

# Integrando Textura e Forma para a Recuperação de Imagens por Conteúdo

André G. R. Balan<sup>1</sup>, Agma J. M. Traina<sup>2</sup>, Caetano Traina Jr.<sup>3</sup>, Paulo M. Azevedo-Marques<sup>4</sup>

<sup>1,2,3</sup>Grupo de Base de Dados e Imagens (GBDI), Instituto de Ciências Matemáticas e de Computação (ICMC), Universidade de São Paulo (USP), Brasil.

<sup>4</sup>Centro de Ciências das Imagens e Física Médica (CCIFM), Faculdade de Medicina de Ribeirão Preto (FMRP), Universidade de São Paulo (USP), Brasil.

**Resumo** - Este artigo apresenta uma nova abordagem de recuperação de imagens por conteúdo que utiliza um conjunto de características extraídas a partir da imagem segmentada com base em textura. O algoritmo de segmentação implementa um processo estocástico e tem como base os Campos Aleatórios de Markov. A segmentação é automática e tem como parâmetro o número de classes com texturas diferentes. As regiões obtidas na segmentação guiam a extração de medidas da imagem segmentada e da imagem original, produzindo um vetor de característica de 30 dimensões, utilizado para caracterizar o conteúdo da imagem. Experimentos realizados revelaram que o conjunto de características proposto tem um alto poder de discriminação quando comparado com técnicas mais tradicionais de distribuição de brilho (histograma).

**Palavras-chave:** CBIR, Segmentação de Imagens, Textura, Forma.

**Abstract** - This paper presents a new approach to retrieve images by content using a set of characteristics extracted from the texture-based segmented image. The segmentation algorithm implements a random process and is based on Markov Random Fields. The segmentation is automatic and has as parameter the number of desired classes with different textures. The regions obtained by the segmentation step lead the extraction of measures from segmented and original images, producing a 30-dimensional feature vector, which is used to characterize the image content. Experiments showed that the proposed feature vector has higher discrimination power when compared with traditional techniques based on image's brightness distribution (histogram).

**Key-words:** CBIR, Image Segmentation, Texture, Shape.

## Introdução

A crescente quantidade de imagens adquiridas em hospitais e em centros médicos tem motivado a pesquisa de algoritmos automáticos para recuperação de imagens baseada em suas características intrínsecas. As técnicas de recuperação de imagens baseada em conteúdo (*Content-Based Image Retrieval* - CBIR) consideram a informação visual da imagem e não apenas uma simples descrição textual sobre a mesma. De fato, existe um certo nível de subjetividade em se caracterizar uma imagem pelo seu conteúdo, visto que diferentes especialistas podem estar interessados em diferentes aspectos de uma mesma imagem. Observa-se também que, a recuperação de imagens baseada na descrição textual e a recuperação baseada em conteúdo são abordagens independentes e ortogonais, mas, de forma prática, podem ser integradas em um único sistema a fim de proporcionar um maior poder de recuperação em Sistemas de Informação Hospitalar (SIH) ou em Sistemas de Comunicação e

Arquivamento de Imagens (*Picture Archiving and Communication Systems* - PACS) [4].

A busca por imagens em uma base de dados realizada por um sistema CBIR, envolve a comparação de pares de imagens e a atribuição de um coeficiente de similaridade a cada par, onde tal medida de similaridade é obtida a partir de uma função de distância pré-definida (dissimilaridade). Informações sobre o conteúdo da imagem são obtidas através da extração sistemática de características que devem ser escolhidas de modo a representar a "essência" do conteúdo das mesmas, tomando-se como base a percepção visual de um especialista da área. Assim, um problema importante na área de CBIR é o desenvolvimento de métodos eficazes e eficientes de extração de características para representação adequada do conteúdo das imagens, proporcionando uma recuperação mais próxima daquela idealizada por um especialista. Comumente as características de imagens são agrupadas em três classes: cor, textura e forma [1, 10]. Idealmente, estas características devem ser integradas para proporcionar melhor eficácia no

processo de comparação. A primeira e mais simples classe de características são as características de cor, ou distribuição de brilho, dada principalmente por histogramas da imagem. A extração de características de forma e textura constitui uma tarefa bem mais complexa e computacionalmente mais cara.

O presente artigo apresenta uma abordagem singular de extração de características da imagem para serem utilizadas em um sistema de recuperação de imagens baseado em conteúdo. A técnica consiste em segmentar automaticamente a imagem com base em textura utilizando um algoritmo estocástico baseado nos campos aleatórios de Markov. A partir da imagem segmentada, são realizadas medidas sobre as regiões encontradas e a imagem original, obtendo um conjunto de 30 características. Os experimentos realizados demonstraram que as características propostas possibilitam uma caracterização apurada do conteúdo da imagem e apresentam um nível de discriminação maior do que a técnicas mais tradicionais baseadas na distribuição de brilho e cor (histogramas) da imagem.

### **Conceitos preliminares**

A maioria dos sistemas de recuperação de imagens utiliza características de distribuição de cor. As técnicas mais populares são baseadas nos histogramas de cor ou brilho [3]. Embora o processo utilizado para se obter os histogramas da imagem tenha custo computacional baixo, já que é obtido em tempo linear considerando o tamanho da imagem, o histograma produz uma representação ambígua da mesma, visto que duas imagens diferentes podem ter o mesmo histograma. Entretanto, os histogramas podem fornecer uma rápida pré-seleção para sistemas de recuperação de imagens, e o resultado da pré-seleção pode ser posteriormente refinado por técnicas mais elaboradas e complexas, e conseqüentemente mais caras. Porém, a aplicação de técnicas mais complexas em um conjunto menor de imagens reduz consideravelmente o custo total da recuperação.

A respeito das imagens médicas, acreditamos que as características baseadas em textura e forma das obtidas das regiões da imagem podem discriminar e separar as imagens de modo mais apurado. Diversos estudos encontrados na literatura apresentam técnicas baseadas em ambas as abordagens, tais como [6, 9] para a textura e [8, 11] para a forma.

Neste artigo é considerada a recuperação das  $N$  imagens mais próximas da imagem de busca, por exemplo: "*encontre as três imagens*

*mais similares à imagem identificada como RM\_torax527*". Tal pergunta pode ser respondida se existir uma função de distância (dissimilaridade) para comparar as características da imagem de busca (*RM\_torax527*) com as características de todas as imagens na base de dados.

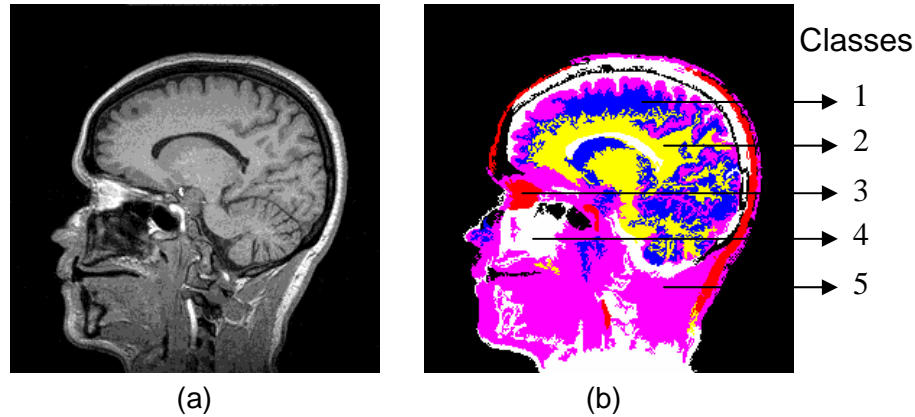
Sempre que uma técnica nova para a extração da característica é proposta, é necessário realizar uma avaliação de sua eficácia. A utilização do gráfico de *Precision x Recall* [2] (Precisão x Revocação) constitui uma abordagem simples e bem conhecida para tal avaliação e, deste modo, foi selecionada para avaliar e comparar a eficácia da técnica proposta. A precisão (*precision*) representa a porção de imagens recuperadas que são relevantes à consulta. Já a revocação (*recall*) corresponde à porção de imagens relevantes que foram recuperadas. Na prática, uma curva mais próxima do alto do gráfico representa maior eficácia na recuperação.

### **A técnica proposta**

Em imagens médicas, as características de textura e forma estão fortemente relacionadas, devido ao fato que os objetos (órgãos, tecidos e anomalias) têm, na maioria das vezes, forma e textura específicas. As diferentes texturas dos objetos da imagem podem ser empregadas para a delimitação dos mesmos, caracterizando assim uma segmentação baseada em textura. Neste tipo de segmentação, cada textura distinta é representada por uma classe. Na imagem segmentada, cada pixel é classificado como pertencente a uma determinada classe e, pixels da mesma classe constituem uma região de textura homogênea. As características propostas neste estudo são obtidas extraindo-se medidas dessas regiões homogêneas. A Figura 1 ilustra uma imagem segmentada utilizando-se cinco classes.

### **O método de segmentação**

A etapa de segmentação baseada em textura é o passo inicial para a extração das características da imagem. O método de segmentação empregado é um algoritmo que implementa um processo estocástico de otimização. Trata-se de um método iterativo, cujo objetivo é minimizar o número previsto de pixels classificados erroneamente. Para se chegar a este objetivo, um campo aleatório de Markov é empregado para modelar/representar o mapa de classificação dos pixels. Nesta representação, cada pixel corresponde a um elemento do campo que, por sua vez, governa a influência que a classificação de um pixel exerce na classificação dos pixels vizinhos.



**Figura 1** – Exemplo de segmentação. (a) imagem original; (b) imagem segmentada em 5 classes.

O processo de segmentação alterna entre duas etapas: uma de classificação dos pixels e outra de reajuste dos parâmetros das classes. Na etapa de classificação, cada pixel é “visitado” um número determinado de vezes, sendo que em cada “visita” é atribuído um valor de classe para o pixel. Tal classificação leva em consideração, basicamente, a classificação anterior do pixel, a classificação dos pixels vizinhos e os parâmetros atribuídos às classes. Os parâmetros das classes correspondem a duas medidas das regiões da imagem consideradas de textura homogêneas: a média e variância da intensidade dos pixels. Na etapa de reajuste, esses parâmetros são reajustados por equações predefinidas, levando-se em consideração a classificação obtida na etapa antecedente. Não cabe aqui apresentar o embasamento teórico que conduz à origem do algoritmo e, por limitação de espaço, maiores detalhes do mesmo tampouco são apresentados. O artigo de Mary Comer e Eduard Delp [5] apresenta um estudo completo sobre o método.

Uma das vantagens de se utilizar um campo aleatório de Markov para modelagem de textura é a demanda relativamente pequena de parâmetros [7]. O parâmetro principal de um MRF é conhecido como parâmetro de interação espacial, ou simplesmente temperatura, e tem a função de ajustar a influência que a classificação de um pixel exerce na classificação de seus vizinhos.

A segmentação de imagens baseada em textura pode ser supervisionada ou não-supervisionada. Na abordagem supervisionada os parâmetros das classes/texturas são obtidos através de um conjunto de treinamento formado por imagens similares previamente segmentadas por um especialista. Na abordagem não-supervisionada os parâmetros assumem valores iniciais aleatórios e são ajustados durante o processo de segmentação. O método utilizado neste estudo é

não-supervisionado e somente o número de classes previstas para a segmentação precisa ser informado de antemão.

### O vetor de características

Após a imagem ter sido segmentada automaticamente em um número de classes pré-definido, várias medidas são extraídas das regiões de textura homogênea (mesma classe). Assim, se a imagem é segmentada em  $L$  classes, então  $L$  regiões de texturas distintas serão utilizadas para a extração de características. As características que nós consideramos para cada classe são: o centro da massa da região, indicado pela Equação (3); a massa da região (tamanho em número de pixels) dada pela Equação (1); uma medida denominada dispersão (Equação (4)); e a média e a variância da região (Equações (5) e (6)). As três primeiras características (massa, centro de massa e dispersão) são computadas diretamente sobre a imagem segmentada, sendo que a média e variância são medidas da imagem original considerando-se a segmentação obtida. A massa da região (Equação (1)) é o tamanho em pixels da região, ou seja o número dos pixels que foram classificados como sendo de uma determinada classe. Suponha que o resultado da segmentação da imagem  $I$  corresponda à imagem  $S$ , e que o par  $(x,y)$  represente uma posição espacial no espaço onde  $I$  e  $S$  estão definidos. Assim,  $I_{x,y}$  e  $S_{x,y}$  denotam, respectivamente, os valores dos pixels de  $I$  e  $S$  na posição  $(x,y)$ . A massa da região de classe  $c$  ( $m_c$ ) após a segmentação é dada por:

$$m_c = \sum_{x,y} h_c(x, y) \quad (1)$$

onde

$$h_c(x, y) = \begin{cases} 1 & \text{se } S_{x,y} = c \\ 0 & \text{se } S_{x,y} \neq c \end{cases} \quad (2)$$

O centro da massa (centróide) da região de classe  $c$  denotado por  $o_c$ , é dado pelas coordenada  $xo_c$  e pelo  $yo_c$ , onde

$$xo_c = \frac{\sum_{x,y} x \cdot h_c(x, y)}{m_c}, \quad yo_c = \frac{\sum_{x,y} y \cdot h_c(x, y)}{m_c} \quad (3)$$

A característica de dispersão é construída através de uma soma das distâncias. Para cada bloco conexo<sup>1</sup>  $i$  de uma região de classe  $c$ , o centróide  $o_{i,c} = (xo_{i,c}, yo_{i,c})$  é computado e suas distâncias ao centróide  $o_c$  são somadas conforme segue:

$$d_c = \sum_i dist(o_c, o_{i,c}) \quad (4)$$

onde

$dist(o_c, o_{i,c}) = \sqrt{(xo_c - xo_{i,c})^2 + (yo_c - yo_{i,c})^2}$  é a distância Euclidiana entre  $o_c$  (o centro de massa da classe  $c$ ) e  $o_{i,c}$  (o centro de massa bloco conexo  $i$  de classe  $c$ ). Finalmente, conforme mencionado, as características de média e variância da classe  $c$  são computados sobre a imagem original  $I$  considerando a segmentação resultante  $S$ , sendo, respectivamente, denotadas por  $\mu_c$  e  $\sigma_c$ , onde

$$\mu_c = \frac{\sum_{x,y} I_{x,y} \cdot h_c(x, y)}{N} \quad \text{e} \quad (5)$$

$$\sigma_c^2 = \frac{\sum_{x,y} (I_{x,y} - \mu_c) \cdot h_c(x, y)}{N} \quad (6)$$

O vetor de características proposto é construído utilizando-se estas cinco características para cada classe. Visto que o centróide tem dois valores (coordenadas  $x$  e  $y$ ), tem-se um total de seis características por classe. A Figura 2 ilustra uma representação gráfica do vetor de característica para uma imagem segmentada em  $L$  classes. A ordem das características de cada classe no vetor

leva em consideração a característica de média ( $\mu_c$ ), isto é, a classe com a menor média tem suas características na parte mais à esquerda do vetor e, a classe com média maior constitui a parte mais à direita do mesmo. Esta configuração do vetor, aliada a um ajuste adequado das escalas das características, permite o uso da distância Euclidiana para a comparação de vetores de imagens distintas.

## Experimentos

Para teste e avaliação da técnica proposta foi utilizada uma base de imagens médicas de diversas categorias. Devido às limitações de espaço, serão exibidos aqui apenas os resultados obtidos com a base de imagens sumarizada na Tabela1.

**Tabela1** - Configuração da base experimental de imagens médicas

Categoria de imagem	Número de imagens
Angiograma	21
RM <sup>2</sup> Bacia Axial	33
RM Cabeça Axial	50
RM Abdômen Coronal	34
RM Cabeça Sagital	38
RM Espinha Sagital	44

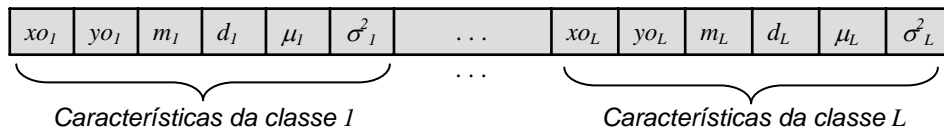
A Figura 3 ilustra o exemplo de uma busca às 24 imagens mais similares a uma imagem de corte axial de cabeça (imagem 1725.jpg), usando o vetor de características proposto. Visto que a segmentação está sendo realizada utilizando-se cinco classes, o vetor final de características possui 30 elementos (características). Como base de comparação, as mesmas consultas foram realizadas utilizando-se como vetores de características, os histogramas de brilho das imagens (o histograma é uma das características mais empregadas nas técnicas de recuperação de imagens por conteúdo). A Figura 4 ilustra os gráficos de *Precision x Recall* para os dois tipos de vetores de características. As curvas, em ambos os gráficos, representam a eficácia das técnicas em recuperar imagens das diversas categorias da base. Para obter essas curvas, cinco imagens de cada categoria foram selecionadas aleatoriamente e as operações de busca foram realizadas utilizando-se ambas as técnicas. Cada curva apresenta uma média das cinco operações de consulta. Observa-se que o vetor de características proposto apresenta uma eficácia maior na recuperação de imagens para todas as categorias da base. Vale também ressaltar que os histogramas

<sup>1</sup> Um bloco conexo consiste em um conjunto de pixels da mesma classe onde cada pixel apresenta uma relação de vizinhança com algum outro pixel do conjunto.

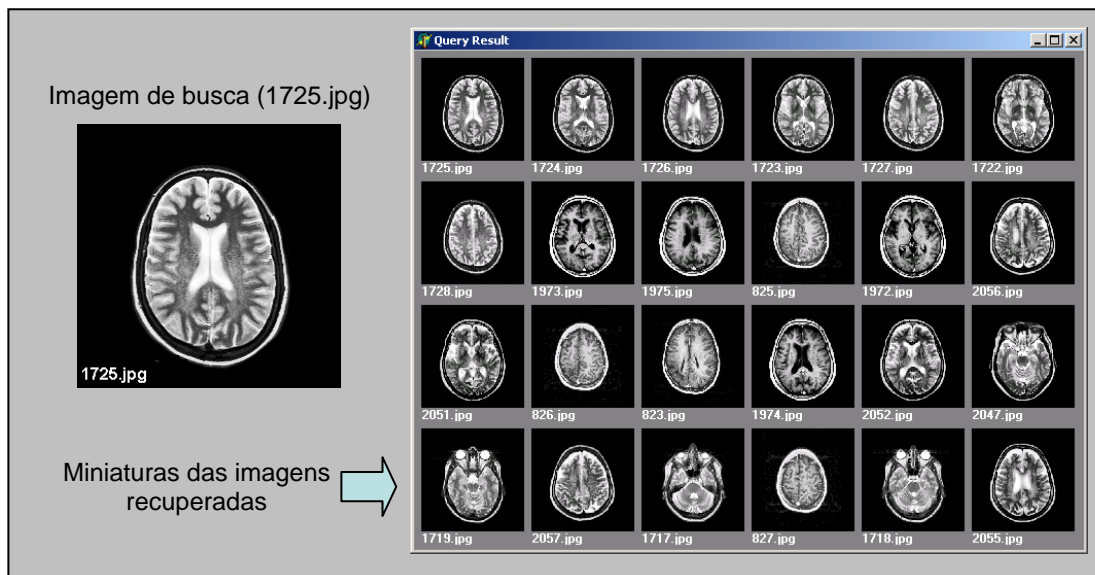
<sup>2</sup> Ressonância Magnética.

são vetores de 256 dimensões enquanto o vetor de características proposto possui apenas 30 elementos. Deste modo, o tempo de resposta às

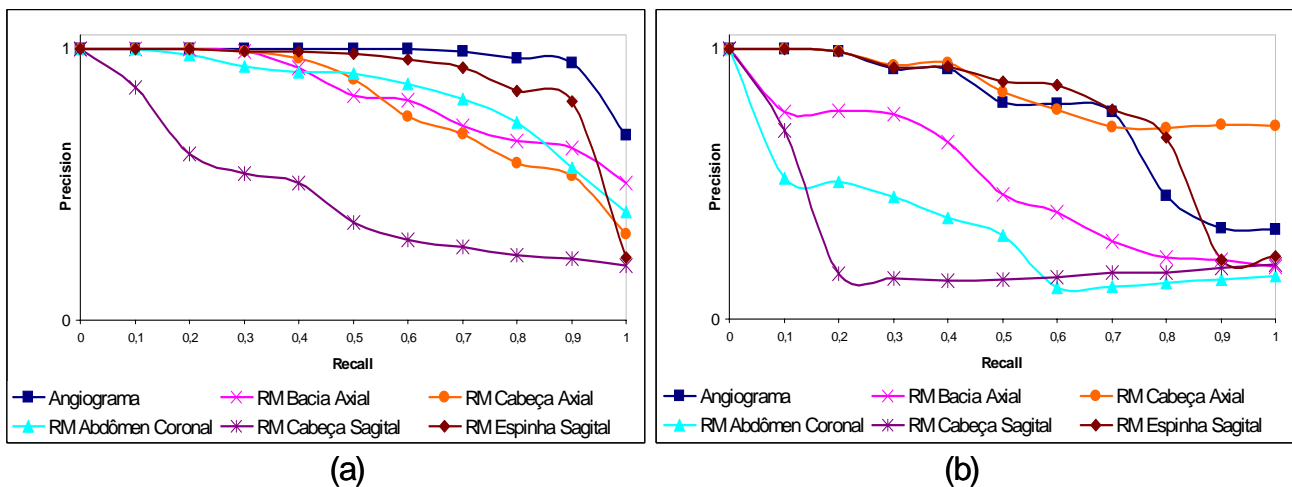
consultas é menor, assim como o espaço de memória necessário para armazenar todos os vetores da base também é menor.



**Figura 2** – Representação gráfica do vetor de características para a segmentação em  $L$  classes



**Figura 3** – Exemplo de uma operação de consulta: recuperação das 24 imagens mais parecidas com uma imagem de corte axial de cabeça (imagem 1725.jpg)



**Figura 4** – Gráficos *Precisão x Recall*. (a) Curvas referentes ao vetor de características proposto; (b) Curvas referentes ao vetor histograma

## Conclusões

A técnica proposta de extração de características utilizando a segmentação de imagens baseada em textura é capaz de lidar adequadamente com o problema de identificação de órgãos e possíveis anomalias presentes em imagens de exames médicos. Um ponto interessante desta abordagem é o uso de um esquema não-supervisionado de segmentação de imagens, que apresenta resultados promissores. As características são extraídas considerando o resultado da segmentação que é obtido automaticamente. O único inconveniente da etapa da segmentação é seu elevado custo computacional. Para segmentar uma imagem de dimensão 300 x 400 pixels, por exemplo, o algoritmo leva de 3 a 5 segundos em uma máquina com processador AMD Athlon XP 2600. Porém, considerando que a geração dos vetores da característica para as imagens da base pode ser realizada "off-line" (em um período de inatividade da máquina), e somente a imagem de consulta precisa ter suas características extraídas durante a operação de consulta, pode-se dizer que este custo não impacta negativamente sua utilização em sistemas CBIR.

## Agradecimentos

Os autores são gratos ao Centro de Ciências das Imagens e Física Médica da Faculdade de Medicina de Ribeirão Preto (USP), que gentilmente forneceu as imagens utilizadas para os experimentos apresentados neste artigo. Esta pesquisa é financiada, em parte, pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) sob os processos de número 02/07318-1 e 03/01769-4; e pelo Conselho Nacional de Pesquisa (CNPq) sob os processos de número 52.1685/98-6, 860.068/00-7 e 35.0852/94-4.

## Referências

- [1] Aslandogan, Y.A., Yu, C.T. (1999), "Techniques and Systems for Image and Video Retrieval", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11, No. 1, Jan/Feb, pp. 56-63.
- [2] Baeza-Yates, R., Ribeiro-Neto (1999), B.A., *Modern Information Retrieval*, Wokingham, UK: Addison-Wesley.
- [3] Brunelli, R., Mich, O. (1999), "On the Use of Histograms for Image Retrieval", *Proc. IEEE Intl. Conf. on Multimedia Computing and*

*Systems (ICMCS)*, Florence, Italy, June 7-11, pp. 143-147.

- [4] Cao, X., Huang, H.K. (2000), "Current Status and Future Advances of Digital Radiography and PACS", *IEEE Engineering in Medicine and Biology Magazine*, Vol. 9, No. 5, Sept-Oct, pp. 80-88.
- [5] Comer, M.L., Delp E.J. (2000), "The EM/MPM Algorithm for Segmentation of Textured Images: Analysis and Further Experimental Results", *IEEE Transactions on Image Processing*, Vol. 9, No. 10, October, pp. 731-744.
- [6] Huang, P.W., Dai, S.K. (2003), "Image retrieval by texture similarity", *Pattern Recognition*, Vol. 36, No. 3, March, pp. 665-679.
- [7] Rosholm (1997), A., *Statistical Methods for Segmentation and Classification of Images*, Ph.D. thesis, Technical University of Denmark.
- [8] Safar M., Shahabi, C., Sun, X. (1999), "Image Retrieval by Shape: A Comparative Study", *Proc. IEEE Int. Conf. On Multimedia and Expo*, New York.
- [9] Sheikholeslami G., Chang, W., Zhang, A. (2002), "SemQuery: Semantic Clustering and Querying on Heterogeneous Features for Visual Data", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 14, No. 5, September/October, pp. 988-1002.
- [10] Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain R. (2000), "Content-Based Image Retrieval at the End of the Early Years", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, December.
- [11] Zhang, D., Lu, G. (2001), "Content-Based Shape Retrieval Using Different Shape Descriptors: A Comparative Study", *Proc. IEEE Conference on Multimedia and Expo (ICME'01)*, Tokyo, Japan, August, pp. 317-320.

## Contatos

André G. R. Balan<sup>1</sup> – [agrbalan@icmc.usp.br](mailto:agrbalan@icmc.usp.br)  
Agma J. M. Traina<sup>2</sup> – [agma@icmc.usp.br](mailto:agma@icmc.usp.br)  
Caetano Traina Jr.<sup>3</sup> – [caetano@icmc.usp.br](mailto:caetano@icmc.usp.br)  
Paulo Azevedo-Marques<sup>4</sup> – [pmarques@fmrp.usp.br](mailto:pmarques@fmrp.usp.br)  
<sup>1,2,3</sup>Instituto de Ciências Matemáticas e de Computação (ICMC – USP) – Av. do Trabalhador São-Carlense, 400 – Centro – Cx. Postal 668 – São Carlos – SP – Brasil, CEP 13560-970.  
<sup>4</sup>Faculdade de Medicina de Ribeirão Preto (FMRP – USP) – Av. Bandeirantes, 3900 – Ribeirão Preto – SP – Brasil, CEP 14049-900.