Capítulo 3 – Recuperação de Imagens em Base de Dados

3.1 Considerações Iniciais

Ao se elaborar um *Software*, para qualquer finalidade, algumas observações devem ser levadas em consideração para que o resultado final do mesmo atinja os objetivos iniciais. Várias dúvidas surgem durante todo o processo de desenvolvimento e conseqüentemente, isso gera uma série de estudos, análises, avaliações e definições até o término do projeto.

Todas essas preocupações durante o processo de desenvolvimento do *Software* afetam no seu resultado final, podendo o mesmo ser finalizado com sucesso ou não. Alguns pontos críticos que devem ser previstos antecipadamente para se evitar imprevistos indesejáveis são: o entendimento completo em relação ao assunto do projeto, ferramentas de desenvolvimento e banco de dados que serão utilizados, domínio da tecnologia adotada, entre outros.

Uma das dificuldades que se enfrenta no dia a dia, apesar do desenvolvimento tecnológico atual é o tempo de processamento e a resposta gerada pelo sistema, onde se está sempre querendo minimizar esse tempo.

Além desse problema de tempo de processamento e resposta, tem-se um outro aspecto a se destacar: o sistema deve fornecer resultados de uma forma simples, clara e objetiva. O fato não é apenas fornecer um resultado ligeiramente rápido, mas, além disso, fornecer resultados confiáveis, precisos e úteis. Muitas vezes, os sistemas gerenciadores de bancos de dados (SGBD) trabalham com uma grande quantidade de informações, sendo essas informações armazenadas em algum espaço físico e para o acesso entre o

solicitado ao banco e o retorno, gera-se um tempo de espera. Uma base de dados mal elaborada também pode contribuir para o aumento do tempo.

Um dos cuidados para se evitar tais problemas é saber gerar uma chave que se consiga identificar no meio de uma grande base de dados a informação certa de forma eficiente. Com uma boa definição dessas chaves, consegue-se aprimorar os problemas citados.

Um sistema CBIR típico não trata somente das várias fontes de informações tais como: texto, imagem, vídeo, mas também da necessidade do usuário. Basicamente analisa as informações da pesquisa, *query* do usuário e combina essas informações para a recuperação dos dados mais relevantes. As funções principais são:

1. Analisar o conteúdo das informações de origem e representar o conteúdo da análise de uma maneira apropriada que será combinada com a *query* informada pelo usuário. Nesse processo consome-se um bom tempo de processamento, pelo fato da pesquisa ser feita seqüencialmente na base de dados. Portanto, isso é feito somente uma vez podendo ser executado *off-line*.
2. Análise da *query* do usuário e representação apropriada para pesquisa no banco de dados. Parte desta etapa é similar à etapa precedente, mas aplicada somente para a imagem de centro de consulta.
3. Definir uma estratégia para combinar as buscas com a informação armazenada em uma base de dados. Recuperar a informação que é relevante de uma forma eficiente.
4. Fazer ajustes necessários no sistema (geralmente parâmetros) baseado no retorno dos usuários e/ou das imagens recuperadas.

3.2 Recuperação de Imagens por Conteúdo em Base de Dados

Tem-se visto um crescimento muito rápido das coleções de imagens digitais. Todos os dias equipamentos militares e civis geram gigabytes de

imagens. Em conseqüência, não se pode acessar ou fazer uso da informação sem que esta esteja organizada, permitindo sua navegação, pesquisa e recuperação de modo eficiente. A recuperação de imagens representa uma área de pesquisa muito ativa desde a década de 1970 e tem recebido a contribuição de duas grandes linhas de pesquisas: o gerenciamento de banco de dados e a visão computacional. Estas estudam a recuperação de imagens sob dois ângulos diferentes, um baseado em textos e o outro baseado no conteúdo visual (RUI; HUANG; CHANG, 1999).

Recuperação de imagens por conteúdo ou CBIR *(Content Based Image Retrieval)* é um método de pesquisa em base de dados, utilizando uma imagem de consulta *(query image)* (JACOBS; FINKELSTEIN; SALESIN, 1995).

A imagem de consulta é o centro de interesse desse método. Os sistemas de Recuperação de Imagens Baseada em seu Conteúdo (*Content Based Image Retrieval* - CBIR) permitem recuperar imagens considerando o conteúdo das mesmas. Dada uma imagem de consulta, o objetivo de um sistema CBIR é procurar numa base de dados e recuperar uma determinada quantidade de imagens similares, mais próximas em comparação com a imagem da consulta, segundo um critério pré-determinado.

Um vetor de características é uma representação numérica sucinta de uma imagem ou parte dela, descrevendo seus detalhes mais representativos, segundo um critério especificado.

Essa nova representação da imagem pode ser armazenada em uma base de dados, e assim, agilizar o processo de recuperação de imagens.

3.3 Trabalhos Correlatos

Nos últimos tempos, aumentou-se o número de imagens disponibilizadas na *Internet* devido ao seu desenvolvimento. A diversidade dessas imagens tem por objetivo atrair pessoas das mais variadas áreas de atuação, tais como: geografia, medicina, arquitetura, informática, publicidade, dentre outras. Entretanto o estudo de recuperação de imagens concentra-se na eficiência e

eficácia em acessar as mais amplas e variadas coleções de imagens e recuperá-las, sendo esse o maior desafio (WANG, 2001).

Os métodos de recuperação de imagens localizam a imagem desejada por uma chave de busca que são as assinaturas para cada uma das imagens (FLANK et al, 1995).

3.3.1 *Query By Image Content* (QBIC)

Os primeiros sistemas de recuperação de imagem baseada em conteúdo foram desenvolvidos no início dos anos 80 (CHANG; FU, 1980), sendo que a maioria das pesquisas nesta área inspirou-se no sistema da IBM (*Query By Image Content* - QBIC) (FLICKNER, 1980).

O QBIC foi o primeiro sistema CBIR comercial e sua concepção teve profundos efeitos nos sistemas de recuperação de imagens que foram desenvolvidos após o seu lançamento. O QBIC suporta consultas baseadas em imagens exemplo, esboços, desenhos e com seleção de cores e texturas, etc. As características de cor utilizadas pelo QBIC são RBG, Yiq, Lab e coordenadas MTM (*Mathematical Transform to Munsell*). Utiliza também os histogramas de cores. Para textura utiliza uma versão aperfeiçoada da representação de Tamura, que de modo geral, é uma combinação de regularidade, contraste e direcionalidade. As características de forma consistem de área, circularidade, excentricidade, orientação do eixo maior e um conjunto de momentos invariantes. O QBIC é um dos poucos sistemas que conta com índices para atributos de alta dimensionalidade (ERPEN, 2004).

No subsistema de índice, utiliza primeiramente a KLT (*Karhunen- Loeve Transform*) para realizar a redução de dimensionalidade e então uma *R-tree* como uma estrutura de índice multi-dimensional. A KLT é um método bastante utilizado para a redução de dimensionalidade e é a base da Análise dos Componentes Principais, empregada amplamente em diversas áreas, tais como o sensoriamento remoto onde os dados possuem dimensionalidade extremamente elevada. Nos novos sistemas, consultas textuais baseadas em

palavras-chave podem ser combinadas com consultas baseadas em similaridade de conteúdo (ERPEN, 2004).

3.3.2 CBIR utilizando Histogramas Locais

Na pesquisa de Wang (2001), a cor é uma das principais características utilizadas em CBIR. Muitos estudos e teses apontaram a cor como uma característica eficiente para a recuperação de imagens coloridas.

Os olhos humanos são sensíveis às cores, pois estas são uma das mais importantes e principais características que as pessoas utilizam para reconhecer uma imagem. As características das cores podem prover informações para categorizar imagens e os Histogramas de Cores são geralmente utilizados para representar as características da cor de uma imagem (PASS; ZABIH; MILLER, 1996).

A técnica utilizada por Wang (2001) para se fazer a Recuperação da Imagem Baseada em Conteúdo resume-se em gerar um Histograma Geral da Imagem. Uma *grid* sobreposta à imagem é utilizada e para cada célula da mesma é gerado um Histograma Local. Além dessas informações têm-se também a extração de características de forma e textura, visto que o histograma já contém as informações globais e locais referentes às cores da imagem. Com as informações dos histogramas mede-se a similaridade dos objetos respectivos da imagem. Todos esses dados são armazenados como um vetor de características e através dele é feita a pesquisa dentro de uma base de dados.

O autor propõe um novo método, o qual nomeou *Harbin*, baseado nas características da cor. Duas técnicas baseadas na cor são o enfoque principal do *Harbin*: Histogramas Global e Local da cor na imagem. Os dois histogramas foram desenvolvidos para minimizar o problema de distância entre os objetos.

Com os dois histogramas, reduz-se o problema para calcular a distância entre os objetos das imagens e o problema de encontrar a mínima distância entre os mesmos.

O autor também relata em seu trabalho que as características da textura referem-se ao padrão visual da região da imagem. Imagens com intensidades de cores similares podem ter muita diferença na textura. A característica da textura pode freqüentemente ser usada para distinguir a distribuição de cores similares entre as imagens, como, por exemplo, o céu e o mar, uma folha e a grama.

O relacionamento espacial entre os objetos pode conter informações importantes para a recuperação da imagem. A estrutura baseada no objeto recupera imagens pela similaridade da posição dos objetos na imagem, mas não são explicitamente armazenados seus relacionamentos. A estrutura representativa inclui uma *Grid, Quad-Tree e R-Tree*, por conseguinte, a estrutura é baseada na relação como uma representação simbólica da imagem, conforme a figura 3.1.

Na figura (a) observa-se a forma do objeto, na figura (b) a representação da textura de uma imagem e por último, na figura (c), observa-se o relacionamento espacial da imagem. Divide-se a imagem em *Grid* e obtêm-se uma informação geral do que contém a imagem.

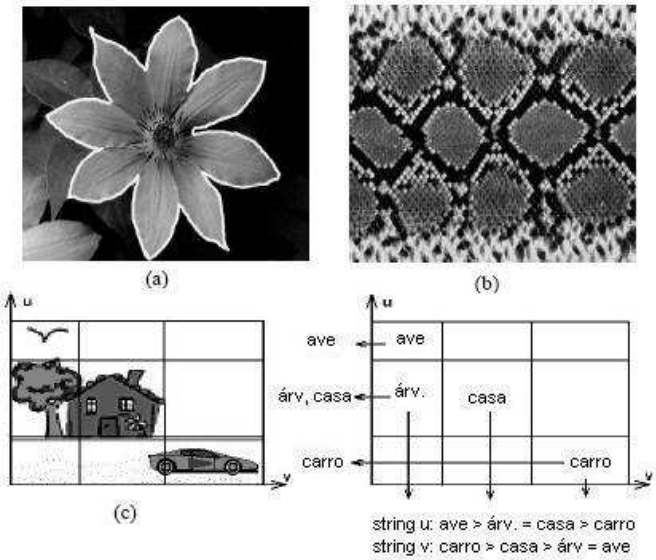

Figura 3.1: (a) Forma, (b) Textura e (c) Relacionamento espacial

A forma de um objeto contém informações importantes de uma imagem. A recuperação pela forma sempre requer uma robusta deteção de borda e algoritmos de segmentação para detectar objetos ou regiões de fronteiras.

Características da textura referem-se ao padrão visual das regiões da imagem. Por essa razão, muitas técnicas de reconhecimento de padrões podem ser usadas para recuperação pela textura. Imagens com intensidade de cores similares podem possuir texturas diferentes; a textura pode ser usada para distinguir entre as imagens com similaridades na distribuição da cor. Algumas das medidas utilizadas para computar a similaridade da textura são o contraste, aspereza e simetria.

A estrutura representativa da *Grid* bem como a *R-Tree* tem como objetivo preservar a representação espacial dos objetos.

Muitas tecnologias baseadas nestas características foram propostas no século XX. Cada uma delas é apropriada para uma consulta específica, ao mesmo tempo, todas estas tecnologias têm suas limitações. Por exemplo, ao procurar-se por aviões usando uma simples imagem de um avião branco, a recuperação pela forma pode retornar bons resultados. Se utilizar a recuperação pela cor, desde que o conjunto de informações de respostas inclua uma variedade de imagens contendo elementos brancos ou prata, os resultados podem ser bem precários. De outra forma, combinando estas características, os sistemas CBIR podem possuir informações mais elaboradas, com mais detalhes. Por exemplo, a *query* "buscar imagens que contenham uma esfera vermelha" pode ser apresentada pela combinação da característica forma como "círculo" e a característica cor como "vermelho". Entretanto é muito difícil encontrar uma maneira geral, usando combinações destas características primitivas (BIMBO, 1999).

Sistemas CBIR requerem métodos que são baseados na primitiva de comparar a similaridade ou diferença entre uma imagem de exemplo (imagem de entrada ou *query-by-example*) e todas as imagens de uma base de dados.

Entretanto, as similaridades ou as diferenças entre as imagens não podem ser quantificadas de uma forma ideal. A extensão em que as imagens são similares mudará quando a *query* variar. Por exemplo, no caso das duas

imagens, uma com o céu azul da manhã e outra com uma montanha verde e com o céu da manhã, quando o levantar do sol é considerado, a similaridade entre as duas imagens pode ser alta, mas se o objeto de interesse for o céu azul, a similaridade entre essas duas imagens pode ser baixa. É muito difícil encontrar um método que meça a similaridade ou a diferença entre as imagens precisamente para todos os tipos de *query*. Em outras palavras, todos os métodos de recuperação têm suas próprias limitações. Um outro exemplo de dificuldade para a recuperação de imagens baseada na cor é diferenciar uma imagem com o céu azul e outra com o mar azul, tudo isso depende da *query* a ser utilizada.

Conseqüentemente, ao avaliar uma tecnologia CBIR, deve-se lembrar que a eficácia da recuperação dessa tecnologia depende dos tipos de exigências da *query* feita.

As imagens utilizadas para os testes no trabalho de Wang (2001) foram as dos tipos: JPEGs, GIFs, BMPs tendo sido utilizado o modelo de cor RGB. Embora só tenham sido realizados testes com o modelo de cor RGB, poder-se-ia implementar e testar em outros modelos conhecidos, para verificar em quais se obteria um melhor resultado.

Na quantização do espaço de cor, reduziu-se o espaço de todas as cores possíveis em um conjunto de cores discretas. Dividiu-se o cubo do modelo de cores RGB em sub-cubos e cada sub-cubo foi representado por uma única cor. Dividiu-se, ao todo, em 64 sub-cubos (4 sub-divisões), separando os 3 eixos: Verde, Vermelho e Azul (R,G e B).

O Histograma da cor é um conjunto de *bins,* onde cada *bin* representa a probabilidade do *pixel* na imagem ser de uma cor específica. O Histograma da cor (H) dada uma imagem é definido como um vetor:

$$H = \{H[0], H[1], H[2], ..., H[i], H[N]\} \quad (3.1)$$

Onde: i representa a cor no Histograma e corresponde a um sub-cubo no modelo RGB. H[i] é o número de *pixels* na cor i na imagem, e N é o número de *bins* no Histograma, isto é, o número das cores adotado no modelo de cor.

Tipicamente, cada *pixel* da imagem será atribuído para um *bin* do Histograma, assim para o Histograma o valor de cada *bin* é o número de *pixels* que têm a mesma correspondência de cor. A fim de comparar imagens de diferentes tamanhos, os Histogramas devem ser normalizados.

Uma quantização ideal do espaço de cor presume que as cores distintas não devam ficar situadas no mesmo sub-cubo e as cores similares devem ser atribuídas ao mesmo sub-cubo. Usar poucas cores diminuirá a possibilidade das cores similares serem atribuídas para *bins* diferentes, mas aumenta a possibilidade das cores distintas serem atribuídas ao mesmo *bin*, e o conteúdo de informação das imagens diminuirá. Os Histogramas da cor com um grande número de *bins* conterão mais informação sobre o conteúdo das imagens, diminuindo, assim, a possibilidade de cores distintas serem atribuídas aos mesmos *bins*, aumentando, entretanto, a possibilidade de cores similares serem atribuídas aos *bins* diferentes.

Wang (2001) utilizou a distância de Minkowski para comparar a distância das cores do mesmo *bin* entre os Histogramas das cores. A figura 3.2 ilustra o processo.

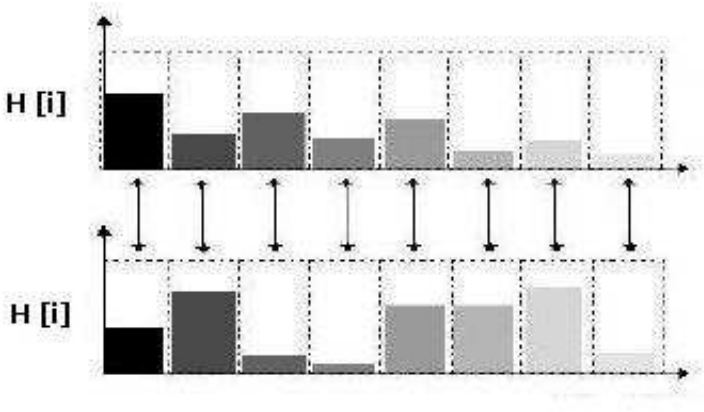

Figura 3.2: Métrica da distância de *Minkowski* (WANG, 2001)

O Histograma Global da Cor representa a imagem total como um simples Histograma e o Histograma Local da Cor divide e fixa a imagem em

blocos, e de cada bloco obtém-se um novo Histograma da Cor. O Histograma Global não captura com detalhes o conteúdo da imagem, mas o Histograma Local contém mais informações e é melhor para distinguir as cores entre as regiões da imagem a serem comparadas.

As imagens da figura 3.3 mostram exemplos de Histogramas Globais.

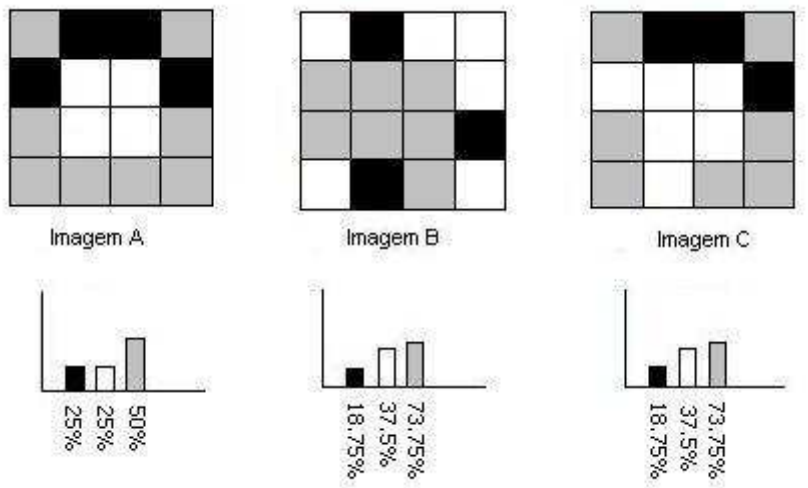

Figura 3.3: Histogramas Globais de Imagens

Como pode-se observar no gráfico da figura 3.3, cada *bin* corresponde a uma cor, na seqüência: preto, branco e cinza. Observa-se que o Histograma da Imagem A possui (25%, 25% e 50%) e o Histograma das Imagens B e C possuem os mesmos valores de (18.75%, 37.5% e 73.75%) *pixels*, apesar de serem imagens diferentes.

O Histograma Global da Cor (*Global Color Histogram* - GCH) é um método tradicional para a Recuperação da Imagem Baseada na Cor, entretanto, não inclui a informação da distribuição da cor nas regiões, sendo esse um ponto fraco dessa metodologia. A distância entre as imagens às vezes pode não mostrar a diferença real entre elas, por exemplo, a distância entre as imagens A e C deve ser menor do que a distância entre as imagens A e B, mas usando o GCH obtém-se a mesma distância. Além disso, no exemplo de um GCH, é possível para duas imagens diferentes terem uma distância muito curta entre seus Histogramas (como no exemplo da figura 3.3 em que a imagem B e a imagem C são tratadas como a mesma).

O Histograma Local da Cor (*Local Color Histogram* - LCH), inclui informações a respeito da distribuição da cor nas regiões. A primeira etapa é segmentar a imagem em blocos e obter um Histograma para cada bloco. Uma imagem será representada por estes Histogramas. Ao se comparar duas imagens, calcula-se a distância usando os Histogramas, entre uma região em uma imagem e uma região na mesma posição na outra imagem. A distância entre as duas imagens será determinada pela soma de todas estas distâncias. No entanto, se houver rotação ou translação dos objetos dentro da imagem, o LCH falha.

3.3.3 Protótipo de um Sistema de Recuperação de Imagens Baseado na Cor

No trabalho de Camillo e outros (2004), uma outra técnica é utilizada para a recuperação de Imagens em Base de Dados com base na cor. Este trabalho armazena informações de regiões com a mesma cor na imagem (regiões cromáticas), o tamanho, posição e limites com outras regiões, sendo utilizado para isso, os conceitos de grafos. A recuperação da imagem é feita através de uma medida de similaridade baseada na comparação dos grafos. O protótipo permite dois tipos de consultas:

- Imagem para consulta, fornecida pelo usuário (há a interação do usuário com o sistema).
- Definição, pelo usuário, das cores de uma imagem para consulta.

O protótipo analisa as regiões cromáticas de uma imagem, suas adjacências e variações espaciais (modelo estatístico para obter a distribuição da cor no domínio espacial). Para a indexação das imagens utilizou-se conceitos da *R-Tree*, estrutura capaz de indexar dados multi-dimensionais.

Os autores dividiram o projeto em cinco partes, como segue:

- **Primeira Parte:** Gerencia o acesso ao arquivo da imagem, realizando conversão do formato da imagem, manipulação do

modelo de cor utilizado e armazenamento das informações na base de imagens.

- **Segunda Parte:** Obtém a assinatura das imagens. Trata-se de uma representação da imagem em um Sistema de Recuperação de Imagens Baseada no conteúdo. Aplica-se um filtro para remover os ruídos que existem próximos das adjacências das regiões cromáticas, tornando as regiões uniformes e as transições entre as regiões mais aguçadas, proporcionando a geração de um grafo mais preciso.

  Cada nó do grafo representa uma cor no modelo RGB da imagem, contendo a quantidade de *pixels* da sua cor e cada aresta entre os nós a e b representará a adjacência espacial de a e b, que significa, o número de *pixels* na cor a que são vizinhos dos *pixels* na cor b, considerando as adjacências de cores baseadas na conectividade oito. Uma janela 3 x 3 é aplicada em todos os *pixels* de cada região e cada *pixel* contado na região vizinha (vizinhança 8) é adicionado ao nó correspondente.

  As informações necessárias para elaborar este grafo são: as cores da imagem, a quantidade de pixels de cada cor e as coordenadas de cada *pixel*, para cada uma das cores. Cada nó contém a própria variância espacial da cor e a variância espacial relativa das duas cores. Cada nó contém a própria variância das cores denotadas e cada borda tem como seus atributos a variância espacial entre as cores denotadas pelos nós ligados.

- **Terceira Parte:** Realiza a indexação dos atributos extraídos no módulo anterior. Geram-se índices utilizando uma *R-Tree* (estrutura de dados capaz de indexar dados multidimensionais).

- **Quarta Parte:** É um servidor de consultas do usuário *on-line* baseado no conteúdo da imagem. O módulo de recuperação recebe do módulo de interface do usuário uma imagem para consulta. O módulo de recuperação faz uso do módulo de indexação para extrair os atributos dessa imagem fornecida como

consulta. A partir da obtenção do vetor de atributos desta imagem, o módulo de recuperação utiliza seu módulo comparador para procurar na base de assinaturas, imagens que correspondam ao vetor de atributos da imagem consultada. Depois de obtidas tais imagens, utiliza-se o módulo de classificação para ordenar os resultados em função do resultado da aplicação de uma métrica de similaridade. Por fim, o módulo de recuperação envia os resultados ao módulo de interface do usuário.

- **Quinta Parte:** É formado por programas clientes que permitem a formulação de consultas baseadas no conteúdo. Estas consultas podem ser realizadas de várias maneiras: o usuário fornece a porcentagem de uma determinada cor na imagem; o usuário fornece uma imagem para consulta e, por último, o usuário constrói uma imagem através do preenchimento de cores.

Seria interessante que o protótipo desenvolvido pelos autores pudesse ser implementado com outros modelos de cores e comparado seus resultados, visto que cada modelo de cor tem sua particularidade própria. Uma idéia seria testar o protótipo com os conceitos e definições das onze cores culturais, a serem utilizadas em nossa pesquisa verificando-se o impacto na redução do número de cores no resultado final.

A figura 3.4 mostra um grafo utilizado para a recuperação da imagem através de uma medida de similaridade baseada na comparação desses grafos.

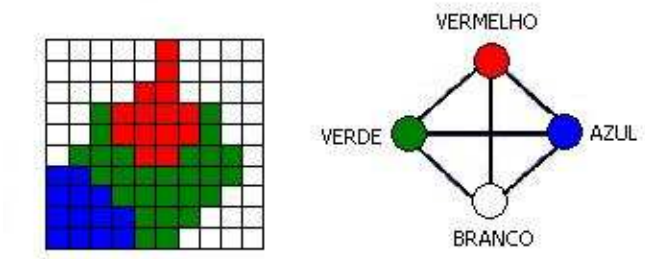

Figura 3.4: Grafo da Imagem (CAMILLO et al, 2004)

Segundo os autores os resultados são bons, salientando que algumas melhorias para o trabalho seriam diminuir o número de cores utilizadas, fazer testes com o modelo de cor HSV e possibilitar outros tipos de *query's*.

3.3.4 Recuperação de imagens com descritores estatísticos da cor

O termo "recuperação de informação" foi inventado em 1952 e ganhou popularidade na comunidade de pesquisa em 1961 (JONES; WILLETT, 1977).

O conceito de um sistema de recuperação de informação nada mais é que armazenar e localizar tais dados de uma forma eficiente e confiável. É composto de um conjunto de componentes de interação, onde projeta-se uma função específica e todos estes componentes são relacionados para conseguir um objetivo, ou seja, recuperar a informação.

O foco geral do trabalho de Jones e Willett(1977) é extrair características da cor de uma imagem e Recuperar a Imagem Baseada no Conteúdo em uma base de dados.

Alguns métodos mencionados por Jones e Willett(1977) para a Recuperação de Imagens Baseada no Conteúdo são:

- ***Query* por Exemplo** (*Query by example* - QBE): Neste tipo de *query* o usuário especifica uma imagem para pesquisa e é comparada com as imagens armazenadas no banco de dados. A imagem *query*, pode ser qualquer tipo de imagem.
- ***Query* por Característica** (*Query by Feature* - QBF): São especificadas as características importantes para se fazer a busca. Exemplo, o usuário solicita que sejam recuperadas as imagens que possuam 25% de *pixels* amarelos no quadrante esquerdo. Para esse tipo de *query* necessita-se de um histograma.

- ***Query* Baseada em Atributos** (*Query by Attribute* - QBA): Esse tipo de representação envolve um nível elevado de abstração e é difícil concluir totalmente com métodos automatizados porque uma imagem contém uma vasta informação que dificulta o uso de palavras chaves.

A extração de características é a base do CBIR. No geral, as características incluem as características baseadas no texto (palavras chaves, anotações) e características visuais (cor, textura, forma, faces). Dentro de um escopo visual, as características podem ser classificadas como de baixo e alto nível. Por causa da subjetividade da percepção, não existe uma única e melhor apresentação para uma característica dada, porque para cada característica dada existem diversas representações que caracterizam a característica para diferentes perspectivas.

A cor é a primeira característica visual para a indexação e recuperação de imagens (SWAIN; BALLARD, 1991).

Uma imagem tipicamente colorida adquirida através de câmera digital ou através de *download* da *Internet* tem normalmente três canais de cor (as imagens com tons de cinza têm somente um canal). Os valores desta tri-dimensionalidade da imagem colorida, entretanto, não descrevem a cor exata na imagem, mas a posição desses *pixels* no espaço de cor. Os *pixels* com valores de (1,1,1) terão cor diferentes em espaços de cor diferentes. Desse modo a descrição de uma imagem colorida, consiste da informação do espaço bi-dimensional que informa a cor do *pixel* no domínio espacial e no domínio tri-dimensional tem-se a informação da cor do *pixel* no espaço da cor.

Os sinais uni ou bidimensionais da cor são usadas extensamente em CBIR nas aplicações onde a robustez na captura de imagens é importante. A informação da Matiz foi usada em aplicações onde somente as diferenças entre materiais dos objetos na cena são importantes. Mostrou-se em Gevers e Smeulders(2000) que a Matiz (H do modelo de cores HSV) é invariante sob realces, sombras, mudanças da geometria e ângulos da iluminação.

Estatisticamente, o histograma é uma forma de aproximar a probabilidade dos valores dos três canais da cor. Para cada *bin*, o número de pontos de uma série de dados (cor dos *pixels* de uma imagem) são contados e normalizados para uma totalização de pontos, que nos dá a probabilidade de um *pixel* que cai nesse *bin*.

No espaço de cor RGB, se dividirmos cada canal R, G e B em 8 intervalos iguais com comprimento de 32: 0-31, 32-63, ... , 224-255 têm-se um histograma de 8 x 8 x 8 *bins* de cores. Um exemplo de histograma é mostrado na figura 3.5 em cada uma das três dimensionalidades do Histograma no modelo de cores RGB. A figura 3.6 mostra um outro exemplo de Histograma da imagem mostrada na figura 3.5, onde dividiu-se a informação da *Matiz* em 32 *bins* iguais.

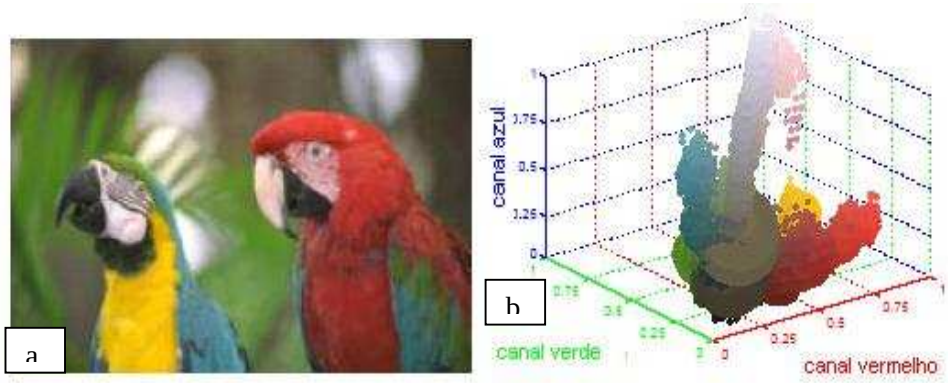

Figura 3.5: (a) imagem colorida; (b) histograma da imagem

Há dois parâmetros importantes que necessitam ser especificados ao se construir um histograma: a largura e a posição do *bin*. Não é muito difícil notar que a escolha da largura do *bin* tem um efeito enorme na aparência do histograma resultante. Escolher o *bin* muito pequeno resulta em um Histograma denteado, ou seja, um bloco separado para cada descrição. Um *bin* muito largo resulta num histograma com apenas um bloco. A posição dos *bins* são muito importantes para a forma do Histograma.

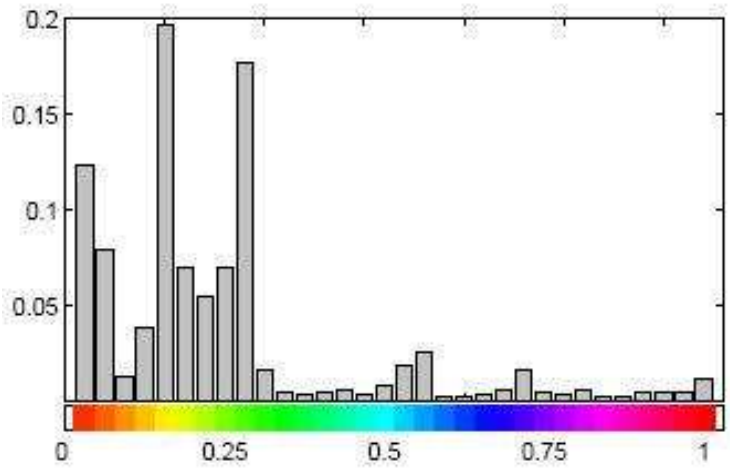

Figura 3.6: Histograma com 32 *bins*

A figura 3.7 é um resumo do trabalho de Jones e Willett(1977) mostrando claramente como cada um dos componentes são inter-relacionados para a obtenção de um resultado final.

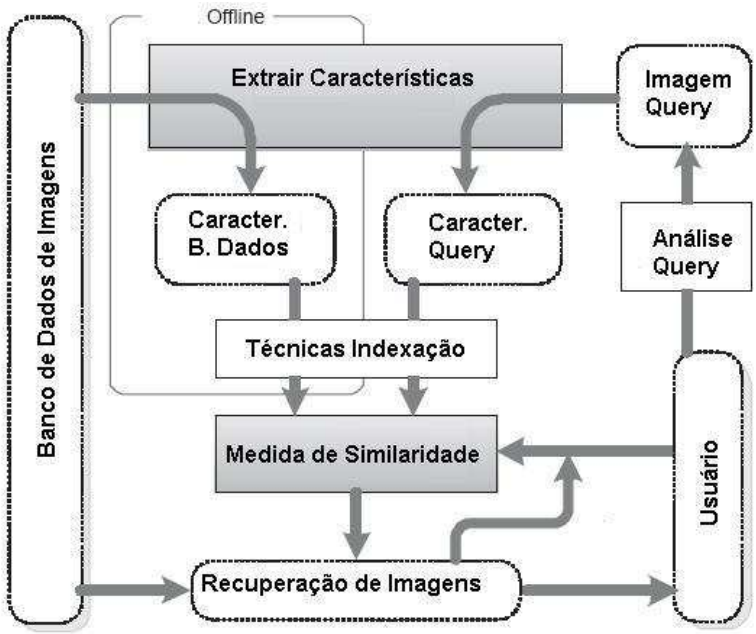

Figura 3.7: Esboço do trabalho de Jones e Willett(1977)

Para os testes foram usados mais de 1.300.000 imagens sendo, segundo os autores, o método desenvolvido é eficiente na recuperação das imagens em base de dados.

3.4 Aparência da Cor

A cor é uma das características mais utilizadas pelos seres humanos para reconhecimento e discriminação visual (BIMBO, 1999). No entanto, é normal as pessoas utilizarem diferentes sistemas nominais de cores em diferentes contextos para descrever a cor de um certo objeto, como por exemplo, associar a cor vermelha a uma maçã, embora possua um tom avermelhado, ou descrever um carro como sendo vermelho escuro, ou marrom avermelhado. A aparência de uma cor em objetos reais, geralmente é alterada pela textura da superfície, iluminação e sombras refletidas de outros objetos, e salientando também, as condições de observação e captura.

Os sistemas de bancos de dados de imagens devem enfrentar esses problemas para obter uma análise automática das cores. A informação de cor é uma componente importante do conteúdo de uma imagem, além de se apresentar robusta à distorção, rotação e translação do objeto. Quando usada em operações de recuperação de dados, permite encontrar imagens (COLOMBO; BIMBO, 1999) com cores especificadas no vetor de características e procurando por cores dos objetos armazenados na base de dados que são similares àquelas da imagem de busca.

A grande dificuldade em se trabalhar com imagens coloridas é de estabelecer o número de cores em relação ao número de características. Por exemplo, se a imagem possui 100 tonalidades de cores diferentes, será que é necessário montar um vetor de características com essas 100 tonalidades de cores?

Um aspecto interessante em se gerar um vetor de características é que o mesmo pode ser usado para a recuperação de conteúdos áudio visuais (GUIMARÃES; ARAÚJO, 2003) e muitas vezes esses conteúdos estão *on-line*. No caso de vídeo, tem-se também o mesmo problema das imagens estáticas,

ou seja, armazenamento, indexação e recuperação de informações. As informações mais utilizadas para a indexação em vídeo são a cor, a textura, a forma, o movimento e as informações de categoria do vídeo, dentre outras. O resultado tem-se mostrado viável com aplicações em jornais, bibliotecas, etc.

3.5 Considerações Finais

O processo de Recuperação de Imagens Baseada em Conteúdo envolve a extração de informações e características discriminantes da imagem. É através dessas características extraídas da imagem que se realiza a pesquisa na base de dados.

O foco principal na geração do vetor de características é estabelecer quais as informações são necessárias, quais as regras são necessárias e quais as mais relevantes para representar a imagem.

Todas as características devem ser bem moldadas, estudadas e bem elaboradas, pois se em uma *query* tiver poucas informações, certamente será problemática a recuperação de uma forma eficiente e, da mesma maneira, se a *query* tiver muitas informações além das necessárias ou mal elaboradas, tem-se também problemas para a recuperação de imagens em uma base de dados.

Os trabalhos citados na literatura mostram a dificuldade em se obter as características que são essenciais para se aprimorar e se confiar em um sistema de reconhecimento. Quando se determinam muitas variáveis, essas podem trazer seqüelas drásticas para a recuperação dos dados, um enorme tempo de processamento e as respostas nem sempre podem ser as mais adequadas. Caso as características não sejam suficientes, o resultado também não será eficiente.

Avaliar as características é fundamental nesse aspecto e, para isso, o estudo e domínio das técnicas são primordiais.

A cor tem-se mostrado de grande valia na recuperação de imagens genéricas e é uma característica discriminante para a maioria dos casos. No

entanto, o problema da alta dimensionalidade gerado pelo uso da cor tem suscitado diversas soluções que tentam otimizar o resultado.

Este trabalho vem no sentido de colaborar com esse espaço, propondo o uso de um Histograma Perceptual de cores na geração do vetor de características.