

### 3 Revisão Bibliográfica

Mineral é todo sólido natural cristalino inorgânico de composição química e de propriedades físicas definidas. Minério é toda rocha constituída de um mineral ou agregado de minerais contendo um ou mais minerais valiosos, que podem ser aproveitados economicamente. Esses minerais valiosos, aproveitáveis como bens úteis, são chamados de minerais-minério. O mineral ou conjunto de minerais não aproveitados de um minério é denominado ganga [2].

#### 3.1. Minério de Ferro

O minério de ferro é a rocha, a partir da qual, pode ser obtido o ferro metálico de maneira economicamente viável [3]. O ferro é um elemento químico de símbolo Fe, número atômico 26 e massa atômica relativa igual a 55,85 [4]. Em temperatura ambiente encontra-se em estado sólido, mas por ser um elemento muito reativo, não é encontrado na natureza na forma metálica, ou seja, sem estar ligado quimicamente a outro elemento [5].

O ferro é um dos elementos mais abundantes da crosta terrestre, representando, em massa, cerca de 5%. Também é encontrado como constituinte, em maior ou menor quantidade, em outras classes minerais. Como a combinação do ferro com outros compostos pode ocorrer de formas variadas, existem muitos minerais de ferro diferentes. Nas condições atuais, os óxidos são utilizados para a extração de ferro. Isso ocorre pela quantidade de ferro presente nesses minerais ou pela concentração desses minerais nas rochas, que formam os corpos de minérios [6].

A mineração é uma indústria global presente em países desenvolvidos e em desenvolvimento. Em países em desenvolvimento, essa atividade tem uma importância significativa para o produto interno bruto (PIB), contribuindo para a redução da pobreza. Por exemplo, segundo relatórios do Conselho Internacional de Mineração e Metais (ICMM), a produção da indústria de mineração e metais é necessária para estimular avanços e progressos tecnológicos em uma sociedade [5].

Os maiores depósitos de minério de ferro, no Brasil, são praticamente todos do tipo hematítico, com altos teores de ferro. As reservas medidas e indicadas de minério de ferro no Brasil alcançam 23 bilhões de toneladas (Tabela 3-1), situando

o país em terceiro lugar em relação às reservas mundiais que são, de 170 bilhões de toneladas [7].

Entretanto, em termos de ferro contido, as reservas brasileiras, com um teor médio de 49,0% de ferro, representam 13,1% das reservas mundiais. A produção brasileira de minério de ferro em 2013 atingiu 386 milhões de toneladas, com um teor médio de 63,6% de ferro. Os principais estados brasileiros detentores de reservas de minério de ferro são: Minas Gerais (72,5% das reservas e teor médio de 46,3% de Fe), Mato Grosso do Sul (13,1% e teor médio de 55,3%) e Pará (10,7% e teor médio de 64,8%) [7].

Entre todos os metais, o ferro é o mais produzido e o que está mais presente em nossa vida. A indústria siderúrgica utiliza o minério de ferro e seus aglomerados como matéria prima básica para a produção de aço, produto essencial para a civilização moderna e cuja produção é considerada a espinha dorsal do desenvolvimento industrial de um país. Cerca de 97% do minério de ferro extraído das minas é utilizado na siderurgia, sendo que, os 3% restantes são utilizados nas indústrias de cimento, química, etc. [8] [9].

Os minerais de ferro brasileiros apresentam texturas muito variadas, devido às diferentes condições de metamorfismo, tectonismo e intemperismo a que foram sujeitos, ou, mesmo, em virtude de sua gênese. Conseqüentemente, dentro de uma mesma amostragem, poderão ocorrer minérios de diferentes características mineralógicas e microestruturais [10], [11], [12].

Os minerais de ferro considerados economicamente exploráveis são agrupados de acordo com a sua composição química, nas classes: óxidos, carbonatos, sulfetos e silicatos, sendo esta última de menor expressão econômica. Cada classe de minerais pode ser representada por um ou mais minerais, sendo que, destes, apenas os minerais de ferro da classe óxido, são explorados economicamente, nas condições atuais [13].

Tabela 3-1: Reservas mundiais de minério de ferro no ano 2013 [14].

Discriminação	Reservas (10 <sup>6</sup> t)	Produção (10 <sup>3</sup> t)		
		2012 (e)	2013 (e)	(%)
Países	2013 (e)	2012 (e)	2013 (e)	(%)
Austrália	35.000,0	521.000	530.000	18,0
Rússia	25.000,0	105.000	102.000	3,5
<b>Brasil</b>	<b>23.126</b>	<b>400.627</b>	<b>386.270</b>	<b>13,1</b>
China	23.000,0	1.310.000	1.320.000	44,7
Índia	8.100,0	144.000	150.000	5,1
Ucrânia	6.500,0	82.000	80.000	2,7
Outros países	49.274	367.373	381.730	12,9
<b>TOTAL</b>	<b>170.000</b>	<b>2.930.000</b>	<b>2.950.000</b>	<b>100,0</b>

(1) reserva lavrável; (2) Estimativa de produção da China baseada em minério bruto;

(e) dados estimados, exceto Brasil.

### 3.2. Hematita

A hematita é o mineral de ferro de maior importância. Isto se deve a sua alta ocorrência em vários tipos de rochas e suas origens diversas [9]. Sua fórmula química é  $Fe_2O_3$ , a qual apresenta uma fração mássica em ferro de 69,94% e uma fração mássica em oxigênio de 30,06% [4].

Este mineral pode ser encontrado com diferentes morfologias como, por exemplo, microcristalina, lobular, granular ou lamelar. A hematita também pode ser encontrada com hábito de magnetita formando uma fase mista chamada de martita [15]. Estas morfologias estão relacionadas a uma condição de cristalização diferenciada, que poderá ter influência no seu processamento [11].

O tamanho, a forma e a distribuição dos cristais de hematita podem influenciar na redutibilidade e resistência mecânica dos aglomerados.

As hematitas compactas aumentam a resistência mecânica dos aglomerados, mas reduzem a porosidade e redutibilidade [1]. Estes tipos de hematita são apresentados na Tabela 3-2.

As hematitas não compactas aumentam a porosidade e redutibilidade dos aglomerados, porém reduzem sua resistência mecânica [1]. Estes tipos de hematita podem ser observados na Tabela 3-3.

Tabela 3-2: Tipos de hematita compacta [16].

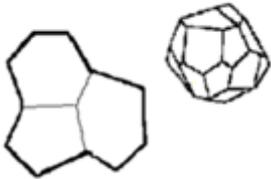
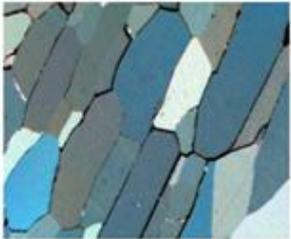
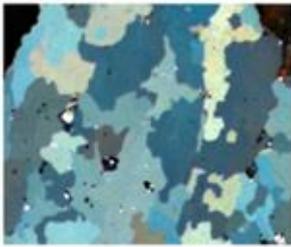
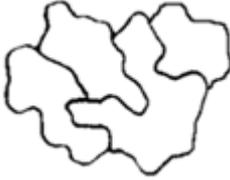
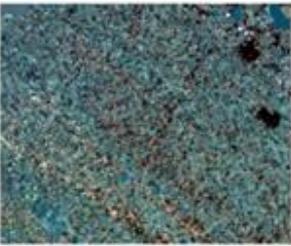
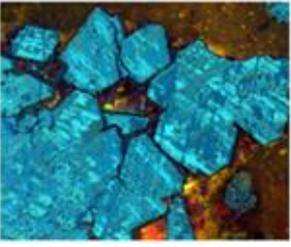
Tipo	Características	Forma textura	Ilustração esquemática
Hematita Granular	<ul style="list-style-type: none"> <li>✓ Formatos regulares equidimensionais.</li> <li>✓ Contatos retilíneos e junções tríplices.</li> <li>✓ Cristais compactos.</li> </ul>		
Hematita Lamelar	<ul style="list-style-type: none"> <li>✓ Cristais inequidimensionais, hábito tabular.</li> <li>✓ Contato retilíneo.</li> <li>✓ Cristais compactos.</li> </ul>		
Hematita Lobular	<ul style="list-style-type: none"> <li>✓ Formatos irregulares inequidimensionais.</li> <li>✓ Contatos irregulares, geralmente imbricados.</li> </ul>		

Tabela 3-3: Tipos de hematita não compacta [16].

Tipo	Características	Forma textura	Ilustração esquemática
Hematita Microcristalina	<ul style="list-style-type: none"> <li>✓ Cristais muito pequenos &lt; 0,01mm.</li> <li>✓ Textura porosa.</li> <li>✓ Contatos pouco desenvolvidos.</li> </ul>		
Martita	<ul style="list-style-type: none"> <li>✓ Hematita com hábito de magnetita.</li> <li>✓ Oxidação segundo os planos cristalográficos de magnetita.</li> <li>✓ Geralmente porosa.</li> </ul>		

A hematita é um material opaco no Microscópio Ótico de luz Refletida (MLR), mas para lamelas muito finas e em luz transmitida é vermelha sanguínea escura, e pode apresentar dicroísmo de vermelho acastanhado a vermelho amarelado. Este mineral tem um coeficiente de reflexão médio de 25%-32% no ar [17].

Para o presente trabalho a propriedade ótica mais importante do mineral de ferro é sua birrefletância devido à forte anisotropia que ele apresenta. Deste modo, sua reflexão e, conseqüentemente, o seu brilho na imagem mudam com diferentes orientações dos cristais com relação ao plano de incidência da luz. Este efeito ainda pode ser acentuado quando é usada luz polarizada no MLR [17].

### **3.3. Microscopia**

Os microscópios são instrumentos que produzem magnificação visual ou imagens fotográficas de objetos muito pequenos que não podem ser vistos a olho nu. Os microscópios precisam realizar três tarefas: produzir ampliação de espécies, separar detalhes na imagem, e gerar detalhes visíveis para os olhos humanos e câmeras [18].

#### **3.3.1. Microscopia Digital**

A microscopia digital pode ser definida como a integração do microscópio e o computador, abarcando as seguintes etapas: aquisição digital de imagens, automação do microscópio e análise de imagens [19].

Os sistemas controlados por softwares e o uso de rotinas que envolvem programação podem propiciar uma automação completa do microscópio. Exemplos de funções que podem se realizar através desses softwares são a motorização da platina, a troca de lentes, a focalização, o ajuste de iluminação e cor, a captura automática de várias imagens e de mosaicos, etc. Em geral, estas funções já vem disponível nos softwares, e em alguns casos, podem ser programadas independentemente [20].Trazendo reprodutibilidade ao processo.

A microscopia digital é utilizada para minimizar os erros humanos já que a visão humana é raramente uma ferramenta confiável. Além disto, a microscopia digital pode introduzir rigor ao processo de obtenção de informações quantitativas, e permite capturar imagens digitais. O objetivo é eliminar ao máximo a interferência do operador na obtenção e processamento das imagens [19].

A Figura 3-1 apresenta o sistema integrado de microscopia ótica digital do Laboratório de Microscopia Digital (LMD) da PUC-Rio. O microscópio do LMD é um microscópio motorizado do tipo Zeiss AxioImager-M2m, com câmera digital Zeiss AxioCam MRc5, automatizados com o software *AxioVision*.

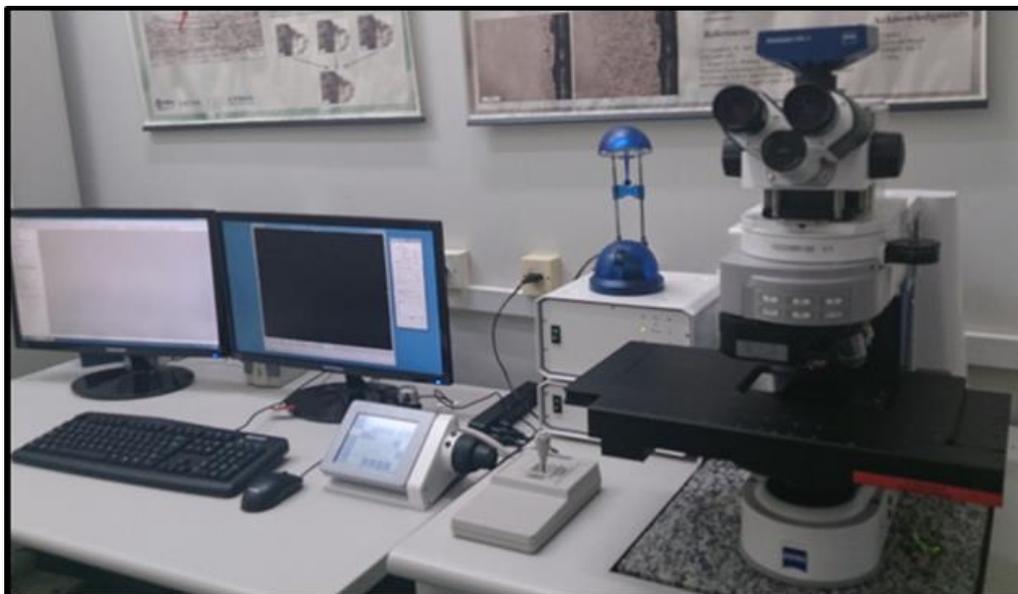


Figura 3-1: Microscópio ótico do LMD da PUC-Rio.

Existem dois modos de funcionamento do microscópio ótico: o modo de luz transmitida, para minerais transparentes, e o de luz refletida, para minerais opacos. Estes dois modos, são provavelmente os métodos de identificação de minerais mais tradicionais que existem. Baseiam-se ambos na interação da luz (geralmente luz branca do espectro visível) com os minerais, e são bastante precisos e flexíveis para análises qualitativas [2].

As análises exigem preparação especial das amostras, em seções delgadas (luz transmitida), polidas (luz refletida), ou delgadas polidas (ambas). No caso de luz refletida é possível analisar fragmentos de rochas ou montagens de material particulado, embutido em resinas (geralmente epóxi, acrílica ou poliéster) [2].

A microscopia ótica de luz refletida possibilita a caracterização de minério de ferro e seus aglomerados (sinter e pelotas) porque as fases mais comuns presentes são facilmente discriminadas através das diferentes refletâncias de cada uma delas [6].

### 3.3.2. Microscopia Ótica de Luz Refletida em Campo Claro

A microscopia ótica de luz refletida permite que materiais opacos (metais, cerâmicos, compósitos, etc.) sejam analisados [21].

Quando um feixe de luz, com comprimentos de onda na faixa da luz visível, incide na amostra a ser analisada, os fótons interagem com os elétrons dos átomos do material. Neste fenômeno, os elétrons são excitados pela luz e reemitem luz, que será capturada ou não pela lente do microscópio, gerando o contraste na imagem formada. Esta interação ocorre de maneira diferente, dependendo do elemento ou do composto em cada parte do material envolvido. Assim, é possível obter refletâncias com características distintas de cada fase da amostra, gerando uma imagem com regiões específicas de cores diferentes [22].

Em campo claro, normalmente, a luz que é absorvida ou desviada pela amostra não é capturada pela lente objetiva. Desta forma, os objetos que causaram esta absorção ou desvio aparecem escuros contra um fundo claro na imagem, como é mostrado no esquema da Figura 3-2 [22].

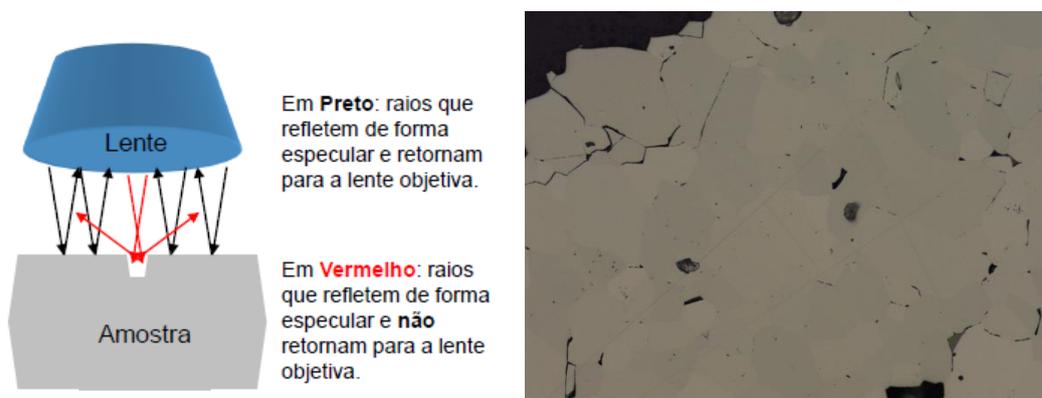


Figura 3-2: Esquema da Luz Refletida em Campo Claro [22].

### 3.3.3. Luz Polarizada

A luz polarizada, vibra num plano só (plano de polarização) e os vetores intensidade do campo elétrico ( $\vec{E}$ ) e intensidade do campo magnético ( $\vec{H}$ ) são normais à direção de propagação [17], como é mostrado no esquema da Figura 3-3.

Amostras contendo cristais anisotrópicos podem apresentar pleocroísmo (alteração da refletância em função da orientação cristalina quando iluminadas

com luz polarizada). Esta alteração de refletância poderá permitir discriminar cristais adjacentes que, em Campo Claro, teriam refletâncias quase idênticas [22].

A configuração mais comum utiliza um polarizador linear fixo para a luz incidente e um analisador giratório para a luz refletida ou transmitida. Variando o ângulo de analisador varia-se o contraste dos diferentes cristais que alteram diferentemente a polarização da luz incidente [22].

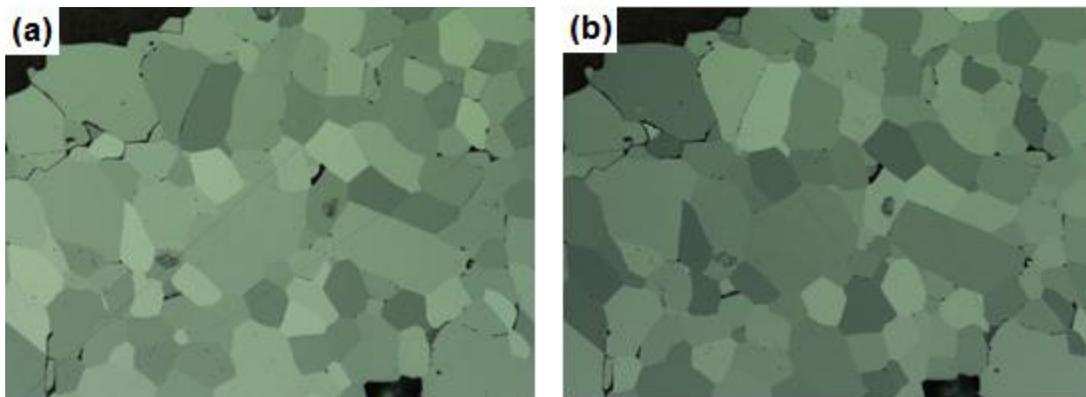
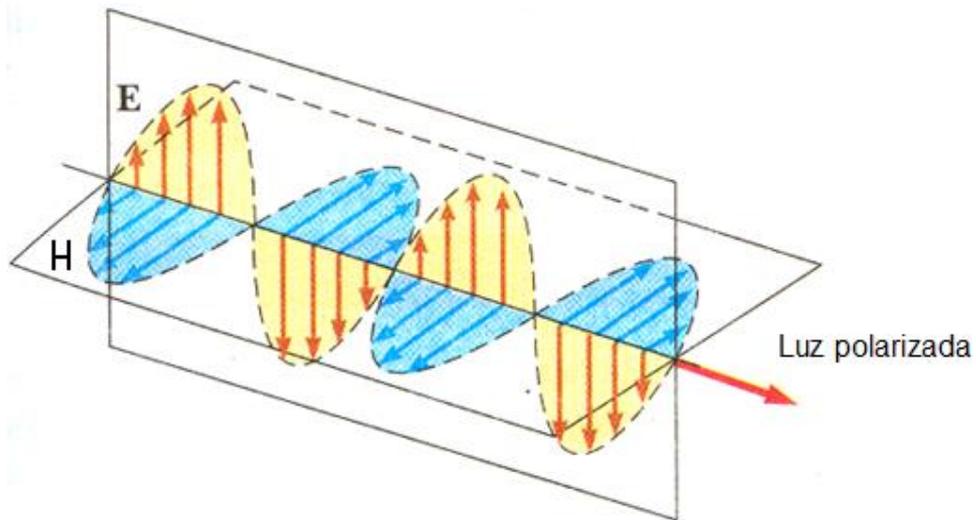


Figura 3-3: Esquema da formação da luz polarizada. (a) Imagem LPOL com polarizador girado  $-10^\circ$  e (b) Imagem LPOL com polarizador girado  $+10^\circ$ .

### 3.3.4. Polarização Circular

A luz circularmente polarizada consiste em uma onda de luz cujo vetor campo elétrico descreve uma trajetória helicoidal como apresenta a Figura 3-4, daí a denominação polarização circular [23].

O sistema ótico de polarização circular do microscópio do DEQM foi montado com um sistema ótico composto por um polarizador, um analisador e duas placas de um quarto de onda, colocados em um refletor. Ambos os lados do

refletor, entrada e saída de luz, são constituídos por uma combinação de polarizador e placa de um quarto de onda, fixados transversalmente [17].

Assim, a luz não-polarizada passa pelo polarizador, tornando-se linearmente polarizada. Em seguida, a luz linearmente polarizada incide na primeira placa de um quarto de onda, orientada com um ângulo de  $45^\circ$  em relação ao plano de polarização da luz incidente. Essa placa divide a luz em duas componentes com uma diferença de fase de  $90^\circ$ . A combinação dessas duas ondas linearmente polarizadas, de mesma amplitude e defasadas  $90^\circ$ , resulta em uma onda circularmente polarizada [22], como mostrado na Figura 3-4.

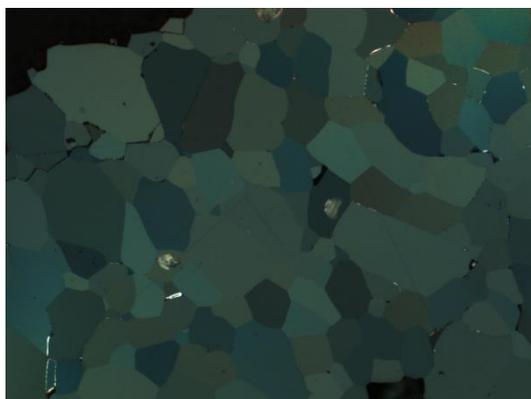
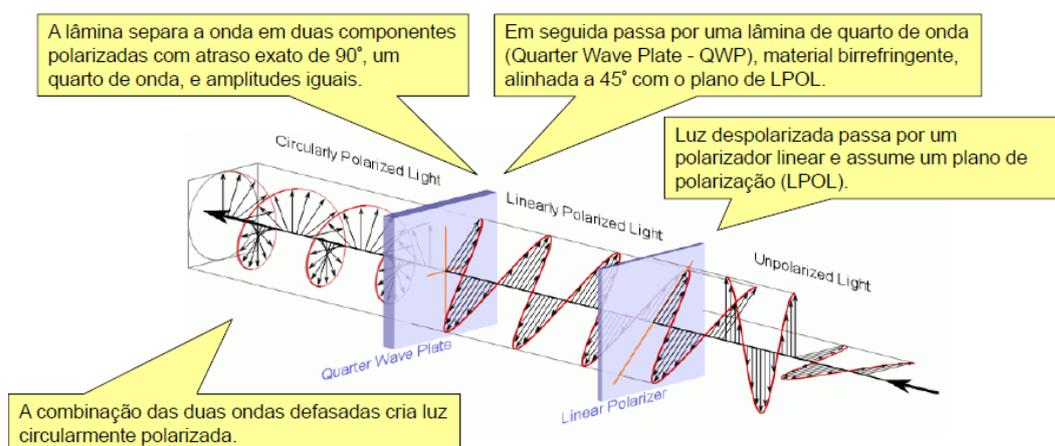


Figura 3-4: Esquema da formação da luz polarizada circular [22].

A luz circularmente polarizada incide na amostra e é refletida. Então, a luz refletida passa pela segunda placa de um quarto de onda, que é orientada ortogonalmente à primeira. A segunda placa converte a luz circularmente polarizada em linearmente polarizada, com uma orientação diferente daquela do polarizador. Em seguida, essa luz passa pelo analisador e segue para a câmera [22].

Na imagem resultante da polarização circular (CPOL), não há ponto de extinção, pois todas as orientações de polarização estão presentes. Esta técnica inovadora de polarização circular em microscopia de materiais, permite melhorar o contraste entre os cristais na imagem. Também, as cores não variam com a rotação da amostra. Dessa forma, objetos que costumavam ser visíveis apenas em uma direção agora podem ser vistas em sua totalidade independente de sua orientação e sem rotação da platina [17].

### **3.3.5. Aquisição Automática de Imagens**

A utilização de microscópios equipados com câmeras coloridas CCD e seu controle através do computador facilitam a aquisição automática de imagens. Esta forma de aquisição permite a captura de inúmeras imagens numa velocidade muito maior, cobrindo grande região da amostra, evitando a ocorrência de erros por parte do operador e também a análise de campos repetidos da amostra [20].

Fazendo uso da Microscopia Digital, é possível realizar a captura de imagens-mosaico. O funcionamento desta técnica consiste na captura de ladrilhos distribuídas uniformemente em uma região demarcada pelo operador. Neste caso, duas posições x-y opostas são fornecidas como dado de entrada. Assim, a rotina determina a área de interesse e a quantidade máxima de ladrilhos igualmente espaçados dentro desta área. A platina vai se deslocar no plano x-y conforme os ladrilhos vão sendo adquiridos em posições previamente calculadas. Na Figura 3-5 (a), pode ser observado o caminho percorrido na captura do mosaico [22].

A automação permite a captura com o recurso de autofoco, de modo a garantir uma imagem resultante de qualidade. O recurso consiste em obter o foco automaticamente para cada ladrilho capturado, mas nem sempre é factível fazer autofoco em cada ladrilho pelo tempo excessivo que demoraria o processo de captura. Para agilizar este processo existe a opção de alternar o autofoco entre ladrilhos, tão espaçados quanto possível, levando em consideração a relação tempo-qualidade [20], [21].

A captura de cada “ladrilho” ocorre de modo que se superponha parcialmente ao “ladrilho” vizinho, para que sejam realizadas posteriormente correções de alinhamento por rotinas de correlação. Todos estes recursos, fruto da automação do microscópio, tornam possível a captura de imagens cobrindo grandes áreas em alta resolução, sem interferência do operador [22].

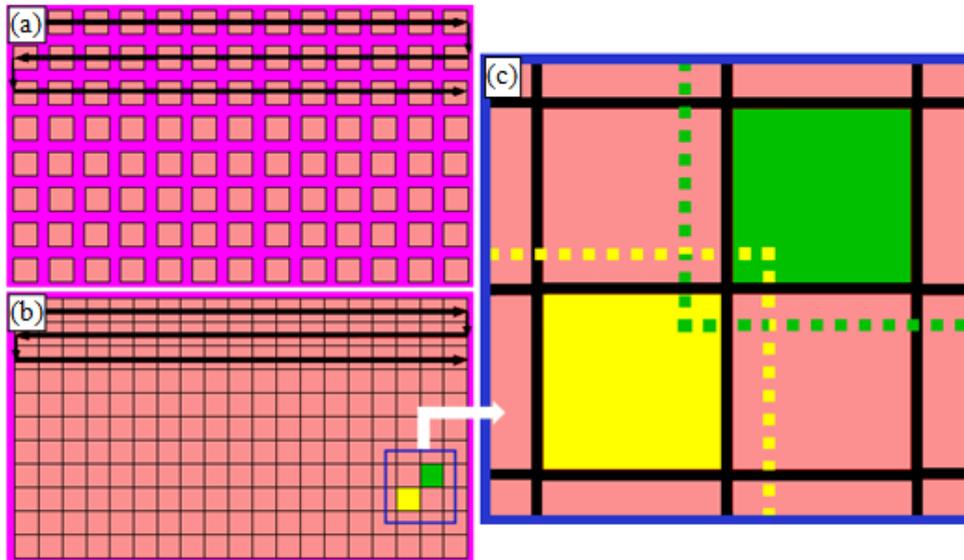


Figura 3-5: Técnica de captura do mosaico; (a) varredura espacial com captura de imagens individuais; (b) concatenação das imagens e criação do mosaico; (c) visão ampliada para exemplificar a superposição dos campos verde e amarelo [20].

### 3.4. Processamento e Análise Digital de Imagens

O processamento Digital de Imagens (PDI) é uma técnica que utiliza operações matemáticas para alterar os valores dos pixels de uma imagem digital, modificando-a para facilitar sua visualização e análise. Por outro lado a Análise Digital de Imagens (ADI) consiste na extração e tratamento de dados quantitativos de imagens digitais. Deste modo é empregado o termo Processamento e Análise Digital de Imagens (PADI) para englobar as duas técnicas anteriores, como é mostrado no esquema da Figura 3-6 [24].

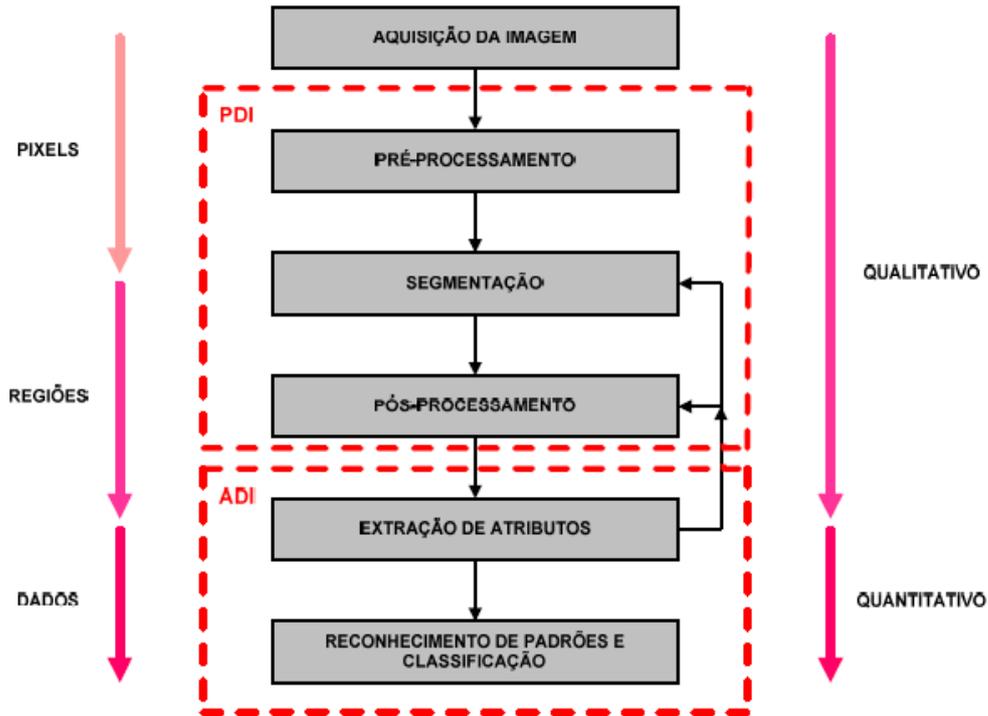


Figura 3-6: Sequência padrão de PADI [24].

### 3.4.1. Registro de Imagem

Quando várias imagens de um mesmo campo são obtidas por diferentes sensores ou com um mesmo sensor, porém sob diferentes condições de captura, é necessário um registro espacial [25]. No presente trabalho isto será importante porque serão capturadas imagens em dois modos: Campo Claro e Polarização Circular.

O registro espacial de imagens, ou simplesmente registro é um importante passo no processamento. O propósito desta técnica é criar imagens geometricamente iguais com uma coerência pixel a pixel.

A característica fundamental do registro consiste em corrigir o alinhamento geométrico de duas ou mais imagens de um mesmo campo. Para isto, toma-se uma imagem como referência enquanto as outras são registradas a ela. O registro baseia-se em encontrar uma transformação capaz de remapear as posições dos pixels da imagem registrada. Isto é feito de modo que a área sobreposta esteja alinhada com a imagem de referência [26].

Esta transformação pode ser composta por uma combinação de seis transformações básicas distintas:

- Translação;
- Rotação;
- Escala;
- Paralelismo;
- Projeção e;
- Outras distorções como curvaturas, distorções locais, etc. [27].

A Figura 3-7 apresenta estas diferentes transformações, mostrando claramente seu efeito em uma imagem exemplo.

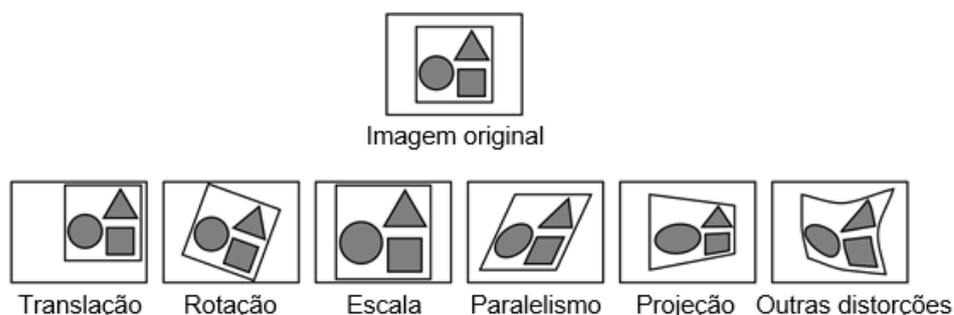


Figura 3-7: Possíveis transformações em registro de imagens [27].

Historicamente as transformações têm sido classificadas como rígidas ou não rígidas. Uma transformação é denominada rígida quando unicamente translações e rotações são permitidas. Este tipo de transformação é também conhecida como euclidiana, já que as distâncias euclidianas são preservadas. Por oposição, as demais transformações são definidas como não rígidas ou elásticas [27]. Neste trabalho são de interesse específico as transformações rígidas, como será visto posteriormente.

A primeira coisa a fazer para registrar uma imagem é encontrar pontos em comum entre a imagem de referência e a imagem a ser registrada. Estes pontos são chamados de pontos de controle (regiões fechadas, arestas, contornos, interseções de linhas, cantos, etc). A partir destes pontos estimam-se os parâmetros do modelo de transformação que irá gerar a imagem registrada que será sobreposta à imagem de referência [17].

Alguns autores afirmam que para conseguir um registro é necessário como mínimo encontrar 3 pontos de controle entre a imagem de referência e a imagem

a ser registrada [28]. Certo é que quanto maior seja a quantidade de pontos, maior será a qualidade do registro como tal [26].

Existe uma técnica de processamento de imagens que permite a detecção e extração automáticas de pontos de controle. Estes pontos de controle são invariáveis a ruído de imagem, rotação, translação, escala, e de forma parcial a mudanças de iluminação e de perspectivas. Esta técnica é chamada de transformada SIFT, ou simplesmente SIFT (*Scale Invariant Feature Transform*).

A SIFT é formada por duas partes distintas: o detector e o descritor. O detector SIFT está baseado em cálculos de diferença de Gaussianas e o descritor SIFT utiliza histogramas de gradientes, orientados para descrever a vizinhança local dos pontos de interesse. A transformada SIFT tem demonstrado ser muito eficiente ao gerar grande número de pontos de controle que conseguem cobrir densamente uma imagem [29].

O *software* livre FIJI, é um dos muitos softwares que dispõem de função para a transformada SIFT no registro de imagens.

Por outra parte, o *software* comercial *AxioVision* emprega outra técnica de registro de imagem, na elaboração de mosaicos. Isto é necessário, já que a técnica de mosaico tem como restrição um alinhamento imperfeito entre as direções x e y da platina do microscópio e da imagem formada pela câmera, fazendo com que os campos capturados fiquem desalinhados [20]. O *AxioVision* resolve este problema através de uma função chamada *Stitching*, a qual se baseia no registro de imagens para alinhar os ladrilhos do mosaico.

### **3.4.2. Correção de Fundo**

Problemas de iluminação são muito comuns em microscopia ótica. Estes problemas podem ser oriundos de defeitos no sistema ótico do microscópio, como desalinhamento de diafragmas ou sujeira nas lentes, ou de problemas da amostra, como uma superfície inclinada com relação ao eixo ótico.

Os defeitos oriundos do microscópio normalmente são corrigidos capturando uma imagem de referência de uma amostra uniforme, sem nenhuma estrutura interna, e subtraindo automaticamente esta imagem de referência de todas as imagens capturadas.

Os defeitos associados à não planicidade da amostra podem ser reduzidos prensando a amostra contra uma massa flexível e uma superfície plana.

Em certos casos específicos, como na captura de imagens no modo de polarização circular, as imagens sempre apresentam um grande problema de iluminação devido às características dos componentes óticos utilizados. Nestes casos, também é possível capturar uma imagem de referência, mas isto é pouco prático porque exigiria mudar, constantemente, entre as referências para captura em campo claro e CPOL. Assim, é comum utilizar processamento de imagens para realizar a correção de iluminação. Para isso, a imagem CPOL é fortemente borrada com um filtro passa-baixa com kernel grande, e esta imagem borrada é subtraída da imagem original. A Figura 3-8 ilustra esta situação [30].

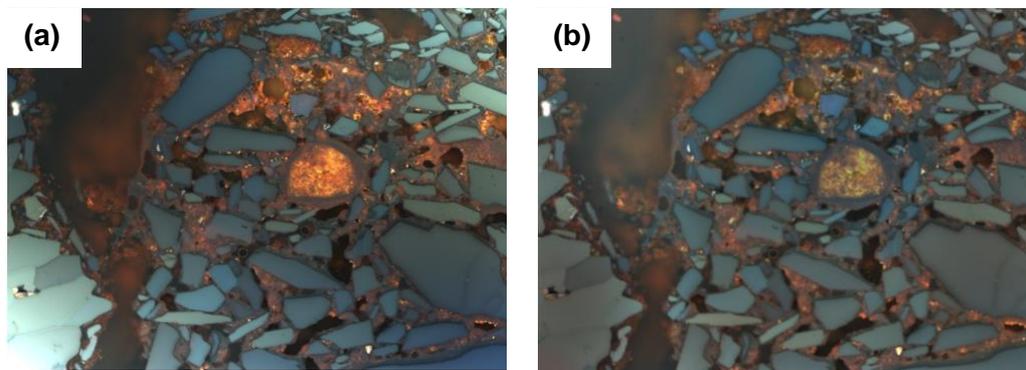


Figura 3-8: Ilustração do processo de correção de iluminação. (a) Imagem original, (b) Imagem corrigida.

### 3.4.3. Filtragem de Ruído

As imagens de microscopia ótica de luz refletida em geral apresentam pouco ruído. Como veremos mais adiante, o aumento da saturação de cor da câmera digital incrementa o contraste entre componentes das imagens, facilitando a segmentação, mas também aumenta a presença de ruído. Nesta situação, deve ser utilizado um filtro passa-baixa.

Dentre os filtros passa-baixa mais eficientes, que reduzem ruído sem borrar a imagem, um dos mais modernos é o Filtro *Non-Local Means* (NLM) [31]. Este filtro reduz o ruído em uma dada região da imagem utilizando um valor de intensidade calculado a partir de uma estatística em outra região da imagem que não contém bordas. Desta maneira, consegue reduzir ruído de forma muito eficiente, sem perda de detalhe das bordas dos objetos. A Figura 3-9 ilustra o funcionamento deste filtro.

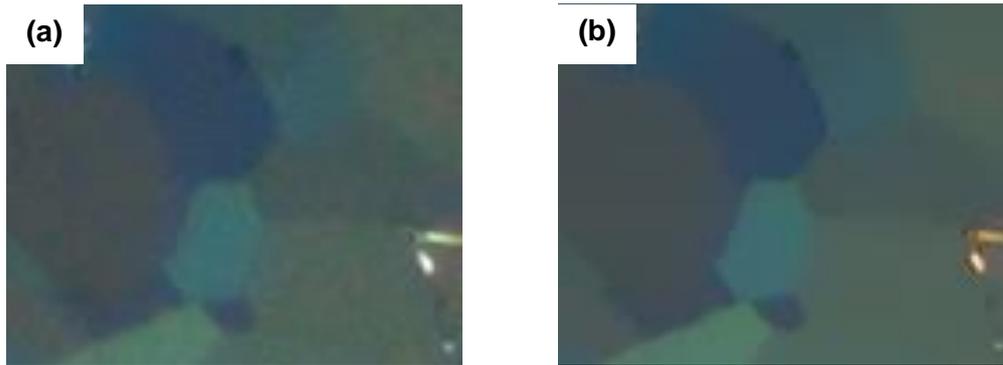


Figura 3-9: Ilustração do efeito do filtro NLM. (a) Imagem original;(b) Imagem depois da aplicação NLM.

#### 3.4.4. Técnicas de segmentação

A segmentação pode ser definida como o processo de separação de uma imagem em suas partes ou objetos constituintes, sendo uma das tarefas mais difíceis na análise de imagens [32].

Não existe um método ideal e comum de segmentação [27]. A segmentação costuma ser a etapa mais crítica da sequência padrão de PADI. Uma segmentação adequada praticamente garante o sucesso no reconhecimento e na identificação dos objetos sobre os quais será feita a análise [33].

Existem diversas técnicas de segmentação [17].As empregadas neste trabalho foram:

- (i) Segmentação por faixa tonal ou limiarização;
- (ii) Segmentação por textura;
- (iii) Segmentação por crescimento de regiões.

##### 3.4.4.1. Limiarização

A técnica de segmentação mais simples e a mais utilizada é a segmentação por faixa tonal, também chamada *thresholding* ou limiarização. A limiarização usa o tom dos pixels para separar objetos (regiões de pixels contíguos com tons dentro de uma faixa tonal delimitada) de outros usando um tom de corte [24].

No caso de diferenciar objetos de um fundo, utiliza-se a limiarização bimodal. Esta segmentação parte da hipótese de que a imagem apresenta um histograma bimodal e, portanto, que os objetos podem ser separados do fundo por uma simples operação que compara os tons dos pixels da imagem com um valor

de limiar ( $L$ ). Supondo que a imagem  $f(x,y)$  corresponda a um histograma bimodal, então a imagem segmentada  $g(x,y)$  seria definida como:

$$g(x,y) = \begin{cases} 1 & \text{Se } (x,y) > L \\ 0 & \text{Se } (x,y) \leq L \end{cases} \quad 3-1$$

O resultado da limiarização seria uma imagem binária, onde os pixels com valor 1 correspondem aos objetos, enquanto os pixels com valor 0 correspondem ao fundo.

A técnica de limiarização também é aplicável se for necessário discriminar mais de uma faixa tonal na imagem, pois ela não está restrita apenas a duas fases. Este tipo de limiarização é conhecido como limiarização multimodal e gera tantas imagens binárias quantas fases sejam segmentadas. Neste caso os pixels brancos da imagem binária formam a fase de interesse, que fica entre os dois tons de corte, e o fundo preto é o resto. Assim, supondo que a imagem  $f(x,y)$  corresponda a um histograma multimodal então, neste caso, a imagem segmentada  $g(x,y)$  seria definida como:

$$g_{(x,y)} = \bigcup_{i=1}^n g_i(x,y) \quad 3-2$$

Onde,

$$g_i(x,y) = \begin{cases} 1 & \text{Se } L_i < (x,y) \leq L_{i+1} \\ 0 & \text{Se } L_i \geq (x,y) > L_{i+1} \end{cases} \quad 3-3$$

Ou seja,  $g_i(x,y)$  seria a  $i$ -ésima imagem binária pertencente à  $i$ -ésima fase.

O resultado final da limiarização multimodal seria uma imagem binária, onde, os pixels com valor 1 corresponderiam às fases de interesse, enquanto os pixels com valor 0 corresponderiam ao resto.

#### 3.4.4.2. Segmentação por Textura

Mesmo que não exista uma definição formal, a textura em PADI pode ser entendida como uma medida de propriedades como suavidade, rugosidade, regularidade e granularidade [24].

As três formas principais utilizadas em PADI para a medição de texturas são a estatística, a estrutural e a transformada. 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.

Por sua vez, as técnicas de transformada 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 [24]. Como neste trabalho a técnica utilizada para caracterização da textura é a estatística, então será estudado este tipo de medida.

As técnicas estatísticas são uma das mais antigas e mais simples técnicas para a descrição de texturas. Os parâmetros estatísticos extraídos estão divididos em dois grupos: os de primeira ordem e os de segunda ordem. Os parâmetros de primeira ordem, tais como intensidade média ou contraste, são calculados de forma simples através dos histogramas de níveis de cinza de uma imagem ou de uma região. Estes parâmetros podem ser classificados como medidas de ocorrência de pixels com certas intensidades nas imagens.

Já 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, onde cada parâmetro fornece uma característica de textura, tais como contraste, homogeneidade, complexidade, etc [24].

#### **3.4.4.2.1. 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. Em vista disso, para trazer essa informação ao processo de análise de texturas, foram propostas as estatísticas baseadas na matriz de co-ocorrência. Onde 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  $N_x \times N_y$ , segundo a notação linha-coluna, e cujos

elementos sejam os valores de tom de cinza dos pixels na posição correspondente. Dessa forma, sua matriz de co-ocorrência é uma matriz quadrada, de dimensão  $N_g \times N_g$ . Assim, o elemento da linha  $i$  e coluna  $j$ , designado por  $(p_d(i,j))$  corresponde à probabilidade de ocorrência de pares de pixels com os tons de cinza  $i$  e  $j$ . Estes pares de pixels se encontram separados por uma certa distância numa dada direção e sentido, determinada por um vetor  $d$ . O vetor  $d$  é definido por suas componentes  $d_x$  e  $d_y$ , como mostrado na Figura 3-10 [24].

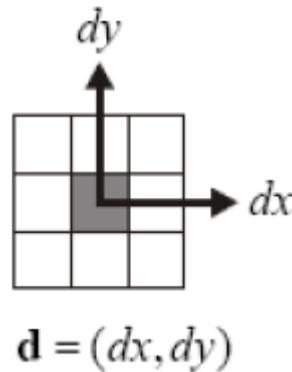


Figura 3-10: Representação gráfica do vetor  $d$  [24].

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

$$p_d(i,j) = \frac{P_d(i,j)}{R} \quad 3-4$$

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  $\alpha$ . Este ângulo só pode adquirir quatro valores possíveis ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ ) correspondentes aos pixels vizinhos mais próximos, isto é, os pixels adjacentes.

$$\alpha = 0^\circ \rightarrow d = (1,0) \quad 3-5$$

$$\alpha = 45^\circ \rightarrow d = (1,1) \quad 3-6$$

$$\alpha = 90^\circ \rightarrow d = (0,1) \quad 3-7$$

$$\alpha = 135^\circ \rightarrow d = (-1,1) \quad 3-8$$

Para que as medidas sejam independentes de rotação, em geral, é utilizada a média desses quatro (3-5, 3-6, 3-7, 3-8) parâmetros [24]. Porém, a escolha do tipo de medida deve ser feita dependendo do caso específico.

A Tabela 3-4 apresenta a lista dos 11 parâmetros de Haralick utilizados neste trabalho, em que cada parâmetro proporciona 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 ideia básica está em caracterizar o “conteúdo” da matriz de co-ocorrência através de diferentes descritores, sendo alguns deles correlatados entre si [6].

Tabela 3-4: Parâmetros de Haralick [6].

Parâmetro	Definição
1) Uniformidade	$f_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [p(i, j)]^2$
2) Contraste	$f_2 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 \cdot p(i, j)$
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	$f_6 = \sum_{k=2}^{2 \cdot N_g} k \cdot p_{x+y}(k)$
7) Variância da Soma	$f_7 = \sum_{k=2}^{2 \cdot N_g} (k - f_6)^2 \cdot p_{x+y}(k)$
8) Entropia da Soma	$f_8 = - \sum_{k=2}^{2 \cdot N_g} p_{x+y}(k) \cdot \log[p_{x+y}(k)]$
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	$f_{11} = - \sum_{k=0}^{N_g-1} p_{x-y}(k) \cdot \log[p_{x-y}(k)]$
Notação:	
$p(i, j) = p_d(i, j)$	$p_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{N_g} \sum_{j=1}^{N_g} p(i, j)$
$p_x(i) = \sum_{j=1}^{N_g} p(i, j)$	$p_{x+y}(k) = \sum_{\substack{i=1 \\  i-j=k }}^{N_g} \sum_{j=1}^{N_g} p(i, j)$
$p_y(j) = \sum_{i=1}^{N_g} p(i, j)$	$\mu_x$ e $\mu_y$ são as médias
$\sigma_x$ e $\sigma_y$ são, respectivamente, os desvios padrões de $p_x$ e $p_y$	

### 3.4.4.3. Crescimento de Regiões

Algoritmos de segmentação por crescimento de regiões agrupam pixels ou sub-regiões em regiões maiores, partindo de um conjunto de pontos iniciais (sementes) que crescem anexando pixels ou regiões adjacentes que possuam propriedades similares, como, por exemplo, textura, cor, etc [17].

Porém, o que define o tipo de crescimento de regiões é o critério de parada do algoritmo. Neste trabalho o critério de parada a ser empregado será a distância espectral entre dois pixels como parâmetro de cor.

Primeiramente vai ser desenvolvido o método tradicional de crescimento de regiões. A formulação básica adotada para este tipo de abordagem é dada considerando  $I$  como uma imagem onde a segmentação é a decomposição de  $I$  em  $n$  regiões  $R_1, R_2, \dots, R_n$ , de tal forma que:

$$I = \bigcup_{i=1}^n R_i \quad 3-9$$

$$R_i \text{ é uma região conexa, } 1, 2, \dots, n \quad 3-10$$

$$R_i \cap R_j = \emptyset \quad \forall i \neq j \quad 3-11$$

$$P(R_i) = \textit{verdadeiro} \quad \forall i \quad 3-12$$

$$P\left(R_i \cap R_j\right) = \textit{falso} \quad \forall i \neq j \quad 3-13$$

Pode existir um número de possíveis partições, mas a seleção de um conjunto adequado de regiões depende da escolha da propriedade  $P$  associada à região, ou seja, do predicado de uniformidade dos pixels da região [17].

A técnica pode-se resumir nas seguintes etapas:

- (i) Escolha dos pixels-sementes (pontos ou simplesmente sementes)
- (ii) Escolha do limiar que separará as regiões
- (iii) Crescimento das regiões

A escolha dos pixels-sementes geralmente é feita baseando-se na natureza do problema. A escolha destes pontos, é importante, pois as regiões crescerão ao redor deles.

$$\text{Se } |P_{(x,y)} - P_{(x_1,y_1)}| < t, \text{ então } P_{(x,y)} \in R_1 \text{ senão,} \quad 3-14$$

$$\text{Se } |P_{(x,y)} - P_{(x_2,y_2)}| < t, \text{ então } P_{(x,y)} \in R_2 \text{ senão,} \quad 3-15$$

...

$$\text{Se } |P_{(x,y)} - P_{(x_n,y_n)}| < t, \text{ então } P_{(x,y)} \in R_n; \quad 3-16$$

Onde  $P(x,y)$  é a intensidade de cinza em um ponto  $(x, y)$  da tabela de intensidades;  $n$  são as sementes;  $R_i$ , com  $1 \leq i \leq n$ , são as regiões e;  $t$  é o valor do limiar.

Para entender bem o método, vai ser utilizado um exemplo. Neste exemplo trabalha-se com dois cenários: o primeiro com duas sementes com os valores mínimo e máximo de tom de cinza da imagem. Já no segundo cenário será utilizada uma terceira semente com o valor médio dos tons de cinza da imagem.

Na Figura 3-11(a) pode-se observar uma imagem em tons de cinza onde são tomados como sementes os pontos  $P_{(1,1)} = 0$  e  $P_{(2,3)} = 7$ . Como se pode observar, estas sementes representam os extremos da faixa tonal da imagem, isso combinado com um limiar  $t=3$  faz com que alguns pixels não sejam incluídos em nenhuma das duas regiões possíveis ( $R_1$  e  $R_2$ ) Figura 3-11(b). Contudo, se for tomada uma terceira semente  $P_{(4,4)} = 3$  e for mantido o mesmo limiar  $t=3$ , seria criada uma terceira região  $R_3$  incluindo os pixels antes rejeitados Figura 3-11(c). É interessante também notar que, se for escolhido um valor de limiar mais alto, por exemplo  $t=8$ , só haveria uma região no exemplo em questão. Esse exemplo serve para mostrar a importância de uma escolha adequada dos limiares, bem como dos pixels-sementes para ter sucesso no método de crescimento de regiões.

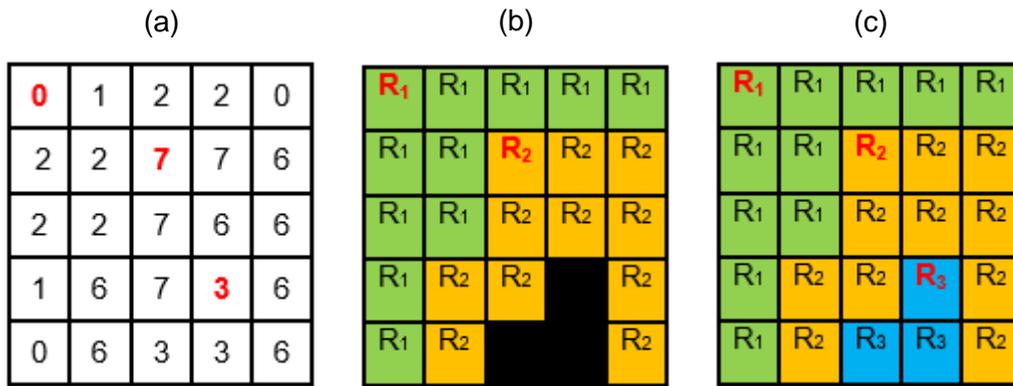


Figura 3-11: Exemplo de crescimento de regiões com  $t=3$ : (a) Imagem em tons de cinza; (b) Duas regiões ( $R_1$  e  $R_2$ ) separadas com  $n=2$  sementes (em vermelho); (c) Três regiões ( $R_1$ ,  $R_2$ , e  $R_3$ ) separadas com  $n=3$  sementes (em vermelho) [17].

### 3.4.5. Espaço de Medidas

Nesta etapa se inicia a análise da imagem, são realizadas medidas sobre a imagem segmentada e/ou sobre a imagem em tons de cinza, extraindo-se atributos característicos, gerando dados quantitativos para o objetivo final. Existem duas classes de medidas, as medidas de campo e as medidas de região.

As medidas de campo, são as medidas feitas na imagem como um todo, a fim de caracterizá-la integralmente. Podem ser feitas também em sub- campos da imagem, caracterizando-os individualmente como imagens diferentes. Algumas das medidas de campo usadas com maior frequência são:

- Contagem de objetos;
- Área total de objetos;
- Fração de área.

A contagem de objetos em uma imagem é uma das medidas mais fáceis de obter digitalmente. A partir da imagem segmentada, as regiões de pixels contíguos com a mesma tonalidade, que correspondem aos objetos, são contadas [24].

A área total de objetos é também uma medida fácil e rápida para o computador, devido a que ela é obtida pela simples contagem dos pixels brancos na imagem binária. Esta área serve para calcular outras medidas importantes, como a fração de área do campo ocupada pelos objetos e a área média dos objetos.

A fração de área é calculada pela razão entre o número de pixels brancos (área total dos objetos) e o número total de pixels (área da imagem) na imagem binária. Ou seja, trata-se de uma grandeza adimensional com valores entre 0 e 1.

Finalmente, a área média dos objetos é calculada pela razão entre a área e o número de objetos.

As medidas de região, são medidas realizadas de forma independente sobre os objetos segmentados da imagem. As medidas de região usadas com maior frequência são:

- Razão de Aspectos;
- Fatores de Forma Circular;
- Convexidade;
- Solidez.

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 Ferets. Os Ferets máximos e mínimos equivalem às projeções máximas e mínimas do objeto, caracterizando assim sua dimensão externa. A razão de aspectos é obtida pela razão entre o Feret mínimo e Feret máximo, como é apresentada em 3-17:

$$RA = \frac{F_{min}}{F_{max}} \quad 3-17$$

Onde Fmin é o Feret mínimo e Fmax é o Feret máximo.

Por sua vez, existem vários tipos de Fatores de Forma Circular (FFC). Estes fatores de forma circular podem ser definidos em função de sua área, perímetro, e/ou Feret máximo, segundo as equações 3-18, 3-19, 3-20 e 3-21 abaixo.

Como pode-se observar, a equação 3-18 parte da relação  $P^2=4\pi A$ , que existe entre o perímetro (P) e a área (A) de um círculo. Esta relação só será uma igualdade para o caso do círculo, já que para qualquer outra geometria do objeto teremos que  $P^2>4\pi A$ .

$$FFC = \frac{4\pi A}{P^2} \quad 3-18$$

Caso se deseje um fator de forma circular mais sensível ao alongamento do objeto, então basta substituir  $P^2$  por  $\pi(F_{max})^2$ , como:

$$FFC_c = \frac{4A}{\pi(F_{max})^2} \quad 3-19$$

Onde A é a área preenchida e Fmax é o Feret máximo. Este novo fator de forma circular é chamado de circularidade (FFCC).

Outras variantes do fator de forma circular são:

$$FFC_m = \frac{4A}{P(F_{max})} \quad 3-20$$

$$FFC_g = \frac{16A^2}{\pi P(F_{max})^3} \quad 3-21$$

Onde  $FFC_m$  é chamado de fator de forma circular modificado e  $FFC_g$  é chamado de fator de forma circular Grum.

As medidas de convexidade servem para descrever se um objeto é convexo ou não convexo. Para isso, existem dois parâmetros para definir a convexidade dos objetos, através da convexidade (C) e da solidez (S), de modo que:

$$C = \frac{P_c}{P} \quad 3-22$$

$$S = \frac{A}{A_c} \quad 3-23$$

Onde  $P_c$  e  $A_c$  são o perímetro e a área convexa, respectivamente.

As medidas de campo e as medidas de região empregam o conceito de pixel contíguo. Os analisadores de imagem geralmente oferecem duas opções de contiguidade entre os pixels, também denominada conectividade: conectividade 4 e conectividade 8 [34].

A conectividade 4 estabelece que somente os 4 pixels adjacentes vertical e horizontal, são considerados contíguos, sem levar em conta os pixels adjacentes da diagonal. Isto significa que se um pixel estiver conectado a outro pixel adjacente pela 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 [23].

A Figura 3-12a apresenta um exemplo de uma pequena imagem de 10x10 pixels, onde o pixel circulado em vermelho é o pixel analisado. A Figura 3-12b e a Figura 3-12c 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 3-12b, o pixel só 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 3-12c, o pixel em questão é vizinho do outro pixel dentro da região de contiguidade, indicando que compõem um mesmo objeto.

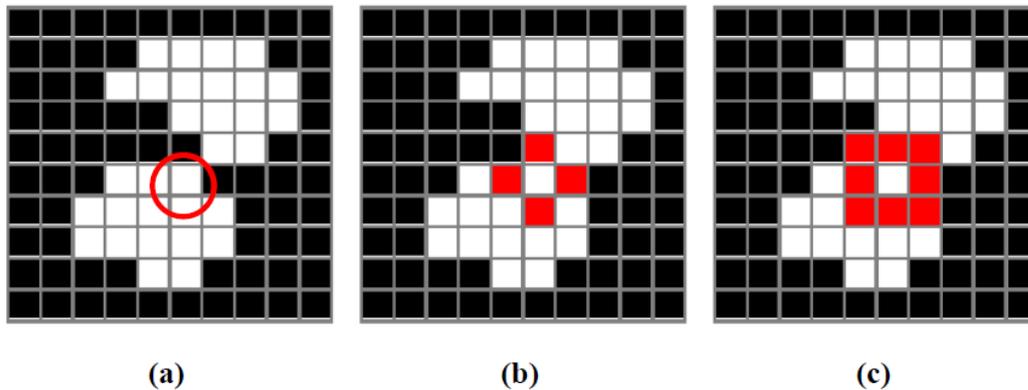


Figura 3-12 Conectividade entre os pixels: (a) Imagem original, onde o pixel circulado é o pixel analisado; (b) Conectividade 4, considerando que são dois objetos separados; (c) Conectividade 8, considerando que os pixels fazem parte de um mesmo objeto [6].

### 3.4.6. Reconhecimento de Padrões e Classificação

A etapa de reconhecimento de padrões e classificação é a etapa final da sequência padrão do PADI.

O reconhecimento pode-se acontecer por identidade ou por semelhança. No reconhecimento por identidade um objeto, previamente conhecido, é identificado. Por outro lado, o reconhecimento por semelhança ocorre quando um objeto é identificado como membro de uma classe através de traços característicos. Deste modo, o reconhecimento pode ser visto como um processo de classificação [9].

As técnicas de reconhecimento de padrões são usadas para classificar objetos através de um conjunto de propriedades ou características comuns a cada classe de objetos [35].

Um padrão pode ser descrito matematicamente como um vetor cujas componentes são características numéricas dos objetos de interesse, as quais são obtidas por meio de um conjunto de observações [36] [37].

O problema de reconhecimento é colocado como uma tarefa de classificação ou categorização dos padrões. Por outro lado, as classes são definidas como um grupo de padrões que compartilham algumas propriedades em comum. Estas classes podem ser identificadas como regiões do espaço de características.

Um espaço de características é construído a partir dos atributos dos objetos, 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.

Os métodos de reconhecimento de padrões podem ser agrupados em duas grandes categorias: supervisionada e não-supervisionada.

- (i) Na classificação supervisionada um conjunto de objetos conhecidos pertencentes a diferentes classes é analisado, escolhe-se os parâmetros ideais para separação, e estabelece-se a faixa de valores dos parâmetros para cada classe. A partir deste conjunto de treinamento, encontra-se a função discriminante que separa as diversas classes [22]. Objetos desconhecidos são fornecidos ao classificador tal como é apresentado na Figura 3-13.
- (ii) Na classificação não-supervisionada o sistema recebe objetos desconhecidos e a partir da medida de diferentes parâmetros tenta encontrar as classes. A identificação de classes pode ser realizada a partir da identificação de agrupamentos (clusters) de objetos no espaço de parâmetros [22].

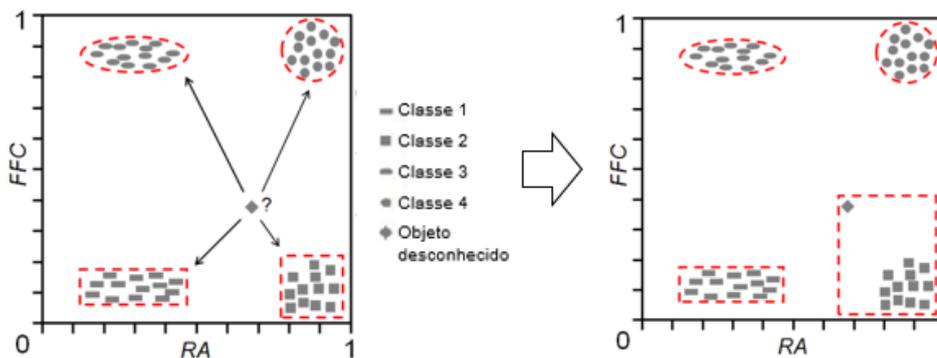


Figura 3-13: Classificação supervisionada utilizando os parâmetros característicos FFC e RA [17].

Este trabalho considera só a classificação supervisionada.

### **3.4.6.1. Classificação Supervisionada**

A classificação supervisionada demanda:

- Um conjunto de treinamento, conjunto de objetos padrões conhecidos, rotulados, para representar as classes.
- Um conjunto de características, conjunto de atributos que define o espaço de características.
- Um classificador, que é o algoritmo que realiza as medidas de similaridade.

#### **3.4.6.1.1. Conjunto de treinamento**

O conjunto de treinamento é um conjunto constituído por objetos padrões conhecidos rotulados pelo usuário. Ele deve ser grande o suficiente para ter uma boa estatística, com o intuito de aumentar a precisão do classificador e deve apresentar grande variabilidade dentro de cada classe para representar bem as diversas classes.

#### **3.4.6.1.2. Conjunto de características**

A escolha do conjunto de características é feita a partir da observação dos atributos que melhor representam as classes no conjunto de treinamento. Deve ser um conjunto de atributos que caracterize bem os objetos, agrupando os objetos similares e separando os distintos. No entanto, o conjunto de características precisa ser robusto para agrupar as diferenças dentro de cada classe e sensível para discriminar essas classes. Assim, os atributos do conjunto de características são escolhidos visando a separação das classes para possibilitar a classificação.

#### **3.4.6.1.3. Classificador**

O classificador é a função que calcula as medidas de similaridade entre os objetos no espaço de características. Ademais os classificadores são funções que utilizam como entrada os padrões desconhecidos, e como saída as classes a que estes padrões provavelmente pertencem [27].