

5 Processamento e Análise Digital de Imagens

A área de processamento digital de imagens vem evoluindo continuamente ao longo dos anos, com um aumento significativo de estudos envolvendo morfologia matemática, redes neurais, processamento de imagens coloridas, compressão de imagens, reconhecimento de imagens e sistemas de análise de imagens baseados em conhecimento (Gonzalez & Woods, 2002).

O processamento de imagens tem como funções facilitar a visualização da imagem ou adequá-la para análises quantitativas através de correções de defeitos ou realces das regiões de interesse nas imagens; e a extração e tratamento de dados quantitativos, feitos pelo próprio computador. (Gomes, 2001).

Conforme as diferentes funções citadas, o processamento de imagens pode ser entendido como o conjunto de duas técnicas, a saber: (a) Processamento Digital de Imagens (PDI), que consiste no preparo da imagem para posteriores análises, em que são realizadas operações matemáticas que alteram os valores dos pixels; e (b) Análise Digital de Imagens (ADI), que compreende a análise quantitativa do processo a partir da qual as regiões, partículas e objetos identificados na imagem são medidos. Ao longo deste trabalho, serão mantidas as expressões PDI e ADI para designar as expressões, respectivamente, “Processamento Digital de Imagens” e “Análise Digital de Imagens”. Além destes termos, será utilizado o termo PADI (Processamento e Análise Digital de Imagens), quando estiver englobando as duas categorias.

A Figura 22 mostra o fluxograma da sequência padrão de PADI, o qual é dividido em três principais etapas: aquisição da imagem, PDI e ADI.

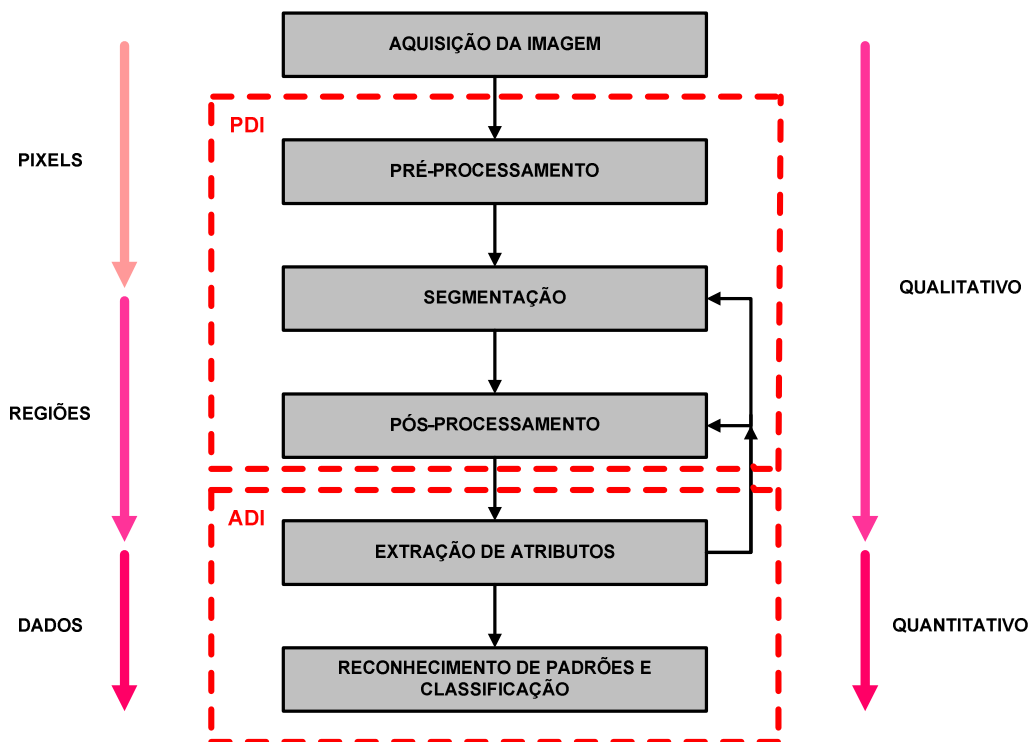


Figura 22 – Sequência padrão de PADI (Gomes, 2001)

O primeiro passo no processo é a aquisição da imagem, com a formação e a digitalização da imagem.

Após a obtenção da imagem digital, o próximo passo consiste no pré-processamento da imagem da etapa anterior. O pré-processamento serve para melhorar a imagem, para que as próximas etapas sejam bem-sucedidas.

Após o pré-processamento, vem a segmentação. O papel básico da segmentação é separar a imagem em partes ou objetos de interesse.

Muitas vezes o resultado da segmentação não é adequado. Para corrigir esses defeitos, realiza-se o estágio de pós-processamento.

A imagem de saída do pós-processamento está pronta para fornecer dados numéricos através da etapa de extração de atributos, que é a parte quantitativa da sequência. Nesta etapa, algumas características dos objetos nas imagens são extraídas, resultando em algum tipo de discriminação entre as classes de objetos.

O último estágio envolve o reconhecimento de padrões e classificação. O reconhecimento atribui uma descrição a um objeto, baseado na informação fornecida pelo seu descritor. A classificação confere um significado para um conjunto de objetos reconhecidos.

As setas ao lado do fluxograma indicam os níveis com que se trabalha em cada etapa. No pré-processamento e na segmentação, as operações são feitas nos pixels das imagens, resultando em imagens com objetos ou regiões de pixels contíguos de valores iguais. As medidas para a geração de dados são feitas na extração de atributos e estas medidas são utilizadas no fim do processo, na etapa de reconhecimento de padrões e classificação, de forma a obter dados mais complexos. Do pré-processamento ao pós-processamento a análise é considerada qualitativa e a partir da extração de atributos, quantitativa (Gomes, 2001).

As etapas de segmentação, pré-processamento e pós-processamento fazem parte do PDI, enquanto as etapas de extração de atributos e reconhecimento de padrões e classificação fazem parte do ADI. Vale ressaltar que a sequência é adaptativa para o problema envolvido, ou seja, não há a obrigação de todas as etapas serem realizadas e, além disso, o fluxograma poderá retroceder a qualquer outra etapa, caso o resultado obtido não seja satisfatório.

Todos estes itens do PADI serão explicados a fundo nos próximos capítulos, dentro do contexto em que foram utilizados no presente trabalho.

5.1. Aquisição da Imagem

O bloco de aquisição da imagem tem como objetivo a obtenção de uma imagem digital. Para que o leitor tenha conhecimento suficiente para o entendimento das demais etapas do PADI, nesta seção serão apresentados conceitos básicos sobre a imagem digital.

Para que uma imagem digital seja adquirida são necessários: um instrumento formador de imagem, neste caso, o próprio microscópio ótico; um dispositivo físico sensível a uma faixa do espectro eletromagnético, capaz de produzir um sinal elétrico proporcional ao nível de energia percebida; e um *digitalizador*, que converterá a saída elétrica do dispositivo de sensoriamento físico para a forma digital (Gonzalez & Woods, 2002).

A tecnologia usada em sensores de imageamento é baseada principalmente em dispositivos de acoplamento de carga (“charge-coupled devices” – CCDs). Este dispositivo é formado por uma matriz de “poços”, que armazenam carga proporcionalmente à quantidade de luz recebida. Posteriormente, as cargas são transferidas para um porto de saída através de registradores de deslocamento. As

câmeras digitais utilizam o circuito CCD e um sistema de digitalização interno para a geração de uma imagem digital, a qual pode ser transferida diretamente para o computador (Paciornik, 2010).

Uma imagem digital pode ser compreendida como uma matriz, cujos índices de linhas e de colunas identificam um ponto na imagem. Cada coordenada (linha, coluna) é um elemento da matriz e possui o valor correspondente ao nível de cinza, ou à cor, naquele ponto da imagem. Esses elementos da matriz são os *pixels*, assim denominados pela abreviação de “picture elements” (elementos de figura).

Uma das formas de apresentar a distribuição de intensidade de pixels numa imagem digital é feita através do *histograma*. O histograma de uma imagem digital com k níveis de cinza é uma função discreta dada pela eq. (1):

$$p(k) = \frac{n_k}{n} \quad (1)$$

Onde:

k = nível de cinza, podendo variar entre 0 (preto) e 255 (branco);

n_k = número de pixels na imagem com o nível de cinza k ;

n = número total de pixels na imagem;

$p(k)$ = estimativa da probabilidade de ocorrência do nível de cinza k . A soma das probabilidades de todos os eventos elementares, isto é, $\sum_k p(k)$, será igual a 1, satisfazendo a teoria das probabilidades.

O gráfico do histograma é plotado com a intensidade dos pixels para 256 tons no eixo horizontal e, no eixo vertical, a probabilidade de ocorrência dos tons de cinza na imagem (Gonzalez & Woods, 2002). Por fornecer uma descrição global da aparência de uma imagem, o histograma é uma das características mais importantes a ser avaliada.

A Figura 23 mostra os histogramas de quatro tipos básicos de imagens em tons de cinza de um mesmo campo da pelota. Na Figura 23-a, os níveis de cinza estão concentrados em direção à extremidade escura do histograma, correspondendo a uma imagem com características predominantemente escuras. Por outro lado, a imagem da Figura 23-b apresenta um histograma contendo os

pixels concentrados na região de níveis com maiores intensidades, caracterizando uma imagem com muito brilho. O histograma da Figura 23-c tem uma forma estreita, indicando uma imagem de baixo contraste, cuja função é dada pelo desvio padrão das intensidades dos pixels. Já a Figura 23-d apresenta um histograma melhor comportado, com um espalhamento significativo, causado por condições de brilho e contraste ideais.

Embora a descrição feita pelo histograma nada diga a respeito do conteúdo da imagem, a informação fornecida é extremamente útil para o seu processamento (Gonzalez & Woods, 1992 *apud* Gomes, 2001). A visão estatística sobre a distribuição dos pixels e dos níveis de brilho e contraste oferecida pelo histograma permite que o operador possa melhorar as condições de captura, para que já sejam adquiridas imagens com a melhor qualidade possível. Caso ainda seja necessário, um histograma mais bem comportado pode ser obtido através de realces nas imagens na etapa de pré-processamento. Além disso, o histograma tem papel importante na etapa de segmentação, visto que os picos de intensidade correspondem às fases presentes, permitindo a separação e/ou quantificação de cada uma delas (Iglesias, 2008). Isso será detalhado nos capítulos a seguir.

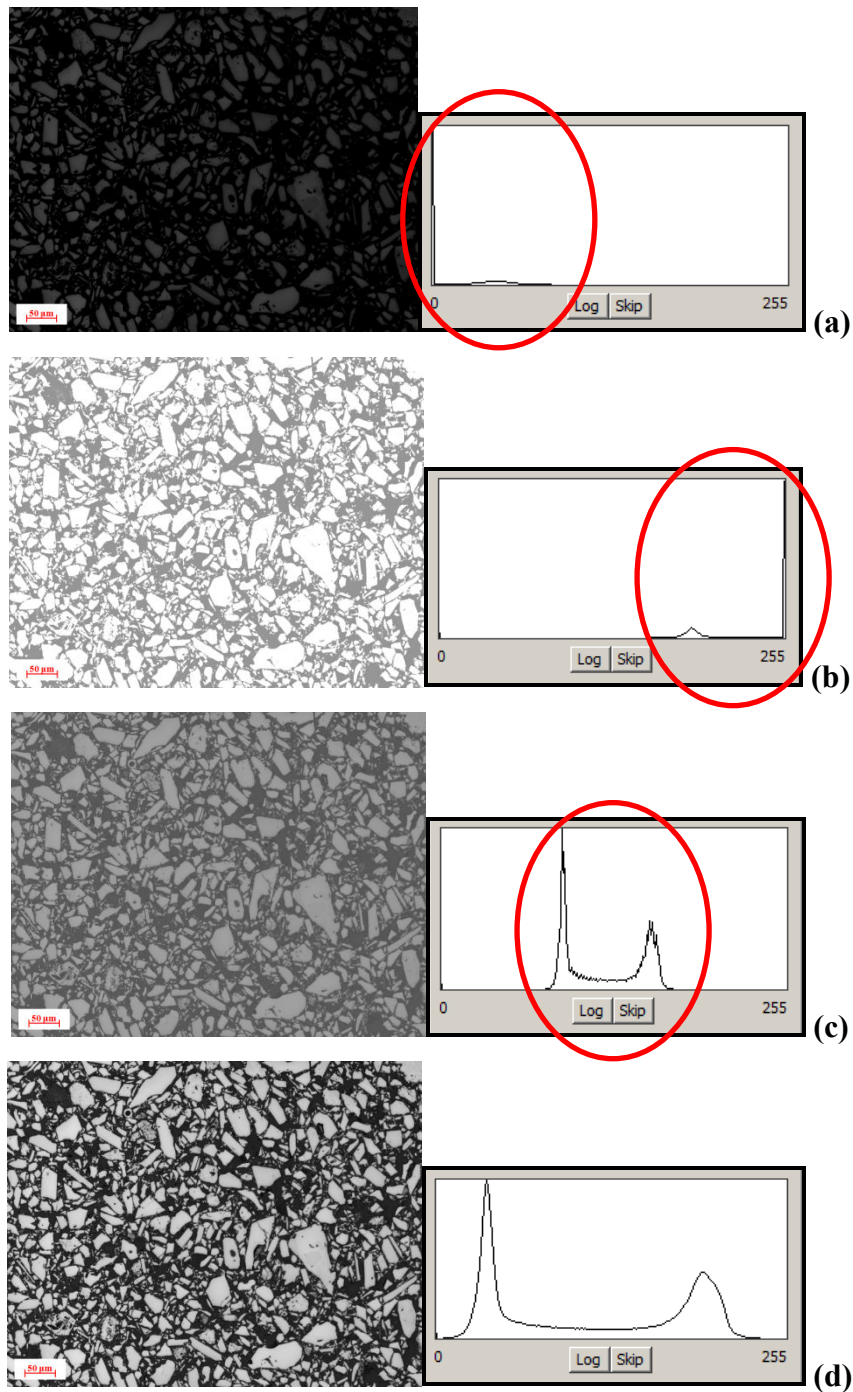


Figura 23 – Histogramas correspondentes a quatro tipos básicos de imagens: (a) Imagem escura (histograma deslocado para a esquerda); (b) imagem clara (histograma deslocado para a direita); (c) imagem de baixo contraste (histograma estreito); e (d) imagem de brilho e contraste ideais (histograma bem espalhado).

5.2. Pré-Processamento

A partir do bloco de aquisição de imagens, a maioria das funções realizadas em cada etapa do PADI pode ser implementada em software, sendo expressa em forma de algoritmo (Gonzalez & Woods, 2002).

O pré-processamento tem como objetivo melhorar a imagem, corrigindo algum defeito proveniente de sua aquisição e/ou realçando detalhes importantes para a análise (Gomes, 2001). Para que a etapa de segmentação (seção 5.3) tenha resultados satisfatórios, é necessário que a imagem esteja com o mínimo de imperfeições, daí a importância da etapa de pré-processamento.

São inúmeros os defeitos decorrentes da aquisição de imagem, assim como as possibilidades para corrigi-los. Esta etapa é caracterizada por soluções específicas, de modo que as técnicas que funcionam bem para o tipo de problema correspondente, podem se mostrar totalmente inadequadas em outro (Gonzalez & Woods, 2002). Como exemplo de procedimentos eventualmente utilizados, podemos citar: melhoras no brilho e no contraste, redução de ruídos, correção de iluminação irregular, realce de bordas, entre outros (Gomes, 2001).

Se o procedimento de captura for realizado de forma cuidadosa, em condições corretas, não se tornam necessárias muitas operações de correção nas imagens adquiridas.

Uma das operações mais comuns de pré-processamento em imagens de microscopia ótica é a correção de fundo. Esta operação pode ser realizada por dois métodos:

- *Online*, que é realizado já na etapa de aquisição de imagens e consiste na subtração entre a imagem de interesse e uma imagem branca de referência obtida de uma amostra padrão com refletância constante, plana, polida e homogênea; ou
- *Offline*, que envolve a subtração entre a imagem de interesse e a imagem correspondente ao fundo obtida pela aplicação de repetidas vezes de um filtro passa-baixa na mesma imagem. O filtro passa-baixa borra a imagem original gerando uma estimativa do fundo irregular (Paciornik & Maurício, 2004).

Daqui por diante, a abordagem será dada à outra função de pré-processamento utilizada neste trabalho, o *delineamento* (ou *delineate*), cujo objetivo principal é o realce de bordas. Em geral, o *delineamento* é bastante aplicado, pois gera um resultado mais apropriado para uma aplicação específica do que a imagem original.

Uma borda é a divisão entre duas regiões com características relativamente distintas de nível de cinza (Gonzalez & Woods, 2002). Dessa forma, as bordas serão mais definidas ou mais abruptas, quanto maior for essa diferença de intensidade entre essas regiões separadas por elas (Gomes, 2001).

No entanto, as bordas das imagens, em geral, são suaves, limitadas pela resolução do sistema de formação da imagem, e, ao contrário dessa diferença de intensidade entre regiões distintas, ocorre uma mudança de intensidade na fronteira com valores intermediários entre os valores esperados para cada fase no fenômeno denominado “efeito halo” (Iglesias, 2008).

Para reduzir o “efeito halo”, a operação de *delineamento* faz uma varredura na imagem, detecta as bordas, isto é, as transições entre fases, através da comparação entre os níveis de cinza de pixels vizinhos, e decide a qual fase esses pixels pertencem. A decisão é baseada na maior proximidade entre o pixel em questão e uma das fases, e, então, seu valor é alterado para o mesmo valor da fase ao qual foi incluído. A Figura 24 mostra exatamente como funciona o *delineamento* e a Figura 25 apresenta a diferença entre a imagem original e a imagem após a aplicação do filtro de *delineamento*. Fica visível que, na imagem da Figura 25-b, as transições entre as fases tornaram-se mais abruptas, dando a sensação de maior nitidez. Além disso, os picos nos histogramas são mais bem definidos, evidenciando a separação entre as fases.

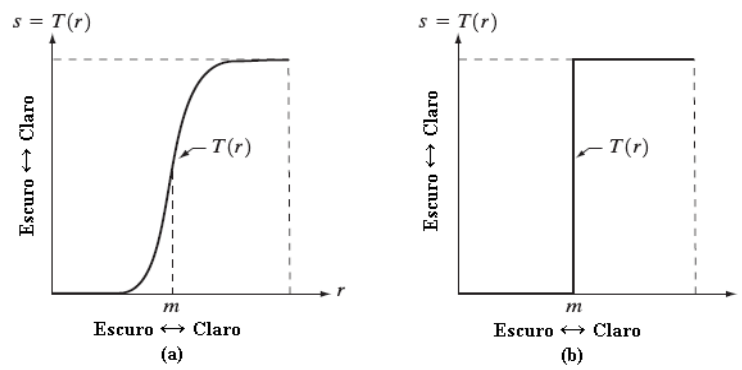


Figura 24 – Função de *delineamento* (transformação de níveis de cinza para o realce de bordas) (Gonzalez & Woods, 2002)

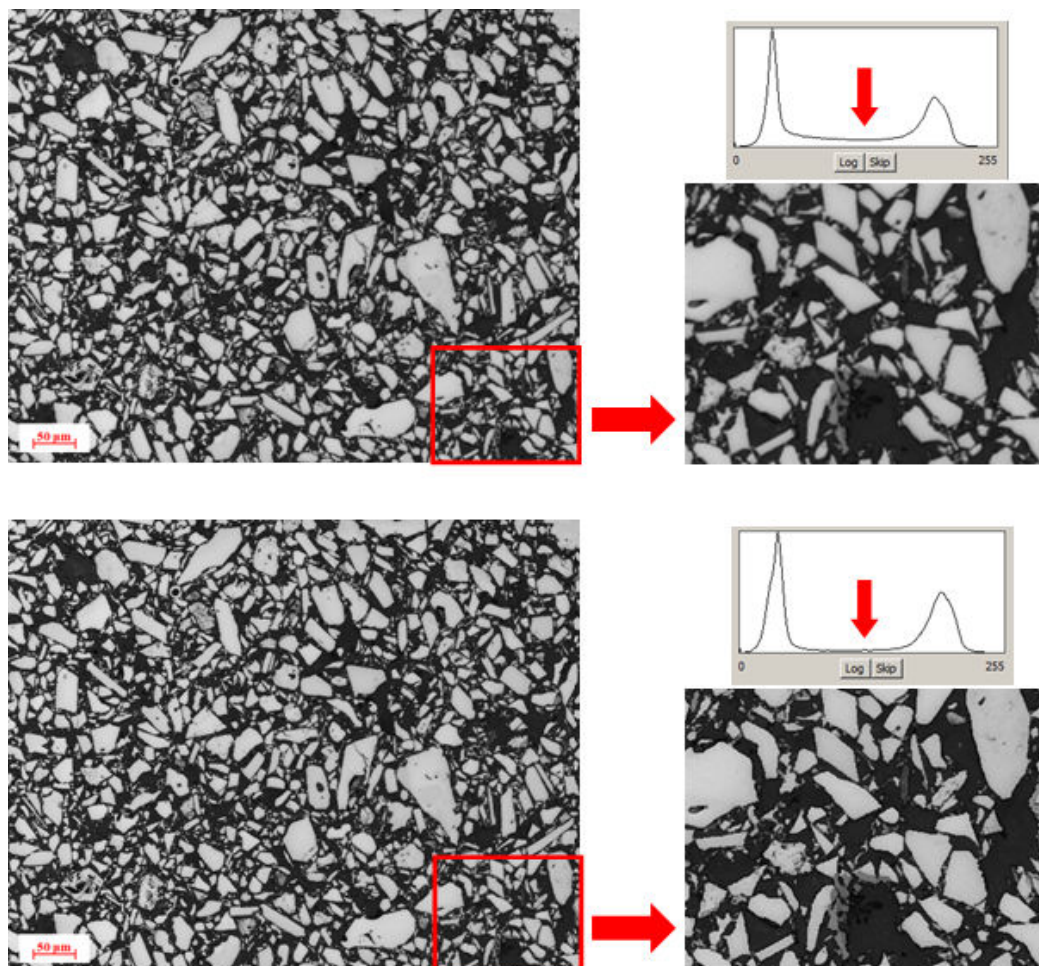


Figura 25 – Delineamento: (a) Imagem original, seu histograma e região ampliada para melhor visualização; e (b) imagem delineada, seu histograma e a mesma região ampliada.

5.3. Segmentação

A segmentação é a etapa crítica do fluxograma de PADI, onde se obtém a imagem a partir da qual alguma informação será extraída. Seu objetivo principal é subdividir uma imagem em regiões ou objetos de interesse, formados pelo agrupamento de pixels contíguos (Paciornik, 2010).

A imagem de saída da segmentação é uma imagem binária, onde os pixels pretos representam o fundo ou objetos que não são de interesse na imagem, e os pixels brancos constituem os objetos de interesse, os quais serão quantificados, ou vice-versa (Iglesias, 2008; Gomes, 2001). Dessa forma, neste trabalho fica designado o valor 0 (pixels pretos) para regiões do fundo e valor 1 (pixels brancos) para os objetos de medida.

Os algoritmos de segmentação são inúmeros e são baseados, em geral, em 2 propriedades básicas de valores de níveis de cinza: descontinuidade e similaridade. A descontinuidade tem como abordagem a detecção de pontos isolados e detecção de linhas e bordas na imagem, ou seja, particiona a imagem conforme mudanças bruscas nos níveis de cinza. Já a similaridade baseia-se nos métodos de limiarização, crescimento de regiões e divisão e fusão de regiões (Gonzalez & Woods, 2002). Como a *limiarização* é o método utilizado no presente trabalho, e, além disso, é uma das mais importantes abordagens na distinção de objetos um dos outros e do fundo, a discussão acerca da etapa de segmentação de imagens será focada nesta técnica.

O fundamento da *limiarização* está na escolha de um ou mais limiares, que permitem separar os diferentes grupos de pixels existentes na imagem. Dado um histograma de uma imagem composta por objetos iluminados sobre um fundo escuro, os pixels do objeto e os do fundo podem ser agrupados de acordo com os níveis de cinza em dois grupos dominantes correspondentes. A divisão da imagem será feita através da seleção de um limiar T , separando-a em fundo (pixels pretos) e objetos (pixels brancos), conforme mostrado na Figura 26-a.

No caso da separação de três ou mais grupos dominantes, por exemplo, dois tipos de objetos iluminados sobre um fundo escuro, a *limiarização* é denominada multimodal. Esse tipo de limiarização é feita com o estabelecimento de múltiplos limiares que isolem efetivamente as regiões de interesse. Uma ilustração de limiarização multimodal pode ser observada na Figura 26-b.

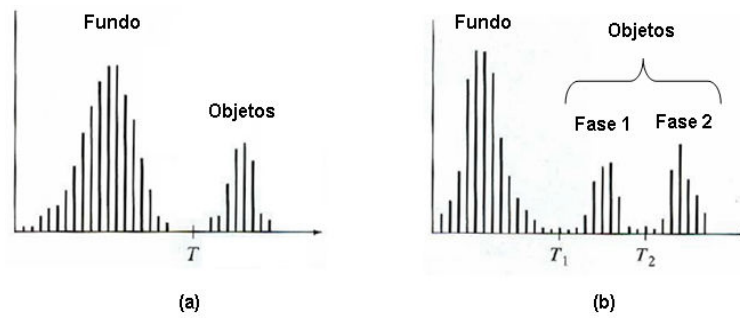


Figura 26 – Histograma de níveis de cinza que podem ser particionados por (a) um limiar único e (b) limiares múltiplos

A Figura 27-b apresenta um exemplo típico de imagem segmentada de pelota de minério de ferro, obtida por microscopia ótica. A imagem original possui 256 tons de cinza, na qual foi realizada uma segmentação bimodal. A imagem segmentada é uma imagem binária, apresentando em pixels brancos todos os pixels da imagem original que continham valor acima do limiar estabelecido. Os demais pixels da imagem original receberam valor zero, aparecendo em preto na imagem de saída da segmentação. O histograma e o tom de corte para a separação das fases na imagem também estão indicados na Figura 27.

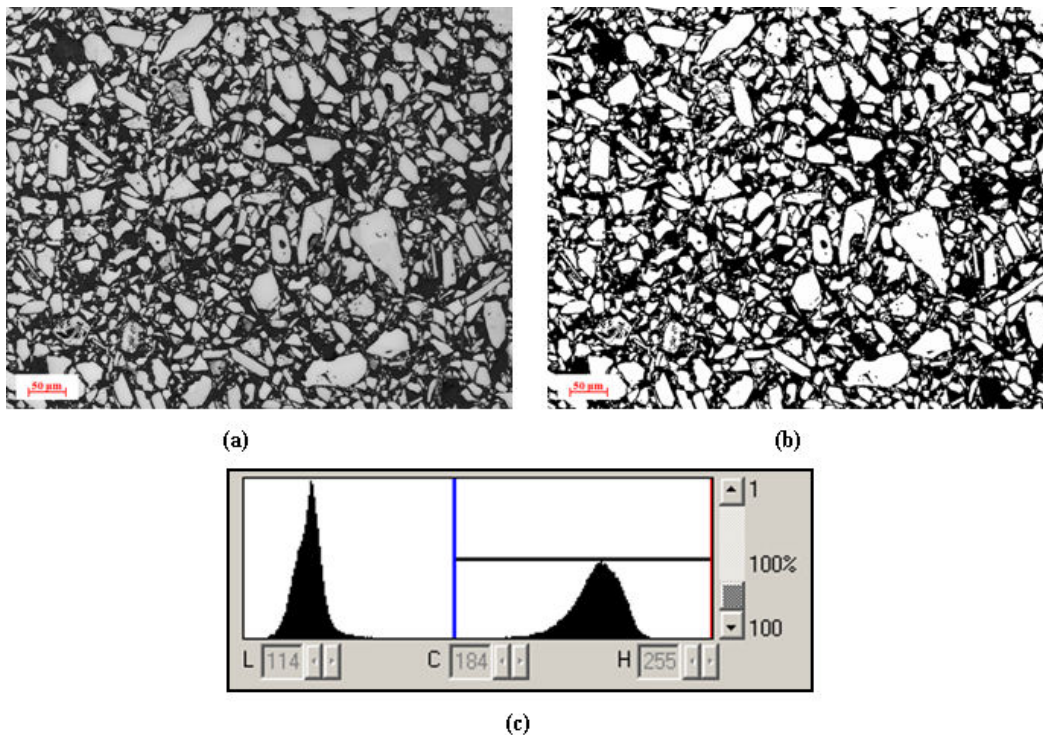


Figura 27 – Limiarização bimodal na imagem de pelota de minério de ferro: (a) imagem original em 256 tons de cinza; (b) imagem binária; (c) tons de corte.

Por ser uma das tarefas mais difíceis do PADI, podendo determinar o sucesso ou o fracasso das análises posteriores, a segmentação precisa ser realizada sob cuidado considerável. Dificilmente a escolha manual é acurada ou reprodutível, fazendo-se necessário uma seleção automática do tom de corte para uma maior robustez no processo. Na segmentação automática, a escolha do limiar é feita baseada na análise dos histogramas das imagens. Um dos métodos de segmentação mais comuns é o que utiliza os mínimos do histograma como decisão de tom de corte entre as fases, ou seja, os limiares correspondem às tonalidades intermediárias entre duas bandas (ou dois picos) (Paciornik, 2010).

Todo o processo aqui discutido em imagens monocromáticas pode ser aplicado também em imagens coloridas. Neste caso, existem três histogramas correspondentes a cada componente R, G e B da imagem colorida. A escolha dos limiares deve ser feita nos três histogramas, de forma que a faixa de cores selecionada faça parte da região segmentada (Iglesias, 2008).

5.4. Pós-Processamento

Muitas vezes o resultado da segmentação não é adequado para que os grupos de pixels segmentados sejam representados e descritos em termo de suas características nas etapas subsequentes. Sendo assim, para a correção de defeitos residuais nas imagens, torna-se necessária a etapa de pós-processamento.

A separação de objetos que se tocam, a eliminação de objetos de que não se deseja extrair nenhuma informação e o agrupamento de objetos para a formação de objetos mais complexos são exemplos de procedimentos realizados na etapa de pós-processamento. Estes procedimentos são realizados através de operações lógicas e morfológicas (Gomes, 2001; Iglesias, 2008).

As operações lógicas são operações pontuais, isto é, são desempenhadas pixel a pixel, gerando uma imagem de saída contendo pixels preservados ou invertidos em relação aos da imagem de entrada. As principais operações lógicas em processamento de imagens são a interseção (AND), o complemento (NOT) e a união (OR), a partir dos quais combinações podem ser feitas para formar qualquer outra operação lógica (Gonzalez & Woods, 2002; Gomes, 2001). A Figura 28 mostra as principais operações lógicas e os resultados correspondentes, dependendo da operação e da imagem de entrada, de cada uma delas.

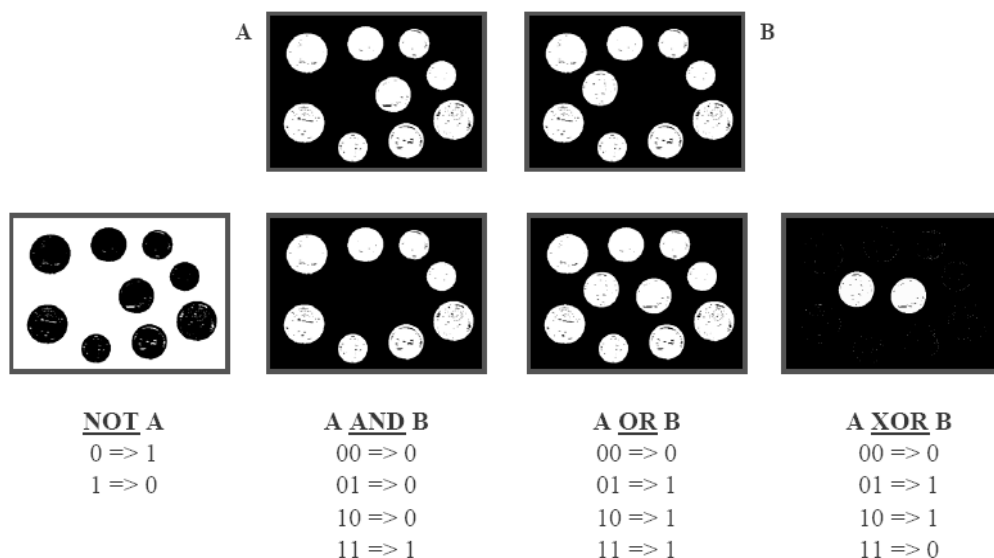


Figura 28 – Operações Lógicas (Paciornik, 2010)

Diferentemente das operações lógicas, as quais são realizadas pontualmente pixel a pixel, as operações morfológicas são operações realizadas localmente, onde um pixel da imagem de saída é função do valor dos pixels numa vizinhança da imagem de entrada (Paciornik, 2010). Em imagens binárias, um pixel será preservado ou invertido em função de ter certo número de vizinhos iguais ou diferentes em posições específicas de sua vizinhança e as posições desses pixels são definidas por uma espécie de máscara, denominada *elemento estruturante*. O *elemento estruturante* consiste numa vizinhança que “varre” a imagem de entrada, preservando ou invertendo o pixel central da vizinhança de acordo com a função aplicada.

A Figura 29 mostra uma das operações de pós-processamento mais realizadas neste trabalho: o NOT. A imagem original é apresentada na Figura 29-a, a imagem da fase clara segmentada na Figura 29-b, e, por fim, o negativo da imagem binária na Figura 29-c.

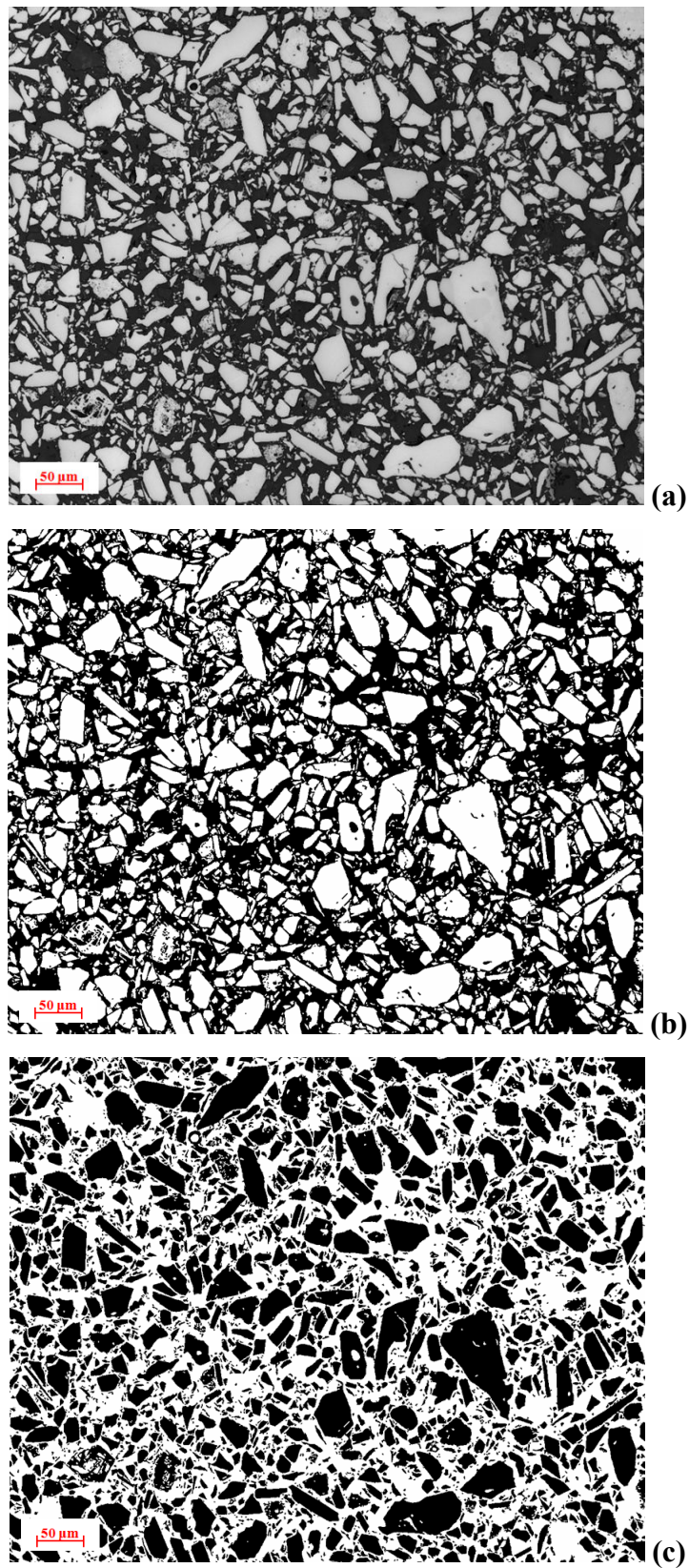


Figura 29 – Exemplo de pós-processamento: (a) Imagem original; (b) imagem binária da segmentação da fase clara; (c) imagem resultante da aplicação de um NOT.

As discussões acerca das operações morfológicas estão baseadas em duas técnicas fundamentais: a erosão e a dilatação. A erosão, conforme o nome sugere, consiste em fazer o objeto encolher. Para isso, cada pixel branco da imagem de entrada é invertido na imagem de saída, caso não exista um pixel branco na vizinhança “coberta” pelo elemento estruturante. Opostamente à erosão, a dilatação cresce objetos em área com o procedimento contrário, isto é, cada pixel preto da imagem de entrada é invertido na imagem de saída, se houver pelo menos um vizinho branco na região “coberta” pelo elemento estruturante (Gomes, 2001).

A partir da combinação destas duas operações morfológicas, outras operações podem ser geradas, tais como: o OPEN (ou *abertura*), que é uma sequência de n ciclos de erosão seguidos de n ciclos de dilatação, e o CLOSE (ou *fechamento*), cuja sequência é o inverso da sequência do OPEN.

Outro tipo de operação morfológica é o Mapa de Distâncias Euclidianas (MDE). O MDE tem como resultado uma imagem na qual a intensidade dos pixels da imagem de saída é proporcional à distância do pixel até a borda do objeto da imagem de entrada. Uma aplicação do MDE bastante utilizada neste trabalho foi na separação de objetos (partículas) que se tocam nas imagens analisadas. O método de separação, denominado *divisores de águas* (ou *watersheds*), é feito pela obtenção do MDE da imagem binária, pela limiarização da imagem obtida no MDE, gerando-se as sementes dos objetos (de apenas 1 pixel correspondente a cada objeto), e por uma *dilatação derradeira* desta última, em que pixels vão sendo agregados às sementes fazendo crescer ilhas de pixels brancos sem que quaisquer duas ilhas voltem a se unir (Gomes, 2001; Paciornik, 2010).

As partículas, que antes estavam conectadas, foram separadas pelo *divisor de águas* no exemplo da Figura 30.

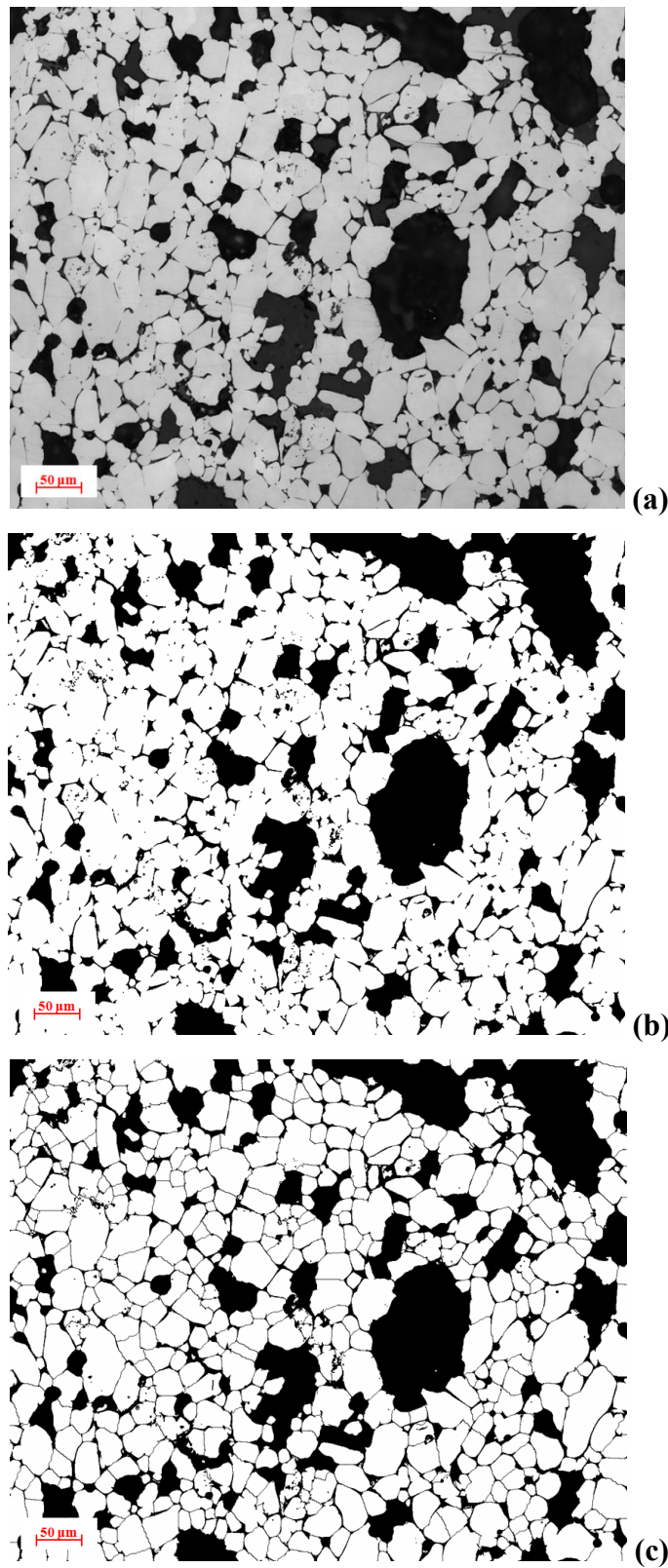


Figura 30 – Método dos divisores de águas (*watersheds*): (a) imagem binária inicial; e (b) imagem binária com objetos separados.

Além de todas estas operações, existe outra bastante utilizada no processamento de imagens, e que não é propriamente derivada das operações aqui citadas, chamada *Scrap*. Esta operação tem como finalidade eliminar objetos que tenham área, isto é, número de pixels agrupados, dentro de uma faixa de valores determinada. A Figura 31 mostra um exemplo de aplicação do *Scrap*, onde os objetos com área menor do que 1000 pixels foram eliminados da imagem segmentada.

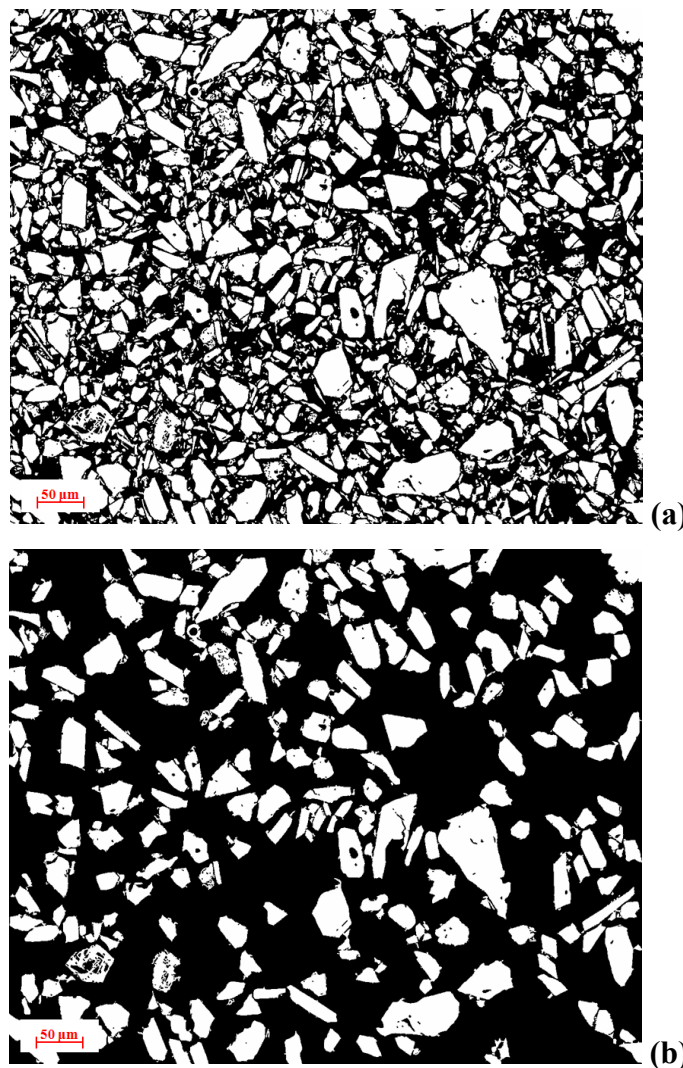


Figura 31 – Eliminação de objetos por área: (a) Imagem segmentada; (b) imagem sem objetos menores do que 1000 pixels.

5.5. Extração de Atributos

A Extração de Atributos é onde se inicia a etapa propriamente dita de análise da imagem. Nesta etapa são realizadas medidas na imagem segmentada ou pós-processada, ou até mesmo na imagem em tons de cinza. Através dessas medidas, os grupos de pixels são descritos por atributos característicos, gerando dados quantitativos para o objetivo final.

Em geral, a Extração de Atributos pode ser dividida em dois tipos de medida: medidas de campo e medidas de região. As medidas de campo se referem ao campo como um todo, como na medição de número de objetos, área total dos objetos e fração de área, gerando como resultado apenas um valor por medida. As medidas de região se referem aos objetos individualmente, ou seja, é extraído um parâmetro de cada objeto na imagem, como por exemplo, tamanho, forma e posição das partículas (Gomes, 2001; Paciornik, 2010).

Excepcionalmente, nesta dissertação, a notação será diferente, visto que os resultados estão sempre relacionados ao campo como um todo. No caso de parâmetros de região, as medidas extraídas em cada imagem são sempre a média de todas as partículas, isto é, a soma de todas as medidas dividida pelo número de partículas de uma imagem (seção 5.5.1.1). Sendo assim, os tipos de medida foram divididos em:

- Medidas de Campo
- Medidas de Tamanho
- Medidas de Textura
- Medidas de Forma

Os atributos cuja extração é de interesse deste trabalho foram separados nestes quatro grupos e cada um deles será explicado separadamente nas seções a seguir. Vale lembrar que a divisão convencional destes atributos é feita somente em dois grupos: parâmetros de campo e parâmetros de região. Em parâmetros de região estão incluídos outros grupos, tais como tamanho e forma. Maiores detalhes acerca da organização tradicional dos diferentes tipos de atributos podem ser encontrados em Gomes (2001).

5.5.1. Medidas de Campo

5.5.1.1. Contagem de Partículas

A contagem de partículas é uma das medidas mais fáceis de obter digitalmente. A partir da imagem segmentada, é feita uma contagem de pixels contíguos, cada qual designará um objeto (neste caso, uma partícula).

A contiguidade entre os pixels, também denominada conectividade, é determinada, em geral, por duas opções: Conectividade 4 e Conectividade 8. A Conectividade 4 estabelece que somente os 4 pixels adjacentes vertical e lateralmente são considerados contíguos, sem levar em conta os pixels adjacentes da diagonal. Isto significa que se um pixel estiver conectado a outro pixel adjacente à sua diagonal, eles serão considerados pertencentes a dois objetos separados. A Conectividade 8 determina que todos os 8 pixels adjacentes são considerados contíguos, fazendo parte do mesmo objeto (Gomes, 2001).

A Figura 32-a apresenta um exemplo de uma pequena imagem de 10x10 pixels, onde o pixel circulado em vermelho é o pixel analisado. A Figura 32-b e a Figura 32-c são exemplos das diferentes considerações sobre a conectividade, que, conseqüentemente, levam à contagem diferente de número de objetos. Os pixels em vermelho nas imagens representam os pixels adjacentes considerados contíguos ao pixel em questão. Para a Conectividade 4 (Figura 32-a), o pixel está conectado diagonalmente a outro pixel, fora da região de contiguidade, significando que os pixels fazem parte de dois objetos diferentes. Para a Conectividade 8 (Figura 32-c), o pixel em questão é vizinho do outro pixel dentro da região de contiguidade, indicando que compõem um mesmo objeto.

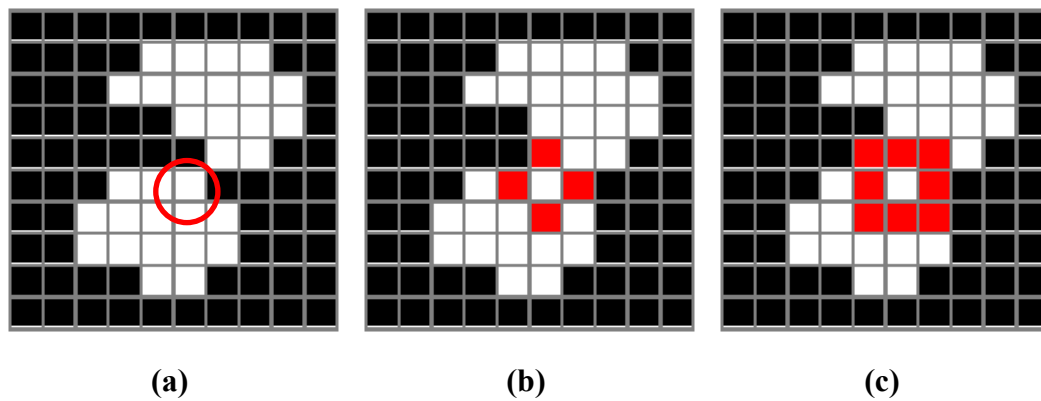


Figura 32 – Conectividade entre os pixels: (a) Imagem original, onde o pixel circulado é o pixel analisado; (b) Conectividade 4, considerando que são dois objetos separados; (c) Conectividade 8, considerando que os pixels fazem parte de um mesmo objeto.

5.5.1.2. Fração de Área

A fração de área também é uma medida fácil e rápida, e é obtida simplesmente pela razão entre números de pixels brancos e o número total de pixels na imagem binária.

A Figura 33 mostra como é medida a fração de área de uma imagem binária de 10x10 pixels.

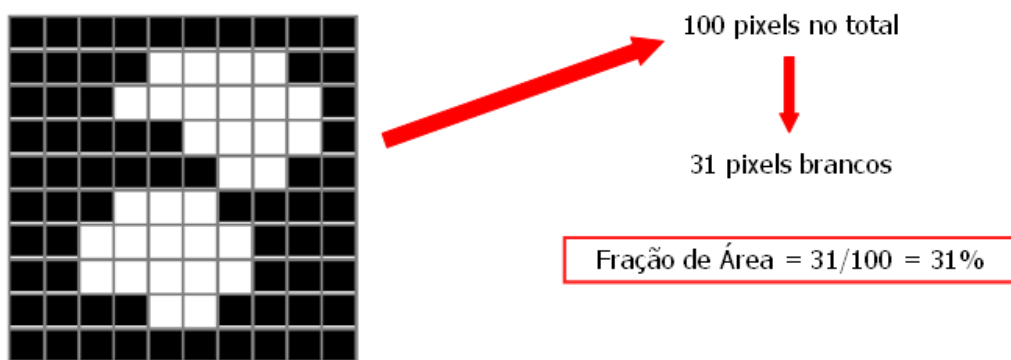


Figura 33 – Cálculo da Fração de Área de uma imagem binária

5.5.2. Medidas de Tamanho

5.5.2.1. Área

A área ocupada de um objeto numa imagem binária corresponde ao número de pixels brancos contidos dentro de sua fronteira. De forma simples, a medida é feita pela contagem de número de pixels.

Dentro do mesmo conceito, existem outras medidas relacionadas à área, tais como área preenchida e área convexa. A área preenchida é a medida de área incluindo buracos internos. A área convexa é a área equivalente a um elástico passado em torno do objeto.

A Figura 34 exemplifica os tipos de medida de área.

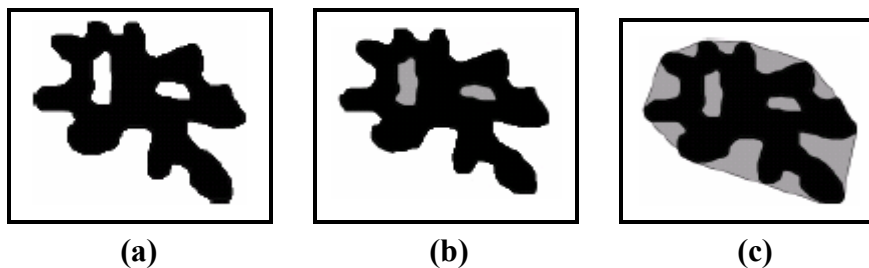


Figura 34 – Medidas de Área: (a) área simples (toda a região em preto); (b) área preenchida (toda a região em preto e em cinza); (c) área convexa (toda a região em preto e em cinza) (Disponível em: <www.dema.puc-rio.br/cursos/ipdi>. Acesso em: 20 jan. 2012)

5.5.2.2. Perímetro

O perímetro é outra medida geométrica básica que indica o tamanho da fronteira dos objetos.

O perímetro pode ser obtido a partir do contorno do objeto, dado pela seguinte expressão:

$$P = N + \sqrt{N_d} \quad (2)$$

onde N é o número de passos horizontais ou verticais de cada pixel para o seguinte e N_d é o número de passos diagonais de cada pixel para o seguinte (Gomes, 2001).

Analogamente às medidas de área, os tipos de perímetro estão mostrados na Figura 35. Note que, na medida do perímetro (Figura 35-a), os perímetros dos “buracos” internos do objeto também são considerados. E os perímetros preenchido (Figura 35-b) e convexo (Figura 35-c) são, respectivamente, os equivalentes aos da área preenchida e convexa.

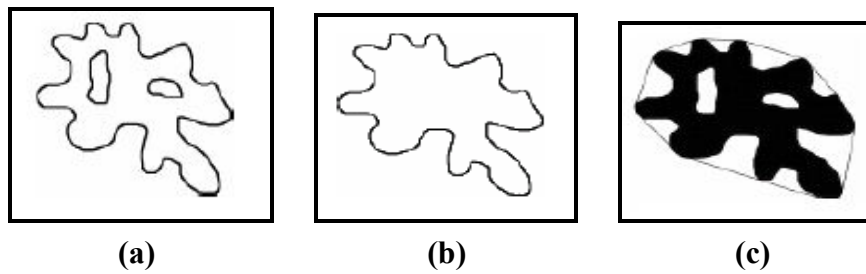


Figura 35 – Medidas de Perímetro: (a) perímetro simples; (b) perímetro preenchido; (c) perímetro convexo (Disponível em: <www.dema.puc-rio.br/cursos/ipdi>. Acesso em: 20 jan. 2012)

5.5.2.3. Diâmetro Circular Equivalente

O diâmetro circular equivalente (*DCE*) também é uma medida de tamanho do objeto e é definido a partir de sua área pela seguinte equação:

$$DCE = \sqrt{\frac{4 \times \text{Área}}{\pi}} \quad (3)$$

Outra forma de se obter o diâmetro circular equivalente é a partir do perímetro do objeto, conforme abaixo:

$$DCE = \frac{\text{Perímetro}}{\pi} \quad (4)$$

5.5.2.4. Interceptos e Interseções

As medidas de interceptos e interseções, em geral, são consideradas medidas de campo, visto que fornecem um único resultado referente ao campo e não aos objetos separadamente. No entanto, como já dito anteriormente, neste trabalho todos os atributos extraídos são somados e divididos pelo número de partículas,

tornando-se sempre, parâmetros de campo. Sendo assim, as medidas de interceptos e interseções foram colocadas dentro do grupo de medidas de tamanho, pois, como será explicado em seguida, essas medidas referem-se à cálculos de distância, seja entre um objeto e outro ou do próprio objeto, baseados em suas fronteiras.

Para facilitar o entendimento desta seção, a diferença entre os conceitos de intercepto e interseção pode ser claramente visualizados na Figura 36, onde várias linhas horizontais de teste foram desenhadas nas imagens. As interseções consistem nos pixels comuns entre a linha e os contornos dos objetos (Figura 36-a) e os interceptos são a sobreposição entre esta linha e os objetos (Figura 36-b). Em outras palavras, as interseções são os pixels extremos de cada intercepto. Dessa forma, para cada intercepto, há duas interseções, exceto no caso da linha não se estender por toda a região (Friel, 2000; Gomes, 2001).

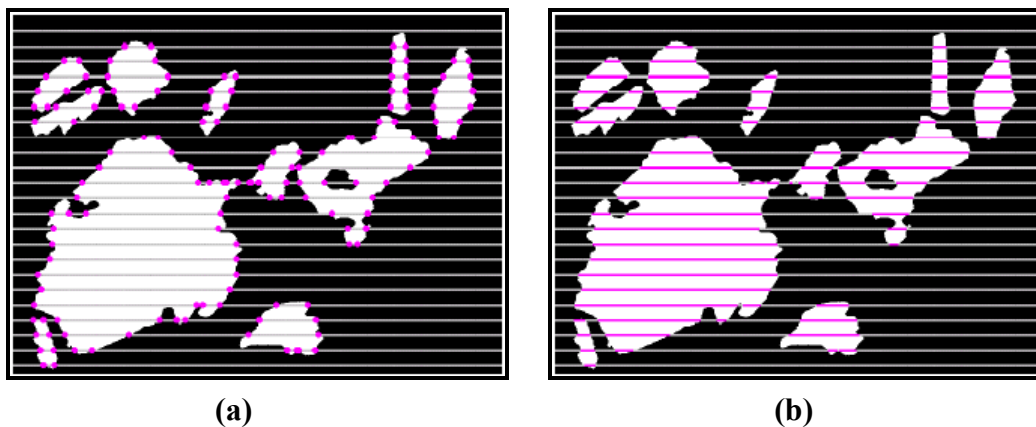


Figura 36 – (a) Interseções; (b) Interceptos (Gomes, 2001)

A partir desses conceitos, muitas medidas podem ser feitas, tais como o comprimento de cada intercepto, o comprimento total obtido pela soma de cada um individualmente e o número de objetos interceptados, calculados pela quantidade de interseções. Dessas medidas, derivam-se a *densidade linear* e a *fração de comprimento*, conforme as Equações 5 e 6, respectivamente.

$$N_L = \frac{n_i}{L} \quad (5)$$

Onde n_i é o número de interceptos e L é o comprimento total das linhas de teste.

$$L_L = \frac{L_i}{L} \quad (6)$$

Onde L_i é o comprimento total de interceptos.

Além dessas medidas, também pode ser calculado, pela eq. (7), o *intercepto linear médio*.

$$\bar{L} = \frac{L_L}{N_L} \quad (7)$$

5.5.3. Medidas de Textura

Outra abordagem para a descrição de imagens é a quantificação de seu conteúdo de textura. Neste caso, as medições são feitas em imagens em tons de cinza, podendo utilizar ou não uma imagem segmentada como máscara, a fim de definir quais regiões devem ter as texturas quantificadas.

Embora não exista uma definição formal, a textura em PADI pode ser entendida como uma medida de propriedades como suavidade, rugosidade, regularidade e granularidade (Russ, 1995; Gonzalez & Woods, 2002 *apud* Gomes, 2001).

As três formas principais utilizadas em PADI para a medição de texturas são a estatística, a estrutural e a espectral. As técnicas estatísticas estão relacionadas a caracterizações de textura como suave, áspera, granular etc. As técnicas estruturais são mais utilizadas para caracterizar imagens com textura muito regular, pois tratam da organização de primitivas de imagens (textels), como no caso de imagens contendo linhas paralelas igualmente espaçadas. As abordagens espectrais são baseadas em propriedades de alguma transformada, como por exemplo, na transformada de Fourier, onde se detecta a periodicidade global em uma imagem, já que as posições desses padrões periódicos são identificadas através de picos de alta-energia no espectro (Gomes, 2001; Gonzalez & Woods, 2002). Como a técnica utilizada para caracterização de textura no presente trabalho é a estatística, daqui por diante será dada ênfase a este tipo de medida.

As técnicas estatísticas são uma das mais antigas e mais simples técnicas para a descrição de texturas e, segundo Gomes (2001), os parâmetros estatísticos

extraídos estão divididos em dois grupos: os de primeira ordem e os de segunda ordem. Os parâmetros de primeira ordem são calculados de forma simples através dos histogramas de níveis de cinza de uma imagem ou de uma região, os quais podem ser classificados como medidas de intensidade (seção 5.5.3.1). Os parâmetros estatísticos de segunda ordem são obtidos a partir da construção da matriz de co-ocorrência de tons de cinza da imagem, considerando o posicionamento relativo entre os pixels. Exemplos deste tipo são os chamados Parâmetros de Haralick (seção 5.5.3.2), onde cada parâmetro fornece uma característica de textura, tais como contraste, homogeneidade e complexidade.

5.5.3.1. Intensidade

Por serem medidas estatísticas descritivas da aparência de uma imagem, tais como brilho e contraste, as medidas de intensidade podem ser consideradas como parâmetros simples de caracterização de textura. Assim como qualquer medida de textura, os parâmetros são calculados sobre a imagem em tons de cinza.

As propriedades de brilho e contraste são definidas, respectivamente, por:

$$B = \sum_{r=0}^{255} r \cdot p(r) \quad (8)$$

$$C = \sqrt{\sum_{r=0}^{255} [r - B]^2 \cdot p(r)} \quad (9)$$

Onde r representa os tons de cinza; e $p(r)$ é o histograma.

Como as medidas de intensidade são definidas através dos momentos no histograma, dois momentos, mais comumente utilizados, podem ser calculados:

- Média de tons de cinza, que está associada ao brilho;
- Desvio padrão da intensidade, que é calculado pelo quadrado do contraste.

Além desses, existem outros tipos de momentos, de ordens mais altas, relacionados à anti-simetria ou ao achatamento e planaridade do histograma, que também fornecem informação quantitativa sobre a textura da imagem (Gomes,

2001), porém não foram utilizados neste trabalho e, por isso, não serão discutidos aqui.

5.5.3.2. Parâmetros de Haralick

O uso de estatísticas simples nem sempre permite discriminar variações mais sutis dos tons de cinza. Além disso, não fornecem informação sobre a posição relativa dos pixels em relação uns aos outros. Assim, para trazer essa informação ao processo de análise de texturas, foram propostas as estatísticas baseadas na matriz de co-ocorrência. A matriz de co-ocorrência, basicamente, verifica quantas combinações diferentes de valores de intensidade de pixels ocorrem em uma imagem. Os parâmetros de Haralick são medidas baseadas nessa matriz de co-ocorrência de tons de cinza da imagem, cuja construção será explicada a seguir.

Dada uma imagem com resolução de N_x pixels na direção horizontal e N_y pixels na direção vertical, contendo N_g tons de cinza, pode-se representá-la como uma matriz I de dimensão $N_x \times N_y$, onde os elementos são os valores de tom de cinza dos pixels na posição correspondente. Dessa forma, sua matriz de co-ocorrência é uma matriz quadrada, de dimensão $N_g \times N_g$, onde os elementos são designados pela probabilidade de ocorrência de pares de pixels, com os tons de cinza i e j , separados por uma certa distância numa dada direção e sentido determinados por um vetor d . O vetor d é definido por suas componentes dx e dy , como mostrado na Figura 37 (Gomes, 2001; Gonzalez & Woods, 2002).

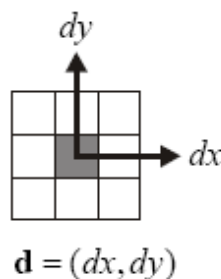


Figura 37 – Vetor d (Gomes, 2001)

Assim, os elementos da matriz de co-ocorrência ($p_d(i, j)$) podem ser definidos por:

$$p_d(i, j) = \frac{P_d(i, j)}{R} \quad (10)$$

Onde $P_d(i, j)$ é o número de vezes que um par de pontos ocorre e R é o número total de pares de pontos, ambos satisfazendo as direções do vetor d .

Para o cálculo dos parâmetros de Haralick, a direção do vetor d é dada pelo ângulo α , que só pode adquirir quatro valores possíveis (0° , 45° , 90° e 135°) correspondentes aos pixels vizinhos mais próximos, isto é, os pixels adjacentes. Para que sejam medidas independentes de rotação, em geral, é utilizada a média desses quatro parâmetros (Gomes, 2001). Porém, a escolha do tipo de medida deve ser feita dependendo do caso específico.

A Tabela 2 lista os 11 parâmetros de Haralick utilizados neste trabalho, em que cada parâmetro fornece uma indicação diferente relativa à resposta da matriz de co-ocorrência. O entendimento do significado físico de alguns parâmetros de Haralick não é simples, mas a idéia básica está em caracterizar o “conteúdo” da matriz de co-ocorrência através de diferentes descritores, sendo alguns deles correlatados entre si.

Tabela 2 – Parâmetros de Haralick

Parâmetro	Definição
1) Segundo Momento Angular	$f_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [p(i, j)]^2$
2) Contraste	$f_2 = \sum_{n=0}^{N_g-1} n^2 \sum_{\substack{i=1 \\ i-j =n}}^{N_g} \sum_{j=1}^{N_g} p(i, j)$
3) Correlação	$f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i \cdot j \cdot p(i, j) - \mu_x \cdot \mu_y}{\sigma_x \cdot \sigma_y}$
4) Variância	$f_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 \cdot p(i, j)$
5) Momento Inverso da Diferença	$f_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)^2} \cdot p(i, j)$
6) Média da Soma	$f_6 = \sum_{i=2}^{2 \cdot N_g} i \cdot p_{x+y}(i)$
7) Variância da Soma	$f_7 = \sum_{i=2}^{2 \cdot N_g} (i - f_6)^2 \cdot p_{x+y}(i)$
8) Entropia da Soma	$f_8 = - \sum_{i=2}^{2 \cdot N_g} p_{x+y}(i) \cdot \log[p_{x+y}(i)]$
9) Entropia	$f_9 = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \cdot \log[p(i, j)]$
10) Variância da Diferença	$f_{10} = \text{variância de } p_{x-y}$
11) Entropia da Diferença	$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \cdot \log[p_{x-y}(i)]$

Notação:

$$p_x(i) = \sum_{j=1}^{N_g} p(i, j)$$

$$p_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{N_g} \sum_{j=1}^{N_g} p(i, j), \quad k = 2, 3, \dots, 2N_g$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j)$$

$$p_{x-y}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{N_g} \sum_{j=1}^{N_g} p(i, j), \quad k = 0, 1, \dots, N_g - 1$$

μ_x e μ_y são as médias

σ_x e σ_y são, respectivamente, os desvios padrões de p_x e p_y

Fonte: Haralick, 1973

5.5.4. Medidas de Forma

Como não existe uma forma genérica para descrever um objeto, para a caracterização de sua morfologia, em geral, quantifica-se o afastamento da forma deste objeto em relação ao de um modelo teórico. Assim, as medidas de forma em PADI são, normalmente, parâmetros adimensionais obtidos pela combinação de parâmetros de tamanho.

Por serem obtidos desta maneira, geralmente variam entre 0 e 1, sendo 1 para formas com geometrias regulares, ou seja, formas padrão, e 0 para formas irregulares (Gomes, 2001; Paciornik, 2010).

A seguir serão explicados os principais parâmetros de forma utilizados neste trabalho.

5.5.4.1. Razão de Aspectos

A *Razão de Aspectos (RA)* é um parâmetro utilizado para medir o alongamento do objeto. Para entender seu conceito, será necessária a apresentação de dois parâmetros de tamanho, denominados *aspectos*, que não foram discutidos anteriormente, pois, neste trabalho, não serviram individualmente como medidas de tamanho.

Os *aspectos*, também chamados *calibres*, *Feret diameters*, ou simplesmente *ferets*, equivalem às sombras dos objetos em diferentes direções. Os principais *ferets* são o *feret* mínimo e o *feret* máximo, mostrados na Figura 38.

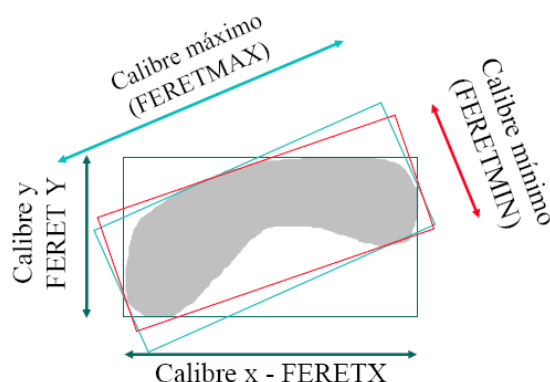


Figura 38 – Calibres (*Ferets*) (Disponível em: <www.dema.puc-rio.br/cursos/ipdi>.

Acesso em: 20 jan. 2012)

A *razão de aspectos*, então, é obtida pela razão entre estes dois *ferets*, de acordo com a seguinte equação:

$$RA = \frac{Fmin}{Fmax} \quad (11)$$

5.5.4.2. Medidas de Convexidade

As medidas de convexidade servem para descrever se um objeto é côncavo ou convexo. Para isso, existem dois parâmetros de convexidade, aqui denominados $Conv_1$ e $Conv_2$, que são definidos, respectivamente, como a razão entre a área (A) e a área convexa (A_c) e como a razão entre o perímetro convexo (P_c) e o perímetro (P), assim:

$$Conv_1 = \frac{A}{A_c} \quad (12)$$

$$Conv_2 = \frac{P_c}{P} \quad (13)$$

Ambos valem 1 para objetos convexos e diminuem para objetos não-convexos. O parâmetro $Conv_2$, por depender do perímetro, é mais sensível ao contorno do objeto, decaindo rapidamente para objetos com ramificações finas e longas. Já o parâmetro $Conv_1$, é menos sensível a este tipo de irregularidade.

5.5.4.3. Fatores de Forma Circular

Existem dois tipos de fatores de forma circular (FFC) e eles podem ser definidos, a partir da área, perímetro ou *feret* máximo, segundo as Equações 14 e 15.

$$FFC_1 = \frac{4.\pi.A}{P^2} \quad (14)$$

$$FFC_2 = \frac{4.A}{\pi.Fmax^2} \quad (15)$$

Assim, os valores de FFC valem 1 para objetos circulares e diminuem para objetos com outras formas. Como FFC_1 é dependente do perímetro, é um parâmetro sensível à irregularidade do contorno do objeto. Já FFC_2 é mais sensível ao alongamento do objeto, que pode ser explicado pelo fato de seu cálculo ser feito a partir do *feret* máximo.

5.5.4.4. Fatores de Ramificação

Os fatores de ramificação podem ser definidos como uma razão de aspectos modificada, onde se substitui um dos parâmetros da razão por outro fator de tamanho, que é o diâmetro do maior círculo circunscrito no objeto (W).

O diâmetro W (Figura 39) é calculado a partir do Mapa de Distâncias Euclidianas do objeto. Como no Mapa de Distâncias Euclidianas a intensidade dos pixels é proporcional à sua distância até a borda do objeto (seção 5.4), o valor do pixel mais brilhante corresponde ao raio do maior círculo circunscrito. Sendo assim, o dobro do pixel de maior intensidade é o diâmetro do maior círculo circunscrito.

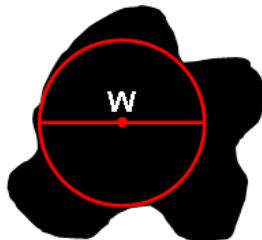


Figura 39 – Diâmetro do maior círculo circunscrito no objeto (W)

Os fatores de ramificação podem ser de dois tipos, aqui denominados de *Modification Ratio* (MR) e *Branching Factor* (BF).

O MR é um parâmetro de medida sensível ao alongamento e às ramificações do objeto, definido pela seguinte equação:

$$MR = \frac{W}{F_{max}} \quad (16)$$

O BF é um parâmetro mais sensível ainda às ramificações do objeto que se deseja medir. É obtido pela razão entre o diâmetro W e o *feret* mínimo, conforme a eq. (17).

$$BF = \frac{W}{F_{min}} \quad (17)$$

5.6. Reconhecimento de Padrões e Classificação

A etapa de reconhecimento de padrões e classificação é a etapa final da sequência padrão do PADI. Um dos principais objetivos da análise de imagens por computador é dotar uma máquina com capacidade similar à dos seres humanos na realização de tarefas.

Um *padrão* é uma descrição quantitativa ou estrutural de um objeto ou qualquer outra região de interesse em uma imagem, que, em geral, é feita por um ou mais descritores, tais como aqueles discutidos na seção 5.5. Uma *classe de padrões* é um grupo de padrões que compartilham algumas propriedades em comum. O reconhecimento de padrões, então, envolve a atribuição dos padrões às suas respectivas classes automaticamente (Gonzalez & Woods, 2002).

Em PADI, esta técnica pode ser usada para classificar os objetos de uma imagem ou, como no caso deste trabalho, toda a região de uma imagem. A partir dos atributos dos objetos, constrói-se um espaço de atributos (espaço de características), onde cada objeto é representado por um vetor. Cada vetor é um ponto neste espaço, composto pelos atributos que o caracterizam. A classificação é feita de acordo com o posicionamento desses vetores no espaço de características, onde grupos formados por pontos com características similares, ou seja, próximos, são identificados (Gomes, 2001).

Os métodos de reconhecimento de padrões podem ser agrupados em duas grandes categorias: supervisionada e não-supervisionada. Na classificação supervisionada (Figura 40), as classes são determinadas de acordo com a sua proximidade, no espaço de características, aos grupos de padrões conhecidos previamente e definidos por um analista. Já na classificação não-supervisionada (Figura 41), nenhuma informação a respeito das classes é fornecida e a definição de cada uma delas é feita através de agrupamentos dos pontos próximos,

chamados *clusters*, em que as classes existentes num conjunto de dados são identificadas por si só (Gomes, 2001).

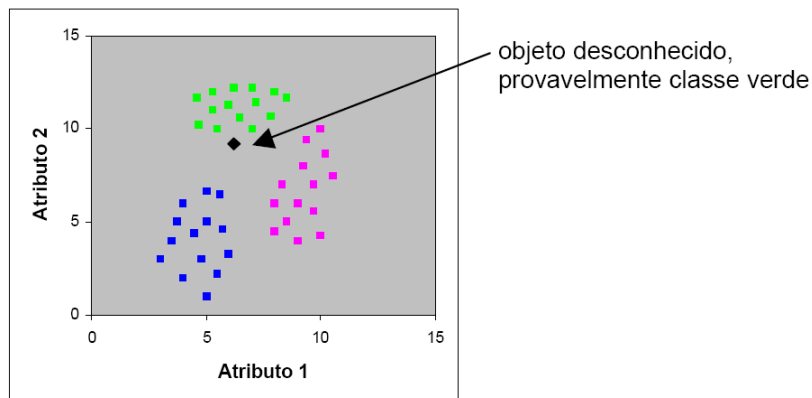


Figura 40 – Classificação supervisionada (Gomes, 2001)

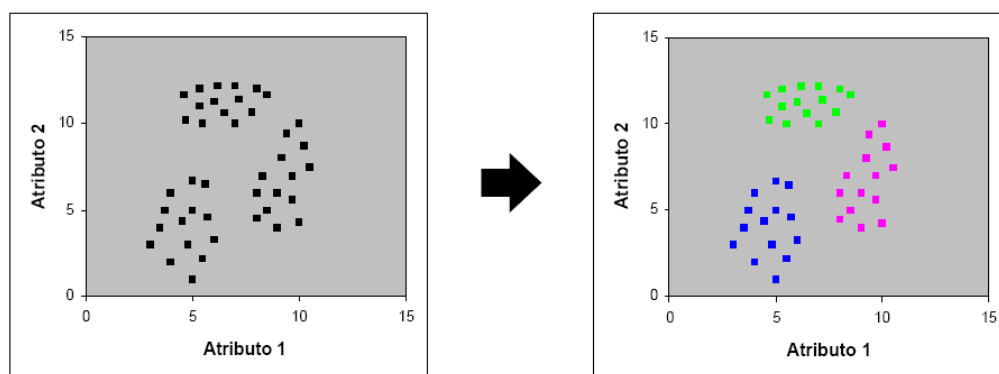


Figura 41 – Classificação não-supervisionada (Gomes, 2001)

Como o método de classificação utilizado neste trabalho é o supervisionado, nas próximas seções a atenção será voltada para este tipo.

5.6.1. Classificação Supervisionada

Conforme já explicado, na classificação supervisionada um objeto desconhecido é classificado através de sua similaridade aos grupos de objetos conhecidos. Para isso, são necessários um conjunto de treinamento, composto por objetos (ou regiões) conhecidos já separados em grupos (ou classes), um conjunto de atributos para a construção do espaço de características, e um classificador, que

consiste num algoritmo para alocar os objetos desconhecidos, baseado em medidas de similaridade, nas classes existentes.

5.6.1.1. Conjunto de Treinamento

O conjunto de treinamento é um conjunto formado por objetos (ou regiões) conhecidos previamente classificados. O ideal é que este conjunto contenha um grande número de objetos, a fim de ter uma boa representatividade, e apresente grande variabilidade dentro de cada classe, para que o sistema seja bastante genérico.

5.6.1.2. Conjunto de Atributos

O conjunto de atributos vai permitir caracterizar os objetos, tornando possível o agrupamento de objetos com características similares e a distinção entre esses grupos. Para facilitar a discriminação entre as classes, em geral, são utilizados vários atributos. Porém nem sempre a adição de mais atributos significa uma melhora na classificação.

Uma das formas de se obter o melhor conjunto de características é por busca exaustiva do conjunto que atinja a melhor taxa de acerto na classificação, onde são feitas todas as combinações possíveis de atributos e a de melhor taxa é selecionada. Além da seleção do conjunto de atributos ideal, existem técnicas de combinação de atributos, tais como a Análise de Componentes Principais (PCA) ou a Análise Discriminante Linear de Fisher (LDA), que constroem novos atributos através da combinação linear de outros, buscando maximizar a variância entre as classes.

A seguir serão explicadas em detalhes como funciona cada técnica aqui citada.

- **Busca Exaustiva do Conjunto com a Melhor Taxa de Acerto**

A busca exaustiva tem como objetivo a redução do número de atributos que irão representar os dados, simplificando a análise.

Assim, avaliam-se todas as possíveis combinações das variáveis (atributos) existentes, com todas as dimensões possíveis, e seleciona-se a que obtiver o maior valor na classificação.

O número de possíveis subconjuntos é dado pela equação abaixo:

$$n_d = \frac{p!}{(p-d)!d!} \quad (18)$$

Onde:

n_d é o número de subconjuntos;

p é o número de atributos;

d é a dimensão.

- **PCA**

No PCA é feita uma transformação linear do dado de entrada x de d dimensões para m dimensões, de forma que m seja menor do que d e preserve o máximo de variância entre os dados. Para que isto seja feito, é desejável que os atributos escolhidos para descrever os objetos numa imagem sejam descorrelacionados. Para fins práticos, a correlação é uma medida de quão bem se pode modelar a relação entre dois atributos através de uma função linear (Figura 42).

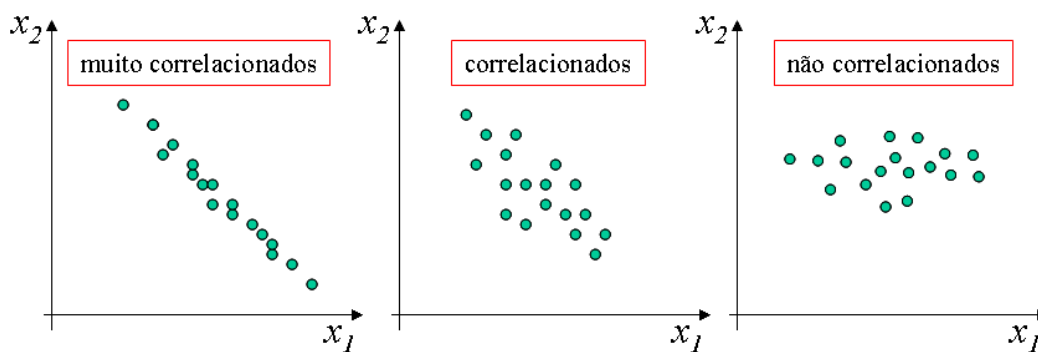


Figura 42 – Correlação entre dois atributos (Feitosa, [200-?])

Para facilitar o projeto e o treinamento do classificador, é necessário que poucos atributos sejam suficientes para discriminar os objetos de uma imagem. Assim, o objetivo do PCA é reduzir o número de atributos por translação e rotação dos eixos, de modo a minimizar o erro de reconstrução. No exemplo da Figura 43, observa-se que só é possível separar as classes utilizando-se os dois atributos X_1 e X_2 . Somente um novo atributo construído a partir de X_1 e X_2 , a componente Y_1 , é capaz de discriminar as classes sozinho.

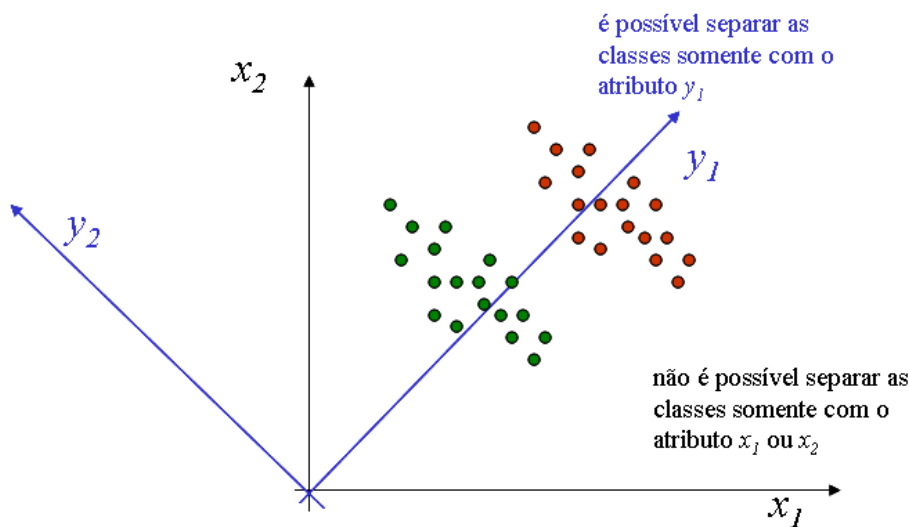


Figura 43 – Atributos para a descrição dos objetos em uma imagem (Feitosa, [200-?])

Do ponto de vista algébrico, o PCA é feito da seguinte forma: a base de dados é ajustada pela subtração da média correspondente a cada dimensão. Em seguida, é calculada a matriz de covariância, dada pela eq. (19), e depois os autovalores e autovetores por meio da eq. (20). O autovetor associado ao maior autovalor é a componente mais relevante. Quanto maior o autovalor, maior a importância do autovetor (componente). Para obter a representação de cada padrão na base do PCA, constrói-se um novo conjunto de dados através da eq. (21).

$$C = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \quad (19)$$

$$Ce = \lambda e \quad (20)$$

Onde:

C é a matriz de covariância amostral;

N é o número de padrões na amostra;

x é o valor do padrão;

μ é a média dada pela seguinte equação: $\mu = \frac{1}{N} \sum_{i=1}^N x_i$;

e é o autovetor;

λ é o autovalor;

$$y = x P \quad (21)$$

Onde:

y é o valor do novo banco de dados;

x é o valor do banco de dados antigo;

P são os autovetores, dado por: $P = [e_1 e_2 \dots e_n]$ e $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

- **LDA**

O método LDA tem como objetivo encontrar um novo conjunto de dados, também através de uma transformação linear de um conjunto de dados de entrada, com máxima separabilidade entre as classes, minimizando paralelamente a variância dentro das classes. A análise é similar ao método PCA, porém o cálculo dos autovalores é executado através das matrizes de dispersão inter-classes e intra-classes.

O LDA pode ser facilmente entendido através do exemplo da Figura 44, em que o caso explicado é unidimensional com duas classes. Enquanto o PCA não se preocupa em discriminar as classes numa população, o LDA minimiza o erro da classificação por meio da diminuição da abertura da distribuição dos dados e do afastamento de seus centróides.

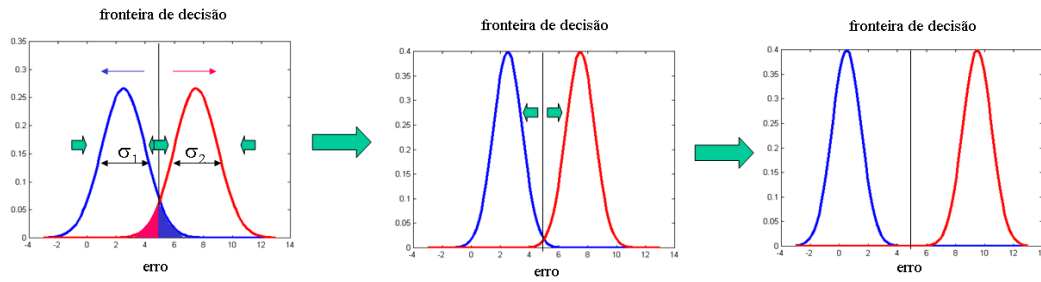


Figura 44 – Interpretação geométrica do LDA: Caso unidimensional com duas classes. (Feitosa, [200-?])

Assim, para diminuir o erro de classificação, procura-se maximizar a razão entre as variâncias inter-grupo e intra-grupo ($\sigma_{inter} / \sigma_{intra}$). As variâncias inter-grupo e intra-grupo do exemplo de duas classes da Figura 44 são calculadas de acordo com as Equações 22 e 23 respectivamente.

$$\sigma_{intra} = \frac{(\sigma_1 + \sigma_2)}{2} \quad (22)$$

$$\sigma_{inter} = \frac{[(\mu_1 - \mu)^2 + (\mu_2 - \mu)^2]}{2} \quad (23)$$

Onde:

σ é a variância;

μ é a média.

No caso multidimensional, procura-se a direção na qual a relação entre as variâncias inter-grupos e intra-grupos seja máxima. Estas direções são dadas pelos autovetores da matriz, calculados pela equação abaixo:

$$C_{intra}^{-1} C_{inter} \quad (24)$$

Onde:

$$C_{intra} = \frac{1}{g} \sum_{j=1}^g C_j \quad (25)$$

$$C_j = \frac{1}{N_j - 1} \sum_{i=1}^{N_j} (x_{ji} - \mu_j)(x_{ji} - \mu_j)^T \quad (26)$$

$$C_{inter} = \frac{1}{g} \sum_{j=1}^g (\mu_j - \mu)(\mu_j - \mu)^T \quad (27)$$

Em que g é o número de grupos e N é o número de elementos contidos no grupo j da amostra.

De modo análogo ao PCA, para obter a representação para cada padrão na base do LDA, tem-se:

$$y = x Q \quad (28)$$

Onde:

y é o valor do novo banco de dados;

x é o valor do banco de dados antigo;

Q são os autovetores, dado por: $Q = [e_1 e_2 \dots e_{g-1}]$ e $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{g-1}$.

A diferença entre PCA e LDA quando aplicados sobre os mesmos dados pode ser visualizada no exemplo da Figura 45.

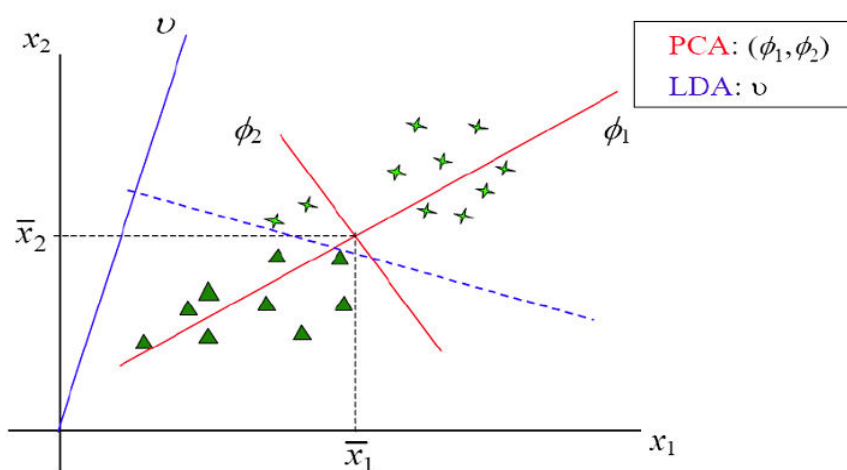


Figura 45 – Diferença entre PCA e LDA (Disponível em:

<<http://www.decom.ufop.br/menotti/rp112/slides/09-ClassificadoresLineares.ppt>>. Acesso em: 08 jun. 2012)

5.6.1.3. Classificador

O classificador é um algoritmo baseado em *funções de decisão* (ou *discriminantes*). O algoritmo mede a similaridade entre os objetos e entre estes e as classes, determina uma fronteira de decisão que separa os grupos e classifica o objeto a partir da função de decisão programada.

Os processos de classificação podem ser de dois tipos: paramétricos e não-paramétricos. Os classificadores paramétricos distinguem as classes através da função densidade de probabilidade que descreve o comportamento dos dados. Em outras palavras, o uso dos classificadores paramétricos deve levar à menor probabilidade de erros de classificação. As funções de decisão deste tipo de classificador são ótimas por minimizarem a perda média de classificações erradas. Porém, para que isto seja verdade, as funções densidade de probabilidade dos padrões de cada classe, bem como a probabilidade de ocorrência de cada classe, devem ser conhecidas. Este tipo de classificador é também chamado de *classificador bayesiano* (Gonzalez & Woods, 2002).

Os classificadores não-paramétricos são os mais simples, cujas fronteiras de decisão para discriminar as classes são resultantes de uma função de distância entre os grupos. Um dos classificadores não-paramétricos mais comuns é o *Classificador de Distância de Mahalanobis*. Este classificador calcula a distância entre o ponto a ser classificado e as classes, considerando o “espalhamento” dos dados de cada grupo, onde o valor central de cada classe é dado pela matriz de covariância e o vetor de média dos dados.

5.6.1.4. Etapas da Classificação Supervisionada

A classificação supervisionada é dividida em três etapas, executadas nesta mesma ordem: treinamento, validação e classificação.

A etapa de treinamento consiste no fornecimento do banco de dados de um conjunto de objetos conhecidos, composto pelo conjunto de características e as respectivas classes ao qual pertencem, para o sistema.

A validação verifica a qualidade do sistema. Esta etapa é realizada com a classificação de objetos previamente conhecidos pelo analista, que podem ser os mesmos fornecidos na etapa de treinamento, caso o número de objetos conhecidos seja pequeno, ou outros diferentes destes. Obviamente a classificação dos mesmos

objetos utilizados para o treinamento do sistema, também chamada de autovalidação, fornece uma estimativa de qualidade inferior ao da classificação de objetos desconhecidos pelo sistema. O resultado da validação é obtido através do cálculo da taxa de acerto da classificação (Gomes, 2001).

A partir destas duas etapas, o sistema está pronto para a classificação. A classificação, em geral, é o objetivo principal da automatização do processo, e consiste na classificação de objetos desconhecidos, tanto pelo analista quanto pelo sistema, cuja confiabilidade dependerá do resultado obtido na etapa de validação.