Sistemas de Visão e Percepção Industrial

5-Processamento a Alto Nível

Parte 1 - Reconhecimento de Imagem: Modelos

Universidade de Aveiro
Vitor Santos, 2022-2023

Sumário

Referências

- Burger, Cap. 17
- Gonzalez, Cap. 12

Introdução

Problemática

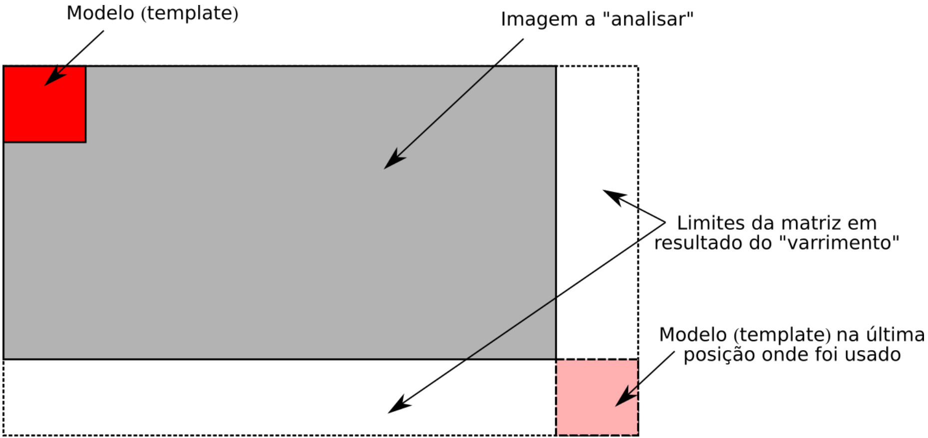
- Reconhecimento de imagens
  - Por comparação com modelos (*templates*)
    - Representações de imagens (subconjuntos de pixels)
  - Por comparação com padrões (*patterns*)
    - "Padrões" são arranjos ou combinações de descritores.
- Como se efetua a comparação?
  - Avaliando uma qualquer espécie de diferença ou relação.
  - Recorrendo a ...
    - ... cálculo da "distância" entre imagens
    - ... cálculo de "distância" entre padrões

Comparar imagens

- Pixel a pixel (diferenças entre valores de pixels)
  - Geralmente impraticável porque:
    - pequenas diferenças individuais podem implicar grandes diferenças globais.
    - As imagens podem não ter o mesmo número de pixels.
- Alternativa
  - Comparar sub-regiões da imagem com uma imagem de referência (modelo – *template*)
  - Encontrar a correspondência do modelo em alguma parte da imagem global
    - *Template matching* (correspondência de modelos)

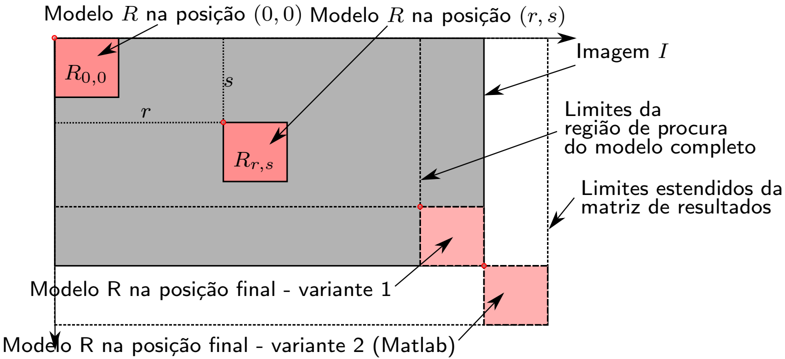
Princípio da operação de "comparação"

- "Varrer" a imagem por linhas e colunas e ...
- ... em cada ponto calcular um indicador de "similaridade" do modelo com a região subjacente na imagem.



Formalização da "comparação"

- Imagem de referência ou modelo – R
- Imagem onde se faz a busca – I



- $R_{r,s}$ designa o modelo "colocado" na posição (r,s) da imagem I.
  - Dependendo do software, (r,s) pode ir até aos limites de I (caso do Matlab), ou apenas até onde R esteja confinado à imagem I.

Distâncias como métricas de comparação

As "distâncias" mais comuns

- Para cada posição (r, s) definem-se as seguintes:
  - Soma das diferenças absolutas:

$$d_A(r,s) = \sum_{(i,j)\in R} |I(r+i,s+j) - R(i,j)|$$

  - Diferenças máximas:

$$d_M(r,s) = \max_{(i,j)\in R} |I(r+i,s+j) - R(i,j)|)$$

  - Soma das diferenças quadráticas (distância Euclidiana N-dimensional):

$$d_E(r,s) = \sqrt{\sum_{(i,j)\in R} [I(r+i,s+j) - R(i,j)]^2}$$

A distância d_E em detalhe

- Para d_E, é suficiente minimizar o seu quadrado para obter a "melhor" correspondência (e reduzir o peso computacional da raiz quadrada):

$$d_E(r,s)^2 = \sum_{(i,j)\in R} [I(r+i,s+j) - R(i,j)]^2 =$$

$$= \underbrace{\sum_{(i,j)\in R} I^2(r+i,s+j)}_{A(r,s)} + \underbrace{\sum_{(i,j)\in R} R^2(i,j)}_{B} - 2 \underbrace{\sum_{(i,j)\in R} I(r+i,s+j)R(i,j)}_{C(r,s)}$$

- O termo B é constante: não depende de (r,s)
- O termo $A(r,s)$ designa a "energia" da imagem de busca "sob" o padrão R em cada instante. Se fosse constante ao longo da imagem, bastaria usar o termo...
- ... $C(r,s)$, também conhecido como "correlação" cruzada entre I e R.

Correlação Cruzada

A distância d_E e a correlação cruzada

- A correlação cruzada entre I e R define-se formalmente por:

$$(I \star R)(r,s) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I(r+i, s+j)R(i,j) = \sum_{(i,j)\in R} I(r+i, s+j)R(i,j)$$

- Pela expressão da d_E, verifica-se que quanto maior a correlação $C(r,s)$, menor será d_E ...
- ..., portanto, o ponto (r,s) onde a correlação for maior, indica a zona da imagem I mais similar a R (menor d_E).
- O problema é que o termo $A(r,s)$ não é em geral constante...
  - logo a correlação cruzada por si só não tem o mesmo efeito de d_E!
  - Para resolver essa questão, a solução passa por "normalizar" a correlação cruzada com outros termos!
- Símbolos comuns para o operador de correlação: $\star$, $\circledast$, $\otimes$

Correlação cruzada normalizada

- Para obviar a dependência anterior, define-se a correlação cruzada normalizada $C_N(r, s)$:

$$C_N(r,s) = \frac{C(r,s)}{\sqrt{A(r,s)}\sqrt{B}} = \frac{\sum\limits_{(i,j)\in R} I(r+i, s+j)R(i,j)}{\sqrt{\sum\limits_{(i,j)\in R} I^2(r+i, s+j)}\sqrt{\sum\limits_{(i,j)\in R} R^2(i,j)}}$$

- Mesmo assim, este valor C_N é uma medida absoluta que depende da intensidade local.
- A solução passará por "aferir" pelas médias de intensidades de I e R ($\bar{I}(r, s)$ e $\bar{R}$) abaixo definidas, sendo K o número de pixels de R:

$$\bar{I}(r,s) = \frac{1}{K}\sum_{(i,j)\in R} I(r+i, s+j) \quad , \quad \bar{R} = \frac{1}{K}\sum_{(i,j)\in R} R(i,j)$$

Correlação cruzada normalizada

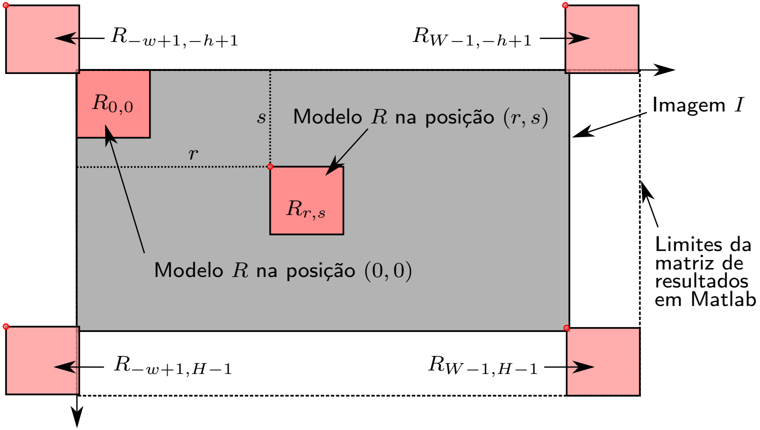
- Em conclusão, para se obter uma medida da correlação cruzada normalizada independente da intensidade média da imagen I (por exemplo, devido a iluminação variável), usa-se a expressão seguinte:

$$C_L(r,s) = \frac{\sum_{(i,j)\in R} [I(r+i,s+j) - \bar{I}(r,s)][R(i,j) - \bar{R}]}{\sqrt{\sum_{(i,j)\in R} [I(r+i,s+j) - \bar{I}(r,s)]^2} \sqrt{\sum_{(i,j)\in R} [R(i,j) - \bar{R}]^2}}$$

- Em matlab este valor obtém-se com a função: `normxcorr2(R,I)`

Limites de (r, s) no cálculo da correlação cruzada

- Na verdade, o cálculo não se inicia para $(r, s) = (0, 0)$
- Mas sim $(r, s) = (-w + 1, -h + 1)$ com w, h as dimensões de R
- Isto origina uma matriz de correlação com as dimensões:
  - $(W + w - 1) \times (H + h - 1)$ sendo W, H as dimensões da imagem I



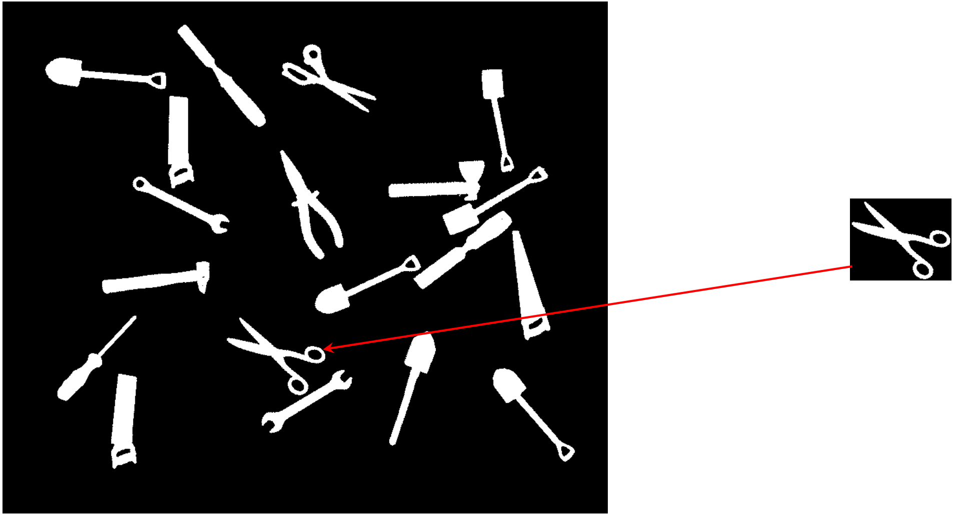
Observações

- A correlação cruzada define-se, e é operacional, para imagens **binárias** e para imagens a níveis de **cinzento**.
- Diminuir d_E (aumentar a similaridade) é equivalente a aumentar a correlação cruzada (ver expressão desenvolvida de d_E)
- Comparada com a distância quadrática (d_E), a correlação cruzada (normalizada) é mais robusta a variação geral de iluminação.

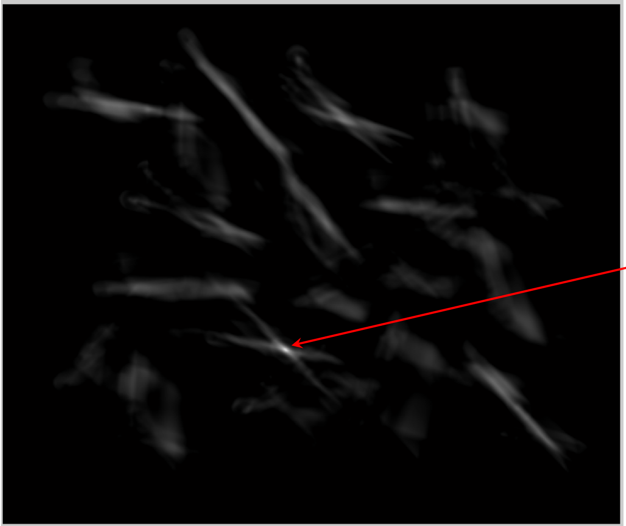
Exemplos

Exemplo de aplicação da correlação

- Na imagem de máscaras (binária) localizar a tesoura por correlação com base no modelo.



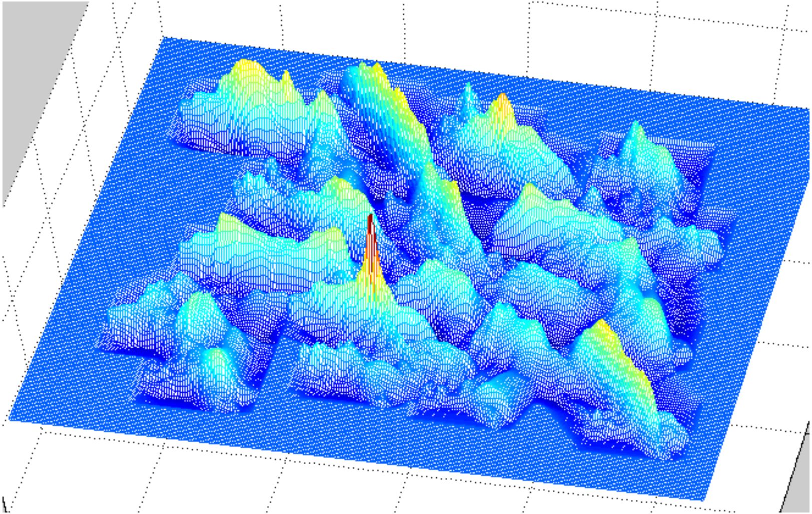
Resultado da correlação da tesoura



Zona do maior pico da correlação

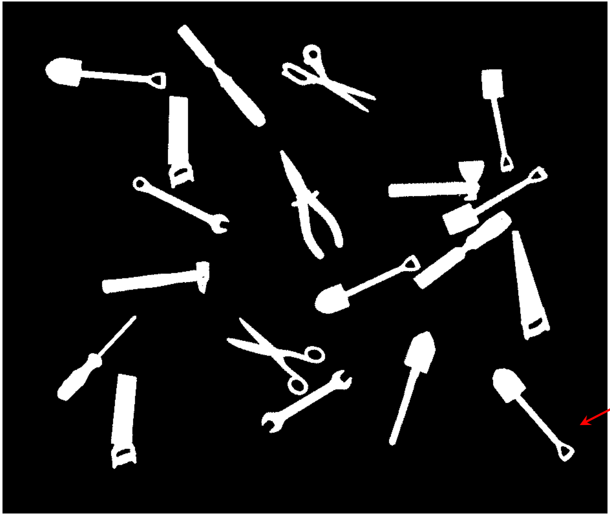
vs_corr1.m

Resultado da correlação visto a 3D

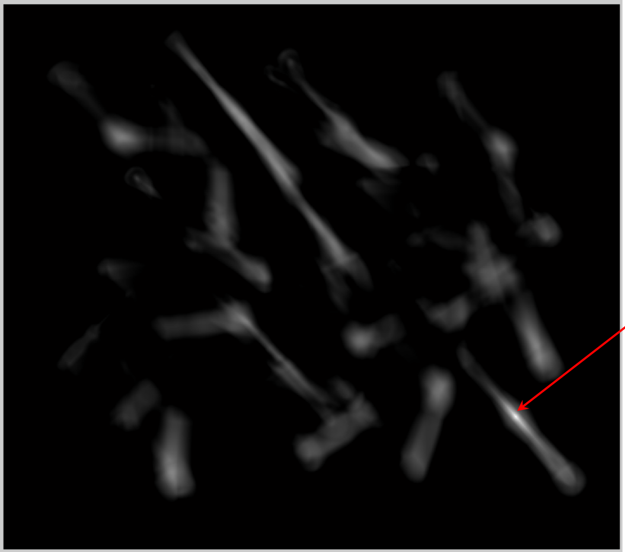


Exemplo para outro objecto

- Na imagem de máscaras (binária) localizar a pá por correlação com base no modelo.

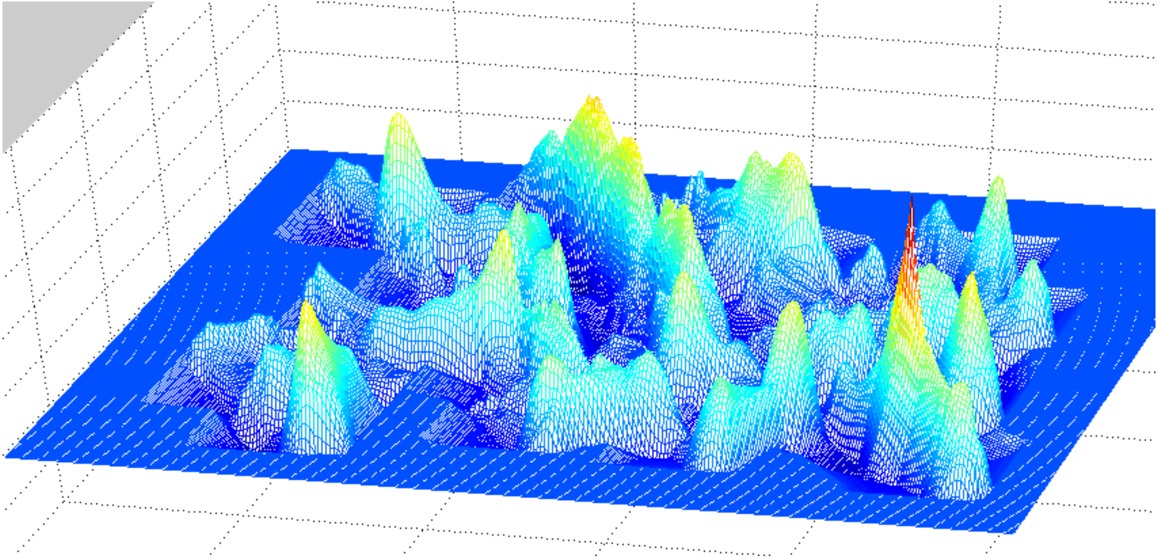


Resultado da correlação da pá



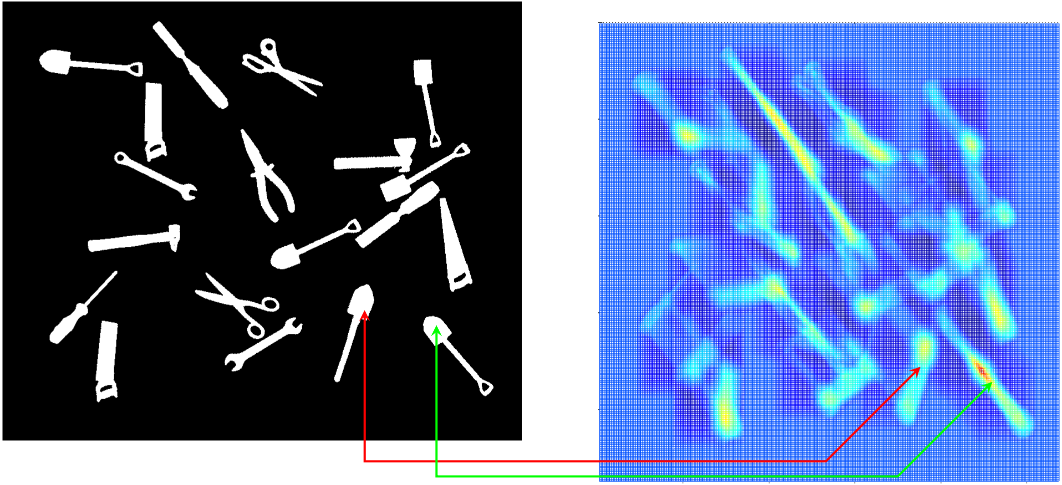
Zona do maior pico da correlação

Resultado da correlação da pá visto a 3D



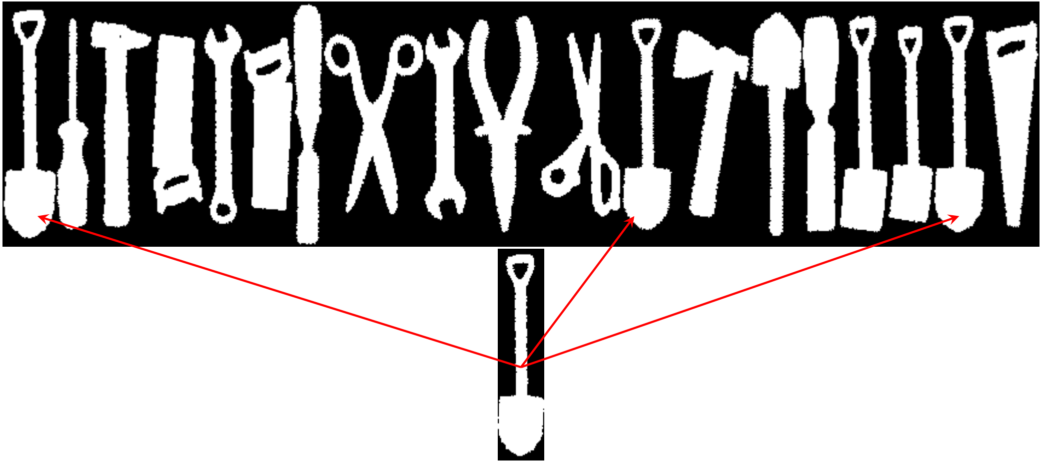
Sobre a generalização da abordagem

- A imagem possui porém outras pás que não foram bem (ou mesmo nada bem) detetadas!
- A solução é sensível à escala e à orientação!



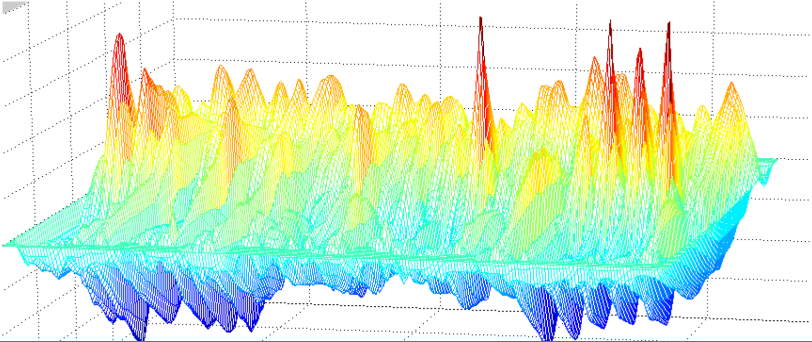
Limitações desta correlação

- Depende da orientação (e escala) dos objetos.
- Possível solução para a **orientação**: reorientar os objetos e usar um padrão orientado.
- **Nota**: esta solução não resolve todavia o problema da escala!



Solução para objetos reorientados

- Para um limiar de 80% encontram-se várias "pás"



Soluções adotadas por diversas aplicações

- Testes dos modelos (*templates*) em:
  - múltiplas escalas;
  - múltiplas orientações.
- Definição de limiares de "acerto" (*matching*) baseados no coeficiente de correlação.