

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**

**TÉCNICAS DE CLUSTERIZAÇÃO
BASEADAS EM CARACTERÍSTICAS
DE COR PARA A CONSULTA EM
BANCOS DE DADOS DE IMAGENS**

DISSERTAÇÃO DE MESTRADO

Juliano Gomes Weber

Santa Maria, RS, Brasil

2009

**TÉCNICAS DE CLUSTERIZAÇÃO BASEADAS EM
CARACTERÍSTICAS DE COR PARA A CONSULTA
EM BANCOS DE DADOS DE IMAGENS**

por

Juliano Gomes Weber

Dissertação apresentada ao Programa de Pós-Graduação em Informática
da Universidade Federal de Santa Maria (UFSM, RS), como requisito
parcial para a obtenção do grau de
Mestre em Computação

Orientador: Prof. Dr. Eng. Marcos Cordeiro d'Ornellas (UFSM)

**Dissertação de Mestrado N°
Santa Maria, RS, Brasil**

2009

**Universidade Federal de Santa Maria
Centro de Tecnologia
Programa de Pós-Graduação em Informática**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**TÉCNICAS DE CLUSTERIZAÇÃO BASEADAS EM
CARACTERÍSTICAS DE COR PARA A CONSULTA EM BANCOS
DE DADOS DE IMAGENS**

elaborada por
Juliano Gomes Weber

como requisito parcial para obtenção do grau de
Mestre em Computação

COMISSÃO EXAMINADORA:

Prof. Dr. Eng. Marcos Cordeiro d'Ornellas (UFSM)
(Presidente/Orientador)

Prof. Dr. Manoel de Araújo Sousa Júnior (INPE)

Prof. Dr. José Antônio Trindade Borges da Costa (UFSM)

Santa Maria, 29 de julho de 2009.

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter chegado até aqui, nunca tive dúvidas que nossa vida é direcionada por Ele. Em segundo lugar ao professor Marcos Cordeiro d'Ornellas pela confiança depositada neste aluno, pelo incentivo, paciência, generosidade e por ter proporcionado todas as condições possíveis para o desenvolvimento deste trabalho, além do brilhante direcionamento do mesmo.

Agradecimento especial também à minha esposa Raquel, nossa filha Morgana, que "entenderam" minha ausência neste período.

Também um agradecimento especial à meus amigos, familiares, colegas e ex-colegas, que sem dúvida foram fundamental importância no decorrer deste trabalho, se eu citasse nomes com certeza cometeria injustiças, deixo a todos vocês minha eterna gratidão.

Agradecimento ao Laboratório de Computação Aplicada (LaCA) do Centro de Tecnologia (CT) da Universidade Federal de Santa Maria (UFSM), aos colegas que estão vinculados a ele, pelo apoio em todos os sentidos e também pela infraestrutura disponibilizada para a elaboração deste trabalho.

É uma graça de Deus começar bem, mas a graça das graças é a possibilidade proporcionada por Ele de não desistir nunca. — AUTOR DESCONHECIDO

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Informática
Universidade Federal de Santa Maria

TÉCNICAS DE CLUSTERIZAÇÃO BASEADAS EM CARACTERÍSTICAS DE COR PARA A CONSULTA EM BANCOS DE DADOS DE IMAGENS

Autor: Juliano Gomes Weber

Orientador: Prof. Dr. Eng. Marcos Cordeiro d'Ornellas (UFSM)

Local e data da defesa: Santa Maria, 29 de julho de 2009.

As tecnologias atuais de aquisição, armazenamento e transmissão de dados digitais geram grandes quantidades de dados. Esse aumento quantitativo é diretamente proporcional à ampliação das bases de dados multimídia, onde se inserem as bases de imagens. Fatores relevantes que contribuem para esta ampliação são o acesso e a geração de dados multimídia, os quais são freqüentemente utilizados pela população através dos meios de comunicação atuais. Desta forma, percebe-se claramente a necessidade existente por sistemas automatizados, capazes de lidar com o armazenamento e a recuperação destes dados em um tempo aceitável para os padrões atuais. Para este fim, são desenvolvidos sistemas de recuperação de imagens por conteúdo, onde este conteúdo é descrito através de suas características visuais de baixo nível, como forma, textura e cor. Para que um sistema deste tipo seja considerado ideal, ele deve ser eficiente e eficaz. A eficácia será resultado da maneira de como foram obtidas as informações de baixo nível das imagens, considerando diferentes condições de foco, oclusão e iluminação. A eficiência é consequência dos resultados obtidos utilizando-se a organização das informações extraídas. Os métodos de agrupamento constituem em uma das técnicas úteis para diminuir a complexidade computacional destes sistemas, uma vez que agrupa informações com características semelhantes, sob determinado critério, porém sem perder a representatividade das informações extraídas.

Este trabalho propõe um método para recuperação de imagens baseada em conteúdo, que utiliza apropriadamente as técnicas de agrupamento, uma técnica de detecção de cantos e um método para normalizar as imagens no aspecto da iluminação, visando através disso obter descritores da imagem que sejam robustos e possam ser aplicados eficientemente em um sistema de recuperação de imagens por conteúdo - CBIR(*Content Based Image Retrieval*).

Palavras-chave: Características de cor, técnicas de agrupamento, técnicas de detecção de cantos, classificação de Imagens por conteúdo.

ABSTRACT

Master's Dissertation
Programa de Pós-Graduação em Informática
Universidade Federal de Santa Maria

TECHNIQUES OF CLUSTER-BASED FEATURES FOR CLASSIFICATION OF COLOR IMAGES

Author: Juliano Gomes Weber
Advisor: Prof. Dr. Eng. Marcos Cordeiro d'Ornellas (UFSM)

The current technologies for acquisition, storage and transmission of digital data, generate large amounts of data. This quantitative increase is directly proportional to the expansion of multimedia databases, where the bases are part of images. Factors contributing to this expansion is the generation of data access and multimedia, which are frequently used by the population through the media today. Thus, we find a clear need exists for automated systems, capable of dealing with the storage and retrieval of data in a time acceptable to the current standards. To this end, systems are designed for content retrieval of images, where the content is described through its low-level visual features such as shape, texture and color. To have such a system is considered ideal, it must be efficient and effective. The effectiveness will result from the way the information was obtained as a low level of images, considering different conditions of focus, lighting and occlusion. The efficiency is a consequence of the results obtained using the organization of information extracted. The methods of grouping are in one of the useful techniques to reduce the computational complexity of these systems, reducing the computational complexity of the methods implemented, but without losing the representation of information extracted.

This work proposes a method for retrieval of images based on content, using appropriate techniques of clustering, a technique for detecting edges and a method to normalize the images in the aspect of enlightenment, to get through it the image descriptors that are robust and can be applied efficiently in a retrieval system for images by content - CBIR (Content Based Image Retrieval).

Keywords: cluster-based, color features, content based image retrieval, corner detection.

LISTA DE TABELAS

Tabela 4.1 – Testes de processamento na fase de agrupamento	45
Tabela 4.2 – Instrução SQL responsável pela classificação das imagens por ordem de semelhança	50
Tabela 4.3 – Instrução SQL responsável pelo processo de recuperação das imagens	51
Tabela 4.4 – Categorias de imagens utilizadas nos testes computacionais	53
Tabela 4.5 – Resultados de todos os métodos envolvidos no comparativo	55

LISTA DE FIGURAS

Figura 2.1 – Arquitetura típica de um sistema de recuperação de imagens por conteúdo. Fonte:(TORRES e FALCÃO,2006)	18
Figura 2.2 – Abordagens em recuperação de imagens.....	19
Figura 2.3 – Imagem com a saturação original e seu respectivo histograma	21
Figura 2.4 – Imagem saturada e seu respectivo histograma.....	22
Figura 2.5 – Imagens com textura	23
Figura 3.1 – Idéia básica do detector de cantos Harris	30
Figura 3.2 – Classificação em função de R	32
Figura 3.3 – Imagem Original(A), valores de $R >$ limiar (B) e os máximos locais ou cantos (C)	33
Figura 3.4 – Etapas de uma tarefa de agrupamento	34
Figura 3.5 – Etapas de uma tarefa de agrupamento	34
Figura 3.6 – Tendência do algoritmo de deslocamento pela média.	38
Figura 4.1 – Arquitetura do método desenvolvido	42
Figura 4.2 – Fases do pré-processamento empregado pelo método desenvolvido ...	43
Figura 4.3 – - Interface desenvolvida em MATLAB utilizada na fase de pré-processamento do método desenvolvido	44
Figura 4.4 – Imagem original e agrupada com os cantos detectados	46
Figura 4.5 – Imagem com iluminação otimizada, agrupada com os cantos detectados	46
Figura 4.6 – Imagem original sua representação 3D no espaço de cores RGB	47
Figura 4.7 – Imagem agrupada por característica de cor e sua representação 3D no espaço de cores RGB	47
Figura 4.8 – Cantos detectados na imagem agrupada e apresentados na imagem original	48
Figura 4.9 – Procedimento de indexação das imagens	49
Figura 4.10 –Projeto físico da tabela MATCH_IMAGE	50
Figura 4.11 –Tela principal da Interface Web Desenvolvida.....	52
Figura 4.12 –Exemplo de retorno de consulta por semelhança	52
Figura 4.13 –Resultados referentes à complexidade computacional do método desenvolvido	54
Figura 4.14 –Representação gráfica dos resultados apresentados na tabela 4.5	56

LISTA DE ABREVIATURAS E SIGLAS

USAN	Univalue Segment Assimilating Nucleus
SQL	Structured Query Language
GUI	Graphical User Interface
SGBD	Sistema Gerenciador de Banco de Dados
CBIR	Content Based Image Retrieval
MAD	Mean Absolute Distance
DBMS	Database Management System
CSS	Curvature Scale Space

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO DA LITERATURA	14
2.1	Imagens e seus Metadados: Em Busca da Semântica	14
2.2	A Recuperação de Imagens por Conteúdo	15
2.2.1	Fusão de Descritores	15
2.2.2	Fusão de Evidências Textuais e Visuais	16
2.2.3	Realimentação de Relevância (<i>relevance feedback</i>)	16
2.2.4	A Importância da Recuperação de Imagens no Contexto Atual	17
2.3	Sistemas de Indexação e Recuperação de Imagens (CBIR)	17
2.3.1	Arquiteturas Existentes	18
2.3.2	Extração de Características	20
2.3.3	Trabalhos relacionados	24
2.3.4	Comparações e Critérios de Similaridade	25
3	TÉCNICAS ATUAIS DE GERAÇÃO DE METADADOS DA IMAGEM ...	27
3.1	A Detecção dos Cantos	27
3.1.1	Abordagens Baseadas em Arestas	28
3.1.2	Abordagens Baseadas na Topologia	28
3.1.3	Abordagens Baseadas em Auto-Correlações	28
3.1.4	Detectores de Cantos Alternativos	29
3.1.5	O Detector de Harris: Fundamentação e Algoritmo	29
3.2	O Agrupamento dos dados	33
3.2.1	Etapas do Processo de Agrupamento	34
3.2.2	Paradigmas de Agrupamento	35
3.2.3	Mean-Shift: Fundamentação e Algoritmo	36
3.3	A Detecção de Cantos e o Agrupamento na Construção de Metadados ...	40
4	MÉTODO DESENVOLVIDO	41
4.1	Arquitetura	41
4.1.1	Pré-Processamento	42
4.1.2	Correção da Iluminação	45
4.1.3	Agrupamento	47
4.1.4	Detecção de Cantos	48
4.1.5	Indexação de Imagens	48
4.1.6	Recuperação de Imagens	50

4.2	Uma Aplicação Cliente-Servidor para a Recuperação de Imagens por Conteúdo	51
4.3	Resultados e conclusões finais	53
4.3.1	Análise dos Resultados	53
5	COMENTÁRIOS FINAIS	57
5.0.2	Conclusões.....	57
5.0.3	Sugestões Para Trabalhos Futuros	58
	REFERÊNCIAS	59

1 INTRODUÇÃO

A proposta deste trabalho é criação um novo método que utiliza de maneira apropriada duas técnicas da área da Visão Computacional presentes no estado da arte da literatura, que originalmente possuem objetivos distintos, para que atuando em conjunto consigam gerar metadados robustos, que posteriormente são utilizados eficientemente em um Sistema de Recuperação de Imagens por Conteúdo - CBIR. A primeira técnica se refere à detecção de cantos proposta por (HARRIS; STEPHENS, 1988) e que está presente em recentes publicações á área da visão Computacional, como por exemplo, (GEVREKCI; GUNTURK, 2009), (WEIJER; GEVERS; GEUSEBROEK, 2005), e é apontada como a mais robusta técnica disponível com este objetivo específico (ZOU et al., 2008). A segunda é a técnica de agrupamento *mean-shift*, uma técnica também presente no estado da arte na área de visão computacional que pode ser observado em (ZHAO; KNIGHT; REID, 2008), (KALPATHY-CRAMER; HERSH, 2008), (SHOTTON; BLAKE; CIPOLLA, 2008), (VENUGOPAL; SUDHAMANI, 2008), (LI; SHIH; HSU, 2004), e que possui características primárias que atendem plenamente o problema a ser resolvido por este trabalho, que é a classificação de imagens coloridas heterogêneas através seu conteúdo de baixo nível.

2 REVISÃO DA LITERATURA

2.1 Imagens e seus Metadados: Em Busca da Semântica

O chamado *semantic gap* ou gargalo semântico é um dos principais fatores que dificultam a recuperação de imagens digitais, que é entendido como a diferença entre a percepção de alto nível que os usuários tem perante as imagens e as características de baixo nível que são extraídas das imagens para caracteriza-las, possibilitando sua recuperação automática através de um sistema CBIR(*Content Based Imagem Retrieval*). Este gargalo semântico ainda pode ser influenciado pela subjetividade humana, onde uma pessoa pode interpretar a mesma imagem de diversas maneiras e conforme o seu (des)conhecimento a priori.

O gargalo semântico ainda pode ser reduzido pela inclusão de informação de alto nível através do usuário em sistemas que já indexam informações de baixo nível, sendo considerada uma prática corrente. Desta maneira, a tarefa de comparação e recuperação de imagens considera não somente as características de baixo nível, extraídas automaticamente das imagens, como também as características de alto nível que são informadas pelo usuário. Estas características podem ser, por exemplo, palavras chave frequentemente usadas em sistemas de recuperação baseados na web ou descrições na forma de pequenas anotações incluídas em um banco de dados de imagens.

Semantic gap é um problema amplamente discutido na literatura, porém ainda é uma questão em aberto nas pesquisas em Visão Computacional. Recentemente duas técnicas baseadas em *relevance feedback* que foram integradas à um sistema CBIR, aumentando a precisão na recuperação das imagens em 42%, quando do uso das texturas da imagem (TRAINA; MARQUES; JR., 2006).

A grande maioria das pesquisas em sistemas de recuperação de imagens por conteúdo, utilizam características visuais de baixo nível, devido à boa eficiência computacional que

se consegue atingir através de métodos de recuperação que utilizam-se estas características.

2.2 A Recuperação de Imagens por Conteúdo

As imagens podem ser descritas através de atributos independentes do seu conteúdo visual, visando à indexação posterior. Dentre estes atributos podemos incluir o nome do arquivo, formato gráfico, tamanho físico, dimensões espaciais, etc. Muito embora estes atributos possam ser eficientemente gerenciados por um sistema gerenciador de banco de dados (SGBD), as consultas restringem-se somente aos atributos disponíveis, os quais não descrevem o conteúdo semântico das imagens.

A descrição textual pode ser eficientemente gerenciada por sistemas de recuperação de informação (LUX; GRANITZER; KLIEBER, 2004), porém, este tipo de técnica requer intervenção humana para obter as descrições textuais de cada imagem individualmente, onde problemas de subjetividade e incompletude podem vir a ocorrer na descrição das imagens, já que a interpretação visual da imagem varia de acordo com o conhecimento, objetivo, experiência e percepção de cada analisador.

Finalmente, é possível utilizar características visuais de baixo nível, tais como textura, formas, posição, distribuição de cores, para descrever, representar e comparar imagens. A Recuperação de imagens baseada em conteúdo é baseada no conceito de busca de imagens por similaridade, ou seja, quando um usuário submeter uma consulta (imagem exemplo), o sistema deve recuperar o conjunto de imagens mais similares à imagem de referência, organizadas em ordem decrescente de similaridade. CBIR é uma área multidisciplinar que envolve, principalmente, técnicas de banco de dados, processamento de imagens, recuperação de informação e reconhecimento de padrões, incluindo também interfaces usuário-máquina.

2.2.1 Fusão de Descritores

Neste tipo de solução, a idéia é descobrir mecanismos para explorar as vantagens de diferentes descritores encontrados na literatura e combiná-los em uma aplicação específica. O problema é que de fato, não existem descritores que apresentam alto desempenho (eficácia e eficiência) em todos os sistemas nos quais são utilizados.

2.2.2 Fusão de Evidências Textuais e Visuais

Esta solução é baseada em anotações textuais, onde as palavras-chaves são associadas às imagens armazenadas na base de dados e as consultas são feitas de forma textual, trazendo problemas de interpretação por parte do usuário e elevado tempo de trabalho. Porém, essa abordagem introduz, de fato, informação semântica ao processo de recuperação de imagens. A idéia é utilizar anotações textuais em conjunto com o conteúdo de baixo nível das imagens, aproveitando assim as qualidades implícitas em ambas as abordagens.

2.2.3 Realimentação de Relevância (*relevance feedback*)

A forma mais usual para fazer uma consulta em um sistema CBIR é informar uma imagem de entrada, e o sistema retornar as imagens mais semelhantes em ordem decrescente (*rank*), baseando-se nos descritores utilizados pelo sistema. Uma tarefa não trivial é permitir que o usuário forneça parâmetros adicionais ao sistema de busca, uma vez que exigiria do mesmo um conhecimento bastante detalhado das características dos descritores utilizados. Para resolver este problema, o *relevance feedback* é usada a realimentação de relevância, que é uma técnica utilizada para recuperação de informações textuais (XU; XIAOWEI; TRESP, 2003). O *relevance feedback* é usado por pesquisadores atuantes na área de recuperação de imagens por conteúdo (ZHOU; HUANG, 2003) e (KIM; CHUNG; BARNARD, 2005). A idéia é permitir que o usuário expresse sua necessidade durante a especificação da consulta, sem a necessidade de preocupar-se com propriedades de baixo nível. As características de alto nível, mais utilizadas na representação de imagens, permitem que o usuário indique, por exemplo, as imagens relevantes dentro das imagens retornadas. A cada consulta, o algoritmo de classificação é realimentado com outros dados provenientes de consultas anteriores produzidas pelo usuário. As propriedades visuais que melhor definem as imagens relevantes são usadas, permitindo que o sistema produza novas consultas a partir da consulta original, produzindo assim, resultados similares, conforme o critério escolhido.

O objetivo do *relevance feedback* é tratar duas questões importantes da recuperação das imagens por conteúdo que são: evitar que o usuário tenha necessidade de conhecer as características de baixo nível utilizada no processo de classificação, necessitando apenas a percepção visual do usuário e atender o caráter subjetivo de diferentes pessoas em diferentes circunstâncias. Com o uso de *relevance feedback* estas duas questões passam

despercebidas pelo usuário.

2.2.4 A Importância da Recuperação de Imagens no Contexto Atual

Os preços cada vez mais acessíveis de equipamentos que permitem a geração e armazenamento de imagens digitais, têm tornado as pesquisas em recuperação de imagens por conteúdo cada vez mais relevantes em nossos dias, devido à sistemas eficazes para gerenciar toda esta informação gerada. Aliado a esses fatores, o crescimento e popularização da Internet torna a disponibilização de imagens digitais bastante abundante. Essas grandes coleções de imagens estão disponíveis nas mais diversas áreas do conhecimento, como: imagens médicas, presentes no trabalho de (ANTANI; LONG; THOMA, 2008) e (CHANG et al., 2008), sensoriamento remoto, conforme (TOBIN et al., 2006), (LIU; HE, 2008) e (XIE; ROBERTS; JOHNSON, 2008). A popularização do acesso a imagens digitais tem permitido que esta área de pesquisa busque soluções específicas para a recuperação de imagens em setores da sociedade não previstos em primeira instância, como por exemplo organização de imagens digitais pessoais (WEN; CHEN; TANG, 2008). Em 2006 algumas estimativas que demonstraram alguns números impressionantes para a produção audiovisual nos anos vindouros. Em todo o mundo, 1-2 hexabytes (bilhões de gigabytes) de conteúdo eletrônico seriam produzidos e 80 bilhões de imagens digitais seriam produzidas. Além disso, mais de um bilhão de imagens relacionadas a transações comerciais estariam disponíveis e deveriam aumentar dez vezes nos dois anos seguintes (KOMPATSIARIS, 2006).

2.3 Sistemas de Indexação e Recuperação de Imagens (CBIR)

Sistemas de recuperação de imagens por conteúdo são focados na idéia de similaridade de imagens, onde, dado um banco de dados de imagens com grande número de imagens, o usuário necessita recuperar imagens similares a uma determinada imagem exemplo ou um esboço de imagem fornecido como entrada. O processo de recuperação é baseado na comparação das imagens por meio de descritores (TORRES; FALCÃO, 2006), que devem armazenar de maneira eficaz determinadas características extraídas das imagens. A arquitetura típica de um sistema de recuperação de imagens por conteúdo é demonstrada na Figura 2.1. A arquitetura básica de um sistema de recuperação de imagens por conteúdo possui duas funcionalidades principais: a inserção de dados e o processamento de consultas.

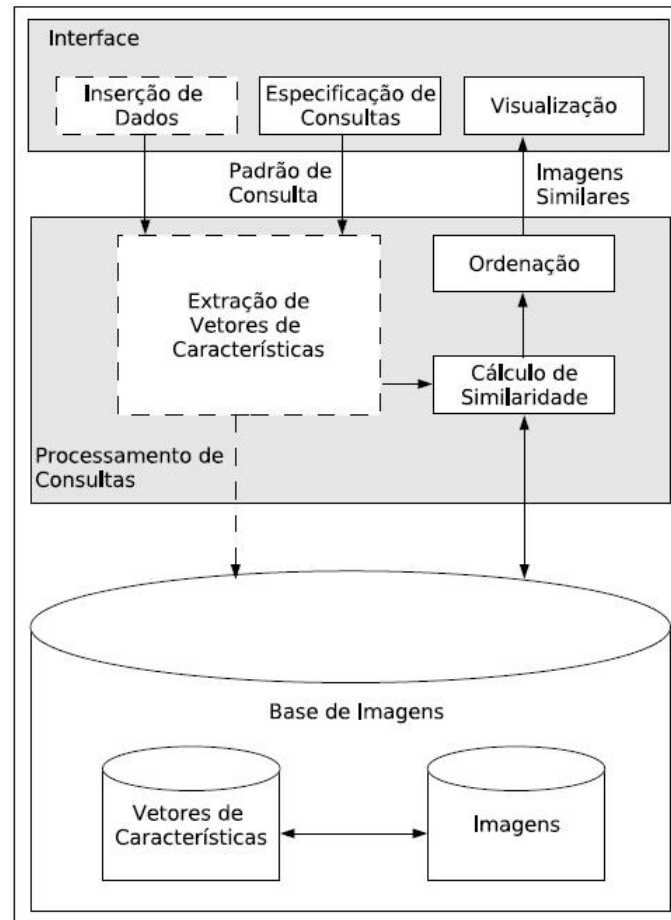


Figura 2.1: Arquitetura típica de um sistema de recuperação de imagens por conteúdo. Fonte:(TORRES e FALCÃO,2006)

2.3.1 Arquiteturas Existentes

Os sistemas CBIR possuem atualmente duas abordagens principais. Uma delas é baseada no conteúdo da imagem e a outra é baseada em texto. Estas abordagens podem ser divididas em 4 categorias, que são baseadas em atributo, anotação, objetos e características. A Figura 2.2 mostra a distribuição das abordagens.

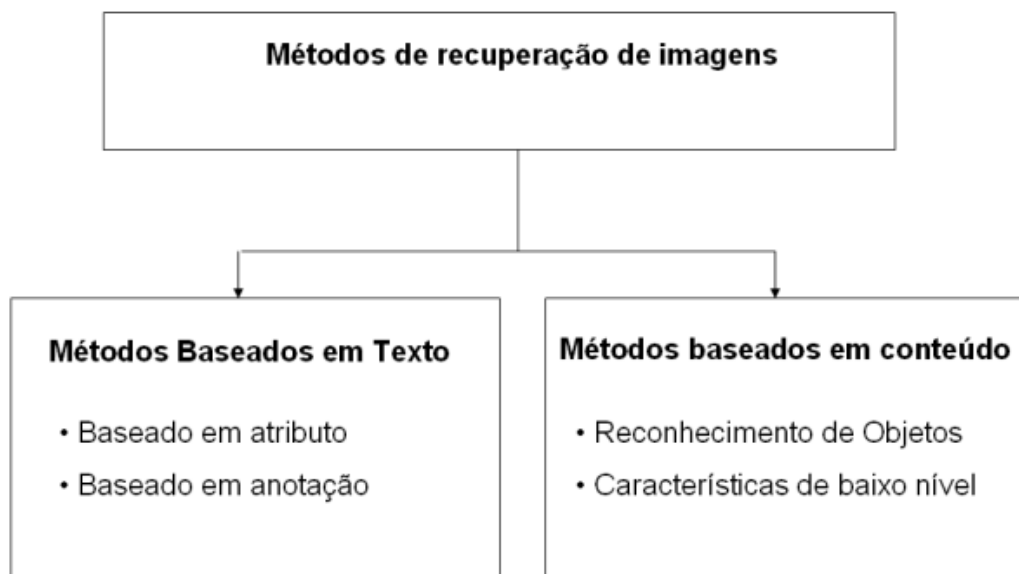


Figura 2.2: Abordagens em recuperação de imagens

A abordagem baseada em atributo representa o conteúdo da imagem como um conjunto de características extraídas manualmente da mesma imagem, como assunto, categoria, autor, origem, data de criação, etc. Esta abordagem faz uso de métodos de indexação oferecidos por sistemas de gerenciamento de bancos de dados, e a SQL (*Structured Query Language*). Devido à subjetividade de quem está extraindo as informações, um dos maiores problemas desta abordagem é que elas nem sempre poderão descrever o conteúdo da imagem de maneira apropriada.

A abordagem baseada em anotação utiliza-se de um texto que descreve o conteúdo da imagem. A recuperação destas imagens é efetuada através de palavras-chaves ou frases, onde são frequentemente utilizados operadores booleanos. Esta abordagem traz como principal vantagem a possibilidade de abstração de uma imagem, porém com a desvantagem de exigir que a anotação seja feita de maneira manual, também trazendo consigo o problema da subjetividade. Reconhecimento de objetos é outra abordagem trabalhada, onde o objetivo é interpretar cenas, pessoas, ações e objetos presentes na imagem.

Por fim, a abordagem que parte das características de baixo nível (cor, textura e forma) realiza de forma automática a indexação das imagens. Este procedimento automático reduz consideravelmente o tempo necessário para indexar uma coleção de imagens, tornando-se assim bastante importante na recuperação em grandes coleções de imagens.

2.3.2 Extração de Características

(AMSALEG; GROS; BERRANI, 2004) afirmam que as imagens são consideradas como dados complexos, pois suas características são representadas por um número usualmente elevado de atributos. Para tratar esse tipo de dado é necessário extrair informações que melhor o represente. Esse processo, denominado extração de características, permite identificar as propriedades inerentes da imagem que serão utilizadas para recuperar imagens de maneira eficiente. As características da imagem podem ser consideradas como globais e locais. Em um processo de busca por imagens similares são retornadas as imagens mais semelhantes globalmente, obtidas por meio de um descritor que codifica as informações globais da imagem. As técnicas de extração de características globais não se preocupam com detalhes de regiões individuais da imagem, podendo ocasionar o aparecimento de informações irrelevantes no resultado de uma consulta.

Contudo, não é possível detectar quando duas imagens contêm objetos similares em locais diferentes, seja em segundo plano, em diferentes pontos de vista ou diferentes variações de iluminação. Por não conseguir capturar aspectos de semântica mais local da imagem, a extração de características pode produzir, assim, uma descontinuidade semântica. Para resolver esse problema, técnicas de processamento de imagens têm focado no reconhecimento de objetos, tipicamente para extrair informações locais contidas em pequenas regiões das imagens. Nas técnicas de extração de características locais, as imagens são segmentadas em regiões menores que podem ser utilizadas para identificar objetos com um alto nível semântico, podendo obter um reconhecimento robusto em relação à rotação e translação da imagem, variação de iluminação e oclusões parciais. Algumas características da imagem podem ser detectadas automaticamente, porém sua precisão depende da complexidade da imagem. As características extraídas automaticamente da imagem são categorizadas como baixo nível (*raw features*) e são baseadas, principalmente, nos atributos de cor, textura e forma. A comparação de imagens efetuadas em sistemas CBIR busca simular a percepção humana comparando as características visuais de cor, textura e forma (KOSARA; MIKSCH, 2002). Apesar de a cor ser um atributo bastante utilizado e eficaz na recuperação de imagens, dependendo da aplicação e do domínio específico, essa característica pode não ser a mais adequada, sendo necessário o uso dos atributos de textura e forma. Deve-se observar que, muitas vezes é necessário o uso de múltiplos atributos para alcançar uma resposta desejada (MIRMEHDI; PERISSAMY, 2002).

2.3.2.1 Usando Cores

As cores presentes em uma imagem possuem um papel significativo na indexação e recuperação. Existem diferentes representações de cores que incluem desde o tradicional modelo RGB (red, green, blue), o mais simples modelo que mapeia diretamente as características físicas do dispositivo de exibição, até o HSI (*hue, saturation, intensity*) que reflete mais precisamente o modelo de cores para a percepção humana.

Muitos trabalhos de extração de características baseadas em distribuição de cores estão concentrados nos histogramas de cor, onde a intersecção de histogramas é utilizada (SWAIN; BALLARD, 1990). Os histogramas de cor são invariantes à translação e rotação das imagens, sendo que, com a normalização dos histogramas, obtém-se também a invariância à escala. De qualquer forma, os histogramas de cor não indicam a localização espacial dos pixels na imagem.

Muitas operações podem ser feitas com os histogramas de cor, mudando parâmetros e limiares, por exemplo, a Figura 2.3 apresenta a imagem original e seu respectivo histograma. Ao se aplicar uma operação de alongamento (*stretch*) sobre a imagem, o histograma ficaria como na Figura 2.4, tendo como resultado uma imagem melhor definida.

Alguns trabalhos, que utilizam a distribuição de cores na imagem para montar o vetor de características baseado nos coeficientes wavelets são apresentados por (ALBUZ; KOCCALAR; KHOKHAR, 2001), conseguindo diminuir o tempo de busca das imagens similares, mas a eficiência da busca não é tão boa (máximo de 35% de eficiência).

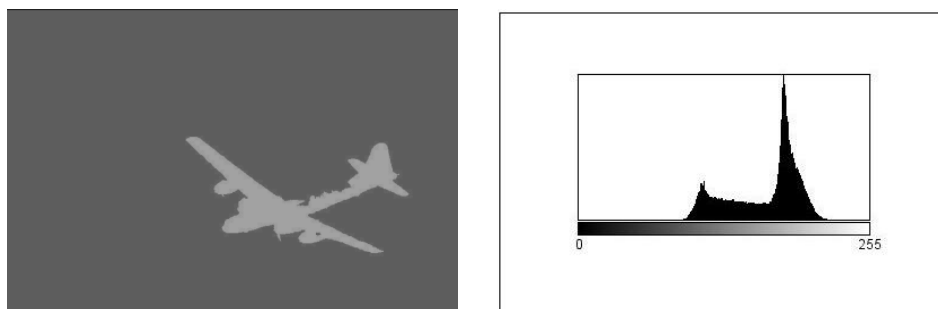


Figura 2.3: Imagem com a saturação original e seu respectivo histograma

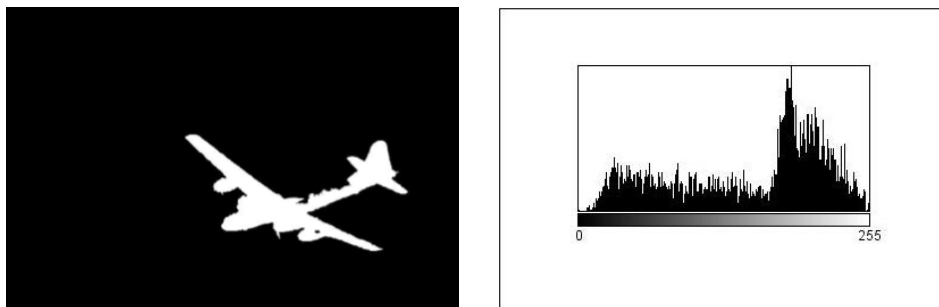


Figura 2.4: Imagem saturada e seu respectivo histograma

Embora existam sistemas comerciais muito populares que utilizam os histogramas de cores em sistemas de recuperação de imagens por conteúdo como por exemplo: (FLICKNER et al., 1995a), (SMITH; CHANG, 1996), (MA; MANJUNATH, 1999a), esta abordagem mostra algumas dificuldades pelo próprio caráter não invariante do histograma de cores. Além disso, se a imagem contém um número alto de cores, então o vetor de características será muito grande e a indexação de vetores com essa dimensão torna-se problemática

2.3.2.2 *Usando Formas*

Deve-se indicar que em sistemas de recuperação de imagens por conteúdo, o atributo forma é o que apresenta maior dificuldade, especialmente pelo fato de ter que segmentar e conhecer o tamanho dos objetos contidos na imagem. Vários esquemas foram propostos na literatura para a recuperação de imagens baseadas na forma. (GONZALES; WOODS R., 1993) propôs uma técnica hierárquica que mistura métodos de momentos invariantes com arco poligonal 2D. O sistema IRUS proposto por (ADORAM; LEW, 1999) utiliza contornos ativos para a segmentação de objetos. Uma análise multiescala de objetos bidimensionais aplicando Gabor wavelets é apresentada em (CESAR-JR., 1997).

Atualmente, os sistemas CBIR que são baseados em formas, procuram explorar o contorno e também o conteúdo do interior destes objetos (ZHANG; LU, 2004), mas estas definições dependem muito do domínio das imagens em questão. (TORRES et al., 2008) afirma que é um desafio encontrar uma descrição ideal para cada tipo de aplicação, considerando informações de contorno, interior ou uma combinação dos métodos para obter uma recuperação de imagens mais eficiente.

2.3.2.3 Usando Texturas

Não existe uma definição clara de textura (TUCERYAN; JAIN, 1993), mas os autores concordam em defini-la como as mudanças na intensidade da imagem que formam determinados padrões repetitivos. Esses padrões podem ser o resultado de propriedades físicas da superfície do objeto (rugosidade), ou ser o resultado de diferenças de reflexão tal como a cor na superfície. A Figura 2.5 mostra alguns exemplos de imagens com textura. Muito embora o reconhecimento de objetos através de suas respectivas texturas seja simples para os seres humanos, esta tarefa é considerada complexa para os sistemas computacionais. Nos sistemas CBIR as características de textura são muito úteis, por esta estar presente na maioria de superfícies naturais exibe texturas. A extração de metadados a partir das texturas vem despertando interesse na comunidade científica, tendo em vista que os descritores de textura podem medir algumas propriedades das imagens como regularidade, suavidade, granularidade, entre outras (TORRES et al., 2008). Aplicações desenvolvidas incluindo a recuperação de imagens por conteúdo, considerando a textura pode ser observados em (LIU et al., 2007).



Figura 2.5: Imagens com textura

A análise de textura tem por objetivo estabelecer o relacionamento de vizinhança dos elementos e seu posicionamento em relação aos demais (conectividade), o número de elementos por unidade espacial (densidade) e a sua regularidade (homogeneidade). A identificação das texturas de imagens visa permitir sua segmentação em determinadas regiões, que possuam a mesma textura, porém da mesma forma que no caso da extração de características por cores, apresentam-se dificuldades com a invariância e a dimensionalidade do vetor de características.

2.3.2.4 Abordagens Híbridadas

É possível imaginar que a combinação de vários descritores de imagens possam ser utilizados em conjunto para caracterizar determinada imagem em um sistema CBIR. Porém, o fato é que cada tipo de descritor tem um comportamento ótimo em aplicações específicas. É um desafio da área de Visão Computacional descobrir quais os descritores que produzem uma determinada combinação entre si, para recuperar imagens de maneira automática eficiente. (TORRES et al., 2008) afirma que não existem descritores que apresentam alto desempenho (eficácia e eficiência) em todas as aplicações, sendo este um dos desafios a ser explorado na pesquisa de sistemas CBIR (LEW et al., 2006).

2.3.3 Trabalhos relacionados

Os sistemas de recuperação de imagens por conteúdo primordiais datam da década de 80 (CHANG; FU, 1980). Os mais importantes começaram a surgir durante a década de 90. Dentre os sistemas comerciais de recuperação de imagens baseados em conteúdo mais conhecidos são o QBIC da IBM (FLICKNER et al., 1995b), que foi pioneiro na área e é um dos mais importantes até hoje e o *Virage* que foi desenvolvido para recuperação de vídeos e imagens. Após o QBIC, outros sistemas foram desenvolvidos, como o *Netra* (MA; MANJUNATH, 1999b) e o *Photobook* (PENTLAND; PICARD; SCLAROFF, 1996), onde foram sendo agregados novos recursos que incluíram características de cor e textura para descrever seu conteúdo. O *Blobworld* (CARSON et al., 1999) introduziu o uso de segmentação de imagens e o *PicHunter* implementou um mecanismo de realimentação de relevância (*relevance feedback*). A abordagem proposta no sistema MARS (RUI; HUANG; MEHROTRA, 1997) faz uso de uma solução de realimentação de relevância e tem destaque perante os demais sistemas pelo envolvimento de múltiplas pesquisas em sua implementação, que incluem visualização computacional, recuperação de informação e sistemas de gerenciamento de base de dados.

Encontram-se disponíveis via Web, sistemas de domínio público como o GIFT (*GNU Image-Finding Tool*) e várias versões de demonstrações disponíveis on-line, como o *Viper* e o *Compass* (*Computer Aided Search System*). Uma descrição mais apurada desses e de outros sistemas CBIR pode ser vista em (VELTKAMP; TANASE, 2002). Tendências das pesquisas relacionadas a sistemas CBIR pode ser visto em: (DATTA et al., 2008).

Recentemente, (VENUGOPAL; SUDHAMANI, 2008) utilizou uma técnica de agrupa-

mento de dados em conjunto com uma técnica de indexação R* Tree em um sistema CBIR. Os resultados foram utilizados como parâmetro de comparação para atestar a eficiência do método desenvolvido e apresentado neste trabalho.

2.3.4 Comparações e Critérios de Similaridade

Para concluir o quanto uma imagem é similar à outra, após a escolha de uma imagem de comparação, é necessário avaliar a melhor função de distância para o cálculo de similaridade entre os metadados coletados das imagens, o que a literatura constantemente chama de vetores de características. Tradicionalmente, sistemas CBIR utilizam medidas Euclidianas, para medir distâncias entre a imagem consultada (suas características extraídas) e as características das outras n imagens já representadas também por suas características. A função de distância é utilizada de acordo com a estrutura e conteúdo dos metadados, podendo ser necessário utilizar mais de uma função de distância. Esta escolha irá depender das especificidades do conjunto de imagens que está sendo usado e os seus respectivos metadados. Algumas das medidas de similaridade mais utilizadas, conhecidas e encontradas na literatura, estão descritas a seguir:

2.3.4.1 Distância de Minkowski

Esta técnica trata-se da generalização da distância entre 2 pontos $X = (x_1, x_2, x_3, \dots, x_n)$ e $Y = (y_1, y_2, y_3, \dots, y_n)$ em um plano. Está descrita em (PLATANIOTIS; ANDROUTSOS; VENETSANOPOULOS, 1998) para recuperação de imagens coloridas baseadas em contexto. Esta métrica é conhecida também como distância L_p . Na Equação (2.1) pode-se observar a sua fórmula matemática, onde p representa a norma:

$$D_{Minkowski} = \sqrt[p]{\sum_{k=0}^n |x_k - y_k|^p} \quad (2.1)$$

2.3.4.2 Distância Euclidiana

Trata-se de outra especialização da distância de Minkowski para norma-2, ou L_2 . É a distância linear entre os pontos $X = (x_1, x_2, x_3, \dots, x_n)$ e $Y = (y_1, y_2, y_3, \dots, y_n)$, com sua fórmula definida pela Equação (2.2) para n dimensões. Pode-se perceber que a fórmula da distância euclidiana para $n = 2$ é equivalente a fórmula da Equação (2.2), que representa a forma de um círculo.

$$D_E = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \dots + (x_n - y_n)^2} \quad (2.2)$$

2.3.4.3 Distância de Mahalanobis

Distância não linear entre pontos $X = (x_1, x_2, x_3, \dots, x_n)$ e suas médias

$$X = (x_1, x_2, x_3, \dots, x_n)$$

. Sua fórmula matemática é representada pela Equação (2.3) em que é calculada a distância do ponto X até o centro do grupo X, C representando a matriz inversa da co-variância entre os pontos. Nessa distância é necessário que já haja um agrupamento prévio, pois a métrica a ser feita nos pontos leva em consideração os pontos já pertencentes a esse grupo.

$$D_M = \sqrt{((X - \bar{X})^T * C^{-1}(X - \bar{X}))} \quad (2.3)$$

A obtenção de C é dada pela Equação (2.4) em que a função E(z) representa o valor esperado. Nota-se que na diagonal principal dessa matriz encontram-se as variâncias e que também essa matriz tem a propriedade de ser espelhada

$$\begin{bmatrix} E[(x_1 - \bar{x}_1)(x_1 - \bar{x}_1)] & E[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)] & \dots & E[(x_1 - \bar{x}_1)(x_n - \bar{x}_n)] \\ E[(x_2 - \bar{x}_2)(x_1 - \bar{x}_1)] & E[(x_2 - \bar{x}_2)(x_2 - \bar{x}_2)] & \dots & E[(x_2 - \bar{x}_2)(x_n - \bar{x}_n)] \\ & & \vdots & \\ E[(x_n - \bar{x}_n)(x_1 - \bar{x}_1)] & E[(x_n - \bar{x}_n)(x_2 - \bar{x}_2)] & \dots & E[(x_n - \bar{x}_n)(x_n - \bar{x}_n)] \end{bmatrix} \quad (2.4)$$

3 TÉCNICAS ATUAIS DE GERAÇÃO DE METADADOS DA IMAGEM

3.1 A Detecção dos Cantos

A extração de determinados pontos de interesse da imagem durante o processamento de imagens digitais é base para diversas áreas do conhecimento. O interesse na detecção de pontos (cantos ou quinas) reside no fato destes pontos possuírem características que variam fortemente em relação aos vizinhos e que os distinguem. Tais características normalmente estão relacionadas ao brilho e intensidade, a análise matemática dessas características permite o cálculo da posição destes pontos na imagem. Por outro lado os cantos têm grande importância em processamento digital de imagens por possuírem características de não serem afetados por rotações da imagem e serem detectados sob diferentes condições de iluminação. É desejável que um detector de cantos atenda aos seguintes requisitos (WANG W., 2004)

- Detecte todos os cantos que realmente existam na imagem;
- Não detecte falsos cantos;
- A localização do canto deve ser a mais fiel possível;
- O detector de cantos deve ser capaz de manter sua eficiência sob diferentes circunstâncias de aplicação;
- O detector deve ser robusto com respeito a ruídos;
- O detector deve ser computacionalmente eficiente.

Desde o desenvolvimento do primeiro detector de cantos por (MORAVEC, 1977), várias outras tentativas têm sido empregadas por diversos autores na busca de um detector

de cantos ideal. Este trabalho divide os métodos de detecção de cantos em quatro grupos: o primeiro grupo diz respeito aos métodos relacionados a arestas, o segundo refere-se aos métodos baseados em topologia, o terceiro aos métodos baseados em autocorrelação e o quarto aos métodos alternativos.

3.1.1 Abordagens Baseadas em Arestas

(KITCHEN; ROSENFELD, 1982) foram os pioneiros na aplicação de operadores de gradiente na detecção de cantos. Eles utilizaram o Laplaciano da função gaussiana para a identificação de locais de máximo sobre bordas detectadas. Os cantos eram identificados com a aplicação da supressão não máxima. Esta metodologia apresenta como fragilidade a alta sensibilidade à presença de ruídos.

3.1.2 Abordagens Baseadas na Topologia

(BEAUDET, 1978) desenvolveu um operador invariante à rotação. A detecção do canto é baseada no cálculo do determinante da matriz Hessiana, a qual inclui derivadas de segunda ordem tornando o método sensível a ruídos e a variação de escala. (DERICHE; GILAUDON, 1992) estenderam o método proposto por (BEAUDET, 1978) pela aplicação de múltiplas escalas na suavização gaussiana. Os cantos são definidos nos pontos de cruzamento por zero, nas diferentes escalas aplicadas. Embora a metodologia apresente ganho quanto à posição relativa dos cantos detectados permanece ainda a alta sensibilidade aos ruídos da imagem.

3.1.3 Abordagens Baseadas em Auto-Correlações

A metodologia desenvolvida em (MORAVEC, 1977) introduz o conceito de ponto de interesse ou pontos que representem um máximo local para a função da imagem considerada. Uma máscara de operador de gradiente centrada nestes pontos de interesse é aplicada para medir a variação da intensidade de brilho em diferentes direções. Moravec implementou esta metodologia através do cálculo de autocorrelação local não normalizada em 4 e 8 diferentes direções (vizinhança 4 e 8), resultando numa resposta anisotrópica altamente sensível a ruídos devido ao uso de máscaras com dimensões entre 3x3 e 5x5. (HARRIS; STEPHENS, 1988) contribuíram para a extensão e aperfeiçoamento do método de Moravec, principalmente com vistas à utilização de um operador gaussiano e aumento da vizinhança de pesquisa de variação da função da imagem. A autocorrelação

local é calculada a partir de derivadas de primeira ordem (gaussiano), produzindo uma resposta isotrópica. A partir do trabalho de (FORSTNER, 1986), um importante melhoramento foi incorporado, relacionado ao uso de máscaras circulares e a decomposição da matriz de autovalores.

(ZHENG; WANG; KHWANG, 1999), motivados por reduzir a complexidade computacional do método de Harris e Stephens, introduziram simplificações no algoritmo do referido método cujas modificações implicaram em otimização computacional e de localização dos cantos, porém tiveram como consequência a diminuição da capacidade de detecção de cantos. (TRIGGS, 2004) denomina o operador resultante dos desenvolvimentos e aprimoramentos desempenhados independentemente por (FORSTNER, 1986) e (HARRIS; STEPHENS, 1988) quanto ao operador de Moravec como operador Forstner-Harris.

3.1.4 Detectores de Cantos Alternativos

(SMITH; BRADY, 1995) desenvolveram o detector de cantos e linhas denominado SUSAN (*Smallest Univalve Segment Assimilating Nucleus*). O algoritmo assume que dentro de uma pequena máscara circular a intensidade de brilho pertencente a diferentes objetos varie pouco, o mesmo calcula também o número de pixels que possuem brilho similar ao brilho do pixel central ou núcleo da máscara. Estes pixels são chamados USAN (*Univalve Segment Assimilating Nucleus*), os cantos são detectados aplicando-se o procedimento descrito a toda a imagem proporcionando o cálculo da mínima USAN. O método possui robustez quanto a ruídos (por não utilizar derivadas parciais), é ágil computacionalmente, porém, sua repetibilidade é baixa. (TRAJKOVIC; HEDLEY, 1998) desenvolveram uma metodologia para detectar cantos a partir do algoritmo SUSAN, porém, assumindo que um canto seja uma variação abrupta de intensidade de brilho em todas as direções. Este método apresenta alta eficiência computacional, entretanto não apresenta adaptabilidade a variações na imagem comparável ao detector de cantos de Harris. O operador CSS (*Curvature Scale Space*) detecta cantos através da busca direta de pontos de máxima curvatura sobre uma linha de contorno, ou seja, busca sobre a linha de contorno as máximas variações de direção (MOKHTARIAN; SUOMELA, 1998).

3.1.5 O Detector de Harris: Fundamentação e Algoritmo

A idéia básica do detector de cantos de Harris (HARRIS; STEPHENS, 1988) consiste em definir um pequeno quadrado dentro da imagem e deslocar em várias direções. Estas

variações de deslocamento provocam pequenas variações de intensidade neste quadro. Dependendo do tipo de variação de intensidade, este quadro contém um canto, uma borda ou mesmo uma região plana. A Figura 3.1 mostra esta característica.

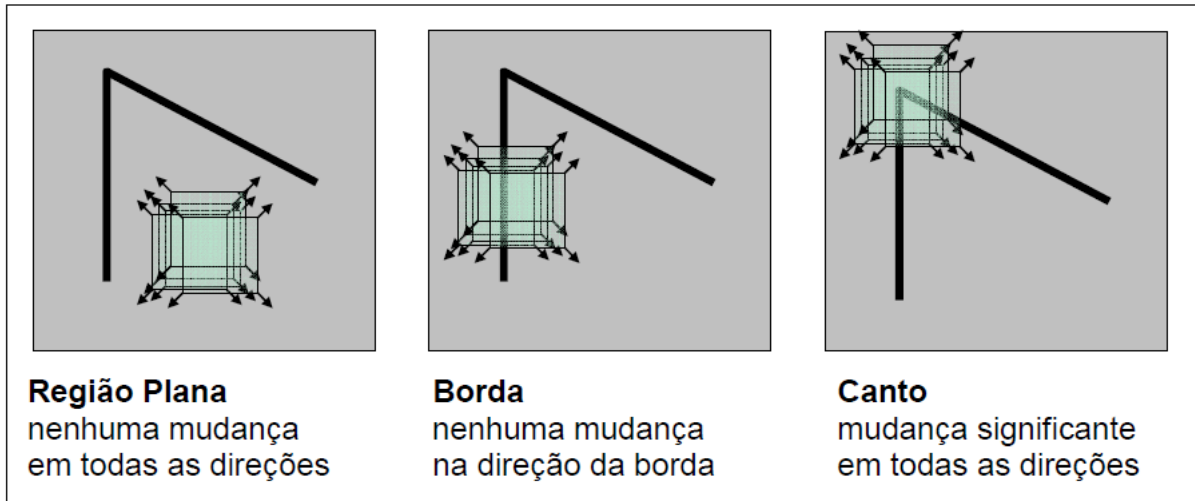


Figura 3.1: Idéia básica do detector de cantos Harris

A especificação matemática é demonstrada como:

$$E(x, y) = \sum w(u, v) \cdot |I_{x+u, y+v} - I_{u, v}|^2 \quad (3.1)$$

onde I é a intensidade da imagem, E a variação de intensidade introduzida pelo pequeno deslocamento da imagem e w especifica o tipo de janela do quadro, sendo a unidade na região retangular especificada e zero nas demais regiões.

Se a janela considerada for pequena (2x2 pixels), então é possível simplificar por uma aproximação bilinear

A especificação matemática é demonstrada como:

$$E(u, v) = [u, v] \cdot M \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.2)$$

onde M é uma matriz 2x2 computada a partir das derivadas da imagem

$$M = \sum W(x, y) \cdot \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} \\ \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} & \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix} \quad (3.3)$$

sendo $I(x, y)$ a intensidade da imagem considerada.

O cálculo das derivadas parciais da imagem é feito da seguinte forma

$$\frac{\partial I}{\partial x} = I(x, y) \otimes (-1, 0, 1) \quad (3.4)$$

$$\frac{\partial I}{\partial y} = I(x, y) \otimes (-1, 0, 1)^T \quad (3.5)$$

ou seja, a convolução com um filtro derivativo horizontal e vertical. Harris ainda propõe que

$$\left(\frac{\partial I}{\partial x}\right)^2, \left(\frac{\partial I}{\partial y}\right)^2 \quad (3.6)$$

e

$$\frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} \quad (3.7)$$

sejam filtrados utilizando uma janela circular gaussiana para diminuir o ruído causado pela janela binário e retangular das derivadas. Segundo o autor, os autovalores (λ_1 e λ_2) da matriz M , representam a característica da região analisada. quando os dois autovalores são pequenos, temos uma região "plana". Se um dos autovalores for maior que o outro, então temos uma borda da imagem. Se os dois autovalores forem grandes e próximos temos um canto da imagem.

Harris propôs a utilização de uma função chamada de *resposta aos cantos* para evitar a decomposição dos autovalores da matriz M , definida como

$$R = \det M - k \cdot (\text{trace}\{M\})^2 \quad (3.8)$$

$$\det M = \lambda_1 \cdot \lambda_2 = \left(\frac{\partial I}{\partial x}\right)^2 \cdot \left(\frac{\partial I}{\partial y}\right)^2 - \frac{\partial I}{\partial x} \cdot \frac{\partial I}{\partial y} \quad (3.9)$$

$$\text{trace}\{M\} = \lambda_1 + \lambda_2 = \left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 \quad (3.10)$$

onde k é uma constante empírica de valores típicos proposto por Harris entre 0,04 e 0,06. Então, dependendo do valor de R temos a classificação da região da imagem em região plana, canto ou borda.

Se a resposta da função for um valor pequeno, será classificada como uma região plana; Se a resposta for um valor grande mas negativo, então será classificada como borda.

E se a resposta for um valor grande mas positivo, será classificada como um ponto característico.

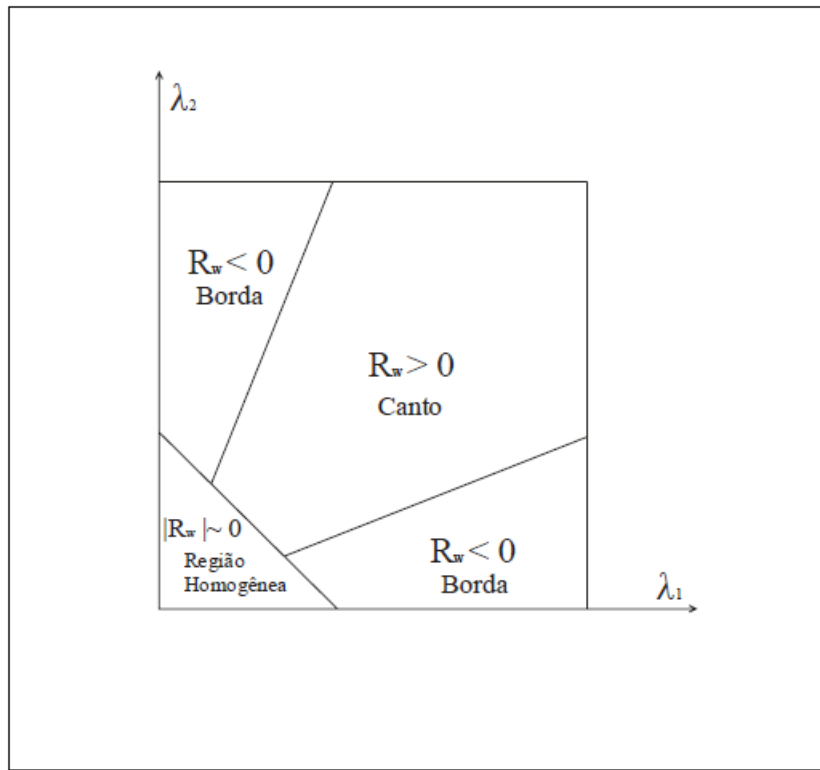


Figura 3.2: Classificação em função de R

Para detectarmos os pontos de canto da imagem (pontos característicos) é necessário encontrar os pontos com alto valor de resposta aos cantos. Apenas as regiões acima de um limiar são analisadas ($R > \text{limiar}$) a fim de eliminar a possibilidade de encontrar pontos característicos em regiões planas. Para a determinação da posição precisa dos pontos característicos, é proposto determinar os máximos locais do resultado como sendo a posição dos pontos. Para a determinação dos máximos locais, uma janela de tamanho fixo é aplicada e o máximo global desta janela é julgado como sendo um máximo local da imagem. Esta janela é deslocada em toda a imagem a fim de encontrar todos os máximos locais (pontos característicos).

Na Figura 3.3 é mostrado um exemplo da detecção de pontos característicos. A partir da imagem original (A) são apresentados apenas os valores de R maiores que um limiar são analisados (B), e por fim, os máximos locais encontrados são determinados como sendo pontos característicos (C).

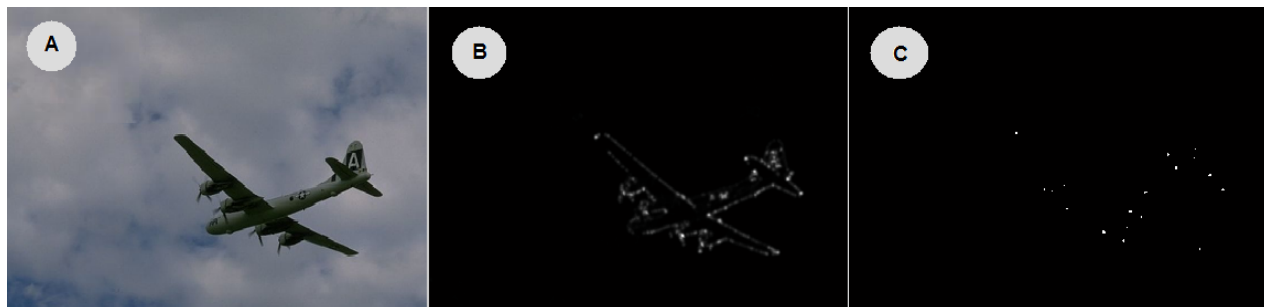


Figura 3.3: Imagem Original(A), valores de $R > \text{limiar}$ (B) e os máximos locais ou cantos (C)

Observações importantes a respeito do detector de cantos de Harris:

- A sensibilidade a ruídos é função da dimensão da máscara gaussiana utilizada; máscaras maiores elevam o custo computacional e podem acarretar borramentos na imagem;
- Quanto a localização dos cantos seu melhor desempenho diz respeito as formas em L;
- Embora por vezes seja relatado na literatura como um operador isotrópico (rotacionalmente invariante), na verdade em função de que os autovalores da matriz M_w são calculados usando gradientes horizontais e verticais, resulta num operador anisotrópico;
- Por efetuar convolução da imagem com máscara gaussiana, ele não é o método com o menor custo computacional, porém, o detector de cantos de Harris é largamente utilizado, está presente no estado da arte no que se refere à técnicas de detecção de cantos e normalmente tem vantagem sobre os demais métodos quando considerados todos os quesitos para um detector de cantos ideal, que pode ser comprovado em recente publicação de (ZOU et al., 2008) onde é demonstrada a superioridade do detector de cantos Harris (HARRIS; STEPHENS, 1988).

3.2 O Agrupamento dos dados

Agrupamento é uma classificação não supervisionada de padrões (observações, itens de dados ou vetores de características) em grupos. Intuitivamente, cada grupo é composto por padrões que são similares entre si e dissimilares em relação aos padrões de outros grupos (JAIN; MURTY; FLYNN, 1999).

Um exemplo de agrupamento é descrito na Figura 3.4. O padrão de entrada é mostrado na Figura 3.4 a) e os grupos são mostrados na Figura 3.4 b). Pontos dentro de um mesmo grupo apresentam um mesmo rótulo.

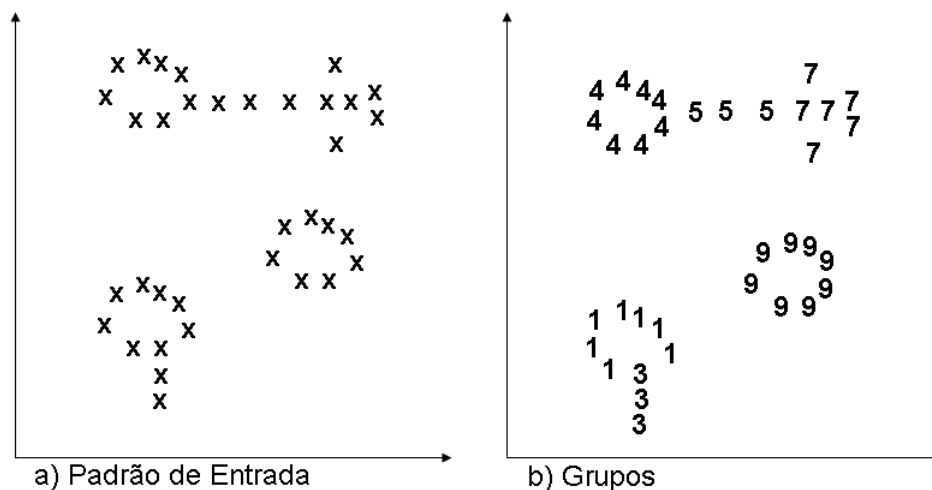


Figura 3.4: Etapas de uma tarefa de agrupamento

3.2.1 Etapas do Processo de Agrupamento

A Figura 3.5 descreve hipoteticamente uma seqüência típica das atividades envolvidas em uma tarefa de agrupamento, incluindo um caminho de retorno pelo qual a saída do processo de agrupamento pode afetar as tarefas seguintes (JAIN; MURTY; FLYNN, 1999). Os padrões representativos se referem ao número de classes, ao número de padrões avaliados e ao número, tipo e escalas de características disponíveis no processo de agrupamento. A seleção de características é o processo de identificar o subconjunto mais eficaz de características para serem utilizadas no agrupamento. A similaridade entre padrões é normalmente medida por meio de uma função de distância definida entre pares de padrões. Por fim, a etapa de agrupamento consiste na técnica utilizada para formar grupos a partir desses valores de similaridade (JAIN; MURTY; FLYNN, 1999).

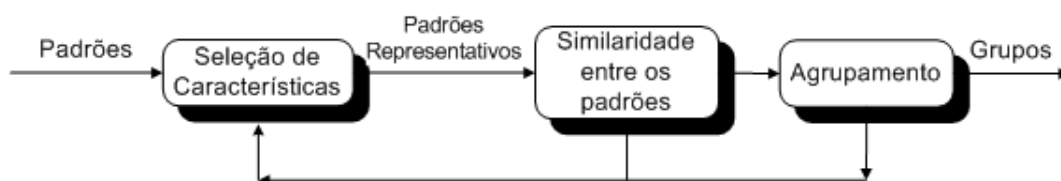


Figura 3.5: Etapas de uma tarefa de agrupamento

3.2.2 Paradigmas de Agrupamento

3.2.2.1 *hierárquico*

Um modelo hierárquico de agrupamento constrói uma hierarquia de grupos ou, em outras palavras, uma árvore de grupos, conhecida como dendrograma. Dessa forma, esse método permite explorar dados sob diferentes níveis de granularidade (BERKHIN, 2002). As estratégias de agrupamento hierárquico podem ser divididas em dois modelos básicos: aglomerativos e divisivos. Uma estratégia aglomerativa começa com cada elemento como um grupo independente e os une sucessivamente em grupos cada vez maiores. Por outro lado, uma estratégia divisiva começa com um grande grupo e o divide sucessivamente em grupos menores (BERKHIN, 2002).

3.2.2.2 *Modelo particional*

Um modelo particional de agrupamento obtém uma única partição dos dados ao invés de uma estrutura de grupos, como o dendrograma produzido pelos métodos hierárquicos (JAIN; MURTY; FLYNN, 1999). *K-means* e *K-medoids* são os dois métodos particionais de agrupamento mais conhecidos (HASTIE; TIBSHIRANI; FRIEDMAN, 2001). *K-means* é voltado para aplicações em que todas as variáveis são quantitativas e as dissimilaridades entre elas podem ser medidas em um espaço Euclidiano. *K-medoids* é uma generalização do *K-means* no sentido que outras funções de distâncias, além da distância Euclidiana, podem ser utilizadas para medir as dissimilaridades entre os elementos (BERKHIN, 2002).

3.2.2.3 *Aplicações de técnicas de agrupamento*

Muitos pesquisadores têm utilizado métodos de agrupamento na recuperação de documentos (SAHOO et al., 2006), (ZHAO et al., 2002). (FERRAGINA; GULLI, 2005) propuseram o SnakeT, um método hierárquico de agrupamento para organizar os resultados de diversos mecanismos de busca sob demanda. Eles utilizaram a hierarquia obtida para complementar a análise sequencial dos resultados retornados pelos mecanismos de busca disponíveis.

Um processo de agrupamento pode ser aplicado em diferentes contextos. Por exemplo, alguns métodos utilizam agrupamento no seu próprio espaço de características (BAEZA-YATES et al., 2003).

(COOPER et al., 2005) propuseram um método de agrupamento baseado na similaridade entre sequências de fotos. Os autores mostraram que fotografias de um mesmo

evento e muito próximas no tempo podem reduzir 33% do tempo de processamento de uma consulta nesse sistema. Kim e Chung (KIM, 2003) utilizaram uma técnica adaptativa de agrupamento para filtrar os resultados de um sistema de recuperação de imagens por conteúdo. (SHYU et al., 2004) introduziram uma estrutura unificada para facilitar o agrupamento em bancos de atributos e recuperar imagens utilizando modelos mediadores de Markov (*MMM Markov Models Mediators*).

(HELLER; GHAHRAMANI, 2005) desenvolveram um algoritmo hierárquico aglomerativo de agrupamento baseado em probabilidades. Entretanto, esse método é inviável para recuperação de imagens por conteúdo, uma vez que é difícil encontrar um modelo probabilístico apropriado para representar imagens.

(KINOSHENKO et al., 2005) propuseram uma técnica para particionar uma imagem em subconjuntos disjuntos. Nesse método, o sistema divide cada consulta em subclasses representativas e encontra a subclasse mais similar armazenada para cada parte da consulta. Dessa forma, a consulta é hierarquicamente classificada em partes. Entretanto, o método escolhe essas classes de maneira fortemente conexa, isto é, as classes das imagens precisam representar uma estrutura hierárquica, por exemplo, o relacionamento presente na imagem de um carro e suas partes. Em geral, é difícil encontrar essas hierarquias em imagens de domínio geral, mas sim, em domínios específicos e bem conhecidos.

3.2.3 Mean-Shift: Fundamentação e Algoritmo

O algoritmo de deslocamento pela média talvez seja o melhor segmentador existente na atualidade (BRASDSKI G. R.; THRUN, 2006). Esta afirmação é confirmada também em recentes publicações encontradas na literatura (ZHAO; KNIGHT; REID, 2008), (KALPATHY-CRAMER; HERSH, 2008), (SHOTTON; BLAKE; CIPOLLA, 2008), (VENUGOPAL; SUDHAMANI, 2008), (WANG et al., 2008), (HAN; PAN, 2008), (RAO; MARTINS; PRÍNCIPE, 2009) que utilizam este método de agrupamento de dados em suas implementações.

O algoritmo de deslocamento pela média *mean-shift* é uma técnica robusta de classificação de agrupamentos que não exige conhecimento a priori sobre a quantidade de grupos, nem restringe o formato destes agrupamentos. O *mean-shift* trata-se de um método iterativo que estima a moda (regiões mais densas) de distribuições multivariadas obtidas no espaço de atributos. A quantidade de agrupamentos é obtida automaticamente pela descoberta dos centros destas modas.

Em virtude destas características, empregou-se o procedimento de (COMANICIU; MEER, 2002) para a etapa de segmentação da imagem colorida. Tendo em vista que o propósito deste trabalho é a classificação de imagens coloridas heterogêneas, logo, não há conhecimento prévio de qual o melhor número de agrupamentos que o método deveria gerar em cada imagem. Métodos de agrupamento importantes como o *K-means*, por exemplo, exigem que sejam informados alguns parâmetros antes da execução da tarefa de agrupamento, o que tornaria inviável seu uso para resolver o problema proposto por este trabalho.

A Equação (2.11) define o vetor de deslocamento pela média, sendo primeiramente apresentada por (FUKUNAGA; HOSTETLER, 1975), como um algoritmo de classificação não-paramétrica, e mostra que a estimativa do gradiente de densidade é dado por:

$$(\hat{\nabla} f(x)) \quad (3.11)$$

e normalizado

$$(\hat{\nabla} f(x)/\hat{f}(x)) \quad (3.12)$$

pode ser calculada através do deslocamento amostral pela média em um kernel uniforme centrado em x , ou seja, com as mesmas dimensões de altura e largura, por exemplo. O deslocamento pela média é um excelente estimador não supervisionado do gradiente de densidade não-paramétrico, e o vetor de deslocamento pela média representa a diferença entre uma média local, calculada em uma janela com raio (banda de passagem) h no espaço d -dimensional, e o centro desta janela.

$$\vec{M}_h(x) = \frac{h^2}{(d+2)} \frac{\hat{\nabla} f(x)}{\hat{f}(x)} \quad (3.13)$$

O vetor de deslocamento pela média aponta sempre na direção do maior incremento na densidade, isto é, descobre o caminho que leva a um máximo local de densidade (moda da densidade). O algoritmo trabalha em forma iterativa, deslocando as amostragens sucessivamente em direção ao máximo local mais adequado. A Figura 3.6 ilustra esta tendência.

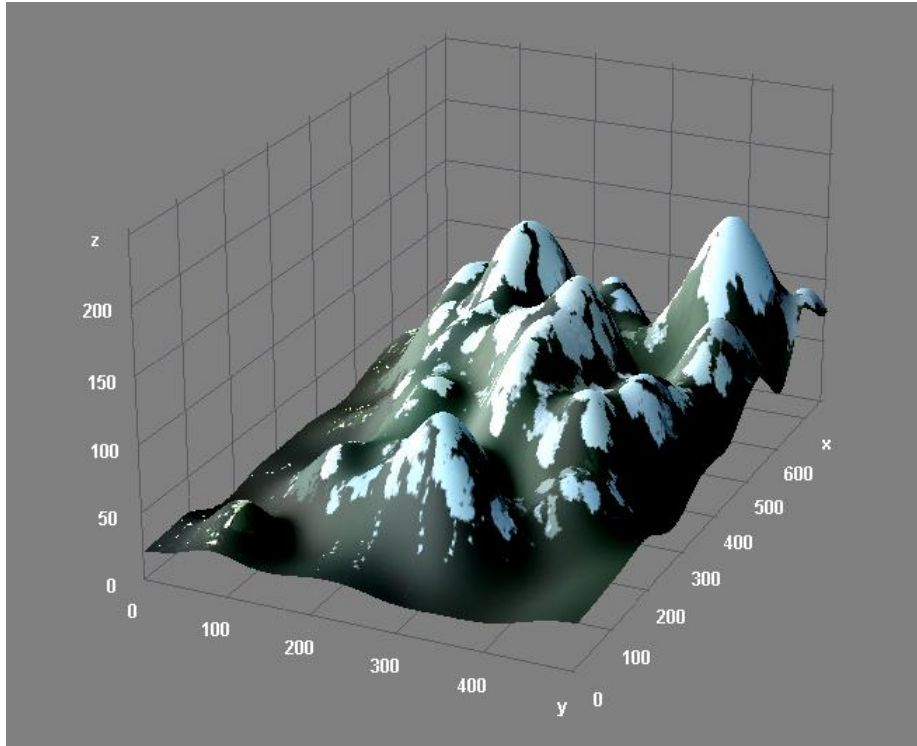


Figura 3.6: Tendência do algoritmo de deslocamento pela média.

O algoritmo de deslocamento pela média pode ser estendido para processamento de imagens levando em consideração a vizinhança espacial de cada pixel selecionado (COMANICIU; MEER, 2002). Estes autores introduzem o domínio conjunto: toda amostra passa a ser descrita pela união do domínio da imagem (características da imagem: tons de cinza, componentes RGB, etc.) com o domínio espacial (localização de um pixel dentro da imagem). Dado:

$$Z_i^m = \{S_i^m, a_i^m\}, \forall i \in [1..n] \quad (3.14)$$

domínio conjunto, onde s representa a informação espacial e a a informação contida na imagem, (COMANICIU; MEER, 2002) criaram o kernel definido pelo produto de dois kernels: K_A para o domínio da imagem, e K_S para o domínio espacial. A estimativa de densidade (suavizada) pode ser então, escrita como:

$$f(\hat{z}_j^m) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_A h_S} K_A \left(\frac{a_j^m - a_i^m}{h_A} \right) K_S \left(\frac{s_j^m - s_i^m}{h_S} \right) \quad (3.15)$$

onde h_A e h_S são as bandas de passagem (janelas) no domínio da imagem e no domínio espacial, respectivamente. Somente as amostras que recaem dentro dos limites

impostos por estas bandas de passagem e Z_j^m é que contribuem para a estimativa da densidade no ponto Z_j^m . O deslocamento das médias é realizado em ambos os espaços separadamente, isto é:

$$a_j^{m+1} = a_j^m + M_{h_A, K_A}(a_j^m) \quad (3.16)$$

e

$$s_j^{m+1} = s_j^m + S_{h_S, K_S}(s_j^m) \quad (3.17)$$

e a convergência ocorre simultaneamente nos dois domínios, quando ambos os vetores de deslocamento pela média ficarem abaixo de um limiar. O algoritmo classifica os pontos segundo os agrupamentos existentes nesses domínios. O algoritmo de deslocamento pela média tem este nome porque, no fundo, não passa de um calculador de média ponderada, deslocando-se segundo os valores desta média. Os pesos dos pixels advêm do kernel empregado. Neste trabalho empregou-se o kernel uniforme, isto é, as médias são médias aritméticas. A definição do tamanho da janela de pesquisa h_A , no espaço de atributos de cores, resulta de medida da "atividade visual" da imagem, isto é, deriva da matriz global de variância e covariância de cores.

O tamanho da janela de pesquisa no domínio de atributos espaciais h_S deve ser obtido empiricamente, ocorrendo um compromisso entre os efeitos indesejáveis de divisão de uma mesma região (tamanho de janela pequeno) ou união de diferentes regiões (tamanho de janela grande). Neste trabalho adotou-se o valor único de 21x21 como sendo o tamanho da janela de pesquisa no âmbito de atributos espaciais.

A janela de pesquisa quer no domínio das cores (três dimensões), quer no domínio espacial (duas dimensões), definem uma hiperesfera de cinco dimensões. As três dimensões do espaço de atributos das cores são iteradas, no algoritmo de deslocamento pela média, assumindo novos valores e, por conseguinte alterando o posicionamento da hiperesfera. A janela de pesquisa no domínio espacial é fixa durante todo o processo. Todo o algoritmo de deslocamento pela média é executado dentro desta hiperesfera semimóvel (parte móvel, parte fixa), com resultado válido para o pixel central da hiperesfera final.

3.3 A Detecção de Cantos e o Agrupamento na Construção de Metadados

A recuperação de imagens por conteúdo é uma das áreas de pesquisa mais importantes da área da visão Computacional na atualidade. Existem diversas técnicas utilizadas e pesquisadas, sendo que as de detecção de cantos são uma das mais importantes nesta área, como pode ser observado em (GEVREKCI; GUNTURK, 2009). As técnicas de agrupamento de imagens, também, têm sido amplamente utilizadas, como pode ser observado em ((ZHAO; KNIGHT; REID, 2008), (KALPATHY-CRAMER; HERSH, 2008), (SHOTTON; BLAKE; CIPOLLA, 2008), (VENUGOPAL; SUDHAMANI, 2008), (LI; SHIH; HSU, 2004), sendo que ampla utilização deve-se ao fato de que, as mesmas, proporcionam uma diminuição da complexidade computacional, já que reduzem a quantidade de informações contidas nas imagens, agrupando-as por determinado critério, porém sem reduzir a qualidade da informação de baixo nível a ser extraída da imagem.

A baixa complexidade computacional é uma característica tão importante em sistemas CBIR quanto a precisão das imagens retornadas. Este trabalho demonstra que o uso em conjunto destas técnicas é viável, tanto no ponto de vista da complexidade computacional, quanto sob o ponto de vista da precisão das imagens recuperadas pelo sistema.

4 MÉTODO DESENVOLVIDO

O método *mean-shift*, além de gerar agrupamentos com bastante qualidade, não exige um conhecimento prévio do número de agrupamentos a ser gerado, o que torna o método bastante viável para este trabalho, uma vez que o objetivo é classificar imagens coloridas heterogêneas, onde nunca se terá conhecimento prévio do "possível melhor" número de agrupamentos a ser gerado. Com o objetivo de otimizar a ação do algoritmo Harris foi empregado uma técnica de baixa complexidade computacional para corrigir a iluminação das imagens durante a fase de pré-processamento. A correção da iluminação tem o objetivo de otimizar a ação do algoritmo Harris (HARRIS; STEPHENS, 1988), evitando que sejam detectados falsos positivos, como por exemplo cantos em sombras, bem como, realçar cantos que poderiam não ser detectados na imagem original.

4.1 Arquitetura

Na Figura 4.1 observa-se o que se pode chamar de duas grandes fases do método desenvolvido, onde há uma fase de pré-processamento, que ocorre de modo off-line e uma fase de indexação, classificação e recuperação das imagens.

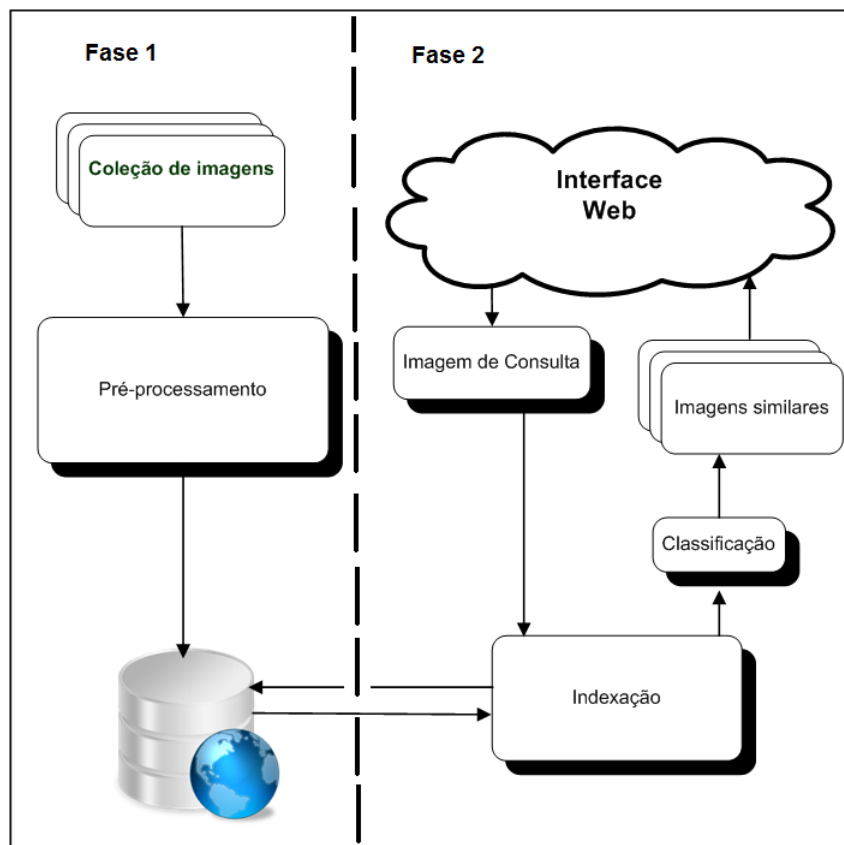


Figura 4.1: Arquitetura do método desenvolvido

4.1.1 Pré-Processamento

O pré-processamento proposto neste trabalho objetiva eliminar a produção de agrupamentos fragmentados, ou seja, evitar a que o agrupamento da imagem seja efetuado várias vezes durante o processo de comparação das imagens, o que traria um acréscimo bastante elevado da complexidade computacional do método. Com o uso do pré-processamento, o método desenvolvido armazena uma vez só na base de dados todos os metadados necessários para a tarefa de indexação e comparação de imagens, baseadas em suas regiões e seus cantos detectados, ou seja, o método aplica o algoritmo de agrupamento e detecção de cantos somente uma vez durante todo o fluxo de execução do método. Todo o pré-processamento é efetuado off-line. Os metadados extraídos durante este processo são armazenados em um banco de dados relacional. A Figura 4.2 ilustra em linhas gerais os passos executados pelo método na fase de pré-processamento, enquanto a Figura 4.3 apresenta a GUI (*Graphical user interface*) desenvolvida para facilitar este procedimento.

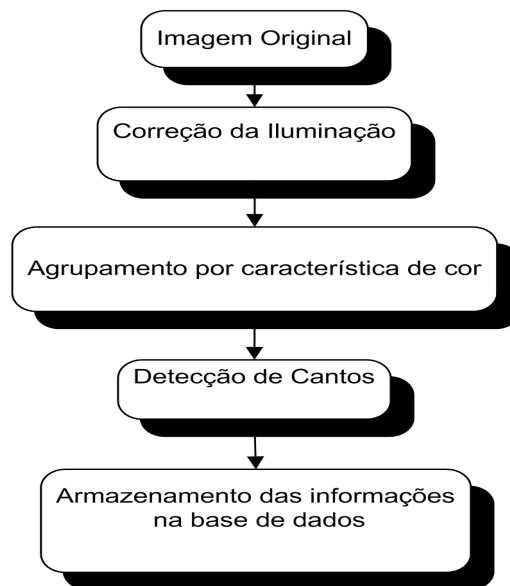


Figura 4.2: Fases do pré-processamento empregado pelo método desenvolvido

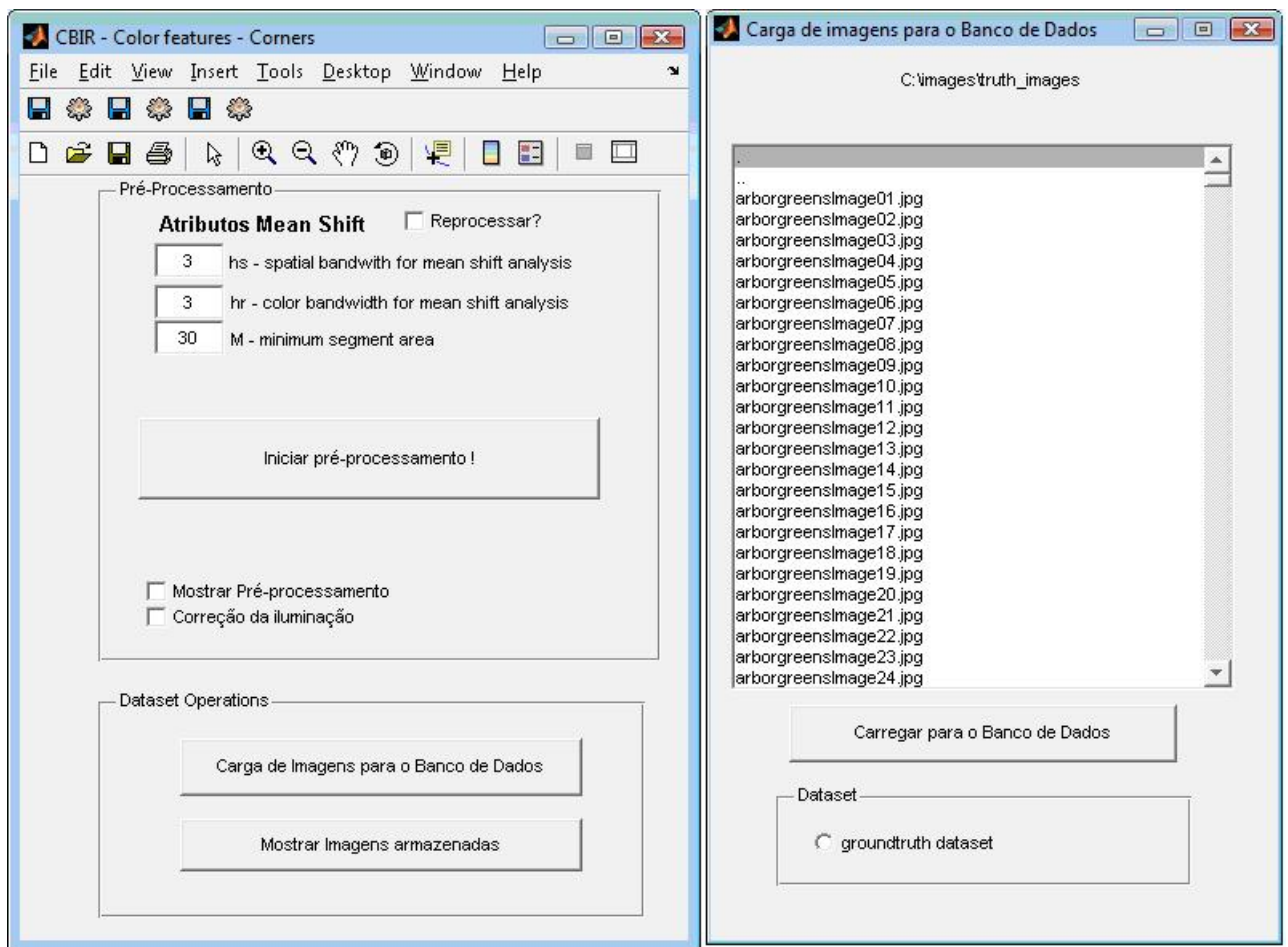


Figura 4.3: - Interface desenvolvida em MATLAB utilizada na fase de pré-processamento do método desenvolvido

A fase de agrupamento de dados é de fundamental importância na solução proposta por este trabalho. Desta forma, foram efetuados vários experimentos para o aperfeiçoamento da fase de segmentação, tentando chegar a um bom nível de segmentação, porém com um tempo de processamento baixo. Considerando que a etapa de agrupamento ocupa em média, cerca de 60% do tempo total do processamento das imagens, esta análise se tornou importante, para que obtivéssemos bons resultados com um baixo tempo de processamento. O algoritmo *mean-shift* permite definir três parâmetros de entrada, que são:

Spatial bandwidth: é um valor que define o tamanho da janela de pesquisa na imagem, e por isso, recebe valores positivos inteiros.

Color bandwidth: Este parâmetro define o valor limite da distância entre duas cores, a partir do qual elas são tomadas como diferentes, ou pertencendo a conjuntos diferentes.

Minimum segment area: a área mínima das regiões obriga a que cada região tenha uma área igual ou superior a esse valor e toma também números inteiros positivos.

Para verificar quais valores de parâmetros produziam melhores resultados, foram efetuados testes com 5 imagens de dimensão 756 x 504 pixels, variando os valores dos parâmetros. A primeira observação é que o parâmetro *color bandwidth* não afeta de maneira importante o tempo de processamento, sendo que a diferença no tempo de processamento alterando os valores deste parâmetro, fica na casa de centésimos de segundos. Por outro lado, o parâmetro *spatial bandwidth* afeta bastante o tempo de processamento. Considerando que este trabalho aplicaria o método de detecção de cantos na imagem agrupada por características de cor, durante os testes foi considerado o número de cantos detectados em função da variação dos parâmetros do método *mean-shift*. A conclusão foi que, na maioria dos casos, os melhores resultados foram obtidos fixando os parâmetros, durante a fase de pré-processamento com os seguintes valores: *Spatial bandwidth* = 3; *Color bandwidth* = 3; *Minimum segment area* = 30; A tabela 4.1 apresenta os testes considerando a variação do parâmetro *spatial bandwidth* e seus respectivos tempos médios de segmentação por imagem.

Tabela 4.1: Testes de processamento na fase de agrupamento

Spatial Bandwidth	Tempo médio de processamento p/ imagem (em segundos)
3	1,8
5	2,5
10	3,2
15	4,1
20	4,8
30	5,2
40	5,6
50	7,1

4.1.2 Correção da Iluminação

Com o objetivo de aperfeiçoar a ação do detector de cantos empregado neste trabalho, no sentido de não considerar sombras da imagem como cantos e também detectar cantos que na imagem original não seriam detectados, foi empregada uma fase de correção de iluminação durante o pré-processamento das imagens. Esta técnica foi proposta por (JAHNE, 1999), baseada em operações básicas da morfologia matemática e possui baixa complexidade computacional e que atendeu perfeitamente as necessidades deste trabalho.



Figura 4.4: Imagem original e agrupada com os cantos detectados



Figura 4.5: Imagem com iluminação otimizada, agrupada com os cantos detectados

Na Figura 4.4 está um exemplo da aplicação do algoritmo de detecção de cantos na imagem agrupada original, ou seja, sem a correção de iluminação aplicada, já Figura 4.5 foi aplicada a detecção de cantos na imagem agrupada com a correção da iluminação. É perceptível que os cantos detectados na imagem representada pela Figura 4.5 possuem cantos detectados com maior precisão, por exemplo, basta observar as "sombrias" que na Figura 4.4 foram consideradas como canto, já na imagem representada pela Figura 4.5, não foram. Outro aspecto importante é a detecção de cantos reais, o que pode ser observada nos pilares dos prédios, que de fato são cantos da imagem, e não haviam sido detectados na imagem 4.4.

4.1.3 Agrupamento

O método de agrupamento utilizado por este trabalho trata-se de um método iterativo que estima a moda (regiões mais densas) de distribuições multivariadas obtidas no espaço de atributos. A quantidade de agrupamentos é obtida automaticamente pela descoberta dos centros destas modas. Uma vez que o objetivo deste trabalho é a classificação de imagens coloridas, o método *mean-shift* se faz ideal, uma vez que nunca será conhecido previamente o número de agrupamentos gerados em uma imagem colorida, quando se faz agrupamento por característica de cor em imagens heterogêneas. Após a aplicação deste método, apenas algumas cores se mantêm na imagem resultado, e é possível extrair informações como número de cantos, tamanho e circunferências destas regiões. Este trabalho priorizou a extração dos cantos das regiões como principal metadado a ser utilizado na comparação entre as imagens. A Figura 4.6 mostra a imagem original e seu respectivo histograma em 3D no espaço de cores RGB. A figura 4.7 apresenta a imagem após passar pelo processo de agrupamento por característica de cor e seu respectivo histograma em 3D representado no espaço de cores RGB. É possível observar na Figura 4.7 que o espaço de busca de cantos diminui, otimizado o método desenvolvido.

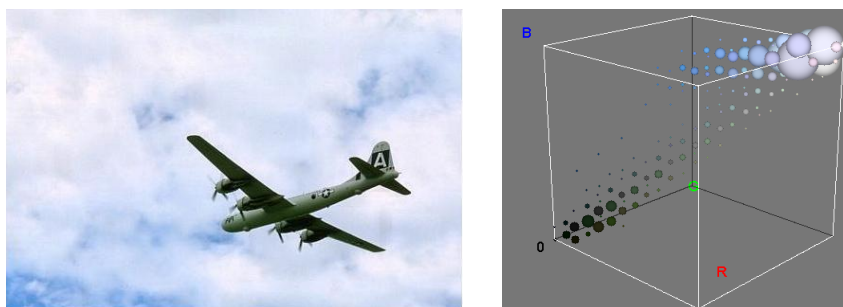


Figura 4.6: Imagem original sua representação 3D no espaço de cores RGB

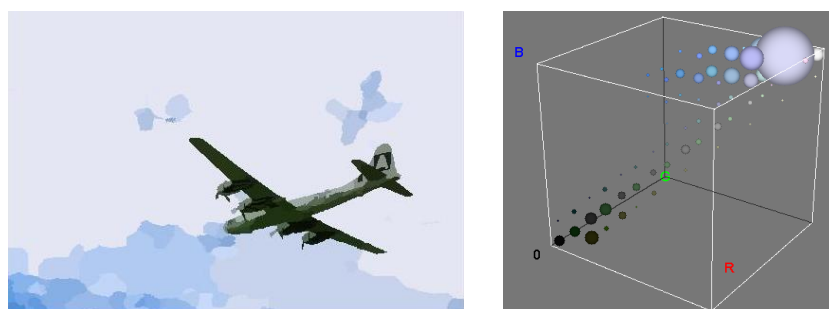


Figura 4.7: Imagem agrupada por característica de cor e sua representação 3D no espaço de cores RGB

4.1.4 Detecção de Cantos

Este trabalho aplica o algoritmo de detecção de cantos desenvolvido por (HARRIS; STEPHENS, 1988) na imagem agrupada por característica de cor e com iluminação corrigida. A escolha desta técnica se deve ao fato deste método estar presente no estado da arte da literatura, com superioridade comprovada, que pode ser visto em recente publicação de (ZOU et al., 2008). A Figura 4.8 demonstra o resultado da aplicação do método de detecção de cantos na imagem agrupada e com sua intensidade de iluminação corrigida, porém apresentados na imagem original, o que atesta a robustez da detecção de cantos, onde mesmo aplicando-o em uma imagem com menor complexidade computacional, os cantos detectados são compatíveis.



Figura 4.8: Cantos detectados na imagem agrupada e apresentados na imagem original

4.1.5 Indexação de Imagens

Os passos efetuados durante o processo de indexação dos metadados extraídos, visando sua utilização posterior na recuperação das imagens por semelhança estão dispostos na Figura 4.9.

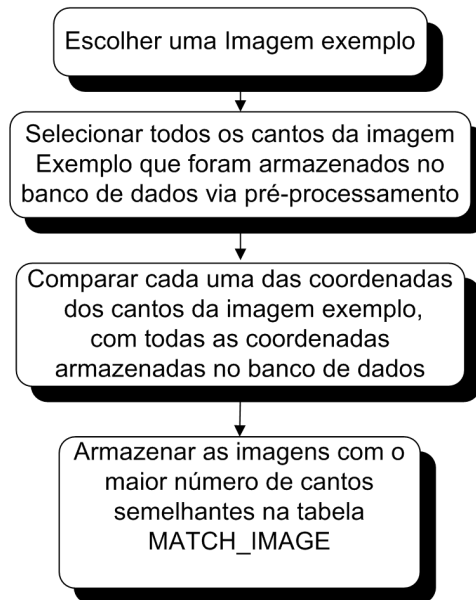


Figura 4.9: Procedimento de indexação das imagens

A tabela 4.2 demonstra a instrução SQL responsável pela indexação dos metadados extraídos pelo método desenvolvido. O objetivo é armazenar os quantificadores necessários na tabela MATCH_IMAGE, conforme Figura4.10.

Esta operação é efetuada via uma instrução SQL. A formulação matemática para o critério de indexação está definida pela equação (3.1).

$$I(x_i, y_i) \begin{cases} 1, & \text{if } I(x_i, y_i) \leq K(x_i - \sigma, y_i - \sigma) \leq K(x_i + \sigma, y_i + \sigma) \\ 0, & \text{if } I(x_i, y_i) \geq K(x_i - \sigma, y_i - \sigma) \geq K(x_i + \sigma, y_i + \sigma) \end{cases} \quad (4.1)$$

Para $i = 1$ até n , onde n é o número de corners armazenados via pré-processamento, e I armazena os corners de todas as imagens armazenadas na base de dados e K são os corners da imagem exemplo os quais estão sendo comparados para verificar sua similaridade. Assim sendo, a imagem mais semelhante será a que retornar valores 1 o maior número de vezes, durante a execução da indexação. Objetivando um maior dinamismo na comparação entre as imagens, foi, adicionalmente, criado um fator de variação σ que

permite que cantos detectados em áreas próximas (pixels vizinhos) aos detectados originalmente na fase de pré-processamento, sejam também considerados como similares aos cantos da imagem exemplo. Foram efetuados testes variando σ de 1 a 30, e os resultados de eficiência armazenados. Na média, melhores resultados apresentados foram com $\sigma = 20$.

Tabela 4.2: Instrução SQL responsável pela classificação das imagens por ordem de semelhança

```

SELECT a.id_image, count(*), a.coord_x,a.coord_y,c.coord_x,c.coord_y
      FROM corner a, image b , corner_origem c
      WHERE a.id_image = b.id_image and
(a.COORD_X >= c.coord_x -  $\sigma$  AND a.COORD_X <= c.coord_x +  $\sigma$ ) AND
(a.COORD_y >= c.coord_y -  $\sigma$  AND a.COORD_y <= c.coord_y +  $\sigma$ )
      GROUP BY a.id_image order by 2 desc

```

O conjunto de dados retornados pelo processamento da consulta SQL demonstrada na tabela 4.2, é armazenado em uma tabela denominada MATCH_IMAGE, que será utilizada na fase de recuperação das imagens.

MATCH IMAGE
ID_MATCH_IMAGE: Int;
ID_IMAGE_IN: Int;
ID_SIMILAR_IMAGE: Int;
NUMBER_CORNER_POINTS: Int;

Figura 4.10: Projeto físico da tabela MATCH_IMAGE

4.1.6 Recuperação de Imagens

A forma mais usual para fazer uma consulta em um sistema CBIR é informar uma imagem de entrada, e o sistema retornar as imagens mais semelhantes em ordem decrescente (*rank*), ou seja, ordenado do maior para o menor, as imagens mais similares. Com este objetivo, é executada uma busca direta na tabela MATCH_IMAGE via SQL (*Structured Query Language*), ordenando o resultado da consulta pela coluna NUMBER_CORNER_POINTS em ordem decrescente, pois ela é na verdade uma estrutura que acumula a quantidade de cantos semelhantes das imagens pré-processada em relação à imagem exemplo informada pelo usuário. Assim o método retornará automaticamente ordenado por *rank*, as imagens mais semelhantes. A tabela 4.3 demonstra a consulta SQL

que efetua a classificação das imagens, retornando uma seqüência de imagens ordenadas por *rank*.

Tabela 4.3: Instrução SQL responsável pelo processo de recuperação das imagens

```
SELECT path_image FROM match_image a, image b
WHERE a.id_similar_image = b.id_image ORDER BY number_corner_points DESC
```

4.2 Uma Aplicação Cliente-Servidor para a Recuperação de Imagens por Conteúdo

A arquitetura do protótipo computacional desenvolvido neste trabalho permite que após o pré-processamento do conjunto de imagens, os seus metadados sejam disponibilizados para qualquer tipo de aplicação que tenha possibilidade de acessar os dados em um banco de dados relacional. Neste contexto, os aplicativos Web, por sua natureza Cliente-Servidor conseguem utilizar-se da técnica desenvolvida neste trabalho com grande facilidade, tornando o protótipo computacional aqui desenvolvido, bastante versátil enquanto disponibilização da ferramenta para os mais diversos conjuntos de imagens possíveis.

Além da GUI (*Graphic User Interface*)- desenvolvida para a fase de pré-processamento, foi desenvolvido uma interface Web, que permite a utilização do método implementado neste trabalho. A Figura 4.11 mostra a interface Web desenvolvida, já a Figura 4.12 mostra o retorno do sistema após uma consulta por imagem exemplo. A imagem assinalada com vermelho indica a imagem consulta submetida.

Categorias

Escolha uma Categoria

- Escolha uma Categoria
- Todas
- arborgreensImage
- australialImage
- barcelonalImage
- cambridge
- campusinfallImage
- cannonbeachImage
- cherriesImage
- footballImage**
- genevalImage
- greenlakelImage
- greenlandgreenland
- indonesialImage
- iranImage
- italyImage
- japanImage

Escolha uma imagem consulta para recuperar imagens similares. Galeria selecionada: **footballImage**

Range x 15 y 15 Range x 15 y 15 Range x 15 y 15 Range x 15 y 15 Range x 15 y 15

Cantos na imagem: 107; Cat: footballImage. Cantos na imagem: 53; Cat: footballImage. Cantos na imagem: 47; Cat: footballImage. Cantos na imagem: 63; Cat: footballImage. Cantos na imagem: 49; Cat: footballImage.

Range x 15 y 15 Range x 15 y 15 Range x 15 y 15 Range x 15 y 15 Range x 15 y 15

Cantos na imagem: 34; Cat: footballImage. Cantos na imagem: 140; Cat: footballImage. Cantos na imagem: 47; Cat: footballImage. Cantos na imagem: 24; Cat: footballImage. Cantos na imagem: 31; Cat: footballImage.

| 1 | 2 | 3 | 4 | 5 | Próxima >

Voltar

Figura 4.11: Tela principal da Interface Web Desenvolvida

Imagens retornadas na consulta: **Galeria selecionada: footballImage**

Range x 25 y 25 Range x 25 y 25 Range x 25 y 25 Range x 25 y 25 Range x 25 y 25

Cantos na imagem: 105; Cantos similares: 387; Distancia Total: 96696; Tempo: 0.0001; segundos. Cat: footballImage. Cantos na imagem: 71; Cantos similares: 145; Distancia Total: 66883; Tempo: 0.0171; segundos. Cat: footballImage. Cantos na imagem: 61; Cantos similares: 128; Distancia Total: 47804; Tempo: 0.0354; segundos. Cat: footballImage. Cantos na imagem: 66; Cantos similares: 124; Distancia Total: 51913; Tempo: 0.0522; segundos. Cat: footballImage. Cantos na imagem: 134; Cantos similares: 124; Distancia Total: 51301; Tempo: 0.0691; segundos. Cat: footballImage.

Range x 25 y 25 Range x 25 y 25 Range x 25 y 25 Range x 25 y 25 Range x 25 y 25

Cantos na imagem: 118; Cantos similares: 124; Distancia Total: 62310; Tempo: 0.0861; segundos. Cat: footballImage. Cantos na imagem: 110; Cantos similares: 123; Distancia Total: 53478; Tempo: 0.1117; segundos. Cat: footballImage. Cantos na imagem: 140; Cantos similares: 120; Distancia Total: 51102; Tempo: 0.1331; segundos. Cat: footballImage. Cantos na imagem: 94; Cantos similares: 118; Distancia Total: 53688; Tempo: 0.1525; segundos. Cat: footballImage. Cantos na imagem: 93; Cantos similares: 106; Distancia Total: 42978; Tempo: 0.1717; segundos. Cat: footballImage.

| 1 | 2 | 3 | 4 | 5 | Próxima >

Figura 4.12: Exemplo de retorno de consulta por semelhança

A base de testes utilizada foi a de denominação *groundtruth*, disponível em <http://www.cs.washington.edu/research/imagedatabase/groundtruth>. Esta base possui 1085 imagens pertencentes à 20 categorias variadas. Todas estas imagens possuem o mesmo tamanho, 756 x 504 pixels, foram pré-processadas e seus metadados armazenados em um banco de dados. As categorias utilizadas nos testes são as seguintes:

Tabela 4.4: Categorias de imagens utilizadas nos testes computacionais

1	arborgreensImage
2	australiaImage
3	barcelona2
4	barcelonaImage
5	cambridge
6	campusinfallImage
7	cannonbeachImage
8	cherriesImage
9	footballImage
10	genevaImage
11	greenlakeImage
12	greenlandgreenland
13	indonesiaImage
14	iranImage
15	italyImage
16	japanImage
17	leaflesstreesimage
18	sanjuansImage
19	springflowersImage
20	swissmountainsImage

4.3 Resultados e conclusões finais

4.3.1 Análise dos Resultados

Este trabalho comparou seu desempenho com os resultados apresentados pela recente publicação de (VENUGOPAL; SUDHAMANI, 2008), que também utiliza o método de agrupamento imagens *mean-shift* como base no processo de extração de metadados das imagens, e por isto foi escolhido como ponto de comparação com os resultados deste trabalho.

4.3.1.1 Complexidade computacional

Para verificar a complexidade computacional do método de indexação proposto e também validar a eficiência dos resultados, o procedimento de testes ocorreu da seguinte

forma: Foram eleitas aleatoriamente 70 imagens (não repetidas) como imagens exemplo, e a média dos resultados obtidos estão apresentadas na Figura 4.13.

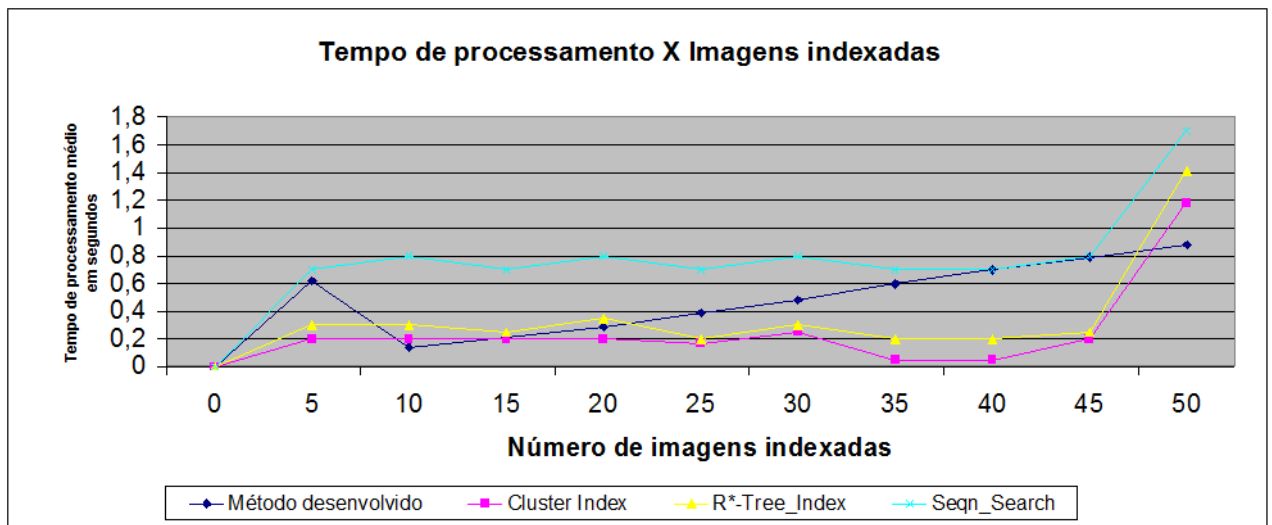


Figura 4.13: Resultados referentes à complexidade computacional do método desenvolvido

4.3.1.2 Precisão dos resultados

O critério de avaliação para a precisão dos retornos obtidos após a execução do método está disposto a seguir:

$$\text{precisao}(k) = c_k = k ; \text{recall}(k) = c_k = M ; \text{onde:}$$

k = número de imagens retornadas;

c_k = número de retornos relevantes entre os k retornos;

M = Total de imagens relevantes encontradas no banco de dados.

Tabela 4.5: Resultados de todos os métodos envolvidos no comparativo

Categorias	Cluster index		R* Tree Index		Sequential Search		Mét. desenv.	
	Precisão	Recall	Precisão	Recall	Precisão	Recall	Precisão	Recall
1	1	1	1	1	1	0,989	0,988	0,971
2	0,9	1	0,875	1	0,9	1	0,89	1
3	0,958	1	0,887	0,989	0,916	1	0,956	0,971
4	0,983	0,986	0,983	0,986	0,99	0,961	0,981	0,953
5	0,98	0,942	1	0,879	1	0,84	0,973	0,892
6	1	0,811	1	0,804	1	0,801	0,989	0,801
7	0,916	1	0,914	1	0,947	1	0,912	0,975
8	1	1	1	1	1	1	0,989	1
9	1	1	1	1	1	1	0,987	1
10	1	0,973	1	0,811	1	0,801	0,998	0,859
11	1	1	1	1	1	1	1	1
12	1	1	1	1	1	0,987	1	1
13	0,981	0,981	0,981	0,981	0,981	0,981	0,961	0,956
14	0,916	1	0,943	0,951	1	0,943	0,933	0,948
15	1	0,954	1	0,954	0,954	0,954	0,982	0,959
16	0,987	0,987	0,987	0,987	0,987	0,997	0,997	0,975
17	0,833	1	0,833	1	0,833	0,833	0,829	0,889
18	0,981	1	0,982	1	0,981	1	0,981	1
19	1	1	1	1	1	1	1	1
20	0,951	1	0,944	0,991	0,958	0,991	0,935	0,969
Média	0,969	0,981	0,966	0,965	0,963	0,953	0,964	0,956

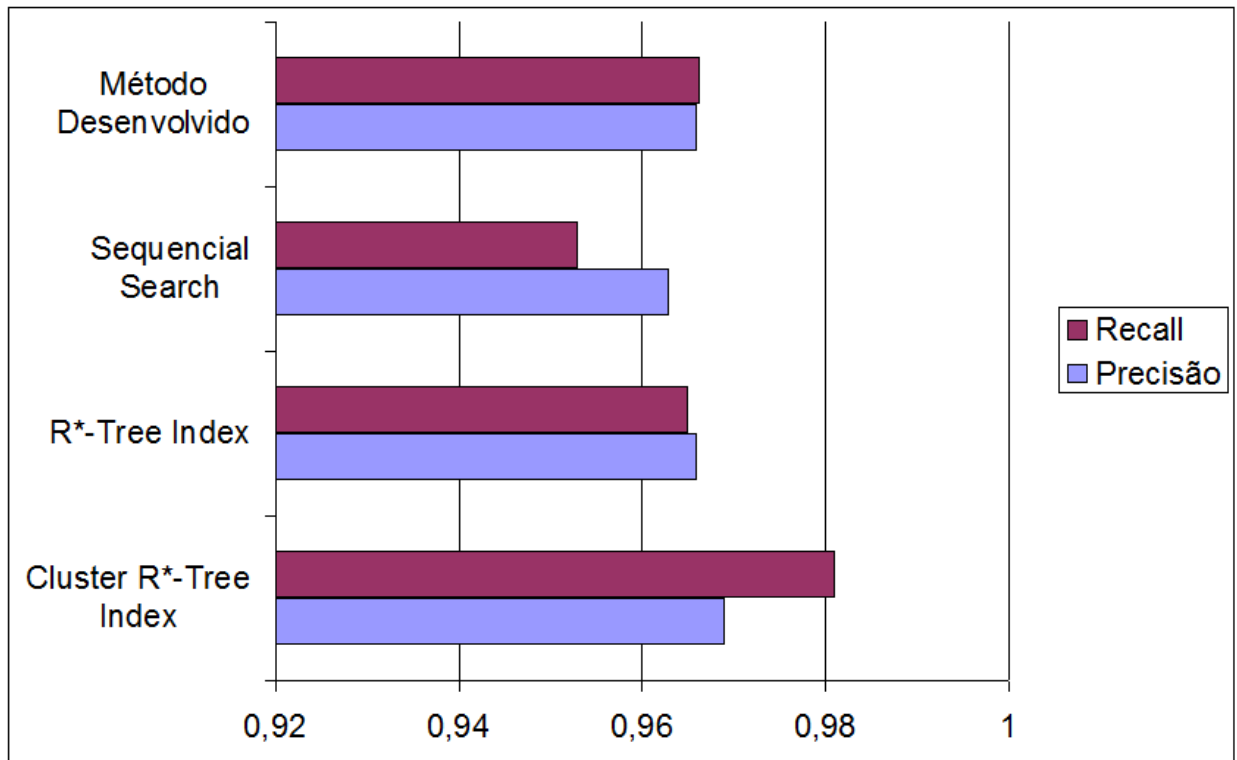


Figura 4.14: Representação gráfica dos resultados apresentados na tabela 4.5

5 COMENTÁRIOS FINAIS

5.0.2 Conclusões

Observa-se através dos experimentos efetuados e a comparação com métodos atuais (VENUGOPAL; SUDHAMANI, 2008), que os resultados são semelhantes no ponto de vista da precisão dos resultados, e superiores em complexidade computacional quando o número de imagens ultrapassa o número de 45 imagens indexadas.

Quanto à eficácia do método no que se refere ao número e qualidade de imagens retornadas por semelhança, também se percebe que os resultados são compatíveis com os resultados presentes nas publicações com mesmo fim, e com uso da mesma técnica de agrupamento de dados. Outro fator importante e que pode ser observado, é que em todos os testes efetuados, a imagem com maior *rank* retornada é a própria imagem de entrada (que está sempre inclusa no conjunto de imagens pré-processadas), o que assegura a robustez do método desenvolvido.

Complementarmente foram efetuados testes com o banco de dados de imagens utilizado na pesquisa de (WANG; LI; WIEDERHOLD, 2001) disponibilizado por *James Z. Wang's Research Group* e os resultados dos testes confirmam a robustez do método desenvolvido, ou seja, o maior número de cantos identificados como semelhantes durante a tarefa de comparação das imagens faz parte da imagem correta, ou seja, fazem parte da imagem mais semelhante. Um aspecto importante percebido durante os testes utilizando o protótipo computacional, é que o método de correção da iluminação auxiliou a classificação das imagens em dois sentidos, primeira, no sentido de evitar que o algoritmo de detecção de cantos encontrasse falsos positivos, como por exemplo, sombra na imagem, em segundo lugar, favoreceu que uns números maiores de cantos verdadeiros da imagem se tornassem melhor perceptíveis pelo algoritmo de detecção de cantos.

5.0.3 Sugestões Para Trabalhos Futuros

Para continuidade deste trabalho, considerando que o mesmo retorna regiões de interesse que são resultado do agrupamento por característica de cor, explorar descritores de imagens que consigam comparar estas regiões sob o ponto de vista da semelhança são uma boa alternativa como, por exemplo, através do uso do algoritmo MAD - *Mean Absolute Distance*, que é amplamente utilizado em processamento de imagens médicas, um exemplo pode ser visto em (ZHENG et al., 2008), onde as informações geradas por este algoritmo podem ser utilizadas de maneira complementar atuando em conjunto com os metadados já gerados pelo detector de cantos empregado neste trabalho.

REFERÊNCIAS

ADORAM, M.; LEW, M. S. Irus: Image retrieval using shape. *Multimedia Computing and Systems, International Conference on*, IEEE Computer Society, Los Alamitos, CA, USA, v. 2, p. 597, 1999. ISSN 1530-2032.

ALBUZ, E.; KOCALAR rturk; KHOKHAR, A. A. Scalable color image indexing and retrieval using vector wavelets. *IEEE Trans. on Knowledge and Data Eng*, v. 13, p. 851–861, 2001.

AMSALEG, L.; GROS, P.; BERRANI, S.-A. Robust object recognition in images and the related database problems. *Multimedia Tools and Applications*, V23, n. 3, p. 221–235, 2004. Disponível em: <<http://dx.doi.org/10.1023/B:MTAP.0000031758.46389.00>>.

ANTANI, S.; LONG, L. R.; THOMA, G. R. Bridging the gap: Enabling cbir in medical applications. *Computer-Based Medical Systems, IEEE Symposium on*, IEEE Computer Society, Los Alamitos, CA, USA, v. 0, p. 4–6, 2008. ISSN 1063-7125.

BAEZA-YATES, R. A. et al. Clustering in metric spaces with applications to information retrieval. In: *Clustering and Information Retrieval*. [S.l.: s.n.], 2003. p. 1–34.

BEAUDET, P. Rotationally invariant image operations. In: *International Joint Conference on Patter Recognition*. [S.l.: s.n.], 1978.

BERKHIN, P. *Survey Of Clustering Data Mining Techniques*. San Jose, CA, 2002. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.3739>>.

BRASDSKI G. R.; THRUN, S. [S.l.], 2006.

CARSON, C. et al. Blobworld: Image segmentation using expectation-maximization and its application to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, p. 1026–1038, 1999.

CESAR-JR., R. M. *Análise Multi-Escala de Formas Bidimensionais*. Tese (Doutorado) — IFSC - USP, São Carlos, 1997.

CHANG, N.-S.; FU, K.-S. Query-by-pictorial-example. *IEEE Trans. Softw. Eng.*, IEEE Press, Piscataway, NJ, USA, v. 6, n. 6, p. 519–524, 1980. ISSN 0098-5589.

CHANG, Y. et al. Cbir of spine x-ray images on inter-vertebral disc space and shape profiles. In: *CBMS '08: Proceedings of the 2008 21st IEEE International Symposium on Computer-Based Medical Systems*. Washington, DC, USA: IEEE Computer Society, 2008. p. 224–229. ISBN 978-0-7695-3165-6.

COMANICIU, D.; MEER, P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 24, n. 5, p. 603–619, 2002. ISSN 0162-8828.

COOPER, M. et al. Temporal event clustering for digital photo collections. *ACM Trans. Multimedia Comput. Commun. Appl.*, ACM, New York, NY, USA, v. 1, n. 3, p. 269–288, 2005. ISSN 1551-6857.

DATTA, R. et al. Image retrieval: Ideas, influences, and trends of the new age. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 40, n. 2, p. 1–60, 2008. ISSN 0360-0300.

DERICHE, R.; GIRAUDON, G. A computational approach for corner and vertex detection. *International Journal of Computer Vision*, v. 10, p. 101–124, 1992.

FERRAGINA, P.; GULLI, A. A personalized search engine based on web-snippet hierarchical clustering. In: *WWW '05: Special interest tracks and posters of the 14th international conference on World Wide Web*. New York, NY, USA: ACM, 2005. p. 801–810. ISBN 1-59593-051-5.

FLICKNER, M. et al. Query by image and video content: the qbic system. *Computer*, v. 28, n. 9, p. 23–32, 1995.

FLICKNER, M. et al. Query by image and video content: The qbic system. *Computer*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 28, n. 9, p. 23–32, September 1995. ISSN 0018-9162. Disponível em: <<http://dx.doi.org/10.1109/2.410146>>.

FORSTNER, W. *A Feature Based Correspondence Algorithm for Image Matching*. [S.l.: s.n.], 1986. III: 150-166 p.

FUKUNAGA, K.; HOSTETLER, L. The estimation of the gradient of a density function, with applications in pattern recognition. *Information Theory, IEEE Transactions on*, v. 21, n. 1, p. 32–40, 1975.

GEVREKCI, M.; GUNTURK, B. Illumination robust interest point detection. v. 113, n. 4, p. 565–571, April 2009.

GONZALES, R. A. W.; WOODS R., . *Digital Image Processing*. [S.l.: s.n.], 1993.

HAN, J.; PAN, Z. Robust moving objects detection in dynamic scenes based on mean shift. In: *CW '08: Proceedings of the 2008 International Conference on Cyberworlds*. Washington, DC, USA: IEEE Computer Society, 2008. p. 271–275. ISBN 978-0-7695-3381-0.

HARRIS, C.; STEPHENS, M. A combined corner and edge detector. In: *The Fourth Alvey Vision Conference*. [S.l.: s.n.], 1988. p. 147–151.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. [S.l.]: Springer, 2001. Hardcover. ISBN 0387952845.

HELLER, K. A.; GHAHRAMANI, Z. Bayesian hierarchical clustering. In: *ICML '05: Proceedings of the 22nd international conference on Machine learning*. New York, NY, USA: ACM, 2005. p. 297–304. ISBN 1-59593-180-5.

JAHNE, B. *Handbook of Computer Vision and Applications: Volume 2: From Images to Features*. Orlando, FL, USA: Academic Press, Inc., 1999. ISBN 0123797721.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 31, n. 3, p. 264–323, 1999. ISSN 0360-0300.

KALPATHY-CRAMER, J.; HERSH, W. Effectiveness of global features for automatic medical image classification and retrieval - the experiences of ohsu at imageclefmed. *Pattern Recogn. Lett.*, Elsevier Science Inc., New York, NY, USA, v. 29, n. 15, p. 2032–2038, 2008. ISSN 0167-8655.

KIM, D.-H.; CHUNG, C.-W.; BARNARD, K. Relevance feedback using adaptive clustering for image similarity retrieval. *J. Syst. Softw.*, Elsevier Science Inc., New York, NY, USA, v. 78, n. 1, p. 9–23, 2005. ISSN 0164-1212.

KIM, D. hwan. Qcluster: Relevance feedback using adaptive clustering for content-based image retrieval. In: *In Proc. of the ACM SIGMOD Int. Conf. on Management of Data*. [S.l.: s.n.], 2003. p. 599–610.

KINOSHENKO, D. et al. Hierarchical partitions for content image retrieval from large-scale database. In: *MLDM*. [S.l.: s.n.], 2005. p. 445–455.

KITCHEN, L.; ROSENFELD, A. Gray level corner detection. v. 1, n. 2, p. 95–102, December 1982.

KOMPATSIARIS, Y. Multimedia semantic analysis technologies and their potential uses. 2006. Disponível em http://www.samt2006.org/presentationsITI_MMAnalysis.pdf Acesso em 2009. 2006.

KOSARA, R.; MIKSCH, S. *Visualization Methods for Data Analysis and Planning In Medical Applications and*. 2002.

LEW, M. S. et al. Content-based multimedia information retrieval: State of the art and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, ACM, New York, NY, USA, v. 2, n. 1, p. 1–19, 2006. ISSN 1551-6857.

LI, C.-Y.; SHIH, M.-C.; HSU, C.-T. Image retrieval with relevance feedback based on graph-theoretic region correspondence estimation. In: *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 3*. Washington, DC, USA: IEEE Computer Society, 2004. p. 842–845. ISBN 0-7695-2128-2.

LIU, H.; HE, G. Shape feature extraction of high resolution remote sensing image based on susan and moment invariant. In: *CISP '08: Proceedings of the 2008 Congress on*

Image and Signal Processing, Vol. 2. Washington, DC, USA: IEEE Computer Society, 2008. p. 801–807. ISBN 978-0-7695-3119-9.

LIU, Y. et al. A survey of content-based image retrieval with high-level semantics. *Pattern Recogn.*, Elsevier Science Inc., New York, NY, USA, v. 40, n. 1, p. 262–282, 2007. ISSN 0031-3203.

LUX, M.; GRANITZER, M.; KLIEBER, W. Caliph & emir: Semantics in multimedia retrieval and annotation. In: CODATA2004. 2004. Disponível em: <<http://www.semanticmetadata.net/publications/mlux-codata2004.pdf>>.

MA, W. ying; MANJUNATH, B. Netra: A toolbox for navigating large image databases. In: *Multimedia Systems*. [S.l.: s.n.], 1999. p. 568–571.

MA, W. ying; MANJUNATH, B. Netra: A toolbox for navigating large image databases. In: *Multimedia Systems*. [S.l.: s.n.], 1999. p. 568–571.

MIRMEHDI, M.; PERISSAMY, R. Perceptual image indexing and retrieval. *Journal of Visual Communication and Image Representation*, v. 13, n. 4, p. 460–475, December 2002. ISSN 1047-3203.

MOKHTARIAN, F.; SUOMELA, R. Robust image corner detection through curvature scale space. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 20, n. 12, p. 1376–1381, 1998. ISSN 0162-8828.

MORAVEC, H. Towards automatic visual obstacle avoidance. In: *Proceedings of the 5th International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 1977.

PENTLAND, A.; PICARD, R.; SCLAROFF, S. *Photobook: Content-based Manipulation of Image Databases*. 1996. Disponível em: <citeseer.ist.psu.edu/pentland95photobook.html>.

PLATANIOTIS, K.; ANDROUTSOS, D.; VENETSANOPOULOS, A. Adaptive multi-channel filters for colour image processing. v. 11, n. 3, p. 171–177, January 1998.

RAO, S.; MARTINS, A. de M.; PRÍNCIPE, J. C. Mean shift: An information theoretic perspective. *Pattern Recognition Letters*, v. 30, n. 3, p. 222 – 230, 2009. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/B6V15-4TN0KR0-1/2/7d4f2bf5f2d85adbcbcaf89dc3fb38d04>>.

RUI, Y.; HUANG, T. S.; MEHROTRA, S. Content-based image retrieval with relevance feedback in mars. In: *In Proc. IEEE Int. Conf. on Image Proc.* [S.l.: s.n.], 1997. p. 815–818.

SAHOO, N. et al. Incremental hierarchical clustering of text documents. In: *CIKM '06: Proceedings of the 15th ACM international conference on Information and knowledge management*. New York, NY, USA: ACM, 2006. p. 357–366. ISBN 1-59593-433-2.

SHOTTON, J.; BLAKE, A.; CIPOLLA, R. Multiscale categorical object recognition using contour fragments. v. 30, n. 7, p. 1270–1281, July 2008.

SHYU, M.-L. et al. A unified framework for image database clustering and content-based retrieval. In: *MMDB '04: Proceedings of the 2nd ACM international workshop on Multimedia databases*. New York, NY, USA: ACM, 2004. p. 19–27. ISBN 1-58113-975-6.

SMITH, J. R.; CHANG, S. fu. Querying by color regions using the visualseek content-based visual query system. In: *in Intelligent Multimedia Information Retrieval*. [S.l.]: AAAI Press, 1996. p. 23–41.

SMITH, S. M.; BRADY, J. M. Susan - a new approach to low level image processing. *International Journal of Computer Vision*, v. 23, p. 45–78, 1995.

SWAIN, M.; BALLARD, D. Indexing via color histograms. In: *Proc. Third International Conference on Computer Vision*. [S.l.: s.n.], 1990. p. 390–393.

TOBIN, K. et al. Automated feature generation in large-scale geospatial libraries for content-based indexing. v. 72, n. 5, p. 531–540, May 2006.

TORRES, R. D. S.; FALCÃO, A. X. *Content-Based Image Retrieval: Theory and Applications: Revista de Informatica Teorica e Aplicada*. 2006. 161-185 p.

TORRES, R. da S. et al. Recuperação de Imagens: Desafios e Novos Rumos. In: *XXXV Seminário Integrado de Software e Hardware (SEMISH)*. Belém: [s.n.], 2008.

TRAINA, A. J. M.; MARQUES, J.; JR., C. T. Fighting the semantic gap on cbir systems through new relevance feedback techniques. In: *CBMS '06: Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems*. Washington, DC, USA: IEEE Computer Society, 2006. p. 881–886. ISBN 0-7695-2517-1.

TRAJKOVIC, M.; HEDLEY, M. Fast corner detection. v. 16, n. 2, p. 75–87, February 1998.

TRIGGS, B. Detecting keypoints with stable position, orientation and scale under illumination changes. In: *European Conference on Computer Vision*. [s.n.], 2004. p. IV 100–113. Disponível em: <<http://lear.inrialpes.fr/pubs/2004/Tri04>>.

TUCERYAN, M.; JAIN, A. K. Texture analysis. World Scientific Publishing Co., Inc., River Edge, NJ, USA, p. 235–276, 1993.

VELTKAMP, R. C.; TANASE, M. *Content-Based Image Retrieval Systems: A Survey*. [S.l.], October 2002.

VENUGOPAL, C. R.; SUDHAMANI, M. V. Image retrieval from databases: an approach using region color and indexing technique. In: *HPCNCS*. [S.l.: s.n.], 2008. p. 82–89.

WANG, J. Z.; LI, J.; WIEDERHOLD, G. Simplicity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 23, p. 947–963, 2001.

WANG, L.-Y. Y. et al. Msb: A mean-shift-based approach for the analysis of structural variation in the genome. *Genome research*, November 2008. ISSN 1088-9051. Disponível em: <<http://dx.doi.org/10.1101/gr.080069.108>>.

WANG W., D. R. Evaluation of image corner detectors for hardware implementation'. In: *Canadian Conference on Electrical and Computer Engineering, May 2-5, 2004, Niagara Falls, Ontario*. [S.l.: s.n.], 2004.

WEIJER, J. van de; GEVERS, T.; GEUSEBROEK, J.-M. Edge and corner detection by photometric quasi-invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Los Alamitos, CA, USA, v. 27, n. 4, p. 625–630, 2005. ISSN 0162-8828.

WEN, F.; CHEN, S.; TANG, X. Easytoon: cartoon personalization using face photos. In: *MM '08: Proceeding of the 16th ACM international conference on Multimedia*. New York, NY, USA: ACM, 2008. p. 1021–1022. ISBN 978-1-60558-303-7.

XIE, Z.; ROBERTS, C.; JOHNSON, B. Object-based target search using remotely sensed data: A case study in detecting invasive exotic australian pine in south florida. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 63, n. 6, p. 647 – 660, 2008. ISSN 0924-2716. Disponível em: <<http://www.sciencedirect.com/science/article/B6VF4-4SM0XFY-2/2/7985dfce04dbf9bc64f4766b483e3ead>>.

XU, Z.; XIAOWEI, Z. X.; TRESP, V. A hybrid relevance-feedback approach to text retrieval. In: *In Proceedings of the 25th European Conference on Information Retrieval Research, Lecture Notes in Computer Science*. [S.l.]: Springer-Verlag, 2003. p. 281–293.

ZHANG, D.; LU, G. Review of shape representation and description techniques. *Pattern Recognition*, Elsevier, v. 37, n. 1, p. 1–19, 2004.

ZHAO, C.; KNIGHT, A.; REID, I. Target tracking using mean-shift and affine structure. In: *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. [s.n.], 2008. p. 1–5. Disponível em: <<http://dx.doi.org/10.1109/ICPR.2008.4761914>>.

ZHAO, Y. et al. Hierarchical clustering algorithms for document datasets. In: *Data Mining and Knowledge Discovery*. [S.l.]: ACM Press, 2002. p. 515–524.

ZHENG, B. et al. 6-dof pose estimation from single ultrasound image using 3d ip models. *Computer Vision and Pattern Recognition Workshop*, IEEE Computer Society, Los Alamitos, CA, USA, v. 0, p. 1–8, 2008.

ZHENG, Z.; WANG, H.; KHWANG. Analysis of gray level corner detection. *Pattern Recognition Letters*, v. 20, n. 2, p. 149–162, February 1999. Disponível em: <[http://dx.doi.org/10.1016/S0167-8655\(98\)00134-2](http://dx.doi.org/10.1016/S0167-8655(98)00134-2)>.

ZHOU, X. S.; HUANG, T. S. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, v. 8, n. 6, p. 536–544, April 2003. Disponível em: <<http://dx.doi.org/10.1007/s00530-002-0070-3>>.

ZOU, L.-h. et al. The comparison of two typical corner detection algorithms. In: *IITA '08: Proceedings of the 2008 Second International Symposium on Intelligent Information Technology Application*. Washington, DC, USA: IEEE Computer Society, 2008. p. 211–215. ISBN 978-0-7695-3497-8.