

ESTRUTURAS PIRAMIDAIRES IRREGULARES DIFUSAS EM SEGMENTAÇÃO DE IMAGENS

GILSON BRAVIANO¹
ANNICK MONTANVERT²

¹Departamento de Expressão Gráfica - CCE - UFSC
CP 476 - 88.010-970 - Florianópolis - SC
Gilson@mbox1.ufsc.br

²Laboratoire de l'Informatique du Parallélisme - ENS Lyon
6, allée d'Italie
69.364 Lyon cedex 07 - França
Annick.Montanvert@lip.ens-lyon.fr

Abstract. In opposite to classic logic, fuzzy logic enables to manipulate other truth values than “true” or “false”. Adjacency and similarity relations between elements composing an image can be represented by means of graphs. This work presents a way to introduce fuzzy edges and vertices in graphs to manipulate uncertainty and imprecision in images. This procedure provide more flexibility in segmentation techniques based on pyramid structures. The order in which the graphs elements of the pyramid levels are evaluated doesn't have any importance. All procedures are parallelisable.

1 Introdução

Dando à máquina a possibilidade de reconhecer formas, o homem certamente transpôs uma importante barreira rumo à automação de suas tarefas cotidianas. Percebe-se este automatismo em campos como o do reconhecimento da voz (tradução em tempo real de línguas estrangeiras, por exemplo) e da escrita (leitura e reconhecimentos de endereços para a triagem automática de cartas).

Pode-se citar também a visão assistida por computador, que contou com um grande avanço nos últimos anos. Dá-se, a seguir, algumas de suas aplicações, compreendendo o tratamento, a análise e a interpretação de imagens: a numeração citológica de preparações microscópicas [4], o controle do movimento de robôs, a planificação de trajetórias, a análise de imagens aéreas e do solo, o reconhecimento automático de impressões digitais, entre outras.

No processo de análise de imagens, a etapa que tem por objetivo extrair os objetos que compõem a cena chama-se segmentação. Uma das maneiras de realizar esta etapa é particionar o suporte da imagem em regiões, de maneira que cada uma delas corresponda a uma entidade na cena analisada.

Neste trabalho aborda-se técnicas piramidais de segmentação, as quais são baseadas no crescimento de regiões e em processos de contração de grafos. Propõe-se a introdução de arestas e vértices difusos nos grafos de adjacência e de similaridade que representam a imagem, objetivando-se com isso obter uma

maior flexibilidade nos processos de segmentação.

A seção 2 deste trabalho enfoca o uso de estruturas piramidais irregulares na segmentação de imagens. Na seção 3 propõe-se um modelo piramidal para a segmentação, baseado na lógica difusa. Resultados são mostrados e comparados com aqueles fornecidos pelo modelo clássico. As conclusões são apresentadas na seção 4.

2 Estruturas piramidais para a segmentação

Uma pirâmide é uma estrutura de dados capaz de representar uma imagem em diferentes níveis de resolução. Da base (que normalmente é a imagem inicial, representada por uma matriz de *pixels* contendo normalmente $2^M \cdot 2^M$ elementos) até seu ápice (que é o último nível da pirâmide), a resolução da imagem decresce, permitindo a eliminação das informações redundantes e inúteis contidas na cena.

A combinação do paralelismo, da recursividade e da multi-resolução faz da estrutura piramidal uma ferramenta de porte na representação e no tratamento de dados. O paralelismo fornece a possibilidade de melhorar consideravelmente o tempo de processamento, a multi-resolução reduz o número de operações necessárias ao tratamento da imagem, e a recursividade tem sua importância na coerência e simplificação dos algoritmos. Análises comparando os diferentes aspectos dos modelos piramidais citados acima, podem ser encontradas em [1, 6, 9].

Pode-se citar como exemplos de aplicação de

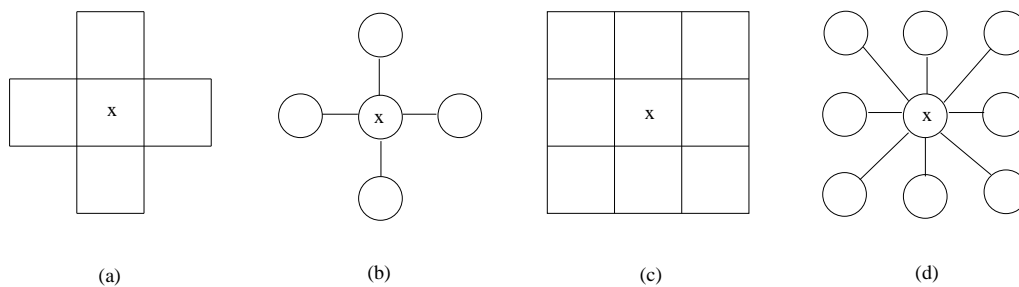


FIGURA 1: Um *pixel* x e suas (a) 4-vizinhança e (c) 8-vizinhança. Os grafos de adjacência gerados por x em relação a (b) 4-conexidade, (d) 8-conexidade.

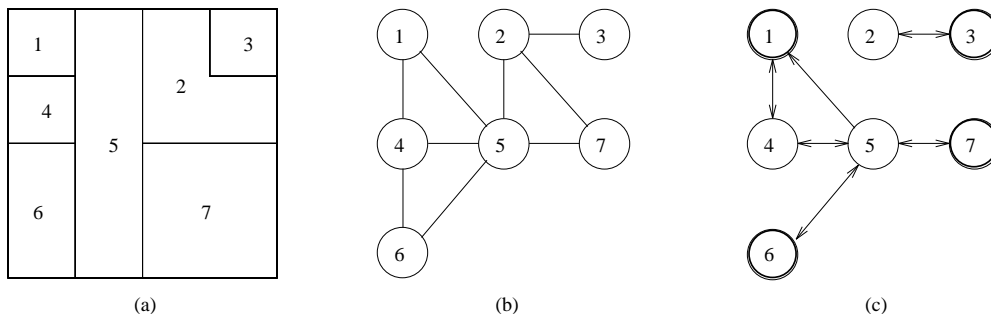


FIGURA 2: Grafos de (b) adjacência e de (c) similaridade associados à configuração de regiões de (a).

técnicas piramidais: a pirâmide laplaciana, que fornece uma decomposição frequencial da imagem [5], a discriminação de texturas [7] e a detecção de zonas de contraste [2] na imagem. Pela simplificação da informação inicial (através da redução de seu volume) a pirâmide pode ser utilizada na compressão de dados.

2.1 Notação e modelagem

A modelagem do nível zero (**base**) de uma pirâmide, usando grafos, é realizada da seguinte forma: a cada *pixel* da imagem, associa-se um vértice, e a aresta (x, y) é criada se e somente se os dois *pixels* representados pelos vértices x e y são vizinhos. Quando se trabalha com um *grid* quadrado, é possível associar a um *pixel* sua 4-vizinhança ou 8-vizinhança, conforme mostra a figura 1. Nos níveis superiores da pirâmide, cada vértice representa um conjunto conexo de vértices da base, chamado **campo receptor**. Em cada nível da pirâmide as relações de vizinhança entre os campos receptores são representadas por um grafo chamado **grafo de adjacência**. A figura 2(b) mostra o grafo de adjacência associado à configuração mostrada em 2(a).

Um nível k da pirâmide é construído a partir do nível precedente $k - 1$, cujos vértices são ditos **vivos**. Um subconjunto desses vértices será escolhido (segundo regras que serão apresentadas em 2.3), formando o conjunto dos **sobreviventes** do nível $k - 1$.

Estes sobreviventes serão então os vértices vivos do nível k . No fim do processo, as ligações entre os h níveis da pirâmide são feitas de maneira que cada elemento de um nível k esteja ligado com os elementos do nível $k - 1$ que ele representa (chamados seus **filhos**) e com um elemento do nível $k + 1$ (chamado de **pai**). Estas relações definem uma estrutura arborescente na pirâmide.

2.2 Grafo de similaridade

A utilização de um limiar de similaridade s_g no grafo de adjacência permite a eliminação das adjacências sem interesse, criando um **grafo de similaridade**. Desta maneira, mesmo que os campos receptores dos vértices x e y sejam vizinhos, a aresta (x, y) só estará presente no grafo de similaridade se a condição seguinte for satisfeita:

$$|NC(x) - NC(y)| \leq s_g \quad (1)$$

onde $NC(x)$ é o nível de cinza da região representada pelo vértice x .

A escolha de s_g é um problema difícil pois se este valor for grande, várias fusões entre regiões vizinhas tendo uma certa similaridade se produzem; o caso inverso induz à detecção de regiões sem interesse. Para evitar esses problemas, a extração de um grafo de similaridade a partir da análise mais detalhada da vizinhança de cada região é proposta em [10].

Esta extração é executada em paralelo em toda a imagem, e divide os vizinhos de cada vértice em dois grupos utilizando um valor de limiar local $s_l(x)$. O critério (1) transforma-se então em:

$$|NC(x) - NC(y)| \leq s_l(x) \quad (2)$$

Somente os vizinhos y de x que satisfizerem (2) serão aceitos. Como este critério não é simétrico, para representar as relações locais de similaridade necessita-se de um grafo orientado, como ilustra a figura 2(c).

2.3 Algoritmo de construção da pirâmide

A construção da pirâmide irregular de uma imagem é feita recursivamente e em paralelo, da base ao ápice, através de um processo de contração de grafos. Em cada nível, as etapas seguintes devem ser realizadas:

- escolha dos sobreviventes;
- os não sobreviventes escolhem um pai;
- atualização dos grafos de adjacência e de similaridade;
- atualização dos atributos dos sobreviventes.

No momento em que o grafo de similaridade no nível k da pirâmide não possui mais arcos, o algoritmo pára pois nenhuma outra fusão é possível.

A escolha dos sobreviventes é feita em paralelo, de maneira adaptativa ou estocástica ¹, segundo os seguintes critérios:

- C1:** Cada não sobrevivente deve ter ao menos um vizinho que sobrevive (para que ele possa ser representado nos níveis seguintes);
- C2:** Dois vizinhos (vértices adjacentes no grafo de similaridade) não podem sobreviver ao mesmo tempo (garantindo assim a convergência do processo).

Na teoria de grafos, um conjunto de vértices que satisfaz **C1** e **C2** ao mesmo tempo é chamado de estável. Em um grafo orientado essa escolha se traduz pela busca de um estável que seja também um conjunto dominante. Este conjunto de vértices, mostrado em destaque no grafo da figura 2(c), denomina-se núcleo. As transformações das condições **C1** e **C2** ao caso orientado, bem como a análise da existência e unicidade de soluções deste problema, são apresentadas em detalhes em [3].

¹Por comodidade de notação e visando simplificar a apresentação, não se distinguirá neste trabalho, as pirâmides irregulares estocásticas das adaptativas. Mais detalhes podem ser encontrados em [3, 8, 10].

3 Introdução da lógica difusa em estruturas piramidais

Foi visto na seção precedente que a escolha do conjunto de sobreviventes no nível k da pirâmide gera um conjunto de vértices mortos. Cada um destes vértices deve fusionar-se com um vizinho sobrevivente. Tal vizinho será então chamado de pai e receberá seus atributos para representá-lo nos próximos níveis da pirâmide. Esta escolha é de extrema importância já que o resultado do processo de segmentação depende fortemente de boas fusões.

A escolha de um pai é fácil quando o vértice morto possui apenas um vizinho sobrevivente, mas por falta de opções, esta escolha forçada pode ser errônea. Pode-se considerar também o caso em que um vértice é identicamente similar a vários vizinhos sobreviventes; nesta situação, ele é obrigado a decidir aleatoriamente. Em vista desses fatores, que podem facilmente gerar fusões indesejáveis na passagem de um nível a outro da pirâmide, seria interessante permitir aos vértices *indecisos* que adiassem sua escolha.

Esta flexibilidade em uma estrutura piramidal permitiria que os outros vértices se regroupassem, formando pouco a pouco novas regiões. Estas últimas, podendo ser mais similares à região indecisa, gerariam a realização de melhores fusões.

Esta seção introduz um modelo que permite o adiamento das decisões relativas à escolha de um pai pelos vértices mortos indecisos.

3.1 Conjuntos difusos

A noção básica da lógica difusa vem da idéia de pertinência parcial de um elemento a várias classes ao mesmo tempo.

Definição 1 *Seja X o conjunto universo. Um conjunto difuso A em X é definido por uma função de pertinência μ_A que associa para cada elemento x de X seu grau de pertinência à A em $[0,1]$:*

$$\mu_A : X \rightarrow [0, 1]. \quad (3)$$

Um conjunto é dito clássico quando μ_A assume somente os valores 0 ou 1.

3.2 Ligações difusas

Uma das restrições impostas ao modelo piramidal pela lógica clássica é a unicidade de pertinência de uma região a um campo receptor. Propõe-se que esta restrição seja relaxada. Utilizando a lógica difusa, pretende-se permitir que um vértice possa selecionar vários pais até o momento de sua tomada de decisão definitiva.

Definição 2 *Um vértice indeciso, que não está preparado para escolher um pai definitivo, é chamado vértice difuso enquanto ele permanece neste estado de indecisão.*

Seja F um vértice que morre no nível k da pirâmide e $\Gamma^k(F) = \{V_1, V_2, \dots, V_v\}$ o conjunto de seus vizinhos. Considere-se o conjunto de pais potenciais de F , $P^k(F) = \{S_1, S_2, \dots, S_p, \dots, S_t\}$, no nível k , de maneira que $s(F, S_1) = s(F, S_2) = \dots = s(F, S_p)$ e $s(F, S_p) > s(F, S_j) \quad \forall j > p$, onde s é a função de similaridade. Como cada um dos pais potenciais é um vértice sobrevivente, tem-se $P^k(F) \subseteq \Gamma^k(F)$, então $t \leq v$. Ora, F não pode escolher seu pai em $P^k(F)$ pois existem vários deles que lhe são identicamente similares, sendo este valor maximal; F torna-se então um vértice difuso.

3.3 Gestão das ligações difusas

Existem diversas maneiras de determinar o conjunto de pais $P^{k+1}(F)$ no nível $k+1$ a partir de $P^k(F)$:

- F mantém a ligação somente com os pais potenciais que sobrevivem no nível $k+1$ e que têm o mesmo grau máximo de similaridade, eliminando os que lhe são menos similares. Neste caso, o conjunto dos pais $P^{k+1}(F)$ do vértice difuso F no nível $k+1$ é tal que $P^{k+1}(F) \subseteq \{S_1, S_2, \dots, S_p\}$.
- F mantém a ligação com todos os elementos de $\{P^k(V_1), P^k(V_2), \dots, P^k(V_v)\}$ independentemente das suas similaridades. O conjunto $P^{k+1}(F)$ conterá $P^k(S_1), P^k(S_2), \dots, P^k(S_t)$, mas não necessariamente $\Gamma^k(F)$, nem $P^k(F)$. No entanto, todo vértice sobrevivente, pai de um vizinho morto de F , fará parte dos pais de F no nível $k+1$. Esta estratégia mantém, na verdade, todas as ligações possíveis entre F e os sobreviventes do nível superior.
- É possível separar os pais potenciais do nível k de um vértice difuso F em dois grupos no nível $k+1$. Esta separação pode ser realizada utilizando um valor de limiar baseado nas diferenças de níveis de cinza entre F e cada elemento de $\{P^k(S_1), P^k(S_2), \dots, P^k(S_m)\}$. Neste caso, $P^k(F) = \{S_1, S_2, \dots, S_t\}$ será um sub-conjunto de $\{P^k(S_1), P^k(S_2), \dots, P^k(S_m)\}$.

Apesar do aspecto interessante da última proposição, a determinação de um valor de limiar é necessária, além dos limiares local e global. Prefere-se a segunda opção pois um pai potencial que não é, até o momento, um dos mais similares, pode sê-lo

$P_i(5)$	1	3	6	7
NC	100	135	110	115
$\mu_{P_i}(5)$ - equ. 5	0.255	0.240	0.255	0.250
$\mu_{P_i}(5)$ - equ. 6	0.300	0.133	0.300	0.267

TABELA 1: Graus de pertinência do vértice 5 a seus pais potenciais.

nas iterações seguintes. Esta estratégia não elimina a possibilidade de fusão entre F e este pai, ao passo que a primeira é muito mais restritiva. Propõe-se então a seguinte regra: *Cada vértice difuso estará parcialmente ligado a seus pais potenciais.* Isto significa que a região que ele representa fará parcialmente parte de cada um dos campos receptores vizinhos na base da pirâmide.

Desta maneira, o grau de pertinência $\mu_{P_i}(F)$ do vértice F a seu pai P_i pode ser dado por:

$$\mu_{P_i}(F) = \frac{s(F, P_i)}{\sum_j s(F, P_j)} \quad \forall i \quad (4)$$

(onde $\sum_j \mu_{P_j}(F) = 1$) e $s(F, P_i)$ é dado por

$$s(F, P_i) = 1 - \frac{|NC(F) - NC(P_i)|}{NC_{max} - NC_{min}} \quad (5)$$

ou alternativamente por:

$$s(F, P_i) = 1 - \frac{|NC(F) - NC(P_i)|}{\sum_j |NC(F) - NC(P_j)|} \quad (6)$$

As equações (5) e (6) sendo aplicadas ao grafo da figura 2(c), cujo vértice 5, de nível de cinza 105, tem 4 pais potenciais, fornecem os resultados da tabela 1.

Percebe-se que com a segunda proposição de calcular dos graus de pertinência, obtém-se resultados mais discriminativos. É através desses resultados que os atributos do vértice difuso serão tomados em consideração nos níveis superiores da pirâmide, até que ele tome sua decisão definitiva. A figura 3(b) ilustra as ligações clássicas ($P(2) = 3, P(4) = 1$ e $P(x) = x \quad \forall x$ sobrevivente) e difusas dadas por (6).

3.4 Algumas considerações

A introdução de arestas e vértices difusos na pirâmide de grafos, cria uma nova topologia nas relações entre os campos receptores na base da pirâmide. Em função disso, as regras para a obtenção dos sobreviventes e a colocação em dia dos grafos de adjacência e similaridade devem ser adaptadas. Em [3] é proposto um

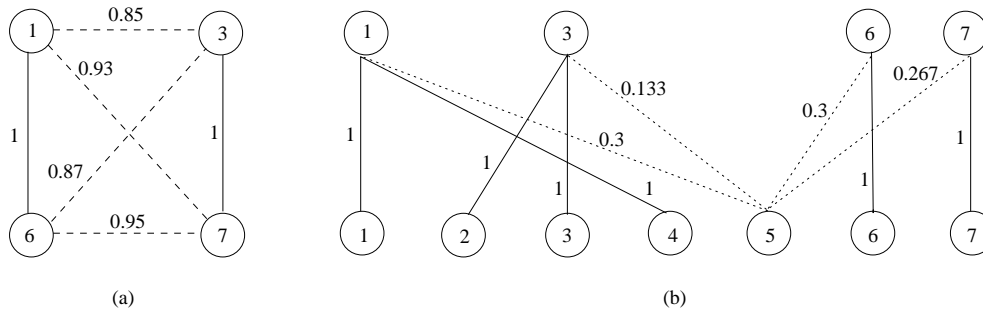


FIGURA 3: Grafos representando (a) adjacências e (b) ligações difusas.

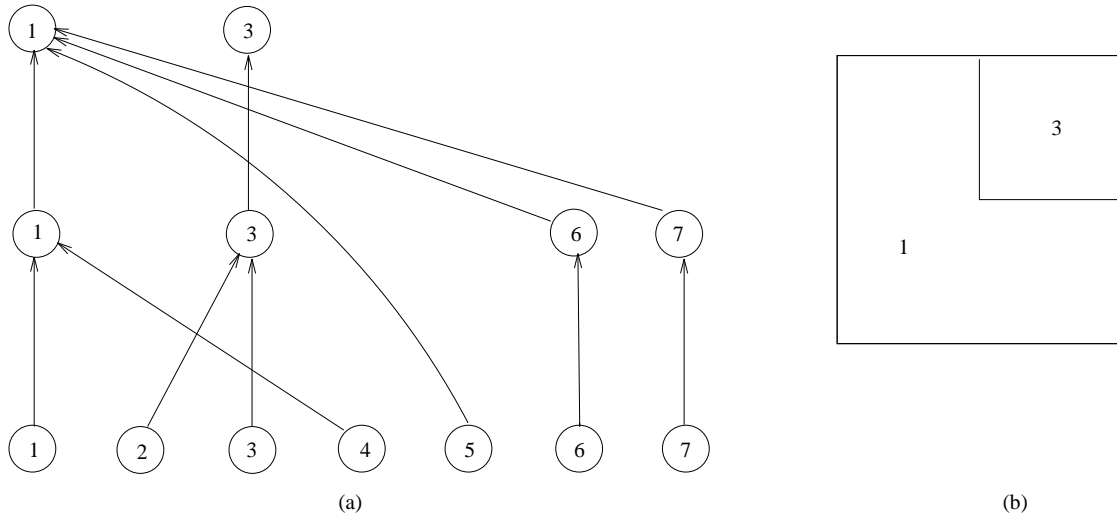


FIGURA 4: (a) Grafo representando as ligações da (b) configuração final.

modelo para resolver estes problemas. A figura 3(a) fornece uma idéia de como ficaria o grafo de adjacência, que contém também arestas difusas, as quais traduzem a incerteza de fusão dos vértices difusos.

Uma das possíveis configurações finais do exemplo adotado neste artigo é mostrada na figura 4(b), representada pela estrutura arborescente de 4(a). Note-se que o vértice difuso 5, quando escolhe seu pai definitivo, liga-se diretamente a ele, sem passar pelos níveis intermediários.

3.5 Resultados e comparações

Mostra-se nesta seção a aplicação das pirâmides irregulares clássica e difusa a dois tipos de imagem: a primeira (figura 5(a)), é uma imagem de ressonância magnética, mostrando um corte dorsal. É uma imagem ruidosa e os diferentes órgãos que a compõem são dificilmente segmentáveis. Seu histograma de níveis de cinza é dado em 5(b). A segunda imagem (figura 7(a)) foi obtida através da radiografia de redes vasculares. Uma das dificuldades na segmentação

desta imagem é a conservação da conectividade dos vasos, devido ao fato do contraste com o fundo variar ao longo de seus trajetos. O histograma de níveis de cinza é mostrado em 7(b).

A figura 5 (c a f) mostra dois resultados da aplicação da pirâmide irregular à imagem mostrada em (a). O primeiro desses resultados, acompanhado da sua carta de contornos (c e d), deixa transparecer a dificuldade do algoritmo para tratar informações ruidosas (obtm-se várias regiões de pequeno tamanho nas zonas de textura e o fundo da imagem não é homogêneo). Aumentando o limiar global de 10 a 12, passa-se de 153 a 71 regiões detectadas (e e f). Este resultado é melhor que o precedente, mas grandes zonas na figura, que poderiam ser representadas em uma só região, não o são, ao passo que o interior da figura começa a se confundir com o fundo da imagem.

Para a mesma imagem, os resultados fornecidos pela pirâmide difusa (figura 6), quando s_g vale 10 (a e b), 11 (f e i) e 12 (d e g, e e h) mostram uma superioridade no tratamento de zonas ruidosas, graças a adaptabilidade no tratamento de informações incer-

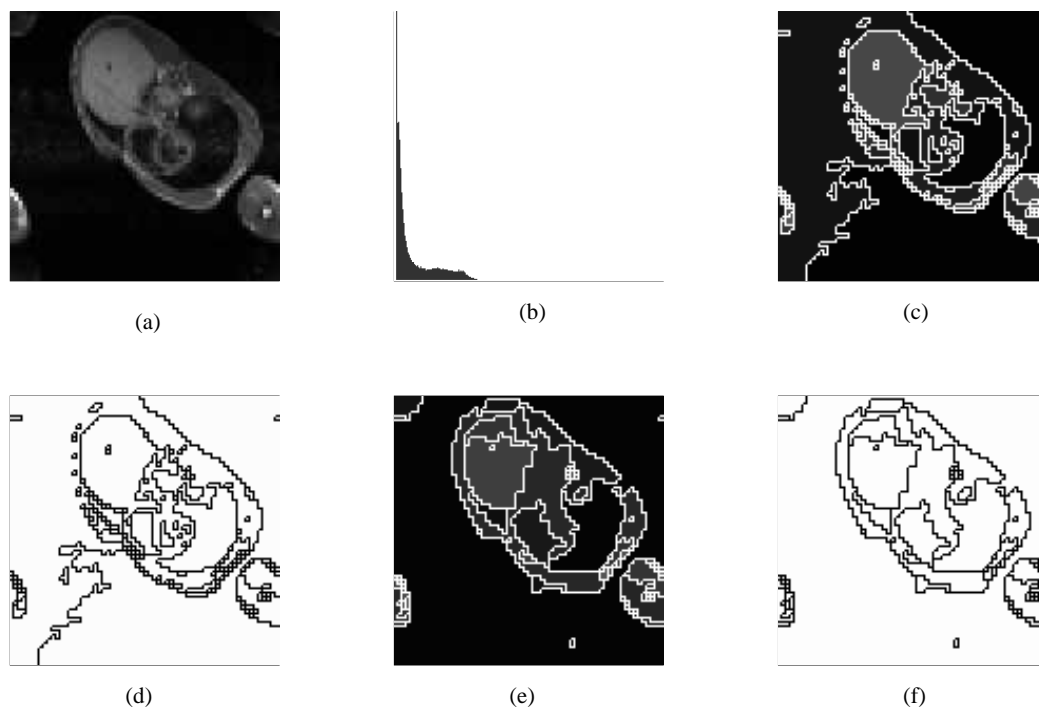


FIGURA 5: Pirâmide irregular clássica aplicada a uma imagem de ressonância magnética.

tas e imprecisas. Vê-se claramente a influência do ruído no oitavo nível da pirâmide (c), o qual é pouco a pouco eliminado, antes do ápice ser atingido no décimo sexto nível (b).

Dois resultados da aplicação da pirâmide irregular clássica à segunda imagem são mostrados na figura 7. Percebe-se em (c e d) a detecção de um fundo bem homogêneo, mas os elementos alongados foram particionados em vários pedaços. Buscando obter a conectividade desses vasos, baixou-se o valor de limiar global. O resultado (e e f) decepciona pois torna-se sobrecarregado de informações não significativas, além de não se obter uma distinção entre o fundo e os vasos sanguíneos, em algumas zonas da imagem.

Os resultados fornecidos pela pirâmide irregular difusa (figura 8) permitem uma melhor detecção das entidades alongadas na imagem. Percebe-se isso em (a e b), imagens nas quais os vasos menos contrastados com o fundo têm uma maior conectividade que no caso clássico. Diminuindo o limiar global, a segmentação perde em qualidade (c e d).

Em geral, para um mesmo valor de limiar global, a pirâmide difusa chega a seu ápice em um número menor de níveis que a clássica. Não existe melhoramento significativo nos resultados fornecidos pela pirâmide difusa, em relação à clássica, quando a imagem inicial é bem contrastada.

4 Conclusão

Apresentou-se neste trabalho uma estrutura piramidal irregular difusa visando prover uma segmentação de imagens mais adaptativa. O método se baseou no princípio de que um vértice não sobrevivente do grafo de adjacência pode se associar a mais de um pai ao mesmo tempo. Propôs-se uma forma de gerenciar as relações pai-filho nos diversos níveis da pirâmide.

Esta técnica apresenta resultados superiores àqueles fornecidos pelo algoritmo tradicional quando as imagens tratadas são ruidosas ou possuem entidades alongadas. Em imagens bem contrastadas não foi observada melhora significativa.

O novo método abre caminho ao estudo de operações do tipo regiões-bordas e à generalização para o caso 3D.

Referências

- [1] M. Bister, J. Cornelis, and A. Rosenfeld. A critical view of pyramid segmentation algorithms. *Pattern Recognition Letters*, 11(9):605–617, September 1990.
- [2] R. P. Blanford and S. L. Tanimoto. Bright-spot detection in pyramids. *Computer Vision, Graphics, and Image Processing*, 43:133–149, 1988.

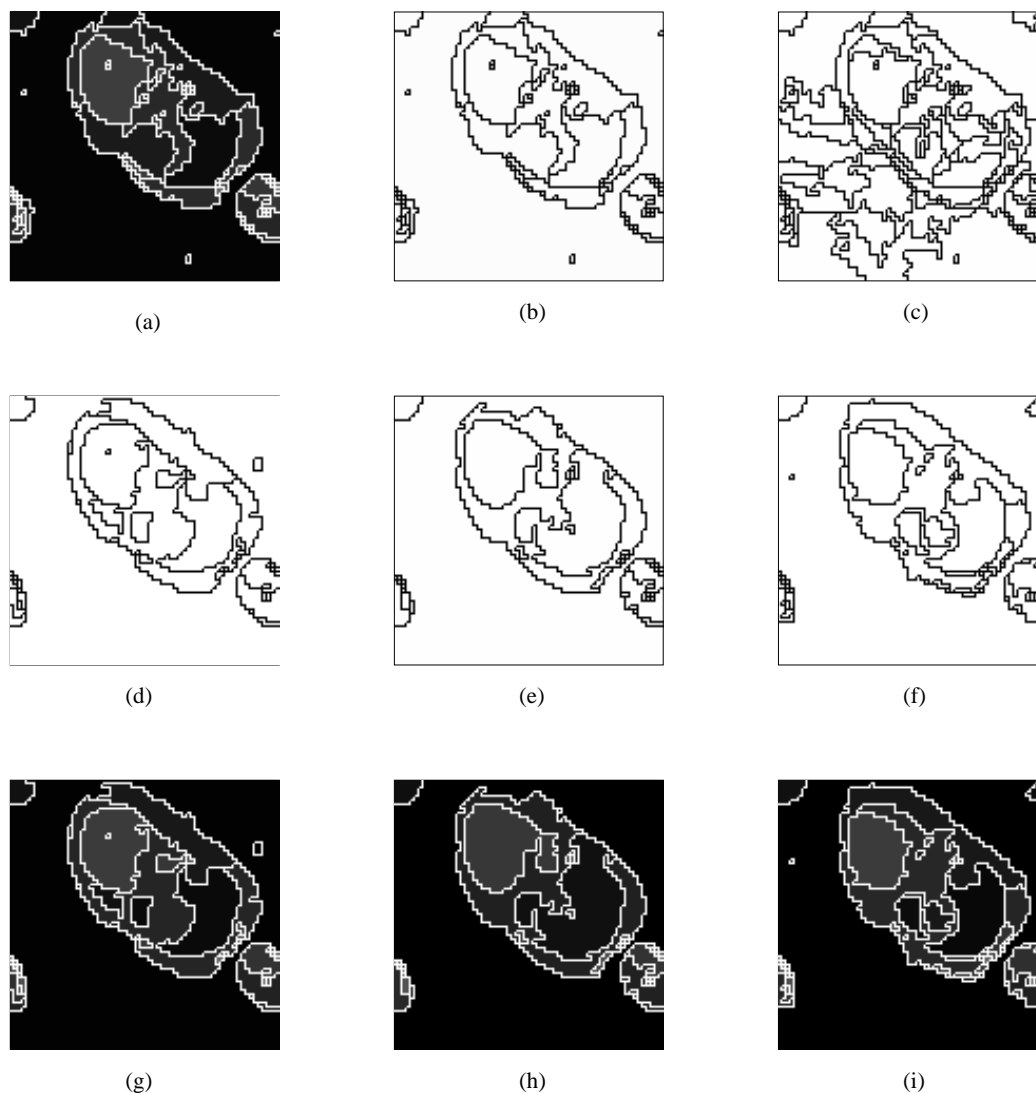


FIGURA 6: Pirâmide irregular difusa aplicada a uma imagem de ressonância magnética.

- [3] G. Braviano. *Logique floue en segmentation d'images : seuillage par entropie et structures pyramidales irrégulières*. PhD thesis, Université Joseph Fourier, Grenoble, 1995.
- [4] G. Braviano and A. Montanvert. Cooperation between fuzzy and nonfuzzy techniques for cytological image segmentation. In *8th SIBGRAPI*, pages 265–271, São Carlos, SP, 1995.
- [5] P.J. Burt. Fast filter transforms for image processing. *Computer Graphics and Image Processing*, 16:20–51, 1981.
- [6] J.M. Jolion. Analyse d'images : le modèle pyramidal. *Traitement du Signal*, 7(1):5–17, 1990.
- [7] S.W.C. Lam and H.H.S. Ip. Structural texture segmentation using irregular pyramid. *Pattern Recognition Letters*, 15(9):691–698, July 1994.
- [8] P. Meer. Stochastic image pyramids. *Computer Vision, Graphics, and Image Processing*, 45:269–294, 1989.
- [9] P. Meer, S. N. Jiang, S. Baugher, and A. Rosenfeld. Robustness of image pyramids under structural perturbations. *Computer Vision, Graphics, and Image Processing*, 44:307–311, 1988.
- [10] A. Montanvert, P. Meer, and A. Rosenfeld. Hierarchical image analysis using irregular tessellations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):307–316, April 1991.

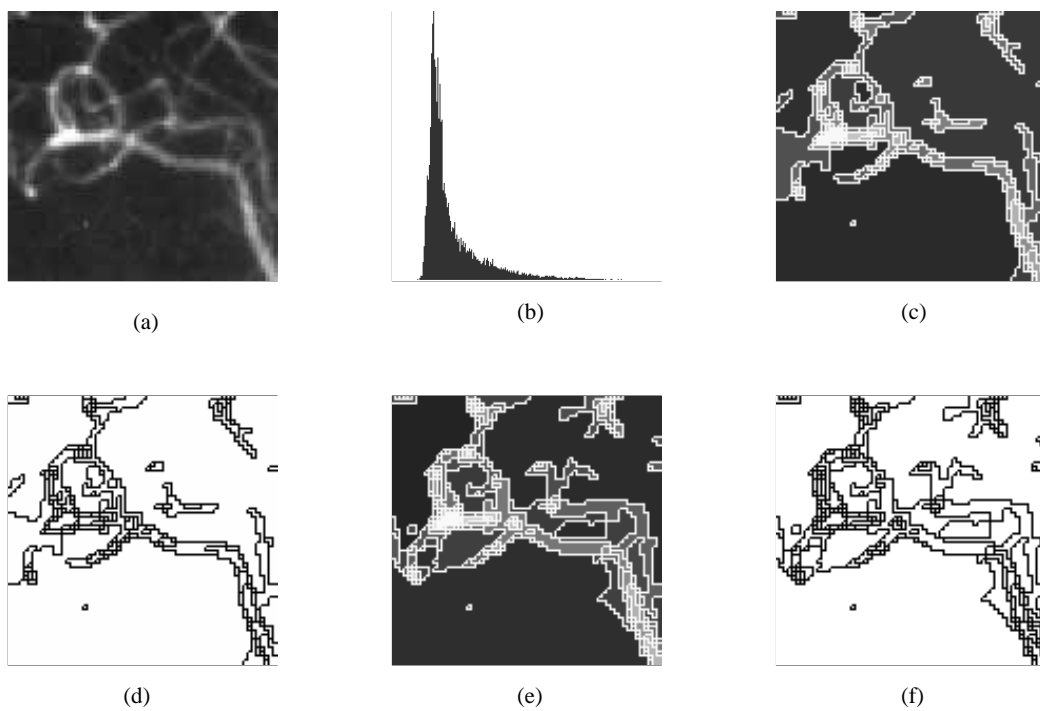


FIGURA 7: Pirâmide irregular clássica aplicada a uma imagem de redes vasculares.

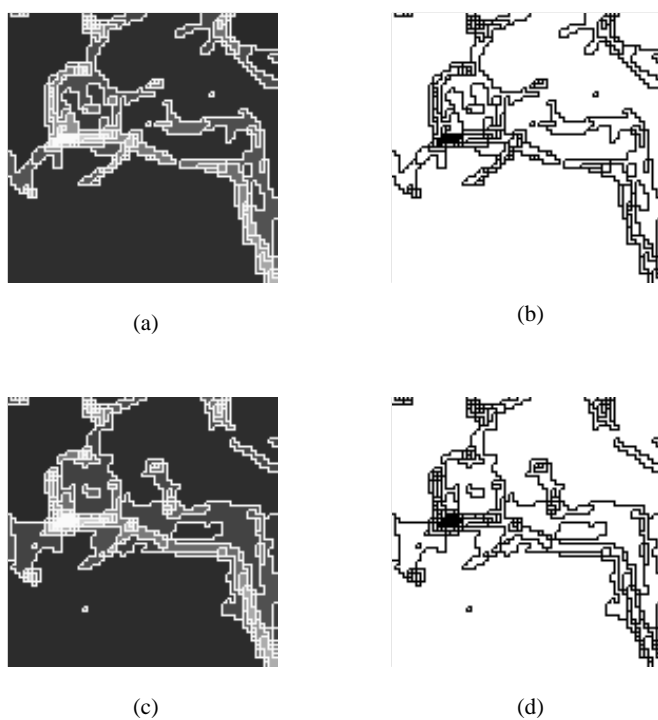


FIGURA 8: Pirâmide irregular difusa aplicada a uma imagem de redes vasculares.