

## 2

### Revisão Bibliográfica

Neste capítulo, são apresentadas primeiramente a importância econômica, formação, composição e estrutura interna do material a ser estudado. A seguir são revisadas as técnicas de MD, relevantes no contexto deste trabalho, no âmbito da microscopia ótica de luz refletida e da análise de imagens.

#### 2.1

##### Minério de Ferro

O ferro é um dos elementos mais abundantes da crosta terrestre (o segundo metal, com aproximadamente 4,2%) e matéria-prima fundamental para a indústria.

O maior desenvolvimento de formação ferrífera ocorreu no período Pré-cambriano tardio da era Proterozóica. Estas formações ocorreram nas rochas sedimentares do tipo Formações Ferríferas Bandadas (BIF - *Banded Iron Formation*).[14, 15] As BIF estão compostas por grandes depósitos de camadas ricas em ferro, alternando com camadas ricas em silício. Elas representam os maiores depósitos de minério de ferro do planeta.[16]

Os maiores depósitos de minério de ferro, no Brasil, são praticamente todos do tipo hematítico, com altos teores de ferro.[17] As reservas medidas e indicadas de minério de ferro no Brasil alcançam 29 bilhões de toneladas (Tabela 2.1), situando o país em segundo lugar em relação às reservas mundiais, de 166 bilhões de toneladas.[18]

Entretanto, em termos de ferro contido, as reservas brasileiras têm teor médio muito alto (aproximadamente 55%) e, embora continue em segundo, fica muito mais próximo do primeiro lugar internacional. Esse fato ocorre devido ao alto teor de ferro encontrado nas hematitas (60% a 67% de ferro), predominante no Pará, e itabirito (50% a 60% de ferro), predominante em Minas Gerais.[19]

Se também são consideradas as reservas inferidas, o Brasil aumenta significativamente o seu potencial, totalizando 73,7 bilhões de toneladas de minério de ferro.[20] Por sua vez, se for considerado que a geologia do território brasileiro é pouco conhecida, quando comparada com a quantidade e qualidade

Tabela 2.1: Reservas mundiais de minério de ferro no ano 2011 ( $10^6$  t).[18]

Países	RM <sup>†</sup>	RM [%]	RC	RC [%]
Austrália	35000	21,08	17000	21,25
Brasil	29000	17,47	16000	20,00
Rússia	25000	15,06	14000	17,50
China	23000	13,86	7200	9,00
Índia	7000	4,22	4500	5,62
Estados Unidos	6900	4,16	2100	2,62
Canadá	6300	3,80	2300	2,88
Ucrânia	6000	3,61	2100	2,62
Venezuela	4000	2,41	2400	3,00
Suécia	3500	2,11	2200	2,75
Cazaquistão	3000	1,81	1000	1,25
Irã	2500	1,51	1400	1,75
Mauritânia	1100	0,66	700	0,88
África do Sul	1000	0,60	650	0,81
México	700	0,42	400	0,50
Outros países	12000	7,23	6000	7,50
<b>TOTAL</b>	166000	100,00	80000	100,00

<sup>†</sup> RM: Reservas Medidas; RC: Reservas Contidas

dos estudos geológicos de outros países com tradição mineral como o Canadá, a Austrália e os Estados Unidos, o Brasil bem poderia estar em uma posição ainda melhor.

Como pode ser observado na Tabela 2.2, as reservas brasileiras estão assim distribuídas:

- (i) Minas Gerais - 67,0%;
- (ii) Pará - 16,0%;
- (iii) Mato Grosso do Sul - 15,5% e;
- (iv) Outros estados (Alagoas, Amazonas, Bahia, Ceará, Goiás, Pernambuco, Rio Grande do Norte e São Paulo) - 1,5%.

Entre todos os metais, o ferro é o mais produzido e o que está mais presente em nossa vida. Cerca de 99,0% do minério de ferro produzido é utilizado na fabricação de aço e ferro fundido. Uma vez que o ferro e o aço estão presentes nos principais setores da indústria moderna, o consumo de minério

Tabela 2.2: Reservas brasileiras de minério de ferro Medidas e Indicadas (em toneladas).[20]

Lugar	Unidade da Federação	Reservas	Reservas [%]	Teor [%]
1	Minas Gerais	19359905311	66,968	51,53
2	Pará	4616877438	15,970	67,37
3	Mato Grosso do Sul	4472348567	15,470	55,09
4	São Paulo	344577533	1,192	31,91
5	Amazonas	71933809	0,249	65,92
6	Ceará	25677321	0,089	35,69
7	Pernambuco	8942804	0,031	60,62
8	Goiás	4269208	0,015	50,00
9	Bahia	2046658	0,007	56,00
10	Distrito Federal	1191610	0,004	50,00
11	Rio Grande do Norte	1086925	0,004	57,91
12	Alagoas	209005	0,001	54,95
<b>TOTAL</b>	————	28909066189	100,000	54,89 <sup>†</sup>

<sup>†</sup> Este valor corresponde à média ponderada.

de ferro é um dos índices considerados na medida de industrialização de um país.[21, 22]

Observando a evolução da produção mundial de aço (Figura 2.1) a partir da década de cinquenta do século XX, um crescimento na produção é notado, seguido, a partir da década de setenta, por uma estabilização. Já no fim dos anos 90, o fator China na economia mundial fez com que a produção de aço fosse puxada novamente para cima em um ritmo vertiginoso. Ávida por *commodities* para sustentar seu crescimento, a China continuará a demandar muito minério de ferro, pois o consumo ainda é crescente e assim deverá permanecer por muitos anos.[23]

A dependência chinesa de recursos naturais externos é de fato a principal barreira a seu crescimento no longo prazo, e o Brasil representa uma fonte segura de abastecimento aos chineses nas próximas décadas. Segundo dados do Conselho Empresarial Brasil-China, entre 1989 e 2004 a China foi responsável por 56% do crescimento do consumo de minério de ferro em todo o mundo. Em 2010, a China importou quase 60% das exportações mundiais de minério de ferro total e produziu cerca de 60% de ferro gusa do mundo.[24]

O Brasil é o segundo maior produtor de minério de ferro, conforme a Conferência das Nações Unidas para o Comércio e o Desenvolvimento (UNCTAD). Em 2010, o país produziu 372 milhões de toneladas, o que

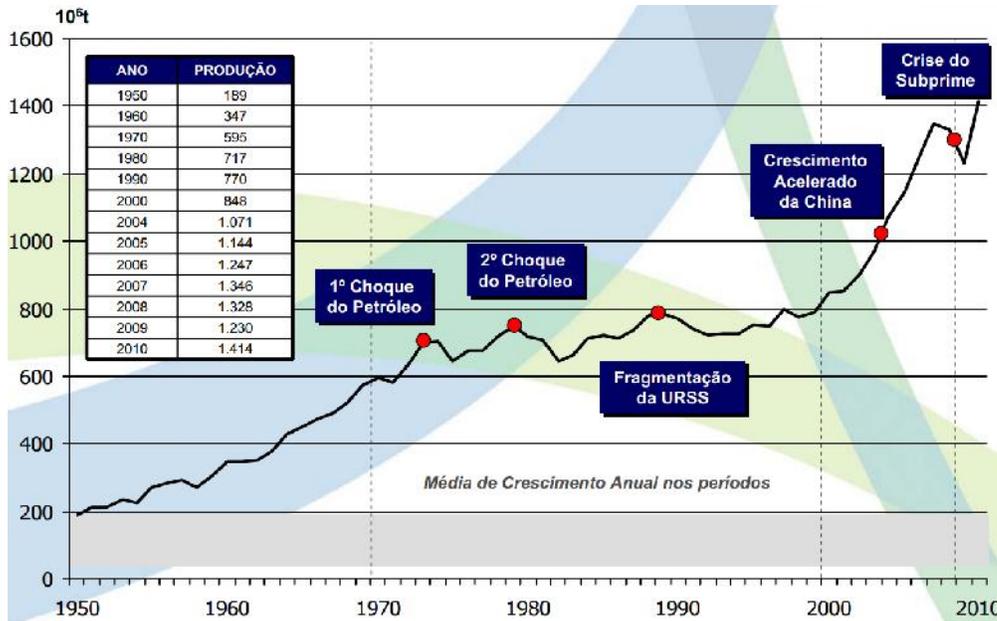


Figura 2.1: Evolução da Produção Mundial de Aço Bruto no período 1950 - 2010.[23]

corresponde a 15,5% da produção mundial (Tabela 2.3).[19] As vantagens das abundantes reservas brasileiras devem-se a suas características naturais, cujas jazidas são de ótima qualidade e de fácil extração, produzindo grandes volumes a baixos custos.

Tabela 2.3: Produção Mineral (em milhões de tons/ano).[19]

Ano	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
PM†	1060	1060	1080	1060	1340	1540	1712	1900	2200	2240	2400
PB	212	237	214	264	262	278	317	350	351	331	372
PB [%]	20,0	22,0	19,8	22,7	19,5	18,0	18,5	18,4	15,9	14,8	15,5
CM	2°	2°	2°	2°	2°	2°	2°	2°	2°	2°	2°
EB	157	156	167	175	211	224	242	269	282	266	311
EB [%]	74,0	65,8	78,0	66,3	80,5	80,6	76,3	76,8	80,3	80,4	83,6

† PM: Produção Mundial; PB: Produção Brasileira; CM: Colocação Mundial; EB: Exportação Brasileira

Os minérios 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. Consequentemente, dentro de uma mesma amostragem, poderão ocorrer minérios de diferentes características mineralógicas e microestruturais.[10, 25, 26]

O ferro é encontrado em diversos minerais, mas apenas alguns são economicamente viáveis como fontes deste elemento, sendo os óxidos os mais importantes.[27, 28] Os principais tipos de texturas de óxidos e hidróxidos de ferro e suas características são apresentadas na Tabela 2.4.[29]

### 2.1.1

#### Hematita

A hematita é o mineral de ferro mais importante devido a sua alta ocorrência em vários tipos de rochas e suas origens diversas.[30] A composição química deste mineral é  $\text{Fe}_2\text{O}_3$ , com uma fração mássica em ferro de 69,9% e uma fração mássica em oxigênio de 30,1%.[31]

Este mineral pode ser encontrado com diferentes morfologias (Tabela 2.4), 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.[2, 29] Estas morfologias estão relacionadas a uma condição de cristalização diferenciada, que poderá ter influência no seu processamento.[25]

O nome hematita vem do grego *haima* que significa sangue. Como pode ser observado na Figura 2.2a e Figura 2.2b a cor da hematita varia do vermelho-sangue até o preto com forte brilho metálico. Por sua vez, na Figura 2.2c, pode ser observado que seu arranjo atômico é trigonal-hexagonal ( $\bar{3}2/m$ ), com parâmetros de  $a = 5,034 \text{ \AA}$  e de  $c = 13,752 \text{ \AA}$  e grupo espacial  $R\bar{3}c$ . [4, 31] Esta estrutura hexagonal pode ser observada facilmente no mineral como mostra a Figura 2.2b.

A hematita também pode ser indexada no sistema romboédrico, cuja cela unitária seria  $a_{tr} = 5,427 \text{ \AA}$  e  $\alpha = 55,3^\circ$ , de tal forma que existem camadas de oxigênio rodeando átomos de ferro formando octaedros ligados pelas faces, conforme mostrado na Figura 2.2d. Esta estrutura consiste de um empacotamento denso de planos de oxigênio empilhados na repetição *ABAB* ao longo da direção [001] (planos de íons paralelos ao plano (001)) com os átomos de *Fe* ocupando as posições octaédricas. Dois terços dos sítios octaédricos estão ocupados por íons  $\text{Fe}^{3+}$  que são arranjados regularmente com dois sítios preenchidos seguidos por um sítio vacante na direção [001] formando anéis sêxtuplos.[4, 32]

O arranjo de cátions produz pares de octaedros  $\text{Fe}(O)_6$ . Cada octaedro compartilha uma face com um octaedro da camada adjacente. Assim, a estrutura real de hematita é distorcida devido à repulsão dos átomos de Fe presentes nos octaedros que partilham a face. Os átomos de Fe, então, se situam mais próximos das faces não partilhadas (Figura 2.2d). A estrutura

Tabela 2.4: Quadro ilustrativo com as principais morfologias de cristais de óxidos/hidróxidos de ferro.[29]

Tipo	Características	Forma Textura	Ilustração Esquemática
Hematita Microcristalina	<ul style="list-style-type: none"> <li>▷ Cristais muito pequenos, &lt; 0.01 mm.</li> <li>▷ Textura porosa.</li> <li>▷ Contatos pouco desenvolvidos.</li> </ul>		
Magnetita	<ul style="list-style-type: none"> <li>▷ Cristais euédricos isolados ou em agregados.</li> <li>▷ Cristais compactos.</li> </ul>		
Martita	<ul style="list-style-type: none"> <li>▷ Hematita com hábito de magnetita.</li> <li>▷ Oxidação segundo os planos cristalográficos da magnetita.</li> <li>▷ Geralmente porosa.</li> </ul>		
Hematita Lobular	<ul style="list-style-type: none"> <li>▷ Formatos irregulares inequidimensionais.</li> <li>▷ Contatos irregulares, geralmente imbricados.</li> </ul>		
Hematita Granular	<ul style="list-style-type: none"> <li>▷ Formatos regulares equidimensionais.</li> <li>▷ Contatos retilíneos e junções triplíceis.</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>		
Hidróxidos de Fe (Goethita-Limonita)	<ul style="list-style-type: none"> <li>▷ Material cripto-cristalino.</li> <li>▷ Estrutura coloforme, hábito botrioidal.</li> <li>▷ Textura porosa.</li> </ul>		

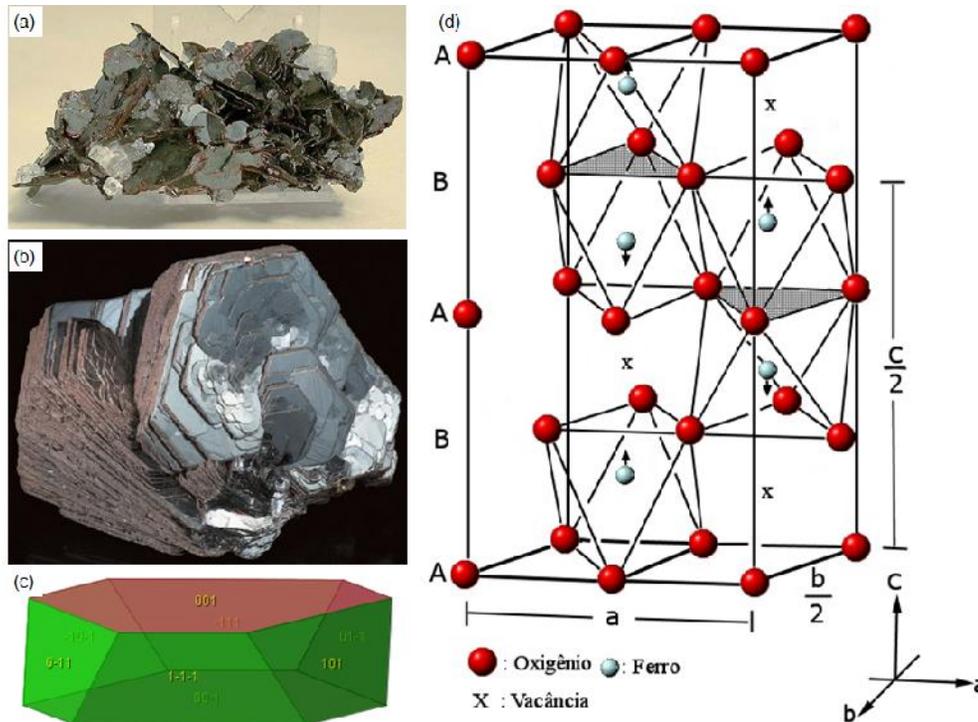


Figura 2.2: Hematita: (a) e (b) Imagens de fragmentos do mineral; (c) Estrutura trigonal-hexagonal visualizando os planos; (d) Modelo de bolas e varetas da cela unitária.[4, 31]

tripla  $Fe - O_3 - Fe$  influencia as propriedades magnéticas da hematita que, diferente da magnetita, é um óxido fracamente ferromagnético à temperatura ambiente.[4, 33]

Existem relacionamentos estruturais entre certos planos na estrutura da hematita e planos em outros óxidos de ferro, especialmente com a magnetita e a goethita (Tabela 2.5). Há, por exemplo, um relacionamento entre o plano (111) da magnetita e o plano (001) da hematita. Com isso algumas vezes pode-se observar nucleação e crescimento de magnetita no plano (001) da hematita. Do mesmo modo, como resultado da afinidade estrutural entre os planos (100) da goethita e o plano (003) da hematita, pode ocorrer um crescimento epitaxial de goethita na hematita.[4]

A hematita é um material opaco no MLR, porém em 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.[34] Contudo, a propriedade ótica mais importante deste mineral para o presente trabalho é sua birrefletância devido à forte anisotropia que ele apresenta. Assim, sua reflexão e, conseqüentemente, o seu brilho na imagem mudam com diferentes orientações dos

Tabela 2.5: Orientações estruturais entre os óxidos de ferro.[4]

Par de Óxidos	Fórmula Química	Plano Cristalográfico	Direção Cristalográfica
Goethita	$FeO(OH)$	(100)(004)(200)	[100]
Hematita	$\alpha - Fe_2O_3$	(003)(110)(100)	[100]
Hematita	$\alpha - Fe_2O_3$	(001)	[100]
Magnetita	$Fe_3O_4$	(111)	[110]
Lepidocrocita	$FeO(OH)$	(100)	[001][051]
Maghemita	$\gamma - Fe_2O_3$	(001)	[110][111]

cristais com relação ao plano de incidência da luz.[5] Este efeito ainda pode ser ampliado quando é usada luz polarizada no MLR.

## 2.2

### Microscopia Digital

Podemos definir a MD como a integração entre o microscópio e o computador, oferecendo automação do microscópio, aquisição e análise digital de imagens. Esta área, além de permitir a automação, fornece novos recursos para a caracterização microestrutural.[35]

Alguns sistemas totalmente controlados por software e com ambiente de programação permitem uma automação completa.[36] Exemplos disto são:

- ▷ Uso de rotinas para o controle da platina motorizada:[37]
- ▷ Troca de lentes;
- ▷ Correção de defeitos na aquisição:[38, 39]
- ▷ Autofoco:[40, 41]
- ▷ Foco estendido:[42, 43]
- ▷ Ajuste automático da iluminação e cor:[44, 45]
- ▷ Controle dos filtros e dos diafragmas;
- ▷ Varredura da amostra com aquisição automática de imagens e;
- ▷ Obtenção de mosaico ou campo estendido.[46]

Algumas destas rotinas vêm incluídas em softwares comerciais de aquisição e análise de imagens, como o KS400, o AxioVision e o SIS, apenas para nomear alguns.[47, 48, 49] O Laboratório de MD da PUC-Rio (LMD)

conta com dois microscópios óticos motorizados, Zeiss AxioPlan 2 e Zeiss AxioImager.M2m, com câmeras digitais, Zeiss AxioCam HRc e Zeiss AxioCam MRc5 respectivamente, automatizados com o software AxioVision.[50]

O desenvolvimento de câmeras coloridas CCD (*Charge-Coupled Device*) e a integração do controle por computador do conjunto câmera/microscópio, facilitou a aquisição automática de imagens.[35, 51] Entenda-se por aquisição automática de imagem uma rotina que faz uma varredura da amostra, capturando imagens de campos adjacentes ou espaçados, de forma automática. Esse tipo de rotina foi implementada em softwares comerciais de aquisição e análise de imagens. Além da maior velocidade e praticidade, o método automático é reprodutível e evita a ocorrência de erros por fadiga do operador, como a repetição e a sobreposição de campos.[52]

A MD, além de permitir certo grau de automação, abriu novas possibilidades para a caracterização microestrutural.[35]

Várias técnicas quantitativas já vinham sendo utilizadas na caracterização de amostras de minério de ferro.[11] Entre estas técnicas encontra-se a difratometria de raios X, a espectroscopia Mössbauer, a análise de imagens obtidas tanto por microscopia ótica como eletrônica, entre outras.[6, 7, 8, 32, 53, 54, 55, 56, 57, 58, 59, 60, 61]

A microscopia ótica de luz transmitida, para minerais transparentes, e de luz refletida, para minerais opacos, é provavelmente o método mais tradicional de identificação mineralógica.[62] Esta técnica é uma ferramenta adequada, uma vez que permite obter dados sobre a porosidade, associações minerais, liberação mineral, forma das partículas, distribuição por tamanho e textura, entre outras.[30]

Recentemente, alguns autores aplicaram uma técnica de análise de textura em imagens adquiridas com luz polarizada para classificar fases de minério de ferro.[58] Embora com algumas limitações, a metodologia desenvolvida por eles foi promissora na identificação de cristais de hematita. Esta metodologia se baseia no processamento de imagens para determinar fronteiras de cristais de hematita onde um conjunto de sete imagens, por campo, é adquirido girando o polarizador em pequenos intervalos angulares.

Outros autores também fizeram uso da luz polarizada na identificação e análise de minerais e rochas.[30, 63, 64, 65, 66] É de particular importância o fato de que este tipo de iluminação produz contraste entre cristais com orientações diferentes no espaço. Isto possibilita a visualização dos cristais, geralmente impossíveis de serem vistos com iluminação convencional em campo claro (Figura 2.3).

A Figura 2.3 mostra duas imagens de um mesmo campo de uma amos-

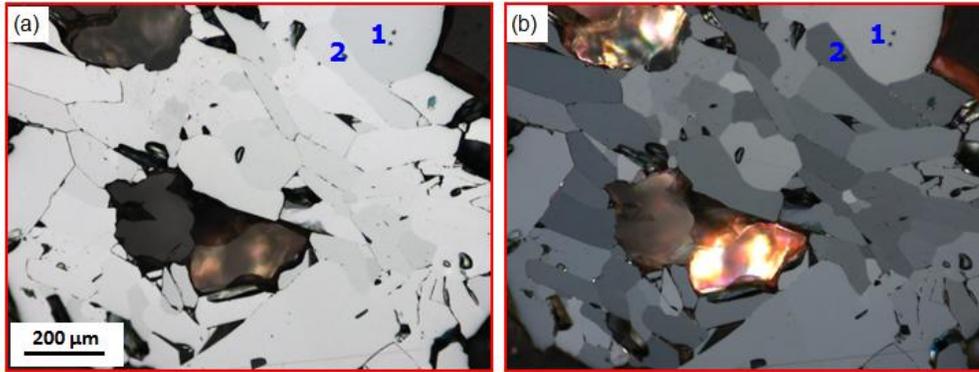


Figura 2.3: Variação de brilho entre dois cristais adjacentes: (a) Imagem em campo claro; (b) Imagem com polarização de um mesmo campo.[67]

tra de minério de ferro adquiridas no MLR, a primeira em campo claro (Figura 2.3a) e a segunda sob luz polarizada (Figura 2.3b). Na imagem sem polarização dois cristais adjacentes são indicados, em azul, como 1 e 2. Neste caso quase não pode ser observada uma diferença de brilho, embora na imagem com polarização a diferença de brilho entre esses mesmos cristais seja muito alta. Assim, fica evidente a importância da luz polarizada, no âmbito da microscopia ótica da luz refletida, para identificar cristais de hematita em minérios de ferro.

### 2.2.1 Polarização da Luz

A luz pode ser tratada como uma onda eletromagnética transversal, ou seja, os vetores intensidade do campo elétrico ( $\vec{E}$ ) e intensidade do campo magnético ( $\vec{H}$ ) são ortogonais entre si. A luz vibra não só num plano, mas em todos os planos simultaneamente, portanto há planos de ondas em todos os ângulos (Figura 2.4). Este tipo de luz é conhecida como luz natural ou luz não polarizada.[68]

A luz polarizada, diferente da luz natural, vibra num plano só (plano de polarização) e o os vetores  $\vec{E}$  e  $\vec{H}$  são normais à direção de propagação como mostra a Figura 2.5. Isto pode ser demonstrado através das equações de Maxwell onde, após algumas considerações e simplificações, pode-se chegar às seguintes equações de ondas:

$$\nabla^2 \vec{E} = \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2} \quad (2-1)$$

$$\nabla^2 \vec{H} = \mu_0 \epsilon_0 \frac{\partial^2 \vec{H}}{\partial t^2} \quad (2-2)$$

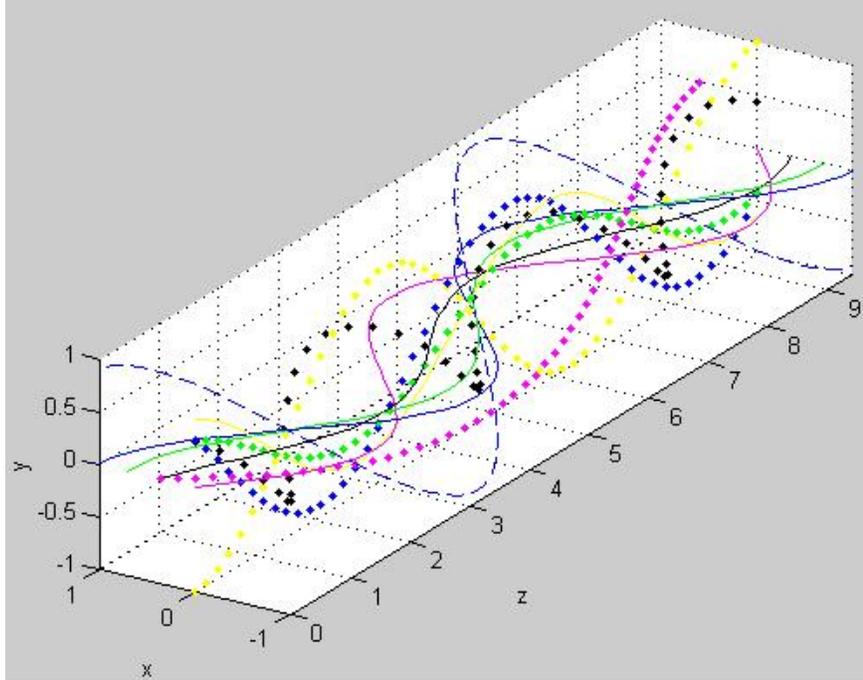


Figura 2.4: Luz natural com os diferentes comprimentos de onda correspondentes a cada cor. Neste caso só foi representada a intensidade do campo elétrico para facilitar a visualização.[69]

sendo que  $\mu_0$  é a permeabilidade magnética do vácuo e  $\epsilon_0$  a permeabilidade elétrica do vácuo.

As soluções das equações (2-1) e (2-2) seriam:

$$\vec{E} = \vec{E}_x + \vec{E}_y \quad (2-3)$$

$$\vec{H} = \vec{H}_x + \vec{H}_y \quad (2-4)$$

$$\vec{E} = E_{0x}\hat{x} \exp[i(\vec{k}\vec{z} - wt)] + E_{0y}\hat{y} \exp[i(\vec{k}\vec{z} - wt)] \quad (2-5)$$

$$\vec{H} = H_{0x}\hat{x} \exp[i(\vec{k}\vec{z} - wt)] + H_{0y}\hat{y} \exp[i(\vec{k}\vec{z} - wt)] \quad (2-6)$$

$$\vec{E} = (E_{0x}\hat{x} + E_{0y}\hat{y}) \exp[i(\vec{k}\vec{z} - wt)] \quad (2-7)$$

tomando a parte real do vetor, temos que:

$$\vec{E} = (E_{0x}\hat{x} + E_{0y}\hat{y}) \cos(\vec{k}\vec{z} - wt) \quad (2-8)$$

$$\vec{E} = \vec{E}_0 \cos(\vec{k}\vec{z} - wt) \quad (2-9)$$

$$\vec{H} = (H_{0x}\hat{x} + H_{0y}\hat{y}) \cos(\vec{k}\vec{z} - wt) \quad (2-10)$$

$$\vec{H} = \vec{H}_0 \cos(\vec{k}\vec{z} - wt) \quad (2-11)$$

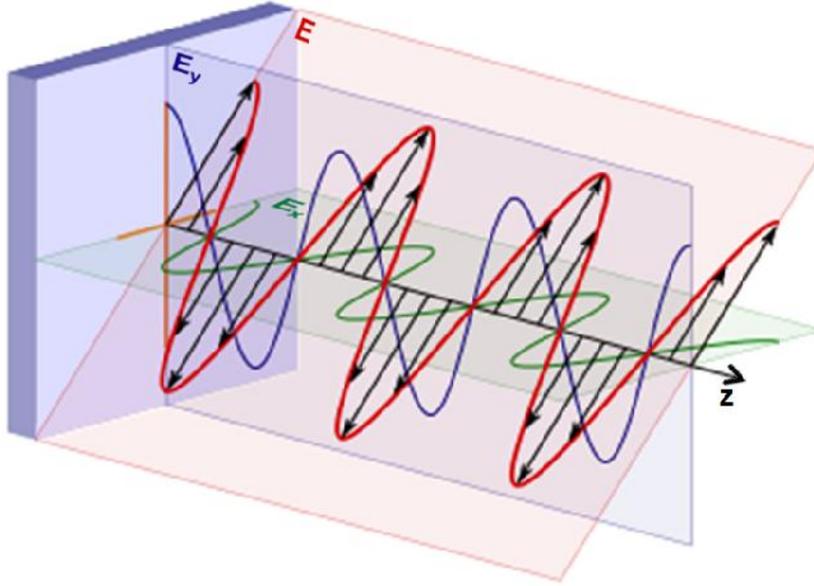


Figura 2.5: Propagação de uma onda plana linearmente polarizada. Neste caso só foi representada a intensidade do campo elétrico para facilitar a visualização.[70]

onde  $\frac{w}{k} = c = \frac{1}{\sqrt{\mu_0\epsilon_0}}$  ( $c$  é a velocidade da luz no vácuo,  $k$  o módulo do vetor de onda,  $w$  a frequência e  $t$  o tempo). Portanto, a onda resultante tem uma amplitude fixa igual a  $\vec{E}_0 = E_{0x}\hat{x} + E_{0y}\hat{y}$  para a intensidade do campo elétrico e  $\vec{H}_0 = H_{0x}\hat{x} + H_{0y}\hat{y}$  para a intensidade do campo magnético ou seja, ela também é linearmente polarizada. Assim, a onda resultante  $\vec{E}$  (e conseqüentemente  $\vec{H}$ ) oscila ao longo de um plano inclinado, como mostra a Figura 2.5, segundo uma cossenoide no tempo.

A luz, além de ser linearmente polarizada, também pode ser circular e elípticamente polarizada. A polarização circular é um caso particular da polarização elíptica porém, esta última, não será detalhada por não ter sido utilizada neste trabalho.

No caso da polarização linear, a projeção do vetor  $\vec{E}$  sobre o plano  $xy$  descreve um segmento de reta. No entanto, quando ambas ondas constitutivas têm a mesma amplitude (ou seja,  $E_{0x} = E_{0y} = E_0$  e conseqüentemente  $H_{0x} = H_{0y} = H_0$ ) a projeção será uma circunferência. Neste tipo de polarização, a soma de dois campos  $\vec{E}_x$  e  $\vec{E}_y$  (e conseqüentemente de  $\vec{H}_x$  e  $\vec{H}_y$ ) se propaga na direção  $z$  com a mesma frequência e vetor de onda, porém com uma defasagem de  $\phi = \pi/2$  (Figura 2.6).

Neste caso, a solução das equações (2-1) e (2-2) seria:

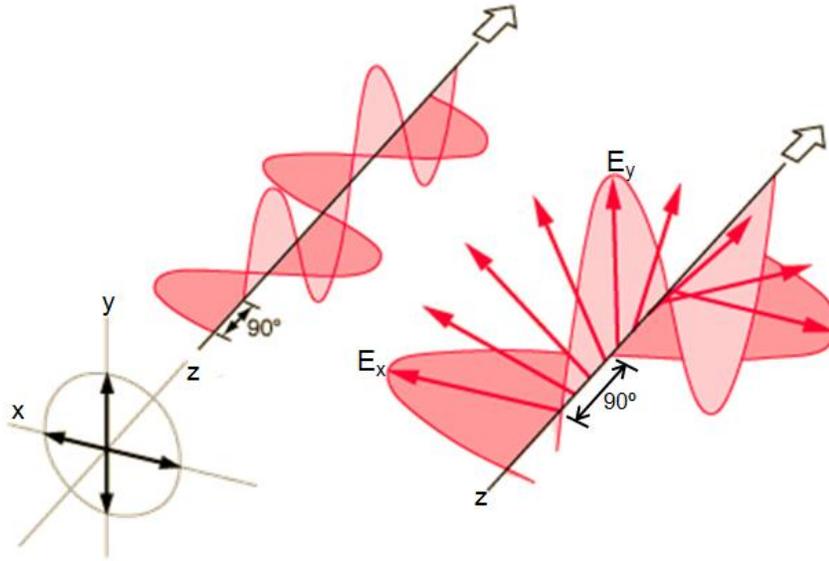


Figura 2.6: Propagação de uma onda com polarização circular. Neste caso só foi representada a intensidade do campo elétrico para facilitar a visualização.[71]

$$\vec{E} = \vec{E}_x + \vec{E}_y \quad (2-12)$$

$$\vec{H} = \vec{H}_x + \vec{H}_y \quad (2-13)$$

$$\vec{E} = E_0 \hat{x} \exp[i(\vec{k}\vec{z} - wt)] + E_0 \hat{y} \exp[i(\vec{k}\vec{z} - wt + \pi/2)] \quad (2-14)$$

$$\vec{H} = H_0 \hat{x} \exp[i(\vec{k}\vec{z} - wt)] + H_0 \hat{y} \exp[i(\vec{k}\vec{z} - wt + \pi/2)] \quad (2-15)$$

tomando a parte real do vetor, temos que:

$$\vec{E} = E_0 \hat{x} \cos(\vec{k}\vec{z} - wt) + E_0 \hat{y} \cos(\vec{k}\vec{z} - wt + \pi/2) \quad (2-16)$$

$$\vec{H} = H_0 \hat{x} \cos(\vec{k}\vec{z} - wt) + H_0 \hat{y} \cos(\vec{k}\vec{z} - wt + \pi/2) \quad (2-17)$$

$$\vec{E} = E_0 \hat{x} \cos(\vec{k}\vec{z} - wt) + E_0 \hat{y} \sin(\vec{k}\vec{z} - wt) \quad (2-18)$$

$$\vec{H} = H_0 \hat{x} \cos(\vec{k}\vec{z} - wt) + H_0 \hat{y} \sin(\vec{k}\vec{z} - wt) \quad (2-19)$$

$$\vec{E}_x = E_0 \hat{x} \cos(\vec{k}\vec{z} - wt) \quad (2-20)$$

$$\vec{H}_x = H_0 \hat{x} \cos(\vec{k}\vec{z} - wt) \quad (2-21)$$

$$\vec{E}_y = E_0 \hat{y} \sin(\vec{k}\vec{z} - wt) \quad (2-22)$$

$$\vec{H}_y = H_0 \hat{y} \sin(\vec{k}\vec{z} - wt) \quad (2-23)$$

$$\|\vec{E}\| = \sqrt{(\vec{E}_x)^2 + (\vec{E}_y)^2} = E_0 \quad (2-24)$$

$$\|\vec{H}\| = \sqrt{(\vec{H}_x)^2 + (\vec{H}_y)^2} = H_0 \quad (2-25)$$

onde  $\hat{x}$  e  $\hat{y}$  são os vetores unitários na direção  $x$  e  $y$ , respectivamente. Pode-se observar que as amplitudes escalares de  $\vec{E}$  e  $\vec{H}$  são contantes. Contudo, as direções dos vetores variam no espaço e tempo, girando no sentido horário, e sua projeção no plano  $xy$  descreve uma circunferência (Figura 2.6).

Microscópios com polarização são poderosos instrumentos analíticos em estudos petrográficos e análise de minérios.[50] Quando um mineral anisotrópico, composto por cristais com orientações diferentes, é observado ao microscópio de luz polarizada, se faz visível a diferença de brilho entre os cristais.

Os microscópios com polarização linear são dotados de dois polarizadores designados de polarizador e analisador. Nos microscópios modernos, estes polarizadores são constituídos de placas de polaróides. Um polaróide é formado por um composto químico orgânico (álcool polivinílico) embebido em iodo e estirado segundo uma certa direção. Desta forma as longas moléculas do material são alinhadas, assim o iodo se liga às cadeias alongadas das moléculas poliméricas. Deste modo, os elétrons dos íons de iodo podem deslocar-se ao longo das cadeias moleculares, tal como num fio condutor.

A componente do campo  $\vec{E}$ , da luz natural incidente, paralela às moléculas executa trabalho sobre os elétrons e é absorvida. Assim, o eixo de transmissão do polaróide é normal à orientação das moléculas. A luz linearmente polarizada incide na amostra e é refletida. Um segundo polaróide (o analisador) deixará passar apenas a componente do campo elétrico que vibra na direção de seu eixo de transmissão. Desta forma, se  $\vec{E}_0$  é a intensidade do campo elétrico determinada pelo polarizador somente sua componente  $\vec{E}_0 \cos(\theta)$ , paralela ao eixo de transmissão do analisador, passará pelo analisador. Neste caso,  $\theta$  é o ângulo entre os eixos de transmissão do polarizador e o analisador. Um exemplo de imagem obtida por polarização linear pode ser observada na Figura 2.7b.

Por outro lado, alguns microscópios óticos modernos, além da polarização linear da luz, permitem também a polarização circular.[72] O microscópio usado no presente trabalho conta 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.

Deste modo, a luz não-polarizada passa pelo polarizador, tornando-se linearmente polarizada. Então, 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

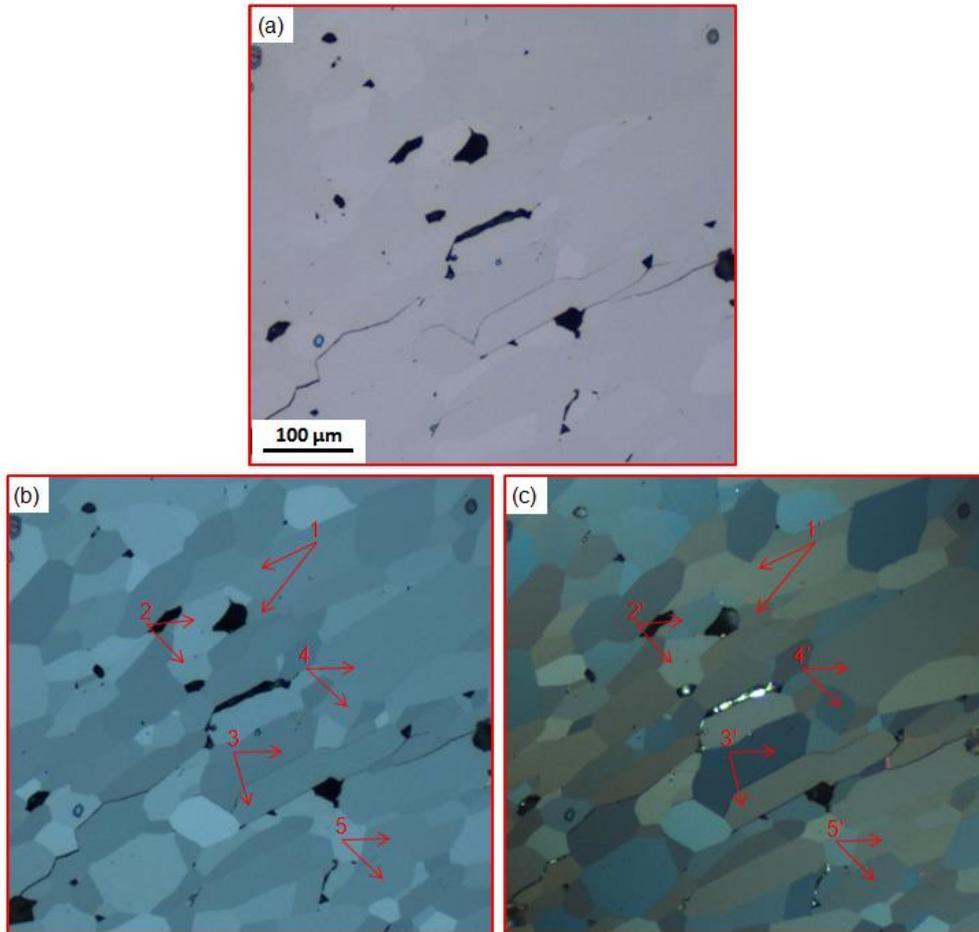


Figura 2.7: Variação de brilho entre cristais adjacentes: (a) Imagem em campo claro; (b) Imagem com polarização linear; (c) Imagem com polarização circular, de um mesmo campo.

componentes com uma diferença de fase de  $90^\circ$ . A combinação dessas duas ondas linearmente polarizadas, de mesma amplitude e defasadas de  $90^\circ$ , resulta em uma onda circularmente polarizada, como mostrado na Figura 2.6.[73]

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 da do polarizador. Em seguida, essa luz passa pelo analisador e segue para a câmera.[73]

Na imagem resultante da polarização circular, não há ponto de extinção, pois todas as orientações de polarização estão presentes.[73] A polarização circular é uma inovadora técnica em microscopia de materiais, permitindo melhorar o contraste na imagem. Além disso, as cores não variam com a rotação da amostra. Assim, objetos que costumavam ser visíveis apenas em uma certa

direção agora podem ser vistas em sua totalidade (Figura 2.7), independente de sua orientação e sem rotação da platina.[50]

Na Figura 2.7 pode-se observar três imagens de um mesmo campo de uma amostra de minério de ferro, a primeira em campo claro (Figura 2.7a), a segunda com polarização linear (Figura 2.7b) e a terceira com polarização circular (Figura 2.7c). Após observar os pares de pontos indicados em vermelho (1-1', 2-2', 3-3' e 4-4'), nas Figuras 2.7b e 2.7c, fica evidente a superioridade de contraste da luz circularmente polarizada. Cristais com o mesmo brilho na Figura 2.7b foram totalmente diferenciados na Figura 2.7c. A formação de contraste entre os cristais de hematita, por luz polarizada, desempenha um papel central no método de segmentação de cristais proposto neste trabalho.

## 2.3

### 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 sua vez, a Análise Digital de Imagens (ADI) consiste na extração e tratamento de dados quantitativos de imagens digitais. Deste modo, utiliza-se o termo Processamento e Análise Digital de Imagens (PADI) para englobar as duas técnicas anteriores.[74]

A Figura 2.8 apresenta a sequência padrão de PADI.[35, 52, 74, 75] Esta sequência está dividida em três blocos básicos:

- (i) A aquisição, que envolve a formação e a digitalização da imagem;
- (ii) O PDI, no qual encontra-se a etapa de pré-processamento, que tem como objetivo melhorar a qualidade da imagem para facilitar sua visualização e análise e;
- (iii) E finalmente a ADI, onde finalmente ocorre a identificação dos objetos dos quais são extraídos e tratados os dados quantitativos. Neste bloco encontram-se as etapas de segmentação, pós-processamento, extração de atributos e reconhecimento e classificação.

Nos lados da sequência padrão aparecem setas que indicam o nível semântico dos dados sobre os quais se trabalha. No pré-processamento e na segmentação, opera-se diretamente sobre os pixels da imagem, gerando-se uma imagem com objetos, representados por regiões de pixels contíguos de mesmo valor. No pós-processamento trabalha-se sobre os objetos que serão medidos na

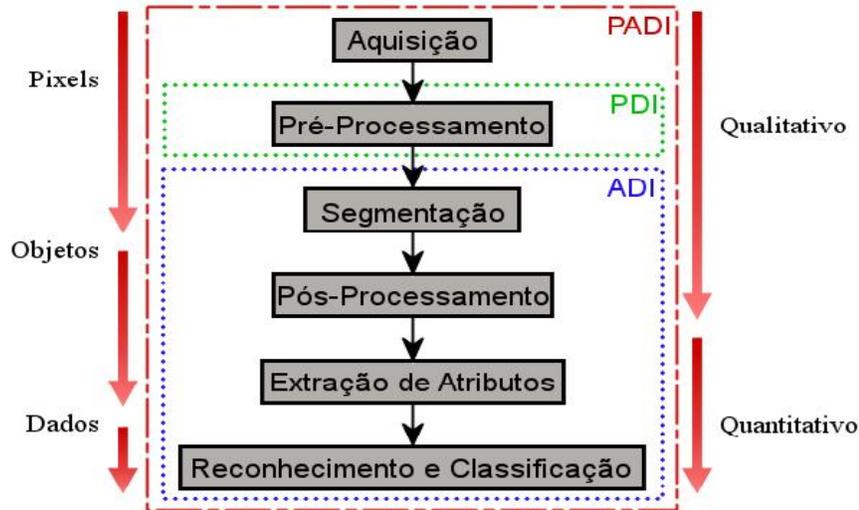


Figura 2.8: Sequência padrão de PADI por etapas.[52]

extração de atributos. A partir daí, na etapa de reconhecimento e classificação, trabalha-se com estas medidas, gerando dados de mais alto nível.

Esta sequência padrão tem sido discutida amplamente em trabalhos prévios do grupo, por isso só as etapas e suas rotinas de interesse para este trabalho serão abordadas a seguir.[35, 52, 74, 75, 76]

### 2.3.1

#### Pré-processamento

A etapa de pré-processamento é o primeiro passo depois da aquisição da imagem. Esta etapa tem como objetivo melhorar a qualidade da imagem, seja corrigindo defeitos gerados durante a aquisição ou seja realçando detalhes de interesse específico.

#### 2.3.1.1

##### Registro de imagem

A primeira coisa a se fazer quando várias imagens de um mesmo campo são capturadas com diferentes sensores ou com um mesmo sensor, porém sob diferentes condições de captura, é um registro espacial das mesmas.[77] O registro é um importante passo, no pré-processamento, a fim de criar imagens geometricamente iguais com uma coerência pixel a pixel.

O registro de imagem é normalmente utilizado em sensoriamento remoto, na medicina, na cartografia, em processamento digital de imagens, entre outras. Durante as últimas décadas, os dispositivos de aquisição de imagens

passaram por um rápido desenvolvimento. Assim, uma crescente e diversa quantidade de imagens estavam sendo obtidas, invocando pesquisas sobre registro automático de imagem.[78] Existe na literatura uma grande variedade de trabalhos relacionados com o registro de imagens, baseados em diferentes princípios e com emprego em diversas aplicações.[77, 78, 79, 80, 81, 82, 83, 84, 85]

O registro basicamente 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 consiste em encontrar uma transformação capaz de remapear as posições dos pixels da imagem registrada de modo que a área sobreposta esteja alinhada com a imagem de referência.[79]

Contudo, a transformação capaz de remapear as posições dos pixels das duas imagens pode ser bastante complexa. Esta transformação pode ser composta por uma combinação de transformações de basicamente seis naturezas distintas:

- ▷ Translação;
- ▷ Rotação;
- ▷ Escala;
- ▷ Paralelismo;
- ▷ Projeção e;
- ▷ Outras distorções como curvas, distorções locais, etc.[52]

A Figura 2.9 apresenta estas diferentes transformações, mostrando claramente seu efeito em uma imagem exemplo.

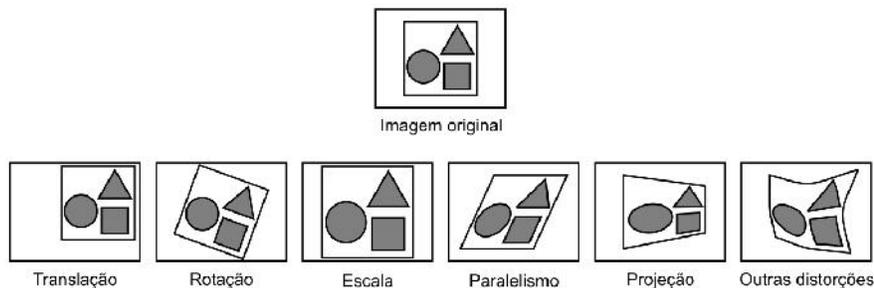


Figura 2.9: Possíveis transformações em registro de imagens.[52]

Historicamente as transformações têm sido classificadas como rígidas ou não-rígidas. Uma transformação é denominada rígida quando apenas

translações e rotações são permitidas. Este tipo de transformação é também conhecida como transformações euclidianas, já que as distâncias euclidianas são preservadas. Por oposição, as demais transformações são definidas como não-rígidas.[52] No caso particular deste trabalho só é de interesse as transformações rígidas.

Não obstante, a primeira coisa a se 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 a gerar a imagem registrada que será sobreposta à imagem de referência.[78]

A Figura 2.10 mostra um exemplo simples de registro de imagem. Nela encontra-se a imagem de referência (Figura 2.10a) e a imagem a ser registrada (Figura 2.10b). Em ambas imagens foram marcados ou desenhados alguns possíveis pontos de controle que poderiam ser usados na geração da imagem registrada (Figura 2.10c). Alguns autores afirmam que para conseguir um registro é necessário pelo menos encontrar 3 pontos de controle entre a imagem de referência e a imagem a ser registrada.[86] Claro está que quanto maior seja a quantidade de pontos, maior será a qualidade do registro em si.[79]

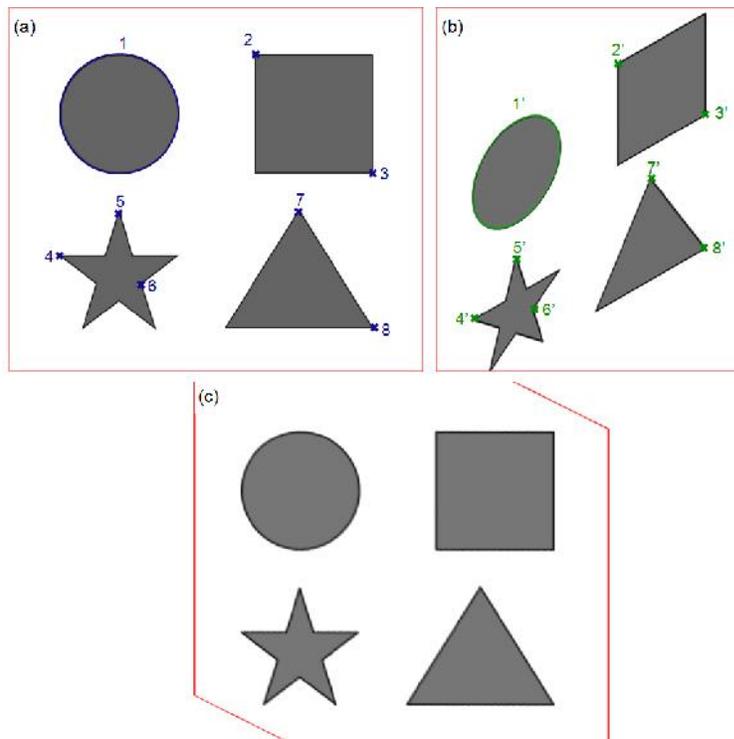


Figura 2.10: Possíveis pontos de controle: (a) Imagem de referência e (b) Imagem a ser registrada; (c) Imagem registrada.

Existe uma técnica de processamento de imagens que permite a detecção e extração de pontos de controle invariáveis a ruído de imagem, rotação, traslação, escala, e de forma parcial a mudanças de iluminação e de perspectivas, chamada transformada SIFT (*Scale Invariant Feature Transform*). A transformada SIFT tem demonstrado ser muito eficiente ao gerar grande número de pontos de controle que conseguem cobrir densamente uma imagem.[87]

Vários softwares usam a transformada SIFT no registro de imagens, entre eles o FIJI. O FIJI é um versátil software livre de processamento digital de imagens, que pode ser descrito como uma distribuição do ImageJ, desenvolvido pelo *National Institutes of Health* dos Estados Unidos. A versatilidade do FIJI está dada pela sua natureza, pois ele é um software composto por *plugins* organizados em uma estrutura de menu coerente. Como software livre ele permite que experientes desenvolvedores em Java possam criar novos *plugins*. Ao mesmo tempo, ele oferece um ambiente de programação mais simples através de uma interface de *script*. [88]

Outros softwares de processamento digital de imagem também usam o registro de imagem entre suas rotinas. É de particular interesse o software comercial AxioVision. Uma das aplicações do registro de imagens por este software é na fabricação de mosaicos. Esta técnica 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.[76] 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. Além disso, este software tem outra função chamada *Geometric Correction* que permite registrar imagens a uma imagem de referência.[48]

### 2.3.1.2 Delineação

Outra operação comum de pré-processamento é a delineação. Imagens obtidas por microscopia ótica apresentam valores de tonalidades intermediários nas bordas dos objetos, devido a restrições de resolução do sistema de formação da imagem. Assim, as fronteiras dos objetos em lugar de serem uma linha de transição abrupta passam a ser uma faixa de transição suave entre as tonalidades vizinhas.[89] Isto traz como consequência que, na hora de segmentar os objetos, seja gerada uma fase espúria em forma de "halo" ao redor dos objetos. Este fenômeno é conhecido como "efeito halo" e dificulta o correto reconhecimento de regiões numa imagem.[52]

Este problema é comumente reduzido utilizando uma técnica de delimitação. A delimitação não é mais que um filtro inteligente que varre a imagem procurando as transições entre fases e virtualmente decidindo a qual fase os pixels pertencem. Para isto ele cria um limiar ( $m$ ) onde todos os pixels abaixo deste limiar pertencem a uma fase e os acima dele pertencem à outra (Figura 2.11). Já os pixels que se encontram fora destas faixas de transição ficam praticamente inalterados.

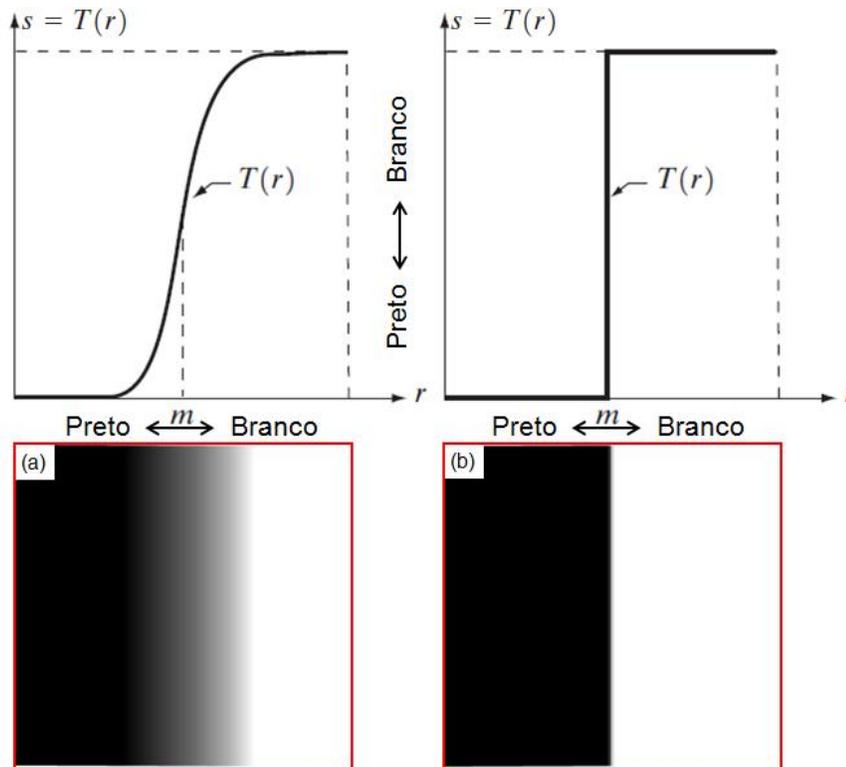


Figura 2.11: Função de transformação para o realce do brilho nas bordas dos objetos.[90]

Um exemplo esquemático disto pode ser observado na Figura 2.11, onde dois objetos vizinhos com tonalidades diferentes (preto e branco) apresentam uma zona de transição (Figura 2.11a). Após aplicar uma delimitação na imagem a transição passa a ser abrupta (Figura 2.11b), como esperado, aumentando o contraste entre os objetos.

Assim, esta técnica é empregada com bastante frequência quando a separação de objetos na etapa de segmentação é crítica. Isto pode ser observado claramente na Figura 2.12 onde uma imagem multimodal de sinter de minério de ferro está prestes a ser delimitada. Da Figura 2.12a é possível notar que o contraste entre as fases com tonalidades mais cinzas (silicatos vs. ferritos e ferritos vs. magnetita) não é tão bom como o contraste entre a fase mais branca

e o resto (silicatos, ferritos e magnetita) vs. hematita. Após a delimitação pode ser observado que os vales entre os picos no histograma das fases mais cinzas diminuíram muito. Ao mesmo tempo, na região ampliada da imagem, percebe-se um aumento do contraste entre as fases.

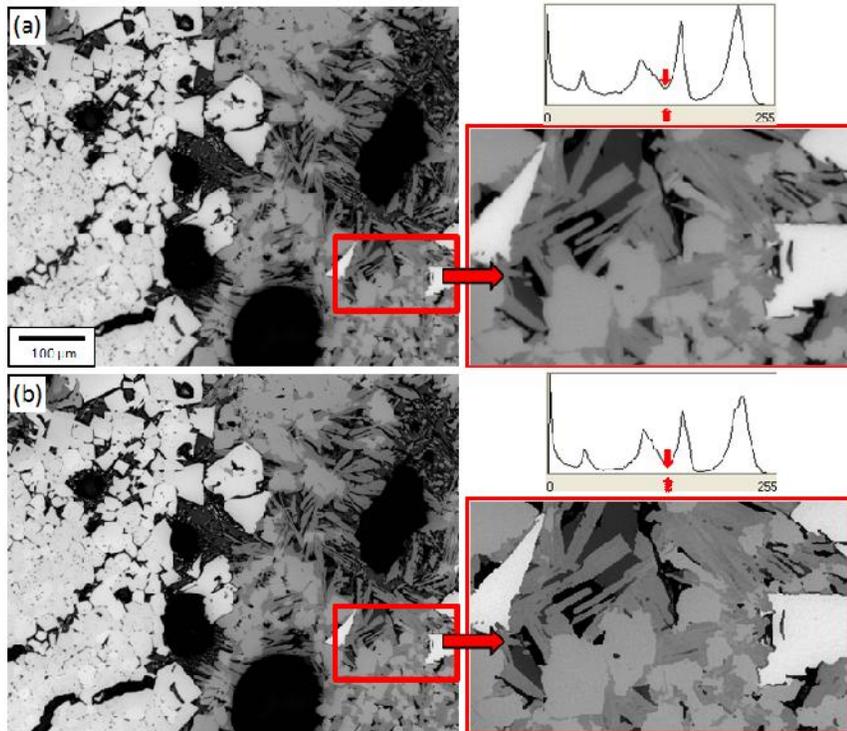


Figura 2.12: Delineação: (a) Imagem original, seu histograma e uma visão ampliada da região demarcada em vermelho; (b) Imagem delineada, seu histograma e a ampliação da mesma região demarcada em vermelho.[76]

### 2.3.1.3 Filtro Mediana

Os ruídos aleatórios aparecem com muita frequência nas imagens. Estes ruídos podem ser gerados por mal funcionamento dos dispositivos de captura ou por falta de iluminação.[91] Os pixels corrompidos ou são alterados para o valor máximo da tonalidade, ou têm só alguns *bits* alterados, causando uma diferença brusca de tons entre estes pixels e seus vizinhos. Quando os pixels são alternadamente modificados para 0 ou o máximo (255 em imagens de 8 bit), este ruído é chamado de ruído sal-e-pimenta, devido a sua aparência.

Para este tipo de ruído o filtro mais eficiente é o filtro mediana.[92] A mediana, além de eliminar eficientemente o ruído aleatório, preserva o contorno dos objetos da imagem.[93] Portanto, este filtro é de grande importância

na etapa de pré-processamento ao eliminar defeitos da aquisição e assim ajudando na segmentação.

A mediana consiste num *kernel* (Figura 2.13) de tamanho  $n \times m$  que ordena os pixels em ordem crescente de intensidade e escolhe como saída o valor mediano (aquele que esta no centro da sequência). Como mostra a Figura 2.13, o pixel central do *kernel* (245) será substituído pelo valor que está no centro da sequência ordenada (98). Neste exemplo observe-se que quando  $n \times m$  é ímpar, a mediana é o próprio elemento central da sequência ordenado. Nos casos em que  $n \times m$  é par, a mediana é calculada pela média aritmética dos dois elementos mais próximos do centro.[94]

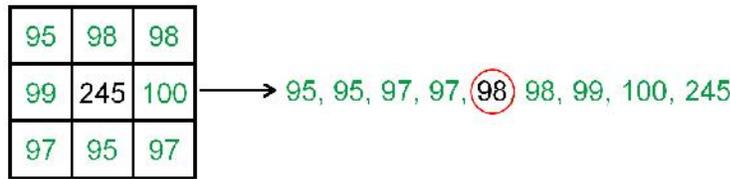


Figura 2.13: Exemplo de um filtro mediana de tamanho 3x3.

O filtro mediana pode ser observado em ação na Figura 2.14. Nesta figura temos uma imagem esquemática com um ruído do tipo sal-e-pimenta (Figura 2.14a). Sobre a mesma é aplicado um filtro mediana de tamanho  $3 \times 3$ , onde os pixels com tonalidades muito diferentes dos vizinhos foram eliminados (Figura 2.14b). Contudo, ainda pode ser observado pequenos defeitos residuais nas bordas dos objetos e da imagem, isto se deve à característica do filtro de não alterar os contornos dos objetos.

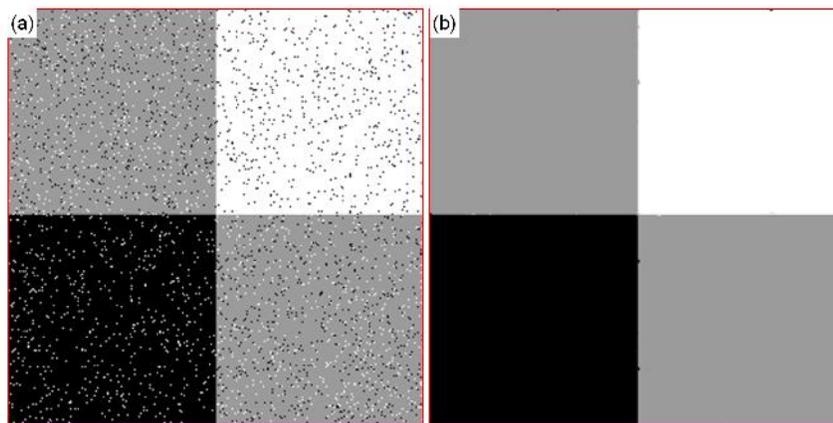


Figura 2.14: Uso do filtro mediana: (a) Imagem esquemática com ruído sal-e-pimenta; (b) Imagem esquemática com ruído reduzido pelo filtro mediana de tamanho 3x3.[80]

### 2.3.2

#### Técnicas de segmentação

Alguns psicólogos alemães no início do século XX, Kbler, Wertheimer e Koffka, introduziram o princípio da segmentação. Eles mostraram que o ser humano, no nível da visão, efetua agrupamentos sobre o que ele percebe, baseados na proximidade, similaridade, e continuidade.[95] Entende-se por segmentação o processo de divisão de uma imagem em suas partes ou objetos constituintes, sendo uma das tarefas mais difíceis na análise de imagens.[96]

A segmentação é uma tarefa básica no processo de análise de imagens, mas não é fácil traduzir para o computador o sofisticado processo de seleção e agrupamento realizado pela visão humana na identificação de regiões semelhantes. Várias dificuldades estão presentes nesse processo como a complexidade da textura, a irregularidade da iluminação e as imprecisões das regiões das bordas. O problema de segmentação torna-se particularmente difícil devido à textura da imagem. Métodos de agrupamento baseados na cor podem alcançar resultados satisfatórios quando as regiões que se deseja segmentar são homogêneas. Porém, cenas naturais são complexas pela riqueza de variações tonais e texturas presentes.[97]

A segmentação de imagens tem sido um problema amplamente discutido no campo de análise de imagens digitais. Existem diversas técnicas de segmentação, dentre as quais se encontram:

- (i) Segmentação por faixa tonal ou limiarização ;[98, 99, 100]
- (ii) Segmentação por contornos;[101, 102]
- (iii) Segmentação por crescimento de regiões;[103, 104]
- (iv) Entre outras, tais como *watershed*, segmentação por textura, segmentação por entropia, etc.[105, 106, 107, 108, 109]

Categoricamente, não existe um método ideal e genérico de segmentação que seja sempre o melhor.[52] A segmentação costuma ser a etapa crítica da sequência padrão de PADI. Uma segmentação adequada praticamente garante o sucesso no reconhecimento e na identificação dos objetos de interesse sobre os quais será feita a análise.[90]

### 2.3.2.1

#### Limiarização

A técnica de segmentação mais simples e a mais utilizada é a segmentação por faixa tonal, também chamada limiarização ou *thresholding*. 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 limiar ou tom de corte.[74] Este método é implementado de forma interativa na maior parte das vezes, porém alguns autores tem se esforçado por criar inovadores métodos de limiarização automática.[100, 110]

No caso de diferenciar objetos de um fundo, utiliza-se a limiarização bimodal, a qual discrimina duas fases na imagem, o fundo e os objetos. 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 o tom 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:

$$g(x, y) = \begin{cases} 1 & \text{se } (x, y) > L \\ 0 & \text{se } (x, y) \leq L \end{cases} \quad (2-26)$$

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 Figura 2.15 mostra um exemplo simples de segmentação por faixa tonal. Uma imagem de minério de ferro (Figura 2.15a), em 256 tons por canal RGB, obtida por microscopia ótica, é segmentada por limiarização bimodal. Gera-se uma imagem binária (Figura 2.15b), através do tom de corte (L=132) indicado em seu histograma (Figura 2.15c), e a fase mais clara (hematita) é separada do resto da imagem. Neste caso específico, por razões metodológicas, o histograma mostrado pertence apenas a um canal da imagem RGB.

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:

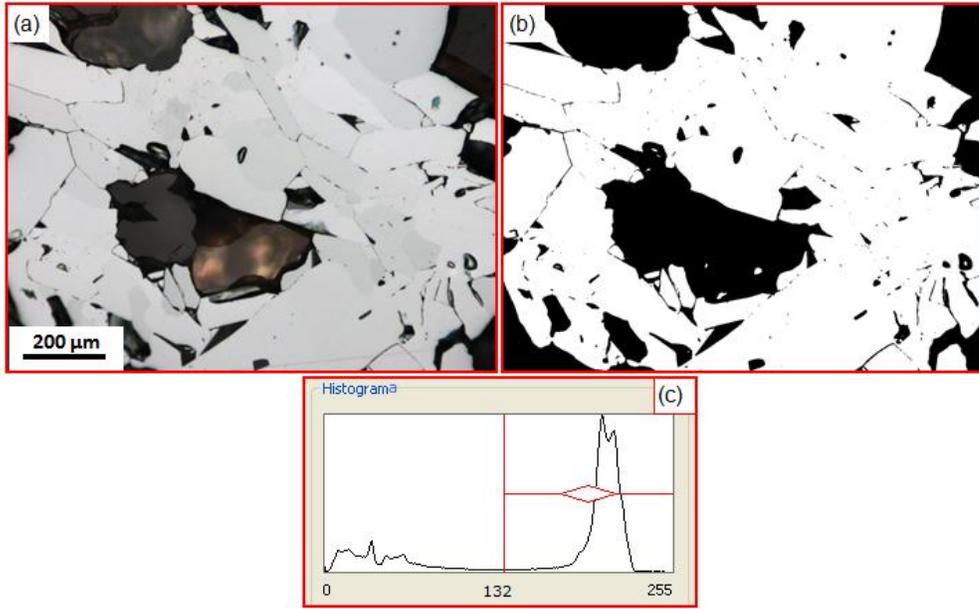


Figura 2.15: Exemplo de limiarização bimodal: (a) Imagem original; (b) Imagem binária; (c) Tom de corte no histograma.

$$g_i(x, y) = \begin{cases} 1 & \text{se } L_i < (x, y) \leq L_{i+1} \\ 0 & \text{se } L_i \geq (x, y) > L_{i+1} \end{cases} \quad (2-27)$$

onde  $g_i(x, y)$  seria a  $i$ -ésima imagem binária pertencente à  $i$ -ésima fase e;

$$g(x, y) = \bigcup_{i=1}^n g_i(x, y) \quad (2-28)$$

o resultado final da limiarização multimodal, onde os pixels com valor 1 correspondem às fases de interesse, enquanto os pixels com valor 0 correspondem ao resto.

A Figura 2.16 mostra um exemplo de limiarização multimodal para distinguir cinco fases. Uma imagem de sinter de minério de ferro (Figura 2.16a) segmentada por limiarização pentamodal, gerando uma imagem com 5 faixas (Figura 2.16b) através dos quatros tons de corte ( $L=32$ ,  $L=74$ ,  $L=134$ ,  $L=170$ ) mostrados em seu histograma (Figura 2.16c). A cada faixa foi atribuída uma cor, para facilitar a visualização, correspondente à cor do histograma entre seus respectivos tons de cortes.

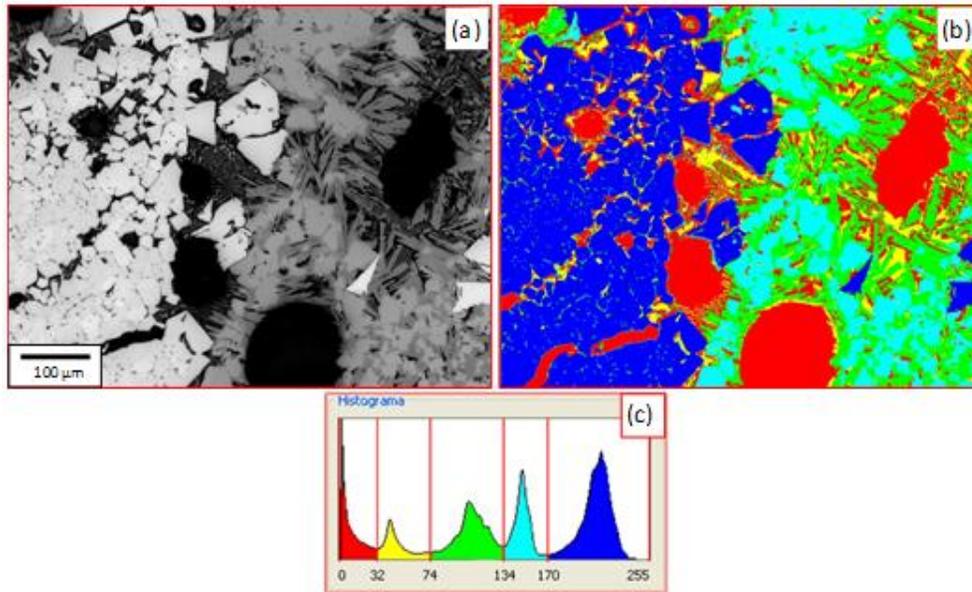


Figura 2.16: Exemplo de limiarização pentamodal: (a) Imagem original em 256 tons de cinza; (b) Imagem quinária com as fases diferenciadas com cores; (c) Tons de corte no histograma.[76]

### 2.3.2.2

#### Segmentação por Contornos

Um contorno é uma mudança brusca da tonalidade entre duas regiões relativamente homogêneas. Ele pode aparecer como uma sequência de pontos, uma linha, um segmento, uma curva ou uma forte variação do nível de cinza médio.[90]

Uma das técnicas de detecção de bordas mais usadas consiste no processamento de uma imagem a partir de um operador de derivada local. Exemplos destes operadores são o gradiente e o laplaciano, operadores de derivada de primeira e segunda ordem, frequentemente utilizados nas técnicas de detecção de bordas.[90]

A segmentação por contornos simula o funcionamento da visão humana. Ela detecta as bordas dos objetos, a partir das quais constrói seus contornos, considerando como objeto a região dentro do contorno.[74]

Entre as técnicas de segmentação por contornos, uma das mais conhecidas é a proposta por Canny.[102] Este autor desenvolveu um processo de detecção de bordas a partir de critérios de quantificação de desempenho de operadores de bordas conhecidos como os critérios de detecção e de localização. Estes critérios de desempenho ainda estão sujeitos ao critério de resposta múltipla, que corresponde ao fato de que deve existir, na saída do operador, uma única

resposta para uma única borda. Para que os critérios sejam aproximadamente atendidos, Canny aproxima o operador ótimo, obtido a partir dos três critérios de desempenho, pela primeira derivada da função Gaussiana. Em complemento a este operador foram propostos dois processos conhecidos como:

- (i) Supressão não máxima (supressão de valores de pixels que não forem máximos locais na direção transversal à borda), que causaria um afinamento da borda, atendendo à injunção de resposta múltipla; e
- (ii) Uma limiarização adaptativa (histerese) com “complementação de bordas”, para eliminar a fragmentação dos contornos das bordas.[111]

A Figura 2.17 mostra um exemplo de detecção de borda. Uma imagem em luz polarizada de minério de ferro (Figura 2.17a) obtida por microscopia ótica, é segmentada pelo método de Canny. Gera-se uma imagem binária (Figura 2.17b) com as bordas dos cristais de hematita. Contudo, este método é sensível aos riscos, às bordas espúrias provenientes de ruído e à textura da imagem, conforme indicado pelas setas verdes nas imagens da Figura 2.17. Por outro lado, o método do Canny gera muitas bordas descontínuas.

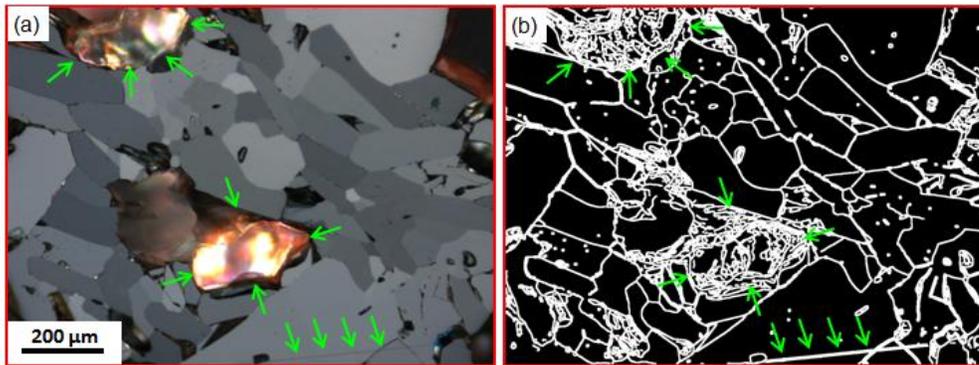


Figura 2.17: Exemplo de detecção de bordas pelo método de Canny: (a) Imagem original; (b) Imagem binária com as bordas dos cristais de hematita.

### 2.3.2.3

#### 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, tom de cinza, textura ou cor.[112] Porém, o que define o tipo de crescimento de regiões é o critério de parada do algoritmo. No presente trabalho serão usados dois tipos de crescimento de regiões:

- (i) Crescimento de regiões usando a distância espectral como parâmetro de cor e;
- (ii) Crescimento de regiões usando a heterogeneidade como parâmetro de textura.

Inicialmente, vai ser explicado 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:

$$I = \bigcup_{i=1}^n R_i \quad (2-29)$$

$$R_i \text{ é uma região conexa, } 1, 2, \dots, n \quad (2-30)$$

$$R_i \cap R_j = \emptyset \quad \forall i \neq j \quad (2-31)$$

$$P(R_i) = \text{verdadeiro} \quad \forall i \quad (2-32)$$

$$P\left(R_i \cap R_j\right) = \text{falso} \quad \forall i \neq j \quad (2-33)$$

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.[90]

A técnica consiste nas seguintes etapas:

- (i) Escolha dos pixels-sementes (pontos ou simplesmente sementes);
- (ii) Escolha do limiar que separará as regiões; e
- (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.

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

Como anteriormente definido, o crescimento das regiões em si consiste em se fazer o agrupamento de pixels por similaridade baseado em alguns critérios como intensidade dos tons de cinza, textura, cor, etc. Aqui, é utilizada

a intensidade dos tons de cinza como exemplo. Então pode-se dizer que:

$$\text{Se } | P(x, y) - P(x_1, y_1) | < t, \text{ então } P(x, y) \in R_1 \text{ senão,} \quad (2-34)$$

$$\text{Se } | P(x, y) - P(x_2, y_2) | < t, \text{ então } P(x, y) \in R_2 \text{ senão,} \quad (2-35)$$

...

$$\text{Se } | P(x, y) - P(x_n, y_n) | < t, \text{ então } P(x, y) \in R_n; \quad (2-36)$$

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.

Um exemplo simples do funcionamento do crescimento de regiões clássico pode ser observado na Figura 2.18. Na Figura 2.18a pode-se observar uma imagem em tom 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 2.18b). Contudo, se for tomada uma terceira semente  $P(4, 4) = 3$  e for mantido com o mesmo limiar  $t = 3$ , seria criada uma terceira região ( $R_3$ ) incluindo os pixels antes rejeitados (Figura 2.18c). 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 o sucesso esperado no método de crescimento de regiões.

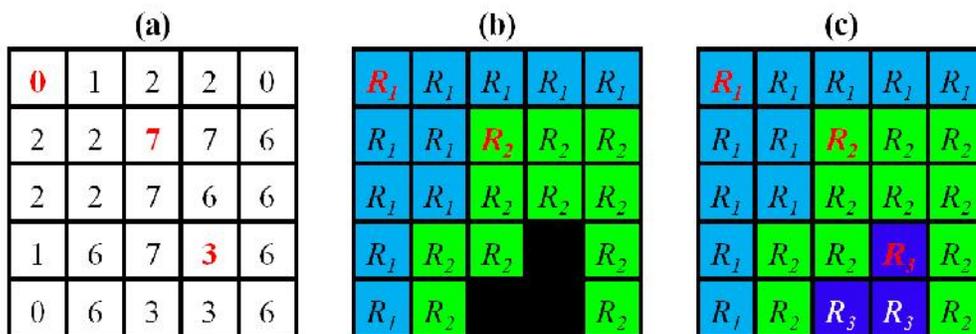


Figura 2.18: Exemplo de crescimento de regiões com  $t = 3$ : (a) Imagem em tons de cinza; (b) Duas regiões ( $R_1$  e  $R_2$ ) separadas com  $n = 2$  sementes (em vermelho); (c) Três regiões ( $R_1$ ,  $R_2$  e  $R_3$ ) separadas com  $n = 3$  sementes (em vermelho).

### 2.3.3

#### Pós-processamento

Frequentemente o resultado da segmentação não é tão bom quanto esperado, sendo necessário aplicar rotinas de pós-processamento para corrigir os defeitos recorrentes da segmentação. Estas rotinas geralmente se baseiam na implementação de operações lógicas e morfológicas.

#### 2.3.3.1

##### Operações Lógicas

Na segmentação a imagem é binarizada, ou seja, cada pixel que compõe a imagem ou bem é branco (com valor 1 ou 255 dependendo do software), ou bem é preto (com valor 0). Assim, resulta fácil imaginar que operações lógicas (também conhecidas como operações booleanas em homenagem a seu criador Boole) possam ser aplicadas entre os pixels das imagens.

As operações lógicas são operações pontuais entre imagens binárias, realizadas por operadores lógicos ( $\otimes$ ) que varrem as imagens de entrada, operando pixel a pixel, gerando uma imagem de saída (Figura 2.19) onde cada pixel é preservado ou invertido.[74] Ou seja:

$$A \otimes B = C \quad (2-37)$$

onde  $A$ ,  $B$  e  $C$  são imagens e  $\otimes$  é um operador lógico (AND, OR, XOR).

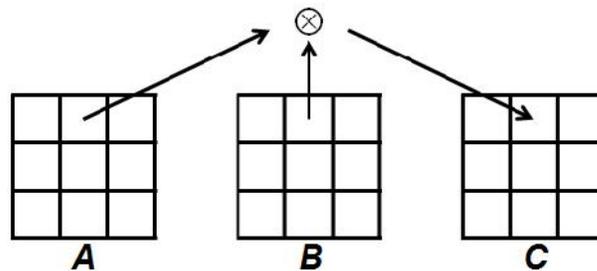


Figura 2.19: Operações lógicas pixel a pixel.

As três operações lógicas básicas são o complemento (NOT), a união (OR) e a intersecção (AND), a partir das quais qualquer outra operação lógica pode ser definida. Na Figura 2.20 pode ser observado como a operação NOT inverte todos os pixels da imagem de entrada, gerando uma imagem de saída que é seu negativo. Ao mesmo tempo, a operação OR realiza a união das duas imagens de entrada, produzindo uma imagem de saída onde são

brancos somente os pixels que são brancos em pelo menos uma das imagens de entrada. Finalmente, a operação AND faz a intersecção entre as duas imagens de entrada, produzindo uma imagem de saída onde são brancos somente os pixels que são brancos em ambas as imagens de entrada.

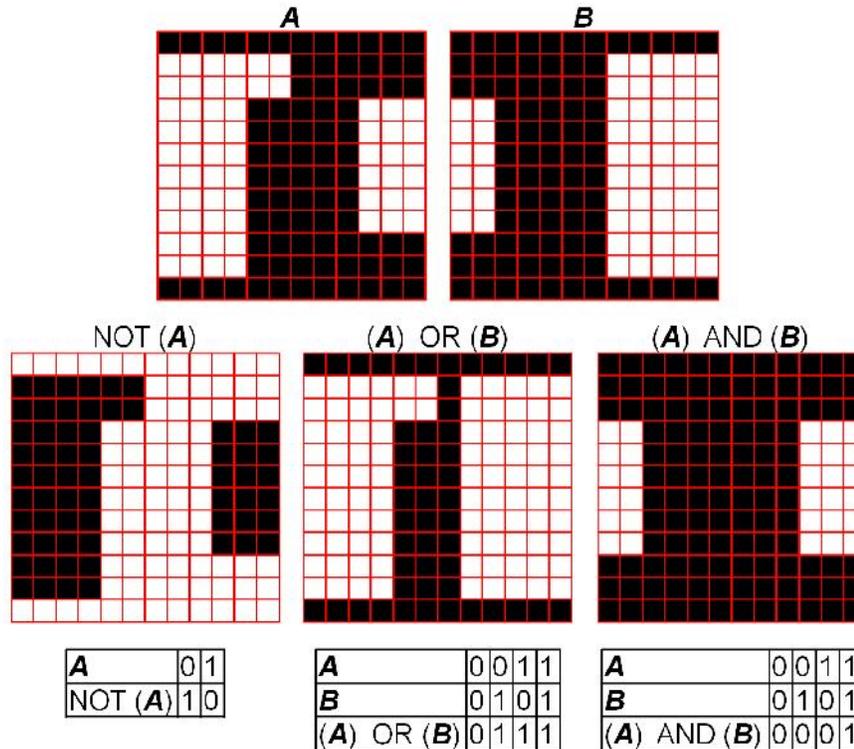


Figura 2.20: Operações lógicas básicas (NOT, OR e AND).

Outra operação lógica, derivada dessas operações básicas, também bastante utilizada, é o OR exclusivo (XOR). Esta operação não é mais que a intersecção entre a saída de um OR e o complemento da saída de um AND. Ou seja:

$$(A) XOR (B) = \{(A) OR (B)\} AND \{NOT [(A) AND (B)]\} \quad (2-38)$$

Na Figura 2.21 pode ser observado como a operação XOR gera uma imagem de saída onde unicamente são brancos aqueles pixels que eram brancos em somente uma das imagens de entrada. Na prática, o operador XOR calcula a diferença entre as duas imagens binárias (não confundir com a operação de subtração).

Como pode ser observado, os operadores AND, OR e XOR requerem duas imagens de entrada e geram uma imagem de saída. Porém, no caso particular

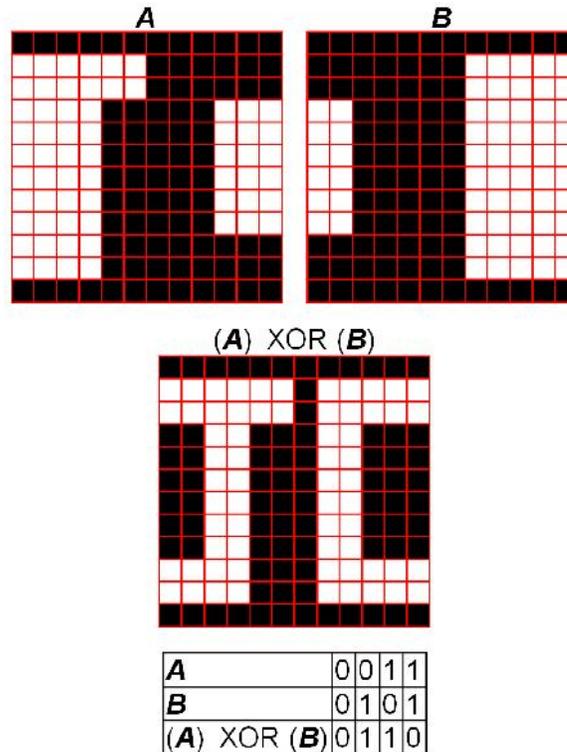


Figura 2.21: Operação lógica XOR.

do operador NOT somente requer uma imagem de entrada para gerar uma imagem de saída. Assim, podem-se criar infinitas combinações lógicas com estes operadores a fim de corrigir defeitos da segmentação em imagens binárias. Contudo, é bom ter presente a importância da ordem dos operadores, deste modo é de vital importância o uso de parêntesis para esclarecer a ordem e prioridade das operações.

Até aqui vimos o uso dos operadores lógicos em imagens binárias, porém eles também podem ser usados em imagens em tom de cinza. O uso de uma imagem binária como máscara para modificar uma imagem em tom de cinza é uma prática comum no processamento de imagens.

### 2.3.3.2 Operações Morfológicas

A morfologia matemática concentra seus esforços no estudo das estruturas geométricas das entidades presentes numa imagem.[113] As operações morfológicas podem ser aplicadas em várias áreas de processamento e análise de imagens, com objetivos tão distintos como realce, filtragem, segmentação, detecção de bordas, esqueletização, afinamento, dentre outras.[93]

Assim como as operações lógicas, as operações morfológicas também têm

suas operações básicas (traslação, reflexão, complemento, erosão e dilatação). Estas operações consistem em extrair as informações relativas à geometria dos objetos de uma imagem, apoiando-se numa pequena imagem binária, denominada elemento estruturante. O elemento estruturante varre a imagem de entrada, preservando ou invertendo o valor do pixel central da vizinhança, na imagem de saída, em função de seus vizinhos.[74] Portanto, a base das operações morfológicas é a teoria de conjuntos. Neste caso particular em que tanto as imagens de entrada como o elemento estruturante são imagens binárias, os conjuntos em questão pertencem ao espaço inteiro bidimensional ( $\mathbb{Z}^2$ ).

Primeiramente procederemos a fazer as definições das operações morfológicas básicas entre os conjuntos  $\mathbf{A}$  e  $\mathbf{B}$ , pertencentes ao espaço  $\mathbb{Z}^2$ , onde  $a$  e  $b$  são suas respectivas componentes. O conjunto  $\mathbf{A}$  é normalmente atribuído à imagem de entrada, quanto o conjunto  $\mathbf{B}$  é normalmente atribuído ao elemento estruturante.

A primeira definição é a traslação de  $\mathbf{A}$  ( $\mathbf{A}_x$ ) em  $x$  gerando a imagem  $\mathbf{C}$ . [90] Ou seja:

$$\mathbf{A}_x = \{c; c = a + x \quad \forall a \in A\} \quad (2-39)$$

A reflexão de  $\mathbf{B}$  ( $\hat{\mathbf{B}}$ ), por sua vez, seria:

$$\hat{\mathbf{B}} = \{c; c = -b \quad \forall b \in B\} \quad (2-40)$$

O complemento de  $\mathbf{A}$  ( $\mathbf{A}^c$ ):

$$\mathbf{A}^c = \{c; c \notin A\} \quad (2-41)$$

A erosão de  $\mathbf{A}$  por  $\mathbf{B}$  ( $\mathbf{A} \ominus \mathbf{B}$ ):

$$\mathbf{A} \ominus \mathbf{B} = \{c; \mathbf{B}_x \subseteq A\} \quad (2-42)$$

o que, em outras palavras, significa que a erosão de  $\mathbf{A}$  por  $\mathbf{B}$  resulta no conjunto dos pontos  $c$ , da imagem de saída, tais que  $\mathbf{B}$ , trasladado ( $\mathbf{B}_x$ ), está contido em  $\mathbf{A}$ .

Finalmente a dilatação de  $\mathbf{A}$  por  $\mathbf{B}$  ( $\mathbf{A} \oplus \mathbf{B}$ ):

$$\mathbf{A} \oplus \mathbf{B} = \left\{ c; \left[ (\hat{\mathbf{B}})_x \cap \mathbf{A} \right] \subseteq \mathbf{A} \right\} \quad (2-43)$$

ou seja, o processo de dilatação consiste em obter primeiramente a reflexão de  $\mathbf{B}$  sobre sua origem e depois sua traslação  $((\hat{\mathbf{B}})_x)$ . Deste modo, a dilatação de  $\mathbf{A}$  por  $\mathbf{B}$  é o conjunto dos pontos  $c$ , da imagem de saída, para os quais a intersecção de  $(\hat{\mathbf{B}})_x$  e  $\mathbf{A}$  está contido em  $\mathbf{A}$ .

A erosão e a dilatação podem ser observadas mais facilmente na Figura 2.22, onde uma imagem esquemática ( $\mathbf{A}$ ), de 12x12 pixel, é erodida ( $\mathbf{C}$ ) e dilatada ( $\mathbf{D}$ ) pelo elemento estruturante ( $\mathbf{B}$ ), de 3x3 pixel.

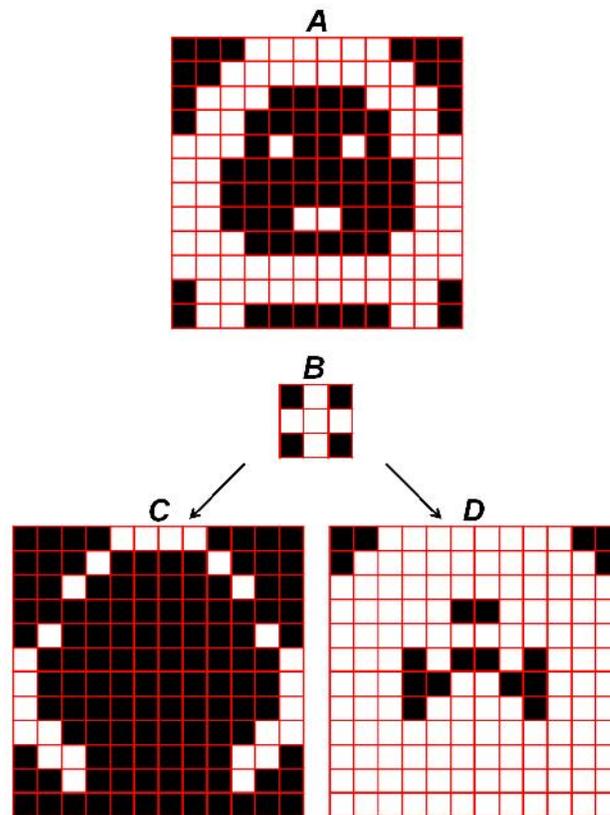


Figura 2.22: Erosão e Dilatação.

É bom ressaltar que alguns softwares de processamento digital de imagens permitem ao usuário criar seus próprios elementos estruturantes.[114] Estes elementos estruturantes podem ter as mais variadas formas e tamanhos, sendo a escolha do mais adequado determinada somente em função do problema. No caso particular deste exemplo (Figura 2.22) o elemento estruturante usado ( $\mathbf{B}$ ) é chamado de elemento de conectividade 4. Isto se deve ao fato de que ele define uma vizinhança 3x3, sendo considerados vizinhos do pixel central somente

os 4 pixels adjacentes a ele na horizontal e na vertical. Um outro elemento estruturante, também muito utilizado e que também define uma vizinhança 3x3, é o elemento de conectividade 8. Neste caso todos os 8 pixels adjacentes ao pixel central são considerados seus vizinhos.

Combinando estas operações morfológicas básicas podem-se criar outras novas e importantes operações morfológicas. Duas das mais importantes são a abertura e o fechamento (Figura 2.23).

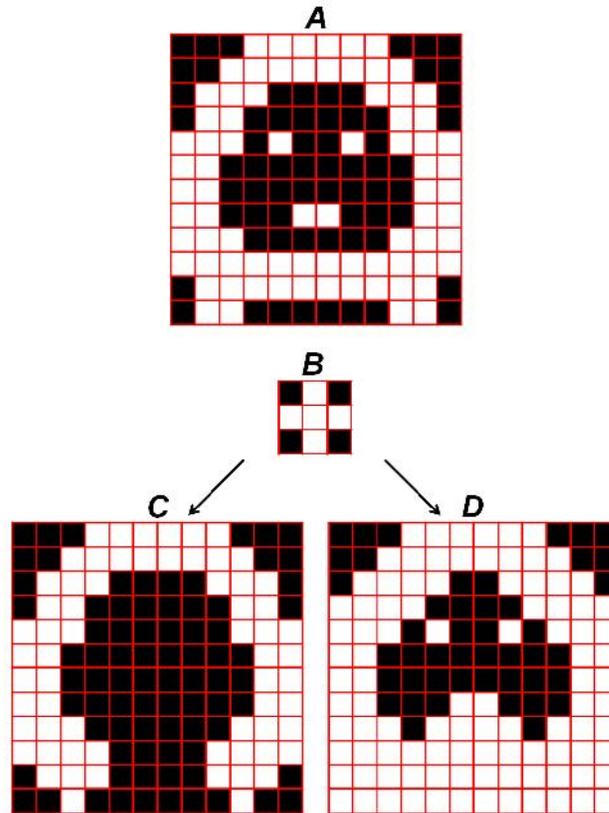


Figura 2.23: Abertura e Fechamento.

A abertura de um conjunto **A** por um elemento estruturante **B** ( $A \circ B$ ) é definida como:

$$A \circ B = (A \ominus B) \oplus B \tag{2-44}$$

o que equivale a dizer que a abertura de **A** por **B** é simplesmente a erosão de **A** por **B** seguida de uma dilatação do resultado por **B**.

O fechamento de um conjunto **A** por um elemento estruturante **B** ( $A \bullet B$ ) é definido como:

$$\mathbf{A} \bullet \mathbf{B} = (\mathbf{A} \oplus \mathbf{B}) \ominus \mathbf{B} \quad (2-45)$$

ou seja, o fechamento não é mais que a dilatação de  $\mathbf{A}$  por  $\mathbf{B}$  seguida da erosão do resultado pelo mesmo elemento estruturante  $\mathbf{B}$ .

Como pode ser observado na Figura 2.23, a abertura em geral suaviza o contorno de uma imagem, elimina faixas estreitas, assim como objetos pequenos. Por sua vez, o fechamento pode preencher trincas ou buracos pequenos e até conectar objetos próximos.

É bom ressaltar que estas operações morfológicas podem ser aplicadas  $k$  vezes sobre uma mesma imagem. Desta forma pode-se, por exemplo, reduzir o tamanho de um objeto numa imagem binária através de múltiplas erosões até chegar ao resultado desejado. Outro exemplo poderia ser aplicar uma abertura de  $k$  passos para eliminar pequenos objetos espúrios.

A erosão derradeira é um bom exemplo de iteração sucessiva de operações morfológicas. Esta técnica consiste em simplesmente erodir os objetos de uma imagem usando a Equação 2-42 até que um próximo passo os eliminaria. Esta técnica pode ser utilizada, por exemplo, para localizar “sementes” de objetos que mais tarde viriam a ser usadas na segmentação por crescimento de regiões.[67]

O preenchimento de buracos é outra das operação morfológicas comumente utilizadas na análises de imagens. Neste caso, entenda-se como buraco o interior de um subconjunto fechado (contorno), de pixels brancos (1) de conectividade 8, que pertence ao conjunto  $\mathbf{A}$ . Assim, partindo de um ponto  $X_0$  situado dentro do contorno, o que se deseja é preencher o interior do mesmo com pixels brancos. Assumindo que todos os pontos que não pertencem ao contorno são pixels pretos (0), será atribuído o valor 1 a  $X_0$  para iniciar o procedimento, e assim por todos os pontos  $X_k$  do interior contorno até preencher a região com pixels brancos. Ou seja:

$$X_k = (X_{k-1} \oplus \mathbf{B}) \cap \mathbf{A}^c \quad (2-46)$$

onde  $k = 1, 2, 3, \dots$  e  $\mathbf{B}$  é o mesmo elemento estruturante da Figura 2.23. O algoritmo termina na  $k$ -ésima iteração se  $X_k = X_{k-1}$ .

O afinamento é similar à erosão derradeira, porém com a condição de não remover pixels que quebrariam o objeto em dois. Assim, o afinamento de um conjunto  $\mathbf{A}$  por um elemento estruturante  $\mathbf{B}$  ( $\mathbf{A} \oslash \mathbf{B}$ ) pode ser definido como:

$$A \oslash B = A \cap [(A \ominus b_1) \cap (A^c \ominus b_2)]^c \quad (2-47)$$

onde  $B=(b_1,b_2)$  sendo que  $b_1$  é o conjunto dos elementos de  $B$  associado com os pixels brancos e  $b_2$  é o conjunto dos elementos de  $B$  associado com os pixels pretos.

Até aqui, todas as operações morfológicas foram definidas usando um elemento estruturante. Contudo, o uso de um elemento estruturante nem sempre é a melhor escolha. O problema é que os elementos estruturantes deformam os objetos, dando-lhes a forma deste elemento, transformando contornos suaves em contornos angulosos. Porém, existem outras formas de criar estes operadores morfológicos a partir do cálculo do Mapa de Distâncias Euclidianas (MDE).

O MDE é uma ferramenta, que diferentemente das operações morfológicas anteriores, só pode ser usada em imagens binárias, produzindo uma imagem em tom de cinza como imagem de saída.[115] A definição desta técnica é bastante simples, ela consiste em atribuir um valor tonal a cada pixel proporcional à sua menor distância da borda do objeto (Figura 2.24b).

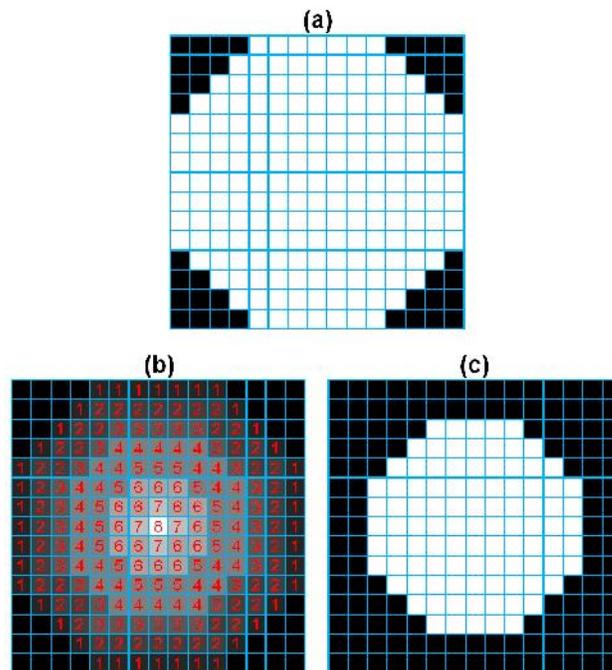


Figura 2.24: Exemplo esquemático do uso de MDE: (a) Imagem binária; (b) Imagem MDE; (c) Erosão de passo 2.

Como pode ser observado na Figura 2.24, a partir da imagem binária de entrada (Figura 2.24a), o cálculo do MDE gera uma imagem de saída

em tons de cinza (Figura 2.24b), mantendo os pixels do fundo em preto e atribuindo, a cada pixel dos objetos, o valor, aproximado ou truncado, da distância euclidiana deste pixel ao pixel mais próximo da borda do objeto. A partir daqui, as operações morfológicas podem ser feitas com uma simples linearização da imagem MDE. Foi assim que uma erosão de passo 2 foi aplicada nela, tomando como limiar inferior e superior da segmentação os valores 1 e 8, respectivamente (Figura 2.24c).

No caso particular da dilatação, precisa-se inverter (NOT) primeiramente a imagem binária, depois calcular o MDE, limiarizar o MDE e finalmente inverter a imagem segmentada. Desta forma, com ajuda da erosão e dilatação euclidianas, podem-se criar os operadores morfológicos restantes.

As operações baseadas no MDE apresentam um custo computacional maior que as anteriores. Contudo, este método realiza a propagação de maneira verdadeiramente radial, afetando menos a forma dos objetos e sendo independentes de sua rotação na imagem. De fato, as operações morfológicas baseadas no MDE são mais acuradas.[74]

Finalmente, outra operação morfológica que se baseia no MDE é o método do divisor de águas.[105] Este método é muito usado para separar objetos que se tocam, ou parcialmente superpostos, na imagem. Este é um problema muito comum, derivado de uma segmentação incompleta, o que faz com que este método seja uma ferramenta muito importante na etapa de pós-processamento.[116]

O método do divisor de águas se baseia no crescimento de sementes dos objetos para visar sua separação. Para isto, calcula-se o MDE (Figura 2.25b) da imagem binária de entrada (Figura 2.25a). Este MDE será usado para obter as sementes dos objetos (Figura 2.25c) através de uma erosão derradeira da imagem binária de entrada. Finalmente, é aplicada uma dilatação derradeira acima das sementes, porém com a condição de que os objetos não voltem a se unir, para assim separar os objetos (Figura 2.25d).

Cada iteração da dilatação será equivalente à erosão anteriormente sofrida, adicionando contornos com a espessura de um pixel à semente, exceto aqueles pixels que eventualmente poderiam unir objetos. Este processo permite que os objetos cresçam até seu tamanho original conservando sua forma e separados por linhas de 1 pixel de espessura.

A ideia do nome “método do divisor de águas” surgiu pela aparente topografia observada numa imagem MDE, que assemelha os objetos com “montanhas”, onde quanto maior a distância de um pixel à borda do objeto, maior será a sua “altura”. Quando dois objetos se tocam, formam-se dois picos, um vale entre eles é uma linha que os separa. A ideia então seria fazer passar

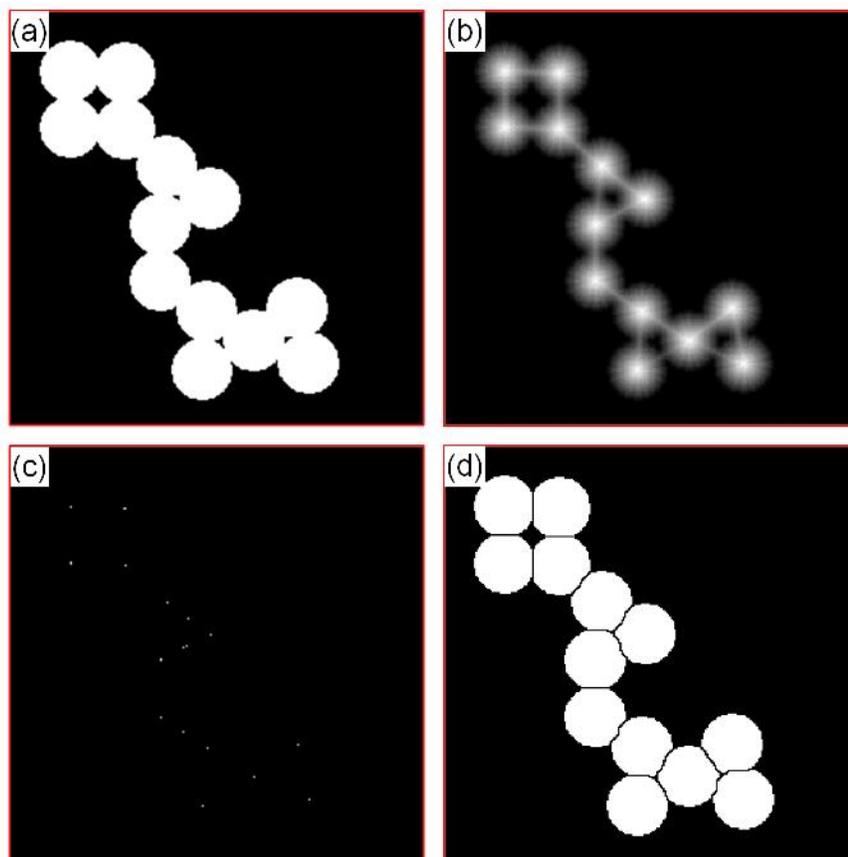


Figura 2.25: Método dos divisores de água: (a) Imagem binária inicial; (b) Imagem MDE; (c) Imagem sementes; (d) Imagem binária com os objetos separados.

água pela linha com o fim de separar as montanhas, daí o nome divisor de águas.

#### 2.3.4

##### Extração de Atributos

Na etapa de extração de atributos é onde fica mais marcada a diferença entre processamento e análise de imagem. Nesta etapa é onde ocorre a extração da informação dos objetos da imagem. Além disso, características dos objetos e da imagem são medidas.[116] Existem basicamente duas classes de medidas:

- (i) Medidas de campo, que são as medidas feitas na imagem como um todo e
- (ii) Medidas de região, que são medidas feitas independentemente em cada objeto.

Algumas das medidas de campo usadas com maior frequência são:

- ▷ Contagem de objetos;
- ▷ Área total de objetos;
- ▷ Fração de área;
- ▷ entre outras.

A contagem de objeto consiste em contar as regiões de pixels contíguos com a mesma tonalidade, que correspondem aos objetos numa imagem binária. Este tipo de medida é o mais fácil de ser implementado, porém problemas de conectividade podem eventualmente gerar erros (Figura 2.26).

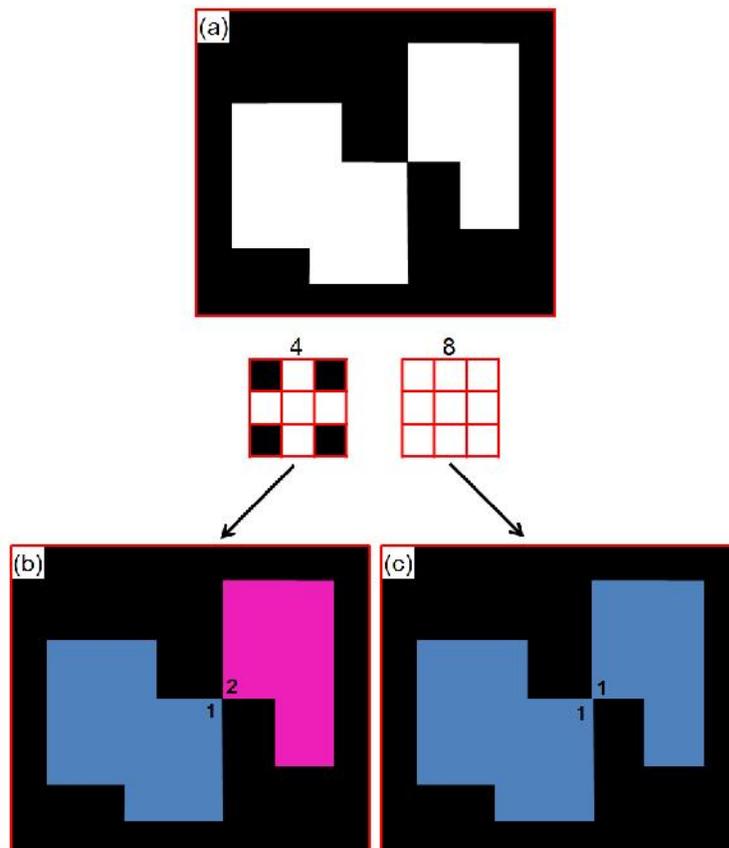


Figura 2.26: Separação de objetos numa imagem binária dependendo do tipo de conectividade entre pixels: (a) Imagem binária inicial; (b) Objetos separados com conectividade 4; (c) Objetos não separados com conectividade 8.[117]

Este problema pode ser observado na Figura 2.26, onde uma imagem binária tem dois objetos que se tocam (Figura 2.26a). Estes objetos são separados em dois objetos diferentes (1 e 2) usando uma contagem de conectividade 4 (Figura 2.26b). Em contrapartida, o uso de uma conectividade 8 leva a contagem de um objeto só (Figura 2.26c). A escolha de qual conectividade usar vai depender somente do problema em questão.

No caso da área total dos objetos, também é uma medida fácil de se implementar, já que ela se baseia na contagem dos pixels brancos na imagem binária. Esta área serve para calcular a fração de área do campo ocupada por objetos. Para isto, calcula-se a razão entre o número de pixels brancos (área total dos objetos) e o número total de pixels (área da imagem) na imagem binária. Como pode ser notado, a fração de área do campo é adimensional com valores entre 0 e 1.

Da mesma forma, nesta etapa são usadas medidas de região, tais como:

- ▷ Área;
- ▷ Posição (Equação (2-48));
- ▷ Perímetro (Equação (2-49));
- ▷ *Ferets* (*feret* máximo e mínimo);
- ▷ Razão de aspectos (Equação (2-50));
- ▷ Fator de forma circular (Equação (2-51));
- ▷ Circularidade (Equação (2-52));
- ▷ Convexidade (Equação (2-55));
- ▷ Solidez (Equação (2-56));
- ▷ entre outras.

A Área ( $A$ ), como nas medidas de campo, é calculada simplesmente contando o número total de pixels de cada objeto da imagem binária. Porém existem outras variantes de áreas, como área preenchida e área convexa ( $A_c$ ). A área preenchida é calculada da mesma forma que a área, só que neste caso, também contam-se os pixels dos poros internos aos objetos. Por sua vez a área convexa consiste na área obtida após tornar o objeto convexo. Esta medida é equivalente à área definida por um “elástico” passado em torno do objeto.

A posição de um objeto numa imagem binária pode ser descrita por várias formas diferentes, umas mais simples de serem calculadas que as outras. Por exemplo, as coordenadas de seu centro de gravidade e as medidas relativas ao retângulo circunscrito, ou seja, as maiores e menores coordenadas dos pixels nas direções horizontais e verticais. Não entanto, as medidas relativas ao retângulo circunscrito podem ser muito sensíveis a possíveis rotações dos objetos, assim as coordenadas do centro de gravidade ( $CG_x, CG_y$ ) descrevem melhor a posição do objeto.[116] Ou seja:

$$CG_x = \frac{\sum_{i=1}^n x_i}{Area} \quad (2-48)$$

$$CG_y = \frac{\sum_{i=1}^n y_i}{Area}$$

onde  $x_i$  e  $y_i$  são as coordenadas do  $i$ -ésimo pixel e  $Area$  é o número total de pixels brancos do objeto.

Por outro lado, o perímetro pode ser calculado somando o número de “passos” na horizontal, vertical ou diagonal dados de cada pixel para o seguinte até percorrer todo o contorno do objeto da imagem binária. Ou seja:

$$P = N + \sqrt{2}N_d \quad (2-49)$$

onde  $N$  é o número de passos horizontais ou verticais e  $N_d$  é o número de passos diagonais. No caso do perímetro, o perímetro dos poros internos do objeto também são considerados. Em contrapartida, no caso do perímetro preenchido só é considerado o perímetro do contorno do objeto. Finalmente no perímetro convexo ( $P_c$ ) só é considerado o perímetro do “elástico” passado em torno do objeto.

Um exemplo do cálculo do perímetro pode ser observado na Figura 2.27. Neste caso  $N = 12 + 4 = 16$  passos e  $N_d = 16 + 4 = 20$  passos, então  $P = 16 + 20\sqrt{2} = 44,3$  passos. Desta forma pode-se dizer que, segundo a Equação (2-49), o perímetro do objeto é de 44,3 pixel.

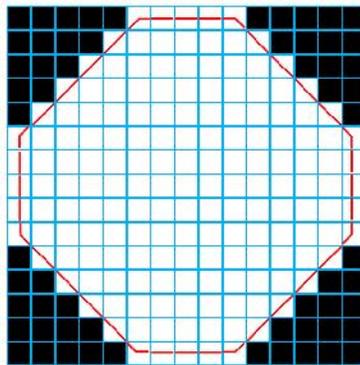


Figura 2.27: Cálculo do perímetro de um objeto através dos passos horizontais, verticais e diagonais.

Os *ferets* máximos e mínimos equivalem às projeções máximas e mínimas do objeto, caracterizando assim sua dimensão externa. Por sua vez, a razão entre o *feret* mínimo e *feret* máximo é uma boa medida do alongamento do objeto. Ou seja:

$$RA = \frac{F_{min}}{F_{max}} \quad (2-50)$$

onde  $F_{min}$  é o *feret* mínimo e  $F_{max}$  é o *feret* máximo.

Por sua vez, o fator de forma circular parte da relação que existe entre o perímetro e a área de um círculo ( $P^2 = 4\pi A$ ). Esta relação só será uma igualdade para o caso do círculo, já para qualquer outra forma geométrica do objeto teremos que  $P^2 > 4\pi A$ . Desta relação surge o fator de forma circular como:

$$FFC = \frac{4\pi A}{P^2} \quad (2-51)$$

que é uma boa medida da suavidade dos contornos dos objetos, devido a sua forte dependência do perímetro. Onde  $A$  é a área preenchida e  $P$  é o perímetro obtido pela Equação (2-49). Caso se deseje um fator de forma circular mais sensível ao alongamento do objeto e menos dependente da suavidade do contorno, então basta substituir  $P^2$  por  $(\pi F_{max})^2$ , como:

$$FFC_c = \frac{4A}{\pi(F_{max})^2} \quad (2-52)$$

onde  $A$  é a área preenchida e  $F_{max}$  é o *feret* máximo. Este novo fator de forma circular é chamado de circularidade ( $FFC_c$ ). Outras variantes do fator de forma circular são:

$$FFC_m = \frac{4A}{P(F_{max})} \quad (2-53)$$

e

$$FFC_g = \frac{16A^2}{\pi P(F_{max})^3} \quad (2-54)$$

onde  $FFC_m$  é chamado de fator de forma circular modificado.[118] Da mesma forma,  $FFC_g$  é chamado de fator de forma circular grum.[119]

Finalmente os objetos podem ser definidos por sua concavidade através da convexidade ( $C$ ) e da solidez ( $S$ ), de modo que:

$$C = \frac{P_c}{P} \quad (2-55)$$

e

$$S = \frac{A}{A_c} \quad (2-56)$$

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

Como já deve ter sido notado, os fatores de forma são adimensionais derivados das medidas geométricas básicas (área, perímetro, *ferets*, etc). Eles geralmente tem valores entre 0 e 1, sendo 1 para formas padrões (geométricas regulares) e valores menores para formas irregulares. No caso do  $FFC$  vale 1 para objetos circulares e apresenta valor menor para objetos com outras formas.

A fim de caracterizar a forma de um objeto a partir do conceito de alongamento são utilizados alguns fatores adimensionais de forma mais genéricos, que não visam comparar sua forma a um modelo específico. Um destes fatores de forma é a  $RA$ , a mesma é inversamente proporcional ao alongamento, sendo igual a 0 quando ele tende ao infinito.

Também é possível descrever um objeto como côncavo ou convexo com os fatores de forma  $C$  e  $S$ , o primeiro dependente do perímetro e o segundo da área. Estes fatores de forma têm valores iguais a 1 quando o objeto é convexo e diminuem com a presença da concavidade. Ao mesmo tempo, por depender do perímetro,  $C$  é mais sensível à suavidade do contorno que  $S$ .

Na Figura 2.28 pode-se observar a influência das propriedades de alongamento e suavidade do contorno em alguns fatores adimensionais de forma.[116] Nota-se que os objetos da esquerda (A, C e E) são facetados, enquanto que os à direita (B, D e F) possuem contornos mais suaves. Assim, os objetos A, C e E apresentaram valores próximos a 0,32 e 0,72 no fator de forma circular e convexidade, respectivamente. Os objetos B, D e F apresentaram valores próximos a 0,47 e 0,81, para esses mesmos parâmetros. Além disso, o alongamento dos objetos diminui de cima para baixo, de modo que há claramente três níveis de razão de aspecto, característicos para os pares A-B, C-D, e E-F.[74, 116]

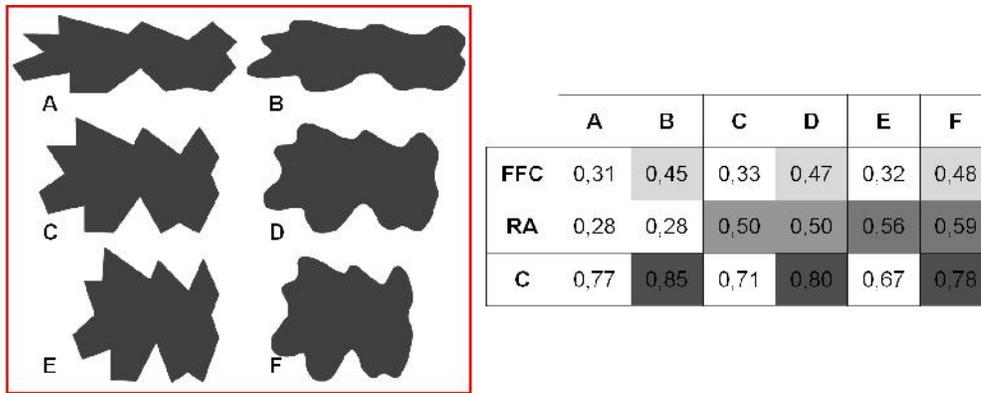


Figura 2.28: Comparação de fatores adimensionais de forma.[116]

### 2.3.5 Reconhecimento e Classificação

A etapa de reconhecimento e classificação é a etapa final da sequência padrão de PADI. Durante essa etapa é realizado o tratamento dos dados quantitativos obtidos na etapa anterior, interpretando-os, de modo a fornecer um resultado de mais alto nível, similar ao processo de reconhecimento de padrões realizado pelo cérebro humano.

Segundo o dicionário Houaiss, a definição da palavra “reconhecer” é tomar conhecimento de novo ou em outra situação, distinguir os traços característicos de alguém ou algo; caracterizar, identificar, distinguir (alguém ou algo) por certos caracteres.[120] 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.[30]

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.[121] O reconhecimento de padrões por computadores é parte importante da inteligência artificial. Estes sistemas bem podem ser usados por uma pessoa para restringir sua atenção a um conjunto de casos selecionados pelo sistema, ou bem para automatizar completamente o processo de tomada de decisão, sem necessidade de intervir pessoalmente.[30]

O crescimento acelerado e disponibilidade de capacidade computacional tornam mais rápido o processamento de grandes conjuntos de dados, além de facilitar o uso de métodos elaborados e variados para análise e classificação destes dados. Computadores, às vezes, são melhores classificadores que as pessoas,

devido a que eles não se distraem com variações aleatórias de parâmetros não críticos e conseguem extrair comportamentos estatísticos significativos para isolar grupos.[116]

Em contrapartida, as pessoas geralmente ainda são mais rápidas no reconhecimento de padrões que os computadores.[116] Ao mesmo tempo, a quantidade de dados a serem processados vem aumentando com o tempo, o que tem gerado também um aumento da demanda por melhorias no desempenho das técnicas de reconhecimento de padrões, tanto em velocidade como em exatidão e custo.

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.[122, 123] O problema de reconhecimento é colocado como uma tarefa de classificação ou categorização dos padrões.

Por outro lado, as classes são definidas como um conjunto de padrões semelhantes entre si. Estas classes podem ser identificadas como regiões do espaço de características onde é muito provável encontrar pontos representativos dos objetos. Isto significa que quando são extraídas as características de um conjunto de objetos pertencentes à mesma classe, os pontos representativos destes no espaço de características devem agrupar-se mantendo entre si distâncias menores do que aquelas que se medem em relação a objetos pertencentes a outras classes. Desta forma são formados agrupamentos, aglomerados ou *clusters*. [30]

Os sistemas de classificação são divididos, conforme a participação ou não do usuário, em supervisionados ou não supervisionados (Figura 2.29).[123] Os sistemas de classificação supervisionados são aqueles em que o usuário define as classes rotulando objetos que lhe são apresentados. Neste caso, as classes são conhecidas ou previamente determinadas pelo usuário do sistema de classificação. Os sistemas de classificação não-supervisionados são aqueles em que a semelhança entre os padrões é estabelecida pelo sistema, o qual identifica quantas e quais são as classes. Este tipo de sistema é usado quando as classes são desconhecidas ou precisa-se de um sistema de classificação sem a intervenção do usuário.[52]

A Figura 2.29a ilustra um sistema supervisionado. Nela pode-se observar um conjunto de objetos representados em um espaço de características bidimensional (razão de aspecto vs. fator de forma circular). Como a legenda indica, estes objetos são conhecidos e foram previamente rotulados (classe 1, classe 2, classe 3 e classe 4). Tanto a morfologia dos objetos quanto o espaço de características bidimensional foram escolhidos intencionalmente assim. Desta

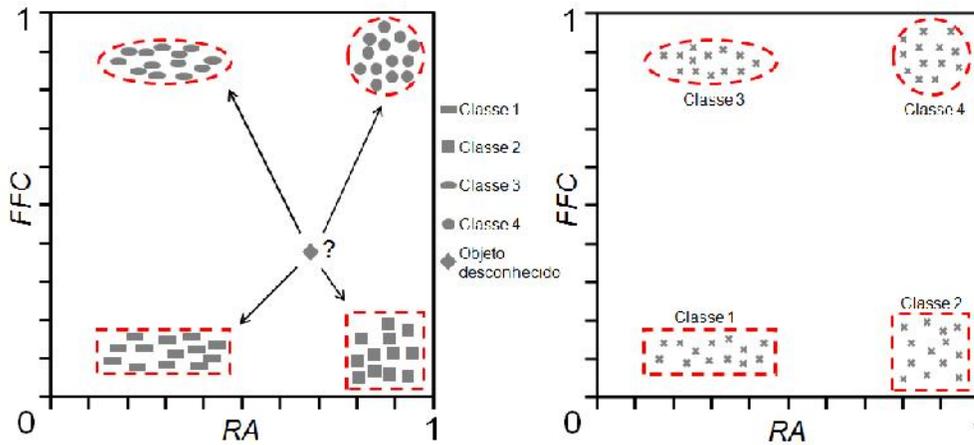


Figura 2.29: Tipos de sistemas de classificação utilizando os parâmetros característicos FFC e RA: (a) Classificação supervisionada; (b) Classificação não-supervisionada.

maneira o objeto desconhecido fica mais próximo da classe 2 que do resto, como indicam os comprimentos das setas (distância à classe 2 < distância à classe 1 < distância à classe 4 < distância à classe 3).

Assim, a partir desta base de conhecimento, pode ser atribuída uma classe a um objeto desconhecido através de um procedimento de classificação supervisionada. Portanto, o objeto desconhecido é classificado como pertencente à classe 2. O critério utilizado é a proximidade, no espaço de características, aos objetos desta classe.

A Figura 2.29b mostra a classificação não-supervisionada. Neste caso, os padrões dos objetos são desconhecidos. Contudo, um procedimento de classificação não-supervisionada procura por objetos similares e os agrupa em classes. Neste caso, apesar de não conhecer a morfologia dos objetos, o classificador consegue separá-los em quatro classes diferentes, como no exemplo anterior. Contudo, os sistemas não-supervisionados escapam do escopo deste trabalho.

### 2.3.5.1

#### Conjunto de treinamento

O conjunto de treinamento é um conjunto de objetos previamente identificado pelo usuário. Este conjunto de treinamento deve ser estatisticamente grande, com o intuito de aumentar a precisão do classificador, além de apresentar grande variabilidade dentro de cada classe para assim poder garantir uma boa representatividade.

Como poderá ser observado mais na frente, o conjunto de treinamento

precisa crescer exponencialmente com o aumento da dimensionalidade do espaço de características.[124] Outros autores definem que o conjunto de treinamento deve aumentar em dez vezes por classe, com o aumento da dimensionalidade do espaço de características.[123] Contudo, nem sempre é possível obter mais objetos conhecidos para ampliar o conjunto de treinamento.

### 2.3.5.2

#### 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. O conjunto de atributos que define o espaço de características deve caracterizar bem os objetos, agrupando os objetos similares e separando os distintos. Este conjunto deve ser sensível o bastante para discriminar todas as classes.

Raramente uma única característica separa mais de duas classes, sendo geralmente utilizadas várias características. Porém, um número de características muito grande em relação ao tamanho do conjunto de treinamento pode acarrear problemas de dimensionalidade. Um espaço de características com dimensionalidade muito alta torna o sistema mais complexo, consumindo maior tempo no treinamento e ainda podendo reduzir a capacidade de generalização do sistema.[118] Assim, é de crucial importância realizar uma cuidadosa escolha do conjunto de características.

Na verdade, a alta dimensionalidade é um problema bastante recorrente no reconhecimento de padrões, chegando a ser chamada de “maldição da dimensionalidade”. Assim, várias técnicas foram desenvolvidas com o intuito de reduzir a dimensionalidade do espaço de características com a esperança de obter problemas mais simples. Um exemplo disto é a Análise Discriminante Linear (LDA).[124]

Fica claro que a dimensionalidade poderia ser reduzida de  $d$  dimensões para 1 dimensão se for possível projetar as  $d$  dimensões numa linha formando *clusters* devidamente separados. Este é exatamente o objetivo da análise discriminante linear, que serve de base para as técnicas de Análise Discriminante de Fisher (FDA) e a Análise de Componentes Principais (PCA), também conhecida como Transformada de Karhunen-Loève (KLT). Estes métodos são dos mais populares usados para a redução da dimensionalidade, porém outros métodos como o Escalonamento Multidimensional (MDS) também podem ser úteis.[121, 125] No presente trabalho foram usadas as técnicas PCA e FDA.

### 2.3.5.3 Classificadores

Os classificadores são funções que utilizam como entrada os padrões desconhecidos, e como saída as classes a que estes padrões provavelmente pertencem. Os classificadores podem ser divididos em dois tipos principais: os estatísticos (paramétricos e não paramétricos) e os conexionistas (redes neurais).[52]

Os classificadores estatísticos não paramétricos são os mais simples. Eles utilizam uma função de distância como medida de similaridade, atribuindo um objeto desconhecido à classe mais próxima dele no espaço de características. Dois dos classificadores não paramétricos comumente empregados são o de Distância Euclideana e o de Distância de Mahalanobis.

Antes de entrar em detalhes acerca dos classificadores estatísticos, serão definidas algumas nomenclaturas e conceitos a serem usados. Seja  $w_1, w_2, \dots, w_W$  um padrão de classe que contem algumas propriedades em comum, onde  $W$  é o número total de classes. Da mesma forma  $\mathbf{x}$  é um padrão de características, que não é mais que um vetor de  $n$  componentes:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (2-57)$$

onde cada componente,  $x_k$ , representa a  $k$ -ésima medida de um total de  $n$  medidas (Seção 2.3.4) associadas ao padrão.

Para finalizar, será definida a função discriminante ( $d$ ) que vai separar as classes ( $w$ ) decidindo a que classe pertence um determinado padrão de características. Assim, a fronteira de separação entre as classes  $w_i$  e  $w_j$  será dada pelo valor  $\mathbf{x}$  para o qual:

$$d_i(\mathbf{x}) = d_j(\mathbf{x}) \quad (2-58)$$

ou seja,

$$d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0 \quad (2-59)$$

Assim, é uma prática comum identificar a fronteira de separação entre duas

classes como:

$$d_{ij}(\mathbf{x}) = d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0 \quad (2-60)$$

Desta forma  $d_{ij}(\mathbf{x}) > 0$  para objetos pertencentes à classe  $w_i$  e  $d_{ij}(\mathbf{x}) < 0$  para objetos pertencentes à classe  $w_j$ . Tendo claro estes conceitos, pode-se detalhar alguns classificadores de interesse.

O classificador de Distância Euclideana, também conhecido como classificador de distância mínima, consiste em primeiramente calcular a média da classe  $w_i$ :

$$\bar{\mathbf{x}}_i = \frac{1}{N_i} \sum_{i=1}^W \mathbf{x}_i \quad (2-61)$$

onde  $N_i$  é o número total de padrões de características pertencentes à classe  $w_i$ . Assim, para determinar a classe a qual pertence um padrão de características desconhecido  $\mathbf{x}$ , calcula-se a distância mínima ou distância euclideana da seguinte forma:

$$D_i(\mathbf{x}) = \|\mathbf{x} - \bar{\mathbf{x}}_i\| \quad (2-62)$$

então:

$$D_i^2(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}}_i)^T (\mathbf{x} - \bar{\mathbf{x}}_i) \quad (2-63)$$

Deste modo, o objeto será atribuído à classe mais próxima dele. Fazendo uso da definição de função discriminante e da Equação 2-63, após algumas simplificações temos que:

$$d_i(\mathbf{x}) = \mathbf{x}^T \bar{\mathbf{x}}_i - \frac{1}{2} \bar{\mathbf{x}}_i^T \bar{\mathbf{x}}_i \quad (2-64)$$

Finalmente, das Equações 2-60 e 2-64, a função discriminante que vai separar as classes  $w_i$  e  $w_j$  através do classificador de distâncias euclideana é:

$$d_{ij}(\mathbf{x}) = \mathbf{x}^T (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) - \frac{1}{2} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)^T (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) = 0 \quad (2-65)$$

Como pode ser observado, o classificador de distância euclideana só dá garantias de bons resultados caso todas as classes tenham a mesma variância

em todas as características. É assim que o classificador de distância de Mahalanobis, embora seja muito parecido ao de distância euclídeana, atende melhor este problema. Para isto, ele normaliza a distância subtraindo cada característica por sua média e dividindo-a pelo desvio padrão, segundo:

$$\bar{x} = \frac{|x - \mu|}{\sigma} \quad (2-66)$$

Desta modo, generalizando para um padrão de características, obtém-se a distância de Mahalanobis de forma similar à Equação 2-63, ou seja:

$$D_i^2(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}}_i)^T \Sigma_i^{-1} (\mathbf{x} - \bar{\mathbf{x}}_i) \quad (2-67)$$

onde  $\Sigma_i$  é a matriz de covariância da classe  $w_i$ , definida como:

$$\Sigma_i = \frac{1}{(N_i - 1)} \sum_{i=1}^W (\mathbf{x} - \bar{\mathbf{x}}_i) (\mathbf{x} - \bar{\mathbf{x}}_i)^T \quad (2-68)$$

onde  $N_i$  é o número total de padrões de características pertencentes à classe  $w_i$ .

Um dos classificadores estatísticos paramétricos mais usados é o classificador de Bayes, também conhecido como classificador bayesiano. Este consiste em minimizar a probabilidade média de erro total na classificação.[90] Assim, seja  $P(w_i)$  a probabilidade a priori de ocorrência da classe  $w_i$  ( $i = 1, 2, \dots, W$ ), tal que:

$$\sum_{i=1}^W P(w_i) = 1 \quad (2-69)$$

e  $p(\mathbf{x}|w_i)$  a densidade de probabilidade de uma características  $\mathbf{x}$  que vem de uma classe  $w_i$  conhecida, então  $p(\mathbf{x})$  seria a densidade de probabilidade de uma características  $\mathbf{x}$  que vem de uma classe desconhecida. Assim temos que:

$$p(\mathbf{x}) = \sum_{i=1}^W p(\mathbf{x}|w_i)P(w_i) \quad (2-70)$$

Por sua vez, conhece-se da teoria clássica das probabilidades que:

$$P(w_i|\mathbf{x}) = \frac{p(\mathbf{x}|w_i)P(w_i)}{p(\mathbf{x})} \quad (2-71)$$

Assim, as regras de classificação de Bayes podem começar a ser definidas como:

$$\begin{aligned} \text{Se } P(w_1|\mathbf{x}) > P(w_2|\mathbf{x}), \quad \mathbf{x} \text{ pertence à classe } w_1 \\ \text{Se } P(w_1|\mathbf{x}) < P(w_2|\mathbf{x}), \quad \mathbf{x} \text{ pertence à classe } w_2 \end{aligned} \quad (2-72)$$

Usando a Equação 2-71, as regras ficam como:

$$\begin{aligned} \text{Se } p(\mathbf{x}|w_1)P(w_1) > p(\mathbf{x}|w_2)P(w_2), \quad \mathbf{x} \text{ pertence à classe } w_1 \\ \text{Se } p(\mathbf{x}|w_1)P(w_1) < p(\mathbf{x}|w_2)P(w_2), \quad \mathbf{x} \text{ pertence à classe } w_2 \end{aligned} \quad (2-73)$$

onde  $p(x)$  não foi levado em consideração já que é o mesmo para todas as classes e isto não afeta à decisão. Assim, finalmente, podemos definir a função discriminante de Bayes como:

$$d_i(\mathbf{x}) = p(\mathbf{x}|w_i)P(w_i) \quad (2-74)$$

Estes são alguns exemplos de classificadores estatísticos comumente utilizados. No final, a escolha de tipo de classificador vai depender do problema em questão e do que estiver disponível no software.

Por outro lado, os classificadores conexionistas são redes neurais artificiais. As redes neurais constituem classificadores mais complexos, inspirados na estrutura neural de organismos inteligentes, que aprendem por experiência.[52] Classificadores baseados nas redes neurais possuem capacidade de se ajustar a qualquer topologia de classes, o que os tornam muito eficientes nos casos onde há classes com distribuições mal-comportadas.[90] Contudo, para os problemas de classificação deste trabalho, os classificadores estatísticos são eficazes. Deste modo, as redes neurais não fazem parte do escopo deste trabalho.