

Índice Invertido para Recuperação de Imagens Baseada em Conteúdo

Tauller A. Matos, Ilmério Silva, Celia A. Barcelos, Patrícia A. Proença

Faculdade de Ciência da Computação, FACOM, UFU

38400-902, Uberlândia, MG

E-mail: tauller@yahoo.com.br, ilmerio@facom.ufu.br, celiazb@ufu.br, patriciaproenca@gmail.com

Resumo: *Apresentamos uma proposta inovadora de utilização do Índice Invertido para recuperação de imagens baseado em conteúdo (CBIR). O objetivo da proposta é acelerar o processamento de consultas, sem perda de qualidade na resposta. Para isto, implementamos dois sistemas. O primeiro, baseia-se nas técnicas para extração das características e para cálculo de similaridade normalmente utilizadas em CBIR, a saber, momentos de cor e distância euclidiana. O segundo, emprega o mesmo extrator de características do primeiro, mas indexa as imagens utilizando uma estrutura originada em recuperação textual conhecida como índice invertido. O índice invertido permite reduzir significativamente o número de imagens que devem ser comparadas com a imagem consulta. Avaliamos a qualidade dos resultados por meio de experimentos, usando a medida Precisão-R. Os resultados mostram que nossa abordagem não tem impactos negativos na qualidade do resultado.*

1 Introdução

O preço acessível dos computadores, máquinas fotográficas e celulares digitais, juntamente com os dispositivos de alta capacidade de armazenamento, as conexões rápidas e os serviços que proporcionam o compartilhamento de informações, são alguns dos fatores que têm motivado uma alta demanda de soluções para a organização e compartilhamento de informações digitais. A informação, disponível, atualmente se caracteriza por um grande volume de dados e por uma diversidade de formatos de armazenamento.

Neste cenário, encontrar informações relevantes, se torna cada vez mais uma tarefa

difícil. Para lidar com o problema de auxiliar o usuário a encontrar informação relevante, denominado "Problema de Recuperação de Informação (RI)", vários modelos têm sido propostos na literatura. Alguns modelos são específicos para bibliotecas digitais com documentos semi-estruturados, outros são genéricos, considerando dados não estruturados, por exemplo, texto e imagens.

Uma abordagem comumente utilizada em grandes coleções como a Web é a recuperação com base na informação textual que descreve a imagem. Esta é a abordagem usada no site google (<http://images.google.com.br>) [18]. Neste caso, a consulta é um conjunto de palavras-chaves (chamada termo) que é usada para recuperar imagens em uma coleção. Entretanto, a informação textual sobre as imagens nem sempre está disponível ou, quando disponível, não consegue descrever todas as suas características.

Outra abordagem utilizada para recuperação de imagens é conhecida como Recuperação de Imagens Baseada em Conteúdo (CBIR). Seu objetivo principal é encontrar imagens relevantes conforme a necessidade do usuário, através de características visuais automaticamente extraídas das imagens. Atualmente, os métodos de representação de características das imagens mais utilizados usam: cor, textura e forma como atributos de indexação, os quais são extraídos da imagem de maneira independente [10, 11]. Uma das diferenças dessa abordagem com a recuperação textual é conter uma imagem exemplo como consulta e não palavras-chaves. Uma vez que as características das imagens tenham sido extraídas e armazenadas em vetores de característica, faz-se necessário medidas que comparam a similaridade entre ve-

tores de característica das imagens do banco de dados com o vetor de característica da imagem exemplo. Medidas de similaridade são normalmente baseadas em medidas de distância e os vetores das imagens devem ter a mesma dimensão.

O Sistema MARS [17] utiliza a distância euclidiana para computar a similaridade de características de textura entre imagens. O Sistema Netra [13], também utiliza a distância euclidiana para similaridade baseada em cor e forma. Além disso, usa *distância city-block* para calcular a similaridade baseada em textura. Em CIRES [14] e *Blobworld* [15] a comparação de características de forma e textura é obtida por meio da distância euclidiana. Em todos estes sistemas, o processo de recuperação inclui o cálculo da distância entre o vetor da imagem de consulta e todos os vetores das imagens do banco de dados, tornando a recuperação computacionalmente cara. Este processo é inviável para grandes coleções de imagens, por exemplo na Web.

Diante disso, propomos uma nova abordagem para o cálculo da similaridade em CBIR. Construímos um sistema com uma estrutura utilizada em recuperação textual chamada Índice Invertido, descrita na Seção 3. O uso desta estrutura em recuperação textual tem resultado em ganho na velocidade de recuperação de documentos sem perda de qualidade [12].

O trabalho consiste na implementação e avaliação experimental de dois sistemas de recuperação. O primeiro é um sistema CBIR que utiliza o extrator de características conhecido como Momentos de Cor [7] e calcula a similaridade baseada na distância euclidiana. Chamaremos este sistema de CBIR-Cor. Já o segundo sistema, se difere do primeiro no processo de indexação das imagens e no cálculo da similaridade. O extrator de característica é o mesmo que do anterior, com a diferença na indexação, pois utilizamos o índice invertido. Este sistema é aqui chamado de CBIR-indexado. Os experimentos foram realizados com a base de dados Corel-1000 - um subconjunto do banco de dados Corel - contendo 1000 imagens.

Este trabalho, tem como objetivo principal, comparar o resultado qualitativo destes dois sistemas. Além de mostrar, através de resultados experimentais, que o ganho de tempo no processamento da consulta usando o índice in-

vertido não provoca perda de qualidade na recuperação. Este artigo está dividido da seguinte forma: na Seção 2 definimos CBIR e apresentamos o CBIR-Cor, na Seção 3 descrevemos o índice invertido e o CBIR-Indexado, na Seção 4 discutimos os resultados experimentais, e por fim, na Seção 5 apresentamos as conclusões e trabalhos futuros.

2 CBIR-Cor

A dificuldade e o custo de produzir anotações textuais ricas e confiáveis para grandes bancos de dados de imagens, como também a subjetividade associada a estas anotações, explicam o fato do CBIR, ser de grande interesse nos dias atuais [4].

As etapas de um sistema de CBIR são resumidas a seguir. Em primeiro lugar, à partir de um banco de dados de imagens digitais, extraí-se características sintáticas e/ou semânticas, obtendo um banco de dados de vetores de características, onde cada vetor representa uma imagem. O mesmo processo de extração das características feito para o banco de dados é aplicado à consulta do usuário, obtendo-se o vetor de características da imagem consulta. Então, compara-se o vetor de características da imagem consulta com os vetores de características das imagens da coleção através de alguma medida de similaridade. Com base nos valores de similaridade um *ranking* é formado e mostrado ao usuário.

O primeiro sistema implementado é aqui chamado de CBIR-Cor. É um sistema de recuperação de imagens por conteúdo utilizando o extrator de característica Momentos de Cor. O cálculo da similaridade é realizado pela formulação da distância euclidiana.

2.1 Modelos de Cor

Imagens digitais em geral são representações de uma imagem em duas dimensões, armazenadas como uma matriz de valores digitais, chamados *pixels*. Um dos modelos de cores mais comuns na recuperação de imagens digitais coloridas é atribuir a cada pixel um valor associado à percentagem de três cores: vermelho, verde e azul. As três componentes primárias formam um espaço tridimensional de cores e cada pixel da imagem é representado por um

ponto neste espaço. Este modelo de representação é conhecido como RGB. Segundo [3], outro modelo de cor comum na recuperação de imagens é o HSV (*Hue, Saturation, Value*) ou (matiz, saturação, intensidade). Uma representação tridimensional do espaço HSV é um cone, onde o eixo vertical central representa a variação da intensidade. A matiz é definida como sendo o ângulo, variando de $[0, 2\pi]$, relativo ao eixo do Vermelho, sendo o ângulo vermelho igual a 0, o verde a $2\pi/3$, o azul $4\pi/3$ e o vermelho novamente em 2π . A saturação indica a pureza da cor, e é medida pela distância radial em relação ao eixo central, tendo como valor 0 no centro e valor 1 na face externa do cone. Um exemplo gráfico do modelo de cor HSV pode ser visualizado na Figura 1 [6].

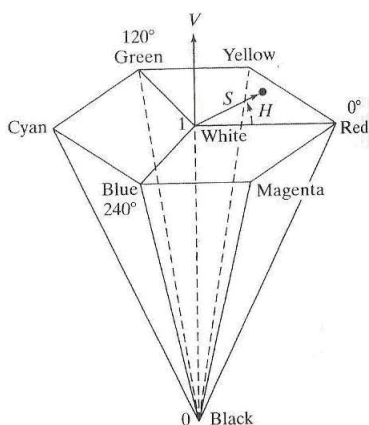


Figura 1: Espaço de cor HSV

Conversão de RGB para HSV Podemos transformar qualquer imagem no espaço RGB para uma imagem no espaço HSV aplicando as seguintes regras: As cores nos modelos HSV são obtidas a partir das informações RGB com respeito aos valores normalizados do vermelho, verde e azul, dados por:

$$r = \frac{red}{255} \quad (1)$$

$$g = \frac{green}{255} \quad (2)$$

$$b = \frac{blue}{255} \quad (3)$$

Após obter os canais red(r), green (g) e blue (b) normalizados, as equações abaixo são utilizadas para o cálculo da matiz (H), saturação (S), intensidade (V).

$$H = \arctan \left(\frac{\sqrt{(r-g) + (r-b)}}{\sqrt{(r-g)^2 + (r-g)(r-b)}} \right) \quad (4)$$

$$S = 1 - \frac{3}{r+g+b} \times \min(r, g, b) \quad (5)$$

$$V = \frac{r+g+b}{3} \quad (6)$$

A vantagem dessa representação de cor está na possibilidade de separar os canais. Esse fato torna o modelo HSV uma ferramenta ideal para o desenvolvimento de algoritmos de processamento de imagens. Os experimentos foram realizados usando o modelo HSV. Além disso, a distribuição de cores nos diversos pixels da imagem foi computada usando Momentos de Cor, conforme descrição a seguir.

2.2 Momentos de Cor

Caracterizam as imagens em termos da distribuição das cores nos pixels da imagem. São normalmente dados por três medidas estatísticas: média, desvio padrão e obliquidade. Os momentos de cor podem ser derivados do espaço de cor HSV da seguinte forma.

Sendo N o número de pixels de uma imagem e p_{ij} o valor do j -ésimo pixel no i -ésimo canal de cor, os três primeiros momentos: média (E_i), desvio padrão (σ_i) e obliquidade (s_i) são computados através das expressões mostradas a seguir:

$$E_i = \frac{1}{N} \sum_{j=1}^1 p_{ij} \quad (7)$$

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^1 (p_{ij} - E_i)^2 \right)} \quad (8)$$

$$s_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^1 (p_{ij} - E_i)^3 \right)} \quad (9)$$

Em resumo, convertemos as imagens do espaço de cor RGB para o espaço de cor HSV. Então, calculamos os três momentos acima para cada canal H, S e V. Com isto, o vetor de característica da imagem terá nove posições distribuídas conforme Tabela 1.

Vetor	Descrição
1	média da matiz H
2	desvio padrão da matiz H
3	obliquidade da matiz H
4	média da saturação S
5	desvio padrão da saturação S
6	obliquidade da saturação S
7	média da intensidade V
8	desvio padrão da intensidade V
9	obliquidade da intensidade V

Tabela 1: Descrição do vetor de característica resultante da extração de característica da imagem usando três momentos do espaço de cor HSV

3 CBIR-Indexado

O segundo sistema implementado, denominado CBIR-Indexado, utiliza o mesmo extrator de características Momentos de Cor descritos na seção 2.2. A diferença deste sistema está no processo de indexação das imagens e no cálculo da similaridade. Esta seção tem como objetivo apresentar o método utilizado para implementação deste segundo sistema. Inicialmente descrevemos o índice invertido da recuperação textual.

3.1 Índice Invertido

Um banco de dados textual é uma coleção de documentos, que pode ser representado por um conjunto de registros. Cada registro contém uma lista de palavras. Uma estratégia de recuperação textual é comparar a lista de palavras de uma consulta com a lista de palavras dos registros. Entretanto, em grandes coleções a comparação direta da consulta com todos os registros é computacionalmente cara. Com o objetivo de acelerar este cálculo, limitando a comparação a um subconjunto de registros, utiliza-se o o índice invertido [16]. A idéia é limitar a comparação somente com o subconjunto de registros que contenham pelo menos um termo da consulta.

O índice invertido possui duas partes principais: uma estrutura de busca, chamada de vocabulário, contendo todos os termos distintos existentes nos textos indexados e, para cada termo, uma lista invertida que armazena os identificadores dos registros contendo o termo,

a saber, documentos onde a palavra ocorre [12], como ilustrado na figura 2.

Consultas nos arquivos invertidos são feitas tomando-se a lista invertida correspondente ao termo procurado. O uso de arquivos invertidos na recuperação textual mostra-se com um grande ganho na velocidade de recuperação de documentos.

Uma busca por documentos com os termos "matemática" e "aplicada" pode ser processada por meio das duas listas invertidas, sem a necessidade de percorrer todos os registros.

3.2 Análise dos vetores de características

A indexação das imagens utilizando o índice invertido necessitou de uma classificação de grupos de valores de características de baixo nível de imagens digitais gerados pelo CBIR. Esse valores foram mapeados para identificadores de indexação da coleção (faixas).

À distribuição dos valores em cada uma das nove posições do vetor (ver Tabela 1, quando observada em um histograma obtido da coleção de teste, tem comportamento semelhante à distribuição normal de probabilidades. Com o intuito de agrupar as imagens semelhantes dentro de uma mesma lista invertida, aplicamos ocorrências dentro de um intervalo de valores [9], separando cada posição do vetor em 8 classes, usando a tabela da distribuição normal padronizada [8], conforme a figura 3.

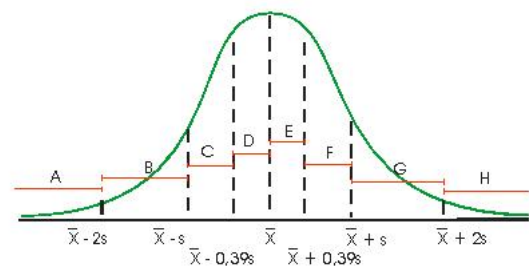


Figura 3: Distribuição adotada para classificação das faixas. Onde \bar{X} é a média e s o desvio padrão.

4 Resultados Experimentais

A base de dados utilizada para extração dos vetores de características das imagens de ambos os sistemas foi a Corel-1000. Esta base é

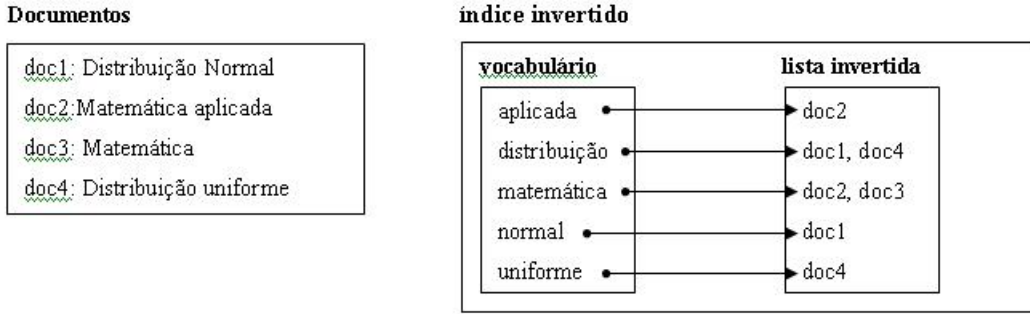


Figura 2: Exemplo de arquivo invertido

um subconjunto do banco de dados Corel, contendo 1000 imagens com resolução 384×256 ou 256×384 pixels. O banco de dados é composto por 10 categorias (África, praia, edifícios, ônibus, dinossauros, elefantes, flores, comidas, cavalos e montanhas), com 100 imagens em cada categoria.

Para avaliar a eficácia de recuperação utilizamos a medida conhecida como Precisão-R. A Precisão-R é dada pela precisão entre as R primeiras imagens do *ranking*, onde R é o número total de imagens relevantes para a consulta em análise. Então, a Precisão-R de uma consulta q é dada por:

$$PR(q) = \frac{Relevantes(R, q)}{R} \quad (10)$$

onde $Relevantes(R, q)$ é a quantidade de imagens relevantes encontradas até a posição R do ranking. Em nossos experimentos R é igual a 100. Pois, dada uma imagem consulta, pertencente ao banco de imagens, o conjunto de imagens relevantes é formado pelas imagens da mesma classe da imagem consulta.

Para verificarmos a perda de qualidade na recuperação, consideramos como imagem consulta algumas imagens escolhidas de forma aleatória. Para cada classe do Corel-1000, escolhemos quatro imagens-consulta, calculamos a Precisão-R de cada sistema e o valor médio, conforme a Tabela 2. No final, calculamos a média final da Precisão-R para a base de dados do Corel.

Como podemos observar, obtivemos médias muito próximas entre os dois sistemas. A diferença na média geral foi de 3,5%.

Classe	CBIR-Indexado	CBIR-Cor
África	51,6%	58%
Praia	26,3%	22,3%
Edifícios	22,3%	18,6%
Dinossauros	55%	75%
Ônibus	44,3%	51,3%
Cavalos	44%	45%
Flores	37%	31,6%
Comidas	37,6%	37,3%
Elefantes	38,6%	42,3%
Montanhas	20%	31%
Média Geral	37,7%	41,2%

Tabela 2: Comparação da Precisão-R nos sistemas CBIR-Indexado e CBIR-Cor

5 Conclusão e Trabalhos Futuros

Visando uma melhoria no desempenho do processamento das consultas em CBIR, este artigo propôs uma nova forma de indexar as imagens de uma coleção. Foram implementados dois sistemas aqui chamados de CBIR-Cor e CBIR-Indexado. O primeiro foi construído com as ideias originais de CBIR. A similaridade foi calculada através da distância euclidiana. Medida esta, que tem um baixo desempenho de processamento em grandes coleções. O segundo programa, foi implementado utilizando a estrutura conhecida como índice invertido. Estrutura esta, que permite um ganho de desempenho no processamento da consulta. Este ganho é obtido porque nem todas as imagens do banco serão comparadas com a consulta, mas somente aquelas que tiverem características em uma das faixas citadas na seção 3.2 da imagem de consulta. Nossos resultados mostram que não houve perda significativa na qualidade da recuperação, o que sugere, que nossa forma de indexar as imagens é satisfatória. Como trabalho

futuro, pretende-se aumentar estes testes para grandes coleções de imagens e avaliarmos os resultados qualitativos, além de expandirmos o vetor de características, utilizando outros extratores, tais como, textura e forma.

Referências

- [1] Jones, K. S. and Willett, P. Readings in Information Retrieval. Morgan Kaufmann Publishers, San Francisco, California, 1997.
- [2] Manning, C., Raghavan, P., Schütze, H., An Introduction to information Retrieval. Cambridge University Press, Cambridge, England, 2007, 365pgs.
- [3] Schettini, R., Ciocca, G., and ZUT, S. A survey on methods for colour image indexing and retrieval in image databases. In R. Luo and L. MacDonald (Eds.), Color Imaging Science: Exploiting Digital Media, Wiley, New York 2001.
- [4] Crucianu, M., Ferecatu, M., and Boujema, N. Relevance Feedback for Image Retrieval: a Short Survey. Le Chesnay Cedex. France, Italy, 2004.
- [5] Gonzalez, R. e Woods, R. E. Processamento de Imagens Digitais. Ed. Edgard Blücher Ltda, São Paulo, 2000.
- [6] Gonzalez, R. C and Woods, R. E. and EDDINS, S. L. Digital Image Processing Using MATLAB. Ed. Pearson Education, 2004.
- [7] Stricker, M. and Orengo, M. Similarity of Color Images. In Proceedings of IS&T and SPIE Storage and Retrieval of Image and Video Databases III, San Jose, USA, 1995.
- [8] Levine, D. M. and Berenson, M. L., Stephan, D. Estatística: Teoria e Aplicações: Usando Microsoft Excel. 1998 tr. It. de Teresa Cristina Padilha de Souza, LTC, Rio de Janeiro, 2000.
- [9] Spiegel, M. R. Scaum's Outline of Theory and Problems of Statistics. Schaum Publishing Co. E.U.A 1970. tr. It. de Pedro Cosentino, ed. McGraw-Hill do Brasil, Rio de Janeiro, 1971.
- [10] Smeulders Awm, Worring M, Gupta A, and Jain R. Content based image retrieval at the end of the early years. IEEE Transactions on Image Processing Intell 22, pages 1349-1380, 2000.
- [11] Ryszard S. Choras, Tomasz Andrysiak, and Michal Choras. Integrated color, texture and shape information for content based image retrieval. Pattern Anal Applic, 2007.
- [12] R. Baeza-Yates and B. Ribeiro-Neto. Modern Information Retrieval. Addison-Wesley Essex, UK, 1999.
- [13] Ma, W. Y. and Manjunath, B. S. Netra: A toolbox for navigating large image databases. Multimedia Systems, 7(3):184-198, 1999.
- [14] Iqbal, Q. and Aggarwal, J. K. CIRES: A System for Content-Based Retrieval in Digital Image Libraries. In Seventh International Conference on Control, Automation, Robotics and Vision, pages 205-210, 2002.
- [15] Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M., and Malik, J. Blobworld: A system for region-based image indexing and retrieval. In Proceedings of the Third International Conference VISUAL, Lecture Notes in Computer Science, Springer, Amsterdam, Netherlands, 1999.
- [16] Zobel J. Inverted files versus signatures files for text indexing. ACM Transaction on Database Systems, Vol. 23, No. 4, December, 1998.
- [17] Rui, Y., T.S.Huang, and Mehrotra, S. . Content-based image retrieval with relevance feedback in MARS. In Proceedings of International Conference on Image Processing, volume 2, pages 815-818, 1997.
- [18] Silva, S. F. Realimentação de Relevância via Algoritmos Genéticos aplicada na Recuperação de Imagens, Faculdade de Computação, Universidade Federal de Uberlândia, UFU, Dissertação de mestrado, 2007.