

# SEGMENTAÇÃO DE IMAGENS ACÚSTICAS

João Rodrigues  
Área Eng. Electrotécnica  
EST/UALG

Hans du Buf  
Vision Laboratory  
UCEH/UALG

## 1. INTRODUÇÃO

O processamento de imagem tem entre outras aplicações a melhoria a codificação e o reconhecimento de padrões, pode-se citar como exemplos de áreas onde se efectua essa aplicação: o diagnostico medico (raios X, ressonância magnética, ultra som) a detecção remota (aviões e satélites) e a geologia (análise dados sonares e sísmicos).

Muitas dessas aplicações necessitam um pre-processamento para estabelecer uma representação intermédia; uma representação que permita uma análise sintáctica/geométrica para uma descrição ou um reconhecimento, até atingirmos a visualização. A segmentação de imagem, ou de dados físicos, em duas ou mais dimensões, é considerado um passo importante nesse processamento.

A segmentação tem por objectivo a divisão da imagem em regiões homogéneas, utilizando p.ex. os níveis de cinzento, as cores, ou a textura. Em geral, este objectivo é constituído por duas partes dependentes: obter um numero reduzido de regiões finais e a melhor qualidade destas. Dado o facto de estes requisitos serem recíprocos, a segmentação é um aspecto importante na investigação. Um requisito adicional, nomeadamente no processamento do dados físicos em três dimensões, é o tempo CPU: a melhor qualidade dentro de um tempo aceitável (p.ex. para uma visualização interactiva dos dados em realidade virtual).

O trabalho apresentado neste artigo está inserido no projecto ISACS (*Integrated System for the Analysis and Characterization of the Seafloor*), financiado pela Comissão Europeia no âmbito do programa MAST III, cujo objectivo é a análise e a classificação do fundo do mar pela exploração dos dados adquiridos a partir de diferentes tipos de sonares e a sua correcta integração [1]. O cooperante Português, UALG, *Vision Laboratory*, está a desenvolver para o projecto um sistema para visualização interactiva dos dados de sonar em 3D [1,2], onde se insere este trabalho.

Em concreto este trabalho tem como objectivo testar um algoritmo de segmentação em 2D que seja robusto e rápido para servir de base para um algoritmo de segmentação em 3D, que por sua vez vai ser aplicado

aos volumes de dados do projecto ISACS (trabalho a ser apresentado numa próxima Tecnovisão). Os resultados em 2D encontram-se apresentados com imagens (de acústica submarina) do projecto ISACS (fig. 1).

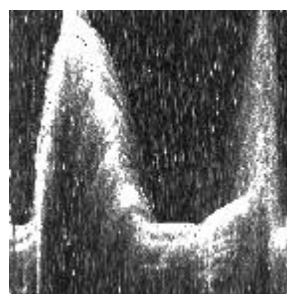


Fig. 1: Imagem "Topas 2D", imagem ISACS 2D típica.

Spann e Wilson [4] mostraram que na segmentação em 2D a aplicação de uma árvore quaternária é muito útil porque reduz o ruído e melhora a separação das classes, permitindo num determinado nível um agrupamento e uma classificação. A resolução é melhorada nível a nível pela projecção para baixo das regiões e dos pixels fronteira. Os últimos são reclassificados tendo por base a informação disponível nos níveis mais baixos na árvore. Mais tarde, Schroeter e Bigün [3] experimentaram com diferentes tipos de algoritmos de agrupamento e melhoraram o processo de estimação da fronteira pela aplicação de filtros adaptativos. Os diferentes algoritmos de agrupamento considerados deram mais ou menos o mesmo resultado e o resultado da segmentação final foi excelente.

O algoritmo aqui apresentado faz uso de uma pirâmide baseada numa árvore quaternária onde a presença de ruído vai ser reduzida a custa da resolução da imagem. Num nível alto da pirâmide é aplicado um algoritmo de agrupamento onde é criada a imagem das etiquetas. A resolução é restaurada pela projecção para baixo das etiquetas, para cada nível (que se desce), surgindo a incerteza na classificação dos pixels fronteira. Esta é minimizada por uma reclassificação dos pixels pertencentes às fronteiras em função da direcção dominante destas.

## 2. SEGMENTAÇÃO EM 2D

O algoritmo 2D pode assim ser dividido em três passos: (1) alisamento por árvore quaternária, (2)

agrupamento e classificação e (3) estimação da fronteira.

### 2.1. ÁRVORE QUATERNÁRIA

O primeiro passo do algoritmo consiste na construção de uma pirâmide chamada árvore quaternária (fig. 2). O nível da árvore quaternária acima da base consiste no conjunto de médias de blocos 2x2 de pixels não sobrepostos da imagem original, este processo pode ser repetido tantas vezes até que o nível mais alto seja alcançado e em que o pico da pirâmide tenha um só pixel cujo valor é o valor médio da imagem inteira.

Matematicamente, considere-se uma imagem  $N \times N$  (não obrigatoriamente quadrada) definida por  $I(i,j)$  para  $0 \leq i,j < N$ , onde  $N = 2^m$ . A árvore quaternária é definida por:

$$I(i,j,L) = \frac{1}{4} \sum_{a=0}^1 \sum_{b=0}^1 I(2i+a, 2j+b, L-1) \quad (1)$$

onde  $0 < L \leq m$ ;  $0 \leq i,j < 2^{m-L}$  e  $I(i,j;0) = I(i,j)$ .

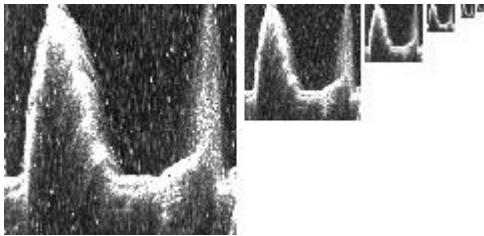


Fig. 2: Árvore quaternária de “Topas 2D”, a partir do nível  $L=0$  (esquerda).

### 2.2. AGRUPAMENTO E CLASSIFICAÇÃO

O segundo passo é a aplicação de um algoritmo de agrupamento num nível alto da pirâmide ( $L_{max}$ ), recorrendo para tal ao algoritmo que faz uso do histograma dos níveis de cinzento dos pixels: Agrupamento Centróide Local. Este algoritmo permite encontrar os centros das classes existentes na imagem no nível em que vai ser aplicado. Para a criação da imagem das etiquetas  $I_g(i,j;L_{max})$ , será aplicada uma classificação distância mínima. Este método é simples e rápido na atribuição das classes aos pixels. No entanto, esporadicamente pode classificar incorrectamente alguns pixels, mas estes serão

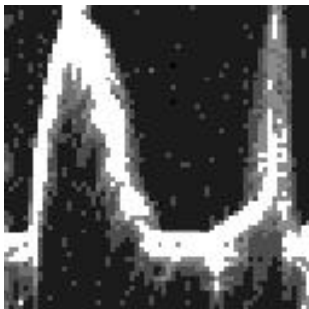


Fig. 3: Imagem após a aplicação do algoritmo de agrupamento e classificação com  $L_{max}=2$  (ampliado).

corrigidos pela restauração da conectividade espacial.

#### 2.2.1. AGRUPAMENTO CENTRÓIDE LOCAL

Entende-se por um agrupamento a selecção de um conjunto de valores de  $I$ , para os quais a gama de valores que ocorrem relativamente muitas vezes estão separados dos que ocorrem poucas vezes. O algoritmo agrupamento centróide local [4] funciona pela contínua actualização do histograma, movendo as massas de probabilidade para a posição do seu centróide local até à inexistência de alterações no histograma, permitindo assim encontrar os centros das classes existentes na imagem no nível em que é aplicado.

Para a criação da imagem das etiquetas  $I_g(i,j,k;L_{max})$  é aplicado uma classificação distância mínima, que consiste no cálculo do valor médio entre os centros das classes consecutivas. Após a determinação destes, atribuí-se a todos os pixels da imagem  $I_g(i,j,k;L_{max})$  compreendidos entre duas distâncias mínimas consecutivas a etiqueta da classe que se encontra no interior dessas distâncias (obtendo-se a imagem da fig. 3).

#### 2.2.2. RESTAURAÇÃO DA CONECTIVIDADE ESPACIAL

Após a conclusão do processo de agrupamento e classificação podem ser observados alguns pixels e grupos de pixels (regiões insignificantes) que estão espacialmente mal classificados devido à não utilização de um critério de conectividade espacial na classificação (fig. 3). Cada pixel isolado ou pertencente a uma região insignificante é atribuído à classe pertencente à sua vizinhança de 8 do pixel, a qual apresenta a distância mínima a esse pixel. A imagem das etiquetas no nível  $L=2$ , após ter sido aplicado o algoritmo de aglomeração e classificação, retirados os pixels isolados e as regiões insignificantes, está apresentada na fig. 4a. Na fig. 4b está representada uma ampliação do canto superior direito da imagem das etiquetas apresentada na fig. 3. Na fig. 4c a imagem anterior depois de terem sido retirados os pixels isolados e na fig. 4d depois de serem retiradas as regiões insignificantes.

#### 2.3. ESTIMAÇÃO DA FRONTEIRA

O último passo consiste em transformar a imagem a partir do nível mais alto da pirâmide  $L_{max}$  até ao nível com a resolução original  $L=0$ .

No nível  $L_{max}$  são determinados os pixels da fronteira. Um pixel é considerado como pixel fronteiro se pelo menos uma etiqueta na vizinhança de 8 for diferente em  $I_g(i,j;L_{max})$ . Depois  $I_g(i,j;L_{max}-1)$  é obtida pela projecção para baixo da etiqueta de cada pixel pai não-fronteiro para os seus respectivos 4 pixels filhos. Os

filhos da fronteira definem a região  $\beta$  (fig. 5) na qual vai ser efectuada a estimação da fronteira. Foram implementados dois métodos de estimação da fronteira que têm por base o trabalho de Schroeter e Bigün [3].

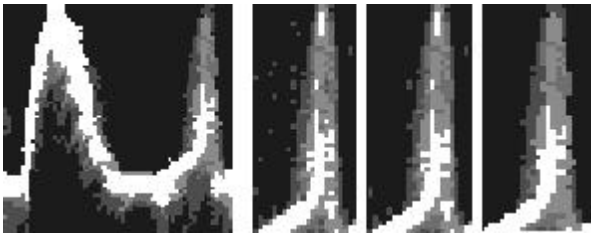


Fig. 4: Da esquerda para a direita a, b, c, d. (a) imagem das etiquetas após terem sido retiradas as regiões insignificantes e os pixels isolados. (b) pormenor do canto superior direito, (c) sem pixels isolados e (d) sem regiões insignificantes.

O **primeiro método** segue por completo o que os autores atrás referidos propõem para diminuir a incerteza no cálculo dos pixels fronteiras.

Para cada pixel da fronteira é determinado o número complexo  $z_1 = (\nabla I_g)^2 * m$ , com  $\nabla I_g$  o gradiente da imagem das etiquetas  $(\partial I_g / \partial x) + i(\partial I_g / \partial y)$  e  $*m$  a convolução com um filtro Gaussiano de tamanho  $7 \times 7$ .

A partir do argumento de  $z_1$  é obtida a orientação dominante  $\Theta$  da fronteira nessa posição.

Para cada orientação local dominante é aplicado um filtro borboleta (fig. 6a), onde:

$$rr = (1-r)/n, \quad (2)$$

$$r = \begin{cases} 0 & d \leq 0 \\ (d-2)/6 & 2 < d < 8 \\ 1 & d \geq 8 \end{cases} \quad (3)$$

com a função de dissimilaridade

$$d = |\mathbf{m}_1 - \mathbf{m}_2| / \sqrt{\mathbf{s}_1^2 + \mathbf{s}_2^2} \quad (4)$$

e onde  $\mathbf{m}_{r,2}$  e  $\mathbf{s}_{r,2}^2$  são a média e a variância de duas classes em ambos os lados da fronteira, onde  $n$  é o número de pesos da máscara diferentes de 0 e  $r$ .

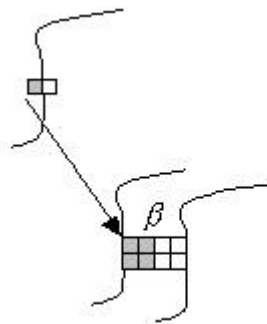


Fig. 5: Em cima: zona fronteira no nível  $L$ , em baixo: região  $\beta$  no nível  $L-1$  (imagem  $I_g$ ).

É ainda necessário rodar o filtro para a direcção local  $\Theta$  dominante. Isto é feito pela rotação da máscara da

fig. 6a ( $\Theta=0$ ) e pela redistribuição dos pesos de forma a corresponderem à nova máscara rodada.

Após a realização de testes considerando um número de direcções dominantes entre 4 e 8, verificou-se que não existem alterações significativas em termos de resultados finais para uma máscara borboleta  $3 \times 3$ . Limitou-se as direcções dominantes a 4 fig. 6b (a, b, c, d). Com esta simplificação diminuiu-se o tempo gasto no cálculo da rotação da máscara da posição horizontal para a direcção estimada  $\Theta$  e o cálculo dos novos coeficientes da máscara em função da direcção.

Os dois alvos  $r_1(i, j, k, L-1)$  e  $r_2(i, j, k, L-1)$ , que correspondem respectivamente à parte direita e esquerda da máscara da fig. 6a, são calculados para a direcção dominante. O pixel fronteiro é atribuído a classe que obtiver o mínimo valor de  $\|\mathbf{m}_\beta - r_u\|$ , com  $1 \leq u, v \leq 2$ .

Diminuiu-se o valor de  $L$  de 1 e repetiu-se o procedimento até atingir o fim da pirâmide. Pode-se observar os resultados obtidos na fig. 7a e 7b, imagem das etiquetas e imagem original com os contornos sobrepostos, respectivamente.

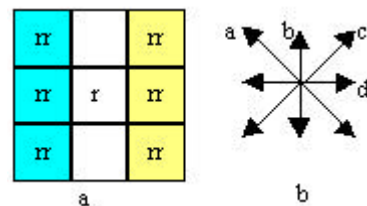


Fig. 6: (a) máscara borboleta; (b) orientações do filtro.

Com o objectivo de diminuir o tempo de estimação da fronteira implementou-se o **segundo método**, onde se determina os coeficientes do filtro  $r$  e  $rr$  aplicando (2) e (3), mas alterou-se o método de cálculo de  $\mathbf{m}_{r,2}$  e  $\mathbf{s}_{r,2}^2$  para a dissimilaridade (4), bem como o modo de aplicação do filtro borboleta aos pixels fronteira.

Para determinar  $\mathbf{m}_{r,2}$  e  $\mathbf{s}_{r,2}^2$ , determina-se as duas classes mais representativas na vizinhança de 8 na imagem  $I_g(i, j; L)$ . Para determinar essas classes foram apenas consideradas representativas aquelas que eram compostas por dois ou mais pixels. Quando existe mais do que duas destas classes com o mesmo grau de representatividade, considera-se as  $m_r$  classes (com  $C$  o número de classes com igual representatividade das “mais representativas”) para o cálculo da dissimilaridade.

Aplica-se simultaneamente todas as orientações (fig. 6b) ao mesmo tempo e para cada filtro efectua-se todas as combinações das  $m_r$  classes (duas a duas entre diferentes classes). De seguida e em separado, determina-se para cada filtro os dois alvos,  $r_1(i, j, k, L-1)$  e  $r_2(i, j, k, L-1)$ . Finalmente calcula-se a distância Euclidiana entre as classes

determinadas e todos os alvos  $\|m_c - r_u\|$ ,  $1 \leq u \leq 4 C(C-1)$ .

Ao pixel fronteiro é atribuído a etiqueta da classe que corresponde à distância mínima.

Por fim diminui-se o valor de  $L$  de 1 e repete-se o procedimento até atingir o fim da pirâmide. Pode-se observar os resultados deste método na fig. 7c (imagem das etiquetas) e na fig. 7d imagem original com os contornos delimitados.

### 3. CONCLUSÕES

Observou-se que o segundo método de estimação da fronteira simplificou o processo de reclassificação dos pixels pertencentes às fronteiras, bem como diminuiu o tempo de cálculo, evitando o cálculo da direcção da fronteira e a rotação da máscara. O segundo método é 1.25 vezes mais rápido que o primeiro, isto para a imagem da fig. 1 com as dimensões 256x256. Observou-se ainda que os dois métodos apresentam resultados muito aproximados. No entanto, o segundo método para apresenta os contornos das regiões com mais pormenor.

O algoritmo de segmentação em 2D aqui proposto com o segundo método de estimação da fronteira, apresenta-se robusto e rápido para as imagem de acústica submarina em 2D provenientes do projecto ISACS e apresenta boas perspectivas para ser extrapolado para 3D. Esta conclusão é devida à qualidade dos resultados da segmentação e também devido a utilizar pouco tempo de cálculo. Note-se que o último factor é muito importante tendo em vista a aplicação aos volumes de dados ISACS, devido às suas grandes dimensões.

### 4. REFERÊNCIAS

[1] Loke, R.E. and du Buf, J.M.H. (1998) Volumetric processing of Topas

underwater acoustic data, Proc. OCEANS'98 IEEE/OES Conf., Nice, pp. 742-746.

[2] Nikolov, S.G., Loke, R.E. and du Buf, J.M.H., (1997) Interactive 3D visualization of sonar data using VRML, Proc. Int. Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging, Rhodes, Greece, pp. 36-39.

[3] Schroeter, P. and Bigün, J. (1995) Hierarchical image segmentation by multi-dimensional clustering and orientation-adaptive boundary refinement, Pattern Recognition, Vol 28, N° 5, pp. 695-709.

[4] Spann, M. and Wilson, R. (1985) A quad-tree approach to image segmentation which combines statistical and spatial information, Pattern Recognition, Vol 18, N° 3/4, pp. 257-269.

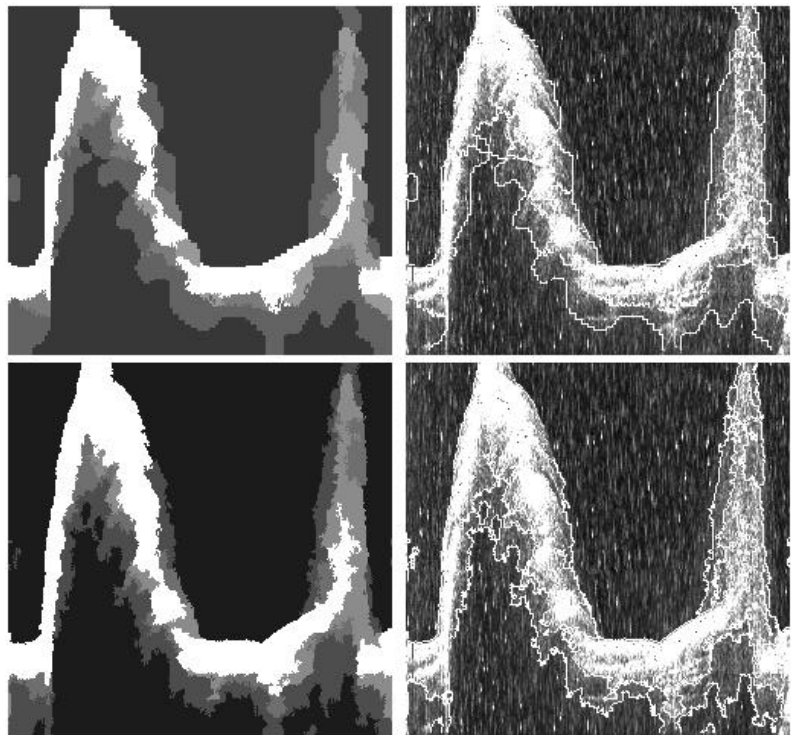


Fig. 7: Em cima da esquerda para a direita: (a) imagem das etiquetas e (b) original com os contornos delimitados para o 1º método de estimação da fronteira. Em baixo: (c) e (d) para o 2º método.