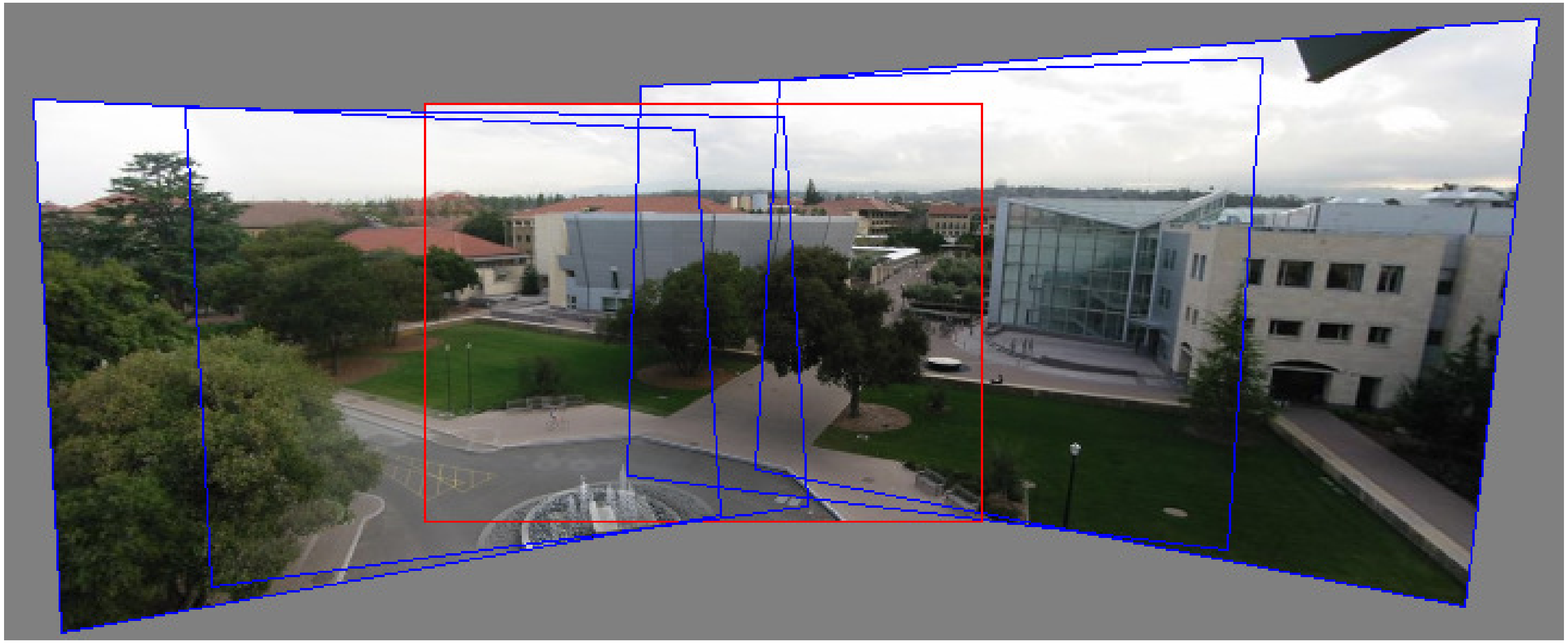
PMR2560 – Visão Computacional
Alinhamento de imagens

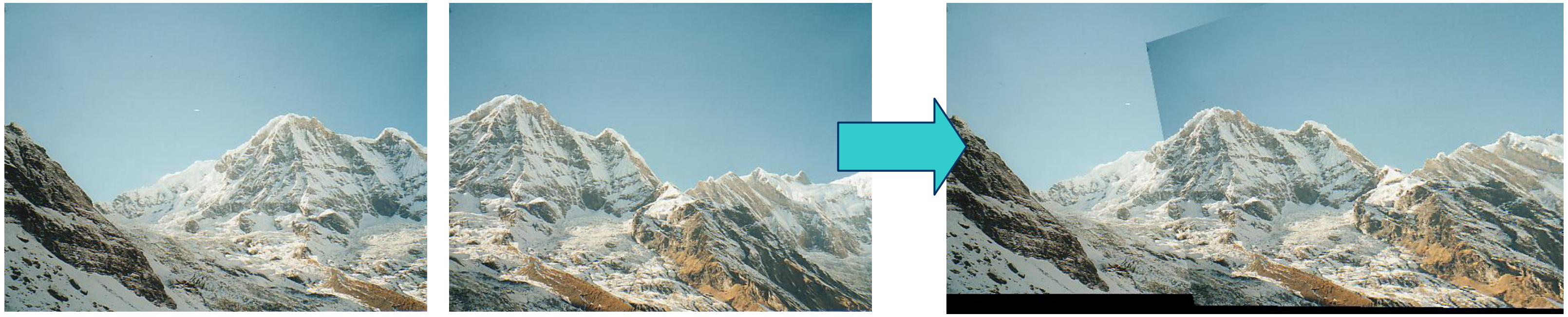
Prof. Eduardo L. L. Cabral

Objetivos

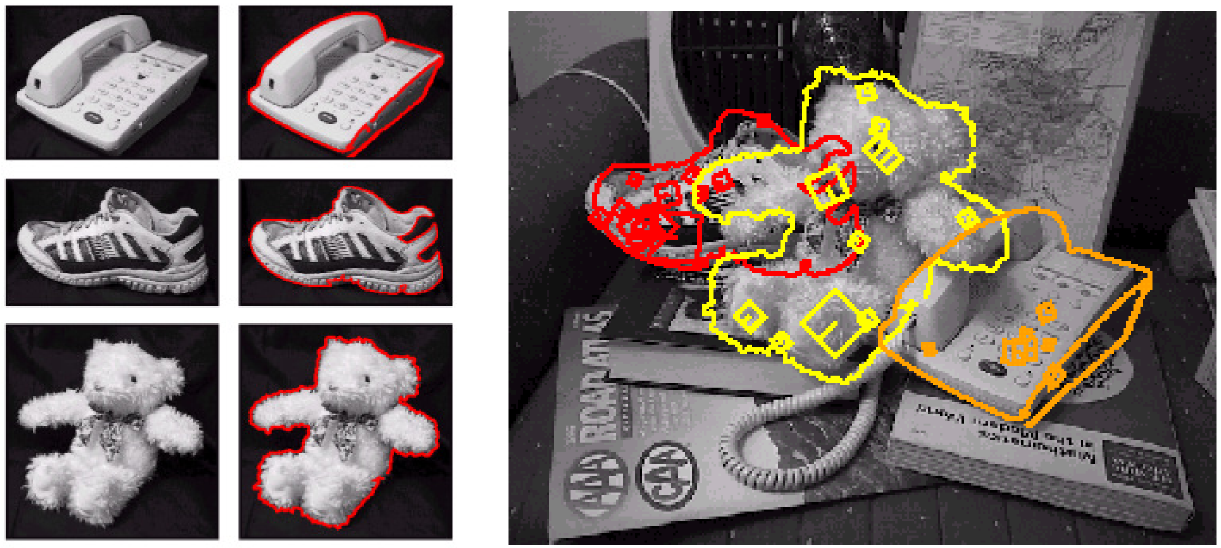
- Alinhamento de imagens
- Mosaico $\Rightarrow$ imagem panorâmica



Motivação

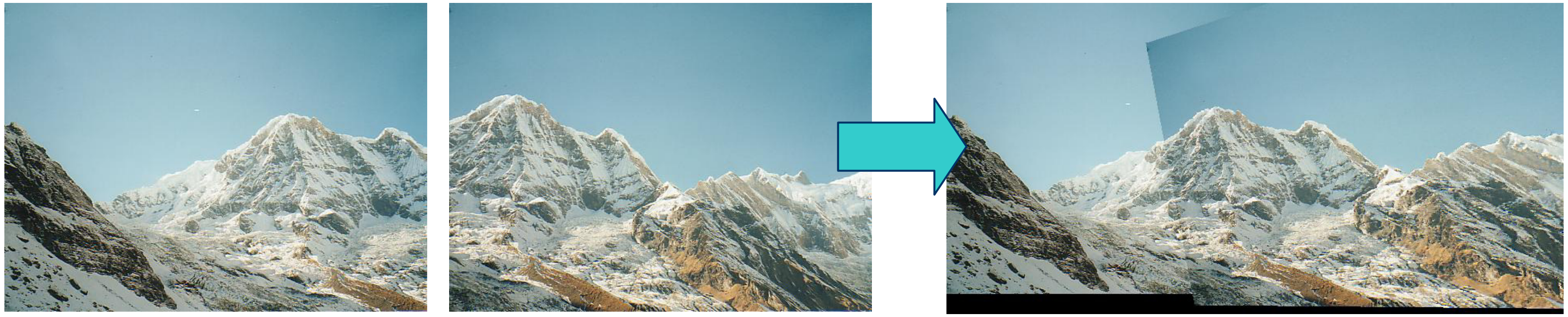


Mosaico – montagem de imagens panorâmicas

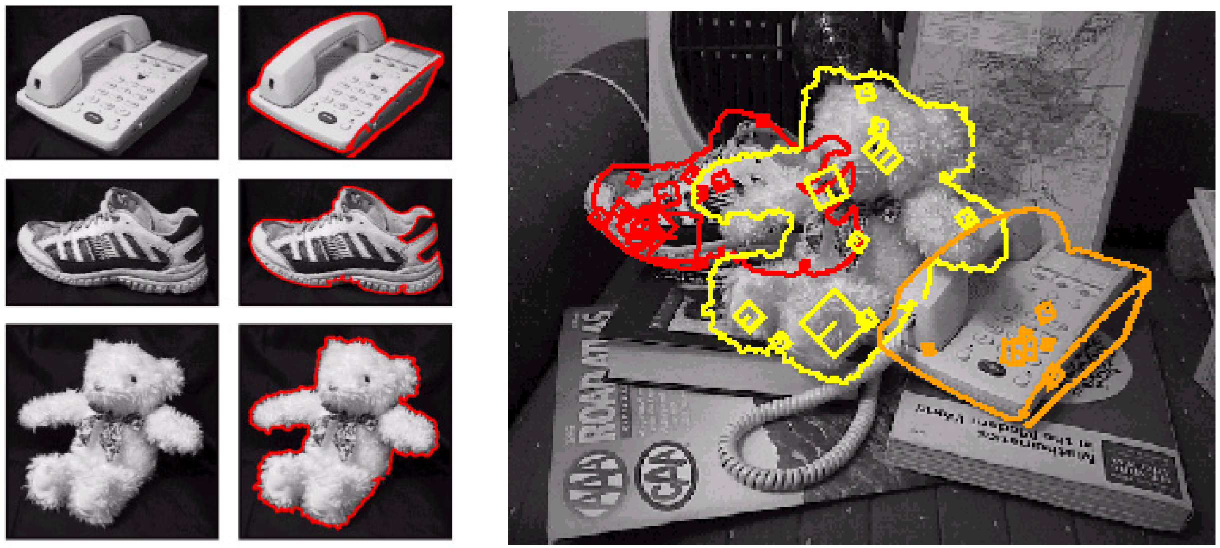


Reconhecimento de objetos

Desafios



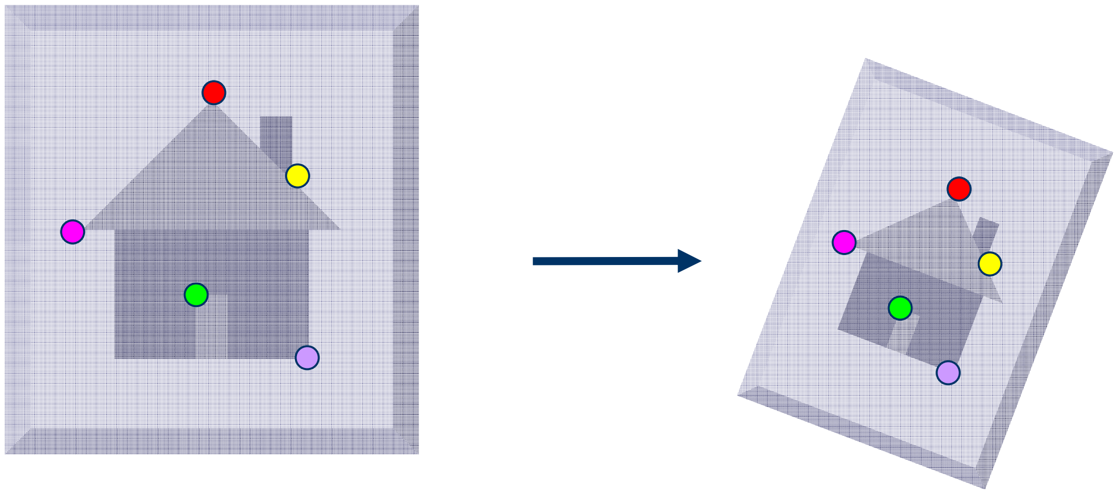
Pequeno grau de sobreposição



Oclusão, deformação, rotação

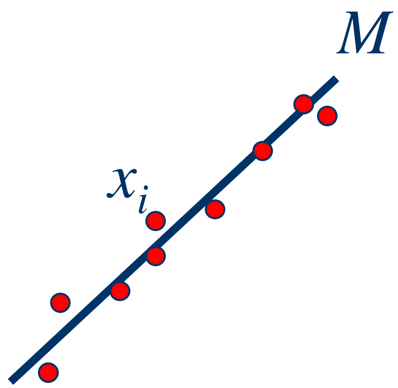
Duas abordagens

- Alinhamento direto (“pixel-based”) ⇒ busca pelo alinhamento onde quase todos os pixels se casam
- Alinhamento baseado em características ⇒ busca pelo alinhamento onde características se casam:
  - Pode ser verificado a posteriori usando alinhamento direto



Alinhamento como ajuste

- Alinhamento de duas imagens $\Rightarrow$ processo de obtenção de um modelo de transformação
- Objetivo do alinhamento $\Rightarrow$ ajustar um modelo para algumas caracterísitcas de uma imagem

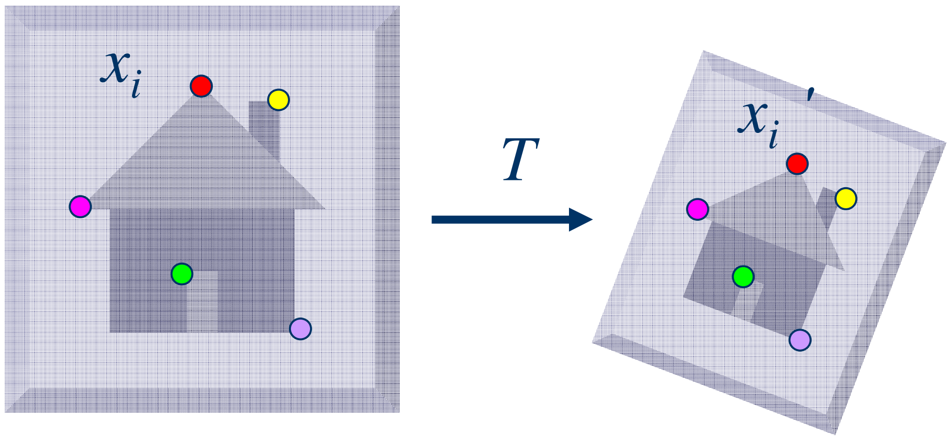


Achar modelo *M* que minimiza:

$$\sum_i \mathrm{residuo}(x_i, M)$$

Alinhamento como ajuste

- Obter um modelo para uma transformação entre dois conjuntos de características correspondentes em duas imagens



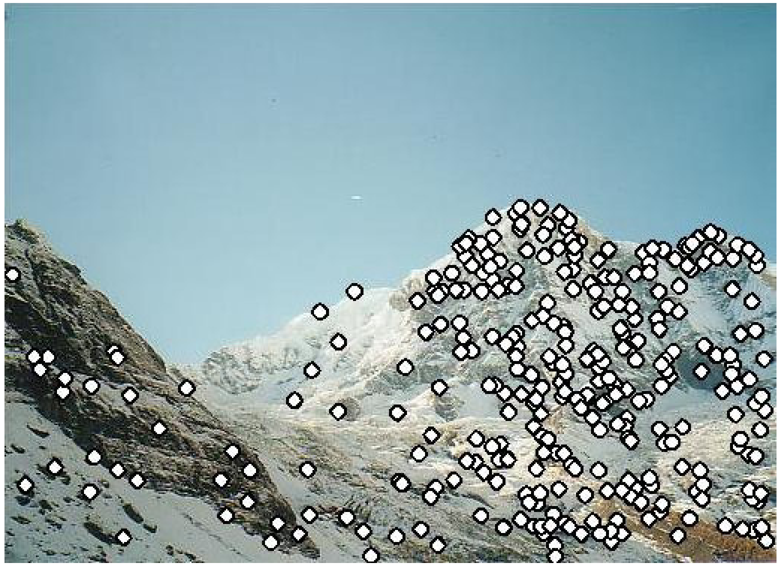
Achar transformação *T* que minimiza

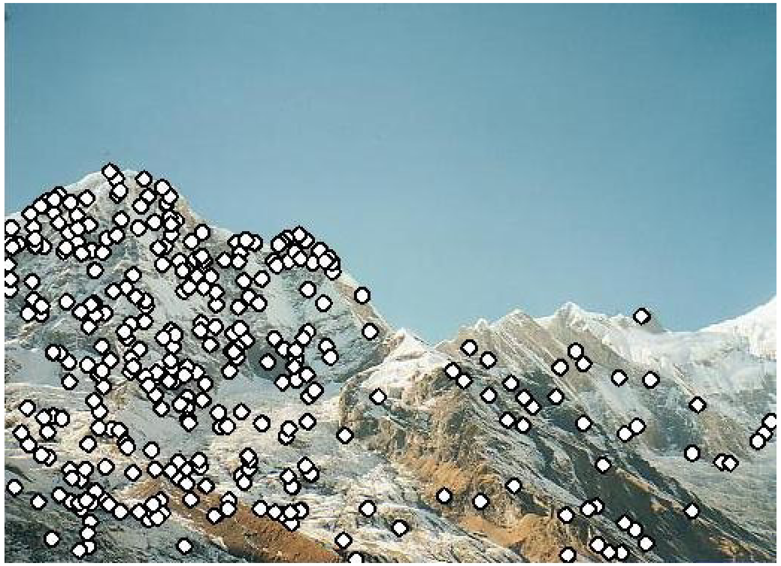
$$\sum_i \text{residuo}(T(x_i), x_i')$$

Alinhamento usando características

- Dadas duas imagens

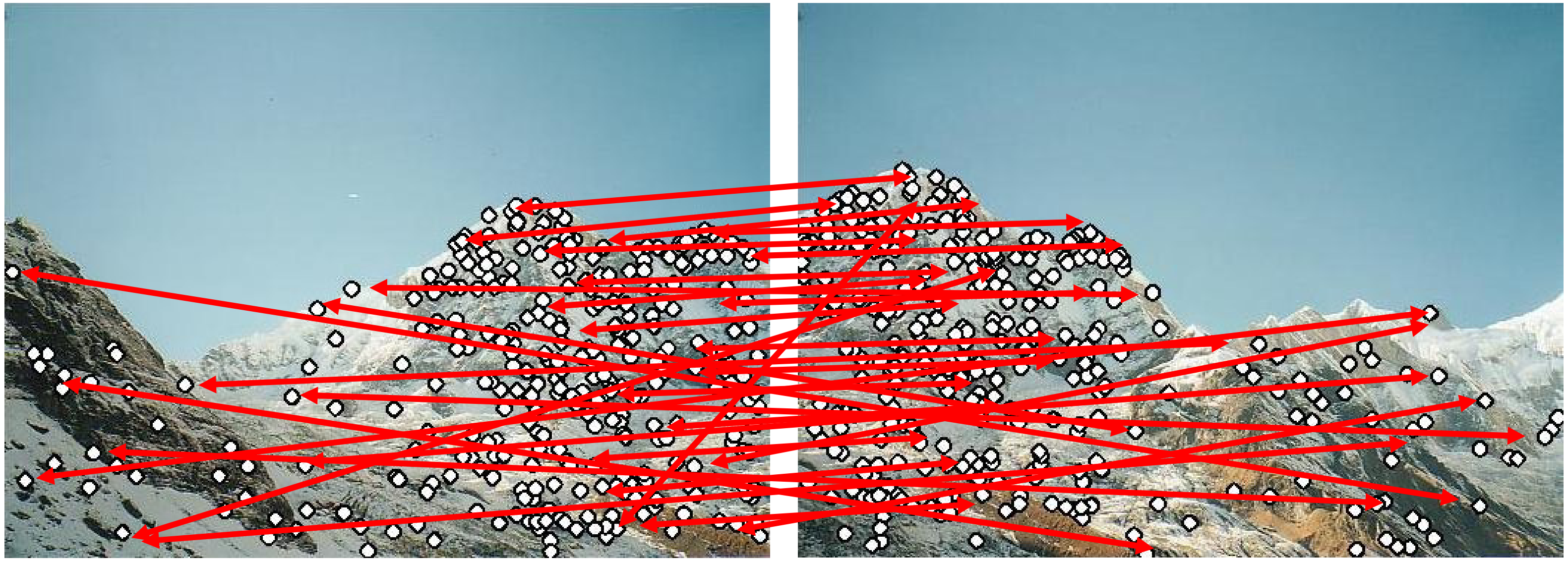
Alinhamento usando características





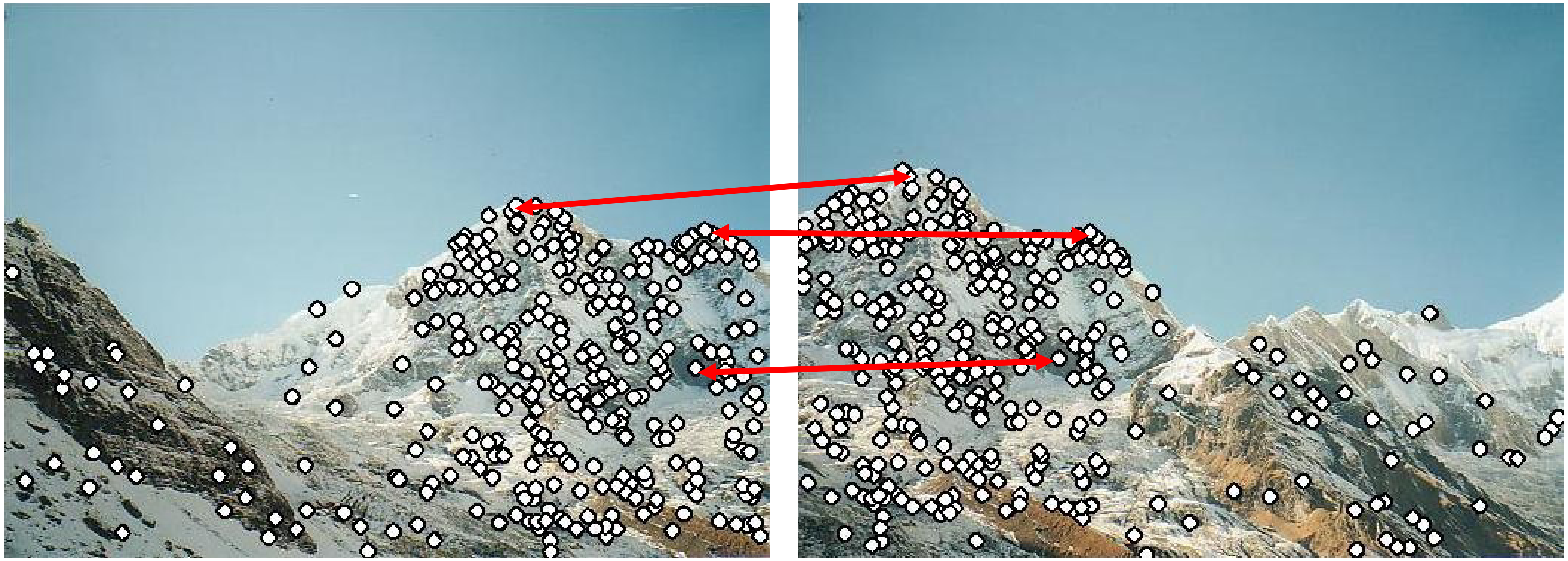
- Extrair cacacterísticas

Alinhamento usando características



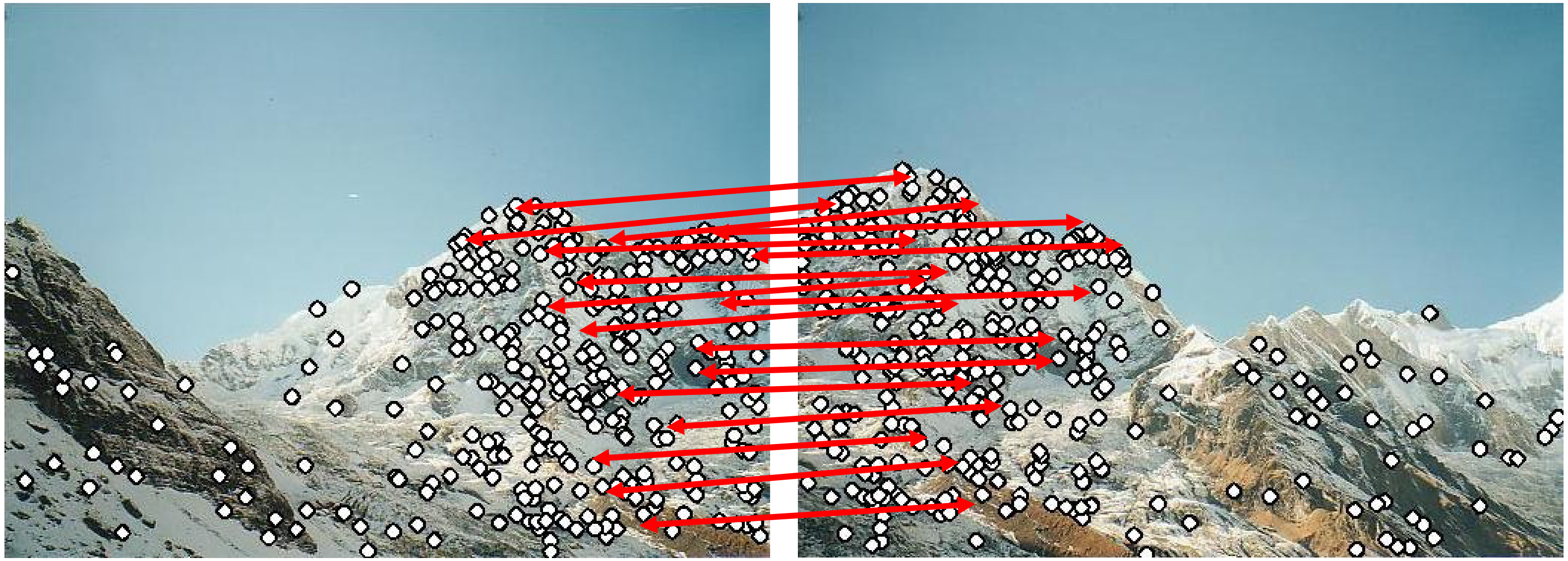
- Extrair caracterísiticas
- Determinar possíveis correspondencias

Alinhamento usando características



- Extrair caracterísiticas
- Determinar possíveis correspondências
- Repetir:
  - *Assumir* transformação *T* (usar *T* que relaciona um subconjunto de características);

Alinhamento usando características



- Extrair caracterísiticas
- Determinar possíveis correspondências
- Repetir:
  - *Assumir* transformação *T* (usar *T* que relaciona um subconjunto de características);
  - *Verificar* transformação (procurar por outras características que são consistentes com *T*)

Alinhamento usando características

- Extrair caracterísiticas
- Determinar possíveis correspondências
- Repetir:
  - *Assumir* transformação *T* (usar *T* que relaciona um subconjunto de características);
  - *Verificar* transformação (procurar por outras características que são consistentes com *T*)

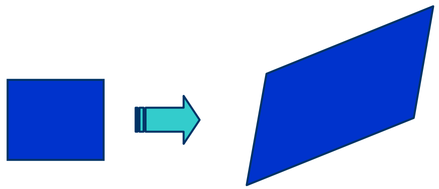
Transformações

- O que acontece quando temos duas imagens da mesma cena e tentamos alinhá-las?
- Possíveis movimentos:
  - Translação
  - Rotação
  - Escala
  - “Affine”
  - Perspective?

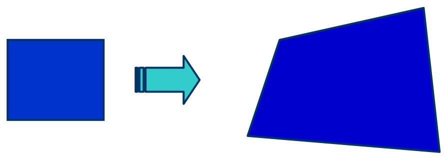
Transformações em 2D

- Similaridade (translação, escala, rotação)

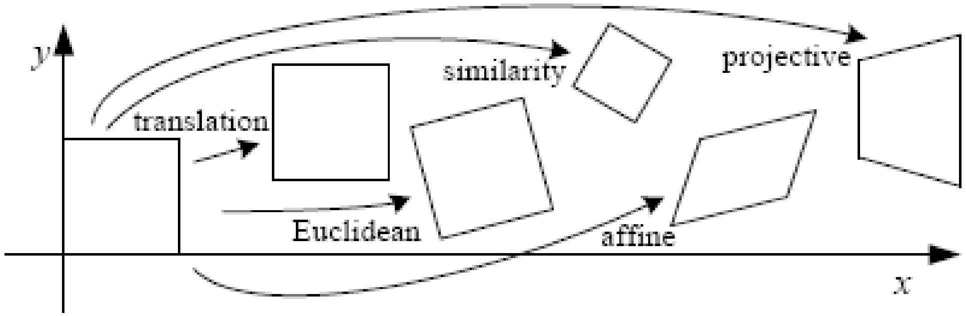
- “Affine”



- Perspectiva (homográfica)



Transformações



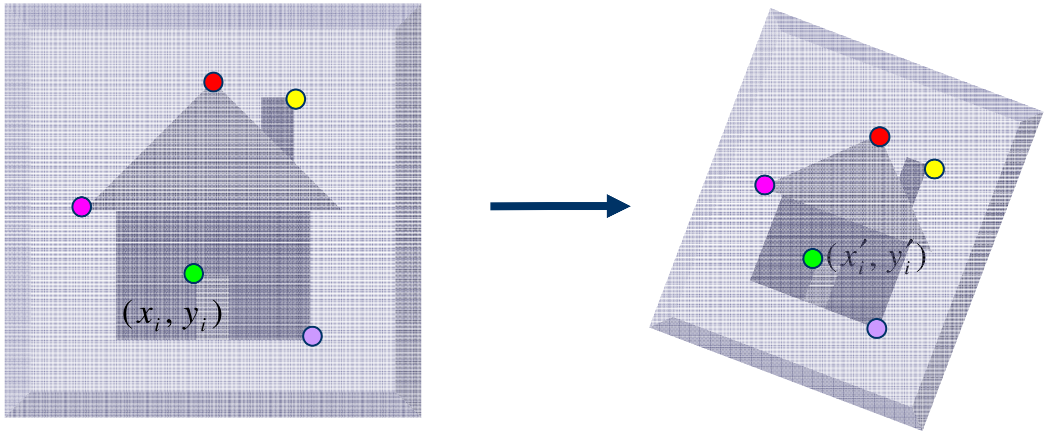
Name	Matrix	# D.O.F.	Preserves:	Icon
translation	$\begin{bmatrix} I \mid t \end{bmatrix}_{2\times 3}$	2	orientation $+\cdots$	
rigid (Euclidean)	$\begin{bmatrix} R \mid t \end{bmatrix}_{2\times 3}$	3	lengths $+\cdots$	
similarity	$\begin{bmatrix} sR \mid t \end{bmatrix}_{2\times 3}$	4	angles $+\cdots$	
affine	$\begin{bmatrix} A \end{bmatrix}_{2\times 3}$	6	parallelism $+\cdots$	
projective	$\begin{bmatrix} \tilde{H} \end{bmatrix}_{3\times 3}$	8	straight lines	

Transformação “affine”

- Procedimento de ajuste simples (método dos mínimos quadrados)
- Aproxima mudanças de ponto de vista da cena para objetos quase planos e câmeras comuns
- Pode ser usado para inicializar ajustes mais complexos

Transformação "affine"

- Assumindo que se conhece os pontos correspondentes $\Rightarrow$ como obter a transformação?



$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} & & \cdots & & & \\ x_i & y_i & 0 & 0 & 1 & 0 \\ 0 & 0 & x_i & y_i & 0 & 1 \\ & & \cdots & & & \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \cdots \\ x_i' \\ y_i' \\ \cdots \end{bmatrix}$$

Transformação “affine”

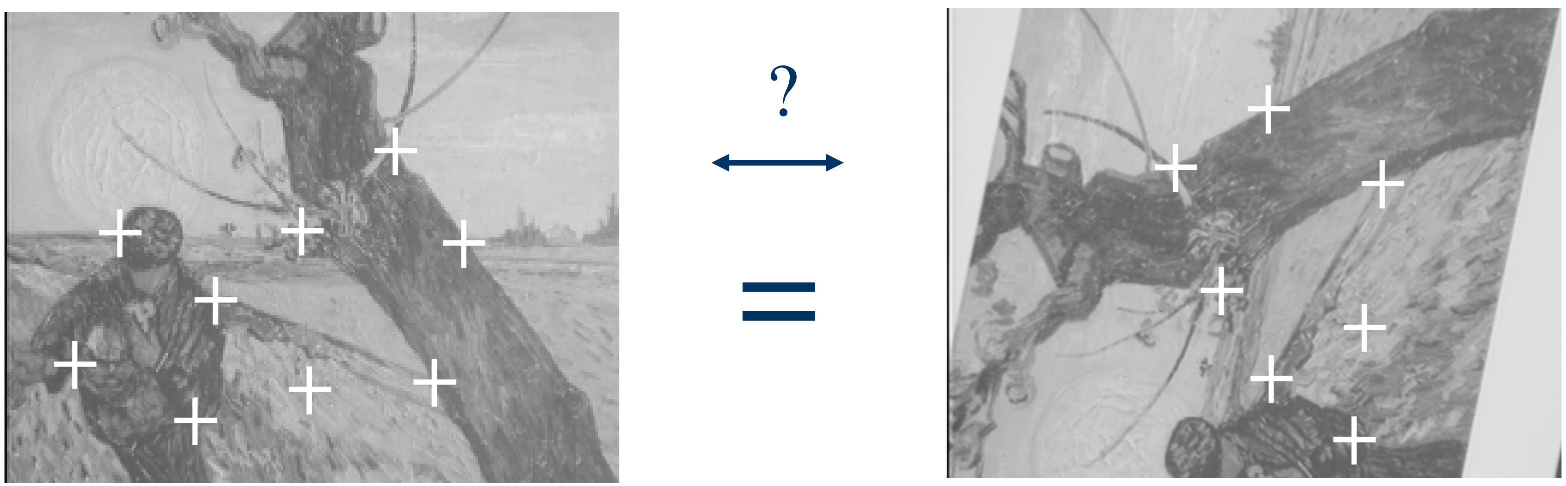
$$\begin{bmatrix} & & \cdots & & & \\ x_i & y_i & 0 & 0 & 1 & 0 \\ 0 & 0 & x_i & y_i & 0 & 1 \\ & & \cdots & & & \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ m_3 \\ m_4 \\ t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \cdots \\ x_i' \\ y_i' \\ \cdots \end{bmatrix}$$

- Sistema linear com 6 variáveis
- Cada característica correspondente fornece 2 equações linearmente independentes $\Rightarrow$ precisa pelo menos 3 características correspondentes para obter a transformação

Transformação “affine”

- Se as correspondências não forem conhecidas?

  ⇒ Precisa determinar características correspondentes por meio de verificação da similaridade



Busca de correspondentes

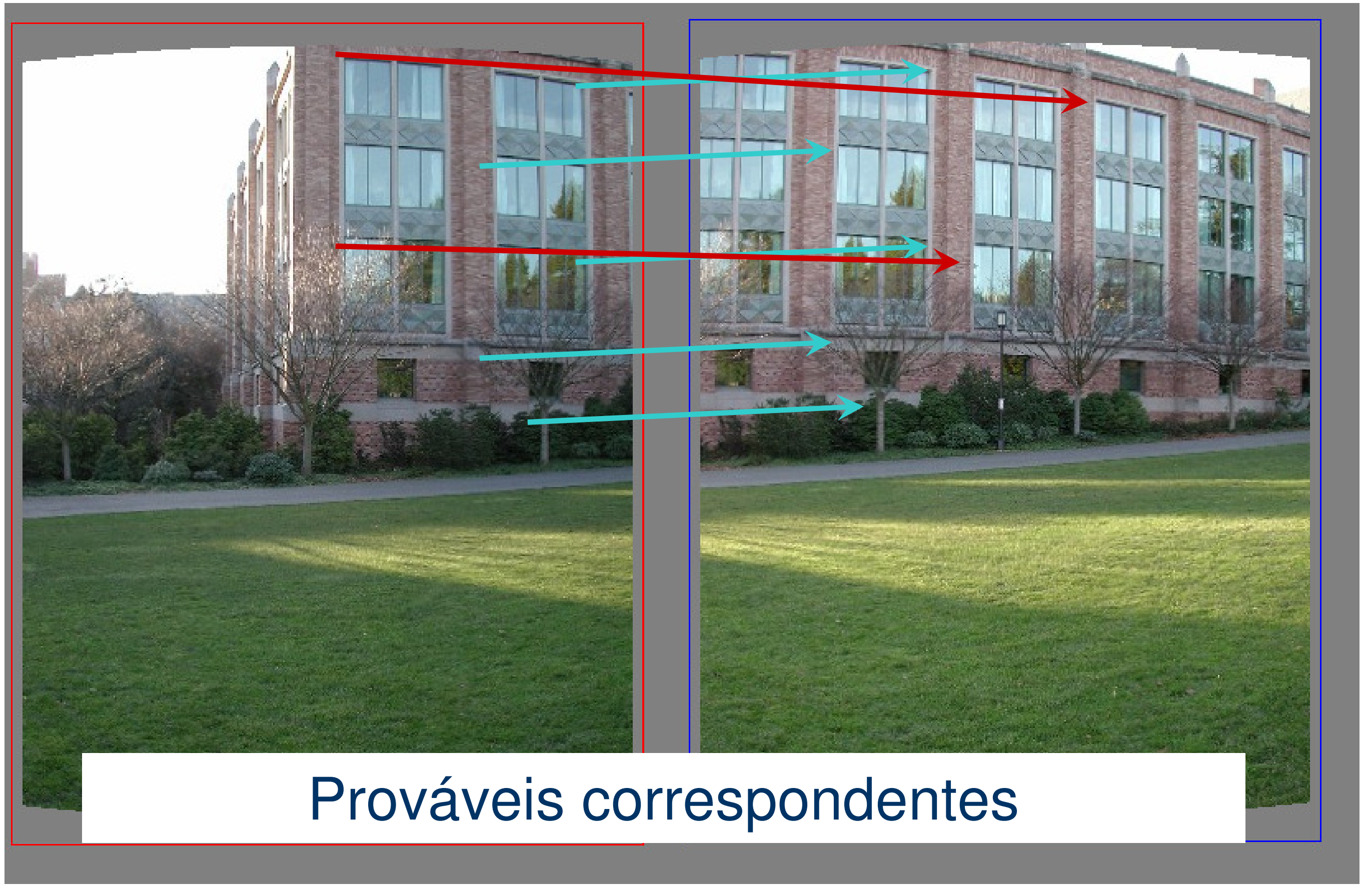
- Métodos para achar características correspondentes:
  - “Random Sample Consensus” (RANSAC)
  - Alinhamento incremental
  - Transformada de Hough
- Cuidado especial para identificar e isolar “outliers”

RANSAC

1. Selecionar aleatoriamente um grupo semente de correspondentes
2. Calcular a transformação desse grupo
3. Determinar as outras características que seguem essa transformação
4. Se o números de correspondentes for grande o suficiente $\Rightarrow$ recalcular a transformação com o grupo maior de correspondentes
5. Repetir o processo diversas vezes para outros grupo iniciais (novas sementes) $\Rightarrow$ guardar a transformação que é obedecida pelo maior número de caracterísiticas

RANSAC

- Considerando somente translação



Prováveis correspondentes

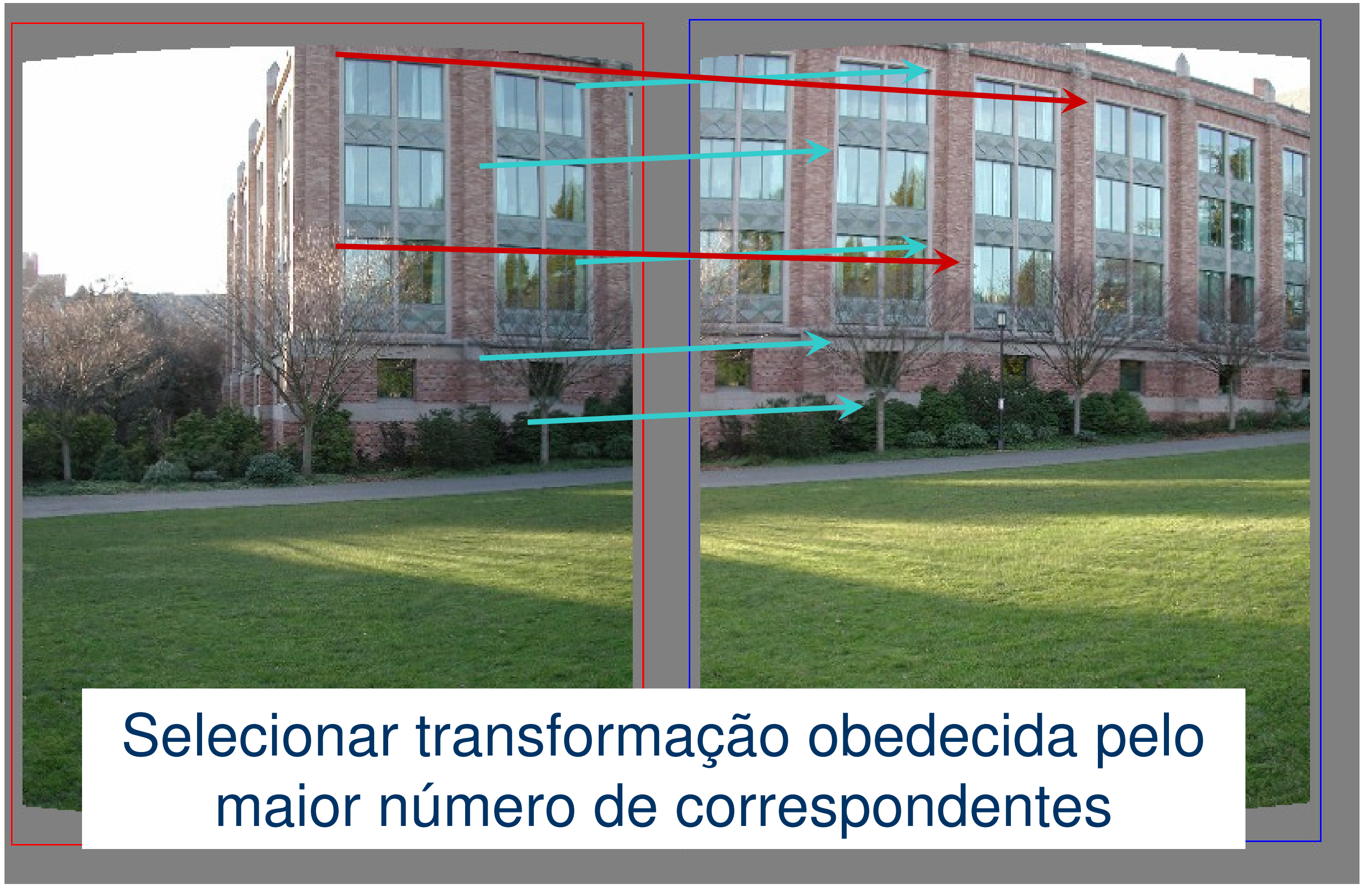
RANSAC

- Considerando somente translação

Selecionar um par, contar quantas outras características obedecem a transformação do par selecionado

RANSAC

- Considerando somente translação



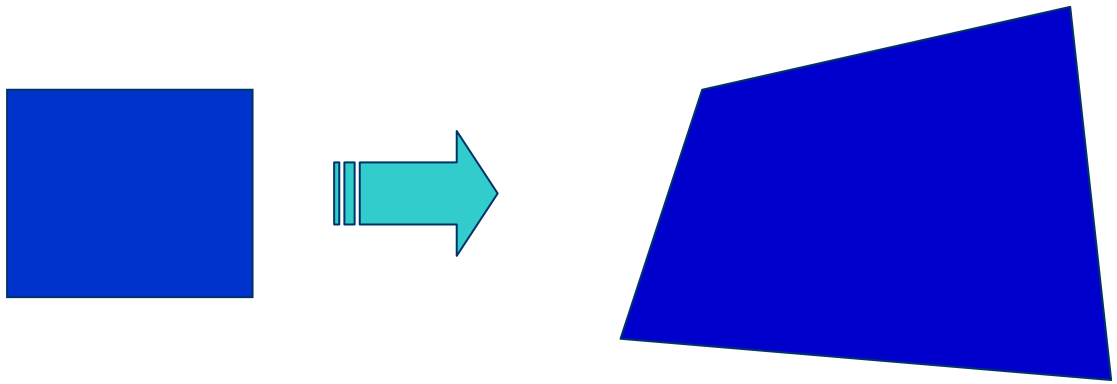
Selecionar transformação obedecida pelo maior número de correspondentes

RANSAC

- Problemas com o RANSAC:
  - Em muitas situações práticas o número de características que não obedecem a transformação é muito grande (90% ou mais)
  - Estratégia alternativa $\Rightarrow$ restringir busca em regiões limitadas da imagem $\Rightarrow$ Alinhamento Incremental

Transformação homográfica

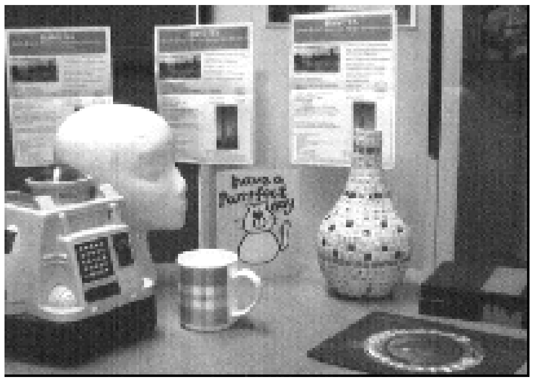
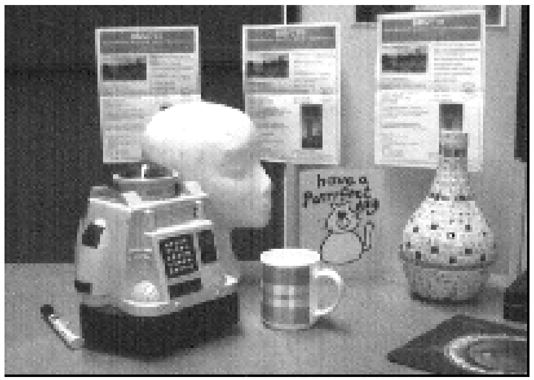
- **Homográfica** $\Rightarrow$ transformação de projeção (transforma um quadrado em um quadrilátero arbitrário)



Transformação homográfica

- Transformação entre duas vistas de uma superfície plana


- Transformação entre imagens obtidas de duas câmeras no mesmo centro

Transformação homográfica

- Transformação homogênea:

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

- Modelo de transformação homográfica:

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Transformação homográfica

- Mais compactamente:

$$\lambda \mathbf{x}_i' = \mathbf{H}\,\mathbf{x}_i = \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \\ \mathbf{h}_3^T \end{bmatrix} \mathbf{x}_i$$

- Vetores $\mathbf{x}'$ e $\mathbf{x}$ são paralelos $\Rightarrow$ produto vetorial igual a zero:

$$\mathbf{x}_i' \times \mathbf{H}\,\mathbf{x}_i = 0 \quad \Rightarrow \quad \mathbf{x}_i' \times \mathbf{H}\,\mathbf{x}_i = \begin{bmatrix} y_i' \mathbf{h}_3^T \mathbf{x}_i - \mathbf{h}_2^T \mathbf{x}_i \\ \mathbf{h}_1^T \mathbf{x}_i - x_i' \mathbf{h}_3^T \mathbf{x}_i \\ x_i' \mathbf{h}_2^T \mathbf{x}_i - y_i' \mathbf{h}_1^T \mathbf{x}_i \end{bmatrix}$$

Transformação homográfica

- Equação da transformação para um par correspondente:

$$\begin{bmatrix} 0^T & -\mathbf{x}_i^T & y_i' \mathbf{x}_i^T \\ \mathbf{x}_i^T & 0^T & -x_i' \mathbf{x}_i^T \\ -y_i' \mathbf{x}_i^T & x_i' \mathbf{x}_i^T & 0^T \end{bmatrix} \begin{pmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{pmatrix} = 0$$

Somente 2 equações são linearmente independentes

Transformação homográfica

- Equação da tranformação homográfica:

$$\begin{bmatrix} 0^T & \mathbf{x}_1^T & -y_1'\,\mathbf{x}_1^T \\ \mathbf{x}_1^T & 0^T & -x_1'\,\mathbf{x}_1^T \\ \cdots & \cdots & \cdots \\ 0^T & \mathbf{x}_n^T & -y_n'\,\mathbf{x}_n^T \\ \mathbf{x}_n^T & 0^T & -x_n'\,\mathbf{x}_n^T \end{bmatrix} \begin{pmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{pmatrix} = 0$$

  - 9 parâmetros, sendo que escala (λ) é arbitrária
  - Um par de características correspondentes fornece duas equações
  - Necessário no mínimo 4 pares de correspondentes para obter uma solução
  - Mais de 4 pares $\Rightarrow$ solução por minimização do erro

Mosaico ⇒ imagem panorâmica

- Unir várias imagens sobrepostas em uma única imagem panorâmica

 +  + … +  =

Mosaico ⇒ imagem panorâmica