

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

**CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA**

**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO**

**UTILIZAÇÃO DE CONDIÇÕES DE CONTORNO PARA  
COMBINAÇÃO DE MÚLTIPLOS DESCRITORES EM  
CONSULTAS POR SIMILARIDADE**

**RODRIGO FERNANDES BARROSO**

**ORIENTADOR: PROF. DR. RENATO BUENO**

São Carlos - SP  
Março/2014

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**UTILIZAÇÃO DE CONDIÇÕES DE CONTORNO PARA  
COMBINAÇÃO DE MÚLTIPLOS DESCRITORES EM  
CONSULTAS POR SIMILARIDADE**

**RODRIGO FERNANDES BARROSO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Engenharia de Software, Banco de Dados e Interação Humano Computador.

Orientador: Prof. Dr. Renato Bueno.

São Carlos - SP  
Março/2014

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária da UFSCar**

B277cc Barroso, Rodrigo Fernandes.  
Utilização de condições de contorno para combinação de múltiplos descritores em consultas por similaridade / Rodrigo Fernandes Barroso. -- São Carlos : UFSCar, 2014.  
78 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2014.

1. Banco de dados. 2. Dados complexos. 3. Consultas por similaridade. 4. Gap semântico. 5. Combinação de múltiplos descritores. I. Título.

CDD: 005.74 (20ª)

**Universidade Federal de São Carlos**  
**Centro de Ciências Exatas e de Tecnologia**  
**Programa de Pós-Graduação em Ciência da Computação**

**“Utilização de Condições de Contorno para  
Combinação de Múltiplos Descritores em  
Consultas por Similaridade ”**

Rodrigo Fernandes Barroso

Dissertação de Mestrado apresentada ao  
Programa de Pós-Graduação em Ciência da  
Computação da Universidade Federal de São  
Carlos, como parte dos requisitos para a  
obtenção do título de Mestre em Ciência da  
Computação

Membros da Banca:



---

**Prof. Dr. Renato Bueno**  
(Orientador - DC/UFSCar)



---

**Prof. Dr. Ricardo Rodrigues Ciferri**  
(DC/UFSCar)



---

**Prof. Dr. Joaquim Cezar Felipe**  
(USP/RP)

São Carlos  
Março/2014

*Dedico este trabalho aos meus pais José Francisco Barroso e Maria Regina Fernandes Barroso, e a minha esposa Michelle Sandrin dos Santos Barroso que estão sempre do meu lado me apoiando e dando força para seguir em frente e lutar pelos meus sonhos. E as minhas irmãs, familiares e amigos que fazem parte e completam minha vida.*

*Em memória de meu pai José Francisco Barroso, sempre do meu lado.*

*Rodrigo Fernandes Barroso*

# AGRADECIMENTO

Agradeço primeiramente a Deus, pela oportunidade de realizar mais esta etapa muito importante na minha vida e por sempre estar ao meu lado. Agradeço também ao meu orientador, professor doutor Renato Bueno por ter acreditado em mim, pela paciência com as minhas limitações, serei eternamente grato pela oportunidade de conhecer e trabalhar juntos. Também aos companheiros alunos e professores do Grupo de Banco de Dados do Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, das quais engrandeceram muito este trabalho com diversas opiniões e pontos de vista. Aos amigos do Departamento de Computação da UFSCar, foi um prazer compartilhar de vários momentos juntos durante os três anos do mestrado. Agradeço também a todos os meus amigos que me apoiaram e acreditaram sempre no meu potencial para realização deste trabalho. Agradeço também pelo apoio financeiro recebido pela CNPq.

# RESUMO

Dados complexos, como imagens, enfrentam problemas semânticos em suas consultas que comprometem a qualidade dos resultados. Esses problemas são caracterizados pela divergência entre a interpretação semântica desses dados e a forma como são representados computacionalmente em características de baixo nível. Nessa representação são utilizados vetores de características que descrevem características intrínsecas (como cor, forma e textura) em atributos qualificadores. Ao analisar a similaridade em dados complexos percebe-se que essas características intrínsecas se complementam na representação do dado, bem como é realizada pela percepção humana e por este motivo a utilização de múltiplos descritores tende a melhorar a capacidade de discriminação dos dados. Nesse contexto, outro fato relevante é que em um conjunto de dados, alguns subconjuntos podem apresentar características intrínsecas específicas essenciais que melhor evidenciam seus elementos do restante dos dados. Com base nesses preceitos, este trabalho propõe a utilização de condições de contorno para delimitar estes subconjuntos e determinar o melhor balanceamento de múltiplos descritores para cada um deles, com o objetivo de diminuir o “gap semântico” nas consultas por similaridade. Em todos os experimentos realizados a utilização da técnica proposta sempre apresentou melhores resultados. Em comparação a utilização de descritores individuais com as mesmas condições de contorno e sem condições de contorno, e também a combinação de descritores para o conjunto todo sem a utilização de condições de contorno.

**Palavras-chave:** Dados Complexos, Consulta por Similaridade, Gap Semântico, Combinação de Múltiplos Descritores.

# ABSTRACT

Complex data, like images, face semantic problems in your queries that might compromise results quality. Such problems have their source on the differences found between the semantic interpretation of the data and its low level machine language. In this representation are utilized feature vectors that describe intrinsic characteristics (like color, shape and texture) into qualifying attributes. Analyzing the similarity in complex data, perceives that these intrinsic characteristics complemented the representation of data, as well as is carried out by human perception and for this reason the use of multiple descriptors tend to improve the ability of discrimination data. In this context, another relevant fact is that in a data set, some subsets may present essential specific intrinsic characteristics to better show their rest of the data elements. Based in such premises, this work proposes the use of boundary conditions to identify these subsets and then use the best descriptor combination balancing for each of these, aiming to decrease the existing “semantic gap” in similarity queries. Throughout the conducted experiments the use of the proposed technique had better results when compared to use individual descriptor using the same boundary conditions and also using descriptors combination for the whole set without the use of boundary conditions.

**Keywords:** Complex Data, Similarity Queries, Semantic Gap, Combination of Multiple Features.

# LISTA DE FIGURAS

Figura 2.1 – Imagens com seus respectivos histogramas.....	23
Figura 2.2 – Consulta por Abrangência.....	24
Figura 2.3 – Consulta por k-vizinhos mais próximos (k=4).....	25
Figura 2.4 – Exemplo de Histograma extraído de uma imagem.....	27
Figura 2.5 – Exemplo de Vetor de Características extraído de uma imagem.....	28
Figura 3.1 – Gráfico da variação da precisão média na combinação de descritores. (BUENO et al., 2009) .....	33
Figura 4.1 – Gráfico da média de precisão do conjunto de dados MRI_704. (BUENO et al.,2011) .....	40
Figura 4.2 – Arquitetura do ambiente de treinamento do método proposto. ....	45
Figura 4.3 – Arquitetura do ambiente de utilização do método proposto. ....	46
Figura 5.1– a) Imagem Original, b) Imagem resultante do janelamento para Histograma Baixo– realça regiões mais brilhantes, c) Imagem resultante do janelamento para Histograma Alto – realça regiões mais escuras (PONCIANO-SILVA et al., 2009).....	50
Figura 5.2 – Amostragem das Classes de Imagens de Exames de Ressonância Magnética.....	51
Figura 5.3 – Curvas de Precisão x Revocação de consultas TC_Pulmão utilizando Textura com as Funções de Distância Minkowski e Canberra.....	53
Figura 5.4 – Curvas de Precisão x Revocação das consultas executadas para o conjunto de dados TC_Pulmão. ....	55
Figura 5.5 – Precisões médias obtidas nas consultas realizadas no conjunto de dados TC_Pulmão.....	56
Figura 5.6 – Curvas de Precisão x Revocação das consultas executadas no conjunto de dados TC_Pulmão.....	57
Figura 5.7 – Precisões médias obtidas nas consultas realizadas no conjunto de dados TC_Pulmão.....	58
Figura 5.8 – Curvas de Precisão x Revocação das consultas executadas no conjunto de dados ROI_TCPulmão (fase de treinamento). ....	60
Figura 5.9 – Precisões médias obtidas nas consultas realizadas no conjunto de dados ROI_TCPulmão(fase de treinamento). ....	60

Figura 5.10 – Precisões médias obtidas nas consultas realizadas no conjunto de dados ROI_TCPulmão (conjunto todo) .....	61
Figura 5.11 – Precisões médias obtidas nas consultas realizadas no conjunto de dados ROI_TCPulmão(conjunto todo). .....	61
Figura 5.12 – Curvas de Precisão x Revocação das consultas executadas no conjunto de dados MRI_704.....	63
Figura 5.13 – Precisões médias obtidas nas consultas realizadas no conjunto de dados MRI_704.....	63
Figura 5.14 – Precisão média combinando linearmente Histograma Baixo – L2 / w Textura - Canberra.....	65
Figura 5.15 – Precisão média combinando linearmente (Histograma Baixo – L2 + (0.11 * Textura – Canberra)) / w Histograma Tradicional - LInf.....	66
Figura 5.16 – Precisão média combinando linearmente (Histograma Baixo – L2 + (0.11 * Textura – Canberra) + (0.10 * Histograma Tradicional – LInf)) / w Histograma Alto – L2.....	67
Figura 5.17 – Curva de Precisão x Revocação da Combinação de Descritores feito por algoritmos exaustivos e a Combinação de Descritores realizado neste experimento 4.....	68
Figura 5.18 –Curvas de Precisão x Revocação das consultas executadas para o conjunto de dados TC_Pulmão comparando com a Combinação por Condição de Contorno partindo do melhor descritor. ....	69
Figura 5.19 –Precisão média das consultas executadas para o conjunto de dados TC_Pulmão comparando com a Combinação por Condição de Contorno partindo do melhor descritor.....	69

# LISTA DE TABELAS

Tabela 5.1 Conjunto de Características Usadas nos experimentos .....	51
Tabela 5.2 Definição dos Descritores.....	54
Tabela 5.3 Métricas por Condição de Contorno .....	54
Tabela 5.4 Definição dos Descritores.....	56
Tabela 5.5 Métricas por Condição de Contorno .....	57
Tabela 5.6 Métricas por Condição de Contorno .....	59
Tabela 5.7 Métricas por Condição de Contorno .....	62

# SUMÁRIO

<b>CAPÍTULO 1 - INTRODUÇÃO.....</b>	<b>12</b>
1.1 Contexto do Trabalho .....	12
1.2 Motivação.....	14
1.3 Objetivo .....	16
1.4 Organização da Monografia .....	17
<b>CAPÍTULO 2 - FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>18</b>
2.1 Estruturação do Capítulo.....	18
2.2 Dados Complexos .....	19
2.3 Espaço Métrico.....	21
2.4 Consulta por Similaridade .....	22
2.5 Funções de Distância .....	25
2.6 Extrator de características.....	27
2.7 Considerações Finais .....	29
<b>CAPÍTULO 3 - TRABALHOS CORRELATOS .....</b>	<b>30</b>
3.1 Estruturação do Capítulo.....	30
3.2 Combinação de Características .....	31
3.2.1 Abordagem Básica .....	31
3.2.2 Combinação de Múltiplos Descritores .....	32
3.3 Condições de Contorno.....	35
3.4 Considerações Finais .....	37
<b>CAPÍTULO 4 - COMBINAÇÃO DE MÚLTIPLOS DESCRITORES UTILIZANDO CONDIÇÕES DE CONTORNO .....</b>	<b>38</b>
4.1 Estruturação do Capítulo.....	38
4.2 Introdução .....	39
4.3 Metodologia.....	42
4.4 Combinação de Múltiplos Descritores Utilizando Condições de Contorno .....	43
4.5 Considerações Finais .....	47
<b>CAPÍTULO 5 - RESULTADOS .....</b>	<b>48</b>

5.1 Estruturação do Capítulo.....	48
5.2 Experimentos .....	49
5.2.1 Definição de descritores .....	52
5.2.2 Conjunto de Imagens de Tomografia Computadorizada de Pulmão (TC_Pulmão) .....	54
5.2.3 Conjunto de Imagens de Áreas de Interesse (ROI) de Tomografia Computadorizada de Pulmão (ROI_TCPulmão) .....	58
5.2.4 Conjunto de Imagens de Ressonância Magnética (MRI_704) .....	62
5.2.5 Combinação de múltiplos descritores partindo do melhor descritor .....	64
5.3 Análise dos resultados obtidos.....	70
<b>CAPÍTULO 6 - CONCLUSÃO .....</b>	<b>71</b>
6.1 Considerações Finais .....	71
6.2 Principais Contribuições .....	72
6.3 Trabalho Futuros .....	72
<b>REFERÊNCIAS.....</b>	<b>73</b>

# Capítulo 1

## INTRODUÇÃO

---

*Este capítulo apresenta o contexto em que este trabalho está inserido e a motivação que deu origem a este projeto de pesquisa. Em seguida são discutidos os objetivos a serem alcançados, finalizando com a descrição da organização desta dissertação de mestrado.*

### 1.1 Contexto do Trabalho

A necessidade de armazenar e manipular dados não tradicionais (como imagens, sons e vídeos) vem se tornando cada vez mais comum em diversos sistemas de computação, exigindo que o gerenciamento e a recuperação desses dados ocorram de forma mais eficiente e efetiva. Esses dados são comumente chamados de dados complexos.

Sistemas de computadores que utilizam modelos de dados relacionais e objeto-relacionais tratam de uma forma geral dados dos domínios de texto curto ou numérico. Para tratar estes domínios de informação os Sistemas de Gerenciamento de Banco de Dados (SGBDs) atuais utilizam-se de uma propriedade chamada de relação de ordem total (ROT) que permite comparar quaisquer pares de elementos do conjunto, determinando a antecedência/precedência entre eles, podendo assim ser denominados de conjunto totalmente ordenáveis.

Porém a necessidade de armazenar e manipular dados mais complexos vem sendo cada vez mais comum em diversas atividades (TRAINA JR., TRAINA, FILHO & FALOUTSOS, 2002). Esses sistemas devem permitir além de armazenar, gerenciar e recuperar os dados de forma eficiente e efetiva (KEIM et al., 2002).

Porém, nem sempre esses tipos de dados irão satisfazer a propriedade de relação de ordem total.

Esses tipos de dados são comumente chamados de dados complexos e em muitos casos não possuem o conceito de dimensão associado, o que impossibilita consultas topológicas ou cardinais.

Em muitos dos sistemas existentes a recuperação desses dados complexos é efetuada sobre a descrição textual sobre o conteúdo (KEIM et al., 2002), como por exemplo, para encontrar uma música pelo nome do compositor ou procurar uma paisagem informando o tipo de imagem desejada. No entanto, esse formato nem sempre retorna um bom resultado, além de não oferecer recursos satisfatórios para interagir com vários tipos de mídia (MORIKAWA & SILVA, 2012).

Para comparar conjuntos de dados complexos utiliza-se a similaridade entre pares de elementos que indica o quão um objeto é parecido ou distinto de outro. Essa comparação proporciona um grau de similaridade (ou dissimilaridade) entre cada par de elementos do conjunto, possibilitando, por exemplo, trazer os  $k$  elementos mais parecidos com um elemento central de consulta. Este tipo de consulta é denominado Consulta por Similaridade.

Usualmente nas consultas por Similaridade não são comparados os elementos de dados complexos em si, e sim características intrínsecas que foram extraídas desses elementos, como exemplo, em imagens pode-se extrair características relativas à textura, forma ou cor e comparar todos os elementos através destas características. Consultas por similaridade que utilizam informações extraídas do conteúdo do dado complexo também são chamadas de Consultas Baseadas em Conteúdo. Este conteúdo é extraído comumente no formato de um vetor de características e representa algum aspecto do elemento.

Após a obtenção desses vetores de características, são utilizadas funções matemáticas sobre eles para calcular a similaridade entre pares de elementos, sendo estas denominadas funções de distância ou funções de similaridade.

Estas funções resultam em um valor numérico que determina o grau de diferença ou semelhança entre elementos, tornando este fator o mais importante e a solução mais adequada para consultas por similaridade. A combinação vetor de característica e função de distância é comumente encontrado na literatura como binômio (BUGATTI, 2008) ou descritor (da SILVA TORRES & FALCÃO, 2006).

## 1.2 Motivação

Cada vez mais, a utilização de sistemas caracterizados pela manipulação com dados mais complexos e variados em relação às aplicações tradicionais comerciais e negócios tem sido explorada e tornou-se de grande interesse para a comunidade científica (da SILVA TORRES & FALCÃO, 2006) (DEEPAK et al., 2012) (GARGANO, NARDELLI, TALAMO, 1991) (LIU et al., 2007) (SILVA et al., 2012) (SINGH & KOTHARI, 2003) (VASCONCELOS & KUNT, 2001).

Um problema enfrentado pelas consultas por similaridade é o “gap semântico” (DESERNO, 2009) existente na forma como o dado é interpretado computacionalmente. Em outras palavras, estes problemas são caracterizados pela divergência entre como os dados são interpretados pela percepção humana realizada a partir de análises de conteúdo de alto nível e a forma como são representados computacionalmente em características de baixo nível.

Tem-se como exemplo, dado uma imagem de um animal como referência, é possível que a pesquisa por similaridade retorne imagens de veículos, paisagens ou qualquer outra imagem semanticamente diferente ao objetivo da consulta e ao resultado esperado porque essas imagens foram consideradas similares pelo algoritmo de busca em sua representação computacional.

Atualmente existem inúmeras pesquisas que atuam sobre dados complexos, bancos de dados multimídia e vários exemplos de aplicabilidade de consultas por similaridades podem ser observadas. Nessas pesquisas é possível observar a definição de uma melhor configuração entre o vetor de características e uma função de distância, tendo essa combinação empregada ao termo descritor ou binômio (BUGATTI, 2008) onde ressalta a importância do inter-relacionamento entre uma função de distância sobre o vetor de características.

Um fator importante a ser levado em consideração é a utilização de diferentes descritores para a pesquisa em subconjuntos distintos. Isso é dado pelo fato de que uma característica pode ser mais relevante para evidenciar um subconjunto do restante dos elementos, e esta mesma característica não ser tão efetiva para outro subconjunto de elementos (PONCIANO-SILVA, 2009) (PONCIANO-SILVA et al., 2009).

Portanto, devem ser encontrados quais descritores melhor evidenciam cada subconjunto do restante dos elementos, e assim utilizá-los em pesquisas direcionadas, certificando que em cada pesquisa está sendo utilizada a melhor característica para evidenciar o elemento e o subconjunto a qual ele possa pertencer.

Considerando como exemplo um banco de imagens médicas com características específicas que delimitem cada subconjunto, seja uma doença ou uma lesão que permita ser caracterizada por uma hipótese de diagnóstico, então a hipótese de diagnóstico será a condição que pré-define todos os subconjuntos da pesquisa. Tendo em vista a hipótese de diagnóstico como delimitador desses subconjuntos, possibilita-se que a consulta por similaridade recorra ao descritor que melhor evidencia seus elementos do restante do conjunto de dados.

Sendo assim, é possível definir que cada subgrupo de elementos, apoiando em condições de contorno, utilize de uma função de distância própria com o intuito de aumentar o grau de precisão das pesquisas (PONCIANO-SILVA, 2009) (PONCIANO-SILVA et al., 2009).

Outro fator comum em estudos recentes que se mostra muito efetivo na diminuição do gap semântico em consultas por similaridade é a utilização de combinação de múltiplos descritores (AREVALILLO-HERRÁEZ, FERRI & DOMINGO, 2010) (BROILO & NATALE, 2010) (BUENO et al., 2010) (BUENO et al. 2011) (BUSTOS et al. 2004) (CAICEDO et al., 2007) (FERREIRA et al. 2008) (FERREIRA et al. 2011) (HEESCH & RÜGER, 2002) pois as características acabam se complementando para representarem o conteúdo da imagem.

Porém, geralmente define-se um único balanceamento entre descritores para atender a relação toda. O balanceamento ou função de distância resultante deste produto pode ter maior grau de precisão média para o conjunto como um todo e não atender de forma efetiva certos grupos de elementos do conjunto.

O presente trabalho propõe que para cada subconjunto definido pelas condições de contorno, exista um balanceamento de descritores relacionado que lhe atenda com um melhor grau de precisão.

Visando executar tais consultas aplicadas a condições de contorno é possível propor um balanceamento otimizado para cada um dos grupos (subconjuntos) existentes no conjunto de dados complexos, delimitados por condições de contorno. Estas condições vinculadas a um balanceamento de descritores criado

especificadamente para atender aquele grupo, seria uma hipótese para melhorar a precisão das consultas por similaridade e diminuir o gap semântico da interpretação dos dados complexos.

### 1.3 Objetivo

Para um mesmo objeto é possível serem extraídos vários conjuntos de características que descrevem diferentes aspectos dos dados, por exemplo, em imagens são comumente utilizados forma, distribuição de cores e textura. Na literatura é possível analisar que consultas usando combinações de características apresentam resultados melhores levando em consideração seu grau de precisão (BUENO et al., 2011).

A utilização de condições de contorno e a combinação de múltiplos descritores se mostraram individualmente com grande potencial para auxiliar na diminuição do gap semântico em consultas por similaridade, se tornando uma grande motivação para serem utilizadas juntas.

As condições de contorno possibilitam utilizar características que melhor representam subconjuntos de imagens que possuem um comportamento semelhante na consulta por similaridade, e a combinação de múltiplos descritores por sua vez permite uma melhor representação de cada elemento já que cada característica se complementa.

A proposta deste trabalho é utilizar condições de contorno para delimitar subconjuntos de dados e então determinar o balanceamento ideal de descritores para cada um deles, possibilitando que consultas por similaridade utilizem a combinação de múltiplos descritores que melhor representa cada um desses subconjuntos, aumentando assim o grau de precisão das consultas.

Uma vez realizado o balanceamento de descritores, este ficará atrelado a uma condição de contorno, sendo possível utilizá-lo ao realizar as consultas por similaridade de forma transparente, a fim de retornar um resultado mais próximo do esperado.

É importante ressaltar que para execução da consulta por similaridade não é necessário conhecimentos específicos da computação. Ao efetuar essas consultas

devem ser informados apenas a imagem de referência para pesquisa e a condição de contorno, o melhor balanceamento será então determinado automaticamente.

Os experimentos para comprovar o método proposto foram realizados em três bancos de dados de imagens médicas pré-classificados e o grau de precisão obtido pelas consultas foram comparados com as técnicas: um descritor individualmente para o conjunto todo, um descritor individualmente para cada classe e a combinação de múltiplos descritores para o conjunto todo.

## 1.4 Organização da Monografia

Além desta introdução, esta monografia está organizada da seguinte forma:

- O Capítulo 2 descreve os fundamentos teóricos necessários para a compreensão deste trabalho. Serão contextualizadas as definições sobre Dados Complexos, Espaço Métrico e Métodos de Acesso Métrico, Consultas por Similaridade, Funções de Distância e Extrator de Características com ênfase para imagens;
- O Capítulo 3 faz uma contextualização sobre a combinação de múltiplos descritores e como é realizado o balanceamento entre eles, como as condições de contorno podem ser aplicadas e como elas agregam a pesquisa. Ainda são citadas algumas pesquisas correlatas e a relação entre elas e o tema proposto;
- O Capítulo 4 apresenta a metodologia para utilização de condições de contorno para determinar a combinação de múltiplos descritores, método este sendo a contribuição inovadora do trabalho;
- O Capítulo 5 descreve os experimentos realizados para validação do método proposto, junto a ele os desafios encontrados e outras sugestões de desenvolvimento; e
- O Capítulo 6 apresenta as conclusões finais do trabalho, juntamente com as contribuições e os trabalhos futuros.

# Capítulo 2

## FUNDAMENTAÇÃO TEÓRICA

---

*Este capítulo descreve os principais conceitos relacionados à esta dissertação de mestrado. O capítulo é iniciado com as definições e conceitos sobre Dados Complexos definições estas primordiais para entender o tipo da informação que é trabalhada na pesquisa. Logo em seguida é contextualizado Espaço Métrico e Métodos de Acesso Métrico (MAM). Na sequência são apresentadas definições sobre Consulta por Similaridade e Funções de Distância, essenciais para o gerenciamento de dados complexos. E por fim os Extratores de Características são definidos e exemplificados, com o foco voltado para o domínio de imagens.*

### 2.1 Estruturação do Capítulo

Este capítulo está estruturado da seguinte forma. Na seção 2.2 são descritos conceitos básicos sobre dados complexos: o que são, e exemplos de onde e como são utilizados. A seção 2.3 explica os conceitos de espaços métricos, que são usados como apoio ao trabalho de pesquisa para indexar e agilizar consultas em dados complexos. As seções 2.4, 2.5 e 2.6 fazem uma introdução à consulta por similaridade, explicando os tipos de consultas mais comuns, o que são funções de distância e vetores de características. O capítulo é finalizado na seção 2.7 com as considerações finais.

## 2.2 Dados Complexos

É cada vez mais comum a necessidade de armazenar e utilizar não apenas dados tradicionais em bancos de dados, mas também dados mais complexos, como por exemplo: textos longos, imagens, arquivos de áudio e vídeo, onde certamente o tipo imagem é o mais utilizado, aplicados em diversas áreas como: astronomia, museus (HACID & ZIGHED 2006) e medicina se destacando no apoio ao diagnóstico e sistemas de suporte a decisão clínica (DEEPAK et al., 2012).

Em algumas aplicações as imagens são os dados mais importantes ou o único dado extraído (EURIPIDES & FALOUTSOS, 1997), como em sensoriamento remoto e astronomia, que coletam grandes quantidades de dados de imagem através de estações terrestres para processamento, análise e arquivamento. Outras necessidades são identificadas em aplicações relacionadas à cartografia e meteorologia, contemplando imagens de mapas, entretenimento, história da arte, publicidade e na indústria (BERRETTI, DEL BIMBO & PALA, 2002). Por fim a medicina, onde um grande número de imagens de diversas modalidades (por exemplo, tomografia computadorizada, ressonância magnética, etc.) são produzidas diariamente e usadas no apoio à decisão clínica ou diagnóstico. (MÜLLER et al., 2004)

No apoio ao diagnóstico por imagens podemos citar o PACS (Picture Archiving and Communication System), como sendo o sistema responsável por receber as imagens de diversas origens ou diversos dispositivos de aquisição como aparelhos de tomografia, raio-x, etc., e disponibilizá-las para que médicos e outros sistemas possam acessá-las (PARE et al., 2005) (DEEPAK et al., 2012). Estes sistemas PACS são exemplos da utilização eficaz do uso de imagens para apoio ao diagnóstico, sendo que devem ter uma resposta rápida e seus resultados atenderem a expectativa do especialista (BUENO et al., 2009).

Embora muitas consultas executadas em sistemas PACS sejam baseadas em metadados, basear-se em características extraídas das imagens, como por exemplo suas características visuais, tem mostrado ser um complemento adequado para estas pesquisas (BUENO et al., 2010). Como estes tipos de dados complexos não possuem características que possibilitem efetuar uma ordenação ou comparação entre eles, no caso da recuperação e cruzamento das informações em dados

complexos, é necessário utilizar características intrínsecas extraídas dos elementos, comumente denominadas vetor de características.

Devido à complexidade dos elementos de um domínio de dados complexos, o mais comum é utilizar vetores de características extraídos para fazer a comparação dos elementos, calculando a distância entre eles, (BUENO, 2009). As consultas por similaridade serão realizadas sobre os vetores de características extraídos dos elementos (recuperação de dados por conteúdo), e serão estas características que vão determinar a distância entre os elementos, determinando a similaridade entre eles.

A partir deste vetor de características, é possível comparar um determinado elemento do tipo complexo com todos os elementos do conjunto, aplicando uma função de distância para determinar quais são os elementos mais parecidos (próximos) com o elemento de pesquisa.

Desta forma, um especialista no domínio da aplicação é o mais indicado para definir como serão comparados os dados (TRAINA & TRAINA JR., 2003). Esta concepção de funções de distância apropriadas entre os elementos é a questão essencial do processo de busca e depende da forma que será aplicada (EURIPIDES & FALOUTSOS, 1997).

Este processo de comparação é denominado Recuperação de Dados por Conteúdo (Content-based Retrieval) ou CBR, em imagens é conhecido por CBIR (Content-based Image Retrieval) sendo uma tecnologia de apoio ao gerenciamento de arquivos de imagens organizadas pelo seu conteúdo visual (VELTKAMP & TANASE, 2002) (DATTA et al., 2008) (AREVALILLO-HERRÁEZ, FERRI & MORENO-PICOT, 2011). Este processo de recuperação por conteúdo de imagens não deve ser tratado como um processo exato, pois raramente duas imagens são idênticas.

Difícilmente uma base de imagens utilizará a igualdade para fazer comparação entre seus elementos, da mesma forma que comparações como "menor que" e "maior que" também não fazem sentido, o mais comum é a busca por elementos similares ao elemento de busca (TRAINA JR., TRAINA, FILHO & FALOUTSOS, 2002), (AMATO et al., 2003).

## 2.3 Espaço Métrico

Dado uma função de distância que calcule e retorne um valor que indica à semelhança entre cada elemento do domínio de acordo com os critérios de busca, a comparação por similaridade de tais dados leva naturalmente a sua representação em espaços métricos ou multidimensionais.

Bancos de dados métricos são bases onde uma função de distância métrica é definida para pares de objetos do banco de dados, em destaque também aos bancos de dados de objetos de um espaço vetorial, onde os vetores são compostos de atributos numéricos (BRAUNMÜLLER et al., 2001).

Um Espaço Métrico  $M$ , é definido por  $M = (S,d)$  onde  $S$  é o conjunto de todos os elementos que atendem a propriedade do domínio, e  $d$  uma função de distância (ou métrica), entre esses elementos, definida como  $d: S \times S \rightarrow \mathbb{R}^+$ . Sendo que tal função determina quão distante um elemento está de outro e esta distância determina o grau de similaridade (ou melhor, dissimilaridade) entre estes elementos.

Dados quaisquer elementos  $s_1, s_2$  e  $s_3 \in S$ , a função de distância deve satisfazer as seguintes propriedades, para ser considerada uma métrica:

- I. Simetria: a distância retornada pela função  $d$  do elemento  $s_1$  ao elemento  $s_2$  deve ser a mesma distância do elemento  $s_2$  ao elemento  $s_1$ :  $d(s_1,s_2) = d(s_2,s_1)$ ;
- II. Não Negatividade: a distância  $d$  entre quaisquer elementos distintos deve ser sempre maior que 0 (zero) e menor ou igual a  $\infty$  (infinito):  $0 \leq d(s_1,s_2) \leq \infty$ , e a distância  $d$  entre um elemento a ele mesmo é 0 (zero):  $d(s_1,s_1) = 0$ ;
- III. Desigualdade Triangular: a distância  $d$  entre um elemento  $s_1$  para um elemento  $s_2$  deve ser sempre menor ou igual à distância entre eles passando por um terceiro elemento:  $d(s_1,s_2) \leq d(s_1,s_3) + d(s_3,s_2)$ ;

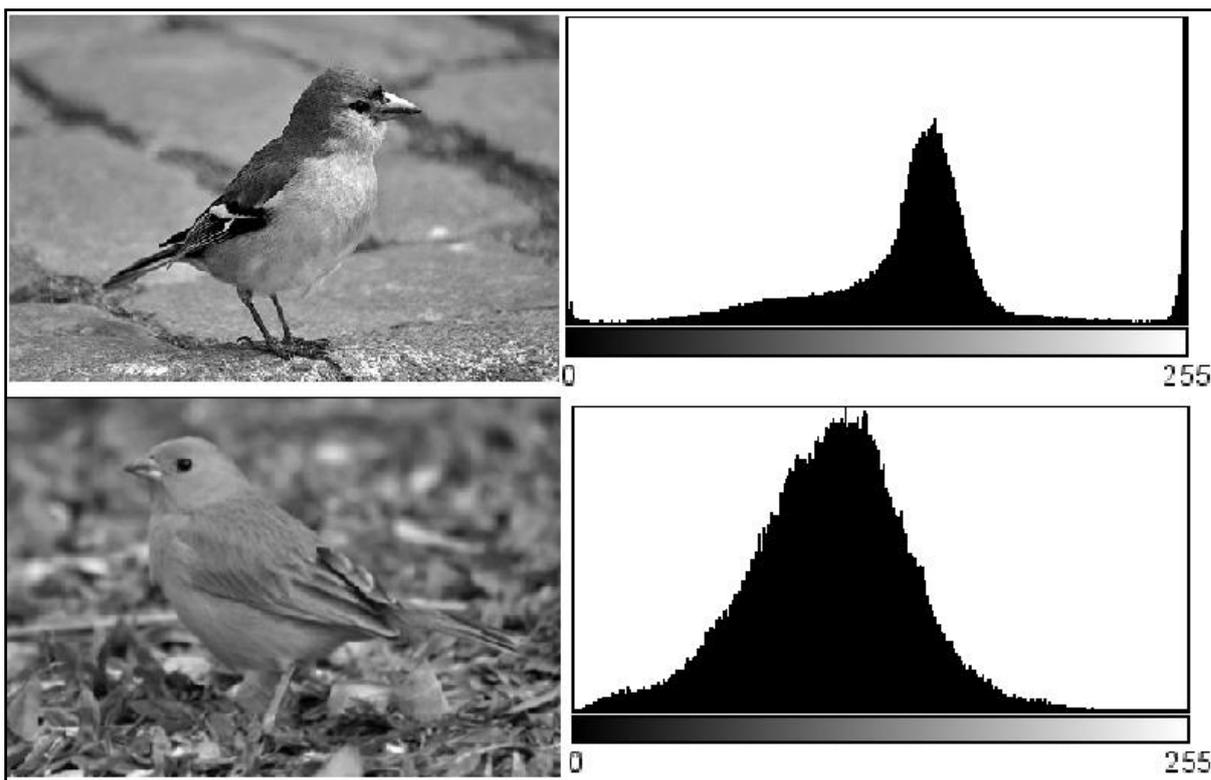
A utilização de dados em domínios métricos permite adotar estruturas de indexação chamadas genericamente de métodos de acesso métrico (MAM). Esses métodos de acesso métrico são estruturas de dados utilizados para indexação de dados complexos e permitem melhorar o desempenho de consultas por similaridade (TRAINA JR., TRAINA, FILHO & FALOUTSOS, 2002) (PETRAKIS & FALOUTSOS, 1997).

Dentre os métodos de acesso métrico existe um grupo que se baseia em armazenamento de dados em estruturas de páginas, blocos de dados, e árvores (AMATO et al., 2003) onde é possível encontrar na literatura diversos exemplos, dentre eles a M-Tree (CIACCIA, PATELLA & ZEZULA, 1997), Slim-Tree (TRAINA JR. et al., 2002), DBM-Tree (VIEIRA et al., 2004), Onion-Tree (CARÉLO et al., 2009) e cx-Sim (SOARES & KASTER, 2013).

## **2.4 Consulta por Similaridade**

As consultas por similaridade são comuns em diversas áreas. Além da recuperação de informação multimídia, também é comum sua aplicação em data mining, reconhecimento de padrões, aprendizado de máquina, visão computacional, compressão de dados e análise estatística dos dados. (AMATO et al., 2003).

Existem várias dificuldades em projetar um sistema de recuperação de imagens: primeiramente, porque a opinião sobre o que é "similar" comparado ao "não similar" pode ser subjetivo ou ter em pequenos detalhes esta diferenciação; o segundo motivo é pelo próprio modo da pesquisa por similaridade que utiliza vetores de características (informações de baixo nível) definindo que dados similares possuem vetores similares e nem sempre esta definição é compatível com a realidade. Podemos observar na Figura 2.1 do lado esquerdo, imagens que podem ser consideradas semanticamente semelhantes acompanhadas de suas respectivas características extraídas em forma de histograma de níveis de cinza. Nesta figura é possível observar que nem sempre imagens semanticamente semelhantes possuem representação computacional semelhantes (neste caso representadas pelo seu histograma).



**Figura 2.1 – Imagens com seus respectivos histogramas.**

Em bancos de dados multimídia, pesquisas por similaridade visam recuperar objetos semelhantes ao objeto de pesquisa com base em suas características visuais (HUANG et al., 2011). Existem dois tipos de consultas principais por similaridade em domínios métricos: consultas por abrangência (“range queries”) e consultas por  $k$ -vizinhos mais próximos (“ $k$ -nearest neighbor queries”). Estas consultas têm um papel importante em aplicações, tais como sistemas multimídia, sistemas de apoio à decisão, e mineração de dados (BRAUNMÜLLER et al., 2001).

Em consultas por abrangência são considerados os elementos que estejam até uma distância máxima do elemento de referência (ou elemento de busca). Por sua vez a consulta aos  $k$  vizinhos mais próximos limita os elementos do conjunto de resposta a um número máximo  $k$ .

Considerando um conjunto de elementos  $S = \{s_1, s_2, \dots, s_n\}$  de um domínio de dados  $\mathcal{S}$ ,  $S \subset \mathcal{S}$  e uma função de distância (métrica)  $d$  entre seus elementos.

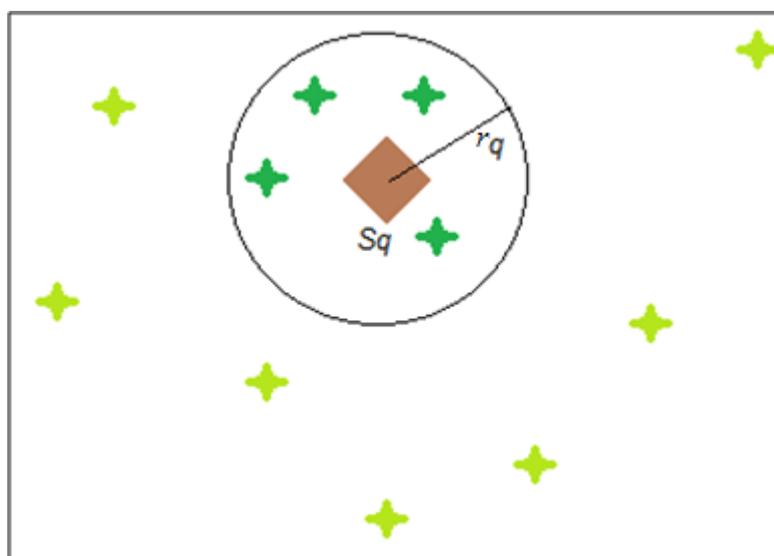
Ambas as consultas recebem um elemento do domínio de dados  $s_q \in \mathcal{S}$ , sendo este o elemento central da consulta, ou também chamado elemento de referência, onde a partir deste elemento será calculado quais objetos são mais semelhantes.

No caso da consulta por abrangência existe também o parâmetro  $r_q \geq 0$ , referente a abrangência (ou grau de dissimilaridade máximo) para a resposta, representando a tolerância para recuperação das informações (TRAINA JR., TRAINA, FILHO & FALOUTSOS, 2002) (AMATO et al. 2003).

Tal consulta obtém todos os elementos da base de dados  $S$  com grau de dissimilaridade máxima  $r_q$  do elemento central de consulta  $s_q$ . Esta função é representada formalmente como:

$$range(s_q, r_q) = \{s_i / s_i \in S, d(s_i, s_q) < r_q\}$$

A consulta por abrangência é ilustrada na Figura 2.2.



**Figura 2.2 – Consulta por Abrangência.**

Já na consulta por k-vizinhos mais próximos, o operador deve receber um parâmetro com o máximo  $k > 0$  de vizinhos que a consulta deve retornar (TRAINA JR., TRAINA, FILHO & FALOUTSOS, 2002). Esta consulta obtém os  $k$  elementos da base de dados  $S$  mais próximos do elemento de consulta  $s_q$ . Ela é representada formalmente por:

$$k-nn(s_q, k) = \{s_i, s_i \in A, A \subseteq S, |A| = k, \forall s_i \in A, s_j \in S - A, d(s_q, s_i) < d(s_q, s_j)\}$$

A consulta por k-vizinhos mais próximos é ilustrada na Figura 2.3, onde foi informado valor para  $k=4$ .

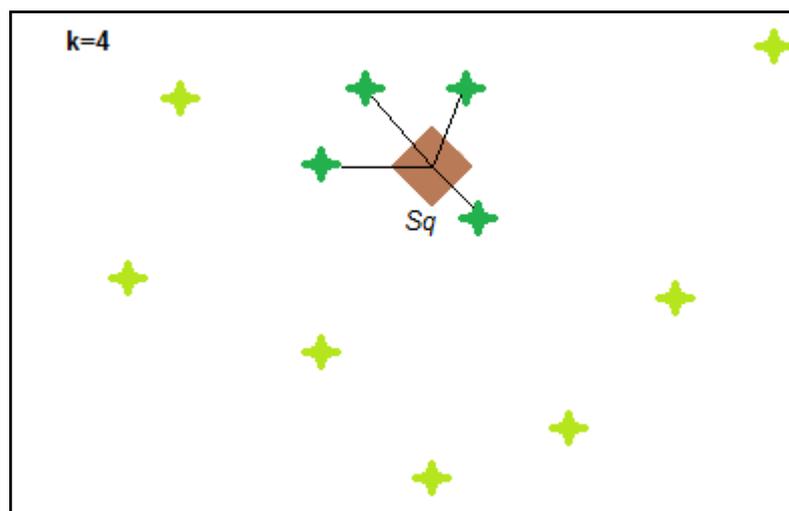


Figura 2.3 – Consulta por k-vizinhos mais próximos (k=4).

A semelhança entre elementos é expressa pela função de distância tal que um baixo valor de distância corresponde a um grau elevado de semelhança, ao passo que dois elementos com um alto valor de distância são considerados com um baixo grau de semelhança. (BRAUNMÜLLER et al., 2001).

As características extraídas são usadas para efetuar a pesquisa e determinar qual o grau de similaridade (ou dissimilaridade) entre os elementos. Os elementos recuperados não são necessariamente iguais ao elemento de busca, ao invés disto, a pesquisa deve ser realizada de forma aproximada, onde todos os elementos parecidos com o elemento de pesquisa serão retornados respeitando um grau de tolerância, ou seja, a representação dos dados utilizando vetores de características permite organizar os objetos de acordo com a semelhança entre eles (BUENO et al, 2011).

## 2.5 Funções de Distância

A função de distância deve mensurar uma distância (dissimilaridade) entre os elementos do conjunto, possibilitando uma classificação de acordo com essa distância de um elemento a um objeto de referência (AMATO et al. 2003). O resultado da função de distância  $d()$  definida para calcular o grau de dissimilaridade entre os elementos deve retornar um valor numérico que quantifica o quão distantes

estes elementos são, onde quanto mais próximo de zero for o valor, mais similares são os elementos, e quanto maior o valor menos similar eles serão (BUGATTI, 2008). Em outras palavras, a função de distância compara um par de elementos do domínio de dados e retorna um valor que representa a diferença entre eles (BUENO et al., 2009).

A escolha da função deve ser bem criteriosa, pois muitas vezes, é possível que a função selecionada não tenha o grau de precisão esperado. Por exemplo, aplicar a função de distância Euclidiana para forma onde funções de distância quadrática podem ter uma maior precisão para um banco de imagens (BRAUNMÜLLER et al., 2001).

Funções de distância amplamente utilizadas são da Família de Distâncias Minkowski, também conhecidas como funções de distância da Família  $L_p$ .

Elas são definidas como:

$$L_p(x, y) = \sqrt[p]{\sum_{i=0}^{d-1} |x_i - y_i|^p}$$

onde  $d$  é a dimensão do espaço.  $x$  e  $y \in \mathbb{S}$  são elementos de dimensão  $d$ .

De acordo com o valor definido a  $p$ , são obtidas as variações da família Minkowski, dentre elas, podemos citar: função de distância  $L_{\infty}$ , chamada de Infinity ou Chebychev, representado pela maior diferença de suas coordenadas ao longo de todas suas dimensões; função de distância  $L_1$ , conhecida como City Block ou Manhattan e corresponde a soma das diferenças entre cada coordenada; por fim a função de distância  $L_2$ , denominada distância euclidiana.

Outra função de distância comum e também utilizada por este trabalho é a Função de Distância Canberra. O seu cálculo é realizado pela diferença absoluta das características dos vetores dividida pela soma absoluta dos mesmos ao longo de todas suas dimensões, o que a torna sensível a pequenas variações e por este motivo é muito utilizada nas áreas de sequências de DNA e bioinformática. Ela é representada matematicamente como:

$$Canberra(x, y) = \sum_{i=0}^{d-1} \frac{|x_i - y_i|}{|x_i| + |y_i|}$$

## 2.6 Extrator de características

Na maioria dos conjuntos de dados complexos, é possível extrair vários conjuntos de características de um mesmo elemento, onde cada domínio de dados complexos possui distintos extratores de características. Por exemplo, em aplicações científicas na área de astronomia é medido um grande volume de características numéricas para cada elemento do banco de dados, tal como, a amplitude emitida em algumas frequências de banda (BRAUNMÜLLER et al., 2001).

No caso de bancos de dados de imagens, podemos exemplificar por algumas características, tais como: o histograma normalizado para cor (GONZALEZ & WOODS, 2010), Haralick para textura (HARALICK, SHANMUGAM & DINSTEIN, 1973) (HARALICK, 1979) e Zernike para forma (KHOTANZAD & HONG, 1990) (AMATO, MAINETTO & SAVINO, 1997).



**Figura 2.4 – Exemplo de Histograma extraído de uma imagem.**

No exemplo ilustrado na Figura 2.4, resultou no vetor da característica histograma apresentado na Figura 2.5:

```

[ 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
  0, 0, 0, 0, 0, 2, 10, 35, 40, 60, 89, 89, 63, 77, 75,
  71, 86, 67, 73, 100, 96, 104, 112, 145, 145, 178, 216, 219, 212, 255,
  221, 229, 223, 285, 296, 310, 309, 317, 304, 361, 368, 351, 413, 409, 429,
  431, 413, 465, 449, 477, 487, 455, 531, 522, 550, 600, 555, 651, 643, 739,
  664, 665, 698, 706, 774, 773, 698, 720, 709, 706, 730, 699, 719, 694, 689,
  674, 713, 704, 713, 735, 656, 635, 694, 642, 610, 653, 618, 570, 589, 554,
  564, 523, 565, 516, 508, 546, 516, 455, 433, 441, 463, 440, 428, 419, 384,
  396, 389, 363, 320, 328, 339, 288, 303, 292, 292, 296, 323, 304, 248, 254,
  246, 250, 222, 236, 216, 200, 245, 225, 207, 204, 199, 203, 189, 176, 158,
  147, 172, 151, 151, 134, 132, 117, 129, 106, 118, 88, 111, 90, 87, 75,
  81, 75, 82, 69, 71, 73, 70, 66, 63, 56, 60, 48, 56, 67, 54,
  60, 47, 46, 53, 51, 73, 54, 51, 43, 47, 51, 47, 53, 43, 37,
  48, 45, 61, 53, 60, 60, 54, 51, 78, 73, 61, 85, 82, 97, 111,
  71, 109, 115, 110, 139, 148, 212, 230, 248, 268, 298, 296, 330, 394, 464,
  465, 219, 143, 114, 105, 125, 148, 180, 271, 415, 784, 2086, 94008, 1045, 467,
  219, 122, 74, 28, 19, 13, 10, 3, 2, 0, 1, 0, 0, 0, 0,
  0]

```

**Figura 2.5 – Exemplo de Vetor de Características extraído de uma imagem.**

A extração das características de imagens ocorre quando as mesmas são submetidas a algoritmos de processamento de imagem, conhecidos como extratores de características, que resultam em um vetor numérico, chamado vetor de características, que descreve tais características em forma de assinatura matemática.

Com esses valores numéricos é possível através das funções de distância compará-los e identificar a semelhança entre os objetos (BUENO et al., 2010). A comparação feita pela consulta de similaridade é baseada diretamente no vetor de características, ao invés de utilizar o elemento (dado complexo) propriamente dito. Vale ressaltar que o binômio formado pelo vetor de características e a função de distância dá-se o nome de descritor.

Dentre as características comuns citadas para imagens temos: textura que pode ser definida como padrões visuais com propriedades de homogeneidade que não resultam da presença de apenas uma cor ou densidade e sua utilização é muito comum na área de medicina; cor é a característica mais comum, pois independe do tamanho ou orientação da imagem, é uma propriedade também muito usada no reconhecimento humano; forma utilizada principalmente no rastreamento de objetos em movimento e reconhecimento de formas.

## **2.7 Considerações Finais**

Neste capítulo foram descritos os principais conceitos sobre dados complexos e alguns recursos utilizados para possibilitar que estes dados sejam recuperados em consultas por similaridade.

O capítulo focou na contextualização sobre dados complexos, consultas por conteúdo, espaço métrico que oferece estruturas de indexação para este tipo de dados, como são realizadas as consultas por similaridade possibilitando entender melhor porque sofrem com problemas de gap semântico e as funções de distância com alguns exemplos, por fim os extratores de características que são feitos para extração do conteúdo intrínseco dos dados, resultando nos vetores de características possibilitando assim, junto com as funções de distância, realizar as consultas por similaridade.

Todo conteúdo deste capítulo se faz necessário para o embasamento teórico utilizado neste trabalho.

# Capítulo 3

## TRABALHOS CORRELATOS

---

*Este capítulo apresenta algumas técnicas utilizadas na combinação de características. No início são realizadas definições sobre essa técnica iniciando com sua abordagem básica e seguindo pela combinação de descritores da forma que é usada por este estudo. Na sequência são apresentados e discutidos trabalhos correlatos que utilizam a combinação de múltiplos descritores. Por fim são apresentados os conceitos da utilização de condições de contorno e de que forma ela colabora para a diminuição do gap semântico em consultas por similaridade.*

### 3.1 Estruturação do Capítulo

Este capítulo está estruturado da seguinte forma. A seção 3.2 inicia-se com a definição da abordagem básica da Combinação de Características e a utilização de “super-vetores”, seguido dos conceitos sobre a Combinação de Múltiplos Descritores e o que eles agregam sobre a abordagem básica. Nesta sessão também são apresentados e discutidos os trabalhos correlatos e as diversas formas de combinar múltiplos descritores. Na sessão 3.3 é apresentado o uso das condições de contorno nas consultas por similaridade e de que forma elas agregam a estas consultas com a finalidade de melhorar a qualidade de seus resultados. O capítulo é finalizado na seção 3.4, com as considerações finais.

## 3.2 Combinação de Características

### 3.2.1 Abordagem Básica

Consultas por similaridade em dados complexos podem tratar um único conjunto de características (e.g. em imagens cor, textura ou forma) ou fazer uma combinação dos vetores de características (e.g. em imagens cor e textura, cor e forma, etc.).

Uma maneira trivial para combinar vários conjuntos de características é concatenar todos os vetores extraídos, formando assim um “super-vetor” e a partir disto utilizar alguma função de distância para comparar os dados. Porém é provável que os conjuntos apresentem atributos com valores bem diferentes, o que necessita que esses valores sejam normalizados, evitando qualquer prevalência das características que possuem valores mais altos sobre o resultado.

O aumento da dimensionalidade também torna os espaços multi-dimensionais mais esparsos, ocasionando na “maldição da dimensionalidade” (KORN, PAGEL & FALOUTSOS, 2001) e o desempenho das estruturas de indexação se degradam rapidamente. Esse aumento indiscriminado no número de características ainda prejudica a precisão dos resultados (BUSTOS et al. 2004).

Outro fator relevante é a relação existente entre a característica e a função de distância, onde conjuntos distintos de características que descrevem o mesmo objeto podem obter melhores resultados utilizando funções de distância distintas (BUGATTI, TRAINA & TRAINA JR., 2008) (BUGATTI, 2008). Por isso a utilização de “super-vetores” com uma única função de distância pode deixar de utilizar todo o poder discriminatório de cada característica, diminuindo a possibilidade de alcançar os melhores resultados, pois se utiliza da mesma métrica para comparar diferentes conjuntos de características.

Por estes motivos, ao invés de apenas concatenar os vetores de características, pode-se realizar a combinação de múltiplos descritores, onde cada descritor é composto pela característica e a função de distância.

### 3.2.2 Combinação de Múltiplos Descritores

Neste método é feita a extração de diversas características intrínsecas no formato de vetores, como por exemplo, as propriedades de imagens normalmente são descritas pela cor, textura e forma extraídos a partir de algoritmos de processamento de imagem. Cada um desses vetores deve ser vinculado à função de distância que melhor se aplique àquela característica formando os descritores.

Feito isso, esses descritores são combinados aplicando-se um peso para cada um dos valores calculados por cada descritor. Este balanceamento tende a melhorar a precisão das consultas, sendo fundamental equilibrar adequadamente a participação de cada descritor à necessidade de cada consulta por similaridade.

Esta prática também requer cuidados perante o aumento descontrolado dos vetores de características, onde o mesmo pode degradar tanto o desempenho quanto a precisão dos resultados (BUENO et al., 2009). Isso ocorre porque conforme há um aumento na dimensionalidade, os espaços multidimensionais tendem a ser mais esparsos e a distribuição dos elementos homogênea, este fenômeno é conhecido como “maldição da dimensionalidade”.

Neste cenário as características que melhor poderiam evidenciar os elementos tem seu poder de discriminação diminuído devido ao aumento na quantidade de outras características, esta por sua vez resulta em valores mais parecidos de distâncias entre seus elementos, ou seja a distribuição dos elementos no espaço torna-se cada vez mais homogênea.

Outro ponto a ser levado em consideração é a ponderação entre os pesos atribuídos a cada descritor, cuidando para não superestimar ou subestimar seus valores, causando problemas com o balanceamento e conseqüentemente comprometendo a precisão das consultas.

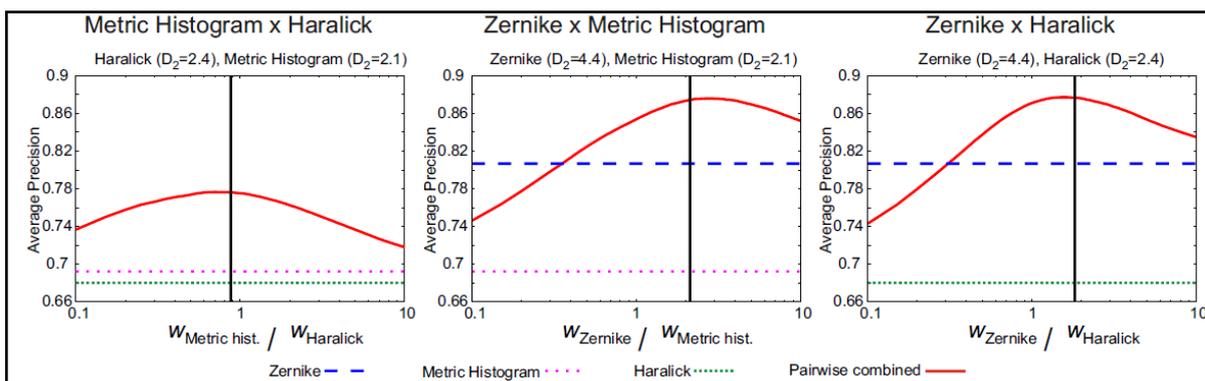
No balanceamento de múltiplos descritores o método comum para agregá-los de forma linear, é dado por: uma imagem  $x$  é representada pelas características  $x_1, \dots, x_n$ , com as métricas  $\delta_1, \dots, \delta_n$ , definidas sobre o domínio dos respectivos conjuntos de características e a natureza das imagens. A composição da função de distância entre a imagem  $x$  e  $y$  é a combinação das distâncias  $\delta_i(x_i, y_i)$  dado por:

$$\delta(x,y) = \sum_{i=1}^n \omega_i \cdot \delta_i(x_i, y_i)$$

, onde  $\omega_j$  é o peso respectivo atribuído a cada descritor.

A utilização deste método permite definir como cada descritor contribui para o cálculo de similaridade.

É possível destacar trabalhos que utilizam a combinação de múltiplos descritores para o conjunto todo. Como (BUENO et al., 2009) que propôs uma técnica de balanceamento não supervisionado para descritores baseado na dimensionalidade intrínseca dos conjuntos de dados. Os resultados podem ser observados na Figura 3.1, que apresenta a variação da precisão média na combinação entre os descritores de acordo com a variação dos pesos atribuídos.



**Figura 3.1 – Gráfico da variação da precisão média na combinação de descritores. (BUENO et al., 2009)**

A linha vertical que corta cada gráfico mostra a estimativa do método proposto para o melhor balanceamento para o conjunto todo. É possível observar que essa precisão não é alcançada pela utilização dos descritores individualmente.

Em (BUSTOS et al. 2004) é proposto um método que calcula dinamicamente o balanceamento entre os descritores para cada consulta, utilizando um conjunto de treinamento com imagens pré-classificadas. Este método faz a combinação de múltiplos descritores de forma não ponderada e ponderada. O que foi mostrado anteriormente em (BUENO et al., 2009), é que a vantagem de ponderar a contribuição de cada descritor nas consultas por similaridade, garante uma maior participação de descritores que melhor caracterizem os elementos do conjunto,

proporcionando assim melhores resultados nas consultas por similaridade. A mesma definição pode ser observada pelo autor, onde os experimentos ponderados alcançaram melhores precisões nas consultas.

Nos trabalhos de (LIU et al., 2007) (LIU, JIA & WANG, 2007) (LIU, JIA & WANG, 2008) é proposto a combinação de descritores de cor, textura e forma em conjuntos de dados de imagens. Para o balanceamento, a técnica utiliza operadores para atribuição adaptativa de pesos baseadas em estatísticas das características extraídas. Através de simulações que exploram essa combinação de múltiplos descritores, são extraídas informações que mostram a significância da participação de cada um nessa combinação e como ela afetou na precisão das consultas por similaridade. Essas informações são utilizadas para adaptar o balanceamento entre os descritores, de forma a utilizar um peso maior ao descritor que melhor colabora para discriminação dos dados.

No estudo de (FANG, HAIMEI & WEI, 2010) é empregado a teoria da análise de correspondência para integrar os recursos de imagem. Neste trabalho são feitas transformações nos elementos de dados onde é realizada uma extração de diversas características intrínsecas dos elementos além das comumente extraídas (cor, textura e forma), tais como, características relacionadas a pontos, linhas e traços das imagens objetivando obter recursos mais precisos da imagem ou que colaborem mais na sua representação. A consulta por similaridade então equipara cada uma dessas transformações formando uma matriz de resultados, essas informações são usadas para fazer o ranking dos elementos.

Em (da SILVA TORRES et al., 2009) (CHEN et al.,2010) são apresentados métodos para combinação de múltiplos descritores que utilizam programação genética (GP) e algoritmos evolutivos. Neste tipo de combinação é realizada uma evolução da métrica utilizada. Esses algoritmos evolutivos são chamados assim porque se comportam de uma forma parecida a genética dos seres vivos, onde as características positivas prevalecem e são mantidas adiante para as próximas gerações. Assim como na genética, estas métricas (chamadas de elementos) vão sofrendo mutações e cruzamentos entre si e as características das funções que se mostrarem mais eficientes prevalecem para as métricas resultantes. Todo esse processo de evolução é realizado para aperfeiçoar a métrica de consulta “mais evoluída” ou em outras palavras, que apresente um melhor poder de discriminação dos dados e conseqüentemente um melhor grau de precisão para as consultas por

similaridade. Dentre as vantagens desta técnica que se utiliza de algoritmos evolutivos é a possibilidade de utilizar funções de balanceamento mais complexas do que a simples combinação linear de descritores, o que pode proporcionar uma melhor discriminação dos dados.

Nas pesquisas de (WANG & ZHU, 2010) (RUI et al., 1998) (ZHUANG & WANG, 2010) são propostos métodos flexíveis, que possuem algoritmos para recálculo dos pesos utilizando realimentação por relevância de acordo com preferências do usuário para encontrar um melhor balanceamento entre os descritores. As técnicas de realimentação por relevância proporcionam interações entre as consultas e o usuário do sistema, onde a cada iteração o usuário indica em uma amostra de dados quais estão mais próximos do esperado. Uma das maiores vantagens apresentadas por esta técnica é justamente a sua forma adaptativa, possibilitando que o algoritmo de busca aprimore a métrica e refaça a pesquisa que tende a obter resultados cada vez mais próximos à expectativa do usuário.

Em (FERREIRA et al., 2011) (SINGH & KOTHARI, 2003) são combinadas as duas estratégias acima. Utiliza-se realimentação por relevância para evoluir os elementos (as métricas) através de algoritmos evolutivos de programação genética. As iterações com o usuário do sistema ocorrem pelo processo de realimentação por relevância, o que possibilita que a métrica seja aprimorada de acordo com as necessidades do usuário.

### **3.3 Condições de Contorno**

A aplicação das condições de contorno foi estudada por (PONCIANO-SILVA, 2009) (PONCIANO-SILVA et al., 2009), onde em seu estudo o autor adicionou mais um parâmetro em suas consultas por similaridade, chamado de parâmetro perceptual. Como o próprio autor define “é um indicador de maior relevância das características visuais de alto nível da imagem durante o processo de reconhecimento e comparação realizado pelo especialista”. Este parâmetro tende a expressar padrões ou características do elemento, e podem assim determinar quais características devem ter maior relevância na busca e comparação com outras imagens.

Na pesquisa o autor utiliza como parâmetro perceptual o “achado” radiológico de imagens médicas de exames radiológicos que analisa: lesões, anormalidades e alterações na estrutura normal de um tecido. Seriam essas as características reconhecidas visualmente na imagem de maior relevância perceptual em uma determinada situação diagnóstica (PONCIANO-SILVA et al., 2009).

O trabalho utilizou a análise de exames torácicos por tomografia computadorizada, com foco em análise de lesões pulmonares. As imagens foram classificadas em seis classes distintas, delimitadas pelas condições de contorno (achados radiológicos) presentes nas imagens. Estes subconjuntos foram: Consolidação, Enfisema, Vidro Fosco, Favo de Mel, Espessamento e Normal. E cada imagem foi classificada exclusivamente em uma das classes.

Foram identificados os relacionamentos entre descritores individuais e as classes, delimitadas pelas condições de contorno, e definidos os descritores que melhor representavam as características visuais dos elementos de cada classe. Em experimentos, o autor mostrou que para cada subconjunto de dados (classes), existia um determinado descritor que apresentava melhores resultados. A partir disto, de acordo com a condição de contorno de cada elemento de consulta é determinado qual o descritor que deve ser usado na consulta por similaridade.

O estudo destaca a contribuição que a utilização dessa “tríade” (vetor de característica, função de distância e parâmetro visual), diminui o gap semântico e melhorou a precisão de consultas por similaridade.

Portanto, podemos concluir sobre este trabalho, a importância que cada característica tem no papel de representação dos dados e o poder de cada uma delas para melhor evidenciar uma determinada classe de elementos do conjunto. Nem sempre uma característica muito relevante para expressar atributos visuais de uma determinada classe tem o mesmo poder de discriminação para evidenciar outra do restante do conjunto.

A relevância com que cada característica irá contribuir na consulta por similaridade deve ser analisada individualmente para cada uma das classes pois cada característica colabora de forma diferente, fazendo-se necessário serem ponderadas de acordo com o poder de representação que cada uma possui para a classe que se deseja evidenciar do restante do conjunto.

Vale ressaltar que os critérios utilizados como condição de contorno não necessariamente devem ser o mesmo que classifica a imagem. Um exemplo

simples: pode ser utilizado o sexo do paciente como condição de contorno, enquanto que as imagens utilizadas foram classificadas de acordo com a doença identificada na imagem. Neste caso, a combinação de múltiplos descritores usado seria o que melhor realça a doença que deve ser encontrada (classes) para cada um dos subconjuntos de imagens identificadas pela condição de contorno (homens e mulheres).

### **3.4 Considerações Finais**

Neste capítulo foram descritos vários métodos para combinação de características e descritores. No decorrer do capítulo foi possível esclarecer o quanto a combinação de descritores colabora com a diminuição do gap semântico em consultas por similaridade.

Foram descritas formas de utilizar a combinação de múltiplos descritores, seja pelo recálculo dinâmico dos pesos em combinações lineares ou utilizando técnicas de realimentação por relevância e algoritmos adaptativos, que aperfeiçoa a métrica utilizada com as necessidades do utilizador do sistema.

Também foi discutida a utilização de condições de contorno em consultas por similaridade. Elas proporcionam a possibilidade de utilizar uma métrica que ressalta melhor um grupo de elementos que possuam um comportamento em comum nas consultas e com isso é possível diminuir o gap semântico.

No próximo capítulo será apresentado método proposto neste trabalho: utilização de condições de contorno para delimitação de subconjuntos e definição do melhor balanceamento para cada subconjunto.

# Capítulo 4

## COMBINAÇÃO DE MÚLTIPLOS DESCRITORES UTILIZANDO CONDIÇÕES DE CONTORNO

---

*Este capítulo apresenta o método proposto que utiliza condições de contorno para determinar a combinação de múltiplos descritores. Neste é apresentada a metodologia e a arquitetura referente à proposta deste trabalho.*

### 4.1 Estruturação do Capítulo

Este capítulo está estruturado da seguinte forma. A sessão 4.2 apresenta o detalhamento do trabalho de pesquisa, os conceitos do método proposto e de que forma ele pretende contribuir para melhorar a precisão de consultas por similaridade. Posteriormente, na sessão 4.3 discorre sobre a metodologia utilizada. Na sessão 4.4 é explicado em detalhes toda metodologia e arquitetura do método. Por fim, a sessão 4.5 traz as considerações finais do capítulo e uma comparação do método proposto com demais práticas apresentadas anteriormente: utilização de descritores individuais, utilização do melhor descritor por condição de contorno (a tríade) (PONCIANO, et al., 2009) e utilização da melhor combinação de descritores para o conjunto todo.

## 4.2 Introdução

Este trabalho tem como objetivo propor maneiras de melhorar a precisão de consultas por similaridade aplicando combinação de múltiplos descritores, com balanceamentos específicos para subconjuntos delimitados do conjunto de dados, definidos por condições de contorno.

A combinação de múltiplos descritores mostra-se com grande capacidade de melhorar os resultados de consultas por similaridade, sendo que tal fato é explicado pelas características serem complementares ao fazer a identificação de atributos visuais das imagens, de forma similar ao comportamento humano para sua interpretação.

Tal fato pode ser explicado pela forma como um ser humano visualiza e interpreta uma imagem. Ao olhar para uma imagem, o comportamento humano não utiliza apenas as cores da imagem, ou o formato, ou utiliza exclusivamente uma de suas características individualmente, e sim todas as características da imagem são utilizadas para interpreta-la.

Experimentos realizados em (BUENO et al., 2011) apresentaram a precisão dos resultados das consultas por similaridade com combinação de características (Zernike e Haralick) variando os pesos de cada descritor. A Figura 4.1, apresenta o gráfico resultante destes experimentos realizados com uma coleção de dados composta por 704 imagens de exames de ressonância magnética (MRI\_704). Neste conjunto as imagens foram classificadas em 40 classes distintas de acordo com a região do corpo, plano de digitalização e posição do corte, conjunto este também utilizado nos experimentos do presente trabalho.

O gráfico inicia-se mais a esquerda com proporção entre os descritores de 10 para Haralick e 1 para Zernike. A diferença na proporção entre os pesos diminui até chegarem ao centro do gráfico, onde a proporção do peso é de 1 para 1 entre os dois descritores percorrendo até a parte mais a direita do gráfico a proporção dos pesos de Zernike aumenta até chegar a 10 vezes maior que Haralick .

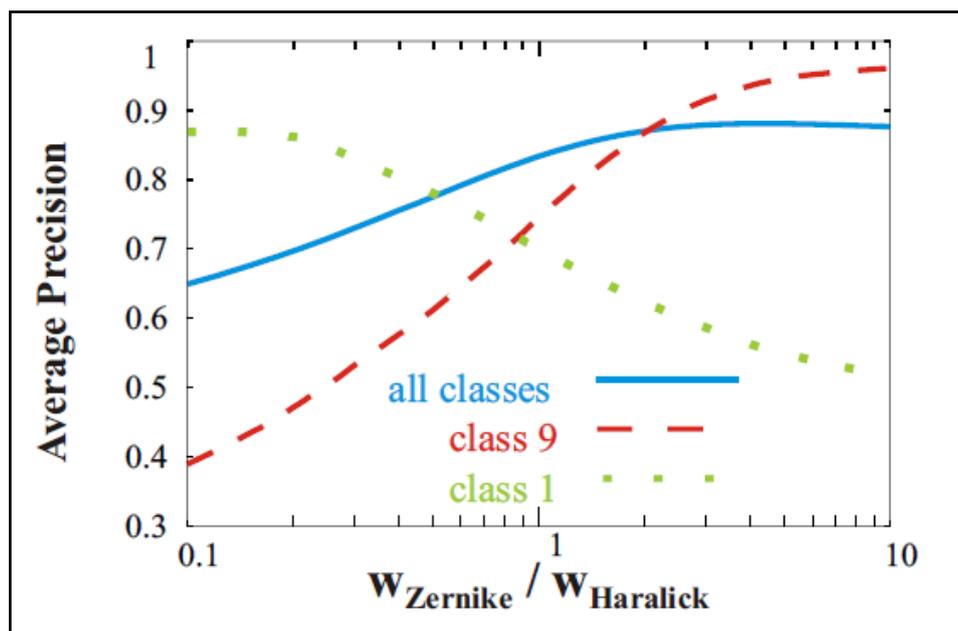


Figura 4.1 – Gráfico da média de precisão do conjunto de dados MRI\_704. (BUENO et al.,2011)

O gráfico possibilita observar a influência da proporção dos pesos de cada descritor no resultado da consulta. O melhor resultado obtido, na média para o conjunto todo, aparece quando temos aproximadamente a proporção do triplo de Zernike sobre Haralick.

Já para consultas sobre elementos da “Classe 1”, os melhores resultados acontecem quando peso de Haralick é aproximadamente 10 vezes maior que o peso de Zernike. Porém, para elementos da “Classe 9”, a melhor proporção variou aproximadamente entre 5 e 10 vezes o peso para Zernike sobre Haralick.

O gráfico ainda mostrou que subconjuntos diferentes são melhor identificados por balanceamentos diferentes. Se fosse utilizado apenas um balanceamento para o conjunto todo, não seriam alcançadas as melhores precisões nas consultas para elementos da “Classe 1” e da “Classe 9”, obtidos com o balanceamento específico para cada uma delas. Além disso, se fosse utilizado um balanceamento ideal para o conjunto como um todo, os resultados para a “Classe 1” seriam ainda mais comprometidos, pois a curva das linhas no gráfico vão em direções opostas.

Quanto às condições acima descritas foi possível indicar uma possível melhoria oferecida pela delimitação de subconjuntos, cada qual com um balanceamento específico, motivando a utilização de condições de contorno para melhorar as consultas por similaridade.

Imagine por exemplo em um banco de imagens médicas, pode ser utilizada a hipótese de diagnóstico como condições de contorno para delimitar esses subconjuntos. Em muitos casos a hipótese de diagnóstico já vem definida no pedido do exame, se tratando de algo que levou o médico a solicitá-lo a partir do que ele identificou na primeira investigação ou pelo histórico do paciente.

Um exemplo deste cenário permite considerar que um paciente chega com queixas que levam o médico a identificar possíveis sintomas de pneumonia, solicita um raio-x do pulmão para verificar e constatar se a hipótese lançada inicialmente é comprovada pelos exames. Neste caso, a hipótese de diagnóstico inicial de “pneumonia” já acompanha o pedido de exame. Essa hipótese pode ser utilizada como uma condição de contorno, visando definir o melhor balanceamento para evidenciar imagens similares, ou seja, imagens de exames com pneumonia.

Portanto, pretende-se vincular uma condição de contorno a um balanceamento otimizado de múltiplos descritores para cada subconjunto de dados delimitado por tal condição. Neste mesmo banco de imagens poderiam existir outros subconjuntos (p.e. “carcinoma”, “enfisema pulmonar”, etc.) identificados por cada uma das hipóteses de diagnóstico (condição de contorno) onde cada subconjunto estaria vinculado a um balanceamento ideal de múltiplos descritores e desta forma poderia ser utilizado no momento da consulta por similaridade.

Vale ressaltar novamente que as condições de contorno não necessariamente devem estar relacionadas às classes dos elementos, e sim a subconjuntos que possuem comportamento semelhante nas consultas por similaridade. Assim como no exemplo visto anteriormente, onde foi utilizado o sexo do paciente como condição de contorno, enquanto que as imagens utilizadas foram classificadas de acordo com a doença identificada na imagem. Tal configuração resultaria em uma combinação de múltiplos descritores que melhor realça a doença que deve ser encontrada (classes) para cada um dos subconjuntos de imagens identificadas pela condição de contorno (homens e mulheres), em outras palavras, resultará em uma melhor combinação de descritores para cada classe/condição de contorno.

### 4.3 Metodologia

A metodologia utilizada por este trabalho incluiu a constante pesquisa sobre o estado da arte em consultas por similaridade, CBIR e combinação de múltiplos descritores. Foram realizados diversos estudos teóricos sobre esses temas por meio da leitura de livros, dissertações de mestrado, teses de doutorado e principalmente artigos e revistas disponibilizados pelas bibliotecas digitais IEEE, ACM, Scopus, Springer, DBLP, Google Acadêmico e Microsoft Academic.

Esses estudos ofereceram a fundamentação para o desenvolvimento das atividades propostas. Ademais, foram realizadas reuniões periódicas com o orientador, além da participação e realização de seminários nas reuniões periódicas do Grupo de Banco de Dados da UFSCar, as quais permitiram a realização de discussões a respeito das soluções propostas para este projeto de mestrado.

Primeiramente, foi realizado um estudo teórico de todos os conceitos fundamentais relacionados a consultas por similaridade, sendo elas: arquiteturas e métodos, recuperação por conteúdo, CBIR, vetores de características, funções de distância e métodos de acesso métrico. Na sequência foram estudados os problemas enfrentados por este tipo de consulta em relação ao gap semântico. Posteriormente, os estudos ficaram focados em métodos que visavam à diminuição do gap semântico como as combinações de múltiplos descritores e a utilização de condições de contorno, técnicas estas utilizadas por este trabalho por se mostrarem eficientes na melhora da precisão em consultas por similaridade.

Finalizando estes estudos, com o método proposto pelo trabalho definido, foram focados os experimentos, realizando a combinação de descritores em três bases de dados de imagens médicas disponibilizados pelo Centro de Ciência das Imagens e Física Médica (CCIFM) do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto – USP (HCFMRP). As bases de imagens foram previamente classificadas por especialistas e foram utilizadas como condições de contorno nos experimentos, que contaram com esta informação também para validação dos resultados obtidos nos experimentos.

Da execução dos experimentos, foram realizados os balanceamentos entre descritores por meio de algoritmos exaustivos que fazem a combinação linear visando à obtenção dos melhores resultados possíveis.

Com os resultados dos experimentos em mãos, os mesmos foram comparados com três práticas da literatura: os descritores utilizados individualmente (apenas um descritor para o conjunto todo), o melhor descritor para cada condição de contorno (tríade), e a combinação de descritores para o conjunto todo (BUENO et al. 2009). Para avaliação dos resultados, foram utilizadas curvas de precisão e revocação das consultas (BAEZA-YATES & RIBEIRO-NETO, 1999) um dos métodos mais eficientes para avaliação de sistemas de busca (MÜLLER et al., 2004).

Na implementação dos algoritmos, foi utilizado à linguagem de programação C++ e os resultados foram plotados em gráficos e tabelas para melhor visualização e análise.

#### **4.4 Combinação de Múltiplos Descritores Utilizando Condições de Contorno**

Vários estudos se baseiam em técnicas para combinação de múltiplos descritores, que tem se mostrado muito eficiente para melhorar a precisão de resultados de consultas por similaridade (BUENO et al., 2009) (BUENO et al., 2011) (BUSTOS et al. 2004) (CHEN et al., 2010) (da SILVA TORRES et al., 2009) (FANG, HAIMEI & WEI, 2010) (FERREIRA et al., 2011) (LIU, JIA & WANG, 2008) (RUI et al., 1998) (SINGH & KOTHARI, 2003) (WANG & ZHU, 2010) (ZHUANG & WANG, 2010). Além disso, os estudos de (PONCIANO-SILVA, 2009) (PONCIANO-SILVA et al., 2009) mostraram que também é possível aumentar a precisão dos resultados de consultas por similaridade através de condições de contorno.

Este trabalho propõe o uso de condições de contorno para delimitar subconjuntos de imagens e assim determinar o balanceamento de múltiplos descritores de forma a usar combinação ideal para cada uma desses subconjuntos de imagens nas consultas por similaridade.

O método se inicia com a definição das condições de contorno, onde cabe ao especialista no domínio da aplicação defini-las, considerando sempre o objetivo das consultas no conjunto de dados. Após definidas, estas são aplicadas ao conjunto de imagens para delimitar subconjuntos que tem comportamento semelhante em consultas por similaridade.

Tomando como exemplo as aplicações de imagens médicas, as condições de contorno podem ser definidas através de dados associados, tais como, a hipótese de diagnóstico determinado pelo médico ao solicitar o exame ou até mesmo partindo de uma análise inicial das imagens fornecidas por um especialista, que pode pré-classificá-las com base em alterações encontradas (achado radiológico).

Com as imagens pré-classificadas de acordo com as condições de contorno definidas pelo especialista do domínio, é realizada a combinação de múltiplos descritores em uma fase de treinamento. São realizados experimentos para combinar descritores variando a razão entre seus pesos (um algoritmo de combinação linear simples) e a partir disto, é possível encontrar o balanceamento ideal para cada condição de contorno.

Nos experimentos apresentados a combinação de descritores é feita de forma linear exaustiva, onde a distância final entre os elementos é calculada através da soma ponderada de cada uma das distâncias, onde as mesmas devem ser divididas pela sua distância máxima correspondente para que haja uma normalização entre a contribuição de cada descritor.

Este método é dado por: uma imagem  $x$  é representada pelas características  $x_1, \dots, x_n$ , com as métricas  $\delta_1, \dots, \delta_n$ , definidas sobre o domínio dos respectivos conjuntos de características. A composição da função de distância entre a imagem  $x$  e  $y$  é a combinação das distâncias  $\delta_i(x_i, y_i)$  dado por:

$$\delta(x,y) = \sum_{i=1}^n \omega_i \cdot \frac{\delta_i(x_i, y_i)}{dmax_i}$$

, onde  $\omega_i$  é o peso respectivo atribuído ao descritor e  $dmax_i$  representa a distância máxima feita pela comparação entre todos os elementos do conjunto e é utilizado para equilibrar a participação de cada descritor ao representar o elemento.

Vale ressaltar que o balanceamento de múltiplos descritores pode ser realizado com outras técnicas de combinação como por exemplo a utilização de programação genética (da SILVA TORRES et al., 2009) (CHEN et al., 2010) e/ou algoritmos de realimentação por relevância (FERREIRA et al., 2011) (RUI et al., 1998) (SINGH & KOTHARI, 2003) (WANG & ZHU, 2010) (ZHUANG & WANG, 2010).

Os passos realizados pelo método proposto são descritos na Figura 4.2. Na figura são ilustrados os passos realizados pela fase de treinamento. No quadro a esquerda da imagem destacado como “PASSO – 1” é o primeiro estágio onde o especialista prepara uma amostragem de imagens pré-classificadas para serem utilizadas de acordo com cada uma das condições de contorno definidas de acordo com a necessidade do especialista. No quadro destacado como “PASSO – 2” são aplicados algoritmos de extração de características nos elementos e estes carregados para uma base de dados. Já com a base de dados alimentada com as características extraídas dos elementos, cada qual vinculada a uma condição de contorno, é realizado o processamento de algoritmos de combinação de múltiplos descritores, destacado na imagem como “PASSO – 3” (quadro mais a direita da figura), objetivando encontrar o melhor balanceamento entre eles para cada uma das condições de contorno.

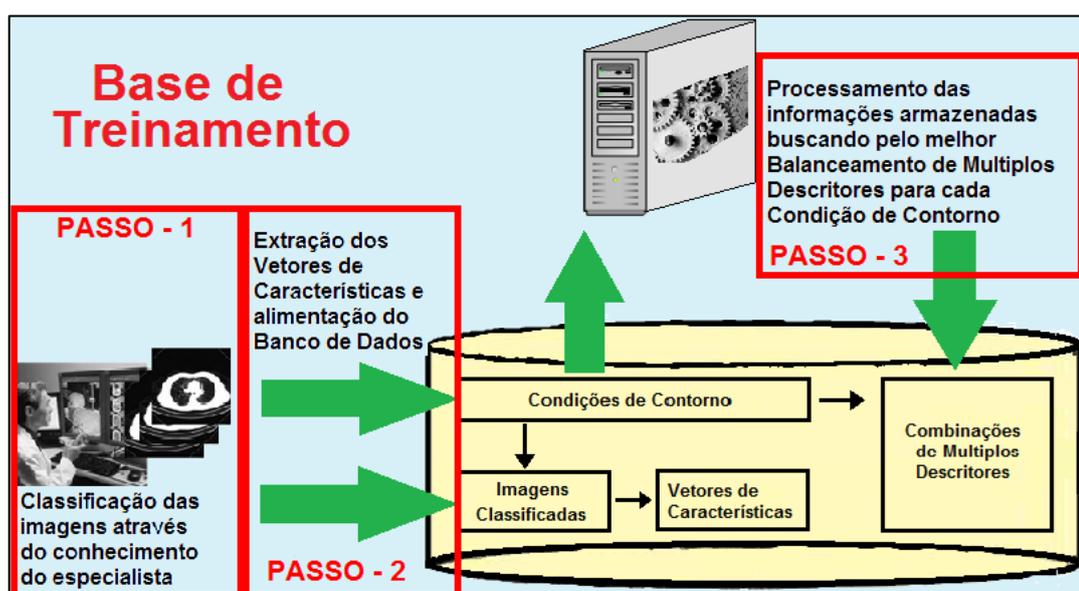


Figura 4.2 – Arquitetura do ambiente de treinamento do método proposto.

Esta primeira fase resulta basicamente em uma tabela  $T$  que relaciona o balanceamento ideal de múltiplos descritores ( $CMD$ ) a cada uma das Condições de Contorno ( $CC$ ) do conjunto, onde temos os elementos  $T = \{(CC_1, CMD_1), \dots, (CC_i, CMD_i)\}$ , onde  $i$  representa a quantidade de condições de contorno.

Com estes resultados gerados, parte-se para o ambiente de utilização do método proposto. Esta arquitetura pode ser visualizada na Figura 4.3.

A figura se inicia no quadro mais a esquerda destacado como “PASSO – 1”. Neste ponto é realizada a importação da tabela resultante da fase de treinamento. Esta tabela se torna o recurso mais importante a ser utilizado pelo método por relacionar o balanceamento de múltiplos descritores ideal a cada condição de contorno, o que possibilita utilizar sempre este balanceamento nas consultas de similaridade.

Seguindo sua estrutura, segue para o quadro destacado como “PASSO 2” da figura, esta etapa esta relacionada com o inicio da consulta por similaridade. Neste devem ser passadas como parâmetros para a consulta a imagem de referência para e a condição de contorno, sendo esta informação fundamental pois é através dela que será recuperado o melhor balanceamento de descritores para ser utilizado pela consulta por similaridade.

Neste cenário a condição de contorno é utilizada então para identificar (ou estimar) a qual subconjunto esta imagem de referência pertence e então utilizar a combinação de múltiplos descritores que melhor o evidencie do restante do conjunto e a partir disto utilizar esta combinação na consulta por similaridade.

Tal operação corresponde ao quadro destacado como “PASSO – 3” da figura. Uma vez selecionado a melhor combinação de múltiplos descritores ela pode ser utilizada na consulta por similaridade, que é executada e então os elementos são ranqueados e retornados como resultado.

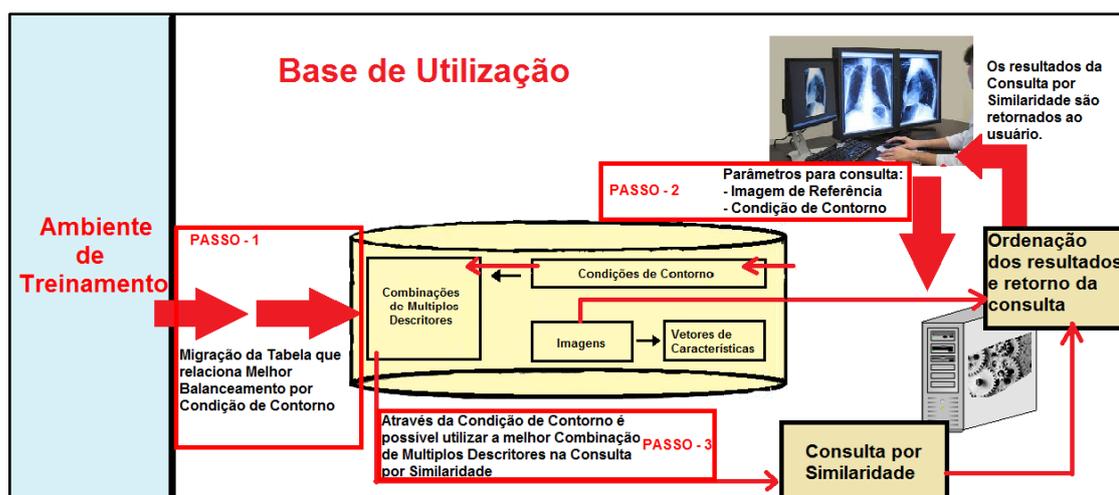


Figura 4.3 – Arquitetura do ambiente de utilização do método proposto.

## 4.5 Considerações Finais

Neste capítulo foi reforçada a importância que as condições de contorno têm em consultas por similaridade que juntamente com a combinação de descritores, tende a diminuir ainda mais a lacuna semântica enfrentada pelas consultas por similaridade.

Portanto, unir estas duas técnicas possibilita utilizar o balanceamento de múltiplos descritores ideal para cada subconjunto elementos do conjunto de dados. No próximo capítulo serão apresentados os experimentos realizados neste trabalho que indicam que esta hipótese é verdadeira.

Vale ressaltar também que para validação do método proposto durante a fase de treinamento, foi utilizado um algoritmo de combinação linear simples, para uma busca exaustiva onde a melhor combinação de descritores pôde ser executada. Apesar disso, o método proposto pode ser usado com outros método de combinação de descritores.

Estruturalmente o método é constituído de duas fases, sendo a primeira de treinamento, onde algoritmos de combinação de descritores resultam em uma relação condição de contorno X combinação de múltiplos descritores ideal. E a segunda fase já em ambiente de utilização, onde ao executar as consultas por similaridade é necessário passar como parâmetro, além da imagem de referência, a condição de contorno possibilitando utilizar a combinação de múltiplos descritores ideal na consulta por similaridade.

# Capítulo 5

## RESULTADOS

---

*Este capítulo descreve os experimentos realizados no presente trabalho. Serão apresentados os conjuntos de dados utilizados e os resultados obtidos pela técnica proposta de combinação de múltiplos descritores utilizando condições de contorno. Posteriormente, o método é comparado a outras técnicas da literatura que se apresentaram eficientes na diminuição do gap semântico em consultas por similaridade. Mediante isto foi possível comprovar a eficácia da técnica proposta que apresentou melhores resultados em todos os experimentos realizados. Por fim é proposto uma nova forma de combinação de descritores no objetivo de tornar os processos de treinamento menos morosos.*

### 5.1 Estruturação do Capítulo

Este capítulo está estruturado da seguinte forma. A seção 5.2 apresenta os experimentos realizados, compartilhando as experiências e os desafios encontrados. A cada experimento foram comparados os resultados obtidos com outros métodos da literatura, como a: utilização de apenas um descritor individual, utilização do melhor descritor individual por classe (tríade) (PONCIANO-SILVA, 2009) (PONCIANO-SILVA et al., 2009) e a utilização de uma combinação de descritores para o conjunto todo. O último experimento em particular, não tem a intenção de validar o método proposto nem de compará-lo com outras práticas da literatura, mas sim propor outra forma para combinação de múltiplos descritores chamado aqui de “Combinando Descritores a Partir do Melhor Descritor”. Finalmente na seção 5.3 são discutidos os resultados obtidos e a validação do método proposto que obteve melhor precisão em todos os experimentos realizados.

## 5.2 Experimentos

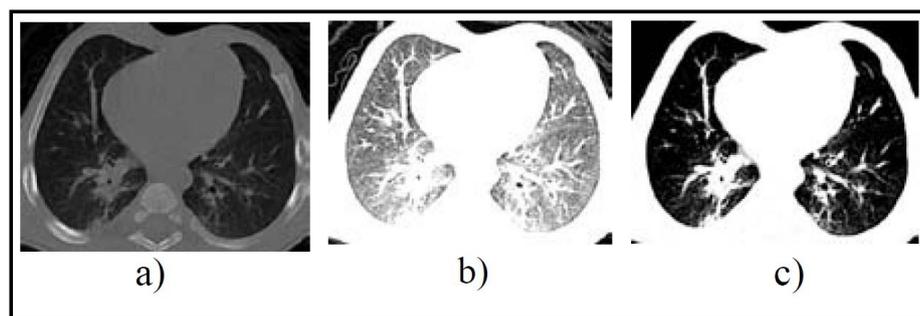
Para a realização dos experimentos foram utilizados três conjuntos de dados de imagens médicas cedidas pelo Centro de Ciência das Imagens e Física Médica (CCIFM) do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Brasil.

O primeiro conjunto chamado de TC\_Pulmão é uma coleção de imagens de tomografia pulmonar composto de 247 imagens, classificadas pelos especialistas e separadas em classes de acordo com a enfermidade de cada uma. Foram assim separadas em 6 subconjuntos com aproximadamente 40 imagens cada, classificadas em: enfisema, consolidação, opacidade em vidro fosco, espessamento do septo interlobular, faveolamento e normal. O grupo de imagens da classe “normal” não foi utilizado como centro de pesquisa nos experimentos e as demais 5 classes definidas foram utilizadas como condição de contorno nos experimentos.

Das imagens relacionadas foram extraídas as seguintes características: Histograma Tradicional que representa a cor e Haralick para representar a textura.

As imagens ainda foram submetidas a operação de “janelamento” sendo este um método espacial de processamento de imagens muito utilizado pelos especialistas na visualização de imagens médicas por obter um contraste maior entre determinados valores de níveis de cinza (PONCIANO-SILVA, 2009). Tal método é também conhecido como “realce por alargamento de contraste”.

Foram extraídos dois “janelamentos” do histograma pelo motivo de cada uma evidenciar melhor uma característica da imagem, sendo eles: Histograma Baixo utilizado para realçar melhor regiões mais brilhantes e Histograma Alto utilizado para realçar regiões mais escuras, como é possível visualizar na Figura 5.1.



**Figura 5.1– a) Imagem Original, b) Imagem resultante do janelamento para Histograma Baixo– realça regiões mais brilhantes, c) Imagem resultante do janelamento para Histograma Alto – realça regiões mais escuras (PONCIANO-SILVA et al., 2009).**

O segundo conjunto de imagens chamado de ROI\_TCPulmão é uma coleção de 3.257 imagens de regiões de interesse (Regions of Interest – ROIs) das imagens do primeiro conjunto, resultante em imagens de tamanho de 64x64 pixels e 256 níveis de cinza. O conjunto também foi dividido em 6 subconjuntos onde um deles apresenta imagens normais e as demais são compostas por imagens que contém diferentes padrões de anomalias que representaram as condições de contorno (neste experimento também foi ignorada a classe de imagens “normal” como referência de pesquisa).

Dessas imagens foram extraídas as características de forma (primeiros 256 momentos de Zernike), histograma referente aos níveis de cinza e Haralick (HARALICK, SHANMUGAM & DINSTEIN, 1973) (HARALICK, 1979) para textura que teve seus descritores de variância, entropia, energia, homogeneidade, momento da 3ª ordem, variância inversa e *step* combinados resultando em um vetor de características de 140 posições.

Para esta coleção foram realizados experimentos em duas fases, onde foram eleitas aleatoriamente 100 imagens de cada classe para a fase de treinamento, resultando em uma base composta de 600 imagens. Posteriormente as métricas obtidas nessa fase de treinamento (com 600 imagens) foram aplicadas ao conjunto todo (com as 3.257 imagens), para analisar como os balanceamentos se comportariam para um conjunto sendo calculadas apenas por uma amostragem.

Outro conjunto de imagens utilizado foi chamado de MRI\_704 e trata-se de um conjunto de imagens composto por 704 imagens de exames de Ressonância Magnética (IRM) e Angiogramas. As imagens foram separadas em 40 classes de

acordo com a região do corpo, plano de digitalização e posição do corte como ilustrado na Figura 5.2.

Dessas 40 classes foram agrupadas em 8 Condições de Contorno de acordo com o tipo do exame, são eles: Angiograma, Abdômen Axial, Abdômen Coronal, Cabeça Axial, Cabeça Coronal, Cabeça Sagital, Pélvis Axial e Espinha Sagital.

Vale reforçar que neste caso as classes de imagens não foram utilizadas como condições de contorno. Neste caso, a combinação de múltiplos descritores usado seria o que melhor realça os elementos para cada tipo de exame.

As imagens apresentam tamanho de 256x256 pixels, tiveram a profundidade da cor reduzida a 8 bits, resultando em 256 níveis de cinza. Das imagens foram extraídas as seguintes características: Histograma para representar os níveis de cinza, 256 primeiros momentos de Zernike representando forma e Haralick para as características de textura com os descritores de variância, entropia, energia, homogeneidade, momento da 3ª ordem, variância inversa e *step* combinados resultando em um vetor de características de 140 posições como no experimento anterior.

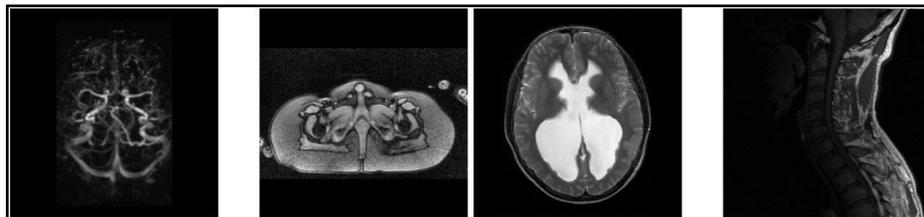


Figura 5.2 – Amostragem das Classes de Imagens de Exames de Ressonância Magnética.

Os detalhes dos conjuntos de características extraídas que foram utilizados nos experimentos podem ser visualizados na tabela 5.1.

Tabela 5.1 Conjunto de Características Usadas nos experimentos

Conjunto	Elementos	Classes	Condições de Contorno	Característica	Dimensão
TC_Pulmão	247	6	5	Histograma Alto	32
				Histograma Baixo	32
				Histograma Trad.	256
				Haralick	26
ROI_TCPulmão	3.257	6	5	Histograma	256
				Zernike	256
				Haralick	140

MRI_704	704	40	8	Histograma	256
				Zernike	256
				Haralick	140

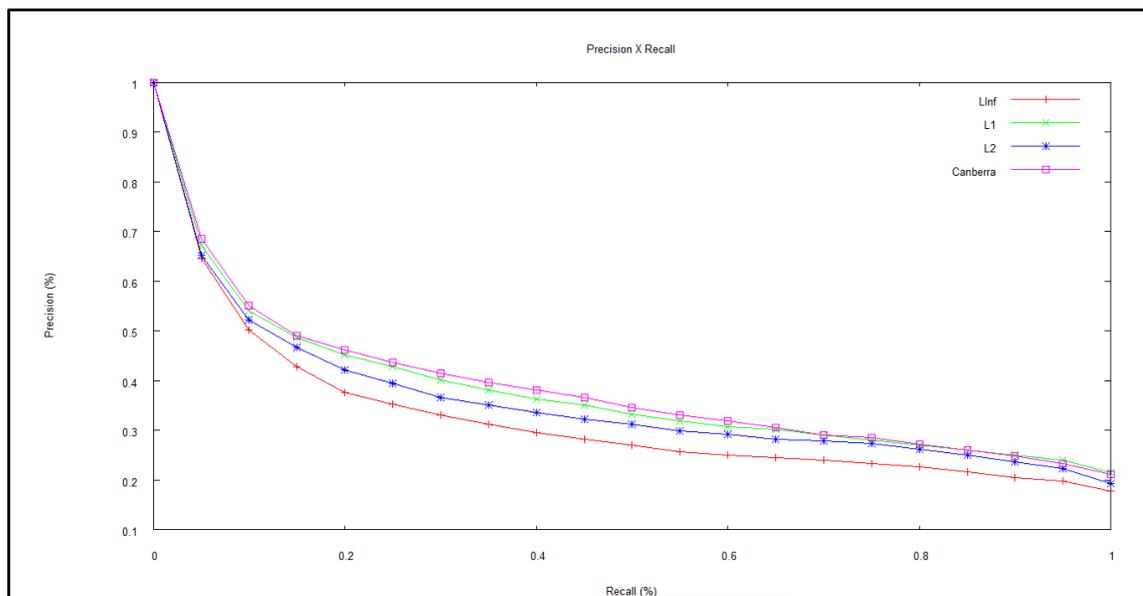
As funções de distância utilizados nos experimentos compreendem as funções da família Minkowski (função de distância L<sub>inf</sub> (Infinity ou Chebychev), função de distância L1 (City Block ou Manhattan) e função de distância L2 (Euclidiana)) e a função de distância Canberra, sendo estas as funções de distâncias utilizadas pelos estudos utilizados nas comparação do presente trabalho.

### 5.2.1 Definição de descritores

Em uma fase que precede ao balanceamento foram definidos os descritores.

Com os vetores de características extraídos devem ser aplicadas as funções de distância a serem utilizadas. Os descritores podem ser definidos por um especialista do domínio das imagens, ou ser calculada através da precisão média obtida por estas consultas, sendo então possível constatar qual função de distância melhor representa cada característica e assim usá-las em conjunto.

Utilizando a base de dados do primeiro experimento, é possível ilustrar como este processo é realizado. Tomando por exemplo a característica textura e realizando consultas por similaridade com todas as funções de distância obtêm-se as seguintes curvas de precisão (BAEZA-YATES & RIBEIRO-NETO, 1999) ilustradas na Figura 5.3.



**Figura 5.3 – Curvas de Precisão x Revocação de consultas TC\_Pulmão utilizando Textura com as Funções de Distância Minkowski e Canberra.**

A figura compreende um gráfico com curvas de precisão e revocação, calculadas através de experimentos realizados neste exemplo com o vetor de características textura, onde o mesmo foi utilizado para consultas por similaridade utilizando as funções de distância da família Minkowski e Canberra através de algoritmos exaustivos, para todos os elementos do conjunto.

A partir das consultas foram obtidas as precisões médias para textura de 39,45% utilizando a função de distância Canberra, 38,77% para  $L_1$ , 36,82 para  $L_2$  e 33,55% para  $L_{Inf}$ .

Através dessas médias foi possível definir a distância Canberra para representar a característica textura, formando então o descritor Textura com Canberra.

E assim são encontrados para cada característica os descritores a serem utilizados na combinação.

Para todos os experimentos, inicialmente foram utilizadas o cruzamento de todas as funções de distância citadas acima com todas as características extraídas para encontrar qual delas obtinha a melhor precisão por característica, definindo então os descritores utilizados pelos experimentos de combinação na fase seguinte. Os descritores resultantes para cada conjunto podem ser visualizados na tabela 5.2.

Tabela 5.2 Definição dos Descritores

Conjunto	Característica	Função de Distância
TC_Pulmão	Histograma Alto (HÁ)	L2
TC_Pulmão	Histograma Baixo (HB)	L2
TC_Pulmão	Histograma Trad. (HT)	LInf
TC_Pulmão	Haralick (Tex)	Canberra
ROI_TCPulmão	Histograma Trad. (HT)	L1
ROI_TCPulmão	Zernike	Canberra
ROI_TCPulmão	Haralick (Tex)	Canberra
MRI_704	Histograma Trad. (HT)	L1
MRI_704	Zernike	L2
MRI_704	Haralick (Tex)	Canberra

### 5.2.2 Conjunto de Imagens de Tomografia Computadorizada de Pulmão (TC\_Pulmão)

Neste primeiro experimento foi realizada a combinação de múltiplos descritores utilizando 5 condições de contorno das 6 classes pré-definidas de imagens. Optou-se por não utilizar a classe “normal” como condição de contorno, devido a mesma não apresentar uma característica visual que pudesse evidenciá-la do restante do conjunto, influenciando assim no resultado dos experimentos. Porém os elementos dessa classe continuaram presentes no conjunto de dados e nas consultas por similaridade.

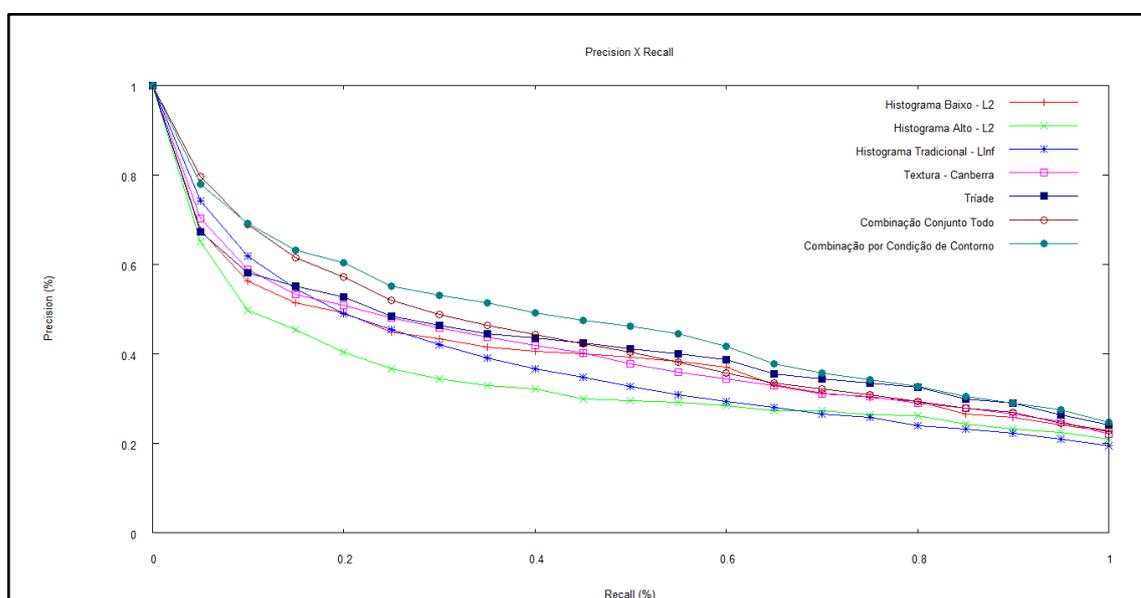
O experimento foi realizado de duas formas, primeiramente avaliou-se a precisão média considerando todos os níveis de revocação para definir, o melhor descritor por classe, a melhor combinação para o conjunto todo e a melhor combinação por classe. Para tal foram utilizadas para cada condição de contorno as respectivas métricas apresentadas na Tabela 5.3:

Tabela 5.3 Métricas por Condição de Contorno

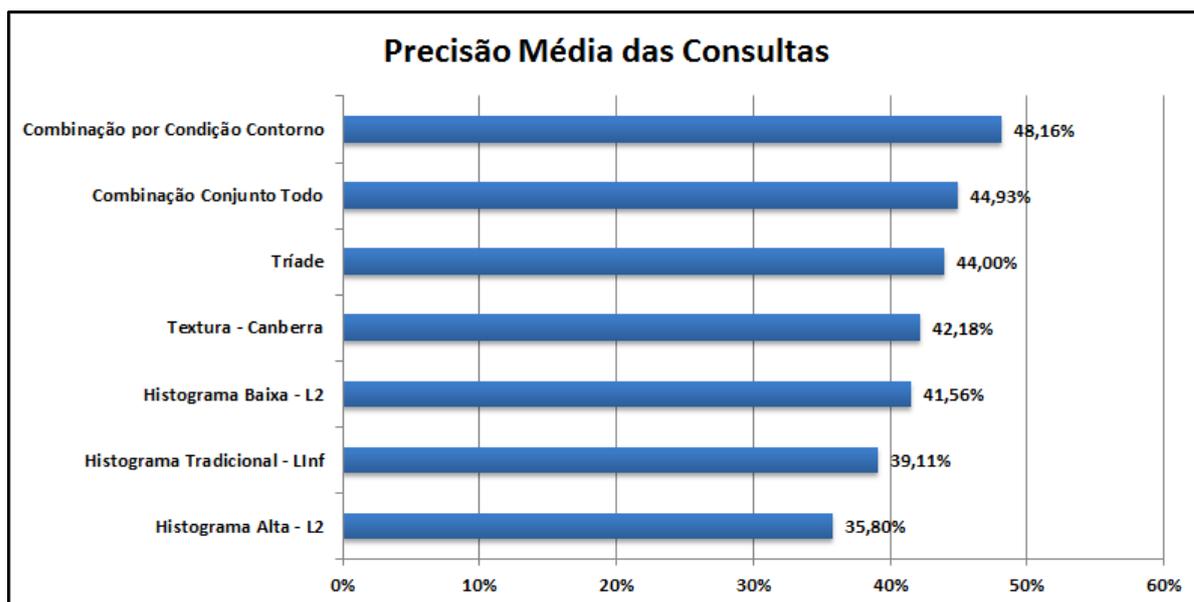
C. Contorno	Tríade	Combinação de Múltiplos Descritores
Consolidação	HB - L2	$(HB-L2)+(HT-LInf*0.1)+(Tex-Canberra*0.11)$
Enfisema	Tex-Canberra	$(HA-L2)+(Tex-Canberra*0.57)$
Vidro Fosco	HB - L2	$(HB-L2)+(HT-LInf*0.27)+(Tex-Canberra*0.44)$
Favo de Mel	HB - L2	$(HA-L2)+(Tex-Canberra*0.67)$
Espessamento	HA - L2	$(HT-LInf)+(HA-L2*10)+(Tex-Canberra*1.74)$

Comb. Única para Conj. Todo	$(HB-L2) + (HT-LInf*1.05) + (HA-L2*1.38) + (Tex-Canberra*3.08)$
-----------------------------	---

Os resultados obtidos são apresentados na Figura 5.4 por curvas de precisão *versus* revocação das consultas por similaridade, comparando utilização dos descritores simples, a tríade, uma combinação de múltiplos descritores para o conjunto todo e a combinação de múltiplos descritores por condição de contorno. A análise dos resultados pode ser visualizada na Figura 5.5, que apresenta um gráfico com os percentuais de precisão média alcançada pelas consultas.



**Figura 5.4 – Curvas de Precisão x Revocação das consultas executadas para o conjunto de dados TC\_Pulmão.**



**Figura 5.5 – Precisões médias obtidas nas consultas realizadas no conjunto de dados TC\_Pulmão.**

A partir desse gráfico de precisão média é possível observar a melhora alcançada pelo método proposto sobre as demais técnicas comparadas. O método proporcionou uma melhora de 14,16% a 34,51% sobre a precisão média da utilização de um único descritor, 9,46% sobre a precisão média da utilização da tríade, e 7,17% sobre a precisão média da utilização de uma combinação de múltiplos descritores para o conjunto todo.

Em uma segunda formulação para o experimento é utilizado a precisão média das primeiras 25% posições de revocação para encontrar estas mesmas métricas. Este método torna-se útil para sistemas onde os primeiros resultados são mais importantes e o retorno ao usuário é analisado apenas para as primeiras imagens retornadas e na prática não seria utilizado um número maior de imagens para análise (PONCIANO-SILVA, 2009).

Desta maneira os descritores, a tríade e as combinações ficaram da seguinte forma:

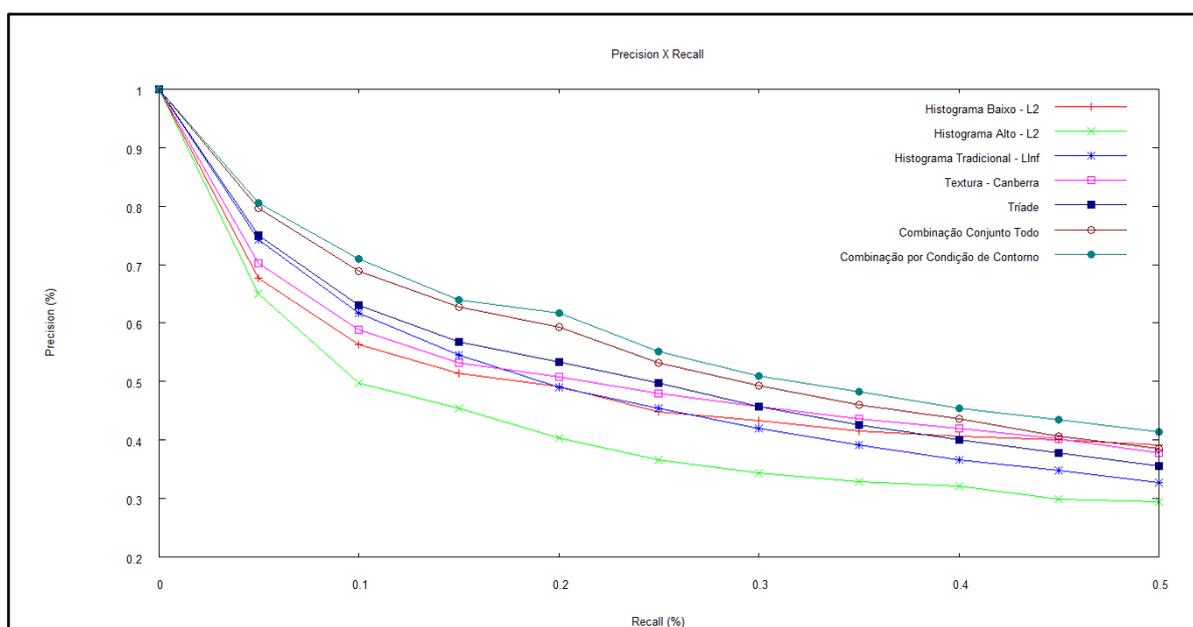
**Tabela 5.4 Definição dos Descritores**

Conjunto	Característica	Função de Distância
TC_Pulmão	Histograma Alto (HÁ)	L2
TC_Pulmão	Histograma Baixo (HB)	L2
TC_Pulmão	Histograma Trad. (HT)	L2
TC_Pulmão	Haralick (Tex)	Canberra

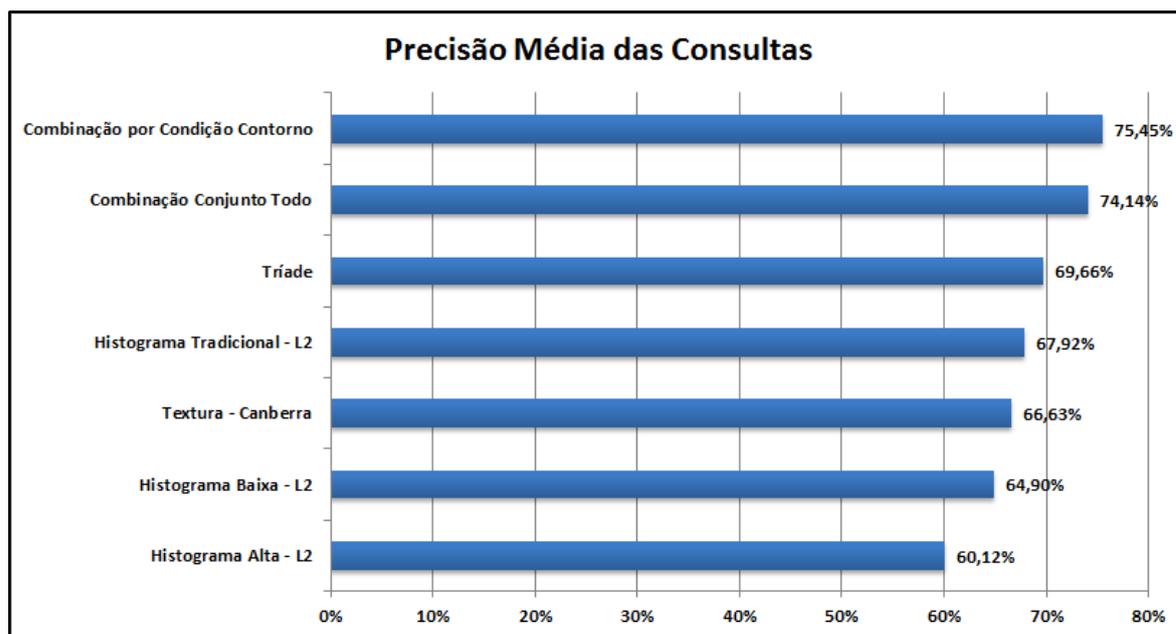
**Tabela 5.5 Métricas por Condição de Contorno**

C. Contorno	Triade	Combinação de Múltiplos Descritores
Consolidação	Tex - Canberra	$(HB-L2)+(HT-L2*1.11)+(HA\_L2*0.15)+(Tex-Canberra* 1.11)$
Enfisema	Tex - Canberra	$(HA-L2)+(Tex-Canberra*1.48)$
Vidro Fosco	HT – L2	$(HB-L2)+(HT-L2*1.25)+(HA-L2*1.6)+(Tex-Canberra* 1.43)$
Favo de Mel	HB – L2	$(HB-L2)+(HT-L2)+(HA-L2*4)+(Tex-Canberra*3.64)$
Espessamento	HT – L2	$(HB-L2)+(HT-L2*0.33)+(HA-L2*10)+(Tex-Canberra* 2.67)$
Comb. Única para Conj. Todo		$(HT-L2) + (HA-L2*1.48) + (Tex-Canberra*1.29)$

As curvas de precisão nesta nova configuração podem ser vistas na Figura 5.6. e analisadas logo na sequência pela Figura 5.7.



**Figura 5.6 – Curvas de Precisão x Revocação das consultas executadas no conjunto de dados TC\_Pulmão.**



**Figura 5.7 – Precisões médias obtidas nas consultas realizadas no conjunto de dados TC\_Pulmão.**

A análise das informações permite concluir uma melhora na precisão média das consultas de 11,08% a 21,07% sobre a utilização de descritores individuais, 8,31% sobre a tríade e 1,77% sobre a combinação de múltiplos descritores.

### **5.2.3 Conjunto de Imagens de Áreas de Interesse (ROI) de Tomografia Computadorizada de Pulmão (ROI\_TCPulmão)**

No segundo experimento foi realizada a combinação de múltiplos descritores utilizando 5 condições de contorno referente as 6 classes pré-definidas de imagens. Como no experimento anterior a classe de imagens normal não foi utilizada como condição de contorno por não apresentar uma característica visual capaz de evidenciá-la do restante do conjunto.

Este experimento tem como diferencial sua base de dados composta por 3.257 elementos, onde foram definidos aleatoriamente 100 elementos de cada classe para compor uma base de treinamento com uma amostragem de 600 elementos.

A partir desta base e treinamento, a configuração das métricas considerando a precisão média de toda a revocação é mostrada na Tabela 5.6.

Tabela 5.6 Métricas por Condição de Contorno

C. Contorno	Tríade	Combinação de Múltiplos Descritores
Classe 1	HT-L1	(Zernike-Canberra) + (Tex-Canberra*10) + (HT-L1*10)
Classe 2	HT-L1	(Zernike-Canberra) + (Tex-Canberra* 0.1) + (HT-L1*5.71)
Classe 3	HT-L1	(Zernike-Canberra) + (Tex-Canberra*1.03) + (HT-L1*2.35)
Classe 4	Tex-Canberra	(Zernike-Canberra) + (Tex-Canberra*2.11) + (HT-L1*2)
Classe 5	HT-L1	(Zernike-Canberra) + (Tex-Canberra*0.13) + (HT-L1*5.71)
Comb. Única para Conj. Todo		(Zernike-Canberra) + (Tex-Canberra*1.11) + (HT-L1*3.33)

Para esta coleção de imagens também foram realizados experimentos considerando apenas as primeiras 25% posições de revocação, porém os resultados foram muito próximos considerando toda a revocação.

Os resultados obtidos por esta fase de treinamento são apresentados na Figura 5.8 em forma de gráfico, com curvas de precisão x revocação das consultas por similaridade da utilização dos descritores simples, a tríade, uma combinação de múltiplos descritores para o conjunto todo e a combinação de múltiplos descritores por condição de contorno. A análise é realizada na Figura 5.9, que apresenta um gráfico com os percentuais de precisão média alcançada pelas consultas.

Para deixar claro estes dois gráficos abaixo (Figura 5.8 e Figura 5.9) são referentes à experimentos realizados com os elementos de amostragem composto pelas 600 imagens (base de treinamento). Os resultados obtidos aplicando estas métricas ao conjunto todo (3.247 imagens) são apresentados e discutidos na sequência (Figura 5.10 e Figura 5.11).

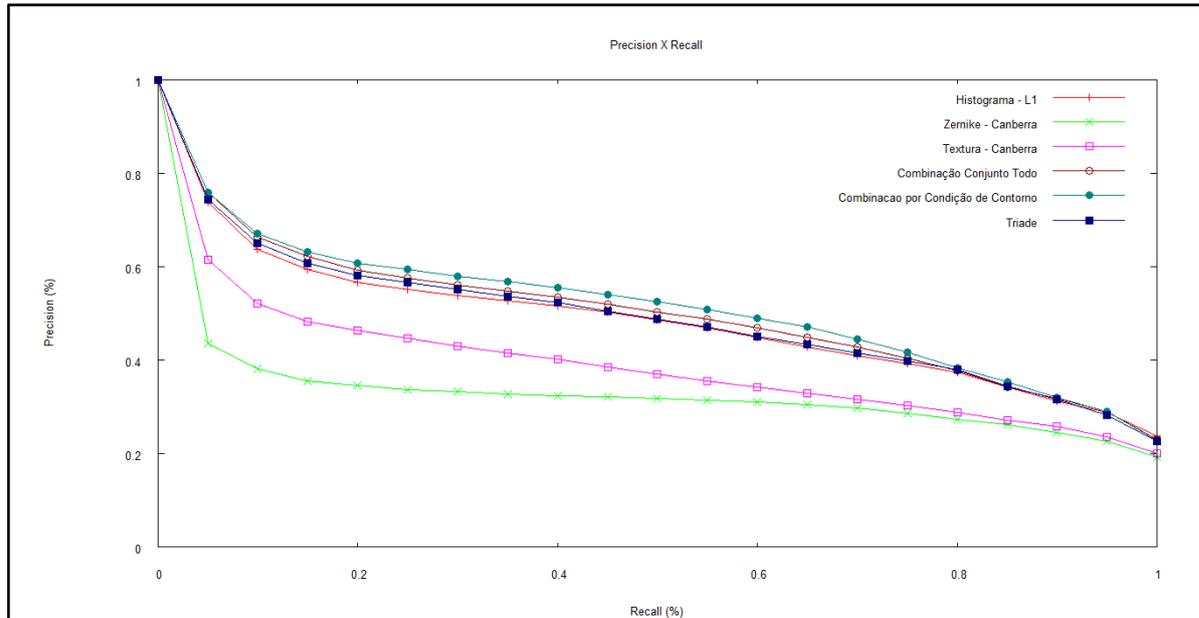


Figura 5.8 – Curvas de Precisão x Revocação das consultas executadas no conjunto de dados ROI\_TCPulmão (fase de treinamento).

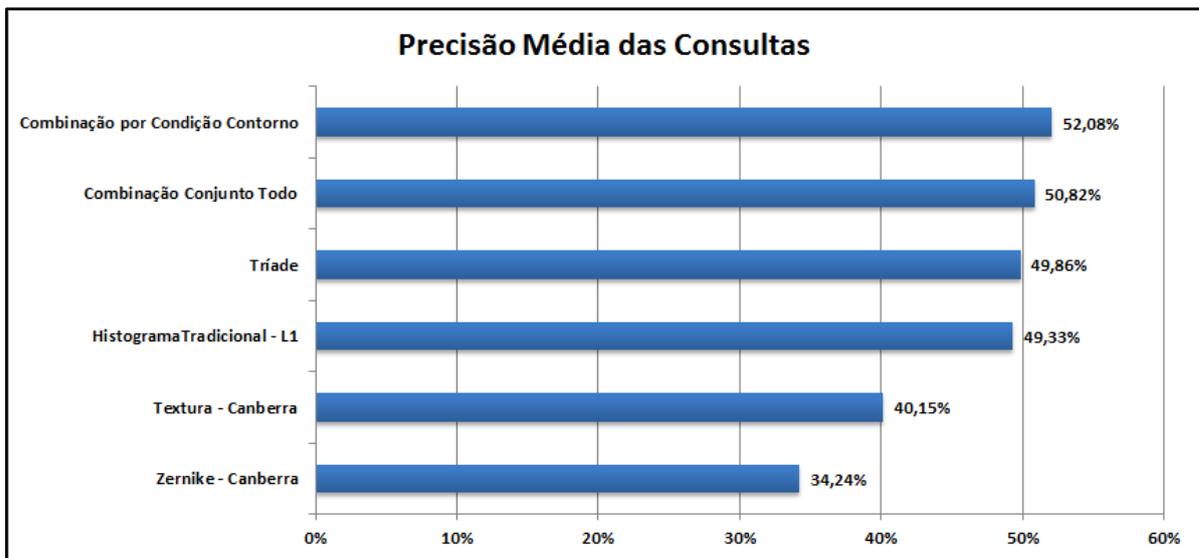


Figura 5.9 – Precisões médias obtidas nas consultas realizadas no conjunto de dados ROI\_TCPulmão(fase de treinamento).

Comparando as informações do gráfico é possível concluir uma melhora na precisão média das consultas de 5,57% a 52,08% sobre a utilização de descritores individuais, 4,45% sobre a tríade e 2,49% sobre uma única combinação de múltiplos descritores para o conjunto todo.

Com as métricas então definidas na fase de treinamento, partindo para o ambiente de utilização, as mesmas foram aplicadas ao conjunto completo composto

pelos 3.247 elementos, onde permaneceram com melhor precisão sobre as demais práticas. Os resultados são apresentados pela Figura 5.10 e a análise é feita na figura 5.11 que apresenta um ganho na precisão média de aproximadamente 1,5% sobre a utilização da combinação para o conjunto todo, 4,26% sobre a tríade, e a partir de 3,35% sobre a utilização dos descritores individuais chegando a até 58,48%.

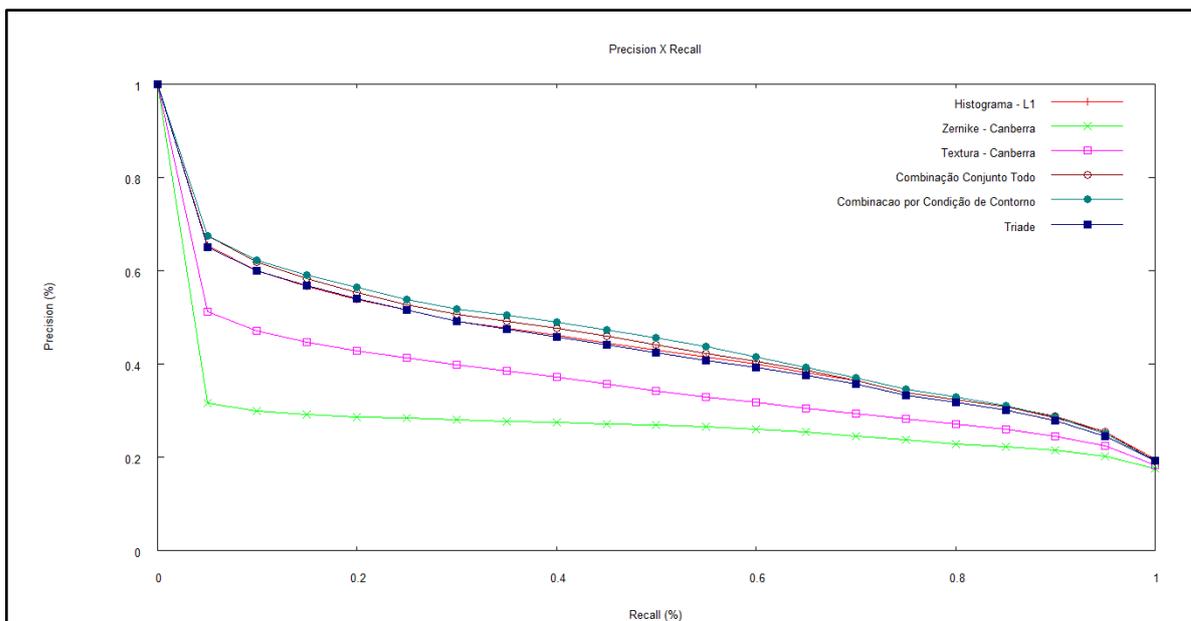


Figura 5.10 – Precisões médias obtidas nas consultas realizadas no conjunto de dados ROI\_TCPulmão (conjunto todo)

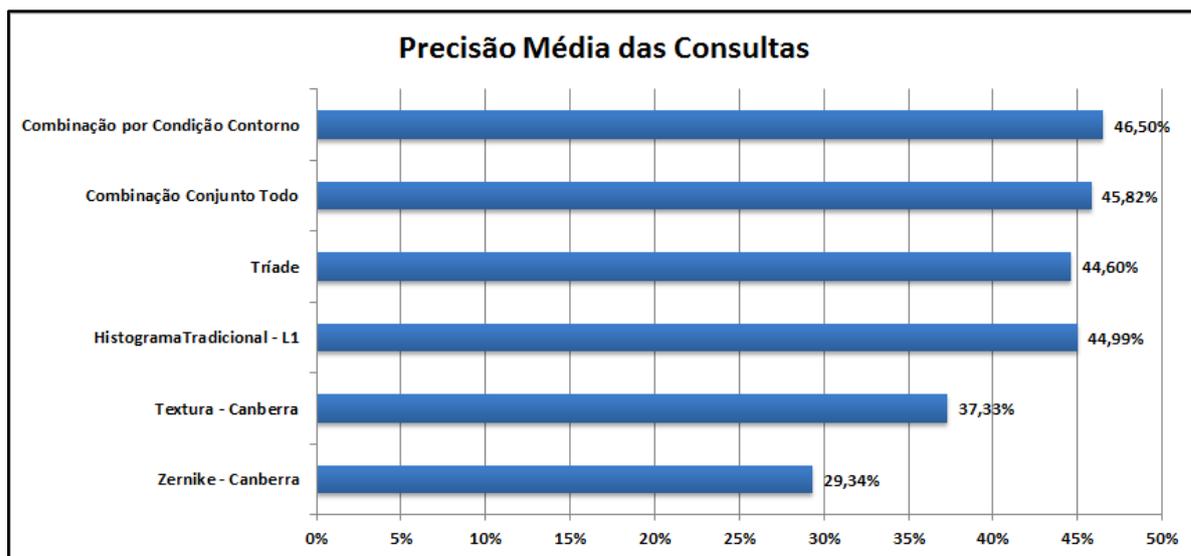


Figura 5.11 – Precisões médias obtidas nas consultas realizadas no conjunto de dados ROI\_TCPulmão (conjunto todo).

Este experimento pode simular em menor proporção o que seria realizado em uma implantação real do método, tendo uma amostragem dos dados para treinamento e enfim a sua utilização em um ambiente real.

Desta forma o método continuou apresentando melhores resultados, com uma maior precisão média sobre as técnicas comparadas.

#### 5.2.4 Conjunto de Imagens de Ressonância Magnética (MRI\_704)

Neste experimento foi feita a combinação de múltiplos descritores utilizando 8 condições de contorno e a precisão das consultas foi avaliada baseando-se em 40 classes pré-definidas de imagens. Os experimentos seguiram os mesmos critérios dos experimentos anteriores. As métricas utilizadas são apresentadas na Tabela 5.7.

**Tabela 5.7 Métricas por Condição de Contorno**

<b>C. Contorno</b>	<b>Triade</b>	<b>Combinação de Múltiplos Descritores</b>
Angiograma	Tex-Canberra	Tex-Canberra + (Zernike-L2*0.19) + (HT-L1*0.21)
Abdômen Axial	Zernike-L2	Tex-Canberra + (Zernike-L2*5.71) + (HT-L1*2.67)
Cabeça Axial	Zernike-L2	Tex-Canberra + (Zernike-L2*2.5) + (HT-L1*1.6)
Pélvis Axial	Zernike-L2	Tex-Canberra + (Zernike-L2*0.5) + (HT-L1*0.4)
Abdômen Coronal	Tex-Canberra	Tex-Canberra + (Zernike-L2*3.33) + (HT-L1*0.57)
Cabeça Coronal	Zernike-L2	Tex-Canberra + (Zernike-L2*5.71) + (HT-L1*0.5)
Cabeça Sagital	Zernike-L2	Tex-Canberra + (Zernike-L2*2.86) + (HT-L1*0.33)
Espinha Sagital	Zernike-L2	Zernike-L2
Comb. Única para Conj. Todo		Tex-Canberra + (Zernike-L2*2) + (HT-L1*0.4)

As curvas de precisão resultantes do experimento são apresentadas na Figura 5.12, sempre comparando com a utilização dos descritores simples, a tríade, uma combinação de múltiplos descritores para o conjunto todo e a combinação de múltiplos descritores por condição de contorno. A síntese dos resultados pode ser visualizada na Figura 5.13 para uma melhor análise e conclusão.

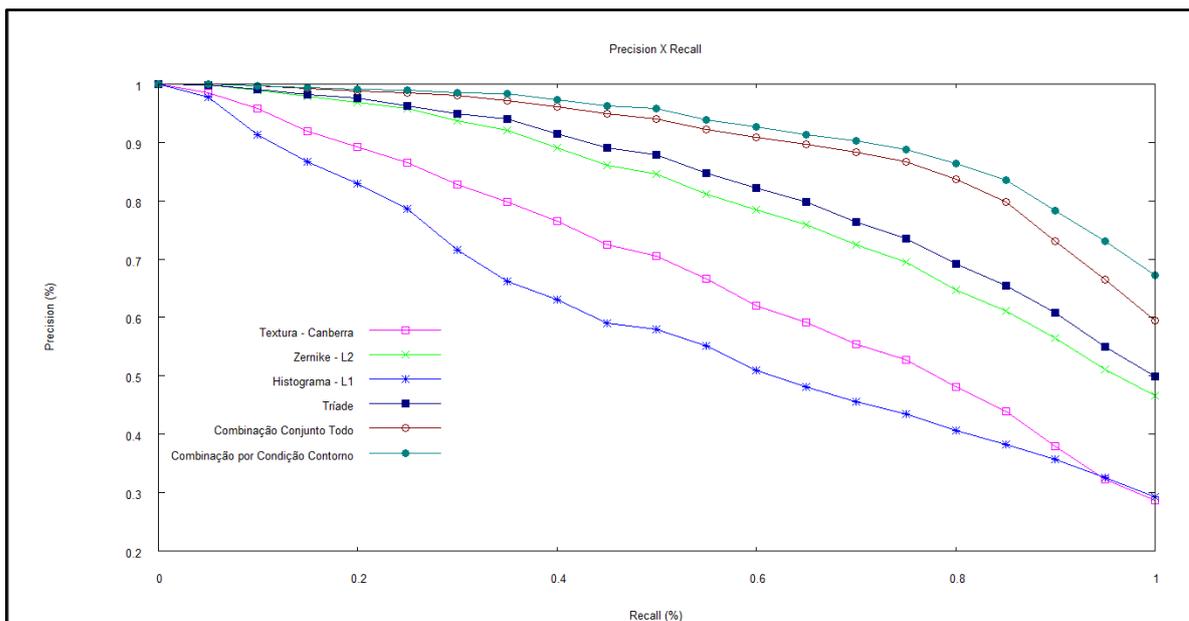


Figura 5.12 – Curvas de Precisão x Revocação das consultas executadas no conjunto de dados MRI\_704.

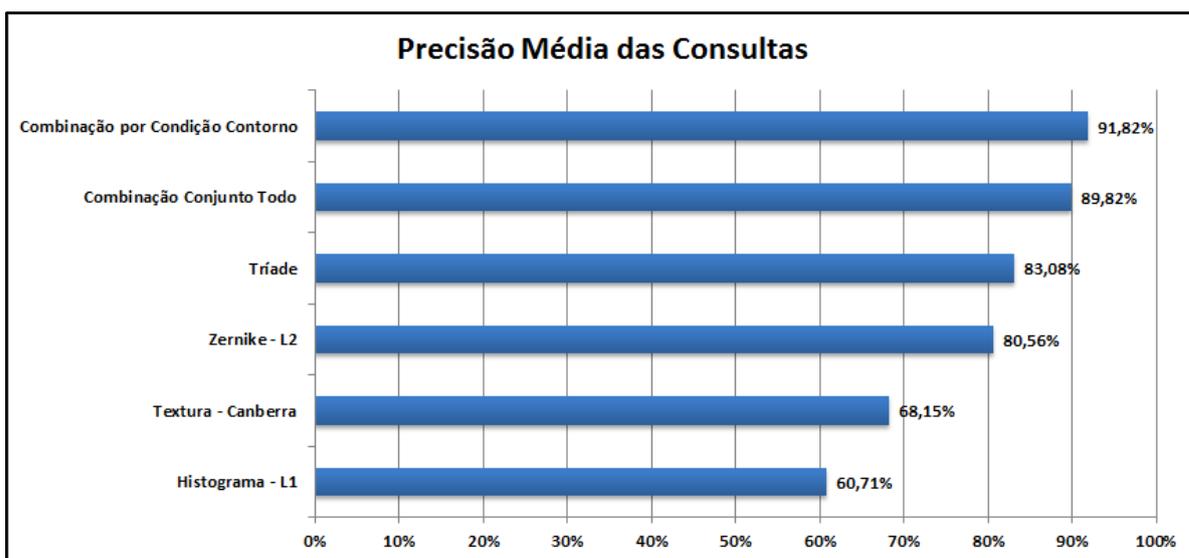


Figura 5.13 – Precisões médias obtidas nas consultas realizadas no conjunto de dados MRI\_704.

A análise das informações permite concluir uma melhora na precisão média das consultas de 13,97% a 51,24% sobre a utilização de descritores individuais, 10,51% sobre a tríade e 2,22% sobre a combinação de múltiplos descritores.

### 5.2.5 Combinação de múltiplos descritores partindo do melhor descritor

Um quarto experimento foi realizado com intuito de diminuir o tempo gasto pelos algoritmos exaustivos para combinação linear. As combinações com todos os pesos possíveis a fim de encontrar a melhor combinação, tornam esses algoritmos com muitas operações de combinação e laços de repetição, deixando o processo demorado e um crescimento exponencial de acordo com o número de descritores utilizados pela quantidade de elementos do conjunto para cada condição de contorno.

Neste quarto experimento foram relacionadas informações coletadas no conjunto de dados TC\_Pulmão para levantar outra forma de combinar os descritores, incluindo na combinação cada um dos descritores partindo daquele que obteve melhor média de precisão. Desta mesma forma foram realizadas as combinações para cada condição de contorno.

Fato curioso que o tempo gasto para executar o algoritmo exaustivo de combinação linear para este conjunto e com os parâmetros definidos conforme sessão 5.2.1 foram gastas cerca de vinte e oito horas para que o algoritmo seja executado e gerado as informações necessárias para a análise.

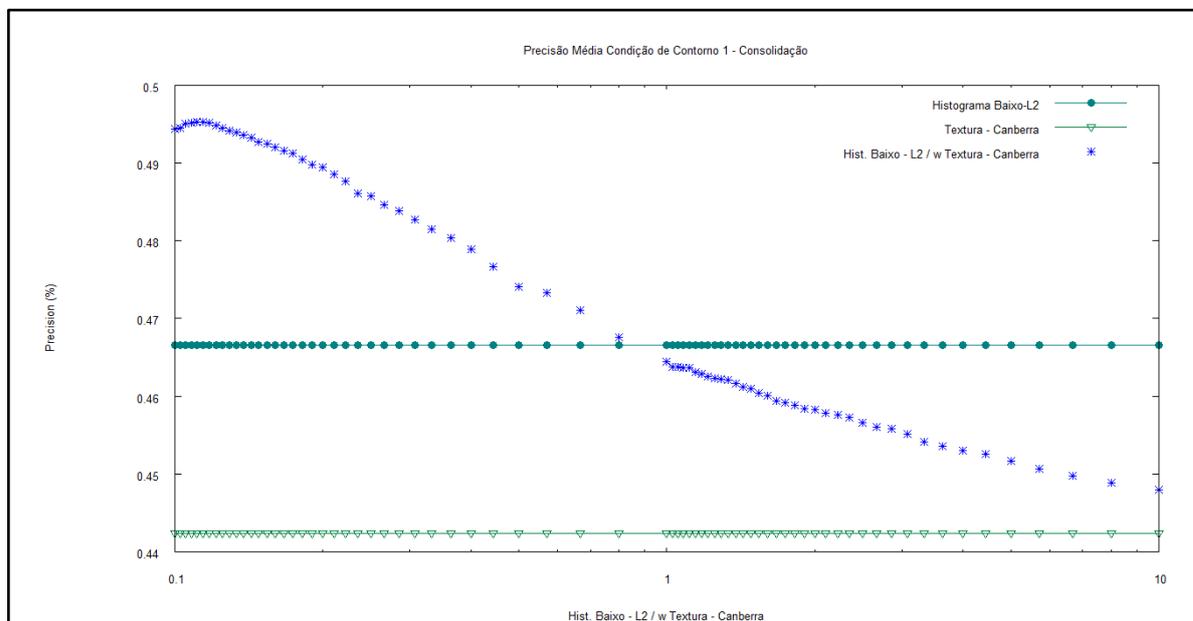
O algoritmo foi implementado em C++ e executado em um notebook com as configurações: Dell Vostro 3500, processador Intel Core i3 M370 2.40 GHz, com 4 Gb de memória RAM e sistema operacional Microsoft Windows 7 Ultimate 64 bits Service Pack 1.

Pelo método proposto na sessão atual utilizando o mesmo equipamento foram gastos apenas cerca de quatro horas para execução do algoritmo e geração dos dados necessários para análise. Lembrando que este tempo cresce exponencialmente de acordo com a quantidade de descritores e quantidade de elementos da base de dados.

Usando como exemplo a primeira condição de contorno para a classe de imagens de Consolidação, os descritores obtiveram as seguintes médias de precisão na sequência da melhor para a pior: Histograma Baixo – L2 46,65%, Textura – Canberra 44,24%, Histograma Tradicional – Llnf 38,26%, Histograma Alto – L2 34,50%.

Com esta informação foi possível fixar o peso 1 para o melhor descritor (Histograma Baixo – L2) e incluir o segundo melhor descritor (Textura – Canberra)

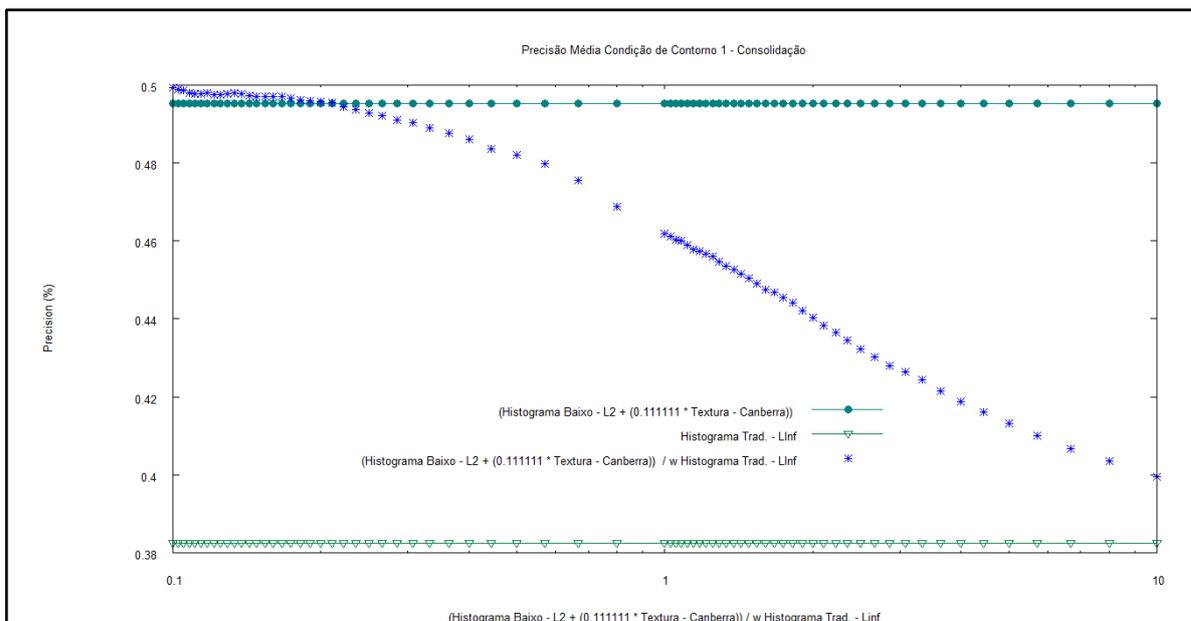
variando seu peso, os resultados são apresentados na Figura 5.14 que mostra como essa variação do peso  $w$  influencia na precisão da consulta através da combinação entre os dois descritores.



**Figura 5.14 – Precisão média combinando linearmente Histograma Baixo – L2 / w Textura - Canberra.**

Analisando as informações do gráfico, é possível concluir que a combinação dos descritores apresenta melhores resultados desde que, o peso de Histograma Baixo – L2 seja pelo menos 25% maior que o peso de Textura – Canberra e o melhor resultado é obtido com o peso de Histograma Baixo – L2 aproximadamente nove vezes maior que o peso de Textura – Canberra, que proporcionou uma precisão média de 49,53%, ou uma melhora de aproximadamente 6,17% na consulta.

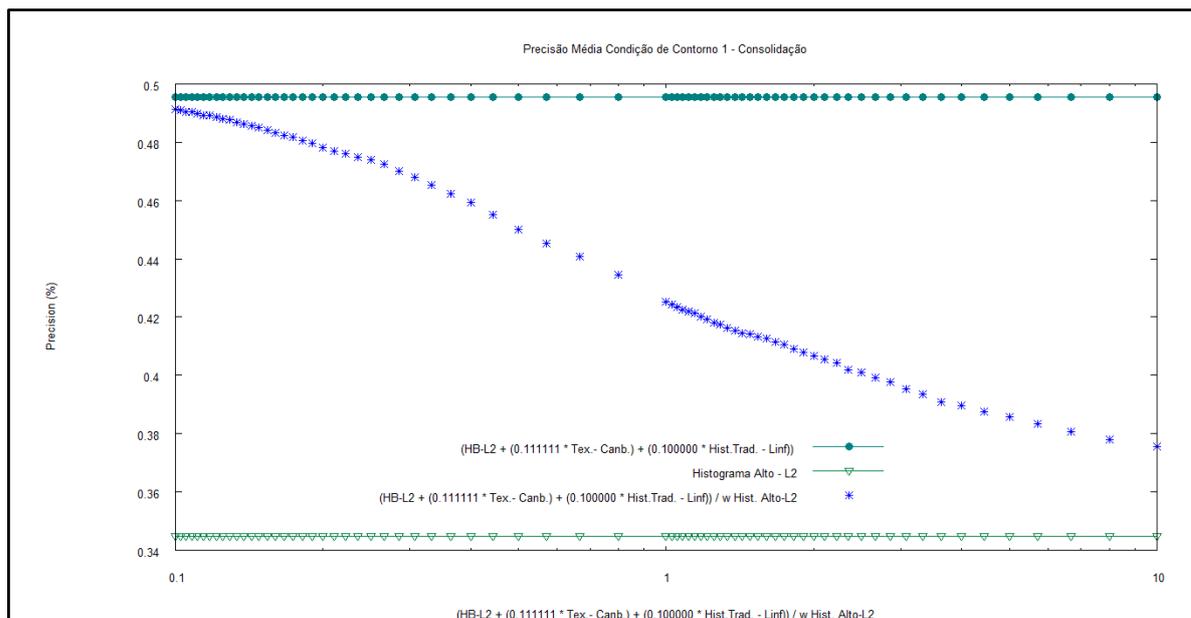
Depois de definido este melhor peso entre o primeiro par de descritores, deve ser incluído o terceiro descritor. Agora é necessário fixar esta combinação (Histograma Baixo – L2 +  $(0.11 * \text{Textura – Canberra})$ ) como sendo uma função de distância, fixando também seu peso e incluir o próximo descritor que é o Histograma Tradicional - LInf. O resultado é apresentado na Figura 5.15.



**Figura 5.15 – Precisão média combinando linearmente (Histograma Baixo – L2 + (0.11 \* Textura – Canberra)) / w Histograma Tradicional - Llnf**

Da mesma forma que o anterior, no gráfico é possível observar que a inclusão do terceiro descritor melhorou a combinação desde que, o peso de Histograma Tradicional – Llnf fique entre 4 a 9 vezes menor quando incluído a combinação dos dois primeiros descritores sendo que, o melhor resultado alcançado é para o peso aproximadamente nove vezes menor, onde possibilitou 49,91% de precisão na consulta melhorando em 0,77%.

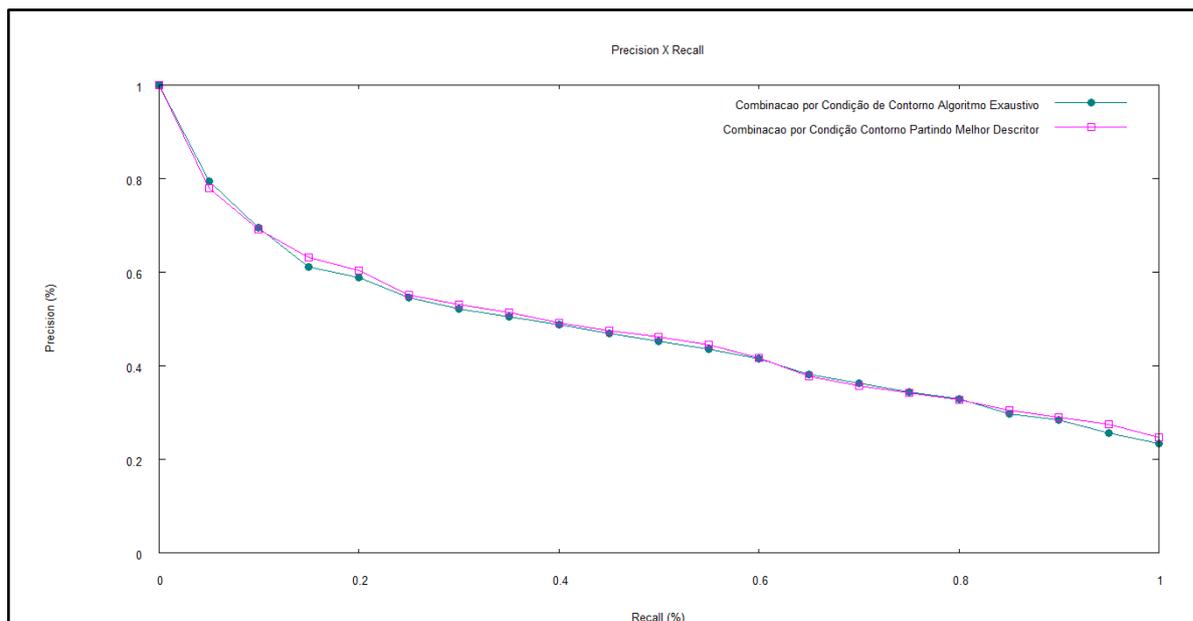
Este processo continua para todos os demais descritores. O descritor Histograma Alto – L2 ao ser incluído na combinação degradou o poder de discriminação e diminuiu a precisão anteriormente alcançado apenas com os três primeiros descritores como podemos ver na Figura 5.16. Então, conseqüentemente ele não será utilizado na combinação para esta condição de contorno.



**Figura 5.16 – Precisão média combinando linearmente (Histograma Baixo – L2 + (0.11 \* Textura – Canberra) + (0.10 \* Histograma Tradicional – LInf)) / w Histograma Alto – L2**

Neste experimento, através da combinação de múltiplos descritores partindo do melhor descritor, a métrica resultante para a condição de contorno “Consolidação” foi: Histograma Baixo – L2 + (0.11 \* Textura – Canberra) + (0.10 \* Histograma Tradicional - LInf).

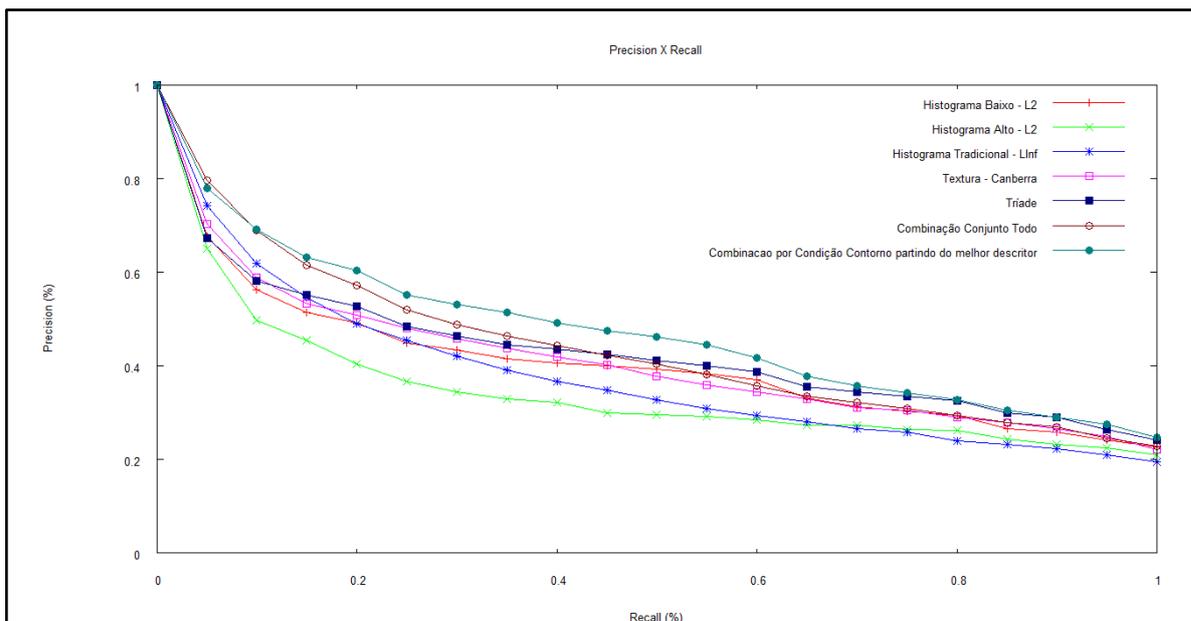
Este mesmo método foi realizado para as demais Condições de Contorno. Porém ao final do experimento, pode-se observar um fator muito relevante, a combinação realizada desta forma obteve um resultado muito parecido com a combinação realizada através de algoritmos exaustivos. Esse resultado pode ser observado na Figura 5.17.



**Figura 5.17 – Curva de Precisão x Revocação da Combinação de Descritores feito por algoritmos exaustivos e a Combinação de Descritores realizado neste experimento 4.**

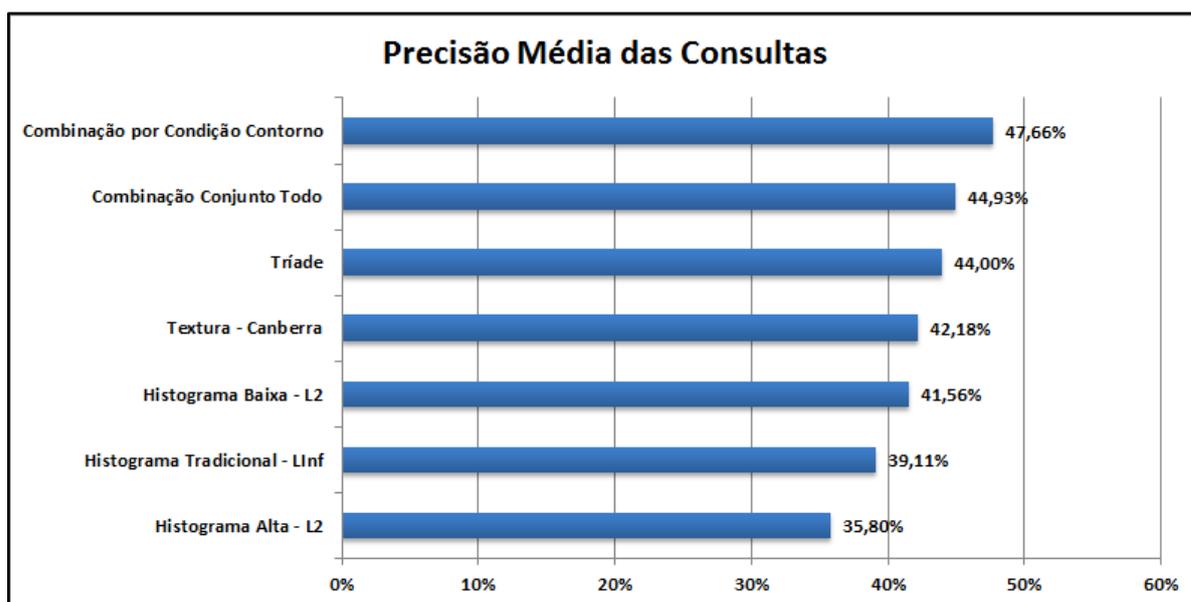
A precisão média alcançada pelos dois métodos empregados foi de 48,16% para a combinação realizada pela combinação linear exaustiva de 47,66% para a combinação deste quarto experimento que parte do melhor descritor, uma diferença de 1,02%. Porém a forma proposta partindo do melhor descritor necessitou de apenas 0,0274% de esforço para os cálculos de distância para cada condição de contorno, ou seja, aproximadamente 3.650 vezes menor.

Além disso, apesar dessa perda de 1,02% o método também conseguiu melhores resultados do que as práticas comparadas por este estudo.



**Figura 5.18 –Curvas de Precisão x Revocação das consultas executadas para o conjunto de dados TC\_Pulmão comparando com a Combinação por Condição de Contorno partindo do melhor descritor.**

Através da Figura 5.18 foi possível identificar que a combinação partindo do melhor descritor obteve melhores resultados. Este ganho na precisão das consultas fica mais claro ao analisar a Figura 5.19, onde a combinação apresentou um ganho de 6,07% sobre a precisão média de uma combinação para o conjunto todo, 8,33% sobre a tríade e até 33,13% de ganho sobre o uso dos descritores individuais.



**Figura 5.19 –Precisão média das consultas executadas para o conjunto de dados TC\_Pulmão comparando com a Combinação por Condição de Contorno partindo do melhor descritor.**

### 5.3 Análise dos resultados obtidos

As condições de contorno se mostraram eficientes para determinar o balanceamento de descritores de forma a proporcionar uma melhora nas consultas por similaridade. O que também ocorreu com a combinação de descritores, que complementou a utilização dessas condições de contorno e proveu uma métrica mais efetiva para cada subgrupo de elementos do conjunto de dados.

O método proposto apresentou melhores resultados em todos os experimentos realizados. Em alguns dos experimentos observa-se uma precisão média entre 5,57% e 52,08% maior que o melhor descritor único, 4,45% e 10,51% sobre a tríade e 1,77% e 7,17% sobre a combinação de múltiplos descritores única para o conjunto todo.

Além das métricas calculadas a partir da média total de precisão, houve experimentos que exploraram apenas a precisão nos primeiros 25% de revocação. Esses casos são utilizados onde as primeiras imagens possuem maior prioridade ou a análise é realizada apenas com alguns primeiros elementos resultantes da consulta. Essa modalidade reforça a precisão neste ponto crítico da consulta e consegue, através do método proposto, obter um ganho de precisão.

Para complementar, foi estudada outra forma de executar o balanceamento linear de múltiplos descritores.

Apesar da combinação feita através de algoritmos exaustivos de combinação linear chegarem a melhor combinação de descritores possível, a combinação partindo do melhor descritor e agregando os demais obteve resultados muito parecidos com os do método exaustivo. No experimento realizado, a prática efetuou 3.650 vezes menos cálculos de distância e apesar de degradar 1,02% a precisão média obtida, ainda obteve melhor precisão as demais técnicas comparadas pelo presente estudo.

# Capítulo 6

## CONCLUSÃO

---

*Este capítulo apresenta as conclusões da presente dissertação de mestrado. Primeiramente é feita uma conclusão sobre o trabalho. Posteriormente são destacadas as contribuições na área de estudos para a combinação de múltiplos descritores e diminuição do gap semântico em consultas por similaridade e possíveis trabalhos futuros.*

### 6.1 Considerações Finais

O principal objetivo deste trabalho foi propor uma metodologia para execução de consultas por similaridade que contribuísse com a diminuição do gap semântico enfrentado por estas consultas. Portanto, foi proposta a utilização de condições de contorno para determinar subconjuntos de dados e então estimar a melhor combinação de múltiplos descritores em consultas por similaridade. Esse parâmetro permitiu que a consulta fosse realizada através do balanceamento ideal para cada conjunto de elementos que se comportam de maneira parecida nessas consultas.

O método proposto apresentou melhores resultados nas consultas por similaridade possibilitando utilizar uma métrica que melhor evidencia cada um dos seus respectivos subconjuntos. A combinação de descritores colaborou no sentido de utilizar as características de forma que uma complemente a outra, dando poder maior de representação para os elementos.

A utilização do método proporcionou ganhos em todos os experimentos executados, chegando a ter um ganho na precisão média de 14,16% maior que o

melhor descritor único, 10,51% sobre a tríade e 7,17% sobre a combinação de múltiplos descritores única para o conjunto todo.

## 6.2 Principais Contribuições

O presente trabalho propôs a utilização de duas técnicas: a utilização de condições de contorno para delimitar subconjuntos de elementos que possuem comportamentos semelhantes na consulta por similaridade e a combinação de múltiplos descritores que proporciona uma melhor representação visual dos elementos. Com a junção das técnicas foi possível utilizar um melhor balanceamento para cada condição de contorno visando a diminuição do gap semântico nas consultas por similaridade.

Nos experimentos realizados o método proposto sempre teve um ganho na precisão média, indicando sua eficiência e motivando sua utilização em consultas por similaridade e sistemas CBIR.

## 6.3 Trabalho Futuros

No decorrer do trabalho, foram levantadas possíveis aplicações em pesquisas que levem ainda mais a diminuição do gap-semântico em consultas por similaridade. Como trabalhos futuros podem ser abordados algumas sugestões como:

- Utilizar as condições de contorno em métodos de algoritmos evolutivos e Realimentação por Relevância, que se mostraram muito eficientes em combinar múltiplos descritores que melhor representem os dados.
- Efetuar mais experimentos com a combinação de múltiplos descritores partindo do melhor descritor para comprovar efetivamente sua eficácia.

# REFERÊNCIAS

---

AMATO, G.; RABITTI, F.; SAVINO, P.; ZEZULA, P.; "**Region Proximity in Metric Spaces and Its Use for Approximate Similarity Search**", ACM Transactions on Information Systems (TOIS), Vol. 21, No. 2, April 2003, Pages 192–227.

AMATO, G.; MAINETTO, G.; SAVINO, P.: "**A Query Language for Similarity-based Retrieval of Multimedia Data**", Advances in Databases and Information Systems 1997; Proceedings of the First East-European Symposium on Advances in Databases and Information Systems, (ADBIS'97), St Petersburg, 2-5 September 1997

AREVALILLO-HERRÁEZ, M.; FERRI, F. J.; MORENO-PICOT, S.: "**Distance-based relevance feedback using a hybrid interactive genetic algorithm for image retrieval.**" Applied Soft Computing 11, 2 (March 2011), 1782-1791

AREVALILLO-HERRÁEZ, M.; FERRI, F. J.; DOMINGO, J.: "**A naive relevance feedback model for content-based image retrieval using multiple similarity measures**", Pattern Recognition 43 (3) (2010) 619–629.

BAEZA-YATES, R.A.; RIBEIRO-NETO, B.A.: **Modern Information Retrieval**. ACM Press / Addison-Wesley (1999)

BERRETTI, S.; DEL BIMBO, A.; PALA, P.; "**Extraction of resource descriptors for distributed content based image retrieval,**" Pattern Recognition, 2002. Proceedings. 16th International Conference on, vol.3, no., pp. 995- 998 vol.3, 2002

BRAUNMÜLLER, B.; ESTER, M.; KRIEGEL, H. P.; SANDER, J.; **Multiple Similarity Queries: A Basic DBMS Operation for Mining in Metric Databases**. IEEE Transactions on Knowledge and Data Engineering (TKDE), Vol. 13, No. 1, 2001.

BROILO, M.; NATALE, F.: "**A stochastic approach to image retrieval using relevance feedback and particle swarm optimization**", Multimedia, IEEE Transactions on 12 (4) (2010) 267–277

BUENO, R.; **Tratamento do tempo e dinamicidade em dados representados em espaços métricos**. 2009. Tese (Doutorado em Ciência da Computação) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2009.

BUENO, R.; KASTER, D. S.; PATERLINI, A. A.; TRAINA, A. J. M.; TRAINA JR., C. T., "**Unsupervised scaling of multi-descriptor similarity functions for medical image datasets,**" In 22th IEEE Intl. Symposium on Computer-Based Medical Systems (CBMS 2009). Albuquerque, New Mexico, USA: IEEE, 2009, pp. 1–8, 24.

BUENO, R.; KASTER, D.S.; RAZENTE, H.L.; BARIONI, M.C.N.; TRAINA, A.J.M.; TRAINA JR., C.; , **"Using Visual Analysis to Weight Multiple Signatures to Discriminate Complex Data,"** Information Visualisation (IV), 2011 15th International Conference on , vol., no., pp.282-287, 13-15 July 2011

BUENO, R.; RIBEIRO, M.X.; TRAINA, A.J.M.; TRAINA JR., C.: **"Improving medical image retrieval through multi-descriptor similarity functions and association rules,"** Symposium on Computer-Based Medical Systems (CBMS 2010), 2010 IEEE 23rd International Symposium on , vol., no., pp.309-314, 12-15 Oct. 2010.

BUGATTI, P. H.: **"Análise da Influencia de funções de distancia para o processamento de consultas por similaridade em recuperação a imagens por conteúdo."** 2008. Dissertação (Mestrado em Ciência da Computação) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2008.

BUGATTI, P. H.; TRAINA, A. J. M.; TRAINA JR., C.: **"Assessing the best integration between distance-function and imagefeature to answer similarity queries"**. In SAC, pages 1225– 1230, Fortaleza, CE, Brazil, 2008. ACM.

BUSTOS, B.; KEIM, D.; SAUPE, D.; SCHRECK, T.; VRANIC, D.: **"Automatic selection and combination of descriptors for effective 3d similarity search,"** in Multimedia Software Engineering. Miami, FL, USA: IEEE, 2004, pp. 514–521.

CAICEDO, J. C.; GONZÁLEZ, F. A.; TRIANA, E.; ROMERO, E.: **"Design of a medical image database with content-based retrieval capabilities."** In Advances in Image and Video Technology, pages 919–931, Santiago, Chile, 2007. Springer.

CARÉLO, C. C.; POLA, I. R.; CIFERRI, R. R.; TRAINA, A. J.; TRAINA-JR, C.; CIFERRI, C. D. A.: **"The Onion-Tree: Quick Indexing of Complex Data in the Main Memory."** In Proceedings of the 13th East European Conference on Advances in Databases and Information Systems (ADBIS), 2009.

CHEN, M.; FU, P.; SUN, Y.; ZHANG, H.: **"Image retrieval based on multi-feature similarity score fusion using genetic algorithm,"** Computer and Automation Engineering (ICCAE), 2010 The 2nd International Conference on , vol.2, no., pp.45-49, 26-28 Feb. 2010

CIACCIA, P.; PATELLA, M.; ZEZULA, P.: **"M-tree: An efficient access method for similarity search in metric spaces."** In VLDB, pages 426–435, Athens, Greece, 1997. Morgan Kaufmann.

da SILVA TORRES, R.; FALCÃO, A. X.: **"Content-Based Image Retrieval: Theory and Applications"**, Revista de Informática Teórica e Aplicada (RITA) 13 (2), 161--185, 2006.

da SILVA TORRES, R.; FALCÃO, A. X.; GONÇALVES, M. A.; PAPA, J. P.; ZHANG, B.; FAN, W.; FOX, E. A.: **"A genetic programming framework for content-based image retrieval."** Pattern Recogn. 42, 2 (February 2009), 283-292.

da SILVA TORRES, R.; FALCÃO, A. X.: "**Content-Based Image Retrieval: Theory and Applications**," Revista de Informática Teórica e Aplicada 13 (2) (2006) 161–185.

DATTA, R.; JOSHI, D.; LI, J.; WANG, J. Z.: "**Image retrieval: Ideas, influences, and trends of the new age**." ACM Computing Survey, 40(2):1–60, 2008.

DEEPAK, K. S.; RAI, H. G. N.; SYED, S.; KRISHNA, P. R.: "**Texture edge statistics for efficient retrieval of biomedical images**." In Proceedings of the 5th ACM COMPUTE Conference: Intelligent & scalable system technologies (2012).

DESERNO, T.M., ANTANI, S., LONG, L.R.: "**Ontology of gaps in content-based image retrieval**." J. Digital Imaging 22 (2009) 202\_215

FANG, D.; HAIMEI, H.; WEI, H.: "**Integrating multi-feature of image based on correspondence analysis**," Industrial Electronics and Applications (ICIEA), 2010 the 5th IEEE Conference on , vol., no., pp.1632-1635, 15-17 June 2010

FERREIRA, C. D.; da SILVA TORRES, R.; GONÇALVES, M. A.; FAN, W.: "**Image retrieval with relevance feedback based on genetic programming**". In SBBD, pages 120–134, 2008.

FERREIRA, C. D.; SANTOS, J. A.; da SILVA TORRES, R.; GONÇALVES, M. A.; REZENDE, R. C.; FAN, W.: "**Relevance feedback based on genetic programming for image retrieval**". Pattern Recognition Letters 32, 1 (January 2011), 27-37.

GARGANO, M.; NARDELLI, E.; TALAMO, M.: "**Abstract data types for the logical modeling of complex data**", presented at Information Systems, 1991, pp.565-583.

GONZALEZ, R. C; WOODS, R. E.: "**Processamento digital de imagens**." - 3ª Edição – São Paulo, Pearson Prentice Hall, 2010.

HACID, H.; ZIGHED, D. A.: "**Content-based retrieval in large image databases**," Granular Computing, 2006 IEEE International Conference on , vol., no., pp. 498. May 2006

HARALICK, R.; SHANMUGAM, K.; DINSTEN, I.: "**Textural features for image classification**." Systems, Man and Cybernetics, IEEE Transactions on SMC-3 (1973) 610\_621

HARALICK, R.M.: "**Statistical and structural approaches to texture**," Proceedings of the IEEE , vol.67, no.5, pp. 786- 804, May 1979

HEESCH, D.; RÜGER, S.: "**Combining features for content-based sketch retrieval – a comparative evaluation of retrieval performance**". In In Proc. 24th BCS-IRSG European Colloquium on IR Research, pages 41–52. Springer, 2002.

HUANG, Z.; SHEN, H.; LIU, J.; ZHOU, X.: **“Effective data co-reduction for multimedia similarity search”**. In Proceedings of the 2011 international conference on Management of data (SIGMOD '11). 2011. ACM, New York, NY, USA, 1021-1032.

KEIM, D.A.; HECZKO, M.; HINNEBURG, A.; WAWRYNIUK, M.: **“Multi-Resolution Similarity Search in Image Databases”**. In Proceedings of Multimedia Information Systems. 2002, 76-85.

KHOTANZAD, A.; HONG, Y.H.: **“Invariant image recognition by Zernike moments,”** Pattern Analysis and Machine Intelligence, IEEE Transaction on Knowledge. and Data Engineering, vol.12, no.5, pp.489-497, May 1990.

KORN, F.; PAGEL, B.-U.; FALOUTSOS, C.: **“On the 'dimensionality curse' and the 'self-similarity blessing’,”** IEEE Transactions on Knowledge and Data Engineering (TKDE), (2001), 13(1):96-111.

LIU, P.; JIA, K.; WANG, Z.; LV, Z.: **“A New and Effective Image Retrieval Method Based on Combined Features,”** Image and Graphics, 2007. ICIG 2007. Fourth International Conference on , vol., no., pp.786-790, 22-24 Aug. 2007

LIU, P.; JIA, K.; WANG, Z.: **“An Effective Image Retrieval Method Based on Color and Texture Combined Features,”** Intelligent Information Hiding and Multimedia Signal Processing, 2007. IIHMSPP 2007. Third International Conference on , vol.1, no., pp.169-172, 26-28 Nov. 2007

LIU, P.; JIA, K.; WANG, Z.: **“A New Image Retrieval Method Based on Combined Features and Feature Statistic,”** Image and Signal Processing, 2008. CISP '08. Congress on , vol.2, no., pp.635-639, 27-30 May 2008

LIU, Y.; ZHANG, D.; LU, G.; MA, W.: **“A survey of content-based image retrieval with high-level semantics”,** Pat-tern Recognition(40):262–282, 2007

MORIKAWA, C.; SILVA, G. C.: **“User interaction techniques for multimedia retrieval”**. In Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments (HCCE '12). 2012. ACM, New York, NY, USA, 68-75

MÜLLER, H.; MICHOUX, N.; BANDON, D.; GEISSBUHLER, A.: **“A review of content-based image retrieval systems in medical applications-clinical benefits and future directions.”** Int. Journal of Medical Informatics, 73(1):1–23, 2004.

PARE, G.; AUBRY, D.; LEPANTO, L.; SICOTTE, C.: **“Evaluating PACS Success: A Multidimensional Model,”** System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference on , vol., no., pp. 147c, 03-06 Jan. 2005

PETRAKIS, E.G.M.; FALOUTSOS, A.: **“Similarity searching in medical image databases,”** Knowledge and Data Engineering, IEEE Transaction on Knowledge. and Data Engineering, vol.9, no.3, pp.435-447, May/June 1997

PONCIANO-SILVA, M.: **“Processamento de consultas por similaridade em imagens médicas visando a recuperação perceptual guiada pelo usuário.”** 2009. Dissertação (Mestrado em Ciência da Computação) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2009.

PONCIANO-SILVA, M.; TRAINA, A.J.M.; AZEVEDO-MARQUES, P.M.; FELIPE, J.C.; TRAINA JR., C.: **“Including the perceptual parameter to tune the retrieval ability of pulmonary CBIR systems.”** In: CBMS. (2009) 1\_8

RUI, Y.; HUANG, T.S.; ORTEGA, M.; MEHROTRA, S.: **“Relevance feedback: a power tool for interactive content-based image retrieval,”** Circuits and Systems for Video Technology, IEEE Transactions on , vol.8, no.5, pp.644-655, Sep 1998

SILVA, A. T.; SANTOS, J. A.; FALCÃO, A. X.; da SILVA TORRES, R.; MAGALHÃES, L.P.: **“Incorporating multiple distance spaces in optimum-path forest classification to improve feedback-based learning.”** Computer Vision and Image Understanding, 116(4):510–523, 2012.

SINGH, R.; KOTHARI, R.: **“Relevance feedback algorithm based on learning from labeled and unlabeled data,”** Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on , vol.1, no., pp. 1- 433-6 vol.1, 6-9 July 2003

SOARES, L. C.; KASTER, D. S.: **“ cx-Sim: A Metric Access Method for Similarity Queries with Additional Conditions.”** Journal of Information and Data Management (JIDM), vol. 4 no. 3, pp. 437-452 (2013).

TRAINA, A. J. M.; TRAINA JR., C.: **“Similarity search in multimedia databases.”**, In Furht, B. e Marques O., editors, Handbook of Video Databases - Design and Applications, v. 1, pp. 711-738. 2003. CRC Press.

TRAINA JR., C.; TRAINA, A. J. M.; FALOUTSOS, C.; SEEGER, B.: **“Fast indexing and visualization of metric datasets using Slim-trees.”** IEEE Transaction on Knowledge. and Data Engineering, 14(2):244–260, 2002.

TRAINA JR., C.; TRAINA, A. J. M.; FILHO, R. S.; FALOUTSOS, C.: **“How to improve the pruning ability of dynamic metric access methods”**. In Proceedings of the eleventh international Conference on Information and knowledge management (CIKM '02). ACM, New York, NY, USA, 219-226. 2002.

VASCONCELOS, N.; KUNT, M.: **“Content-based retrieval from image databases: current solutions and future directions,”** Image Processing, 2001. Proceedings. 2001 International Conference on , vol.3, no., pp.6,9 vol.3, 2001a

VELTKAMP, R.; TANASE, M.: **“Content-based image retrieval systems: A survey,”** Department of Computing Science, Utrecht University. 2002. Technical Report.

VIEIRA, M. R.; TRAINA-JR, C.; CHINO, F. J. T.; TRAINA, A. J. M.: **“DBM-Tree: A Dynamic Metric Access Method Sensitive to Local Density Data.”** In Proceedings

---

of the Brazilian Symposium on Databases (SBBD'04) Brasília, Brazil, pp. 163–177, pages 163–177, 2004.

WANG, J. Y.; ZHU, Z.: "**Image retrieval system based on multi-feature fusion and relevance feedback**," Machine Learning and Cybernetics (ICMLC), 2010 International Conference on , vol.4, no., pp.2053-2058, 11-14 July 2010.

ZHUANG, D.; WANG, S.: "**Content-Based Image Retrieval Based on Integrating Region Segmentation and Relevance Feedback**," Multimedia Technology (ICMT), 2010 International Conference on , vol., no., pp.1-3, 29-31 Oct. 2010.