

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



**RECUPERAÇÃO DE IMAGENS DIGITAIS COM BASE NA
DISTRIBUIÇÃO DE CARACTERÍSTICAS DE BAIXO NÍVEL
EM PARTIÇÕES DO DOMÍNIO UTILIZANDO ÍNDICE
INVERTIDO**

PATRÍCIA APARECIDA PROENÇA

Uberlândia - Minas Gerais

2010

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



PATRÍCIA APARECIDA PROENÇA
Orientador: PROF. DR. ILMÉRIO REIS DA SILVA
Co-orientadora: PROF. DRA. CELIA A. ZORZO BARCELOS

**RECUPERAÇÃO DE IMAGENS DIGITAIS COM BASE NA
DISTRIBUIÇÃO DE CARACTERÍSTICAS DE BAIXO NÍVEL
EM PARTIÇÕES DO DOMÍNIO UTILIZANDO ÍNDICE
INVERTIDO**

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

Área de concentração: Banco de Dados.

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

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da dissertação intitulada “**Recuperação de Imagens Digitais com Base na Distribuição de Características de Baixo Nível em Partições do Domínio Utilizando Índice Invertido**” por **Patrícia Aparecida Proença** como parte dos requisitos exigidos para a obtenção do título de **Mestre em Ciência da Computação**.

Uberlândia, 29 de Março de 2010

Orientador:

Prof. Dr. Ilmério Reis da Silva
Universidade Federal de Uberlândia/Minas Gerais

Co-orientadora:

Prof. Dra. Celia A. Zorzo Barcelos
Universidade Federal de Uberlândia/Minas Gerais

Banca Examinadora:

Prof. Dr. Luciano Romero Soares de Lima
Hospital Sarah/Belo Horizonte/Minas Gerais

Prof. Dra. Denise Guliato
Universidade Federal de Uberlândia/Minas Gerais

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

Data: Março de 2010

Autor: **Patrícia Aparecida Proença**
Título: **Recuperação de Imagens Digitais com Base na Distribuição de Características de Baixo Nível em Partições do Domínio Utilizando Índice Invertido**
Faculdade: **Faculdade de Computação**
Grau: **Mestrado**

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

Autor

O AUTOR RESERVA PARA SI QUALQUER OUTRO DIREITO DE PUBLICAÇÃO DESTE DOCUMENTO, NÃO PODENDO O MESMO SER IMPRESSO OU REPRODUZIDO, SEJA NA TOTALIDADE OU EM PARTES, SEM A PERMISSÃO ESCRITA DO AUTOR.

Agradecimentos

A Deus que sem dúvidas foi o grande facilitador para esta trajetória, abrindo os caminhos em direção ao sucesso e possibilitando que eu encontrasse as pessoas certas para me auxiliar nesta jornada.

A minha família, pelo amor, carinho e encorajamento que sempre me deram em todos os momentos da minha vida. Agradeço, especialmente, aos meus pais José Márcio e Elza, que sempre dedicaram suas vidas aos filhos. Ao meu irmão Márcio, pelo companheirismo.

Ao meu noivo Sandro, com quem sempre posso contar verdadeiramente, uma pessoa especial que sempre me apoiou e incentivou em várias decisões de minha vida.

Ao meu querido orientador Prof. Dr. Ilmério Reis da Silva, a quem tenho grande estima por sua paciência, amizade, compreensão e pelos preciosos conhecimentos compartilhados.

Professora Célia, pelo apoio e cooperação na realização deste trabalho.

A Carina e Isabela por sofrerem e sorrirem juntas comigo nos momentos tristes e alegres destes anos em Uberlândia.

A todos os colegas e amigos do programa de pós-graduação pela boa vontade em ajudar: Nádia, Fabíola, Rodrigo, Juliana, Jean, Núbia e Eduardo. Em especial, ao amigo Tauller cuja colaboração foi valiosa na realização deste trabalho.

Aos professores e funcionários do Programa de Pós Graduação em Ciência da Computação da Universidade Federal de Uberlândia, sempre prestativos e atenciosos.

Agradeço a Universidade Federal de Uberlândia por colocar a minha disposição a sua estrutura e a FAPEMIG pelo apoio financeiro.

Resumo

O principal objetivo de um sistema de recuperação de imagens é obter imagens de uma coleção que atendam a uma necessidade do usuário. Para atingir esse objetivo, em geral, os sistemas de recuperação de imagens calculam a similaridade entre a necessidade do usuário, representada por uma consulta, e representações das imagens da coleção. Tal objetivo é difícil de ser alcançado devido à subjetividade do conceito de similaridade entre imagens, visto que uma mesma imagem pode ser interpretada de formas diferentes por pessoas distintas. Na tentativa de resolver este problema os sistemas de recuperação de imagens por conteúdo exploram as características de baixo nível cor, forma e textura no cálculo da similaridade entre as imagens. Um problema desta abordagem é que na maioria dos sistemas o cálculo da similaridade é realizado comparando-se a imagem de consulta com todas as imagens da coleção, tornando o processamento difícil e lento. Considerando a indexação de características de baixo nível de partições de imagens digitais mapeadas para um índice invertido, este trabalho busca melhorias no desempenho do processamento de consultas e ganho na precisão considerando o conjunto de imagens recuperadas em grandes bases de dados. Utilizamos uma abordagem baseada em índice invertido, que é aqui adaptada para imagens particionadas. Nesta abordagem o conceito de termo da recuperação textual, principal elemento da indexação, é utilizado no trabalho como característica de partições de imagens para a indexação. Experimentos mostram ganho na qualidade da precisão usando duas coleções de imagens digitais.

Palavras chave: Recuperação de informação, Recuperação de imagens, Índice Invertido, Modelo Vetorial, Peso do Termo, Cosseno, Distância Euclidiana, Partições

Abstract

The main goal of a images retrieval system is to obtain images from a collection that assist a need of the user. To achieve this objective, in generally, the systems of retrieval of images calculate the similarity between the user's need represented by a query and representations of the images of the collection. Such an objective is difficult of being obtain due to the subjectivity of the similarity concept among images, because a same image can be interpreted in different ways by different people. In the attempt of solving this problem the content based image retrieval systems explore the characteristics of low level color, forms and texture in the calculation of the similarity among the images. A problem of this approach is that in most of the systems the calculation of the similarity is accomplished being compared the query image with all of the images of the collection, turning the difficult and slow processing. Considering the indexation of characteristics of low level of partitions of digital images mapped to an inverted index, this work looks for improvements in the acting of the processing of queries and improve in the precision considering the group of images retrieval in great bases of data. We used an approach based in inverted index that is here adapted for partitions images. In this approach the concept of term of the retrieval textual, main element of the indexation, it is used in the work as characteristic of partitions of images for the indexation. Experiments show improvement in the quality of the precision using two collections of digital images.

Keywords: Information Retrieval, Image Retrieval, Inverted Index, Vector Space, Weight of the Term, Cosine, Euclidean Distance, Partitions

Sumário

Lista de Figuras	ix
Lista de Tabelas	xi
Lista de Abreviaturas e Siglas	xiv
1 Introdução	1
1.1 Motivação	1
1.2 Objetivos	2
1.3 Trabalhos Relacionados	3
1.4 Estrutura do Trabalho	6
2 Conceitos Básicos	7
2.1 Representação de Imagens	7
2.1.1 Atributo COR	8
2.1.2 Atributo Textura	13
2.1.3 Atributo Forma	15
2.1.4 Considerações Finais	18
2.2 Recuperação de Imagens	18
2.2.1 Recuperação de Imagens Baseada em Conteúdo Textual	19
2.2.2 Recuperação de Imagens baseada em Conteúdo Visual	21
2.2.3 <i>Bag-of-features</i>	22
2.2.4 Considerações Finais	23
2.3 Recuperação de Informação em Textos	23
2.3.1 Arquivo Invertido	25
2.3.2 Modelo Vetorial	28
2.3.3 Métricas de Avaliação	32
2.3.4 Considerações Finais	35
3 Proposta de um Método para CBIR Baseado em Índice Invertido com Partições do Domínio	36
3.1 Introdução	36

3.2	CBIR-Índice	37
3.2.1	Indexação das Imagens	38
3.2.2	Cálculo da Similaridade	41
3.3	Sistema CBIR-I	41
3.4	CBIR-Índice-Particionado	42
3.4.1	Esquema Geral	43
3.4.2	Particionamento das Imagens	44
3.4.3	Extração de Características	45
3.4.4	Construção do Índice Invertido	50
3.4.5	Similaridade	52
3.4.6	Adaptação dos Sistema CBIR-I e CBIR-Índice	53
3.4.7	Considerações Finais	53
4	Resultados Experimentais	55
4.1	Base de dados	55
4.2	Avaliação Experimental	57
4.2.1	Avaliação Experimental BD-110	58
4.2.2	Avaliação Experimental Corel-1000	66
4.2.3	Avaliação Experimental BD-10000	71
4.3	Comparação do número de operações Aritméticas	75
4.4	Considerações Finais	75
5	Conclusão e Trabalhos Futuros	77
	Referências Bibliográficas	79

Lista de Figuras

2.1	Modelo de cor RGB.	9
2.2	Esquema de cubo do espaço de cor RGB.	10
2.3	Espaço de cor HSV [Gonzales and Woods, 2002].	11
2.4	Exemplos de textura (a) suave, (b) áspera, (c) regular. Fonte: [Gonzales and Woods, 2002].	13
2.5	Exemplos de textura.	14
2.6	a) Imagem em nível de cinza. b) Vizinhaça oito do segundo <i>pixel</i> da segunda linha.	15
2.7	Classificação geral das técnicas de representação de formas [Zhang and Lu, 2004].	16
2.8	Operadores de Sobel	18
2.9	Processo de Extração de Características.	21
2.10	Componentes de um sistema de recuperação de informação [Cardoso, 2000].	24
2.11	Exemplo arquivo invertido.	26
2.12	Indexação dos termos, parte um.	27
2.13	Indexação dos termos, parte dois.	27
2.14	Exemplo consulta utilizando arquivo invertido.	28
2.15	Exemplo modelo vetorial com três termos de indexação.	29
2.16	Exemplo modelo vetorial, vetores dos documentos e das consultas [Santos, 2003].	30
2.17	Algoritmo para calcular o cosseno usando índice invertido.	32
2.18	Algoritmo para calcular o cosseno.	33
3.1	Distribuição adotada para classificação das faixas. Onde $\bar{X}$ é a média e s o desvio padrão.	38
3.2	Vetor de característica correspondente a imagem Africa1.jpg.	39
3.3	Vetor de característica gerado pelo sistema correspondente a uma imagem exemplo.	40
3.4	(a) Banco de dados de Característica (b) Estrutura de busca (vocabulário) e a lista invertida.	40

3.5	Fluxograma básico do sistema CBIR-Índice-Particionado proposto neste trabalho.	43
3.6	Particionamento de uma imagem exemplo I em quatro partições Ω_1 , Ω_2 , Ω_3 e Ω_4	45
3.7	Vetor de características gerado pelo sistema CBIR-Índice correspondente a imagem da figura 3.6.	46
3.8	Vetor de características obtidos de cada partição da imagem da figura 3.6.	47
3.9	Vetor de característica da imagem da figura 3.6 utilizando o CBIR-Índice-Particionado.	48
3.10	Distribuição adotada para classificação das faixas. Onde $\bar{X}$ é a média e s o desvio padrão.	49
3.11	Vetor com as características de uma imagem exemplo utilizando o CBIR-Índice-Particionado. Noventa posições do vetor estão relacionadas à cor, sendo dez para cada momento, oitenta relacionadas a textura e oitenta relacionadas a forma.	51
4.1	Amostra do Banco de Dados BD-10000.	56
4.2	Amostra do Banco de Dados BD-110 e Corel 1000.	56
4.3	Curva precisão x revocação dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando a característica cor.	62
4.4	Curva 11-pt dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando as características cor, forma e textura.	65
4.5	Curva 11-pt dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando a característica cor.	68
4.6	Curva 11-pt dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando as características cor, forma e textura.	70
4.7	Curva <i>precision x recall</i> dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando a característica cor.	72
4.8	Curva <i>precision x recall</i> dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando as características cor, forma e textura.	74

Lista de Tabelas

3.1	Descrição do vetor de característica resultante da extração de característica da imagem usando os três momentos de cor do espaço HSV.	37
3.2	Construção do vetor de Características.	39
3.3	Descrição do vetor de característica resultante da extração de característica da imagem usando histograma de direção de bordas do espaço HSV.	48
3.4	Descrição do vetor de característica resultante da extração de característica da imagem usando vizinhança de textura do espaço HSV.	50
4.1	Precisão média no sistema CBIR-I utilizando a característica cor. Coleção BD-110	59
4.2	Precisão média no sistema CBIR-Índice utilizando a característica cor. Coleção BD-110	59
4.3	Precisão média no sistema CBIR-Índice-Particionado utilizando 2 partições e a característica cor. Coleção BD-110	59
4.4	Precisão média no sistema CBIR-Índice-Particionado utilizando 4 partições e a característica cor. Coleção BD-110	60
4.5	Precisão média no sistema CBIR-Índice-Particionado utilizando 8 partições e a característica cor. Coleção BD-110	60
4.6	Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e a característica cor. Coleção BD-110	60
4.7	Precisão média no sistema CBIR-Índice-Particionado utilizando 32 partições e a característica cor. Coleção BD-110	61
4.8	Comparação do CBIR-Índice-Particionado utilizando partições com tamanho 8 e 16 com os sistemas CBIR-I e CBIR-Índice utilizando a característica cor. Coleção BD-110	61
4.9	Precisão média no sistema CBIR-I utilizando cor, textura e forma. Coleção BD-110	63
4.10	Precisão média no sistema CBIR-Índice utilizando cor, textura e forma. Coleção BD-110	63
4.11	Precisão média no sistema CBIR-Índice-Particionado utilizando 2 partições e as características cor, textura e forma. Coleção BD-110	63

4.12	Precisão média no sistema CBIR-Índice-Particionado utilizando 4 partições e as características cor, textura e forma. Coleção BD-110	64
4.13	Precisão média no sistema CBIR-Índice-Particionado utilizando 8 partições e as características cor, textura e forma. Coleção BD-110	64
4.14	Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e as características cor, textura e forma. Coleção BD-110	64
4.15	Precisão média no sistema CBIR-Índice-Particionado utilizando 32 partições e as características cor, textura e forma. Coleção BD-110	65
4.16	Comparação do CBIR-Índice-Particionado utilizando 16 partições com os sistemas CBIR-I e CBIR-Índice e as características cor, textura e forma. Coleção BD-110	66
4.17	Precisão média no sistema CBIR-I utilizando a característica cor. Coleção Corel-1000	67
4.18	Precisão média no sistema CBIR-Índice utilizando a característica cor. Coleção Corel-1000	67
4.19	Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e a característica cor. Coleção Corel-1000	67
4.20	Comparação do CBIR-Índice-Particionado utilizando 16 partições e a característica cor com os sistemas CBIR-I e CBIR-Índice . Coleção Corel-1000	68
4.21	Precisão média no sistema CBIR-I utilizando as características cor, forma e textura. Coleção Corel-1000	69
4.22	Precisão média no sistema CBIR-Índice utilizando as características cor, forma e textura. Coleção Corel-1000	69
4.23	Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e as características cor, forma e textura. Coleção Corel-1000	69
4.24	Comparação do CBIR-Índice-Particionado utilizando 16 partições e as características cor, forma e textura com os sistemas CBIR-I e CBIR-Índice. Coleção Corel-1000	70
4.25	Precisão média no sistema CBIR-I utilizando a característica cor. Coleção BD-10000	71
4.26	Precisão média no sistema CBIR-Índice utilizando a característica cor. Coleção BD-10000	71
4.27	Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e a característica cor. Coleção BD-10000	72
4.28	Comparação do CBIR-Índice-Particionado utilizando 16 partições e a característica cor com os sistemas CBIR-I e CBIR-Índice. Coleção BD-10000	72
4.29	Precisão média no sistema CBIR-I utilizando as características cor, forma e textura. Coleção BD-10000	73

4.30	Precisão média no sistema CBIR-Índice utilizando as características cor, forma e textura. Coleção BD-10000	73
4.31	Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e as características cor, forma e textura. Coleção BD-10000	73
4.32	Comparação do CBIR-Índice-Particionado utilizando 16 partições com os sistemas CBIR-I e CBIR-Índice. Coleção BD-10000	74

Lista de Abreviaturas e Siglas

BD	Banco de dados
CBIR	Recuperação de Imagens Baseada por Conteúdo
cf	Frequência na coleção
df	Frequência de documentos
fn	Falso negativo
fp	Falso positivo
HSV	Tonalidade, Saturação e Brilho
ID	Identificador do documento
idf	Frequência Inversa dos Documentos
LSI	Indexação semântica latente
NR	Conjuntos dos documentos não relevantes e não recuperados
PLN	Processamento de Linguagem Natural
Rec	Conjunto dos documentos recuperados
Rel	Conjunto dos documentos relevantes
RGB	Vermelho, Verde e Azul
RI	Recuperação de Informação
RR	Conjunto dos documentos relevantes recuperados
SGBD	Sistemas de Gerenciamento de Banco de Dados
SRI	Sistemas de Recuperação de Informação
tf	Frequência do termo
tn	Verdadeiro negativo
tp	Verdadeiro positivo

Termos do CBIR-Índice-Particionado

Cor:	MHA, MHB, ..., MHJ:	faixas A-J da Média da Matiz
	DHA, DHB, ..., DHJ:	faixas A-J do Desvio Padrão da Matiz
	IHA, IHB, ..., IHJ:	faixas A-J da Inclinação da Matiz
	MSA, MSB, ..., MSJ:	faixas A-J da Média da Saturação
	DSA, DSB, ..., DSJ:	faixas A-J do Desvio Padrão da saturação

	ISA, ISB, ..., ISJ:	faixas A-J da Inclinação da Saturação
	MVA, MVB, ..., MVJ:	faixas A-J da Média da Intensidade
	DVA, DVB, ..., DVJ:	faixas A-J do Desvio Padrão da Intensidade
	IVA, IVB, ..., IVJ:	faixas A-J da Inclinação da Intensidade
Forma:	H0A, H0B, ..., H0J:	faixas A-J do Histograma de Direção de Bordas 0°
	H45A, H45B, ..., H45J:	faixas A-J do Histograma de Direção de Bordas 45°
	H90A, H90B, ..., H90J:	faixas A-J do Histograma de Direção de Bordas 90°
	H135A, H135B, ..., H135J:	faixas A-J do Histograma de Direção de Bordas 135°
	H180A, H180B, ..., H180J:	faixas A-J do Histograma de Direção de Bordas 180°
	H225A, H225B, ..., H225J:	faixas A-J do Histograma de Direção de Bordas 225°
	H270A, H270B, ..., H270J:	faixas A-J do Histograma de Direção de Bordas 270°
	H315A, H315B, ..., H315J:	faixas A-J do Histograma de Direção de Bordas 315°
Textura:	V0A, V0B, ..., V0J:	faixas A-J da Vizinhaça de Textura 0°
	V45A, V45B, ..., V45J:	faixas A-J da Vizinhaça de Textura 45°
	V90A, V90B, ..., V90J:	faixas A-J da Vizinhaça de Textura 90°
	V135A, V135B, ..., V135J:	faixas A-J da Vizinhaça de Textura 135°
	V180A, V180B, ..., V180J:	faixas A-J da Vizinhaça de Textura 180°
	V225A, V225B, ..., V225J:	faixas A-J da Vizinhaça de Textura 225°
	V270A, V270B, ..., V270J:	faixas A-J da Vizinhaça de Textura 270°
	V315A, V315B, ..., V315J:	faixas A-J da Vizinhaça de Textura 315°

Capítulo 1

Introdução

O uso de imagens em nossa comunicação é bastante comum. Diversas são as áreas que contribuem para a geração de imagens digitais, por exemplo, entretenimento, bibliotecas digitais e educação. Nos últimos anos, o crescimento do número de imagens produzidas em meio digital foi imenso, sendo que o custo de processamento e armazenamento desse tipo de imagem decresceu consideravelmente.

A facilidade de distribuição advinda da popularização da Internet e o aumento do poder computacional disponível foram decisivos para crescimento do número de imagens produzidas em meio digital. Um dos principais objetivos nesse contexto é armazenar imagens de forma compacta em um banco de imagens, com possibilidades de recuperar as imagens para atender consultas efetuadas pelos usuários [Zhang, 2002]. Nesse cenário, muitas vezes o usuário tem dificuldades em localizar imagens específicas, surge então a necessidade crescente por melhores algoritmos de busca, métodos de indexação e técnicas de classificação de imagens digitais.

1.1 Motivação

O acesso e uso de imagens se torna uma atividade difícil de ser realizada quando não estão organizadas [Rui et al., 1997]. Um usuário final geralmente lida com um repositório de imagens cujo conteúdo é complicado e em parte desconhecido por ele. Tais situações são comuns, por exemplo, ao usar bancos de dados de imagem públicos na *Web*.

Um dos principais problemas encontrados na organização de uma coleção de imagens está relacionado em descrever os elementos presentes em uma imagem, seja uma pessoa, seja objeto, seja local, seja, literalmente, qualquer coisa. Uma imagem é um dado, que, embora tenha um significado visual para as pessoas, não possui nenhum significado para um sistema digital ou sistema computadorizado.

Dentre os métodos para facilitar a busca de imagens, pode-se citar o método de indexação das imagens por meio de palavras-chaves que descrevam as imagens [Lew, 2001]. Porém, este método é dispendioso em termos de tempo gasto com as descrições, pois estas

em geral são inseridas manualmente. Em alguns casos são extraídas por algoritmos complicados, ineficientes e imprecisos em relação à classificação das imagens, pois a descrição pode ser subjetiva [Datta et al., 2008].

Outro método muito utilizado é o de busca de imagens baseado no contexto do documento em que a imagem está inserida, processo utilizado pela máquina de busca Google [Google, 2010]. O processo se baseia no casamento de uma palavra-chave que descreve a imagem. O sistema procura imagens em arquivos de hipertexto que possuam a palavra-chave. Esta abordagem apresenta um bom desempenho no tempo de processamento, mas uma baixa eficiência na qualidade de recuperação, proporcionada pelo fato de que na maioria das vezes, as imagens inseridas nos hipertextos genéricos não necessariamente correspondem ao texto próximo a uma imagem.

Atualmente tem crescido a técnica de busca de imagens pelo seu próprio conteúdo digital denominada sistemas CBIR (Recuperação de Imagens Baseada por Conteúdo). O sistema CBIR engloba tecnologias que ajudam a organizar arquivos de imagens digitais pelas características visuais extraídas automaticamente da própria imagem [Datta et al., 2008]. Característica pode ser definida como uma função de uma ou mais medidas, calculadas de forma que quantifiquem alguma propriedade da imagem [Bueno, 2002].

Nas técnicas usadas num processo de CBIR, a consulta é feita tendo como base uma imagem exemplo, e não um termo ou palavra-chave. A vantagem desse método sobre os demais está no fato de as informações serem extraídas automaticamente das imagens e de um curto tempo gasto na indexação, por não serem geradas anotações textuais. Por outro lado, uma desvantagem é o cálculo da similaridade, pois muitos sistemas de recuperação calculam a similaridade da imagem de consulta com todas as imagens do banco de dados, tornando a recuperação demasiadamente lenta para grandes coleções como a *Web* [Datta et al., 2008].

Com o objetivo de melhorar o desempenho e o tempo de recuperação de imagens por conteúdo em grandes coleções, técnicas de indexação têm sido utilizadas na recuperação de imagens.

1.2 Objetivos

Em [Matos et al., 2008a] e [Matos et al., 2008b] são apresentadas propostas de adaptação de CBIR para a utilização da técnica existente no campo de recuperação de informação textual denominada Índice Invertido. O objetivo geral dos trabalhos é apresentar um método para classificar grupos de valores de características de baixo nível de imagens digitais, baseado na característica cor, que possibilite o mapeamento para um identificador de indexação da coleção para utilização do invertido invertido. Com esta abordagem obteve-se ganho no desempenho do processamento das consultas, mas a qualidade do conjunto recuperado não foi melhorada.

O objetivo deste trabalho é estender os trabalhos propostos por [Matos et al., 2008a] e [Matos et al., 2008b] contribuindo na qualidade e quantidade dos termos utilizados na indexação por meio do particionamento das imagens e aplicá-lo dentro do ambiente de recuperação por conteúdo para grandes coleções, visando não somente o ganho de desempenho no processamento das consultas, mas também ganho na precisão no conjunto de imagens recuperadas. Além do particionamento das imagens, outro diferencial do trabalho foi o acréscimo de duas novas características, forma e textura, nos grupos de valores de características das imagens. O uso dessas características trouxe ganho na precisão do conjunto resposta.

1.3 Trabalhos Relacionados

Um grande número de sistemas de recuperação de imagens para grandes coleções tem sido propostos nos últimos anos. Como resultado, vários sistemas CBIR acadêmicos e comerciais foram desenvolvidos. Nesta seção, alguns sistemas são introduzidos brevemente. O desempenho global das implementações atuais permanece ainda muito modesto quanto à tarefa de desenvolver um sistema de recuperação de imagem comparável aos sistemas de recuperação textuais [Kherfi et al., 2004].

Sistemas de Recuperação de Imagens Baseado no Conteúdo

Os sistemas de Recuperação de Imagens Baseada no Conteúdo são implementados em diversos ambientes. Estes sistemas operam com banco de dados de imagens predeterminados, oposto aos sistemas de busca de imagens da Web. O mais conhecido sistema de recuperação de imagens baseado no conteúdo é provavelmente o sistema da IBM chamado de Query By Image Content (QBIC) [Niblack et al., 1993]. Ele foi o primeiro sistema comercial de CBIR. Este sistema foi muito utilizado para avaliação e comparação de sistemas de recuperação de imagens baseado no conteúdo. Avaliações de características incluem cor, textura e forma. Utiliza a distância euclidiana para o cálculo da similaridade.

O sistema MARS [Rui et al., 1997] foi implementado com a característica de baixo nível cor, com o espaço de cor HSV, além da textura e forma. Faz uso da distância euclidiana para computar a similaridade de características de textura entre imagens. Outros sistemas usando a característica cor são apresentados em [Laaksonen, 2000], [Datta et al., 2008] e [Rui et al., 1999].

Os sistemas de Recuperação de Imagens Baseada no Conteúdo apresentam bons resultados de avaliação. Mas apresentam problemas em relação ao tempo de processamento para grandes coleções como a Web. Pois o processo de recuperação inclui o cálculo da distância entre o vetor da imagem de consulta e todos os vetores das imagens do banco de dados, tornando a recuperação computacionalmente cara. Por este motivo sistemas

CBIR não são usualmente utilizados em motores de busca da Web.

Sistemas de Busca de Imagens na Web

A recuperação de informações visuais e de multimídia a partir da Web é uma área que vem sendo muito explorada [Datta et al., 2008]. Os atuais motores de busca de textos, como o AltaVista e Google, não são adequados para CBIR. Motores de busca de imagens enfrentam os mesmos problemas que o texto com base nos mecanismos de busca, como o enorme tamanho da coleção, a diversidade e a natureza dinâmica da Web [Koskela, 1999].

Dentre os sistemas de busca de imagens da Web, destaca-se o site Google, baseado no texto adjacente à imagem, presente na legenda da imagem e dezenas de outros fatores que determinam o conteúdo da imagem. A pesquisa de Imagens do Google tem mais de 1 bilhão de imagens indexadas e disponíveis para visualização [Google, 2010].

O sistema *ImageRover* [Sclaroff et al., 1997] foi desenvolvido no Departamento de Ciência da Computação da Universidade de Boston. *ImageRover* combina características textuais e visuais em um único índice para busca baseada no conteúdo de uma coleção de imagens da Web. O *ImageRover* representa as imagens analisando o conteúdo HTML de onde elas se encontram e analisando histogramas de cor, histogramas de orientação e textura das imagens.

O Diogenes [Aslandogan and Yu, 2000b] e [Aslandogan and Yu, 2000a], desenvolvido na Universidade de Illinois, EUA, é um sistema específico para busca de imagens de faces de celebridades. O usuário primeiro informa o nome da celebridade textualmente, o sistema então realiza a busca em sistemas de busca textual, como o Google, e os resultados são avaliados levando em conta o conteúdo HTML de onde as imagens foram encontradas e o conteúdo visual delas, a partir de um sistema detector de faces.

O AtlasWISE, projeto do Departamento de Matemática e Ciência da Computação da Universidade de Sherbrooke, Canadá, faz a coleta de dados da Web navegando por páginas de índices populares, como o Google e o Yahoo. O sistema utiliza ambas as características, visual e textual, para descrever e indexar as imagens. As características visuais são avaliadas por histogramas de cor e histogramas de orientação de bordas. As características textuais são captadas de suas legendas e tags, títulos das páginas e os textos ao seus arredores. O sistema também possui um módulo de realimentação de relevantes, no qual o usuário pode informar exemplos positivos e negativos na formulação de consultas [Kherfi et al., 2004].

Representação de Imagens Digitais baseada em *Bag-of-features*

Representações de *bag-of-features* se tornaram populares recentemente para classificação de imagem baseada em conteúdo devido à sua simplicidade e bom desempenho. A idéia básica é descrever uma imagem como uma coleção não-ordenada de características

locais em torno de regiões de pontos de interesse. Para uma representação compacta, um vocabulário visual é geralmente construído por meio do agrupamento (*clustering*) dessas características. Cada grupo (*cluster*) de pontos de interesse é tratado como uma palavra visual no vocabulário. Por meio do mapeamento dos pontos de interesse em uma imagem para palavras no vocabulário visual, pode-se descrever a imagem como um vetor de características de acordo com a presença ou contagem de cada palavra visual [Jiang et al., 2007]. Esse vetor de características é o *bag-of-features* de uma imagem e um classificador pode ser aplicado aos vetores de características de determinado conjunto de imagens para determinar à qual categoria cada imagem pertence.

Na literatura pode-se encontrar diversos trabalhos que utilizam *Bag-of-features*, dentre eles podemos citar: [Lopes et al., 2009a], [Lopes et al., 2009b], [Nowak et al., 2006a] e [Csurka et al., 2004].

Contextualização e Contribuição do Trabalho

Os sistemas Google, *ImageRover*, Diogenes, *AtlasWise* e *ImageScape* apresentam bons resultados na velocidade de recuperação, mas apresentam o problema em relação a sua fraca eficiência e qualidade de recuperação, proporcionada pelo fato de que na maioria das vezes o texto próximo a uma figura não a descreve fielmente. Isso motiva nosso trabalho com CBIR.

O sistema QBIC, assim como alguns outros sistemas, os quais o processo de recuperação inclui o cálculo da distância entre o vetor da imagem de consulta e todos os vetores das imagens do banco de dados, apresentam problemas em relação ao tempo de processamento para grandes coleções como a Web. Pelo fato de compararem a imagem de consulta, com todas as imagens da coleção, torna a recuperação computacionalmente cara. Por este motivo, esses sistemas não são usualmente utilizados em motores de busca da Web.

Com o intuito de aproveitar a vantagem dos sistemas de recuperação baseado no conteúdo visual, que conseguem um bom desempenho na qualidade de recuperação juntamente com a vantagem dos sistemas de recuperação de imagens da Web, que apresentam um alto desempenho na velocidade de recuperação, novas abordagens, tanto para indexação das imagens quanto para o cálculo da similaridade em CBIR vem sido propostas. Os trabalhos com *bag-of-features* são uma alternativa. Uma abordagem semelhante foi usada em [Matos et al., 2008a].

Em [Matos et al., 2008a] é apresentada uma proposta de construção do Índice Invertido para recuperação de imagens baseado em conteúdo (CBIR). O objetivo é acelerar o processamento de consultas, sem perda de qualidade na resposta. O sistema indexa as imagens através de uma estrutura muito utilizada na recuperação textual, conhecida como índice invertido e a similaridade entre a imagem de consulta e a coleção de imagens é medida através do cálculo do cosseno entre vetores representantes da imagem de consulta

e das imagens do banco de dados.

Uma desvantagem da proposta de [Matos et al., 2008a] é o uso exclusivo da característica cor para a descrição e indexação das imagens. Com o intuito de aproveitar a vantagem desse sistema propomos uma abordagem que visa contribuir na qualidade e quantidade dos termos utilizados na indexação por meio das partições das imagens, e do uso de textura e forma, e aplicá-la dentro do ambiente de recuperação por conteúdo para grandes coleções, visando não somente o ganho de desempenho no processamento das consultas, mas também ganho na precisão no conjunto de imagens recuperadas.

A principal contribuição deste trabalho é a proposta de um sistema de recuperação de imagens batizado no conteúdo, indexado com base nas três principais características de baixo nível utilizadas na literatura, a saber, cor, textura e forma, extraídas de partições de imagens digitais. Essas características são mapeadas para um índice invertido e o sistema é avaliado experimentalmente, mostrando ganho na precisão da recuperação das imagens, em relação a sistema semelhante baseado em cor e outro baseado na distância Euclidiana. Além disso, o sistema proposto mantém um bom desempenho no processamento.

1.4 Estrutura do Trabalho

Esta dissertação está organizada da seguinte forma:

- No Capítulo 2 são apresentados os conceitos básicos utilizados no desenvolvimento deste trabalho.
- No Capítulo 3 é apresentado o sistema de recuperação de imagem por conteúdo proposto, bem como as modificações realizadas nos sistemas de [Matos et al., 2008a] e [Matos et al., 2008b], a fim de utilizá-los como base de avaliação da proposta deste trabalho.
- No Capítulo 4 são apresentadas as avaliações experimentais: comparação da nossa técnica com as técnicas apresentadas em [Matos et al., 2008a] e [Matos et al., 2008b].
- No Capítulo 5 são apresentadas as conclusões e as propostas de trabalhos futuros.

Capítulo 2

Conceitos Básicos

2.1 Representação de Imagens

O uso de imagens é antigo. Povos pré-históricos desenhavam cenas de seu cotidiano em paredes de cavernas como uma forma de registrar seus feitos. Um fato notável é que ainda hoje, milhares de anos depois, somos capazes de inferir fatos da vida de pessoas que viveram a tanto tempo, por meio da observação de ilustrações. Imagens transmitem informações, histórias e humor, auxiliam na divulgação de cultura, na educação em outras coisas mais. À medida que a humanidade evoluiu, muitas coisas aconteceram, fatos históricos, inventos, descobertas, surgindo a necessidade e possibilidade destes serem registrados por meio de imagens.

Dentre as diversas formas de representação de imagens está a imagem digital. Uma imagem digital é uma função bidimensional $f(x, y)$ discretizada tanto em coordenadas espaciais quanto em brilho. Esta pode ser considerada como sendo uma matriz cujos índices de linhas e de colunas identificam um ponto na imagem e o correspondente valor de elemento da matriz identifica a informação de cor naquele ponto. Quando os valores de x e y juntamente com os valores da amplitude f são todos finitos, a imagem é chamada de imagem digital. Os elementos dessa matriz digital são chamados de elementos da imagem, elementos da figura, *pixels* ou *pels*, estes dois últimos, abreviações de "*picture elements*" (elementos de figura). Neste trabalho é utilizado o termo *pixel* [Gonzales and Woods, 2002].

A matriz de *pixels* quando interpretada pelo dispositivo de exibição produz o efeito visual perceptível à visão humana, mas não é adequada para comparação ou classificação de imagens. Com esse objetivo, um dos principais problemas é encontrar a melhor representação numérica que sintetize a essência da imagem através de um vetor de características. Assim, uma imagem pode ser considerada um ponto no espaço vetorial definido por suas características, e esses vetores podem ser utilizados para comparação e classificação de imagens.

As características podem ser de baixo ou alto nível. Baixo nível são as representações matemáticas das características visuais das imagens como cor, textura e forma. Alto nível refere-se à semântica, como objetos, interpretações de cenas, ações. Por meio de algoritmos automáticos de processamento de imagens, são criados vetores de características que representam propriedades das imagens [Datta et al., 2008]. O número de características dependerá do domínio de imagens e das propriedades que se deseja caracterizar, esse número determinará a dimensão do vetor de características [Gonzales and Woods, 2002].

As características mais importantes das imagens são aquelas que permitem ao próprio sistema de visão humana identificar e posteriormente descrever de forma semântica os objetos de uma determinada imagem. Tais atributos podem ser extraídos de forma automática através de um processo computacional. Os principais atributos discutidos aqui são cor, forma e textura.

2.1.1 Atributo COR

A cor é uma das características visuais mais utilizadas na recuperação de imagens e possui um papel bastante significativo na indexação e recuperação. A utilização da cor no processamento de imagem é motivada por dois fatores principais. Primeiramente, é um poderoso descritor que geralmente simplifica a identificação e extração de um objeto em uma cena. O outro fator é que humanos podem discernir milhares de tons de cores e intensidades, em comparação com apenas cerca de duas dezenas de tons de cinza. Esse segundo fator é particularmente importante na análise manual de imagem (ou seja, feita por humanos) [Gonzales and Woods, 2002].

Existem diferentes representações de cores, sendo as mais usadas o RGB (*red, green, blue*) e o HSV (*hue, saturation, value ou intensity*). O RGB é o modelo mais simples que mapeia diretamente as características físicas do dispositivo de exibição. Já o HSV reflete mais precisamente o modelo de cores para a percepção humana.

Neste trabalho foram abordados esses dois modelos de cor: RGB e HSV pois são sistemas projetados para interagir com o usuário e com isso são apropriados para CBIR.

Modelo RGB

O modelo de cor mais utilizado em representação gráfica é o RGB. Os monitores do computador que se baseiam na tecnologia de tubos catódicos (CRT), por exemplo, usam fósforos vermelhos, verdes e azuis para criar as cores em suas telas.

O modelo RGB, baseado na mistura de luz colorida, é chamado frequentemente de modelo aditivo, no qual o vermelho, o verde e o azul (usados em modelos aditivos de luzes) são combinados de várias maneiras para reproduzir outras cores.

O nome do modelo vem das três cores primárias *Red, Green* e *Blue*, sua origem está no desenvolvimento tecnológico de tubos de raios catódicos, com os quais foi possível fazer

a exibição de cores ao invés de uma fosforescência monocromática (incluindo a escala de cinza), como no filme preto e branco e nas imagens de televisão antigas, [Gonzales and Woods, 2002].

Uma cor no modelo de cores RGB pode ser descrita pela intensidade de cada cor primária: vermelho, verde e azul. Cada uma pode variar entre o mínimo (completamente escuro) e máximo (completamente intenso). Quando todas as cores estão no mínimo, o resultado é preto. Se todas estão no máximo, o resultado é branco. Onde quaisquer duas destas cores sobrepõem-se, uma cor secundária é criada. Por exemplo, onde as luzes vermelhas e verdes se sobrepõem, amarelo é criado. Variando o brilho do vermelho, verde e azul, é possível criar milhões de cores diferentes, Figura 2.1.

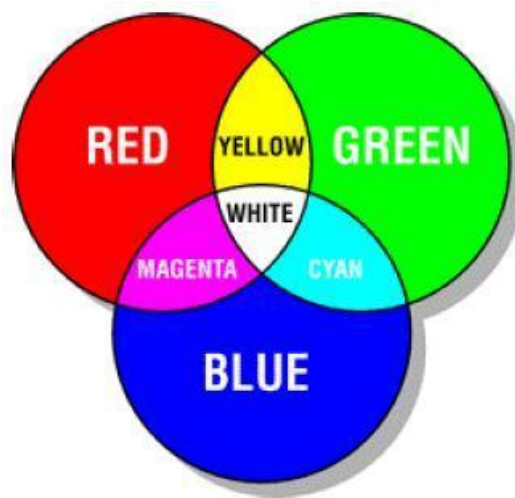


Figura 2.1: Modelo de cor RGB.

Uma Imagem colorida RGB é um arranjo de $M \times N \times 3$ *píxels* onde cada *píxel* é uma tripla correspondente às respectivas cores do modelo em uma localização espacial específica definida por M e N . Este modelo é baseado em um sistema de coordenadas cartesianas, mostrado na Figura 2.2. Podem ser observadas as coordenadas das cores vermelho, verde, azul, ciano, magenta e amarelo. O preto está na origem $(0,0,0)$ e o branco em $(1,1,1)$, e na diagonal do cubo ligando estes pontos temos o degradê de níveis de cinza.

Modelo HSV

Outro modelo de definição da cor é o HSV, que define as cores em função dos valores de três importantes atributos: tonalidade ou matiz, saturação e brilho, sendo que este último representa a luminosidade ou o brilho de uma cor.

O modelo HSV é a concepção intuitiva da técnica utilizada por um artista ao misturar cores básicas para obtenção de outros tons compostos. A seleção e obtenção de cores no

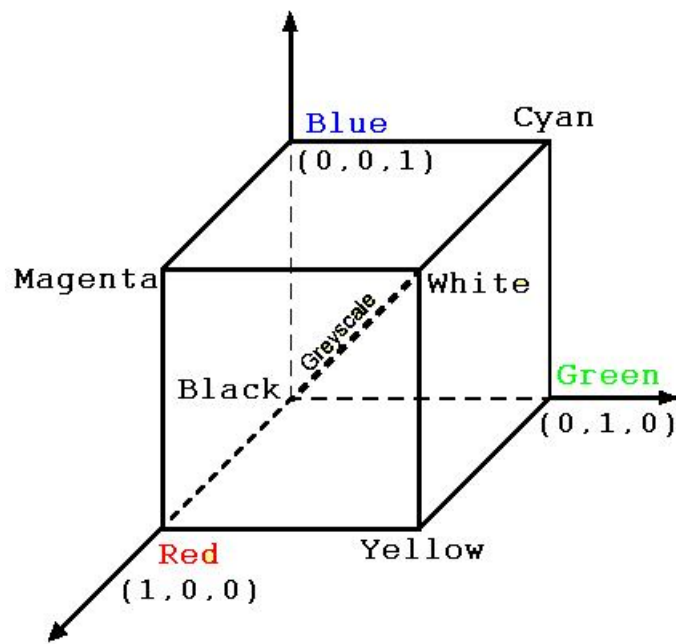


Figura 2.2: Esquema de cubo do espaço de cor RGB.

modelo HSV é muito mais intuitiva que no modelo RGB. O modelo HSV baseia-se no controle dos valores de H, S e V.

Matiz (H) é a componente que seleciona a “tinta” a ser utilizada, sendo controlada pela posição angular de um ponteiro numa roda de cores definida de 0° a 360° graus. Ou seja, é a cor pura com saturação e luminosidade máximas, permitindo fazer a distinção das várias cores puras.

Saturação (S) é a componente que determina a pureza da cor selecionada na matiz. Todos os tons de cinza possuem saturação igual a 0 e todas as matizes puras possuem saturação igual a 1. Uma cor saturada ou pura não contém a cor preta nem a branca. Desta forma a saturação é utilizada para descrever quão viva ou pura é a cor e em termos técnicos descreve a quantidade de cinzas em uma cor.

Brilho (V) traduz a luminosidade ou o brilho de uma cor, isto é, se uma cor é mais clara ou mais escura, indicando a quantidade de luz que a mesma contém. O termo luminosidade está relacionado com a luz refletida, enquanto que o termo brilho está relacionado com a luz emitida. A cor preta possui brilho zero para qualquer valor de matiz ou saturação. O valor 1 de brilho determina uma intensidade pura de matiz e saturação. A Figura 2.3 ilustra a representação gráfica deste modelo.

O sistema de representação HSV busca de forma mais próxima da manipulação natural de cor utilizada pelo homem. Nesse sistema, torna-se mais natural a busca por uma determinada cor, dada sua matiz, variando a saturação e o brilho de forma independente.

O modelo de cor HSV é considerado o mais próximo do sistema RGB. Ele é utilizado aqui pelo fato de decompor a cor em sua tonalidade predominante e pureza (o que permite

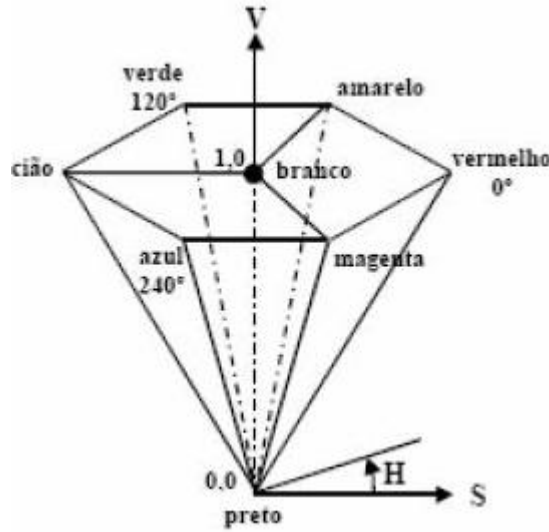


Figura 2.3: Espaço de cor HSV [Gonzales and Woods, 2002].

uma função de comparação de características baseadas em cor (uma discriminação adequada de tonalidades semelhantes) e na componente brilho, que pode variar em diferentes cenas ou condições.

Seja uma cor definida por RGB, onde R, G e B estão normalizados entre 0 e 1, correspondendo, respectivamente, ao maior e ao menor valor possível para cada atributo. A transformação para os parâmetros (H, S, V) dessa cor pode ser determinada pelas seguintes relações [Gonzales and Woods, 2002]:

$$H = \arctan \sqrt{\frac{(R - G) + (R - B)}{(R - G)^2 + (R - G)(G - B)}} \quad (2.1)$$

$$S = 1 - \frac{3}{R + G + B} \times \min(R, G, B) \quad (2.2)$$

$$V = \frac{(R + G + B)}{3} \quad (2.3)$$

Na conversão de uma imagem digital no modelo RGB para o modelo HSV, o processo é aplicado em cada *pixel* da imagem, ou seja, é calculado H, S e V para cada *pixel*.

A vantagem dessa representação de cor está na possibilidade de separar os atributos H, S e V. Esse fato torna o modelo HSV uma ferramenta ideal para o desenvolvimento de algoritmos de processamento de imagens.

O próximo passo é utilizar um descritor de cor para descrever a informação de cor da imagem. Como, até o momento, têm-se 3 atributos de cor (H,S,V), estes serão utilizados pelo descritor de cor, momentos de cor, que será abordado a seguir.

Descritores de cor

Um descritor de cor é utilizado para sumarizar a informação de cor da imagem a partir dos *pixels* da imagem digital. Na literatura são apresentados descritores de cor comumente usados, alguns deles são: momentos de cor [Stricker and Orengo, 1995], histograma de cores [Swain and Ballard, 1991], vetor de coerência de cores [Pass and Zabith, 1996] e correlograma de cores [Huang and Zabih, 1997].

Momentos de cor [Stricker and Orengo, 1995] caracterizam as imagens em termos da distribuição das cores. Estes têm sido utilizados com sucesso em muitos sistemas CBIR, sendo o mais famoso, o QBIC [Flickner et al., 1995] da IBM.

Momentos de cor são normalmente obtidos por três medidas estatísticas: média, desvio padrão e obliquidade, calculadas sobre cada canal de cor de uma imagem definido pelas cores primitivas do modelo. No modelo HSV os momentos de cor obtêm nove características da imagem utilizando todos *pixels* da imagem, conforme descrito a seguir.

Sendo N o número de *pixels* de uma imagem e p_{ij} o valor do j -ésimo *pixel* no i -ésimo canal de cor, os três primeiros momentos: média, desvio padrão e obliquidade são dados pelas seguintes expressões, respectivamente:

$$E_i = \frac{1}{N} \sum_{j=1}^N p_{ij} \quad (2.4)$$

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^2 \right)} \quad (2.5)$$

$$s_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^3 \right)} \quad (2.6)$$

Devido a sua alta compactabilidade e por não considerar informações espaciais, momentos de cor têm poder de discriminação reduzido. Porém sua performance pode ser significativamente melhorada aplicando-o a regiões de imagem devido à descrição de características locais [Laaksonen, 2000]. Esta foi a motivação para uso deste descritor no trabalho.

A utilização deste descritor gera um vetor de características de nove dimensões, representando os três primeiros momentos de cor descritos acima (Equações 2.4, 2.5 e 2.6) para cada canal de cor. Ou seja, média, desvio padrão e obliquidade do canal H (matiz ou tonalidade); média, desvio padrão e obliquidade do canal S (saturação) e média, desvio padrão e obliquidade do canal V (brilho).

2.1.2 Atributo Textura

O descritor de textura fornece medidas de propriedades como suavidade, rugosidade e regularidade (Figura 2.4). A textura é um descritor importante para segmentação das imagens em objetos ou regiões distintas, como também para classificação ou reconhecimento de materiais de superfície [Gonzales and Woods, 2002].

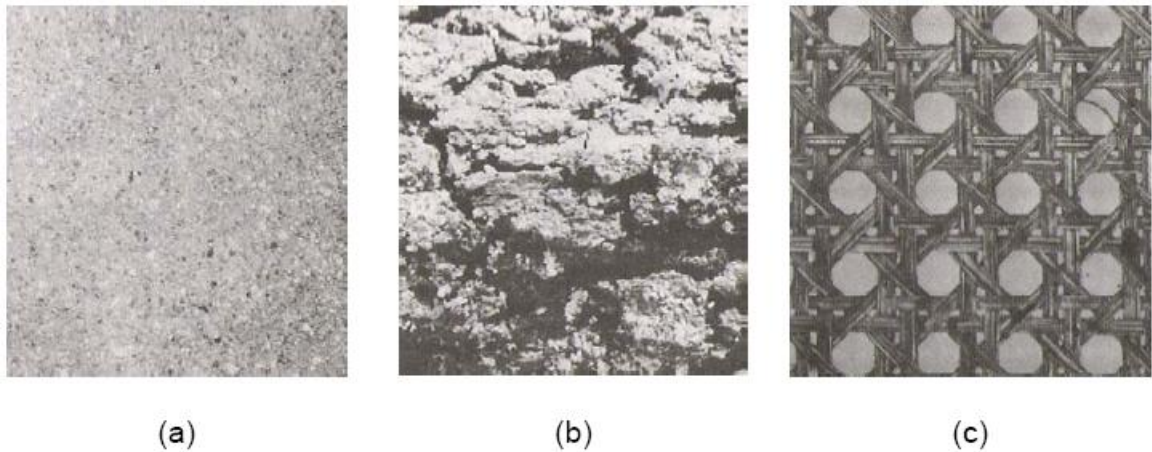


Figura 2.4: Exemplos de textura (a) suave, (b) áspera, (c) regular. Fonte: [Gonzales and Woods, 2002].

Textura é uma propriedade presente em praticamente todas as estruturas, como nuvens, vegetação, paredes, cabelo e outros. Ela contém informação importante sobre o arranjo estrutural da superfície e sua relação com o ambiente [Rui et al., 1997]. Em outras palavras, textura é composta por elementos de intensidade uniforme e de forma simples que se repetem num intervalo. Por ser um padrão útil e importante em visão computacional e reconhecimento de padrão, é uma área de pesquisa rica em trabalhos nas últimas décadas. Atualmente, são encontrados muitos trabalhos de pesquisas na área de recuperação de informação visual baseada em textura.

A análise de texturas é um dos aspectos fundamentais da visão humana, com a qual é possível discriminar entre superfícies e objetos. De modo semelhante, utilizando-se da visão computacional, pode-se tirar proveito das pistas fornecidas pela textura de superfícies para distinguir e reconhecer objetos. Logo, a textura representa uma fonte de informação visual muito rica acerca da natureza e estrutura tridimensional dos objetos físicos.

Embora o reconhecimento de textura seja fácil para as pessoas, isso não ocorre com procedimentos automáticos, já que essa tarefa geralmente precisa de técnicas computacionais complexas. O tratamento de textura difere do realizado sobre as cores, devido ao fato de que as texturas são definidas sobre janelas ou regiões da imagem e não sobre *pixels* individuais como as cores. A segmentação de uma imagem utilizando textura determina

quais regiões da imagem possuem textura uniforme [Bueno, 2002]. A figura 2.5 apresenta alguns exemplos de textura.



Figura 2.5: Exemplos de textura.

De acordo com [Long et al., 2000], modelos computacionais de característica de textura podem ser agrupados em três categorias: estrutural, estatístico e espectral. Modelos estruturais caracterizam texturas de acordo com o relacionamento local entre pixels de imagens, ou seja, lidam com o arranjo espacial de primitivas estruturais. Modelos estatísticos categorizam texturas de acordo com medidas estatísticas de característica visual, tais como grossura, granularidade, regularidade, entre outros. Este modelo considera também a distribuição dos tons de cinza e o inter-relacionamento entre eles. Por fim, modelos espectrais caracterizam textura como propriedades da transformada de Fourier ou nos resultados de filtragem das texturas por filtros apropriados.

A análise de textura tem por objetivo estabelecer o relacionamento de vizinhança dos elementos de textura e seu posicionamento em relação aos demais (conectividade), o número de elementos por unidade espacial (densidade) e a sua regularidade (homogeneidade) [Gonzales and Woods, 2002]. O uso dessa análise pode ser encontrado em diversas áreas computacionais, como reconhecimento de padrões e processamento de imagens.

Algumas das mais conhecidas técnicas de extração de textura são as *Wavelets*, os filtros de Gabor e aquelas baseadas nas matrizes de co-ocorrência. A partir dessas técnicas é possível computar medidas de periodicidade, granularidade, direcionalidade e regularidade das regiões das imagens.

Os experimentos neste trabalho foram realizados com os descritores de vizinhança de textura definidos a seguir. A escolha se deu ao fato de que a medida utilizada por essa técnica é fortemente baseada nos aspectos da percepção humana de textura [Laaksonen, 2000].

Vizinhança de textura

No método de vizinhança de textura a imagem é representada em níveis de cinza. Um descritor de dimensão oito (V_0 , V_{45} , V_{90} , V_{135} , V_{180} , V_{225} , V_{270} e V_{315}) é gerado da seguinte forma: para cada *pixel* da imagem são analisados os *pixels* de vizinhança oito, se os valores dos *pixels* dentre os da vizinhança forem maiores ou iguais ao valor do *pixel* central, então

os contadores direcionais indicando as direções dos *pixels* de valores maiores ou iguais são incrementados [Laaksonen, 2000].

Considere o segundo *pixel* da segunda linha da Figura 2.6a, os *pixels* de sua vizinhança oito a serem analisados são os *pixels* destacados da Figura 2.6b. Dentre os oito *pixels* em análise os únicos que possuem valores maiores ou iguais são o primeiro *pixel* da primeira linha e o terceiro da terceira linha, assim os descritores nas direções V_{135} e V_{315} são incrementados. Todos os *pixels* da imagem serão analisados desta mesma forma.

255	200	200	200
200	255	200	200
200	200	255	200
200	200	200	255

a)

255	200	200	200
200	255	200	200
200	200	255	200
200	200	200	255

b)

Figura 2.6: a) Imagem em nível de cinza. b) Vizinhança oito do segundo *pixel* da segunda linha.

2.1.3 Atributo Forma

A forma de um objeto é uma característica importante para seu reconhecimento, da mesma maneira que a cor e a textura também são importantes. De fato é possível identificar diferentes objetos em uma imagem analisando apenas suas formas.

Apesar da importância, a recuperação de imagens baseada em forma é um dos problemas mais difíceis de serem tratados pelos sistemas de recuperação de imagens. Isto se deve, principalmente, à dificuldade de segmentar os objetos de uma imagem, sendo que esta recuperação de formas deve ser tipicamente limitada aos poucos objetos melhores discriminados que estão presentes na imagem [Bueno, 2002].

A segmentação se refere ao processo de dividir uma imagem digital em múltiplas regiões (conjunto de *pixels*) ou objetos, com o objetivo de simplificar e/ou mudar a representação de uma imagem para facilitar a sua análise. Segmentação de imagens é tipicamente usada para localizar objetos e formas (linhas, curvas, etc) em imagens. O nível até o qual essa subdivisão deve se realizada, assim como a técnica utilizada, depende do problema que está sendo resolvido [Gonzales and Woods, 2002].

A identificação dessas características é muito complexa quando estas não possuem formas pré-definidas. [Bueno, 2002] afirma que o pré-processamento de filtros e algoritmos para identificação de objetos depende do domínio da aplicação.

O processo de recuperação de imagens baseado em forma pode ser dividido em duas categorias: os baseados em contorno e os baseados em região. A primeira categoria adota o reconhecimento dos objetos baseado na sua forma, isto é, reconhecimento de forma

retangular, triangular, circular e outras formas. A segunda categoria adota o reconhecimento dos objetos baseado em regiões, ou seja, reconhecimento de semelhanças entre determinadas regiões de uma imagem. Em cada um destes grupos os diferentes métodos são divididos em abordagens estruturais e abordagens globais, Figura 2.7 [Zhang and Lu, 2004].

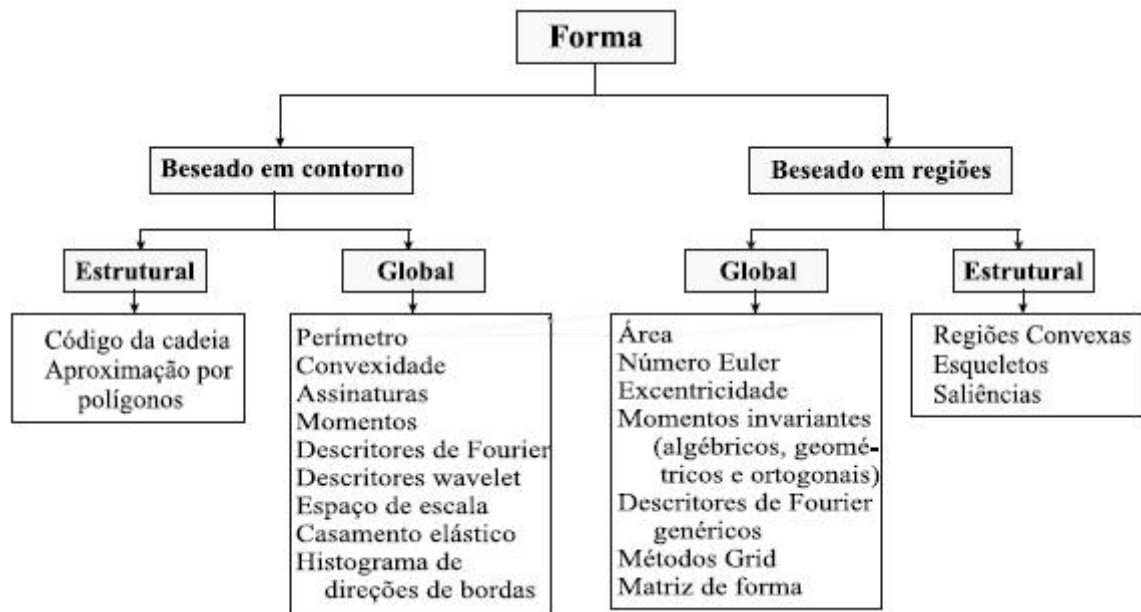


Figura 2.7: Classificação geral das técnicas de representação de formas [Zhang and Lu, 2004].

Métodos baseados em região

As técnicas baseadas em região levam em conta não só a borda da forma do objeto, como também os detalhes internos da forma do mesmo, ou seja, todos os *pixels* presentes no interior da forma do objeto analisado são utilizados para obtenção da representação da mesma [Safar et al., 1999]. Similarmente, os métodos baseados em regiões podem ser divididos em métodos globais e métodos estruturais.

Os métodos globais tratam a forma como um todo. Estes métodos medem a distribuição dos *pixels* de regiões e nesta categoria incluem: área, circularidade e excentricidade.

Métodos estruturais decompõem as regiões em partes que são usadas para as descrições. Os métodos estruturais baseados em regiões mais conhecidos são fecho convexo e saliências [Zhang and Lu, 2004].

Métodos baseados em contorno

Os métodos baseados em contorno usam apenas o contorno (ou bordas) da forma de um objeto e ignoram completamente seu interior. Há dois tipos de abordagens para a modelagem do contorno de formas: a abordagem contínua (global) e a abordagem discreta (estrutural) [Zhang and Lu, 2004].

Na abordagem global, um vetor de característica numérico derivado de toda a borda é usado para descrever a forma. Medidas de similaridade entre formas são usualmente tomadas por alguma métrica de distância.

Mais detalhes sobre diversas técnicas de extração de forma podem ser encontradas em [Zhang, 2002] e [Zhang and Lu, 2004].

Em [Brandt, 1999], [Laaksonen, 2000] e [da Silva, 1999] os autores utilizaram o método global baseado em contorno Histograma de Direção de Bordas e obtiveram bons resultados. Baseado nesses resultados, nos experimentos realizados neste trabalho também foi utilizado o método global baseado em contorno Histograma de Direção de Bordas, definido a seguir.

Histograma de Direção de Bordas

O histograma de direção de bordas é um dos descritores de forma mais populares. Sua probabilidade de sucesso é atribuída à sua simplicidade e ao fato de não necessitar de segmentação de objetos. O histograma de direção de bordas modela a similaridade entre imagens em termos da distribuição das direções de bordas [Brandt, 1999].

Primeiramente uma imagem é pré-processada para extrair as bordas e suas direções. Operadores de Canny, de Sobel, ou derivações Gaussianas podem ser usados na detecção de bordas. Depois deste pré-processamento um histograma é formado pela frequência das direções de borda. Histogramas de bordas são representações compactas, invariantes às operações de translação e de escala.

Em [Brandt, 1999] e [da Silva, 1999] os autores utilizaram operadores de Sobel na detecção de bordas das imagens e através de experimentos os autores obtiveram bons resultados. Considerando esses resultados também utilizaremos neste trabalho o operador de Sobel de oito direções.

Primeiramente as imagens são convertidas para o espaço de cor HSV, onde somente os canais S e V são usados na identificação das direções de bordas. A razão do descarte do canal H se dá pelo fato deste negligenciar bordas importantes em regiões ou objetos com a mesma tonalidade de cor. Operadores de Sobel de oito direções (ver Figura 2.8) são aplicados sobre os dois canais e é definido um descritor representando as oito direções ($H_0^\circ, H_{45}^\circ, H_{90}^\circ, H_{135}^\circ, H_{180}^\circ, H_{225}^\circ, H_{270}^\circ, H_{315}^\circ$). Cada *pixel*, de cada canal de cor, é analisado segundo as 8 máscaras de Sobel e o maior valor obtido define a direção da borda naquele *pixel* e o contador correspondente àquela direção é incrementado.

<table><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>0</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr></table> 0°	-1	0	1	-2	0	2	-1	0	1	<table><tr><td>0</td><td>1</td><td>2</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>-2</td><td>-1</td><td>0</td></tr></table> 45°	0	1	2	-1	0	1	-2	-1	0	<table><tr><td>1</td><td>2</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>-1</td><td>-2</td><td>-1</td></tr></table> 90°	1	2	1	0	0	0	-1	-2	-1	<table><tr><td>2</td><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>0</td><td>-1</td><td>-2</td></tr></table> 135°	2	1	0	1	0	-1	0	-1	-2
-1	0	1																																					
-2	0	2																																					
-1	0	1																																					
0	1	2																																					
-1	0	1																																					
-2	-1	0																																					
1	2	1																																					
0	0	0																																					
-1	-2	-1																																					
2	1	0																																					
1	0	-1																																					
0	-1	-2																																					
<table><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>2</td><td>0</td><td>-2</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr></table> 180°	1	0	-1	2	0	-2	1	0	-1	<table><tr><td>0</td><td>-1</td><td>-2</td></tr><tr><td>1</td><td>0</td><td>-1</td></tr><tr><td>2</td><td>1</td><td>0</td></tr></table> 225°	0	-1	-2	1	0	-1	2	1	0	<table><tr><td>-1</td><td>-2</td><td>-1</td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>2</td><td>1</td></tr></table> 270°	-1	-2	-1	0	0	0	1	2	1	<table><tr><td>-2</td><td>-1</td><td>0</td></tr><tr><td>-1</td><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td><td>2</td></tr></table> 315°	-2	-1	0	-1	0	1	0	1	2
1	0	-1																																					
2	0	-2																																					
1	0	-1																																					
0	-1	-2																																					
1	0	-1																																					
2	1	0																																					
-1	-2	-1																																					
0	0	0																																					
1	2	1																																					
-2	-1	0																																					
-1	0	1																																					
0	1	2																																					

Figura 2.8: Operadores de Sobel

2.1.4 Considerações Finais

Nesta seção foram apresentadas as características utilizadas na Recuperação de Imagens. As características mais importantes das imagens são aquelas que permitem ao próprio sistema de visão humana identificar e posteriormente descrever de forma semântica os objetos de uma determinada imagem. Estas características podem ser de baixo nível ou alto nível. Baixo nível são as representações matemáticas das características visuais das imagens como cor, textura e forma. Alto nível refere-se à semântica, como objetos, interpretações de cenas, ações. Aqui foram discutidas apenas as características de baixo nível.

Após a extração de características das imagens, é possível o reconhecimento de seus elementos, etapa crucial para o processo de recuperação de imagens através de similaridades.

Neste trabalho utilizamos como descritores momentos de cor, vizinhança de textura e operadores de sobel de oito direções, conforme descrito nesta seção.

2.2 Recuperação de Imagens

A recuperação de imagens pode ser analisada sob dois ângulos diferentes, um baseado no conteúdo textual e o outro baseado no conteúdo visual. O processo de recuperação de imagens baseado em conteúdo textual refere-se ao conceito de Anotação de Imagens (*Image Annotation*), que apresenta algumas técnicas com a tentativa de solucionar o problema de como descrever os elementos presentes numa imagem. O processo de recuperação de imagens baseado em conteúdo visual envolve basicamente duas etapas: Extração de Características de Imagens e Registro de Imagens [Rui et al., 1997].

Ao longo desta seção serão apresentados os conceitos relativos a cada uma das etapas. Primeiramente, serão apresentadas as etapas envolvidas no processo de recuperação de imagens baseada em conteúdo textual e posteriormente as etapas envolvidas no processo de recuperação de imagens baseada em conteúdo visual, foco deste trabalho.

2.2.1 Recuperação de Imagens Baseada em Conteúdo Textual

O crescimento muito rápido das coleções de imagens digitais ocasionou a necessidade de um repositório para armazenamento destas imagens. Com isso é necessário um melhor gerenciamento destas imagens, qual seja: armazenamento, edição, publicação, visualização e pesquisa destas imagens. Essa necessidade de gerenciamento tende a aumentar e conseqüentemente, requerer sistemas de gerenciamento de imagens cada vez mais robustos e eficientes [Rodden, 1999].

A Anotação de Imagens é um meio de indexação semântico de imagens através de um conjunto de palavras (termos-chave ou palavras-chave) que são associadas à imagem. Anotação de Imagens provê formas de identificar termos que proporcionam algum significado à imagem e permitem sua recuperação por meio da utilização desses termos [Rui et al., 1997].

A Anotação de Imagem tem chamado a atenção de pesquisadores com objetivo de criar técnicas mais eficientes e precisas para a indexação semântica de imagens como, por exemplo, os trabalhos de [Shneiderman and Kang, 2000], [Lew, 2001] e [Flickner et al., 1995]. As técnicas de anotação de imagens podem ser classificadas em três grupos: manual, semi-automática e automática [Rui et al., 1997].

Anotação Manual

Anotação Manual é a técnica mais simples, consiste no reconhecimento e na descrição dos elementos presentes em uma imagem por parte do usuário. A técnica não provê nenhum mecanismo de reconhecimento automático dos elementos.

A aplicação da Anotação Manual é dada mediante uma interface entre o usuário e as imagens de interesse. Os elementos específicos da imagem serão descritos através de um conjunto de palavras-chave que serão armazenadas em uma base de dados.

A técnica permite maior flexibilidade ao usuário de identificar livremente os elementos, porém é uma técnica lenta, outro problema presente nesta técnica é que usuários diferentes, podem ter diferentes percepções na descrição de uma mesma imagem.

Anotação Automática

A técnica mais abordada em estudos científicos e grupos de pesquisa e desenvolvimento é a Anotação Automática. Esta técnica surgiu como uma alternativa para realizar o trabalho de Anotação de Imagem sem a necessidade de interação com o usuário. Utilizando-se de algoritmos matemáticos complicados, torna-se possível a identificação dos elementos presentes em uma imagem e a posterior relação desses elementos com palavras-chave, armazenados em uma base de dados que descreve tais elementos com exatidão.

Assim, a aplicação da técnica da Anotação Automática elimina a descrição manual dos elementos, realizando este procedimento automatizado em que os elementos identificados

são descritos através de um conjunto de palavras-chave armazenados em uma base de dados. Entretanto, essa técnica é muito suscetível a erros de reconhecimento e descrição dos elementos de uma imagem. No caso de falha no processo de reconhecimento dos elementos, não existem meios para o usuário eliminar ou modificar os elementos incorretos, sendo isso uma grande desvantagem do método.

Anotação Semi-Automática

Uma outra técnica existente é denominada de Anotação Semi-automática, que consiste na junção de mecanismos de Anotação Automática com Anotação Manual, isto é, a técnica realiza um processo automático de reconhecimento dos elementos presentes em uma imagem e após o reconhecimento permite que o usuário insira e/ou modifique os elementos identificados.

Embora semelhante ao processo de Anotação Manual, no que diz respeito ao processo de anotação através de uma interface gráfica, difere daquela no que tange ao auxílio no reconhecimento dos elementos presentes em uma imagem, que serão descritas através de um conjunto de palavras-chave e armazenadas em uma base de dados.

O reconhecimento automático dos elementos é dado mediante a recuperação de imagens baseada no conteúdo visual onde a identificação dos elementos é realizada através do processo de comparação de similaridades entre as imagens.

Essa técnica é suscetível a erros de reconhecimento e descrição dos elementos de uma imagem, pois a imagem pode conter distorções diversas, influenciando de forma negativa o reconhecimento dos elementos na imagem. Além disso, para a descrição dos elementos reconhecidos, existe a necessidade de uma base de dados que contenha as palavras-chave relacionadas aos elementos, ou seja, o reconhecimento automático necessita da criação previamente, de uma base de dados na qual será consultada as similaridades existentes entre as imagens para identificação dos elementos.

Entretanto, essa técnica é mais flexível em relação às outras técnicas de anotação, uma vez que auxilia no processo de reconhecimento e descrição dos elementos de imagens, além de deixar à escolha do usuário o que identificar e alterar nos elementos reconhecidos dentro de uma imagem. Assim, em caso de falha na detecção dos elementos, o usuário pode eliminar ou modificar os elementos incorretos.

A subjetividade da percepção e anotações imprecisas podem causar perdas no processo de recuperação de imagens. Por causa do crescimento em larga escala das coleções de imagens, as dificuldades encontradas pelas Anotações Manual, Automática e Semi-automática tornaram-se mais agudas. A fim de superar estas dificuldades, volta-se para a recuperação de imagens baseada no conteúdo visual a partir de 1990. Ou seja, em vez de serem anotadas manualmente por chaves baseadas em texto, as imagens são indexadas pelo seu próprio conteúdo visual, tais como, cor, forma e textura [Rui et al., 1997].

2.2.2 Recuperação de Imagens baseada em Conteúdo Visual

Segundo [Rui et al., 1997], até 1990 a recuperação de imagens baseada em texto era a maneira mais utilizada. Embora muitos avanços tenham sido alcançados ao longo das pesquisas, tais como modelagem de dados, indexação multidimensional e avaliação de consultas, ainda existem duas grandes dificuldades, especialmente quanto ao tamanho das coleções de imagens. A primeira dificuldade é a quantidade de trabalho requerido na anotação manual das imagens. A segunda dificuldade, que é mais importante, resulta do rico conteúdo das imagens e a subjetividade da percepção humana.

A subjetividade da percepção humana e anotações imprecisas, podem causar perdas no processo de recuperação. O processo de consulta por similaridades entre características visuais da imagem como cor, textura e forma, tem como objetivo superar esses problemas, dado o fato de que as consultas são realizadas sobre as características das imagens e não sobre textos e anotações.

Processo de Extração de Características

Para o emprego do processo de Recuperação de Imagens por similaridade, existe a necessidade da identificação das características da imagem, dado o fato de que estas características são o meio de comparar duas ou mais imagens para identificação de semelhanças, [Ramos, 2004]. Logo, o primeiro passo para este processo de Recuperação de Imagens é a extração das características das imagens.

Conforme a Figura 2.9, extraída de [Bueno, 2002], o processo de extração de características de imagens consiste no mapeamento das propriedades relevantes de uma imagem. Em outras palavras, a imagem original é analisada por um motor de extração de características que obtém as principais propriedades de uma imagem e as armazena em um espaço de características.



Figura 2.9: Processo de Extração de Características.

Para a extração das características das imagens, essas podem ser segmentadas em

regiões utilizando-se algoritmos que particionam a imagem em regiões correspondentes às áreas de interesse do ponto de vista da aplicação. Uma região é definida como um conjunto de pixels contíguos, com espalhamento bidimensional, os quais apresentam uniformidade em relação a determinado atributo. Área, forma, elementos estatísticos e textura são exemplos de atributos que podem ser obtidos e utilizados na análise da imagem.

Após a imagem ser segmentada em regiões, resultante do processo de extração de características, inicia-se o processo de representação e descrição do conjunto de pixels segmentados em uma forma adequada de processamento [Bueno, 2002].

Segundo [Bueno, 2002], há duas formas de representar uma região: baseando-se nas características externas (suas fronteiras), ou nas características internas (os pixels contidos na região). Para análise das características morfológicas ou formas presentes na imagem, utilizam-se as características externas, e para análise das características refletidas como, por exemplo, cor e textura, utilizam-se as características internas.

A análise de imagem baseada no seu conteúdo tem seu processo modelado como uma hierarquia de abstrações, isto é, em um primeiro momento estão os pixels da imagem, com a informação sobre cores ou textura associada ao elemento. Em um segundo momento estão os atributos como bordas, cantos, linhas, curvas e regiões de cores. Em um terceiro momento, procura-se combinar e interpretar os atributos do momento anterior, relacionando os objetos que possuam tais características. Em um último momento aproxima-se do mapeamento humano, a partir do qual se busca compreender o relacionamento entre os objetos presentes na imagem [Traina, 2001].

2.2.3 *Bag-of-features*

A recuperação de imagens pode ser realizada por meio do reconhecimento de categorias de objetos. Nesse caso, um detector especializado em uma classe de interesse é aplicado a uma dada imagem, determinando se um objeto dessa classe está presente ou não. O paradigma consiste em coletar manualmente um grande conjunto de bons exemplares da categoria do objeto desejado (conjunto de treinamento), treinar um classificador nesse conjunto e então avaliar o modelo obtido em novas imagens (conjunto de teste). O método é denominado *bag-of-features*. O *bag-of-features* é uma analogia às representações *bag-of-words* usadas em recuperação de informação em textos, na qual cada documento é representado por um vetor das palavras que ocorrem nele [Nowak et al., 2006b].

A idéia básica desse método é descrever uma imagem como uma coleção de características locais em torno de regiões de pontos de interesse. Para uma representação compacta, um vocabulário visual é geralmente construído por meio do agrupamento (*clustering*) dessas características. Cada grupo (*cluster*) de pontos de interesse é tratado como uma palavra visual no vocabulário. Por meio do mapeamento dos pontos de interesse em uma imagem ao vocabulário visual, pode-se descrever a imagem como um vetor de ca-

racterísticas de acordo com a presença ou contagem de cada palavra visual [Jiang et al., 2007]. Esse vetor de características é o bag-of-features de uma imagem e um classificador pode ser aplicado aos vetores de características de determinado conjunto de imagens para determinar à qual categoria cada imagem pertence.

Novas abordagens, tanto para indexação das imagens quanto para o cálculo da similaridade em CBIR têm sido propostas na literatura. Os trabalhos com bag-of-features são uma alternativa. Essas abordagens tem como objetivo aproveitar a vantagem dos sistemas de recuperação baseado no conteúdo visual, que conseguem um bom desempenho na qualidade de recuperação juntamente com a vantagem dos sistemas de recuperação de imagens da Web, que apresentam um alto desempenho na velocidade de recuperação.

2.2.4 Considerações Finais

Como visto no decorrer desta seção, embora existam alguns problemas no processo de análise de imagens, por exemplo na segmentação, o processo de Recuperação de Imagens baseado em características visuais tende a ser mais eficiente e confiável em relação a baseada em Conteúdo Textual, uma vez que, independentemente da subjetividade humana quanto à compreensão e descrição de uma imagem, o processo de consulta por similaridade avalia as características das imagens.

Após a extração de características das imagens, é possível o reconhecimento dos elementos presentes nestas, etapa crucial para o processo de recuperação de imagens através de similaridades. Isso porque na consulta de uma imagem, o processo de recuperação por similaridades avalia as semelhanças entre a imagem consultada e as armazenadas em um repositório, onde um dos pontos avaliados são os elementos presentes em ambas as imagens. Assim, as que contenham elementos iguais à imagem consultada são mais relevantes.

Considerando os problemas descritos com os sistemas de recuperação de imagens, este trabalho propõe uma abordagem para recuperação de imagens, visando não somente o ganho no desempenho do processamento das consultas, mas também no ganho da precisão no conjunto de imagens recuperadas, baseando-se em técnicas originais de recuperação de informação textual que são discutidas na próxima seção.

2.3 Recuperação de Informação em Textos

Recuperação de informação é uma subárea da Ciência da Computação que estuda o armazenamento e recuperação automática de documentos, que são objetos de dados, geralmente textos. Um sistema de Recuperação de Informação (SRI) pode ser estruturado conforme a Figura 2.10 [Cardoso, 2000].

O processo de indexação basicamente extrai todas as palavras de cada documento e

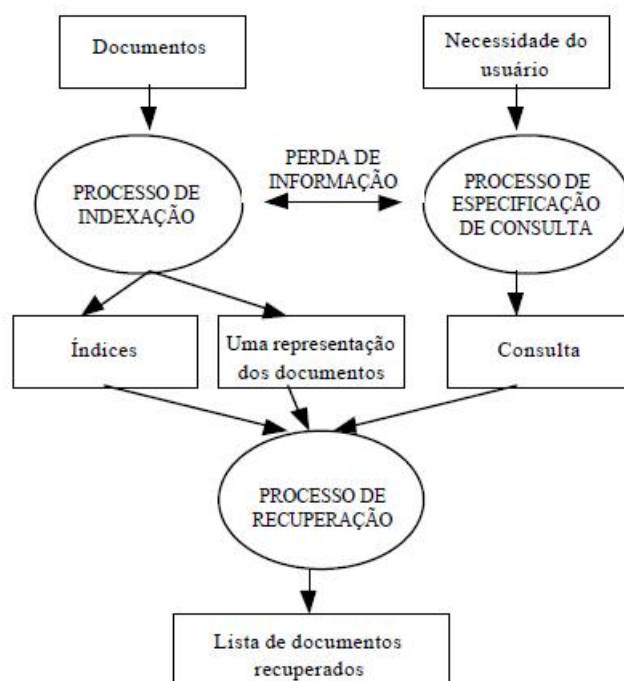


Figura 2.10: Componentes de um sistema de recuperação de informação [Cardoso, 2000].

armazena dados estatísticos sobre a ocorrência das palavras nos documentos, ou seja, é a criação de estruturas de dados associados à parte textual dos documentos, por exemplo, as estruturas de arranjos de sufixos (PAT *arrays*) e arquivos invertidos, descritas em [Frakes and Baeza-Yates, 1992]. A estrutura de dados Arquivo Invertido é a principal estrutura utilizada neste trabalho e é apresentada em detalhes na seção 2.3.1.

O processo de especificação da consulta geralmente é uma tarefa difícil. Há frequentemente uma distância semântica entre a real necessidade do usuário e o que ele expressa na consulta formulada. Essa distância é gerada pelo limitado conhecimento do usuário sobre o universo de pesquisa e pelo formalismo da linguagem de consulta, [Cardoso, 2000].

Finalmente, o processo de recuperação consiste na geração de uma lista de documentos recuperados para responder à consulta formulada pelo usuário. Uma das limitações desse processo é a perda de informações gerada por processos de indexação e especificação de consulta.

Os modelos clássicos de recuperação de informação apresentam estratégias de busca de documentos relevantes para uma consulta. Ou seja, são formas sistemáticas que definem como um sistema faz o cálculo da relevância dos documentos. Tanto a consulta feita pelo usuário, quanto os documentos que compõem a coleção a ser pesquisada, são representados pelos seus termos.

Um dos principais problemas de sistemas de recuperação de informação é prever quais documentos são relevantes e quais não são. Esta decisão depende do algoritmo de

cálculo de relevância utilizado, considerado como o núcleo do sistema, no qual é definida uma função de ordenação da lista de documentos recuperados, resultado final do processo de recuperação 2.10. Esta função é uma relação que associa um número real ao par formado pela consulta e um item de resposta. Tal função depende do modelo utilizado que em geral define uma ordem entre os itens de resposta de modo que os itens que estejam no topo sejam considerados mais relevantes.

Os modelos clássicos de RI são: booleano, vetorial e probabilístico. Neste trabalho foi utilizado o modelo vetorial devido a sua simplicidade, a facilidade que ele provê de se computar similaridades com eficiência e o fato de que o modelo se comporta bem com coleções genéricas, [Cardoso, 2000]. A seção 2.3.2 detalha o modelo vetorial.

2.3.1 Arquivo Invertido

A primeira solução sistemática para o problema de recuperação de informação em uma grande coleção de dados, foi desenvolvida há cerca de 4000 anos por bibliotecários, que passaram a organizar os livros catalogando-os por autor e título. Uma das soluções que surgiu, a construção dos índices, vem sendo utilizada até hoje. Contudo agora os índices são construídos automaticamente. Existem várias estruturas de índice, porém a mais empregada em recuperação de informação é a estrutura de arquivos invertidos [Santos, 2003], por ser a forma mais intuitiva de modelar o acesso aos dados. Segundo [Faloutsos, 1985] e [Salton and McGill, 1986] arquivo invertido é um mecanismo para indexar coleções textuais utilizando palavras. A função principal de um arquivo invertido é retornar uma lista de documentos na qual ocorre um termo, tendo como idéia básica o armanejamento do mapeamento inverso de termos para documentos.

A estrutura do arquivo invertido é composta por dois elementos: o vocabulário, que consiste no conjunto de todos os termos distintos que ocorrem na coleção; e as listas invertidas, que são listas contendo todos os documentos da coleção nos quais o termo ocorre, como ilustrado na Figura 2.11. A primeira tabela da Figura 2.11 ilustra uma coleção de documentos e os termos que aparecem nos respectivos documentos. A segunda tabela da Figura 2.11 mostra o arquivo invertido correspondente à coleção, sendo que para cada termo, há uma lista de documentos no qual o termo aparece.

A lista invertida contém a relação de documentos onde a palavra aparece. A informação presente em cada lista invertida é uma lista de apontadores para os documentos onde a palavra ocorre e, geralmente, são adicionadas outras informações que podem ser úteis durante o processamento de consultas, como por exemplo, a frequência do termo no documento. Esse valor é chamado na literatura [Baeza-Yates and Ribeiro-Neto, 1999] de *Term-Frequency* (tf). Além da frequência, a lista de posições de cada termo em cada documento pode ser também adicionada a lista invertida, para que seja possível processar consultas por frases exatas ou consultas por proximidade.

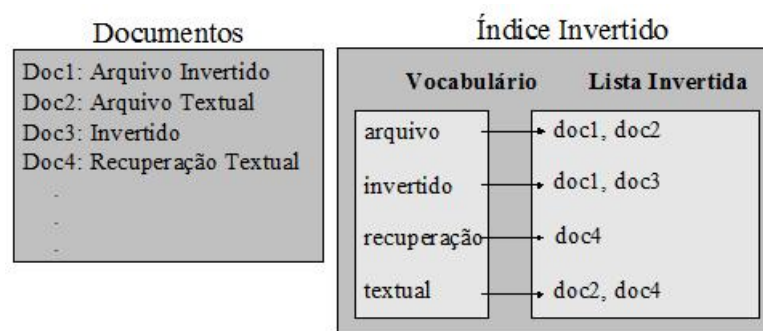


Figura 2.11: Exemplo arquivo invertido.

Um arquivo invertido é construído por meio do processamento de cada documento da coleção. Este processamento é chamado indexação e é responsável por identificar todos os termos existentes na coleção, inserindo-os no vocabulário juntamente com o ponteiro para a lista invertida.

As principais tarefas para a construção de um índice invertido serão descritas a seguir. Esta explicação é baseada em [Manning, 2007].

Devidamente escolhida a coleção de documentos a ser indexada, é necessário assumir que cada documento tenha um identificador do documento (ID) único na coleção. A cada documento encontrado é atribuído um número serial. Os dados inseridos no processo de indexação são uma lista de termos contidos no documento, que podem também ser vistos como uma lista com dois dados possíveis, que são o termo e o ID do documento, conforme ilustrado na Figura 2.12.

Na primeira coluna da Figura 2.12 os termos são armazenados na ordem em que eles ocorrem no documento juntamente com o ID do documento correspondente. Termos que ocorrem mais de uma vez em um mesmo documento, nesta fase, aparecem quantas vezes eles estiverem presentes no documento (coluna 1 da Figura 2.12). A seguir, os termos são ordenados por ordem alfabética (coluna 2 da Figura 2.12). Para finalizar esta etapa de construção, as várias ocorrências do mesmo termo, a partir do mesmo documento são agrupadas, e uma coluna extra é adicionada para gravar a frequência do termo no documento (coluna 3 da Figura 2.12).

Finalmente as instâncias do mesmo termo são agrupadas e o resultado é dividido em um dicionário e uma lista invertida, como ilustrado na Figura 2.13. A lista invertida da Figura 2.13 é representada pelo par (documento, frequência do termo). Considerando-se que um termo geralmente está presente em diversos documentos, este tipo de estrutura reduz o número de comparações necessárias para o processamento de consultas, pois somente os documentos que contiverem os termos da consulta serão examinados. Sem dúvidas, a estrutura de índice invertido é a melhor e mais eficiente estrutura para suportar uma busca de texto *ad-hoc* [Frakes and Baeza-Yates, 1992].

DOC 1
O homem medíocre espera tudo dos outros. O homem é covarde.

DOC 2
O homem de bem exige tudo de si próprio. O homem é corajoso.

Coluna 1		Coluna 2		Coluna 3		
Termo	DocID	Termo	DocID	Termo	DocID	Freq. Termo
o	1	bem	2	bem	2	1
homem	1	corajoso	2	corajoso	2	1
medíocre	1	covarde	1	covarde	1	1
espera	1	de	2	de	2	2
tudo	1	de	2	dos	1	1
dos	1	dos	1	é	1	1
outros	1	é	1	é	2	1
o	1	é	2	espera	1	1
homem	1	espera	1	exige	2	1
é	1	exige	2	homem	1	2
covarde	1	homem	1	homem	2	2
o	2	homem	1	medíocre	1	1
homem	2	homem	2	o	1	2
de	2	homem	2	o	2	2
bem	2	medíocre	1	outros	1	1
exige	2	o	1	próprio	2	1
tudo	2	o	1	si	2	1
de	2	o	2	tudo	1	1
si	2	o	2	tudo	2	1
próprio	2	outros	1			
o	2	próprio	2			
homem	2	si	2			
é	2	tudo	1			
corajoso	2	tudo	2			

Figura 2.12: Indexação dos termos, parte um.

Termo	DocID	Freq. Termo	Dicionário	Lista Invertida
bem	2	1	bem	1,1
corajoso	2	1	corajoso	2,1
covarde	1	1	covarde	1,1
de	2	2	de	2,2
dos	1	1	dos	1,2
é	1	1	é	1,2 → 2,2
é	2	1	espera	1,2
espera	1	1	exige	2,2
exige	2	1	homem	1,2 → 2,2
homem	1	2	medíocre	1,2
homem	2	2	o	1,2 → 2,2
medíocre	1	1	outros	1,2
o	1	2	próprio	2,2
o	2	2	si	2,2
outros	1	1	tudo	1,2 → 2,2
próprio	2	1		
si	2	1		
tudo	1	1		
tudo	2	1		

Figura 2.13: Indexação dos termos, parte dois.

Quando uma consulta contendo mais de um termo é submetida a um sistema de recuperação de informação, o processo de busca é responsável por localizar a lista invertida para cada termo da consulta. Dependendo dos operadores booleanos utilizados na formulação da consulta, a resposta será gerada mediante a interseção ou união dos conjuntos de documentos existentes em cada lista, sendo esse um processo rápido.

Considerando o arquivo invertido da Figura 2.11, supondo que desejamos executar uma consulta que encontre todos os documentos que contenham o termo “arquivo” ou o termo “invertido”. O primeiro passo é localizar na lista invertida de todos os documentos que contêm o termo “arquivo” e a de todos documentos do termo “invertido”. Como na consulta o operador utilizado foi união, uma lista com todos os documentos distintos que ocorrem um dos termos ou ambos é encontrado através da união dos conjuntos de documentos dos dois termos, Figura 2.14.

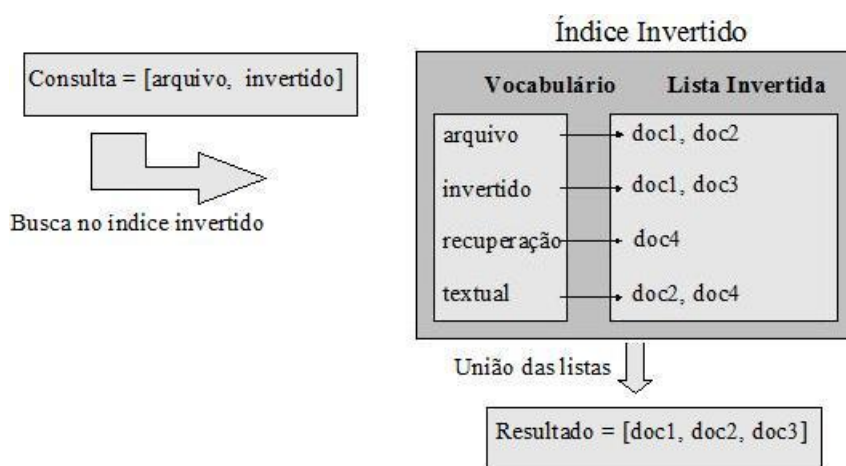


Figura 2.14: Exemplo consulta utilizando arquivo invertido.

Devido a esse rápido tempo de resposta obtido na técnica de busca textual, estudos vêm sendo feitos para a utilização da técnica textual em recuperação de imagens baseada em conteúdo, pois esta, por sua vez, quando usando uma medida de similaridade baseada na distância euclidiana, percorre todo o banco de imagens para o cálculo da similaridade, o que gera um alto custo computacional e longo tempo de resposta.

2.3.2 Modelo Vetorial

O modelo vetorial foi desenvolvido por Gerard Salton, para ser utilizado num SRI chamado SMART [Salton and McGill, 1986] e [Baeza-Yates and Ribeiro-Neto, 1999].

No modelo vetorial, cada documento é representado por um vetor com n termos, no qual n é o número total de termo distintos encontrados em todos os documentos de uma

coleção. Cada termo possui um valor associado que indica o grau de importância (peso - *weight*) do termo no documento.

Cada elemento do vetor de termos é considerado uma coordenada dimensional. Assim, os documentos podem ser colocados em um espaço Euclidiano de n dimensões (onde n é o número de termos) e a posição do documento em cada dimensão é dada pelo seu peso. A Figura 2.15 ilustra um documento tridimensional com três termos (termo1, termo2 e termo3).

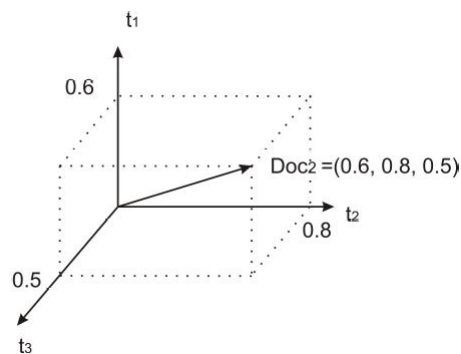


Figura 2.15: Exemplo modelo vetorial com três termos de indexação.

Peso

O cálculo dos pesos dos termos pode variar em função do modelo utilizado pelo sistema de recuperação de informações, porém a forma mais utilizada se baseia em dois fatores. O primeiro fator está relacionado à frequência do termo, ou seja, ao número de vezes que este termo aparece no documento. Este fator é denominado *tf* (*term frequency*). O segundo valor está relacionado à frequência dos documentos, que corresponde ao número de documentos que possuem o termo. Este fator é denominado *df* (*document frequency*). Estas duas medidas têm suas utilidades no cálculo da função de ordenação. Para o *tf*, quanto maior a frequência de um termo maior será o seu peso de ordenação. O *df* é útil para indicar o valor da função de ordenação dos termos que ocorrem em muitos documentos. Esses tipos de termos são considerados poucos úteis, pois não servem para distinguir se um documento é relevante ou não. Assim sendo, o fator *df* contribui inversamente para o cálculo do peso, podendo ser caracterizado como o inverso da frequência do documento, recebendo assim a denominação de *idf* (*inverse document frequency*).

A frequência inversa do documento (*idf*) de um termo t é obtida pela equação 2.7:

$$idf_t = \log \frac{N}{df_t} \quad (2.7)$$

na qual N é o número total de documentos da coleção, e df_t é o número de documentos da coleção que o termo t ocorre. Com isso, o *idf* de um termo com frequência baixa é

elevado, enquanto que o idf de termos com frequência alta decresce em escala logarítmica.

A medida $tf_{t,d}$ reflete características intra-documentos e idf_t dá uma medida de distinções inter-documentos. Pesos da abordagem $tf-idf_{t,d}$ são baseados no produto $tf_{t,d} \times idf_t$. Estes servem para produzir um peso composto para cada termo em cada documento. Na abordagem $tf-idf_{t,d}$, o peso do termo t no documento d é dado por

$$w_{d,t} = tf_{d,t} \times \log \frac{N}{df_t} \quad (2.8)$$

A consulta do usuário também é representada por um vetor. Desta forma, os vetores dos documentos podem ser comparados com o vetor da consulta e o grau de similaridade entre cada um deles pode ser identificado. Os documentos mais similares (mais próximos no espaço) à consulta são considerados relevantes para o usuário e retornados como resposta.

A Figura 2.16 mostra exemplo de documentos e consultas no espaço vetorial.

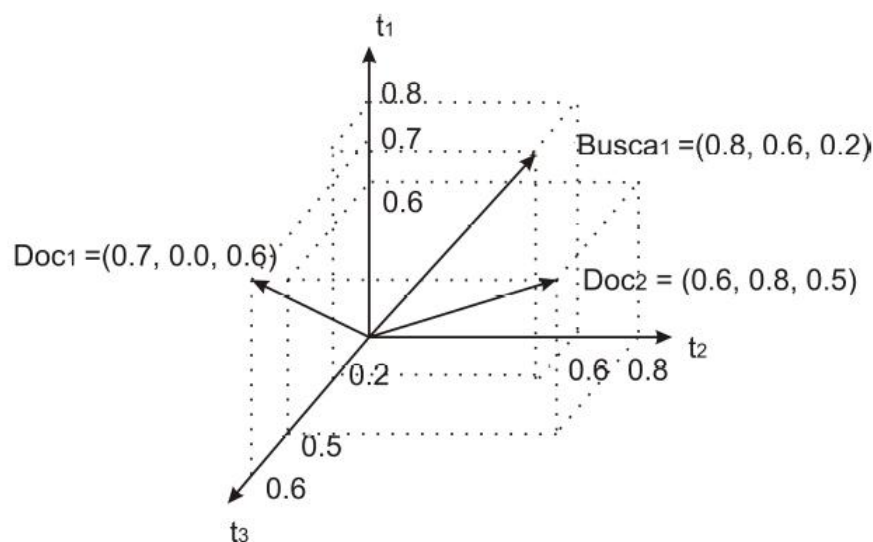


Figura 2.16: Exemplo modelo vetorial, vetores dos documentos e das consultas [Santos, 2003].

Similaridade

Uma vez que as características de uma imagem tenham sido extraídas e armazenadas em vetores de característica, faz-se necessário medidas que comparam a similaridade entre vetores de característica. Medidas de similaridade são normalmente baseadas em medidas de distância.

Uma das formas de calcular a proximidade entre os vetores é analisar o ângulo entre estes vetores. Geralmente essa correlação é quantificada pelo co-seno do ângulo entre esse dois vetores. Em um espaço vetorial de dimensão n , a similaridade (sim) entre uma

consulta q e um documento d_j , representada pelos vetores $\vec{q}$ e $\vec{d}_j$ é dada pelo cosseno do ângulo formado por estes vetores, através da seguinte relação [Manning, 2007]:

$$\text{sim}(q, d_j) = \frac{\sum_{i=1}^n w_{iq} w_{ij}}{\sqrt{\sum_{i=1}^n w_{iq}^2} \sqrt{\sum_{i=1}^n w_{ij}^2}} \quad (2.9)$$

na qual, $\vec{q}$ é o vetor de termos da consulta q ; $\vec{d}_j$ é o vetor de termos do documento d_j ; w_{iq} é o peso do termo i da consulta q e w_{ij} é o peso do termo i no documento d_j .

Depois de terem sido calculados os graus de similaridade, é possível montar uma lista ordenada (*ranking*) de todos os documentos e seus respectivos graus de relevância à consulta.

A escolha de uma função de distância adequada é de extrema importância. Há alternativas ao cosseno como função de similaridade, um bom estudo sobre essas alternativas está em [Zobel and Moffat, 1998]. No entanto, não existe uma formalização definida e a função é identificada por meio de heurísticas dependente das características dos dados [Zobel and Moffat, 1998].

Uma das vantagens do modelo vetorial está no fato de ser uma estratégia que permite a recuperação de informações com base em combinações parciais dos documentos em relação a uma consulta. Isto permite que documentos que estão parcialmente relacionados com uma consulta também sejam considerados. Outra vantagem está no fato de que o modelo vetorial é uma estratégia que permite classificar os documentos relevantes em relação a uma consulta, estabelecendo um *ranking* de similaridade. Devido a estas características o modelo foi escolhido para ser utilizado no desenvolvimento deste trabalho.

Lista Invertida Aplicada ao Cálculo do Cosseno

A estratégia utilizada para o cálculo da medida de similaridade é uma extensão da implementação usual para o modelo vetorial, conforme descrito no algoritmo da Figura 2.17.

Para o algoritmo da figura 2.17, consideramos que a estrutura lista invertida já esteja construída. Inicialmente se cria uma estrutura para o armazenamento dos valores (acumuladores) para as similaridades parciais dos documentos. Posteriormente, para cada termo da consulta, recupera-se toda a sua lista invertida e para cada entrada nesta lista calcula-se a sua contribuição para a similaridade dos documentos (medida do cosseno) que é adicionada ao acumulador do documento. Em seguida, os valores dos acumuladores são normalizados no passo 3, usando a divisão pela norma, definida na equação 2.11. O passo final consiste em selecionar os r maiores valores para os acumuladores e retornar os documentos correspondentes.

A contribuição de um termo t para a similaridade entre uma consulta q e um docu-

1. Crie e inicialize uma estrutura de acumuladores (A).
2. Para cada termo t da consulta faça.
 - a. recupere a lista invertida do termo t do disco;
 - b. para cada par $[d, f_{d,t}]$ na lista invertida t faça:
 Se $A_d \notin A$ então $A_d = 0$;
 $A_d = A_d + w_{q,t} * w_{d,t} \text{ (sim}_{q,d,t}\text{)}$;
3. Divida cada acumulador A_d por W_d ;
4. Recupere os r maiores valores para os acumuladores e retorne os documentos relevantes;

Figura 2.17: Algoritmo para calcular o cosseno usando índice invertido.

mento d é dada por:

$$\text{sim}_{q,d,t} = w_{q,t} \times w_{d,t} \quad (2.10)$$

no qual, $w_{q,t}$ é o peso do termo t na consulta q e $w_{d,t}$ o peso do termo t no documento d .

A norma do documento d é dada por:

$$W_d = \sqrt{\sum w_{d,t}^2} \quad (2.11)$$

no qual, $w_{d,t}$ é dado pela equação 2.12

$$w_{d,t} = tf_{d,t} \times \log \frac{N}{df_t} \quad (2.12)$$

A Figura 2.18 ilustra o algoritmo utilizado no cálculo do cosseno sem utilizar a lista invertida. Para cada documento da coleção recupera-se a lista de termos e para cada um desses termos que esteja presente na consulta calcula-se a sua contribuição para a similaridade dos documentos (medida do cosseno). O passo final consiste em selecionar os r maiores valores para os acumuladores e retornar os documentos correspondentes. No algoritmo é possível observar que todos os documentos da coleção são recuperados no passo 2.

2.3.3 Métricas de Avaliação

Os sistemas de Recuperação de Informação, em geral, são avaliados sob os aspectos de capacidade de desempenhar as funcionalidades especificadas, tempo de resposta a uma solicitação, espaço utilizado em memória e eficácia de recuperação. A eficácia de recuperação está relacionada à capacidade do sistema retornar somente os documentos

1. Crie e inicialize uma estrutura de acumuladores (A).
2. Para cada documento d
 - a. recupere a lista de termos de d ;
 - b. $A_d = 0$;
 - c. para todo termo $t \in q$ e $t \in d$

$$A_d = A_d + w_{q,t} * w_{d,t} \text{ (sim}_{q,d,t}\text{)};$$
3. Divida cada acumulador A_d por W_d ;
4. Recupere os r maiores valores para os acumuladores e retorne os documentos relevantes;

Figura 2.18: Algoritmo para calcular o cosseno.

relevantes existentes na coleção para uma dada consulta. Um documento é considerado relevante para uma dada consulta, se ele atende à necessidade de informação do usuário, que é expressa por meio de uma consulta.

Para comparar a eficácia de recuperação de diferentes configurações de sistemas de RI, diversas medidas podem ser utilizadas, como por exemplo, Precisão, Revocação, Medida F, Medida E e Fallout [Cardoso et al., 2007]. A decisão de quais as medidas utilizar numa avaliação depende da aplicação, havendo sempre discussões sobre a confiabilidade de tais medidas. O artigo [Gwizdka et al., 1999], por exemplo, discute como avaliar máquinas de busca.

A avaliação do sistema de RI, proposto neste trabalho, está centrada na eficácia da recuperação medida por meio da precisão e no desempenho em número de operações necessárias para efetuar a busca. As principais métricas de avaliação da eficácia são: precisão e revocação, definidas a seguir.

Precisão

A precisão de um conjunto de documentos recuperados para uma consulta q é a razão entre a quantidade de documentos relevantes recuperados divididos pelo número total de documentos recuperados isto é:

$$Pr(q) = \frac{DRR_{(q)}}{DR_{(q)}} \quad (2.13)$$

em que $DRR_{(q)}$ é o número de documentos relevantes recuperadas para a consulta q e $DR_{(q)}$ o total de documentos recuperados para a consulta q .

Por exemplo, se para uma busca realizada por um SRI forem recuperados 20 docu-

mentos e destes apenas 10 forem realmente relevantes, a precisão do sistema é 50%.

Revocação

Uma outra medida padrão de efetividade de um sistema de recuperação é a revocação. O valor da revocação é a razão entre número total de documentos relevantes recuperados pelo sistema e o número total de documentos que podem ser considerados relevantes na coleção. Podem existir muitos documentos na coleção que o usuário considere relevantes, mas só alguns são recuperados pelo sistema.

$$Re(q) = \frac{DRR_{(q)}}{DRC_{(q)}} \quad (2.14)$$

em que $DRR_{(q)}$ é o número de documentos relevantes recuperados para a consulta q e $DRC_{(q)}$ o total de documentos da coleção considerados relevantes para a consulta q .

Para um sistema ser efetivo tanto a precisão quanto a revocação devem ser maximizadas. Por exemplo, para maximizar a revocação, basta recuperar todos os documentos indexados, entretanto isso levaria a uma baixa precisão. Enquanto para maximizar ambas as medidas o sistema deve recuperar todos os documentos relevantes, e não recuperar documentos irrelevantes.

Cabe destacar que se o SRI retorna um *ranking* de documentos, as medidas Precisão e Revocação podem ser calculadas em diversas posições do ranking.

Precisão em Rankings

Embora os sistemas de recuperação retornem ao usuário uma grande quantidade de documentos, geralmente, o usuário observa somente as primeiras posições do *ranking*. Assim sendo, a medida de precisão a seguir visa avaliar os documentos que efetivamente são observados pelo usuário. A Precisão de uma consulta q em uma posição X do *ranking*, aqui chamada de Precisão- X é dada pela seguinte relação:

$$Pr(q, x) = \frac{DRR_{(q,x)}}{X} \quad (2.15)$$

onde $DRR_{(q,x)}$ é a quantidade de documentos relevantes encontrados até a posição X do *ranking*.

Por exemplo, a Precisão-5, $Pr(q,5)$ ou sucintamente $P5$, é o cálculo da precisão na quinta posição do *ranking* da consulta q , ou seja, é o número de documentos relevantes encontrados até a quinta posição do ranking dividido por cinco.

A curva de precisão x revocação que resulta da média dos resultados para várias consultas e é geralmente utilizada para comparar a performance na recuperação de algoritmos distintos [Manning, 2007].

A curva de precisão versus revocação é geralmente baseada em 11 níveis de revocação,

que são 0%, 10%, 20%, ..., 100%, tal curva é conhecida como 11-pt. Uma vez que os níveis de revocação para cada consulta podem ser distintos dos 11 níveis de revocação, frequentemente é necessário um processo de interpolação. Os cálculos de precisão nos 11 níveis de revocação são interpolados como se segue [Manning, 2007].

Seja r_j , $j \in \{0, 1, 2, \dots, 10\}$, uma referência para o j -ésimo nível de revocação, a precisão interpolada é dada pela seguinte expressão

$$P(r_j) = \max_{r_j \leq r \leq r_{j+1}} P(r) \quad (2.16)$$

na qual indica que a precisão interpolada no j -ésimo nível de revocação é a máxima precisão conhecida em qualquer nível de revocação entre o j -ésimo nível de revocação e o $(j+1)$ -ésimo nível de revocação [Manning, 2007].

2.3.4 Considerações Finais

Nesta seção foram apresentadas as principais técnicas envolvidas no processo de recuperação de informação. Foram descritos: a técnica arquivo invertido para representação de dados, o modelo vetorial de recuperação e algumas medidas de avaliação dos sistemas de recuperação.

Capítulo 3

Proposta de um Método para CBIR Baseado em Índice Invertido com Partições do Domínio

3.1 Introdução

Os sistemas clássicos de CBIR tendem a ser lentos para grandes coleções de imagens, por utilizarem, na maioria dos casos, a distância Euclidiana no cálculo da similaridade em processo de recuperação, que inclui o cálculo da distância entre o vetor da imagem de consulta e todos os vetores das imagens do banco de dados, tornando a recuperação computacionalmente cara.

Neste capítulo é apresentada uma proposta de CBIR chamada CBIR-Índice-Particionado que estende o trabalho de [Matos et al., 2008a] e [Matos et al., 2008b] chamado CBIR-Índice. Com o objetivo de obter um ganho de desempenho e qualidade na recuperação de imagens. A extensão consiste em particionamento do domínio das imagens para a extração das características, ou seja, as imagens foram particionadas em diversas regiões e as características foram extraídas de cada região. Além do particionamento, a extensão também acrescentou duas novas características, forma e textura.

Inicialmente é descrita a proposta de [Matos et al., 2008a] e [Matos et al., 2008b], que utiliza o Índice Invertido para recuperação de imagens baseada em conteúdo. Para avaliar a proposta foram desenvolvidos os sistemas CBIR-I e CBIR-Índice. O primeiro sistema, usado como base de comparação, utiliza técnicas de extração de características e cálculo de similaridade, normalmente utilizadas em CBIR, a saber, momentos de cor e distância Euclidiana respectivamente. O outro sistema, o CBIR-Índice, indexa as imagens utilizando-se Índice Invertido. A similaridade desse último sistema é medida por meio do cálculo do cosseno entre vetores representantes da imagem de consulta e das imagens do banco de dados. Ambos os sistemas serão melhor descritos neste capítulo.

3.2 CBIR-Indice

O sistema CBIR-Indice é uma adaptação do sistema CBIR tradicional (CBIR-I), para que possa ser utilizado o índice invertido.

A idéia do CBIR-Indice é definir termos de indexação associados a faixa de valores extraídos da imagem com base em características de baixo nível, especificamente momentos de cor. Este caracteriza as imagens em termos da distribuição das cores nos *pixels* da imagem. São obtidos por três medidas estatísticas: média, desvio padrão e obliquidade, seção 2.1.1. Os momentos de cor são derivados do espaço de cor HSV, correspondendo aos canais matiz, saturação e intensidade, seção 2.1.1.

Em resumo, as imagens foram convertidas do espaço de cor RGB para o espaço de cor HSV. Então, foram calculados os três momentos acima para cada canal H, S e V. Com isto, foi obtido o vetor de característica da imagem com nove posições distribuídas conforme Tabela 3.1.

Posição do Vetor	Descrição
1º	Corresponde a Média da Matiz(H)
2º	Corresponde ao Desvio Padrão da Matiz(H)
3º	Corresponde a Inclinação da Matiz(H)
4º	Corresponde a Média da Saturação(S)
5º	Corresponde ao Desvio Padrão da Saturação(S)
6º	Corresponde a Inclinação da Saturação(S)
7º	Corresponde a Média da Intensidade(V)
8º	Corresponde ao Desvio Padrão da Matiz(V)
9º	Corresponde a Inclinação da Intensidade(V)

Tabela 3.1: Descrição do vetor de característica resultante da extração de característica da imagem usando os três momentos de cor do espaço HSV.

Uma análise estatísticas de valores da tabela 3.1 obtidos em uma ampla variedade de imagens constata que a distribuição dos valores em cada posição do vetor é semelhante a uma distribuição normal da estatística e probabilidade. Com isso, as regras e propriedades desta distribuição foram aplicadas às características para obterem a classificação destas em 10 faixas, conforme a Figura 3.1

Desta forma, cada intervalo tem a seguinte porcentagem aproximada de imagens:

- 15% dos casos estão incluídos no intervalo $(\bar{X} - 0.39s, \bar{X}]$ (correspondente a faixa E da Figura 3.1) e 15% está no intervalo $(\bar{X}, \bar{X} + 0.39s]$ (correspondente a faixa F da Figura 3.1);
- 13% dos casos estão incluídos no intervalo $(\bar{X} - 0.78s, \bar{X} - 0.39s]$ (faixa D) e 13% está no intervalo $(\bar{X} + 0.39s, \bar{X} + 0.78s]$ (faixa G)

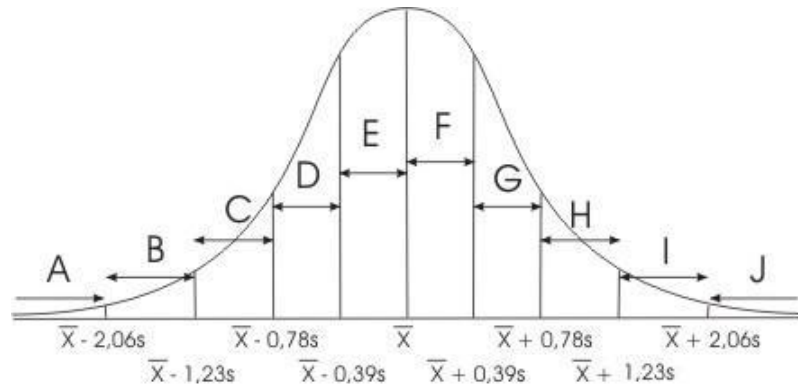


Figura 3.1: Distribuição adotada para classificação das faixas. Onde $\bar{X}$ é a média e s o desvio padrão.

- 11% dos casos estão incluídos no intervalo $(\bar{X} - 1,23s, \bar{X} - 0,78s]$ (faixa C) e 11% está no intervalo $(\bar{X} + 0,78s, \bar{X} + 1,23s]$ (faixa H)
- 09% dos casos estão incluídos no intervalo $(\bar{X} - 2,06s, \bar{X} - 1,23s]$ (faixa B) e 09% está no intervalo $(\bar{X} + 1,23s, \bar{X} + 2,06s]$ (faixa I)
- 02% dos casos estão incluídos no intervalo $(0, \bar{X} - 2,06s]$ (faixa A) e 02% está no intervalo $(\bar{X} + 2,06s, 1]$ (faixa J)

O vetor de características da imagem irá representar as 9 características da imagem (Tabela 3.1), e estas características estão agrupadas em faixas, figura 3.1. Cada índice do vetor de características da imagem foi dividido em 10 faixas de valores que variam de A até J de acordo com o seu valor.

Para que as técnicas de recuperação textual, tais como, índice invertido e modelo vetorial, possam ser utilizadas, são necessários termos de indexação para coleção. Definiu-se então as seguintes analogias entre recuperação de imagens e recuperação textual: **i)** imagens correspondem aos documentos ; **ii)** as faixas encontradas através da análise dos vetores de características da Figura 3.1 são os termos; **iii)** imagens exemplos ou imagens de consulta são as consultas ou *query* na recuperação textual.

3.2.1 Indexação das Imagens

No CBIR-Indice após a extração das características, é gerado um banco de características, em que consta o nome da imagem e os nove valores correspondentes a média, desvio padrão e obliquidade dos canais de cores H, S e V. Posteriormente os nove valores são mapeados para faixas de acordo com a sua ocorrência determinada pelas propriedades da distribuição normal, Figura 3.1. A construção de um desses vetores pode ser visualizada pela Tabela 3.2. Supondo que os valores da Tabela 3.2 correspondam a primeira característica do vetor, que é a Média da Tonalidade (MH). Considerando as 10 faixas

descritas na seção anterior, o valor do MH será mapeado para um dos seguintes termos: (MHA, MHB, MHC, MHD, MHE, MHF, MHG, MHH, MHI, MHJ). No qual MH significa a média (M) dos *pixels* do canal de cor tonalidade (H) e A até J é o valor correspondente a faixa onde a características estiver presente.

De	a	Faixa	Termo
0	0.086547	A	MHA
0.086548	0.197300	B	MHB
0.197301	0.257346	C	MHC
0.257347	0.309386	D	MHD
0.309387	0.361427	E	MHE
0.361428	0.413467	F	MHF
0.413468	0.465507	G	MHG
0.465508	0.525554	H	MHH
0.525555	0.636306	I	MHI
0.636307	1	J	MHJ

Tabela 3.2: Construção do vetor de Características.

Considerando a Figura 3.2, o valor da primeira posição do vetor de característica é 0.317627. Analisando a Tabela 3.2, é possível notar que este valor corresponde a faixa E, sendo assim, esta característica será mapeada para o termo MHE ao invés de 0.317627.

MH	DH	IH	MS	DS	IS	MV	DV	IV
0.317627	0.267757	0.232397	0.391507	0.195728	0.119390	0.426673	0.242813	0.181251

Figura 3.2: Vetor de característica correspondente a imagem Africa1.jpg.

As outras oito posições do vetor de característica são construídas de forma análoga a esta. Portanto, o sistema CBIR-Índice terá um vetor de características com 90 posições, situações em que o valor for igual a 1 corresponde a presença do termo correspondente ao momento/canal/faixa e o valor igual a 0 a ausência do termo. Como toda imagem terá um vetor original com 9 características, o vetor mapeado será formado por 9 valores iguais a 1 e 81 valores iguais a 0. Assim, o vetor de característica para uma imagem exemplo é representado na Figura 3.3.

Portanto, os 90 termos criados pelo mapeamento de faixas são utilizados para indexação das imagens através da estrutura Índice Invertido descrito na seção 2.3.1. Como visto anteriormente, o arquivo invertido possui duas estruturas principais: uma estrutura de busca, chamada vocabulário, contendo todos os termos distintos existentes nas imagens indexadas, e, para cada termo, uma lista invertida que armazena os identificadores das imagens que contém o termo, como ilustrado na Figura 3.4

0	0	0	0	1	0	0	0	0	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ

0	0	0	1	0	0	0	0	0	0
DHA	DHB	DHC	DHD	DHE	DFH	DHG	DHH	DHI	DHJ

0	0	0	0	0	0	0	1	0	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ

0	0	0	0	0	1	0	0	0	0
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ

0	0	0	0	1	0	0	0	0	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ

0	0	0	0	0	1	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ

0	0	0	0	1	0	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ

0	0	0	1	0	0	0	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ

0	0	0	0	0	1	0	0	0	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

Figura 3.3: Vetor de característica gerado pelo sistema correspondente a uma imagem exemplo.

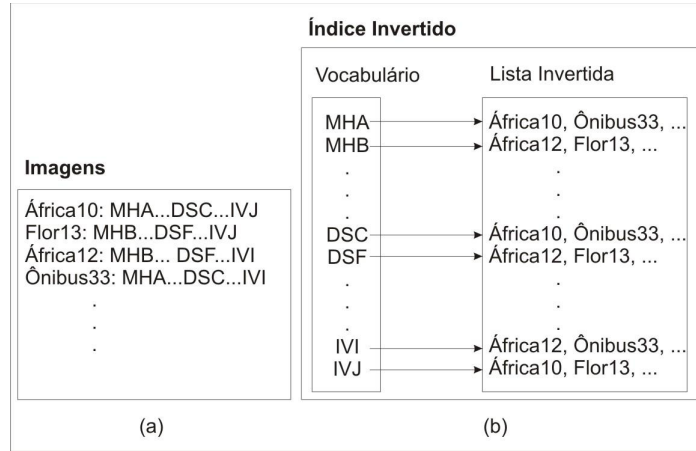


Figura 3.4: (a) Banco de dados de Característica (b) Estrutura de busca (vocabulário) e a lista invertida.

Consultas nos arquivos invertidos são feitas tomando-se a lista invertida correspondente aos termos procurados conforme o algoritmo da Figura 2.17. Por exemplo, caso a imagem de consulta tivesse os seguintes termos: MHA, DSC e IVJ, apenas as listas invertidas destes três termos seriam percorridas.

Em relação ao peso dos termos, a abordagem utilizada pelos autores foi o $tf - idf_{t,d}$, descrito na seção 2.3.2, o qual pode ser obtido pela expressão 2.8. Na imagem I_j , tf_{ij} é considerado o número de ocorrências do termo como binário, ou seja, se o termo i estiver presente na imagem I_j terá ocorrência igual a 1; caso contrário a ocorrência é 0. A

frequência na coleção é dada pelo idf, apresentado na seção 2.3.2. Logo o peso dos termos é dado pela expressão 3.1.

$$w_{ij} = tf_{ij} \times \log \frac{N}{df_i} \quad (3.1)$$

em que $tf_{ij} \in \{0, 1\}$, N é o número total de imagens da coleção e df_i é o número de imagens da coleção em que o termo i ocorre.

3.2.2 Cálculo da Similaridade

As imagens de consulta utilizadas pelo sistema de [Matos et al., 2008a] e [Matos et al., 2008b] são representadas da mesma forma que as imagens indexadas pelo sistema.

A utilização de uma mesma representação para documentos e consultas permite o cálculo da similaridade entre uma consulta q e uma imagem I . Este cálculo é realizado através da correlação entre os vetores que os representam, quantificada pelo cosseno do ângulo formado por $\vec{I}$ e $\vec{q}$. Esta métrica é conhecida como medida de similaridade do cosseno conforme descrito na seção 2.3.2.

Os valores da similaridade entre uma expressão de busca e cada uma das imagens da coleção são utilizados na ordenação das imagens resultantes. Portanto, no modelo vetorial os resultados são ordenados de acordo com a medida de similaridade do cosseno formando o ranking.

A vantagem desse processo de busca do sistema CBIR-Índice está no fato de que apenas as imagens que contiverem no mínimo um dos termos de busca da consulta são recuperadas. Em relação ao cálculo do cosseno, os termos que não tiverem presentes receberão peso zero e, portanto, seus produtos não serão calculados.

3.3 Sistema CBIR-I

Com o propósito de validar a proposta descrita na seção 3.2, os autores desenvolveram um sistema para o qual atribuíram o nome de CBIR-I. O sistema CBIR-I também foi utilizado com o objetivo de validação da abordagem proposta neste trabalho.

O sistema CBIR-I utiliza o mesmo paradigma do CBIR-Índice, porém se diferem no processo de indexação tradicional das imagens e no cálculo da similaridade.

O CBIR-I indexa as imagens com os valores encontrados nos três momentos de cor para cada canal de cor do modelo HSV. Diferente do sistema CBIR-Índice o qual agrupa os valores encontrados em cada característica e indexa as faixas. Assim o vetor de características do sistema terá nove posições distribuídas conforme a Tabela 3.1.

O cálculo da similaridade é realizado com base na distância Euclidiana entre os vetores,

dada pela equação 3.2.

$$d(\vec{x}, \vec{y}) = \sqrt{\sum_{k=1}^n (|x_k - y_k|^2)} \quad (3.2)$$

em que x_k é o valor correspondente a um elemento da k -ésima posição do vetor $\vec{x}$ que representa uma imagem do banco de dados de característica e y_k é a k -ésima posição do vetor de característica $\vec{y}$ que representa a imagem de consulta.

O algoritmo para o cálculo da similaridade baseado nesta forma é semelhante ao algoritmo da Figura 2.18 e, quando se trabalha com grandes coleções, torna computacionalmente cara a recuperação das imagens. Pois o processo de recuperação é baseado no cálculo da distância entre o vetor da imagem de consulta e todos os vetores das imagens do banco de dados, tornando assim, o processo inviável para grandes coleções.

A idéia de usar a distância Euclidiana do CBIR-I foi para comparar a precisão obtida entre os rankings gerados pelo cosseno e por esta distância, sendo que a primeira possibilita o uso da lista invertida.

3.4 CBIR-Índice-Particionado

Na seção anterior foi descrito o trabalho de [Matos et al., 2008a] e [Matos et al., 2008b] no qual os autores apresentam uma proposta de utilização do índice invertido para recuperação de imagens baseado em conteúdo utilizando a característica cor, que aqui denominaremos CBIR-Índice. Com o intuito de aproveitar as vantagens do sistema CBIR-Índice, que obteve um ganho significativo no tempo de processamento das consultas sem impactos negativos na qualidade da recuperação, este trabalho propõe uma abordagem semelhante que utiliza o mesmo processo de indexação e o mesmo cálculo da similaridade, porém, acrescentando duas novas características, forma e textura.

Outro diferencial do trabalho proposto é utilizar partições do domínio das imagens para extração das características, ou seja, as imagens foram particionadas em diversas regiões e as características são extraídas de cada região. O vetor da imagem será composto pelas características extraídas em cada região. Tal fato é considerado, pois, segundo [Datta et al., 2008], a quantidade de termos relacionados às imagens contribuem na precisão da recuperação, uma vez que podem descrever de forma mais detalhada o conteúdo de uma imagem.

Esta seção apresenta o sistema construído neste trabalho de pesquisa. Inicialmente é dada uma visão geral do sistema proposto, seus módulos principais e seu funcionamento. Posteriormente, inicia-se a apresentação dos componentes do sistema: os descritores e a medida de similaridade.

Com o propósito de validar a proposta deste trabalho, as duas novas características

utilizadas neste trabalho, a saber forma e textura, também foram implementadas nos sistemas CBIR-Índice e CBIR-I. Ao longo deste capítulo tais modificações são melhores descritas.

3.4.1 Esquema Geral

O objetivo deste trabalho é obter uma contribuição na qualidade e quantidade dos termos utilizados na indexação por meio das partições das imagens, e aplicá-lo dentro do ambiente de recuperação por conteúdo para grandes coleções de imagens, usando índice invertido.

A Figura 3.5 mostra a estrutura geral do sistema desenvolvido nesta pesquisa. Inicialmente, considera-se um banco de dados de imagens onde são armazenadas todas as imagens do universo de busca do sistema. As imagens deste banco são divididas em diversas partições formando um banco de dados de imagens particionadas. O banco de dados de imagens particionadas alimenta o módulo de extração de características, que produz como saída, estruturas contendo o código de identificação das imagens e os vetores de características de cada partição do banco de dados. Estes dados (código de identificação + vetores de características) são armazenadas no banco de dados de características. O índice invertido é construído a partir do banco de dados de características obtido. Este processo de caracterização das imagens do banco de dados, ilustrado pela setas pontilhadas, é feito em fase de pré-processamento (*off-line*).

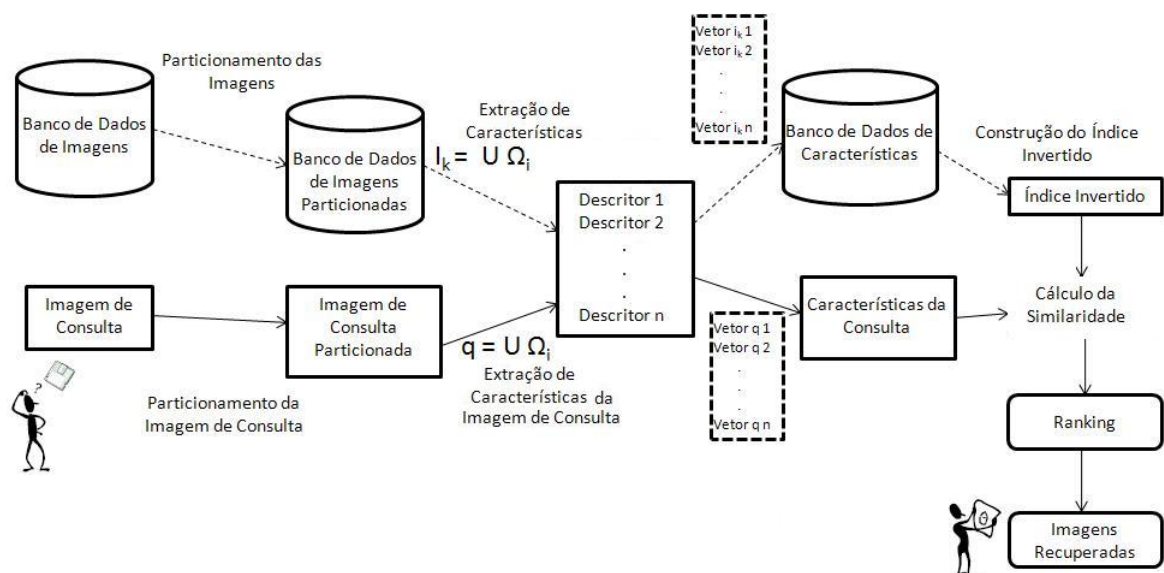


Figura 3.5: Fluxograma básico do sistema CBIR-Índice-Particionado proposto neste trabalho.

Quando o usuário faz uma busca (*on-line*), a imagem de consulta é particionada e vetores de características são extraídos de cada partição usando os mesmos módulos da

fase de pré-processamento. O cálculo da similaridade com as características das imagens do banco de dados é feito utilizando o Índice Invertido. O módulo cálculo da similaridade retorna um valor de similaridade para cada imagem retornada pelo Índice Invertido, em relação à imagem de consulta. Com base nestes valores as imagens são colocadas em ordem decrescente de similaridade (*ranking*) e mostradas ao usuário.

As próximas seções descrevem cada módulo do sistema apresentado no esquema da Figura 3.5.

3.4.2 Particionamento das Imagens

Uma das etapas utilizadas na proposta deste trabalho é o particionamento das imagens. Diversos trabalhos na literatura utilizam a técnica de particionamento ou divisão das imagens em regiões, como por exemplo [Laaksonen, 2000] e [Penatti and Torres, 2007].

Em [Laaksonen, 2000] as imagens são divididas em cinco regiões e as características visuais são extraídas de cada região separadamente. O particionamento utilizado pelos autores consiste em dividir a imagem de forma que uma das partições seja um círculo posicionado no centro da imagem, e as demais quatro partições dividem o restante da imagem em tamanhos semelhantes.

[Penatti and Torres, 2007] apresenta um novo descritor de relacionamento espacial para recuperação de imagens por conteúdo. O novo descritor baseia-se no particionamento do espaço em análise de quadrantes e na contagem da ocorrência de pontos do objeto de interesse em cada quadrante.

O particionamento utilizado neste trabalho consiste em dividir as imagens em partes de tamanhos semelhantes, seguindo o particionamento apresentado em [Penatti and Torres, 2007]. Porém cada partição é analisada separadamente e depois de extraídas as características é feita uma soma dos vetores de cada partição. Diversos números de partições foram utilizados e alguns experimentos serão mostrados no próximo capítulo. As partições possuem tamanhos semelhantes pelo fato das imagens terem tamanhos diversos e esses podem não serem múltiplos do número de partições. Por exemplo, considere uma imagem com dimensões 100x120 pixels, se a imagem for particionada utilizando 8 partições teremos 11 partições com tamanho 12x15 e 1 partição com tamanho 16x15.

Considerando uma imagem I no espaço de cor RGB, como uma função $I : \Omega \subset \mathbb{R}^2 \rightarrow [a, b] \times [a, b] \times [a, b] \subset \mathbb{R}^3$ e $\Omega = \Omega_1 \cup \Omega_2 \cup \dots \cup \Omega_i$, de modo que $\Omega_i \cap \Omega_j = \emptyset$ para $i \neq j$, o CBIR-Índice-Particionado utiliza as diversas partições Ω_i da imagem I . Em que $\mathbf{a}$ e $\mathbf{b}$ representam o menor e maior valor, respectivamente, de R, G, e B.

Considere a imagem da Figura 3.6a. A imagem original I foi particionada em quatro partições Ω_1 , Ω_2 , Ω_3 e Ω_4 , mostradas na Figura 3.6b. Em que $\Omega_1 \cap \Omega_2 = \emptyset$, $\Omega_1 \cap \Omega_3 = \emptyset$, $\Omega_1 \cap \Omega_4 = \emptyset$, $\Omega_2 \cap \Omega_3 = \emptyset$, $\Omega_2 \cap \Omega_4 = \emptyset$ e $\Omega_3 \cap \Omega_4 = \emptyset$.



Figura 3.6: Particionamento de uma imagem exemplo I em quatro partições Ω_1 , Ω_2 , Ω_3 e Ω_4 .

3.4.3 Extração de Características

Nesta seção, serão apresentados três descritores, de cor, forma e textura, utilizados na caracterização das imagens. São eles, momentos de cor, histograma de direção de bordas e vizinhança de textura.

Momentos de Cor

Como foi mencionado no início deste capítulo, este trabalho propõe uma extensão ao trabalho de [Matos et al., 2008a] e [Matos et al., 2008b], o qual denominamos CBIR-Índice, descrito no capítulo anterior. A forma de extração da característica cor é a mesma utilizada no CBIR-Índice.

As partições das imagens foram convertidas do espaço de cor RGB (*Red, Green and Blue*) para o espaço de cor HSV (*Hue, Saturation and Value*). Então, foram calculados os três primeiros momentos: média, desvio padrão e obliquidade (dados pelas expressões 2.4, 2.5 e 2.6) para cada canal de cor H, S e V, obtendo assim um vetor de características com nove posições. Com isto, atribuiu-se os termos da Tabela 3.1 a cada partição Ω_i da imagem I, construindo um vetor contendo todos os termos encontrados em todas as partições Ω'_i s.

Conforme a Tabela 3.2, cada índice do vetor de características da imagem é dividido em 10 faixas de valores que variam de A até J.

A fim de melhor expressar o objetivo do trabalho proposto, considere a imagem I da Figura 3.6 antes do particionamento. Utilizando o sistema CBIR-Índice, proposto no capítulo anterior, o vetor de termos da imagem exemplo é mostrada na Figura 3.7.

Utilizando o particionamento das imagens proposto neste trabalho, os vetores de característica são obtidos sobre cada partição da Figura 3.6(b). A Figura 3.8 ilustra os vetores obtidos no processo de extração da característica cor utilizando a proposta deste

0	0	0	0	0	0	1	0	0	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
0	0	0	0	1	0	0	0	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	1	0	0	0	0	0	0	0	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	0	1	0
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	0	1	0	0	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	0	1	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	1	0	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
0	0	1	0	0	0	0	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	1	0	0	0	0	0	0	0	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

Figura 3.7: Vetor de características gerado pelo sistema CBIR-Índice correspondente a imagem da figura 3.6.

trabalho. Logo em seguida é realizada a soma dos vetores das partições obtendo o vetor ilustrado na Figura 3.9. Esse vetor representa todos os termos distintos, juntamente com sua frequência, conforme são encontrados nas partições.

Observa-se que a proposta deste trabalho obtém um vetor com características que podem ter frequências maiores do que um, diferentemente da proposta do CBIR-Índice em que cada característica aparece no máximo uma vez. Analisando o vetor da Figura 3.9 é possível verificar que muitas posições do vetor tem frequência maior do que um, por exemplo, as característica MHE, IHC, MSH, MVE, dentre outras. Isso possibilita uma analogia com o fator *tf* do peso de termos no modelo vetorial clássico de recuperação de informação textual.

Histograma de Direção de Bordas

Uma das contribuições deste trabalho em relação ao trabalho de [Matos et al., 2008a] é a utilização do histograma de direção de bordas na definição da característica forma.

O histograma de direção de bordas [Brandt, 1999] modela a similaridade entre imagens em termos da similaridade da distribuição das direções de bordas. As imagens são convertidas para o espaço de cor HSV (*Hue, Saturation and Value*), onde somente os canais S e V são utilizados na identificação das direções de bordas. O canal H é descartado pelo fato deste negligenciar bordas importantes em regiões ou objetos com a mesma tonalidade de cor.

Operadores de Sobel de oito direções são aplicados sobre os dois canais e é definido

0	0	0	0	0	0	0	0	1	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
1	0	0	0	0	0	0	0	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	1	0	0	0	0	0	0	0	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	0	0	1
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	1	0	0	0	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	0	1	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	1	0	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
1	0	0	0	0	0	0	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	0	1	0	0	0	0	0	0	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

(a) Vetor de Termos da Partição 1

0	0	0	0	0	0	0	0	1	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
0	1	0	0	0	0	0	0	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	0	1	0	0	0	0	0	0	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	0	1	0
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	0	1	0	0	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	0	1	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	0	1	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
0	0	1	0	0	0	0	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	0	0	0	0	1	0	0	0	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

(b) Vetor de Termos da Partição 2

0	0	0	0	1	0	0	0	0	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
0	0	1	0	0	0	0	0	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	0	1	0	0	0	0	0	0	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	1	0	0
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	0	0	0	1	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	1	0	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	1	0	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
0	0	0	0	0	1	0	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	0	0	0	0	0	0	1	0	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

(c) Vetor de Termos da Partição 3

0	0	0	0	1	0	0	0	0	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
0	0	0	0	0	0	0	0	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	0	1	0	0	0	0	0	0	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	1	0	0
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	0	0	0	1	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	1	0	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	1	0	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
0	0	0	0	0	1	0	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	0	0	0	0	0	0	1	0	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

(d) Vetor de Termos da Partição 4

Figura 3.8: Vetor de características obtidos de cada partição da imagem da figura 3.6.

um descritor representando as oito direções (H_0^o , H_{45}^o , H_{90}^o , H_{135}^o , H_{180}^o , H_{225}^o , H_{270}^o , H_{315}^o). Cada *pixel*, de cada canal de cor, é analisado segundo as oito máscaras de Sobel, definidas na seção 2.5, e o maior valor obtido define a direção da borda naquele *pixel*. Cada vez que a condição acima for atendida, o contador correspondente àquela direção é incrementado. Finalmente o histograma de direção de bordas é normalizado pelo número de *pixels* presentes na borda, [Brandt, 1999]. Com isto, foi obtido o vetor de característica da imagem com oito posições distribuídas conforme Tabela 3.3.

A característica forma foi indexada da mesma forma que os descritores de cor, ou seja, com base na distribuição normal e na faixa de valores encontrados para cada posição.

Os valores dos vetores de características foram analisados depois de extraídos de coleções com uma ampla variedade de imagens. Ao analisar matematicamente a dis-

0	0	0	0	2	0	0	1	1	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
1	1	1	0	0	0	0	1	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	1	2	0	0	0	0	0	1	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	2	1	1
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	1	1	1	1	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	2	2	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	3	1	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
1	0	1	0	0	1	1	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	0	1	0	0	1	0	1	1	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ

Figura 3.9: Vetor de característica da imagem da figura 3.6 utilizando o CBIR-Índice-Particionado.

tribuição dos valores em cada posição do vetor, foi observado que as distribuições eram semelhantes à distribuição normal da estatística e probabilidade. Com isso, as regras e propriedades desta distribuição foram aplicadas as características para obterem a classificação destas em 10 faixas, conforme a Figura 3.10

Posição do Vetor	Descrição
1º	Histograma de Direção de Bordas 0º (H0)
2º	Histograma de Direção de Bordas 45º (H45)
3º	Histograma de Direção de Bordas 90º (H90)
4º	Histograma de Direção de Bordas 135º (H135)
5º	Histograma de Direção de Bordas 180º (H180)
6º	Histograma de Direção de Bordas 225º (H225)
7º	Histograma de Direção de Bordas 270º (H270)
8º	Histograma de Direção de Bordas 315º (H315)

Tabela 3.3: Descrição do vetor de característica resultante da extração de característica da imagem usando histograma de direção de bordas do espaço HSV.

Desta forma, cada intervalo tem a mesma porcentagem aproximada de imagens descritas no início da seção 3.2.

O vetor de características da imagem irá representar as 8 características da imagem, só que estas características estão agrupadas em faixas. Cada índice do vetor de características da imagem foi dividido em 10 faixas de valores, que variam de A até J de acordo com o seu valor.

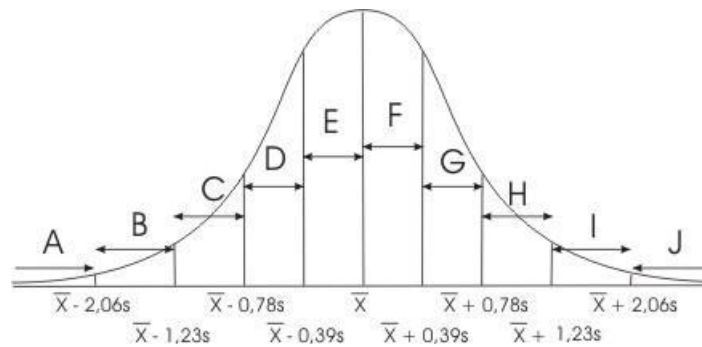


Figura 3.10: Distribuição adotada para classificação das faixas. Onde $\bar{X}$ é a média e s o desvio padrão.

Como descrito na característica cor, utilizando o particionamento das imagens proposto neste trabalho, os vetores de característica são obtidos sobre cada partição da imagem. Logo em seguida é realizada a soma dos vetores das partições obtendo um único vetor que represente todos os termos distintos, juntamente com sua frequência, conforme são encontrados nas partições.

Vizinhança de Textura

Na definição da característica textura foi utilizada vizinha de textura. Esta característica é outro diferencial em relação a proposta de [Matos et al., 2008a].

No método de vizinhança de textura [Laaksonen, 2000] a imagem é representada em níveis de cinza. Um descritor de dimensão oito (V_0^o , V_{45}^o , V_{90}^o , V_{135}^o , V_{180}^o , V_{225}^o , V_{270}^o , V_{315}^o), como definido na seção 2.4, é gerado da seguinte forma: se o valor de um *pixel* na vizinhança-8 for maior que o *pixel* central, então o contador direcional indicando a respectiva direção é incrementado. Isto é feito para todos os *pixels* na imagem. Com isto, foi obtido o vetor de característica da imagem com oito posições distribuídas conforme Tabela 3.4.

Da mesma forma que os descritores de cor, as imagens foram indexadas com os atributos de textura, com base na distribuição normal e na faixa de valores encontrados para cada posição.

Os valores dos vetores de características foram analisados depois de extraídos de coleções com uma ampla variedade de imagens. Ao analisar matematicamente a distribuição dos valores em cada posição do vetor, foi observado que as distribuições eram semelhantes à distribuição normal da estatística e probabilidade. Com isso, as regras e propriedades desta distribuição foram aplicadas as características para obterem a classificação destas em 10 faixas da figura 3.10.

Posição do Vetor	Descrição
1º	Corresponde a vizinhança 0º (V0)
2º	Corresponde a vizinhança 45º (V45)
3º	Corresponde a vizinhança 90º (V90)
4º	Corresponde a vizinhança 135º (V135)
5º	Corresponde a vizinhança 180º (V180)
6º	Corresponde a vizinhança 225º (V225)
7º	Corresponde a vizinhança 270º (V270)
8º	Corresponde a vizinhança 315º (V315)

Tabela 3.4: Descrição do vetor de característica resultante da extração de característica da imagem usando vizinhança de textura do espaço HSV.

Desta forma, cada intervalo tem a mesma porcentagem aproximada de imagens descritas no início da seção 3.2.

O vetor de características da imagem irá representar as 8 características da imagem, só que estas características estão agrupadas em faixas. Cada índice do vetor de características da imagem foi dividido em 10 faixas de valores que variam de A até J de acordo com o seu valor.

Como descrito nas características cor e forma, utilizando o particionamento das imagens proposto neste trabalho, os vetores de característica são obtidos sobre cada partição da imagem. Logo em seguida é realizada a soma dos vetores das partições obtendo um único vetor que represente todos os termos distintos, juntamente com sua frequência, conforme são encontrados nas partições.

A Figura 3.11 ilustra o vetor final depois de serem acrescentadas as duas novas características (textura e forma) utilizando a proposta deste trabalho e utilizando quatro partições.

3.4.4 Construção do Índice Invertido

Para que o modelo de recuperação textual índice invertido possa ser utilizado é necessário definir os termos de indexação para coleção. Assim como na proposta do CBIR-Índice, definiu-se as seguintes analogias entre recuperação de imagens e recuperação textual: **i)** imagens correspondem aos documentos ; **ii)** as faixas encontradas através da análise dos vetores de características são os termos; **iii)** imagens exemplos ou imagens de consulta são as consultas.

Analisando os três descritores utilizados na proposta deste trabalho, pode-se observar que o descritor cor continuará tendo o mesmo número de possíveis termos. Ou seja, tendo um vetor com nove posições onde cada posição pode possuir 10 possíveis faixas (A a J), tendo assim noventa termos distintos. A diferença em relação a proposta de [Matos et al., 2008a] é o uso dos descritores de textura e forma, que, como mencionado anteriormente, são categorizados com base na faixa de valores encontrados, sendo assim também cada

0	0	0	0	2	0	0	1	1	0
MHA	MHB	MHC	MHD	MHE	MHF	MHG	MHH	MHI	MHJ
1	1	1	0	0	0	0	1	0	0
DHA	DHB	DHC	DHD	DHE	DHF	DHG	DHH	DHI	DHJ
0	1	2	0	0	0	0	0	1	0
IHA	IHB	IHC	IHD	IHE	IHF	IHG	IHH	IHI	IHJ
0	0	0	0	0	0	0	2	1	1
MSA	MSB	MSC	MSD	MSE	MSF	MSG	MSH	MSI	MSJ
0	0	0	0	0	1	1	1	1	0
DSA	DSB	DSC	DSD	DSE	DSF	DSG	DSH	DSI	DSJ
0	2	2	0	0	0	0	0	0	0
ISA	ISB	ISC	ISD	ISE	ISF	ISG	ISH	ISI	ISJ
0	0	0	0	3	1	0	0	0	0
MVA	MVB	MVC	MVD	MVE	MVF	MVG	MVH	MVI	MVJ
1	0	1	0	0	1	1	0	0	0
DVA	DVB	DVC	DVD	DVE	DVF	DVG	DVH	DVI	DVJ
0	0	1	0	0	1	0	1	1	0
IVA	IVB	IVC	IVD	IVE	IVF	IVG	IVH	IVI	IVJ
0	1	2	0	0	0	0	0	1	0
H0A	H0B	H0C	H0D	H0E	H0F	H0G	H0H	H0I	H0J
0	0	0	0	3	1	0	0	0	0
H45A	H45B	H45C	H45D	H45E	H45F	H45G	H45H	H45I	H45J
1	1	1	0	0	0	0	1	0	0
H90A	H90B	H90C	H90D	H90E	H90F	H90G	H90H	H90I	H90J
1	0	1	0	0	1	1	0	0	0
H135A	H135B	H135C	H135D	H135E	H135F	H135G	H135H	H135I	H135J
0	0	0	0	2	0	0	1	1	0
H180A	H180B	H180C	H180D	H180E	H180F	H180G	H180H	H180I	H180J
0	2	2	0	0	0	0	0	0	0
H225A	H225B	H225C	H225D	H225E	H225F	H225G	H225H	H225I	H225J
1	1	1	0	0	0	0	1	0	0
H270A	H270B	H270C	H270D	H270E	H270F	H270G	H270H	H270I	H270J
0	1	2	0	0	0	0	0	1	0
H315A	H315B	H315C	H315D	H315E	H315F	H315G	H315H	H315I	H315J
0	0	0	0	3	1	0	0	0	0
V0A	V0B	V0C	V0D	V0E	V0F	V0G	V0H	V0I	V0J
1	1	1	0	0	0	0	1	0	0
V45A	V45B	V45C	V45D	V45E	V45F	V45G	V45H	V45I	V45J
0	2	2	0	0	0	0	0	0	0
V90A	V90B	V90C	V90D	V90E	V90F	V90G	V90H	V90I	V90J
0	0	0	0	2	0	0	1	1	0
V135A	V135B	V135C	V135D	V135E	V135F	V135G	V135H	V135I	V135J
0	0	1	0	0	1	0	1	1	0
V180A	V180B	V180C	V180D	V180E	V180F	V180G	V180H	V180I	V180J
0	0	0	0	3	1	0	0	0	0
V225A	V225B	V225C	V225D	V225E	V225F	V225G	V225H	V225I	V225J
0	0	1	0	0	1	0	1	1	0
V270A	V270B	V270C	V270D	V270E	V270F	V270G	V270H	V270I	V270J
0	0	0	0	2	0	0	1	1	0
V315A	V315B	V315C	V315D	V315E	V315F	V315G	V315H	V315I	V315J

Figura 3.11: Vetor com as características de uma imagem exemplo utilizando o CBIR-Índice-Particionado. Noventa posições do vetor estão relacionadas à cor, sendo dez para cada momento, oitenta relacionadas a textura e oitenta relacionadas a forma.

posição pode possuir 10 possíveis faixas (A a J). Como descrito na seção anterior os descritores de textura e forma possuem oito posições cada um, assim categorizando esses descritores terá oitenta posições distintas, ou seja, oitenta termos distintos.

Assim teremos um vocabulário contendo os noventa termos do descritor cor, oitenta do descritor textura e oitenta da forma. Totalizando 250 termos no vocabulário do índice invertido. Cabe lembrar que na proposta de [Matos et al., 2008a] são utilizados somente 90 termos.

Da mesma forma que no índice invertido da recuperação textual, cada termo irá apontar para as imagens onde aquele termo está presente. Neste caso, as imagens que contêm aquele termo juntamente com a frequência que o termo aparece na imagem (termo : nome-da-imagem, frequência ; nome-da-imagem, frequência ; nome-da-imagem frequência ;....).

3.4.5 Similaridade

Consultas nos arquivos invertidos são feitas tomando-se a lista invertida correspondente aos termos procurados. Por exemplo, caso a imagem de consulta tivesse os seguintes termos: MHA (Média da Matiz na faixa A), DSC (Desvio Padrão da Saturação na faixa C) e IVJ (Inclinação da Intensidade na faixa J). Apenas as lista invertidas destes três termos seriam percorridas.

As imagens de consulta são representadas da mesma forma que as imagens indexadas pelo sistema. Com isso é possível calcular a similaridade entre uma consulta q e uma imagem I .

O cálculo da similaridade é executado da mesma forma que a proposta do CBIR-Índice descrito no capítulo anterior. Ou seja, é realizado através da correlação entre os vetores que representam a imagem de consulta e a imagem em questão, quantificada pelo cosseno do ângulo formado por $\vec{I}$ e $\vec{q}$.

Em relação ao peso dos termos, a abordagem utilizada também foi a mesma do CBIR-Índice o $tf - idf_{t,d}$, descrito na seção 2.3.2, o qual pode ser obtido pela expressão 2.8. Na imagem é considerado o número de ocorrências do termo naquela imagem. A frequência na coleção é dada pelo idf, apresentado na seção 2.3.2.

Os valores da similaridade encontrados pela equação 2.9 entre uma imagem consulta e cada uma das imagens da coleção são utilizados na ordenação das imagens resultantes. Logo, no modelo vetorial os resultados são ordenados de acordo com a medida de similaridade do cosseno formando o *ranking*.

A vantagem desse processo de busca utilizado pelos autores nos trabalhos [Matos et al., 2008a] e [Matos et al., 2008b] está no fato de que apenas as imagens que contiverem no mínimo um dos termos de busca da consulta são recuperadas. Em relação ao cálculo do cosseno, os termos que não tiverem presentes receberão peso zero e, portanto, seus produtos não serão calculados.

3.4.6 Adaptação dos Sistema CBIR-I e CBIR-Índice

Com o objetivo de validar a proposta deste trabalho, os sistemas CBIR-I e CBIR-Índice propostos por [Matos et al., 2008a] e [Matos et al., 2008b] foram modificados, adicionando-se dois novos descritores. Ou seja, além do descritor cor utilizado na indexação das imagens, os dois descritores utilizados no sistema CBIR-Índice-Particionado, textura e forma, foram também implementados nos sistemas CBIR-Índice e CBIR-I.

Em ambos os sistemas as metodologias de extração de características, indexação e cálculo da similaridade permaneceram. A única mudança foi um aumento no número de características utilizadas na indexação das imagens, aumentando assim o tamanho do vetor de características. Em ambos os sistemas o tamanho do vetor de características será o mesmo do sistema aqui proposto, ou seja, os noventa termos do descritor cor, oitenta do descritor textura e oitenta da forma. Totalizando 250 termos no vocabulário do índice invertido.

As formas de extração das características que foram acrescentadas seguem a mesma metodologia do sistema proposto neste trabalho.

3.4.7 Considerações Finais

Um dos problemas da maioria dos sistemas CBIR tradicionais que usam a distância Euclidiana no cálculo da similaridade entre a imagem de consulta e as imagens da coleção é percorrer a coleção toda para calcular a similaridade entre a imagem de consulta e cada imagem da coleção.

Na tentativa de contornar este problema foi apresentada uma proposta que utiliza uma técnica de recuperação textual, índice invertido, na indexação das imagens. Esta técnica varre somente a lista de termos que estão presentes na imagem de consulta.

O particionamento das imagens foi adicionado nesta proposta, tal fato foi considerado pois, segundo [Datta et al., 2008], a quantidade de termos relacionados às imagens contribuem na precisão da recuperação, uma vez que podem descrever de forma mais completa o conteúdo de uma imagem.

A complexidade do cálculo do ranking baseado na distância Euclidiana é função do número de imagens da coleção, já a complexidade de nossa proposta, é função do subconjunto de imagens que têm pelo menos um termo comum à imagem consulta.

Neste capítulo foram descritos os sistemas CBIR-I, CBIR-Índice e CBIR-Índice-Particionado. O primeiro trabalho foi implementado com as características tradicionais de sistemas CBIR, com o intuito de ser utilizado como base de comparação tanto na qualidade quanto na velocidade do processo de recuperação com o CBIR-Índice. O sistema CBIR-Índice foi implementado utilizando técnicas tradicionais da recuperação de informação textual, tais como, índice invertido, peso do termo e modelo vetorial, com o intuito de obter um ganho de tempo no processo de recuperação sem grandes perdas na qualidade de

recuperação. O sistema CBIR-Índice-Particionado propõe uma abordagem semelhante ao CBIR-Índice, utiliza o mesmo processo de indexação e o mesmo cálculo da similaridade, porém, acrescenta duas novas características, forma e textura. Outro diferencial do CBIR-Índice-Particionado é utilizar partições do domínio das imagens para extração das características.

O próximo capítulo descreve os experimentos realizados neste trabalho bem como os resultados obtidos.

Capítulo 4

Resultados Experimentais

Este capítulo apresenta uma avaliação experimental do sistema proposto, o CBIR-Índice-Particionado. Os experimentos aqui realizados têm três objetivos principais: i) comparar os resultados do CBIR-Índice-Particionado com aqueles obtidos pelos sistemas CBIR-I e CBIR-Índice utilizando somente a característica cor; ii) comparar os resultados do CBIR-Índice-Particionado com aqueles obtidos pelos sistemas CBIR-I e CBIR-Índice utilizando as características cor, textura e forma; iii) comparar os sistemas em relação ao número de operações aritméticas efetuadas para o cálculo da similaridade, entre a imagem de consulta e a base de dados.

Para comparar estes sistemas foram utilizadas duas coleções (Corel-1000 [Database, 2010] e BD-10000), as quais serão descritas a seguir. Inicialmente foi utilizada uma base de treinamento para definir o número de partições a ser utilizado. A coleção escolhida para esta etapa foi a BD-110 descrita a seguir.

Nos experimentos a medida de avaliação da qualidade do *Ranking* foi a Precisão em pontos específicos do ranking e a curva de Precisão x Revocação.

4.1 Base de dados

Na condução dos experimentos foram utilizados três banco de dados. Uma descrição seguida de uma amostra de cada um destes bancos de dados é apresentada a seguir:

- Corel-1000: o banco de dados é composto por 10 categorias (África, Praia, Edifícios, Ônibus, Dinossauros, Elefantes, Flores, Comidas, Cavalos e Montanhas), com 100 imagens coloridas em cada. As imagens são coloridas e possuem uma resolução de 384x256 ou de 256x384 *pixels* [Database, 2010]. A Figura 4.2 contém uma amostra deste banco de dados. Na realização dos experimentos foram selecionadas 10 imagens de consulta de cada cada classe de forma aleatória.
- BD-10000: é um banco de dados de 10000 imagens reais cobrindo uma ampla variedade de categorias semânticas (textura uniforme, pôr-do-sol, nuvens, oceanos,

animais, mamíferos, peixes, pássaros, répteis, plantas, flores, paisagens naturais, edifícios, motocicletas, carros, aviões, etc) sendo que 1000 dessas imagens são transportadas da coleção Corel-1000. As outras 9000 imagens foram coletadas de bases de dados públicas disponíveis na Web, principalmente de [CalPhotos, 2010], e pertencem a categorias diferentes daquelas da coleção Corel-1000. A Figura 4.1 contém uma amostra deste banco de dados. Na realização dos experimentos foram selecionadas 20 imagens de consulta das três classes classificadas.

- BD-110: é um banco de dados de 110 imagens coloridas, sendo 11 imagens extraídas aleatoriamente de cada uma das categorias do Corel-1000. As 110 imagens do BD-110 estão disponíveis provisoriamente no laboratório de BD da UFU. De forma a avaliar os três sistemas, todas as imagens da base de dados BD-110 foram utilizadas como consulta.



Figura 4.1: Amostra do Banco de Dados BD-10000.

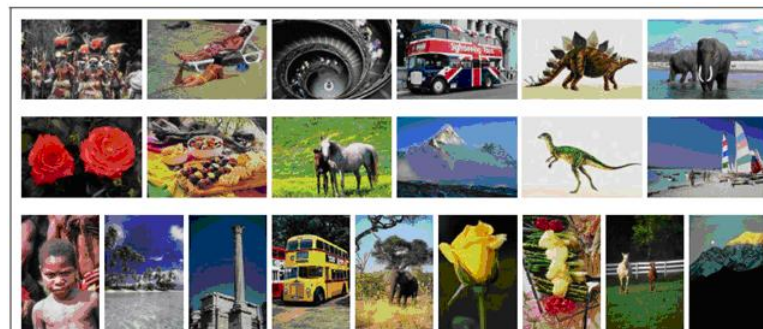


Figura 4.2: Amostra do Banco de Dados BD-110 e Corel 1000.

4.2 Avaliação Experimental

Para avaliar o desempenho do nosso sistema de recuperação de imagens, uma série de experimentos foram realizados. A qualidade de recuperação foi avaliada por meio da precisão medida em várias posições do *ranking* e pela curva de precisão x revocação, chamada 11-pt.

Dada uma consulta e um *ranking* resposta, a precisão na posição X do *ranking* é definida como a fração entre o número de imagens relevantes recuperadas, sobre o número total de imagens recuperadas (equação 2.15).

Para definir as posições para a medida da precisão baseou-se em uma estatística encontrada em [Manning, 2007]. Os autores constataram que os usuários de sistemas de recuperação de informação, navegam na grande maioria das vezes até a terceira página de resultados. Portanto, é de fundamental importância, que os resultados da requisição do usuário mais relevantes estejam nas primeiras posições do *ranking*. Assim, calculou-se a Precisão em quatro posições aqui chamadas P5, P10, P20 e P100, onde P5 é o cálculo da precisão na quinta posição do *ranking*, analogamente P10, P20 e P100 são a décima, vigésima e centésima respectivamente.

Para avaliar os experimentos também foram gerados gráficos de precisão x revocação dos três sistemas analisados: CBIR-I, CBIR-Índice e CBIR-Índice-Particionado. Precisão (Precisão) indica a proporção de imagens relevantes em relação ao total de imagens recuperadas e Recall (Revocação) indica a proporção de imagens relevantes recuperadas em relação ao total de imagens relevantes. Essas medidas foram descritas em detalhes na seção 2.3.3.

Uma imagem da coleção é considerada relevante à imagem de consulta se ambas pertencerem à mesma classe. Esta definição de relevância por classe, foi adotada a fim de evitar um dos maiores problemas na análise de resultados em CBIR, a subjetividade humana. Pois considera-se que as imagens nos bancos dados de imagens foram corretamente classificadas.

Na realização dos testes foram escolhidas, de forma aleatória, algumas imagens da coleção como imagem de consulta. Para a coleção BD-110 foram utilizadas todas as imagens da coleção como imagens consulta. A imagem consulta é retirada do banco no momento de calcular a precisão, técnica conhecida como leave-one-out. Para a coleção Corel-1000 e BD-10000 foram escolhidas como consulta 100 imagens e 80 imagens, respectivamente.

Uma das etapas utilizadas na proposta deste trabalho é o particionamento das imagens. O particionamento utilizado consiste em dividir a imagem em partições tamanhos semelhantes. Diversos número de partições foram utilizados neste trabalho.

Nos experimentos utilizando a base de dados de treinamento, ou BD-110, foram utilizadas cinco tamanhos de partições: 2, 4, 8, 16 e 32. Ou seja, as imagens foram parti-

cionadas em cinco tamanhos diferentes, sendo que as partes têm dimensões semelhantes. Logo em seguida, é feita uma análise de qual partição obteve o melhor resultado na coleção de treinamento e assim esse tamanho de partição foi utilizado nos demais experimentos.

4.2.1 Avaliação Experimental BD-110

Nos experimentos são comparados três sistemas de recuperação de imagens. O primeiro é o CBIR-I, que é um sistema tradicional de CBIR baseado na distância euclidiana entre os vetores de características. O segundo sistema, CBIR-Índice, é o sistema descrito na seção 3.2, utilizando o Índice Invertido. E o terceiro sistema, CBIR-Índice-Particionado, é o sistema proposto neste trabalho, o qual utiliza a indexação de características de baixo nível em partições de imagens digitais mapeadas para um índice invertido.

Para a realização deste primeiro experimento foi utilizada uma base de dados de treinamento que, neste caso, foi a BD-110. De forma a avaliar os três sistemas, todas as imagens da base de dados foram testadas como consulta e para cada uma delas foram calculadas as curvas 11-pt e a precisão em diversas posições do ranking.

No primeiro experimento consideramos somente a característica cor para a indexação e recuperação das imagens, e nos demais experimentos consideramos as três características (cor, forma e textura), as quais foram implementadas no sistema CBIR-Índice-Particionado e nos demais sistemas utilizados neste trabalho.

Avaliação Experimental Utilizando a Característica Cor

Nos sistemas CBIR-I e CBIR-Índice a característica utilizada na indexação e recuperação das imagens é a cor. Assim, nos experimentos descritos nesta seção somente a característica cor foi utilizada na indexação e recuperação das imagens para verificar o ganho obtido com o sistema CBIR-Índice-Particionado utilizando cinco tamanhos de partições (2, 4, 8, 16 e 32) em relação aos outros dois, CBIR-I e CBIR-Índice.

A Tabela 4.1 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-I juntamente com o valor médio da precisão de 110 consultas da coleção. Já na Tabela 4.2 encontram-se os resultados para o sistema CBIR-Índice. Observe que a última coluna da tabela, o P100, representa a precisão a 100% de revocação.

As Tabelas 4.3, 4.4, 4.5, 4.6 e 4.7 mostram os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-Índice-Particionado. Em que cada uma das tabelas apresenta um tamanho de partição do domínio utilizado.

De acordo com os resultados pode-se observar que o sistema CBIR-Índice-Particionado apresentou um ganho na qualidade de recuperação, em relação ao sistema CBIR-I na maioria das posições do rankings, houve uma pequena perda somente nas posição P5 e P100 utilizando 8 e 16 partições respectivamente. Em relação ao sistema CBIR-Índice os resultados do CBIR-Índice-Particionado sempre foram superiores.

Tabela 4.1: Precisão média no sistema CBIR-I utilizando a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	41,82%	33,64%	29,55%	10,00%
Cavalo	74,55%	54,55%	32,73%	10,00%
Comida	50,91%	43,64%	37,27%	10,00%
Construção	38,18%	32,73%	25,91%	10,00%
Dinossauro	100,00%	80,00%	50,00%	10,00%
Elefante	54,55%	35,45%	23,18%	10,00%
Flor	49,09%	34,55%	21,82%	10,00%
Montanha	45,45%	37,27%	27,73%	10,00%
Ônibus	54,55%	41,82%	31,82%	10,00%
Praia	20,00%	20,00%	18,64%	10,00%
Média	52,91%	41,36%	29,86%	10,00%

Tabela 4.2: Precisão média no sistema CBIR-Índice utilizando a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	23,64%	17,27%	15,45%	9,27
Cavalo	40,00%	30,91%	21,82%	8,91
Comida	29,09%	27,27%	25,00%	9,09
Construção	25,45%	22,73%	16,82%	7,82
Dinossauro	74,55%	56,36%	40,45%	10,00
Elefante	29,09%	27,27%	20,45%	9,09
Flor	30,91%	24,55%	19,55%	8,55
Montanha	18,18%	13,64%	15,91%	8,55
Ônibus	36,36%	30,91%	24,55%	9,27
Praia	18,18%	19,09%	14,09%	8,18
Média	32,55%	27,00%	21,41%	8,87%

Tabela 4.3: Precisão média no sistema CBIR-Índice-Particionado utilizando 2 partições e a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	21,82%	23,64%	24,55%	9,91
Cavalo	56,36%	45,45%	33,18%	10,00
Comida	40,00%	43,64%	30,00%	10,00
Construção	27,27%	20,91%	17,73%	9,18
Dinossauro	89,09%	75,45%	45,91%	10,00
Elefante	36,36%	22,73%	21,36%	9,91
Flor	36,36%	32,73%	25,00%	9,45
Montanha	40,00%	34,55%	27,73%	9,91
Ônibus	47,27%	41,82%	30,00%	10,00
Praia	20,00%	18,18%	16,36%	10,00
Média	41,45%	35,91%	27,18%	9,84%

As principais vantagens do CBIR-Índice-Particionado estão relacionadas ao ganho de tempo no processo de recuperação em relação ao CBIR-I e na obtenção da melhoria na qualidade em relação ao CBIR-I e CBIR-Índice, sendo de acordo com nossos experimentos, uma proposta melhor que as duas.

Analisando os resultados nos pontos do *ranking* em geral é possível perceber que o sistema CBIR-Índice-Particionado sempre obteve melhores resultados em relação aos de-

Tabela 4.4: Precisão média no sistema CBIR-Índice-Particionado utilizando 4 partições e a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	34,55%	29,09%	29,09%	10,00
Cavalo	45,45%	40,00%	32,27%	9,91
Comida	54,55%	40,91%	33,18%	10,00
Construção	27,27%	23,64%	16,36%	9,64
Dinossauro	96,36%	85,45%	50,00%	10,00
Elefante	36,36%	30,00%	23,18%	10,00
Flor	47,27%	37,27%	28,64%	9,82
Montanha	23,64%	20,91%	22,73%	9,91
Ônibus	47,27%	42,73%	32,73%	10,00
Praia	20,00%	19,09%	18,18%	9,91
Média	43,27%	36,91%	28,64%	9,92%

Tabela 4.5: Precisão média no sistema CBIR-Índice-Particionado utilizando 8 partições e a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	45,45%	40,00%	31,36%	10,00
Cavalo	74,55%	60,00%	40,00%	10,00
Comida	65,45%	53,64%	39,55%	10,00
Construção	25,45%	20,91%	19,09%	10,00
Dinossauro	100,00%	96,36%	50,00%	10,00
Elefante	29,09%	29,09%	22,27%	10,00
Flor	63,64%	51,82%	33,18%	10,00
Montanha	23,64%	26,36%	23,18%	10,00
Ônibus	65,45%	49,09%	35,91%	10,00
Praia	20,00%	25,45%	23,64%	10,00
Média	51,27%	45,27%	31,82%	10,00%

Tabela 4.6: Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	34,55%	36,36%	31,36%	10,09
Cavalo	81,82%	66,36%	41,82%	10,00
Comida	81,82%	62,73%	41,82%	10,00
Construção	49,09%	37,27%	25,45%	9,64
Dinossauro	100,00%	100,00%	50,00%	10,00
Elefante	30,91%	28,18%	20,45%	10,00
Flor	60,00%	43,64%	32,27%	10,00
Montanha	23,64%	20,00%	16,36%	9,82
Ônibus	67,27%	50,91%	34,09%	10,00
Praia	14,55%	19,09%	18,64%	10,00
Média	54,36%	46,45%	31,23%	9,95%

mais sistemas. Nos pontos P5 e P10 o CBIR-Índice-Particionado obteve melhor resultado utilizando o número de partições igual a 16. Nos demais pontos o número de partição 8 obteve melhor resultado.

Em geral os números de partições igual a 8 e 16 obtiveram os melhores resultados na coleção de treinamento. A Tabela 4.8 apresenta o ganho obtido com o sistema CBIR-Índice-Particionado em relação aos outros sistemas, CBIR-I e CBIR-Índice. Em relação

Tabela 4.7: Precisão média no sistema CBIR-Índice-Particionado utilizando 32 partições e a característica cor. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	38,18%	35,45%	31,36%	10,09
Cavalo	69,09%	57,27%	39,09%	10,00
Comida	78,18%	61,82%	42,27%	10,00
Construção	54,55%	40,00%	29,09%	9,36
Dinossauro	100,00%	100,00%	50,00%	10,00
Elefante	29,09%	25,45%	21,36%	9,82
Flor	52,73%	45,45%	36,82%	10,00
Montanha	20,00%	19,09%	16,82%	9,73
Ônibus	78,18%	54,55%	34,55%	10,00
Praia	18,18%	17,27%	15,45%	10,18
Média	53,82%	46,64%	31,68%	9,92%

ao CBIR-I e ao CBIR-Índice, o CBIR-Índice-Particionado obteve um ganho significativo nos diversos pontos do *ranking*. Em relação ao CBIR-Índice, o CBIR-Índice-Particionado sempre demonstrou um ganho superior na qualidade da recuperação.

De acordo com os resultados pode-se observar que somente o sistema CBIR-I sempre atinge 10% na posição P100 dos rankings. Esse fato ocorre devido ao sistema CBIR-I fazer a comparação entre a consulta e toda a coleção e os demais sistemas somente com as imagens da coleção que tem pelo menos uma característica em comum a consulta. Portanto na lista de rankings do sistema CBIR-I sempre estará todas as imagens da coleção e nas dos demais sistemas somente as imagens com pelo menos uma característica em comum a consulta.

Tabela 4.8: Comparação do CBIR-Índice-Particionado utilizando partições com tamanho 8 e 16 com os sistemas CBIR-I e CBIR-Índice utilizando a característica cor. Coleção BD-110

	P5	P10	P20	P100
Média CBIR-I	52,91%	41,36%	29,86%	9,99
Média CBIR-Índice	32,55%	27,00%	21,41%	8,87
Média CBIR-Índice-Particionado-8-partições	51,27%	45,27%	31,82%	10,00
Média CBIR-Índice-Particionado-16-partições	54,36%	46,45%	31,23%	9,95
Razão CBIR-Índice-Particionado-8/ CBIR-I	-3,09%	9,45%	6,54%	0,09%
Razão CBIR-Índice-Particionado-8/ CBIR-Índice	57,54%	67,68%	48,62%	12,70%
Razão CBIR-Índice-Particionado-16/ CBIR-I	2,75%	12,31%	4,57%	-0,36%
Razão CBIR-Índice-Particionado-16/ CBIR-Índice	67,04%	72,05%	45,86%	12,19%

Para analisar os resultados, também foram aplicadas as curvas 11-pt. As médias dos valores obtidos foram usadas para gerar curvas 11-pt mostradas na Figura 4.3, onde podemos observar que a precisão se aproxima de 0% quando a revocação se aproxima de 100%. Ao analisar as expressões para o cálculo da precisão e revocação, podemos perceber que a medida que percorremos a lista de rankings, o número de imagens recuperadas aumenta e com isso diminui o valor da precisão, pois somente uma parte das imagens da

coleção são relevantes. A precisão zero a 100% de revocação indica que o sistema não conseguiu recuperar todas as imagens relevantes para cada consulta.

Como pode ser visto nos gráficos, os resultados obtidos com o sistema CBIR-Indice-Particionado utilizando partições com tamanhos 16 e 32 foram os que obtiveram os melhores resultados em relação aos demais sistemas. Entretanto, usando 32 partições observa-se uma perda de precisão nas regiões de alta precisão (até 20% de revocação). Diante disso a escolha do número de regiões foi 16, ou seja, como em geral o número de partições igual a 16 obteve melhores resultados na coleção de treinamento, nas demais coleções utilizaremos esse número de partição na execução dos testes que utilizam a característica cor.

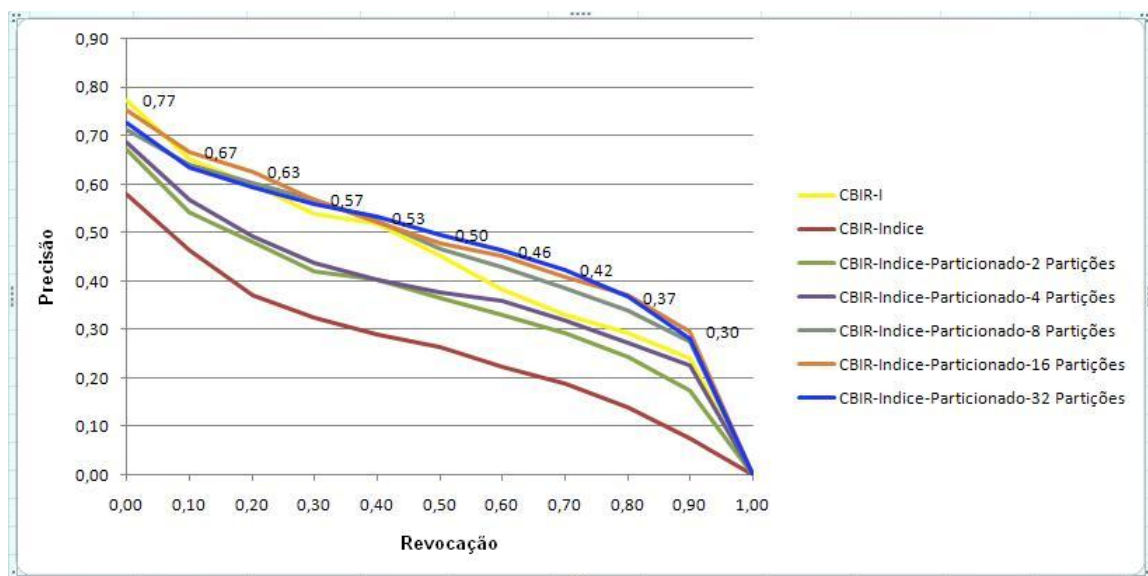


Figura 4.3: Curva precisão x revocação dos sistemas CBIR-I, CBIR-Indice e CBIR-Particionado utilizando a característica cor.

Avaliação Experimental Utilizando as Características Cor, Textura e Forma

Nos experimentos descritos nesta seção as três características (cor, forma e textura) foram utilizadas na indexação e recuperação das imagens, visando obter uma melhor descrição das imagens em relação à utilização somente da característica cor.

A Tabela 4.9 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-I relativo à média de 110 consultas de cada classe, juntamente com o valor médio da precisão de 110 consultas da coleção. Já na Tabela 4.10 encontram-se estes mesmos resultados para os sistemas CBIR-Indice.

As Tabelas 4.11, 4.12, 4.13, 4.14 e 4.15 mostram os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-Indice-Particionado. Em que cada uma das tabelas apresenta um tamanho de partição do domínio utilizado.

Analisando os resultados nos pontos de *ranking*, em geral é possível perceber que o sistema CBIR-Indice-Particionado sempre obteve melhores resultados em relação aos

Tabela 4.9: Precisão média no sistema CBIR-I utilizando cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	32,73%	30,00%	23,64%	10,00
Cavalo	60,00%	43,64%	32,27%	10,00
Comida	50,91%	38,18%	29,09%	10,00
Construção	47,27%	34,55%	21,36%	10,00
Dinossauro	94,55%	87,27%	49,09%	10,00
Elefante	52,73%	48,18%	36,36%	10,00
Flor	58,18%	40,00%	26,82%	10,00
Média	52,91%	42,45%	29,45%	10,00
Montanha	29,09%	25,45%	20,00%	10,00
Ônibus	70,91%	58,18%	38,18%	10,00
Praia	32,73%	19,09%	17,73%	10,00
Média	52,91%	42,45%	29,45%	10,00%

Tabela 4.10: Precisão média no sistema CBIR-Índice utilizando cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	27,27%	20,00%	16,82%	9,82
Cavalo	47,27%	40,00%	25,00%	9,91
Comida	25,45%	31,82%	24,55%	10,00
Construção	29,09%	23,64%	16,82%	8,82
Dinossauro	83,64%	66,36%	41,82%	10,00
Elefante	38,18%	30,00%	20,00%	9,82
Flor	38,18%	29,09%	21,36%	9,27
Média	37,45%	31,45%	22,68%	9,65
Montanha	18,18%	18,18%	14,55%	9,82
Ônibus	54,55%	45,45%	33,18%	10,00
Praia	12,73%	10,00%	12,73%	9,09
Média	37,45%	31,45%	22,68%	9,55%

Tabela 4.11: Precisão média no sistema CBIR-Índice-Particionado utilizando 2 partições e as características cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	20,00%	21,82%	20,00%	9,82
Cavalo	54,55%	40,91%	30,45%	10,00
Comida	36,36%	32,73%	26,82%	10,00
Construção	36,36%	28,18%	17,73%	8,91
Dinossauro	98,18%	81,82%	48,18%	10,00
Elefante	38,18%	29,09%	21,36%	9,91
Flor	45,45%	35,45%	27,27%	9,82
Montanha	18,18%	14,55%	14,55%	9,55
Ônibus	72,73%	60,00%	42,27%	10,00
Praia	14,55%	15,45%	15,00%	9,73
Média	43,45%	36,00%	26,36%	9,77%

demais sistemas. Em todos os pontos do ranking o CBIR-Índice-Particionado obteve melhor resultado utilizando o número de partições igual a 16. Resultado análogo àqueles obtidos utilizando somente a característica cor.

Os conceitos de precisão/revocação também foram utilizados para analisar os resulta-

Tabela 4.12: Precisão média no sistema CBIR-Índice-Particionado utilizando 4 partições e as características cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	29,09%	30,00%	25,91%	10,00
Cavalo	56,36%	42,73%	31,36%	10,00
Comida	40,00%	31,82%	31,82%	10,00
Construção	52,73%	38,18%	28,64%	9,55
Dinossauro	100,00%	91,82%	50,00%	10,00
Elefante	40,00%	28,18%	23,18%	10,00
Flor	56,36%	44,55%	30,45%	10,00
Montanha	29,09%	19,09%	15,45%	9,82
Ônibus	83,64%	67,27%	47,27%	10,00
Praia	21,82%	19,09%	16,36%	9,82
Média	50,91%	41,27%	30,05%	9,92%

Tabela 4.13: Precisão média no sistema CBIR-Índice-Particionado utilizando 8 partições e as características cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	29,09%	34,55%	29,55%	10,00
Praia	30,91%	22,73%	23,64%	9,91
Construção	41,82%	40,00%	29,55%	9,64
Ônibus	90,91%	76,36%	45,91%	10,00
Dinossauro	100,00%	98,18%	50,00%	10,00
Elefante	36,36%	31,82%	23,64%	10,00
Flor	74,55%	58,18%	35,91%	10,00
Comida	56,36%	45,45%	35,91%	10,00
Cavalo	63,64%	57,27%	39,55%	10,00
Montanha	21,82%	27,27%	19,09%	9,91
Média	54,55%	49,18%	33,27%	9,95%

Tabela 4.14: Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e as características cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa	40,00%	35,45%	29,09%	10,00
Cavalo	78,18%	65,45%	42,73%	10,00
Comida	65,45%	55,45%	36,36%	10,00
Construção	65,45%	43,64%	33,64%	9,91
Dinossauro	100,00%	100,00%	50,00%	10,00
Elefante	34,55%	29,09%	24,55%	10,00
Flor	72,73%	54,55%	37,73%	10,00
Montanha	25,45%	23,64%	18,64%	10,00
Ônibus	80,00%	65,45%	45,45%	10,00
Praia	29,09%	25,45%	23,18%	10,00
Média	59,09%	48,82%	34,14%	9,99%

dos. As médias dos valores obtidos foram usadas para gerar curvas 11-pt mostradas na Figura 4.4. Como pode ser visto nos gráficos, os resultados obtidos com o sistema CBIR-Índice-Particionado utilizando partições com tamanhos 16 e 32 foram os que obtiveram os melhores resultados em relação aos demais sistemas. No ponto 0.80 o CBIR-I foi o melhor, já no ponto 0.45 o CBIR-Índice-Particionado o tamanho de partição 32 foi o melhor e nos

Tabela 4.15: Precisão média no sistema CBIR-Indice-Particionado utilizando 32 partições e as características cor, textura e forma. Coleção BD-110

Classe	P5	P10	P20	P100
Africa%	40,00%	37,27%	33,18%	10,09
Cavalo%	74,55%	60,91%	41,36%	10,00
Comida%	58,18%	55,45%	39,09%	10,00
Construção%	67,27%	48,18%	32,27%	10,09
Dinossauo%	100,00%	100,00%	50,00%	10,00
Elefante%	29,09%	27,27%	24,09%	9,91
Flor%	60,00%	51,82%	37,73%	10,00
Montanha%	21,82%	20,91%	16,82%	9,73
Ônibus%	78,18%	64,55%	42,27%	10,00
Praia%	21,82%	19,09%	17,73%	9,91
Média	55,09%	48,55%	33,45%	9,97%

demais pontos o tamanho de partição 16 foi melhor.

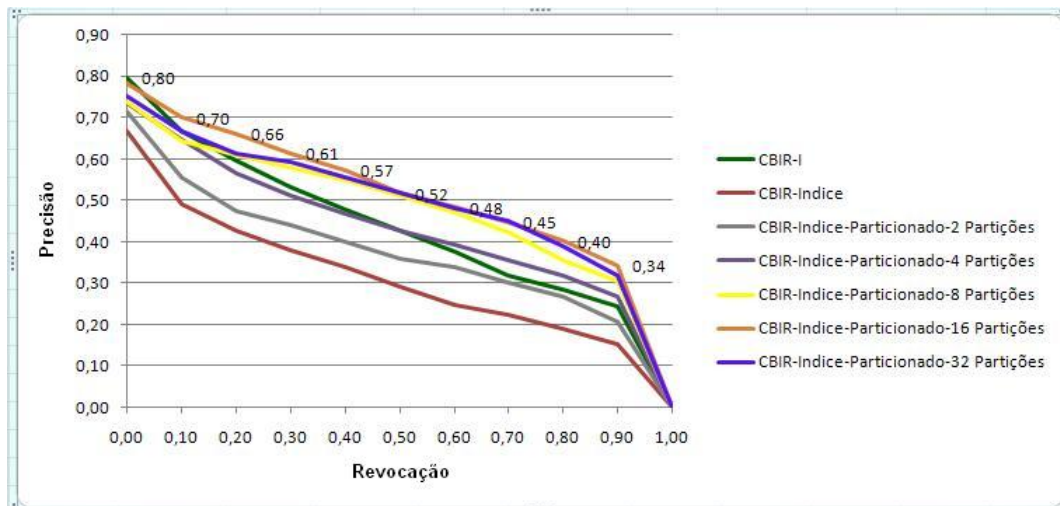


Figura 4.4: Curva 11-pt dos sistemas CBIR-I, CBIR-Indice e CBIR-Particionado utilizando as características cor, forma e textura.

Assim como utilizando somente a característica cor, utilizando as características (cor, forma e textura), em geral o número de partições igual a 16 obteve melhores resultados na coleção de treinamento. Nas demais coleções utilizaremos esse número de partição na execução dos testes.

Analisando os resultados obtidos é possível observar que o sistema CBIR-Indice-Particionado continua tendo resultados superiores aos demais sistemas em análise, mesmo depois das características cor, forma e textura terem sido acrescentadas nas propostas CBIR-I e CBIR-Indice. Ou seja, em todas as posições de *ranking* os valores obtidos aplicando a proposta CBIR-Indice-Particionado foi maior do que os resultados obtidos com a proposta dos sistemas CBIR-I e CBIR-Indice.

A Tabela 4.16 apresenta o ganho obtido com o sistema CBIR-Indice-Particionado em relação aos outros sistemas, CBIR-I e CBIR-Indice. Em relação ao CBIR-I somente na

posição P100 do ranking não houve diferença nas médias, nos demais pontos sempre obteve um ganho significativo. Analisando em relação ao CBIR-Índice, o CBIR-Índice-Particionado sempre demonstrou um ganho superior na qualidade da recuperação.

Tabela 4.16: Comparação do CBIR-Índice-Particionado utilizando 16 partições com os sistemas CBIR-I e CBIR-Índice e as características cor, textura e forma. Coleção BD-110

	P5	P10	P20	P100
Média Geral CBIR-I	52,91%	42,45%	29,45%	9,83%
Média Geral CBIR-Índice	37,45%	31,45%	22,68%	9,65%
Média Geral CBIR-Índice-Particionado-16	59,09%	49,82%	34,14%	9,99%
Razão CBIR-Índice-Particionado-16/CBIR-I	11,68%	17,34%	15,90%	1,67%
Razão CBIR-Índice-Particionado-16/CBIR-Índice	57,77%	58,38%	50,50%	3,48%

4.2.2 Avaliação Experimental Corel-1000

Os mesmos sistemas comparados na seção 4.2.1 são aqui utilizados, mas nesta seção o banco de dados de imagens utilizado é o Corel-1000 e o número de partições usadas no CBIR-Índice-Particionado foi fixado em 16.

Em alguns experimentos consideramos somente a característica cor para a indexação e recuperação das imagens. E nos demais experimentos consideramos as três características (cor, forma e textura), as quais foram implementadas no sistema CBIR-Índice-Particionado e nos demais sistemas utilizados neste trabalho.

Avaliação Experimental Utilizando a Característica Cor

Nos experimentos descritos nesta seção, somente a característica cor foi utilizada na indexação e recuperação das imagens.

A Tabela 4.17 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-I. O banco de dados utilizado foi o Corel-1000 relativo à média de 10 consultas de cada classe, juntamente com o valor médio da precisão de 100 consultas da coleção. Já na Tabela 4.18 encontram-se estes mesmos resultados para os sistemas CBIR-Índice.

A Tabela 4.19 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-Índice-Particionado utilizando 16 partições.

Analisando os resultados obtidos é possível observar que o sistema CBIR-Índice-Particionado continua tendo resultados superiores aos demais sistemas em análise. Ou seja, em todas as posições de *ranking* os valores obtidos aplicando a proposta CBIR-Índice-Particionado foi maior do que os resultados obtidos com a proposta dos sistemas CBIR-I e CBIR-Índice.

A Tabela 4.20 apresenta o ganho obtido com o sistema CBIR-Índice-Particionado em relação aos outros sistemas, CBIR-I e CBIR-Índice na coleção Corel-1000. Analisando

Tabela 4.17: Precisão média no sistema CBIR-I utilizando a característica cor. Coleção Corel-1000

Classe	P5	P10	P20	P100
Africa	90,00%	75,00%	57,50%	42,50%
Cavalo	80,00%	70,00%	67,50%	56,00%
Comida	70,00%	60,00%	57,50%	34,00%
Construção	30,00%	25,00%	22,50%	23,50%
Dinossauro	100,00%	100,00%	100,00%	80,00%
Elefante	60,00%	55,00%	47,50%	32,50%
Flor	50,00%	55,00%	60,00%	27,00%
Montanha	40,00%	25,00%	27,50%	20,50%
Ônibus	50,00%	60,00%	47,50%	34,50%
Praia	30,00%	35,00%	25,00%	18,50%
Média	60,00%	56,00%	51,25%	36,90%

Tabela 4.18: Precisão média no sistema CBIR-Índice utilizando a característica cor. Coleção Corel-1000

Classe	P5	P10	P20	P100
Africa	48,00%	45,00%	43,00%	28,80
Cavalo	66,00%	57,00%	48,50%	31,40
Comida	34,00%	35,00%	36,50%	26,50
Construção	40,00%	35,00%	29,50%	20,20
Dinossauro	88,00%	87,00%	86,50%	62,30
Elefante	66,00%	58,00%	44,50%	28,20
Flor	46,00%	43,00%	40,50%	25,20
Montanha	28,00%	24,00%	18,00%	17,90
Ônibus	60,00%	55,00%	49,00%	29,70
Praia	36,00%	32,00%	29,00%	20,70
Média	51,20%	47,10%	42,50%	29,09%

Tabela 4.19: Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e a característica cor. Coleção Corel-1000

Classe	P5	P10	P20	P100
Africa	58,00%	56,00%	53,00%	42,80
Cavalo	94,00%	92,00%	86,00%	57,60
Comida	64,00%	63,00%	57,50%	38,20
Construção	58,00%	58,00%	53,00%	34,70
Dinossauro	100,00%	100,00%	100,00%	96,30
Elefante	40,00%	36,00%	32,00%	23,30
Flor	82,00%	74,00%	62,50%	42,70
Montanha	48,00%	40,00%	36,50%	23,90
Ônibus	76,00%	73,00%	76,50%	55,70
Praia	30,00%	28,00%	27,50%	23,20
Média	65,00%	62,00%	58,43%	43,84%

em relação ao CBIR-Índice, o CBIR-Índice-Particionado sempre demonstrou um ganho superior na qualidade da recuperação.

Uma análise dos resultados utilizando a curva de precisão x revocação também foi realizada. As médias dos valores obtidos foram usadas para gerar curvas 11-pt mostradas na Figura 4.5.

Como pode ser visto nos gráficos, os resultados obtidos com o sistema CBIR-Índice-

Tabela 4.20: Comparação do CBIR-Indice-Particionado utilizando 16 partições e a característica cor com os sistemas CBIR-I e CBIR-Indice . Coleção Corel-1000

	P5	P10	P20	P100
Média Geral CBIR-I	60,00%	56,00%	51,25%	36,90%
Média Geral CBIR-Indice	51,20%	47,10%	42,50%	29,09%
Média Geral CBIR-Indice-Particionado	65,00%	62,00%	58,45%	43,84%
Razão CBIR-Indice-Particionado-16/CBIR-I	8,33%	10,71%	14,05%	18,81%
Razão CBIR-Indice-Particionado/CBIR-Indice	26,95%	31,63%	37,53%	50,70%

Particionado utilizando partições com tamanho 16 obteve os melhores resultados em relação aos demais sistemas. Somente no ponto 0.87 o CBIR-I foi o melhor, nos demais pontos o CBIR-Indice-Particionado foi melhor em relação aos demais.

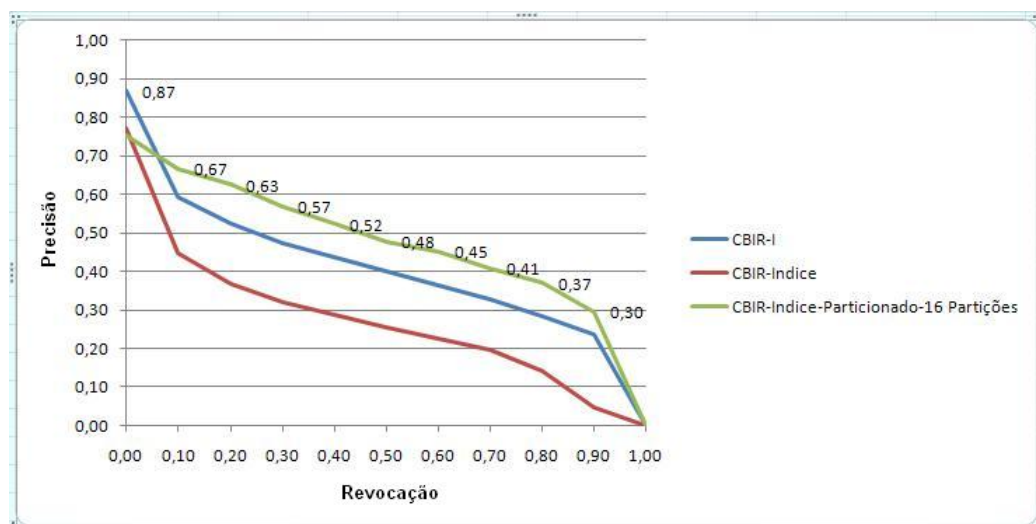


Figura 4.5: Curva 11-pt dos sistemas CBIR-I, CBIR-Indice e CBIR-Particionado utilizando a característica cor.

Avaliação Experimental Utilizando as Características Cor, Textura e Forma

Nos experimentos descritos nesta seção as três características (cor, forma e textura) foram utilizadas na indexação e recuperação das imagens.

A Tabela 4.21 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-I. O banco de dados utilizado foi o Corel-1000 relativo à média de 10 consultas de cada classe, juntamente com o valor médio da precisão de 100 consultas da coleção. Já na Tabela 4.22 encontram-se estes mesmos resultados para os sistemas CBIR-Indice.

A Tabela 4.23 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-Indice-Particionado.

Novamente, os resultados obtidos com o sistema CBIR-Indice-Particionado continuam tendo resultados superiores aos demais sistemas em análise, mesmo depois das caracterís-

Tabela 4.21: Precisão média no sistema CBIR-I utilizando as características cor, forma e textura. Coleção Corel-1000

Classe	P5	P10	P20	P100
Africa	70,00%	60,00%	57,50%	33,50%
Cavalo	80,00%	55,00%	57,50%	40,50%
Comida	60,00%	40,00%	45,00%	34,00%
Construção	80,00%	75,00%	67,50%	36,50%
Dinossauro	100,00%	100,00%	100,00%	93,00%
Elefante	90,00%	90,00%	70,00%	40,50%
Flor	50,00%	50,00%	55,00%	28,50%
Montanha	10,00%	10,00%	10,00%	12,00%
Ônibus	90,00%	75,00%	82,50%	50,00%
Praia	50,00%	75,00%	62,50%	28,00%
Média	68,00%	63,00%	60,75%	39,65%

Tabela 4.22: Precisão média no sistema CBIR-Índice utilizando as características cor, forma e textura. Coleção Corel-1000

Classe	P5	P10	P20	P100
Africa	58,00%	46,00%	40,00%	24,30%
Cavalo	74,00%	64,00%	52,00%	33,50%
Comida	48,00%	35,00%	32,00%	24,50%
Construção	50,00%	45,00%	40,50%	24,20%
Dinossauro	98,00%	99,00%	97,00%	76,60%
Elefante	66,00%	55,00%	42,00%	26,70%
Flor	64,00%	57,00%	51,50%	28,20%
Montanha	24,00%	24,00%	22,50%	17,10%
Ônibus	94,00%	79,00%	73,50%	46,20%
Praia	40,00%	37,00%	31,50%	18,60%
Média	61,60%	54,10%	48,25%	31,99%

Tabela 4.23: Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e as características cor, forma e textura. Coleção Corel-1000

Classe	P5	P10	P20	P100
Africa	70,00%	67,00%	61,00%	42,20
Cavalo	90,00%	90,00%	83,00%	58,40
Comida	50,00%	51,00%	51,50%	34,20
Construção	74,00%	68,00%	63,50%	43,40
Dinossauro	100,00%	100,00%	100,00%	96,70
Elefante	46,00%	39,00%	38,00%	25,90
Flor	82,00%	77,00%	75,00%	53,30
Montanha	36,00%	34,00%	27,50%	24,20
Ônibus	90,00%	87,00%	88,50%	68,70
Praia	44,00%	43,00%	39,00%	28,90
Média	68,20%	65,60%	62,70%	47,59%

tivas cor, forma e textura terem sido acrescentadas às proposta CBIR-I e CBIR-Índice. Ou seja, em várias posições de *ranking* os valores obtidos aplicando a proposta CBIR-Índice-Particionado foi maior do que os resultados obtidos com a proposta dos sistemas CBIR-I e CBIR-Índice.

Analisando os resultados nos pontos de *ranking* em geral é possível perceber que

o sistema CBIR-Índice-Particionado sempre obteve melhores resultados em relação aos demais sistemas.

A Tabela 4.24 apresenta o ganho obtido com o sistema CBIR-Índice-Particionado em relação aos outros sistemas, CBIR-I e CBIR-Índice na coleção Corel-1000 utilizando as características cor, textura e forma. Analisando em relação aos sistemas CBIR-I e CBIR-Índice, o CBIR-Índice-Particionado sempre demonstrou um ganho superior na qualidade da recuperação.

Tabela 4.24: Comparação do CBIR-Índice-Particionado utilizando 16 partições e as características cor, forma e textura com os sistemas CBIR-I e CBIR-Índice. Coleção Corel-1000

	P5	P10	P20	P100
Média Geral CBIR-I	68,00%	63,00%	60,75%	39,65%
Média Geral CBIR-Índice	61,60%	54,10%	48,25%	31,99%
Média Geral CBIR-Índice-Particionado-16	68,20%	65,60%	62,70%	47,59%
Razão CBIR-Índice-Particionado-16/CBIR-I	0,29%	4,13%	3,21%	20,03%
Razão CBIR-Índice-Particionado-16/CBIR-Índice	10,71%	21,26%	29,95%	48,77%

Os conceitos de *precision/recall* também foram realizados para a análise dos resultados. As médias dos valores obtidos foram usadas para gerar curvas *precision x recall* mostradas na Figura 4.6.

Como pode ser visto nos gráficos, os resultados obtidos com o sistema CBIR-Traducional-Particionado utilizando partições com tamanho 16 obteve os melhores resultados em relação aos demais sistemas. Somente no ponto 0.85 o CBIR-I foi o melhor, nos demais pontos o CBIR-Índice-Particionado foi melhor em relação aos demais.

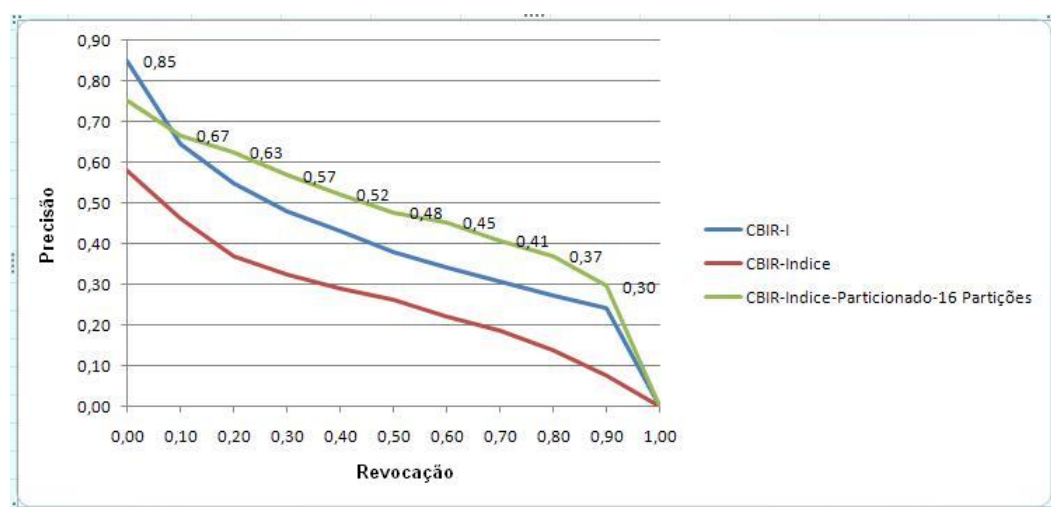


Figura 4.6: Curva 11-pt dos sistemas CBIR-I, CBIR-Índice e CBIR-Particionado utilizando as características cor, forma e textura.

4.2.3 Avaliação Experimental BD-10000

A coleção BD-10000 foi utilizada nesta seção para comparar os mesmos sistemas comparados nas seções 4.2.1 e 4.2.2.

A coleção BD-10000 é uma coleção de imagens reais cobrindo uma ampla variedade de categorias semânticas. Dentre estas categorias três estão classificadas: avião, carro e moto. Vinte imagens de consulta de cada classe foram escolhidas aleatoriamente.

Em alguns experimentos, consideramos somente a característica cor para a indexação e recuperação das imagens, nos demais experimentos consideramos as três características (cor, forma e textura), as quais foram implementadas no sistema CBIR-Índice-Particionado e nos demais sistemas utilizados neste trabalho.

Avaliação Experimental Utilizando a Característica Cor

Nos experimentos descritos nesta seção, somente a característica cor foi utilizada na indexação e recuperação das imagens.

A Tabela 4.25 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-I. O banco de dados utilizado foi o BD-10000 relativo a média de 20 consultas de cada classe, juntamente com o valor médio da precisão de 60 consultas da coleção. Já na Tabela 4.26 se encontra estes mesmos resultados para os sistemas CBIR-Índice.

A Tabela 4.27 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-Índice-Particionado.

Tabela 4.25: Precisão média no sistema CBIR-I utilizando a característica cor. Coleção BD-10000

Classe	P5	P10	P20	P100
Avião	77,5%	71,3%	68,8%	42,5%
Carro	62,5%	58,8%	56,3%	46,3%
Moto	72,5%	57,5%	46,9%	30,8%
Média Geral	70,80%	62,50%	57,30%	39,90%

Tabela 4.26: Precisão média no sistema CBIR-Índice utilizando a característica cor. Coleção BD-10000

Classe	P5	P10	P20	P100
Avião	75%	73,8%	62,5%	41,6%
Carro	62,5%	55%	49,4%	38,4%
Moto	60%	41,3%	38,8%	29,5%
Média Geral	65,80%	56,70%	50,20%	36,50%

Nota-se através destes resultados que as qualidades de recuperação continuam bem similares aos resultados encontrados nas seções anteriores.

A Tabela 4.28 apresenta uma análise do ganho obtido com o sistema CBIR-Índice-Particionado em relação aos outros sistemas, CBIR-I e CBIR-Índice na coleção BD-10000.

Tabela 4.27: Precisão média no sistema CBIR-Indice-Particionado utilizando 16 partições e a característica cor. Coleção BD-10000

Classe	P5	P10	P20	P100
Avião	82%	76,5%	72,26%	42,40%
Carro	66,80%	64,70%	61,60%	46,90%
Moto	69,70%	50,8%	39,40%	38%
Média Geral	72,83%	64,00%	57,75%	42,43%

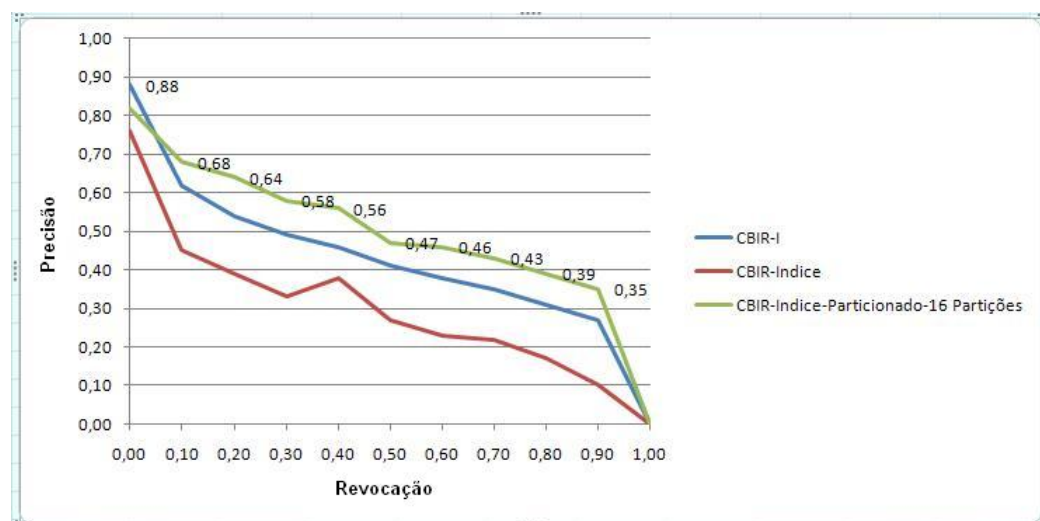
Em ambos os sistemas CBIR-I e CBIR-Indice, o CBIR-Indice-Particionado sempre demonstrou um ganho superior na qualidade da recuperação.

Tabela 4.28: Comparação do CBIR-Indice-Particionado utilizando 16 partições e a característica cor com os sistemas CBIR-I e CBIR-Indice. Coleção BD-10000

	P5	P10	P20	P100
Média Geral CBIR-I	70,80%	62,50%	57,30%	39,90%
Média Geral CBIR-Indice	65,80%	56,70%	50,20%	36,50%
Média Geral CBIR-Indice-Particionado	72,83%	64,00%	57,75%	42,43%
Razão CBIR-Indice-Particionado/CBIR-I	2,87%	2,40%	0,79%	6,34%
Razão CBIR-Indice-Particionado/CBIR-Indice	10,68%	12,87%	15,04%	16,25%

Na análise dos resultados os conceitos de *precision/recall* também foram utilizados. As médias dos valores obtidos foram usadas para gerar curvas *precision x recall* mostradas na Figura 4.7.

Como pode ser visto nos gráficos, os resultados obtidos com o sistema CBIR-Indice-Particionado utilizando partições com tamanho 16 obteve os melhores resultados em relação aos demais sistemas. Somente no ponto 0.88 o CBIR-I foi o melhor, nos demais pontos o CBIR-Indice-Particionado foi melhor em relação aos demais.

Figura 4.7: Curva *precision x recall* dos sistemas CBIR-I, CBIR-Indice e CBIR-Particionado utilizando a característica cor.

Avaliação Experimental Utilizando as Características Cor, Textura e Forma

Nos experimentos descritos nesta seção, as três características (cor, forma e textura) foram utilizadas na indexação e recuperação das imagens.

A Tabela 4.29 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-I. O banco de dados utilizado foi o BD-10000 relativo a média de 20 consultas de cada classe, juntamente com o valor médio da precisão de 60 consultas da coleção. Já na Tabela 4.30 se encontra estes mesmos resultados para os sistemas CBIR-Índice.

A Tabela 4.31 mostra os valores encontrados das precisões P5, P10, P20 e P100 do sistema CBIR-Índice-Particionado.

Tabela 4.29: Precisão média no sistema CBIR-I utilizando as características cor, forma e textura. Coleção BD-10000

Classe	P5	P10	P20	P100
Avião	79,9%	73%	72,6%	44,5%
Carro	64,8%	59,9%	58,6%	47,8%
Moto	74,6%	57,5%	47,8%	30,6%
Média Geral	73,1%	63,46%	59,66%	40,96%

Tabela 4.30: Precisão média no sistema CBIR-Índice utilizando as características cor, forma e textura. Coleção BD-10000

Classe	P5	P10	P20	P100
Avião	80,5%	69,7%	73%	40,8%
Carro	65,4%	70%	54,6%	43,5%
Moto	68%	45,8%	40,7%	35,8%
Média Geral	71,3%	61,83%	56,1%	40,03%

Tabela 4.31: Precisão média no sistema CBIR-Índice-Particionado utilizando 16 partições e as características cor, forma e textura. Coleção BD-10000

Classe	P5	P10	P20	P100
Avião	92,10%	86,50%	70,85%	37,30%
Carro	67,80%	79,70%	70,38%	56,90%
Moto	72,60%	52,35%	46,40%	46,80%
Média Geral	77,50%	72,85%	62,54%	47,00%

Novamente, os resultados obtidos com o sistema CBIR-Índice-Particionado continuam sendo superiores aos demais sistemas em análise, mesmo depois das características cor, forma e textura terem sido acrescentadas nas propostas CBIR-I e CBIR-Índice. Ou seja, na maioria das posições de *ranking* os valores obtidos aplicando a proposta CBIR-Índice-Particionado foi maior do que os resultados obtidos com a proposta dos sistemas CBIR-I e CBIR-Índice.

A Tabela 4.32 apresenta uma análise do ganho obtido com o sistema CBIR-Índice-Particionado em relação aos outros sistemas, CBIR-I e CBIR-Índice na coleção BD-10000. Analisando em relação aos sistemas CBIR-I e CBIR-Índice, o CBIR-Índice-Particionado

sempre demonstrou um ganho superior na qualidade da recuperação em relação aos demais.

Tabela 4.32: Comparação do CBIR-Indice-Particionado utilizando 16 partições com os sistemas CBIR-I e CBIR-Indice. Coleção BD-10000

	P5	P10	P20	P100
Média Geral CBIR-I	73,10%	63,46%	59,66%	40,96%
Média Geral CBIR-Indice	71,30%	61,83%	56,10%	40,03%
Média Geral CBIR-Indice-Particionado	77,50%	72,85%	62,54%	47,00%
Razão CBIR-Indice-Particionado-16/CBIR-I	6,02%	14,85%	4,83%	14,75%
Razão CBIR-Indice-Particionado-16/CBIR-Indice	8,70%	17,82%	11,48%	17,41%

Na análise dos resultados os conceitos de *precision/recall* também foram utilizados. As médias dos valores obtidos foram usadas para gerar curvas *precision x recall* mostradas na Figura 4.8.

Como pode ser visto nos gráficos, os resultados obtidos com o sistema CBIR-Traducional-Particionado utilizando partições com tamanho 16 obteve os melhores resultados em relação aos demais sistemas. Aqui, inclusive no ponto 0.87 de precisão, o CBIR-Indice-Particionado obteve bom resultado, equiparado ao CBIR-I.

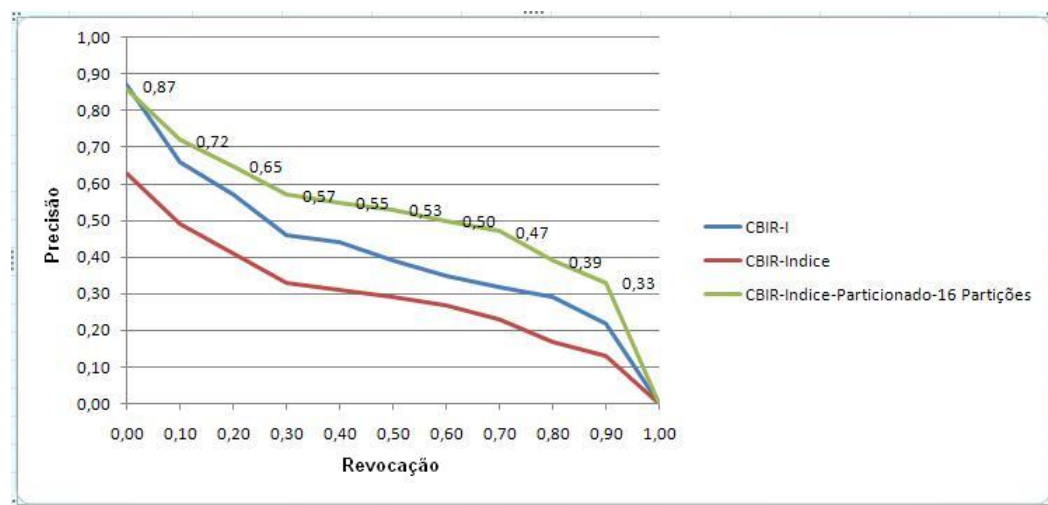


Figura 4.8: Curva *precision x recall* dos sistemas CBIR-I, CBIR-Indice e CBIR-Particionado utilizando as características cor, forma e textura.

Pode-se concluir então que, a abordagem utilizando particionamento contribui para o melhoramento da qualidade de recuperação comparada com a abordagem sem particionamento, o CBIR-Indice. O CBIR-Indice-Particionado, apresentou na maioria dos testes uma qualidade de recuperação superior ao CBIR-I.

4.3 Comparação do número de operações Aritméticas

Em [Matos et al., 2008a] foi feita uma comparação dos sistemas CBIR-I e CBIR-Índice em relação ao número de operações aritméticas efetuadas para o cálculo da similaridade, entre a imagem de consulta e a base de dados.

No sistema CBIR-I, de acordo com a equação da distância Euclidiana, e o vetor de característica da Tabela 5.1, para cada consulta são feitas as seguintes operações: i) 9 subtrações; ii) 9 multiplicações; iii) 8 adições; iv) 1 raiz quadrada. Como mencionado anteriormente o CBIR-I, faz a comparação do vetor de consulta com todas as imagens da coleção, com o objetivo de se obter o *ranking* de similaridade final. Tendo a base de dados 1000 imagens ao final de uma consulta totaliza-se 27.000 operações aritméticas [Matos et al., 2008a].

No CBIR-Índice, por meio da lista invertida não é necessário varrer toda coleção, mas somente aquelas que casarem com pelo menos um termo da consulta. Para cada característica (termo) presente faremos as seguintes operações: i) $z * 9$ somas, em que z é o tamanho médio das listas invertidas; ii) y divisões, onde y é o número total de imagens encontradas nas listas. Nos experimentos z e y tiveram valores médios de 1038 adições e 677 divisões respectivamente. Totalizando 1715 operações aritméticas por consulta. Valor bem abaixo dos 27.000 encontrados no CBIR-I.

No sistema CBIR-Índice-Particionado, o número de operações é semelhante ao sistema CBIR-Índice, porém pelo fato de particionar as imagens antes da indexação aumentou o tamanho das listas invertidas. A quantidade é relativamente pequena em relação ao alto número de operações do sistema CBIR-I. Nos experimentos, os valores de z e y descritos anteriormente tiveram valores médios de 3633 adições e 2572. Totalizando 6025 operações aritméticas por consulta. Mesmo o valor sendo 3.5 vezes maior do que o número de operações do sistema CBIR-Índice, o valor corresponde a aproximadamente 4.5 vezes menor do que o número de operações encontradas no CBIR-I. Mostrando portanto um ganho de desempenho.

Essas análises foram feitas considerando somente a característica cor para cada um dos sistemas. Quando analisamos as demais, o número de operações crescem nos sistemas CBIR-Índice e CBIR-Índice-Particionado, mas o valor acrescido ainda continua sendo inferior ao CBIR-I. Basicamente é duas vezes menor a proporção entre os sistemas CBIR-I e CBIR-Índice-Particionado.

4.4 Considerações Finais

Neste capítulo se verificou que o método de recuperação de imagens proposto produziu bons resultados quando comparados com os resultados de um método tradicional de recuperação de imagens CBIR-I e do método CBIR-Índice.

Foram mostradas seis avaliações experimentais. As duas primeiras foram utilizadas a coleção BD-110 como treinamento e definição do número de partições para os testes. Nesta avaliação, pode-se notar que a qualidade de recuperação do sistema proposto no topo do *Ranking* é superior ao sistema CBIR tradicional. Já na segunda e terceira avaliação, foi utilizada a coleção de teste Corel-1000 e BD-10000, respectivamente. Os resultados nestas coleções se mantiveram similar ao da coleção de treinamento. Na primeira posição do *ranking* o sistema proposto apresenta melhor qualidade de recuperação e a medida que se avança no *ranking* o sistema proposto perde um pouco na qualidade.

Para compensar a perda da qualidade em algumas posições do *ranking*, foi mostrado que a abordagem proposta é bem mais rápida no cálculo da similaridade da imagem de consulta com a coleção de imagens, pelo fato de realizar menos operações aritméticas.

O próximo capítulo traz as considerações finais desta dissertação além das perspectivas futuras deste trabalho.

Capítulo 5

Conclusão e Trabalhos Futuros

Nos últimos anos, o crescimento do número de imagens produzidas em meio digital foi imenso, com isso tornou-se necessária a criação de mecanismos capazes de prover, de maneira rápida e eficaz, a informação requisitada pelo usuário. Nesse cenário, muitas vezes o usuário tem dificuldades em localizar imagens específicas, surge então a necessidade crescente por melhores algoritmos de busca, métodos de indexação e técnicas de classificação de imagens digitais.

Na tentativa de resolver o problema de recuperação de imagens, os sistemas de recuperação de imagens por conteúdo exploram as características de cor, forma e textura no cálculo da similaridade entre as imagens. O problema desta abordagem é que na maioria dos sistemas o cálculo da similaridade é realizado comparando-se a imagem de consulta com todas as imagens da coleção de referência.

Foi considerando a necessidade de recuperação de imagens em grandes coleções que esta dissertação estendeu uma abordagem que faz uso das técnicas tradicionais de Recuperação de Informação textual com a eficácia na qualidade de recuperação dos sistemas de recuperação de imagens baseado no conteúdo.

Neste trabalho foi apresentado um sistema de recuperação de imagens baseado no conteúdo. Considerando a indexação de características de baixo nível de partições de imagens digitais mapeadas para um índice invertido, este trabalho obteve melhorias no desempenho do processamento de consultas e ganho na precisão considerando o conjunto de imagens recuperadas em grandes bases de dados. O índice invertido, análogo ao utilizado em recuperação textual, foi adaptado para recuperação de imagens particionadas baseado em conteúdo.

Na determinação do número de partições a ser utilizado por este trabalho foi utilizada uma base de treinamento, a coleção BD-110. Assim como utilizando somente a característica cor, utilizando as características cor, forma e textura, em geral o número de partições igual a 16 obteve melhores resultados na coleção de treinamento. Nas demais coleções foi utilizado esse número de partição na execução dos testes.

Assim comparou-se o desempenho entre três sistemas. O primeiro utiliza a medida da

distância euclidiana para o cálculo da similaridade, o segundo calcula através da medida do cosseno, além de indexar as imagens através do índice invertido e o terceiro além de utilizar as mesmas características do segundo, utiliza o particionamento do domínio das imagens em diversas partições de mesmo tamanho e o acréscimo de duas novas características, forma e textura. Mostrou-se, que o terceiro sistema, tem um ganho significativo no tempo de processamento por realizar menos operações aritméticas, pois não há necessidade de se varrer toda a base de dados. Além disso esse sistema obteve um ganho significativo em relação a qualidade da recuperação, 10% em relação ao sistema CBIR-I e 40% com relação ao CBIR-Índice.

A principal contribuição deste trabalho foi propor o CBIR-Índice-Particionado que considera o uso de 16 partições da imagem e momentos obtidos das características cor, forma e textura dessas partições, indexando-as por meio de arquivo invertido e mostrar por meio de experimentos que essa proposta obteve ganho de qualidade, bom desempenho no tempo de processamento. Em relação ao CBIR-I obteve um ganho médio de 7% com a característica cor e 10% utilizando as características cor, forma e textura. E considerando o sistema CBIR-Índice obteve um ganho médio de 33% com a característica cor e 38% utilizando todas.

Como trabalhos futuros destacam-se:

- Implementar vários descritores de cor, forma e textura e utilizar a abordagem de algoritmos para decidir os melhores ou a importância de cada um destes no contexto de uma busca particular do usuário, utilizando o índice invertido.
- Realizar uma série de experimentos com usuários reais, com o intuito de verificar e comparar a performance dos sistemas.
- Realizar novas adaptações da recuperação de informação textual para recuperação de imagens, por exemplo algoritmo de poda de listas invertidas.
- Realizar comparações com abordagens que utilizam bag-of-features.
- Implementação da técnica LSI para facilitar a busca.

Referências Bibliográficas

- [Aslandogan and Yu, 2000a] Aslandogan, Y. A. and Yu, C. T. (2000a). Diogenes:a web search agent for content based indexing of personal images. In *In Proceedings of the ACM SIGIR*. ACM.
- [Aslandogan and Yu, 2000b] Aslandogan, Y. A. and Yu, C. T. (2000b). Evaluating strategies and systems for content based indexing of person images on the web. In *MULTIMEDIA '00: Proceedings of the eighth ACM international conference on Multimedia*, pages 313–321, New York, NY, USA. ACM.
- [Baeza-Yates and Ribeiro-Neto, 1999] Baeza-Yates, R. A. and Ribeiro-Neto, B. A. (1999). *Modern Information Retrieval*. Addison-Wesley. Essex, UK.
- [Brandt, 1999] Brandt, S. (1999). Use of shape features in content-based image retrieval. Master’s thesis, Master’s thesis, Helsinki University of Technology, Department of Engineering Physics e Mathematics, Espoo, Finland.
- [Bueno, 2002] Bueno, J. M. (2002). Suporte à recuperação de imagens médicas baseada em conteúdo através de histogramas métricos. Tese de doutorado, Universidade de São Paulo.
- [CalPhotos, 2010] CalPhotos (2010). University of california, berkeley.
- [Cardoso et al., 2007] Cardoso, N., Martins, B., Gomes, D., and Silva, M. J. (2007). *WPT 03: a primeira coleção pública proveniente de uma recolha da web portuguesa*, chapter 23. IST Press.
- [Cardoso, 2000] Cardoso, O. (2000). Recuperação de informação. *INFOCOMP Journal of Computer Science*, 2(1):33–38.
- [Csurka et al., 2004] Csurka, G., Dance, C. R., Fan, L., Willamowski, J., and Bray, C. (2004). Visual categorization with bags of keypoints. In *In Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22.
- [da Silva, 1999] da Silva, S. F. (1999). Realimentação de relevância via algoritmos genéticos aplicada na recuperação de imagens. Master’s thesis, Dissertação de Mestrado, Faculdade de Computação, Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil.
- [Database, 2010] Database, C. (2010). Corel corporation, corel gallery 3.0.
- [Datta et al., 2008] Datta, R., Dhiraj Joshi, J. L., and Wang, J. Z. (2008). Image retrieval: Ideas, influences, and trends of the new age. *ACM Transactions on Computing Surveys*, 40(2):1–60.

- [Faloutsos, 1985] Faloutsos, C. (1985). Access methods for text. *ACM Comput. Surv.*, 17(1):49–74.
- [Flickner et al., 1995] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., and Yanker, P. (1995). Query by image and video content: the qbic system. *Computer*, 28(9):23–32.
- [Frakes and Baeza-Yates, 1992] Frakes, W. B. and Baeza-Yates, R. A., editors (1992). *Information Retrieval: Data Structures & Algorithms*. Prentice-Hall.
- [Gonzales and Woods, 2002] Gonzales, R. C. and Woods, R. E. (2002). *Digital Image Processing*, volume Second Edition. Pretenci Hall.
- [Google, 2010] Google (2010).
- [Gwizdka et al., 1999] Gwizdka, J., , Gwizdka, J., and Chignell, M. (1999). Towards information retrieval measures for evaluation of web search engines.
- [Huang and Zabih, 1997] Huang, J., K.-S. R. M. M. Z. W. and Zabih, R. (1997). Image indexing using color correlogram. In *IEEE International Conference on Computer Vision and Evolutionary Computation and Pattern Recognition, Puerto Rico*.
- [Jiang et al., 2007] Jiang, Y.-G., Ngo, C.-W., and Yang, J. (2007). Towards optimal bag-of-features for object categorization and semantic video retrieval. In *CIVR '07: Proceedings of the 6th ACM international conference on Image and video retrieval*, pages 494–501, New York, NY, USA. ACM.
- [Kherfi et al., 2004] Kherfi, M. L., Ziou, D., and Bernardi, A. (2004). Image retrieval from the world wide web: Issues, techniques, and systems. *ACM Comput. Surv.*, 36(1):35–67.
- [Koskela, 1999] Koskela, M. (1999). Content-based image retrieval with self-organizing maps. Master’s thesis, PhD thesis, Helsinki University of Technology.
- [Laaksonen, 2000] Laaksonen, J. (2000). Analyzing low-level visual features using content-based image retrieval. In *Proceedings of the 7th International Conference on Neural Information Processing, Teajon, Korea*.
- [Lew, 2001] Lew, M. S., editor (2001). *Principles of visual information retrieval*. Springer-Verlag, London, UK.
- [Long et al., 2000] Long, H. Z., Leow, W. K., and Chua, F. K. (2000). Perceptual texture space for content-based image retrieval. In *In Proc. Int. Conf. on Multimedia Modeling (MMM 2000*, pages 167–180.
- [Lopes et al., 2009a] Lopes, A. P. B., Avila, S. E. F. d., Peixoto, A. N. A., Oliveira, R. S., Coelho, M. d. M., and Araújo, A. d. A. (2009a). Nude detection in video using bag-of-visual-features. In *SIBGRAPI '09: Proceedings of the 2009 XXII Brazilian Symposium on Computer Graphics and Image Processing*, pages 224–231, Washington, DC, USA. IEEE Computer Society.
- [Lopes et al., 2009b] Lopes, A. P. B., Oliveira, R. S., Almeida, J. M. d., and Araújo, A. d. A. (2009b). Spatio-temporal frames in a bag-of-visual-features approach for human

- actions recognition. In *SIBGRAPI '09: Proceedings of the 2009 XXII Brazilian Symposium on Computer Graphics and Image Processing*, pages 315–321, Washington, DC, USA. IEEE Computer Society.
- [Manning, 2007] Manning, C., R. P. S. H. (2007). *An Introduction to information Retrieval*. Cambridge University Press, Cambridge, England.
- [Matos et al., 2008a] Matos, T., Silva, I. R., Barcelos, C., and Proença, P. (2008a). Avaliação de Índice invertido em busca de imagens por conteúdo. *XXXIV Conferencia Latinoamericana de Informática, 2008, Santa Fe-Argentina. Anais da XXXIV Conferencia Latinoamericana de Informática. Buenos Aires-Argentina : Sociedade Argentina de Informatica*, 1:1–10.
- [Matos et al., 2008b] Matos, T., Silva, I. R., Barcelos, C. A. Z., and Proença, P. (2008b). *Índice Invertido para Recuperação de Imagens Baseada em Conteúdo*.
- [Niblack et al., 1993] Niblack, C. W., Barber, R. J., Equitz, W. R., Flickner, M. D., Glasman, D., Petkovic, D., and Yanker, P. C. (1993). The qbic project: Querying image by content using color, texture, and shape. pages 173–187.
- [Nowak et al., 2006a] Nowak, E., Jurie, F., and Triggs, B. (2006a). Sampling strategies for bag-of-features image classification. In *European Conference on Computer Vision*. Springer.
- [Nowak et al., 2006b] Nowak, E., Jurie, F., and Triggs, B. (2006b). Sampling strategies for bag-of-features image classification. In *Proceedings of the European Conference on Computer Vision*. Springer.
- [Pass and Zabith, 1996] Pass, G. and Zabith, R. (1996). Histogram refinement for content-based image retrieval. In *IEEE Workshop on Applications of Computer Vision*, pages 96–102.
- [Penatti and Torres, 2007] Penatti, O. B. and Torres, R. S. (2007). Descritores de relacionamento espacial baseado em partições. *Revista Eletrônica de Iniciação Científica*, VII(3).
- [Ramos, 2004] Ramos, F. R. (2004). Recuperação de informação baseada em conteúdo: Analisando imagens priorizando a característica cor. Master's thesis, Programa de Pós-Graduação em Informática Aplicada da PUC-Paraná.
- [Rodden, 1999] Rodden, K. (1999). How do people organise their photographs? In *Proceedings of the BCS IRSG Colloquium*.
- [Rui et al., 1997] Rui, Y., Huang, T. S., and fu Chang, S. (1997). Image retrieval: Past, present, and future. In *Journal of Visual Communication and Image Representation*, pages 1–23.
- [Rui et al., 1999] Rui, Y., Huang, T. S., and fu Chang, S. (1999). Image retrieval: Current techniques, promising directions and open issues. *Journal of Visual Communication and Image Representation*, 10:39–62.

- [Safar et al., 1999] Safar, M., Shahabi, C., and Sun, X. (1999). Image retrieval by shape: A comparative study. In *In Proceedings of IEEE International Conference on Multimedia and Exposition ICME*, pages 141–144.
- [Salton and McGill, 1986] Salton, G. and McGill, M. J. (1986). *Introduction to Modern Information Retrieval*. McGraw-Hill, Inc., New York, NY, USA.
- [Santos, 2003] Santos, K. S. (2003). Dependência entre termos no modelo vetorial. Master’s thesis, Dissertação de Mestrado, Universidade Federal de Uberlândia, Departamento de Ciência da Computação, Uberlândia, Minas Gerais, Brasil.
- [Sclaroff et al., 1997] Sclaroff, S., Taycher, L., and Cascia, M. L. (1997). Imagerover: A content-based image browser for the world wide web. In *In Proc. IEEE Workshop on Content-based Access of Image and Video Libraries*, pages 2–9.
- [Shneiderman and Kang, 2000] Shneiderman, B. and Kang, H. (2000). Direct annotation: A drag-and-drop strategy for labeling photos. In *CiteSeerX*, pages 88–95.
- [Stricker and Orengo, 1995] Stricker, M. and Orengo, M. (1995). Similarity of color images. In *Proceedings of IS&T and SPIE Storage and Retrieval of Image and Video Databases III*.
- [Swain and Ballard, 1991] Swain, M. J. and Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, 7:11–32.
- [Traina, 2001] Traina, A. J. M. (2001). Suporte à visualização de consultas por similaridade em imagens médicas através de estrutura de indexação métrica. Master’s thesis, Tese de livre docência apresentada ao ICMC-USP.
- [Zhang, 2002] Zhang, D. (2002). *Image Retrieval Based on Shape*. PhD thesis, Faculty of Information Technology Monash University.
- [Zhang and Lu, 2004] Zhang, D. and Lu, G. (2004). Review of shape representation and description techniques. *Pattern Recognition*, 37(1):1–19.
- [Zobel and Moffat, 1998] Zobel, S. and Moffat, A. (1998). Exploring the similarity spacial. In *ACM SIGIR Forum*.