Recuperação de Imagens na Web por meio de Características de Cores

Eduardo Habib Bechelane Maia
Geraldo Augusto Massahud dos Santos Rodrigues

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
{habib,massahud}@dcc.ufmg.br

Resumo

A cor é uma característica importante da maioria das imagens disponíveis na WEB. Os algoritmos de casamento de imagens baseados em conteúdo levam em consideração esta característica. Entretanto, a eficiência e robustes deste tipo de algoritmo é um problema que ainda se encontra em aberto. O objetivo deste trabalho é apresentar alguns algoritmos para indexação de imagens que levam em consideração as cores presentes na imagem. Um algoritmo de casamento de imagens baseado no conteúdo foi implementado e comparado com outros algoritmos. Experimentos estão sendo realizados e espera-se em breve, apresentá-los.

1 Introdução

Nos últimos anos, muita pesquisa tem sido feita para a definição de ferramentas eficientes para fazer busca de imagens em uma base de dados. O uso da cor tem sido muito experimentado em:

1. Casamento de cores, para encontrar imagens com uma mesma proporção de cores;
2. Pesquisa por similaridades, para encontrar imagens que se assemelham a uma imagem apresentada pelo usuário;
3. Pesquisa por região, para encontrar imagens contendo regiões com a mesma coloração da imagem especificada no busca;
4. Pesquisa por objeto específico, para encontrar imagens que possuem um determinado objeto apresentado pelo usuário (um cachorro, por exemplo);
5. Categorização de imagens, para armazenar as imagens de uma dada base de dados em grupos(Ex.: desenhos ou fotografia)
6. Encontrar imagens com determinado efeito.

Todas estas pesquisas dependem da definição de características robustas e eficientes que podem representar a imagem. Entretanto, não existe uma melhor representação de características de cores e sim representações que se adequam melhor a uma determinada situação. Entretanto, qualquer que seja o contexto, as seguintes propriedades básicas devem ser representadas:

- Porcentual de similaridade: As características de duas imagens somente são muito diferentes se as imagens não são iguais;
- Eficiência: Elas podem ser rapidamente computadas;
- Economia: Sua dimensão é pequena para não afetar a eficiência do retorno.
- Escalabilidade: A performance do sistema não pode ser influenciada pelo tamanho da base de dados.
- Robustês: Mudanças nas condições da imagem não podem afetar o retorno.

Ao montar a base de dados onde serão armazenadas as imagens, corre uma indexação destas imagens. Ou seja, a cada imagem é atribuída um conjunto de características ou índices que serão usados pelo sistema no casamento do padrão.
Quando é feita a busca por uma imagem, as suas características são extraídas e ela passa pelo mesmo processo que as imagens armazenadas no banco de dados passaram. Existe chance de a imagem procurada ser uma imagem do banco de dados se as características de ambas são iguais ou pelo menos similares. As imagens são retornadas em uma lista e as que o sistema julgou serem mais similares à imagem procurada são retornadas no início da lista.
Exceto algumas poucas exceções, a busca levando-se em consideração características de cores só é eficiente caso as cores utilizadas para representar uma imagem sejam reduzidas. Para isso, normalmente um algoritmo faz um pré-processamento da imagem, procurando quais as n cores mais significativas da imagem e é com estas cores que a imagem passa a ser representada.
Este texto será organizado da seguinte maneira: A seção 2 fala sobre algoritmos baseados unicamente na medição das cores da imagem, a seção 3 leva em consideração as características espaciais da imagem, a seção 4 descreve um método de casamento de imagens baseado em *wavelets* e a seção 5 demonstra um algoritmo apresentado em [1].

2 Índices de cores

Após uma redução significativa das cores da imagem, ela é codificada de forma a ter seu conteúdo representado. Existem muitas maneiras diferentes de se codificar uma imagem. Normalmente, as cores são representadas em histogramas: Neste caso, é a estratégia de casamento que distingue um método de outro. Um histograma de cores, H, é um vetor em que cada posição H_i contém o número de pixels que possuem a cor i na imagem.

Alguns algoritmos de casamento de imagens simplesmente verificam se o número de pixels encontrado para as cores da imagem procurada se assemelham com o das imagens da base de dados.

Outros algoritmos de casamento de imagens simplesmente fazem um "shift" nas cores da imagem(Figura 1) procurada para verificar se conseguem realizar a busca com sucesso. As vezes, um simples shift nas cores da imagem, faz com que a imagem procurada seja encontrada. Isto pode ocorrer principalmente devido a condições diferentes de luminosidade na hora da geração da imagem.

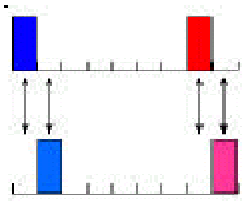

Figura 1 - Shift nas cores das imagens

Em [2] é proposto um algoritmo que leva em consideração uma diferença de cores. Se uma cor estiverem até a uma distância x da cor original esperada, o algoritmo considera que esta é a mesma cor. O problema é que este algoritmo pode ser muito ineficiente e, além disso, ele possui um grande custo computacional.

3 Índice de cores espaciais

O maior problema dos métodos descritos anteriormente é que eles não levam em consideração a informação espacial da imagem. Por exemplo, todas os quadrados demonstrados na Figura 2 possuem a mesma proporção, mas diferentes distribuições espaciais. A aparência é

claramente diferente, o que implica que, obviamente, não podemos esperar que somente as características de cores levem a uma boa representação de uma imagem.

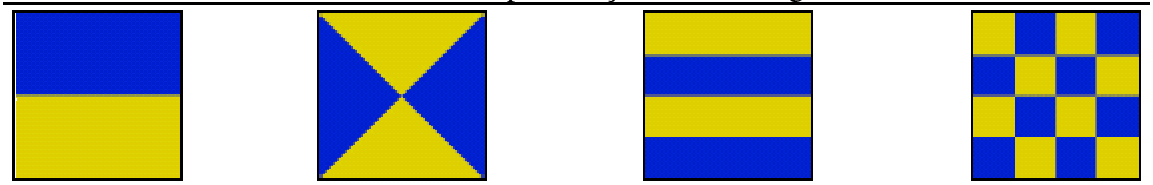

Figura 2 - Alguns padrões com a mesma proporção de cor

Este problema nos leva a crer que os métodos anteriormente citados podem não retornar bons resultados em uma base de dados grande. Existem algumas maneiras de se armazenar a informação espacial de uma imagem e estas maneiras serão demonstradas abaixo.

A maneira mais simples de se prover esta informação é dividir a imagem em sub-imagens e então indexar separadamente cada uma delas.

Em [3], por exemplo, apresentado um modelo em que a imagem é dividida em 9 partes iguais e cada sub-área da imagem é representada por um histograma diferente (Figura 3). Entretanto, embora simples, este método, além de não prover uma boa informação espacial de cor, possui um alto custo computacional e de armazenamento.

Também em [3], é demonstrado um outro modelo que divide a imagem em uma região oval no meio e outras 4 regiões nos cantos (Figura 3). Após esta divisão, o sistema analisa as características de cores de todas as regiões e combina as características dessas sub-imagens, atribuindo um peso maior às características da região central.

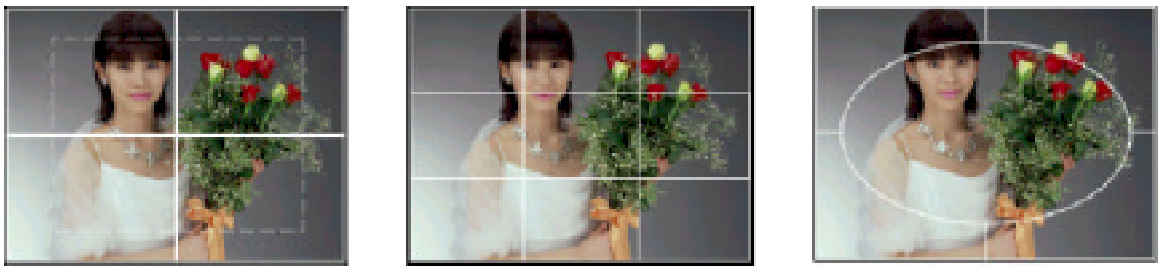

Figura 3 - Exemplos de possíveis sub-indexação de imagens

4 Análise Multi-Resolução e Wavelets

A análise multi-resolução é uma estratégia de processamento de sinais onde é utilizado um conjunto de filtros especializados em extrair as informações do sinal, como as freqüências presentes nesse e sua localização no tempo de duração do sinal, em diferentes resoluções [4]. O uso deste ferramental se faz útil em sistemas de classificação de imagem, segundo diversos estudos sobre o sistema visual humano [5]. Estes estudos indicaram que o córtex visual humano detecta as imagens em diferentes resoluções, onde as informações sobre as freqüências e a orientação dos tons presentes na imagem são interpretadas separadamente. A análise da imagem é feita através de um mapeamento desta para um conjunto de imagens em diferentes resoluções, a partir de uma transformada matemática.

As transformadas matemáticas são utilizadas em um conjunto de dados para se obter informações adicionais, não disponíveis no modelo de dados primitivos. Neste caso, é necessário o uso de uma transformada que detecte a variação dos tons de cores da vizinhança de um pixel e a localização espacial deste e transponha isto de forma eficiente em um espaço multi-resolução [4].

Métodos tradicionais de análise de sinal, baseados na transformada de Fourier, como o TDC, transformada discreta de cossenos, detectam a variação de tons presentes na imagem, mas não

geram a informação sobre a localização espacial destas, uma vez que estas técnicas consideram que o sinal é periódico, ou seja, uma imagem que não possui informações de bordas ou texturas.

Uma variação da transformada de Fourier, a transformada de Fourier em Janela, foi elaborada para efetuar uma localização espacial das freqüências dos tons presentes na imagem, mas a informação sobre estas freqüências não é precisa. Dessa forma, não se tem o controle da resolução dos objetos da imagem, apenas sua localização [4].

As transformadas de wavelets suprem estas falhas presentes na abordagem da transformada de Fourier para a análise de sinais não periódicos. As wavelets foram desenvolvidas com base na matemática, física quântica e estatística, e são funções que decompõem os sinais em diferentes componentes de freqüência e analisa cada componente em diferentes resoluções. Outras aplicações comuns das wavelets são a remoção de ruído de sinal, a compressão de imagens, suavização de imagens e análise fractal [4].

Diferente da transformada de Fourier, a transformada de wavelets não possui apenas uma função base e sim uma classe delas, todas obedecendo aos requisitos de localização tempo/freqüência do sinal, onde cada função possui características próprias, que implicam numa melhor ou pior distinção das freqüências analisadas e da localização destas.

Dependendo da aplicação desejada e da natureza das imagens, a escolha da função base é um diferencial [4].

4.1.1 Algoritmo que utiliza Wavelets

O algoritmo da transformada de wavelets é feito da seguinte forma: o sinal é passado por dois filtros, um passa-baixas e um passa-altas, os quais irão dividir o sinal em dois sinais menores, um de freqüências maiores que um limiar e outro de freqüências menores. Este processo é repetido sobre o sinal de menores freqüências até um nível predeterminado e os sinais de maior freqüência são armazenados a cada iteração com os filtros. Estes sinais de maior freqüência são conhecidos como coeficientes de wavelet.

A aplicação destes filtros em uma imagem é feita pela iteração entre linhas e colunas da matriz de imagem, o que gera a cada nível de iteração 3 sub-imagens, que são os coeficientes de wavelets armazenados, em diferentes orientações, cada uma responsável por um ângulo na variação dos tons. Dessa forma, o resultado da transformada é composto de sub-imagens organizadas em diferentes níveis de resolução, cada um responsável por espectro de freqüências, e em cada nível, têm–se a angulação destas texturas, obtendo-se assim, uma descrição das variações dos tons de cinza e a localização destes.

4.1.2 Vetor de características

Para a geração do vetor de características, é efetuada uma análise estatística em cada uma das sub-imagens, onde cada elemento do vetor corresponde a uma descrição de cada sub-imagem, de maneira que o conteúdo presente na imagem seja corretamente descrito por este vetor. Para efetuar essa análise estatística, foram testados desvio padrão da distribuição dos coeficientes, a média e a energia.

Depois de ocorrida esta extração de características, é possível, a partir de alguma métrica, como distância euclidiana ou de malahobilis, identificar imagens ou regiões de imagens como semelhantes, dado um certo grau de aproximação de seus vetores. Esta técnica foi utilizada na identificação de grãos de pólen, para que seja efetuada uma procura por seus padrões. Assim, após serem definidas quais são as regiões mais significativas da imagem é possível distingui-las. Estes vetores de características foram amostrados para uma rede neural que determinou quais deles poderiam identificar melhor a imagem. Abaixo alguns exemplos de identificação de padrões de regiões nas imagens de pólen.

5 MCAG + SVG

Em [1] é apresentado um algoritmo de reconhecimento de imagens baseado nas características de cores. Ele pode ser dividido em 3 etapas:

1. **aplicação do filtro majoritário** : Inicialmente, ele faz um pré-processamento da imagem removendo 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. No pré-processamento adotado este ruído é retirado preenchendo os pixels onde ele ocorre com a cor predominante da janela(Figura 4).

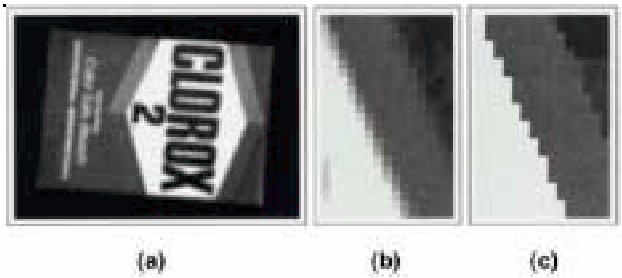

Figura 4 - Aplicação do filtro majoritário. (a) Imagem Real; (b)Região superior direita; (c) resultado da aplicação do filtro em (b)

2. **obtenção do MCAG**, que descreve as regiões cromáticas da imagem e o tamanho das adjacências entre as mesmas. Cada nodo do grafo representa uma cor no modelo RGB da imagem, contendo a quantidade de pixel da sua cor e cada aresta entre os nodos *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, cada pixel contado, na região vizinha (vizinhança 8), é adicionado ao nodo correspondente (Figura 3);

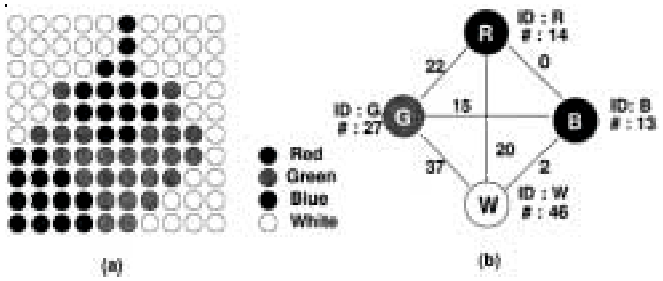

Figura 5 - Exemplo simples da geração do MCAG. (a) Imagem 10 X 10; (b) Resultado do MCAG

3. **obtenção do SVG,** que descreve a distribuição de cor no domínio espacial da imagem. As informações necessárias para elaborar este grafo são: as cores da imagem, a contagem de cor da imagem e coordenadas de cada pixel, para cada cor. Cada nodo contém a própria variância espacial da cor e a variância espacial relativa de duas cores. Cada nodo contém a própria variância das cores denotadas e cada borda tem como seus atributos a variância espacial entre as cores denotadas pelos nodos ligados.

6 Avaliação de algoritmos

Na avaliação de algoritmos resolvemos utilizar um sistema de pesquisa de imagens por conteúdo[6] desenvolvido no NPDI. Este sistema já possui uma implementação da técnica *MCAG*

+ *SVG*, resolvemos então implementar uma técnica de indexação de cores, mais simples, para verificar o quanto é melhor um algoritmo mais atual em relação aos primeiros propostos.

O algoritmo que implementamos faz a indexação dos canais RGB, com 8 níveis cada um (os 3 bits mais significativos das cores). Para gerar os histogramas, percorremos todos os pixels da imagem, e considerando os 3 bits significativos de cada canal, somamos um ao índice de cada canal e depois dividimos todos valores pelo número de pixels, para que a soma de todos os índices de um determinado canal seja 1.

Após gerar os histogramas dos canais RGB, a comparação para verificação da similaridade é feita através da soma dos quadrados das diferenças da quantidade de pixels. Assim quanto menor for esta soma, maior é a similaridade entre as imagens.

Infelizmente até a finalização deste texto não terminamos a execução dos testes para avaliação dos dois algoritmos. Esperamos que na data da apresentação estas avaliações estejam terminadas e as apresentaremos. O que podemos dizer de antemão é que esperamos que o método *MCAG + SVG* retorne resultados mais corretos que o método de indexação.

7 Conclusão:

Nos últimos anos, houve um grande crescimento de coleções de imagens digitais. Com isso, a existência de um sistema eficiente que permita uma navegação, pesquisa e recuperação destas imagens se tornou necessária.

Concluímos que a recuperação de imagens baseadas em descrições textuais era maneira mais utilizada no início das pesquisas nesta área, mas apresentava dois grandes problemas. O primeiro é que a anotação manual da informação das imagens ocasionaria em um grande trabalho. A segunda dificuldade, resulta da subjetividade humana de percepção e do rico conteúdo das imagens. Como pessoas diferentes podem notar detalhes diferentes em imagens idênticas, existe a probabilidade de se fazer uma descrição incompleta e/ou incorreta o que causaria perdas irrecuperáveis nas bases de dados. Assim, vários estudos foram feitos no intuito de se desenvolver algoritmos que conseguissem fazer uma busca levando-se em conta características da própria imagem.

Na seção 2 apresentamos algoritmos que se baseiam unicamente nas coloração das imagens. Nos algoritmos apresentados, há a montagem de um histograma e no fim, ele é analisado e a imagem é comparada com a imagem fornecida pelo usuário. Entretanto, vimos que estes algoritmos não levam em consideração características espaciais da figura.

Na seção 3 vimos alguns algoritmos que levam em consideração estas características espaciais simplesmente com a divisão da imagem em várias e a análise separada de cada parte da imagem.

Na seção 4 foram apresentados algoritmos que levam em consideração transformadas matemáticas (de Fourier e wavelets).

Na seção 5 foi apresentado um dos algoritmos mais atuais de pesquisa de imagens, que utiliza representação das características da imagem em grafos híbridos (MCAG e SVG), representando características de cor e espacial simultaneamente.

Na seção 6 apresentamos um algoritmo simples implementado por nós, com a intenção de verificar as melhoras existentes entre os primeiros algoritmos de pesquisa de imagens por conteúdo (indexação de cores) e um dos algoritmos mais atuais (MCAG + SVG).

Esperamos ter os resultados da avaliação até o dia da apresentação oral do seminário.

8 Referências Bibliográficas

[1] Park, I. K., Yun, I. D., Lee, S. U., "*Color image retrieval using hibryd graph representation*", Image and Vision Computing, 17(7), 465-474, (1999).

[2] Hafner J., Sawhney H.S., Esquitz W., FlicknerM. Niblack W., *"Efficient color histogram indexing for quadratic form distance function"*. IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 17, pp. 729-736, 1995.

[3] Schettini, R., Ciocca, G., Zuffi, S; *"A survey of methods for colour image indexing and retrieval in image Databases"*

[4] Castleman., K. R. "*Digital Imaging Processing"*, 1 ed, Prentice Hal l(1996).

[5] Bradley, A. P. *"A wavelet visible difference predictor"*, IEEE Transactions On Image Processing 8 (1999), no 5, 717 –730

[6] Gomes, D. R., Araújo, A. A. "*Implementação de um Sistema de Recuperação de Imagens para a Web.*", Relatório Final de POC 2, Março de 2003.