

# Sistemas de Visão e Percepção Industrial

## 5-Processamento a Alto Nível

### Parte 2 - Reconhecimento de Imagem: Padrões

- 1 Generalidades e definições
- 2 Interpretação de distâncias
- 3 Exemplos numéricos

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

# Generalidades e definições

# Reconhecimento de imagens

- Por comparação com padrões (*patterns*)
  - "Padrões" são arranjos ou combinações de descritores (*features*)
  - Recurso à distâncias entre padrões:
    - Euclidiana
    - Mahalanobis
- A definição dos padrões mais adequados para um dado problema pode ser complexo e muito dependente do problema.
  - Serão apresentados exemplos generalizáveis e indicações adicionais para o fazer.

# Norma de um vetor

- Seja um vetor  $\mathbf{x}$  com  $n$  componentes:

$$\vec{x} = \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1 \quad x_2 \quad \cdots \quad x_n]^\top$$

- A sua norma (Euclidiana) é dada por:  $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

- Ou, na forma matricial:  $\|\mathbf{x}\| = \sqrt{[x_1 \quad x_2 \quad \cdots \quad x_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}} = \sqrt{\mathbf{x}^\top \mathbf{x}}$

# "Distância" entre padrões

- Seja um padrão genérico com, por exemplo, 3 descritores:

$$\mathbf{x} = [x_1 \quad x_2 \quad x_3]^\top$$

- Seja um padrão de referência da mesma classe:

$$\boldsymbol{\mu}_x = [\mu_1 \quad \mu_2 \quad \mu_3]^\top$$

- A diferença entre os dois pode ser dada pela distância euclidiana (por vezes, usa-se o quadrado por questões computacionais):

$$d_E(\mathbf{x} - \boldsymbol{\mu}_x) = \|\mathbf{x} - \boldsymbol{\mu}_x\| = \sqrt{(\mathbf{x} - \boldsymbol{\mu}_x)^\top (\mathbf{x} - \boldsymbol{\mu}_x)}$$

# Distância de Mahalanobis

- É uma generalização da distância Euclidiana para acomodar a covariância entre componentes do vetor (os descritores):

$$D_M(\mathbf{x} - \boldsymbol{\mu}_x) = \|\mathbf{x} - \boldsymbol{\mu}_x\|_M = \sqrt{(\mathbf{x} - \boldsymbol{\mu}_x)^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}_x)}$$

- Onde  $\Sigma$  é a matriz de covariâncias e que, para vetores com 3 componentes, é dada por:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix}$$



# Casos particulares da Mahalanobis

- Se a matriz  $\Sigma$  for identidade, então tem-se o caso da distância Euclidiana
- Se a matriz  $\Sigma$  for diagonal (inexistência de termos de covariância):

$$\Sigma_{diag} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix} \quad \text{e também} \quad \Sigma_{diag}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} \end{bmatrix}$$

- Tem-se a chamada distância Euclidiana normalizada:

$$D_{E_N}(\mathbf{x}, \boldsymbol{\mu}_x) = \sqrt{\sum_{i=1}^N \frac{(x_i - \mu_{x_i})^2}{\sigma_i^2}}$$

- Neste exemplo seria  $N = 3$  na fórmula anterior.

## Definição de covariância de duas V. A.

- Sejam  $X$  e  $Y$  duas variáveis aleatórias (V.A.):

$$\sigma_{XY} = \text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) = E(X \cdot Y) - E(X)E(Y)$$

- $E(X)$  é a esperança matemática (na prática, é a média):

$$E(X) = \sum_{i=1}^N x_i p_i \hat{=} \frac{1}{N} \sum_{i=1}^N x_i \quad , \quad E(X \cdot Y) \hat{=} \frac{1}{N} \sum_{i=1}^N (x_i y_i)$$

$\hat{=}$  Significa "estimado por..."

- Se  $X$  e  $Y$  forem independentes, a sua covariância é 0
- Por outro lado, a  $\text{Cov}(X, X)$  é a variância de  $X$

$$\sigma_{XX} = \sigma_X^2 = \text{Cov}(X, X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 = E(X^2) - [E(X)]^2$$

# Demonstração da expressão da covariância

$$\begin{aligned}\sigma_{XY} &= \text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) = \\&= \frac{1}{N} \sum_{i=1}^N (x_i y_i - \bar{x} y_i - x_i \bar{y} + \bar{x} \bar{y}) = \\&= \bar{x} \bar{y} + \frac{1}{N} \left[ \sum_{i=1}^N (x_i y_i) - \bar{x} \sum_{i=1}^N y_i - \bar{y} \sum_{i=1}^N x_i \right] = \\&= \bar{x} \bar{y} - \bar{x} \bar{y} - \bar{y} \bar{x} + \frac{1}{N} \sum_{i=1}^N (x_i y_i) = \frac{1}{N} \sum_{i=1}^N (x_i y_i) - \bar{x} \bar{y} = \\&= E(X \cdot Y) - E(X) E(Y)\end{aligned}$$

# Interpretação de distâncias

# Como interpretar a distância $D(\mathbf{x}_n, \boldsymbol{\mu}_m)$ ?

- A expressão  $D(\mathbf{x}_n, \boldsymbol{\mu}_m)$  traduz a distância da amostra  $\mathbf{x}_n$  a um padrão de referência  $\boldsymbol{\mu}_m$
- Duas possibilidades de interpretação surgem:
  - Todas as classes possíveis para  $\mathbf{x}_n$  são conhecidas...?
  - ... ou pode haver classes desconhecidas? Isto é, não há padrões de referência  $\boldsymbol{\mu}$  para todas as classes!
- Ou seja:
  - Caso 1: Se as possibilidades de classificação são todas conhecidas, então, para uma dada amostra  $\mathbf{x}_n$ , procura-se o menor valor de  $D(\mathbf{x}_n, \boldsymbol{\mu}_j)$  para todos os valores de  $j$ ;
  - Caso 2: Se houver classes desconhecidas (possibilidade de objetos estranhos), pode-se definir um limiar  $L$  de distância acima do qual não se considera que haja reconhecimento, i.e.,  $\mathbf{x}_n$  não pertence a nenhuma classe (objeto desconhecido) porque a sua distância é "demasiado" grande a qualquer dos padrões de referência definidos!

# Formalização da interpretação de $D(\mathbf{x}_n, \boldsymbol{\mu}_m)$

- Caso 1: Não se admitem objetos estranhos:

$$\text{class}(\mathbf{x}_n) = \text{class} \left( \boldsymbol{\mu}_k : D(\mathbf{x}_n, \boldsymbol{\mu}_k) = \min_j D(\mathbf{x}_n, \boldsymbol{\mu}_j) \right)$$

- Caso 2: Pode haver objetos desconhecidos
  - Estabelecer um limiar relativo  $L$  (sendo  $L < 1$ )

$$\text{class}(\mathbf{x}_n) = \text{class} \left( \boldsymbol{\mu}_k : \frac{D(\mathbf{x}_n, \boldsymbol{\mu}_k)}{\max_j D(\mathbf{x}_n, \boldsymbol{\mu}_j)} < L \right)$$

- mas isto pode levar a conclusões múltiplas porque pode haver vários  $\boldsymbol{\mu}_k$  que verifiquem esta condição!!
- A complexidade do problema pode estar na definição de  $L$ .
  - Dependendo do problema, podem usar-se limites de  $\approx 1\%$  ou outros bem diferentes, a avaliar caso a caso.

# Exemplos numéricos

# Ilustração com um exemplo

- Sejam dois objetos que se sabem pertencer a duas categorias A e B (duas classes).
- Os objetos são identificados por padrões com 3 descritores, por exemplo:  $[f \ s \ e]^T$ 
  - Fator de forma ( $f$ )
  - Solidez ( $s$ )
  - Excentricidade ( $e$ )
- Dada uma lista de objetos, e com base no padrão que os representam, indicar a que categoria cada um deles apresenta mais semelhança usando os conceitos de distância Euclidiana e de distância de Mahalanobis.



## Exemplo – Padrões de referência

- Sejam os seguintes objetos de referência e os respectivos descritores  $[f \ s \ e]^T$ :

- Objeto A (pá)



- $p_A = [0.1983 \ 0.5354 \ 0.9871]^T$

- Objeto B (Serrote)



- $p_B = [0.3783 \ 0.8289 \ 0.9712]^T$

- A qual das duas categorias (se alguma delas) pertencem os objetos na figura seguinte?

# Exemplo – Padrões dos objetos todos



$$p_A = \begin{bmatrix} 0.1983 \\ 0.5354 \\ 0.9871 \end{bmatrix}$$



$$p_B = \begin{bmatrix} 0.3783 \\ 0.8289 \\ 0.9712 \end{bmatrix}$$

# Exemplo – Distâncias Euclidianas

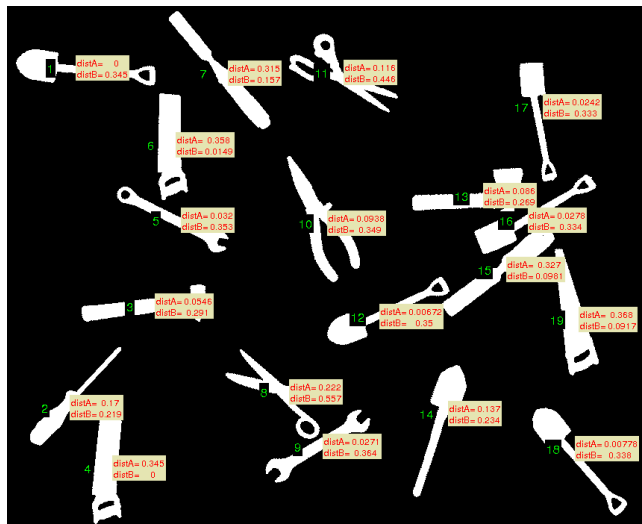
- Usando a distância Euclidiana simples ter-se-á:

Obj	Ffactor	Solidity	Eccentr.	PattA	PattB
1	0.1983	0.5354	0.9871	0.00	0.34
2	0.1982	0.7058	0.9914	0.17	0.22
3	0.2198	0.5855	0.9887	0.05	0.29
4	0.3783	0.8289	0.9712	0.34	0.00
5	0.1687	0.5458	0.9935	0.03	0.35
6	0.3908	0.8372	0.9711	0.36	0.01
7	0.2243	0.8493	0.9954	0.32	0.16
8	0.0998	0.3508	0.9118	0.22	0.56
9	0.1718	0.5296	0.9890	0.03	0.36
10	0.1316	0.584	0.9425	0.09	0.35
11	0.1209	0.4664	0.9340	0.12	0.45
12	0.1918	0.5336	0.9871	0.01	0.35
13	0.2173	0.6142	0.9585	0.09	0.27
14	0.2049	0.6723	0.9846	0.14	0.23
15	0.2847	0.8504	0.9913	0.33	0.10
16	0.1831	0.5587	0.9888	0.03	0.33
17	0.1870	0.5568	0.9888	0.02	0.33
18	0.2045	0.5400	0.9870	0.01	0.34
19	0.3076	0.8869	0.9774	0.37	0.09

- Se, por exemplo, se definir o limiar de semelhança  $L$  como 0.02 os objetos 4 e 6 são bem reconhecidos como serrotes, e os 1, 12, 17 e 18 como pás!
- Se se alargasse o limiar  $L$  para 0.03 ainda se detetaria o objeto 16 como do tipo A (embora não seja o mesmo)...
- ...mas não se conseguiria alargar mais  $L$  para detetar o objeto 14 como pá sem antes detetar outros objetos completamente não relacionados...
- A maioria dos objetos não é de tipo A nem B, o que se reflete nas distâncias.

## Exemplo – Distâncias Euclidianas: 2

- Os padrões usados como referências foram dos objetos 1 e 4.
- Quanto menor  $distA$  ou  $distB$ , mais se aproxima o objeto do respectivo padrão de referência A ou B.



## Exemplo – Distâncias de Mahalanobis

- Para se usar a distância de Mahalanobis, é preciso dispor da matriz de covariâncias.
- Para isso, porém, é preciso informação estatística sobre os descritores, o que só é possível se se dispuser de múltiplas amostras.
- Neste caso, podemos usar o que temos para a obter:
  - 6 objetos do tipo A – pás (embora um deles um pouco diferente dos restantes)
  - 3 objetos do tipo B – serrotes (embora um deles um pouco diferente dos restantes)

# Estatísticas para os padrões de referência

- Os valores relevantes para o estudo estatístico são:

	<b>Obj</b>	<b>Ffactor</b>	<b>Solidity</b>	<b>Eccentr.</b>
	pa 1	0.1983	0.5354	0.9871
	pa 12	0.1918	0.5336	0.9871
	pa 14	0.2049	0.6723	0.9846
	pa 16	0.1831	0.5587	0.9888
	pa 17	0.1870	0.5568	0.9888
	pa 18	0.2045	0.5400	0.9870
serrote	4	0.3783	0.8289	0.9712
serrote	6	0.3908	0.8372	0.9711
serrote	19	0.3076	0.8869	0.9774

- A partir destes dados é possível obter as matrizes de covariância dos três descritores para cada tipo de objeto.

## Exemplo – As matrizes de covariância

- Por cálculo manual (Cf. slides anteriores), ou usando o Matlab (função `cov()`), podemos obter as seguintes matrizes de covariância para os padrões de referência A e B:

$$\Sigma_A = \begin{bmatrix} 0.0828 & 0.1933 & -0.0119 \\ 0.1933 & 2.8194 & -0.0586 \\ -0.0119 & -0.0586 & 0.0024 \end{bmatrix} \times 10^{-3}$$

$$\Sigma_B = \begin{bmatrix} 2.0129 & -1.3552 & -0.1612 \\ -1.3552 & 0.9831 & 0.1121 \\ -0.1612 & 0.1121 & 0.0131 \end{bmatrix} \times 10^{-3}$$

- De notar que as matrizes têm determinantes muito pequenos e, portanto, a inversa pode originar imprecisões por se estar próximo da singularidade.
- Isto é devido sobretudo a muito poucas amostras e à sua limitada coerência, em especial no caso B (3 amostras)!

## Exemplo – Sobre Mahalanobis

- Agora já se pode aplicar a expressão da distância de Mahalanobis apresentada antes:

$$D_M(x - \mu_x) = \sqrt{(x - \mu_x)^\top \Sigma^{-1} (x - \mu_x)}$$

- Como referido antes, e para normalizar o resultado final (e eliminar fatores de escala), deve-se dividir o conjunto das distâncias obtidas pelo máximo de todas elas e assim obter um limiar relativo (percentual).
- Bastará de seguida definir um limiar para concluir das semelhanças e a que categoria pertencem os objetos.



# A função mahal() no Matlab

- $d = \text{mahal}(Y, X)$

		Variáveis					Variáveis		
Y=	Amostras a testar	$a_1$	$b_1$	$c_1$	X=	Amostras para definir o padrão de referência	$A_1$	$B_1$	$C_1$
		$a_2$	$b_2$	$c_2$			$A_2$	$B_2$	$C_2$
		$a_3$	$b_3$	$c_3$			$A_3$	$B_3$	$C_3$
		$\dots$	$\dots$	$\dots$			$A_4$	$B_4$	$C_4$
		$a_{N-1}$	$b_{N-1}$	$c_{N-1}$			$A_5$	$B_5$	$C_5$
		$a_N$	$b_N$	$c_N$			$\dots$	$\dots$	$\dots$
							$A_M$	$B_M$	$C_M$

- Y – matriz de amostras a testar ( $N$  amostras com descritores  $[a \ b \ c]$ )
- X – matriz de amostras para referência (para um dado padrão) de onde se calculará o padrão X e a respetiva matriz de covariâncias ( $M$  amostras dos descritores a usar para calcular a referência  $[A \ B \ C]$ )
- d – distâncias de Mahalanobis (ao quadrado) de cada uma das  $N$  amostras testadas com o padrão calculado de X:  $[\bar{A} \ \bar{B} \ \bar{C}]$  (médias das  $M$  amostras usadas para criar um padrão de referência)

## Exemplo – Comparação de métodos

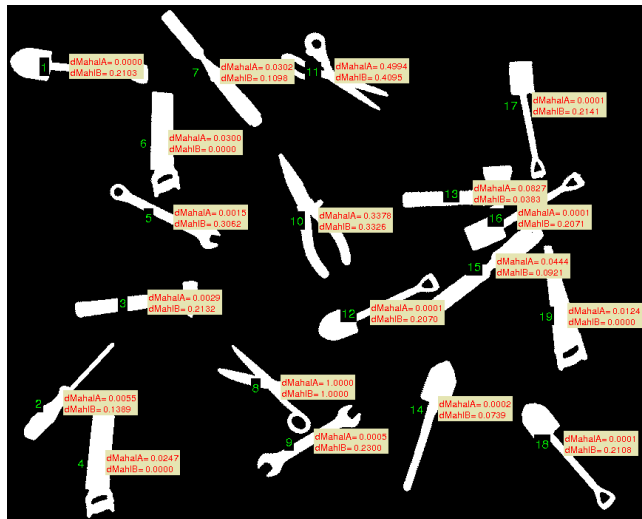
- A comparação final usando distância Euclidiana e distância de Mahalanobis (ambas **normalizadas**) fica do seguinte modo:

Obj	Ffactor	Solidity	Eccentr.	PattAEuc	PattBEuc	PattAMah	PattBMah
1	0.1983	0.5354	0.9871	0.0000	0.6193	0.0000	0.2103
2	0.1982	0.7058	0.9914	0.4628	0.3938	0.0055	0.1389
3	0.2198	0.5855	0.9887	0.1483	0.5228	0.0029	0.2132
4	0.3783	0.8289	0.9712	0.9362	0.0000	0.0247	0.0000
5	0.1687	0.5458	0.9935	0.0869	0.6342	0.0015	0.3062
6	0.3908	0.8372	0.9711	0.9731	0.0268	0.0300	0.0000
7	0.2243	0.8493	0.9954	0.8558	0.2826	0.0302	0.1098
8	0.0998	0.3508	0.9118	0.6040	1.0000	1.0000	1.0000
9	0.1718	0.5296	0.9890	0.0736	0.6542	0.0005	0.2300
10	0.1316	0.5840	0.9425	0.2548	0.6267	0.3378	0.3326
11	0.1209	0.4664	0.9340	0.3161	0.8017	0.4994	0.4095
12	0.1918	0.5336	0.9871	0.0182	0.6283	0.0001	0.2070
13	0.2173	0.6142	0.9585	0.2336	0.4828	0.0827	0.0383
14	0.2049	0.6723	0.9846	0.3724	0.4206	0.0002	0.0739
15	0.2847	0.8504	0.9913	0.8873	0.1762	0.0444	0.0921
16	0.1831	0.5587	0.9888	0.0755	0.5999	0.0001	0.2071
17	0.1870	0.5568	0.9888	0.0657	0.5985	0.0001	0.2141
18	0.2045	0.5400	0.9870	0.0212	0.6064	0.0001	0.2108
19	0.3076	0.8869	0.9774	1.0000	0.1648	0.0124	0.0000

- Na distância Euclidiana com um limiar de 3% identificam-se (bem) **3 pás e 2 serrotes**.
- Aumentando o limiar, o próximo objeto detetado seria o 9 que claramente não é uma pá.
- Na distância de Mahalanobis, os resultados são perfeitos (**as 6 pás e os 3 serrotes**) usando um limiar abaixo dos 0.03%!
- À parte o objeto 9 detetado nos 0.05%, os restantes ficam muito mais distantes!

# Exemplo – Distâncias Mahalanobis

- Os dois padrões usados como referências foram os mesmos;
- mas agora todos os da mesma categoria foram usados para definir as covariâncias;
- assim se compreendem os melhores resultados!



## Nota sobre descritores do padrão

- Não é garantido que os descritores usados no exemplo anterior sejam os melhores possíveis.
- Por exemplo, a excentricidade tem pouca variação e provavelmente não seria um bom descritor para este problema.
- Em teoria, qualquer outro descritor numérico definido numa escala similar poderia ser adicionado ao padrão.
- Por outro lado, por exemplo, o Número de Euler não seria um bom descritor a juntar a estes, nem será o mais indicado em geral para um tratamento estatístico integrado num padrão como estes.
- Ainda, não faria muito sentido juntar num mesmo padrão a "Área" e "Momentos de Hu" porque a respetiva gama de variação é tão diferente que pequenas flutuações relativas num deles esconderiam ou seriam escondidas para quaisquer valores absolutos do outro!

- A distância de Mahalanobis deu muito melhores resultados do que a distância Euclidiana simples.
- A distância de Mahalanobis permite a extensão da definição de padrões, abarcando objetos com alguma variabilidade nos descritores.
- Os descritores usados são robustos à orientação e à escala. Se isso não fosse forçoso, seria possível definir outros descritores interessantes e robustos para o problema em causa.
- Os exemplos realizados apenas usaram duas categorias (pá e serrote).
- Para classificar todos os objetos na sua categoria seria preciso definir os padrões de referência também para as restantes categorias.