

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



**HQ FEATURE: DESCRITORES DE FORMA UTILIZANDO A  
CURVA DE HILBERT**

WALTER ALEXANDRE A. DE OLIVEIRA

Uberlândia - Minas Gerais

2011

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO



WALTER ALEXANDRE A. DE OLIVEIRA

## **HQ FEATURE: DESCRITORES DE FORMA UTILIZANDO A CURVA DE HILBERT**

Dissertação de Mestrado apresentada à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como parte dos requisitos exigidos para obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Banco de Dados.

Orientadora:

Prof<sup>a</sup>. Dr<sup>a</sup>. Denise Guliato

Uberlândia, Minas Gerais

2011

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da dissertação intitulada “**Hq Feature: Descritores de Forma Utilizando a Curva de Hilbert**” por **Walter Alexandre A. de Oliveira** como parte dos requisitos exigidos para a obtenção do título de **Mestre em Ciência da Computação**.

Uberlândia, 25 de Fevereiro de 2011

Orientadora:

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Denise Guliato  
Universidade Federal de Uberlândia

Banca Examinadora:

---

Prof. Dr. Ricardo Rodrigues Ciferri  
Universidade Federal de São Carlos

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Célia Aparecida Zorzo Barcelos  
Universidade Federal de Uberlândia

UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Data: Fevereiro de 2011

Autor: **Walter Alexandre A. de Oliveira**  
Título: **Hq Feature: Descritores de Forma Utilizando a Curva de Hilbert**  
Faculdade: **Faculdade de Computação**  
Grau: **Mestrado**

Fica garantido à Universidade Federal de Uberlândia o direito de circulação e impressão de cópias deste documento para propósitos exclusivamente acadêmicos, desde que o autor seja devidamente informado.

---

Autor

# Dedicatória

*Aos meus pais Walter Silvério e Tânia Maria.*

# Agradecimentos

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) e ao INCT-MACC (Instituto Nacional de Ciência e Tecnologia - Medicina Assistida por Computação Científica) pelo apoio financeiro.

Ao Dr. Rangaraj M. Rangayyan, da Universidade de Calgary, por ter gentilmente cedido a base de imagens de tumores de câncer de mama utilizada em experimentos neste trabalho.

Ao prof. Dr. Caetano Traina Jr., da Universidade de São Paulo, que muito ajudou através de sugestões e comentários.

Ao colega Glauco Pedrosa, por ter fornecido seus resultados do teste *bull's eye* utilizados para comparação com experimentos realizados neste trabalho.

Aos professores e funcionários da pós-graduação. Sempre se mostraram prestativos quando recorri a sua ajuda.

Aos colegas e amigos que tanto contribuíram com sugestões e apoio nos momentos de dificuldade ao longo de meu caminho percorrido durante o mestrado.

Por fim, porém não menos importante, à professora Dra. Denise Guliato, por seu profissionalismo, apoio, paciência, orientação e amizade durante todo o tempo.

# Resumo

A classificação e recuperação por similaridade baseadas em descritores de formas são processos de grande importância em tarefas de reconhecimento de padrões e recuperação de formas baseada em conteúdo. O desenvolvimento de descritores eficientes para caracterizar a forma ainda é um desafio. Este trabalho se concentra no desenvolvimento de novos extratores para caracterizar formas no espaço 2D e 3D usando curvas de preenchimento de espaço de Hilbert e decomposição da transformada discreta de wavelet. Diversos experimentos são realizados utilizando bases de dados de silhuetas 2D e 3D. Os descritores propostos são invariantes a transformações de translação e alguns experimentos para avaliar a robustez quanto a rotação, a escala, a oclusões, a articulações e a deformações foram realizados e obtiveram excelentes resultados.

**Palavras chave:** extração de características, descritores de forma, curva de hilbert, transformada wavelet, recuperação de imagens baseada em conteúdo, classificação de padrões, busca por similaridade, casamento de formas.

# Abstract

The similarity classification and retrieval based on shape descriptors are processes of great importance in pattern recognition and content-based shape retrieval tasks. The development of efficient shape descriptors is still a challenge. This work focuses on the development of new extractors to characterize shapes in 2D and 3D space using Hilbert space-filling curves and discrete wavelet transform decomposition. Several experiments are conducted using 2D and 3D silhouettes databases. The descriptors proposed are invariant under translation transformations and some experiments to assess the robustness on scale and rotation transformations and on occlusion, articulation and deformation effects were performed and obtained excellent results.

**Keywords:** feature extraction, shape descriptors, hilbert curve, wavelet transform, content-based image retrieval, pattern recognition, similarity searching, shape matching.

# Sumário

|                                                                  |             |
|------------------------------------------------------------------|-------------|
| <b>Lista de Figuras</b>                                          | <b>x</b>    |
| <b>Lista de Tabelas</b>                                          | <b>xiii</b> |
| <b>Lista de Algoritmos</b>                                       | <b>xiv</b>  |
| <b>1 Introdução</b>                                              | <b>15</b>   |
| 1.1 Considerações Iniciais . . . . .                             | 15          |
| 1.1.1 Classificação de Padrões . . . . .                         | 17          |
| 1.1.2 Sistemas de Recuperação de Imagem por Conteúdo . . . . .   | 17          |
| 1.2 Objetivos . . . . .                                          | 18          |
| 1.3 Organização do Trabalho . . . . .                            | 19          |
| <b>2 Fundamentação Teórica</b>                                   | <b>20</b>   |
| 2.1 Transformada Discreta de Wavelet . . . . .                   | 20          |
| 2.2 Redes Neurais Artificiais . . . . .                          | 22          |
| 2.3 Medidas de Avaliação de Desempenho . . . . .                 | 24          |
| 2.3.1 Curva ROC e Área Sobre a Curva ROC . . . . .               | 24          |
| 2.3.2 Curva Precisão-Revocação . . . . .                         | 26          |
| 2.3.3 Ganho Acumulativo Descontado . . . . .                     | 27          |
| 2.4 Curvas de Preenchimento de Espaço . . . . .                  | 28          |
| 2.5 Interpolação de <i>Slices</i> . . . . .                      | 30          |
| 2.6 Considerações Finais . . . . .                               | 32          |
| <b>3 Descritores de Forma</b>                                    | <b>33</b>   |
| 3.1 Métodos para Caracterização de Formas no espaço 2D . . . . . | 33          |
| 3.1.1 Extratores de Forma Baseados em Contorno . . . . .         | 33          |
| 3.1.2 Extratores de Forma Baseados em Região . . . . .           | 36          |
| 3.2 Métodos para Caracterização de Formas no Espaço 3D . . . . . | 38          |
| 3.3 Considerações Finais . . . . .                               | 41          |

|          |                                                                                           |           |
|----------|-------------------------------------------------------------------------------------------|-----------|
| <b>4</b> | <b><i>2Hq Feature: Um Novo Método Para Descrição de Formas</i></b>                        | <b>46</b> |
| 4.1      | Introdução . . . . .                                                                      | 46        |
| 4.2      | <i>2Hq Feature: proposta de um novo descritor de formas 2D</i> . . . . .                  | 47        |
| 4.2.1    | Redimensionamento . . . . .                                                               | 47        |
| 4.2.2    | Procedimento de decomposição . . . . .                                                    | 49        |
| 4.2.3    | Aplicação da curva de Hilbert . . . . .                                                   | 50        |
| 4.2.4    | Subamostragem . . . . .                                                                   | 51        |
| 4.2.5    | Normalização dos dados . . . . .                                                          | 52        |
| 4.3      | <i>3Hq Feature: Estendendo 2Hq para a Terceira Dimensão</i> . . . . .                     | 53        |
| 4.4      | Considerações Finais . . . . .                                                            | 55        |
| <b>5</b> | <b><i>WESS: Framework para Reconstrução de Objetos 3D</i></b>                             | <b>56</b> |
| 5.1      | Ambiente Para Segmentação de Imagens Tridimensionais Baseadas em <i>Slices</i>            | 56        |
| 5.2      | Métodos Desenvolvidos Para Interpolação de <i>Slices</i> . . . . .                        | 58        |
| 5.2.1    | Interpolação Baseada em Operações Morfológicas e Lógicas . . . . .                        | 59        |
| 5.2.2    | Interpolação Usando Transformação de Distância . . . . .                                  | 60        |
| 5.2.3    | Discussão de Resultados . . . . .                                                         | 62        |
| 5.3      | Considerações Finais . . . . .                                                            | 69        |
| <b>6</b> | <b>Experimentos de Classificação</b>                                                      | <b>70</b> |
| 6.1      | Introdução . . . . .                                                                      | 70        |
| 6.2      | Primeiro Experimento: Classificação Usando as Bases Kimia e MPEG-7 . .                    | 71        |
| 6.3      | Segundo Experimento: Rotação . . . . .                                                    | 74        |
| 6.4      | Terceiro Experimento: Classificação de Tumores de Câncer de Mama . . .                    | 77        |
| 6.5      | Quarto Experimento: Classificação de Objetos 3D Espiculados e Arredonda-<br>dos . . . . . | 79        |
| 6.6      | Considerações Finais . . . . .                                                            | 81        |
| <b>7</b> | <b>Experimentos de Recuperação de Formas Baseada em Conteúdo</b>                          | <b>82</b> |
| 7.1      | Introdução . . . . .                                                                      | 82        |
| 7.2      | Primeiro Experimento: Recuperação Usando a Base MPEG-7 . . . . .                          | 83        |
| 7.3      | Segundo Experimento: Recuperação Usando a PSB . . . . .                                   | 85        |
| 7.4      | Terceiro Experimento: Recuperação Usando a Base W3BC-111 . . . . .                        | 89        |
| 7.5      | Considerações Finais . . . . .                                                            | 90        |
| <b>8</b> | <b>Conclusões e Trabalhos Futuros</b>                                                     | <b>92</b> |
| 8.1      | Conclusões . . . . .                                                                      | 92        |
| 8.2      | Trabalhos Futuros . . . . .                                                               | 93        |
|          | <b>Referências Bibliográficas</b>                                                         | <b>95</b> |

# Lista de Figuras

|      |                                                                                                                                                                                             |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1  | a) Esquema de uma pirâmide de imagens. b) Principais componentes de um sistema de codificação e decodificação de sub-banda com duas bandas.                                                 | 21 |
| 2.2  | Decomposição wavelet em até três níveis.                                                                                                                                                    | 22 |
| 2.3  | Modelo de um neurônio artificial.                                                                                                                                                           | 23 |
| 2.4  | Exemplo de uma rede neural artificial.                                                                                                                                                      | 24 |
| 2.5  | Demonstração da curva ROC                                                                                                                                                                   | 26 |
| 2.6  | Exemplo de curva precisão-revocação.                                                                                                                                                        | 27 |
| 2.7  | Curvas de preenchimento de espaço de Hilbert.                                                                                                                                               | 29 |
| 2.8  | Método alternativo para geração da curva de Hilbert.                                                                                                                                        | 29 |
| 2.9  | Curvas tridimensionais de Hilbert para níveis 1 e 2.                                                                                                                                        | 30 |
| 2.10 | Contornos a serem interpolados.                                                                                                                                                             | 31 |
| 2.11 | a) Imagem binária $I_1$ . b) Imagem binária $I_2$ . c) Imagem $D_1$ resultante da transformação de distância sobre $I_1$ . d) Imagem $D_2$ resultante da multiplicação de $I_2$ por $D_1$ . | 32 |
| 4.1  | Exemplos de imagens de silhuetas.                                                                                                                                                           | 46 |
| 4.2  | Fluxograma de execução do <i>2Hq Feature</i> .                                                                                                                                              | 47 |
| 4.3  | Redimensionamento de uma imagem.                                                                                                                                                            | 48 |
| 4.4  | Efeito do processo de redimensionamento em objetos de tamanhos diferentes.                                                                                                                  | 49 |
| 4.5  | Processo de decomposição da imagem.                                                                                                                                                         | 50 |
| 4.6  | Aplicação da curva de Hilbert.                                                                                                                                                              | 50 |
| 4.7  | Aplicação da curva de Hilbert sobre região $R_1$ da imagem.                                                                                                                                 | 51 |
| 4.8  | Exemplo numérico da aplicação da transformada <i>wavelet</i> .                                                                                                                              | 52 |
| 4.9  | Redução da quantidade de dados através da etapa de subamostragem.                                                                                                                           | 52 |
| 4.10 | Exemplo do processo de normalização aplicado a um conjunto de valores.                                                                                                                      | 52 |
| 4.11 | Normalização dos dados obtidos até a etapa anterior.                                                                                                                                        | 53 |
| 4.12 | Etapa de redimensionamento aplicada a um objeto tridimensional.                                                                                                                             | 54 |
| 4.13 | Processo de decomposição de um objeto tridimensional.                                                                                                                                       | 54 |
| 5.1  | Janela principal do ambiente EMISS.                                                                                                                                                         | 57 |
| 5.2  | Janela de edição para segmentação de um <i>slice</i> no ambiente EMISS.                                                                                                                     | 57 |

|      |                                                                                                                                                                             |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 5.3  | Menu de opções da janela de edição do ambiente EMISS. . . . .                                                                                                               | 58 |
| 5.4  | <i>Slices</i> obtidos usando o algoritmo de interpolação SIMOL. . . . .                                                                                                     | 60 |
| 5.5  | Contornos e suas imagens de transformação de distância. . . . .                                                                                                             | 61 |
| 5.6  | Exemplo de aplicação da Equação 5.1. . . . .                                                                                                                                | 61 |
| 5.7  | Resultado da equação 5.1 aplicada às Imagens 5.5b e 5.5d. . . . .                                                                                                           | 63 |
| 5.8  | <i>Slices</i> obtidos pelo algoritmo de interpolação SIDITRANS. . . . .                                                                                                     | 63 |
| 5.9  | Conjunto de <i>Slices</i> obtidos usando o algoritmo SIMOL. . . . .                                                                                                         | 64 |
| 5.10 | Conjunto de <i>Slices</i> obtidos usando o algoritmo SIDITRANS. . . . .                                                                                                     | 64 |
| 5.11 | Volumes obtidos com os métodos propostos: (a-c) Diferentes vistas do volume obtido utilizando SIMOL. (d-f) Diferentes vistas do volume obtido utilizando SIDITRANS. . . . . | 65 |
| 5.12 | Conjunto de <i>slices</i> obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS. . . . .                                                                           | 65 |
| 5.13 | Diferença entre os <i>slices</i> originais e os interpolados obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS. . . . .                                        | 66 |
| 5.14 | a) Volume original. b) Volume obtido usando algoritmo SIMOL. c) Volume obtido usando algoritmo SIDITRANS. . . . .                                                           | 66 |
| 5.15 | Conjunto de <i>slices</i> obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS. . . . .                                                                           | 67 |
| 5.16 | Diferença entre os <i>slices</i> originais e os interpolados obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS. . . . .                                        | 67 |
| 5.17 | a) Volume original. b) Volume obtido usando algoritmo SIMOL. c) Volume obtido usando algoritmo SIDITRANS. . . . .                                                           | 68 |
| 6.1  | Exemplo do esquema de uma rede neural recebendo representações $2Hq_i$ em suas quatro entradas. . . . .                                                                     | 71 |
| 6.2  | a) Conjunto de imagens da base Kimia-99. b) Conjunto de imagens da base Kimia-216. c) Amostra do conjunto de imagens da base MPEG-7. . .                                    | 72 |
| 6.3  | Curvas ROC obtidas para as bases de dados: a) Kimia-99. b) Kimia-216. .                                                                                                     | 73 |
| 6.4  | Curvas ROC obtidas para as bases de dados MPEG-7 CE Shape 1 Part-B. .                                                                                                       | 74 |
| 6.5  | AUC obtida para cada ângulo de rotação. . . . .                                                                                                                             | 75 |
| 6.6  | Amostra dos objetos contidos na base de imagens W-720. . . . .                                                                                                              | 75 |
| 6.7  | Curva ROC obtida na classificação da base W-720. . . . .                                                                                                                    | 76 |
| 6.8  | Subconjunto da base de imagens com contornos de tumores de câncer [Alto et al. 2005]. . . . .                                                                               | 78 |
| 6.9  | Curva ROC obtida na classificação de calcificações de câncer de mama. . .                                                                                                   | 79 |
| 6.10 | a) Subconjunto dos objetos presentes na base W3BC-111 com características visuais semelhantes aos contornos em (b). b) Subconjunto de lesões de mama base BC-111. . . . .   | 80 |

|      |                                                                                                                                          |    |
|------|------------------------------------------------------------------------------------------------------------------------------------------|----|
| 6.11 | Curva ROC obtida na classificação de objetos da base W3BC-111. . . . .                                                                   | 80 |
| 7.1  | Resultados do teste <i>Bull's eye</i> para os extratores <i>2Hq Fature</i> e D1. . . . .                                                 | 83 |
| 7.2  | Consultas realizadas na base MPEG-7 (linhas correspondem a consultas, primeira coluna exibe objeto de consulta). . . . .                 | 84 |
| 7.3  | Amostra dos objetos da base PSB utilizados no segundo experimento. . . . .                                                               | 85 |
| 7.4  | Curvas precisão-revocação comparando o <i>3Hq Feature</i> com os extratores definidos em [Shilane et al. 2004]. . . . .                  | 86 |
| 7.5  | Curvas precisão-revocação comparando o <i>3Hq Feature</i> com os extratores definidos em [Dutagaci et al. 2010]. . . . .                 | 87 |
| 7.6  | Resultado de consultas realizadas na PSB (linhas correspondem a consultas, primeira coluna exibe objetos de consulta). . . . .           | 87 |
| 7.7  | Valores de NDCG obtidos para cada classe usando o <i>3Hq Feature</i> . . . . .                                                           | 89 |
| 7.8  | Curva precisão-revocação obtida na aplicação do <i>3Hq Feature</i> sobre a base W3BC-111. . . . .                                        | 90 |
| 7.9  | Resultado de consultas realizadas na base W3BC-111 (linhas correspondem a consultas, primeira coluna exibe objetos de consulta). . . . . | 90 |

# Lista de Tabelas

|     |                                                                                                                            |    |
|-----|----------------------------------------------------------------------------------------------------------------------------|----|
| 2.1 | Exemplos de filtros wavelets. . . . .                                                                                      | 21 |
| 3.1 | Métodos para extração de característica no espaço 2D e suas principais características. A continuar na Tabela 3.2. . . . . | 42 |
| 3.2 | Métodos para extração de característica no espaço 2D e suas principais características. Continuação da Tabela 3.1. . . . . | 43 |
| 3.3 | Métodos para extração de característica no espaço 3D e suas principais características. . . . .                            | 44 |
| 5.1 | Erro de interpolação no segundo experimento. Valores são dados em porcentagem de <i>voxels</i> incorretos. . . . .         | 68 |
| 5.2 | Tempo de execução dos experimentos de interpolação usando os algoritmos SIMOL e SIDITRANS. . . . .                         | 69 |
| 6.1 | Descrição das bases de imagens utilizadas no primeiro experimento . . . . .                                                | 71 |
| 6.2 | Parâmetros utilizados no experimento para cada base de imagens. . . . .                                                    | 71 |
| 6.3 | Valores de AUC obtidos no primeiro experimento. . . . .                                                                    | 73 |
| 6.4 | Parâmetros utilizados no primeiro cenário do experimento. . . . .                                                          | 74 |
| 6.5 | Parâmetros utilizados no segundo cenário do experimento. . . . .                                                           | 76 |
| 6.6 | Parâmetros utilizados no terceiro experimento. . . . .                                                                     | 78 |
| 6.7 | Valores de AUC obtidos no terceiro experimento. . . . .                                                                    | 78 |
| 6.8 | Parâmetros utilizados no quarto experimento. . . . .                                                                       | 79 |
| 7.1 | Parâmetros utilizados no primeiro experimento de recuperação. . . . .                                                      | 83 |
| 7.2 | Resultado da recuperação de cada classe no teste <i>Bull's eye</i> aplicado ao extrator <i>2Hq Feature</i> . . . . .       | 84 |
| 7.3 | Parâmetros utilizados no segundo experimento. . . . .                                                                      | 86 |
| 7.4 | Valores de DCG obtidos no segundo experimento. . . . .                                                                     | 87 |
| 7.5 | Classes com melhor desempenho, em termos de valores NDCG, usando o <i>3Hq Feature</i> . . . . .                            | 88 |
| 7.6 | As 40 classes com os menores valores de NDCG obtidos no experimento. . . . .                                               | 88 |
| 7.7 | Parâmetros utilizados no terceiro experimento de recuperação. . . . .                                                      | 90 |

# Lista de Algoritmos

|     |                                                                  |    |
|-----|------------------------------------------------------------------|----|
| 5.1 | Algoritmo SIMOL para interpolação de <i>slices</i> . . . . .     | 59 |
| 5.2 | Algoritmo para cálculo da transformação de distância. . . . .    | 60 |
| 5.3 | Algoritmo SIDITRANS para interpolação de <i>slices</i> . . . . . | 62 |

# Capítulo 1

## Introdução

### 1.1 Considerações Iniciais

Com os avanços em tecnologias de aquisição e armazenamento de imagens, observados nas últimas décadas, o volume de informação digital tem crescido muito rápido, sobretudo em áreas como entretenimento, engenharia, medicina e robótica. O gerenciamento eficiente de dados multimídia em grandes bancos de dados ainda constitui um grande desafio e uma necessidade para imagens [Tangelder e Veltkamp 2008, Torres e Falcão 2006, Veltkamp e Hagedoorn 2001]. A anotação manual de imagens utilizada em máquinas de buscas baseadas em texto é uma solução considerada inadequada e impraticável nos dias atuais. Uma alternativa para este problema consiste em realizar a rotulação e recuperação automáticas usando descritores [Dutagaci et al. 2010, Bustos et al. 2005].

As áreas relacionadas à recuperação de imagens por conteúdo e ao reconhecimento de padrões têm recebido exaustiva atenção nos últimos 20 anos. Diversas técnicas para análise de imagens têm sido exploradas com o objetivo de obter conjuntos de características que sejam capazes de descrever propriedades intrínsecas das imagens. Neste sentido, os métodos de extração de características são utilizados para extrair tais propriedades das imagens [Fernandes 2009].

Segundo Chanda e Majumder em [Chanda e Majumder 2006], a extração de características consiste em reduzir a quantidade de dados encontrando propriedades capazes de distinguir objetos ou partes de objetos. A meta de um extrator de característica é descrever um objeto usando medidas cujos valores são próximos, quando objetos são similares e pertencentes a uma mesma categoria. A tarefa de extração de características é fortemente dependente da aplicação e requer conhecimento sobre o domínio em que é aplicado [Duda et al. 2000]. Os extratores de características mais comuns são baseados em cor, textura e forma.

Extratores de cor, em geral, apresentam baixo custo computacional. Um exemplo clássico de método para extração de características de cor é o chamado histograma de

cor obtido pela contagem do número de *pixels* para cada cor (ou grupo de cores) em uma imagem. Porém, algumas desvantagens podem ser apontadas: o histograma de cor não mantém informações sobre a localização das cores na imagem; e a grande quantidade de espaço de memória necessário para o armazenamento [Razente 2009]. Algumas tentativas em contornar estes problemas podem ser encontradas em métodos como *color correlogram* [Huang et al. 1997], *color distribution entropy* [Sun et al. 2006], *metric histogram* [Traina et al. 2003] e *cell histogram* [Stehling et al. 2003].

Segundo Pratt em [Pratt 2007], a textura é frequentemente descrita de forma qualitativa por sua aspereza (no sentido de um remendo em um tecido de lã ser mais áspero, ou grosseiro, do que em um tecido de seda, sob as mesmas condições de visualização). O quanto uma textura é áspera é dado por um valor relacionado ao período de repetição da estrutura local, de modo que quanto maior o período, mais grosseira é a textura. Diversos extratores de textura têm sido desenvolvidos no sentido de medir uma textura quantitativamente, usando, para isto, ferramentas como filtros Gabor [Sandler e Lindenbaum 2006, Vyas e Rege 2006, Grigorescu et al. 2002], transformadas wavelet [Zhang et al. 2010, Mor e Aladjem 2005] e de Fourier [Arof e Deravi 1998, Wen e Acharya 1998], autocorrelação [Canlin et al. 2008], matriz de coocorrência [Xie et al. 2010, Park e Chen 2001, Gotlieb e Kreyszig 1990], decomposição de valor singular [Selvan et al. 2009, Selvan e Ramakrishnan 2007] e campos aleatórios de Markov [Razente 2009, Blunsden e Atallah 2005] [Pratt 2007, Pedrini e Schwartz 2007].

Os descritores de forma têm sido amplamente utilizados nas áreas de reconhecimento de padrões (veja Seção 1.1.1) e recuperação de imagens por similaridade (veja Seção 1.1.2). Segundo Manjunath em [Manjunath 2002], os seres humanos são capazes de reconhecer objetos característicos apenas por sua forma, pois a forma geralmente carrega informação semântica. As características de forma são, em geral, fortemente ligadas à funcionalidade e identidade dos objetos. Estas propriedades dão destaque a este tipo de descritor, tornando-se um dos tipos de características mais importantes em sistemas de recuperação de imagens por conteúdo [Zhang e Lu 2003, Manjunath 2002]. Após uma imagem ter sido segmentada, a região de interesse obtida (ou seu contorno) pode ser representada por um descritor de forma [Gonzalez e Woods 2008].

Atualmente, uma grande quantidade de métodos extratores de forma pode ser encontrada na literatura. Alguns destes descritores são desenvolvidos para caracterizar formas bidimensionais outros para descrever formas tridimensionais. O Capítulo 3 deste trabalho apresenta diversos exemplos de métodos para extração de características de formas, tanto no espaço 2D, quanto no 3D.

### 1.1.1 Classificação de Padrões

A classificação de padrões tem por objetivo rotular elementos de um conjunto de dados com base nas propriedades extraídas dos mesmos. Assim, amostras que compartilham propriedades em comum são ditas similares e devem ser classificadas com um mesmo rótulo, ou seja, devem pertencer a uma mesma classe ou categoria. Os algoritmos responsáveis por realizar esta tarefa são denominados *algoritmos de classificação* ou simplesmente *classificadores* [Fernandes 2009].

O processo de classificação pode ser supervisionado ou não-supervisionado. Na classificação supervisionada os agrupamentos no espaço de características são pré-determinados, ou seja, as classes são previamente conhecidas [Jähne 1997]. Antes de a classificação ser realizada, uma etapa de treinamento é executada, a partir de amostras, para se obter parâmetros que caracterizem cada classe. Para estas amostras, denominadas conjunto de treinamento, são previamente conhecidos os seus descritores (obtidos por um extrator de características) e a classe a que pertencem [Pedrini e Schwartz 2007].

Já a classificação não-supervisionada não dispõe de nenhum conhecimento prévio sobre as classes e os objetos a serem classificados. Assim, todas as informações necessárias devem ser obtidas a partir dos próprios objetos a serem classificados. Como acontece na classificação supervisionada, objetos que compartilham características similares devem pertencer a uma mesma classe [Pedrini e Schwartz 2007, Jähne 1997].

Exemplos de métodos de classificação, que podem ser encontrados na literatura, são: classificador Bayesiano, janela de Parzen, k-vizinhos mais próximos, redes neurais, aprendizado de Boltzmann, árvores de decisão, *support vector machine*, algoritmos de agrupamento hierárquicos e aglomerativos (dos quais os dois últimos são não-supervisionados, enquanto que os demais são supervisionados) [Pedrini e Schwartz 2007, Russ 2006, Duda et al. 2000]. Na Seção 2.3 deste trabalho, são descritos alguns dos métodos utilizados para avaliação de desempenho de classificadores.

### 1.1.2 Sistemas de Recuperação de Imagem por Conteúdo

Um sistema de recuperação de imagem baseado em conteúdo (também conhecido pela sigla CBIR - *Content-Based Image Retrieval*) é constituído, essencialmente, de um banco de dados de imagem, um descritor e uma estrutura de dados para indexação [Torres e Falcão 2006]. O objetivo de um sistema CBIR consiste em realizar a busca eficiente por imagens relevantes em grandes bancos de dados, usando características visuais, tais como cor, forma e/ou textura, extraídas de forma automática das próprias imagens [Eakins e Graham 1999].

Em um sistema CBIR, cada imagem pode ser interpretada por um ponto em um espaço métrico  $n$ -dimensional, onde cada dimensão corresponde a uma característica [Torres e Falcão 2006]. A relevância (ou “distância”) entre uma imagem de consulta (*query*) e uma

imagem alvo é quantificada de acordo com uma medida de similaridade calculada a partir das características visuais obtidas neste espaço métrico [Tangelder e Veltkamp 2008, Chen e Wang 2002]. Assim, para uma dada imagem de consulta, o sistema CBIR procura retornar imagens que sejam semelhantes à imagem de consulta tão rápido quanto possível.

Muitos dos trabalhos encontrados na literatura têm dado foco à etapa de extração de características de modo a obter descritores os mais discriminativos possíveis. Alguns dos aspectos que ajudam a avaliar ou até mesmo comparar os descritores obtidos são [Tangelder e Veltkamp 2008, Veltkamp e Tanase 2002]:

- *Eficiência*: o descritor da imagem de consulta deve ser calculado num tempo rápido o suficiente para permitir uma busca interativa.
- *Poder discriminativo*: o descritor deve ser capaz de capturar as características que diferenciam entre uma classe ou outra de imagens/objetos. Quais características devem ser observadas é uma questão a ser definida para cada tipo de aplicação.
- *Robustez e sensibilidade*: é comumente desejável que o descritor seja robusto contra ruídos e artefatos. Também é esperado que pequenas mudanças na imagem/objeto acarretem pequenas mudanças no descritor, levando assim a uma boa sensibilidade e capacidade de discriminação.
- *Normalização de pose*: é desejável que descritores seja invariantes com relação a transformações geométricas tais como translação, rotação e escala. Entretanto, nem todos os descritores alcançam tal invariância, sendo, portanto, necessário submeter as imagens/objetos a um procedimento de normalização de pose.

Apesar de ser considerado um dos aspectos mais difíceis na busca baseada por conteúdo, as características de forma, em particular, são amplamente usadas em sistemas CBIR [Veltkamp e Hagedoorn 2001]. Este trabalho se concentra principalmente no desenvolvimento de extratores de características de forma e sua efetividade para classificação e recuperação de formas baseada em conteúdo.

## 1.2 Objetivos

Este trabalho propõe o desenvolvimento de dois novos descritores de forma: 1) um descritor de formas bidimensionais denominado *2Hq Feature*; 2) um descritor de formas tridimensionais denominado *3Hq Feature*. Ambos os descritores definem etapas de redimensionamento e decomposição e utilizam a curva de preenchimento de espaço de Hilbert e a transformada discreta de wavelet. Os extratores propostos neste trabalho são invariantes a transformações geométricas de translação e robustos a rotações e escalas. Tarefas de classificação e de recuperação de imagens por conteúdo são executadas, a fim de avaliar a eficiência dos extratores propostos.

Este trabalho apresenta também dois novos métodos para reconstrução de objetos 3D a partir de contornos 2D. Um *framework* foi desenvolvido para a realização de segmentação manual e obtenção dos contornos.

## 1.3 Organização do Trabalho

A fim de alcançar os objetivos acima citados, o restante deste trabalho é organizado da seguinte maneira: o Capítulo 2 apresenta os conceitos teóricos e suas principais características, que são de grande importância para o desenvolvimento deste trabalho. Incluem-se neste capítulo assuntos tais como curva de preenchimento de espaço, transformada wavelet, redes neurais artificiais e métodos para a avaliação de desempenho de classificadores ou sistemas CBIR. No Capítulo 3 é apresentada uma revisão da literatura sobre descritores de forma. O capítulo aborda tanto métodos aplicados a imagens 2D quanto a objetos 3D e resume, ao final, suas principais características. No Capítulo 4 são introduzidos os métodos *2Hq Feature* e *3Hq Feature* propostos neste trabalho para extração de características de formas 2D e 3D, respectivamente. O Capítulo 5 descreve o *framework* desenvolvido para segmentação e reconstrução de imagens tridimensionais, a fim de facilitar a aplicabilidade do método para imagens reais em três dimensões. Neste capítulo é introduzido o ambiente EMISS desenvolvido para segmentação de contornos, e são propostas duas novas abordagens para reconstrução tridimensional, denominadas SIMOL e SIDITRANS. Para ilustrar os ganhos obtidos com o uso dos extratores *2Hq Feature* e *3Hq Feature*, experimentos e resultados são discutidos nos Capítulos 6 e 7. Os experimentos abordados nestes capítulos incluem a utilização de redes neurais artificiais para classificação e recuperação de imagens por conteúdo em diversas bases de imagens (públicas e de acesso restrito). Finalizando este trabalho, o Capítulo 8 aponta as conclusões obtidas e lista alguns possíveis trabalhos futuros.

# Capítulo 2

## Fundamentação Teórica

Este capítulo tem como objetivo descrever brevemente os conceitos teóricos necessários para o entendimento deste trabalho, apontando suas principais características.

### 2.1 Transformada Discreta de Wavelet

Nos últimos anos, a transformada wavelet tem-se mostrado uma importante ferramenta para a solução de problemas em diversas áreas, como processamento digital de imagens, processamento de sinais, análise de formas, reconhecimento de padrões, visão computacional, compressão de dados, análise numérica, equações diferenciais, teoria da aproximação, teoria da amostragem, ótica quântica, medicina, engenharia biomédica, computação gráfica e muitas outras. A transformada wavelet fornece uma maneira de descrever tanto informações espaciais quanto informações de frequência de uma imagem [Gonzalez e Woods 2008, Pedrini e Schwartz 2007, Debnath 2002].

Em [Mallat 1987], Mallat apresenta a wavelet como a base da teoria da multirresolução. A teoria da multirresolução se concentra na representação e análise de sinais (ou imagens) em diversas resoluções. A teoria incorpora técnicas de outras disciplinas, incluindo a codificação em sub-banda (*subband coding*) do processamento de sinais, os bancos de filtros QMF (*quadrature mirror filtering*) de reconhecimento de voz e processamento de imagem piramidal. A principal motivação da teoria da multirresolução é que características que podem não ser detectadas em uma resolução podem ser facilmente detectadas em outra resolução [Gonzalez e Woods 2008, Pedrini e Schwartz 2007].

Um banco de filtros QMF é um conjunto de filtros que divide a entrada, seja um sinal ou uma imagem, em bandas distintas (assim, os subsinais resultantes podem ser compactados de forma mais eficiente, o que é realizado através de subamostragem). No caso das wavelets, o banco de filtros é formado por filtros passa-baixa e passa-alta, os quais formam as bases para a transformada de wavelet [Pedrini e Schwartz 2007, Strang e Nguyen 1996]. Exemplos de bancos de filtros são apresentados na Tabela 2.1 que ilustra os filtros de Haar e de Daubechies [Pedrini e Schwartz 2007]. Outros filtros usados pela transformada

Tabela 2.1: Exemplos de filtros wavelets.

| Nome       | Filtro passa-baixa (l)                                                                                                                               | filtro passa-alta (h)                                                                                                                                  |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| Haar       | $\left[\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}}\right]$                                                                                           | $\left[\frac{1}{\sqrt{2}} \quad -\frac{1}{\sqrt{2}}\right]$                                                                                            |
| Daubechies | $\left[\frac{1+\sqrt{3}}{4\sqrt{2}} \quad \frac{3+\sqrt{3}}{4\sqrt{2}} \quad \frac{3-\sqrt{3}}{4\sqrt{2}} \quad \frac{1-\sqrt{3}}{4\sqrt{2}}\right]$ | $\left[\frac{1-\sqrt{3}}{4\sqrt{2}} \quad -\frac{3-\sqrt{3}}{4\sqrt{2}} \quad \frac{3+\sqrt{3}}{4\sqrt{2}} \quad -\frac{1+\sqrt{3}}{4\sqrt{2}}\right]$ |

wavelet, que podem ser encontrados na literatura, incluem Coiflets, Symmlets, Morlet e Meyer discreta. Porém, os filtros de Haar são os mais simples, e seu cálculo o menos dispendioso computacionalmente [Chui 1992].

Uma pirâmide de imagens consiste de uma coleção de imagens em resoluções decrescentes organizadas em uma pirâmide (veja a Figura 2.1a). A base da pirâmide contém a representação em alta resolução, enquanto que o topo contém a aproximação em baixa resolução da imagem [Gonzalez e Woods 2008].

Na codificação em sub-bandas, uma entrada  $f(n)$  é decomposta em um conjunto de bandas. A Figura 2.1b ilustra um sistema composto de dois bancos de filtros (um para análise e outro para síntese). Na imagem os blocos  $h0$  e  $h1$  correspondem aos *filtros de análise passa-baixa e passa-alta*, respectivamente. Os blocos  $g0$  e  $g1$  indicam *filtros de síntese passa-baixa e passa-alta*, respectivamente, usados para a reconstrução de  $f(n)$ . Os blocos  $2\downarrow$  e  $2\uparrow$  indicam operações de subamostragem e superamostragem, respectivamente. Ao final da etapa de análise, a entrada  $f(n)$  é dividida em *coeficientes de aproximação*, resultantes do filtro passa-baixa, e em *coeficientes de detalhe*, resultantes da aplicação do filtro passa-alta. A reconstrução de  $f(n)$  é realizada utilizando um banco de filtros de síntese (filtros passa-baixa  $g0$  e passa-alta  $g1$ ). Em imagens digitais, os coeficientes de aproximação enfatizam objetos grandes e com baixas variações de intensidade, os coeficientes de detalhe enfatizam detalhes finos como bordas ou arestas, por exemplo, [Gonzalez e Woods 2008, Pedrini e Schwartz 2007, Weeks 2007].

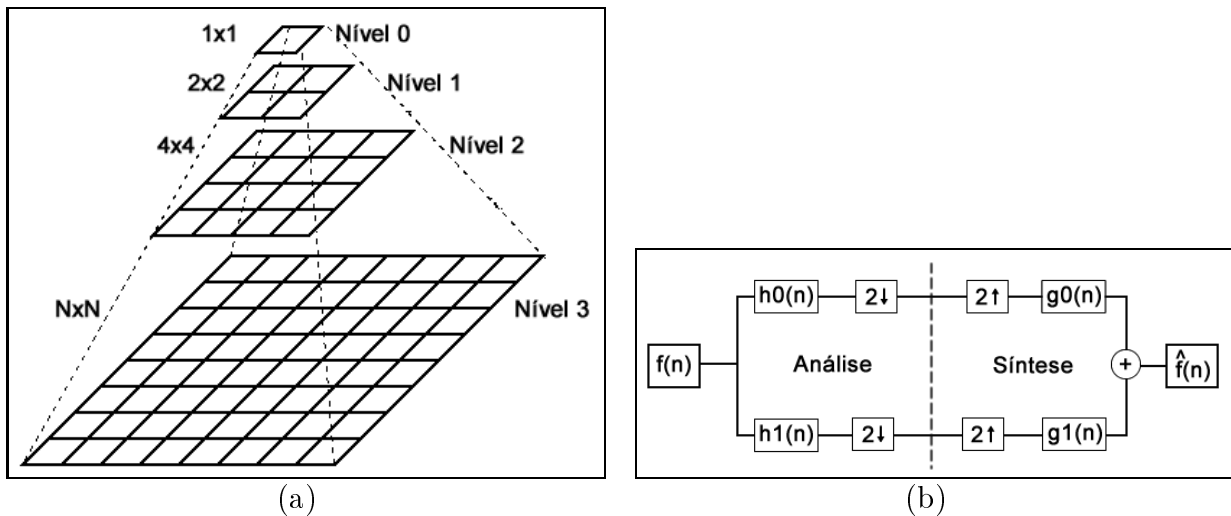


Figura 2.1: a) Esquema de uma pirâmide de imagens. b) Principais componentes de um sistema de codificação e decodificação de sub-banda com duas bandas.

A transformada wavelet de uma imagem pode ser obtida ao aplicar os filtros  $h0$  e  $h1$

em cada dimensão da imagem separadamente, gerando um nível de decomposição e produzindo quatro sub-bandas:  $LL$ ,  $LH$ ,  $HL$  e  $HH$ . As sub-bandas  $LL$  e  $HH$  descrevem as frequências baixas e altas, respectivamente, enquanto que as sub-bandas  $LH$  e  $HL$  descrevem frequências intermediárias. A sub-banda  $LL$  corresponde à aproximação da imagem, já as demais sub-bandas correspondem às imagens de detalhe [Pedrini e Schwartz 2007].

A aplicação dos filtros QMF em uma imagem é chamada *decomposição wavelet* e pode ser realizada recursivamente na sub-banda  $LL$  [Pedrini e Schwartz 2007, Weeks 2007]. A Figura 2.2 ilustra a decomposição em até três níveis.

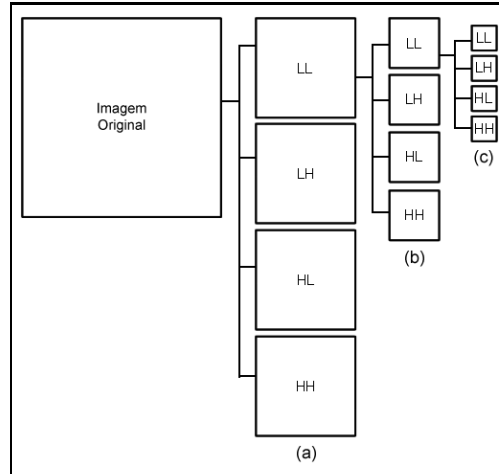


Figura 2.2: a) Decomposição em um nível. b) Decomposição em dois níveis. c) Decomposição em três níveis.

Neste trabalho, a aplicação da transformada discreta de wavelet é utilizada para obter uma subamostragem dos dados, reduzindo, assim, a quantidade de informação a ser processada (para este propósito, somente serão utilizados os coeficientes de aproximação resultantes da transformada).

## 2.2 Redes Neurais Artificiais

Uma rede neural artificial (NN) é um sistema dinâmico de processamento de informações que tenta modelar a capacidade de processamento do cérebro humano [Rojas 1996]. As NNs têm estado presentes em quase todas as áreas onde há um conjunto de dados disponíveis e a necessidade de uma boa solução [Kasabov 1998]. Em especial, áreas como a robótica, processamento de imagens e de sinais e inteligência artificial têm feito grande uso desta ferramenta [Fausett 1993].

As principais características de uma NN são sua robustez, sua capacidade de generalização, seu forte potencial de paralelismo, sua capacidade de realizar *matching* parcial e, uma das características que mais tem chamado a atenção, a sua capacidade de aprendizado. Dada uma tarefa a resolver e uma classe de funções  $F$ , o aprendizado consiste

em utilizar um conjunto de observações para encontrar  $f^* \in F$  que resolva a tarefa de maneira ótima [Guliato et al. 2010, Kasabov 1998].

Uma das mais simples tarefas em que uma rede neural pode ser treinada é o reconhecimento de padrões [Harvey 1994]. Um problema típico em reconhecimento de padrões poderia ser formulado da seguinte forma: dado um conjunto de  $n$  padrões conhecidos e um novo padrão desconhecido (*query*), descobrir quais dos padrões já conhecidos mais se assemelham ao padrão *query*. Imagens, assinaturas ou um conjunto de características são exemplos de padrões [Kasabov 1998].

Uma NN consiste de elementos de processamento, denominados neurônios, e de conexões entre os neurônios, tendo pesos associados [Kasabov 1998]. Um neurônio artificial, por sua vez, é constituído de conexões de entradas, uma função de entrada, uma função de ativação e uma função de saída. As entradas do neurônio são valores numéricos, provenientes do meio externo ou de outros neurônios, e comumente possuem valores de pesos associados. A função de entrada é responsável por combinar as várias entradas de um neurônio, geralmente através da soma dos produtos das entradas pelos pesos. Uma função de ativação calcula o nível de ativação do neurônio, restringindo a amplitude do valor de saída. A saída do neurônio transfere o valor final para um próximo neurônio ou de volta para o meio externo [Ribeiro 2009, Kasabov 1998, Rojas 1996]. A Figura 2.3 mostra o neurônio artificial e seus componentes.

Assim, uma rede neural artificial constitui a ligação entre diversos neurônios em uma estrutura de rede [Fausett 1993]. A Figura 2.4 ilustra uma rede neural simples. Conforme pode ser observado na figura, a topologia da rede define três camadas de neurônios: de entrada, intermediária e de saída. Aqui, cada neurônio de uma camada é conectado a todos os outros neurônios da camada vizinha.

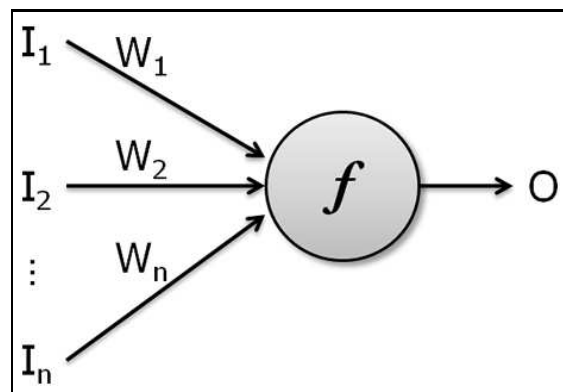


Figura 2.3: Modelo de um neurônio artificial. As entradas, pesos e saída são representadas por  $I$ ,  $W$  e  $O$ , respectivamente.

Neste trabalho, redes neurais são utilizadas para dois fins: primeiramente, são utilizadas para realizar a classificação de imagens e objetos no Capítulo 6. Já no Capítulo

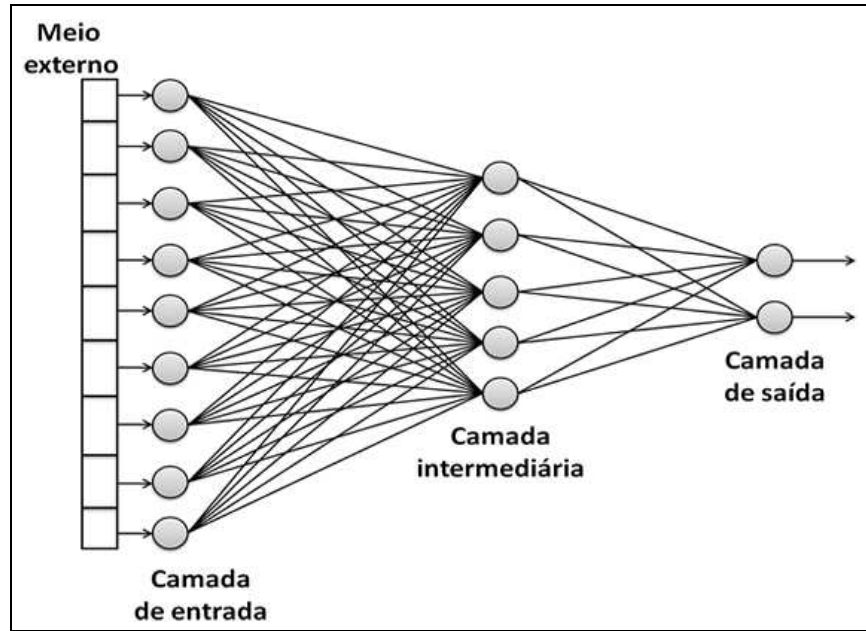


Figura 2.4: Exemplo de uma rede neural artificial. Neurônios são representados por círculos e conexões entre neurônios por arestas.

7, as redes neurais são utilizadas com o propósito de redução da dimensionalidade dos dados.

## 2.3 Medidas de Avaliação de Desempenho

Nesta seção, são discutidas algumas das medidas de desempenho comumente utilizadas em comparações entre sistemas de classificação ou entre sistemas de recuperação de imagens por conteúdo.

### 2.3.1 Curva ROC e Área Sobre a Curva ROC

A curva denominada *Receiver Operating Characteristics* (ROC) é uma técnica para visualização, organização e seleção de classificadores baseados em sua performance. Dado um classificador e um conjunto de instâncias (o conjunto de testes), uma matriz de confusão 2x2 (também chamado tabela de contingência) pode ser construída, representando as disposições do conjunto de instâncias [Fawcett 2005]. Curvas ROC são gráficos bidimensionais em que a taxa de verdadeiros positivos (também conhecida como sensibilidade) é traçada no eixo vertical e a taxa de falsos positivos (igual a 1 - especificidade) é traçada no eixo horizontal. Uma curva ROC descreve a relação entre custos (falsos positivos) e benefícios (verdadeiros positivos) [Davis e Goadrich 2006].

$$TPR = \frac{TP}{TP+FN} \quad FPR = \frac{FP}{FP+TN} \quad (2.1)$$

As Equações em 2.1 apresentam as fórmulas para o cálculo das taxas de verdadeiros

positivos (TPR) e falsos positivos (FPR). As variáveis TP, FP, TN e FN correspondem a verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos, respectivamente [Davis e Goadrich 2006]. Verdadeiros positivos são instâncias classificadas corretamente como positivas, enquanto que falsos positivos são instâncias erroneamente classificadas como positivas. Analogamente, verdadeiros negativos são instâncias classificadas corretamente como negativas, e falsos negativos são instâncias classificadas incorretamente como negativas [Manning et al. 2009].

Alguns classificadores, tais como o Bayesiano ingênuo ou a rede neural, produzem naturalmente um valor numérico que representa o grau com que uma instância pertence a uma classe. Nestes classificadores, se o valor produzido está acima de um limiar  $t$ , o classificador produz uma resposta binária: a amostra pertence, ou não, a uma determinada classe. Cada valor de limiar produz um ponto diferente no espaço ROC. Qualquer curva ROC gerada de um conjunto finito de instâncias é na verdade uma função degrau que aproxima uma curva verdadeira, à medida que o número de instâncias se aproxima de infinito, e pode ser pensada como uma representação completa do desempenho de um classificador, à medida que o limiar  $t$  de classificação varia [Krzanowski e Hand 2009, Fawcett 2005].

Uma estratégia comum na avaliação do desempenho de um classificador é reduzir a curva ROC a um valor escalar. Isto pode ser feito calculando-se a área abaixo da curva ROC (abreviado por AUC - *area under curve*) [Bradley 1997, Hanley e McNeil 1982]. Uma vez que a AUC é uma porção da área de um quadrado unitário, seu valor estará sempre entre zero e um. A AUC pode ser vista como uma medida global de separabilidade entre as distribuições de resultados (*scores*) para populações positiva e negativa. Uma característica importante: a medida não requer a escolha de um limiar  $t$ , pois sumariza todos as possíveis escolhas deste parâmetro [Krzanowski e Hand 2009].

A Figura 2.5 exhibe exemplos de curva ROC, ilustrando também a separação entre as classes (obtida por um classificador), as taxas de positivos verdadeiros (TPR) e falsos positivos (FPR) para um determinado limiar de teste (na imagem as cores azul, ciano, vermelho e magenta correspondem às variáveis TN, FP, TP e FN da Equações em 2.1). Também é indicado, nos exemplos, o valor da área sob a curva ROC (observe que a curva da Figura 2.5a possui valor de AUC superior a curva da Figura 2.5b).

Classificadores cuja curva ROC coincide com o segmento de reta que liga os pontos (0,0) e (1,1) do espaço ROC são ditos classificadores aleatórios. Observa-se que, quanto melhor é a classificação, mais próxima a curva passa do ponto (0,1) do espaço ROC (na classificação ideal a curva ROC passa por este ponto).

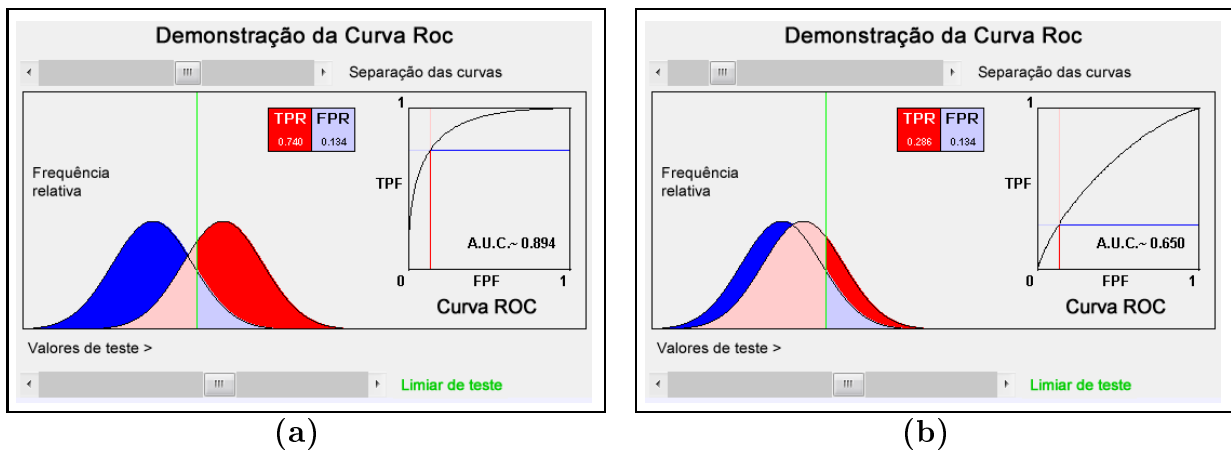


Figura 2.5: Demonstração da curva ROC (imagens obtidas a partir de <http://www.anaesthetist.com/mnm/stats/roc/Findex.htm>).

### 2.3.2 Curva Precisão-Revocação

Outra medida de avaliação bastante comum é a curva Precisão-Revocação (PR), muito utilizada quando as bases de dados apresentam classes com quantidades de objetos altamente desbalanceadas. Um gráfico Precisão-Revocação exibe valores de precisão no eixo vertical e valores de revocação no eixo horizontal [Manning et al. 2009, Davis e Goadrich 2006]. A Equação 2.2 ilustra como são calculados os valores de precisão e revocação (observe que a revocação é calculada de maneira idêntica à taxa de verdadeiros positivos da curva ROC).

$$\text{Precisão} = \frac{TP}{TP+FP} \quad \text{Revocação} = \frac{TP}{TP+FN} \quad (2.2)$$

Davis e Goadrich em [Davis e Goadrich 2006] mostram que existe um relacionamento entre curvas ROC e curvas Precisão-Revocação. Em seu trabalho, eles provam o seguinte teorema: “para um dado banco de dados de valores positivos e negativos, existe uma correspondência de um para um entre uma curva no espaço ROC e uma curva no espaço PR, de modo que estas curvas possuam exatamente as mesmas matrizes de confusão, se o valor de revocação é diferente de zero”. Uma consequência deste teorema é a possibilidade de se converter uma curva ROC em uma curva Precisão-Revocação, e *vice-versa*.

A Figura 2.6 exibe um exemplo de curva precisão-revocação. Uma possível interpretação para esta curva é a seguinte: em uma tarefa de recuperação de imagens por conteúdo, ao recuperar 30% das imagens relevantes (revocação igual a 0.30), qual a precisão dos resultados obtidos? Para uma recuperação perfeita (ideal), observa-se que a curva se torna um segmento de reta que liga os pontos (0,1) e (1,1) no espaço precisão-revocação (em outras palavras, a precisão é sempre de 100% para qualquer valor de revocação).

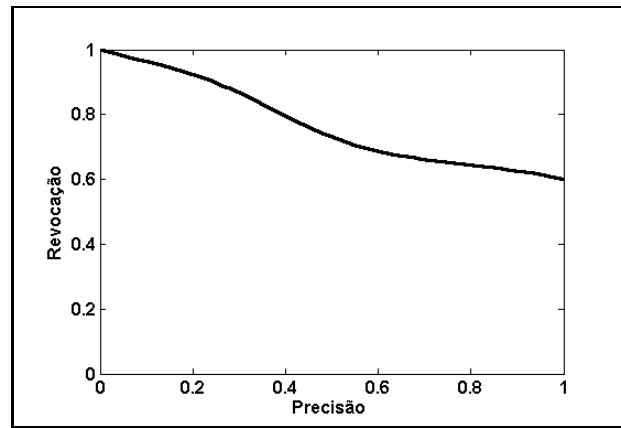


Figura 2.6: Exemplo de curva precisão-revocação.

### 2.3.3 Ganho Acumulativo Descontado

Uma medida de avaliação de desempenho, muito utilizada em aplicações de recuperação de informação e em algoritmos de busca na *Web* e que tem ganho bastante aceitação nos últimos anos, especialmente em trabalhos de recuperação de objetos tridimensionais, é a chamada Ganho Acumulativo Descontado (DCG – *Discounted Cumulative Gain*) [Croft et al. 2009]. A medida é projetada para ponderar mais fortemente resultados corretos no início de uma lista de *ranking*, enquanto resultados corretos que aparecem mais ao final da lista contribuem menos para o valor da medida [Manning et al. 2009] [Shilane et al. 2004]. Esta medida estatística sumarizada parte do princípio de que é menos provável que um usuário considere instâncias mais distantes dos primeiros resultados obtidos em uma busca [Shilane et al. 2004].

A medida é calculada a partir da lista *ranking*, obtida em uma busca. Dado um objeto de consulta (*query*), o DCG mede a utilidade (o ganho) de cada resposta, baseando-se na posição desta resposta, na lista de classificação (*ranking*), ponderando mais fracamente os objetos relevantes que aparecem em posições mais baixas na lista. Suponha que  $n$  candidatos sejam recuperados para uma dada consulta e um candidato na  $i$ -ésima posição da lista de classificação é associado a um valor  $rel_i$ , que indica a relevância deste candidato. Assim, o valor *DCG* em uma posição particular  $k$  da classificação é definido como [Croft et al. 2009]:

$$DCG_k = rel_1 + \sum_{i=2}^k \frac{rel_i}{\log_2 i} \quad (2.3)$$

com  $rel_i \in \{0, 1\}$  em aplicações onde o grau de relevância de um candidato é binário, ou seja, pode ser dito como não relevante ou relevante [Croft et al. 2009].

Contudo, para que seja possível comprar o desempenho de diferentes algoritmos, os valores de DCG, obtidos para cada algoritmo, devem ser normalizados [Croft et al. 2009, Manning et al. 2009, Shilane et al. 2004]. Para uma dada consulta, o valor normalizado do ganho acumulativo descontado (NDCG), na posição  $k$  da classificação obtida, é definido

como:

$$NDCG_k = \frac{DCG_k}{IDCG_k} \quad IDCG_k = 1 + \sum_{i=2}^{|C|} \frac{1}{\log_2 i} \quad (2.4)$$

onde o IDCG corresponde ao valor ideal do DCG (obtido com uma lista de classificação perfeita), e  $|C|$  é o número de elementos da classe do objeto de consulta.

O valor do NDCG é, portanto, definido no intervalo  $[0, 1]$ . Os valores de NDCG obtidos em todas as consultas podem ser tomados em uma média aritmética, e seu resultado pode ser tomado como medida média de desempenho do algoritmo correspondente [Croft et al. 2009].

As medidas descritas nesta seção são utilizadas nos Capítulos 6 e 7, para avaliar o desempenho dos descritores propostos no Capítulo 4.

## 2.4 Curvas de Preenchimento de Espaço

Uma curva de preenchimento de espaço pode ser definida como um caminho contínuo que visita todos os pontos de uma grade  $n$ -dimensional uma única vez e nunca se cruza [Prusinkiewicz e Lindenmayer 2004]. Ela pode ser aplicada a imagens e volumes (neste caso, a matriz de *pixels* ou o volume de *voxels* são as grades que a curva irá percorrer).

Uma característica bastante interessante e útil é a sua capacidade em preservar a localização da informação. Esta característica produz uma alta correlação dos dados obtidos na sequência determinada pela curva. Em outras palavras, diz-se que características próximas na imagem tendem, tanto quanto possível, a permanecer próximas na representação unidimensional obtida pela curva. Dentre as diversas curvas de preenchimento de espaço, a de Hilbert é a que mais fortemente apresenta esta característica [Armstrong et al. 2009, Ebrahim et al. 2008].

A curva de Hilbert é definida a partir da seguinte gramática:

$$\begin{aligned} S &\rightarrow +BF - SFS - FB+ \\ B &\rightarrow -SF + BFB + FS- \\ + &\rightarrow \text{Virar para esquerda} \\ - &\rightarrow \text{Virar para direita} \\ F &\rightarrow \text{Seguir em frente} \end{aligned} \quad (2.5)$$

A notação acima pode ser interpretada segundo a geometria tartaruga apresentada em [Prusinkiewicz e Lindenmayer 2004]: considere a tupla  $(x, y, \alpha)$ , onde  $(x, y)$  representa a posição e  $\alpha$  a direção para a qual a tartaruga está voltada. Dado um tamanho de passo  $d$  e um incremento de ângulo  $\delta$ , a tartaruga então responde a comandos de acordo com os símbolos da gramática. Para a curva de preenchimento de espaço de Hilbert, os valores

$d = 1$  e  $\delta = 90^\circ$  são considerados e os comandos são dados pelos símbolos ‘+’, ‘-’ e ‘F’. Observe que os comandos ‘+’ e ‘-’ apenas modificam a orientação da tartaruga, sua posição será alterada somente pelo símbolo ‘F’.

Uma curva de preenchimento de espaço pode ser gerada em diversos níveis [Armstrong et al. 2009]. Na gramática apresentada na Equação 2.5, para gerar uma curva de nível  $n$  aplica-se as regras de produção até no máximo  $n$  níveis de recursão, ignorando assim quaisquer comandos presentes em uma recursão posterior a  $n$ . A Figura 2.7 ilustra as curvas de Hilbert para os níveis de 1 a 4.

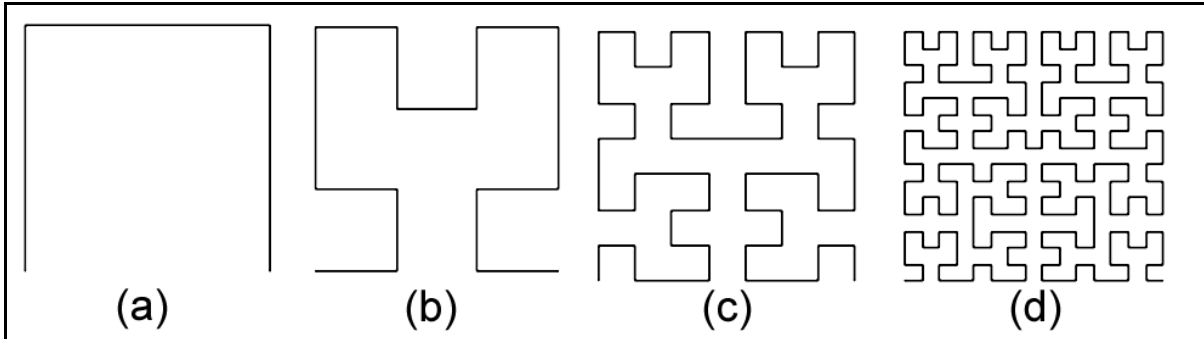


Figura 2.7: Curvas de preenchimento de espaço de Hilbert para os níveis de um a quatro.

Um método alternativo para a criação da curva de Hilbert é ilustrado na Figura 2.8. Na figura, uma curva de nível 1 (Figura 2.8a) é copiada, e suas cópias, em cinza claro, são transladadas como mostra a Figura 2.8b. Em seguida, na Figura 2.8c, duas instâncias da curva de nível 1, destacadas em cinza claro, são rotacionadas 90 graus em sentidos horário e anti-horário. Finalmente, na Figura 2.8d, as curvas são conectadas a fim de se obter a curva de Hilbert de nível 2. Este procedimento pode ser utilizado para gerar a curva de Hilbert em qualquer nível a partir de uma curva de nível anterior [Ebrahim et al. 2008].

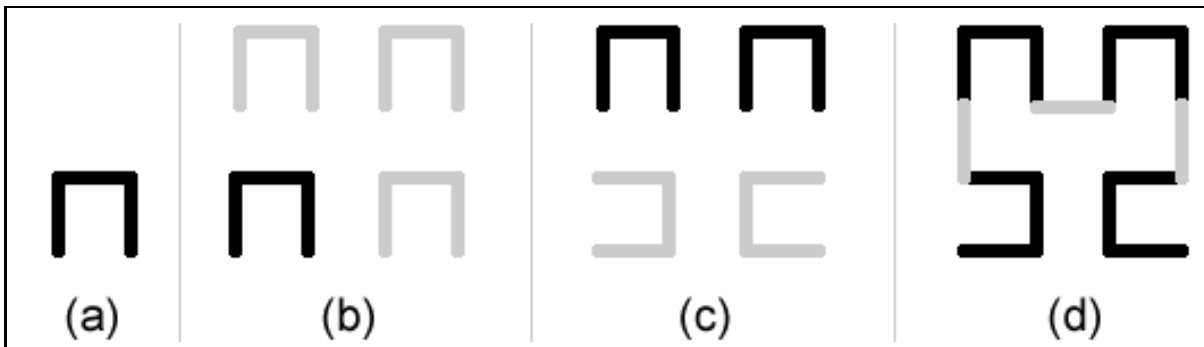


Figura 2.8: Método alternativo para geração da curva de Hilbert.

A curva de Hilbert pode ser estendida para três dimensões. Para isto, considere um sistema de coordenadas cujos eixos  $\vec{H}$ ,  $\vec{L}$  e  $\vec{U}$  indicam as direções para onde a cabeça da tartaruga está voltada, para seu lado esquerdo e a para cima (direção do casco),

respectivamente. Considere também os valores  $d = 1$  e  $\delta = 90^\circ$  para incrementos de posição e orientação, respectivamente [Prusinkiewicz e Lindenmayer 2004]. A seguinte gramática é capaz de gerar a curva tridimensional de Hilbert:

$$\begin{aligned}
 S &\rightarrow B - F + CFC + F - D \& F \wedge D - F + \& \& CFC + F + B // \\
 B &\rightarrow S \& F \wedge CFB \wedge F \wedge D \wedge \wedge - F - D \wedge | F \wedge B | FC \wedge F \wedge S // \\
 C &\rightarrow | D \wedge | F \wedge B - F + C \wedge F \wedge S \& \& FS \& F \wedge C + F + B \wedge F \wedge D // \\
 D &\rightarrow | CFB - F + B | FS \& F \wedge S \& \& FB - F + B | FC // \\
 + &\rightarrow \text{Girar } \delta \text{ ao redor do eixo } \vec{U} \\
 - &\rightarrow \text{Girar } -\delta \text{ ao redor do eixo } \vec{U} \\
 | &\rightarrow \text{Girar } 2\delta \text{ ao redor do eixo } \vec{U} \\
 \& &\rightarrow \text{Girar } \delta \text{ ao redor do eixo } \vec{L} \\
 \wedge &\rightarrow \text{Girar } -\delta \text{ ao redor do eixo } \vec{L} \\
 / &\rightarrow \text{Girar } \delta \text{ ao redor do eixo } \vec{H} \\
 F &\rightarrow \text{Seguir em frente}
 \end{aligned} \tag{2.6}$$

A Figura 2.9 exibe as curvas tridimensionais de Hilbert para os níveis um e dois.

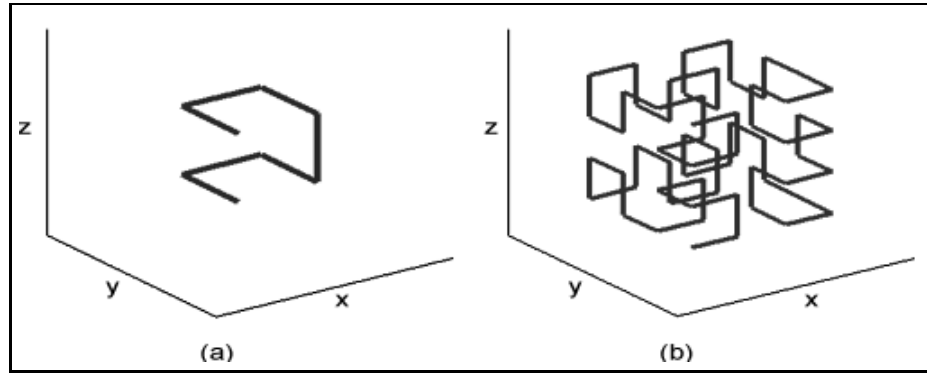


Figura 2.9: Curvas tridimensionais de Hilbert para níveis 1 e 2.

Neste trabalho, as curvas de preenchimento de espaço de Hilbert são utilizadas para a obtenção dos descritores de características de forma, propostos no Capítulo 4.

## 2.5 Interpolação de *Slices*

Durante as últimas décadas, muitos algoritmos têm sido desenvolvidos para reconstrução tridimensional a partir de um conjunto de imagens 2D de seções transversais de um objeto 3D. Muitos destes métodos utilizam informação de contorno da forma a ser reconstruída, criando modelos precisos para regiões de interesse no espaço 3D [Nonato et al. 2005]. Aplicações em medicina envolvendo imagens de tomografia computadorizada (CT) e de ressonância magnética (MRI) frequentemente fazem uso destes algoritmos. Em geral, as seções transversais de um objeto (também conhecidas como *slices*) não

são próximas o bastante, havendo um *gap* de informação entre os *slices*. Assim uma interpolação se faz necessária, de modo a preencher esta lacuna de informação [Migeon et al. 1998].

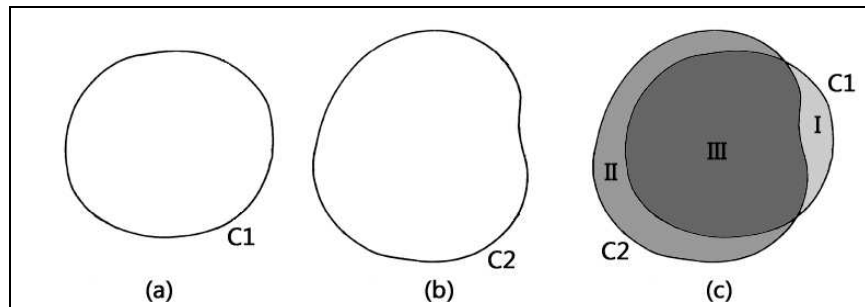


Figura 2.10: a) Contorno  $C_1$  a ser interpolado. b) Contorno  $C_2$  adjacente a  $C_1$ . c) Sobreposição dos contornos  $C_1$  e  $C_2$  delimitando as regiões I, II e III.

Diferentes abordagens para a interpolação de *slices* têm sido desenvolvidas. Guo et al. em [Guo et al. 1995] descrevem um método de interpolação usando operações morfológicas. Os autores sobrepõem os contornos  $C_1$  e  $C_2$  de *slices* adjacentes, formando o que eles denominam “Imagem Diferença Morfológica”, a fim de identificar 3 tipos de regiões: *I*) região pertencente apenas ao contorno  $C_1$ ; *II*) região pertencente apenas ao contorno  $C_2$ ; *III*) região pertencente a ambos os contornos. A Figura 2.10 exemplifica dois contornos e exibe as regiões da Imagem Diferença Morfológica. Os autores definem as operações morfológicas como dilatação e erosão ponderadas. O tamanho do elemento estruturante (em forma de disco) varia conforme a sua posição no contorno e é proporcional à distância entre os dois contornos nesta posição. O método consiste em, a partir da região definida pelo contorno  $C_1$ , erodir parte da região *I* que não faça limite com a região *III* e dilatar a fronteira da região *III* com a região *II* (usando erosão e dilatação ponderadas, respectivamente). Assim, as regiões *I* e *II*, que definem o contorno  $C_1$ , se transformam gradualmente nas regiões *II* e *III*, que definem o contorno  $C_2$ . Uma vantagem do método, devido ao uso das operações morfológicas ponderadas, é a possibilidade de se controlar a quantidade desejada de *slices* a serem interpolados entre  $C_1$  e  $C_2$ . Porém, segundo os próprios autores, o método não consegue interpolar corretamente quando algum contorno possui forma espiculada.

Uma abordagem diferente é utilizada por Luo e Hancock em [Luo e Hancock 1997]. A transformação de distância, definida neste método, atribui a cada *pixel* de um *slice* o valor da menor distância entre este *pixel* e o contorno da região de interesse (a distância será positiva, se o *pixel* estiver dentro do contorno; negativa, se estiver fora; e zero, se o *pixel* for um ponto do contorno). Assim, considere duas imagens binárias,  $I_1$  e  $I_2$ , a serem interpoladas, correspondentes a dois *slices* adjacentes (veja as Figuras 2.11a-b para um exemplo). Seja  $D_1$  a imagem resultante da transformação de distância aplicada a  $I_1$  (ilustrada na Figura 2.11c). Seja  $D_2$  a imagem obtida pela multiplicação *pixel a pixel*

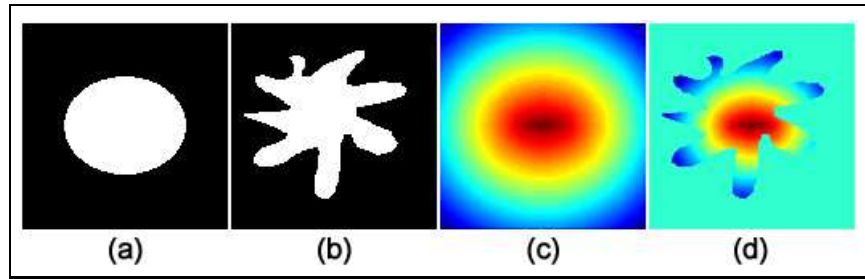


Figura 2.11: a) Imagem binária  $I_1$ . b) Imagem binária  $I_2$ . c) Imagem  $D_1$  resultante da transformação de distância sobre  $I_1$ . d) Imagem  $D_2$  resultante da multiplicação de  $I_2$  por  $D_1$ .

entre  $D_1$  e  $I_2$  ( $I_2$  é utilizada como máscara sobre  $D_1$  para obter  $D_2$ ) e ilustrada na Figura 2.11d. A quantidade de *slices* a serem interpolados entre  $I_1$  e  $I_2$  é determinada pelo maior valor absoluto em  $D_2$ . Deste modo, é realizada, a partir de  $D_2$ , e de maneira iterativa, a erosão morfológica dos pontos de maior intensidade positiva e a dilatação morfológica dos pontos de maior intensidade negativa. A interpolação obtida pelo método é mais suave que o método de [Guo et al. 1995]. Entretanto, há uma desvantagem: o método não permite ao usuário determinar a quantidade desejada de *slices* a serem interpoladas (esta característica pode distorcer o objeto, uma vez que a resolução *intraslice* não se mantém constante para todo o objeto, o que, dependendo da aplicação, pode ser um efeito indesejável). Em outras palavras, considere três *slices*  $I_1$ ,  $I_2$  e  $I_3$ , sendo  $I_1$  adjacente a  $I_2$  e  $I_2$  adjacente a  $I_3$ . A quantidade de *slices* interpolados entre  $I_1$  e  $I_2$  pode ser diferente da quantidade obtida na interpolação entre  $I_2$  e  $I_3$ .

## 2.6 Considerações Finais

Neste capítulo foram brevemente revisados alguns dos conceitos teóricos a serem utilizados neste trabalho. Em cada tópico, foram apresentadas suas principais características, as quais se mostrarão úteis no decorrer dos capítulos seguintes deste trabalho.

O capítulo 5 apresenta dois novos métodos de interpolação desenvolvidos e inspirados nas abordagens descritas acima (veja seção 2.5). O método SIMOL, inspirado na abordagem de Guo et al. [Guo et al. 1995], simplifica a implementação, pois utiliza operações morfológicas clássicas (em contraste com as operações ponderadas definidas em [Guo et al. 1995]) e permite a interpolação correta entre formas espiculadas. Por outro lado, o método SIDITRANS, inspirado no trabalho de Luo e Hancock em [Luo e Hancock 1997], dispensa as operações morfológicas e permite ao usuário especificar a quantidade desejada de *slices* a serem interpolados.

# Capítulo 3

## Descritores de Forma

O contorno de uma região ou objeto 3D é o componente de maior importância na representação e análise de imagens digitais. Diversos trabalhos têm sido desenvolvidos com o objetivo de quantificar propriedades de uma imagem, tais como a irregularidade e a complexidade da forma de uma região ou objeto, a fim de melhorar o processo de classificação ou recuperação de imagens [Chanda e Majumder 2006].

Neste capítulo descrevemos alguns dos trabalhos encontrados nesta área de pesquisa aplicados a formas bidimensionais e a formas tridimensionais.

### 3.1 Métodos para Caracterização de Formas no espaço 2D

De um modo geral, os descritores de forma podem ser divididos em duas grandes categorias: os descritores de forma baseados em contorno e os descritores de forma baseados em região [Zhang e Lu 2002].

#### 3.1.1 Extratores de Forma Baseados em Contorno

Gonzalez e Woods em [Gonzalez e Woods 2008] descrevem diversos descritores clássicos, como perímetro, diâmetro, excentricidade e curvatura de um contorno. O perímetro pode ser definido, de forma aproximada, como comprimento de um contorno obtido pela contagem de seus *pixels*. O diâmetro é obtido calculando-se a maior distância entre dois pontos do contorno. A excentricidade é definida como a razão entre o eixo maior e o eixo menor, onde o eixo maior é compreendido pelo segmento de reta que liga os dois pontos mais distantes em um contorno, enquanto o eixo menor é o segmento de reta perpendicular ao eixo maior. O comprimento do eixo menor é definido pela largura de um retângulo básico (uma caixa passando através dos quatro pontos exteriores da interseção da fronteira com os dois eixos completamente dentro desta fronteira). A curvatura de

um contorno é estabelecida como a taxa de mudança de inclinação e, em geral, obtida a partir de um modelo poligonal do contorno. Estes descritores são simples de calcular e invariantes à escala (exceto perímetro e diâmetro), translação e rotação, porém possuem um baixo poder discriminativo.

Descritores de Fourier são exemplos clássicos para representação de contornos. Zhang e Lu em [Zhang e Lu 2003] utilizam descritores de Fourier aplicados a diferentes assinaturas de formas (funções unidimensionais para representar regiões ou contornos). Em seu trabalho, os autores consideram quatro assinaturas: distância central (distância dos pontos do contorno até o centro de massa da forma); coordenadas complexas (transforma a posição dos pontos do contorno em números complexos); função de curvatura (derivada segunda do contorno ou derivada primeira da tangente do contorno); e função angular cumulativa (representa ângulos da fronteira num intervalo de  $[-\pi; \pi]$  ou  $[0; 2\pi]$ ). Os descritores de Fourier são então constituídos dos coeficientes obtidos pela transformada discreta de Fourier. Estes descritores representam a forma do objeto no domínio da frequência. Descritores com baixas frequências contêm informações de características gerais da forma, ao passo que as altas frequências descrevem detalhes da forma. Resultados mostram que, das quatro assinaturas utilizadas, somente a função de distância central obteve resultados satisfatórios.

Bartolini et al. em [Bartolini e Ciaccia 2005] usam descritores de Fourier para adquirir coeficientes de fase. A informação de fase fornece uma descrição mais precisa da fronteira do objeto, quando comparada ao usar apenas os coeficientes de amplitude de Fourier, mas introduz um deslocamento de fase nos coeficientes de Fourier, quando uma imagem é rotacionada ou tem seu ponto inicial de contorno alterado. Assim, para alcançar a invariância de rotação e de ponto inicial, os autores propõem derivar termos de compensação apropriados e adicioná-los aos coeficientes de fase originais, resultando, assim, em um espectro de fase modificado. Para realizar a busca por similaridade, é introduzida uma distância de similaridade chamada *Time Warping* [Berndt e Clifford 1994]. Entretanto, segundo Marzal et al. em [Marzal et al. 2006], a assinatura resultante é ambígua, de modo que formas bastante diferentes podem ser descritas com uma mesma assinatura e, consequentemente, levando a erros no processo de análise de imagens. Outra desvantagem do método é a dependência de uma métrica de distância específica, a *Time Warp*.

Guliato et al. em [Guliato et al. 2008a, Guliato et al. 2008b] propõem métodos para derivar valores, tais como Índice de Espículos, Índice de Convexidade e Dimensão Fractal a partir de uma assinatura da *Turning Angle Function* (TA) aplicada a um contorno. A TA é uma função que relaciona o comprimento e o ângulo de um segmento de um contorno. A função será crescente, quando aplicada em regiões convexas do contorno, e decrescente em regiões côncavas. A assinatura da *Turning Angle Function* elimina, da TA do contorno, informações irrelevantes para descrever a forma. O método foi empregado em contornos de lesões de mama, a fim de classificá-los como benignos ou malignos, tendo

alcançado boa taxa de precisão nos testes apresentados. A extensão deste método para objetos tridimensionais não é trivial.

Rangayyan e Nguyen em [Rangayyan e Nguyen 2007] utilizam a análise fractal na classificação de lesões de mama. Os autores apresentam quatro métodos para cálculo da dimensão fractal, incluindo o método da régua, aplicado a representações unidimensionais do contorno, e o método *box counting*, usado em representações 2D. Os métodos são aplicados a uma base de dados com 111 contornos de lesões de mama, e os resultados são avaliados em termos de curva ROC. Os autores observaram que o método baseado em análise fractal complementa outros fatores de forma (a combinação da dimensão fractal e a concavidade fracionária obteve melhores resultados quando comparados a outros descritores isoladamente).

Torres e Falcão em [Torres e Falcão 2006] propõem dois descritores de forma baseados em saliências. Eles fornecem um método robusto para encontrar saliências côncavas, que eles utilizam para obter seu descritor *Contour Saliences*. O segundo descritor, *Segment Saliences*, é uma variação do descritor anterior e utiliza valores de saliência de segmentos de contorno em vez de pontos isolados. Os dois descritores propostos são comparados com outros, como *Convex Contour Saliences*, *Curvature Scale Space* e *Beam Angle Statistics*, usando um banco de dados de peixes com 11.000 imagens organizadas em 1.100 categorias. Os resultados reportados pelos autores (avaliados em termos de curvas precisão-revocação e de curvas de separabilidade multiescala) apontam o método *Segment Saliences* como sendo mais efetivo que os métodos *Contour Saliences* e outros três métodos utilizados para comparação (*Convex Counter Saliences*, *Curvature Scale Space* e *Beam Angle Statistics*). Os autores também relatam que a incorporação de saliências côncavas no descritor *Counter Saliences* melhora os resultados deste descritor.

Pedrosa et al. em [Pedrosa et al. 2008] introduzem um novo algoritmo para estimação da similaridade entre duas saliências de contorno, explorando a relação entre um contorno e seu esqueleto. As saliências são obtidas a partir do esqueleto de um contorno usando a técnica denominada *Image Foresting Transform* (IFT). Experimentos realizados num banco de dados com 665 imagens binárias mostram resultados mais eficientes, avaliados no uso de curvas de precisão-revocação, quando comparado com o uso da IFT em conjunto com dimensão fractal multiescala e saliências de contorno proposto em [Torres et al. 2004].

Shen et al. em [Shen et al. 1994] utilizam características de forma, como compacidade, momentos (um novo conjunto definido pelos autores) e descritores de Fourier, na classificação de contornos de calcificações em mamogramas. Os autores obtêm os contornos de calcificações através de uma técnica de crescimento de regiões. Os três fatores de formas são combinados em um vetor de características. A classificação dos contornos em benigno ou maligno é realizada com uso do método dos vizinhos mais próximos e a validação cruzada *leave-one-out*. Os autores utilizaram uma base de dados com 148 contornos, provenientes de 18 biópsias, e relataram ter obtido uma classificação ideal (100%

de acerto) quando combinados os três descritores.

### 3.1.2 Extratores de Forma Baseados em Região

Alguns descritores de forma clássicos, baseados em região, citados por Gonzalez e Woods em [Gonzalez e Woods 2008], são: a área da região (definida pelo número de *pixels* da região); compacidade (obtida pela razão entre o quadrado do perímetro e a área); razão de circularidade (razão entre a área da região e a área de um círculo de mesmo perímetro). Novamente, estes descritores são simples de se calcular, porém possuem baixo poder discriminativo. A invariância à translação, rotação e escala é verificada nestes descritores (exceto a área da região, que é sensível à escala).

Bai e Latecki em [Bai e Latecki 2008] propõem um método para reconhecimento de formas baseado em esqueletos. O método é aplicado a silhuetas de objetos bidimensionais. Embora seja baseado em grafos, não considera a estrutura topológica do grafo, pois, segundo os autores, esqueletos visualmente similares podem ter estruturas topológicas completamente diferentes. O *matching* entre grafos é feito comparando-se os caminhos geodésicos entre pontos extremos do esqueleto, usando a *Optimal Subsequence Bijection* (uma medida de similaridade entre sequências de tamanhos diferentes). Seus experimentos mostram que, apesar de superar o desempenho (em termos de taxa de objetos recuperados corretamente) de outros métodos baseados em grafos, quando é usada a base de imagens Kimia-216 [Sebastian et al. 2004], seus resultados mostram-se inferiores ao usar a base de imagens Kimia-99 [Sebastian et al. 2004]. Segundo os autores, este comportamento se deve à presença de protuberâncias nas formas. O método também não é indicado para formas cujo grafo esqueleto não tenha pontos extremos, como no caso de objetos em forma de toro (também conhecido como *donut*). Outra desvantagem do método é a sua complexidade computacional, que, embora seja menor que a de outros métodos baseados em grafos, é quadrática com relação ao número de nós dos grafos.

Ebrahim et al. em [Ebrahim et al. 2008] utilizam a curva de preenchimento de espaço de Hilbert para representar silhuetas. Eles definem, de modo supervisionado, um *template* chamado *Key Feature Points*, onde são indicados os pontos de menor desvio padrão para uma determinada classe de silhuetas. A similaridade entre duas silhuetas é obtida com o uso da distância euclidiana entre duas imagens, usando os valores indicados pelos *templates*. Testes realizados com as bases de imagens Kimia-99 e Kimia-216 e a base *MPEG-7 CE-Shape 1 Part-B* [Latecki et al. 2000] mostram melhores resultados, em termos de curvas precisão-revocação, quando comparados a métodos baseados em rasterização; porém, seus resultados mostraram-se inferiores com relação a métodos baseados em grafos *shock*. Em [Armstrong et al. 2009] é proposto o uso do PCA para alinhamentos das silhuetas, com o objetivo de alcançar a invariância à rotação no método apresentado por Ebrahim et al. Também é avaliada, no trabalho de Armstrong et al., que o uso da

curva de Hilbert leva a uma maior eficácia dos resultados, quando comparada ao uso das curvas de preenchimento de espaço de Moore e Z-order na representação de silhuetas.

Zhang e Lu em [Zhang e Lu 2002] descrevem um descritor genérico de Fourier (GFD) derivado da transformação bidimensional de Fourier em imagens amostradas por rasterização polar (baseada no sistema de coordenadas polares). Diferentemente de outros métodos baseados em Fourier, o GFD é aplicado a uma região, e não ao contorno da forma. Porém, a aplicação da transformada 2D de Fourier não garante à invariância a translação, rotação e escala. Para isso, os autores realizam uma normalização para as transformações geométricas e aplicam a transformada polar de Fourier, denominada por PFT. Para derivar a PFT, tanto as informações do plano da imagem quanto as informações do plano espectral são convertidas para o espaço polar. O método captura características tanto na direção radial, quanto na direção circular. Segundo os autores, o descritor é genérico, robusto e independente de aplicação. Apesar de mostrar resultados superiores ao uso de um descritor baseado em momentos de Zernike, em termos de teste *bull's eye*, sua precisão cai consideravelmente na presença de distorções como inclinação (*skew*) ou alongamento (*stretching*), de mudança de escala, rotação, ou transformações de perspectiva. Testes realizados com a base *MPEG-7 Region Shape Database CE-2* [Kim e Kim 2000] apresentam uma precisão média para classificação em torno de 80.2%.

Rangayyan et al. em [Rangayyan et al. 1997] propõem uma medida baseada em região da acutância do perfil da borda da imagem que caracteriza a transição na densidade da região de interesse (ROI), ao longo das normais da ROI em cada pixel da fronteira. Os autores utilizam esta medida na discriminação entre tumores mamográficos benignos e malignos. Fatores de forma, como compacidade, descritores de Fourier, momentos e estatísticas de comprimento de corda, também são utilizados para distinguir entre formas circunscritas e espiculadas. A classificação utiliza a abordagem *leave-one-out* em conjunto com a distância de Mahalanobis, obtendo bons resultados da medida de acutância, em uma base de dados composta por 39 imagens obtidas pelo *Mammographic Image Analysis Society* [Suckling et al. 1994] mais 15 outras imagens de caso locais, ao classificar calcificações como malignas ou benignas (classificação usando informações de diagnóstico). A combinação da medida de acutância com os outros fatores de forma não apresentou melhorias significativas na classificação benigno/maligno, mas sim quando classificou como benigno circunscrito, maligno circunscrito, benigno espiculado e maligno espiculado (classificação usando informações de diagnóstico e forma).

Khotanzad e Hong em [Khotanzad e Hong 1990] introduzem os momentos de Zernike para descrição de formas. Momentos de Zernike são obtidos mediante polinômios complexos, que formam um conjunto completo de bases ortogonais definidas no interior de um círculo unitário. Os momentos de Zernike são as projeções de uma função da imagem nestas funções de bases ortogonais. Para calcular os momentos de Zernike de uma imagem, o centro da imagem é tomado como origem e as coordenadas de *pixels* são mapeadas para o

intervalo de um círculo unitário (*pixels* localizados fora desta faixa são ignorados). Para obter características invariantes a rotação, os momentos de Zernike são mapeados em coordenadas polares, mantendo, deste modo, inalterada a magnitude dos momentos para formas rotacionadas. A magnitude é então tomada como característica da imagem. Experimentos realizados por El-ghazal et al. em [El-ghazal et al. 2007], utilizando a base de imagens *MPEG-7 CE-Shape 1 Part-B*, mostram que o uso dos momentos de Zernike como descritores tem baixa precisão média (resultado obtido foi de 66.8%), não distinguindo muito do resultado obtido no uso dos descritores de Fourier (69.1%).

Uma combinação de métodos baseado em contorno e métodos baseado em região pode ser encontrado em [Bai et al. 2010]. Neste trabalho, Bai et al. combinam informações de contorno e região, tanto locais quanto globais, para obter um classificador de forma (denominado CS & SP). Do contorno é calculada uma simplificação poligonal, usando o algoritmo DCE (*Discrete Contour Evolution*) [Latecki e Lakamper 2000], e seus vértices são extraídos. Uma assinatura é gerada para cada contorno. Em seguida, usando a informação de região, é extraído o esqueleto da forma, empregando o método descrito em [Bai et al. 2007]. Para combinar as informações de contorno e esqueleto é utilizada uma média ponderada, na qual o peso de cada informação é obtido por um processo de aprendizagem que visa minimizar o erro de classificação. Segundo os autores, seu método é invariante a translação, rotação e escala e tem baixa sensibilidade a articulações e transformações não-rígidas. Testes usando as bases de Kimia e MPEG-7 mostram que, com informações de contorno e esqueleto combinadas, o resultado é mais efetivo, de um modo geral, do que somente com uso de contorno ou esqueleto (algumas classes de imagens não apresentaram ganhos ao combinar informações de contorno e esqueleto).

## 3.2 Métodos para Caracterização de Formas no Espaço 3D

Muito tem sido pesquisado no sentido de classificar e recuperar objetos em “imagens” tridimensionais com objetos do cotidiano, lesões em imagens de tomografia computadorizada (CT) ou de ressonância magnética (MRI), entre outros. Alguns dos trabalhos relacionados a descritores para objetos tridimensionais são apresentados a seguir.

Pan et al. em [Pan et al. 2011] indicam um descritor de forma denominado Histograma de Poisson. O método utiliza a equação de Poisson para definir uma assinatura de forma tridimensional. Em seguida, é derivado um descritor de forma baseado em histograma, que acumula os valores da assinatura obtida, montando o histograma. O método é robusto com relação a diferentes transformações geométricas, e seu descritor possui uma baixa dimensão, permitindo uma indexação eficiente. Os experimentos realizados utilizam a base de dados McGill [Zhang et al. 2005] de objetos 3D alcançando resultados superiores,

em termos de curvas precisão-revocação, aos descritores *D2 Histogram*, *Shell Histogram* e *Spherical Harmonics Descriptor*. Segundo os autores, o método é bom para recuperar formas com estruturas óbvias, mas outros métodos podem ser melhores para discriminar formas com estruturas não óbvias.

Dutagaci et al. em [Dutagaci et al. 2010] exploram abordagens baseadas em subespaços para a classificação e recuperação de formas tridimensionais. Os autores desenvolveram um esquema de recuperação usando três técnicas populares de subespaços: análise de componentes principais (PCA), análise de componentes independentes (ICA) e fatoração de matriz não negativa (NMF). O método foi aplicado a representações volumétricas de objetos 3D e requer normalização de pose, o que é obtido por meio do CPCA (*Continuous PCA*) [D.V. Vranić 2001]. Os resultados alcançados em experimentos realizados sobre objetos da *Princeton Shape Benchmark* [Shilane et al. 2004], avaliados em termos de valores DCG e curvas de Precisão-Revocação, indicam uma eficiência equivalente aos métodos DSR (combinação de descritores baseados em *buffer* de profundidade, silhuetas e *ray tracing*) [Vranić 2004], DBF (*density-based framework* derivado de funções de densidade de probabilidade) [Akgül et al. 2009, Akgül et al. 2007] e LFD (*LightField Descriptor*) [Chen et al. 2003].

Kazhdan et al. em [Kazhdan et al. 2003] apresentam uma ferramenta matemática denominada *Spherical Harmonic Representation* para transformar descritores dependentes de rotação em descritores invariantes a rotação. O método substitui o alinhamento de objetos para um sistema de coordenadas canônico, em geral obtido pelo uso da Análise de Componentes Principais (PCA). A principal ideia por trás do método é descrever uma função esférica em termos da quantidade de energia em diferentes frequências. Conforme os autores, sua representação melhora o desempenho do *matching* de diversos extratores (mesmo quando comparado a descritores que utilizam uma normalização de rotação) e também reduz a dimensionalidade do descritor, resultando em uma representação mais compacta. O método pode ser visto como uma generalização do descritor de Fourier para o caso de funções esféricas. Uma limitação do método está no fato de somente poder ser aplicado a descritores que podem ser vistos como uma função esférica ou como função de uma grade de *voxels*. Segundo Tangelder e Veltkamp em [Tangelder e Veltkamp 2008], uma desvantagem apresentada pelo método é a sua incapacidade em distinguir entre objetos que diferem apenas por uma rotação ocorrida no seu interior.

Dois objetos tridimensionais são similares, quando eles se parecem similares em todos os ângulos de vista. Partindo-se deste princípio, Chen et al. em [Chen et al. 2003] desenvolveram um extrator de características baseado em geometria, denominado *LightField Descriptor*, utilizando um conjunto de câmeras ao redor de um objeto 3D. As câmeras são posicionadas nos vértices de um dodecaedro regular que envolve o objeto tridimensional. Embora os resultados sejam bons (superando métodos como *Spherical Harmonics*, *MPEG-7 Shape 3D Descriptors* e *MPEG-7 Multiple View Descriptor*), o custo computa-

cional associado é bastante alto, uma vez que a similaridade entre dois objetos, segundo os autores, deve ser calculada para todos os 20 possíveis ângulos de vista de cada uma das 60 possíveis rotações dos dodecaedros de ambos os objetos.

Vranic em [Vranić 2003] define um vetor de características invariante a rotação baseado em esferas concêntricas. O autor translada o centro de massa do objeto para a origem e dispara raios (*ray cast*) da origem a todos os pontos de intersecção com o objeto. Estes pontos são utilizados para definir uma coleção de funções esféricas, descrevendo a distância máxima do centro de massa do objeto em função do ângulo e raio dos pontos de intersecção com o objeto 3D. O procedimento desenvolvido pelo autor alcança a invariância à rotação em características que podem ser vistas como funções em uma esfera. Seus resultados mostram que o método aumenta a efetividade da busca por modelos 3D, em comparação com o uso da técnica PCA (Principal Component Analysis) para alinhamento de objetos; contudo, de acordo com Tangelder em [Tangelder e Veltkamp 2008], o método é instável e ruídos podem deslocar o centro das esferas concêntricas.

Novotni e Klen em [Novotni e Klein 2004] defendem o uso de momentos 3D de Zernike como descritores de características para recuperação de formas no espaço 3D. Seu descritor é, na verdade, uma extensão dos *spherical harmonics*. Os autores aplicaram o método aos objetos da base *Princeton Shape Benchmark* e obtiveram resultados superiores (porém não muito) ao *Spherical Harmonic*, embora a quantidade de coeficientes utilizada seja bem superior (cerca de 5 vezes mais que no método *Spherical Harmonic*). Este método possui as mesmas desvantagem do *Spherical Harmonic*.

Horn em [Horn 1984] expõe um extrator estatístico denominado *Extended Gaussian Image* (EGI). Trata-se de um histograma que mapeia vetores normais de um objeto 3D para uma esfera unitária (também denominada esfera gaussiana). A esfera gaussiana é dividida em células, para que seja possível representar a informação usando um computador. Idealmente, esta discretização deve seguir alguns critérios: todas as células devem ter a mesma área e mesma forma; as células devem ter formas regulares e compactas; a divisão deve ser fina o bastante para se obter uma boa resolução angular; para algumas rotações as células devem ser induzidas a se coincidirem. Segundo o autor, uma cúpula geodésica derivada de um dodecaedro pentakis com 240 células é suficiente para a maioria dos propósitos práticos. Assim, em um objeto formado por uma malha de triângulos, cada triângulo contribui com o *bin* correspondente a sua direção normal e com um peso igual à área do triângulo. Kang e Ikeuchi em [Kang e Ikeuchi 1994] estenderam o método de tal modo que cada *bin* armazene também a distância dos pontos do triângulo à origem. Esta modificação é conhecida como *Complex Extended Gaussian Image* (CEGI). Ambos EGI e CEGI requerem normalização de orientação do objeto. Outra desvantagem é que objetos diferentes podem produzir histogramas iguais. Embora sejam descritores clássicos para representação de forma de objetos tridimensionais, resultados obtidos com a base *Princeton Shape Benchmark* mostram o EGI e o CEGI como ótimas escolhas para diferenciar

objetos naturais de objetos feitos pelo homem. No entanto, para uma diferenciação mais específica, os métodos mostram resultados inferiores a outros comparados em [Shilane et al. 2004], em termos de curva precisão-revocação.

Osada et al. em [Osada et al. 2002] descrevem o uso de distribuições de forma para representar objetos. Estas distribuições são baseadas em propriedades baseadas em medidas de ângulo, área, volume e distância entre pontos aleatórios da superfície de um objeto 3D. O *matching* entre dois objetos é feito comparando-se distribuições de probabilidade. O método é relativamente rápido e robusto com relação à translação, rotação, escala, espelhamento, tesselações, simplificações e degenerações, mas seus autores o descrevem como um método de pré-classificação em um sistema de análise e recuperação baseada em conteúdo. Entre as diversas distribuições de forma descritas em [Osada et al. 2002], a D2, que mede a distância entre pontos aleatórios, foi a que apresentou melhores resultados. Ip et al. em [Ip et al. 2002] estendem a distribuição D2, classificando pares de pontos como IN (se a linha reta que liga os pontos reside completamente no interior do objeto), OUT (se a linha que liga os pontos está completamente fora do objeto) ou MIXED (se a linha que liga os pontos tem parte dentro e parte fora do objeto). A dissimilaridade entre dois objetos passa a ser medida por uma combinação ponderada das distribuições D2, IN, OUT e MIXED. Entretanto, nessa extensão, é necessário que o modelo defina um volume, não podendo ser aplicado a uma malha poligonal (a extensão é apropriada para uso em sistemas CAD). Testes utilizando as bases *Princeton Shape Database*, *MPEG-7 3D model database* [Zaharia e Prêteux 2001], *Utrecht database* [Tangelder e Velkamp 2003] e *Osada database* [Osada et al. 2001] mostram resultados inferiores (em termos das medidas NDCG, *first tier*, *second tier* e *e-measure*) a dez dentre outros onze métodos comparados em [Shilane et al. 2004], confirmando a sugestão dos autores de utilizar o método como um pré-classificador.

### 3.3 Considerações Finais

Neste capítulo foram apresentados alguns dos diversos descritores de forma para os espaços 2D e 3D, incluindo alguns descritores clássicos como os de Fourier, e momentos de Zernike para imagens 2D, e os métodos EGI, CEGI e D2 para objetos tridimensionais. Como pode ser visto, existem várias abordagens possíveis (baseado em características, grafos ou geometria; aplicáveis a contorno/superfície ou a área/grade de *voxels*; espectrais ou espaciais).

Contudo, o que se verifica, no geral, é que um descritor que mostra excelentes resultados para um tipo de aplicação nem sempre será uma boa escolha em outras tarefas (apesar da tentativa do GFD de se tornar um descritor de uso geral). As Tabelas 3.1, 3.2 e 3.3 listam as principais características dos métodos contidos neste capítulo.

Tabela 3.1: Métodos para extração de característica no espaço 2D e suas principais características. A continuar na Tabela 3.2.

| Métodos                                                                                  | Principais características                                                                                   |
|------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|
| Perímetro, diâmetro, excentricidade e curvatura [Gonzalez e Woods 2008]                  | Invariante a transformações geométricas.                                                                     |
|                                                                                          | Simples de se calcular.                                                                                      |
|                                                                                          | Baixo poder discriminativo.                                                                                  |
| Área, compacidade e circularidade [Gonzalez e Woods 2008]                                | Invariante a transformações geométricas.                                                                     |
|                                                                                          | Cálculo simples.                                                                                             |
|                                                                                          | Baixo poder discriminativo.                                                                                  |
| Descritores de Fourier [Zhang e Lu 2003]                                                 | Representa a forma do objeto no domínio da frequência.                                                       |
|                                                                                          | Descreve características gerais ou detalhadas, se as frequências são baixas ou altas, respectivamente.       |
|                                                                                          | Sensível ao tipo de assinatura utilizado.                                                                    |
|                                                                                          | Baixo poder discriminativo.                                                                                  |
| Descritor de fase de Fourier [Bartolini e Ciaccia 2005]                                  | Fornece uma descrição mais precisa do que usar somente coeficientes de amplitude da transformada de Fourier. |
|                                                                                          | Representação ambígua.                                                                                       |
|                                                                                          | Dependente da métrica <i>Time Warp</i> .                                                                     |
| <i>Assinatura da Turning Angle Function</i> [Guliato et al. 2008a, Guliato et al. 2008b] | A função é crescente em regiões convexas e decrescente em regiões côncavas de um contorno.                   |
|                                                                                          | Possível obter diversos índices a partir da função.                                                          |
|                                                                                          | Dependente de um modelo poligonal de contorno.                                                               |
|                                                                                          | Invariante a rotação, translação e escala.                                                                   |
|                                                                                          | Não é facilmente estendido para objetos 3D.                                                                  |
| Análise fractal [Rangayyan e Nguyen 2007]                                                | Aplicado a representações 1D e 2D de contornos.                                                              |
|                                                                                          | Pode ser estendido para representações 3D.                                                                   |
|                                                                                          | Invariante a transformações geométricas.                                                                     |
| <i>Contour Saliences</i> (CS) e <i>Segment Saliences</i> (SS) [Torres e Falcão 2006]     | Utiliza saliências côncavas.                                                                                 |
|                                                                                          | SS não necessita da localização dos pontos salientes ao longo do contorno.                                   |
|                                                                                          | Robusto com relação a rotação e escala.                                                                      |
| [Pedrosa et al. 2008]                                                                    | Invariante a transformações de rotação e escala.                                                             |
|                                                                                          | Explora a relação entre um contorno e seu esqueleto.                                                         |
|                                                                                          | Apresenta um novo algoritmo de <i>matching</i> para saliências.                                              |
| Compacidade, momentos e descritores de Fourier [Shen et al. 1994]                        | Combina os três fatores de forma em um vetor de características.                                             |
|                                                                                          | Invariante a transformações de rotação, translação e escala.                                                 |

Tabela 3.2: Métodos para extração de característica no espaço 2D e suas principais características. Continuação da Tabela 3.1.

| Métodos                                         | Principais características                                                                                                                 |
|-------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| [Bai e Latecki 2008]                            | Método baseado em grafos.                                                                                                                  |
|                                                 | Não utiliza informação da estrutura topológica dos grafos.                                                                                 |
|                                                 | Utiliza a medida de similaridade <i>Optimal Subsequence Bijection</i> .                                                                    |
|                                                 | Não indicado para formas que contenham protuberâncias ou cujo esqueleto não contenha extremos.                                             |
|                                                 | Complexidade computacional quadrática no número de nós.                                                                                    |
| <i>Key Feature Points</i> [Ebrahim et al. 2008] | Utiliza curva de preenchimento de espaço.                                                                                                  |
|                                                 | O método é supervisionado.                                                                                                                 |
|                                                 | Sensível a rotação, translação e escala.                                                                                                   |
| GFD [Zhang e Lu 2002]                           | Imagens são amostradas usando uma rasterização polar.                                                                                      |
|                                                 | A transformada 2D de Fourier é aplicada a uma região de interesse.                                                                         |
|                                                 | Não invariante a translação, rotação e escala.                                                                                             |
|                                                 | O método é independente de aplicação.                                                                                                      |
|                                                 | Sensível a distorções de alongamento, inclinação e perspectiva.                                                                            |
| Momentos de Zernike [Khotanzad e Hong 1990]     | Invariante a rotação de objetos.                                                                                                           |
|                                                 | Requer normalização de translação e escala.                                                                                                |
|                                                 | Baixo poder discriminativo.                                                                                                                |
| CS & SP [Bai et al. 2010]                       | Combina métodos baseados em contorno e baseados em região.                                                                                 |
|                                                 | Utiliza grafos esqueletos.                                                                                                                 |
|                                                 | Invariante a translação, rotação e escala.                                                                                                 |
|                                                 | Necessidade de criar um modelo poligonal do contorno.                                                                                      |
|                                                 | É necessário que as regiões da imagem sejam conectadas.                                                                                    |
| Medidas de acutância [Rangayyan et al. 1997]    | Caracteriza a transição de densidade de uma ROI.                                                                                           |
|                                                 | Complementa o uso de fatores de forma, quando utilizado em classificação de tumores de mama, usando informações de forma e de diagnóstico. |

Tabela 3.3: Métodos para extração de característica no espaço 3D e suas principais características.

| Métodos                                         | Principais características                                                                                                            |
|-------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| Histograma de Poisson [Pan et al. 2011]         | Robusto com relação a transformações geométricas.                                                                                     |
|                                                 | Baixa dimensionalidade do descritor.                                                                                                  |
| [Dutagaci et al. 2010]                          | Utiliza técnicas de subespaço PCA, ICA e NMF.                                                                                         |
|                                                 | Requer normalização de pose.                                                                                                          |
|                                                 | Aplicado a volumes de <i>voxels</i> .                                                                                                 |
| <i>LightField Descriptor</i> [Chen et al. 2003] | O método é aplicado a um conjunto de imagens 2D representando diferentes pontos de vista de um objeto 3D.                             |
|                                                 | Alto custo computacional.                                                                                                             |
| <i>Ray Cast</i> [Vranić 2003]                   | Invariante a rotação.                                                                                                                 |
|                                                 | Aplicável apenas a extratores que podem ser vistos como funções de uma esfera.                                                        |
|                                                 | O método é instável e sensível a ruídos.                                                                                              |
| <i>Spherical Harmonic</i> [Kazhdan et al. 2003] | Descreve funções esféricas em termos de quantidade de energia em diferentes frequências.                                              |
|                                                 | Transforma descritores sensíveis a rotação em descritores invariantes a rotação.                                                      |
|                                                 | Reduz a dimensionalidade do descritor.                                                                                                |
|                                                 | Pode ser visto como uma generalização do descritor de Fourier usando funções esféricas ou como função de uma grade de <i>voxels</i> . |
|                                                 | Não distingue modelos que diferem por uma rotação no interior de um objeto.                                                           |
|                                                 | Aplicável a descritores definidos como uma coleção de funções esféricas ou como uma grade de <i>voxel</i> .                           |
| Momentos 3D de Zernike [Novotni e Klein 2004]   | Pode ser visto como uma extensão do <i>Spherical Harmonic</i> .                                                                       |
| EGI e CEGI [Horn 1984, Kang e Ikeuchi 1994]     | Histograma de vetores normais.                                                                                                        |
|                                                 | Diferencia bem entre objetos naturais e objetos feitos pelo homem.                                                                    |
|                                                 | Requer normalização de orientação do objeto.                                                                                          |
|                                                 | Pode gerar representações ambíguas.                                                                                                   |
|                                                 | Poder discriminativo fraco para distinguir entre classes mais específicas de objetos.                                                 |
| D2 [Osada et al. 2002]                          | Baseado em distribuição de probabilidades.                                                                                            |
|                                                 | Baixo poder discriminativo.                                                                                                           |
| D2 IN OUT MIXED [Ip et al. 2002]                | Extensão do método D2.                                                                                                                |
|                                                 | Método desenvolvido para sistemas CAD.                                                                                                |
|                                                 | Não indicado para malhas poligonais.                                                                                                  |

O método para extração de características de silhueta 2D proposto por Ebrahim et al. em [Ebrahim et al. 2008] utiliza a curva de preenchimento de espaço de Hilbert para percorrer determinada imagem contendo uma silhueta. O trabalho é limitado a silhueta de objetos 2D e não é invariante a rotação ou escala. É exigida uma etapa de treinamento, onde são gerados “*templates*” para cada classe de objetos armazenados na base de dados e utilizados na comparação entre duas silhuetas.

No Capítulo 4 apresentamos uma proposta fundamentada no trabalho de Ebrahim que é aplicada a silhueta de objetos 2D e é estendida para objetos 3D. Nosso método não necessita da criação de algum *template* e é altamente competitivo com outros descritores de forma encontrados na literatura.

## Capítulo 4

# *2Hq Feature*: Um Novo Método Para Descrição de Formas

### 4.1 Introdução

Neste capítulo propomos dois extratores de forma, utilizando a curva de preenchimento de espaço de Hilbert e a transformada discreta de wavelet. O primeiro extrator, denominado *2Hq Feature*, é numa modificação do método proposto por Ebrahim et al. em [Ebrahim et al. 2008], enquanto o segundo consiste numa extensão do *2Hq Feature* para representação de formas de objetos 3D e é denominado *3Hq Feature*. Ambos os métodos propostos neste capítulo são aplicados a silhuetas.

Neste trabalho, silhueta é entendida como uma imagem binária composta de uma forma 2D (região de interesse) preenchida na cor branca com o fundo da imagem preenchido na cor preta. Alguns exemplos de silhueta podem ser observados na Figura 4.1. Similarmente, uma silhueta 3D é, neste trabalho, interpretada como um volume binário de *voxels* contendo uma forma 3D preenchida na cor branca e um fundo (demais *voxels* não pertencentes a forma 3D) preenchido na cor preta.

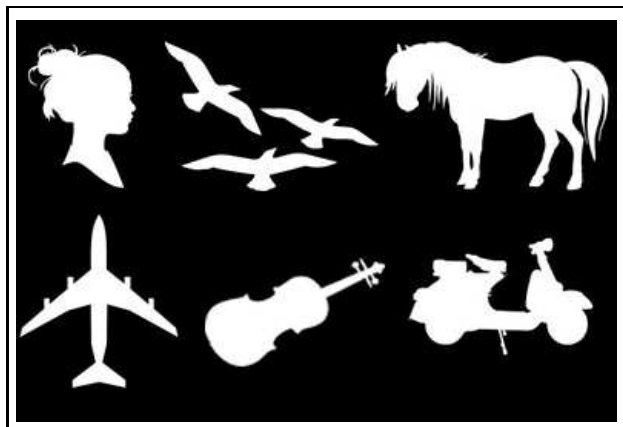


Figura 4.1: Exemplos de imagens de silhuetas.

## 4.2 2Hq Feature: proposta de um novo descritor de formas 2D

Inicialmente, para derivar o *2Hq Feature*, a imagem contendo a silhueta de uma forma é dividida em regiões. Cada região é então preenchida com a curva de preenchimento de espaço de Hilbert. O conjunto dos valores de níveis de cinza, derivados de uma região percorrida pelo caminho determinado pela curva de Hilbert, é obtido em uma representação unidimensional, que será amostrada e normalizada para formar o *2Hq Feature*.

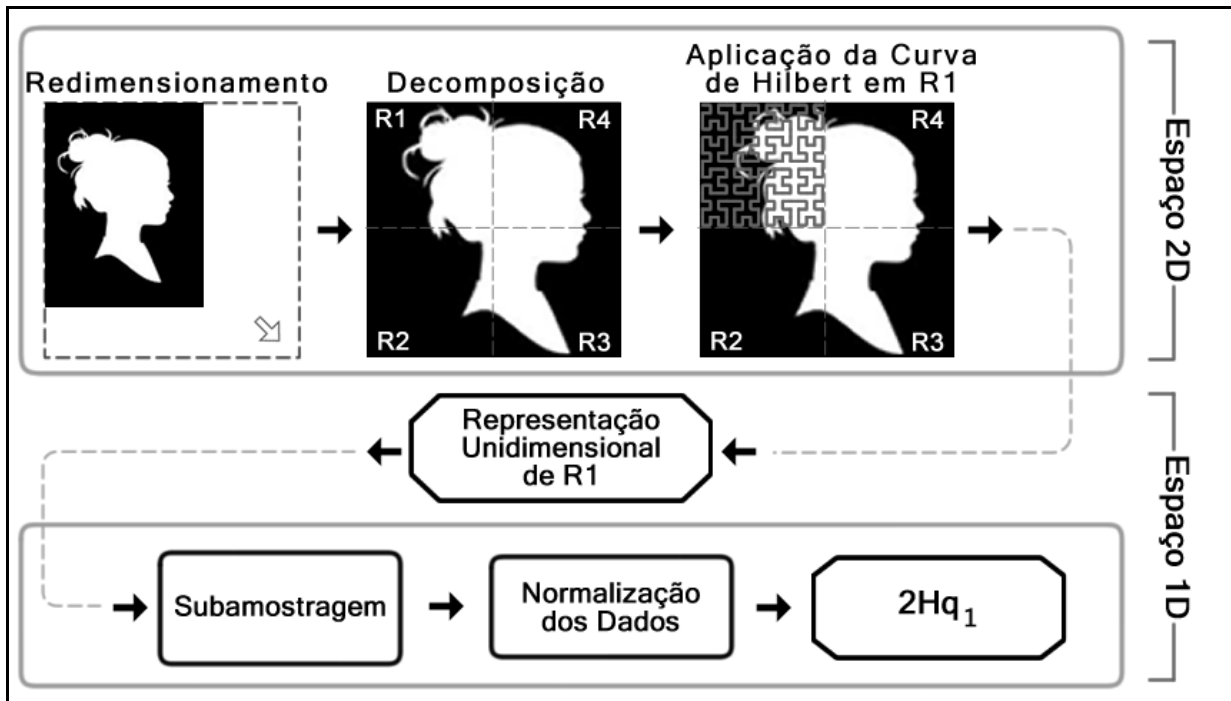


Figura 4.2: Fluxograma de execução do *2Hq Feature* exemplificando o processo empregado para obter o  $2Hq_1$  de uma imagem de silhueta.

Considere  $I$  uma imagem e  $L$  o nível de decomposição da imagem. Considere que, no nível  $L$ ,  $I$  é dividida em  $4^L$  sub-regiões  $R_i, i = 1, \dots, 4^L$ . O *2Hq Feature* obtido para cada sub-região  $R_i$  é aqui denominado por  $2Hq_i$ . O vetor de característica final que descreve  $I$  é composto por todos os  $2Hq_i, i = 1, \dots, 4^L$ . O processo para obter o  $2Hq_i$  para a sub-região  $R_1$  é ilustrado na figura 4.2. Neste exemplo a imagem é dividida em quatro regiões através do processo de decomposição com nível  $L = 1$ . Ao final do processo, o descritor  $2Hq_1$  é obtido para a região  $R_1$ .

As etapas que compõem o processo de obtenção do *2Hq Feature* são detalhadas nas subseções a seguir.

### 4.2.1 Redimensionamento

No espaço 2D, a curva de Hilbert preenche completamente uma imagem digital com dimensões  $2^N \times 2^N$  pixels, onde  $N$  é o nível da curva utilizada (veja a Seção 2.4). Porém,

uma estratégia deve ser tomada com relação às imagens cujas dimensões não se encaixam neste padrão. Uma possível solução consiste em aplicar uma transformação de escala para ampliar uma curva de Hilbert, de modo que a curva e a imagem passem a ter as mesmas dimensões (esta solução é sugerida em [Ebrahim et al. 2008]). Contudo, uma desvantagem desta abordagem é o fato de que nem todos os *pixels* da imagem serão percorridos pela curva, havendo assim uma perda de informação (esta perda pode ser significativa dependendo das dimensões da imagem). Uma segunda desvantagem: em uma coleção de imagens com dimensões arbitrárias, será necessário gerar uma curva para cada imagem, podendo levar a um aumento considerável no tempo necessário para a obtenção do descritor.

Outra solução para o problema descrito acima é redimensionar as imagens, de modo que, em uma base de imagens, todas tenham o mesmo tamanho e com dimensões apropriadas ( $2^N \times 2^N$  *pixels*). Assim, uma única curva de Hilbert pode ser aplicada a todas as imagens da base de dados, sem ocorrer perda de informações. Entretanto, uma simples mudança de escala pode distorcer as imagens, como pode se vê na Figura 4.3. A Figura 4.3a exibe uma imagem original com dimensão arbitrária de  $p \times q$  *pixels*, enquanto a Figura 4.3b mostra a imagem distorcida devido a uma aplicação direta de uma transformação de escala à imagem original (a Figura 4.3b tem dimensões  $2^N \times 2^N$  *pixels*, onde  $N$  é um inteiro positivo).

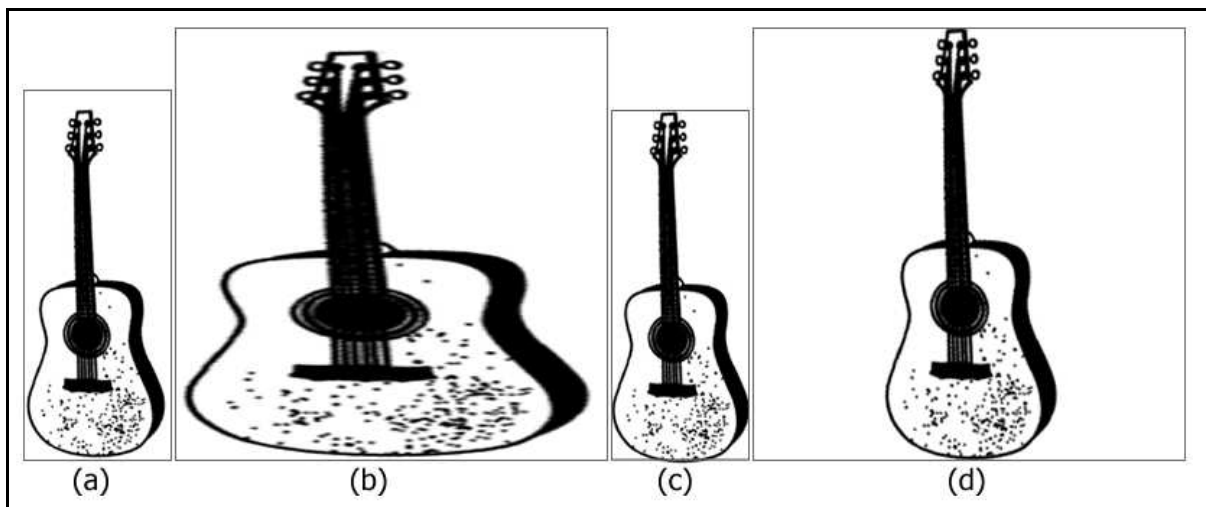


Figura 4.3: Redimensionamento de uma imagem. (a) Imagem original. (b) Imagem distorcida por mudança de escala. (c) Recorte da região de interesse. (d) Imagem ampliada sem distorção.

Para redimensionar as imagens sem causar distorções, um simples procedimento pode ser realizado: primeiramente, recorta-se a região de interesse da imagem (esta operação, no contexto do processamento de imagens, é conhecida como *cropping* e seu resultado ilustrado na Figura 4.3c, agora com dimensões  $r \times s$  *pixels*). Em seguida, o fundo da imagem (*background*) é estendido de modo a se obter uma imagem quadrada com dimensões

$r \times r$  se  $r \geq s$  ou  $s \times s$ , caso contrário. Por fim, realiza-se uma mudança de escala (ampliação ou redução) na imagem para que a mesma tenha as dimensões  $2^N \times 2^N$  (convém observar que a mudança de escala não causa distorções quando a largura e a altura de uma imagem têm a mesma medida). O resultado deste procedimento pode ser visto na Figura 4.3d. Observe que o método de redimensionamento descrito acima faz com que silhuetas semelhantes tendam a possuir dimensões semelhantes ao final do redimensionamento (se duas silhuetas se diferenciam apenas por um fator de escala, igualmente atribuído a cada dimensão, ambas as silhuetas terão exatamente o mesmo tamanho após o processo de redimensionamento, conforme ilustrado na Figura 4.4).

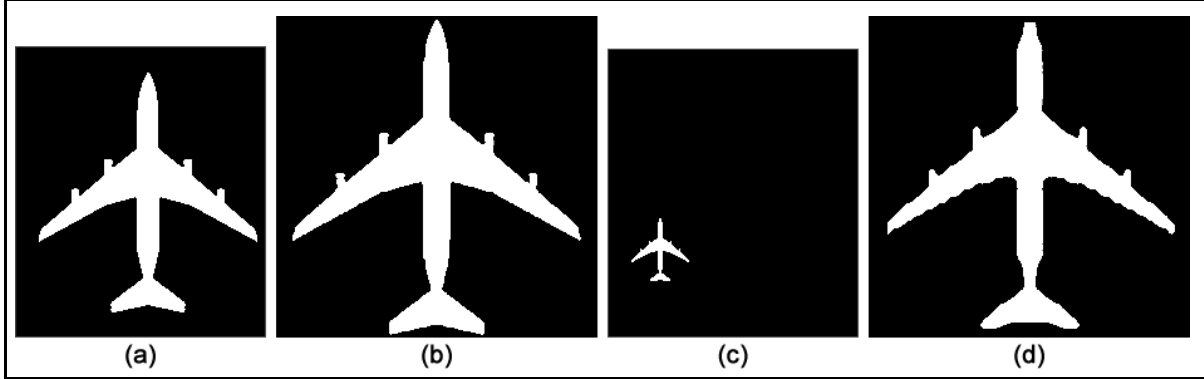


Figura 4.4: Efeito do processo de redimensionamento em objetos de tamanhos diferentes. a) Imagem original. b) Imagem de (a) redimensionada. c) Imagem original. d) Imagem de (c) redimensionada.

Convém lembrar que o uso desta etapa de redimensionamento no processo de obtenção do descritor *2Hq Feature* oferece duas vantagens: a garantia de que não haverá perda de informação quando a imagem for percorrida por uma curva de preenchimento de espaço; a redução no tempo de processamento para obtenção do extrator, uma vez que não é necessário gerar uma curva para cada imagem de uma base de dados.

### 4.2.2 Procedimento de decomposição

Nesta etapa, a imagem é decomposta em sub-regiões. Esta decomposição é realizada dividindo-se a imagem em duas metades ao longo de cada dimensão, resultando em quatro sub-regiões. Cada sub-região pode ser novamente decomposta em outras quatro sub-regiões, o que é aqui denominado de segundo nível de decomposição (o nível de decomposição será representado pela letra  $L$ ). Com  $L = 2$ , soma-se um total de dezesseis sub-regiões (para um nível  $L = n$  haverá um total de  $4^n$  sub-regiões). A Figura 4.5 ilustra o processo descrito acima.

As sub-regiões obtidas nesta etapa serão processadas de maneira independente, nas etapas seguintes, para obter o *2Hq Feature* da silhueta de forma. Convém observar que todas as sub-regiões possuem o mesmo tamanho.

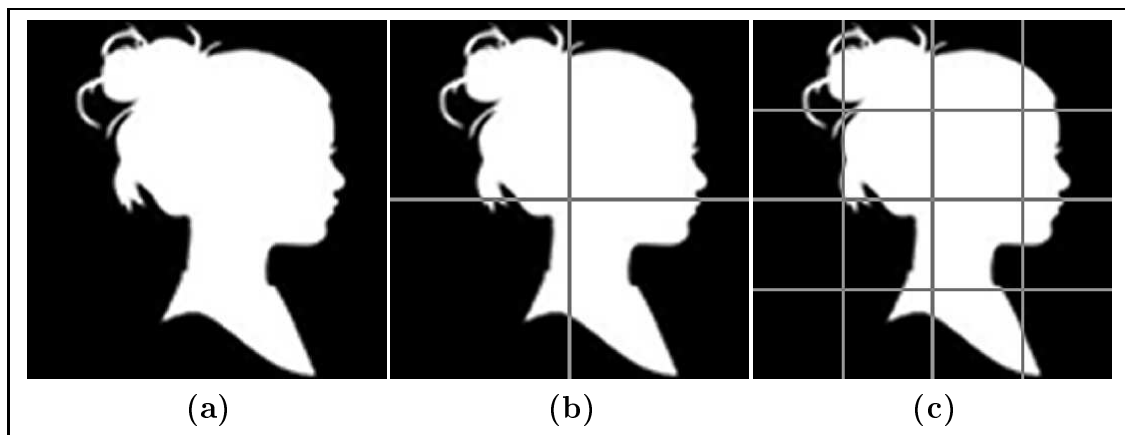


Figura 4.5: Processo de decomposição da imagem. a) Imagem original. b) Imagem decomposta em um nível ( $L = 1$ ). c) Decomposição em dois níveis ( $L = 2$ , cada sub-região do nível anterior é dividida em outras quatro).

### 4.2.3 Aplicação da curva de Hilbert

Em processamento digital de imagens, uma imagem é, em geral, percorrida linha por linha sistematicamente. O uso da curva de preenchimento de espaço de Hilbert oferece uma alternativa à maneira como a imagem é percorrida. A geração da curva é feita conforme descrito na Seção 2.4. Um exemplo numérico de como uma imagem é percorrida pela curva é exibido na Figura 4.6, onde os valores representam intensidades de níveis de cinza. A Figura 4.7 ilustra a aplicação da curva de Hilbert sobre uma região de uma imagem (para uma melhor visualização, foi utilizada uma curva de baixa resolução).

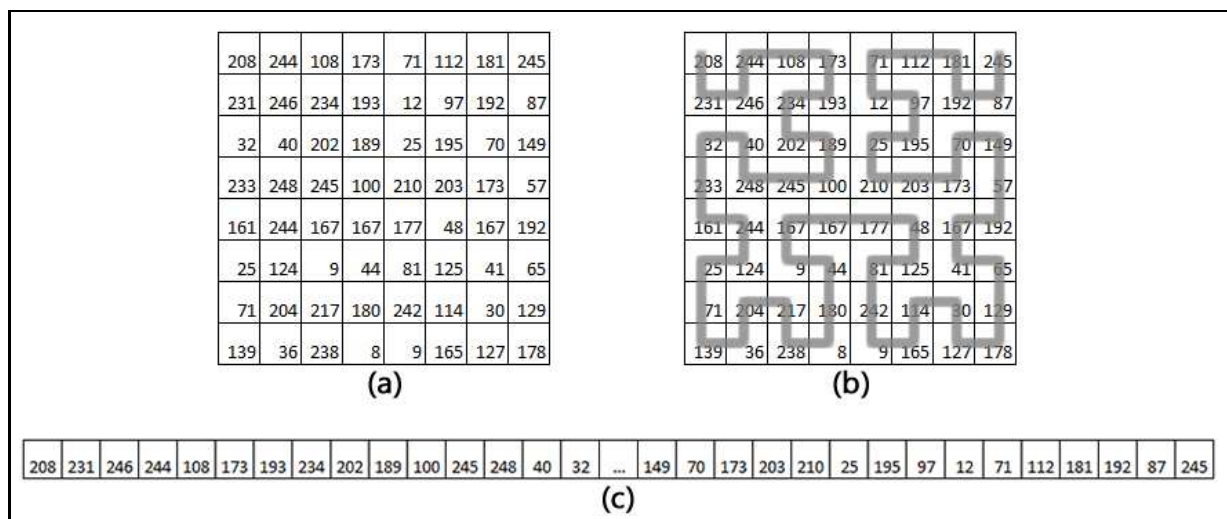


Figura 4.6: Aplicação da curva de Hilbert. a) Intensidade de níveis de cinza de uma imagem hipotética (os valores foram gerados aleatoriamente). b) A sobreposição da curva de Hilbert indica a ordem que os valores são tomados. c) Representação unidimensional dos valores de (a).

O resultado desta etapa é uma representação unidimensional de cada sub-região da imagem contendo os valores de níveis de cinza dos *pixels* percorridos pela curva. Diferentemente do que ocorre, ao se percorrer a imagem linha por linha (estratégia conhecida

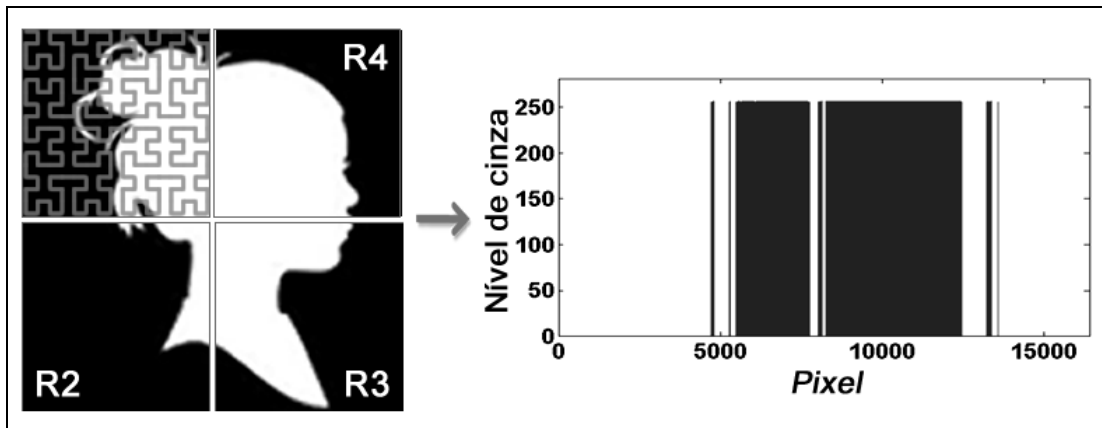


Figura 4.7: Aplicação da curva de Hilbert sobre região  $R_1$  da imagem.

como *raster scan*), a representação usando a curva de Hilbert possui a vantagem de manter, tanto quanto possível, a localização espacial dos *pixels*, ou seja, valores localizados próximos na imagem tendem a se localizar próximos na representação obtida nesta etapa, mantendo, deste modo, uma correlação entre os dados.

Observa-se que, devido à etapa de redimensionamento, em uma coleção de imagens todas as regiões, em todas as imagens desta coleção, irão utilizar uma mesma curva de Hilbert neste processo.

#### 4.2.4 Subamostragem

Como pode ser visto na Figura 4.7, a representação unidimensional de cada região pode chegar a tamanhos consideráveis (no exemplo da imagem, a representação tem tamanho superior a 15000 valores). A fim de reduzir a quantidade de dados, a representação obtida na etapa anterior é subamostrada usando a *transformada discreta wavelet*.

A aplicação da transformada *wavelet* é realizada a partir da representação unidimensional obtida ao final da etapa anterior e pode ser reaplicada sucessivamente aos coeficientes de aproximação resultante de cada aplicação da transformada (os coeficientes de detalhe são descartados). A Figura 4.8 mostra um exemplo de aplicações sucessivas da transformada *wavelet* sobre um conjunto de valores gerados randomicamente. Cada aplicação da transformada é indicada, na figura, pelo bloco DWT e tem o vetor seguinte como resultado de seus coeficientes de aproximação. A Figura 4.9 ilustra a subamostragem aplicada à representação da sub-região  $R_1$  exibida na Figura 4.7. Observe que o tamanho da representação original foi reduzido de cerca de 15 mil valores de *pixels* para aproximadamente 250 valores. Apesar da grande redução, os dados obtidos nesta etapa apresentam um comportamento semelhante aos dados originais.

Esta etapa é aplicada a cada sub-região de uma imagem, até que a quantidade de dados tenha sido reduzida ao tamanho desejado (o tamanho  $n$  final da representação é um parâmetro empírico escolhido pelo usuário). Note que, ao final desta etapa, cada sub-região terá a mesma quantidade de dados (para todas as imagens do banco de dados),

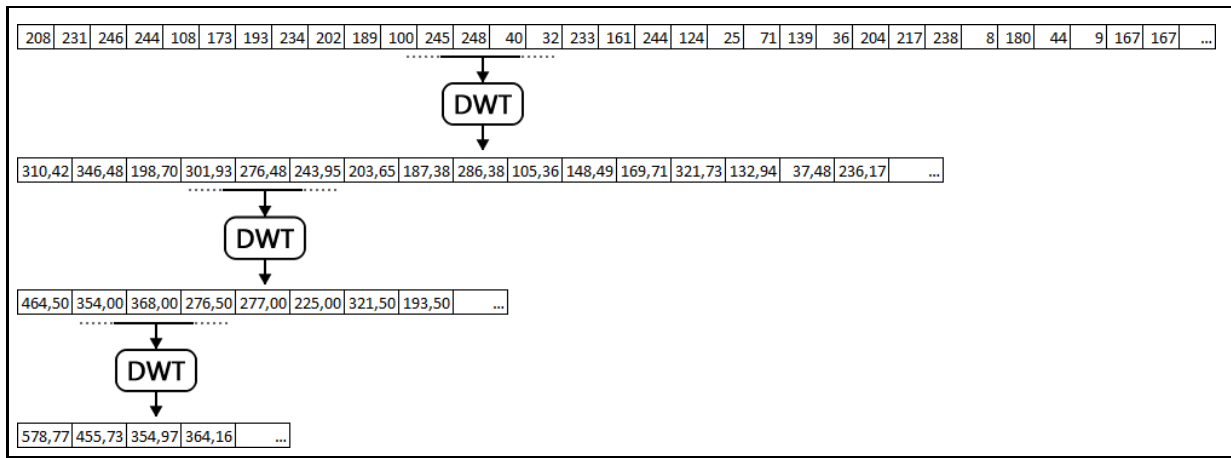


Figura 4.8: Exemplo numérico da aplicação da transformada *wavelet*.

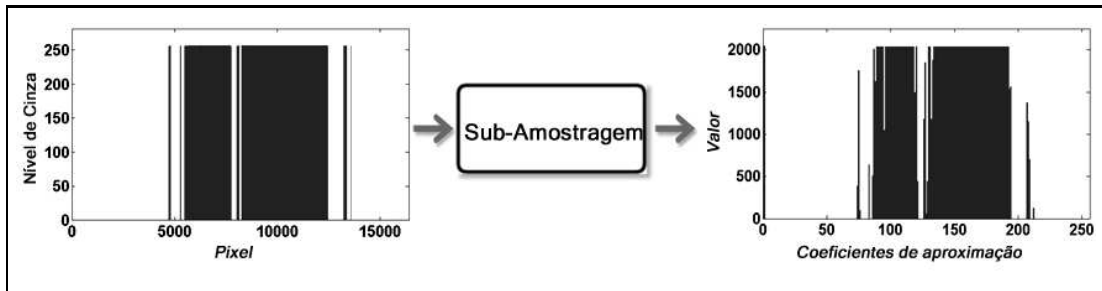


Figura 4.9: Redução da quantidade de dados através da etapa de subamostragem.

não sendo assim necessário um método complementar de subamostragem, como ocorre em [Ebrahim et al. 2008].

#### 4.2.5 Normalização dos dados

Nesta última etapa do processo de obtenção do descritor *2Hq Feature*, os dados obtidos anteriormente são normalizados para valores no intervalo  $[0, 1]$ . Em uma lista de valores  $V$  com  $k$  valores no intervalo  $[0, n]$ , os valores resultantes da normalização  $N$  são dados por:

$$N_i = V_i / MAX(V)$$

onde  $MAX(V)$  se refere ao maior valor pertencente ao conjunto de valores de  $V$ . A Figura 4.10 ilustra um exemplo numérico desta etapa de normalização.

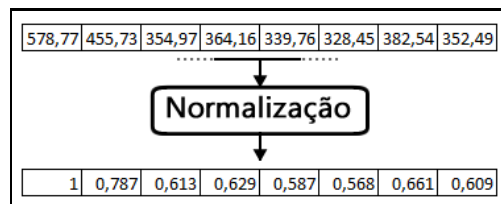


Figura 4.10: Exemplo do processo de normalização aplicado a um conjunto de valores.

A Figura 4.11 exibe o resultado da normalização aplicada à representação de uma sub-região obtida após a decomposição *wavelet* na Figura 4.9. Note que a representação, após a execução da etapa anterior, consistia de valores no intervalo de 0 a aproximadamente 2000.

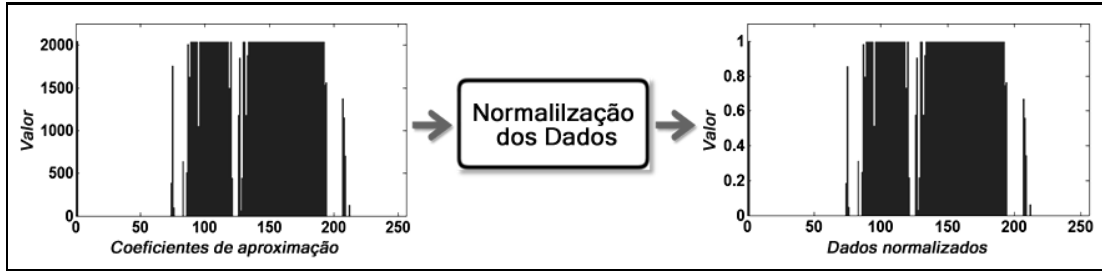


Figura 4.11: Normalização dos dados obtidos até a etapa anterior.

Ao final deste processo, haverá um descritor  $2Hq_i$  para cada região  $R_i, i = 1, 2, \dots, 4^L$ . O  $2Hq$  Feature de uma imagem  $I$  é então constituído pelo conjunto de todos os  $2Hq_i$ , ou seja,  $2HqF = [2Hq_1, 2Hq_2, \dots, 2Hq_{4^L}]$  (assim, a imagem  $I$  passa a ser representada pelo  $2HqF$  obtido). É importante notar que, após a decomposição da imagem, o processo para a obtenção do  $2Hq_i$  é executado de maneira independente para cada sub-região  $R_i$ .

### 4.3 $3Hq$ Feature: Estendendo $2Hq$ para a Terceira Dimensão

O método  $2Hq$  Feature, exposto na seção anterior, pode ser estendido para obter descritores de forma no espaço tridimensional. Para isto, é necessário que as imagens sejam definidas como volumes de *voxels*. Para a extensão 3D do método, alguns ajustes são necessários nas etapas iniciais:

- *Redimensionamento*: a etapa de redimensionamento é similar à descrita na seção anterior. Para objetos tridimensionais o volume de *voxels* deve ter dimensões  $2^N \times 2^N \times 2^N$ , sendo  $N$  um número inteiro positivo. Assim a curva de Hilbert irá visitar todos os *voxels* do volume, não havendo perda de informações e garantindo que os descritores obtidos, para todos os volumes em uma base de dados, tenham o mesmo tamanho. A Figura 4.12 ilustra o funcionamento desta etapa.
- *Decomposição*: a etapa de decomposição irá dividir o volume de *voxels* em oito regiões (subvolumes com dimensões iguais). No espaço 3D, um nível de decomposição  $L = n$  irá produzir  $8^n$  sub-regiões. A Figura 4.13 ilustra a decomposição em um nível de um objeto 3D.
- *Aplicação da curva de Hilbert*: nesta etapa o procedimento é igual ao realizado no caso 2D, exceto que, neste caso, é utilizada a curva tridimensional de preenchimento

de espaço de Hilbert (veja Seção 2.4). A curva de Hilbert é aplicada a cada sub-volume.

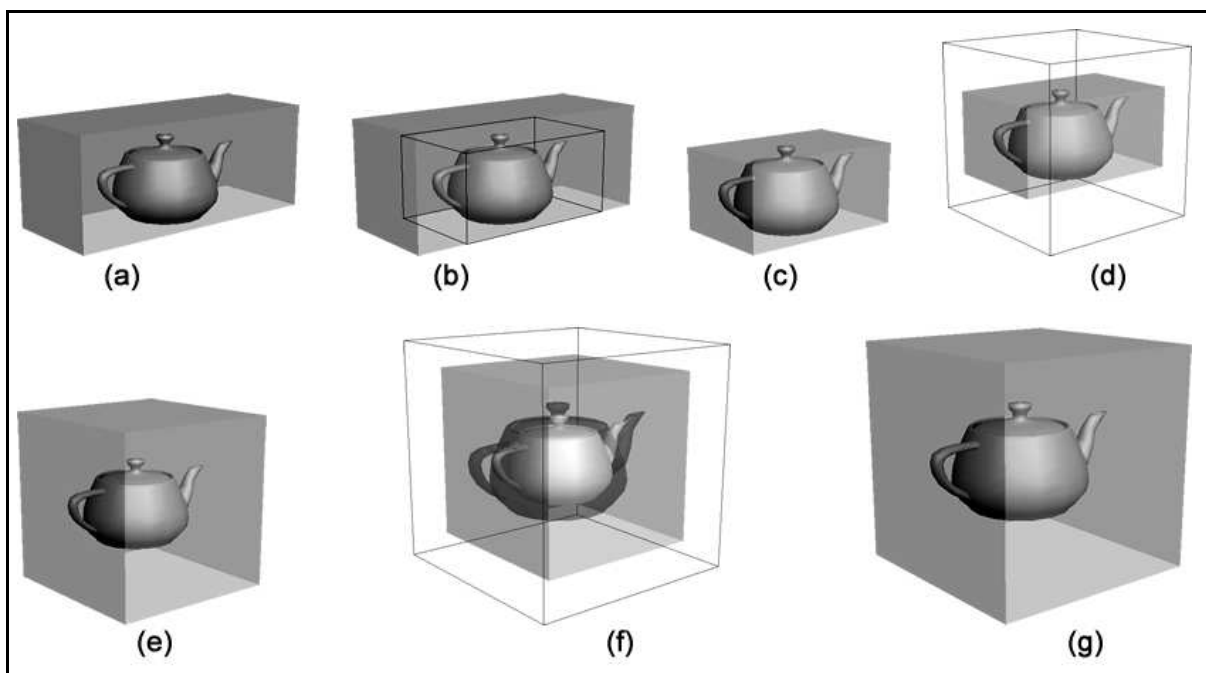


Figura 4.12: Etapa de redimensionamento aplicada a um objeto tridimensional. a) Volume original com dimensões  $225 \times 175 \times 190$ . b) Ilustra a redução do volume após a operação de *cropping*. c) Volume obtido após a operação de *cropping*. As dimensões do volume agora são  $180 \times 128 \times 128$ . d) Indica a expansão da região do *background* do volume. e) Volume com o *background* estendido. As dimensões do volume passam a ser  $180 \times 180 \times 180$ . f) Indica o tamanho final obtido após a operação de mudança de escala. Tanto o volume quanto o objeto serão afetados. g) Volume final com dimensões  $256 \times 256 \times 256$ .

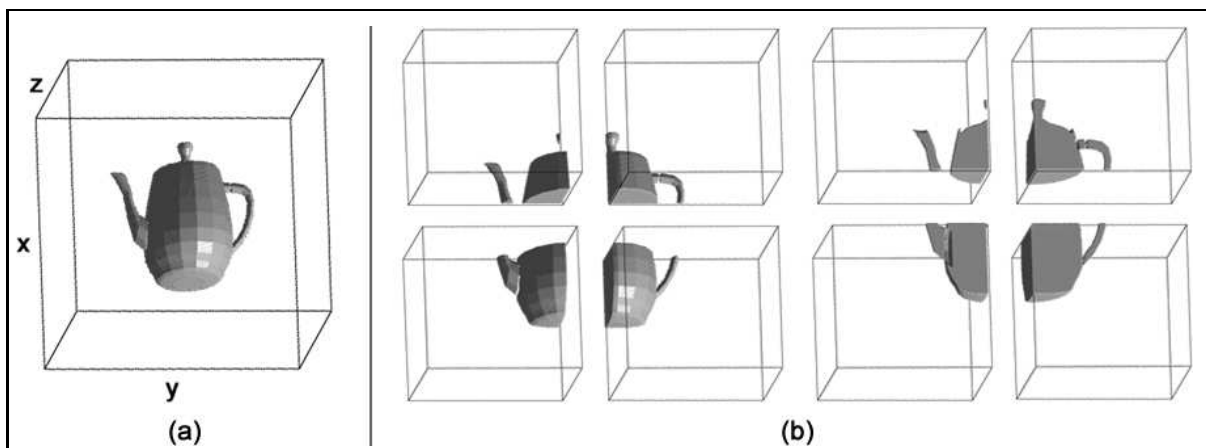


Figura 4.13: Processo de decomposição de um objeto tridimensional. a) Volume original com dimensões  $256 \times 256 \times 256$ . b) Volume decomposto. Cada subvolume tem dimensões  $128 \times 128 \times 128$ .

O resultado da aplicação da curva 3D de Hilbert será uma representação unidimensional da região (assim como ocorre no processo de obtenção do *2Hq Feature*). Portanto,

a partir deste ponto, as demais etapas (subamostragem e normalização) são realizadas de maneira idêntica ao que foi descrito na seção anterior. Ao final do processo, cada sub-volume obtido na etapa de decomposição será descrito por um  $3Hq_i$ . O  $3Hq$  Feature que descreve todo o volume de *voxels* é formado pelo conjunto dos  $3Hq_i$  obtidos para cada sub-volume, ou seja,  $3HqF = [3Hq_1, 3Hq_2, \dots, 3Hq_{8^L}]$ .

## 4.4 Considerações Finais

Neste capítulo foi detalhado o processo de obtenção do descritor  $2Hq$  Feature e de sua extensão  $3Hq$  Feature para o espaço 3D. Como pode ser visto, a extensão do método aplicada a imagens volumétricas é bastante simples e é obtida com poucas modificações.

A complexidade de tempo do método de obtenção do  $2Hq$  Feature é linear com relação ao número de *pixels* (ou *voxels* no caso do  $3Hq$  Feature), pois cada uma de suas etapas é executada em tempo linear exceto a etapa de decomposição, que é executada em tempo constante. A complexidade de espaço também é linear, uma vez que não é necessário manter os dados de todas as etapas (somente da etapa atual), e a quantidade de memória necessária diminui ao longo do processo de obtenção do descritor.

Algumas características do método proposto podem ser apontadas como vantajosas com relação a outros métodos. A divisão da imagem em sub-regiões permite que cada sub-região seja processada de forma independente, possibilitando uma paralelização do processo de obtenção do descritor proposto. A representação das sub-regiões é unidimensional e compacta (a redução da quantidade de dados beneficia principalmente aplicações no espaço 3D, em que imagens volumétricas ocupam uma grande quantidade de memória).

Nos Capítulos 6 e 7 serão descritos experimentos realizados com os descritores  $2Hq$  Feature e  $3Hq$  Feature, e os resultados obtidos serão comparados com outros métodos, a fim de se avaliar o desempenho dos métodos propostos neste capítulo.

# Capítulo 5

## WESS: *Framework* para Reconstrução de Objetos 3D

Neste capítulo é apresentado o WESS, um *framework* desenvolvido para reconstrução de imagens volumétricas formadas de um conjunto de *slices*. O objetivo deste *framework* é proporcionar uma solução simples para segmentar as imagens tridimensionais e obter o volume de *voxels* que represente a forma tridimensional da região de interesse segmentada, permitindo assim a aplicação do *3Hq Feature*, apresentado no Capítulo 4.

### 5.1 Ambiente Para Segmentação de Imagens Tridimensionais Baseadas em *Slices*

Neste trabalho, foi desenvolvido um ambiente para segmentação manual de contornos 2D em imagens volumétricas baseadas em conjuntos de *slices*. O aplicativo, denominado como EMISS (*Environment for Manual Image Segmentation via Slices*), gerencia os *slices* de um objeto tridimensional e permite a um usuário traçar o contorno de regiões de interesse (ROIs) em imagens 2D.

O aplicativo gerencia duas janelas: uma janela principal, contendo os *slices* que compõem uma estrutura tridimensional; e uma janela de edição, que permite o usuário segmentar manualmente um determinado *slice*. A Figura 5.1 exibe a janela principal com um conjunto de miniaturas dos *slices* de uma ressonância magnética de mama. Nesta janela o usuário pode carregar imagens em formatos BMP, DICOM ou GE (formato utilizado em equipamentos da *General Electric*®). Os *slices* são automaticamente organizados em uma grade de miniaturas, e o usuário pode alterar o tamanho e a quantidade de miniaturas exibidas por linha da grade.

Para abrir a janela de edição, o usuário deverá selecionar, usando um dispositivo apontador, alguma das miniaturas exibidas na janela principal. A Figura 5.2 mostra a janela de edição com a imagem de um *slice* carregado. Nesta janela é permitido desenhar

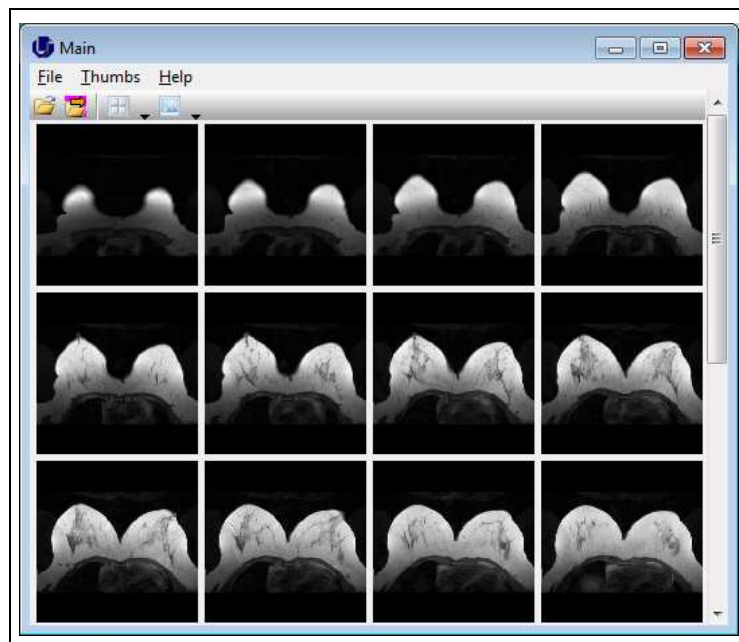


Figura 5.1: Janela principal do ambiente EMISS.

manualmente um contorno utilizando um dispositivo apontador (a Figura 5.2 exibe um contorno em vermelho, deixado como exemplo).



Figura 5.2: Janela de edição para segmentação de um *slice* no ambiente EMISS.

Usando um *mouse* de computador, o usuário, ao clicar com o botão auxiliar (geralmente o botão direito) do dispositivo, torna disponível uma barra de ferramentas, como ilustrado na Figura 5.3. Dentre as diversas opções disponíveis, o usuário pode alterar o tamanho de exibição da imagem, alterar a cor do contorno, fechar automaticamente um contorno aberto, salvar ou carregar um contorno salvo e alterar o contraste da imagem, além de selecionar opções para desenhar ou apagar um contorno. Os contornos são salvos em dois formatos: bitmap, contendo o contorno criado pelo usuário (a imagem é binária e, portanto, não contém informação de cores); e BND, um arquivo de texto contendo,

na primeira linha, as dimensões da imagem e a quantidade de pontos do contorno e, nas linhas subsequentes, todas as coordenadas do contorno listadas no sentido anti-horário (neste formato o contorno deve ser fechado, e cada arquivo BND deve conter um único contorno).

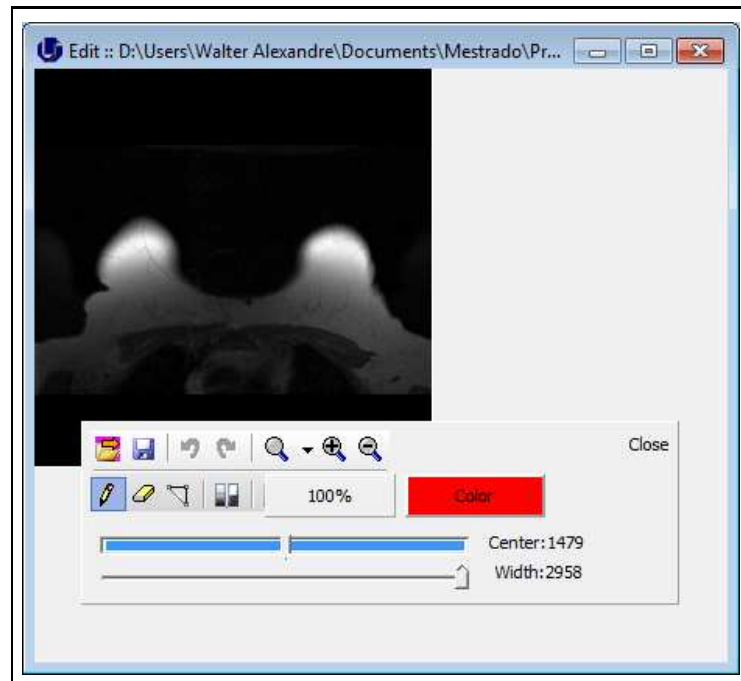


Figura 5.3: Menu de opções da janela de edição do ambiente EMISS.

A partir da coleção de contornos segmentados usando o ambiente EMISS, é possível reconstruir, para o espaço 3D, a ROI via interpolação de contornos. Para isso, são propostos neste trabalho dois métodos de interpolação de contornos, que são apresentados na próxima seção. Uma vez tendo o volume de *voxels* que representa a região de interesse, o descritor *3Hq Feature*, apresentado no capítulo anterior, pode ser utilizado para caracterizar a forma da ROI.

## 5.2 Métodos Desenvolvidos Para Interpolação de *Slices*

Imagens tridimensionais podem ser visualizadas a partir de um conjunto de *slices*. Para transformar esse conjunto de imagens 2D de seções transversais de um objeto 3D em um volume de *voxels*, é necessária a aplicação de um método de reconstrução tridimensional.

Nesta seção são propostos dois métodos para interpolação de contornos. O primeiro deles, baseado em operações de morfologia matemática e lógica, é inspirado no método de Guo et al. introduzido em [Guo et al. 1995] (veja seção 2.5). O segundo método de interpolação proposto neste capítulo tem por fundamento o trabalho de Luo e Hancock apresentado em [Luo e Hancock 1997] e utiliza a transformação de distância para obter os *slices* interpolados.

|    |                                                           |
|----|-----------------------------------------------------------|
|    | <b>Entrada:</b> Imagens $I_1$ e $I_2$                     |
|    | <b>Saída:</b> Conjunto $SI$ de <i>slices</i> interpolados |
| 1  | <b>Início</b>                                             |
| 2  | $SI_1 \leftarrow I_1$                                     |
| 3  | $DL \leftarrow I_2$                                       |
| 4  | $EL \leftarrow I_1 \text{ AND } I_2$                      |
| 5  | $i \leftarrow 2$                                          |
| 6  | <b>Loop</b>                                               |
| 7  | $a \leftarrow \text{Dilata}(SI_{i-1}) \text{ AND } DL$    |
| 8  | $b \leftarrow \text{Erode}(SI_{i-1}) \text{ OR } EL$      |
| 9  | $c \leftarrow a \text{ OR } b$                            |
| 10 | <b>Se</b> $c = SI_{i-1}$ <b>Então</b>                     |
| 11 | Finaliza loop                                             |
| 12 | <b>Senão</b>                                              |
| 13 | $S_i \leftarrow c$                                        |
| 14 | $i \leftarrow i + 1$                                      |
| 15 | <b>Fim</b>                                                |
| 16 | <b>Fim</b>                                                |
| 17 | <b>Retorne</b> $SI$                                       |
| 18 | <b>Fim</b>                                                |

**Algoritmo 5.1:** Algoritmo SIMOL para interpolação de *slices*.

### 5.2.1 Interpolação Baseada em Operações Morfológicas e Lógicas

Este primeiro método de interpolação de *slices*, denominado SIMOL (*Slice Interpolation by MORphology and LOGic operations*), utiliza operações clássicas de morfologia matemática (dilatação e erosão), em contraste com o método, no qual se baseia, de Guo et al. (veja a Seção 2.5), em que definem operações de dilatação ponderada e erosão ponderada.

Para realizar a interpolação entre dois *slices*,  $I_1$  e  $I_2$  são representados como imagens binárias, ou seja, com o fundo da imagem preenchido na cor preta, enquanto a região de interesse é preenchida na cor branca. A partir das imagens  $I_1$  e  $I_2$ , o método SIMOL, descrito no Algoritmo 5.1, pode ser executado.

No Algoritmo 5.1, a variável  $DL$  é utilizada para limitar a dilatação, enquanto  $EL$  serve para limitar a erosão. O laço que compreende as linhas 6 a 16 será executado, sempre que for possível obter um *slice* diferente do anterior. No final do algoritmo, a variável  $SI$  irá conter os *slices* obtidos pelo método (incluindo os originais utilizados como parâmetros de entrada do algoritmo). A Figura 5.4 mostra os *slices* obtidos com o uso do Algoritmo SIMOL (os *slices* de número 1 e 32 na figura são as regiões de interesse  $ROI_1$  e o  $ROI_2$ , respectivamente). O volume de *voxels* se consegue pela concatenação dos *slices* obtidos no processo.

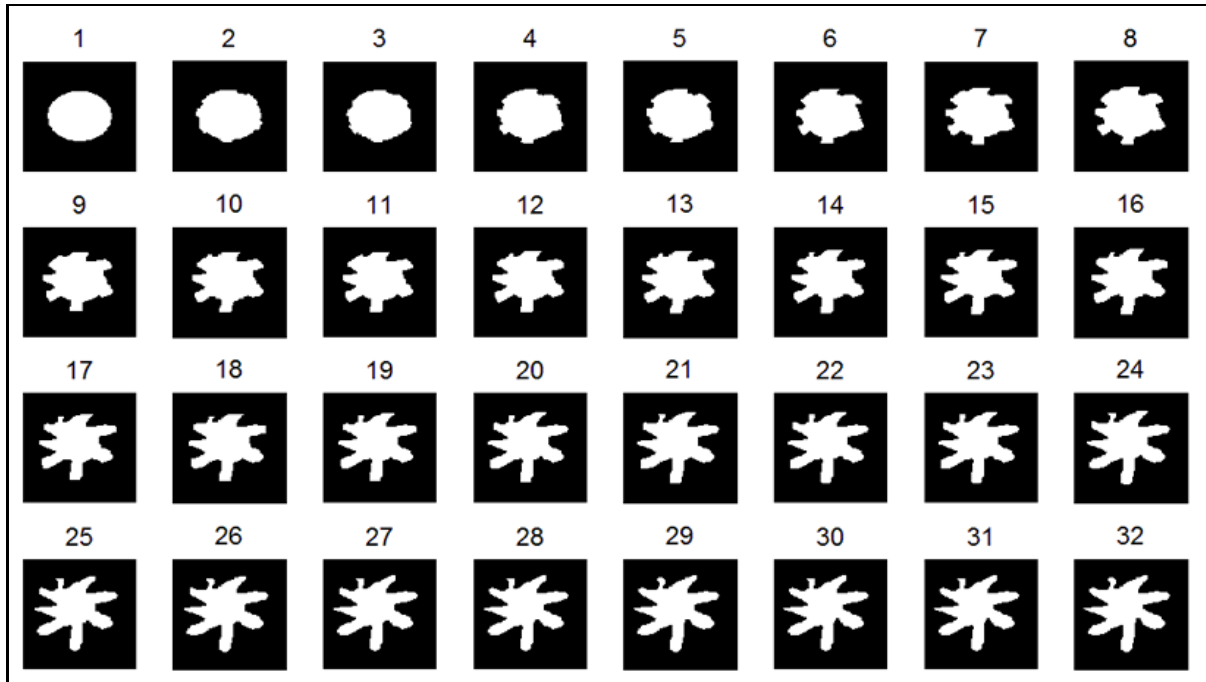


Figura 5.4: *Slices* obtidos usando o algoritmo de interpolação SIMOL.

**Entrada:**  $I$ ,  $Contour$

**Saída:** Imagem transformação de distância (DTI)

```

1  Função DT(I, Contour)
2      Para Cada pixel  $p(x, y) \in I$  Faça
3          // verificando se  $p(x, y)$  pertence à região de interesse
4          Se  $p(x, y) == white$  Então
5              |  $DTI(x, y) \leftarrow$  menor distância entre  $p(x, y)$  e os pontos de Contour
6          Fim
7          Senão
8              |  $DTI(x, y) \leftarrow -1 * (\text{menor distância entre } p(x, y) \text{ e os pontos de Contour})$ 
9          Fim
10     Retorne  $DTI$ 
11 Fim

```

Algoritmo 5.2: Algoritmo para cálculo da transformação de distância.

### 5.2.2 Interpolação Usando Transformação de Distância

O método proposto nesta seção é denominado SIDITRANS (*Slice Interpolation by DIstance TRANSform*) e é baseado no método de Luo e Hancock, descrito na Seção 2.5. Considere dois *slices* adjacentes,  $I_1$  e  $I_2$ , representados como imagens binárias. De posse da lista de pontos de contorno das regiões de interesse  $Contour_1$  e  $Contour_2$ , é obtida a imagem de transformação de distância  $DTI$  para cada *slice*, conforme o procedimento descrito no Algoritmo 5.2. Exemplos de contornos das regiões de interesse e as imagens obtidas pela transformação de distância são ilustrados na Figura 5.5.

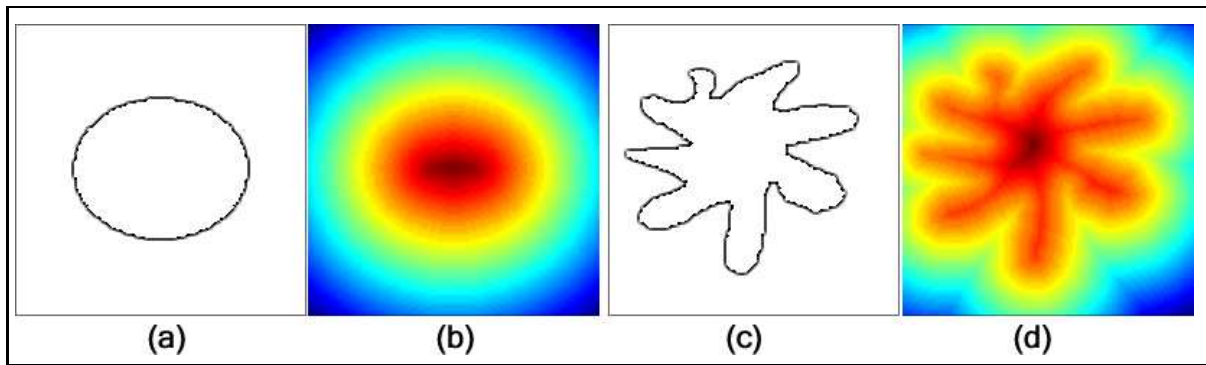


Figura 5.5: Contornos e suas imagens de transformação de distância. a) Contorno  $Contour_1$  de  $I_1$ . b)  $DTI_1$  obtida pela transformação de distância aplicada a  $I_1$ . c) Contorno  $Contour_2$  de  $I_2$ . d)  $DTI_2$  obtida pela transformação de distância aplicada a  $I_2$ .

Em seguida, é realizada uma interpolação numérica entre  $DTI_1$  e  $DTI_2$ , de acordo com a seguinte equação:

$$SI_i(x, y) = \frac{[(n + 1 - i) * DTI_1(x, y)] + [i * DTI_2(x, y)]}{n + 1}, i = 1, \dots, n \quad (5.1)$$

onde  $SI_i$  é o  $i$ -ésimo *slice* interpolado e  $n$  é a quantidade desejada de *slices* a serem interpolados entre  $I_1$  e  $I_2$ . Esta equação corresponde à interpolação linear (alternativamente, outro tipo de interpolação numérica pode ser utilizado, tal como a *spline* ou interpolação de Hermite, por exemplo).

Um exemplo numérico desta equação pode ser visualizado na Figura 5.6. Neste exemplo, a primeira e a última matriz foram criadas com valores aleatórios, enquanto as demais foram obtidas por interpolação usando a Equação 5.1.

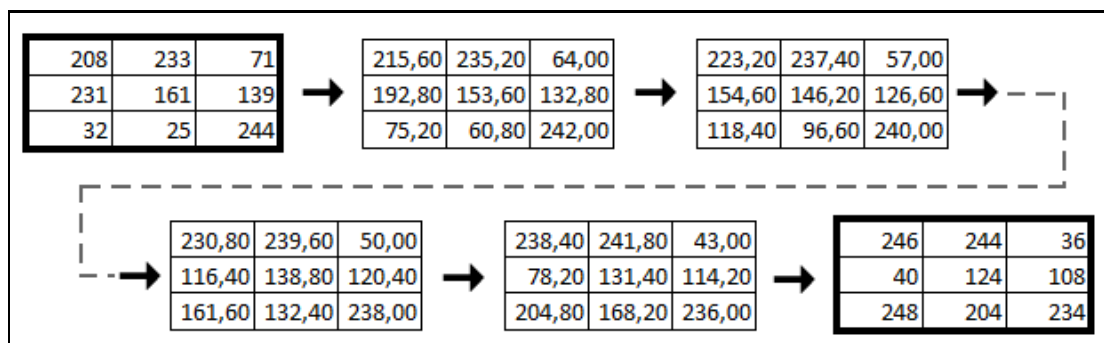


Figura 5.6: Exemplo de aplicação da Equação 5.1. A equação é aplicada à primeira e última matriz para obter as quatro demais (interpoladas).

Por fim, os *slices* obtidos passam por um processo de binarização, onde os *pixels* com valor negativo se tornam preto (fundo da imagem) e os demais se tornam branco (região de interesse). O volume de *voxels* é obtido pela concatenação dos *slices* obtidos após a binarização. O Algoritmo 5.3 ilustra o procedimento executado pelo método de interpolação SIDITRANS (o conjunto de *slices* retornado pelo procedimento inclui os dois *slices* originais passados como parâmetro de entrada do algoritmo).

|                                                                              |
|------------------------------------------------------------------------------|
| <b>Entrada:</b> $I_1, I_2, Contour_1, Contour_2, n$                          |
| <b>Saída:</b> Conjunto $SI$ de <i>slices</i> interpolados                    |
| <b>1 Início</b>                                                              |
| <b>2</b> $DTI_1 \leftarrow DT(I_1, Contour_1)$                               |
| <b>3</b> $DTI_2 \leftarrow DT(I_2, Contour_2)$                               |
| <b>4</b> $SI_1 \leftarrow I_1$                                               |
| <b>5</b> <b>Para</b> $i \leftarrow 2$ <b>to</b> $n + 1$ <b>Faça</b>          |
| <b>6</b> $SI_i \leftarrow ((n + 2 - i) * DTI_1 + (i - 1) * DTI_2) / (n + 1)$ |
| // procedimento de binarização                                               |
| <b>7</b> <b>Para</b> Cada pixel $p(x, y) \in SI_i$ <b>Faça</b>               |
| <b>8</b> <b>if</b> $p(x, y) \geq 0$ <b>then</b>                              |
| <b>9</b> $p(x, y) \leftarrow 1$                                              |
| <b>10</b> <b>end</b>                                                         |
| <b>11</b> <b>else</b>                                                        |
| <b>12</b> $p(x, y) \leftarrow 0$                                             |
| <b>13</b> <b>end</b>                                                         |
| <b>14</b> <b>Fim</b>                                                         |
| <b>15</b> <b>Fim</b>                                                         |
| <b>16</b> $SI_{n+2} \leftarrow ROI_2$                                        |
| <b>17</b> <b>Retorne</b> $SI$                                                |
| <b>18 Fim</b>                                                                |

**Algoritmo 5.3:** Algoritmo SIDITRANS para interpolação de *slices*.

A Figura 5.7 exibe o resultado da equação 5.1 aplicada às imagens 5.5b e 5.5d (a equação foi aplicada com  $n = 32$  para obter as imagens de 2 a 31, de modo que a primeira e a última correspondem a  $DTI_1$  e  $DTI_2$ ). Na Figura 5.8 é mostrado o conjunto de *slices* obtidos ao final do Algoritmo 5.3 (o primeiro e o último correspondem aos originais fornecidos por um usuário, enquanto os demais foram obtidos pelo processo de interpolação).

### 5.2.3 Discussão de Resultados

Para ilustrar o resultado de ambos os métodos de interpolação propostos neste capítulo, dois experimentos foram realizados. No primeiro, o objetivo é verificar a capacidade do método em interpolar quando alguma das formas é espiculada (uma limitação do método de Guo et al., descrito em [Guo et al. 1995], é a sua inabilidade em interpolar este tipo de forma). Neste experimento os *slices* originais utilizados têm dimensões  $128 \times 128$ . As Figuras 5.9 e 5.10 mostram os *slices* interpolados obtidos com os métodos SIMOL e SIDITRANS, respectivamente (o primeiro e último *slices* nas figuras são os originais fornecidos como parâmetro para os algoritmos). As Figuras 5.11a-c exibem, sob três pontos de vista diferentes, o volume de *voxels* obtido pela concatenação dos *slices* na Figura

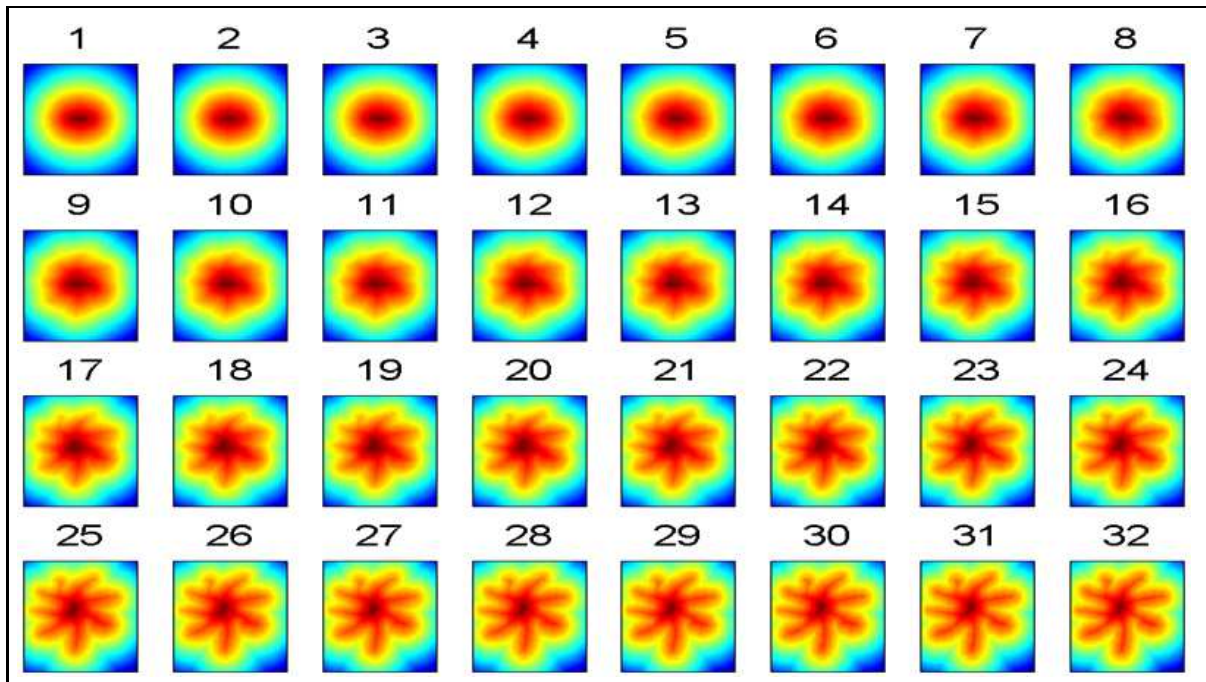


Figura 5.7: Resultado da equação 5.1 aplicada às Imagens 5.5b e 5.5d.

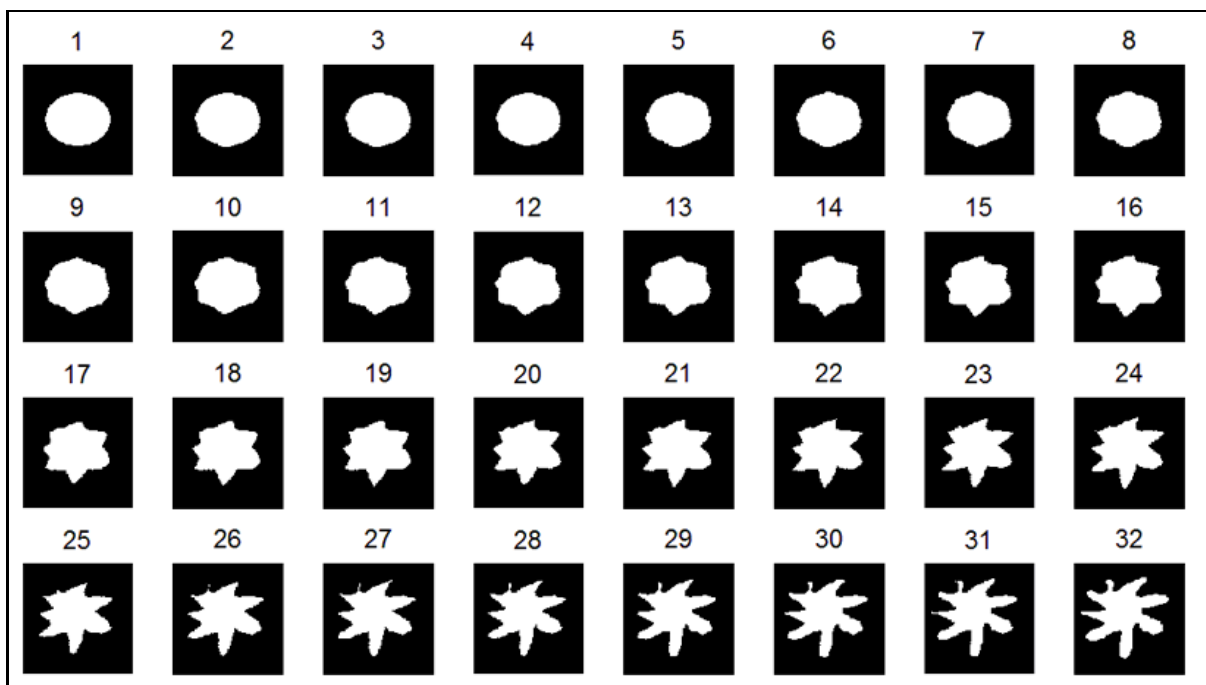


Figura 5.8: *Slices* obtidos pelo algoritmo de interpolação SIDITRANS.

5.9. Analogamente, as Figuras 5.11d-f mostram, também em três ângulos diferentes de visualização, o volume de *voxels* obtido com os *slices* da Figura 5.10.

No segundo experimento, foi verificado o quanto a interpolação se aproxima do volume real. Neste experimento foram utilizados dois pares de *slices* com dimensões  $256 \times 256$ . Com o primeiro par de *slices* foram gerados cinco *slices* interpolados, utilizando-se os métodos SIMOL e SIDITRANS. Os *slices* obtidos em ambos os métodos são exibidos na Figura 5.12 (o primeiro e último *slices* são os originais, e os demais são interpolados).

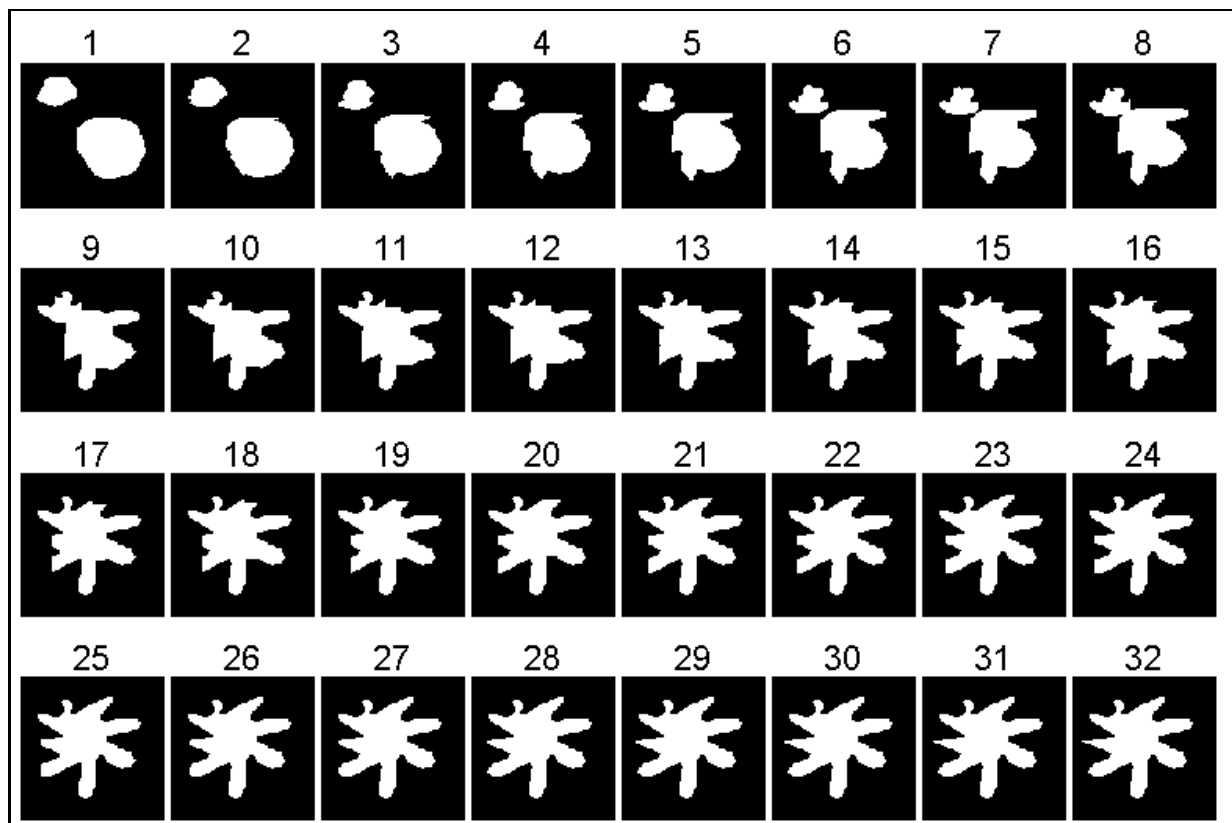


Figura 5.9: Conjunto de *Slices* obtidos usando o algoritmo SIMOL.

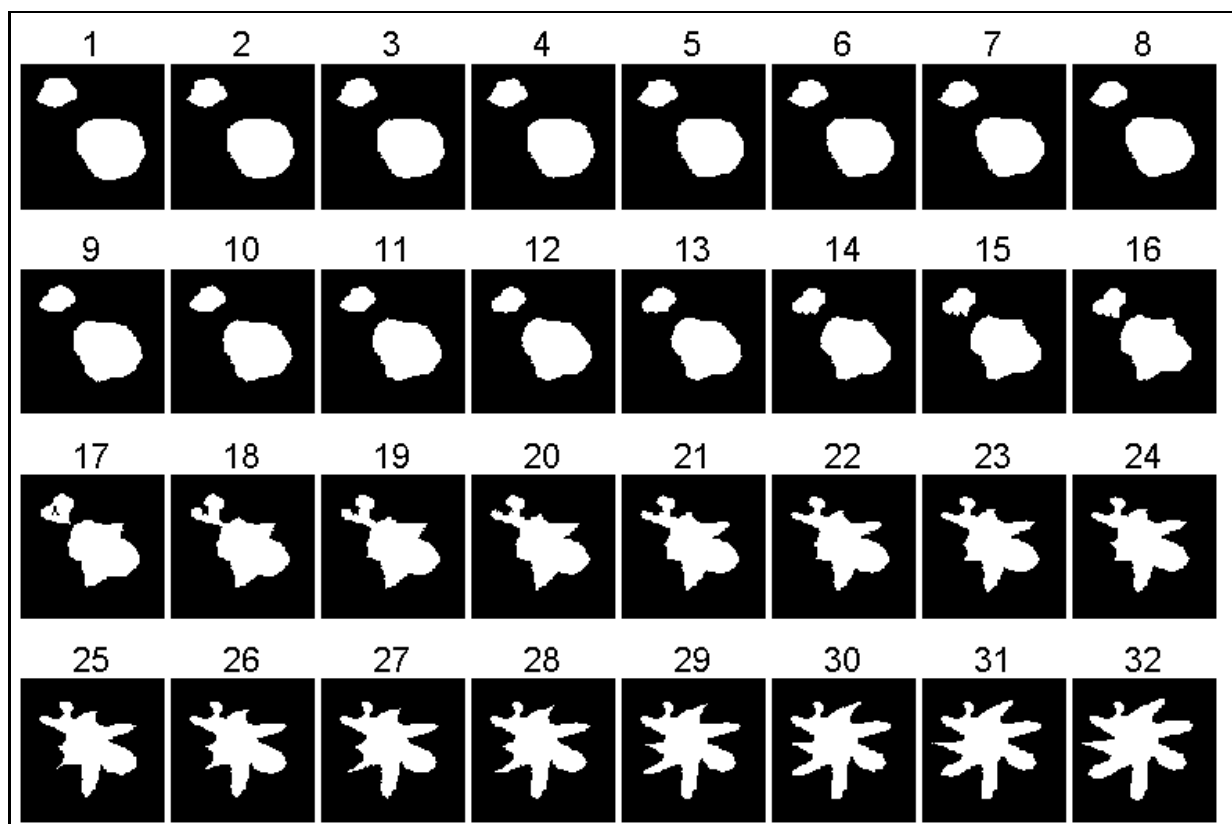


Figura 5.10: Conjunto de *Slices* obtidos usando o algoritmo SIDITRANS.

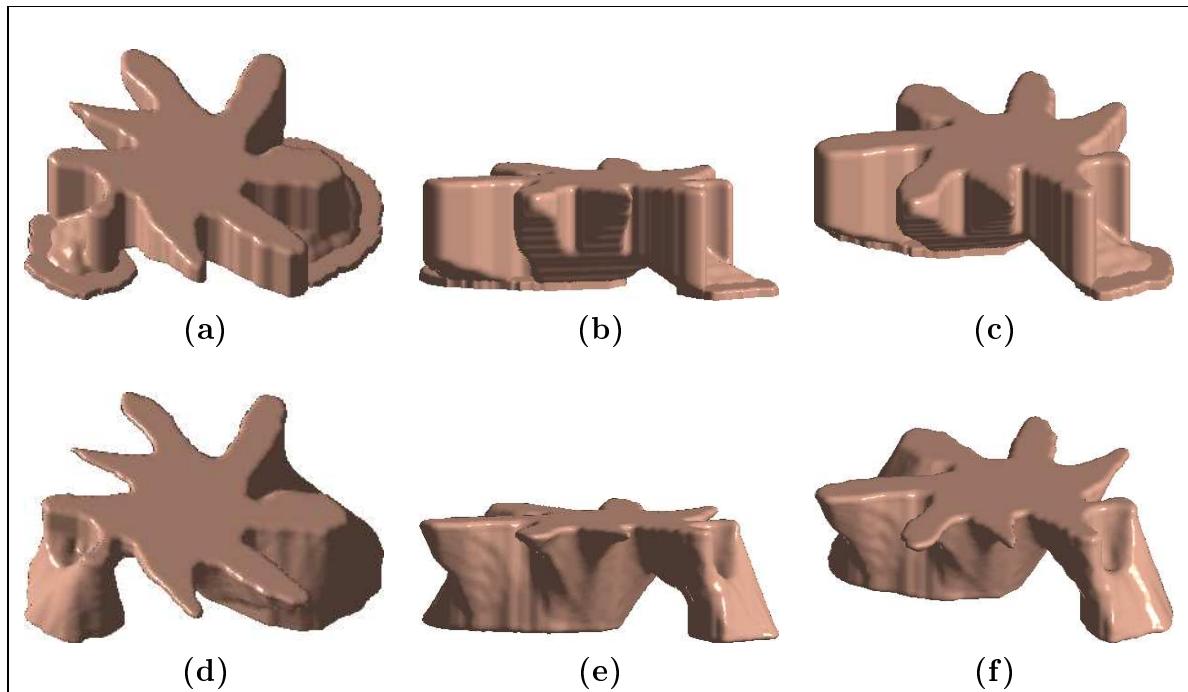


Figura 5.11: Volumes obtidos com os métodos propostos: (a-c) Diferentes vistas do volume obtido utilizando SIMOL. (d-f) Diferentes vistas do volume obtido utilizando SIDI-TRANS.

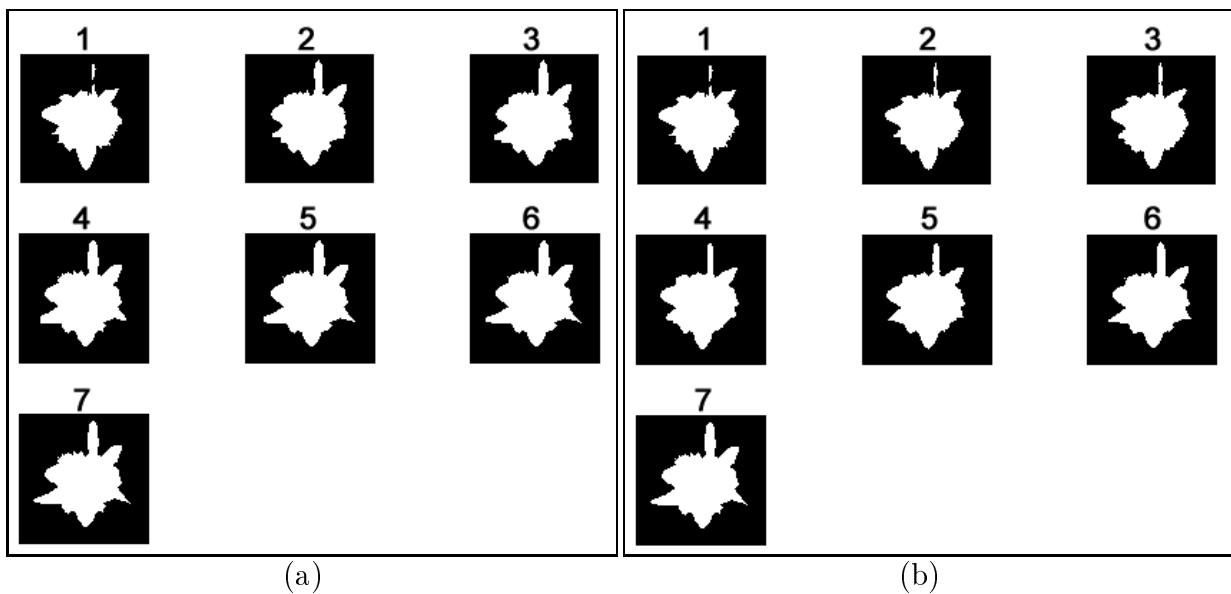


Figura 5.12: Conjunto de *slices* obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDI-TRANS.

A Figura 5.13 exibe as imagens de diferença entre os *slices* do volume original e os *slices* do volume interpolado para cada um dos métodos propostos. Nestas imagens, a cor preta indica que houve uma correspondência entre o volume original e o volume interpolado, ao passo que a cor branca indica uma discordância entre os dados originais e interpolados (observe que o primeiro e último *slices* são os originais, enquanto os demais são interpolados). Note que o SIDI-TRANS apresenta maior aproximação do *slice* interpolado com o *slice* obtido do modelo original.

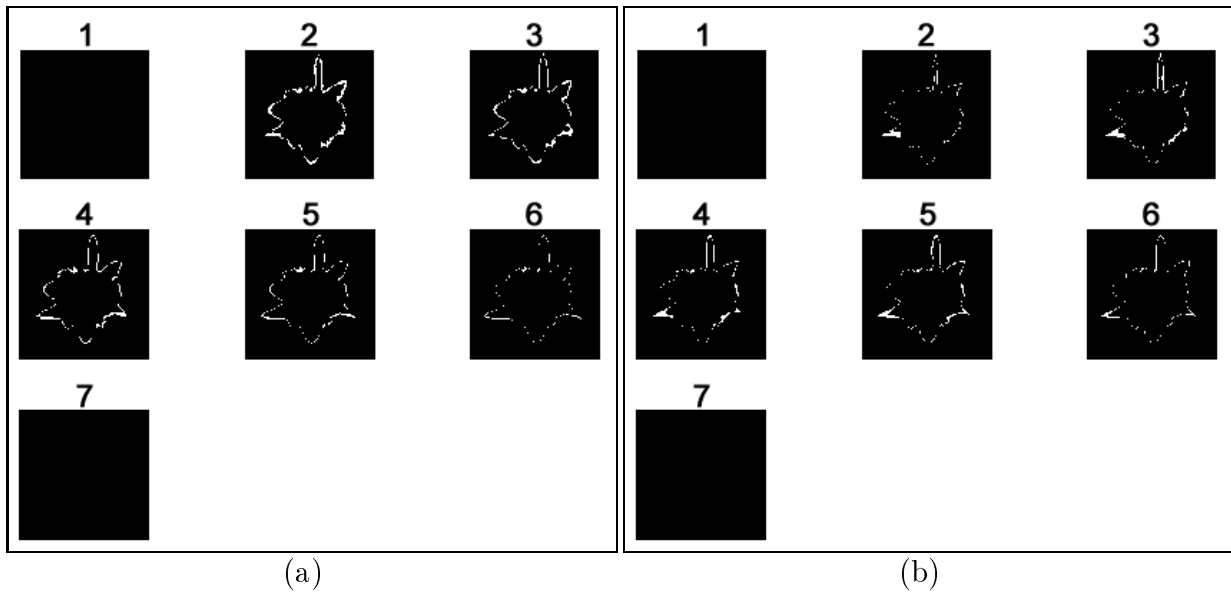


Figura 5.13: Diferença entre os *slices* originais e os interpolados obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS.

Na Figura 5.14 são exibidos os volumes original e interpolados (para uma melhor visualização os volumes são exibidos com uma ampliação na dimensão do eixo  $Z$ ). Os volumes interpolados ilustrados nestas figuras correspondem aos *slices* da Figura 5.12.

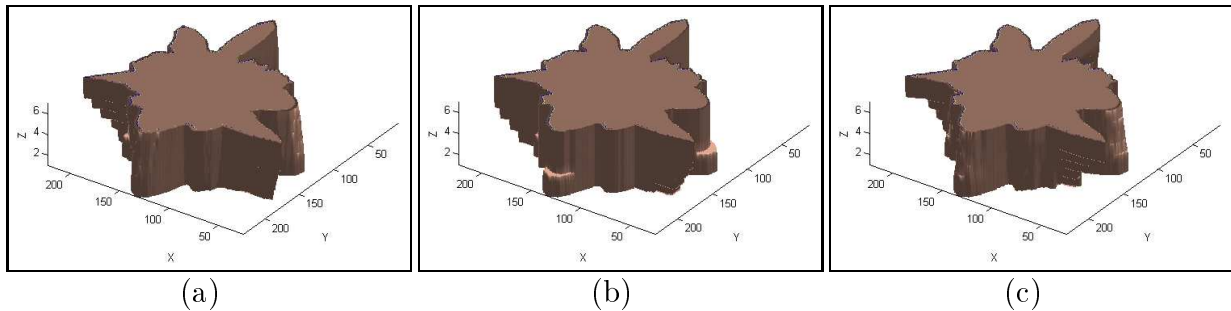


Figura 5.14: a) Volume original. b) Volume obtido usando algoritmo SIMOL. c) Volume obtido usando algoritmo SIDITRANS.

Um segundo par de *slices* foi utilizado no segundo experimento, porém, desta vez, gerando dez *slices* interpolados em ambos os métodos propostos neste capítulo. A Figura 5.15 exibe os *slices* obtidos com ambos os métodos. Na Figura 5.16 são dadas as imagens diferenças entre os dados originais e os dados interpolados. Por fim, os volumes original e interpolados são exibidos na Figura 5.17 (os volumes interpolados correspondem aos *slices* obtidos na Figura 5.15).

A Tabela 5.1 exibe as taxas de erro de interpolação neste segundo experimento. Repare que, para o primeiro par de *slices* do experimento, os *slices* 1 e 7 são os originais fornecidos como parâmetros para os métodos de interpolação, de modo que a taxa de erro para estes *slices* é zero (o mesmo acontece com os *slices* 1 e 12 do segundo par de *slices*). O cálculo do erro médio leva em conta somente os *slices* interpolados, ou seja, os originais não são considerados.

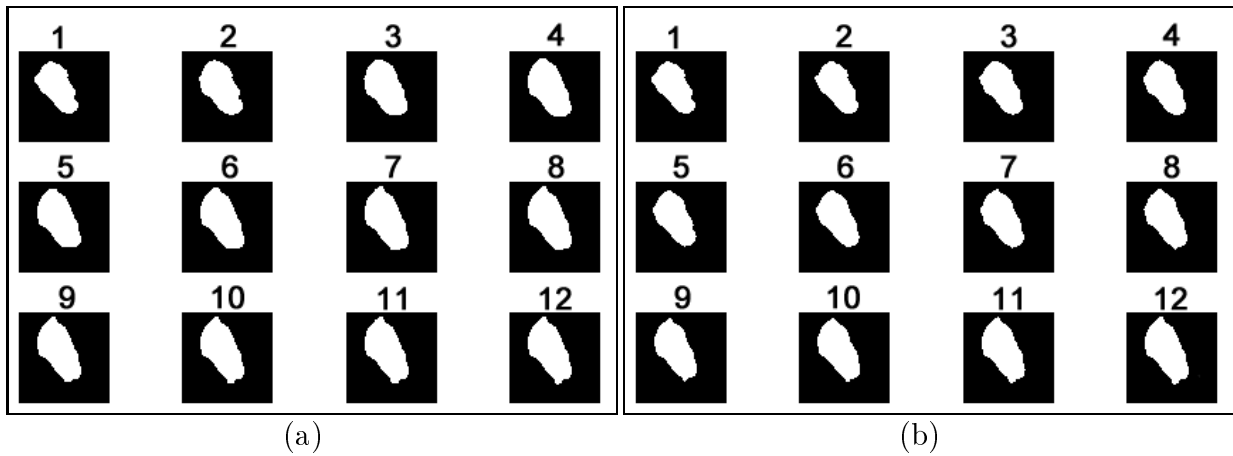


Figura 5.15: Conjunto de *slices* obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS.

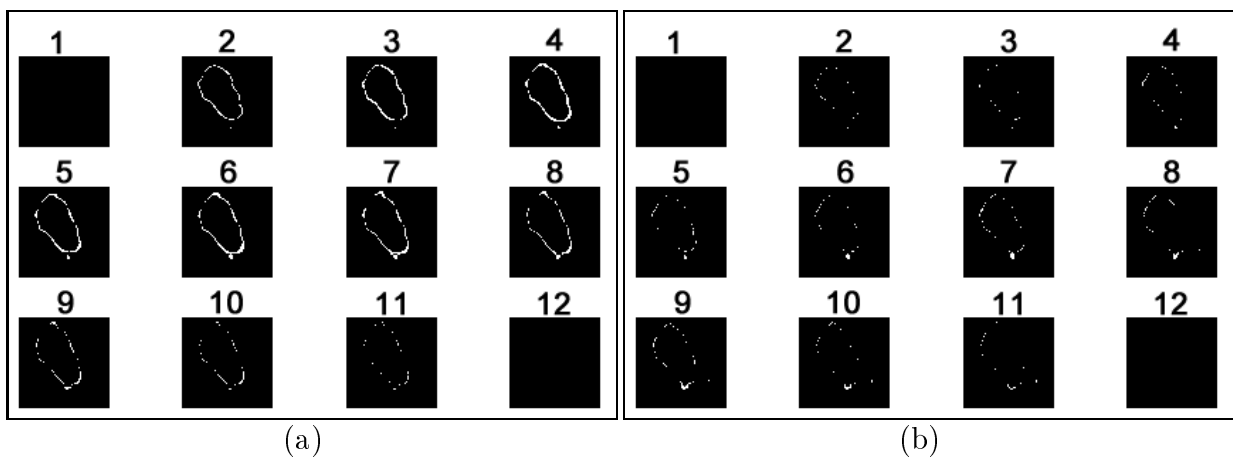


Figura 5.16: Diferença entre os *slices* originais e os interpolados obtidos usando: a) O algoritmo SIMOL. b) O algoritmo SIDITRANS.

É importante notar que os métodos SIMOL e SIDITRANS são definidos para imagens binárias, uma vez que neste trabalho está sendo considerada apenas a informação de forma (porém, para imagens em escala de cinza, ou colorida, a interpolação pode ser realizada apenas aplicando-se a equação 5.1).

Também é importante notar que o método SIMOL é capaz de interpolar formas espiculadas (como pode ser observado nos resultados do primeiro experimento), ao contrário do método de Guo et al., também baseado em morfologia matemática e descrito na Seção 2.5. Além disso, o fato de usar operações clássicas de morfologia matemática (dilatação e erosão) e operações lógicas simplifica a implementação do método. Por outro lado, o algoritmo SIMOL não permite configurar a quantidade de *slices* a serem interpolados (para um conjunto de *slices*  $S_1$ ,  $S_2$  e  $S_3$ , onde  $S_1$  é adjacente a  $S_2$  que, por sua vez, é adjacente a  $S_3$ , a quantidade de *slices* interpolados entre  $S_1$  e  $S_2$  pode diferir da quantidade obtida entre  $S_2$  e  $S_3$ ). Contudo, uma amostragem da quantidade de *slices* obtidos (descartando ou repetindo alguns) pôde facilmente resolver este problema nos experimentos realizados. Esta desvantagem, também presente no método de Luo e Hancock baseado na transfor-

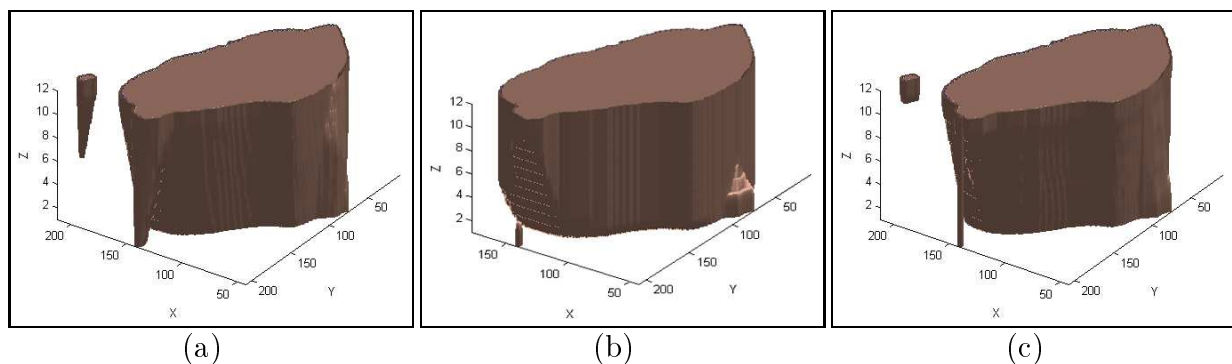


Figura 5.17: a) Volume original. b) Volume obtido usando algoritmo SIMOL. c) Volume obtido usando algoritmo SIDITRANS.

Tabela 5.1: Erro de interpolação no segundo experimento. Valores são dados em porcentagem de *voxels* incorretos.

| Primeiro par de <i>slices</i> |          |           | Segundo par de <i>slices</i> |          |           |
|-------------------------------|----------|-----------|------------------------------|----------|-----------|
|                               | Erro     |           |                              | Erro     |           |
| <i>Slices</i>                 | Simol    | Siditrans | <i>Slices</i>                | Simol    | Siditrans |
| 1                             | 0,000000 | 0,000000  | 1                            | 0,000000 | 0,000000  |
| 2                             | 0,033188 | 0,014572  | 2                            | 0,019394 | 0,003754  |
| 3                             | 0,029144 | 0,020386  | 3                            | 0,027344 | 0,004089  |
| 4                             | 0,025345 | 0,019699  | 4                            | 0,030701 | 0,004974  |
| 5                             | 0,018112 | 0,018219  | 5                            | 0,030121 | 0,006119  |
| 6                             | 0,009766 | 0,012741  | 6                            | 0,029617 | 0,007217  |
| 7                             | 0,000000 | 0,000000  | 7                            | 0,026016 | 0,008575  |
|                               |          |           | 8                            | 0,020950 | 0,008743  |
|                               |          |           | 9                            | 0,016159 | 0,008255  |
|                               |          |           | 10                           | 0,010788 | 0,006683  |
|                               |          |           | 11                           | 0,005844 | 0,005051  |
|                               |          |           | 12                           | 0,000000 | 0,000000  |
| Média                         | 0,023111 | 0,017123  | Média                        | 0,021693 | 0,006346  |

mação de distância, é contornada no algoritmo SIDITRANS. Outra vantagem do algoritmo SIDITRANS, em comparação com o método de Luo e Hancock, é sua simplificação (uma vez que o método não necessita mais da realização de operações morfológicas). Em suma, pode-se verificar que ambos os métodos de interpolação, apresentados neste capítulo, superam as limitações dos métodos de Guo et al. e de Luo e Hancock, mostrados na Seção 2.5, e simplificam o processo de implementação dos algoritmos.

Ao comparar ambos os métodos desenvolvidos, é possível perceber visualmente que o método SIDITRANS apresentou uma interpolação mais suave, conforme se observa na figura 5.11. Também é possível constatar, a partir das Figuras 5.13 e 5.16 e da Tabela 5.1, que o método SIDITRANS apresentou resultados mais próximos dos dados originais. Por outro lado, o algoritmo SIMOL executou em menos tempo. A Tabela 5.2 apresenta o tempo para execução em segundos dos algoritmos SIMOL e SIDITRANS em ambos os experimentos. Como pode ser observado o algoritmo SIMOL foi mais rápido em todos os

experimentos. O algoritmo SIDITRANS levou um tempo 16,6 vezes maior para executar o primeiro experimento e, em média, 21,2 vezes maior no segundo experimento. Ao aumentar as dimensões dos *slices* utilizados (de  $128 \times 128$  no primeiro experimento para  $256 \times 256$  no segundo) o método SIMOL levou, em média, 3,29 vezes mais tempo para executar, enquanto o método SIDITRANS tomou, em média, 4,15 vezes mais tempo.

Tabela 5.2: Tempo de execução dos experimentos de interpolação usando os algoritmos SIMOL e SIDITRANS.

| Experimento                               | Tempo de Execução (em segundos) |           |
|-------------------------------------------|---------------------------------|-----------|
|                                           | SIMOL                           | SIDITRANS |
| Exper. 1                                  | 0,3934                          | 06,5200   |
| Exper. 2 (primeiro par de <i>slices</i> ) | 1,4762                          | 28,2702   |
| Exper. 2 (segundo par de <i>slices</i> )  | 1,1134                          | 25,7980   |

Em suma, o algoritmo SIDITRANS, apesar de consumir uma maior quantidade de tempo de execução, apresenta uma interpolação mais suave e mais próxima dos dados originais (o algoritmo SIMOL apresentou *voxels* incorretos em uma quantidade 2,3840 vezes maior que o método SIDITRANS).

### 5.3 Considerações Finais

Neste capítulo foi apresentado o *framework* WESS composto pelo aplicativo EMISS, uma solução desenvolvida para segmentar as imagens volumétricas, e pelos algoritmos de interpolação SIMOL e SIDITRANS, capazes de obter o volume de *voxels* que represente a forma tridimensional da região de interesse segmentada no EMISS.

O uso do EMISS, em conjunto com um dos algoritmos de interpolação, também apresentados neste capítulo, permite, por exemplo, a criação de volumes de *voxels* a partir de imagens médicas provenientes de exames de ressonância magnética ou tomografia computadorizada. Assim, este exemplo ilustra a possibilidade de uso do descritor *3Hq Feature*, definido no capítulo anterior, em aplicação médica.

# Capítulo 6

## Experimentos de Classificação

### 6.1 Introdução

A fim de mostrar o potencial e validar a eficiência dos descritores *2Hq Feature* e *3Hq Feature* apresentados no Capítulo 4, alguns experimentos foram realizados, envolvendo tarefas de classificação.

O sistema de classificação utilizado nos experimentos tem por base o uso de redes neurais artificiais (NN). Para isto, foi empregada uma rede *feed-forward* multicamada com *back propagation* projetada com  $4^L$  entradas (para experimentos usando o *2Hq Feature*) ou  $8^L$  entradas (para experimentos usando o *3Hq Feature*), onde  $L$  é o nível de decomposição utilizado no descritor. Cada  $2Hq_i$  (ou  $3Hq_i$ ) é associado a uma entrada da rede neural, conforme ilustrado na Figura 6.1. O número de saídas da rede depende do número de categorias da base de dados utilizada.

A rede neural é treinada em *batch mode*, de modo que seus pesos e *bias* são atualizados somente depois de todo o conjunto de treinamento ter sido processado pela rede. Para evitar o efeito de supertreinamento (*over-fitting*), foi utilizada a estratégia *Early Stopping* para validação da rede (esta estratégia determina um ponto de parada no treinamento, tentando encontrar um ponto ótimo de generalização). A função de transferência usada é a tangente sigmoide hiperbólica.

As seções a seguir descrevem os experimentos realizados e exibe os resultados obtidos, avaliados em termos de curvas ROC e área sob a curva ROC (AUC). O conjunto de dados em cada experimento é dividido randomicamente em três partes: 60% das amostras são atribuídas ao conjunto de treinamento, 20% atribuídas ao conjunto de validação e 20% atribuídas ao conjunto de testes. Validação cruzada foi utilizada nos experimentos. Convém esclarecer que, nos experimentos apresentados a seguir, um mesmo classificador baseado em redes neurais é utilizado tanto para o *2Hq Feature*, quanto para os demais métodos utilizados para comparação.

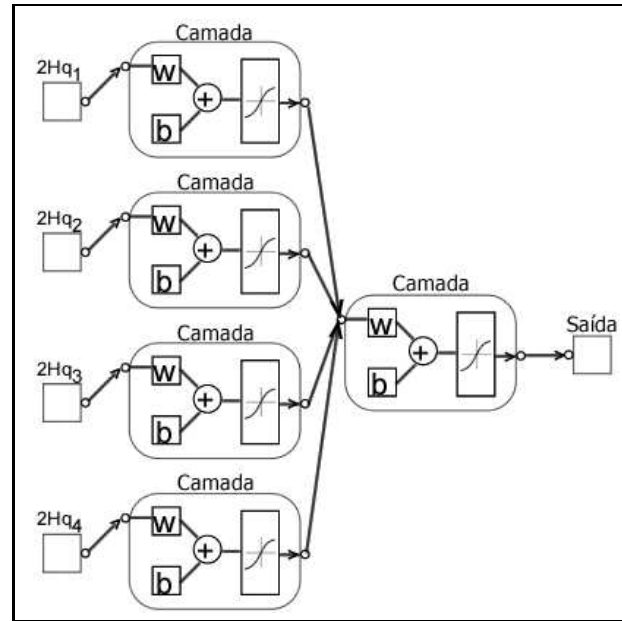


Figura 6.1: Exemplo do esquema de uma rede neural recebendo representações  $2Hq_i$  em suas quatro entradas.

## 6.2 Primeiro Experimento: Classificação Usando as Bases Kimia e MPEG-7

O primeiro experimento realizado utiliza o descritor *2Hq Feature* para classificação de silhuetas bidimensionais. Para isto, foram utilizadas as bases de dados *MPEG-7 CE-Shape 1 Part-B* [Latecki et al. 2000], Kimia-99 (formada a partir de diferentes fontes na Internet, inclui imagens com oclusões, articulações e deformações) e Kimia-216 (formada a partir de imagens da base MPEG-7) [Sebastian et al. 2004]. A Tabela 6.1 lista as características das bases de dados utilizadas neste experimento. A Figura 6.2 exibe uma amostra dos objetos presentes na base MPEG-7 e o conjunto de objetos das bases Kimia-99 e Kimia-216.

Tabela 6.1: Descrição das bases de imagens utilizadas no primeiro experimento

| Base de dados      | Kimia-99     | Kimia-216    | MPEG-7       |
|--------------------|--------------|--------------|--------------|
| Tipo de imagens    | silhuetas 2D | silhuetas 2D | silhuetas 2D |
| Número de classes  | 9            | 18           | 70           |
| Imagens por classe | 11           | 12           | 20           |
| Total de imagens   | 99           | 216          | 1400         |

Tabela 6.2: Parâmetros utilizados no experimento para cada base de imagens.

| Parâmetro                        | Valor    |           |        |
|----------------------------------|----------|-----------|--------|
|                                  | Kimia-99 | Kimia-216 | MPEG-7 |
| Nível de decomposição ( $L$ )    | 1        | 1         | 2      |
| Família Wavelet                  | Haar     | Haar      | Haar   |
| Tamanho do $2Hq_i$               | 256      | 256       | 256    |
| Neurônios nas camadas escondidas | 20       | 20        | 90     |

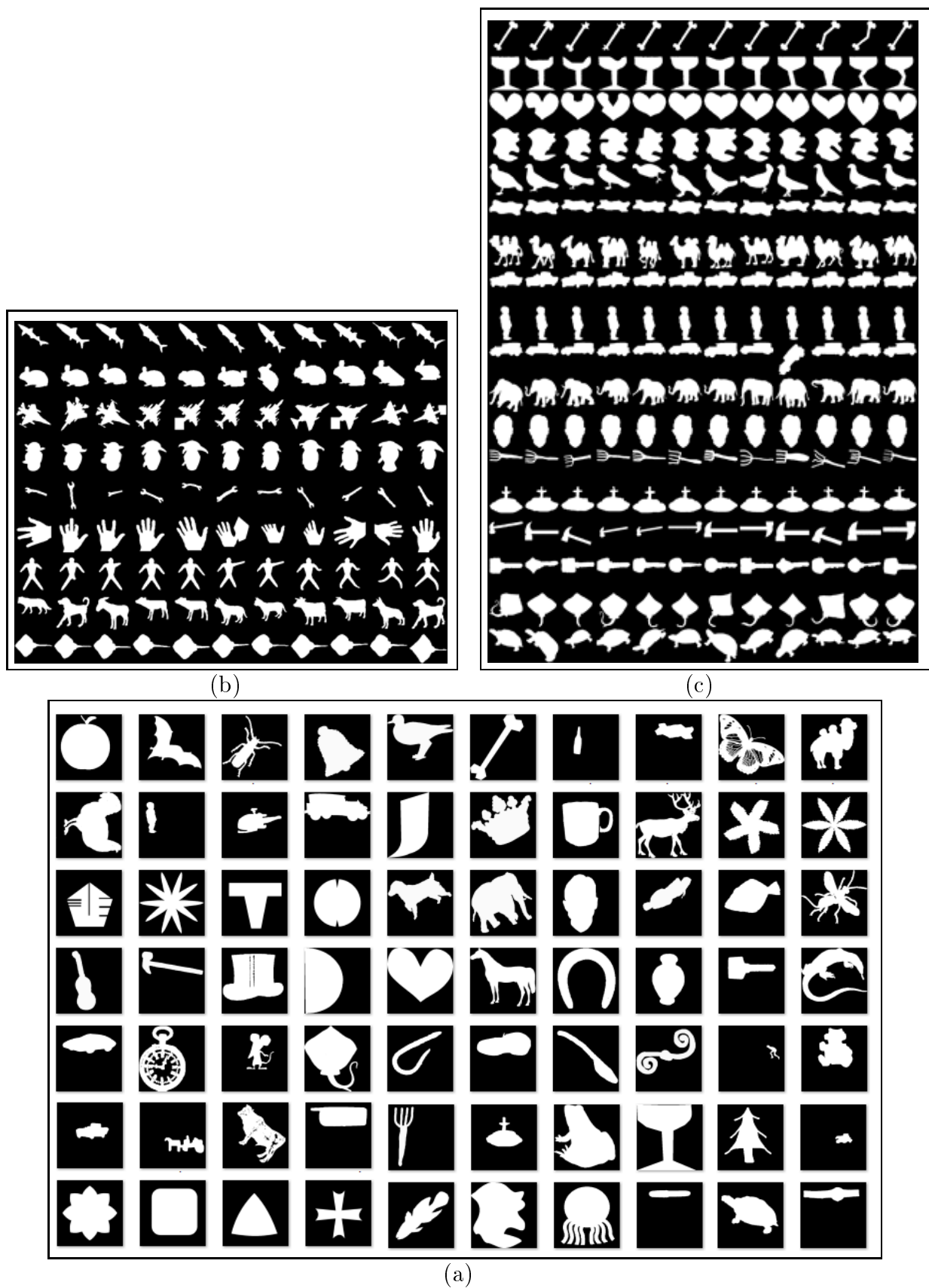


Figura 6.2: a) Conjunto de imagens da base Kimia-99. b) Conjunto de imagens da base Kimia-216. c) Amostra do conjunto de imagens da base MPEG-7.

Os resultados alcançados neste experimento são comparados aos obtidos pelo trabalho de Ebrahim et al. em [Ebrahim et al. 2008]. Os parâmetros utilizados para o cálculo do descritor *2Hq Feature* são apresentados na Tabela 6.2. As imagens usadas têm dimensões  $256 \times 256$  nas três bases de imagens deste experimento. A Tabela 6.3 exibe os valores de AUC obtidos para as três bases de imagens deste experimento. A Figura 6.3 expõe as curvas ROC alcançadas com as bases de imagens Kimia utilizadas neste experimento. A Figura 6.4 mostra as curvas ROC providas base de dados MPEG-7. O ganho com relação ao método de Ebrahim et al., em termos de área sob a curva ROC, foi de 0,07, 0,08 e 0,14 para as bases Kimia-99, Kimia-216 e MPEG-7, respectivamente, obtendo uma classificação ótima para as duas bases Kimia (o resultado ideal obtido nas bases Kimia pode ser explicado pela pequena quantidade de dados em ambas as bases de imagens).

Tabela 6.3: Valores de AUC obtidos no primeiro experimento.

| Base de imagens | Área sob a curva ROC |                |
|-----------------|----------------------|----------------|
|                 | <i>2Hq Feature</i>   | Ebrahim et al. |
| Kimia-99        | 1.00                 | 0.93           |
| Kimia-216       | 1.00                 | 0.92           |
| MPEG-7          | 0.96                 | 0.82           |

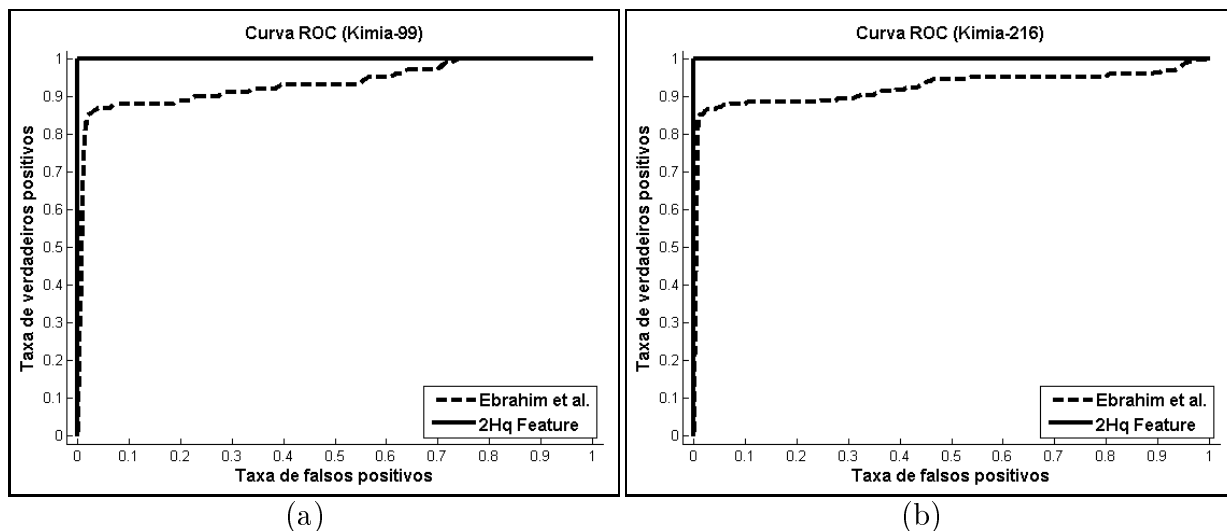


Figura 6.3: Curvas ROC obtidas para as bases de dados: a) Kimia-99. b) Kimia-216.

Neste experimento, foram utilizadas bases de silhuetas 2D para avaliar o desempenho do extrator *2Hq Feature*. Os resultados apontam que o extrator proposto supera, em termos de curva ROC e valores AUC, o método de Ebrahim et al. no qual se inspirou.

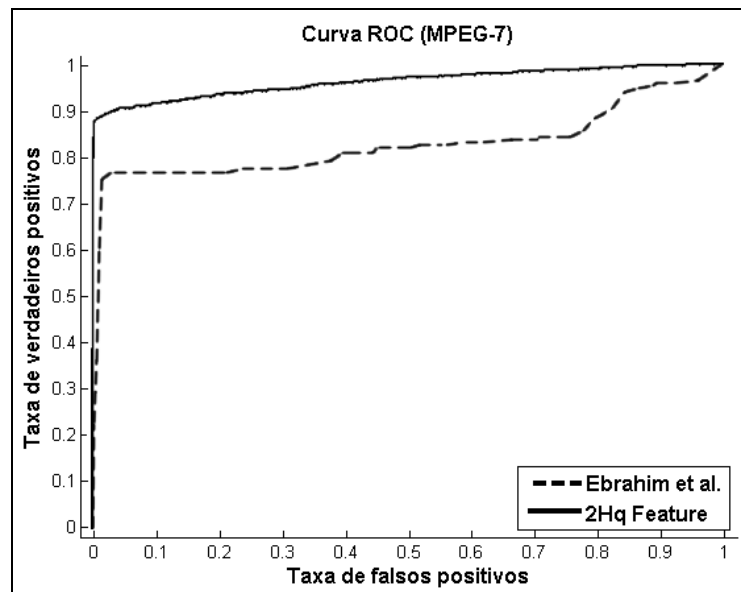


Figura 6.4: Curvas ROC obtidas para as bases de dados MPEG-7 CE Shape 1 Part-B.

### 6.3 Segundo Experimento: Rotação

Para avaliar a sensibilidade do método *2Hq Feature* com relação a transformações de rotação nas silhuetas, foram considerados dois cenários: primeiramente, apenas o objeto de consulta (*query*) é rotacionado; na segunda situação, o método proposto é aplicado a uma base de imagens contendo objetos com diversos graus de rotação.

Para a realização do experimento, considerando-se o primeiro cenário descrito, a base de dados *MPEG-7 CE Shape 1 Part-B* foi utilizada para treinamento. As imagens de consulta (e somente estas) foram rotacionadas em ângulos múltiplos de cinco graus conseguindo assim um total de 72 rotações para cada objeto consultado.

A Tabela 6.4 lista os parâmetros utilizados neste experimento, que consiste no uso do extrator *2Hq Feature* para classificação das silhuetas utilizando a rede neural descrita na Seção 6.1. Neste experimento, a rede neural foi treinada com as imagens não rotacionadas, ou ângulo zero de rotação, e testada com as imagens rotacionadas. As imagens utilizadas têm dimensões  $256 \times 256$ . Novamente, os resultados são comparados ao método de Ebrahim et al. A Figura 6.5 exibe a variação da área sob a curva ROC obtida ao rotacionar os objetos de consulta.

Tabela 6.4: Parâmetros utilizados no primeiro cenário do experimento.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 2     |
| Família Wavelet                  | Haar  |
| Tamanho do $2Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 100   |

Para uma situação em que a base de dados contém objetos rotacionados em diversos ângulos, correspondente ao segundo cenário descrito no início desta seção, foi criada uma

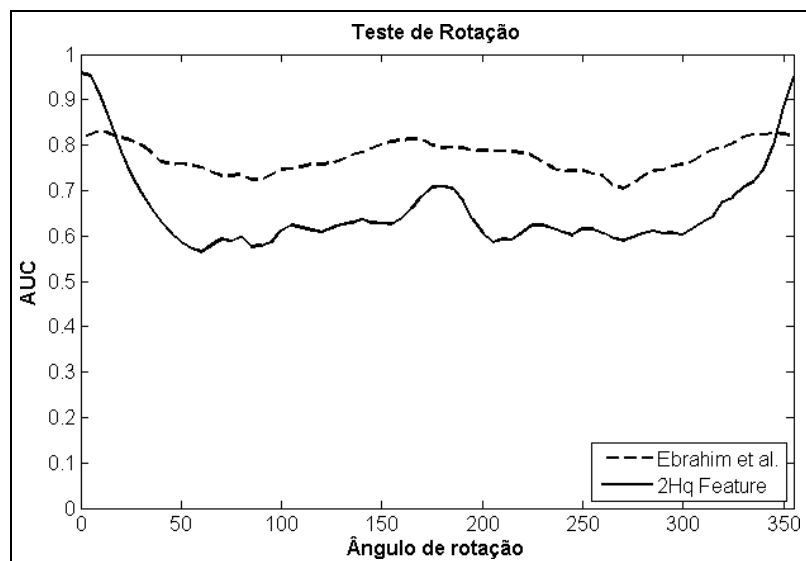


Figura 6.5: AUC obtida para cada ângulo de rotação.

base de imagens, denominada W-720<sup>1</sup>, formada a partir de dez imagens, de classes diferentes, tomadas a partir da base *MPEG-7 CE Shape 1 Part-B*. Cada uma das dez imagens foi rotacionada em ângulos múltiplos de cinco graus, totalizando 72 rotações para cada objeto. Assim, a base W-720 é constituída por 10 classes com 72 imagens cada, resultando em 720 silhuetas. A Figura 6.6 ilustra uma amostra da base de imagens criada.

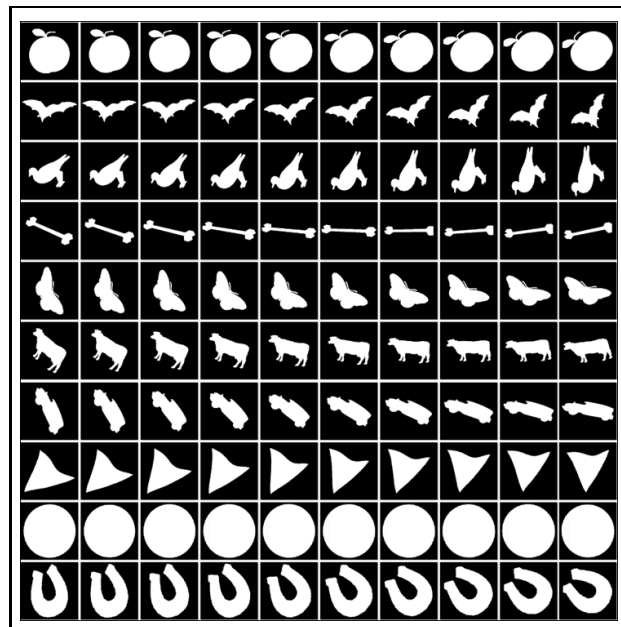


Figura 6.6: Amostra dos objetos contidos na base de imagens W-720.

A Tabela 6.5 lista os parâmetros utilizados no segundo cenário previsto neste experimento. Novamente, o extrator *2Hq Feature* é utilizado na classificação das silhuetas, conforme descrito na seção 6.1. As imagens utilizadas têm dimensões  $256 \times 256$  e os

<sup>1</sup>A base W-720 encontra-se disponível através do endereço eletrônico <http://www.facom.ufu.br/~walter/hqf/database/>.

resultados são comparados ao método de Ebrahim et al. A curva ROC obtida a partir dos resultados encontrados é ilustrada na Figura 6.7. A área sob a curva ROC obtida foi de 0.990 para o método *2Hq Feature* e de 0.892 para o método de Ebrahim et al.

Tabela 6.5: Parâmetros utilizados no segundo cenário do experimento.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 2     |
| Família Wavelet                  | Haar  |
| Tamanho do $2Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 20    |

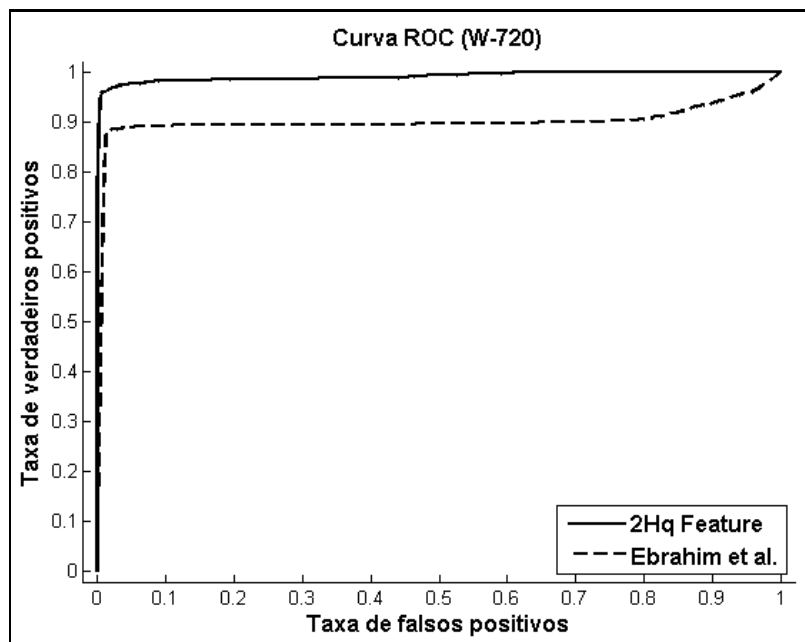


Figura 6.7: Curva ROC obtida na classificação da base W-720.

Como pode ser observado neste experimento, o extrator *2Hq Feature* sofreu mais efeito das transformações de rotação quando a base de imagens não continha objetos rotacionados, mas somente os objetos de consulta são rotacionados, embora tenha tido um bom desempenho para rotações de até dez graus. Para o caso onde há rotações diversas, tanto nas imagens da base de dados quanto nas imagens de consulta, o método proposto sofreu menos o efeito das transformações aplicadas, com relação ao método de Ebrahim et al.

Apesar do desempenho sofrido no primeiro cenário deste experimento, é possível utilizar métodos de normalização de pose e, assim, tornar o método proposto invariante a rotações (o método PCA tem sido tradicionalmente utilizado para este fim).

## 6.4 Terceiro Experimento: Classificação de Tumores de Câncer de Mama

Massas anormais geralmente aparecem em mamogramas como regiões mais densas. Em geral, tumores malignos têm um contorno mais áspero, espiculado ou microlobulado, enquanto massas benignas têm contornos mais suaves, arredondados/ovalados ou macrolobulados. Entretanto casos atípicos também ocorrem como massas benignas espiculadas ou macrolobuladas e tumores malignos em formas mais suaves e arredondadas, tornando a tarefa de classificação um verdadeiro desafio [Guliato et al. 2010, Rangayyan et al. 2000, American College of Radiology, Reston, VA 1998, Homer 1997, Rangayyan et al. 1997].

Neste terceiro experimento, o extrator *2Hq Feature* é utilizado para classificar calcificações de câncer de mama em massas benignas ou tumores malignos. As imagens que compreendem esta base de dados provêm de duas fontes diferentes. O primeiro conjunto contém vinte contornos de mamogramas obtidos a partir do *Screen Test: the Alberta Program for the Early Detection of Breast Cancer* [Alberta Cancer Board, Alberta, Canada 2004]. Estes mamogramas foram digitalizados através de um escaner Lumiscan 85 com resolução de  $50\mu\text{m}$  e 12 *bits* por *pixel*. O conjunto possui 57 regiões de interesse (ROIs), dos quais 37 correspondem a massas benignas e 20 a tumores malignos [Guliato et al. 2010].

O segundo conjunto de imagens presente na base de dados deste experimento foi obtido a partir do banco de dados de *Mammographic Image Analysis Society* (MIAS, UK) [MIA 2004] e da biblioteca de ensino do *Foothills Hospital* em Calgary, Canadá. As imagens do MIAS foram digitalizadas com resolução de  $50\mu\text{m}$ , e as imagens do *Foothills Hospital* digitalizados em resolução de  $62\mu\text{m}$ . Este conjunto inclui 28 massas benignas e 26 tumores malignos.

Os contornos de cada massa foram manualmente desenhados por um radiologista especializado em mamografia. O banco de dados combinado pelos dois conjuntos de dados totaliza 111 contornos, incluindo formas típicas e atípicas de massas benignas (65) e tumores malignos (46), e é referenciado neste trabalho como BC-111<sup>2</sup>. A classificação dos contornos é baseada em resultados de biópsias. A Figura 6.8 exibe os contornos contidos na base de dados utilizada neste experimento.

A Tabela 6.6 lista os parâmetros usados neste experimento. As imagens utilizadas têm dimensões  $256 \times 256$ . Os resultados obtidos são comparados com os extratores *Spiculation Index* ( $SI_{TA}$ ) [Guliato et al. 2008b], *Fractal Dimension* ( $FD_{TA}$ ) [Nguyen e Rangayyan 2006] e *Convexity Index* ( $CI_{TA}$ ) [Guliato et al. 2008a], aplicados sobre a assinatura *Turning Angle Function* ( $TA$ ).

Os valores de área sob a curva ROC obtidos neste experimento são mostrados na

---

<sup>2</sup>A BC-111 é uma base de acesso restrito, não estando disponível publicamente para *download*.

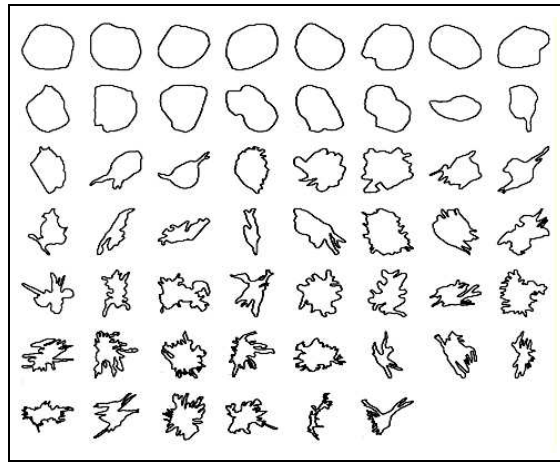


Figura 6.8: Subconjunto da base de imagens com contornos de tumores de câncer [Alto et al. 2005].

Tabela 6.6: Parâmetros utilizados no terceiro experimento.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 1     |
| Família Wavelet                  | Haar  |
| Tamanho do $2Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 20    |

Tabela 6.7. A Figura 6.9 exibe as curvas ROC obtidas com os extratores  $2Hq$  Feature,  $SI_{TA}$ ,  $CI_{TA}$  e  $FD_{TA}$  descritos em [Guliatto et al. 2008b, Guliatto et al. 2008a, Nguyen e Rangayyan 2006]. A Tabela 6.7 relaciona os valores de área sob a curva ROC, obtidos neste experimento, onde se nota um ganho mínimo de 0,04 com relação aos métodos comparados.

Tabela 6.7: Valores de AUC obtidos no terceiro experimento.

| Extrator      | Área sob a curva ROC |
|---------------|----------------------|
| $2Hq$ Feature | 0.98                 |
| SI            | 0.93                 |
| CX            | 0.92                 |
| FD            | 0.94                 |

Neste experimento, pudemos verificar que o extrator  $2Hq$  Feature é capaz de distinguir entre massas benignas e tumores malignos, usando contornos de calcificações de mama. Os resultados apontam que, em termos de curva ROC e AUC, o extrator proposto supera outros extratores do estado-da-arte.

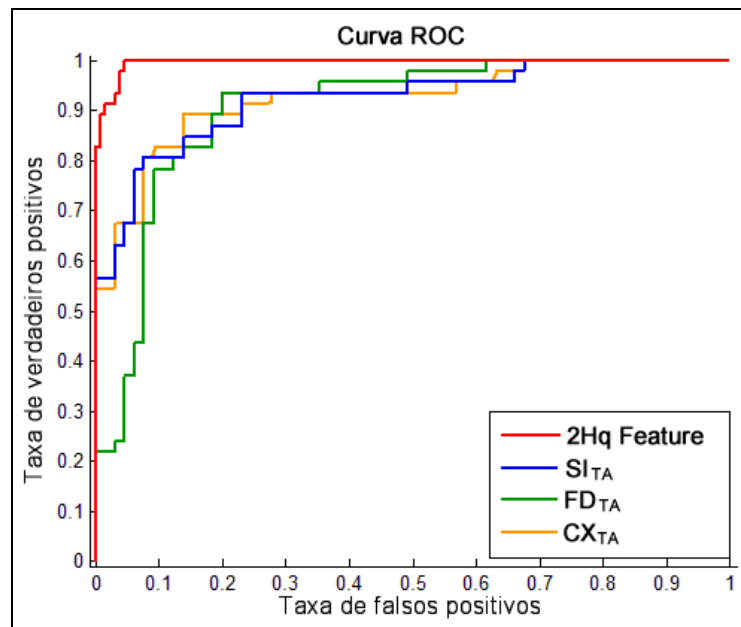


Figura 6.9: Curva ROC obtida na classificação de calcificações de câncer de mama.

## 6.5 Quarto Experimento: Classificação de Objetos 3D Espiculados e Arredondados

Como foi visto no experimento anterior calcificações de mama podem ser classificadas analisando-se suas formas. Embora casos atípicos ocorram, a distinção entre objetos com formatos espiculados ou arredondados pode auxiliar no diagnóstico destas calcificações.

Com isso em mente, foi criada a base de dados sintética W3BC-111 contendo objetos tridimensionais espiculados e arredondados. Os objetos foram criados com a intenção de manter características visuais presentes nos contornos da base utilizada no experimento anterior. Para uma breve comparação, a Figura 6.10 exhibe alguns dos objetos 3D da base W3BC-111 juntamente com alguns contornos da base utilizada no experimento anterior.

Os objetos da base W3BC-111 são divididos em duas classes, espiculados e arredondados, tendo 64 e 47 objetos, respectivamente, totalizando 111 objetos tridimensionais. Os objetos são definidos como volumes de *voxels* com dimensões  $256 \times 256 \times 256$  e disponíveis no endereço eletrônico <http://www.pos.facom.ufu.br/~walter/Database/>.

Neste experimento, o extrator *3Hq Feature* foi utilizado na tarefa de classificação dos objetos da base, conforme descrito na Seção 6.1. Os parâmetros são dados na Tabela 6.8. Os resultados deste experimento são ilustrados na Figura 6.11.

Tabela 6.8: Parâmetros utilizados no quarto experimento.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 1     |
| Família Wavelet                  | Haar  |
| Tamanho do $3Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 20    |

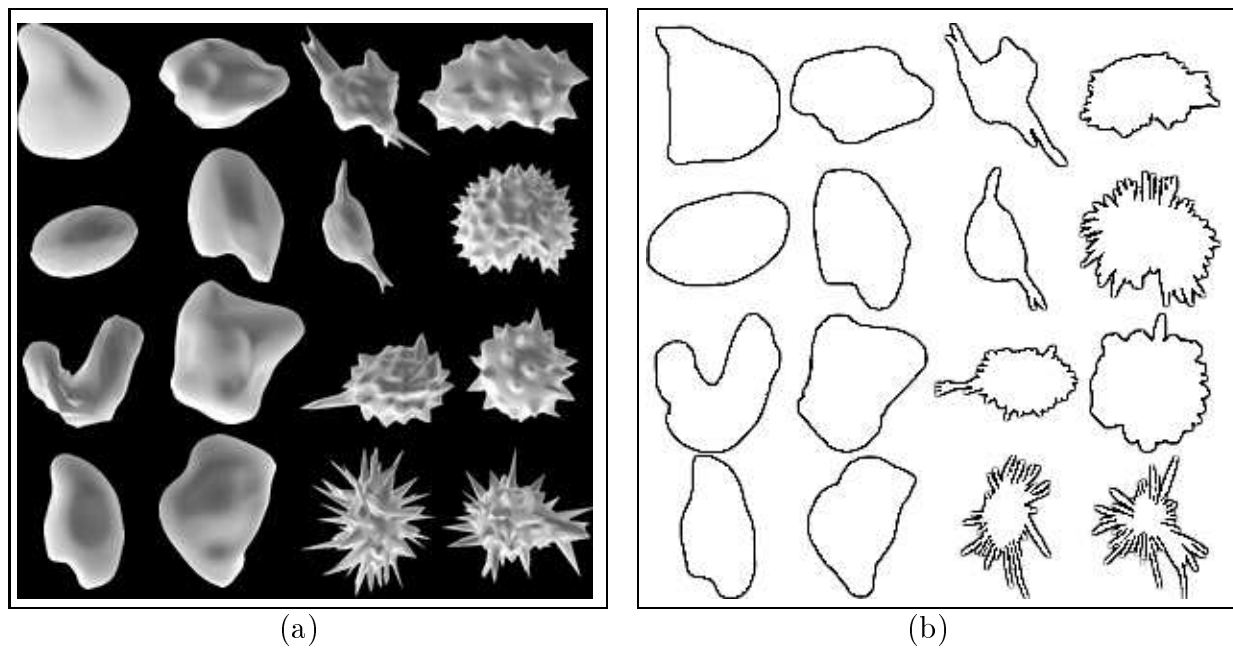


Figura 6.10: a) Subconjunto dos objetos presentes na base W3BC-111 com características visuais semelhantes aos contornos em (b). b) Subconjunto de lesões de mama base BC-111.

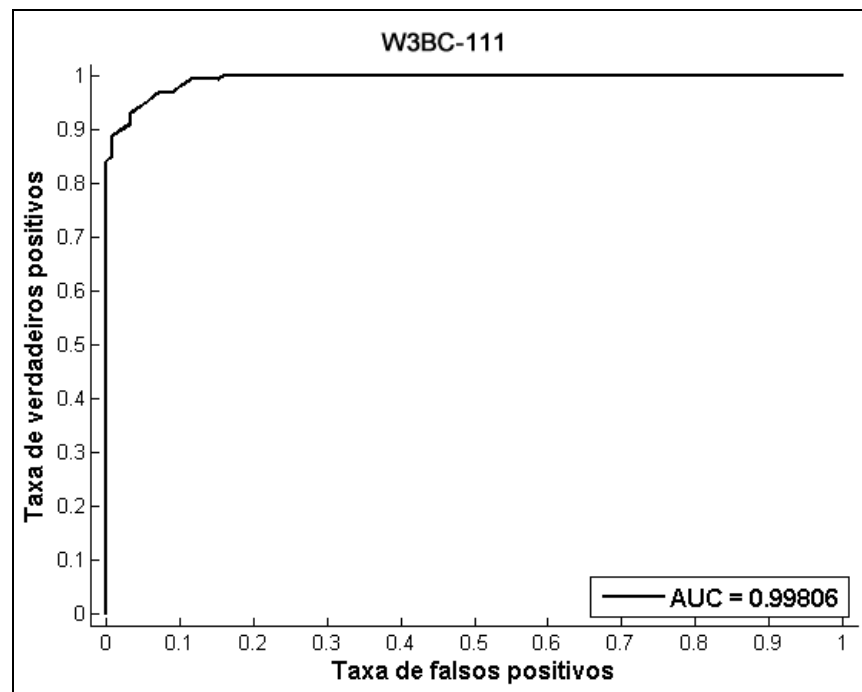


Figura 6.11: Curva ROC obtida na classificação de objetos da base W3BC-111.

Como pode ser visto na Figura 6.11, o extrator *3Hq Feature* apresentou alta taxa de precisão para distinguir entre formas arredondadas e espiculadas, o que o torna potencialmente útil para classificar lesões de câncer de mama em imagens de tomografia computadorizada ou ressonância magnética.

## 6.6 Considerações Finais

Neste capítulo foram apresentados quatro experimentos envolvendo classificação de padrões realizados com o objetivo de avaliar o desempenho dos extratores de forma *2Hq Feature* e *3Hq Feature*, descritos no Capítulo 4.

Tendo a curva ROC e o valor de área sob a curva ROC como medidas de avaliação de desempenho, constatamos que os extratores propostos são robustos na classificação de objetos contendo oclusões, articulações e deformações, como as silhuetas da base Kimia-99, e de objetos rotacionados, conforme verificado no segundo experimento utilizando a base W-720. Também pode ser observado o ótimo desempenho dos referidos extratores em distinguir tanto entre classes de objetos em geral, tais como as silhuetas das bases Kimia-99, Kimia-216, MPEG-7, como objetos de domínios mais específicos, como o caso da base de contornos de calcificações de mama e da base W3BC-111 de objetos espiculados e arredondados.

Em suma, avaliando os quatro experimentos abordados neste capítulo, verificamos que os extratores *2Hq Feature* e *3Hq Feature* alcançaram resultados superiores, em termos de curvas ROC e valores AUC, na classificação de forma de imagens e objetos 3D com relação a alguns métodos do estado-da-arte.

No capítulo seguinte são apresentados experimentos realizados no contexto da recuperação de imagens (ou objetos) baseada em conteúdo.

# Capítulo 7

## Experimentos de Recuperação de Formas Baseada em Conteúdo

### 7.1 Introdução

Neste capítulo os extratores de forma propostos neste trabalho *2Hq Feature* e *3Hq Feature*, introduzidos no Capítulo 4, são avaliados em experimentos envolvendo a recuperação de formas baseada em conteúdo.

O sistema de recuperação de imagens/objetos por conteúdo utilizado neste trabalho foi inspirado no trabalho proposto por Barcelos et al. em [Barcelos et al. 2008]. Este sistema se baseia no uso de redes neurais artificiais, sendo dividido em três fases:

- **Treinamento:** Uma rede neural, projetada conforme descrito na Seção 6.1 do capítulo anterior, é utilizada para realizar o treinamento dos dados.
- **Mapeamento do *HqFeature* para um vetor de características de alto nível:** A rede neural treinada é utilizada para mapear o *2HqFeature* ou *3HqFeature*, obtido a partir de cada objeto do conjunto de dados de teste em um vetor cujo número de elementos é igual ao número de categorias da base de dados utilizada. As representações mapeadas do *Hq* de todos os objetos no conjunto de teste são armazenados para uso no processo de busca por similaridade.
- **A busca por similaridade:** A dissimilaridade entre dois objetos é determinada pela distância entre suas representações mapeadas na fase anterior. Para isto, é utilizada a distância euclidiana.

Nas seções a seguir são descritos os experimentos, e são exibidos os seus resultados, avaliados em termos de curvas Precisão-Revocação (curva média das consultas) e de valores de ganho acumulativo descontado normalizado (NDCG) (veja Seção 2.3.3). No primeiro experimento a avaliação é feita pelo teste *Bull's eye*.

## 7.2 Primeiro Experimento: Recuperação Usando a Base MPEG-7

No primeiro experimento de recuperação baseada em conteúdo, foi utilizada a base de dado *MPEG-7 CE Shape 1 Part-B*, descrita anteriormente na Seção 6.2. Os parâmetros para este experimento são fornecidos na Tabela 7.1. As imagens utilizadas têm dimensões  $256 \times 256$ .

Neste experimento, os resultados são comparados ao trabalho de Pedrosa et al., exposto em [Pedrosa et al. 2011], em que propõem um extrator de forma baseado em pontos de saliência, denominado D1. Os extratores *2Hq Feature* e D1 foram avaliados de acordo com o teste *Bull's eye* [Latecki et al. 2000], comumente utilizado na avaliação do desempenho em recuperação de imagens por conteúdo com a base de dados MPEG-7. Neste tipo de teste, cada imagem da base é utilizada como consulta, e para cada consulta, conta-se a quantidade de imagens relevantes entre os 40 objetos mais similares retornados. Para cada consulta, tem-se um máximo de 20 imagens relevantes (acertos). O teste *Bull's eye* mede a razão entre o número de objetos relevantes recuperados e a quantidade máxima possível de acertos que é de 28.000 (1.400 consultas, cada uma com um máximo de 20 acertos). A Figura 7.1 exibe, para cada classe da base MPEG-7, a taxa de recuperação do *2Hq Feature* e do D1 de acordo com teste *Bull's eye* aplicado aos extratores. A média alcançada para a base inteira foi de 0,5136, para o extrator D1, e de 0,8096 para o *2Hq Feature* (o ganho foi de 0,2960 em termos de precisão do teste *bull's eye*).

Tabela 7.1: Parâmetros utilizados no primeiro experimento de recuperação.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 2     |
| Família Wavelet                  | Haar  |
| Tamanho do $3Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 90    |

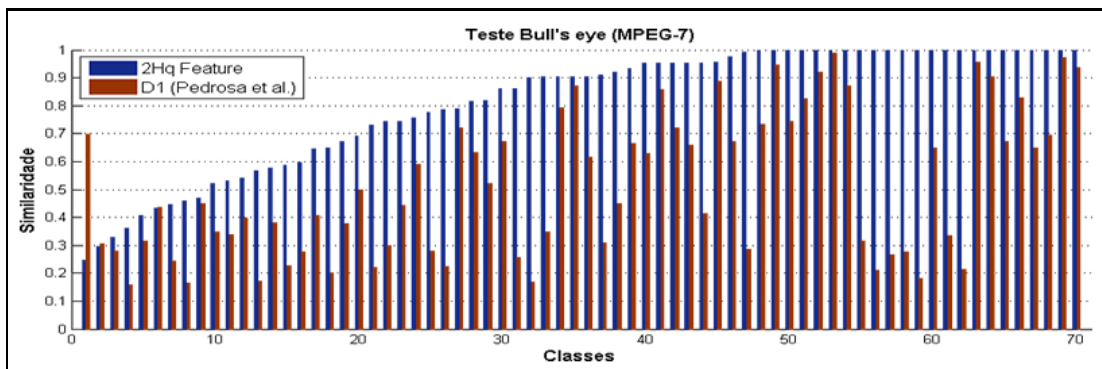


Figura 7.1: Resultados do teste *Bull's eye* para os extratores *2Hq Feature* e D1.

Na Figura 7.2 estão os 19 primeiros resultados de consultas realizadas neste experimento. Na primeira coluna de cada linha aparece o objeto da consulta. Os objetos de

consulta pertencem às classes *bone*, *hcircle*, *beetle*, *cattle*, *device2*, *flatfish*, *personal car* e *turtle*.

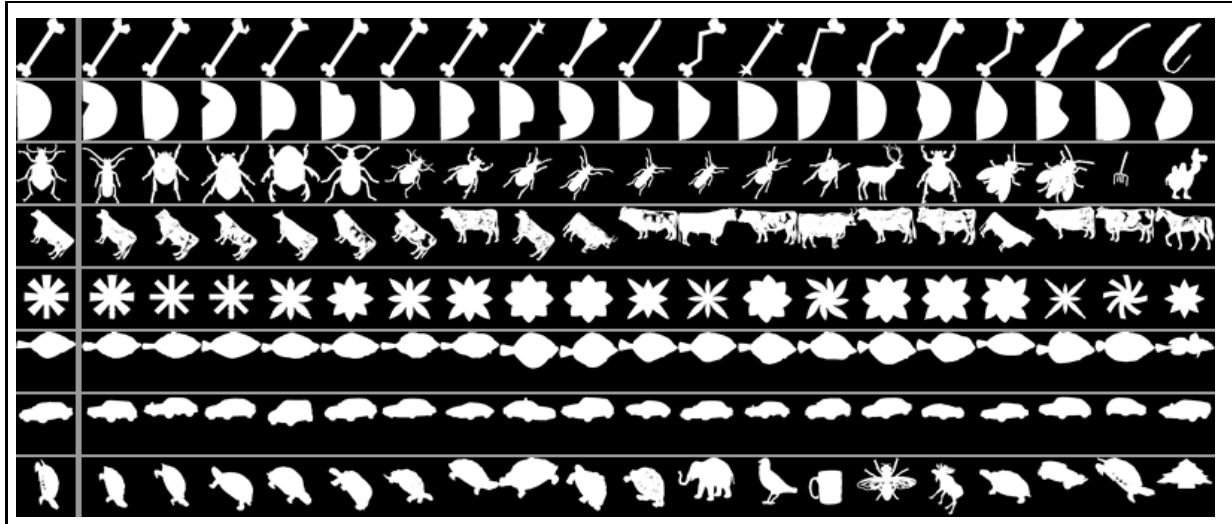


Figura 7.2: Consultas realizadas na base MPEG-7 (linhas correspondem a consultas, primeira coluna exibe objeto de consulta).

Como pode ser visto, o extrator *2Hq Feature* conseguiu excelentes resultados em algumas classes de imagens da base MPEG-7, além de resultados inferiores em outras classes. A Tabela 7.2 lista os valores obtidos para cada classe no teste *Bull's eye* aplicado ao extrator *2Hq Feature*. De um modo geral, o extrator teve um bom desempenho ao buscar por imagens similares, conforme mostrado na Figura 7.2. Na comparação com o extrator D1, proposto em [Pedrosa et al. 2011], verificou-se que o *2Hq Feature* superou o D1 em 67 classes.

Tabela 7.2: Resultado da recuperação de cada classe no teste *Bull's eye* aplicado ao extrator *2Hq Feature*.

|          |        |         |        |        |        |          |        |
|----------|--------|---------|--------|--------|--------|----------|--------|
| bootle   | 1,0000 | heart   | 1,0000 | rat    | 0,9050 | lmfish   | 0,5950 |
| brick    | 1,0000 | shoe    | 1,0000 | watch  | 0,9050 | ray      | 0,5875 |
| car      | 1,0000 | stef    | 1,0000 | crown  | 0,9000 | octopus  | 0,5775 |
| carriage | 1,0000 | teddy   | 1,0000 | apple  | 0,8600 | elephant | 0,5675 |
| cellular | 1,0000 | truck   | 1,0000 | dog    | 0,8600 | jar      | 0,5400 |
| children | 1,0000 | device5 | 0,9925 | pocket | 0,8200 | camel    | 0,5300 |
| chopper  | 1,0000 | pencil  | 0,9750 | bone   | 0,8150 | turtle   | 0,5225 |

Continua na próxima página

Continuação da página anterior

|          |        |              |        |         |        |           |        |
|----------|--------|--------------|--------|---------|--------|-----------|--------|
| device0  | 1,0000 | misk         | 0,9575 | comma   | 0,7900 | spoon     | 0,4675 |
| device1  | 1,0000 | flatfish     | 0,9525 | frog    | 0,7850 | chicken   | 0,4600 |
| device3  | 1,0000 | glas         | 0,9525 | classic | 0,7750 | sea snake | 0,4475 |
| device4  | 1,0000 | horseshoe    | 0,9525 | cup     | 0,7575 | deer      | 0,4325 |
| device6  | 1,0000 | key          | 0,9525 | guitar  | 0,7425 | butterfly | 0,4075 |
| device7  | 1,0000 | personal car | 0,9525 | hat     | 0,7425 | lizzard   | 0,3625 |
| device8  | 1,0000 | fork         | 0,9325 | fish    | 0,7300 | horse     | 0,3275 |
| device9  | 1,0000 | cattle       | 0,9200 | bat     | 0,6925 | bird      | 0,2950 |
| face     | 1,0000 | device2      | 0,9100 | spring  | 0,6725 | fly       | 0,2475 |
| fountain | 1,0000 | bell         | 0,9050 | tree    | 0,6500 |           |        |
| hcircle  | 1,0000 | hammer       | 0,9050 | beetle  | 0,6450 |           |        |

### 7.3 Segundo Experimento: Recuperação Usando a PSB

Para o segundo experimento deste capítulo foi utilizada a base de objetos do *Princeton Shape Benchmark* (PSB) [Shilane et al. 2004], formada por objetos adquiridos em diversas fontes pela Internet. A base contém 1.814 objetos tridimensionais, divididos em um total de 161 classes de diferentes tamanhos.

Os objetos da PSB são definidos como malhas de triângulos; portanto, tiveram que ser convertidos em volumes de *voxels* a fim de se aplicar o extrator *3Hq Feature*. Cada objeto convertido passou a ter então as dimensões  $256 \times 256 \times 256$  *voxels*. A Figura 7.3 apresenta uma amostra dos objetos pertencentes a PSB (já convertidos a volumes de *voxels*). A Tabela 7.3 lista os parâmetros utilizados neste experimento.

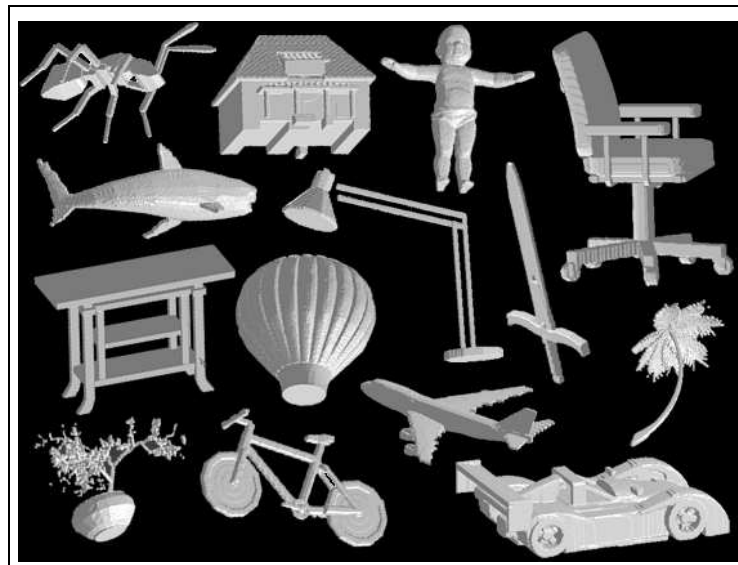


Figura 7.3: Amostra dos objetos da base PSB utilizados no segundo experimento.

Tabela 7.3: Parâmetros utilizados no segundo experimento.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 1     |
| Família Wavelet                  | Haar  |
| Tamanho do $3Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 180   |

O PSB já fornece os conjuntos de treinamento e de teste, em que divide a base de objetos igualmente. Para este experimento, usando o sistema CBIR projetado como descrito na Seção 7.1, o conjunto de validação, que soma 20% da base, foi retirado do conjunto de testes fornecido pela PSB. Os resultados obtidos na recuperação de objetos são comparados a outros métodos do estado-da-arte. A Figura 7.4 traz a curva de precisão-revocação obtida na recuperação com o uso do extrator  $3Hq$  Feature e outros extratores reportados em [Shilane et al. 2004]. Na Figura 7.5, a comparação é feita com os métodos NMF, DSR, DBF, ICA e PCA descritos em [Dutagaci et al. 2010]. Na tabela 7.4 são exibidos os valores de NDCG obtidos para cada extrator. O ganho mínimo verificado, em termos de valores NDCG, foi de 0.103.

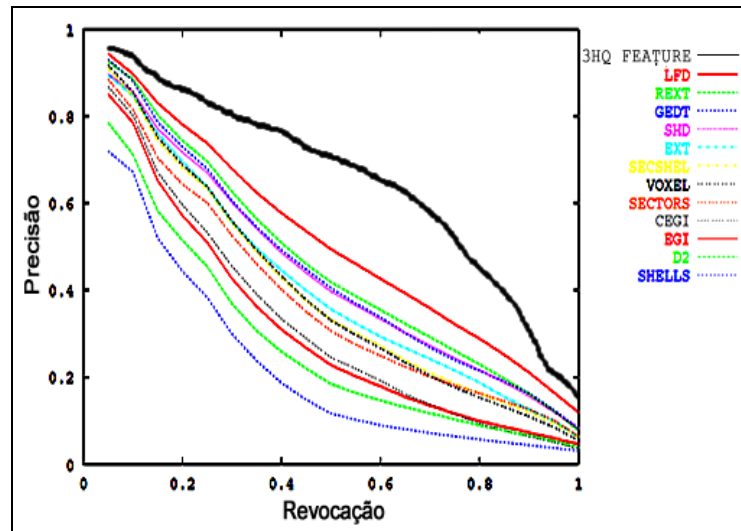


Figura 7.4: Curvas precisão-revocação comparando o  $3Hq$  Feature com os extratores definidos em [Shilane et al. 2004].

A Figura 7.6 ilustra os dez primeiros resultados para as consultas realizadas no experimento. Cada linha corresponde a uma consulta realizada. Na primeira coluna é representada a imagem consultada.

Como pode ser visto, o extrator  $3Hq$  Feature mostrou um desempenho, em termos de curvas de Precisão-Revocação e valores NDCG, melhor que outros métodos do estado-da-arte. A Tabela 7.5 lista as 40 classes em que o extrator proposto teve melhor desempenho. A Tabela 7.6 relaciona as 40 classes com os menores valores de NDCG obtidos. A Figura 7.7 ilustra os valores de NDCG alcançados para cada classe da base PSB usando o extrator  $3Hq$  Feature.

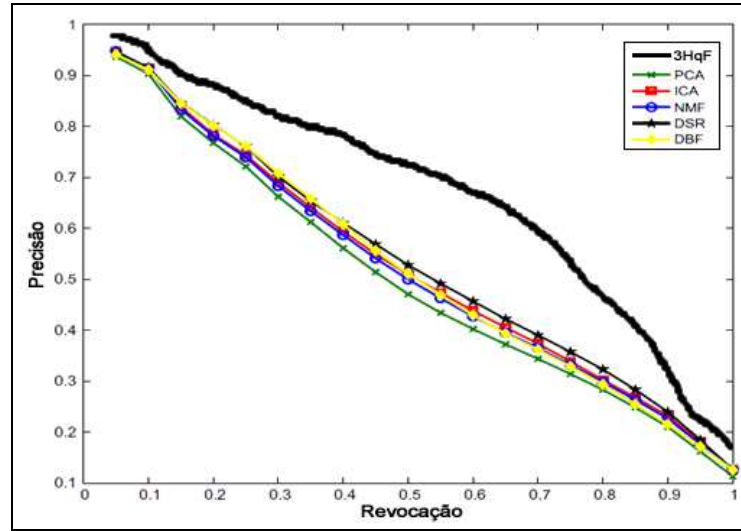


Figura 7.5: Curvas precisão-revocação comparando o *3Hq Feature* com os extratores definidos em [Dutagaci et al. 2010].

Tabela 7.4: Valores de DCG obtidos no segundo experimento.

| Método extrator                            | NDCG  |
|--------------------------------------------|-------|
| <b>3Hq Feature</b>                         | 0.771 |
| CRSP [Papadakis et al. 2007]               | 0.668 |
| DSR [Vranić 2004]                          | 0.665 |
| DBF [Akgül et al. 2009, Akgül et al. 2007] | 0.659 |
| ICA [Dutagaci et al. 2010]                 | 0.655 |
| NMF [Dutagaci et al. 2010]                 | 0.650 |
| LFD [Chen e Ouhyoung 2002]                 | 0.643 |
| PCA [Dutagaci et al. 2010]                 | 0.634 |
| REXT [Vranić 2003]                         | 0.601 |

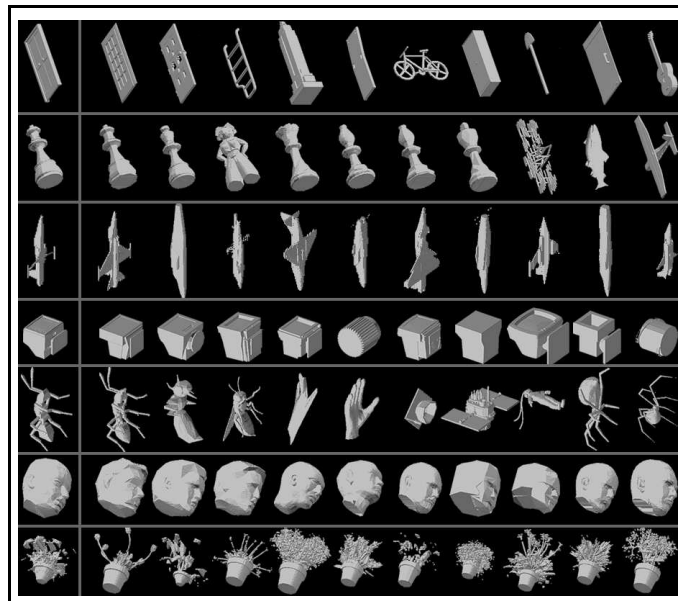


Figura 7.6: Resultado de consultas realizadas na PSB (linhas correspondem a consultas, primeira coluna exibe objetos de consulta).

Tabela 7.5: Classes com melhor desempenho, em termos de valores NDCG, usando o  $3Hq$  Feature.

|                                  |        |                          |        |
|----------------------------------|--------|--------------------------|--------|
| vase liquid container            | 1,0000 | wrench tool              | 0,9610 |
| axe blade                        | 1,0000 | chess piece              | 0,9606 |
| gear wheel                       | 1,0000 | covered wagon vehicle    | 0,9584 |
| screwdriver tool                 | 1,0000 | desk lamp lamp           | 0,9540 |
| submarine sea vessel             | 1,0000 | ice cream                | 0,9505 |
| book                             | 1,0000 | church building          | 0,9482 |
| slot machine                     | 1,0000 | sink                     | 0,9482 |
| dome church building             | 0,9960 | desk with hutch desk     | 0,9413 |
| fireplace                        | 0,9919 | butcher knife blade      | 0,9378 |
| hat                              | 0,9909 | chess set                | 0,9361 |
| pickup truck vehicle             | 0,9865 | glider airplane          | 0,9295 |
| glass with stem liquid container | 0,9834 | dragon fantasy animal    | 0,9283 |
| bench seat                       | 0,9826 | train car train          | 0,9270 |
| newtonian toy                    | 0,9814 | eyeglasses               | 0,9237 |
| tv display device                | 0,9781 | school desk furniture    | 0,9132 |
| hot air balloon balloon vehicle  | 0,9775 | gazebo building          | 0,9111 |
| computer monitor display device  | 0,9774 | helicopter aircraft      | 0,9102 |
| potted plant plant               | 0,9673 | palm tree                | 0,9099 |
| barren tree                      | 0,9653 | round table              | 0,9030 |
| mailbox                          | 0,9621 | electrical guitar guitar | 0,9024 |

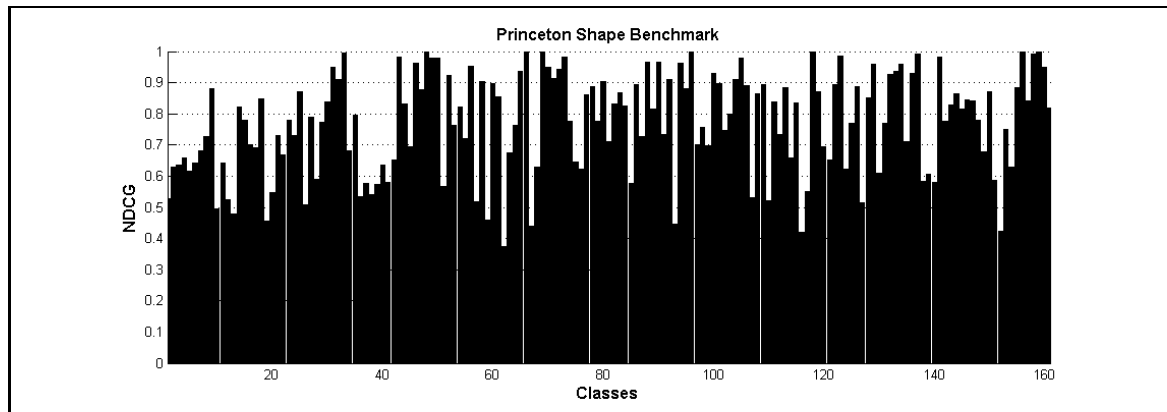
Tabela 7.6: As 40 classes com os menores valores de NDCG obtidos no experimento.

|                    |        |                             |        |
|--------------------|--------|-----------------------------|--------|
| snake animal       | 0,6410 | sailboat with oars sailboat | 0,5508 |
| one peak tent tent | 0,6361 | walking human               | 0,5476 |
| butterfly insect   | 0,6359 | skyscraper building         | 0,5410 |
| bee insect         | 0,6303 | two story home building     | 0,5332 |
| tire wheel         | 0,6300 | enterprise like spaceship   | 0,5326 |
| desktop computer   | 0,6275 | ant insect                  | 0,5272 |
| dining chair chair | 0,6237 | dog quadruped               | 0,5254 |
| race car car       | 0,6229 | space shuttle spaceship     | 0,5221 |

Continua na próxima página

Continuação da página anterior

|                         |        |                           |        |
|-------------------------|--------|---------------------------|--------|
| standing bird bird      | 0,6147 | streetlight lamp          | 0,5192 |
| semi vehicle            | 0,6093 | antique car car           | 0,5136 |
| ladder                  | 0,6072 | hand body part            | 0,5077 |
| torso body part         | 0,5887 | shark underwater creature | 0,4952 |
| snowman                 | 0,5852 | feline quadruped          | 0,4798 |
| helmet hat              | 0,5834 | acoustic guitar guitar    | 0,4603 |
| multiple peak tent tent | 0,5814 | human arms out human      | 0,4573 |
| city                    | 0,5813 | shovel tool               | 0,4462 |
| barn building           | 0,5775 | wheel                     | 0,4397 |
| bush plant              | 0,5764 | phone handle              | 0,4246 |
| lighthouse building     | 0,5720 | large sail boat sailboat  | 0,4187 |
| microchip               | 0,5673 | rifle gun                 | 0,3743 |

Figura 7.7: Valores de NDCG obtidos para cada classe usando o *3Hq Feature*.

## 7.4 Terceiro Experimento: Recuperação Usando a Base W3BC-111

Para o terceiro experimento, o extrator *3Hq Feature* foi aplicado à base sintética W3BC-111 de objetos espiculados e arredondados, descrita anteriormente na Seção 6.5. Os parâmetros utilizados são dados na Tabela 7.7. Os objetos da base têm dimensões  $256 \times 256 \times 256$ . A curva Precisão-Revocação obtida neste experimento é dada na Figura 7.8. O valor NDCG foi de 0.9971.

A Figura 7.9 ilustra os dez primeiros resultados para consultas na base de dados W3BC-111. O objeto de consulta é exibido na primeira coluna da figura.

Como se vê neste experimento, o extrator *3Hq Feature* conseguiu obter uma precisão ideal até pouco mais de 80% de revocação. O valor de NDCG obtido neste experimento foi de 0,9971. Os resultados verificados sugerem o *3Hq Feature* como uma boa ferramenta para a recuperação de objetos espiculados e circunscritos.

Tabela 7.7: Parâmetros utilizados no terceiro experimento de recuperação.

| Parâmetro                        | Valor |
|----------------------------------|-------|
| Nível de decomposição ( $L$ )    | 1     |
| Família Wavelet                  | Haar  |
| Tamanho do $3Hq_i$               | 256   |
| Neurônios nas camadas escondidas | 20    |

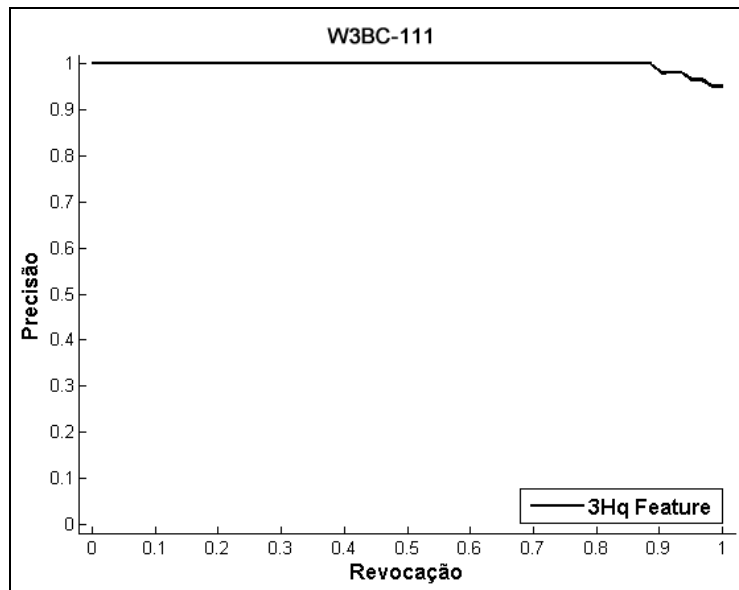
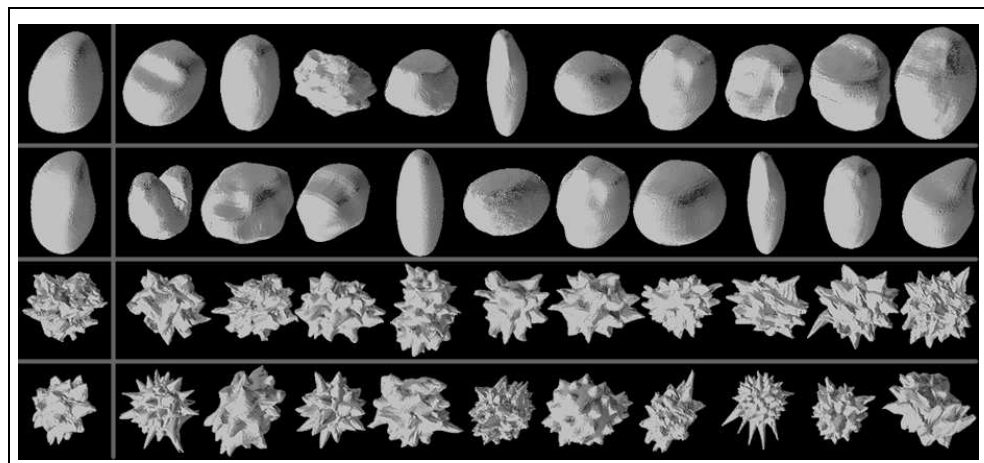
Figura 7.8: Curva precisão-revocação obtida na aplicação do  $3Hq$  Feature sobre a base W3BC-111.

Figura 7.9: Resultado de consultas realizadas na base W3BC-111 (linhas correspondem a consultas, primeira coluna exibe objetos de consulta).

## 7.5 Considerações Finais

Neste capítulo foram apresentados alguns experimentos realizados no contexto da recuperação baseada em conteúdo. Os resultados foram avaliados em termos de curvas Precisão-Revocação e de valores NDCG, mostrando um bom desempenho dos extratores propostos. O primeiro e segundo experimentos mostraram que tais extratores são capazes

de um bom desempenho na recuperação de objetos em geral. Na base de imagens MPEG-7, 23 classes foram capazes de recuperar todas as imagens relevantes no teste *Bull's eye*, e somente 16 classes obtiveram uma recuperação abaixo de 0,60. Para a base PSB percebeu-se que, dentre as 40 classes com melhor desempenho na recuperação, a maioria consiste de objetos manufaturados (em contraste com objetos encontrados na natureza). No terceiro experimento, usando uma base de dados com objetos típicos de algumas aplicações médicas, podemos notar um excelente desempenho do extrator *3Hq Feature*.

Pode-se concluir, portanto, que os extratores *2Hq Feature* e *3Hq Feature*, em conjunto com o sistema de recuperação baseado em redes neurais, mostram bons desempenhos e até superam, em termos de curvas Precisão-Revocação e valores DCG, alguns métodos do estado-da-arte.

# Capítulo 8

## Conclusões e Trabalhos Futuros

Neste capítulo, são apresentadas as conclusões obtidas à partir dos experimentos realizados ao longo do presente trabalho, no que diz respeito ao *framework* e aos extratores de forma desenvolvidos. Ao final do capítulo, são sugeridos possíveis novos experimentos e linhas de investigação futuras.

### 8.1 Conclusões

Este trabalho apresentou diversas contribuições: primeiramente, foi desenvolvido um *framework*, denominado WESS, para realizar desenhos de contornos e reconstrução tridimensional. O *framework* consiste do aplicativo EMISS, que permite desenhar e salvar contornos em imagens, e dos métodos de interpolação SIMOL e SIDITRANS, úteis na tarefa de reconstrução tridimensional dos contornos obtidos. Foi visto que os métodos de interpolação propostos apresentaram bons resultados e superaram as limitações presentes nos dois métodos tidos como inspiração para o desenvolvimento do SIMOL e do SIDITRANS. Também foi visto que, entre os métodos de interpolação propostos, o SIMOL, embora seja mais rápido (executou em um tempo 16,6 menor), apresentou mais erros de interpolação (cerca de 2,38 vezes mais de erro) com relação ao método SIDITRANS.

Outra contribuição deste trabalho, e de maior peso, foi o desenvolvimento de extratores *2Hq Feature* e *3Hq Feature*, aplicados para a caracterização de formas em silhuetas 2D e 3D, respectivamente. Os descritores definem operações de redimensionamento e decomposição e fazem uso da curva de Hilbert e da transformada discreta de wavelet.

Também foi possível observar bons resultados nos experimentos utilizando a base de contornos de lesões de mama e a base sintética W3BC-111 de objetos 3D criados a partir da base BC-111 (nos experimentos realizados os métodos propostos superaram outros métodos do estado-da-arte). Estes experimentos tornam promissor o uso dos extratores de forma, propostos neste trabalho, para caracterizar formas, tanto bi quanto tridimensionais, com o objetivo de apoiar o diagnóstico de câncer de mama.

Outro fator que pôde ser verificado durante a realização dos experimentos é da quantidade de neurônios nas camadas escondidas da rede neural que, para cada experimento, variou conforme o número de classes da base de dados utilizada. Assim, bases com relativamente poucas classes, como a Kimia-99, Kimia-216, BC-111 e W3BC-111, tiveram ótimos resultados utilizando 20 neurônios em cada camada escondida da rede neural. Por outro lado, para as bases MPEG-7 e PSB, bons resultados surgiram quando a quantidade de neurônios nas camadas escondidas aumentou para 90 e 180, respectivamente.

É importante notar também o ganho obtido com relação à redução da quantidade de dados, principalmente nos experimentos com bases de objetos 3D. O uso da rede neural contribuiu ainda mais com a redução de dados dos descritores, obtendo uma representação bastante compacta para os objetos e imagens usados nos experimentos.

## 8.2 Trabalhos Futuros

A análise de resultados obtidos nos experimentos realizados neste trabalho mostrou a superioridade, com relação às medidas de avaliação utilizadas, dos métodos extratores propostos em comparação com outros métodos do estado-da-arte. Porém algumas melhorias podem ser realizadas em trabalhos futuros.

Os extratores propostos são invariantes a transformação de translação e, como pôde ser visto nos experimentos realizados, embora o *2Hq Feature* seja sensível a transformações de rotação, o uso de algum método de normalização de pose pode tornar os métodos propostos invariantes à rotação. Os experimentos realizados mostram a robustez do *2Hq Feature* com relação a efeitos de oclusão, articulação e deformações. Também pode ser considerada a criação de novas bases de dados para testes de escalabilidade<sup>1</sup>.

Os extratores propostos realizam a decomposição de imagens, ou objetos 3D, em diversas sub-regiões, que são processadas de maneira independente. Uma possível melhoria no tempo de execução para a obtenção dos extratores consiste no uso de técnicas de processamento paralelo. Uma vez que cada sub-região seja processada em paralelo, o ganho no tempo de execução pode ser bastante proveitoso no caso do *3Hq Feature*, devido à grande quantidade de dados presente em um volume de *voxels*.

O uso da rede neural para a redução da dimensionalidade (abordagem utilizada nos experimentos de classificação apresentados no Capítulo 7) faz com que os extratores propostos se tornem métodos supervisionados. Em uma tentativa de se obterem métodos não-supervisionados, pode-se verificar o efeito do uso de uma rede neural não supervisionada.

Uma vez que o extrator *3Hq Feature* não trabalha diretamente com a representação geométrica da forma, pode ser estudada a aplicação do extrator para a caracterização de

---

<sup>1</sup>A palavra “escalabilidade” não existe oficialmente no idioma Português, porém é utilizada aqui como tradução de *scalability*.

texturas. Para isto, serão necessárias pequenas modificações nos métodos de reconstrução tridimensional, a fim de levar em conta informações de textura presente nos *slices*. Também pode ser estudado, nesta possível abordagem para caracterização de texturas, o uso dos coeficientes de detalhes obtidos na transformada discreta de wavelet (no trabalho desenvolvido nesta dissertação foram utilizados apenas os coeficientes de aproximação para efeito de subamostragem de dados).

Analisando os experimentos realizados neste trabalho, nota-se que, para um bom desempenho nos testes, pode-se variar a quantidade de neurônios necessária nas camadas escondidas da rede neural. Embora este trabalho tenha utilizados valores empíricos para este parâmetro, trabalhos futuros certamente serão conduzidos com o objetivo de encontrar uma maneira de calcular o parâmetro em questão (acreditando que o valor ideal seja dependente da quantidade de classes e/ou elementos presentes em cada base de dados).

# Referências Bibliográficas

- [MIA 2004] (accessed June, 2004). The Mammographic Image Analysis Society digital mammogram database.
- [Akgül et al. 2007] Akgül, C., Sankur, B., Yemez, Y., e Schmitt, F. (2007). Density-Based 3D Shape Descriptors. *Advances in Signal Processing*, 2007.
- [Akgül et al. 2009] Akgül, C., Sankur, B., Yemez, Y., e Schmitt, F. (2009). 3D Model Retrieval Using Probability Density-Based Shape Descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(6):1117–1133.
- [Alberta Cancer Board, Alberta, Canada 2004] Alberta Cancer Board, Alberta, Canada (2004). Screen Test: Alberta Program for the Early Detection of Breast Cancer – 2001/03 Biennial Report.
- [Alto et al. 2005] Alto, H., Rangayyan, R., e Desautels, J. (2005). Content-Based Retrieval and Analysis of Mammographic Masses. *Journal of Electronic Imaging*.
- [American College of Radiology, Reston, VA 1998] American College of Radiology, Reston, VA (1998). Illustrated Breast Imaging Reporting and Data System BI-RADS<sup>TM</sup>.
- [Armstrong et al. 2009] Armstrong, J., Ahmed, M., e Chau, S.-C. (2009). A Rotation-Invariant Approach to 2D Shape Representation Using the Hilbert Curve. *Image Analysis and Recognition*, 5627:594–603.
- [Arof e Deravi 1998] Arof, H. e Deravi, F. (1998). Circular Neighborhood and 1-D DFT Features for Texture Classification and Segmentation. *Vision Image and Signal Processing*, 145(3):167–172.
- [Bai e Latecki 2008] Bai, X. e Latecki, L. (2008). Path Similarity Skeleton Graph Matching. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(7):1282–1292.
- [Bai et al. 2007] Bai, X., Latecki, L., e Liu, W.-Y. (2007). Skeleton Pruning by Contour Partitioning With Discrete Curve Evolution. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(3):449–462.
- [Bai et al. 2010] Bai, X., Liu, W., e Tu, Z. (2010). Integrating Contour and Skeleton for Shape Classification. *IEEE Computer Vision Workshops*, pp. 360–367.
- [Barcelos et al. 2008] Barcelos, C., Ribeiro, E., e Batista, M. (2008). Image Characterization via Multilayer Neural Networks. *International Conference on Tools with Artificial Intelligence*, pp. 325–332.

- [Bartolini e Ciaccia 2005] Bartolini, I. e Ciaccia, P. (2005). WARP - Accurate Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance. *IEEE Trans. Pattern Anal. Mach. Intell.*
- [Berndt e Clifford 1994] Berndt, D. e Clifford, J. (1994). Using Dynamic Time Warping to Find Patterns in Time Series. *Knowledge Discovery in Databases*, pp. 359–370.
- [Blunsden e Atallah 2005] Blunsden, S. e Atallah, L. (2005). Investigating the Effects of Scale in MRF Texture Classification. *IEEE Conf. on Visual Information Engineering*.
- [Bradley 1997] Bradley, A. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition Letters*.
- [Bustos et al. 2005] Bustos, B., Keim, D., Saupe, D., Schreck, T., e Vranić, D. (2005). Feature-Based Similarity Search in 3D Object Databases. *ACM Computing Surveys*, 37(4):345–387.
- [Canlin et al. 2008] Canlin, M., Min, C., e Shaoyi, G. (2008). Texture Synthesis with Self-Correlation Distinguishing and Its Applications. *Computer Research and Development*.
- [Chanda e Majumder 2006] Chanda, B. e Majumder, D. (2006). *Digital Image Processing And Analysis*. Prentice-Hall.
- [Chen e Ouhyoung 2002] Chen, D.-Y. e Ouhyoung, M. (2002). A 3D Object Retrieval System Based on Multi-Resolution Reeb Graph. *Proc. of Computer Graphics Workshop*, pp. 16–20.
- [Chen et al. 2003] Chen, D.-Y., Tian, X.-P., Shen, Y.-T., e Ouhyoung, M. (2003). On Visual Similarity Based 3D Model Retrieval. *Computer Graphics Forum*, 22(3):223–232.
- [Chen e Wang 2002] Chen, Y. e Wang, J. (2002). A Region-Based Fuzzy Feature Matching Approach To Content-Based Image Retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.*
- [Chui 1992] Chui, C. (1992). *An Introduction to Wavelets*. Academic Press, San Diego.
- [Croft et al. 2009] Croft, W., Metzler, D., e Strohman, T. (2009). *Search Engines: Information Retrieval in Practice*. Addison Wesley.
- [Davis e Goadrich 2006] Davis, J. e Goadrich, M. (2006). The Relationship Between Precision-Recall and ROC Curves. *International Conference on Machine Learning*.
- [Debnath 2002] Debnath, L. (2002). *Wavelet Transforms and Their Applications*. Birkhäuser.
- [Duda et al. 2000] Duda, R., Hart, P., e Stork, D. (2000). *Pattern Classification*. John Wiley & Sons, Inc., 2nd edition.
- [Dutagaci et al. 2010] Dutagaci, H., Sankur, B., e Yemez, Y. (2010). Subspace Methods for Retrieval of General 3D Models. *Computer Vision and Image Understanding*, 114(8):865–886.

- [D.V. Vranić 2001] D.V. Vranić, D. Saupe, J. R. (2001). Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics. *Multimedia Signal Processing*, pp. 293–298.
- [Eakins e Graham 1999] Eakins, J. e Graham, M. (1999). Content-based Image Retrieval. Technical Report, University of Northumbria at Newcastle.
- [Ebrahim et al. 2008] Ebrahim, Y., Ahmed, M., Abdelsalam, W., e Chau, S.-C. (2008). Shape representation and description using the Hilbert curve. *Pattern Recognition Letters*, 30(4):348–358.
- [El-ghazal et al. 2007] El-ghazal, A., Basir, O., e Belkasim, S. (2007). A New Shape Signature Fof Fourier Descriptors. *IEEE International Conference on Image Processing*, pp. 161–164.
- [Fausett 1993] Fausett, L. (1993). *Fundamentals of Neural Networks: Architectures, Algorithms And Applications*. Prentice Hall.
- [Fawcett 2005] Fawcett, T. (2005). An introduction to ROC analysis. *Pattern Recognition Letters*.
- [Fernandes 2009] Fernandes, B. (2009). Segmentação e Classificação de Padrões Visuais Baseados em Campos Receptivos e Inibitórios. Master's thesis, Universidade Federal de Pernambuco.
- [Gonzalez e Woods 2008] Gonzalez, R. e Woods, R. (2008). *Digital Image Processing*. Pearson Prentice Hall, 3rd edition.
- [Gotlieb e Kreyszig 1990] Gotlieb, C. e Kreyszig, H. (1990). Texture Descriptors Based on Co-occurrence Matrices. *Computer Vision, Graphics, and Image Processing*, 51(1):70–86.
- [Grigorescu et al. 2002] Grigorescu, S., Petkov, N., e Kruizinga, P. (2002). Comparison of Texture Features Based on Gabor Filters. *IEEE Transactions On Image Processing*, 11(10):1160–1167.
- [Guliatto et al. 2010] Guliatto, D., Oliveira, W., e Traina Jr., C. (2010). A New Feature Descriptor Derived From Hilbert Space-Filling Curve to Assist Breast Cancer Classification. *Computer-Based Medical Systems*.
- [Guliatto et al. 2008a] Guliatto, D., Rangayyan, R., Carvalho, J., e Santiago, S. (2008a). Feature Extraction From the Turning Angle Functoin for the Classification of Contours of Breast Tumors. *Journal of Digital Imaging*, 21(2).
- [Guliatto et al. 2008b] Guliatto, D., Rangayyan, R., Carvalho, J., e Santiago, S. (2008b). Polygonal Modeling of Contours of Breast Tumors With the Preservation of Spicules. *Journal of Electronic Imaging*, 17(2).
- [Guo et al. 1995] Guo, J.-F., Cai, Y.-L., e Wang, Y.-P. (1995). Morphology-Based Interpolation for 3D Medical Image Reconstruction. *Comuterized Medical Imaging and Graphics*, 19(3):267–279.

- [Hanley e McNeil 1982] Hanley, J. e McNeil, B. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*.
- [Harvey 1994] Harvey, R. (1994). *Neural Network Principles*. Prentice Hall.
- [Homer 1997] Homer, M. (1997). *Mammographic Interpretation: A Practical Approach*. McGraw-Hill, Boston, MA, 2nd edition.
- [Horn 1984] Horn, B. (1984). Extended Gaussian Images. *Proc. of the IEEE*, 72:1671–1686.
- [Huang et al. 1997] Huang, J., Kumar, S., Mitra, M., Zhu, W.-J., e Zabih, R. (1997). Image Indexing Using Color Correlograms. *EEE Conference on Computer Vision and Pattern Recognition*, pp. 762–768.
- [Ip et al. 2002] Ip, C., Lapadat, D., Sieger, L., e Regli, W. (2002). Using Shape Distributions to Compare Solid Models. *ACM symposium on Solid modeling and applications*, pp. 273–280.
- [Jähne 1997] Jähne, B. (1997). *Practical Handbook On Image Processing For Scientific Applications*. CRC Press LLC.
- [Kang e Ikeuchi 1994] Kang, S. e Ikeuchi, K. (1994). The Complex EGI: A New Representation for 3-D Pose Determination. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(7):707–721.
- [Kasabov 1998] Kasabov, N. (1998). *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*. The MIT Press.
- [Kazhdan et al. 2003] Kazhdan, M., Funkhouser, T., e Rusinkiewicz, S. (2003). Rotation Invariant Spherical Harmonic Representation of 3D Shape Descriptors. *Symposium on Geometry Processing*.
- [Khotanzad e Hong 1990] Khotanzad, A. e Hong, Y. (1990). Invariant Image Recognition by Zernike Moments. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(5):489–497.
- [Kim e Kim 2000] Kim, H.-K. e Kim, J.-D. (2000). Region-Based Shape Descriptor Invariant to Rotation, Scale and Translation. *Signal Processing: Image Communication*, 16:87–93.
- [Krzanowski e Hand 2009] Krzanowski, W. e Hand, D. (2009). *ROC Curves for Continuous Data*. CRC Press.
- [Latecki e Lakamper 2000] Latecki, L. e Lakamper, R. (2000). Shape Similarity Measure Based on Correspondence of Visual Parts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(10):1185–1190.
- [Latecki et al. 2000] Latecki, L., Lakamper, R., e Eckhardt, U. (2000). Shape Descriptors for Non-rigid Shapes with a Single Closed Contour. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1:424–429.
- [Luo e Hancock 1997] Luo, B. e Hancock, E. (1997). Slice Interpolation Using the Distance Transform and Morphing. *Digital Signal Processing Proceedings*, 2:1083–1086.

- [Mallat 1987] Mallat, S. (1987). A Compact Multiresolution Representation: The Wavelet Model. In *CVWS87*.
- [Manjunath 2002] Manjunath, B. (2002). *Introduction to MPEG-7: Multimedia Content Description Interface*. John Wiley & Sons Ltd.
- [Manning et al. 2009] Manning, C., Raghavan, P., e Schütze, H. (2009). *An Introduction to Information Retrieval*. Cambridge University Press.
- [Marzal et al. 2006] Marzal, A., Palazón, V., e Peris, G. (2006). Shape Retrieval Using Normalized Fourier Descriptors Based Signatures and Cyclic Dynamic Time Warping. *Lecture Notes in Computer Science*.
- [Migeon et al. 1998] Migeon, B., Charreyron, R., Deforge, P., e marché, P. (1998). Improvement of Morphology-Based Interpolation. *International Conference of the IEEE Engineering in Medicine and Biology Society*, 20(2):585–587.
- [Mor e Aladjem 2005] Mor, E. e Aladjem, M. (2005). Boundary refinements for wavelet-domain multiscale texture segmentation. *Image and Vision Computing*, 23(13):1150–1158.
- [Nguyen e Rangayyan 2006] Nguyen, T. e Rangayyan, R. (2006). Shape analysis of breast masses in mammograms via the fractal dimension. *Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- [Nonato et al. 2005] Nonato, L., Vargas, A., Minghim, R., e de Oliveira, M. (2005). Beta-Connection: Generating a Family of Models from Planar Cross Sections. *ACM Transactions on Graphics*, 24(4):1239–1258.
- [Novotni e Klein 2004] Novotni, M. e Klein, R. (2004). Shape Retrieval Using 3D Zernike Descriptors. *Computer-Aided Design*, 36:1047–1062.
- [Osada et al. 2001] Osada, R., Funkhouser, T., Chazelle, B., e Dobkin, D. (2001). Matching 3D Models with Shape Distributions. *Shape Modeling International*, pp. 154–166.
- [Osada et al. 2002] Osada, R., Funkhouser, T., Chazelle, B., e Dobkin, D. (2002). Shape Distributions. *ACM Transactions on Graphics*, 21(4):807–832.
- [Pan et al. 2011] Pan, X., You, Q., Liu, Z., e Chen, Q. (2011). 3D Shape Retrieval by Poisson Histogram. *Pattern Recognition Letters*.
- [Papadakis et al. 2007] Papadakis, P., Pratikakis, I., Perantonis, S., e Theoharis, T. (2007). Efficient 3D Shape Matching and Retrieval Using a Concrete Radialized Spherical Projection Representation. *Pattern Recognition Letters*, 40:2437–2452.
- [Park e Chen 2001] Park, B. e Chen, Y. R. (2001). Co-occurrence Matrix Texture Features of Multi-spectral Images on Poultry Carcasses. *J. Agric. Engng Res.*, 78(2):127–139.
- [Pedrini e Schwartz 2007] Pedrini, H. e Schwartz, W. (2007). *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. Editora Thomson Learning Edições Ltda.

- [Pedrosa et al. 2011] Pedrosa, G., Barcelos, C., e Batista, M. (2011). An Image Retrieval System Using Shape Saliency Points. *In IEEE International Symposium on Circuits and Systems*.
- [Pedrosa et al. 2008] Pedrosa, G., Santos, C., Batista, M., Fernandes, H., e Barcelos, C. (2008). An Effective Saliency-based Algorithm for Shape Retrieval. *Advanced Concepts for Intelligent Vision Systems*, pp. 818–827.
- [Pratt 2007] Pratt, W. (2007). *Digital Image Processing - PIKS scientific Inside*. John Wiley & Sons, Inc., 4th edition.
- [Prusinkiewicz e Lindenmayer 2004] Prusinkiewicz, P. e Lindenmayer, A. (2004). *The Algorithmic Beauty of Plants*. Springer-Verlag.
- [Rangayyan et al. 1997] Rangayyan, R., El-Faramawy, N., Desautels, J., e Alim, O. (1997). Measures of Acutance and Shape for Classification of Breast Tumors. *IEEE Transactions on Medical Imaging*, 16(6):799–810.
- [Rangayyan et al. 2000] Rangayyan, R., Mudigonda, N., e Desautels, J. (2000). Boundary Modelling and Shape Analysis Methods for Classification of Mammographic Masses. *Medical and Biological Engineering and Computing*.
- [Rangayyan e Nguyen 2007] Rangayyan, R. e Nguyen, T. (2007). Fractal Analysis of Contours of Breast Masses in Mammograms. *Journal of Digital Imaging*, 20(3):223–237.
- [Razente 2009] Razente, H. (2009). *Adequando Consultas Por Similaridade Para Reduzir A Descontinuidade Semântica Na Recuperação De Imagens Por Conteúdo*. PhD thesis, Universidade de São Paulo.
- [Ribeiro 2009] Ribeiro, E. (2009). *Caracterização de Imagens Utilizando Redes Neurais Artificiais*. Master's thesis, Universidade Federal de Uberlândia.
- [Rojas 1996] Rojas, R. (1996). *Neural Networks - A Systematic Introduction*. Springer-Verlag.
- [Russ 2006] Russ, J. (2006). *The Image Processing Handbook*. CRC Press, 5th edition.
- [Sandler e Lindenbaum 2006] Sandler, R. e Lindenbaum, M. (2006). Gabor Filter Analysis for Texture Segmentation. *Computer Vision and Pattern Recognition Workshop*.
- [Sebastian et al. 2004] Sebastian, T., Klein, P., e Kimia, B. (2004). Recognition of Shapes by Editing Their Shock Graphs. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(5):550–571.
- [Selvan et al. 2009] Selvan, S., Kumar, J., Manickam, I., e Sundareshbaran, A. (2009). Texture Classification Using SVD and Wavelet Transform. *RTCS*, pp. 217–221.
- [Selvan e Ramakrishnan 2007] Selvan, S. e Ramakrishnan, S. (2007). SVD-Based Modeling for Image Texture Classification Using Wavelet Transformation. *IEEE Trans. On Image Processing*, 16(11):2688–2696.

- [Shen et al. 1994] Shen, L., Rangayyan, R., e Desautels, J. (1994). Application of Shape Analysis to Mammographic Calcifications. *IEEE Trans. On Medical Imaging*, 13(2):263–274.
- [Shilane et al. 2004] Shilane, P., Min, P., Kazhdan, M., e Funkhouser, T. (2004). The Princeton Shape Benchmark. *Shape Modeling International*.
- [Stehling et al. 2003] Stehling, R. O., Nascimento, M. A., e Falcão, A. X. (2003). Cell Histograms Versus Color Histograms For Image Representation And Retrieval. *Knowledge and Information Systems*, 5(3):315–336.
- [Strang e Nguyen 1996] Strang, G. e Nguyen, T. (1996). *Wavelets and Filter Banks*. Wellesley Cambridge.
- [Suckling et al. 1994] Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, J., Doggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S., Taylor, P., Betal, D., e Savage, J. (1994). The Mammographic Image Analysis Society Digital Mammogram Database.
- [Sun et al. 2006] Sun, J., Zhang, X., Cui, J., e Zhou, L. (2006). Image Retrieval Based on Color Distribution Entropy. *Pattern Recognition Letters*, 27(10):1122–1126.
- [Tangelder e Veltkamp 2003] Tangelder, J. e Veltkamp, R. (2003). Polyhedral Model Retrieval Using Weighted Point Sets. *Shape Modeling International*, pp. 119–129.
- [Tangelder e Veltkamp 2008] Tangelder, J. e Veltkamp, R. (2008). A Survey of Content Based 3D Shape Retrieval Methods. *Multimed Tools Appl*, 39(3):441–471.
- [Torres e Falcão 2006] Torres, R. e Falcão, A. (2006). Contour Saliency Descriptors For Effective Image Retrieval And Analysis. *Image And Vision Computing*, 25(1):3–13.
- [Torres et al. 2004] Torres, R., Falcão, A., e da F. Costa, L. (2004). A Graph-based Approach for Multiscale Shape Analysis. *Pattern Recognition Letters*, 37(6):1163–1174.
- [Traina et al. 2003] Traina, A., Traina-Jr, C., Bueno, J., Chino, F., e Marques, P. (2003). Efficient Content-Based Image Retrieval Through Metric Histograms. *World Wide Web Journal*, 6(2):157–185.
- [Veltkamp e Hagedoorn 2001] Veltkamp, R. e Hagedoorn, M. (2001). *Advances in Pattern Recognition - Principles of Visual Information Retrieval*, chapter State of the Art in Shape Matching, pp. 87–117. Springer.
- [Veltkamp e Tanase 2002] Veltkamp, R. e Tanase, M. (2002). Content-Based Image Retrieval Systems: A Survey. Technical report, Utrecht University.
- [Vranić 2003] Vranić, D. (2003). An Improvement of Rotation Invariant 3D-Shape Descriptor Based on Functions on Concentric Spheres. *IEEE International Conference on Image Processing*, 3:757–760.
- [Vranić 2004] Vranić, D. (2004). *3D Model Retrieval*. PhD thesis, Universität Leipzig.
- [Vyas e Rege 2006] Vyas, V. e Rege, P. (2006). Automated Texture Analysis with Gabor filter. *GVIP Journal*, 6(1):35–41.

- [Weeks 2007] Weeks, M. (2007). *Digital Signal Processing Using MATLAB® and Wavelets*. Infinity Science Press LLC.
- [Wen e Acharya 1998] Wen, C. e Acharya, R. (1998). Self-Similar Texture Characterization Using a Fourier Domain Maximum Likelihood Estimation Method. *Pattern Recognition Letters*, 19(8):735–739.
- [Xie et al. 2010] Xie, Z., Liu, G., He, C., e Wen, Y. (2010). Texture Image Retrieval Based on Gray Level Co-Occurrence Matrix and Singular Value Decomposition. *Int. Conf. on Multimedia Technology*.
- [Zaharia e Prêteux 2001] Zaharia, T. e Prêteux, F. (2001). 3D Shape-Based Retrieval Within the MPEG-7 Framework. *SPIE Conf. on Nonlinear Image Processing and Pattern Analysis XIII*, 4304:133–145.
- [Zhang e Lu 2002] Zhang, D. e Lu, G. (2002). Shape-Based Image Retrieval Using Generic Fourier Descriptor. *Signal Processing: Image Communication*, 17:825–848.
- [Zhang e Lu 2003] Zhang, D. e Lu, G. (2003). A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures. *Visual Communication and Image Representation*, 14(1):41–60.
- [Zhang et al. 2010] Zhang, J., Palmer, S., e Wang, X. (2010). Identification of Animal Fibers with Wavelet Texture Analysis. *World Congress on Engineering*, pp. 742–747.
- [Zhang et al. 2005] Zhang, J., Siddiqi, K., Macrini, D., Shokoufandeh, A., e Dickinson, S. (2005). Retrieving Articulated 3-D Models Using Medial Surfaces and their Graph Spectra. *Machine Vision and Applications*, pp. 285–300.