

1. Quais os parâmetros intrínsecos da câmara?
  - a) **Ponto principal, distâncias focais em pixels, fator de “skewing”**
  - b) Posição do referencial, distâncias focais em unidades métricas
  - c) Posição do referencial, ponto principal e fator de “skewing”
  - d) Ponto principal, distâncias focais em pixels e posição do referencial
2. Qual o operador de deteção de arestas é **mais rápido**, mas muito **mais sensível ao ruído**?
  - a) **Roberts**
  - b) *Prewitt*
  - c) *Sobel*
  - d) *Canny*
3. Qual a vantagem do plano virtual da imagem relativamente ao plano real?
  - a) **Objetos com a orientação alinhada como a imagem analisar**
  - b) Objetos com orientação diferente da imagem a analisar
  - c) Objetos com a orientação alinhada como a imagem a analisar, mas maiores
  - d) Objetos com a orientação diferente da imagem a analisar, mas maiores
4. Qual a vantagem do filtro de mediana em relação à média?
  - a) **Não cria novos valores**
  - b) Cria novos valores
  - c) É um filtro dito linear
  - d) É um filtro de convulsão
5. O processo de calibração de uma câmara:
  - a) **Calcula-se através do produto dos seus parâmetros intrínsecos e extrínsecos**
  - b) Depende exclusivamente da matriz intrínseca da câmara
  - c) Não depende da distância focal da câmara
  - d) Não é influenciado pelo tamanho da imagem que foi fotografada
6. Num filtro de média 3x3, o peso de cada componente é:
  - a) **1/9**
  - b) 1/16
  - c) 1/4
  - d) 1/8
7. A operação num histograma que permite uniformizar a distribuição dos níveis de cinzento numa imagem é:
  - a) **Equalização**
  - b) Expansão
  - c) Contração
  - d) *Thresholding*

- **Equalização** (normalizar o histograma)
    - alteração do valor de cada pixel para uniformizar a distribuição de níveis de cinzento na imagem. Útil em casos de imagens adquiridas em circunstâncias diferentes de iluminação.
8. Das seguintes técnicas de deteção de bordas, qual delas usa gradiente de 2ª ordem?
  - a) **Log**
  - b) Sobel
  - c) Kirsh

d) Prewitt

- Técnicas baseadas em gradiente de primeira ordem
  - usam dois filtros de convolução (kernels)
  - Básico, Sobel / Prewitt, Robberts
  - Canny
  - ...
- Técnicas que usam gradientes de 2<sup>a</sup> ordem
  - Zero Crossing (ou Marr edge detector)
  - Usa laplacianos, LoG, DoG, etc.
  - ...

9. Relativamente ao espaço de cores, assinale a opção correta:

- a) Sistema de cores RGB varia entre 0-255  
 b) Sistema de cores HSV varia entre 0-255  
 c) Sistema de cores RGB varia entre 0-1  
 d) Sistema de cores HSV varia entre 0-100

RGB varia [0,255]  
 HSV varia [0,1]

10. O operador morfológico fecho (*closing*):

- a) Suaviza o objeto pelo exterior do seu contorno  
 b) Tem como funcionalidade aplicar uma erosão seguida de uma dilatação  
 c) Suaviza o objeto pelo interior do seu contorno  
 d) Tem a mesma funcionalidade do *Top-Hat*

Fecho (*closing*)

- Operação de dilatação seguida de erosão
- Um resultado comum interessante é o de fundir objetos que inicialmente só estariam ligados por um único pixel.
- "Suaviza" o objeto pelo exterior do seu contorno

11. A suavidade e o terceiro momento do histograma tem em comum o facto de:

- a) Serem normalizados dividindo por  $(L-1)^2$   
 b) Serem 0 para histogramas simétricos  
 c) Serem 0 para regiões de intensidade constante  
 d) Serem ambos positivos para histogramas desviados para a direita

Suavidade (smoothness)
<ul style="list-style-type: none"> <li>• A expressão:  <math display="block">R = \frac{\sigma^2}{1 + \sigma^2} = 1 - \frac{1}{1 + \sigma^2}</math></li> <li>• Mede a suavidade relativa da intensidade na região.</li> <li>• <math>R</math> é zero para uma região de intensidade constante.</li> <li>• <math>R</math> aproxima-se de 1 para regiões com grandes excursões dos valores das suas intensidades.</li> <li>• Na prática esta medida é usada <u>normalizada</u> na gama <math>[0; 1]</math> dividindo-a por <math>(L-1)^2</math></li> </ul>

Terceiro momento do histograma
<ul style="list-style-type: none"> <li>• A expressão:  <math display="block">\mu_3 = \sum_{i=0}^{L-1} (i - m)^3 h(i)</math></li> <li>• Mede a assimetria do histograma.</li> <li>• É zero para histogramas simétricos.</li> <li>• Positivo para histogramas desviados para a direita (em relação à média)</li> <li>• Negativo para histogramas desviados para a esquerda</li> <li>• Também se costuma <u>normalizar</u> dividindo-o por <math>(L-1)^2</math></li> </ul>

12. Qual a operação morfológica que define o *Hit-and-Miss*:

- a)  $D = A \otimes (B, C)$   
 b)  $D = A \ominus B$   
 c)  $C = A \oplus B$

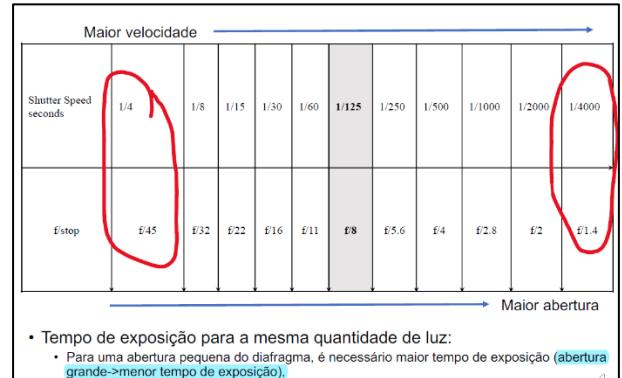
Usado na busca de "padrões" de pixels.  
 Recorre a dois elementos estruturantes:  

- $A \otimes (E, F) = (A \ominus E) \cap (A^C \ominus F)$

d)  $A \setminus B = A - B$

13. A aquisição de imagem pode ser condicionada de várias formas. Uma delas é uma função do movimento/ luz do objeto em estudo. Qual a relação entre o *f/stop* e o *shutter speed*?

- a) O Shutter speed diminui e o f/stop aumenta
- b) São inversamente proporcionais
- c) Dependem das dimensões da imagem
- d) Nenhuma das anteriores



14. Sendo P-Perímetro e A-Área, qual é a expressão para o cálculo do fator de forma?

a)  $FF = \frac{4\pi A}{P^2}$

b)  $FF = \frac{2\pi A}{P}$

c)  $FF = \frac{P^2}{4\pi A}$

d)  $FF = \frac{P^2}{4\pi A^2}$

$$ff = \frac{4\pi A}{P^2}$$

15. Quais os descritores cujos valores são independentes da escala e da rotação?

- a) Solidez, Excentricidade, Número de Euler
- b) Solidez, Área, Excentricidade
- c) Número de Euler, Excentricidade, Retângulo envolvente
- d) Número de Euler, Polígono convexo, Centroide

16. A gama de valores **240-300 (graus)** de **H** nos valores de HSV corresponde:

- a) Aos azuis
- b) Aos amarelos
- c) Às magentas
- d) Aos verdes

Os valores de HSV aparecem nas seguintes gamas:

- H em graus [0, 360]
  - 0-60 vermelhos
  - 60-120 amarelos
  - 120-180 verdes
  - 180-240 cians
  - 240-300 azuis
  - 300-360 magentas
- S e V na gama [0, 1]

17. Qual dos seguintes algoritmos deteta pixels fronteira?

- a) Kirsh
- b) Open

- c) Extent  
d) Shrink
18. Imagine uma imagem que contém as seguintes formas geométricas,  , que descriptor usaria para o reconhecimento eficiente desses objetos?
- a) Fator de forma  
b) Solidez  
c) Área  
d) Perímetro
19. O *Top-Hat* tem robustez:
- a) Às variações de iluminação  
b) Às variações de orientação  
c) Às variações de tamanho  
d) Nenhuma das anteriores
- O top-hat tem robustez às variação de iluminação.  
Exemplo com imagem original com intensidade duplicada
20. A operação *shrink* do Matlab equivale ao *thinning* onde as afirmações:
- a) C1+C2 é o esqueleto reduzido a linhas fechadas se houver objetos com buracos. No algoritmo de "thinning" há diversas combinações das condições C1, C2 e C3.  
Objetos sem buracos desaparecem
- b) C1+C2 traduz o esqueleto completo tradicional
- c) C1+C2 é o esqueleto reduzido a 1 pixel
- d) C1+C2 é o esqueleto constituído por pixels isolados e para objetos com "buracos" e por contornos fechados para objetos com "buracos"
- Só condição C1
    - Esqueleto reduzido a um pixel
  - Só condição C2
    - Esqueleto reduzido a linhas fechadas se houver objetos com buracos.  
Objetos sem buracos desaparecem.
  - C1 + C2
    - O esqueleto é constituído por pixels isolados para objetos sem "buracos" e por contornos fechados para objetos com "buracos".
  - C1 + C2 + C3
    - Ter-se-á o esqueleto completo tradicional
21. Se a imagem for trimodal, que binarização será a mais indicada? Assinale a opção correta.
- a) Binarização por histograma trimodal  
b) Binarização pela mediana  
c) Binarização pelo método de *Otsu*  
d) Binarização pelo algoritmo de isodados
22. Uma imagem com um único pixel branco corresponde a uma transformada de *Hough* com:
- a) Uma única curva que traduz todas as retas imaginárias que podem passar nesse ponto  
b) Mais do que uma curva que traduzem todas as retas imaginárias que podem passar nesse ponto  
c) Uma reta imaginária que passa nesse ponto  
d) Nenhuma das anteriores
- É um conjunto de curvas (discretizadas) ...  
• uma curva para cada pixel branco da imagem  
... que representam todas as retas possíveis que passam em cada um desses pontos.  
A uma imagem com um único pixel branco corresponde uma TH com uma única curva (que traduz todas as retas imaginárias que podem passar nesse ponto):
23. Uma correlação cruzada define-se e é operacional para que tipo/s de imagens?
- a) Binária e a níveis de cinzento  
b) Binária  
c) Qualquer tipo de imagem  
d) RGB
- A correlação cruzada define-se, e é operacional, para imagens binárias e para imagens a níveis de cinzento.

24. Na conversão de RGB para níveis de cinzento qual das seguintes cores é mais sensível ao olho humano e consequentemente tem maior proporção?

- a) Verde
- b) Vermelho
- c) Azul
- d) Preto

- Na conversão de RGB para níveis de cinzento...
  - ... poder-se-ia pensar numa média igual das três componentes..
  - ... mas o olho humano é mais sensível ao verde.
  - A proporção aproximada é:  

$$I = 0.30 \times R + 0.59 \times G + 0.11 \times B$$
  - Em Matlab pode-se usar a função:  
`rgb2gray()`

25. Qual das seguintes características **não se adequa** aos momentos invariantes da imagem?

- a) Sensíveis à mudança de escala
- b) Insensíveis à translação
- c) Insensíveis à reflexão
- d) Insensíveis à rotação

Os momentos de *Hu* são insensíveis à translação, mudança de escala, reflexão e rotação.

26. Qual o fator de forma de um quadrado, de uma circunferência e de um triângulo equilátero, respetivamente, considerando L o lado do quadrado, L o diâmetro da circunferência e L o lado do triângulo.

- a)  $\frac{\pi}{4}, 1, \frac{\sqrt{3}\pi}{9}$
- b)  $\frac{\pi}{2}, \frac{\pi}{2}, \frac{\sqrt{3}\pi}{3}$
- c)  $\frac{\pi}{4}, \frac{\pi}{2}, \frac{\sqrt{3}\pi}{3}$
- d)  $\frac{\pi}{2}, 1, \frac{\sqrt{3}\pi}{9}$

$ff = \frac{4\pi A}{P^2} \Rightarrow = 1$  para círculos;  $< 1$  para outros casos.

Quadrado

- $A = L^2$ ;  $P = 4L$
- $ff = \frac{4\pi L^2}{16L^2} = \frac{4\pi}{16} \approx 0.785$

Triângulo

- $A = \frac{\sqrt{3}L^2}{4}$ ;  $P = 3L$
- $ff = \frac{4\pi\sqrt{3}L^2}{9L^2} = \frac{\pi\sqrt{3}}{9} \approx 0.605$

27. Qual dos seguintes modelos de cor são os indicados para impressoras?

- a) CMYK
- b) RGB
- c) HSV
- d) YUV

**CMYK** (Cyan, Magenta, Yellow, Black)

- para impressão e impressoras – sistemas absorsores de luz (ou subtractivos)

28. A operação *shrink* do Matlab é equivalente a:

- a) C1+C2, ou seja, thinning
- b) Ao *thinning* com C1 e C3
- c) Ao *thinning* com C2 e C3
- d) Ao *thinning* com C1, C2 e C3

Exemplo da operação de "shrink" em Matlab

- Equivale ao "thinning" com C1 e C2  
`bw2=bwmorph(bw1, 'shrink', inf)`

29. A limitação através do método de *Otsu*, o limiar corresponde a:

- a) Valor que minimiza a variância intra-classes
- b) Valor que minimiza a variância inter-classes
- c) Valor que maximiza a variância intra-classes
- d) Nenhuma das anteriores

A ideia principal é de descobrir o limiar de binarização  $T$  que minimize a variância intra-classes W dada por:

30. Nos operadores morfológicos a variante *thinning* de esqueletização é **barrada** na erosão condicionada. Qual das seguintes condições **não** faz parte das combinações inerentes a este algoritmo?

- a) Um pixel não é erodido se o produto com SE=ones(3,3) for igual a 9
- b) Um pixel não é erodido se o produto com SE=ones(3,3) for igual a 1
- c) Um pixel não é erodido se afetar a conectividade
- d) Um pixel não é erodido se for um pixel terminal

Um pixel não é erodido se: (erosão condicionada)

- For isolado: C1
- Removê-lo afetar a conectividade: C2
- Removê-lo afetar o comprimento da linha: C3

31. Igual à 29.

32. Relativamente a sensores de imagem:

- a) O CCD tem maior sensibilidade à luz
- b) O CMOS é mais caro
- c) O CCD dá origem a imagens de menor qualidade
- d) O CMOS consome mais energia

#### CCD (Charge-Coupled Devices):

- **Array de condensadores**, a carga dos condensadores é proporcional à intensidade de luz, a carga do condensador é convertida em tensão,
- Imagens com **maior qualidade**,
- Menos suscetível ao ruído,
- Maior número de pixels,
- Maior sensibilidade à luz;

#### CMOS (Complementary Metal Oxide Semiconductor):

- **Array de transistores e foto diodos**, os foto diodos transformam a luz em tensão,
- **Menor custo**,
- Mais fácil construir,
- Menor consumo;

33. Qual das fórmulas descreve a distância euclidiana entre 2 pixels  $p$  e  $q$ ?

a)  $D(p, q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$

b)  $D(p, q) = (p_x - q_x)^2 + (p_y - q_y)^2$

c)  $D(p, q) = \max(|p_x - q_x|, |p_y - q_y|)$

d)  $D(p, q) = |p_x - q_x| + |p_y - q_y|$

#### Euclidiana

$$D_E(p, q) = \sqrt{(p_x - q_x)^2 + (p_y - q_y)^2}$$

34. Na operação morfológica de esqueletização define-se esqueleto como sendo:

- a) **Objeto filiforme que passa pelo "meio" do objeto e que preserva a topologia do objeto original**
- b) Objeto uniforme que passa pelo "meio" do objeto e que preserva a topologia do objeto original
- c) Objeto filiforme que passa pelo lado do objeto e que preserva a topologia do objeto original
- d) Estrutura de sustentação que passa pelo "meio" do objeto

#### Operadores Morfológicos - Esqueletização

- Operação de determinação do esqueleto
- Definição de esqueleto:
  - Objeto filiforme (1 pixel de largura) ...
  - ... que passa pelo "meio" do objeto
  - ... e que preserva a topologia do objeto original

35. As transformadas de *Hough* aplicam-se na:

- a) Determinação de formas geométricas, facilmente parametrizadas em imagens binarizadas
- b) Deteção de bordas dos objetos em imagens binarizadas
- c) Deteção de pontos de interseção dos objetos em imagens binarizadas
- d) Determinação de qualquer forma geométrica em imagens binarizadas

36. Um contorno diz-se minimamente ligado se:

- a) Todos os seus pontos exatamente 2 vizinhos que não são adjacentes em vizinhança de 4
- b) Nenhum dos pontos seja adjacente à sua vizinhança
- c) Todos os seus pontos tiverem mais que 2 vizinhos que não sejam adjacentes em vizinhança de 2

37. No espaço  $\rho\theta$  de uma imagem 1024x768, uma reta paralela ao eixo dos yy que passe por  $P=(200,100)$ :

- a) Tem  $\theta=0^\circ$
- b) Tem  $\theta=90^\circ$
- c) Tem  $\rho=1800$
- d) Tem  $\rho=0$

38. As câmaras termográficas fornecem imagem onde o valor do pixel é:

- a) Proporcional à radiação infravermelha
- b) Proporcional à radiação ultravioleta
- c) Desproporcional à radiação IV
- d) Desproporcional à radiação UV

39. No modelo HSV, as três componentes principais são:

- a) Tom, Saturação, Intensidade
- b) Luminância, Tom, Saturação
- c) Luminância, Saturação e Crominância
- d) Tom, Intensidade e Crominância

## Câmaras termográficas

- Fornecem imagem onde o valor do pixel é proporcional à intensidade de radiação infravermelha (e à temperatura dos corpos)
- Uso de pseudo-cor
- Análise similar à da imagem no visível em níveis de cinzento
- Tem aplicações em:
  - Medicina
  - Segurança
  - Inspeção de edifícios
  - Monitorização de máquinas
  - Etc.

### HSV:

- Hue- Tom
- Saturation- Saturação
- Value- Intensidade

40. A variante *thinning* da esqueletização é baseada na erosão condicionada, conhecida por ser:

- a) Menos exigente computacionalmente
- b) Mais exigente computacionalmente
- c) Condicionada à morfologia do objeto
- Condicionada ao esqueleto do objeto

A variante "thinning" da esqueletização é baseada na erosão condicionada (menos exigente computacionalmente)

41. Qual das seguintes afirmações está correta em relação ao filtro da mediana?

- a) É um filtro não linear
- b) É a média dos pixels da sua vizinhança
- c) É um filtro linear
- d) Gera novos valores

N.B. Embora se use a expressão "Filtro" quando se refere a operação de mediana, não se trata de um verdadeiro filtro de convolução como os anteriores — a operação pertence à categoria dos filtros ditos não lineares.

42. O fator de forma de um círculo tem o valor de:

- a) 1  
 b) 0.9  
 c) 705  
 d) 605

$$ff = \frac{4\pi A}{P^2} \Rightarrow = 1 \text{ para círculos;}$$

43. O método de *Chow & Kaneko* é distinguido por:

- a) Dividir a imagem e usar histogramas locais  
 b) Dividir os seus histogramas em duas regiões dominantes  
 c) Calcular distribuições estatísticas de intensidades de pixels para cada região  
 d) Determinar no ponto do histograma o ponto de maior frequência

Limiar dinâmico ou adaptativo  
 • Método de *Chow & Kaneko*  
 • Divide a imagem em regiões e usa histogramas parciais  
 • Médias locais (mean-C)  
 • Cálculo para cada pixel baseado na sua vizinhança e num limiar

44. No contexto das transformações no plano, a matriz T tem simetria axial tem a forma:

- a)  $T = [-1 \ 0; 0 \ 1]$   
 b)  $T = [1 \ 0; 0 \ 1]$   
 c)  $T = [1 \ 0; 0 \ -1]$   
 d)  $T = [-1 \ 0; 0 \ -1]$

Casos particulares – simetria axial

$$T = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

45. O sensor PSD é um detetor analógico baseado em duas correntes geradas e que dependem da posição de incidência da luz. Se a incidência for no centro, as correntes medidas são:

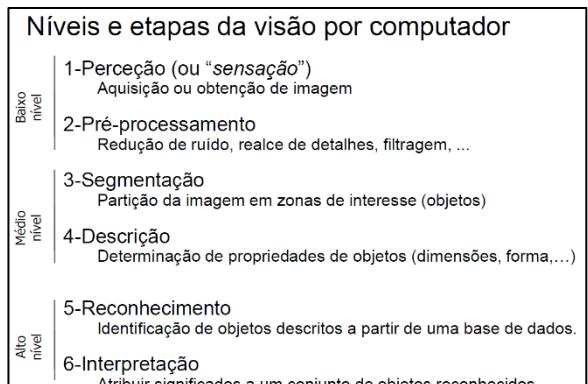
- a) Iguais  
 b) Diferentes  
 c) Assimétricas  
 d) Simétricas

46. Os descritores de imagem são frequentemente utilizados para identificação de objetos, em que consiste o “fator de forma” de um determinado objeto?

- a) Quociente que envolve a área e o perímetro do objeto  
 b) Um descritor utilizado para diferenciar objetos em função do seu diâmetro equivalente  
 c) Relação entre área da elipse maior envolvente do objeto e a sua área superficial  
 d) Descritor obtido unicamente em função do perímetro do objeto

47. Segmentação e descrição são áreas da visão por computador:

- a) Médio nível  
 b) Baixo nível  
 c) Baixo nível e médio, respetivamente  
 d) Nenhuma das anteriores



48. Em que condições um filtro de média e um de mediana de 3x3 são o mesmo resultado para um dado pixel?

- e) Nunca poderão dar o mesmo valor porque um filtro é linear e o outro não  
 f) Se os 9 pixels da vizinhança tiverem todos o valor de 0.25

- g) Em qualquer situação desde que o valor do pixel central seja o 5º mais elevado por 9 pixels
- h) Sempre que o pixel central tiver um valor 8 vezes maior do que qualquer dos vizinhos
- i) Se os pixels de  $N_4$  forem iguais aos de  $N_D$ , mas diferentes do pixel central
49. Qual das seguintes afirmações é verdadeira relativamente à operação  $C = \text{?}$

- e) Tem o mesmo efeito que  $C = C = A \oplus (B \ominus B)$ .
- f) Elimina pontos isolados pretos e brancos numa imagem
- g) A operação pode ser repetida indefinidamente sem se alterar o resultado, qualquer que seja B**

- h) Tem o mesmo efeito que  $C = C = A \ominus (B \oplus B)$ .**
- i) Elimina pontos pretos isolados no meio de objetos brancos

50. Em comparação dos modelos HSV com RGB, qual das seguintes afirmações é verdadeira?
- e) Se uma das componentes em RGB for nula e as outras forem máximas, em HSV isso representa um cinzento
- f) Para uma dada cor, se os valores de RGB forem todos iguais, a saturação é máxima
- g) Nenhuma das restantes respostas está correta
- h) Há cores em HSV que têm mais de uma correspondência em RGB
- i) Para valores RGB de 0 a 255 e valores HSV no intervalo [0; +1], tem-se o seguinte:  $\text{RGB}(255,0,0) = \text{HSV}(0,1,1)$**

51. Qual o significado de  $T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  ?

**e) Rotação de -90º**

f) Simetria axial inversa

g) Rotação de +90º

h) Simetria central

i) Translação de uma unidade em z

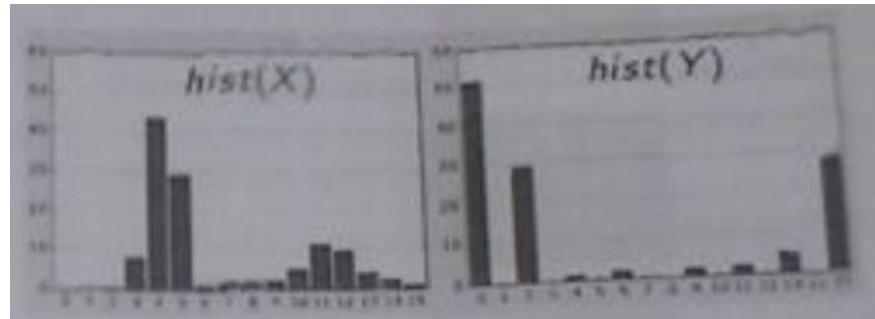
$$T_2 = \text{Rot}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

52. Se  $T_1$  e  $T_2$  forem duas transformações geométricas genéricas no plano, qual das seguintes opções é verdadeira?

- e)  $T_1$  e  $T_2$  são sempre comutativas para simetrias axiais e rotações
- f)  $T_1$  e  $T_2$  são sempre comutativas para rotações e translações
- g)  $T_1 \cdot T_2 \neq T_2 \cdot T_1$  se ambas forem translações
- h)  $T_1$  e  $T_2$  são sempre comutativas para simetrias axiais e translações
- i)  $T_1 \cdot T_2 = T_2 \cdot T_1$  se ambas forem rotações**

53. Uma imagem X foi transformada noutra imagem Y, usando  $Y = 15 \cdot \frac{X - \min(X)}{\max(X) - \min(X)}$ , cujos histogramas se ilustram. Quantos níveis de cinzento da imagem X foram passados para valores

extremos (preto ou branco)?

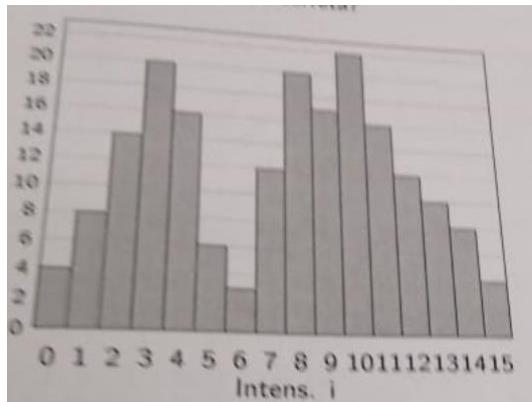


- e) 6 à esquerda e 6 à direita do histograma
- f) 2 à esquerda e 2 à direita do histograma
- g) 8 à esquerda e 4 à direita do histograma
- h) 2 à esquerda e 8 à direita do histograma
- i) 4 à esquerda e 4 à direita do histograma

54. Sabendo que  $S = \begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{matrix}$  e  $S_1 = \begin{matrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{matrix}$ , qual das expressões indicadas NÃO é verdadeira?

- e)  $D = (C \circ S) \setminus (A \otimes S_1)$
- f)  $D = (C \circ S) \setminus (C \ominus S)$
- g)  $D = A \cap [C \cap (A \circ S)]$
- h)  $D = A \setminus [C \setminus (C \ominus S)]$  é esta!
- i)  $D = A \cap [(A \otimes S_1) \oplus |_c S]$

55. Para uma imagem A com o seguinte histograma, qual das afirmações é correta?



- e) A imagem A tem 7 pixels com nível de intensidade 12
- f) A imagem A pode ser quadrada (largura e altura iguais)
- g) Pelo algoritmo *isodata*, se  $T_0=6$  então  $T_1=10$
- h) A intensidade média da imagem A é menor do que 8
- i) O histograma normalizado terá dois níveis de intensidade com valor de 0.2

56. Sejam os seguintes descritores: f=fator de forma, e=excentricidade, p=perímetro, a=área, s=solidez, n=número de Euler. Pretende-se criar um descritor composto D que permita

distinguir objetos diferentes e que seja invariante à escala, à translação e à rotação. Qual das propostas satisfaz esses requisitos em mais casos?

e)  $D = f * (s + 1) + e$

f)  $D = n * \frac{f}{e} * a$

g)  $D = n * (f + s + e)$

h)  $D = \frac{a}{p} + f + n$

i)  $D = f + s + a + n$

57. À imagem da esquerda (X) é aplicado um conjunto de operações por forma a obter a imagem

da direita (Y). Considerar os seguintes elementos estruturantes:  $N4 = \begin{matrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{matrix}$ ,  $ND = \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{matrix}$ , e  $NB = N4 \cup ND$ . Indicar qual das seguintes opções é verdadeira.



e)  $\underline{Y = (X \oplus N4) \oplus N4}$  é esta!

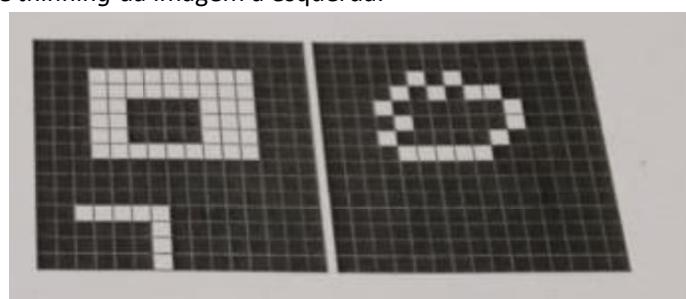
f)  $\underline{Y = (X \oplus N8) \oplus N8}$

g)  $\underline{Y = (X \oplus N8) \oplus N4}$

h)  $\underline{Y = (X \oplus ND) \oplus ND}$

i)  $\underline{Y = (X \oplus N4) \oplus ND}$

58. A variante *thinning* da esqueletização é baseada na erosão condicionada por uma ou mais das seguintes condições, que indicam que um pixel não deve ser erodido se: o pixel foi isolado (C1); a remoção do pixel quebrar a conectividade (C2); a remoção do pixel afetar o comprimento da linha (C3). Indicar quais as condições usadas de forma a obter a imagem à direita a partir de uma operação de *thinning* da imagem à esquerda.



- e) C2
  - f) C1 e C3
  - g) C1
  - h) C2 e C3
  - i) C1 e C2 e C3
59. Qual dos seguintes parâmetros não é levado em conta no dimensionamento de um sistema de visão industrial?
- e) Distância de trabalho
  - f) Velocidade do obturador
  - g) Velocidade do autofocus
  - h) Profundidade de campo
  - i) Campo ou área de visão
60. Relativamente a uma *Smart Camera* industrial, qual das seguintes afirmações é verdadeira?
- e) A câmara que usa um sensor laser 2D para se focar
  - f) A câmara que não precisa de ser calibrada intrínseca ou extrinsecamente
  - g) O programa de visão artificial executa sempre dentro da própria câmara
  - h) Há sempre iluminação integrada na própria câmara
  - i) A câmara que usa um sensor *Polaroid* de ultrassom para se calibrar
61. Nas operações morfológicas em que A e B são imagens binárias, qual das seguintes igualdades é verdade?
- e)  $B \cup A = \max(A, \min(A, B))$
  - f)  $B \cap A = \min(A, \max(A, B))$
  - g)  $B \cap A = \max(A, \max(B, A))$
  - h)  $B \cup A = \min(B, \min(A, B))$
  - i)  $B \cap A = \min(A, \min(A, B))$
62. Nos algoritmos de binarização de imagens quando é que o método de Otsu é matematicamente aplicado e o do *Isodata* não?
- e) Se o histograma foi multimodal
  - f) Se a imagem for a componente de ... representação HSV
  - g) Se a representação dos valores dos ... diferente de 8 bits/pixel
  - h) Não há nenhuma situação onde se ... um e não se possa aplicar...
  - i) Em imagens já binarizadas

1-A imagem da esquerda X é aplicado um conjunto de operações por forma a obter a imagem da direita Y. Considerar os seguintes elementos estruturais

primeira operação a aplicar é a operação estruturante N4 =  $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$  e ND =  $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$  e N8 = N4UND  
Indicar a opção verdadeira.

A)  $Y = (X \oplus N4) \ominus N4$

B)  $Y = (X \oplus N8) \oplus N8$

C)  $Y = (X \oplus N8) \ominus N4$

D)  $Y = (X \oplus ND) \oplus ND$

E)  $Y = (X \oplus N4) \oplus ND$

2- Qual das seguintes afirmações é verdadeira relativamente à operação

Operação  $C = A \ominus B$

- A) Tem o mesmo efeito que  $C = A \ominus (B \ominus B)$  WTF
- B) A operação pode ser repetida indefinidamente sem se alterar o resultado para qualquer B
- C) Elimina pontos isolados pretos e brancos numa imagem WTF
- D) Elimina pontos pretos isolados no meio de objetos brancos Falsa
- E) Tem o mesmo efeito que  $C = A \oplus (B \ominus B)$  WTF

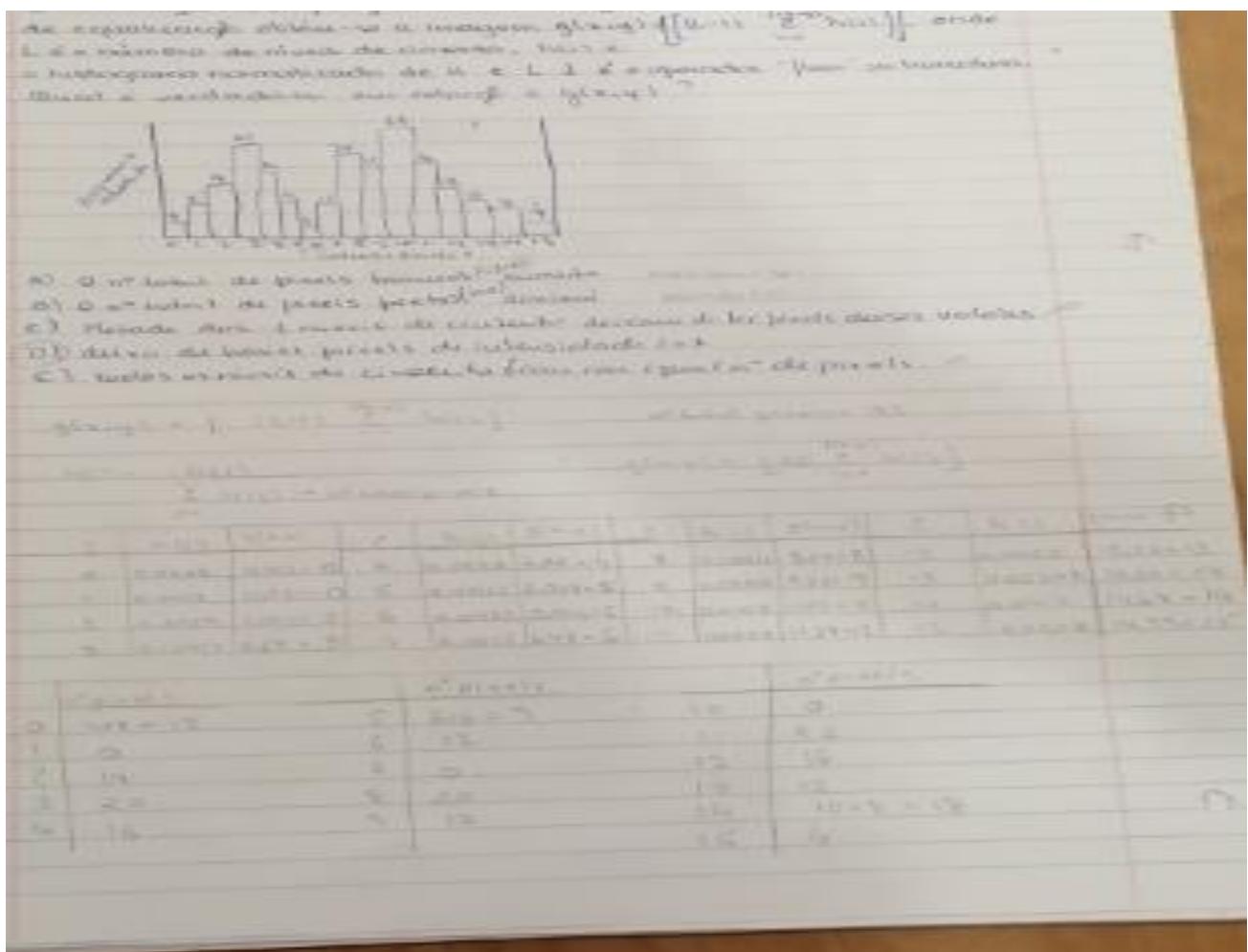
3-Quais condições C1,C2 ou C3?

Q) Quais condições C1, C2 ou C3?

C1 - só apaga pixels  
C2 - só apaga pixels  
isolados brancos  
C3 - só apaga pixels  
e pixels pertencentes  
à borda.

C2

4- A imagem  $f(x,y)$  tem o histograma  $H$ . Após a operação de equalização obtém-se a imagem  $g(x,y)$



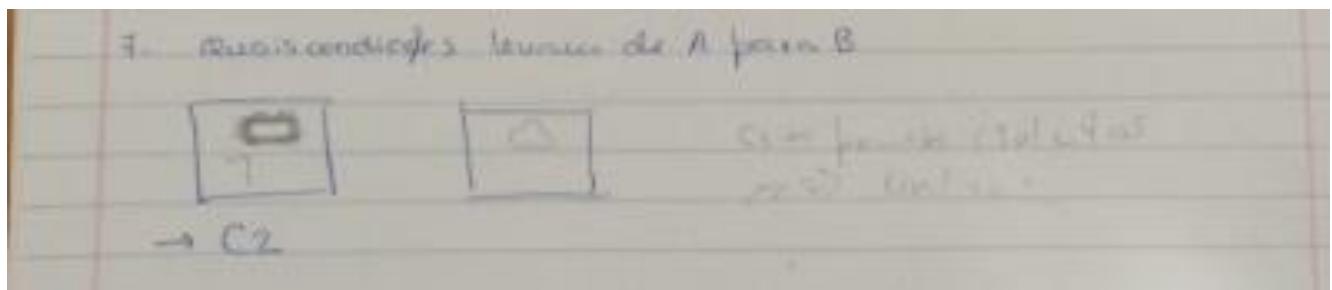
5- Para uma imagem A com o histograma anterior, qual a correta?

- A) A imagem A tem 3 pixels com nível de intensidade 12
- B) A imagem A pode ser quadrada (largura e altura iguais)
- C) Pelo algoritmo isodata se  $T_2 = 6$  então  $T_1 = 10$
- D) A intensidade média da imagem A é menor que 8
- E) O histograma normalizado tem 2 níveis de intensidade com valor 0,2

## 6 Distancia de Mahalandais

$\Sigma A = \begin{bmatrix} 4.4715 & 5.9830 \\ 5.9830 & 8.0990 \end{bmatrix} \times 10^{-4}$	$\Sigma B = \begin{bmatrix} 5.7641 & 11.3428 \\ 11.3428 & 22.3621 \end{bmatrix} \times 10^{-4}$																																																		
<table border="1"> <thead> <tr> <th></th> <th>A</th> <th>A</th> <th>B</th> <th>B</th> <th>objetos desconhecidos</th> </tr> <tr> <th></th> <th>d<sub>1</sub></th> <th>d<sub>2</sub></th> <th>d<sub>1</sub></th> <th>d<sub>2</sub></th> <th></th> </tr> </thead> <tbody> <tr> <td>Amostras</td> <td>0.1613</td> <td>0.5007</td> <td>0.1936</td> <td>0.4836</td> <td>Num</td> </tr> <tr> <td></td> <td>0.1614</td> <td>0.5084</td> <td>0.1965</td> <td>0.4561</td> <td>1</td> <td>0.1412</td> <td>0.4415</td> </tr> <tr> <td></td> <td>0.2046</td> <td>0.5613</td> <td>0.2296</td> <td>0.5680</td> <td>2</td> <td>0.1706</td> <td>0.4687</td> </tr> <tr> <td></td> <td>0.1639</td> <td>0.5051</td> <td>0.1871</td> <td>0.4927</td> <td>3</td> <td>0.1241</td> <td>0.4828</td> </tr> <tr> <td>média</td> <td>0.11925</td> <td>0.519215</td> <td>0.1942</td> <td>0.500125</td> <td></td> <td></td> <td></td> </tr> </tbody> </table>		A	A	B	B	objetos desconhecidos		d <sub>1</sub>	d <sub>2</sub>	d <sub>1</sub>	d <sub>2</sub>		Amostras	0.1613	0.5007	0.1936	0.4836	Num		0.1614	0.5084	0.1965	0.4561	1	0.1412	0.4415		0.2046	0.5613	0.2296	0.5680	2	0.1706	0.4687		0.1639	0.5051	0.1871	0.4927	3	0.1241	0.4828	média	0.11925	0.519215	0.1942	0.500125				
	A	A	B	B	objetos desconhecidos																																														
	d <sub>1</sub>	d <sub>2</sub>	d <sub>1</sub>	d <sub>2</sub>																																															
Amostras	0.1613	0.5007	0.1936	0.4836	Num																																														
	0.1614	0.5084	0.1965	0.4561	1	0.1412	0.4415																																												
	0.2046	0.5613	0.2296	0.5680	2	0.1706	0.4687																																												
	0.1639	0.5051	0.1871	0.4927	3	0.1241	0.4828																																												
média	0.11925	0.519215	0.1942	0.500125																																															

## 7- Quais condições levam de A para B



## 8-Seja imagem A e B

8 Seja imagem A e B =  $\begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix}$   $x = A \setminus (A \circ B)$

→ Top Mat

## 10- Os parâmetros extinsecos de uma câmera:

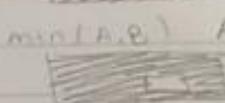
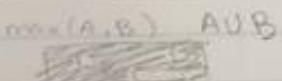
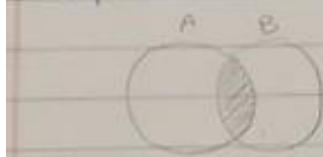
- a) representam-se numa matriz que em geral não tem inversa
- b) incluem a distância focal e o fator de skewness
- c) são iguais aos parâmetros intrínsecos se a câmera estiver estatica
- d) variam se varia o dot-pitch do CCD
- e) variam se a câmera estiver em movimento no referencial global

11- Relativamente a uma smart câmara industrial, qual é verdadeira?

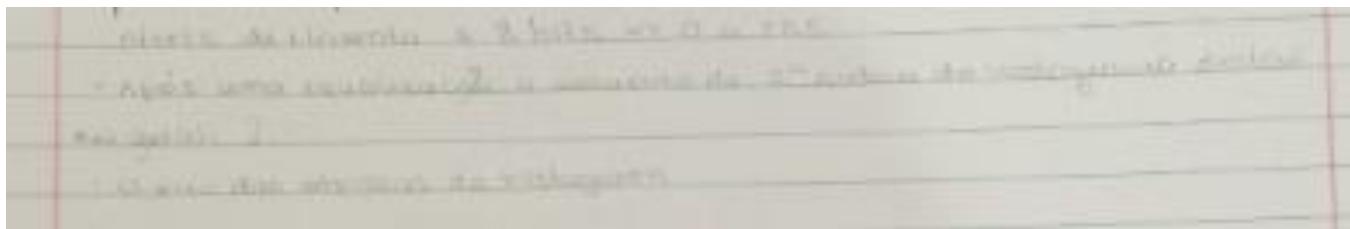
- a) o programa de visão artificial executa sempre dentro da própria câmara
- b) a câmara que usa um sensor laser 2D para se focar
- c) a câmara que usa um sensor polaroid de ultrassom para se calibrar
- d) a câmara que não precisa de ser calibrada intrinsecamente ou extrinsecamente
- e) há sempre iluminação integrada na própria câmara ✓

12- Nas operações morfológicas em que A e B são imagens binárias qual é verdadeira?

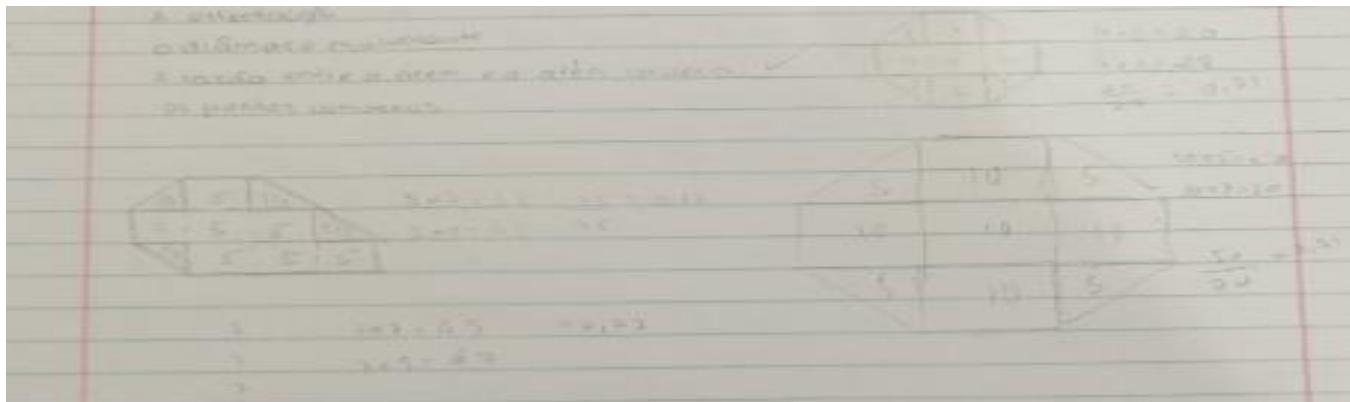
- a)  $B \cap A = \max(A, \max(B, A)) = A \cap B$
- b)  $A \cup B = \min(B, \min(A, B)) = A \cup B$
- c)  $B \cup A = \max(A, \min(A, B)) = A$
- d)  $A \cap B = \min(A, \min(A, B)) = A \cap B$
- e)  $B \cap A = \min(A, \max(A, B)) = A$



13- Se os pixéis de uma imagem em níveis de cinzento forem representados por valores de 8 bits, qual é verdadeira?



14- Qual das seguintes propriedades de uma região numa imagem binária é considerada à escala e rotação



15 O que é verdade sobre o fecho

11. O que é verdade sobre o Rácio?

Unicos pontos coincidentes pertencentes a figura

Somente objeto para exterior da corrente

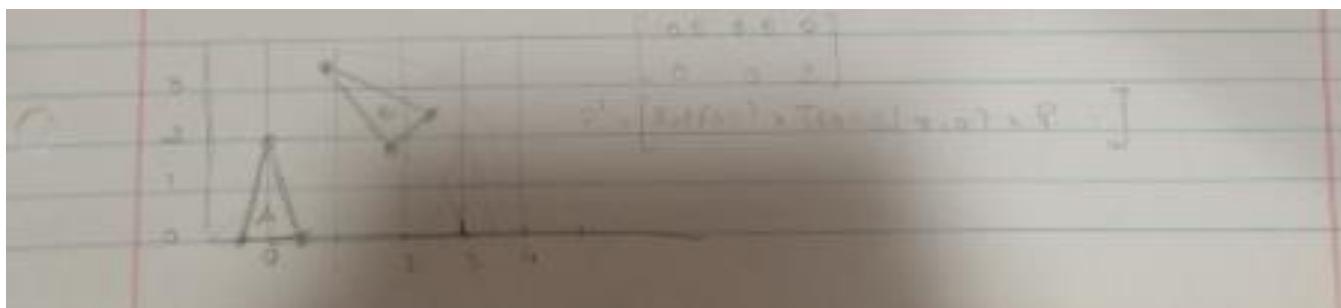
é uma operação identificativa

$$C = A \cdot R = (A \cup B) \cap R \rightarrow \text{Diâmetro-Função}$$

12 O que é o fator de forma?

$$ff = \frac{4\pi A}{p^2}$$

13 Quais as transformações que o polígono sofreu?



O significado de T1=

a) rotação de  $90^\circ$  no eixo x

-1	0	1
-1	0	1
-1	0	1

b) reflexão -  $90^\circ$

c) simetria axial horizontal

d) rotação de  $+90^\circ$

e) simetria radial

f) translação de uma unidade para a

Se T1 e T2 forem duas transformações geométricas no plano, qual das opções é verdadeira?

20. Se T1 e T2 forem duas transformações geométricas no plano, qual das opções é verdadeira?

a) T1 e T2 são sempre comutativas para simetrias axiais e rotacões

b) T1 e T2 são sempre comutativas para rotacões e translações

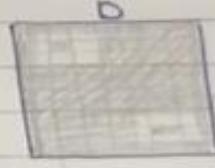
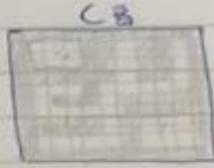
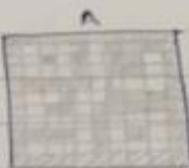
c) T1T2 ≠ T2T1 se ambas forem translações

d) T1 e T2 são sempre comutativas para simetrias axiais e translações

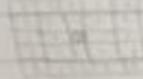
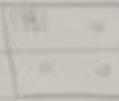
e) T1T2 = T2T1 se ambas forem rotacões

Sabendo que

2.2. Sabendo que  $S = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$  e que  $S_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ , qual das expressões é falsa?



- a)  $D = (\bar{C} \otimes S) \setminus (A \otimes S_1)$   
b)  $D = (\bar{C} \otimes S) \setminus (\bar{C} \otimes S)$   
c)  $D = A \cap [C \cap (\bar{A} \otimes S)]$   
d)  $D = A \setminus [C \setminus (\bar{C} \otimes S)]$   
e)  $D = A \cap [(A \otimes S_1) \oplus \bar{C} \otimes S]$



Qual das afirmações é verdadeira?

16. Qual das afirmações é verdadeira?

- a) Se uma das componentes em RGB for nula e as outras forem máximas em HSV, isso representa um cíntento.
- b) Para uma dada cor, se os valores RGB forem todos iguais, a saturação é máxima.
- c) Nenhuma é correta.
- d) Há cores em HSV que têm mais de uma correspondência em RGB.
- e) Para valores RGB de 0 a 255 e valores HSV no intervalo  $[0, 1]$  tem-se o seguinte:  $RGB(255, 0, 0) = HSV(0, 1, 1)$ . Verdadeiro!

Em que condições um filtro média e um mediana  $3 \times 3$  dão o mesmo resultado para um dado pixel

17. Em que condições um filtro de média e um de mediana  $3 \times 3$  dão o mesmo resultado para um dado pixel?
- a) Nunca poderão dar o mesmo valor porque um filtro é linear e o outro não.
  - b) Se os 9 pixels da vizinhança tiverem todos o valor de 0,25  $\pm$  0,05.
  - c) Em qualquer situação desde que o valor do pixel central seja o 5º mais elevado dos 9 pixels.
  - d) Sempre que o pixel central tiver um valor 3 vezes maior do que qualquer dos vizinhos.
  - e) Se os pixels de N4 forem iguais aos de N0 mas diferentes do pixel central.

Qual é a verdadeira relativamente a

18. Qual é verdadeira, relativamente a  $C = A \oplus B$ ? Abertura = Erosão  $\rightarrow (A \oplus B) \ominus B$
- a) Tem o mesmo efeito que  $C = A \ominus (B \ominus B)$  LOL
  - b) Elimina pontos isolados pretos e brancos numa imagem.
  - c) A operação pode ser repetida indefinidamente sem se alterar o resultado qualquer que seja B YES! Abertura e Fecho são idempotentes.
  - d) Tem o mesmo efeito que  $C = A \ominus (B \oplus B)$  LOL
  - e) Elimina pontos pretos isolados no meio de objetos brancos  
Fecho  $\Rightarrow$  elimina pontos pretos em objeto branco  
Abertura  $\Rightarrow$  elimina pontos brancos isolados no fundo preto

## QUIZ

Em reconhecimento de imagem qual a relação entre modelos e padrões?

Conforme questão A15Q1 da aula

Selecione uma opção de resposta:

- a. Modelos e padrões são sinônimos da mesma entidade.
- b. O modelo é a imagem onde se procura e o padrão é a imagem de referência.
- c. Modelos são descritores e padrões são imagens.
- d. Modelos são imagens e padrões são descritores.

A resposta correta é: Modelos são imagens e padrões são descritores.

Em que circunstância não é possível usar a distância de Mahalanobis para reconhecimento de padrões?

Conforme questão A16Q1 da aula

Selecione uma opção de resposta:

- a. Se o padrão tiver 5 elementos e houver 10 amostras distintas para o definir.
- b. É sempre possível de a usar para reconhecimento de padrões.
- c. Se as amostras para definir um padrão de referência forem todas iguais.
- d. Se a matriz de covariâncias for diagonal e não singular.

A resposta correta é: Se as amostras para definir um padrão de referência forem todas iguais.

Na classificação por distância entre padrões qual das afirmações é verdadeira?

Conforme questão A16Q2 da aula

Selecione uma opção de resposta:

- a. Um padrão pode não ser classificável se houver classes por definir.
- b. Quando maior a distância entre dois padrões, maior é a semelhança entre eles.
- c. Os descritores de um padrão tem de ser necessariamente invariantes às dimensões e à orientação.
- d. Um padrão com 5 descritores tem sempre mais robustez do que um padrão com 3 descritores.

A resposta correta é: Um padrão pode não ser classificável se houver classes por definir.

A que modelo de representação de cor mais se assemelha o sistema humano de percepção primária da cor?

---

Conforme questão A17Q1 da aula

Selecione uma opção de resposta:

- a. HSV
  - b. YUV
  - c. CMY
  - d. RGB ✓
- 

A resposta correta é: RGB

Porque é que o sistema subtrativo de cores tem 4 cores em vez de 3?

Conforme questão A17Q2 da aula

Selecione uma opção de resposta:

- a. A cor K permite economizar o consumo das outras 3 cores. ✓
  - b. Não há nenhum sistema com 4 cores.
  - c. Porque 3 cores primárias não chegam para representar todas as cores
  - d. Porque em vídeo digital é necessário ter uma cor para o estado de desligado.
- 

A resposta correta é: A cor K permite economizar o consumo das outras 3 cores.

Quais das indicadas são células sensíveis à luz no olho humano?

Conforme questão A17Q3 da aula

Selecione uma opção de resposta:

- a. Cones dos tipos SML ✓
  - b. Reticulos
  - c. Cones dos tipos HGM
  - d. Besoneiros e cilíndrios
- 

A resposta correta é: Cones dos tipos SML.

Na conversão entre os modelos de cor RGB e HSV, qual das afirmações é verdadeira?

---

Conforme questão A18Q1 da aula

Selecione uma opção de resposta:

- a. Uma cor com saturação máxima em HSV nunca terá nenhuma componente RGB nula. ✗
  - b. É possível ter diferentes representações em HSV que equivalem à mesma representação em RGB.
  - c. É possível ter diferentes representações em RGB que equivalem à mesma representação em HSV.
  - d. Um valor de cinzento em HSV equivale a R=0, G=127 e B=255.
- 

A resposta correta é: É possível ter diferentes representações em HSV que equivalem à mesma representação em RGB.

Na conversão de modelos de cor, o que significa a expressão  
 $Y=0.3R+0.59G+0.11B$ ?

Conforme questão A18Q2 da aula

Selecione uma opção de resposta:

- a. Resultado de um dos termos da conversão de YCbCr para RGB.
  - b. Significa que Y terá sempre o valor máximo de brilho.
  - c. Componente de crominância do sistema NTSC.
  - d. Cálculo do valor de luminância. ✓
- 

A resposta correta é: Cálculo do valor de luminância.

Numa imagem em representação HSV, qual das expressões melhor representa uma máscara das regiões vermelhas dessa imagem?

Conforme questão A18Q3 da aula

Selecione uma opção de resposta:

- a.  $V < 0.1 \& H > 0.5$
  - b.  $S > 0.8 \& H > 0.1$  ✗
  - c.  $H < 0.0 \& S > 0.5$
  - d.  $H < 0.1 \& S > 0.9$
- 

A resposta correta é:  $H < 0.1 \& S > 0.9$

Se numa imagem binária tivermos apenas 4 pixels brancos, qual das afirmações é verdadeira?

Conforme questão A19Q1 da aula

Selecionar uma opção de resposta:

- a. Se apenas 3 dos 4 pontos forem colineares, só será possível definir 3 retas distintas.
- b. Para avaliar quantos ternos de pontos são colineares podem ser precisos  $4 \times 4 = 16$  testes.
- c. O número total de retas possíveis de definir estará entre 1 e 6. ✓
- d. Como a imagem é discreta, nunca será possível ter-se uma única reta.

A resposta correta é: O número total de retas possíveis de definir estará entre 1 e 6.

No espaço de parâmetros p e teta o que significam estas variáveis?

Conforme questão A19Q2 da aula

Selecionar uma opção de resposta:

- a. Afastamento da origem e ângulo da reta com o eixo dos xx. ✗
- b. Comprimento do segmento de reta e o simétrico do seu ângulo com o eixo dos yy.
- c. Distância da reta à origem e ângulo do pé da perpendicular.
- d. Duas grandezas cujo produto é igual a ab no espaço de parâmetros associado.

A resposta correta é: Distância da reta à origem e ângulo do pé da perpendicular.

Numa imagem preta com apenas dois pixels brancos em circunstância as duas curvas da respectiva transformada de Hough se intersetam em  $\theta = -90^\circ$

Conforme questão A19Q3 da aula

Selecionar uma opção de resposta:

- a. Se os dois pixels brancos estiverem ao longo da diagonal que passa no canto superior direito da imagem.
- b. Se os dois pixels brancos estiverem na primeira coluna da imagem.
- c. Se os dois pixels brancos estiverem na primeira linha da imagem. ✓
- d. Se os dois pixels brancos estiverem ao longo da diagonal que passa no canto superior esquerdo da imagem.

A resposta correta é: Se os dois pixels brancos estiverem na primeira linha da imagem.

Na transformada de Hough para linhas retas, o que é um acumulador?

Conforme questão A20Q1 da aula

Selecione uma opção de resposta:

- a. Uma célula usada para contar o número de pontos por onde passa uma dada reta.
- b. Um contador que acumula o número de interseções de retas.
- c. Um registo que conta o número de retas que passa num dado ponto. ✗
- d. Um somatório acumulado do número de segmentos de uma imagem de arestas.

A resposta correta é: Uma célula usada para contar o número de pontos por onde passa uma dada reta.

Para poder atualizar o valor de um acumulador na transformada de Hough para retas, o que é necessário fazer?

Conforme questão A20Q2 da aula

Selecione uma opção de resposta:

- a. Encontrar a correspondência entre o espaço de parâmetros ab e o espaço xy.
- b. Para cada novo ponto, ajustar a resolução no espaço de parâmetros. ✗
- c. Normalizar as distâncias entre pontos no espaço xy.
- d. Utilizar a equação polar da reta.

A resposta correta é: Utilizar a equação polar da reta.

Se uma transformada de Hough para retas de uma imagem tiver 10 picos com mais de 100 pontos, o que é verdade?

Conforme questão A20Q3 da aula

Selecione uma opção de resposta:

- a. Não é possível ter 10 picos a menos que a resolução angular do acumulador seja de 1 grau.
- b. Existem pelo menos 10 retas potenciais na imagem com 100 pontos cada.
- c. Existem pelo menos 100 retas potenciais na imagem com 10 pontos cada.
- d. Há pelo menos 1000 pontos colineares dentro das tolerâncias da resolução dos acumuladores.

A resposta correta é: Existem pelo menos 10 retas potenciais na imagem com 100 pontos cada.

De qual das seguintes propriedades não goza a dilatação morfológica?

Conforme questão A22Q1 da aula

Selecione uma opção de resposta:

- a. Associatividade
- b. Comutatividade
- c. Idempotência ✓
- d. Distributividade em relação à união

A resposta correta é: Idempotência

O que resulta da seguinte operação  $C = A \ominus (A \oplus B)$  onde  $B$  é matriz  $3 \times 3$  só com 1s?

Conforme questão A22Q2 da aula

Selecione uma opção de resposta:

- a. Uma imagem toda a 1s
- b. Um contorno ✗
- c. A interseção entre A e B
- d. Conjunto vazio

A resposta correta é: Conjunto vazio

Para fazer a abertura de uma imagem A com um elemento estruturante B com 10 pontos quantas operações de translação morfológica sobre A seriam presisas?

Conforme questão A22Q3 da aula

Selecione uma opção de resposta:

- a. 2 operações
- b. 10 operações ✗
- c. 20 operações
- d. 4 operações

A resposta correta é: 20 operações

Qual é a operação base para realizar uma reconstrução como faz por exemplo a função mreconstruct em Matlab?

Conforme questão A23Q1 da aula

Selecione uma opção de resposta:

- a. Esqueletização e abertura
- b. Fecho repetido
- c. Dilatação condicionada ✓
- d. Erosão recursiva

A resposta correta é: Dilatação condicionada

Das operações indicadas, quais são necessárias para fazer uma abertura em níveis de cinzento?

Conforme questão A23Q2 da aula

Selecione uma opção de resposta:

- a. max() e min() ✓
- b. sum() e and()
- c. and() e or()
- d. Translações morfológicas

A resposta correta é: max() e min()

Qual das seguintes é uma característica da operação TopHat ?

Conforme questão A23Q3 da aula

Selecione uma opção de resposta:

- a. Só se deve aplicar na parte superior de uma imagem.
- b. Perde toda a eficácia se a iluminação diminuir para metade.
- c. Obter contornos em níveis de cinzento.
- d. Destacar objetos do fundo que os rodeia. ✓

A resposta correta é: Destacar objetos do fundo que os rodeia.

Relativamente ao sensor ilustrado e à sua resposta , o que se pode afirmar ?

Conforme questão A26Q1 da aula

Selecione uma opção de resposta:

- a. A saída do sensor é feita em corrente de 4 a 24 mA.
- b. A distância mínima de medição é de 80 cm.
- c. A frequência de medição é cerca de 1/25 Hz.
- d. Existe ambiguidade de medição para algumas medidas abaixo de 10 cm.

A resposta correta é: Existe ambiguidade de medição para algumas medidas abaixo de 10 cm.

Qual a afirmação verdadeira para as imagens obtidas em simultâneo por uma câmara termográfica?

Conforme questão A26Q2 da aula

Selecione uma opção de resposta:

- a. As imagens representam os 3 canais independentes de infravermelho para formar a imagem térmica final.
- b. A imagem do centro tem mais informação do que a de níveis de cinzento, mas menos do que a da direita.
- c. As imagens com pseudo-cor têm ambas a mesma informação e que, por ser a cores, é maior do que a de níveis de cinzento.
- d. São todas aproximadamente equivalentes em termos de quantidade de informação nos pixels.

A resposta correta é: São todas aproximadamente equivalentes em termos de quantidade de informação nos pixels.

Qual das indicadas não é uma característica dos sensores ou da percepção com SONAR?

Conforme questão A26Q3 da aula

Selecione uma opção de resposta:

- a. Especularidade das reflexões.
- b. Múltiplos caminhos na medição.
- c. Interferência entre sensores próximos. ✗
- d. Diagrama de radiação uniforme.

A resposta correta é: Diagrama de radiação uniforme.

Qual das linhas é uma limitação num LiDAR baseado em tempo de vôo?

Conforme questão A27Q1 da aula

Selecione uma opção de resposta:

- a. Distância limitada pelo comprimento de onda do sinal modulado.
- b. Não opera em ambientes noturnos. ✗
- c. Especularidade em superfícies polidas.
- d. Requer alvos negros para não dispersar o feixe de retorno.

A resposta correta é: Especularidade em superfícies polidas.

No scan LiDAR ilustrado, a que se devem as medições mais longas?

Conforme questão A27Q2 da aula

Selecione uma opção de resposta:

- a. Presença de um ou mais pilares.
- b. Limite de medição do sensor. ✓
- c. Presença de uma ou mais árvores.
- d. Paredes dos edifícios.

A resposta correta é: Limite de medição do sensor.

Relativamente a figura qual das afirmações é verdadeira? (carro autônomo do dem)

Conforme questão A27Q3 da aula

Selecione uma opção de resposta:

- a. Estão presentes LiDARs 1D, 2D e 3D parcial. ✓
- b. O número total de planos de medição nos LiDARs presentes é 4.
- c. Só há percepção LiDAR no espaço à frente do veículo.
- d. Com o arranjo ilustrado, é possível medir o perfil de túneis e pontes.

A resposta correta é: Estão presentes LiDARs 1D, 2D e 3D parcial.

Qual o significado do valor do pixel numa imagem de disparidade estéreo?

Selecione uma opção de resposta:

- a. Valor do pixel mais intenso dos dois pixels obtidos nas duas câmaras.
- b. Valor da intensidade do pixel mais escuro dos dois obtidos nas duas câmaras.
- c. Diferença de profundidade dos dois pixels correspondentes nas duas câmaras.
- d. Diferença entre coordenadas de pixels correspondentes nas duas imagens. ✓

A resposta correta é: Diferença entre coordenadas de pixels correspondentes nas duas imagens.

Qual a solução mais adotada para resolver a ambiguidade na deteção de pontos num sistema de luz estruturada?

- a. Projeção de uma matriz de pontos com cores diferentes.
- b. Projeção de um padrão de linhas no ultravioleta.
- c. Projeção de uma matriz uniforme de cruzes assimétricas.
- d. Projeção de um padrão com distribuição pseudo-aleatória de pontos.

resposta correta é: Projeção de um padrão com distribuição pseudo-aleatória de pontos.

Qual das indicadas é uma afirmação VERDADEIRA em relação ao sensor ilustrado?



- a. As suas câmaras têm uma resolução igual para uma correta informação RGB-D.
- b. Possui 3 câmaras: duas de infravermelho e uma de RGB.
- c. Usa o princípio de TOF (time-of-flight) para fazer medições. ✗
- d. Precisa de iluminação externa infravermelha para fazer medições de distância.

A resposta correta é: As suas câmaras têm uma resolução igual para uma correta informação RGB-D.

A operação LoG quando aplicada a uma imagem ?

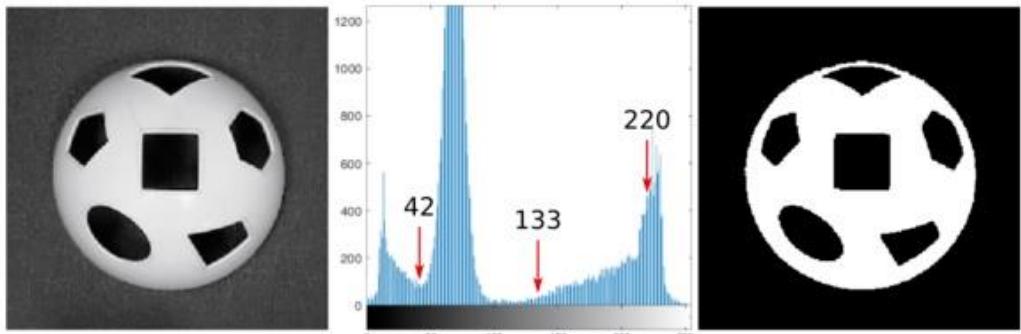
A operação LoG, quando aplicada a uma imagem, significa:

Selecione uma opção de resposta:

- a. Aplicação de um filtro gaussiano sobre um gradiente de 2<sup>a</sup> ordem dessa imagem.
- b. Cálculo de um gradiente de 2<sup>a</sup> ordem sobre uma imagem filtrada com um filtro de suavização.
- c. O mesmo que detetar pontos isolados nessa imagem.
- d. Cálculo do logaritmo do valor dos pixels da imagem.

Que limiar de binarização resulta na imagem da direita?

Que limiar de binarização resulta na imagem da direita?



Selecione uma opção de resposta:

- a. Não há nenhum possível.
- b. 220
- c. 133
- d. 42



[Limpar a minha escolha](#)

Qual das soluções indicadas permite fazer a reconstrução morfológica?

Qual das soluções indicadas permite fazer a reconstrução morfológica de uma região poligonal sem furos?

Selecione uma opção de resposta:

- a. Uma imagem com um ponto isolado e outra com uma máscara que o contém.
- b. Uma imagem com um anel e outra com um ponto isolado situado no interior do anel.
- c. Uma imagem com uma máscara e outra com um contorno sem pontos comuns.
- d. Uma imagem com um esqueleto e outra com uma semente que se intersetam.

Em relação a esqueletização

Em relação à esqueletização e "thinning" qual das afirmações é FALSA?

Selecione uma opção de resposta:

- a. Há uma variante da esqueletização que gera pontos isolados e anéis.
- b. Em certos objetos, o esqueleto pode ser igual ao próprio objeto.
- c. Há uma variante de esqueletização que gera pontos isolados.
- d. O esqueleto de um objeto tem sempre mais pixels do que o próprio objeto.

[Limpar a minha escolha](#)

Qual dos seguintes elementos não está presente na matriz instrinseca?

Qual dos seguintes elementos não está presente na matriz instrínseca de uma câmera?

Selecione uma opção de resposta:

- a. Coeficiente de distorção radial.
- b. Ponto principal.
- c. Fator de "skewing".
- d. Distância focal.

Qual dos seguintes não é um descritor de contornos?

Selecionar uma opção de resposta:

- a. Códigos de Freeman
- b. Dipum
- c. Assinaturas
- d. Descritores de Fourier

[Limpar a minha escolha](#)



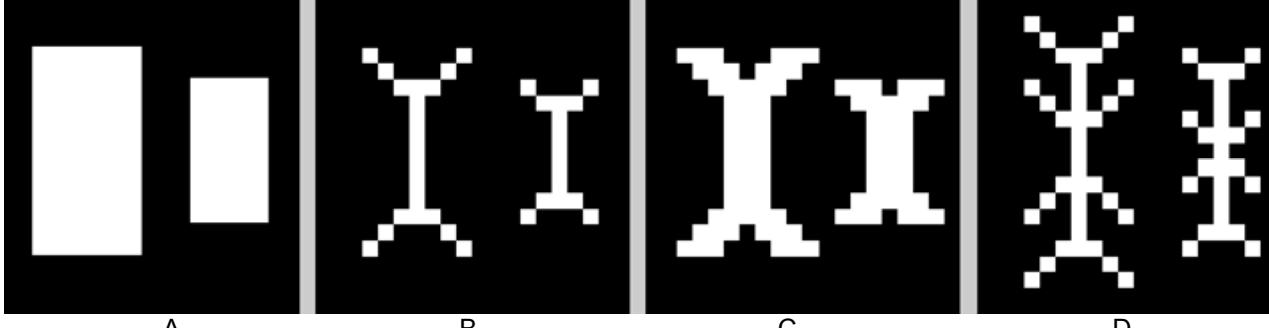
# Sistemas de Visão e Percepção Industrial

Exame de Época de Recurso

9 de Julho de 2008

Mestrado Integrado em Engenharia Mecânica  
Mestrado em Engenharia de Automação Industrial

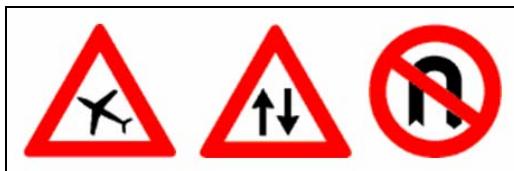
- 1- Sejam as seguintes imagens binárias de 20x20 pixels com os pontos de fundo a preto (valor 0).



A                      B                      C                      D

- A imagem B foi obtida a partir da imagem A por erosão condicionada. Indicar genericamente as condições a respeitar para efectuar (ou não efectuar) a erosão de cada pixel da imagem A para obter a imagem B.
- Seja Y a imagem que contém apenas os quatro pontos extremos de cada uma dos dois objectos de B. Indicar um filtro F de 3x3 tal que, quando aplicado à imagem B, resultando em  $Z = \text{filter2}(F, B)$ , permite obter  $Y = (Z == 7)$ .
- Obter a mesma matriz Y da alínea anterior mas agora através da operação morfológica *Hit-and-miss*. Indicar todos os elementos estruturantes necessários.
- A imagem C foi obtida pelo seguinte modo:  $C = B \oplus S_H$  onde  $S_H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ ; de modo similar, tem-se que  $D = B \oplus S_V$ . Determinar e representar  $S_V$ . Obs: A origem do elemento estruturante é o seu centro geométrico.
- Indicar e demonstrar a veracidade da seguinte afirmação: "a operação de esqueletização de uma imagem binária preserva sempre os centros de massa dos diversos objectos dessa imagem".
- Qual o resultado da operação de abertura morfológica da imagem D com o elemento estruturante de defeito [ones(3,3)]? Justificar a resposta.

- 2- Seja a imagem seguinte, T, que contém três sinais de trânsito onde a orla de cada um é de cor vermelha saturada, o fundo é branco e os símbolos internos são de cor preta. A ordem ou distribuição dos três sinais pode ser variável na imagem e diferente do representado. A linha do caixilho não faz parte da imagem, serve apenas para delimitar a área a processar.



**Recomendação:** nas respostas às questões, atender à natureza das imagens onde são aplicadas as diversas operações (RGB, níveis de cinzento, binária, etc...); operações em imagens desapropriadas podem tornar a resposta incompleta ou mesmo errada. Deve-se portanto acautelar as menções às conversões de formato de imagem sempre que necessário. A imagem original está em formato RGB. É permitido o recurso a funções de alto nível do Matlab, se for necessário.

- Quais os passos necessários para separar os três sinais (, , ) em três imagens distintas A, B e C também no formato RGB?
- Admitindo que se consegue obter as máscaras *mask1*, *mask2* e *mask3* para as três imagens, como se determina qual delas é a máscara da imagem com o sinal circular? , , .
- Quais as operações para, a partir de A, B e C, extraer de cada uma o respectivo símbolo? Ou seja, os seguintes

elementos:  $X_{1S} = \begin{bmatrix} \text{pedestrian} \\ \text{up-down} \\ \text{circle} \end{bmatrix}$ ,  $X_{2S} = \begin{bmatrix} \text{up-down} \\ \text{circle} \end{bmatrix}$ ,  $C_S = \begin{bmatrix} \text{pedestrian} \end{bmatrix}$ .

- Admitindo que se obtém as imagens dos símbolos da alínea anterior, como se podem distinguir os sinais triangulares entre si, ou seja  $X_{1S}$  de  $X_{2S}$ ? Indicar todas as operações necessárias.
- Se por alguma razão um dos sinais surgisse no bordo da imagem, que operações seriam necessárias para o eliminar, incluindo todo o seu conteúdo? Pode-se tomar o seguinte caso como exemplo: passar de



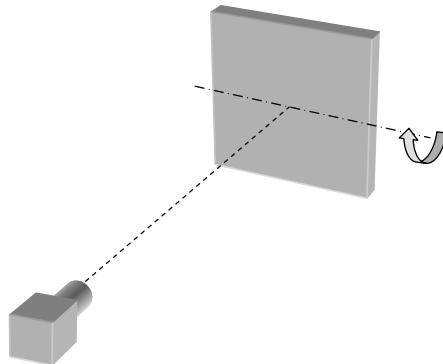
para esta:



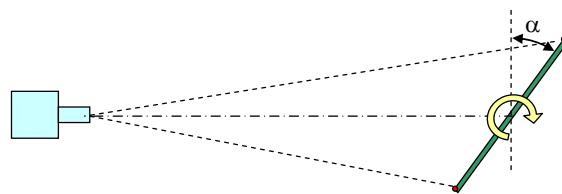
, mas não para esta:



- 3- Uma câmara de 1024x768 pixels com distância focal de 16 mm e *dot-pitch* de 160 pixels/mm (na horizontal e na vertical) está colocada a uma distância de 4 m de um painel quadrado de 50 cm de lado. De início o painel está paralelo ao plano focal mas depois reorienta-se em torno do seu eixo médio.



- a) Estabelecer a equação da transformação de perspectiva para um ponto genérico do painel. Ou seja, estabelecer as coordenadas na imagem  $Q=[u \ v \ w]$  (com  $x_{pix}=u/w$  e  $y_{pix}=v/w$ ) em função das coordenadas de um ponto genérico do painel  $P=[x_s \ y_s \ z_s]^T$ , e dos parâmetros intrínsecos da câmara. Considerar que o centro geométrico do CCD coincide com a intersecção do eixo óptico com o plano focal.
- b) A partir do factor de forma ( $ff$ ) da imagem do painel que se gera na câmara, e dos dados do problema, determinar genericamente o ângulo  $\alpha$  de inclinação do painel. Testar para  $ff=\sqrt{2}/2$ . N.B. neste contexto, o factor forma usado é o quociente da altura sobre a largura do objecto visto na imagem, e que será igual a 1 para  $\alpha=0$ ;



#### Cotação:

Questão 1 – 8 Valores. Questão 2 – 7 Valores. Questão 3 – 5 Valores.

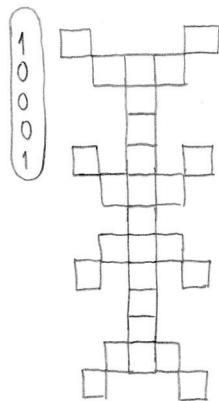
#### Breve formulário

$$\text{Matriz intrínseca da câmara: } \mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Relações trigonométricas:  $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ ,  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$

A imagem C foi obtida pelo seguinte modo:  $C = B \oplus S_H$  onde  $S_H = \boxed{1 \ 1 \ 1}$ ; de modo similar, tem-se que  $D = B \oplus S_V$ .

Determinar e representar  $S_V$ . (Obs: a origem do elemento estruturante é o seu centro geométrico).



Então temos que  $S_V =$

1
0
0
0
1

) Indicar e demonstrar a veracidade da seguinte afirmação: "a operação de esqueletização de uma imagem binária preserva sempre os centros de massa dos diversos objectos numa imagem".

A esqueletização consiste em:

- operação de determinação do esqueleto.
  - definição do esqueleto:  $\rightarrow$  objecto filiforme (1 pixel de largura).
    - $\rightarrow$  passa pelo meio do objecto.
    - $\rightarrow$  preserva a topologia do objecto original.
- } Logo a afirmação é VERDADEIRA.

) Qual o resultado da operação de abertura morfológica da imagem D com o elemento estruturante de defeito [ones(3,3)]? Justificar a resposta.

$$SE = \text{ones}(3,3);$$

$$\text{Imagen} = \text{imopen}(C, SE)$$

/. Elemento estruturante

$$/. \text{Abertura morfológica} \longrightarrow C \circ SE = (C \ominus SE) \oplus SE$$

$\downarrow$  DILATACÃO  
 $\downarrow$  EROSÃO

O resultado final será uma imagem de fundo branco, uma vez que a abertura morfológica consiste numha erosão seguida de uma dilatação. Como em D, só temos imagens com 1 pixel de largura (esqueleto), os objectos desaparecem logo na erosão.

Exercício 2 Seja a imagem seguinte, T, que contém 3 sinais de trânsito onde a cor de cada um é de cor vermelha saturada, o fundo é branco e os símbolos internos são de cor preta. A ordem ou distribuição dos três sinais pode ser variável na imagem e diferente de representado. A linha do caixilho não faz parte da imagem, serve apenas para delimitar a área a processar.

a) Quais os passos necessários para separar os três sinais em três imagens distintas A,B e C também no formato RGB?

```

>> A = imread('Sinais.png');
>> A = im2double(A);
>> B = im2bw(A);
>> B = ~B;
>> C = imfill(B, 'holes');
>> [L num] = bwlabel(C);
>> Sinal1 = A;
>> Sinal2 = A;
>> Sinal3 = A;
    
```

(...) (continuação)

```

>> Sinal1 (~repmat((L)==1, [1 1 3])) = 0;
>> Sinal2 (~repmat((L)==2, [1 1 3])) = 0;
>> Sinal3 (~repmat((L)==3, [1 1 3])) = 0
    
```

1) Se por alguma razão um dos símbolos surgisse no borda da imagem, que operações seriam necessárias para o eliminar, incluindo todo o seu conteúdo? Pode-se tomar o seguinte caso como exemplo: passar de  para esta:



, mas não para esta:



$A = \text{imread}('Sinais 2.png');$

$A = \text{im2double}(A);$

$B = \text{im2bw}(A);$

$B = \sim B;$

$C = \text{imfill}(B, 'holes');$

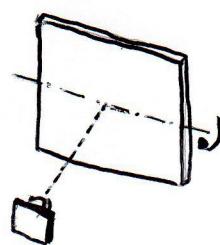
$C(1:\text{end}, 1, :) = 1;$

$C = \text{imfill}(C, 'holes');$

$D = \text{imclearborder}(C);$

$A(\sim \text{repmat}((D == 1, [1 1 3]))) = 0$

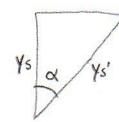
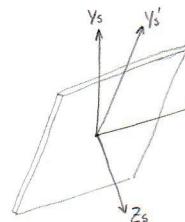
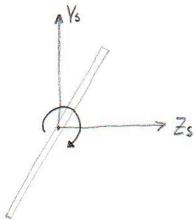
Exercício 3 Uma câmara de  $1024 \times 768$  pixels com distância focal de 16 mm e dot-pitch de 160 pixels/mm (na horizontal e na vertical) está colocada a uma distância de 4m de um painel quadrado de 50 cm de lado. De inicio o painel está paralelo ao plano focal mas depois rocionta-se em torno do seu eixo médio.



1) Estabelecer a equação da transformação de perspectiva para um ponto genérico do painel. Ou seja, estabelecer as coordenadas na imagem  $Q = [u \ v \ w]$  (com  $x_{pix} = u/w$  e  $y_{pix} = v/w$ ) em função das coordenadas de um ponto genérico do painel  $P = [x_s \ y_s \ z_s]^T$ , e dos parâmetros intrínsecos da câmara. Considerar que o centro geométrico do CCD coincide com a intersecção do eixo óptico com o plano focal.

$$\begin{cases} x_{pix} = x_0 + k_x \cdot x_i \\ y_{pix} = y_0 + k_y \cdot y_i \end{cases} \Rightarrow \begin{cases} x_{pix} = x_0 + k_x \cdot f \cdot (x_s/z_s) \\ y_{pix} = y_0 + k_y \cdot f \cdot (y_s/z_s) \end{cases} \Rightarrow \begin{cases} z_s \cdot x_{pix} = z_s \cdot x_0 + d_x \cdot x_s \\ z_s \cdot y_{pix} = z_s \cdot y_0 + d_y \cdot y_s \end{cases} \Rightarrow \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} d_x & 0 & x_0 \\ 0 & d_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$$

$$\boxed{\frac{x_s}{z_s} = \frac{x_i}{f}; \frac{y_s}{z_s} = \frac{y_i}{f}}$$



$$\Rightarrow \cos \alpha = \frac{y_s}{y'_s} \Leftrightarrow y'_s = \frac{y_s}{\cos \alpha}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} d_x & 0 & x_0 \\ 0 & d_y & y_0/\cos \alpha \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s/\cos \alpha \\ z_s \end{bmatrix} \Rightarrow \begin{bmatrix} x_{pix} \\ y_{pix} \\ 1 \end{bmatrix} = \begin{bmatrix} d_x & 0 & x_0 \\ 0 & d_y & y_0/\cos \alpha \\ 0 & 0 & z_s \end{bmatrix} \begin{bmatrix} x_s \\ y_s/\cos \alpha \\ z_s \end{bmatrix} \cdot \frac{1}{z_s}$$

2) A partir do factor de forma (ff) da imagem do painel que se gera na câmara, e dos dados do problema, determinar genericamente o ângulo  $\alpha$  de inclinação do painel. Testar para  $ff = \sqrt{2}/2$ . N.B. neste contexto, o factor de forma usado é o quociente da altura sobre a largura da objecto visto na imagem, e que será igual a 1 para  $\alpha = 0$ .

$$ff = \frac{y_{pix}}{x_{pix}} \Leftrightarrow ff = \frac{[d_y \cdot (y_s/\cos \alpha) + (y_0/\cos \alpha) \cdot z_s] / z_s}{[d_x \cdot x_s + x_0 \cdot z_s] / z_s}$$

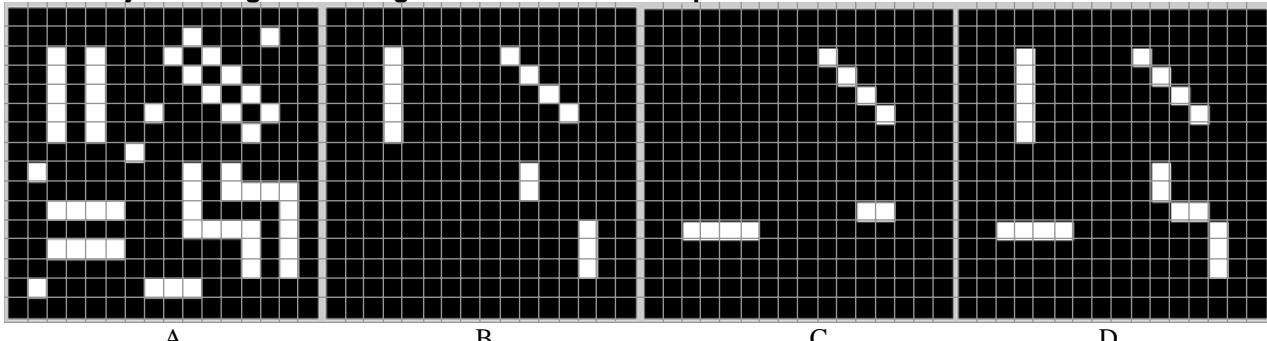


# Sistemas de Visão e Percepção Industrial

Exame de Época Normal - 30 de Junho de 2010

Mestrado Integrado em Engenharia Mecânica; Mestrado em Engenharia de Automação Industrial  
Minor em Automação da Licenciatura em Matemática  
Programa Doutoral em Informática, MAP-I

## 1. Sejam as seguintes imagens binárias de 16x16 pixels.



- Indicar um filtro F1 (3x3) que quando aplicado sobre A permita obter B, e explicar o resto do procedimento para obter B. **N.B.** Admitir uma operação em Matlab do género  $X=filter2(F1,A)$  e depois explicar como obter B a partir de X.
- Indicar um filtro F2 (que pode ser inspirado no filtro F1 da alínea anterior) para obter a imagem C, e indicar a operação para obter D com base em B e C. **N.B.** Todas as imagens representadas são binárias (0 ou 1).
- Determinar e representar a imagem que resulta da aplicação de um filtro de mediana de 3X3 na imagem A.

## 2. Considerar as operações morfológicas binárias.

- Através da definição (Cf. formulário), calcular, e representar, todos os pixels do resultado  $P=M \oplus S$ , onde M e S são como indicado a seguir. Indicam-se com ( as "origens" (ou pontos de referência) da imagem e do elemento estruturante.

$$M =$$

1		1	
		1	
1		1	
1			
1			
X			

$$S =$$

1	X	1
---	---	---

- Restringindo-se apenas às dimensões ilustradas para M e S, qual seria o resultado da operação  $Q=M \ominus S^C$ ? Justificar com a indicação do cálculo.
- Referir qual o nome pela qual é usualmente conhecida a operação “ $\otimes$ ” e indicar, justificando, qual o resultado da operação  $R=M \otimes (S1, S2)$  com:

$$S1 =$$

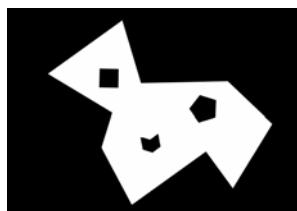
		1
	X	

$$S2 =$$

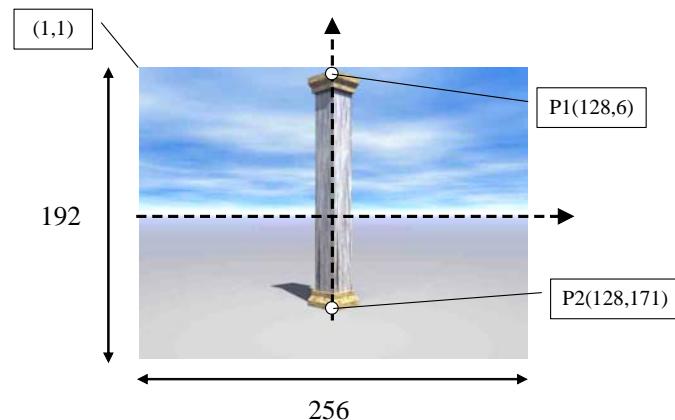
	1	
1	X	

## 3. Considerar os descritores de regiões.

- Atendendo à definição de factor de forma, que é unitário para um círculo, e igual a  $\pi/4$  para uma reigão quadrada, determinar o factor de forma teórico para uma região hexagonal regular.
- Considerar que se dispõe de uma imagem como a ilustrada a seguir, e da qual se pretende calcular o número de Euler, mas não se dispõe de uma função que o dê directamente. Baseando-se apenas em operações morfológicas e/ou de pré-processamento, indicar que operações se deveriam levar a cabo para calcular o número de Euler da imagem ( $N_E=N_O-N_B$ ). **N.B.** Se se entender útil, pode-se complementar a descrição das operações com os correspondentes em Matlab. Admitir que o objecto é feito de pixels de valor 1 e o fundo 0.



4. Seja uma imagem de 256X192 pixels, oriunda de um CCD com 150 pixels/mm e obtida com uma lente com distância focal de 6 mm. Nessa imagem está uma coluna vertical que se sabe estar paralela ao plano de imagem e que tem 2 metros de altura.



- Com base nos dados, indicar o valor numérico da matriz intrínseca da câmara.
- Sabendo que os pixels P1 e P2 indicados na imagem são os extremos da coluna, determinar a distância a que está a coluna do plano de imagem. Representar, antes de tudo, a expressão da transformação geométrica matricial que relaciona os pontos no mundo físico PW1 e PW2 com os respectivos pontos (pixels) na imagem (P1 e P2), e só depois disso concluir o cálculo da distância a que se encontra a coluna.

**Cotação:** Questão 1 – 5 Valores. Questão 2 – 6 Valores. Questão 3 – 4 Valores. Questão 4 – 5 Valores

## Breve formulário

Momentos de imagens:

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}, \quad m_{01} = \sum_x \sum_y y \cdot f(x, y), \quad m_{10} = \sum_x \sum_y x \cdot f(x, y)$$

Expressões para operar/analisar histogramas:

$$\mu_n = \sum_{i=0}^{L-1} (i - \mu_0)^n h(i), \quad \mu_0 = \sum_{i=0}^{L-1} i h(i), \quad \text{com } h(i) \text{ normalizado, i.e., } 0 \leq h(i) < 1, \forall i \in \{0, 1, 2, \dots, L-1\}$$

$$\text{Matriz intrínseca da câmara: } \mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Morfologia:

$$A_h = \{p \in \mathbf{Z}^2 : p = x + h, x \in A\}$$

$$A^c = \overline{A} = \{p \in \mathbf{Z}^2 : p \notin A\}$$

$$C = A \oplus B = \{c \in \mathbf{Z}^2 : c = a + b, a \in A \wedge b \in B\} = \bigcup_{h \in B} A_h$$

$$C = A \ominus B = \{c \in \mathbf{Z}^2 : c + b \in A, \text{ para todos } b \in B\} = \{c \in \mathbf{Z}^2 : B_c \subseteq A\} = \bigcap_{h \in B} A_{-h}$$

$$D = A \otimes (B, C) = (A \ominus B) \cap (A^c \ominus C)$$

$$\bigcup_i A \otimes (B_i, C_i) = \bigcup_i [(A \ominus B_i) \cap (A^c \ominus C_i)]$$

# SISTEMAS DE VISÃO E PERCEPÇÃO INDUSTRIAL

Exame de Época Normal - 30 de Junho de 2010

Exercício 1) Sejam as seguintes imagens binárias de  $16 \times 16$  pixels.

a) Indicar um filtro  $F1 (3 \times 3)$  que quando aplicado sobre A permite obter B, e explicar o resto do procedimento para obter B. N.B: admitir uma operação em Matlab do género  $X = filter2 (F1, A)$  e depois explicar como obter B a partir de X.

Princípio de utilização de um filtro: o novo pixel é função do pixel original e do filtro.

$$g(x,y) = \sum_{i=1}^N w_i p_i$$

→ pertence à vizinhança do ponto em causa  
→ pode ser o valor directo do novo pixel, ou ser usado de forma indireta e assim permitir calcular o seu novo valor

Os coeficientes de um filtro  $3 \times 3$  são:

$w_1$	$w_2$	$w_3$
$w_4$	$w_5$	$w_6$
$w_7$	$w_8$	$w_9$

Neste caso será do tipo:

0	0	0
2	1	2
0	0	0

»  $F1 = [0 \ 0 \ 0; 2 \ 1 \ 2; 0 \ 0 \ 0];$

»  $X = filter2(F1, A);$  % Aplicação do filtro F1 à imagem A.

»  $B = zeros (size(A));$  % Criação de uma imagem toda preta (zeros) com o tamanho de A ( $16 \times 16$ ).

»  $B (X == 4) = 1$  % O resultado da aplicação do filtro é apresentado na imagem B.

b) Indicar um filtro  $F2$  (que pode ser inspirado no filtro  $F1$  da alínea anterior) para obter a imagem C, e indicar a operação para obter D com base em B e C. N.B: Todas as imagens representadas são binárias (0 ou 1).

Os coeficientes de um filtro  $3 \times 3$  são:

$w_1$	$w_2$	$w_3$
$w_4$	$w_5$	$w_6$
$w_7$	$w_8$	$w_9$

Neste caso será do tipo:

0	2	0
0	1	0
0	2	0

»  $F2 = [0 \ 2 \ 0; 0 \ 1 \ 0; 0 \ 2 \ 0];$

»  $X = filter2(F2, A);$  % Aplicação do filtro F2 à imagem A

»  $C = zeros (size(A));$  % Criação de uma imagem toda preta (zeros) com o tamanho de A ( $16 \times 16$ )

»  $C (X == 1) = 1$  % O resultado da aplicação do filtro é apresentado na imagem C.

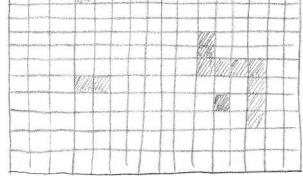
Para obter D:

»  $D = B + C$  % Representação no Matlab da união de dois conjuntos ( $B \cup C$ )

c) Determinar e representar a imagem que resulta da aplicação de um filtro de mediana  $3 \times 3$  na imagem A.

Filtro de mediana: → move os valores moves para os pixels

→ substituir cada pixel pela mediana da distribuição dos níveis de todos os pixels da sua vizinhança (incluindo o próprio).



IMPORTANTE: sempre que um pixel tem 5 ou mais pixels brancos à sua volta é afectado pelo filtro de média.

NOTA: representação com cores trocadas

### Exercício 2 | Considerem as operações morfológicas binárias.

a) Através da definição (Cf. formulário), calcular, e representar, todos os pixels do resultado  $P = M \oplus S$ , onde  $M$  e  $S$  são como indicado a seguir. Indicam-se com (as "origens" (ou pontos de referência) da imagem e do elemento estruturante.

$$M = \begin{array}{|c|c|c|c|c|} \hline & & & & \\ \hline & 1 & 1 & & \\ \hline & 1 & & & \\ \hline & 1 & 1 & & \\ \hline & 1 & & & \\ \hline & 1 & & & \\ \hline \text{X} & & & & \\ \hline 0 & 1 & 2 & 3 & 4 \\ \hline \end{array} \quad S = \boxed{1 \mid X \mid 1} \circ$$

$$P = M \oplus S = \{ p \in \mathbb{Z}^2 : c = m + s, m \in M \wedge s \in S \}$$

↳ dilatação da imagem  $M$  com o elemento estruturante  $S$

$$M = \{(1,1); (1,2); (1,3); (1,5); (2,4); (3,3); (3,5)\}$$

$$S = \{(-1,0); (1,0)\}$$

NOTA PONTOS: (coluna, linha)

Então, se  $P = M \oplus S$ :

$$P = \{(0,1); (0,2); (0,3); (0,5); (1,4); (2,3); (2,5); (2,1); (2,2); (2,3); (2,5); (3,4); (4,3); (4,5)\}$$

NOTA: pontos repetidos reduzem-se a um só ponto no conjunto final.

$$P = \begin{array}{|c|c|c|c|c|} \hline & & & & \\ \hline & 1 & 1 & 1 & \\ \hline & 1 & 1 & 1 & \\ \hline & 1 & 1 & 1 & \\ \hline & 1 & 1 & 1 & \\ \hline \text{X} & & & & \\ \hline 0 & 1 & 2 & 3 & 4 \\ \hline \end{array}$$

b) Restringindo-se apenas às dimensões ilustradas para  $M$  e  $S$ , qual seria o resultado da operação  $Q = M \ominus S^c$ ? Justificar com a indicação do cálculo.

$$Q = M \ominus S^c$$

↳ erosão da imagem  $M$  com o elemento estruturante  $S$  complementar ( $S^c = \boxed{0 \mid X \mid 0} = \boxed{\square \mid \times \mid \square}$ )

$$M = \{(1,1); (1,2); (1,3); (1,5); (2,4); (3,3); (3,5)\}$$

$$S^c = \{(0,0); (0,0)\}$$

Então se  $Q = M \ominus S^c$ :

$$Q = \{(1,1); (1,2); (1,3); (1,5); (2,4); (3,3); (3,5)\}$$

$$Q = \begin{array}{|c|c|c|c|c|} \hline & & & & \\ \hline & 1 & 1 & & \\ \hline & & 1 & & \\ \hline & 1 & 1 & & \\ \hline & & 1 & & \\ \hline & & & 1 & \\ \hline \text{X} & & & & \\ \hline 0 & 1 & 2 & 3 & 4 \\ \hline \end{array}$$

1) Referir qual é o nome pela qual é usualmente conhecida a operação "⊗" e indicar, justificando, qual o resultado da operação

$$R = M \otimes (S_1, S_2) \text{ com:}$$

$$S_1 = \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & \times & \\ \hline & & \\ \hline \end{array}$$

$$S_2 = \begin{array}{|c|c|c|} \hline & 1 & \\ \hline 1 & \times & \\ \hline & & \\ \hline \end{array}$$

Operação Hit-and-Miss ⊗: ↳ é usada na busca de padrões de pixels.

↳ recebe dois elementos estruturantes:  $A \otimes (B, C) = (A \ominus B) \cap (A^c \ominus C)$

No nosso caso:  $M \otimes (S_1, S_2) = (M \ominus S_1) \cap (M^c \ominus S_2)$

↳ analisar onde se verifica a correspondência do elemento  $S_1$  em  $M$

1º Passo:

$$\begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & \\ \hline \times & & \\ \hline \end{array}$$

2º Passo:

$$\begin{array}{|c|c|c|} \hline & & 0 \\ \hline & 0 & \\ \hline \times & & \\ \hline \end{array}$$

3º Passo:

$$\begin{array}{|c|c|c|} \hline & & 1 \\ \hline & 1 & \\ \hline \times & & \\ \hline \end{array}$$

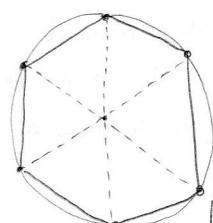
1º Passo: verificar os pontos de  $M$  onde se verifica o padrão de  $S_1$ .

2º Passo: complementar a imagem obtida no passo anterior (0 passa a 1 e 1 passa a 0); aplicar  $S_2$  nos pontos obtidos no 1º passo e verificar o padrão obtido.

3º Passo: encontrar os padrões que tenham  $S_1$  mas não tenham  $S_2$ .

**Exercício 3** Considerar os descritores de regiões.

a) Atendendo à definição do factor de forma, que é unitário para um círculo, e igual a  $\pi/4$  para uma região quadrada, determinar o factor de forma teórico para uma região hexagonal regular.



$$ff = \frac{4\pi A}{P^2} \Rightarrow ff = \frac{4\pi \cdot \left(\frac{6\sqrt{3}L^2}{4}\right)}{(6L)^2} \Rightarrow ff = \frac{\sqrt{3}\pi}{6} \Rightarrow ff = 0,9069$$

(C) A área do hexágono é dada pela soma das áreas dos 6 triângulos equiláteros no qual o hexágono se divide:

$$A_{\Delta} = \frac{b \times h}{2} = \frac{L \cdot \sqrt{3}L/2}{2} = \frac{\sqrt{3}L^2}{4}$$

$$\text{NOTA: } \sin 60^\circ = \frac{\sqrt{3}}{2}$$

$$\text{Como são 6 triângulos temos que } A_{\Delta} = \frac{6\sqrt{3}L^2}{4}$$

b) Considerar que se dispõe de uma imagem como a ilustrada a seguir, e da qual se pretende calcular o número de Euler, mas não se dispõe de uma função que o faça directamente. Baseando-se apenas em operações morfológicas e/ou de pré-processamento, indicar que operações se deveriam levar a cabo para calcular o número de Euler da imagem ( $N_E = N_o - N_b$ ). N.B: Se se entender útil, pode-se complementar a descrição das operações com os correspondentes em Matlab. Admitir que o objecto é feito de pixels de valor 1 e o fundo 0.

Nº Euler: nº de objectos muros + nº de buracos - nº muros objectos.

Passos a considerar: → limpar os objectos dos bordos

→ oriar uma semente

→ propagar a semente até à máscara

→ inverter

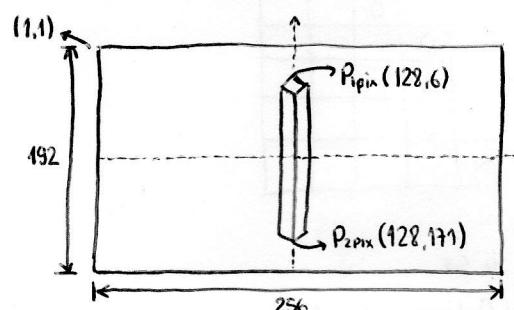
→ fazer uma erosão ( $C_1 \oplus C_2$ ) ⇒ contar os pontos que dão o nº de objectos ( $N_o$ )

→ somar propagação com a máscara

→ inverter e fazer uma erosão ( $C_1 \oplus C_2$ ) ⇒ contar os pontos que dão o nº de buracos ( $N_b$ )

$$\rightarrow N_E = N_o - N_b$$

Exercício 4: Seja uma imagem de  $256 \times 192$  pixels, oriunda de um CCD com 150 pixels/mm e obtida com uma lente com distância focal de 6 mm. Nessa imagem está uma coluna vertical que se sabe estar paralela ao plano da imagem e que tem 2m de altura.



a) Com base nos dados, indicar o valor numérico da matriz intrínseca da câmara.

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$$

em que  $K = \begin{bmatrix} 900 & 0 & 128 & 0 \\ 0 & 900 & 96 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(C.A)

$$\begin{aligned} \alpha_x &= f \cdot K_x = 6 \cdot 150 = 900 \\ \alpha_y &= f \cdot K_y = 6 \cdot 150 = 900 \\ x_0 &= 256/2 = 128 \\ y_0 &= 192/2 = 96 \end{aligned}$$

→ coordenadas reais

→ matriz K (matriz própria da câmara)

→ coordenadas homogêneas da imagem

b) Sabendo que os pixels P1 e P2 indicados na imagem são os extremos da coluna, determinar a distância a que está a coluna do plano da imagem. Representar, antes de tudo, a expressão da transformação geométrica matricial que relaciona os pontos no mundo físico PW1 e PW2 com os respectivos pontos (pixels) na imagem (P1 e P2), e só depois disso concluir o cálculo da distância a que se encontra a coluna.

MÉTODO 1

Sistema de equações:  $\begin{bmatrix} u \\ v \\ w \end{bmatrix} = K \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \rightarrow z_s \begin{bmatrix} x_{pix} \\ y_{pix} \\ 1 \end{bmatrix} = K \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$

$$P_{1pix} = \begin{bmatrix} 128 \\ 192-6 \\ 1 \end{bmatrix} = \begin{bmatrix} 128 \\ 186 \\ 1 \end{bmatrix}$$

$$P_{2pix} = \begin{bmatrix} 128 \\ 192-171 \\ 1 \end{bmatrix} = \begin{bmatrix} 128 \\ 21 \\ 1 \end{bmatrix}$$

Temos então:

$$P_1 = \begin{bmatrix} x_{s1} \\ y_{s1} \\ z_{s1} \end{bmatrix} \Rightarrow \text{Ponto 1: } Z_s \begin{bmatrix} 128 \\ 186 \\ 1 \end{bmatrix} = \begin{bmatrix} 900 & 0 & 128 \\ 0 & 900 & 96 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s1} \\ y_{s1} \\ z_{s1} \end{bmatrix} \Rightarrow \begin{cases} 128z_s = 900x_{s1} + 128z_s \\ 186z_s = 900y_{s1} + 96z_s \\ z_s = z_s \end{cases}$$

$$P_2 = \begin{bmatrix} x_{s2} \\ y_{s2} \\ z_{s2} \end{bmatrix} \Rightarrow \text{Ponto 2: } Z_s \begin{bmatrix} 128 \\ 21 \\ 1 \end{bmatrix} = \begin{bmatrix} 900 & 0 & 128 \\ 0 & 900 & 96 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s2} \\ y_{s2} \\ z_{s2} \end{bmatrix} \Rightarrow \begin{cases} 128z_s = 900x_{s2} + 128z_s \\ 21z_s = 900y_{s2} + 96z_s \\ z_s = z_s \end{cases}$$

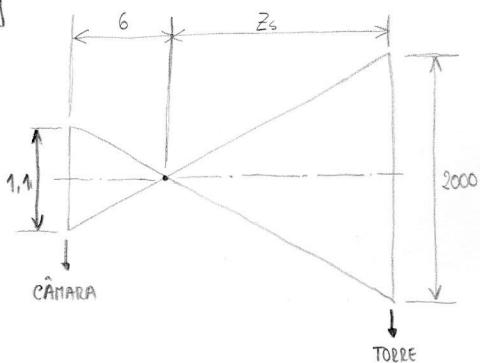
Passando para um sistema de equações:

$$\begin{aligned} 1^{\text{a}} \text{ equação: } & 186z_s = 900y_{s1} + 96z_s \\ 2^{\text{a}} \text{ equação: } & 21z_s = 900y_{s2} + 96z_s \\ 3^{\text{a}} \text{ equação: } & y_{s1} - y_{s2} = 2000 \end{aligned} \Rightarrow \begin{cases} 90z_s = 900y_{s1} \\ -75z_s = 900y_{s2} \\ y_{s1} - y_{s2} = 2000 \end{cases} \Rightarrow \begin{cases} y_{s1} = \\ y_{s2} = \\ z_s = 10909 \text{ mm} \end{cases}$$

(Altura da coluna em mm)

A coluna está colocada, portanto, a  $z_s = 10909 \text{ mm} = 10,91 \text{ m}$  do plano da imagem.

TODO 2



$$\left. \begin{aligned} P_1 \rightarrow \frac{x-x'}{150} &= \frac{128-128}{150} = \begin{bmatrix} 0 \\ 0,64 \end{bmatrix} \text{ mm} \\ P_2 \rightarrow \frac{x-x'}{150} &= \frac{128-128}{150} = \begin{bmatrix} 0 \\ -0,5 \end{bmatrix} \text{ mm} \end{aligned} \right\} \text{Coordenadas no CCD}$$

$$\textcircled{A} \quad d = \sqrt{(x_{s1}-x_{s2})^2 + (y_{s1}-y_{s2})^2} = \sqrt{(0-0)^2 + (0,64+0,5)^2} = 1,1 \text{ mm}$$

$$\text{Temos então: } \boxed{\frac{d}{f} = \frac{D}{z_s}} \Leftrightarrow \frac{1,1}{6} = \frac{2000}{z_s} \Leftrightarrow z_s = 10909 \text{ mm} = 10,91 \text{ m}$$



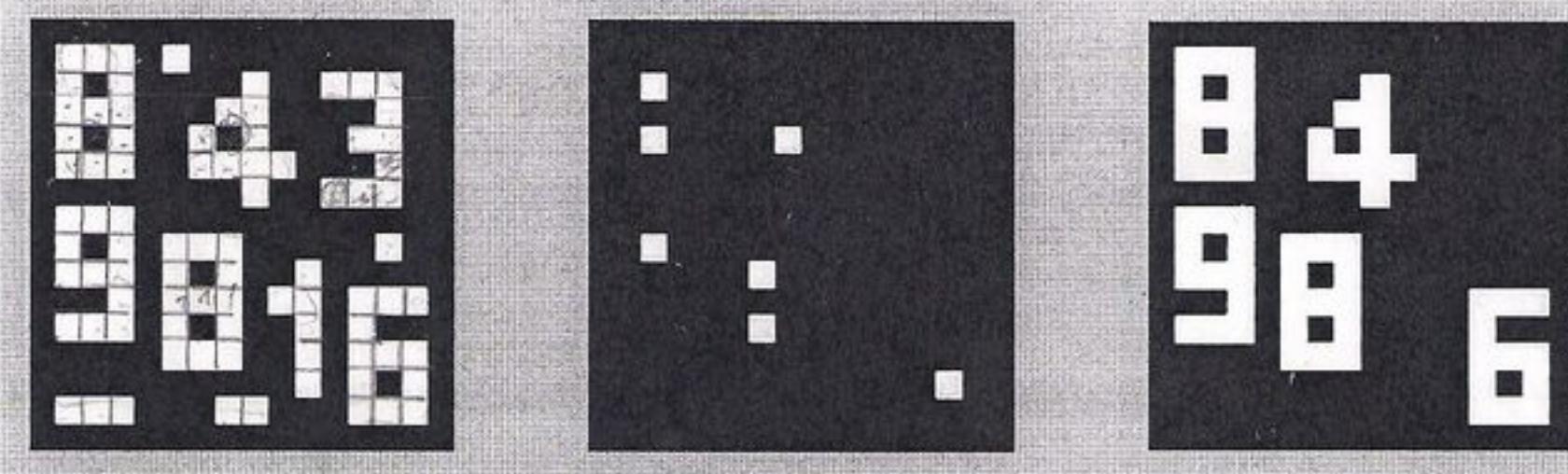
# Sistemas de Visão e Percepção Industrial

Exame de Época Especial

13 de Outubro de 2009

Mestrado Integrado em Engenharia Mecânica  
Mestrado em Engenharia de Automação Industrial

1. Sejam as seguintes imagens binárias de 16x16 pixels.

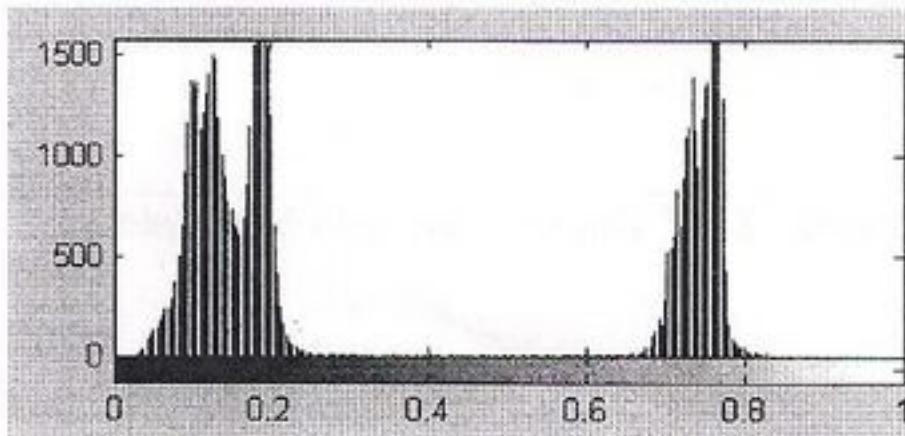


A

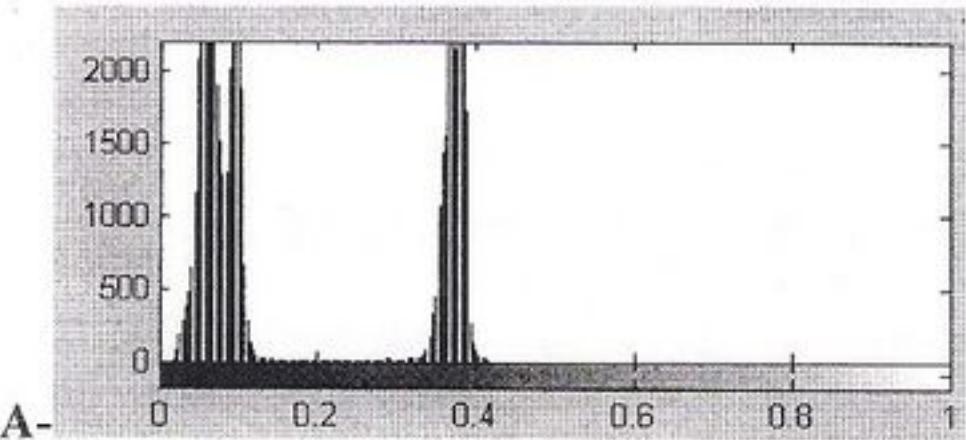
B

C

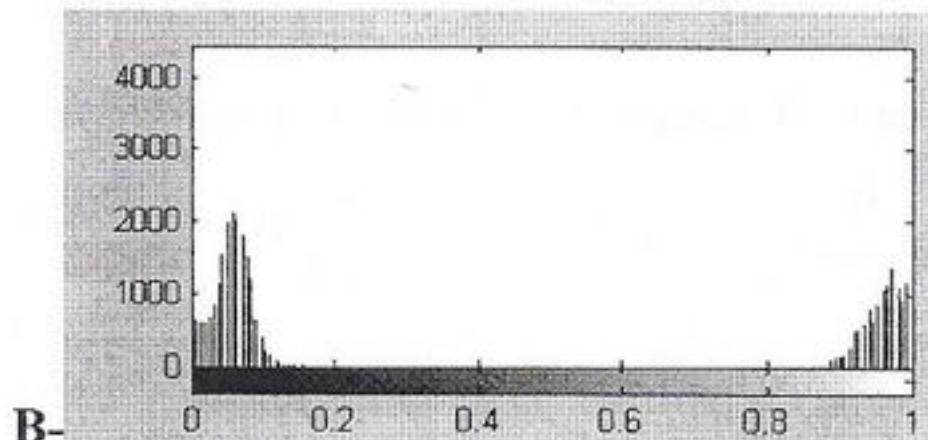
- Indicar um filtro de convolução de 3x3 que permita extrair apenas os pontos interiores (buracos) dos objectos (algarismos) da imagem A para gerar a imagem B; mostrar os cálculos ou expressões para obter a imgem B. (Sugestão: atender a vizinhanças diferentes da usual N<sub>8</sub>).
  - Indicar e descrever os procedimentos para obter a imagem C a partir de A e B. Indicar as operações morfológicas usadas e quais os respectivos argumentos.
  - Indicar e descrever um procedimento para obter uma imagem só com os algarismos 1 e 3 e tudo o resto, incluindo o ruído, deve ficar excluído. Indicar operações e parâmetros.
2. Seja uma dada imagem  $f(x, y)$  em níveis de cinzento ( $f(x, y) \in [0; 1]$ ), com as dimensões de N linhas  $\times$  M colunas, e que apresenta o seguinte histograma:



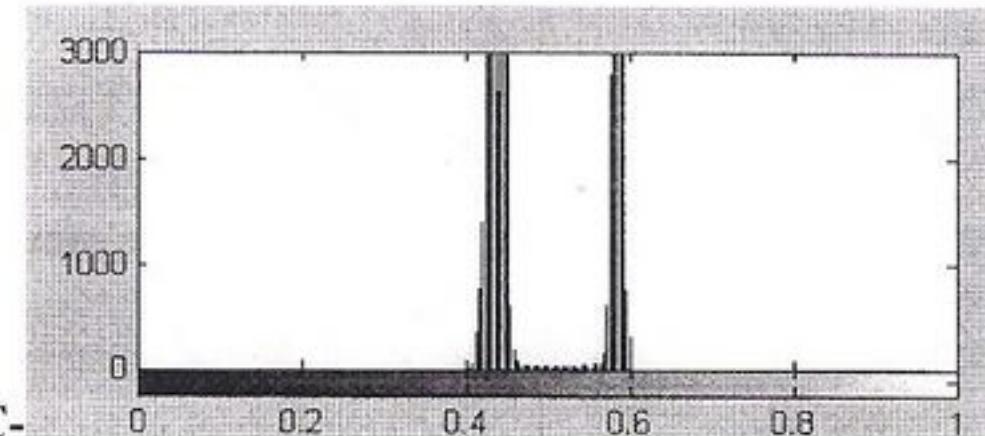
- A quatro cópias dessa imagem foram aplicadas quatro transformações que resultaram nos seguintes histogramas: A, B, C e D. Indicar justificando, que operações/transformações poderiam ter sido aplicadas a cada uma das cópias para obter os referidos histogramas.



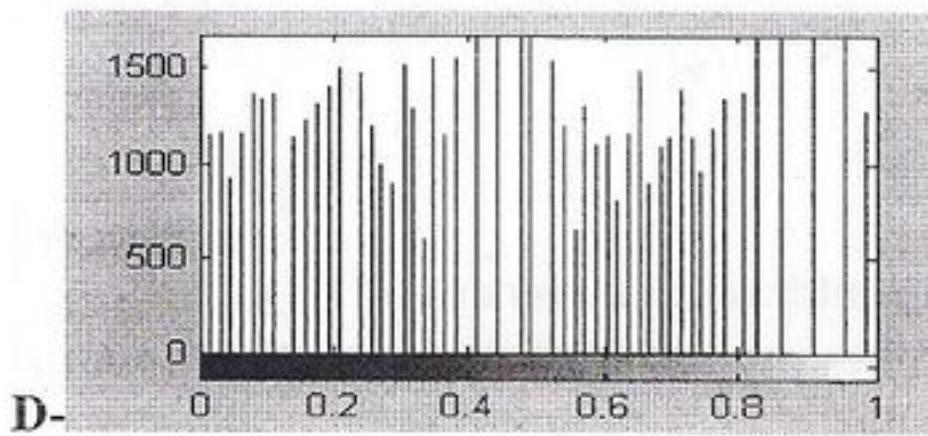
A-



B-



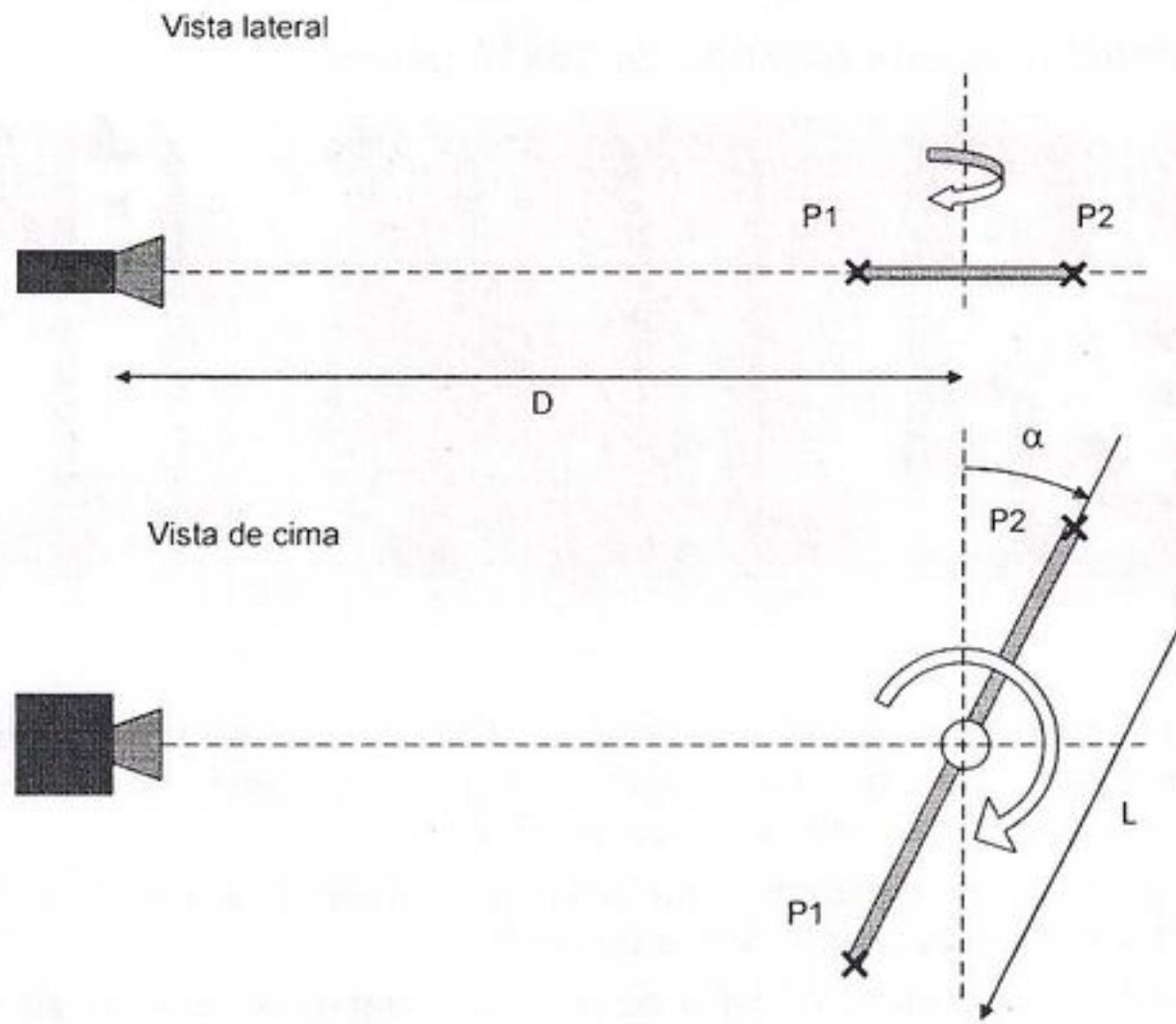
C-



D-

- O que representa a imagem  $g(x, y) = 1 - f(x, y)$  e como será o seu histograma em relação a  $f(x, y)$ ?
- A partir da imagem original obtém-se a imagem  $k(u, v) = f(M - x, y)$ . O que representa a imagem  $k(u, v)$  e como será o seu histograma em relação a  $f(x, y)$ ? (N.B:  $u, x \in \{0, 1, \dots, M - 1\}$  e  $v, y \in \{0, 1, \dots, N - 1\}$ )

3. Seja uma câmara ideal com distância focal  $f$ , um *dot pitch*  $d$  e uma resolução de  $W \times H$  pixels. No campo visual da câmara, uma barra de comprimento  $L$  gira sobre o seu eixo como ilustrado. A secção da barra é de dimensões desprezáveis para este problema. São definidos dois pontos  $P_1$  e  $P_2$  nas extremidades da barra. O eixo de rotação da barra está a uma distância  $D$  do plano focal da câmara. O centro óptico do CCD é o seu centro geométrico  $(W/2, H/2)$ . O ângulo  $\alpha$  pode variar entre  $+90^\circ$  e  $-90^\circ$ .



- d) Com base nos dados genéricos indicados ( $W, H, L, f, d, D, \alpha$ ), e formalizando as expressões usando a matriz intrínseca da câmara, obter as coordenadas das imagens de  $P_1$  e  $P_2$  em pixels, ou seja, obter as expressões genéricas para  $p_{1pix} = [x_{pix1} \ y_{pix1}]^T$  e  $p_{2pix} = [x_{pix2} \ y_{pix2}]^T$ .
- e) Determinar a expressão da distância  $D_{min}$  a que se pode aproximar a câmara do objecto de modo a que ele seja sempre integralmente visível na imagem.

**Cotação:** Questão 1 – 7 Valores. Questão 2 – 7 Valores. Questão 3– 6 Valores.

### Breve formulário

Momentos de imagens:

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}, \quad m_{01} = \sum_x \sum_y y \cdot f(x, y), \quad m_{10} = \sum_x \sum_y x \cdot f(x, y)$$

Expressões para operar/analisar histogramas:

$$\mu_n = \sum_{i=0}^{L-1} (i - \mu_0)^n h(i), \quad \mu_0 = \sum_{i=0}^{L-1} i h(i), \quad \text{com } h(i) \text{ normalizado, i.e., } 0 \leq h(i) < 1, \forall i \in \{0, 1, 2, \dots, L-1\}$$

$$m_k^b = \frac{\sum_{i=0}^{T_{k-1}-1} i H(i)}{\sum_{i=0}^{T_{k-1}-1} H(i)}, \quad m_k^f = \frac{\sum_{i=T_{k-1}}^{L-1} i H(i)}{\sum_{i=T_{k-1}}^{L-1} H(i)}, \quad g(x, y) = (L-1) \frac{f(x, y) - \min[f(x, y)]}{\max[f(x, y)] - \min[f(x, y)]}$$

$$\text{Matriz intrínseca da câmara: } \mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Equação polar da recta:  $x \cos \theta + y \sin \theta = \rho$

Relações trigonométricas:  $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ ,  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$

# SISTEMAS DE VISÃO E PERCEPÇÃO INDUSTRIAL

Exame de Época Especial - 13 de Outubro de 2009

Exercício 1 Sejam as seguintes imagens binárias de  $16 \times 16$  pixels.

a) Indicar um filtro de convolução de  $3 \times 3$  que permita extrair apenas os furos interiores (buracos) dos objectos (algazarmos) da imagem A para gerar a imagem B; mostrar os cálculos ou expressões para obter a imagem B. (Sugestão: atender a vizinhanças diferentes da usual N8).

Princípio de utilização de um filtro: o novo pixel é função do pixel original e do filtro.

$$g(x,y) = \sum_{i=1}^N w_i \cdot p_i$$

└ pertence à vizinhança do ponto em causa  
└ pode ser o valor directo do novo pixel, ou ser usada de forma indirecta e assim permitir calcular o seu novo valor.

Os coeficientes de um filtro  $3 \times 3$  são:

w <sub>1</sub>	w <sub>2</sub>	w <sub>3</sub>
w <sub>4</sub>	w <sub>5</sub>	w <sub>6</sub>
w <sub>7</sub>	w <sub>8</sub>	w <sub>9</sub>

Neste caso será de tipo:

-1	-1	-1
-1	8	-1
-1	-1	-1

» Filter =  $[-1 -1 -1; -1 8 -1; -1 -1 -1]$ ;

» aux = filter2(Filter, A);      / Usa o filtro definido na imagem A

» B = A;                          / Mudança de variável

» B(aux <=-7) = 1

b) Indicar e descrever os procedimentos para obter a imagem C a partir de A e B. Indicar as operações morfológicas usadas e quais os respectivos argumentos.

Preparação (imreconstruct()): operação de dilatação sucessiva de um objecto "semente" (por exemplo um esqueleto) até aos limites dadas por uma imagem "máscara", com base num dado elemento estruturante.

• Máscara usada: imagem A

• Semente usada: imagem B

» C = imreconstruct(B, A);

» C = C - B

c) Indicar e descrever um procedimento para obter uma imagem só com os algarismos 1 e 3 e tudo o resto, incluindo o ruído, deve ficar excluído. Indicar operações e parâmetros.

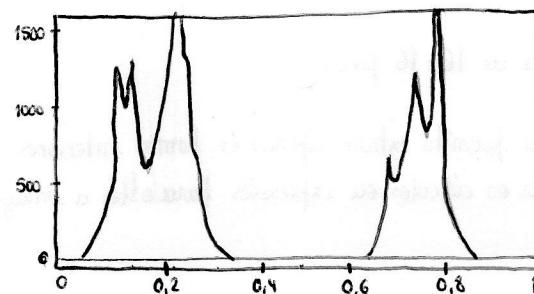
» D = A - C;

» Filtro = [20 20 20; -2 -1 -2; 20 20 20];

» aux2 = filter(Filtro, D);

» D (aux2 == -5 and aux2 == -3 and aux2 == -1) = 0

Exercício 2: Seja uma dada imagem  $f(x,y)$  em níveis de cinzento ( $f(x,y) \in [0;1]$ ), com as dimensões de  $N$  linhas  $\times M$  colunas, e que apresenta o seguinte histograma:



a) A quatro cópias dessa imagem foram aplicadas quatro transformações que resultaram nos seguintes histogramas: A, B, C e D. Indicar, justificando, que operações/transformações poderiam ter sido aplicadas a cada uma das cópias para obter os referidos histogramas.

Histograma A: afetar linearmente o bicho (ex: dividir o valor dos pixels por 2).

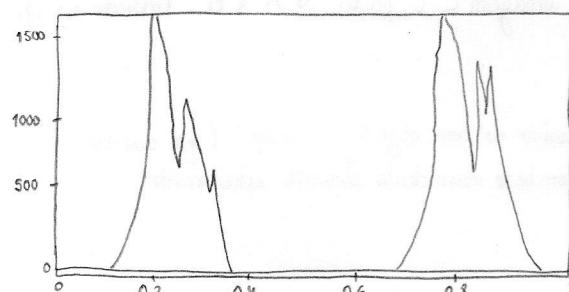
Histograma B: expansão do histograma (operação que altera os valores dos pixels de forma linear e de modo a ocupar gamas diferentes de níveis); afeta o contraste da forma global. Exemplo de expansão [15% 70%] para [0% 100%].

Histograma C: contração do histograma. Exemplo de "expansão" [0% 80%] para [40% 60%].

Histograma D: equalização do histograma (operação que procura distribuir uniformemente o nº de pixels pelos níveis); obriga a eliminar alguns níveis do histograma original; tende a facilitar a comparação posterior de imagens; em geral aumenta o contraste local.

b) O que representa a imagem  $g(x,y) = 1 - f(x,y)$  e como será o seu histograma em relação a  $f(x,y)$ ?

$g(x,y) = 1 - f(x,y) \Rightarrow$  representa o negativo da imagem, o que provoca a inversão do histograma.

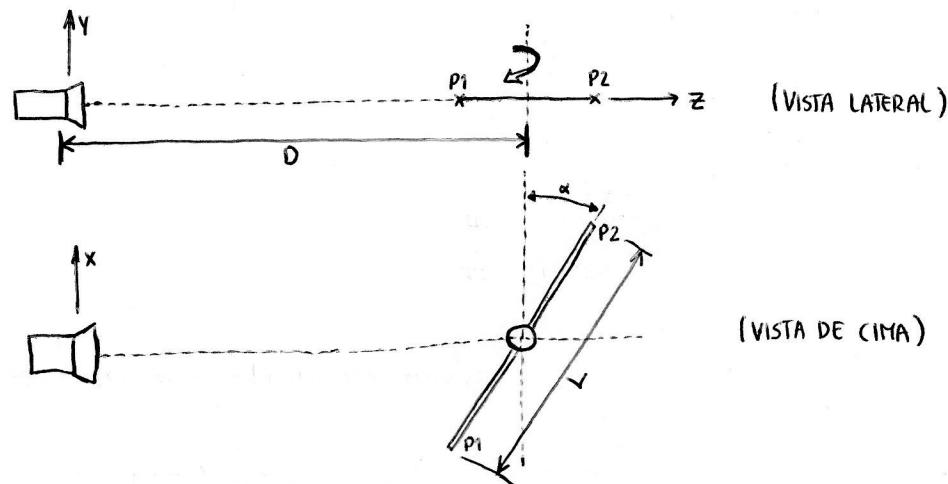


c) A partir da imagem original obtém-se a imagem  $k(u,v) = f(M-x,y)$ . O que representa a imagem  $k(u,v)$  e como será o seu histograma em relação a  $f(x,y)$ ? (N.B:  $u,x \in \{0,1,\dots,M-1\}$  e  $v,y \in \{0,1,\dots,N-1\}$ ).

• M representa o nº de colunas.

$k(u,v) = f(M-x,y)$  origina uma inversão da imagem segundo o eixo x. O histograma será, por isso, igual ao histograma original.

Exercício 3: Seja uma câmara ideal com distância focal  $f$ , um dot-futch  $d$  e uma resolução de  $W \times H$  pixels. No campo visual da câmara, uma barra de comprimento  $L$  gira sobre o seu eixo como ilustrado. A secção da barra é de dimensões desprezáveis para este problema. São definidos dois pontos  $P_1$  e  $P_2$  nas extremidades da barra. O eixo de rotação da barra está a uma distância  $D$  do plano focal da câmara. O centro óptico do CCD é o seu centro geométrico  $(W/2, H/2)$ . O ângulo  $\alpha$  pode variar entre  $+90^\circ$  e  $-90^\circ$ .



a) Com base nos dados genéricos indicados ( $W, H, L, f, d, D, \alpha$ ), e formulando as expressões usando a matriz intrínseca da câmara, obter as coordenadas das imagens de  $P_1$  e  $P_2$  em pixels, ou seja, obter as expressões genéricas para  $P_{1\text{pix}} = [x_{1\text{pix}} \ y_{1\text{pix}}]^T$  e  $P_{2\text{pix}} = [x_{2\text{pix}} \ y_{2\text{pix}}]^T$ .

$$Z_s \begin{bmatrix} x_{\text{pix}} \\ y_{\text{pix}} \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha x & 0 & x_0 \\ 0 & \alpha y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} \Leftrightarrow Z_s \begin{bmatrix} x_{\text{pix}} \\ y_{\text{pix}} \\ 1 \end{bmatrix} = \begin{bmatrix} f.d & 0 & W/2 \\ 0 & f.d & H/2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix}$$

$$\cos \alpha = \frac{x_s}{L/2} \Rightarrow x_s = \frac{L}{2} \cdot \cos \alpha$$

$$Z_s = D - \frac{1}{2} \operatorname{sen} \alpha$$

$$y_s = 0$$

b) Determinar a expressão da distância  $D_{\min}$  a que se pode aproximar a câmara do objeto de modo a que ele seja sempre integralmente visível na imagem.



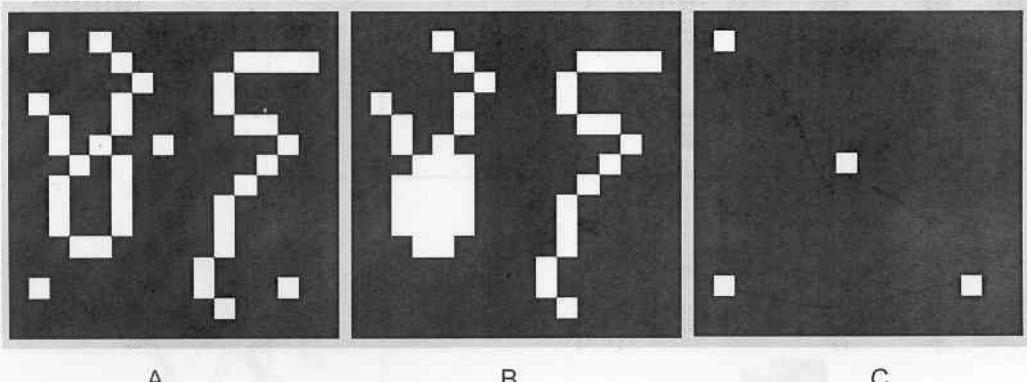
# Sistemas de Visão e Percepção Industrial

Exame de Época de Recurso

8 de Julho de 2009

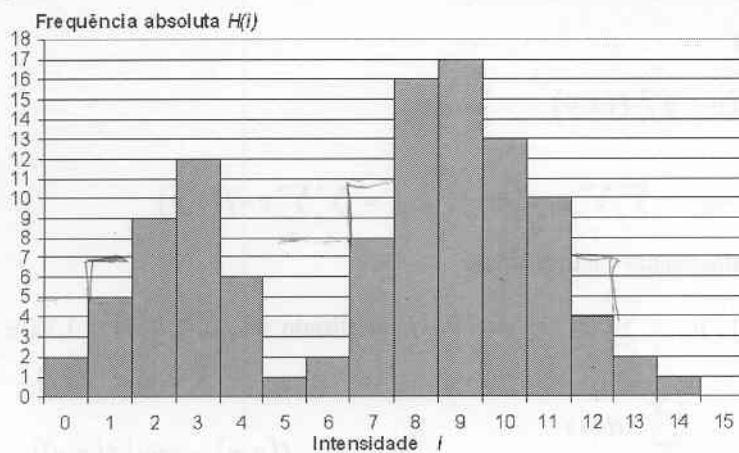
Mestrado Integrado em Engenharia Mecânica  
Mestrado em Engenharia de Automação Industrial

## 1. Sejam a seguintes imagens binárias de 16x16 pixels.



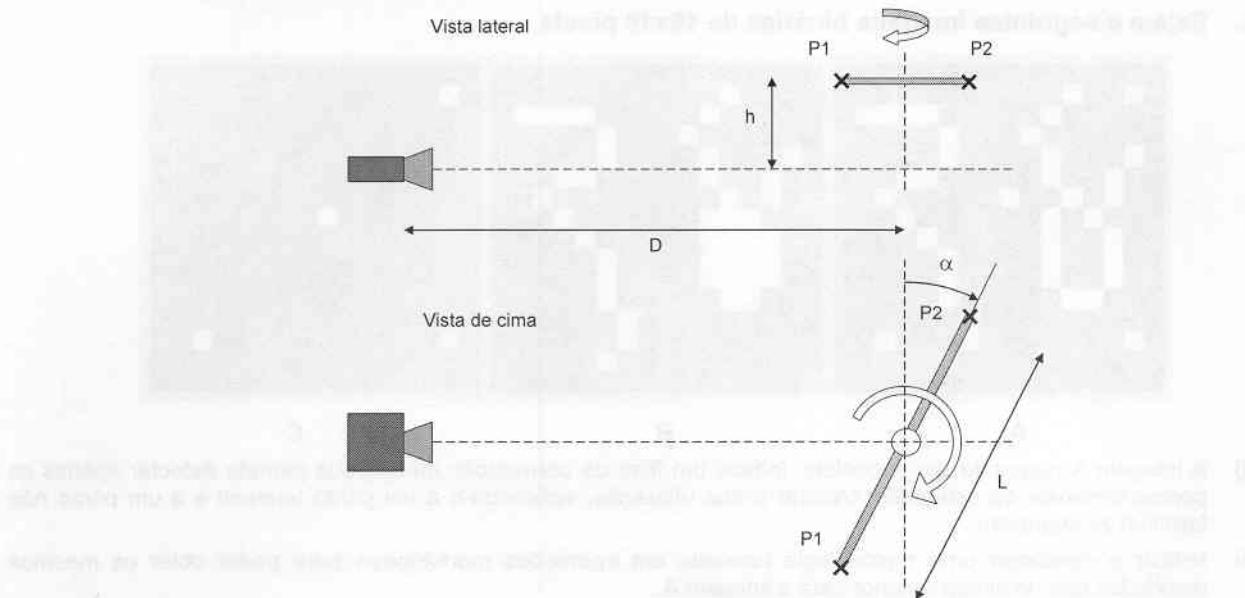
- A imagem A representa um esqueleto. Indicar um filtro de convolução de 3x3 que permita detectar apenas os pontos terminais do esqueleto. Mostrar a sua utilização, aplicando-o a um ponto terminal e a um ponto não terminal do esqueleto.
- Indicar e descrever uma metodologia baseada em operações morfológicas para poder obter os mesmos resultados que na alínea anterior para a imagem A.
- Descrever uma metodologia com operadores concretos (sem usar o `fill(...,'holes',...)` do Matlab) para poder obter a imagem B a partir da imagem A. Devem enumerar-se todas as operações e os seus parâmetros ou argumentos.
- Sabendo que a transformada de Hough de uma imagem binária num espaço de parâmetros para linhas rectas ( $\rho, \theta$ ) são sinusóides nesse espaço ( $\rho = k_i \sin(\theta + \varphi_i)$   $i = 1, 2, \dots, N$ ), teoricamente (i.e., sem limites de resolução), indicar quantas sinusóides distintas tem a transformada de Hough da imagem C e esboçar num gráfico ( $\rho, \theta$ ) os diversos pontos de intersecção dessas sinusóides, apontando também o seu significado.

## 2. Seja uma imagem a 16 níveis de cinzento com o histograma absoluto $H(i)$ indicado.



- Admitindo que a imagem tem uma relação de 4x3 (largura X altura), indicar, justificando, as suas dimensões.
- Considere o algoritmo de isodados (*isodata*) para a binarização de imagem. Definindo como o valor inicial para esse algoritmo iterativo ( $T_0$ ) a intensidade média ponderada dos pixels da imagem, calcular, com a indicação do cálculo, o valor da iteração seguinte do algoritmo ( $T_1$ ).
- Obter um indicador formal da simetria do histograma (momento de terceira ordem) e comentar o seu valor face ao desvio que apresenta em relação ao valor central.
- Numa operação de redefinição de limites de intensidades, definiu-se que grupos de 3 ou menos pixels (grupos 0, 5 e 14 na imagem) passariam a ter a intensidade do grupo mais próximo (no sentido da média). De seguida, nessas novas condições, faz-se a expansão do contraste (*contrast stretching*) para a gama dos 16 níveis. Que novo valor passará a um pixel que na imagem original tinha intensidade 9? Indicar as operações para o cálculo.

3. Seja uma câmara ideal com distância focal  $f$ , um *dot pitch*  $d$  e uma resolução de  $W \times H$  pixels. No campo visual da câmara, uma barra de comprimento  $L$  gira sobre o seu eixo a uma altura  $h$  do eixo óptico como ilustrado. A secção da barra é de dimensões desprezáveis para este problema. São definidos dois pontos  $P_1$  e  $P_2$  nas extremidades da barra. O eixo de rotação da barra está a uma distância  $D$  do plano focal da câmara. O centro óptico do CCD é o seu centro geométrico ( $W/2, H/2$ ). O ângulo  $\alpha$  pode variar entre  $+90^\circ$  e  $-90^\circ$ .



- a) Com base nos dados genéricos indicados ( $W, H, f, d, D, h, \alpha$ ), e formalizando as expressões usando a matriz intrínseca da câmara, obter as coordenadas das imagens de  $P_1$  e  $P_2$  em pixels, ou seja, obter as expressões genéricas para  $p_{1pix} = [x_{pix1} \ y_{pix1}]^T$  e  $p_{2pix} = [x_{pix2} \ y_{pix2}]^T$ . Sugestão: para ajudar a verificar a expressão pode-se testá-la em casos particulares com  $h=0$  e  $\alpha=0$ , por exemplo.
- b) Se, por outro lado, na imagem da barra se souber que o seu comprimento é de  $X$  pixels, qual será a expressão para calcular o valor do ângulo  $\alpha$  em função de  $X$  e dos restantes parâmetros.

**Cotação:** Questão 1 – 8 Valores. Questão 2 – 8 Valores. Questão 3– 4 Valores.

## Breve formulário

Momentos de imagens:

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}, \quad m_{01} = \sum_x \sum_y y \cdot f(x, y), \quad m_{10} = \sum_x \sum_y x \cdot f(x, y)$$

Expressões para operar/analisar histogramas:

$$\mu_n = \sum_{i=0}^{L-1} (i - \mu_0)^n h(i), \quad \mu_0 = \sum_{i=0}^{L-1} i h(i), \quad \text{com } h(i) \text{ normalizado, i.e., } 0 \leq h(i) < 1, \forall i \in \{0, 1, 2, \dots, L-1\}$$

$$m_k^b = \frac{\sum_{i=0}^{T_{k-1}-1} i H(i)}{\sum_{i=0}^{T_{k-1}-1} H(i)}, \quad m_k^f = \frac{\sum_{i=T_{k-1}}^{L-1} i H(i)}{\sum_{i=T_{k-1}}^{L-1} H(i)}, \quad g(x, y) = (L-1) \frac{f(x, y) - \min[f(x, y)]}{\max[f(x, y)] - \min[f(x, y)]}$$

$$\text{Matriz intrínseca da câmara: } \mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Equação polar da recta:  $x \cos \theta + y \sin \theta = \rho$

Relações trigonométricas:  $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ ,  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$

# SISTEMAS DE VISÃO E PERCEPÇÃO INDUSTRIAL

Exame de Época de Recurso - 8 de Julho de 2009

**Exercício 1:** Sejam as seguintes imagens binárias de  $16 \times 16$  pixels.

a) A imagem A representa um esqueleto. Indicar um filtro de convolução de  $3 \times 3$  que permita detectar apenas os pontos terminais do esqueleto. Montar a sua utilização, aplicando-o a um ponto terminal e a um ponto não terminal do esqueleto.

Princípio de utilização de um filtro: o novo pixel é função do pixel original e do filtro.

$$g(x,y) = \sum_{i=1}^N w_i \cdot p_i$$

└ pertence à vizinhança do ponto em causa  
└ pode ser o valor directo do novo pixel, ou ser usado de forma indireta e assim permitir calcular o seu novo valor.

Os coeficientes de um filtro  $3 \times 3$  são:

w <sub>1</sub>	w <sub>2</sub>	w <sub>3</sub>
w <sub>4</sub>	w <sub>5</sub>	w <sub>6</sub>
w <sub>7</sub>	w <sub>8</sub>	w <sub>9</sub>

» Filtro =  $[1 \ 1 \ 1; 1 \ -8 \ 1; 1 \ 1 \ 1]$  % Quando o resultado da aplicação do filtro é -7 o ponto é extremidade.

Aplicação a um ponto terminal:

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -8 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 1 - 8 = -7 \Rightarrow \text{extremidade}$$

Aplicação a um ponto não terminal:

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -8 & 0 \\ 0 & 1 & 0 \end{bmatrix} = 1 - 8 + 1 = -6 \Rightarrow \text{não é extremidade}$$

b) Indicar e descrever uma metodologia baseada em operações morfológicas para poder obter os mesmos resultados que na alínea anterior para a imagem A.

c) Descrever uma metodologia com operadores concretos (sem usar o `fill(...)` do Matlab) para poder obter a imagem B a partir da imagem A. Devem enumerar-se todas as operações e os seus parâmetros ou argumentos.

```
>A = imread ('ImagenA.png');
>A = im2double (A);
>A = im2bw (A);
>Filtro = [1 1 1; 1 -8 1; 1 1 1];
>X = filter2 (A, Filtro);
>A (X == -8) = 0;
>Aux = zeros (16);
>Aux (1,1) = 1;
>A = ~A;
>B = imreconstruct (Aux, A);
>B = ~B
```

1) Sabendo que a transformada de Hough de uma imagem binária num espaço de parâmetros para linhas rectas ( $P, \theta$ ) são sinusóides nesse espaço ( $P = k_i \cdot \sin(\theta + \phi_i)$ ,  $i = 1, 2, \dots, N$ ), tecnicamente (i.e. sem limites de resolução), indicar quantas sinusóides distintas tem a transformada de Hough da imagem C e esboçar num gráfico ( $P, \theta$ ) os diversos pontos de intersecção dessas sinusóides, apontando também o seu significado.

Transformada de Hough: conjunto de curvas discretizadas (uma para cada pixel branco da imagem) que representam todas as rectas possíveis que passam em cada um desses pontos.

Exercício 2] Seja uma imagem a 16 níveis de cinzento com o histograma absoluto  $H(i)$  indicado.

a) Admitindo que a imagem tem uma relação de  $4 \times 3$  (largura  $\times$  altura), indicar, justificando, as suas dimensões.

$$\text{Nº Total de pixels} = 2+5+9+12+6+1+2+8+16+17+13+10+4+2+1 = 108$$

$$\begin{cases} l \times h = 108 \\ \frac{l}{h} = \frac{4}{3} \end{cases} \Rightarrow \begin{cases} l = 9 \\ h = 12 \end{cases}$$

b) Considerar o algoritmo de isodades (isodata) para a binarização de imagem. Definindo como o valor inicial para esse algoritmo iterativo ( $T_0$ ) a intensidade média ponderada dos pixels da imagem, calcular, com a indicação do cálculo, o valor da iteração seguinte do algoritmo ( $T_1$ ).

$$T_0 = \frac{(0 \times 2) + (1 \times 5) + (2 \times 9) + (3 \times 12) + (4 \times 6) + (5 \times 1) + (6 \times 2) + (7 \times 8) + (8 \times 16) + (9 \times 17) + (10 \times 13) + (11 \times 10) + (12 \times 4) + (13 \times 2) + (14 \times 1) + (15 \times 0)}{108}$$

↳ intensidade média

$$\Leftrightarrow T_0 = 7,08 \Leftrightarrow T_0 = 7$$

Binarização (algoritmo de isodades) : obtenção do limiar  $T$  de binarização de forma iterativa.

• Iteração inicial : ponto central da gama dinâmica  $T_0 = 2^{B-1}$  (em que  $B = \text{nº bits}$ )

$$\text{Média objecto} : m_x^o = \frac{\sum_{i=T_{k-1}}^{L-1} i \cdot H(i)}{\sum_{i=T_{k-1}}^{L-1} H(i)}$$

$$\text{Média fundo} : m_x^f = \frac{\sum_{i=0}^{T_{k-1}-1} i \cdot H(i)}{\sum_{i=0}^{T_{k-1}-1} H(i)}$$

• Nova iteração :  $T_k = \frac{m_x^f - m_x^o}{2} \rightarrow$  o processo é repetido até que o limiar deixe de variar :  $T_k = T_{k-1}$

i	H(i)	$h(i) = \frac{H(i)}{\text{Nº Total pixels}}$
0	2	0,019
1	5	0,046
2	9	0,083
3	12	0,111
4	6	0,056
5	1	0,009
6	2	0,019
7	8	0,074
8	16	0,148
9	17	0,157
10	13	0,120
11	10	0,093
12	4	0,037
13	2	0,019
14	1	0,009

$$m_x^o = \frac{(0 \times 2) + (1 \times 5) + (2 \times 9) + (3 \times 12) + (4 \times 6) + (5 \times 1) + (6 \times 2)}{2+5+9+12+6+1+2} = 2,70$$

$$m_x^f = \frac{(7 \times 8) + (8 \times 16) + (9 \times 17) + (10 \times 13) + (11 \times 10) + (12 \times 4) + (13 \times 2) + (14 \times 1)}{8+16+17+13+10+4+2+1} = 9,37$$

$$T_1 = \frac{m_x^o + m_x^f}{2} = \frac{2,70 + 9,37}{2} = 6$$

c) Obter um indicador formal da simetria do histograma (momento de terceira ordem) e comentar o seu valor face ao desvio que apresenta em relação ao valor central.

$$m = \sum_{i=0}^{L-1} i \cdot h(i) = 8,978$$

↳ intensidade média da imagem

3º momento do histograma:

- mede a assimetria do histograma
- é 0 para histogramas simétricos
- positivo para histogramas desviados para a direita (em relação à média)
- negativo para histogramas desviados para a esquerda
- também se costuma normalizar dividindo-o por  $(L-1)^2$

Numa operação de redefinição de limites de intensidades, definiu-se que grupos de 3 ou menos pixels (grupos 0,5 e 14 na imagem) passariam a ter a intensidade do grupo mais próximo (no sentido da média). De seguida, nessas mesmas condições, faz-se a expansão do contraste (contrast stretching) para a gama dos 16 níveis. Que novo valor passará a um pixel que na imagem original tinha intensidade 9? Indicar as operações para o cálculo.

$$9 - (0,5) + (14,5) = 9,5 - 0,5 = 9$$

$$9 - 0,5 + 14,5 = 9 + 14 = 23$$

$$23 - 0,5 + 14,5 = 23 + 14 = 37$$

Original	Novo
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100



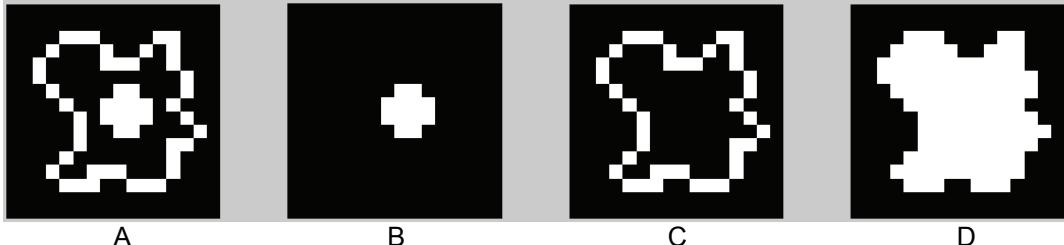
# Sistemas de Visão e Percepção Industrial

Exame de Época Normal

24 de Junho de 2009

Mestrado Integrado em Engenharia Mecânica  
Mestrado em Engenharia de Automação Industrial

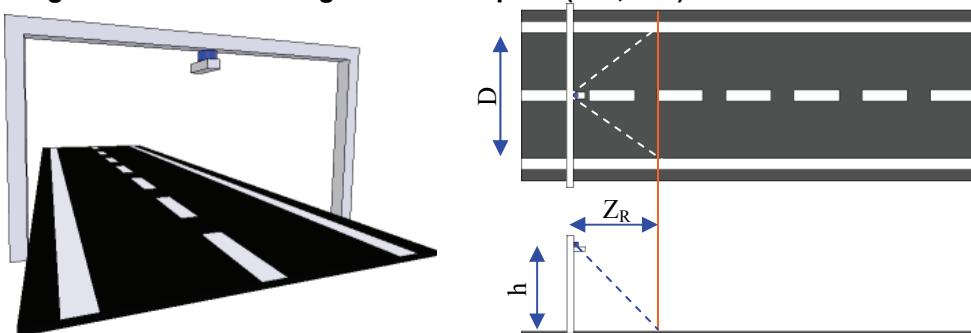
## 1. Sejam as seguintes imagens binárias de 16x16 pixels.



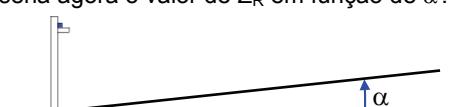
- Indicar um filtro X (3x3) que quando aplicado a A permite obter a imagem B. **Sugestão:** o procedimento para a execução desta tarefa pode ser:  $A_1 = \text{Filt}(X, A)$ ;  $B = \{p : p \in A_1 \cap A_1(p) > 11\}$  ( $p$  designa as coordenadas do pixel genérico da imagem e  $\text{Filt}()$  a operação de filtragem ou convolução).
- Calcular e indicar o factor de forma (ou circularidade) do objecto da imagem B.
- Indicar justificando os números de Euler das regiões presentes na imagem A.
- Sendo o objecto da imagem C um contorno, representar o respectivo código de cadeia usando um código de Freeman a 8 direcções e representar o código na sua forma normalizada. **Obs:** Os códigos de Freeman iniciam-se em zero na direcção horizontal para a direita, e evoluem em sentido anti-horário.
- Indicar as operações para obter a imagem D a partir de outras anteriores sem utilizar a operação de preenchimento automático de buracos (como o *imfill(..., 'holes')* no Matlab).
- Seja  $SE_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$  e  $SE_2 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ . Usando a propriedade da associatividade da operação de dilatação morfológica, indicar quantos pixels terá o objecto resultante da seguinte operação:  $B_2 = (B \oplus SE_1) \oplus SE_2$ . **Obs:** A origem do elemento estruturante é o seu centro geométrico.

- Qual é o número mínimo de erosões com o elemento estruturante  $\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$  necessário para eliminar todos os pixels da imagem D? **Obs:** A origem do elemento estruturante é o seu centro geométrico.

## 2. Seja uma câmara ideal com distância focal $f=12$ mm, um *dot pitch* $d=90$ pixels/mm e com uma resolução máxima de 721X577 pixels. A câmara está suspensa a uma altura $h=4$ m, centrada sobre uma estrada, e com o seu eixo óptico paralelo à linha tracejada central, como ilustrado na figura. A distância interior entre as linhas laterais da estrada é $D=6$ m e a estrada pode considerar-se recta e infinita em comprimento. Admite-se que o centro óptico da câmara é o centro geométrico do rectângulo do CCD: pixel (360, 288).



- Nas condições referidas, e em relação à câmara, determinar a distância  $Z_R$  dos pontos interiores das duas linhas laterais da estrada mais próximos da câmara que ainda são visíveis na imagem.
- Determinar qual é a máxima distância na estrada tal que as duas linhas laterais ainda surjam separadas por um pixel na imagem.
- Se a estrada não fosse paralela em relação ao eixo óptico da câmara, mas tivesse com ele uma inclinação  $\alpha$  como ilustrado a seguir, qual seria agora o valor de  $Z_R$  em função de  $\alpha$ ?



**Cotação:** Questão 1 – 14 Valores. Questão 2 – 6 Valores.

---

## Breve formulário

Matriz intrínseca da câmara:  $\mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Equação polar da recta:  $x\cos\theta + y\sin\theta = \rho$

Relações trigonométricas:  $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ ,  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$

Algumas propriedades de operações morfológicas:

$$A \oplus B = B \oplus A$$

$$(A \oplus B) \oplus C = A \oplus (B \oplus C)$$

$$A \oplus (B \cup C) = (A \oplus B) \cup (B \oplus C)$$

$$(A \ominus B) \ominus C = A \ominus (B \oplus C)$$

$$A \ominus (B \cup C) = (A \ominus B) \cap (B \ominus C)$$

$$(A \cap B) \ominus C = (A \ominus C) \cap (B \ominus C)$$

$$A \circ B = (A \ominus B) \oplus B$$

$$A \bullet B = (A \oplus B) \ominus C$$

$$(A \circ B) \circ B = A \circ B$$

$$(A \bullet B) \bullet B = A \bullet B$$

$$A \otimes (B, C) = (A \ominus B) \cap (A^C \ominus C)$$

$$\bigcup_i A \otimes (B_i, C_i) = \bigcup_i (A \ominus B_i) \cap (A^C \ominus C_i)$$

## Momentos de uma imagem

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y)$$

### Momentos Centrais:

$$m_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

$$m_{01} = \sum_x \sum_y y \cdot f(x, y)$$

$$m_{10} = \sum_x \sum_y x \cdot f(x, y)$$

# SISTEMAS DE VISÃO E PERCEPÇÃO INDUSTRIAL

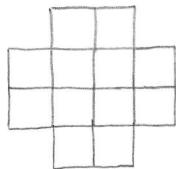
Exame de Época Normal - 24 de Junho de 2009

**Exercício 1** Sejam as seguintes imagens binárias de  $16 \times 16$  pixels.

- a) Indicar um filtro  $\times (3 \times 3)$  que quando aplicado a A permite obter a imagem B. Sugestão: o procedimento para a execução deste filtro pode ser:  $A1 = \text{Filt}(X, A)$ ;  $B = \{p \in A1 \mid A1(p) > 11\}$  ( $p$  designa as coordenadas do pixel genérico da imagem e  $\text{Filt}()$  a operação de filtragem ou convolução).

» Filtro =  $[1 \ 1 \ 1; 1 \ 1 \ 1; 1 \ 1 \ 1]$ ;  $\therefore$  Analisa cada pixel vizinho do analisado e conta-se se for igual a 1.  
»  $B = \text{filter2}(\text{Filtro}, A)$ ;  $\therefore$  Aplicação do filtro à imagem A  
»  $B(B \leq 5) = 0$   $\therefore$  Os pixels cujo nº de vizinhos brancos (1) é  $\leq 5$  é eliminado ( $= 0$ ).

- b) Calcular e indicar o factor de forma (ou circularidade) do objecto da imagem B.



- Perímetro = 16
- Área =  $(2 \times 4) + (1 \times 2) + (1 \times 2) = 12$

$$\text{Factor Forma: } ff = \frac{4\pi A}{p^2} = \frac{4\pi \cdot 12}{16^2} = 0,59$$

- c) Indicar justificando os números de Euler das regiões fechadas na imagem A.

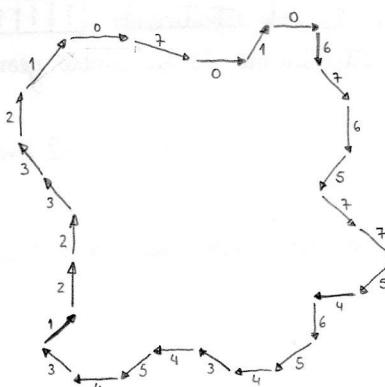
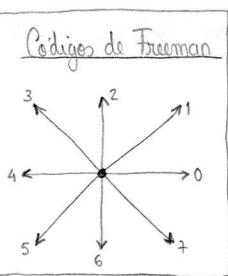
O número de Euler corresponde ao nº de objectos simples e nº de buracos nos objectos. Na imagem A, temos que:

$$Nº\text{ Euler} = 2\text{ objectos} - 1\text{ buraco} = 1$$

- d) Sendo o objecto da imagem C um contorno, representar o respetivo código de cadeia usando um código de Freeman a 8 direções e representar o código na sua forma normalizada. Obs: os códigos de Freeman iniciam-se em zero na direcção horizontal para a direita, e evoluem em sentido anti-horário.

Codificação de Freeman:  $\rightarrow$  para não definir o ponto de inicio, considerar uma cadeia circular e reordenar o ponto de começo da cadeia de forma a formar um inteiro o menor possível.

$\rightarrow$  para normalizar a orientação, usar a 1ª diferença de códigos numa vez dos previous códigos.



Cadeia: 0106765775465434543122332107 (normalizada pelo seu inicio)

2) Indicar as operações para obter a imagem D a partir de outras anteriores sem utilizar a operação de preenchimento automático de buracos (como o `IMFILL(..., 'holes')` no Matlab).

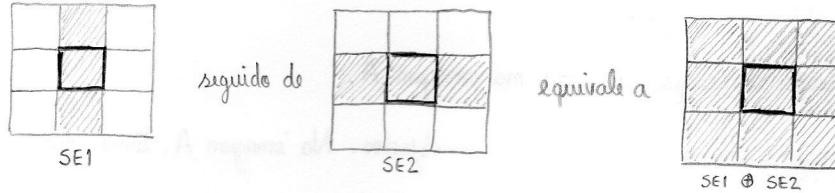
Propagação (`imreconstruct`(1)): operação de dilatação sucessiva de um objecto "semente" (por exemplo um esqueleto) até aos limites dados por uma imagem "máscara", com base num dado elemento estruturante.

- Máscara usada: imagem C
  - Semente usada: imagem B
- } em geral, o processo resulta numa máscara "cheia".

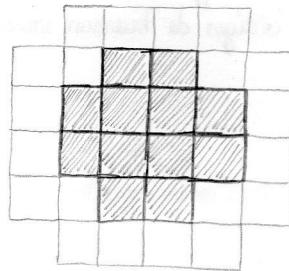
$D = \text{imreconstruct}(B, C)$

F) Seja  $SE_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$  e  $SE_2 = [1 \ 1 \ 1]$ . Usando a propriedade da associatividade da operação de dilatação morfológica, indicar quantos pixels terá o objecto resultante da seguinte operação:  $B_2 = (B \oplus SE_1) \oplus SE_2$ . Obs: a origem do elemento estruturante é o seu centro geométrico.

A associatividade permite, por exemplo, aplicar duas dilatações de seguida sobre uma imagem e obter o mesmo resultado. A dilatação, em geral, permite as dimensões do objecto.



A operação desejada é:  $B_2 = (B \oplus SE_1) \oplus SE_2 \Leftrightarrow B_2 = B \oplus (SE_1 \oplus SE_2)$



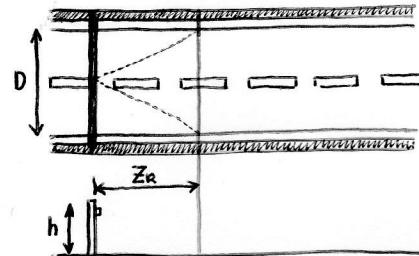
TOTAL DE PIXELS: 32

g) Qual é o número mínimo de erosões com o elemento estruturante  $[1 \ 1 \ 1 \ 1 \ 1]$  necessário para eliminar todos os pixels da imagem D? Obs: a origem do elemento estruturante é o seu centro geométrico.

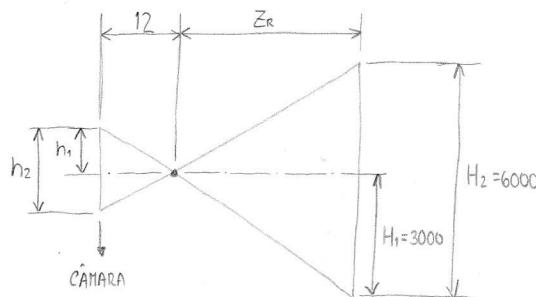
A zona mais larga do objecto da imagem D, tem 10 pixels de largura e 12 pixels de altura. O elemento estruturante tem 5 pixels de largura.

Contando que é necessário usar 2 vezes o SE em cada linha, temos que  $2 \times 12 = 24$  erosões (valor mínimo).

**Exercício 2** Seja uma câmara ideal com distância focal  $f = 12\text{ mm}$ , um dot-fitch  $d = 90\text{ pixels/mm}$  e com uma resolução máxima de  $721 \times 577$  pixels. A câmara está suspensa a uma altura  $h = 4\text{ mm}$ , centrada sobre uma estrada, e com o seu eixo óptico paralelo à linha tracjada central, como ilustrado na figura. A distância interior entre as linhas laterais da estrada é  $D = 6\text{ m}$  e a estrada pode considerar-se recta e infinita em comprimento. Admite-se que o centro óptico da câmara é o centro geométrico do rectângulo do CCD: pixel (360, 288).

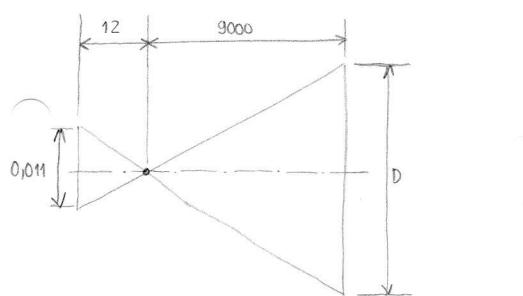


a) Nas condições referidas, e em relação à câmara, determinar a distância  $Z_R$  dos pontos interiores das duas linhas laterais da estrada mais próximas da câmara que ainda são visíveis na imagem.



$$\begin{aligned} \bullet h_2 &= \frac{721}{90} = 8\text{ mm} \\ \bullet h_1 &= \frac{360}{90} = 4\text{ mm} \\ \bullet \frac{12}{4} &= \frac{Z_R}{3000} \Leftrightarrow Z_R = 9000\text{ mm} \end{aligned}$$

b) Determinar qual é a máxima distância da estrada tal que as duas linhas laterais ainda sejam separadas por um pixel na imagem.



$$\begin{array}{l} 90\text{ pixels} \longrightarrow 1\text{ mm} \\ 1\text{ pixel} \longrightarrow x \end{array} \Rightarrow x = 0,011\text{ mm}$$

$$\frac{12}{0,011} = \frac{9000}{D} \Rightarrow D = 8,25\text{ mm}$$

c) Se a estrada não fosse paralela em relação ao eixo óptico da câmara, mas tivesse com ele uma inclinação  $\alpha$  como ilustrado a seguir, qual seria agora o valor de  $Z_R$  em função de  $\alpha$ ?





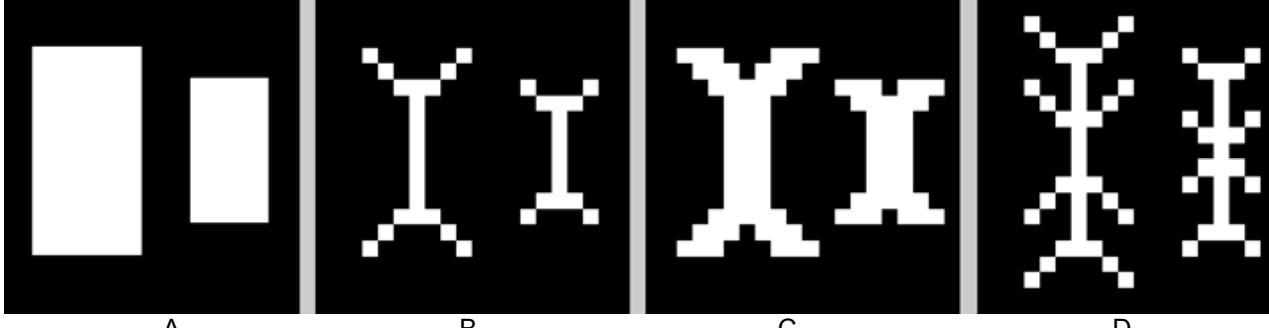
# Sistemas de Visão e Percepção Industrial

Exame de Época de Recurso

9 de Julho de 2008

Mestrado Integrado em Engenharia Mecânica  
Mestrado em Engenharia de Automação Industrial

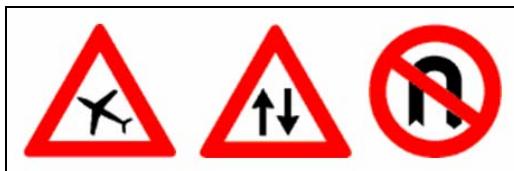
- 1- Sejam as seguintes imagens binárias de 20x20 pixels com os pontos de fundo a preto (valor 0).



A                      B                      C                      D

- A imagem B foi obtida a partir da imagem A por erosão condicionada. Indicar genericamente as condições a respeitar para efectuar (ou não efectuar) a erosão de cada pixel da imagem A para obter a imagem B.
- Seja Y a imagem que contém apenas os quatro pontos extremos de cada uma dos dois objectos de B. Indicar um filtro F de 3x3 tal que, quando aplicado à imagem B, resultando em  $Z = \text{filter2}(F, B)$ , permite obter  $Y = (Z == 7)$ .
- Obter a mesma matriz Y da alínea anterior mas agora através da operação morfológica *Hit-and-miss*. Indicar todos os elementos estruturantes necessários.
- A imagem C foi obtida pelo seguinte modo:  $C = B \oplus S_H$  onde  $S_H = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ ; de modo similar, tem-se que  $D = B \oplus S_V$ . Determinar e representar  $S_V$ . Obs: A origem do elemento estruturante é o seu centro geométrico.
- Indicar e demonstrar a veracidade da seguinte afirmação: "a operação de esqueletização de uma imagem binária preserva sempre os centros de massa dos diversos objectos dessa imagem".
- Qual o resultado da operação de abertura morfológica da imagem D com o elemento estruturante de defeito [ones(3,3)]? Justificar a resposta.

- 2- Seja a imagem seguinte, T, que contém três sinais de trânsito onde a orla de cada um é de cor vermelha saturada, o fundo é branco e os símbolos internos são de cor preta. A ordem ou distribuição dos três sinais pode ser variável na imagem e diferente do representado. A linha do caixilho não faz parte da imagem, serve apenas para delimitar a área a processar.



**Recomendação:** nas respostas às questões, atender à natureza das imagens onde são aplicadas as diversas operações (RGB, níveis de cinzento, binária, etc...); operações em imagens desapropriadas podem tornar a resposta incompleta ou mesmo errada. Deve-se portanto acautelar as menções às conversões de formato de imagem sempre que necessário. A imagem original está em formato RGB. É permitido o recurso a funções de alto nível do Matlab, se for necessário.

- Quais os passos necessários para separar os três sinais (, , ) em três imagens distintas A, B e C também no formato RGB?
- Admitindo que se consegue obter as máscaras *mask1*, *mask2* e *mask3* para as três imagens, como se determina qual delas é a máscara da imagem com o sinal circular? , , .
- Quais as operações para, a partir de A, B e C, extraer de cada uma o respectivo símbolo? Ou seja, os seguintes

elementos:  $X_{1S} = \begin{bmatrix} \text{pedestrian} \\ \text{up-down} \\ \text{circle} \end{bmatrix}$ ,  $X_{2S} = \begin{bmatrix} \text{up-down} \\ \text{circle} \end{bmatrix}$ ,  $C_S = \begin{bmatrix} \text{pedestrian} \end{bmatrix}$ .

- Admitindo que se obtém as imagens dos símbolos da alínea anterior, como se podem distinguir os sinais triangulares entre si, ou seja  $X_{1S}$  de  $X_{2S}$ ? Indicar todas as operações necessárias.
- Se por alguma razão um dos sinais surgisse no bordo da imagem, que operações seriam necessárias para o eliminar, incluindo todo o seu conteúdo? Pode-se tomar o seguinte caso como exemplo: passar de



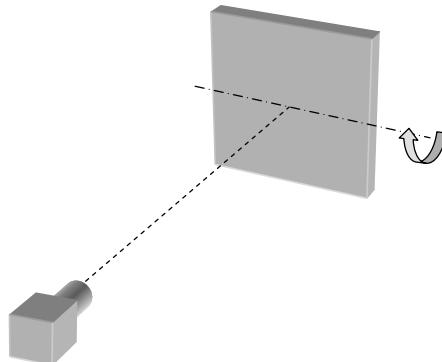
para esta:



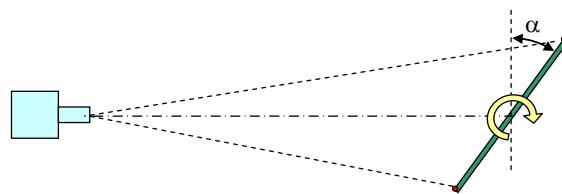
, mas não para esta:



- 3- Uma câmara de 1024x768 pixels com distância focal de 16 mm e *dot-pitch* de 160 pixels/mm (na horizontal e na vertical) está colocada a uma distância de 4 m de um painel quadrado de 50 cm de lado. De início o painel está paralelo ao plano focal mas depois reorienta-se em torno do seu eixo médio.



- a) Estabelecer a equação da transformação de perspectiva para um ponto genérico do painel. Ou seja, estabelecer as coordenadas na imagem  $Q=[u \ v \ w]$  (com  $x_{pix}=u/w$  e  $y_{pix}=v/w$ ) em função das coordenadas de um ponto genérico do painel  $P=[x_s \ y_s \ z_s]^T$ , e dos parâmetros intrínsecos da câmara. Considerar que o centro geométrico do CCD coincide com a intersecção do eixo óptico com o plano focal.
- b) A partir do factor de forma ( $ff$ ) da imagem do painel que se gera na câmara, e dos dados do problema, determinar genericamente o ângulo  $\alpha$  de inclinação do painel. Testar para  $ff=\sqrt{2}/2$ . N.B. neste contexto, o factor forma usado é o quociente da altura sobre a largura do objecto visto na imagem, e que será igual a 1 para  $\alpha=0$ ;



#### Cotação:

Questão 1 – 8 Valores. Questão 2 – 7 Valores. Questão 3 – 5 Valores.

#### Breve formulário

$$\text{Matriz intrínseca da câmara: } \mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Relações trigonométricas:  $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ ,  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$

# SISTEMAS DE VISÃO E PERCEPÇÃO INDUSTRIAL

Exame de Época de Recurso - 9 de Julho de 2008

Exercício 1 Sejam as seguintes imagens binárias de  $20 \times 20$  pixels com os pontos de fundo a preto (valor 0).

a) A imagem B foi obtida a partir da imagem A por erosão condicionada. Indicar genericamente as condições a respeitar para efectuar (ou não efectuar) a erosão de cada pixel da imagem A para obter a imagem B.

Os operadores morfológicos servem para lidar com as partes da imagem divididas em "objecto" e "fundo". A função BWMORPH() e operadores associados permitem a realização de operações morfológicas na imagem binária. Esta função usa como elemento estruturante o N8. O Matlab assume sempre "ones(3,3)" para as operações com BWMORPH().

$$N_8 = \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 1 & 1 & 1 \\ \hline 0 & 1 & 0 \\ \hline \end{array}$$

$\gg B = bwmorph(A, 'skel', 'inf')$  %. Obtensão do esqueleto (skel) da imagem A; a operação é feita infinitas (inf) vezes.

A esquelitização consiste em:

- operação de determinação do esqueleto.
- definição do esqueleto:
  - objecto filiforme (1 pixel de largura).
  - passa pelo meio do objecto.
  - preserva a topologia do objecto original.

b) Seja Y a imagem que contém apenas os quatro pontos extremos de cada um dos dois objectos de B. Indicar um filtro F de  $3 \times 3$  tal que, quando aplicado à imagem B, resultando em Z = filter2(F,B), permite obter Y = (Z == 7).

$\gg F = [-1 -1 -1; -1 8 -1; -1 -1 -1]$ ; %. Filtro de pontos isolados

$\gg Z = \text{filter2}(F, B)$ ;

$\gg Y = (Z == 7)$

c) Obter a mesma matriz Y da alínea anterior mas agora através da operação morfológica Hit-and-Miss. Indicar todos os elementos estruturantes necessários.

Operação Hit-and-Miss:  $\rightarrow$  é usada na busca de padrões de pixels.

$\rightarrow$  recebe a dois elementos estruturantes:  $A \otimes (B, C) = (A \ominus B) \cap (A^c \ominus C)$

$$\text{Hit1} = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array}$$

$$\text{Hit2} = \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

$$\text{Hit3} = \begin{array}{|c|c|c|} \hline 0 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array}$$

$$\text{Hit4} = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 0 \\ \hline \end{array}$$

$$\text{Miss1} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 0 & 1 \\ \hline 1 & 1 & 0 \\ \hline \end{array}$$

$$\text{Miss2} = \begin{array}{|c|c|c|} \hline 0 & 1 & 1 \\ \hline 1 & 0 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$$\text{Miss3} = \begin{array}{|c|c|c|} \hline 1 & 1 & 0 \\ \hline 1 & 0 & 1 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

$$\text{Miss4} = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 0 & 1 \\ \hline 0 & 1 & 1 \\ \hline \end{array}$$

$\gg H1 = bwhitmiss(B, \text{Hit1}, \text{Miss1})$ ;

$\gg H2 = bwhitmiss(B, \text{Hit2}, \text{Miss2})$ ;

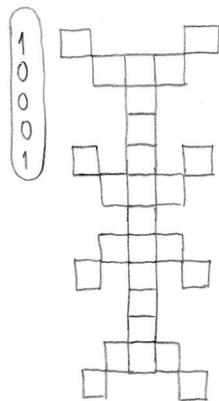
$\gg H3 = bwhitmiss(B, \text{Hit3}, \text{Miss3})$ ;

$\gg H4 = bwhitmiss(B, \text{Hit4}, \text{Miss4})$ ;

$\gg HH = H1 + H2 + H3 + H4$

A imagem C foi obtida pelo seguinte modo:  $C = B \oplus S_H$  onde  $S_H = \boxed{1 \ 1 \ 1}$ ; de modo similar, tem-se que  $D = B \oplus S_V$ .

Determinar e representar  $S_V$ . (Obs: a origem do elemento estruturante é o seu centro geométrico).



Então temos que  $S_V =$

1
0
0
0
1

) Indicar e demonstrar a veracidade da seguinte afirmação: "a operação de esqueletização de uma imagem binária preserva sempre os centros de massa dos diversos objectos numa imagem".

A esqueletização consiste em:

- operação de determinação do esqueleto.
- definição do esqueleto:  $\rightarrow$  objecto filiforme (1 pixel de largura).
  - $\rightarrow$  passa pelo meio do objecto.
  - $\rightarrow$  preserva a topologia do objecto original.

} Logo a afirmação é VERDADEIRA.

) Qual o resultado da operação de abertura morfológica da imagem D com o elemento estruturante de defeito [ones(3,3)]? Justificar a resposta.

$$SE = \text{ones}(3,3);$$

$$\text{Imagen} = \text{imopen}(C, SE)$$

/. Elemento estruturante

$$/. \text{Abertura morfológica} \longrightarrow C \circ SE = (C \ominus SE) \oplus SE$$

$\begin{matrix} \swarrow & \searrow \\ \text{DILATACÃO} & \\ \text{EROSÃO} & \end{matrix}$

O resultado final será uma imagem de fundo branco, uma vez que a abertura morfológica consiste numha erosão seguida de uma dilatação. Como em D, só temos imagens com 1 pixel de largura (esqueleto), os objectos desaparecem logo na erosão.

Exercício 2 Seja a imagem seguinte, T, que contém 3 sinais de trânsito onde a cor de cada um é de cor vermelha saturada, o fundo é branco e os símbolos internos são de cor preta. A ordem ou distribuição dos três sinais pode ser variável na imagem e diferente de representado. A linha do caixilho não faz parte da imagem, serve apenas para delimitar a área a processar.

a) Quais os passos necessários para separar os três sinais em três imagens distintas A,B e C também no formato RGB?

```

>> A = imread('Sinais.png');
>> A = im2double(A);
>> B = im2bw(A);
>> B = ~B;
>> C = imfill(B,'holes');
>> [L num] = bwlabel(C);
>> Sinal1 = A;
>> Sinal2 = A;
>> Sinal3 = A;
    
```

(...) (continuação)

```

>> Sinal1(~repmat((L)==1,[1 1 3])) = 0;
>> Sinal2(~repmat((L)==2,[1 1 3])) = 0;
>> Sinal3(~repmat((L)==3,[1 1 3])) = 0
    
```

a) Admitindo que se consegue obter as máscaras mask1, mask2 e mask3 para as três imagens, como se determina qual delas é a máscara da imagem com o sinal circular?

Deve-se usar as propriedades dos objectos (REGIONPROPS). Tendo os objectos separados, é possível estudar as propriedades individuais, tais como:

área  
perímetro

rectângulo envolvente

eixos de elipse envolvente

extremos (pixels extremos)

fator de forma

As propriedades a usar neste caso seriam a área e o perímetro para posteriormente achar o fator de forma.

> rp = regionprops (L,'Area','Perimeter');

> idxA = find([rp.Area] > ...);

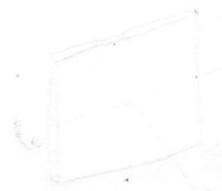
> idxB = find([rp.Perimeter] > ...);

> ff = (4\*pi\*A/P^2) % Calcular o fator de forma.

> A = ismember(L,idxA) & ismember(L,idxB)

c) Quais as operações para, a partir de A, B e C, extrair de cada uma o respetivo símbolo? Ou seja, os seguintes elementos:

X1S = ; X2S = ; CS = .



>> A = imread('Sinais.png');

>> A = im2double(A);

>> HSV = rgb2HSV(A); % Conversão da imagem A para o modelo HSV

>> H = HSV(:,:,1); % Tom

>> S = HSV(:,:,2); % Saturação

>> V = HSV(:,:,3); % Valor (intensidade)

>> Red\_mask = (H<... | H>...) & S>... & V>...;

>> B = regionprops(Red\_mask,[1 1 3])

d) Admitindo que se obtém as imagens dos símbolos da alínea anterior, como se podem distinguir os sinais triangulares entre si, ou seja, X1S de X2S? Indicar todas as operações necessárias.

1) Se por alguma razão um dos símbolos surgisse no borda da imagem, que operações seriam necessárias para o eliminar, incluindo todo o seu conteúdo? Pode-se tomar o seguinte caso como exemplo: passar de  para esta:



, mas não para esta:



$A = \text{imread}('Sinais 2.png');$

$A = \text{im2double}(A);$

$B = \text{im2bw}(A);$

$B = \sim B;$

$C = \text{imfill}(B, 'holes');$

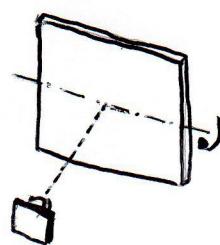
$C(1:\text{end}, 1, :) = 1;$

$C = \text{imfill}(C, 'holes');$

$D = \text{imclearborder}(C);$

$A(\sim \text{repmat}((D == 1, [1 1 3]))) = 0$

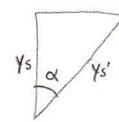
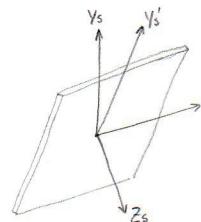
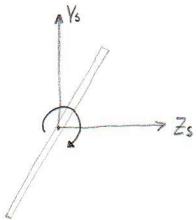
Exercício 3 Uma câmara de  $1024 \times 768$  pixels com distância focal de 16 mm e dot-pitch de 160 pixels/mm (na horizontal e na vertical) está colocada a uma distância de 4m de um painel quadrado de 50 cm de lado. De inicio o painel está paralelo ao plano focal mas depois rocionta-se em torno do seu eixo médio.



1) Estabelecer a equação da transformação de perspectiva para um ponto genérico do painel. Ou seja, estabelecer as coordenadas na imagem  $Q = [u \ v \ w]$  (com  $x_{pix} = u/w$  e  $y_{pix} = v/w$ ) em função das coordenadas de um ponto genérico do painel  $P = [x_s \ y_s \ z_s]^T$ , e dos parâmetros intrínsecos da câmara. Considerar que o centro geométrico do CCD coincide com a intersecção do eixo óptico com o plano focal.

$$\begin{cases} x_{pix} = x_0 + k_x \cdot x_i \\ y_{pix} = y_0 + k_y \cdot y_i \end{cases} \Rightarrow \begin{cases} x_{pix} = x_0 + k_x \cdot f \cdot (x_s/z_s) \\ y_{pix} = y_0 + k_y \cdot f \cdot (y_s/z_s) \end{cases} \Rightarrow \begin{cases} z_s \cdot x_{pix} = z_s \cdot x_0 + d_x \cdot x_s \\ z_s \cdot y_{pix} = z_s \cdot y_0 + d_y \cdot y_s \end{cases} \Rightarrow \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} d_x & 0 & x_0 \\ 0 & d_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$$

$$\boxed{\frac{x_s}{z_s} = \frac{x_i}{f}; \frac{y_s}{z_s} = \frac{y_i}{f}}$$



$$\Rightarrow \cos \alpha = \frac{y_s}{y_s'} \Leftrightarrow y_s' = \frac{y_s}{\cos \alpha}$$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} d_x & 0 & x_0 \\ 0 & d_y & y_0/\cos \alpha \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_s \\ y_s/\cos \alpha \\ z_s \end{bmatrix} \Rightarrow \begin{bmatrix} x_{pix} \\ y_{pix} \\ 1 \end{bmatrix} = \begin{bmatrix} d_x & 0 & x_0 \\ 0 & d_y & y_0/\cos \alpha \\ 0 & 0 & z_s \end{bmatrix} \begin{bmatrix} x_s \\ y_s/\cos \alpha \\ z_s \end{bmatrix} \cdot \frac{1}{z_s}$$

2) A partir do factor de forma (ff) da imagem do painel que se gera na câmara, e dos dados do problema, determinar genericamente o ângulo  $\alpha$  de inclinação do painel. Testar para  $ff = \sqrt{2}/2$ . N.B. neste contexto, o factor de forma usado é o quociente da altura sobre a largura da objecto visto na imagem, e que será igual a 1 para  $\alpha = 0$ .

$$ff = \frac{y_{pix}}{x_{pix}} \Leftrightarrow ff = \frac{[d_y \cdot (y_s/\cos \alpha) + (y_0/\cos \alpha) \cdot z_s] / z_s}{[d_x \cdot x_s + x_0 \cdot z_s] / z_s}$$



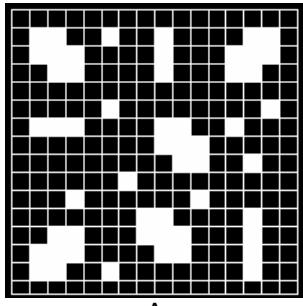
# Sistemas de Visão e Percepção Industrial

Exame de Época Normal

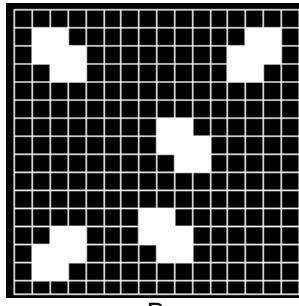
23 de Junho de 2008

Mestrado Integrado em Engenharia Mecânica  
Mestrado em Engenharia de Automação Industrial

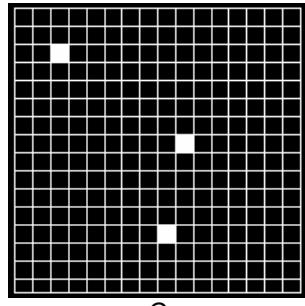
- 1- Sejam as seguintes imagens binárias de 16X16 pixels onde os pontos de objecto são brancos (valor 1). Nas questões, as letras A, B, C e D referem-se às respectivas imagens.



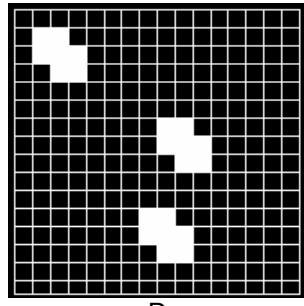
A



B



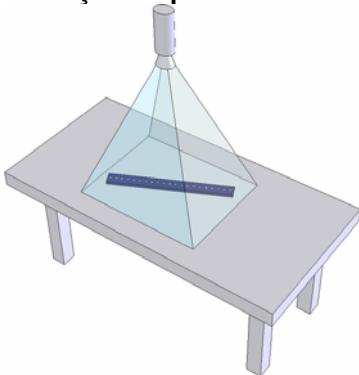
C



D

- Determinar um filtro 3x3 através do qual se possa obter a imagem B a partir da imagem A, e indicar de forma resumida as operações necessárias para levar a cabo o procedimento.
- Determinar os elementos estruturantes da operação morfológica *Hit-and-miss* para obter a imagem C a partir da imagem B, ou seja, indicar  $S=(S_1, S_2)$  em  $C = B \otimes S$ .
- Para obter D a partir de C usou-se a seguinte expressão  $D = C \oplus S$  (dilatação). Determinar e representar S.
- Qual seria o resultado X da operação  $X = TopHat(A, S) = A - (A \circ S)$  onde  $S=ones(3,3)$ ? O operador representado é a “abertura” morfológica.
- Se a partir da imagem C se quisesse criar uma imagem com 3 linhas horizontais de um pixel em toda a largura da imagem, qual seria a operação morfológica e o menor elemento estruturante possível que se poderia usar?

- 2- Uma régua de comprimento indefinido, mas com furos de centímetro em centímetro, está pousada horizontalmente sobre uma mesa e é inteiramente visível através de uma câmara com as seguintes características: distância focal de 8 mm, *dot-pitch* do CCD de 80 pixels/mm e dimensão de imagem de 640X480 pixels. A câmara está colocada verticalmente sobre a mesa e com o plano focal paralelo ao tampo. A espessura da régua é desprezável e pode ter qualquer orientação no plano da mesa. A imagem na câmara (A) tem 256 níveis de cinzento.



Sistema físico

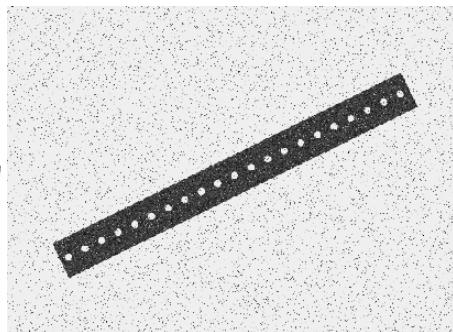


Imagen A (8 bits)

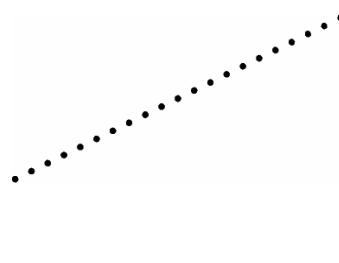
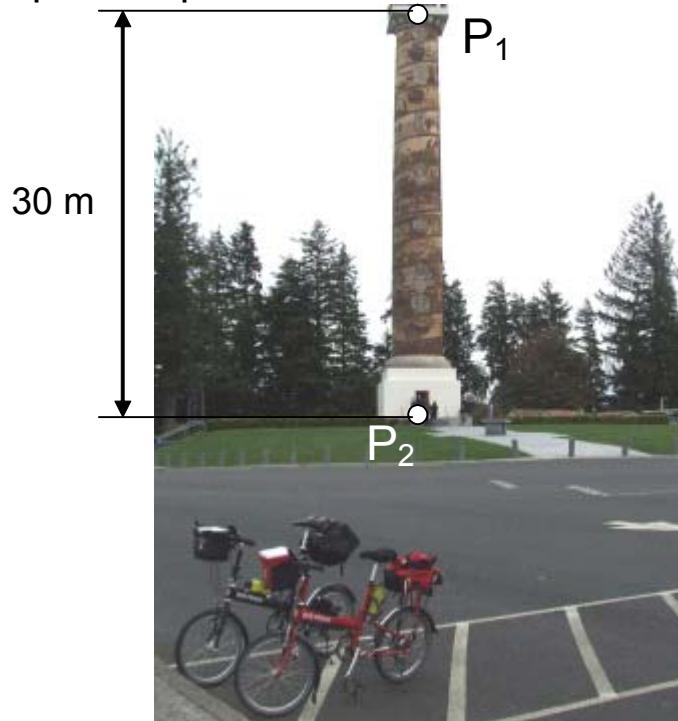


Imagen B (binária)

- Indicar e explicar que operações se deveriam aplicar genericamente sobre a imagem A para obter a imagem B.
- A partir da imagem B, indicar um procedimento para contar o número de furos, mas sem invocar funções de alto nível como o “bwlabel” ou similar do MatLab ou outros. Ou seja, deve apenas usar-se procedimentos básicos de morfologia e manipulação de pixels de imagem.
- Indicar um procedimento algorítmico para obter as coordenadas dos centros de massa dos furos dos extremos da régua. Pode complementar-se a resolução com instruções Matlab, se necessário.
- Sabendo que as coordenadas, em pixels, se medem a partir do canto superior esquerdo da imagem, se as coordenadas encontradas para os furos nos extremos da régua tiverem os seguintes valores em pixels  $P1(86,368)$  e  $P2(558,128)$ , determinar a que altura da mesa está o plano focal da câmara.

- 3- A coluna Astoria nos EUA tem cerca de 30 metros de altura entre os pontos assinalados na imagem como P<sub>1</sub> e P<sub>2</sub>. A imagem tem dimensões 1596x2160 possuindo um *dot-pitch* de 216 pixels/mm. A fotografia foi tirada com a câmara colocada com o seu eixo óptico paralelo ao plano do pavimento e com o plano focal paralelo à coluna. A distância focal usada foi de 24 mm.



- a) Se as coordenadas em pixels de P<sub>1</sub> e P<sub>2</sub> forem P<sub>1</sub>(792, 60) e P<sub>2</sub>(792, 1236), com o pixel (1,1) correspondendo ao canto superior esquerdo da imagem, determinar as coordenadas físicas (em milímetros) dos respectivos pontos no CCD da câmara admitindo que o centro geométrico do CCD é usado para origem do seu sistema de coordenadas (0,0).
  - b) Com base nos dados, determinar a que distância da coluna foi tirada a fotografia.
  - c) A que distância do solo estava a câmara na altura da fotografia?
  - d) Descrever um procedimento (com o detalhe necessário a uma implementação, por exemplo, em Matlab do qual se podem usar as funções) que permita contar a área em pixels do espaço verde da relva em torno da coluna. Sugestão: admitir que a saturação do verde da relva é superior ao de qualquer outra região e é maior de 75%.
- 4- Seja uma imagem binária em que os pixels brancos têm coordenadas genéricas ( $x_i, y_i$ ). Considere-se um caso particular de uma imagem com apenas dois pixels P<sub>1</sub>=( $x_1, y_1$ ) e P<sub>2</sub>=( $x_1, y_1$ ).
- a) Demonstrar que a transformada de Hough dessa imagem num espaço de parâmetros para linhas rectas ( $\rho, \theta$ ) são duas sinusóides nesse espaço,  $\rho = k_i \sin(\theta + \varphi_i)$   $i = 1, 2$ , e determinar as suas amplitudes  $k_i$  e fases  $\varphi_i$  em função das coordenadas de P<sub>1</sub> e P<sub>2</sub>.
  - b) Se P<sub>1</sub>=(5, 10) e P<sub>2</sub>=(10, 5) determinar em que ponto do espaço ( $\rho, \theta$ ) se intersectam essas sinusóides.

#### Cotação:

Questão 1 – 6 Valores. Questão 2 – 5 Valores. Questão 3 – 5 Valores. Questão 4 – 4 Valores

#### Breve formulário

Matriz intrínseca da câmara.  $\mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Equação polar da recta:  $x \cos \theta + y \sin \theta = \rho$

Relações trigonométricas:  $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$ ,  $\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$

# SISTEMAS DE VISÃO E PERCEPÇÃO INDUSTRIAL

Exame de Época Normal - 23 de Junho de 2008

**Exercício 1** Sejam as seguintes imagens binárias de  $16 \times 16$  pixels onde os pontos de objecto são brancos (valor 1). Nas questões, as letras A, B, C & D referem-se às respectivas imagens.

a) Determinar um filtro  $3 \times 3$  através do qual se possa obter a imagem B a partir da imagem A, e indicar de forma resumida as operações necessárias para levar a cabo o procedimento.

Princípio de utilização de um filtro: o novo pixel é função do pixel original e do filtro.

$$g(x,y) = \sum_{i=1}^N w_i \cdot p_i$$

↳ pertence à vizinhança do ponto em causa.  
↳ pode ser o valor directo do novo pixel, ou ser usado de forma indirecta e assim permitir calcular o seu novo valor.

Os coeficientes de um filtro  $3 \times 3$  são:

w <sub>1</sub>	w <sub>2</sub>	w <sub>3</sub>
w <sub>4</sub>	w <sub>5</sub>	w <sub>6</sub>
w <sub>7</sub>	w <sub>8</sub>	w <sub>9</sub>

» Filtro =  $[1\ 1\ 1; 1\ 1\ 1; 1\ 1\ 1]$ ;

» aux = filter2(Filtro, A);

/. Usa o filtro definido na imagem A

» B = A;

» B(aux <= 3) = 0

/. Mudança de variável

b) Determinar os elementos estruturantes da operação morfológica Hit-and-Miss para obter a imagem C a partir da imagem B, ou seja, indicar  $S = (S_1, S_2)$  em  $C = B \otimes S$ .

Operação Hit-and-Miss: ↳ é usada na busca de padrões de pixels.

↳ recebe dois elementos estruturantes:  $A \otimes (B, C) = (A \ominus B) \cap (A^c \ominus C)$

Hit =

1	1	0
1	1	1
0	1	1

Miss =

0	0	1
0	0	0
1	0	0

» Hit =  $[1\ 1\ 0; 1\ 1\ 1; 0\ 1\ 1]$ ; /. Partes do padrão que devem ter valor 1.

» Miss =  $[0\ 0\ 1; 0\ 0\ 0; 1\ 0\ 0]$ ; /. Partes do padrão que devem ter valor 0.

» C = bwhitmiss(B, Hit, Miss)

c) Para obter D a partir de C usou-se a seguinte expressão  $D = C \oplus S$  (dilatação). Determinar e representar S.

Dilatação: em geral aumenta as dimensões do objecto.

» SE =  $[1\ 1\ 0; 1\ 1\ 1; 0\ 1\ 1]$ ; - /. Elemento estruturante

» D = imdilate(C, SE) /. Dilatação de C com base no elemento estruturante SE.

SE =

1	1	0
1	1	1
0	1	1

1) Qual seria o resultado X da operação  $X = \text{TopHat}(A, S) = A - (A \circ S)$  onde  $S = \text{ones}(3, 3)$ ? O operador representado é a "abertura" morfológica.

Operação Top-Hat:   
→ diferença entre a imagem original e a sua abertura.   
→ usada para realçar os baixos contrastes.   
→ o efeito principal é o de destacar localmente os pixels brancos com fundos variáveis.

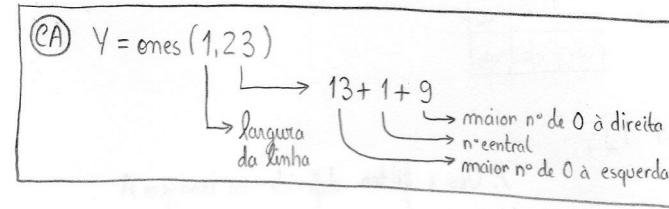
ABERTURA (opening):  $A \circ B = (A \ominus B) \oplus B$    
FECHO (closing):  $A \bullet B = (A \oplus B) \ominus C$   $\rightarrow$  são ambos idempotentes  $\begin{cases} (A \circ B) \circ B = A \circ B \\ (A \bullet B) \bullet B = A \bullet B \end{cases}$

•  $A \circ S \rightarrow \underbrace{\text{erosão}}_{\text{imagem}} + \underbrace{\text{dilatação em } S}_{\text{imagem preta}} \quad \xrightarrow{\text{ENTÃO}} A - \text{"Imagem preta"} = A$

2) Se a partir da imagem C se quisesse criar uma imagem com 3 linhas horizontais de um pixel em toda a largura da imagem, qual seria a operação morfológica e o menor elemento estruturante possível que se poderia utilizar?

⇒  $Y = \text{ones}(1, 23);$

⇒ Imagem = imdilate(C, Y)



JOTA: no MATLAB teríamos  $Y = \text{ones}(1, 26)$  em que 26 é o maior intervalo de 0  $\Rightarrow 13 \times 2 = 26$ .

**Exercício 2** Uma régua de comprimento indefinido, mas com furos de 1cm em 1cm, está pousada horizontalmente sobre uma mesa e é inteiramente visível através de uma câmara com as seguintes características: distância focal de 8mm, dot-pitch do CCD de 80 pixels/mm e dimensão de imagem de  $640 \times 480$  pixels. A câmara está colocada verticalmente sobre a mesa e tem o plano focal paralelo ao tempo. A espessura da régua é desprezável e pode ter qualquer orientação no plano da mesa. A imagem na câmara (A) tem 256 níveis de cinzentos.

1) Indicar e explicar que operações se deveriam aplicar genericamente sobre a imagem A para obter a imagem B.

```
> A = imread('ImagemA.bmp');           % leitura da imagem A
> A = im2double(A);                   % binarização da imagem A
> BW = im2bw(A, graythresh(A));       % procura a melhor binarização
> B = medfilt2(BW);                  % filtro de mediana
> C = bwmorph(BW, 'shrink', 'inf');   % encontra objectos para pontos
SE = [1 1 1; 1 -8 1; 1 1 1];          % elemento estruturante (filtro de pontos isolados)
F = filter2(SE, C);                   % usa o filtro definido na imagem C; filtro de pontos isolados
D = (abs(F)) == 8;                     % reconstroi a imagem D
E = sum(sum(D));                      % soma os pixels da imagem D
F = imreconstruct(D, A);              % reconstrói a imagem B
F = ~F                                % negação de F
```

- % Reconstrução morfológica da imagem D sob a imagem A  
% Negação de F.

A partir da imagem B, indicar um procedimento para contar o número de furos, mas sem invocar funções de alto nível como BWLABEL() ou similar do Matlab ou outros. Ou seja, deve apenas usar-se procedimentos básicos da morfologia e manipulação de pixels de imagem.

```

A = imread ('ImagemA.bmp');
A = im2double (A);
BW = im2bw (A, graythresh (A));
B = medfilt2 (BW);
C = bwmorph (BW, 'shrink', 'inf');
Furos = sum (sum (C == 0))

```

) Indicar um procedimento algorítmico para obter as coordenadas dos centros de massa dos furos dos extremos da régua. Pode complementar-se a resolução com instruções Matlab, se necessário.

Denominando por  $\begin{cases} r \rightarrow \text{coordenadas das linhas de cada ponto branco na imagem C} \\ c \rightarrow \text{coordenadas das colunas de cada ponto branco na imagem C} \end{cases}$

```

>> [r c] = find(C);
>> P1 = [r(1) c(1)]; % Primeiro ponto
>> Pend = [r(end) c(end)] % Último ponto

```

d) Sabendo que as coordenadas, em pixels, se medem a partir do canto superior esquerdo da imagem, se as coordenadas encontradas para os furos nos extremos da régua tiverem os seguintes valores em pixels  $P_1(86, 368)$  e  $P_2(558, 128)$ , determinar a que altura da mesa está o plano focal da câmara.

### MÉTODO 1:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \quad \text{em que } K = \begin{bmatrix} 640 & 0 & 320 & 0 \\ 0 & 640 & 240 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(P)  $\alpha_x = f \cdot k_x = 8.80 = 640$   
 $\alpha_y = f \cdot k_y = 8.80 = 640$   
 $x_0 = 640/2 = 320$   
 $y_0 = 480/2 = 240$

↳ coordenadas reais

→ matriz K (matriz própria da câmara)

→ coordenadas homogêneas da imagem

Sabe-se então que:  $\begin{cases} u = z_s \cdot x_{pix} \\ v = z_s \cdot y_{pix} \\ w = z_s \end{cases} \Rightarrow \begin{cases} u = (k_x \cdot f) \cdot x_s + x_0 \cdot z_s \\ v = (k_y \cdot f) \cdot y_s + y_0 \cdot z_s \\ w = z_s \end{cases}$

Para  $P_1$ :

$$\begin{bmatrix} 86 \\ 368 \\ 1 \end{bmatrix} = \begin{bmatrix} 640 & 0 & 320 \\ 0 & 640 & 240 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s1} \\ y_{s1} \\ z_s \end{bmatrix} \Rightarrow \begin{cases} x_{s1}/z_s = -0,37 \\ y_{s1}/z_s = 0,2 \end{cases}$$

Para  $P_2$ :

$$\begin{bmatrix} 558 \\ 128 \\ 1 \end{bmatrix} = \begin{bmatrix} 640 & 0 & 320 \\ 0 & 640 & 240 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s2} \\ y_{s2} \\ z_s \end{bmatrix} \Rightarrow \begin{cases} x_{s2}/z_s = 0,37 \\ y_{s2}/z_s = -0,175 \end{cases}$$

CA) O tamanho da régua D é dado por:

$$\sqrt{86^2 + 368^2} = 377,91$$

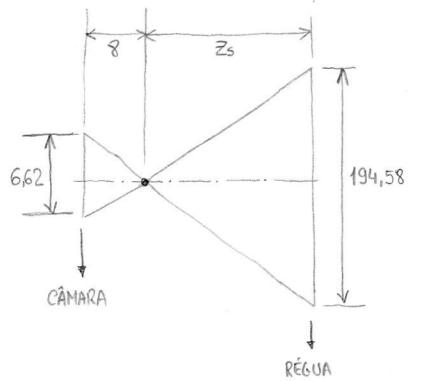
$$\sqrt{558^2 + 128^2} = 572,49$$

$$D = 572,49 - 377,91 = 194,58 \text{ mm}$$

MÉTODO 2

$$P_1 \rightarrow (86, 368) \quad P_2 \rightarrow (558, 128) \quad D = \sqrt{(86 - 558)^2 + (368 - 128)^2} \Leftrightarrow D = 529,5 \text{ pixels}$$

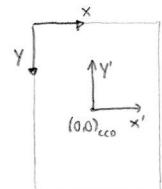
$$\text{Dividindo } D \text{ pelo valor do CCD temos: } \frac{529,5 \text{ pixels}}{80 \text{ pixels/mm}} = 6,62 \text{ mm}$$



$$6,62 \text{ mm} \longrightarrow 8 \text{ mm} \quad z_s = \frac{194,58 \times 8}{6,62} \Leftrightarrow z_s = 235,14 \text{ mm}$$

Exercício 3) A coluna Astoria nos EUA tem cerca de 30m de altura entre os pontos assinalados na imagem como P1 e P2. A imagem tem dimensões 1596 x 2160 pixels com um dot-pitch de 216 pixels/mm. A fotografia foi tirada com a câmara colocada com o seu eixo óptico paralelo ao plano de pavimento e com o plano focal paralelo à coluna. A distância focal usada foi de 24mm.

1) Se as coordenadas em pixels de P1 e P2 forem P1(792,60) e P2(792,1236), com o pixel (1,1) correspondendo ao canto superior esquerdo da imagem, determinar as coordenadas físicas (em mm) dos respectivos pontos no CCD da câmara admitindo que o centro geométrico do CCD é usado para origem do seu sistema de coordenadas (0,0).



$$(0,0)_{CCD} = \left( \frac{1596}{2}, \frac{2160}{2} \right) = (798, 1080)$$

$$P_1 \rightarrow \frac{\begin{bmatrix} x - x' \\ y' - y \end{bmatrix}}{216} = \frac{\begin{bmatrix} 792 - 798 \\ 1080 - 60 \end{bmatrix}}{216} = \begin{bmatrix} -0,027 \\ 4,722 \end{bmatrix} \text{ mm} \quad \left. \begin{array}{l} \\ \end{array} \right\} \text{Coordenadas no CCD}$$

$$P_2 \rightarrow \frac{\begin{bmatrix} x - x' \\ y' - y \end{bmatrix}}{216} = \frac{\begin{bmatrix} 792 - 798 \\ 1080 - 1236 \end{bmatrix}}{216} = \begin{bmatrix} -0,027 \\ -0,722 \end{bmatrix} \text{ mm}$$

Com base nos dados, determinar a que distância da coluna foi tirada a fotografia.

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{s1} \\ y_{s1} \\ z_s \\ 1 \end{bmatrix}$$

$$\text{em que } K = \begin{bmatrix} 5184 & 0 & 798 & 0 \\ 0 & 5184 & 1080 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(P)  $\alpha_x = f \cdot k_x = 24.216 = 5184$   
 $\alpha_y = f \cdot k_y = 24.216 = 5184$   
 $x_0 = 1596/2 = 798$   
 $y_0 = 1080/2 = 1080$

coordenadas reais

→ matriz  $K$  (matriz própria da câmara)

→ coordenadas homogêneas da imagem

obse-se então que:

$$\begin{cases} u = z_s \cdot x_{pix} \\ v = z_s \cdot y_{pix} \\ w = z_s \end{cases} \Leftrightarrow \begin{cases} u = (k_x \cdot f) \cdot x_s + x_0 \cdot z_s \\ v = (k_y \cdot f) \cdot y_s + y_0 \cdot z_s \\ w = z_s \end{cases}$$

Para P1:

$$z_s \begin{bmatrix} 792 \\ 60 \\ 1 \end{bmatrix} = \begin{bmatrix} 5184 & 0 & 798 \\ 0 & 5184 & 1080 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s1} \\ y_{s1} \\ z_s \end{bmatrix} \Leftrightarrow \begin{cases} x_{s1} = -0,0012 z_s \\ y_{s1} = -0,2 z_s \end{cases}$$

Para P2:

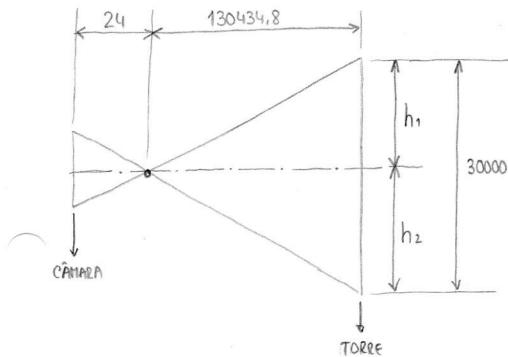
$$z_s \begin{bmatrix} 792 \\ 1236 \\ 1 \end{bmatrix} = \begin{bmatrix} 5184 & 0 & 798 \\ 0 & 5184 & 1080 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{s2} \\ y_{s2} \\ z_s \end{bmatrix} \Leftrightarrow \begin{cases} x_{s2} = -0,0012 z_s \\ y_{s2} = 0,03 z_s \end{cases}$$

$$D = \sqrt{(x_{s1} - x_{s2})^2 + (y_{s1} - y_{s2})^2}$$

$$\Leftrightarrow (30 \times 10^3)^2 = (-0,0012 z_s - (-0,0012 z_s))^2 + (-0,2 z_s - 0,03 z_s)^2$$

$$\Leftrightarrow z_s = 130434,8 \text{ mm}$$

2) A que distância do solo estava a câmara na altura da fotografia?



$$\begin{cases} \frac{4,51}{24} = \frac{h_1}{130434,8 + 24} \\ \frac{0,7}{24} = \frac{h_2}{130434,8 + 24} \end{cases} \Rightarrow \begin{cases} h_1 = 24515 \text{ mm} \\ h_2 = 3805 \text{ mm (altura ao solo)} \end{cases}$$

1) Descrever um procedimento (com o detalhe necessário a sua implementação, por exemplo, em Matlab da qual se podem usar as funções) que permita contar a área em pixels do espaço verde da relva em torno da coluna. Sugestão: admitir que a saturação do verde da relva é superior ao de qualquer outra região e é maior de 75%.

```

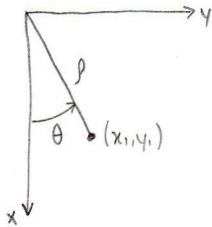
>> A = imread('ficheiro.png');
>> A = im2double(A);
>> B = rgb2HSV(A); % Conversão da imagem A para o modelo HSV
>> H = B(:,:,1); % Tom
>> S = B(:,:,2); % Saturação
>> V = B(:,:,3); % Valor (intensidade)
>> Verde = (H>0,2 & H<0,45 & S>0,2 & V>0,2);
>> Verde2 = regionprops(Verde, [1 1 3]);
>> A(~Verde2) = 0;
>> C = im2bw(A, graythresh(A));
>> D = sum(sum(A));

```

**Exercício 4** Seja uma imagem binária em que os pixels brancos têm coordenadas genéricas  $(x_i, y_i)$ . Considerar-se um caso particular de uma imagem com apenas dois pixels  $P1 = (x_1, y_1)$  e  $P2 = (x_2, y_2)$ .

a) Demonstrar que a transformada de Hough dessa imagem num espaço de parâmetros para linhas rectas  $(P, \theta)$  são duas sinusoides nesse espaço,  $P = K_i \cdot \sin(\theta + \varphi_i)$   $i=1,2$ , e determinar as suas amplitudes  $K_i$  e fases  $\varphi_i$  em função das coordenadas de  $P1$  e  $P2$ .

Transformada de Hough: conjunto de curvas discretizadas (uma para cada pixel branco da imagem) que representam todas as rectas possíveis que passam em cada um desses pontos.



$$P = \sqrt{x_1^2 + y_1^2}$$

$$\theta = \arctan(y_1/x_1)$$

$$\begin{aligned} x_1 \cdot \cos \theta + y_1 \cdot \sin \theta &= P \\ \Leftrightarrow x_1 \cdot \cos \theta + y_1 \cdot \sin \theta &= K_1 \cdot \sin(\theta + \varphi_1) \\ \Leftrightarrow x_1 \cdot \cos \theta + y_1 \cdot \sin \theta &= K_1 \cdot (\sin \theta \cdot \cos \varphi_1 + \cos \theta \cdot \sin \varphi_1) \end{aligned}$$

$$\Leftrightarrow \begin{cases} y_1 = K_1 \cdot \cos \varphi_1 \\ x_1 = K_1 \cdot \sin \varphi_1 \end{cases} \Leftrightarrow \begin{cases} K_1 = y_1 / \cos \varphi_1 \\ x_1 = y_1 \cdot \sin \varphi_1 / \cos \varphi_1 \end{cases} \Leftrightarrow \begin{cases} \text{---} \\ \arctan(x_1/y_1) = \varphi_1 \end{cases}$$