

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

**Indexação e Recuperação de Imagens
por Cor e Estrutura**

por

YANDRE MALDONADO E GOMES DA COSTA

Dissertação submetida à avaliação,
como requisito parcial para a obtenção do grau de Mestre
em Ciência da Computação

Prof. Dr. Jacob Scharcanski
Orientador

Porto Alegre, julho de 2002.

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Costa, Yandre Maldonado e Gomes da

Indexação e Recuperação de Imagens por Cor e Estrutura / por Yandre Maldonado e Gomes da Costa. – Porto Alegre: PPGC da UFRGS, 2002.

77 f.: il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR-RS, 2002. Orientador: Scharcanski, Jacob.

1. Indexação de Imagens. 2. Recuperação de Imagens. 3. Cor e Estrutura. I. Scharcanski, Jacob. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora: Profa. Wrana Panizzi

Pró-Reitor de Ensino: Prof. José Carlos Ferraz Hennemann

Pró-Reitor Adjunto de Pós-Graduação: Prof. Jaime Evaldo Fensterseifer

Diretor do Instituto de Informática: Prof. Philippe Olivier Alexandre Navaux

Coordenador do PPGC: Prof. Carlos Alberto Heuser

Bibliotecária-Chefe do Instituto de Informática: Beatriz Bastos Haro

Aos meus pais Sebastião e Lucia.

Agradecimentos

Agradeço em primeiro lugar a Deus, sem o qual este trabalho não teria sido realizado. Foi Ele que me capacitou e me deu forças para lutar até o último instante.

Em segundo lugar, agradeço à minha família, em especial aos meus pais, Sebastião e Lúcia, que sempre apostaram em mim. É a eles que dedico este trabalho.

Gostaria de agradecer também ao meu orientador, o Prof. Dr. Jacob Scharcanski, que, apesar das dificuldades impostas pela distância, sempre esteve disposto a contribuir para que este trabalho se realizasse.

Agradeço ainda a todas as pessoas que contribuíram direta ou indiretamente para com a realização deste trabalho. Dentre elas agradeço em especial àquelas que estiveram intercedendo através de suas orações para que este trabalho pudesse ser concluído, e ao amigo Ailton Sergio Bonifácio, companheiro em boa parte das horas mais difíceis enfrentadas desde o início desta jornada. Agradeço também aos amigos Luciano, pela presteza em colaborar como pôde; Álvaro e Giuliana, pela hospitalidade; Heitor, pelo apoio técnico; e à Carolina que apareceu no “apagar das luzes” me incentivando e inspirando.

Finalmente, gostaria de agradecer à Universidade Paranaense – UNIPAR, e à Faculdade de Filosofia, Ciências e Letras de Mandaguari – FAFIMAN, pelo apoio prestado à realização deste trabalho.

Sumário

Lista de Abreviaturas	07
Lista de Figuras	08
Lista de Tabelas	10
Resumo	11
Abstract	12
1 Introdução	13
2 Recuperação de imagens	15
2.1 Indexação de Imagens	16
2.1.1 Cor	16
2.1.2 Layout de Cor	16
2.1.3 O uso da cor em recuperação de imagens	17
2.1.4 Integração de aspectos cromático e estrutural	18
2.2 Descritores de conteúdo da imagem	18
2.2.1 Descrição de conteúdo cromático	19
2.2.2 Descrição de conteúdo estrutural	21
2.3 Conclusões	22
3 Recuperação de imagens pela distância euclidiana ponderada	23
3.1 Medida de similaridade entre imagens	23
3.1.1 O uso de diferentes espaços de cores	24
3.1.2 Espaço de cores RGB	25
3.1.3 Espaço de cores XYZ	26
3.1.4 Espaço de cores $L^*u^*v^*$	27
3.1.5 Espaço de cores $L^*a^*b^*$	27
3.1.6 As matrizes de distâncias entre cores $A_{RGB}, A_{LUV}, A_{LAB}$	28
3.1.7 A matriz de distâncias A_E	29
3.1.8 Uma medida de similaridade entre imagens por cor e estrutura	29
3.2 Esquema conceitual da recuperação de imagens por conteúdo	30
3.3 Conclusões	31
4 Sistema RICE – Sistema para a Recuperação de Imagens por Cor e Estrutura	32
4.1 Conclusões	37
5 Resultados experimentais	38
5.1 Desempenho com padrões cromáticos e estruturais artificiais	38
5.2 Desempenho com imagens reais	43
5.2.1 Avaliação de desempenho do sistema RICE	46
5.2.2 O uso de curvas de “recall x precision”	56
5.2.2.1 Variando o tamanho das zonas	56
5.2.2.2 Alterando o espaço de representação das cores	62
5.2.2.3 Alterando os pesos dos aspectos cromático e estrutural	65
5.3 Conclusões	72

6 Conclusão	73
6.1 Trabalhos Futuros	74
Bibliografia	75

Lista de Abreviaturas

CIE	Comission Internacionale L'Eclairage
DBMS	Data Base Management System
RGB	Red, Green, Blue
RICE	Recuperação de Imagem por Cor e Estrutura
UFRGS	Universidade Federal do Rio Grande do Sul

Lista de Figuras

FIGURA 2.1 - Divisão da imagem em zonas com $n=3$ e $m=3$	19
FIGURA 2.2 - Formação do histograma de cores a partir de um conjunto de pixels	20
FIGURA 2.3 - Formação do histograma de estrutura a partir de um conjunto de pixels..	22
FIGURA 3.1 - Extração de medidas de similaridade entre as zonas correspondentes das imagens que estão sendo comparadas.	24
FIGURA 3.2 - Imagem com: (a) 2^{24} bits cores; (b) 256 cores; (c) 16 cores; (d) monocromática (preto e branco)	25
FIGURA 3.3 - Espaço de cores RGB	26
FIGURA 3.4 – Esquema conceitual da recuperação de imagens	31
FIGURA 4.1 - Interface principal do sistema RICE	33
FIGURA 4.2 - Menu “configuração”	33
FIGURA 4.3 - Janela que exhibe as imagens cadastradas no banco de dados.....	34
FIGURA 4.4 - Janela para cadastro de padrões de descritores de imagem.....	35
FIGURA 4.5 - Menu processamento	35
FIGURA 4.6 - Janela que permite a comparação entre duas imagens	36
FIGURA 4.7 - Janela que permite a recuperação de uma imagem no repositório	37
FIGURA 5.1 - Padrões cromáticos e estruturais artificiais	38
FIGURA 5.2 - Exemplo de curva de “recall x precision”	45
FIGURA 5.3 - Foto de uma face humana e o conjunto de imagens relevantes durante a sua recuperação.....	47
FIGURA 5.4 - Foto de uma paisagem litorânea e o conjunto de imagens relevantes durante a sua recuperação.....	48
FIGURA 5.5 - Foto de uma paisagem urbana com predominância de edificações e o conjunto de imagens relevantes durante a sua recuperação	49
FIGURA 5.6 - Foto de uma paisagem retratando uma montanha e o conjunto de imagens relevantes durante a sua recuperação	50
FIGURA 5.7 - Foto de um bosque e o conjunto de imagens relevantes durante a sua recuperação.....	51
FIGURA 5.8 - Foto de uma onda no mar e o conjunto de imagens relevantes durante a sua recuperação.....	52
FIGURA 5.9 - Foto de um conjunto de nuvens e o conjunto de imagens relevantes durante a sua recuperação.....	53
FIGURA 5.10 - Foto de uma estrada entre árvores e o conjunto de imagens relevantes durante a sua recuperação.....	54
FIGURA 5.11 - Foto da água do mar e o conjunto de imagens relevantes durante a sua recuperação.....	55
FIGURA 5.12 - Foto de dunas e o conjunto de imagens relevantes durante a sua recuperação.....	56
FIGURA 5.13 – Curvas das consultas 1x1, com 50% de cor e 50% de estrutura utilizando o espaço de representação de cores RGB	57
FIGURA 5.14 – Imagem de consulta (1x1, 50% cor, 50% estrutura, RGB)	57
FIGURA 5.15 - As nove imagens mais similares recuperadas na consulta (1x1).....	58
FIGURA 5.16 - Curvas das consultas 3x3, com 50% de cor e 50% de estrutura utilizando o espaço de representação de cores RGB	59
FIGURA 5.17 – Imagem de consulta (3x3, 50% cor, 50% estrutura, RGB)	59
FIGURA 5.18 - As nove imagens mais similares recuperadas na consulta (3x3).....	60

FIGURA 5.19 - Curvas das consultas 10x10, com 50% de cor e 50% de estrutura utilizando o espaço de representação de cores RGB	61
FIGURA 5.20 – Curvas representantes das médias das consultas 1x1, 3x3 e 10x10	61
FIGURA 5.21 – Curvas das consultas 3x3, com 100% do aspecto cromático e espaço de cores RGB.....	63
FIGURA 5.22 - Curvas das consultas 3x3, com 100% do aspecto cromático e espaço de cores $L^* a^* b^*$	63
FIGURA 5.23 - Curvas das consultas 3x3, com 100% do aspecto cromático e espaço de cores $L^* u^* v^*$	64
FIGURA 5.24 - Curvas representantes das médias das consultas 3x3, com 100% do aspecto cromático para os espaços de representação de cores RGB, $L^* a^* b^*$ e $L^* u^* v^*$	65
FIGURA 5.25 – Curvas das consultas 3x3 com 100% do aspecto estrutural	66
FIGURA 5.26 – Curvas das médias das consultas 3x3 com 100% do aspecto cromático, 50% para cromático e 50% para estrutural, e 100% para estrutural	67
FIGURA 5.27 – Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem das dunas	68
FIGURA 5.28 – Curvas obtidas na consulta da imagem das dunas com pesos específicos para cada zona da imagem	69
FIGURA 5.29 - Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem da onda.....	69
FIGURA 5.30 - Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem das praias	70
FIGURA 5.31 - Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem das prédios	70
FIGURA 5.32 – Curva da consulta média da imagem da onda com pesos específicos para as zonas comparada com outras configurações de consulta para a mesma imagem	70
FIGURA 5.33 - Curva da consulta média da imagem da praia com pesos específicos para as zonas comparada com outras configurações de consulta para a mesma imagem	71
FIGURA 5.34 - Curva da consulta média da imagem da prédios com pesos específicos para as zonas comparada com outras configurações de consulta para a mesma imagem	71

Lista de Tabelas

TABELA 5.1 - Resultados obtidos em comparações entre imagens artificiais.....39

Resumo

Este trabalho descreve um conjunto de técnicas para a recuperação de imagens baseada nos aspectos cromático e estrutural das mesmas. A abordagem aqui descrita utiliza mecanismos que permitem a preservação de informação espacial referente aos conteúdos extraídos da imagem de forma que a sua precisão possa ser ajustada de acordo com a necessidade da consulta. Um outro importante aspecto aqui considerado, é a possibilidade de se optar por um dos seguintes espaços de cores para a verificação de distâncias entre cores no momento da recuperação: RGB, $L^*u^*v^*$, ou $L^*a^*b^*$. Com estas diferentes possibilidades de espaços de cores, será verificada a influência que os mesmos podem provocar no processo de recuperação de imagens baseado em aspectos cromáticos. O conjunto de técnicas para a recuperação de imagens abordadas neste trabalho levou à construção do sistema RICE, um ambiente computacional através do qual pode-se realizar consultas a partir de um repositório de imagens. Para a verificação do desempenho dos diferentes parâmetros ajustáveis na recuperação de imagens aqui descrita e implementada no sistema RICE, foram utilizadas curvas de “Recall x Precision”.

Palavras-chave: Indexação de Imagens, Recuperação de Imagens, Cor, Estrutura.

TITLE: “IMAGE INDEXING AND RETRIEVAL BY COLOR AND SHAPE”

Abstract

This work describes a set of image retrieval techniques by color and shape similarity. The approach presented here allows to preserve spacial relationships of the contents extracted from the image. And it can be adjusted accordingly to the query needs. Another important feature considered here, is the possibility of choosing between the RGB, $L^*u^*v^*$, and $L^*a^*b^*$ color spaces to compute color distances during the image retrieval operation. With these three options of color spaces, the influence of each one in the image retrieval process based in chromatic contents will be verified. The set of techniques for image retrieval described here led to development of the RICE system, a computational environment for image retrieval by color and shape similarity. Furthermore, the recall x precision graph was applied in order to verify the performance of the RICE system in several configuration modes of image retrieval.

Keywords: Image Indexing, Image Retrieval, Color, Shape.

1 Introdução

As imagens digitais têm sido aplicadas em um número de situações que cresce a cada dia. Isto se deve ao fato de que o formato digital tem se mostrado o formato ideal para o armazenamento e manipulação das mesmas [GOM 94]. Com este crescimento, surge naturalmente a necessidade de se estabelecer técnicas que permitam uma melhor manipulação destes conteúdos. Uma das questões que vêm chamando a atenção de pesquisadores já há algum tempo é a criação de descritores de imagens. Inicialmente, muitos trabalhos surgiram propondo a descrição textual de imagens. Estas propostas logo se depararam com duas dificuldades. Uma é a necessidade de muito trabalho manual para representar imagens em forma de texto. A outra dificuldade, mais importante, resulta do conteúdo rico presente nas imagens e a subjetividade da percepção humana. Isto é, pessoas diferentes podem perceber características diferentes pertencentes ao conteúdo de uma mesma imagem. Considerando que a recuperação de imagens é uma das aplicações mais frequentes para estes descritores, a subjetividade da percepção e a imprecisão da descrição podem causar falhas incorrigíveis nos processos de recuperação posteriores [RUI 97].

Desta forma, a recuperação de imagens baseada em conteúdo surgiu para tentar superar estas dificuldades. Isto é, ao invés de serem descritas por palavras-chaves, as imagens passaram ser indexadas pelo seu próprio conteúdo visual, através de propriedades como cor, estrutura, textura, forma entre outras [PET 93].

Este trabalho descreve um conjunto de técnicas para a recuperação de imagens baseada em aspectos cromático e estrutural das mesmas. Os conteúdos cromático e estrutural das imagens são características imediatamente percebidas pelo observador humano, e desta forma são elementos importantes para diferenciá-las, este fator motivou o emprego destas características no desenvolvimento deste trabalho. Juntamente às atividades de pesquisa desenvolvidas neste trabalho, construiu-se o sistema RICE, um sistema para a recuperação de imagens por cor e estrutura no qual foram implementadas as técnicas de recuperação de imagens aqui abordadas. O Sistema RICE extrai histogramas de cor e bordas das imagens para compará-las através de uma métrica para a verificação da similaridade entre os mesmos, a Distância Euclidiana Ponderada. A representação de ambas as classes de características através de histogramas permite uma uniformidade de estruturas de dados de cor e orientação de bordas dentro do sistema [ASS 2000]. Além disso, o sistema é capaz de preservar informações espaciais relacionadas aos elementos cromáticos e estruturais, podendo o usuário ajustar a precisão destas informações determinando a divisão da mesma em zonas antes da formação de seus descritores. No desenvolvimento do trabalho houve também uma preocupação voltada ao espaço de representação de cores utilizado no cálculo de diferenças entre cores durante o processo de recuperação de imagens. De acordo com [BIM 99], o espaço de representação de cores pode influenciar significativamente nos resultados obtidos em muitas técnicas de recuperação de imagens. Desta forma, foram selecionados os espaços de representação de cores RGB, $L^*u^*v^*$ e $L^*a^*b^*$ para serem aplicados nas técnicas que serão aqui descritas. O espaço RGB é o espaço no qual a maioria das imagens digitais encontra-se originalmente descrita, os espaços $L^*a^*b^*$ e $L^*u^*v^*$ foram escolhidos por serem espaços perceptualmente uniformes. Assim, poderá ser feita a opção por qualquer um destes espaços para a realização dos cálculos de diferenças entre cores no processo de recuperação.

Na verificação do desempenho do sistema RICE foram utilizadas as curvas de “recall x precision”. Com isto, foram analisadas as influências de alguns fatores tais como preservação de informações espaciais, ponderação dos aspectos cromático e estrutural, e espaço de cores utilizado durante a consulta.

Este trabalho encontra-se organizado da seguinte forma: no capítulo 2 será discutida a recuperação de imagens envolvendo, principalmente, questões relacionadas ao seu contexto histórico e a criação de descritores para indexá-las a partir dos seus conteúdos cromático e estrutural. No capítulo 3 será descrito o processo de recuperação de imagens através do cálculo da Distância Euclidiana Ponderada para mensurar a similaridade entre imagens. Dentre outros aspectos a serem abordados, será descrita a possibilidade de se integrar os conteúdos cromático e estrutural e de se utilizar diferentes espaços para a representação das cores durante a recuperação. No capítulo 4 será apresentado o sistema RICE – um sistema para a recuperação de imagens baseado em cor e estrutura que foi construído a partir da implementação das técnicas de recuperação discutidas ao longo do trabalho.

No capítulo 5 serão descritos os resultados experimentais obtidos através de testes realizados com o sistema RICE a partir de um repositório de imagens com conteúdos semânticos heterogêneos [COH 99]. Esta avaliação se dará através da elaboração de curvas “recall x precision” e buscará, inclusive, verificar a diferença do desempenho da recuperação de imagens com o uso de diferentes espaços para a representação das cores (RGB, $L^*u^*v^*$ e $L^*a^*b^*$) e com o uso de diferentes aspectos visuais (cromático e estrutural). No capítulo 6 serão apresentados os resultados conclusivos do trabalho, e no capítulo 7 a bibliografia.

2 Recuperação de imagens

A recuperação de imagens tem sido um tópico de pesquisa que vem chamando a atenção da comunidade acadêmica nos últimos anos. Neste capítulo, serão introduzidos alguns aspectos relevantes que nortearam estas pesquisas ao longo dos últimos anos. Desta forma, pretende-se abordar a representação e extração de características de imagens, bases fundamentais da Recuperação de Imagens Baseada em Conteúdo e que podem posicionar o leitor acerca do tema.

Nos últimos anos observou-se um rápido aumento no tamanho das coleções de imagens digitais. Todos os dias, equipamentos civis e militares geram gigabytes de imagens, com uma quantidade enorme de informações. Mas, não se pode ter acesso a essas informações sem um sistema eficiente de busca e recuperação. Recuperação de imagens tem sido uma área de pesquisa ativa desde os anos 70, com o impulso de duas comunidades de pesquisa, o Gerenciamento de Bancos de Dados e a Visão Computacional. Inicialmente, estas duas áreas estudaram a Recuperação de Imagens de diferentes ângulos, baseadas em texto e conteúdo visual [RUI 97].

A recuperação de imagens baseada em texto parte da descrição textual das imagens e do uso de Sistemas de Gerenciamento de Bancos de Dados (DBMS) para fazer a recuperação das mesmas. Ao longo do tempo em que estas pesquisas se desenvolveram muitos avanços ocorreram, como modelagem de dados, indexação multi-dimensional, avaliação de consulta e outros. Entretanto, ainda existem duas dificuldades que não foram superadas pela recuperação de imagens baseada em descrição textual: uma é a necessidade de muita mão-de-obra para representar imagens em forma de texto. A outra dificuldade, mais importante, resulta do conteúdo rico nas imagens e a subjetividade da percepção humana. Isto é, pessoas diferentes podem perceber características diferentes pertencentes ao conteúdo de uma mesma imagem. A subjetividade da percepção e a imprecisão da descrição podem causar falhas incorrigíveis nos processos de recuperação posteriores [RUI 99].

Nos últimos dez anos, em decorrência do surgimento de coleções de imagens em grande escala, as duas dificuldades citadas anteriormente tornaram-se cada vez mais agudas. A Recuperação de Imagens Baseada em Conteúdo surgiu para tentar superar estas dificuldades. Isto é, ao invés de serem descritas por palavras-chaves, as imagens passaram a ser indexadas pelo seu próprio conteúdo visual, através de cor, textura, e forma entre outras características. Desde então, muitas técnicas neste sentido foram desenvolvidas e muitos sistemas de recuperação de imagens baseados em conteúdo visual foram construídos [VEN 2001]. Os avanços nesta área receberam contribuição principalmente da Visão Computacional.

A abordagem baseada em conteúdo visual estabeleceu um *framework* geral de Recuperação de Imagens a partir de uma nova perspectiva. Entretanto, existem ainda muitas questões abertas relacionadas à esta pesquisa a serem resolvidas antes que possam ser colocadas em prática. Com relação à Recuperação de Imagens, há a necessidade de pesquisar o que já foi obtido num passado recente e quais são as perspectivas para o futuro para que se possa lidar com aplicações convincentes [RUI 97].

Existem três bases fundamentais para Recuperação de Imagens Baseada em Conteúdo: extração de características, indexação e projeto de sistemas de recuperação. Neste capítulo, serão discutidos alguns aspectos relacionados à indexação de imagens por cor e estrutura.

2.1 Indexação de Imagens

A indexação de imagens envolve extração de características (conteúdo) das mesmas, e constitui a base da recuperação de imagens. Em um sentido amplo, características podem ser desde descrições textuais (palavras-chave, anotações), quanto características visuais (cor, textura, forma, faces). Contudo, já existem literaturas ricas em extração de características baseadas em texto nas áreas de Gerenciadores de Bancos de Dados (DBMS) e Recuperação de Imagens. Desta forma, a atenção deste trabalho está relacionada à extração de características visuais, relacionadas aos aspectos cromático e estrutural da imagem. Dentro do escopo de características visuais, pode-se classificá-las como características gerais e específicas do domínio. A primeira pode incluir cor, textura, estrutura e forma, enquanto a segunda é dependente da aplicação e pode incluir, por exemplo, faces humanas e impressões digitais. As características específicas do domínio são melhor estudadas nas literaturas de Reconhecimento de Padrões. Neste trabalho, será dado enfoque às características gerais que podem ser utilizadas em uma grande variedade de aplicações.

Devido à subjetividade da percepção, não existe uma única e melhor apresentação para uma dada característica. Assim, é importante salientar que para uma dada característica (como cor ou forma), existem muitas possíveis representações a partir de diferentes perspectivas. A seguir, serão descritos alguns importantes aspectos inerentes à indexação por conteúdo cromático e estrutural.

2.1.1 Cor

A cor é uma das características visuais mais usadas em recuperação de imagens. É independente do tamanho da imagem e da sua orientação. Em recuperação de imagens, Histograma de Cores é a representação mais usada para esta característica. Este tipo de descritor denota estatisticamente a probabilidade das intensidades de cores.

Algumas técnicas têm aplicado uma transformação do espaço de cores utilizado para um espaço de cores perceptualmente uniforme, e pela distribuição dos valores destas cores pelo histograma. Isto é um fato importante quando se pensa em verificação de similaridade de imagens, pois os espaços de cores uniformes fazem distinção das cores de uma forma mais próxima àquela realizada pelos humanos.

2.1.2 Layout de Cor

Embora a característica de cor global seja simples para calcular e tenha um poder de discriminação razoável em recuperação de imagens, ela tende a fornecer muitas imagens não similares como resposta no caso de coleções de imagens muito grandes.

Muitas pesquisas sugerem o uso de um layout de cores (característica de cor e seus relacionamentos espaciais) como uma solução melhor para recuperação de imagens. Para estender uma característica de cor global para uma local, uma abordagem natural é dividir a imagem em zonas e extrair a cor de cada uma delas. Embora conceitualmente simples, esta abordagem pode trazer ganhos significativos nos processos de recuperação. Uma abordagem mais sofisticada é segmentar a imagem em regiões com características salientes de cor pela retroprojeção de conjuntos de cores e armazenar a posição e a característica do conjunto de cor de cada região para suportar consultas posteriores. A vantagem é o aumento da precisão e a desvantagem é o difícil problema da segmentação.

2.1.3 O uso da cor em recuperação de imagens

A cor é uma característica visual que é imediatamente percebida quando se observa uma imagem [BIM 99]. No entanto, a recuperação de imagens baseada em similaridade de cores exige o uso de modelos de representação de cores. Especialmente modelos que façam a distinção entre cores diferentes da forma mais próxima possível aquela feita pelo sistema visual humano.

A recuperação de imagens baseada em cor deve se encontrar, segundo [BIM 99], em uma das seguintes situações:

- A busca por imagens contendo uma cor específica em uma determinada proporção;
- A busca por imagens cujas cores sejam similares às de uma determinada imagem;
- A busca por imagens contendo regiões de cores conforme as especificadas em uma imagem de consulta;
- A busca por imagens contendo um objeto previamente conhecido baseado nas suas propriedades de cor;
- A busca por imagens com certos efeitos de cores induzidos.

Consultas podem ser expressas tanto através de texto quanto através de exemplos visuais. A especificação textual de cores é uma forma simples de expressar consultas sobre a presença de uma cor. Contudo, é necessária uma correspondência (difícil de se alcançar, mesmo em contextos culturalmente homogêneos) comumente aceita entre os nomes das cores e os estímulos provocados pelas mesmas.

Consultas através de exemplos visuais baseadas em conteúdo, que é o enfoque deste trabalho, caracterizam uma forma mais efetiva de consulta a propriedades perceptuais de imagens coloridas, como distribuição de cor ou regiões de cores. Nesta abordagem, as imagens pesquisadas são aquelas cujas cores têm proximidade com aquelas de uma imagem de consulta, ou cores que aparecem numa proporção ou localização igual ou similar àquelas da imagem de consulta. Exemplos visuais são também úteis para encontrar imagens contendo um objeto previamente conhecido com certas propriedades de cor. Em consultas através de exemplos de imagens, a cor pode ser combinada com outras características como textura e estrutura entre outras. Um peso diferente pode ser associado a cada uma destas características [JAI 96], indicando sua importância relativa dentro do contexto.

2.1.4 Integração de aspectos cromático e estrutural

O aspecto estrutural de uma imagem é uma característica que está fortemente relacionada à identificação de formas e texturas na mesma. Segundo [GOO 2000]: forma, cor e textura são indubitavelmente características visuais extremamente importantes para a representação de imagens. Ou ainda de acordo com [ANT 2000]: o reconhecimento de formas em imagens é considerado muito importante, já que o reconhecimento de objetos pelo homem se faz primeiramente pelas formas. Assim sendo, optou-se neste trabalho por incorporar o estudo da extração do conteúdo estrutural da imagem, além do conteúdo cromático, quando da sua indexação.

A integração de características visuais na representação de imagens [SCH 94] e no processo de recuperação das mesmas pode contribuir significativamente para a obtenção de melhores resultados. Neste ponto reside um outro importante fator que motivou o uso tanto de características cromáticas quanto estruturais, já que, de acordo com [SMI 95] a percepção de imagens pelos humanos é um processo muito complexo. Durante este processo, as células nervosas presentes na retina do olho humano decompõem imediatamente as imagens em componentes separadas, que são contornos, texturas e cores. Estes pedaços fragmentados de informação são então remontados no cérebro formando uma única imagem. Assim, as pessoas não costumam classificar as imagens que vêm a partir de características isoladas, tais como cor, textura e forma. Nos computadores, por outro lado, as imagens têm sido indexadas nas mais diversas aplicações a partir da decomposição das imagens em componentes de texturas, formas, contornos e cores. Desta forma, os computadores são um tanto limitados no que diz respeito à capacidade de reintegrar estas características em uma única imagem. Sem a compreensão da imagem, um computador ou um sistema automatizado não pode atingir um nível de identificação de imagem natural e conveniente para os humanos. Contudo, será visto ao longo deste trabalho que, na prática, estas características de baixo nível podem melhorar significativamente as operações relacionadas à recuperação de imagens a partir de repositórios de imagens.

A representação de características estruturais que será descrita neste trabalho, será baseada no cálculo de operadores de gradientes para os pixels da imagem. Cada pixel terá seu conteúdo estrutural representado por um par de valores, sendo que um destes valores representará a intensidade (magnitude) do contorno identificado no pixel, enquanto que o outro valor representará a direção (ângulo de inclinação) do contorno representado pelo pixel. Para o cálculo destes valores para cada pixel, serão utilizados também os valores de alguns de seus pixels vizinhos.

2.2 Descritores de conteúdo da imagem

Neste trabalho, são utilizados dois tipos de descritores de imagens, um para descrever o conteúdo cromático das mesmas e outro para descrever o conteúdo estrutural. Os descritores de conteúdo de imagem utilizados nas técnicas de recuperação de imagens aqui estudadas constituem-se de conjuntos não vazios de histogramas [FAC 2001] em ambos os casos.

Os descritores estão organizados de tal forma que a imagem original pode ser dividida em zonas a fim de se preservar informações espaciais referentes aos conteúdos presentes na mesma, conforme sugerido em [ASS 2000]. Desta forma, a imagem de consulta utilizada na recuperação pode ser dividida estabelecendo o número de faixas horizontais e verticais segundo as quais as zonas retangulares de imagem deverão ser criadas. Esta divisão em zonas feita na imagem de consulta implicará em uma igual divisão nas imagens que serão comparadas com a mesma.

Com o objetivo de permitir que se possa alterar a precisão das zonas a serem criadas, possibilitando que estas configurações sejam adequadas à imagem que se pretende recuperar, estabeleceu-se a possibilidade de que a divisão da imagem em zonas seja ajustada antes de iniciar o processo de recuperação de uma imagem.

Para isto, definiu-se que o número mínimo de faixas verticais e horizontais utilizados para dividir a imagem em zonas é igual a 1, o que corresponderia à imagem sem divisões, representada inteiramente em uma única zona, enquanto que o máximo é igual a 10. Desta forma, a imagem se divide em $n \times m$ zonas onde n corresponde ao número de faixas horizontais e m ao número de faixas verticais, podendo-se extrair um histograma descritor de conteúdo cromático e um histograma descritor de conteúdo estrutural a partir de cada uma destas zonas. Na (figura 2.1), descrita a seguir, ilustra-se a formação de zonas em uma imagem com $n=3$ e $m=3$.

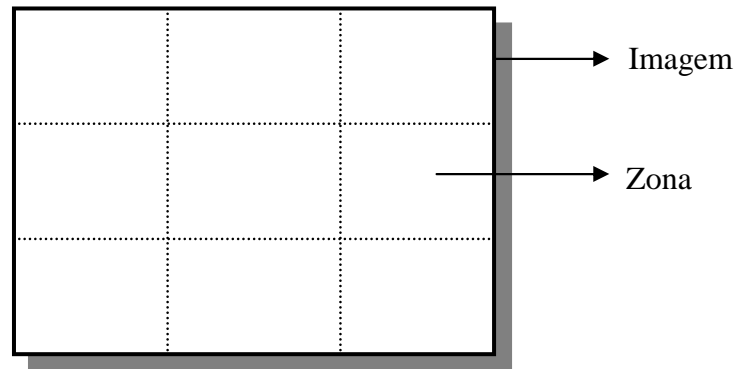


FIGURA 2.1 - Divisão da imagem em zonas com $n=3$ e $m=3$

Portanto, dada uma imagem genérica I , e seu particionamento em $q=n \times m$ (com $1 \leq n \leq 10$ e $1 \leq m \leq 10$) zonas conforme descrito anteriormente, o conteúdo cromático da imagem pode ser expresso através de um conjunto de q histogramas $\{H_i(I)\}_{i=1}^q$, onde o i -ésimo histograma $H_i(I)$ representa o conteúdo cromático da i -ésima zona da imagem I .

2.2.1 Descrição de conteúdo cromático

A partir de cada zona da imagem, conforme descrito na seção anterior, deverá ser extraído um histograma de cores [CHU 98]. Para isto, o espaço de cores através do qual a imagem foi originalmente representada é quantizado para um espaço composto por quatro valores de vermelho, quatro valores de verde, e quatro valores de azul, totalizando 64 possíveis combinações de cores. O número de cores existentes no espaço de cores depois da quantização deve ser um número que viabilize o tratamento

computacional do processo de recuperação de imagens, e que ao mesmo tempo tenha um razoável poder de representação do conteúdo cromático da mesma. Com base nos trabalhos de [ASS 2000] e [SCL 96] pode-se verificar que a divisão do espaço de cores em 64 valores diferentes uniformemente distribuídos constitui-se em uma boa alternativa para este processo de recuperação de imagens. Esta quantização é uma medida bastante importante para que a manipulação dos descritores cromáticos seja mais viável em termos computacionais. Esta importância pode ser observada se levarmos em consideração o fato de que é muito comum o uso e a manipulação de imagens digitais formadas a partir de um espaço de representação de cores com 256 valores de vermelho, 256 valores de verde e 256 valores de azul. Isto implica em um espaço de cores com um total de 16.777.216 possibilidades diferentes para a formação das cores. Somando-se a isto o fato de que cada descritor cromático é representado computacionalmente por meio de um vetor, pode-se constatar a necessidade de se utilizar o processo de quantização do espaço de cores. Para este tipo de imagem (256^3), pode-se realizar a quantização das cores para um espaço com 64 valores diferentes através de:

$$\begin{aligned} R_q &= [R/64] \\ G_q &= [G/64] \\ B_q &= [B/64] \end{aligned} \quad (2.1)$$

onde R_q , G_q e B_q são os valores das componentes R, G e B respectivamente para o espaço quantizado.

Adicionalmente, é importante observar que foi adotada uma quantização uniforme do espaço de cores [SOB 98]. Isto se deve ao fato de que o conjunto padrão de cores para o qual a quantização é dirigida deve ser igual para todas as imagens a serem comparadas. Portanto, a quantização não-uniforme (adaptativa) não seria uma técnica adequada para a construção dos descritores a serem aplicados neste processo de recuperação de imagens.

O histograma é formado por pilhas, uma para cada cor do espaço quantizado, que somam o número de ocorrências de cada cor na zona de imagem correspondente. Na (figura 2.2), descrita a seguir, está ilustrada a formação de um histograma de cores a partir de um pequeno conjunto de pixels. Isto feito, o histograma de cada zona é normalizado, dividindo-se o valor encontrado em cada pilha pelo número total de pixels presentes na zona de imagem. Desta forma, obtêm-se um percentual da ocorrência de cada cor na zona. Assim, o descritor encontrado pode ser confrontado com outros descritores de imagem independentemente da ocorrência de eventuais diferenças entre as dimensões das imagens a serem comparadas.

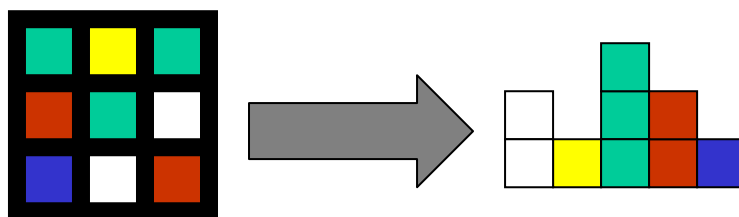


FIGURA 2.2 - Formação do histograma de cores a partir de um conjunto de pixels

2.2.2 Descrição de conteúdo estrutural

O conteúdo estrutural de uma imagem também será descrito através de histogramas, sendo extraído um histograma de cada zona da imagem. A partir de cada pixel da imagem discreta será calculado um par de valores $\langle \rho, \nu \rangle$ que será utilizado para compor o histograma. O par $\langle \rho, \nu \rangle$ será um indicador de características estruturais do seu ponto correspondente na imagem. O valor de ν indica a direção da borda presente no ponto, enquanto que o valor de ρ representa a intensidade da borda neste ponto.

Para isto, são computados gradientes nas direções x e y (δ^x, δ^y) para cada pixel p_{ij} da imagem, conforme descrito a seguir [ASS 2000] [FAC 2000].

$$\delta_{i,j}^x = \frac{p_{i+1,j} - p_{i-1,j}}{\sqrt{2(p_{i+1,j}^2 + p_{i-1,j}^2)}} \quad \delta_{i,j}^y = \frac{p_{i,j+1} - p_{i,j-1}}{\sqrt{2(p_{i,j+1}^2 + p_{i,j-1}^2)}} \quad (2.2)$$

Onde $p_{i,j}$ representa a intensidade de cinza presente no pixel da i -ésima linha e j -ésima coluna.

A orientação de borda e intensidade são respectivamente

$$\nu_{i,j} = \tan^{-1} \left(\frac{\delta_{i,j}^x}{\delta_{i,j}^y} \right) \quad \rho_{i,j} = \sqrt{(\delta_{i,j}^x)^2 + (\delta_{i,j}^y)^2} \quad (2.3)$$

O resultado da fase de detecção de bordas para cada zona da imagem é um conjunto de bordas $\mathcal{E} = \{B_1, \dots, B_{\vartheta}\}$, onde ϑ corresponde ao número de pontos contidos na zona. Cada borda B_i é descrita por um par $\langle \rho, \nu \rangle$ onde ρ é a intensidade da margem e ν sua orientação. O valor da intensidade é discretizado em l ($l=128$) níveis, pois com este número conseguiu-se um poder razoável de representação das intensidades sem comprometer o desempenho do sistema. Depois, um histograma de intensidades $E \equiv \langle e_1, \dots, e_l \rangle$ é computado.

Para descrever um conteúdo estrutural de uma imagem genérica, somente bordas relevantes da imagem devem ser consideradas. Para isto, um limite mínimo τ descarta todas aquelas bordas de imagens cujas intensidades não o ultrapassem. Assim, uma nova imagem é produzida, preservando somente a informação de orientação sobre bordas $B_i \equiv \langle \rho_i, \nu_i \rangle$ para as quais $\lfloor \rho_i \rfloor \geq \tau$. O valor de τ é calculado de forma que a razão entre o número de bordas preservadas e o número total de bordas esteja próximo a 15 % [ASS 2000], pois com esta razão conseguiu-se, na média, bons resultados no que diz respeito à representação do aspecto estrutural relevante da imagem:

$$\tau: \frac{\sum_{k=\tau}^l e_k}{\sum_{k=1}^l e_k} \cong 0.15 \quad (2.4)$$

A imagem resultante é então formada por $q = m \times n$ zonas retangulares R_i ($i=1, \dots, q$). O espaço de orientação é discretizado em t ($t=4$) níveis e um histograma de

orientação $T(R_i)$ é computado para cada zona R_i . Os quatro valores presentes no espaço de orientação discretizado são suficientes para representar inclinações de bordas correspondentes direções dos ângulos 0° e 180° , 45° e 225° , 90° e 270° , e 135° e 315° . Estas direções encontram-se uniformemente distribuídas no círculo trigonométrico.

Assim, cada histograma descritor de conteúdo estrutural $T(R_i)$ será composto por um total de quatro colunas, cada uma delas correspondendo a uma das orientações possíveis para v em cada ponto da imagem. Desta forma, para cada ocorrência de borda com um determinado ângulo de inclinação v e com intensidade suficiente para ser considerada relevante ($\rho \geq \tau$), a pilha correspondente a este ângulo de inclinação v deverá ser incrementada em uma unidade. O conjunto de histogramas $T = \{T(R_1), \dots, T(R_q)\}$ descreve a estrutura da imagem I . A (figura 2.3), descrita a seguir, ilustra a formação de um histograma de estrutura para um pequeno conjunto de pixels. Observe que nesta figura os pixels correspondentes a bordas relevantes estão representados em preto, enquanto que os não relevantes estão representados em cinza.

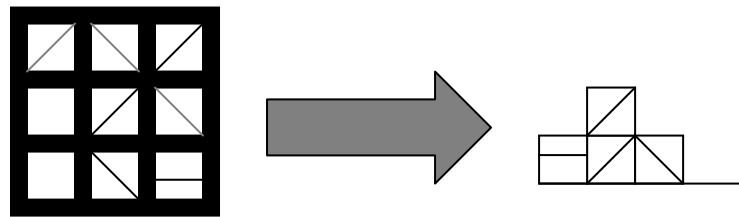


FIGURA 2.3 - Formação do histograma de estrutura a partir de um conjunto de pixels

A elaboração do histograma de estruturas é concluída com a normalização do mesmo. Para isto, divide-se o valor encontrado em cada pilha pelo número total de pixels correspondentes a bordas relevantes presentes no zona de imagem. Desta forma, obtêm-se uma frequência de ocorrência de bordas para cada uma das quatro orientações possíveis. Assim, o descritor encontrado pode ser confrontado com outros descritores de imagem independentemente da ocorrência de eventuais diferenças entre as dimensões das imagens a serem comparadas.

2.3 Conclusões

Neste capítulo foram discutidas algumas questões importantes relacionadas à evolução das técnicas de recuperação de imagens. O aspecto mais importante que se pôde observar está relacionado aos descritores utilizados na indexação de imagens.

Neste sentido, falou-se inicialmente da evolução dos descritores de imagem e pôde-se perceber a importância do surgimento dos descritores baseados em conteúdo visual. Em seguida, foram descritas duas técnicas para indexação de imagens baseadas nos aspectos visuais de cor e estrutura através da criação de histogramas. Ao longo desta discussão, também foi colocada a possibilidade de se preservar, durante a indexação, informações referentes ao relacionamento espacial existente entre os conteúdos da imagem (por meio das zonas retangulares).

3 Recuperação de imagens pela distância euclidiana ponderada

Neste capítulo será descrita a métrica de comparação de descritores de imagens utilizada no processo de recuperação de imagens aqui estudado.

3.1 Medida de similaridade entre imagens

Depois de extraídos os descritores de conteúdo de cada zona presente na imagem, conforme descrito no capítulo anterior, pode-se compará-los aos descritores de zonas extraídos de outra imagem a fim de mensurar a diferença perceptual entre os mesmos através do cálculo da Distância Euclidiana Ponderada, descrita em [BIM 99] e [SCH 2001]:

$$D_{n \times m}(H'_{n \times m}, H''_{n \times m}) = (H'_{n \times m} - H''_{n \times m})^T A (H'_{n \times m} - H''_{n \times m}) \quad (3.1)$$

onde:

- $H'_{n \times m}$ representa o histograma (de cor ou estrutura), da zona posicionada na n-ésima linha e m-ésima coluna de uma das imagens a serem comparadas (imagem 1);
- $H''_{n \times m}$ representa o histograma, também da zona posicionada na n-ésima linha e m-ésima coluna da outra imagem a ser comparada (imagem 2);
- A representa uma matriz de distâncias simétrica, que será apresentada em maiores detalhes nas seções 3.1.6 e 3.1.7;

Cada elemento a_{ij} da matriz de distâncias A representa o quanto a diferença entre a i-ésima e a j-ésima pilha do histograma contribuem para D . Neste caso, como a similaridade entre as imagens pode ser verificada através do conteúdo cromático ou estrutural, constrói-se as matrizes de distância distintas A_{RGB} , A_{LUV} , A_{LAB} e A_E . Estas matrizes são definidas para a computação da distância entre histogramas de cor e histogramas de estrutura (orientação da borda), sendo que a matriz A_{RGB} deve ser utilizada no cálculo da similaridade cromática baseado nas diferenças entre as cores no espaço RGB; A matriz A_{LUV} deve ser utilizada no cálculo da similaridade cromática baseado nas diferenças entre as cores no espaço $L^*u^*v^*$. A matriz A_{LAB} deve ser utilizada no cálculo da similaridade cromática baseado nas diferenças entre as cores no espaço $L^*a^*b^*$. Estas matrizes de distâncias de cores são matrizes simétricas $C \times C$ (onde C é o número de cores de referência, neste caso $C=64$) cujos elementos a_{ij} codificam a similaridade de percepção entre a i-ésima e a j-ésima cor de referência calculada nos seus respectivos espaços de cores. Maiores detalhes sobre a formação destas matrizes podem ser encontrados na seção 3.1.6. A matriz de distâncias de estrutura A_E é uma matriz simétrica 4×4 (considerando 4 valores de orientação de referência) cujos elementos a_{ij} codificam a similaridade de percepção entre a i-ésima e a j-ésima orientação de referência [ASS 2000], maiores detalhes sobre a sua formação serão descritos na seção 3.1.7.

Para a extração de uma medida final das diferenças dos conteúdos cromáticos globais das duas imagens que estão sendo comparadas, é feita uma média aritmética simples das diferenças encontradas entre os histogramas das zonas correspondentes em cada imagem, conforme descrito a seguir e ilustrado na (figura 3.1).

$$D(I_1, I_2) = \frac{\sum_{n=1}^N \sum_{m=1}^M D_{n \times m}(H'_{n \times m}, H''_{n \times m})}{N \times M} \quad (3.2)$$

onde:

- I_1 e I_2 representam as duas imagens que estão sendo comparadas;
- H' e H'' representam os histogramas descritores de conteúdo das imagens I_1 e I_2 respectivamente;
- N e M são, respectivamente, o número de linhas e colunas das zonas componentes das imagens, ajustados para uma determinada comparação entre imagens;

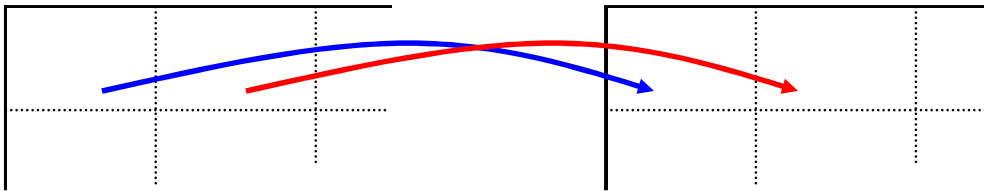


FIGURA 3.1 - Extração de medidas de similaridade entre as zonas correspondentes das imagens que estão sendo comparadas.

3.1.1 O uso de diferentes espaços de cores

Muitos autores, como [BIM 99], [ASS 2000], [PAR 99] e [GEV 9?] têm discutido o uso de outros espaços de cores, que não o RGB, para uma melhor verificação da diferença entre cores. Considerando que a técnica de recuperação de imagens aqui descrita também se baseia no conteúdo cromático da imagem, deve haver uma preocupação também neste sentido.

A maioria das imagens digitais tem suas respectivas cores originalmente descritas através do sistema RGB. Entretanto, outros espaços de cores, como o $L^*u^*v^*$ e o $L^*a^*b^*$ têm se mostrado mais eficientes em alguns casos pelo fato de se caracterizarem como modelos perceptualmente uniformes. Desta forma, optou-se neste trabalho por descrever uma técnica de recuperação de imagens na qual se possa escolher qualquer um entre os três modelos de representação de cores supracitados para a realização dos cálculos de diferenças cromáticas que acontecem durante o processo. Nas próximas seções, estes três espaços de cores serão descritos em maiores detalhes, e além deles, também será brevemente apresentado o espaço de cores XYZ, que é um espaço para o qual as cores no sistema RGB devem ser transformadas antes de serem finalmente convertidas para os modelos $L^*u^*v^*$ e $L^*a^*b^*$. No capítulo 5 deste trabalho serão descritos os resultados experimentais, onde poderá se confirmar ou não a influência significativa do espaço de representação de cores utilizado nesta técnica de recuperação de imagens.

3.1.2 Espaço de cores RGB

O hardware geralmente exibe a cor através do modelo RGB (red, green, and blue), muito utilizado para a representação de cores, mas que não é perceptualmente uniforme; neste modelo, um pixel particular deve estar associado a este vetor tridimensional (r, g, b) que determina as intensidades de cor respectivas; onde (0,0,0) é preto, (k,k,k) é branco, (k,0,0) é vermelho "puro", e assim por diante. Aqui, k representa a quantização máxima para cada primária (normalmente 255). Isto implica num espaço de cores de k^3 cores distintas (2^{24} se $k=255$) que nem todos os dispositivos de exibição, particularmente os mais antigos, podem exibir. Por esta razão, para o propósito de exibição, é comum que seja especificado um subconjunto deste espaço para ser utilizado; este subconjunto é freqüentemente chamado de palheta de cores [SON 99].

Pode-se observar a importância das cores em uma imagem digital, e a informação contida nas mesmas, através da (figura 3.2), que mostra uma mesma imagem representada com diferentes quantidades de cores.

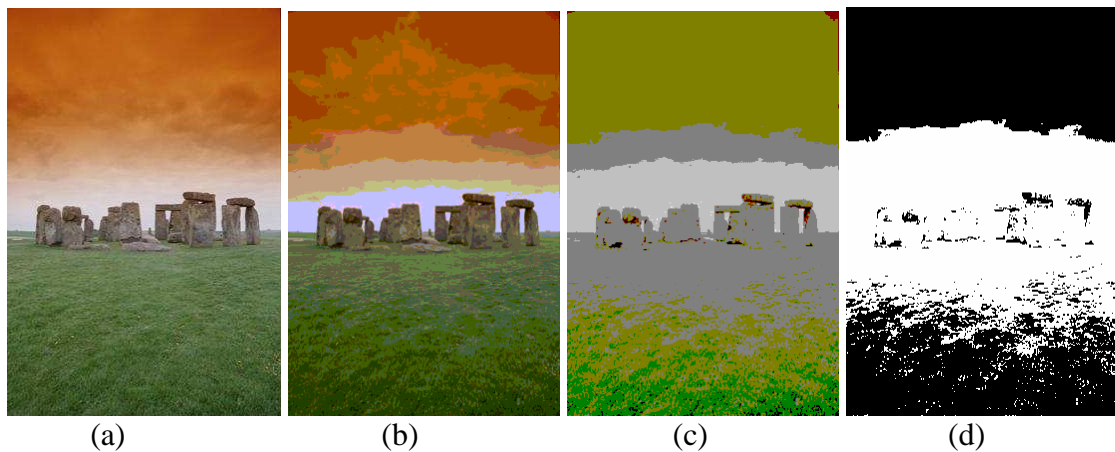


FIGURA 3.2 - Imagem com: (a) 2^{24} bits cores; (b) 256 cores; (c) 16 cores; (d) monocromática (preto e branco)

O modelo RGB pode ser pensado como um espaço de cores tridimensional (figura 3.3); observe as cores secundárias que são combinações de duas primárias puras.

Muitos sensores de captura de imagens fornecem dados de acordo com este modelo; a imagem pode ser capturada por vários sensores, cada um sensível a uma faixa específica de comprimentos de onda, e a função da imagem na saída do sensor é da seguinte forma: cada banda espectral é digitalizada independentemente e é representada por uma função de imagem digital individual como se fosse uma imagem monocromática. Algumas vezes, as imagens são distribuídas utilizando-se uma abordagem similar, mas com diferentes componentes espectrais - por exemplo, existem satélites que transmitem imagens digitalizadas em cinco bandas espectrais desde a próxima a ultravioleta até o infravermelho [GOM 94].

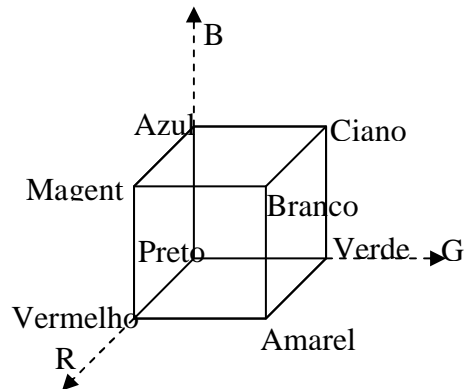


FIGURA 3.3 - Espaço de cores RGB

Principais características do modelo RGB:

- Dependente do dispositivo de exibição;
- Não possui uniformidade perceptual;
- Dependente da direção de visualização, da geometria do objeto, da direção da iluminação, da intensidade e cor da iluminação.

3.1.3 Espaço de cores XYZ

Este sistema de cores é baseado na mistura aditiva de três primárias imaginárias X, Y e Z introduzidas pela CIE (Comissão Internacional de Iluminação). Estas primárias não podem ser vistas por um olho humano ou produzidas, porque são muito saturadas. O fato de estas primárias serem imaginárias não é importante, já que qualquer cor perceptível pode ser descrita matematicamente pela soma destas primárias. Uma outra propriedade importante é que a luminância é determinada somente pelo valor de Y. Isto porque o sistema XYZ é uma combinação linear dos valores R, G e B. O sistema de cores XYZ herda todas as dependências das condições da imagem do sistema de cores RGB [GOM 94].

A transformação do sistema RGB para o sistema XYZ se dá através de:

$$\begin{aligned} X &= 0.490R + 0.310G + 0.200B \\ Y &= 0.177R + 0.813G + 0.010B \\ Z &= 0.000R + 0.010G + 0.990B \end{aligned} \quad (3.3)$$

Principais características do modelo XYZ:

- Dependente do dispositivo de exibição;
- Não possui uniformidade perceptual;
- Dependente da direção de visualização, da geometria do objeto, da direção da iluminação, da intensidade e cor da iluminação.

3.1.4 Espaço de cores $L^*u^*v^*$

Este é um modelo de descrição de cores independente do dispositivo usado na exibição da imagem e diferencia as cores de forma parecida com a realizada pelos humanos [WES 99]. O espaço $L^*u^*v^*$ tem sido recomendado pela CIE (Comissão Internacional de Iluminação) para quantificar diferenças em monitores de vídeo e em condições onde a imagem é formada pela adição de luzes. O sistema $L^*u^*v^*$ faz uma transformação do sistema XYZ, de modo a obter uniformidade perceptual no plano de cromaticidade [BIM 99]. Este modelo foi definido para tornar mais fácil a avaliação das distâncias perceptuais entre cores, e estas distâncias podem ser encontradas por:

$$(\Delta s)^2 = (\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2 \quad (3.4)$$

onde L é a luminosidade (“lightness”) e é baseada em medidas perceptuais de brilho ao invés de medidas físicas, e u^* e v^* são coordenadas cromáticas. Valores do sistema XYZ podem ser transformados em coordenadas $L^*u^*v^*$ através das seguintes equações:

$$\begin{aligned} L^* &= 116(Y/Y_0)^{1/3} - 16 && \text{para } Y/Y_0 > 0.008856 \\ L^* &= 903.3(Y/Y_0) && \text{para } Y/Y_0 \leq 0.008856 \\ u^* &= 13L^*(u' - u'_0) && \text{onde } u' = 4X/(X + 15Y + 3Z) \\ v^* &= 13L^*(v' - v'_0) && \text{onde } v' = 9Y/(X + 15Y + 3Z) \end{aligned} \quad (3.5)$$

onde u'_0 , v'_0 e Y_0 são valores de u' , v' e Y para o branco de referência.

Alguns trabalhos recentes têm mostrado que a escolha por este espaço de cores tem influenciado positivamente na eficiência da recuperação de imagens a cores porque neste espaço, as distâncias Euclidianas são de acordo com as distâncias perceptivas entre as cores [ASS 2000].

3.1.5 Espaço de cores $L^*a^*b^*$

Este modelo é recomendado pelo CIE para quantificar diferenças entre cores com iluminação próxima à luz do dia. O sistema $L^*a^*b^*$ estabelece coordenadas uniformes no espaço tridimensional de cor. Este sistema, juntamente ao $L^*u^*v^*$ apresenta o atrativo da uniformidade perceptual, mas, em contrapartida, apresenta custo computacional significativo para a transformação das cores para sistemas padrão como RGB e XYZ. Este modelo também foi definido para tornar mais fácil a avaliação das distâncias perceptuais entre cores, e estas distâncias podem ser encontradas por:

$$(\Delta s)^2 = (\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2 \quad (3.6)$$

Valores do sistema XYZ podem ser transformados em coordenadas $L^*a^*b^*$ através das seguintes equações:

$$\begin{aligned} L^* &= 116(Y/Y_0)^{1/3} - 16 && \text{para } Y/Y_0 > 0.008856 \\ L^* &= 903.3(Y/Y_0) && \text{para } Y/Y_n \leq 0.008856 \\ a^* &= 500 (f(X/X_0) - f(Y/Y_0)) \\ b^* &= 200 (f(Y/Y_0) - f(Z/Z_0)) \end{aligned} \quad (3.7)$$

onde X_0 , Y_0 e Z_0 são os valores XYZ do branco de referência e

$$\begin{aligned} f(t) &= (t)^{1/3} && \text{para } t > 0.008856 \\ f(t) &= 7.787t + 16/116 && \text{para } t \leq 0.008856 \end{aligned} \quad (3.8)$$

3.1.6 As matrizes de distâncias entre cores A_{RGB} , A_{LUV} , A_{LAB}

A métrica descrita na equação 3.1 para a verificação de similaridade de imagens exige o uso de uma matriz de distâncias A . Para os propósitos deste trabalho, serão estabelecidas quatro diferentes matrizes de distâncias, sendo uma para a comparação de conteúdos cromáticos das imagens levando-se em consideração o espaço de cores RGB, outra baseada no espaço de cores $L^*a^*b^*$, outra baseada no espaço de cores $L^*u^*v^*$, e ainda uma quarta matriz para a comparação de imagens com base nos seus respectivos conteúdos estruturais, que será descrita na próxima seção.

As matrizes A_{RGB} , A_{LUV} e A_{LAB} são matrizes que armazenam as distâncias entre as cores presentes no espaço quantizado sobre o qual os histogramas de cores para as zonas descritoras da imagem foram formadas. Entretanto, é importante observar que este espaço reduzido de cores foi formado a partir da discretização do espaço segundo o qual as cores estavam originalmente descritas nos arquivos de imagem, o espaço RGB com 256^3 possibilidades de cores.

Estas matrizes são matrizes simétricas, de dimensão 64×64 , que armazenam em cada uma de suas células a_{ij} a similaridade perceptual entre a i -ésima e a j -ésima cor. As matrizes A_{LUV} e A_{LAB} são construídas para permitir que se possa utilizar qualquer um dos três espaços de cores descritos na seção 3.1.1 durante a comparação dos conteúdos cromáticos das imagens. Para isto, são utilizadas primeiramente as equações 3.5 e 3.7 para realizar a conversão das cores para os espaços $L^*u^*v^*$ e $L^*a^*b^*$ respectivamente. Em seguida, calcula-se as diferenças d_{ij} entre as cores i e j através de 3.4 e 3.6. Finalmente, as três matrizes de distâncias de cores são construídas colocando-se em cada célula a_{ij} o valor da diferença d_{ij} , dividido pelo valor máximo de distância entre duas cores, subtraído de 1, conforme mostra a próxima equação. Desta forma, a_{ij} varia entre 0 e 1, já que mede similaridade.

$$a_{ij} = 1 - (d_{ij} / \max(d_{ij})) \quad (3.9)$$

A introdução de diferentes espaços de cores se deve ao fato de que o espaço RGB não oferece uma verificação de diferenças perceptuais entre cores de maneira uniforme quando comparada àquela realizada pelos humanos. Isto porque as diferenças entre as cores calculadas neste espaço introduzem distorções em relação à diferença entre cores percebida pelos humanos. Desta forma, julga-se oportuna a introdução dos espaços $L^*u^*v^*$ e $L^*a^*b^*$, que são perceptualmente uniformes. Este é um dos importantes aspectos que serão avaliados no capítulo 5, quando será descrita uma análise de desempenho da técnica de recuperação de imagens aqui descrita através de curvas de “recall x precision”.

3.1.7 A matriz de distâncias A_E

Conforme descrito anteriormente, além das três matrizes discutidas na seção anterior, que apóiam o processo de verificação da similaridade entre imagens baseada nos seus conteúdos cromáticos, também será construída uma matriz para apoiar o processo de verificação da similaridade entre imagens com base nos seus conteúdos estruturais, esta será a matriz A_E .

Os histogramas descritores de conteúdos estruturais da imagem possuem quatro colunas, cada uma representando uma das possíveis inclinações de borda que podem ser encontradas pela equação 2.3. A matriz A_E deve ter em suas células a_{ij} a medida de similaridade perceptual entre as direções i e j . Para isto, a matriz A_E construída será uma matriz simétrica 4x4 cujas células a_{ij} terão os seus valores encontrados através de:

$$a_{ij}=1-(d_{ij}/90) \quad (3.10)$$

onde d_{ij} é a diferença perceptual entre as direções i e j , e pode ser calculado pela seguinte equação:

$$\begin{aligned} d_{ij} &= |i-j| && \text{para } |i-j| \leq 90 \\ d_{ij} &= 180-i+j && \text{para } |i-j| > 90 \text{ e } i > j \\ d_{ij} &= 180-j+i && \text{para } |i-j| > 90 \text{ e } j > i \end{aligned} \quad (3.11)$$

onde i e j são ângulos descritos em graus no intervalo $[0^\circ, 180^\circ]$.

3.1.8 Uma medida de similaridade entre imagens por cor e estrutura

Uma forma de realizar a recuperação de imagens baseada tanto em conteúdo cromático quanto em conteúdo estrutural é através do cálculo da média aritmética ponderada entre as medidas de similaridade cromática e estrutural previamente extraídas de acordo com os cálculos descritos nas seções anteriores. Entretanto, a técnica de recuperação aqui descrita permitirá que antes do início da recuperação de uma imagem os pesos dos aspectos cromático e estrutural sejam ajustados individualmente para cada uma das zonas segundo as quais a imagem foi dividida.

Para isto, será permitido que seja ajustado um valor de peso para o aspecto estrutural p_e e um valor de peso para o aspecto cromático p_c para cada uma das zonas da imagem. Os valores destes pesos devem ser complementares, de forma que a soma entre os dois seja igual a 1. Assim, pode-se calcular diferença entre duas zonas correspondentes das imagens I_1 e I_2 através de:

$$D_i(I_1, I_2) = D_{C_i}(I_1, I_2)p_{c_i} + D_{E_i}(I_1, I_2)p_{e_i} \quad (3.12)$$

onde:

- $D_i(I_1, I_2)$ é a diferença entre a i -ésima zona da imagem I_1 e a i -ésima zona da imagem I_2 , com i variando de 1 a q , sendo q o número máximo de zonas;
- $D_{C_i}(I_1, I_2)$ é a diferença cromática entre a i -ésima zona da imagem I_1 e a i -ésima zona da imagem I_2 , com i variando de 1 a q , sendo q o número máximo de zonas;

- $D_{Ei}(I_1, I_2)$ é a diferença estrutural entre a i -ésima zona da imagem I_1 e a i -ésima zona da imagem I_2 , com i variando de 1 a q , sendo q o número máximo de zonas;
- p_{ci} é o peso ajustado para o aspecto cromático na i -ésima zona da imagem, com i variando de 1 a q , sendo q o número máximo de zonas;
- p_{ei} é o peso ajustado para o aspecto estrutural na i -ésima zona da imagem, com i variando de 1 a q , sendo q o número máximo de zonas;

Depois de calculadas as diferenças por conteúdo (cromático e estrutural) entre cada uma das zonas da imagem, pode-se calcular a medida final de diferença entre as imagens $D(I_1, I_2)$ através do cálculo da média aritmética simples entre as diferenças de cada uma das zonas através de:

$$D(I_1, I_2) = \frac{\sum_{i=1}^q D_i(I_1, I_2)}{q} \quad (3.13)$$

onde:

- $D_i(I_1, I_2)$ é a diferença entre a i -ésima zona da imagem I_1 e a i -ésima zona da imagem I_2 (descrito em 3.12), com i variando de 1 a q , sendo q o número máximo de zonas;

3.2 Esquema conceitual da recuperação de imagens por conteúdo

A recuperação de imagens segundo a técnica aqui descrita parte de uma imagem de consulta. Para isto, o mecanismo de recuperação deverá extrair descritores de características visuais (cor e estrutura) conforme descrito no capítulo 2. Os descritores resultantes da imagem de consulta são utilizados para fazer a busca de imagens mais similares no repositório. Para isto, percorre-se o repositório de descritores de conteúdos extraídos de uma coleção de imagens e compara-os ao descritor da imagem de consulta previamente calculado. A extração de descritores de conteúdo deverá se dar de acordo com os parâmetros previamente ajustados pelo usuário. Por exemplo, se o usuário ajustou peso 1 para o aspecto cromático, e peso 0 para o aspecto estrutural em todas as zonas da imagem, o sistema fará somente a extração do descritor cromático a fim de poupar tempo na realização da recuperação.

A comparação realizada entre os descritores de imagens mencionada no parágrafo anterior, se dará conforme descrito na seção 3.1.8, sendo que para isto será considerado inclusive o número de zonas previamente definido, tanto na extração dos descritores quanto na realização do cálculo de similaridade.

Depois de calculadas as medidas de similaridade entre a imagem de consulta e cada uma das imagens do repositório, a recuperação pode ser caracterizada pela exibição das imagens ordenadas de forma decrescente pelas suas respectivas medidas de similaridade encontradas. Este esquema conceitual da recuperação de imagens encontra-se ilustrado na (figura 3.4), descrita a seguir. A fim de avaliar resultados experimentais acerca das técnicas de recuperação de imagens aqui descritas, foi desenvolvido ao longo

deste trabalho um ambiente computacional para recuperação de imagens baseado nestas técnicas, este sistema será apresentado no próximo capítulo.

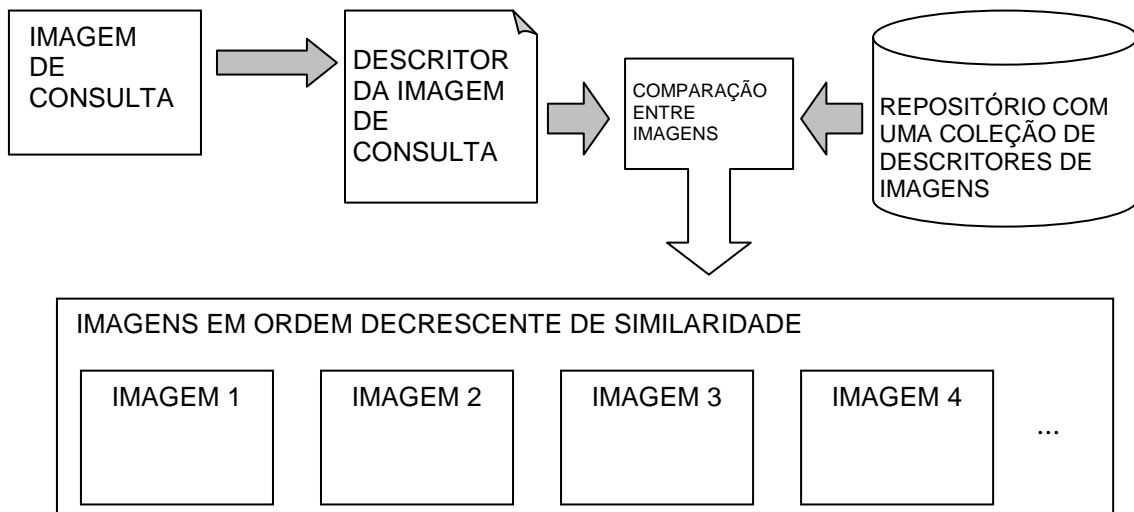


FIGURA 3.4 – Esquema conceitual da recuperação de imagens

3.3 Conclusões

Neste capítulo foi abordado o uso da Distância Euclidiana Ponderada como métrica de similaridade entre imagens a partir dos histogramas descritores das mesmas. Nesta discussão foi observada a possibilidade de que esta métrica seja utilizada dentro do processo de recuperação de imagens. Dentre as conclusões daqui extraídas, é importante ressaltar a possibilidade de se utilizar esta métrica aplicando-se diferentes espaços de cores para o cálculo das diferenças cromáticas; diferentes padrões de divisões das imagens em zonas; e ainda diferentes ponderações para os aspectos visuais cor e estrutura durante o processo de recuperação.

4 Sistema RICE – Sistema para a Recuperação de Imagens por Cor e Estrutura

Os estudos realizados ao longo deste trabalho e o conjunto de técnicas até aqui descritas levaram ao desenvolvimento do sistema RICE, um sistema para recuperação de imagens baseada em cor e estrutura. O sistema foi desenvolvido no ambiente de programação Delphi 6.0 e utilizou o banco de dados Inter Base versão 1.0.1.334 para o armazenamento dos descritores de conteúdo da imagem. A aplicação produzida é compatível com sistema operacional Windows 9x ou superior e é necessário que o Inter Base Server esteja instalado no equipamento onde ela será utilizada. Neste capítulo este sistema será descrito com ênfase em sua interface e em seus aspectos operacionais. O desenvolvimento deste sistema foi de extrema importância para esta pesquisa, pois através dele pôde-se viabilizar a análise de desempenho das técnicas de recuperação de imagens anteriormente descritas através de curvas de “recall x precision”, que serão apresentadas no capítulo 5.

Para apoiar o desenvolvimento do sistema RICE e viabilizar o processo de recuperação de imagens nele implementado, foi desenvolvido um banco de dados para armazenar descritores de imagens. Esta necessidade se deve ao fato de que os cálculos envolvidos no processo de extração dos descritores das imagens consome um tempo de processamento significativo, fato que poderia inviabilizar o uso do sistema. A ciência da existência deste banco de dados é muito importante para a compreensão dos aspectos inerentes ao sistema RICE que serão descritos a seguir.

A (figura 4.1), descrita a seguir, mostra a interface principal do sistema RICE. Nela pode-se observar a presença dos menus “Configuração” e “Processamento”, cujas opções podem ser vistas na (figura 4.2) e na (figura 4.5) respectivamente. As opções oferecidas no menu “Configuração” estão relacionadas ao armazenamento e organização dos descritores de imagens no banco de dados, enquanto que as opções oferecidas no menu “Processamento” permitem que se realize a recuperação de imagens cujos descritores estão armazenados no repositório, ou permite ainda a simples comparação entre duas imagens.

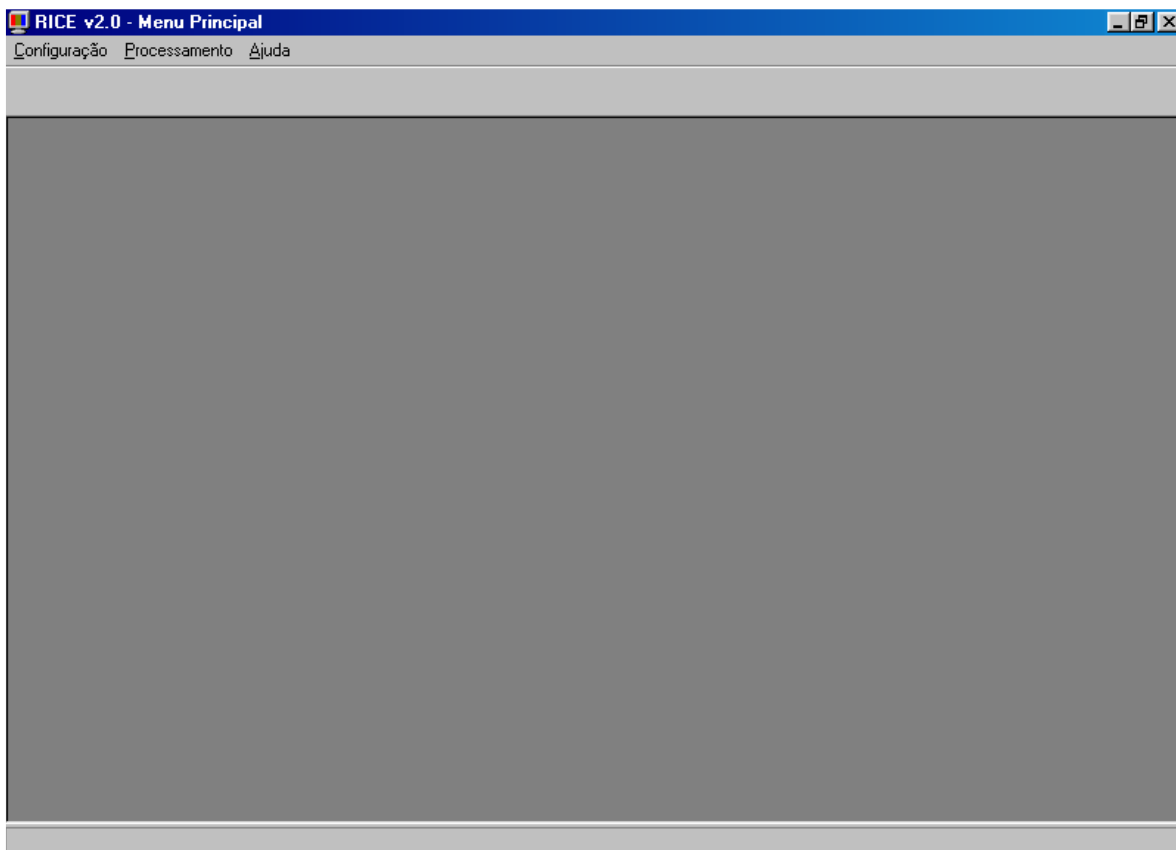


FIGURA 4.1 - Interface principal do sistema RICE

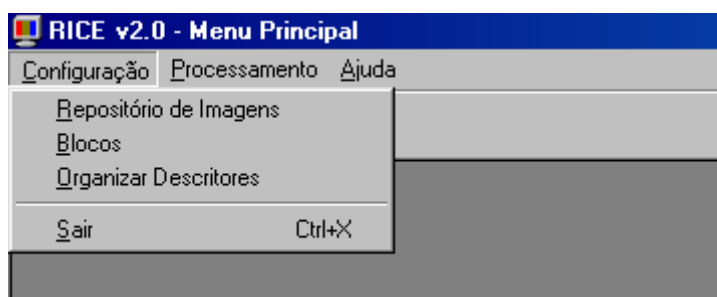


FIGURA 4.2 - Menu "configuração"

A opção "Repositório de imagens", do menu "Configuração", ativa uma janela para a visualização das imagens cujos descritores (tanto cromáticos quanto estruturais) já foram cadastrados no banco de dados. Esta janela pode ser vista na (figura 4.3), mostrada a seguir.

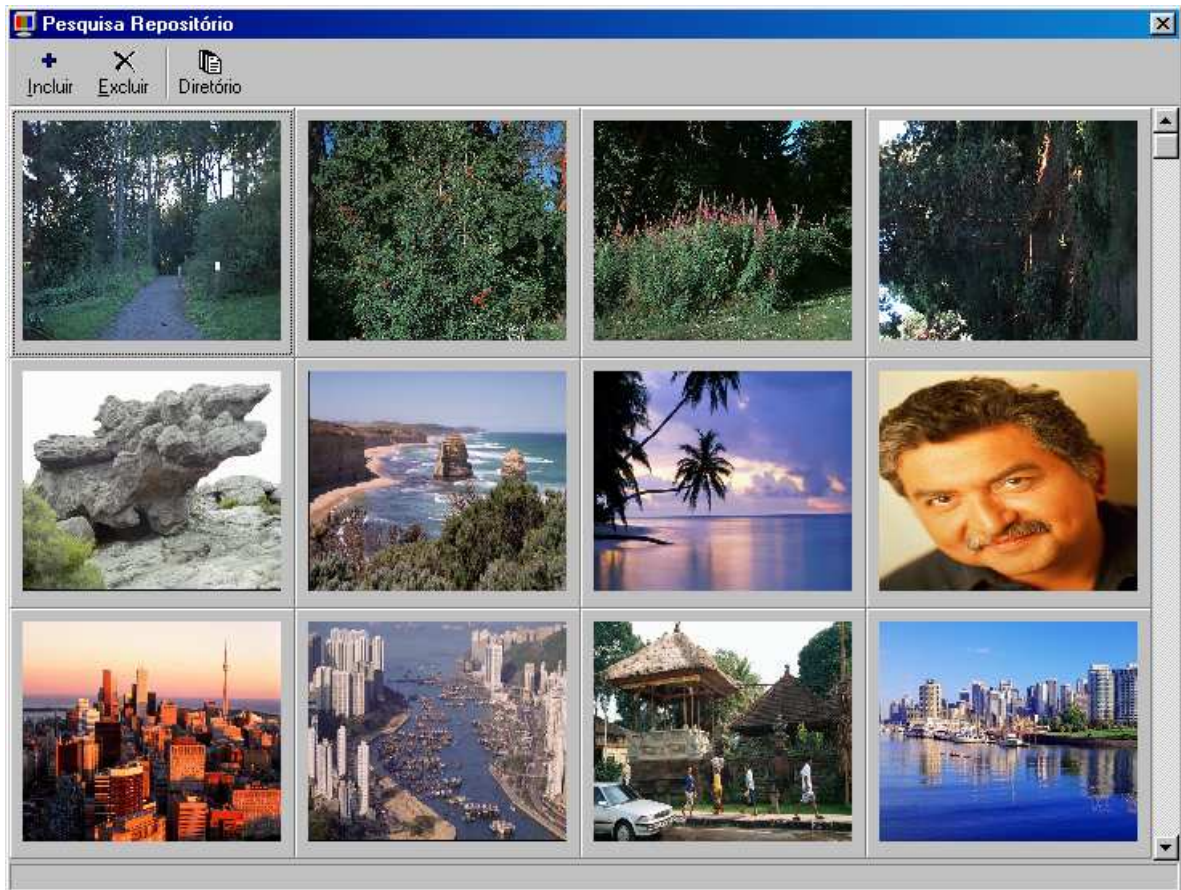


FIGURA 4.3 - Janela que exhibe as imagens cadastradas no banco de dados

Na janela exibida acima, pode-se verificar a presença de botões que permitem inserir novas imagens no repositório ou ainda excluir imagens previamente cadastradas. O cadastro de descritores de imagens pode ser feito individualmente para cada imagem ou pode-se selecionar um diretório e determinar que sejam extraídos e cadastrados descritores para todos os arquivos de imagem nele presentes.

A opção “Blocos” do menu “Configuração” ativa a janela descrita a seguir, na (figura 4.4). Através desta janela é possível verificar quais são os padrões de descritores de imagem, no que diz respeito à divisão das imagens em zonas, que estão cadastrados no sistema. Na figura mostrada a seguir, pode-se identificar que os padrões de divisão da imagem em zonas 1x1, 10x10, 5x5 e 2x2 já estão cadastrados no repositório. Este controle de padrões de descritores se deve ao fato de que diferentes padrões de divisão da imagem em zonas produzem descritores diferentes. No sistema RICE, para cada imagem inserida no repositório são armazenados todos os padrões de descritores previamente cadastrados. Novos padrões de descritores podem ser inseridos no sistema a qualquer momento, a partir da janela exibida na (figura 4.4), assim como padrões previamente cadastrados podem ser excluídos. Quando se cadastra um novo padrão de descritor, são calculados descritores referentes a este novo padrão para todas as imagens já inseridas no banco de dados. Este cálculo pode consumir um tempo de processamento razoável, dependendo da quantidade de imagens já cadastradas no banco de dados. Em contrapartida, estes descritores estarão disponíveis futuramente para a recuperação de imagens, não necessitando mais serem recalculados.

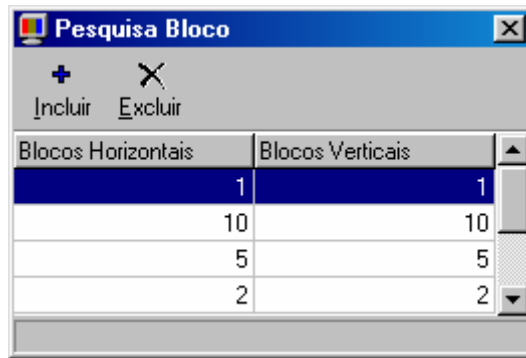


FIGURA 4.4 - Janela para cadastro de padrões de descritores de imagem

A seguir serão descritos os aspectos de interface do sistema RICE relacionados aos processos de comparação entre duas imagens ou recuperação de uma imagem a partir das imagens presentes no repositório. Conforme pode-se observar na (figura 4.5), o menu processamento, presente no sistema RICE, oferece estas duas possibilidades ao usuário.

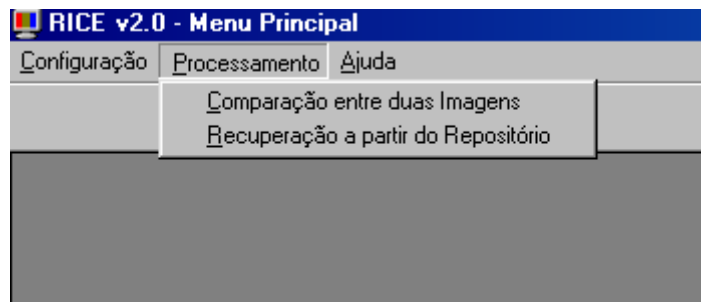


FIGURA 4.5 - Menu processamento

Quando acionada a opção “comparação entre duas imagens”, o sistema exibe a janela descrita na (figura 4.6). Nesta janela, pode-se selecionar as duas imagens a serem comparadas através dos botões “Adquirir imagem 1” e “Adquirir imagem 2”, posicionados na parte superior da mesma. Na parte inferior da janela, são oferecidas opções ao usuário que permitem ajustar parâmetros que podem influenciar significativamente no resultado obtido no processo de comparação entre imagens. Dentre estas opções, pode-se configurar a divisão da imagem segundo a qual serão formadas as zonas e conseqüentemente os descritores das mesmas. Além de configurar a divisão da imagem em zonas, o usuário pode ainda optar por escolher um peso único para os aspectos cromáticos e estrutural que valeriam para todas as zonas, ou estabelecer pesos específicos para os aspectos cromático e estrutural em cada uma das zonas formados na divisão da imagem. O usuário pode também, nos casos em que a comparação tem peso para o conteúdo cromático diferente de zero, optar por um dos seguintes espaços de cores RGB, $L^*u^*v^*$, ou $L^*a^*b^*$. Desta forma, os cálculos de diferenças cromáticas empregados na verificação da similaridade das imagens pode se basear em diferentes espaços de cores.

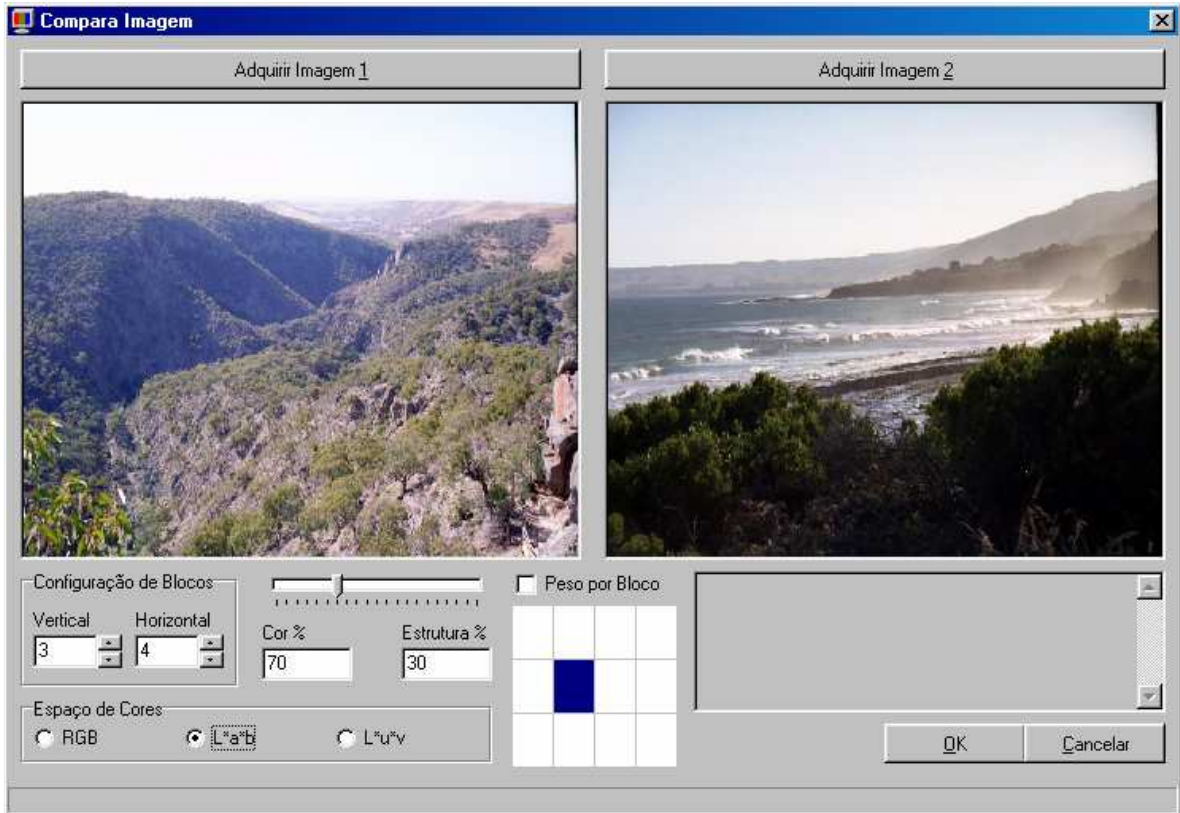


FIGURA 4.6 - Janela que permite a comparação entre duas imagens

A segunda e última opção oferecida no menu processamento é a opção “Recuperação a partir do repositório”, a partir dela ativa-se a janela mostrada na (figura 4.7). Antes de iniciar o processo de recuperação, o usuário deverá optar por um padrão de divisão da imagem em zonas que já esteja previamente cadastrado no sistema. Isto feito, o usuário poderá ajustar as demais configurações (peso do aspecto cromático e estrutural e espaço de cores) da mesma forma como foi descrito no modo de comparação entre duas imagens. Depois de selecionada a imagem de consulta, através do botão “Adquirir imagem 1”, e ajustadas as configurações, o usuário poderá disparar o processo de recuperação da imagem, que será realizado sobre as imagens previamente cadastradas no repositório.

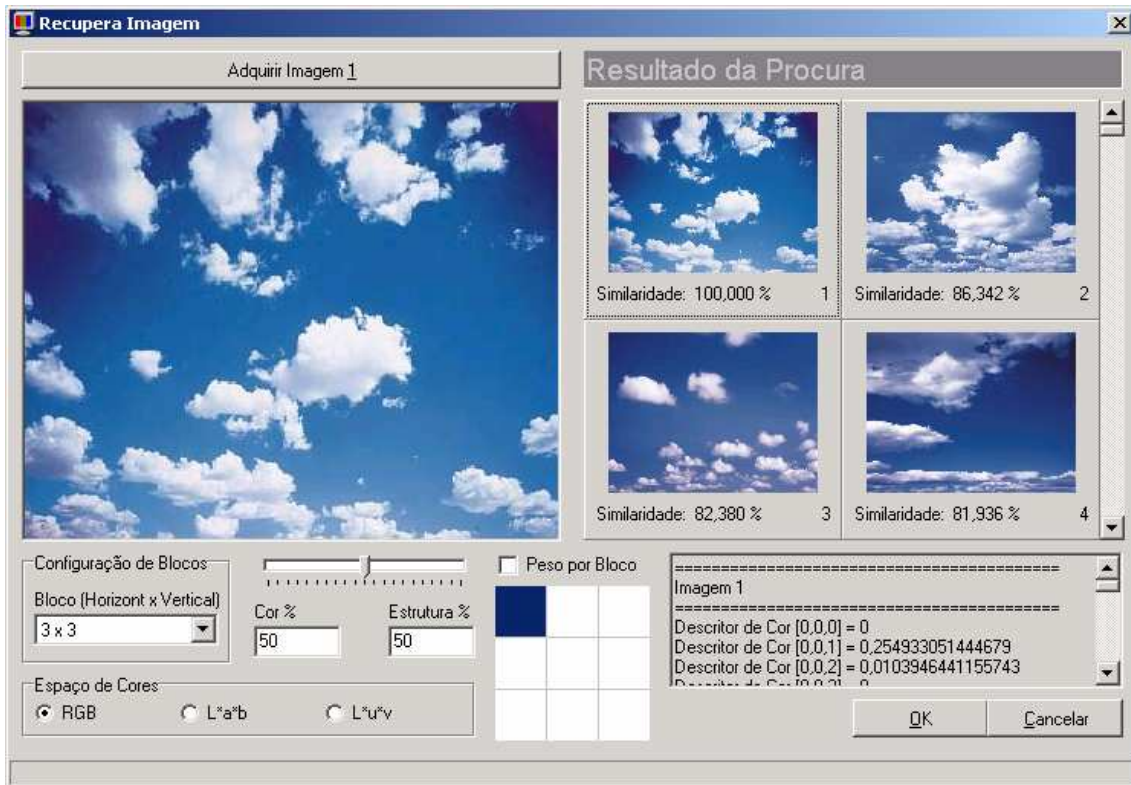


FIGURA 4.7 - Janela que permite a recuperação de uma imagem no repositório

No próximo capítulo, as técnicas de recuperação de imagens até aqui descritas e implementadas no sistema RICE terão seu desempenho avaliado através de curvas de “recall x precision”.

4.1 Conclusões

Neste capítulo foi descrito um ambiente computacional, projetado e construído ao longo do desenvolvimento desta pesquisa, totalmente baseado nas técnicas de recuperação de imagens abordadas neste trabalho. A importância do desenvolvimento deste sistema se deve ao fato de que o mesmo foi utilizado como ferramenta para avaliação do desempenho das técnicas aqui estudadas. Esta avaliação será descrita no próximo capítulo.

5 Resultados experimentais

Conforme descrito no capítulo anterior, o sistema RICE permite que o usuário recupere uma imagem a partir de um repositório, ou ainda permite que o mesmo faça simplesmente a comparação entre duas imagens. Este capítulo encontra-se organizado em duas seções que dizem respeito à descrição de resultados experimentais. Na seção 5.1 será descrito o comportamento do sistema RICE no modo de comparação entre pares de imagens extraídos de uma coleção de padrões cromáticos e estruturais artificiais ilustrada na (figura 5.1). Na seção 5.2 serão descritos os resultados experimentais obtidos com o sistema RICE na recuperação de imagens a partir de uma coleção heterogênea de imagens no que diz respeito aos seus conteúdos semânticos.

5.1 Desempenho com padrões cromáticos e estruturais artificiais

Alguns dos padrões artificiais descritos na (figura 5.1) possuem características cromáticas e estruturais que, quando comparadas entre si, podem ser úteis no sentido de demonstrar o comportamento das técnicas de verificação de similaridade entre imagens descritas neste trabalho e implementadas no sistema RICE. Entretanto, não se tem como objetivo neste momento demonstrar o comportamento do sistema em situações reais de uso, que será descrito na próxima seção. A seguir serão descritos resultados obtidos na comparação entre alguns destes padrões artificiais.

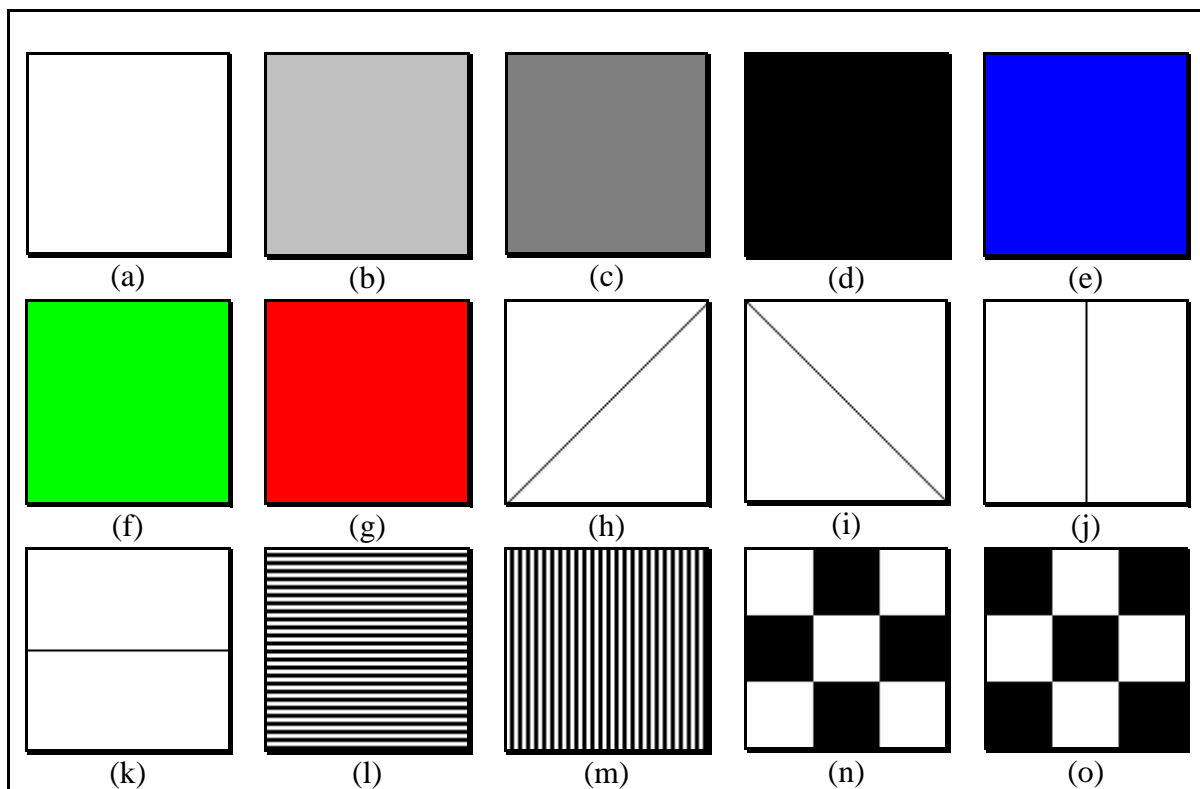


FIGURA 5.1 - Padrões cromáticos e estruturais artificiais

A (tabela 5.1) descrita a seguir é uma esquematização que procura demonstrar resultados obtidos em algumas comparações realizadas entre as imagens mostradas na figura anterior através do sistema RICE. Em cada linha da tabela será descrita uma comparação entre duas imagens sendo que: na segunda e terceira coluna estão identificadas as imagens a serem comparadas; na quarta e quinta coluna estão descritos respectivamente os números de faixas de zonas verticais e horizontais ajustados no sistema; na sexta e sétima coluna estão descritos os pesos (em percentual) estabelecidos para os aspectos cromático e estrutural respectivamente; e, finalmente, na última coluna é apresentado o resultado obtido na comparação realizada. É importante observar que neste experimento foi utilizado o espaço de cores $L^*u^*v^*$ para o cálculo das diferenças entre as cores.

TABELA 5.1 - Resultados obtidos em comparações entre imagens artificiais

	Imagem 1	Imagem 2	Nº de zonas verticais	Nº de zonas horizontais	Peso Cor (%)	Peso Estrutura	Similaridade (%)
1	5.1(a)	5.1(d)	1	1	100	0	0
2	5.1(a)	5.1(d)	1	1	0	100	100
3	5.1(a)	5.1(d)	1	1	50	50	50
4	5.1(h)	5.1(i)	1	1	100	0	99,98
5	5.1(h)	5.1(i)	1	1	0	100	0
6	5.1(j)	5.1(k)	1	1	0	100	0
7	5.1(j)	5.1(h)	1	1	0	100	29,29
8	5.1(l)	5.1(m)	1	1	100	0	100
9	5.1(l)	5.1(m)	1	1	0	100	0
10	5.1(k)	5.1(m)	1	1	0	100	0
11	5.1(n)	5.1(o)	1	1	0	100	100
12	5.1(n)	5.1(o)	1	1	100	0	88,88
13	5.1(n)	5.1(o)	3	3	100	0	0
14	5.1(n)	5.1(o)	2	2	100	0	88,88
15	5.1(n)	5.1(o)	3	3	0	100	100
16	5.1(n)	5.1(o)	2	2	0	100	100
17	5.1(f)	5.1(e)	1	1	100	0	6,09
18	5.1(f)	5.1(g)	1	1	100	0	7,81
19	5.1(e)	5.1(g)	1	1	100	0	13,67
20	5.1(g)	5.1(a)	1	1	100	0	9,63
21	5.1(g)	5.1(d)	1	1	100	0	27,40
22	5.1(d)	5.1(c)	1	1	100	0	29,42
23	5.1(d)	5.1(b)	1	1	100	0	13,45
24	5.1(a)	5.1(b)	1	1	100	0	49,90

Na tabela descrita acima, existe um índice posicionado na primeira coluna a esquerda para identificar cada uma das comparações entre imagens realizada. A seguir, serão feitas algumas observações relacionadas a cada uma destas comparações, sendo que cada uma delas será identificada pelo seu índice.

1. Nesta comparação uma das imagens é constituída apenas por pixels brancos, enquanto que a outra imagem é constituída apenas por pixels pretos. Neste caso,

como a similaridade considerou apenas o aspecto cromático a similaridade entre ambas foi igual a 0%.

2. Aqui foram utilizadas as mesmas imagens da comparação anterior. Entretanto, só se considerou neste caso o aspecto estrutural. Como não se formam bordas em nenhuma das imagens, pois cada uma destas imagens é formada por pixels de mesma cor, estas imagens são iguais em relação ao aspecto estrutural. Desta forma, a similaridade obtida foi igual a 100%.
3. Novamente a comparação se dá entre as imagens 5.1(a) e 5.1(d). Neste caso, foram estabelecidos pesos iguais para os aspectos cromático e estrutural (50% para cada). Assim, o resultado obtido na medida de similaridade foi igual a 50%, obtido a partir do cálculo da média entre as similaridades encontradas nas comparações 1 e 2.
4. Esta comparação é realizada entre as imagens 5.1(h) e 5.1(i). Observe que do ponto de vista cromático estas imagens estão bastante próximas, de forma que o valor de similaridade encontrado é igual a 99,98%. Isto se deve ao fato de que estas imagens são formadas apenas por pixels pretos e brancos e em quantidades aproximadamente iguais em cada uma das imagens.
5. Novamente a comparação se dá entre as imagens 5.1(h) e 5.1(i). Entretanto, nesta comparação é considerado apenas o aspecto estrutural das imagens. Assim, pode-se observar que os pixels que formam borda na primeira imagem possuem todos um mesmo ângulo de direção. Em outras palavras, pode-se dizer que 100% dos pixels formadores de borda presentes na imagem 1 possuem o mesmo ângulo de direção. Adicionalmente, pode-se notar que os pixels formadores de borda na imagem 2 também possuem todos um mesmo ângulo de direção, mas neste caso um ângulo de direção que forma uma borda perpendicular à borda formada pelos pixels da imagem 1. Neste caso, pode-se dizer que todos os pixels formadores de bordas da imagem 1 formam uma borda perpendicular à borda formada por todos os pixels da imagem 2, caracterizando uma similaridade estrutural de 0% entre as imagens.
6. Esta comparação se dá entre as imagens as imagens 5.1(j) e 5.1(k). Neste caso, novamente todos os pixels formadores de borda de uma imagem formam uma borda perpendicular à borda formada por todos os pixels da outra imagem (como no caso anterior). Assim, como apenas foi considerado o aspecto estrutural na comparação, a similaridade verificada mais uma vez foi igual a 0%.
7. As imagens 5.1(j) e 5.1(h) são comparadas em relação ao aspecto estrutural. Neste caso, todos os pixels formadores de borda em uma imagem possuem um ângulo de inclinação com uma diferença (inferior à diferença máxima de 90°) em relação ao ângulo dos pixels formadores de borda da outra imagem. Desta forma, observa-se uma similaridade igual a 29,29%, maior do que as verificadas nas comparações 5 e 6.
8. Comparando as imagens 5.1(l) e 5.1(m), considerando apenas o aspecto cromático, observa-se uma similaridade de 100% entre as imagens. Isto se deve ao fato de que cada uma destas imagens possui 50% dos pixels pretos e 50% dos

pixels brancos. Assim, em termos de aspectos cromáticos elas podem ser consideradas iguais.

9. Nesta situação, as imagens 5.1(l) e 5.1(m) são novamente comparadas, mas neste caso o peso máximo (100%) é dado ao aspecto estrutural. Desta forma, acontece o mesmo relatado na comparação 6, ou seja, todos os pixels indicadores de borda de uma imagem possuem um ângulo de inclinação perpendicular ao ângulo de todos os pixels formadores de borda da outra imagem. Desta forma, a similaridade verificada é igual a 0%.
10. A comparação entre as imagens 5.1(k) e 5.1(m) mostra que embora a imagem 5.1(m) tenha muito mais pixels formadores de borda, o descritor estrutural final extraído desta imagem é normalizado, possuindo a informação de que 100% dos pixels formadores de borda presentes na imagem possuem um ângulo de inclinação perpendicular ao ângulo dos pixels formadores de borda da imagem 5.1(k). Assim, a similaridade encontrada é de 0%, exatamente como na comparação 6.
11. A comparação entre as imagens 5.1(n) e 5.1(o) aqui realizada considera apenas o aspecto estrutural. Desta forma, as bordas extraídas de cada uma destas imagens são exatamente iguais, e conseqüentemente seus descritores também. Assim, a similaridade encontrada foi igual a 100%.
12. Esta comparação considera novamente as imagens 5.1(n) e 5.1(o). Entretanto, nesta situação o peso máximo (100%) foi atribuído ao aspecto cromático. Desta forma, a similaridade verificada entre as imagens foi de 88,88%, já que a soma dos pixels pretos na imagem 5.1(o) é um pouco superior a soma dos pixels pretos encontrados na imagem 5.1(n), caracterizando uma pequena diferença de similaridade entre as mesmas.
13. Nesta comparação, utiliza-se as imagens 5.1(n) e 5.1(o) com peso máximo (100%) atribuído ao aspecto cromático novamente. Entretanto, esta é a primeira comparação na qual os valores dos números de zonas verticais e horizontais (presentes na terceira e quarta coluna respectivamente) foram ajustados para valores diferentes de 1. Com estes números ajustados para o valor igual a 3, pode-se verificar que esta divisão sobre cada uma das imagens provoca uma comparação entre zonas retangulares de imagem que não possuem nenhuma similaridade cromática. Observe por exemplo que a zona superior esquerda da imagem 5.1(n) apresenta somente pixels brancos, enquanto que a zona situada na mesma posição na imagem 5.1(o) possui apenas pixels pretos. Esta diferença cromática se repete para todas as demais zonas formadas na divisão das imagens. Desta forma, a média entre as diferenças cromáticas entre cada zona da primeira imagem e sua zona respectiva presente na outra imagem determina a medida de similaridade final entre as imagens, que é igual a 0%. Esta situação mostra como o uso desta divisão da imagem em zonas pode preservar informações espaciais influenciando de forma bastante significativa no valor de similaridade encontrado.
14. A diferença entre esta comparação e a última descrita está na divisão da imagem em zonas. Neste caso as imagens foram divididas em duas faixas verticais e duas

faixas horizontais para a formação de zonas. Considerando-se novamente apenas o aspecto cromático a medida de similaridade obtida foi igual a 88,88%. Isto mostra novamente, como a divisão em zonas, e o número de zonas que se determina para a divisão da imagem podem influenciar muito na medida de similaridade obtida.

15. Nesta comparação temos as mesmas imagens e um ajuste dos números de zonas verticais e horizontais como na comparação 13. Contudo, foi atribuído peso máximo ao aspecto estrutural, e neste caso, pode-se verificar que esta divisão em zonas não influenciou no valor de similaridade obtido na comparação 11, permanecendo a medida de similaridade em 100%.
16. Esta é a última comparação aqui realizada entre as imagens 5.1(n) e 5.1(o). A única diferença entre esta comparação e a anterior é que nesta o número de zonas verticais e horizontais foi ajustado para 2. Observe no entanto, que neste caso esta diferença entre o número de zonas ajustado não provocou mudança no valor final de similaridade encontrado na comparação anterior, que foi novamente igual a 100%.
17. Esta comparação, realizada entre as imagens 5.1(f) e 5.1(e) leva em consideração apenas o aspecto cromático. A similaridade aqui verificada se dá entre uma imagem com verde “puro” e azul “puro”, e é igual a 6,09%. Este valor será oportunamente comparado com os valores de similaridade obtidos entre verde “puro” e vermelho “puro”, e azul “puro” e vermelho “puro”, ainda a serem comentados.
18. A similaridade verificada nesta comparação, entre as imagens 5.1(f) e 5.1(g), é igual a 7,81%. Considerando que nesta comparação apenas foi considerado o aspecto cromático, e que ela se dá entre verde “puro” e vermelho “puro”, é importante observar que o valor de similaridade obtido é diferente do valor obtido na comparação 17. Isto se deve ao fato de que os cálculos de diferenças entre cores utilizados pelo sistema RICE nesta comparação entre imagens, utilizou o espaço de cores $L^*u^*v^*$, que calcula a diferença entre cores de forma parecida com aquela realizada pelos humanos, fato que não aconteceria se fosse utilizado apenas o sistema RGB, que produziria valores de similaridade iguais para as comparações 17 e 18.
19. Esta comparação se dá entre as imagens 5.1(e) e 5.1(g), ou seja, entre azul “puro” e vermelho “puro”. Aqui é atribuído peso máximo (100%) ao aspecto cromático e a similaridade entre as imagens encontrada é igual a 13,67%. Desta forma, pode-se observar mais uma vez a influência do uso de um espaço de cores uniforme perceptualmente durante a realização dos cálculos de diferença entre cores.
20. Na comparação entre as imagens 5.1(g) e 5.1(a), ou seja, vermelho “puro” e branco, considerando-se apenas o aspecto cromático a similaridade verificada foi de 9,63%.

21. Na comparação entre as imagens 5.1(g) e 5.1(d), ou seja, vermelho “puro” e preto, considerando-se apenas o aspecto cromático a similaridade verificada foi de 27,40%.
22. A comparação entre as imagens 5.1(d) e 5.1(c), levando-se em consideração apenas o aspecto cromático resultou em uma similaridade igual a 29,42%.
23. A comparação entre as imagens 5.1(d) e 5.1(b), levando-se em consideração apenas o aspecto cromático, resultou em uma similaridade igual a 13,45%. Observe que nesta situação a imagem 5.1(b), comparada com o preto, caracteriza um tom de cinza mais distante do preto se comparado com o tom de cinza da imagem 5.1(c), utilizada na comparação anterior. Por isto, a medida de similaridade aqui encontrada é menor.
24. Nesta comparação, entre as imagens 5.1(a) e 5.1(b), com peso máximo atribuído ao aspecto cromático, a similaridade verificada entre as imagens é igual a 49,90%.

5.2 Desempenho com imagens reais

Nesta seção serão descritos os resultados experimentais obtidos através do sistema RICE na recuperação de imagens. Para isto, foi utilizado um conjunto de 287 imagens com conteúdos semânticos bastante variados. As bibliotecas utilizadas encontram-se disponíveis em [COM 2002], e possuem imagens que mostram desde paisagens naturais, como montanhas, bosques, nuvens, vales, paisagens litorâneas, até fotos de pessoas e paisagens urbanas, que ilustram palácios, prédios, igrejas e carros entre outros.

A partir desta base de dados, com imagens de diferentes categorias, verificou-se o desempenho do sistema de recuperação de imagens desenvolvido, no que diz respeito à efetividade dos processos de recuperação de imagens oferecidos pelo mesmo. Para cumprir este propósito, foram utilizadas curvas de “recall x precision”, um método para verificação de desempenho de sistemas de recuperação utilizado em vários trabalhos tais como [RAO 99] e [SMI 9?] e que se encontra descrito em [BEA 99], [SAL 83] e [SMI 2002].

Considere uma consulta Q dirigida à uma sistema de recuperação de imagens que fará a busca em um conjunto de imagens, o repositório I . Considere ainda que, para a consulta Q existe um conjunto de imagens R , subconjunto de I , que é formado pelas imagens relevantes para esta consulta. Considere $|R|$ como sendo o número de imagens pertencentes ao conjunto R .

Um sistema de recuperação de imagens deve possuir uma estratégia segundo a qual as imagens retornadas devem ser apresentadas. Geralmente, apresenta-se as imagens do repositório em ordem decrescente de similaridade em relação à imagem de consulta. Isto também acontece no sistema RICE, entretanto, com o propósito de avaliar o desempenho das opções de recuperação oferecidas pelo sistema, estabeleceu-se um limite L abaixo do qual as imagens não são consideradas imagens retornadas pela consulta. O valor escolhido para L foi 70%, assim, somente as imagens com

similaridade superior ou igual a este valor são consideradas as imagens retornadas em uma consulta. A partir de agora, este conjunto de imagens retornadas por uma consulta será chamado de conjunto A . Adicionalmente, $|A|$ será igual ao número de imagens pertencentes ao conjunto A , e $|Ra|$ será o número de imagens que pertencem tanto ao conjunto R quanto ao conjunto A .

Recall é a porcentagem de documentos relevantes (o conjunto R) que foram recuperados pela consulta, e pode ser obtido por:

$$\text{Recall} = \frac{|Ra|}{|R|} \quad (5.1)$$

Precision é a porcentagem dos documentos recuperados (o conjunto A) que são relevantes, e pode ser obtido por:

$$\text{Precision} = \frac{|Ra|}{|A|} \quad (5.2)$$

Para demonstrar a construção de um gráfico de “recall x precision”, será utilizado um exemplo baseado em [BEA 99]. Considere inicialmente que para a consulta Q , há o seguinte conjunto R_q de imagens relevantes:

$$R_q = \{d_3, d_5, d_9, d_{25}, d_{39}, d_{44}, d_{56}, d_{71}, d_{89}, d_{123}\}$$

O conjunto de imagens relevantes deve ser previamente estabelecido, e isto deve ser feito por um especialista.

Considere agora um algoritmo de recuperação de imagens que retorne para a consulta Q as seguintes imagens na seguinte ordem decrescente de similaridade:

d_{123} *
 d_{84}
 d_{56} *
 d_6
 d_8
 d_9 *
 d_{511}
 d_{129}
 d_{187}
 d_{25} *
 d_{38}
 d_{48}
 d_{250}
 d_{113}
 d_3 *

As imagens marcadas com asteriscos são as imagens pertencentes ao conjunto R_q , isto é, são imagens consideradas relevantes para a consulta Q . Ao se examinar o conjunto de imagens retornadas, a partir do primeiro documento, pode-se fazer as

seguintes observações. Primeiramente, a imagem d_{123} , que ficou posicionada em primeiro lugar na recuperação, é relevante. Esta imagem corresponde a 10% do total de documentos relevantes (o conjunto R_q). Desta forma, pode-se dizer que se tem 100% de “precision” em 10% de “recall”. Continuando a análise, verifica-se que o documento d_{56} é o próximo documento relevante na seqüência e ficou posicionado em terceiro lugar. Neste ponto, pode-se dizer que se tem aproximadamente 66% de “precision” em 20% de “recall”. Continuando esta análise, pode-se traçar a curva mostrada na (figura 5.2). Observe que a medida de “precision” para valores de “recall” superiores a 50% declina para 0%. Isto acontece porque nem todos os documentos tidos como relevantes foram retornados pela consulta.

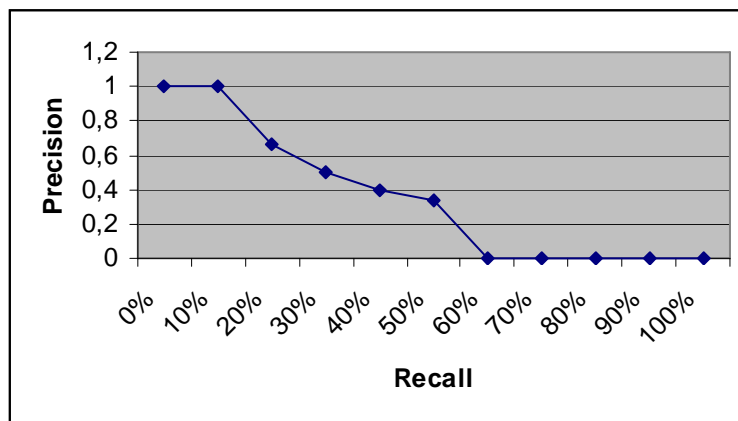


FIGURA 5.2 - Exemplo de curva de “recall x precision”

O exemplo ilustrado anteriormente mostra a construção da curva de “recall x precision” para uma única consulta. Entretanto, para a avaliação de um algoritmo de recuperação de imagens recomenda-se a realização de várias consultas. Neste ponto, deve-se observar a possibilidade de duas situações distintas. A primeira é a realização de várias consultas de forma que cada uma delas tenha o mesmo número de elementos no conjunto de imagens relevantes. Neste caso, constrói-se uma curva “recall x precision” calculando-se as médias das medidas de “precision” para cada medida de “recall”.

O segundo caso acontece quando as consultas possuem diferentes números de imagens relevantes. Neste caso, adota-se um padrão de níveis de “recall” segundo os quais serão tomadas as medidas de “precision”. Geralmente, adota-se um padrão com 11 níveis de “recall”, que são os seguintes: 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% e 100%. Isto feito, as medidas de “precision” são tomadas para cada consulta de acordo com suas respectivas quantidades de imagens relevantes e depois estas medidas são interpoladas de forma que sejam descritas no padrão adotado para todas as consultas (11 níveis de “recall”).

Para exemplificar a interpolação descrita anteriormente, considere o padrão de níveis de “recall” descrito no parágrafo anterior, que é o mesmo utilizado na consulta exemplificada e descrita na (figura 5.2). Considere ainda uma consulta cujas imagens retornadas sejam as mesmas que foram retornadas pela consulta Q descrita anteriormente, mas com o seguinte conjunto de imagens relevantes:

$$R_q = \{d_3, d_{56}, d_{129}\}$$

Neste caso, a primeira imagem relevante no retorno da consulta Q é d_{56} , que proporciona um nível de “recall” de 33,33%, com “precision” também igual a 33,33%. A segunda imagem relevante é a d_{129} , que proporciona um nível de “recall” de 66,66%, com “precision” igual a 25%. O terceiro documento relevante, é o d_3 , que proporciona um nível de “recall” de 100%, com “precision” igual a 20%. A interpolação que coloca os três níveis de “recall” descritos anteriormente no padrão com 11 níveis de “recall” se dá da seguinte forma. Nos níveis de “recall” 0%, 10%, 20% e 30% (que são menores do que 33,33%) as medidas de “precision” são iguais a 33,33%, já que este é o valor de “precision” encontrado para o “recall” de 33,33%. Para os níveis de “recall” de 40%, 50% e 60% (que são menores do que 66,66%) a medida de “precision” interpolada é 25%, que é a medida de “precision” para o nível de “recall” 66,66%. Finalmente, nos níveis de “recall” 70%, 80%, 90% e 100% a medida de “precision” interpolada é 20%, que é a medida de “precision” em nível de “recall” 100%.

5.2.1 Avaliação de desempenho do sistema RICE

Para realizar os testes de desempenho do sistema RICE, selecionou-se 10 imagens dentre as 287 presentes no repositório utilizado. Para cada uma destas imagens selecionadas foi escolhido um conjunto de imagens que seriam consideradas relevantes quando as mesmas fossem consultadas. Desta forma, realizou-se os testes de desempenho, que serão descritos a seguir, fazendo-se 10 consultas para cada configuração dos parâmetros ajustáveis do sistema que foi testada. As próximas figuras mostram alguns conjuntos de imagens presentes no repositório utilizado agrupadas por similaridade. Em cada uma destas figuras a primeira imagem corresponde à imagem que foi utilizada como imagem de consulta durante a realização dos testes, enquanto que as demais correspondem às imagens consideradas relevantes no processo de recuperação da primeira pelo fato de apresentarem similaridade no que diz respeito ao conteúdo semântico. É importante lembrar ainda que nos testes realizados, o conjunto de imagens recuperadas foi formado por aquelas que apresentaram similaridade igual ou superior a 70% em relação à imagem de consulta.

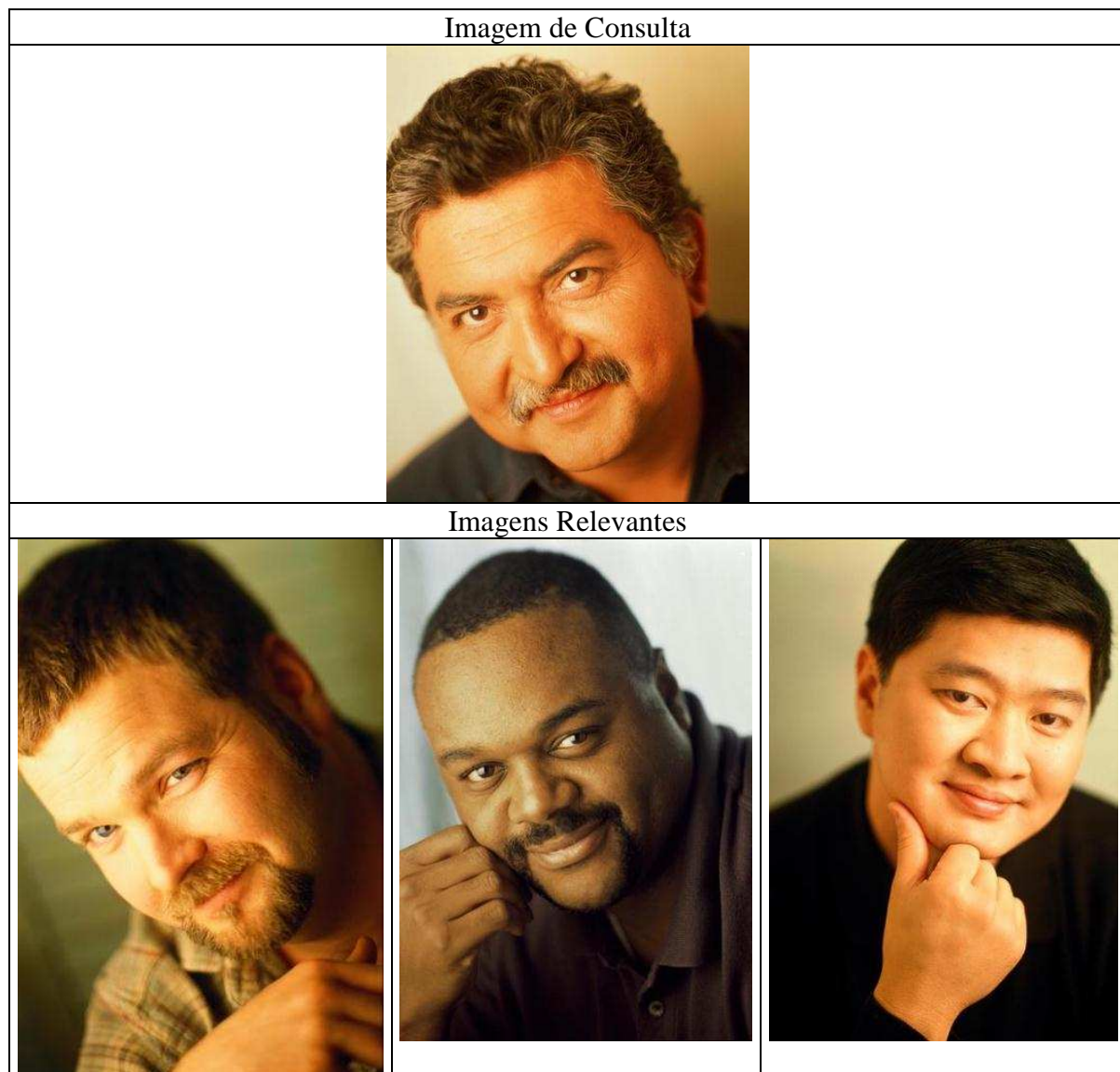


FIGURA 5.3 - Foto de uma face humana e o conjunto de imagens relevantes durante a sua recuperação

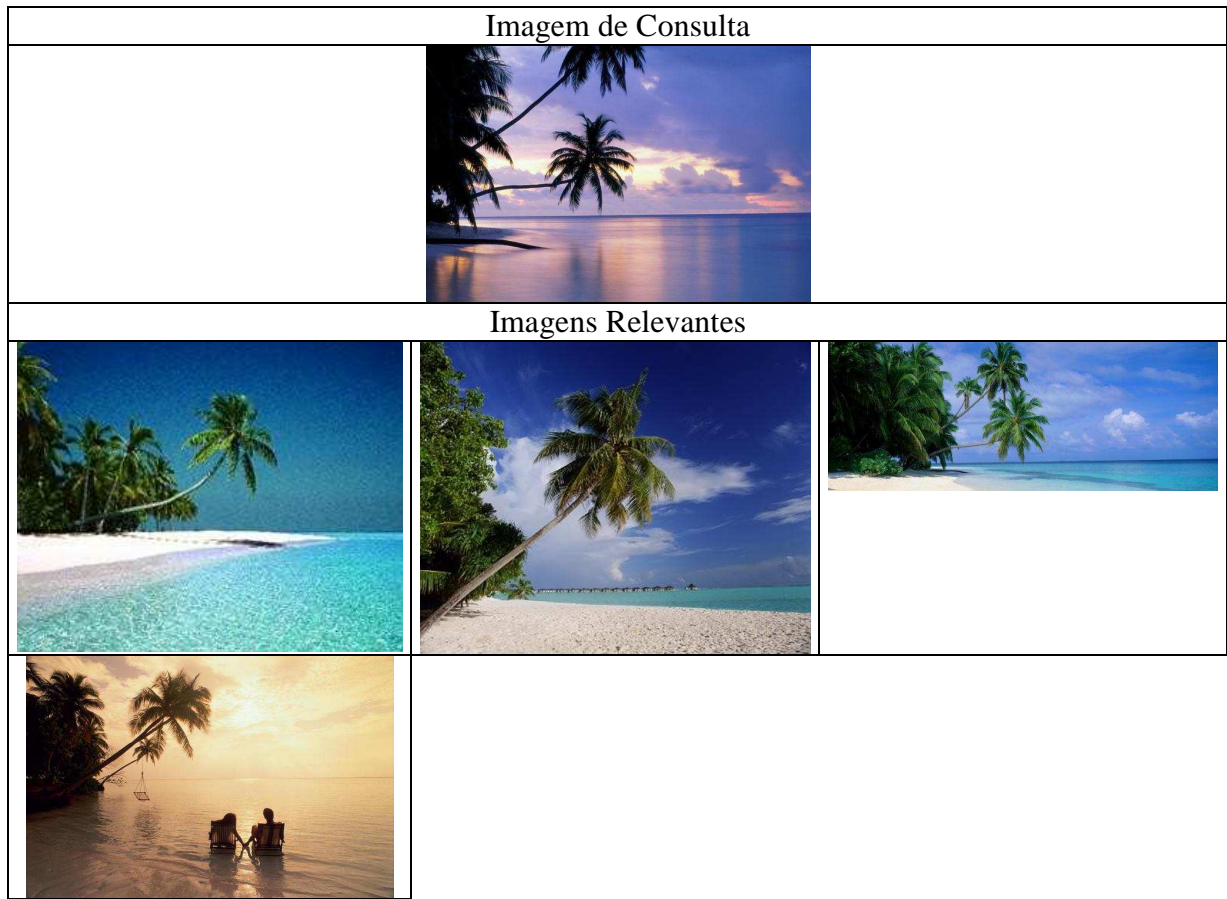


FIGURA 5.4 - Foto de uma paisagem litorânea e o conjunto de imagens relevantes durante a sua recuperação

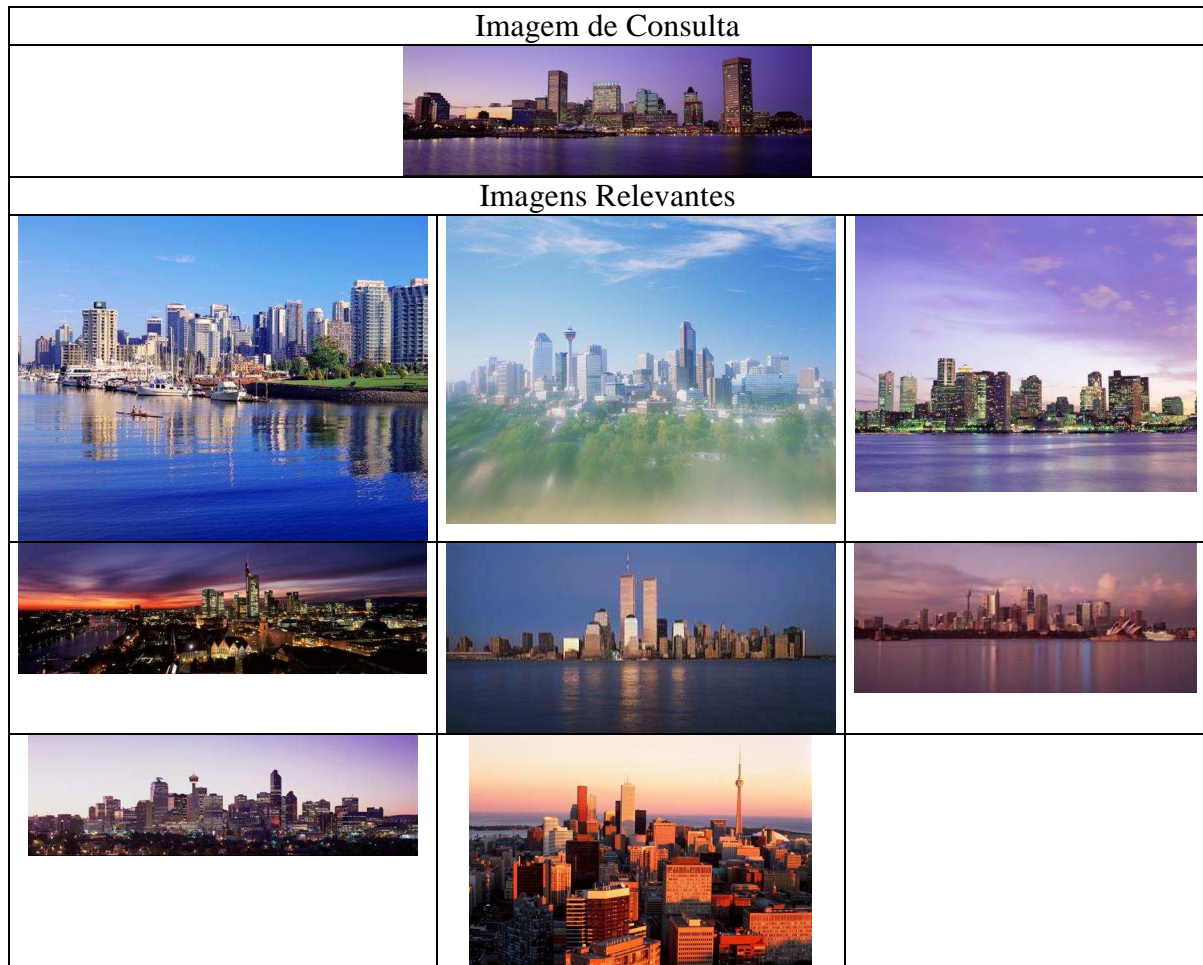


FIGURA 5.5 - Foto de uma paisagem urbana com predominância de edificações e o conjunto de imagens relevantes durante a sua recuperação

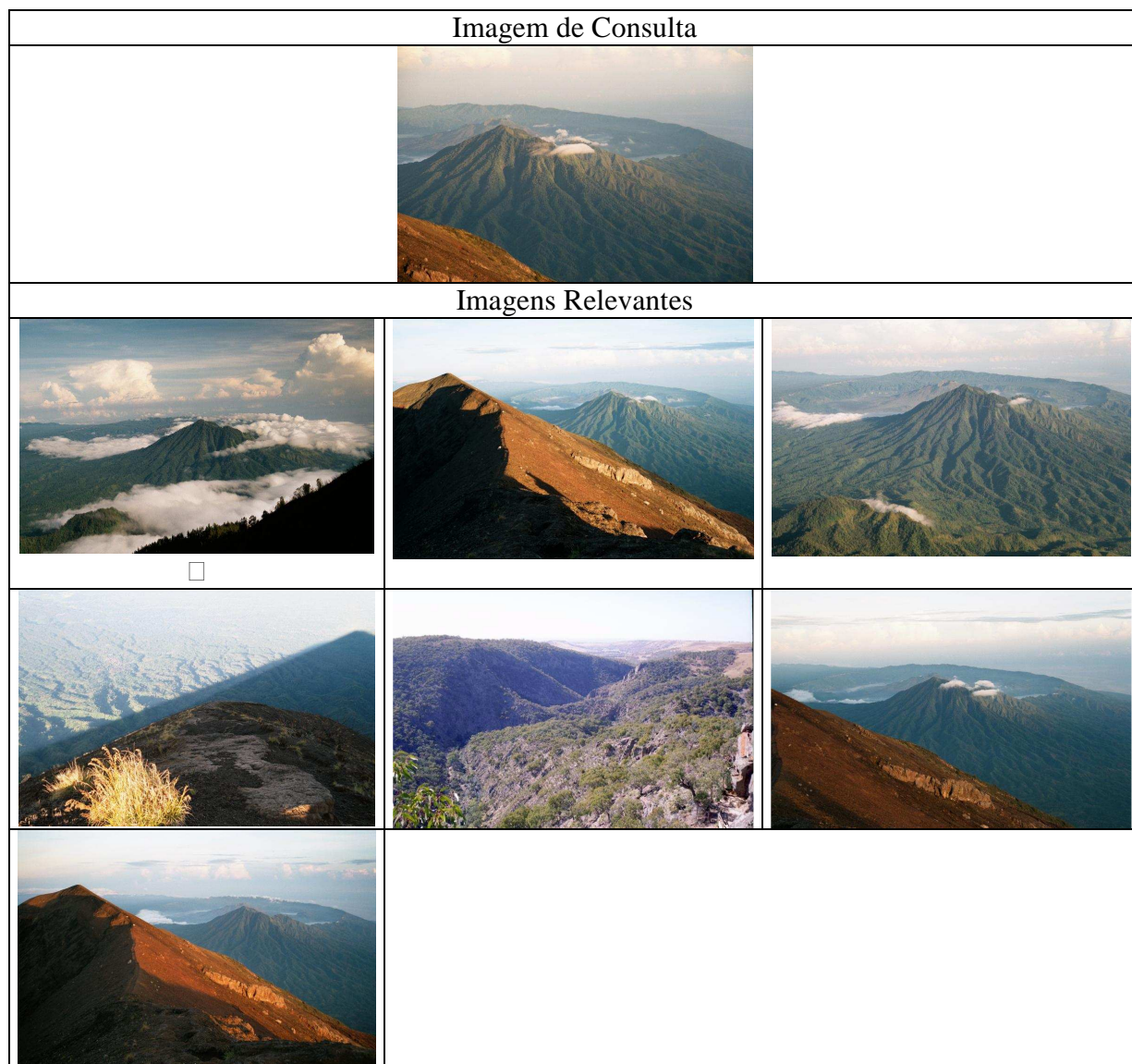


FIGURA 5.6 - Foto de uma paisagem retratando uma montanha e o conjunto de imagens relevantes durante a sua recuperação

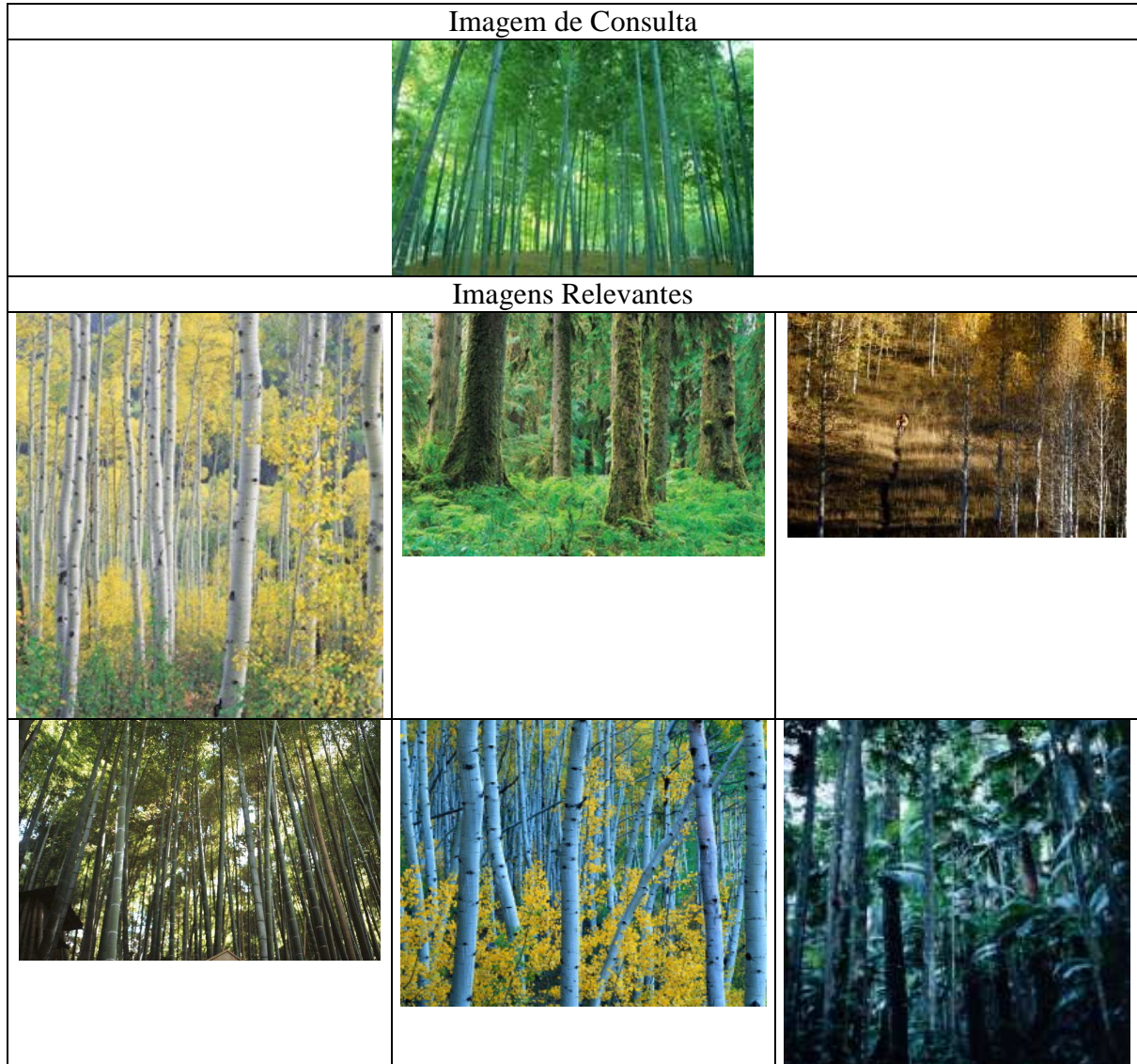


FIGURA 5.7 - Foto de um bosque e o conjunto de imagens relevantes durante a sua recuperação

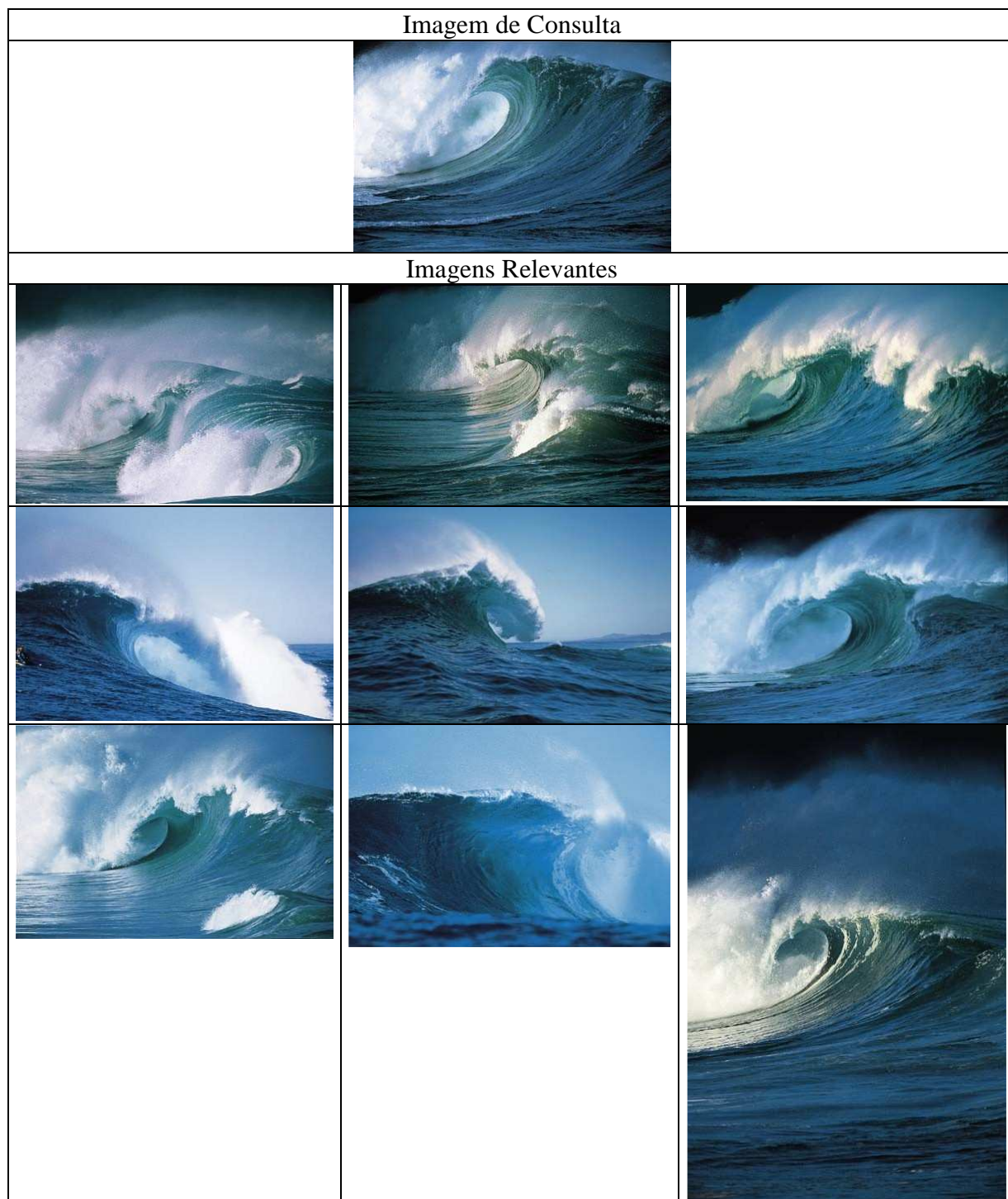


FIGURA 5.8 - Foto de uma onda no mar e o conjunto de imagens relevantes durante a sua recuperação

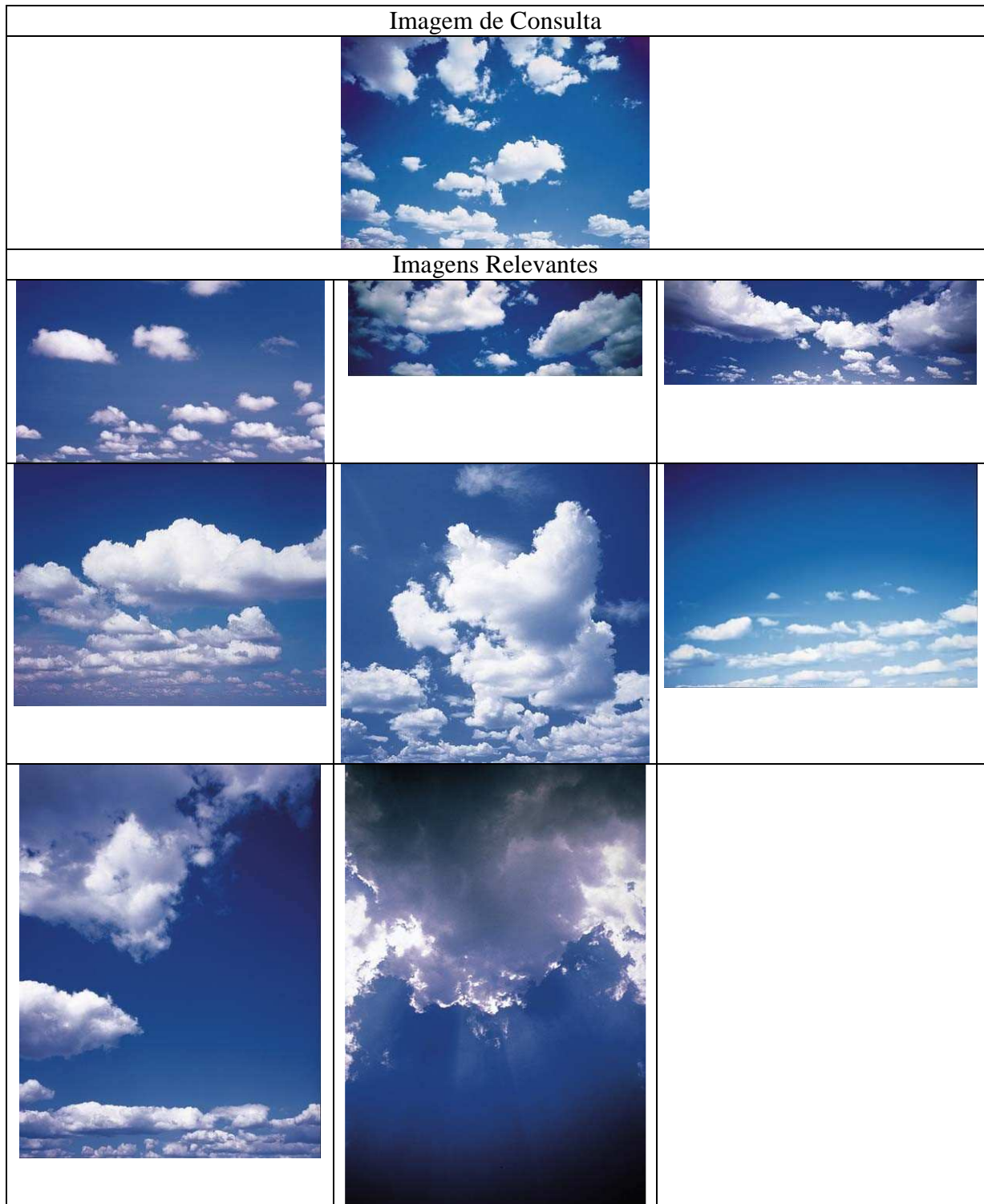


FIGURA 5.9 - Foto de um conjunto de nuvens e o conjunto de imagens relevantes durante a sua recuperação



FIGURA 5.10 - Foto de uma estrada entre árvores e o conjunto de imagens relevantes durante a sua recuperação

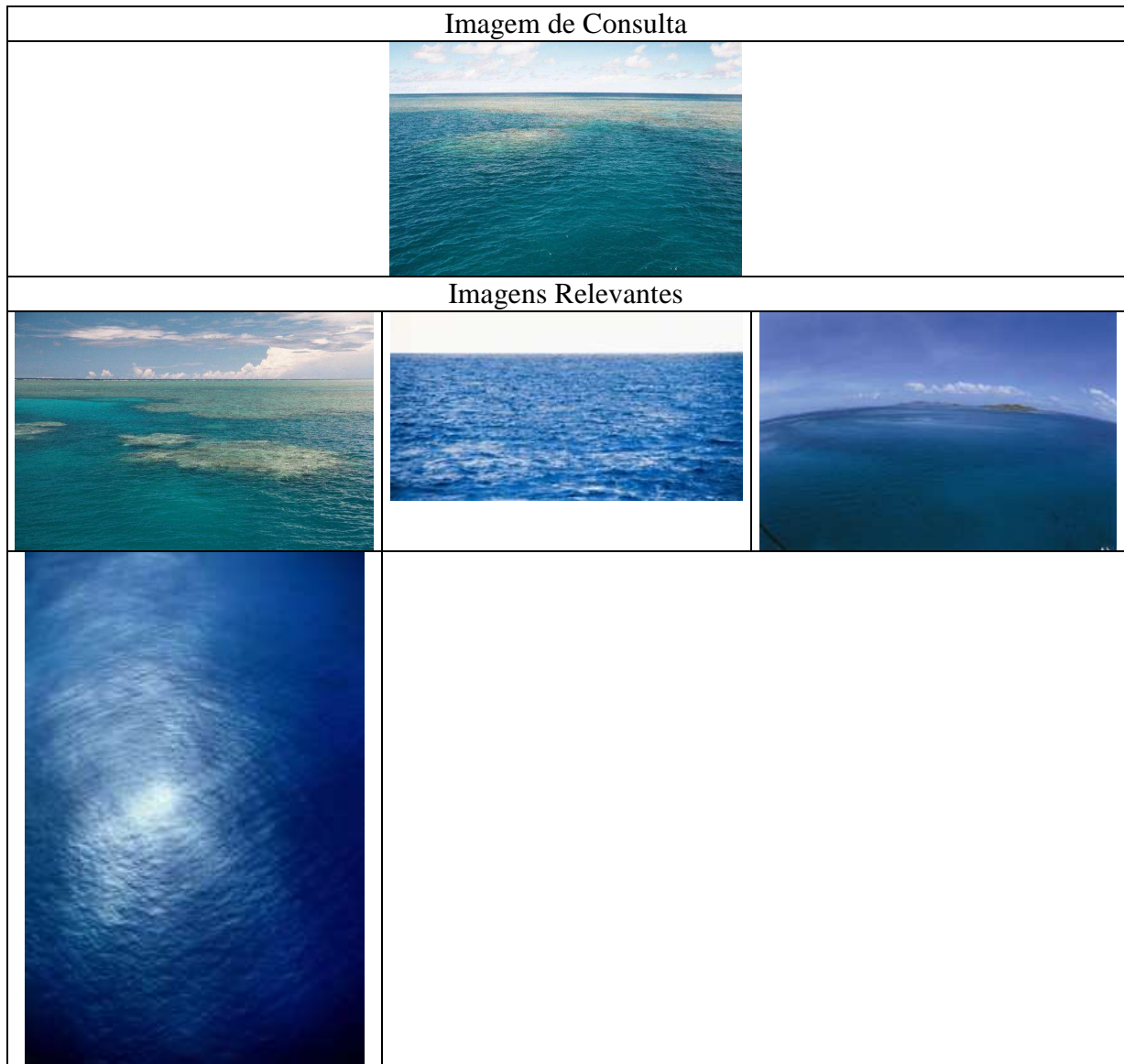


FIGURA 5.11 - Foto da água do mar e o conjunto de imagens relevantes durante a sua recuperação

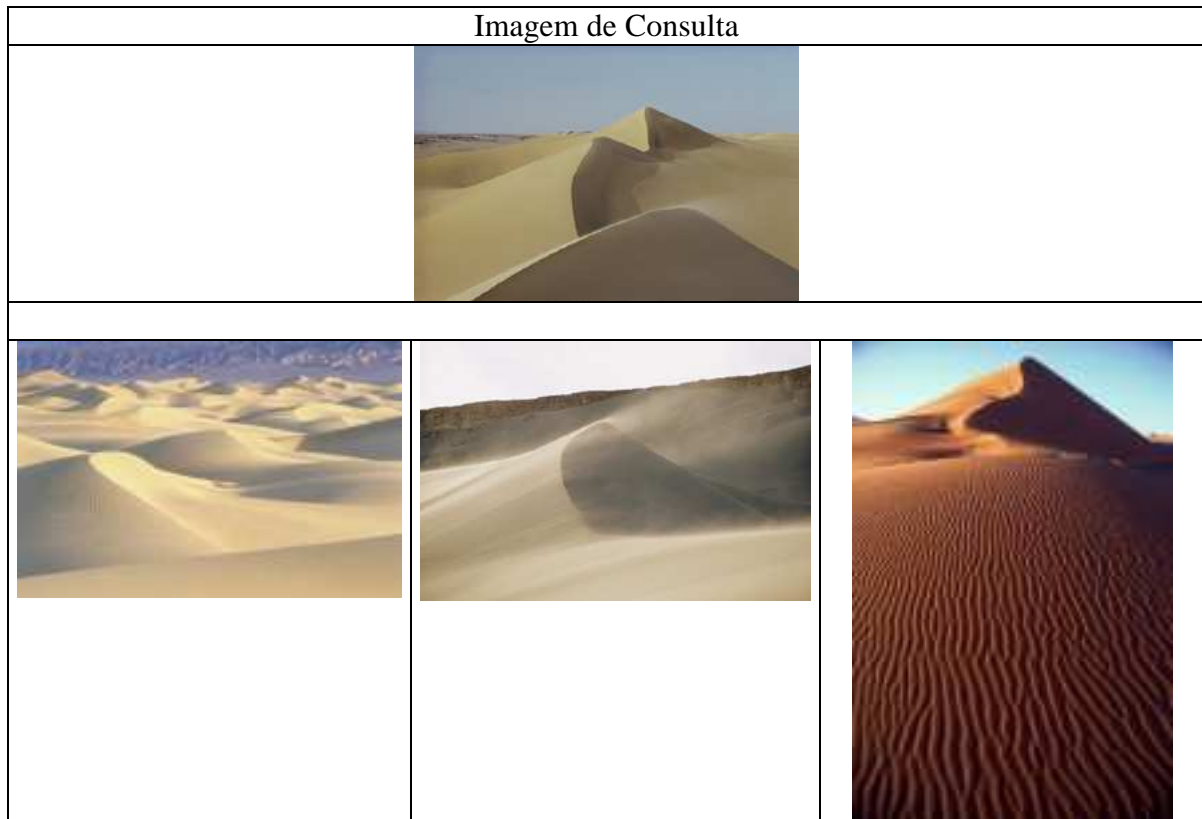


FIGURA 5.12 - Foto de dunas e o conjunto de imagens relevantes durante a sua recuperação

5.2.2 O uso de curvas de “recall x precision”

Considerando os conjuntos de imagens similares descritos na seção anterior, foram realizadas consultas para as dez imagens de referência de cada conjunto. As consultas foram realizadas no sistema RICE com diferentes configurações dos parâmetros ajustáveis. Assim, através de curvas de “recall x precision” construídas durante estas consultas, pode-se avaliar o desempenho das técnicas que foram empregadas a medida em que as configurações de número de zonas, peso do aspecto cromático e estrutural, e espaço de representação de cores são modificadas.

5.2.2.1 Variando o tamanho das zonas

As primeiras consultas tiveram como objetivo avaliar o desempenho do sistema RICE de acordo com o padrão de divisão de zonas estabelecido durante a consulta. Para isto, foram utilizadas as dez imagens de consulta descritas na seção 5.2.1, e a partir da média das dez consultas realizadas foi traçada a curva de “recall x precision” para cada padrão de divisão de zonas previamente definido. Neste caso, foram realizadas consultas nos padrões 1x1, 3x3 e 10x10. Esta avaliação é bastante oportuna no sentido de que favorece a análise da influência da precisão da preservação de informações espaciais durante uma consulta, já que a medida em que se estabelece um número maior de zonas esta precisão aumenta.

Na (figura 5.13), pode-se verificar as curvas de “recall x precision” traçadas para cada uma das dez consultas, utilizando-se a divisão em zonas 1x1, com peso 0,5 tanto para o aspecto cromático quanto para o aspecto estrutural e utilizando-se o espaço RGB para o cálculo da matriz de distâncias de cores.

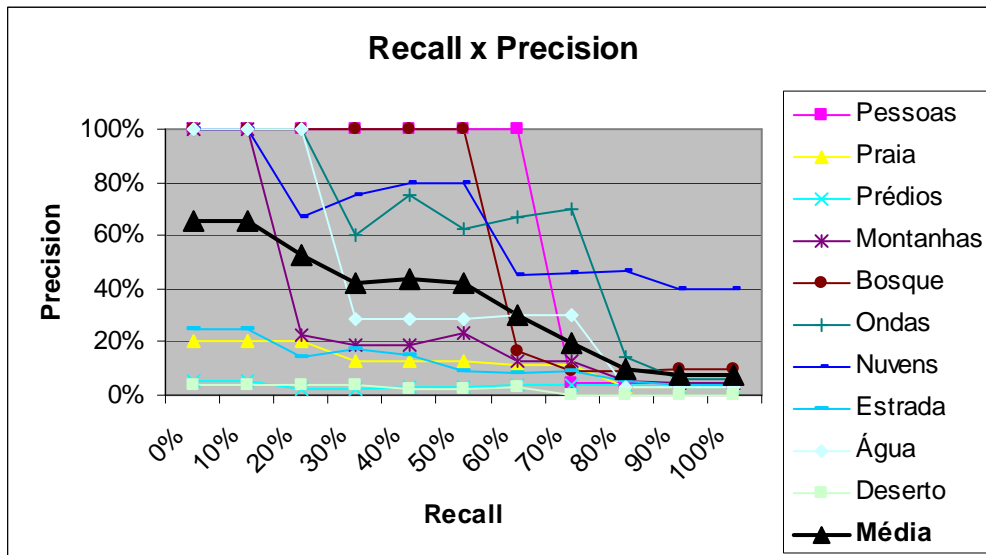


FIGURA 5.13 – Curvas das consultas 1x1, com 50% de cor e 50% de estrutura utilizando o espaço de representação de cores RGB

Neste gráfico, aparece em destaque a curva que representa a média das dez consultas. Esta curva representará então o comportamento do sistema RICE para consultas realizadas com as configurações nelas utilizadas.

A seguir, serão ilustradas as nove imagens de maior similaridade retornadas pelo sistema RICE, com estas configurações de consulta, para uma destas consultas que foram realizadas com esta configuração. Nesta recuperação, a imagem de consulta utilizada foi a que se encontra descrita na (figura 5.14).

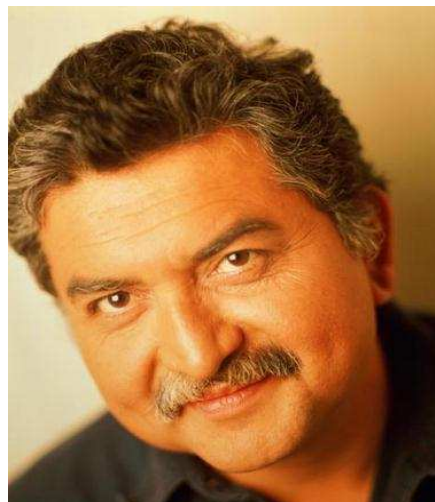


FIGURA 5.14 – Imagem de consulta (1x1, 50% cor, 50% estrutura, RGB)

Na (figura 5.15), são apresentadas em ordem decrescente de similaridade as nove imagens melhor posicionadas nesta consulta.

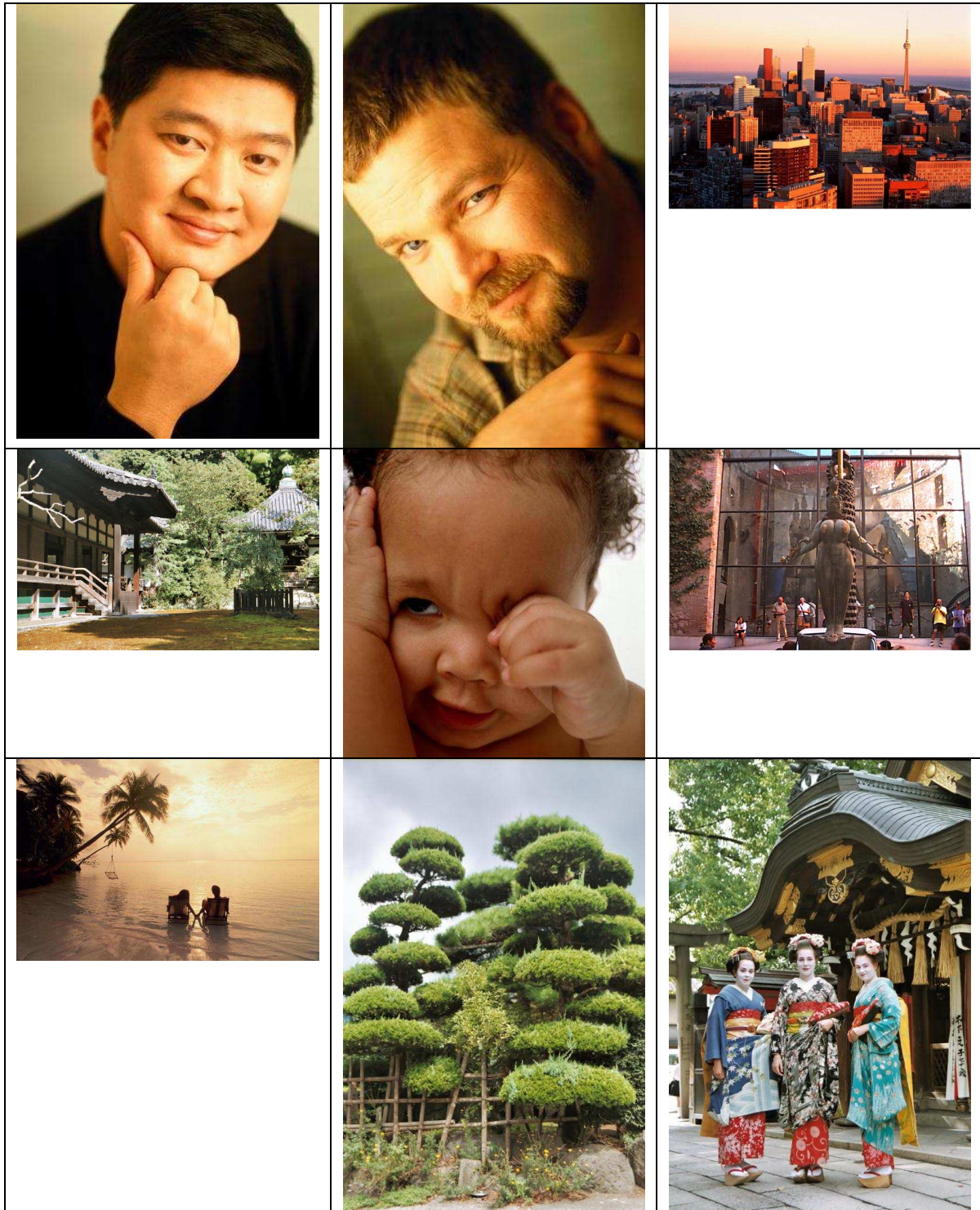


FIGURA 5.15 - As nove imagens mais similares recuperadas na consulta (1x1)

Para as próximas consultas a serem descritas foram traçadas as curvas de “recall x precision” mostradas na (figura 5.16). Neste caso, foi utilizada a divisão em zonas 3x3,

com peso 0,5 tanto para o aspecto cromático quanto para o aspecto estrutural e utilizou-se o espaço RGB para o cálculo da matriz de distâncias de cores.

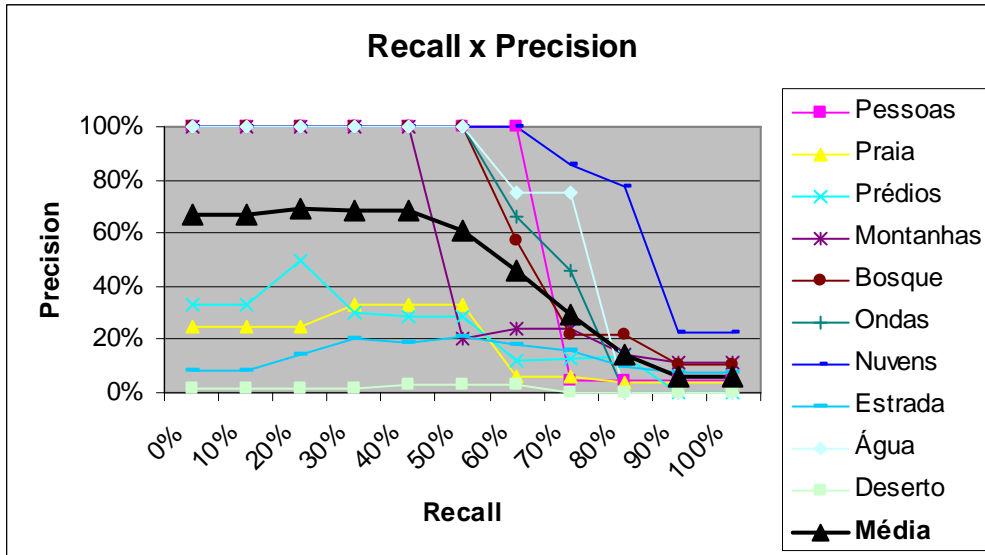


FIGURA 5.16 - Curvas das consultas 3x3, com 50% de cor e 50% de estrutura utilizando o espaço de representação de cores RGB

Pode-se observar novamente em destaque a curva representante da média das dez consultas realizadas. Assim, esta curva representará o desempenho do sistema RICE para consultas com padrão de divisão em zonas 3x3, com peso 0,5 tanto para o aspecto cromático quanto para o estrutural, e utilizando o espaço de cores RGB para o cálculo da matriz de distâncias cromáticas.

A seguir, serão exibidas as nove imagens de maior similaridade retornadas pelo sistema RICE, com estas configurações para recuperação de imagens, em uma das consultas que foram descritas na figura anterior, que utilizou como imagem de consulta a imagem mostrada na (figura 5.17).



FIGURA 5.17 – Imagem de consulta (3x3, 50% cor, 50% estrutura, RGB)

Na (figura 5.18), são apresentadas em ordem decrescente de similaridade as nove imagens melhor posicionadas nesta consulta.



FIGURA 5.18 - As nove imagens mais similares recuperadas na consulta (3x3)

As próximas consultas, ilustradas na (figura 5.19), tiveram suas curvas de “recall x precision” traçadas com um ajuste dos parâmetros configuráveis no qual as imagens foram divididas em 10x10 zonas, com peso 0,5 para os aspectos cromático e estrutural, e novamente com o uso do espaço de representação de cores RGB para o cálculo da matriz de distâncias cromáticas.

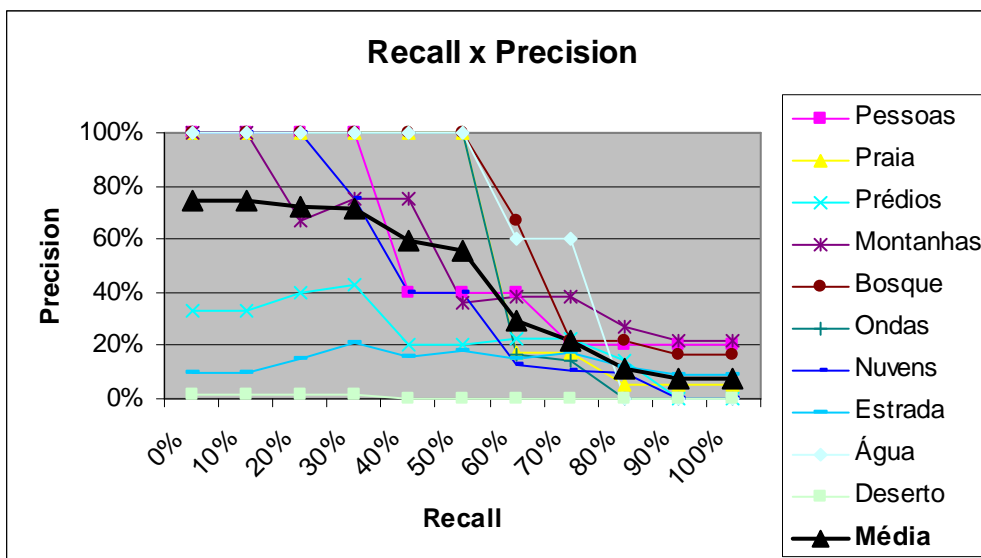


FIGURA 5.19 - Curvas das consultas 10x10, com 50% de cor e 50% de estrutura utilizando o espaço de representação de cores RGB

Pode-se observar novamente em destaque a curva representante da média das dez consultas realizadas. Assim, esta curva representará o desempenho do sistema RICE para consultas com padrão de divisão em zonas 10x10, com peso 0,5 tanto para aspecto cromático quanto para estrutural, e utilizando o espaço de cores RGB para o cálculo da matriz de distâncias cromáticas.

Depois de traçadas as curvas para os diferentes padrões de divisão das imagens em zonas, descritas anteriormente, faz-se oportuno descrever em um único gráfico as três curvas representantes de cada um deles a fim de permitir uma melhor comparação entre os mesmos. Este gráfico está ilustrado na (figura 5.20), descrita a seguir.

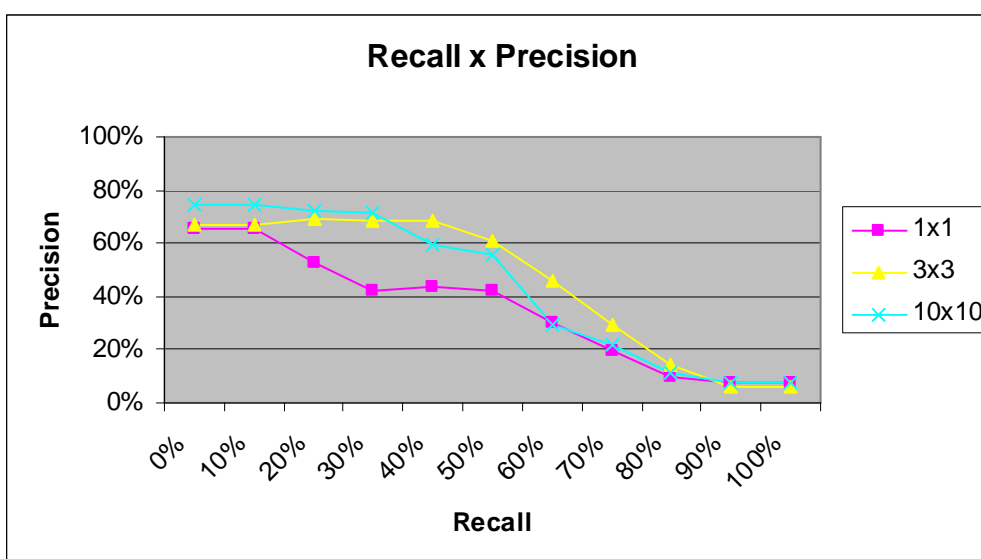


FIGURA 5.20 – Curvas representantes das médias das consultas 1x1, 3x3 e 10x10

As curvas descritas nesta figura mostram que o padrão 1x1 para a divisão da imagem em zonas perde em desempenho para os outros padrões analisados. Isto ilustra que a preservação de informações espaciais, propiciada pela divisão da imagem em zonas, tem uma importância significativa no desempenho das técnicas aqui estudadas. Além disso, pode-se observar que as curvas representantes dos padrões de divisão de zonas 3x3 e 10x10 têm desempenhos semelhantes. Nos primeiros níveis de “recall” as consultas com divisão de zonas no padrão 10x10 apresentaram uma precisão superior, enquanto que a partir dos níveis intermediários de “recall” o padrão de divisão em zonas 3x3 passou a apresentar um melhor desempenho. Isto mostra que o padrão de divisão em zonas 10x10 recupera um maior percentual das imagens relevantes entre aquelas imediatamente mais similares à imagem de consulta, enquanto que o padrão de divisão em zonas 3x3 recupera um maior percentual de imagens relevantes a medida em que a similaridade vai decrescendo.

5.2.2.2 Alterando o espaço de representação das cores

As próximas consultas a serem analisadas tiveram como objetivo avaliar o desempenho do sistema RICE de acordo com o espaço de representação de cores utilizado na construção da matriz de distâncias entre cores empregada no cálculo da medida de similaridade cromática entre as imagens. Para isto, foram usadas as dez imagens de consulta descritas na seção 5.2.1, e a partir da média das dez consultas realizadas foi traçada a curva de “recall x precision” para cada um dos espaços de representação de cores. Neste caso, foram realizadas consultas com matrizes construídas a partir dos espaços para a representação de cores RGB, $L^*u^*v^*$ e $L^*a^*b^*$. Esta avaliação é bastante oportuna no sentido de permitir que se verifique qual é a influência da escolha do espaço de cores nas técnicas de recuperação de imagens aqui estudadas.

Na (figura 5.21), pode-se verificar o gráfico traçado para cada uma das dez consultas, utilizando-se o espaço de cores RGB. Nestas consultas foi considerado apenas o aspecto cromático das imagens (peso 1), e estas foram divididas em zonas no padrão 3x3.

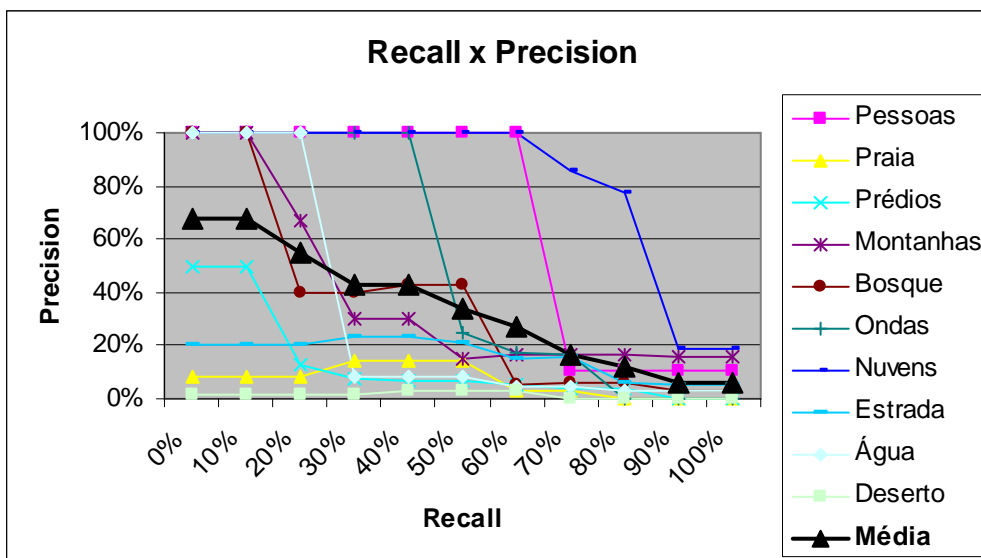


FIGURA 5.21 – Curvas das consultas 3x3, com 100% do aspecto cromático e espaço de cores RGB

Neste gráfico, aparece em destaque a curva que representa a média das dez consultas. Esta curva representará então o comportamento do sistema RICE para consultas realizadas com as configurações nelas utilizadas.

Para as próximas consultas a serem descritas foram traçadas as curvas de “recall x precision” mostradas na (figura 5.22). Neste caso, foi utilizada a divisão em zonas 3x3, com peso 1 para o aspecto cromático e foi utilizado o espaço $L^*a^*b^*$ para o cálculo da matriz de distâncias de cores.

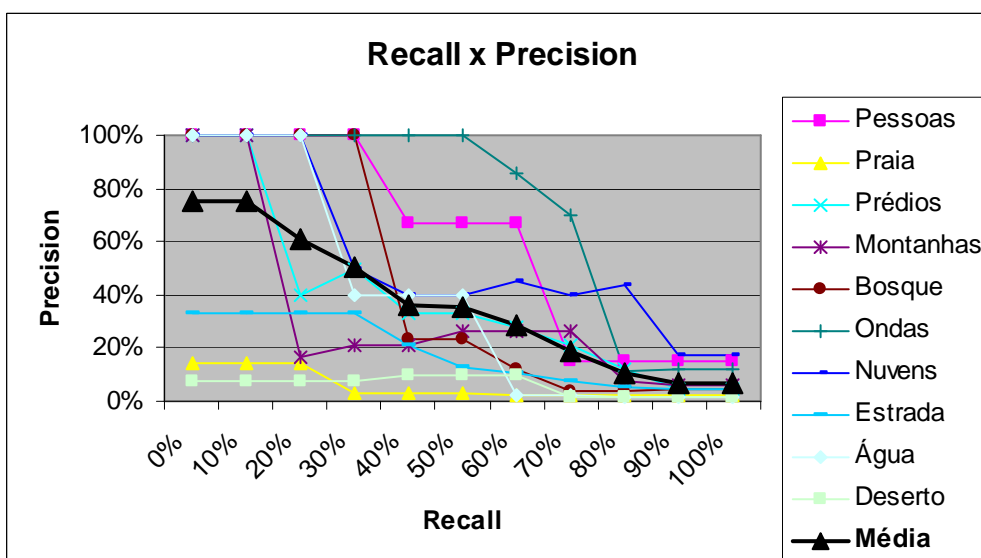


FIGURA 5.22 - Curvas das consultas 3x3, com 100% do aspecto cromático e espaço de cores $L^*a^*b^*$

Pode-se observar novamente em destaque a curva representante da média das dez consultas realizadas. Assim, esta curva representará o desempenho do sistema RICE para consultas com padrão de divisão em zonas 3x3, com peso 1 para o aspecto cromático, e utilizando o espaço de cores $L^*a^*b^*$ para o cálculo da matriz de distâncias cromáticas.

As próximas consultas, ilustradas na (figura 5.23), tiveram suas curvas de “recall x precision” traçadas com um ajuste dos parâmetros configuráveis no qual as imagens foram divididas em 3x3 zonas, com peso 1 para o aspecto cromático e com o uso do espaço de representação de cores $L^*u^*v^*$ para o cálculo da matriz de distâncias cromáticas.

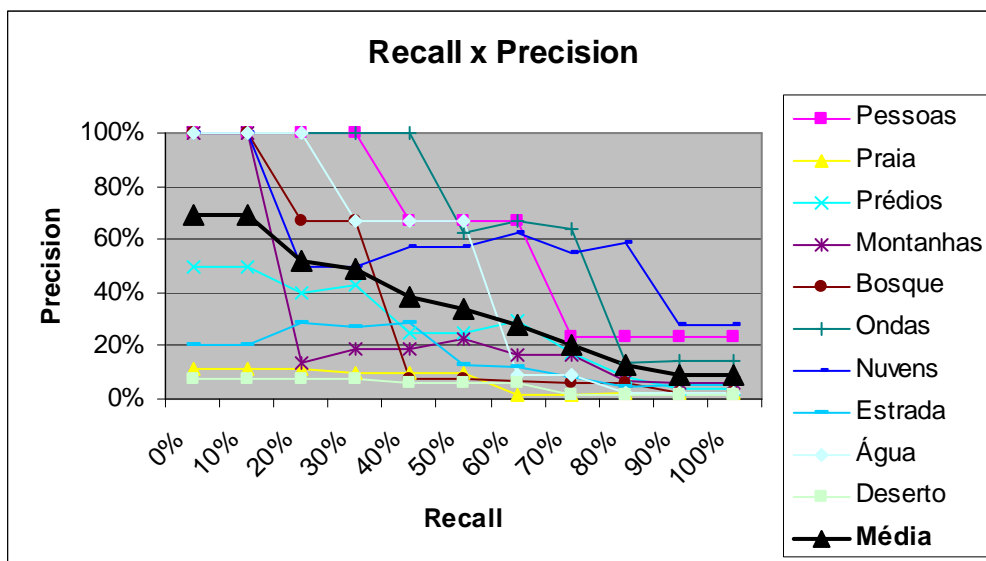


FIGURA 5.23 - Curvas das consultas 3x3, com 100% do aspecto cromático e espaço de cores $L^*u^*v^*$

Pode-se observar novamente em destaque a curva representante da média das dez consultas realizadas. Assim, esta curva representará o desempenho do sistema RICE para consultas com padrão de divisão em zonas 3x3, com peso 1 para o aspecto cromático, e utilizando o espaço de cores $L^*u^*v^*$ para o cálculo da matriz de distâncias cromáticas.

Depois de traçadas as curvas baseadas no uso de diferentes espaços de representação de cores, faz-se oportuno descrever em um único gráfico as três curvas representantes de cada um deles a fim de permitir uma melhor comparação entre os mesmos. Este gráfico está ilustrado na (figura 5.24), descrita a seguir.

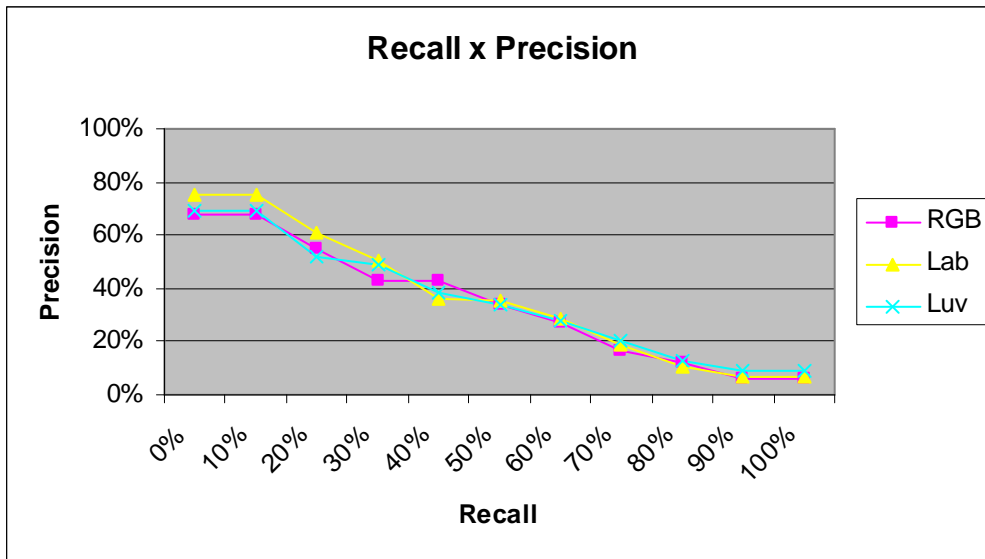


FIGURA 5.24 - Curvas representantes das médias das consultas 3x3, com 100% do aspecto cromático para os espaços de representação de cores RGB, $L^*a^*b^*$ e $L^*u^*v^*$

As curvas descritas neste gráfico mostram que houve uma ligeira superioridade no desempenho do sistema nos primeiros níveis de “recall” quando foi utilizado o espaço de cores $L^*a^*b^*$. Isto mostra que o uso do espaço de representação de cores $L^*a^*b^*$ recupera um maior percentual das imagens relevantes entre aquelas imediatamente mais similares à imagem de consulta. Entretanto, é importante observar que, de uma forma geral, os desempenhos dos três espaços de representação de cores aqui estudados foram muito semelhantes.

A proximidade dos desempenhos dos três espaços de cores aqui estudados pode ser explicada pelo fato de que quando são escolhidas as imagens relevantes para cada consulta, é levado em consideração o conteúdo semântico das imagens. Desta forma, muitas imagens tidas como similares podem não apresentar necessariamente conteúdos cromáticos similares. Este problema é um obstáculo enfrentado por muitas técnicas de recuperação de imagens e uma discussão acerca do mesmo pode ser encontrada em [SME 2000].

5.2.2.3 Alterando os pesos dos aspectos cromático e estrutural

Em todas as zonas

Uma outra importante análise a ser considerada, diante das técnicas estudadas neste trabalho, é aquela que procura verificar a importância dos aspectos visuais utilizados no processo de recuperação de imagens, neste caso os aspectos cromático e estrutural. Para isto, foram traçadas curvas de “recall x precision” para consultas baseadas exclusivamente no aspecto cromático das imagens, outras para consultas baseadas exclusivamente no aspecto estrutural da imagem, e outras para consultas que envolvem estes dois aspectos visuais na comparação entre imagens.

A (figura 5.21), descrita anteriormente, ilustra as curvas de “recall x precision” traçadas a partir das dez imagens de consultas empregadas nas consultas anteriormente

descritas neste trabalho. Neste caso, as consultas foram realizadas levando-se em consideração apenas o conteúdo cromático das imagens e sendo que estas foram divididas em 3x3 zonas para a construção de seus respectivos descritores. Durante a realização destas consultas foi utilizado o espaço de cores RGB para a construção da matriz de distâncias cromáticas. O gráfico apresenta em destaque a linha correspondente à média das dez consultas realizadas. Esta linha representará o comportamento do sistema RICE para as consultas que sejam baseadas apenas no conteúdo cromático das imagens.

A seguir será descrita a (figura 5.25), que ilustra o comportamento do sistema RICE nos processos de recuperação de imagens baseados apenas no conteúdo estrutural da imagem. A linha em destaque no gráfico representa a média das dez consultas realizadas com a imagem sendo dividida em 3x3 zonas e considerando-se apenas o aspecto estrutural. Desta forma, esta linha representará o comportamento do sistema quando utilizando sob esta configuração.

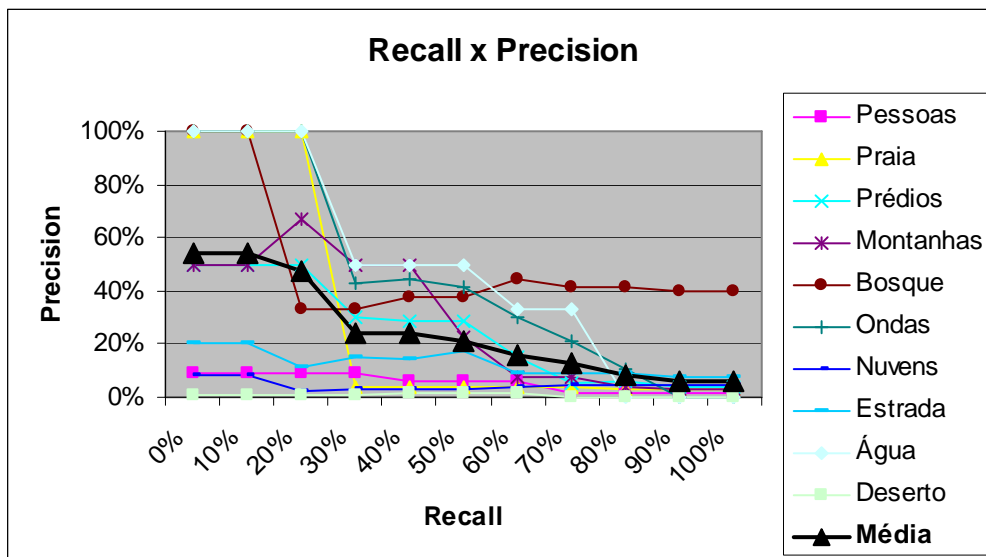


FIGURA 5.25 – Curvas das consultas 3x3 com 100% do aspecto estrutural

Conforme já descrito e ilustrado na (figura 5.19), foram realizadas consultas dividindo as imagens em 3x3 zonas para a formação de seus descritores com pesos 0,5 tanto para o aspecto cromático, quanto para o aspecto estrutural. Desta forma, seria oportuno traçar agora um gráfico para ilustrar as curvas representantes dos três tipos de consultas mencionados nesta seção, cujas configurações se diferem no que diz respeito aos pesos dos aspectos visuais cor e estrutura no processo de recuperação da imagem. Este gráfico é exibido na (figura 5.26).

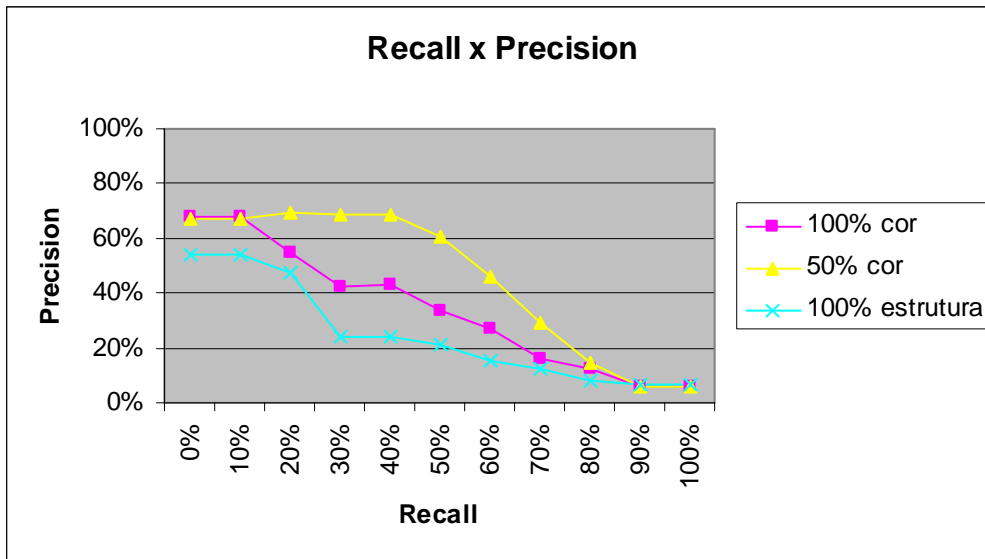


FIGURA 5.26 – Curvas das médias das consultas 3x3 com 100% do aspecto cromático, 50% para cromático e 50% para estrutural, e 100% para estrutural

A partir do gráfico exibido acima, pode-se observar que os resultados obtidos através das avaliações aqui descritas confirmam as expectativas em torno da importância da integração de diferentes aspectos visuais na recuperação de imagens, conforme foi colocado na seção 2.1.4. As linhas traçadas no gráfico mostram que a combinação dos aspectos cromático e estrutural leva uma vantagem significativa em termos de desempenho quando comparada ao uso destas mesmas características visuais isoladamente. Constatou-se que com a combinação destas duas características a precisão da recuperação é mantida em níveis mais altos até os níveis intermediários de “recall”, fato que não acontece quando as características visuais são utilizadas isoladamente. Este fato indica que a combinação dos dois aspectos visuais aqui estudados proporciona resultados nos quais as imagens relevantes ficam melhor posicionadas em relação à imagem de consulta, já que os valores de precisão são maiores. As curvas referentes às recuperações baseadas exclusivamente em um aspecto visual mostram que o uso apenas do conteúdo cromático durante o processo de recuperação produz resultados melhores do que o uso apenas de conteúdo estrutural. Observou-se durante a realização das consultas que isto acontece pelo fato de que o conteúdo estrutural tende a produzir bons resultados apenas em situações onde as imagens possuem uma forte predominância de bordas que seguem uma mesma orientação.

Em algumas zonas

A última análise a ser descrita neste trabalho, leva em consideração a possibilidade de que o usuário do sistema RICE ajuste pesos diferentes para os aspectos cromático e estrutural em cada zona criada na imagem para a formação do seu descritor. Este tipo de ajuste para os pesos das características visuais se baseia na descrição presente na seção 3.1.8. Nas consultas que serão descritas a seguir foram utilizadas as imagens de consulta ilustradas em (figura 5.4), (figura 5.5), (figura 5.8) e (figura 5.12). É importante observar o ajuste dos pesos para os aspectos cromático e estrutural para cada zona da imagem devem ser estabelecidos levando-se em consideração particularidades referentes ao conteúdo (cromático e estrutural) da imagem de consulta que se pretende recuperar. Este fato exige que o usuário tenha uma certa habilidade para identificar os

pontos da imagem onde cada um destes aspectos deve receber um peso maior. Adicionalmente, é importante colocar que um ajuste mal estabelecido para estes pesos pode comprometer seriamente o processo de recuperação, derrubando significativamente a precisão dos resultados obtidos. Estes ajustes devem ser feitos a partir de características da imagem de consulta, isto fez com que fossem selecionadas apenas algumas imagens e os resultados obtidos fossem comparados com resultados obtidos em outras consultas realizadas com a mesma imagem, mas com ajustes diferentes dos pesos dos aspectos visuais.

A fim de verificar o quanto este tipo de ajuste pode influenciar no processo de recuperação foram realizadas as consultas que serão descritas a partir de agora. Inicialmente, foi tomada como imagem de consulta aquela contendo dunas, ilustrada na primeira posição na (figura 5.12). Observando os conteúdos cromático e estrutural desta imagem, e considerando a divisão da mesma em 3x3 zonas, foram estabelecidos os pesos descritos na (figura 5.27) para as zonas da imagem.

Cor 50% Est. 50%	Cor 25% Est. 75%	Cor 50% Est. 50%
Cor 75% Est. 25%	Cor 25% Est. 75%	Cor 75% Est. 25%
Cor 75% Est. 25%	Cor 25% Est. 75%	Cor 75% Est. 25%

FIGURA 5.27 – Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem das dunas

Estabelecidos estes pesos, foi realizada uma consulta no sistema RICE e o resultado obtido foi comparado com a recuperação da mesma imagem em outras três situações em que a imagem também foi dividida em 3x3 zonas: peso 1 para o conteúdo cromático em todas as zonas (somente cor); peso 1 para o conteúdo estrutural em todas as zonas (somente estrutura); peso 0,5 para cor e estrutura em todas as zonas da imagem. Os resultados obtidos estão ilustrados na (figura 5.28), descrita a seguir.

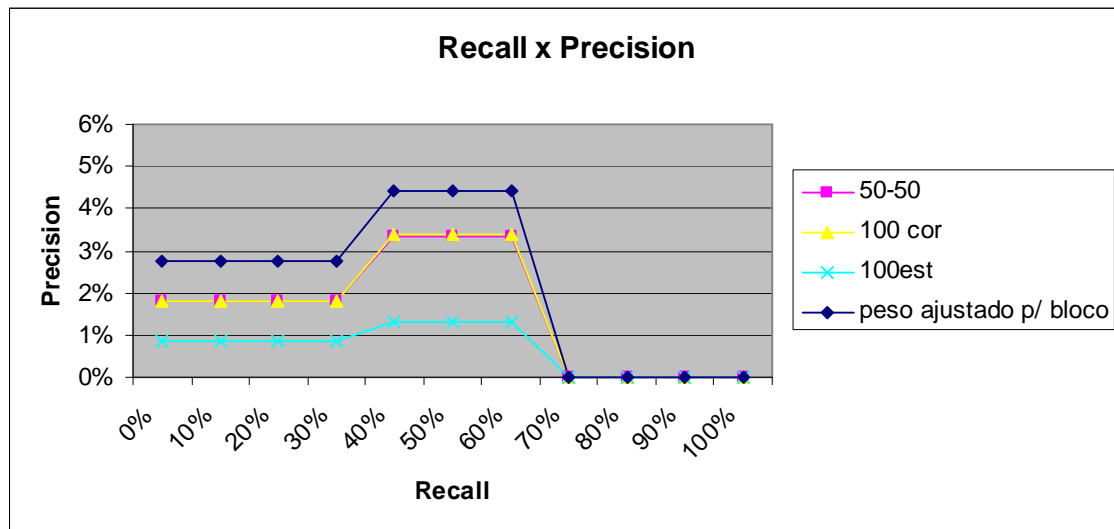


FIGURA 5.28 – Curvas obtidas na consulta da imagem das dunas com pesos específicos para cada zona da imagem

Neste caso, pode-se observar pela curva “recall x precision” descrita em azul marinho, referente à consulta em que os pesos foram atribuídos tal como apresentado na (figura 5.27), que o resultado obtido apresentou um desempenho superior aos demais desde os baixos níveis de “recall” até os intermediários, igualando-se aos demais nos últimos níveis de “recall”.

As próximas consultas seguiram o modelo da consulta anterior. Entretanto, as imagens utilizadas como imagem de consulta foram as imagens de consulta apresentadas nas primeiras posições de (figura 5.8) (onda), (figura 5.4) (praias) e (figura 5.5) (prédios). Para estas imagens, foram adotadas as ponderações de aspecto cromático e estrutural descritas em (figura 5.29), (figura 5.30) e (figura 5.31) respectivamente.

Cor 75% Est. 25%	Cor 75% Est. 25%	Cor 75% Est. 25%
Cor 50% Est. 50%	Cor 75% Est. 25%	Cor 50% Est. 50%
Cor 50% Est. 50%	Cor 50% Est. 50%	Cor 50% Est. 50%

FIGURA 5.29 - Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem da onda

Cor 30% Est. 70%	Cor 30% Est. 70%	Cor 60% Est. 40%
Cor 30% Est. 70%	Cor 30% Est. 70%	Cor 60% Est. 40%
Cor 50% Est. 50%	Cor 50% Est. 50%	Cor 50% Est. 50%

FIGURA 5.30 - Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem das praias

Cor 70% Est. 30%	Cor 70% Est. 30%	Cor 70% Est. 30%
Cor 30% Est. 70%	Cor 30% Est. 70%	Cor 30% Est. 70%
Cor 70% Est. 30%	Cor 70% Est. 30%	Cor 70% Est. 30%

FIGURA 5.31 - Distribuição dos pesos para os aspectos visuais utilizados na consulta da imagem das prédios

Com estas configurações de pesos descritas nas figuras anteriores estabelecidas para as imagens de consulta 5.8, 5.4 e 5.5, os resultados obtidos foram os seguintes, descritos respectivamente em (figura 5.29), (figura 5.30) e (figura 5.31).

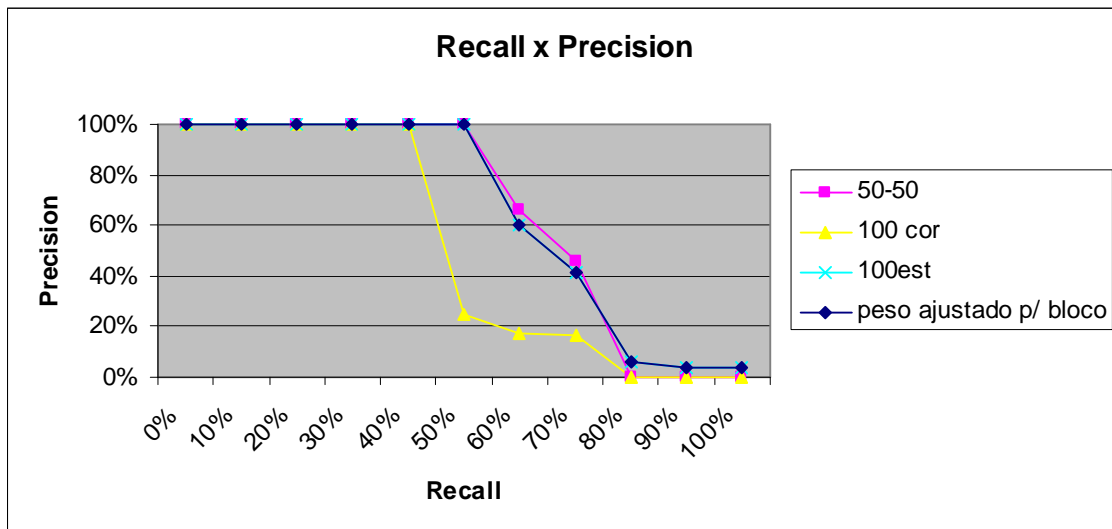


FIGURA 5.32 – Curva da consulta média da imagem da onda com pesos específicos para as zonas comparada com outras configurações de consulta para a mesma imagem

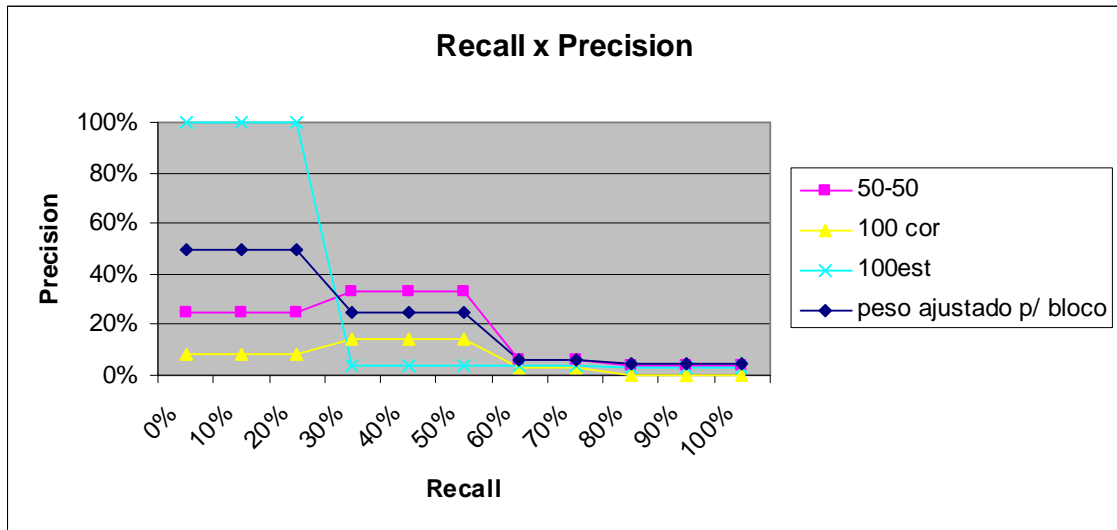


FIGURA 5.33 - Curva da consulta média da imagem da praia com pesos específicos para as zonas comparada com outras configurações de consulta para a mesma imagem

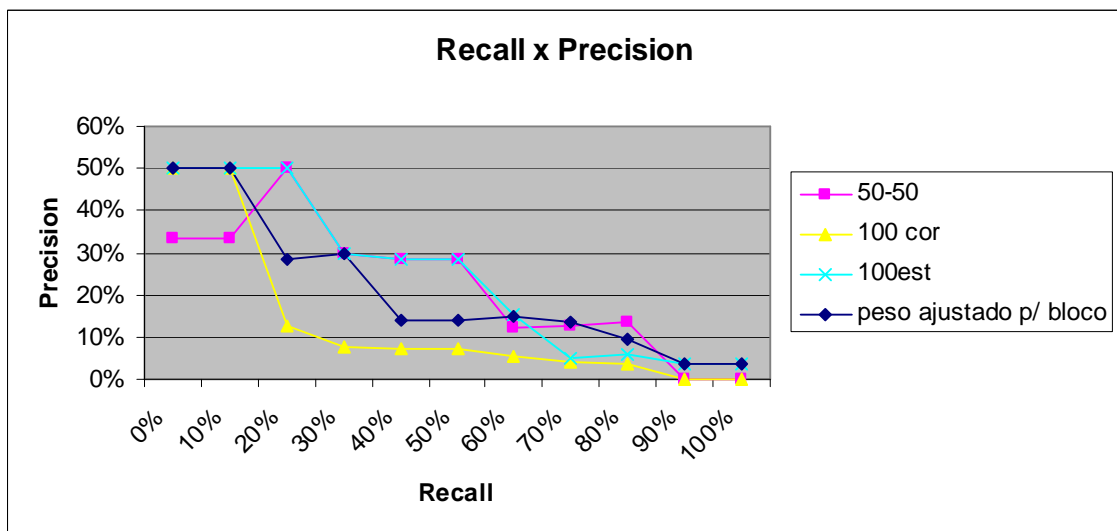


FIGURA 5.34 - Curva da consulta média da imagem de prédios com pesos específicos para as zonas comparada com outras configurações de consulta para a mesma imagem

Nas três últimas consultas descritas pode-se observar que, de uma forma geral, não houve ganho de desempenho com o uso de pesos específicos para os aspectos cromático e estrutural em cada zona da imagem. Isto mostra que o ajuste dos pesos de cor e estrutura em cada zona da imagem não é uma tarefa simples já que em muitas situações não proporciona melhoria dos resultados. A percepção dos conteúdos cromáticos e estruturais para o elemento humano é algo bastante subjetivo, isto faz com que muitas vezes não se atinja os resultados esperados em consultas deste tipo.

5.3 Conclusões

Neste capítulo foram apresentados resultados experimentais obtidos através do sistema RICE. Para a realização destes experimentos, utilizou-se uma coleção com 15 padrões cromáticos e estruturais artificiais, e uma coleção com 287 imagens reais com conteúdo semântico heterogêneo.

Os experimentos realizados com as imagens artificiais procuraram apenas demonstrar o comportamento das técnicas de recuperação de imagens aqui descritas diante de algumas situações específicas. Já os experimentos realizados com imagens reais, procuraram avaliar, através de curvas de “recall x precision”, o desempenho destas técnicas em situações reais de uso.

As principais conclusões aqui apontadas estão relacionadas ao desempenho das técnicas de recuperação de imagens com o uso de: diferentes espaços de cores, diferentes padrões de divisão da imagem em zonas, e diferentes ponderações para os aspectos visuais (cor e estrutura).

6 Conclusão

Neste capítulo serão apresentadas as considerações finais do trabalho. Serão discutidos os pontos positivos e negativos das técnicas aqui estudadas, bem como serão descritos possíveis trabalhos futuros.

Este trabalho envolveu o estudo da indexação de imagens por seus conteúdos cromáticos e estruturais. Com este estudo foi identificada a possibilidade de se descrever tanto o aspecto cromático quanto o aspecto estrutural das imagens através do uso de histogramas. Diante da homogeneidade destas estruturas, foi possível a aplicação de uma única métrica de verificação da similaridade entre imagens para ambos os casos. Para isto, foi empregada a distância euclidiana ponderada de forma que fosse possível ajustar alguns elementos importantes que podem influenciar no processo de recuperação de imagens. Entre estes elementos encontra-se o espaço de representação de cores utilizado para o cálculo de distâncias cromáticas durante o processo de recuperação da imagem; o nível de precisão da preservação das informações espaciais da imagem durante o processo de recuperação, que pode ser ajustado através da divisão da imagem em uma quantidade maior ou menor de zonas; e os pesos dos aspectos visuais utilizados nas técnicas aqui descritas, que são cor e estrutura.

A principal contribuição deste trabalho veio no sentido de empregar um método para avaliar o desempenho das técnicas de recuperação de imagens aqui descritas diante das diferentes possibilidades de ajuste dos elementos supracitados, pois, embora algumas destas técnicas sejam bastante conhecidas, não se encontra na literatura muitos trabalhos destinados especificamente a testá-las. Para isto, foi implementado o sistema RICE, um sistema para a recuperação de imagem por cor e estrutura, a fim de que fosse possível realizar os referidos testes e comparar resultados. Este sistema foi desenvolvido no ambiente de programação Delphi 6.0 e utilizou o banco de dados Inter Base para o armazenamento dos descritores de conteúdo das imagens.

A partir do sistema RICE, construído ao longo deste trabalho, foi possível a realização de testes envolvendo a recuperação de imagens utilizando curvas de “recall x precision”, uma metodologia para avaliação de desempenho de sistemas de recuperação de informação amplamente difundida inclusive na comunidade voltada à pesquisa em recuperação de imagens. Para isto, foi utilizada uma coleção contendo 287 imagens com conteúdos semânticos bastante diversificados. Dentre estas imagens, dez foram escolhidas para serem utilizadas como imagens de consulta durante a realização dos testes. Para cada uma destas imagens foram relacionadas aquelas que seriam consideradas relevantes quando da sua recuperação, já que isto é imprescindível para a construção das curvas de “recall x precision”.

Estabelecidas as imagens de consulta, o sistema foi testado através da realização de consultas com diferentes ajustes dos seus parâmetros configuráveis para cada uma delas. Dentre estas diferentes configurações, foi variado o espaço de cores utilizado no cálculo das diferenças cromáticas; o número de zonas criadas para preservar as informações espaciais do conteúdo da imagem; e o peso dos aspectos cromático e estrutural durante a recuperação das imagens.

Traçadas as curvas de “recall x precision” para estas diferentes possibilidades de consulta, elas foram organizadas de forma que se pudesse comparar entre si: o desempenho dos diferentes espaços de cores; a influência da precisão da preservação de informações espaciais; e a importância da combinação de diferentes aspectos visuais durante uma consulta. Com esta organização, as principais observações que foram extraídas indicaram que a combinação de diferentes aspectos visuais e a preservação de informações espaciais durante a construção dos descritores de imagens são fatores que podem influenciar positivamente de forma significativa o processo de recuperação de imagens baseada na métrica aqui estudada. Estes fatos vieram de encontro as expectativas que haviam no início do trabalho. Por outro lado, os resultados experimentais mostraram que o uso de espaços de representação de cores perceptualmente uniformes ($L^*u^*v^*$ e $L^*a^*b^*$) não apresentaram desempenho significativamente superior ao desempenho apresentado pelo espaço de representação de cores RGB na recuperação de imagens descrita neste trabalho. Este fato contrariou as expectativas que haviam no início do trabalho. Adicionalmente, foi verificado o desempenho da técnica de recuperação aqui abordada diante da aplicação de pesos diferentes para os aspectos cromático e estrutural em cada zona obtida na divisão da imagem. Os resultados obtidos mostraram que a subjetividade da percepção humana pode influenciar decisivamente no estabelecimento dos pesos para cada zona, sendo um fator que pode impedir o bom desempenho do sistema nestas condições.

6.1 Trabalhos Futuros

Este trabalho teve como principal objetivo a realização de avaliações de desempenho no que diz respeito à efetividade das técnicas de recuperação aqui descritas. Contudo, existe um outro aspecto muito importante inerente às técnicas de recuperação de imagens relacionado ao tempo gasto durante uma consulta. Desta forma, trabalhos futuros poderiam ser desenvolvidos nos sentido de se avaliar a eficiência das técnicas de recuperação envolvendo o fator tempo de consulta.

O desenvolvimento de estudos voltados a um melhor desempenho das técnicas de recuperação de imagens em termos de tempo de consulta pode, inclusive, contribuir com o desenvolvimento de aplicações de recuperação de imagens para a Web, já que, nesta, o tempo é um elemento de extrema importância e a quantidade de imagens disponíveis cresce rapidamente.

Bibliografia

- [ANT 2000] ANTUNES, Eurico Jardim. **Um Estudo sobre as Técnicas de Compressão e Indexação de Imagens**. 2000. Trabalho Individual (Mestrado em Ciência da Computação) – Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre.
- [ASS 2000] ASSFALG, Jüergen; PALA, Pietro. Querying by Photographs: A VR Metaphor for Image Retrieval. **IEEE Multimedia**, Los Alamitos v. 7, n. 1, p. 52-59, 2000.
- [BEA 99] BEAZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. [S.l.]: Addison-Wesley, 1999. p. 74-81.
- [BIM 99] BIMBO, Alberto del. **Visual Information Retrieval**. San Francisco: Morgan Kaufmann, 1999. p. 81-113.
- [CHU 98] CHUA, Tat-Seng et al. **Relevance Feedback Techniques for Color-based Image Retrieval**. Lausanne: IEEE Press, 1998. p. 24-31.
- [COH 99] COHEN, Scott. **Finding Color and Shape Patterns in Images**. 1999. Disponível em: <<http://robotics.stanford.edu/~scohen>>. Acesso em: 25 maio 2002.
- [COM 2002] COMPUTER Vision Test Images: Disponível em: <<http://www.cs.washington.edu/research/imagedatabase/groundtruth/>>. Acesso em: 10 jun. 2002.
- [FAC 2000] FACON, Jaques. **Princípios Básicos da Visão por Computador e Processamento de Imagens**. Disponível em: <<ftp://ftp.ppgia.pucpr.br>>. Acesso em: 14 ago. 2000.
- [FAC 2001] FACON, Jaques. **Processamento e Análise de Imagens**. Curitiba: Pontifícia Universidade Católica do Paraná, 2001. Disponível em: <<ftp://ftp.ppgia.pucpr.br>>. Acesso em: 04 fev. 2001.
- [GEV 199?] GEVERS, Theo. **Color in Image Databases. Intelligent Sensory Information Systems**. Amsterdam: University of Amsterdam, [199?]. Disponível em: <<http://carol.wins.uva.nl/~gevers/GeversCV.html>>. Acesso em: 15 abr. 2002.
- [GOM 94] GOMES, Jonas; VELHO, Luiz. **Computação Gráfica: imagem**. Rio de Janeiro: IMPA/SBM, 1994. p. 63-127.
- [GOO 2000] GOODRUM, Abby A. Image Informations Retrieval: An overview of current research. **Informing Science**, [S.l.], v. 3, n. 2, p. 63-67, 2000.

- [JAI 96] JAIN, Anil K.; VAILAYA, Aditya. Image Retrieval using Color and Shape. **Pattern Recognition**, Oxford, v. 29, n. 8, p. 1233-1244, 1996.
- [PAR 99] PARK, Du-Sik et al. **Image Indexing using Color Histogram in the CIELUV color Space**. Disponível em: <<http://citeseer.nj.nec.com/272360.html>>. Acesso em: 17 mar. 2002.
- [PET 93] PETRAKIS, Euripides G. M. **Image Representation, Indexing and Retrieval Based on Spatial Relationships and Properties of Objects**. 1993. Disponível em: <<http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/p/Petrakis:Euripides>>. Acesso em: 12 dez. 2001.
- [RAO 99] RAO, Aibing et al. Spatial Color Histograms for Content-based Image Retrieval. In: IEEE INTERNATIONAL CONFERENCE ON TOOLS WITH ARTIFICIAL INTELIGENCE, ICTAI, 11., 1999, Chicago, Illinois. **Proceedings...** [S.l.: s.n.], 1999. p. 183-186.
- [RUI 97] RUI, Yong et al. Image Retrieval: Past, Present, and Future. In: SYMPOSIUM ON MULTIMEDIA INFORMATION PROCESSING, 12., 1997, Taipei, Taiwan. **Proceedings...** [S.l.: s.n.], 1997.
- [RUI 99] RUI, Yong et al. Image Retrieval: Current Techniques, Promising Directions and Open Issues. **Journal of Visual Communication and Image Representation**, [S.l.], v. 10, p. 36-62, 1999.
- [SAL 83] SALTON, Gerard; MCGILL, Michael J. **Introduction to Modern Information Retrieval**. New York: McGraw-Hill, 1983. p. 157-191.
- [SCL 96] SCLAROFF, S. et al. Photobook: Content-based Manipulation of Image Databases. **International Journal of Computer Vision**, [S.l.], v. 18, n. 03, p. 233-254, 1996.
- [SCH 94] SCHARCANSKI, Jacob et al. Representing the color aspect of texture images. **Pattern Recognition Letters**, Amsterdam, v. 15, n. 02, p. 191-197, 1994.
- [SCH 2001] SCHETTINI, Raimondo et al. **A Survey of Methods for Colour Image Indexing and Retrieval in Image Databases**. 2001. Disponível em: <<http://www.itim.mi.cnr.it/Staff/Schettini/publicazioni.htm>>. Acesso em: 30 maio 2002.
- [SME 2000] SMEULDERS, Arnold W. M. et al. Content-Based Image Retrieval at the End of the Early Years. **IEEE Transactions on PAMI**, [S.l.], v. 22, n. 12, p. 1349-1380, 2000.
- [SMI 95] SMITH, John R.; CHANG, Shih-Fu. Automated Image Retrieval Using Color and Texture. Submetido ao IEEE Transactions on PAMI em 1995.

- [SMI 96] SMITH, John R.; CHANG, Shih-Fu. **Tools and Techniques for Color Image Retrieval**. 1996. Disponível em: <<http://www.ctr.columbia.edu/~jrsmith/html/pubs/tatfeir/color.html>>. Acesso em: 14 dez. 1999.
- [SMI 9?] SMITH, John R.; CHANG, Shih-Fu. Tools and Techniques for Color Image Retrieval. **IS&T/SPTE Proceedings**, [S.l.], v. 2670, p. 1-12, [199?].
- [SMI 2002] SMITH, John R. **Retrieval Effectiveness**. Disponível em: <<http://disney.ctr.columbia.edu/jrsthesis/node72.html>>. Acesso em: 05 jun. 2002.
- [SOB 98] SOBREIRO, Marcos Vinícius Rayol. **Quantização de Imagens**. Florianópolis: UFSC, 1998. p. 1-18.
- [SON 99] SONKA, Milan et al. **Image Processing, Analysis, and Machine Vision**. Pacific Grove: ITP, 1999. p. 1-55.
- [VEN 2001] VENTERS, Colin C.; COOPER, Matthew. **A Review of Content-Based Image Retrieval Systems**. [S.l.]: University of Manchester, [2001]. Disponível em: <<http://www.jtap.ac.uk/reports/htm/jtap-054.html>>. Acesso em: 26 jul. 2001.
- [WES 99] WESOLKOWSKI, Slawo; JERNIGAN, Ed. Color Edge Detection in RGB using Jointly Euclidean Distance and Vector Angle. In: VISION INTERFACE, 1999, Trois-Rivières. **Proceedings...** [S.l.: s.n.], 1999. p. 9-16.