

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

MONICA MARCUZZO

**Quantificação de Impressões Diagnósticas
em Imagens de Cintilografia Renal**

Dissertação apresentada como requisito parcial
para a obtenção do grau de
Mestre em Ciência da Computação

Prof. Dr. Jacob Scharcanski
Orientador

Porto Alegre, junho de 2007

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Marcuzzo, Monica

Quantificação de Impressões Diagnósticas em Imagens de Cintilografia Renal / Monica Marcuzzo. – Porto Alegre: PPGC da UFRGS, 2007.

74 f.: il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR-RS, 2007. Orientador: Jacob Scharcanski.

1. Processamento de imagens. 2. Imagens médicas. 3. Cintilografia renal. 4. Quantificação de informações visuais. I. Scharcanski, Jacob. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora: Prof^a. Wrana Maria Panizzi

Vice-Reitor de Ensino: Prof. José Carlos Ferraz Hennemann

Pró-Reitora de Pós-Graduação: Prof^a. Valquíria Linck Bassani

Diretor do Instituto de Informática: Prof. Flávio Rech Wagner

Coordenadora do PPGC: Prof^a. Luciana Porcher Nedel

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	5
LISTA DE FIGURAS	6
LISTA DE TABELAS	8
RESUMO	9
ABSTRACT	10
1 INTRODUÇÃO	11
1.1 Impressões diagnósticas em cintilografias renais	12
1.2 Estrutura da dissertação	13
2 TÉCNICAS DE PROCESSAMENTO DE IMAGENS PARA ANALISAR CINTILOGRAFIAS RENAI	15
2.1 Processamento e Análise de Imagens utilizando técnicas da Morfologia Matemática	15
2.1.1 Definições e notação	15
2.1.2 Elemento estruturante	16
2.1.3 Operadores básicos	16
2.1.4 Composição de operadores básicos	20
2.1.5 Geodésia	22
2.1.6 Máximos e mínimos	25
2.1.7 Transformações h -máximo e h -máximo estendido	27
2.2 Métodos de Segmentação de imagens	27
2.2.1 Classificação de técnicas de segmentação	27
2.2.2 Limiarização	28
2.2.3 Detecção de bordas	28
2.2.4 Crescimento de regiões	28
2.2.5 Agrupamento	29
2.3 Extração de Feições de Imagens	32
2.3.1 Definição de Feição	32
2.3.2 Organização de Feições	32
2.4 Análise de Componentes Principais	37
2.4.1 Definição	37
2.4.2 Covariância	38
2.4.3 Matriz de covariância	38
2.4.4 Autovetores e autovalores	38

2.4.5	Componentes principais	39
2.4.6	Aplicação da PCA em problemas de processamento de imagens	40
3	UMA PROPOSTA PARA QUANTIFICAR IMPRESSÕES DIAGNÓSTICAS EM IMAGENS DE CINTILOGRAFIA RENAL	41
3.1	O exame de cintilografia renal	41
3.1.1	Características das imagens de cintilografia renal	42
3.1.2	Metodologia de coleta das imagens	43
3.2	Segmentação das imagens de cintilografia renal	43
3.2.1	Aplicação de métodos existentes na literatura	43
3.2.2	Explicação do método desenvolvido	44
3.3	Extração de feições das imagens de cintilografia renal	49
3.3.1	Feições de baixo nível	51
3.3.2	Feições de nível intermediário	52
3.3.3	Feições de alto nível	54
4	ANÁLISE E DISCUSSÃO DOS RESULTADOS	59
4.1	Comparação dos resultados com a avaliação de especialistas	59
4.2	Discussão dos resultados obtidos	60
4.2.1	Captação do rádiofarmaco	60
4.2.2	Comportamento dos contornos	60
4.2.3	Forma	61
4.2.4	Simetria	62
4.2.5	Posição relativa	62
4.3	Resumo dos resultados	63
4.4	Aplicação das feições para representar estudos	63
5	CONSIDERAÇÕES FINAIS	68
5.1	Sugestão de trabalhos futuros	69
	REFERÊNCIAS	70

LISTA DE ABREVIATURAS E SIGLAS

CSS	<i>Curvature Scale-Space</i>
DICOM	<i>Digital Imaging and Communications in Medicine</i>
DMSA	<i>Dimercapto-succinic Acid</i> – Ácido Dimercapto-sucínico
EM	<i>Expectation-Maximization</i>
HCPA	Hospital de Clínicas de Porto Alegre
PCA	<i>Principal Component Analysis</i> – Análise de Componentes Principais
pdf	<i>Probability Density Function</i> – Função Densidade de Probabilidade

LISTA DE FIGURAS

Figura 2.1:	Exemplo de elemento estruturante do tipo diamante com tamanho 7x7.	16
Figura 2.2:	Erosão binária.	17
Figura 2.3:	Erosão em níveis de cinza.	18
Figura 2.4:	Dilatação binária.	19
Figura 2.5:	Dilatação em níveis de cinza.	19
Figura 2.6:	Abertura binária.	20
Figura 2.7:	Fechamento binário.	21
Figura 2.8:	Exemplo de reconstrução morfológica.	24
Figura 2.9:	Decomposição por limiarização da imagem vista como um relevo, gerando uma pilha de conjuntos.	25
Figura 2.10:	Reconstrução morfológica da máscara I pelo conjunto J .	25
Figura 2.11:	Determinação dos máximos regionais de uma imagem por meio da reconstrução morfológica.	26
Figura 2.12:	Transformação h -máximo e h -máximo estendido.	27
Figura 2.13:	Orientação do maior eixo de uma forma.	34
Figura 2.14:	Localização do centróide de uma região.	35
Figura 2.15:	Cálculo das distâncias do centróide à borda.	36
Figura 2.16:	Perfil de distâncias centroidais.	36
Figura 2.17:	Aplicação da PCA a dados de 2 dimensões.	40
Figura 3.1:	Localização e anatomia dos rins.	42
Figura 3.2:	Características das imagens de cintilografia renal.	42
Figura 3.3:	Comparação de métodos de segmentação.	45
Figura 3.4:	Resultados parciais do processo de segmentação.	47
Figura 3.5:	Curva de crescimento das áreas em função dos níveis de cinza.	48
Figura 3.6:	Processo de segmentação das imagens de cintilografia renal.	50
Figura 3.7:	Aplicação do algoritmo EM para classificação dos <i>pixels</i> .	50
Figura 3.8:	Passo de otimização da segmentação utilizando o algoritmo EM.	50
Figura 3.9:	Representação do conteúdo de imagens de cintilografia renal em três níveis de abstração.	51
Figura 3.10:	Exemplo da variação da orientação do maior eixo.	51
Figura 3.11:	Extração dos perfis de distâncias centroidais.	53
Figura 3.12:	Extração dos pontos da borda a partir do ponto no ângulo 0° .	53
Figura 3.13:	Comportamento da curvatura frente ao aumento do σ .	54
Figura 3.14:	Comparação de perfis de distâncias centroidais como medida de simetria.	55
Figura 3.15:	Modelo de forma usual de rins obtido por meio da PCA.	56

Figura 3.16:	Posição relativa entre os rins.	57
Figura 3.17:	Bordas utilizadas para avaliar a captação do radiofármaco.	58
Figura 3.18:	Avaliação da captação do radiofármaco por meio das bordas internas e externas.	58
Figura 4.1:	Exemplos da avaliação de casos quanto à homogeneidade.	61
Figura 4.2:	Exemplos da avaliação de casos quanto ao comportamento das bordas.	61
Figura 4.3:	Exemplos da avaliação de casos quanto à forma.	62
Figura 4.4:	Exemplos da avaliação de casos quanto à simetria dos rins.	62
Figura 4.5:	Exemplos da avaliação de casos quanto à posição relativa entre os rins.	63
Figura 4.6:	Quadro que mostra como as feições podem ser utilizadas para descrever os casos de cintilografia renal.	64
Figura 4.7:	Gráfico do resultado da classificação da feição de homogeneidade da captação do radiofármaco.	65
Figura 4.8:	Gráfico do resultado da classificação da feição de comportamento das bordas do rim.	65
Figura 4.9:	Gráfico do resultado da classificação da feição de forma dos rins.	66
Figura 4.10:	Gráfico do resultado da classificação da feição de simetria entre os rins.	66
Figura 4.11:	Gráfico do resultado da classificação da feição de posição relativa entre os rins.	67

LISTA DE TABELAS

Tabela 1.1:	Expressões encontradas em laudos de cintilografia renal.	13
Tabela 4.1:	Resultados da classificação.	64

RESUMO

A cintilografia renal é um exame amplamente utilizado para a avaliação visual do funcionamento do córtex renal. Ele permite visualizar a concentração do radiofármaco, o tamanho, a forma, a simetria e a posição dos rins. No entanto, a avaliação visual das impressões diagnósticas dessas imagens tende a ser um processo subjetivo. Isso faz com que ocorra uma significativa variabilidade entre as interpretações feitas por diferentes especialistas. Assim, este trabalho tem como objetivo propor medidas quantitativas que refletem impressões diagnósticas comumente observadas por especialistas nas imagens de cintilografia renal. São atribuídos valores numéricos a essas impressões, o que, potencialmente, reduz a subjetividade e a variabilidade da interpretação das descobertas. A fim de permitir a extração dessas medidas, um método de segmentação específico para essas imagens também é proposto. Os resultados indicam que as medidas propostas atingem níveis de concordância de no mínimo 90% dos casos quando comparadas com a avaliação visual de especialistas. Esses resultados sugerem que as medidas podem ser usadas para reduzir a subjetividade na avaliação das imagens, já que elas fornecem uma alternativa quantitativa e objetiva para reportar as impressões diagnósticas das imagens de cintilografia renal.

Palavras-chave: Processamento de imagens, imagens médicas, cintilografia renal, quantificação de informações visuais.

Quantification of Diagnostic Impressions in Renal Scintigraphy

ABSTRACT

Renal scintigraphy is a well established functional technique for the visual evaluation of the renal cortical mass. It allows the visualization of the radiopharmaceutical tracer distribution, the size, the shape, the symmetry, and the position of the kidneys. However, the visual diagnostic impressions for these images tend to be a subjective process. It causes significant variability in the interpretation of findings. Thus, this work aims at proposing quantitative measures that reflect common diagnostic impressions for those images. These measures can potentially minimize the inter-observer variability. In order to make possible the extraction of these measures, a specific segmentation method is also proposed. The results indicate that our proposed features agree in at least 90% of the cases with the specialists visual evaluation. These results suggest that the features could be used to reduce the subjectivity in the evaluation of the images, since they provide a quantitative and objective alternative to report the diagnostic impressions.

Keywords: Image processing, medical images, renal scintigraphy, visual information description.

1 INTRODUÇÃO

A computação tem possibilitado o desenvolvimento de inúmeras aplicações na área médica. Em ambientes médico-hospitalares, existem desde sistemas responsáveis pelo gerenciamento e recuperação de informações textuais dos pacientes até ferramentas para o processamento e a análise de imagens (DUNCAN; AYACHE, 2000).

A utilização desses sistemas tem por objetivo facilitar o acesso às informações disponíveis no ambiente médico. Os médicos, por exemplo, podem usar as informações textuais do paciente juntamente com informações extraídas das imagens dos exames para auxiliar no diagnóstico e no monitoramento de doenças. O imageamento médico também exerce um papel importante em atividades relacionadas ao ensino e à pesquisa em saúde, como o treinamento de profissionais e a avaliação de drogas (LEHMANN et al., 2005).

Em vista disso, nota-se o importante papel que as imagens desempenham na área médica. No entanto, para que a imagem seja realmente útil, é necessário que se compreenda o que nela está representado. Em termos computacionais, isso significa entender e descrever o conteúdo da imagem. Para se chegar a esse nível de entendimento, é preciso estabelecer quais aspectos relevantes devem ser analisados de acordo com o tipo da imagem. A partir disso, métodos de processamento e análise de imagens podem ser utilizados para medir esses aspectos computacionalmente. Essas medições devem refletir o que é analisado visualmente por um observador humano.

Uma abordagem para descrever o conteúdo de uma imagem é baseada em feições (SMEULDERS et al., 2000). As feições são obtidas por meio de algoritmos e devem refletir o conteúdo da imagem e/ou as necessidades de informação do usuário. Assim, as feições possuem definições bastante amplas, extensíveis para acomodar casos particulares de aplicação.

Entretanto, na maior parte dos casos, a associação de conhecimento a imagens enfrenta uma grande dificuldade: os sistemas computadorizados não são capazes de compreender o conteúdo visual associado à imagem. Grandes esforços têm sido devotados à definição clara do relacionamento entre feições visuais de baixo nível e a verdadeira semântica do conteúdo visual. Em domínios específicos, como na análise de imagens médicas, o problema pode ser simplificado em função das necessidades das aplicações e do conhecimento que se possui sobre o comportamento delas. De acordo com Chang (CHANG, 2002), a análise de imagens e de vídeos médicos é uma área de pesquisa promissora, pois as tarefas são bem definidas. Isso se deve ao fato de que os médicos procuram analisar imagens e vídeos para obter diagnósticos, prognósticos e para fins de pesquisa. Além disso, a semântica dos dados é completamente definida por uma perspectiva clínica, na qual as situações normais e anormais são diferenciadas e classificadas em categorias específicas.

Desse modo, o escopo das aplicações desenvolvidas nessa área tende a ser limitado,

visando tirar proveito de um conhecimento bem fundamentado em um domínio específico. Nessa perspectiva, este trabalho utiliza imagens provenientes de exames de cintilografia renal estática. Esse tipo de exame fornece imagens funcionais dos rins, representando, basicamente, seus aspectos morfológicos.

1.1 Impressões diagnósticas em cintilografias renais

A análise visual das imagens de cintilografia renal feita por um especialista detém-se em alguns aspectos considerados relevantes. Esses aspectos são avaliados e categorizados em impressões diagnósticas. Geralmente, os laudos quantificam as impressões diagnósticas de cinco aspectos observados nas imagens de cintilografia renal:

1. Posição relativa entre os dois rins;
2. Simetria entre o rim esquerdo e o direito;
3. Forma de cada um dos rins;
4. Captação do radiofármaco em cada um dos rins;
5. Comportamento dos contornos de cada um dos rins.

Nos laudos, cada um desses conceitos é descrito por uma expressão verbal que representa o aspecto observado. Em muitas situações, os termos usados qualificam o grau de severidade, porém não o quantificam. Assim, a avaliação feita pelo especialista pode ser considerada um processo subjetivo. Além disso, termos equivalentes são usados com frequência nos laudos. Alguns exemplos de expressões encontradas frequentemente em laudos são apresentados na Tabela 1.1.

Essas características são prejudiciais em muitos aspectos, por exemplo, na comparação de diferentes casos e na avaliação da evolução de uma doença ao longo do tempo (comparação de um mesmo caso, mas em momentos distintos). Isso ocorre porque não existem recursos que permitam a comparação. Termos como “levemente” e “um pouco” são equivalentes? Qual grau de normalidade está presente em uma “forma quase usual”? Quantas formas diversas são categorizadas em “quase usuais”?

A falta de padronização de critérios e de terminologia causa também variabilidade na interpretação das avaliações feitas por observadores distintos (JAKSIC et al., 2005; DE GUEVARA et al., 2001; DE SADELEER et al., 2000). Diferentes especialistas observando a mesma feição podem discordar na interpretação de 9% a 72% dos casos (JAKSIC et al., 2005). Essa variabilidade de interpretação é chamada de reproducibilidade interobservador e afeta a confiabilidade de um exame. Se a reproducibilidade é baixa significa que a interpretação dos resultados é subjetiva, variando de acordo com o especialista. Entretanto, quando termos e critérios são bem estabelecidos, a variabilidade diminui consideravelmente, aumentando a reproducibilidade, que atinge taxas próximas a 84,4% (DUNCAN; AYACHE, 1993).

Desse modo, julga-se que critérios e termos consistentemente definidos e métodos precisos de cálculo de medidas reduzem a variabilidade interobservador e fornecem uma análise quantitativa. Assim, este trabalho tem como objetivo fornecer medidas quantitativas que refletem impressões diagnósticas geralmente observadas por especialistas. São atribuídos valores numéricos a essas impressões, o que, potencialmente, reduz a subjetividade na interpretação das descobertas médicas. Os resultados encontrados indicam que

Tabela 1.1: Expressões encontradas em laudos de cintilografia renal.

Feição	Expressões
Posição relativa	<i>rins tópicos</i> <i>rim pélvico</i>
Simetria	<i>assimétricos</i> <i>simétricos</i> <i>assimétricos em suas dimensões</i> <i>pouco assimétricos</i> <i>levemente assimétricos</i>
Forma	<i>forma usual</i> <i>forma não usual</i> <i>forma quase usual</i>
Captação	<i>moderada redução da captação do traçador</i> <i>leve redução da captação do traçador</i> <i>concentração cortical heterogênea do radiofármaco</i> <i>concentração homogênea do radiotraçador</i> <i>rim esquerdo tem menor concentração cortical do radiotraçador</i> <i>hipocaptação severa do radiofármaco</i>
Contornos	<i>contornos regulares</i> <i>contornos irregulares</i> <i>contornos pouco irregulares</i> <i>contornos levemente irregulares</i>

as medidas tendem a concordar com a análise visual dos especialistas em taxas elevadas (no mínimo 90%). Por sua vez, os objetivos específicos são:

1. Desenvolver e testar tais métodos computacionais que detectam aspectos visualmente observados em exames de cintilografia renal; e
2. Assegurar que os resultados tenham alta correlação quando comparados com as impressões diagnósticas de especialistas em medicina nuclear.

Além disso, para poder extrair as medidas das imagens de cintilografia renal, um método de segmentação dessas imagens é proposto, utilizando fundamentalmente técnicas da Morfologia Matemática. As imagens de cintilografia renal possuem características peculiares, que fazem com que métodos clássicos de segmentação falhem consideravelmente.

É importante destacar que este trabalho foi desenvolvido com o auxílio de especialistas do Serviço de Medicina Nuclear do Hospital de Clínicas de Porto Alegre (HCPA). Esses especialistas, também vinculados ao Grupo de Pesquisa e Pós-Graduação do HCPA, forneceram os referenciais teóricos e fundamentos que guiam esse trabalho no campo da Medicina. As imagens de teste foram obtidas junto ao HCPA e também à Clínica Cardionuclear, vinculada ao Instituto de Cardiologia de Porto Alegre.

1.2 Estrutura da dissertação

O texto está organizado em cinco capítulos, incluindo esta Introdução. O Capítulo 2 apresenta os referenciais teóricos pertinentes à realização do trabalho. Nesse capítulo, são introduzidos conceitos referentes ao processamento e análise de imagens usando técnicas

da Morfologia Matemática, segmentação de imagens, extração de feições em imagens e análise de componentes principais. A seguir, no Capítulo 3, é apresentada a proposta para quantificar impressões diagnósticas em imagens de cintilografia renal. No Capítulo 4, os resultados encontrados são apresentados, discutidos e interpretados. Por fim, no Capítulo 5, as considerações finais são delineadas.

2 TÉCNICAS DE PROCESSAMENTO DE IMAGENS PARA ANALISAR CINTILOGRAFIAS RENAI

Este capítulo apresenta os referenciais teóricos utilizados para analisar as imagens de cintilografia renal. São abordados temas ligados à segmentação e à extração de feição de imagens. A Seção 2.1 descreve os fundamentos da Morfologia Matemática, utilizados na segmentação das imagens de cintilografia renal. A seguir, na Seção 2.2 são apresentados métodos clássicos para segmentação de imagens, incluindo a descrição detalhada do algoritmo de agrupamento utilizado para refinamento da segmentação. A Seção 2.3 trata da extração de feições de imagens, discutindo sua importância e enfocando em algumas feições comumente utilizadas. Por fim, a Seção 2.4 apresenta a Análise de Componentes Principais, detalhando seu funcionamento e usos na área de processamento de imagens. Esses referenciais servem de base para a proposta de quantificação de impressões diagnósticas apresentada no próximo Capítulo 3.

2.1 Processamento e Análise de Imagens utilizando técnicas da Morfologia Matemática

A Morfologia Matemática é uma abordagem do processamento de imagens que busca analisar a estrutura e a forma de elementos presentes em imagens (DOUGHERTY; LOTUFO, 2004). Foi inicialmente desenvolvida pelos pesquisadores Georges Matheron e Jean Serra (MATHERON, 1975; SERRA, 1982) e atualmente é bastante utilizada no processamento e na análise de imagens binárias e em níveis de cinza. A literatura também apresenta estudos que mostram a extensão dos conceitos para imagens coloridas (CRESPO; SCHAFFER, 1994; COMER; DELP, 1999; ORTIZ et al., 2002). As técnicas baseadas na morfologia são usadas para, por exemplo, realce, filtragem e segmentação.

Alguns conceitos referentes à teoria da morfologia matemática utilizados para realce, filtragem, quantificação e segmentação de imagens de cintilografia renal são apresentados a seguir.

2.1.1 Definições e notação

Uma imagem I é um mapeamento de um conjunto retangular finito D_I do plano discreto \mathbb{Z}^2 em um conjunto discreto $\{0, 1, \dots, N - 1\}$ de níveis de cinza. No caso de imagens binárias $N = 2$, então somente os valores 0 ou 1 são assumidos. Dado que G é uma grade de qualquer tipo, normalmente quadrada, $N_G(p)$ é o conjunto dos vizinhos do *pixel* p , associado à grade G .

Na morfologia matemática, é comum que as imagens em níveis de cinza sejam consideradas em três dimensões, sendo que os eixos x e y representam as posições dos *pixels* e

o eixo z representa a intensidade de cada *pixel*. Assim, os valores de intensidade representam elevações, tal como uma representação topográfica. As áreas de maior intensidade, chamadas de picos, e as áreas de menor intensidade, os vales, são úteis para a morfologia porque geralmente determinam objetos relevantes na imagem.

2.1.2 Elemento estruturante

A morfologia matemática examina a estrutura geométrica das imagens utilizando um conjunto conhecido e determinado, o elemento estruturante (DOUGHERTY; LOTUFO, 2004), denotado aqui por B . A maneira como o elemento estruturante interage com a imagem baseia-se em operações entre conjuntos. Os objetos presentes na imagem podem ser considerados conjuntos desconhecidos que são explorados por meio do elemento estruturante. Como o elemento estruturante é conhecido em forma e tamanho, a análise da imagem é feita pela quantificação da maneira na qual ele está (ou não) contido no conjunto da imagem.

Muitas operações da morfologia recebem como entrada a imagem e o elemento estruturante. A maioria delas requer que se defina um ponto de origem do elemento estruturante, que permite o seu posicionamento em um determinado *pixel* da imagem para a realização da operação. Assim, se o elemento B está posicionado em x , denotado por B_x , tem-se que sua origem está posicionada em x .

Os principais formatos de elementos estruturantes planares presentes nos pacotes de programas para morfologia são: cruz, disco, linha (vertical, horizontal e diagonal), diamante, caixa e círculo. Esses elementos estão disponíveis em tamanhos variados. A Figura 2.1 ilustra um exemplo de elemento estruturante do tipo diamante de tamanho 7×7 . Sua origem é o próprio centro físico, localizado nas coordenadas (4,4) e destacado na figura.

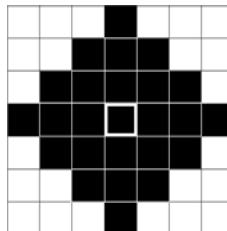


Figura 2.1: Exemplo de elemento estruturante do tipo diamante com tamanho 7×7 .

2.1.3 Operadores básicos

A morfologia matemática é sustentada por dois operadores básicos: a dilatação e a erosão. Esses operadores são combinados para formar outras transformações mais complexas. Utilizando-se de uma analogia, Soille (SOILLE, 1999) trata a erosão e a dilatação como sendo as letras do alfabeto morfológico. Essas letras podem ser combinadas para criar palavras da linguagem morfológica. E a resolução de problemas de análise de imagens consiste na construção de frases com sentido utilizando o vocabulário morfológico.

2.1.3.1 Erosão

A erosão de um conjunto X por um elemento estruturante B , denotada por $X \ominus B$, é definida da seguinte maneira (DOUGHERTY; LOTUFO, 2004), considerando que a

origem de B é posicionada no *pixel* x :

$$X \ominus B = \{x \mid B_x \subseteq X\} \quad (2.1)$$

Essa equação permite estabelecer que o resultado da erosão é dado por todos os pontos x quando B está inteiramente contido em X .

A equação 2.1 pode ser reescrita em termos de uma intersecção de translações de conjuntos (SOILLE, 1999), sendo que as translações são determinadas pelo elemento estruturante B :

$$\varepsilon_B(X) = \bigcap_{b \in B} X_{-b} \quad (2.2)$$

A equação 2.2 pode ser estendida para imagens em níveis de cinza (SOILLE, 1999). Sendo a imagem uma função f , a erosão de f pelo elemento estruturante B é denotada por $\varepsilon_B(f)$ e definida de acordo com a equação abaixo:

$$\varepsilon_B(f) = \bigwedge_{b \in B} f_{-b} \quad (2.3)$$

A Figura 2.2 ilustra uma operação de erosão binária de um conjunto X por um elemento estruturante B , do tipo diamante com tamanho 9×9 .



Figura 2.2: Imagem original e imagem resultante da erosão binária pelo elemento estruturante do tipo diamante de tamanho 9×9 .

Pode-se notar, nesta figura, que foram eliminadas as partículas de tamanho menor que o elemento estruturante. As pequenas esferas presentes na imagem foram totalmente eliminadas. Além disso, todos os elementos presentes na imagem tiveram seus tamanhos diminuídos.

Observando a Figura 2.2, pode-se notar que os efeitos da erosão em imagens binárias são os seguintes:

- diminuir as partículas;
- eliminar partículas de tamanho inferior ao tamanho do elemento estruturante;
- aumentar os buracos existentes na imagem;
- separar partículas que estão próximas.

A Figura 2.3 mostra os efeitos da erosão de uma imagem em níveis de cinza pelo elemento estruturante diamante de tamanho 9×9 . Pode-se observar que a imagem resultante é mais escura e menos rica em detalhes que a original. Isso ocorre porque as partes

escuras são aumentadas, enquanto que as partes claras são reduzidas, e algumas vezes até eliminadas. Além disso, quando duas partes escuras estão próximas, elas tendem a ser conectadas em uma só área, enquanto que as partes claras são afastadas.



Figura 2.3: Imagem original e imagem resultante da erosão em níveis de cinza pelo elemento estruturante diamante de tamanho 9×9 .

Os efeitos da erosão em imagens em níveis de cinza, evidenciados na Figura 2.3, são os seguintes, considerando a imagem vista como um relevo:

- escurecer a imagem;
- alargar e engordar os padrões escuros (vales);
- conectar vales próximos;
- reduzir e, possivelmente, eliminar padrões claros (picos);
- separar picos próximos.

2.1.3.2 Dilatação

A dilatação de um conjunto X por um elemento estruturante B , denotada por $X \oplus B$, é definida da seguinte maneira (DOUGHERTY; LOTUFO, 2004), sendo que a origem de B é posicionada no *pixel* x :

$$X \oplus B = \{x \mid B_x \cap X \neq \emptyset\} \quad (2.4)$$

A equação 2.4 pode ser reescrita em termos de uma união de translações de conjuntos (SOILLE, 1999):

$$X \oplus B = \bigcup_{b \in B} X_{-b} \quad (2.5)$$

A partir da equação 2.5, é possível definir a dilatação de uma imagem f em níveis de cinza por um elemento estruturante B . Essa operação é denotada por $\delta_B(f)$ e é definida pela equação 2.6:

$$\delta_B(X) = \bigvee_{b \in B} f_{-b} \quad (2.6)$$

A Figura 2.4 mostra a imagem original e a imagem resultante da operação de dilatação pelo elemento estruturante de formato diamante e de tamanho 9×9 . Nesta imagem é possível perceber com nitidez que todos os elementos da imagem tiveram seus tamanhos aumentados. Além disso, algumas partículas que se encontravam próximas umas das outras foram conectadas, formando uma única região.



Figura 2.4: Imagem original e imagem resultante da dilatação binária pelo elemento estruturante do tipo diamante de tamanho 9×9 .

Pode-se perceber, pela observação da Figura 2.4, que os efeitos da dilatação em imagens binárias são os seguintes:

- aumentar as partículas;
- preencher os pequenos buracos existentes na imagem;
- conectar partículas que estão próximas.

A Figura 2.5 mostra os efeitos da dilatação em uma imagem em níveis de cinza, utilizando o elemento estruturante do formato diamante e de tamanho 9×9 .



Figura 2.5: Imagem original e imagem resultante da dilatação em níveis de cinza pelo elemento estruturante do tipo diamante de tamanho 9×9 .

Em imagens em níveis de cinza, como a da Figura 2.5, pode-se afirmar que os efeitos da dilatação são:

- clarear a imagem;
- alargar e engordar os padrões claros (picos);
- conectar picos próximos fisicamente;
- reduzir e possivelmente eliminar padrões escuros;
- separar vales próximos.

É importante salientar que a operação de erosão não é o inverso da dilatação, e vice-versa. Pode ocorrer, em alguns casos, que a dilatação de um conjunto erodido retorne a imagem original, e que o mesmo ocorra no caso inverso (SOILLE, 1999). No entanto, em geral, as duas operações são duais, no sentido de o complemento de uma erosão é o

mesmo que uma dilatação do complemento da imagem pelo elemento estruturante refletido, conforme mostra a equação 2.7, e vice-versa:

$$A \ominus B = [A^c \oplus (-B)]^c \quad (2.7)$$

2.1.4 Composição de operadores básicos

Os operadores de erosão e dilatação podem ser combinados para gerar outras transformadas morfológicas. As seções seguintes apresentam algumas dessas operações.

2.1.4.1 Abertura

A erosão e a dilatação podem corrigir defeitos de uma imagem, porém os conjuntos resultantes da aplicação destas operações não mantêm os tamanhos dos conjuntos originais. Os conjuntos resultantes são sempre menores ou maiores, conforme for uma erosão ou uma dilatação, respectivamente. Porém, por meio da propriedade da iteratividade (DOUGHERTY; LOTUFO, 2004), pode-se realizar a filtragem de uma imagem sem a modificação das características de forma e tamanho dos conjuntos relevantes.

A operação de abertura é uma dessas operações que elimina partículas indesejáveis, mas não modifica o tamanho dos outros conjuntos. A abertura consiste na aplicação de uma erosão imediatamente seguida por uma dilatação utilizando o mesmo elemento estruturante. A equação 2.8 demonstra a operação matematicamente, sendo que $X \circ B$ representa a abertura da imagem X pelo elemento estruturante B :

$$X \circ B = (X \ominus B) \oplus B \quad (2.8)$$

A Figura 2.6 mostra a operação de abertura binária pelo elemento estruturante do tipo caixa (ou *box*) de tamanho 7x7.



Figura 2.6: Imagem original e imagem resultante da abertura binária pelo elemento estruturante do tipo caixa de tamanho 7x7.

Nessa figura, pode-se perceber que a abertura não retorna o conjunto original, apesar de manter as entidades da imagem quase idênticas às originais. Os contornos dos conjuntos são nivelados pelo interior, o que pode levar à separação de algumas partículas. A abertura também elimina os conjuntos que possuem tamanho inferior ao elemento estruturante. Além disso, a imagem resultante é mais regular e menos rica em detalhes do que a imagem original (FACON, 1996).

Para imagens em níveis de cinza, a notação utilizada para a abertura é a seguinte:

$$\gamma_B = \delta_B \epsilon_B \quad (2.9)$$

Os efeitos da abertura em imagens em níveis de cinza, seguindo a analogia da imagem vista como um relevo topográfico, são os seguintes:

- separar picos próximos;
- eliminar picos inferiores em tamanho ao elemento estruturante;
- conservar vales afastados;
- emendar vales próximos;
- deixar a imagem mais regular que a original;
- eliminar detalhes da imagem.

2.1.4.2 Fechamento

O fechamento é a operação inversa da abertura pelo fato de realizar primeiro a dilatação e depois a erosão. Matematicamente, a operação pode ser expressa por meio da seguinte equação:

$$X \bullet B = (X \oplus B) \ominus B \quad (2.10)$$

A Figura 2.7 reproduz os efeitos do fechamento aplicado a uma imagem binária utilizando elemento estruturante do tipo caixa de tamanho 7x7. Nessa figura, pode-se observar que a imagem resultante é mais regular e menos rica em detalhes do que a imagem original. O fechamento produz o efeito de emendar as partículas que se encontram próximas fisicamente, além de preencher os buracos pelo interior das partículas de tamanho inferior ao elemento estruturante.



Figura 2.7: Imagem original e imagem resultante do fechamento em níveis de cinza pelo elemento estruturante do tipo caixa de tamanho 7x7.

O fechamento em níveis de cinza é denotado como

$$\phi_B = \epsilon_B \delta_B \quad (2.11)$$

Os efeitos do fechamento em imagens em níveis de cinza são os seguintes:

- separar vales próximos;
- eliminar vales inferiores em tamanho ao elemento estruturante;
- conservar picos afastados;
- emendar picos próximos;
- deixar a imagem mais regular que a original;
- eliminar detalhes da imagem.

2.1.4.3 Gradiente morfológico

A informação do gradiente é bastante utilizada para detectar bordas em imagens. Na morfologia matemática, existem várias implementações para o cálculo do gradiente. A abordagem aqui apresentada usa simultaneamente as operações de dilatação e erosão (DOUGHERTY; LOTUFO, 2004).

O gradiente morfológico $grad(X, B)$ de um conjunto X pelo elemento estruturante B é dada por:

$$grad(X, B) = (X \oplus B) - (X \ominus B) \quad (2.12)$$

Essa equação indica que a imagem erodida deve ser subtraída da dilatada. Quando aplicado em imagens binárias, o gradiente resulta em um contorno ao redor do objeto, detectando suas bordas. Em imagens em níveis de cinza, o resultado são bordas mais claras e escurecimento da região interna dos objetos. Em ambos os casos, a largura desses contornos depende do tamanho do elemento estruturante utilizado.

2.1.5 Geodésia

Os operadores apresentados até agora consideram as imagens como sendo conjuntos indivisíveis. No entanto, muitas aplicações precisam restringir uma região específica da imagem e atuar somente nesta região. Isso ocorre no processamento das imagens de cintilografia renal, pois os dois rins devem ser processados separadamente devido ao fato de, em geral, possuírem características diferentes.

Para conseguir processar isoladamente um conjunto da imagem, pode-se utilizar os operadores de dilatação ou erosão condicionais (SOILLE, 1999). Essas operações são baseadas nos operadores elementares, mas utilizam uma imagem de marcadores que limita a região de atuação da transformada. Por exemplo, a dilatação condicional de um subconjunto Z da imagem X pelo elemento estruturante B é definida como a união da própria imagem X com a dilatação de Z por B ($Z \oplus B \cap X$). Porém, o fato de processar de forma condicional um subconjunto da imagem não fornece sempre resultados satisfatórios, porque o processo pode invadir de forma indevida outro subconjunto da imagem (FACON, 1996). Isso ocorre em função de não existir uma noção de distância entre os conjuntos. Essa característica é fornecida pela geodésia.

A geodésia, por definição, é a ciência que envolve o levantamento e a representação da forma e da superfície da terra. Essa ciência deu origem a alguns conceitos que são usados para medição e cálculo de superfícies curvas empregando métodos semelhantes àqueles usados nas superfícies curvas da terra.

2.1.5.1 Distância geodésica

A distância geodésica $d_S(x, y)$ entre dois pontos x e y de um subconjunto S é o comprimento do menor caminho que liga os dois pontos e está completamente contido em S . Em geodésia, a definição de um subconjunto onde se aplica a métrica, no caso o subconjunto S , é necessária, diferentemente dos conceitos Euclidianos.

2.1.5.2 Dilatação geodésica

Como mencionado anteriormente, a dilatação condicional trabalha com uma imagem marcadora para definir o local de atuação do operador. A dilatação geodésica também usa o mesmo princípio. Existe uma imagem marcadora J e uma imagem máscara I . As duas possuem o mesmo domínio e $J \subseteq I$. O que ocorre é que a imagem marcadora é dilatada

por um elemento estruturante elementar e então é utilizada para restringir a propagação da dilatação da imagem marcadora.

Dada uma imagem X (máscara) e uma imagem Y (marcadora), sendo que $Y \subseteq X$, a dilatação geodésica $\delta_X^{(n)}(Y)$ de tamanho $n \geq 0$ de Y dentro de X é o conjunto dos *pixels* de X cuja distância geodésica de Y é menor ou igual a n (VINCENT, 1993):

$$\delta_X^{(n)}(Y) = \{p \in X | d_X(p, Y) \leq n\} \quad (2.13)$$

Pela definição pode-se notar que $Y \subseteq \delta_X^{(n)}(Y)$, o que confere à esta transformação a propriedade da extensividade (DOUGHERTY; LOTUFO, 2004). Portanto, a dilatação geodésica de tamanho n pode ser obtida por meio da iteração de n dilatações geodésicas elementares (VINCENT, 1993)

$$\delta_X^{(n)}(Y) = \underbrace{\delta_X^{(1)} \circ \delta_X^{(1)} \circ \dots \circ \delta_X^{(1)}}_{n \text{ vezes}}(Y) \quad (2.14)$$

A definição de uma dilatação geodésica elementar pode ser obtida por meio da dilatação elementar seguida por uma intersecção com o conjunto X

$$\delta_X^{(1)}(Y) = (Y \oplus B) \cap X \quad (2.15)$$

Essa definição pode ser erroneamente avaliada como equivalente a equação da dilatação condicional apresentada anteriormente. Entretanto, a definição anterior não obedece ao critério da métrica geodésica.

Dessa maneira, a aplicação sucessiva de dilatações geodésicas elementares de um conjunto Y dentro de uma máscara X faz com que os componentes conectados de X que tem intersecção não-vazia com Y sejam progressivamente revelados. Assim, pode-se estabelecer formalmente o conceito de reconstrução.

2.1.5.3 Reconstrução binária

A reconstrução de X a partir de Y , sendo que $Y \subseteq X$, é obtida pela iteração de dilatações geodésicas elementares de Y dentro de X até que a estabilidade seja atingida (VINCENT, 1993)

$$\rho_X(Y) = \bigcup_{n \geq 1} \delta_X^n(Y) \quad (2.16)$$

A Figura 2.8 ilustra o processo de reconstrução das letras “a” da palavra por meio dos marcadores, mostrados na Figura 2.8(b) e sobrepostos à imagem na Figura 2.8(c) para melhor visualização da posição dos mesmos. O resultado da reconstrução está na Figura 2.8(d), que mostra a reconstrução somente das letras “a”. Nesse caso, X é a palavra inteira, e Y são os marcadores. Esse exemplo demonstra o potencial de aplicação da reconstrução. As filtragens clássicas permitem eliminar padrões indesejáveis na imagem, mas, em muitos casos, deformam também as estruturas que devem ser mantidas.

2.1.5.4 Reconstrução em níveis de cinza

A reconstrução morfológica para imagens em níveis de cinza é definida como seqüência lógica da versão binária. O objetivo é similar, ou seja, tentar reconstruir uma região da imagem. Entretanto, na reconstrução em níveis de cinza, a imagem não será perfeitamente



Figura 2.8: Exemplo de reconstrução morfológica: (a) imagem original; (b) imagem de marcadores; (c) sobreposição dos marcadores na imagem original para fins de visualização da posição; e (d) resultado da reconstrução, na qual somente as letras “a” foram reconstruídas.

reconstruída. Essa característica da transformação permite introduzir outras técnicas de processamento (FACON, 1996).

Qualquer transformação binária crescente ψ pode ser estendida a imagens em níveis em níveis de cinza (DOUGHERTY; LOTUFO, 2004). A transformação ψ é dita crescente se

$$Y \subseteq X \Rightarrow \psi(Y) \subseteq \psi(X), \quad \forall X, Y \subset \mathbb{Z}^2 \quad (2.17)$$

Para estender essa transformação ψ , basta considerar as sucessivas limiarizações $T_k(I)$ da imagem I , no seu domínio D_I , para $k \in [0, N]$

$$T_k(I) = \{p \in D_I | I(p) \geq k\} \quad (2.18)$$

$T_k(I)$ constitui a decomposição por limiarização da imagem I e os conjuntos obtidos satisfazem a relação de inclusão seguinte

$$T_k(I) \subseteq T_{k-1}(I), \quad \forall k \in [1, N - 1] \quad (2.19)$$

Essa relação é demonstrada na Figura 2.9, que evidencia que cada conjunto um nível acima ($T_k(I)$) é subconjunto do nível de cinza logo abaixo ($T_{k-1}(I)$).

Quando a transformação ψ é aplicada a cada um desses conjuntos, a relação de inclusão se mantém. Dessa maneira, pode-se estender a transformação ψ para imagens em níveis de cinza da seguinte maneira

$$\forall p \in D_I, \quad \psi(I)(p) = \max \{k \in [0, N - 1] \mid p \in \psi(T_k(I))\} \quad (2.20)$$

Assim, a reconstrução geodésica binária ρ é uma transformação crescente que satisfaz a relação (VINCENT, 1993)

$$Y_1 \subseteq Y_2, X_1 \subseteq X_2, Y_1 \subseteq X_1, Y_2 \subseteq X_2 \Rightarrow \rho_{X_1}(Y_1) \subseteq \rho_{X_2}(Y_2) \quad (2.21)$$

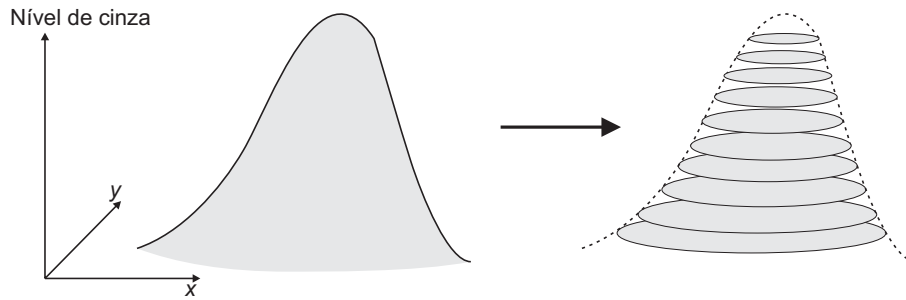


Figura 2.9: Decomposição por limiarização da imagem vista como um relevo, gerando uma pilha de conjuntos.

Com base no princípio da decomposição por limiarização, pode-se definir o conceito de reconstrução em níveis de cinza. Sejam J e I duas imagens em níveis de cinza, sob o mesmo domínio D_I , com valores definidos no intervalo discreto $[0, 1, \dots, N - 1]$ e tal que $J \leq I$. Então, a reconstrução em níveis de cinza $\rho_I(J)$ da máscara I pela imagem marcadora J é dada por (VINCENT, 1993)

$$\forall p \in D_I, \quad \rho_I(J)(p) = \max \{k \in [0, 1, \dots, N - 1] \mid p \in \rho_{T_k(I)}(T_k(J))\} \quad (2.22)$$

Como no caso da reconstrução binária, na qual são extraídos os subconjuntos marcados, a reconstrução em níveis de cinza permite extrair os picos da imagem que são marcados pela imagem de marcadores. A Figura 2.10 exemplifica essa transformação em uma imagem vista como um perfil.

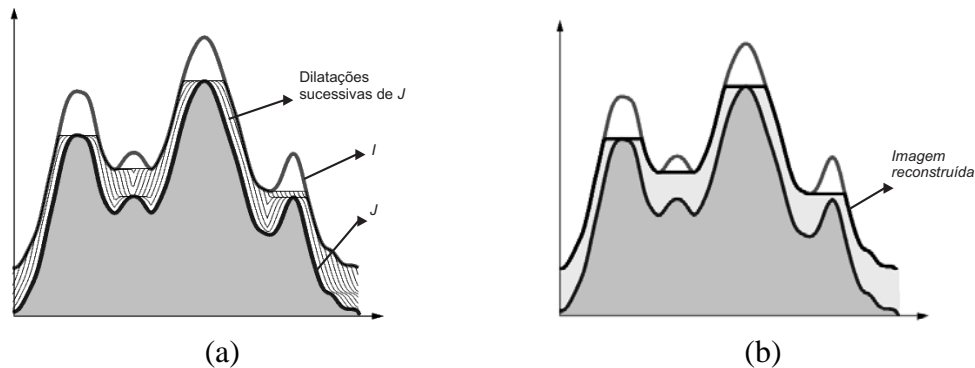


Figura 2.10: Reconstrução morfológica da máscara I pelo conjunto J : (a) conjunto I original, marcador J (em cinza escuro) e dilatações geodésicas sucessivas; (b) resultado da reconstrução (em cinza claro).

2.1.6 Máximos e mínimos

Como já foi exposto anteriormente, imagens em níveis de cinza podem ser consideradas em três dimensões, sendo que os eixos x e y representam as posições dos *pixels* e o eixo z representa a intensidade de cada *pixel*. As áreas de maior intensidade (picos) e as áreas de menor intensidade (vales) são úteis para a morfologia porque geralmente determinam objetos relevantes na imagem. Dessa maneira, encontrar essas regiões de

máxima ou mínima intensidade é uma tarefa bastante útil para muitas aplicações, como, por exemplo, criar imagens de marcadores para usar na reconstrução morfológica.

Um máximo regional M de uma imagem em níveis de cinza I é definido pelo conjunto de *pixels* conexos que tem um determinado valor h , tal que cada um dos seus *pixels* vizinhos tem um valor estritamente menor (VINCENT, 1993).

De maneira análoga, um mínimo regional M de uma imagem em níveis de cinza I é definido pelo conjunto de *pixels* conexos que tem um determinado valor h , tal que cada um dos seus *pixels* vizinhos tem um valor estritamente maior (VINCENT, 1993).

Pode-se estabelecer que os conceitos de mínimos e máximos regionais são globais e não-locais (SBEH et al., 2001). Não se pode afirmar que um componente é um máximo (ou mínimo) regional somente examinando uma pequena vizinhança do componente, como ocorre nos conceitos locais. Essa característica torna-se vantajosa na medida que não se consideram pontos que são vistos como máximos locais devido a uma vizinhança plana, mas que não são verdadeiramente máximos. Além disso, somente um conjunto de máximo regional será detectado em um máximo com área plana, ao invés de vários pontos.

Para determinar o máximo de uma imagem, pode-se usar a reconstrução em níveis de cinza. Dada uma imagem I , o conjunto M que determina seu máximo regional é dado por

$$M(I) = I - \rho_I(I - 1) \quad (2.23)$$

Essa equação demonstra que para se obter o máximo regional pode-se reconstruir morfológicamente a imagem I a partir de $I - 1$ e subtrair o resultado da própria imagem. A Figura 2.11 ilustra esse processo.

O conjunto dos máximos regionais de uma imagem I , simbolizado por $RMAX(I)$, é definido da seguinte maneira:

$$RMAX(I) = \{p \in D \mid (I - \rho_I(I - 1))(p) > 0\} \quad (2.24)$$

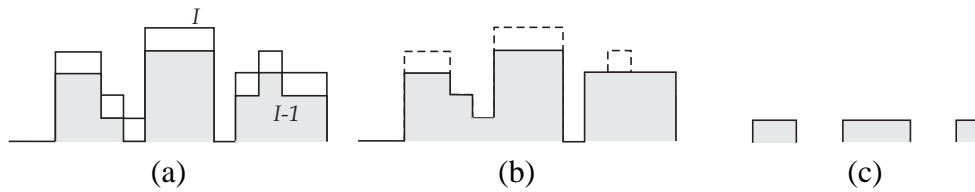


Figura 2.11: Determinação dos máximos regionais de uma imagem por meio da reconstrução morfológica: (a) representação em perfil de I e de $I - 1$; (b) reconstrução morfológica da imagem I pela imagem de marcadores $I - 1$; e (c) resultado da diferença, que mantém somente os máximos regionais.

De maneira semelhante, define-se o conjunto dos mínimos regionais, denotados por $RMIM(I)$. Nesse caso, usa-se o complemento I^C de I no processo de reconstrução. I^C é dado por

$$I^C(p) = (N - 1) - I(p) \quad (2.25)$$

Então, o conjunto $RMIM(I)$ fica assim definido

$$RMIM(I) = RMAX(I^C) \quad (2.26)$$

2.1.7 Transformações h -máximo e h -máximo estendido

Até esse ponto, as operações apresentadas permitem encontrar pontos de início para todas as regiões mais claras de uma imagem. Agora, os conceitos são estendidos para encontrar conjuntos de maior área. A transformação que realiza essa operação é a h -máximo, que permite filtrar uma imagem com base em um critério de contraste.

A transformação h -máximo suprime todos os máximos que tenham profundidade menor que um dado valor de intensidade, determinado por h , que indica a altura do pico. Utilizando a reconstrução morfológica, a transformação h -máximo, denotada por $HMAX_h$, pode ser definida como (SOILLE, 1999)

$$HMAX_h(I) = \rho_I(I - h) \quad (2.27)$$

Já a operação h -máximo estendido corresponde a encontrar os máximos regionais da transformação h -máximo, ou seja

$$EMAX_h(I) = RMAX(HMAX_h(I)) \quad (2.28)$$

O resultado da operação h -máximo estendido é uma imagem binária que determina as localizações dos máximos regionais da imagem. No caso da operação h -máximo, o resultado é uma imagem alterada, mas em níveis de cinza, como mostra a Figura 2.12.

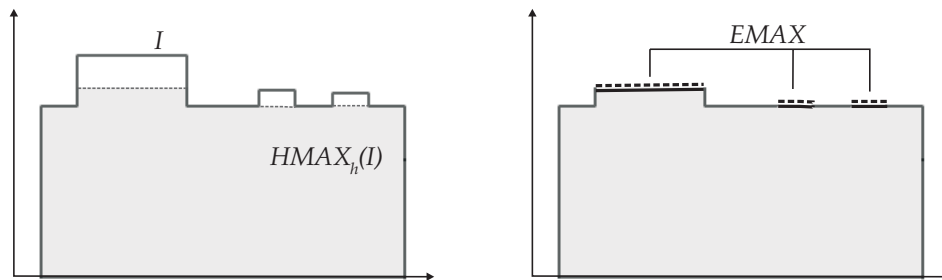


Figura 2.12: Transformação h -máximo e h -máximo estendido: (a) perfil do conjunto I e do conjunto $HMAX_h(I)$, resultante da transformação h -máximo; (b) resultado da transformação h -máximo estendido, dado pelo conjunto $EMAX_h(I)$.

2.2 Métodos de Segmentação de imagens

O processo de segmentação de imagens consiste em agrupar *pixels* em regiões homogêneas. Por região homogênea, entende-se um grupo de *pixels* conectados que compartilham de uma mesma característica (BOVIK; GIBSON; BOVIK, 2000), que pode ser cor, textura, luminância, entre outras. Além disso, é desejável que essas regiões tenham um possível significado semântico associado (FORSYTH; PONCE, 2002).

Alguns métodos tradicionais de segmentação de imagens são apresentados nessa seção.

2.2.1 Classificação de técnicas de segmentação

Geralmente a segmentação é o primeiro passo da resolução de problemas em análise de imagens. Não é eficiente manipular diretamente os *pixels*, pois eles constituem informações de baixo nível das imagens. Assim, encontrar regiões homogêneas fornece uma representação de mais alto nível.

A segmentação desempenha um papel central em muitos sistemas que envolvem processamento de imagens ou vídeos. Isso ocorre porque a maioria dos sistemas parte da segmentação para computar feições, então, a confiabilidade dessas feições é dependente da confiabilidade que se tem no esquema utilizado para a segmentação.

Existem muitas técnicas de segmentação de imagens, como, por exemplo, técnicas baseadas na limiarização de histogramas, por agrupamento ou crescimento de regiões ou ainda em detecção de bordas. A categorização das técnicas existentes é uma tarefa bastante complexa, condição comprovada pelo fato de cada autor apresentar um esquema diferente e possivelmente divergente dos demais. Para o propósito desta dissertação, são consideradas técnicas baseadas em limiarização, detecção de bordas, crescimento de regiões e agrupamento.

2.2.2 Limiarização

Nas técnicas baseadas na limiarização, é definido um limiar (*threshold*) e a partir dele a imagem é transformada em uma imagem binária (BOVIK; GIBSON; BOVIK, 2000). Por exemplo, em uma imagem em níveis de cinza, com valores que variam de 0 a 255, a definição de um limiar igual a 170 faz com que todos os *pixels* que estiverem abaixo desse valor recebam a intensidade 1, e os que tem valor superior recebam a intensidade 0. O limiar é considerado efetivo quando os níveis de intensidade do objeto a ser segmentado sejam bastante diferentes da intensidade do fundo da imagem. A escolha do limiar pode basear se em histogramas. A referência (GLASBEY, 1993) apresenta uma comparação de métodos dessa natureza.

2.2.3 Detecção de bordas

Nos métodos baseados em detecção de bordas, procuram-se descontinuidades na imagem. A maneira mais comum de encontrar essas descontinuidades é utilizar filtros direcionais que analisam a variação das intensidades dos *pixels* (GONZALEZ; WOODS, 2000). Os filtros são úteis porque analisam não só o valor de um *pixel*, mas também a sua vizinhança. Alguns exemplos de filtros amplamente utilizados para esse propósito são Sobel, Prewitt, Roberts e Laplaciano de Gaussiano. Mais detalhes sobre esses filtros podem ser encontrados na referência (BOVIK; GIBSON; BOVIK, 2000).

Outro método clássico de detecção de bordas é o Canny (CANNY, 1987). Esse método encontra as bordas procurando por máximos locais no gradiente da imagem original. O gradiente pode ser calculado pela derivada de um filtro Gaussiano, por exemplo. O método usa dois limiares para detectar bordas “fracas” e “fortes” e inclui as bordas fracas na imagem de saída somente se elas estiverem conectadas a bordas fortes. Por esse motivo, ele é menos suscetível ao ruído.

2.2.4 Crescimento de regiões

As técnicas baseadas no crescimento de regiões realizam o particionamento das imagens em regiões conectadas por meio do agrupamento de *pixels* vizinhos com intensidades semelhantes. O objetivo final é produzir regiões coerentes com o maior tamanho possível. No caso das bordas serem borradas ou pouco nítidas, ocorre o problema da sobre-segmentação, que produz uma grande fragmentação da imagem. Como exemplos de métodos baseados em crescimento de regiões estão a segmentação por *Watersheds* (VINCENT; SOILLE, 1991) e o modelo Mumford-Shah (MUMFORD; SHAHA, 1989).

O método de segmentação por *Watersheds* baseia-se nos princípios da Morfologia

Matemática. A idéia do algoritmo é considerar a imagem como um relevo topográfico que é inundado com água, permitindo a formação de bacias de captação e de linhas divisoras de água (*watersheds*), particionando a imagem em regiões (VINCENT; SOILLE, 1991).

Já o algoritmo de Mumford-Shah (MUMFORD; SHAHA, 1989) usa a idéia de que as regiões são grupos de *pixels* com um “elástico” ao redor. Assim, as regiões crescem enquanto o elástico puder ser esticado, sendo que a elasticidade diminui à medida que a variação entre os *pixels* dentro de uma região aumenta.

2.2.5 Agrupamento

A classificação por agrupamento (*clustering*) é uma técnica de aprendizagem não-supervisionada que busca encontrar uma estrutura em uma coleção de dados não-rotulados (THEODORIDIS; KOUTROUMBAS, 2003). A aplicação dos algoritmos de agrupamento é bastante extensa. Por exemplo, analistas de mercado podem encontrar grupos de consumidores com comportamento similar a partir de uma grande base de dados de registros de compras. Na área de processamento e análise de imagens, esses algoritmos auxiliam na segmentação em regiões, pois busca-se agrupar *pixels* que devem ficar juntos.

Em alguns algoritmos de agrupamento, os dados são agrupados de uma maneira única, de tal forma que se um certo objeto pertence a um grupo definido, então ele não pode ser incluído em nenhum outro grupo. Um exemplo de algoritmo desse grupo é o *K-means* (MACQUEEN, 1967). Porém, outros algoritmos permitem sobreposições, de maneira que um ponto pode pertencer a mais de um grupo, com diferentes critérios para ser membro de cada um. Um exemplo de algoritmo dessa classe é o *Fuzzy C-means* (BEZ-DEK, 1981).

Entretanto, nenhum desses métodos envolve o uso de modelos de probabilidade explícitos para classificar os dados. Eles utilizam essencialmente as informações disponíveis de similaridade dos dados, mas nenhuma suposição probabilística é feita. Essa abordagem é utilizada em métodos que se baseiam em modelos de distribuição dos dados. Nesses métodos, modelos probabilísticos são usados e tenta-se encontrar o melhor ajustamento entre os dados e os modelos. Dentro dessa categoria de algoritmos de agrupamento encontra-se o *Expectation Maximization* (EM) que foi utilizado para o refinamento do resultado da segmentação dos rins e é apresentado a seguir.

2.2.5.1 *Expectation Maximization*

O algoritmo *Expectation Maximization* (EM), proposto por Dempster, Laird e Rubin (DEMPSTER; LAIRD; RUBIN, 1977), consiste em um processo iterativo envolvendo dois passos: previsão e estimação. Ele permite calcular estimativas de máxima verossimilhança a partir de dados incompletos (THEODORIDIS; KOUTROUMBAS, 2003).

O algoritmo EM é frequentemente utilizado na segmentação de imagens em problemas de visão computacional. Nesses problemas, geralmente há várias fontes contribuindo para a geração dos dados. Por exemplo, o valor de um *pixel* pode vir de um objeto, da iluminação ou de ruído na captação. Se a fonte que gera o valor do *pixel* for conhecida, a segmentação é facilitada. Porém, essa fonte que gera o dado (ou a mistura de fontes) é desconhecida. A falta de informação para poder atribuir um dado a uma fonte faz com que o conjunto de dados seja considerado incompleto. Problemas com essa característica são conhecidos como problemas com conjunto de dados incompletos ou faltantes (*incomplete data set problem* ou *missing data problem*).

O objetivo do EM é separar os subconjuntos de dados. Entretanto, não se conhece como cada fonte contribui para a geração dos dados. Não se sabe se um *pixel* é escuro

porque tem pouca iluminação ou se a região na qual ele está contido pertence a um objeto realmente escuro. A situação ideal seria fazer medidas e determinar uma descrição razoável do contexto que gera os dados. Mas mesmo essas medidas adicionam incerteza ao processo, pois podem ser ambíguas. Para lidar com essas incertezas, o mecanismo mais adequado é tratar com probabilidades (FORSYTH; PONCE, 2002). Estima-se as preferências e tendências do cenário, como, por exemplo, se é mais comum existir regiões escuras da imagem por problemas de iluminação ou devido a existência de objetos escuros.

Então, agora é preciso tratar o problema levando em consideração aspectos probabilísticos. Mais especificamente, é preciso estimar as funções densidade de probabilidade (*probability density function* – pdf) de cada fonte utilizando os dados disponíveis, já que que se desconhece essas funções.

De acordo com as informações existentes, algumas abordagens podem ser utilizadas para estimar a pdf desconhecida. Algumas vezes, se conhece o tipo de pdf, que pode ser uma Gaussiana ou uma Rayleigh, mas não se conhecem certos parâmetros, como a média e a variância. Em outros casos, não se sabe o tipo de pdf, mas existem informações disponíveis sobre alguns parâmetros estatísticos (THEODORIDIS; KOUTROUMBAS, 2003). Então, dependendo da informação disponível, diferentes abordagens podem ser adotadas, geralmente por meio de inferência.

No caso de imagens, é razoável estimar que cada fonte de dados pode ser matematicamente representada por uma distribuição paramétrica, como uma Gaussiana ou uma Poisson (FORSYTH; PONCE, 2002), então o conjunto de todos os dados pode ser modelado como uma mistura dessas distribuições. Essa forma de modelar uma pdf desconhecida é chamada de modelo de mistura (*mixture model*) (THEODORIDIS; KOUTROUMBAS, 2003). Formalmente, uma $p(x)$ desconhecida pode ser modelada por meio de uma combinação linear de funções de densidade da forma

$$p(x) = \sum_{j=1}^J p(x|j)P_j \quad (2.29)$$

onde

$$\sum_{j=1}^J P_j = 1, \quad \int_x p(x|j)dx = 1 \quad (2.30)$$

Ou seja, assume-se que existe J distribuições contribuindo para a formação de $p(x)$ (THEODORIDIS; KOUTROUMBAS, 2003). Dessa maneira, esse modelo assume que cada ponto x pode ser gerado por qualquer um das J distribuições com probabilidade $P_j, j = 1, 2, \dots, J$.

Uma maneira de resolver o problema de encontrar os dados que estão faltando e também estimar os parâmetros consiste em formalizar uma idéia intuitiva de lidar com dados incompletos, que é a seguinte:

1. substituir os valores que estão faltando por valores estimados;
2. estimar os parâmetros com base no palpite feito para os valores;
3. reestimar os valores faltando considerando que os novos parâmetros são corretos;
4. reestimar os parâmetros;

5. repetir os passos anteriores até que um critério de convergência seja atingido.

Esses passos são a base de funcionamento do algoritmo EM. Formalmente, assume-se um espaço de dados completo Y , com amostras y , tal que $y \in Y \subseteq \mathbb{R}^m$. A pdf correspondente é dada por $p_y(y; \theta)$, onde θ é um vetor de parâmetros desconhecidos. Entretanto, as amostras y não podem ser diretamente observadas. O que se observa são as amostras $x = g(y) \in X_{ob} \subseteq \mathbb{R}^l, l < m$. Ou seja, existe um mapeamento muitos-para-um g do espaço Y para o espaço X . Existe perda de informação nesse mapeamento. No caso de imagens, o espaço completo seria formado pelos valores de cada *pixel* e pelo conjunto de variáveis que indicam de que fonte de dados cada *pixel* vem. Durante o mapeamento, perde-se a informação das fontes dos dados, o que torna o espaço incompleto. Assim, para esse caso, os passos seriam: 1) estimar de que fonte de dados vem um *pixel*, usando um vetor de parâmetros θ ; e 2) atualizar θ , utilizando a estimativa.

Sendo $Y(x) \subseteq Y$ o subconjunto de todos os y 's que correspondem a um x específico, então a pdf dos dados incompletos é dada por

$$P_x(x; \theta) = \int_{Y(x)} p_y(y; \theta) dy \quad (2.31)$$

A estimativa de máxima verossimilhança de θ é dada pela seguinte equação (THEODORIDIS; KOUTROUMBAS, 2003):

$$\hat{\theta}_{ML} : \sum_k \frac{\partial \ln(p_y(y_k; \theta))}{\partial \theta} = 0 \quad (2.32)$$

Entretanto, os valores dos y 's não estão disponíveis. Então, o algoritmo EM maximiza a previsão da função de log-verossimilhança, condicionada às amostras observadas e a estimativa da iteração corrente de θ . Assim, os dois passos do algoritmo são:

- Passo *E*: No passo $(t+1)$ da iteração, onde $\theta(t)$ está disponível, computa-se o valor esperado de

$$Q(\theta; \theta(t)) \equiv E \left[\sum_k \ln(p_y(y_k; \theta | X; \theta(t))) \right] \quad (2.33)$$

Esse é o passo de previsão do algoritmo (*expectation step* ou *E-step*).

- Passo *M*: Computa-se a próxima $(t+1)$ -ésima estimativa de θ , maximizando $Q(\theta; \theta(t))$, ou seja

$$\theta(t+1) : \frac{\partial Q(\theta; \theta(t))}{\partial \theta} = 0 \quad (2.34)$$

Esse é o passo de maximização do algoritmo (*maximization step* ou *M-step*). A diferenciabilidade é assumida.

O algoritmo começa com uma estimativa inicial $\theta(0)$ e cessa as iterações quando uma condição de parada $\|\theta(t+1) - \theta(t)\| \leq \epsilon$ for assumida, adotando-se uma norma vetorial e um ϵ apropriados.

2.3 Extração de Feições de Imagens

O sistema visual humano é capaz de diferenciar objetos, localizá-los no espaço, perceber deslocamentos, tamanhos, cores e texturas. De maneira análoga, sistemas de visão computacional tentam realizar tarefas semelhantes. Para isso, esses sistemas procuram mapear aspectos da percepção visual humana para algoritmos de processamento e análise de imagens. Uma estratégia para realizar esse mapeamento é recorrer ao uso de feições.

Essa seção aborda conceitos referentes à extração de feições, detalhando como elas podem ser organizadas e extraídas.

2.3.1 Definição de Feição

Uma feição de imagem pode ser definida como uma característica ou atributo de uma imagem (LEW, 2001). Geralmente, as feições são definidas em função de características úteis para a descrição do conteúdo da imagem. O conceito de feição é bastante amplo, podendo assumir tanto o formato de resultado de uma computação como o de critério de avaliação de um especialista. Na perspectiva de um sistema de visão computacional, o conteúdo de uma imagem pode ser definido como o conjunto de todas as feições possíveis ou a combinação de feições.

2.3.2 Organização de Feições

Cada aplicação possui características diferentes. Essas características variam de acordo com a natureza do conteúdo da imagem, as informações que o usuário precisa e a maneira como os usuários humanos interpretam o conteúdo da imagem. Dessa maneira, para atender às demandas da aplicação, sistemas de visão computacional podem recorrer a estratégias que envolvem feições de imagem que não são intuitivas para a percepção humana, como histogramas, ou então desenvolver novos algoritmos para refletir a percepção visual. Esses fatos demonstram que a definição de feições de imagem é bem flexível, podendo acomodar casos particulares de aplicação.

Em vista disso, sistemas que usam feições para descrever ou recuperar imagens possuem formas de organização bastante diversas (SMEULDERS et al., 2000). O número de níveis abstração varia em cada aplicação, de acordo com a necessidade. Em geral, o conteúdo da imagem pode ser compreendido em três níveis de abstração com complexidade crescente e alvos diferentes (EAKINS; GRAHAM, 1999):

1. feições primitivas, como cor, textura e forma são úteis para algoritmos de visão de máquina;
2. feições lógicas, que descrevem objetos e são úteis para usuários humanos; e
3. feições abstratas, como a semântica da imagem são utilizados para compreender a imagem como um todo.

As feições primitivas são extraídas diretamente da imagem e são frequentemente de propósito geral. Já as feições lógicas e as abstratas resultam da composição de medidas primitivas e são específicas, pois dependem da aplicação. Por essa razão, a maioria dos sistemas usa somente feições primitivas (LEW et al., 2006). Os sistemas não oferecem interpretação das imagens, nem mesmo conceitos de médio nível para descrever os objetos. A falta de informação que separa a imagem da representação por feições é chamada de lacuna semântica (*semantic gap*) (SMEULDERS et al., 2000). Mesmo com essa lacuna, muitas aplicações atingem bons resultados na descrição e recuperação de imagens. Os

melhores resultados são atingidos em sistemas especializados em um domínio limitado, como em um determinado tipo de imagem médica (MULLER et al., 2004). Quanto mais especializado for o sistema, menor a lacuna entre a imagem e as feições.

As próximas subseções apresentam e discutem algumas feições de baixo e médio nível de abstração.

2.3.2.1 Feições Primitivas

Feições primitivas são extraídas por meio de algoritmos diretamente das imagens. Elas geralmente são usadas para representar forma, cor e textura. Alguns exemplos dessas medidas são: circularidade, retangularidade, momentos centrais, excentricidade, orientação do maior eixo, histograma de intensidades e matriz de co-ocorrência. Aqui são apresentados os conceitos referentes a momentos centrais, orientação do maior eixo e centróide, que são utilizados na análise das imagens de cintilografia renal.

2.3.2.1.1 Momentos

Objetos identificados em uma imagem podem ser descritos por um conjunto de momentos m_{pq} (VELTKAMP, 2001). Assim, dada uma imagem binária $f(x, y)$, os momentos de ordem $(p + q)$ são definidos como (HU, 1962)

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y), \quad (2.35)$$

onde $p, q = 0, 1, 2, \dots, \infty$

Os momentos centrais de $f(x, y)$ são definidos como

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y), \quad (2.36)$$

onde

$$\bar{x} = m_{10}/m_{00}, \quad \bar{y} = m_{01}/m_{00} \quad (2.37)$$

são as coordenadas do centróide da imagem. O momento m_{00} representa a área do objeto binário.

Os momentos são úteis para descrever regiões depois da segmentação (CHEN et al., 2005). Algumas propriedades da região que podem ser encontradas por meio dos momentos incluem área, centróide e orientação.

Os momentos centrais são computados usando o centro de gravidade do objeto como origem e, por essa razão, são invariantes à translação. No caso da área estar normalizada para uma constante c , os momentos também são invariantes à escala. Porém, eles não são invariantes à rotação.

Para representar forma, o momento é descrito por somente um número, podendo ser facilmente indexado para recuperação. Entretanto, o poder de discriminação de somente um momento não é bom porque ele não é único. Duas formas diferentes podem ter o mesmo valor de momento. Dessa maneira, é necessário um grande número de momentos para distinguir formas similares (VELTKAMP, 2001).

2.3.2.1.2 Orientação do Maior Eixo

A orientação de um objeto é definida como o ângulo entre o eixo x e o eixo ao redor do qual o objeto pode ser rotacionado com o mínimo de inércia (JAHNE, 1997). Esse

é o maior eixo da elipse que tem o mesmo segundo momento central normalizado que a região (S. X. LIAO, 1996), como pode ser visto na Figura 2.13.

Então, a orientação do maior eixo é dada por:

$$\theta = \frac{1}{2} \arctan \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}, \quad (2.38)$$

A orientação é útil porque as formas podem ser rotacionadas de maneira a serem alinhadas com o eixo vertical, fazendo com que os objetos fiquem com orientação de 90° . Dessa maneira, comparações invariantes à rotação podem ser feitas, porque os ângulos de todos os objetos têm a mesma correspondência. Essa condição é válida para formas que possuem uma dimensão mais alongada que as demais.

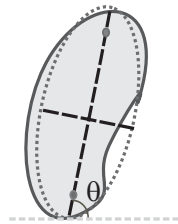


Figura 2.13: Orientação do maior eixo de uma forma, encontrada a partir da elipse que tem o mesmo segundo momento normalizado que a região.

2.3.2.1.3 Centróide

Dada uma imagem $f(x, y)$, as coordenadas do centróide $c = (c_x, c_y)$ são determinadas por (JAHNE, 1997):

$$c_x = \sum_{i=1}^n x_i/n, \quad c_y = \sum_{i=1}^n y_i/n, \quad (2.39)$$

onde n é o número total de *pixels* da imagem.

O centróide também pode ser encontrado por meio dos momentos da imagem, tal como mostrado na Equação 2.37.

Para calcular o centróide de um objeto previamente segmentado, basta utilizar somente os *pixels* do objeto em questão. O valor de n deve ser igual à área do objeto. Além disso, esse cálculo utiliza os *pixels* como valores binários e não as intensidades em níveis de cinza. Se fossem utilizados os valores das intensidades, a medida obtida seria o centro de massa do objeto. O centro de massa sofre influência da variação entre regiões homogêneas e heterogêneas.

A localização do centróide oferece uma estimativa para o centro de um objeto (Figura 2.14), fornecendo um ponto de referência que minimiza as distâncias entre esse ponto e as bordas da forma.

2.3.2.2 Feições Intermediárias

As feições intermediárias têm um maior nível de abstração e são formadas pela combinação das medidas primitivas. A seguir são apresentados os conceitos referentes a duas feições intermediárias: o perfil de distâncias do centróide à borda e a curvatura. Essas duas feições são utilizadas para a representar informações de forma, tamanho e outras características dos rins extraídos das imagens de cintilografia renal.



Figura 2.14: Localização do centróide c de uma região.

2.3.2.2.1 Distância do Centróide à Borda

Para representar a forma de um objeto, podem-se usar todos os pontos da borda. Entretanto, algumas formas possuem bordas com um número grande de pontos. Além disso, objetos de mesma natureza possuem bordas bastante diferentes e que, conseqüentemente, possuem tamanhos diferentes. Para contornar essa questão, pode-se optar por usar um número fixo de pontos da borda, que devem ser amostrados da mesma maneira em todos os objetos. Assim, realiza-se um mapeamento dos pontos das bordas para uma seqüência reduzida de pontos, que mantém as propriedades das bordas e serve para sua representação. Uma maneira de se chegar a essa representação é usar as distâncias que existem entre o centróide do objeto e as bordas. A seqüência de valores obtida é denominada perfil de distâncias centroidais ou simplesmente distâncias do centróide à borda (PARK; D., 1996).

Definindo M como o número de pontos de amostra $b = (b_x, b_y)$ ao longo da borda do objeto $B = (b_1, b_2, \dots, b_M)$, a distância Euclidiana d entre o centróide c e cada um dos pontos de amostra é computada da seguinte maneira:

$$d(k) = \sqrt{(c_x - b_{kx})^2 + (c_y - b_{ky})^2}, \quad k = 1 \dots M \quad (2.40)$$

O comprimento da seqüência é determinado pelo número de pontos amostrados ao longo da borda. Essa amostragem pode ser feita tanto em intervalos eqüidistantes como eqüiangulares (PARK; D., 1996).

Para obter a seqüência de amostragem eqüiangular, dada por (d_0, d_1, \dots, d_M) , pode-se rotacionar sucessivamente um raio a partir do centróide, em sentido anti-horário, em um ângulo fixo $\Delta\theta$, onde $\Delta\theta = 2\pi/M$, para um M positivo. A cada passo da rotação, guarda-se a distância do ponto de intersecção do raio com a borda, como pode ser visto na Figura 2.15.

Esse método de representação e descrição da forma de um objeto é adequado para objetos convexos (LIN et al., 2004). Para objetos que contêm concavidades, o raio pode ter intersecção com mais de um ponto. Se somente o ponto de menor distância for usado, algumas saliências podem ser perdidas. Assim, opta-se por usar a maior distância de intersecção.

A alternativa de usar intervalos eqüiangulares é bastante interessante porque é possível enxergar o perfil de distâncias como uma maneira de representar a borda em coordenadas polares, onde se tem uma distância (dada por d) e um ângulo, que pode ser obtido facilmente pela posição do ponto na seqüência. Dado que um ponto está na posição i , seu ângulo em radianos é dado por:

$$\theta_i = \frac{2\pi i}{M} \quad (2.41)$$

O perfil pode ser mostrado como uma função unidimensional das distâncias, como pode ser visto na Figura 2.16.

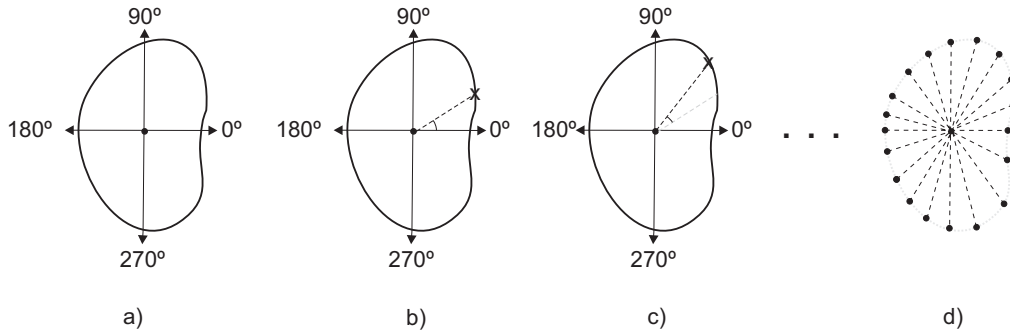


Figura 2.15: Cálculo das distâncias do centróide à borda: a) alinhamento da forma com o eixo vertical; b) primeiro raio traçado a partir do centróide, determinando um ponto de intersecção com a borda; c) rotação do raio; e d) pontos de intersecção encontrados após a rotação completa do raio.

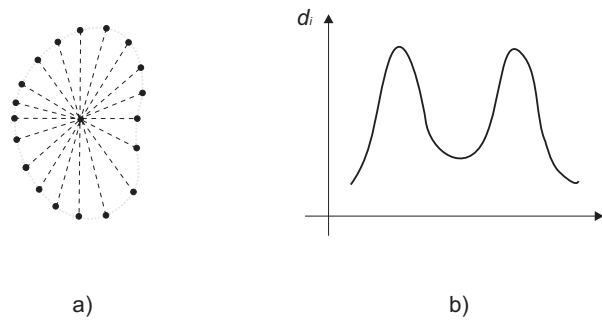


Figura 2.16: Perfil de distâncias centroidais.

2.3.2.2.2 Curvatura

A curvatura de uma função é definida como a derivada do vetor tangente à curva (ABBASI; MOKHTARIAN; KITTLER, 1999). Considerando uma equação paramétrica de um vetor para a curva $\Gamma(u) = (x(u), y(u))$, onde u é um parâmetro arbitrário, a função da curvatura pode ser expressa como:

$$k(u) = \frac{x'(u)y''(u) - x''(u)y'(u)}{(x'^2(u) + y'^2(u))^{3/2}} \quad (2.42)$$

Se cada componente da curva Γ for convoluído com uma máscara Gaussiana unidimensional $g(u, \sigma)$, de tamanho σ , então $X(u, \sigma)$ e $Y(u, \sigma)$ representam os componentes da curva resultante, dada por Γ_σ :

$$X(u, \sigma) = x(u) * g(u, \sigma), \quad Y(u, \sigma) = y(u) * g(u, \sigma) \quad (2.43)$$

Aplicando as propriedades da convolução, a primeira e a segunda derivadas dos componentes são dadas por:

$$X_u(u, \sigma) = x(u) * g_u(u, \sigma) \quad (2.44)$$

e

$$X_{uu}(u, \sigma) = x(u) * g_{uu}(u, \sigma) \quad (2.45)$$

$Y_u(u, \sigma)$ and $Y_{uu}(u, \sigma)$ podem ser calculadas de maneira similar. Então, a curvatura de Γ_σ pode ser expressa como:

$$k(u, \sigma) = \frac{X_u(u, \sigma)Y_{uu}(u, \sigma) - X_{uu}(u, \sigma)Y_u(u, \sigma)}{(X_u(u, \sigma)^2 + Y_u(u, \sigma)^2)^{3/2}} \quad (2.46)$$

A medida que o desvio padrão σ aumenta, a curva convoluída muda. A representação por curvaturas no espaço-escala (*Curvature Scale-Space* – CSS) utiliza-se dessa propriedade para representar formas em mais de uma escala (MOKHTARIAN; MACKWORTH, 1992). Essa maneira de representação permite comparar uma forma em diversos níveis de suavização. Em uma escala menor, com σ pequeno, pequenas variações na forma podem determinar grandes diferenças na comparação. Entretanto, em uma escala maior, a forma suavizada pode ser comparada com menos detalhes de representação. Quando σ torna-se suficientemente alto, a curva torna-se convexa.

A curvatura de uma forma é útil porque cada contorno corresponde a uma concavidade ou convexidade do objeto analisado. Qualquer deformação local na forma causa uma mudança no contorno correspondente da curvatura.

2.3.2.3 Feições de alto nível

As feições de alto nível dependem da aplicação em questão. Elas são diretamente relacionadas à semântica da imagem. Dessa forma, cada aplicação tem o seu conjunto particular de feições de alto nível, que podem ser construídas a partir de feições primitivas e intermediárias.

2.4 Análise de Componentes Principais

A Análise Multivariada é um ramo da Estatística responsável por examinar como variáveis dependentes relacionam-se em um determinado conjunto de dados. Existem muitas técnicas de análise multivariada, porém todas têm como objetivo principal a redução ou simplificação da complexidade do problema em análise.

A análise de componentes principais (*Principal Component Analysis* – PCA), também conhecida como transformada de *Hotelling* ou transformada de *Karhunen-Loève* (KLT), é uma técnica da Análise Multivariada que tem aplicações em campos como reconhecimento de faces e compressão de imagens (JAIN; DUIN; MAO, 2000).

Essa seção apresenta detalhes de funcionamento da PCA.

2.4.1 Definição

A PCA é uma técnica analítica que transforma um grupo de variáveis correlacionadas em um grupo de variáveis não-correlacionadas. Dessa maneira, ela permite reduzir a dimensionalidade de um conjunto de dados (JOLLIFFE, 2002).

Dado um conjunto de variáveis, a PCA busca reescrever as coordenadas dessas variáveis em outro sistema, que deve reter o máximo possível da informação contida nos dados originais.

A partir de n variáveis a PCA gera, por meio de combinações lineares, n componentes principais. Esses componentes são ortogonais e são fornecidos em ordem decrescente de variância. Dessa maneira, a primeira componente detém mais informação estatística que a segunda componente e assim sucessivamente. Frequentemente, um pequeno número de componentes pode ser usado para representar as variáveis originais. Assim, a técnica

permite a redução da dimensionalidade dos dados, porque é comum obter mais de 90% da informação em apenas duas ou três das primeiras componentes .

O processamento da PCA pode partir da matriz de variâncias e covariâncias ou da matriz de correlação (JOLLIFFE, 2002). A seguir são explicados os passos para o cálculo da PCA utilizando a matriz de covariâncias.

2.4.2 Covariância

Medidas estatísticas como o desvio padrão e a variância são aplicáveis a dados unidimensionais. Para dados de dimensões maiores, essas medidas só podem ser aplicadas se forem calculadas independentemente para cada dimensão. Entretanto, analisar o comportamento de uma variável em relação à outra é bastante útil em muitos casos. A medida que fornece tal poder de análise é a covariância.

A covariância é medida sempre entre 2 dimensões. Se a dimensão for a mesma, ou seja, se for calculada a covariância entre uma dimensão e ela mesma, o resultado será a variância. Dado que X é a primeira dimensão e Y a segunda, a fórmula da covariância entre X e Y é:

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n - 1)} \quad (2.47)$$

onde \bar{X} e \bar{Y} são as médias e n é o número de itens de dados. Quando os dados representam uma amostra de um conjunto, usa-se $n - 1$ no denominador. Quando os dados representam o conjunto total de dados, usa-se n no denominador.

2.4.3 Matriz de covariância

Quando mais de duas dimensões estiverem sendo analisadas, é preciso calcular a covariância entre cada par de dimensões. Para organizar essa medida, usa-se uma matriz de covariâncias. Para três dimensões, dadas por x , y e z , a matriz de covariância C pode ser organizada como:

$$C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{pmatrix} \quad (2.48)$$

A diagonal principal de C contém as variâncias e essa matriz é simétrica, já que $\text{cov}(a, b) = \text{cov}(b, a)$.

2.4.4 Autovetores e autovalores

Autovalores e autovetores são conceitos importantes da matemática e possuem aplicações práticas em diversas áreas, como processamento de imagens e mecânica dos sólidos (JOLLIFFE, 2002).

Considerando uma matriz quadrada A de ordem n , o vetor coluna x de tamanho n é autovetor dessa matriz se a multiplicação da matriz A por x resulta em um escalar λ múltiplo de x , ou seja, se

$$Ax = \lambda x \quad (2.49)$$

Se essa condição for verdadeira, então λ é o autovalor de A associado ao autovetor x .

Considerando I a matriz identidade de ordem n , λx pode ser reescrito como $\lambda x = \lambda Ix$. Substituindo essa expressão na Equação 2.49, fica

$$Ax = \lambda Ix \quad (2.50)$$

Ordenando e agrupando, resulta em

$$\lambda Ix - Ax = 0 \quad (2.51)$$

$$x(\lambda I - A) = 0 \quad (2.52)$$

A função $\det(\lambda I - A)$ é denominada função característica da matriz A . Para solução não nula da Equação 2.52, deve-se ter a seguinte condição

$$\det(\lambda I - A) = 0 \quad (2.53)$$

Resolvendo essa equação, obtêm-se os valores de λ que, substituídos na Equação 2.49, permitem a determinação dos autovetores x .

Todos os autovetores de uma matriz são ortogonais entre si, independentemente do número de dimensões. Essa propriedade é útil porque permite expressar os dados em termos dos autovetores encontrados, ao invés de usar o sistema de coordenadas cartesianas (x, y, z) (SMITH, 2002). Além disso, os autovetores podem ter sua escala modificada para qualquer comprimento, sem alterar suas propriedades, já que a direção permanece a mesma. O que muda é o autovalor.

2.4.5 Componentes principais

Como mencionado anteriormente, a PCA fornece componentes ordenadas de forma decrescente de variância. Essa característica da PCA permite expressar os dados de maneira a ressaltar suas similaridades e diferenças (THEODORIDIS; KOUTROUMBAS, 2003). Isso é bastante útil principalmente em dados com grandes dimensões, nos quais não se pode fazer uso de uma representação gráfica.

Além de fornecer uma maneira diferente de expressar os dados, a PCA também permite comprimi-los. Isso pode ser feito pela redução do número de dimensões, sem a perda de muita informação, já que as primeiras componentes detêm grande parte da informação do conjunto de dados.

Para calcular as componentes principais de um conjunto de dados, primeiramente deve-se obter um conjunto de dados a serem analisados, de d dimensões. Depois, calcula-se a média de cada uma das dimensões separadamente e subtrai-se essa média de cada uma dessas dimensões. Esse passo faz com que o conjunto de dados fique com média zero. Depois, o cálculo da matriz de covariância cov é realizado, conforme discutido na subseção 2.4.3. A matriz cov será de ordem d . O próximo passo consiste em calcular os autovetores e autovalores da matriz cov . Agora, os autovetores podem ser ordenados pelos seus autovalores associados de maneira decrescente. Esse ordenamento irá fornecer as componentes em ordem de significância. O autovetor com maior autovalor associado corresponde à componente principal dos dados.

Depois desses passos, pode-se usar as componentes para visualização, compressão (descartando componentes de menor significância), reconhecimento de padrões e outras aplicações.

A Figura 2.17 mostra um exemplo de aplicação da PCA em dados de 2 dimensões. A componente principal, dada pelo eixo pontilhado $P1$, é a dimensão que representa melhor a distribuição dos dados.

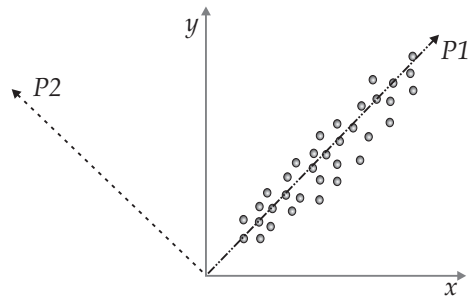


Figura 2.17: Aplicação da PCA a dados de 2 dimensões: o eixo $P1$ mostra a distribuição principal dos dados, que tem o autovetor com maior autovalor associado. $P2$ representa a segunda componente principal, perpendicular à $P1$.

2.4.6 Aplicação da PCA em problemas de processamento de imagens

Devido ao potencial de redução da dimensionalidade dos dados e do fato de ressaltar similaridades e diferenças em um conjunto de dados, a PCA encontra inúmeras aplicações nas áreas do processamento de imagens e da visão computacional. Uma das aplicações mais recorrentes está na detecção de faces (ZHAO et al., 2003). Entretanto, o mesmo princípio pode ser aplicado a outros tipos de imagens, como em imagens médicas (ROWE; HOFFMANN, 2006), ou aplicações variadas, como a classificação de pedestres em imagens (MUNDER; GAVRILA, 2006) ou a análise de vídeos (SAHOURIA; ZAKHOR, 1999). A compressão de imagens é outra aplicação comum da transformada (OLIVEIRA et al., 2000).

3 UMA PROPOSTA PARA QUANTIFICAR IMPRESSÕES DIAGNÓSTICAS EM IMAGENS DE CINTILOGRAFIA RENAL

Este capítulo apresenta o processo utilizado para obtenção das feições das imagens de cintilografia renal. A seção 3.1 apresenta detalhes sobre o exame de cintilografia e sobre as imagens resultantes desse exame, além de descrever a metodologia de obtenção dos dados. Depois, o processo de obtenção das feições é apresentado. Esse processo pode ser dividido em duas partes: a segmentação da imagem e a extração das feições. A primeira etapa é descrita na seção 3.2, que explica o método desenvolvido e faz comparações com outros métodos de segmentação. Depois, a etapa de extração das feições é detalhada na seção 3.3.

3.1 O exame de cintilografia renal

A cintilografia é um exame que capta imagens de processos fisiológicos usando isótopos radioativos. Esses isótopos, chamados de radiotraçadores ou radiofármacos, são retidos em determinado órgão e emitem radiação gama. A radiação é captada por uma Gama-câmera, gerando as imagens de cintilografia (HILSON, 2003).

Esse exame pode ser realizado em diversos órgãos, como no miocárdio, nos pulmões, nos ossos e nos rins. No caso da cintilografia renal, o exame fornece imagens funcionais dos rins (Figura 3.1), ou seja, mostra como as células estão desempenhando suas funções. Devido à permanência prolongada do traçador no córtex renal, o exame fornece informações principalmente relativas à concentração do mesmo no parênquima renal, morfologia e dimensões dos rins.

Existem duas modalidades de cintilografia renal: a estática e a dinâmica. Os estudos estáticos são realizados após a injeção intravenosa de pequena dose de radiofármacos que são retidos na camada cortical dos rins. As imagens são obtidas no momento de maior concentração do radiotraçador nos rins (SANDLER et al., 1996). O radiotraçador comumente usado é o ácido dimercapto-sucínico (^{99m}Tc -DMSA). Ele é fixado pelos túbulos do córtex renal separadamente para cada rim. Essa fixação se dá em proporção linear ao grau de funcionamento do rim. Dessa maneira, doenças com comprometimento de apenas um rim podem ser diagnosticadas e acompanhadas na sua evolução. Nesse trabalho, são utilizadas somente cintilografas estáticas. Na cintilografia renal dinâmica, a gravação das imagens é feita desde o momento da injeção intravenosa do radiofármaco (em geral ^{99m}Tc -DTPA, MAG3), permitindo acompanhamento não invasivo da sua rota de excreção metabólica desde os rins até a bexiga.

A cintilografia auxilia no diagnóstico de pielonefrites, definição de anomalias, locali-

zação de rim ectópico, avaliação de trauma e lesões (TAYLOR JR; NALLY, 1995).

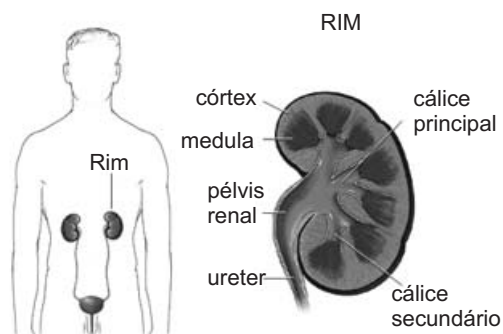


Figura 3.1: Localização e anatomia dos rins.

3.1.1 Características das imagens de cintilografia renal

Nas imagens de cintilografia renal estática, os rins aparecem como regiões claras sobre um fundo escuro. A forma dos rins varia bastante de uma imagem para outra, como pode ser visto nas Figuras 3.2(a) e 3.2(b). Os rins aparecem como regiões com intensidades não homogêneas, geralmente com mais de um máximo regional (Figura 3.2(c)). Além disso, podem apresentar bordas difusas (Figura 3.2(d)).

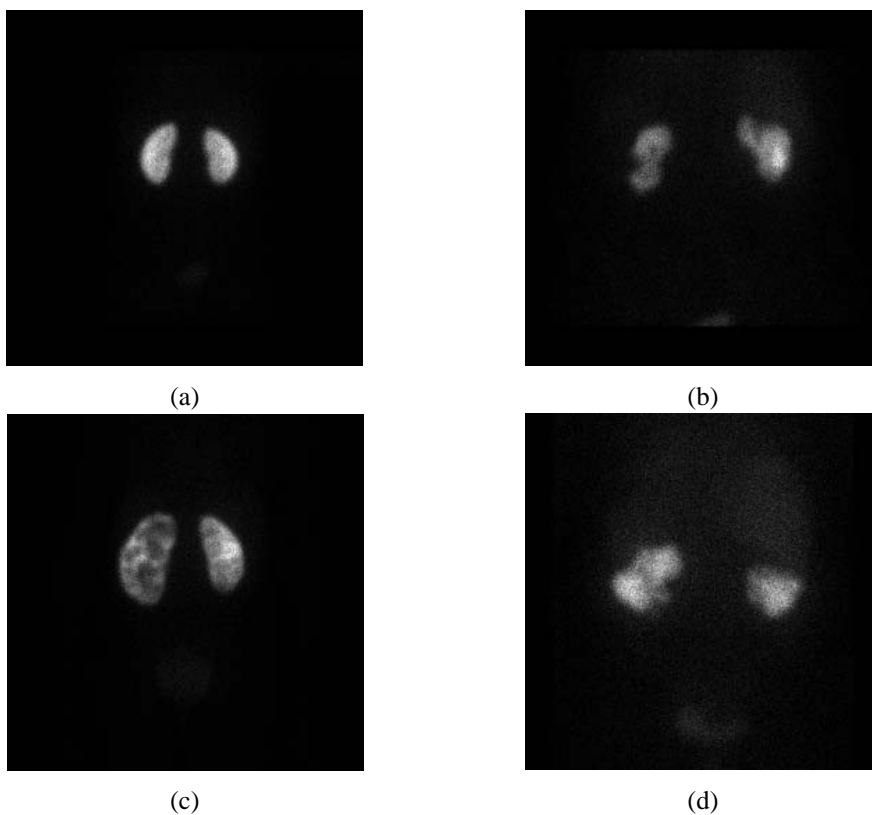


Figura 3.2: Características das imagens de cintilografia renal: (a) rins com formas usuais; (b) rins com formas não usuais; (c) diferença de intensidade dentro do mesmo rim; e (d) bordas difusas.

3.1.2 Metodologia de coleta das imagens

A coleta das imagens foi feita em um momento posterior à realização do exame, quando as elas já estavam armazenadas no servidor e já haviam sido avaliadas por um especialista.

As imagens utilizadas para extração de características foram fornecidas no formato DICOM (*Digital Imaging and Communications in Medicine*), sendo utilizada somente a vista posterior de cada caso. Os respectivos laudos dos exames também foram utilizados. Entretanto, tanto as informações do laudo quanto das imagens foram anonimizadas antes de serem manipuladas, ou seja, qualquer informação pessoal, como nome, idade e sexo foram excluídas do arquivo.

As imagens foram coletadas semanalmente, durante cinco semanas não-consecutivas. Todos os exames realizados a cada semana foram coletados, sem nenhum critério de escolha. Isso foi feito porque o HCPA armazena por um curto período as imagens em formato digital. Dessa maneira, o conjunto de dados contém tanto casos anormais como normais sem proporções definidas, e, naturalmente, desiguais.

3.2 Segmentação das imagens de cintilografia renal

A primeira etapa necessária para permitir a extração das feições das imagens de cintilografia renal consiste em segmentar a região dos rins do fundo da imagem. Como mencionado anteriormente, na seção 3.1.1, essas imagens possuem características bem específicas, que dificultam a utilização de métodos clássicos de segmentação.

Existem referências na literatura que tratam da segmentação de imagens com características parecidas, ou seja, que tentam segmentar regiões claras não-homogêneas em um fundo escuro com intensidades variadas. Problemas desse tipo são bastante comuns, pois diversos tipos de imagens possuem essas características. Exemplos incluem microcalcificações em mamografias (THANGAVEL et al., 2005), genes de *microarray* (SIDDIQUI; HERO; SIDDIQUI, 2002), drusas em imagens de fundo de olho (SBEH et al., 2001) entre outras aplicações, principalmente na área médica.

Alguns métodos existentes foram testados para a segmentação das imagens de cintilografia renal. Esses métodos incluem técnicas clássicas, como detecção de bordas, limiarização e *watersheds* e também o método desenvolvido por Sbeh et al. (SBEH et al., 2001) para imagens de fundo de olho, devido à existência de características semelhantes entre os dois tipos de imagens.

3.2.1 Aplicação de métodos existentes na literatura

O método desenvolvido por Sbeh et al. (SBEH et al., 2001) é baseado em uma função adaptativa para a detecção do contraste localmente. Ele utiliza métodos da morfologia matemática, analisando a imagem como um relevo topográfico. O método obtém bons resultados na segmentação de drusas em imagens de fundo de olho. Entretanto, existe a necessidade de ajuste de parâmetros por parte do usuário. Esses parâmetros estão relacionados à intensidade dos *pixels* e à distância entre as estruturas a serem detectadas. Dessa forma, esse método torna-se pouco atraente para a segmentação de um grande número de casos por não ser totalmente automatizado. Além disso, um usuário da área médica pode ter dificuldade para ajustar esses valores. A Figura 3.3(a) mostra o resultado da aplicação desse método para uma imagem de cintilografia renal, mostrando as visões bidimensionais e tridimensionais da segmentação. Observando essas imagens, pode-se chegar a

conclusão de que o método consegue detectar as regiões mais claras dos rins, mas ele sofre influência das regiões mais escuras no interior do rim, levando a um resultado insatisfatório.

A segunda técnica testada, que usa a detecção de cruzamentos por zero depois da aplicação do filtro Laplaciano de Gaussiano, apresenta resultados satisfatórios em muitos casos. Entretanto, observando as imagens em três dimensões, é possível verificar que algumas bordas cortam os rins em níveis de cinza acima do desejado, fazendo com que o rim pareça menor do que realmente é. A Figura 3.3(b) mostra o resultado da aplicação desse método para o mesmo caso da Figura 3.3(a). A segmentação pode ser melhorada eliminando as bordas não-fechadas, mas, mesmo assim, o método falha ao detectar o interior do rim como outra região.

A detecção de bordas pelo método Canny, terceiro método testado, apresenta resultados semelhantes aos da detecção de cruzamentos por zero depois da aplicação do filtro Laplaciano de Gaussiano. Isso se deve ao fato de que esses dois métodos têm por objetivo a detecção de bordas, o que nem sempre leva a uma segmentação em regiões. Os resultados da aplicação do Canny podem ser vistos na Figura 3.3(c).

O resultado do quarto método testado, que consiste na segmentação por limiarização utilizando o método de Otsu (OTSU, 1979), pode ser visto na Figura 3.3(d). Esse método busca encontrar um limiar que minimiza a variância intraclasse dos *pixels* escuros e claros. No entanto, por atuar na imagem globalmente, ou seja, com base no histograma de toda a imagem, ele leva a resultados insatisfatórios na segmentação dos dois rins simultaneamente.

A aplicação da segmentação por *watersheds* gera sobre-segmentação da imagem, como pode ser visto na Figura 3.3(e). Uma análise cuidadosa desse resultado mostra, no entanto, que existem linhas *watersheds* sobrepondo a maioria das bordas reais dos rins. Pode-se pensar que um pós-processamento é capaz de eliminar essas linhas excedentes. Porém, o rim esquerdo apresenta uma linha cruzando o seu interior, o que tornaria difícil definir um critério para remoção das linhas.

Com base nesses resultados, pode-se afirmar que esses métodos não são adequados para segmentar as imagens de cintilografia renal satisfatoriamente. Uma das principais inferências que podem ser feitas após analisar os resultados desses métodos é que os rins precisam ser considerados separadamente durante a segmentação, pois muitos casos possuem características bastante diversas.

É provável que grande parte dos casos apresente resultados aceitáveis utilizando algum desses métodos clássicos. Entretanto, a escolha do método deve levar em consideração que os resultados da segmentação têm impacto direto sobre as feições extraídas. Dessa forma, um método alternativo de segmentação foi desenvolvido, visando minimizar os erros de segmentação para maximizar a fidelidade das feições calculadas. O resultado obtido por esse método pode ser visto na Figura 3.3(f). A próxima seção apresenta detalhes desse método.

3.2.2 Explicação do método desenvolvido

As imagens de cintilografia renal são freqüentemente ruidosas e podem ter grande variação de intensidade, apresentando tanto picos como vales no interior da região dos rins. Para que se atinja resultados de segmentação aceitáveis, é preciso reduzir o ruído e tornar as regiões dos rins um pouco mais homogêneas, para que o processo de segmentação não seja influenciado por oscilações que pertencem naturalmente ao interior do rim.

Inicialmente, as regiões escuras presentes no interior dos rins devem ser removidas

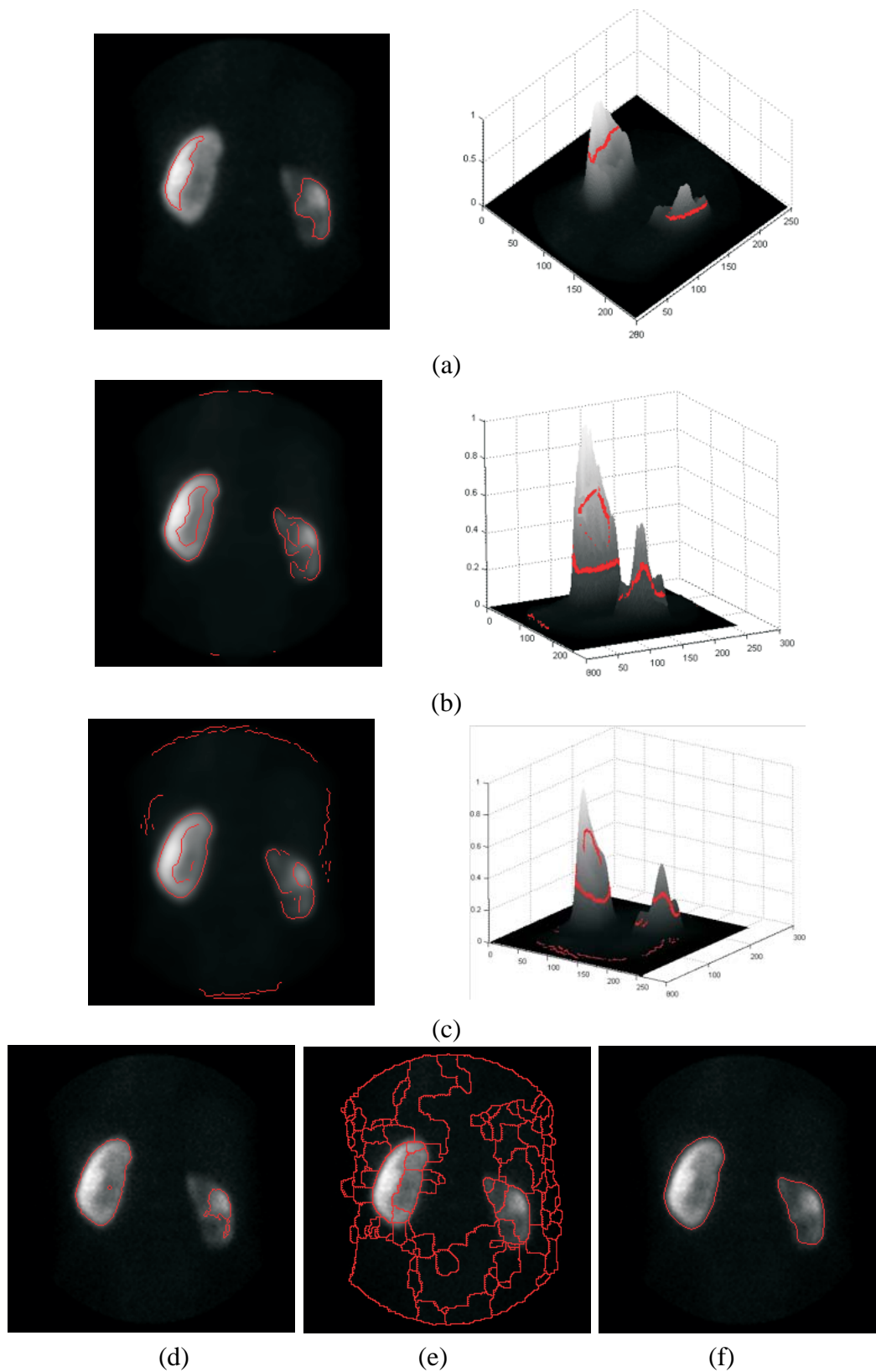


Figura 3.3: Comparação entre métodos de segmentação: (a) resultado da segmentação pelo método proposto por Sbeh et al.; (b) resultado utilizando a detecção de cruzamentos por zero depois da aplicação do filtro Laplaciano de Gaussiano; (c) resultado utilizando o método de Canny para detecção de bordas; (d) limiarização pelo método proposto por Otsu; (e) segmentação por *watersheds* e (f) resultado do método proposto.

para evitar que, durante o processo de segmentação, elas sejam interpretadas como fundo da imagem. A justificativa para esse procedimento é dada pelo fato de que um rim nunca será um objeto “furado”, ou seja, nunca existirão buracos dentro do órgão que deixem aparecer o fundo da imagem. O que pode ocorrer é não haver captação em determinadas regiões, mas, para fins de segmentação, essas regiões pertencem ao interior do rim.

Para preencher buracos em uma imagem, as regiões mais escuras cercadas por regiões mais claras passam a ter suas intensidades iguais às das regiões claras. Dessa forma, os mínimos regionais que não são conectados às bordas da imagem serão removidos. O procedimento para preencher buracos usa a técnica da reconstrução morfológica (Seção 2.1.5.4). A imagem marcadora J tem intensidade igual a 0 onde não for borda e 1 nas bordas e a máscara é igual ao complemento da imagem original I . Assim, a imagem resultante do preenchimento de buracos I_F é dada por

$$I_F = [\rho_{I^c}(J)]^c \quad (3.1)$$

O resultado dessa operação, aplicado à Figura 3.4(a), pode ser visto na Figura 3.4(b).

Para tornar a imagem I_F mais homogênea e menos ruidosa, um filtro Gaussiano G de desvio padrão σ é aplicado

$$I_G = I_F * G_\sigma \quad (3.2)$$

Esse passo leva ao resultado mostrado na Figura 3.4(c).

Agora que os buracos estão preenchidos e suavizados, é necessário remover essas regiões de mínimos regionais formadas pelo preenchimento de buracos, ou seja, é preciso remover esses vales. A imagem I_G possui regiões planas, achatadas, porém ainda escuras. Assim, é aplicado um fechamento em níveis de cinza

$$I_C = \phi(I_G) \quad (3.3)$$

O fechamento é aplicado pois, conforme os efeitos listados na Seção 2.1.4.2, esse operador possui a propriedade de eliminar vales inferiores em tamanho ao elemento estruturante, conectando picos próximos. O resultado dessa operação é mostrado na Figura 3.4(d).

Dependendo das características da imagem I_C , do formato e do tamanho do elemento estruturante utilizado, as regiões de interação entre a imagem e o elemento estruturante podem ficar muito marcadas, fazendo com que a imagem exiba um padrão quadriculado, ou com outro formato, de acordo com a situação. Para suavizar esse efeito, a imagem I_C é novamente convoluída por uma máscara Gaussiana

$$I_{G2} = I_C * G_\sigma \quad (3.4)$$

O resultado dessa operação é mostrado na Figura 3.4(e).

Até esse passo pode ser considerado o pré-processamento da imagem. A partir desse ponto, todos os procedimentos adotados são etapas da segmentação da imagem.

No caso ideal, cada região de máximo corresponde a um rim na imagem. Contudo, mesmo com os processamentos anteriores, alguns rins podem apresentar mais de um pico (máximo) em seu interior. Além disso, regiões extra-renais, como a bexiga podem apresentar elevada captação do radiofármaco. Por isso, algumas restrições devem ser impostas para a detecção de máximos, de forma que somente as regiões relevantes, isto é, potenciais rins, sejam detectadas. Então, aplica-se a operação de h -máximo estendido para identificar grupos de *pixels* que têm suas intensidades significativamente maiores que os seus

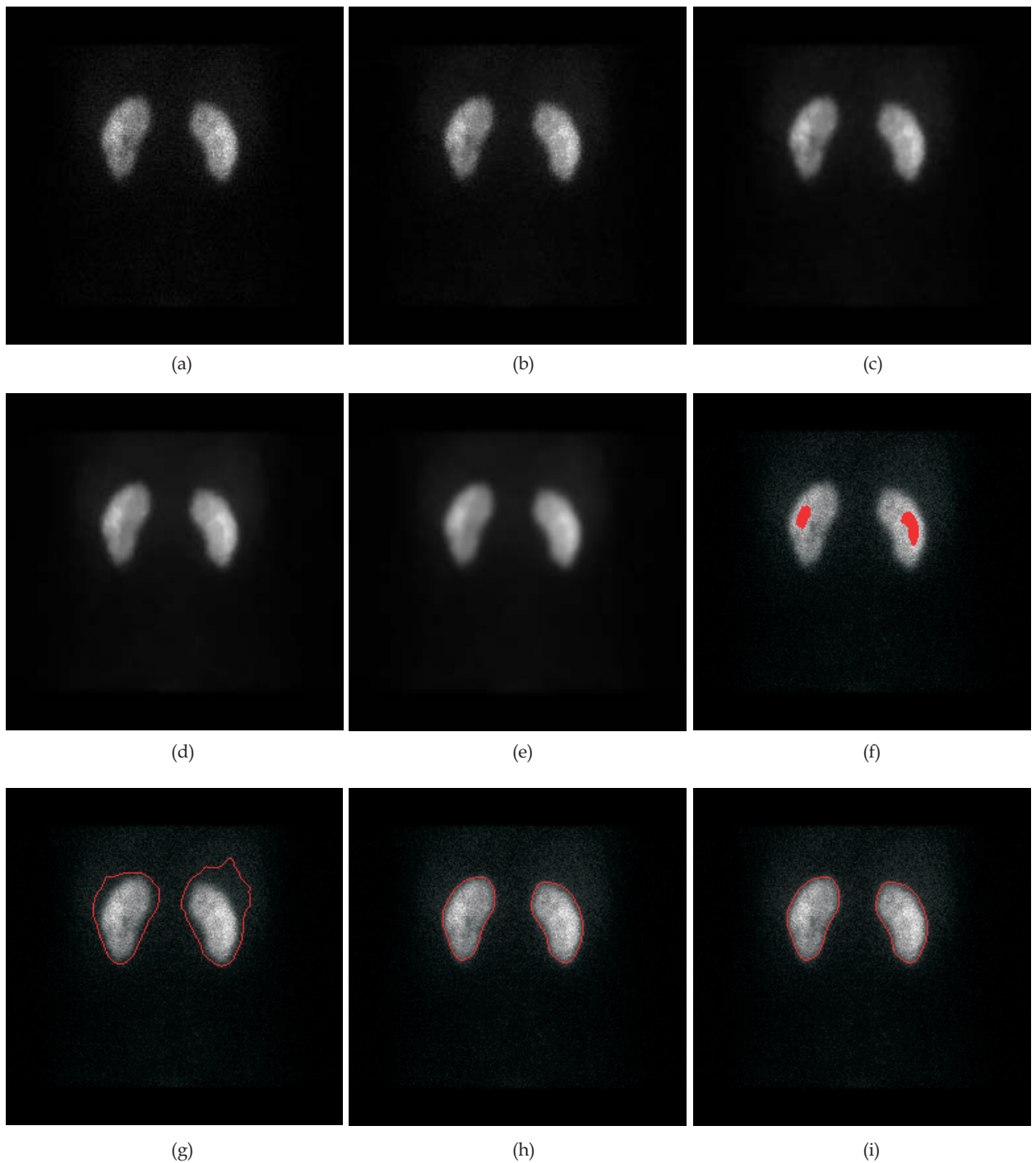


Figura 3.4: Processo de segmentação: (a) imagem original; (b) imagem I_F , resultado do preenchimento de buracos; (c) imagem I_G , resultante da filtragem Gaussiana; (d) imagem I_C , resultado do fechamento morfológico; (e) I_{G2} , segunda filtragem Gaussiana; (f) $EMAX_h$ em I_{G2} , demarcando os máximos; (g) limiarização por b ; (h) limiarização por med ; e (i) resultado otimizado pelo EM, gerando a segmentação final da imagem.

arredores imediatos. Ou seja, para ser considerado um rim em potencial, a região deve exceder um limiar local de contraste h em relação ao fundo da imagem. Isso implica que regiões de máximo que apresentem uma diferença maior que h entre o valor máximo de intensidade (topo do rim) e o valor de intensidade do fundo (base do rim) serão mantidas.

Como visto na Seção 2.1.7, o resultado da operação de h -máximo estendido $EMAX_h$ é uma imagem binária que determina as localizações dos máximos regionais da imagem. Dessa forma, a aplicação de $EMAX_h$ em I_{G^2} irá resultar em conjuntos de pontos que correspondem às regiões mais claras do rim, como pode ser visto na Figura 3.4(e), na qual o resultado de $EMAX_h$ foi sobreposto à imagem original. Esses conjuntos são bons pontos de partida para o processo de segmentação, pois marcam a existência de potenciais rins. O valor de h utilizado é constante para todos os casos avaliados.

A partir de agora, cada região $EMAX_h$ encontrada será processada separadamente, com o objetivo de analisar cada rim de forma independente. Assim, para cada região $EMAX_h$, calcula-se o valor de intensidade mínimo m . Sendo A_k o número de *pixels* (área) cuja intensidade é maior que k , onde $k = m, m - 1, \dots, b$ e $b < m$. Então, são computados sucessivos valores de área A_k correspondentes aos níveis de cinza do intervalo $[m, b]$. Esse processo é ilustrado na Figura 3.6, que mostra visões bidimensional e tridimensional do processo de segmentação. Em resumo, considerando a imagem como um relevo de intensidades, o processo de segmentação começa computando valores de área A_k do topo do rim (nível de cinza m) e desce em direção à base do rim (nível de cinza b).

Para estimar o valor de b , é analisado o comportamento da seqüência $s = \{A_k\}$, $k \in [m, b]$. É razoável assumir que s é uma curva que aumenta de forma lenta e gradual, exceto quando k atinge o valor de intensidade b , que corresponde ao fundo da imagem. Nesse momento, a área aumenta de maneira considerável, pois ultrapassa a região do rim, pegando grande parte do fundo. Assim, freqüentemente, b pode ser determinado pelo primeiro pico significativo na curva s , como mostra a Figura 3.5. Nesse caso, são mostradas as áreas (normalizadas entre 0 e 1) para cada nível de cinza da imagem (no eixo x do gráfico), que vão de 0 a 200. O pico localiza-se próximo à intensidade 150.

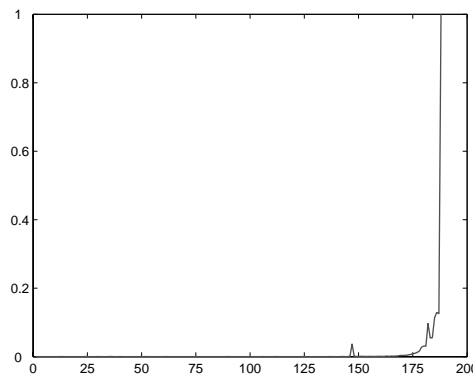


Figura 3.5: Curva de crescimento das áreas em função dos níveis de cinza.

Então, um limiar t em s é utilizado para estimar b . A estimativa inicial das bordas dos rins é obtida pelo uso do valor de intensidade med que corresponde à mediana de $[m, b]$ como limiar t para realizar a segmentação do rim. Como cada rim é processado separadamente, pode-se obter valores diferentes de med para os rins, ao invés de um

valor global para toda a imagem.

Entretanto, nem sempre o valor de med permite uma segmentação perfeita da forma do rim, isto é, a borda pode parecer um pouco “folgada”, como na Figura 3.8(a). Nesse caso, é necessário um passo de otimização. O algoritmo EM (Seção 2.2.5.1) é aplicado à uma faixa larga E em torno da estimativa inicial de bordas do rim para classificar os *pixels* como sendo pertencentes aos rins ou ao fundo da imagem.

O conjunto de *pixels* E é encontrado por meio da aplicação do gradiente morfológico (Seção 2.1.4.3) na imagem binária de bordas. Os níveis de cinza correspondentes aos *pixels* de E são utilizados como entrada para o algoritmo EM.

O EM deve classificar os *pixels* como pertencentes ao fundo ou ao rim. Para isso, ele estima as pdf's de cada uma delas. As duas classes são estimadas inicialmente como tendo probabilidades de ocorrência iguais, ou seja, 0.5 para cada. As variâncias e médias dessas distribuições também são estimadas. A Figura 3.7(a) mostra o histograma da imagem (em azul) e as duas distribuições iniciais (em verde). A curva em vermelho é o resultado do modelo de mistura. Na Figura 3.7(b) está o resultado do EM, o qual mostra um melhor ajuste das duas distribuições ao histograma.

A Figura 3.8(b) mostra a região de atuação do EM e a Figura 3.8(c) apresenta o resultado obtido com esse passo de otimização da segmentação.

3.3 Extração de feições das imagens de cintilografia renal

A extração das feições de interesse é feita a partir das imagens segmentadas, utilizando somente as regiões correspondentes aos rins. As feições foram organizadas em uma estrutura com três níveis de abstração: baixo, médio e alto (ver Seção 2.3).

O nível mais baixo é formado por feições primitivas extraídas diretamente das imagens. Essas feições incluem as bordas e outras medidas de localização espacial. No nível intermediário, medidas de mais alto nível de abstração são formadas pela combinação de feições primitivas. Por fim, as medidas de médio nível são combinadas para descrever os dados observados em um nível mais alto de abstração. A Figura 3.9 ilustra como estão organizadas as camadas dessa arquitetura de extração de feições.

Na prática, é mais fácil computar feições de baixo nível do que de alto nível que sejam ligadas ao conteúdo semântico. Isto porque relacionar conceitos abstratos humanos para feições de baixo nível é complexo. Dessa maneira, a complexidade das feições computadas é justificada pela relação direta com o conteúdo semântico da imagem.

Como mencionado no Capítulo 1, cinco feições de alto nível foram selecionadas para descrever o conteúdo das imagens de cintilografia renal. Recapitulando, as feições são as seguintes:

1. Posição relativa entre os dois rins;
2. Simetria entre o rim esquerdo e o direito;
3. Forma de cada um dos rins;
4. Captação do radiofármaco em cada um dos rins;
5. Comportamento dos contornos de cada um dos rins.

As próximas seções abordam como cada feição é computada, desde as de baixo nível até as de alto nível de abstração.

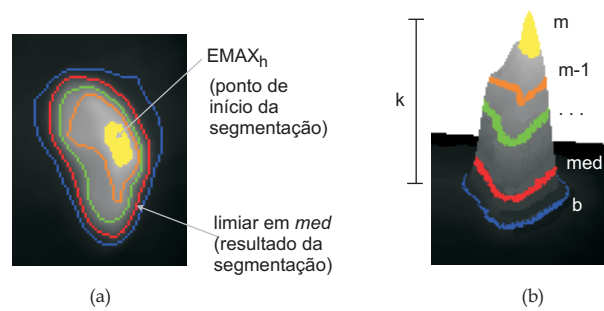


Figura 3.6: Processo de segmentação das imagens de cintilografia renal: (a) visão bidimensional; e (b) visão tridimensional. O conjunto de *pixels* no nível de cinza m (em amarelo) é o ponto inicial da segmentação. A intensidade é decrementada até chegar no limiar b (em azul). Então, o limiar med é calculado a partir da mediana do conjunto k , gerando a borda do rim (em vermelho).

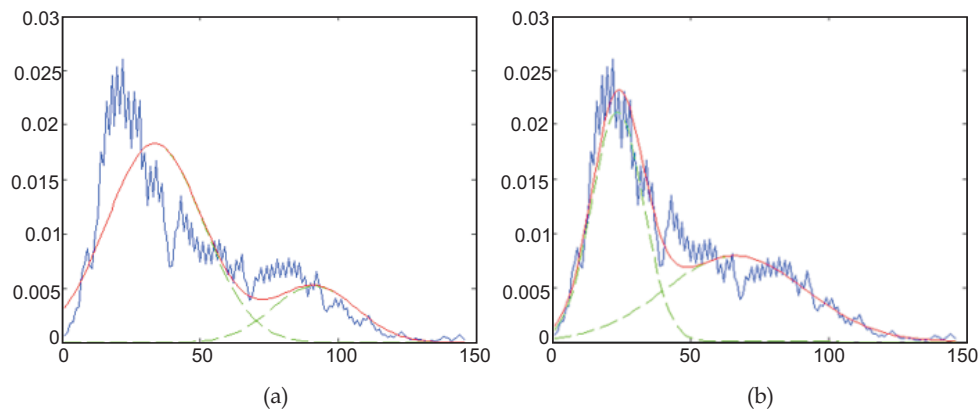


Figura 3.7: Aplicação do algoritmo EM para classificação dos *pixels*: (a) estimativa inicial das pdf's (em verde), sobrepostas ao histograma da imagem (em azul); (b) distribuição de probabilidades resultantes do EM, demonstrando melhor ajuste ao conjunto de dados.

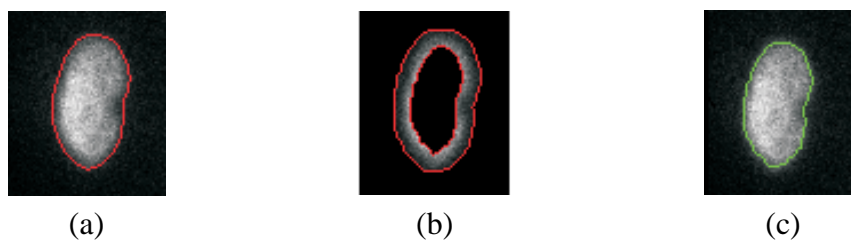


Figura 3.8: Passo de otimização da segmentação utilizando o algoritmo EM: (a) bordas que poderiam estar melhor ajustadas à forma do rim; (b) região onde o EM é aplicado; e (c) resultado final depois da aplicação do EM.

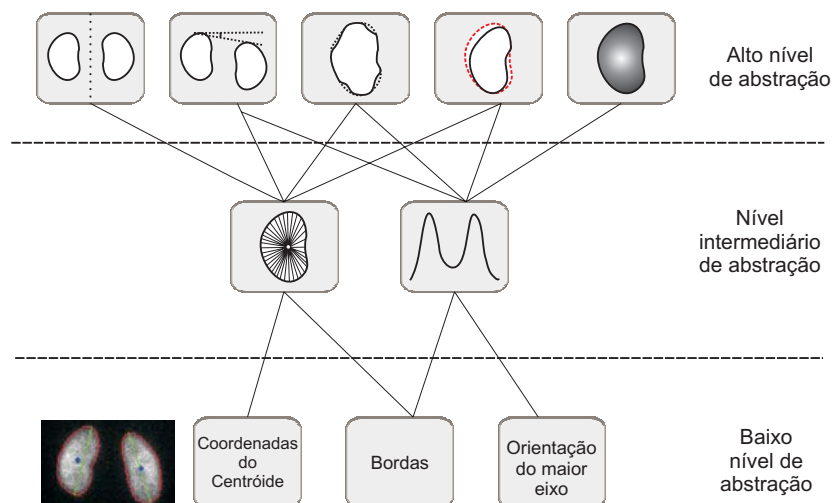


Figura 3.9: Representação do conteúdo de imagens de cintilografia renal em três níveis de abstração.

3.3.1 Feições de baixo nível

3.3.1.1 Bordas dos rins

As bordas dos rins são extraídas de acordo com o algoritmo apresentado na Seção 3.2. Nesse nível de abstração, todos os pontos da borda são considerados. O que está disponível é uma imagem binária das bordas previamente segmentadas de cada um dos casos.

3.3.1.2 Orientação do maior eixo

A orientação do maior eixo é calculada de acordo com o que foi apresentado na Seção 2.3.2.1.2. A informação armazenada é a medida, em graus, do ângulo formado entre o maior eixo do rim e o eixo horizontal.

A orientação, no caso das imagens de cintilografia renal, é útil porque todos os rins podem ser rotacionados, isto é, alinhados com o eixo vertical (90°). Dessa forma, comparações invariantes à rotação podem ser feitas, pois existe uma correspondência interna entre os ângulos de todos os rins.

A Figura 3.10 mostra quatro casos e os seus respectivos ângulos de orientação do maior eixo. O caso da Figura 3.10(a) tem orientação de 79° ; na Figura 3.10(b), a orientação é de 56° ; o caso da Figura 3.10(c) tem ângulo de orientação de 88° , aproximando-se muito da posição vertical; por fim, o caso da Figura 3.10(d) tem orientação de 17° , aproximando-se da posição horizontal. Esses exemplos mostram como as orientações podem variar consideravelmente de acordo com o caso.

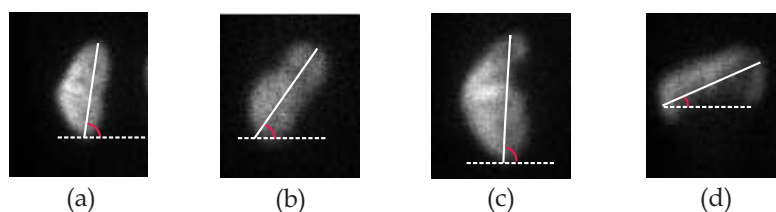


Figura 3.10: Exemplo da variação da orientação do maior eixo: (a) ângulo de 79° ; (b) ângulo de 56° ; (c) ângulo de 88° ; e (d) ângulo de 17° .

3.3.1.3 Centróide

O centróide dos rins é calculado utilizando a imagem de bordas preenchida, ou seja, a imagem binária do rim, conforme detalhado na Seção 2.3.2.1.3. O cálculo da posição do centróide fornece um centro geométrico para os rins. O centróide também serve como um ponto de referência que minimiza as distâncias até as bordas da forma.

A posição do centróide é usada principalmente como uma medida de localização relativa ao próprio rim, e não em relação à imagem inteira. A localização dos rins nas imagens varia, pois depende da posição em que o paciente estava na hora da aquisição. Dessa maneira, as coordenadas do centróide também variam consideravelmente.

3.3.2 Feições de nível intermediário

As três feições de baixo nível apresentadas são combinadas de maneira a gerar duas feições intermediárias: o perfil de distâncias centroidais e a curvatura. Essas feições são detalhadas a seguir.

3.3.2.1 Perfil de distâncias centroidais

A seqüência de distâncias do centróide à borda é calculada de acordo com os detalhes apresentados na seção 2.3.2.2.1, usando uma amostragem em intervalos equiângulares. A cada 5° um ponto da borda é amostrado, resultando em um total de 72 pontos que geram o vetor de distâncias $D = (d_1, d_2, \dots, d_{72})$.

O perfil de distâncias é usado nas feições de alto nível como representante das coordenadas das bordas dos rins. Conforme demonstrado na seção 2.3.2.2.1, o perfil centroidal substitui com vantagens a utilização de todos os pontos da borda, principalmente em comparações entre rins. A seqüência pode ser armazenada em um vetor, sendo, assim, uma representação bastante compacta.

Além disso, é trivial obter as coordenadas cartesianas (x, y) dos pontos das bordas do rim a partir do vetor de distâncias D . Basta utilizar a forma de conversão de coordenadas polares para cartesianas, sendo θ_i o ângulo implícito pela posição i da distância no vetor e ρ_i a própria distância d_i :

$$x_i = \cos(\theta_i)\rho_i, \quad y_i = \text{sen}(\theta_i)\rho_i \quad (3.5)$$

A Figura 3.11 mostra três exemplos de perfis centroidais, juntamente com as imagens de bordas que foram utilizadas para obtê-los. As Figuras 3.11(a), 3.11(b) e 3.11(c) mostram as imagens originais e abaixo de cada uma delas está a seqüência de distâncias. Os pontos marcam as distâncias e as linhas delineiam o perfil.

3.3.2.2 Curvatura

As curvaturas das bordas dos rins são extraídas a partir dos princípios detalhados na Seção 2.3.2.2.2. A partir da segmentação, extrai-se os pontos das bordas em sentido anti-horário partindo do ponto que forma um ângulo de 0° com o centróide, como mostra a Figura 3.12. Esses pontos são usados como entrada para o cálculo da curvatura. Entretanto, como a quantidade de pontos de borda varia a cada caso, eles são reamostrados para 200 pontos. Com a quantidade de pontos uniforme, as comparações são facilitadas. Assim, a curvatura das bordas do rim é dada por

$$k(u, \sigma) = \frac{X_u(u, \sigma)Y_{uu}(u, \sigma) - X_{uu}(u, \sigma)Y_u(u, \sigma)}{(X_u(u, \sigma)^2 + Y_u(u, \sigma)^2)^{3/2}} \quad (3.6)$$

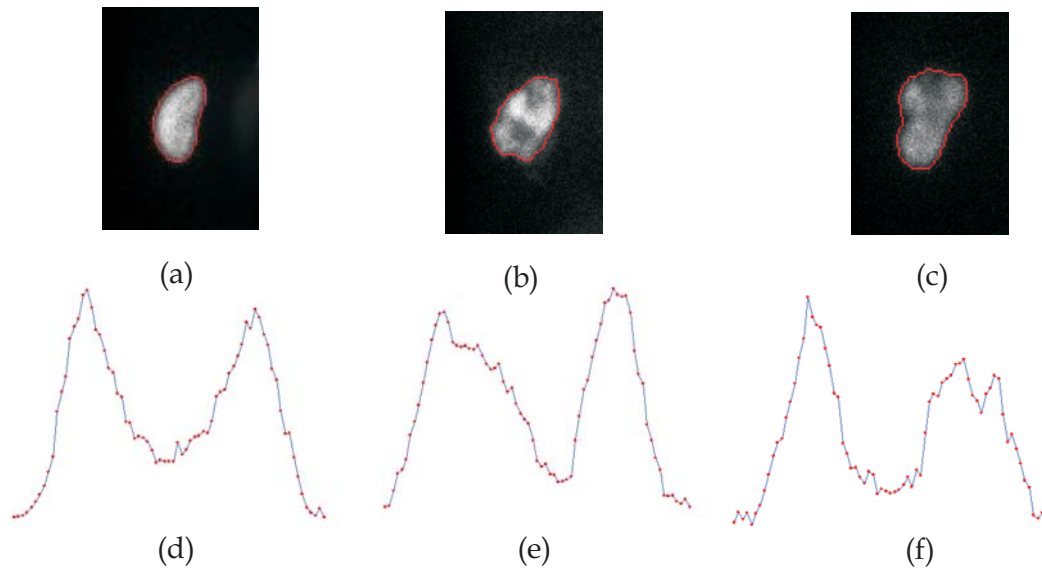


Figura 3.11: Extração dos perfis de distâncias centroidais: as figuras (a), (b) e (c) mostram as imagens originais seguidas pelos seus respectivos perfis em (d), (e) e (f). Os pontos marcam as distâncias e as linhas delineiam o perfil.

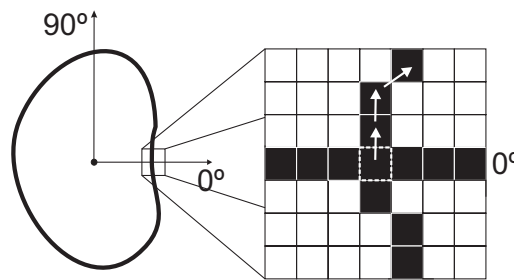


Figura 3.12: Os pontos da borda são extraídos ordenadamente a partir do ponto que forma um ângulo de 0° com o centróide. O ponto central, marcado por bordas pontilhadas, é o primeiro da lista de pontos, seguido pelos demais na ordem apontada pelas setas.

A medida que aumenta o valor de σ , a forma da curva convoluída muda. Dessa maneira, é preciso encontrar uma relação de equilíbrio entre uma forma suavizada, com menos detalhes e oscilações, mas que ainda preserve suas concavidades e convexidades. Analisando os casos existentes, testes experimentais indicaram que o valor que suaviza a forma, mas mantém os contornos é $\sigma = 20$. A Figura 3.13 mostra a mesma forma de um rim sendo convoluída por máscaras Gaussianas com diferentes valores de σ . Abaixo de cada forma está a curvatura correspondente, exceto a forma original, que apresenta o seu perfil centroidal. Analisando essas figuras, pode-se concluir que o uso de um valor de σ menor que 10 faz com que a curvatura, e também a forma, tenha muitas oscilações pequenas. Com $\sigma=10$, a forma está bastante suavizada, porém a curvatura ainda apresenta variações que não correspondem diretamente à concavidades ou convexidades na forma. Já o valor de $\sigma = 20$ compensa essas variações na curvatura sem deformar a forma. Para $\sigma=30$, a forma perde concavidades importantes, ficando suavizada demais. Dessa maneira, o valor escolhido para o cálculo das curvaturas de todos os casos foi $\sigma = 20$.

A diferença entre essa representação e a representação por curvaturas no espaço-escala (CSS) está no fato de se usar somente uma escala, ou seja, um desvio padrão σ . Utilizar mais de uma escala implicaria em fazer comparações ou outras computações em mais de

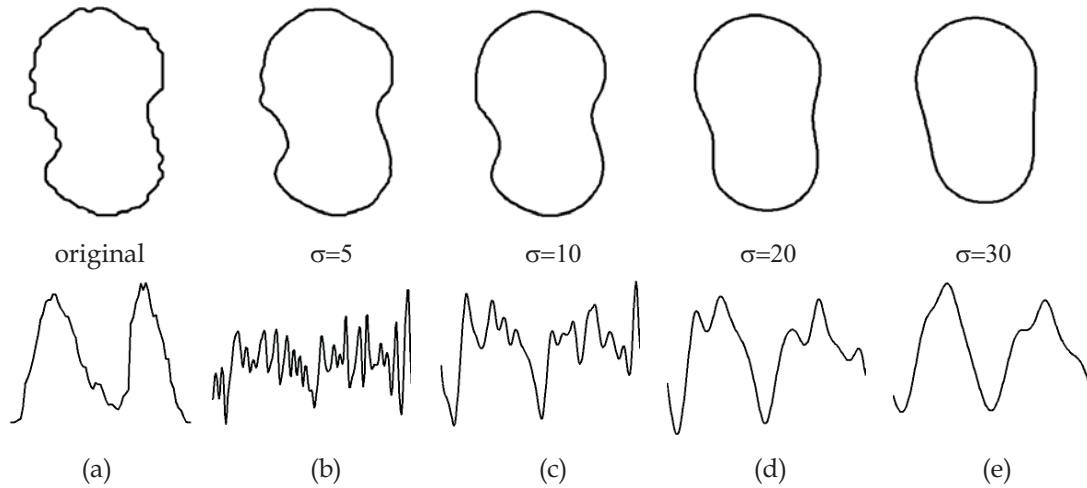


Figura 3.13: Comportamento da curvatura frente ao aumento do σ : (a) forma original e seu perfil centroidal; (b) forma suavizada com $\sigma = 5$ e curvatura correspondente; (c) $\sigma = 10$; (d) $\sigma = 20$; e (e) $\sigma = 30$.

uma escala, o que seria uma tarefa mais custosa. Utilizando o mesmo desvio padrão para todas as formas, tem-se uma relação direta entre elas. Além disso, o fato da curvatura ser calculada tendo o primeiro ponto sempre na mesma posição em todas as formas evita que comparações circulares sejam necessárias, como no caso de representações CSS comuns.

Encontrar a curvatura da forma do rim é útil porque as mudanças na curvatura correspondem a concavidades e convexidades nas bordas dos rins.

3.3.3 Feições de alto nível

Com base nas feições apresentadas anteriormente, pode-se definir as feições que representam o conteúdo de uma imagem de cintilografia renal em alto nível de abstração. Esses conceitos descrevem impressões diagnósticas que os especialistas observam na avaliação de um exame.

3.3.3.1 Simetria entre o rim esquerdo e o direito

De acordo com os especialistas, os rins esquerdo e direito são considerados simétricos se existe uma correspondência direta entre as formas e os tamanhos dos dois. Assim, utilizar os perfis de distâncias centroidais é adequado, visto que eles contêm informações tanto da forma, quanto do tamanho dos rins.

A simetria entre os rins é computada com base na diferença absoluta dif entre as distâncias dos perfis centroidais d_E e d_D dos rins esquerdo e direito, respectivamente

$$dif_{ED} = |d_E - d_D| \quad (3.7)$$

A medida final de simetria é estimada por

$$sim_{ED} = \frac{\sum(dif_{ED})}{\max(dif_{ED})}. \quad (3.8)$$

A Figura 3.14 ilustra o motivo de a medida anterior refletir o conceito de simetria. A Figura 3.14(a) mostra um caso em que os dois rins são simétricos. Suas formas e tamanhos são muito parecidos, considerando os perfis de distâncias. Já a Figura 3.14(b),

mostra um caso no qual os rins são assimétricos, fato evidenciado pela região central dos perfis de ambos os rins.

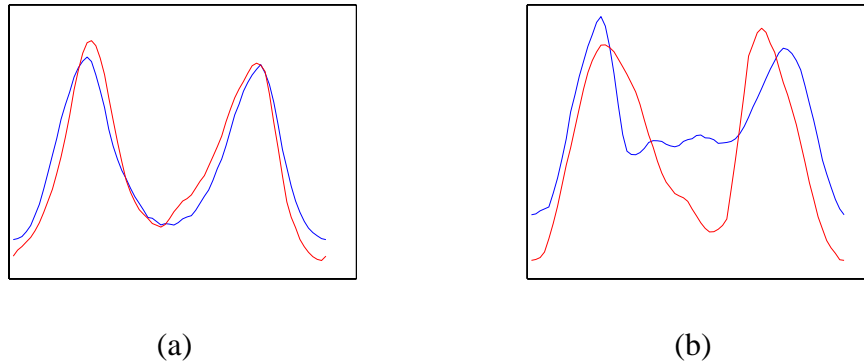


Figura 3.14: Comparação de perfis de distâncias centroidais como medida de simetria: (a) caso simétrico, pequenas diferenças entre as duas curvas; e (b) caso assimétrico, diferenças consideráveis entre as duas curvas.

3.3.3.2 Forma de cada um dos rins

A forma considerada pelos especialistas como usual para um rim lembra o formato de um grão de feijão. Caso a forma seja diferente desse padrão, ela é considerada não-usual. Com base nessa informação, uma alternativa viável para determinar como se comporta a forma de um rim é compará-la com um modelo de forma usual.

Para estimar um modelo de forma usual, a PCA foi utilizada devido à sua propriedade de encontrar padrões em um conjunto de dados (Capítulo 2.4). Para isso, essa transformação foi aplicada a um conjunto de casos considerados como tendo forma usual pelos especialistas, gerando um modelo que foi usado para comparações.

Foram selecionados 20 casos de rins esquerdos e 20 de direitos considerados como tendo forma usual pelos especialistas. Cada um desses casos teve o seu perfil de distâncias centroidais $D_i = (di_1, di_2, \dots, di_{72})$ calculado e normalizado para valores entre 0 e 1. Então, eles foram organizados em uma matriz da forma

$$C_{dados} = \begin{pmatrix} d1_1 & d1_2 & \dots & d1_{72} \\ d2_1 & d2_2 & \dots & d2_{72} \\ \vdots & \vdots & \ddots & \vdots \\ d20_1 & d20_2 & \dots & d20_{72} \end{pmatrix} \quad (3.9)$$

A matriz de covariância cov de C_{dados} foi calculada. A matriz cov contém as covariâncias entre cada par de dimensões, tendo, portanto, tamanho igual a 20×20 . Depois, encontra-se os autovetores e autovalores de cov . O autovetor que tiver o maior autovalor associado corresponde à primeira componente.

Como a PCA ordena as componentes principais de acordo a significância, a primeira componente concentra grande parte da informação do conjunto de dados, fornecendo uma aproximação razoável da forma considerada normal. O fato de a primeira componente concentrar 96,93% da variância total do conjunto de dados do rim esquerdo e 96,59% do rim direito comprova essa hipótese.

Os coeficientes da primeira componente geram, então, o perfil de distâncias centroidais do modelo de um rim que possui forma usual. Para converter o perfil em coordenadas

cartesianas, é adequado usar a conversão apresentada na Equação 3.5. Essa conversão foi utilizada na Figura 3.15, que mostra os modelos de forma para os rins esquerdo e direito. As linhas finas representam as bordas dos rins que foram usados para construir o modelo. Já a linha mais grossa marca o modelo de forma obtido.

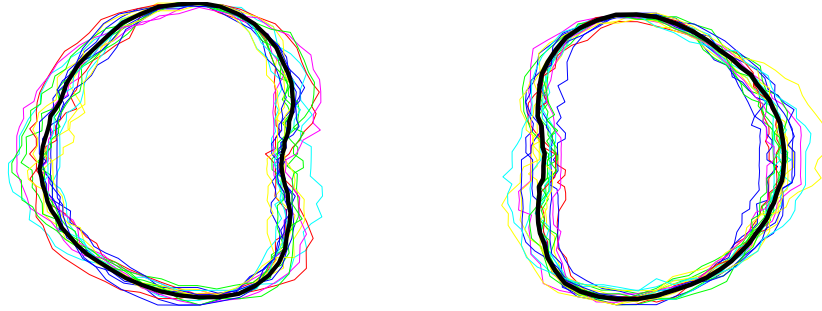


Figura 3.15: Modelo de forma usual de rins obtido por meio da PCA. As linhas finas representam as bordas dos rins que foram usados para construir o modelo e a linha mais grossa representa o modelo de forma obtido: (a) modelo de forma para rins esquerdos; e (b) modelo de forma para rins direitos.

De posse desse modelo, a forma de um rim qualquer S_a pode ser comparada com o modelo de forma S_m . Para tornar a medida invariante à escala, os tamanhos de ambas as formas S_a e S_m são normalizados para 1. Depois disso, a primeira derivada da curvatura de ambas as formas é calculada, ou seja

$$k(u, \sigma)' = \partial(k(u, \sigma))/\partial(u) \quad (3.10)$$

A curva $k(u, \sigma)'$ evidencia onde ocorrem as maiores variações nas duas formas, porque, quanto maior a inclinação, maior o coeficiente de variação da curva (derivada). Depois é calculada a soma das diferenças absolutas entre as derivadas encontradas, obtendo-se a medida de similaridade de forma do rim sh_a :

$$sh_a = \sum |(k_a(u, \sigma))' - (k_m(u, \sigma))'| \quad (3.11)$$

3.3.3.3 Posição relativa entre os dois rins

A posição relativa visa medir o relacionamento entre a localização do rim esquerdo e do direito. Comumente, o rim direito está localizado na mesma altura ou um pouco abaixo do rim esquerdo (por causa do lóbulo direito do fígado). A posição relativa entre eles permite investigar se os rins são tópicos (estão na posição esperada) ou ectópicos (um deles está localizado muito abaixo ou muito acima da posição considerada normal).

Para calcular essa medida, usa-se a localização dos pólos superiores dos dois rins. Então, o ângulo q formado pelo vetor horizontal tangente ao pólo superior do rim esquerdo u e pelo vetor que passa pelos dois pólos (conforme mostra a Figura 3.16) é calculado:

$$\arccos(q) = \frac{u \cdot v}{|u| \cdot |v|} \quad (3.12)$$

No caso de o rim esquerdo estar acima do direito, quanto maior o ângulo q , maior a tendência de ser ectópico. No caso do rim esquerdo estar abaixo do direito, mesmo que o

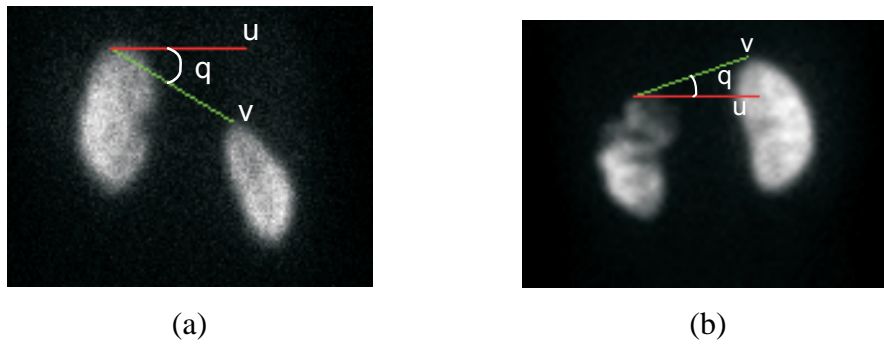


Figura 3.16: Posição relativa entre os rins, dada pelo ângulo formado pelos vetores tangentes aos pólos superiores: (a) caso em que o rim direito está abaixo do esquerdo; e (b) caso anormal, no qual o rim direito está acima do esquerdo.

ângulo de inclinação q seja pequeno, a probabilidade de ser considerado ectópico é maior, porque ele está fora do padrão usual. Então, existe uma penalidade a ser considerada nesse caso.

3.3.3.4 Captação do radiofármaco em cada um dos rins

A determinação da captação do radiofármaco nos rins baseada nas contagens é uma medida bastante comum na avaliação dos exames de cintilografia renal. Normalmente, o especialista seleciona uma região de interesse em um programa de visualização e, com base nas intensidades dos níveis de cinza, um programa de análise calcula a quantidade de contagens naquela região. Essa medida é utilizada inclusive no cálculo da função renal, juntamente com variáveis como idade e sexo.

Além dessa medida, o aspecto de como ocorreu a captação ao longo do rim é uma importante impressão diagnóstica (PIEPSZ; HAM, 2006). Rins normais, ou homogêneos, têm pequenas áreas de hipocaptação (regiões escuras) no interior. Entretanto, se o rim tiver regiões escuras próximas às bordas, fazendo com que as mesmas fiquem difusas, então esse rim possui captação irregular do radiofármaco, o que faz com que ele seja heterogêneo. Convém ressaltar que, nesse contexto, a homogeneidade não tem relação com o conceito de homogeneidade ligado à textura.

Assim, a quantificação da captação do radiofármaco concentra-se nas informações próximas às bordas do rim. Considerando a imagem como um relevo, e focando no esquema de segmentação desenvolvido (detalhado na Seção 3.2), o comportamento das bordas é investigado por meio da comparação das curvaturas de duas bordas. A primeira delas é a borda original O_{out} , obtida pelo limiar med e otimizada pelo EM. A segunda borda O_{in} localiza-se z níveis de cinza acima de med , como mostra a Figura 3.17 em vistas bidimensionais e tridimensionais. A borda O_{out} explora a presença de regiões escuras, indicando a presença de áreas de hipocaptação próximas à borda.

A equação abaixo sintetiza como a medida é calculada:

$$hom_a = \frac{\sum |k_{O_{out}}(u, \sigma) - k_{O_{in}}(u, \sigma)|}{\max(\sum k_{O_{out}}(u, \sigma), \sum k_{O_{in}}(u, \sigma))}. \quad (3.13)$$

Quando o rim tem captação homogênea, as duas bordas tem comportamentos similares, como mostrado na Figura 3.18(a). Caso contrário, as bordas externas apresentam distorções, o que provoca diferenças substanciais na comparação, como mostra a Figura 3.18(b).

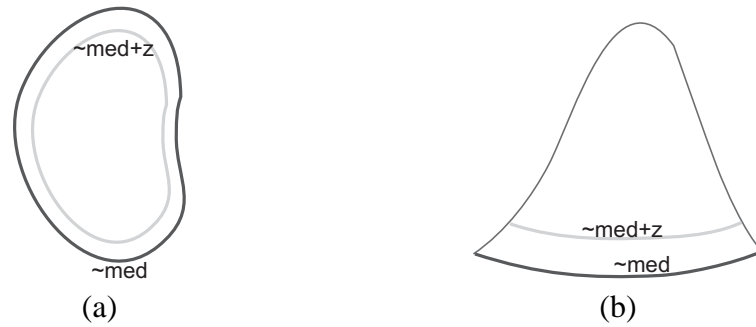


Figura 3.17: Bordas utilizadas para avaliar a captação do radiofármaco: (a) visão bidimensional, mostrando a borda interna em cinza claro e a externa em cinza escuro; e (b) visão tridimensional, mostrando os limiares utilizados.



Figura 3.18: Avaliação da captação do radiofármaco por meio das bordas internas e externas: (a) caso em que o rim é homogêneo, não apresentando diferenças significativas entre as bordas internas e externas; e (b) caso heterogêneo, no qual as bordas internas (em vermelho) apresentam deformações quando comparadas às bordas externas (em amarelo).

3.3.3.5 Comportamento dos contornos de cada um dos rins

As bordas de um rim normal são convexas na parte externa (nas “costas” do rim) e são côncavas na parte interna. Algumas vezes, a parte interna aparece plana (achatada) devido à rotação do órgão, que se volta mais para dentro da cavidade abdominal.

Existe uma forte relação entre a presença de concavidades e a homogeneidade da captação dos rins. Com a metodologia utilizada para avaliar a captação do radiofármaco, analisa-se fundamentalmente a região das bordas do rim. Se existem reentrâncias na borda, esse rim é considerado heterogêneo, pois há grandes diferenças entre $k_{O_{out}}(u, \sigma)$ e $k_{O_{in}}(u, \sigma)$. Dessa forma, pode-se assumir que rins homogêneos não possuem reentrâncias em suas bordas, apresentando comportamento normal. Entretanto, rins heterogêneos podem ou não apresentar reentrâncias. Então, é necessário verificar se as bordas desse tipo de rim apresentam concavidades. Isso é feito pela comparação da sua curvatura com a curvatura do rim modelo, pois o modelo fornece a localização de onde não pode haver concavidades. A Equação 3.13 é usada para verificar se o rim pode ser considerado homogêneo. Se ele não for homogêneo, a mesma equação é utilizada, só que a comparação é feita com a curvatura do modelo.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Nesse Capítulo são apresentados os resultados da implementação dos métodos propostos no Capítulo 3.

As implementações foram feitas em MATLAB. O conjunto de testes é composto por 58 imagens de cintilografia renal, perfazendo um total de 108 rins, pois 8 casos apresentam nefrectomia (ausência de um rim). Esse conjunto contém tanto casos normais como anormais, em proporções desiguais.

Para cada imagem, e para cada feição que representa uma impressão diagnóstica (feições de alto nível), os especialistas foram solicitados a classificar o conteúdo das imagens em duas classes: normal ou anormal. Essa forma de classificação foi escolhida porque dessa maneira não existe a possibilidade do especialista tender à indecisão. No caso de existirem categorias de classificação intermediárias (por exemplo, três classes: normal, quase normal e anormal), a tendência é que os casos duvidosos, quando não se tem certeza quanto ao diagnóstico, sejam classificados na classe intermediária. A maneira de classificação adotada faz com que o especialista escolha qual é a maior tendência do caso em questão. Se o caso tender mais para a normalidade, será classificado como normal; caso contrário, será anormal. Se essa abordagem não fosse utilizada, as feições computacionais propostas seriam prejudicadas quando comparadas às avaliações dos especialistas. Isso porque o sistema procura calcular um valor real que corresponde a uma escala de normalidade: quanto mais baixo o valor, maior a anormalidade e quanto mais alto, maior a normalidade. Por exemplo, considerando que a feição de forma tem seus valores de avaliação contidos no intervalo entre 0 e 1, se um caso recebe valor 0.8, ele pode ser considerado “mais normal” que um caso que tem valor 0.3.

Quatro avaliações diferentes foram realizadas pelo mesmo especialista. Entretanto, a fim de se levar em conta a variabilidade interobservador, no caso de haver divergência entre as diferentes avaliações, o caso em questão não foi mais utilizado nas comparações. Essa estratégia foi adotada para evitar que o sistema fosse penalizado no caso de haver divergência até mesmo entre os especialistas.

4.1 Comparação dos resultados com a avaliação de especialistas

Para avaliar se as feições computadas refletem a percepção dos especialistas, a classificação feita pelos profissionais foi comparada com os valores obtidos pelos métodos computacionais. Para isso, os resultados das medidas computacionais foram normalizados para valores entre 0 e 1. Entretanto, as medidas feitas pelos especialistas possuem somente dois valores possíveis: 0 (anormal) e 1 (normal). Então, a fim de tornar a comparação possível, o que foi analisado foi a tendência de cada feição, ou seja, se o valor tende mais para um caso normal ou anormal. Os valores das medidas foram representados em

um gráfico que relaciona o número do caso (de 1 a 58) ao valor da feição em questão (com valores de 0 a 1). A partir desse gráfico, pode-se separar visualmente as medidas das feições em duas classes (normal e anormal) analisando somente a distribuição espacial dos valores numéricos.

Essa maneira de classificação dispensa o uso de um classificador que potencialmente introduziria erros adicionais de classificação. Além disso, essa abordagem foi utilizada somente para permitir a comparação das feições com a avaliação de especialistas. Forçar o sistema a produzir um valor binário de classificação (normal ou anormal) seria um retrocesso em relação ao objetivo do trabalho, que é fornecer uma avaliação quantitativa que permite reduzir a subjetividade e não tem pretensão de prover um diagnóstico.

4.2 Discussão dos resultados obtidos

As seções a seguir relatam os resultados obtidos, apresentando as distribuições dos valores obtidos, as taxas de concordância com a avaliação do especialista e relatando casos de sucesso e insucesso na avaliação.

4.2.1 Captação do rádiofarmaco

A Figura 4.7 mostra o gráfico da distribuição dos valores da feição de homogeneidade da captação do rádiofarmaco. A Figura 4.7(a) refere-se aos rins esquerdos e a Figura 4.7(b) mostra a classificação dos rins direitos. Nesses gráficos, os valores das feições (eixo y) foram representados em relação ao número do caso (no eixo x). Os pontos em verde, significam que a avaliação dos especialistas classificou esse caso como normal; em vermelho, estão os casos classificados como anormais. Dessa forma, pode-se visualizar quando existe um caso classificado como normal em uma região de valores de casos anormais. A linha em azul marca o que seria o limiar de separação das classes, determinado visualmente. Pelo gráfico da Figura 4.7(a), pode-se visualizar que dois casos normais foram classificados erroneamente como anormais e um caso anormal foi erroneamente classificado como normal. Já para os rins do lado direito, existe somente um caso normal classificado de maneira equivocada e três anormais. Dessa forma, de um total de 102 casos, 95 foram classificados corretamente, gerando uma taxa de concordância entre as feições computacionais e o especialista de 0.93.

A Figura 4.1 mostra dois casos, sendo o da direita um caso que foi corretamente avaliado como tendo captação heterogênea do rádiofarmaco e o caso da esquerda mostra uma avaliação incorreta, de um caso heterogêneo avaliado como tendo captação homogênea.

4.2.2 Comportamento dos contornos

A Figura 4.8 mostra o gráfico da distribuição dos valores da feição do comportamento das bordas dos rins. A Figura 4.8(a) refere-se ao rins do lado esquerdo e a Figura 4.8(b) mostra a classificação para o lado direito. Esses gráficos seguem a mesma simbologia dos gráficos da Figura 4.7. Dessa maneira, pelo gráfico da Figura 4.8(a), pode-se visualizar que somente um caso normal foi classificado erroneamente como anormal e dois anormais foram classificados como normais. O gráfico da Figura 4.8(b) mostra que somente um caso anormal foi classificado erroneamente como normal. Assim, de um total de 103 casos, 99 foram classificados corretamente, gerando uma taxa de concordância entre as feições computacionais e o especialista de 0.96.

A Figura 4.2 mostra exemplos da avaliação. Os três primeiros casos foram erroneamente classificados. Os dois primeiros representam casos que contém concavidades inde-



Figura 4.1: Exemplos da avaliação de casos quanto à homogeneidade: (a) caso heterogêneo avaliado corretamente; e (b) caso homogêneo avaliado incorretamente como heterogêneo.

vidas, mas foram classificados como normais. O caso da Figura 4.2(c) mostra a avaliação contrária: um caso normal que foi classificado como contendo concavidades. O último caso foi corretamente avaliado como contendo concavidades indevidas.

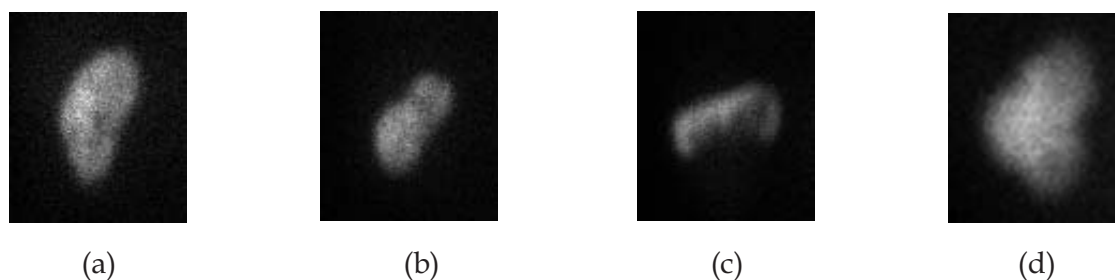


Figura 4.2: Exemplos da avaliação de casos quanto ao comportamento das bordas: (a) e (b) casos com concavidades avaliados incorretamente como normais; (c) caso normal avaliado incorretamente como anormal; e (d) caso corretamente classificado como tendo comportamento anormal das bordas.

4.2.3 Forma

Para a feição de forma, a Figura 4.9 mostra a distribuição das medidas. A Figura 4.9(a) mostra o resultado para o lado esquerdo, no qual um caso normal é classificado erroneamente como anormal. Já para os rins do lado direito, o gráfico da Figura 4.9(b) mostra que um caso anormal é classificado erroneamente como normal. Dessa forma, a taxa de concordância entre o especialista e as feições computacionais é de 0.98, visto que de 101 casos, 99 foram avaliados corretamente.

A Figura 4.3 mostra exemplos da avaliação. As duas primeiras figuras mostram os casos erroneamente avaliados. Os dois são casos normais, porém foram avaliados como anormais. O último caso, da Figura 4.3(c), foi corretamente classificado. Os dois primeiros casos foram possivelmente avaliados de forma incorreta devido às regiões escurecidas, que podem ter influenciado na classificação. O especialista consegue abstrair essas regiões e analisar a forma independentemente dessas alterações. O sistema, entretanto, não consegue realizar esse tipo de compensação.

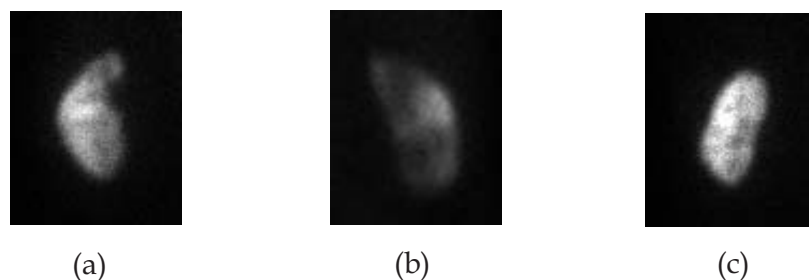


Figura 4.3: Exemplos da avaliação de casos quanto à forma: (a) e (b) casos erroneamente classificados como tendo forma não-usual; e (c) caso corretamente avaliado como tendo forma usual.

4.2.4 Simetria

A feição de simetria avalia a relação entre os dois rins, portanto não existe distinção entre valores para rins esquerdos ou direitos. A Figura 4.10 mostra a distribuição de valores para essa feição. Pode-se notar que quatro casos simétricos são incorretamente avaliados como assimétricos. Nenhum caso assimétrico foi avaliado incorretamente. Outras separações entre as duas classes são possíveis, se a linha azul for deslocada mais acima no gráfico. No entanto, a linha nessa posição maximiza o acerto para os casos assimétricos, o que pode ser vantajoso tendo em vista que, para o paciente, é melhor um diagnóstico falso positivo do que falso negativo. Assim, a taxa de concordância é de 0.90, visto que dos 42 casos disponíveis, 38 foram corretamente classificados.

A Figura 4.4 mostra os casos que foram incorretamente avaliados. O caso da Figura 4.4(a) evidencia que a diferença de tamanho entre os rins fez a medida tender para a assimetria. Os casos das Figuras 4.4(a) e (b) apresentam irregularidades que devem ter influenciado a medida computacional. Já o caso da Figura 4.4(d) não apresenta diferenças visuais significativas, porém foi classificado incorretamente.

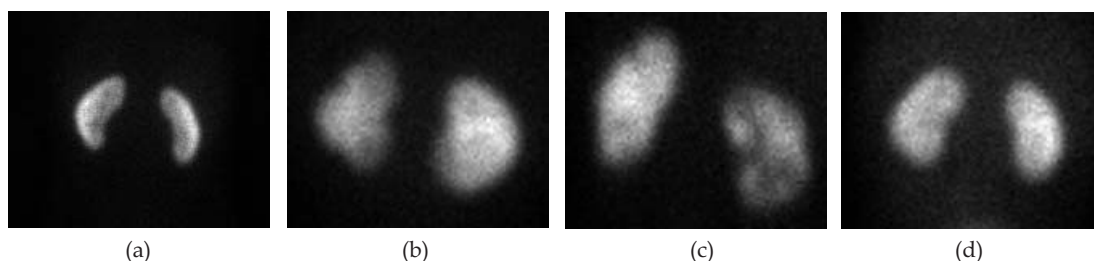


Figura 4.4: Exemplos da avaliação de casos quanto à simetria dos rins: casos avaliados incorretamente como assimétricos.

4.2.5 Posição relativa

A distribuição dos valores para a feição que mede a posição relativa entre os rins pode ser vista na Figura 4.11(a). Esse gráfico foi representado com os valores reais dos ângulos que existem entre o pólo superior do rim esquerdo e do direito, ou seja, os valores da feição não foram normalizados entre 0 e 1. Isso foi feito para mostrar as duas regiões onde os rins são considerados ectópicos: quando possuem ângulos de inclinação elevados

(aproximadamente acima de 40°) e ângulos negativos. A Figura 4.11(b) mostra o detalhe ampliado da região marcada pelo retângulo pontilhado na Figura 4.11(a), evidenciando que existe uma diferença considerável entre os valores, o que permite separá-los em duas classes. Dessa forma, somente um caso é classificado incorretamente, mostrado na Figura 4.5(a). A Figura 4.5(b) mostra um caso corretamente classificado como ectópico e a Figura 4.5(c) mostra um caso corretamente avaliado como tópico. Assim, dos 50 casos analisados, 49 foram classificados corretamente, gerando uma taxa de concordância de 0.98.

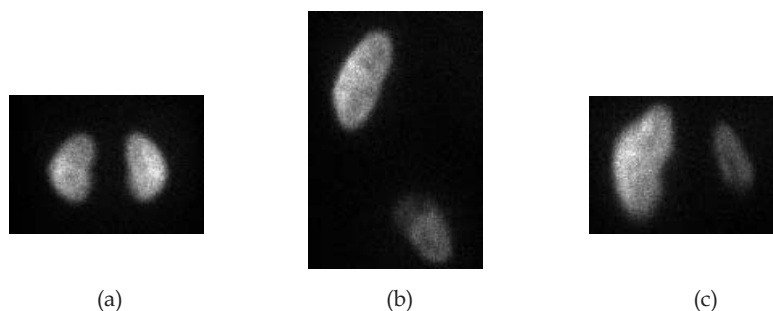


Figura 4.5: Exemplos da avaliação de casos quanto à posição relativa entre os rins: (a) caso avaliado incorretamente como ectópico; (b) caso avaliado corretamente como ectópico; e (c) caso avaliado corretamente como tópico.

4.3 Resumo dos resultados

A Tabela 4.1 sumariza todos os resultados obtidos. A coluna que indica o número total de casos contém somente aqueles que não têm divergência na avaliação dos especialistas. O número de casos corretamente classificados refere-se aos casos em que as feições computacionais e os especialistas concordam na avaliação. E a taxa de concordância é dada pela razão entre o número de casos corretamente classificados e o número total de casos.

Como pode ser visto na tabela, todas as medidas calculadas concordam em pelo menos 90% dos casos com a avaliação dos especialistas, indicando um grau de concordância considerável. Dessa maneira, pode-se dizer que as medidas possuem grande probabilidade de refletirem a opinião do especialista, fornecendo, assim, uma medida quantitativa para orientar a avaliação.

4.4 Aplicação das feições para representar estudos

As feições extraídas das imagens podem ser usadas para representar estudos. Dessa forma, a avaliação de casos pode ser descrita por informações quantitativas que descrevem as características de cada caso. A Figura 4.6 mostra três exemplos de casos acompanhados de suas respectivas descrições. Para cada caso, existe o valor calculado para cada feição, a classificação pelo sistema e uma possível correção do especialista. Esse tipo de abordagem pode ser utilizado para busca, recuperação e comparação de casos utilizando métodos vetoriais ou Booleanos. Além disso, essas informações podem ser adicionadas ao Prontuário Eletrônico do Paciente.

Tabela 4.1: Resultados da classificação.

Feição		Número total de casos	Número de casos corretamente classificados	Taxa de concordância
Captação DMSA	Normal	69	66	0.96
	Anormal	33	29	0.88
	<i>Total</i>	102	95	0.93
Comportamento das Bordas	Normal	79	78	0.99
	Anormal	24	21	0.87
	<i>Total</i>	103	99	0.96
Forma	Normal	91	90	0.99
	Anormal	10	9	0.90
	<i>Total</i>	101	99	0.98
Simetria	Normal	15	11	0.73
	Anormal	27	27	1.00
	<i>Total</i>	42	38	0.90
Posição Relativa	Normal	42	42	1.00
	Anormal	8	7	0.87
	<i>Total</i>	50	49	0.98

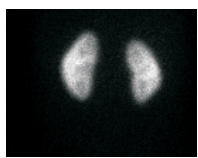
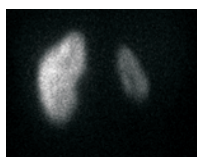
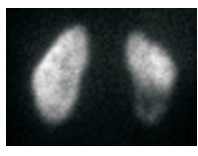
	Feição	Simetria	Posição relativa	Forma		Contornos		Captação	
				Esq	Dir	Esq	Dir	Esq	Dir
	Feição	0,97	12°	0,78	0,67	0,68	0,79	0,85	0,87
	Classificação	1	1	1	1	1	1	1	1
	Especialista	1	1	1	1	1	1	1	1
	Feição	0,61	6°	0,64	0,21	0,83	0,07	0,72	0,57
	Classificação	0	1	1	1	1	0	1	0
	Especialista	0	1	1	1	1	0	1	0
	Feição	0,75	6°	0,72	0,24	0,85	0,19	0,87	0,30
	Classificação	0	1	1	1	1	0	1	0
	Especialista	0	1	1	1	1	0	1	0

Figura 4.6: Quadro que mostra como as feições podem ser utilizadas para descrever os casos de cintilografia renal.

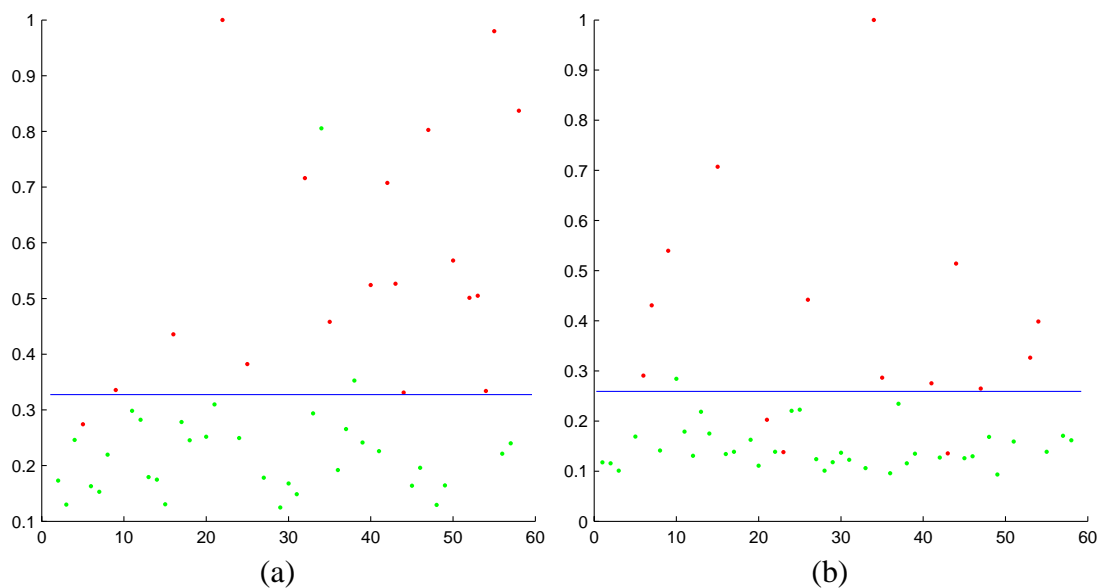


Figura 4.7: Gráfico do resultado da classificação da feição de homogeneidade da captação do radiofármaco: (a) rim esquerdo; e (b) rim direito. Em verde estão os casos classificados pelo especialista como tendo captação do radiofármaco normal. Em vermelho, os que tem captação anormal. A reta em azul indica o limiar de separação das duas classes.

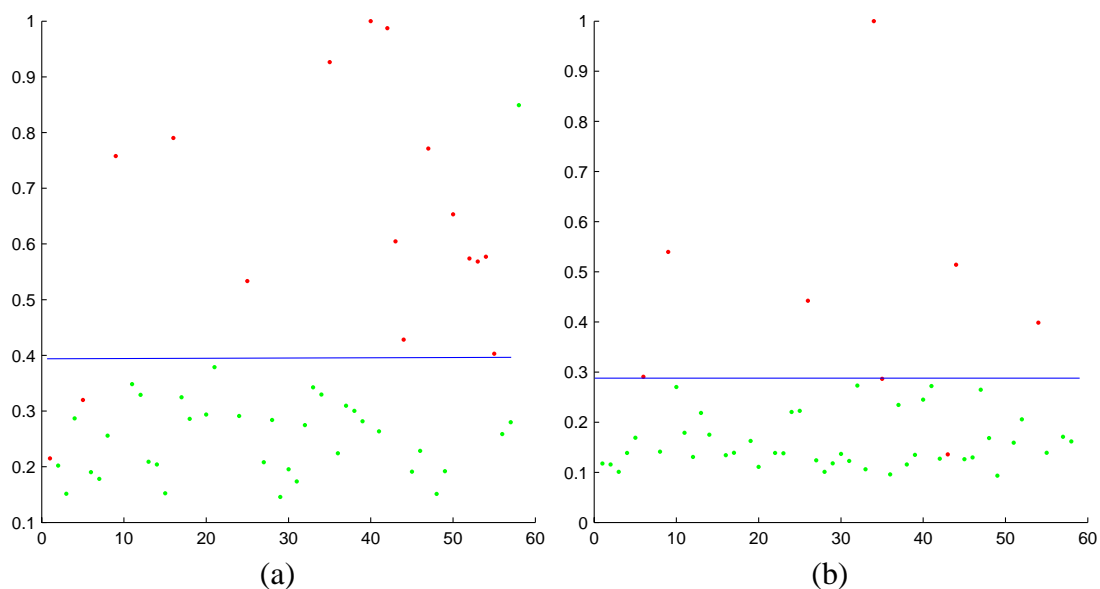


Figura 4.8: Gráfico do resultado da classificação da feição de comportamento das bordas do rim: (a) rim esquerdo; e (b) rim direito. Em verde estão os casos classificados pelo especialista como tendo comportamento das bordas normal. Em vermelho, os que tem comportamento anormal. A reta em azul indica o limiar de separação das duas classes.

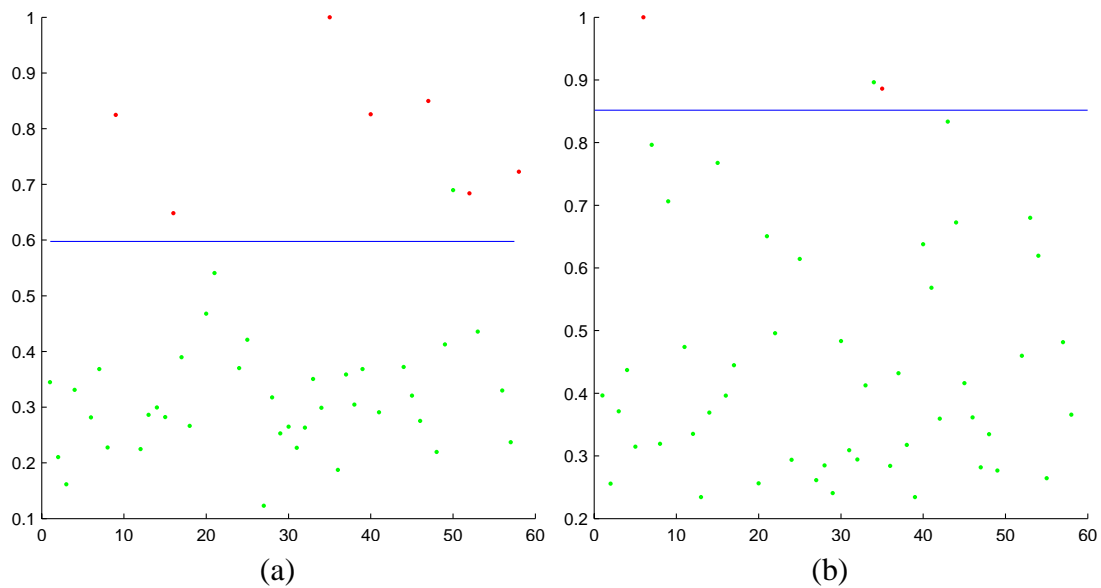


Figura 4.9: Gráfico do resultado da classificação da feição de forma dos rins: (a) rim esquerdo; e (b) rim direito. Em verde estão os casos classificados pelo especialista como tendo forma usual. Em vermelho, os que tem forma não-usual. A reta em azul indica o limiar de separação das duas classes.

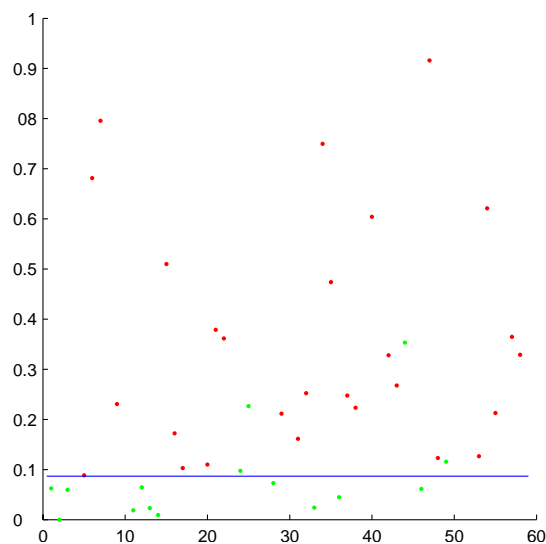


Figura 4.10: Gráfico do resultado da classificação da feição de simetria entre os rins. Em verde estão os casos classificados pelo especialista como simétricos. Em vermelho, os assimétricos. A reta em azul indica o limiar de separação das duas classes.

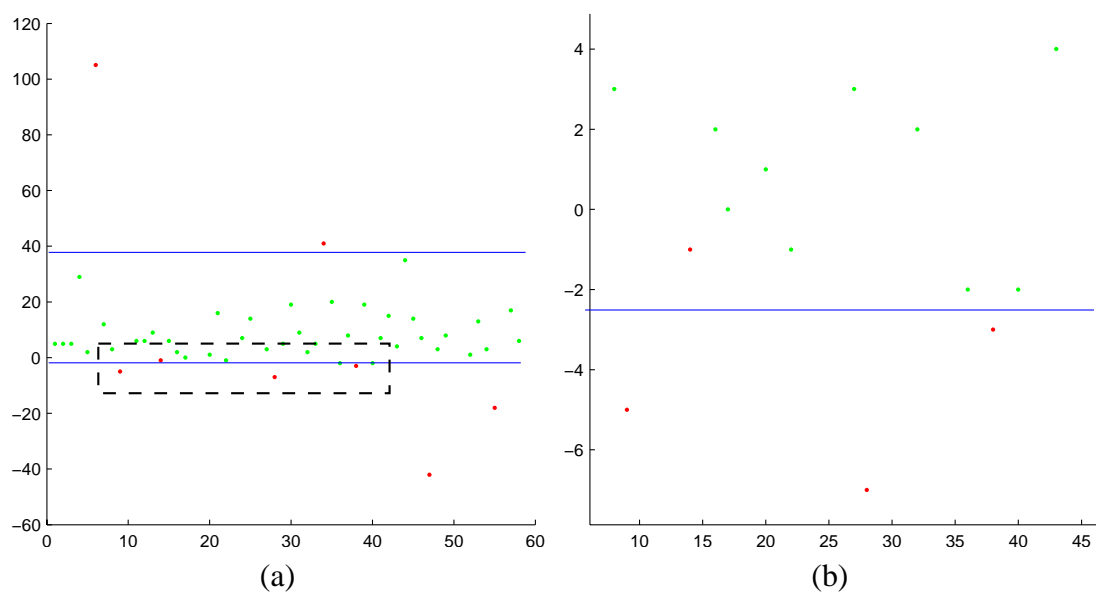


Figura 4.11: (a) Gráfico do resultado da classificação da feição de posição relativa entre os rins. Em verde estão os casos classificados pelo especialista como tópicos. Em vermelho, os ectópicos. A reta em azul indica o limiar de separação das duas classes. (b) Detalhe ampliado da região marcada pelo retângulo pontilhado na figura (a).

5 CONSIDERAÇÕES FINAIS

Este trabalho apresenta um conjunto de medidas quantitativas que refletem impressões diagnósticas geralmente observadas por especialistas em imagens de cintilografia renal. A avaliação visual dessas imagens é permeada por questões subjetivas, que levam a altos índices de variabilidade intra e interobservador. Assim, busca-se reduzir a subjetividade e, conseqüentemente, a variabilidade por meio de medidas quantitativas extraídas das imagens.

Como as medidas estão fortemente relacionadas às impressões diagnósticas que os médicos avaliam nos exames, elas refletem o conteúdo das imagens. Ao contrário das feições de baixo nível, que pouco ou nada informam sobre a semântica, essas medidas exprimem conceitos abstratos, relacionados ao conteúdo da imagem.

Para permitir a extração das feições, um método de segmentação foi proposto. Esse método apresentou resultados satisfatórios para as imagens de cintilografia renal. Assim, pode-se imaginar que esse método pode ser estendido e aplicado a outras imagens com características semelhantes. Para a segmentação das imagens de cintilografia renal, esse método não precisou de ajuste de parâmetros, mantendo a mesma configuração para todos os casos, o que configura uma grande vantagem em relação a outros métodos. Além disso, a abordagem de considerar cada rim separadamente leva a bons resultados, visto a variabilidade existente entre os rins de um mesmo caso.

Os resultados atingidos indicam que as feições propostas concordam em, pelo menos, 90% dos casos com a avaliação visual dos especialistas. Esses resultados tornam as medidas confiáveis do ponto de vista médico. Além disso, a utilização dessas medidas pode fazer com que a reproducibilidade interobservador seja reduzida, pois elas fornecem um modo mais preciso e objetivo de reportar as impressões diagnósticas.

Outra característica importante das feições encontradas é que elas podem ser utilizadas em buscas, recuperações e comparação de casos de imagens de cintilografia renal. Essa característica abre um vasto leque de aplicações, principalmente para busca de casos relacionados e avaliação da evolução dos casos.

A maioria dos sistemas de recuperação de informações visuais baseada em conteúdo usam feições de baixo nível, enquanto que o ideal seria recuperar essas informações com base na semântica das imagens. Esta dissertação busca resolver exatamente esse problema com uma aplicação em um escopo limitado na área médica. A metodologia proposta permite melhor utilização e acesso às informações armazenadas em imagens de cintilografia renal. Nesse contexto, os métodos implementados preenchem a lacuna semântica entre as informações de baixo nível e o conteúdo das imagens de uma parte do universo de aplicações existentes. Entretanto, a investigação de como o conteúdo é associado às imagens, de como os observadores humanos interpretam essas imagens e de que maneira as informações estão interligadas pode trazer benefícios a outras áreas de aplicação.

Desse modo, as contribuições desta dissertação são as seguintes:

- Levantamento das características relevantes nas imagens de cintilografia renal, do ponto de vista de impressões diagnósticas;
- Proposição de um método de segmentação das imagens, que leva a resultados satisfatórios;
- Proposição de métodos de extração de feições que representam conceitos abstratos (conteúdo semântico) das imagens, que foram avaliados por especialistas; e
- Caráter de inovação, visto que não existem registros na literatura relacionados ao tema abordado na dissertação.

5.1 Sugestão de trabalhos futuros

Uma das aplicações que desponta nos sistemas de apoio à medicina tem como base a integração de diversas fontes de dados para facilitar o diagnóstico, avaliação de tratamentos, pesquisa, treinamento e ensino. Assim, surgem sistemas “aumentados”, que reúnem casos avaliados, suas feições, diagnósticos, materiais de apoio à decisão, tais como artigos relacionados e casos semelhantes. Nesse contexto, as feições propostas podem ser integradas a sistemas do tipo, gerando dados que permitem comparação e são mais adequados que feições numéricas de baixo nível.

Além disso, o sistema pode ter métodos para extração de outras feições de baixo e médio nível. Assim, um sistema de mineração de dados poderia varrer essa base em busca de outras associações de feições, para descobrir como elas se relacionam, refletindo com mais precisão as percepções visuais dos observadores humanos. Também se pode descobrir padrões de ocorrência de doenças e feições.

Essa mesma linha de raciocínio remete à questão de que existe uma grande quantidade de informação disponível que devem ser acessíveis. Recuperar essas informações de forma eficiente e que resulte em informações concisas e relevantes ainda é um desafio. Nesse sentido, surge o conceito de sumarização de documentos, que visa condensar o que é realmente relevante de acordo com a aplicação. As feições numéricas propostas podem servir também a esse propósito, visto que fornecem descritores das impressões diagnósticas das imagens.

As feições extraídas abrangem uma parte da avaliação das imagens de cintilografia renal. Assim, outros estudos podem ser realizados para quantificação de outros aspectos presentes nas imagens, por exemplo, a quantificação e a localização de lesões.

REFERÊNCIAS

ABBASI, S.; MOKHTARIAN, F.; KITTLER, J. Curvature Scale Space Image in Shape Similarity Retrieval. **Multimedia Syst.**, Secaucus, NJ, USA, v.7, n.6, p.467–476, 1999.

BEZDEK, J. C. **Pattern Recognition with Fuzzy Objective Function Algorithms**. Norwell, MA, USA: Kluwer Academic Publishers, 1981.

BOVIK, A. C.; GIBSON, J. D.; BOVIK, A. (Ed.). **Handbook of Image and Video Processing**. Orlando, FL, USA: Academic Press, Inc., 2000.

CANNY, J. A Computational Approach to Edge Detection. In: FISCHLER, M. A.; FIRSCHEIN, O. (Ed.). **Readings in Computer Vision: issues, problems, principles, and paradigms**. Los Altos, CA.: Kaufmann, 1987. p.184–203.

CHANG, S. The holy grail of content-based media analysis. **IEEE Multimedia**, Los Alamitos, CA, USA, v.9, n.2, p.6–10, April-June 2002.

CHEN, Y.; ZHANG, M.; LU, P.; WANG, Y. Local Moment Invariant Analysis. In: INTERNATIONAL CONFERENCE ON COMPUTER GRAPHICS, IMAGING AND VISUALIZATION, CGIV, 2005, Washington, DC, USA. **Proceedings...** [S.l.]: IEEE Computer Society, 2005. p.137–140.

COMER, M.; DELP, E. Morphological Operations for Color Image Processing. **Journal of Electronic Imaging**, [S.l.], v.8, n.3, p.279–289, 1999.

CRESPO, J.; SCHAFER, R. The flat zone approach and color images. In: SERRA, J.; SOILLE, P. (Ed.). **Mathematical morphology and its applications to image processing**. Dordrecht: Kluwer Academic Publishers, 1994. p.85–92.

DE GUEVARA, D. L. et al. Interobserver reproducibility in reporting on ^{99m}Tc-DMSA scintigraphy for detection of late renal sequelae. **Journal of Nuclear Medicine**, Reston, VA, USA, v.42, n.4, p.564–566, 2001.

DE SADELEER, C. et al. A multicenter trial on interobserver reproducibility in reporting on ^{99m}Tc-DMSA planar scintigraphy: a belgian survey. **Journal of Nuclear Medicine**, Reston, VA, USA, v.41, n.1, p.23–26, 2000.

DEMPSTER, A.; LAIRD, N.; RUBIN, D. Maximum likelihood from incomplete data via the EM algorithm. **J. Royal Statistical Society, Series B**, London, UK, v.39, n.1, p.1–38, 1977.

- DOUGHERTY, E.; LOTUFO, R. **Hands-on Morphological Image Processing**. Portland: Book News, 2004.
- DUNCAN, J. C.; AYACHE, N. Intra- and interobserver variability in interpretation of DMSA scans using a set of standardized criteria. **Journal Pediatric Radiology**, Berlin, v.23, n.7, p.506–509, 1993.
- DUNCAN, J. C.; AYACHE, N. Medical Image Analysis: progress over two decades and the challenges ahead. **IEEE Trans. Pattern Anal. Mach. Intell.**, Washington, DC, USA, v.22, n.1, p.85–106, 2000.
- EAKINS, J. P.; GRAHAM, M. E. **Content-based image retrieval**. [S.l.: s.n.], 1999. Report to JISC Technology Applications Programme.
- FACON, J. **Morfologia Matemática: teoria e exemplos**. Curitiba: Ed. Universitária Champagnat da Pontífica Universidade Católica do Paraná, 1996.
- FORSYTH, D. A.; PONCE, J. **Computer Vision: a modern approach**. Upper Saddle River, NJ: Prentice Hall, 2002.
- GLASBEY, C. A. An analysis of histogram-based thresholding algorithms. **CVGIP: Graph. Models Image Process.**, Orlando, FL, USA, v.55, n.6, p.532–537, 1993.
- GONZALEZ, R. C.; WOODS, R. E. **Processamento de Imagens Digitais**. São Paulo: E. Blücher, 2000.
- HILSON, A. J. W. Functional renal imaging with nuclear medicines. **Journal Abdominal Imaging**, New York, USA, v.8, p.176–179, 2003.
- HU, M. K. Visual Pattern Recognition by Moment Invariants. **IEEE Transactions on Information Theory**, Washington, DC, USA, v.IT-8, p.179–187, February 1962.
- JAHNE, B. **Digital image processing: concepts, algorithms and scientific applications**. 4th ed. New York: Springer-Verlag, 1997.
- JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical Pattern Recognition: a review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Washington, DC, USA, v.22, n.1, p.4–37, 2000.
- JAKSIC, E. et al. Variability in interpretation of static renal scintigraphy findings. **Vojnosanitetski pregled - Military-medical and pharmaceutical review**, Serbia and Montenegro, v.62, n.3, p.189–193, 2005.
- JOLLIFFE, I. **Principal component analysis**. 2nd ed. New York: Springer, 2002.
- LEHMANN, T. M. et al. Augmented Medical Image Management. In: AITTH CONFERENCE, 2005, Minsk, Belarus. **Proceedings...** [S.l.]: IEEE Computer Society, 2005. v.1, p.7–11.
- LEW, M. S. (Ed.). **Principles of visual information retrieval**. London, UK: Springer-Verlag, 2001.

LEW, M. S. et al. Content-based multimedia information retrieval: state of the art and challenges. **ACM Trans. Multimedia Comput. Commun. Appl.**, New York, NY, USA, v.2, n.1, p.1–19, 2006.

LIN, H.J. et al. A Study of Shape-Based Image Retrieval. In: INTERNATIONAL CONFERENCE ON DISTRIBUTED COMPUTING SYSTEMS WORKSHOPS - W7: EC ICDCSW 24., 2004, Washington, DC, USA. **Proceedings...** Los Alamitos: IEEE Computer Society, 2004. p.118–123.

MACQUEEN, J. B. Some Methods for Classification and Analysis of MultiVariate Observations. In: BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, 1967. **Proceedings...** Berkeley: University of California Press, 1967. v.1, p.281–297.

MATHERON, G. **Random Sets and Integral Geometry**. New York: John Wiley & Sons, 1975.

MOKHTARIAN, F.; MACKWORTH, A. K. A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves. **IEEE Trans. Pattern Anal. Mach. Intell.**, Washington, DC, USA, v.14, n.8, p.789–805, 1992.

MULLER, H. et al. A Review of Content-Based Image Retrieval Systems in Medical Applications - Clinical Benefits and Future Directions. **International Journal of Medical Informatics**, Ireland, n.73, p.1–23, 2004.

MUMFORD, D.; SHAHA, J. Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems. **Comm. Pure Applied Math.**, Hoboken, NJ, v.42, p.577–685, 1989.

MUNDER, S.; GAVRILA, D. An Experimental Study on Pedestrian Classification. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Washington, DC, USA, v.28, n.11, p.1863–1868, November 2006.

OLIVEIRA, P. R. et al. Techniques for Image Compression: a comparative analysis. In: BRAZILIAN SYMPOSIUM ON NEURAL NETWORKS, SBRN, 2000, Rio de Janeiro. **Proceedings...** [S.l.]: IEEE Computer Society, 2000. p.249–254.

ORTIZ, F.; TORRES, F.; JUAN, E. D.; CUENCA, N. Colour mathematical morphology for neural image analysis. **Real-Time Imaging**, London, UK, UK, v.8, n.6, p.455–465, 2002.

OTSU, N. A threshold selection method from gray-level histograms. **IEEE Transactions on Systems, Man and Cybernetics**, Washington, DC, USA, v.9, n.1, p.62–66, January 1979.

PARK, K.. Recognition and Localization of a 3D Polyhedral Object using a Neural Network. In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 1996, Minneapolis, MN. **Proceedings...** [S.l.]: IEEE Computer Society, 1996. v.4, p.3613–3618.

PIEPSZ, A.; HAM, H. R. Pediatric Applications of Renal Nuclear Medicine. **Seminars in Nuclear Medicine**, USA, v.36, n.1, p.16–35, 2006.

- ROWE, D. B.; HOFFMANN, R. G. Multivariate statistical analysis in fMRI. **IEEE Eng. Med. Biol. Mag.**, Milwaukee, USA, v.25, n.2, p.60–64, 2006.
- S. X. LIAO, M. P. On Image Analysis by Moments. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 1996, Los Alamitos, v.18, n.3, p.254-266, Mar. 1996.
- SAHOURIA, E.; ZAKHOR, A. Content Analysis of Video using Principal Componets. **IEEE Trans. on Circuits and Systems for Video Technology**, Washington, DC, USA, v.9, n.8, p.1290–1298, 1999.
- SANDLER, M. P. et al. (Ed.). **Diagnostic Nuclear Medicine**. New York, USA: Lippincott Williams & Willkinks, 1996.
- SBEH, Z.; COHEM, L.; MIMOUN, G.; COSCAS, G. A new approach of geodesic re-construction for drusen segmentation in eye fundus images. **IEEE Transactions on Medical Imaging**, Washington, DC, USA, v.20, n.12, p.1321–1333, Dec. 2001.
- SERRA, J. **Image Analysis and Mathematical Morphology**. London: Academic Press, 1982.
- SIDDIQUI, K. I.; HERO, A.; SIDDIQUI, M. Mathematical Morphology applied to Spot Segmentation and Quantification of Gene Microarray Images. In: ASILOMAR CONFERENCE ON SIGNALS AND SYSTEMS, 2002, Pacific Grove, CA. **Proceedings...** [S.l.]: IEEE Computer Society, 2002. v.1, p.926– 930.
- SMEULDERS, A. W. M. et al. Content-Based Image Retrieval at the End of the Early Years. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Washington, DC, USA, v.22, n.12, p.1349–1380, Dec. 2000.
- SMITH, L. **A tutorial on Principal Component Analysis**. [S.l.: s.n.], 2002. Disponível em: <http://www.cs.otago.ac.nz/cosc453/student_tutorialsprincipal_components.pdf>. Acesso em: abril 2007.
- SOILLE, P. **Morphological Image Analysis: principles and applications**. Berlin, Germany: Springer-Verlag, 1999.
- TAYLOR JR, A.; NALLY, J. Clinical Applications of Renal Scintigraphy. **American Journal of Radiology**, Reston, VA, n.164, p.31–41, 1995.
- THANGAVEL, K. et al. Automatic Detection of Microcalcification in Mammograms - A Review. **International Journal on Graphics Vision and Image**, [S.l.], v.5, p.31–61, 2005.
- THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition, Second Edition**. Orlando, FL, USA: Academic Press, 2003.
- VELTKAMP, R. C. Shape Matching: similarity measures and algorithms. In: INTERNATIONAL CONFERENCE ON SHAPE MODELING AND APPLICATIONS, 2001, Washington, DC, USA. **Proceedings...** [S.l.]: IEEE Computer Society, 2001. p.188.
- VINCENT, L. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. **IEEE Transactions on Image Processing**, Washington, DC, USA, v.2, n.2, p.176–201, April 1993.

VINCENT, L.; SOILLE, P. Watersheds in Digital Spaces: an efficient algorithm based on immersion simulations. **IEEE Trans. Pattern Anal. Mach. Intell.**, Washington, DC, USA, v.13, n.6, p.583–598, 1991.

ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P. J.; ROSENFELD, A. Face recognition: a literature survey. **ACM Comput. Surv.**, New York, NY, USA, v.35, n.4, p.399–458, 2003.