

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

**Um Mecanismo de Realimentação de Relevâncias
para Recuperação de Informações Visuais
Utilizando Feições Extraídas a partir de Imagens JPEG**

IRAÇÚ OLIVEIRA SANTOS

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, dezembro de 2002

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Santos, Iraçú Oliveira

Um Mecanismo de Realimentação de Relevâncias para Recuperação de Informações Visuais Utilizando Feições a partir de Imagens JPEG / Iraçú Oliveira Santos. – Porto Alegre: PPGCC da UFRGS, 2002.

53f.: il.

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

1. Recuperação de informação visual. 2. Recuperação de imagens através de conteúdo. 3. Realimentação de relevâncias. I. Scharcanski, Jacob. II Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitora: Profa. Wrana Maria Panizzi

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

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

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

Coordenador do PPGC: Prof. Carlos Alberto Heuser

Bibliotecária – chefe do Instituto de Informática: Beatriz Regina Bastos Haro

Agradecimentos

“Para qualquer problema por maior que pareça, Deus reservou uma solução.
É preciso buscá-la.”

(Extraído do livro 'Sementes de Felicidade')

Meu agradecimento principal é dirigido ao professor Jacob Scharcanski, por sua orientação e extrema paciência e compreensão no andamento dos resultados finais deste trabalho.

Outros agradecimentos também são direcionados a todos aqueles amigos que no momento oportuno, principalmente quando o desânimo, devido às atribulações da vida pessoal, interferiu de alguma forma no andamento do Mestrado, ensinaram-me a ver a realidade de que Deus sempre nos reserva uma solução para todo e qualquer problema, bastando apenas que tenhamos fé, determinação e boa vontade para buscá-la.

Agradecimentos em especial...

... ao Prof. Dr. João dos Santos Protázio, professor do Departamento de Matemática da Universidade Federal do Pará, pelo seu auxílio na demonstração da expressão obtida para normalização dos valores de similaridade e ao Prof. Dr. Hermínio Simões Gomes, também professor do Departamento de Matemática da UFPA quanto ao seu auxílio na explicação de conceitos relacionados ao reescalonamento de espaços.

.... ao Prof. Dr. Francisco Pereira Assunção, Diretor do Centro de Ciências Exatas e Naturais da UFPA, por seu apoio durante o período de afastamento para realização do Mestrado, assim como na liberação dos recursos financeiros possibilitando o deslocamento até a cidade de Porto Alegre, quando isso se fez necessário.

... ao Prof. Dr. Antonio Benedito Coimbra Sampaio, idealizador do projeto MINTERCC – UFPA/UFRGS, através do qual foi possível a minha participação no Mestrado.

Finalmente, dirijo este agradecimento àqueles que mais merecem toda a gratidão do mundo, que são minha esposa Socorro, e aos meus filhos Luciana Carolina, Ana Luiza e Felipe Augusto, que estiveram um longo período longe de mim, durante os estágios obrigatórios em Porto Alegre, além dos momentos em que a distância se deu em função da dedicação direcionada mais para o trabalho de conclusão do Mestrado do que para a vida familiar.

Aos meu pais,
Iraçú e Mariana.
Os maiores responsáveis por tudo de bom que sou e aprendi na vida.

Sumário

Lista de Abreviaturas	6
Lista de Figuras	7
Lista de Tabelas	8
Resumo	9
Abstract	10
1 Introdução	11
2 Sistemas de Recuperação de Informações Visuais	14
2.1 Sistemas de Recuperação de Informação Visual Estática	14
2.2 Recuperação de Informação Baseada em Similaridade	15
2.3 Interação Baseada em Conteúdo	16
2.4 Consulta Através de Exemplo Visual e Realimentação de Relevâncias	18
3 Modelo de Recuperação de Informação Visual	20
3.1 Realimentação de Relevâncias para Recuperação de Informações Visuais	22
3.2 Estratégias para Realimentação de Informações	23
3.3 Realimentação de Informações Relevantes (Positivas)	24
3.4 Realimentação de Informações Não Relevantes (Negativas)	24
4 Representação de Imagens para Recuperação de Informações Visuais	25
5 Mecanismo de Realimentação de Relevâncias (MRR)	29
5.1 Elementos do MRR	29
5.2 Diagrama de Representação de Relacionamento entre Elementos do MRR	30
5.3 Descrição do MRR	31
5.3.1 Algoritmo do MRR	33
6 Resultados Experimentais	34
6.1 Recuperação de Imagens através das Cores	35
6.2 Recuperação de Imagens através das Texturas	37
6.3 Avaliação de Desempenho de Recuperação	40
7 Conclusões e Trabalhos Futuros	44
Anexo 1 Coleção 1 de Imagens – Recuperação de Imagens através das Cores	45
Anexo 2 Coleção 2 de Imagens – Recuperação de Imagens através das Texturas	47
Bibliografia	49

Lista de Abreviaturas

1-D	Uma Dimensão
2-D	Duas Dimensões
CBIR	Content-Based Image Retrieval
CBVIR	Content-Based Visual Information Retrieval
DCT	Discrete Cosine Transform (Transformada Discreta do Cosseno)
IR	Information Retrieval
JPEG	Joint Photographic Expert Group
MEV	Modelo de Espaço Vetorial
MRR	Mecanismo de Realimentação de Relevâncias
RGB	Red, Green e Blue (Vermelho, Verde e Azul)
RR	Realimentação de Relevâncias
VS	Valor de Similaridade
Z^+	Conjunto dos Números Interiores Positivos

Lista de Figuras

FIGURA 2.1 - Consulta Através de Exemplo.....	18
FIGURA 2.2 - Realimentação Positiva (Imagens Relevantes) e Realimentação Negativa (Imagens Irrelevantes).....	19
FIGURA 3.1 - Representação Conceitual do Modelo Vetorial de Recuperação de Informação Visual.....	21
FIGURA 4.1 - Bloco 8x8 de Coeficientes DCT.....	26
FIGURA 4.2 - Imagem DC.....	27
FIGURA 5.1 - Diagrama de Relacionamento entre Elementos do MRR.....	30
FIGURA 5.2 - Espaço de Feições em Relação à Imagem de Consulta.....	32
FIGURA 6.1 - Partição de uma Coleção de Imagens em Relação a uma Imagem de Consulta q.....	35
FIGURA 6.2 - Resultado de Consulta Através de Cor sem RR.....	36
FIGURA 6.3 - Resultado de Consulta Através de Cor com RR – 1a. Iteração...	36
FIGURA 6.4 - Resultado de Consulta Através de Cor com RR – 2a. Iteração...	37
FIGURA 6.5 - Resultado de Consulta Através de Textura sem RR.....	38
FIGURA 6.6 - Resultado de Consulta Através de Textura com RR – 1a. Iteração.....	38
FIGURA 6.7 - Resultado de Consulta Através de Textura com RR – 2a. Iteração.....	39
FIGURA 6.8 - Resultado de Consulta Através de Textura com RR – 3a. Iteração.....	39
FIGURA 6.9 - Gráfico Comparativo de Recuperação de Imagens Através do Conteúdo de Cor.....	41
FIGURA 6.10 - Gráfico Comparativo de Recuperação de Imagens Através do Conteúdo de Textura.....	42

Lista de Tabelas

TABELA 6.1 - Médias de Recall e Precision para Recuperação por Cor.....	40
TABELA 6.2 - Médias de Recall e Precision para Recuperação por Textura.....	41

Resumo

Esta dissertação propõe e discute um mecanismo de realimentação de relevâncias (i. e. “Relevance Feedback”). A técnica de realimentação de relevâncias foi introduzida inicialmente em meados dos anos 60, como uma estratégia para refinamento de consultas para a recuperação de informações. Como uma técnica de refinamento de consultas, foi aplicada inicialmente em sistemas de recuperação de informações textuais. Neste caso, os termos ou expressões consideradas importantes, são utilizados na formulação de uma nova consulta.

Ao surgirem os sistemas de recuperação de informação visual baseada em conteúdo (CBVIR), houve a necessidade de serem introduzidos novos elementos associados a esse processo de reformulação de consultas, de tal forma que fossem utilizados não apenas as informações de alto nível, como os termos e expressões. Esses novos elementos passaram a considerar também a subjetividade de percepção humana em relação ao conteúdo visual. Neste trabalho, apresenta-se um processo de extração e representação desse conteúdo, através da utilização de feições (conteúdo) de cor e textura, extraídos de imagens JPEG, uma vez que no processo de compressão de imagens nesse formato, utiliza-se coeficientes da Transformada Discreta do Cosseno (DCT), sendo, portanto esses coeficientes utilizados como elementos que possuem as informações associadas a cor e textura na imagem. Além da DCT é utilizada a Transformação Mandala [YSH 83] no processo de agrupamento de somente 10 coeficientes, com o objetivo de produzir 10 imagens com resolução menor que a imagem original, mas que representam cada uma, o conteúdo de uma frequência particular da imagem original. A escolha por uma representação como essa, é a garantia de uma redução significativa na quantidade de dados a serem processados. Entretanto, a representação obtida nesse formato para as imagens, é com base em conteúdo global de cor e textura, o que pode produzir resultados insatisfatórios.

A introdução de um mecanismo de realimentação de relevâncias, associado à representação utilizada, permite contornar a dificuldade apontada acima, através da obtenção de consultas subseqüentes, selecionando os objetos mais relevantes, assim como menos objetos não relevantes, utilizando o conhecimento do usuário de forma interativa no refinamento de consultas para recuperação de informações visuais.

Palavras-Chave: Recuperação de Informação Visual, Recuperação de Imagens Através de Conteúdo, Realimentação de Relevâncias.

TITLE: “A RELEVANCE FEEDBACK MECHANISM FOR VISUAL INFORMATION RETRIEVAL USING FEATURES EXTRACTED FROM JPEG IMAGES”

Abstract

This dissertation proposes and discusses a relevance feedback mechanism. The relevance feedback technique was introduced initially in middle of years 60, as a strategy of query refinement for information retrieval. As a query refinement technique, was applied initially in textual information retrieval systems. In this case, the terms or expressions considered important are used to formulate a new query.

When appearing the content-based visual information retrieval systems (CBVIR), it had the necessity to be introduced new elements associated to this process of query refinement, of such form that the information of high level were used not only, as the terms and expressions. These new elements had started to also consider the human subjectivity perception in relation to the visual content. In this work, a process of extraction and representation of this content is presented, through the use of features (content) of color and texture, extracted from JPEG images, since that in the process of compression of images in this format, it is used coefficients of the Discrete Cosine Transform (DCT), being therefore these coefficients the elements that possess the information associated to color and texture in the image. Beyond the DCT is used the Mandala Transformation [YSH 83] in the process of grouping of only 10 coefficients, with the objective to produce 10 images with lesser resolution than the original image, but that represent each one, the content of a particular frequency of the original image. The choice for a representation like this is the guarantee of a significant reduction in the amount of data to be processed. However, the representation gotten in this format for the images is based on global content of color and texture, what it can produce unsatisfactory results.

The introduction of a mechanism of relevance feedback, associated to the used representation, allows skirting the difficulty pointed above through the adoption of subsequent queries, selecting objects most relevant, as well as less not relevant objects, using the knowledge of the user of interactive form in the refinement of queries to retrieve visual information.

Keywords: Content-Based Image Retrieval, Visual Information Retrieval, Relevance Feedback.

1 Introdução

A aplicação da técnica de realimentação de relevâncias (RR) pode ser considerada como uma das estratégias mais utilizadas na reformulação de consultas. Em termos gerais o processo de realimentação de relevâncias ocorre a partir da apresentação de uma lista de documentos recuperados ao usuário, que após examiná-los, destaca aqueles que considere relevantes. Um dos sistemas pioneiros na recuperação de informações visuais oferecendo interatividade com o usuário, foi o sistema comercial QBIC (Query by Image Content) [FLI 95], que oferece recuperação baseada em conteúdo utilizando imagens estáticas, assim como recuperação baseada em conteúdo e visualização para vídeos.

Pode-se destacar três aspectos fundamentais relacionados ao processo de RR, que determinam a vantagem da sua aplicação em relação a outras estratégias para refinamento de consultas:

- 1) O usuário não necessita conhecer os detalhes do processo de reformulação de uma consulta, pois a única coisa que é exigida do mesmo é a escolha dos objetos que considere relevantes ou não;
- 2) Permite que todo o processo de busca seja dividido em várias seqüências de busca menores, que facilitam o processo de recuperação de informação;
- 3) Todo o processo de realimentação é feito de forma controlada, onde é possível se enfatizar os itens que sejam relevantes, assim como desenfatar aqueles que sejam não relevantes.

A recuperação de informações ("Information Retrieval" - IR) requer um modelo de objeto (documentos), um modelo de consulta e um modelo para calcular a similaridade entre objetos e consultas. No capítulo 2 serão abordados os modelos utilizados neste trabalho, no processo de recuperação de informações visuais. O modelo vetorial tem sido um dos mais populares [SAL 87], para o qual várias técnicas de Realimentação de Relevância foram desenvolvidas para a recuperação de informações visuais.

Os sistemas de recuperação de informações visuais têm como paradigma tradicional a utilização de palavras chaves na identificação de imagens. Entretanto, algumas dificuldades estão relacionadas ao uso deste paradigma, entre elas estão: a) a quantidade de informação necessária no processo de recuperação, registrada de forma manual; b) a interpretação e descrição do conteúdo de uma imagem, que pode ser feita de diferentes maneiras, dependendo do usuário que fizer a análise da imagem; c) a inconsistência que pode surgir em relação ao uso de palavras-chave, considerando-se diferentes formas de indexação. Essas dificuldades podem vir a se tornar cada vez maiores na medida em que aumentar o número de imagens armazenadas.

Nos últimos anos os avanços tecnológicos na área da informática e da Internet, motivaram em grande parte o desenvolvimento de novas técnicas para facilitar a recuperação de informações visuais e gerenciar o conteúdo de imagens, armazenadas em grandes bancos de dados contendo milhares de objetos. Com isso, surgiu a partir de 1990 uma nova proposta para recuperação de informações visuais, onde se passou a considerar não apenas palavras-chave, mas também o conteúdo visual das imagens, através de feições de cor, textura e forma. Cada um desses sistemas permite a Recuperação de Imagens Baseada em Conteúdo ("Content-Based Image Retrieval" - CBIR), utilizando informações extraídas a partir das próprias imagens e auxiliando o usuário na composição de consultas [DIS 99]. As pesquisas existentes nesta área deram origem a alguns sistemas, alguns deles com características acadêmicas, e outros já com aplicação em nível comercial. Em ambos os casos pode-se identificar aqueles baseados em conteúdo ou texto, ou mesmo aqueles que apresentem os dois princípios inseridos na sua concepção. Entre os mais conhecidos destacamos:

- QBIC (Query by Image Content) – propriedade da IBM Corporation. (<http://www.qbic.almaden.ibm.com/tmdemo/>).
- Virage Search Engine – propriedade da Virage Corporation.
- Visual Retrievalware – propriedade da Excalibur Technologies Corporation. (<http://www.excalib.com/products/vrw/vrweustomers.html>)
- VisualSEEK – desenvolvido na Universidade de Colúmbia por Smith e Chang [SMI 97]. <http://www.ctr.columbia.edu/jrsmith/VisualSEEK/VisualSEEK.html>
- MARS (Multimedia Analysis and Retrieval System) – desenvolvido na Universidade de Illinois por Thomas S. Huang e Young Rui [RUI 97]. (<http://www.db.ies.uci.edu:8004/marsApplet.html>)

Em cada um desses sistemas é permitido ao usuário durante um processo de interação, formular exemplos que permitam uma comparação de similaridade entre imagens. Entretanto, vários problemas podem dificultar essa interação, relacionados com questões como o tamanho do banco de dados de imagens, a heterogeneidade do conteúdo das imagens que o mesmo possa apresentar, ou ainda com a complexidade da tarefa de recuperação, o que pode tornar difícil para o usuário a busca ou mesmo encontrar uma imagem que seja similar a da consulta logo na primeira iteração de recuperação de informação. Outra dificuldade pode estar relacionada à definição da informação que se necessita, formulada de uma maneira vaga, como por exemplo: “encontrar todas as imagens contendo paisagens”. Em qualquer um dos casos citados, o julgamento de uma imagem como sendo relevante pode ser muito diferente. O tratamento das imagens realizado de uma mesma maneira, feito através da análise de feições de imagens relevantes a fim de se formular uma nova consulta, ou atualizar medidas de similaridade, pode produzir resultados que não sejam satisfatórios. Associado a isso, imagens relevantes podem apresentar algumas feições como cor, que sejam casualmente similares. Com isso, o sistema deve possuir recursos que possibilitem diferenciar estas feições e com isso evitar iterações durante o processo de recuperação de informação que sejam equivocadas.

Com base no que foi exposto acima, pode-se afirmar que as técnicas de refinamento de consultas na recuperação de informações visuais, muitas vezes dependem de outros campos de pesquisa, tais como: Recuperação de Informações, Modelagem e Representação de Dados Visuais, Processamento e Análise de Imagem/Vídeo, Reconhecimento de Padrões e Visão Computacional, Organização de Banco de Dados Multimídia, Indexação Multidimensional, Modelamento Psicológico do Comportamento do Usuário e Visualização de Dados, que juntos ou separadamente podem contribuir para a eficiência da recuperação de informações visuais.

Este trabalho apresenta uma proposta de mecanismo de realimentação de relevâncias para recuperação de informações visuais através da utilização dos conteúdos de cor e textura, e cujo objetivo principal é refinar o resultado de uma consulta através de exemplos. No Capítulo 2 são abordados conceitos iniciais, que auxiliam na compreensão da área de recuperação de informações visuais. O Capítulo 3 aborda conceitos relacionados ao modelo de recuperação de informações visuais, procurando posicionar a técnica de realimentação de relevâncias dentro desta área de estudo.

No Capítulo 4 é apresentado o método adotado no processo de extração das informações de cor e textura de imagens estáticas 2-D. Essas informações são utilizadas juntamente com medidas para a determinação de similaridade na busca e recuperação de informações visuais de imagens no formato JPEG. São utilizadas com esse objetivo duas coleções de imagens. A primeira contendo 150 imagens coloridas de flores ornamentais, e a segunda 195 imagens de texturas coloridas.

O Capítulo 5 corresponde à apresentação e descrição do mecanismo de realimentação de relevâncias proposto. Em seguida, no capítulo 6 são apresentados os resultados, assim

como a análise de desempenho dos resultados obtidos, utilizando-se critérios de recuperação (i. e. “recall”) e precisão (i. e. “precision”) [BAE 99].

Finalmente, o Capítulo 7 apresenta algumas considerações finais relacionadas ao método utilizado para representação e recuperação das imagens, assim como das conclusões em relação ao mecanismo de realimentação de relevâncias proposto, levando em consideração a sua possível aplicação em outros métodos de representação e recuperação de imagens ou melhoramento em trabalhos futuros.

2 Sistemas de Recuperação de Informações Visuais

Na recuperação de informações visuais consideramos inicialmente dois tipos básicos de informações [DEL 97]:

- **Metadados independentes de conteúdo** – São informações da imagem que não estão diretamente associadas ao seu conteúdo. O metadado é alfanumérico e geralmente a este tipo de informação estão relacionados o formato, nome do autor, data, localização, proprietário, etc.
- **Feições visuais ou informações relacionadas ao conteúdo visual** – As feições visuais são obtidas através de processos computacionais - tipicamente processamento de imagem, visão computacional, e rotinas de geometria computacional; – executadas sobre o objeto visual. Neste caso dois aspectos podem ser diferenciados:
 - Feições de baixo nível, associadas ao conteúdo de cor, textura, forma, relações espaciais, movimento e relações entre esses conteúdos. Estas feições são denominadas também de *metadados dependentes do conteúdo*. Para imagens estáticas, pode-se relacionar feições associadas aos conteúdos de cor, forma e textura, enquanto que para vídeo, além destas três feições associa-se o movimento.
- Semântica do conteúdo, chamado também de *metadado descritivo do conteúdo*. Associam-se elementos contidos no objeto visual com elementos do mundo real, eventos relacionados ao tempo, emoções e significados associados a símbolos visuais e cenas.

2.1 Sistemas de Recuperação de Informações Visuais Estáticas

Na recuperação de imagens estáticas 2-D, consideram-se propriedades perceptuais tais como cor, textura, forma e relações espaciais entre objetos na imagem. Além dessas propriedades pode-se considerar também abstrações relacionadas a objetos e cenas presentes na imagem, assim como impressões, emoções e significados associados à combinação de feições que sejam percebidas.

A recuperação de imagens através de propriedades de percepção, utiliza um paradigma de recuperação, segundo o qual pode-se obter previamente para cada imagem, um conjunto de feições distintas, chamadas de parâmetros de modelo de representação [ARIS 2000]. São através destes parâmetros que serão formuladas consultas *através de exemplos visuais*, fornecidos pelo usuário, ou extraídos de amostras de imagens. A consulta formulada pelo usuário é construída em termos das feições e uma faixa de parâmetros que sejam julgados como *relevantes*, e associado a isso deve ser feito a escolha de uma medida de similaridade (similaridade baseada em recuperação). A partir dessa consulta o sistema realiza uma primeira verificação de similaridade entre o conteúdo visual expresso na consulta e cada item armazenado no banco de imagens.

Uma vez obtidos através da consulta, os resultados podem não ser completamente satisfatórios, com isso pode-se adotar algum mecanismo de refinamento que possibilite a obtenção de um número maior de informações relevantes. Neste trabalho é apresentada a técnica de Realimentação de Relevâncias, que através de um mecanismo de RR, permite a interação do usuário com o sistema, possibilitando o refinamento das consultas subsequentes e dessa forma a obtenção de informações mais relevantes.

2.2 Recuperação de Informações Através de Similaridade

A recuperação de objetos visuais baseada em similaridade consiste basicamente em um processo contínuo de reorganização do banco de dados de imagens, feita através da utilização do valor de similaridade assumido por cada um dos elementos do banco de dados, medido em relação à consulta exemplo. Com isso, a cada processo de reorganização é produzido um conjunto de imagens, posicionadas dentro de um determinado grupo de imagens com conteúdos visuais similares, mesmo que não haja imagens próximas à consulta exemplo, ou seja, mesmo que não se considere uma imagem específica do banco de dados. Nesse caso, é fornecido um critério de similaridade que deve servir de parâmetro na organização do grupo de imagens dentro do banco de dados. Uma das conseqüências desse processo é o fato de que a resposta para qualquer consulta é todo o banco de dados. Por esse motivo, deve-se associar ao processo um mecanismo que organize as imagens em uma ordem de acordo com os valores de similaridade, exibindo somente um número limitado de imagens que estejam mais próximas do objetivo especificado. No trabalho desenvolvido, apresentam-se resultados diferenciados a partir da exibição de um número de imagens que varia de 10 a 55 imagens. O papel do usuário neste processo é o de sempre analisar as respostas do sistema, definindo os novos critérios de relevância (relevante ou não relevante) para cada imagem, permitindo assim a construção de uma nova *consulta através de associação* cuja característica principal é possibilitar o refinamento do processo de recuperação de informação visual. É esta característica que posiciona a técnica de RR como sendo de grande importância nos sistemas de recuperação de informações visuais.

Vários modelos de similaridade já foram propostos, tornando com isso, difícil a aplicação de um único modelo que reproduza a forma como o homem acessa a similaridade entre objetos visuais diferentes. Devido a isso, o usuário tende a utilizar uma combinação de diferentes medidas de similaridade, associada a um determinado domínio. Sem entrarmos em detalhes quanto ao conteúdo teórico, destacamos a seguir quatro modelos de similaridade descritos em [DEL 99]:

- **Modelo Métrico**
- **Modelo de Espaço Métrico Virtual**
- **Modelo de Tversky**
- **Distâncias Transformacionais**

Ao se trabalhar com similaridade, uma das conseqüências é a possibilidade de se criar consultas cada vez mais complexas. Uma das formas de tornar essa tarefa mais simples é a utilização de sucessivos refinamentos. Parte-se assim, de uma consulta inicial associada a um determinado critério de similaridade, que no trabalho apresentado, relaciona cores e texturas. O resultado dessa consulta é constituído, como descrito anteriormente, por uma porção do banco de dados onde os objetos visuais apresentam maior similaridade em relação à primeira consulta. Em seguida essa porção de imagens é utilizada na composição de uma nova consulta de acordo com um segundo critério de similaridade, e assim por diante. O critério de similaridade é tipicamente definido utilizando-se a distância entre pontos das imagens em um espaço de feições multidimensional, dado por alguma métrica, onde imagens próximas a um determinado exemplo de consulta são consideradas “similares” a essa consulta. O objetivo principal dos sistemas que utilizam este conceito, é o de retornar imagens que sejam similares à consulta de acordo com a percepção e julgamento de relevância por parte do usuário.

A primeira geração de sistemas que utilizavam o conceito de similaridade na recuperação de imagens não utilizava as técnicas hoje disponíveis em áreas como visão computacional e processamento de imagens. Esses sistemas utilizavam anotações manuais

(informações), associadas à imagem e definidas através da interação com o usuário. As informações eram utilizadas em técnicas de consulta textual que utilizavam o conceito de similaridade para recuperação de imagens relevantes, através da associação existente entre a consulta do usuário e essas anotações [RIC 94], [SRI 94], [SME 96] e [PON 98].

Várias dificuldades foram surgindo com a utilização de sistemas baseados em textos, e uma das maiores encontradas, foi a grande quantidade de anotações exigidas quando se trabalhava com um grande volume de dados de imagem ou vídeo, o que provocou o surgimento dos sistemas de recuperação baseados em conteúdo, que basicamente utilizam a noção de que uma imagem pode ser recuperada a partir de seu conteúdo visual. Os primeiros trabalhos propostos já utilizavam representações para as imagens através de feições visuais, extraídas após a aplicação de algoritmos de processamento de imagens [CAR 97], [CHA 97], [MES 99], [FLI 95] e [PEN 96]. A partir desses e de outros sistemas que surgiram, várias medidas de similaridade foram sendo propostas, muitas delas de acordo com o tipo de problema ou abordagem utilizada no processo de recuperação. O estado da arte quanto a essas abordagens encontra-se atualmente ligado a temas como: Informação Contextual para Recuperação de Informações Visuais [GRE 2001], Técnicas de Aprendizado para Recuperação de Imagens Baseada em Conteúdo [FOU 2002] e [ZHG 2002], Lógica Fuzzy para Recuperação de Informações Visuais [KUS 2002] e [NIK 2002], Modificação Automática de Consultas em Sistemas de Recuperação de Imagens [AGG 2002], Recuperação de Imagens Baseada em Semântica [DEL 99] e [HAN 2002], Utilização de fractal para extração de feições em imagens [NIK 2002].

Finalmente, acrescentamos que quanto ao desempenho dos sistemas que utilizam uma ou outra abordagem na recuperação de informações visuais, o que existe é uma dependência em relação à medida de similaridade utilizada para determinar o grau de similaridade ou dissimilaridade entre duas imagens [DOU 2002], pois essas medidas podem variar de acordo com a subjetividade de percepção humana, conteúdo visual ou tipo de indexação (indexação através de cor, textura ou forma).

2.3 Interação Baseada em Conteúdo

O processo de busca de informações pelo usuário associado à visualização permite uma visão geral do *espaço de informação visual* (descrito através de um modelo de recuperação de informações), no qual as imagens são representadas, sendo esse definido na unidade 3, através do modelo de espaço vetorial (MEV). Além disso, pode-se executar a busca de informações visuais de forma localizada, em um ou mais pontos de interesse específicos, associados a objetos visuais. Dentre esses pontos de interesse, destacam-se todas as características que sejam comuns a um determinado tipo de imagem e que possam ser rapidamente identificadas, como por exemplo, características de textura (forma, direção e organização espacial). O resultado de consultas pode ser fornecido através de ferramentas de visualização, que podem exibir as imagens em dimensões consideravelmente reduzidas em relação a imagem original (thumbnails), posicionadas em uma ordem de similaridade, ou ainda, através da exibição de todo o espaço de visualização. Além disso, as características visuais observadas nas imagens, pelo usuário, podem estar disponíveis para seleção em uma interface contendo opções de escolha para texturas, cores ou formas, com o objetivo de formular consultas de busca que levem em consideração características associadas a essas feições. O papel do sistema é o de identificar todas as imagens que possuam um alto grau de similaridade em relação às características (feições) solicitadas. Assim sendo, podemos considerar a interação baseada

em conteúdo como um processo de verificação de similaridade em relação a feições associadas à cor, textura e forma, podendo ser dividida em três tipos básicos [BAE 97]:

- **Consultas Exatas** – Tem como objetivo localizar cópias exatas ao da imagem exemplo, descrita através de uma consulta. Neste caso, o grau de similaridade entre duas imagens é máximo.
- **Consultas de classificação** - São consultas que definem uma ordem de classificação para cada tipo de feição, posicionando dentro de uma determinada classe todas as imagens que se encontrem dentro de um determinado espaço de feições.
- **Consultas de vizinhança k** - Neste tipo de consulta o valor numérico associado a k representa o número de itens que devem ser recuperados, ou exibidos ao usuário como resultado da consulta. Neste trabalho atribui-se valor para k variando de 10 a 55 nos experimentos realizados. O resultado da consulta é obtido através da classificação de valores em ordem decrescente de similaridade, de acordo com uma determinada métrica de similaridade.

Com o objetivo de otimizar o tempo de localização de uma imagem, após a execução de uma determinada consulta, assim como evitar processos subjetivos de rotulação de imagens de forma manual, a maioria dos bancos de imagens utiliza atualmente técnicas de recuperação de imagens baseada em conteúdo, através da indexação das imagens utilizando feições de baixo nível, tais como cor, textura e forma [ZHG 2002]. Entretanto, de acordo com alguns resultados obtidos a partir de estudos recentes em [ROD 2000] e [HAN 2002], pode-se verificar também a necessidade de utilização de sistemas que permitam a formulação de consultas baseadas em conceitos de alto nível, como por exemplo: “Localizar flores vermelhas entre várias amarelas e violetas”, e não apenas feições de baixo nível (40% de vermelho, 10% violetas e 50% de amarelo). Com isso, passou-se a buscar outras formas de representação das informações visuais, que partissem de feições visuais de baixo nível para uma representação através de conceitos de alto nível, de forma dinâmica e interativa com o usuário [RUI 99]. É com base nisso também, que a técnica de realimentação de relevância pode ser introduzida, permitindo que haja uma interação entre as feições visuais e as medidas de similaridade, de tal forma que os conceitos de alto nível possam ser mais bem especificados, a partir dos exemplos visuais fornecidos pelo usuário [WAN 2002].

2.4 Consulta Através de Exemplo Visual e Realimentação de Relevâncias

A recuperação de objetos visuais (imagem/vídeo) utiliza propriedades de percepção, baseadas em um paradigma de recuperação. Para tal, o qual deve-se obter previamente para cada imagem um conjunto de feições distintas, chamadas de parâmetros de modelo, que auxiliarão na elaboração das consultas visuais através de exemplos (Figura 2.1). A técnica de realimentação de relevâncias possui um papel importante neste processo, na medida em que auxilia o usuário na tarefa de análise das respostas fornecidas pelo sistema e na indicação, para cada imagem recuperada, do grau de relevância ou de irrelevância. Após o registro desses resultados, o usuário deve realimentar o sistema, utilizando um processo de realimentação positiva (imagens relevantes) ou realimentação negativa (imagens irrelevantes) (Figura 2.2) a fim de que a consulta seja refinada e dessa forma, produza resultados mais próximos do esperado.

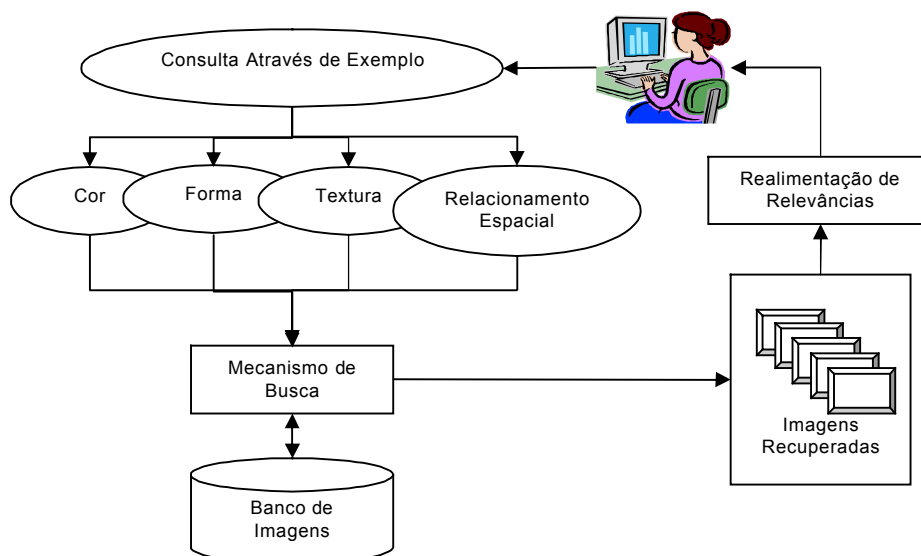


FIGURA 2.1 – Consulta de Imagens Através de Exemplo

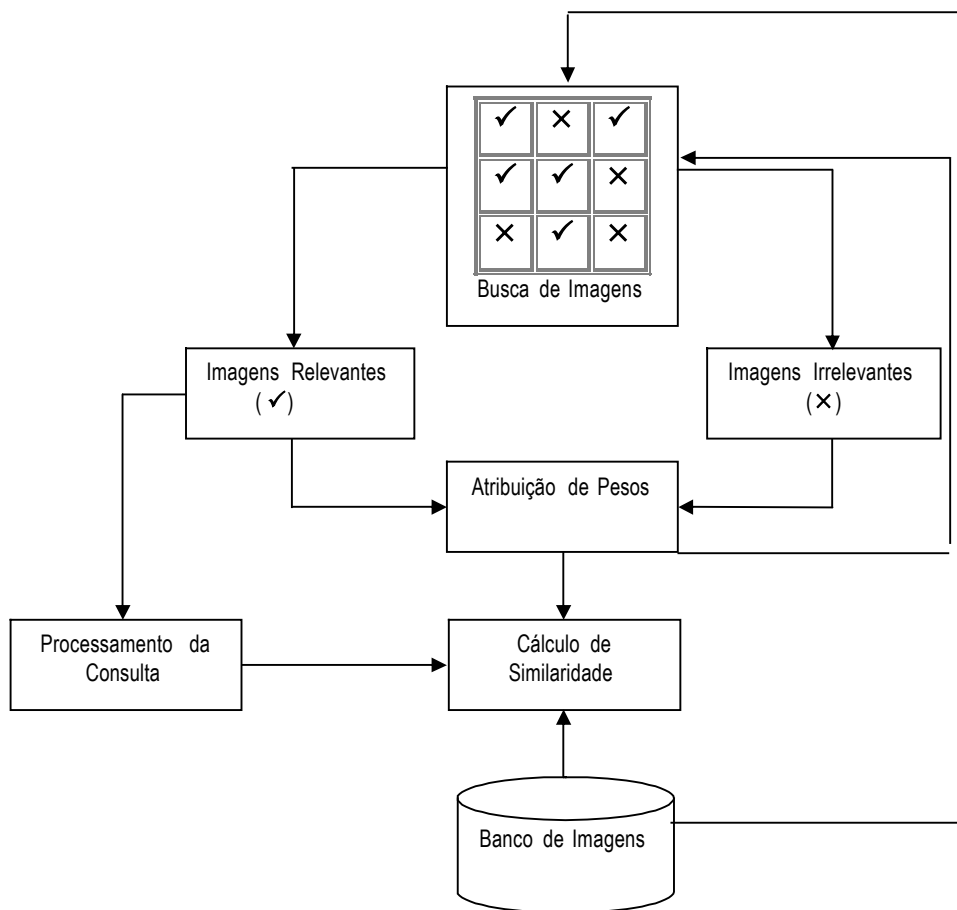


FIGURA 2.2 – Realimentação Positiva (Imagens Relevantes) e Realimentação Negativa (Imagens Irrelevantes) [CIO 2001]

3 Modelo de Recuperação de Informação Visual

Um modelo de recuperação de informações consiste de um modelo de objeto, um modelo de consulta e um modelo que calcule a similaridade entre objetos e consultas, de tal forma que a especificação de cada um desses elementos definirá o modelo de recuperação. Entre os principais modelos desenvolvidos para recuperação de informações destacam-se: o Modelo Booleano, Modelo Estatístico (que pode incluir o Modelo Vetorial e Probabilístico), Modelo Baseado em Conhecimento (que inclui modelos baseados em Lógica Fuzzy e Redes Neurais) e o Modelo Baseado em Linguística (Modelo de Indexação Semântica) [BAE 99]. O modelo de recuperação utilizado neste trabalho é baseado no Modelo de Espaço Vetorial (MEV).

O MEV é constituído de um método de indexação que gera vetores de descrição de objetos (imagens) e uma função de recuperação. Os vetores representam as feições de cor (vetor de feição de cor) ou de textura (vetor de feição de textura), e a função de recuperação mede a similaridade entre imagens através da comparação de vetores de feição do mesmo tipo, utilizando-se pesos associados a cada uma das feições (Figura 3.1).

Consideremos um banco de imagens contendo um conjunto D de imagens i a serem recuperadas. Esse conjunto de imagens armazenadas necessita de um processo de indexação que permita um processo de busca não apenas seqüencial, mas também a filtragem seletiva somente de imagens relevantes, através da verificação do conteúdo das imagens em relação à consulta formulada pelo usuário. Com isso será possível a ordenação das imagens recuperadas, baseada na similaridade em relação à consulta. A indexação da consulta e das imagens $i_j \in D$ é obtida através de um relacionamento de feições indexadas, e da atribuição de pesos a essas feições, com o objetivo de se obter a descrição da consulta e das imagens. Com isso obtêm-se dois conjuntos de feições indexadas através de pesos, que representam as descrições da consulta e das imagens. Os elementos desses conjuntos são representados através de vetores, denominados respectivamente de vetores de descrição das imagens e da consulta, expressos como:

$$\begin{aligned} \mathbf{i}_j &= (a_{0j}, a_{1j}, \dots, a_{ij}, \dots, a_{m-1,j}) \\ \mathbf{q} &= (b_0, b_1, \dots, b_i, \dots, b_{m-1}) \end{aligned} \quad (3.1)$$

onde a_{ij} e b_i são os pesos atribuídos a cada feição usada na indexação.

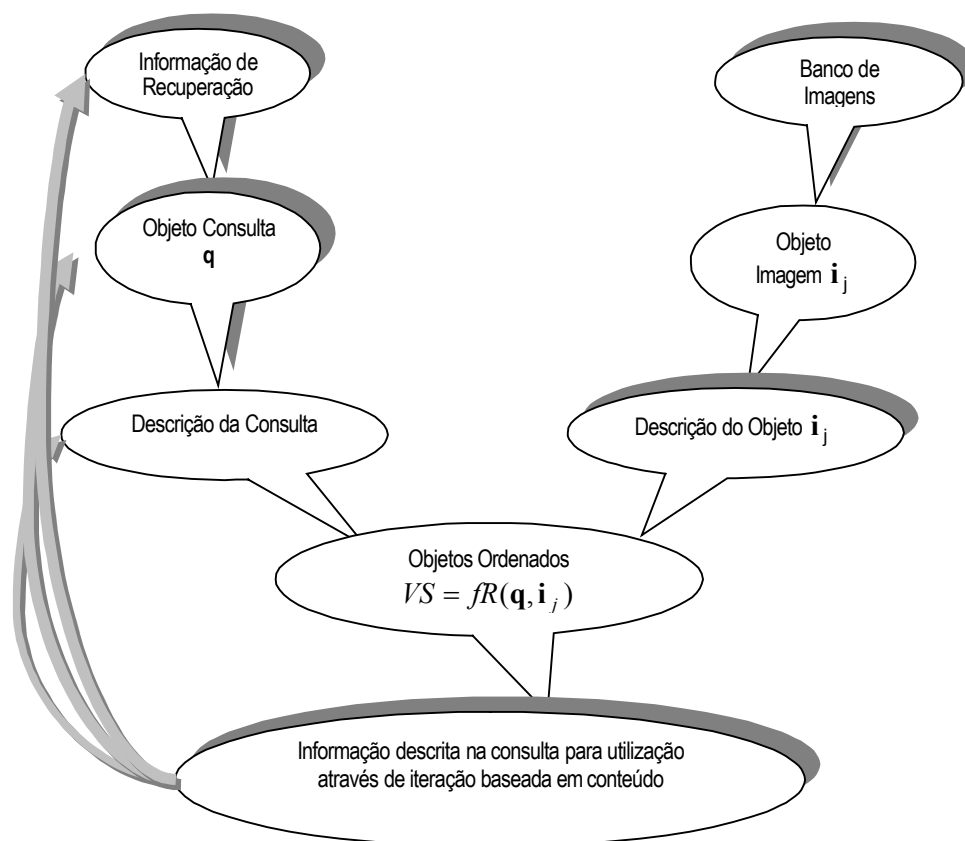


FIGURA 3.1 – Representação Conceitual do Modelo Vetorial de Recuperação de Informação Visual

Quando uma determinada feição indexada não é considerada na consulta ou na imagem, é atribuído o valor zero aos pesos correspondentes. Estas duas descrições são comparadas e em seguida é gerado um número associado ao resultado obtido, denominado de valor de similaridade (VS). A cada consulta que o usuário apresenta ao sistema é retornado um grupo de imagens que são exibidos em ordem decrescente de seus VS. Após isto, o usuário analisa somente um subconjunto de D , a partir do início da relação exibida, na busca de imagens que considere relevantes, onde em geral este é um número k de imagens (no nosso caso, k é escolhido entre 10 e 55 imagens, vizinhas ao exemplo apresentado). A realimentação do usuário a partir destas informações é que possibilitará a reformulação do vetor de descrição da consulta. Esta técnica é flexível, e possibilita a recuperação de imagens de acordo com o VS, mesmo que não contenham feições indexadas relacionadas à consulta original [ARIS 2000].

3.1 Realimentação de Relevâncias para Recuperação de Informações Visuais

A realimentação de relevâncias é uma técnica tradicional, onde o usuário e o sistema interagem com o objetivo de melhorar a qualidade da recuperação em ordem de relevância, através de um sistema de recuperação de informações [SAL 90].

A relevância pode ser considerada como uma medida da proximidade entre um objeto e um exemplo num processo interativo (Saracevic 1975). Podemos considerar relevância como sendo uma variável aleatória que assume o valor 1, quando um objeto for relevante e 0, quando um objeto não for relevante. Podemos ainda dividir o conceito de relevância em dois tipos: Objetiva e Subjetiva. A Relevância Objetiva modela o processo de decisão de um grupo de peritos que fazem uma consulta apresentando um exemplo. O resultado depende da consulta e do objeto exemplo, mas não depende da informação original solicitada pela pessoa que formulou a consulta, uma vez que essa informação pode não estar de acordo com o exemplo apresentado. A Relevância Subjetiva modela o processo de decisão de uma pessoa que formula uma consulta, ou seja, o que deve ser escolhido como relevante ou não relevante. O resultado depende do objeto fornecido como exemplo e da informação solicitada pela pessoa, como por exemplo, a recuperação de imagens com um determinado padrão de textura. Entretanto, o resultado não pode depender da consulta, que neste caso pode ter uma formulação que não seja adequada para a recuperação de imagens a partir de uma determinada informação solicitada. Como exemplo podemos citar o caso de uma consulta que utilize informações de textura de imagens, mas que, entretanto possuem somente padrões de cores associadas a formas, sem texturas propriamente definidas.

A utilização de RR em sistemas de recuperação de informações se justifica pelo fato de melhorar a qualidade da recuperação de informações, pois a interação com o usuário permite que os mesmos se tornem mais eficientes, além disso, ajuda a definir consultas melhores através de sucessivos refinamentos, utilizando a subjetividade de percepção do usuário.

Os algoritmos baseados na técnica de RR podem ser classificados basicamente em dois tipos: Geométricos e Probabilísticos [AKS 2000].

1 - Geométricos – a métrica de similaridade é baseada em critérios geométricos, onde uma medida de distância é utilizada na comparação entre imagens;

2 - Probabilísticos – utilizam funções de distribuição de probabilidade, que são definidas sobre o espaço de feições de imagens e aplicadas sobre regiões em torno da imagem, indicando assim, aquelas com uma maior probabilidade de serem relevantes.

Além disto, estudos recentes na área de busca baseada em conteúdo, tentam impor uma visão mais estruturada dessa área que auxilie na aplicação da técnica de RR. Entre esses estudos destacam-se [NAS 99]:

- **Busca Objetiva** – relaciona-se à tarefa do usuário em buscar uma imagem específica em um banco de dados.
- **Busca através de Categoria** – consiste na busca de uma ou mais imagens pertencentes a uma determinada categoria. Neste caso a subjetividade de semântica associada a uma categoria pode provocar uma grande complexidade à tarefa de busca.
- **Navegação open-ended** - neste caso a tarefa de navegação pode ser considerada como útil na busca de informações, assim como bastante complexa na medida em

que envolve vários aspectos relacionados à gerência de informações visuais, mais especificamente ao problema de recuperação.

A tarefa do usuário no processo de realimentação de dados visuais para o sistema de recuperação de informações, depende em parte da sua habilidade em reconhecer de maneira relativamente rápida, imagens que tenham particularidades iguais à de uma determinada consulta. Essa habilidade permite que o usuário realmente o sistema (feedback) com imagens retornadas a partir de uma consulta, indicando as **informações relevantes** que serão utilizadas no refinamento de consultas subseqüentes. Procedimentos como este foram implementados inicialmente no projeto VisualSEEK [SMI 96], onde o usuário poderia requisitar mais de uma informação de imagens, obtidas a partir de um único resultado da consulta inicial.

Outra forma que o usuário possui para obter informações é marcar as imagens consideradas como sendo exemplos **positivos (relevantes)** ou **negativos (não relevantes)**, baseado na natureza da imagem exemplo, realimentando assim o sistema, de tal forma que melhore a recuperação de informações. Este procedimento realizado de forma **interativa** com o sistema permite que o mesmo “aprenda” a partir de exemplos do usuário. Além disso, produz um refinamento de forma **iterativa** que pode fornecer como resultado um conjunto de informações a respeito de um objeto visual ou cena, que podem ser armazenadas para utilização posterior. As informações armazenadas podem ser utilizadas posteriormente em cálculos como o de similaridade entre duas ou mais imagens, uma vez que a percepção humana de similaridades é bastante subjetiva e dependente de outras tarefas [MUL 2000]. Devido a isso, a realimentação do usuário em sistemas de recuperação de informações visual é tida como um procedimento essencial, na medida em que, associando a uma consulta, feições que sejam relevantes e/ou não relevantes, pode-se produzir outras consultas que representem melhor os requisitos das informações solicitadas pelo usuário.

3.2 Estratégias para Realimentação de Informações

A utilização de realimentação do usuário em recuperação de informações visuais é bastante recente, comparado à recuperação de informações textuais, que possui trabalhos editados desde a década de 60, quando a realimentação com **informações relevantes positivas** (relevantes) já era utilizada na recuperação de informação textual. Nesta mesma década, outros trabalhos mostraram que a realimentação com **informações relevantes negativas** (não relevantes), também poderia ser utilizada em sistemas de recuperação de informações visuais. As estratégias que utilizam ambos os componentes positivo e negativo de forma ponderada, através de “pesos”, são mais usadas [ROC 71].

Existem dois tipos básicos de estratégias de realimentação utilizadas na recuperação de informações visuais [RUI 98], [MUL 2000]:

- Utilizar consultas separadas para cada imagem realimentada e em seguida juntar os resultados dessas consultas;
- Criar uma pseudo-imagem a partir da associação de várias outras, obtendo-se com isso uma consulta através de **busca associativa**. Essa associação de imagens pode incluir imagens relevantes ou não relevantes e é baseada na similaridade existente entre as imagens. Neste trabalho, apresenta-se uma abordagem que utiliza este tipo de estratégia.

Na sua forma geral, a busca associativa pode ser descrita da seguinte forma [ALB 98]:

A estrutura de dados de um objeto é uma $(k+1)$ -tupla, $(r_1, r_2, \dots, r_k, I)$, onde r_k representa uma componente e I é uma informação adicional que sofrerá uma alteração em seu valor inicial. Se considerarmos as imagens como sendo os objetos em questão, então a $(k+1)$ -tupla será dada por $(f_1, f_2, \dots, f_k, VS)$, onde f_k representa uma feição visual e VS o valor de similaridade que se altera a cada processo de busca. Dessa forma, o problema de busca associativa consistirá em se localizar e separar todas as imagens que satisfaçam a condição $n-1 < VS < n$, onde $n \in \mathbb{Z}^+$ e $[(n-1), n]$ representa um intervalo de valores de similaridade normalizados.

Como a utilização de RR em recuperação de informações visuais baseadas em conteúdo ainda é recente, poucas estratégias de realimentação tem sido estudadas, incluindo as que estejam associadas à realimentação utilizando informações não relevantes (realimentação negativa). Além disso, ressaltamos também a necessidade de trabalhos que considerem medidas de desempenho e validação, utilizando aspectos qualitativos e quantitativos daquilo que é recuperado [MUL 2000].

3.3 Realimentação de Informações Relevantes (Positivas)

Neste caso as imagens que retornam possuem os maiores valores de relevância e são exibidas em uma ordem decrescente de relevância, devendo todas elas apresentar várias feições atributos em comum. As imagens que sejam consideradas não relevantes devem ser novamente ordenadas no passo seguinte. Sendo assim, observada toda a etapa de realimentação, devem-se selecionar como relevantes todas as imagens obtidas a partir da primeira consulta e que sejam consideradas relevantes pelo usuário. Pode-se considerar ainda como um bom número de imagens a serem exibidas para escolha e realimentação, por exemplo, as 20 ou 50 primeiras por ordem de VS , o que pode ser justificado pela limitação de exibição do número de imagens na tela. Um estudo comparativo quanto ao número de imagens escolhidas, com respectivas porcentagens de melhoramento de performance utilizando realimentação do usuário é apresentado em [MUL 2000].

3.4 Realimentação de Informações Não Relevantes (Negativas)

Apesar de informações não-relevantes também refinarem consideravelmente o resultado das consultas, de acordo com o que foi comentado na seção anterior, deve-se ter o cuidado em utilizar um número de imagens correto como realimentação não relevante, a fim de não provocar a perda de feições que sejam consideradas importantes. Sendo assim, em função da limitada capacidade humana em organizar várias informações visuais de uma só vez, uma boa estratégia seria escolher, por exemplo, entre as 20 ou 50 primeiras imagens para realimentação relevante, e para realimentação não relevante as duas primeiras e as duas últimas não relevantes dentro do grupo de imagens, a partir da reposta da consulta. Esta estratégia combinada, pode contribuir para melhorar de forma automática a consulta, através da eliminação de imagens consideradas como realimentação não relevante. Uma proposta de estratégia combinada é apresentada neste trabalho no Capítulo 5.

4 Representação de Imagens para Recuperação de Informações Visuais

Neste trabalho as imagens são representadas através de feições de cor e textura extraídas diretamente no domínio da Transformada Discreta do Cosseno (DCT), cuja formulação matemática é apresentada a seguir [GON 93] :

A DCT direta 1-D é definida como:

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \quad (4.2)$$

para $u=0, 1, 2, \dots, N-1$.

Similarmente, a CDT inversa 1-D é definida como:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \quad (4.3)$$

para $x = 0, 1, 2, \dots, N-1$. Nas equações (4.2) e (4.3) o valor de α é:

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}}, p/u = 0 \\ \sqrt{\frac{2}{N}}, p/u = 1 \end{cases} \quad (4.4)$$

A DCT direta 2-D é dada por:

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \quad (4.5)$$

para $u, v = 0, 1, 2, \dots, N-1$.

Enquanto a DCT inversa 2-D pode ser calculada como:

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v) C(u, v) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right] \quad (4.6)$$

para $x, y = 0, 1, 2, \dots, N-1$, sendo que α é dado de acordo com a equação (4.4).

Com base nessa formulação, parte-se da mesma abordagem apresentada em [NGO 2001] onde são utilizados somente os 10 primeiros coeficientes da DCT em cada bloco 8x8 de uma imagem JPEG. Sendo assim, para cada imagem de dimensão 512x512 utilizada, obtém-se um total de 4096 blocos 8x8 (Figura 4.1), a partir dos quais são

extraídos 10 agrupamentos de coeficientes. O agrupamento desses coeficientes é obtido através da aplicação de outra transformação, denominada de transformação Mandala [YSH 83].

F ₀₀	F ₀₁	F ₀₂	F ₀₃	F ₀₄	F ₀₅	F ₀₆	F ₀₇
F ₁₀	F ₁₁	F ₁₂	F ₁₃	F ₁₄	F ₁₅	F ₁₆	F ₁₇
F ₂₀	F ₂₁	F ₂₂	F ₂₃	F ₂₄	F ₂₅	F ₂₆	F ₂₇
F ₃₀	F ₃₁	F ₃₂	F ₃₃	F ₃₄	F ₃₅	F ₃₆	F ₃₇
F ₄₀	F ₄₁	F ₄₂	F ₄₃	F ₄₄	F ₄₅	F ₄₆	F ₄₇
F ₅₀	F ₅₁	F ₅₂	F ₅₃	F ₅₄	F ₅₅	F ₅₆	F ₅₇
F ₆₀	F ₆₁	F ₆₂	F ₆₃	F ₆₄	F ₆₅	F ₆₆	F ₆₇
F ₇₀	F ₇₁	F ₇₂	F ₇₃	F ₇₄	F ₇₅	F ₇₆	F ₇₇

FIGURA 4.1 – Bloco 8x8 de coeficientes DCT

Cada agrupamento mandala é denotado por $I_{x^u y^v}$, onde $u, v = 0, 1, \dots, 7$ são os índices associados a cada coeficiente da DCT. A aplicação da Transformação Mandala resulta na obtenção de 64 agrupamentos, dispostos segundo a seguinte ordem: $I, I_y, I_x, I_{x^2}, I_{xy}, I_{y^2}, I_{y^3}, I_{xy^2}, I_{x^2y}, I_{x^3}, \dots, I_{x^7y^7}$. A partir disso são extraídos os 10 primeiros agrupamentos (mandalas), onde cada um representa uma imagem de dimensão 64x64 com resolução inferior a da imagem original, mas que representa um conteúdo de frequência específico da mesma. O primeiro agrupamento ($u=v=0, I_{x^u y^v} = I_{x^0 y^0} = I$), contendo informações de cor, recebe a denominação de imagem DC (Figura 4.2) e os demais de imagens AC.

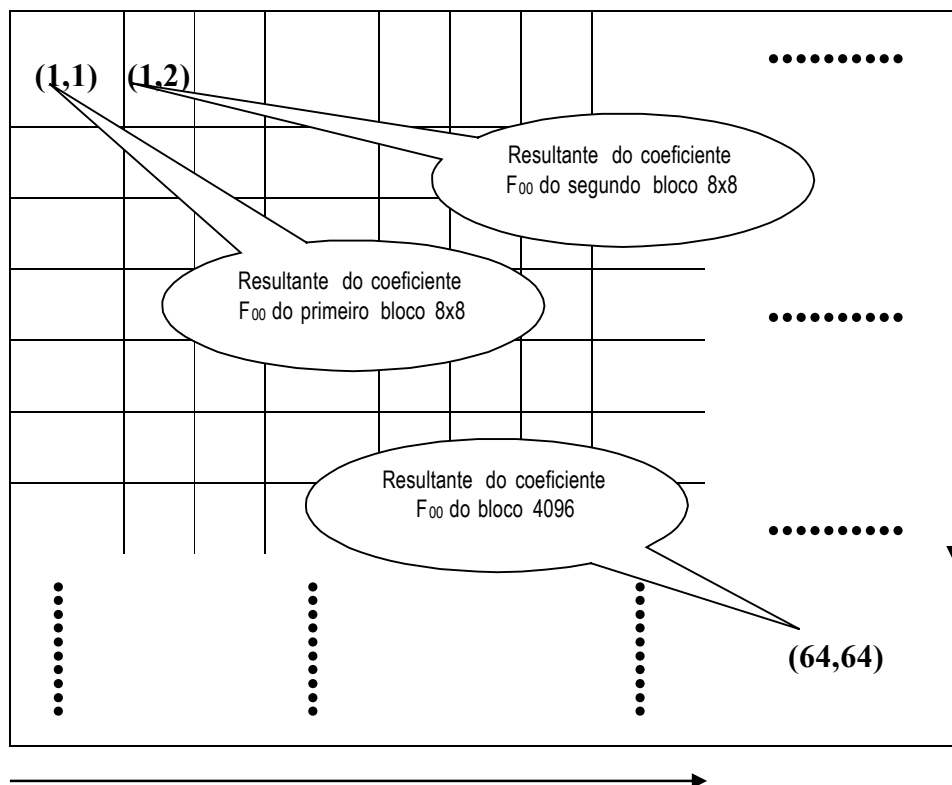


FIGURA 4.2 – Imagem DC

A partir de cada agrupamento são extraídas feições para descrição de propriedades de cor e textura da imagem original. Os coeficientes DC representam a média das intensidades dos pixels de cada bloco 8x8 no domínio espacial, assim como a composição cromática da imagem de forma simplificada. Após obter-se a imagem DC, calcula-se o conteúdo cromático. As propriedades de textura são descritas com base em valores de variância obtidos dos agrupamentos 2 a 10.

O que se obtém com essa abordagem é uma descrição global do conteúdo de cor e textura das imagens. Em função disso é que se propõe a introdução de um mecanismo de realimentação de relevâncias, com o objetivo de refinar o resultado no processo de recuperação de imagens.

Como um dos objetivos do trabalho é o de testar a eficiência de um mecanismo de realimentação de relevâncias no processo de refinamento de consultas de imagens, optou-se inicialmente por uma representação do conteúdo cromático menos precisa que a proposta em [NGO 2001], que utiliza histograma de cores. Por essa razão, escolheu-se a média de cores baseada no sistema de cores RGB, que pode representar o conteúdo cromático de uma imagem de forma menos precisa que o histograma de cores e dessa forma avaliar os efeitos do MRR utilizando-se essa representação. Sendo assim, para cada imagem DC obtida é extraído um vetor de feição de cores definido com base nas cores dominantes presentes na imagem e expresso através de um vetor médio de cores $\bar{x} = (R_{med}, G_{med}, B_{med})$:

$$R_{med} = (1/n) \sum_{p=1}^n R(p), \quad (4.7)$$

$$G_{med} = (1/n) \sum_{p=1}^n G(p), \quad (4.8)$$

$$B_{med} = (1/n) \sum_{p=1}^n B(p) \quad (4.9)$$

onde n representa o número de coordenadas (x,y) em cada um dos canais RGB e $R(p)$, $G(p)$ e $B(p)$ são os valores assumidos para cada canal red (vermelho), green (verde) e blue (azul) na imagem DC.

A similaridade entre duas imagens utilizando o critério de cor é medida com base na distância Euclidiana entre dois vetores médios de cores $\bar{\mathbf{x}}$ e $\bar{\mathbf{y}}$, através da seguinte equação:

$$d_m^2(\bar{\mathbf{x}}, \bar{\mathbf{y}}) = (\bar{\mathbf{x}} - \bar{\mathbf{y}})^T \cdot (\bar{\mathbf{x}} - \bar{\mathbf{y}}) \quad (4.10)$$

Os coeficientes AC representam a variação de energia em diferentes bandas do domínio da DCT [JAI 89]. Como a DCT compacta a energia em coeficientes de ordem mais baixa, consideram-se somente os nove primeiros coeficientes AC.

A feição de textura é descrita através da seguinte equação:

$$S_{uv} = E[I_{x^u y^v}^2] - E[I_{x^u y^v}]^2 \quad (4.11)$$

onde S_{uv} e $E[I_{x^u y^v}]$ representam a variância e o valor esperado associados a um determinado agrupamento mandala $I_{x^u y^v}$. O vetor de feição de textura resultante é obtido após o cálculo de variância aplicado em cada um dos nove agrupamentos mandala $I_y, I_x, I_{x^2}, I_{xy}, I_{y^2}, I_{y^3}, I_{xy^2}, I_{x^2y}, I_{x^3}$, obtendo-se com isso para cada imagem um vetor de feições $\mathbf{x} = (S_{01}, S_{10}, S_{20}, S_{11}, S_{02}, S_{03}, S_{12}, S_{21}, S_{30})$.

A similaridade entre uma imagem e a imagem exemplo de consulta utilizando o critério de textura é medida com base na distância de Mahalanobis, através da seguinte equação [HON 2000]:

$$d_{text}^2 = (\mathbf{x} - \mathbf{q})^T C^{-1} (\mathbf{x} - \mathbf{q}) \quad (4.12)$$

onde \mathbf{q} representa o vetor da imagem de consulta e C^{-1} é a matriz de covariância em relação as nove feições de textura no domínio da DCT, associadas ao vetor de consulta.

5 Mecanismo de Realimentação de Relevâncias (MRR)

O mecanismo de realimentação de relevâncias proposto neste trabalho é baseado em algoritmos geométricos, e tem como objetivo principal a atualização de parâmetros da métrica de similaridade (distância entre vetores de feição), mais precisamente do valor de similaridade (VS) de forma iterativa. Os itens relevantes (positivos) são ajustados ao mesmo tempo em que exclui os não relevantes (negativos) do conjunto de imagens recuperadas a cada iteração. Com isso, pretende-se obter uma estratégia que combine realimentação de imagens relevantes positivas e negativas. O processo de realimentação ocorre de forma interativa entre o usuário e o sistema, que necessita indicar apenas quais das imagens recuperadas considera como relevantes em relação à consulta exemplo utilizada.

5.1 Elementos do MRR

I – Imagem

m – Número total de imagens

q – Imagem de consulta

F = $\{f_i\}$ – Conjunto de feições (cor, textura), onde $i=1$ ou $i=2$;

R = $\{r_{ij}\}$ – Conjunto de representações (média de cores, variância)

A representação r_{ij} de uma determinada feição é expressa através de vetores do tipo $\mathbf{r}_{ij} = (r_{ij1}, r_{ij2}, \dots, r_{ijk})$, onde k é a dimensão do vetor \mathbf{r} .

M = $\{m_{ij}\}$ – Conjunto de métricas de similaridade (distância euclidiana, distância de Mahalanobis).

S = Medida de Similaridade obtida através da métrica m_{ij}

VS = Valor de similaridade, que possui um valor inicial dado por S e que vai sendo atualizado a cada iteração durante o processo de realimentação de relevâncias.

R = $\{I_1, I_2, \dots, I_{nrec}\}$ – Conjunto de imagens recuperadas, onde $nrec$ é o número de imagens recuperadas em cada iteração.

$nrel$ = Número de imagens relevantes em **R**.

5.2 Diagrama de Representação de Relacionamento entre Elementos do MRR

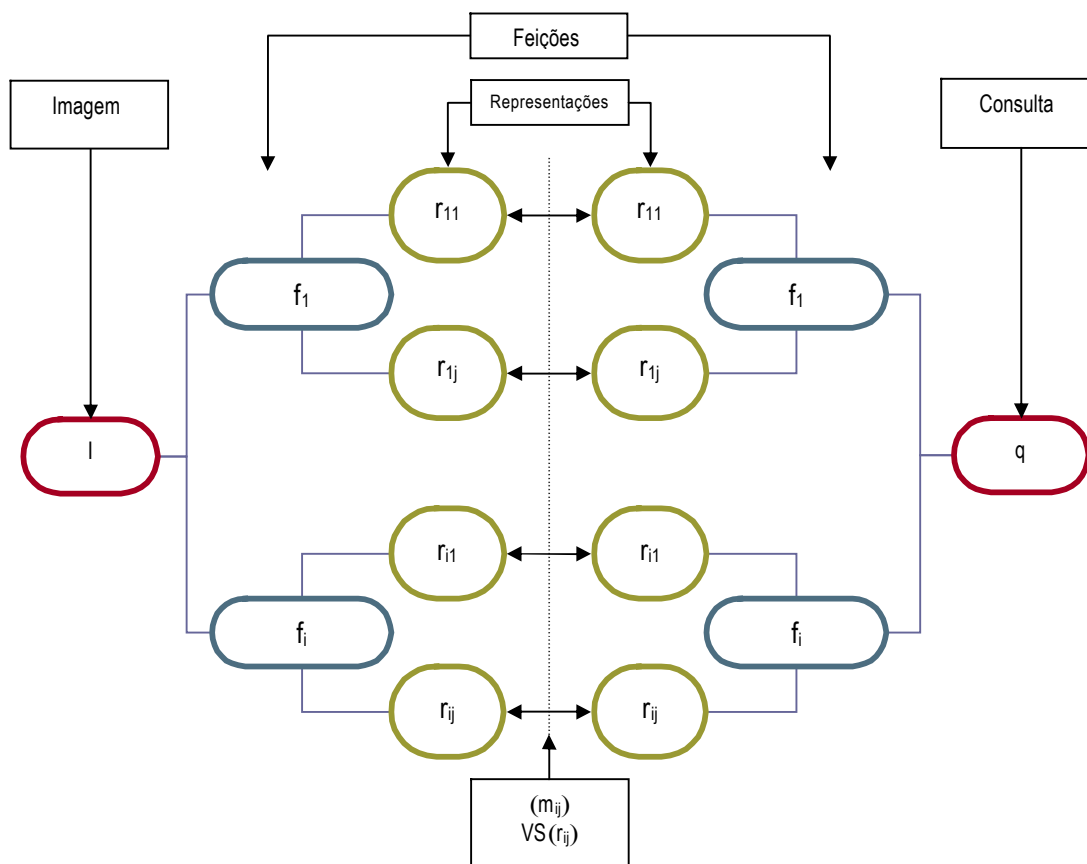


FIGURA 5.1 – Diagrama de Relacionamento entre Elementos do MRR [RUI 98]

Na figura anterior, a informação contida em q se desloca no sentido de I e vice-versa. A linha pontilhada representa o ponto onde é aplicada a métrica (m_{ij}) para o cálculo dos valores de similaridade $(VS(r_{ij}))$ entre q e cada imagem I .

5.3 Descrição do MRR

O mecanismo funciona através de um sistema de atualização automática de pesos, definidos da seguintes forma:

$p = 1 \rightarrow$ peso atribuído às imagens relevantes

$p = 2 \rightarrow$ peso atribuído às imagens não relevantes

O mecanismo tem como objetivo, enfatizar as imagens com maior relevância, através da atualização de pesos a cada iteração, aproximando-as cada vez mais do início do rank de ordenação por similaridade. As imagens não relevantes são afastadas para o final da coleção, de forma permanente, a não ser que outra consulta seja solicitada.

A proposta de MRR deste trabalho difere de outras abordagens mais utilizadas como em [RUI 97] e [HON 2000], pois executa a atualização de pesos diretamente no valor de similaridade obtido através da métrica de distância. Para isso, houve a necessidade de se definir uma expressão de normalização dos valores de similaridade, com o objetivo de fixar todas as imagens relevantes dentro de um intervalo que esteja relacionado à imagem de consulta, através de um valor de tolerância correspondente ao tamanho do intervalo dentro do espaço de feições. A representação do intervalo de tolerância pode ser visualizado através da Figura 5.2 como sendo uma esfera fechada [LIM 77], vizinhança esférica ou bola de raio $\varepsilon=1$ onde estão contidas todas as imagens relevantes, através de suas respectivas representações do vetor de feições. Nos experimentos realizados utilizou-se um intervalo de comprimento $\varepsilon=1$.

Para cada feição f_i está associado um conjunto \mathbf{R} de representações r_{ij} . No nosso caso, a média de cores é a única representação associada à feição de cor, e a variância é a representação associada à feição de textura. O valor de similaridade entre uma imagem I e a consulta q , considerando uma determinada representação r_{ij} é calculada de acordo com uma medida m_{ij} , definida no trabalho através da distância euclidiana para cor e a distância de Mahalonobis para a textura.

Para uma determinada feição f_i tem-se as seguintes expressões definidas para o cálculo do valor de similaridade:

$$S(f_i) = \sum_j p_{ij} S(r_{ij}) \quad (5.1)$$

$$S = \sum_i p_i S(f_i) = \sum_i \sum_j p_{ij} S(r_{ij}) \quad (5.2)$$

onde $S(f_i)$, é o VS obtido em função de uma única feição i e S (VS total entre uma imagem I e a consulta q , se consideradas mais de uma feição no processo de recuperação). Como nos experimentos realizados considerou-se a recuperação de imagens somente através do critério de cor ou textura, utilizou-se somente a equação (5.1), onde $p_{ij}=1$ em ambos os casos (cor ou textura), pois foi considerada somente uma representação, tanto para cor como para textura. A razão para tal escolha parte da observação de que determinadas imagens não apresentam padrões de textura propriamente definidos, mas somente padrões de cores que podem estar associadas a uma determinada forma, como as apresentadas em [CIO 2001]. Essa seria a principal razão em aplicar-se inicialmente, de forma separada, o MRR somente a cor ou textura.

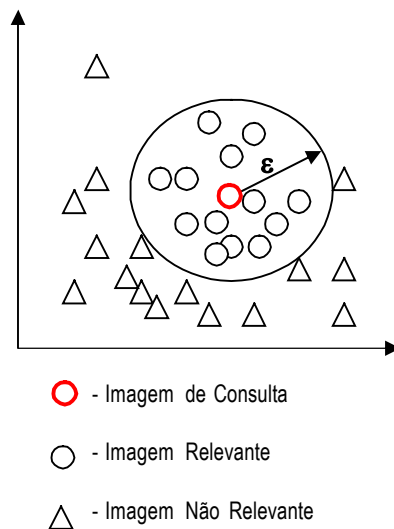


FIGURA 5.2 – Espaço de feições em relação à imagem de consulta

No início do processo de refinamento, ou seja, na primeira iteração com realimentação, o usuário destaca todas as imagens consideradas relevantes. Essas imagens recebem, automaticamente, através do algoritmo, um peso $p=1$ associado aos seus respectivos valores de similaridade (VS). Em seguida esses valores são normalizados para valores contidos dentro do intervalo $[(n-1),n]$, para o qual foi definido o valor $n=1$, a fim de que se possa garantir uma distância mínima onde se obtenha bons resultados de imagens recuperadas. Para o processo de normalização, obteve-se a expressão abaixo, demonstrada através do método de indução finita [LIM 76]:

$$S = n - 1/(S + 1) \quad (5.3)$$

Demonstração:

Seja $S > 0$ e defina $S_0 = S$. Então a sequência definida por $S_{k+1} = n - 1/(S_k + 1)$ satisfaz a desigualdade $n - 1 < S_k < n$, para todo n natural e todo $k = 1, 2, 3, 4, \dots$

Prova (indução sobre k):

- Mostremos inicialmente que a desigualdade vale para $k = 1$. Para isto, temos que $S_1 = n - 1/(S_0 + 1) = n - 1/(S + 1) < n$, uma vez que $0 < 1/(S+1) < 1$. Por outro lado, $S_1 - (n - 1) = n - 1/(S + 1) - n + 1 = 1 - 1/(S + 1) = S/(S + 1) > 0$. Logo, $n - 1 < S_1 < n$.
- Suponhamos que a propriedade vale para k , ou seja, que $n - 1 < S_k < n$. Devemos mostrar que a propriedade vale para $k + 1$, ou seja, $n - 1 < S_{k+1} < n$. Com efeito,

temos que por hipótese de indução: $n - 1 < S_k < n \Rightarrow n < S_k + 1 < n + 1 \Rightarrow 1/(n + 1) < 1/(S_k + 1) < 1/n \Rightarrow -1/n < -1/(S_k + 1) < -1/(n + 1) \Rightarrow n - 1/n < n - 1/(S_k + 1) < n - 1/(n + 1) \Rightarrow n - 1 < (n - 1) + (1 - 1/n) < S_{k+1} < n - 1/(n + 1) < n$. Logo $n - 1 < S_{k+1} < n$.

onde S representa o VS de cada imagem relevante que será posicionado dentro do intervalo de valores crescentes $[0,1]$, sendo que os valores de similaridade das imagens mais relevantes aproximam-se de 0 pela esquerda. Dessa forma, admite-se que a ordem dos valores de similaridade seja decrescente, apesar de pertencerem a um intervalo crescente de números reais. As imagens não relevantes também recebem um peso de valor $p=2$ associado aos seus respectivos VS, sem que ocorra a normalização dos valores resultantes. Com essa estratégia de atualização de pesos, o que ocorre durante o processo de realimentação das imagens não relevantes é um progressivo deslocamento do VS das mesmas para fora do intervalo $[0,1]$ de forma permanente. Esse deslocamento das imagens não relevantes provoca ao mesmo tempo um reposicionamento de outras imagens para mais próximo da imagem de consulta, através de uma nova ordenação (i. e. “ranking”) dos VS. Se alguma dessas imagens fizer parte do conjunto de imagens recuperadas por RR, então poderá ser submetida ao processo de realimentação em uma próxima iteração, atribuindo-se $p=1$ ao seu respectivo valor de similaridade.

5.3.1 Algoritmo de RR

- (1) Definir o tipo de relevância (relevante ou não relevante)
- (2) Faça para $i=1$ até $nrec$
 - Se I é relevante
 - $VS=1-1/(VS+1)$; (normalização para o intervalo $[0,1]$)
 - Senão
 - $p=2$; (afasta a imagem irrelevante)
 - $VS=p.VS$;
- (3) Reordenar todos os m (número de imagens) valores de similaridade (VS);
- (4) Exibir o conjunto R de imagens recuperadas;
- (5) Fim

Na implementação do algoritmo proposto devem estar definidos previamente os valores da variável m e da variável $nrec$. A primeira é necessária no processo de reordenação de todos os valores de similaridade associados às imagens do banco de imagens e a segunda no processo de verificação de relevância e atualização de VS somente das imagens pertencentes ao conjunto R .

6 Resultados Experimentais

O conjunto de imagens teste é constituído de 345 imagens coloridas, sendo que 195 são utilizadas em testes de recuperação de imagens com textura colorida, e 150 na recuperação de imagens utilizando somente o critério de cor. As imagens foram todas obtidas a partir do repositório disponível para “download” na Home Page da Universidade de Stanford em <http://www-db.stanford.edu/IMAGE/>. As dimensões são variadas, entretanto, durante o processo de implementação do método para representação e recuperação, todas as imagens foram redimensionadas para o tamanho 512x512. A implementação do método para representação e recuperação das imagens foi realizada utilizando-se a ferramenta de prototipação MATLAB 5.3.

O objetivo principal dos experimentos realizados não consistiu na comparação do método de representação e recuperação de imagens com outros já existentes, mas sim na comparação de desempenho do processo de recuperação sem RR e com RR. O método de representação já havia sido implementado em [NGO 2001], sem a utilização de um mecanismo de realimentação de relevâncias que pudesse refinar o resultado do processo de recuperação de imagens. Foram consideradas medidas de precisão (*precision*) e recuperação (*recall*), para avaliar o desempenho da recuperação de informações. Recall está relacionada à proporção de imagens relevantes em relação ao conjunto R , e Precision se refere à proporção de imagens recuperadas (*nrec*) que são efetivamente relevantes [BAE 99]. Com base na Figura 6.1, que ilustra uma coleção qualquer de imagens de teste, e a partição que distingue as imagens entre relevantes e não relevantes, o cálculo dos valores de precision (π) e recall (ρ), foram obtidos a partir das seguintes expressões:

$$\pi = RR / nrec \quad (6.1)$$

$$\rho = RR / (NRR + RR), \quad (6.2)$$

onde $\rho = \pi \times nrec / (NRR + RR) = \pi \times nrec / NTR$ e $NTR = NRR + RR$ é o número total de imagens relevantes na coleção [BAE 99].

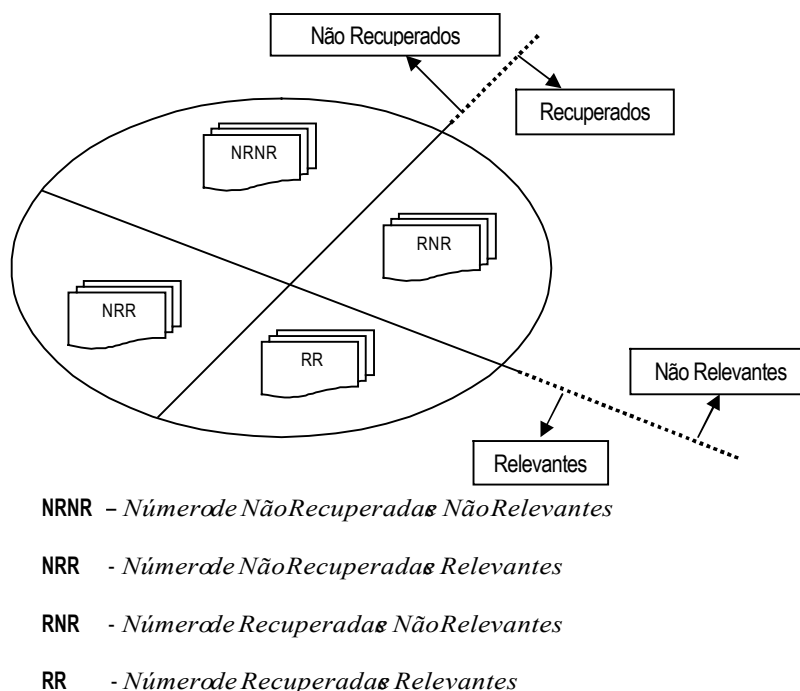


FIGURA 6.1 – Partição de uma Coleção de Imagens em relação a uma imagem de consulta q

6.1 Recuperação de Imagens Através das Cores

Utilizou-se na implementação uma coleção com 150 imagens coloridas, constituída de imagens onde o conteúdo apresentava flores ornamentais, animais e paisagens (Anexo 1). A justificativa para essa escolha foi em função da necessidade de se testar o desempenho do MRR em diferentes situações de heterogeneidade quanto ao conteúdo cromático. Em ambos os casos se considerou um conjunto R onde $nrec$ variou de 10 a 55. Para avaliação de desempenho do processo de recuperação tanto de cor como de textura adotou-se os mesmos valores de $nrec$.

A Figura 6.2 exibe o resultado de uma consulta através de exemplos sem realimentação de relevâncias, utilizando a coleção de imagens apresentada no Anexo 1 (nesse exemplo utilizou-se somente 100 imagens), enquanto que as Figuras 6.3 e 6.4, exibem os resultados do processo de realimentação após 2 iterações. A imagem 1 (flor amarela) foi escolhida como consulta, sendo que para o usuário são retornadas somente as 20 (vinte) primeiras imagens, exibidas por ordem decrescente de VS. Admitiu-se como relevantes para a primeira iteração as imagens 1, 2, 3, 4, 5, 6, 7, 8, 10, 11 e 15.



FIGURA 6.2 – Resultado de Consulta Através de Cor sem RR



FIGURA 6.3 – Resultado de Consulta Através de Cor com RR – 1ª Iteração
Relevantes: 1, 2, 3, 4, 5, ..., 15 e 18



FIGURA 6.4 – Resultado de Consulta Através de Cor com RR – 2ª Iteração
Relevantes: 1, 2, 3, 4, 5, ..., 20

6.2 Recuperação de Imagens Através das Texturas

O Anexo 2 exibe uma coleção de 195 imagens de texturas coloridas utilizadas nos testes de implementação. A Figura 6.5 mostra o resultado de uma consulta de textura sem realimentação de relevâncias, enquanto que as Figuras 6.6, 6.7 e 6.8 exibem os resultados após três iterações para refinamento da consulta. Assim como na recuperação através de cor, o resultado da consulta mostra somente as 20 primeiras imagens ao usuário, onde a primeira imagem exibida representa a imagem de consulta.

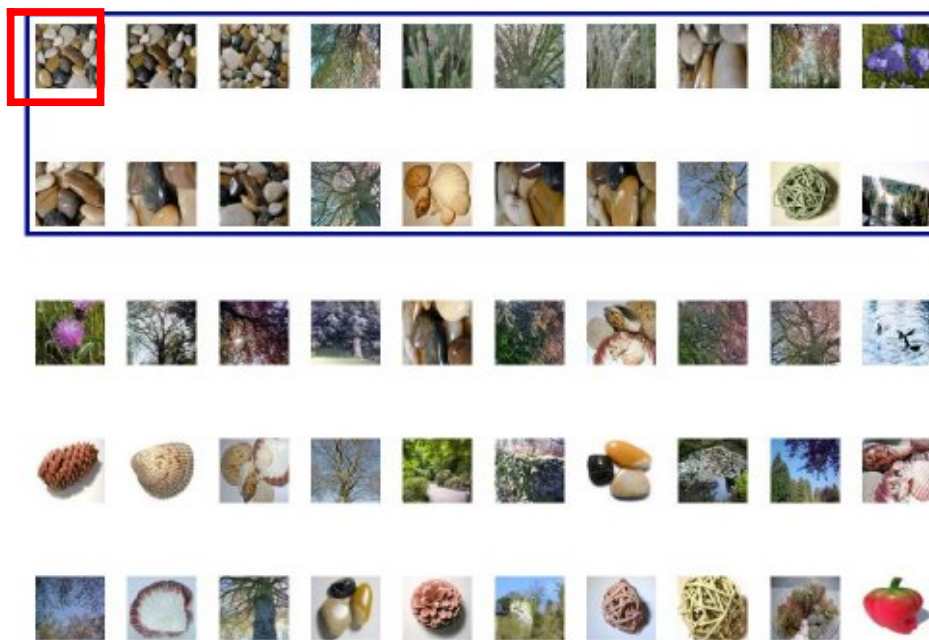


FIGURA 6.5 – Resultado de Consulta Através de Textura sem RR
Relevantes: 1, 2, 3, 8, 11, 12, 13, 16 e 17

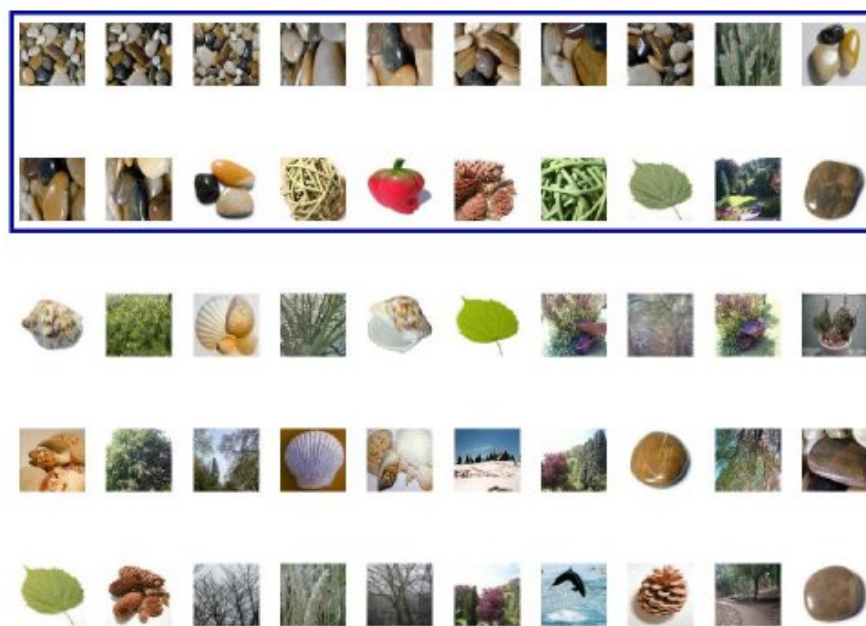


FIGURA 6.6 – Resultado de Consulta Através de Textura com RR – 1ª Iteração
Relevantes: 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12 e 13

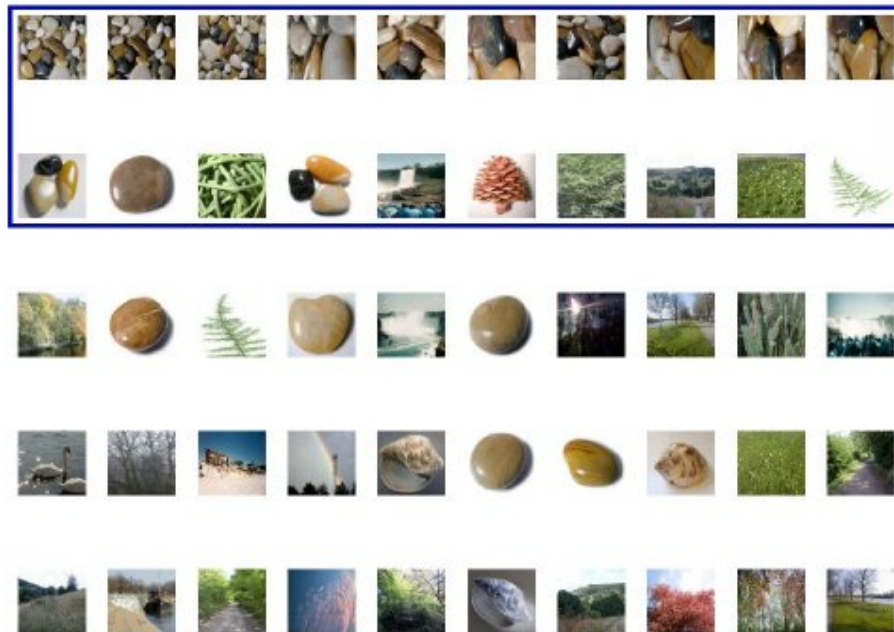


FIGURA 6.7 – Resultado de Consulta Através de Textura com RR – 2ª Iteração
Relevantes: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 e 14

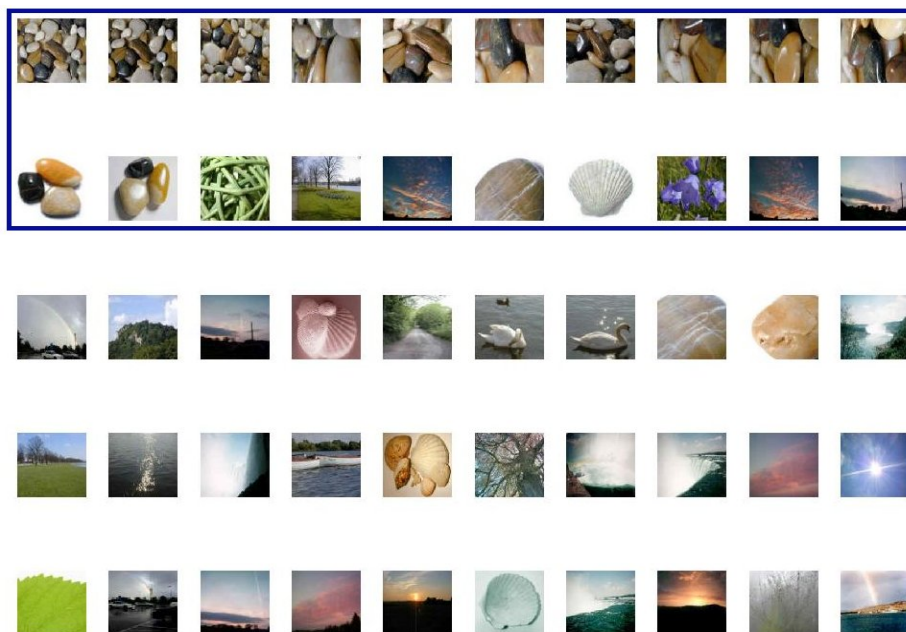


FIGURA 6.8 – Resultado de Consulta Através de Textura com RR – 3ª Iteração
Relevantes: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 e 13

6.3 Avaliação de Desempenho de Recuperação

O método de avaliação de desempenho adotado foi baseado nas medidas de precisão e recuperação, descritos no início do Capítulo 6. O objetivo principal do processo de avaliação consiste na análise de gráficos comparativos, exibindo a relação existente entre o resultado de uma consulta sem RR, e o resultado de refinamento dessa consulta após a aplicação do mecanismo de realimentação de relevâncias. Utiliza-se para tal a proporção de documentos recuperados e relevantes (medida de precisão), pertencentes ao conjunto R (conjunto de documentos recuperados). Com esse objetivo, executou-se uma série de 100 consultas, divididas segundo o número de imagens recuperadas a cada interação do usuário com o sistema, ou seja, 5 (cinco) consultas para cada valor $nrec$. Com base nesses resultados, construiu-se duas tabelas, uma para recuperação de imagens por cor (Tabela 6.1) e a outra para recuperação através de textura (Tabela 6.2), contendo as médias das medidas de recall e precision. A Figura 6.9 exibe o gráfico comparativo para cor e a Figura 6.10 o gráfico comparativo para textura.

TABELA 6.1 – Médias de Recall e Precision para Recuperação por Cor

No. de Consultas = 50								
Sem Realimentação de Relevância			Com Realimentação de Relevância					
			Recall Médio			Precision Médio		
Número de Imagens Recuperadas	Recall Médio	Precision Médio	1a. Iteração	2a. Iteração	3a. Iteração	1a. Iteração	2a. Iteração	3a. Iteração
10	0,35	0,60	0,37	0,42	0,42	0,65	0,70	0,70
15	0,40	0,47	0,44	0,44	0,53	0,50	0,50	0,57
20	0,44	0,38	0,47	0,56	0,58	0,40	0,45	0,48
25	0,44	0,30	0,58	0,63	0,79	0,38	0,40	0,48
30	0,47	0,27	0,58	0,74	0,90	0,32	0,38	0,47
35	0,47	0,23	0,63	0,86	0,98	0,29	0,39	0,44
40	0,54	0,23	0,72	0,91	0,98	0,29	0,36	0,39
45	0,58	0,21	0,84	0,93	1,00	0,29	0,33	0,36
50	0,58	0,20	0,93	0,98	1,00	0,29	0,31	0,32
55	0,58	0,18	0,98	1,00	1,00	0,28	0,29	0,29

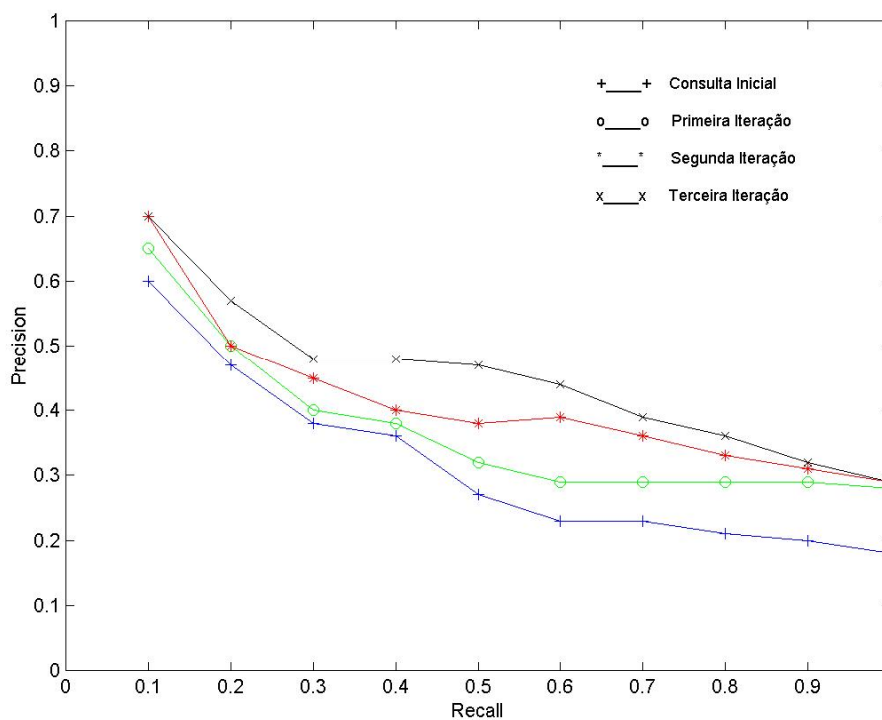


FIGURA 6.9 – Gráfico Comparativo de Recuperação de Imagens Através do Conteúdo de Cor

TABELA 6.2 – Médias de Recall e Precision para Recuperação por Textura

Sem Realimentação de Relevância		Com Realimentação de Relevância						
		Recall Médio			Precision Médio			
Número de Imagens Recuperadas	Recall Médio	Precision Médio	1a. Iteração	2a. Iteração	3a. Iteração	1a. Iteração	2a. Iteração	3a. Iteração
10	0,33	0,75	0,38	0,41	0,41	0,80	0,85	0,85
15	0,41	0,60	0,48	0,46	0,54	0,67	0,63	0,73
20	0,49	0,55	0,54	0,61	0,69	0,58	0,63	0,73
25	0,54	0,48	0,64	0,71	0,74	0,54	0,60	0,62
30	0,62	0,47	0,69	0,79	0,81	0,50	0,57	0,58
35	0,69	0,44	0,81	0,82	0,84	0,50	0,51	0,53
40	0,72	0,40	0,82	0,85	0,92	0,45	0,48	0,53
45	0,76	0,37	0,84	0,90	0,93	0,41	0,46	0,48
50	0,76	0,33	0,89	0,92	0,97	0,40	0,42	0,45
55	0,82	0,33	0,92	0,97	0,97	0,38	0,41	0,41

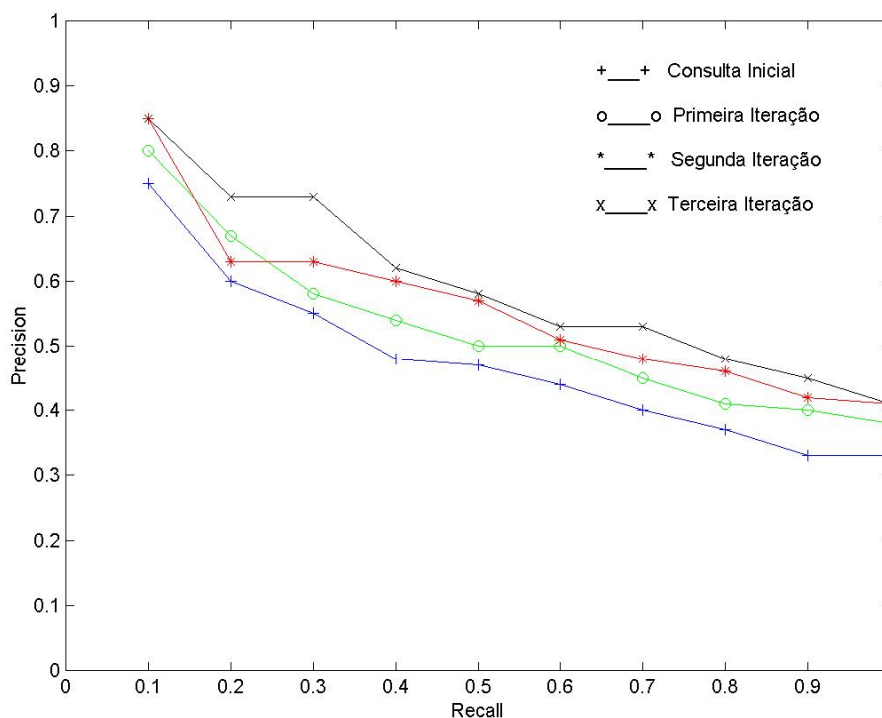


FIGURA 6.10 – Gráfico Comparativo de Recuperação de Imagens Através do Conteúdo de Textura

O que se observou nos dois casos avaliados, através dos gráficos obtidos, é o fato de que os valores de precisão e recall aumentam de acordo com o processo de realimentação de relevâncias. Com isso, consegue-se mostrar que o ganho conseguido (refinamento) em termos de resultados, no processo de recuperação das imagens, se deu exclusivamente à aplicação do MRR proposto, pois nenhum outro elemento foi acrescentado ou alterado durante esse processo, como por exemplo a alteração do tipo de representação utilizado para as imagens, que pode retornar um número mais preciso de imagens em termos de similaridade na primeira iteração, ou ainda a métrica no cálculo de similaridade. O equipamento utilizado durante a execução das consultas foi uma máquina com processador Pentium MMX 266 com 64 MB de memória RAM e 2MB de memória de vídeo, obtendo-se um tempo de resposta de 3 (três) a 5 (cinco) segundos em média, de acordo com o número de imagens a serem exibidas de uma só vez no processo de recuperação, o que variou de 10 a 50 imagens. Além disso, o número de iterações depende basicamente de dois fatores:

- 1) Número de imagens ($nrec$) definido para o conjunto de imagens a serem recuperadas (R) em uma consulta (consulta k-vizinhança) – neste caso utilizou-se valores entre 10 e 55 para avaliação de resultados.
- 2) Eficiência e precisão do método utilizado para representação e recuperação das imagens – quanto mais eficiente e preciso, maior será o número de imagens recuperadas em uma primeira consulta sem realimentação. Com isso, o número de iterações exigido para o processo de refinamento de resultados será menor.

Outros fatores que influenciaram na precisão dos resultados obtidos, foram:

- A utilização de imagens com características de homogeneidade variada, principalmente na recuperação através de cor, onde o conteúdo cromático foi obtido dentro do espaço de cor RGB, sem levar em consideração o arranjo espacial do conteúdo de cor extraído.

- A subjetividade de percepção por parte do usuário, uma vez que o julgamento daquilo que seja considerado como relevante ou não, pode variar de um usuário para outro.

Comparado aos resultados com refinamento de consulta apresentados em [RUI 98] e [CIO 2001], utilizados como referências na apresentação de uma proposta de refinamento, observou-se que o método funciona dentro dos mesmos valores de precisão se considerado um número de imagens recuperadas somente entre 10 e 20 imagens, entretanto, esse valor tende a sofrer um aumento consideravelmente maior em relação aos dois trabalhos citados, quando o número de imagens recuperadas tende a aumentar. Esse valor de precisão também se mostrou bem maior na medida em que as imagens de uma determinada coleção apresentavam maior heterogeneidade em termos de cor e textura.

Apesar dos resultados quanto à precisão serem menores que os apresentados nos dois trabalhos citados, acreditamos que a utilização de outras feições além de cor e textura, assim como a combinação dessas feições, possa ser capaz de melhorar consideravelmente os resultados até então obtidos, uma vez que em nossa proposta de MRR adotou-se uma simplificação na representação das imagens, utilizando-se somente cor e textura sem algum tipo de combinação, o que se acredita ter contribuído para a obtenção de resultados menos precisos.

7 Conclusões e Trabalhos Futuros

Neste trabalho apresentou-se uma proposta de mecanismo de realimentação de relevâncias. Para isso foram introduzidos inicialmente conceitos relacionados à área de Recuperação de Informações Visuais Baseada em Conteúdo (CBVIR). Em seguida descreveu-se o problema de recuperação de informações visuais apresentado em [NGO 2001], com modificação da abordagem quanto a representação do conteúdo cromático. Propôs-se então, a aplicação de um mecanismo de realimentação, já que o mesmo também não foi implementado no referido trabalho. Para o desenvolvimento do mecanismo apresentado, baseou-se em parte dos trabalhos desenvolvidos por [RUI 98], [ALB 98], [AKS 2000] e [CIO 2001], obtendo-se um mecanismo associado exclusivamente às feições de baixo nível. Quanto ao posicionamento do trabalho em relação ao estado da arte, o que falta é uma associação das feições de baixo nível a conceitos de alto nível, o que já vem sendo apresentado em trabalhos como em [ZHA 2002] e [WAN 2002]. Propõe-se também a combinação de feições de cor e textura em uma única consulta, uma vez que os resultados do trabalho apresentado foram obtidos com base na implementação das consultas utilizando somente o conteúdo de cor ou textura de forma separada. Qualquer implementação proposta de forma combinada, pode ser feita a partir da representação descrita na seção 5.3, utilizando-se as equações (5.1) e (5.2).

Como o objetivo principal do trabalho não visava a elaboração de um estudo comparativo de métodos para realimentação de relevâncias, optou-se então por uma avaliação final de desempenho, que demonstrasse a eficiência do mecanismo em melhorar o resultado de recuperação de informações visuais para o problema apresentado.

Com base nas conclusões obtidas, pode-se sugerir como trabalhos futuros, a utilização de outras representações para as imagens, que capturem o conteúdo cromático e de textura através de um número mais detalhado de informações (feições locais), com o objetivo de um estudo comparativo mais geral de desempenho do mecanismo de realimentação de relevâncias proposto, que possibilite a sua validação para mais de uma representação de imagens. Associado a isso se propõe também outros estudos, que melhorem a descrição e possibilitem a implementação do MRR em outros ambientes de programação, como JAVA e C, assim como testar a sua aplicação em grandes bancos de dados de imagens.

Anexo 1

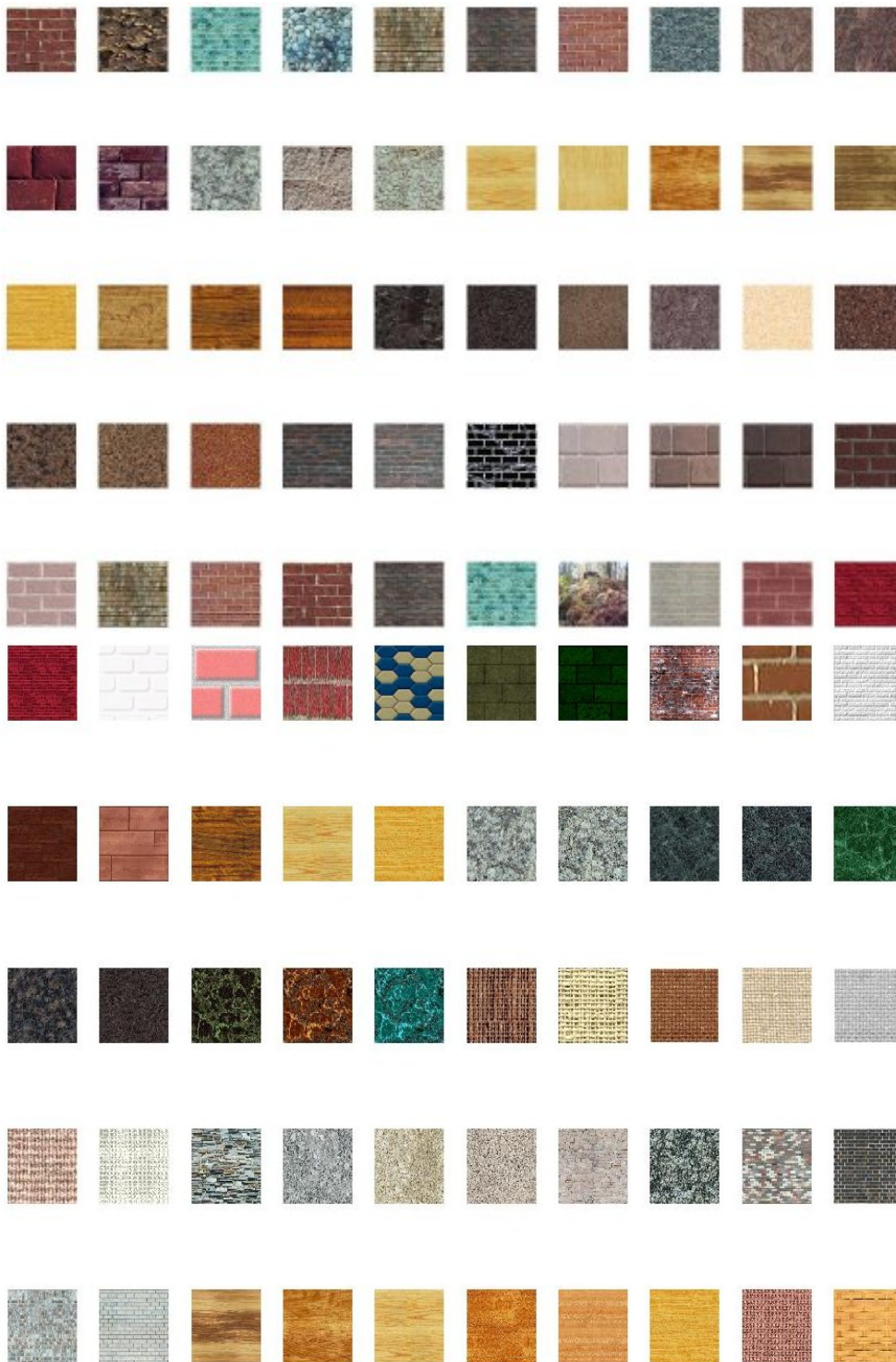
Coleção 1 de Imagens – Recuperação de Imagens Através das Cores

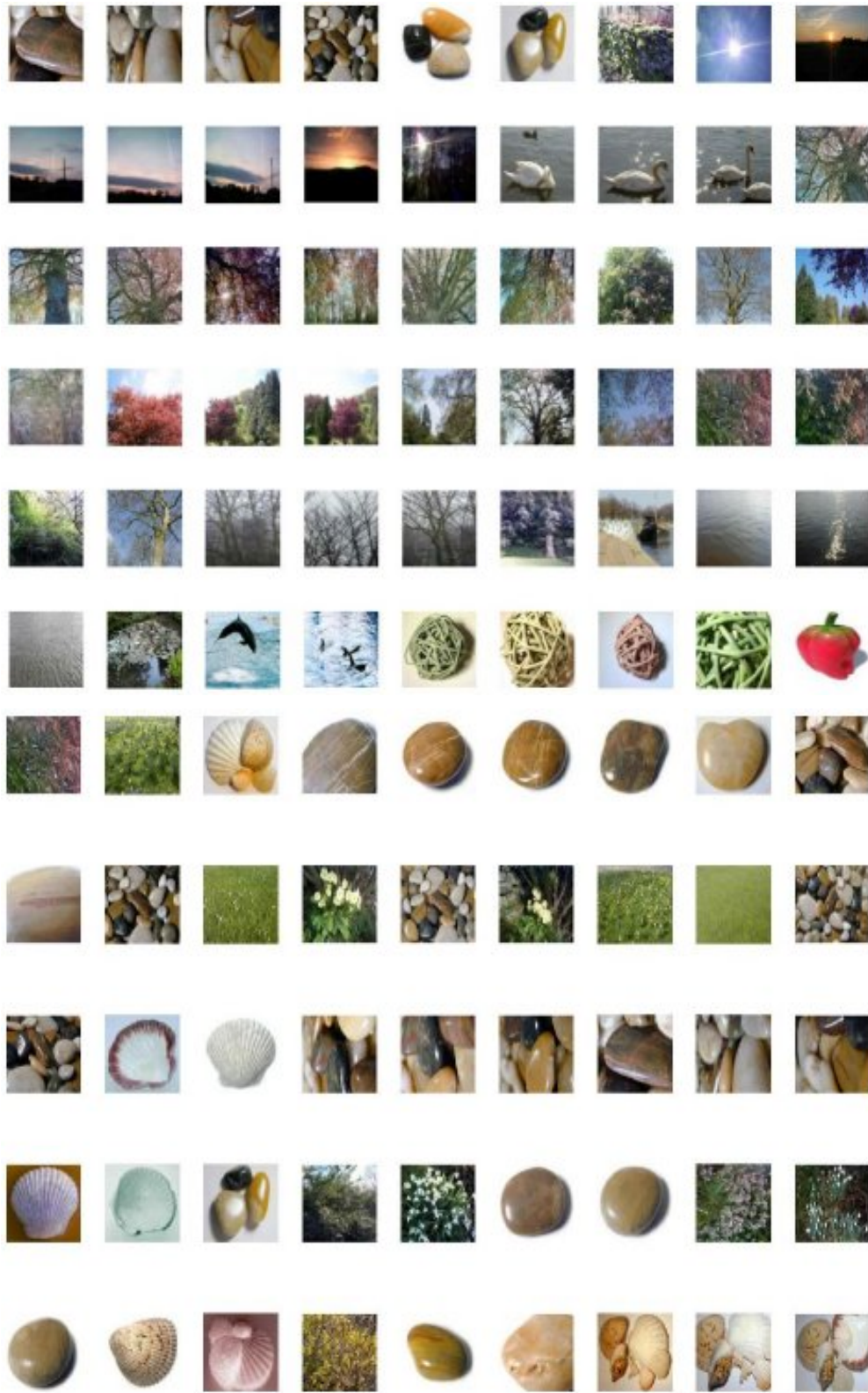




Anexo 2

Coleção 2 de Imagens – Recuperação de Imagens Através das Texturas





Bibliografia

- [ARIS 2000] ARISAWA, H.; CATARCI, T. **Advances in Visual Information Management: Visual Database Systems**. Boston: Kluwer Academic 2000.
- [AGG 2002] AGGARWAL, G.; GHOSAL, S. An Image Retrieval System With Automatic Query Modification. **IEEE Transactions on Multimedia**, Piscataway, v.4, n.2, p.201-214, June 2002.
- [AKS 2000] AKSOY, S.; HARALICK, R. M. Probabilistic vs. Geometric Similarity Measure for Image Retrieval. In: INTERNATIONAL CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. **Proceedings...** South Caroline: IEEE Computer, 2000. v.2, p.13-15.
- [ALB 98] ALBRECHT, R.; HOOK, R. N.; BUSHOUSE, H. A. **A Multidimensional Binary Search Tree for Star Catalog Correlations**. [S.l.]: Astronomical Society of the Pacific, 1998. (Asp Conference Series, v.145).
- [ASH 2001] ASHWIN, T. V.; GHOSAL, S.; JAIN, N. Improving Image Retrieval Performance with Negative Relevance Feedback. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING AND SYSTEMS, **Proceedings...** [S.l.]: IEEE Computer, 2001. p.1637-1640.
- [ASL 99] ASLANDOGAN, Y. A.; YU, C. T. Techniques and Systems for Image and Video Retrieval. **IEEE Transactions on Knowledge and Data Engineering** Piscataway v.11, n.1, p.56-63, January 1999.
- [BAE 99] BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. Reading: Addison-Wesley, 1999.
- [BUC 95] BUCKLEY, C.; SALTON, G. Optimization of Relevance Feedback Weights. In: ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 18., 1995. **Proceedings...** New York: ACM Press, 1995. p.351-357.
- [CAR 97] CARSON, C. et al. Region-Based Image Querying. In: WORKSHOP ON CONTENT-BASED ACCESS OF IMAGE AND VIDEO LIBRARIES. **Proceedings...** New York: IEEE Comp. Soc., 1997. p.42-49.
- [CHA 97] CHANG, S. et al. Videoq: An Automated Content-Based Video Search System Using Visual Cues. In: ACM MULTIMEDIA, 1997. **Proceeding...** New York: ACM, 1997. p.313-324.
- [CIO 99] CIOCCA, G.; SCHETTINI, R. Using a Relevance Feedback Mechanism to Improve Content-based Image Retrieval. In: INTERNATIONAL CONFERENCE ON VISUAL INFORMATION SYSTEMS, VISUAL 3., 1999. **Visual Information and Information Systems**. Berlin: Springer-Verlag, 1999. p.107-114. (Lecture Notes in Computer Science, v.1614).
- [CIO 2001] CIOCCA, G.; SCHETTINI, R. Content-Based Similarity Retrieval of Trademarks Using Relevance Feedback. **The Journal of Recognition Society**, [S.l.], v.34, p.1639-1655, 2001.

- [DEL 99] DEL BIMBO, A. **Visual Information Retrieval**. [S.l.]: Morgan Kaufmann, 1999. p.1-132.
- [DEL 97] DEL MIMBO, A. et al. Image Retrieval by Color Regions. In: INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING **Proceeding...** London: Springer Verlag, 1997. p.180-187. (Lecture Notes in Computer Science, v.1165).
- [DEL 2000] DEL BIMBO, A. Semantics in Visual Information Retrieval. **IEEE MULTIMEDIA**, Los Alamitos, v.6, n.3, 1999.
- [DIS 99] DI SCIASCIO, E.; MINGOLLA, G.; MONGIELLO, M. Content-Based Image Retrieval Over the WEB Using Query by Sketch and Relevance Feedback. In: INTERNATIONAL CONFERENCE ON VISUAL INFORMATION SYSTEMS, 3., 1999. **Visual Information and Information Systems**. Berlin: Springer-Verlag, 1999. p.123-130. (Lecture Notes in Computer Science, v.1614).
- [DOU 2000] DOULAMIS, N.; DOULAMIS, A.; KOLLIAS, D. Non-Linear Relevance Feedback: Improving the Performance of Content-Based Retrieval Systems. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, ICME, 2000. **Proceedings...** New York: IEEE Computer, 2000. p.331-334.
- [DOU 2002] DOULAMIS, N.; DOULAMIS, A. Optimal Recursive Similarity Measure Estimation for Interactive Content-Based Image Retrieval. In: INTERNATIONAL CONFERENCE IN IMAGE PROCESSING, 2000. **Proceedings...** [S.l.:s.n.], 2002. p.972-975.
- [FLI 95] FLICKNER, M.; SAWHNEY, H.; NIBLACK, W. et al. Query by image and video content: the QBIC system. **IEEE Computer Magazine**, [S.l.], v.28, n.9, p.23-32, 1995.
- [FOU 2002] FOURNIER, J.; CORD, M. long-Term Similarity Learning in Content-Based Image Retrieval. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2000. **Proceedings...** [S.l.:s.n.], 2002. p.441-444.
- [GUO 2002] GUO, G.; JAIN, A. K. Learning Similarity Measure for Natural Image Retrieval With Relevance Feedback. **IEEE Transactions on Neural Networks**, New York, v.13, n.4, July 2002.
- [GON 93] GONZALES, R. C.; WOOD, R. E. **Digital Image Processing**. Reading: Addison-Wesley, 1993.
- [GRE 2001] GREGORY, L.; KITTLER, J. Using Contextual Information for Image Retrieval. In: INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING, 11., 2001. **Proceedings...** Los Alamitos: IEEE Computer Society, 2001. p.230-235.
- [HAN 2002] HAN, J.; GUO, L. A New Image Retrieval System Supporting Query by Semantics and Exemple. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2002. **Proceedings...** [S.l.:s.n.], 2002. v.3, p.953-956.

- [HON 2000] HONG, P.; TIAN, Q.; HUANG, T. S. Incorporate Support Vector Machines to Content-Based Image Retrieval With Relevance Feedback. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, ICI, 2000, **Proceedings...** Vancouver, Canadá: IEEE Computer. 2000. v.3, p.750-753.
- [HOU 92] HOU, H. S.; TRETTER, D. R.; VOGUEL, M. J. Interesting Properties of the Discrete Cosine Transform. **Journal of Visual Communication and Image Representation**, [S.l.], v.3, p.73-83, 1992.
- [JAI 89] JAIN, A. K. **Fundamentals of Digital Image Processing**. Englewoods Cliffs: Prentice Hall, 1989.
- [KUS 2002] KUSHKI, A.; ANDROUTSOS, P. Fuzzy Aggregation of Image Features in Content-Based Image Retrieval. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 2002. **Proceedings...** [S.l.:s.n.], 2002. v.3, p.909-912.
- [LIM 76] LIMA, E. L. **Curso de Análise**. Rio de Janeiro: IMPA, 1976. v.1, p.1-76.
- [LIM 77] LIMA, E. L. **Espaços Métricos**. Rio de Janeiro: IMPA/CNPq, 1977. p.8-9.
- [MAN 96] MANJUNATH, B. S.; MA, W. Y. Texture Features for Browsing and Retrieval of Image Data. **IEEE Transactions on Pattern Analysis and Machine Intelligence (Special Issue on Digital Libraries)**, [S.l.], v.18, n.8, p.837-842, 1996.
- [MES 99] MESSER, K.; KITTLER, J. A Region-Based Image Database System Using Colour and Texture. **Pattern Recognition Letters**, [S.l.], p.1323-1330, November 1999.
- [MUL 2000] MÜLLER, H.; MÜLLER, W.; MARCHAND-MAILLET, S.; PUN, T. Strategies for Positive and Negative Relevance Feedback in Image Retrieval. In: INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, 15., 2000. **Proceedings...** Barcelona, Spain: IEEE Computer, 2000. p.1043-1046.
- [NAS 99] NASTAR, C.; MEILHAC, C. Relevance Feedback and Category Search in Image Databases. In: INTERNATIONAL CONFERENCE ON MULTIMEDIA COMPUTING AND SYSTEMS, 3., 1999. **Proceedings...** Florence: IEEE Computer, 1999. p.512-517.
- [NAS 98] NASTAR, C.; MITSCHKE, M.; MEILHAC, C. Efficient Query Refinement for Image Retrieval. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, CVPR. **Proceedings...** Santa Barbara: IEEE Computer, 1998. p.547-552.
- [NGO 2001] NGO, C-W; PONG, T-C; CHIN, R. T. Exploiting Image Indexing Techniques in DCT Domain. **The Journal of the Pattern Recognition Society**, [S.l.], v.34, p.1841-1851, 2001.
- [NIK 2002] NIKOLAU, N.; PAPAMARKOS, N. Image Retrieval Using a Fractal Signature Extraction Technique. In: INTERNATIONAL CONFERENCE ON DIGITAL SIGNAL PROCESSING, 14., 2002. **Proceedings...** [S.l.:s.n.], 2002. v.2, p.1215-1218.

- [PEN 99] PENG, J.; BHANU, B.; QING, S. Probabilistic Feature Relevance Learning for Content-Based Image Retrieval. **Computer Vision and Image Understanding**, [S.l.], 1999. p.150-164.
- [PEN 96] PENTLAND, A.; PICARD, R.; SCLAROFF, S. Photobook: Content-Based Manipulation of Image Databases. **International Journal of Computer Vision**, [S.l.], v.8, n.3, p.233-254, June 1996.
- [PON 98] PONTE, I.; CROFT, W. A Language Modeling Approach to Information Retrieval. In: ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL, 21., 1998. **Proceedings...** [S.l.:s.n], 1998. p.275-281.
- [POR 99] PORKAEW, K. et al. Similarity Search Using Multiple Examples in Mars. In: INTERNATIONAL CONFERENCE ON VISUAL INFORMATION SYSTEMS. **Proceedings...** Amsterdam, Netherlands: Springer-Verlag, 1999. p.2-4.
- [RIC 94] RICHARDSON, R.; SMEATON, A. F.; MURPHY, J. **Using Wordnet as a Knowledge base for Measuring Semantic Similarity Between Words**. Dublin, Ireland: [s.n.], 1994. (Technical Report CA-1294).
- [ROC 71] ROCCHIO, J. J. Relevance Feedback in Information Retrieval. In: SALTON, G. **The SMART Retrieval System: Experiments in Automatic Document Processing**. Englewood Cliffs, New Jersey: Prentice Hall, 1971. p.313-323.
- [ROD 2000] RODDEN, K.; et al. Does Organization by Similarity Assist Image Browsing. In: CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS. **Proceedings...** Seattle, Washington: ACM Compute-Human Interaction, 2001. p.190-197.
- [RUI 99] RUI, Y.; HUANG, T. S.; CHANG, S. Image Retrieval: Current Techniques, Promising Directions, and Open Issues. **Visual Communication and Image Representation**, [S.l.], v.10, p.39-62, March 1999.
- [RUI 97] RUI, Y.; HUANG T. S.; MEHROTRA, S. Content-Based Image Retrieval with Relevance Feedback in Mars. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, ICIP, 1997. **Proceedings...** Santa Barbara, CA: IEEE Computer, 1977. p.815-818(II).
- [RUI 98] RUI, Y.; HUANG, G. S. T.; ORTEGA, M. Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval. **Transactions on Circuits and Systems for Video Technology**, [S.l.], v.8, n.5, p.644-655, September 1998.
- [SAL 90] SALTON, G.; BUCKLEY, C. **Improving Retrieval Performance by Relevance Feedback**. New York: Department of Computer Science, Cornell University; Ithaca: John Wiley & Sons, 1990.
- [SAL 87] SALTON, G.; MCGILL, M. J. **Introduction to Modern Information Retrieval**. [S.l.]: McGraw-Hill, 1987.
- [SHAW 95] SAHAW, W. M. Term-relevance Computations and Perfect Performance. **Information Processing and Management**, [S.l.], v.31, n.4, p.491-498, 1995.

- [SME 96] SMEATON, A. F.; QUIGLEY, J. Experiments on Using Semantic Distances Between Words in Image Caption Retrieval. In: CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL SIGIR, 19., 1996. **Proceedings...** Zurich, Switzerland: ACM SPECIAL INTEREST GROUP ON INFORMATION RETRIEVAL, 1996. p.174-180.
- [SMI 96] SMITH, J. R.; CHANG, Shih-Fu Chang. VisualSEEK: a Fully Automated Content-Based Image Query System. In: INTERNATIONAL MULTIMEDIA CONFERENCE AND EXHIBITION, 4., 1996. **Proceedings...** Boston, MA: ACM Multimedia Systems, 1996. p.87-98.
- [SRI 94] SRIHARI, R. K.; BURHANS, D. T. Visual Semantics: Extracting Visual Information from Text Accompanying Pictures. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, 12., 1994. **Proceedings...** Seattle, Washington: AAAI, 1994. p.793-798.
- [STR 95] STRICKER, M.; ORENKO, M. Similarity of Color Images. In: **Storage and Retrieval for Image and Video Databases** [S.l.]: SPIE Digital Library, 1995. p.381-392.
- [SWA 91] SWAIN, M. J.; BALLARD, D. H. Color Indexing. **International Journal of Computer Vision**, [S.l.], v.7, n.1, p.11-32, 1991.
- [WAN 2002] WANG, T.; et al. Adaptive Tree Similarity Learning for Image Retrieval. **ACM Multimedia Systems Journal**, [S.l.], v.9, n.2, 2003.
- [WOO 98] WOOD, M. E. J.; CAMPBELL, N. W.; THOMAS, B. T. Interactive Refinement by Relevance Feedback in Content-Based Digital Image Retrieval. In: INTERNATIONAL MULTIMEDIA CONFERENCE, 6., 1998. **Proceedings...** Bristol, UK: ACM Press, 1998. p.13-20.
- [YSH 83] HSU, Y. S.; et al. Pattern Recognition Experiments in the Mandala/Cosine Domain. **IEEE Trans. Pattern Anal.**, New York, v.5, n.5, p.512-520, March 1998.
- [ZHA 2002] ZHANG, A.; WU, Y. A Feature Re-Weighting Approach for Relevance Feedback in Image Retrieval. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING. **Proceedings...** [S.l.:s.n.], 2002. p.581-584.
- [ZHG 2002] ZHANG, C.; CHEN, T. An Active Learning Framework for Content-Based Information Retrieval. **IEEE Transactions on Multimedia**, [S.l.], v.4, n.2, p.260-268, June 2002.