

MINISTÉRIO DA DEFESA  
EXÉRCITO BRASILEIRO  
DEPARTAMENTO DE CIÊNCIA E TECNOLOGIA  
INSTITUTO MILITAR DE ENGENHARIA  
CURSO DE MESTRADO EM SISTEMAS E COMPUTAÇÃO

SEIMOU HAMILTON OSHIRO

UM AMBIENTE INTELIGENTE PARA RECUPERAÇÃO DE  
IMAGENS

Rio de Janeiro  
2008

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

**INSTITUTO MILITAR DE ENGENHARIA**

**SEIMOU HAMILTON OSHIRO**

**UM AMBIENTE INTELIGENTE PARA RECUPERAÇÃO DE IMAGENS**

Dissertação de Mestrado apresentada ao Curso de Mestrado em Sistemas e Computação do Instituto Militar de Engenharia, como requisito parcial para obtenção do título de Mestre em Sistemas e Computação.

Orientador: Prof. Alex Vasconcelos Garcia - D.Sc.

Orientador: Prof. Ronaldo Ribeiro Goldschmidt - D.Sc.,

Rio de Janeiro  
2008

INSTITUTO MILITAR DE ENGENHARIA  
Praça General Tibúrcio, 80-Praia Vermelha  
Rio de Janeiro-RJ CEP 22290-270

Este exemplar é de propriedade do Instituto Militar de Engenharia, que poderá incluí-lo em base de dados, armazenar em computador, microfilmar ou adotar qualquer forma de arquivamento.

É permitida a menção, reprodução parcial ou integral e a transmissão entre bibliotecas deste trabalho, sem modificação de seu texto, em qualquer meio que esteja ou venha a ser fixado, para pesquisa acadêmica, comentários e citações, desde que sem finalidade comercial e que seja feita a referência bibliográfica completa.

Os conceitos expressos neste trabalho são de responsabilidade do autor e do orientador.

082 Oshiro, S. H.  
Um Ambiente Inteligente para Recuperação de  
Imagens/ Seimou Hamilton Oshiro.  
– Rio de Janeiro: Instituto Militar de Engenharia, 2008.  
117 p.: il., tab.

Dissertação (mestrado) – Instituto Militar de  
Engenharia – Rio de Janeiro, 2008.

1. Recuperação de Imagens 2. Recuperação baseada  
em Sentenças. 3. Recuperação baseada em Conteúdo.  
I. Um Ambiente Inteligente para Recuperação de  
Imagens. II. Instituto Militar de Engenharia.

CDD 621.367

**INSTITUTO MILITAR DE ENGENHARIA**  
**SEIMOU HAMILTON OSHIRO**  
**UM AMBIENTE INTELIGENTE PARA RECUPERAÇÃO DE IMAGENS**

Dissertação de Mestrado apresentada ao Curso de Mestrado em Sistemas e Computação do Instituto Militar de Engenharia, como requisito parcial para obtenção do título de Mestre em Sistemas e Computação.

Orientador: Prof. Alex Vasconcelos Garcia - D.Sc.

Orientador: Prof. Ronaldo Ribeiro Goldschmidt - D.Sc.,

Aprovada em 21 de Julho de 2008 pela seguinte Banca Examinadora:

---

Prof. Alex Vasconcelos Garcia - D.Sc. do IME - Presidente

---

Prof. Ronaldo Ribeiro Goldschmidt - D.Sc., do FAETEC/RJ

---

Prof. Emmanuel Piseces Passos - D.Sc., do PUC/RJ

---

Profa. Maria Cláudia Reis Cavalcanti - D.Sc., do IME

Rio de Janeiro  
2008

Dedico esta à minha mãe. Pelo seu apoio incondicional, seu exemplo, sua luta e seu amor. De um filho que sempre deu trabalho e quem sabe um pouco de orgulho.

*In Memoriam* a minha Oba, que nos deixou grandes ensinamentos e ótimas demonstrações de superação, garra e força.

## AGRADECIMENTOS

Agradeço a todas as pessoas que contribuíram com o desenvolvimento desta dissertação de mestrado, tenha sido por meio de críticas, idéias, apoio, incentivo ou qualquer outra forma de auxílio. Em especial, desejo agradecer às pessoas citadas a seguir.

A minha família, que sempre me deu suporte, até nos momentos mais difíceis das minhas vidas. Sempre pude contar com o meus pais, meus irmãos, meus tios e meus primos. Vão sempre representar a árvore fundamental de minha vida.

Aos meus amigos, que estão esparramados pelo quatros cantos desse Brasil, mas que sempre vão ter minha consideração. Alguns sentirei muita falta ao meu lado, pois aprendi muito com essas pessoas e espero que a vida nos dê chance de nos esbararmos para ter aquela excelente conversa.

Aos meus colegas de mestrado, que compartilharam de momentos sacrificantes e provatórios dessa etapa de nossas vidas. Conheci pessoas extraordinárias, que procurarei dentro de mim espelhá-las.

Aos meus companheiros de república, que conviveram comigo nessa estadia no Rio de Janeiro. Foram pessoas fantásticas que me fizeram evoluir como pessoa, um aprendizado que não trocaria por nada.

Aos meus oponentes de jogos, os quais me acolheram quando estava muito solitário e indeciso. Homens ímpares que me ajudaram e muito a acostumar com o Rio de Janeiro e me aconselham com sua vivência e experiência.

Aos meus colegas de trabalho, que partilharam dos meus sonhos, das minhas frustrações e das concretizações. Pessoas das mais diversas opiniões e posturas, que me despertaram para a importância da inteligência emocional.

Aos órgãos de fomentos CNPq e CAPES, que me auxiliaram e muito com a implementação das bolsas de estudo, com as quais não conseguiria terminar o mestrado sem esta ajuda de custo.

Aos grandes pesquisadores, Christian Paz-Trillo e Claudio Passos, que me ajudaram no desenvolvimento desse sistema, com a colaboração e compartilhamento de alguns conhecimentos. Talvez sem essa ajuda não seria possível a realização deste trabalho.

Por fim, a todos os professores e funcionários do Departamento de Engenharia de Sistemas (SE/8) do Instituto Militar de Engenharia.

Espero não ter esquecido de ninguém, pois acredito que a gratidão seja a base para todas as outras qualidades de um homem.

*Seimou Hamilton Oshiro*



”Uma inteligência modesta aliada a muito trabalho, frequentemente pode mais que uma inteligência brilhante e vadia”

## SUMÁRIO

LISTA DE ILUSTRAÇÕES .....	11
LISTA DE TABELAS .....	13
LISTA DE ABREVIATURAS E SÍMBOLOS .....	15
<b>1 INTRODUÇÃO .....</b>	<b>18</b>
1.1 Posicionamento e Motivação .....	18
1.2 Objetivos e Contribuições Esperadas .....	19
1.3 Organização do Texto .....	20
<b>2 FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>21</b>
2.1 Considerações Iniciais .....	21
2.2 Ontologias .....	21
2.3 Linguagens para Representação de Ontologias .....	23
2.3.1 RDF .....	23
2.3.2 OWL .....	24
2.4 Recuperação de Informação .....	27
2.4.1 Introdução .....	27
2.4.2 Processo de Indexação .....	28
2.4.2.1 Remoção de Afixos .....	29
2.4.2.2 Expansão de termos do índice .....	30
2.4.2.3 Cálculo dos Pesos .....	30
2.4.3 Processo de Recuperação .....	30
2.5 Processamento de Imagens .....	32
2.5.1 Representação de Imagens .....	32
2.5.2 Descritores de Imagens .....	33
2.5.2.1 Descritores de Cor .....	35
2.5.2.2 Descritores de Textura .....	42
<b>3 TRABALHOS RELACIONADOS .....</b>	<b>54</b>
3.1 Recuperação de Imagens Baseada em Sentenças .....	54
3.1.1 OnAIR .....	54

3.1.2	Photo-RDF .....	55
3.1.3	Medusa .....	56
3.2	Recuperação de Imagens Baseada em Conteúdo .....	57
3.2.1	BIRAM .....	58
3.2.2	QBIC .....	59
3.2.3	VisualSEEK .....	61
3.2.4	Photobook .....	61
3.3	Análise Comparativa .....	63
<b>4</b>	<b>ABORDAGEM PROPOSTA .....</b>	<b>65</b>
4.1	Considerações Iniciais .....	65
4.2	Recuperação de Imagens Baseada em Sentenças .....	65
4.2.1	Pré-Processamento .....	68
4.2.2	Indexação .....	70
4.2.3	Recuperação .....	71
4.3	Recuperação de Imagens Baseada em Conteúdo .....	76
4.3.1	Pré-Processamento .....	78
4.3.2	Indexação .....	82
4.3.3	Recuperação .....	82
4.4	Combinação das Abordagens de Recuperação de Imagens .....	84
<b>5</b>	<b>PROTÓTIPO, EXPERIMENTOS E RESULTADOS .....</b>	<b>86</b>
5.1	Protótipo .....	86
5.1.1	Módulo da Recuperação Baseada em Sentença .....	86
5.1.2	Módulo da Recuperação Baseada em Conteúdo .....	90
5.1.3	Módulo de Recuperação de Imagens Baseado em Sentenças e em Conteúdo ..	91
5.2	Avaliação do Sistema .....	92
5.2.1	Metodologia dos Testes .....	92
5.2.2	Contexto da Aplicação .....	93
5.2.3	Metodologia de Testes para Recuperação Baseada em Sentenças .....	93
5.2.4	Metodologia de Testes para Recuperação Baseada em Conteúdo .....	95
5.2.5	Metodologia de Testes para Recuperação Baseada em Sentenças e Recuperação Baseada em Conteúdo .....	96
5.3	Análise dos Resultados .....	96

5.3.1	Avaliação da Recuperação Baseada em Sentenças .....	96
5.3.2	Avaliação da Recuperação Baseada em Conteúdo .....	103
5.3.3	Avaliação da Recuperação Baseada em Sentenças e em Conteúdo .....	105
<b>6</b>	<b>CONSIDERAÇÕES FINAIS</b> .....	108
6.1	Retrospecto .....	108
6.2	Contribuições Obtidas .....	109
6.3	Trabalhos Futuros .....	110
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	111
<b>8</b>	<b><u>ANEXOS</u></b> .....	115
8.1	Anexo A .....	115
8.2	Anexo B .....	115

## LISTA DE ILUSTRAÇÕES

FIG.2.1	Grafo de RDF .....	24
FIG.2.2	Modelo de representação de RDF .....	24
FIG.2.3	Namespaces declarados em uma ontologia OWL .....	25
FIG.2.4	Cabeçalho OWL .....	26
FIG.2.5	Declaração de classe e subclasse em OWL .....	26
FIG.2.6	Declaração de propriedade em OWL .....	26
FIG.2.7	Visão Geral de um Sistema de Recuperação de Informação .....	27
FIG.2.8	Índice Invertido .....	29
FIG.2.9	(a) imagem binária, (b) imagens em níveis de cinza e (c) imagem colorida onde cada pixel é composto por uma tupla do sistema de cores RGB .....	33
FIG.2.10	Visão Conceitual de um descritor de imagem .....	34
FIG.2.11	Cubo RGB .....	37
FIG.2.12	Imagem colorida e seu histograma de cores .....	39
FIG.2.13	Modelagem conceitual do processo de cálculo de similaridades entre imagens utilizando-se Histogramas de Cor .....	41
FIG.3.1	Diagrama do Sistema Photo-RDF .....	56
FIG.3.2	Entrada de dados do sistema photo-RDF .....	57
FIG.3.3	- Meta-informações do algoritmo((MORENO, 2006)) .....	58
FIG.3.4	Interface de Especificação de Busca de uma Determinada Imagem((MORENO, 2006)) .....	59
FIG.3.5	Interface de apresentação de resultados do sistema Query By Image Content - QBIC .....	60
FIG.3.6	Tela de especificação de consulta por disposição espacial do Sistema VisualSEEK .....	62
FIG.3.7	Apresentação de Resultados de Recuperação de Imagens do Sistema Photobook .....	63
FIG.4.1	Visão Geral de um Sistema de Recuperação de Informação .....	66
FIG.4.2	Visão Geral da Recuperação baseada por Sentenças .....	67
FIG.4.3	Remoção das Palavras de Parada .....	68

FIG.4.4	Tratamento de Sinônimos .....	69
FIG.4.5	Remoção de Afixos - Fonte: (ORENGO, 2001) .....	70
FIG.4.6	Processo de Indexação .....	71
FIG.4.7	Processo de Recuperação .....	72
FIG.4.8	Algoritmo de Expansão da Consulta (PAZ-TRILLO, 2001) .....	73
FIG.4.9	Algoritmo de Similaridade de Termos (PAZ-TRILLO, 2001) .....	74
FIG.4.10	Algoritmo de Recuperação (PAZ-TRILLO, 2001) .....	75
FIG.4.11	Recuperação Baseada em Conteúdo - Visão Geral .....	77
FIG.4.12	Pré-Processamento da Imagem .....	79
FIG.4.13	Algoritmo para Cálculo de Descritores .....	79
FIG.4.14	Algoritmo para cálculo de descritores de textura .....	80
FIG.4.15	Algoritmo para cálculo de descritores de Cor .....	81
FIG.4.16	Algoritmo para cálculo de descritores de Cor .....	82
FIG.4.17	Processo de Indexação .....	83
FIG.4.18	Processo de Recuperação .....	83
FIG.4.19	Algoritmo para cálculo de similaridade de duas imagens .....	84
FIG.4.20	Recuperação Combinando as Duas Abordagens .....	85
FIG.5.1	Associação das Imagens .....	87
FIG.5.2	Geração dos Arquivos de Índice Invertido .....	88
FIG.5.3	Recuperação de Imagens .....	90
FIG.5.4	Resultado da Recuperação .....	91
FIG.5.5	Resultado da Recuperação .....	92
FIG.5.6	K-NN para $K = 1$ .....	96

## LISTA DE TABELAS

TAB.2.1	Histograma quantificado a partir da figura 2.11 .....	39
TAB.2.2	Histograma de Cor da Imagem A .....	40
TAB.2.3	Histograma de Cor da Imagem D1 .....	40
TAB.2.4	Histograma de Cor da Imagem D2 .....	40
TAB.2.5	Cálculo da distância L1 entre a imagem exemplo e a imagem D1 do banco de dados de imagens .....	41
TAB.2.6	Cálculo da distância L1 entre a imagem exemplo e a imagem D2 do banco de dados de imagens .....	42
TAB.2.7	Matriz de Co-Ocorrência, cada elemento da matriz diz respeito às transições de níveis de cinza .....	44
TAB.2.8	Imagem 3x3 pixels e com quatro níveis de cinza .....	44
TAB.2.9	Imagem 4x4 pixels e com quatro níveis de cinza .....	46
TAB.2.10	Relacionamento entre os níveis de cinza da imagem A .....	46
TAB.2.11	Relacionamento de vizinhança entre os tons de cinza da imagem A com $d \leq 1$ e $\theta=0^\circ$ . .....	47
TAB.2.12	Relacionamento de vizinhança entre os tons de cinza da imagem A com $d \leq 1$ e $\theta=90^\circ$ . .....	47
TAB.2.13	Cálculo do SMA para $\theta = 0$ e $d \leq 1$ . .....	48
TAB.2.14	Cálculo do SMA para $\theta = 90^\circ$ e $d \leq 1$ . .....	49
TAB.2.15	Cálculo do ENTROPIA para $\theta = 0^\circ$ e $d \leq 1$ . .....	50
TAB.2.16	Cálculo do ENTROPIA para $\theta = 90^\circ$ e $d \leq 1$ . .....	51
TAB.2.17	Cálculo da CONTRASTE para $\theta = 0^\circ$ e $d \leq 1$ . .....	52
TAB.2.18	Cálculo da CONTRASTE para $\theta = 90^\circ$ e $d \leq 1$ . .....	53
TAB.3.1	Tabela Comparativa dos Trabalhos Relacionados .....	63
TAB.3.2	Mnemônicos da Comparação dos Trabalhos .....	64
TAB.5.1	Tabela de Distribuições .....	94
TAB.5.2	Precisão para as Primeiras 5 Figuras Recuperadas .....	97
TAB.5.3	Precisão para as Primeiras 10 Figuras Recuperadas .....	98
TAB.5.4	Cobertura para as Primeiras 5 Figuras Recuperadas .....	99
TAB.5.5	Cobertura para as Primeiras 10 Figuras Recuperadas .....	100

TAB.5.6	Medida F para as Primeiras 5 Figuras Recuperadas .....	101
TAB.5.7	Medida F para as Primeiras 10 Figuras Recuperadas .....	102
TAB.5.8	Taxa de Erro para as Imagens Classificadas .....	104
TAB.5.9	Taxa de Erro para as Imagens Classificadas .....	106
TAB.5.10	Consultas Seleccionadas para essa Recuperação .....	107
TAB.5.11	Imagens Associadas a Consulta .....	107
TAB.5.12	Precisão para 40 Imagens Recuperadas .....	107
TAB.8.1	Consultas Seleccionadas para essa Recuperação .....	116



## LISTA DE ABREVIATURAS E SÍMBOLOS

### ABREVIATURAS

IME	-	<i>Instituto Militar de Engenharia</i>
RI	-	<i>Recuperação de Informação</i>
CBIR	-	<i>Content-Based Image Retrieval</i>
IST	-	<i>Instituto Superior de Tecnologia</i>
RDF	-	<i>Resource Description Framework</i>
RDFS	-	<i>Resource Description Framework-Schema</i>
OWL	-	<i>Web Ontology Language</i>
W3C	-	<i>World Wide Web Consortium</i>
XML	-	<i>Extensible Markup Language</i>
DAML+OIL	-	<i>DARPA Agent Markup Language + Ontology Inference Layer</i>
URI	-	<i>Universal Resource Identifier</i>
INCOR	-	<i>Instituto do Coração</i>
FAETEC	-	<i>Fundação de Apoio à Escola Técnica</i>
JNI	-	<i>Java Native Interface</i>
TREC	-	<i>Text Retrieval Conference</i>
DLL	-	<i>Dynamic-Link Library</i>
API	-	<i>Application Programming Interface</i>
RGB	-	<i>Red Green Blue</i>
SMA	-	<i>Segundo Momento Angular</i>
GLCM	-	<i>Grey Level Co-Occurrence Matrix</i>
MIT	-	<i>Massachusetts Institute of Technology</i>
CGI	-	<i>Common Gateway Interface</i>
QBIC	-	<i>Query by Image Content</i>
IBM	-	<i>International Business Machine</i>
JDBC	-	<i>Java Database Connectivity</i>
BIRAM	-	<i>Base de Imagens Relacional de Algoritmos e Métricas</i>
DC	-	<i>Dublin Core</i>

## RESUMO

Com o crescimento no volume de informações, devido aos constantes avanços na área da Tecnologia de Informação, a Recuperação de Informação(RI) vem crescendo muito. RI visa desenvolver mecanismos computacionais que localizem e recuperem informações. Este presente trabalho vem abordar a Recuperação de Imagens, que se mostra de grande importância na atualidade, devido aos grandes acervos de imagens nas mais diferentes áreas e aplicações. Serão abordadas a recuperação de imagens baseada em sentenças(que faz a recuperação por meio de uma sentença em linguagem natural), recuperação de imagens baseada em conteúdo(que faz a recuperação por meio de características de imagens) e é proposta uma combinação dessas abordagens.

Um protótipo foi desenvolvido para atender o IST-RIO, unidade de ensino superior da FAETEC, com intuito de preservar a história da instituição e ser acoplado ao sistema de gestão de conhecimento desenvolvido pelo instituto. Os resultados obtidos com este protótipo foram avaliados pela sua efetividade, ou seja, por resultados relevantes. Ao fim, foram apresentadas conclusões sobre as abordagens de recuperação de imagens avaliadas e propostas sugestões de trabalhos futuros.

## ABSTRACT

With the increasing volume of information being divulged as a result of the constant advances in the area of Information Technology, Information Retrieval (IR) is consequently one of the areas that has grown considerably. IR aims at developing computational mechanisms that locate and retrieve information. This present work seeks to dwell on Image Retrieval, which has proven to be of paramount importance today, due to the large quantities of images in the most diverse areas and applications. The proposition of a fusion of Annotation-based image retrieval (which takes place in natural language sentences) and Content-based image retrieval will be discussed.

A prototype was developed in order to provide support for IST-RIO, a higher education unit at FAETEC, and in so doing, preserve the history of the institution and connect to the knowledge management system developed by the institute. The results obtained by means of this prototype were evaluated on its effectiveness, that is, on its excellent results. Finally, conclusions were drawn on evaluated image retrieval approaches and suggestions for future projects proposed.

# 1 INTRODUÇÃO

## 1.1 POSICIONAMENTO E MOTIVAÇÃO

Os constantes avanços na área da Tecnologia da Informação, ocorridos sobretudo na última década, têm propiciado um crescimento exponencial do volume de informações armazenadas. Estas informações têm se apresentado em diferentes mídias e formatos tais como textos, imagens, vídeos, hipertextos, dentre outros. Diante deste cenário, a área da Recuperação de Informação (RI) vem despertando grande interesse tanto na academia quanto na indústria (TORRES, 2007). Esta área consiste essencialmente em se desenvolver e aplicar mecanismos computacionais que permitem localizar e recuperar informações de maneira eficaz (resultados relevantes) e eficiente (rapidamente).

Mais especificamente, no contexto de recuperação de informações sob a forma de imagens, foco da presente dissertação, são inúmeras as áreas onde a quantidade de imagens armazenadas vem aumentando significativamente nos últimos anos (TORRES, 2006). Mecanismos para recuperar imagens específicas a partir de grandes acervos de imagens têm se mostrado de grande utilidade em diversas áreas e aplicações tais como: Medicina, Sensoriamento Remoto, Segurança, Agricultura, Biologia, História, Jornalismo, dentre outros (TORRES, 2006).

Em síntese, dois tipos de abordagens têm sido utilizados na recuperação de imagens (TORRES, 2007):

- Recuperação de imagens baseada em Sentenças (PAZ-TRILLO, 2001)(PHOTO-RDF, 2002);
- Recuperação de imagens baseada em Conteúdo (MORENO, 2006)(SMITH, 1996)(ZHANG, 2002).

Na recuperação de imagens baseada em Sentenças, o usuário pode consultar um banco de imagens através de uma sentença em linguagem natural, como é feito em uma busca na Internet. O usuário tem liberdade de entrar com a sentença que lhe convier, podendo ser uma palavra apenas ou frases completas, cabendo ao sistema, que utiliza dessa recuperação, tratar esta sentença. Geralmente são usados textos descritivos, palavras-chave e ontologia para fazer a recuperação dessas imagens.

Na recuperação de imagens baseada em conteúdo, o usuário ao invés de entrar com uma sentença, vai escolher uma imagem, que vai servir de imagem-exemplo para fazer a consulta. Esta imagem passa por um tratamento que extrai características discriminantes tais como cor, textura e forma. Uma vez calculados os valores das características da imagem-exemplo, estes são utilizados na recuperação de imagens que sejam mais similares em relação à imagem apresentada (TORRES, 2007). Estes valores são conhecidos como descritores de imagens, sendo categorizados em cor, textura e forma.

Os sistemas que utilizam esta abordagem de conteúdo são mais conhecidos como sistemas de CBIR(Content-Based Image Retrieval). Exemplares desse tipo de sistema são: BIRAM, QBIC, VisualSeek e Photobook, entre outros. Maiores detalhes a respeito desses sistemas estão explicitadas no capítulo sobre trabalhos relacionados.

As abordagens mencionadas acima podem ser encontradas na literatura de forma isolada, não tendo sido identificada nenhuma tentativa de combinação entre elas. Cabe ressaltar, no entanto, que a existência de mecanismos que permitam recuperar imagens a partir de sentenças textuais e de exemplos de imagens, podem ser de grande utilidade em inúmeras aplicações práticas (TORRES, 2006).

## 1.2 OBJETIVOS E CONTRIBUIÇÕES ESPERADAS

Diante do exposto na seção anterior, esta dissertação tem como objetivos: pesquisar, propor, desenvolver e avaliar um ambiente computacional inteligente que possa ser utilizado em processos envolvendo recuperação de imagens. Esse ambiente deverá permitir a recuperação de imagens em três cenários: (a) consultas formuladas em linguagem natural; (b) recuperação de imagens a partir de exemplos; (c) combinação dos cenários anteriores.

Considerando o bom desempenho da utilização de ontologias em aplicações de recuperação de informação baseada em texto (GUARINO, 1999), o ambiente proposto deverá ser baseado no conceito de ontologias para orientar a anotação de características e auxiliar na interpretação de consultas em linguagem natural. Deverão também ser utilizados parâmetros quantitativos para representar cores e texturas das imagens e para a recuperação por conteúdo baseada em conceitos de similaridade.

A contribuição esperada é conseguir desenvolver um ambiente que possa tratar das recuperações das imagens a contento, com o auxílio dessas duas abordagens. Esperando também solucionar um problema do IST(Instituto Superior de Tecnologia)-Rio, que se

encontra com um acervo considerável de imagens e não conseguem gerir de maneira conveniente.

### 1.3 ORGANIZAÇÃO DO TEXTO

O texto da dissertação segue disposto em mais cinco capítulos. O capítulo 2 trata da recuperação de informação, contendo a fundamentação teórica para o desenvolvimento deste trabalho. O capítulo 3 trata dos trabalhos relacionados, descrevendo em detalhes como a questão da recuperação de imagens foi abordada em outras pesquisas. O capítulo 4 trata da abordagem proposta onde é apresentada conceitualmente a solução adotada para o problema. O capítulo 5 descreve o protótipo, a implementação e a avaliação do sistema, apresentando algumas considerações a respeito de seu desempenho. E, finalmente, o capítulo 6 que traz as considerações finais contendo uma conclusão sobre todo o trabalho e alternativas de trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 CONSIDERAÇÕES INICIAIS

O presente capítulo tem como objetivos descrever os principais pontos de fundamentação teórica necessária à formulação e ao desenvolvimento da abordagem proposta pela presente dissertação.

A seção 2.2 apresenta definições e conceitos sobre ontologias, necessários à recuperação de imagens baseada em sentenças. Ainda nesta seção são apresentados de forma resumida as linguagens RDF e OWL para representação de ontologias.

A seção 2.4 descreve os principais pontos usualmente envolvidos nos processos de armazenamento e recuperação de informação.

Detalhes sobre processamento de imagens necessários à formulação e à implementação da recuperação de imagens baseada em conteúdo estão descritos na seção 2.5.

### 2.2 ONTOLOGIAS

Ontologia, segundo Karin Breitman (BREITMAN, 2005), vem do grego *ontos* (ser) + *logos* (palavra). A definição do dicionário Merriam Webster, disponível on line no site [www.m-w.com](http://www.m-w.com) é:

- a) Ramo da metafísica dedicado à natureza e aos relacionamentos do ser;
- b) Teoria sobre a natureza do ser ou daqueles que existem.

Na Filosofia, ontologia é a ciência do que é, dos tipos de estruturas dos objetos, propriedades, eventos, processos e relacionamentos em todas as áreas da realidade. Como disciplina da área de Filosofia, a Ontologia tem como objetivo criar sistemas de classificação para facilitar a organização da realidade.

A definição mais citada de ontologia na literatura da Web Semântica, é a proposta por (GRUBER, 1993), a qual define ontologia como: "Uma especificação formal e explícita de uma conceitualização compartilhada". Neste contexto, "conceitualização" quer dizer que a ontologia representa um modelo abstrato de algum fenômeno e os conceitos relevantes para o mesmo. "Explícita" quer dizer que os elementos e suas restrições estão claramente

definidos, a expressão "formal" significa que a ontologia deve ser definida de forma a permitir um processamento automático, e "compartilhada" significa que uma ontologia representa um conhecimento definido por um consenso de pessoas.

O consórcio W3C(World Wide Web Consortium) define uma ontologia como "a definição dos termos utilizados na descrição e na representação de uma área de conhecimento". Este mesmo órgão coloca que ontologias devem prover descrições para os seguintes tipos de conceito:

- Classes (ou "coisas") nos vários domínios de interesse;
- Relacionamentos entre essas "coisas";
- Propriedades (ou atributos) que essas "coisas" devem possuir.

Levando-se em consideração estas definições de ontologia, este trabalho considerará ontologia como um vocabulário estruturado, que possui relações e restrições entre seus termos.

Tais relacionamentos podem ser hierárquicos ou não. Embora os hierárquicos sejam os mais comuns, pode-se definir uma relação de afinidade, por exemplo, a qual diria que o homem tem afinidade com futebol, ou que a disciplina de Matemática tem afinidade com a disciplina de Lógica.

Como já foi citado, além de relacionamentos, as ontologias possuem também restrições, as quais são definidas como axiomas. Por exemplo, numa ontologia de animais, pode-se ter um subgrupo de humanos, o qual teria como restrição as capacidades de falar e raciocinar, pois todo ser humano tem a capacidade de falar e raciocinar, e é exatamente isso que o difere dos demais animais.

Após ter sido feita a definição de uma ontologia, se torna possível deduzir alguns conceitos com base em outros já existentes, visto que os relacionamentos existentes entre os termos tornam possível descobrir termos semelhantes dentro do vocabulário. Por exemplo, uma ontologia que considere um relacionamento "ser parente de" como sendo mais geral do que "ser primo". Sendo assim, se soubermos que João é primo de José, podemos deduzir que João também é parente de José, sem que isto esteja definido de forma explícita. Isto quer dizer que não seria necessário entrar com a informação de que João é parente de José, uma vez que já se sabe que ele é primo de José.



## 2.3 LINGUAGENS PARA REPRESENTAÇÃO DE ONTOLOGIAS

Partindo do pressuposto de que XML (Extensible Markup Language)(XML, 2006) só expressa sintaxe, foram criadas linguagens de representação de ontologias para que fosse possível modelar um domínio descrevendo não só seu vocabulário, mas também relacionamentos entre os termos e ainda possibilitar algum poder de inferência. Entre as linguagens de representação criadas, pode-se destacar a evolução de linguagens como RDF (Resource Description Framework) (RDFS, 2004), RDF-S (Resource Description Framework-Schema)(RDFS, 2004), DAML+OIL (DARPA Agent Markup Language + Ontology Inference Layer)(DAML+OIL, 2001) e OWL (Web Ontology Language)(OWL, 2004). Todas têm por base a idéia proveniente de RDF, onde o "mundo" é descrito por meio de declarações na forma de triplas, ou seja, as definições possuem dois objetos e uma relação entre eles. O RDF e a OWL serão explicados mais detalhadamente nas próximas seções.

### 2.3.1 RDF

O objetivo do RDF é fornecer a interoperabilidade e a semântica para metadados de uma forma que facilite a busca por recursos na *Web*. Antes da especificação do mesmo, a qual foi feita pela W3C em fevereiro de 1999, os recursos podiam ser procurados somente por mecanismos de busca textuais simples.

Os três princípios fundamentais do RDF são: recursos, propriedades e frases. Os recursos são os objetos dos quais se quer fazer alguma definição. Estes possuem um identificador denominado URI (*Universal Resource Identifier*) - pode ser um endereço na *Web* - e podem ser exemplificados por pessoas, livros, cidades, etc. As propriedades são as descrições dos relacionamentos entre estes recursos, "editado por", "sexo", "localização". A junção de dois recursos, sendo eles um sujeito e outro objeto, e a definição de relacionamento entre eles, forma a frase.

Por exemplo, na frase "João escreveu *Web Semântica*", poderia ser definida como uma fórmula lógica  $P(x,y)$ , onde o predicado binário "escreveu" relaciona o objeto  $x$  ao objeto  $y$ . Neste caso, o sujeito seria *Web Semântica*, a propriedade seria a característica "escrito por" e o objeto seria o valor (objeto) João.

Outra forma de representar esta frase é a utilização de um grafo, a qual é a forma mais usual de representar sentenças em RDF. Este é um grafo direcionado, com os arcos

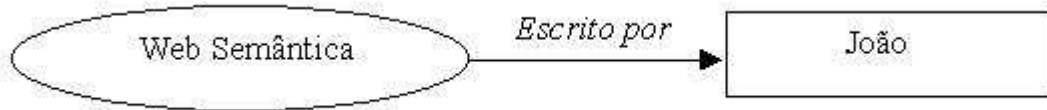


FIG. 2.1: Grafo de RDF

```

<rdf:RDF>
  <rdf:Description about: "Web Semântica">
    <f:escrito>
      João
    </f:escrito>
  </rdf:Description>
</rdf:RDF>
  
```

FIG. 2.2: Modelo de representação de RDF

apontando do recurso para o valor (do sujeito para o objeto). Como foi mostrado na figura 2.1.

Existe uma outra forma de representação para esta mesma sentença do exemplo acima, a qual se baseia em uma sintaxe XML. Neste modelo de representação, um documento RDF utiliza um elemento XML com etiqueta `rdf:RDF`. As descrições desse elemento, utilizam a etiqueta `rdf:Description`. Como por exemplo, na figura 2.2.

### 2.3.2 OWL

Com base em uma revisão da linguagem DAML+OIL(DAML+OIL, 2001), acrescentadas de características aprendidas no desenvolvimento e nas aplicações desta, criou-se a OWL. Esta linguagem é ideal para representação de informações contidas em documentos que precisam ser processados por aplicações, pelo fato de fornecer um vocabulário adicional juntamente com uma semântica formal. Além disso, com ela é mais fácil expressar a semântica. A OWL possui três sub-linguagens, que foram projetadas conforme o grupo de implementadores e usuários: OWL Lite, OWL DL e OWL Full(OWL, 2004).

OWL Lite - sua finalidade principal é dar suporte para que classificações hierárquicas com restrições simples sejam criadas. Por exemplo, restrições de cardinalidade são permitidas para apenas valores iguais a zero ou um.

```

xmlns:vin="http://www.w3.org/TR/2004/REC-owl-guide-20040210/wine#"
xmlns:food="http://www.w3.org/TR/2004/REC-owl-guide-20040210/food#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#">

```

FIG. 2.3: Namespaces declarados em uma ontologia OWL

OWL DL - sua finalidade é prover um maior grau de expressividade onde todas as conclusões são computáveis (completude) e todas as computações terminam em tempo finito. OWL DL é baseada na lógica descritiva.

OWL Full - sua finalidade é prover o máximo de expressividade e liberdade sintática. Com OWL Full, os usuários podem aumentar o vocabulário pré-definido de RDF ou OWL, porém sem nenhuma garantia computacional. Assim, máquinas de inferência poderão não derivar conclusões ou poderão não ser computáveis em tempo finito.

Uma ontologia OWL inicia com uma declaração, em RDF, contendo os namespaces que serão utilizados na ontologia. Basicamente, a ontologia começa da seguinte forma, como é demonstrado na figura 2.3.

As duas primeiras declarações indicam o namespace associado à ontologia que está sendo desenvolvida, a qual refere-se a vinhos e comidas. Na terceira linha, é indicado o namespace específico da OWL e nas linhas subseqüentes têm-se as referências aos vocabulários do RDF, RDF Schema e XML Schema, respectivamente.

Uma vez que os namespaces estão definidos, inclui-se, através da tag <owl:Ontology>, informações como comentários, controle de versões e inclusão de outras ontologias, caso esteja havendo extensão de uma já existente. Na Figura 2.4, apresenta-se um código referente a estas informações.

Têm-se na figura acima a representação de um cabeçalho OWL contendo informações diversas representadas por:

- rdfs:comment - apresenta algum comentário referente à ontologia modelada;
- owl:priorVersion - é uma tag padrão utilizada para o controle de versões da ontologia;
- owl:imports - indica a importação de uma ontologia já existente definida pelo recurso

```

<owl:Ontology rdf:about="">
  <rdf:comment>An example OWL ontology</rdf:comment>
  <owl:priorVersion rdf:resource="http://www.w3.org/2001/sw/WebOnt/guide-src/wine-112102.owl"/>
  <owl:imports rdf:resource="http://www.w3.org/2001/sw/WebOnt/guide-src/food.owl"/>
  ...

```

FIG. 2.4: Cabeçalho OWL

```

1. <owl:Class rdf:ID="Curso" />
2. <owl:Class rdf:ID="Graduação">
3.   <rdfs:subClassOf rdf:resource="#Curso" />
4.   ...
5. </owl:Class>

```

FIG. 2.5: Declaração de classe e subclasse em OWL

apresentado em seguida.

As classes e subclasses OWL são definidas, respectivamente, através dos elementos `owl:Class` e `rdfs:subClassOf`. Como por exemplo, na figura 2.5:

Na figura acima, tem-se a declaração de duas classes: `Curso` e `Graduação`. Sendo que, na linha 3, é identificado que `Graduação` é subclasse de `Curso`.

Através do elemento `owl:ObjectProperty` define-se as propriedades das classes e subclasses. Utiliza-se, em conjunto com a tag de propriedades, o elemento `rdfs:domain` para especificar a qual classe a referida propriedade está sendo atribuída. Como por exemplo na figura 2.6:

A propriedade `nome` está sendo declarada na primeira linha (figura acima) e identificada que pertence à classe `Aluno` na linha 2.

```

1. <owl:ObjectProperty rdf:ID="nome" />
2.   <rdfs:domain rdf:resource="#Aluno" />
3. </owl:ObjectProperty>

```

FIG. 2.6: Declaração de propriedade em OWL

A OWL oferece vários elementos como owl:Restriction, owl:cardinality, owl:minCardinality, owl:maxCardinality, owl:inverseOf, entre diversos outros, os quais permitem determinar, respectivamente, as restrições, cardinalidades, cardinalidade mínima, cardinalidade máxima, relação inversa, de modo a obter uma ontologia mais completa e melhor modelada.

## 2.4 RECUPERAÇÃO DE INFORMAÇÃO

### 2.4.1 INTRODUÇÃO

Segundo(JACKSON, 2002) Recuperação de Informação é a maneira com que é realizada a aquisição, organização, armazenamento, recuperação e distribuição da informação. Ou seja, o usuário através de uma consulta, tem acesso a objetos relevantes a sua necessidade. Tudo isso é realizado por meio de um sistema de Recuperação de Informação(RI).

A visão geral de um sistema de Recuperação de Informação é mostrada na figura 2.7. Onde a linha contínua representa o processo de indexação, a preparação das imagens para uma futura recuperação, e a linha tracejada indica o processo de recuperação.

A dificuldade maior desse sistema, mencionado anteriormente, é trazer documentos relevantes, separando aqueles que serão relevantes a seu usuário e ignorando os supérfluos. Essa situação ideal só é consolidada por meio de uma boa metodologia de teste e avaliação.

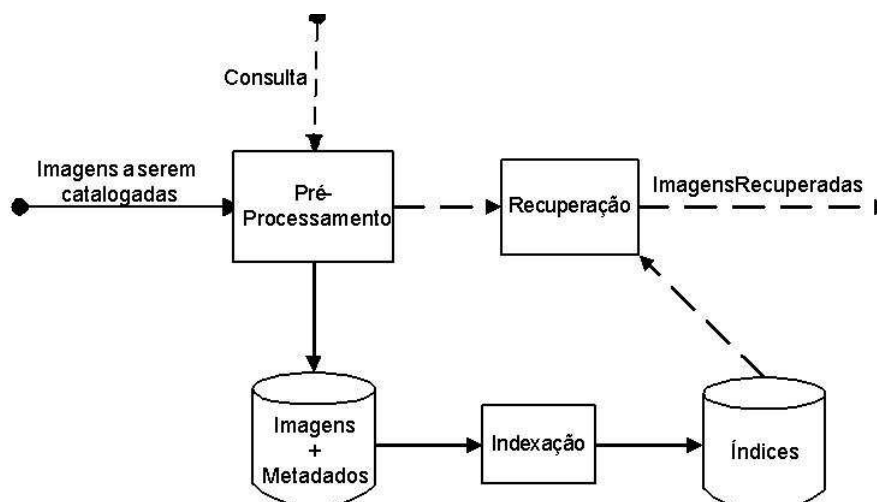


FIG. 2.7: Visão Geral de um Sistema de Recuperação de Informação

Os sistemas de busca booleana, em seu primórdio, eram feitos por intermédio das palavras-chave, ou seja, dada uma consulta, é procurado um documento que contenha as palavras desta consulta. Esse método é suscetível à falhas, como as citadas logo abaixo:

- Sinônimos: a busca fica restrita à palavra passada para a consulta. Se houver algum documento que contenha algum sinônimo de uma palavra da consulta, esse documento não é considerado pelo sistema;
- Polissemia: Caso uma palavra tenha vários significados, uma consulta esta palavra, pode ser ruim para um sistema de RI, pois são recuperados documentos com esta palavra, entretanto, não levando em conta os seus diferentes significados.

Todos esses problemas tratados anteriormente são conhecidos como "barreiras das palavras-chave" (MAULDIN, 1991). Um sistema de RI, é definido geralmente por duas etapas: processo de indexação e processo de recuperação. A primeira é responsável por preparar os documentos, através de um pré-processamento nas informações, para que depois elas sejam recuperadas pelo segundo processo.

#### 2.4.2 PROCESSO DE INDEXAÇÃO

Trata-se do primeiro processo de um sistema de RI. Este processo é composto de três componentes, listadas logo a seguir:

- a) Coleção de documentos: representa os documentos que serão recuperados pelo sistema;
- b) Dicionário: representa as palavras que ocorrem no banco de dados;
- c) Arquivo de índice invertido: representa as ocorrências da palavra do dicionário e permite o acesso ao banco de dados eficientemente.

Sendo que a coleção de documentos é uma entrada para o processo de indexação, enquanto que o dicionário e o arquivo de índice invertido compõem a sua saída.

Segundo (JACKSON, 2002), um arquivo de índice invertido é composto dos seguintes itens, listados na figura 2.8.

- *Token*: Pode ser uma palavra, ou um radical desta;
- Freqüência dos documentos: a quantidade de documentos em que o token está inserido;
- Freqüência total: a quantidade com que o token aparece na coleção de documentos;

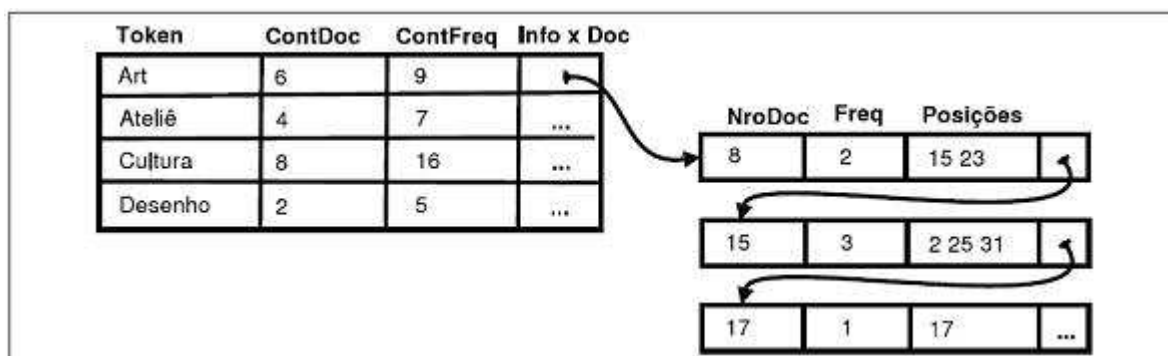


FIG. 2.8: Índice Invertido

- Informações do documento: as informações relativas ao documento em relação à ocorrência de um token. Cada documento referenciado por um *token*, é composto de dois itens: frequência e posição. A primeira representa a quantidade de vezes que o *token* aparece no documento e o segundo trata da(s) posição(ões) em que o *token* aparece no documento.

As palavras de parada (*stopwords*) são aqueles termos que aparecem demasiadamente em todo/qualquer texto, podendo ser desconsideradas em uma consulta, por não terem muita importância seletiva para indexação. Preposições, pronomes, artigos e alguns verbos, são exemplos de palavras de parada.

#### 2.4.2.1 REMOÇÃO DE AFIXOS

Remoção de afixos ou *stemming* é responsável em obter o radical das palavras, ou seja, removendo os afixos referentes, por exemplo, de plural, advérbio, aumentativo, etc. Utilizando esse processo, ganha em desempenho, pois dessa maneira, o armazenamento dos tokens é reduzido. Com este recurso, existe uma agrupamento de palavras com mesmo radical de palavras, nas consultas do sistema de RI.

O algoritmo de remoção de afixos geralmente se utiliza de heurísticas para extração destes. As heurísticas obedecem às regras estabelecidas, especificadas com o tipo de afixo a ser removido e as exceções às regras. Em (ORENGO, 2001) foi desenvolvido um stemmer para o português, com regras estabelecidas e executadas sequencialmente.

#### 2.4.2.2 EXPANSÃO DE TERMOS DO ÍNDICE

Na criação do arquivo de índice invertido, durante o processo de indexação, podem ser criados *tokens* com palavras que não se encontram na coleção de documentos, ou seja, expandindo os termos do índice. Estes podem ser obtidos através de uma estrutura associada à coleção de documentos, por exemplo, uma ontologia, mostrado anteriormente na seção 2.2. Entretanto, é preferível que esse processo seja feito durante o processo de recuperação, para que dessa maneira não seja sobrecarregado o desempenho do sistema de RI, através do aumento no índice.

#### 2.4.2.3 CÁLCULO DOS PESOS

Para definir a relevância de um documento, em uma dada consulta, são calculados pesos que relacionam os termos do arquivo de índice invertido e os seus documentos. Sendo estes calculados durante a consulta ou armazenados no arquivo de índice invertido, variando conforme o sistema de RI.

#### 2.4.3 PROCESSO DE RECUPERAÇÃO

É parte destinada à busca, feita pelo o usuário do sistema de RI. Esta etapa é dividida em quatro partes: o processamento da consulta, o cálculo da relevância dos documentos, ordenação destes documentos e apresentação ao usuário dos resultados.

Na literatura existem três tipos de enfoques (JACKSON, 2002) para recuperação: busca booleana, recuperação ordenada e recuperação probabilística.

##### **Busca Booleana**

A busca é feita antes da recuperação, e é o tipo de busca que realiza a consulta utilizando conectores lógicos(*and*, *or* e *not*), ou seja, conecta as palavras da consulta com esses conectores. Também é muito utilizada por motores de busca, que recuperam documentos referentes à consulta.

Podem ser usados nesta busca, recursos como expansão da consulta, esse processo é semelhante ao que foi apresentado na seção 2.4.2.2. Juntamente com a remoção de afixos, detalhadamente apresentada na seção 2.4.2.1.

Contudo, essa busca têm muitos limitantes. Dentre eles, a recuperação muito grande de documentos e desordenados, pois não considera a relevância dos dois. Seu desempenho é muito comprometido em coleções de documentos muito pequenas e muito



diferentes(JACKSON, 2002).

### **Recuperação Ordenada**

É um tipo de busca baseada na freqüência dos termos de uma consulta sobre uma coleção de documentos. O usuário entra com um consulta em linguagem irrestrita, e o sistema de RI trata das palavras de parada(*stopwords*) e remoção de afixos.

Os documentos e a consulta são vetores do espaço multidimensional, embasados nos termos da coleção de documentos. A freqüência de um termo é associada ao seu documento correspondente dentro de vetores. Através desses vetores é obtida a relevância entre a consulta e o documento, que no final, vai retornar os documentos em ordem de relevância.

### **O Modelo de Espaços Vetoriais**

Considere  $d_1, d_2, \dots, d_n$ , uma coleção de documentos,  $t_1, t_2, \dots, t_n$ , os termos da coleção, sendo que  $n$  define a dimensão dos vetores e  $.$ . O documento  $d$  é representado por um vetor  $d = (w_{1,d}, w_{2,d}, \dots, w_{m,d})$ . Onde  $w_{i,d}$  representa o peso do  $i$ -ésimo termo no documento, ou seja, quanto o termo é relevante para o documento. O mesmo vale para a consulta que é representado por  $q = (w_{1,q}, w_{2,q}, \dots, w_{m,q})$ , onde  $w_{i,q}$  representa o peso do  $i$ -ésimo termo na consulta.

O valor de similaridade pode ser calculado de diversas formas, uma delas é o produto escalar obtido por (SALTON, 1975), mostrada na equação 2.1:

$$sim(t, q) = \sum_{t_i \in q} w_{i,q} \times sim(t, t_i) \quad (2.1)$$

Podem existir casos em que os termos são muitos comuns. Nessa situação os valores dos pesos para esses termos podem ser recalculados, já que os termos são menos pertinentes para a consulta. Assim, pode-se usar o cálculo de pesos TF-IDF (JACKSON, 2002):

$$W_{t,d} = tf_{t,d} \times idf_{t,d} \quad (2.2)$$

onde,  $tf_{t,d}$ (term frequency) é a freqüência do termo  $t$  no documento  $d$ ,  $idf_{t,d} = \log(n/n_t + 1)$ (inverse document frequency),  $n_t$  é a quantidade de documentos em que o termo aparece, e mede a escassez do termo na coleção de documentos, assim obtem-se a noção da importância do termo na recuperação.

Deve-se ter uma maneira de calcular a similaridade, uma das formas utilizadas é a medida da similaridade do coseno(HARMAN, 1992). Esta calcula o ângulo cosseno dos vetores consulta e documento, sua fórmula é:

$$sim(d, q) = \frac{\sum_{t=1}^m w_{t,d} \times w_{t,q}}{\sqrt{\sum_{t=1}^m w_{t,d}^2 \times \sum_{t=1}^m w_{t,q}^2}} \quad (2.3)$$

Essa medida de similaridade é muito utilizada em muitos projetos e importantes de recuperação de informação.

### **Recuperação Probabilística**

A recuperação Probabilística é uma técnica de recuperação ordenada embasada na teoria da probabilidade, ou seja, na probabilidade do documento ser relevante a uma dada consulta(JACKSON, 2002)(DEERWESTER, 1999).

## **2.5 PROCESSAMENTO DE IMAGENS**

### **2.5.1 REPRESENTAÇÃO DE IMAGENS**

Uma imagem pode ser definida como sendo a representação visual de um objeto. Do ponto de vista matemático, uma imagem é considerada uma função bidimensional  $f(x,y)$  onde  $x$  e  $y$  são coordenadas planas. A amplitude de  $f$  em qualquer par de coordenadas  $(x,y)$  é chamada de intensidade ou nível de cinza da imagem no referido ponto. Quando  $(x,y)$  e a amplitude de  $f$  variam dentro de um conjunto de valores finitos, ou discretos, a imagem é chamada de imagem digital (GONZALEZ e WOODS, 2001).

Em outras palavras, uma imagem digital pode ser representada através de uma matriz  $n \times m$  onde cada elemento da matriz corresponde à intensidade ou nível de cinza  $f(x,y)$  em um determinado ponto da imagem. Estes elementos de imagens digitais são os pixels, que na prática possuem uma intensidade que representa uma cor em um determinado ponto da imagem. Para imagens binárias (em preto e branco) os valores dos pixels podem assumir os valores 0 e 1. Para imagens em tons de cinza estes valores podem variar de 0 a 255 e, finalmente, para imagens coloridas, tem-se que o valor do pixel é representado por três valores variando de 0 a 255 cada um.

Como exemplo, verifique a disposição de três imagens, na figura 2.9, compostas por 9 pixels (3 linhas e 3 colunas) bem como sua respectiva matriz de representação digital:

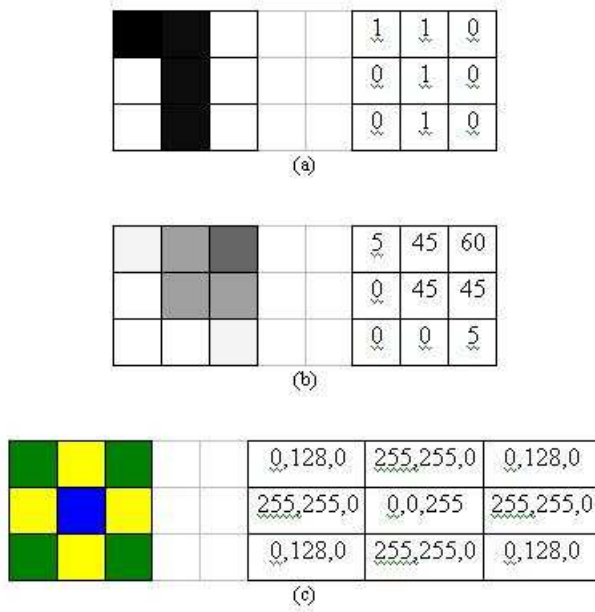


FIG. 2.9: (a) imagem binária, (b) imagens em níveis de cinza e (c) imagem colorida onde cada pixel é composto por uma tupla do sistema de cores RGB

## 2.5.2 DESCRITORES DE IMAGENS

Descritores de imagens podem ser definidos como um par composto por um vetor de características extraídas e uma função de distância. O vetor de características representa um conjunto de propriedades de uma determinada imagem (cor, textura e forma, por exemplo) enquanto que a função de distância diz respeito a dissimilaridades entre duas imagens tendo como base características das mesmas. Estas distâncias podem ser mensuradas através de métricas como a Euclidiana (TORRES, 2006).

Mais formalmente, um descritor de imagem  $D$  pode ser definido como uma tupla  $(\epsilon_D, \delta_D)$  onde  $\epsilon_D$  é a função que extrai um vetor de características  $\vec{v}$  de  $n$  dimensões de uma imagem -  $n$  é o número de características utilizadas para descrever a imagem em questão - e  $\delta_D$  é uma função de similaridade baseada em uma métrica de distância, que calcula a similaridade entre duas imagens.

A fim de ilustrarmos as formalizações apresentadas anteriormente, consideremos o exemplo da figura 2.10, onde temos uma visão conceitual do funcionamento de um descritor de imagens para análise de similaridades entre duas imagens.

Inicialmente, duas imagens  $I_Q$  e  $I_D$  são submetidas a uma função de extração de características  $\epsilon_D$  que extrai os vetores de características  $\vec{v}_{I_Q}$  e  $\vec{v}_{I_D}$  das respectivas imagens.

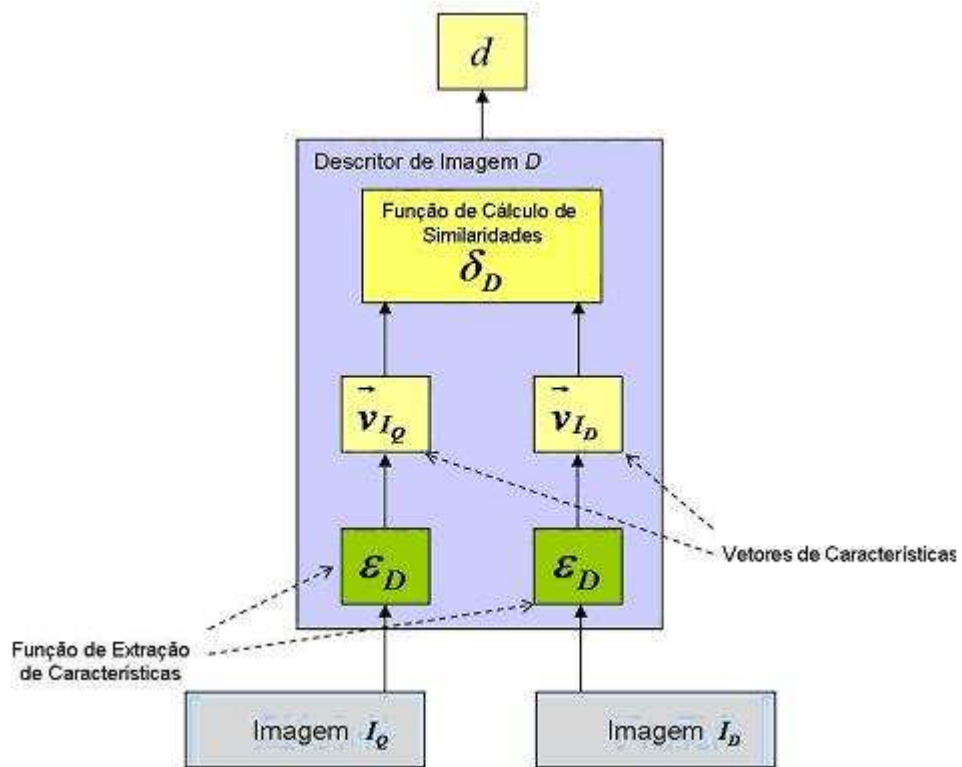


FIG. 2.10: Visão Conceitual de um descritor de imagem

Em seguida, a função  $\delta_D$  é utilizada para a obtenção do valor de similaridade  $d$  entre as imagens a serem comparadas.

Para o exemplo, podemos representar os vetores de características das imagens como:  $\vec{v}_{I_Q} = (a_1, a_2, \dots, a_{n-1}, a_n)$  e  $\vec{v}_{I_D} = (b_1, b_2, \dots, b_{n-1}, b_n)$  e a função de cálculo de similaridades, como o cálculo da distância entre as características de ambas as imagens:

$$\delta_D(I_Q, I_D) = \sqrt{\sum_{i=1}^n (b_i - a_i)^2} \quad (2.4)$$

Os modelos propostos para a recuperação de imagens por conteúdo dependem inteiramente do processo de extração de características. É a partir deste processo que imagens são indexadas na base de dados. Para a tarefa de recuperação de imagens, geralmente diversos descritores são combinados a fim de se aumentar a eficiência desta técnica.

Muitas são as propriedades que podem ser recuperadas a partir de uma imagem e estas podem ser características gerais ou de domínio específico. As características gerais são aquelas que podemos encontrar facilmente em qualquer imagem, tais como cor, textura e forma, bem como uma combinação destas. As características de domínio específico, como o próprio nome sugere, são aquelas que dependem da área de aplicação do sistema em questão. Como exemplo, podemos citar características referentes a impressões digitais e faces humanas. Para a presente dissertação serão abordadas apenas características de domínio geral, mais especificamente cor e textura.

Dentre as propriedades mais utilizadas para a representação em vetor de características podemos citar a cor e a textura. A seguir serão introduzidos conceitos a respeito de cada uma delas.

#### 2.5.2.1 DESCRITORES DE COR

Quando se utiliza a cor como característica para a recuperação de imagens, o objetivo é utilizá-la para recuperar imagens que possuam uma composição de cor similar a imagens da base de dados, mesmo que elas sejam de contextos diferentes. A cor é uma das características mais largamente utilizadas em sistemas de recuperação de imagens baseados em conteúdo (TORRES, 2006) por ser relativamente independente quanto ao tamanho, orientação e resolução da imagem e é computacionalmente menos cara quando comparada aos demais descritores (BENDER, 2003). Quando esta característica

é utilizada, podemos querer obter o seu histograma de cores, que representa a distribuição global de cores, ou considerar apenas uma determinada parte da imagem.

Para se entender como os descritores de cor são utilizados nos sistemas de recuperação de imagens é necessário definir previamente o conceito de cor e listar algumas de suas características. A cor pode ser definida como um fenômeno perceptual da luz quando incide e é refletida sobre uma superfície ou passa através de um meio, não totalmente opaco à luz como, por exemplo, um líquido ou um vidro colorido.

Como se pode observar, o conceito de cor é diretamente relacionado ao conceito de luz (GONZALEZ, 2001). Portanto, ao se referir ao conceito de cor deve-se mencionar, obrigatoriamente, a luz. A luz atravessa um determinado espaço, ora se comportando como onda, ora se comportando como partícula. Pode-se então definir luz como uma radiação eletromagnética que possui uma determinada energia para cada comprimento de onda. Para o estudo de cores é relevante apenas analisar o comportamento de luz como onda. Assim, de agora em diante neste texto, sempre que houver uma referência a este fenômeno, deve-se subentender, na verdade, a luz como onda eletromagnética.

Da interação entre a energia luminosa e o meio, a cor pode se formar através de (a) processo aditivo, (b) processo subtrativo e (c) formação por pigmentação (BENDER, 2003). No processo aditivo ocorre a combinação entre raios de luz com frequências diferentes através da soma de energia dos fótons. No processo de formação subtrativo a luz é transmitida através de um filtro que absorve a radiação luminosa de um determinado comprimento de onda. Alternativamente a luz também pode incidir através de um corante, que é constituído por filtros que podem absorver a radiação luminosa de um determinado comprimento de onda. Finalmente, no processo de formação por pigmentação os pigmentos podem absorver, refletir ou transmitir a radiação luminosa.

Dois conceitos são particularmente importantes para o entendimento do conceito de percepção de cor. São eles: a luminância e a cromaticidade. A luminância contém a informação da quantidade das cores pretas e brancas presentes em uma imagem. O cérebro humano interpreta essa informação como a quantidade de cinza presente na cor (ou brilho). A cromaticidade informa a respeito da tonalidade de uma cor. É a frequência dominante do raio de luz. A combinação destes dois conceitos, em diferentes proporções, permite ao cérebro perceber o espectro de cores visível em uma determinada imagem ou cena. Para representar as cores existem modelos ou sistemas de cores (GONZALES 2001), como por exemplo RGB ("red, green, blue"), CMY ("cyan, magenta, yellow"), HSI ("hue, saturation, intensity"), etc.

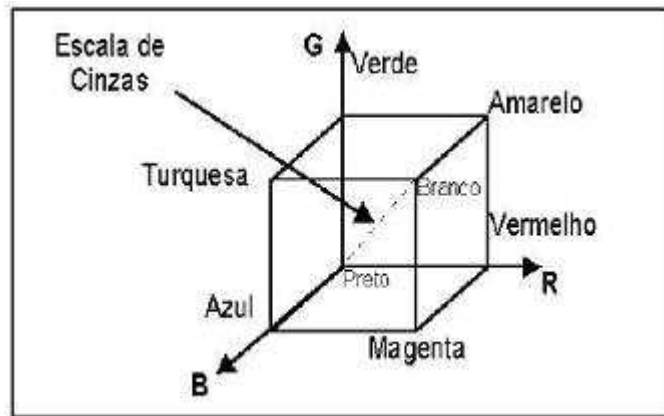


FIG. 2.11: Cubo RGB

saturação, intensidade”), YIQ(luminância, em-fase, quadratura), LAB(luminosidade, verde-vermelho, azul-amarelo) e MTM.

### RGB

O sistema Red, Green, Blue é um sistema de representação de cor (BIMBO, 1999) aditivo e se baseia na teoria dos três estímulos proposta por Young-Helmholtz (1983). Segundo esta teoria, o olho humano percebe a cor através do estímulo de três pigmentos visuais presentes nos cones da retina, que possuem sensibilidades para alguns comprimentos de onda, como por exemplo, 630 nanômetros (vermelho - red), 530 nanômetros (verde - green) e 450 nanômetros (azul - blue). Uma representação deste sistema pode ser formulada através de um cubo com os eixos R, G e B como ilustra a figura 2.11.

Podemos dizer que a origem representa a cor preta, os vértices das coordenadas (1,1,1) representam o branco, os vértices que estão sobre os eixos representam as cores primárias e os vértices restantes representam os complementos das primárias. É importante observar que cada ponto no interior do cubo corresponde a uma cor, representada por uma tupla  $(R, G, B)$ , com os valores de R, G e B variando de 0 a 1. Os tons de cinza são representados ao longo da diagonal principal do cubo, que se inicia no ponto de origem até o vértice que foi apontado como a cor branca. Então se pode afirmar que cada tom de cinza presente no cubo é formado por contribuições iguais das cores primárias que o compõem. Um exemplo de tom de cinza médio poderia ser representado pela tupla  $(0.5, 0.5, 0.5)$ . Algumas cores podem não ser reproduzidas através da sobreposição de três primárias uma vez que a resposta do olho humano a estímulos espectrais não é linear. Isso significa que algumas cores existentes na natureza não podem ser reproduzidas neste sistema. Por

isso, um fenômeno natural colorido como o processo de formação das rochas não pode ser reproduzido precisamente (BIMBO, 1999).

### **Histogramas de Cor**

Conforme visto anteriormente, a cor de um determinado objeto pode ser representada em diferentes espaços de cor, como por exemplo, no sistema RGB. É comum que uma determinada imagem possua diversas cores, portanto, descritores de cores normalmente representam a distribuição das cores da imagem levando ou não em consideração sua informação espacial. O processo de recuperação baseado em cores é largamente utilizado em sistemas de recuperação de imagens dentre outros motivos por que esta é uma característica local de identificação, e é um atributo independente da resolução da imagem. Dentre os descritores de cor podemos citar o Histograma de Cores que é uma característica que não leva em consideração a distribuição espacial das cores, e é o descritor mais largamente utilizado em recuperação de imagens (TORRES, 2006).

Na computação gráfica, um histograma de cor é a representação da distribuição de cores em uma imagem derivada da contagem dos pixels de cada subconjunto de cores de um espaço de cor tipicamente bidimensional ou tridimensional. Um histograma é uma descrição estatística padrão da distribuição em termos da frequência da ocorrência de diferentes classes de eventos; para cor, as classes ou regiões de evento são os espaços de cor. Pode-se afirmar que histogramas de cor são construções flexíveis que podem ser elaborados a partir de imagens em vários sistemas de cores de qualquer dimensão, como o RGB e etc. O histograma de uma imagem é produzido primeiramente através da amostragem das cores deste objeto em um número de índice de cores. Em seguida, é realizada a contagem do número de pixels da imagem em para cada índice.

O processo de formação de um histograma se baseia na construção de diversas pilhas também conhecidas como regiões de cores, uma para cada cor presente na imagem, dentro de um espaço de cores. O que ocorre a seguir é o somatório da ocorrência de uma cor na imagem, incrementando, assim, as pilhas já mencionadas. Como exemplo observe a figura 2.12 onde é ilustrada a formação de um histograma de cores baseado em um conjunto de nove pixels. Uma vez formado, o histograma deve agora ser normalizado, dividindo-se o valor total encontrado na pilha pelo número total de pixels. Desta forma encontraremos o percentual de ocorrência de uma cor em uma determinada imagem, como está ilustrado na tabela 2.1.

A grande vantagem da utilização do valor percentual de ocorrência de uma cor em



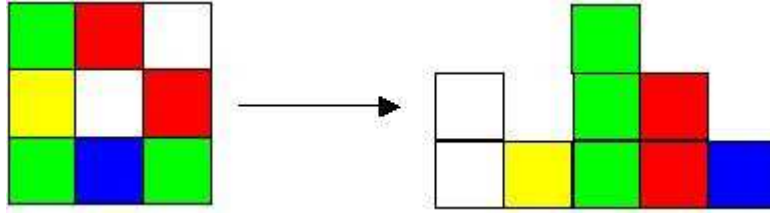


FIG. 2.12: Imagem colorida e seu histograma de cores

Cor	Pilha	Pixels	Percentual
Verde	3	9	33%
Vermelho	2	9	22%
Branco	2	9	22%
Azul	1	9	11%
Amarelo	1	9	11%

TAB. 2.1: Histograma quantificado a partir da figura 2.11

uma imagem está no fato de que o descritor encontrado pode ser confrontado com outros descritores de imagem independentemente da ocorrência de eventuais diferenças entre as dimensões das imagens a serem comparadas.

### Cálculo de Similaridade usando Histogramas de Cor

Existem diversos métodos de avaliação de similaridades entre imagens utilizando-se histogramas de cor; este mecanismo, além de ser o mais largamente utilizado em sistemas de Recuperação de Imagens Baseados em Conteúdo, é extremamente útil no momento de se aferir à frequência de cores em um objeto. Diversos métodos de cálculo de similaridade utilizam esta abordagem, que também se faz muito útil para procedimentos indexação de imagens e extração de outras características como a textura. Histogramas de Cores serão largamente utilizados neste projeto, tanto para indexação de imagens, quanto para a tarefa de recuperação.

A avaliação da similaridade de cores utilizando histogramas, pode ser realizada através do cálculo das distâncias  $L_1$  ou  $L_2$  entre duas imagens (BIMBO, 1999). As distâncias  $L_1$  e  $L_2$  entre o histograma de uma imagem exemplo  $H(I_Q)$  e os histogramas das imagens do banco de dados  $H(I_D)$  podem ser respectivamente definidas como:

$$D_H(I_Q, I_D) = \sum_{j=1}^n |(H(I_Q, j) - H(I_D, j))| \quad (2.5)$$

Onde o  $j$  é o índice de uma região no histograma (para histogramas de cores estas

Cor	Pixels Associados a Cor	Total de Pixels	Percentual
Verde	4	9	44%
Vermelho	3	9	33%
Branco	2	9	22%

TAB. 2.2: Histograma de Cor da Imagem A

Cor	Pixels Associados a Cor	Total de Pixels	Percentual
Verde	5	9	55%
Vermelho	3	9	33%
Branco	1	9	11%

TAB. 2.3: Histograma de Cor da Imagem D1

constituem a pilha de uma cor)  $I_Q$  é a imagem exemplo e  $I_D$  é uma imagem contida na base de dados de imagens. As imagens mais similares são aquelas que minimizam essas distâncias.

Para ilustrar os conceitos abordados considere uma imagem exemplo A que após o processo de extração de histogramas, apresenta os seguintes resultados na tabela 2.2.

Consideremos agora a existência de uma base de dados de imagens a ser utilizada neste experimento, que contenha as imagens D1 e D2, com os respectivos histogramas de cores extraídos nas tabelas 2.3 e 2.4 respectivamente:

A similaridade entre as imagens será dada pelo cálculo das distâncias entre a imagem A (imagem exemplo) e todas as imagens da base de dados, que para o exemplo em questão são D1 e D2. Todo processo de cálculo de similaridade é demonstrado na figura 2.13.

Finalmente é realizado o cálculo das distâncias L1 ou L2 entre o histograma da imagem exemplo e os histogramas das imagens presentes no banco de dados. Para este experimento, realizaremos apenas o cálculo da distância L1. Note que j será valorado como 3, uma vez que cores que não estão presentes no histograma da imagem A são irrelevantes para este cálculo de similaridade.

Logo:

Cor	Pixels Associados a Cor	Total de Pixels	Percentual
Verde	1	9	11%
Vermelho	6	9	66%
Branco	1	9	11%
Amarelo	1	9	11%

TAB. 2.4: Histograma de Cor da Imagem D2

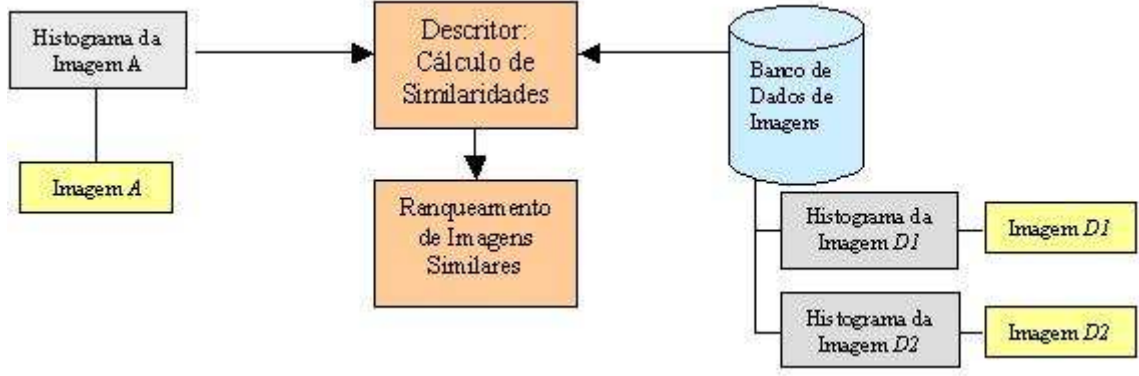


FIG. 2.13: Modelagem conceitual do processo de cálculo de similaridades entre imagens utilizando-se Histogramas de Cor

Cor	j	$H(I_A, j)$	$H(I_{D1}, j)$	$(H(I_A, j) - H(I_{D1}, j))$
Verde	1	4	5	1
Vermelho	2	3	5	1
Branco	3	2	1	1
Distância				1,41

TAB. 2.5: Cálculo da distância L1 entre a imagem exemplo e a imagem D1 do banco de dados de imagens

$$D_H(I_Q, I_D) = \sqrt{\sum_{j=1}^n |(H(I_Q, j) - H(I_D, j))|} \quad (2.6)$$

Para o cálculo das distância entre  $H_A$  e  $H_{D1}$ , temos:

$$D_H(I_A, I_{D1}) = \sqrt{\sum_{j=1}^n |(H(I_A, j) - H(I_{D1}, j))|} \quad (2.7)$$

Para o cálculo das distância entre  $H_A$  e  $H_{D2}$ , temos:

$$D_H(I_A, I_{D2}) = \sqrt{\sum_{j=1}^n |(H(I_A, j) - H(I_{D2}, j))|} \quad (2.8)$$

Os resultados obtidos a partir do cálculo da distância L1 entre as imagens exemplo e as do banco de dados de imagens, demonstrados nas tabelas 2.5 e 2.6, revelam que existe uma similaridade maior entre a imagem A e a imagem D1, uma vez que a distância entre os seus histogramas é inferior à distância obtida para as imagens A e D2.

Cor	j	$H(I_A, j)$	$H(I_{D1}, j)$	$(H(I_A, j) - H(I_{D1}, j))$
Verde	1	4	1	9
Vermelho	2	3	6	9
Branco	3	2	1	1
Distância				4,36

TAB. 2.6: Cálculo da distância L1 entre a imagem exemplo e a imagem D2 do banco de dados de imagens

### 2.5.2.2 DESCRITORES DE TEXTURA

Juntamente com as cores, a textura é uma forte característica de discriminação de imagens, presente em praticamente qualquer coisa na natureza. A pele humana e até mesmo o tronco de uma árvore são exemplos vivos da presença de textura. Segundo (BIMBO, 1999) é praticamente impossível expressar texturas em palavras. Uma representação efetiva pode ser baseada, no entanto, sobre propriedades espectrais, estatísticas e estruturais de padrões de brilho. No contexto de sistemas de recuperação de imagens por conteúdo, estas propriedades têm sido bastante utilizadas especialmente nas etapas de indexação e recuperação.

Não existe um consenso a respeito de uma definição formal de textura. (AKSOY, 1998), por exemplo, conceituaram textura como sendo a distribuição estatística das propriedades da dependência dos níveis de cinza. (ZHONG, 1999) defendem que uma região de uma imagem é texturizada se esta possuir algumas repetições do padrão de intensidade. A textura também pode ser definida como sendo a propriedade presente em imagens, caracterizada pela existência de primitivas básicas, às quais a distribuição espacial cria alguns padrões visuais definidos em termos de granularidade, direcionalidade e repetitividade (TORRES, 2006).

#### **Abordagens para a Descrição de Texturas**

Gonzalez e Woods (2001) citam a existência de três principais abordagens para se descrever texturas: a abordagem espectral, estrutural e estatística. A abordagem espectral diz respeito a propriedades baseadas no espectro de Fourier, onde pode ser detectada a existência de padrões periódicos ou semiperiódicos (KULKARNI, 1994). A abordagem estrutural representa textura como sendo formada pela repetição de padrões que obedeçam alguma regra de posicionamento para a sua geração (BALLARD, 1982)(BALLARD, BROWN, 1982). Já na abordagem estatística, medidas são utilizadas para descrever texturas. Esta abordagem é geralmente recomendada para texturas

que não apresentam boa regularidade (BALLARD, 1982)(BALLARD, BROWN, 1982). O presente projeto concentrará seus esforços em descritores baseados na abordagem estatística.

### **Abordagem Estatística**

A abordagem estatística de descrição de texturas considera o relacionamento, a distribuição dos tons de cinza e o inter-relacionamento entre eles. A seguir serão apresentados descritores de texturas que utilizam esta abordagem, porém, antes será introduzida a matriz de Co-ocorrência, onde são extraídos os descritores de textura utilizado neste trabalho.

**Matriz de Co-Ocorrência** A matriz de co-ocorrência de níveis de cinza (Grey Level Co-Occurrence Matrix - GLCM) leva em consideração a posição de pixels que possuem níveis de cinza. Uma matriz de co-ocorrência é uma tabulação sobre o número de combinações distintas de valores de intensidade que ocorrem em uma imagem (SCHWARTZ, 2003).

A idéia que sustenta uma matriz de co-ocorrência é a realização de uma varredura na imagem, mantendo-se a par do comportamento da periodicidade dos pixels que diferem na intensidade e que são separados por uma distância  $d$  fixa. Normalmente, a direção entre dois pixels também é importante para classificação a partir de textura, e esta é implementada por matrizes múltiplas, sendo uma para cada região de interesse. Em geral, são utilizadas quatro direções: a vertical, horizontal e duas diagonais. Para cada valor de  $d$  temos quatro matrizes, cada uma com uma dimensão  $256 \times 256$  pixels, para uma imagem original com 256 tons de cinza. Ao analisar essas matrizes deve-se calcular alguns valores discretos simples, de forma a encapsular a informação e descrever a textura de uma forma inteligível para o computador. Toda informação pode ser apresentada na forma de matrizes de frequências relativas com a qual dois pixels vizinhos separados pela distância  $d$ , orientação  $q$  ocorrem na imagem, o primeiro pixel tendo seu nível de cinza representado por  $i$  e o seu vizinho por  $j$ . Como resultado teremos matrizes simétricas que podem ser normalizadas através da divisão de cada termo de entrada na matriz pelo número de pixels vizinhos utilizados para computar essa matriz.

As matrizes de co-ocorrência levam em consideração a relação entre dois pixels por vez, sendo o primeiro chamado de pixel de referência e o segundo chamado de pixel vizinho. A quantidade de linhas e colunas nesta matriz é proporcional ao número de tons de cinza presentes na textura. Finalmente tem-se que o elemento  $P(m,n)$  da matriz de co-ocorrência representa o número de transições entre os níveis de cinza  $m$  e  $n$  presentes

	0	1	2	3
0	P(0,0)	P(0,1)	P(0,2)	P(0,3)
1	P(1,0)	P(1,1)	P(1,2)	P(1,3)
2	P(2,0)	P(2,1)	P(2,2)	P(2,3)
3	P(3,0)	P(3,1)	P(3,2)	P(3,3)

TAB. 2.7: Matriz de Co-Ocorrência, cada elemento da matriz diz respeito às transições de níveis de cinza

0	1	2
3	3	0
2	2	1

TAB. 2.8: Imagem 3x3 pixels e com quatro níveis de cinza

na textura (SCHWARTZ, 2005).

Antes da formação da matriz, são determinados os pixels e relacionamentos espaciais a serem considerados. Constrói-se um conjunto S, no qual cada elemento é composto de dois pares ordenados denotando cada pixel envolvido na relação espacial. Em seguida é utilizada a relação para determinar o número de transições que ocorrem entre cada par de tons de cinza contidos na textura, onde  $f(x,y)$  indica o tom de cinza do pixel localizado em  $(x,y)$ .

Onde  $(k,l)$  e  $(i,j)$  são as coordenadas dos respectivos pixels.

Como exemplo, consideremos que o pixel vizinho pode estar à direita de cada pixel referência. Isto pode ser exposto como uma relação  $(1,0)$ : 1 pixel na direção i, 0 pixels na direção j. Cada pixel na imagem se torna um pixel referência, iniciando no canto superior esquerdo indo até o canto inferior direito. Note que os pixels situados na margem direita não possuem vizinhos na sua direita. Portanto estes não serão utilizados para esta contagem.

A tabela 2.7 abaixo exemplifica uma matriz de co-ocorrência para 4 níveis de cinza.

Um elemento  $P(m,n)$  é armazenado em uma linha i e coluna j da matriz, e consiste no número de vezes que um pixel i possui um vizinho de valor j à esquerda ou à direita para situações onde  $\theta=0^\circ$ , em uma distância d estabelecida.

Consideremos uma imagem I 3x3 pixels, com 4 níveis de cinza nas direções 0 e 90 graus e distância  $|d| \leq 1$ , conforme mostra a tabela 2.8.

Sua matriz de co-ocorrência  $M(I,\theta,d)$  para a distância e ângulos especificados, seria dado pelas seguintes matrizes:

$$M(I, 0^\circ, 1) = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \quad M(I, 90^\circ, 1) = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

Segundo (SCHWARTZ, 2005), uma vez que obtivemos as matrizes de co-ocorrência, estas precisam ser normalizadas para que possamos extrair descritores. Esta normalização é realizada a partir da divisão de cada elemento da matriz original pela soma dos seus componentes, como podemos observar na equação abaixo.

$$P(m, n) = \frac{P(m, n)}{\sum_{i=1}^{H_g} \sum_{j=1}^{H_g} P(i, j)} \quad (2.9)$$

Onde  $H_g$  é o maior nível de cinza presente na imagem.

Com o objetivo de melhorar o processo de recuperação por textura, geralmente são criadas diversas matrizes de co-ocorrência para as direções de 0,45,90 e 135 graus, além de se utilizar distâncias d variadas.

Conforme exposto anteriormente uma matriz de co-ocorrência é o ponto de partida para a obtenção de diversos descritores estatísticos. A seguir daremos continuidade ao nosso estudo de descritores baseados nestas matrizes, abordando descritores estatísticos propostos por (HARALICK, 1973).

### **Descritores de Haralick**

(HARALICK, 1973) propuseram quatorze descritores, para representar características presentes em texturas. Estes descritores são obtidos através de cálculos de domínio espacial e a abordagem estatística de segunda ordem, baseada na premissa de que toda a informação de textura de uma imagem está contida total ou parcialmente em relações espaciais entre os seus tons de cinza. São obtidas diversas matrizes de co-ocorrência, a fim de se conhecer a dependência espacial e probabilidade de distribuição, e em seguida são extraídos um ou mais descritores dentre os 14 propostos. (FELIPE, 2002) afirmam que a escolha dos descritores a serem adotados para uma solução de recuperação de imagens por conteúdo se baseia em testes empíricos sobre um domínio específico, levando-se em consideração os que produzirem os melhores resultados.

0	1	2	2
3	3	0	0
2	2	1	1
2	2	1	1

TAB. 2.9: Imagem 4x4 pixels e com quatro níveis de cinza

	0	1	2	3
0	P(0,0)	P(0,1)	P(0,2)	P(0,3)
1	P(1,0)	P(1,1)	P(1,2)	P(1,3)
2	P(2,0)	P(2,1)	P(2,2)	P(2,3)
3	P(3,0)	P(3,1)	P(3,2)	P(3,3)

TAB. 2.10: Relacionamento entre os níveis de cinza da imagem A

A seguir, os descritores propostos por (HARALICK, 1973) serão apresentados em detalhes.

### Segundo Momento Angular e Energia

O segundo momento angular (SMA) denota a homogeneidade da imagem e é caracterizado pela soma dos quadrados dos elementos da matriz de co-ocorrência. Valores elevados de SMA ocorrem quando uma imagem está bastante ordenada (STEIN, 2005).

O descritor de energia pode ser obtido a partir do SMA, para isso, deve-se simplesmente obter a sua raiz quadrada. Como está demonstrada na equação 2.10.

$$ENERGIA = \sqrt{SMA} \quad (2.10)$$

Para ilustrar o estudo dos descritores estatísticos de (HARALICK, 1973), consideremos uma imagem A 4x4 pixels, com quatro níveis de cinza nas direções 0 e 90 graus e distância, conforme ilustra a tabela 2.9.

Inicialmente, obtém-se a matriz de co-ocorrência da imagem A para a distância e ângulo pré-estabelecidos, este processo é ilustrado nas tabelas 2.10 e 2.11. Observe que a matriz é quadrada e o seu número de linhas e colunas é dado pelo número de tons de cinza presentes na imagem.

Portanto para  $|d| \leq 1$  e  $\theta=0^\circ$ , tem-se a tabela 2.11

Finalmente tem a  $M(A,0^\circ,1)$ :



	0	1	2	3
0	2	1	0	1
1	1	4	3	0
2	0	3	6	0
3	1	0	0	2

TAB. 2.11: Relacionamento de vizinhança entre os tons de cinza da imagem A com  $d \leq 1$  e  $\theta=0^\circ$ .

$$M(A, 0^\circ, 1) = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 4 & 3 & 0 \\ 0 & 3 & 6 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$$

Agora se deve obter uma nova matriz, desta vez com a direção  $90^\circ$ . Para isso basicamente, repete-se o processo utilizado para a matriz anterior, atentando para a nova direção a ser avaliada, neste caso  $90^\circ$ , como é mostrado na tabela 2.13. Desta forma extrai-se a seguinte matriz resultante:

	0	1	2	3
0	0	2	2	1
1	2	4	0	1
2	2	0	4	2
3	1	1	2	0

TAB. 2.12: Relacionamento de vizinhança entre os tons de cinza da imagem A com  $d \leq 1$  e  $\theta=90^\circ$ .

Logo:

$$M(A, 90^\circ, 1) = \begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 4 & 0 & 1 \\ 2 & 0 & 4 & 2 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

Uma vez computadas as matrizes de co-ocorrência, deve-se agora partir para a normalização das matrizes. O passo seguinte é o cálculo dos descritores estatísticos de textura em questão. Portanto, considerando as matrizes obtidas, calcula-se o segundo momento angular das matrizes de co-ocorrência da imagem A da seguinte maneira:

$$SMA = \sum_i^{H_g} \sum_j^{H_g} (p(i, j, d, \theta))^2 \quad (2.11)$$

$$M(A, 0^\circ, 1) = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 4 & 3 & 0 \\ 0 & 3 & 6 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$$

Para  $\theta = 0^\circ$  e  $d \leq 1$ :

i	j	$p(i, j, 1, 0^\circ)$	$p(i, j, 1, 0^\circ)^2$	$\sum_i \sum_j (p(i, j, d, 0^\circ))^2$
0	0	2	4	82
0	1	1	1	
0	2	0	0	
0	3	1	1	
1	0	1	1	
1	1	4	16	
1	2	3	9	
1	3	0	0	
2	0	0	0	
2	1	3	9	
2	2	6	36	
2	3	0	0	
3	0	1	1	
3	1	0	0	
3	2	0	0	
3	3	2	4	

TAB. 2.13: Cálculo do SMA para  $\theta = 0$  e  $d \leq 1$ .

Finalmente para  $\theta = 90^\circ$  e  $d \leq 1$ , teremos:

$$M(A, 90^\circ, 1) = \begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 4 & 0 & 1 \\ 2 & 0 & 4 & 2 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

Para se obter o descritor de energia deve-se basicamente calcular a raiz quadrada dos valores computados para o Segundo Momento Angular. Portanto:

i	j	$p(i,j,1,90^\circ)$	$p(i,j,1,90^\circ)^2$	$\sum_i \sum_j (p(i,j,d,90^\circ))^2$
0	0	0	0	58
0	1	2	4	
0	2	2	4	
0	3	1	1	
1	0	2	2	
1	1	4	16	
1	2	0	0	
1	3	1	1	
2	0	2	4	
2	1	0	0	
2	2	4	16	
2	3	2	4	
3	0	1	1	
3	1	1	1	
3	2	2	4	
3	3	0	0	

TAB. 2.14: Cálculo do SMA para  $\theta = 90^\circ$  e  $d \leq 1$ .

$$ENERGIA(A, 0^\circ, 1) = \sqrt{82} = 9.05$$

$$ENERGIA(A, 90^\circ, 1) = \sqrt{58} = 7.61$$

A fim de se apresentar um modelo invariante quanto à rotação de imagens, (HARALICK, 1973), propôs o uso de uma média de cada descritor obtido das quatro direções com uma distância escolhida, geralmente 1 (GOLDSCHMIDT, 2001). Portanto para o cálculo de similaridades entre imagens exemplo, deveremos encontrar a média entre os valores dos descritores, descobertos a partir das suas matrizes de co-ocorrência. Uma vez obtidos os descritores, parte-se para a utilização de um conceito de distância como, por exemplo, a Euclidiana, para se obter o grau de similaridade entre duas imagens.

### Entropia

A entropia representa o grau de dispersão dos níveis de cinza de uma imagem. Este descritor, juntamente com o SMA, pode denotar homogeneidade. A entropia fornece o tamanho e a probabilidade de um sistema ser fechado, no sentido de que à maior entropia corresponde a menor ordem de organização das estruturas, e também a maior probabilidade de ocorrência de um determinado nível de cinza (GUALBERTO, 2007).

$$ENTROPIA = - \sum_i^{H_g} \sum_j^{H_g} p(i, j, d, \theta) \log p(i, j, d, \theta) \quad (2.12)$$

Para exemplificar o cálculo do descritor de entropia, consideremos ainda, o exemplo da imagem A e suas respectivas matrizes de co-ocorrências  $M(A, 0^\circ, 1)$  e  $M(A, 90^\circ, 1)$ .

Para  $\theta = 0^\circ$  e  $d \leq 1$ :

i	j	$p(i,j,1,0^\circ)$	$\log p(i,j,1,0^\circ)$	$p(i,j,d,0^\circ)\log p(i,j,d,0^\circ)$
0	0	2	0,301	0,602
0	1	1	0	0
0	2	0	-	-
0	3	1	0	0
1	0	1	0	0
1	1	4	0,602	2,408
1	2	3	0,477	1,431
1	3	0	-	0
2	0	0	-	0
2	1	3	0,477	1,431
2	2	6	0,778	4,668
2	3	0	-	0
3	0	1	0	0
3	1	0	-	0
3	2	0	-	0
3	3	2	0,301	0,602

TAB. 2.15: Cálculo do ENTROPIA para  $\theta = 0^\circ$  e  $d \leq 1$ .

Logo:

$$ENTROPIA(A, 0^\circ, 1) = - \sum_i^3 \sum_j^3 p(i, j, 1, 0^\circ) \log p(i, j, 1, 0^\circ) = -11,142$$

Para  $\theta = 90^\circ$  e  $d \leq 1$ :

Logo:

$$ENTROPIA(A, 90^\circ, 1) = - \sum_i^3 \sum_j^3 p(i, j, 1, 90^\circ) \log p(i, j, 1, 90^\circ) = -8,428$$

### Contraste

O descritor de contraste representa a diferença de níveis de cinza (diferença de momento) dos pixels de uma imagem. Em termos de matriz de co-ocorrência, valores baixos para este descritor ocorrem quando há uma grande concentração de elementos significativos na diagonal principal (STEIN, 2005).

$$CONTRASTE = \sum_{i=0}^{H_g-1} \sum_j^{H_g-1} |i - j|^2 p(i, j, d, \theta) \quad (2.13)$$

i	j	p(i,j,1,90°)	logp(i,j,1,90°)	p(i,j,d,90°)logp(i,j,d,90°)
0	0	0	-	0
0	1	2	0,301	0,602
0	2	2	0,301	0,602
0	3	1	0	0
1	0	2	0,301	0,602
0	3	1	0	0
1	0	2	0,301	0,602
1	1	4	0,602	2,408
1	2	0	-	0
1	3	1	0	0
2	0	2	0,301	0,602
2	1	0	-	0
2	2	4	0,602	2,408
2	3	2	0,301	0,602
3	0	1	0	0
3	2	2	0,301	0,602
3	3	0	-	0

TAB. 2.16: Cálculo do ENTROPIA para  $\theta = 90^\circ$  e  $d \leq 1$ .

Para ilustrar o cálculo deste descritor, mais uma vez será utilizado o exemplo proposto onde tem-se uma imagem A e suas respectivas matrizes de co-ocorrência e :

$$M(A, 0^\circ, 1) = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 1 & 4 & 3 & 0 \\ 0 & 3 & 6 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \quad M(A, 90^\circ, 1) = \begin{bmatrix} 0 & 2 & 2 & 1 \\ 2 & 4 & 0 & 1 \\ 2 & 0 & 4 & 2 \\ 1 & 1 & 2 & 0 \end{bmatrix}$$

Para  $\theta = 0^\circ$  e  $d \leq 1$  tem a tabela 2.17.

Logo:

$$CONTRASTE(A, 1, 0^\circ) = \sum_{i=0}^3 \sum_j^3 |i - j|^2 p(i, j, 1, 0^\circ) = 40 \times 24 = 960$$

Para  $\theta = 90^\circ$  e  $d \leq 1$  tem a tabela 2.18

Finalmente:

$$CONTRASTE(A, 1, 90^\circ) = \sum_{i=0}^3 \sum_j^3 |i - j|^2 p(i, j, 1, 90^\circ) = 40 \times 22 = 880$$

### Homogeneidade

A representação de homogeneidade assume valores altos quando uma dada textura apresenta uma baixa variação entre os tons de cinza da imagem.

i	j	$p(i,j,1,0^\circ)$	$ i-j $	$ i-j ^2$	$\sum_i^3 \sum_j^3  i-j ^2$	$\sum_i^3 \sum_j^3 p(i,j,1,0^\circ)$
0	0	2	0	0	40	24
0	1	1	1	1		
0	2	0	2	4		
0	3	1	3	9		
1	0	1	1	1		
1	1	4	0	0		
1	2	3	1	1		
1	3	0	2	4		
2	0	0	2	4		
2	1	3	1	1		
2	2	6	0	0		
2	3	0	1	1		
3	0	1	3	9		
3	1	0	2	4		
3	2	0	1	1		
3	3	2	0	0		

TAB. 2.17: Cálculo da CONTRASTE para  $\theta = 0^\circ$  e  $d \leq 1$ .

$$HOMOGENEIDADE = \sum_{i=0}^{H_g-1} \frac{1}{1 + (i-j)^2} p(i, j, d, \theta) \quad (2.14)$$

### Cálculo de Similaridade utilizando Descritores de Haralick

Uma vez que foram calculados os descritores de (HARALICK, 1973) para duas imagens  $I_A$  e  $I_B$ , pode-se aplicar uma medida de distância, como a euclidiana, para mensurar a similaridade entre duas imagens com base nestas características. Para isso calcula-se:

$$d = \sqrt{(D(I_B) - D(I_A))^2} \quad (2.15)$$

i	j	$p(i,j,1,90^\circ)$	$ i - j $	$ i - j ^2$	$\sum_i^3 \sum_j^3  i - j ^2$	$\sum_i^3 \sum_j^3 p(i, j, 1, 90^\circ)$
0	0	0	0	0	40	24
0	1	2	1	1		
0	2	2	2	4		
0	3	1	3	9		
1	0	2	1	1		
1	1	4	0	0		
1	2	0	1	1		
1	3	1	2	4		
2	0	2	2	4		
2	1	0	1	1		
2	2	4	0	0		
2	3	2	1	1		
3	0	1	3	9		
3	1	1	2	4		
3	2	2	1	1		
3	3	0	0	0		

TAB. 2.18: Cálculo da CONTRASTE para  $\theta = 90^\circ$  e  $d \leq 1$ .

### 3 TRABALHOS RELACIONADOS

Este capítulo tem por finalidade apresentar trabalhos relacionados à presente dissertação. Alguns destes trabalhos nortearam a elaboração da abordagem proposta. Em função da estrutura da solução descrita neste trabalho, este capítulo foi dividido em duas partes conforme as abordagens de recuperação de imagens: Recuperação baseada em Sentenças e baseada em Conteúdo. A primeira compreende trabalhos que envolvem a recuperação de imagens a partir de sentenças formuladas em linguagem natural. A segunda utiliza imagens exemplo como referência na busca por imagens que sejam semelhantes às apresentadas.

Ao final deste capítulo é apresentada uma análise comparativa entre os trabalhos pesquisados e a abordagem proposta nesta dissertação.

#### 3.1 RECUPERAÇÃO DE IMAGENS BASEADA EM SENTENÇAS

Nesta seção são mostrados sistemas capazes de recuperar imagens através de consultas em sentença de linguagem natural. Os trabalhos a seguir abordam como a utilização de informações descritivas sobre as imagens, como fonte de comparação no processo de recuperação dos dados.

##### 3.1.1 ONAIR

Em (PAZ-TRILLO, 2001) é apresentado um sistema para recuperação de vídeos através de sentenças escritas em linguagem natural. A coleção de vídeos utilizada foi compilada de um conjunto de entrevistas com a artista brasileira Ana Teixeira. Essas entrevistas foram transcritas em textos, para que assim pudessem ser tratadas como informações das imagens e, posteriormente serem recuperadas através destas. O domínio da entrevista desenvolvida pelo sistema é o da arte contemporânea. Esse domínio foi utilizado para desenvolver uma ontologia, que serviu de suporte para a recuperação das imagens em vídeo.

Nesse trabalho, o usuário primeiramente entra com uma consulta em linguagem natural e o sistema retorna trechos de vídeos, os quais são pertinentes à consulta. O problema



da sinonímia é tratado por uma representação de conhecimento, neste caso, a ontologia, a qual trata de um domínio específico.

Basicamente, este sistema consiste de dois processos: indexação e recuperação. A primeira parte trata de preparar os vídeos e as informações para uma eventual recuperação, são utilizados procedimentos como remoção de afixos e palavras de parada(stopwords), e geração de índices invertidos. Estes são gerados através dos textos transcritos e palavras-chave associadas às imagens, com o auxílio da ontologia para tratar de sinônimos e peso para os índices. E a última parte, a recuperação propriamente dita, usa algumas das operações da indexação, como remoção de afixos e de palavras de parada, para tratar da sentença de consulta, juntamente com um adicionamento da expansão de consulta e a recuperação dos vídeos. Os trechos de vídeos são mostrados através de um player que permite visualizá-los individualmente ou todos em uma sequência.

Esse trabalho é muito parecido com o nosso trabalho proposto por esta dissertação, na parte de recuperação por sentenças, tanto que muitas idéias deste foram agregadas ao nosso projeto. Porém, há uma diferença significativa para o nosso sistema, possui claramente uma recuperação para vídeos, imagens dinâmicas, enquanto o nosso manipula fotos, imagens estáticas.

### 3.1.2 PHOTO-RDF

O trabalho apresentado em (PHOTO-RDF, 2002) é um projeto de descrever e recuperar fotos digitalizadas com metadados escritos em RDF(seção 2.3.1). É usada esta linguagem para caracterizar uma imagem, através de um programa de entrada que cataloga as fotos com seus metadados, sendo implementado em Java. Um servidor específico chamado JigSaw foi desenvolvido para recuperar a foto ou as informações RDF, dependendo da requisição *Web*. O funcionamento geral desse sistema é ilustrado na figura a seguir, na figura 3.1:

O programa de entrada possui três tipos esquemas de metadados: Dublin Core(SOUZA, 2000), Técnico(informações à respeito da câmera fotográfica ou digital) e de Contexto(palavras-chave a respeito da foto). Os três têm os seus requisitos bem definidos e restando ao usuário apenas inserir os dados que julgar pertinente. Esses metadados estão destacados em forma de abas, DC(Dublin Core) e Technical(Técnico), na figura 3.2, logo a seguir, mostrando como é a entrada desse sistema:

Foi averiguada que esse sistema contém muitas limitações. Como foi citado

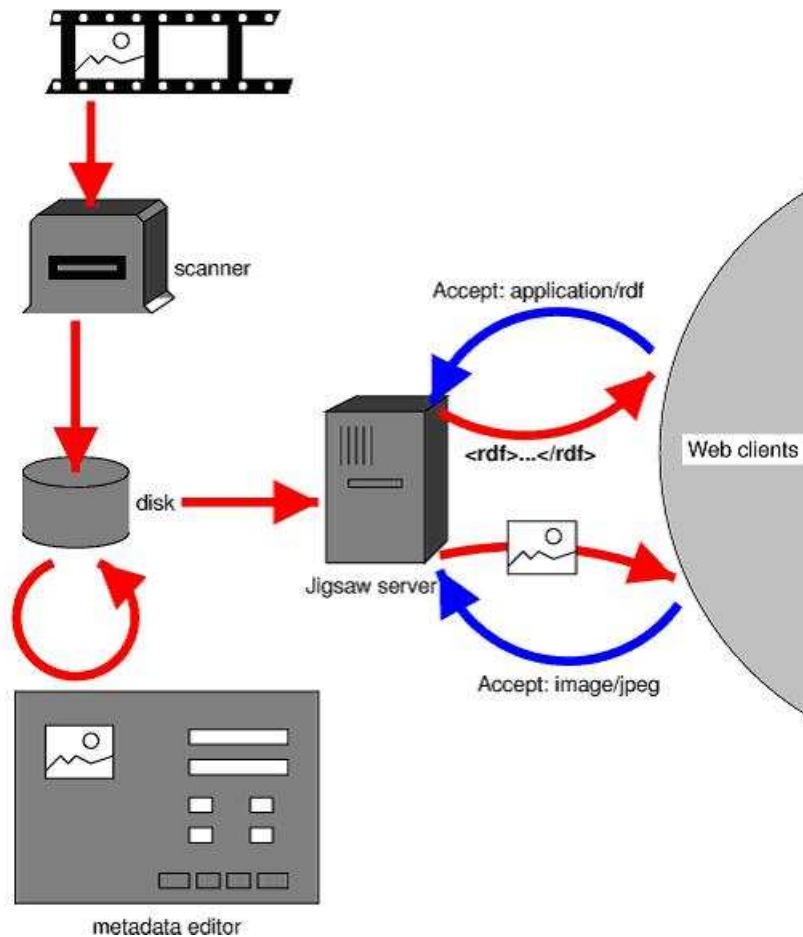


FIG. 3.1: Diagrama do Sistema Photo-RDF

anteriormente, o fato dos requisitos de entrada serem pré-estabelecidos, não dando a oportunidade de inserir mais campos ou alterar alguns deles. Outra observação é que por se tratar de metadados, representado por Rdf, a expressividade fica restrita a esta linguagem. O que se verificou ser limitante ao presente trabalho.

### 3.1.3 MEDUSA

Medusa(BROWN, 1995) é um sistema de recuperação de imagens dinâmicas, mais precisamente vídeos de noticiário. Este sistema implementa métodos estatísticos para recuperação textual de documentos multimídia. A recuperação é feita por sobre textos elaborados pela produção de um jornal, que servem de pontos eletrônicos para os apresentadores. Esses textos passam por um tratamento, como por exemplo, a retirada de palavras de parada(stopwords) e depois são armazenados. O sistema disponibiliza uma consulta que o usuário requisita, as notícias que possuem similaridades serão relacionadas

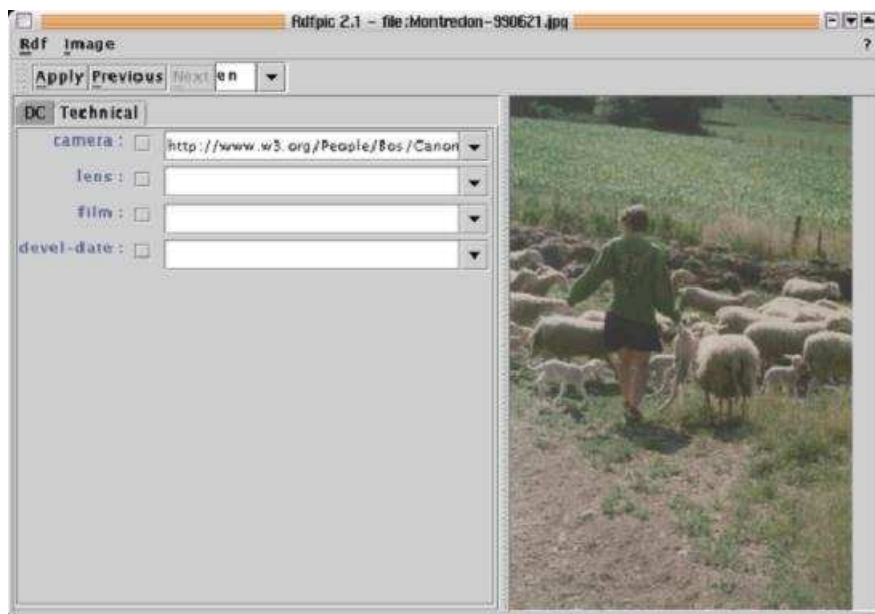


FIG. 3.2: Entrada de dados do sistema photo-RDF

com o seu tempo de duração.

As limitações desse sistema são de tratar apenas de vídeos e do texto do ponto eletrônico. A primeira é a mesma limitação já citada anteriormente na seção 3.1.1, não atende bem ao presente trabalho, porém algumas idéias como a remoção de palavras de parada e afixos podem ser aproveitadas. A segunda é que não há nenhum tratamento para palavras sinônimas e polissêmicas, problemas que já foram apontadas anteriormente na seção 2.4.1.

### 3.2 RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO

O interesse por sistemas CBIR(Content-Based Image Retrieval) surgiu do grande crescimento na produção de imagens digitais e da necessidade de recuperá-las em grandes bases de dados com maior facilidade. Pesquisas sobre novas tecnologias nesta área têm sido feitas com grande intensidade e cada vez mais aplicações têm deixado o meio acadêmico em forma de produtos comerciais, como é o caso de aplicações como o Query by Image Content - QBIC da IBM e Virage(BIMBO, 1999). Nesta seção apresentaremos uma visão geral de alguns dos principais sistemas de CBIR existentes na atualidade.

### 3.2.1 BIRAM

O sistema BIRAM (Base de Imagens Relacional de Algoritmos e Métricas) (MORENO, 2006) é um protótipo que foi desenvolvido pela Universidade de São Paulo - USP. O objetivo principal dos pesquisadores foi desenvolver uma ferramenta flexível capaz de resolver os problemas de recuperação de informações na crescente base de dados do INCOR (Instituto do Coração).

Este sistema foi originalmente desenvolvido em Java utilizando uma base de dados Microsoft Access cuja comunicação é feita via interface Java Database Connectivity - JDBC. Esta ferramenta tem como característica a possibilidade de adicionar novos algoritmos de cálculo de descritores. Este procedimento é realizado através do envio de um arquivo jar (*Java<sup>TM</sup> Archive*) para o servidor, onde as meta-informações do algoritmo são armazenadas no banco de dados, estas estão ilustradas na figura 3.3. Atualmente a ferramenta somente possui algoritmos como cálculo de momento, histograma, histograma métrico e centro de massa(seção 2.5.2).



ALGORITHM
ALG_NAME
ALG_DESCRIPTION
ALG_AUTHOR
ALG_CLASS_JAVA
ALG_BODY_PART
<b>ALG_ALGORITHM_ID</b>
ALG_MOD
ALG_BITS
ALG_VERSION
ALG_ROOT_PATH

FIG. 3.3: - Meta-informações do algoritmo((MORENO, 2006))

O processo de busca de imagens semelhantes pode ser feito tanto de forma de sentenças como através de conteúdo. Este último é feito através do fornecimento de uma imagem de entrada que é processada por um dos algoritmos armazenados que produzem como resultado um descritor. Em seguida este descritor é comparado com todos os já cadastrados no banco de dados através de um método de cálculo de distância presente no algoritmo utilizado. Com o resultado das distâncias calculadas é possível ordenar o resultado da pesquisa de forma que os resultados mais similares apareçam primeiro. A interface principal do sistema é mostrada na figura 3.4 abaixo:

Esta ferramenta possui um futuro promissor, entretanto possui alguns pontos negativos

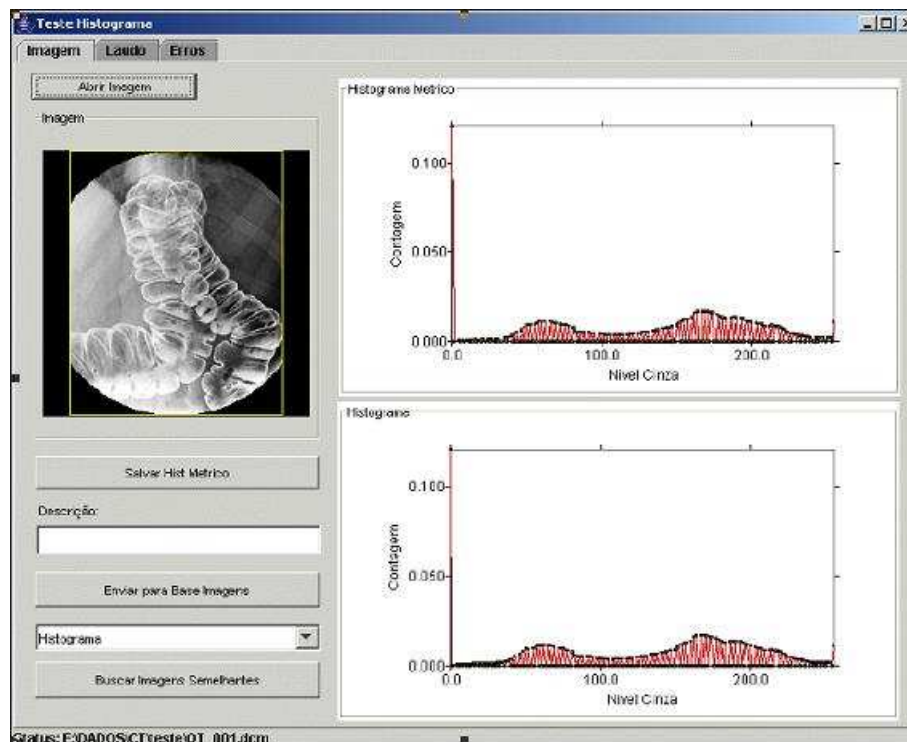


FIG. 3.4: Interface de Especificação de Busca de uma Determinada Imagem((MORENO, 2006))

como o fato de ser inteiramente voltado para a área médica e não possuir nenhum método de indexação das imagens armazenadas, fazendo com que a cada pesquisa o sistema tenha que acessar todas as imagens armazenadas no banco para encontrar as mais similares, o que se torna inviável em grandes bases de dados. Ao contrário da nossa abordagem que visa fazer uma indexação de todas as imagens. Evitando assim o acesso a todas as imagens.

### 3.2.2 QBIC

A ferramenta Query by Image Content, popularmente conhecida como QBIC, foi um dos primeiros sistemas comerciais de recuperação de imagens por conteúdo. É, também, um dos mais famosos. Desenvolvido pela International Business Machines - IBM, este software é fornecido como uma aplicação independente ou como parte integrante de outros produtos desta empresa, como por exemplo, o banco de dados DB2(EAKINS, 1999).

Devido a muitos anos de pesquisa, esta ferramenta hoje se apresenta como um dos principais sistemas de recuperação de imagens e também como sendo um dos mais completos, por possuir uma vasta lista de técnicas para recuperação de imagens e

vídeos(EAKINS, 1999). Dentre suas principais características, podemos citar suporte à recuperação por combinações de descritores de cor, textura, forma, assim como descrição textual. A interface de recuperação do sistema é mostrada na figura 3.5 abaixo:



FIG. 3.5: Interface de apresentação de resultados do sistema Query By Image Content - QBIC

Para recuperação por cor as características calculadas nesta ferramenta são a média dos sistemas de cor (R, G, B), (Y, i, q), (L, a, b) e MTM (transformações matemáticas de Munsell) e histogramas com base no algoritmo de k-element. As características relacionadas à textura utilizadas são aspereza, contraste e direcionalidade. Já para forma são usados cálculos de áreas de forma, circularidade, excentricidade, maior eixo de orientação e um conjunto de momentos de invariância algébricos.

Além de permitir a possibilidade de fornecer recuperação através de imagem exemplo, o QBIC também disponibiliza a pesquisa por determinados tipos de padrões de textura e cor, podendo inclusive determinar a quantidade das cores desejadas nas imagens a serem recuperadas. O usuário também pode efetuar buscas de determinadas imagens que obedeçam as formas definidas em um esboço desenhado por ele.

Esta ferramenta mostra-se madura, com muita flexibilidade no tocante à realização de consultas, entretanto, por ser uma ferramenta proprietária não fornece a possibilidade de incorporar novos algoritmos de cálculos de descritores. A parte que trata da recuperação textual se restringe a palavras-chave, que representam a um campo de interesse, ou seja,

pode ser considerada bem limitada.

### 3.2.3 VISUALSEEK

Desenvolvido pela Universidade de Columbia(Estados Unidos), esta aplicação surgiu com a finalidade de facilitar a busca de imagens via WEB. Este projeto deu grande ênfase a determinadas funções como a automação de extração de características locais de regiões de uma imagem; possibilidade de buscas tanto por características das imagens quanto por informações referentes à disposição espacial de objetos; desenvolvimento de técnicas de indexação mais eficientes e de uma interface que possibilitasse variadas opções de busca ao usuário (SMITH, 1996).

Sua arquitetura consiste de uma interface construída na linguagem Java que coleta a consulta feita pelo usuário e envia para servidor via Common Gateway Interface (CGI). Este servidor responde a consulta acessando os metadados, que descrevem as imagens, armazenados estruturalmente em arquivos de indexação (SMITH, 1996).

Esta aplicação tem suporte à extração de características de cor, textura, assim como a disposição espacial destas. Além disso, o usuário conta com funcionalidades como especificar sua consulta apontando as características espaciais da imagem ao qual deseja, apontando em quais regiões da imagem deve ser pesquisada uma determinada cor ou textura e de efetuar uma nova busca utilizando uma imagem retornada de uma consulta feita previamente como base. Quanto aos resultados das pesquisas, estes são listados em forma de *thumbnails* ordenados em ordem decrescente de similaridade. Na figura 3.6 é mostrada a maneira como a consulta é atribuída:

Esta ferramenta ainda se encontra restrita ao meio acadêmico, porém apresenta interessantes opções de recuperação. Segundo (SMITH, 1996) a aplicação ainda sofrerá algumas melhorias como a inclusão da possibilidade de se pesquisar utilizando descritores de forma e efetuar pesquisas por descrição textual, entretanto não consta nesta lista de melhorias a possibilidade de incorporação dinâmica de novos cálculos de descritores.

### 3.2.4 PHOTOBOOK

Este sistema foi desenvolvido pelo laboratório de mídia do MIT (Massachusetts Institute of Technology) e consiste em um conjunto de ferramentas interativas para navegar e recuperar imagens ou seqüências de imagens em uma determinada base de imagens(ZHANG, 2002). A idéia básica desta ferramenta é uma estrutura de compressão



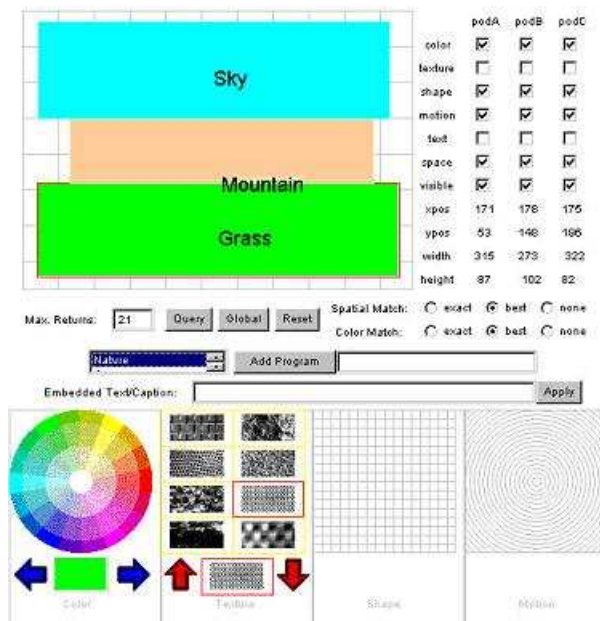


FIG. 3.6: Tela de especificação de consulta por disposição espacial do Sistema VisualSEEK

que permite armazenar um conjunto de coeficientes de significados perceptuais da imagem sem que haja perda de semântica (PENTLAND, 1995).

O Photobook permite a recuperação através de descrição textual e da comparação de características de cor, textura e forma associadas à imagem fornecida pelo usuário com as imagens armazenadas na base de imagens. Ele também provê vários algoritmos de comparação de similaridade, dentre eles cálculos de distância Euclidiana, Mahalanobis, histograma, divergência e wavelet tree. Além disso, este sistema ainda fornece a possibilidade de recuperação por faces que inclusive é utilizada por alguns departamentos de polícia nos Estados Unidos.

A interface deste sistema ainda permite que o usuário possa delimitar sua pesquisa a uma determinada categoria e que também possa buscar uma anotação textual presente na imagem desejada. Isto tudo sendo apoiado por um banco de dados embarcado e orientado a objetos chamado Framer(HAASE, 1993). O resultado das consultas é exibido no formato de thumbnails ordenados por ordem decrescente de similaridade sendo passível de navegação por entre as páginas do resultado. A figura 3.7 abaixo mostra a recuperação do sistema:

Este sistema se mostra bem completo, funciona exclusivamente sobre a plataforma UNIX/Linux e inclui a possibilidade de se adicionar novos algoritmos de busca a partir de sua versão 6. Porém, esse sistema faz a consulta direcionada para determinados banco



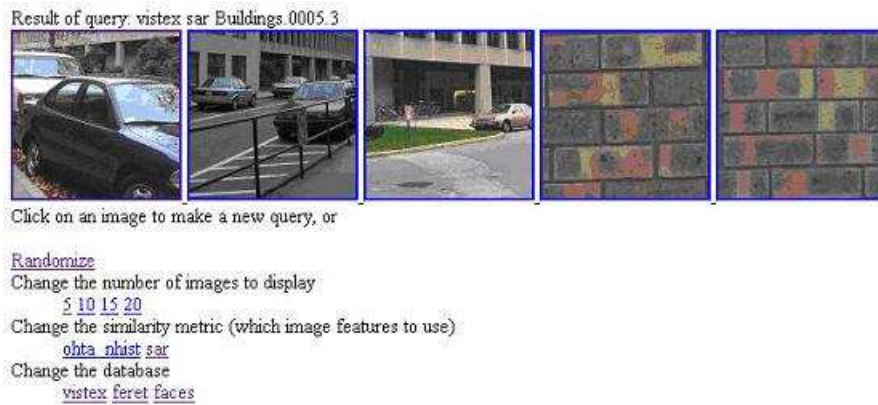


FIG. 3.7: Apresentação de Resultados de Recuperação de Imagens do Sistema Photobook de imagens, por exemplo: "faces de pessoas", "ferramentas", "peixes", "ventrículos", etc. O objetivo do nosso sistema está em não ser tão restrito assim.

### 3.3 ANÁLISE COMPARATIVA

A tabela 3.1 apresenta um quadro em que os trabalhos pesquisados são enquadrados segundo características e funcionalidades identificadas e consideradas relevantes para fins comparativos.

	IE	ID	MD	Ont	Ind	PC	DT	CI	BD	EX	OS
OnAIR		X	X	X	X	X	X			X	
Photo-RDF	X		X			X	X				
Medusa		X	X				X		X		
BIRAM	X		X				X	X	X		X
QBIC	X	X				X		X	X		
VisualSeek	X		X		X			X		X	
Photobook	X	X					X	X	X		
Abordagem Proposta	X		X	X	X	X	X	X		X	X

TAB. 3.1: Tabela Comparativa dos Trabalhos Relacionados

A tabela 3.2 contém uma legenda explicativa para os mnemônicos utilizados na tabela 3.1:

Convém ressaltar que, embora com caráter acadêmico, o ambiente proposto nesta dissertação congrega a maioria das características e funcionalidades relevantes em sistemas de recuperação de imagens.

Cabe ainda mencionar que dos trabalhos estudados, nenhum deles apresenta formas de combinação entre as abordagens de recuperação de imagens baseada em sentenças e

Sigla	Características/Funcionalidade
IE	Imagens Estatísticas, ou seja, que trabalham com fotos
ID	Imagens Dinâmicas, ou seja, sistemas que trabalham com vídeos ou sequências de imagens
MD	Meta-Dados, sistemas que se utilizam de meta-dados
Ont	Ontologia, sistemas que se utilizam de ontologia
Ind	Indexação, sistemas que fazem indexação das imagens
PC	Palavras-Chave, sistemas que se utilizam de palavras-chave
DT	Descrição Textual, sistemas que se utilizam de descrição textual
CI	Características de Imagens, sistemas que trabalham com descritores de imagens
BD	Banco de Dados, sistemas que se utilizam de banco de dados
EX	Extensível, sistemas que podem extensíveis a outras informações
OS	OpenSource, sistemas que possuem código aberto

TAB. 3.2: Mnemônicos da Comparação dos Trabalhos

em conteúdo, principal característica inovadora da proposta da presente dissertação. E só dois trabalhos se utilizam de ontologia para realizar a recuperação.

## 4 ABORDAGEM PROPOSTA

### 4.1 CONSIDERAÇÕES INICIAIS

Este capítulo tem por finalidade fornecer uma descrição conceitual detalhada do ambiente inteligente para recuperação de imagens proposto pela presente dissertação. Este ambiente foi concebido com base nas seguintes abordagens de recuperação de imagens existentes na literatura específica:

- (a) Recuperação de Imagens Baseada em Sentenças;
- (b) Recuperação de Imagens Baseada em Conteúdo.

O ambiente a ser descrito incorpora de forma isolada as abordagens mencionadas, além de permitir uma combinação entre elas. Esta combinação caracteriza-se pela interseção entre os conjuntos de imagens produzidas pelas abordagens (a) e (b).

Em quaisquer das abordagens, é necessário um pré-processamento prévio das imagens a fim de extrair metadados com informações relevantes para o processo de recuperação. A abordagem (a) pressupõe a existência de textos descritivos e palavras-chave associadas às imagens armazenadas. A abordagem (b) envolve apenas o conteúdo gráfico inerente às próprias imagens.

Na figura 4.1 logo abaixo, tem-se a idéia conceitual da recuperação, para todas as abordagens. As setas contínuas indicam o fluxo de indexação e armazenamento das imagens. As setas tracejadas mostram o fluxo de recuperação das imagens a partir de consultas formuladas.

### 4.2 RECUPERAÇÃO DE IMAGENS BASEADA EM SENTENÇAS

Nesta seção é descrita a abordagem de Recuperação de Imagens Baseada em Sentenças. Essa abordagem é composta basicamente de dois processos: indexação e recuperação. Na indexação, o sistema gera arquivos de índices invertidos que serão utilizados para a recuperação das imagens. Na recuperação, é feito um tratamento na consulta para que definitivamente possa retornar as imagens relevantes a esta. Para uma melhor

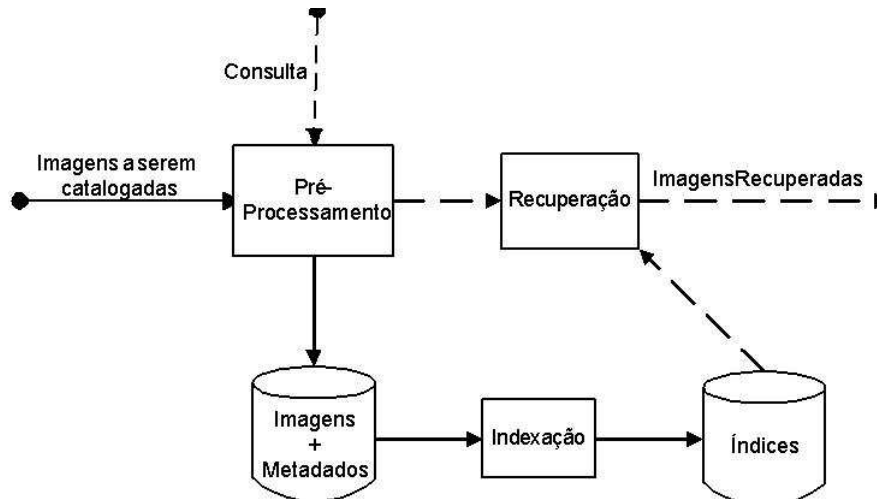


FIG. 4.1: Visão Geral de um Sistema de Recuperação de Informação

compreensão da abordagem, as etapas de pré-processamento, indexação e recuperação serão detalhadas nas seções logo a seguir.

Convém destacar que a abordagem de recuperação de imagens baseada em sentenças utiliza recursos de ontologia com o objetivo de identificar sinônimos e palavras relacionadas aos termos utilizados tanto na descrição das imagens quanto na formulação das consultas para recuperação das imagens. Assim sendo, esta abordagem tem como premissa a existência de uma ontologia de domínio no contexto da aplicação.

Cabe ainda enfatizar que a recuperação de imagens baseada em sentenças está estruturada de forma a permitir a busca de imagens tanto a partir de consultas formuladas em linguagem natural quanto a partir de palavras-chave previamente associada às imagens.

Conforme mostrado no diagrama conceitual da figura 4.2, a abordagem da recuperação baseada em sentenças possui essencialmente duas etapas principais: uma que envolve a indexação das figuras e a outra que trata da recuperação destas. Os itens de 1 a 4, mostrados abaixo, descrevem a etapa de indexação (fluxo de informação em linhas contínuas) e os demais, a etapa de recuperação (fluxo de informação em linhas tracejadas):

- (1) Relaciona todas as figuras envolvidas no contexto, que posteriormente serão recuperadas pelo sistema;
- (2) Com as figuras devidamente relacionadas, é feita uma associação destas com os textos descritivos e as palavras-chave, resultando em um arquivo de metadados (XML);

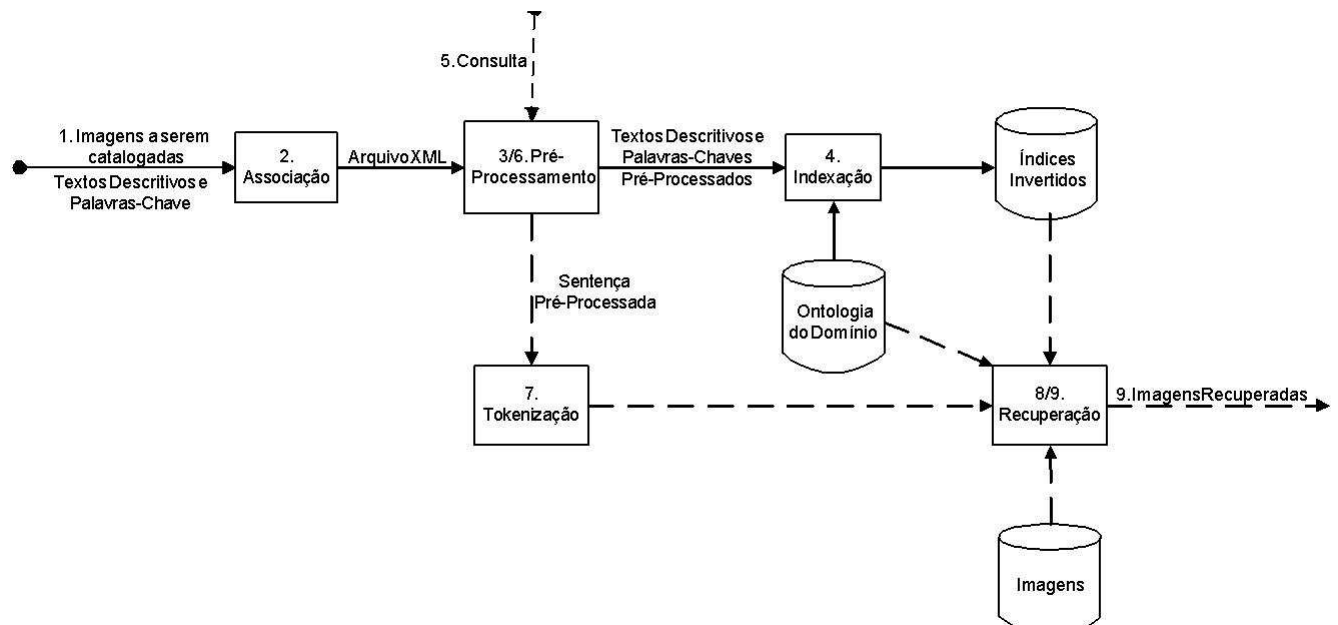


FIG. 4.2: Visão Geral da Recuperação baseada por Sentenças

- (3) Com o arquivo de metadados é feito o pré-processamento dos textos descritivos e das palavras-chave. São realizadas operações como remoção de palavras de paradas (stopwords) e de remoção de afixos (stemming) explicados em detalhe no capítulo 2;
- (4) Com os textos descritivos e as palavras-chave pré-processados, e ainda mais a ontologia de domínio (capítulo 2) são criados os arquivos de índices invertidos, que são formados pelos termos componentes dos textos descritivos e palavras-chave pré-processados. A cada termo, há um peso correspondente à sua relevância no domínio;
- (5) Para iniciar uma consulta, é elaborada uma sentença, que pode ser uma frase ou palavra-chave, como foi citado anteriormente;
- (6) A sentença escolhida é passada pelo pré-processamento, resultando uma sentença pré-processada;
- (7) A sentença pré-processada é passada por uma tokenização dos seus termos;
- (8) Cada termo na sentença é procurado no(s) índice(s) invertido(s) na busca de um item correspondente. Todos os itens correspondentes são catalogados em uma lista;

- (9) Os itens correspondentes são ordenados e depois são mostradas as figuras correspondentes nessa mesma ordem.

#### 4.2.1 PRÉ-PROCESSAMENTO

Esta seção apresenta o tratamento dos textos descritivos e das palavras-chave, onde estes são passados por remoção de palavras de parada(*stopwords*), tratamento de sinônimos e remoção de afixos(*stemming*)(seção 2.4.2.1). Na remoção de palavras de parada, é feita a retirada de todas as palavras que sejam irrelevantes para serem usadas na recuperação, tais como por exemplo, preposições, interjeições, pronomes, etc. No tratamento de sinônimos é verificado se existe um sinônimo definido para um determinado termo. A remoção de afixos consiste em tirar sufixos que compõem os termos, ou seja, reduzir cada termo somente ao seu radical para que possa assim armazenar menos informações na indexação.

##### Remoção de Palavras de Parada

O sistema analisa todos os textos descritivos, palavras-chave e a consulta a fim de remover as palavras de parada(*stopwords*). As palavras de parada são irrelevantes para uma futura consulta, como nos casos de preposições, interjeições, pronomes, etc. Por exemplo, os termos como "de", "a", "o", "que", "e", "do", "da", "em", "um", etc.. O funcionamento desse processo está demonstrado na figura 4.3 logo abaixo:

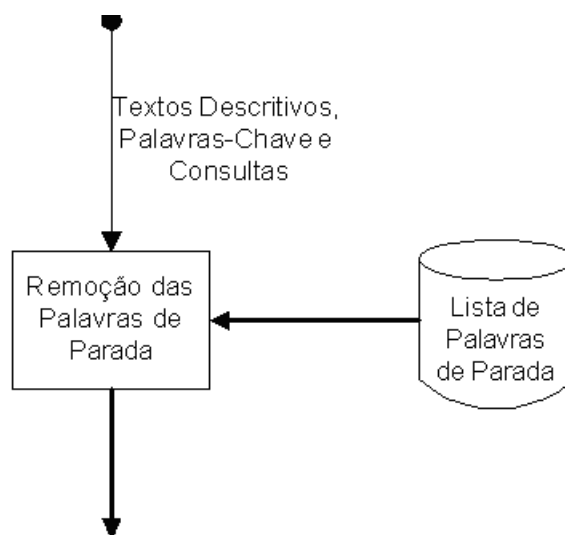


FIG. 4.3: Remoção das Palavras de Parada

##### Tratamento de Sinônimos

O sistema verifica se existem sinônimos para os termos dos textos descritivos, palavras-

chave e consultas, associando todos os seus sinônimos correspondentes. A partir deste ponto, os sinônimos passarão pelo mesmo tratamento dos termos originais, e terão o mesmo peso em todas as etapas da recuperação.

Os sinônimos dos termos estão relacionados na ontologia de domínio, pelo uso dos comentários(*comments*) associados aos conceitos da referida ontologia. Se o termo foi definido como conceito do domínio, poderá ou não ter um sinônimo correspondente. O funcionamento desta etapa pode ser visualizado na figura 4.4 logo abaixo:

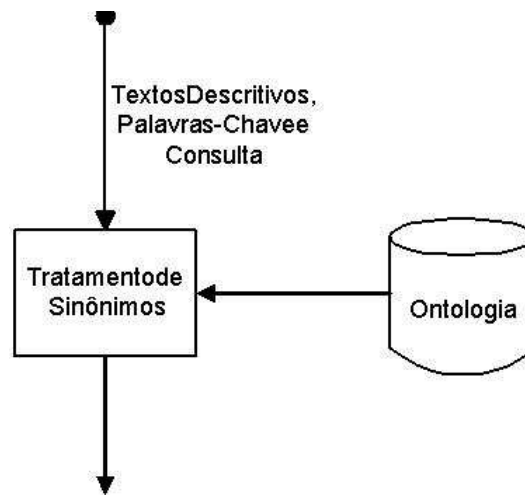


FIG. 4.4: Tratamento de Sinônimos

### Remoção de Afixos

O sistema analisa os textos descritivos, palavras-chave e consulta sem as palavras de parada e realiza uma remoção de afixos(stemming). Essa etapa é responsável por extrair o radical de cada termo, ou seja, obter a raiz da palavra que indica a família de palavras a qual pertence, através da remoção de dígitos, sinais de pontuações e afixos que indiquem advérbio, aumentativo, substantivo, gênero, plural, vogal temática, etc. Por exemplo, o termo "alunos" passado por uma remoção de afixos resultará em "alun". Nesse caso foi retirado o afixo "s" que representa um plural e "o" que representa uma vogal temática.

O algoritmo aplicado em nosso sistema foi o mesmo utilizado por (ORENGO, 2001), que se trata de um algoritmo de remoção que utiliza regras estabelecidas para tratar cada caso de afixo para a Língua Portuguesa. A estrutura de funcionamento desse algoritmo está exibida na figura 4.5 logo abaixo:

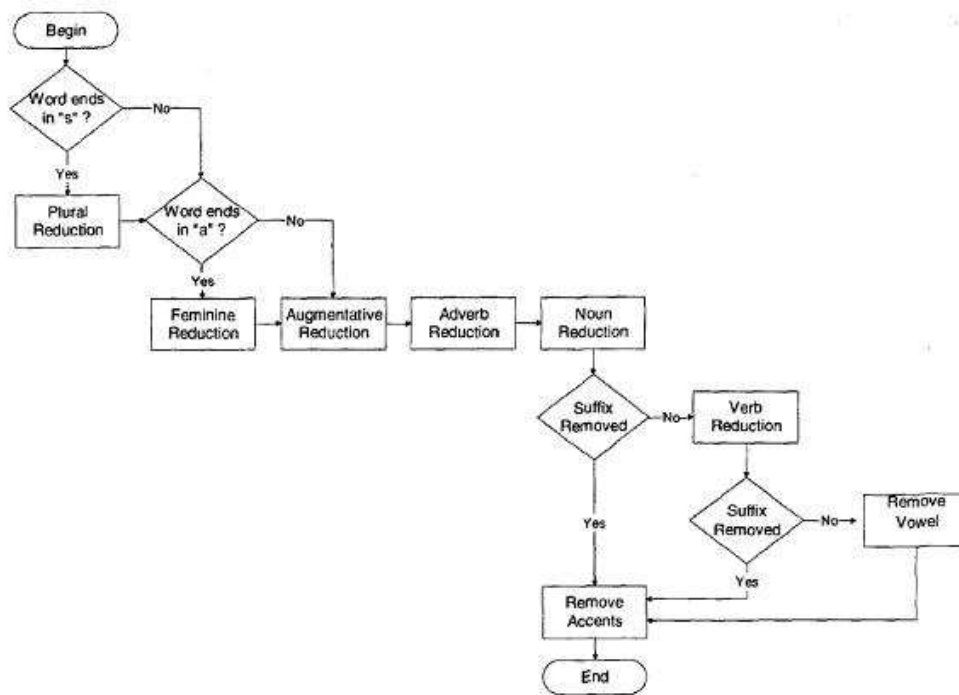


FIG. 4.5: Remoção de Afixos - Fonte: (ORENGO, 2001)

#### 4.2.2 INDEXAÇÃO

Esta seção trata da geração dos arquivos de índices invertidos a partir das imagens, textos descritivos e palavras-chaves, que foram adquiridos após o pré-processamento. Tais arquivos podem ser concebidos exclusivamente para os textos descritivos e as palavras-chaves, ou para ambos. Basicamente, os textos e as palavras-chave, passam pelo mesmo tratamento.

Os termos dos textos descritivos ou as palavras-chave, dependendo do tipo escolhido para gerar os índices, como foi citado anteriormente, passam por um cálculo de pesos. Primeiro o processo de indexação verifica se é um termo relevante dentro do contexto. Caso seja um termo integrante do domínio da aplicação, é atribuído um peso maior a esse termo. Também é atribuída uma importância maior, caso esse termo seja mais raro nos textos descritivos e palavras-chave, pois tem um poder discriminatório maior.

Ao fim, todos os termos são indexados e associados às figuras a que pertencem, gerando seus índices invertidos. Esses índices serão consultados posteriormente para a realização de uma consulta. O funcionamento deste processo está demonstrado na figura 4.6 logo abaixo:



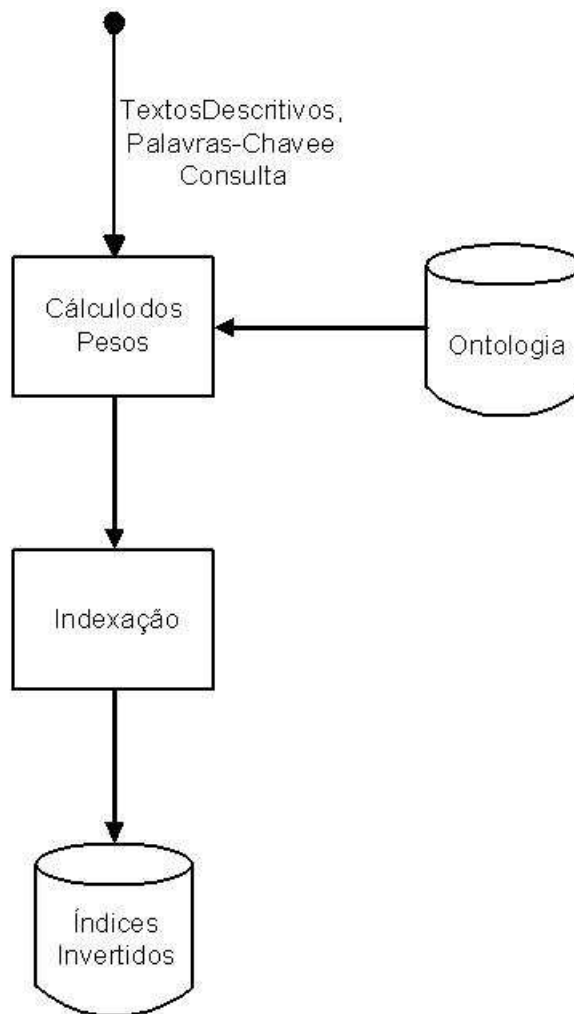


FIG. 4.6: Processo de Indexação

### 4.2.3 RECUPERAÇÃO

Esta seção apresenta a recuperação propriamente dita, permitindo ao usuário consultar por meio de sentenças em Linguagem Natural e visualizar uma lista de resultados com as imagens referentes a esta consulta. Existindo a opção do usuário do sistema de definir com qual(is) arquivo(s) de índices invertidos deseja trabalhar: com o dos textos descritivos, das palavras-chaves ou ambos. Caso seja escolhido a opção ambos, então deve ser especificado pelo usuário um peso para cada tipo de índice. A soma dos pesos deve estar compreendida entre 0 e 1. A estrutura usada para este processo é demonstrada na figura 4.7.

#### **Pré-Processamento**

O sistema realiza um pré-processamento análogo ao descrito para o processo de indexação.

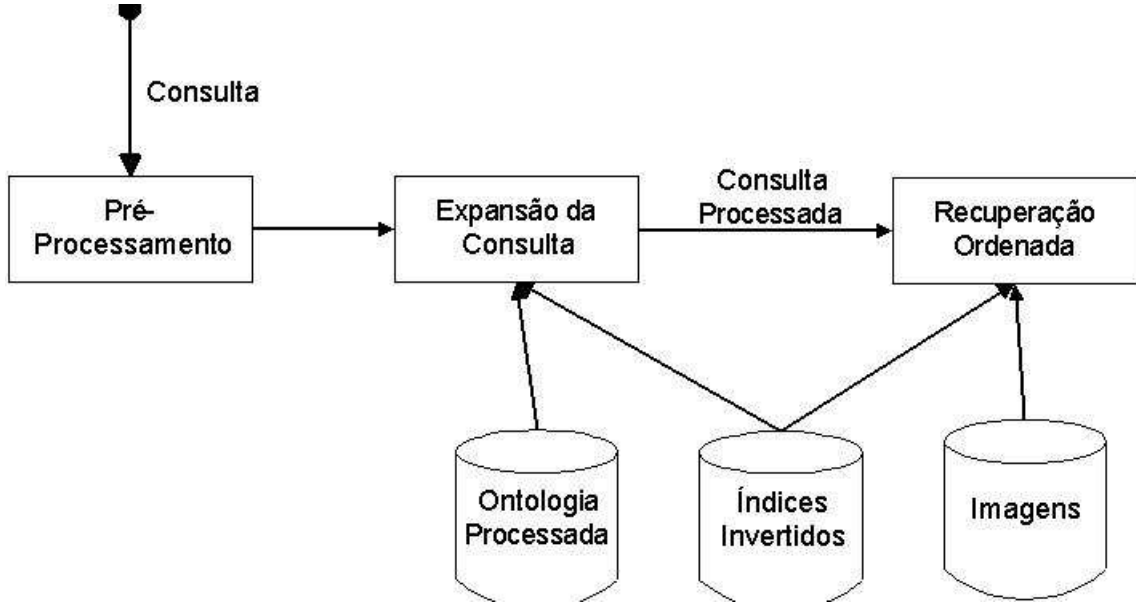


FIG. 4.7: Processo de Recuperação

### Expansão da Consulta

Nessa etapa o sistema atualiza os termos e os pesos da consulta. Verifica se há no arquivo de índice invertido, algum termo que seja similar ao da consulta, caso isso aconteça, este é acrescido na consulta juntamente com o seu peso. Contudo, se o termo da consulta estiver contido no índice, somente o seu peso é atualizado. E serão escolhidos os termos que obtiveram melhores pesos (os  $\max\_exp1$  termos dos termos similares).

A similaridade de cada termo  $t$  da consulta  $q$  é calculada da seguinte maneira:

$$sim(t, q) = \sum_{t_i \in q} w_{i,q} \times sim(t, t_i) \quad (4.1)$$

onde  $w_{i,q}$  é o peso do termo  $t_i$  (termo do arquivo de índice invertido) na consulta  $q$  e  $sim(t_i, t)$  é a similaridade calculada entre  $t_i$  e  $t$ . O cálculo da similaridade dos dois termos consiste em diagnosticar se são termos iguais ou parentes. O novo peso dos termos expandidos é calculado da seguinte maneira:

$$w_{ex}(t, q) = \frac{sim(t, q)}{\sum_{t_i \in q} w_{i,q}} \quad (4.2)$$

No cálculo da similaridade, representado pela função `SimTermos` (Figura 4.9) no algoritmo de expansão da consulta (Figura 4.8), é retornado 1, caso os termos sejam

**Algoritmo** Expansão\_Consulta(*consulta*, *indice*, *ontologia*)

1. *peso\_consulta* = 0;
2. **Para todos os termos** *u* **da** *consulta*
  - (a) *peso\_consulta* += *consulta*(*u*).*peso*;
3. **Para todos os termos** *t* **do** *indice*
  - (a) *sim*(*t*) = 0;
  - (b) **Para todos os termos** *u* **da** *consulta*
    - i. *sim*(*t*) += *consulta*(*u*).*peso* × **SimTermos**(*t*, *u*, *ontologia*, *indice*);
  - (c) *w<sub>ex</sub>*(*t*).*peso* = *sim*(*t*) / *peso\_consulta*;
4. **Para todos os termos** *t* **da** *consulta*
  - (a) **Se** *w<sub>ex</sub>*(*t*) **é** um dos *max\_exp* termos com maior peso
    - i. *consulta\_expandida*(*t*) = *consulta*(*t*);
    - ii. *consulta\_expandida*(*t*).*peso* += *w<sub>ex</sub>*(*t*).*peso*;
5. **Retorna** *consulta*;

FIG. 4.8: Algoritmo de Expansão da Consulta (PAZ-TRILLO, 2001)

sinônimos e 0 se um dos termos não estiver presente na ontologia. Entretanto, caso as condições anteriores não sejam satisfeitas, o algoritmo calcula dois fatores de similaridade: sobre a hierarquia dos conceitos e relacionamentos da ontologia. Como está ilustrada na equação 4.3, logo abaixo:

$$sim(t_1, t_2) = \beta \times \frac{2 \times \log_{sup}}{\log_{t_1} + \log_{t_2}} + (1 - \beta) \times \frac{props(t_1, t_2) + props(t_2, t_1)}{props(t_1) + props(t_2)} \quad (4.3)$$

O primeiro fator que foi proposto por (LIN, 1998) considera que *sup* é a frequência total das classes que possuem termo *t<sub>1</sub>* e o termo *t<sub>2</sub>* como sub-classes dentro da ontologia, e *t<sub>1</sub>* e *t<sub>2</sub>* as frequências dos termos. Enquanto, o segundo fator mede o relacionamento entre os termos por meio de suas propriedades. Onde *props*(*t<sub>1</sub>*, *t<sub>2</sub>*) representa a quantidade de propriedades que têm *t<sub>1</sub>* como domínio da propriedade e *t<sub>2</sub>* como imagem, *props*(*t<sub>2</sub>*, *t<sub>1</sub>*) o inverso, *props*(*t<sub>1</sub>*) e *props*(*t<sub>2</sub>*) com a quantidade de propriedades que contenham *t<sub>1</sub>* e *t<sub>2</sub>* respectivamente.

O valor de *b* é 0,5, portanto os dois fatores têm a mesma importância nesse cálculo,

**Algoritmo** SimTermos(*term1*, *term2*, *onto*, *indice*)

1. **Se** *term1* == *term2* **Retorna** 1.0;
2. *classes1* = *onto*.pegaClasses(*term1*), *classes2* = *onto*.pegaClasses(*term2*);
3. **Se** *classes1* é vazio ou *classes2* é vazio **Retorna** 0.0;
4. **Se** *classes1*  $\cap$  *classes2* não é vazio **Retorna** 1.0;
5. *classes\_sup* = *onto*.pegaClassesSuperioresComuns(*classes1*, *classes2*);
6. **Se** *classes\_sup* não é vazio
  - (a)  $probterm1 = \frac{indice.pegaFrequencia(classes1)}{indice.totalDocs}$ ;
  - (b)  $probterm2 = \frac{indice.pegaFrequencia(classes2)}{indice.totalDocs}$ ;
  - (c)  $probtermsup = \frac{indice.pegaFrequencia(classes\_sup)}{indice.totalDocs}$ ;
  - (d) **Se**  $\log(probterm1 \times probterm2) \neq 0$ 
    - i.  $atermo = \frac{2 \times \log(probtermsup)}{\log(probterm1 \times probterm2)}$ ;
7. (*totalProps1*, *propsRelacionadas1*) = *onto*.contarProps(*classes1*, *classes2*);
8. (*totalProps2*, *propsRelacionadas2*) = *onto*.contarProps(*classes2*, *classes1*);
9. **Se** *totalProps1* + *totalProps2* > 0
  - (a)  $btermo = \frac{propsRelacionadas1 + propsRelacionadas2}{totalProps1 + totalProps2}$ ;
10. **Retorna** ( $\beta \times atermo$ ) + (( $1 - \beta$ )  $\times btermo$ );

FIG. 4.9: Algoritmo de Similaridade de Termos (PAZ-TRILLO, 2001)

pois seus valores representam metade dessa medida.

### Recuperação Ordenada

Na etapa de recuperação, o sistema retorna uma lista de figuras, relativas a consulta em Linguagem Natural. Aqui é usado o modelo estendido de vetores, que foi apresentado no capítulo 2. Ou seja, cada termo tem armazenado em um vetor espacial, um peso agregado, sendo este de uma consulta ou dos documentos. É utilizado um algoritmo de recuperação (PAZ-TRILLO, 2001), ilustrado logo abaixo na figura 4.10, que é chamado uma vez para cada índice invertido (textos descritivos ou palavras-chaves).

Os valores retornados para cada chamada desse algoritmo para os índices invertidos,

**Algoritmo Recupera(*consulta*, *indice*)**

1. *peso\_total\_consulta* = 0;
2. **Para todos os termos** *termo* **na consulta**
  - (a) *entrada* = *indice.pegarEntradaTermo(termo)*;
  - (b) **Se** *entrada* **é nula** **processar seguinte termo**
  - (c) *peso\_consulta* = (*consulta(termo).peso*)<sup>2</sup>;
  - (d) *peso\_total\_consulta* += *peso\_consulta*;
  - (e) **Para todos os documentos** *doc* **na entrada**
    - i. *docs\_analisados(doc).cos* += *peso\_consulta* × *entrada(doc).peso*;
3. **Para todos os documentos analisados** *doc* **em docs\_analisados**
  - (a) *entrada* = *indice.pegarEntradaDoc(doc)*;
  - (b) **Para todos os termos** *termo* **em entrada**
    - i. *docs\_analisados(doc).peso\_doc* += (*entrada(termo).peso*)<sup>2</sup>;
  - (c) *docs\_analisados(doc).sim* = 
$$\frac{\text{docs\_analisados}(doc).cos}{\sqrt{\text{docs\_analisados}(doc).peso\_doc \times \text{peso\_total\_consulta}}}$$
;
4. **Retorna** *docs\_analisados*;

FIG. 4.10: Algoritmo de Recuperação (PAZ-TRILLO, 2001)

são atualizados e depois atribuídos um peso para cada um deles. O *kwweight*, o peso para o índice invertido para palavra-chave, assinalado no xml, serve para atribuir um valor de importância para os valores obtidos pelo índice invertido de palavra-chave e texto descritivo, ou seja, se um valor de 0.5 é atribuído para *kwweight*, quer dizer que o índice de palavra-chave tem um grau de importância de 50%, enquanto, o de textos descritivos tem a outra parte, os 50% restante.

O *cos*(cosseno do ângulo formado pelo vetor da consulta e vetor dos documentos) calculado no Algoritmo de Recuperação é representado da seguinte forma:

$$\cos_{doc,q} = \sum_{t_i \in doc} w_{i,q} \times w_{i,doc} \quad (4.4)$$

onde  $w_{i,q}$  é o peso do termo  $i$  da consulta  $q$  e  $w_{i,doc}$  é o peso de cada documento  $doc$  que contenha o termo  $i$ . E para a medida da similaridade é formulada da seguinte maneira:

$$sim_{doc,q} = \frac{\cos_{doc,q}}{\sqrt{\sum_{t_i \in doc} w_{i,doc} \times \sum_{t_j \in q} w_{j,q}}} \quad (4.5)$$

onde  $w_{j,q}$  é o peso total da consulta, ou seja, a soma de todos os pesos da consulta e  $w_{i,doc}$  é o peso de cada documento.

E, por fim, é exibida a listagem de max.imagens(número máximo de imagens relacionados), valor de parâmetro do sistema. Melhor dizendo, os max.imagens que forem mais pertinentes pelo modelo de espaço de vetores(SALTON, 1975).

### 4.3 RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO

Nesta seção será descrita uma outra abordagem para recuperação de imagens. A idéia nesta abordagem é permitir consultas, com uso de características de imagens(Cor e Textura), por meio de figuras-exemplo. Nesta abordagem também são necessárias três etapas: pré-processamento, indexação e recuperação. No pré-processamento são extraídas as características de imagem através de algoritmos de cálculo de descritores. Na indexação, as informações calculadas são anexadas à imagem correspondente em um arquivo XML. Na recuperação são listados os resultados da consulta.

Conforme ilustrado no diagrama conceitual da figura 4.11, a abordagem de recuperação baseada em Conteúdo é composta de duas etapas principais como a anterior: indexação e recuperação. Os itens de 1 a 3, mostrados abaixo, indicam a etapa de indexação e o restante a etapa de recuperação, sendo que a linha contínua representa a indexação e a tracejada a recuperação:

- (1) Relaciona todas as figuras envolvidas no contexto, que posteriormente serão recuperadas pelo sistema;

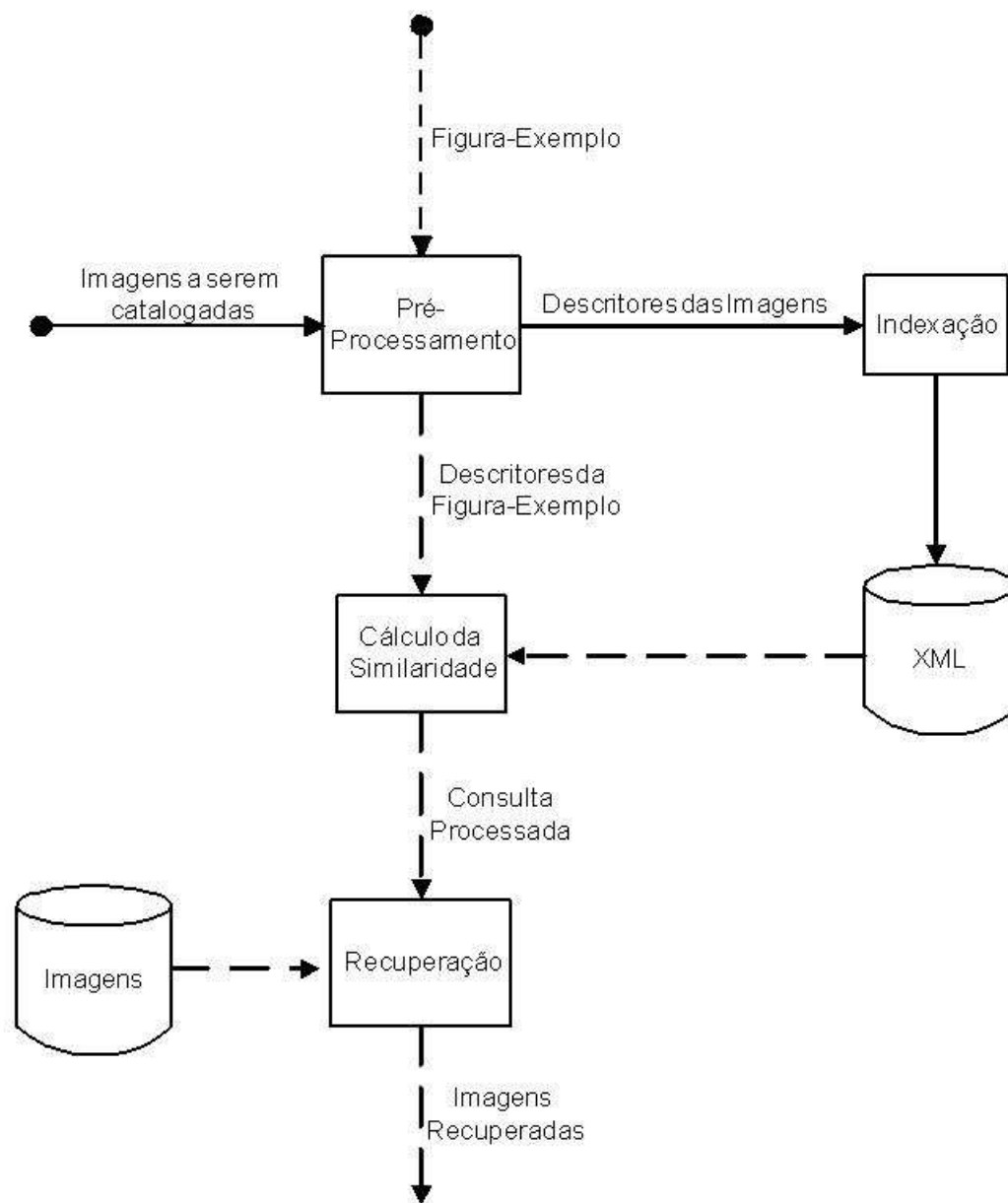


FIG. 4.11: Recuperação Baseada em Conteúdo - Visão Geral

- (2) Com as figuras devidamente relacionadas, é feito um pré-processamento destas. Extraindo características das imagens através de descritores de cores e textura (seção 2.5.2);
- (3) Com os valores dos descritores, a imagem é indexada. Neste caso, as imagens e os respectivos descritores são armazenados no arquivo XML;
- (4) Para iniciar uma consulta, é escolhida uma figura-exemplo, como foi citado anteriormente;
- (5) A figura-exemplo escolhida é submetida a um pré-processamento, que extrai as suas características de imagem;
- (6) Com a figura-exemplo pré-processada, é feito o cálculo de similaridade entre esta e as figuras indexadas, que estão armazenadas no arquivo XML. O cálculo é obtido pela distância euclidiana;
- (7) Com o cálculo de similaridade feito, os resultados são passados em uma lista. Depois são ordenados pela menor distância, ou seja, aquelas que forem mais parecidas serão mostradas primeiras.

#### 4.3.1 PRÉ-PROCESSAMENTO

Nesta etapa é realizado o tratamento da imagem, tanto para o processo de indexação quanto para o processo de consulta. Aqui é feita uma extração das características de imagens, representadas por descritores de imagens (seção 2.5.2), mais especificamente de cores e texturas. Os descritores escolhidos para cores foram Média, Desvio Padrão, Obliquidade e Curtose (seção 2.5.2.1) e para textura foram Energia, Contraste, Entropia, Homogeneidade, Inverso da Contraste e Distorção (seção 2.5.2.2). O funcionamento dessa etapa é ilustrado na figura 4.12 que se encontra logo abaixo:

Para calcular os descritores é usado um algoritmo de cálculo de descritores. Este algoritmo é ilustrado na figura 4.13 logo abaixo:

O procedimento MatrizCoOcorrência, na linha 2 do algoritmo, serve para calcular a Matriz de Co-Ocorrência (Grey Level Co-Occurrence Matrix - GLCM) (seção 2.5.2.2). Uma matriz de Co-Ocorrência é uma tabulação sobre o número de combinações distintas de valores de intensidade que ocorrem em uma imagem, levando em consideração a posição de pixels que possuem níveis de cinza (SCHWARTZ, 2003). Depois de calculada a matriz



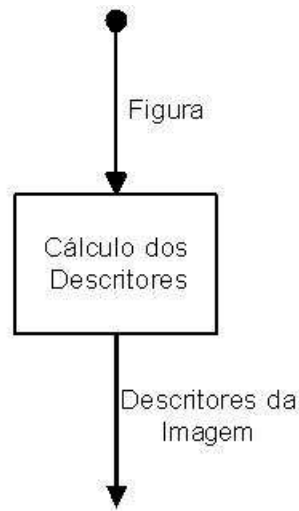


FIG. 4.12: Pré-Processamento da Imagem

Algoritmo `GetDescritores(path)`

```

1. descritores = "";
2. matriz = MatrizCoOcorrência(path, 1, 1, 256);
3. descritores = GetTextura(matriz, 256);
4. descritores += GetCores(path);
5. retorna descritores;
  
```

FIG. 4.13: Algoritmo para Cálculo de Descritores

de Co-Ocorrência, podem ser calculados os descritores, na função `GetTextura`, na linha 3 do algoritmo. Abaixo se encontra o algoritmo `GetTextura` na figura 4.14, que calcula os descritores relacionados a textura:

Os descritores de textura(seção 2.4.2.1.2) são os seguintes:

- Energia: é obtida pela soma dos quadrados dos elementos da matriz de Co-Ocorrência. É usado para denotar a homogeneidade da imagem;

$$Energia = \sum_i \sum_j (p(i, j, d, \theta))^2 \quad (4.6)$$

- Contraste: representa a diferença de níveis de cinza(diferença de momento) dos pixels de uma imagem;

$$Contraste = \sum_i \sum_j (i - j)^2 (p(i, j, d, \theta))^2 \quad (4.7)$$

```

Algoritmo GetTextura(matrizCoOcorrência , tonsCinza)

1. descritoresdeTextura = "";
2. para r de 0 até tonsCinza
  a. para c de 0 até tonsCinza
    i. Energia+=matrizCoOcorrência[r][c] *
      matrizCoOcorrência[r][c];
    ii. Contraste+=(r-c)2*matrizCoOcorrência[r][c];
    iii. Se matrizCoOcorrência[r][c] é maior que 0 então
      I. Entropia+=matrizCoOcorrência[r][c] *
        log(matrizCoOcorrência[r][c]);
      II. Homogeneidade+=matrizCoOcorrência[r][c] /
        (1+abs(r-c));
      III. Distorção+=(r-c)3*matrizCoOcorrência[r][c];
      IV. InversodoContraste+=matrizCoOcorrência[r][c] /
        (1+(abs(r-c))2);
3. descritoresdeTextura = Energia + "##" + Contraste +
  "##" + Entropia + "##" + Homogeneidade + "##" + Distorção +
  "##" + InversodoContraste;
4. retorne descritoresdeTextura;

```

FIG. 4.14: Algoritmo para cálculo de descritores de textura

- Entropia: representa o grau de dispersão dos níveis de cinza de uma imagem. Este descritor, juntamente com SMA, pode denotar homogeneidade;

$$Entropia = - \sum_i \sum_j (p(i, j, d, \theta)) \log(p(i, j, d, \theta)) \quad (4.8)$$

- Homogeneidade: a representação de homogeneidade assume valores altos quando uma dada textura apresenta uma baixa variação entre tons de cinza;

$$Homogeneidade = \sum_i \sum_j \frac{(p(i, j, d, \theta))}{1 + \|i - j\|} \quad (4.9)$$

- Distorção: Momento de terceira ordem, representa a distorção da imagem;

$$Distorcao = \sum_i \sum_j (i - j)^3 (p(i, j, d, \theta)) \quad (4.10)$$

- Inverso da Contraste: representa a variância inversa.

$$InversodaContraste = \sum_i \sum_j \frac{(p(i, j, d, \theta))}{(i - j)^2} \quad (4.11)$$

O procedimento GetHistograma na linha 2 do algoritmo GetCores, ilustrado na figura 4.15, calcula um vetor que contém os valores do histograma de cores, ou seja, a frequência dos pixels das cores pertencente ao padrão RGB. A partir desse histograma é possível calcular os descritores de cor, os quais são:

```

Algoritmo  GetCores(path)

1.  DescritoresdeCor = Média = DesvioPadrão =
    Obliquidade = Curtose = "";
2.      GetHistograma(path);
3.      Média = GetMédia(histograma);
4.      DesvioPadrão = GetDesvioPadrão(histograma);
5.      Obliquidade += GetObliquidade(histograma);
6.      Curtose += GetCurtose(histograma);
7.      DescritoresdeCor = Média + "###" + DesvioPadrão
    + "###" + Obliquidade + "###" + Curtose;
8.      retorne DescritoresdeCor;

```

FIG. 4.15: Algoritmo para cálculo de descritores de Cor

- Média: representa a média dos pixels encontrados no histograma;

$$Media = \frac{\sum_{i=1}^n x_i}{n} \quad (4.12)$$

- Desvio Padrão: representa o desvio padrão dos pixels encontrados no histograma;

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - Media)^2}{n}} \quad (4.13)$$

- Obliquidade: representa a distribuição das frequências dos pixels no histograma;

$$Obliquidade = \frac{m_3}{\sigma^3} \quad (4.14)$$

- Curtose: representa o achatamento da curva da função de distribuição no histograma.

$$Curtose = \frac{m_4}{\sigma^4} - 3 \quad (4.15)$$

### 4.3.2 INDEXAÇÃO

Esta etapa é responsável por armazenar as características de imagens, tais como os descritores textura e cor. A complexidade dessa tarefa não é tão árdua como a indexação da abordagem descrita na seção 4.2.2, pois a cada inserção de nova figura são registrados seus descritores, e assim sendo colocados diretamente no arquivo xml. Na figura 4.16 logo abaixo, a parte que compõe o rótulo descritores, são os valores dos descritores da imagem separados por "##". Os descritores são ordenados da seguinte maneira: Energia, Contraste, Entropia, Homogeneidade, Distorção, Inverso da Contraste, Média, Desvio Padrão, Curtose e Obliquidade. Sendo, respectivamente, os seis primeiros são de textura e os restantes de cor.

```
<clip id="0" tipo="figura" nome="Fábio Chagas"

nomearquivo="D:\Eclipse\workspace\Prototipo\
Informativos\inf0205_fig1.bmp"

arquivotexto="D:\Eclipse\workspace\Prototipo\
Informativos\inf0205_fig1.txt"

descritores = "12403140##18845041##253652,860932256##
37428,8517259489##5615193##33798,1153021927##
1,31298059895833E+002##6,74806083008972E+001##
2,44896698430931E+000##-4,23891259320257E-001"
figuraorigem="null"/>
```

FIG. 4.16: Algoritmo para cálculo de descritores de Cor

A figura 4.17 mostra logo abaixo como é conceitualmente o processo de indexação para a recuperação baseada em conteúdo.

### 4.3.3 RECUPERAÇÃO

Nesta etapa é desenvolvida uma aplicação que realiza a recuperação das imagens por meio de uma figura-exemplo, ou seja, permitindo ao usuário definir uma figura em sua consulta e visualizar uma lista de resultados com as imagens similares à apresentada por imagens armazenadas. A busca de similares é efetuada calculando a distância (nesse caso, a euclidiana) entre os descritores das imagens armazenadas e as da figura-exemplo.

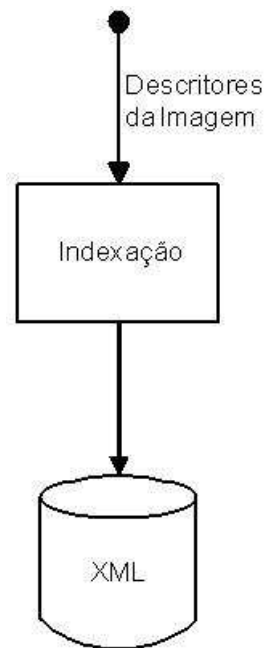


FIG. 4.17: Processo de Indexação

As imagens que retornarem a menor distância entre estas, serão as mais relevantes. A representação desse processo é melhor ilustrada na figura 4.18, logo a seguir.

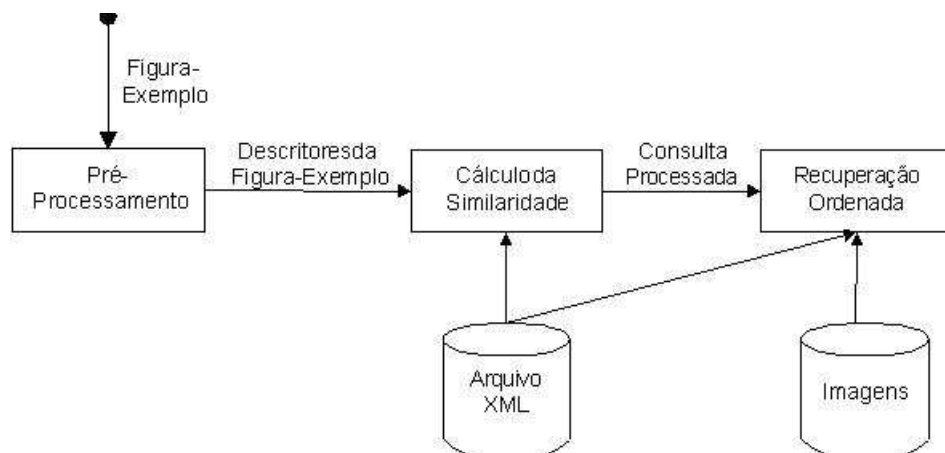


FIG. 4.18: Processo de Recuperação

### **Pré-Processamento**

É usado o mesmo pré-processamento da indexação.

### **Cálculo da Similaridade**

Nessa etapa, o sistema pega as informações no arquivo xml a respeito dos descritores de imagens e carrega estes para que assim possa calcular a distância (euclidiana) entre estas e a figura-exemplo. Há um cuidado para associar o descritor da figura-exemplo com

os seus respectivos descritores das imagens, tanto que é utilizado o mesmo algoritmo de extração de descritores para a indexação e recuperação. A distância escolhida para efetuar este cálculo foi à euclidiana, como é exemplificada na equação 4.16.

$$Distancia_{i,j} = \sum_{v=1}^n \sqrt{(D_{i,v} - D_{j,v})^2} \quad (4.16)$$

onde i e j são as imagens e v todos os descritores.

Para cálculo da distância entre a figura-exemplo e a cada uma das imagens do banco de imagens é usado o algoritmo CalcularDistância, que é apresentado logo abaixo:

```
Algoritmo CalcularDistancia( DescritoresFiguraExemplo ,
DescritoresFigura )
```

1. Distância = i = diferença = 0;
2. enquanto houver descritor
  - a. diferença += DescritoresFiguraExemplo[i] - DescritoresFigura[j];
  - b. Distância += diferença \* diferença;
3. Distância = raiz\_quadrada(Distância);
4. Retorna Distância;

FIG. 4.19: Algoritmo para cálculo de similaridade de duas imagens

### Recuperação Ordenada

Na etapa de recuperação, o sistema retorna uma lista de imagens, relativas a consulta. Retorna e ordena as imagens de acordo com a distância calculada no processo anterior. As imagens que possuírem as características de imagens mais próximas serão apresentadas logo no início da lista. As imagens são apresentadas em ordem crescente de distância em relação à imagem de exemplo apresentada como consulta.

## 4.4 COMBINAÇÃO DAS ABORDAGENS DE RECUPERAÇÃO DE IMAGENS

A principal contribuição do presente trabalho é conseguir combinar as duas abordagens citadas anteriormente, ou seja, o sistema ter a capacidade de recuperar imagens tanto baseadas por sentenças quanto por conteúdo. Tentar de alguma forma atacar um dos grandes desafios da recuperação de imagens(TORRES, 2007), desenvolvimento de novos algoritmos de fusão de dados que suporta recuperação baseada em texto e conteúdo

combinando informações de diferentes formatos. A maneira como é executada essa abordagem é mostrada na figura 4.20. Na abordagem proposta, a combinação das abordagens é realizada por meio da interseção de dois conjuntos de imagens: as imagens recuperadas a partir da abordagem de recuperação baseada em sentenças e as imagens recuperadas a partir da abordagem de recuperação baseada em conteúdo.

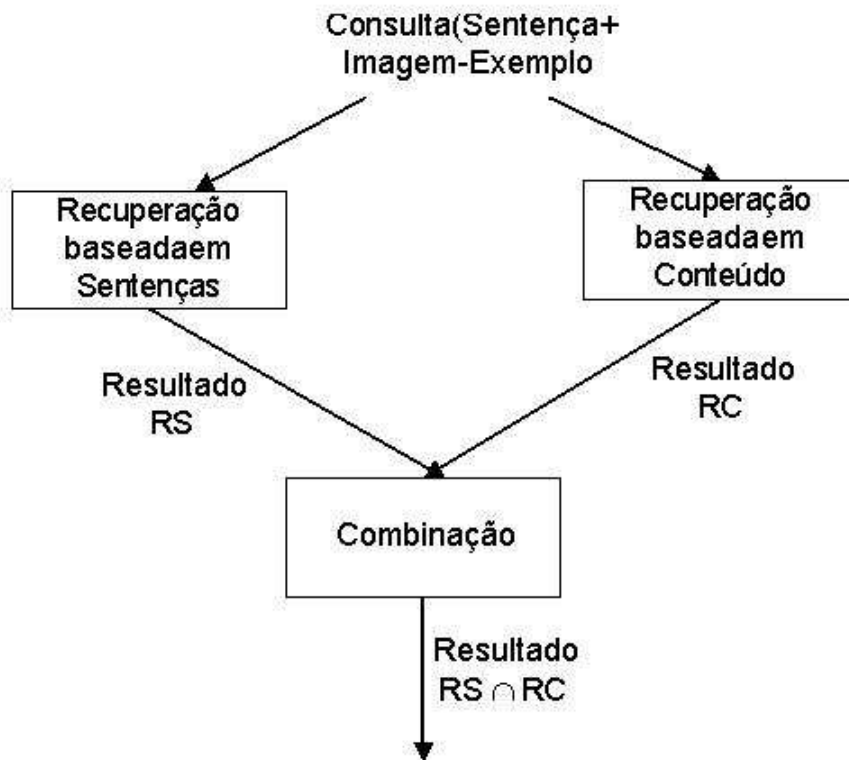


FIG. 4.20: Recuperação Combinando as Duas Abordagens

## 5 PROTÓTIPO, EXPERIMENTOS E RESULTADOS

Este capítulo tem por objetivo descrever em detalhe o protótipo implementado, a metodologia de avaliação da solução, assim como os resultados obtidos.

### 5.1 PROTÓTIPO

O protótipo implementado encontra-se estruturado em três módulos: módulo de recuperação de imagens baseada em sentenças, módulo de recuperação de imagens baseada em conteúdo e módulo de recuperação de imagens baseada em sentenças e em conteúdo.

A linguagem de programação adotada no desenvolvimento dos três módulos do protótipo foi a linguagem Java.

#### 5.1.1 MÓDULO DA RECUPERAÇÃO BASEADA EM SENTENÇA

O desenvolvimento do módulo de recuperação de imagens baseada em sentenças foi fortemente influenciado pelo trabalho descrito em (PAZ-TRILLO, 2001). Como comentado no capítulo de trabalhos relacionados, (PAZ-TRILLO, 2001) desenvolveu um sistema para recuperação de vídeos a partir de consultas formuladas em linguagem natural. A implementação desenvolvida por (PAZ-TRILLO, 2001) foi utilizada como ponto de partida para a implementação do protótipo desta dissertação. Diversos foram os pontos de adaptação e melhorias implementados. Entre os principais, podem ser citados: o tratamento de imagens ao invés de vídeos, o novo domínio da aplicação e o tratamento de palavras compostas.

Na etapa de indexação, foi usada a linguagem Java para escrever e registrar os arquivos, usando métodos de entrada e saída(Java.io), principalmente nos arquivos de índices invertidos. Para escrita e manipulação do arquivo xml, foi usado o analisador SAXParser, um pacote Java para tratar de arquivos xml. Para criar a ontologia foi utilizado o ambiente Protégé(PROTÉGÉ, 2002). Para navegar sobre a ontologia foi usado uma API Java, chamada (JENA, 2002), específica para manipular ontologias. O Jena foi usado tanto para processar a ontologia e calcular os pesos para os termos indexados, explicitados anteriormente no capítulo 4. Na recuperação foram usados os mesmos recursos(pacotes e



API) do Java utilizadas na indexação.

A figura 5.1 mostra tela do protótipo responsável pela etapa de indexação, ou seja, a associação das imagens com os seus recursos. O usuário pode criar e manipular um arquivo xml, contendo tal associação de forma a utilizá-lo posteriormente na etapa de recuperação. Permite introduzir uma imagem por vez, preenchendo os itens requeridos para formar esse arquivo. As imagens também podem ser excluídas uma a uma. No final de todas as operações, pode ser salvo o arquivo, armazenando as alterações desejadas.

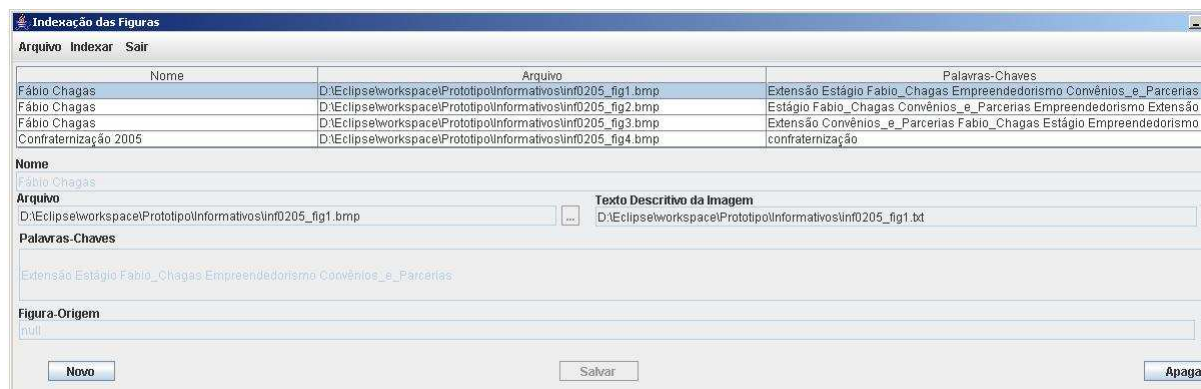


FIG. 5.1: Associação das Imagens

Na etapa de associação das imagens é feita a ligação das imagens com os seus respectivos recursos: nome da imagem, arquivo de imagem, arquivo de texto descritivo, palavras-chaves e figura de origem. A seguir uma descrição de cada um destes recursos.

### Nome da Imagem

É um identificador para a imagem, mais especificamente um nome para ela, a título de apresentação. Um campo que pode ser preenchido por um especialista no domínio da aplicação para rotular uma imagem.

### Arquivo de Imagem

Trata-se do endereço e do nome do arquivo de armazenamento da imagem em questão. O campo deve ser preenchido por um caminho, onde se encontra localizado o arquivo.

### Arquivo de Texto Descritivo

Trata-se do arquivo onde se encontra a descrição textual da imagem, ou seja, aquelas informações acerca da imagem, coletadas de um especialista no domínio da aplicação e apresentadas em linguagem natural. Como o item anterior, é representado por um caminho, onde se encontra armazenado. Esse arquivo é necessário para que posteriormente seja realizado o processo indexação por texto completo, conforme descrito no capítulo 4.

## Palavras-Chaves

São as palavras-chaves que o especialista deve dar entrada, para que assim possa fazer a indexação por palavras-chaves(capítulo 4). A palavra-chave pode ser composta, ou seja, possuir mais de um termo. Nesse caso, a palavra deve ser escrita com " \_ " nos lugares do espaço vazio(" ").

## Figura de Origem

É uma referência a um arquivo que contém uma imagem, de onde foi extraída a imagem que está sendo trabalhada nessa associação, ou seja, é uma sub-imagem de outra. Esta informação foi adicionada para permitir a recuperação de imagens contidas em outras imagens, funcionalidade necessária ao módulo de recuperação de imagens baseada em sentenças e em conteúdo, descrito mais à frente.

A figura 5.2 exposta logo abaixo, contém a tela do protótipo responsável pela última etapa da indexação, que trata da geração dos arquivos de índices invertidos. O especialista opta como deseja realizar esse processo. Escolhendo o tipo de índice que deseja produzir(de palavras-chave ou de descrição textual), remoção ou não de afixos e a ontologia que vai ser considerada para representar o domínio das imagens.

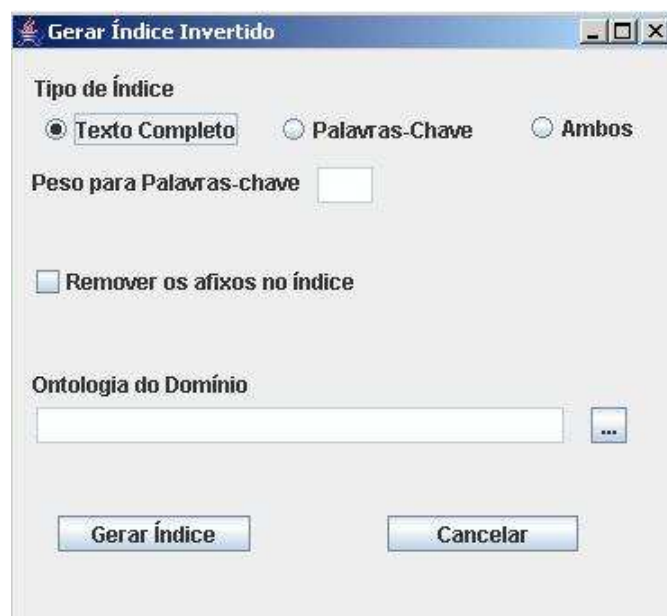


FIG. 5.2: Geração dos Arquivos de Índice Invertido

## Tipo de Índice

O usuário deve escolher entre "Texto Completo", "Palavras-Chave" e "Ambos", que tratam da indexação por descrições textuais, palavras-chave ou pelos dois métodos

simultaneamente. O campo "Peso para Palavras-Chave", atribui um valor que quantifica a importância do índice de "Palavras-Chave" na recuperação. Este campo depende do tipo de índice escolhido. Quando é escolhido "Texto Completo" o campo recebe automaticamente o valor 0. No caso de "palavras-chave" o valor atribuído automaticamente é 1. E no caso de "ambos" o campo é habilitado para que o usuário possa entrar com o valor que for pertinente, entre 0 e 1.

### **Remover Afixos no Índice**

É uma alternativa que o usuário tem, para utilização do índice na recuperação das imagens. Podendo optar por índices sem alteração nenhuma ou com uma remoção de afixos, para obter um radical da palavra para assim ser indexada. Esse procedimento foi detalhado anteriormente no capítulo 4.

### **Ontologia de Domínio**

Nesta parte é escolhida uma ontologia que será processada e considerada na etapa de indexação. Aplicações como cálculo de pesos dos termos, tratamento de sinônimos e de termos compostos são realizadas com auxílio dessa ontologia.

As saídas desse processo são três arquivos: um formado por índices invertidos para descrição textual, um para palavras-chave e uma ontologia processada. O primeiro contém todos os termos das descrições textuais e suas associações com as respectivas imagens. O segundo recebe os índices das palavras-chave, com suas respectivas imagens. O último armazena uma ontologia processada, ou seja, a ontologia de domínio trabalhada e modificada para facilitar o processamento da recuperação.

Na figura 5.3 é mostrada a tela que permite a recuperação das imagens. O usuário deve se ater aos itens consulta e "tipo de consulta", para fazer uma recuperação baseada em sentença. A opção figura-exemplo é utilizada pela abordagem de recuperação baseada em conteúdo. Na consulta, o usuário entra com uma sentença em linguagem natural que é o ponto de partida para a recuperação. O "tipo de consulta" serve para escolher com qual arquivo de índice, o usuário deseja recuperar as imagens, ou seja, tentar fazer um casamento dos termos da consulta com os termos do índice escolhido. As imagens recuperadas serão expostas em uma tabela que contém os campos Nome, Arquivo e Palavras-Chaves, sendo estes o nome da imagem, caminho do arquivo da imagem e as palavras-chave da imagem, respectivamente. Existe um limite máximo de imagens a serem recuperadas por consulta.

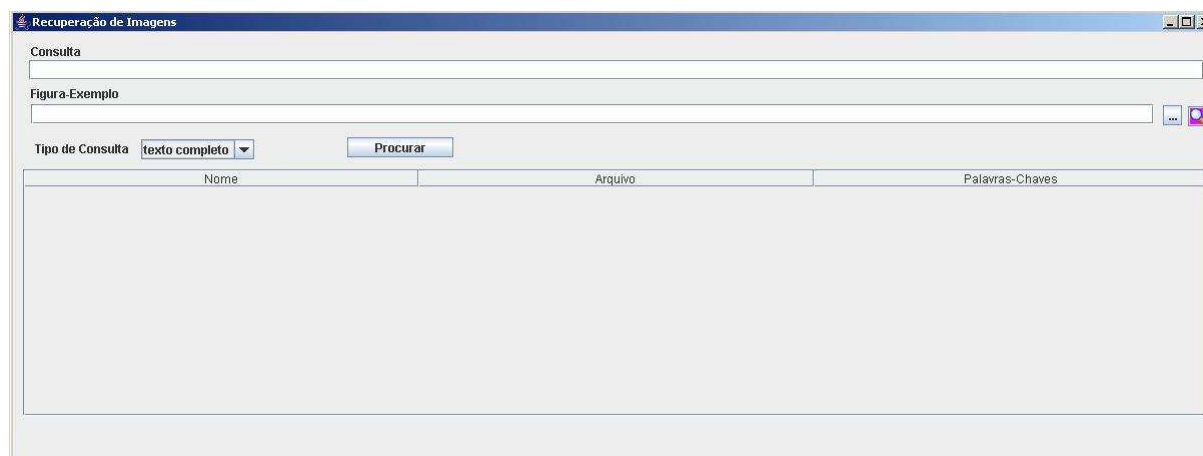


FIG. 5.3: Recuperação de Imagens

### 5.1.2 MÓDULO DA RECUPERAÇÃO BASEADA EM CONTEÚDO

O desenvolvimento do módulo de recuperação de imagens baseada em Conteúdo foi feito em linguagem Java integrado a métodos de cálculo de descritores implementados em Pascal. Tal integração foi realizada com auxílio do pacote JNI (Java Native Interface) que permite ao Java acionar métodos disponíveis em dll's (dynamic-link library). Assim, os métodos de cálculo de descritores foram primeiramente incorporados em uma dll específica e depois integrados à aplicação Java.

Cabe ressaltar que, por restrição da dll gerada, o protótipo implementado encontra-se restrito à manipulação de imagens armazenadas no formato bmp (bitmap).

A indexação feita nesse módulo é transparente para o usuário, ou seja, o protótipo faz o cálculo dos descritores e associação com as imagens de maneira automática. Quando o usuário insere uma nova imagem são calculados os seus descritores prontamente, sendo estes depois associados com a imagem inserida. Somente quando é confirmado o registro da imagem, é feita a inserção efetiva da imagem com seus descritores no arquivo xml.

Na recuperação, o usuário deve se ater aos campos Figura-exemplo e o "tipo de consulta" da tela mostradas na figura 5.3. Nelas deve dar uma entrada a um caminho que represente esta figura-exemplo e o "tipo de consulta" deve ser igual a "imagem". Entretanto, deve ser lembrado que a imagem que representa a figura-exemplo tem de estar ajustada, com a sua luminosidade padronizada, para evitar futuras distorções no cálculo da distância entre ela e as outras imagens, como foi mencionado no capítulo 4. Nesta recuperação, não há uma preocupação com a limitação de imagens a serem exibidas

e nem uma distância mínima a ser considerada para a filtragem dos resultados.

Os descritores implementados nessa abordagem foram: (a) De textura: Energia, Contraste, Entropia, Homogeneidade, Inverso do Contraste e Distorção; (b) De cor: Média, Desvio Padrão, Obliquidade e Curtose.

O campo "figura-origem" encontrado na tela da figura 5.1 serve para o especialista entrar com uma imagem que origina a figura que está sendo requisitada o preenchimento desse campo. Quando o sistema recuperar uma imagem que contenha esse valor fornecido, a imagem que tiver sendo indicada será considerada como uma figura candidata da recuperação.

A figura 5.4 logo abaixo mostra a tela com os resultados obtidos por uma consulta. Cada linha da tabela representa uma imagem recuperada. Quando uma delas é selecionada, o sistema apresenta a imagem correspondente em uma nova janela.

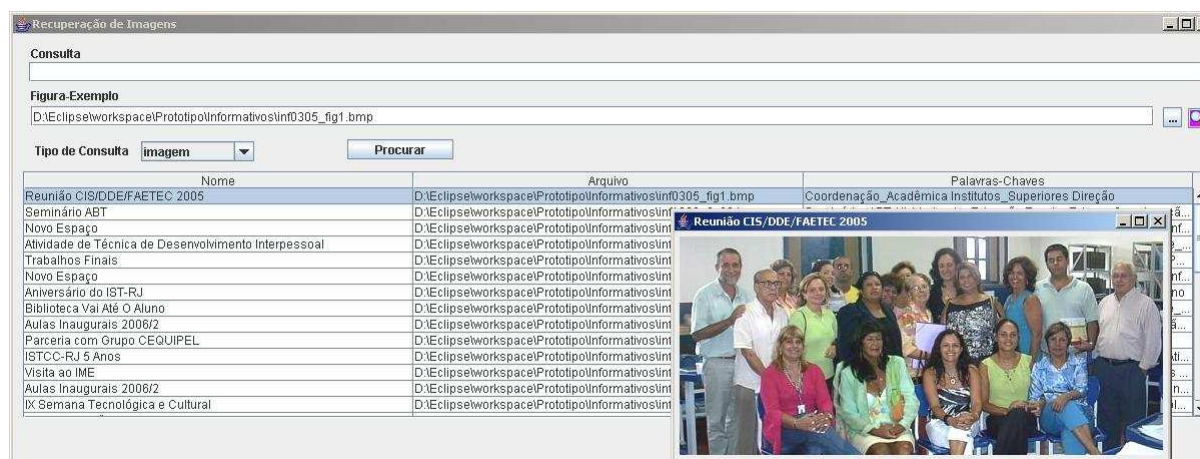


FIG. 5.4: Resultado da Recuperação

### 5.1.3 MÓDULO DE RECUPERAÇÃO DE IMAGENS BASEADO EM SENTENÇAS E EM CONTEÚDO

Conforme descrito no capítulo 4, a abordagem de recuperação de imagens baseada em sentenças e em conteúdo consiste da combinação das abordagens de recuperação de imagens baseada em sentenças e recuperação de imagens baseada em conteúdo. Tal combinação é materializada com a interseção dos conjuntos de imagens recuperadas por cada abordagem individualmente.

Neste módulo, a etapa de indexação segue os mesmos passos descritos para as

abordagens anteriores. Vide figuras 5.1 e 5.2. Na etapa de recuperação, o módulo permite que sejam informadas uma consulta em linguagem natural e uma figura-exemplo, conforme ilustra a figura 5.5.

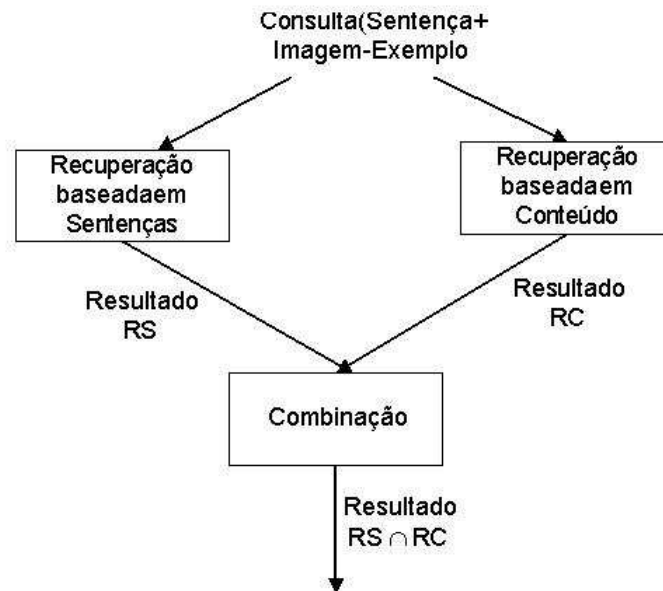


FIG. 5.5: Resultado da Recuperação

## 5.2 AVALIAÇÃO DO SISTEMA

Nesta seção são apresentados a metodologia de testes adotada assim como os resultados obtidos e uma análise sobre eles.

### 5.2.1 METODOLOGIA DOS TESTES

Nesta seção encontram-se delineados os testes das três abordagens de recuperação de imagens desenvolvidas na presente dissertação: baseada em sentença, baseada em conteúdo e em ambas. Em função da natureza distinta das abordagens, cada uma delas foi avaliada de uma forma distinta.

O domínio da aplicação sobre o qual todos os testes do protótipo foram realizados encontra-se descrito a seguir.

## 5.2.2 CONTEXTO DA APLICAÇÃO

Embora desenvolvido de forma independente do domínio da aplicação, o protótipo foi configurado e avaliado para atender às necessidades de uma instituição de ensino, o IST-Rio (Instituto Superior de Tecnologia em Ciência da Computação do Rio de Janeiro), unidade de ensino superior pertencente à rede FAETEC (Fundação de Apoio à Escola Técnica do Estado do Rio de Janeiro). O IST-Rio possui quase 1500 fotos acerca de suas realizações, eventos, colaborações e corpo de trabalho. Contudo, essas imagens não se encontram muito bem catalogadas, o que dificulta muito a sua recuperação.

Assim sendo, parte do presente trabalho consistiu em organizar, padronizar e tratar essas imagens, para que estas pudessem, posteriormente, ser recuperadas pelo protótipo desenvolvido. Entrevistas junto a diversos funcionários e a coleta de informações em informativos da instituição foram realizadas. Também houve um estudo prévio da estrutura organizacional da instituição e da memória de todas as ações realizadas pelo IST-Rio.

## 5.2.3 METODOLOGIA DE TESTES PARA RECUPERAÇÃO BASEADA EM SENTENÇAS

Geralmente, a avaliação de um sistema de recuperação de informação é dividida em duas partes (DOSSANTOS, 2005):

- a) Análise de eficiência: custo e tempo;
- b) Análise de efetividade: precisão e cobertura.

A eficiência é utilizada basicamente para medir o custo e o tempo. Por exemplo, para avaliar o tempo gasto com operações, como indexação e recuperação, e também a complexidade dos algoritmos de busca e consumo requerido de memória.

A efetividade, segundo (SALTON, 1983), é uma medida para mensurar a capacidade de um sistema de recuperação de informação retornar documentos relevantes dentro da recuperação. Ou seja, um sistema é bem efetivo quando satisfaz os anseios do usuário.

A maneira escolhida para avaliar a abordagem de recuperação de imagens baseada em sentenças foi a efetividade, sendo que esta é de uso comum e de fácil entendimento. Para explicar melhor o conceito de efetividade, a tabela 5.1, apresenta as associações

	Relevante	Não-Relevante	
Recuperado	$A \cap B$	$\bar{A} \cap B$	B
Não-Recuperado	$A \cap \bar{B}$	$\bar{A} \cap \bar{B}$	$\bar{B}$
	A	$\bar{A}$	N

TAB. 5.1: Tabela de Distribuições

entre *imagens relevantes*  $X$  *imagens não-relevantes* e *imagens recuperadas*  $X$  *imagens não-recuperadas*.

Tem-se que:

$$Precisao = \frac{|A \cap B|}{|B|} \quad (5.1)$$

$$Cobertura = \frac{|A \cap B|}{|A|} \quad (5.2)$$

A precisão é o número de documentos recuperados e considerados relevantes sobre o número de documentos recuperados, ou seja, a proporção de documentos recuperados que são relevantes. Por outro lado, a cobertura é o número de documentos recuperados e considerados relevantes sobre o número total de documentos relevantes, ou seja, proporção de todos os documentos relevantes que são recuperados.

As medidas precisão e cobertura são inversamente proporcionais, ou seja, quando há um incremento em uma delas, o contrário se sucede com a outra (DOSSANTOS, 2005). E também, quando um sistema tem a precisão máxima, pode ocorrer que um documento é retornado apenas e o qual seja relevante, no entanto, a cobertura pode ser muito baixa. No caso da cobertura máxima, pode ocorrer que o sistema devolva todos os documentos para quaisquer consultas e a precisão seja desconsiderada. A fim de evitar problemas dessa natureza, para fins de avaliação será utilizada a média harmônica entre as duas medidas, denominada medida F (RIJSBERGEN, 1979):

$$F = \frac{1}{\frac{\alpha}{Precisao} + \frac{1-\alpha}{Cobertura}} \quad (5.3)$$

sendo definido  $\alpha = 0,5$ , para ter a mesma importância relativa de precisão e cobertura, e  $0 < \alpha < 1$ . Tem-se que a média harmônica é igual:



$$F = \frac{2 \times Cobertura \times Precisao}{Cobertura + Precisao} \quad (5.4)$$

Na avaliação do protótipo foram escolhidas 25 consultas como em (VOORHEES, 2000), tendo como base a quantidade adotada pelo comitê TREC (Text Retrieval Conference), acerca da avaliação de sistemas de recuperação de informação. Em todas as consultas, o protótipo foi configurado com similaridade igual a 0,1 e número máximo de imagens a serem apresentadas igual a 5 e 10. Os resultados de cada consulta foram expressos a partir da Medida F onde a relevância das imagens recuperadas foi aferida por um especialista do domínio da aplicação. Cabe ressaltar que para os experimentos nesta abordagem foram utilizados 360 imagens do IST-Rio. Todas elas tendo sido documentadas por meio de textos em linguagem natural elaborados por pessoas do domínio de aplicação.

#### 5.2.4 METODOLOGIA DE TESTES PARA RECUPERAÇÃO BASEADA EM CONTEÚDO

Na abordagem de recuperação de imagens baseada em conteúdo foi adotado o método de validação denominado "leave-one-out" (GOLDSCHMIDT, 2005). Este método consiste em dividir um conjunto de dados com N elementos em N subconjuntos unitários (um elemento por conjunto). O processo de avaliação se repete N vezes da seguinte forma: cada subconjunto é considerado como teste e os N-1 restantes considerados para construção de um modelo de conhecimento. Em seguida, o elemento no subconjunto de teste é submetido ao modelo de conhecimento e o resultado é avaliado.

Neste trabalho, foram utilizadas trinta e nove imagens de objetos do IST-Rio tais como mesas, cadeiras, data-shows, computadores e lousa eletrônica. O processo de avaliação consistiu em apresentar cada uma destas imagens ao algoritmo K-NN(K-vizinhos mais próximos) e verificar se os K vizinhos apresentados como resposta eram imagens do mesmo tipo(classe) da imagem apresentada. Os experimentos foram realizados com K = 1, 3 e 5, conforme diversos trabalhos relacionados(OLIVEIRA, 2001). A figura 5.6 ilustra este processo para K=1. As imagens utilizadas encontram-se no anexo A, na seção 8.1.

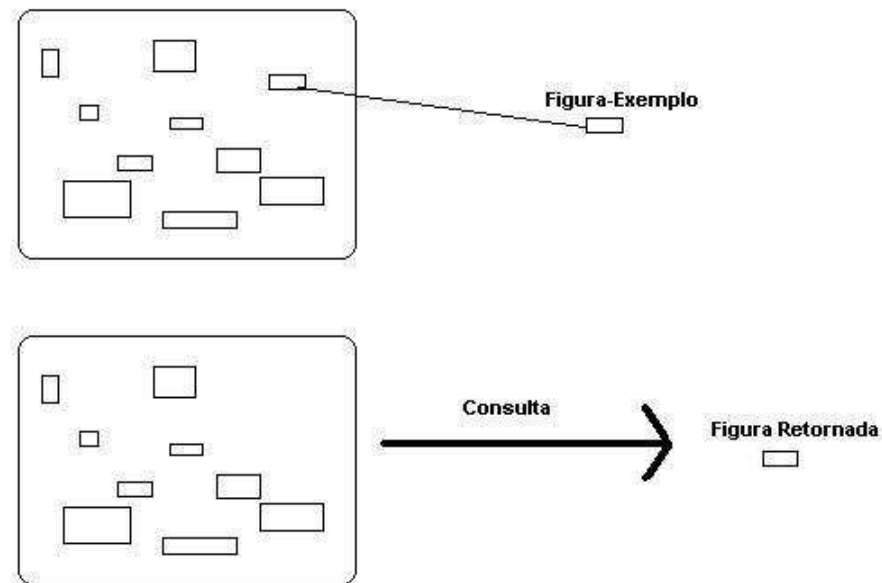


FIG. 5.6: K-NN para  $K = 1$

### 5.2.5 METODOLOGIA DE TESTES PARA RECUPERAÇÃO BASEADA EM SENTENÇAS E RECUPERAÇÃO BASEADA EM CONTEÚDO

Nessa abordagem foi definido a precisão 5.2.3, a maneira encontrada para fazer essa avaliação. Foram utilizadas por volta de 400 imagens, a combinação das imagens utilizadas nas duas abordagens. Infelizmente, só foi possível testar dez consultas, pelo fato, de precisar casar a sentença e a imagem-exemplo das consultas. Caso contrário, a recuperação seria vazia.

## 5.3 ANÁLISE DOS RESULTADOS

Esta seção contém os resultados obtidos em cada abordagem além dos comentários e análises sobre cada caso.

### 5.3.1 AVALIAÇÃO DA RECUPERAÇÃO BASEADA EM SENTENÇAS

Na avaliação de abordagem da recuperação de imagens baseada em sentenças foi utilizada a medida F. Os experimentos consideraram 5 e 10 como limites de imagens recuperadas na avaliação da precisão. A escolha de tais limites buscaram reproduzir o comportamento da maioria das ferramentas de busca que, em geral, não processam todos os resultados

Consulta	Texto Completo(P5)	Palavras-Chave(P5)	Ambos(P5)
1	0.6	0	1
2	0.4	1	1
3	0	0.2	0.2
4	0	0	0
5	0.6	0.2	0.8
6	1	0.8	0.8
7	0.4	0.2	0.4
8	0.4	1	1
9	0.2	0.6	0.4
10	0.6	0	0
11	0.6	1	0.8
12	1	1	1
13	0.6	1	1
14	0	0	0
15	0.8	0	0.8
16	0	1	1
17	0.2	0	0
18	0.8	0.4	0.2
19	0.2	0.8	0.6
20	0.8	0	0.2
21	0.2	0	0.2
22	0.4	0	0
23	0.2	0.4	0.8
24	1	1	1
25	0	0	0
26	1	0.4	0.2
Média	0.46	0.42	0.52

TAB. 5.2: Precisão para as Primeiras 5 Figuras Recuperadas

da consulta, normalmente apresentando os resultados mais relevantes na primeira página. Assim, para cada tipo de índice (texto completo, palavras-chave e ambos), foi calculada a Precisão, a Cobertura e a Medida F para as 5 e 10 primeiras imagens retornadas. Os resultados encontram-se nas tabelas 5.2, 5.3, 5.4, 5.5, 5.6 e 5.7.

Analisando as Tabelas 5.2 e 5.3 pode-se perceber que a precisão média das respostas às consultas aumenta quando as consultas em linguagem natural são submetidas em conjunto com palavras-chave. Tal fato parece estar associado à maior especificidade decorrente da combinação das consultas com palavras-chave específicas.

A cobertura média para as consultas em linguagem natural apresentou-se superior às consultas por palavras-chave e às consultas que combinam as duas alternativas. Acredita-

Consulta	Texto Completo(P10)	Palavras-Chave(P10)	Ambos(P10)
1	0.3	0	1
2	0.2	1	1
3	0	0	0.1
4	0	0.4	0.3
5	0.4	0.1	0.6
6	0.7	0.5	0.6
7	0.5	0.1	0.6
8	0.3	1	1
9	0.2	0.3	0.3
10	0.6	0	0
11	0.6	1	0.8
12	0.7	1	1
13	0.8	1	1
14	0	0	0
15	0.7	0	0.5
16	0	0.9	0.8
17	0.7	0	0
18	0.7	0.7	0.2
19	0.1	0.9	0.9
20	0.5	0	0.1
21	0.1	0	0.1
22	0.3	0	0.2
23	0.2	0.3	0.6
24	0.7	1	1
25	0.1	0	0
26	0.6	0.6	0.6
Média	0.38	0.42	0.51

TAB. 5.3: Precisão para as Primeiras 10 Figuras Recuperadas

Consulta	Texto Completo(C5)	Palavras-Chave(C5)	Ambos(C5)
1	0.09	X	0.15
2	0.67	0.18	0.18
3	0	0	0.5
4	0	0	0
5	0.3	1	0.11
6	0.34	0.27	0.27
7	0.29	0.09	0.18
8	0.11	0.16	0.13
9	0.1	0.6	0.4
10	0.06	0	0
11	0.09	0.09	0.08
12	0.23	0.16	0.16
13	0.1	0.07	0.07
14	0	X	0
15	0.06	X	0.27
16	0	0.29	0.29
17	0.14	0	0
18	0.25	0.04	0.02
19	0.11	0.11	0.09
20	0.5	0	0.13
21	1	X	1
22	0.22	X	0
23	0.04	0.12	0.08
24	0.5	0.42	0.42
25	0	0	0
26	0.45	0.18	0.18
Média	0.22	0.18	0.18

TAB. 5.4: Cobertura para as Primeiras 5 Figuras Recuperadas

Consulta	Texto Completo(C10)	Palavras-Chave(C10)	Ambos(C10)
1	0.09	X	0.30
2	0.67	0.36	0.36
3	0	0	0.5
4	0	0.31	0.23
5	0.4	1	0.16
6	0.47	0.33	0.4
7	0.71	0.09	0.09
8	0.17	0.31	0.26
9	0.2	0.6	0.6
10	0.11	0	0
11	0.19	0.19	0.13
12	0.32	0.31	0.31
13	0.26	0.14	0.14
14	0	X	0
15	0.1	X	0.34
16	0	0.53	0.47
17	1	0	0
18	0.44	0.15	0.04
19	0.11	0.26	0.26
20	0.63	0	0.13
21	1	X	1
22	0.33	X	0.22
23	0.08	0.18	0.24
24	0.7	0.83	0.83
25	0.83	0	0
26	0.55	0.55	0.55
Média	0.33	0.29	0.29

TAB. 5.5: Cobertura para as Primeiras 10 Figuras Recuperadas

Consulta	Texto Completo(F5)	Palavras-Chave(F5)	Ambos(F5)
1	0.15	X	0.26
2	0.5	0.3	0.3
3	X	X	0.28
4	X	X	X
5	0.4	0.33	0.19
6	0.5	0.4	0.4
7	0.2	0.12	0.25
8	0.17	0.27	0.23
9	0.13	0.6	0.4
10	0.1	X	X
11	0.16	0.17	0.14
12	0.37	0.27	0.27
13	0.17	0.13	0.13
14	X	X	X
15	0.11	X	0.4
16	X	0.45	0.45
17	0.17	X	X
18	0.38	0.7	0.04
19	0.14	0.2	0.15
20	0.62	X	0.15
21	0.33	X	0.33
22	0.29	X	X
23	0.67	0.28	0.14
24	0.67	0.59	0.59
25	0	X	X
26	0.62	0.25	0.19
Média	0.31	0.34	0.26

TAB. 5.6: Medida F para as Primeiras 5 Figuras Recuperadas

Consulta	Texto Completo(F10)	Palavras-Chave(F10)	Ambos(F10)
1	0.14	X	0.47
2	0.31	0.53	0.53
3	X	X	0.17
4	X	0.35	0.25
5	0.4	0.18	0.26
6	0.56	0.4	0.48
7	0.5	0.09	0.16
8	0.21	0.47	0.41
9	0.2	0.4	0.4
10	0.19	X	X
11	0.28	0.32	0.22
12	0.44	0.48	0.48
13	0.39	0.25	0.25
14	X	X	X
15	0.17	X	0.4
16	X	0.67	0.59
17	0.82	X	X
18	0.54	0.24	0.07
19	0.1	0.4	0.4
20	0.55	X	0.11
21	0.18	X	0.18
22	0.31	X	0.24
23	0.11	0.22	0.31
24	0.7	0.9	0.91
25	0.09	X	X
26	0.57	0.57	0.57
Média	0.35	0.40	0.36

TAB. 5.7: Medida F para as Primeiras 10 Figuras Recuperadas



se que a diversidade de termos nas consultas textuais tenha propiciado uma busca mais ampla na base de imagens.

Cabe mencionar que embora a precisão das consultas em linguagem natural tenha reduzido quando foi aumentado o número de imagens recuperadas de 5 para 10, a cobertura destas consultas aumentou quando foram consideradas 10 imagens. Tal fato parece ser justificado pois aumentando o número de imagens recuperadas, aumenta o número de oportunidades para imagens relevantes na base de imagens que devam ser recuperadas.

Tomando por base os valores das medidas F, que estão postados nas tabelas 5.6 e 5.7, observa-se que a recuperação que retorna 10 figuras tem um melhor desempenho.

Para cada tipo de índice (texto completo, palavras-chave e ambos), pode ser observado que o melhor resultado foi obtido pelo índice de "palavras-chaves". Contudo este possui o maior número de consultas inválidas (assinaladas com "X"), ou seja, as que não retornaram nenhuma figura relevante na consulta. Os outros tipos de índice apresentam desempenho similar e o número de consultas inválidas.

Contudo, cabe ressaltar que, no geral, os resultados estão bons para a aplicação, pois a medida F chegou próximo a 35%, o que, analisando a literatura, pode ser considerado um resultado bem razoável.

Houve algumas consultas que geraram resultados surpreendentes. Casos em que era esperado um resultado e não retornaram nenhuma imagem. Após um exame detalhado dessas consultas descobriu-se que não era problema de recuperação, e sim humano, pois houve alguns erros de grafia, como "meio-ambiente" e "data-show" nas palavras-chave e descrições textuais.

### 5.3.2 AVALIAÇÃO DA RECUPERAÇÃO BASEADA EM CONTEÚDO

Nessa abordagem foi aplicado o método de K-NN(GOLDSCHMIDT, 2005), para os valores de k igual a 1, 3 e 5. Os descritores usados para representar as figuras foram: Energia, Contraste, Entropia, Homogeneidade, Distorção, Inverso da Contraste, Média, Desvio Padrão, Curtose e Obliquidade. O conjunto de testes para essa abordagem foi diferente da anterior, pois para essa avaliação é preciso classificar as imagens. Assim, decidiu-se por classificar e separar as fotos por classes de objetos(cadeira, mesa, datashow, quadro, monitor, estabilizador e máquina de bingo). Os desempenhos obtidos pelos testes estão demonstrados na tabela 5.8 logo abaixo:

Sub-Imagem	Erro(k=1)	Erro(k=3)	Erro(k=5)
quadro1.bmp	1	1	1
quadro2.bmp	1	1	1
quadro3.bmp	1	1	1
quadro4.bmp	1	1	1
mesa1.bmp	1	1	1
mesa2.bmp	1	1	1
mesa3.bmp	0	1	1
mesa4.bmp	0	1	1
mesa5.bmp	1	1	1
cadeiraazul1.bmp	1	1	1
cadeiraazul2.bmp	1	1	1
cadeiraazul3.bmp	1	1	1
cadeiraazul4.bmp	1	1	1
cadeiraazul5.bmp	1	1	1
cadeirapreta1.bmp	1	1	1
cadeirapreta2.bmp	1	1	1
cadeirapreta3.bmp	1	1	1
cadeirapreta4.bmp	1	1	1
cadeirapreta5.bmp	1	1	1
datashow1.bmp	1	1	1
datashow2.bmp	1	1	1
datashow3.bmp	1	1	1
datashow4.bmp	1	1	1
datashow5.bmp	1	1	1
maquina1.bmp	1	1	1
maquina2.bmp	1	1	1
maquina3.bmp	1	1	1
maquina4.bmp	1	1	1
maquina5.bmp	1	1	1
monitor1.bmp	1	0	0
monitor2.bmp	0	1	1
monitor3.bmp	0	0	0
monitor4.bmp	0	0	0
monitor5.bmp	1	1	1
monitorlcd1.bmp	1	1	1
monitorlcd2.bmp	1	1	1
monitorlcd3.bmp	0	1	1
monitorlcd4.bmp	0	1	1
nobreak1.bmp	1	1	1
nobreak2.bmp	1	1	1
nobreak3.bmp	1	1	1
nobreak4.bmp	1	1	1
nobreak5.bmp	1	1	1
Taxa de Erro	0.84	0.94	0.94

TAB. 5.8: Taxa de Erro para as Imagens Classificadas

Conforme pode ser observado, os resultados não foram satisfatórios, pois as taxas de erro foram muito altas, indicando que os descritores não funcionaram bem na discriminação das imagens. Os descritores utilizados não foram bem com seus valores reais. Portanto, decidiu-se normalizar estes valores. O método de normalização adotado foi pelo valor máximo de cada descritor. Com essa nova representação, foram obtidos os resultados, ilustrados na Tabela 5.9.

Conclui-se que com a normalização, os resultados para  $k = 1, 3$  e  $5$ , foram melhorados. Cabe ressaltar ainda que os melhores resultados foram obtidos para  $k = 1$ . Mesmo nesses casos, o desempenho da abordagem de recuperação baseada em conteúdo não foi satisfatório, indicando necessidade de melhorias futuras.

### 5.3.3 AVALIAÇÃO DA RECUPERAÇÃO BASEADA EM SENTENÇAS E EM CONTEÚDO

Na avaliação da abordagem de recuperação de imagens baseada em sentenças e em conteúdo foram consideradas as consultas e as imagens das tabelas 8.1 e 5.11.

A tabela 5.12 abaixo apresenta os resultados obtidos:

Os resultados dessa abordagem foram muito abaixo do esperado, com alguns resultados vazios (assinalados com "X") e 0% de precisão, ou seja, sendo mais da metade dos resultados. Esses resultados foram muito afetados pela abordagem da recuperação baseada em Conteúdo, que também não tiveram uma avaliação boa.

Sub-Imagem	Erro(k=1)	Erro(k=3)	Erro(k=5)
quadro1.bmp	1	1	1
quadro2.bmp	1	1	1
quadro3.bmp	1	1	1
quadro4.bmp	1	1	1
mesa1.bmp	0	0	1
mesa2.bmp	0	0	1
mesa3.bmp	1	1	1
mesa4.bmp	1	1	1
mesa5.bmp	1	1	0
cadeiraazul1.bmp	1	1	1
cadeiraazul2.bmp	0	0	1
cadeiraazul3.bmp	1	1	1
cadeiraazul4.bmp	0	0	0
cadeiraazul5.bmp	0	0	0
cadeirapreta1.bmp	1	1	1
cadeirapreta2.bmp	0	1	1
cadeirapreta3.bmp	0	1	1
cadeirapreta4.bmp	0	1	1
cadeirapreta5.bmp	0	1	1
datashow1.bmp	1	1	1
datashow2.bmp	0	0	0
datashow3.bmp	0	0	0
datashow4.bmp	0	0	0
datashow5.bmp	0	0	0
maquina1.bmp	1	1	1
maquina2.bmp	0	1	1
maquina3.bmp	1	1	1
maquina4.bmp	1	1	1
maquina5.bmp	1	1	1
monitor1.bmp	1	1	1
monitor2.bmp	1	1	1
monitor3.bmp	0	0	0
monitor4.bmp	0	1	1
monitor5.bmp	0	0	0
monitorlcd1.bmp	1	1	1
monitorlcd2.bmp	1	1	1
monitorlcd3.bmp	0	1	1
monitorlcd4.bmp	0	1	1
nobreak1.bmp	1	1	1
nobreak2.bmp	0	0	0
nobreak3.bmp	0	0	1
nobreak4.bmp	0	0	0
nobreak5.bmp	0	0	0
Taxa de Erro	0.47	0.65	0.72

TAB. 5.9: Taxa de Erro para as Imagens Classificadas

	Consulta
1	Recuperar fotos de ações do IST-Rio em projetos de inclusão digital.
2	Apresentar imagens que exemplifiquem alguns dos investimentos do Estado do Rio de Janeiro em infra-estrutura e educação.
3	Mostrar eventos relacionados ao Dicionário Técnico de Libras ou ao Projeto de Transformação de Máquinas em Computadores.
4	Apresentar imagens de máquinas caça-níqueis transformadas em computadores.
5	Listar fotos com recursos de hardware pertencente ao Instituto Superior de Tecnologia em Ciência da Computação do Rio de Janeiro.
6	Obter fotos do novo espaço de identidade e pesquisa do ISTCC-RJ.
7	Obter imagens de encontros marcados com o prof. Emmanuel Passos.
8	Recuperar imagens do evento de inauguração do novo espaço de identidade e pesquisa do instituto.
9	Apresentar fotos que destaquem recursos físicos do IST-Rio tais como equipamentos e mobiliários.
10	Selecionar imagens de eventos promovidos com o apoio ou com a participação da Biblioteca do IST-Rio.

TAB. 5.10: Consultas Seleccionadas para essa Recuperação

Consulta	Imagem
1	maquinareco3.bmp
2	quadro1reco.bmp
3	maquinareco5.bmp
4	maquinareco5.bmp
5	datashowreco5.bmp
6	datashowreco3.bmp
7	mesareco4.bmp
8	mesareco3.bmp
9	cadeirapretareco2.bmp
10	cadeirapretareco3.bmp

TAB. 5.11: Imagens Associadas a Consulta

Consulta	Imagem	Precisão(40)
1	maquinareco3.bmp	X
2	quadro1reco.bmp	X
3	maquinareco5.bmp	0.42
4	maquinareco5.bmp	0.42
5	datashowreco5.bmp	0.19
6	datashowreco3.bmp	X
7	mesareco4.bmp	0
8	mesareco3.bmp	X
9	cadeirapretareco2.bmp	0.23
10	cadeirapretareco3.bmp	0

TAB. 5.12: Precisão para 40 Imagens Recuperadas

## 6 CONSIDERAÇÕES FINAIS

### 6.1 RETROSPECTO

Os avanços na área de Tecnologia de Informação têm propiciado um aumento nos volumes de informações armazenadas. Estas informações vêm sendo apresentadas em diferentes mídias e formatos, como por exemplo, texto, imagens, vídeos, hipertextos, etc. Esse contexto tem fomentado o desenvolvimento da área de pesquisa com a "Recuperação de Informação (RI)", tal área vem despertando um interesse cada vez maior da indústria e academia (TORRES, 2006).

O foco principal desse presente trabalho foi trabalhar com imagens; estudando maneiras de recuperá-las de forma eficiente e eficaz. Sendo a aplicação muito importante para diversas áreas, como Biologia, História, Jornalismo, Medicina, etc.

Em síntese, dois tipos de abordagens têm sido utilizadas na recuperação de imagens:

- Recuperação de imagens baseada em Sentenças;
- Recuperação de imagens baseada em Conteúdo.

Essas abordagens se encontram na literatura de maneira isolada, ou seja, sem nenhuma tentativa de combinação entre elas. Essa tentativa de combinação pode ser grande utilidade como é mencionado em (TORRES, 2006).

Assim sendo, esta dissertação teve como objetivos: pesquisar, propor, desenvolver e avaliar um ambiente computacional inteligente que pudesse ser utilizado em processo envolvendo recuperação de imagens. Esse ambiente permite a recuperação de imagens em três cenários: (a) consultas parcialmente formuladas em linguagem natural, orientadas a partir da construção de filtros; (b) recuperação de imagens a partir de exemplos; (c) combinação dos cenários anteriores.

A abordagem de recuperação de imagens baseada em sentença foi desenvolvida de forma a ampliar consultas formuladas em linguagem natural a partir de dados disponíveis em ontologias de domínio.

Por outro lado, a abordagem de recuperação de imagens em conteúdo foi desenvolvida com base no conceito de similaridade calculado a partir de descritores de cor e textura extraídos das imagens armazenadas.

A abordagem de recuperação de imagens baseada em sentença e em conteúdo foi concebida de forma a combinar os resultados obtidos a partir das duas abordagens mencionadas por meio da operação de interseção entre os conjuntos de imagens resultantes das referidas abordagens.

O protótipo do ambiente de recuperação de imagens foi desenvolvido em linguagem Java, com o auxílio de vários plugins e pacotes tais como Jigloo e JNI(Java Native Interface). Cerca 22000 linhas de código foram codificadas.

Como estudo de caso e avaliação do ambiente de recuperação de imagens desenvolvido, este ambiente foi configurado de forma a tratar a base de imagens do IST-Rio, instituição de ensino pertencente à rede FAETEC com grande base de imagens e sem mecanismos que facilitassem a recuperação destas imagens.

Na avaliação do protótipo desenvolvido, foram utilizadas cerca de 360 imagens e seus respectivos textos descritivos. Para a abordagem de recuperação de imagens baseada em sentenças, uma ontologia de domínio foi integralmente elaborada. Para abordagem de recuperação de imagens baseada em conteúdo, foram implementadas 10 descritores de imagens, dentre os quais estão Energia, Contraste, Entropia, Homogeneidade, Inverso do Contraste, Distorção, Média, Desvio Padrão, Obliquidade e Curtose.

## 6.2 CONTRIBUIÇÕES OBTIDAS

Entre as principais contribuições proporcionadas por esta dissertação, podem ser destacadas:

- Uma proposta de integração das tradicionais abordagens de recuperação de imagens baseada em sentença e em conteúdo;
- A criação de um ambiente voltado a recuperação de imagens, que integra mas também executa de forma individual as abordagens de recuperação de imagens baseada em sentença e em conteúdo;
- O desenvolvimento de uma ontologia 8.2 para representação dos termos de domínio do IST-Rio. Cabe ressaltar que o processo de elaboração da ontologia proporcionou ao próprio instituto uma reflexão quanto à terminologia usualmente adotada em suas práticas cotidianas;

- Uma documentação descritiva de todas as imagens armazenadas na base de imagens do IST-Rio;
- A disponibilização no IST-Rio do ambiente de recuperação de imagens implementado. Tal ambiente tem sido utilizado na prática pelos membros do instituto em suas buscas por imagens de interesse.

### 6.3 TRABALHOS FUTUROS

Baseado nos desempenhos da implementação da abordagem de recuperação de imagens baseada em conteúdo, sugere-se como trabalho futuro a incorporação de novos descritores de cor, de textura e de forma. Recomenda-se uma especial atenção à implementação de descritores de forma, visto que nos experimentos realizados não foi envolvido nenhum descritor dessa natureza.

Um outro trabalho recomendado refere-se à incorporação de novas imagens na aplicação do IST-Rio. Recomenda-se uma ampliação da base de imagens a fim de investigar a influência da inclusão de novas imagens nas abordagens de recuperação implementadas.

A aplicação do ambiente desenvolvido em outro domínio de aplicação também se mostra como um trabalho bastante oportuno.

A fim de aprimorar os resultados da abordagem da recuperação de imagens baseada em sentenças, recomenda-se realizar mais testes nesta abordagem permitindo o retorno de 15, 20 e 30 imagens. Cabe ressaltar que estes são os valores usualmente utilizados em testes desse tipo de abordagem. Tais testes buscariam avaliar se em uma recuperação mais abrangente o resultado poderia ser melhorado.

Um outro trabalho futuro compreende testes/avaliação com outros operadores de conjunto (além da interseção).



## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- AKSOY, S. e HARALICK, R. Textural features for image database retrieval. Em *IEEE Workshop on Content Based Access of Image and Video Libraries*, Washington, USA, 1998.
- BALLARD, D. H. e BROWN, C. M. *Computer Vision*. **Prentice Hall**, New Jersey, USA, 1982.
- BENDER, T. Classificação e recuperação de imagens por cor utilizando técnicas de inteligência artificial. Dissertação de Mestrado, **UNISINOS**, São Leopoldo, RS, 2003.
- BIMBO, A. D. *Visual Information Retrieval*. **Morgan Kaufmann Publishers**, 1999.
- BREITMAN, K. *Web Semântica: a Internet do Futuro*. **LTC**, 2005.
- BROWN, M. G., FOOTE, J. T., JONES, G. J. F., JONES, K. S. e YOUNG, S. J. Automatic content-based retrieval of broadcast news. Em *ACM Multimedia 95*, San Francisco, USA, 1995.
- DAML+OIL. Daml+oil reference description, 2001. URL <http://www.w3.org/TR/daml+oil-reference>.
- DEERWESTER, S., DUMAIS, S., LANDAUER, T., FURNAS, G. e HARSHMAN, R. Indexing by latent semantic analysis. *American Society of Information Science*, págs. 391–407, 1999.
- DOS SANTOS, W. B. Recuperação de informação com auxílio de extratos automáticos. Dissertação de Mestrado, **UFSCAR**, São Carlos, SP, 2005.
- EAKINS, J. e GRAHAM, M. Content-based image retrieval. Dissertação de Mestrado, **University of Northumbria**, Newcastle, UK, 1999.
- FELIPE, J. C. e TRAINA, A. J. M. Utilizando características de textura para identificação de tecidos em imagens médicas. Dissertação de Mestrado, **Universidade de São Paulo**, São Paulo, Brasil, 2002.
- GOLDSCHMIDT, R. R. Texture based classifiers applied to a cultural heritage problem. *SPPRA 2001 - Signal Processing, Pattern Recognition and Applications*, 2001, págs. 58–61, 2001.
- GOLDSCHMIDT, R. R. e PASSOS, E. *Data Mining - UM GUIA PRÁTICO*. **Editora Campus**, 2005.
- GONZALEZ, R. e WOODS, R. *Digital Image Processing*. **Prentice Hall**, 2001.

- GRUBER, T. What is an ontology?, 1993. URL <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>.
- GUALBERTO, L. P. S. e FILHO, P. W. M. S. Reconhecimento de ambientes costeiros a partir de análise textural de imagens sar r99b (banda l - polarização vv). Em *XIII Simpósio Brasileiro de Senroriamento Remoto*, Florianópolis, SC, 2007.
- GUARINO, N., MASOLO, C. e VETERE, G. Ontoseek: Content-based access to the web. *IEEE INTELLIGENT SYSTEMS* 1094-7167, 1999.
- HAASE, K. Framer: A portable persistent representation library. Em *AAAI Workshop on AI in Systems and Support*, 1993.
- HARALICK, R. M., SHANMUGAN, K. e DINSTEN, I. Textural features for image classification. *IEEE Transaction on Systems, Man and Cybernetics*, 1973.
- HARMAN, D. Ranking algorithms. *Information Retrieval*, 1992.
- JACKSON, P. e MOULINIER, I. *Natural Language Processing for Online Applications: Text Retrieval, Extraction, and Categorization*. **John Benjamins Publishing Co**, 2002.
- JENA. Jena - a semantic web framework for java, 2002. URL <http://jena.sourceforge.net/>.
- KULKARNI, A. D. *Artificial Neural Networks for Image Understanding*. **John Wiley & Sons**, 1994.
- LIN, D. An information-theoretic definition of similarity. Em *15th International Conference on Machine Learning*, 1998.
- MAULDIN, M. *Conceptual Information Retrieval A case study in adaptive partial parsing*. **Kluwer Academic Publishers**, 1991.
- MORENO, R. A. e FURUIE, S. S. Biram: Sistema de recuperação de imagens por conteúdo. *Revista SBC, Anais do X Congresso Brasileiro de Informática em Saúde*, 2006.
- OLIVEIRA, C. J. S. Classificação de imagens coletadas na web. Dissertação de Mestrado, **UFMG**, 2001.
- ORENGO, V. e HUYCK, C. A stemming algorithm for the portuguese language. Em *8th International Symposium on String Processing and Information Retrieval(SPIRE)*, 2001.
- OWL. Owl web ontology language semantics and abstract syntax, 2004. URL <http://www.w3.org/TR/owl-absyn/>.
- PAZ-TRILLO, C. e WASSERMAN, R. Recuperação de vídeos indexados por conceito. Dissertação de Mestrado, **Universidade de São Paulo**, 2001.

- PENTLAND, A., PICARD, R. W. e SCLARO, S. Photobook: Content based manipulation of image databases. *Perceptual Computing Section, The Media Laboratory, Massachusetts Institute of Technology*, 1995.
- PHOTO-RDF. Describing and retrieving photos using rdf and http, 2002. URL <http://www.w3.org/TR/photo-rdf/>.
- PROTÉGÉ. Protégé 2000 documentation, 2002. URL <http://protege.stanford.edu/publications/UserGuide.pdf>.
- RDFS. Rdf vocabulary description language 1.0: Rdf schema, 2004. URL <http://www.w3.org/TR/rdf-schema/>.
- RIJSBERGEN, C. J. V. *Information Retrieval*. **Butterworths**, 1979.
- SALTON, G. e MCGILL, M. J. *Introduction to modern information retrieval*. **McGraw-Hill**, 1983.
- SALTON, G., WONG, A. e YANG, C. S. A vector space model for automatic indexing. *Commun.ACM*, 1975.
- SCHWARTZ, W. R. Segmentação de imagens baseada em dependência espacial utilizando o campo aleatório de markov associado com características de texturas. Dissertação de Mestrado, **UFPR**, 2005.
- SCHWARTZ, W. R. e PEDRINI, H. Método para classificação de imagens baseada em matrizes de co-ocorrências utilizando características de textura. *Colóquio Brasileiro de Pesquisas Geodésicas*, 2003.
- SMITH, J. R. e CHANG, S. Tools and techniques for color image retrieval. Dissertação de Mestrado, **Columbia University**, 1996.
- SOUZA, M. I. F., VENDRUSCULO, L. G. e MELO, G. C. Metadados para descrição de recursos de informação eletrônica: utilização do padrão dublin core. *Ci. Inf.*, 2000.
- STEIN, T. Avaliação de descritores de textura para segmentação de imagens. Dissertação de Mestrado, **UFPR**, 2005.
- TORRES, R. Recuperação de vídeos indexados por conteúdo. *Computação Brasil*, 2007.
- TORRES, R. e FALCÃO, A. Content-based image retrieval. *Theory and Applications*, 2006.
- VOORHEES, E. M. Special issue: The sixth text retrieval conference (trec-6). *Information Processing and Management*, 2000.
- XML. Extensible markup language (xml) 1.0 (fourth edition), 2006. URL <http://www.w3.org/TR/2006/REC-xml-20060816/>.
- ZHANG, Y. On the use of cbir in image mosaic generation. Dissertação de Mestrado, **University of Alberta**, 2002.

ZHONG, Y. e JAIN, A. K. Object localization using color, texture and shape pattern recognition. Dissertação de Mestrado, **Michigan State University**, 1999.

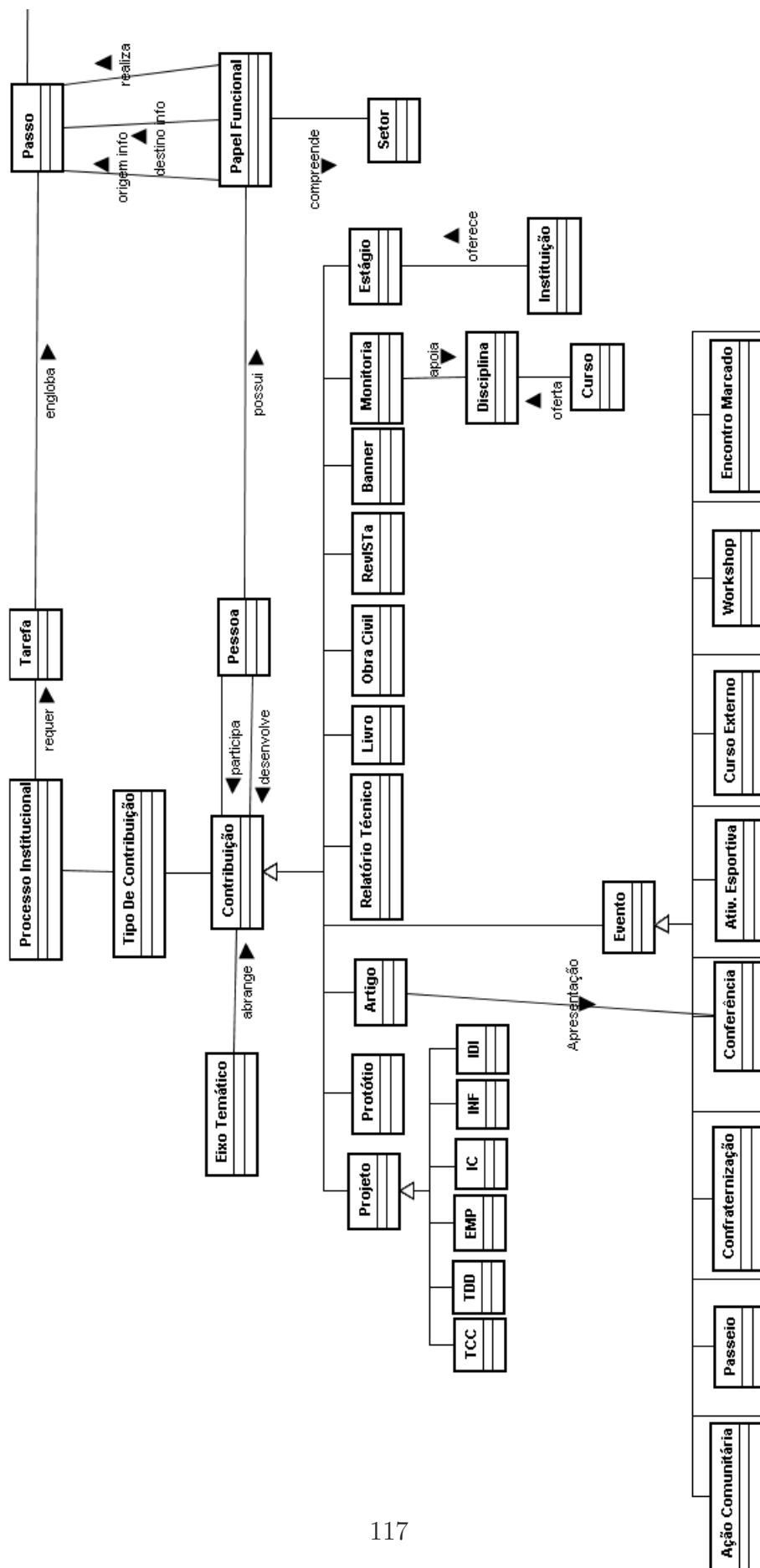
## 8 ANEXOS

### 8.1 ANEXO A

### 8.2 ANEXO B

	Consulta
1	Obter imagens que ilustrem a atuação do IST-Rio em ações sociais.
2	Recuperar fotos de ações do IST-Rio em projetos de inclusão digital.
3	Recuperar fotos de ações do IST-Rio em projetos de inclusão social.
4	Ilustrar o compromisso do IST-Rio com atividades de pesquisa aplicada.
5	Apresentar imagens que exemplifiquem alguns dos investimentos do Estado do Rio de Janeiro em infra-estrutura e educação.
6	Mostrar eventos relacionados ao Dicionário Técnico de Libras ou ao Projeto de Transformação de Máquinas em Computadores.
7	Apresentar imagens de máquinas caça-níqueis transformadas em computadores.
8	Recuperar imagens relacionadas à produção institucional do IST-Rio.
9	Listar fotos com recursos de hardware pertencente ao Instituto Superior de Tecnologia em Ciência da Computação do Rio de Janeiro.
10	Mostrar participações da equipe do IST-Rio em eventos externos.
11	Obter fotos do novo espaço de identidade e pesquisa do ISTCC-RJ.
12	Selecionar fotos com membros da equipe do IST-Rio atuando no novo prédio.
13	Ilustrar o espírito festivo do IST-Rio por meio de imagens de confraternizações de sua comunidade interna.
14	Recuperar imagens da semana tecnológica e cultural cuja temática foi saúde e tecnologia.
15	Obter imagens de encontros marcados com o prof. Emmanuel Passos.
16	Selecionar fotos relacionadas a lançamentos de livros produzidos por membros da equipe do IST-Rio.
17	Comprovar a realização de eventos de cunho artístico e cultural realizados no espaço de identidade e pesquisa do instituto.
18	Comprovar ações de integração envolvendo ensino, pesquisa e extensão no IST-Rio.
19	Recuperar imagens associadas aos projetos de infra-estrutura referentes à construção e às obras do novo espaço de identidade e pesquisa do IST-Rio.
20	Apresentar imagens de eventos envolvendo a divulgação do livro Escola Mandala que descreve o projeto pedagógico institucional do IST-Rio.
21	Obter fotos que ilustrem ações do IST-Rio na preservação do meio-ambiente.
22	Fornecer imagens de visitas técnicas realizadas com a participação de discentes e de docentes do IST-Rio.
23	Recuperar imagens do evento de inauguração do novo espaço de identidade e pesquisa do instituto.
24	Recuperar fotos de workshops e encontros marcados realizados no IST-Rio.
25	Apresentar fotos que destaquem recursos físicos do IST-Rio tais como equipamentos e mobiliários.
26	Selecionar imagens de eventos promovidos com o apoio ou com a participação da Biblioteca do IST-Rio.

TAB. 8.1: Consultas Selecionadas para essa Recuperação



# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)



[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)