

Extração de Características e Segmentação

A área de reconhecimento em imagens não se restringe aos objetos. Existe um campo de aplicações muito mais amplo, que envolve a busca de qualquer tipo de **padrões** detectáveis. Padrões em reconhecimento, por imagens digitais ou não, é um termo genérico que se refere a qualquer elemento que possa ser definido **quantitativamente** mesmo que sujeito a variações (algo que possa ser mais ou menos padronizado).

O termo padrão pode caracterizar praticamente qualquer coisa. Podem se relacionar à detecção de tumores em imagens médicas. Neste caso, os tumores são regiões **diferenciadas** da imagem, que devem ser reconhecidas de acordo com características extraídas que estejam fora do padrão normal. O reconhecimento de padrões é muito aplicado no sensoriamento remoto em identificação de regiões em imagens aéreas, de radar (SAR) ou de satélite, neste caso podem-se conhecer os padrões procurados e todas as regiões da imagem podem ser agrupadas em **classes predefinidas**. O reconhecimento de padrões é também muito útil na indústria, por exemplo, na inspeção de tecidos, onde as diferentes tramas e mesmo as estampas podem ser analisadas como padrões: uma vez reconhecidas no tecido, é possível avaliar a existência de defeitos e mesmo catalogá-los em relação os tipos de fios, máquinas e cores.

Como sugerido nos exemplos acima o reconhecimento pode ser feito por diferenciação ou por classificação (ou mesmo ambos). Para realizar qualquer tipo de reconhecimento por classificação é necessário ter inicialmente uma descrição o mais detalhada possível das classes padrão que se deseja reconhecer. Desta descrição serão projetadas formas possíveis de se extrair as informações quantitativas a serem obtidas analisando-se as imagens. Tais informações consistem nas chamadas **características** que serão extraídas da imagem. Em geral é necessário descrever os padrões com base em **características** que tenham propriedades invariantes a translação, escala e rotação. Podem ser utilizados diversos **tipos de descritores** para caracterizar o objeto/padrão. Cada tipo de descritor será mais adequado a **determinado aspecto** como: forma, dimensões cor, textura, entre outros.

Uma vez que os padrões possuem descrições, que os caracterizam e que são possíveis de serem obtidas da imagem, procede-se a etapa de classificação, que é a etapa

final do processo de reconhecimento. Esta etapa consiste em determinar em que classe a região da imagem se enquadra. A classificação depende das características que discriminam cada padrão dos demais. Para realizar a classificação torna-se necessário utilizar alguma técnica de decisão. Existem várias abordagens para isso, indo de critérios bem simples como distância mínima e funções discriminantes ao uso de algoritmos genéticos e até técnicas de inteligência artificial como redes nebulosas (fuzzy) ou neurais artificiais.

Na Figura 1 apresenta-se um esquema de todo o processo de reconhecimento.

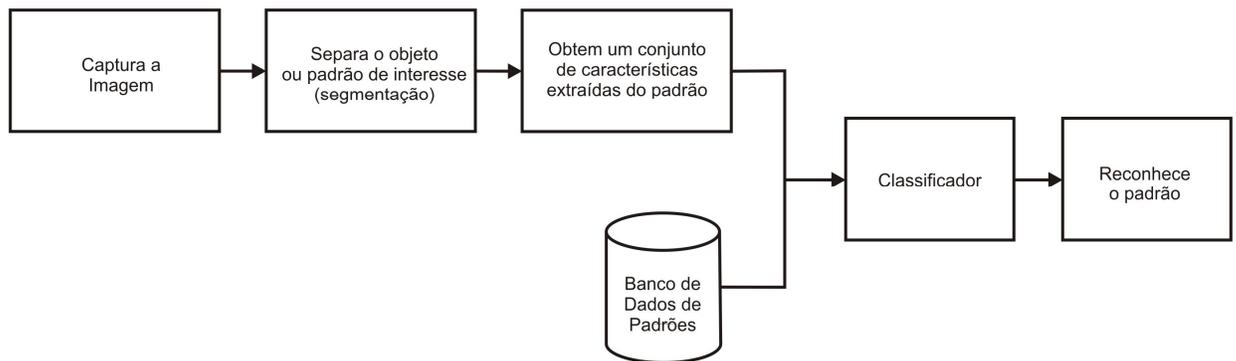


Figura 1 – Etapas de um sistema de reconhecimento de padrões.

De uma forma mais genérica, pode-se considerar que o objeto ou qualquer aspecto que está sendo reconhecido é um padrão. Deste modo quando se descrever uma técnica de reconhecimento de padrões, estes passam a incluir objetos e qualquer elemento predefinido em seu escopo. Todos os aspectos relacionados a este reconhecimento, da segmentação e extração das características às suas classificações, serão considerados neste capítulo.

A textura é um aspecto que está presente na maioria das imagens naturais, sendo fundamental para o reconhecimento pela visão humana e servindo como excelente descritor, contribuindo, em muitos casos, na melhoria da exatidão do processo de reconhecimento, descrição e classificação de imagens. Mas na análise de imagens digitais, o reconhecimento de texturas, realizado intuitivamente pela visão humana, se reveste de grande complexidade. Em casos em que o padrão é uma textura e não um objeto, ou ainda quando a textura do objeto é uma característica relevante para o seu reconhecimento, deve-se utilizar técnicas especiais de extração de características que permitam realizar o reconhecimento de texturas. Este tipo de características, pelas suas particularidades não serão tratadas neste capítulo, mas sim terão o próximo totalmente dedicado a elas.

A identificação dos padrões/objetos em uma imagem, em muitas aplicações, também depende de uma fase prévia: **a segmentação**. Só é possível reconhecer um dado objeto se este estiver “separado” dos demais elementos contidos na imagem. Observa-se

que esta etapa é fundamental nestas aplicações e consiste no primeiro estágio do processo de reconhecimento. A próxima seção aborda esse assunto.

1. Segmentação

Pode-se considerar que o primeiro passo para análise e entendimento da imagem consiste em particioná-la de modo que seja possível explicitar regiões representativas. Estas partições devem ser obtidas a partir das características dos valores de tom de cinza (ou cores) dos pixels da imagem, de texturas semelhantes, de formas pré-definidas, etc.

A segmentação é muito natural e fácil para o ser humano, porém complexa para o computador. Segmentar para os olhos humanos é simples, pois os sistemas biológicos reconhecem e interpretam os objetos assim que os vêm a partir de um conjunto de informações que são combinadas e processadas em paralelo no cérebro. Não há visão humana sem interpretação. Para o ser humano ver, ele precisa interpretar o significado das variações de intensidade de luz e cor que chegam a seus olhos. A literatura oftalmológica apresenta relatos de casos de recém nascidos, portadores de problemas visuais que tiveram seus problemas de visão corrigido, devido ao progresso da medicina, depois de estarem adultos. Estes adultos embora tecnicamente podendo enxergar, continuaram “cegos” pois já tinham passado pela idade possível de seus cérebros aprenderem a distinguir os significados das mudança de intensidade da luz captadas pelos seus olhos e, que por consequência, não conseguem distinguir os objetos a sua volta: eles.

O processo de segmentação consiste em dividir uma imagem em regiões que possuem o mesmo conteúdo no contexto de uma aplicação. O objetivo da segmentação é fazer com que os objetos e as áreas de interesse em uma imagem tenham os seus pixels agrupados e destacados dos demais. Por exemplo, na identificação de queimadas na Amazonia, a análise da área com nível de calor muito alto ou coberto por fumaça pode ser feita por diferentes tipos sensores de Landsat. As queimadas, neste caso, são as áreas de interesse (ou padrões) que precisamos ter os pixels identificados na imagem. Na maioria das aplicações, a segmentação de imagens é um estágio prévio de processamento para reconhecimento ou identificação das imagens.

Existem diversas formas de segmentar imagens. A segmentação pode ser baseada nas **descontinuidades e nas similaridades** dos diferentes aspectos da imagem.. Podem ser baseadas em **limites** (ou **bordas**) ou em **áreas (regiões)**. As descontinuidades são representadas pelas mudanças bruscas nos níveis de tons, cores e texturas. As similaridades baseiam-se no quanto esses aspectos podem ser comuns pelo diversos pixels (figura 7.2).

Quando a característica usada é a cor ou o tom de cinza da imagem a segmentação baseada nas **descontinuidades** pode ser resolvida com o uso de **filtros de detecção de pontos, linhas e contornos** da imagem (veja capítulo de Filtragem) e a segmentação baseada nas **similaridades** corresponde à busca de **limiars** adequados de **binarização ou multiníveis** (veja seção de limiarização da imagem).

Pixels (e mesmo os elementos de texturas ou texels tratados no próximo capítulo) são associados a regiões usando-se algum critério, atributo ou predicado que os distinguem dos demais pixels ou texels da imagem. Os dois princípios envolvidos neste processo são: similaridade de valor dos atributos ou predicados e proximidade espacial. Dois pixels (ou texels) podem ser associados a uma mesma região se estes possuem características (de intensidade de cinza, cores ou qualquer predicado pré-definido) semelhantes e se estão próximos. Ou seja, satisfizerem a critérios de vizinhança e tiverem as mesmas características associadas à região.

A idéia geral em muitos dos métodos de segmentação é agrupar, de alguma forma, pixels ou grupo de pixels com mesma propriedade. Por exemplo, pode-se dividir a imagem em regiões de textura correspondente à área de campos, tipos de plantação ou floresta para medir, por exemplo, a extensão de uma região e fazer previsões de colheitas em uma aplicação para a agricultura. Por isso, a segmentação adequada é um assunto muito importante nas áreas de Análise de Imagens, Visão Computacional e Reconhecimento de Padrões. É fundamental identificar sub-imagens que representem elementos de interesse na imagem.

Existem dificuldades inerentes ao processo de segmentação automática de imagens, e por isso, deve-se considerar que as fronteiras ou bordas das regiões podem não possam não ser muito nítidas e são muitas vezes irregulares e imprecisas. Na segmentação em larga escala especialmente por processos automáticos, existe uma grande necessidade em controlar a iluminação do ambiente de captura. Ambientes bem controlados, com grandes contrastes, tendem a facilitar a interpretação de imagens. Ambientes externos em geral, apresentam mais dificuldades, pois estão sujeitos a variação de iluminação. A existência de sombras muda os aspectos da região e tende a dar uma falsa impressão acerca da região a ser segmentada.

Quando as segmentações são baseadas em conjunto de características ou atributos é geralmente associada a identificação de grupos de pixels ou áreas da imagem, é denominado de técnica baseada em regiões ou **segmentação orientada a regiões**. Uma **região** em uma imagem é um **grupo de pixels conectados** com propriedades similares. Regiões são importantes para a interpretação de uma imagem, pois podem corresponder a objetos ou aspectos relevantes de uma cena. Uma imagem pode conter vários objetos, e, cada objeto pode conter várias regiões correspondentes a partes diferentes do objeto. Neste caso todos os pixels semelhantes são considerados pertencentes a uma mesma região sendo agrupados em conjunto e demarcados para indicar que pertencem a uma determinada região.

1.1. Segmentação Baseada em Regiões

A partição da imagem baseada no conteúdo de grupos de pixels é chamada de segmentação orientada ou baseada em regiões. De uma maneira ideal, uma partição será associada a um padrão, objeto ou parte dele numa imagem. Portanto, para uma imagem

ser interpretada corretamente, é necessário particioná-la em regiões que correspondam aos objetos ou padrões. O processo de segmentação baseada em regiões consiste na separação destas regiões. A associação destas aos objetos é uma etapa seguinte: a **interpretação** ou identificação (*label*). Em alguns casos, no entanto, principalmente se técnicas de inteligência artificial estiverem sendo usadas, os dois passos podem não ser associados.

Os segmentos de uma imagem devem ser uniformes e homogêneos com relação a característica considerada (cor, textura, etc). As regiões devem ser simples e não apresentar outras em seu interior. Diferenças entre as regiões adjacentes devem ser significativas, dentro da tolerância especificada. As bordas ou limites devem ser nítidos. As regiões devem possuir algumas propriedades ou predicados diferentes para ser possível segmentá-los. Esta propriedade deve permitir, por exemplo, destacar uma superfície que representa um osso dentro de uma radiografia, ou uma peça dentre várias sendo submetida a um controle de qualidade ou ainda uma área de estradas de uma imagem de satélite usada para planejamento urbano.

Alguns aspectos relevantes devem ser considerados no processo de segmentação de imagens em regiões. A **formulação básica** usada para realizar a segmentação parte das seguintes premissas:

- (1) a região representada pelos pixels a serem agrupados deve ser homogênea dentro de algum fator de tolerância e quanto a algumas características (predicados) predefinidas;
- (2) só se consideram regiões fechadas, delimitadas por fronteiras (bordas) contínuas que separam seu interior do dos outros segmentos (regiões da imagem);
- (3) cada interior destas regiões possui pontos conectados que só pertencem a uma região, ou em outras palavras os **predicados** devem ser adequados para caracterizar univocamente a região; e
- (4) o conjunto de todas as regiões deve formar a imagem.

A forma mais comum utilizadas para representar as regiões são as notações e os diagramas (de Venn) da teoria matemática de conjuntos. Sendo p_n os pixels representativos de uma região R_n , contendo N pixels, um segmento ou região pode ser descrito matematicamente como a união de todos os pixels conectados que satisfaçam a certa propriedades, ou seja os predicados : $P(pi)$:

$$R_n = \bigcup_{i=1}^N P_i \quad \text{onde } P(pi) = \text{Verdade para } i = 1, 2, \dots, N$$

Como as regiões adjacentes não devem possuir pixels em comum, ou seja, as bordas não são consideradas pertencentes à região e sim aos limites tem-se para cada i, j :

$$R_i \cap R_j = \emptyset$$

$$R_i \cap R_j = \phi$$

Formalmente, a segmentação de imagens é definida como um tratamento que visa particionar uma imagem f em um subconjunto composto de n regiões R_i tais que:

$$0 < i \leq n \quad \forall i \quad R_i \neq \emptyset$$

$$\forall i, j; i \neq j; R_i \cap R_j = \emptyset$$

$$f = \bigcup_i R_i$$

Na prática todos os 4 aspectos mencionados acima na **formulação básica** são utópicos por que áreas homogêneas são abstrações em imagens reais e as bordas em geral são irregulares e pouco nítidas. Além do mais, as regiões adjacentes tendem a se confundir e perder as bordas. Como regra geral a identificação de segmentos é específica e típica para cada aplicação. Na segmentação, a imagem é dividida em regiões com propriedades comuns (intensidade, cor, textura, etc). O nível de subdivisão da imagem depende do objetivo que se pretende alcançar. A segmentação se encerra quando as partes constituintes de interesse forem isoladas

Em geral, em uma abordagem clássica a segmentação em região de interesse é baseada em técnicas de **subdivisão e fusão da imagem** em regiões homogêneas ou **crescimento de regiões por agregação de pixels** (Gonzalez e Woods, 2000, Watt e Policarpo, 1998). Analisa-se a seguir estas técnicas e outras duas que se mostram cada vez mais utilizadas nos sistemas mais modernos: **a segmentação baseada em agrupamentos** (*clustering*) e em **janelas** (*windows*).

A técnica conhecida como Crescimento de Regiões, por exemplo, agrupa iterativamente regiões com propriedades similares (i.e., nível de cinza, textura, cor etc.) a partir de um grupo inicial de pixels (semente) em regiões maiores, até alcançar um determinado critério de parada.

1.2- Técnicas de segmentação baseadas em crescimento de regiões

Crescimento de regiões é um processo iterativo de agrupamento de pixels ou regiões com **predicados** em comuns. Regiões homogêneas em relação ao **predicado** e adjacentes no espaço são agrupadas.

O processo de segmentação inicia a partir de um pixel ou um conjunto de pixels (denominado de “semente”). Pode haver qualquer número de sementes. Para cada semente avalia-se os **predicados** dos pixels vizinhos (ou região). Como exemplo, de **predicado pode-se ter**: cor em (RGB) com mais ou menos 5% de variação, e textura definida pela probabilidade máxima das matrizes de co-ocorrência da textura estar a horizontal. A agregação das regiões é feita quando o critério de similaridade ou de

decisão do predicado for verdadeiro.

Alguns problemas precisam ser considerados no processo de crescimento de regiões. Inicialmente é necessário que a seleção de “sementes” represente adequadamente as regiões de interesse. Também é preciso selecionar adequadamente as propriedades para a inclusão de pontos nas várias regiões durante o processo de crescimento.

Outro problema é o estabelecimento de um critério de parada. O crescimento de uma região deveria parar quando não houvesse mais pixel que satisfizesse os critérios de inclusão naquela região.

Uma vez que critérios como: intensidade, textura e cor são locais e estáticos por natureza, não considerando o histórico do processo de crescimento de região. Critérios adicionais que utilizam o conceito de tamanho, similaridade entre o pixel candidato e os pixels acrescidos até aquele momento e a forma da região sendo operada são desejáveis para aumentar o poder do algoritmo de crescimento de regiões.

Exemplo de opções que aumentam a capacidade do algoritmo são a inclusão das fatias anteriores e posteriores em imagens médicas de ressonância magnética e tomografia, ou o uso das informação dos quadros posteriores e anteriores em análise de vídeos ou imagens em movimento.

Normalmente, o crescimento de regiões emprega um conjunto de descritores baseados em intensidade e em propriedades espaciais (como descritores de texturas) de uma única fonte de imagens. Porém, a informação de conectividade ou de adjacência deve ser considerada no processo de crescimento de regiões, pois o emprego destes descritores isoladamente pode conduzir a resultados enganosos (GONZALEZ; WOODS, 2000).

No caso de não ser possível a definição de um ponto semente inicial, todas as regiões da imagem pode ser consideradas sementes em relação ao predicado. No caso de serem obtidas diversas segmentações diferentes algum critério de decisão sobre a melhor forma final pode ser incluído.

1.3 - Técnicas de segmentação baseadas em divisão e fusão de regiões

Divisão e fusão é um método que subdivide uma imagem em quatro blocos e testa cada um destes blocos verificando se os pixels (ou *textels*) pertencentes aos mesmos atendem a algum critério de homogeneidade.

Os blocos que atenderem ao critério não serão mais divididos. O bloco que não atender será subdividido em blocos menores. Esse processo é repetido iterativamente até um critério de parada ser satisfeito. Em seguida, é realizada a junção dos blocos vizinhos que sejam homogêneos (GONZALEZ; WOODS, 2000). Esse método geralmente é associado a uma estrutura *quadtree* (ou seja, uma árvore em que cada nó possui exatamente quatro descendentes) que possibilita decompor e agrupar partes de uma imagem (seção 4.47 do volume 1). Mas outras técnicas podem se usadas como por uso de algoritmo genético para adaptar o processo de segmentação. Uma outra abordagem para segmentação de imagens utilizando algoritmo genético (AG) pode ser vista em (Chun, 1996), onde é proposto uma função *fuzzy*, para medir o grau de separação entre regiões segmentadas e a intensidade da aresta ao longo das fronteiras da região.

Juntamente a esta função utilizam um AG para encontrar uma boa e utilizável segmentação de região, a qual deve maximizar a qualidade de segmentação das regiões geradas durante o processo de divisão e fusão. Para tanto o projeto dos mecanismos de *crossover* e mutação visa a divisão e fusão de regiões. A Figura 2, fornece uma rápida idéia deste processo.

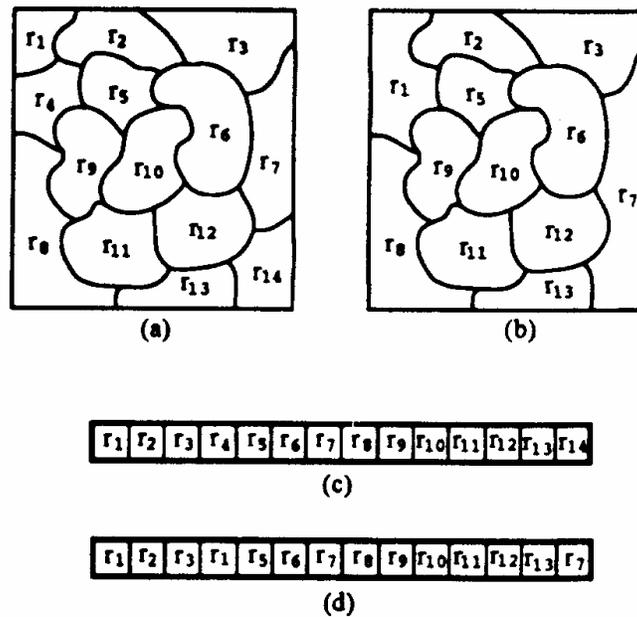


Figura 2 - Processo de divisão e fusão de regiões.

A Figura 2 apresenta em (a) e (b) uma idéia da aparência do processo de divisão de regiões, assim como o que acontece quando ocorre uma fusão de regiões. Acompanhando as modificações dos cromossomos em (c) e (d), observa-se que a posição 4 em torna-se r_1 pois houve uma fusão de r_1 com r_4 e a posição 14 torna-se r_7 pois houve uma de r_7 com r_{14} .

Ainda considerando AG na segmentação Andrey e Tarroux(1993), apresentam uma abordagem de análise de imagens baseada em uma modificação simplificada de sistemas de classificação, onde a função de classificação é implementada tal qual um conjunto espacialmente estruturado de regras produzidas por codificação binária. Desta maneira a classificação pode ser iterativamente modificada usando um algoritmo

genético distribuído. Esta abordagem permite a utilização do AG não como uma técnica de otimização alternativa, e sim de maneira a gerar automaticamente um programa de segmentação de imagem.

1.4 Clusterização ou Agrupamentos

Dado um conjunto X , um problema de Clusterização consiste em se agrupar os objetos (elementos) de X de modo que objetos mais similares fiquem no mesmo cluster (grupo) e os objetos menos similares sejam alocados para clusters distintos, figura 3.

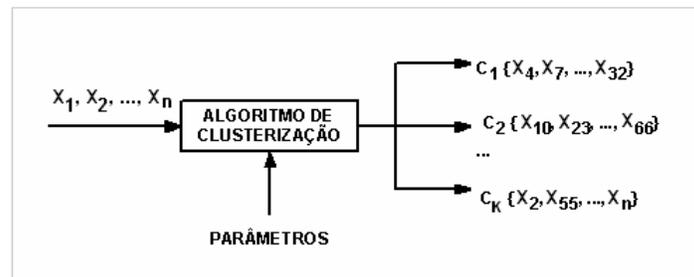


FIGURA 3 - Esquema funcional de um algoritmo de clusterização

Problemas de clusterização podem ser definidos formalmente da seguinte maneira (OCHI, 2004):

Dado um conjunto com n elementos $X = \{X_1, X_2, \dots, X_n\}$, o problema de clusterização consiste na obtenção de um conjunto de k clusters, $C = \{C_1, C_2, \dots, C_k\}$, tal que haja uma maior similaridade entre os elementos contidos em um cluster C_i do que qualquer um destes com os elementos de um dos demais clusters do conjunto C . O conjunto C é considerado uma clusterização com k clusters caso as condições das equações sejam satisfeitas:

$$\bigcup_{i=1}^k C_i = X$$

$$C_i \neq \emptyset, \text{ para } 1 \leq i \leq k$$

$$C_i \cap C_j = \emptyset, \text{ para } 1 \leq i, j \leq k, i \neq j$$

Se o valor de k for conhecido, o problema é referenciado na literatura como “problema de k -clusterização”. Caso contrário, o problema é referenciado como “problema de clusterização automática” e a obtenção do valor de k faz parte do processo de solução.

A clusterização é uma ferramenta útil para o estudo e compreensão do comportamento de dados em diferentes **situações e tem sido** empregadas na solução de diversos problemas nas mais variadas áreas do conhecimento e em especial na segmentação por textura (Nunes, 2006).

O objetivo da clusterização é agrupar os elementos de um conjunto de tal forma que os grupos formados sejam constituídos por elementos que possuam maior similaridade com os elementos do mesmo grupo do que com qualquer dos elementos de outros grupos.

Para medir o quanto um elemento é similar a outro, a fim de determinar se devem pertencer ou não a um mesmo cluster, são utilizadas medidas de similaridade. O critério de similaridade mais comum quando se utilizam atributos numéricos baseia-se nas funções de distância.

Para empregar estas funções é preciso representar cada elemento como um vetor no espaço n dimensional das características. Neste caso, quanto menor for a distância entre um par de elementos, maior é a similaridade entre eles. Espera-se que a distância entre objetos de um mesmo agrupamento seja significativamente menor do que a distância entre objetos de agrupamentos diferentes.

Dentre as medidas de distância mais utilizadas, encontram-se a distância “city-block ” (para $r = 1$) e a distância euclidiana (para $r = 2$),

$$d(X_i, X_j) = \left[\sum_{l=1}^n |x_{il} - x_{jl}|^r \right]^{\frac{1}{r}}$$

A busca pela melhor solução no espaço de soluções possíveis torna o processo de clusterização num problema NP-Difícil. A utilização de métodos exatos para obtenção da solução ótima fica impraticável, uma vez que a verificação exaustiva de todas as configurações de agrupamentos possíveis é computacionalmente inviável. Há

aproximadamente $\frac{1}{k!} \sum_{i=0}^k (-1)^i \binom{k}{i} (k-i)^k$ maneiras de repartir n elementos em k

agrupamentos.

Por exemplo, existem os seguintes números de soluções possíveis para se combinar 10 elementos, 100 elementos e 1000 elementos em 2 clusters, respectivamente: 512 soluções, $6,33825 \times 10^{29}$ soluções e $5,3575 \times 10^{300}$ soluções.

A fim de reduzir a complexidade na solução do problema, utilizam-se metaheurísticas capazes de fornecer soluções sub-ótimas em tempo satisfatório. Entretanto, as metaheurísticas são normalmente desenvolvidas para certas classes de problemas, não sendo genéricas o suficiente para obter bons resultados em toda gama de aplicações de clusterização existentes.

Na literatura são encontradas diversas metaheurísticas para a solução de problemas de clusterização (BERKHIN, 2002), basicamente elas podem ser classificadas em **métodos hierárquicos** e métodos de **particionamento** (FASULO, 1999).

Os Métodos Hierárquicos se dividem em Métodos Hierárquicos Aglomerativos (*bottom-up*) e Métodos Hierárquicos Divisivos (*top-down*).

Os Métodos Hierárquicos Aglomerativos (*bottom-up*) iniciam com n clusters de um elemento e os clusters vizinhos são agregados até que o número desejado de clusters seja alcançado. Os Métodos Hierárquicos Divisivos (*top-down*) iniciam com um cluster de n componentes que é dividido sucessivamente até que o número de cluster desejado seja alcançado.

Os Métodos Hierárquicos geram uma hierarquia de clusters, normalmente representada através de uma estrutura em árvore, conforme exemplificado na figura 4.16.

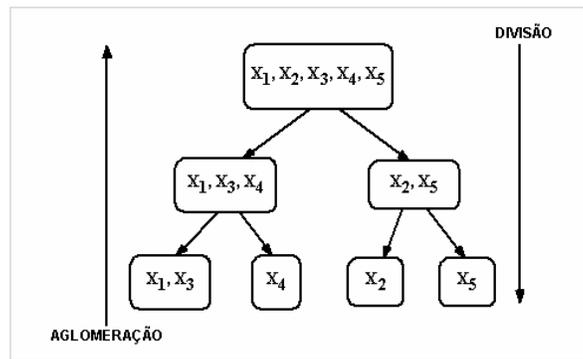


FIGURA 4 – Exemplo de árvore de clusters na clusterização hierárquica

Os algoritmos de clusterização hierárquica possuem a vantagem de lidarem com qualquer medida de similaridade e são aplicáveis a qualquer tipo de atributo (numérico ou categórico). Porém, tem como desvantagem o fato de não revisitarem os clusters construídos. Ou seja, no método hierárquico aglomerativo, uma vez realizado a fusão de dois objetos dentro de um cluster, estes não mais serão separados, permanecendo até o final do procedimento juntos no mesmo cluster. De igual forma, no método hierárquico

divisivo, uma vez separados dois objetos, eles não mais serão agrupados em um mesmo cluster (BERKHIN, 2002). Em imagens de satélite, devido ao grande número de objetos (pixels) presentes, esta variação raramente é utilizada por ser mais custosa computacionalmente (RICHARDS, 1995).

Os Métodos de Partição Baseados em Recolocação tem por objetivo particionar um conjunto de dados com n elementos em k grupos distintos de forma a minimizar um critério escolhido. Estes métodos tentam descobrir novos clusters realocando interativamente pontos entre os subconjuntos, de forma a melhorar os clusters gradualmente, o que não ocorre nos métodos hierárquicos.

Freqüentemente, os métodos de particionamento encontram clusters com qualidade superior (maior similaridade interna) aos dos encontrados pelos métodos hierárquicos (NG; HAN, 1994). Devido a este melhor desempenho, os algoritmos de particionamento normalmente são mais empregados. Dentre eles, os que são baseados em um ponto central (média dos atributos dos objetos - *k-means*) (ZHANG; HSU; DAYAL, 2001) ou em um objeto representativo para o cluster (*k-medoids*) (KAUFMAN; ROUSSEEUW, 1990).

A qualidade do resultado obtido com os métodos de particionamento depende da seleção coerente das seguintes variáveis: escolha dos atributos, homogeneização das variáveis, medidas de similaridade, critérios de agrupamento, escolha do algoritmo, e definição do número de *clusters*.

1.4.1 - Implementação com Algoritmo K-Means

Nesta seção, são apresentados os aspectos computacionais do algoritmo *K-Means* proposto para o problema de classificação de textura em imagens multiespectrais.

K-means é um método amplamente difundido, existindo muitas variações propostas na literatura e diversos nomes (*K-médias*, *isodata*, ou migração de médias). Em Sensoriamento Remoto, ele é bastante utilizado para executar procedimentos de classificação não supervisionada de imagens de satélite (SCHOWENGERDT, 1997).

K-means é um método de partição baseado em recolocação que necessita da definição a priori do número de agrupamentos k . O critério de custo a ser minimizado é definido em função da distância dos elementos em relação aos centros dos agrupamentos.

Usualmente, este critério é a soma residual dos quadrados das distâncias, geralmente a distância euclidiana. Entende-se por soma residual dos quadrados, a soma dos quadrados das distâncias dos elementos ao centróide do seu cluster. Para toda partição, a soma residual dos quadrados, será a soma de todas as somas dos quadrados das distâncias de cada grupo. Quanto menor for este valor, mais homogêneos serão os objetos dentro de cada grupo e melhor será a partição (COELHO; EBECKEN, 2001).

O elemento representativo de um cluster é o seu centróide, que possui um valor médio para os atributos considerados, relativos a todos os elementos do cluster. A utilização do centróide como elemento representativo de um cluster é conveniente apenas para atributos numéricos e possui um significado geométrico e estatístico claro, podendo, entretanto, receber muita influência de um único elemento que se encontre próximo à fronteira do cluster.

A partir de uma estimativa inicial das coordenadas dos centros dos agrupamentos (centróides), o algoritmo calcula a distância de cada ponto do conjunto a estes centróides. A seguir, o algoritmo aloca cada elemento do conjunto em um grupo, de acordo com a menor distância ao centróide correspondente. A nova estimativa das coordenadas dos centróides é calculada pela média aritmética das coordenadas dos pontos associados a cada grupo, veja figura abaixo

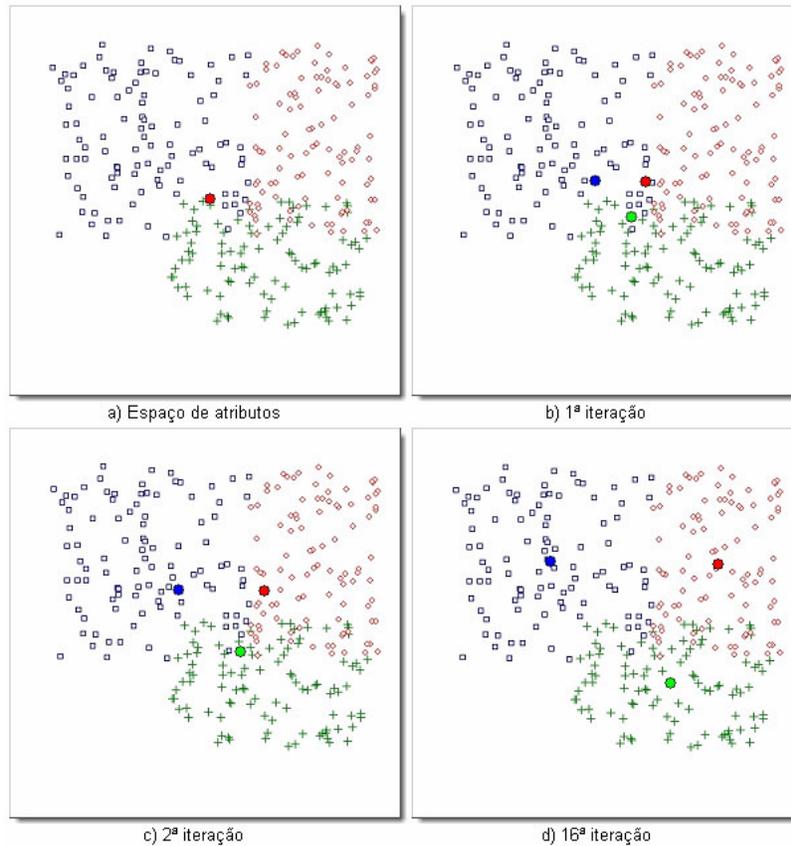


FIGURA 5 - Aplicação do método k-means

O método *K-means* é sensível ao particionamento inicial realizado, em virtude da escolha das coordenadas dos k centróides dos *clusters* ser feita inicialmente de forma aleatória. A partir deste primeiro particionamento, o algoritmo realiza uma busca de um ponto de máximo para o seu critério de parada. Não há garantias de que o algoritmo encontre o máximo global, sendo possível encontrar distintas soluções em diferentes execuções do algoritmo.

O *K-means* é um tipo de algoritmo de classificação, que pode ser utilizado para a classificação não-supervisionada. Abaixo é apresentado um algoritmo básico:

- (1) Determinar as posições iniciais dos k centróides dos *clusters*;
- (2) Alocar cada elemento ao cluster do centróide mais próximo;
- (3) Recalcular os centros dos *clusters* a partir dos elementos alocados;
- (4) Repetir os passos de 2 a 4 segundo algum critério de convergência.

Na implementação desenvolvida, inicialmente é necessário selecionar um conjunto de amostras de treinamento para cada região de interesse. Essas amostras são regiões da

imagem obtidas por janelas de tamanho $M \times M$ (3×3 , 5×5 , ..., 21×21) definido pelo usuário. Cada amostra selecionada passa a representar um ponto no espaço euclidiano tridimensional. As coordenadas dos pontos são definidas pelo valor do CVE da região de amostra calculados para cada canal RGB (CVE_{RED} , CVE_{GREEN} , CVE_{BLUE}). Então, as amostras do conjunto de treinamento são agrupadas por meio do algoritmo de clusterização *K-means*.

A fim de evitar um número excessivo de amostras, somente amostras com coordenadas significativas são incluídas no grupo de amostras de treinamento. Entende-se por valores significativos, uma amostra cuja coordenada provoque uma alteração na configuração do cluster. Em seguida, é realizada a clusterização das amostras e os centróides das classes são calculados por meio do algoritmo *K-Means*, veja figura abaixo:

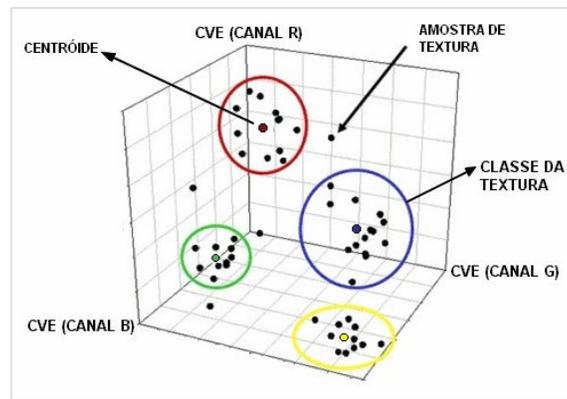


FIGURA 6 – Clusterização das amostras baseado no valor do CVE em cada banda espectral

Finalmente, os valores dos centróides (CVE's para os canais RGB) são utilizados na classificação do restante da imagem, figura abaixo

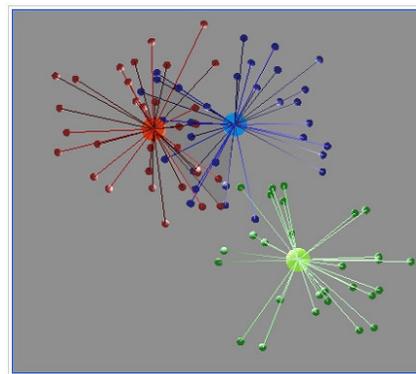


FIGURA 7 – Centróides das amostras de treinamento no espaço euclidiano tridimensional

O método proposto utiliza um processo de classificação híbrida, com a realização de uma classificação supervisionada inicial das amostras de treinamento para obtenção dos centróides dos clusters, seguido de uma classificação não supervisionada do restante da imagem, baseado na distância das amostras de teste a estes centróides.

Para segmentar somente textura de interesse na imagem, é verificado se a distância da amostra de teste é maior do que uma distância de tolerância. Caso esta distância seja menor, o pixel será rotulado com a cor da classe correspondente. Caso contrário, ele permanecerá com sua cor original. O critério de tolerância adotado foi a maior distância da amostra de treinamento ao centro do seu cluster.

1.4- Segmentação baseada em janelas (*windows*).

O uso de janelas para características baseadas em áreas é muito comum em análises baseadas em multiresolução ou texturas, pois pixels sozinho dificilmente seriam representativos nestes casos. A maior desvantagem desta técnica é que os limites entre as regiões pode apenas ser determinado dentro de um distância de w pixels, onde w representa o número de pixels que constituem a janela (Parker, 1997).

O método pode ser descrito por:

- para cada pixel da imagem, substitua seu valor pelo do predicado em análise, por exemplo, o desvio padrão dos tons da região de $w \times w$ pixels centrada nele;
- depois divide a imagem em regiões baseado em valores do desvio padrão

O uso de apenas um predicado pode não funcionar bem muitas vezes, mas o método pode ser generalizado para qualquer numero de predicados e tecnica de desição quanto aos limiares das regiões.

1.5 - Outras Técnicas de Segmetação

Atualmente, existe uma grande variedade de técnicas de segmentação de imagens que podem ser classificadas de diferentes formas. Dentre as técnicas mais comuns empregadas na segmentação de imagens estão:

- Domínio espacial.
- Transformação para um espaço de medida específico.
- Técnicas de segmentação baseadas em morfologia matemática.
- Contornos ativos.
- Domínio da Frequência

No grupo do **domínio espacial** executa-se a segmentação na própria imagem, sem utilizar a transformação linear, baseado em diversas medidas calculadas sobre a imagem. Um exemplo é o uso de tons de cinza ou cores para segmentar imagens.

O enfoque voltado ao espaço de **medida específico de interesse** é o mais simples destes. Neste, a imagem é considerada representada em um espaço Euclidiano. Executa-se uma transformação linear para outro espaço de representação, processa-se a imagem e executa-se a transformação inversa. A segmentação é feita na forma final no "espaço de domínio" da imagem. Um exemplo disso é o uso de transformadas de Hough na segmentação.

No terceiro grupo, com base na morfologia matemática a principal ferramenta de segmentação é baseada na transformação Watershed, também conhecida como divisor de águas. Devido ao amplo espectro de aplicações práticas e sua simplicidade de implementação, a Morfologia Matemática tornou-se uma importante metodologia para análise de imagens, no entanto não será tratada neste livro. Para segmentação baseada na Morfologia Matemática, recomenda-se a consulta a outras obras (FACON, 1996, MEYER, 1993; SOILE, 1991).

Os modelos de contornos ativos são técnicas que visam à extração das bordas dos objetos da cena. Estas técnicas se caracterizam pelo ajuste de uma curva (*spline*) sobre uma imagem definindo o contorno do objeto segmentado. Geralmente são aplicados conjuntamente com técnicas de filtragem para detecção de pontos de bordas. A inicialização é realizada com um contorno de configuração arbitrária que evolui até se ajustar ao objeto de interesse. Devido ao seu comportamento dinâmico, os contornos ativos são também conhecidos como "modelos deformáveis". Para mais detalhes deste modelos veja os trabalhos de (GIRALDI, 2000, VENETSANOPOULOS, 2001 e GIRALDI; STRAUSS e OLIVEIRA, 2001).

O último grupo trabalha da mesma forma que o primeiro, ou seja, em um outro espaço que não o da imagem propriamente dita. Neste caso, o Espaço de Frequência onde consideramos a imagem como um sinal e as suas variações no espaço da imagem são utilizadas como frequências. Isto nada mais é do que restringir o primeiro método ao uso de transformadas de Fourier.

2. Propriedades do Pixel

Uma representação adequada de um pixel é como uma área retangular ou quadrada. Três aspectos são muito importantes neste contexto representação: a vizinhança, a conectividade e as distâncias

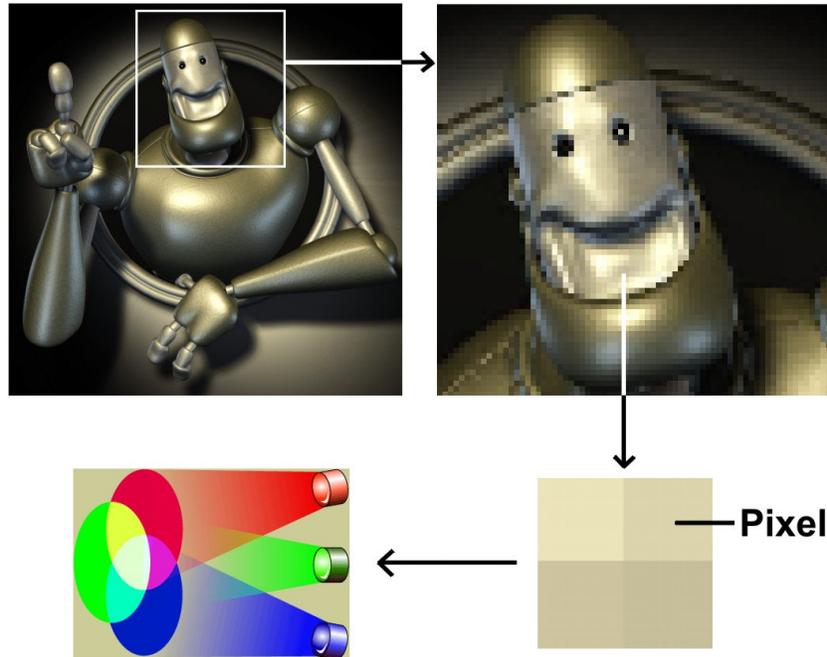


Figura 8 – Em imagem coloridas tem-se 3 cores associada a cada pixel da imagem. Pixels vizinhos com características semelhantes representarão os objetos, regiões ou padrões das imagens. Na figura acima, tons semelhantes caracterizarão a região da boca do personagem.

2.1. Vizinhaça em Pixel (4-Vizinhaça e 8-Vizinhaça)

A amostragem uniformemente espaçada define pixels quadrados. Um problema que aparece nesse momento é: quais são os vizinhos de um determinado pixel? Essa pergunta é fundamental para algoritmos de preenchimento de áreas e algoritmos de determinação de continuidade de objetos na imagem.

Um pixel p nas coordenadas (x,y) tem quatro vizinhos horizontais e verticais cujas coordenadas são dadas por $(x+1,y)$, $(x-1,y)$, $(x,y-1)$, $(x,y+1)$. Este conjunto de pixels, ditos estar na 4-vizinhaça de p , recebem a denotação $N_4(p)$. Os quatro pixels vizinhos das diagonais de p têm coordenadas $(x+1,y+1)$, $(x+1,y-1)$, $(x-1,y+1)$, $(x-1,y-1)$ e recebem a notação $N_D(p)$. Junto com a 4-vizinhaça, este conjunto é chamado 8-vizinhaça e recebe a notação $N_8(p)$.

A figura 9 mostra estas vizinhaças de um pixel. As vizinhaças N_4 e N_D descrevem elementos equidistantes do pixel considerando se os pixels forem quadrados.

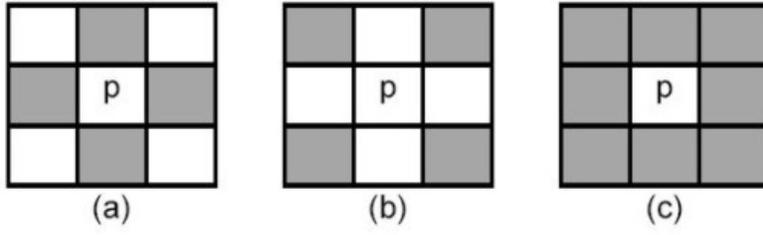


Figura 9 – Conceitos de (a) 4-vizinhança de p, (b) vizinhança diagonal de p, (c) 8-vizinhança de p