



**Centro Brasileiro de Pesquisas Físicas - CBPF**

ANDRÉ PERSECHINO AMÉRICO DE OLIVEIRA

INFLUÊNCIA DA DIFUSÃO NÃO-LINEAR NA DETECÇÃO DE  
CALCIFICAÇÕES EM IMAGENS MAMOGRAFICAS

Rio de Janeiro - RJ  
2015



Centro Brasileiro de Pesquisas Físicas - CBPF

# INFLUÊNCIA DA DIFUSÃO NÃO-LINEAR NA DETECÇÃO DE CALCIFICAÇÕES EM IMAGENS MAMOGRÁFICAS

André Persechino Américo de Oliveira

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Física pelo programa de Física com ênfase em Instrumentação Científica do Centro Brasileiro de Pesquisas Físicas.

**Orientador:** Márcio Portes de Albuquerque

Rio de Janeiro - RJ  
2015

ANDRÉ PERSECHINO AMÉRICO DE OLIVEIRA

INFLUÊNCIA DA DIFUSÃO NÃO-LINEAR NA DETECÇÃO DE  
CALCIFICAÇÕES EM IMAGENS MAMOGRÁFICAS

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Física pelo programa de Física com ênfase em Instrumentação Científica do Centro Brasileiro de Pesquisas Físicas.

APROVADA: 04 de Setembro de 2015.

---

Prof. Márcio Portes de Albuquerque  
(Orientador)  
(CBPF)

---

Prof. Marcelo Lobato Martins  
(UFV)

---

Prof. Sérgio José Barbosa Duarte  
(CBPF)

# Agradecimentos

A lista das pessoas que contribuíram direta ou indiretamente para a realização deste trabalho é imensa e não seria razoável colocá-la aqui. Contudo, algumas pessoas e instituições não poderiam deixar de ser citadas: agradeço à Universidade Federal de Viçosa por me proporcionar a formação básica em ciência e me abrir o caminho a ser trilhado. Em particular, agradeço a Oswaldo Monteiro Del Cima e Álvaro José Magalhães Neves. Professor Álvaro: obrigado por acreditar em mim.

Agradeço aos amigos Marcos Costa e Silva e Juliana Assunção, com quem tive a sorte de conviver. Vocês são os melhores companheiros de laboratório que alguém pode ter. À Equipe de Judo da UFV e aos colegas de alojamento, minha gratidão.

Este trabalho não seria realizado sem a ajuda valiosa dos integrantes do Laboratório de Física Aplicada ao Radiodiagnóstico da UNESP de Botucatu e dos radiologistas Sergio Marrone Ribeiro e André Trindade. Agradeço também a Francisco Abdala do Instituto Nacional de Câncer pela boa vontade e iniciativa no início deste projeto. A José Apolinário Jr., do Instituto Militar de Engenharia, agradeço pelos ensinamentos em teoria de sinais.

Agradeço o Centro Brasileiro de Pesquisas Físicas pela oportunidade de desenvolver as atividades deste trabalho. Em particular, agradeço Sérgio Barbosa Duarte por ter viabilizado o contato com a equipe da UNESP e a Yuri Müller, cujo *insight* descompromissado ajudou bastante na computação do projeto. Agradeço a Márcio Portes de Albuquerque pela amizade, confiança e orientação. À Coordenação de Atividades Técnicas do CBPF: obrigado pelo bom convívio! Embora não tenha me tornado carioca como planejaram, sei que vocês agora gostam mais dos paulistas!

Agradeço à família Ushimaru pela confiança e amizade entregues ao longo destes quase vinte anos. Agradeço também a Bruna Barbosa Lopes, pelos incentivos a ingressar na física. Se não fosse seu apoio, tudo isso ainda seria uma idéia. Meus agradecimentos também à família Barbosa pela amizade e respeito.

Agradeço a minha família por ter me apoiado sempre. A meu irmão Flávio Américo Persechino de Oliveira e a Luis Fernando Pena apresento este trabalho como forma de mostrar que o esforço valeu a pena: mais um passo dado!

Por fim, dedico este trabalho a Ana Maria Persechino, a leoa que foi além da definição e conseguiu fazer o impossível para que seus filhos pudessem crescer em segurança. Ainda que não esteja mais neste mundo, a lembrança de seu amor e de seu esforço continuarão em mim. Cada vitória que vier, será sua também.

A seed can lie dormant for hundreds and thousands of years. Dried and frozen in a glacier. Place it in moist earth and it sprouts. There is a power. The green force of the flower. What is the power of the human being?

[...]

In a rapid succession of images he saw a blade of grass break through the seed casing and struggle toward the surface, saw a sunflower strive toward the sky, turning to the light, saw a small child pull itself to its feet, hold its arms out, jubilant, and everything lives and is drawn to the light, and he saw...

[...]

Everything is effort, work.

(John Ajvide Lindqvist)

## Resumo

É descrito neste trabalho o desenvolvimento de um sistema computacional dedicado à detecção de (micro) estruturas em imagens mamográficas. O núcleo deste sistema baseia-se no processo de difusão não-linear modelado pela Equação de Perona-Malik, cuja estrutura permite realizar o realce e segmentação de objetos sem comprometer o conteúdo relevante da imagem. Profunda revisão dos métodos clássicos de processamento de imagens é feita e todo o aparato auxiliar necessário desenvolvido é descrito, tais como as etapas de segmentação, realce, contagem e caracterização. Por fim, é feita análise da influência direta da difusão não-linear sobre os indicadores característicos das detecções: é mostrado que, embora apresente incremento médio relativo de até 35% nas eficiência das detecções, a incorporação da Equação de Perona-Malik segundo a cadeia de processamento proposta ainda demanda aperfeiçoamentos específicos relativos à utilização de métodos numéricos mais eficientes, melhor determinação de seus parâmetros e, principalmente, segmentação.

## **Abstract**

It is described in this work the development of a computational system dedicated to detection of (micro) structures on mammographic images. The kernel of the system is based on the nonlinear diffusion process modeled by the Perona-Malik Equation, whose structure allows to perform both enhancement and segmentation in images without degrading its meaningful content. Extensive review on the classic image processing techniques is performed and all the development of auxiliary apparatus is described, such as the steps of segmentation, enhancement, counting and characterization. Lastly, an analysis of direct influence of nonlinear diffusion on the detection's indicators is carried out, showing that despite improving relative mean detection efficiency up to 35%, the use of Perona-Malik Equation in the proposed chain still demands specific improvements related to efficient numerical methods, better derivation of the Equation's parameters and, mainly, segmentation.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>Mamografia</b>	<b>4</b>
2.1	Câncer de mama do ponto de vista de saúde pública . . . . .	4
2.2	Produção de raios-X . . . . .	5
2.3	Interação dos raios-X com a matéria . . . . .	9
2.4	Imagem mamográfica . . . . .	12
<b>3</b>	<b>Processamento de imagens</b>	<b>17</b>
3.1	Conceitos introdutórios . . . . .	17
3.2	Filtragem . . . . .	22
3.3	Segmentação de imagens . . . . .	35
3.4	Operações de alto nível: labeling e clustering . . . . .	40
3.5	Pós-processamento: morfologia matemática . . . . .	52
<b>4</b>	<b>Modelos derivados de Equações Diferenciais</b>	<b>58</b>
4.1	Problemática do realce em imagens . . . . .	58
4.2	Descrição multiescala, filtragem Gaussiana e equação do calor . . . . .	59
4.3	Equação de Perona-Malik . . . . .	63
<b>5</b>	<b>Metodologia</b>	<b>76</b>
5.1	Materiais . . . . .	76
5.2	Algoritmos desenvolvidos . . . . .	79
<b>6</b>	<b>Resultados</b>	<b>103</b>
6.1	Efetividade da técnica de segmentação das ROIs . . . . .	103
6.2	Mascaramento da ROI pelo preenchimento morfológico . . . . .	105
6.3	Análise dos thresholds . . . . .	109
6.4	Influência da difusão não-linear . . . . .	111
<b>7</b>	<b>Conclusões</b>	<b>121</b>
	<b>Bibliografia</b>	<b>124</b>
<b>A</b>	<b>Demonstração do resultado para o PVI da condução do calor</b>	<b>131</b>
<b>B</b>	<b>Obtenção do threshold para binarização</b>	<b>133</b>

# Capítulo 1

## Introdução

A mamografia é tida como o método mais eficiente para a detecção do câncer de mama, que representa a um problema sério de saúde pública em todo o mundo. Dentre os diversos sinais radiológicos relevantes para diagnóstico do câncer de mama, a presença de (micro) calcificações constitui um desafio importante, já que calcificações malignas tipicamente têm dimensões micrométricas. Em adição à baixa dimensionalidade das calcificações, há o fato de que, devido a características do exame e da mama, a imagem mamográfica apresenta tipicamente baixo contraste das estruturas retratadas. Nesse sentido, processamento de imagens pode ser empregado com vistas a auxiliar nas detecções.

Dentre a miríade de técnicas de processamento existentes, a classe de modelos derivados de equações diferenciais parciais (EDPs) chama atenção por sua abrangência: com elas é possível realizar realce, morfologia matemática e segmentação. Em particular, a Equação de Perona-Malik apresenta características desejáveis no processamento de imagens mamográficas, já que evidencia bordas e uniformiza regiões entre-bordas. Nesse sentido, sua implementação pode, teoricamente, auxiliar nas detecções de estruturas suspeitas em mamogramas.

O objetivo principal deste trabalho consiste em analisar a influência da incorporação da Equação de Perona-Malik em uma cadeia de processamento dedicada à detecção de (micro) calcificações. Para embasar o desenvolvimento da cadeia de processamento, é feita inicialmente uma revisão das técnicas básicas de filtragem, seguida de estudo das características fundamentais da Equação de Perona-Malik.

A dissertação está organizada segundo a seguinte sequência:

No Capítulo 2 são apresentadas informações básicas sobre mamografia e câncer de mama. Enfoque é dado no entendimento do processo de interação dos raios-X com a matéria e formação da imagem mamográfica, de forma que se consiga fazer compreender as dificuldades intrínsecas relativas à detecção em mamogramas (baixo contraste e baixa dimensionalidade dos achados). Os conceitos são expostos de forma bastante resumida, apontando os pontos principais e construindo uma visão geral do assunto. Para aqueles que desejam aprofundamento nos tópicos de física médica, sugere-se os livros de Podgoršak [66] e Hobbie e Roth [37]. Para revisões sobre o desenvolvimento da mamografia, sugere-se os trabalhos de Law [49] e James [44].

No Capítulo 3 são apresentados os conceitos fundamentais da área de processamento de imagens. Isso se faz necessário para que as técnicas utilizadas no trabalho sejam minimamente contextualizadas. Com base nessa premissa, os conceitos são apresentados de forma sucinta, pois constituem tópicos largamente discutidos na literatura. Inicialmente, são apresentados os conceitos básicos sobre representação de uma imagem digital, seguido

de uma discussão de técnicas usuais de filtragem e, finalmente, o capítulo é encerrado com discussões sobre processamento de informações de alto nível (labeling e clustering) e pós-processamento via morfologia matemática. Para aqueles que desejam aprofundamento nos tópicos discutidos, um curso completo de processamento de sinais pode ser encontrado em Diniz *et al.* [27] e um curso específico em processamento de imagens pode ser encontrado no livro de Pedrini e Schwartz [62].

No Capítulo 4 são apresentados os conceitos da descrição multiescala de sinais, arcabouço teórico da Equação de Perona-Malik, que corresponde à técnica mais profundamente estudada neste trabalho. O capítulo é iniciado com a colocação da problemática em realce de imagens, em que são apresentados os requisitos de um sistema de realce dito ideal. Na sequência, é apresentada a abordagem de descrição multiescala de sinais e sua relação com a filtragem Gaussiana e equação do calor. Uma vez apresentadas as características da descrição multiescala linear de sinais, a Equação de Perona-Malik é introduzida como um recurso não-linear de aperfeiçoamento do processo de realce. Questões de cunho prático, como discretização das equações e escolha dos parâmetros são discutidas na sequência. Por fim, são feitas breves considerações sobre a ausência de solução formal da Equação de Perona-Malik.

No Capítulo 5 é realizada a descrição completa do desenvolvimento da cadeia de processamento utilizada neste trabalho. Este capítulo baseia-se massivamente nos conceitos estudados no Capítulo 3 e deve ser interpretado como a parte descritiva do trabalho referente ao desenvolvimento do instrumental utilizado. Detalhes de ordem prática, tais como escolha e definição da base de imagens mamográficas a ser analisada e recursos computacionais também são abordados. Por fim, o capítulo é encerrado com a apresentação da metodologia desenvolvida para contagem e caracterização dos achados, ilustrando-a com exemplos variados.

O Capítulo 6 traz a descrição completa dos experimentos executados e faz uma análise dos diversos indicadores das detecções em ambos os cenários: com e sem enhancement via difusão não-linear. Os resultados evidenciados neste capítulo servem de base para que no Capítulo 7 as conclusões do trabalho sejam apresentadas.

Por fim, há dois apêndices: o Apêndice A traz a demonstração do problema de valor inicial para equação do calor em uma dimensão, provando a relação entre filtragem Gaussiana e difusão linear. Por sua vez, o Apêndice B traz uma apresentação breve dos thresholds usados na segmentação das imagens, bem como apresenta um código computacional para a obtenção destes.

Embora esse texto represente um trabalho de mestrado, o autor quis – na medida do possível – apresentar os conteúdos de forma didática, principalmente aqueles referentes a processamento de imagens e à Equação de Perona-Malik. Embora isso tenha refletido no volume do texto, o autor acredita que o material de revisão aqui contido possa ajudar estudantes iniciantes na área de processamento de imagens que, frequentemente, se deparam com um conjunto imenso de informações e acabam se desestimulando a prosseguir seus estudos nessa área tão abrangente.

## Capítulo 2

# Mamografia

### 2.1 Câncer de mama do ponto de vista de saúde pública

O que denominamos genericamente como câncer corresponde na verdade a um conjunto grande de doenças que têm em comum o crescimento anormal e descontrolado de células em determinado organismo. Esse crescimento desordenado pode ocorrer em diversas partes do organismo e de maneiras distintas. De um modo geral, o câncer corresponde a um grave problema de saúde pública em âmbito mundial, mesmo com avanços significativos em prevenção, diagnóstico e tratamento [14]. O Brasil não é, dentro do contexto global, uma exceção e apresenta índices significativos de incidência de câncer. Estimativas do Instituto Nacional de Câncer (INCA) apontaram para o biênio 2014/2015 a ocorrência de aproximadamente 576 mil novos casos de câncer [14]. Nesse mesmo relatório, é mostrado que, à exceção do câncer de pele não-melanoma, o câncer de mama seria o mais incidente sobre a população feminina.

Frente a essa realidade, políticas têm sido propostas no sentido de combater o câncer em diversas linhas de enfrentamento, tais como prevenção, detecção e tratamento. Em particular, temos no Brasil a Política Nacional de Atenção Oncológica, parte integrante da Portaria do Ministério da Saúde número 2 048 de 2009, que aprova o regulamento do Sistema Único de Saúde (SUS) [12]. A Política Nacional de Atenção Oncológica prevê ações integradas de combate ao câncer e alguns dos objetivos apontados versam sobre universalização do atendimento ao paciente oncológico; aperfeiçoamento de técnicas e métodos e incorporação/geração de novas tecnologias.

Dentre os diversos tipos de câncer existentes, o câncer de mama vem, há tempos, chamando a atenção de governos e sociedade, dado que por si só representa um problema de saúde pública de grande magnitude. Segundo documento de consenso elaborado pelo INCA [13], as intervenções no âmbito de saúde pública no Brasil para o câncer de mama devem ser direcionadas à detecção precoce, dado que há dificuldades em implementação de ações de prevenção. Várias condutas foram propostas, como rastreamento de população assintomática, políticas educacionais etc. De um modo geral, nota-se forte referência quanto à importância da mamografia nas etapas de detecção do câncer de mama, sendo essa modalidade diagnóstica tida como a forma mais eficaz para detecção e identificação de anormalidades na mama.

### 2.1.1 O papel da instrumentação científica

Explicitada a importância da mamografia, convém ressaltar que a *qualidade* da imagem mamográfica tem papel fundamental dentro de toda a estrutura direcionada à prevenção e combate ao câncer de mama. Diversos fatores influenciam diretamente a qualidade da imagem, desde desajustes técnico-operacionais até a ocorrência de artefatos na imagem [17]. Nesse sentido, paralelamente à formação e ao aperfeiçoamento dos recursos humanos envolvidos em mamografia, atenção deve ser dada às tecnologias associadas. Sobre a Política Nacional de Atenção Oncológica [12], é dito que esta deve

[...]fomentar, coordenar e executar projetos estratégicos de incorporação tecnológica, mediante estudos de custo-efetividade, eficácia e qualidade e avaliação tecnológica da atenção oncológica no Brasil.

Portanto, existe todo um cenário de atuação em termos de desenvolvimento tecnológico em medicina, e conseqüentemente em mamografia. Mantendo a vanguarda, a física atua com competência dentro das mais diversas vertentes da tecnologia, proporcionando ferramentas e técnicas inovadoras com vistas à melhoria dos procedimentos existentes. Apenas como exemplo, vale lembrar que a própria radiologia teve início em um laboratório de física, tal como será apresentado na Seção 2.2.

Dentro deste escopo, a instrumentação científica desempenha papel importante, uma vez que existe uma demanda contínua para a resolução de problemas. Em particular, a área de processamento de sinais e imagens vem contribuindo há anos com os mais diversos ramos da tecnologia (militar, médica, astronômica etc.). Dentro do contexto da mamografia, as técnicas de processamentos de imagens vêm sendo empregadas como recursos auxiliares em diagnóstico e detecção (CAD)<sup>1</sup>, em que modelos computacionais são propostos com vistas a auxiliar os profissionais em saúde envolvidos no diagnóstico. Esse é o caso do trabalho em questão, em que são investigados modos de detecção automática de calcificações na mama. Nas próximas seções serão apresentados os conceitos básicos da produção da imagem mamográfica e será feita breve exposição sobre as calcificações mamárias.

## 2.2 Produção de raios-X

Em 1895, W. C. Röntgen, então professor da Universidade alemã de Würzburg, conduzia experimentos sobre descargas elétricas em tubos altamente evacuados (ampola de Crookes). Em um de seus ensaios, Röntgen – que à época estudava a luz ultravioleta produzida pelas descargas por meio de sensibilização de um filme de platinocianeto de bário – notou que seus filmes apresentavam fluorescência mesmo com o tubo coberto com papel opaco. Ao inserir materiais diversos entre o tubo e o filme, Röntgen ainda observava alguma fluorescência. Desconhecendo a natureza destes agentes de fluorescência dotados de alto poder de penetração, Röntgen os denominou de raios-X.

Muitas pesquisas ocorreram no sentido de caracterizar os raios-X: investigações posteriores mostraram que os raios-X não sofrem curvatura quando expostos a campos externos e que podem ser difratados e polarizados. Com base nestas constatações, percebeu-se então que os raios-X nada mais são que radiação eletromagnética (REM), diferindo da luz visível apenas no que tange ao comprimento de onda. Nesse sentido, o aparato formal da

<sup>1</sup>Do inglês, *Computer Aided Diagnosis*.

teoria eletromagnética clássica pôde ser aplicado ao estudo dos raios-X. Entretanto, diversas constatações experimentais contradiziam as expectativas teóricas: a forma do espectro (existência de um comprimento de onda mínimo para ocorrência de raios-X e o surgimento de linhas características) e a alteração do comprimento de onda no espalhamento inelástico dos raios-X são dois exemplos. Estas questões só puderam ser elucidadas com o advento das hipóteses da mecânica quântica. A seguir serão feitas algumas exposições com vistas à compreensão da fenomenologia da produção de raios-X e de sua interação com a matéria, de forma que possamos chegar a um entendimento básico sobre a formação da imagem mamográfica.

### Análise clássica da produção de raios-X

A dinâmica da produção de raios-X pode ser compreendida com o auxílio da Figura 2.1: no tubo altamente evacuado, um catodo é aquecido e emite elétrons (emissão termo-iônica). Estes elétrons são acelerados em função do campo existente gerado pela diferença de potencial (ddp) entre os terminais do tubo<sup>2</sup> e rumam em direção a um alvo (anodo). Estes elétrons experimentam uma desaceleração ao adentrar o anodo, emitindo REM nos comprimentos de onda da região dos raios-X.

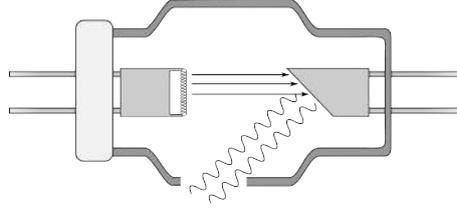


Figura 2.1: Diagrama esquemático de um tubo de raios-X: um catodo aquecido emite elétrons e estes são acelerados devido a uma diferença de potencial imposta ao tubo. Ao colidirem com o anodo, os elétrons são desacelerados, emitindo radiação. A janela do tubo é geralmente composta por material transparente aos raios-X, tal como berílio.

A descrição formal desta dinâmica pode ser realizada via teoria eletromagnética clássica, e está ilustrada na Figura 2.2: a trajetória do elétron sofre uma deflexão em função da repulsão Coulombiana do núcleo atômico com o qual interage. Essa alteração na aceleração do elétron ocasiona a emissão de REM, cujos campos elétrico e magnético (no vácuo) são dados, respectivamente por<sup>3</sup>

$$\begin{cases} \vec{E}(\vec{r}, t) &= -\frac{e}{4\pi\epsilon_0} \frac{\vec{s}}{(\vec{s}\cdot\vec{u})^3} [(c^2 - v^2)\vec{u} + \vec{s} \times (\vec{u} \times \vec{a})]; \\ \vec{B}(\vec{r}, t) &= \frac{1}{c} \hat{s} \times \vec{E}(\vec{r}, t), \end{cases} \quad (2.1)$$

em que  $e$  corresponde ao valor absoluto da carga do elétron;  $\epsilon_0$  é a permitividade elétrica no vácuo;  $\vec{s}$  é o vetor posição retardada<sup>4</sup>;  $c$  é o módulo da velocidade da luz no vácuo;

<sup>2</sup>Em radiologia diagnóstica as tensões de operação dos tubos de raios-X variam tipicamente entre 20 e 120 KV [58].

<sup>3</sup>No que segue, a norma de um vetor qualquer  $\vec{x}$  será representada apenas por  $x$ , salvo observação em contrário. Adicionalmente, o versor associado a  $\vec{x}$  será denotado por  $\hat{x}$ .

<sup>4</sup> $\vec{s}$  corresponde ao vetor posição do ponto de análise dos campos em relação à posição retardada (avaliada em tempo passado) do elétron.

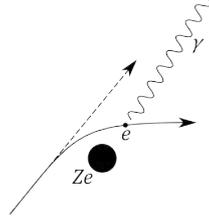


Figura 2.2: Ilustração da interação elétron-núcleo do ponto de vista clássico. A interação Coulombiana entre o núcleo e o elétron faz com que este último sofra um desvio de sua trajetória inicial (linha tracejada), levando à emissão de REM, denotada pela letra  $\gamma$ .

$\vec{v}$  e  $\vec{a}$  correspondem à velocidade e aceleração retardadas, respectivamente. O vetor  $\vec{u}$  é definido como

$$\vec{u} = c\hat{s} - \vec{v}. \quad (2.2)$$

Uma rápida manipulação na expressão de  $\vec{E}$  na Equação (2.1) leva a

$$\vec{E}(\vec{r}, t) = -\frac{e}{4\pi\epsilon_0 u^2 (\hat{s} \cdot \hat{u})^3} \left[ \frac{c^2 - v^2}{s^2} \hat{u} + \frac{1}{s} \hat{s} \times (\hat{u} \times \vec{a}) \right],$$

em que fica explícita a dependência do primeiro e segundo termos de  $\vec{E}$  com  $1/s^2$  e  $1/s$ , respectivamente. Estes termos são chamados usualmente de campos de velocidade e aceleração. No que diz respeito à *potência de radiação*, apenas o campo de aceleração contribui significativamente, uma vez que o campo de velocidade decai mais intensamente com a distância. Por esse motivo, o campo de aceleração também é chamado de *campo de radiação*:

$$\vec{E}_{\text{rad}}(\vec{r}, t) = -\frac{e}{4\pi\epsilon_0} \frac{s}{(\vec{s} \cdot \vec{u})^3} \vec{s} \times (\vec{u} \times \vec{a}). \quad (2.3)$$

A potência de radiação se dá pela integração do vetor de Poynting:

$$P = \frac{1}{\mu_0} \oint (\vec{E} \times \vec{B}) \cdot \hat{n} dS,$$

em que  $\mu_0$  corresponde à permeabilidade magnética no vácuo. Partindo da hipótese que a carga encontra-se em repouso em certo instante de tempo retardado, chega-se à conhecida potência de Larmor:

$$P = \frac{\mu_0 e^2 a^2}{6\pi c}. \quad (2.4)$$

A expressão apresentada na Equação (2.4) é válida apenas para  $v = 0$ ; contudo, esta mesma expressão corresponde a boa aproximação para casos em que o elétron não possui velocidade relativística [32]. A Equação (2.4) mostra que, a todo instante, um elétron acelerado irradia parte de sua energia, elucidando assim a questão da emissão de REM por cargas aceleradas.

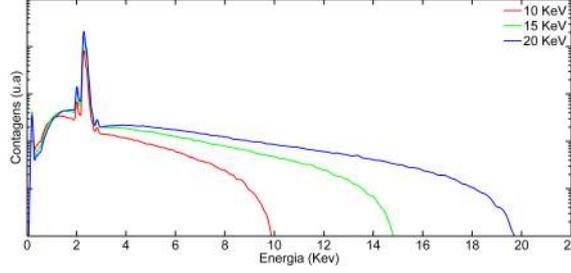


Figura 2.3: Espectro de raios-X para o elemento molibdênio ( $Z = 42$ ) segundo diversas tensões de operação. Verifica-se a existência de um comprimento de onda mínimo para ocorrência dos raios-X (apresentado como ponto de energia máxima no eixo horizontal), além da presença de raios bem definidas (raios-X característicos). Dados simulados via método de Monte-Carlo pelo software DTSA II [71].

### Falha na análise clássica

Embora a teoria eletromagnética clássica logre êxito quanto à previsão de que cargas elétricas aceleradas emitem REM, uma série de características mensuráveis dos raios-X que contradizem as previsões teóricas foram sendo observadas e relatadas ao longo do tempo. A composição espectral dos raios-X, ilustrada na Figura 2.3, mostra as duas principais constatações não explicadas pela teoria clássica.

Em primeiro lugar, nota-se que não há produção de raios-X para valores abaixo de certos comprimentos de onda mínimos, que são tão menores quanto mais intensas forem as ddps no tubo. Além disso, superpostos ao espectro contínuo, há raios bem definidas que guardam profunda relação com o material do anodo. Devido a essa forte dependência, as raios são chamadas de raios-X característicos. Estes fenômenos só puderam ser compreendidos com o advento da mecânica quântica.

A curva contínua do espectro pode ser parcialmente compreendida pelos argumentos clássicos expostos na seção anterior: o elétron que adentra o anodo sofre sucessivas desacelerações, emitindo parte de sua energia cinética na forma de REM. Entretanto, a obtenção da expressão analítica do perfil espectral da radiação de frenamento (*bremstrahlung*) não é trivial, exigindo certas considerações especiais; vide, por exemplo, o trabalho de Kramers [48]. No que diz respeito à existência de um comprimento de onda mínimo para ocorrência de raios-X, devemos recorrer à hipótese da quantização da energia:

$$E = h\nu, \quad (2.5)$$

em que  $E$  é a energia da partícula,  $h$  é a constante de Planck e  $\nu$  é a frequência característica para a energia em questão. A hipótese de quantização da energia é suficiente para explicar a existência do comprimento de onda mínimo: considerando que o elétron mais energético do feixe proveniente do catodo seja completamente freado, temos

$$E_{\max} = h\nu_{\max} = eV \rightarrow h\frac{c}{\lambda_{\min}} = eV \Rightarrow \lambda_{\min} = \frac{hc}{eV}, \quad (2.6)$$

em que  $E_{\max}$  é a energia do elétron e  $V$  é a ddp existente entre os terminais do tubo. A Equação (2.6) apresenta a chamada lei de Duane-Hunt, em que fica explícita a dependência entre o comprimento de onda mínimo para ocorrência de raios-X com o inverso da tensão de operação do tubo.

Os raios-X característicos são compreendidos quando levamos em consideração a possibilidade de ionização de um átomo submetido à ação de um elétron externo: se este possuir

energia suficiente, um elétron de determinada camada pode ser arrancado. Este desarranjo põe o átomo em estado excitado e, em busca de um estado de equilíbrio, transições entre as camadas eletrônicas ocorrem. Por sua vez, estas transições ocorrem por meio de liberação de fótons com energias correspondentes às diferenças energéticas entre as camadas. Uma característica importante destas emissões está associada ao fato de que elas ocorrem segundo *regras de seleção* específicas, envolvendo os números quânticos que descrevem o sistema. Nesse sentido, são observadas *estruturas finas* nos raios-X característicos, uma vez que transições entre sub-camadas podem ocorrer. Portanto, é possível realizar análises espectroscópicas de materiais por meio de suas emissões de raios-X característicos.

### 2.3 Interação dos raios-X com a matéria

Explicitadas a natureza e as características básicas dos raios-X, faz-se necessária uma discussão mínima sobre a sua interação com a matéria. Dentre os diversos mecanismos de interação do fóton com um átomo, será dada ênfase nos fenômenos mais relevantes em radiodiagnóstico: espalhamento coerente, efeito fotoelétrico, efeito Compton e produção de pares.

#### Atenuação dos raios-X

Uma grandeza de interesse geral em física dos raios-X corresponde ao *coeficiente de atenuação linear*,  $\mu$ , do material absorvedor. Esta grandeza pode ser encarada como uma estimativa da probabilidade por unidade de comprimento de quanto o fóton interagirá com o absorvedor [66]. Uma primeira forma de obtenção do coeficiente de atenuação linear seria realizar o experimento esquematizado na Figura 2.4: um feixe monocromático deve ser lançado sobre absorvedores de diferentes espessuras e a intensidade da radiação transmitida medida. Sendo  $I(x)$  a intensidade medida dos raios-X em função da espessura

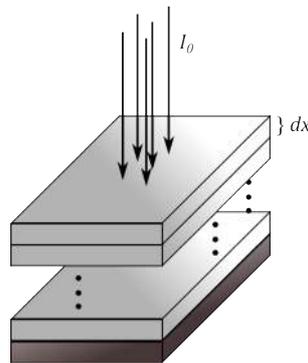


Figura 2.4: Experimento para determinação do coeficiente de atenuação de um determinado absorvedor: um feixe monocromático e colimado incide, inicialmente, diretamente sobre um detector (objeto marrom). Em seguida, aumenta-se a espessura dos absorvedores interpostos entre o feixe e o detector. Verifica-se nesse caso um decaimento exponencial da intensidade dos raios-X em função da espessura dos absorvedores.

$x$  do absorvedor, verifica-se experimentalmente [24, 66] que

$$\frac{dI}{I} = -\mu dx. \quad (2.7)$$

A Equação (2.7) mostra que a intensidade relativa da radiação detectada decai linearmente com a espessura deste. Por integração, obtemos então que

$$I(x) = I_0 \exp(-\mu x), \quad (2.8)$$

em que  $I_0$  corresponde à intensidade do feixe sem absorvedor.

Deve-se ter em mente que a lei de decaimento expressa em (2.8) é restrita em suas hipóteses: o feixe incidente deve ser monocromático – diferentemente daquele gerado diretamente por tubos convencionais – e efeitos de espalhamento não foram levados em consideração. Efetivamente, o coeficiente  $\mu$  guarda forte dependência com a energia dos fótons, fazendo com que a relação de decaimento da intensidade de feixe com a profundidade não seja tão simples quanto à modelada pelas Equações 2.7 e 2.8. Ademais, cabe ressaltar que é possível expressar  $\mu$  de maneira independente da densidade do material, bastando dividi-lo pela densidade do meio:

$$\mu_m = \frac{\mu}{\rho}.$$

A esta grandeza damos o nome de coeficiente de atenuação mássico. Sua utilidade reside na pouca variabilidade de seus valores entre absorvedores distintos. Usando um exemplo de Compton e Allison [24]: o coeficiente de atenuação linear da água para um certo feixe é muito maior do que o coeficiente linear do vapor de água para o mesmo feixe, ao passo que os coeficientes mássicos de ambos são semelhantes.

Uma outra observação relativa à Equação (2.8) remete ao absorvedor: a referida expressão foi deduzida considerando-se o absorvedor como sendo um material puro. De um modo geral, o coeficiente de atenuação mássico para uma *mistura* de  $N$  elementos é dado por

$$\mu_{\text{mist}} = \sum_{i=1}^N w_i \left( \frac{\mu}{\rho} \right)_i, \quad (2.9)$$

em que  $w_i$  corresponde à fração em massa do  $i$ -ésimo composto. Conforme apontado por Okuno e Yoshimura [58], na derivação da expressão contida em (2.9) não são consideradas alterações nas configurações eletrônicas dos constituintes.

Para concluir a discussão sobre coeficientes de atenuação, é importante ressaltar que – conforme apontado no início da seção – estas grandezas são estimativas das probabilidades de interação dos fótons com o meio absorvedor. Nesse sentido, o coeficiente  $\mu$  de um certo absorvedor deve carregar informação sobre as probabilidades de interação dos raios-X segundo todos os fenômenos possíveis (espalhamento coerente, efeito fotoelétrico etc.). Tratando estes eventos como estatisticamente independentes uns dos outros, temos [58, 37]:

$$\mu_m = \frac{\mu}{\rho} = \frac{N_A}{A} {}_a\sigma_{\text{tot}}, \quad (2.10)$$

em que  $N_A$  é o número de Avogadro e  $A$  é o número de massa. A grandeza  ${}_a\sigma_{\text{tot}}$  corresponde à seção de choque *atômica* total, isto é,

$${}_a\sigma_{\text{tot}} = {}_a\sigma_{\text{coe}} + {}_a\sigma_{\text{C}} + {}_a\tau + {}_a\kappa, \quad (2.11)$$

em que os termos da soma correspondem às seções de choque atômicas do espalhamento coerente, efeito Compton, efeito fotoelétrico e produção de pares, respectivamente. Uma descrição sucinta sobre estes fenômenos será apresentada nas próximas seções.

### Espalhamento coerente

O espalhamento coerente<sup>5</sup> caracteriza-se pelo espalhamento do fóton pelos elétrons mais externos do material absorvedor sem transferência de energia apreciável (espalhamento elástico). Do ponto de vista clássico, o espalhamento coerente pode ser compreendido como a imposição de uma oscilação sobre os elétrons do átomo em função do campo oscilante da onda incidente. Nesse contexto, espera-se que as frequências de oscilação do elétron e da onda emitida sejam as mesmas. Para este efeito, o ângulo de espalhamento do fóton é relativamente pequeno e a energia do fóton espalhado é muito próxima da energia original. Embora não apresente influência no meio devido à não-transferência de energia, o espalhamento coerente tem importância na qualidade da imagem médica [66].

A expressão formal para a seção de choque atômica do espalhamento coerente,  ${}_a\sigma_{\text{coe}}$ , é obtida pela integração de sua seção de choque diferencial por unidade de ângulo espalhado<sup>6</sup>. Experimentalmente, verifica-se para certas frequências  $\nu'$  uma proporcionalidade da forma

$${}_a\sigma_{\text{coe}} \propto \left( \frac{Z}{h\nu'} \right)^2, \quad (2.12)$$

em que  $\nu'$  é definida como uma frequência em que as seções de choque do espalhamento coerente e do efeito Compton se igualam [66] e  $Z$  é o número atômico do absorvedor.

### Efeito fotoelétrico

O efeito fotoelétrico é caracterizado pela completa absorção do fóton pelo átomo, seguida da emissão de um elétron (fotoelétron) e recuo do átomo excitado como um todo. Embora pequeno, o recuo do átomo deve ser considerado para que as conservações de energia e momento sejam mantidas [58, 66]. Em geral, o efeito fotoelétrico ocorre entre o fóton e um elétron fortemente ligado, das camadas mais internas do átomo. Com essa ejeção foto-eletrônica, o átomo tende a se desexcitar via emissão de raios-X característicos ou por emissão de um elétron Auger<sup>7</sup>.

A seção de choque atômica do efeito fotoelétrico,  ${}_a\tau$ , não possui expressão analítica, uma vez que apresenta descontinuidades para energias iguais às energias de ligação dos elétrons. Estas descontinuidades são chamadas bordas de absorção e são identificadas de acordo com sua camada característica (bordas K, L, M etc.) [58, 66]. Contudo, verifica-se [66, 37] para energias relativamente distantes das bordas de atenuação que existe uma proporcionalidade da forma

$${}_a\tau \propto \frac{Z^4}{(h\nu)^3}. \quad (2.13)$$

### Efeito Compton

O efeito Compton é caracterizado pelo espalhamento *inelástico* entre um fóton e um elétron livre, ou muito-fracamente ligado. A cinemática do efeito Compton é simples: um fóton de energia inicial  $h\nu$  é espalhado com energia  $h\nu' < h\nu$  e arranca o elétron, que deixa o

<sup>5</sup>Também chamado de espalhamento Rayleigh.

<sup>6</sup>Chamamos de seção de choque diferencial por unidade de ângulo espalhado a grandeza  $\frac{d\sigma}{d\theta}$ , que quantifica o espalhamento ocorrido dentro de um cone delimitado entre  $\theta$  e  $\theta + d\theta$  [66].

<sup>7</sup>No efeito Auger, a energia que seria liberada na forma de fóton é transferida para um elétron de alguma camada superior (menor energia de ligação), levando à ejeção deste [66].

átomo com energia  $K$ . É interessante notar que o efeito Compton prevê retroespalhamento, evidenciando mais uma vez o comportamento corpuscular dos fótons. A seção de choque atômica total do efeito Compton,  ${}_a\sigma_C$ , é dada pela multiplicação do número atômico  $Z$  do absorvedor pela seção de choque diferencial eletrônica, que é obtida pela chamada *fórmula de Klein-Nishina* e tem a forma

$${}_a\sigma_C = 2\pi r_e^2 Z \left[ \frac{1 + \varepsilon}{\varepsilon^2} \left( 2 \frac{1 + \varepsilon}{1 + 2\varepsilon} - \frac{\ln(1 + 2\varepsilon)}{\varepsilon} \right) + \frac{\ln(1 + 2\varepsilon)}{2\varepsilon} - \frac{1 + 3\varepsilon}{(1 + 2\varepsilon)^2} \right], \quad (2.14)$$

em que  $\varepsilon = \frac{h\nu}{m_e c^2}$  – sendo  $m_e c^2$  a energia de repouso do elétron – e  $r_e$  é o raio clássico do elétron.

### Produção de pares

Se a energia do fóton é maior que o dobro da energia de repouso do elétron; isto é,  $h\nu > 2m_e c^2 = 1,02$  MeV, a produção de uma par elétron-pósitron é viabilizada. O fenômeno de produção de pares evidencia os princípios de conservação da energia, carga elétrica e momento linear: sendo a carga do fóton nula, a soma total das cargas das partículas geradas deve ser nula e a energia do fóton é convertida totalmente nas energias das partículas. Contudo, a conservação do momento ocorre apenas se o fóton estiver sobre influência de um campo Coulombiano (elétron ou núcleo), uma vez que o momento inicial do fóton,  $p = \frac{h\nu}{c}$  não é completamente transferido para o par criado [66]. Nesse sentido, o centro gerador do campo elétrico age como um receptor de parte dessa quantidade. Se o centro gerador do campo corresponde ao núcleo do absorvedor, um par elétron-pósitron é criado; caso o centro gerador de campo / receptor do momento for um elétron orbital, o recuo deste último pode não ser desprezível, caracterizando um processo de ionização. Com isso, o produto final da interação é a emissão de um pósitron e *dois* elétrons (triplete).

Na Figura 2.5 podem ser visualizadas seções de choque para o molibdênio, evidenciando as características discutidas até agora.

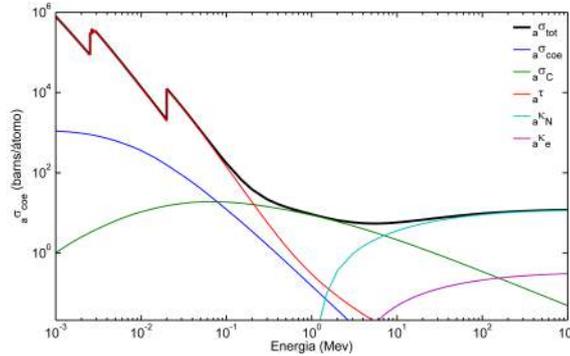


Figura 2.5: Seções de choque atômicas para os mecanismos de interação dos raios-X com molibdênio ( $Z = 42$ ). A seção de choque para produção de pares é mostrada individualmente para a produção de par elétron-pósitron ( $a_{\kappa N}$ ) e produção de tripleto ( $a_{\kappa e}$ ). Dados obtidos da base XCOM [8].

## 2.4 Imagem mamográfica

A mamografia é uma modalidade diagnóstica baseada na *transmissão* de raios-X pela mama e posterior sensibilização de filme ou sensor. Ela é tida como a melhor opção em

detecção precoce do câncer de mama e sua operacionalização assemelha-se à radiografia convencional: em função dos diferentes coeficientes de atenuação das componentes da mama, ocorrem sensibilizações mais ou menos intensas, gerando assim a representação por diferenças de contraste das estruturas retratadas. Um esquema simplificado do arranjo usado em mamografia pode ser visto na Figura 2.6.

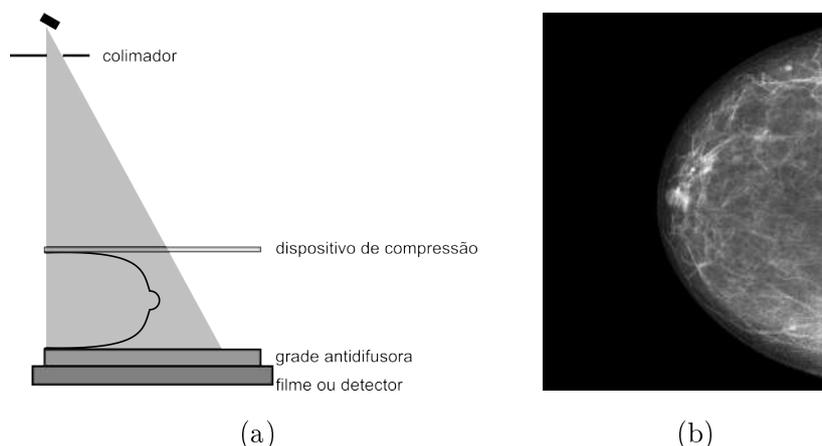


Figura 2.6: Esquema simplificado do arranjo utilizado em mamografia (a) e imagem real (b), adquirida segundo a modalidade radiografia computadorizada (CR).

Dois características particulares da mamografia dizem respeito ao aparato para a aquisição da imagem: há um dispositivo de compressão, que leva – simultaneamente – à redução da espessura da mama (diminuindo o espalhamento dos fótons) e ao distanciamento entre suas estruturas internas, permitindo a visualização de detalhes que outrora seriam obstruídos. A força de compressão deve variar, idealmente, entre 11 e 18 Kgf [15]. A segunda característica é a presença (opcional) de uma grade antidifusora. A função desta grade é óbvia: impedir que feixes não-paralelos às suas estrias sejam transmitidos. Isso é feito para que fótons espalhados – principalmente por espalhamento Compton – não sensibilizem inutilmente o detector e diminuam as diferenças locais de contraste, degradando a imagem final. Existe um compromisso quanto ao uso da grade: a dose sobre o paciente deve aumentar, já que fótons potencialmente úteis na formação da imagem podem ser absorvidos pelo anteparo [37]. Segundo Hobbie e Roth [37], o uso da grade antidifusora melhora em 40% o contraste da imagem, mas às expensas de uma dose duas ou até três vezes maior que a usual.

Outro recurso usado em radiografia diagnóstica consiste no emprego de *filtros* para a modificação dos espectros. Uma aplicação típica consiste em remover as mais baixas energias do espectro contínuo, pois fótons pouco energéticos podem não atravessar o tecido, configurando um aumento da dose sem melhoria alguma na imagem. Em mamografia a situação é mais complexa: as estruturas constituintes da mama possuem coeficientes de atenuação muito próximos entre si, e só há uma separação clara entre eles (e portanto das estruturas associadas) para baixas energias. Para contornar esta condição, é usual em mamografia o uso de filtros de borda K, que tendem a evidenciar os raios-X característicos reativos a transições dessa borda [37, 58, 40]. Isso é feito para que se favoreçam interações via efeito fotoelétrico, permitindo uma diferenciação significativa dos elementos devido à forte dependência da seção de choque  $\sigma_a$  com o número atômico do absorvedor, tal como mostrado na Equação (2.13). Esse recurso é eficaz quando se quer diferenciar calcificações

(pontos de acúmulo de cálcio,  $Z = 20$ ) de tecido mole, por exemplo. A Figura 2.7 ilustra a pouca separação entre os coeficientes de atenuação da água e do cálcio.

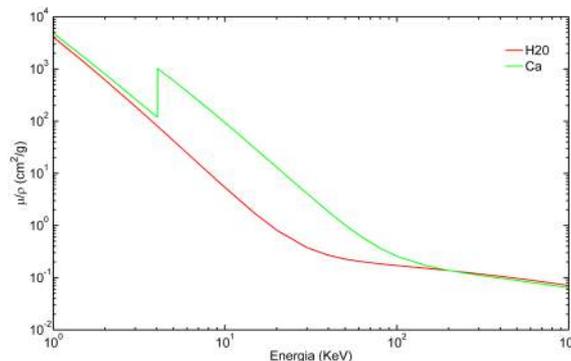


Figura 2.7: Coeficientes de atenuação mássicos da água e cálcio. Nota-se que só há uma separação clara entre os dois coeficientes para baixas energias. Dados obtidos da base XCOM [8].

É importante ressaltar que existem diferentes modalidades de aquisição da imagem mamográfica. O método mais usual corresponde ao sistema tela-filme, em que o fóton de raios-X sensibiliza uma emulsão fotográfica e gera a imagem diretamente. Contudo, a probabilidade de interação dos fótons com a emulsão é relativamente baixa [37] e frequentemente é feito uso de *telas intensificadoras*, que são produzidas de forma que os fótons interajam fortemente com elas, produzindo luz visível. É justamente essa luz visível que sensibiliza o filme outrora não sensibilizado e produz o registro físico da imagem. Essa combinação tela-filme é acondicionada em *cassetes* que impedem a entrada de luz externa no conjunto, de forma a preservar a informação de interesse. Embora usual, os sistemas tela-filme carregam inconvenientes, tais como ruídos e perda de resolução espacial em função do espalhamento da luz visível sobre o filme. Ademais, para fins de manipulação da imagem, é necessário que imagens adquiridas em sistemas tela-filme sejam submetidas a digitalização, em que há – inevitavelmente – alterações na estrutura da informação, pois um sinal até certo ponto contínuo passa por quantização e amostragem (vide Seção 3.1). Uma alternativa para minimizar os inconvenientes do sistema tela-filme está associada à mamografia *digital*. Dentre as diferentes sub-modalidades de mamografia digital, apenas a *radiografia computadorizada*, CR<sup>8</sup>, será discutida. Discussões especializadas em mamografia digital podem ser encontradas no trabalho de James [44] e no *report 82* da ICRU [40].

Na modalidade CR, a absorção dos fótons ocorre em uma placa composta de material cerâmico fotoluminescente, cujos átomos são excitados em função da energia recebida e depois de algum tempo decaem, processo esse acompanhado de emissão de luz visível. Ou seja, uma vez absorvidos os fótons oriundos da mama, forma-se uma imagem latente no fósforo, a qual é lida posteriormente por um sistema de varredura a laser. Nesse sistema, os fótons do laser levam à emissão dos fótons que caracterizam o registro físico da imagem. Esses fótons são coletados em posições *discretas* por um fotodetector, formando a imagem digital. Analogamente ao sistema tela-filme, o detector em CR é acondicionado em cassetes.

Embora seja uma modalidade intrinsecamente digital, a CR não possui resolução espacial maior que sistemas tela-filme. Entretanto, imagens CR possuem informação derivada diretamente do sinal original, diferentemente de imagens tela-filme digitalizadas. Em um

<sup>8</sup>Do inglês, *Computed Radiography*.

cenário de processamento de imagens, isso se torna relevante.

### 2.4.1 Calcificações

Dentre os diversos sinais radiológicos associados ao câncer de mama, as *calcificações* demandam especial atenção. Ocorre que existem dificuldades sérias quanto à observação e classificação desse tipo de estrutura. Algumas delas são:

**Baixa dimensionalidade:** Segundo classificação adotada na literatura médica [54, 47, 76], classificam-se como *microcalcificações* partículas com diâmetro menor ou igual a 0,50 mm;

**Número:** A forma de agrupamento de um conjunto de calcificações sugere a malignidade (ou não) do achado. Quão maior o número de calcificações em um agrupamento, maior a suspeita de malignidade;

**Densidade:** Embora as calcificações malignas sejam, em média, fortes atenuadoras dos raios-X, parece não haver consenso se a densidade destes achados corresponde a bom descritor de malignidade [54]; isto é, mesmo calcificações malignas podem ser parcialmente radio-transparentes.

Existe ainda uma subjetividade grande na descrição dos achados, o que motivou a elaboração de um léxico específico, chamado BI-RADS<sup>®9</sup> pelo Colégio Americano de Radiologia (ACR). Associe esta subjetividade ao baixo contraste típico das imagens mamográficas e à baixa dimensionalidade das calcificações e logo percebe-se que a detecção destas estruturas constitui tarefa não-trivial. Na Figura 2.8 são mostrados alguns casos de ocorrência de calcificações.

O objetivo deste trabalho foi propor um modelo para detecção de (micro)calcificações, devido à criticalidade desta tarefa. Nesse sentido, o trabalho aqui descrito pode ser caracterizado como um sistema de auxílio à detecção, já que o objetivo principal é a detecção de formas suspeitas, e não sua descrição. Para tanto, foram empregadas técnicas diversas, oriundas de diferentes quadros teóricos. Nos próximos capítulos serão descritas as técnicas e a cadeia de processamento será detalhada.

---

<sup>9</sup>Do inglês, *Breast Imaging Reporting and Data System*.

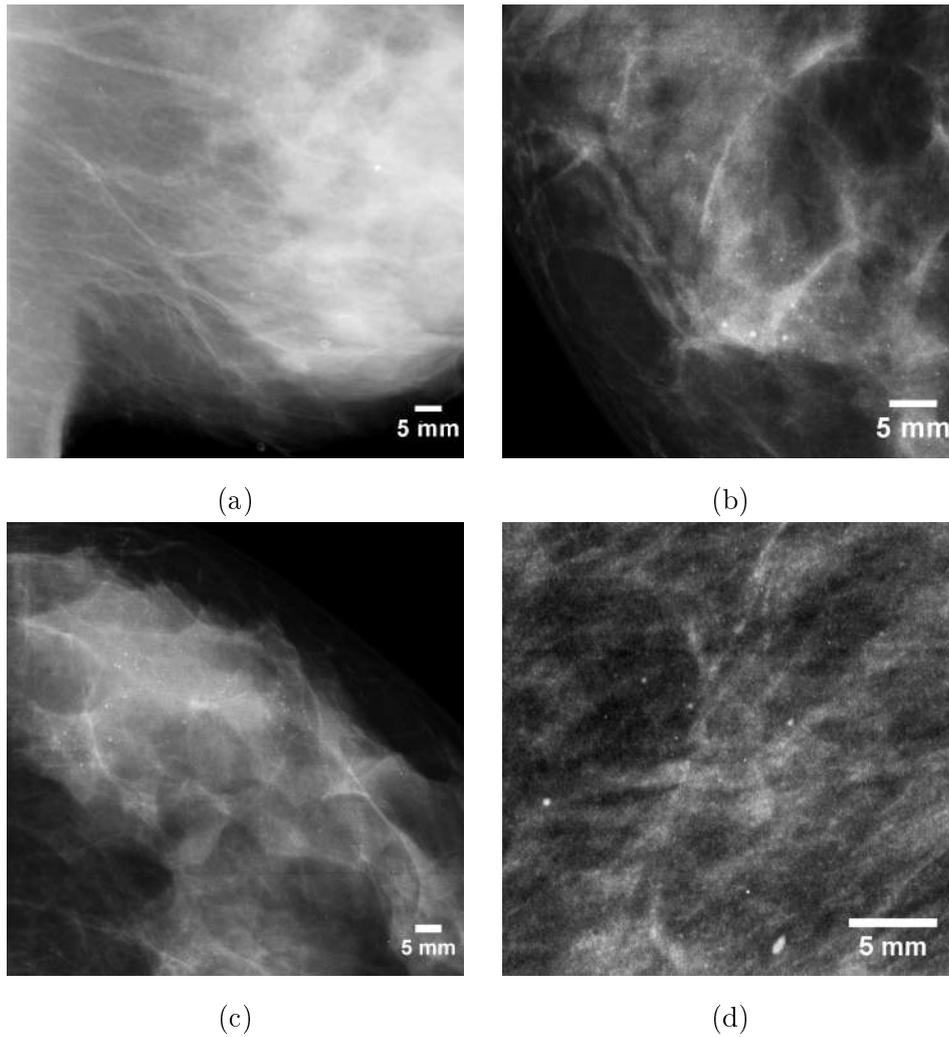


Figura 2.8: Fragmentos de imagens mamográficas contendo calcificações. Verifica-se por estes exemplos a baixa dimensionalidade característica destes achados. Verifica-se também a dificuldade em se diferenciar as calcificações dos tecidos que as circundam, principalmente no fragmento exposto em (a). Os fragmentos (b), (c) e (d) apresentam agrupamentos de microcalcificações.

## Capítulo 3

# Processamento de imagens

### 3.1 Conceitos introdutórios

Uma imagem pode ser compreendida como um sinal representativo de alguma estrutura, seja ela de natureza física, artística, conceitual etc. De uma forma geral, uma imagem pode ser tomada como um sinal (normalmente multidimensional) que carrega informação sobre a estrutura retratada. Um primeiro modelo matemático para uma imagem seria a de uma função escalar que atribui a cada ponto do plano um valor dado por

$$\begin{cases} u(x, y) = I(x, y)R(x, y), \\ 0 \leq I(x, y) < \infty \\ 0 \leq R(x, y) < 1. \end{cases} \quad (3.1)$$

em que  $I(x, y)$  e  $R(x, y)$  referem-se à iluminância e à reflectância, respectivamente. A iluminância corresponde a uma medida da quantidade de luz incidente sobre a cena ou estrutura representada, ao passo que a reflectância está associada à quantidade de luz refletida pelas estruturas representadas [62].

O modelo apresentado na Equação (3.1) triunfa por sua simplicidade. Contudo, dentro de um cenário de processamento *digital* de imagens, ele deve ser adaptado. Uma imagem digital  $u$  com  $M$  linhas e  $N$  colunas pode ser compreendida como uma matriz (que pode ser multidimensional<sup>1</sup>), em que cada entrada  $u_{ij}$  está associada à intensidade luminosa na posição  $(i, j)$ . Ou seja,

$$u = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{M1} & u_{M2} & \cdots & u_{MN} \end{bmatrix}. \quad (3.2)$$

Cada elemento  $u_{ij}$  da imagem digital é chamado de *pixel*, abreviação de *picture element*.

Nas discussões seguintes, a imagem digital mostrada na Figura 3.1 será utilizada para exemplificar vários conceitos expostos.

---

<sup>1</sup>Neste caso, à entrada  $u_{ij}$  associa-se um conjunto ordenado de valores  $\{u_{ij1}, u_{ij2}, \dots, u_{ijk}\}$ , ao invés de apenas um. Esse é o caso de uma imagem colorida.

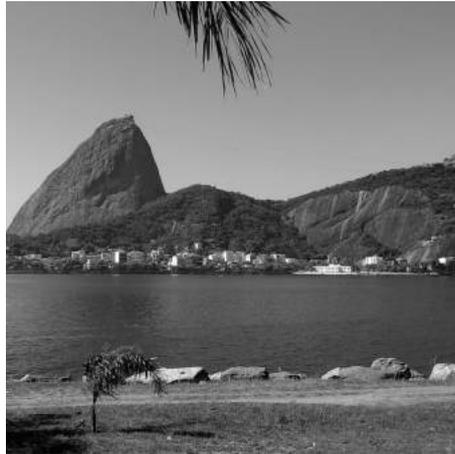


Figura 3.1: Imagem-teste a ser processada segundo as técnicas discutidas neste capítulo.

### Amostragem e quantização

Deve-se ter em mente que o processo de digitalização não pode ser realizado de qualquer maneira: o domínio da função imagem deve ser corretamente *amostrado* e seus valores devem ser adequadamente *quantizados*. A amostragem do domínio da imagem deve ser tal que a *frequência de amostragem*,  $f_s$ , seja maior que o dobro da frequência máxima presente no sinal original (teorema da amostragem). Caso este critério não seja atendido, poderá ocorrer o fenômeno de *aliasing*, que corresponde ao surgimento de componentes espectrais que não existiam originalmente, causando a corrupção ou degradação do sinal [62, 27]. Uma unidade de medida típica para resolução espacial é o ppi: pixels por polegada (*pixels per inch*). Um exemplo de reamostragem de uma imagem pode ser visualizado na Figura 3.2.



Figura 3.2: Imagem original (à esquerda) e imagem reamostrada com frequência  $f_s = 1/32\text{px}^{-1}$  em ambas direções espaciais. A imagem reamostrada foi mantida com as mesmas dimensões em pixels da imagem original para evidenciar os efeitos da amostragem.

Por sua vez, a quantização está associada à quantidade de bits usados para representar os valores de intensidade luminosa da imagem digital. Em geral, usam-se quantidades de bits que são potências de 2, dessa forma o intervalo de valores possíveis para uma imagem

quantizada em  $N$  bits varia de 0 a  $2^N - 1$ . Por exemplo, uma imagem quantizada em 8 bits tem entradas que podem assumir valores entre 0 e 255. O efeito de diferentes quantizações sobre uma mesma imagem pode ser visualizado nas Figuras 3.3, 3.4, 3.5 e 3.6.



Figura 3.3: Imagem original, quantizada em 8 bits. Os valores de suas entradas podem variar de 0 a 255.

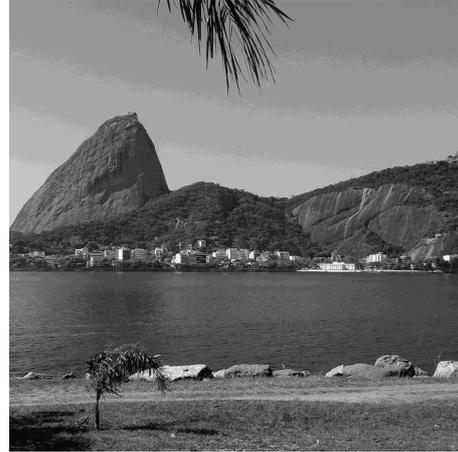


Figura 3.4: Imagem quantizada em 4 bits. Os valores de suas entradas podem variar de 0 a 15.

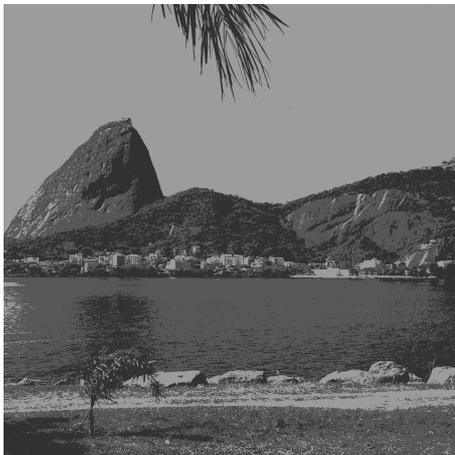


Figura 3.5: Imagem quantizada em 2 bits. Os valores de suas entradas podem variar de 0 a 3.

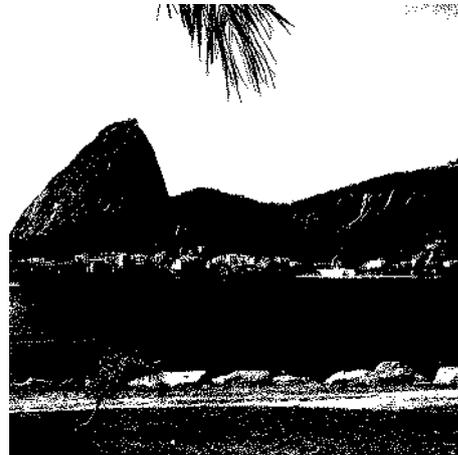


Figura 3.6: Imagem quantizada em 1 bit. Os valores de suas entradas podem variar entre 0 e 1.

Podemos concluir das discussões sobre amostragem e quantização que quão mais refinada for a amostragem e mais largo o intervalo de bits usados na quantização, maior espaço será necessário para o armazenamento de uma imagem e mais ineficiente será sua - eventual - transmissão. Ainda sobre a questão da amostragem, há o fato de que uma imagem com um número elevado de entradas demandará maior esforço computacional em processamento. Questões como essas devem ser consideradas previamente ao estabelecimento do sistema de aquisição e processamento.

### Coordenadas espaciais e vizinhanças

Usualmente, o sistema de coordenadas espaciais segue a orientação sugerida pela representação matricial mostrada na Equação (3.2) e tal como mostrado na Figura 3.7.

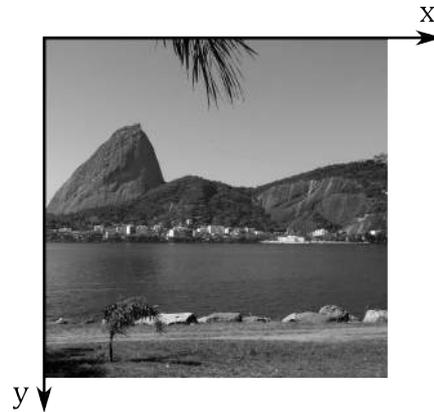


Figura 3.7: Sistema de coordenadas usual em processamento de imagens.

Conforme será discutido adiante, há técnicas de processamento de imagens que se baseiam em relações entre pixels de uma determinada região. Nesse sentido, é importante definir *a priori* qual a *vizinhança* adotada no esquema de processamento. Dada uma imagem digital  $u$ , temos que – à exceção dos pixels de fronteira – cada pixel  $u_{ij}$  tem 8 vizinhos imediatos: dois na horizontal, dois na vertical e quatro nas diagonais. Se estes oito vizinhos se relacionam com o pixel  $u_{ij}$ , sua relação de vizinhança com os demais é chamada  $N_8$ . Por outro lado, se o pixel  $u_{ij}$  só se relacionar com os vizinhos na vertical e horizontal, a vizinhança é dita ser  $N_4$ . O que define este relacionamento é a forma como as operações sobre vizinhança são executadas. A Figura 3.8 ilustra o que foi dito.

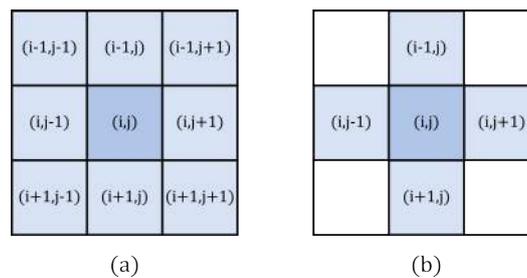


Figura 3.8: Vizinhanças (a)  $N_8$  e (b)  $N_4$  do pixel  $u_{ij}$ .

É importante ressaltar que a definição da vizinhança tem importância fundamental quando se processa uma imagem via modelos derivados de equações diferenciais parciais (Capítulo 4), em que a aproximação das derivadas parciais se dá por diferenças finitas sobre os pixels da região processada.

### Histograma de uma imagem

O histograma de uma imagem digital  $u$  de dimensões  $M \times N$  é definido como sendo a sequência

$$h_k = \frac{p_k}{M \times N}. \quad (3.3)$$

em que  $\frac{p_k}{M \times N}$  é a probabilidade de ocorrência do  $k$ -ésimo nível de intensidade.

Na forma apresentada na Equação (3.3), o histograma  $h$  constitui uma aproximação à distribuição de probabilidades associada aos níveis de intensidade de  $u$ . Nesse sentido, nenhuma informação espacial é mantida quando o histograma de uma imagem é obtido. Por esse motivo, um mesmo histograma pode estar associado a mais de uma imagem.

Do ponto de vista de *análise* de imagens, o histograma guarda informações sobre o contraste da imagem a ele associada. Imagens com boa relação de contraste apresentam histogramas mais distribuídos ao longo dos valores possíveis, ao passo que imagens de baixo contraste apresentam maior concentração em torno da média. As Figuras 3.9 e 3.10 apresentam uma ilustração deste fato.

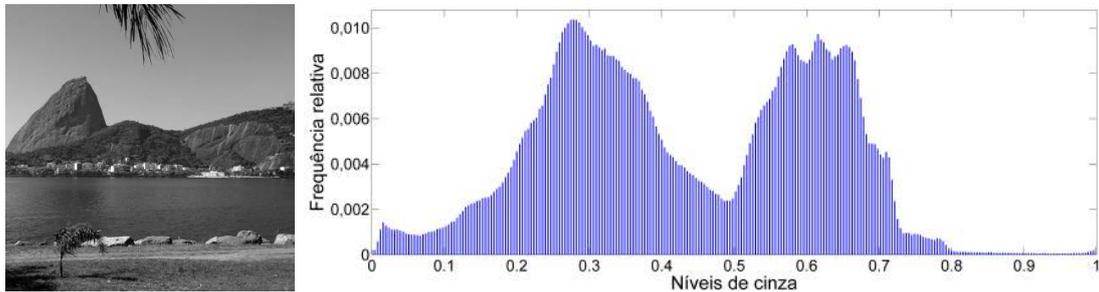


Figura 3.9: Figura original à esquerda e seu histograma, estimado com 256 divisões de níveis de cinza. O histograma fornece informações acerca da distribuição dos níveis de intensidade, mas não de sua localização. Os valores de intensidade (eixo horizontal) foram normalizados para o intervalo  $[0,1]$ .

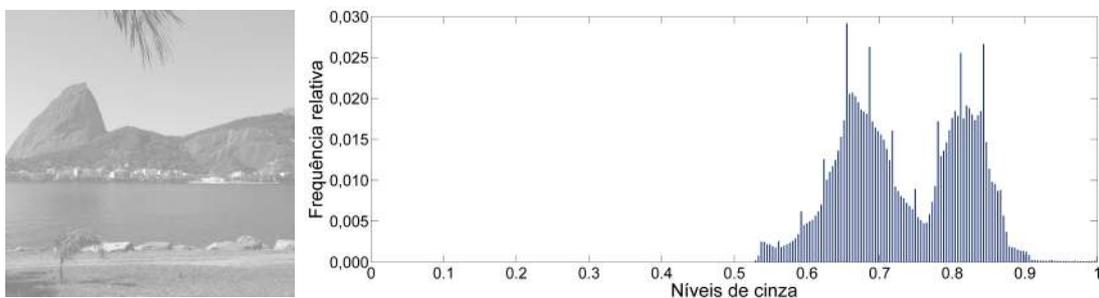


Figura 3.10: Figura com baixa relação de contraste (à esquerda) e seu histograma. Nota-se nesse caso, quando comparado com a Figura 3.9, que o perfil do histograma associado a uma imagem com pouco contraste apresenta maior concentração espacial, refletindo o fato de que os valores de intensidade estão menos dispersos em torno da média. Os valores de intensidade (eixo horizontal) foram normalizados para o intervalo  $[0,1]$ .

Alterações na distribuição dos níveis de cinza de uma imagem podem ser executadas segundo as técnicas de filtragem – a serem discutidas mais adiante –, ou mesmo por meio de transformações sobre um único pixel, cuja forma geral é

$$w_{ij} = T(u_{ij}), \quad (3.4)$$

em que  $T(\cdot)$  corresponde a uma transformação qualquer, linear ou não.

Um exemplo típico desse tipo de transformação consiste na operação de *alargamento linear de contraste*<sup>2</sup>,  $T_{\text{ALC}}(\cdot)$ . Essa operação redistribui – linearmente – os níveis de cinza

<sup>2</sup>Usualmente conhecido como *linear stretching*. Não confundir com *equalização do histograma*.

da imagem, que estão dentro do intervalo  $[u_{\min}, u_{\max}]$  original de forma a abranger todos os valores possíveis de uma nova escala, cujos limites são  $v_{\min}$  e  $v_{\max}$ . A expressão para esta operação pode ser vista na Equação (3.5):

$$T_{\text{ALC}}(u) = \frac{v_{\max} - v_{\min}}{u_{\max} - u_{\min}}(u - u_{\min}) + v_{\min}. \quad (3.5)$$

Para finalizar a discussão sobre histogramas, é importante notar que este sinal possibilita a extração de atributos estatísticos importantes, tais como valor médio, desvio padrão e variância, energia, entropia etc.

## 3.2 Filtragem

Apresentados os conceitos básicos sobre representação de imagens digitais, faz-se necessário introduzir um conjunto de conceitos relativos às operações mais básicas definidas sobre estes objetos. Especificamente, a classe das operações aritméticas desempenha importante papel, pois leva naturalmente ao conceito de operação orientada à vizinhança. Esta, por sua vez, leva ao conceito de filtragem, que permeará todo este trabalho.

### Operações aritméticas básicas

Uma primeira classe de operações sobre imagens consiste em manipularmos duas ou mais imagens por meio de operações aritméticas básicas (adição, subtração, multiplicação, divisão). Ao efetuarmos operações desse tipo sobre imagens, devemos nos atentar ao fato de que os valores resultantes podem estar fora do intervalo de quantização. Por exemplo, ao somarmos duas imagens quantizadas a 8 bits, é possível que algum pixel da imagem resultante tenha valor maior que 255. Valores negativos podem ocorrer e deve-se atentar ao fato de que em divisão entre imagens pode ocorrer divisão por zero. Para contornar a questão do intervalo de quantização, normalmente submete-se a imagem resultante a um processo de normalização para que os valores estejam dentro do intervalo requerido. Quanto à divisão por zero, geralmente soma-se 1 no divisor.

Inicialmente, as operações aritméticas são aplicadas pixel a pixel, isto é, dada uma operação aritmética  $f$  a ser aplicada sobre duas imagens –  $u$  e  $v$  –, temos

$$w_{ij} = f(u_{ij}, v_{ij}).$$

Ou seja, cada pixel da imagem resultante  $w$  é dado pela aplicação de  $f$  sobre os pixels correspondentes de  $u$  e  $v$ .

Como exemplo de aplicação de operações aritméticas sobre imagens, são mostrados na Figura 3.12 os resultados de diversas manipulações de duas imagens-teste mostradas na Figura 3.11.

### Operações orientadas a uma vizinhança

Conforme dito antes, as operações aritméticas relacionam diretamente um pixel de uma imagem ao pixel correspondente de outra imagem a ser operada. Contudo, estas operações podem ser estendidas para que sejam realizadas sobre uma vizinhança de certo pixel, daí a importância de se estabelecer *a priori* a estrutura de vizinhanças de uma imagem, tal como apontado ao final de Seção 3.1.

A dinâmica das operações orientadas a vizinhança pode ser compreendida como a seguinte sequência:

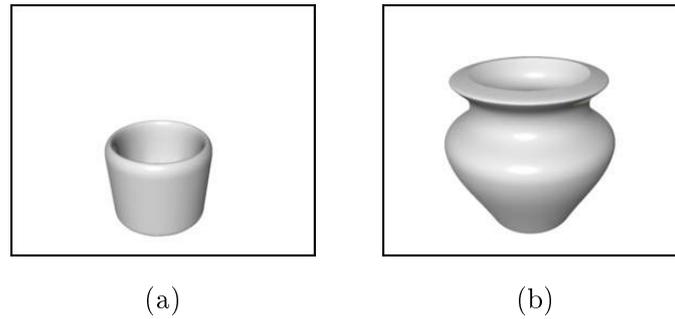


Figura 3.11: Imagens  $u$  (copo) e  $v$  (vaso) a serem usadas nas ilustrações de operações aritméticas sobre imagens.

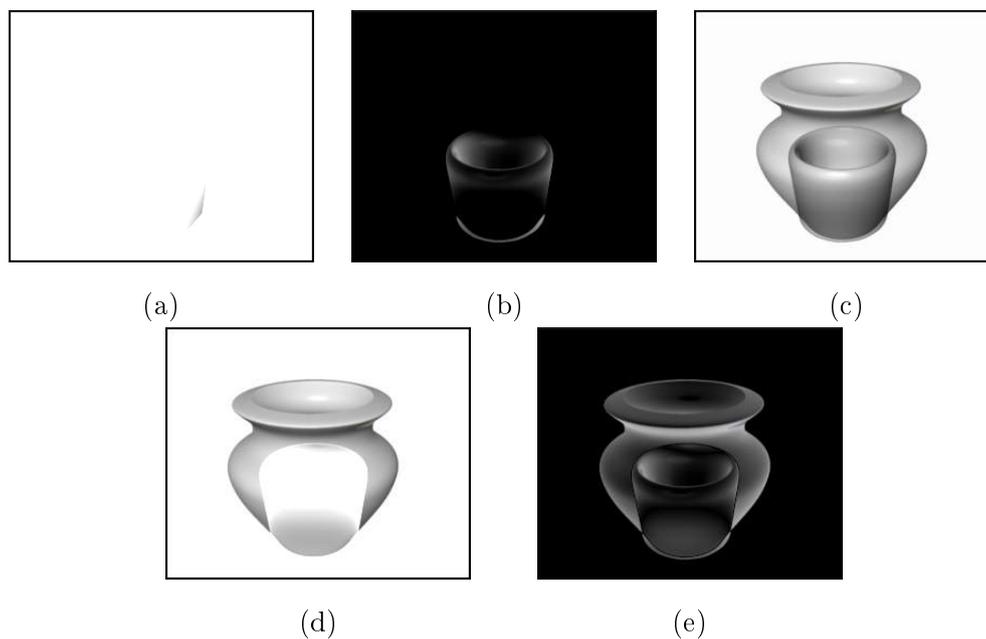


Figura 3.12: Resultados de operações aritméticas entre as imagens constantes na Figura 3.11. Em (a) é mostrado  $w = u + v$ , em (b)  $w = u - v$ , em (c)  $w = u \times v$ , em (d)  $w = \frac{u}{v}$  e em (e) é mostrado  $w = |u - v|$ .

1. Fixe uma janela de tamanho fixo sobre o pixel  $u_{ij}$ , estabelecendo uma vizinhança  $\eta(i, j)$  a ser levada em consideração na operação;
2. atribua ao pixel resultante  $w_{ij}$  o resultado da operação descrita na Equação (3.6):

$$w_{ij} = \sum_{k,l \in \eta(i,j)} f_{kl} \cdot u_{ij}, \quad (3.6)$$

em que  $f$  é uma *máscara* cujos coeficientes definirão qual operação está sendo realizada <sup>3</sup>; e

3. passe ao pixel seguinte e reinicie o processo até varrer todos os pixels da imagem original.

Uma ilustração desta dinâmica pode ser visualizada na Figura 3.13.

Convém ressaltar que os pixels de fronteira não possuem a mesma estrutura de vizinhança de pixels internos. Esse fato demanda uma adaptação de vizinhos para pixels localizados nas fronteiras da imagem. Duas abordagens usuais são: complementar as posições ausentes com zeros (*zero padding*) ou replicar os pixels. Adicionalmente, os valores

<sup>3</sup>Conforme será visto adiante, a máscara corresponde, de fato, ao *núcleo* de um certo operador

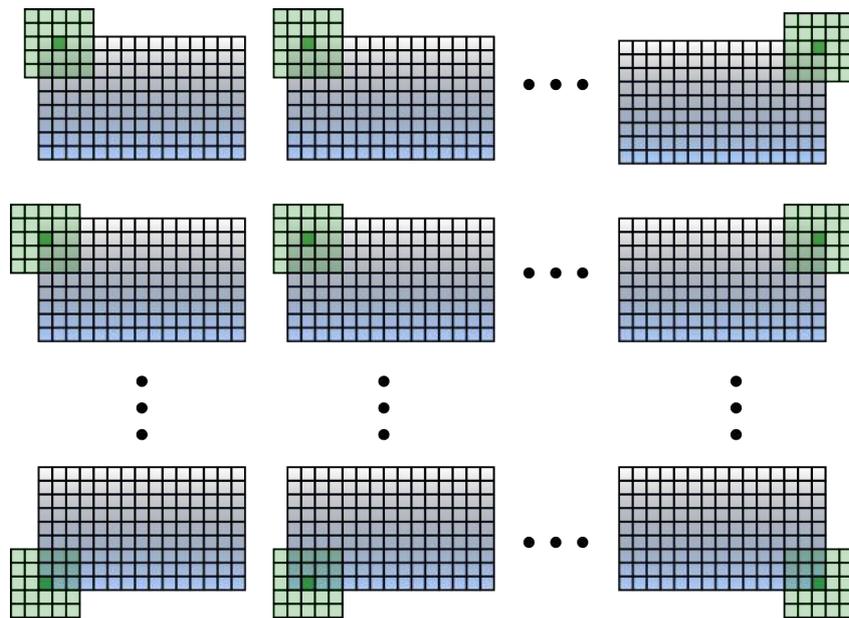


Figura 3.13: Ilustração da dinâmica de operação orientada a vizinhança: uma máscara (em verde) varre a imagem original (em azul). Uma vez centralizada sobre um determinado pixel, uma operação sobre os pixels delimitados é efetuada.

atualizados dos pixels processados devem ser armazenados em outra imagem, de forma que a vizinhança não seja alterada durante o processamento dela própria.

Um exemplo básico de operação orientada à vizinhança consiste em atribuir ao pixel  $u_{ij}$  a média aritmética dos pixels delimitados pela máscara. Um exemplo da aplicação da média sobre uma vizinhança de 400 pixels pode ser visualizado na Figura 3.14.



Figura 3.14: Imagem original não-processada (à esquerda) e submetida à operação de média aritmética sobre uma vizinhança de 400 pixels. Note-se a perda de detalhes finos na imagem processada.

Neste exemplo, a máscara a ser usada é da forma

$$f = \frac{1}{400} \cdot \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}_{20 \times 20}.$$

Resultados completamente distintos podem ser atingidos alterando-se os coeficientes de  $f$  na equação (3.6). Essa é a base da filtragem de imagens, a ser discutida a seguir.

Em primeiro lugar, é conveniente rescrever a Equação (3.6) de forma explícita [62]:

$$w_{ij} = \sum_{k=\lfloor -\frac{m}{2} \rfloor}^{\lfloor \frac{m}{2} \rfloor} \sum_{l=\lfloor -\frac{n}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} f_{kl} \cdot u_{i+k, j+l}, \quad (3.7)$$

para uma máscara de dimensão  $m \times n$ .

Analisando a Equação (3.7), percebemos que uma operação orientada a vizinhança corresponde, de fato, à operação de *correlação discreta* entre a máscara  $f$  e a imagem  $u$ . De maneira análoga, é possível que a operação de correlação torne-se uma operação de *convolução discreta*:

$$w_{ij} = (\tilde{f} * u)_{ij} = \sum_{k=\lfloor -\frac{m}{2} \rfloor}^{\lfloor \frac{m}{2} \rfloor} \sum_{l=\lfloor -\frac{n}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} \tilde{f}_{kl} \cdot u_{i-k, j-l}, \quad (3.8)$$

em que o asterisco representa a convolução entre os dois sinais e  $\tilde{f}$  corresponde a  $f$  rotacionada em  $180^\circ$  (caso  $f$  seja simétrica, as operações de correlação e convolução são equivalentes).

Do ponto de vista conceitual, o interesse em representar uma operação por meio da convolução reside no fato de que esta está intimamente ligada ao produto das transformadas de Fourier de  $f$  e  $u$ , permitindo assim uma análise no domínio da frequência. Em uma dimensão e em domínio contínuo, temos:

$$\mathcal{F}(w(x)) = \mathcal{F}((f * u)(x)) = \mathcal{F}(f(x)) \cdot \mathcal{F}(u(x)).$$

Ou seja, a transformada de Fourier da convolução entre  $f(x)$  e  $u(x)$  é o produto das transformadas individuais:

$$\mathcal{W}(\omega) = \mathcal{F}(\omega) \cdot \mathcal{U}(\omega) \quad (3.9)$$

Este é um resultado bastante conhecido [27, 62, 53] e pode ser estendido para o caso bidimensional:

$$\mathcal{W}(\omega_1, \omega_2) = \mathcal{F}(\omega_1, \omega_2) \cdot \mathcal{U}(\omega_1, \omega_2), \quad (3.10)$$

em que  $\omega_1$  e  $\omega_2$  correspondem às frequências espaciais.

Levadas em conta as devidas questões sobre discretização dos sinais, todas as características da transformada de Fourier podem ser exploradas no contexto de imagens digitais. Ou seja, analogamente ao processamento de sinais unidimensionais, pode-se processar imagens no domínio da frequência. Apesar de não ter sido usada neste trabalho, a análise no

domínio das frequências é útil para se entender o princípio de filtros passa-baixa, passa-alta e passa-banda. Por isso, as discussões a seguir fazem uso destes conceitos.

Uma descrição alternativa de uma imagem *digital*  $u$  consiste em tomarmos sua Transformada de Fourier Discreta (DFT<sup>4</sup>). Nesta representação, podemos avaliar a composição espectral do sinal original; isto é, como este é composto em termos de frequências. Transições abruptas - tais como bordas bem definidas - correspondem a componentes de alta frequência, ao passo que regiões de pouca - ou nenhuma - variação espacial estão associadas a componentes de baixa frequência.

### 3.2.1 Filtros passa-baixas

Como primeira abordagem para o processo de filtragem no domínio das frequências, tomemos o caso dos filtros passa-baixas. Esse tipo de filtro remove do sinal original as componentes de alta frequência, removendo do sinal ruídos impulsivos, mas sob a pena de remover também detalhes finos. A expressão de um filtro passa-baixas ideal no domínio das frequências é da forma

$$|H(\omega_1, \omega_2)| = \begin{cases} 1; & |r| \leq r_c \\ 0; & \text{caso contrário.} \end{cases} \quad (3.11)$$

Na Equação acima,  $r = \sqrt{\omega_1^2 + \omega_2^2}$ .

Percebe-se que a magnitude do filtro  $H$  consiste em um disco de raio  $r_c$  centrado na origem. Como ilustração, um filtro passa-baixas semelhante ao representado pela Equação (3.11) é aplicado sobre uma imagem teste, tal como mostrado nas Figuras 3.15 e 3.16.



Figura 3.15: Imagem original (à esquerda) e seu espectro discreto de potência, dado pela magnitude da DFT. Para melhor visualização do espectro, este foi submetido a uma transformação logarítmica. A origem do sistema de coordenadas do espectro está localizada no centro da imagem.

Conforme ilustrado na Figura 3.16, pode-se verificar que a aplicação do filtro passa-baixas representado pela Equação (3.11) conserva as componentes espectrais de baixa frequência, levando a uma suavização generalizada da imagem. As ondulações presentes na imagem filtrada são devidas ao *fenômeno de Gibbs* [28, 27, 53], que está associado,

<sup>4</sup>Do inglês, *Discrete Fourier Transform*.



Figura 3.16: Espectro discreto de potências da imagem filtrada por um passa-baixas ideal (à esquerda) e imagem reconstruída por inversão da DFT (à direita). Nota-se a suavização das bordas e a consequente perda dos detalhes finos. As ondulações devem-se à transição abrupta do filtro (vide texto). A origem dos eixos das frequências está localizado no centro da imagem do espectro.

em parte, ao fato de descontinuidades completamente abruptas só poderem ser representadas com um número infinito de componentes espectrais. Como uma representação com infinitos termos é impossível na prática, o truncamento sobre o número de componentes espectrais leva ao surgimento das referidas ondulações. Uma maneira de minimizar esse fenômeno consiste em usar janelas nos sinais [27, 53].

### 3.2.2 Filtros passa-altas

O caso oposto ao filtro passa-baixas é o filtro passa-altas, cuja forma ideal tem a representação

$$|H(\omega_1, \omega_2)| = \begin{cases} 1; & |r| \geq r_c \\ 0; & \text{caso contrário.} \end{cases} \quad (3.12)$$

Nesse caso, o disco delimitado por  $r_c$  define quais frequências internas a ele serão anuladas. Isso acarreta a evidenciação das regiões de maior variação espacial (alta frequência). Devido ao fato de que bordas apresentam transições razoavelmente intensas entre tons de cinza distintos, estes detalhes são preservados, ao passo que regiões de maior uniformidade espacial são suprimidas. Uma ilustração da aplicação deste tipo de filtro pode ser vista na Figura 3.17. Analogamente ao exemplo anterior, o fenômeno de Gibbs pode ser observado na imagem filtrada, dada a variação abrupta do em torno de  $r_c$ .

### 3.2.3 Filtros passa-banda

É possível selecionar um intervalo fechado específico de frequências, diferentemente dos filtros passa-altas e passa-baixas. Isto é, seleciona-se uma região em forma de anel no plano das frequências. Tal tipo de filtro é chamado de filtro passa-banda e seu caso ideal tem a forma

$$|H(\omega_1, \omega_2)| = \begin{cases} 1; & r_{c1} \leq |r| \leq r_{c2} \\ 0; & \text{caso contrário.} \end{cases} \quad (3.13)$$

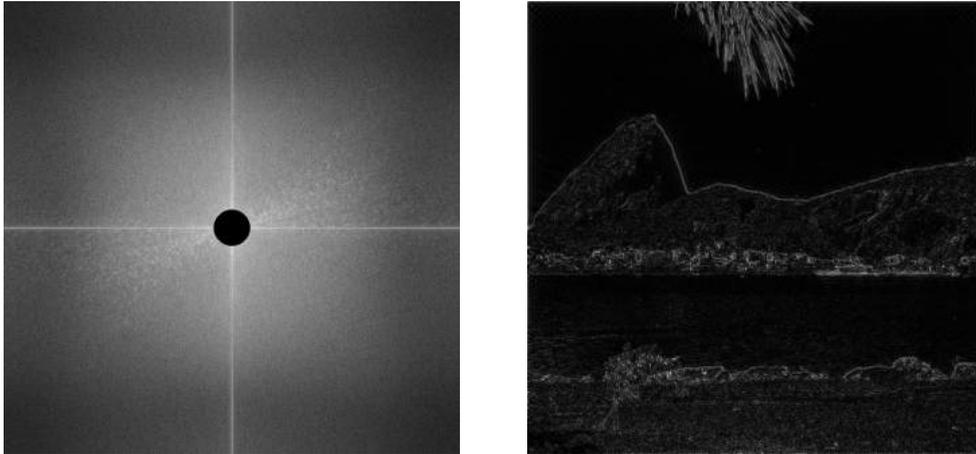


Figura 3.17: Espectro discreto de potências da imagem filtrada por um passa-altas ideal (à esquerda) e imagem reconstruída por inversão da DFT (à direita). Nota-se o realce dos detalhes finos e a evidênciação de bordas. A origem dos eixos das frequências está localizado no centro da imagem do espectro.

Com essa configuração de anel com largura  $|r_{c1} - r_{c2}|$ , o filtro passa banda faz com que o sinal filtrado mantenha características intermediárias quando comparado com os dois filtros anteriores. Isto é, a supressão de ruídos não é extrema, bem como não o é a evidênciação de bordas. Uma ilustração da aplicação deste tipo de filtro pode ser visualizada na Figura 3.18.



Figura 3.18: Espectro discreto de potência do produto entre a DFT da imagem original e o filtro passa-banda (esquerda) e a imagem reconstruída por inversão da DFT (à direita). Nesse caso, há a preservação de um intervalo fechado de frequências, permitindo a associação das características básicas dos filtros passa-baixas e passa-altas. A origem dos eixos das frequências está localizado no centro da imagem do espectro.

### 3.2.4 Filtros Gaussiano e diferenciador

Uma vez explicitados os conceitos fundamentais de filtragem sobre imagens, convém apresentar dois tipos de filtro de especial interesse: Gaussiano e diferenciador. Um filtro Gaussiano corresponde a um passa-baixas cuja transição entre as frequências a serem mantidas e suprimidas não é abrupta. Por sua vez, o diferenciador constitui um filtro

passa-altas, proporcionando aproximações para as derivadas direcionais da imagem original. A seguir, será feita uma discussão sobre os dois filtros listados acima, devido à sua utilização massiva neste trabalho.

### Filtro Gaussiano

Uma função Gaussiana, unidimensional, centrada na média  $\mu$  e com variância  $\sigma^2$  tem a forma

$$G(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x - \mu)^2}{2\sigma^2}\right), \quad (3.14)$$

em que  $\sigma$  corresponde ao *desvio-padrão* da função. O desvio padrão  $\sigma$  está associado à largura da Gaussiana, tal como ilustrado na Figura 3.19.

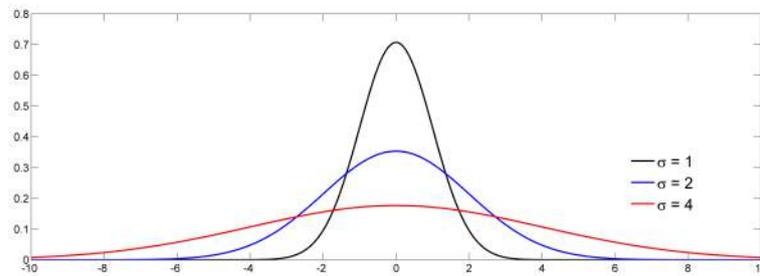


Figura 3.19: Três Gaussianas centradas na origem com diferentes valores de desvio padrão.

Em sua versão bidimensional, uma Gaussiana centrada na posição  $\vec{r}_0 = (x_0, y_0)$  com variância  $\sigma^2$  tem a forma

$$G(\vec{r}, \vec{r}_0, \sigma^2) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-\|\vec{r} - \vec{r}_0\|^2}{2\sigma^2}\right), \quad (3.15)$$

em que  $\vec{r} = (x, y)$  e o operador  $\|\cdot\|$  remete à norma Euclidiana, isto é,  $\|\vec{r} - \vec{r}_0\| = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ . Na Figura 3.20 podem ser visualizadas projeções no plano de três Gaussianas com diferentes valores de  $\sigma$ . Para melhor visualização, as amplitudes foram normalizadas.

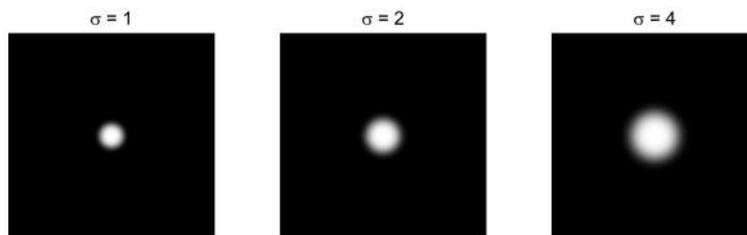


Figura 3.20: Projeções no plano de três Gaussianas bidimensionais centradas na origem com diferentes valores de desvio padrão. Para melhor visualização as Gaussianas tiveram suas amplitudes normalizadas.

A implementação digital de um filtro Gaussiano – isto é, a obtenção dos coeficientes de seu operador – pode ocorrer por amostragem de valores em sua forma contínua, ou seja:

$$\mathbf{G}_n = \frac{1}{2^{2n}} \cdot [g]_{n \times n},$$

em que

$$g_{ij} = \exp\left(\frac{-(i^2 + j^2)}{2\sigma^2}\right) \text{ com } i, j \in \left[-\frac{n-1}{2}, \frac{n-1}{2}\right]. \quad (3.16)$$

Outra possibilidade consiste em representar o filtro Gaussiano 2-D como um produto de filtros Gaussianos 1-D, refletindo o fato de que este é um operador separável [62]:

$$G_n = \frac{1}{2^{2n}} \cdot \begin{bmatrix} \frac{n!}{0!(n-0)!} \\ \frac{n!}{1!(n-1)!} \\ \vdots \\ \frac{n!}{n!(n-n)!} \end{bmatrix} \cdot \begin{bmatrix} \frac{n!}{0!(n-0)!} & \frac{n!}{1!(n-1)!} & \cdots & \frac{n!}{n!(n-n)!} \end{bmatrix} \quad (3.17)$$

Em ambas Equações (3.16) e (3.17), o fator  $\frac{1}{2^{2n}}$  é uma constante de normalização correspondente à soma dos elementos das matrizes em questão.

Uma característica fundamental dos filtros Gaussianos é a sua *isotropia* [62]. Isto é, o filtro opera igualmente em todas as direções, sem privilegiar alguma direção em particular.

Devido ao fato de a transformada de Fourier de uma função Gaussiana também ser uma Gaussiana (vide Apêndice A), constata-se que esse filtro pertence à classe dos filtros passa-baixas. Diferentemente do filtro ideal, a transição entre as componentes espectrais toleradas e suprimidas é suave. Na Figura 3.21 pode ser visualizada a aplicação de um filtro Gaussiano com  $n = 60$  e com  $\sigma = 8$  e  $\sigma = 16$  sobre a imagem-teste apresentada na Figura 3.1.



Figura 3.21: Aplicação de filtros Gaussianos de tamanho  $n = 60$ ,  $\sigma = 8$  (imagem à esquerda) e  $\sigma = 16$  sobre a imagem-teste apresentada na Figura 3.1. Nota-se em ambos os casos a supressão de detalhes finos, característica de filtros passa-baixa. Contudo, devido ao perfil espectral suave do filtro, o fenômeno de Gibbs é praticamente inexistente.

Para finalizar a discussão sobre filtros Gaussianos, cabe ressaltar que há uma outra maneira de se implementar esse tipo de filtro: ele pode ser implementado iterativamente devido à sua relação com processos de difusão lineares. Esta relação será vista em detalhes no Capítulo 4.

### Diferenciador

Uma operação típica em sinais consiste em obtermos suas derivadas. No caso de sinais multidimensionais, é necessária a especificação da direção em que a derivada será tomada. Uma primeira aproximação das derivadas de um sinal (contínuo)  $f(x)$  consiste em truncar sua série de Taylor:

$$f(x+h) = \sum_{n=0}^{\infty} \frac{h^n}{n!} f^{(n)}(x). \quad (3.18)$$

Na Equação (3.18) o termo  $f^{(n)}(x)$  corresponde à  $n$ -ésima derivada. O termo em que a série será truncada definirá o erro associado a este procedimento. Duas formas imediatas de aproximação da derivada do sinal  $f(x)$  se dão pelas diferenças *progressivas*,  $\Delta_+ [f(x)] = f(x+h) - f(x)$ , e *regressivas*,  $\Delta_- [f(x)] = f(x) - f(x-h)$ :

$$f'(x) \approx \frac{1}{h} \Delta_+ [f(x)] \quad \text{e} \quad f'(x) \approx \frac{1}{h} \Delta_- [f(x)]. \quad (3.19)$$

Pode-se mostrar [25] que a *diferença centrada*,  $\delta [f(x)] = f(x+h) - f(x-h)$ , constitui melhor aproximação para a derivada  $f'(x)$ :

$$f'(x) \approx \frac{1}{2h} \delta [f(x)] \quad (3.20)$$

Seguindo este procedimento, a estimativa da derivada segunda  $f''(x)$  é obtida por

$$f''(x) \approx \frac{f(x+2h) - 2f(x) + f(x-2h)}{(2h)^2} = \delta_x^2 [f(x)].$$

Uma forma mais frequente para a aproximação de  $f''(x)$  usa o incremento  $\frac{h}{2}$ . Com isso, chega-se em

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}. \quad (3.21)$$

A generalização para funções de mais variáveis é imediata, bastando operar as diferenças em função da variável de interesse ( $x, y, z$  etc.). Com isso, o operador *Laplaciano*,  $\nabla^2(\cdot) \equiv \partial_{xx} + \partial_{yy}$ , pode ser aproximado por

$$\begin{aligned} \nabla^2 f(x, y) &\approx \frac{f(x+h_x, y) - 2f(x, y) + f(x-h_x, y)}{h_x^2} + \\ &\quad \frac{f(x, y+h_y) - 2f(x, y) + f(x, y-h_y)}{h_y^2} \\ &= \frac{1}{h_x^2} \delta_x^2 [f(x, y)] + \frac{1}{h_y^2} \delta_y^2 [f(x, y)]. \end{aligned} \quad (3.22)$$

Os conceitos de diferenças finitas aqui apresentados podem ser aplicados a sinais discretos, tal como imagens digitais. A única diferença fundamental remete ao fato de que o refinamento da malha é limitado. Ou seja, o incremento  $h$  não pode tender a zero indefinidamente. Tipicamente, usa-se  $h = 1$ . Dessa forma, a vizinhança de um pixel segue a notação discutida na Seção 3.1 (vide Figura 3.8). Com base nessas informações, podemos definir uma série de filtros diferenciadores para imagens.

A diferença progressiva direcional pode ser aproximada por uma matriz cujas entradas correspondem aos coeficientes das aproximações via diferenças finitas:

$$\mathbf{D}_{x_+} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{D}_{y_+} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad (3.23)$$

De maneira análoga, é possível definir filtros de diferenças regressivas e centradas:

$$\mathbf{D}_{x_-} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{D}_{y_-} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad (3.24)$$

$$\mathbf{D}_{x_c} = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{D}_{y_c} = \frac{1}{2} \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \quad (3.25)$$

Por sua vez, o filtro Laplaciano tem a forma

$$\mathbf{L} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (3.26)$$

Os coeficientes nulos nos operadores foram inseridos para proporcionar uma matriz quadrada de dimensão ímpar, visando facilitar a centralização do pixel a ser operado.

Cabe aqui uma observação importante: os filtros diferenciadores mostrados nas Equações (3.23), (3.24), (3.25) e (3.26) foram definidos segundo um esquema de vizinhança  $N_4$  (vide Seção 3.1). Essa constatação surge do fato que os pixels nas diagonais do pixel central não são levados em consideração no cômputo das diferenças. Ao se adotar um esquema de vizinhança  $N_8$ , as diagonais devem ser consideradas.

Dois filtros diferenciadores que adotam vizinhança  $N_8$  e que foram usados neste trabalho são os filtros de Prewitt [62, 69],  $\mathbf{P}$ , e Sobel [62],  $\mathbf{S}$ :

$$\mathbf{P}_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{P}_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \quad (3.27)$$

$$\mathbf{S}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{S}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \quad (3.28)$$

Para ilustrar a aplicação dos filtros diferenciadores, usaremos a imagem sintética retratada na Figura 3.22. Este tipo de figura, com padrões espaciais bem definidos, é ideal para ilustrar a ação de filtros passa-alta, dado que as bordas constituem regiões de variação espacial abrupta. Adicionalmente ao padrão espacial, foi inserido na imagem-teste um gradiente de transição suave no fundo.

Nota-se pelas Figuras 3.23 e 3.24 que, de fato, os filtros apresentados nas Equações (3.27) e (3.28) constituem boa aproximação para operadores diferenciais em ambas direções espaciais. A aplicação destes dois filtros é importante na detecção de bordas e contornos,

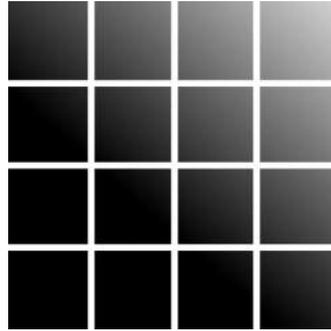


Figura 3.22: Imagem sintética usada para ilustrar a aplicação dos filtros diferenciadores. Esta imagem é constituída por um padrão espacial retangular ideal para supressão segundo a filtragem por diferenciadores.

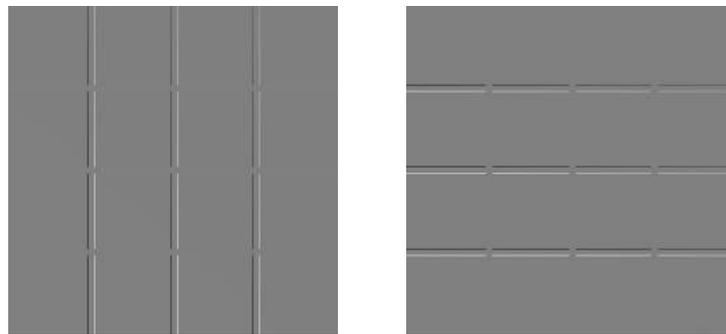


Figura 3.23: Aplicação do filtro de Prewitt sobre a imagem-teste nas direções  $x$  (imagem à esquerda) e  $y$ . Nota-se a evidência das bordas verticais e horizontais, respectivamente. O gradiente de fundo da imagem original tem pouca influência pois sua variação é suave, não permanecendo nas imagens filtradas.

sendo esta uma das principais abordagens em segmentação de imagens, a ser discutida em detalhes na Seção 3.3.

Convém notar que é possível obtermos uma estimativa da *magnitude do gradiente* da imagem,  $\|\vec{\nabla}f(x, y)\|$ :

$$\|\vec{\nabla}f\| = \sqrt{\mathbf{D}_x^2[f] + \mathbf{D}_y^2[f]}, \quad (3.29)$$

em que  $\mathbf{D}$  representa algum operador diferencial direcional de primeira ordem. Este recurso é usado tipicamente na evidência de bordas da imagem original. Além disso, conhecido  $\|\vec{\nabla}f(x, y)\|$ , é possível determinar a direção em que os níveis de cinza crescem localmente. A Figura 3.25 exemplifica a determinação de  $\|\vec{\nabla}f\|$  para a imagem-teste apresentada na Figura 3.22. As derivadas direcionais foram estimadas pelo operador de Prewitt.

Para finalizar a discussão sobre filtros diferenciadores, é importante ressaltar que estes filtros constituem passa-altas. Portanto, são bastante sensíveis a ruídos (principalmente pontuais). Isso deve ser levado em consideração quando da utilização destes operadores. A Figura 3.26 traz a estimativa para o gradiente da Figura 3.1 como exemplo de gradiente de imagens reais, mais suscetíveis a ruídos diversos.

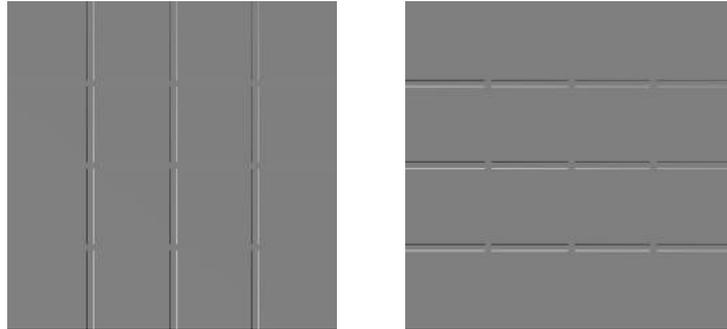


Figura 3.24: Aplicação do filtro de Sobel sobre a imagem-teste na direções  $x$  (imagem à esquerda) e  $y$ . Nota-se a evidência das bordas verticais e horizontais, respectivamente. O gradiente de fundo da imagem original tem pouca influência pois sua variação é suave, não permanecendo nas imagens filtradas.

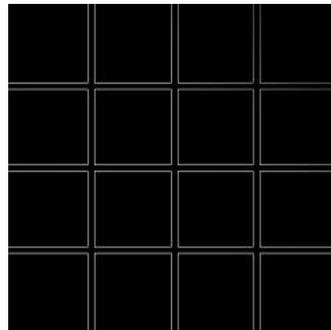


Figura 3.25: Estimação do gradiente da imagem-teste apresentada na Figura 3.22. As derivadas direcionais foram aproximadas pelo operador de Prewitt. Nota-se o destaque eficiente das bordas da imagem, quase não sendo perceptível o efeito do gradiente (suave) de fundo. Para melhor visualização, a imagem resultante foi submetida a um alargamento linear de contraste.

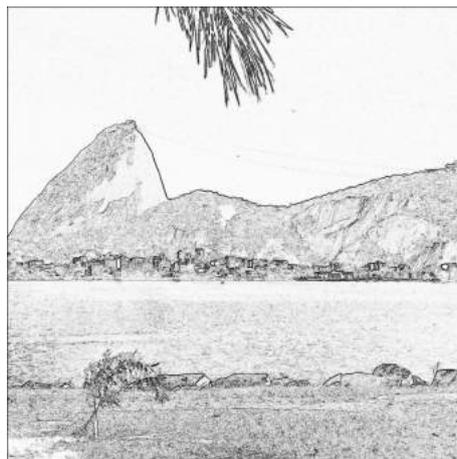


Figura 3.26: Estimação do gradiente da imagem-teste apresentada na Figura 3.1. As derivadas foram aproximadas pelo operador Sobel. Nota-se o destaque eficiente das bordas, mas com a presença de ruídos relativos aos detalhes naturais da imagem, tais como a água ou a grama. Para melhor visualização, a imagem resultante foi submetida a inversão (negativo) e a um alargamento linear de contraste.

### 3.3 Segmentação de imagens

Todo o ferramental discutido anteriormente sobre operações em imagens é comumente direcionado à *segmentação* de imagens. De uma maneira geral, segmentar uma imagem consiste em particioná-la em suas componentes constituintes. Geralmente, o resultado obtido é o destaque dos *objetos* da imagem em relação ao *fundo*. De um ponto de vista formal, uma imagem  $u$  *idealmente* segmentada pode ser compreendida como a união de  $N$  regiões  $R_i$  disjuntas entre si, ou seja

$$u = \bigcup_{i=1}^N R_i \quad \text{e} \quad R_i \cap R_j = \emptyset, \text{ se } i \neq j.$$

Há uma subjetividade intrínseca na questão da segmentação de imagens: a decomposição da imagem em componentes depende da definição destes. Decorre disto que processos de segmentação que atendem requisitos para certa classe de imagens podem funcionar muito mal na segmentação de outra classe. Em síntese, não há, até hoje, uma *teoria da segmentação de imagens*. Isto quer dizer que os processos desenvolvidos para segmentar imagens guardam relações profundas com a estrutura destas, não permitindo generalizações.

Conforme dito, a subjetividade acerca das definições de objeto e fundo representa uma dificuldade natural do processo de segmentação. Entretanto, este ponto é mais profundo: podemos dizer que o processo de segmentação é a etapa mais crítica do processamento de imagens, pois é o resultado da segmentação que será processado e avaliado por etapas posteriores. O diagrama retratado na Figura 3.27 ilustra a organização básica de um sistema de processamento de imagens.

Há duas abordagens básicas para segmentação de imagens [1] :

- Técnicas baseadas em similaridade entre pixels; e
- Técnicas baseadas em descontinuidade entre pixels.

Estas duas abordagens podem ser combinadas. A seguir será feita discussão sobre duas técnicas básicas destas duas classes. A primeira delas, limiarização, faz uso da semelhança entre os valores dos pixels; a segunda, detecção de bordas, faz uso de descontinuidades entre pixels.

#### 3.3.1 Limiarização

Conforme discutido na Seção 3.1, o histograma de uma imagem apresenta informações sobre a distribuição de seus valores de intensidade. Este perfil de distribuição pode ser utilizado para definir duas classes na imagem: objeto e fundo. Para tanto, é aplicada sobre a imagem uma transformação da forma

$$T(u_{ij}) = \begin{cases} 0, & \text{se } u_{ij} \leq L \\ 1, & \text{caso contrário} \end{cases} \quad (3.30)$$

Na Equação (3.30), o parâmetro  $L$  corresponde ao *limiar* de corte <sup>5</sup>, daí o nome da técnica. Devido à forma da transformação  $T$ , esta operação é frequentemente chamada de binarização.

<sup>5</sup>Usualmente denominado por sua forma inglesa, *threshold*.

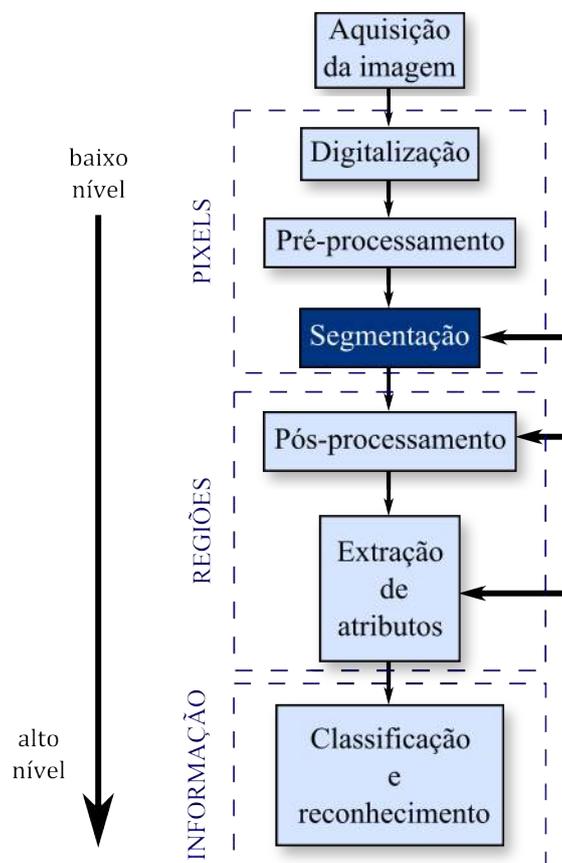


Figura 3.27: Diagrama representativo das etapas de um sistema de processamento de imagens. As etapas de pré-processamento e segmentação operam sobre os pixels da imagem. A transformação de pixels (dados) em informação se dá ao longo das etapas de pós-processamento e extração de atributos, que podem envolver realimentação, isto é, estas etapas são interdependentes juntamente com a segmentação. Adaptado de Albuquerque *et al.* [1].

A técnica de binarização funciona bem desde que o histograma da imagem apresente um perfil aproximadamente *bimodal*, pois isto representa o fato de que os valores de intensidade da imagem estão estatisticamente separados em, pelo menos, duas classes distintas. A Figura 3.28 ilustra essa característica.

Naturalmente, um histograma perfeitamente bimodal constitui uma idealização. Contudo, há situações em que essa aproximação é viável, tal como pode ser visualizado na Figura 3.9 (Seção 3.1). Entretanto, não raro lidamos com imagens cujo histograma em nada se aproxima a uma distribuição bimodal, tal como a imagem apresentada na Figura 3.29.

Justamente pelo fato de lidar-se frequentemente com histogramas sem perfil bimodal, a escolha de um ou mais limiares constitui tarefa não trivial.

Dentre as mais diversas abordagens para determinação do limiar ótimo, uma característica bastante frequente destaca-se: a busca por um limiar  $L_0$  que otimize (maximize ou minimize) alguma função critério. A função critério pode corresponder, por exemplo, a uma distância entre duas distribuições (método do erro mínimo), à entropia do histograma (métodos entrópicos) ou à razão entre variâncias (método de Otsu). Revisões sobre técnicas de limiarização podem ser encontradas nos trabalhos de Sahoo *et al.* [74], Weszka [92] e Chang *et al.* [21]. No Apêndice B constam descrições dos métodos utilizados neste

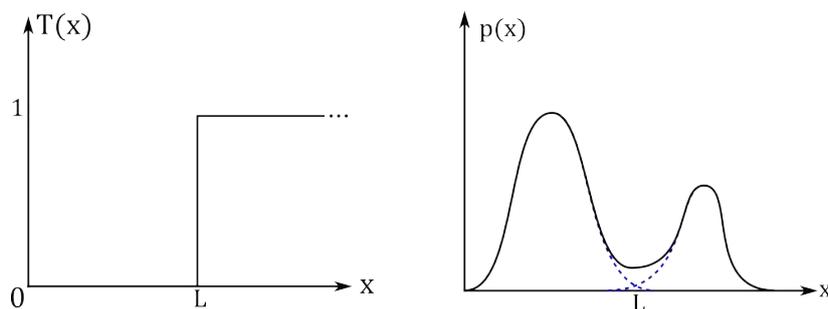


Figura 3.28: Transformação global para binarização (à esquerda). Esta transformação anula os pixels cujos valores estão abaixo do limiar  $L$ , enquanto conserva os demais. O limiar  $L$  é facilmente obtido quando o histograma apresenta perfil bimodal (figura à direita), pois corresponde ao valor de separação entre os dois pontos de máximo no histograma. Em pontilhado, é mostrado que um histograma bimodal pode ser interpretado como a composição de duas distribuições.

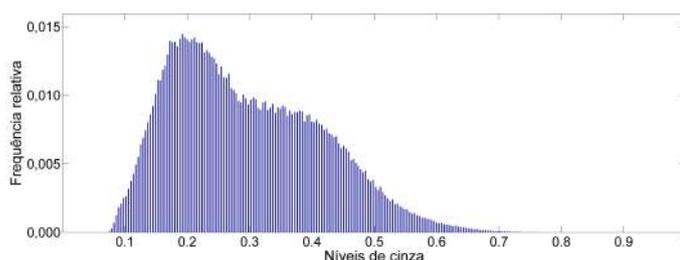
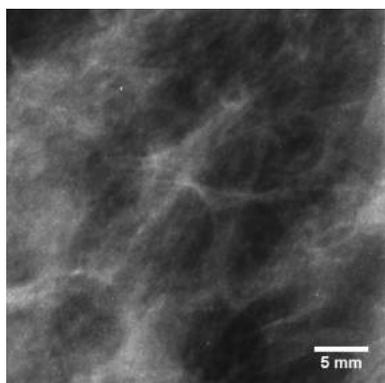


Figura 3.29: Fragmento de uma mamografia (à esquerda) e seu histograma. Nota-se que o histograma não possui perfil bimodal, dificultando a escolha inicial do limiar.

trabalho. Na Figura 3.30 pode ser vista a aplicação de três binarizações sobre a imagem apresentada na Figura 3.29. Os limiares foram calculados pelo métodos do valor médio, Otsu e da máxima entropia.

Para finalizar a discussão sobre a técnica de limiarização, é importante ressaltar que a abordagem usada neste trabalho consiste em uma limiarização *global*, no sentido que o histograma é obtido sobre o conjunto de todos os pixels da imagem em questão. Embora isso simplifique o processo, atributos ou características locais não são considerados. Isso pode acarretar a segmentação inadequada de estruturas na imagem devido a diferenças locais. Nesse sentido, é possível que sejam aplicadas técnicas de limiarização *local*, que consistem basicamente em analisar atributos locais dos pixels para a definição de limiares para cada região delimitada.

### 3.3.2 Detecção de bordas

Diferentemente da limiarização, que opera via semelhança entre pixels, a detecção de bordas funciona com base em *descontinuidades* entre pixels. Idealmente, uma borda corresponde a uma transição completamente abrupta ao longo de certa direção da imagem. A aproximação para a borda ideal é a função degrau, tal como ilustrado na Figura 3.31. O princípio de detecção de bordas já foi discutido na Seção 3.2: descontinuidades são eviden-

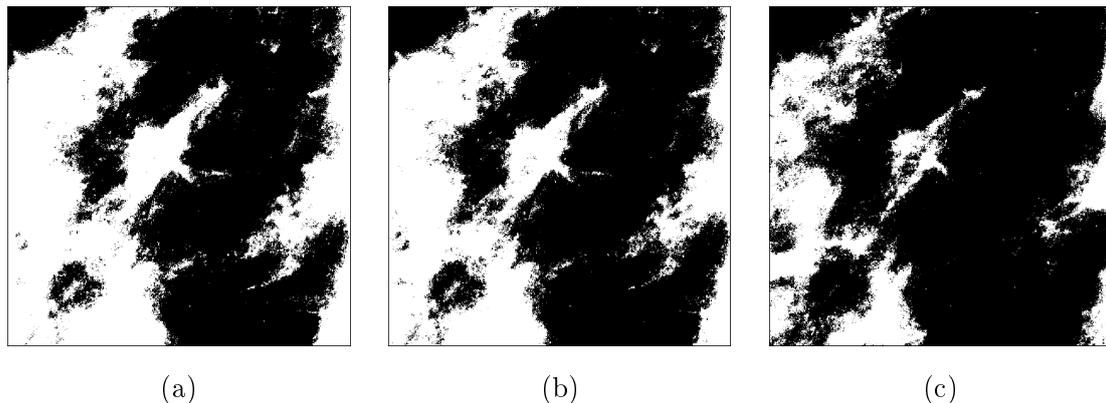


Figura 3.30: Aplicação de limiarização global sobre o fragmento de imagem mamográfica da Figura 3.29. Em (a), o limiar foi estimado pelo valor médio da distribuição ( $L_0 = 0,2955$ ); em (b), pelo método de Otsu ( $L_0 = 0,3137$ ) e em (c) pelo método da máxima entropia ( $L_0 = 0,3843$ ). Nota-se que, para esta imagem em particular, os métodos da média e de Otsu proporcionaram resultados semelhantes. Os valores de intensidade e de limiar foram normalizados para o intervalo  $[0,1]$ .

ciadas por meio de filtros passa-alta. Nesse sentido, os filtros diferenciadores desempenham papel importante, pois permitem a evidenciação das bordas, e quando combinados com a operação de limiarização, permitem sua segmentação.

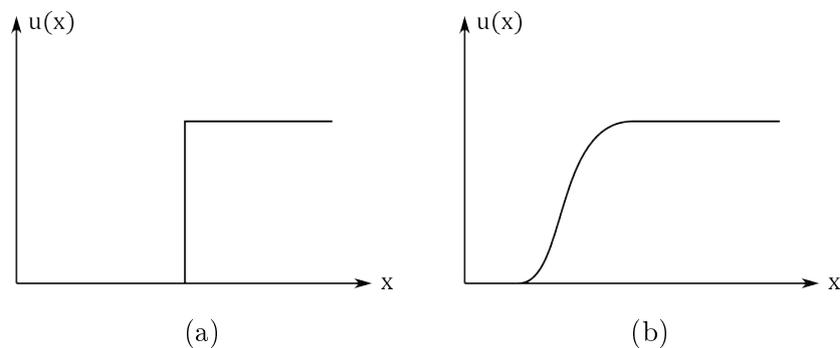


Figura 3.31: Modelos para borda em uma dimensão. Em (a) é mostrado a transição entre níveis de intensidade uma borda ideal, completamente descontínua. Em (b) é mostrado um modelo mais realista, que leva em consideração a possibilidade de a transição entre níveis de intensidade não ocorrer de forma abrupta, diminuindo a efetividade dos detectores.

Uma dificuldade intrínseca em detectar bordas via diferenciação está associada ao fato de que bordas abruptas não costumam ser frequentes em imagens naturais<sup>6</sup>. Nesse sentido, quão mais suave for a transição entre os níveis de intensidade (vide Figura 3.31 (b)), menos eficaz será um filtro diferenciador. No caso limite de não haver transição alguma (níveis de intensidade constantes ao longo de certa direção), o resultado da filtragem via diferenciador será nulo.

Para melhor ilustrar a segmentação por detecção de bordas, tomemos o fragmento de imagem mamográfica mostrado na Figura 3.32. Este fragmento pode ser considerado um

<sup>6</sup>Entenda-se *natural* no sentido de não-sintética, de forma que padrões espaciais regulares tais como os representados na Figura 3.22 não sejam predominantes.

caso emblemático das dificuldades em se detectar estruturas na mama: além da baixa dimensionalidade das estruturas suspeitas, a imagem possui baixo contraste.

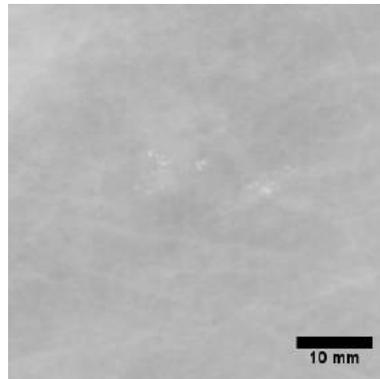


Figura 3.32: Fragmento de imagem mamográfica contendo um agrupamento de microcalcificações nas imediações do centro da imagem. Note-se que esta imagem possui baixa relação de contraste, motivando a aplicação da operação de alargamento linear de contraste.

Em primeiro lugar, é interessante expandir o histograma desta imagem, isto é, aumentar o contraste da imagem. Isso pode ser feito através do alargamento linear de contraste, definido na Equação (3.5). Feito isso, submete-se a imagem resultante a duas filtragens – uma em cada direção – segundo um filtro diferenciador. Isto é feito visando a obtenção da magnitude do gradiente, tal como definido na Equação (3.29). A Figura 3.33 ilustra esta ação. O filtro diferenciador usado foi o de Prewitt, tal como definido na Equação (3.27).

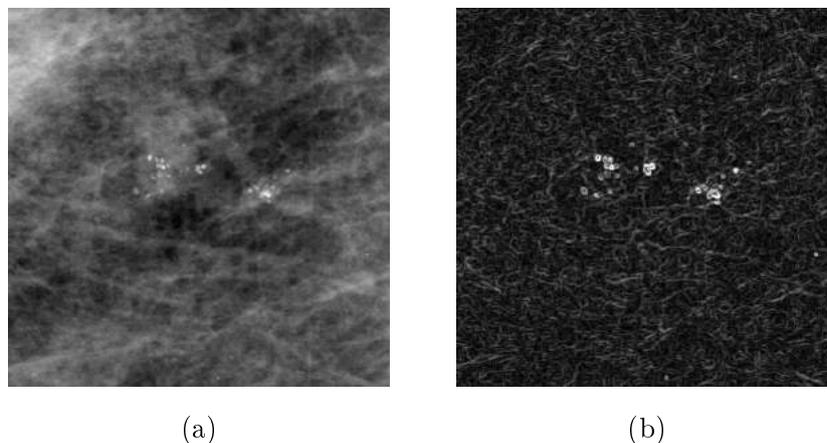


Figura 3.33: Etapas da segmentação por detecção de bordas da imagem retratada na Figura 3.32. Em (a) é mostrado o resultado da operação de alargamento linear de contraste da imagem original e em (b) é mostrada a magnitude do gradiente, calculada usando-se o filtro de Prewitt. Nota-se que, ao final destas duas etapas, as microcalcificações estão significativamente mais evidenciadas do que na imagem original.

O resultado da detecção de bordas mostrado na Figura 3.33 (b) mostra que, de fato, filtros diferenciadores são eficientes em evidenciação de discontinuidades. Contudo, a *segmentação* completa somente será atingida após a limiarização da imagem processada. Isto se faz necessário de forma a proporcionar a separação objeto/fundo, tal como discutido no início desta Seção.

A criticalidade do processo de limiarização fica evidente quando comparamos as imagens resultantes de diferentes formas de cálculo do limiar, tal como ilustrado na Figura 3.34. Nela fica evidente o papel desempenhado pelo *threshold* de separação. Vê-se que o cômputo deste limiar pode variar bastante segundo a técnica escolhida, e isso traz consequências importantes sobre o sinal resultante. Conforme será visto adiante, o resultado da segmentação é utilizado para as etapas de extração de informação (tal como esquematizado na Figura 3.27). Portanto, a obtenção de um sinal ruidoso ou inadequado na etapa de segmentação prejudica a eficiência das etapas posteriores. Pode-se, sem sombra de dúvida, afirmar que a segmentação é a etapa mais importante da cadeia de processamento de uma imagem.

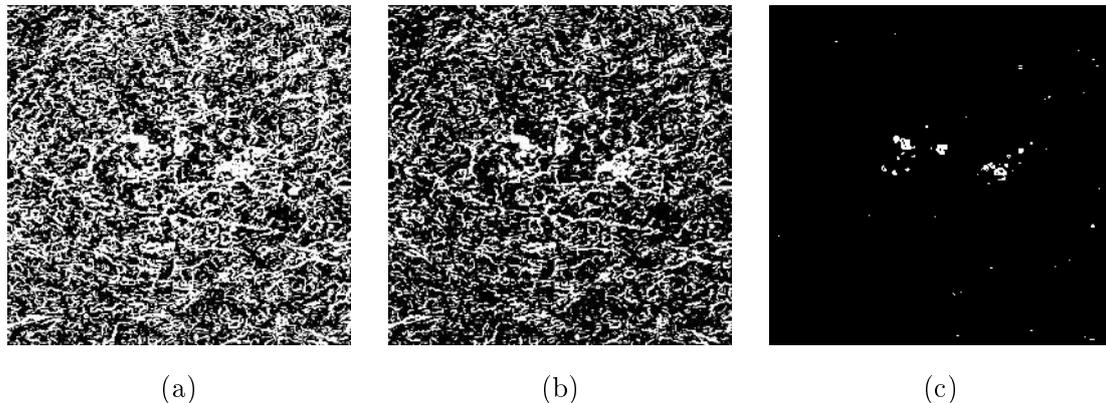


Figura 3.34: Resultados da binarização da imagem com bordas realçadas mostrada na Figura 3.33: em (a), o *threshold* foi calculado segundo o valor médio da distribuição. Em (b), foi calculado pelo método de Otsu e em (c) pelo método da máxima entropia. Nota-se claramente os efeitos de uma escolha inadequada na forma de cômputo do limiar: enquanto os métodos do valor médio e de Otsu evidenciam uma quantidade significativa de ruído, o método da máxima entropia permite uma segmentação muito mais eficiente dos objetos de interesse.

## 3.4 Operações de alto nível: labeling e clustering

Os tópicos discutidos até agora enquadram-se como processamento de baixo nível (vide esquema da Figura 3.27), pois as técnicas discutidas fazem uso da informação contida nos próprios pixels da imagem. As operações a serem discutidas nesta seção devem ser encaradas como operações de *alto-nível*, pois fazem uso de informação em nível mais abstrato do que aquela proporcionada pelos pixels da imagem. Estas operações, *labeling* e *clustering*, operam sobre a imagem segmentada a fim de proporcionar informações sobre as formas presentes na imagem, e como estas se agrupam. A seguir, serão apresentados os conceitos básicos destas classes de operações. Referências sobre clustering podem ser acessadas nos trabalhos de Pedrini e Schwartz [62] e principalmente nos trabalhos de Jain *et al.* [42] e Jain e Dubes [43]. Referências sobre labeling podem ser encontradas nos trabalhos de Rosenfeld e Kak [72], Gonzalez *et al.* [30], He *et al.* [35] e Gotoh *et al.* [31].

### 3.4.1 Labeling

De posse de uma imagem segmentada, o passo natural a ser seguido é a obtenção de informações sobre as formas presentes nos objetos constantes nesta imagem. Medidas

diversas podem ser realizadas, tais como cômputo de áreas, rugosidade, perímetro etc. Estas medidas fazem parte do conjunto dos *descritores* de forma dos objetos estudados e serão discutidas mais adiante. O que se deve compreender é que, para que qualquer medida seja realizada sobre objetos internos à imagem, estes devem ser reconhecidos. Isto é, os objetos presentes na imagem devem ser inicialmente *rotulados*. Usualmente, usamos o termo em inglês, *labeling*. A seguir, será feita breve discussão sobre o processo de labeling segundo a técnica básica de *two-raster scan* [72, 35] por sua simplicidade didática.

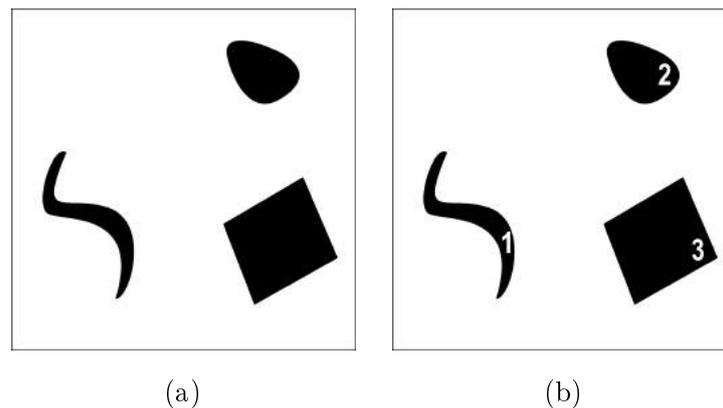


Figura 3.35: Exemplificação do conceito de *labeling*: em (a) é apresentada uma imagem segmentada contendo três formas arbitrárias. Em (b) é dado um *rótulo* para cada uma destas formas, de modo que seja possível em etapas posteriores do processamento identificar estas estruturas e extrair informações descritivas diversas.

Para exemplificar o processo de labeling, tomemos a imagem apresentada na Figura 3.36. Nesta figura temos, além da imagem propriamente dita, sua representação matricial, em que pixels em branco correspondem ao valor 1 e pixels pretos correspondem a 0. Nota-se que há quatro formas distintas na imagem. Conforme discutido nos parágrafos anteriores, o objetivo é rotular estas formas.

O processo de labeling pode ser entendido como uma varredura sobre as linhas da imagem com uma máscara previamente definida, cujo objetivo é delimitar regiões da imagem e permitir a rotulação dos elementos evidenciados. Contudo, pode ocorrer de um pixel ser atribuído a mais de um rótulo, ou *label*. Por isso, uma segunda varredura é feita, corrigindo estas ambiguidades. Convém ressaltar que a forma da máscara varia conforme a vizinhança é definida: se a vizinhança for do tipo  $N_4$  a máscara usada para varredura abrange o pixel em questão, seu vizinho imediato à esquerda e seu vizinho imediato superior; ao passo que em um esquema de vizinhança  $N_8$ , são abrangidos –adicionalmente aos pixels à esquerda e acima– os pixels das diagonais superiores. Estas máscaras estão ilustradas na Figura 3.37.

Conforme discutido antes, o processo de labeling consiste em realizar uma varredura na imagem atribuindo labels aos pixels. Na Figura 3.38 são mostrados os resultados da primeira varredura, em ambos esquemas  $N_4$  e  $N_8$ , para labeling da imagem retratada na Figura 3.36. Cada label é representado por um caractere que o difere de outros rótulos.

Pode ser verificado na Figura 3.38 que ocorre uma equivalência entre os labels b e c no esquema  $N_4$ , mostrando a necessidade da segunda varredura: é nesta etapa que as equivalências são resolvidas e o processo é finalizado. A forma como a segunda varredura ocorre determina a eficiência do algoritmo envolvido [72, 31, 35].

Uma vez rotulados, os objetos segmentados da imagem podem ser usados para extração de informação quantitativa, proporcionada pelo uso de descritores, a serem discutidos a

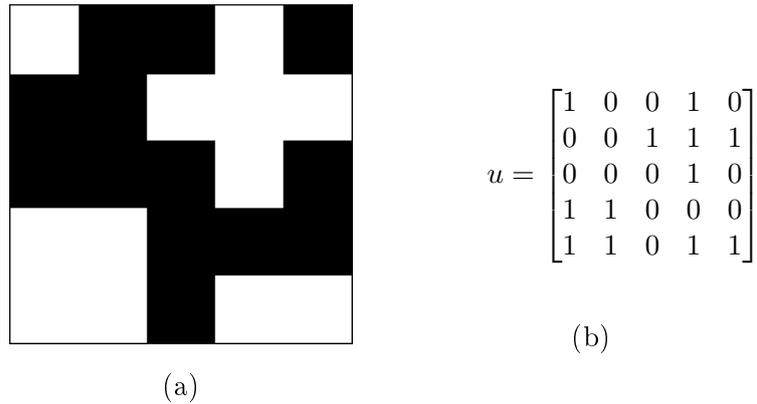


Figura 3.36: Em (a): Imagem a ser usada para exemplificação de labeling sobre vizinhanças  $N_4$  e  $N_8$ . Em (b) é apresentada sua forma matricial, em que 1 equivale a branco e 0 a preto.

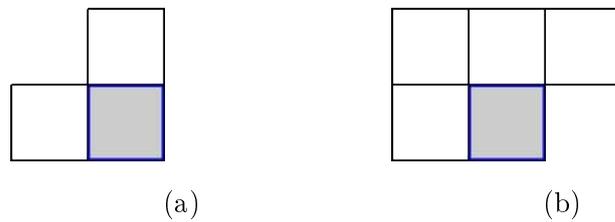


Figura 3.37: Máscaras para varredura no processo de labeling. Em (a), é mostrada a máscara usada para varreduras em um esquema de vizinhança  $N_4$ , em que apenas os pixels superior e à esquerda do pixel atual são abrangidos. Em (b), é mostrada a máscara para um esquema  $N_8$ , em que –adicionalmente aos pixels abrangidos no esquema anterior– são abrangidos os pixels diagonais superiores. As posições associadas ao pixel principal estão destacadas em ambas as máscaras.

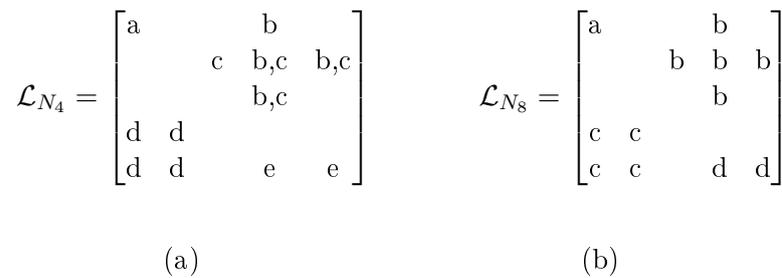


Figura 3.38: Primeira varredura no processo de labeling para a imagem retratada na Figura 3.36: em (a) é mostrada a varredura para um esquema de vizinhança  $N_4$ , ao passo que em (b) é mostrado o resultado da varredura para um esquema  $N_8$ . Nota-se que há uma equivalência entre os rótulos b e c no esquema  $N_4$ , pois estes dizem respeito aos mesmos elementos conectados. Já no caso do esquema  $N_8$ , não há equivalência entre labels, e o processo pode ser finalizado.

seguir.

### 3.4.2 Descritores

Conforme discutido previamente, para que informações quantitativas possam ser extraídas das formas segmentadas previamente, é necessário que estas formas sejam identificadas por meio de rótulos. O passo seguinte consiste na extração de informação propriamente dita. Ou seja, conhecida uma forma qualquer dentro da imagem, que atributos podemos levantar sobre este objeto?

Um *descriptor* de uma forma pode ser compreendido como um atributo representativo da forma analisada. Usualmente, divide-se os descritores em duas grandes classes:

- Descritores de borda;
- Descritores de região.

A primeira classe diz respeito a descrição da *forma* do objeto analisado, levando-se em conta seu *contorno*. Já a segunda, a *área* da forma também é levada em consideração no cômputo das grandezas de interesse. Neste trabalho não foi feito uso de descritores de forma baseados em borda, de modo que esta classe não será abordada. Adicionalmente, há uma variedade enorme de descritores de forma que não será discutida aqui. Para aprofundamento, recomenda-se os trabalhos de Zhang e Lu [95], Pedrini e Schwarz [62] e Pavlidis [61].

Para exemplificar o uso de descritores, será usada a imagem sintética mostrada na Figura 3.39. Esta imagem contém 3 objetos: um círculo de raio igual a 200 px; um quadrado de lado igual a 270 px e uma forma arbitrária. A imagem possui resolução de 20px/mm, o que permite a conversão de medidas em pixels para medidas em milímetros.



Figura 3.39: Imagem sintética a ser usada para exemplificar a utilização de descritores baseados em região.

A seguir serão definidos três descritores básicos a serem usados no exemplo.

**Área:** Dada uma imagem *binária*,  $u$ , com dimensões  $M \times N$ , sua área é definida por:

$$A(u) = \sum_{i=1}^M \sum_{j=1}^N u_{ij} \quad (3.31)$$

No caso da imagem mostrada na Figura 3.39, a área total é igual a  $229328 \text{ px}^7$ . A Equação (3.31) anterior pode ser particularizada para regiões internas à imagem, dado que a área

<sup>7</sup>Deve-se notar que, tal como definida na Equação (3.31), a área de uma região qualquer possui unidades de pixel, ao invés de pixel quadrado ( $\text{px}^2$ ). Isso é devido ao fato de que o pixel, em uma imagem binária, pode ser tomado como um elemento de área. Entretanto, quando conversões entre unidades reais e pixels ocorrerem, deve-se atentar para a potência omitida. Por exemplo, um quadrado de lado 200 px tem sua área igual a 40 000 px; em uma imagem com resolução de 2px/mm, ele teria uma área real de 10 000  $\text{mm}^2$ .

total da imagem é a soma das áreas que a compõem. Sendo assim, cada forma presente em uma imagem pode ter sua área estimada, bastando aplicar a expressão (3.31) sobre os índices que compreendem a estrutura de interesse. Para fins de tabulação, rotulemos o círculo como forma I, o quadrado como forma II e a forma arbitrária como III. Na tabela 3.1 são mostradas as áreas calculadas para as formas presentes na imagem retratada na Figura 3.39. Como a resolução da imagem é conhecida, é possível estimar a área real dos objetos na imagem.

ID forma	Área (px)	Área (mm <sup>2</sup> )
I	125420	313,55
II	72900	182,25
III	31008	77,52

Tabela 3.1: Cômputo das áreas dos objetos constantes na imagem da Figura 3.39.

As formas I e II podem ter suas áreas estimadas facilmente, devido à sua simplicidade. Quando comparamos os valores calculados via Equação (3.31) com os previstos matematicamente, vemos que - de fato - a estimativa está dentro de uma faixa pequena de erro. O valor previsto para a área da forma I (círculo) é de aproximadamente 125663 px, que difere cerca de 0,19% do valor estimado via descritor de área. A área da forma II (quadrado) é de 72900 px, valor idêntico ao estimado pelo descritor de área.

Uma aplicação típica do descritor área é a chamada *filtragem por área*. Este procedimento mantém (ou elimina) formas na imagem que estejam dentro de um certo intervalo. Por exemplo, é possível remover de maneira simples a maior forma presente em uma imagem.

Para finalizar as considerações sobre o descritor área, deve-se ressaltar que este descritor é invariante a translações e rotações das formas<sup>8</sup>, mas não é invariante sobre transformações de escala (ampliação, redução, deformação etc.).

**Retângulo envolvente:** Dada uma forma  $\mathcal{S}$  qualquer pertencente à imagem  $u$ , definimos o retângulo envolvente, ou *bounding box*, de  $\mathcal{S}$  como sendo o menor retângulo que a contém. Na Figura 3.40 pode ser vista uma ilustração do uso deste descritor.

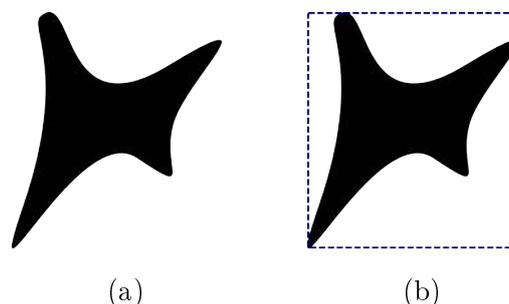


Figura 3.40: Ilustração do uso do descritor retângulo envolvente: em (a) é mostrada uma forma  $\mathcal{S}$  arbitrária, e em (b) é mostrado também seu retângulo envolvente. De um modo geral, esse descritor pode ser caracterizado pelas magnitudes de sua altura e largura.

<sup>8</sup>Os valores do descritor área podem variar um pouco em função de rotações devido à natureza discreta dos dados considerados.

Na Figura 3.41 pode ser vista a superposição dos retângulos envolventes às formas apresentadas na imagem da Figura 3.39

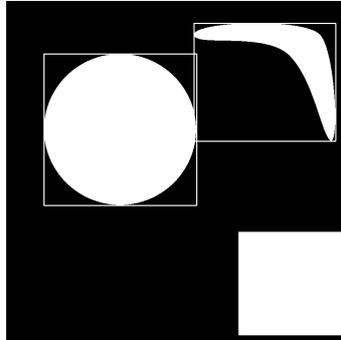


Figura 3.41: Superposição dos retângulos envolventes sobre as formas apresentadas na Figura 3.39. Naturalmente, o retângulo envolvente à forma II (quadrado) coincide com as suas bordas.

O descritor retângulo envolvente possui como quantidades de interesse as suas coordenadas<sup>9</sup> e suas largura e altura. O que este descritor proporciona é a transformação de uma geometria qualquer (tal como a forma III na Figura 3.39) em uma geometria retangular. Quando a forma do objeto não é relevante para a análise a ser realizada, esta simplificação facilita o processo. Isto é particularmente útil quando tratamos da *localização* de objetos na cena. Um exemplo desse tipo de transformação pode ser visualizado na Figura 3.42.

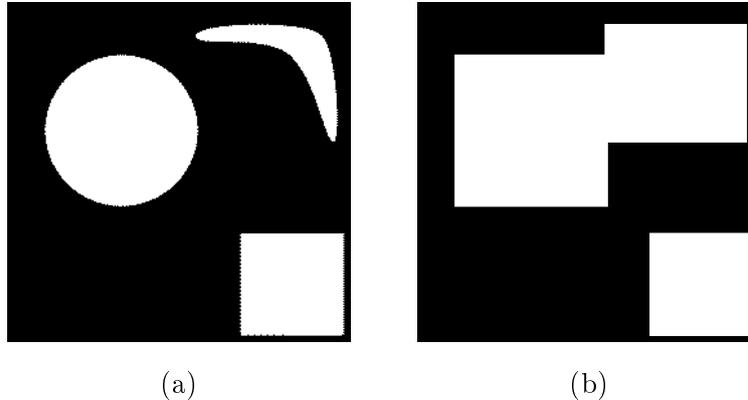


Figura 3.42: Exemplo de transformação de formas: em (a) é mostrada a imagem a ter seus objetos delimitados. Em (b) é mostrada a transformação das formas constantes em (a) para uma geometria mais simples. Este tipo de aplicação é particularmente útil em detecção de objetos em uma cena.

Como o descritor retângulo envolvente leva em consideração as dimensões da forma associada, bem como sua localização, é fácil perceber que ele não é invariante sobre quaisquer transformações de escala, translação, rotação ou deformação.

**Projeções:** Seja  $u$  uma imagem com dimensões  $M \times N$ . Definimos suas projeções nas

<sup>9</sup>As coordenadas podem ser de um dos vértices do retângulo ou de seu centróide.

direção  $x$  e  $y$ , respectivamente, por:

$$P_x(u) = \left( \sum_{i=1}^M u_{i1}, \sum_{i=1}^M u_{i2}, \dots, \sum_{i=1}^M u_{iN} \right) \text{ e } P_y(u) = \left( \sum_{j=1}^N u_{1j}, \sum_{j=1}^N u_{2j}, \dots, \sum_{j=1}^N u_{Mj} \right). \quad (3.32)$$

O que a Equação (3.32) mostra é que a projeção da imagem em uma direção consiste em um sinal 1-D cujas entradas carregam informação sobre as variações da imagem naquela direção. A descrição de uma forma por suas projeções difere de outras técnicas pois gera um sinal a ser analisado, e não um simples escalar (tal como o descritor de área visto anteriormente). Sua utilidade reside no fato de que o sinal resultante da projeção pode ser compreendido como um tipo de assinatura unidimensional de uma forma plana<sup>10</sup>. Na Figura 3.43 podem ser visualizadas as projeções em  $x$  e em  $y$  das formas constantes na imagem apresentada na Figura 3.39.

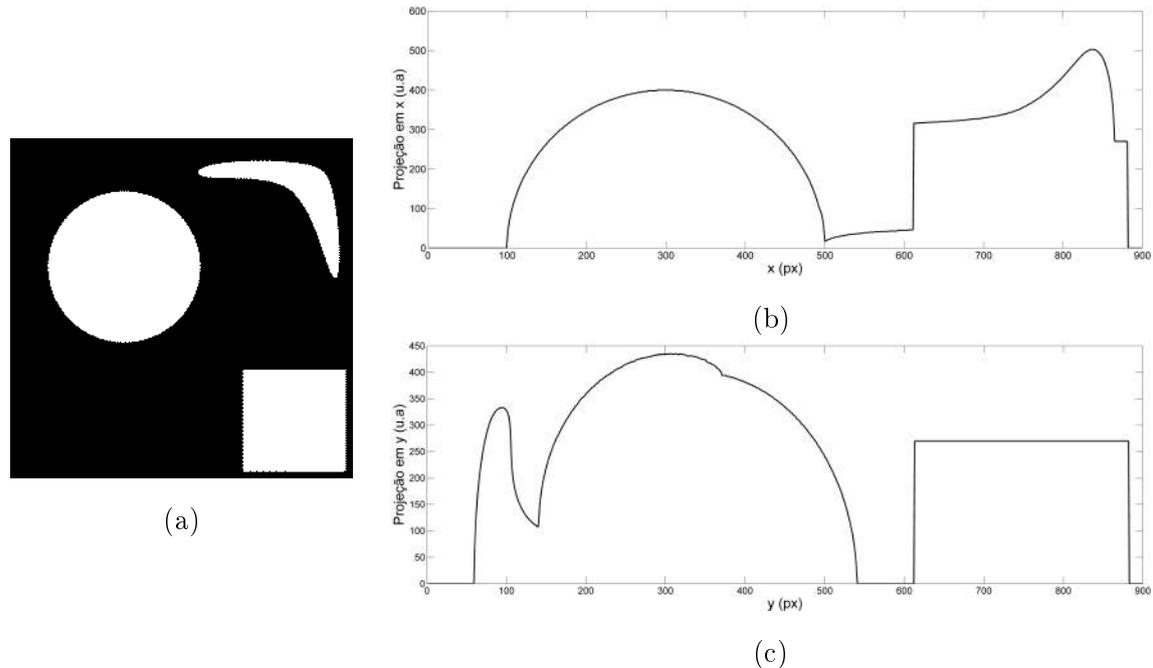


Figura 3.43: Exemplo de obtenção das projeções de uma imagem: em (a) é apresentada a imagem a ser analisada. Em (b) é mostrada sua projeção sobre a direção  $x$  (horizontal) e em (c) é mostrada a projeção em  $y$  (vertical). Diferentemente de outros descritores, a projeção de uma imagem gera um sinal unidimensional, cujas variações correspondem a assinaturas das formas constantes na imagem. Contudo, como pode ser visto pelo exemplo, superposições podem ocorrer, de forma que as assinaturas carregam contribuições de mais de um objeto na cena.

### 3.4.3 Clustering

As duas seções anteriores tiveram por objetivo explicitar como podemos *identificar* formas em uma imagem e como podemos *extrair/medir* atributos particulares destas for-

<sup>10</sup>Existem outros descritores específicos para obtenção de *assinaturas* de uma forma. Para maior aprofundamento, vide Pedrini e Schwartz [62].

mas. O passo seguinte consiste em entendermos como as formas constantes em um imagem podem ser *agrupadas*.

O primeiro passo para se estabelecer um processo de clustering sobre um conjunto de dados consiste em se definir quais *atributos* da forma serão comparados. Por exemplo, é preciso definir se as formas serão agrupadas por semelhança de área, por proximidade de seus centros de massa etc. Chamamos de *vetor de características* a grandeza  $x = (x_1, x_2, \dots, x_k)$  composta por  $k$  medidas sobre os objetos de interesse. É importante notar que cada amostra a ser avaliada terá um vetor de características associado.

A representação espacial das características de cada amostra se dá pela formação do *espaço de características*. Se cada vetor de características possuir  $k$  coordenadas, o espaço gerado será  $k$ -dimensional. Um caso simples e que permite visualização imediata é o caso 2-dimensional, i.e., em que cada vetor de características é da forma  $x = (x_1, x_2)$ . A Figura 3.44 ilustra este conceito. Deve-se ter em mente que os vetores de características podem nada ter a ver com a representação física da imagem. Nesse sentido, o espaço de características não guarda informação da configuração espacial das formas nas imagens, exceto quando esse tipo de informação é usado como característica.

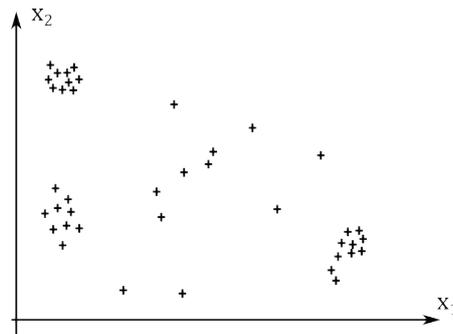


Figura 3.44: Espaço de característica bidimensional. Esse tipo de representação permite a análise que leva os elementos analisados a serem agrupados. Nota-se que há regiões mais compactadas, sugerindo um agrupamento inicial.

É evidente que o processo de clustering ocorre via medidas de semelhança. Portanto, é necessário que se defina, *a priori*, qual métrica será usada. Dentre as diversas opções, a distância Euclidiana surge como opção mais natural. Dados dois vetores de características,  $x_\alpha = (x_{\alpha,1}, \dots, x_{\alpha,k})$  e  $x_\beta = (x_{\beta,1}, \dots, x_{\beta,k})$ , a distância Euclidiana entre eles é definida por

$$d(x_\alpha, x_\beta) = \sqrt{\sum_{l=1}^k (x_{\alpha,l} - x_{\beta,l})^2} = \|x_\alpha - x_\beta\|_2. \quad (3.33)$$

### Formas de agrupamento

Estabelecido o espaço de características, é necessário definir a forma – ou classe de algoritmo – a ser usada para proporcionar o agrupamento dos dados. Duas abordagens básicas para clustering são os algoritmos hierárquicos e os particionais [42]. Neste trabalho apenas algoritmos hierárquicos foram usados, de forma que será dada ênfase nesta categoria.



1. Inicialize a matriz  $\mathcal{D}_{N \times N}$ ;
2. Encontre o par de clusters  $(a, b)$  tal que  $d(a, b) = \min\{d(i, j)\}$  sobre todas as entradas  $ij$  de  $\mathcal{D}$ ;
3. Agrupe o par de clusters  $(a, b)$ :  $(a, b) \rightarrow \{a, b\}$ ;
4. Atualize a matriz  $\mathcal{D}$ : remova a linha  $a$  e a coluna  $b$  e calcule  $d(\{a, b\}, j)$  segundo a norma adotada;
5. Se todos os objetos estiverem agrupados no mesmo cluster, pare. Caso contrário, vá para o passo 2.

Figura 3.46: Algoritmo básico para agrupamento hierárquico via atualização da matriz de distâncias.

Uma vez definido o espaço de características, é mandatório definir o critério de escolha do *representante de grupo*, que é o dado quantitativo que definirá a atualização da matriz de distâncias (passo 4 do algoritmo). Para algoritmos hierárquicos, há três abordagens básicas para determinação do representante de grupo [62, 42, 43], a serem descritas a seguir.

**Ligação simples:** A distância entre dois grupos,  $\mathcal{A}$  e  $\mathcal{B}$  é tomada como sendo a menor distância entre os pares formados por elementos dos dois grupos, i.e.,

$$d(\mathcal{A}, \mathcal{B}) = \min_{i, j \in \mathcal{A} \times \mathcal{B}} \{d(i, j)\}, \quad (3.35)$$

em que  $\mathcal{A} \times \mathcal{B}$  representa o produto cartesiano entre os conjuntos  $\mathcal{A}$  e  $\mathcal{B}$ .

**Ligação completa:** A distância entre dois grupos,  $\mathcal{A}$  e  $\mathcal{B}$ , é tomada como sendo a maior distância entre os pares formados por elementos dos dois grupos, i.e.,

$$d(\mathcal{A}, \mathcal{B}) = \max_{i, j \in \mathcal{A} \times \mathcal{B}} \{d(i, j)\}, \quad (3.36)$$

**Centróide:** A distância entre dois grupos,  $\mathcal{A}$  e  $\mathcal{B}$  é definida como sendo a distância Euclidiana entre seus centros de massa, tal como apresentado na Equação (3.33) :

$$d(\mathcal{A}, \mathcal{B}) = \|\mathbf{C}(\mathcal{A}) - \mathbf{C}(\mathcal{B})\|_2, \quad (3.37)$$

em que

$$\mathbf{C}(\mathcal{X}) = \frac{1}{n} \cdot \left( \sum_{i=1}^n x_{i,1}, \sum_{i=1}^n x_{i,2}, \dots, \sum_{i=1}^n x_{i,k} \right) \quad (3.38)$$

é o centro de massa de um dado cluster  $\mathcal{X}$  contendo  $n$  elementos.

Para fins de exemplificação, submetamos a imagem retratada na Figura 3.47 a um processo de clustering hierárquico aglomerativo. Agruparemos as formas presentes na imagem usando como vetores de características as coordenadas dos centros de massa dos objetos na cena.

Na Tabela 3.2 são mostradas algumas grandezas de interesse para as formas constantes na imagem apresentada na Figura 3.47. Os vetores de característica usados correspondem aos pares  $(x_{CM}, y_{CM})$ .

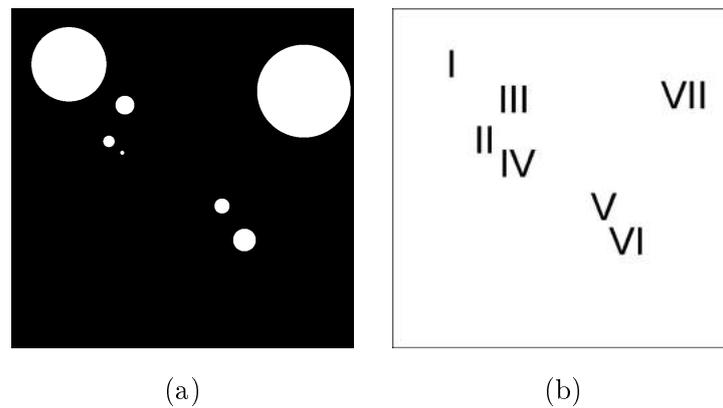


Figura 3.47: Imagem a ser analisada no processo de clustering: em (a) constam 7 formas circulares com raios distintos e em posições tais que não há superposição entre elas. Em (b) são mostrados os rótulos de cada forma. A imagem possui 900x900 pixels e resolução de 20px/mm.

ID forma	Área (px)	$x_{CM}$ (px)	$y_{CM}$ (px)
I	30405	152	149
II	717	257	353
III	1941	299	257
IV	89	292	383
V	1237	554	524
VI	2801	613	614
VII	47409	769	220

Tabela 3.2: Atributos de área e posição dos centros de massa das formas constantes na imagem apresentada na Figura 3.47.

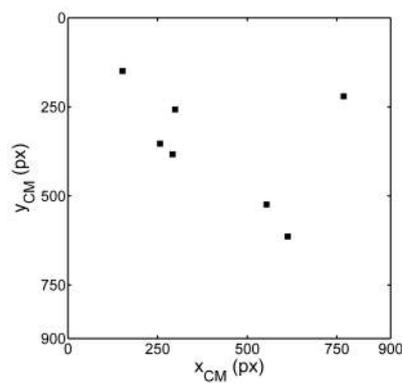


Figura 3.48: Espaço de características da imagem mostrada na Figura 3.47, cujos vetores de característica são as coordenadas dos centros de massa das formas presentes na imagem. Os valores destes vetores podem ser consultados na Tabela 3.2.

Com base na definição de matriz de distâncias e dos dados constantes nos vetores de



O processo de clustering pode ser melhor visualizado por meio de um *dendrograma*, que consiste em uma representação gráfica em forma de árvore. Para este exemplo em particular, o dendrograma tem a forma apresentada na Figura 3.49.

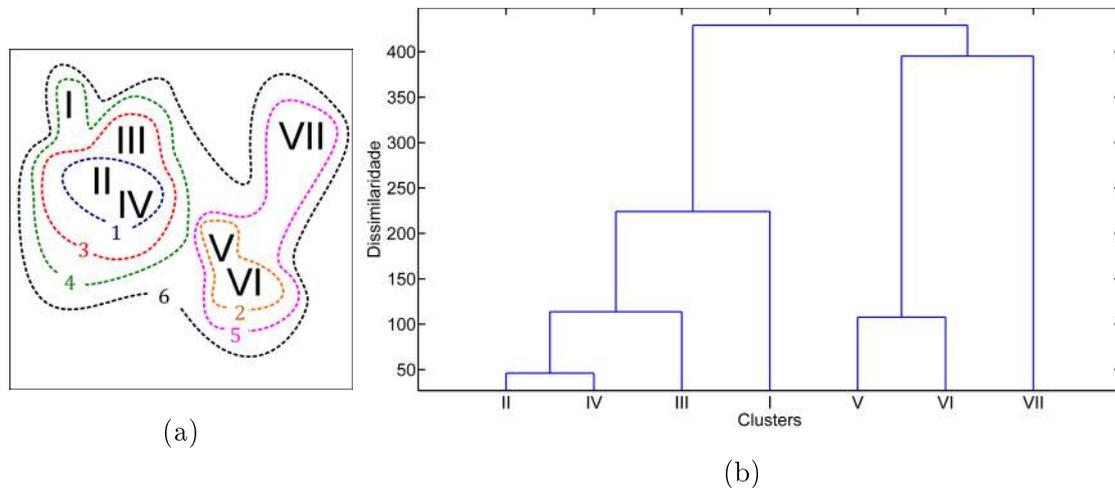


Figura 3.49: Resultado do processo de clustering sobre as formas da imagem mostrada na Figura 3.47, cujos vetores de características são compostos pelos centros de massa das formas em questão. Em (a) é mostrada uma ilustração sobre os rótulos de cada forma, indicando os sucessivos agrupamentos sofridos pelos objetos. Os números indicam a etapa em que ocorreu o agrupamento. Em (b) é mostrado o dendrograma do processo de clustering em questão. O eixo vertical representa a dissimilaridade (não-semelhança) entre os clusters. Nota-se que quão mais agrupada se torna a coleção de dados, maior a dissimilaridade entre seus constituintes. Isto indica a necessidade do estabelecimento de um critério de parada para evitar que objetos discrepantes entre si sejam agrupados.

Conforme pode ser observado pelas discussões anteriores e pelo exemplo estudado, fica evidente que é necessário o estabelecimento de um *critério de parada para o algoritmo hierárquico aglomerativo*. Caso contrário, amostras que não guardam relação significativa serão agrupadas. Uma possibilidade intuitiva consiste em fazer com que o processo de clustering ceda quando a dissimilaridade for maior que um certo limiar. Ou seja, caso a dissimilaridade – ou distância – entre dois objetos seja maior ou igual que certo valor, estes objetos não são mais agrupados.

### 3.5 Pós-processamento: morfologia matemática

Para finalizar este capítulo, apresentaremos algumas informações básicas sobre técnicas de processamento derivadas da morfologia matemática. Podemos dizer, de maneira bastante imprecisa, que a morfologia matemática consiste em um conjunto de técnicas que proporcionam um processamento voltado à forma das estruturas analisadas. Em particular, daremos ênfase ao caráter de pós-processamento das técnicas descritas. Isto é, discutiremos algumas técnicas básicas apropriadas para extração de pontos, dilatação de bordas, preenchimento de regiões não-simplesmente conexas etc. Por simplicidade, consideraremos as operações morfológicas como ocorrendo sobre imagens binárias. Ao final da seção, será apresentada a generalização dos conceitos apresentados para imagens monocromáticas.

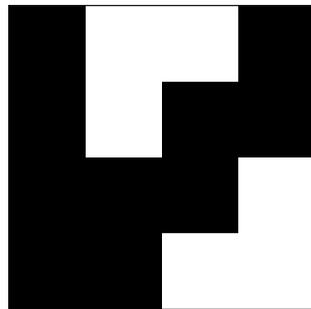
### 3.5.1 Conceitos fundamentais

Uma imagem binária  $u$  pode ser representada como um subconjunto de  $\mathbb{Z}^2$  composto pelas coordenadas em que ela assume o valor 1, isto é,

$$u = \{(x, y) \in \mathbb{Z}^2 | u(x, y) = 1\}. \quad (3.40)$$

Para fins de exemplificação, são mostradas na Figura 3.50 uma imagem binária de dimensões 4x4 e sua representação matricial juntamente com os eixos coordenados de  $\mathbb{Z}^2$ . A representação para essa imagem, em termos de um conjunto seguindo a definição dada na Equação (3.40) é dada por

$$u = \{(0, 0), (0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 1), (2, 2), (3, 0), (3, 1)\}.$$



(a)

	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>
<b>0</b>	1	0	0	1
<b>1</b>	1	0	1	1
<b>2</b>	1	1	1	0
<b>3</b>	1	1	0	0

(b)

Figura 3.50: Imagem binária com dimensões 4x4 pixels (a) e sua representação matricial (b). As entradas em negrito separadas por linhas representam as coordenadas no plano  $\mathbb{Z}^2$  sobre o qual a imagem está definida.

Conforme pode ser verificado pela definição dada por (3.40) e visualizado na representação (b) da Figura 3.50, a representação da imagem depende, necessariamente, do arranjo do sistema de coordenadas. Em particular, a escolha da origem do sistema de referência tem influência nas operações básicas em morfologia.

Podemos manipular imagens binárias sem recorrer à abordagem apresentada na Seção 3.2. Sendo definidas agora como conjuntos, imagens binárias podem ser operadas segundo operações básicas derivadas deste arcabouço teórico. Sejam duas imagens binárias,  $u$  e  $v$  e  $\beta \in \mathbb{Z}$ . Derivam diretamente da teoria de conjuntos as seguintes operações:

<b>União</b>	$u \cup v = \{w   w \in u \text{ ou } w \in v\}$
<b>Interseção</b>	$u \cap v = \{w   w \in u \text{ e } w \in v\}$
<b>Translação</b>	$u + \beta = \{w + \beta   w \in u\}$
<b>Reflexão</b>	$\hat{u} = \{-w   w \in u\}$
<b>Complemento</b>	$u^C = \{w   w \notin u\}$

Tabela 3.3: Operações básicas em morfologia matemática derivadas da teoria de conjuntos.

Dadas duas imagens binárias,  $u$  e  $v$ , podemos definir uma série de operações baseadas naquelas apresentadas na Tabela 3.3. Em geral, se operamos  $u$  com base em  $v$ , chamamos esta última imagem de *elemento estruturante*. As duas operações mais básicas em

morfológica são a *dilatação* e a *erosão*. Definimos [62, 34] a dilatação e erosão de  $u$  por  $v$ , respectivamente, como

$$\begin{cases} D(u, v) = u \oplus v = \bigcup_{w \in v} (u + w); \\ E(u, v) = u \ominus v = \bigcap_{w \in v} (u - w) = \bigcap_{w \in \hat{v}} (u + w). \end{cases} \quad (3.41)$$

Em termos práticos, a dilatação  $u \oplus v$  possibilita o espessamento de bordas de  $u$ , segundo a forma do elemento estruturante  $v$ . Por outro lado, a erosão  $u \ominus v$  causa um afinamento das bordas de  $u$  segundo a forma de  $v$ . Uma ilustração destas operações pode ser visualizada na Figura 3.51. Neste exemplo, o elemento estruturante foi tomado como sendo o produto cartesiano  $[-4, 4] \times [-4, 4]$ , ou seja,  $v = \{(i, j) \in \mathbb{Z}^2 | i, j \in [-4, 4]\}$ .

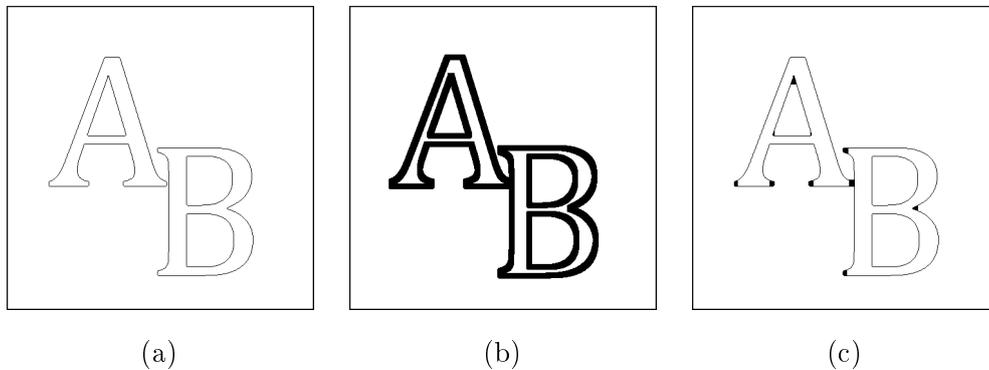


Figura 3.51: Ilustração das operações de dilatação e erosão. Em (a) é mostrada uma imagem sintética cujas bordas foram detectadas e segmentadas. Em (b) é mostrado o resultado da dilatação da imagem binária (a) segundo o elemento estruturante  $v = [-4, 4] \times [-4, 4]$ . Em (c) é mostrada a erosão da imagem dilatada (b) segundo o mesmo elemento estruturante  $v$ . Nota-se claramente o espessamento de bordas quando da aplicação da dilatação, assim como o refino dos contornos quando da aplicação da erosão. Para fins de visualização, as imagens mostradas correspondem aos negativos dos resultados das operações.

As operações de dilatação e erosão gozam de diversas propriedades algébricas, tais como associatividade, comutatividade (não-comutatividade e não-associatividade no caso da erosão), dualidade etc. Para maior aprofundamento sugere-se o livro de Pedrini e Schwartz [62] ou o trabalho de Haralick *et al.* [34].

### 3.5.2 Operações derivadas

#### Abertura e fechamento

Uma classe grande de operações morfológicas pode ser derivada das diferentes combinações entre dilatação e erosão, tal como definidas na Equação (3.41). Nos ateremos a expor três operações derivadas – *abertura*, *fechamento* e *preenchimento de regiões* – pois correspondem às técnicas usadas neste trabalho.

Dados uma imagem  $u$  e um elemento estruturante  $v$ , definimos a abertura e fechamento, respectivamente, por

$$\begin{cases} u \circ v = (u \ominus v) \oplus v \\ u \bullet v = (u \oplus v) \ominus v \end{cases} \quad (3.42)$$

Tais como definidas na Equação (3.42), as operações de abertura e fechamento podem ser entendidas como removedores de ruídos, dependendo da convenção adotada (0 para preto e 1 para branco ou 1 para preto e 0 para branco). A remoção de ruídos pode ser entendida analisando-se a operação de abertura: uma vez erodida por  $v$ , a imagem  $u$  perderá pequenos detalhes que não surgirão quando se dilatar o resultado. A interpretação para o fechamento é análoga. O que permite analisar as operações de abertura e fechamento como quase-equivalentes é a sua relação de dualidade, dada por [62, 34]

$$\begin{cases} (u \circ v)^C = u^C \bullet \hat{v}; \\ (u \bullet v)^C = u^C \circ \hat{v}. \end{cases} \quad (3.43)$$

Um exemplo de aplicação de abertura e fechamento com vistas à remoção de ruídos é mostrado na Figura 3.52

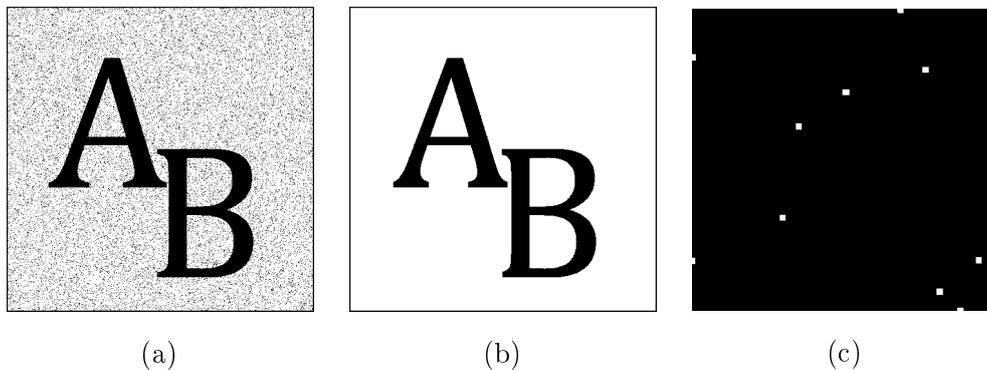


Figura 3.52: Ilustração do conceito de abertura e fechamento morfológicos. Em (a) é mostrada uma imagem binária corrompida por ruído. Em (b) é mostrado o resultado da abertura de (a) segundo o elemento estruturante  $v = [-4, 4] \times [-4, 4]$ , e em (c), o resultado do fechamento de (a) segundo o mesmo elemento estruturante. Nota-se a efetividade da abertura quanto à remoção de ruídos. Nesse exemplo, o fechamento privilegiou a junção dos pontos de ruído, mas operaria igualmente bem na remoção de ruído caso a imagem original fosse invertida (dualidade). Para melhor visualização, as imagens mostradas são apresentadas como os negativos dos resultados.

#### Preenchimento de regiões

A morfologia matemática permite que regiões não-simplesmente conexas – isto é, com “buracos” – sejam preenchidas através de um processo iterativo. Uma realização possível

[62] é tomar a imagem  $u$  como sendo o resultado do processo iterativo

$$u_{t+1} = (u_t \oplus v) \cap u_0^C, \quad (3.44)$$

em que  $u_0$  é a imagem inicial.

A interpretação do processo iterativo descrito em (3.44) é bastante simples: dilata-se a imagem e obtêm-se a interseção com o complemento da imagem inicial para que não ocorra deformação das bordas. O processo é encerrado quando  $u_{t+1} = u_t$ . Na Figura 3.53 pode ser visualizada a aplicação do preenchimento morfológico sobre uma imagem sintética.

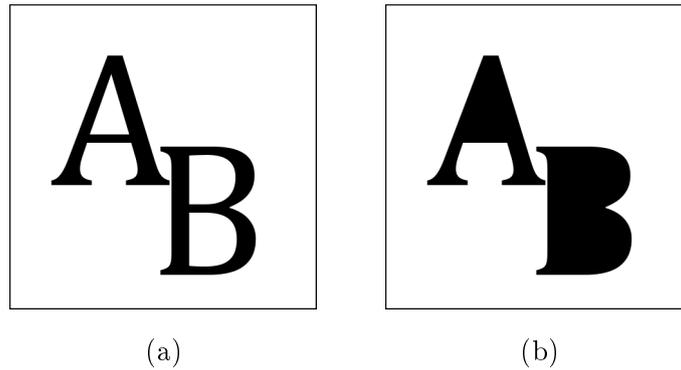


Figura 3.53: Ilustração do preenchimento morfológico. Em (a) é mostrada uma imagem sintética de dois caracteres com regiões a serem preenchidas. Em (b) é mostrado o resultado do preenchimento, ilustrando que o processo iterativo descrito na Equação (3.44) opera segundo o esperado, fechando as regiões. Para fins de visualização, as imagens mostradas são tomadas como sendo o negativo das originais.

### 3.5.3 Extensão a imagens monocromáticas

As operações de dilatação e erosão podem ser estendidas para imagens monocromáticas. Definimos [62, 34] a dilatação e erosão, respectivamente, de  $u$  por  $v$  por

$$\begin{cases} (u \oplus v)(x, y) = \max_{(m,n) \in v} \{u(x-m, y-n) + v(m, n)\} \\ (u \ominus v)(x, y) = \min_{(m,n) \in v} \{u(x-m, y-n) + v(m, n)\} \end{cases} \quad (3.45)$$

Uma vez redefinidas as operações básicas de dilatação e erosão, as operações de abertura, fechamento e preenchimento são redefinidas automaticamente, pois seguem a mesma definição que suas versões para imagens binárias. Uma ilustração da aplicação de operações morfológicas sobre imagens monocromáticas pode ser vista na Figura 3.54.

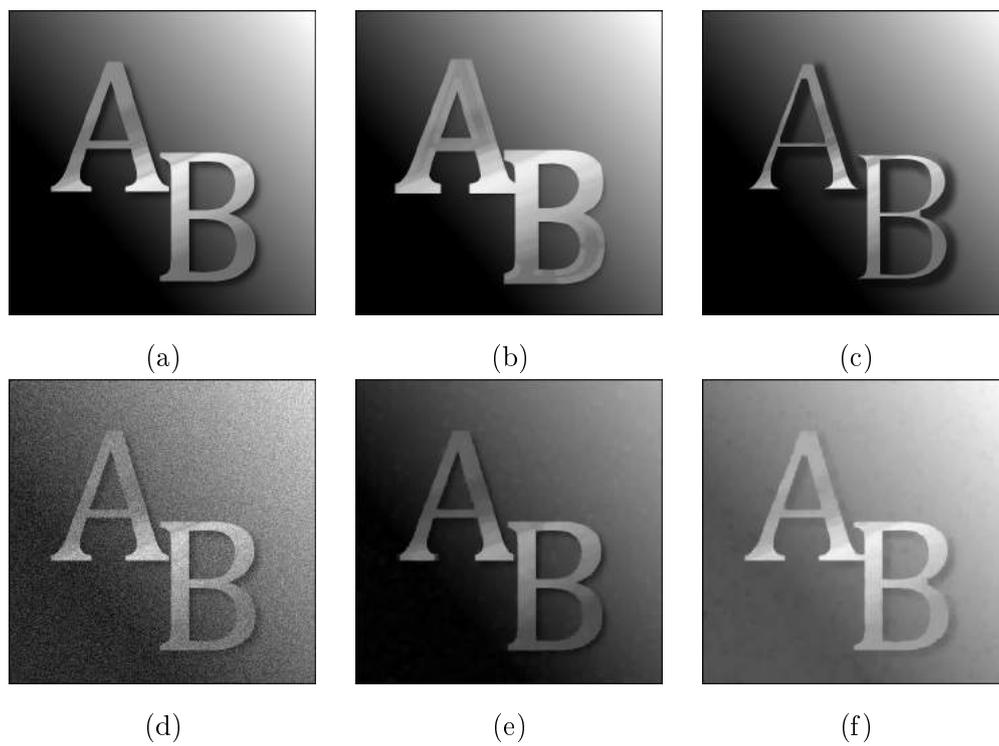


Figura 3.54: Ilustração da extensão da morfologia matemática a imagens monocromáticas. Em (a) é mostrada uma imagem sintética a ser processada. Em (b) e (c) são mostradas sua dilatação e erosão, respectivamente, pelo elemento estruturante  $v = \{v(x, y) = 1 \forall (x, y) \in [-4, 4] \times [-4, 4]\}$ . Em (d) é mostrada uma versão de (a) corrompida com ruído, e em (e) e (f) são mostradas sua abertura e fechamento, respectivamente, com o mesmo elemento estruturante anterior.

## Capítulo 4

# Modelos de processamento derivados de Equações Diferenciais

### 4.1 Problemática do realce em imagens

O propósito desta Seção é evidenciar a problemática do realce em imagens, etapa esta que corresponde ao passo que antecede a segmentação (vide Seção 3.3, p. 35). Segundo a estrutura básica de processamento lá retratada, as informações destacadas na etapa de realce serão aquelas que serão submetidas à segmentação e que, portanto, constituirão a informação de alto nível a ser extraída e interpretada. Logo, a definição das características a serem evidenciadas, bem como as técnicas escolhidas para tal tarefa devem ser criteriosamente discutidas e mensuradas.

Em processamento de imagens, denominamos o realce de alguma característica particular da imagem como sendo um processo de *enhancement*. Promover o *enhancement* de certa característica em uma imagem passa necessariamente pela definição desta. Dentro desse contexto, o processo de *enhancement* se assemelha à segmentação, em que não há uma teoria que apresente a solução para qualquer requisito.

A problemática do *enhancement* pode ser enunciada da seguinte maneira:

*Definidas as características de interesse a serem extraídas via processamento, como as evidenciar eficientemente e comprometer minimamente o restante da imagem?*

Em seu artigo clássico, Prewitt ([69], p.101) apresenta este panorama de forma bastante clara:

In general, it is desirable to deemphasize or discard irrelevant material and *at the same time* to emphasize or clarify features and objects of interest. Moreover, it is desirable to do this without having to identify or recognize specific objects, insofar as possible. Deemphasis is usually achieved by using smoothing or integrating operators, but these tend also to blur detail. Emphasis is usually accomplished by sharpening or differentiating operators, or by feature enhancers or contrast enhancers, but these also tend to accentuate noise. However, smoothing *without* obliteration of the relevant and sharpening *without* amplification of the irrelevant are the desired ends. Thus a judicious selection of techniques is required to balance these counter-tendencies and give optimal enhancement.

As operações citadas acima referem-se aos operadores lineares, discutidos na Seção 3.2. Já foi discutido que filtros passa-baixa são bons removedores de ruídos, mas sacrificam bordas. Por sua vez, filtros passa-alta são bons detectores de bordas, mas acabam evidenciando ruídos diversos. Devido a essa limitação dos filtros lineares, processamentos baseados em processos não-lineares foram concebidos, de forma que se conseguisse atingir um enhancement efetivamente seletivo. Em particular, modelos baseados em equações diferenciais parciais (EDPs) permitem atingir minimamente este objetivo.

Este capítulo tem por objetivo apresentar os conceitos fundamentais sobre a filtragem de imagens via EDPs não-lineares. Em particular, será feita uma discussão sobre a relação intrínseca entre descrição multiescala e difusão linear de um sinal. Esta relação e suas limitações nos levarão à formulação de um processo de difusão não-linear, em que a equação de Perona-Malik será explicitada. Uma vez explicitado o modelo teórico, questões acerca de seus parâmetros e sua implementação computacional também serão tratadas.

## 4.2 Descrição multiescala, filtragem Gaussiana e equação do calor

Nesta Seção exemplificaremos a problemática do enhancement, tal como apresentada na Seção anterior. Essa exemplificação nos guiará à interpretação em termos de difusão linear da filtragem por um passa-baixas Gaussiano e permitirá analisar o processo de filtragem de uma imagem segundo um processo de evolução.

Em princípio, consideremos um sinal unidimensional em domínio espacial contínuo, isto é,  $u = u(x)$ . Este sinal será – mesmo que minimamente – alterado em função de processos de aquisição e transmissão [33]. De um modo geral, podem ocorrer deterioração das bordas do sinal (*blurring*) em associação a um acréscimo de ruído [5]<sup>1</sup>. Para fins de exemplificação, consideremos um trem de pulsos sintético corrompido com ruído Gaussiano aditivo de média nula, tal como ilustrado na Figura 4.1. A relação sinal-ruído –  $\text{SNR}^2$  – é de 3,1 dB. Operaremos sobre o sinal corrompido visando demonstrar as dificuldades já apresentadas quanto ao enhancement.

Devemos entender o enhancement do sinal ruidoso retratado na Figura 4.1 como um processo de evidenciação das bordas originais e de redução do ruído aditivo. Seguiremos duas abordagens básicas típicas dos processos lineares: inicialmente, podemos diferenciar o sinal, com vistas a detectar suas bordas. Alternativamente, podemos usar um filtro passa-baixas Gaussiano sobre o sinal, devido a sua eficiência em remover ruídos. Os resultados das filtrações por diferenciador e passa-baixas Gaussiano –  $u_d(n)$  e  $u_g(n)$ , respectivamente – podem ser visualizados na Figura 4.2. Nota-se que, para  $u_d(n)$ , houve uma completa degradação do sinal, com a evidenciação de ruído e perda completa das bordas. Por outro lado, o sinal  $u_g(n)$  apresenta uma forma mais próxima do sinal original, evidenciando a eficiência do filtro Gaussiano quanto a remoção de ruídos. Contudo, essa redução de ruídos vem às expensas da desconfiguração parcial das bordas. Nesse exemplo, o desvio padrão usado foi  $\sigma = 10$ . A desconfiguração das bordas será tão mais intensa quanto maior for o valor de  $\sigma$ . A influência da magnitude de  $\sigma$  sobre a filtragem de um sinal é ilustrada na Figura 4.3.

Embora a filtragem via Gaussiana apresente o inconveniente de descaracterizar com-

---

<sup>1</sup>A deterioração de bordas via filtragem linear e acréscimo de ruído Gaussiano aditivo constitui um modelo, não correspondendo necessariamente à degradação real sofrida pelo sinal. Contudo, este modelo simples permite uma abordagem avançada para a *reconstrução* do sinal via métodos variacionais. Para

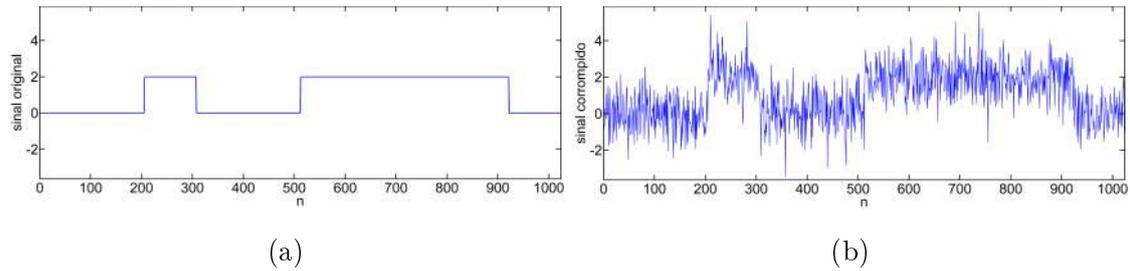


Figura 4.1: Trem de pulsos sintético  $u(n)$  (a) e sua versão corrompida (b), cuja SNR é igual a 3,1dB.

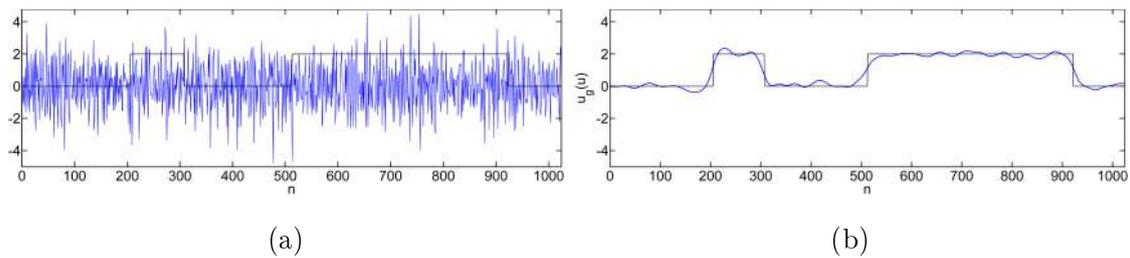


Figura 4.2: Resultado da filtragem linear sobre o sinal ruidoso representado na Figura 4.1: em (a) é mostrado o resultado da filtragem via diferenciador –  $u_d(n)$  – e em (b) o resultado da filtragem via Gaussiana –  $u_g(n)$  –, cujo desvio padrão,  $\sigma$ , é igual a 10. Nota-se que a filtragem via diferenciador do sinal ruidoso não opera de maneira adequada, pois evidencia ruídos diversos, desconfigurando totalmente a forma do sinal em questão. No segundo caso, uma redução significativa dos ruídos é realizada, embora as bordas sofram deterioração, que será tão maior quanto maior for o desvio-padrão  $\sigma$ . Tanto em (a) quanto em (b) o sinal original (sem ruídos) é mostrado em pontilhado para permitir comparação visual.

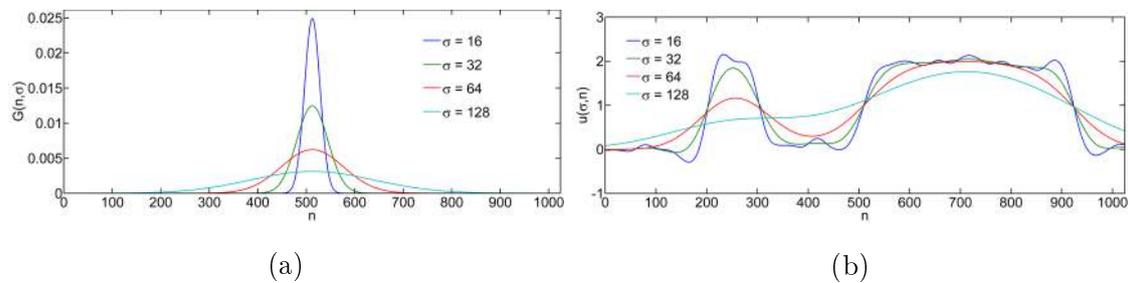


Figura 4.3: Diferentes graus de deterioração do sinal em função da filtragem por Gaussiana. Em (a) são mostradas as representações para filtros Gaussianos com diferentes desvios-padrão,  $\sigma$ . Em (b) é mostrado o resultado da filtragem do sinal ruidoso retratado na Figura 4.1 por estes filtros. Nota-se claramente que, conforme  $\sigma$  aumenta, o sinal perde suas características originais, rumando a uma forma completamente descaracterizada.

pletamente o sinal original conforme o valor de seu desvio padrão cresça, ela permite o estabelecimento de uma forma distinta de se descrever o sinal em questão. Essa é a chamada descrição multiescala e leva naturalmente ao conceito de difusão de um sinal. A seguir, será feita uma discussão sobre os conceitos fundamentais envolvidos nesse tipo de descrição.

No início da Seção 3.1 foi dito que uma imagem consiste em um sinal representativo

aprofundamento, recomenda-se o livro de Aubert e Kornprobst [5].

<sup>2</sup> Do inglês, *signal-to-noise ratio*.

de alguma estrutura física, artística ou conceitual, mas nada foi dito sobre a complexidade da estrutura retratada, pois isto varia segundo a própria estrutura. Contudo, há algumas considerações pertinentes a qualquer imagem, relativas a sua *hierarquia*, que devem ser analisadas. Entenda-se por hierarquia de uma imagem a organização dos detalhes representados nela própria. Usando o exemplo dado por Koenderink [46] e Perona e Malik [63]: ao observarmos uma árvore, podemos nos ater aos detalhes finos, como as nervuras das folhas. Nesta *escala*, a copa da árvore não existe, tampouco a floresta. Em um outro nível de descrição, a folha complexa é substituída por uma região simples. Se a escala de análise for expandida, o “objeto folha” perderá significado, pois observaríamos a copa da árvore como uma única estrutura. Uma ilustração desse tipo de análise segundo diferentes escalas pode ser vista na Figura 4.4.



Figura 4.4: Ilustração do conceito de análise de uma imagem segundo diferentes escalas: seguindo da esquerda para a direita e de cima para baixo, vemos que a análise do objeto retratado por uma imagem obedece uma hierarquia natural. Isto é, em cada escala, detalhes particulares são evidenciados e outras estruturas simplesmente não existem: se observarmos as nervuras da folha, não conseguiremos enxergar o caule, por exemplo.

Seguindo a motivação de análise de imagens em diversas escalas distintas, uma abordagem alternativa para descrição qualitativa de imagens foi concebida, sendo denominada descrição *multiescala* [93, 46, 63]. Essa descrição consiste em se obter uma família de sinais derivada do sinal original, mas com um nível de detalhamento mais pobre. O sinal será simplificado cada vez mais conforme uma certa variável de escala,  $t$  por exemplo, aumente. Os sinais que compõem esta família são obtidos iterativamente por meio de alguma filtragem que simplifique a descrição do sinal cada vez mais. Ou seja, o sinal inicial  $u(x, t = 0)$  é iterativamente filtrado, levando a descrições cada vez mais simplificadas tal como mostrado na Figura 4.5.

A descrição multiescala pode, sem maiores impedimentos, ser estendida a sinais multidimensionais, ou seja, com  $u = u(\vec{r}, t)$ , em que fica explícita a dependência da forma do sinal com a variável de escala,  $t$ . No que segue, toda a discussão será erguida sobre sinais multidimensionais, salvo observação em contrário.

Em princípio, qualquer filtro passa-baixas poderia ser usado para se obter as representações sucessivas em  $t$  de  $u$ . Contudo, se introduzirmos duas hipóteses ao processo de descrição do sinal, poderemos traçar um importante paralelo entre a descrição multiescala e a equação do calor. As hipóteses são as seguintes [46]:

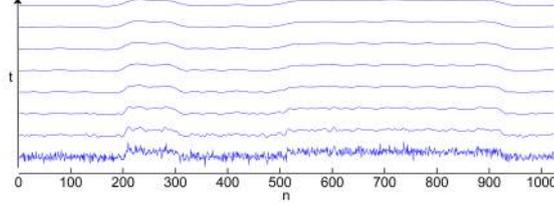


Figura 4.5: Descrição multiescala do sinal ruidoso apresentado na Figura 4.1: conforme a variável de escala  $t$  aumenta, o nível de detalhamento diminui.

1. **Causalidade:** quaisquer características de  $u(\vec{r}, t_1)$  (bordas, inflexões, descontinuidades etc.) devem ter sua origem, necessariamente, em algum  $t_0 < t_1$ . Intuitivamente, isto significa que qualquer detalhe notado em certa resolução deve ter uma *causa* (não necessariamente única) anterior. Esta hipótese impede a geração de ruídos espúrios.
2. **Homogeneidade e isotropia:** a degradação sofrida pelo sinal deve ser invariante a deslocamentos sobre o plano do sinal. Intuitivamente, essa hipótese corresponde à exigência de que o filtro opere sobre o sinal sem privilegiar alguma direção em particular.

Koenderink [46] mostra que caso estas duas hipóteses sejam estabelecidas, a descrição multiescala de  $u(\vec{r}, t)$  é análoga a uma evolução regida pela *equação do calor*<sup>3</sup>:

$$\begin{cases} \partial_t u(\vec{r}, t) = \nabla^2 u(\vec{r}, t) \\ u(\vec{r}, 0) = u_{\text{inicial}}, \end{cases} \quad (4.1)$$

cujas solução é dada pelo produto de convolução

$$u(\vec{r}, t) = u(\vec{r}, 0) * G(\vec{r}, \sigma^2(t)). \quad (4.2)$$

Na Equação (4.2),  $G(\vec{r}, \sigma^2(t))$  corresponde a uma Gaussiana cuja variância depende de  $t$  na forma  $\sigma^2(t) = 2t$ . Para  $\vec{r} = (x, y)$ , temos:

$$G(\vec{r}, \sigma^2(t)) = \frac{1}{4\pi t} \exp\left(-\frac{\|\vec{r}\|^2}{4t}\right). \quad (4.3)$$

Uma demonstração para o caso unidimensional do resultado expresso em (4.2) e (4.3) é apresentada no Apêndice A.

O que as Equações (4.2) e (4.3) nos dizem – dentro do contexto de processamento de sinais – é que uma descrição multiescala que atenda as hipóteses de causalidade e isotropia é completamente equivalente a submeter o sinal em questão a um processo de *difusão linear*, cuja solução  $u(\vec{r}, t)$  é uma versão suavizada do sinal inicial por uma Gaussiana de variância  $\sigma^2(t) = 2t$ . Nesse sentido, a variável de escala  $t$  pode ser tomada como sendo o tempo ao longo do qual o sistema é posto a difundir.

<sup>3</sup>Rigorosamente falando, a Equação (4.1) corresponde a um problema de valor inicial para a equação de *difusão*. Neste trabalho, os termos equação de difusão e equação do calor serão usados para descrever o mesmo fenômeno, salvo observação em contrário.

### 4.3 Equação de Perona-Malik

Em uma primeira análise, poderíamos nos questionar sobre a utilidade em se filtrar uma imagem segundo um processo de difusão. Efetivamente, a difusão de  $u$  segundo a Equação (4.1) corresponde a uma filtragem por Gaussiana, que – conforme visto – apresenta eficiência limitada na restauração de um sinal, dado que deteriora suas bordas. Contudo, isso abre portas para que possamos analisar variações sobre a *maneira* que o sinal é posto a difundir. Uma primeira abordagem rumo a essas variações seria considerar a *equação reversa do calor*:

$$\begin{cases} \partial_t u(\vec{r}, t) &= -\nabla^2 u(\vec{r}, t) \\ u(\vec{r}, 0) &= u_{\text{inicial}}. \end{cases} \quad (4.4)$$

A justificativa para se cogitar o uso da EDP apresentada na Equação (4.4) é bastante simples: se a equação do calor suaviza o sinal de maneira cada vez mais intensa conforme  $t$  aumenta, espera-se que a equação em tempo reverso possa – ao menos minimamente – desfazer essa suavização. De fato, essa restauração ocorre, mas durante um intervalo de tempo bastante curto, até que a solução divirja completamente. Isso se deve ao fato de que a equação reversa do calor corresponde a um problema mal-posto <sup>4</sup>. A Figura 4.6 ilustra o que foi dito.

Com esse exemplo, fica claro que a equação reversa do calor permite, mesmo que em um intervalo estreito de  $t$ , um enhancement efetivo do sinal. Efetivamente, o *coeficiente de difusão* (ou *difusibilidade*) desempenha papel fundamental na eficiência da técnica. Esse ponto foi abordado por Perona e Malik [63] (apud PERONA, MALIK, 1987), e será discutido a seguir.

Inicialmente, deve-se notar que as Equações (4.1) e (4.4) são casos particulares da equação

$$\partial_t u(\vec{r}, t) = \vec{\nabla} \cdot [c(\vec{r}, t) \vec{\nabla} u(\vec{r}, t)], \quad (4.5)$$

em que  $c(\vec{r}, t)$  é a função difusibilidade. A difusibilidade  $c(\vec{r}, t)$  é igual a 1 para a Equação (4.1) e igual a -1 na Equação (4.4). Perona e Malik observaram que a difusibilidade não necessita ser constante: ela pode variar no espaço e no tempo de forma a se atender algum requisito particular. Especificamente, a função  $c(\vec{r}, t)$  deve ser tal que o sinal evolua de acordo com os seguintes requisitos:

1. Nas regiões entre-bordas, a evolução deve levar à máxima homogeneização possível;
2. Nas regiões de borda, a evolução deve ocorrer de modo a ressaltar, ou pelo menos conservar estas estruturas.

Uma função apropriada para estes requisitos é o módulo do gradiente do sinal processado,  $\|\vec{\nabla} u\|$ , devido ao fato de ele poder ser usado como detector de bordas, tal como tratado na Seção 3.2. Logo, a função difusibilidade deve ser da forma  $c(\vec{r}, t) = f(\|\vec{\nabla} u\|)$ . Com isso, chega-se à chamada **Equação de Perona-Malik** [63, 67]:

$$\partial_t u = \vec{\nabla} \cdot \left( f(\|\vec{\nabla} u\|) \vec{\nabla} u \right). \quad (4.6)$$

<sup>4</sup>Um problema mal-posto é aquele em que existência, unicidade ou dependência contínua nos dados iniciais ou de contorno não estão garantidos [5, 39, 25]

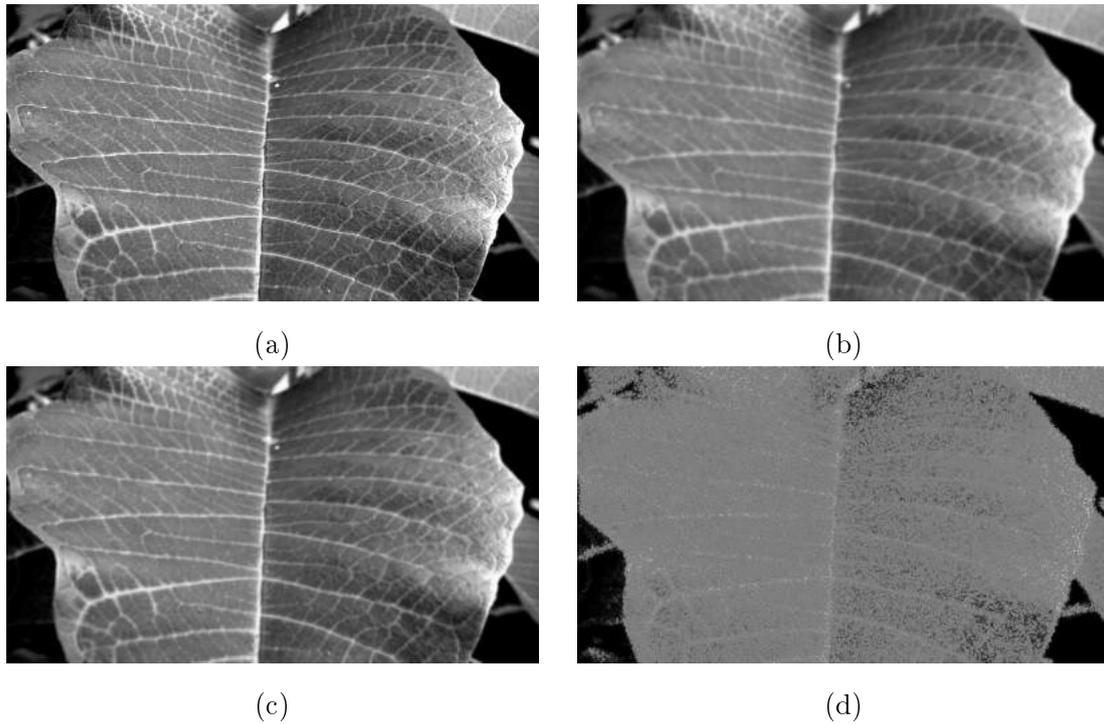


Figura 4.6: Ilustração da filtragem de uma imagem segundo a equação reversa do calor: em (a) é mostrada a imagem a ser filtrada. Em (b) é mostrada sua versão borrada devido a uma evolução pela Equação (4.1). Em (c) é mostrada a restauração proporcionada pela evolução segundo a Equação (4.4) após 60 iterações e em (d) é mostrado o resultado após 70 iterações. Nota-se que para um número pequeno de iterações, a filtragem em tempo reverso opera como desejado, evidenciando as bordas outrora degradadas. Contudo, para um número não muito maior de iterações, a solução diverge, evidenciando o caráter instável da solução da equação reversa do calor.

De acordo com os requisitos discutidos, duas informações básicas que qualificam a função  $f(\|\vec{\nabla}u\|)$  são naturalmente evidenciadas:

1.  $f(0) = 1$ ;
2.  $\|\vec{\nabla}u\| \rightarrow +\infty \Rightarrow f(\|\vec{\nabla}u\|) \rightarrow 0$ .

Ou seja, em pontos de maior uniformidade, a difusão ocorre visando a suavização da região, ao passo que em pontos cujo  $\|\vec{\nabla}u\|$  é acentuado, a difusão é enfraquecida. Originalmente, foram propostas duas funções difusibilidade, não negativas e monotonicamente decrescentes [63]:

$$f_1(\|\vec{\nabla}u\|) = \exp\left(-\left(\frac{\|\vec{\nabla}u\|}{K}\right)^2\right) \quad e \quad (4.7)$$

$$f_2(\|\vec{\nabla}u\|) = \frac{1}{1 + \left(\frac{\|\vec{\nabla}u\|}{K}\right)^2}, \quad (4.8)$$

em que  $K$  está associado a um limiar de separação entre difusão direta ou reversa. Para visualizar essa característica, fazemos a seguinte análise:

Consideremos o caso unidimensional da Equação (4.6):

$$\partial_t u(x, t) = \partial_x (F(\partial_x u(x, t))),$$

em que  $F(\partial_x u) \equiv f(\|\vec{\nabla}u\|)\partial_x u$  é chamada *função fluxo* [63, 28]. Desenvolvendo a derivada espacial sobre a função fluxo, obtemos

$$\partial_t u = F'(\partial_x u)\partial_{xx}u. \quad (4.9)$$

A Equação (4.9) nada mais é que a equação do calor, cujo coeficiente é dado pela derivada da função fluxo,  $F'(\partial_x u)$ . É uma simples questão de cálculo mostrar que

$$\begin{cases} F'_1(\partial_x u) > 0 & \text{se } |\partial_x u| \leq \frac{K}{\sqrt{2}}, \\ F'_1(\partial_x u) < 0 & \text{caso contrário} \end{cases} \quad \text{e} \quad \begin{cases} F'_2(\partial_x u) > 0 & \text{se } |\partial_x u| \leq K, \\ F'_2(\partial_x u) < 0 & \text{caso contrário.} \end{cases} \quad (4.10)$$

Na Equação (4.10),  $F'_1$  e  $F'_2$  são as derivadas das funções fluxo associadas às difusibilidades expressas nas Equações (4.7) e (4.8), respectivamente. A Figura 4.7 mostra qualitativamente esse comportamento.

Segundo a análise feita, vemos que a equação de Perona-Malik proporciona uma difusão seletiva, em que regiões com grandes valores de  $\|\vec{\nabla}u\|$  são realçadas devido à difusão reversa, ao passo que regiões de pouca variação de  $\|\vec{\nabla}u\|$  são suavizadas segundo uma difusão direta. Como o módulo do gradiente opera como um detector de bordas, temos então que o modelo de Perona-Malik apresenta as duas características desejáveis de um sistema de enhancement: realce de bordas/detalhes finos e supressão de regiões entre-bordas.

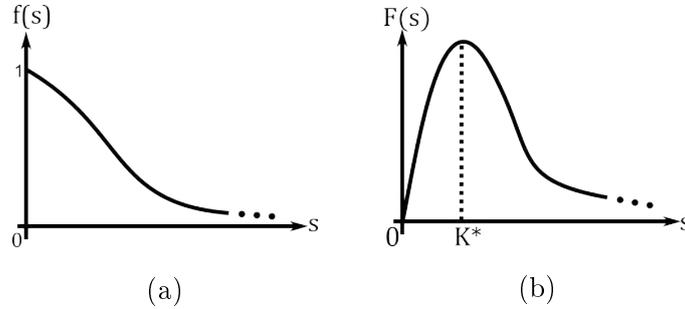


Figura 4.7: Visão qualitativa da difusibilidade proposta por Perona e Malik (a) e a respectiva função fluxo (b). A difusão do sinal muda segundo determinado valor  $K^*$  (dependente da expressão da difusibilidade), permitindo tanto suavizar uma região quanto realçá-la.

Fica claro que o parâmetro  $K$  que aparece em (4.7) e (4.8) desempenha papel fundamental no processo de difusão não-linear de Perona-Malik. Uma discussão sobre algumas formas de cálculo de  $K$  será realizada na Seção 4.3.2.

### 4.3.1 Discretização e implementação computacional

Todo o desenvolvimento anterior se deu para funções de domínio e tempo (escala) contínuos. Visando a implementação computacional da filtragem não-linear proposta por Perona e Malik, um esquema em espaço e tempo discretos deve ser elaborado. O problema de valor inicial que queremos discretizar é a equação de Perona-Malik com condições de

fronteira de Neumann:

$$\begin{cases} \partial_t u(\vec{r}, t) = \vec{\nabla} \cdot \left( f(\|\vec{\nabla} u(\vec{r}, t)\|) \vec{\nabla} u(\vec{r}, t) \right) & \text{com } u(\vec{r}, t) \in \Omega \times ]0, \infty[ \\ u(\vec{r}, 0) = u_{\text{inicial}} \\ \partial_{\vec{n}} u(\vec{r}, t) = 0 & \text{com } \partial_{\vec{n}} u(\vec{r}, t) \in \partial\Omega \times ]0, \infty[ \end{cases} \quad (4.11)$$

Na Equação (4.11),  $\Omega$  é a região que contém a imagem,  $\vec{n}$  é o vetor normal a  $\Omega$  e  $\partial\Omega$  é o bordo de  $\Omega$ . Com a condição de fronteira  $\partial_{\vec{n}} u = 0$  em  $\partial\Omega$ , o esquema discreto a ser analisado a seguir apresenta importantes propriedades [7, 91] que serão abordadas na Seção 4.3.2.

Para as considerações seguintes, usaremos o pixel como unidade fundamental da imagem, ou seja, os incrementos  $h_x$  e  $h_y$  são iguais a 1.

### Caso unidimensional

O ponto de partida é a equação escrita em termos do fluxo,  $\partial_t u = \partial_x(F(\partial_x u))$ . Supondo que o sinal  $u(x, t)$  seja discretizado espacialmente em  $N$  entradas, as derivadas espaciais podem ser aproximadas via diferenças finitas progressivas. Assim, obtemos:

$$\partial_t u_n = F(u_{n+1}^t - u_n^t) - F(u_n^t - u_{n-1}^t) ; \text{ com } n = 1, 2, \dots, N-1. \quad (4.12)$$

A discretização da variável  $t$  é imediata, e a Equação (4.12) torna-se completamente discretizada permitindo sua resolução pelo método de Euler [77]:

$$u_n^{t+1} = u_n^t + \Delta t (F(u_{n+1}^t - u_n^t) - F(u_n^t - u_{n-1}^t)). \quad (4.13)$$

Quanto ao cômputo do fluxo, devemos nos atentar à forma como a difusibilidade é obtida. Uma abordagem simples [9, 91, 88, 85] consiste em calculá-las como sendo uma média sobre a vizinhança imediata da  $n$ -ésima entrada, levando à resolução da Equação (4.13) segundo a seguinte forma explícita:

$$u_n^{t+1} = u_n^t + \Delta t \left( \frac{f_{n+1}^t + f_n^t}{2} \cdot (u_{n+1}^t - u_n^t) - \frac{f_n^t + f_{n-1}^t}{2} \cdot (u_n^t - u_{n-1}^t) \right), \quad (4.14)$$

em que  $f_n^t$  remete à  $n$ -ésima entrada do sinal difusibilidade, que deve ser calculado segundo um esquema de diferenças centradas:

$$f_n^t = f \left( \left| \frac{u_{n+1}^t - u_{n-1}^t}{2} \right| \right). \quad (4.15)$$

Uma implementação em MATLAB<sup>®</sup> da Equação (4.14) pode ser visualizada no Código 4.1. Nesta *function*, os parâmetros de entrada são a imagem a ser filtrada (**x**), o número de iterações (**numIter**), o valor de  $K$  e uma string (**strDifusi**) que define qual força de difusão será aplicada (**exp** para a força apresentada na Equação (4.7) ou **rac** para a força apresentada na Equação (4.8)). Uma ilustração da aplicação da difusão de Perona-Malik sobre o sinal ruidoso apresentado na Figura 4.1 pode ser vista na Figura 4.8.

```

function y = FILTRO_PERONA_MALIK_1D(x,numIter,K,strDifusi)
    y = x;
    dt = 0.1; % incremento temporal
    L = length(y);

    % deslocamentos a serem usados nos calculos das diferencas finitas
    DeslProgressivo = [2:L L]; DeslRegressivo = [1 1:L-1];

    for k=1:numIter
        DifProgressiva = y(DeslProgressivo)-y;
        DifRegressiva = y-y(DeslRegressivo);
        DifCentrada = 0.5*(y(DeslProgressivo)-y(DeslRegressivo));

        Difusibilidade = difusi(abs(DifCentrada),K, strDifusi);

        DifusiProgressiva = 0.5*(Difusibilidade(DeslProgressivo)+...
            Difusibilidade);

        DifusiRegressiva = 0.5*(Difusibilidade(DeslRegressivo)+...
            Difusibilidade);

        FluxoProgressivo = DifProgressiva.*DifusiProgressiva;
        FluxoRegressivo = DifRegressiva.*DifusiRegressiva;

        y = y+dt*(FluxoProgressivo-FluxoRegressivo);
    end
end

function y = difusi(x,K, strDifusi)
    switch lower(strDifusi)
        case 'exp'
            y = exp(-(x./K).^2);
        case 'rac'
            y = 1./(1+(x./K).^2);
    end
end

```

Código 4.1: Implementação da Equação de Perona-Malik 1-D.

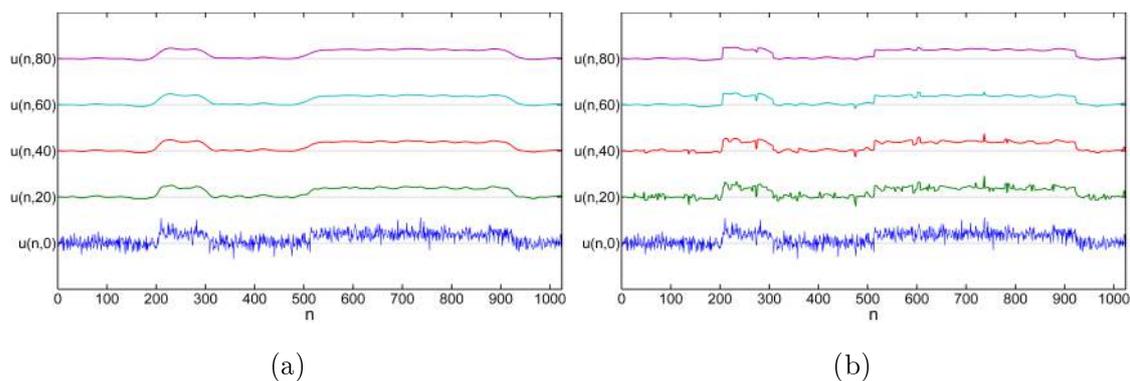


Figura 4.8: Comparação entre difusão linear (a) e não-linear de Perona-Malik (b). Em ambos os gráficos são mostrados os resultados para as difusões com diferentes tempos de parada. Nota-se que o modelo de Perona-Malik opera preservando melhor as bordas, tal como previsto. Os tempos de difusão foram calculados com incremento temporal  $\Delta t = 0,1$ . A força de difusão usada foi a expressa pela Equação (4.8) e o limiar foi tomado como sendo a décima parte da média do sinal inicial, isto é,  $K = \frac{1}{10N} \sum_{n=1}^N u_n^0$ .

**Caso bidimensional**

O desenvolvimento para imagens 2-D demanda apenas o ajuste das diferenças, de forma que estas sejam calculadas sobre uma vizinhança de pixels. Usando um esquema de vizinhança  $N_4$  (vide Seção 3.1), as diferenças finitas direcionais têm a forma

$$\begin{aligned} D_1(u_{i,j}^t) &= u_{i+1,j}^t - u_{i,j}^t, \\ D_2(u_{i,j}^t) &= u_{i,j}^t - u_{i-1,j}^t, \\ D_3(u_{i,j}^t) &= u_{i,j+1}^t - u_{i,j}^t, \text{ e} \\ D_4(u_{i,j}^t) &= u_{i,j}^t - u_{i,j-1}^t. \end{aligned}$$

Com isso, a versão 2-D da Equação (4.14) fica na forma

$$\begin{aligned} \frac{u_{i,j}^{t+1} - u_{i,j}^t}{\Delta t} &= \left( \frac{f_{i+1,j}^t + f_{i,j}^t}{2} \cdot (u_{i+1,j}^t - u_{i,j}^t) - \frac{f_{i-1,j}^t + f_{i,j}^t}{2} \cdot (u_{i,j}^t - u_{i-1,j}^t) \right) \\ &+ \left( \frac{f_{i,j+1}^t + f_{i,j}^t}{2} \cdot (u_{i,j+1}^t - u_{i,j}^t) - \frac{f_{i,j-1}^t + f_{i,j}^t}{2} \cdot (u_{i,j-1}^t - u_{i,j}^t) \right) \quad (4.16) \end{aligned}$$

Uma possível implementação em MATLAB<sup>®</sup> da Equação (4.16) pode ser vista no Código 4.2. Os argumentos de entrada são os mesmos do Código 4.1 (à exceção de  $\mathbf{x}$ , que agora trata-se de uma imagem).

Para finalizar a discussão sobre a discretização da Equação de Perona-Malik, cabe dizer que existem diversas maneiras de fazê-la. Uma discussão sobre métodos numéricos para difusão não-linear pode ser encontrada nos trabalhos de Weickert [91] e Weeratunga e Kamath [88].

```

function y = FILTRO_PERONA_MALIK_2D(x,numIter,K,strDifusi)
y = x;
dt = 0.1; % incremento temporal
[L,C] = size(y); % num. de linhas e colunas da imagem

% deslocamentos a serem usados nos calculos das diferencas finitas
DeslProgLinha = [2:L L]; DeslRegrLinha = [1 1:L-1];
DeslProgColuna = [2:C C]; DeslRegrColuna = [1 1:C-1];

for k=1:numIter
% diferencas direcionais
D1 = y(DeslProgLinha,:)-y;
D2 = y-y(DeslRegrLinha,:);
D3 = y(:,DeslProgColuna)-y;
D4 = y-y(:,DeslRegrColuna);

% diferencas centradas
DX = 0.5*(y(:,DeslProgColuna)-y(:,DeslRegrColuna));
DY = 0.5*(y(DeslProgLinha,:)-y(DeslRegrLinha,:));
% aproximacao da magnitude do gradiente
MagGrad = sqrt(DX.^2+DY.^2);

% difusibilidades centrada e direcionais
DifusiCentrada = difusi(MagGrad,K,strDifusi);

Difusi1 = 0.5*(DifusiCentrada(DeslProgLinha,:)+DifusiCentrada);
Difusi2 = 0.5*(DifusiCentrada(DeslRegrLinha,:)+DifusiCentrada);
Difusi3 = 0.5*(DifusiCentrada(:,DeslProgColuna)+DifusiCentrada);
Difusi4 = 0.5*(DifusiCentrada(:,DeslRegrColuna)+DifusiCentrada);

% fluxos
F1 = Difusi1.*D1; F2 = Difusi2.*D2;
F3 = Difusi3.*D3; F4 = Difusi4.*D4;

y = y+dt*(F1-F2 + F3-F4);
end
end

function y = difusi(x,K,strDifusi)
switch lower(strDifusi)
case 'exp'
y = exp(-(x./K).^2);
case 'rac'
y = 1./(1+(x./K).^2);
end
end
end

```

Código 4.2: Implementação da Equação de Perona-Malik 2-D.

### 4.3.2 Parâmetros do modelo

Uma vez apresentados os conceitos fundamentais da Equação de Perona-Malik e de sua discretização, faz-se necessário discutir o papel de seus parâmetros, especialmente do limiar  $K$ . O propósito desta Seção é de apresentar algumas informações básicas acerca da escolha dos parâmetros e fazer alguns apontamentos sobre a dificuldade do estabelecimento destes.

#### Tempo de difusão

A primeira questão que surge quando se usa um sistema baseado em um modelo de difusão é: ao longo de quanto tempo o sinal deve ser difundido. Para respondermos essa questão, lembremos que – no caso de difusão linear – o tempo de difusão  $t$  está vinculado

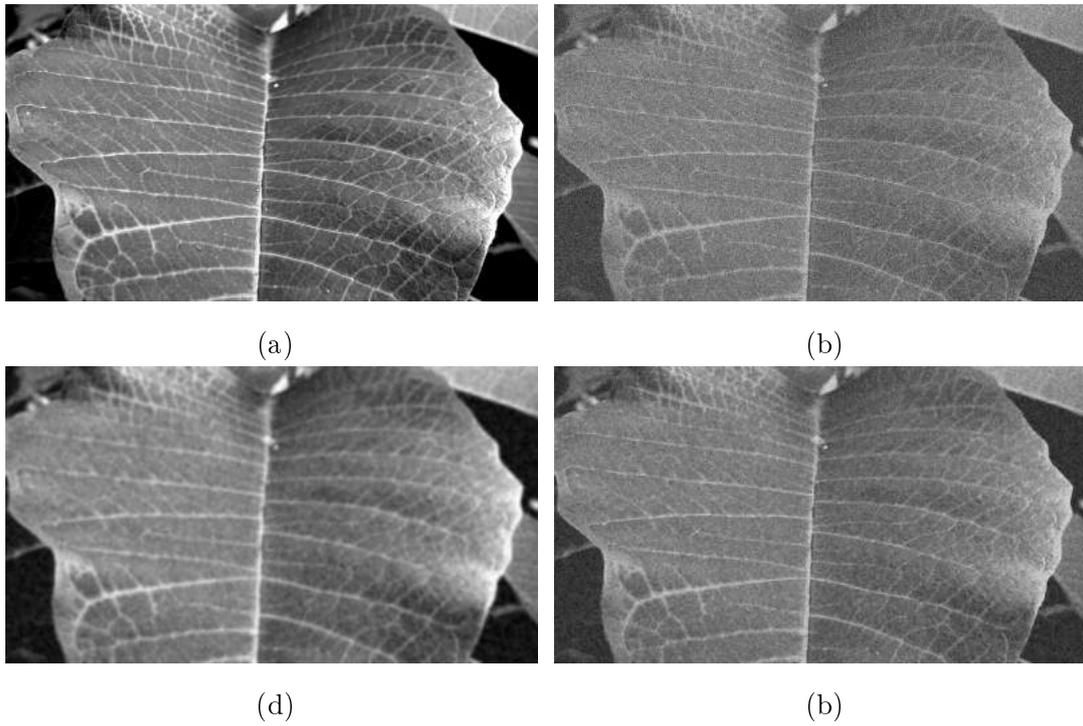


Figura 4.9: Comparação entre difusão linear e não-linear de Perona-Malik. Em (a) é mostrada a imagem original e em (b) sua versão corrompida com ruído gerado segundo distribuição uniforme, sendo a SNR é igual a -3,44 dB. Em (c) é mostrado o resultado da difusão linear sobre a imagem (b) ao longo de 100 iterações e em (d) é mostrado o resultado da difusão segundo o modelo de Perona-Malik, com  $K$  sendo a vigésima parte da média da imagem inicial e a força de difusão sendo dada pela Equação (4.8). Verifica-se que a filtragem via difusão não-linear opera de forma que as bordas são melhor preservadas quando comparada à difusão linear.

ao desvio padrão  $\sigma$  da Gaussiana equivalente segundo a relação

$$\sigma^2(t) = 2t. \quad (4.17)$$

Como a magnitude de  $\sigma$  está associada à capacidade de deterioração dos detalhes do sinal (vide Figura 4.3, por exemplo), temos então que para se deteriorar detalhes da mesma ordem de grandeza espacial de  $\sigma$ , basta difundir o sinal por um tempo

$$t_{\max} = \frac{\sigma_{\max}^2}{2}, \quad (4.18)$$

em que  $\sigma_{\max}$  corresponde a uma estimativa das dimensões espaciais máximas a serem deterioradas ao longo da difusão linear. No que diz respeito à difusão de Perona-Malik, podemos usar a estimativa dada pela Equação (4.18) sem maiores problemas, dado que – no pior cenário possível – as duas difusões se assemelham na deterioração de detalhes.

A Equação (4.18) é contínua em  $t$ . Em um cenário de discretização, ela deve ser reformulada para

$$n_{\max} = \frac{\sigma_{\max}^2}{2\Delta t}, \quad (4.19)$$

em que  $n_{\max}$  é o número máximo de iterações e  $\Delta t$  é elemento de discretização da variável  $t$ .

### Função difusibilidade

Conforme visto, Perona e Malik propuseram originalmente as duas funções difusibilidade expressas nas Equações (4.7) e (4.8):

$$f_1(\|\vec{\nabla}u\|) = \exp\left(-\left(\frac{\|\vec{\nabla}u\|}{K}\right)^2\right) \quad \text{e} \quad f_2(\|\vec{\nabla}u\|) = \frac{1}{1 + \left(\frac{\|\vec{\nabla}u\|}{K}\right)^2},$$

cujos gráficos das respectivas funções fluxo podem ser visualizados na Figura 4.10. Conforme

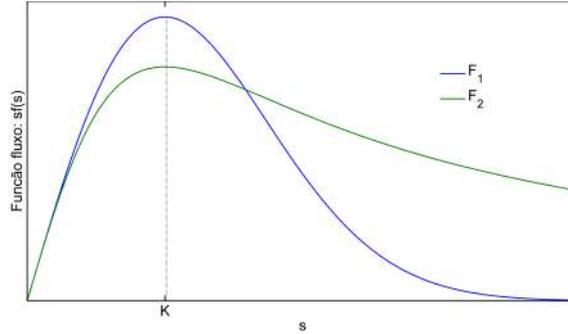


Figura 4.10: Funções fluxo associadas às difusibilidades  $f_1$  e  $f_2$  propostas por Perona-Malik. Nesta ilustração os dois fluxos possuem o mesmo ponto de máximo em  $K$ .

pode ser visto na Figura 4.10, as duas difusibilidades proporcionam fluxo mais ( $f_2$ ) ou menos ( $f_1$ ) intensos para valores acima do limiar  $K$ <sup>5</sup>. De acordo com o discutido, quão mais intenso for o *decaimento* do fluxo, mais conservadas serão as estruturas cujo módulo do gradiente seja maior que o limiar. Portanto, é natural que se busquem outras expressões para as difusibilidades e, conseqüentemente, para o fluxo. Ao longo dos anos, autores propuseram variações para a função difusibilidade. Dois exemplos são os modelos em estatística robusta analisados por Black e Sapiro [10] e a interessante Equação de Difusão Inversa Estabilizada (SIDE)<sup>6</sup>, introduzida por Pollak, Willsky e Krim [68], cuja função fluxo é descontínua.

Uma possibilidade de se obter funções fluxo com decaimento mais intenso para valores acima do limiar consiste em adicionar um parâmetro de curvatura em alguma função difusibilidade usual. Esta foi a técnica adotada por Benhamouda [7] para analisar o contraste proporcionado por diferentes fluxos. Tal abordagem inspirou o procedimento adotado neste trabalho, pois permite a avaliação da influência da forma da difusibilidade no resultado do processamento sem a necessidade de novas definições (mas às expensas de um novo parâmetro). A difusibilidade usada tem a forma

$$f_\alpha(\|\vec{\nabla}u\|) = \exp\left(-\frac{1}{\alpha} \left(\frac{\|\vec{\nabla}u\|}{K}\right)^\alpha\right), \quad \alpha > 0. \quad (4.20)$$

Uma ilustração da influência de  $\alpha$  sobre o fluxo associado à difusibilidade  $f_\alpha$  pode ser visualizada na Figura 4.11

<sup>5</sup>Note-se que, em geral, os pontos de máximo para duas funções fluxo distintas são distintos entre si. Na Figura 4.10 os pontos de máximo de  $f_1$  e  $f_2$  foram propositalmente postos coincidentes visando favorecer a observação das curvas em si.

<sup>6</sup>Do inglês, *Stabilized Inverse Diffusion Equation*.

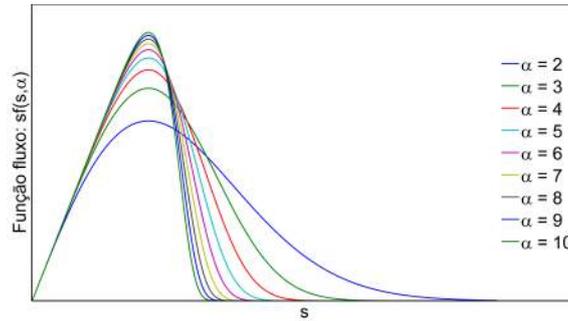


Figura 4.11: Funções fluxo da forma descrita pela Equação (4.20) para diferentes valores de  $\alpha$ . Nota-se claramente que conforme  $\alpha$  aumenta, o fluxo é mais intensamente suprimido para valores acima do limiar.

### O limiar K

O limiar K presente nas difusibilidades  $f_1$ ,  $f_2$  e  $f_\alpha$  corresponde, conforme já discutido, a um ponto de separação entre difusões direta e reversa. É fácil mostrar que fluxos  $F_\alpha(\|\vec{\nabla}u\|)$  associados às difusibilidades  $f_\alpha$  têm esse ponto de separação exatamente em K, independentemente de  $\alpha$ . Contudo, a grande questão sobre este limiar consiste em se obter um valor de K que propicie a melhor difusão possível, em termos de enhancement. Esta questão pode ser tomada como a mais profunda sobre o modelo de Perona-Malik, pelo menos do ponto de vista prático. Os próprios Perona e Malik ([63], p. 637) concluem sobre a dificuldade de se obter um valor ótimo de K:

[...]the abscissa of K of the peak of the flux function  $\Phi(\cdot)$ <sup>7</sup> has to be set according to the typical contrast value, if this changes considerably through the image the value of K has to be set locally. To tackle these difficulties anisotropic diffusion<sup>8</sup> should be implemented using local contrast and noise estimates.

Em princípio, há poucas informações que nos ajudem a definir um valor de K. Sabemos de antemão, por exemplo, que K não deve ultrapassar a escala dinâmica do sinal: se K for muito pequeno, não ocorrerá difusão e o sinal permanecerá praticamente intacto. Por outro lado, caso K seja muito grande, a difusão ocorrerá análoga à linear, suavizando completamente o sinal.

A escolha de um valor ótimo de K se assemelha à busca por um limiar de binarização: muitos métodos têm sido propostos para classes particulares de imagens, mas nada que chegue perto de uma generalização. Como exemplo de propostas podemos citar o uso de p-normas por Voci *et al.* [87] e o uso de abertura/fechamento morfológicos apresentados no mesmo trabalho; o estimador de ruído proposto por Canny [18] usado originalmente por Perona e Malik [63]; as ferramentas de estatística robusta usadas por Black *et al.* [10]; o ajuste polinomial sobre a média do gradiente proposta por Shao e Zou [79]. Daremos atenção ao estimador de ruído de Canny e às p-normas, evidenciando as particularidades destas estimativas.

<sup>7</sup>No trabalho original, Perona e Malik denotam a função fluxo por  $\Phi(\cdot)$ , ao passo que neste trabalho usamos a notação  $F(\cdot)$

<sup>8</sup>Perona e Malik adotaram o termo *difusão anisotrópica* para denominar sua técnica. Neste trabalho esse termo não é usado, sendo substituído por difusão não-linear.

**p-Normas**

Seja  $u = (u_1, u_2, \dots, u_N)$  um elemento genérico de um espaço vetorial tal que valha  $\sum_{i=1}^N |u_i|^p < \infty$  para  $1 \leq p < \infty$ . Definimos sua p-norma como sendo

$$\|u\|_p = \left( \sum_{i=1}^N |u_i|^p \right)^{\frac{1}{p}}. \quad (4.21)$$

Note-se que para  $p = 2$  obtemos a norma Euclidiana. Voci *et al.* sugerem em [87] fazer uso de uma p-norma sobre a imagem  $u^t$  para se obter o valor de K. Mais especificamente,

$$K^t = \frac{\sigma}{MN} \|u^t\|_p, \quad \text{com } \sigma \propto \|u^t\|_1, \quad (4.22)$$

em que M e N correspondem ao número de linhas e colunas de  $u$ , respectivamente.

Como uma soma sobre toda a imagem pode gerar valores muito maiores que os tolerados dentro da faixa dinâmica do sinal, o termo  $\frac{\sigma}{MN}$  aparece como um fator de normalização. Note que para  $\sigma = \|u^t\|_1$  esse fator de normalização é igual à média de  $u^t$  (caso as entradas de  $u^t$  sejam positivas). Podemos destacar dois pontos de forte fragilidade na proposta de Voci *et al.*:

1. O fator  $\frac{1}{MN}$  presente na Equação (4.22) faz com que o cálculo de K leve em consideração o *tamanho* da imagem, mesmo que essa grandeza nada tenha a ver com as relações de contraste que definem regiões de borda ou regiões planas. Isso implica que imagens semelhantes mas de dimensões distintas teriam evoluções discrepantes simplesmente por serem de tamanhos diferentes, o que não é razoável em um cenário de operações locais;
2. Não há qualquer menção sobre qual valor de  $p$  deva ser usado no cômputo de K, o que resulta na inserção de mais um parâmetro na técnica.

Para fins de comparação com outras formas de cômputo de K, usamos neste trabalho  $p = 3$  e  $\sigma = \|\vec{\nabla}u\|_1^t$ . A escolha de se operar  $\|\vec{\nabla}u\|$  difere da originalmente proposta, mas no nosso entendimento é mais condizente com o fato de que quaisquer informações sobre contraste devem ser obtidas do sinal que, efetivamente, corresponde ao detector de bordas da técnica. Com isso, obtemos

$$K^t = \frac{\|\vec{\nabla}u\|_1^t}{MN} \|\vec{\nabla}u\|_3^t. \quad (4.23)$$

Para encerrar a discussão sobre a forma de cálculo de K proposta por Voci, é necessário que sejam feitas duas observações:

**Observação 1.** As p-normas  $\|u^t\|_p$  para  $p \geq 1$  decrescem conforme t aumenta. Esse resultado vale também para um esquema de discretização da forma expressa pelas Equações (4.13) e (4.16) [89, 91]. Em particular, a *energia* de  $u$ , dada por  $\|u^t\|_2$ , decresce no tempo.

**Observação 2.** O valor médio de  $u$  permanece constante ao longo de toda a evolução do sinal. Em espaço e tempo contínuos isso significa que

$$\int_{\Omega} u(\vec{r}, t) dv = \int_{\Omega} u(\vec{r}, 0) dv = \text{constante}, \quad (4.24)$$

desde que as condições de fronteira sejam as mesmas que as expressas na Equação (4.11) ( $\partial_{\vec{n}}u(\vec{r},t) = 0$ ). Novamente,  $\Omega$  corresponde ao domínio da imagem.

Para demonstrar a preservação do valor médio, tomemos a derivada temporal da integral expressa em (4.24) e usemos do fato que os dois operadores comutam:

$$\frac{d}{dt} \int_{\Omega} u dv = \int_{\Omega} \partial_t u dv = \int_{\Omega} \vec{\nabla} \cdot (f(\|\vec{\nabla}u\|)\vec{\nabla}u) dv.$$

Usando o teorema de Gauss, obtemos

$$\int_{\Omega} \vec{\nabla} \cdot (f(\|\vec{\nabla}u\|)\vec{\nabla}u) dv = \int_{\partial\Omega} f(\|\vec{\nabla}u\|)\vec{\nabla}u \cdot \hat{n} da = 0,$$

pois  $\vec{\nabla}u \cdot \hat{n} = \partial_{\vec{n}}u = 0$ . Logo, temos que  $\frac{d}{dt} \int_{\Omega} u(\vec{r},t) dv = 0$ , o que prova a conservação do valor médio de  $u$ . Benhamouda [7] e Weickert [91] mostram que os esquemas de diferenças finitas mostrados nas Equações (4.13) e (4.16) obedecem à conservação do valor médio.

De acordo com a Observação 1, vemos que a proposta de Voci para o cômputo de  $K$  apresenta comportamento decrescente ao longo da evolução. Essa parece ser a única característica desejável da proposta, frente à inconveniência de se levar em conta o tamanho da imagem e de não considerar absolutamente nenhum atributo relacionado a contraste. Por fim, a Observação 2 mostra que qualquer forma de cômputo de  $K$  que leve em consideração *exclusivamente* o valor médio de  $u$  não variará no tempo, o que não é interessante, dado que espera-se que  $K$  seja obtido em função das características de  $u$  em cada instante  $t$ .

### Estimador de ruído de Canny

Em seu artigo clássico, Canny [18] propôs sua famosa técnica para detecção de bordas. Esta baseia-se na suavização da imagem inicial por filtros Gaussianos para posterior obtenção da magnitude do gradiente. De posse desse sinal, são usados dois thresholds para detectar as bordas e uma operação para fortalecer bordas é aplicada.

Um dos parâmetros do método é a estimativa da quantidade de ruído na imagem. Para tanto, Canny propõe que essa quantidade seja estimada como sendo igual ao valor de intensidade correspondente ao 80ºpercentil da distribuição de níveis de intensidade da imagem. Para se obter esse valor, faz-se uso da função distribuição acumulada (f.d.a) associada à distribuição de níveis de intensidade da imagem [26]. Para uma imagem digital  $u$  com histograma  $h$  (vide Seção 3.1), sua soma acumulada é dada por

$$T(L) = \sum_{i=1}^L h_i,$$

em que  $L$  é um índice pré-estabelecido.

Perona e Malik [63] propõem calcular  $K$  como sendo o 90ºpercentil da distribuição de níveis de intensidade da magnitude do gradiente da imagem a cada instante  $t$ . A escolha de se usar o estimador de ruído de Canny sobre  $\|\vec{\nabla}u\|^t$  decorre do fato de que o modelo de Perona-Malik usa essa grandeza como detector de bordas. Em síntese, dado o histograma  $h^t = \{h_1^t, h_2^t, \dots, h_N^t\}$  de  $\|\vec{\nabla}u\|^t$ , cuja abcissa é dada pelo vetor  $x^t = \{x_1^t, x_2^t, \dots, x_N^t\}$ , o valor de  $K$  calculado por Perona-Malik é obtido como sendo

$$K^t = x_L^t = \arg \left\{ \sum_{i=1}^L h_i^t \geq 0,9 \right\}, \quad (4.25)$$

em que  $L$  é o menor índice pertencente a  $[1, N]$  tal que a desigualdade seja satisfeita.

No que se refere à proposta original de Canny, há de se notar que ele sugere o estimador de ruído com base na hipótese de que o ruído possui uma distribuição Gaussiana, o que não necessariamente é verdade. Isso deve ser levado em consideração ao se fazer uso deste estimador.

### 4.3.3 Considerações finais sobre a Equação de Perona-Malik

De acordo com todo o exposto neste capítulo, vemos que a Equação de Perona-Malik constitui uma ferramenta interessante no que diz respeito ao enhancement e reconstrução de sinais, pois cumpre simultaneamente tarefas mutuamente conflituosas: realce de bordas e suavização. Contudo, o preço que se paga por essa possibilidade é a forte dependência da técnica em relação a seus parâmetros ( $t_{\max}$ ,  $K$  e  $f(\|\vec{\nabla}u\|)$ ). Em particular, a escolha do limiar de separação  $K$  ainda representa a maior dificuldade prática quanto ao uso da difusão não-linear.

Um ponto mais sensível a ser analisado diz respeito à solução da Equação de Perona-Malik. Como visto, é possível que a difusão ocorra análoga à equação reversa do calor (Equação (4.4)), promovendo o realce das estruturas submetidas a esse regime. Contudo, foi dito que a equação reversa do calor constitui um problema mal-posto, em que a unicidade da solução e sua convergência não estão garantidos. Nesse sentido, poderíamos nos questionar sobre a viabilidade de se usar a difusão não-linear de Perona-Malik: se o problema é mal-posto, duas imagens muito semelhantes postas a difundir sob as mesmas condições poderiam gerar resultados bastante distintos entre si. Pior do que isso, poderíamos lidar com resultados distintos associados ao processamento da mesma imagem. Com vistas à resolução dessa fragilidade, Catté, Lions, Morel e Coll propuseram em 1992 [19] uma variação da Equação de Perona-Malik, que consiste em inserir uma regularização no argumento da difusibilidade, ou seja,

$$\begin{cases} \partial_t u(\vec{r}, t) &= \vec{\nabla} \cdot \left( f(\|\vec{\nabla}(G_\sigma * u)(\vec{r}, t)\|) \vec{\nabla} u(\vec{r}, t) \right) \\ u(\vec{r}, 0) &= u_{\text{inicial}}. \end{cases} \quad (4.26)$$

Na Equação (4.26), o argumento da função difusibilidade agora é uma versão suavizada de  $\|\vec{\nabla}u\|$ <sup>9</sup> segundo uma Gaussiana  $G_\sigma(\vec{r}, t)$  cujo desvio-padrão é  $\sigma$ . Verifica-se [19, 5] que a Equação (4.26) é bem-posta.

Sendo a variação da Equação de Perona-Malik proposta por Catté *et al.* um problema bem-posto, parece razoável adotar-se a regularização baseada na filtragem com  $G_\sigma$ . Contudo, conforme visto exhaustivamente, a filtragem por Gaussiana provoca perda de detalhes da imagem. Dentro do contexto de imagens mamográficas, em que alguns achados malignos correspondem aproximadamente a pontos na imagem, uma pré-filtragem por Gaussiana (ou qualquer outro passa-baixas) poderia extinguir essa informação e levar a um incremento no número de falsos-positivos. Tendo isso em vista, não adotamos neste trabalho o modelo de Catté *et al.* Contudo, há o fato de que o esquema de diferenças finitas usado para discretizar a Equação de Perona-Malik (Equações (4.14) e (4.16)) faz com que o problema torne-se bem-posto [90, 89], dispensando – em princípio – a necessidade de regularizações.

---

<sup>9</sup>Note que  $\vec{\nabla}(G_\sigma * u) = G_\sigma * \vec{\nabla}u$ .

# Capítulo 5

## Metodologia

### 5.1 Materiais

#### 5.1.1 Base de imagens mamográficas

Dada a natureza deste trabalho, fez-se necessário construir uma base de imagens mamográficas para que as técnicas estudadas pudessem ser testadas e validadas. Entenda-se por base de imagens um conjunto de imagens mamográficas acompanhadas de algum tipo de informação dada por profissional médico. Isto é necessário para que as detecções realizadas pelos métodos desenvolvidos possam ser comparadas com as detecções do profissional radiologista.

Inicialmente, foram cogitadas duas bases de imagens mamográficas:

- mini-MIAS: (mini) Mammographic Image Analysis Society Digital Mammogram Database; e
- Bancoweb: Base de Imagens Mamográficas para Auxílio em Avaliações de esquemas CAD.

A seguir serão feitas algumas observações sobre estas duas bases.

#### mini-MIAS

A base mini-MIAS [83] corresponde a uma versão reduzida do trabalho original de Suckling *et al.* [82]. Ela é composta por 322 imagens *digitalizadas*, correspondentes a 161 exames (duas imagens por paciente, uma para cada mama) provenientes do United Kingdom National Breast Screening Programme. As imagens são disponibilizadas em regiões de 1024x1024 pixels e todas são quantizadas em 8 bits. O tamanho físico dos pixels das amostras dessa base é de 0,200 mm (alternativamente, as imagens têm uma resolução espacial de 5,00 px/mm ou 127,00 ppi). Cada amostra é classificada de acordo com o tipo de tecido mamário predominante:

- gorduroso;
- gorduroso-glandular; ou
- denso-glandular.

As anormalidades são classificadas em:

- calcificação;
- massa bem-definida;
- massa espiculada;
- distorção arquitetural;
- assimetria; ou
- normal.

No que diz respeito à descrição dos achados, são disponibilizados dois descritores quantitativos (coordenadas  $(x_{CM}, y_{CM})$  do centro de massa do achado e seu raio aproximado) e um qualitativo (um caractere indicativo) que refere-se à malignidade ou benignidade da estrutura descrita.

### **Bancoweb**

A base Bancoweb originou-se do trabalho de mestrado de B. R. N. Matheus [55]. A base conta com 1703<sup>1</sup> imagens *digitalizadas* oriundas dos Hospitais das Clínicas de Ribeirão Preto (HC-USP/RP), da Faculdade de Medicina da UNESP de Botucatu (FMB-UNESP), Santa Casa de Misericórdia de São Carlos e Hospital São Paulo (HSP - UNIFESP). As imagens não possuem nenhum tipo de segmentação, tal como os recortes das regiões do mini-MIAS. A profundidade, em bits, da grande maioria das imagens (99,82% do total) é de 12 bits. Em 1425 das imagens, o pixel tem tamanho físico de 0,150 mm (resolução de 6,67 px/mm ou 169,42 ppi). Em 275 das imagens, o pixel possui tamanho físico de 0,075 mm (13,33 px/mm ou 338,67 ppi). Nas três imagens restantes esta informação não está clara. Cada amostra é classificada quanto a diversos critérios. Um deles é o tipo de tecido predominante:

- mama densa; ou
- mama liposubstituída.

Em caso de mamas com achados, as mamas são classificadas segundo o tipo de achado. Alguns exemplos são:

- atrofia lobular;
- calcificações;
- densidades assimétricas;
- fibroses;
- linfonodos;
- nódulos, dentre outros.

O Bancoweb conta com um sistema de busca razoável, que permite filtrar resultados segundo diversos critérios, inclusive segundo a classificação BI-RADS<sup>®</sup>.

---

<sup>1</sup>Em princípio, o Bancoweb permite atualizações; portanto, os números podem ser alterados. Os números apresentados no texto referem-se a um levantamento feito em Abril/2015.

## BIM e BIM2

Quando comparamos as bases mini-MIAS e Bancoweb, vemos que esta última se mostra mais robusta que a primeira: além de conter um sistema de buscas que acelera a localização de imagens dotadas de características semelhantes, a qualidade das imagens é melhor (maior resolução espacial e maior profundidade em bits). Contudo, o Bancoweb possui algumas inconsistências importantes que inviabilizaram sua utilização como base para teste dos algoritmos de detecção. Há imagens duplicadas, imagens não condizentes com os atributos a elas associados e arquivos corrompidos. Portanto, a confiabilidade atribuída a esta base foi diminuída, colocando-a como recurso de teste para algoritmos com menor complexidade e criticalidade, tal como a segmentação das regiões de interesse das mamas. Com isso em mente, foi necessário a utilização de outra base de imagens mamográficas.

Com o auxílio do grupo de Física Aplicada ao Rádiodiagnóstico da Universidade Estadual Paulista (LAFAR-UNESP), compusemos uma base de imagens mamográficas (BIM) constituída de 51 imagens. Destas, 20 são provenientes do Bancoweb e 31 provenientes da Faculdade de Medicina da Unesp de Botucatu. Este segundo conjunto de imagens corresponde a amostras obtidas por meio de *radiografia computadorizada*, apresentando maior resolução espacial (20 px/mm, ou pixels com 0,050 mm). A quantização deste segundo conjunto é idêntica a das imagens do Bancoweb, isto é, 12 bits.

Dois radiologistas analisaram as imagens, contornando os achados e classificando-os quanto sua natureza (calcificação, linfonodo, assimetria etc.). Um dos radiologistas validou as 20 imagens provenientes do Bancoweb e o outro analisou as restantes, da FMB. Com isso, foi providenciada uma base robusta, com imagens diversas cujos diferentes achados foram devidamente descritos quantitativa e qualitativamente. Para fins de testes rápidos, construímos uma terceira base, chamada BIM2, constituída de todas as regiões de interesse das imagens da BIM. A Tabela 5.1 sintetiza as características básicas de cada base.

Base	Número de amostras	Resolução (px/mm)	Quantização (bits)	Informações sobre os achados
mini-MIAS	322	5,00	8	Coordenadas do centro de massa e raio médio ( $x_{CM}, y_{CM}, r$ ); descrição qualitativa.
Bancoweb	1703	6,77 13,33	12	Descrição qualitativa
BIM/BIM2	51/68	6,77 13,33 20,00	12	Contorno dos achados e descrição qualitativa.

Tabela 5.1: Comparação entre as bases cogitadas (mini-MIAS e Bancoweb) e as bases construídas para este trabalho (BIM e BIM2).

### 5.1.2 Recursos computacionais

Para a realização deste trabalho, foi usada a linguagem de computação científica MATLAB<sup>®</sup> [84], versão R2014. Ela foi escolhida por proporcionar ambiente adequado para programação de dados organizados matricialmente, além de possuir extenso conjunto de ferramentas auxiliares em processamento de sinais e visualização científica. Adicionalmente, alguma análise esporádica ocorreu por meio do software livre ImageJ [70], que

também corresponde a uma poderosa ferramenta na área de análise de imagens. Em termos de hardware, foi feito uso de um dos computadores do cluster da Coordenação de Atividades Técnicas do CBPF, operando com sistema operacional Microsoft<sup>®</sup> Windows 8. A máquina usada possui 128 GB de RAM e 64 núcleos de processamento com frequência máxima de 3,20 GHz.

## 5.2 Algoritmos desenvolvidos

### 5.2.1 Cadeia de processamento

A tarefa de detecção e análise em imagens envolve mais do que simplesmente a segmentação. É necessário segmentar as regiões de interesse, realizar o procedimento desejado e quantificar as informações medidas. Neste trabalho, desenvolvemos uma cadeia de processamento organizada segundo o esquema mostrado na Figura 5.1. O propósito desta Seção é detalhar esta organização, uma vez que a cadeia proposta operou neste trabalho como o instrumento de medidas. Nas próximas seções, será feito o detalhamento de cada etapa do processo.

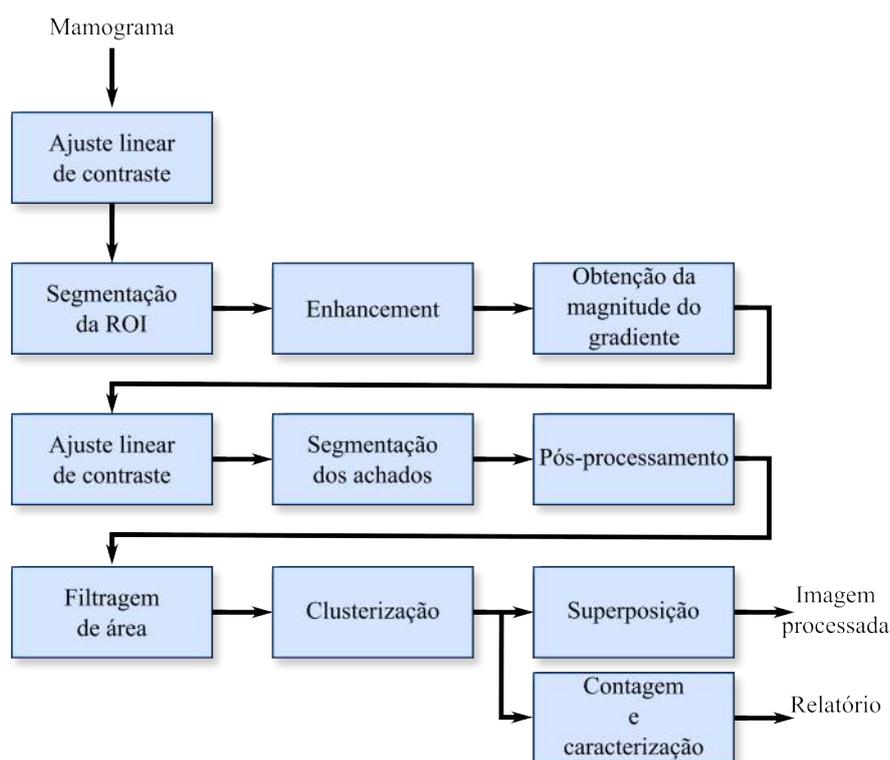


Figura 5.1: Cadeia de processamento desenvolvida neste trabalho para detecção e análise de estruturas em imagens mamográficas.

### Segmentação das Regiões de Interesse

A dinâmica de processamento da cadeia desenvolvida inicia-se com o carregamento do mamograma da base desejada. Efetivamente, é carregada uma estrutura contendo o mamograma, a região de interesse demarcada pelo médico (ROI-M) e o tamanho físico do

pixel. O primeiro passo consiste em submeter o mamograma inteiro a um ajuste linear de contraste, tal como definido na Equação (3.5). Em sequência, a mama é detectada dentre todos os elementos constituintes da imagem e o fragmento contendo apenas a mama é tomado como a ROI efetivamente processada ao longo de toda a cadeia<sup>2</sup>. A descrição detalhada do processo de segmentação da ROI é exposta a seguir.

Deve-se atentar para o fato de que um mamograma contém mais informação visual do que aquela pertinente à mama: há identificações do tipo de incidência; identificação do paciente; artefatos de digitalização ou exposição etc. A Figura 5.2 ilustra estas ocorrências. Do ponto de vista de processamento e detecção *na região compreendida pela mama*, todo o resto pode ser encarado como ruído, já que há óbvia alteração do histograma e inserção de informação adicional irrelevante. Com base nessa constatação, é preciso preparar as imagens para o processamento efetivo.

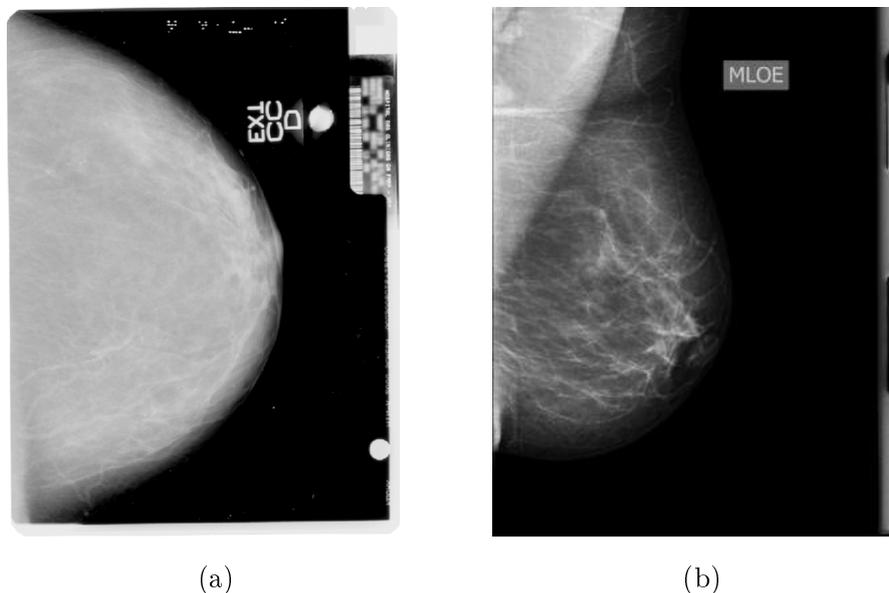


Figura 5.2: Presença de objetos/artefatos diversos em mamogramas: em (a) há identificação da paciente (borrada para não exposição), áreas não sensibilizadas no filme e informações do tipo de incidência. Em (b) há informações sobre a incidência e é possível visualizar detalhes de regiões do cassete. Toda essa informação adicional pode ser encarada como ruído.

O algoritmo desenvolvido realiza a segmentação da mama em duas etapas consecutivas:

1. **ROI-1:** Segmentação da região *que contém a mama*; e
2. **ROI-2:** Segmentação, dentro da ROI-1, da mama propriamente dita.

A Figura 5.3 ilustra esta sequência sobre o mamograma mostrado na Figura 5.2 (a).

### ROI-1

A técnica desenvolvida para a segmentação da região que contém a mama baseia-se na análise das projeções vertical e horizontal do mamograma, tais como definidas na Equação

<sup>2</sup>Embora apenas a região da mama seja processada, a imagem final é reconstruída totalmente.

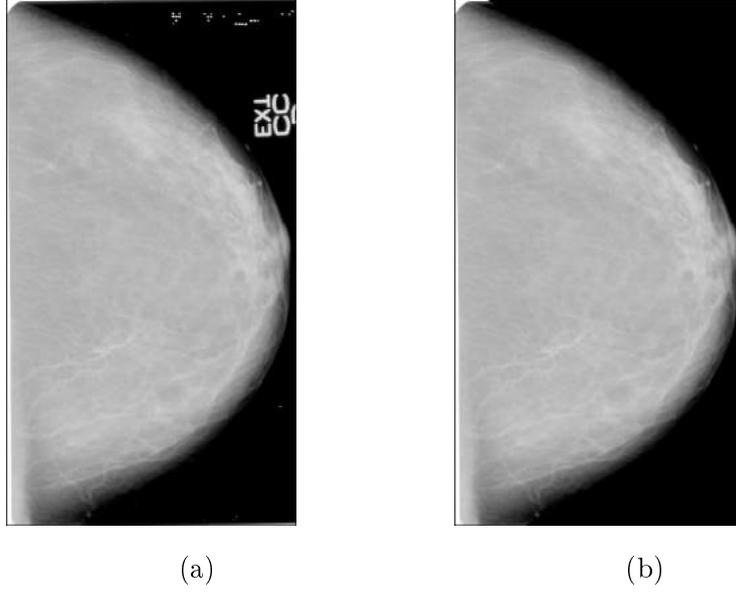


Figura 5.3: Segmentação da ROI sobre o mamograma mostrado na Figura 5.2 (a). Em (a) é segmentada a região que contém a mama (ROI-1) e em (b) é segmentada a mama propriamente dita (ROI-2).

(3.32):

$$P_x(u) = \left( \sum_{i=1}^M u_{i1}, \sum_{i=1}^M u_{i2}, \dots, \sum_{i=1}^M u_{iN} \right) \text{ e } P_y(u) = \left( \sum_{j=1}^N u_{1j}, \sum_{j=1}^N u_{2j}, \dots, \sum_{j=1}^N u_{Mj} \right),$$

em que  $u$  é a imagem analisada.

Tal como discutido na Seção 3.4.2, as projeções  $P_x(\cdot)$  e  $P_y(\cdot)$  fornecem informações sobre as variações de intensidade em cada direção da imagem. Com base nessa característica, buscou-se localizar pontos em ambas as direções da imagem com vistas a delimitar um setor dela que contenha a mama. A Figura 5.4 ilustra essa segmentação.

A obtenção dos pontos de delimitação da ROI-1 ocorre por meio da análise dos *mínimos* das projeções. Na direção horizontal (colunas), o ponto de corte é tomado como sendo o mínimo de  $P_x(u)$  mais próximo do ponto médio horizontal da imagem. A razão para esta escolha é que, de um modo suave ou não, a mama – que geralmente não ocupa toda a extensão horizontal do mamograma – apresenta uma transição para a região de fundo, caracterizando um ponto de mínimo. Como podem ocorrer vários pontos de mínimo para além da região da mama, escolhe-se dentro do conjunto aquele que está mais próximo do meio da imagem. Na direção vertical (linhas), procedimento semelhante é realizado. Contudo, a projeção  $P_y(u)$  é particionada em três segmentos de comprimentos iguais<sup>3</sup>. A região central corresponde àquela que contém – necessariamente – a mama. As regiões superior e inferior são analisadas e seus respectivos mínimos são detectados. O ponto de corte superior é tomado como sendo o ponto de mínimo mais próximo de  $1/3$  do número de linhas da imagem, ao passo que o ponto de corte inferior é tomado como sendo o ponto de mínimo mais próximo de  $2/3$  do número de linhas da imagem. A razão para este procedimento é a mesma do caso da horizontal: a mama apresenta borda de transição nestas regiões para o fundo da imagem.

<sup>3</sup>Ou aproximadamente iguais, dada a natureza discreta do sinal.

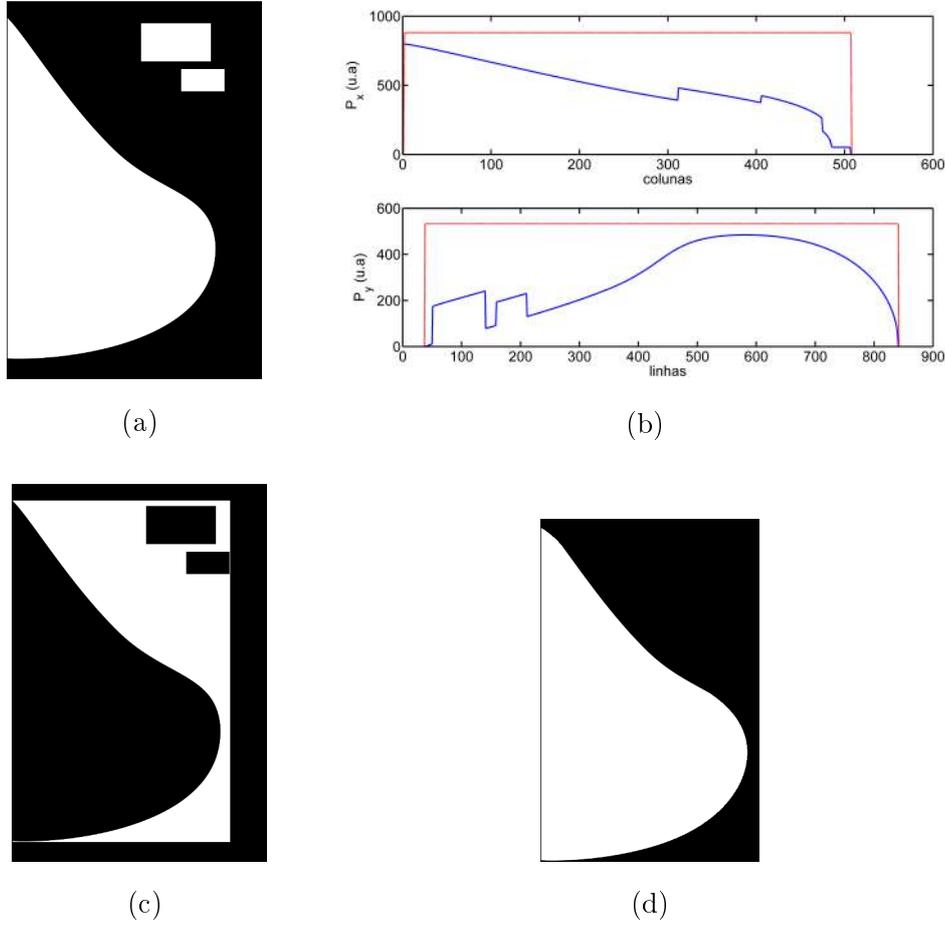


Figura 5.4: Ilustração do conceito de segmentação das ROIs 1 e 2. Em (a) é mostrado um modelo de mamograma, contendo informação adicional à mama. Em (b) são mostradas suas projeções, cujos pontos de delimitação foram detectados e servem para demarcar a ROI-1, compreendida entre as regiões internas às linhas em vermelho. Em (c) é mostrado o modelo de mamograma superposto a uma máscara delimitadora construída com as informações de (b) e em (d) é mostrada a ROI-2, já sem nenhum detalhe adicional à mama.

Para fins de exemplificação, realizemos a análise das projeções do mamograma mostrado na Figura 5.5. O mamograma original possui dimensões de  $1495 \times 1140$  pixels e suas projeções são mostradas na mesma figura. As coordenadas estimadas dos pontos de corte horizontal,  $y_{h_i}$  e  $y_{h_f}$ , e vertical,  $x_v$ , para este mamograma são  $(y_{h_i}, y_{h_f}, x_v) = (226, 1309, 503)$  px.

De posse de  $(y_{h_i}, y_{h_f}, x_v)$ , o mamograma é particionado em seis diferentes regiões, tal como mostrado na Figura 5.6. Os setores superiores e inferiores são descartados, pois – segundo a proposta – correspondem a regiões alheias à mama. Os *candidatos* a ROI-1 são, necessariamente, os setores intermediários ( $R_{esq}$  e  $R_{dir}$ ).

Uma vez obtidos os candidatos a ROI-1, é preciso escolher a região que contenha efetivamente a mama. Isso é feito via análise do descritor área das respectivas regiões, que é definida segundo a Equação (3.31):

$$A(u) = \sum_{i=1}^M \sum_{j=1}^N u_{ij}.$$

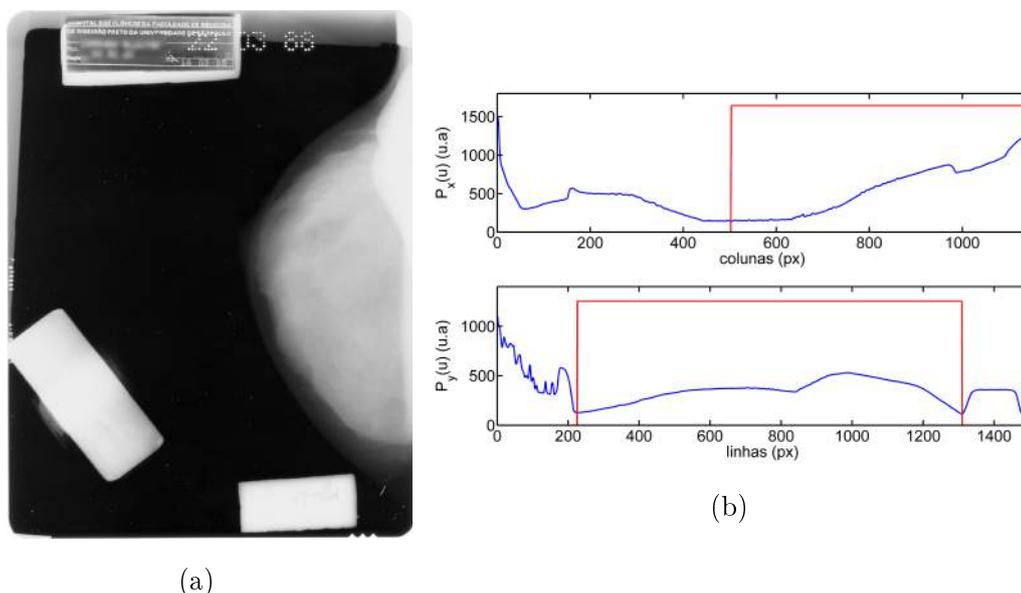


Figura 5.5: Em (a) é mostrado o mamograma  $u$  a ser utilizado na exemplificação do processo de segmentação das ROIs 1 e 2 (a identificação da paciente foi corrompida para preservação de identidade). Em (b) são mostradas suas projeções cujos pontos de corte (vide texto) estão representados em vermelho.

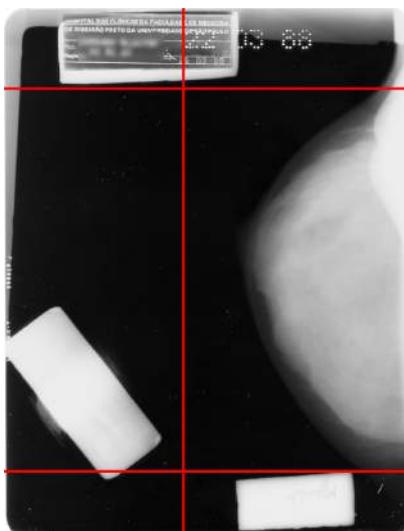


Figura 5.6: Mamograma mostrado na Figura 5.5 (a) particionado em seis diferentes regiões segundo análise automática de suas projeções.

Para tanto, as duas regiões passam por um processo de binarização, com o limiar sendo estimado pelo valor médio dos níveis de intensidade (vide Equação (B.1), Apêndice B). A região que contém a mama é definida como sendo, dentre as duas regiões devidamente binarizadas via valor médio, aquela que possui maior área. Este procedimento se baseia na constatação de que a região que não contém a mama possui muito mais informação *de fundo* em sua composição. Isto é: em média, os pixels da região que não contém a mama são *mais escuros* que os da outra região em questão <sup>4</sup>. Isso reflete, necessariamente em sua

<sup>4</sup>Note que para que esta hipótese seja válida, a classificação dos pixels da imagem deve ser do tipo preto

área. Na Figura 5.7 podem ser visualizadas as regiões avaliadas no exemplo, bem como suas versões binarizadas e seus histogramas.

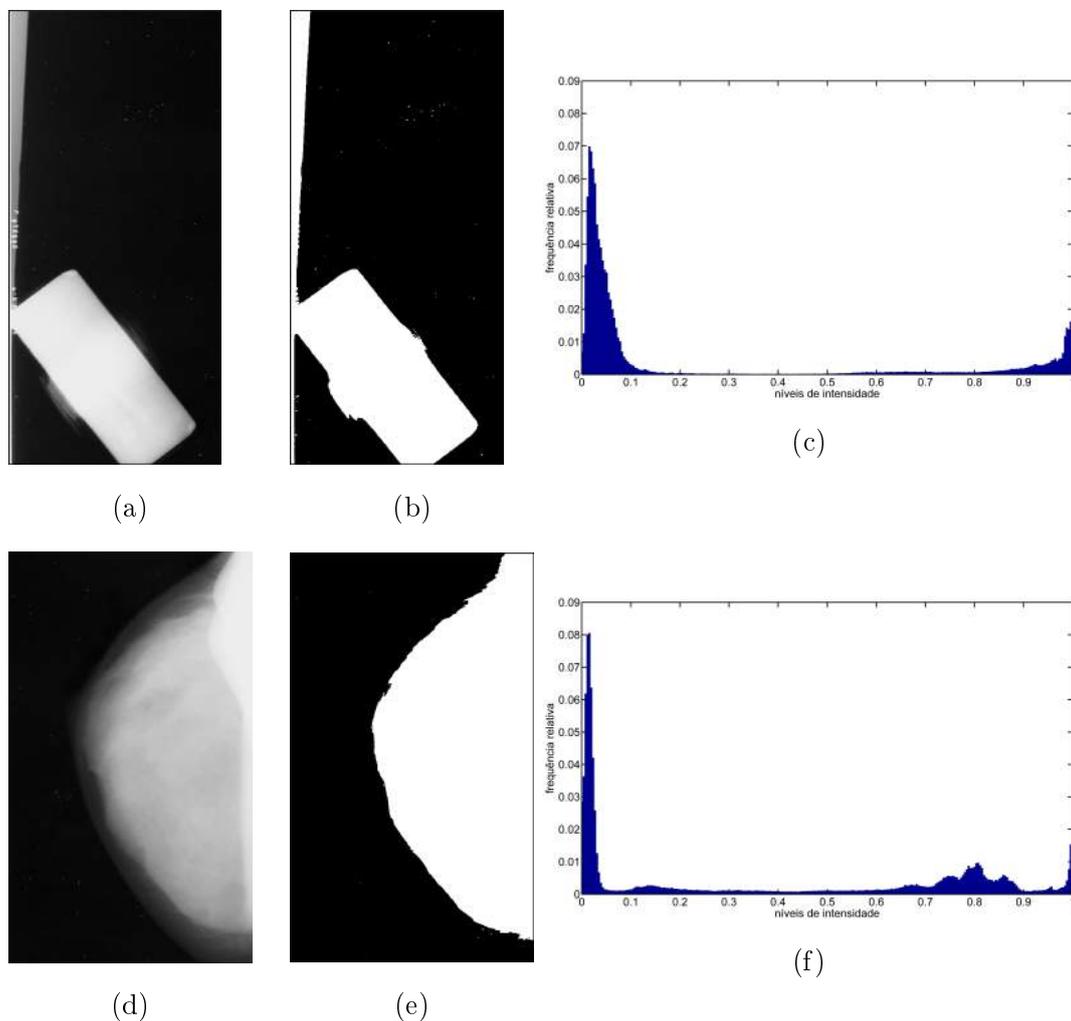


Figura 5.7: Determinação das áreas das regiões candidatas a ROI-1: em (a) e (d) são mostradas as regiões avaliadas, em (b) e (e) são mostradas suas versões binárias obtidas com threshold por valor médio. Em (c) e (f) constam os histogramas de (a) e (d), respectivamente. Nota-se claramente que o histograma (c) representa uma imagem com mais informação de fundo (baixa intensidade luminosa), dada a maior largura desta parcela da distribuição, associada a contagens menos numerosas referentes aos maiores valores de intensidade luminosa.

Para o mamograma usado no exemplo, temos que as áreas calculadas para as regiões mostradas nas Figuras 5.7 (b) e (e) são iguais a 126 670 e 301 328 pixels, respectivamente. Este resultado está de acordo com o discutido e permite a definição da ROI-1, a ser processada com vistas à definição da ROI-2.

## ROI-2

Uma vez definida a ROI-1, é preciso refinar a imagem de forma que, idealmente, apenas a mama conste no sinal. Isso é feito com uso de processamento de alto nível: dentre os = 0 e branco = 1 (ou equivalente para outros intervalos de quantização).

objetos remanescentes na ROI-1, a mama é tomada como sendo aquele com maior área; ou seja, a ROI-1 é submetida a um filtro de área. Antes que essa operação seja realizada, a ROI-1 é submetida a uma filtragem Gaussiana intensa para remoção de detalhes espúrios e extensão das bordas<sup>5</sup>, seguida de binarização por valor médio, com vistas a permitir processos de labeling. O filtro Gaussiano usado para esta tarefa possui dimensão de  $180 \times 180$  pixels e desvio padrão  $\sigma$  igual a 60 px. O resultado desta filtragem e limiarização sobre a ROI-1 do mamograma usado no exemplo pode ser visualizado na Figura 5.8

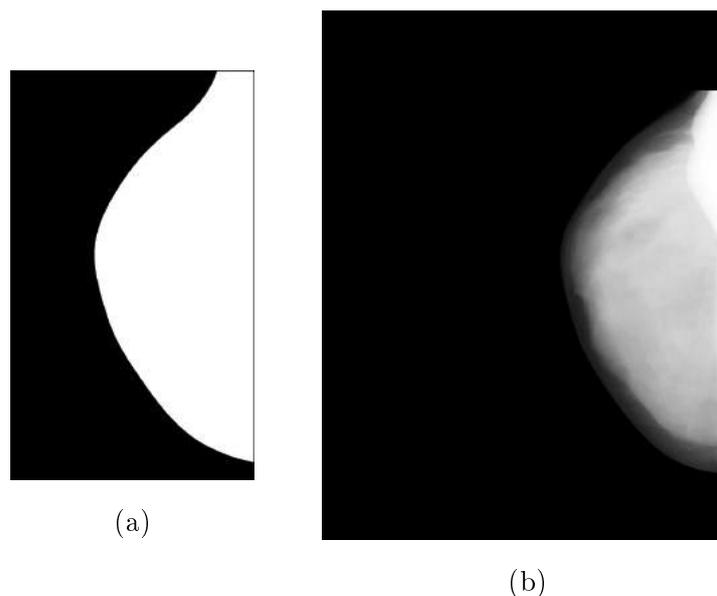


Figura 5.8: Resultado do processamento da ROI: em (a) é mostrada a ROI-2, obtida por filtragem Gaussiana e binarização por valor médio sobre a ROI-1. Para este caso em particular, há apenas um objeto em cena, correspondente à própria mama. Em (b) é mostrado o resultado final sobre o mamograma mostrado na Figura 5.5. Nota-se que há a efetiva remoção de objetos irrelevantes.

De posse do sinal suavizado e binarizado, um pós-processamento morfológico é realizado para que buracos sejam preenchidos e o filtro de área citado no parágrafo anterior possa ser aplicado. O resultado é uma máscara a ser aplicada sobre a ROI-1<sup>6</sup>. Com isso, obtém-se a ROI-2, que corresponde ao sinal efetivamente processado ao longo de toda a cadeia de processamento. Uma estrutura algorítmica sintetizando todos os passos para obtenção das ROIs 1 e 2 pode ser vista na Figura 5.9.

### Enhancement via Perona-Malik

Seguindo a sequência proposta, a ROI segmentada no segundo passo da cadeia é submetida ao processo de enhancement investigado: a filtragem via Equação de Perona-Malik. Tal como visto na Seção 4.3.2, essa filtragem exige a determinação prévia dos seguintes parâmetros:

- Tempo de difusão;

<sup>5</sup>Como visto no Capítulo 4, a filtragem Gaussiana deteriora bordas, apresentando um efeito visual semelhante a uma diluição. Logo, este filtro mostra-se eficaz no que diz respeito à uniformização e prolongamento de regiões.

<sup>6</sup>Aplicar uma máscara sobre uma imagem corresponde a realizar uma multiplicação ponto a ponto.

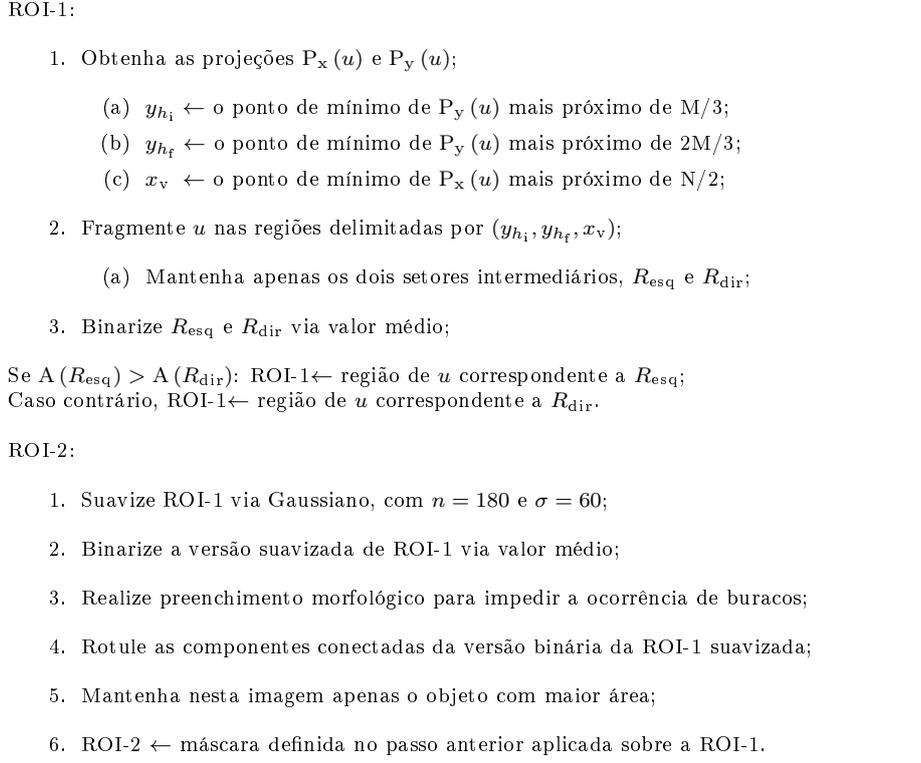


Figura 5.9: Estrutura do algoritmo para definição das ROIs 1 e 2.

- Função difusibilidade; e
- Forma de cômputo do limiar de separação.

De acordo com a Equação (4.19), o número máximo de iterações está relacionado – pelo menos no caso linear – às dimensões máximas a serem deterioradas na imagem:

$$n_{\max} = \frac{\sigma_{\max}^2}{2\Delta t}.$$

Sendo  $h_{\text{pix}}$  o tamanho físico do pixel do mamograma processado, o número máximo de iterações tal que detalhes da ordem de  $\sigma_{\max}$  sejam eventualmente deteriorados é dado por

$$n_{\max} = \left\lceil \frac{1}{2\Delta t} \left( \frac{\sigma_{\max}}{h_{\text{pix}}} \right)^2 \right\rceil. \quad (5.1)$$

Em todo o processamento, o incremento  $\Delta t$  foi tomado como sendo igual a 0,1. Para cada imagem processada, o número de iterações é calculado segundo a Equação (5.1). Para imagens de alta resolução (20 px/mm, por exemplo), há o estabelecimento de um número de iterações relativamente grande, o que influencia a performance computacional do processo. Se, por exemplo, aceitarmos deteriorar detalhes da ordem de 0,50 mm em uma imagem cujos pixels possuam tamanho físico de 0,05 mm, o número de iterações será ajustado em 50. Já para uma imagem cujos pixels têm tamanho físico de 0,20 mm, a difusão ocorreria ao longo de 31 iterações.

No que diz respeito à função difusibilidade, usamos a família de curvas dada pela Equação (4.20):

$$f_{\alpha}(\|\vec{\nabla}u\|) = \exp\left(-\frac{1}{\alpha}\left(\frac{\|\vec{\nabla}u\|}{K}\right)^{\alpha}\right), \quad \alpha > 0. \quad (5.2)$$

Por questões metodológicas o parâmetro de curvatura,  $\alpha$ , foi tomado como sendo um número inteiro maior ou igual a 2.

Quanto à forma do cômputo do limiar de separação entre difusão direta e inversa, usamos três variações:

1. o estimador de ruído de Canny, dado pela Equação (4.25):  $K^t = x_L^t = \arg\left\{\sum_{i=1}^L h_i^t \geq 0,9\right\}$ ;
2. o estimador baseado em p-normas, dado pela Equação (4.23):  $K^t = \frac{\|\vec{\nabla}u\|_1^t}{MN}$ ; e
3. um estimador baseado na média global da magnitude do gradiente:

$$K^t = \frac{\|\vec{\nabla}u\|_1^t}{MN} \quad (5.3)$$

A proposta de acrescentar a forma do cômputo de  $K$  segundo a Equação (5.3) ocorreu com vistas a proporcionar uma variabilidade maior sobre este parâmetro dentro da cadeia de processamento.

Para fins de exemplificação, tomemos o fragmento de imagem mamográfica mostrado na Figura 5.10. Este fragmento é inicialmente submetido a um alargamento linear de contraste para que mais estruturas sejam destacadas e operadas pelo enhancement via Equação de Perona-Malik.

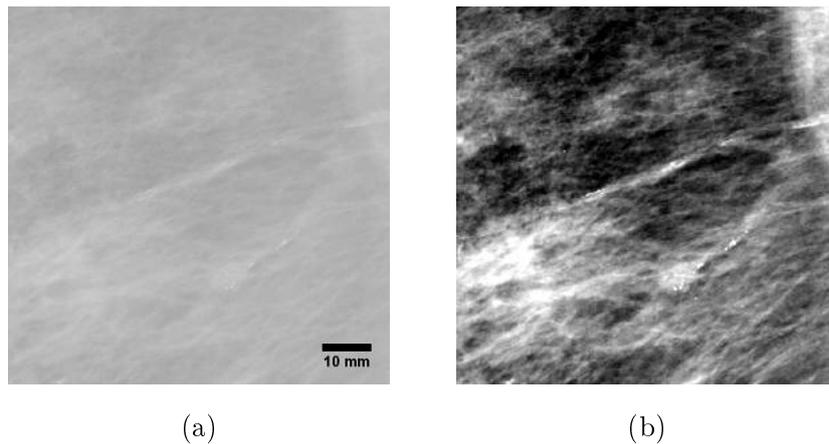


Figura 5.10: Fragmento de imagem mamográfica a ser utilizado na exemplificação dos processos de Enhancement por Perona-Malik e segmentação. Em (b) é mostrada sua versão submetida a alargamento linear de contraste, tal como proposto no início da cadeia de processamento.

Para o processamento deste fragmento, foi utilizado  $\alpha = 2$  e  $K^t$  foi calculado via estimador de ruído de Canny a cada passo de iteração. A degradação máxima tolerada foi de  $\sigma_{\max} = 0,700$  mm e o tamanho físico do pixel da imagem em questão é  $h_{\text{pix}} = 0,150$  mm.

Com isso, o número de iterações, calculado pela Equação (5.1) é igual a 108. A imagem submetida à difusão não-linear pode ser vista na Figura 5.11. Nesta mesma figura é mostrado o resultado da difusão linear ao longo das mesmas 108 iterações.

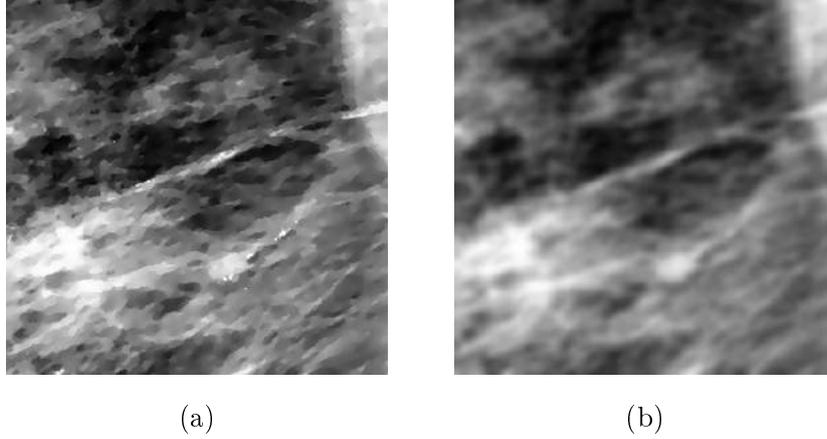


Figura 5.11: Fragmento de mamograma analisado submetido a enhancement via Equação de Perona-Malik (a) e difusão linear (b). Ambas as difusões ocorreram ao longo de 108 iterações. Para a difusão não-linear,  $\alpha = 2$  e o limiar de separação foi estimado pelo estimador de ruído de Canny. Nota-se que, efetivamente, a difusão não-linear conserva bordas e detalhes finos de uma maneira muito mais eficiente que o caso linear.

### Segmentação e pós-processamento

A próxima etapa prevista na cadeia é a segmentação propriamente dita. É nesta etapa, conforme discutido ao longo do trabalho, que há a separação da *informação* a ser analisada por processos de mais alto nível de abstração. É importante ressaltar que, segundo a metodologia proposta, não é o mamograma o sinal a ser segmentado, mas sim a magnitude de seu gradiente. Isto é, embora a saída da etapa de enhancement seja uma versão aperfeiçoada do mamograma, o sinal que é operado na segmentação corresponde às bordas e detalhes finos do sinal. Neste trabalho, as derivadas direcionais da imagem foram estimadas pelo filtro de Sobel, tal como apresentado na Equação (3.28):

$$\mathbf{S}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{S}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}.$$

Aplicados os filtros sobre o sinal de saída da etapa de enhancement, a magnitude do gradiente da imagem,  $\|\vec{\nabla}u\|$ , é dada pela Equação (3.29):

$$\|\vec{\nabla}u\| = \sqrt{\mathbf{S}_x^2[u] + \mathbf{S}_y^2[u]}.$$

Realizada a detecção de bordas via magnitude do gradiente, o sinal resultante deve ser submetido a um alargamento linear de contraste. Esta etapa se justifica da seguinte forma: dado o arranjo discretizado dos dados, mesmo bordas abruptas podem ser representadas de maneira pouco perceptível (baixa intensidade luminosa), devido à natureza discreta

da diferenciação numérica<sup>7</sup>. Em outras palavras, a aplicação do alargamento linear de contraste sobre a magnitude do gradiente visa trazer este sinal à escala dinâmica da imagem inicial. No caso de a magnitude do gradiente já possuir alto contraste, a operação é inócua. Posteriormente ao alargamento de contraste, é realizada a segmentação via limiarização global da magnitude do gradiente. Foram usados três formas de cômputo do threshold: máxima entropia, método de Otsu e valor médio. Na Figura 5.12 podem ser visualizados a magnitude do gradiente do sinal oriundo da etapa de enhancement por Perona-Malik, bem como o resultado de uma binarização via máxima entropia, Otsu e média.

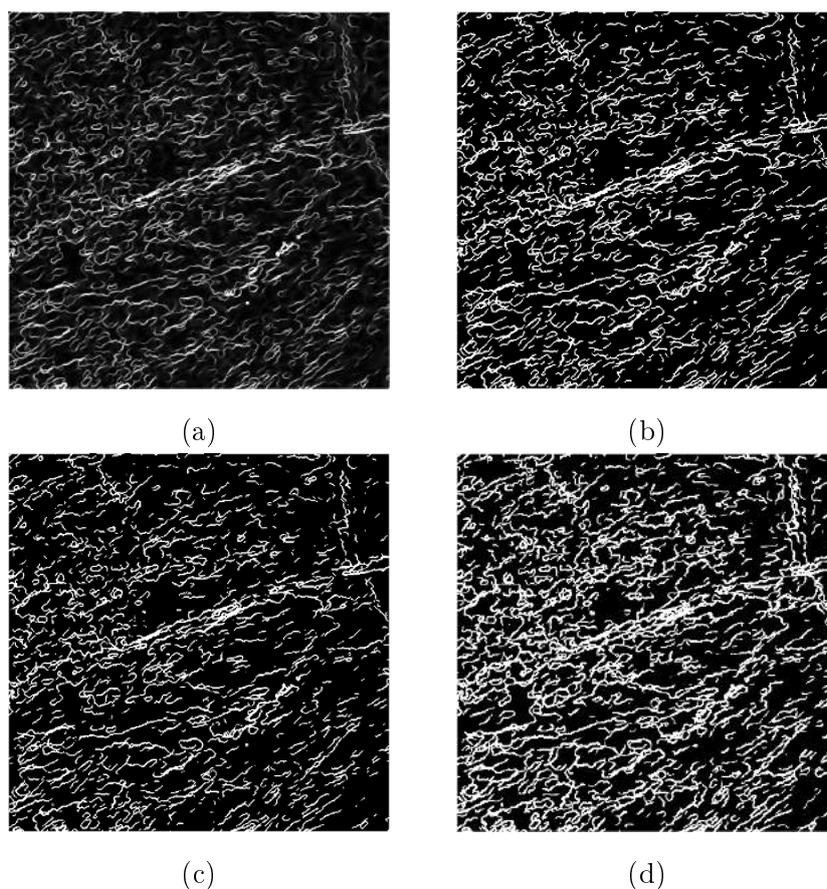


Figura 5.12: Segmentação do fragmento de imagem mamográfica analisado. Em (a) é mostrada a magnitude do gradiente do sinal oriundo da etapa de enhancement via Equação de Perona-Malik. Em (b) é mostrado o resultado da binarização via máxima entropia, em (c) o resultado da binarização pelo método de Otsu e em (d) pelo valor médio.

Uma vez gerado o sinal binário, o processo de segmentação é finalizado com pós-processamento via morfologia matemática, em que “buracos” são preenchidos. Isso se dá com vistas às etapas de alto-nível: quando as estruturas passarem por agrupamento e forem comparadas com parâmetros dimensionais reais, é importante que suas *áreas* sejam

<sup>7</sup>Uma ilustração simples desse fato seria a de, em domínio contínuo, filtrarmos uma borda do tipo função degrau: ao ser submetida a um diferenciador, esta responderia – independentemente de sua altura – segundo uma delta de Dirac. Ou seja, mesmo bordas do tipo degrau com amplitudes arbitrariamente pequenas responderiam impulsivamente a um diferenciador, ao passo que em um cenário discreto a resposta depende fortemente da diferença local.

compatíveis com estes parâmetros. A discussão detalhada sobre o processamento de alto-nível será feita mais adiante. Na Figura 5.13 pode ser visto o resultado do preenchimento morfológico sobre a segmentação anterior do fragmento de mamograma analisado.

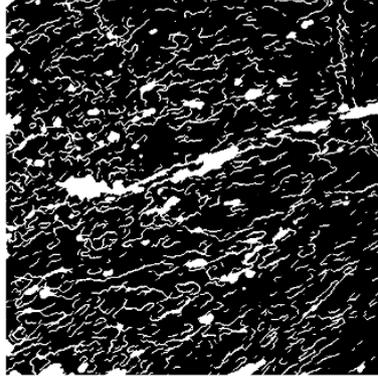


Figura 5.13: Resultado do preenchimento morfológico do resultado da segmentação do fragmento de mamograma analisado. Nota-se claramente o preenchimento da região associada às calcificações ductais no fragmento original da Figura 5.10.

O sinal resultante do fechamento morfológico já possui as informações relevantes para as etapas de caracterização da detecção. Contudo, é aplicado sobre este sinal um filtro de área, tal como descrito na Seção 3.4.2. De posse dos descritores de área dos objetos constituintes da imagem processada até esta etapa, é realizada a remoção de objetos que estejam fora de um intervalo específico. Neste trabalho, adotamos as seguintes áreas mínima e máxima, respectivamente:

$$\begin{cases} A_{\min} = (0,20 \times 0,20) \text{ mm}^2 = 0,04 \text{ mm}^2; \\ A_{\max} = (10,00 \times 10,00) \text{ mm}^2 = 100,00 \text{ mm}^2. \end{cases} \quad (5.4)$$

O estabelecimento destes valores extremos para a filtragem de área se baseia em análise da literatura médica: Segundo Santos [76] e Kopans [47], registra-se em um mamograma a ocorrência de *microcalcificações* se houver um acúmulo de cálcio com extensão (linear) menor que 500  $\mu\text{m}$  (0,500 mm); como *calcificações* se o acúmulo possuir extensão linear entre 0,500 mm e 2,00 mm e como *calcificações grosseiras* as estruturas com extensão igual ou maior que 2,00 mm. Os extremos mostrados na Equação (5.4) abrangem – segundo os valores de referência – desde microcalcificações (limitadas a 0,20 mm de extensão linear) até calcificações grosseiras contidas em um quadrado de lado igual a 1,00 cm. Obviamente, há achados que ultrapassam estes extremos, mesmo sendo classificados como calcificações. Contudo, em termos de processamento, é fundamental limitar as áreas a serem processadas. A Figura 5.14 mostra o resultado da imagem pós-processada por morfologia mostrada na Figura 5.13.

O resultado da filtragem por área mostrado na Figura 5.14 demanda análise mais profunda: à primeira vista, surpreende o fato de as áreas associadas às calcificações ductais serem removidas e diversos segmentos de linha serem preservados; idealmente, o contrário deveria ter ocorrido. Entretanto, é importante notar que o filtro de área corresponde, efetivamente, à primeira etapa de processamento de alto-nível da cadeia de processamento proposta: informações *físicas* (as dimensões toleradas) entram em cena como parâmetros do processamento e sua conversão para o domínio menos abstrato dos pixels traz consigo

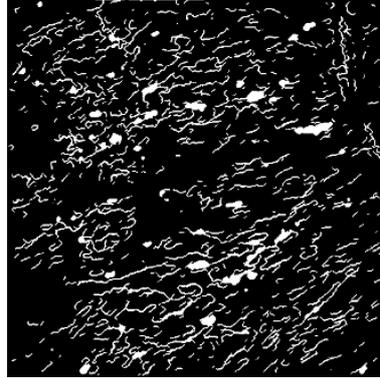


Figura 5.14: Resultado da filtragem por área segundo os valores extremos mostrados na Equação (5.4) do fragmento pós-processado por morfologia mostrado na Figura 5.13.

distorções conceituais que devem ser consideradas. Uma análise atenta sobre a Equação (3.31) revela que, no domínio da imagem, mesmo segmentos de linha possuem área, já que esta grandeza é simplesmente uma soma sobre os índices que identificam o objeto. Isso contraria a intuição, já que encaramos elementos de linha como tendo área nula. Dito isto, esclarece-se a preservação destas estruturas na filtragem mostrada na Figura 5.14.

Uma maneira de minimizar a questão da preservação de segmentos de linha que possuem área dentro dos intervalos especificados é realizar uma abertura morfológica – tal como definido na Equação (3.42):

$$u \circ v = (u \ominus v) \oplus v$$

– sobre o sinal oriundo da filtragem por área. O elemento estruturante  $v$  usado neste trabalho corresponde a  $v = [-2, 2] \times [-2, 2]$ , e pode ser implementado por

$$v = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & \textcircled{1} & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad (5.5)$$

em que o círculo destaca o elemento situado na origem do sistema de coordenadas. Na Figura 5.15 pode ser visualizado o resultado da abertura morfológica sobre o sinal apresentado na Figura 5.14.

### Agrupamento

Findos a segmentação e o pós-processamento, obtém-se o sinal efetivamente segmentado. Este sinal corresponde à *região de interesse detectada pelo processamento* (ROI-P) e é encaminhado para processamento de alto-nível de abstração: agrupamento, contagem e caracterização. O processo de agrupamento tem por finalidade agrupar objetos nas ROIs P e M, tal que achados fragmentados dentro de uma certa região específica sejam unificados e tratados como um único achado pelo processo de contagem e caracterização. Neste trabalho, o espaço de características foi construído sobre os centróides dos objetos presentes nas ROIs, já que o sistema desenvolvido é voltado à detecção de achados, e não para



Figura 5.15: Fragmento segmentado e filtrado por área após abertura morfológica com elemento estruturante  $v = [-2, 2] \times [-2, 2]$ . Quando comparado com a Figura 5.14, fica evidente a eficiência desta operação morfológica no que diz respeito à remoção dos segmentos de linha outrora preservados.

análise de suas formas. O algoritmo usado para o agrupamento corresponde ao hierárquico aglomerativo, analisado na Seção 3.4.3. Dentro desse contexto, o representante de grupo de cada cluster corresponde ao seu centróide<sup>8</sup> e a distância entre grupos é calculada via distância Euclidiana, tal como mostrado nas Equações (3.37) e (3.38):

$$d(\mathcal{A}, \mathcal{B}) = \|\mathcal{C}(\mathcal{A}) - \mathcal{C}(\mathcal{B})\|_2 \quad \text{e} \quad \mathcal{C}(\mathcal{X}) = \frac{1}{n} \cdot \left( \sum_{i=1}^n x_{i,1}, \sum_{i=1}^n x_{i,2}, \dots, \sum_{i=1}^n x_{i,k} \right),$$

respectivamente. A Figura 5.16 apresenta uma representação do espaço de características da ROI-P mostrada na Figura 5.15. Por sua vez, a ROI-M associada ao fragmento analisado, bem como seu espaço de características são mostrados na Figura 5.17.

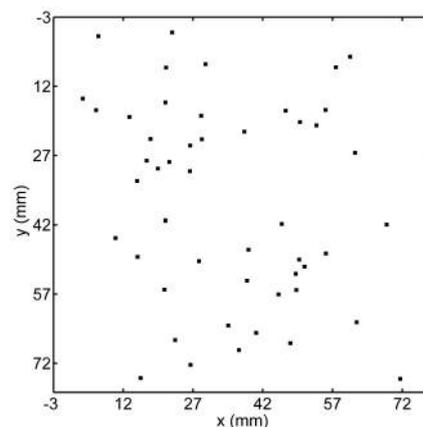


Figura 5.16: Espaço de características da ROI-P mostrada na Figura 5.15, composto pelos centróides das estruturas nela representadas.

Conforme discutido na Seção 3.4.3, é preciso definir um critério de parada para o algoritmo hierárquico aglomerativo. Neste trabalho adotamos a distância máxima entre

<sup>8</sup>Atente-se para o fato de que usar o centróide dos objetos da imagem como vetor de características não implica o uso do centróide *dos clusters* como representante de grupo no algoritmo aglomerativo. Esta última escolha ainda está associada à natureza das detecções: nos parece razoável usar o centro de massa de um conjunto de achados recém-agrupados para analisar os agrupamentos posteriores.

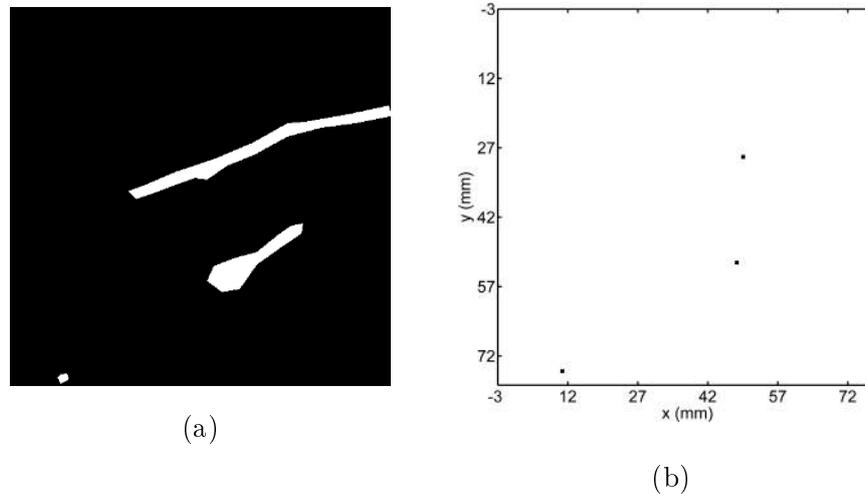


Figura 5.17: ROI-M associada ao fragmento de mamograma mostrado na Figura 5.10 e seu espaço de características, obtido pelos centros de massa das formas presentes nas demarcações do médico.

clusters como sendo igual a

$$d_{\max} = 10\sqrt{2}\text{mm}. \quad (5.6)$$

Esta medida corresponde à diagonal do quadrado definido na literatura médica como sendo a região de abrangência de um conjunto de calcificações tal que este corresponda a um agrupamento, como apontado na Seção 5.2.1, Equação (5.4).

Uma vez estabelecidos os parâmetros do agrupamento, as duas ROIs são clusterizadas. Em sequência, os objetos pertencentes a um mesmo cluster são transformados em uma única forma retangular; isto é, cada cluster é transformado em um retângulo cujas dimensões abrangem toda a extensão dos objetos agrupados. Estas *áreas* correspondem à informação de mais alto-nível do sistema: elas são entregues a uma etapa de contagem e caracterização, que por sua vez, apresentará os indicadores da detecção. A Figura 5.18 ilustra esse procedimento.

### 5.2.2 Contagem e caracterização dos achados

Conforme visto na última seção, as ROIs P e M passam por um processo de agrupamento e transformação, em que cada cluster é representado por sua área de abrangência. Nesse sentido, cada achado P e M é *descrito* agora por um vetor de três componentes, sendo as duas primeiras as coordenadas de seu centro de massa e a terceira o seu valor de área<sup>9</sup>. A lista destes vetores compõe as tabelas de achados-P e achados-M, a serem analisadas pela cadeia de contagem e caracterização. As grandezas de interesse para quantificação das detecções podem ser visualizadas na Tabela 5.2. A seguir, será feita uma descrição detalhada destas grandezas.

<sup>9</sup>Cabe reforçar: os centros de massa e as áreas dos achados correspondem às áreas resultantes de todo o procedimento até este ponto. As ROIs P e M originais não são usadas diretamente para processamento de alto-nível.

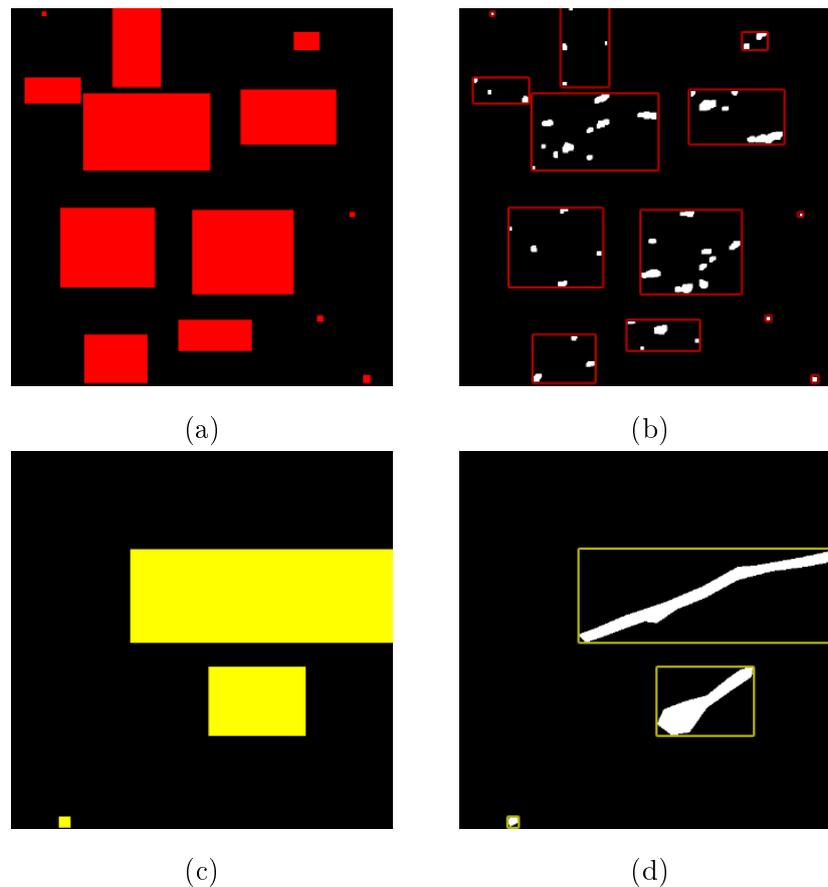


Figura 5.18: Ilustração do processo de agrupamento das ROIs segundo a cadeia proposta: em (a) é mostrada a ROI-P clusterizada e em (b) é mostrada a superposição das bordas deste sinal sobre a ROI-P. Em (c) é mostrada a ROI-M clusterizada e em (d) é mostrada a superposição deste sinal sobre a ROI-M original. Embora o processo de agrupamento ocorra sobre os centros de massa dos objetos, os sinais gerados levam em consideração os retângulos envolventes dos objetos pertencentes a cada cluster. O sinal resultante desse processo de agrupamento é encaminhado para a etapa de contagem e caracterização. As cores vermelho e amarelo foram usadas neste trabalho para destacar achados P e M, respectivamente.

Símbolo/Abreviação	Descrição
$m$	Número de achados na ROI-M
$n$	Número de achados na ROI-P
$N_{VP}$	Número de Verdadeiros-Positivos
$N_{FP}$	Número de Falsos-Positivos
$N_{FN}$	Número de Falsos-Negativos
<b>DPM</b>	Matriz de distâncias entre achados-P e achados-M
<b>MCA</b>	Matriz de medidas de comparação entre áreas
<b>MPF</b>	Matriz de medidas de posição e forma

Tabela 5.2: Grandezas de interesse na quantificação das detecções.

### Matrizes DPM e $DPM_{\log}$

De posse das coordenadas dos achados P e M, faz-se necessário computar as distâncias entre grupos destes achados, isto é, a distância entre cada par de achados P e M. Estas

informações são obtidas através da matriz **DPM**, definida como

$$\mathbf{DPM} = \begin{bmatrix} d(P_1, M_1) & d(P_2, M_1) & \dots & d(P_n, M_1) \\ d(P_1, M_2) & d(P_2, M_2) & \dots & d(P_n, M_2) \\ \vdots & \vdots & \ddots & \vdots \\ d(P_1, M_m) & d(P_2, M_m) & \dots & d(P_n, M_m) \end{bmatrix}, \quad (5.7)$$

em que  $d(\cdot)$  corresponde a qualquer forma de cômputo de distância. Nesta etapa, fizemos novamente uso da distância Euclidiana,  $d(a, b) = \|a - b\|_2$ . Obviamente, **DPM** existirá apenas se  $m, n \neq 0$ , isto é, apenas se ambos o médico e o processamento detectarem algo nas imagens.

Uma vez construída a matriz **DPM**, é feita uma verificação de quais distâncias estão dentro de um certo *raio de abrangência*. Esta verificação ocorre com vistas a identificar quais pares de achados (P,M) estão próximos suficiente para que análises posteriores sejam conduzidas. A verificação dessa proximidade dá origem à *matriz lógica de distâncias entre achados-P e achados-M*, **DPM<sub>log</sub>**, definida como

$$\mathbf{DPM}_{\log ij} = \alpha_{ij} = \begin{cases} 1 & \text{se } \mathbf{DPM}_{ij} \leq d_{\max}; \\ 0 & \text{caso contrário.} \end{cases} \quad (5.8)$$

Na Equação (5.8),  $d_{\max}$  corresponde ao mesmo valor adotado para o processo de agrupamento dos achados das ROIs P e M, dado pela Equação (5.6):  $d_{\max} = 10\sqrt{2}$  mm.

### Matrizes MCA e MPF

As matrizes **DPM** e **DPM<sub>log</sub>** carregam informações sobre as distâncias entre os achados P e M, bem como sobre quais achados dos dois grupos estão próximos o suficiente para que possam ser encarados como detecções do *mesmo achado*. O passo seguinte na avaliação destes achados é verificar se as áreas são compatíveis com a área deste último. Ou seja, é preciso verificar se um ou mais achados-P que estejam dentro do raio de abrangência de um certo achado-M possuem áreas comparáveis com a área deste último. Esta análise é inicialmente proporcionada pela definição da *matriz de medidas de comparação entre áreas*, **MCA**:

$$\mathbf{MCA} = \begin{bmatrix} \frac{A(P_1)}{A(M_1)} & \frac{A(P_2)}{A(M_1)} & \dots & \frac{A(P_n)}{A(M_1)} \\ \frac{A(P_1)}{A(M_2)} & \frac{A(P_2)}{A(M_2)} & \dots & \frac{A(P_n)}{A(M_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A(P_1)}{A(M_m)} & \frac{A(P_2)}{A(M_m)} & \dots & \frac{A(P_n)}{A(M_m)} \end{bmatrix}, \quad (5.9)$$

em que  $A(\cdot)$  corresponde à área de um achado, tal como definido na Equação (3.31).

A matriz **MCA**, tal como definida na Equação (5.9), proporciona comparação entre todos os pares de achados (P,M). Contudo, nem todas essas comparações são úteis, visto que há achados que estão fora do raio de abrangência e que, portanto, não devem ser considerados na avaliação quantitativa de um achado Verdadeiro-Positivo. Para que apenas os achados dentro do raio de abrangência sejam avaliados, definimos a *matriz de medidas de posição e forma*, **MPF**:

$$\mathbf{MPF}_{ij} = \mathbf{DPM}_{\log ij} \cdot \mathbf{MCA}_{ij}, \quad (5.10)$$

ou seja,

$$\mathbf{MPF} = \begin{bmatrix} \alpha_{11} \frac{A(P_1)}{A(M_1)} & \alpha_{12} \frac{A(P_2)}{A(M_1)} & \dots & \alpha_{1n} \frac{A(P_n)}{A(M_1)} \\ \alpha_{21} \frac{A(P_1)}{A(M_2)} & \alpha_{22} \frac{A(P_2)}{A(M_2)} & \dots & \alpha_{2n} \frac{A(P_n)}{A(M_2)} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} \frac{A(P_1)}{A(M_m)} & \alpha_{m2} \frac{A(P_2)}{A(M_m)} & \dots & \alpha_{mn} \frac{A(P_n)}{A(M_m)} \end{bmatrix}. \quad (5.11)$$

Tal como definida nas Equações (5.10) e (5.11)<sup>10</sup>, vemos que a matriz **MPF** carrega informações sobre *quais* pares de achados (P,M) devem ser analisados e *quão* semelhantes eles são. A partir desta grandeza, várias outras são obtidas. A seguir serão apresentados os meios usados para contagem de  $N_{VP}$ ,  $N_{FP}$  e  $N_{FN}$ .

### Contagem dos Verdadeiros-Positivos

Para que possamos contabilizar um ou mais achados-P como um verdadeiro-positivo (VP) associados a um mesmo achado-M, os seguintes requisitos devem ser atendidos:

*O conjunto de achados-P deve estar dentro do raio de abrangência em relação a um determinado achado-M e a soma de suas áreas deve ser maior que uma certa fração da área do achado-M em questão.*

Isto pode ser verificado por meio das linhas da matriz **MPF**: se a soma da *i*-ésima linha da matriz **MPF** for maior ou igual a um certo limiar  $\chi > 0$ , isto é,

$$\left( \sum_{j=1}^n \alpha_{ij} \frac{A(P_j)}{A(M_i)} \right) \geq \chi, \quad (5.12)$$

então o *i*-ésimo achado-M foi detectado. Portanto,  $N_{VP}$  deve ser incrementado. Especificamente, dado o conjunto  $M_d$  das linhas de **MPF** que atendem a Equação (5.12),

$$M_d = \left\{ \left[ \text{MPF}_{k1} \quad \text{MPF}_{k2} \quad \dots \quad \text{MPF}_{kn} \right]; k \in I_{VP} \subseteq \{1, 2, \dots, m\} \left| \left( \sum_{j=1}^n \text{MPF}_{kj} \right) \geq \chi \right. \right\}, \quad (5.13)$$

definimos o número de verdadeiros-positivos,  $N_{VP}$  como sendo o número de elementos de  $M_d$ <sup>11</sup>. Na Equação (5.13),  $I_{VP}$  é um conjunto de índices contido em  $\{1, 2, \dots, m\}$ .

Para finalizar a discussão sobre o cômputo de  $N_{VP}$ , é importante ressaltar o papel do limiar  $\chi$ : se seu valor for muito grande, apenas os achados muito semelhantes serão computados, o que levará inevitavelmente a uma redução de  $N_{VP}$ ; por outro lado, valores muito pequenos para  $\chi$  permitirão que ruídos (detecções espúrias) sejam registrados como verdadeiros-positivos, tornando o processo qualitativamente pior. Portanto, devemos encarar  $\chi$  como um parâmetro da cadeia de processamento.

<sup>10</sup>Deve-se atentar que o produto apresentado na Equação (5.10) é um produto entrada-a-entrada, diferindo completamente do produto matricial usual. O produto entrada-a-entrada entre duas matrizes de dimensões idênticas é chamado *produto de Hadamard* e goza de propriedades algébricas próprias [50, 4]

<sup>11</sup>Rigorosamente, temos que  $N_{VP}$  corresponde à *cardinalidade* do conjunto  $M_d$ .

### Contagem dos Falsos-Positivos e Falsos-Negativos

Um falso-positivo (FP) corresponde a um achado-P sem achado-M correspondente, isto é, um achado que não é comparado com nenhum outro. Segundo a construção das matrizes  $\mathbf{DPM}$  e  $\mathbf{DPM}_{\log}$ , *um achado-P que não está dentro do raio de abrangência de nenhum achado-M corresponde a um falso-positivo*. Especificamente, temos que o número de falsos-positivos,  $N_{FP}$ , é dado pela quantidade de colunas completamente nulas na matriz  $\mathbf{DPM}_{\log}$ . Formalmente, temos: dado o conjunto  $P_e$  das colunas nulas da matriz  $\mathbf{DPM}_{\log}$ ,

$$P_e = \left\{ \left[ \begin{array}{c} \alpha_{1l} \\ \alpha_{2l} \\ \vdots \\ \alpha_{ml} \end{array} \right]; l \in I_{FP} \subseteq \{1, 2, \dots, n\} \mid \alpha_{il} = 0 \quad \forall i \in \{1, 2, \dots, m\} \right\}, \quad (5.14)$$

definimos  $N_{FP}$  como sendo o número de elementos de  $P_e$ . Na Equação (5.14),  $I_{FP}$  é um conjunto de índices contido em  $\{1, 2, \dots, n\}$ .

O cômputo do número de falsos-negativos  $N_{FN}$  é obtido de maneira análoga, pois *um achado-M que não possui achado-P associado corresponde a um falso-negativo*. Em termos práticos,  $N_{FN}$  é computado como sendo o número de linhas completamente nulas da matriz  $\mathbf{DPM}_{\log}$ . Ou seja, dado o conjunto  $M_{nd}$  das linhas nulas da matriz  $\mathbf{DPM}_{\log}$ ,

$$M_{nd} = \left\{ \left[ \alpha_{k1} \quad \alpha_{k2} \quad \dots \quad \alpha_{kn} \right]; k \in I_{FN} \subseteq \{1, 2, \dots, m\} \mid \alpha_{kj} = 0 \quad \forall j \in \{1, 2, \dots, n\} \right\}, \quad (5.15)$$

definimos  $N_{FN}$  como sendo o número de elementos de  $M_{nd}$ . Na Equação (5.15),  $I_{FN}$  é um conjunto de índices contido em  $\{1, 2, \dots, m\}$ .

É importante notar que embora a metodologia desenvolvida permita a imediata estimativa de  $N_{FP}$  e  $N_{FN}$ , um problema ocorre quando o médico ou o processamento não detectam algo ( $m$  ou  $n$  nulo). No caso em que não há achados-M mas há achados-P ( $m = 0$ ,  $n \neq 0$ ), o valor de  $N_{FP}$  é igual a  $n$  (todas as detecções são espúrias). No caso em que há achados-M e não há achados-P ( $m \neq 0$  e  $n = 0$ ), temos que  $N_{FN} = m$ , pois todos os achados-M deveriam ter sido detectados mas não o foram. Com vistas a sintetizar todo o processo de contagem e caracterização, uma estrutura algorítmica para a quantificação dos achados pode ser visualizada na Figura 5.19.

Para facilitar a compreensão da metodologia desenvolvida para quantificação dos achados da cadeia de processamento, serão apresentados a seguir alguns exemplos baseados na ROI-M mostrada na Figura 5.20. Os descritores de interesse para essa ROI podem ser consultados na Tabela 5.3.

**Exemplo 1.** *Detecções perfeitas.* Supondo que a cadeia de processamento tenha detectado perfeitamente todos os achados da ROI-M em questão, temos nesse caso uma ROI-P idêntica à descrita na Tabela 5.3. Usando  $d_{\max} = 40$  px, temos

$$\mathbf{DPM} = \begin{bmatrix} 0,00 \cdot 10^0 & 2,00 \cdot 10^2 & 8,94 \cdot 10^2 & 1,44 \cdot 10^3 \\ 2,00 \cdot 10^2 & 0,00 \cdot 10^0 & 7,21 \cdot 10^2 & 1,28 \cdot 10^3 \\ 8,94 \cdot 10^2 & 7,21 \cdot 10^2 & 0,00 \cdot 10^0 & 5,66 \cdot 10^2 \\ 1,44 \cdot 10^3 & 1,28 \cdot 10^3 & 5,66 \cdot 10^2 & 0,00 \cdot 10^0 \end{bmatrix} \quad \text{e} \quad \mathbf{DPM}_{\log} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

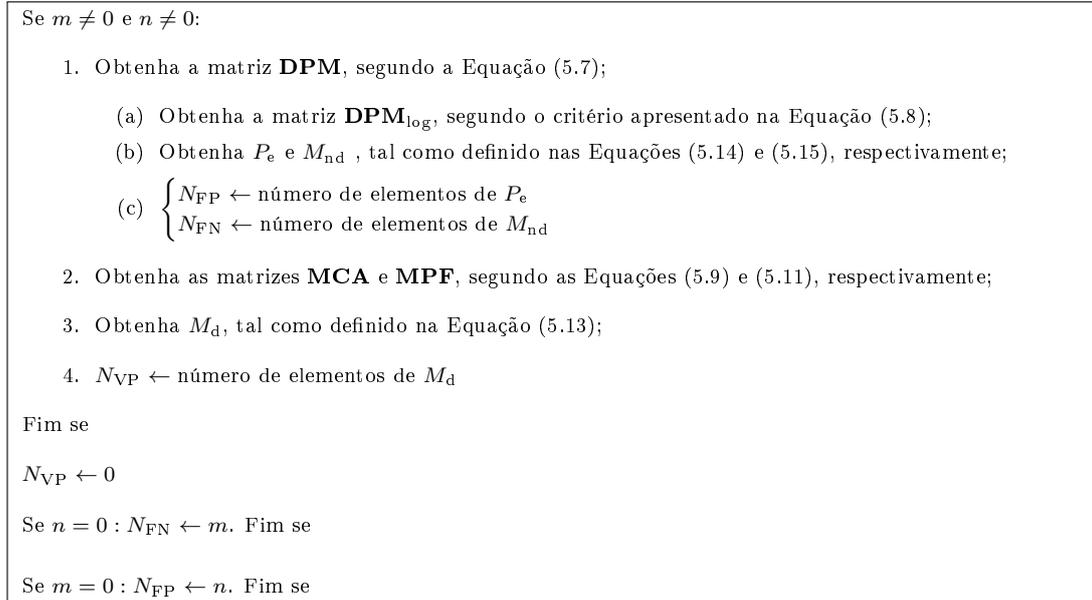
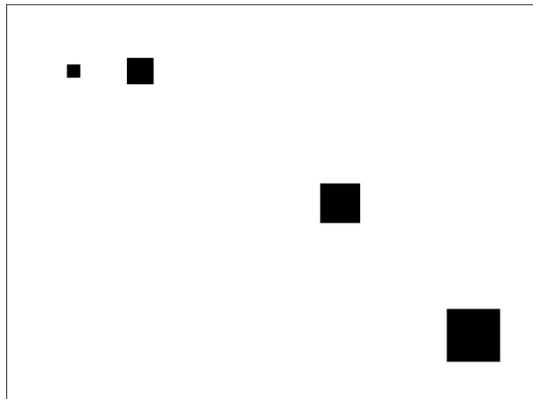


Figura 5.19: Estrutura do algoritmo para quantificação dos achados da cadeia de processamento.



ROI-M			
ID	$y_{CM}$ (px)	$x_{CM}$ (px)	$A$ (px)
1	$2,00 \cdot 10^2$	$2,00 \cdot 10^2$	$1,60 \cdot 10^3$
2	$4,00 \cdot 10^2$	$2,00 \cdot 10^2$	$6,40 \cdot 10^3$
3	$1,00 \cdot 10^3$	$6,00 \cdot 10^2$	$1,44 \cdot 10^3$
4	$1,40 \cdot 10^2$	$1,00 \cdot 10^2$	$2,56 \cdot 10^3$

Tabela 5.3: Descritores de interesse para a ROI-M apresentada na Figura 5.20.

Figura 5.20: ROI-M sintética a ser usada nos exemplos de quantificação dos resultados das detecções.

A matriz **MCA** fica

$$\mathbf{MCA} = \begin{bmatrix} 1,00 \cdot 10^0 & 4,00 \cdot 10^0 & 9,00 \cdot 10^0 & 1,60 \cdot 10^1 \\ 2,50 \cdot 10^{-1} & 1,00 \cdot 10^0 & 2,25 \cdot 10^0 & 4,00 \cdot 10^0 \\ 1,11 \cdot 10^{-1} & 4,44 \cdot 10^{-1} & 1,00 \cdot 10^0 & 1,78 \cdot 10^0 \\ 6,25 \cdot 10^{-2} & 2,50 \cdot 10^{-1} & 5,63 \cdot 10^{-1} & 1,00 \cdot 10^0 \end{bmatrix}.$$

Com isso, obtemos finalmente a matriz **MPF** :

$$\mathbf{MPF} = 1 \cdot 10^0 \begin{bmatrix} 1,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 1,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 1,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 1,00 \end{bmatrix}.$$

Segundo a metodologia proposta, devemos identificar os conjuntos  $P_e$ ,  $M_{nd}$  e  $M_d$ . Efetivamente, devemos obter o número de elementos destes conjuntos. Fixando um valor de  $\chi$  ( $\chi = 0, 10$ , por exemplo), obtemos:

$$\begin{cases} I_{FP} = \emptyset \\ I_{FN} = \emptyset \\ I_{VP} = \{1, 2, 3, 4\} \end{cases} \Rightarrow \begin{cases} N_{FP} = 0 \\ N_{FN} = 0 \\ N_{VP} = 4 \end{cases} .$$

Naturalmente, se as detecções foram perfeitas, o resultado não poderia ser diferente: há quatro achados-M detectados corretamente; portanto, há quatro VPs e nenhum FP ou FN.

**Exemplo 2.** *Detecções espúrias: falsos-positivos.* Neste exemplo, ilustraremos o uso da metodologia proposta em uma situação em que há mais achados-P do que achados-M e estes achados-P estão fora do raio de abrangência de quaisquer achados reais. A ROI-P sintética a ser comparada com a ROI-M mostrada na Figura 5.20 e sua tabela correspondente podem ser visualizadas na Figura 5.21 e na Tabela 5.4, respectivamente. De posse das

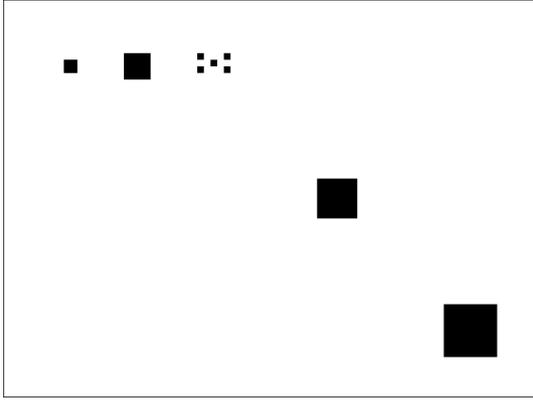


Figura 5.21: ROI-P sintética contendo FPs quando comparada com a ROI-M retratada na Figura 5.20.

ROI-P contendo FPs			
ID	$y_{CM}$ (px)	$x_{CM}$ (px)	$A$ (px)
1	$2,01 \cdot 10^2$	$2,01 \cdot 10^2$	$1,60 \cdot 10^3$
2	$4,01 \cdot 10^2$	$2,01 \cdot 10^2$	$6,40 \cdot 10^3$
3	$5,91 \cdot 10^2$	$1,71 \cdot 10^2$	$4,00 \cdot 10^2$
4	$5,91 \cdot 10^2$	$2,11 \cdot 10^2$	$4,00 \cdot 10^2$
5	$6,31 \cdot 10^2$	$1,91 \cdot 10^2$	$4,00 \cdot 10^2$
6	$6,71 \cdot 10^2$	$1,71 \cdot 10^2$	$4,00 \cdot 10^2$
7	$6,71 \cdot 10^2$	$2,11 \cdot 10^2$	$4,00 \cdot 10^2$
8	$1,00 \cdot 10^3$	$6,01 \cdot 10^2$	$1,44 \cdot 10^4$
9	$1,40 \cdot 10^3$	$1,00 \cdot 10^3$	$2,56 \cdot 10^4$

Tabela 5.4: Descritores de interesse para a ROI-P apresentada na Figura 5.21.

ROIs, temos:

$$\mathbf{MPF} = 1 \cdot 10^0 \begin{bmatrix} 1,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 1,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 1,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 1,00 \end{bmatrix} .$$

Com isso, obtemos:

$$\begin{cases} I_{FP} = \{3, 4, 5, 6, 7\} \\ I_{FN} = \emptyset \\ I_{VP} = \{1, 2, 8, 9\} \end{cases} \Rightarrow \begin{cases} N_{FP} = 5 \\ N_{FN} = 0 \\ N_{VP} = 4 \end{cases} .$$

**Exemplo 3.** *Miscelânea: falsos-positivos e falsos-negativos na mesma imagem.* Neste exemplo é ilustrado o caso em que um grupo de achados-M não possui achados-P correspondentes (caracterizando ocorrência de FNs) em adição à presença de detecções espúrias



Com isso, temos:

$$\begin{cases} I_{FP} = \{1, 2, 3, 4, 5, 6, 8, 10, 11, 13\} \\ I_{FN} = \emptyset \\ I_{VP} = \{7, 9, 12\} \end{cases} \Rightarrow \begin{cases} N_{FP} = 10 \\ N_{FN} = 0 \\ N_{VP} = 3 \end{cases} .$$

O resultado do processamento pode ser visualizado na Figura 5.23.

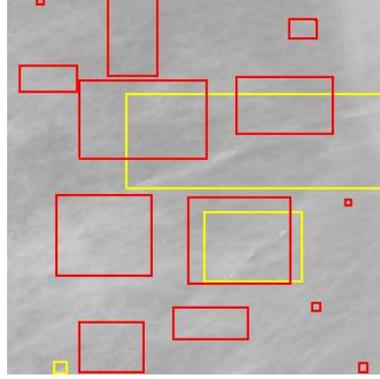


Figura 5.23: Resultado do processamento segundo a cadeia proposta sobre o fragmento de imagem mamográfica mostrado na Figura 5.10. As detecções do processamento são mostradas em vermelho e a ROI-M é mostrada, para comparação, em amarelo. Nota-se a efetividade do processamento na detecção do agrupamento de microcalcificações presente na imagem. A calcificação ductal foi parcialmente detectada e uma microcalcificação isolada no canto inferior esquerdo não foi diretamente detectada. Contudo, de acordo com os parâmetros usados, os três achados-M (calcificações ductais, cluster de microcalcificações e microcalcificação isolada) foram detectados. Nota-se também a forte presença de detecções espúrias, configurando falsos-positivo.

Verificamos por meio dos exemplos que a metodologia proposta é efetiva quanto ao levantamento de  $N_{VP}$ ,  $N_{FP}$  e  $N_{FN}$ , mesmo em situações em que não ocorrem detecções. Obviamente, a introdução do limiar  $\chi$  gera mais um parâmetro na cadeia de processamento desenvolvida, mas se faz necessária para que achados sejam classificados. O estabelecimento deste parâmetro influencia diretamente os indicadores do processamento, e deve ocorrer de maneira a não tornar o processo extremamente sensível a ruídos, mas ainda assim capaz de considerar achados significativos. Neste trabalho esse limiar foi tomado sempre como

$$\chi = 0, 10. \quad (5.16)$$

### Resumo: parâmetros da cadeia de processamento

Conforme descrito ao longo desta seção, a cadeia de processamento desenvolvida neste trabalho possui diversas instâncias que operam em diferentes níveis de abstração. Há de se destacar que buscamos – tanto quanto possível – a automação dos processos constituintes desta cadeia. Contudo, foi preciso utilizar algum conhecimento *a priori*, como por exemplo a distância  $d_{\max}$ . Em adição, o pós-processamento envolve filtragem por áreas e abertura morfológica, o que demanda reflexão sobre as áreas permitidas e sobre os elementos estruturantes usados. Também demanda análise a escolha do centróide como representante de grupo nos processos de agrupamento, já que houve motivação física, mas esta escolha não corresponde à única disponível para tal procedimento. Por fim, o enhancement via Equação de Perona-Malik guarda forte dependência sobre seus parâmetros.

---

Com base nesse cenário, foi preciso restringir vários dos parâmetros usados na cadeia, de forma que pudéssemos investigar mais a fundo a etapa a que nos propusemos. Em síntese, deve-se compreender que as várias etapas da cadeia de processamento desenvolvida operam de maneira praticamente independente umas das outras, o que demandaria milhares de arranjos de configuração distintos. Do ponto de vista metodológico, analisar todos os arranjos de configuração possíveis não seria possível em termos de computação e de análise. Mais importante é que essa análise generalizada descaracterizaria o foco do trabalho, que corresponde a quantificar a influência da etapa de enhancement via Equação de Perona-Malik sobre os indicadores da detecção.

## Capítulo 6

# Resultados

### 6.1 Efetividade da técnica de segmentação das ROIs

A primeira verificação realizada sobre a cadeia de processamento desenvolvida ocorreu sobre a segmentação das ROIs. Obviamente, se esta etapa não funcionar bem, todo o restante do processo será comprometido. Nesse sentido, a metodologia de segmentação das ROIs descrita na Seção 5.2.1 foi aplicada sobre duas bases, BIM e Bancoweb, para que se pudesse verificar se as ROIs estavam sendo segmentadas corretamente. Entenda-se por segmentação correta da ROI a obtenção de um sinal que contenha apenas a mama, mesmo que com seus *contornos externos ligeiramente deteriorados*. Essa tolerância à degradação dos contornos externos da mama guarda certo grau de subjetividade, já que o mamograma retrata naturalmente os contornos da mama com baixo contraste, devido à configuração do exame de mamografia, tal como mostrado na Seção 2.1.

Inicialmente, a técnica desenvolvida foi aplicada sobre um conjunto de 472 imagens do Bancoweb, apresentando 14 imagens caracterizadas como erro – aproximadamente 3% do total. Estes erros englobam a escolha da região errada do mamograma, a presença de artefatos alheios à mama e a subtração de parte significativa do tecido mamário da imagem. Do ponto de vista da cadeia de processamento, a escolha da região errada corresponde ao pior tipo de erro possível, uma vez que o sinal operado não terá relação alguma com a mama. Embora indesejáveis, os outros dois tipos de erro são menos graves. Na Figura 6.1 podem ser visualizados exemplos destas ocorrências.

Em sequência, a técnica foi aplicada sobre as 51 imagens que compõem a BIM, apresentando 3 erros – aproximadamente 6% do total. Diferentemente do Bancoweb, os erros ocorridos na BIM foram todos relativos à remoção de parte significativa da mama, tal como mostrado na Figura 6.2. Há de se ressaltar que os três erros ocorreram em imagens CR.

Os erros por escolha da região errada do mamograma ocorreram sobre imagens com áreas não sensibilizadas, que aparecem como platôs de muito alta intensidade luminosa na imagem final (vide Figura 6.1 (a)). Portanto, esta é uma característica praticamente exclusiva da montagem dos cassetes nos sistemas tela-filme ou CR (embora só tenha ocorrido em imagens tela-filme). Por sua vez, os erros por remoção de parte significativa da mama ocorreram todos em mamas com fortes diferenças de contraste nas regiões suprimidas. Isto se explica quando se leva em consideração a heterogeneidade da mama: não raro lidamos com imagens com regiões extremamente radiotransparentes adjacentes a regiões densas. Do ponto de vista da técnica, essas transições acabaram funcionando ou como mínimos detectados nas projeções ou como regiões de transição, suavizadas pelo passa-baixas Gaussiano. Por fim, o erro relativo à manutenção de artefatos alheios à mama tem origem

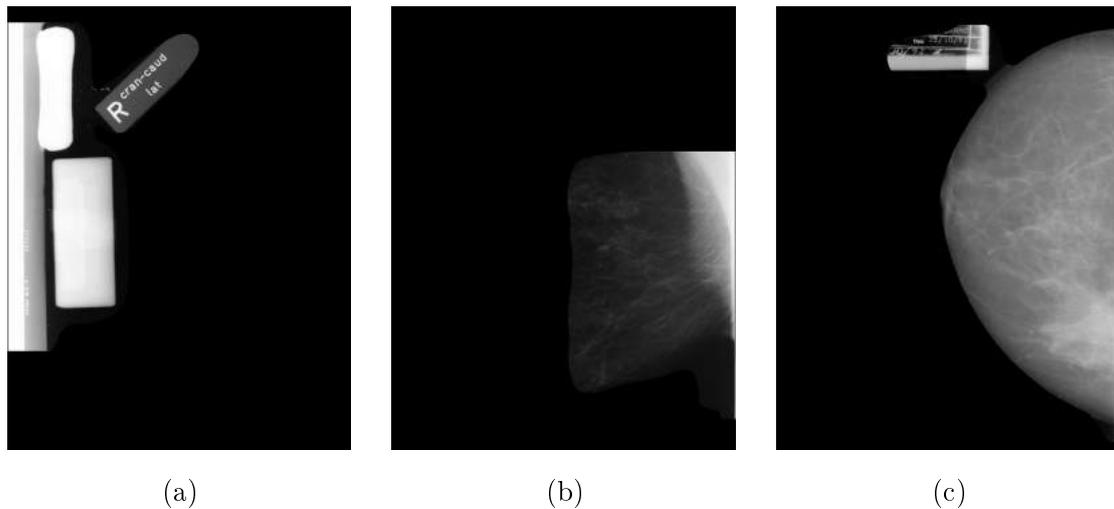


Figura 6.1: Exemplos de erros ocorridos na segmentação da ROI. Em (a) é mostrado o erro mais grave, que corresponde à segmentação da região errada do mamograma. Em (b) é mostrado um caso em que parte significativa da mama foi removida e em (c) é mostrado um caso em que artefatos alheios à mama permanecem na imagem.

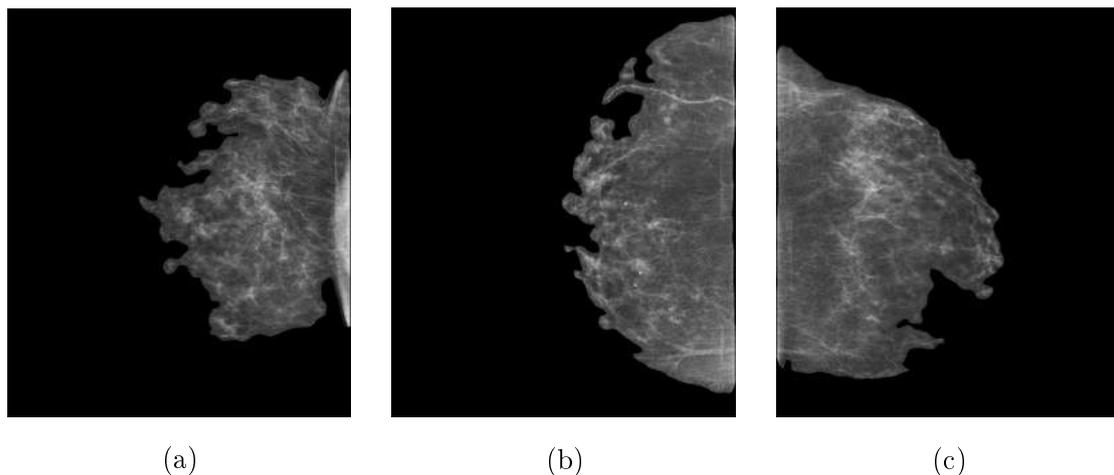


Figura 6.2: Exemplos de erros ocorridos na segmentação da ROI sobre as imagens da BIM. Os únicos três erros ocorreram sobre imagens de modalidade CR.

prática: os técnicos em radiologia que operaram na aquisição das imagens anexaram os identificadores muito próximos das mamas. Essa proximidade levou a uma conexão destes objetos com a mama, devido ao forte passa-baixas ao qual as imagens foram submetidos dentro da técnica. Ou seja, do ponto de vista de informação de alto-nível, os artefatos e a mama constituem o mesmo objeto. Note que este problema inexiste em imagens CR, já que as informações ficam armazenadas no próprio arquivo ou há um controle muito maior de onde o texto descritivo será mostrado na imagem.

De um modo geral, a segmentação das ROIs segundo a técnica desenvolvida mostrou-se muito robusta, operando com êxito em imagens muito distintas umas das outras. Os percentuais de acerto, tanto para o Bancoweb quanto para a BIM ficaram perto de 100% (97% e 94%, respectivamente). Esse percentuais podem ser melhorados ainda mais se houver boa conduta dos técnicos ao fazer a aquisição das imagens, pois isso minimizaria

os erros por artefato e escolha do lado errado do mamograma. Os erros de supressão de regiões da mama merecem mais atenção, mas um primeiro passo para minimizá-los seria reduzir o raio da Gaussiana usada na suavização juntamente com uma transformação em níveis de cinza para aumentar a intensidade nas regiões de transição.

## 6.2 Mascaramento da ROI pelo preenchimento morfológico

Ao longo dos testes, chamou a atenção o fato de várias das imagens testadas quase não apresentarem detecções, independentemente da qualidade destas. Uma análise passo-a-passo da cadeia desenvolvida acabou evidenciando uma característica indesejável da etapa de pós-processamento via morfologia matemática. A etapa de preenchimento, que foi incorporada para preencher pequenas estruturas e otimizar as detecções, pode levar ao preenchimento de toda a região compreendida pela mama (ROI-2). A Figura 6.3 ilustra essa ocorrência.

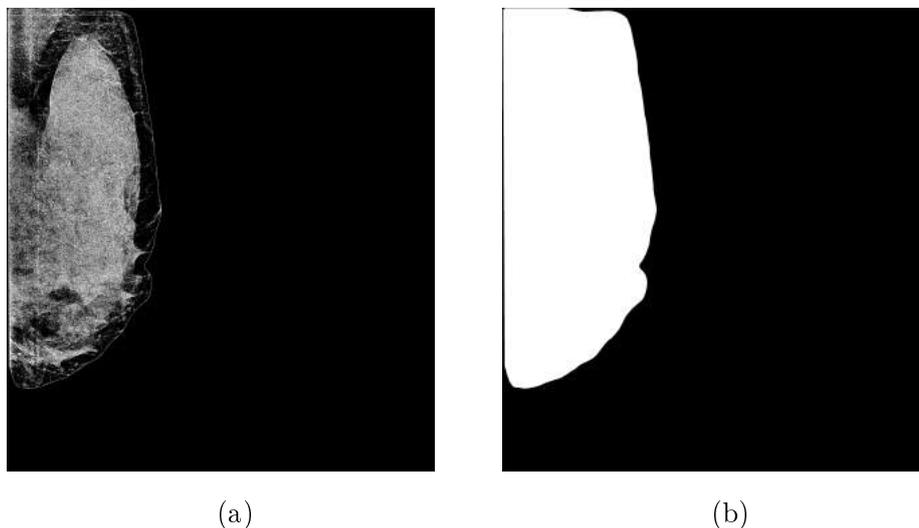


Figura 6.3: Ilustração do mascaramento da ROI: em (a) é mostrada a ROI-2 já segmentada de uma das imagens da BIM. Em (b) é mostrado o resultado do preenchimento morfológico, evidenciando o total preenchimento da ROI, impossibilitando quaisquer detecções.

Obviamente, a remoção da etapa de preenchimento morfológico mostrou-se necessária, mas houve complicações posteriores que demandaram análise mais acurada: uma vez que o recurso de preenchimento morfológico não faz mais parte do processo, *contornos* passam a ser relevantes para as detecções, já que – conforme discutido na Seção 5.2.1 – elementos de linha possuem área e estas são afetadas pela filtragem de área e pela abertura morfológica. Esta última apresenta, *a priori*, uma influência bastante negativa para o processo neste novo cenário: não é razoável descartar elementos de linha na imagem se estes são a fonte de informação para o processamento de alto-nível (agrupamento, contagem e caracterização). Para confirmar esta hipótese foram conduzidos quatro experimentos consecutivos:

**Experimento I:** Processamento *com* preenchimento morfológico e *sem* abertura;

**Experimento II:** Processamento *com* preenchimento morfológico e *com* abertura;

**Experimento III:** Processamento *sem* preenchimento morfológico e *sem* abertura;

**Experimento IV:** Processamento *sem* preenchimento morfológico e *com* abertura.

Em cada um destes experimentos foram levantados os valores médios para os indicadores das detecções ( $n$ ,  $N_{VP}$ ,  $N_{FP}$  e  $N_{FN}$ )<sup>1</sup>. Por uma questão metodológica, estas médias foram estimadas dividindo-se a BIM em dois grupos distintos: 20 imagens digitalizadas (tela-filme) oriundas do BancoWeb mais 31 imagens de radiografia computadorizada (CR) oriundas da FMB-UNESP. Essa divisão se deve às diferentes resoluções espaciais: as imagens CR possuem maior densidade de pixels, e era esperado que isso influísse nas detecções. Dentro de cada grupo, as imagens foram testadas segundo diferentes formas de cálculo do threshold para binarização. Isso foi feito para viabilizar a posterior análise da influência dos thresholds na cadeia. Os resultados são apresentados nas Tabelas 6.1, 6.2, 6.3 e 6.4.

A primeira informação que vem à tona nestes experimentos remete ao número médio de achados-P,  $\bar{n}$ : os processamentos sem preenchimento morfológico (III e IV) ocasionam um aumento do número médio de achados em relação aos experimentos I e II, tanto para o grupo tela-filme quanto para o grupo CR. Isso vai de acordo com o que foi discutido sobre o mascaramento da ROI: uma vez que o resultado da segmentação não é mais corrompido pelo mascaramento, estes objetos passam às etapas posteriores de contagem e caracterização. Portanto, é de se esperar um aumento na contagem (há casos em que  $\bar{n}$  aumentou em mais de uma ordem de grandeza). Uma ilustração do efeito do mascaramento da ROI sobre  $n$  pode ser visto na Figura 6.4.

EXPERIMENTO I											
TELA-FILME					CR						
MÉDIA		OTSU		ENTROPIA		MÉDIA		OTSU		ENTROPIA	
$\bar{n}$	0,65	$\bar{n}$	2,95	$\bar{n}$	13,55	$\bar{n}$	0,03	$\bar{n}$	0,19	$\bar{n}$	6,45
$\bar{N}_{VP}$	0,05	$\bar{N}_{VP}$	0,20	$\bar{N}_{VP}$	1,10	$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,13
$\bar{N}_{FP}$	0,60	$\bar{N}_{FP}$	2,75	$\bar{N}_{FP}$	12,05	$\bar{N}_{FP}$	0,03	$\bar{N}_{FP}$	0,19	$\bar{N}_{FP}$	6,23
$\bar{N}_{FN}$	2,40	$\bar{N}_{FN}$	2,25	$\bar{N}_{FN}$	1,35	$\bar{N}_{FN}$	1,23	$\bar{N}_{FN}$	1,23	$\bar{N}_{FN}$	1,10
DESVIOS MÉDIOS											
MÉDIA		OTSU		ENTROPIA		MÉDIA		OTSU		ENTROPIA	
$n$	0,91	$n$	3,13	$n$	12,22	$n$	0,06	$n$	0,35	$n$	10,59
$N_{VP}$	0,10	$N_{VP}$	0,34	$N_{VP}$	1,25	$N_{VP}$	0,00	$N_{VP}$	0,00	$N_{VP}$	0,28
$N_{FP}$	0,84	$N_{FP}$	2,95	$N_{FP}$	10,77	$N_{FP}$	0,06	$N_{FP}$	0,35	$N_{FP}$	10,15
$N_{FN}$	1,60	$N_{FN}$	1,58	$N_{FN}$	1,46	$N_{FN}$	1,17	$N_{FN}$	1,17	$N_{FN}$	1,17

Tabela 6.1: Valores e desvios médios dos indicadores das detecções para o Experimento I sobre as imagens digitalizadas (tela-filme) e da modalidade radiografia computadorizada (CR).

Deve ficar claro que o mascaramento da ROI é indesejável independentemente dos procedimentos posteriores; afinal, há nesse caso uma completa descaracterização das formas segmentadas. Contudo, a resolução desse empecilho revelou que a cadeia é razoavelmente sensível a ruídos, qualquer que seja o threshold usado.

### 6.2.1 Necessidade de remoção da etapa de abertura morfológica

Conforme dito no início da Seção 6.2, trabalhamos com a hipótese de que a abertura morfológica corresponde a uma etapa prejudicial à cadeia de processamento *desprovida da*

<sup>1</sup>Os valores médios das referidas grandezas serão descritos pelos mesmos símbolos, a menos de uma barra superior.

EXPERIMENTO II											
TELA-FILME						CR					
MÉDIA		OTSU		ENTROPIA		MÉDIA		OTSU	ENTROPIA		
$\bar{n}$	0,00	$\bar{n}$	0,55	$\bar{n}$	8,60	$\bar{n}$	0,00	$\bar{n}$	0,16	$\bar{n}$	5,90
$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,10	$\bar{N}_{VP}$	0,95	$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,10
$\bar{N}_{FP}$	0,00	$\bar{N}_{FP}$	0,50	$\bar{N}_{FP}$	7,45	$\bar{N}_{FP}$	0,03	$\bar{N}_{FP}$	0,16	$\bar{N}_{FP}$	5,81
$\bar{N}_{FN}$	2,60	$\bar{N}_{FN}$	2,50	$\bar{N}_{FN}$	1,60	$\bar{N}_{FN}$	1,23	$\bar{N}_{FN}$	1,23	$\bar{N}_{FN}$	1,13
DESVIOS MÉDIOS											
MÉDIA			OTSU			ENTROPIA					
$n$	0,00	$n$	0,94	$n$	9,88	$n$	0,00	$n$	0,30	$n$	9,26
$N_{VP}$	0,00	$N_{VP}$	0,19	$N_{VP}$	1,24	$N_{VP}$	0,00	$N_{VP}$	0,00	$N_{VP}$	0,18
$N_{FP}$	0,00	$N_{FP}$	0,85	$N_{FP}$	8,69	$N_{FP}$	0,00	$N_{FP}$	0,30	$N_{FP}$	9,11
$N_{FN}$	1,80	$N_{FN}$	1,85	$N_{FN}$	1,68	$N_{FN}$	1,19	$N_{FN}$	1,19	$N_{FN}$	1,17

Tabela 6.2: Valores e desvios médios dos indicadores das detecções para o Experimento II sobre as imagens digitalizadas (tela-filme) e da modalidade radiografia computadorizada (CR).

EXPERIMENTO III											
TELA-FILME						CR					
MÉDIA		OTSU		ENTROPIA		MÉDIA		OTSU	ENTROPIA		
$\bar{n}$	6,45	$\bar{n}$	33,55	$\bar{n}$	37,45	$\bar{n}$	6,00	$\bar{n}$	25,52	$\bar{n}$	36,94
$\bar{N}_{VP}$	0,25	$\bar{N}_{VP}$	1,85	$\bar{N}_{VP}$	2,15	$\bar{N}_{VP}$	0,13	$\bar{N}_{VP}$	0,58	$\bar{N}_{VP}$	0,94
$\bar{N}_{FP}$	6,05	$\bar{N}_{FP}$	30,95	$\bar{N}_{FP}$	34,10	$\bar{N}_{FP}$	5,71	$\bar{N}_{FP}$	24,65	$\bar{N}_{FP}$	35,19
$\bar{N}_{FN}$	2,10	$\bar{N}_{FN}$	0,40	$\bar{N}_{FN}$	0,30	$\bar{N}_{FN}$	0,97	$\bar{N}_{FN}$	0,48	$\bar{N}_{FN}$	0,26
DESVIOS MÉDIOS											
MÉDIA			OTSU			ENTROPIA					
$n$	7,13	$n$	14,72	$n$	13,39	$n$	4,88	$n$	6,98	$n$	11,22
$N_{VP}$	0,43	$N_{VP}$	1,35	$N_{VP}$	1,45	$N_{VP}$	0,22	$N_{VP}$	0,67	$N_{VP}$	0,97
$N_{FP}$	6,67	$N_{FP}$	14,94	$N_{FP}$	14,23	$N_{FP}$	4,66	$N_{FP}$	6,62	$N_{FP}$	11,01
$N_{FN}$	1,63	$N_{FN}$	0,56	$N_{FN}$	0,48	$N_{FN}$	1,00	$N_{FN}$	0,59	$N_{FN}$	0,38

Tabela 6.3: Valores e desvios médios dos indicadores das detecções para o Experimento III sobre as imagens digitalizadas (tela-filme) e da modalidade radiografia computadorizada (CR).

*etapa de preenchimento.* Os experimentos III e IV serviram para testarmos essa hipótese, que acabou mostrando-se verdadeira. A análise foi realizada comparando-se os indicadores do experimento IV em relação aos indicadores do experimento III. Os dados utilizados são aqueles expressos nas Tabelas 6.3 e 6.4 e os resultados das comparações podem ser vistos na Tabela 6.5.

Do ponto de vista do número de achados, o uso da abertura leva a uma redução importante sobre  $\bar{n}$ , chegando quase a reduzir esse número em uma ordem de grandeza quando do uso do threshold por Otsu no grupo tela-filme, por exemplo. Esta seria uma característica desejável para o processo como um todo se não fosse acompanhada de um aumento de  $\bar{N}_{FN}$ , independentemente do threshold ou do tipo de imagem. Isto é, juntamente com achados espúrios, a abertura promove a remoção de objetos relevantes para as detecções, incrementando a contagem de falsos-negativos. Paralelamente, o número médio de verdadeiros-positivos cai quando do uso de abertura, independentemente do tipo de imagem ou do threshold usado.

EXPERIMENTO IV																	
TELA-FILME					CR												
MÉDIA		OTSU		ENTROPIA		MÉDIA		OTSU		ENTROPIA							
$\bar{n}$	0,00	$\bar{n}$	3,55	$\bar{n}$	15,25	$\bar{n}$	0,32	$\bar{n}$	8,19	$\bar{n}$	23,00						
$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,25	$\bar{N}_{VP}$	1,05	$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,00	$\bar{N}_{VP}$	0,52						
$\bar{N}_{FP}$	0,00	$\bar{N}_{FP}$	3,20	$\bar{N}_{FP}$	13,75	$\bar{N}_{FP}$	0,32	$\bar{N}_{FP}$	8,00	$\bar{N}_{FP}$	22,23						
$\bar{N}_{FN}$	2,45	$\bar{N}_{FN}$	2,05	$\bar{N}_{FN}$	1,15	$\bar{N}_{FN}$	1,23	$\bar{N}_{FN}$	1,03	$\bar{N}_{FN}$	0,52						
DESVIOS MÉDIOS																	
MÉDIA			OTSU			ENTROPIA			MÉDIA			OTSU			ENTROPIA		
$n$	0,00	$n$	3,87	$n$	8,03	$n$	0,50	$n$	3,52	$n$	9,35						
$N_{VP}$	0,00	$N_{VP}$	0,43	$N_{VP}$	1,17	$N_{VP}$	0,00	$N_{VP}$	0,00	$N_{VP}$	0,77						
$N_{FP}$	0,00	$N_{FP}$	3,50	$N_{FP}$	7,60	$N_{FP}$	0,50	$N_{FP}$	3,35	$N_{FP}$	9,02						
$N_{FN}$	1,65	$N_{FN}$	1,67	$N_{FN}$	0,94	$N_{FN}$	1,19	$N_{FN}$	1,01	$N_{FN}$	0,63						

Tabela 6.4: Valores e desvios médios dos indicadores das detecções para o Experimento IV sobre as imagens digitalizadas (tela-filme) e da modalidade radiografia computadorizada (CR).

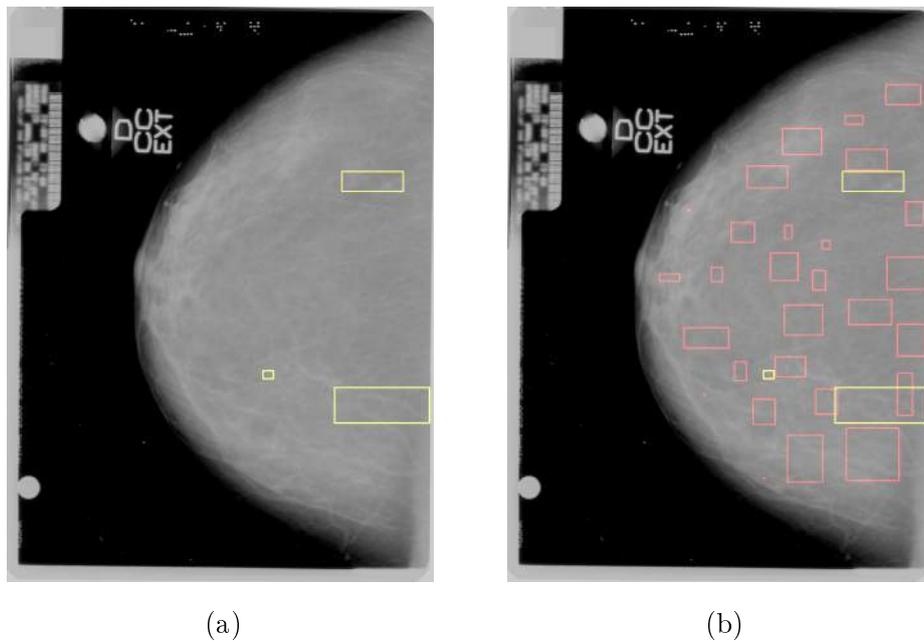


Figura 6.4: Influência do preenchimento morfológico em uma imagem do grupo tela-filme segmentada com threshold calculado por média. Em (a) é mostrado o resultado do processamento da amostra no experimento I (com preenchimento), e em (b) é mostrado o resultado para a mesma amostra no experimento III (sem preenchimento). Nota-se o claro aumento do número de achados-P.

O aumento de  $\bar{N}_{FN}$  e a redução de  $\bar{N}_{VP}$  têm uma origem em comum: a abertura remove objetos outrora segmentados, causando o aumento de  $\bar{N}_{FN}$  e dificultando o estabelecimento de verdadeiros-positivos (menor número de achados adequados/relevantes). Note-se que, de fato, a abertura tem influência direta sobre a *qualidade* das detecções. Nesse sentido, justifica-se a remoção da abertura da cadeia de processamento.

Comparação entre experimentos IV e III (desvios percentuais)											
TELA-FILME					CR						
MÉDIA		OTSU		ENTROPIA		MÉDIA		OTSU		ENTROPIA	
$\bar{n}$	-100	$\bar{n}$	-89	$\bar{n}$	-59	$\bar{n}$	-95	$\bar{n}$	-68	$\bar{n}$	-38
$\bar{N}_{VP}$	-100	$\bar{N}_{VP}$	-86	$\bar{N}_{VP}$	-51	$\bar{N}_{VP}$	-100	$\bar{N}_{VP}$	-100	$\bar{N}_{VP}$	-45
$\bar{N}_{FP}$	-100	$\bar{N}_{FP}$	-90	$\bar{N}_{FP}$	-60	$\bar{N}_{FP}$	-94	$\bar{N}_{FP}$	-68	$\bar{N}_{FP}$	-37
$\bar{N}_{FN}$	17	$\bar{N}_{FN}$	413	$\bar{N}_{FN}$	283	$\bar{N}_{FN}$	27	$\bar{N}_{FN}$	113	$\bar{N}_{FN}$	100

Tabela 6.5: Desvios relativos entre os indicadores das detecções do experimento IV em relação aos do experimento III. Quantidades mostradas em percentuais.

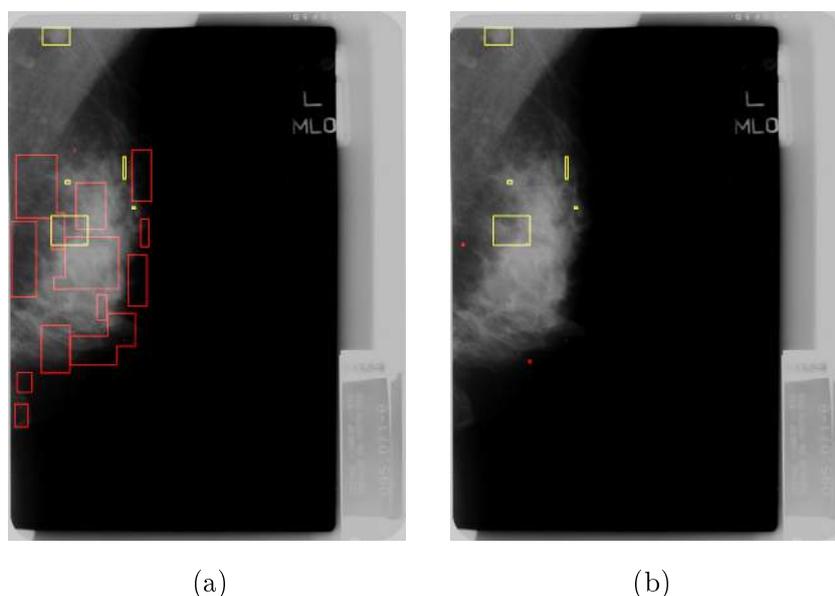


Figura 6.5: Efeito da abertura morfológica sobre as detecções: em (a) é mostrado o resultado do processamento de uma amostra segmentada por threshold entrópico sem uso da abertura. Em (b) é mostrado o resultado de processamento da mesma amostra sem uso da abertura e com o mesmo threshold. Nota-se nesse caso uma redução desejável no número de achados-P mas mal-caracterizada, pois os achados remanescentes em nada se relacionam com os verdadeiros achados-M.

### 6.3 Análise dos thresholds

Os experimentos descritos na Seção 6.2 evidenciaram as inconveniências do preenchimento e abertura na cadeia de processamento. Tal como lá discutido, estas duas operações morfológicas levaram respectivamente à descaracterização dos objetos segmentados e à remoção massiva destes. Uma vez removidas as operações morfológicas, foi revelada a alta sensibilidade da cadeia a ruídos, afirmação esta confirmada pelos indicadores mostrados na Tabela 6.3. Uma análise destes indicadores mostra que a cadeia de processamento – quando operada *sem enhancement* – gera um número inaceitavelmente grande de achados e um número médio de verdadeiros-positivos muito pequeno. Como parâmetros de comparação, usamos os indicadores de *detecções perfeitas*; isto é, quando as ROIs demarcadas pelos médicos são comparadas com *elas próprias* segundo a metodologia exposta na Seção 5.2.2. Estes indicadores podem ser vistos na Tabela 6.6

A Figura 6.6 traz dois gráficos mostrando os números médios de achados e de verdadeiros-

DADOS DE REFERÊNCIA					
GRUPO TELA-FILME			GRUPO CR		
INDICADORES	DESV. MÉDIOS		INDICADORES	DESV. MÉDIOS	
$\bar{m}$	3,85	2,27	$\bar{m}$	1,58	0,88
$\bar{n}$	2,60	1,80	$\bar{n}$	1,23	1,19
$\bar{N}_{VP}$	2,60	1,80	$\bar{N}_{VP}$	1,23	1,19
$\bar{N}_{FP}$	0,00	0,00	$\bar{N}_{FP}$	0,00	0,00
$\bar{N}_{FN}$	0,00	0,00	$\bar{N}_{FN}$	0,00	0,00

Tabela 6.6: Tabela com os valores médios de referência, que foram obtidos avaliando-se as ROIs-M com elas próprias segundo procedimento de contagem e caracterização desenvolvido. Como esperado, não há falsos-positivos ou negativos, e o número de verdadeiros-positivos corresponde ao número de achados. Note-se que o número de achados demarcados pelos médicos não corresponde ao número de achados detectados. Isso é devido ao processo de agrupamento descrito na Seção 5.2.2.

positivos tanto para os dados de referência quanto para o Experimento III.

De posse dos indicadores de detecção, podemos definir duas grandezas básicas para avaliação dos experimentos:

*Eficiência de detecção.* Definimos a eficiência de uma detecção como sendo a razão entre o número de verdadeiros-positivos e o número de achados-P, isto é,

$$\varepsilon = \frac{N_{VP}}{n}; \quad \text{desde que } n \neq 0. \quad (6.1)$$

*Eficiência média de detecção.* Definimos a eficiência média de um conjunto de  $\mathcal{N}$  detecções como sendo a média aritmética das eficiências individuais:

$$\bar{\varepsilon} = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \varepsilon_i \quad (6.2)$$

Tais grandezas são definidas para que se possa quantificar a ação de certa configuração sobre cada imagem individualmente ( $\varepsilon$ ) ou sobre um conjunto de imagens ( $\bar{\varepsilon}$ ). Note-se que a eficiência de detecção definida na Equação (6.1) é indeterminada quando não existem achados-P. Para fins de computação dos valores médios, tais casos não foram considerados.

EFICIÊNCIA MÉDIA DE DETECÇÃO - $\bar{\varepsilon}$					
GRUPO TELA-FILME			GRUPO CR		
Média	3,70%		Média	2,50%	
Otsu	7,21%		Otsu	2,16%	
Máxima Entropia	7,73%		Máxima Entropia	2,43%	
DESVIOS MÉDIOS					
GRUPO TELA-FILME			GRUPO CR		
Média	5,69%		Média	3,74%	
Otsu	5,52%		Otsu	2,67%	
Máxima Entropia	5,78%		Máxima Entropia	2,86%	

Tabela 6.7: Percentuais de eficiência média para os processamentos ocorridos no Experimento III.

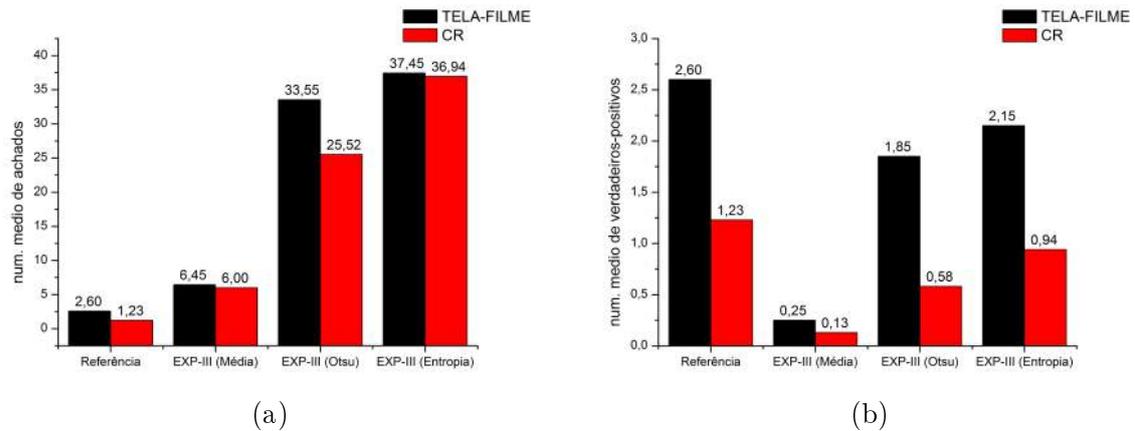


Figura 6.6: Comparações entre valores de referência e experimento III. Em (a) é feita a comparação entre  $\bar{n}$  e em (b) é feita comparação de  $\bar{N}_{VP}$ . Nota-se em (a) a discrepância entre o número médio de achados esperado e os efetivamente produzidos pelo processamento sem enhancement. No caso do thresholds por Otsu e máxima entropia, o número de detecções é uma ordem de grandeza maior que o esperado. Em contrapartida, o número médio de verdadeiros-positivos gerado pelo processamento foi sempre menor do que o esperado. Entretanto, no caso de threshold calculado via máxima entropia, esse número teve um desvio de apenas 17% em relação ao esperado.

Nota-se pela Tabela 6.7 que o threshold calculado via média apresenta baixa eficiência média de detecção para o grupo tela-filme, mesmo em relação aos baixos percentuais dos outros dois thresholds. Já no grupo CR a situação se inverte: o threshold via média proporcionou maior eficiência média, embora supere em apenas 2,9% aquela produzida pelo threshold entrópico. Face a essa constatação, optou-se pelo descarte dos thresholds calculados via média e Otsu. Das informações levantadas nestes experimentos, pudemos definir a configuração ótima segundo qual a cadeia de processamento deve operar (agora com enhancement):

- Sem preenchimento morfológico;
- Sem abertura morfológica;
- Threshold de segmentação obtido pelo método da máxima entropia.

As investigações posteriores ocorreram no sentido de verificar se a cadeia de processamento contendo a etapa de enhancement via Equação de Perona-Malik apresentou indicadores melhores que os do Experimento III – doravante chamado de Grupo de Controle – e o quão próximos dos indicadores de referência eles chegaram. Nas próximas seções serão expostos e discutidos os resultados.

## 6.4 Influência da difusão não-linear

Os experimentos finais do projeto ocorreram com vistas a confrontar seus indicadores com os do grupo de controle e grupo de referência, tal como antecipado no final da seção anterior. Entretanto, os dados analisados não cobrem toda a BIM: por uma questão de tempo computacional (que será discutida na Seção 6.4.3), cinco imagens do grupo CR não puderam ser processadas. Isso alterou os indicadores de ambos os grupos, referência e controle, constantes nas Tabelas 6.6 e 6.3, respectivamente. Os resultados atualizados

para comparação utilizadas nas próximas análises podem ser visualizados nas Tabelas 6.8, 6.9 e 6.10. A compilação dos indicadores de todos os processamentos relativos à Equação de Perona-Malik pode ser visualizada na Tabela 6.17, na página 120.

INDICADORES DO GRUPO DE REFERÊNCIA					
GRUPO TELA-FILME			GRUPO CR		
INDICADORES		DESV. MÉDIOS	INDICADORES		DESV. MÉDIOS
$\bar{m}$	3,85	2,27	$\bar{m}$	1,46	0,92
$\bar{n}$	2,60	1,80	$\bar{n}$	1,04	1,20
$\bar{N}_{VP}$	2,60	1,80	$\bar{N}_{VP}$	1,04	1,20
$\bar{N}_{FP}$	0,00	0,00	$\bar{N}_{FP}$	0,00	0,00
$\bar{N}_{FN}$	0,00	0,00	$\bar{N}_{FN}$	0,00	0,00

Tabela 6.8: Tabela com os valores médios do grupo de referência .

INDICADORES DO GRUPO DE CONTROLE						
TELA-FILME				CR		
MÉDIA		OTSU	ENTROPIA	MÉDIA		ENTROPIA
$\bar{n}$	6,45	$\bar{n}$ 33,55	$\bar{n}$ 37,45	$\bar{n}$ 6,35	$\bar{n}$ 25,54	$\bar{n}$ 36,04
$\bar{N}_{VP}$	0,25	$\bar{N}_{VP}$ 1,85	$\bar{N}_{VP}$ 2,15	$\bar{N}_{VP}$ 0,12	$\bar{N}_{VP}$ 0,50	$\bar{N}_{VP}$ 0,81
$\bar{N}_{FP}$	6,05	$\bar{N}_{FP}$ 30,95	$\bar{N}_{FP}$ 34,10	$\bar{N}_{FP}$ 6,12	$\bar{N}_{FP}$ 24,77	$\bar{N}_{FP}$ 34,38
$\bar{N}_{FN}$	2,10	$\bar{N}_{FN}$ 0,40	$\bar{N}_{FN}$ 0,30	$\bar{N}_{FN}$ 0,81	$\bar{N}_{FN}$ 0,42	$\bar{N}_{FN}$ 0,23
DESVIOS MÉDIOS						
MÉDIA		OTSU	ENTROPIA	MÉDIA		ENTROPIA
$n$	7,13	$n$ 14,72	$n$ 13,39	$n$ 4,77	$n$ 7,43	$n$ 11,03
$N_{VP}$	0,43	$N_{VP}$ 1,35	$N_{VP}$ 1,45	$N_{VP}$ 0,21	$N_{VP}$ 0,70	$N_{VP}$ 0,95
$N_{FP}$	6,67	$N_{FP}$ 14,94	$N_{FP}$ 14,23	$N_{FP}$ 4,62	$N_{FP}$ 6,94	$N_{FP}$ 10,51
$N_{FN}$	1,63	$N_{FN}$ 0,56	$N_{FN}$ 0,48	$N_{FN}$ 1,02	$N_{FN}$ 0,56	$N_{FN}$ 0,34

Tabela 6.9: Valores e desvios médios dos indicadores das detecções para o grupo de controle .

EFIC. MÉDIA DO GRUPO DE CONTROLE			
GRUPO TELA-FILME		GRUPO CR	
Média	3,70%	Média	2,28%
Otsu	7,21%	Otsu	1,85%
Máxima Entropia	7,73%	Máxima Entropia	2,19%
DESVIOS MÉDIOS			
GRUPO TELA-FILME		GRUPO CR	
Média	5,69%	Média	3,74%
Otsu	5,52%	Otsu	2,67%
Máxima Entropia	5,78%	Máxima Entropia	2,86%

Tabela 6.10: Percentuais de eficiência média para os processamentos ocorridos no grupo de controle .

#### 6.4.1 Influência da força de difusão

A primeira análise foi realizada para que se levantassem informações acerca das variações dos indicadores médios das detecções com a forma das forças. Nesse sentido,

foram levantados os indicadores para diversos valores de  $\alpha$ , que corresponde ao parâmetro de curvatura das difusibilidades, tal como expresso na Equação (4.20). Embora pudesse ser posto para variar continuamente sobre qualquer intervalo positivo, restringimos os valores de  $\alpha$  para inteiros variando de 2 até 20, em passos de três unidades; isto é,  $\alpha \in \{2, 5, 8, 11, 14, 17, 20\}$ . Esta escolha foi feita devido a limitações em tempo de processamento.

### Variação do número médio de achados

A Figura 6.7 mostra a evolução do número médio de achados,  $\bar{n}$ , para ambos os grupos tela-filme e CR, segundo as três formas de cômputo do limiar de separação na difusão, K.

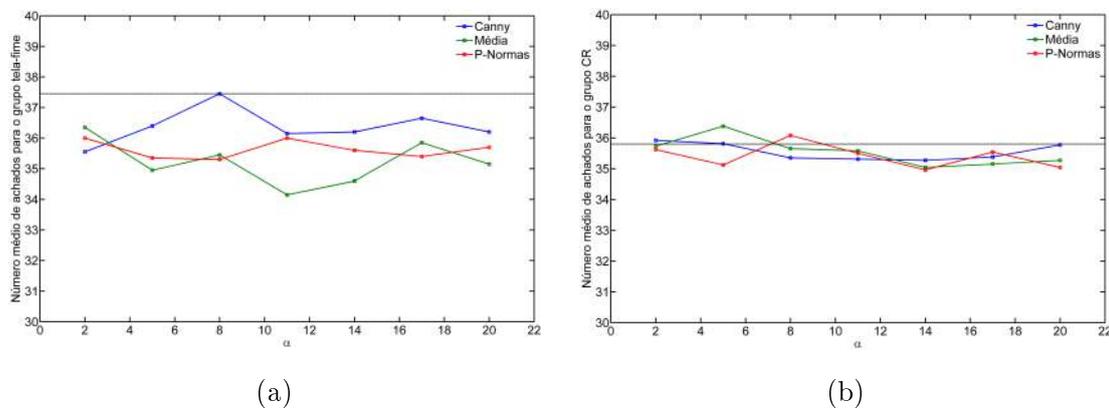


Figura 6.7: Evolução de  $\bar{n}$  com  $\alpha$  para as três formas de cômputo do limiar K. Em (a) são mostrados os pontos para o indicadores do grupo tela-filme e em (b) os do grupo CR. As linhas pretas tracejadas marcam o número médio de achados do grupo de controle. As linhas contínuas entre os pontos servem apenas para visualização.

Verifica-se graficamente que os processamentos do grupo tela-filme produziram – a menos de um ponto – números médios de achados menores que o respectivo indicador no grupo de controle, com aparente tendência de aumento para  $\alpha > 10$ . Por outro lado, no grupo CR os número desviaram muito pouco do valor de referência. Ainda nesse grupo, nota-se aparente tendência de queda.

Foram calculados para cada caso os coeficientes de correlação de Pearson, já que esta grandeza quantifica a tendência de evolução linear conjunta entre dois conjuntos de dados [75]. Os coeficientes ( $r$ ) são mostrados na Tabela 6.11. À exceção do número médio de detecções para as imagens do grupo tela-filme com K calculado pelo estimador de ruído de Canny, todos os grupos mostraram ligeira tendência de *queda* na relação  $\alpha \times \bar{n}$ . Em particular, o processamento do grupo CR com K calculado via média da magnitude do gradiente possui  $r = -0,76$ , indicando uma tendência explícita de queda, tal como mostrado na Figura 6.7 (b).

COEFICIENTES DE CORRELAÇÃO ENTRE $\alpha$ e $\bar{n}$			
GRUPO TELA-FILME		GRUPO CR	
<b>Canny</b>	0,16	<b>Canny</b>	-0,39
<b>Média</b>	-0,27	<b>Média</b>	-0,76
<b>P-Normas</b>	-0,13	<b>P-Normas</b>	-0,39

Tabela 6.11: Coeficientes de correlação de Pearson entre  $\alpha$  e  $\bar{n}$ .

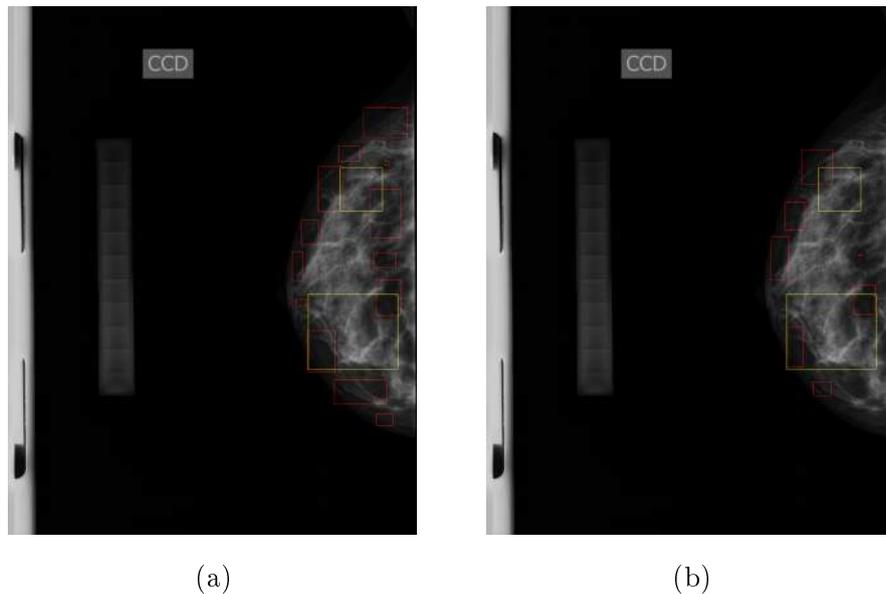


Figura 6.8: Ilustração da tendência de queda do número de achados-P com o aumento de  $\alpha$ . Em (a) é mostrada uma amostra do grupo CR processada com  $\alpha = 2$  e em (b) é mostrada a mesma amostra, mas processada com  $\alpha = 20$ . A redução de  $n$  nesse caso foi de 13 para 7. Em ambos os casos,  $N_{VP} = 1$  e  $N_{FN} = 1$ . Em termos de eficiência, houve um salto – para esta amostra – de  $\varepsilon = 7,69\%$  para  $14,29\%$ .

Do ponto de vista do sistema de detecção, quão menor for o número de achados-P, mais eficiente será o processamento. Entretanto, a análise mostrou que a tendência de queda de  $\bar{n}$  com  $\alpha$  é incipiente dentro do intervalo estudado. Chama a atenção ainda o fato de as imagens do grupo tela-filme apresentarem menor correlação do que as do grupo CR. A Figura 6.8 ilustra uma ocorrência de redução do número de achados em função do aumento da curvatura da função difusibilidade.

### Verdadeiros-positivos e eficiência média

Conforme visto na seção anterior, há uma indicação de que é possível reduzir o número de achados-P por meio do aumento do parâmetro  $\alpha$ . Contudo, essa hipotética redução de  $n$  deveria vir acompanhada, idealmente, de um aumento no número de verdadeiros-positivos, configurando um aumento na eficiência de detecção. A investigação dessa possibilidade se dá pelo levantamento de  $\bar{N}_{VP}$  juntamente com  $\bar{n}$ . Esta análise conjunta é melhor expressa em termos da eficiência média de detecção ( $\bar{\varepsilon}$ ) em função de  $\alpha$ . Na Figura 6.9 são mostrados os gráficos dos grupos tela-filme e CR contendo as variações das referidas grandezas.

Por análise gráfica, verifica-se que o aumento de  $\alpha$  proporciona um aumento efetivo na eficiência média nos grupos tela-filme e CR quando do uso de p-normas para o cômputo do limiar K. Por outro lado, chamou a atenção a forte correlação entre  $\alpha$  e  $\bar{\varepsilon}$  quando do uso da média da magnitude do gradiente ( $r = 0,85$ ), tal como mostrado na Tabela 6.12. No grupo CR, o estimador de ruído de Canny praticamente não influenciou nos indicadores, uma vez que as eficiências oscilaram muito pouco em torno do valor do grupo de controle.

A análise seguinte trata sobre  $\bar{N}_{VP}$  propriamente dito: embora a eficiência média das detecções tenha mostrado sinais de aumento, o número médio de verdadeiros-positivos não o fez. Efetivamente, para o grupo CR, os valores de  $\bar{N}_{VP}$  pouco variaram conforme a evolução de  $\alpha$ . Ademais, este grupo apresentou desempenho inferior ao grupo de controle

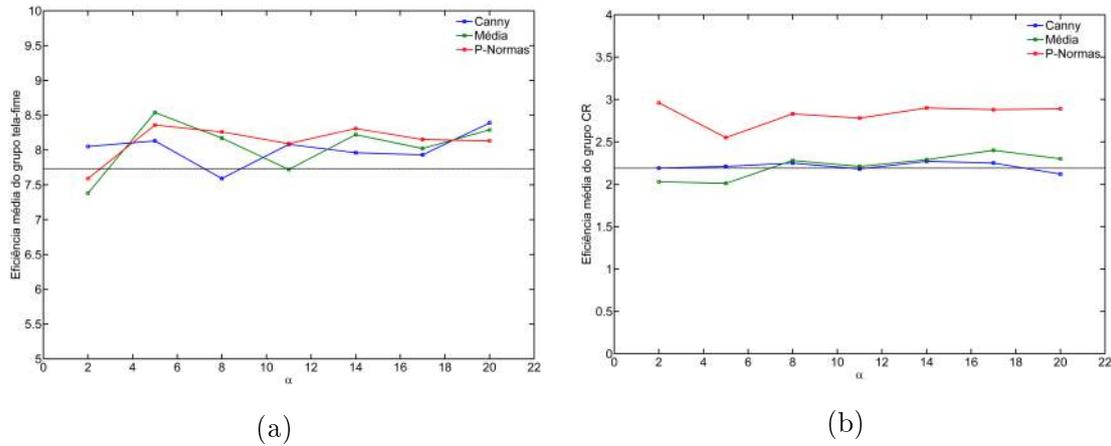


Figura 6.9: Evolução de  $\bar{\varepsilon}$  com  $\alpha$  para as três formas de cômputo do limiar K. Em (a) são mostrados os pontos para o indicadores do grupo tela-filme e em (b) os do grupo CR. As linhas pretas tracejadas marcam a eficiência média do grupo de controle. As linhas contínuas entre os pontos servem apenas para visualização.

COEFICIENTES DE CORRELAÇÃO ENTRE $\alpha$ e $\bar{\varepsilon}$			
GRUPO TELA-FILME		GRUPO CR	
<b>Canny</b>	0,31	<b>Canny</b>	-0,14
<b>Média</b>	0,34	<b>Média</b>	0,85
<b>P-Normas</b>	0,38	<b>P-Normas</b>	0,29

Tabela 6.12: Coeficientes de correlação de Pearson entre  $\alpha$  e  $\bar{\varepsilon}$ .

para os processamentos com K sendo computado pelo estimador de ruído de Canny e pela média da magnitude do gradiente. O grupo tela-filme apresentou índices um pouco melhores, principalmente para o processamento por p-normas. Entretanto, a melhora foi sutil, em termos de  $\bar{N}_{VP}$ . Para se ter uma idéia, o valor máximo de  $\bar{N}_{VP}$  para o grupo tela-filme e K calculado pelas p-normas é igual a 2,30. Este valor está 12% abaixo dos 2,60 de referência, mesmo estando 7% acima de seu correspondente no grupo de controle. Por fim, verificou-se que existe uma correlação de aproximadamente 0,40 entre  $\alpha$  e  $\bar{N}_{VP}$  no grupo tela-filme.

A Figura 6.10 mostra os resultados dos levantamentos de  $\bar{N}_{VP}$  em função de  $\alpha$  e na Tabela 6.13 podem ser vistos os coeficientes de correlação entre os dois conjuntos de dados.

COEFICIENTES DE CORRELAÇÃO ENTRE $\alpha$ e $\bar{N}_{VP}$			
GRUPO TELA-FILME		GRUPO CR	
<b>Canny</b>	0,41	<b>Canny</b>	-0,47
<b>Média</b>	0,46	<b>Média</b>	0,29
<b>P-Normas</b>	0,46	<b>P-Normas</b>	0,00

Tabela 6.13: Coeficientes de correlação de Pearson entre  $\alpha$  e  $\bar{N}_{VP}$ .

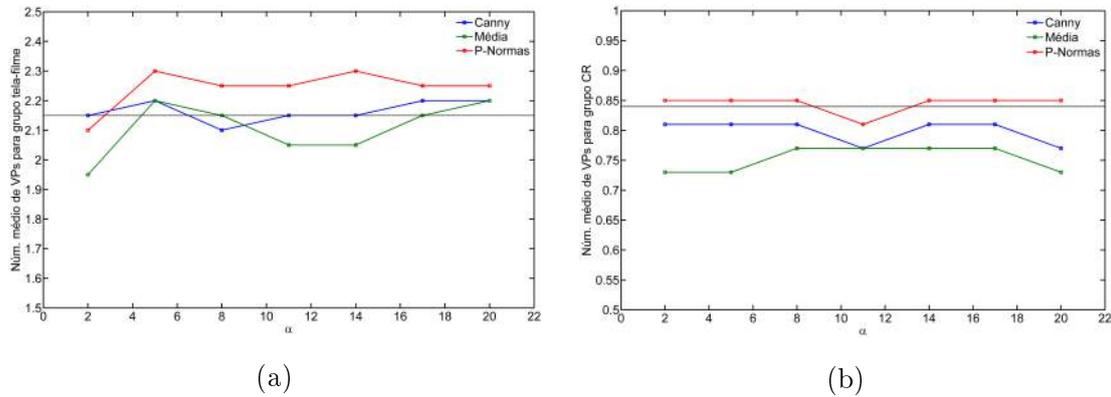


Figura 6.10: Evolução de  $\bar{N}_{VP}$  com  $\alpha$  para as três formas de cômputo do limiar K. Em (a) são mostrados os pontos para o indicadores do grupo tela-filme e em (b) os do grupo CR. As linhas pretas tracejadas marcam o número médio de VPs do grupo de controle. As linhas contínuas entre os pontos servem apenas para visualização.

### Falsos-positivos e falsos-negativos

O levantamento do número médio de falsos-positivos deu base para que fosse constatada mais uma vez a forte sensibilidade da cadeia de processamento a ruídos, uma vez que quase todos os achados foram considerados detecções espúrias. Conforme mostrado na Figura 6.11, os valores de  $\bar{N}_{FP}$  pouco mudaram em relação aos diferentes valores de  $\alpha$ . Mais do que isso, eles permaneceram muito próximos dos valores de seus respectivos grupos de controle. Embora tenham apresentado – à exceção do grupo tela-filme com K obtido pelo estimador de Canny – tendência de queda (vide Tabela 6.14), não foi observada redução significativa nos falsos-positivos.

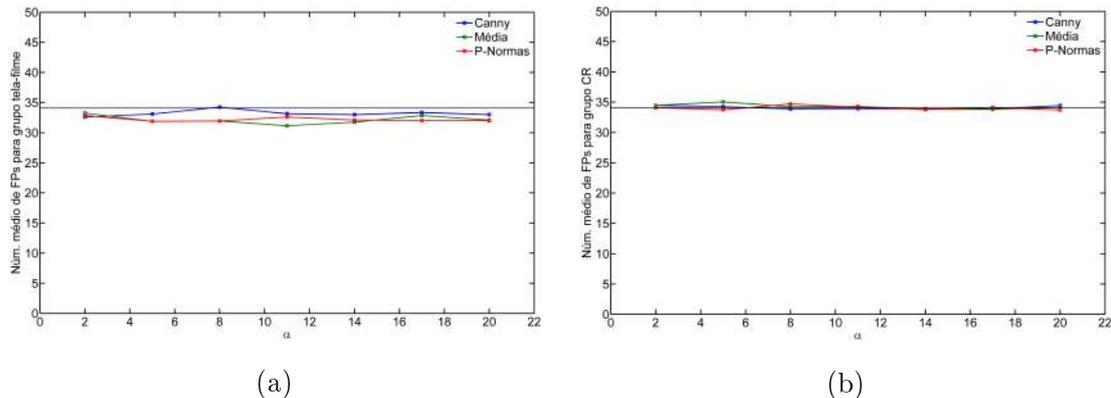


Figura 6.11: Evolução de  $\bar{N}_{FP}$  com  $\alpha$  para as três formas de cômputo do limiar K. Em (a) são mostrados os pontos para o indicadores do grupo tela-filme e em (b) os do grupo CR. As linhas pretas tracejadas marcam o número médio de FPs do grupo de controle. As linhas contínuas entre os pontos servem apenas para visualização.

Caso completamente análogo acontece com os números médios de falsos negativos. Tal como pode ser visto na Figura 6.12, as médias ficaram sempre muito próximas de seus correspondentes no grupo de controle, quando idealmente deveriam ir a zero. Surpreendentemente, o grupo CR processado com K via p-normas não produziu flutuação alguma em torno de seu valor médio, 0,19. Há de se ressaltar que nesse caso específico, houve uma redução de 21% no número médio de falsos-positivos.

COEFICIENTES DE CORRELAÇÃO ENTRE $\alpha$ e $\tilde{N}_{FP}$			
GRUPO TELA-FILME		GRUPO CR	
Canny	0,07	Canny	-0,23
Média	-0,19	Média	-0,71
P-Normas	-0,46	P-Normas	-0,29

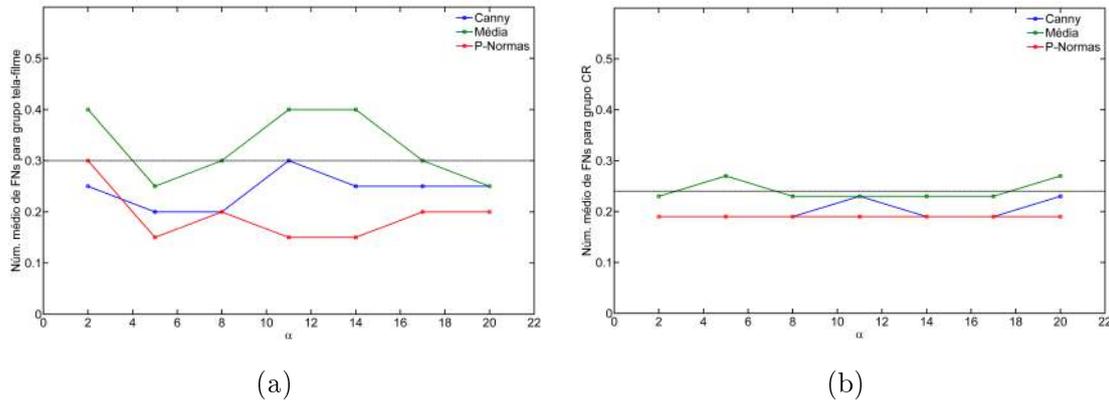
Tabela 6.14: Coeficientes de correlação de Pearson entre  $\alpha$  e  $\tilde{N}_{FP}$ .

Figura 6.12: Evolução de  $\tilde{N}_{FN}$  com  $\alpha$  para as três formas de cômputo do limiar K. Em (a) são mostrados os pontos para o indicadores do grupo tela-filme e em (b) os do grupo CR. As linhas pretas tracejadas marcam o número médio de FNs do grupo de controle. As linhas contínuas entre os pontos servem apenas para visualização.

COEFICIENTES DE CORRELAÇÃO ENTRE $\alpha$ e $\tilde{N}_{FN}$			
GRUPO TELA-FILME		GRUPO CR	
Canny	0,34	Canny	0,47
Média	-0,28	Média	-0,16
P-Normas	-0,36	P-Normas	—

Tabela 6.15: Coeficientes de correlação de Pearson entre  $\alpha$  e  $\tilde{N}_{FN}$ . Note-se o coeficiente indeterminado para o grupo CR com K calculado via p-normas. Isso ocorreu pois não houve variação de  $\tilde{N}_{FN}$  com  $\alpha$ .

#### 6.4.2 Análise qualitativa da influência do limiar K

Juntamente com as funções difusibilidade, o processamento ocorreu segundo variações na forma do cômputo do limiar de separação, K. Este é um parâmetro fundamental da técnica, tal como descrito na Seção 4.3.2, sendo determinante sobre alguns resultados já apresentados na Seção 6.4.1. Efetivamente, os experimentos descritos naquela Seção permitiram formar juízo simultaneamente sobre a influência de  $\alpha$  e K sobre os resultados. Por exemplo, foi mostrado que o cômputo de K pelas p-normas da magnitude do gradiente apresentam ganho importante no número médio de verdadeiros positivos, independentemente do  $\alpha$  adotado nas medidas.

A análise posterior ocorreu sobre K, de maneira qualitativa<sup>2</sup>. Conforme pode ser visto na Tabela 6.16, os percentuais médios para o grupo tela-filme foram todos maiores que 8,30%. No caso particular das amostras difundidas com K obtido por média, a eficiência média máxima chegou a 8,54%, superando em aproximadamente 10,5% o grupo de controle.

<sup>2</sup>A referida análise é caracterizada como qualitativa pois não formamos juízo sobre as *causas* das melhorias dos índices, mas somente apontamos *qual* forma de cômputo proporciona os melhores indicadores.

EFICIÊNCIA MÉDIA MÁXIMA			
GRUPO TELA-FILME		GRUPO CR	
Canny	8,39%	Canny	2,27%
Média	8,54%	Média	2,40%
P-Normas	8,36%	P-Normas	2,96%
EFICIÊNCIA MÉDIA MÍNIMA			
GRUPO TELA-FILME		GRUPO CR	
Canny	7,59%	Canny	2,12%
Média	7,38%	Média	2,01%
P-Normas	7,59%	P-Normas	2,55%

Tabela 6.16: Eficiências medidas máximas e mínimas para os grupos tela-filme e CR.

Analisando os dados em questão, verifica-se que o cômputo de  $K$  por média proporciona um incremento na eficiência média do grupo tela-filme, ao passo que no grupo CR as p-normas levam a melhores resultados. Para o primeiro, há um aumento percentual de 10,5%, enquanto para o segundo o aumento é de 35,16%. Para o grupo tela-filme, mesmo as eficiências médias mínimas se aproximam muito dos respectivos indicadores de controle. Nesse sentido, verifica-se que a introdução da Equação de Perona-Malik é benéfica para o processamento, promovendo um aumento da eficiência média de detecção.

No que diz respeito – de fato – aos limiares, percebemos que o estimador de ruído de Canny não proporcionou bons resultados em relação aos outros dois limiares. Adicionalmente, cabe ressaltar que a média e a p-norma da magnitude do gradiente são diretamente proporcionais, cuja proporcionalidade é dada por  $\|\vec{\nabla}u\|_3$ , tal como mostrado nas Equações (4.23) e (5.3). Isso indica que a média da magnitude do gradiente corresponde a uma estimativa relevante em termos de separação entre difusões diretas e reversas, mas não fica claro porque a alteração de  $K$  igual à média da magnitude do gradiente por uma 3-norma levou a melhores resultados. A única constatação diz respeito ao aumento da magnitude de  $K$  com as p-normas, tal como esperado. Esse aumento de magnitude de  $K$  faz com que estruturas preservadas com  $K$  calculado via média sejam difundidas com  $K$  via p-normas. Uma ilustração desta ocorrência pode ser visualizada nas Figuras 6.13 e 6.14.

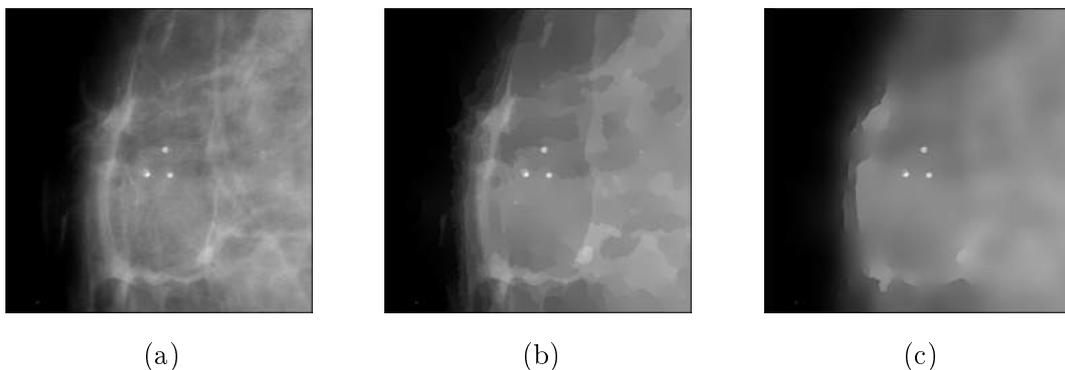


Figura 6.13: Filtragem via Equação de Perona-Malik sobre um fragmento do mamograma. Em (a) é mostrado o fragmento original contendo um agrupamento de calcificações grosseiras. Em (b) e (c) são mostrados os resultados da difusão não linear ao longo de 1000 iterações para  $K$  calculado via média e p-norma da magnitude do gradiente, respectivamente.

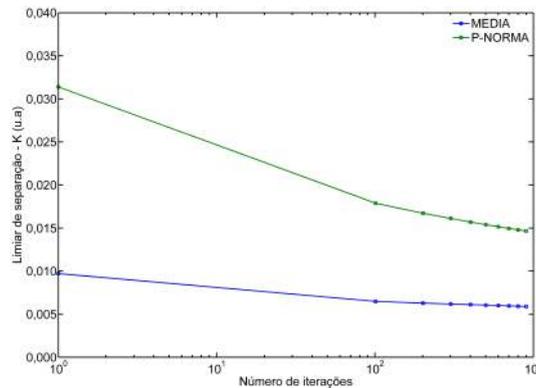


Figura 6.14: Evolução dos limiares calculados segundo média e 3-norma da magnitude do gradiente da amostra mostrada na Figura 6.13 (a) ao longo de 1000 iterações. As linhas contínuas servem apenas para visualização.

### 6.4.3 Tempo computacional

Uma das maiores dificuldades enfrentadas ao longo do trabalho correspondeu ao alto tempo de processamento da cadeia proposta. Por meio de observação sobre várias das amostras, percebemos que a etapa de *enhancement* via Equação de Perona-Malik é a mais custosa em termos computacionais, seguida das etapas de alto-nível (*labeling* e *clustering*). Observamos que imagens de muito alta-definição (alta densidade de pixels) e grandes dimensões consumiam um tempo computacional menor mas próximo de 1 h para completo processamento. Ressalte-se que esta é uma estimativa básica, sem maiores compromissos com acurácia.

Dentro deste cenário de longos tempos de processamento, fomos forçados a restringir os intervalos da variação de  $\alpha$ , tal como apresentado na Seção 6.4.1. Isso tolheu qualquer possibilidade de excursão em intervalos maiores deste parâmetro ou qualquer outro arranjo de configurações. Para se ter uma idéia, um dos primeiros experimentos – realizado *serialmente* – correspondia a processar as imagens da BIM em 12 diferentes configurações<sup>3</sup>. Esta rodada de processamento levou 30 dias para ser concluída. Com isso, pudemos estimar o tempo médio de processamento:  $\mathcal{T}_{\text{ser}} \approx 70$  min/imagem. Como chegamos a operar cada imagem em até 60 arranjos de configurações, tempos de processamento da ordem de meses começaram a ser cogitados. Eis a grande restrição sobre os arranjos possíveis da cadeia de processamento que forçou as análises prévias antes da utilização de etapa de *enhancement*.

Uma solução sugerida foi a de *paralelizar* o processamento. Fizemos uso da boa capacidade do computador usado no projeto e executamos simultaneamente cinco instâncias da cadeia de processamento sobre grupos de cinco imagens da BIM. Isto é, 25 imagens eram lançadas ao processamento, e após dois lances restava apenas uma última a ser processada ( $25+25+1$ ). Com a paralelização do processo, processávamos cada imagem com  $\mathcal{T}_{\text{par}} \approx 50$  min/imagem; uma redução de quase 30%. Embora o tempo de computação ainda tenha permanecido alto, pudemos finalmente variar os parâmetros das técnicas para realizar as análises desejadas.

<sup>3</sup>Esta foi a primeira tentativa de se obter os dados relativos aos experimentos I e II, descritos na Seção 6.2.

CANNY														
TELA-FILME						CR								
médias			desvios médios			médias			desvios médios					
$\alpha$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$	$\epsilon$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$	$\epsilon$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$
2	35,55	2,15	32,60	0,25	8,05%	13,62	1,45	14,22	0,45	2,19%	11,17	0,93	10,81	0,31
5	36,40	2,20	33,10	0,20	8,13%	14,02	1,44	14,63	0,34	2,21%	11,76	0,93	11,08	0,31
8	37,45	2,10	34,25	0,20	7,59%	13,59	1,51	14,35	0,34	2,25%	11,24	0,93	10,74	0,31
11	36,15	2,15	33,15	0,30	8,08%	13,21	1,47	13,70	0,48	2,18%	11,26	0,89	10,78	0,36
14	36,20	2,15	33,00	0,25	7,96%	12,86	1,47	13,50	0,40	2,27%	10,90	0,93	10,60	0,31
17	36,65	2,20	33,35	0,25	7,93%	12,65	1,52	13,36	0,40	2,25%	10,99	0,93	10,52	0,31
20	36,20	2,20	33,00	0,25	8,39%	13,56	1,52	14,20	0,40	2,12%	11,36	0,89	10,97	0,36

MÉDIA DA MAGNITUDE DO GRADIENTE														
TELA-FILME						CR								
médias			desvios médios			médias			desvios médios					
$\alpha$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$	$\epsilon$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$	$\epsilon$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$
2	36,35	1,95	33,25	0,40	7,38%	12,99	1,56	13,70	0,60	2,03%	11,17	0,93	10,81	0,31
5	34,95	2,20	31,90	0,25	8,54%	12,05	1,40	12,18	0,43	2,01%	11,76	0,93	11,08	0,31
8	35,45	2,15	31,95	0,30	8,17%	12,34	1,35	12,84	0,51	2,28%	11,24	0,93	10,74	0,31
11	34,15	2,05	31,15	0,40	7,72%	12,48	1,36	12,58	0,68	2,21%	11,26	0,89	10,78	0,36
14	34,60	2,05	31,75	0,40	8,22%	11,96	1,36	12,58	0,64	2,29%	10,90	0,93	10,60	0,31
17	35,85	2,15	32,85	0,30	8,02%	12,62	1,37	13,31	0,51	2,40%	10,99	0,93	10,52	0,31
20	35,15	2,20	32,10	0,25	8,29%	11,88	1,42	12,33	0,43	2,30%	11,36	0,89	10,97	0,36

P-NORMAS DA MAGNITUDE DO GRADIENTE														
TELA-FILME						CR								
médias			desvios médios			médias			desvios médios					
$\alpha$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$	$\epsilon$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$	$\epsilon$	$\bar{n}$	$\bar{N}_{VP}$	$\bar{N}_{FP}$	$\bar{N}_{FN}$
2	36,00	2,10	32,85	0,30	7,59%	12,50	1,51	13,02	0,51	2,96%	10,00	0,98	9,54	0,31
5	35,35	2,30	31,90	0,15	8,36%	11,56	1,50	12,47	0,27	2,55%	9,74	0,98	9,81	0,31
8	35,30	2,25	31,95	0,20	8,26%	12,02	1,45	12,53	0,34	2,83%	9,85	0,98	9,75	0,31
11	36,00	2,25	32,60	0,15	8,09%	12,30	1,48	13,08	0,27	2,78%	9,73	0,93	9,54	0,31
14	35,60	2,30	32,10	0,15	8,31%	12,12	1,50	12,73	0,27	2,90%	8,66	0,98	8,63	0,31
17	35,40	2,25	32,00	0,20	8,15%	11,94	1,45	12,80	0,34	2,88%	9,30	0,98	9,05	0,31
20	35,70	2,25	32,00	0,20	8,13%	11,41	1,45	12,40	0,34	2,89%	9,49	0,98	9,28	0,31

Tabela 6.17: Compilação dos dados dos experimentos com a filtragem via Equação de Perona-Malik segundo os três métodos de obtenção do limiar de separação.

## Capítulo 7

### Conclusões

O presente trabalho teve por finalidade investigar a influência da introdução da difusão não-linear baseada na Equação de Perona-Malik (EPM) em uma cadeia de detecção de calcificações em imagens mamográficas. Para tanto, a cadeia em si foi completamente desenvolvida, operando as imagens desde a segmentação da mama até a caracterização dos achados.

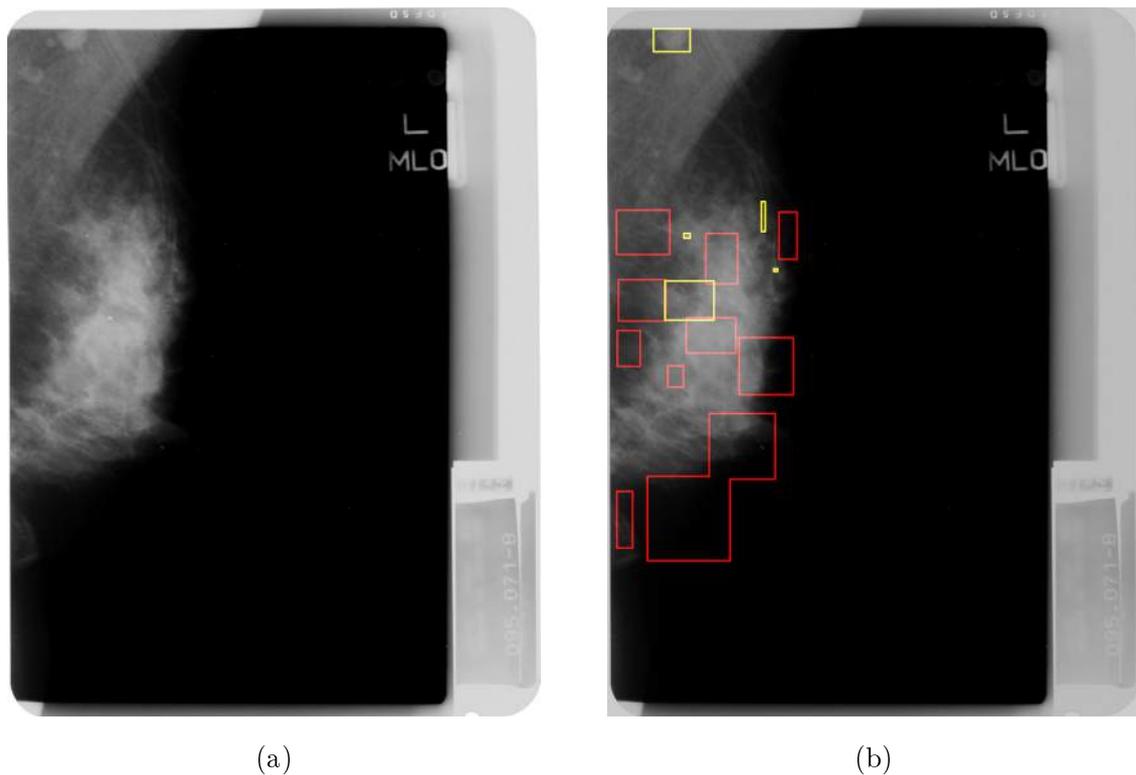


Figura 7.1: Ilustração da aplicação da Equação de Perona-Malik quando incorporada na cadeia de processamento. Em (a) é mostrado o mamograma original e em (b) o processado. As ROIs em amarelo correspondem às demarcações do médico, enquanto as ROIs em vermelho mostram as detecções da técnica. Os indicadores desta detecção são:  $m = 5$ ,  $n = 11$ ,  $N_{VP} = 4$ ,  $N_{FP} = 7$ ,  $\epsilon = 36,4\%$ . Efetivamente, apenas o linfonodo axilar não foi detectado.

Verificamos que a incorporação da EPM na cadeia de processamento trouxe melhorias ao processo de detecção: não houve ocorrência de imagens processadas sem ao menos 1

verdadeiro-positivo, quando aplicável. Nesse sentido, tanto a cadeia desenvolvida quanto o uso da EPM podem encontrar aplicação prática, desde que minimizados os efeitos dos ruídos. Uma ilustração desse tipo de ocorrência (boas contagens de  $N_{VP}$  acompanhadas de ruído) pode ser visualizado na Figura 7.1.

No que diz respeito à eficiência média de detecção, a introdução da difusão não-linear proporcionou um percentual máximo de 8,54% para o grupo de imagens tela-filme; para as imagens do grupo CR, a eficiência média máxima foi de 2,96%. Estes dois percentuais estão muito aquém do ideal (100%) e refletem a alta sensibilidade da cadeia de processamento a ruídos. Entretanto, convém observar que, em relação aos grupos de controle, o aumento máximo relativo foi de 10,5% para o grupo tela-filme e 35,16% para o grupo CR, mostrando que existe um *ganho real* em detecções quando do uso do enhancement via EPM. Contudo, a forte dependência da técnica sobre seus parâmetros constitui um empecilho em sua utilização efetiva.

Dificuldades em escolher a forma da função difusibilidade e do limiar K de separação foram evidenciadas: não foi observado comportamento monotônico para os indicadores quando da mudança da forma da difusibilidade, a menos da contagem de verdadeiros-positivos para imagens tela-filme, que mostrou explícita tendência de aumento, com coeficientes de correlação de Pearson de aproximadamente 0,40. Por outro lado, o número médio de falsos-positivos pouco mudou para todas as configurações.

No que diz respeito à forma do cômputo de K, verificou-se que o uso de p-normas da magnitude do gradiente levou a melhores eficiências médias do que as outras formas. Deve-se salientar que a forma utilizada é uma alteração da originalmente proposta, uma vez que as p-normas são calculadas sobre a magnitude do gradiente da imagem, e não sobre a imagem em si. Nesse sentido, análise posterior deve ocorrer para comprovar a eficiência desta forma de cálculo de K. A mesma análise se aplica à média da magnitude do gradiente, já que esta também apresentou bons resultados.

Outro detalhe relativo à utilização da EPM diz respeito ao longo tempo de processamento. A filtragem de imagens de grandes dimensões demandou tempos da ordem de 1h/imagem. Logo, a implementação de técnicas numéricas eficientes, e.g. métodos implícitos, deve ser uma das linhas futuras de trabalho.

Em termos globais, a inserção da EPM no processo parece não valer a pena frente aos inconvenientes citados. Entretanto, a EPM teve sua influência medida por meio dos resultados do processamento por toda a cadeia. Nesse sentido, a parametrização da cadeia foi extremamente influente nos resultados; efetivamente, a etapa de segmentação foi determinante para os resultados modestos levantados.

Devido à remoção da abertura morfológica do processo, o número de detecções espúrias subiu consideravelmente, diminuindo sua eficiência. Ocorre que muitos dos objetos segmentados correspondem na verdade a segmentos de linha, tal como ilustrado na Figura 7.2. De posse dessa constatação, percebe-se que o processo de segmentação deve, necessariamente, ser repensado: é preciso que este processo não gere como resultado as bordas relativas às diferenças de intensidade luminosa do tecido adjacente aos achados de interesse. A opção aparentemente mais natural a ser tomada consiste em se implementar um processo de segmentação *local*, em que uma certa vizinhança do pixel analisado seja usada para determinar o limiar. Com isso, minimiza-se a influência de bordas intensas alheias aos achados no cômputo do threshold.

A questão do pior desempenho da cadeia de processamento sobre imagens do grupo CR permaneceu em aberto. A única hipótese que nos pareceu razoável diz respeito à resolução espacial destas imagens: uma vez que sua resolução é grande comparativamente ao grupo

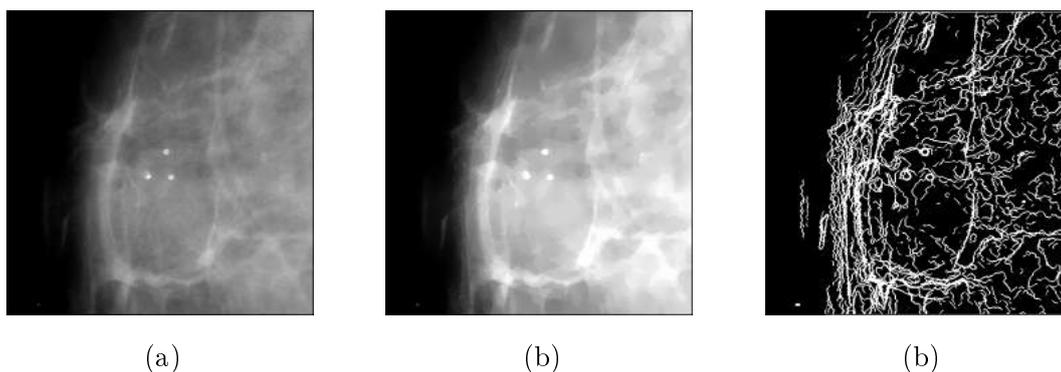


Figura 7.2: Ilustração do enhancement via Equação de Perona-Malik e segmentação por threshold entrópico sobre a magnitude do gradiente (não mostrada). Em (a) é mostrada a versão original, em (b) a filtrada e em (c) a versão binária.

tela-filme, estruturas pontuais acabavam sendo segmentados. Segundo esta hipótese, a redução da resolução destas imagens seria um recurso adicional, já que há muita informação espúria nativa sobre estas amostras. Dentro do escopo de descrição multiescala, a aplicação de filtragem via Gaussiana cumpre perfeitamente este papel, garantindo a manutenção da informação original<sup>1</sup>. Com isso, a utilização da Equação de Catté-Lions-Morel-Coll (Seção 4.3.3) tornar-se-ia justificada, uma vez que esta opera com uma regularização Gaussiana sobre o gradiente da imagem original. Outra alternativa diz respeito ao aumento do número de iterações, ou do tempo de difusão: um exemplo claro dessa alteração é aquele mostrado na Figura 6.13, em que os detalhes relativos às adjacências das calcificações foram *completamente* suprimidos, restando apenas a informação desejada. Entretanto, aumentar o tempo de difusão influi no tempo de processamento, sendo esta uma característica crítica, tal como discutido anteriormente.

Como propostas futuras, podemos citar a necessidade de otimização dos métodos numéricos e a readequação do processo de segmentação, além de um estudo envolvendo alteração na resolução das imagens. Adicionalmente, as técnicas aqui desenvolvidas/avaliadas poderiam ser aplicadas sobre imagens simuladas, pois nestes ambientes há um controle absoluto dos parâmetros de interesse, tais como densidade do tecido, número de achados e posição destes. Ademais, o uso de simulação serviria para confrontar os indicadores do presente trabalho e ajudar a caracterizar quantitativamente a base de imagens usada, já que aqueles são indicadores médios em relação a uma base específica. Ainda sobre a base usada (BIM), cabe dizer que esta é diminuta em relação a outras bases investigadas, o que deve ser considerado nas estatísticas.

<sup>1</sup>Informação original no sentido de não-corrompida por efeitos de *aliasing*.

# Bibliografia

- [1] ALBUQUERQUE, M. P.; ALBUQUERQUE, M. P.; MELLO, A. G.; CANER, E. *Análise de imagens e visão computacional*. In: RODRIGUES, L. M. C. S. (Coord.); REIS, A. C. dos; CURADO, E. M.; TERRA, J.; ALVES JR.; N. A. (Edits.). **Anais da V Escola do CBPF**. Rio de Janeiro: CBPF, 2005. p. 145–176.
- [2] ALVAREZ, L.; MAZORRA, L. *Signal and image restoration using shock filters and anisotropic diffusion*. **SIAM Journal on Numerical Analysis**. v. 31. n. 2. 590–605. 1994.
- [3] ARANCIBIA H., P.; TAUB E., T.; GRAZIA K., J. A. de; DÍAZ C., M. L.; LÓPEZ P., A.; SAÉZ T. C.; CERMENATI B., T. *Microcalcificaciones mamarias: revisión de los descriptores y categorías BI-RADS*. **Revista chilena de obstetricia y ginecología**. v. 78. n. 5. p. 383–394. 2013.
- [4] ARFKEN, G. B.; WEBER, H. J. *Mathematical methods for physicists*. 6 ed. Oxford: Elsevier, 2005. 1182 p.
- [5] AUBERT, G.; KORNPORST, P. *Mathematical problems in image processing: partial differential equations and the calculus of variations*. (Applied mathematical sciences . v. 147). New York: Springer, 2001. 286 p.
- [6] BARCELOS, C.A.Z. *Restauração e análise de imagens via equações diferenciais parciais*. **Tendências em matemática aplicada e computacional**. v.3. n. 2. 1–13. 2002.
- [7] BENHAMOUDA, B. *Parameter adaptation for nonlinear diffusion in image processing*. Dissertation (Magister of Scientiae in Mathematics). Department of Mathematics, University of Kaiserslautern. 1994.
- [8] BERGER, M.J.; HUBBELL, J.H.; SELTZER, S.M.; CHANG, J. S.; COURSEY, J.S.; SUKUMAR, R.; ZUCKER, D.S.; OLSEN, K. *XCOM: Photon Cross Sections Database*. NIST Standard Reference Database 8 (XGAM). Gaithersburg, Maryland, USA. Disponível em <http://www.nist.gov/pml/data/xcom/index.cfm>. Acesso em Julho/2015.
- [9] BIDDLE, H. *Nonlinear diffusion filtering on surfaces*. Dissertation (Master of Scientiae in Mathematical Modelling and Scientific Computing). Hertford College, University of Oxford. 2011.
- [10] BLACK, M. J.; SAPIRO, G.; MARIMONT, D.H.; HEEGER, D. *Robust anisotropic diffusion*. **IEEE Transactions on Image Processing**. v. 7. n. 3. 421–432. 1998.

- [11] BOZEK, J.; MUSTRA, M.; DELAC, K.; GRGIC, M. (Edit.) *A survey of image processing algorithms in digital mammography*. **Recent advances in Multimedia Signal Processing and Communications**. v. 231. 631–657. 2009.
- [12] BRASIL. MINISTÉRIO DA SAÚDE. *Portaria MS 2 048/2009, de 03/09/2009 - Regulamentação do Sistema Único de Saúde*. Disponível em [http://bvsms.saude.gov.br/bvs/saudelegis/gm/2009/prt2048\\_03\\_09\\_2009.html](http://bvsms.saude.gov.br/bvs/saudelegis/gm/2009/prt2048_03_09_2009.html). Acesso em Julho/2015.
- [13] BRASIL. MINISTÉRIO DA SAÚDE. INSTITUTO NACIONAL DE CÂNCER. *Controle do câncer de mama: documento de consenso*. Disponível em <http://www1.inca.gov.br/publicacoes/ConsensoIntegra.pdf>. Acesso em Julho/2015.
- [14] BRASIL. MINISTÉRIO DA SAÚDE. INSTITUTO NACIONAL DE CÂNCER. *Estimativa 2014: Incidência de Câncer no Brasil*. Rio de Janeiro: INCA, 2014. 124 p.
- [15] BRASIL. MINISTÉRIO DA SAÚDE. INSTITUTO NACIONAL DE CÂNCER. *Mamografia: da prática ao controle (recomendações para profissionais de saúde)*. Disponível em [www1.inca.gov.br/inca/Arquivos/qualidade\\_em\\_mamografia.pdf](http://www1.inca.gov.br/inca/Arquivos/qualidade_em_mamografia.pdf). Acesso em Julho/2015.
- [16] CALAS, M. J. G.; GUTFILEN, B.; PEREIRA, W. C. A. *CAD and mammography: why use this tool?*. **Radiologia Brasileira**. v. 45. n. 1. 46–52. 2012.
- [17] CALDAS, F. A. A.; ISA, H. L. V. R.; TRIPPIA, A. C.; BÍSCARO, A. C. F. P. J.; SOUZA, E. C. C.; TAJARA, L. M. *Controle de qualidade e artefatos em mamografia*. **Radiologia brasileira**. v. 38. 295–300. 2005.
- [18] CANNY, J. *A computational approach to edge detection*. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. v. 8. n. 6. 679–698. 1986.
- [19] CATTÉ, F.; LIONS, P.L.; MOREL, J.M.; COLL, T. *Image selective smoothing and edge detection by nonlinear diffusion*. **SIAM Journal on Numerical Analysis**. v. 29. n. 1. pp. 182–193. 1992.
- [20] CERDA, J.; CIFUENTES, L. *Uso de curvas ROC en investigación clínica. Aspectos teórico-prácticos*. **Revista Chilena de Infectología**. v. 29. n. 2. 138–141. 2012.
- [21] CHANG, C. I.; DU, Y.; WANG, J.; GUO, S. M.; THOUIN, P. D. *Survey and comparative analysis of entropy and relative entropy thresholding techniques*. **IEEE proceedings on vision, image and signal processing**. v. 153. n. 6. 837–850. 2006
- [22] CHAO, S.; TSAI, D.; LI, W.; CHIU, W. *A generalized anisotropic diffusion for detection in low-contrast surfaces*. **20th International Conference on Pattern Recognition (ICPR)**. p. 4408–4411. 2010.
- [23] CHENG, H. D.; CAI, X.; CHEN, X.; HU, L.; LOU, X. *Computer-aided detection and classification of microcalcifications in mammograms: a survey*. **Pattern Recognition**. v. 36. n. 12. 1967–2991. 2003.
- [24] COMPTON, A. H.; ALLISON, S. K. *X-rays in theory and experiment*. 2 ed. New York: D. Van Nostrand Company, Inc., 1935. 828 p.

- [25] CUMINATO, J.A.; MENEGUETE JR., M. *Discretização de equações diferenciais parciais*. 1 ed. Rio de Janeiro: SBM, 2013. 379p.
- [26] DeGROOT, M.H.; SCHERVISH, M. J. *Probability and statistics*. 4 ed. Boston: Pearson, 2012. 893 p.
- [27] DINIZ, P.S.R.; SILVA, E.A.B. da; NETTO, S.L. *Digital signal processing - system analysis and design*. 1 ed. UK: Cambridge University Press. 2002.
- [28] FIGUEIREDO, D.G. de. *Análise de Fourier e equações diferenciais parciais*. 4 ed. Rio de Janeiro: IMPA, 2012. 274p.
- [29] GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. 2 ed. New Jersey: Prentice Hall, 2002. 793 p.
- [30] GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. *Digital image processing using MATLAB<sup>®</sup>*. 2 ed. USA: Gatesmark Publishing, 2009. 827 p.
- [31] GOTOH, T.; OHTA, Y.; YOSHIDA, M.; SHIRAI, Y. *Component labeling algorithm for video rate processing*. **Proceedings SPIE, Advances in Image Processing**. v. 0804. 217–224. 1987.
- [32] GRIFFITHS, D. J. *Introduction to electrodynamics*. 3 ed. New Jersey: Prentice Hall, 1999. 576 p.
- [33] GUICHARD, F.; MOISAN, L.; MOREL, J.M. *A review of P.D.E models in image processing and image analysis*. **Journal de Physique IV (Proceedings)**. v. 12. n. 1. 137–154. 2002.
- [34] HARALICK, R. M.; STERNBERG, S. R.; ZHUANG, X. *Image analysis using mathematical morphology*. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. v. 9. n. 4. 532–550. 1987.
- [35] HE, L.; CHAO, Y.; SUSUKI, K.; WU, K. *Fast connected-component labeling*. **Pattern Recognition**. v. 42. n. 9. 1977–1987. 2009.
- [36] HIGHNAM, R.; BRADY, M. *Mammographic image analysis*. (Computational Imaging and Vision Series. v. 14). Springer, 2012. 379 p.
- [37] HOBBIIE, R. K.; ROTH, B. J. *Intermediate physics for medicine and biology*. 4 ed. New York: Springer, 2006. 616 p.
- [38] HOHEISEL, M. *Review of medical imaging with emphasis on X-ray detectors*. **Nuclear Instruments and Methods in Physics Research A**. v. 563. n. 1. 215–224. 2006.
- [39] IÓRIO JR., R.; IÓRIO, V. de. *Equações diferenciais parciais: uma introdução*. 3 ed. Rio de Janeiro: IMPA, 2013. 343p.
- [40] INTERNATIONAL COMMISSION ON RADIATION UNITS AND MEASUREMENTS. *Mammography - assessment of image quality*. **Journal of the ICRU (ICRU Report 82)**. v. 9. n. 2. 2009.

- [41] INTERNATIONAL COMMISSION ON RADIATION UNITS AND MEASUREMENTS. *Receiver Operating Characteristic Analysis in medical imaging*. **Journal of the ICRU** (ICRU Report 79). v. 8. n. 1. 2008.
- [42] JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. *Data Clustering: A Review*. **ACM Computing Surveys**. v. 31. n. 3. 264–323. 1999.
- [43] JAIN, A.K.; DUBES, R.C. *Algorithms for clustering data*(Prentice-Hall advanced reference series). 1 ed. New Jersey: Prentice-Hall. 1988.
- [44] JAMES, J.J. *The current status of digital mammography*. **Clinical Radiology**. v. 59. n. 1. 1–10. 2004.
- [45] KAPUR, J. N.; SAHOO, P. K.; WONG, A. K. C. *A new method for gray-level picture thresholding using the entropy of the histogram*. **Computer vision, graphics and image processing**. v. 29. 273-285. 1985.
- [46] KOENDERINK, J. J. *The structure of images*. **Biological cybernetics**. v. 50. pp. 363–370. 1984.
- [47] KOPANS, D. B. *Diagnóstico por imagem da mama*. 3 ed. Rio de Janeiro: Guanabara Koogan, 2008. 1048 p.
- [48] KRAMERS, H. A. *On the theory of X-ray absorption and the continuous X-ray spectrum*. **Philosophical Magazine** (Series 6). v. 46. n. 275. 836–871. 1923.
- [49] LAW, J. *The development of mammography*. **Physics in Medicine and Biology**. v. 51. n. 13. R 155.
- [50] LIMA, E. L. *Álgebra linear*.(Coleção Matemática Universitária). 8 ed. Rio de Janeiro: IMPA. 2011. 357 p.
- [51] LINDENBAUM, M.; FISCHER, M.; BRUCKSTEIN, A. *On Gabor’s contribution to image enhancement*. **Patter Recognition**. v. 27. n; 1. pp. 1–8. 1994.
- [52] LINGURARU, M. G.; BRADY, M. *An anisotropic diffusion approach for early detection of breast cancer*. **Acta Universitatis Cibiniensis, Technical Series**. v. XLIII. pp. 49–60. 2001.
- [53] LYONS, R. G. *Understanding digital signal processing*. 1 ed. Massachusetts: Pearson Education, 1997. 517p.
- [54] MARTINS, L. A. L.; BARRA, A. A.; LUCENA, C. E. M. *Microcalcificações mamárias suspeitas de malignidade*. **Revista Brasileira de Cancerologia**. v. 56. n. 2. 251–258. 2010.
- [55] MATHEUS, B. R. N. *BancoWeb: base de imagens mamográficas para auxílio em avaliações de esquemas CAD*. Dissertação (Mestrado - Programa de Pós-graduação em Engenharia Elétrica). Escola de Engenharia de São Carlos, Universidade de São Paulo. 2010.
- [56] MIKULA, K. *Image processing with partial differential equations*. In: BOURLIOUX, A.; GANDER, M.J.; SABIDUSSI, G. (Edits.). **Modern methods in scientific computing and applications** (NATO Science Series. II - Mathematics, Physics and Chemistry). v. 75. pp. 283–321. Dodrecht: Kluwer Academic Publishers. 2002.

- [57] MORFU, S. *Image processing using diffusion processes*. **Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)**.11811–1814. 2010.
- [58] OKUNO, E.; YOSHIMURA, E. M. *Física das radiações*. 1 ed. São Paulo: Oficina de Textos, 2010. 296 p.
- [59] OSHER, S.; RUDIN, L. I. *Feature-oriented image enhancement using shock filters*. **SIAM Journal on Numerical Analysis**. v. 27. n. 4. 919–940. 1990.
- [60] OTSU, N. *A threshold selection method from gray-level histograms*. **IEEE transactions on systems, man, and cybernetics**. v. 9. n. 1. 62–66. 1979.
- [61] PAVLIDIS, T. *Algorithms for shape analysis of contours and waveforms*. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. v. 2. n. 4. 301–312. 1980.
- [62] PEDRINI, H.; SCHWARTZ, W.R. *Análise de imagens digitais*. São Paulo: Thomson Learning, 2008. 508 p.
- [63] PERONA, P.; MALIK, J. *Scale-space and edge detection using anisotropic diffusion*. **IEEE Transactions on Pattern Analysis and Machine Intelligence** v. 12. n.7. 629–639. 1990.
- [64] PERONA, P.; SHIOTA, T.; MALIK, J. *Anisotropic diffusion*. In: ROMENY, B.M. t H. (Edit.) **Geometry-driven diffusion in computer vision**(Computational imaging and vision, v.1). Netherlands: Springer, 1994. pp. 73–92.
- [65] PERSECHINO, A. *Estudo e implementação de técnicas de processamento de imagens aplicadas à segmentação de imagens mamográficas*. Monografia - Graduação em física. Universidade Federal de Viçosa. 2012.
- [66] PODGORŠAK, E. B. *Radiation physics for medical physicists* (Biological and medical physics, biomedical engineering series). 2 ed. Berlin: Springer-Verlag, 2010. 745 p.
- [67] POLLAK, I. *Segmentation and restoration via nonlinear multiscale filtering*. **IEEE Signal Processing Magazine**. v. 19. 26–36. 2002.
- [68] POLLAK, I.; WILLSKY, A.S.; KRIM, H. *Image Segmentation and Edge Enhancement with Stabilized Inverse Diffusion Equations* . **IEEE Transactions on Image Processing**. v. 9. n. 2. 256–266. 2000.
- [69] PREWITT, J. M. S. *Object enhancement and extraction*. In: LIPKIN, B. S.; ROSENFELD, A. (Edits.). **Picture processing and psychopictorics**. New York: Academic, 1970. p. 75–149.
- [70] RASBAND, W. S. *ImageJ*. National Institutes of Health, Bethesda, Maryland, USA, <http://imagej.nih.gov/ij/>, 1997-2014. Acesso em Novembro/2014.
- [71] RITCHIE, N. W. M. *NIST DTSA II - power tool for X-ray microanalysis*. Gaithersburg, Maryland, USA. Disponível em <http://www.cstl.nist.gov/div837/837.02/epq/dtsa2/index.html>. Acesso em Julho /2015.

- [72] ROSENFELD, A.; KAK, A.C. *Digital picture processing* (Computer science and applied mathematics series). vols. 1 e 2. New York: Academic Press. 1982.
- [73] SÄBEL, M.; AICHINGER, H. *Recent developments in breast imaging. Physics in Medicine and Biology*. v. 41. n. 3. 315–368. 1996.
- [74] SAHOO, P. K.; SOLTANI, S.; WONG, A. K. C. *A survey of thresholding techniques. Computer vision, graphics, and image processing*. v. 41. 2 ed. 233–260. 1988.
- [75] SANTORO, A.; MAHON, J. R.; OLIVEIRA, J. U. C. L. de; MUNDIM FILHO, L. M.; SILVA, W. L. P. da; OGURI, V. (org.). *Estimativas e erros em experimentos de física*. 2 ed. Rio de Janeiro: EDUERJ, 2008. 131 p.
- [76] SANTOS, A. F. *Física médica em mamografia*. 1 ed. Rio de Janeiro: Revinter, 2010. 480 p.
- [77] SCHERER, C. *Métodos computacionais da física*. 1 ed. São Paulo: Editora Livraria da Física, 2005. 284 p.
- [78] SEPTIANA, L.; LIN, K. *X-ray image enhancement using a modified anisotropic diffusion. IEEE International Symposium on Bioelectronics and Bioinformatics (ISBB)*. 2014.
- [79] SHAO, H.; ZOU, H. *Threshold Estimation Based on Perona-Malik Model. International Conference on Computational Intelligence and Software Engineering, 2009. CiSE 2009*. pp. 1–4. 2009.
- [80] SILVA JR., E. C. *Modelo de processamento de imagens mamográficas para detecção de agrupamento de microcalcificações*. Tese (Doutorado - Programa de Pós-graduação em Engenharia Elétrica). Escola de Engenharia de São Carlos, Universidade de São Paulo. 2009.
- [81] SPORRING, J.; WEICKERT, J. *Information measures in scale-spaces. IEEE transactions on information theory*. v. 45. n. 3. 1051–1058. 1999.
- [82] SUCKLING, J.; PARKER, J.; DANCE, D.R.; ASTLEY, S.; HUTT, I.; BOGGIS, C.R.M.; RICKETTS, I.; STAMATAKIS, E.; CERNEAZ, N.; KOK, S.L.; TAYLOR, P.; BETAL, D.; SAVAGE, J. *The Mammographic Image Analysis Society Digital Mammogram Database. Exerpta Medica. International Congress Series*. v. 1069. pp. 375–378. 1994.
- [83] SUCKLING, J.; PARKER, J.; DANCE, D.R.; ASTLEY, S.; HUTT, I.; BOGGIS, C.R.M.; RICKETTS, I.; STAMATAKIS, E.; CERNEAZ, N.; KOK, S.L.; TAYLOR, P.; BETAL, D.; SAVAGE, J. *The mini-MIAS database of mammograms*. Department of Physics, Royal Marsden Hospital, London, UK. <http://peipa.essex.ac.uk/info/mias.html>. Acesso em Abril/2015.
- [84] THE MATHWORKS, INC. *MATLAB®. The language of technical computing*. Natick, Massachusetts, USA. Disponível em <http://mathworks.com/products/matlab>, 1994-2015. Acesso em Junho/2015.
- [85] TREMBLAIS, B.; AUGEREAU, B. *A fast n-dimensional non-linear filter. Proceedings of SPIE, Medical Imaging*. v. 5370. 2130–2138. 2004.

- [86] TSLOTSIOS, C.; PETROU, M. *On the choice of the parameters for anisotropic diffusion in image processing*. **Pattern Recognition**. v. 46. 5 ed. 1369–1381. 2013.
- [87] VOGLI, F.; EIHO, S.; SUGIMOTO, N.; SEKIBUCHI, H. *Estimating the gradient threshold in the Perona-Malik equation*. **IEEE Signal Processing Magazine**. v. 21. 3 ed. 39–65. 2004.
- [88] WEERATUNGA, S.K.; KAMATH, C. *PDE-based non-linear diffusion techniques for denoising scientific and industrial images: an empirical study*. Disponível em <http://www.llnl.gov/tid/Library.html>. Acesso em março/2014.
- [89] WEICKERT, J. *Anisotropic diffusion in image processing*. B.G. Teubner: Stuttgart, 1998.
- [90] WEICKERT, J.; BENHAMOUDA, B. *Why the Perona-Malik filter works?* Technical report. Department of Computer Science, University of Copenhagen. 1997.
- [91] WEICKERT, J.; ROMENY, B.M.T.H.; VIERGEVER, M.A. *Efficient and reliable schemes for nonlinear diffusion filtering*. **IEEE Transactions on Image Processing**, v. 17. n. 3. 398–410. 1998.
- [92] WESZKA, J. S. *Survey of Threshold Selection Techniques*. **Computer Graphics and Image Processing**. v. 7. n. 2. 259–265. 1978.
- [93] WITKIN, A.P. *Scale-space filtering: a new approach to multi-scale description*. **Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84**, v. 9. 150–153. 1984.
- [94] YOSHIMURA, E. M. *Física das Radiações: interação da radiação com a matéria*. **Revista Brasileira de Física Médica**. v. 3. n. 1. 57–67. 2009.
- [95] ZHANG, D.; GUOJUN, L. *Review of shape representation and description techniques*. **Pattern Recognition**. v. 37. 1 ed. 1–19. 2004
- [96] ZIOU, D.; TABBONE, S. *Edge Detection Techniques - an overview*. **International Journal of Pattern Recognition and Image Analysis**. v. 8. 537–559. 1998.

## Apêndice A

# Demonstração do resultado para o problema de valor inicial da equação do calor

Analisemos o seguinte problema de condução de calor em uma dimensão:

$$\begin{aligned}\partial_t u(x, t) &= K \partial_{xx} u(x, t); \text{ com } K \in \mathbb{R}^+ \\ u(x, 0) &= p(x),\end{aligned}\tag{A.1}$$

Em que  $p(x)$  é contínua e limitada. Queremos verificar que a solução da Equação (A.1) é dada por

$$u(x, t) = u(x, 0) * G(x, t),\tag{A.2}$$

em que

$$G(x, t) = \frac{1}{\sqrt{2Kt}} e^{-\left(\frac{x}{2\sqrt{Kt}}\right)^2}.$$

Na Equação (A.2) o asterisco remete ao produto de convolução entre as funções  $u(x, t)$  e  $G(x, t)$ . Para verificar que esta é realmente a solução, tomemos a transformada de Fourier na Equação (A.1):

$$\begin{aligned}\partial_t \mathcal{U}(\omega, t) &= -K\omega^2 \mathcal{U}(\omega, t) \\ \mathcal{U}(\omega, 0) &= \mathcal{P}(\omega).\end{aligned}\tag{A.3}$$

A equação que aparece em (A.3) é de fácil resolução e seu resultado é dado por:

$$\mathcal{U}(\omega, t) = \mathcal{P}(\omega) e^{-K\omega^2 t}.\tag{A.4}$$

Esta equação pode ser trazida de volta para o domínio espacial por meio da transformada inversa de Fourier:

$$u(x, t) = u(x, 0) * \mathcal{F}^{-1}\left(e^{-K\omega^2 t}\right),\tag{A.5}$$

em que  $\mathcal{F}^{-1}(\cdot)$  denota a transformada inversa de Fourier. O trabalho agora resume-se a calcular  $\mathcal{F}^{-1}\left(e^{-K\omega^2}\right)$ . Da definição de transformada inversa de Fourier, temos que

$$\mathcal{F}^{-1}\left(e^{-K\omega^2}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-K\omega^2 t} e^{j\omega x} d\omega; \text{ com } j = \sqrt{-1}.\tag{A.6}$$

Podemos completar o quadrado nos expoentes:

$$-K\omega^2 t + j\omega x = -\left(\sqrt{Kt}\omega - \frac{jx}{2\sqrt{Kt}}\right)^2 + \left(\frac{jx}{2\sqrt{Kt}}\right)^2.$$

Com isso, a inversa toma a forma

$$\mathcal{F}^{-1}\left(e^{-K\omega^2 t}\right) = \frac{1}{\sqrt{2\pi}} e^{\left(\frac{jx}{2\sqrt{Kt}}\right)^2} \int_{-\infty}^{+\infty} e^{-\left(\sqrt{Kt}\omega - \frac{jx}{2\sqrt{Kt}}\right)^2} d\omega. \quad (\text{A.7})$$

Para resolver a expressão contida em (A.7), fazemos uso da troca de variáveis  $u = \sqrt{Kt}\omega - \frac{jx}{2\sqrt{Kt}}$  e do fato de que  $\int_{-\infty}^{+\infty} e^{-u^2} du = \sqrt{\pi}$ . Com isso, chegamos em

$$\mathcal{F}^{-1}\left(e^{-K\omega^2 t}\right) = G(x, t) = \frac{1}{\sqrt{2Kt}} e^{-\left(\frac{x}{2\sqrt{Kt}}\right)^2}, \quad (\text{A.8})$$

que prova a solução (A.2).

Para o caso particular de  $K = 1$ , temos  $G(x, t) = \frac{1}{\sqrt{2t}} e^{-\left(\frac{x}{2\sqrt{t}}\right)^2}$ , que evidencia o valor da variância para a difusão linear,  $\sigma^2 = 2t$ .

## Apêndice B

# Obtenção do *threshold* para binarização

Neste apêndice serão descritos os três métodos para estimativa do *threshold* para binarização de uma imagem. Ao final é apresentado um código em MATLAB<sup>®</sup> que implementa uma função para determinação do *threshold* de uma imagem qualquer.

O ponto de partida é o histograma de uma imagem, tal como discutido na Seção 3.1. Recordando a Equação (3.3), temos:

$$h(k) = \frac{p_k}{M \times N}.$$

em que  $\frac{p_k}{M \times N}$  é a probabilidade de ocorrência do  $k$ -ésimo nível de intensidade. Ou seja, o histograma da imagem é uma aproximação à função de densidade de probabilidade (discreta) dos seus níveis de intensidade. Dito isto, temos que todos os parâmetros estatísticos de interesse podem ser obtidos a partir do histograma.

### Valor médio

Dado o histograma  $h$  de uma imagem com  $N$  níveis de cinza, o *valor médio* para os níveis de intensidade é definido como

$$\bar{x} = \sum_{i=1}^N ih(i). \tag{B.1}$$

Tal como definido na Equação (B.1), percebemos que  $\bar{x}$  corresponde, de fato, ao centro de massa da distribuição aproximada por  $h$ . Nesse sentido, se o histograma for simétrico em relação ao ponto  $i_0$  – isto é,  $h(i_0 - k) = h(i_0 + k)$  –, teremos  $\bar{x} = i_0$ . Portanto, o uso do valor médio como *threshold* é apropriado para distribuições simétricas de níveis de cinza.

### Método de Otsu

Em 1979, N. Otsu propôs uma técnica de limiarização global erguida sobre a hipótese de que o histograma da imagem pode ser dividido em duas classes distintas [60]. Este método calcula o *threshold*  $L$  como sendo o nível de intensidade que maximiza uma função

critério  $\eta(L)$  definida como sendo a razão entre a variância entres classes e a variância global, isto é:

$$\eta(L) = \frac{\sigma_B^2}{\sigma_T^2}, \quad (\text{B.2})$$

em que as variâncias entre-classes ( $\sigma_B^2$ ) e total ( $\sigma_T^2$ ) têm a forma

$$\sigma_B^2 = \omega_0\omega_1(\mu_0\mu_1)^2 \quad \text{e} \quad \sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 h_i. \quad (\text{B.3})$$

Na Equação (B.3),  $\omega_0$  e  $\omega_1$  são as probabilidades de ocorrência das classes 0 e 1, e são dadas por

$$\omega_0 = \sum_{i=1}^L h_i \quad \text{e} \quad \omega_1 = 1 - \omega_0.$$

As médias das classes 0 ( $\mu_0$ ) e 1 ( $\mu_1$ ) são dadas por

$$\mu_0 = \frac{\mu_S}{\omega_0} \quad \text{e} \quad \mu_1 = \frac{\mu_T - \mu_S}{\omega_1},$$

em que

$$\mu_T = \sum_{i=1}^N i h_i \quad \text{e} \quad \mu_S = \sum_{i=1}^L i h_i.$$

O threshold deve ser escolhido como sendo o valor  $L^*$  que maximiza a função critério dada pela Equação (B.2), ou seja,

$$L^* = \operatorname{argmax} \{ \eta(L) \}. \quad (\text{B.4})$$

Conforme dito antes, este método parte da premissa que a imagem é formada por duas classes distintas (0 e 1, segundo a notação adotada). Contudo, histogramas puramente bimodais não são frequentes e este fato deve ser observado, pois influenciará o resultado da limiarização.

## Método da máxima entropia

A entropia<sup>1</sup> do histograma  $h$  é definida como

$$S = - \sum_{i=1}^N h_i \ln(h_i). \quad (\text{B.5})$$

A entropia de uma fonte fornece uma medida da taxa com que informação é produzida por essa fonte. As abordagens entrópicas para limiarização operam no sentido de estimar

<sup>1</sup>Essa definição é conhecida como entropia de Shannon, que é a equivalente em teoria da informação à entropia de Boltzmann na termodinâmica. Há outras entropias, tais como Entropia Generalizada, Entropia de Tsallis etc.

um threshold que maximize a entropia, promovendo assim a máxima transferência de informação possível da fonte em questão (a imagem).

Uma maneira simples de se obter o threshold pela maximização da entropia foi proposto por Kapur *et al.* [45]: deve-se maximizar uma função critério dada pela soma das entropias das duas classes (A e B, segundo a notação original adotada):

$$S_T(L) = S_A + S_B, \quad (\text{B.6})$$

em que, segundo a Equação (B.5), temos

$$S_A = - \sum_{i=1}^L \frac{h_i}{P_A} \ln \left( \frac{h_i}{P_A} \right) \quad \text{e} \quad S_B = - \sum_{i=L+1}^N \frac{h_i}{P_B} \ln \left( \frac{h_i}{P_B} \right). \quad (\text{B.7})$$

Na Equação (B.7),  $P_A$  e  $P_B$  são as probabilidades de ocorrência das classes A e B, respectivamente:

$$P_A = \sum_{i=1}^L h_i \quad \text{e} \quad P_B = 1 - P_A.$$

De maneira análoga ao método de Otsu, o threshold  $L^*$  é escolhido como sendo aquele que maximiza a função critério  $S_T$ , tal como definida na Equação (B.6):

$$L^* = \operatorname{argmax} \{S_T(L)\}. \quad (\text{B.8})$$

## Implementação computacional

```

% |-----|
% |          Centro Brasileiro de Pesquisas Fisicas - CBPF
% |          Coordenacao de Atividades Tecnicas - CAT
% |
% |          limiar.m
% |
% |   Descricao:   Retorna o valor de limiar para binarizacao
% |               segundo o metodo escolhido. Os parametros de
% |               entrada sao a imagem x (double) e a string str
% |               com o metodo de obtencao do threshold:
% |               - 'med' calcula o valor medio da imagem;
% |               - 'ent' calcula o valor que maximiza a entropia
% |               de Shannon da distribuicao da imagem;
% |               - 'ots' calcula o limiar segundo o metodo de
% |               Otsu.
% |-----|

function y = limiar(x, str)
    s = size(x);
    N_bins = 256;

    h = imhist(x, N_bins);
    h = h/(s(1)*s(2));

    switch lower(str)
        case 'med'
            y = Media(h);
        case 'ent'
            y = MaxEntropia(h);
        case 'ots'
            y = Otsu(h);
        otherwise
            y = Media(h);
    end

end

function y = Media(h)
    N_bins = length(h);
    y = linspace(0, 1, N_bins)*h;
end

function y = MaxEntropia(h)
    N_bins = length(h);
    S = zeros(1, N_bins); % funcao criterio

    % os proximos passos garantem que nao ocorra entrada nula no histograma
    % avaliado, para evitar divergencias. As entradas nulas sao trocadas
    % pelo menor valor nao-nulo da distribuicao.
    hMin = min(h(h~=0));

    hEnt = h;
    hEnt(hEnt==0) = hMin;
    hEnt = hEnt/sum(hEnt);

    for k=1:N_bins
        PA = sum(hEnt(1:k)); % probab. da classe A
        PB = 1-PA; % probab. da classe B

        % entropia classe A
        SA = -hEnt(1:k)*log(hEnt(1:k)/PA)/PA;
        % entropia classe B
        SB = -hEnt(k+1:N_bins)*log(hEnt(k+1:N_bins)/PB)/PB;

        S(k) = SA+SB;
    end
end

```

```

    arg = find(S==max(S));
    y = (arg-1)/(N_bins-1);
end

function y = Otsu(h)
    N_bins = length(h);
    C = zeros(1,256); % funcao criterio

    for k=1:N_bins
        w0 = sum(h(1:k)); % probab. da classe 0
        w1 = 1-w0; % probab. da classe 1
        u0 = ((0:k-1)*h(1:k))/w0; % media da classe 0
        u1 = ((k:N_bins-1)*h(k+1:N_bins))/w1; % media da classe 1
        ut = linspace(0,1,N_bins)*h; % media total

        desv_med_quad = ((0:N_bins-1)-ut).^2;

        sb = w0*(u0-ut)^2+w1*(u1-ut)^2; % variancia entre-classes
        st = desv_med_quad*h; % varianca total

        C(k) = sb/st;
    end
    arg = find(C==max(C));
    y = (arg-1)/(N_bins-1);
end

```

Código B.1: Função para obtenção do threshold para binarização de uma imagem segundo o método de Otsu, Máxima Entropia e Valor Médio.