Análise da influência de funções de distância para o processamento de consultas por similaridade em recuperação de imagens por conteúdo

Pedro Henrique Bugatti

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito: 25 de Março de 2008

Assinatura : \_

### Análise da influência de funções de distância para o processamento de consultas por similaridade em recuperação de imagens por conteúdo<sup>1</sup>

Pedro Henrique Bugatti

Orientador: Profa. Dra. Agma Juci Machado Traina

> Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para a obtenção do título de Mestre em Ciências de Computação e Matemática Computacional.

USP - São Carlos Abril/2008

<sup>&</sup>lt;sup>1</sup>Este trabalho tem o apoio financeiro da **FAPESP** (Fundação de Amparo à Pesquisa do Estado de São Paulo), processo 06/54462-1.

## Agradecimentos

Ao meu pai Ildeberto de Genova Bugatti, exemplo de conduta, caráter, dedicação e dignidade.

À minha adorada mãe Ivani Aparecida da Silva Bugatti por seu amor incondicional, carinho, apoio contínuo e constante, por me ensinar a persistir e ser determinado perante todas as dificuldades.

À minha namorada Priscila Tiemi Maeda Saito por estar ao meu lado em todos os momentos, por dividirmos juntos cada conquista, cada realização, felicidades e infelicidades, problemas e soluções, transpormos todos os empecilhos, por ser a razão de tudo, por todo seu carinho e Amor.

Ao meu irmão e melhor amigo, Fábio Henrique Bugatti pelos agradáveis e divertidos momentos de devaneio, por sua amizade e companheirismo.

À minha orientadora Profa. Dra. Agma Juci Machado Traina por ter acreditado em mim e dado a grande oportunidade e honra de trabalhar ao seu lado, por transmitir tanto conhecimento, sabedoria, delicadeza e incentivo em todos os momentos.

Ao professor Dr. Caetano Traina Júnior pelo apoio, valiosas sugestões e revisões.

A todos os amigos do Grupo de Bases de Dados e Imagens (GBdI).

À minha tia Dirce e primas Télia e Fernanda por terem estendido a mão quando mais precisei.

À FAPESP pelo apoio financeiro.

E por fim a todo e qualquer leitor do presente trabalho pelo interesse manifestado.

ii

### Resumo

A recuperação de imagens baseada em conteúdo (*Content-based Image Retrieval - CBIR*) embasa-se sobre dois aspectos primordiais, um extrator de características o qual deve prover as características intrínsecas mais significativas dos dados e uma função de distância a qual quantifica a similaridade entre tais dados. O grande desafio é justamente como alcançar a melhor integração entre estes dois aspectos chaves com intuito de obter maior precisão nas consultas por similaridade. Apesar de inúmeros esforços serem continuamente despendidos para o desenvolvimento de novas técnicas de extração de características, muito pouca atenção tem sido direcionada à importância de uma adequada associação entre a função de distância e os extratores de características.

A presente Dissertação de Mestrado foi concebida com o intuito de preencher esta lacuna. Para tal, foi realizada a análise do comportamento de diferentes funções de distância com relação a tipos distintos de vetores de características. Os três principais tipos de características intrínsecas às imagens foram analisados, com respeito a distribuição de cores, textura e forma. Além disso, foram propostas duas novas técnicas para realização de seleção de características com o desígnio de obter melhorias em relação à precisão das consultas por similaridade. A primeira técnica emprega regras de associação estatísticas e alcançou um ganho de até 38% na precisão, enquanto que a segunda técnica utilizando a entropia de Shannon alcançou um ganho de aproximadamente 71% ao mesmo tempo em que reduz significantemente a dimensionalidade dos vetores de características.

O presente trabalho também demonstra que uma adequada utilização das funções de distância melhora efetivamente os resultados das consultas por similaridade. Conseqüentemente, desdobra novos caminhos para realçar a concepção de sistemas CBIR.

## Abstract

The retrieval of images by visual content relies on a feature extractor to provide the most meaningful intrinsic characteristics (features) from the data, and a distance function to quantify the similarity between them. A challenge in this field supporting content-based image retrieval (CBIR) to answer similarity queries is how to best integrate these two key aspects. There are plenty of researching on algorithms for feature extraction of images. However, little attention have been paid to the importance of the use of a well-suited distance function associated to a feature extractor.

This Master Dissertation was conceived to fill in this gap. Therefore, herein it was investigated the behavior of different distance functions regarding distinct feature vector types. The three main types of image features were evaluated, regarding color distribution, texture and shape. It was also proposed two new techniques to perform feature selection over the feature vectors, in order to improve the precision when answering similarity queries. The first technique employed statistical association rules and achieve up to 38% gain in precision, while the second one employing the Shannon entropy achieved 71%, while significantly reducing the size of the feature vector.

This work also showed that the proper use of a distance function effectively improves the similarity query results. Therefore, it opens new ways to enhance the acceptance of CBIR systems.

vi

## Sumário

Li	sta c	le Figuras	xi
Li	sta d	le Tabelas	xv
Li	sta d	le Abreviaturas e Siglas	cvii
1	Intr	rodução	1
	1.1	Considerações Iniciais	1
	1.2	Motivação	2
	1.3	Objetivos	3
	1.4	Organização do documento	4
<b>2</b>	Rec	superação de Imagens Baseada em Conteúdo	7
	2.1	Introdução	7
	2.2	Consultas por Similaridade e Espaços Métricos	8
		2.2.1 Consulta por Abrangência	10
		2.2.2 Consulta aos Vizinhos Mais Próximos	10
		2.2.3 Espaços Métricos	10
	2.3	Extração de Características	12
	2.4	Medidas de Similaridade	13
	2.5	Métodos de Acesso	13
	2.6	Avaliação de Eficácia	15
	2.7	Considerações Finais	17
3	Bin	ômio: Relacionamento entre Função de Distância e Características	
	Intr	ínsecas dos Dados	19
	3.1	Introdução	19
	3.2	Extratores de Características	21
		3.2.1 Histograma Tradicional de Níveis de Cinza	21

3.2.3       Matrizes de Co-ocorrência         3.2.4       Momentos de Zernike         3.2.5       Variação do Método EM/MPM         3.3       Funções de Distância         3.3.1       Família de Distâncias Minkowski         3.3.2       Funções de Distâncias Consolidadas         3.3.3       Propostas de Funções de Distâncias         3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordageus para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StarMiner - Mineração de Regras de Associação Estatísticas - 5.5.1         5.5       Experimentos         5.5.1       Resultados         6       Um Novo Método para Redução da Dimensionalidade de Histogramu         6.1       Introdução <th></th> <th></th> <th>3.2.2 Histogramas Métricos</th> <th>22</th>			3.2.2 Histogramas Métricos	22
3.2.4       Momentos de Zernike         3.2.5       Variação do Método EM/MPM         3.3       Funções de Distâncias         3.3.1       Família de Distâncias Consolidadas         3.3.2       Funções de Distâncias Consolidadas         3.3.3       Propostas de Funções de Distâncias         3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas -         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histogramz         6.3.1       Conjuntos de Imag			3.2.3 Matrizes de Co-ocorrência	24
3.2.5       Variação do Método EM/MPM         3.3       Funções de Distância         3.3.1       Família de Distâncias Consolidadas         3.3.2       Funções de Distâncias Consolidadas         3.3.3       Propostas de Funções de Distâncias         3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas -         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Resultados      <			3.2.4 Momentos de Zernike	25
3.3       Funções de Distância         3.3.1       Família de Distâncias Consolidadas         3.3.2       Funções de Distâncias Consolidadas         3.3.3       Propostas de Funções de Distâncias         3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método         6.3       Experimentos         6.3.1 </th <th></th> <th></th> <th>3.2.5 Variação do Método EM/MPM</th> <th>28</th>			3.2.5 Variação do Método EM/MPM	28
<ul> <li>3.3.1 Família de Distâncias Minkowski</li> <li>3.3.2 Funções de Distâncias Consolidadas</li> <li>3.3.3 Propostas de Funções de Distâncias</li> <li>3.4 Considerações Finais</li> <li>4 Buscando a Melhor Configuração do Binômio: Função de distância Características</li> <li>4.1 Introdução</li> <li>4.2 Abordagens para a Extração de Características</li> <li>4.3 Conjuntos de Imagens</li> <li>4.4 Experimentos</li> <li>4.4.1 Resultados</li> <li>4.5 Considerações Finais</li> </ul> 5 Melhorando a Recuperação de Imagens Médicas por meio da Ponder ção de Características 5.1 Introdução <ul> <li>5.2 Conceitos Relacionados</li> <li>5.3 Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas</li> <li>5.4 Método Proposto</li> <li>5.5 Experimentos</li> <li>5.5.1 Resultados</li> <li>5.6 Considerações Finais</li> </ul> 6 Um Novo Método para Redução da Dimensionalidade de Histograma <ul> <li>6.1 Introdução</li> <li>6.2 Método Proposto</li> <li>6.3 Experimentos</li> <li>6.3.1 Conjuntos de Imagens</li> <li>6.4 Considerações Finais</li> </ul>		3.3	Funções de Distância	29
3.3.2       Funções de Distâncias Consolidadas         3.3       Propostas de Funções de Distâncias         3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMine - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.4       Considerações Finais         7.2			3.3.1 Família de Distâncias Minkowski	29
3.3.3       Propostas de Funções de Distâncias         3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5.1       Resultados         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.4       Considerações Finais         7.2       Principais Contribuições         7.3       Propostas para Trabalhos Futuros			3.3.2 Funções de Distâncias Consolidadas	31
3.4       Considerações Finais         4       Buscando a Melhor Configuração do Binômio: Função de distância Características         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder ção de Características         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.3.2       Resultados <th></th> <th></th> <th>3.3.3 Propostas de Funções de Distâncias</th> <th>33</th>			3.3.3 Propostas de Funções de Distâncias	33
4       Buscando a Melhor Configuração do Binômio: Função de distância         Características       4.1         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.3       Resultados         6.4       Considerações Finais         7.2       Principais Contribuições         7.3       Propostas para Trabalhos Futuros		3.4	Considerações Finais	34
Características         4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4       Experimentos         4.4       Experimentos         4.4       Experimentos         4.4       Experimentos         4.4       Experimentos         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.3.2       Resultados         6.4       Considerações Finais         7.2	4	Bus	cando a Melhor Configuração do Binômio: Função de distância e	
4.1       Introdução         4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.3.4       Considerações Finais         6.4       Considerações Finais         7.2       Principais Contribuições         7.3       Propostas para Trabalhos Futuros		Car	acterísticas	37
4.2       Abordagens para a Extração de Características         4.3       Conjuntos de Imagens         4.4       Experimentos         4.4       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.3.2       Resultados         6.3.1       Considerações Finais         7.2       Ornicusões         7.1       Considerações Finais         7.2       Principais Contribuições         7.3       Propostas para Trabalhos Futuros		4.1	Introdução	37
<ul> <li>4.3 Conjuntos de Imagens</li> <li>4.4 Experimentos</li> <li>4.4.1 Resultados</li> <li>4.5 Considerações Finais</li> <li>5 Melhorando a Recuperação de Imagens Médicas por meio da Ponder ção de Características</li> <li>5.1 Introdução</li> <li>5.2 Conceitos Relacionados</li> <li>5.3 Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas</li> <li>5.4 Método Proposto</li> <li>5.5 Experimentos</li> <li>5.6 Considerações Finais</li> <li>6 Um Novo Método para Redução da Dimensionalidade de Histograma</li> <li>6.1 Introdução</li> <li>6.2 Método Proposto</li> <li>6.3 Experimentos</li> <li>6.3.1 Conjuntos de Imagens</li> <li>6.3.2 Resultados</li> <li>6.4 Considerações Finais</li> <li>7 Conclusões</li> <li>7.1 Considerações Finais</li> <li>7.2 Principais Contribuições</li> <li>7.3 Propostas para Trabalhos Futuros</li> </ul>		4.2	Abordagens para a Extração de Características	37
<ul> <li>4.4 Experimentos</li></ul>		4.3	Conjuntos de Imagens	39
4.4.1       Resultados         4.5       Considerações Finais         5       Melhorando a Recuperação de Imagens Médicas por meio da Ponder         ção de Características       5.1         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.4       Considerações Finais         7       Conclusões         7.1       Considerações Finais         7.2       Principais Contribuições         7.3       Propostas para Trabalhos Futuros		4.4	Experimentos	41
<ul> <li>4.5 Considerações Finais</li> <li>5 Melhorando a Recuperação de Imagens Médicas por meio da Ponder ção de Características</li> <li>5.1 Introdução</li> <li>5.2 Conceitos Relacionados</li> <li>5.3 Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas</li> <li>5.4 Método Proposto</li> <li>5.5 Experimentos</li> <li>5.5.1 Resultados</li> <li>5.6 Considerações Finais</li> <li>6 Um Novo Método para Redução da Dimensionalidade de Histograma</li> <li>6.1 Introdução</li> <li>6.2 Método Proposto</li> <li>6.3 Experimentos</li> <li>6.3.1 Conjuntos de Imagens</li> <li>6.3.2 Resultados</li> <li>6.4 Considerações Finais</li> <li>7 Conclusões</li> <li>7.1 Considerações Finais</li> <li>7.2 Principais Contribuições</li> <li>7.3 Propostas para Trabalhos Futuros</li> </ul>			4.4.1 Resultados	42
<ul> <li>5 Melhorando a Recuperação de Imagens Médicas por meio da Ponder ção de Características</li> <li>5.1 Introdução</li></ul>		4.5	Considerações Finais	49
ção de Características         5.1       Introdução         5.2       Conceitos Relacionados         5.3       Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas         5.4       Método Proposto         5.5       Experimentos         5.6       Considerações Finais         5.6       Considerações Finais         6       Um Novo Método para Redução da Dimensionalidade de Histograma         6.1       Introdução         6.2       Método Proposto         6.3       Experimentos         6.3.1       Conjuntos de Imagens         6.3.2       Resultados         6.4       Considerações Finais         7       Conclusões         7.1       Considerações Finais         7.2       Principais Contribuições         7.3       Propostas para Trabalhos Futuros	<b>5</b>	Mel	lhorando a Recuperação de Imagens Médicas por meio da Pondera-	
<ul> <li>5.1 Introdução</li></ul>		ção	de Características	51
<ul> <li>5.2 Conceitos Relacionados</li></ul>		5.1	Introdução	51
<ul> <li>5.3 Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas .</li> <li>5.4 Método Proposto</li></ul>		5.2	Conceitos Relacionados	52
<ul> <li>5.4 Método Proposto</li></ul>		5.3	Algoritmo Sta RMiner - Mineração de Regras de Associação Estatísticas 	53
<ul> <li>5.5 Experimentos</li></ul>		5.4	Método Proposto	56
<ul> <li>5.5.1 Resultados</li></ul>		5.5	Experimentos	58
<ul> <li>5.6 Considerações Finais</li></ul>			5.5.1 Resultados	58
<ul> <li>6 Um Novo Método para Redução da Dimensionalidade de Histograma</li> <li>6.1 Introdução</li></ul>		5.6	Considerações Finais	67
<ul> <li>6.1 Introdução</li></ul>	6	$\mathbf{Um}$	Novo Método para Redução da Dimensionalidade de Histogramas	69
<ul> <li>6.2 Método Proposto</li></ul>		6.1	Introdução	69
<ul> <li>6.3 Experimentos</li></ul>		6.2	Método Proposto	70
6.3.1       Conjuntos de Imagens		6.3	Experimentos	72
6.3.2 Resultados       6.3.2 Resultados         6.4 Considerações Finais			6.3.1 Conjuntos de Imagens	73
<ul> <li>6.4 Considerações Finais</li></ul>			6.3.2 Resultados	73
<ul> <li>7 Conclusões</li> <li>7.1 Considerações Finais</li></ul>		6.4	Considerações Finais	80
<ul> <li>7.1 Considerações Finais</li></ul>	7	Con	nclusões	81
<ul> <li>7.2 Principais Contribuições</li></ul>		7.1	Considerações Finais	81
7.3 Propostas para Trabalhos Futuros		7.2	Principais Contribuições	82
		7.3	Propostas para Trabalhos Futuros	83

#### Referências Bibliográficas

## Lista de Figuras

2.1	Visão geral dos principais componentes que integram um sistema de recu-	
	peração de imagens baseada em conteúdo.	8
2.2	Consultas por similaridade em um domínio de pontos bidimensionais uti-	
	lizando a função de distância Euclidiana como função de dissimilaridade.	
	O objeto "central" preto $O_q$ é o objeto de consulta enquanto os respecti-	
	vos objetos cinza constituem as respostas. (a) Ilustra uma consulta por	
	abrangência de raio $r_q$ e (b) ilustra uma consulta pelos 5 vizinhos mais	
	próximos	11
2.3	Ilustração Desigualdade Triangular	11
2.4	<i>Slim-tree</i> armazenando 17 objetos e possuindo <i>fan-out</i> igual a 3	15
2.5	Conjunto de resultados possíveis para uma dada operação de busca	16
2.6	Exemplo de um gráfico de medidas de precisão e revocação	17
3.1	Histograma normalizado com pontos de controle de máximo e mínimo local	
	$\langle bx, hx \rangle$ os quais definem os <i>buckets</i> correspondentes ao seu <b>histograma</b>	
	<b>métrico</b> [Bueno, 2002]	23
3.2	Exemplos de matrizes de co-ocorrência. (a) matrix de pixels da imagem;	
	(b) matriz de co-ocorrência referente ao ângulo $0 e d = 1$ ; (c) matriz de	
	co-ocorrência referente ao ângulo $135^{\circ}$ e d = 1	25
3.3	Representação das coordenadas polares de um ponto $T$ sobre o plano eu-	
	clidiano $\Re^2$ cuja origem é denotada por $O$	26
3.4	Exemplo de segmentação realizada pela variação do método EM/MPM.	
	(a) imagem original; (b) imagem segmentada com $L = 5$ ; (c) região de	
	classe 1 (fluido cérebro-espinhal); (d) região de classe 2 (massa cinzenta);	
	(e) região de classe 3 (massa branca); (f) região de classe 4 (dura, medula	
	óssea, gordura) [Balan et al., 2005] $\ldots$	28
3.5	Vetor de características das imagens: $6(L - 1)$ valores reais	
	[Balan et al., 2005]	29

3.6	Representação das formas geométricas geradas conforme a distância $L_p$ utilizada e em relação aos pontos equidistantes à distância $r_q$ a partir do objeto centro $o_q$ .	31
4.1	Exemplos de imagens de cada classe pertencentes ao conjunto de teste da base de imagens de ressonância magnética.	40
4.2	Exemplos de imagens de cada classe pertencentes ao conjunto de teste da base de imagens mamográficas de regiões de interesse ( <i>Regions of Interest</i> - <i>ROIs</i> ) de nódulos tumorais: ( <i>a</i> ) nódulo benigno; ( <i>b</i> ) nódulo maligno	40
4.3	Exemplos de imagens de cada classe pertencentes ao conjunto de teste da base de imagens mamográficas classificadas segundo o padrão <i>BI-RADS</i> , contendo nódulos tumorais de mama típicos: (a) Classe 2; (b) Classe 3; (c) Classe 4; (d) Classe 5	41
4.4	Gráfico de curvas de precisão e revocação, ilustrando e comparando o com- portamento das buscas realizadas para as três bases utilizadas para expe- rimentos; comparando o extrator de características de <b>histogramas de</b> <b>níveis de cinza</b> com as funções de distância: $(a)$ base de imagens de res- sonância magnética; $(b)$ base de imagens de regiões de interesse $(ROIs)$ provenientes de exames de mamografia; $(c)$ base de imagens provenientes	
4.5	de exames de mamografia classificadas segundo o padrão $BI$ - $RADS$ Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas para as três bases utilizadas para experimentos; comparando o extrator de características de <b>textura</b> com as funções de distância: (a) base de imagens de ressonância magnética; (b) base de imagens de regiões de interesse ( $ROIs$ ) provenientes de exames de mamografia; (c) base de imagens provenientes de exames de mamografia	43
4.6	classificadas segundo o padrão <i>BI-RADS</i>	45
4.7	Gráfico de curvas de precisão e revocação, ilustrando e comparando o com- portamento das buscas realizadas para a base de imagens de ressonância magnética utilizada para experimento; comparando o extrator de caracte- rísticas baseado em segmentação dos objetos da imagem (EM/MPM) com	11
	as diversas funções de distância.	48
4.8	Ilustração das comparações entre as funções e os extratores de características.	48

5.1	Regiões de rejeição [Ribeiro et al., 2005]	55
5.2	Diagrama ilustrativo das fases do método proposto: fase de treinamento e	
	fase de teste	57
5.3	Gráficos de precisão e revocação utilizando as funções de distância (a) $L_1$ , (b) $L_2$ e (c) $L_{\infty}$ obtidos sobre as características de textura extraídas do conjunto de imagens de ressonância magnética, empregando: característi- cas não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando $q = 0$ (removendo as características irrelevantes) e o método proposto, utilizando $q = 1$	60
5.4	Gráficos de precisão e revocação utilizando as funções de distância (a) $L_1$ , (b) $L_2$ e (c) $L_{\infty}$ obtidos sobre as características de forma extraídas do con- junto de imagens de ressonância magnética, empregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando $q = 0$ (removendo as características irrelevantes) e o método proposto, utilizando $q = 1$	61
5.5	Exemplo da execução de uma consulta k-NN $(k=8)$ , utilizando a imagem situada no canto esquerdo superior como centro de consulta e função de distância $L_2$ ; (a) utilizando as características originais; (b) utilizando o método proposto com $q = 0$ . As imagens envolvidas pela moldura vermelha referem-se a falsos positivos.	63
5.6	Gráficos de precisão e revocação utilizando as funções de distância (a) $\chi^2$ , (b) Canberra e (c) Divergência de Jefrrey obtidos sobre as características de textura extraídas do conjunto de imagens de ressonância magnética, em- pregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando $q = 0$ (removendo as caracte- rísticas irrelevantes) e o método proposto, utilizando $q = 1, \dots, \dots, \dots$	64
5.7	Gráficos de precisão e revocação utilizando as funções de distância (a) $\chi^2$ , (b) Canberra e (c) Divergência de Jefrrey obtidos sobre as características de forma extraídas do conjunto de imagens de ressonância magnética, em- pregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando $q = 0$ (removendo as caracte- rísticas irrelevantes) e o método proposto, utilizando $q = 1$	66
6.1	Método proposto para redução da dimensionalidade de histogramas	72
6.2	Gráfico de curvas de precisão e revocação, ilustrando e comparando o com- portamento das buscas realizadas sobre a base de imagens de ressonância magnética, utilizando as características extraídas por meio do histograma convencional de níveis de cinza, histogramas métricos, e o novo método proposto para vários fatores de aglutinação $t \in \{4, 8, 16, 32, 64, 128\}$ , utili-	
	zando a função de distância Euclidiana	75

6.3 Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens de ressonância magnética, utilizando as características extraídas por meio dos momentos de Zernike e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana. . . .

76

77

78

- 6.4 Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas de regiões de interesse (*Regions of Interest - ROIs*) de nódulos tumorais, utilizando as características extraídas por meio do histograma convencional de níveis de cinza, histogramas métricos, e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.
- 6.5 Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas de regiões de interesse (*Regions of Interest - ROIs*) de nódulos tumorais, utilizando as características extraídas por meio dos momentos de Zernike e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana. . . .
- 6.6 Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas classificadas segundo o padrão *BI-RADS*, utilizando as características extraídas por meio do histograma convencional de níveis de cinza, histogramas métricos, e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana. . . . 79

## Lista de Tabelas

3.1	Descritores de Haralick amplamente utilizados	25
4.1	Distribuição das imagens no conjunto de teste	39
4.2	Distribuição das imagens no conjunto de teste	41
4.3	Tabela Comparativa das Funções de Distância	49
5.1	Valores Críticos de Z [Ribeiro et al., 2005]	55

## Lista de Abreviaturas e Siglas

AID	Attribute Interaction Distances.
ACR	American College of Radiology.
<b>BI-RADS</b>	Breast-Imaging Reporting and Data System.
$\operatorname{cbPACS}$	Content-Based Picture Archiving and Communication System.
CCIFM	Centro de Ciências de Imagens e Física Médica.
CBIR	Content-based Image Retrieval.
$\mathbf{CT}$	Tomografia Computadorizada.
DBM-Tree	Density-Based Metric Tree.
DF-Tree	Distance Fields Tree.
DKL	Divergência de Kullback-Leibler.
DM	Distância Métrica.
DNA	Ácido Desoxirribonucléico.
$\mathbf{E}\mathbf{M}$	Expectation and Maximization.
$\operatorname{GBdI}$	Grupo de Bases de Dados e Imagens - ICMC - USP.
GWD	Global Warp Distance.
HCRP-USP	Hospital das Clínicas de Ribeirão Preto - USP.
ICMC-USP	Instituto de Ciências Matemáticas e de Computação - USP.
k - NN	Consulta aos k-Vizinhos mais próximos (k-Nearest Neighbors Query).
$\mathbf{M}\mathbf{A}$	Método de Acesso.
MAM	Metric Access Method (Método de Acesso Métrico).
PACS	Picture Archiving and Communication System.
PCA	Principal Component Analysis.
RM	Ressonância Magnética.
ROIs	Regions of Interest.
SAID	Strong Attribute Interaction Distance.
$\mathbf{SAM}$	Spatial Access Method (Método de Acesso Espacial).

Sistemas Gerenciadores de Bases de Dados.
Sistemas de Informações Hospitalares.
Structure Query Language.
Statistical Association Rule Miner.
Universidade de São Paulo.
Weak Attribute Interaction Distance.

## Capítulo 1 Introdução

#### 1.1 Considerações Iniciais

s Sistemas de Informações Hospitalares (SIH) processam informações relativas aos pacientes, destacando-se, entre elas, seu estado de saúde, os exames realizados e os procedimentos médicos adotados. Atualmente, a maioria desses sistemas armazenam dados textuais e numéricos, mas, em alguns casos, já é possível a associação das informações textuais com informações gráficas, provenientes de exames como raio-X, tomografia (CT ou RM), ultrasom, entre outros. Recentemente, com a introdução dos sistemas PACS (*Picture Archiving and Communication System*) [Furuie et al., 1999] [Müller et al., 2004], cresceu o interesse por integrar num só sistema todas as informações dos pacientes (textos, imagens, gráficos, dados temporais, entre outros).

Infelizmente, os poucos sistemas PACS comerciais oferecidos atualmente são extremamente caros, e não contemplam todas as necessidades do centro médico [Sung et al., 2000]. Dessa forma, o estado da arte na área resume-se ao desenvolvimento de soluções locais, muitas vezes restritas. Alguns centros médicos integrados a unidades de pesquisa optam, via de regra, por adquirir apenas alguns módulos do sistema, desenvolvendo e adaptando outros módulos e criando soluções próprias conforme sua necessidade. Assim, o custo é reduzido e contemplam-se necessidades e características próprias de cada ambiente.

Como os sistemas PACS organizam as imagens provenientes de exames efetuados sobre os pacientes, seria muito interessante que por meio desses sistemas fosse possível realizar consultas e recuperar tais imagens baseando-se apenas no conteúdo das mesmas. Por exemplo, pode ser necessário *"recuperar todas as imagens obtidas em exames de tomo*grafia computadorizada de coração, similares a uma imagem dada". Nesse caso, o cruzamento das informações é independente das chaves de busca tradicionais, baseadas em dados numéricos e textuais, as quais associam exames e pacientes em um SIH tradicional. A facilidade de recuperação de dados baseada em seu conteúdo é um dos recursos mais almejados para incorporação aos sistemas PACS [Müller et al., 2004] [Korn et al., 1996] [Lima et al., 1998]. Através dessa facilidade, podem ser encontrados registros de pacientes cujas imagens associadas apresentam aspectos de similaridade (utilizando padrões gráficos), além de se obter correlações utilizando sintomas previamente registrados, tratamentos efetuados, etc [Marsh, 1997]. Pode-se notar que um sistema PACS é potencialmente uma ferramenta bastante eficaz no auxílio ao diagnóstico médico.

Um sistema PACS deve então ser capaz de responder a consultas por similaridade, isto é, questões do tipo: "quais são as 10 imagens mais semelhantes (próximas) de uma dada imagem" ou "quais são as imagens que diferem de até 5 unidades de uma imagem padrão". Vale notar que tanto as imagens quanto a função de distância (dissimilaridade) definida são usualmente dependentes do domínio dos dados e do interesse do que se considera "semelhante".

A similaridade entre imagens pode ser medida de várias formas. Características como formato, cor e textura podem ser extraídas de imagens em uma base de dados, e são informações que podem ser utilizadas em cálculos de distância [Aslandogan & Yu, 1999].

#### 1.2 Motivação

Ao efetuar uma consulta por similaridade em um banco de imagens, o sistema retorna o conjunto de imagens que respondem à solicitação efetuada. A consulta é realizada através da estrutura de índices que suporta o armazenamento das imagens. Dessa forma, o conjunto de características extraídas das imagens e a função de distância utilizada têm papel preponderante nessa operação de consulta.

Uma pergunta que surge quando se fala em consultas por similaridade em imagens é, na avaliação dos resultados, se eles correspondem aos esperados pelo usuário. Isto é, como saber, por exemplo, ao solicitar as 5 imagens mais semelhantes a uma dada imagem, se a proximidade ou similaridade retornada pelo sistema realmente mapeia o que é desejado pelo usuário? Ou, mais precisamente, como implementar e avaliar a semântica esperada pelo usuário?

Atualmente são utilizados, de forma generalizada, os modelos  $L_p$  (descritos no capítulo 2) como funções de distância entre vetores de características extraídas das imagens. Porém, tais funções não conseguem representar diretamente o sentimento que o analista humano tem ao comparar duas imagens. Essas funções são gerais e não captam aspectos particulares das imagens em questão. Assim, os resultados de comparações entre imagens deverão ser mais apropriados e efetivos, se for feito uso de funções ajustadas às características mais específicas do domínio de aplicação, no caso, as imagens médicas.

Muitas vezes os próprios médicos divergem quanto à interpretação semântica de imagens médicas. Esse é um dos problemas mais conhecidos e discutidos sobre o diagnóstico por imagens. Imagens patologicamente semelhantes podem ter sido classificadas de formas divergentes. O intuito desta pesquisa é o de auxiliar no diagnóstico, fornecendo ao médico um conjunto de casos semelhantes ao de um paciente, com os diagnósticos já realizados, já que eles também estão associados à imagem no banco de dados. Desta forma o médico poderá analisar comparativamente o tipo de tratamento dado ao paciente e sua resposta ao mesmo. Esse tipo de abordagem não poderá substituir o médico na análise da imagem, mas poderá apontar alternativas para o tratamento do paciente. Atualmente é impossível realizar esse tipo de pesquisa sem um conhecimento prévio sobre casos clínicos semelhantes (ou seja, sem que o médico saiba muito bem onde procurar referências) e, considerando o grande volume de exames realizados em hospitais, essa é uma tarefa bastante árdua.

Imagens médicas são um tipo de dado complexo que traz por si só uma série de desafios. Devem ser tratadas por técnicas que não permitam a perda de informação, o que muitas vezes é relevado em algoritmos para extração de características de imagens tradicionais. Os sistemas de recuperação de imagens por conteúdo (*Content-based Image Retrieval* - CBIR) podem ser uma ferramenta muito valiosa para auxiliar ao processamento de Consultas por Similaridade, e auxiliar o diagnóstico médico, pois retornará as imagens ou casos mais semelhantes àquele que o especialista médico está analisando.

Por mais de uma década, pesquisadores têm explorado a área de análise e recuperação de imagens por conteúdo. Entretanto, a maioria dos trabalhos encontrados preocupam-se com a proposta de novos extratores de características e negligenciam o inter-relacionamento entre as características e as funções de distância que os comparam, indicando o grau de similaridade entre as imagens. Desse modo, fazem uso das mais conhecidas e amplamente utilizadas funções de distância, tais como a distância Euclidiana e outros membros da família  $L_p$ , sem explorar adequadamente se tais funções suportam a recuperação das imagens por conteúdo atendendo às expectativas dos usuários.

#### 1.3 Objetivos

Existe uma grande diversidade de funções de distância para a realização de consultas por similaridade em imagens médicas.

O desenvolvimento de funções de distância depende de quais serão as características da imagem que serão utilizadas, o que leva à elaboração e aperfeiçoamento de desenvolvimento paralelo entre os dois aspectos. Os Métodos de Acesso Métricos (MAMs) atuais utilizam basicamente as distâncias  $L_p$  e  $L_{Edit}$ , sendo que, até o momento, foram efetuados testes sobre imagens médicas utilizando-se apenas distância Euclidiana ( $L_2$ ) [Bozkaya & Özsoyoglu, 1999]. A proposta contempla o estudo e análise de novas funções distância para cálculo de similaridade em imagens médicas

O objetivo do presente projeto é o estudo e a análise do inter-relacionamento entre as funções de distância e os extratores de características de imagens médicas, com o objetivo de obter a melhor integração entre esses dois aspectos chaves para um sistema de recuperação de imagens baseada em conteúdo, visando melhorias e conseqüentemente obtendo resultados mais precisos em relação às consultas por similaridade.

A seguir a estrutura *Slim-tree* foi estendida e adequada para o tratamento de imagens, incluindo as funções de distância estudadas. Foi tomado especial cuidado com o custo computacional das funções implementadas, devido ao fato de o número de cálculos de distâncias efetuadas para as comparações serem usualmente bastante altos. A estrutura *Slim-tree* foi utilizada com o intuito de acelerar o processamento das consultas por similaridade, pois permiti a poda pela propriedade de desigualdade triangular (ver seção 2.2.3). A prova de tal poda pelo uso da propriedade de desigualdade triangular pode ser obtida em [Burkhard & Keller, 1973].

#### 1.4 Organização do documento

Este trabalho de pesquisa apresenta a seguinte organização:

- No presente capítulo foram apresentadas considerações iniciais, a motivação para o desenvolvimento do trabalho, assim como os objetivos pertinentes para seu desenvolvimento e a organização do documento;
- No capítulo 2 são introduzidos os principais conceitos envolvidos no processo de recuperação de imagens baseada em conteúdo (CBIR), envolvendo fundamentalmente a extração de características, métricas de comparação, tipos de consulta, estruturas de indexação adequadas ao contexto abordado, assim como a avaliação de desempenho.
- No capítulo 3 são apresentados o contexto da melhor integração entre função de distância e extratores de características. Uma revisão sobre as principais funções de distância utilizadas no trabalho é realizada, bem como dos extratores de características e suas aplicações.
- No **capítulo 4** são explicitados os experimentos e resultados obtidos com o intuito de definir o melhor par função de distância e extrator de característica com relação ao ganho de precisão em consultas por similaridade em imagens médicas.
- No **capítulo 5** é apresentada a proposta e desenvolvimento do método de ponderação de características baseado na junção entre a mineração das mesmas e a técnica de *continuous feature selection*, assim como os experimentos e resultados obtidos.
- No **capítulo 6** é apresentada a proposta e desenvolvimento de um novo método para redução da dimensionalidade de histogramas convencionais de níveis de cinza, sendo seu desempenho testado e avaliado por meio dos esperimentos realizados.

• No **capítulo 7** são apresentadas as considerações finais, bem como uma síntese das principais contribuições do presente trabalho e propostas para trabalhos futuros.

# Capítulo 2

## Recuperação de Imagens Baseada em Conteúdo

#### 2.1 Introdução

*P* ara realizar o processo de recuperação de imagens, duas abordagens podem ser adotadas. A primeira abordagem utiliza informações textuais e numéricas para catalogar as imagens de uma base. Neste caso, a recuperação é realizada através de comandos interpretados pelo sistema de gerenciamento do banco de dados (por exemplo, comandos em linguagem SQL - *Structured Query Language*). Se por um lado esta abordagem apresenta certa praticidade e simplicidade, por outro lado ela traz determinadas limitações. Uma dessas limitações consiste no fato de que é praticamente inviável fazer anotações manuais completas em todas as imagens do banco, considerandose uma aplicação de grande porte, como é o caso de bases de imagens médicas obtidas por hospitais ou centros médicos. Além disso, a essência semântica do conteúdo visual das imagens, principalmente aquelas que possuem mais detalhes, dificultam a descrição textual das mesmas [Zhang & Lu, 2002]. Outro fator desvantajoso a ser ressaltado diz respeito a diferentes indivíduos fornecerem diferentes descrições de uma mesma imagem, ou seja, as informações inseridas são subjetivas, gerando dessa forma inconsistências e resultados insatisfatórios para uma operação de consulta.

Objetivando transpor as dificuldades e desvantagens apresentadas pela descrição textual das imagens, surgiu a segunda abordagem de recuperação de imagens utilizando seu conteúdo pictórico. O principal desígnio foi o de estender as técnicas tradicionais de recuperação de imagens em bancos de dados multimídia, incluindo também informações visuais inerentes às imagens, possibilitando que essas pudessem ser obtidas por meio de um processo automático. Isto é, as imagens a partir desse momento poderiam ser recuperadas por meio de seu conteúdo visual e não mais textual/descritivo. A essa abordagem, foi dado o nome de "Recuperação de Imagens Baseada em Conteúdo" (*Content-Based Image Retrieval - CBIR*). O objetivo principal da abordagem de recuperação de imagens baseada em conteúdo consiste em extrair informações das imagens que possam caracterizar adequadamente (idealmente univocamente) o conteúdo das mesmas.

Grande parte dos sistemas de recuperação de imagens baseada em conteúdo tem arquitetura similar, e envolve algumas etapas primordiais. Entre as principais etapas de um sistema CBIR estão o armazenamento, indexação e extração de características, e o cálculo de similaridade realizado por meio de uma métrica de similaridade (ou função de distância). Tais etapas são integradas por um mecanismo de recuperação de imagens que possibilita dessa forma, a realização de consultas baseadas em conteúdo [Müller et al., 2004]. A Figura 2.1 ilustra uma visão geral dos principais componentes de um sistema CBIR.



Figura 2.1: Visão geral dos principais componentes que integram um sistema de recuperação de imagens baseada em conteúdo.

Inicialmente as imagens são processadas e as características visuais das mesmas são extraídas. Para a construção da estrutura de índices, a similaridade entre as imagens da base é avaliada. As características da imagem de consulta são extraídas e comparadas com as características das demais imagens da base para a execução de uma consulta.

No presente capítulo são introduzidos e discutidos os conceitos envolvidos em cada etapa do processo de recuperação de imagens baseada em conteúdo, iniciando pela seção subseqüente a qual introduz o conceito de consultas por similaridade.

#### 2.2 Consultas por Similaridade e Espaços Métricos

Os Sistemas Gerenciadores de Bases de Dados (SGBDs) tradicionais manipulam domínios de dados que possuem relação de ordem entre os seus elementos (números, textos pequenos, entre outros). Portanto, permitem identificar e organizar elementos por meio de suas grandezas em relação a uma unidade de ordem ou igualdade. Porém, quando se utilizam SGBDs para armazenar dados complexos, os quais na sua grande maioria não possuem relação de ordem, as buscas por igualdade (exemplo: "Procure a cidade cujo nome é Osaka") ou baseadas em ordem (exemplo: "Procure as pessoas que recebem auxílio institucional entre o dia 5 e o dia 20") não se aplicam ou simplesmente têm pouca utilidade. Portanto surge a questão de como ordenar tais dados complexos (imagens, áudio, vídeo, séries temporais e outros) pois não se consegue estabelecer uma relação de ordem para esses dados (exemplo: "Dadas duas imagens, qual é a "menor" entre elas?"), no entanto pode-se dizer o quanto uma imagem é similar (parecida) com outra imagem.

A busca no domínio de dados complexos, a qual é a base para o processo de recuperação de imagens baseada em conteúdo, é fundamentada no conceito de busca por similaridade (*similarity search*) e surge naturalmente como a maneira mais adequada para realizar consultas sobre tais dados. Esse gênero de consulta é denominado **consulta por similaridade**.

As consultas por similaridade consistem em procurar por objetos em um conjunto que, segundo algum critério de similaridade, sejam mais "parecidos" ou mais "distintos" com/de um determinado objeto. Em outras palavras, consultas por similaridade consistem na comparação de todos os objetos do conjunto com um objeto escolhido, selecionando apenas os elementos que atendam a um certo critério de similaridade. Por exemplo, visando a obtenção de um diagnóstico mais fidedigno sobre determinado tumor cerebral de um paciente, e até mesmo tentando encontrar subsídios para decidir como encaminhar o tratamento, um médico estaria interessado em encontrar as tomografias de cérebro que apresentem um tumor similar àquele em estudo. Esse tipo de consulta, que leva em consideração a informação inerente ao objeto de referência, é chamado de **Consulta Baseada em Conteúdo** (*Content-Based Query*), ou seja, a consulta é feita a partir dos dados contidos no objeto, e não a partir de informações descritivas provavelmente associadas aos mesmos (metadados).

Para se determinar a similaridade entre os objetos (no caso em questão imagens médicas), estes podem ser comparados diretamente ou a partir de vetores de características extraídos desses objetos. Nos dois casos, uma função de distância d() é definida para calcular o grau de similaridade, retornando um valor numérico que quantifica o quão similar estes objetos são entre si.

Existem vários tipos de consultas por similaridade [Traina & Traina Jr., 2003], porém as mais comuns são as buscas por abrangência (limitam os resultados a um certo nível de similaridade) e as buscas pelos vizinhos mais próximos (limitam as buscas a um certo número de objetos mais similares). Esses dois tipos de consultas são detalhados nas seções a seguir, bem como o conceito de espaços métricos, primordial para o completo entendimento de tais consultas e outros conceitos que serão introduzidos posteriormente.

#### 2.2.1 Consulta por Abrangência

As Consultas por Abrangência (*Range query*) consistem na busca de objetos que tenham até um determinado nível de similaridade com o objeto de referência. Isto é, dado um conjunto de objetos  $O = \{o_1, o_2, o_3, ..., o_n\}$  pertencentes a um domínio  $\mathfrak{D}$ , uma função de distância d(), um objeto de consulta  $o_q$  também pertencente ao domínio e uma distância máxima  $r_q$ , o resultado da busca por abrangência é dado por:

$$range(o_q, r_q) = A = \{a | a \in O, d(a, o_q) \le r_q\}$$
(2.1)

Indagações tais como "Quais são as nebulosas que diferem até X da Nebulosa de Órion?" e "Quais são as galáxias que ficam até Y anos luz da Via Láctea?" são exemplos desse tipo de consulta.

Na Figura 2.2 (a) é ilustrada uma resposta de uma consulta por abrangência em um domínio bidimensional utilizando a função de distância Euclidiana  $(L_2)$  - ver seção 3.3.1 do capítulo 3.

#### 2.2.2 Consulta aos Vizinhos Mais Próximos

A Consulta aos k Vizinhos mais Próximos (k-Nearest-Neighbor query) consiste na busca dos k objetos mais similares ao objeto de consulta. Ou seja, dado um conjunto de objetos  $O = \{o_1, o_2, o_3, ..., o_n\}$  pertencentes a um domínio  $\mathfrak{D}$ , uma função de distância d(), um objeto de consulta  $o_q$  também pertencente ao domínio e um número inteiro k, o resultado da busca pelos vizinhos mais próximos é dado por:

$$k - NN(o_q, k) = A = \{a | a \in O, \forall o \in O - A, d(o_q, a) \le d(o_q, o), |A| = k\}$$
(2.2)

Indagações tais como "Quais são as N nebulosas mais parecidas com a Nebulosa de Órion?" e "Quais são as M galáxias mais próximas da Via Láctea?" são exemplos desse tipo de consulta. No primeiro, k é igual a N e no segundo k é M.

Na Figura 2.2 (b) é ilustrada uma resposta de uma consulta pelos vizinhos mais próximos (no caso, k = 5) em um domínio bidimensional empregando a função de distância Euclidiana ( $L_2$ ).

#### 2.2.3 Espaços Métricos

Um espaço métrico M é um par  $\langle \mathfrak{D}, d() \rangle$  onde  $\mathfrak{D}$  é o domínio dos objetos e d():  $\mathfrak{D} \times \mathfrak{D} \to \mathfrak{R}^+$  é uma função que associa uma distância a cada par ordenado de elementos  $o_1, o_2, o_3 \in \mathfrak{D}$ . Então, o par  $\langle \mathfrak{D}, d() \rangle$  é denominado espaço métrico sempre que a função d() atender aos seguintes axiomas:

1. Simetria:  $d(o_1, o_2) = d(o_2, o_1)$ 



Figura 2.2: Consultas por similaridade em um domínio de pontos bidimensionais utilizando a função de distância Euclidiana como função de dissimilaridade. O objeto "central" preto  $O_q$  é o objeto de consulta enquanto os respectivos objetos cinza constituem as respostas. (a) Ilustra uma consulta por abrangência de raio  $r_q$  e (b) ilustra uma consulta pelos 5 vizinhos mais próximos.

2. Não negatividade:  $0 < d(o_1, o_2) < \infty$  se  $o_1 \neq o_2$  e  $d(o_1, o_1) = 0$ 

#### 3. Desigualdade Triangular: $d(o_1, o_2) \leq d(o_1, o_3) + d(o_3, o_2)$

A condição 3 (Desigualdade Triangular) tem origem no fato de que, no plano cartesiano, o comprimento de um dos lados de um triângulo não excede a soma dos outros dois (ver Figura 2.3).



Figura 2.3: Ilustração Desigualdade Triangular

As condições que definem um espaço métrico, principalmente a desigualdade triangular, possibilitam o desenvolvimento de técnicas de indexação eficientes para responder a consultas por similaridade.

Os elementos de um espaço métrico podem ser de natureza bastante distintas, tais como números, pontos, vetores, matrizes, funções, conjuntos entre outros.

A função de distância métrica d() pode ser interpretada como uma função de cálculo de dissimilaridade entre os objetos do domínio ao qual está inserida, onde, quanto mais similares dois objetos são entre si, menor o valor retornado. Conseqüentemente valores maiores são retornados para objetos muito distintos, tornando dessa maneira as operações de busca por similaridade naturais dentro de espaços métricos.

Uma importante característica dos espaços métricos é a possibilidade de englobarem os espaços vetoriais, assim como os espaços adimensionais<sup>1</sup>, como por exemplo conjunto

 $<sup>^1\</sup>mathrm{Domínios}$ onde os objetos não podem ser identificados por um conjunto de coordenadas em eixos ortogonais

de imagens, palavras, sons ou cadeias de DNA, desde que, para isto, haja uma função de distância adequada.

#### 2.3 Extração de Características

A primeira etapa realizada para a análise de uma imagem médica é a extração de características. Na recuperação de imagens baseada em conteúdo uma característica diz respeito a uma medida numérica que tem por objetivo capturar uma determinada propriedade visual da imagem, podendo ser de escopo global, ou seja para a imagem como um todo, ou local onde focam-se pequenas regiões da imagem (regiões de interesse). Em meio às características mais utilizadas para descrever uma imagem de modo sucinto e comumentemente empregadas em CBIR estão, de fato, as características definidas como primitivas [Aslandogan & Yu, 1999] que são aquelas derivadas de três elementos fundamentais da imagem que são: distribuições de intensidades (cores), textura e forma.

Embora as características obtidas a partir da distribuição de intensidades de certa maneira obtenha resultados satisfatórios com relação a recuperação de imagens, existem situações onde o uso apenas desta informação não é suficiente para uma correta representação das mesmas, sendo assim, faz-se necessário o uso de outras características inerentes à imagem, como características de textura e/ou forma com o intuito de refinar o processo de recuperação de imagens. Além desse fato, por muitas vezes, sistemas que empregam apenas um único tipo de característica podem gerar resultados insatisfatórios, ou seja, não alcançam níveis de recuperação adequados. Dessa forma, os sistemas CBIR procuram utilizar múltiplas características das imagens para realizar a indexação e recuperação das mesmas [Zhang et al., 2000]. Uma descrição mais detalhada dos extratores de características é apresenta na seção 3.2 do capítulo 3.

Idealmente os extratores de características deveriam realizar a extração de características relevantes de uma dada imagem, de maneira similar a que um observador humano realizaria. Entretanto, apesar do amplo progresso obtido atualmente, tal preceito permanece ainda distante, devido ao ínfimo conhecimento referente à visão, cognição e emoção humana.

Um fato que contribui de maneira relevante para tal problema diz respeito a uma determinada imagem poder trazer significados variáveis se apresentada a observadores distintos [Neumann & Gegenfurtner, 2006], causando ambigüidades e gerando resultados insatisfatórios para uma operação de consulta. Tal fato pode ser explicado em decorrência do processo de percepção visual humana, em que o mundo o qual se enxerga, na verdade é uma construção realizada pelo cérebro a partir de informações exteriores. Deste modo, hipóteses são construídas, originárias de um histórico de informações, gerando assim subjetividade ao conceito de similaridade. Tais hipóteses por sua vez podem gerar *erros* (enganos), pois uma construção baseada em informações a *priori* é realizada. Portanto, a percepção é algo subjetivo e como dito anteriormente, usualmente a utilização de uma

única característica e representação não é suficiente [Zhang et al., 2000].

Inúmeros algoritmos de extração de características primitivas têm sido desenvolvidos. Porém, um ponto importante a ser ressaltado diz respeito a diversas pesquisas na área de recuperação por conteúdo indicarem que quanto mais especializada for a aplicação de um extrator de características para um certo domínio de imagens, menor será o *gap* semântico gerado. Tal *gap* ocorre pois as características de baixo nível extraídas automaticamente das imagens podem não representar de maneira satisfatória a interpretação semântica destas. Dessa forma, gerando resultados insatisfatórios para uma consulta por similaridade, pois tais resultados não condizem com a expectativa do usuário [Smeulders et al., 2000].

#### 2.4 Medidas de Similaridade

Um dos componentes fundamentais para um sistema com a finalidade de armazenamento e manipulação de imagens é a escolha das medidas de similaridade (ou funções de distância) pois, ao invés de buscas exatas, a recuperação baseada em conteúdo fundamenta-se no cálculo de similaridade entre uma dada imagem de consulta e um conjunto de imagens.

Como anteriormente citado, para a determinação do grau de similaridade entre os objetos, estes podem ser comparados diretamente ou a partir de vetores de características extraídos dos mesmos. Nos dois casos, uma função de distância d() é definida para calcular o grau de similaridade, retornando um valor numérico que quantifica o quão similar estes objetos são. Esses valores são sempre maiores ou iguais a zero, tendo valores próximos de zero para objetos muito similares (zero para objetos iguais) e valores maiores para objetos menos similares. Esta função é chamada de **função de dissimilaridade** ou de **função de distância** e deveria ser sempre definida por um especialista no domínio em questão.

Na realidade, a função de distância indica a dissimilaridade entre dois objetos. Para simplificar a compreensão de um texto, muitas vezes os autores estabelecem que a função de distância indica a similaridade, já que esta é obtida pela dissimilaridade. No entanto, tais funções de nada servem se as características das imagens (ou objetos complexos) não forem previamente extraídas, pois para o cálculo de similaridade, os objetos são representados por meio de suas características.

#### 2.5 Métodos de Acesso

Para que o processo de recuperação das imagens para uma dada consulta por similaridade seja otimizado são utilizados mecanismos para indexar as características previamente extraídas as quais irão representá-las. Os métodos de acesso (MA) compreendem estruturas de dados que permitem recuperar rapidamente uma dada informação armazenada. Considerando dados complexos, como imagens por exemplo, os métodos de acesso convencionais e já presentes nos Sistemas Gerenciadores de Bases de Dados (SGBD) não permitem efetuar recuperação por conteúdo e responder consultas por similaridade. Isso acontece pelo fato de não haver relação de ordem total em tais domínios de dados.

Para domínios de dados complexos como imagens, vídeos e outros dados multimídia, surgiu a necessidade do desenvolvimento de estruturas de indexação mais adequadas e que permitam a realização de consultas baseadas na similaridade dos dados (descritas na seção 2.2).

As primeiras estruturas de indexação foram desenvolvidas para operar com dados com os quais se pode estabelecer relação de ordem total, tais como a família de árvores B-tree [Bayer & McCreight, 1972], [Knuth, 1973], [Comer, 1979], [Folk et al., 1998] ou que tivessem relativamente poucos dados, por exemplo, as estruturas de *hashing*. As primeiras estruturas de indexação destinadas a suportar dados espaciais foram propostas para espaços de poucas dimensões. Tais estruturas são amplamente descritas na literatura, e são chamadas de **Métodos de Acesso Espaciais - SAM** (*Spatial Access Methods*). O desenvolvimento dessas estruturas iniciou-se com o trabalho pioneiro sobre as R-Trees [Guttman, 1984] e prosseguiu com numerosas variações e adaptações [Sellis et al., 1987] [Beckmann et al., 1990], [Hellerstein et al., 1995], [Papadias et al., 1999] (ou veja [Gaede & Günther, 1998] para obter uma descrição e comparação geral sobre estruturas espaciais). No entanto, para dados em espaços de altas dimensões ou mesmo adimensionais, nenhuma dessas estruturas mostrou-se adequada.

As estruturas de indexação para dados métricos, ou os **Métodos de Acesso Métricos** - **MAM** (*Metric Access Methods*), suportam naturalmente consultas por proximidade ou similaridade além de serem eficientes para dados de dimensões altas. Dessa forma, tem-se mostrado bastante apropriado utilizar um MAM para indexar imagens ou, mais propriamente, as características que foram extraídas das imagens suportando busca por similaridade. Tais características, ou descritores, podem ser atributos de forma, textura, histograma de cores [Lew et al., 2006, Aslandogan & Yu, 1999], resultados de transformações como por exemplo *singular value decomposition*, ou *Karhunen-Loeve* [Faloutsos, 1996]. Em [Torres & Falcão, 2006] é proposta a definição de descritores como sendo o par extrator de características e função de distância.

As características extraídas de antemão das imagens são utilizadas pelo MAM para construir a estrutura de índices. Durante essa construção são calculadas as distâncias entre tais características a serem inseridas na estrutura. Logo, as características e suas distâncias são basicamente as únicas informações que a árvore necessita das imagens.

Como pode ser observado, o principio básico dos MAMs é a representação de dados complexos em um espaço métrico. Formalmente, um espaço métrico é definido por um par  $\langle S, d() \rangle$ , onde S representa um domínio de dados e d() uma função de distância que necessariamente deve ser uma função de distância métrica, ou seja, deve respeitar as propriedades de **simetria**, **não negatividade** e **desigualdade triangular** que foram detalhadas na seção 2.2.3.

Dentre as principais estruturas de indexação para espaços métricos encontradas na literatura consideradas dinâmicas, ou seja, que permitem atualizações nos dados sem a necessidade de reconstrução da estrutura, estão a *M-tree* [Ciaccia et al., 1997], a *Slim-tree* [Traina et al., 2000a], a *DF-tree* [Traina et al., 2002b], a *DBM-tree* [Vieira & Traina Jr., 2004], e a *OMNI-Family* [Traina et al., 2007].

O método de acesso métrico *Slim-tree* será utilizado como suporte para esse projeto. Trata-se de uma estrutura balanceada e dinâmica que permite inserções posteriores à criação da árvore. Porém, quando comparada com a *M-tree* nas mesmas condições, sempre a sobrepujou, tanto em termos de número de acessos a disco quanto em termos de número de distâncias calculadas para responder a consultas por abrangência e portanto também em tempo total de execução.

A *Slim-tree* possui nós internos que são chamados de *indexnodes* e nós folha que são os *leafnodes*. O tamanho da página que armazena cada nó é de tamanho fixo, e cada nó pode armazenar um número máximo de objetos C (Capacidade). Além desses fatores a *Slim-tree* possui um novo algoritmo de divisão de nós (*splitting*) e métodos de avaliação e diminuição do grau de sobreposição entre os nós.

A Figura 2.4 (a) ilustra a representação dos nós no espaço métrico. Os círculos brancos referem-se aos nós folhas, e os de cor cinza aos nós índice, que são proporcionais ao raio de cobertura. Já na Figura 2.4 (b) é ilustrada a representação da estrutura de dados gerada pela construção da respectiva *Slim-tree* onde os pontos pretos situados em determinados nós são referentes aos objetos representantes do respectivo nó. Pode-se notar que o nó raiz não possui representante.



Figura 2.4: *Slim-tree* armazenando 17 objetos e possuindo *fan-out* igual a 3.

#### 2.6 Avaliação de Eficácia

Existem diversas maneiras de se analisar a recuperação de dados. A avaliação de eficácia em um sistema que possui a finalidade de recuperação de dados por conteúdo, é realizada principalmente por gráficos de **precisão** (*precision*) versus **revocação** (*recall*) [Baeza-Yates & Ribeiro-Neto, 1999], os quais são também utilizados nos sistemas de recuperação de informação tradicionais.

Além desses gráficos, existem variações que combinam a precisão e a revocação como a **medida E** e a **média harmônica** [Baeza-Yates & Ribeiro-Neto, 1999].

Para uma dada consulta, considera-se que R seja o conjunto de elementos relevantes em uma base de dados e A o conjunto de elementos retornados por tal consulta. Na Figura 2.5 é ilustrado o relacionamento entre os elementos recuperados na consulta (conjunto A) e os elementos relevantes existentes na base (conjunto R). A intersecção dos conjuntos R e A compreende os elementos relevantes que foram recuperados em resposta a uma dada consulta. A cardinalidade dos conjuntos R,  $A \in R \cap A$  é denotada respectivamente como |R|,  $|A| \in |R_A|$ . Pode-se observar que quanto maior a intersecção entre os conjuntos relevantes e recuperados,  $R \in A$  respectivamente, maior será a precisão do método.



Figura 2.5: Conjunto de resultados possíveis para uma dada operação de busca.

A medida de revocação pode ser entendida como a porção do conjunto de elementos relevantes (R) os quais foram recuperados segundo a requisição de uma dada consulta. Já a medida de precisão é a porção do conjunto de elementos recuperados (A) que são relevantes a uma dada consulta.

Formalmente as medidas de precisão e revocação são definidas como:

$$Revocação = \frac{|R_A|}{|R|} \tag{2.3}$$

$$Precisão = \frac{|R_A|}{|A|} \tag{2.4}$$

Um gráfico que mostra a relação entre precisão e revocação ilustra a precisão da resposta retornada após uma operação de busca considerando as características do dado complexo e a função de distância utilizada.

Quando um gráfico deste tipo atinge 100% de revocação, o método não apresenta falsos negativos, e quanto maior a precisão no gráfico, menor é a quantidade de falsos positivos.

Existem várias formas de se gerar tais gráficos. Uma delas é gerar para cada consulta seu gráfico de precisão e revocação, obtendo o gráfico resultante calculando-se a média das precisões pela fórmula  $\overline{P}(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}$ , ou seja, a precisão  $\overline{P}$  num dado nível de revocação r será a média aritmética das  $N_q$  precisões  $P_i$  de cada gráfico daquele nível.

Por exemplo, seja o conjunto  $R = \{o_1, o_4, o_6, o_8\}$  de objetos (imagens) relevantes e após a execução de uma dada consulta o conjunto  $A = \{o_1, o_2, o_3, o_4, o_5, o_6, o_7, o_8\}$  de objetos de fato retornados.

Observando-se o conjunto A dos elementos retornados, verifica-se que o primeiro ele-
mento  $(o_1)$  trata-se de um dos elementos pertencentes ao conjunto R, portanto relevante à consulta. Nesse caso a medida precisão possui valor de 100% em virtude de todos os elementos analisados, por enquanto apenas o primeiro no caso em questão, serem relevantes à consulta. Por outro lado neste ponto o valor da medida de revocação é de 25% pois um elemento relevante, dentre os quatro elementos relevantes (conjunto R) pertencentes à base foi recuperado. O próximo elemento relevante é o quarto  $(o_4)$ . Nesse ponto a medida precisão da consulta é 50%, pois dois dentre os quatro elementos da base (conjunto A) foram analisados até o momento e a medida de revocação é 50%, pois dois dos quatro elementos relevantes da base foram recuperados. A análise prossegue dessa maneira até que todos os elementos relevantes sejam verificados. Os valores de precisão e revocação são então visualizados graficamente, segundo o exemplo ilustrado na Figura 2.6.

Caso todos os elementos relevantes forem recuperados, o gráfico atinge revocação na marca de 100%. Como várias consultas geram gráficos com diferentes coordenadas atribuídas, é necessário realizar-se a interpolação em cada gráfico, onde a precisão interpolada num dado nível de revocação será  $P(r_j) = \max_{r_j \leq r \leq r_{j+1}} P(r)$ .



Figura 2.6: Exemplo de um gráfico de medidas de precisão e revocação.

# 2.7 Considerações Finais

A recuperação de imagens tem sido uma área de pesquisa muito ativa em banco de dados e visão computacional. Devido à crescente utilização de sistemas gerenciadores de bases de dados visando armazenar dados complexos, como informações multimídia, no caso em questão imagens médicas, e à contínua geração de grandes quantidades de imagens bem como o armazenamento das mesmas, culminaram na necessidade de se responder a consultas embasadas no conteúdo intrínseco de tais dados, nas quais apenas a similaridade entre os objetos pode ser considerada, além do desenvolvimento de métodos que permitissem a recuperação rápida e eficiente de tais dados. Apesar da proposta de muitos sistemas com o intuito de tratar do armazenamento e da recuperação de imagens de forma eficiente, uma solução ótima ainda está longe de ser alcançada.

O presente capítulo apresentou os principais conceitos e processos relacionados à recu-

peração de imagens baseada em conteúdo (CBIR) incluindo os principais tipos de consulta, extração de características, cálculo de similaridade, métodos de acesso para dados complexos, bem como a avaliação de desempenho de consultas com relação à precisão dos resultados retornados para um dado nível de revocação da base de imagens. O trabalho desenvolvido nesta pesquisa está intimamente ligado a todos estes conceitos, sendo que os resultados obtidos foram analisados segundo os gráficos de precisão e revocação gerados a partir da busca aos k-vizinhos mais próximos, bem como as características extraídas das imagens obtidas por meio de alguns extratores de características foram comparadas entre si utilizando inúmeras funções de distância (ambos conceitos descritos no capítulo subseqüente) e indexadas pelo método de acesso métrico *Slim-tree*.

E importante ressaltar que um ponto de extrema importância para a recuperação de imagens baseada em conteúdo diz respeito a dois componentes da estrutura de um sistema CBIR, que são os extratores de características e as funções de distância. Para que resultados satisfatórios sejam retornados considerando um dada consulta por similaridade é imperativo que seja estudada e analisada qual a melhor integração entre esses dois componentes. É desse ponto específico que o próximo capítulo irá tratar.

Capítulo 3

# Binômio: Relacionamento entre Função de Distância e Características Intrínsecas dos Dados

# 3.1 Introdução

recuperação de dados complexos depende intimamente de um extrator de características (*features*) o qual provê as características inerentes dos dados complexos, e de um medida (função de distância) a qual é utilizada para quantificar a similaridade entre esses dados. Um grande desafio ainda existente refere-se a como integrar esses dois aspectos chaves, visando a melhoria e aperfeiçoamento dos dados retornados para uma dada consulta por similaridade.

Atualmente o volume de dados complexos gerados e/ou manipulados vem crescendo em um ritmo acelerado. Portanto, o desafio de lidar com tais dados, fundamenta-se principalmente em dois aspectos:

- Como organizar (indexar) os dados objetivando um acesso rápido e eficiente;
- Como retornar os dados em concordância com a expectativa do observador humano.

Para lidar com o primeiro aspecto referente à indexação de dados complexos, os quais são freqüentemente multi-dimensionais ou até mesmo adimensionais, são empregados métodos de acesso para dados complexos provendo assim uma maneira rápida de recuperação de tais dados [Chávez et al., 2001]. Quando se deseja a indexação e recuperação de dados complexos, as estruturas tradicionais não se aplicam ou simplesmente têm pouca utilidade, pois tais dados na sua grande maioria não apresentam relação de ordem. Portanto surge a questão de como ordenar tais dados, pois não se consegue estabelecer uma relação de ordem entre estes, no entanto pode-se dizer o quanto um dado complexo, por exemplo uma imagem, é similar (parecida) com outra imagem.

Os métodos de acesso mais adequados para a realização de consultas por similaridade (ver Seção 2.2) são os Métodos de Acesso Métrico (MAMs), pois sua estrutura de construção é baseada nas distâncias entre elementos [Chávez et al., 2001]. Para que uma função de distância possa ser empregada juntamente com um MAM ela deve satisfazer aos axiomas métricos (não-negatividade, simetria e desigualdade triangular - ver Seção 2.2). O desenvolvimento de funções de distância ou a respectiva melhoria das já existentes depende de quais características extraídas dos dados serão utilizadas para a comparação de similaridade. Portanto, esse fato corrobora que é necessário haver o desenvolvimento e aprimoramento conjunto entre ambos os aspectos, função de distância e extração das características mais relevantes das imagens, o que chamamos de **binômio de recuperação por conteúdo**.

O segundo aspecto refere-se ao processo inicial da extração das características mais significantes, segundo um critério específico, para realizar a comparação entre dois dados complexos. Os dados são representados por estas características obtidas, as quais são organizadas em vetores de características. Posteriormente, esses dados são indexados, recuperados e comparados por meio de seus vetores de características. Para a realização dessa comparação é empregada um função de distância, a qual irá mensurar o quão similar um dado é em relação a outro. Dessa maneira, existe um forte inter-relacionamento entre as características extraídas e a função de distância utilizada, visando adequar-se à expectativa do observador humano com relação a tal comparação.

Por mais de uma década, pesquisadores têm explorado a área de análise e recuperação de dados complexos (imagens, áudio, vídeo, séries temporais e outros). Entretanto, a maioria dos trabalhos encontrados na literatura que tratam da indexação e recuperação de dados complexos negligencia tal inter-relacionamento e acaba por fazer uso das mais conhecidas e amplamente utilizadas funções de distância, tais como a distância Euclidiana e outros membros da família  $L_p$  (descritos na seção 3.3.1).

Porém, um problema ainda permanece sem solução adequada: como mensurar/atingir a similaridade perceptual entre dois objetos.

Para conseguir uma recuperação eficaz é necessário quantificar e caracterizar o melhor possível a similaridade. No entanto, muitas vezes as funções de distância não evidenciam imagens obviamente similares.

No presente capítulo é abordado e demonstrado que uma escolha cuidadosa de determinadas funções de distância agregadas a determinados extratores de características, resulta em uma considerável melhoria na qualidade dos dados retornados para uma dada consulta por similaridade.

## **3.2** Extratores de Características

O poder de discriminação das características extraídas dos dados multimídia, no caso em questão imagens médicas, é um dos aspectos chaves relacionados a comparação de similaridade entre dois dados complexos (imagens, vídeos, sons, séries temporais, seqüências de DNA, entre outros). As características extraídas de cada imagem são organizadas em um vetor denominado vetor de características, o qual pode ser entendido como um ponto n-dimensional em um espaço vetorial.

Na presente seção são introduzidas e explicitadas as técnicas de extração de características de imagens médicas mais relevantes e que são utilizadas neste projeto, as quais visam abstrair as informações relevantes contidas em uma determinada imagem.

## 3.2.1 Histograma Tradicional de Níveis de Cinza

As características baseadas em cores são as mais utilizadas para recuperação por conteúdo, principalmente por apresentarem custo computacional reduzido. A identificação de uma imagem por meio da característica relativa à distribuição global de cor é geralmente realizada pela construção de um histograma de cor. No histograma são calculados o número de pixels da imagem para cada cor (ou nível de cinza em imagens monocromáticas). Um tipo de medida é a análise dos tons de cinza, muito utilizada para imagens médicas, onde a imagem é analisada e seu histograma (também conhecido como gray-level histogram ou brightness histogram) é gerado e utilizado para comparação em buscas por similaridade. Para comparar computacionalmente duas imagens, utilizam-se seus respectivos histogramas e a comparação usualmente é realizada por meio da utilização da função de distância  $L_1$  ou variações da norma  $L_p$  (ver seção 3.3.1).

Em CBIR, o uso de histogramas como passo inicial na seleção de imagens, justifica-se ainda mais devido ao fato de algoritmos que lidam com outras características intrínsecas - textura e forma - apresentarem freqüentemente um custo computacional alto, além de serem dependentes do domínio da aplicação. Histogramas, por outro lado, são obtidos com custo linear em relação ao número de pixels da imagem. Portanto, um procedimento eficiente para recuperação de imagens seria o uso de histogramas para filtrar e reduzir o conjunto de imagens candidatas, como um passo preliminar do processo, seguindo-se de métodos mais precisos e custosos [Alferez et al., 2005].

No entanto, os histogramas tradicionais de níveis de cinza apresentam alguns pontos a serem considerados. Quando analisadas duas imagens distintas entre si e seus respectivos histogramas, estes podem possuir alta semelhança mesmo tratando-se de imagens completamente distintas, uma vez que o método é estatístico e não apresenta informação sobre a distribuição espacial das cores. Outro ponto a ser considerado diz respeito à elevada dimensão dos vetores de características gerados, freqüentemente em torno de 256 posições considerando uma imagem médica com 256 níveis de cinza. Esse fato da elevada dimensionalidade dos vetores recai sobre o problema da chamada maldição da alta dimensionalidade (*dimensionality curse*) [Beyer et al., 1999], tornando-se um grande problema principalmente para as estruturas de indexação.

## 3.2.2 Histogramas Métricos

Em [Bueno, 2002] foi proposto um novo método denominado **Histograma Métrico**, o qual configura-se como uma nova abordagem de extração de características baseada em intensidade de cor. Tal método comparativo conserva a curva original do histograma de brilho de uma dada imagem mantendo de maneira eficiente a informação deste e reduzindo o número de *bins* do histograma, dessa forma, otimizando o custo computacional da operação de busca.

Obter um histograma normalizado de imagens não é uma operação altamente custosa. O histograma normalizado é invariante a transformações geométricas, bem como a transformações de luminosidade. Dessa forma, tais histogramas permitem comparações de imagens de tamanhos variados, portanto transformações geométricas realizadas nas imagens fontes resultam em um histograma análogo.

Um histograma normalizado é composto por uma quantidade específica de *bins*, a qual depende da resolução de luminosidade da imagem, apresentando uma cardinalidade fixa. Já em um **histograma métrico**, o equivalente ao *bin* do histograma tradicional é denominado *bucket*, onde estes correspondem a um segmento de linha na aproximação do histograma normalizado, ou seja, os *buckets* na verdade correspondem a um subconjunto de *bins* do histograma original. Portanto dessa maneira é possível limitar com caráter confiável o formato do histograma e também reduzir a dimensionalidade deste. Na Figura 3.1 são ilustrados os *bins* e os *buckets* de um dado histograma.

Entretanto resultante de tal redução, observa-se que o número de *buckets* dos histogramas métricos, bem como sua largura é variável e dependente de características inerentes de cada imagem. Surge assim um problema, pois normalmente o cálculo de distância entre histogramas é dado pela somatória da diferença entre os *bins* de dois histogramas. Já os histogramas métricos devido a tal característica de variação do número e distribuição dos *buckets* não podem ser comparados pelas funções de distâncias tradicionais, ou seja, da mesma maneira que é realizada com os histogramas convencionais. Portanto surgiu a necessidade da definição de uma nova função de distância denominada **Distância Métrica** (DM) a qual se baseia no cálculo de diferença de áreas entre dois histogramas [Traina et al., 2002a].

A função de distância entre dois **histogramas métricos**, de maneira simplista, é dada pela área não sobreposta entre as curvas que representam os respectivos histogramas métricos, ou seja, pode ser definida pelo grau de sobreposição das áreas das distribuições de freqüência, baseado no cálculo da diferença entre as áreas destes. Tal função de distância, denominada DM, entre dois **histogramas métricos**  $M_H(A) \in M_H(B)$  respectivamente, onde  $bm = \max(b_{N_A-1}, b_{N_B-1}) \in M_H(Imagem, x)$  é a função contínua que representa o histograma métrico, e é definida formalmente como:

$$DM\left(M_{H}\left(A\right), M_{H}\left(B\right)\right) = \int_{x=0}^{bm} \left|M_{H}\left(A_{\langle bx, hx \rangle}\right) - M_{H}\left(B_{\langle bx, hx \rangle}\right)\right| dx$$
(3.1)

Sendo que o histograma métrico de uma dada imagem A é definido como  $M_H(A) = \{N_A, \langle bx, hx \rangle | 0 \le x \le N_A\}$ , que é um conjunto de  $N_A$  buckets formados por pares  $\langle bx, hx \rangle$  consecutivos, onde  $b_x$  indica largura e  $h_x$  altura.

O histograma métrico apresenta um ganho efetivo de desempenho, considerando o processo de indexação e recuperação de imagens, cerca de 4 vezes mais eficiente se comparado aos histogramas convencionais, mantendo a eficiência da recuperação de imagens.



Figura 3.1: Histograma normalizado com pontos de controle de máximo e mínimo local  $\langle bx, hx \rangle$  os quais definem os *buckets* correspondentes ao seu **histograma métrico** [Bueno, 2002].

Em [Campo & Traina, 2003] foi testada e avaliada efetivamente a característica do **histograma métrico** ser invariante a transformações de brilho, característica esta de suma importância devido a imagens de um mesmo paciente e exame ao serem adquiridas de maneiras distintas tendem a apresentar tais variações. Sendo assim, ao utilizar histogramas convencionais estes podem não retornar as imagens semelhantes quando há variações dessa natureza.

Em [Bueno, 2002] são apresentados detalhadamente os **histogramas métricos** e suas respectivas aplicações em recuperação de imagens.

Outra função de distância para histogramas métricos, proposta em [Felipe et al., 2005], foi denominada *Global Warp Distance* (GWD), a qual emptrega o método *Warp*, o qual permite realizar um ajuste de um dado histograma em relação a outro histograma durante o cálculo de distância. O método *Warp* fundamenta-se na técnica denominada *Dynamic Time Warping* [Keogh & Ratanamahatana, 2005] [Keogh, 2002]. Aplicado aos histogramas métricos em associação com as características globais obtidas dos mesmos, tais como, simetria, curtose e número de picos, obtém-se a função de distância GWD. Estas características globais são extremamente significantes em relação aos histogramas obtidos a partir de imagens médicas de diferentes regiões do corpo humano, pois geralmente essas imagens apresentam características globais específicas nos seus histogramas. Portanto, diferentes categorias de imagens usualmente apresentam características globais peculiares. Por exemplo, imagens de cérebro que possuem alguma anomalia geralmente apresentam um pico adicional no seu respectivo histograma em comparação com histogramas de imagens de cérebro normais.

## 3.2.3 Matrizes de Co-ocorrência

Em meio às características intrínsecas utilizadas para representar imagens médicas em sistemas CBIR, os extratores baseados em textura destacam-se pela sua eficácia em discriminar tecidos. As medidas de textura capturam essencialmente a granularidade e padrões repetitivos na distribuição de pixels. Por exemplo, se um radiologista tem em mãos uma imagem que apresenta um nódulo de mama e outra imagem que apresenta um nódulo de pulmão, essas imagens serão diferenciadas entre si tanto pela suavidade da textura quanto pela repetição de padrões.

Entre algumas das técnicas para extração de características baseadas em textura estão os filtros de Gabor [Santini & Jain, 1996], as transformadas wavelets [Santini & Gupta, 2001]. Esses métodos tentam capturar partes da imagem com relação à mudança de direção e escala, e são muito úteis para imagens ou regiões com texturas homogêneas. Em [Zegarra et al., 2008] são avaliadas algumas medidas de similaridade para um extrator baseado em wavelet.

Uma das mais populares técnicas de extração de características baseadas em textura que tem sido largamente utilizada, são as matrizes de co-ocorrência [Haralick et al., 1973] [Haralick, 1979].

As matrizes de co-ocorrência possibilitam uma abordagem estatística para descrição das texturas da imagem, gerando resultados satisfatórios com custo computacional reduzido [Sonka et al., 1998]. Nas matrizes de co-ocorrência os índices das linhas e colunas representam os diferentes valores de níveis de cinza, computando-se a freqüência com que os mesmos ocorrem dois a dois em uma certa direção e uma certa distância.

Dada uma imagem h com um conjunto discreto de tons de cinza G, define-se a matriz de co-ocorrência  $P_{d,\theta}(i,j)$ , onde cada elemento (i,j) corresponde a um número inteiro que indica quantas vezes um pixel  $p_1$  de nível cinza i aparece distante de um pixel  $p_2$  de intensidade j por uma distância d e um ângulo  $\theta$ . Formalmente, cada elemento (i,j) de  $P_{d,\theta}(i,j)$  indica quantas vezes ocorre  $p_1 = p_2(d\cos\theta, d\sin\theta)$ , onde  $G(p_1) = i$  e  $G(p_2) = j$ .

As matrizes de co-ocorrência são matrizes quadradas e simétricas em relação a diagonal principal. Assim, tem-se que  $P_{d,\theta}(i,j) = P_{d,\theta}(j,i)$ . As dimensões das matrizes  $P_{d,\theta}$  é definida pela quantidade de níveis de cinza da imagem. Devido a esse fato, computar inúmeras matrizes de grandes dimensões torna-se praticamente inviável, sendo assim faz-se necessária a utilização de algoritmos para redução da quantidade de tons de cinza objetivando a obtenção de matrizes reduzidas, pois é necessário gerar um número considerável de matrizes de co-ocorrência utilizando diferentes ângulos  $\theta$  e distâncias d para que haja

0	0	1	1	P	$\begin{bmatrix} 4\\ 2 \end{bmatrix}$	$\frac{2}{4}$	$\begin{array}{c} 1 \\ 0 \end{array}$	0 0	$P_{1,135^{\circ}} =$	2 1	$\frac{1}{2}$	$\frac{3}{1}$	0 0
0	2	2	2	<b>1</b> 1,0 <sup>5</sup> -	1	0	6	1	1,100	3	1	0	2
2	2	3	3		0	0	1	2 -		0	Ű	2	
(a)						(l	)				$(\epsilon$	:)	

Figura 3.2: Exemplos de matrizes de co-ocorrência. (a) matrix de pixels da imagem; (b) matriz de co-ocorrência referente ao ângulo 0 e d = 1; (c) matriz de co-ocorrência referente ao ângulo  $135^{\circ} e d = 1$ .

uma melhoria na capacidade de diferenciação de texturas realizada pelo método.

Inúmeras medidas podem ser extraídas das matrizes de co-ocorrência. Em [Haralick et al., 1973] um conjunto de 14 descritores é proposto, dentre estes descritores, os mais utilizados segundo [Felipe et al., 2003] são apresentados na Tabela 3.1, onde  $P_{d,\theta}$  indica uma matriz de co-ocorrência.

Descritor	Equação	Semântica		
Variância	$\sum_{i,j} (i,j)^2 P_{d,\theta}(i,j)$	nível de contraste da imagem		
Step	$\sum_{i,j} P_{d,\theta}(i,j)$	intensidade		
Entropia	$-\sum_{i,j} P_{d,\theta}(i,j) log P_{d,\theta}(i,j)$	suavidade da imagem		
Energia	$\sum_{i,j} P_{d,\theta}^2(i,j)$	uniformidade da imagem		
Homogeneidade	$\sum_{i,j} P_{d,\theta}(i,j)/1 + (i-j)^2$	homogeneidade da distribuição dos pixels		
Momento de $3^a$ Ordem	$\sum_{i,j} (i,j)^3 P_{d,\theta}(i,j)$	nível de distorção da imagem		
Variância Inversa	$\sum_{i,j} P_{d,\theta}(i,j)/(i-j)^2$	nível inverso de contraste		

Tabela 3.1: Descritores de Haralick amplamente utilizados.

Além dos descritores citados na Tabela 3.1, em [Felipe et al., 2003] é apresentado um novo descritor denominado *Gradiente*, bem como também indica uma combinação ótima de três descritores (gradiente, homogeneidade e entropia) para recuperar imagens por similaridade, apresentando bons resultados com relação a precisão.

A seleção dos descritores a serem adotados geralmente é baseada em testes empíricos sobre um dado domínio, comparando-se os resultados que os mesmos produzem, sendo elencado o descritor que gerou o resultado mais satisfatório.

### 3.2.4 Momentos de Zernike

Os momentos de Zernike são utilizados para representar formas complexas, compostas de várias regiões disjuntas. Os polinômios de Zernike, dizem respeito a um conjunto de polinômios complexos que formam uma base ortogonal no interior de um circulo unitário  $x^2 + y^2 \leq 1$  [Khotanzad & Hong, 1990].

Estes polinômios podem ser formalmente denotados por meio das coordenadas polares como:

$$V_{n,m}(x,y) = V_{n,m}(\rho \cos(\theta), \rho \sin(\theta)) = R_{n,m}(\rho)e^{im\theta}$$
(3.2)

onde n é um valor inteiro, maior ou igual a zero, que define a ordem do polinômio; e m, trata-se também de um valor inteiro, que pode assumir valores tanto positivos quanto negativos, o qual descreve a dependência angular, ou rotação, do polinômio. Considerando os valores n e m, estes devem satisfazer às seguintes condições:

$$n - |m|$$
 deve ser par (3.3)

$$-n \le m \le n \tag{3.4}$$

Para realizar o cálculo dos momentos de Zernike, a imagem (ou região de interesse) deve primeiramente ser mapeada das coordenadas do plano cartesiano para o disco unitário por meio das coordenadas polares, onde o centro da imagem é a origem do disco unitário. Os pixels da imagem mapeados para fora do disco não são utilizados no cálculo dos momentos. Para realizar tal mapeamento das coordenadas cartesianas (x, y) para o sistema polar, calcula-se:  $\rho = \sqrt{x^2 + y^2}$  e  $\theta = \arctan(\frac{y}{x})$ , onde  $\rho$  diz respeito à distância do ponto (x, y)à origem, e  $\theta$  denota o ângulo formado entre o vetor de norma (comprimento)  $\rho$  e o eixo x, no sentido anti-horário. A Figura 3.3 ilustra tal representação, onde a origem é denotada por O, o par (x, y) são as coordenadas cartesianas do ponto T e  $(\rho, \theta)$  suas coordenadas polares, sendo que  $x = \rho * \cos\theta$  e  $y = \rho * \sin\theta$ .



Figura 3.3: Representação das coordenadas polares de um ponto T sobre o plano euclidiano  $\Re^2$  cuja origem é denotada por O.

A função  $R_{n,m}(\rho)$  é denominada polinômio radial e pode ser definida formalmente como:

$$R_{n,m}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2}-s\right)! \left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s}$$
(3.5)

Pode-se notar que  $R_{n,m}(\rho) = R_{n,-m}(\rho)$ . Caso as condições descritas pelas restrições 3.3 e 3.4 não sejam satisfeitas, o polinômio  $R_{n,m}(\rho)$  é nulo. Os seis polinômios radiais são:

$$R_{0,0}(\rho) = 1 \qquad R_{1,1}(\rho) = \rho$$
$$R_{2,0}(\rho) = 2\rho^2 - 1 \qquad R_{2,2}(\rho) = \rho^2$$
$$R_{3,1}(\rho) = 3\rho^3 - 3\rho \qquad R_{3,3}(\rho) = \rho^3$$

Basicamente, os momentos de Zernike são uma projeção de uma dada imagem f(x, y)sobre a base ortogonal formada pelos polinômios  $V_{n,m}(x, y)$ . O momento de Zernike  $Z_{n,m}$ de ordem n e repetição (ou dependência angular) m, calculado para uma dada imagem digital, é definido formalmente como:

$$Z_{n,m} = \frac{n+1}{\pi} \sum_{x} \sum_{y} f(x,y) \ V_{n,m}^*(x,y), \qquad x^2 + y^2 \le 1$$

onde  $V_{n,m}^*$  denota o conjugado do valor complexo de  $V_{n,m}$ .

Observa-se que segundo as condições 3.3 e 3.4 para cada ordem n existem  $\frac{n}{2}$  momentos com repetição, ou dependência angular m, distintos. Dessa forma, existem (n+1)(n-1)/2 momentos de ordem inferior ou igual a n. Outro ponto a ser ressaltado diz respeito à relação entre os momentos de ordem n:  $Z_{n,m}^* = Z_{n,-m}$ .

Comparações entre diversos extratores de forma usualmente empregados em CBIR foram realizadas em [Zhang & Lu, 2001], analisando suas limitações, robustez, complexidade computacional e mensuração da similaridade perceptual obtida. Considerando a avaliação geral de tais análises os momentos de Zernike apresentaram os melhores resultados. Tais momentos mostraram-se bastante eficazes na representação de formas complexas, compostas de várias regiões disjuntas [Twa et al., 2003].

Vale ressaltar que os momentos de Zernike são intrinsecamente invariantes à transformações de rotação nas imagens. Porém, nada garante que sejam invariantes a transformações de translação e escala. Para que os momentos de Zernike sejam invariantes a operações de translação é necessário que a origem do disco unitário coincida com o centro de massa da imagem no momento do mapeamento. Já a invariância à escala é obtida escalonando-se a imagem de modo que a sua massa  $(m_{0,0})$  passe a assumir um valor  $\beta$ definido a priori. Sendo que tal fato pode ser realizado por meio da seguinte formulação:

$$h(x,y) = f\left(\frac{x}{\alpha} + x_o, \frac{y}{\alpha} + y_o\right)$$
 onde  $\alpha = \sqrt{\frac{\beta}{m_{0,0}}}$ 

sendo que para tal formulação o par de coordenadas  $(x_o, y_o)$  denota o centro de massa do objeto. Tal centro de massa é calculado a partir da formulação  $m_{p,q} = \sum_x \sum_y x^p y^q$  tal que:

$$x_o = \frac{m_{1,0}}{m_{0,0}}$$
 e  $y_o = \frac{m_{0,1}}{m_{0,0}}$ 

Portanto, os momentos de Zernike invariantes às transformações da imagem com relação a rotação, translação e escala, são computados a partir da função h(x, y).

## 3.2.5 Variação do Método EM/MPM

Uma variação do método EM/MPM [Comer & Delp, 2000] foi proposta em [Balan et al., 2005] a qual representa as imagens extraindo medidas estatísticas de regiões obtidas por meio da segmentação de uma dada imagem. Para tanto, o método proposto, como o próprio nome explicita, utiliza a combinação do algoritmo EM (*Expectation and Maximization*) para realizar o ajuste de um Modelo de Mistura Gaussiano (*Gaussian Mixture Model*) no histograma da imagem com um classificador baseado nos Campos Aleatórios de Markov, o qual incorpora o conceito de coerência espacial entre os pixels [Balan et al., 2005].

A variação introduzida baseia-se na aplicação da técnica Annealing ao algoritmo de segmentação EM/MPM, a qual realiza um aumento gradativo do valor do parâmetro de interação espacial do modelo de Markov ( $\beta$ ) durante o processo de segmentação de uma dada imagem. Para um detalhamento aprofundado do algoritmo consultar [Balan, 2007].

Na Figura 3.4 é ilustrado um exemplo dos resultados obtidos por meio da aplicação do método EM/MPM para segmentação de uma imagem de cérebro, utilizando-se 5 classes, incluindo a classe de pixels do plano de fundo. O número de classes escolhido deve ser o mais próximo possível do número médio de tecidos que o exame de ressonância magnética é capaz de diferenciar a partir do contraste de níveis de cinza da imagem.



Figura 3.4: Exemplo de segmentação realizada pela variação do método EM/MPM. (a) imagem original; (b) imagem segmentada com L = 5; (c) região de classe 1 (fluido cérebroespinhal); (d) região de classe 2 (massa cinzenta); (e) região de classe 3 (massa branca); (f) região de classe 4 (dura, medula óssea, gordura) [Balan et al., 2005]

Após as imagens serem segmentadas pela textura conforme o número de classes (L)especificado, para cada região segmentada são computadas cinco medidas, as quais serão utilizadas para representar estas regiões. As cinco medidas são: (1) a dimensão fractal  $D_2$ ; (2) o centro de massa, ou centróide  $x_o e y_o$ ; (3) a massa, ou tamanho m; (4); o nível de cinza médio  $(\mu)$  e o coeficiente linear utilizado para estimar  $D_2$  (b). Todas as medidas são valores reais e, sendo que o centróide da região é dado por duas coordenadas, cada região é representada por seis valores reais. Como a região de plano de fundo (*background*) da imagem segmentada não é considerada na caracterização da imagem o tamanho fixo do vetor de características proposto por Balan é definido por 6(L-1) valores reais, onde Lé o número de classes para segmentação. As características são agrupadas por região no vetor e dispostas segunda a ilustração da Figura 3.5.

Devido às propriedades estocásticas do algoritmo, uma ordenação auxiliar é sempre



Figura 3.5: Vetor de características das imagens: 6(L-1) valores reais [Balan et al., 2005].

realizada para garantir que os grupos de características de cada classe estejam ordenados pela característica de nível de cinza médio ( $\mu$ ). Ou seja, o primeiro grupo de seis características no vetor está associado à classe de nível de cinza mais escuro, enquanto os demais grupos estão ordenados de maneira crescente de acordo com o nível médio de cinza correspondente. Esta ordenação é importante porque o nível de cinza médio em imagens de ressonância magnética, conforme já mencionado, está ligado à própria identificação dos tecidos na imagem. Assim, a ordenação dos grupos proporciona a correspondência das características com relação ao tecido que elas representam. Ainda que esta correspondência dos tecidos não ocorra entre imagens de categorias diferentes, cérebro e espinha, por exemplo, ela ainda é importante para manter o agrupamento no espaço vetorial das imagens similares de mesma categoria.

A variação do método EM/MPM proposta em [Balan et al., 2005] recai sobre a proposta de um vetor de características compacto, de tamanho fixo e que apresenta-se eficiente para consultas por similaridade em imagens.

## 3.3 Funções de Distância

Na literatura, pode-se verificar uma crescente necessidade pelo desenvolvimento de funções de distância, pois estas são empregadas nas mais diversas áreas do conhecimento. Especificamente para a recuperação de imagens com o intuito de atingir a máxima eficácia na recuperação, muitas deixam a desejar com relação ao problema do **gap semântico** [Smeulders et al., 2000].

## 3.3.1 Família de Distâncias Minkowski

Um aspecto de suma importância pertencente aos espaços métricos diz respeito à capacidade destes englobarem os espaços vetoriais, assim como espaços adimensionais, que não apresentam uma dimensão fixa. No entanto para que o emprego de tal característica seja admissível é necessário a existência de uma função de distância adequada.

A família de distâncias Minkowski [Wilson & Martinez, 1997] também conhecida como  $L_p$ , é comumente empregada em espaços vetoriais. Um espaço é denominado **Espaço Ve**torial se os objetos de um dado domínio  $\mathfrak{D}$  correspondem a valores numéricos estruturados em vetores, desta forma as características extraídas das imagens possuem dimensões fixas podendo assim serem passivas de indexação por métodos de acesso métricos, caso seja associado ao domínio de dados uma função de distância métrica. Os objetos pertencentes a um espaço vetorial de dimensão n (n-dimensional) são representados por n coordenadas de valores reais { $x_1,..., x_n$ }. Sendo assim o conjunto de funções de distância  $L_p$  é definido como:

$$L_p((x_1, ..., x_n), (y_1, ..., y_n)) = \sqrt[p]{\sum_{i=1}^n |x_i - y_i|^p}$$
(3.6)

Note que a função de distância Euclidiana é a  $L_2$  definida dessa forma. De acordo com o valor atribuído a p obtém-se as variações da família  $L_p$ .

A função de distância  $L_1$ , também denominada *City Block* ou *Manhattan*, corresponde ao somatório do módulo das diferenças entre as coordenadas. Nesse caso, o conjunto de pontos equidistantes a um dado raio de abrangência  $r_q$  forma um losango com os diâmetros paralelos aos eixos das coordenadas. A  $L_1$  é definida formalmente como:

$$d_{L_1}(X,Y) = \sum_{i=1}^{n} |x_i - y_i|$$
(3.7)

A função de distância  $L_2$  denominada Euclidiana, corresponde à função comumente utilizada para o cálculo de distância entre vetores. O conjunto de pontos de mesma distância para  $L_2$  forma uma circunferência. Tal distância é definida formalmente como:

$$d_{L_2}(X,Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$
(3.8)

Com o intuito de otimizar o tempo de cálculo, a  $L_2$  é utilizada na sua forma quadrática, ou seja, sem a extração da raíz quadrada. Essa modificação preserva a distância entre os pontos, retornando o mesmo conjunto resposta.

Calculando-se o limite da equação 3.6 quando p tende ao infinito, obtém-se a função de distância  $L_{\infty}$ , conhecida também como *Infinity* ou *Chebychev*, na qual o conjunto de pontos equidistantes formam um quadrado com os lados paralelos aos eixos das coordenadas. A  $L_{\infty}$  é definida formalmente como:

$$L_{\infty}((x_1, ..., x_n), (y_1, ..., y_n)) = \max_{i=1}^n |x_i - y_i|$$
(3.9)

Na Figura 3.6 são ilustradas as formas geométricas das regiões com mesma distância geradas pelas funções  $L_1$  Figura 3.6 (a),  $L_2$  Figura 3.6 (b) e  $L_{\infty}$  Figura 3.6 (c) em um espaço bidimensional. A sobreposição das representações geométricas das funções na Figura 3.6 (d) distingue os subespaços distintos as quais as diferentes funções de distância englobam.

#### 3.3.1.1 Distância Minkowski Ponderada

A distância Minkowski ponderada é uma variação da família de Distâncias Minkowski, a qual atribui-se pesos distintos para cada atributo inerente a uma dada imagem. A



Figura 3.6: Representação das formas geométricas geradas conforme a distância  $L_p$  utilizada e em relação aos pontos equidistantes à distância  $r_q$  a partir do objeto centro  $o_q$ .

utilização de um vetor de ponderação dá-se quando houver diferentes influências entre os atributos sobre o resultado da similaridade. Portanto, diferentes vetores de ponderação podem ser estipulados a contextos específicos, exercendo um papel de calibragem da Distância Minkowski para determinados contextos, levando-se em consideração que determinados atributos podem exercer maior ou menor relevância em relação a outros de acordo com o contexto.

A Distância Minkowski ponderada é definida formalmente como:

$$d_{L_p}(X,Y) = \sqrt[p]{\sum_{i=1}^{n} w_i (x_i - y_i)^p}$$
(3.10)

onde  $w_i$  é o vetor de ponderação  $w = (w_1, w_2, ..., w_n)$ .

### 3.3.2 Funções de Distâncias Consolidadas

Ao invés da realização de uma equiparação/comparação exata, a busca baseada em conteúdo calcula a similaridade entre uma dada imagem de consulta e imagens contidas em uma base de dados. Desta forma, o resultado retornado pelo processo de recuperação não se trata de uma única imagem, mas sim de uma coleção de imagens classificadas em relação aos seus fatores de similaridade com a respectiva imagem de consulta.

Muitas medidas de similaridade têm sido desenvolvidas para recuperação de imagens baseada em estimativas empíricas da distribuição de características. Porém a escolha das funções de distância empregadas muitas vezes são feitas sem critério, ou seja, sem uma fundamentação teórica e não levando em consideração as características explícitas e também implícitas a um determinado contexto específico, como as particularidades semânticas do ambiente ao qual devem ser inseridas, a distribuição estatística intrínseca a este, bem como o modelo de extratores de características e o gênero de dados obtidos por estes, entre outras. Conseqüentemente, diferentes medidas de similaridade e/ou funções de distância poderão afetar de maneira significativa a eficácia da recuperação de imagens.

Além das funções mais utilizadas como as da família **Minkowski** - (Lp), existem também algumas outras comumente utilizadas, as quais serão introduzidas na presente seção.

Na presente seção são considerados dois vetores de características  $X = \{x_1, ..., x_n\}$ e  $Y = \{y_1, ..., y_n\}$  respectivamente referentes a uma dada imagem de consulta X e uma imagem Y pertencente a uma base de dados, sendo que a distância entre tais imagens é denotada por d(X, Y).

#### 3.3.2.1 Divergência de Kullback-Leibler e Divergência de Jeffrey

A divergência de Kullback-Leibler (DKL) do ponto de vista da teoria da informação, mensura quão ineficiente, em média, seria codificar um determinado vetor de características a partir de outro, utilizado como referência. No entanto, a DKL é não-simétrica e sensível, por exemplo, à variação da divisão dos *bins*<sup>1</sup> de um histograma.

A divergência de Kullback-Leibler é definida formalmente como:

$$d_{KL}(X,Y) = \sum_{i=1}^{n} y_i \log \frac{y_i}{x_i}$$
(3.11)

Derivada empiricamente, a partir da DKL, a divergência de Jeffrey, é uma modificação desta. Em contraste com a DKL a divergência de Jeffrey é simétrica e apresenta um comportamento numericamente mais estável e robusto em relação a ruídos e ao tamanho dos *bins* de um dado histograma [Rubner & Tomasi, 2001].

Tal divergência é definida formalmente como:

$$d_J(X,Y) = \sum_{i=1}^n \left( y_i \log \frac{y_i}{m_i} + x_i \log \frac{x_i}{m_i} \right)$$
(3.12)

onde  $m_i = \frac{y_i + x_i}{2}$ .

## **3.3.2.2** Valor Estatístico $\chi^2$

O valor estatístico  $\chi^2$  é calculado encontrando-se a diferença entre cada freqüência observada e teórica para cada resultado possível, elevando-as ao quadrado, e dividindo cada uma pela freqüência teórica, e por fim realizando a soma dos resultados. Para tanto é definida formalmente como:

$$d_{\chi^2}(X,Y) = \sum_{i=1}^n \frac{(y_i - m_i)^2}{m_i}$$
(3.13)

onde novamente  $m_i = \frac{y_i + x_i}{2}$ .

Essa função valoriza as elevadas discrepâncias existentes entre dois vetores de características comparados, e mensura quão improvável é a distribuição.

 $<sup>^1\</sup>mathrm{Um}\ bin$ diz respeito à medida de freqüência de cada nível de intensidade mape<br/>ado pelo histograma

#### 3.3.2.3 Distância Quadrática

A distância Minkowski trata todos os *bins* de um histograma de características de maneira independente, não levando em consideração o fato de que certos pares de *bins* correspondem a características as quais são perceptualmente mais similares do que outros pares. Para solucionar tal problemática a Distância Quadrática é introduzida:

$$d_A(X,Y) = \sqrt{(x-y)^T A (x-y)}$$
(3.14)

onde  $A = [a_{ij}]$  é uma matriz  $N \times N$  de elementos  $a_{ij}$  que denotam a similaridade entre os bins i e j, e x e y dizem respeito a vetores de atributos (características). O valor de  $a_{ij}$  é dado por:

$$a_{ij} = 1 - \frac{d_{i,j}}{max [d_{i,j}]}$$
 onde  $d_{i,j} = |x_i - y_i|$ 

A Distância Quadrática tem sido utilizada em muitos sistemas de recuperação baseada em conteúdo para histogramas de cores. Foi mostrado que esta distância conduz a resultados perceptualmente mais desejáveis do que a distância Euclidiana e o método de Intersecção de Histogramas devido ao fato de considerar a similaridade entre os *bins* do histograma de cores [Long et al., 2003].

#### 3.3.2.4 Distância Canberra

A distância Canberra pode ser entendida e comparada com a distância Manhattan  $(L_1)$ , pois na sua computação a diferença absoluta dos valores das características é dividida pela soma absoluta dos mesmos. Isto a torna a distância mais restritiva entre as apresentadas, a qual é definida formalmente como:

$$d_C(X,Y) = \sum_{i=1}^{n} \frac{|x_i - y_i|}{|x_i| + |y_i|}$$
(3.15)

Esta distância é muito sensível a pequenas variações, e tem sito empregada em diferentes áreas como por exemplo para seqüências de DNA na bioinformática, assim como para detecção de intrusão [Emran & Ye, 2001].

### 3.3.3 Propostas de Funções de Distâncias

Em [Vasconcelos & Lippman, 2000] é traçada uma análise das funções matemáticas mais representativas utilizadas para similaridade entre imagens, correlacionando-as de acordo com a aplicação mais apropriada de cada uma.

Existem ainda propostas interessantes para histogramas de cores [Ling & Okada, 2006] [Traina et al., 2003] [Felipe & Traina, 2005], séries temporais multidimensionais [Lee et al., 2000] [Keogh & Ratanamahatana, 2005], além de associação com seleção de características [Santini & Jain, 1999] [Hiransakolwong et al., 2005]. Um trabalho interessante realizado por Santini [Santini & Jain, 1999] propõe uma métrica baseada na aplicação de lógica nebulosa para a análise dos atributos, na qual medidas nebulosas são utilizadas para tratar dependências entre os mesmos. Resultados obtidos com testes experimentais realizados em seres humanos comprovaram sua eficácia.

A família de distâncias AID leva em consideração os efeitos da interação entre os vetores de características extraídos das respectivas imagens, trabalhando com o conceito de interação fraca ou forte entre dois atributos, sendo assim propostas duas funções de distância a *Weak Attribute Interaction Distance* (WAID) e a *Strong Attribute Interaction Distance* (SAID) [Felipe & Traina, 2005]. Esta nova família de distâncias, tem como principal característica analisar como atributos afetam a percepção humana de similaridade quando seus valores variam em conjunto, permitindo dessa maneira com que o usuário estabeleça parâmetros e ajuste a influência entre atributos, aproximando-se da percepção humana.

Porém, algumas funções de distância utilizadas com o intuito de modelar a percepção humana em relação à similaridade, como é o caso da família AID, por muitas vezes contradizem em diferentes maneiras os axiomas métricos [Santini & Jain, 1999]. Em [Tversky, 1977] [Tversky & Gati, 1982] é apresentada uma análise da percepção humana na qual acredita-se que os axiomas métricos são extremamente restritivos no contexto de consulta por similaridade, propondo então um **modelo de contraste** com propriedades diferentes dos axiomas métricos.

Para que uma função de distância d() seja denominada de **função de distância métrica** esta deve satisfazer aos três axiomas: Simetria, não negatividade e Desigualdade triangular (ver Seção 2.2.3 do Capítulo 2).

Justamente, um dos axiomas mais criticados, diz respeito à propriedade da desigualdade triangular, sendo considerado o axioma vital na área de indexação métrica [Ashby & Perrin, 1988], pois este é de extrema importância para o processo de poda nas árvores métricas. Em determinadas situações, a desigualdade triangular não condiz com a percepção que se tem da inter-relação entre os objetos.

## 3.4 Considerações Finais

Este capítulo apresentou os conceitos nos quais o presente trabalho está fundamentado. Tais conceitos referem-se e estão envolvidos na análise do binômio função de distância e extratores de características, onde foi apresentada uma descrição detalhada de tal sustentáculo primordial para os sistemas de recuperação de imagens baseada em conteúdo.

Como dito anteriormente, a tarefa de recuperação de dados complexos está intimamente relacionada com a utilização de um extrator de características o qual provê as características inerentes dos dados, bem como de uma função de distância utilizada para quantificar a similaridade entre esses dados. Porém um grande desafio ainda permanece: como integrar esses dois aspectos chaves, visando a melhoria e aperfeiçoamento dos dados retornados para uma dada consulta por similaridade.

Justamente é nesse ponto focal em que o capítulo subseqüente fundamenta-se, aonde experimentos utilizando tanto os extratores de características descritos no presente capítulo, como também as funções de distância também aqui descritas são realizados utilizando bases de imagens médicas referentes a domínios específicos com o intuito de definir o melhor par função de distância e extrator de características com relação à precisão das consultas por similaridade. Portanto, auxiliando e contribuindo de maneira essencial para a melhoria do processo de recuperação de imagens médicas baseada em conteúdo.

Os conceitos apresentados no presente capítulo não só embasam o capítulo subseqüente, como também os conceitos aqui explicitados são utilizados no capítulo 5 e no capítulo 6.

Capítulo 4

# Buscando a Melhor Configuração do Binômio: Função de distância e Características

## 4.1 Introdução

o presente Capítulo são explicitados os resultados obtidos visando à obtenção da melhor integração entre o binômio função de distância e extratores de características, com o intuito de responder a consultas por similaridade. É também apresentada uma descrição detalhada dos experimentos desenvolvidos com as técnicas de extração de características baseadas em cor, textura e forma descritas no capítulo anterior (seção 3.2) em conjunto com as funções de distância estudadas, analisadas e implementadas, descritas nas seções 3.3.1 e 3.3.2 do mesmo capítulo. Especificamente, descrevem-se os conjuntos de dados utilizados nos experimentos, as abordagens adotadas na extração de características, além das conclusões obtidas a partir dos resultados.

O intuito do presente capítulo é abordar e demonstrar que uma escolha cuidadosa de determinadas funções de distância agregadas a determinados extratores de características, resulta em uma considerável melhoria na qualidade dos dados retornados para uma dada consulta por similaridade.

# 4.2 Abordagens para a Extração de Características

Na presente seção, são explicitadas as técnicas de extração de características de imagens médicas, mais especificamente com relação à geração dos respectivos vetores de características, abstraindo as informações relevantes inerentes às imagens. Para a extração de características foi adotada uma abordagem específica para cada extrator.

Considerando o extrator baseado em textura (descritores de Haralick) foi adotada uma abordagem fundamentada em 3 passos principais para a geração dos vetores de características, os quais são descritos a seguir:

- No primeiro passo as matrizes de co-ocorrência referentes a cada imagem são obtidas. As imagens têm seus níveis de cinza reduzidos de 256 para 16, permitindo dessa forma uma redução na ordem das matrizes. Para o cálculo das matrizes de co-ocorrência, 4 ângulos (θ) 0°, 45°, 90° e 135° foram considerados, bem como 5 distâncias (d) 1, 2, 3, 4 e 5. Portanto, para cada imagem são geradas 20 matrizes de co-ocorrência de ordem 16.
- No segundo passo são obtidos os valores dos descritores de Haralick escolhidos (definidos na Tabela 3.1 do capítulo 3). Para cada matriz de co-ocorrência, o valor de cada descritor é calculado, sendo que para cada imagem os valores resultantes da aplicação dos descritores sobre as matrizes de co-ocorrência são armazenados em uma matriz, denominada matriz de descritores, na qual as linhas representam cada uma das 4 angulações utilizadas e as colunas representam cada uma das 5 distâncias.
- No terceiro passo são gerados os vetores de características, tais vetores são compostos a partir das matrizes de descritores, calculando-se as médias dos valores obtidos com diferentes distâncias (d), para cada direção (θ).

Para o extrator baseado nos momentos de Zernike, foram utilizados polinômios de Zernike de ordem 0 até ordem 30, onde todos os momentos gerados para uma dada imagem foram organizados em um único vetor de características. Portanto, os vetores de características gerados possuem 256 posições.

Considerando o extrator de histogramas convencionais de níveis de cinza, os vetores de características gerados assim como os momentos de Zernike apresentam 256 posições, pois todas as imagens pertencentes às três bases utilizadas para experimentos, as quais serão descritas na seção subseqüente, possuem pixels de profundidade de 8 bits, resultando portanto em 256 tons de cinza o que leva a histogramas possuindo 256 *bins*.

Ao utilizar a extensão do método EM/MPM [Balan et al., 2005] apresentado na seção 3.2.5 do capítulo 3, primeiramente as imagens foram segmentadas utilizando a técnica citada, onde seis características para cada região segmentada são geradas: a massa m; centróide ( $x_o \, e \, y_o$ ); média de níveis de cinza ( $\mu$ ), dimensão fractal (D) e coeficiente linear (b). Nos experimentos, as imagens foram segmentadas em cinco regiões, gerando portanto um vetor de características composto de 30 elementos.

# 4.3 Conjuntos de Imagens

Visando a realização da análise do binômio função de distância e extrator de características, foram utilizadas três bases de imagens médicas distintas.

A primeira base de imagens utilizada contém 704 imagens que correspondem a exames de ressonância magnética (RM) e angiograma, obtidas de vários tipos de exames realizados no Hospital das Clínicas de Ribeirão Preto (HCRP). Os exames de ressonância magnética ressaltam melhor os tecidos moles (músculos, gordura, entre outros), enquanto que os angiogramas são exames radiológicos em que se utiliza um contraste que permite a visualização dos vasos sangüíneos de alguma região do corpo do paciente, neste caso, foram utilizados angiogramas de coração. A profundidade dos pixels da imagem é de 8 bits, resultando em 256 tons de cinza, e as dimensões das imagens são 256 x 256. A base de imagens foi dividida em 8 classes de acordo com a região do corpo examinada e o tipo de corte especificado. A Tabela 4.1 mostra a distribuição das imagens nas respectivas 8 classes.

Classe	Quantidade de Imagens
Angiograma	36
RM Axial Pélvis	86
RM Axial Cabeça	155
RM Axial Abdômen	51
RM Coronal Abdômen	23
RM Coronal Cabeça	36
RM Sagital Cabeça	258
RM Sagital Espinha	59
Total	704

Tabela 4.1: Distribuição das imagens no conjunto de teste.

Na Figura 4.1 podem ser observados exemplos de imagens pertencentes a cada classe da primeira base de imagens utilizada.

As duas bases de imagens subseqüentes utilizadas nos experimentos caracterizam-se por possuírem imagens de mamografias. Na Figura 4.2 é ilustrado um exemplo de duas imagens, provenientes de exames de mamografia, estas imagens consistem em regiões de interesse (*Regions of Interest - ROIs*) contendo nódulos tumorais, de uma base de 250 imagens de exames de mamografia obtidas em ambiente hospitalar. Todos os casos foram previamente analisados e classificados por médicos radiologistas e eventualmente confirmados por exames complementares. A profundidade dos pixels da imagem é de 8 bits, resultando em 256 tons de cinza, e as dimensões das imagens são variáveis. A base de imagens de regiões de interesse foi dividida em 2 classes, nódulos benignos e malignos, cada uma contendo respectivamente 104 e 146 imagens, totalizando 250 imagens.

A Figura 4.2 mostra um exemplo de imagens, consistindo de regiões de interesse provenientes de exames de mamografia, explicitando tanto os nódulos tumorais benignos,



Figura 4.1: Exemplos de imagens de cada classe pertencentes ao conjunto de teste da base de imagens de ressonância magnética.

Figura 4.2 (a), quanto malignos, Figura 4.2 (b).

Características de forma e textura das imagens podem ser usadas para separá-las em classes distintas pré-definidas. Essas classes, no caso em questão imagens médicas, podem ser utilizadas pelos radiologistas para auxiliar no diagnóstico. A avaliação de nódulos tumorais detectados em mamogramas é um exemplo de tal auxílio. Estas imagens podem ser classificadas pelo radiologista, em uma avaliação inicial, como benignas ou malignas, de acordo com a forma do contorno apresentada por lesões detectadas. Nódulos malignos, em geral, infiltram o tecido adjacente, resultando em um contorno irregular ou de difícil distinção, enquanto nódulos benignos apresentam um contorno bem definido e contínuo, como pode ser observado tanto na Figura 4.2 quanto na Figura 4.3.





Figura 4.2: Exemplos de imagens de cada classe pertencentes ao conjunto de teste da base de imagens mamográficas de regiões de interesse (*Regions of Interest - ROIs*) de nódulos tumorais: (a) nódulo benigno; (b) nódulo maligno.

A Figura 4.3 mostra um exemplo de imagens, da segunda base de imagens provenientes de exames de mamografia utilizada, as quais foram classificadas segundo o padrão de lesões de mama *BI-RADS* (*Breast-Imaging Reporting and Data System*) internacionalmente conhecido, desenvolvido em conjunto com vários comitês internacionais, liderados pelo *American College of Radiology* (ACR). O BI-RADS classifica as lesões da mama em seis categorias:

- Categoria 0: necessidade de avaliações adicionais de imagens;
- Categoria 1: tecido normal;

- Categoria 2: lesões benignas;
- Categoria 3: alta probabilidade de lesões benignas;
- Categoria 4: alguma probabilidade de malignidade, recomenda-se biópsia;
- Categoria 5: alta probabilidade de malignidade.



Figura 4.3: Exemplos de imagens de cada classe pertencentes ao conjunto de teste da base de imagens mamográficas classificadas segundo o padrão *BI-RADS*, contendo nódulos tumorais de mama típicos: (a) Classe 2; (b) Classe 3; (c) Classe 4; (d) Classe 5.

A profundidade dos pixels das imagens da base é de 8 bits, resultando em 256 tons de cinza, e as dimensões das imagens são 256 x 256. A base de imagens de regiões de interesse foi dividida em 4 classes, segundo a classificação *BI-RADS*. A Tabela 4.2 mostra a distribuição das imagens nas respectivas 4 classes, bem como a quantidade de imagens em cada uma destas, totalizando 103 imagens.

Classe	Quantidade de Imagens
BI-RADS 2	1
BI-RADS 3	20

45

37

103

BI-RADS 4

BI-RADS 5

Total

Tabela 4.2: Distribuição das imagens no conjunto de teste.

# 4.4 Experimentos

Na presente seção serão apresentados os resultados experimentais obtidos com relação às comparações realizadas com as funções de distância descritas nas seções 3.3.1 e 3.3.2 e os extratores de características descritos na seção 3.2. Estes experimentos foram realizados com o intuito de obter a melhor combinação entre as características extraídas e as funções de distância utilizadas para a realização de consultas por similaridade.

## 4.4.1 Resultados

As imagens pertencentes às bases utilizadas foram processadas, extraindo suas características gerando os respectivos vetores de características de cada imagem. Os vetores obtidos foram indexados utilizando o Método de Acesso Métrico (MAM) *Slim-tree* com intuito de acelerar o processamento das consultas por similaridade. Após a indexação dos vetores de características, consultas por similaridade aos k-vizinhos mais próximos, k-NN (knearest neighbor), foram aplicadas às "imagens"<sup>1</sup> variando aleatoriamente as imagens de consulta e os valores de k, para todos os conjuntos de características gerados considerando cada extrator utilizado, histograma de níveis de cinza, histogramas métricos, textura (Haralick) e momentos de Zernike. Além desses extratores foi utilizado também o método para a segmentação de imagens EM/MPM proposto em [Traina et al., 2004].

Uma consulta k-NN consiste em recuperar as k imagens mais similares a uma dada imagem de consulta, comparando seus vetores de características utilizando uma determinada função de distância para realizar tal cálculo de similaridade. A partir de cada base utilizada, todas as imagens pertencentes a cada uma destas foram utilizadas como centros de consulta. Para uma descrição detalhada sobre consultas k-NN ver seção 2.2.

Foram gerados gráficos de precisão e revocação, referentes à cada base de imagens, para os extratores de histograma de níveis, textura e momentos de Zernike, utilizando as funções de distância implementadas e observando qual destas adequa-se melhor a um extrator específico. Além dos gráficos já citados, foram gerados também os resultados referentes à precisão e revocação obtidas pelos histogramas métricos e comparados por meio das funções de distância DM e GWD explicitadas na seção 3.2.2 do capítulo 3.

E importante ressaltar que os pontos dos gráficos de curvas de precisão e revocação consistem de médias das consultas k-NN realizadas.

Nas Figuras 4.4 (a), (b) e (c) são ilustrados o comportamento das buscas realizadas para as três bases utilizadas para experimentos, já citadas, comparando o extrator de características de histogramas de níveis de cinza com as diversas funções de distância, juntamente com os histogramas métricos utilizando as funções de distância DM e GWD.

Considerando a base de imagens de ressonância magnética (Base RM), a partir da análise dos gráficos da Figura 4.4 (a), é possível constatar que a distância DM claramente apresenta o melhor ganho com relação à precisão, sendo que esta é em média quatro vezes mais rápida de se computar em relação à função de distância  $L_1$ . As funções de distância  $\chi^2$ , Quadrática, Divergência de Jeffrey e  $L_1$  apresentaram praticamente o mesmo comportamento para um nível de revocação de até 25% da base. A diferença na precisão obtida pelas funções DM e  $L_{\infty}$ , alcança valores de aproximadamente 75%, corroborando o fato da grande diferença de precisão obtida pela modificação da função de distância utilizada. A função de distância Canberra apresenta inicialmente níveis baixos de precisão, porém seu comportamento melhora para níveis de revocação acima de 35%.

Para a base de imagens de regiões de interesse (Base Mama ROIs), provenientes de

<sup>&</sup>lt;sup>1</sup>Nesse ponto não mais imagens propriamente ditas mas sim vetores de características.



Figura 4.4: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas para as três bases utilizadas para experimentos; comparando o extrator de características de **histogramas de níveis de cinza** com as funções de distância: (a) base de imagens de ressonância magnética; (b) base de imagens de regiões de interesse (*ROIs*) provenientes de exames de mamografia; (c) base de imagens provenientes de exames de mamografia classificadas segundo o padrão *BI-RADS*.

exames de mamografia, Figura 4.4 (b), as funções de distância DM e GWD apresentam praticamente o mesmo comportamento para um nível de revocação de até 25%, porém para níveis de revocação superiores a precisão da função de distância DM decai, chegando a igualar-se com a precisão obtida pela função  $L_{\infty}$  a qual apresenta a pior precisão. A função de distância Canberra apresenta a segunda melhor precisão, cerca de 65% para uma revocação de 25% da base e mantém esse padrão de precisão durante praticamente níveis de revocação de até 85%. Novamente as funções de distância  $L_1$ , Divergência de Jeffrey e Quadrática apresentaram comportamentos muito semelhantes. Vale ressaltar que até o momento as funções de distância pertencentes à família  $L_p$  amplamente utilizadas na literatura não apresentaram as melhores precisões.

Pode-se observar que para a base de imagens também provenientes de exames de mamografia, classificadas segundo o padrão *BI-RADS*, Figura 4.4 (c), as precisões de todas as funções de distância utilizadas foram praticamente as mesmas, apresentando uma variação considerável apenas no intervalo de 10% a 25% de revocação, onde pode ser observado que a função de distância Quadrática apresentou a melhor precisão, até 52% para um nível de revocação de 15%.

As Figuras 4.5 (a), (b) e (c) ilustram o comportamento das buscas realizadas para as três bases utilizadas para experimentos, já citadas, comparando o extrator de textura com as diversas funções de distância.

Observa-se que para a base de imagens de ressonância magnética, Figura 4.5 (a), a função de distância que apresentou melhor ganho de precisão foi a distância Canberra apresentando uma precisão de aproximadamente 90% para um nível de 20% de revocação da base, seguida logo após pela distância  $\chi^2$  a qual apresenta uma precisão de 81% considerando o mesmo nível de revocação. Pode-se observar que novamente as funções de distância Quadrática,  $L_1$  e Divergência Jeffrey apresentaram resultados muito próximos, quase idênticos. A função de distância Euclidiana ( $L_2$ ) e  $L_{\infty}$ , comumentemente utilizadas na literatura apresentaram os piores resultados com relação à precisão. A diferença de precisão alcança valores de 92% quando as funções Canberra e  $L_{\infty}$  são comparadas. Esta alta discrepância gera uma enorme diferença no conjunto de respostas retornado aos usuários. Vale ressaltar que as distâncias pertencentes à família  $L_p$  permaneceram abaixo das distâncias Canberra e  $\chi^2$ .

Para a base de imagens de regiões de interesse (ROI), provenientes de exames de mamografia, Figura 4.5 (b), as precisões de todas as funções de distância utilizadas apresentaram comportamentos semelhantes, mostrando uma variação entre o intervalo de revocação de 5% a 10%, sendo que acima de 40% de revocação todas as funções praticamente igualaramse. A função de distância Canberra foi a que apresentou melhor precisão, cerca de 72% para um nível de revocação de 15% da base. A função de distância  $L_{\infty}$  apresentou a pior precisão, cerca de 10% abaixo quando comparada com a precisão obtida pela distância Canberra.

Na Figura 4.5 (c) observa-se que para a base de imagens também provenientes de



Figura 4.5: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas para as três bases utilizadas para experimentos; comparando o extrator de características de **textura** com as funções de distância: (a) base de imagens de ressonância magnética; (b) base de imagens de regiões de interesse (ROIs) provenientes de exames de mamografia; (c) base de imagens provenientes de exames de mamografia classificadas segundo o padrão BI-RADS.

exames de mamografia, classificadas segundo o padrão *BI-RADS*, as precisões de todas as funções de distância utilizadas, novamente apresentaram-se muito próximas, variando muito pouco entre os níveis de revocação 5% e 40% e praticamente iguais daí em diante. Para os níveis de revocação que apresentaram uma pequena variação na precisão, a função de distância que apresentou melhor precisão foi a  $L_{\infty}$  cerca de 50% de precisão aproximadamente, seguida pela função  $\chi^2$ .

Nas Figuras 4.6 (a), (b) e (c) são ilustrados o comportamento das buscas realizadas para as três bases utilizadas para experimentos, já citadas, comparando o extrator de momentos de Zernike com as diversas funções de distância.

Analisando-se os gráficos da Figura 4.6 (*a*), observa-se que as funções de distância praticamente empatam, com exceção da função  $L_{\infty}$  a qual apresentou a pior precisão. Pode-se observar que a distância Canberra apresentou-se como uma boa escolha tanto para as características obtidas por meio dos momentos de Zernike quanto para as características baseadas em textura, sendo que apresenta também um baixo custo computacional.

Tanto para base de imagens de regiões de interesse, provenientes de exames de mamografia, Figura 4.6 (b), quanto para a base de imagens também provenientes de exames de mamografia, Figura 4.6 (c), classificadas segundo o padrão BI-RADS, as precisões de todas as funções de distância utilizadas mostraram-se praticamente as mesmas, apresentando uma ínfima variação, cerca de 5% para determinados níveis de revocação.

Além dos resultados obtidos por meio da utilização dos extratores de histograma de níveis de cinza, textura e momentos de Zernike, também foram realizados experimentos utilizando o método de segmentação EM/MPM proposto em [Traina et al., 2004].

A Figura 4.7 ilustra o comportamento das buscas realizadas para base de imagens de ressonância magnética, utilizando o método de segmentação EM/MPM com as diversas funções de distância e as comparando. O método não foi aplicado às outras bases, pois este tem por finalidade segmentar e diferenciar imagens de ressonância magnética para diferentes cortes e partes do corpo humano, e não a segmentação de nódulos tumorais de exames de mamografia.

Observa-se que a função de distância Canberra, apresentou a melhor precisão, até 97% para um nível de revocação de 60% da base. Novamente as funções de distância  $L_1$ , Quadrática e Divergência de Jeffrey apresentaram comportamentos praticamente idênticos, sendo que estas apresentaram-se como a segunda melhor opção logo após da distância Canberra. Vale ressaltar que as funções  $L_2$  e  $L_{\infty}$  amplamente empregadas na literatura apresentaram as piores precisões.

Considerando a base de imagens de ressonância magnética, a Figura 4.8 sumariza os resultados obtidos por meio da comparação entre as funções de distância e as características extraídas com relação a avaliação de precisão das consultas por similaridade.

Além dos resultados obtidos, o estudo realizado referente às funções de distância, possibilitou a criação da Tabela 4.3 a qual ilustra as principais características de algumas funções de distância, obtidas a partir do estudo da literatura. Nela são descritas caracte-



Figura 4.6: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas para as três bases utilizadas para experimentos; comparando o extrator de características de **momentos de Zernike** com as funções de distância: (a) base de imagens de ressonância magnética; (b) base de imagens de regiões de interesse (ROIs) provenientes de exames de mamografia; (c) base de imagens provenientes de exames de mamografia classificadas segundo o padrão BI-RADS.



Figura 4.7: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas para a base de imagens de ressonância magnética utilizada para experimento; comparando o extrator de características baseado em segmentação dos objetos da imagem (EM/MPM) com as diversas funções de distância.

	Funções de Distância						
I	Maior Precisão <del>&lt;</del> .					Menor Precisão	
	+			1		-	
Textura	Canberra	DM	X <sup>2</sup>	L <sub>1</sub> Quadrática Jeffrey	L <sub>2</sub>	L <sub>co</sub>	
Histograma	DM	X <sup>2</sup>	L <sub>1</sub> Quadrática Jeffrey	Canberra	L <sub>2</sub>	L∞	
Zernike	<i>X</i> <sup>2</sup>	Canberra	L <sub>2</sub>	L <sub>1</sub> Quadrática Jeffrey	L <sub>oo</sub>	DM	
EM/MPM	Canberra	L <sub>1</sub> Quadrática Jeffrey	X <sup>2</sup>	L <sub>2</sub>	L	DM	

Figura 4.8: Ilustração das comparações entre as funções e os extratores de características.

rísticas peculiares de cada uma, tais como:

- Domínio de Dados: refere-se ao domínio de dados ao qual a função de distância realiza os cálculos, se trata-se de um espaço de características dimensional ou adimensional;
- Aplicação: diz respeito à função de distância ser métrica ou não métrica, ou seja, atender às propriedades de simetria, não-negatividade e desigualdade triangular, propriedades estas de extrema importância para a indexação dos dados utilizando

os Métodos de Acesso Métricos, principalmente a propriedade de desigualdade triangular.

- Utilização: refere-se, como o próprio nome demonstra, à utilização das funções de distância para sistemas de busca/recuperação de dados, com relação à suas características específicas. Por exemplo, a função  $L_1$ , a qual tem maior utilização para o cálculo de similaridade entre séries temporais ou histogramas. Já a função  $L_2$  é amplamente utilizada para o cálculo de similaridade entre diversos tipos de dados, fato esse que leva ser classificada como de utilização geral.
- Característica dos atributos: diz respeito ao inter-relacionamento que a função de distância considera em relação aos atributos comparados para o cálculo de similaridade. Por exemplo, as funções de distância  $L_1$ ,  $L_2$  e  $L_{inf}$  não levam em consideração a interdependência entre atributos, ou seja, dessa forma cada atributo contribui de forma independente para o cálculo de distância. Sendo assim, o valor de um atributo não gera influência no valor de outros. Já funções como a Quadrática e a Mahalanobis levam em consideração respectivamente a similaridade entre os *bins* de um dado histograma e a correlação entre os atributos. As funções SAID e WAID propostas em [Felipe et al., 2006] empregam o conceito de interação fraca e forte entre dois atributos, com o objetivo de aproximar-se da percepção humana.

Função	Domínio de Dados Aplicação		Utilização	Característica dos Atributos	
$L_1$	dimensional	métrica	série temporal/ histogramas	independência	
$L_2$	dimensional	métrica	geral	independência	
$L_{\infty}$	dimensional	métrica	mais abrangente	independência	
Ponderada	dimensional	métrica	ênfase específica	relevância	
Quadrática	dimensional	métrica	histograma de cor	bins similares	
DKL	dimensional	não métrica	histograma/textura	assimétrica	
Jeffrey	dimensional	métrica	histograma	robusta a ruídos	
L <sub>edit</sub>	adimensional	métrica	textos curtos	alto custo	
SAID	dimensional	não métrica	mapear percepção	perceptual	
WAID	dimensional	métrica	mapear percepção	perceptual	

Tabela 4.3: Tabela Comparativa das Funções de Distância.

## 4.5 Considerações Finais

No presente capítulo foi demonstrada a importância da melhor escolha e adequação de uma dada função de distância, em conjunto com um tipo de característica. Para tal análise foram gerados gráficos de precisão e revocação, ilustrando o relacionamento entre o binômio explicitado. A partir de tais experimentos pôde-se notar claramente que tal relacionamento é de fato sólido e consistente, ao mesmo tempo em que é extremamente sensível. As funções de distância  $L_1$ , Quadrática e Divergência de Jeffrey apresentaram por muitas vezes comportamentos praticamente idênticos. Este fato pode ser explicado para as funções  $L_1$  e Quadrática, devido à distância Quadrática tratar-se de uma função de distância ponderada, onde esta ponderação é definida por uma matriz de similaridade  $A = [a_{ij}]$  (ver seção 3.3.2.3). Para os experimentos foi empregada a função de distância  $L_1$  para computar tal matriz de similaridade, devido ao seu baixo custo computacional. As funções  $L_1$  e Divergência de Jeffrey apresentam uma representação espacial similar, portanto este fato corrobora seus comportamentos semelhantes.

Considerando as características de forma e textura, a distância Canberra apresentou bons resultados com relação à precisão, apresentando resultados insatisfatórios apenas para as características oriundas dos histogramas de níveis de cinza. É importante ressaltar que as distâncias  $L_2$  e  $L_{\infty}$ , amplamente empregadas segundo a literatura, apresentaram os piores valores de precisão para praticamente todas as características.

A distância Canberra, não comumentemente e amplamente empregada na recuperação de dados multimídia, apresentou resultados promissores com relação à precisão das consultas por similaridade. Além disso, esta função apresenta custo computacional tão baixo quanto a largamente utilizada função de distância  $L_1$ .

Capítulo 5

# Melhorando a Recuperação de Imagens Médicas por meio da Ponderação de Características

# 5.1 Introdução

mineração de imagens tem sido um foco de pesquisa por parte de muitos pesquisadores na área de mineração de dados e recuperação de informação nos dias atuais. Um dos desafios da mineração de imagens é o de relacionar as características de baixo nível (extraídas automaticamente das imagens) com a semântica e subjetividade da percepção humana.

Com o crescimento exponencial de repositórios de imagens atualmente, as técnicas de recuperação de imagens baseada em conteúdo têm sido intensamente pesquisadas. Geralmente as técnicas CBIR utilizam características visuais intrínsecas das imagens (cor, textura e forma), produzindo vetores de características com centenas ou até mesmo milhares de características. Esta elevada cardinalidade das características representa um grande problema, pois leva à chamada maldição da alta dimensionalidade [Malcok et al., 2006] onde as estruturas de indexação degradam-se e o poder de representatividade de cada característica diminui consideravelmente, tornando os processos de indexação e recuperação das imagens muito mais lentos e insatisfatórios [Aggarwal, 2001]. Além disso, freqüentemente muitas características são correlacionadas entre si trazendo desse modo informação redundante, não adicionando nenhuma informação considerável sobre as imagens, deteriorando a habilidade do sistema em distingüí-las de maneira fidedigna. Com o intuito de evitar o problema da maldição da alta dimensionalidade, as técnicas de seleção de características podem ser empregadas visando à redução do tamanho dos vetores de características.

Outro problema considerado é o já citado *gap* semântico onde as características de baixo nível extraídas automaticamente das imagens não representam de maneira satisfatória a interpretação semântica destas.

Neste capítulo é introduzido um novo método que lida com a maldição da alta dimensionalidade e o problema do *gap* semântico, utilizando o algoritmo StARMiner (*Statistical Association Rule Miner*) proposto em [Ribeiro et al., 2005] juntamente com a técnica denominada **seleção de características contínuas** (*continuous feature selection*) a qual atribui pesos a cada característica inerente à uma dada imagem segundo seu grau de relevância.

O método aplica mineração de regras de associação estatísticas para integrar as características de baixo nível com o conhecimento do especialista sobre as imagens, com o intuito de reduzir o *gap* semântico existente entre a representação destas e sua respectiva interpretação pelo especialista humano. Estas regras de associação são empregadas para realizar a ponderação das características, de acordo com sua relevância. A redução da dimensionalidade é realizada pelo descarte das características irrelevantes (aquelas que apresentam peso nulo). Os pesos obtidos são utilizados para calcular a similaridade entre as imagens durante o processo de busca por similaridade.

# 5.2 Conceitos Relacionados

A tarefa de associação de características [Agrawal & Srikant, 1994] é uma das áreas de mineração de dados mais exploradas. Ela foi inicialmente motivada por aplicações comercias tais como, análise de mercado, análise de cesta de compras e classificação de clientes. Entretanto, a extração de regras de associação também tem sido intensamente usada em outras aplicações, tais como sumarização e classificação de imagens.

A mineração de imagens demanda a extração de características destas segundo um critério específico. Após extraídas, essas características são submetidas ao processo de mineração. O foco deste trabalho são imagens médicas, mais especificamente sobre os vetores de características utilizados para comparar e recuperar as imagens quanto a sua similaridade. O intuito é ponderar as características de acordo com a contribuição de cada uma para a interpretação semântica, peneirando as características que contém as informações mais relevantes das imagens, tornando dessa forma o processo de recuperação das mesmas mais eficiente. Um ponto a ser ressaltado é que essa metodologia proposta pode ser diretamente estendida para outros tipos de dados complexos além de imagens, pois consultas por similaridade são empregadas sobre diversos tipos de dados complexos. Como já pudemos observar anteriormente, a similaridade entre imagens é computada por meio da utilização de funções de distância aplicadas aos vetores de características.

Com relação aos processos de redução da dimensionalidade, os quais tratam de outro fator de extrema importância, estes podem ser baseados em técnicas supervisionadas
ou não supervisionadas. Enquanto que as técnicas supervisionadas necessitam de um conjunto de treinamento pré-classificado, as técnicas não supervisionadas não necessitam de um conjunto de treinamento. O processo de redução da dimensionalidade também é classificado quanto a métodos de seleção de características (*feature selection*) ou transformação de características (*feature trasnformation*). O ponto chave de diferença entre essas duas técnicas é que a primeira consiste em selecionar um subconjunto das características originalmente extraídas que melhor representem a imagem, já a segunda técnica consiste em gerar a partir das características originalmente extraídas um novo conjunto de características que melhor representem uma dada imagem. Portanto, o processo de seleção de características não realiza nenhum tipo de transformação sobre as características originais, ele apenas remove aquelas que apresentam certa redundância ou irrelevância, preservando a semântica do conjunto de características original. Em [Cheng et al., 2006] foi demonstrado que a seleção automática de características melhora a classificação das imagens médicas.

As técnicas de seleção de características podem ainda serem divididas em binárias e contínuas. A seleção de características contínuas (*continuous feature selection*) como dito anteriormente atribui pesos a cada característica em particular, enquanto que a binária atribui pesos binários a cada característica.

Nas seções subseqüentes, serão explicitadas e introduzidas as técnicas utilizadas para realização da proposta do novo método, bem como o método em si.

## 5.3 Algoritmo StaRMiner - Mineração de Regras de Associação Estatísticas

A essência do método utilizado é o conceito de mineração de regras de associação. Tradicionalmente, o problema da mineração de regras de associação consiste em encontrar relacionamentos da forma  $A \rightarrow B$ , onde  $A \in B$  são um conjunto de itens, indicando que  $A \in B$  ocorrem freqüentemente juntos em uma transação de banco de dados. Portanto, se A ocorrer existe uma grande probabilidade de B também ocorrer. Estes tipos de regras obtém bons resultados quando se lida com dados categóricos (nominais). Entretanto, quando lidamos com imagens, as quais são representadas por características de baixo nível, que por sua vez são organizadas em vetores de características que as descrevem quantitativamente, ou seja, consistem de dados contínuos, faz-se necessário a utilização de um tipo de regra de associação que considere valores contínuos. Uma classe de regras de associação contínuas que vem sendo pesquisada recentemente são as **regras de associação estatísticas**, as quais tratam-se de regras geradas por meio da utilização de testes estatísticos.

O método proposto emprega o algoritmo denominado StARMiner proposto em [Ribeiro et al., 2005] o qual faz uso de tais regras de associação baseadas na utilização de testes estatísticos. O objetivo do algoritmo StARMiner é implementar a mineração de regras de associação estatísticas para encontrar as características que melhor discriminam a imagem em categorias, dado um conjunto de treinamento. As regras mineradas são utilizadas pelo método para ponderar as características de acordo com a relevância de cada uma, alcançando dessa forma uma nova e eficiente representação das imagens.

Seja uma base de imagens médicas T;  $x_j$  uma categoria de imagens;  $T_{x_j} \in T$  o subconjunto de imagens pertencentes a categoria  $x_j$ ; e  $f_i$  a *i*-ésima característica do vetor de características F. Sejam  $\mu_{f_i}(L)$  (equação 5.1) e  $\sigma_{f_i}(L)$  (equação 5.2), respectivamente, a média e o desvio padrão dos valores da característica  $f_i$  no subconjunto de imagens L e  $(f_i)_m$  o valor da característica  $f_i$  da imagem m.

$$\mu_{f_i}(L) = \frac{\sum_{m \in L} (f_i)_m}{|L|}$$
(5.1)

$$\sigma_{f_i}(L) = \sqrt{\frac{\sum_{m \in L} (f_i)_m - \mu_{f_i}(L)^2}{|L|}}$$
(5.2)

O algoritmo utiliza três restrições definidas pelo usuário, as quais são:

- Δµ<sub>min</sub> um parâmetro de entrada que indica a diferença mínima permitida entre a média dos valores de f<sub>i</sub> para as imagens da categoria x<sub>j</sub> e a média dos valores de f<sub>i</sub> para as imagens restantes;
- $\Delta \sigma_{max}$  um parâmetro de entrada que indica o máximo de desvio padrão permitido aos valores de  $f_i$  para as imagens da categoria  $x_j$ ;
- $\gamma_{min}$  um parâmetro de entrada que indica a confiança mínima com a qual uma hipótese H0 deve ser rejeitada.

O objetivo do algoritmo StARMiner é implementar a mineração de regras de associação estatísticas para encontrar as características que melhor discriminam a imagem em categorias. O formato geral da regra é  $x_j \to f_i$ , significando que para a categoria  $x_j$ a característica  $f_i$  é relevante. As regras mineradas pelo algoritmo devem satisfazer as seguintes condições:

- A característica  $f_i$  deve apresentar um comportamento distinto para as imagens da categoria  $x_j$  em relação às demais imagens;
- A característica  $f_i$  deve apresentar um comportamento uniforme para todas as imagens da categoria  $x_j$ .

Essas condições são implementadas no algoritmo StARMiner por meio da incorporação das restrições de interesse já discutidas. Portanto, para que o algoritmo StARMiner minere as regras na forma  $x_j \rightarrow f_i$ , é necessário satisfazer as condições dadas pelas Equações 5.3, 5.4 e 5.5.

$$\mu_{f_i}(T_{x_j}) - \mu_{f_i}(T - T_{x_j}) \ge \Delta \mu_{min} \tag{5.3}$$



Figura 5.1: Regiões de rejeição [Ribeiro et al., 2005].

$$\sigma_{f_i}(T_{x_j}) \le \Delta \sigma_{max} \tag{5.4}$$

$$H_0: \mu_{f_i}(T_{x_j}) = \mu_{f_i}(T - T_{x_j}) \quad e \quad H_1: \mu_{f_i}(T_{x_j}) \neq \mu_{f_i}(T - T_{x_j})$$
(5.5)

Na equação 5.5, H0 deve ser rejeitada com confiança maior ou igual a  $\gamma_{min}$  em favor da hipótese H1. Rejeitando H0, com confiança  $\gamma_{min}$ , implica que as médias  $\mu_{f_i}(T_{x_j})$  e  $\mu_{f_i}(T - T_{x_j})$  são estatisticamente diferentes.

Para rejeitar H0, o valor Z (teste estatístico de hipóteses), cujo cálculo é apresentado na Equação 5.6, deve estar na região de rejeição explicitada pela Figura 5.1. Os valores críticos z1 e z2 dependem de  $\gamma_{min}$ , como pode ser observado na Tabela 5.1.

$$Z = \frac{\mu_{f_i}(T_{x_j}) - \mu_{f_i}(T - T_{x_j})}{\frac{\sigma_{f_i}(T_{x_j})}{\sqrt{|T_{x_j}|}}}$$
(5.6)

Tabela 5.1: Valores Críticos de Z [Ribeiro et al., 2005].

$\gamma_{min}$	0.9	0.95	0.99
z1	1.64	1.96	2.58
z2	-1.64	-1.96	-2.58

Uma regra  $x_j \to f_i$  é retornada pelo algoritmo, relacionando a característica  $f_i$  com a classe  $x_j$ , onde valores de  $f_i$  apresentam um comportamento peculiar em imagens da classe  $x_j$ . Esta propriedade indica que  $f_i$  é uma característica importante para distinguir imagens pertencentes a classe  $x_j$  das imagens restantes.

Portanto, o intuito do algoritmo StARMiner é identificar as características que manifestam um alto poder de discriminação, pois estas apresentam um comportamento particular e uniforme em imagens de uma dada categoria. Tal fato é de extrema importância pois as características que apresentam um comportamento uniforme para todas as imagens da base, independentemente da categoria, não contribuem para discriminá-las e devem ser eliminadas. Portanto, as regras geradas pelo StARMiner são muito proveitosas para revelar a relevância das características das imagens. As regras obtidas são empregadas para realizar a técnica denominada *continuous feature selection* segundo o método proposto descrito na seção subseqüente.

## 5.4 Método Proposto

Uma questão que surge, diz respeito a como utilizar as regras de associação para ponderar as características das imagens. Supondo que as imagens estejam classificadas em m classes  $X = x_1, x_2, \dots, x_m$ . Para cada característica  $f_i$  o algoritmo StARMiner tenta achar regras da forma  $x_j \to f_i$ , ou seja, o algoritmo visa a relacionar cada característica  $f_i$  com uma classe  $x_i$ . Caso a regra  $x_i \to f_i$  seja encontrada, isso significa que a característica  $f_i$  discrimina de maneira satisfatória as imagens da classe  $x_i$ . Conseqüentemente, as características  $f_i$ que melhor discriminam as imagens são aquelas que geram regras da forma  $x_i \to f_i$  para todo  $x_j \in X$ , denotando que essas características discriminam de maneira satisfatória todas as imagens. De maneira análoga, as características irrelevantes, ou seja, aquelas que não diferenciam bem as imagens, são aquelas para as quais nenhuma regra foi gerada, o que significa que elas apresentam um comportamento uniforme entre todas as classes, tornando-se dessa maneira não úteis para diferenciar as imagens. Portanto, para ponderar a característica  $f_i$ , o método proposto utiliza a quantidade de regras mineradas onde  $f_i$ ocorre. Considerando uma constante q que pode assumir valores 0 ou 1, as equações 5.7 e 5.8 definem a ponderação atribuída a cada característica  $f_i$ , segundo o número de regras geradas e de acordo com o valor atribuído a q, respectivamente 0 ou 1.

$$w_i = \alpha \times r_i \tag{5.7}$$

$$w_i = \alpha \times r_i + q \tag{5.8}$$

Nas equações 5.7 e 5.8,  $\alpha$  é um fator de ponderação que pode ser especificado pelo usuário;  $r_i$  é o número de regras mineradas onde a característica  $f_i$  aparece. Quando se deseja remover as características para as quais não foram geradas nenhuma regra de associação, ou seja, aquelas com menor poder de distinguir as imagens, a constante qassume o valor 0 e a equação 5.7 é utilizada para ponderação das características. Por outro lado, quando se deseja manter as características que não geraram nenhuma regra, a constante q assume o valor 1 e se utiliza a equação 5.8 para ponderar as características. Portanto, quando q = 0 é realizada a redução da dimensionalidade do vetor de características, juntamente com a ponderação das características segundo a relevância de cada uma e quando q = 1, todas as características são ponderadas também de acordo com suas respectivas relevâncias. Por meio de inúmeros experimentos e análises dos resultados foi possível determinar o melhor valor atribuído ao fator de ponderação  $\alpha$ , ou seja, aquele para o qual a precisão das consultas alcançou os melhores níveis, sendo que no caso em questão foi utilizado um fator de ponderação igual a 10. Dessa maneira as equações 5.7 e 5.8 apresentaram respectivamente as seguintes formulações:

$$w_i = 10 \times r_i \tag{5.9}$$

$$w_i = 10 \times r_i + q \tag{5.10}$$

Substituindo a variável de ponderação  $w_i$  dada pela equação 5.10 na equação da função de distância Minkowski Ponderada (equação 3.10) descrita na seção 3.3.1.1 do capítulo 3, é obtida a fórmula da função de distância utilizada no método proposto, quando se deseja ponderar todas as características (q = 1), mesmo aquelas para as quais não foram geradas regras de associação, ou seja, aquelas consideradas irrelevantes:

$$d_{L_p}(F,G) = \sqrt[p]{\sum_{i=1}^n (10 \times r_i + q) (f_i - g_i)^p}$$
(5.11)

O método proposto é um processo supervisionado que lida os com dois principais empecilhos para os sistemas CBIR, o *gap* semântico e a alta dimensionalidade dos vetores de características. O método transpõe esses empecilhos por meio da utilização das regras de associação, relacionando o conhecimento do especialista e as características de baixo nível das imagens, ponderando essas características de acordo com suas relevâncias. Os passos realizados pelo método são ilustrados na Figura 5.2.



Figura 5.2: Diagrama ilustrativo das fases do método proposto: fase de treinamento e fase de teste.

O método é dividido em duas fases principais: a fase de treinamento e a fase de testes. A fase de treinamento é composta por três passos: (1) seleção de características; (2) mineração das regras de associação; (3) emprego da técnica *continuous feature selection*. A fase de teste emprega os pesos encontrados na fase de treinamento para realizar as consultas por similaridade.

#### 5.5 Experimentos

Esta seção apresenta os resultados experimentais da utilização do método proposto, baseado na técnica *continuous feature selection*, com o intuito de melhorar o processo de recuperação de imagens similares segundo uma dada imagem de consulta em uma dada base de imagens, por meio da ponderação das características extraídas das imagens.

As imagens pertencentes à base de imagens de exames de ressonância magnética e angiogramas descrita na seção 4.3 do capítulo 4 foram processadas, extraindo suas características gerando os respectivos vetores de características de cada imagem, os vetores obtidos foram indexados utilizando o Método de Acesso Métrico (MAM) *Slim-tree* (ver seção 2.5) com o intuito de acelerar o processo de recuperação das mesmas. Após a indexação dos vetores de características, consultas por similaridade aos k-vizinhos mais próximos, k-NN (k-nearest neighbor) foram aplicadas às "imagens" variando aleatoriamente as imagens de consulta e os valores de k, para todos os conjuntos de características gerados considerando cada extrator utilizado.

Para a realização dos experimentos foram empregados dois extratores de características: (a) um extrator baseado em textura e (b) um extrator baseado em forma. O extrator baseado em textura utilizado faz uso dos descritores de Haralick (ver seção 3.2.3 do capítulo 3) obtidos a partir das matrizes de co-ocorrência oriundas do conjunto de imagens utilizado. Os descritores de Haralick utilizados para os experimentos foram variância, entropia, energia, homogeneidade, momento de  $3^a$  ordem, variância inversa e *step*. Todos esses descritores foram combinados gerando um vetor de características composto de 140 elementos.

O extrator baseado em forma utilizado foi a variação do método EM/MPM [Balan et al., 2005] apresentado na seção 3.2.5 do capítulo 3. As imagens foram primeiro segmentadas utilizando a técnica citada, onde seis características para cada região segmentada são geradas: a massa m; centróide ( $x_o \in y_o$ ); média de níveis de cinza ( $\mu$ ), dimensão fractal (D) e coeficiente linear (b). Nos experimentos as imagens foram segmentadas em cinco regiões, gerando portanto um vetor de características composto de 30 elementos.

#### 5.5.1 Resultados

Para a realização dos experimentos o conjunto de imagens foi dividido em dois subconjuntos: um conjunto de treinamento composto de 176 imagens (25% da base de imagens) e o conjunto de teste composto de 528 imagens (75% da base de imagens).

Os gráficos de precisão e revocação das Figuras 5.3 e 5.4 correspondem, respectivamente aos experimentos realizados utilizando o conjunto de imagens representado por meio do extrator de características baseado em textura (descritores de Haralick) e do extrator de características baseado em forma (variação do método EM/MPM). Nas Figuras 5.3 e 5.4 os gráficos (a), (b) e (c) correspondem aos resultados obtidos quando utilizadas as funções de distância  $L_1$ ,  $L_2$  e  $L_{\infty}$  respectivamente, comparando o novo método de ponderação das características proposto em relação aos resultados obtidos pela não utilização de tal ponderação.

As curvas de precisão e revocação presentes nos gráficos das Figuras 5.3 e 5.4 foram obtidas por meio da execução de consultas por similaridade empregando: (1) características não ponderadas; (2) características selecionadas pelo algoritmo StARMiner; (3) o método proposto, utilizando q = 0 (removendo características irrelevantes) e (4) o método proposto, utilizando q = 1 (ponderando as características com relação a relevância de cada uma). No método proposto, a utilização da constante q = 0 resulta na redução da dimensionalidade do vetor de características, removendo as características consideradas irrelevantes para o processo de cálculo de similaridade. Analisando o conjunto de imagens representadas pelo extrator baseado em textura, ou seja, pelas características oriundas de tal extração, a utilização de q = 0 resulta em uma redução de 20% no tamanho do vetor de características já que 112 características foram selecionadas como relevantes, sendo assim 28 dimensões foram eliminadas. Considerando o conjunto de imagens representado pelas características oriundas do extrator baseado em forma, a utilização de q = 0 resulta em uma redução de 27% na dimensionalidade do vetor de características, pois do total de 30 características, 22 foram selecionadas como relevantes, ou seja, 8 dimensões foram eliminadas.

A Figura 5.3 ilustra os gráficos de precisão e revocação obtidos por meio da utilização das características de textura. Analisando o gráfico da Figura 5.3 (a) pode-se observar que o método proposto claramente apresenta um ganho com relação a precisão das consultas por similaridade. A melhor precisão foi obtida por meio da utilização da constante q = 1, ou seja, considerando todas as características extraídas e ponderando-as.

As características ponderadas apresentaram um considerável ganho com relação a precisão, cerca de 30% acima da precisão obtida pelas características não ponderadas, considerando um nível de revocação de 40%. Utilizando q = 0, o método atinge um ganho de 20% na precisão para o mesmo nível de revocação citado. É importante ressaltar que utilizando q = 0 além do ganho de precisão, é realizada também uma redução da dimensionalidade de 20% no tamanho do vetor de características, fato que colabora para a redução do custo de processamento e memória utilizados.

A partir da análise dos gráficos da Figura 5.3 (b) é possível notar que o método proposto novamente apresenta um ganho evidente com relação à precisão. Pode-se observar que a redução da dimensionalidade realizada utilizando o algoritmo StARMiner (ver Figura 5.3 (b) curva  $L_2$  StARMiner) causou uma diminuição na precisão. Entretanto, quando o método proposto é aplicado aos mesmos vetores de características observa-se um aumento da precisão, alcançando um ganho de até 20% utilizando q = 0 para um nível de revocação de 35%, bem como um ganho de até 38% quando q = 1. Estes resultados convalidam que o método proposto melhora a precisão das consultas por similaridade, até mesmo quando realiza a redução da dimensionalidade dos vetores de características.



Figura 5.3: Gráficos de precisão e revocação utilizando as funções de distância (a)  $L_1$ , (b)  $L_2$  e (c)  $L_{\infty}$  obtidos sobre as características de textura extraídas do conjunto de imagens de ressonância magnética, empregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando q = 0 (removendo as características irrelevantes) e o método proposto, utilizando q = 1.



Figura 5.4: Gráficos de precisão e revocação utilizando as funções de distância (a)  $L_1$ , (b)  $L_2$  e (c)  $L_{\infty}$  obtidos sobre as características de forma extraídas do conjunto de imagens de ressonância magnética, empregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando q = 0 (removendo as características irrelevantes) e o método proposto, utilizando q = 1.

Analisando os gráficos da Figura 5.3 (c) pode-se observar que o algoritmo StARMiner gerou a pior precisão mesmo em comparação com as características originais obtidas por meio do extrator baseado em textura. Utilizando o método proposto com q = 0, ou seja, ponderando as características selecionadas pelo StARMiner, a precisão aumenta, chegando a igualar-se com a precisão obtida pelas características originais. Vale ressaltar que apesar de igualarem-se, o método proposto, como já citado, realiza também uma redução da dimensionalidade dos vetores de características, tornando o processamento de comparação entre os mesmos mais rápido devido à diminuição do tamanho do vetor de características. Aplicando o método proposto com q = 1 foi obtida um considerável ganho na precisão, até 67% para um nível de revocação de 30% em comparação com o algoritmo StARMiner, bem como um ganho de cerca de 20% em comparação com a precisão obtida pelo método utilizando q = 0 e a obtida pelas características originais.

Como já citado, além dos testes realizados utilizando as características baseadas em textura, também foram realizados testes com as características obtidas por meio do extrator de forma (ver Figura 5.4).

A Figura 5.4 ilustra os gráficos de precisão e revocação obtidos por meio das características baseadas em forma para representar o conjunto de imagens. Analisando a Figura 5.4 (a), pode-se observar que o método proposto, para ambos os valores de q = 0 e q = 1, gerou um ganho na precisão de até 5% em comparação com a precisão obtida pelas características originais, bem como para as características selecionadas pelo StARMiner (sem ponderação). Observando-se a Figura 5.4 (b), novamente nota-se que o método apresentou um ganho na precisão de aproximadamente 14% para um nível de revocação de 45% da base de imagens, novamente em comparação com as características originalmente obtidas e àquelas selecionadas pelo algoritmo StARMiner. O mesmo ocorre quando observamos a Figura 5.4 (c), a qual ilustra um ganho de até 16% na precisão para um nível de revocação de 35% quando utilizado q = 0 na qual a dimensionalidade do vetor de características apresentar 30 posições, porém apenas 22 dessas características foram consideradas relevantes para o processo de comparação entre as imagens.

A Figura 5.5 ilustra um exemplo do processamento de uma consulta k-NN (k=8), utilizando a imagem situada no canto esquerdo superior como centro de consulta e a função de distância  $L_2$ . Na Figura 5.5(a) pode-se observar o resultado obtido utilizando as características originais baseadas em forma. Já a Figura 5.5 (b) ilustra o resultado utilizando o método proposto com q = 0 (características reduzidas). As imagens ressaltadas pela borda vermelha são referentes a falsos positivos<sup>1</sup>. Pode-se observar claramente uma significante melhora nos resultados obtidos quando o método proposto é aplicado, com a diminuição de falsos positivos.

Considerando os resultados obtidos, podemos argumentar que o método proposto para ponderação das características apresenta um considerável ganho com relação a precisão

 $<sup>^1\</sup>mathrm{Um}$ falso positivo é uma imagem retornada a qual é pertencente a uma classe que difere da classe da imagem de consulta.



Figura 5.5: Exemplo da execução de uma consulta k-NN (k=8), utilizando a imagem situada no canto esquerdo superior como centro de consulta e função de distância  $L_2$ ; (a) utilizando as características originais; (b) utilizando o método proposto com q = 0. As imagens envolvidas pela moldura vermelha referem-se a falsos positivos.

de consultas por similaridade em imagens médicas, além de convalidar a utilização da técnica *continuous feature selection*. Outro ponto importante a ser ressaltado refere-se à possibilidade do método ser empregado para outras funções de distância. Experimentos foram realizados nessa direção e os resultados obtidos são discutidos nos parágrafos subseqüentes.

As funções de distância para as quais o método proposto foi também empregado foram: a distância distância  $\chi^2$ , distância Canberra e a distância Divergência de Jeffrey, explicitadas na seção 3.3.2 do capítulo 3.

Os gráficos de precisão e revocação das Figuras 5.6 e 5.7 correspondem, respectivamente, aos experimentos realizados sobre o conjunto de imagens representadas por meio do extrator baseado em textura (Haralik) e do extrator baseado em forma (EM/MPM) já citados. Nas Figuras 5.6 e 5.7 os gráficos (a), (b) e (c) são referentes aos resultados obtidos por meio dos novos experimentos, utilizando respectivamente as funções de distância  $\chi^2$ , Canberra e Divergência de Jeffrey, comparando a técnica de ponderação proposta em relação às convencionais.

Para a geração das curvas de precisão e revocação ilustradas nos gráficos das Figuras 5.6 e 5.7 foi adotada a mesma metodologia utilizada para a comparação dos resultados anteriormente explicitados, ou seja, foram realizadas consultas por similaridade utilizando, as características não ponderadas, as características selecionadas pelo algoritmo StARMiner, bem como o método proposto, utilizando valores de q = 0 (removendo características irrelevantes) e q = 1 (ponderando as características de acordo com a relevância de cada uma).

Na Figura 5.6 são ilustrados os gráficos de precisão e revocação obtidos por meio da utilização das características de textura. Analisando o gráfico da Figura 5.6 (a) é possível observar que o método proposto apresenta um ganho considerável com relação a precisão das consultas por similaridade. A melhor precisão foi obtida por meio da utilização da constante q = 1, ou seja, considerando todas as características extraídas e ponderando-as. Tais características apresentaram um ganho de até 25% em comparação com a precisão obtida pelas características originalmente extraídas, considerando um nível de revocação

de 30%. Já, utilizando o método proposto com q = 0, além de reduzir a dimensionalidade dos vetores de características em 20% a precisão ainda apresentou um ganho de até 12% em relação às características originais para o mesmo nível de revocação de 30% da base.



Figura 5.6: Gráficos de precisão e revocação utilizando as funções de distância (a)  $\chi^2$ , (b) Canberra e (c) Divergência de Jefrrey obtidos sobre as características de textura extraídas do conjunto de imagens de ressonância magnética, empregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando q = 0 (removendo as características irrelevantes) e o método proposto, utilizando q = 1.

Um fato que pode ser observado é que a precisão alcançada pelo algoritmo StARMiner (ver Figura 5.6 (a) curva  $\chi^2$  StARMiner) apresentou a precisão mais baixa, até mesmo em relação à precisão obtida por meio da utilização das características originalmente extraídas. Porém quando o método proposto é aplicado, ou seja, quando as características são ponderadas um ganho na precisão é atingido tanto para q = 0 quanto para q = 1.

Analisando os gráficos da Figura 5.6 (b) nota-se que o método proposto utilizando tanto q = 0, como q = 1 praticamente igualaram-se, apresentando novamente um ganho de precisão em relação às características originais e àquelas selecionadas pelo algoritmo StARMiner, aproximadamente 3% de ganho na precisão considerando uma revocação de 45% da base. Vale ressaltar que mesmo com o ganho de precisão baixo obtido, o método proposto utilizando q = 0 concomitantemente reduziu a dimensão dos vetores de características em 20% e ainda assim houve ganho na precisão das consultas por similaridade.

Na Figura 5.6 (c), a partir da análise das curvas de precisão e revocação obtidas, podese concluir que novamente o método proposto utilizando q = 1, bem como para q = 0apresenta um considerável ganho de precisão nas consultas por similaridade, respectivamente cerca de 32% e 14% em comparação com as características originais, considerando um nível de revocação de 45% da base. As características selecionadas pelo algoritmo StARMiner, em relação a precisão obtida pelas características originais, também apresentaram um ganho de precisão de até 9%.

Dessa forma tais resultados corroboram ainda mais a eficácia do método proposto para a melhoria do processo de recuperação de imagens médicas por conteúdo, uma vez que pôde também ser estendido para outras funções de distância além da família Minkowski, apresentando ganho com relação à precisão das consultas por similaridade.

Como anteriormente explicitado, além dos experimentos realizados utilizando as características de textura extraídas das imagens da base, utilizando as funções de distância  $\chi^2$ , Canberra e Divergência de Jeffrey, experimentos utilizando as características de forma também foram realizados. A Figura 5.7 ilustra os gráficos de precisão e revocação obtidos por meio de tais experimentos.

Observando-se as curvas de precisão e revocação presentes na Figura 5.7 (a), (b) e (c) novamente obteve-se respectivamente um ganho na precisão de até 12%, 7% e 6% em relação tanto às características originais, como para àquelas selecionadas pelo algoritmo StARMiner, considerando um nível de revocação de 45%. Pode-se constatar que as precisões do método proposto utilizando tanto q = 0, como q = 1 apresentaram-se muito próximas, quase igualando-se como ocorreu nos gráficos da Figura 5.4. Uma possível explicação para tal fato seria devido à homogeneidade entre as precisões obtidas também pelas características originalmente extraídas e aquelas selecionadas pelo algoritmo StARMiner.



Figura 5.7: Gráficos de precisão e revocação utilizando as funções de distância (a)  $\chi^2$ , (b) Canberra e (c) Divergência de Jefrrey obtidos sobre as características de forma extraídas do conjunto de imagens de ressonância magnética, empregando: características não ponderadas; características selecionadas pelo StARMiner; o método proposto, utilizando q = 0 (removendo as características irrelevantes) e o método proposto, utilizando q = 1.

## 5.6 Considerações Finais

No presente capítulo foi apresentado um novo método supervisionado que emprega a técnica *continuous feature selection*. O método proposto emprega regras de associação estatísticas para reduzir o *gap* semântico, um problema intrínseco a sistemas CBIR, como também realiza a redução da dimensionalidade, minimizando a o problema da "maldição da alta dimensionalidade".

Os experimentos realizados demonstraram que o método apresenta um ganho de até 38% na precisão de consultas por similaridade, ou seja, sempre subjugando a precisão obtida pelas características originalmente extraídas, ao mesmo tempo em que diminui o custo de processamento envolvido. Portanto, os resultados obtidos convalidam a eficácia do método proposto no aperfeiçoamento da precisão de consultas por similaridade, utilizando regras de associação estatísticas juntamente com a técnica *continuous feature selection* para ponderação de características extraídas a partir de imagens médicas.

Trabalhos futuros poderão ser direcionados na verificação de diferentes estratégias de ponderação, assim como analisando quais funções de distância melhor se adequam com as características ponderadas.

Capítulo 6

# Um Novo Método para Redução da Dimensionalidade de Histogramas

## 6.1 Introdução

evido à contínua expansão da demanda por armazenamento de dados complexos, no caso em questão imagens, os algoritmos de extração de características, com intuito de realizar consultas mais precisas, fazem uso cada vez mais de um número elevado de características. A busca por características que possam representar de modo fiel imagens, por exemplo, leva muitas vezes à coleta de um elevado número de descritores, acarretando o aumento da complexidade do processo de consulta e recuperação de imagens e a dificuldade em separá-las, o qual é denominado pela comunidade científica de *maldição da alta dimensionalidade* (*dimensionality curse*) [Berchtold et al., 1998]. Para resolver tal problema, além do desenvolvimento de métodos de acesso para bases de dados contendo alta dimensionalidade, outros métodos para a realização de redução da dimensionalidade também com o intuito de acelerar a performance do processo de consulta têm sido desenvolvidos [Chávez et al., 2001].

O conceito de redução da dimensionalidade é baseado na seleção de um número reduzido de atributos suficientemente relevantes para determinar as características de determinado conjunto de dados visando sua representação. A indexação de tais dados por sua vez é também realizada neste espaço reduzido [Jolliffe, 1986] [Faloutsos & Lin, 1995] [Pagel et al., 2000]. Geralmente o conjunto resposta retornado poderá conter alguns falsos positivos. Existem na literatura inúmeros métodos para a realização da redução de dimensionalidade, por exemplo, a utilização da dimensão intrínseca (fractal) do conjunto de dados para a realização de tal redução é um deles. A dimensão fractal é utilizada como uma forma de mensuração, a qual caracteriza a irregularidade dos objetos naturais e tem sido utilizada em [Souza & Traina Jr., 2006] [Arantes et al., 2003] [Traina et al., 2000b] como uma aproximação da dimensão intrínseca ou seja, a dimensão de um dado conjunto de dados a qual diz respeito à dimensão espacial dos objetos representados por este, visando a redução de atributos necessários para representar e analisar objetos complexos sem perda de informação, otimizando assim o processamento dos mesmos. Além da dimensão fractal um outro método muito utilizado é denominado PCA (*Principal Component Analysis*) [Jolliffe, 1986] o qual transforma pontos de um espaço original de alta dimensão em outro espaço, geralmente de menor dimensão.

De modo geral o processo de redução da dimensionalidade pode ser aplicado ao conjunto de dados antes deste ser indexado, com intuito de reduzir os efeitos da maldição da alta dimensionalidade na estrutura de indexação. A redução da dimensionalidade naturalmente gera perdas, conseqüentemente a precisão das respostas é afetada. Tal perda de informação depende especificamente do método utilizado e do conjunto de dados em foco. Em [Korn et al., 2001] são apresentados alguns dos métodos discutidos sobre redução da dimensionalidade.

No presente capítulo é explicitado um novo método que combate a maldição da alta dimensionalidade por meio da redução da dimensão dos vetores de características, ao mesmo tempo que mantém a capacidade de retornar de maneira satisfatória as imagens mais similares para uma dada consulta. O método é empregado tanto a histogramas tradicionais de níveis de cinza obtidos a partir dos valores dos pixels, como também a vetores de características obtidos por meio do extrator de características baseado nos momentos de Zernike (Seção 3.2.4), pois ambos geram um número elevado de características (vetores de características com 256 dimensões). Essa nova abordagem proposta utiliza a entropia de Shannon [Shannon, 2001] para realizar tal redução.

## 6.2 Método Proposto

A elevada utilização de histogramas em sistemas CBIR (*Content-Based Image Retrieval*) justifica-se devido ao fato de algoritmos que lidam com outras características intrínsecas, tais como textura e forma, apresentarem freqüentemente um alto custo computacional, além de serem dependentes do domínio de aplicação. Já os histogramas são obtidos com um custo linear em relação ao número de pixels da imagem.

Outros dois fatores que contribuem para a considerável utilização de histogramas em sistemas CBIR, dizem respeito a pequenas alterações de movimentação na imagem geralmente não afetarem histogramas e também por ser observado que freqüentemente objetos distintos apresentam histogramas diferentes.

Os histogramas de níveis de cinza obtidos a partir de imagens médicas geralmente geram vetores de características de alta dimensionalidade, pois tais imagens são freqüentemente obtidas e representadas por no mínimo 8 bits por pixel, resultando portanto, em um histograma de 256 bins. Em [Beyer et al., 1999] foi demonstrado que devido a alta dimensionalidade dos vetores de características obtidos a partir da utilização de histogramas de níveis de cinza convencionais, que o processo de indexação e recuperação de imagens torna-se relativamente custoso, pois os vetores de características são indexados ao invés das imagens propriamente ditas, gerando além desse aumento no custo, também à degradação da precisão das consultas CBIR, pois os vetores de características são utilizados no processo de busca por similaridade.

Com o intuito de lidar com esta restrição, foi proposto o presente método, o qual utiliza a entropia de Shannon sobre histogramas de níveis cinza das imagens. A entropia de Shannon acumulada é calculada para um intervalo fixo de *t bins*, que foi denominado como **fator de aglutinação**. Para um histograma convencional de níveis de cinza, obtido a partir de uma imagem específica. O processo de aglutinar os *bins* do histograma convencional utilizando um fator de aglutinação gerou o que foi denominado *buckets*.

A entropia de Shannon acumulada pode ser definida como:

$$S = -\sum_{i=0}^{256/t} \left( \sum_{j=i\times t}^{(i\times t)+t} p_j ln(p_j) \right)$$
(6.1)

onde t é o fator de aglutinação e  $p_j$  é o valor do bin j referente ao histograma convencional de níveis de cinza.

Considerando um histograma contendo 8 bits de níveis de cinza, ou seja 256 *bins*, e um fator de aglutinação igual a 16, o número de características obtidas será 16, portanto foram gerados 16 *buckets*, comprovando dessa maneira, uma considerável redução da dimensionalidade (256 dimensões para 16). Modificando o fator de aglutinação t para 32, 64, e 128 será obtido respectivamente 8, 4, e 2 características.

Conseqüentemente a cardinalidade do vetor de características reduzido, obtido a partir da aplicação do método proposto, pode ser definida como:

$$Size_{vector} = \frac{n_{bins}}{t} \tag{6.2}$$

onde,  $n_{bins}$  é igual ao número de *bins* do histograma de níveis de cinza;  $0 < t \le n_{bins}$  é o fator de aglutinação o qual será sempre potência de 2.

Concomitantemente ao fato da considerável redução da dimensionalidade, o novo método proposto, ao invés de gerar uma perda de precisão, apresenta um ganho importante desta quando comparado com a precisão do histograma convencional de níveis de cinza, como apresentado na seção de experimentos.

Na Figura 6.1 é ilustrado o método proposto e suas etapas. Processando a imagem de entrada, a etapa 1 obtém o respectivo histograma convencional de níveis de cinza. Na etapa 2 são computados os *buckets* utilizando o fator de aglutinação especificado e então, calcula-se a entropia de Shannon para cada *bucket*. Na etapa 3 é realizada a redução da dimensionalidade onde são gerados os vetores de características reduzidos, que serão



Figura 6.1: Método proposto para redução da dimensionalidade de histogramas.

utilizados na etapa 4 no processo de consulta por similaridade. Portanto, cada posição do vetor de características reduzido contém a entropia de Shannon acumulada calculada para cada *bucket* do histograma. Portanto o tamanho do vetor de características é igual ao número de *buckets*, além disso, outro fator importante a ser ressaltado diz respeito ao **fator de aglutinação**. Este fator fornece o número exato de quantas vezes a dimensionalidade do vetor de características foi reduzida, por exemplo, um fator de aglutinação t = 32, significa uma redução da dimensionalidade em 32 vezes.

### 6.3 Experimentos

Esta seção apresenta os resultados experimentais da utilização do método proposto, baseado na entropia de Shannon, para recuperar imagens similares a uma dada imagem de consulta, presente em uma dada base de imagens.

Foram realizados diversos testes com o novo método proposto com algumas bases de imagens. Os resultados obtidos a partir deste método foram comparados com os obtidos pelo histogramas convencionais de níveis de cinza (seção 3.2.1), histogramas métricos (seção 3.2.2) e com os vetores de características obtidos a partir dos momentos de Zernike (seção 3.2.4).

Para os experimentos, o histograma convencional de níveis de cinza, os vetores de características obtidos por meio dos momentos de Zernike, bem como o novo método

baseado na entropia de Shannon, foram comparados utilizando-se a função de distância Euclidiana  $(L_2)$ . Já os histogramas métricos foram comparados utilizando-se a distância Métrica (DM) [Traina et al., 2002a] e a distância Global Warp (GWD) [Felipe et al., 2005] (seção 3.2.2) para os histogramas métricos, sendo que esses se mostraram como os mais efetivos para esse tipo de histograma reduzido.

As imagens pertencentes às bases utilizadas foram processadas, extraindo-se suas características e gerando os respectivos vetores de características de cada imagem Os vetores obtidos foram indexados utilizando o Método de Acesso Métrico (MAM) *Slim-tree* (ver capítulo 2). Após a indexação dos vetores de características, para cada conjunto de imagens, consultas por similaridade aos **k vizinhos mais próximos**, *k*-NN (*k-nearest neighbor*) foram aplicadas às "imagens" do conjunto, variando aleatoriamente as imagens de consulta e os valores de **k**, para todos os conjuntos de características gerados pelo histograma convencional de níveis de cinza, os momentos de Zernike e para o novo método proposto. Todas as imagens do conjunto foram utilizadas como centro de consulta.

Para a avaliação do método proposto foram gerados gráficos de precisão e revocação (ver capítulo 2), obtidos a partir de consultas por similaridade realizadas sobre cada cada base de imagens, para o histograma convencional de níveis, os vetores obtidos a partir dos momentos de Zernike, bem como para o novo método proposto, utilizando variados fatores de aglutinação. Os pontos dos gráficos de curvas de precisão e revocação consistem de médias das consultas k-NN realizadas.

#### 6.3.1 Conjuntos de Imagens

Além das duas bases de imagens provenientes de exames de mamografia explicitadas na seção 4.3 do capítulo 4, uma outra base contendo 210 imagens de ressonância magnética e angiogramas foi também utilizada. A profundidade dos pixels de cada imagem é de 8 bits, resultando em 256 tons de cinza, e as dimensões das imagens utilizadas é de 256 x 256. A base de imagens foi dividida em 7 classes de acordo com a região do corpo examinada e o tipo de corte especificado. As regiões do corpo, bem como o corte tomado para aquisição das imagens foram: Angiogramas, corte Axial da Pélvis, corte Axial da Cabeça, corte Coronal do Abdômen, corte Coronal da Cabeça, corte Sagital da Cabeça, corte Sagital da Espinha.

#### 6.3.2 Resultados

Na presente seção pode-se analisar respectivamente as curvas de precisão e revocação geradas pela recuperação das imagens a partir da consulta aos k-vizinhos mais próximos, para a base de imagens de ressonância magnética, a base de imagens provenientes de exames de mamografia com regiões de interesse (*Regions of Interest - ROIs*) contendo nódulos tumorais, e para base de imagens classificadas segundo o padrão *BI-RADS*, também obtidas a partir de exames de mamografia. Foram extraídas características utilizando o histograma convencional de níveis de cinza, momentos de Zernike, histogramas métricos e o novo método proposto.

Para todas as bases de imagens o novo método foi aplicado utilizando os seguintes fatores de aglutinação:

- $4 \rightarrow$  gerando 64 buckets;
- $8 \rightarrow$  gerando 32 buckets;
- $16 \rightarrow$  gerando 16 buckets;
- $32 \rightarrow$  gerando 8 buckets;
- $64 \rightarrow \text{gerando } 4 \text{ buckets};$
- $128 \rightarrow \text{gerando } 2 \text{ buckets.}$

Devido ao fato do número de *buckets* ser igual ao número de posições no vetor de características reduzido, obtido a partir da utilização do método proposto, corrobora a considerável redução da dimensionalidade realizada pelo novo método proposto.

As Figuras 6.2 e 6.3 illustram os gráficos de curvas de precisão e revocação, em intervalos de 5 níveis de revocação. Esses gráficos mostram o comportamento do método considerando a média de consultas k-NN realizadas sobre a base de imagens de ressonância magnética, utilizando cada imagem pertencente à base como centro de consulta, onde k equivale ao número de imagens contidas na base.

Portanto, para a base de imagens de ressonância magnética, foram realizadas 210 consultas para cada imagem utilizada como centro de consulta e a média dos valores de precisão e revocação obtidos foi calculada e utilizada para gerar os gráficos das Figuras 6.2 e 6.3.

A Figura 6.2 ilustra os resultados obtidos a partir da utilização dos histogramas convencionais comparados por meio da função de distância Euclidiana, histogramas métricos comparados por meio da função de distância Métrica (DM) e Global Warp (GWD) e do método proposto para fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$  comparados também por meio da função Euclidiana.

Analisando os gráficos da Figura 6.2, pode ser observado que o novo método proposto utilizando os fatores de aglutinação 4, 8, 16, 32, 64 e 128 já citados anteriormente e respectivamente com vetores de características possuindo 64, 32, 16, 8, 4 e 2 posições, levaram a um aumento considerável da precisão em comparação com a obtida pelo histograma convencional de níveis de cinza, até mesmo com um fator de aglutinação de 128, o qual gera 2 *buckets*, ou seja, uma redução de 128 vezes da dimensionalidade o método proposto apresenta um ganho ainda considerável de precisão. Os fatores de aglutinação 8, 16, 32 e 64 apresentaram os melhores ganhos de precisão, cerca de 39% para um nível de revocação de 40% e para níveis acima de 50% os fatores de aglutinação 8 e 16 apresentaram melhores



Figura 6.2: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens de ressonância magnética, utilizando as características extraídas por meio do histograma convencional de níveis de cinza, histogramas métricos, e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.

precisões em relação aos fatores 32 e 64, uma variação de cerca de 10% na diferença das precisões entre eles.

E interessante ressaltar que a função de distância DM para valores de revocação abaixo de 50% apresenta precisão maior que o fator de aglutinação t = 128 e para valores de revocação acima de 50% (que são menos significativos na prática) sua precisão decai. Já a função de distância GWD, mantém praticamente a todo instante uma precisão maior que o fator t = 128, sendo que para valores de revocação entre 45% e 95% praticamente empata com o fator de aglutinação t = 4.

O método proposto para uma revocação de 40% obtém um ganho de precisão de até 40% em comparação com os histogramas de níveis de cinza e 30% utilizando fatores de aglutinação t = 16, também em comparação com os histogramas métricos obtidos, utilizando para estes, respectivamente a função de distância DM e a função de distância GWD.

A Figura 6.3 ilustra os gráficos de precisão e revocação obtidos a partir das características referentes aos momentos de Zernike sobre a base de imagens de ressonância magnética. Nesse caso, pode ser observado que a técnica de redução proposta claramente supera a precisão dos momentos de Zernike apenas quando é utilizado um fator de aglutinação t = 4. Para esse caso o método apresentou um ganho de 9% na precisão em comparação com os momentos de Zernike, considerando uma revocação de 20% e ao mesmo tempo mantendo uma considerável redução da dimensionalidade de 75%. Portanto, o fator de aglutinação t = 4 seria uma boa escolha para esse caso em questão. O método proposto



Figura 6.3: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens de ressonância magnética, utilizando as características extraídas por meio dos momentos de Zernike e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.

praticamente empata com a precisão obtida pelos momentos de Zernike quando se utiliza um fator de aglutinação t = 8. Entretanto, pode-se notar que utilizando-se um fator de aglutinação t = 8, o método reduz em até 8 vezes (32 *buckets*) a dimensionalidade do vetor original de Zernike que continha 256 posições, ao mesmo tempo que mantém a precisão quase igual à obtida pelos momentos de Zernike.

Nas Figuras 6.4 e 6.5 são ilustrados os gráficos de curvas de precisão e revocação, em intervalos de 5 níveis de revocação, explicitando o comportamento de 250 consultas k-NN realizadas sobre a base de imagens mamográficas de regiões de interesse (*Regions* of Interest - ROIs) de nódulos tumorais. A mesma metodologia para realização dos experimentos explicitada anteriormente para a base de ressonância magnética foi utilizada.

Analisando os gráficos da Figura 6.4 observa-se que mais uma vez o novo método proposto utilizando os fatores de aglutinação 4, 8, 16, 32, 64 e 128 respectivamente com vetores de características possuindo 64, 32, 16, 8, 4 e 2 posições levou a um aumento da precisão, no caso em questão de até aproximadamente 11% para uma revocação de 20% em comparação com a precisão obtida pelo histograma convencional de níveis de cinza. Novamente, o método proposto obtém um ganho de precisão também em relação aos histogramas métricos para as duas funções de distância utilizadas, de até aproximadamente 14% considerando uma revocação de 40%.

O fator de aglutinação t = 4, o qual reduz a dimensionalidade em até 4 vezes apresentou o melhor ganho de precisão cerca de até 11%.



Figura 6.4: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas de regiões de interesse (*Regions of Interest - ROIs*) de nódulos tumorais, utilizando as características extraídas por meio do histograma convencional de níveis de cinza, histogramas métricos, e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.

A Figura 6.5 ilustra os gráficos de precisão e revocação obtidos a partir das características referentes aos momentos de Zernike sobre a base de imagens de imagens mamográficas de regiões de interesse de nódulos tumorais. Pode ser observado que novamente para as características extraídas a partir dos momentos de Zernike, o método proposto com um fator de aglutinação t = 4 apresentou o melhor ganho com relação a precisão, cerca de 13% para uma revocação de 20%. Logo em seguida, novamente aparece o fator de aglutinação t = 8 como segunda melhor precisão, muito próxima do fator de aglutinação t = 4.

As Figuras 6.6 e 6.7 ilustram os gráficos de curvas de precisão e revocação, explicitando o comportamento de inúmeras consultas k-NN realizadas sobre a base de imagens mamográficas classificadas segundo o padrão BI-RADS. Para a realização dos experimentos foi adotada a mesma metodologia dos experimentos anteriores.

A Figura 6.6 ilustra os resultados obtidos a partir da utilização dos histogramas convencionais comparados por meio da função de distância Euclidiana, histogramas métricos comparados por meio da função de distância Métrica (DM) e Global Warp (GWD) e do método proposto para fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$  comparados também por meio da função Euclidiana.

A partir da análise dos gráficos da Figura 6.6, pode ser observado que o método proposto proporciona um ganho na precisão de até 30% quando comparada com a precisão obtida pelo histograma convencional de níveis de cinza. Pode-se observar, segundo os gráficos da Figura 6.6, que o método proposto utilizando um fator de aglutinação t = 4



Figura 6.5: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas de regiões de interesse (*Regions of Interest - ROIs*) de nódulos tumorais, utilizando as características extraídas por meio dos momentos de Zernike e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.

apresentou o melhor ganho de precisão, cerca de 30% considerando uma revocação de 15%. Logo em seguida o fator de aglutinação t = 8 apresentou um ganho de aproximadamente 13% quanto à precisão, também com relação ao histograma convencional de níveis de cinza. O método proposto novamente apresentou um ganho na precisão (até 17%) quando comparado aos histogramas métricos utilizando as funções DM e GWD, as quais são as melhores distâncias disponíveis para este tipo de histograma.

A Figura 6.7 ilustra os resultados da aplicação de consultas aos k-vizinhos mais próximos realizadas sobre as características obtidas a partir dos momentos de Zernike a base de imagens mamográficas classificadas segundo o padrão *BI-RADS*, bem como do método proposto para diferentes fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ . Ambos os métodos foram comparados por meio da função de distância Euclidiana.

Analisando os gráficos da Figura 6.7, pode-se concluir que todos os fatores de aglutinação apresentaram um ganho considerável na precisão em relação à precisão obtida pelos características baseadas nos momentos de Zernike. Por exemplo, o método alcançou aproximadamente um ganho de 71% na precisão, considerando uma revocação de 20% e utilizando o fator de aglutinação t = 4 (i.e. 64 *buckets*). Os gráficos da Figura 6.7 demonstram claramente que nosso método em todos os casos melhora a precisão das consultas por similaridade, utilizando qualquer fator de aglutinação, quando comparado com a precisão obtida pelos momentos de Zernike. É importante notar que quando o método é aplicado aos momentos de Zernike, os melhores resultados são obtidos com o fator de aglutinação t = 4.



Figura 6.6: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas classificadas segundo o padrão *BI-RADS*, utilizando as características extraídas por meio do histograma convencional de níveis de cinza, histogramas métricos, e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.



Figura 6.7: Gráfico de curvas de precisão e revocação, ilustrando e comparando o comportamento das buscas realizadas sobre a base de imagens mamográficas classificadas segundo o padrão *BI-RADS*, utilizando as características extraídas por meio dos momentos de Zernike e o novo método proposto para vários fatores de aglutinação  $t \in \{4, 8, 16, 32, 64, 128\}$ , utilizando a função de distância Euclidiana.

Considerando os resultados obtidos e suas respectivas análises, podemos verificar que o método proposto não apenas reduz consideravelmente a dimensionalidade dos vetores de características, bem como provê um ganho significante na precisão das consultas por similaridade e sendo passível de aplicação a vetores de características que apresentam dimensões elevadas. A desvantagem do método diz respeito à necessidade da base de imagens estar classificada previamente à execução do método, sendo que conforme as imagens estejam distribuídas na base a precisão pode sofrer alterações. Outro fator de suma importância, diz respeito ao baixo custo computacional do método, o qual apresenta complexidade linear com relação ao tamanho do vetor de características.

## 6.4 Considerações Finais

A redução de dimensionalidade obtida pelo método proposto, propicia um importante ganho de precisão. Mesmo com a utilização de variados fatores de aglutinação, a precisão obtida apresentou-se sempre maior, até 39%, que a do histograma convencional de níveis de cinza, contendo 256 dimensões.

Esses resultados evidenciam que a maldição da alta dimensionalidade realmente prejudica os resultados das consultas, e também confirmam que o método proposto é efetivo em aumentar a precisão das consultas nos sistemas de busca por conteúdo.

Portanto, o método proposto além de aumentar consideravelmente a precisão da busca por conteúdo em imagens médicas, também contribui para solucionar um problema crítico de sistemas de busca por conteúdo, a elevada dimensão dos vetores de características obtidos a partir da utilização de histogramas de níveis cinza.

Além disso, o método pode ser aplicado a outros vetores de características que apresentam dimensão elevada. Os experimentos realizados com os vetores de características obtidos a partir dos momentos de Zernike, apresentaram resultados muito satisfatórios. Os experimentos demonstraram que não só ocorreu uma significante redução da dimensionalidade original (256 dimensões) dos vetores de características, mas também um ganho considerável com relação a precisão, cerca de 100% em comparação com a precisão obtida pelos momentos de Zernike. Dessa forma, possibilita que novos estudos e experimentos sejam realizados, utilizando o método proposto para ser aplicado a outros extratores de características, como por exemplo, extratores baseados em Wavelets ou até mesmo de textura, os quais também geram um número elevado de características.

Outro ponto a ser considerado, observando a considerável redução da dimensionalidade alcançada pelo método e o ganho na precisão, diz respeito à utilização dessas características reduzidas em conjunto com outros extratores, aderindo mais informação a outros extratores de características.

# Capítulo 7 Conclusões

## 7.1 Considerações Finais

proposta desse trabalho de Mestrado foi contribuir para a área de recuperação por conteúdo de imagens médicas, onde o ponto focal da presente dissertação esteve relacionado com a análise do inter-relacionamento entre funções de distância e vetores de características objetivando a melhor integração entre esses dois componentes primordiais para um sistema de recuperação de imagens baseada em conteúdo. Objetivou também alcançar melhores resultados quanto a consultas por similaridade. Nesse sentido, o principal resultado da pesquisa foi com relação a análise do binômio função de distância e vetores de características, apresentada no Capítulo 4, o qual gerou uma publicação intitulada "Assessing the Best Integration between Distancefunction and Image-Feature to Answer Similarity Queries" no  $23^{rd}$  ACM Symposium of Applied Computing - SAC 2008 (Qualis A internacional).

Além do tópico central da dissertação, o trabalho também derivou outras importantes ramificações gerando duas novas propostas de métodos para auxiliar o processo de recuperação de imagens baseada em conteúdo. Uma delas fundamentada na utilização de regras de associação estatísticas juntamente com a ponderação de características com intuito de melhorar a recuperação de imagens médicas por meio desta ponderação, apresentada no Capítulo 5, a qual gerou uma publicação no 21<sup>th</sup> IEEE International Symposium on Computer-Based Medical Systems - CBMS 2008 (Qualis A internacional) intitulada "Content-based Retrieval of Medical Images by Continuous Feature Selection". A segunda proposta, explicitada no Capítulo 6, refere-se a um novo método para a redução da dimensionalidade de histogramas, a qual gerou uma publicação intitulada "A New Method to Efficiently Reduce Histogram Dimensionality" apresentada e publicada nos anais do Symposium on Medical Imaging - SPIE 2008, Computer-Aided Diagnosis track. Diante de tais fatos, foram alcançados resultados promissores e de extrema relevância para a área cuja pesquisa está inserida.

## 7.2 Principais Contribuições

Resumidamente, as principais contribuições do presente trabalho para a área de CBIR são descritas a seguir:

- Análise do inter-relacionamento entre as características extraídas dos dados e a função de distância utilizada em sua comparação, visando a obtenção da melhor integração entre esse binômio, com o intuito de responder a consultas por similaridade. Os experimentos realizados determinaram para um determinado tipo de base de imagens qual função de distância juntamente com um dado extrator de características gerou melhor resultado com relação à precisão das consultas por similaridade e demonstrando dessa maneira que as funções da família  $L_p$  (descritas no capítulo 3) as quais são amplamente utilizadas não geram resultados satisfatórios. Além, de demonstrar a existência de um forte inter-relacionamento entre as características extraídas dos dados e a funções de distância utilizada em sua comparação e também evidenciando que as funções de distância não devem ser utilizadas sem um prévio conhecimento das características extraídas e do contexto ao qual estão inseridas. Sem tal estudo pode-se não alcançar a melhor discriminação das imagens.
- Um novo método para redução da dimensionalidade de histogramas tradicionais de níveis de cinza, o qual pode ser estendido para outros vetores de características de alta dimensionalidade. A qualidade dos resultados obtidos foi analisada segundo gráficos de precisão e revocação considerando operações de consultas por similaridade. O método proposto apresentou um considerável ganho na precisão em relação aos histogramas tradicionais de níveis de cinza e às características extraídas por meio dos momentos de Zernike.
- Uma nova técnica para ponderação de características com intuito de melhorar a recuperação de imagens médicas. Por meio dos experimentos realizados foi possível verificar que a técnica proposta apresenta um ganho considerável com relação a precisão das consultas por similaridade, bem como pode ser estendida para outras funções de distância encontradas na literatura.
- Extensão e adequação da estrutura *Slim-tree* incluindo as funções de distância implementadas utilizadas para os experimentos, ampliando dessa maneira o leque de opções com relação às métricas de similaridade as quais poderão ser utilizadas para futuros estudos e experimentos dentro do Grupo de Base de Dados e Imagens do ICMC-USP.

A partir da implementação do novo conjunto de funções de distância e dos respectivos experimentos realizados para definir a melhor integração entre o binômio função de distância e extrator de características, foi possível ampliar a gama de funções de distância a serem incluídas no sistema de recuperação de imagens baseado em conteúdo e auxílio a diagnóstico, denominado cbPACS (*Content-Based Picture Archiving and Communication System*) atualmente em desenvolvimento colaborativo entre o Grupo de Bases de Dados e Imagens do ICMC (GBdI) - USP e o Centro de Ciências de Imagens e Física Médica (CCIFM) do Hospital das Clínicas de Ribeirão Preto (HCRP) - USP, bem como uma melhor calibragem entre características extraídas e as métricas de comparação, auxiliando na recuperação de imagens médicas baseada em conteúdo.

## 7.3 Propostas para Trabalhos Futuros

A seguir estão apresentadas algumas sugestões de tópicos que podem ser abordados em estudos futuros.

- Avaliação de uma gama ainda maior de extratores de características e funções de distância, analisando a precisão obtida pelo binômio com relação à busca por similaridade em imagens médicas, bem como análises com usuários especialistas no domínio de dados em questão.
- Estudo e proposta de novas técnicas de ponderação das características extraídas das imagens para o método proposto no Capítulo 5;
- Suporte ao desenvolvimento e aprimoramento de algoritmos de mineração de imagens, os quais dependem fortemente da qualidade das características de imagens obtidas por extratores que consigam representar fielmente as propriedades mais relevantes das imagens;
- Um estudo aprofundado de outras formulações além da entropia de Shannon para o método de redução da dimensionalidade de histogramas de níveis de cinza, proposto no Capítulo 6;
- Proposta e desenvolvimento de novas funções de distância específicas que permitam a comparação entre imagens, mais especificamente imagens médicas, mantendo a semântica natural dessa comparação e preocupando-se com o custo computacional, para que não seja alto, já que essas distâncias serão utilizadas em estruturas de indexação de dados.

## **Referências Bibliográficas**

- [Aggarwal, 2001] Aggarwal, C. C. (2001). On the effects of dimensionality reduction on high dimensional similarity search. In ACM Symposium on Principles of Database Systems (PODS), pp. 256–266, Santa Barbara, CA. ACM Press.
- [Agrawal & Srikant, 1994] Agrawal, R. e Srikant, R. (1994). Fast algorithms for mining association rules. In *International Conference on Very Large Databases (VLDB)*, pp. 487–499, Santiago de Chile, Chile.
- [Alferez et al., 2005] Alferez, R., Wang, Y.-F., e Jiao, L. (2005). An affine-invariant tool for retrieving images from homogeneous databases. *Multimedia Tools and Applications Journal (MTAJ)*, 25(1):133–159.
- [Arantes et al., 2003] Arantes, A. S., Vieira, M. R., Traina, A. J. M., e Traina Jr., C. (2003). The fractal dimension making similarity queries more efficient. In Second Workshop on Fractals and Self-similarity in Data Mining: Issues and Approaches (in conjunction with 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining), pp. 12–17, Washington, DC. ACM Press.
- [Ashby & Perrin, 1988] Ashby, F. G. e Perrin, M. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review*, 95(1):124–150.
- [Aslandogan & Yu, 1999] Aslandogan, Y. A. e Yu, C. T. (1999). Techniques and systems for image and video retrieval. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):56–63.
- [Baeza-Yates & Ribeiro-Neto, 1999] Baeza-Yates, R. A. e Ribeiro-Neto, B. (1999). Modern Information Retrieval. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- [Balan, 2007] Balan, A. G. R. (2007). Métodos adaptativos de segmentação aplicados à recuperação de imagens por conteúdo. Tese de doutorado, Universidade de São Paulo.
- [Balan et al., 2005] Balan, A. G. R., Traina, A. J. M., e Marques, P. M. d. A. (2005). Fractal analysis of image textures for indexing and retrieval by content. In 18th IEEE International Symposium on Computer-Based Medical Systems, pp. 56–63, Dublin, Ireland.

- [Bayer & McCreight, 1972] Bayer, R. e McCreight, E. M. (1972). Organization and maintenance of large ordered indices. In ACM SIGFIDET Workshop on Data Description and Access, pp. 107–141, Houston, Texas. ACM Press.
- [Beckmann et al., 1990] Beckmann, N., Kriegel, H.-P., Schneider, R., e Seeger, B. (1990). The r\*-tree: an efficient and robust access method for points and rectangles. In ACM SIGMOD International Conference on Management of Data, pp. 322–331, New York, NY, USA. ACM Press.
- [Berchtold et al., 1998] Berchtold, S., Bohm, C., e Kriegel, H.-P. (1998). The pyramidtree: Breaking the curse of dimensionality. In ACM SIGMOD International Conference on Management of Data, pp. 142–153, Seattle, WA.
- [Beyer et al., 1999] Beyer, K., Godstein, J., Ramakrishnan, R., e Shaft, U. (1999). When is "nearest neighbor" meaningful? In Beeri, C. e Buneman, P., editors, International Conference on Database Theory (ICDT), v. 1540 of Lecture Notes in Computer Science, pp. 217–235, Jerusalem, Israel. Springer Verlag.
- [Bozkaya & Özsoyoglu, 1999] Bozkaya, T. e Özsoyoglu, Z. M. (1999). Indexing large metric spaces for similarity search queries. ACM Transactions on Database Systems (TODS), 24(3):361–404.
- [Bueno, 2002] Bueno, J. M. (2002). Suporte à Recuperação de Imagens Médicas baseada em Conteúdo através de Histogramas Métricos. Tese de doutorado, Universidade de São Paulo.
- [Burkhard & Keller, 1973] Burkhard, W. A. e Keller, R. M. (1973). Some approaches to best-match file searching. *Communications of the ACM (CACM)*, 16(4):230–236.
- [Campo & Traina, 2003] Campo, C. Y. e Traina, A. J. M. (2003). Uma abordagem eficiente para recuperação de imagens médicas. In Oliveira, K. M. d. e Jr., A. R., editors, 3° Workshop de Informática Médica - WIM'2003 - in CR-ROM, pp. in CD-ROM, Fortaleza, CE.
- [Cheng et al., 2006] Cheng, T.-H., Wei, C.-P., e Tseng, V. S. (2006). Feature selection for medical data mining: Comparisons of expert judgment. In the 19th IEEE International Symposium on Computer-based Medical Systems, pp. 165–170, Salt Lake City, Utah, USA. IEEE Computer Society.
- [Chávez et al., 2001] Chávez, E., Navarro, G., Baeza-Yates, R. A., e Marroquín, J. L. (2001). Searching in metric spaces. ACM Computing Surveys, 33(3):273–321.
- [Ciaccia et al., 1997] Ciaccia, P., Patella, M., e Zezula, P. (1997). M-tree: An efficient access method for similarity search in metric spaces. In Jarke, M., editor, *Proceedings* of 23rd International Conference on Very Large Data Bases (VLDB), pp. 426–435, Athens, Greece. Morgan Kaufmann Publishers.
- [Comer, 1979] Comer, D. (1979). The ubiquitous b-tree. ACM Computing Surveys, 11(2):121–137.
- [Comer & Delp, 2000] Comer, M. L. e Delp, E. J. (2000). The em/mpm algorithm for segmentation of textured images: Analysis and further experimental results. *IEEE Transactions on Image Processing*, 9(10):1731–1744.

- [Emran & Ye, 2001] Emran, S. M. e Ye, N. (2001). Robustness of canberra metric in computer intrusion detection. In *IEEE Workshop on Information Assurance and Security*, West Point, NY, USA.
- [Faloutsos, 1996] Faloutsos, C. (1996). Searching Multimedia Databases by Content. Kluwer Academic Publishers, Boston, MA.
- [Faloutsos & Lin, 1995] Faloutsos, C. e Lin, K.-I. D. (1995). Fastmap: A fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. In Carey, M. J. e Schneider, D. A., editors, ACM SIGMOD International Conference on Management of Data, pp. 163–174, San Jose, CA. ACM Press.
- [Felipe & Traina, 2005] Felipe, J. C. e Traina, A. J. M. (2005). Desenvolvimento de métodos para extração, comparação e análise de características intrínsecas de imagens médicas, visando à recuperação perceptual por conteúdo. Tese de doutorado, Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, Brasil.
- [Felipe et al., 2006] Felipe, J. C., Traina, A. J. M., Balan, A. G. R., Marques, P. M. d. A., e Traina Jr., C. (2006). Comparing images with distance functions based on attribute interaction. In 21st Annual ACM Symposium on Applied Computing (SAC'06), pp. 1398–1399, Dijon, France.
- [Felipe et al., 2003] Felipe, J. C., Traina, A. J. M., e Traina Jr., C. (2003). Retrieval by content of medical images using texture for tissue identification. In 16th IEEE Symposium on Computer-based Medical Systems, pp. 175–180, New York. IEEE Computer Society.
- [Felipe et al., 2005] Felipe, J. C., Traina, A. J. M., e Traina Jr., C. (2005). Global warp metric distance: Boosting content-based image retrieval through histograms. In Tsai, J. J. P., Muhlhauser, M., e Takizawa, M., editors, *IEEE International Symposium on Multimedia-ISM2005*, pp. 295–302, Irvine, CA. IEEE Press.
- [Folk et al., 1998] Folk, M. J., Zoellick, B., e Riccardi, G. (1998). *File structures: An Object-Oriented Approach with C++*. Addison-Wesley, Reading, USA, 3 edition.
- [Furuie et al., 1999] Furuie, S., Bertozoo, N., Figueiredo, J., e Yamaguti, M. (1999). Archiving and retrieving long-term cineangiographic images in a pacs. *Computers in Cardiology*, pp. 435–438.
- [Gaede & Günther, 1998] Gaede, V. e Günther, O. (1998). Multidimensional access methods. ACM Computing Surveys, 30(2):170–231.
- [Guttman, 1984] Guttman, A. (1984). R-tree : A dynamic index structure for spatial searching. In ACM International Conference on Management of Data, pp. 47–57, Boston, MA. ACM PRess.
- [Haralick, 1979] Haralick, R. M. (1979). Statistical and structural approaches to texture. Proceedings of the IEEE, 67(5):786–804.
- [Haralick et al., 1973] Haralick, R. M., Shanmugam, K., e Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621.

- [Hellerstein et al., 1995] Hellerstein, J. M., Naughton, J. F., e Pfeffer, A. (1995). Generalized search trees for database systems. In *International Conference on Very Large Databases (VLDB)*, pp. 562–573, Zurich, Switzerland.
- [Hiransakolwong et al., 2005] Hiransakolwong, N., Hua, K. A., Koompairojn, S., Vu, K., e Lang, S.-D. (2005). An adaptive distance computation technique for image retrieval systems. In SAC '05: Proceedings of the 2005 ACM symposium on Applied computing, pp. 1195–1199, New York, NY, USA. ACM Press.
- [Jolliffe, 1986] Jolliffe, I. T. (1986). Principal Component Analysis. Springer-Verlag, New York, NY, USA.
- [Keogh, 2002] Keogh, E. (2002). Exact indexing of dynamic time warping. Proceedings of the 28th international conference on Very Large Data Bases, pp. 406–417, Hong Kong, China.
- [Keogh & Ratanamahatana, 2005] Keogh, E. e Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and Information Systems*, 7(3):358–386.
- [Khotanzad & Hong, 1990] Khotanzad, A. e Hong, Y. H. (1990). Invariant image recognition by zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 12(5):489–497.
- [Knuth, 1973] Knuth, D. E. (1973). The Art of Computer Programming Sorting and Searching, v. 3. Addison-Wesley, Reading, USA.
- [Korn et al., 2001] Korn, F., Pagel, B.-U., e Faloutsos, C. (2001). On the "Dimensionality Curse" and the "Self-Similarity Blessing". *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 13(1):96–111.
- [Korn et al., 1996] Korn, F., Sidiropoulos, N., Faloutsos, C., Siegel, E. L., e Protopapas, Z. (1996). Fast nearest neighbor search in medical image databases. In *International Conference on Very Large Databases (VLDB)*, pp. 215–226, Bombay, India. Morgan Kaufmann.
- [Lee et al., 2000] Lee, S.-L., Chun, S.-J., Kim, D.-H., Lee, J.-H., e Chung, C.-W. (2000). Similarity search for multidimensional data sequences. In *ICDE '00: Proceedings of the* 16th International Conference on Data Engineering 599, p., Washington, DC, USA. IEEE Computer Society.
- [Lew et al., 2006] Lew, M. S., Sebe, N., Djeraba, C., e Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. ACM Trans. Multimedia Comput. Commun. Appl., 2(1):1–19.
- [Lima et al., 1998] Lima, L. R. S. d., Laender, A. H. F., e Ribeiro-Neto, B. A. (1998). A hierarchical approach to the automatic categorization of medical documents. In Proceedings of the 7th International Conference on Information and Knowledge Management (CIKM), pp. 132–139, Bethesda, Maryland.
- [Ling & Okada, 2006] Ling, H. e Okada, K. (2006). Diffusion distance for histogram comparison. In CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 246–253, Washington, DC, USA.
- [Long et al., 2003] Long, F., Zhang, H., e Feng, D. D. (2003). Multimedia Information Retrieval and Management - Technological Fundamentals and Applications, Capitulo "Fundamentals of Content-Based Image Retrieval". Springer, Berlin.
- [Malcok et al., 2006] Malcok, M., Aslandogan, Y., e Yesildirek, A. (2006). Fractal dimension and similarity search in high-dimensional spatial databases. In *IEEE International Conference on Information Reuse and Integration (IRI 2006)*, pp. 380–384, Waikoloa, Hawaii, USA. IEEE Computer Society.
- [Marsh, 1997] Marsh, A. (1997). Euromed the creation of a telemedical information society. In 10th IEEE Symposium on Computer Based Medical Systems, pp. 86–91, Maribor, Slovenia.
- [Müller et al., 2004] Müller, H., Michoux, N., Bandon, D., e Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1):1–23.
- [Neumann & Gegenfurtner, 2006] Neumann, D. e Gegenfurtner, K. R. (2006). Image retrieval and perceptual similarity. ACM Transactions on Applied Perception, 3(1):31– 47.
- [Pagel et al., 2000] Pagel, B.-U., Korn, F., e Faloutsos, C. (2000). Deflating the dimensionality curse using multiple fractal dimensions. 16th International Conference on Data Engineering (ICDE), 00:589–599.
- [Papadias et al., 1999] Papadias, D., Mamoulis, N., e Theodoridis, Y. (1999). Processing and optimization of multiway spatial joins using r-trees. In ACM Symposium on Principles of Database Systems (PODS), pp. 44–55, Philadelphia, PA. ACM Press.
- [Ribeiro et al., 2005] Ribeiro, M. X., Balan, A. G. R., Felipe, J. C., Traina, A. J. M., e Traina Jr., C. (2005). Mining statistical association rules to select the most relevant medical image features. In *First International Workshop on Mining Complex Data* (*IEEE MCD'05*), pp. 91–98, Houston, USA. IEEE Computer Society.
- [Rubner & Tomasi, 2001] Rubner, Y. e Tomasi, C. (2001). Perceptual Metrics for Image Database Navigation. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers.
- [Santini & Gupta, 2001] Santini, S. e Gupta, A. (2001). A wavelet data model for image databases. In *IEEE International Conference on Multimedia and Expo*, pp. 1103–1106, Tokyo, Japan. IEEE Computer Society.
- [Santini & Jain, 1996] Santini, S. e Jain, R. (1996). Gabor space and the development of preattentive similarity. In 13th International Conference on Pattern Recognition, v. 01, pp. 40–44, Los Alamitos, CA, USA. IEEE Computer Society.
- [Santini & Jain, 1999] Santini, S. e Jain, R. (1999). Similarity measures. IEEE Transactions on Patterns Analysis and Machine Intelligence (TPAMI), 21(9):871–883.
- [Sellis et al., 1987] Sellis, T. K., Roussopoulos, N., e Faloutsos, C. (1987). The R+-tree: A dynamic index for multi-dimensional objects. In *International Conference on Very Large Databases (VLDB)*, pp. 507–518, Brighton, England. Morgan Kaufmann.
- [Shannon, 2001] Shannon, C. E. (2001). A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review, 5(1):3–55.

- [Smeulders et al., 2000] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., e Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions* on Pattern Analysis and Machine Intelligence (TPAMI), 22(12):1349–1380.
- [Sonka et al., 1998] Sonka, M., Hlavac, V., e Boyle, R. (1998). *Image Processing: Analysis and Machine Vision*. Chapman & Hall, 2 edition.
- [Souza & Traina Jr., 2006] Souza, E. P. M. e Traina Jr., C. (2006). *Identificação de correlações usando a Teoria dos Fractais*. Tese de doutorado, Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, Brasil.
- [Sung et al., 2000] Sung, M., Kim, M., Sung, M.-W., Kim, E., e Yoo, J. (2000). Comed: A real-time collaborative medicine system. In 13th IEEE Symposium on Computer-Based Medical Systems (CBMS), pp. 215–220, Houston, Texas. IEEE Computer Society.
- [Torres & Falcão, 2006] Torres, R. d. S. e Falcão, A. X. (2006). Content-Based Image Retrieval: Theory and Applications. *Revista de Informática Teórica e Aplicada*, 13(2):161– 185.
- [Traina et al., 2004] Traina, A. J. M., Balan, A. G. R., Bortolotti, L. M., e Jr., C. T. (2004). Content-based image retrieval using approximate shape of objects. In 17th IEEE Symposium on Computer-Based Medical Systems (CBMS), pp. 91–96, Washington, DC, USA. IEEE Computer Society.
- [Traina & Traina Jr., 2003] Traina, A. J. M. e Traina Jr., C. (2003). Similarity search in multimedia databases. In Furht, B. e Marques, O., editors, *Handbook of Video Databases - Design and Applications*, v. 1, pp. 711–738. CRC Press.
- [Traina et al., 2003] Traina, A. J. M., Traina Jr., C., Bueno, J. M., Chino, F. J. T., e Marques, P. M. d. A. (2003). Efficient content-based image retrieval through metric histograms. World Wide Web Journal (WWWJ), 6(2):157–185.
- [Traina et al., 2002a] Traina, A. J. M., Traina Jr., C., Bueno, J. M., e Marques, P. M. d. A. (2002a). The metric histogram: A new and efficient approach for content-based image retrieval. In Zhou, X. e Pu, P., editors, Sixth IFIP Working Conference on Visual Database Systems, v. 216 of IFIP Conference Proceedings, pp. 297–311, Brisbane, Australia. Kluwer Academic Publishers.
- [Traina et al., 2007] Traina, Caetano, J., Santos Filho, R. F., Traina, A. J. M., Vieira, M. R., e Faloutsos, C. (2007). The omni-family of all-purpose access methods: A simple and effective way to make similarity search more efficient. *The International Journal* on Very Large Databases, 16(4):483–505.
- [Traina et al., 2002b] Traina, Caetano, J., Traina, A. J. M., Santos, Roberto Figueira, F., e Faloutsos, C. (2002b). How to improve the pruning ability of dynamic metric access methods. In *Proceedings of the 11th International Conference on Information* and Knowledge Management (CIKM), pp. 219–226, McLean, USA. ACM Press.
- [Traina et al., 2000a] Traina, Caetano, J., Traina, A. J. M., Seeger, B., e Faloutsos, C. (2000a). Slim-trees: High performance metric trees minimizing overlap between nodes. In Zaniolo, C., Lockemann, P. C., Scholl, M. H., e Grust, T., editors, *International Conference on Extending Database Technology (EDBT)*, v. 1777 of *Lecture Notes in Computer Science*, pp. 51–65, Konstanz, Germany. Springer.

- [Traina et al., 2000b] Traina, Caetano, J., Traina, A. J. M., Wu, L., e Faloutsos, C. (2000b). Fast feature selection using fractal dimension. In Medeiros, C. M. B. e Becker, K., editors, *Brazilian Symposium on Databases (SBBD)*, pp. 158–171, João Pessoa, PB.
- [Tversky, 1977] Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4):327–352.
- [Tversky & Gati, 1982] Tversky, A. e Gati, I. (1982). Similarity, separability and the triangle inequality. *Psychological Review*,, 89(1):123–154.
- [Twa et al., 2003] Twa, M. D., Parthasarathy, S., Raasch, T. W., e Bullimore, M. A. (2003). Decision tree classification of spatial data patterns from videokeratography using zernike polynomials. In *SIAM International Conference on Data Mining*, pp. 3–12, San Francisco, CA. SIAM.
- [Vasconcelos & Lippman, 2000] Vasconcelos, N. e Lippman, A. (2000). A unifying view of image similarity. In 15th International Conference on Pattern Recognition (ICPR), pp. 38–41, Washington, DC, USA. IEEE Computer Society.
- [Vieira & Traina Jr., 2004] Vieira, M. R. e Traina Jr., C. (2004). *DBM-tree: Método de acesso métrico sensível à densidade local*. Dissertação de mestrado, Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, Brasil.
- [Wilson & Martinez, 1997] Wilson, D. R. e Martinez, T. R. (1997). Improved heterogeneous distance functions. *Journal of Artificial Intelligence Research*, 6:1–34.
- [Zegarra et al., 2008] Zegarra, J. P. M., Leite, N. J., e Torres, R. d. S. (2008). Waveletbased Feature Extraction for Fingerprint Image Retrieval. *Journal of Computational* and Applied Mathematics.
- [Zhang et al., 2000] Zhang, C., Meng, W., Wu, Z., e Zhang, Z. (2000). Webssql a query language for multimedia web documents. In Papazoglou, M. P. e Sheth, A., editors, *IEEE Advances in Digital Libraries 2000 (ADL'00)*, pp. 58–67, Washington, D.C. IEEE Computer Society.
- [Zhang & Lu, 2001] Zhang, D. S. e Lu, G. (2001). Content-based shape retrieval using different shape descriptors: A comparative study. In *IEEE Conference on Multimedia* and Expo (ICME'01), pp. 317–320, Tokyo, Japan. IEEE Computer Society.
- [Zhang & Lu, 2002] Zhang, D. S. e Lu, G. (2002). An integrated approach to shape based image retrieval. In *Fifth Asian Conference on Computer Vision*, pp. 652–657, Melbourne, Australia.

Este documento foi preparado com o formatador de textos  ${\rm I\!\!\!\!A} T_{\rm E} X$  .