

Ficha de Trabalho 3 – Segmentação

Docente: Miguel Tavares Coimbra

1. Introdução à segmentação

- a. Aplique uma operação de *thresholding*, que segmente a imagem, usando como limiar k o valor 4. Represente o resultado final em forma de matriz binária.

Uma operação de *thresholding* é uma operação ponto-a-ponto que resulta num valor igual a 1 para um ponto se este tiver um valor original maior ou igual ao limiar k . Caso contrário o valor resultante é 0. Assim sendo, o resultado para os primeiros pontos da matriz seria:

- $f(0,0) = 1 \Rightarrow 1 < 4 \Rightarrow g(0,0) = 0$
- $f(1,0) = 8 \Rightarrow 8 \geq 4 \Rightarrow g(1,0) = 1$
- etc

0	0	0	0	0	0
1	1	1	1	1	1
0	1	1	1	0	0
0	1	1	1	1	0
0	0	1	1	1	0

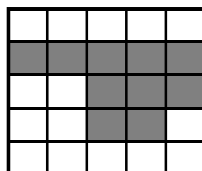
- b. Que valores de k permitem obter o mesmo resultado da alínea anterior? Motive esta escolha desenhando um histograma da imagem.

1	1	2	2	3	3
8	6	6	7	7	9
2	8	8	8	1	2
3	10	11	12	13	3
2	3	13	14	14	3

Se criarmos o histograma da imagem vemos que não há nenhum pixel com valor 4 e 5. Ao escolhermos um limiar igual a 4, sabemos que todos os valores superiores a 4 ficam no grupo de valor 1. O mesmo resultado é então obtido se escolhermos 5 ou 6. Se escolhermos 3 ou 7, já temos valores que se alteram: os 3 passam de 0 para 1, ou os 7 passam de 1 para 0. Cuidado com os ‘números mágicos’!

2. Morfologia matemática

- a. Considere o seguinte resultado de uma segmentação:



Esta alínea tem uma resposta profundamente interessante 😊

- b. Aplique uma operação de erosão morfológica, utilizando o seguinte *kernel*:



Uma operação de erosão morfológica consiste na aplicação da máscara sobre cada ponto da matriz (como a máscara tem dimensão 2x2 vamos considerar que o ponto a sobrepor com a imagem é o canto superior esquerdo), sendo o resultado igual a 1 se todos os pontos ‘iguais a 1’ da máscara estiverem sobre pontos ‘iguais a 1’ da imagem. Caso contrário o resultado é igual a zero. Ao aplicar a máscara em alguns pontos da matriz temos:

Ponto (1,1):

- A sub-matriz da imagem onde vai ser sobreposta a máscara é a zona a cinzento.

0	0	0	0	0
1	1	1	1	1
0	0	1	1	1
0	0	1	1	0
0	0	0	0	0

- Todos os pontos da sub-matriz são iguais a 1? Não. Então o resultado para este ponto (1,1) é igual a zero:

?	?	?	?	?
?	0	?	?	?
?	?	?	?	?
?	?	?	?	?
?	?	?	?	?

Para os outros pontos basta deslizar a máscara e repetir a operação. Note que, tal como nos filtros digitais, temos o problema de fronteira! (o que é isto?).

E se fosse uma dilatação morfológica? O resultado seria 1 se no mínimo um ponto igual a 1 da máscara estiver sobre um ponto igual a 1 da imagem. O resultado para o ponto (1,1) seria então igual a 1, pois 3 dos 4 pontos são iguais a 1.

Resultado final da erosão morfológica (problema de fronteira nos pixels a cinzento):

0	0	0	0	?
0	0	1	1	?
0	0	1	0	?
0	0	0	0	?
?	?	?	?	?

c. Repare na influência que o *kernel* têm no resultado desta operação, utilizando desta vez:



0	0	0	0	?
0	0	1	1	?
0	0	1	1	?
0	0	0	0	?
?	?	?	?	?

Thresholding

- Consiste em dividir uma imagem em duas zonas:
 - 1, se $f(x,y) > K$
 - 0, se $f(x,y) \leq K$
- Não é fácil de achar um número mágico K ideal!
- Provavelmente a técnica mais popular de segmentação
 - Simples
 - Razoavelmente eficaz

U.PORTO MIM 2009 - PSI - Capítulo IV – Processamento de Imagem e Vídeo 6

Dilatação, Erosão

- Implicam dois conjuntos:
 - Imagem
 - Kernel morfológico
- Dilatação (D)
 - União do *kernel* com todo o conjunto da imagem
 - Aumento da área resultante
- Erosão (E)
 - Intersecção
 - Diminuição da área resultante

U.PORTO MIM 2009 - PSI - Capítulo IV – Processamento de Imagem e Vídeo 21

3. Processamento vídeo

- a. Porque é que os standards de compressão de imagem conseguem rácios de compressão de 10 para 1, enquanto que os de compressão vídeo conseguem de 100 para 1?

A redundância temporal de um vídeo de 25 imagens por segundo é muito elevada. O que é que isto quer dizer? Quer dizer que a informação ‘nova’ que aparece numa imagem é muito reduzida dado que a maior parte desta é igual à anterior. Em vez de mandar uma imagem nova inteira mandamos apenas as partes que mudaram (diferenças), conseguindo assim grande compressão. Para melhorar isto, mandamos também ‘vectores de movimento’, que permitem dizer que partes da imagem são iguais, e que apenas estão deslocadas. Note que estes números são para standards ‘com perdas’.

- b. Calcule a taxa de dados por segundo de uma transmissão de vídeo digital com resolução 640x480, 10 imagens por segundo, a preto e branco (8 bits por pixel) e sem nenhuma compressão vídeo.

O cálculo é directo: $640 \times 480 \times 10 \times 8 = 24.57 \text{ Mbits/s} = 3 \text{ Mb/s}$

- c. Caracterize qualitativamente (baixa, alta, média) a redundância espacial das seguintes imagens:



A ‘redundância espacial’ quantifica qual a quantidade nova de informação que temos por pixel, por comparação com os seus vizinhos ou outros pontos da mesma imagem. Podemos pensar no quanto um ponto é ‘previsível’, em comparação com os seus vizinhos. Nesta perspectiva, mesmo intuitivamente, a imagem da esquerda tem a maior parte dos seus pontos negros, com alguma estrutura interessante próximo do semi-círculo. Por contraste, a da direita é riquíssima em termos de cor, texturas e formas. Podemos então arriscar que a imagem da esquerda tem alta redundância espacial (levando a boa compressão), enquanto que a da direita tem baixa (levando a má compressão).

Necessidade da compressão vídeo

- **Problema: Vídeo digital ocupa demasiado espaço!**
 - VGA: 640x480, 3 bytes por pixel -> 920KB por imagem.
 - Cada segundo de vídeo: 23 MB
 - Cada hora de vídeo: 83 GB
- **Solução: Compressão Vídeo**
 - Rádios de compressão superiores a 100:1!
- **Resultado: Explosão da popularidade do vídeo digital.**
 - DVD, máquinas digitais, web-streaming, telemóveis 3G, etc.

Incomportável!

Redundância espacial

- **Correlação entre pixels vizinhos.**
 - Ex: Uma linha branca pode ser codificada com 2 números (nr.pixels, cor).
- **Formas matemáticas de a explorar:**
 - Lossless – Ex: LZW
 - Lossy – Ex: DCT

