

# Segmentação de Imagens Digitais: Uma Revisão

Marcus F. S. Saldanha<sup>1</sup>, Dr<sup>a</sup> Corina da Costa Freitas<sup>1</sup>

<sup>1</sup>Divisão de Processamento de Imagens – Instituto Nacional de Pesquisas Espaciais (INPE)

Av dos Astronautas, 1758 – 12.227-010 – São José dos Campos – SP – Brazil

marcus@dpi.inpe.br, corina@dpi.inpe.br

**Abstract.** *One of the main challenges in the processing of digital images is the rapid extraction of information from the images for analysis. A key issue in the extraction is to simplify the image, reducing the information to homogeneous regions, allowing a more appropriate representation. This task is the problem of segmentation. Due to the diversity of use and distinct types of data, there is no single solution for all applications of segmentation, which leads to various techniques and methods. Within this context, this study aims to present a brief review of theoretical foundations and the more usual approach adopted for the segmentation of images.*

**Resumo.** *Um dos grandes desafios na área de processamento de imagens digitais é a extração rápida de informações a partir das imagens para a realização de análise. Uma questão fundamental na extração é a simplificação da imagem, reduzindo as informações a regiões homogêneas, permitindo uma representação mais adequada. Essa tarefa consiste o problema da segmentação. Devido à diversidade de uso e tipos distintos de dados, não existe uma solução única para todas as aplicações da segmentação, o que dá origem a diversas técnicas e métodos. Dentro deste contexto, este trabalho tem por objetivo apresentar uma breve revisão dos fundamentos teóricos e das abordagens mais usuais adotadas para a segmentação de imagens.*

## 1. Introdução

Atualmente, diversos campos da atividade humana têm recorrido a imagens para armazenar, exibir e fornecer informações sobre o mundo que nos rodeia. O desafio que fica para os pesquisadores, engenheiros e empresários, é extrair rapidamente informação valiosa a partir de dados brutos de imagens. Este é o objetivo primário do processamento de imagens, ou seja, converter imagens em informação (RSI, 2005). De forma geral, funções de processamento que incluam reconhecimento e interpretação estão associadas com aplicações de análise de imagens nas quais o objetivo é a extração automática – ou mesmo parcialmente automática – de informação a partir de uma imagem (Gonzales e Woods, 1987).

Uma questão central na extração de informações a partir de uma imagem digital é a redução das informações dessa imagem em regiões mais ou menos homogêneas,

visando permitir uma análise mais apropriada da imagem. Isto consiste o problema da segmentação, e têm dado origem ao desenvolvimento de várias soluções.

Dentro deste enfoque, este trabalho tem como objetivos apresentar uma discussão a respeito de alguns fatores considerados relevantes para a conceituação e fundamentação da segmentação, bem como, realizar uma breve revisão dos algoritmos mais utilizados na segmentação de imagens digitais.

## **2. Segmentação de Imagens: Conceito e Considerações**

A obtenção de melhores resultados proveniente da análise de informações extraídas de uma imagem esta intimamente ligada à qualidade do tratamento feito sobre a fonte dos dados. Segundo Gonzalez e Woods (1987), as técnicas de extração de informação a partir de imagens pertencem à área do processamento referida como análise de imagens, que para ser executada tem como passo inicial a realização de uma segmentação.

A definição para a segmentação de imagens esta diretamente relacionada à área na qual será aplicada. Dentro da área de visão computacional, a segmentação refere-se ao processo de decomposição de uma imagem digital em vários segmentos (regiões) que a formam (Jain, 1989). Já para a área de processamento digital de imagens de sensoriamento remoto a segmentação de imagem é a parte da análise de imagem que trata da definição de objetos geográficos ou regiões em uma imagem (Moik, 1980).

Os objetos que compõem uma imagem possuem duas características básicas: (1) Eles exibem alguma uniformidade interna em relação a uma propriedade da imagem, e (2) eles contrastam em relação a sua vizinhança (Moik, 1980). Essas características entretanto sofrem influência do ruído e por isso não podem ser deterministicamente modeladas.

A segmentação de imagens tem basicamente dois objetivos: (1º) decompor a imagem em partes para posterior análise; (2º) realizar uma mudança de representação (Shapiro e Stockman, 2001). O resultado final esperado de um processo de segmentação é um número finito de regiões homogêneas, que individualizam as diferentes regiões contidas numa imagem. Sendo assim, a imagem seria formada por um conjunto de regiões, conectadas pelas suas bordas e que não se sobrepõe, de tal forma que cada pixel da imagem pertence unicamente a uma determinada região (Acharya e Ray, 2005).

O conceito de segmentação da forma com que é apresentado idealiza o mundo real de forma conveniente para diversas aplicações. Entretanto, cabe lembrar que é uma invenção da mente humana e que embora seja uma boa abordagem para materializar soluções para trabalhos de análise de imagens, pode em algumas situações levar a resultados insatisfatórios (Davis, 2004).

Pode-se dizer que a elaboração de algoritmos visando à automatização de processos de segmentação é uma das tarefas mais difíceis dentro da área de processamento de imagens (Gonzalez e Woods, 1987). Pois, a segmentação de imagem é basicamente um problema de percepção psicofísica, e, portanto, não susceptível de uma solução puramente analítica (Nikhil e Sankar, 1993). Dessa forma, para obter um resultado satisfatório, qualquer algoritmo de segmentação, não importando modelo matemático adotado, deve ser complementado por heurísticas que envolvam tanto informação

semântica como conhecimento a priori sobre as imagens que estão sendo processadas (Yong et al, 2004).

O desenvolvimento dos algoritmos leva em consideração duas características básicas dos pixels que envolvem dois conceitos: descontinuidade e similaridade. Por meio do uso desses conceitos, conforme estruturado por Gonzalez e Woods (1987), a segmentação é feita usando-se detecção de bordas e detecção de similaridades.

Por sua vez, os métodos mais comuns empregados na segmentação podem ser divididos em três grupos principais, de acordo com a forma como vêm a imagem (Mezaris et al, 2003): Métodos de segmentação no domínio espacial; Métodos de segmentação no espaço de medidas; e, Métodos de segmentação no domínio da frequência.

Cabe lembrar, que a qualidade de segmentação de uma imagem influencia diretamente no resultado da fase de análise e, portanto a escolha de uma técnica de segmentação adequada implica no conhecimento dos detalhes das técnicas existentes.

### **3. Métodos Usados na Segmentação de Imagens**

#### **3.1. Detecção de Descontinuidades**

As descontinuidades encontradas em uma imagem podem ser pontuais, linhas ou os limites (bordas) de um objeto. Essas feições, sobressaem numa imagem, seja por possuir tons de cinza distintos a região na quais estão inseridas (caso de pontos e linhas) ou por assinalarem mudanças bruscas de tons de cinza entre regiões (caso de bordas e linhas). Os algoritmos desenvolvidos para detectar essas descontinuidades usualmente usam a convolução, implicando no uso de máscaras.

Os métodos de detecção de descontinuidades, mais particularmente os de detecção de linhas e de bordas, apresentam geralmente como resultados falhas de detecção. Portanto, esses métodos devem ser seguidos de processamentos visando sanar essas falhas. As técnicas de processamento que realizam esse tipo tratamento não serão abordadas aqui, maiores detalhes a respeito ver Gonzalez e Woods (1987) e Hough (1962).

##### ***Detecção de pontos e linhas***

A idéia que esta por trás dos algoritmos para a detecção de pontos e linhas é que essas feições podem possuir tom de cinza muito diferente quando comparada aos seus vizinhos. Sendo assim, a detecção pode ser realizada por meio da aplicação uma máscara que forneça uma resposta  $R$  e um adequado limiar não negativo  $T$ . Ao se aplicar a máscara sobre a imagem, os pontos que retornarem valores acima do limiar estabelecido são detectados, ou seja,

$$|R| > T . \quad (3.4)$$

##### ***Detecção de bordas***

A detecção de bordas é de longe, a abordagem mais comum para detecção de descontinuidades (Gonzales e Woods, 1987). Em uma imagem a borda é definida como sendo o limite entre duas regiões com diferentes propriedades. Devido ao processo de amostragem em uma imagem digital, a borda aparece com uma pequena suavização em torno ao verdadeiro limite das regiões.

Essa transição pode ser representada por uma função  $f(x)$ . Dessa forma, a detecção de bordas pode ser realizada por meio da aplicação de operadores de diferenciação (Figura 01). A primeira derivada da função  $f(x)$ , apresenta um ponto de mínimo no ponto correspondente à borda. A magnitude do valor dessa derivada, pode indicar a intensidade da borda. A segunda derivada da função  $f(x)$ , por sua vez, apresenta um cruzamento por zero no ponto correspondente a borda. O sinal dessa derivada pode ser usado para determinar se um pixel está localizado na parte clara ou escura da borda.

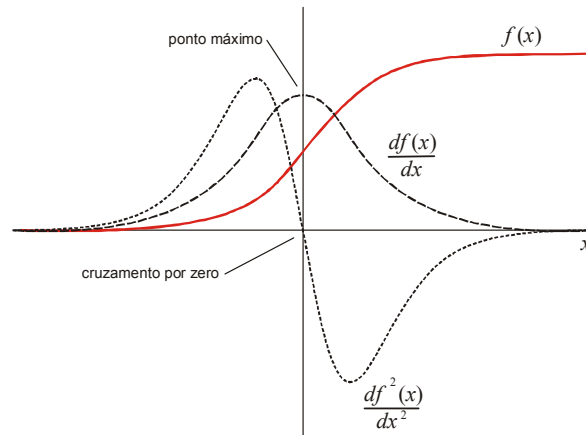


Figura 3.5 - Operadores de diferenciação aplicados sobre uma função que modela a mudança brusca de níveis de cinza em um perfil (espaço unidimensional) de uma imagem.

Para o caso de uma imagem, espaço bidimensional, a função assume a forma  $f(x,y)$ , sendo as duas direções representadas pelas derivadas parciais em função de  $x$  (direção vertical) e de  $y$  (direção horizontal). Nessa situação utilizam-se operadores diferenciais de gradiente, que correspondem à primeira derivada, e o laplaciano, correspondente a segunda derivada.

### 3.2. Detecção de Similaridades

A detecção de similaridade tem como fundamento a observação do interior dos objetos e não as fronteiras que os delimitam. Para tanto, parte da idealização que os pixels que compõem um objeto têm propriedades similares enquanto que pixels de objetos distintos têm propriedades distintas.

A formulação básica adotada para este tipo de abordagem é dada por Fu e Mui (1981). Segundo o autor, se considerarmos  $R$  como sendo uma imagem, a segmentação é a decomposição de  $R$  em  $n$  regiões  $R_1, R_2, \dots, R_n$  de tal forma que:

- a)  $\bigcup_{i=1}^n R_i = R$
- b)  $R_i$  é conectada,  $i = 1, 2, \dots, n$
- c)  $R_i \cap R_j = \emptyset, \forall i \neq j$
- d)  $Pu(R_i) = verdadeiro \forall i$
- e)  $Pu(R_i \cap R_j) = falso \forall i \neq j$

Pode existir um número de possíveis partições, mas a seleção de um conjunto adequado de regiões depende da escolha da propriedade  $P_u$  associada à região, ou seja, do predicado de uniformidade dos pixels da região (Pavlidis, 1977).

### ***Crescimento de regiões***

O princípio do funcionamento da técnica de crescimento por regiões é agrupar pixels ou sub-regiões em regiões maiores (Gonzales e Woods, 1987). Seu início se dá com a adoção de um conjunto de pixels “sementes”, a partir das quais é realizado o crescimento das regiões por meio da inclusão, de pixels vizinhos que tenham atributos similares, tais como, intensidade, textura, cor, etc. O crescimento ocorre iterativamente até que cada pixel seja processado e sejam formadas diferentes regiões cujas fronteiras possam ser definidas por polígonos fechados e os pixels internos a essas compartilhem de certa similaridade.

Por trás de seu funcionamento simples o crescimento de regiões esconde algumas dificuldades relacionadas à sua fundamentação. Uma dessas, diz respeito justamente ao início do processo, que é a escolha de sementes. Outra dificuldade está ligada à natureza da aplicação e ao tipo de imagem disponível, e diz respeito à definição do critério de similaridade a ser usado. Um outro problema que apresenta o crescimento de regiões é o critério de parada.

### ***Pirâmides***

A idéia que esta por trás da abordagem piramidal é dividir a imagem em elementos de área arbitrários e disjuntos para então, a partir desses, realizar a junção ou divisão das regiões. A abordagem tradicional para essa técnica é subdividir a imagem em quadrantes usando convenientemente a estrutura de dados denominada *quadtree* (abreviatura da expressão em inglês *quadratic picture tree*). Nessa estrutura a imagem é representada por um nodo pai (nível  $n=0$ ), enquanto que os quatro quadrantes são representados por quatro nodos filho (nível  $n=1$ ), e assim por diante em uma ordem pré-determinada.

### ***Detecção de Bacias***

A detecção de bacias pode ser visto como uma combinação entre detecção de bordas e junção de regiões (Erthal, 2008). Essa técnica baseia-se no princípio de “inundação de relevos topográficos” (Russ, 1998). Essa abordagem vem do princípio de que uma imagem em níveis de cinza pode ser vista como um relevo topográfico, formado por vales, que representam as regiões mais escuras da imagem, e por montanhas, que representam as porções mais claras (Roerdink e Meijster, 2000). É possível visualizar a “inundação” de duas maneiras distintas: a água vinda de cima, como se fosse chuva; ou vinda de baixo, como se o relevo estivesse perfurado nos pontos de altitude mínima e fosse imerso em um lago (Roerdink e Meijster, 2000, e Wangenheim, 2009).

Conforme as bacias vão sendo inundadas, águas provenientes de diferentes bacias se denominadas *watersheds*. Quando a água atinge o maior pico no relevo, o processo termina (Meyer, 1994). O resultado da inundação é a separação das regiões que compõe a imagem delimitada pelas linhas divisórias das bacias. A Figura 3.9 ilustra o nível de inundação partindo das regiões mais baixas para as mais altas até o limite considerado satisfatório para a distinção das bacias.

### 3.3. Segmentação no Espaço de Atributos

#### *Limiarização*

A limiarização é a forma mais simples de segmentar imagens (Erthal, 2008) e uma das abordagens mais importantes para a segmentação (Gonzales e Woods, 1987). A idéia que está por trás dessas técnicas é a que um objeto pode ser entendido como uma região formada por pixels contíguos que tenham em comum uma faixa de intensidades. Dessa forma, a limiarização usa a intensidade dos pixels para distinguir as regiões.

O processo todo se baseia na análise do histograma de frequências da imagem, onde as faixas de intensidade correspondem às bandas do histograma. Um caso simples de ser entendido é o de um histograma bimodal. Nessa situação o histograma apresenta dois picos separados por um vale entre eles, sendo assim, esses picos podem representar duas regiões distintas, como por exemplo, um fundo e um objeto. Seguindo o mesmo raciocínio, se o histograma de uma imagem for tri-modal poderíamos então obter a representação de três regiões distintas, como dois objetos e um fundo.

#### *Agrupamento (Clustering)*

Uma das idéias principais das técnicas de segmentação é encontrar e individualizar as regiões na imagem que tenham feições similares. Intuitivamente, padrões pertencentes a um dado grupo devem ser mais similares entre si do que em relação a padrões pertencentes a outros grupos (Moscato e Zuben, 2002). Técnicas de agrupamento (do termo original em inglês *clustering*), encontradas na literatura de reconhecimento de padrões possuem objetivos similares e, sendo assim, segundo Jain (1989), podem ser aplicadas para a segmentação de imagens.

Neste tipo de segmentação um conjunto de dados é substituído por grupos ou *clusters*. A definição de clusters, segundo Everitt et al. (2001), parte da idéia que valores de atributos ( $\mathbf{x}$ ) são vistos como pontos no espaço  $l$ -dim, sendo assim, “*clusters* são regiões contínuas deste espaço e que contém uma alta densidade de pontos, separadas de outras regiões densas, por regiões com baixa densidade de pontos”.

A partir desse ponto duas abordagens para fundamentação dos algoritmos podem ser encontradas. A primeira leva em consideração o conceito de agrupamento rígido (dos termos em inglês, *hard*, *crisp*) no qual cada ponto pertence a um único grupo. A segunda considera o conceito de agrupamento nebuloso (do termo em inglês *fuzzy*) onde cada ponto pertence a todos os grupos até certo grau.

A abordagem mais adequada para fins de aplicação na segmentação de imagens entretanto, é a primeira, pois não fere os princípios de segmentação por regiões estipulados por Fu e Mui (1981). A seguir é abordado o algoritmo de agrupamento mais conhecido e comumente usado para a segmentação de imagens, o  $k$ -médias.

#### **- Algoritmo $k$ -médias**

O  $k$ -médias é um algoritmo não supervisionado usado em reconhecimento de padrões e segmentação de imagens, que gera clusters de  $n$  objetos baseados em atributos a partir de  $k$  partições, sendo  $k < n$ . O algoritmo presume que os atributos dos objetos formam um espaço vetorial, e tem por fundamento principal minimizar a variância intra-agrupamento total, ou seja, a função dada por

$$V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2, \quad (3.20)$$

onde existem  $k$  clusters  $S_i$ ,  $i = 1, 2, \dots, k$ , e  $\mu_i$  é o centróide ou ponto médio de todos os pontos  $x_j \in S_i$ .

O algoritmo inicia pela partição dos pontos de entrada em  $k$  conjuntos iniciais, de forma randômica ou usando alguma heurística. Calcula então o ponto médio, ou centróide, de cada agrupamento. Constrói uma nova partição pela associação de cada ponto ao centróide mais próximo. Os centróides são re-calculados para os novos agrupamentos, e o algoritmo é repetido pela aplicação alternada desses dois passos até convergência, que é obtida quando os pontos não mudarem mais de agrupamentos ou alternativamente se os centróides não mudarem de posição. Também pode ser adotado como critério de parada para o algoritmo um número máximo de iterações. A Figura 3.14 ilustra o funcionamento do  $k$ -médias mostrando o deslocamento dos centróides de acordo com as iterações.

Para fins de classificação o valor de  $k$  representa o número de classes definidas *a priori* pelo usuário. A principal desvantagem do  $k$ -médias é que em termos de desempenho o algoritmo não garante o resultado global ótimo. Pois, a qualidade da solução final depende muito do dos conjuntos iniciais de clusters, podendo na prática, vir a se afastar muito do ótimo global.

Outro inconveniente do algoritmo é que como o número de agrupamentos é um parâmetro de entrada uma escolha inapropriada de  $k$  pode retornar em resultados pobres. O algoritmo também assume que a variância é uma medida adequada do espalhamento do cluster o que para algumas situações pode não refletir a verdade.

#### 4. Conclusões

A segmentação é de fundamental importância para a fase de análise, pois por meio dela é possível se extrair informações valiosas a partir das imagens digitais. Conforme foi citado no início deste trabalho e como pode ser notado durante o seu desenvolvimento, a escolha de uma técnica de segmentação adequada está intimamente ligada a natureza dos dados e ao tipo de aplicação. Sendo assim, o conhecimento dos detalhes relacionados a cada técnica de segmentação é de extrema importância para uma boa utilização dessas em aplicações humanas.

Cabe lembrar, entretanto que as técnicas vistas aqui e a maneira com foram apresentadas devem servir somente de referência para estudos iniciais a respeito do tema. Sendo necessário, para um maior aprofundamento, consultas as fontes bibliográficas usadas para confeccionar este trabalho, bem como, consultas a periódicos especializados na área de processamento digital de imagens, visão computacional e reconhecimento de padrões.

#### Referências Bibliográficas

Gonzales, R. C.; Woods, R. E. Digital image processing. New York: Addison-Wesley Publishing Company, Inc, 1987.

Jahne, B. Digital image processing. 5. ed. Berlin: Springer-Verlag, 598 p. p. 2002.

- Jain, A. K. Fundamentals of digital image processing. New Jersey: Prentice-Hall, Inc. 1989.
- Moik, J. G. Digital processing of remotely sensed images. Washington DC: Government Printing Office. 1980.
- Davis, E. R. Machine vision. 3. ed. Royal Holloway, University of London. Morgan Kaufmann Publishers. 2004.
- Shapiro, L. G.; Stockman, G. C. Computer vision. New Jersey: Prentice-Hall, 2001.
- Acharya, T.; Ray, A. K. Image processing, principles and applications. New Jersey: John Wiley and Sons, Inc. 2005.
- Fu, K. S.; Mui, J. K. A survey on image segmentation, Pattern Recognition 13. 1981.
- Pavlidis, T. Structural Patter Recognition. New York: Springer Verlag. 1977.
- Nikhil, R. P.; Sankar, K. P. A review on image segmentation techniques. Pattern Recognition, 26 (9). 1993.
- Yong, X.; Feng, D.; Rongchun, Z.; Optimal selection of image segmentation algorithms based on performance prediction. Sydney: Australian Computer Society, Inc. 2004.
- Mezaris, V.; Kompatsiaris, I.; Strintzis, M. G. Still image objective segmentation evaluating using ground truth. 2003.
- Wangenheim, A.; Mantelli, S. Técnicas de segmentação simples de imagens. Seminário Introdução à Visão Computacional. Florianópolis: Universidade Federal de Santa Catarina. Obtido a partir de <http://www.inf.ufsc.br/~visao/threshold.html>, consultado em 02 de abril de 2009.
- RSI. Image Processing in IDL. Ed. 2005
- Erthal, G. J. Notas de aula da disciplina reconhecimento de padrões. INPE, 2008.
- Silva, F. J. V.; Alves, C. H. F. Aplicação de Técnicas de Processamento de Imagens Digitais em Imagens Geradas dor Ultra-Som. 8º Encontro Regional de Matemática Aplicada e Computacional. 2008.
- Canny, J. A Computational Approach to Edge Detection. IEEE Transaction on Pattern Analysis and machine Intelligence, 8(6). 1986.
- Facon, J. Processamento e Análise de Imagens. Pontificia Universidade Católica do Paraná, Paraná. 2001.
- Marr, D.; Hildreth. Theory of Edge Detection. Proc. R. Soc. Lond. Vol. B207. 1980.
- Hough, P. V. C.. Methodos and Means for Recognizing Complex Patterns. U.S. Patent. 1962.
- Benz 2004 Ecognition IJPRS
- [22] <http://www.inf.unisinos.br/~ari/estrut/quad/Quadtree.htm>
- Horowitz, S. L.; Pavlidis, T. Picture Segmentation by a Tree Traversal Algorithm. Journal of the ACM, Volume23, 368-388. 1976.
- Russ, J. C.. The Image Processing Handbook. CRC Press, Boca Raton. 1998.



- Roerdink, J. B. and Meijster, A.. The watershed transform: Definitions, algorithms and parallelizations strategies. *Fundamenta Informaticae*, 41:187–228. 2000.
- Wangenheim, A.. Técnicas de segmentação de imagens por crescimento de regiões (region-growing image segmentation). Seminário Introdução à Visão Computacional. Florianópolis: Universidade Federal de Santa Catarina. Obtido a partir de <http://www.inf.ufsc.br/~visao/threshold.html>, consultado em 02 de abril de 2009.
- Meyer, F. Topographic distance and watershed lines. *Signal Processing*, 38:113–125. 1994.
- Zuben, F. J. V.; Moscato, P. A. Tópicos em Engenharia de Computação V. DCA/FEEC/Unicamp. Obtido a partir de [ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia368\\_02/](ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia368_02/), consultado em 04 de maio de 2009.
- Teodoridis, S.; Koutroumbas, K. *Pattern Recognition. Second Edition.* Academic Press. Elsevier. 2003.
- Marques de Sá, J. P. *Pattern Recognition: Concepts, Methods and Applications.* Springer-Verlag London Limited. 2001.
- Wu X.; Kumar, V.; Quinlan, J. R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G. J.; Ng, A.; Liu, B.; Yu, P. S.; Zhou, Z.-H.; Steinbach, M.; Hand, D. J.; Steinberg, D. *Top 10 algorithms in data mining.* Springer-Verlag London Limited. 2007.