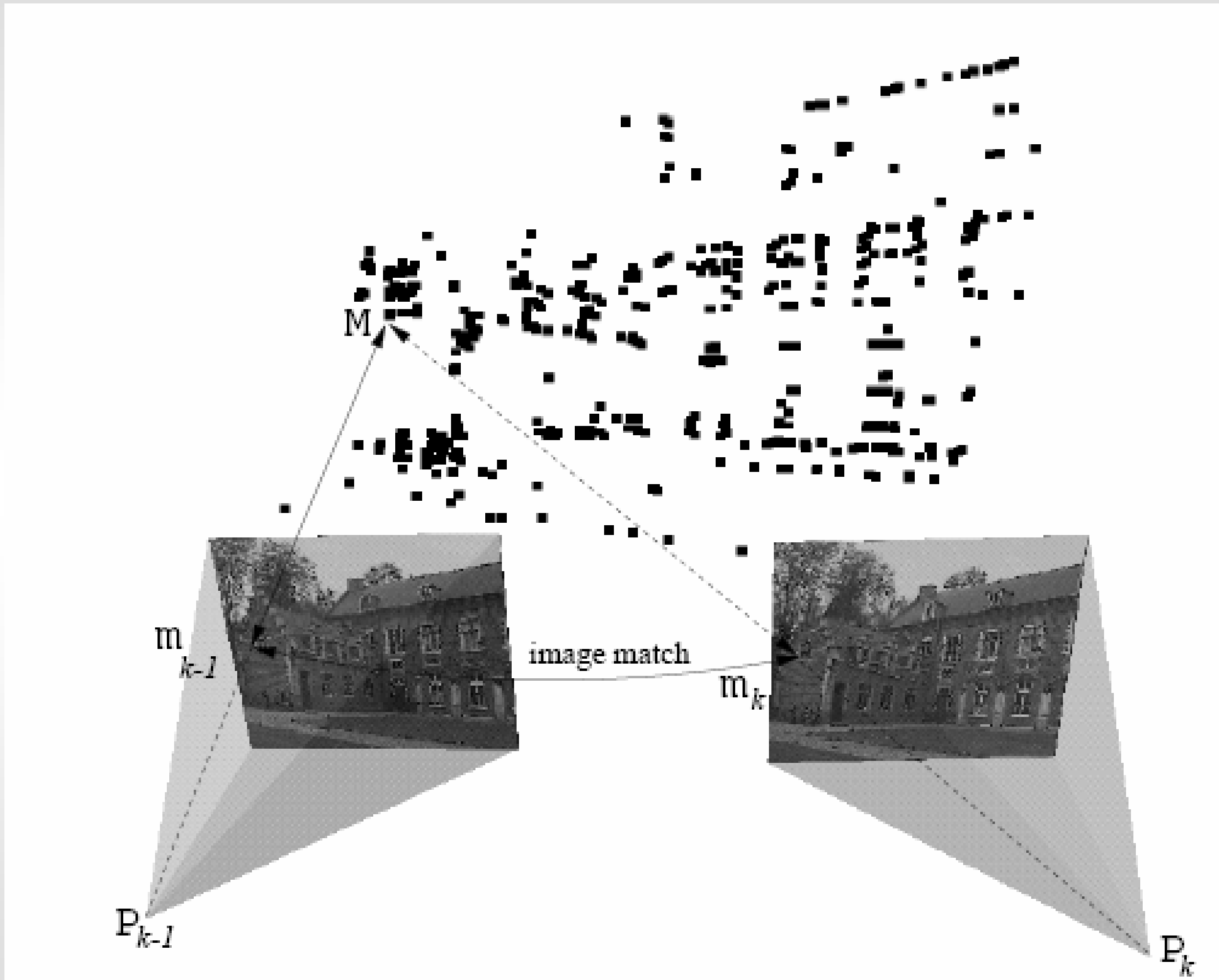
Mapeamento 2D-3D

- Uma simples resolução de sistema linear nos dá como saída a pose de um determinado quadro com apenas 6 correspondências

Refinamento

- Conhecido como Bundle Adjustment
- Minimização não linear do erro de reprojeção.
- Ajusta tanto as poses quanto os pontos 3D
- Extremamente necessário no pipeline de recontrução
- Implementado pelo SBA (sparse bundle adjustment)

Resultado reconstrução esparsa



Poses e pontos: Fim?

- Podemos ficar satisfeitos se a nossa finalidade for descobrir uma malha esparsa de pontos capaz de nos dar noção de planos e algumas estruturas, além de termos as poses de todas as câmeras.
- Aplicação: Realidade aumentada
- E se eu quiser um modelo mais detalhado?

Reconstrução Densa

- Técnicas baseadas em discrepância fotomérica
  - PMVS2
- Técnicas baseadas em retificação e estereoscopia
  - Pollefeys

Conclusão

- Visão geral dos conceitos chaves para a reconstrução 3D de uma cena utilizando uma seqüência de imagens
- Abordagem utilizada no GRVM

Referências

Referências

- R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. New York, NY, USA: Cambridge University Press, 2003
- M. Pollefeys, “Self-calibration and metric 3d reconstruction from uncalibrated image sequences,” Ph.D. dissertation, ESAT-PSI, K.U.Leuven,1999.