

Introdução ao Processamento de Imagem Digital (MO443/MC920)

Prof. Alexandre Xavier Falcão

Segundo semestre de 2003

1 Introdução à segmentação de imagens

Segmentar uma imagem consiste em particioná-la em regiões de pixels relevantes para uma dada aplicação (i.e. objetos e fundo). A segmentação é uma das principais etapas na maioria das aplicações e representa um dos maiores desafios em processamento de imagens. A principal razão desta dificuldade está na falta de informação sobre os objetos nas imagens.

A segmentação consiste de duas tarefas básicas: identificação e delineamento. A identificação indica a localização aproximada do objeto e o delineamento extrai sua extensão na imagem. Seres humanos executam a primeira tarefa com relativa facilidade, mas o computador é capaz de executar a segunda com muito mais precisão do que os seres humanos. A dificuldade da máquina na identificação de objetos se deve a falta de uma descrição global dos objetos na forma de um modelo matemático, pois as decisões automáticas usam normalmente propriedades locais extraídas em torno dos pixels. Por exemplo, sabendo que o objeto de interesse tem a forma triangular, como incorporar esta informação no modelo de segmentação para que o algoritmo procure por objetos triangulares?

Métodos de segmentação podem ser classificados pelo tipo de representação que extraem, como baseados em borda ou em região. Métodos baseados em borda procuram extrair contornos fechados, onde um contorno fechado é um caminho de pixels 4- ou 8-adjacentes que separa o interior do exterior do objeto. Isto é, todo caminho 4-conexo vindo de dentro para fora, ou vice-versa, deve cruzar a borda em um pixel. Se o caminho for 8-conexo, então deve cruzar a borda em pelo menos um vértice de pixel. Algumas abordagens também definem orientação para o contorno. Isto é, se caminharmos ao longo da borda teremos sempre um dos lados (direito ou esquerdo) no interior e o outro no exterior do objeto. Métodos baseados em região extraem o conjunto de pixels que representa o interior do objeto, incluindo os pixels de fronteira. Algumas abordagens classificadas como híbridas utilizam estratégias baseadas em bordas e em região simultaneamente, independente da representação final.

A segmentação de uma imagem $\hat{I} = (D_I, I)$ baseada em região pode ser vista como um mapeamento que associa para todo pixel $p \in D_I$ um inteiro $L(p)$, denominado rótulo, cujo valor é diferente para cada objeto (incluindo o fundo). Neste caso, a segmentação é dita *hard* porque cada pixel p só pertence a um único objeto. Algumas abordagens estendem este

conceito para *fuzzy*, onde cada pixel p pertence a todos os objetos com diferentes graus de pertinência. Por exemplo, calcula-se um grau de pertinência $0 \leq M(p, o) \leq 1$ do pixel p com o objeto o , tal que $\sum_{o'=1}^k M(p, o') = 1$ onde k é o número de objetos. Observe que também podemos entender uma borda de objeto como um conjunto de pixels em torno de uma fronteira e aplicarmos a abordagem *fuzzy*.

Métodos de segmentação também podem ser classificados como interativos ou automáticos. Métodos automáticos evitam a intervenção do usuário, mas nem sempre garantem o resultado desejado. Métodos interativos variam de técnicas manuais, onde o usuário pinta as regiões ou delinea as bordas dos objetos, a métodos que procuram minimizar o envolvimento e o tempo total do usuário na segmentação. Em abordagens interativas, uma sugestão é atribuir ao usuário a tarefa de identificação e ao computador a tarefa de delineamento.

1.1 Técnicas baseadas em região

A forma mais simples de particionar uma imagem é dividi-la em duas regiões, objeto e fundo, aplicando uma classificação pixel a pixel. Esta classificação pode ser baseada em uma única característica (e.g. o brilho) ou em um conjunto de características (e.g. brilho, gradiente, matiz) dos pixels. A classificação também se aplica a múltiplas regiões (objetos) que formam aglomerados (*clusters*) no espaço de características. Exemplos dessas técnicas são limiarização (*thresholding*), classificação estatística, redes neurais, e *clustering*. Entre essas, limiarização e *clustering* são as mais usadas para classificação de pixels.

A limiarização estabelece, por exemplo, um intervalo disjunto de brilho para cada região, e classifica o pixel cujo brilho está em um dado intervalo como pertencente à região correspondente. Sua extensão para múltiplas características, onde os intervalos se transformam em hipercubos, recebe o nome de método do paralelepípedo. Técnicas de *clustering* podem ser divididas entre particionais (e.g. algoritmo *k-means*) e hierárquicas (e.g. algoritmo *single-linkage*), sendo as hierárquicas divididas entre aglomerativas e divisivas. Técnicas particionais dividem o espaço de características em um número k de regiões satisfazendo um dado critério de dissimilaridade entre elas, enquanto os métodos hierárquicos tratam simultaneamente todos as possíveis partições. Métodos aglomerativos partem da partição onde cada pixel forma uma região para o caso onde todos os pixels formam a mesma região, e os métodos divisivos fazem o caminho inverso.

Um aspecto importante na segmentação de imagens é a conectividade entre pixels que pertencem a um mesmo componente. As técnicas mencionadas acima não exploram a conectividade, exceto alguns algoritmos de *clustering* que podem ser aplicados na imagem usando a relação de adjacência em vez de serem aplicados no espaço de características. A conectividade é explorada em técnicas de crescimento de regiões e técnicas que usam divisão e conquista de regiões conexas. O crescimento de regiões usa um conjunto de pixels sementes e um critério de parada. Em algumas abordagens, as regiões crescem a partir de sementes marcadas em um único objeto até atingirem o critério de parada (e.g. páre de crescer a região R se $I(q) < T_1$ ou $I(q) > T_2$, para $T_1 < T_2$, $q \in A(p)$, e $p \in R$), idealmente na fronteira do objeto com o fundo. Outras abordagens usam sementes em vários objetos (incluindo o fundo) e o critério de parada passa a ser definido pelo choque entre as regiões. Técnicas de divisão e conquista iniciam com a imagem

representando uma única região R com um predicado $P(R)$ (e.g. 80% dos pixels possuem o mesmo brilho) associado, e aplicam divisões e uniões sucessivas. Uma divisão de R em regiões menores é aplicada sempre que $P(R)$ for falso. Uma união de duas regiões vizinhas é aplicada sempre que o predicado da região resultante for verdadeiro. O processo simplifica a imagem à medida que obtém regiões com predicado verdadeiro e pára quando nenhuma divisão e união forem mais possíveis. Ao final aplica-se uma limiarização para completar a segmentação.

1.2 Técnicas baseadas em borda

A abordagem mais simples é por classificação de pixels. Aplica-se um filtro de suavização, seguido de um realce de bordas (e.g. magnitude de gradiente, Laplaciano) e depois uma classificação binária (borda/fundo) para decidir quais pixels pertencem à borda de um objeto. Esta abordagem deixa “buracos” na borda. Critérios locais e globais são aplicados para unir pixels que pertencem a uma mesma borda (técnicas de *edge linking*). Critérios locais buscam por segmentos próximos a cada segmento que possam pertencer a mesma borda (e.g. limiarização por histeresis) e os critérios globais assumem que segmentos de uma mesma borda satisfazem uma dada equação (e.g. Transformada de Hough).

Uma outra forma de abordar o problema é evitar a binarização da imagem de bordas realçadas, transformar a imagem em um grafo, e aplicar um algoritmo de busca por caminhos ótimos no grafo, onde cada caminho é um segmento de borda. Técnicas baseadas em busca heurística usando o algoritmo A^* e programação dinâmica são as mais populares. Essas técnicas costumavam impor restrições topológicas e geométricas para a borda e não consideravam todos os arcos de modo que nem sempre garantiam uma solução. Como veremos mais adiante, a IFT generaliza essas técnicas eliminando esses problemas.

Um aspecto negativo nas abordagens acima é a falta de informação global sobre o objeto no modelo de segmentação. Métodos baseados em contornos deformáveis procuram resolver o problema no framework de Equações Parciais e Diferenciais (PDE). A idéia é partir de um contorno inicial que deforma-se para minimizar um funcional de energia, o qual deve ser mínimo quando o contorno adere à borda do objeto. Na maioria das técnicas, porém, a informação relevante para extrair o objeto não é incorporada no funcional de energia e o método falha na segmentação. Por exemplo, o funcional assume que o contorno é suave quando existem pontos de alta curvatura. Como resultado o contorno não adere às indentações e protusões da borda.

1.3 Técnicas Híbridas

Existem várias formas de combinar técnicas baseadas em região com técnicas baseadas em borda. Uma idéia interessante é modelar a fronteira das regiões em abordagens de crescimento de regiões como um contorno deformável, fazendo com que o funcional de energia influencie no crescimento das regiões. O crescimento de regiões também poderia ser aplicado para inicializar o contorno deformável próximo à borda desejada, ou podemos ainda aplicar dois contornos deformáveis de dentro para fora e de fora para dentro do objeto como sementes iniciais para o crescimento de regiões.