

Workshop de Computación Gráfica, Imágenes y Visualización (WCGIV)

Avaliação de Técnicas para Recuperação de Imagens baseada nos aspectos Cromático e Estrutural por curvas “Recall x Precision”

Yandre M. G. da Costa

Diego B. Gonçalves

Gabriel C. Silva

Hereck F. Santin

Universidade Paranaense – Campus Paranavaí

Paranavaí – PR, Brasil, CEP 87706-490

yandre@unipar.br

Resumo

Este trabalho descreve a avaliação, através de curvas “recall x precision”, de um conjunto de técnicas para recuperação de imagens baseadas nos conteúdos cromático e estrutural das mesmas. Durante a aplicação destas técnicas foram utilizadas diferentes combinações de valores que permitiram ajustar importantes fatores, tais como o peso empregado para os aspectos cromático e estrutural no processo de recuperação e a precisão da divisão das imagens para preservação de informações espaciais no processo de recuperação. As diferentes combinações supracitadas foram aplicadas à um conjunto de imagens de conteúdo semântico heterogêneo, previamente selecionadas, através de um sistema de recuperação de imagens implementado ao longo desta pesquisa. Os resultados experimentais obtidos estão descritos por meio de curvas “recall x precision”.

Palavras chave: Recuperação de Imagens, Conteúdo Cromático, Conteúdo Estrutural, “Recall x Precision”.

Abstract

This work describes a set of image retrieval techniques by color and structural similarity. These techniques are evaluated through recall x precision curves. During the experiments done, some values were adjusted, allowing the color weight variation, the structural weight variation and the spatial relationships preservation of the image contents. The set of techniques for image retrieval described here led to development of a computational environment for image retrieval by color and structural similarity. Furthermore, recall x precision curved lines was applied in order to verify the effectiveness of the computational system in several configuration modes of image retrieval.

Key words: Image Retrieval, Color Content, Structural Content, Recall x Precision.

1. Introdução

As aplicações para imagens digitais têm crescido a cada dia. Isto se deve ao fato de que o formato digital tem se mostrado o formato ideal para o armazenamento e manipulação das mesmas [6]. Com este crescimento, surge naturalmente a necessidade de se estabelecer técnicas que permitam uma melhor manipulação destes conteúdos. Uma das questões que vêm chamando a atenção de pesquisadores já há algum tempo é a criação de descritores de imagens. Inicialmente, muitos trabalhos surgiram propondo a descrição textual de imagens. Estas propostas logo se depararam com duas dificuldades. Uma é a necessidade de muito trabalho manual para representar imagens em

forma de texto. A outra dificuldade, mais importante, resulta do conteúdo rico presente nas imagens e a subjetividade da percepção humana. Isto é, pessoas diferentes podem perceber características diferentes pertencentes ao conteúdo de uma mesma imagem. Considerando que a recuperação de imagens é uma das aplicações mais frequentes para estes descritores, a subjetividade da percepção e a imprecisão da descrição podem causar falhas incorrigíveis nos processos de recuperação posteriores [9].

Com isto, surgiram as técnicas de recuperação de segunda geração, que deram origem ao que é conhecido como recuperação de imagens baseada em conteúdo. Nestas técnicas, as imagens passaram a ser indexadas pelo seu próprio conteúdo visual, através de algoritmos capazes de extrair das mesmas características tais como: cor, estrutura, textura e forma entre outras.

Este artigo descreve o emprego de um conjunto de técnicas que visam a extração de conteúdo cromático e estrutural de imagens para, posteriormente, aplicá-los em um sistema de recuperação de imagens. Os conteúdos cromático e estrutural das imagens são características imediatamente percebidas pelo observador humano, e desta forma são elementos importantes para diferenciá-las, este fator motivou o emprego destas características no desenvolvimento deste trabalho. O sistema experimental desenvolvido a partir das técnicas aqui abordadas extrai histogramas de cor e estrutura das imagens para, num momento posterior, compará-los através de um cálculo de similaridade aplicado entre os mesmos. A representação de ambas as classes de características através de histogramas permite uma uniformidade de estruturas de dados de cor e orientação de bordas dentro do sistema [1]. Adicionalmente, o sistema é capaz de preservar informações espaciais relacionadas aos conteúdos cromáticos e estruturais extraídos das imagens. A precisão da preservação destas informações pode ser ajustada pelo usuário no momento da consulta.

Para avaliar os resultados obtidos, foram utilizadas curvas “recall x precision”, um método amplamente empregado na verificação do desempenho da efetividade de sistemas de recuperação.

Na seção 2 deste artigo serão descritas as técnicas utilizadas na elaboração dos descritores de conteúdo da imagem, tanto o cromático quanto o estrutural. Na seção 3 será descrito o processo de verificação da similaridade entre duas imagens, com base nos descritores já mencionados. Na seção 4 será descrito o esquema conceitual do processo de recuperação, a partir de uma imagem de consulta, em um repositório de imagens. Na seção 5, será descrita uma avaliação dos testes realizados com o ambiente computacional RICE (Recuperação de Imagens por Cor e Estrutura), implementado ao longo da realização deste trabalho. Esta avaliação foi feita através de curvas “recall x precision”. Na seção 6, são descritas as conclusões obtidas até o momento e trabalhos futuros.

2. Descritores de Conteúdo da Imagem

Na abordagem aqui descrita, os conteúdos visuais extraídos das imagens – descritores das imagens – serão representados através de histogramas. Um histograma pode ser armazenado e manipulado computacionalmente de forma bastante simples, e tem sido apresentado por vários autores – como [14], [3] e [1] - como um meio através do qual se pode descrever o conteúdo de imagens.

2.1 Descritor de Conteúdo Cromático

Dada uma imagem digital, seu conteúdo cromático pode ser descrito através de um histograma de cores. O histograma de cores de uma imagem pode ser descrito computacionalmente através de uma estrutura de dados linear que armazena o número de ocorrências de cada possível valor cromático em uma imagem. O tamanho desta estrutura deve ser igual ao número de cores possíveis no espaço cromático utilizado para representar a imagem. Considerando que muitas imagens digitais têm as cores representadas num espaço de cores com 256^3 possíveis valores diferentes, é necessário fazer uma redução deste número para que os processos de indexação, e posteriormente de recuperação, sejam viáveis computacionalmente. Para isto, optou-se por fazer a quantização das cores utilizadas na representação das imagens durante a formação dos histogramas de cores. Assim, foram

utilizados apenas quatro valores diferentes para representar cada canal do espaço de cores RGB (Red, Green and Blue), o que proporcionou um número total de 4^3 valores possíveis no espaço cromático utilizado na formação destes descritores. Com isto, o tempo de processamento destes histogramas quando da ocorrência do processo de recuperação de imagens reduz significativamente. O histograma é formado por pilhas, uma para cada cor do espaço quantizado, que somam o número de ocorrências de cada cor na zona de imagem correspondente. Na figura 1, descrita a seguir, é ilustrada a formação de um histograma de cores a partir de um pequeno conjunto de pixels. Depois de formado o histograma, ele deve ainda ser normalizado, dividindo-se o valor encontrado em cada pilha pelo número total de pixels presentes na imagem. Desta forma, obtêm-se um percentual da ocorrência de cada cor na zona. Assim, o descritor encontrado pode ser confrontado com outros descritores de imagem independentemente da ocorrência de eventuais diferenças entre as dimensões das imagens a serem comparadas.

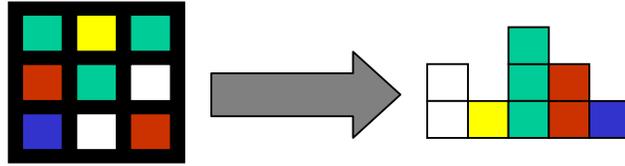


Figura 1. Formação do histograma de cores a partir de um conjunto de pixels

2.2 Descritor de Conteúdo Estrutural

O conteúdo estrutural de uma imagem também será descrito através de histogramas. A partir de cada pixel da imagem discreta será calculado um par de valores $\langle \rho, v \rangle$ que será utilizado para compor o histograma. O par $\langle \rho, v \rangle$ será um indicador de características estruturais do seu ponto correspondente na imagem. O valor de v indica a orientação da borda presente no ponto, enquanto que o valor de ρ representa a intensidade da borda neste ponto.

Para isto, são computados gradientes nas direções x e y (δ^x , δ^y) para cada pixel p_{ij} da imagem, conforme descrito na figura 2 [1] e [5].

$$\delta_{i,j}^x = \frac{p_{i+1,j} - p_{i-1,j}}{\sqrt{2(p_{i+1,j}^2 + p_{i-1,j}^2)}} \quad \delta_{i,j}^y = \frac{p_{i,j+1} - p_{i,j-1}}{\sqrt{2(p_{i,j+1}^2 + p_{i,j-1}^2)}}$$

Figura 2. Equações para cálculo dos gradientes

Onde $p_{i,j}$ representa a intensidade de cinza presente no pixel da i -ésima linha e j -ésima coluna.

A orientação de borda e intensidade são calculadas, respectivamente, pelas equações descritas na figura 3.

O resultado da fase de detecção de bordas da imagem é um conjunto de bordas $\mathcal{E} = \{B_1, \dots, B_\vartheta\}$, onde ϑ corresponde ao número de pontos de borda contidos na imagem. Cada borda B_i é

$$v_{i,j} = \tan^{-1} \left(\frac{\delta_{i,j}^x}{\delta_{i,j}^y} \right) \quad \rho_{i,j} = \sqrt{(\delta_{i,j}^x)^2 + (\delta_{i,j}^y)^2}$$

Figura 3. Equações para cálculo da orientação e intensidade da borda

descrita por um par $\langle \rho, v \rangle$ onde ρ é a intensidade da margem e v sua orientação. O valor da intensidade é discretizado em l ($l=128$) níveis, pois com este número conseguiu-se um poder razoável de representação das intensidades sem comprometer o desempenho do sistema implementado. Depois, um histograma de intensidades $E \equiv \langle e_1, \dots, e_l \rangle$ é computado.

Para descrever um conteúdo estrutural de uma imagem genérica, somente bordas relevantes da imagem devem ser consideradas. Para isto, um limite mínimo τ descarta todas aquelas bordas de imagens cujas intensidades não o ultrapassem. Assim, uma nova imagem é produzida, preservando somente a informação de orientação sobre bordas $B_i \equiv \langle \rho_i, v_i \rangle$ para as quais $\lfloor \rho_i \rfloor \geq \tau$. O valor de τ é calculado de forma que a razão entre o número de bordas preservadas e o número total de bordas esteja próximo a 15 % (conforme figura 4) [1], pois com esta razão conseguiu-se, na média, bons resultados no que diz respeito à representação do aspecto estrutural relevante da imagem:

O espaço de orientação é discretizado em t ($t=4$) níveis e um histograma de orientação é computado para a imagem. Os quatro valores presentes no espaço de orientação discretizado são suficientes

para representar inclinações de bordas correspondentes às direções dos ângulos 0° e 180° , 45° e 225° , 90° e 270° , e 135° e 315° . Estas direções encontram-se uniformemente distribuídas no círculo trigonométrico.

$$\tau: \frac{\sum_{k=\tau}^l e_k}{\sum_{k=1}^l e_k} \cong 0.15$$

Figura 4. Equação para determinação das bordas relevantes

Assim, cada histograma descritor de conteúdo estrutural será composto por um total de quatro colunas, cada uma delas correspondendo a uma das orientações possíveis para v em cada ponto de borda da imagem. Desta forma, para cada ocorrência de borda com um determinado ângulo de inclinação v e com intensidade suficiente para ser considerada relevante ($\rho \geq \tau$), a pilha correspondente a este ângulo de inclinação v deverá ser incrementada em uma unidade. A figura 5, descrita a seguir, ilustra a formação de um histograma de estrutura para um pequeno conjunto de pontos de borda. Observe que nesta figura os pixels correspondentes a bordas relevantes estão representados em preto, enquanto que os não relevantes estão representados em cinza e não são contabilizados na formação do histograma.

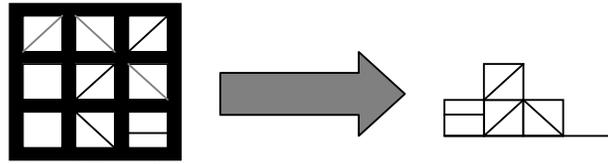


Figura 5. Formação do histograma de estrutura a partir de um conjunto de pixels

A elaboração do histograma de estruturas é concluída com a normalização do mesmo. Para isto, divide-se o valor encontrado em cada pilha pelo número total de pixels correspondentes a bordas relevantes presentes na zona de imagem. Desta forma, obtêm-se uma frequência de ocorrência de bordas para cada uma das quatro orientações possíveis. Assim, o descritor encontrado pode ser confrontado com outros descritores de imagem independentemente da ocorrência de eventuais diferenças entre as dimensões das imagens a serem comparadas.

2.3 Preservação de Informações Espaciais (Divisão em Zonas)

Nesta abordagem, foi introduzida uma estratégia para permitir a preservação de informações espaciais referentes aos conteúdos extraídos das imagens. Para isto, a imagem original pode ser dividida em zonas. Desta forma, determina-se de que forma será dividida a imagem de consulta utilizada na recuperação estabelecendo o número de faixas horizontais e verticais segundo os quais as zonas retangulares de imagem deverão ser criadas. Esta divisão em zonas feita na imagem de consulta implicará em uma igual divisão nas imagens que serão comparadas com a mesma.

Com esta estratégia, a imagem será dividida em $n \times m$ zonas onde n corresponde ao número de faixas horizontais e m ao número de faixas verticais, extraíndo-se um histograma descritor de conteúdo cromático e um histograma descritor de conteúdo estrutural de cada uma destas zonas. Na figura 6, descrita a seguir, pode-se verificar a formação de zonas em uma imagem com $n=3$ e $m=3$.

Assim, dada uma imagem genérica I , e seu particionamento em $q = n \times m$ (com $1 \leq n \leq 10$ e $1 \leq m \leq 10$) pedaços conforme descrito anteriormente, os conteúdos cromático e estrutural da imagem pode ser expresso através de um conjunto de q histogramas $\{H_i(I)\}_{i=1}^q$, onde o i -ésimo histograma $H_i(I)$ representa o conteúdo cromático ou estrutural da i -ésima zona da imagem I .

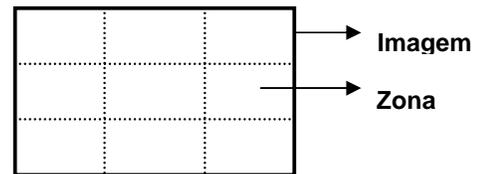


Figura 6. Divisão da imagem em zonas com $n=3$ e $m=3$

3. Verificação da Similaridade Entre Imagens

Depois de extraídos os descritores de conteúdo de cada zona presente na imagem, conforme descrito na seção anterior, pode-se compará-los aos descritores de zonas extraídos de outra imagem a fim de

$$D_{n \times m}(H'_{n \times m}, H''_{n \times m}) = (H'_{n \times m} - H''_{n \times m})^T A (H'_{n \times m} - H''_{n \times m})$$

Figura 7. Equação para cálculo da distância entre histogramas

mensurar a diferença perceptual entre os mesmos através do cálculo da Distância Euclideana Ponderada, mencionada em [3] e [11] e descrito na equação da figura 7.

Onde:

- $H'_{n \times m}$ representa o histograma (de cor ou estrutura), da zona posicionada na n-ésima linha e m-ésima coluna de uma das imagens a serem comparadas (imagem 1);
- $H''_{n \times m}$ representa o histograma, também da zona posicionado na n-ésima linha e m-ésima coluna da outra imagem a ser comparada (imagem 2);
- A representa uma matriz de distâncias simétrica, que será apresentada em maiores detalhes nas seções 3.1 e 3.2;

Cada elemento a_{ij} da matriz de distâncias A representa o quanto a diferença entre a i-ésima e a j-ésima pilha do histograma contribuem para D . Neste caso, como a similaridade entre as imagens pode ser verificada através do conteúdo cromático ou estrutural, constrói-se as matrizes de distância distintas A_C e A_E . Estas matrizes são definidas para a computação da distância entre histogramas de cor e histogramas de estrutura (orientação da borda) respectivamente. Estas matrizes de distâncias de cores são matrizes simétricas $C \times C$ (onde C é o número de cores de referência, neste caso $C=64$) cujos elementos a_{ij} codificam a similaridade de percepção entre a i-ésima e a j-ésima cor de referência. A matriz de distâncias de estrutura A_E é uma matriz simétrica 4×4 (considerando 4 valores de orientação de referência) cujos elementos a_{ij} codificam a similaridade de percepção entre a i-ésima e a j-ésima orientação de referência [1].

Para a extração de uma medida final das diferenças dos conteúdos globais das duas imagens que estão sendo comparadas, é feita uma média aritmética simples das diferenças encontradas entre os histogramas das zonas correspondentes em cada imagem, conforme ilustrado na figura 9 e cujo cálculo está descrito na figura 8.

Onde:

- I_1 e I_2 representam as duas imagens que estão sendo comparadas;
- H' e H'' representam os histogramas descritores de conteúdo das imagens I_1 e I_2 respectivamente;
- N e M são, respectivamente, o número de linhas e colunas de zonas componentes das imagens, ajustados para uma determinada comparação entre imagens;

$$D(I_1, I_2) = \frac{\sum_{n=1}^N \sum_{m=1}^M D_{n \times m}(H'_{n \times m}, H''_{n \times m})}{N \times M}$$

Figura 8. Equação para cálculo da diferença média entre os histogramas das zonas

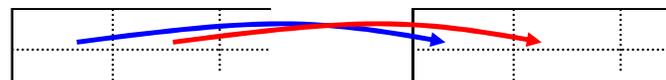


Figura 9. Zonas correspondentes das imagens que estão sendo comparadas.

3.1 A Matriz de Distâncias A_C

A métrica descrita na equação, mostrada na figura 7, para a verificação de similaridade de imagens exige o uso de uma matriz de distâncias A .

A matriz A_C é uma matriz que armazena as distâncias entre as cores presentes no espaço quantizado sobre o qual os histogramas de cores, referentes às zonas, descritores da imagem foram formados. Entretanto, é importante observar que este espaço reduzido de cores foi formado a partir da discretização do espaço segundo o qual as cores estavam originalmente descritas nos arquivos de imagem, o espaço RGB com 256^3 possibilidades de cores.

Esta matriz é uma matriz simétrica, de dimensão 64×64 , que armazena em cada uma de suas células a_{ij} a similaridade perceptual entre a i-ésima e a j-ésima cor. Em seguida, calcula-se as diferenças d_{ij} entre as cores i e j através da equação mostrada na figura 10. Finalmente, a matriz de distâncias entre cores é construída colocando-se em cada célula a_{ij} o valor da

$$a_{ij} = 1 - (d_{ij} / \max(d_{ij}))$$

Figura 10. Equação para cálculo da diferença entre cores

diferença d_{ij} , dividido pelo valor máximo de distância entre duas cores, subtraído de 1, conforme mostra a próxima equação. Desta forma, a_{ij} varia entre 0 e 1, já que mede similaridade entre duas cores.

3.2 A Matriz de Distâncias A_E

Conforme descrito anteriormente, além da matriz discutida na seção anterior, que apóia o processo de verificação da similaridade entre imagens baseada nos seus conteúdos cromáticos, também será construída uma matriz para apoiar o processo de verificação da similaridade entre imagens com base nos seus conteúdos estruturais, esta será a matriz A_E .

Os histogramas descritores de conteúdos estruturais da imagem possuem quatro colunas, cada uma representando uma das possíveis inclinações de borda. A matriz A_E deve ter em suas células a_{ij} a medida de similaridade perceptual entre as orientações i e j .

Para isto, a matriz A_E construída será uma matriz simétrica 4x4 cujas células a_{ij} terão os seus valores encontrados através da equação descrita na figura 11.

$$a_{ij} = 1 - (d_{ij}/90)$$

Figura 11. Equação para cálculo da similaridade perceptual entre duas orientações de borda

Onde d_{ij} é a diferença perceptual entre as direções i e j , e pode ser calculado pelas equações da figura 12. Sendo que, nesta figura, i e j são ângulos descritos em graus, múltiplos de 45°, no intervalo [0°, 180°].

$$\begin{aligned} d_{ij} &= |i-j| && \text{para } |i-j| \leq 90 \\ d_{ij} &= 180 - i + j && \text{para } |i-j| > 90 \text{ e } i > j \\ d_{ij} &= 180 - j + i && \text{para } |i-j| > 90 \text{ e } j > i \end{aligned}$$

Figura 12. Equações para cálculo da diferença perceptual entre duas orientações de borda

3.3 Uma Medida de Similaridade Composta (Cor e Estrutura)

A abordagem adotada neste trabalho permite ainda a extração de uma medida de similaridade entre imagens composta, sendo baseada nas medidas de diferença de cor e estrutura descritas anteriormente. Para isto, basta que o usuário determine valores aos pesos p_c e p_e , para a similaridade por cor e para a similaridade por estrutura respectivamente. A soma dos pesos p_c e p_e deve ser igual a 1, e para o cálculo da medida de similaridade entre as imagens composta, calcula-se uma média aritmética ponderada entre as medidas de diferença por cor e por estrutura, descritas nas seções anteriores, com seus respectivos pesos p_c e p_e . A seguinte equação, ilustrada na figura 13, descreve o cálculo diferença baseado nos aspectos cromático e estrutural com pesos atribuídos para ambos. É importante ressaltar que a similaridade é o complemento da diferença.

$$D(I_1, I_2) = D_C(I_1, I_2)p_c + D_E(I_1, I_2)p_e$$

Figura 13. Cálculo da diferença entre imagens baseada nos aspectos cromático e estrutural com pesos atribuídos para ambos

Onde:

- $D(I_1, I_2)$ é a diferença entre as imagens I_1 e I_2 ;
- $D_C(I_1, I_2)$ é a diferença cromática entre as imagens I_1 e I_2 ;
- $D_E(I_1, I_2)$ é a diferença estrutural entre as imagens I_1 e I_2 ;

4. Recuperação de Imagens Através de Descritores de Conteúdo

4.1 Esquema conceitual da recuperação de imagens por conteúdo

A recuperação de imagens segundo a técnica aqui descrita parte de uma imagem de consulta. Para isto, o mecanismo de recuperação deverá extrair descritores de características visuais (cor e estrutura) conforme descrito na seção 2. Os descritores resultantes da imagem de consulta são utilizados para fazer a busca de imagens mais similares no repositório. Para isto, percorre-se o repositório de descritores de conteúdos extraídos de uma coleção de imagens e compara-os ao descritor da imagem de consulta previamente calculado. A extração de descritores de conteúdo deverá se dar de acordo com os parâmetros previamente ajustados pelo usuário. Por exemplo, se o usuário ajustou peso 1 para o aspecto cromático, e peso 0 para o aspecto estrutural em todas as

zonas da imagem, o sistema fará somente a extração do descritor cromático a fim de poupar tempo na realização da recuperação.

A comparação realizada entre os descritores de imagens mencionada no parágrafo anterior, se dará conforme descrito na seção 3.3, sendo que para isto será considerado inclusive o número de zonas previamente definido, tanto na extração dos descritores quanto na realização do cálculo de similaridade.

Depois de calculadas as medidas de similaridade entre a imagem de consulta e cada uma das imagens do repositório, a recuperação pode ser caracterizada pela exibição das imagens ordenadas de forma decrescente pelas suas respectivas medidas de similaridade encontradas. Este esquema conceitual da recuperação de imagens encontra-se ilustrado na figura 14, descrita a seguir. A fim de avaliar resultados experimentais acerca das técnicas de recuperação de imagens aqui descritas, foi desenvolvido

ao longo deste trabalho um ambiente computacional para recuperação de imagens baseado nestas técnicas, este sistema será apresentado na próxima seção.

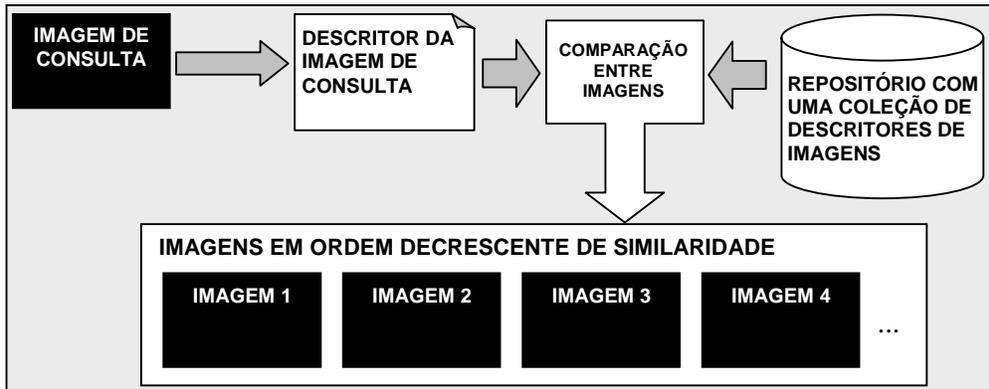


Figura 14. Esquema conceitual de um sistema de recuperação de imagens

5. Avaliação de Resultados Experimentais

A fim de avaliar os resultados obtidos pelo conjunto de técnicas empregadas no desenvolvimento deste trabalho, foi desenvolvido o ambiente computacional RICE (Recuperação de Imagens por Cor e Estrutura). Neste ambiente é possível realizar a recuperação a partir de uma imagem de consulta ou ainda comparar a similaridade entre pares de imagens. O modo principal da interface do sistema RICE (no modo de recuperação) pode ser visto na figura 15.

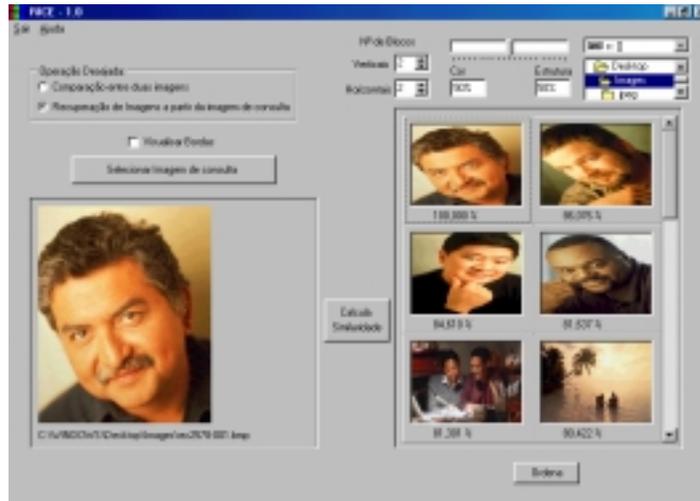


Figura 15. Interface do sistema RICE no modo de recuperação de imagens

5.1 Padrões Artificiais

Inicialmente, foram escolhidos alguns padrões artificiais, mostrados na figura 16, a fim de verificar a coerência dos resultados obtidos na comparação entre pares de imagens. Os padrões artificiais adotados, têm características cromáticas e estruturais que, quando comparadas, podem contribuir significativamente nesta verificação.

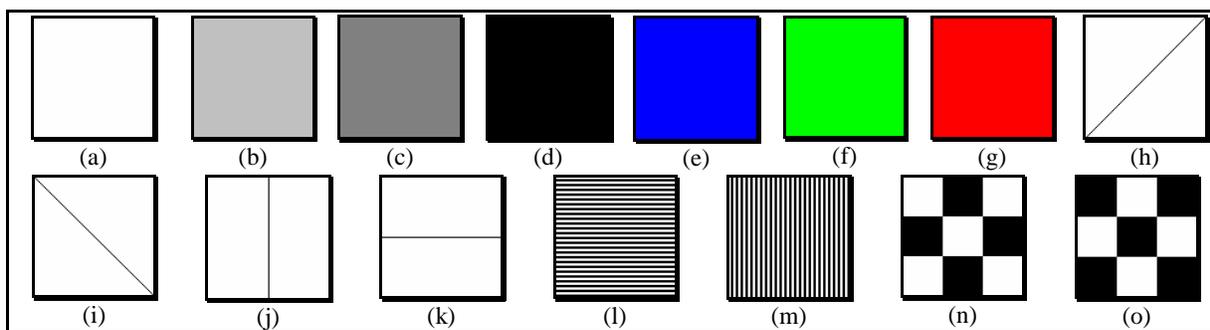


Figura 16. Padrões artificiais utilizados nos testes

A tabela, descrita a seguir na figura 17, é uma esquematização que procura demonstrar resultados obtidos em algumas comparações realizadas entre as imagens mostradas na figura anterior através do sistema RICE. Em cada linha da tabela será descrita uma comparação entre duas imagens sendo que: na segunda e terceira coluna estão identificadas as imagens a serem comparadas (de acordo com legenda da figura 16); na quarta e quinta coluna estão descritos respectivamente os números de faixas de zonas verticais e horizontais ajustados no sistema; na sexta e sétima coluna estão descritos os pesos (em percentual) estabelecidos para os aspectos cromáticos e estrutural respectivamente; e, finalmente, na última coluna é apresentado o resultado obtido na comparação realizada.

	Imagem 1	Imagem 2	Nº de zonas verticais	Nº de zonas horizontais	Peso Cor (%)	Peso Estrutura	Similaridade (%)
1	6(a)	6(d)	1	1	100	0	0
2	6(a)	6(d)	1	1	0	100	100
3	6(a)	6(d)	1	1	50	50	50
4	6(h)	6(i)	1	1	100	0	99,98
5	6(h)	6(i)	1	1	0	100	0
6	6(j)	6(k)	1	1	0	100	0
7	6(j)	6(h)	1	1	0	100	29,29
8	6(l)	6(m)	1	1	100	0	100
9	6(l)	6(m)	1	1	0	100	0
10	6(k)	6(m)	1	1	0	100	0
11	6(n)	6(o)	1	1	0	100	100
12	6(n)	6(o)	1	1	100	0	88,88
13	6(n)	6(o)	3	3	100	0	0
14	6(n)	6(o)	2	2	100	0	88,88
15	6(n)	6(o)	3	3	0	100	100
16	6(n)	6(o)	2	2	0	100	100
17	6(f)	6(e)	1	1	100	0	6,09
18	6(f)	6(g)	1	1	100	0	7,81
19	6(e)	6(g)	1	1	100	0	13,67
20	6(g)	6(a)	1	1	100	0	9,63
21	6(g)	6(d)	1	1	100	0	27,40
22	6(d)	6(c)	1	1	100	0	29,42
23	6(d)	6(b)	1	1	100	0	13,45
24	6(a)	6(b)	1	1	100	0	49,90

Figura 17. Resultados obtidos em comparações entre padrões artificiais no sistema RICE

5.2 Imagens Reais

Nesta seção serão descritos os resultados experimentais obtidos através do sistema RICE na recuperação de imagens. Para isto, foi utilizado um conjunto de 287 imagens com conteúdos semânticos bastante variados. As bibliotecas utilizadas encontram-se disponíveis em [4], e possuem imagens que mostram desde paisagens naturais, como montanhas, bosques, nuvens, vales, paisagens

litorâneas, até fotos de pessoas e paisagens urbanas, que ilustram palácios, prédios, igrejas e carros entre outros.

A partir deste repositório, com algumas imagens reunidas em grupos por similaridade de conteúdo semântico (figura 20), verificou-se o desempenho do sistema de recuperação de imagens desenvolvido, no que diz respeito à efetividade dos processos de recuperação de imagens oferecidos pelo mesmo. Para cumprir este propósito, foram utilizadas curvas de “recall x precision”, um método para verificação de desempenho de sistemas de recuperação utilizado em vários trabalhos tais como [7], [8] e [13] e que se encontra descrito em [2], [10] e [12].

5.2.1 Recall x Precision

Considere uma consulta Q dirigida à uma sistema de recuperação de imagens que fará a busca em um conjunto de imagens, o repositório I. Considere ainda que, para a consulta Q existe um conjunto de imagens R, subconjunto de I, que é formado pelas imagens relevantes para esta consulta. Considere |R| como sendo o número de imagens pertencentes ao conjunto R.

Um sistema de recuperação de imagens deve possuir uma estratégia segundo a qual as imagens retornadas devem ser apresentadas. Geralmente, apresenta-se as imagens do repositório em ordem decrescente de similaridade em relação à imagem de consulta. A partir de agora, este conjunto de imagens retornadas por uma consulta será chamado de conjunto A. Adicionalmente, |A| será igual ao número de imagens pertencentes ao conjunto A, e |Ra| será o número de imagens que pertencem tanto ao conjunto R quanto ao conjunto A.

Recall é a porcentagem de documentos relevantes (o conjunto R) que foram recuperados pela consulta, e pode ser obtido pela equação da figura 18.

$$Recall = \frac{|Ra|}{|R|}$$

Figura 18. Equação para cálculo de recall

Precision é a porcentagem dos documentos recuperados (o conjunto A) que são relevantes, e pode ser obtido pela equação da figura 19.

$$Precision = \frac{|Ra|}{|A|}$$

Figura 19. Equação para cálculo de precision

As curvas “recall x precision” mostram o comportamento da medida de precision na medida em que evolui o valor de recall para uma determinada consulta.

5.2.2 Avaliação de Resultados

A figura 20 mostra quatro das imagens de consulta utilizadas nos experimentos realizados juntamente de seus respectivos conjuntos de imagens similares. Ao todo foram utilizadas dez imagens de consulta sendo que o número de imagens presentes em seus respectivos conjuntos similares era variável. As imagens similares são escolhidas por um especialista humano, que procura identificar e agrupar imagens que tenham conteúdo semântico próximos entre si.

Imagem de Consulta	Imagens Relevantes			
				
				
				



Figura 20. Imagens de consulta e seus respectivos conjuntos de imagens similares

A partir das imagens de consulta escolhidas, foram realizadas consultas com diferentes ajustes para os pesos dos conteúdos cromático e estrutural, e da precisão da preservação de informações espaciais. As curvas “recall x precision” descritas a seguir mostram o desempenho da efetividade do processo de recuperação colhido durante os experimentos.

A figura 21, mostra a curva “recall x precision” para algumas consultas realizadas. Nestas consultas foram feitas variações nos pesos dos aspectos cromático e estrutural das imagens a fim de se verificar em que medida os mesmos influenciam nos resultados obtidos pelas técnicas utilizadas neste trabalho. As imagens foram divididas em nove zonas (3x3) a fim de se preservar informações especiais durante a extração dos descritores das mesmas. Os resultados exibidos na figura 21 referem-se às curvas médias obtidas pelas consultas realizadas com as dez imagens de consulta mostradas na figura 20.

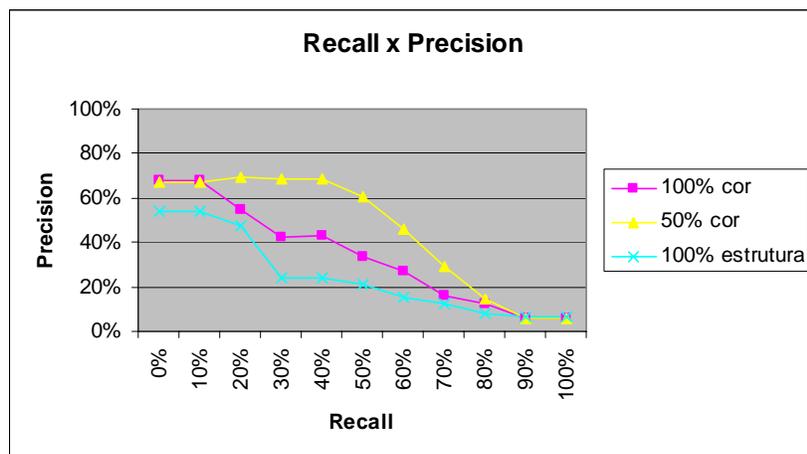


Figura 21. Curvas das médias das consultas 3x3 com variação dos pesos dos aspectos cromático e estrutural

Os resultados exibidos na figura 21 referem-se às curvas médias obtidas pelas consultas realizadas com as dez imagens de consulta mostradas na figura 20.

As curvas obtidas mostram que a combinação dos aspectos visuais de cor e estrutura trouxe melhores resultados, já que a curva “recall x precision” para as consultas onde foi empregado peso 50% tanto para o aspecto cromático quanto para o aspecto estrutural mantém níveis de precision mais altos na medida em que se aumenta o valor de recall.

Além dos testes descritos anteriormente, foram realizadas também consultas com o objeto de se verificar em que medida as divisões das imagens em zonas, para extração dos descritores, podem influenciar nos resultados das consultas. Nestas consultas foram feitas variações nos padrões de divisão da imagem em zonas. Os padrões utilizados foram 1x1 (apenas uma zona), 3x3 (total de nove zonas) e 10x10 (total de 100 zonas). Os pesos dos aspectos cromático e estrutural foram ajustados para 50% durante a realização destas consultas. Os resultados exibidos na figura 22 referem-se às curvas médias obtidas pelas consultas realizadas com as dez imagens de consulta mostradas na figura 20.

As curvas obtidas mostram que a divisão da imagem em zonas, com conseqüente preservação de informações espaciais, trouxe melhores resultados ao processo de recuperação, já que as curvas “recall x precision” para as consultas onde foram empregadas as estratégias para divisão da imagem em zonas, tanto 3x3 quanto 10x10, mantêm níveis de precision mais altos na medida em que se aumenta o valor de recall.

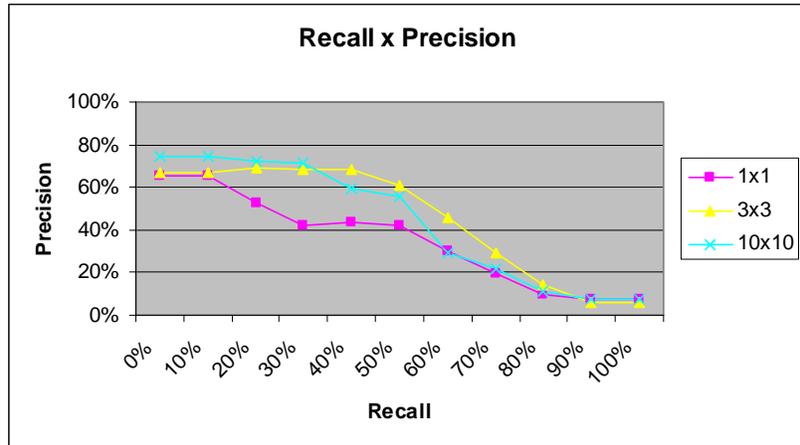


Figura 22. Curvas das médias das consultas com variação das divisões das zonas para preservação de informações espaciais

6. Conclusões e Trabalhos Futuros

A recuperação de imagens baseada em conteúdo é um tópico de pesquisa aberto à investigações. Neste trabalho, foram abordadas e avaliadas técnicas para a extração automática de descritores de imagens para, posteriormente, utiliza-los em processos de recuperação. Com a realização deste trabalho verificou-se que a possibilidade de se preservar informação espacial referente aos conteúdos (cromáticos e estruturais) identificados nas imagens, e ainda a possibilidade de se ajustar a os pesos dos aspectos visuais (cor e estrutura) empregados no processo de recuperação de imagens podem influenciar significativamente nos resultados obtidos no processo de recuperação de imagens.

Além das verificações realizadas neste trabalho, pretende-se ainda utilizar estratégias que permitam o uso de espaços de cores perceptualmente uniformes. Pretende-se também realizar experimentos que envolvam técnicas de segmentação de imagens por regiões durante a extração dos histogramas descritores da imagem. Adicionalmente, pretende-se empregar Earth Mover’s Distance como métrica para verificação de similaridade entre as imagens. Os resultados a serem obtidos por estes experimentos deverão ser descritos através de curvas “recall x precision” a fim de serem comparados com os resultados já obtidos até o momento.

Agradecimentos

Agradecemos de maneira especial à Universidade Paranaense – UNIPAR – por nos apoiar no desenvolvimento deste trabalho.

Referências Bibliográficas

- [1] ASSFALG, Jürgen; PALA, Pietro. *Querying by Photographs: A VR Metaphor for Image Retrieval*. IEEE Multimedia, Los Alamitos v. 7, n. 1, p. 52-59, 2000.
- [2] BEAZA-YATES, R.; RIBEIRO-NETO, B. *Modern Information Retrieval*. [S.l.]: Addison-Wesley, 1999. p. 74-81.
- [3] BIMBO, Alberto del. *Visual Information Retrieval*. San Francisco: Morgan Kaufmann, 1999. p. 81-113.

- [4] COMPUTER Vision Test Images: Disponível em: <<http://www.cs.washington.edu/research/imagedatabase/groundtruth/>>. Acesso em: 10 out. 2004.
- [5] FACON, Jaques. *Processamento e Análise de Imagens*. Curitiba: Pontifícia Universidade Católica do Paraná, 2001. Disponível em: <<ftp://ftp.ppgia.pucpr.br>>. Acesso em: 04 fev. 2001.
- [6] GOMES, Jonas; VELHO, Luiz. *Computação Gráfica: imagem*. Rio de Janeiro: IMPA/SBM, 1994. p. 63-127.
- [7] RAO, Aibing et al. Spatial Color Histograms for Content-based Image Retrieval. In: *IEEE International Conference on Tools with Artificial Intelligence, ICTAI*, 11., 1999, Chicago, Illinois. Proceedings... [S.l.: s.n.], 1999. p. 183-186.
- [8] RIKXOORT, Eva M. Content Based Image Retrieval: Utilizing, Color. Nijmegen: Faculty of Social Sciences, Radboud University Nijmegen, 2005.
- [9] RUI, Yong et al. Image Retrieval: Past, Present, and Future. In: *Symposium on Multimedia Information Processing*, 12., 1997, Taipei, Taiwan. Proceedings... [S.l.: s.n.], 1997.
- [10] SALTON, Gerard; MCGILL, Michael J. *Introduction to Modern Information Retrieval*. New York: McGraw-Hill, 1983. p. 157-191.
- [11] SCHETTINI, Raimondo et al. *A Survey of Methods for Colour Image Indexing and Retrieval in Image Databases*. 2001. Disponível em: <<http://www.itim.mi.cnr.it/Staff/Schettini/publicazioni.htm>>. Acesso em: 30 maio 2002.
- [12] SMITH, John R. *Retrieval Effectiveness*. Disponível em: <<http://disney.ctr.columbia.edu/jrsthesis/node72.html>>. Acesso em: 05 jun. 2002.
- [13] SMITH, John R.; CHANG, Shih-Fu. Tools and Techniques for Color Image Retrieval. *IS&T/SPTTE Proceedings*, [S.l.], v. 2670, p. 1-12, [199?].
- [14] SONKA, Milan et al. *Image Processing, Analysis, and Machine Vision*. Pacific Grove: ITP, 1999. p. 1-55.