

Reconhecimento de padrões de texturas em imagens digitais usando uma rede neural artificial híbrida

Wonder Alexandre Luz Alves, Sidnei
Alves de Araújo, André Felipe Henriques
Librantz
Departamento de Ciências Exatas. Uninove. São Paulo – SP
[Brasil]
wonder@woniz.com

Neste trabalho, apresenta-se um mecanismo de indexação de imagens, com base em características texturais, que utiliza redes neurais artificiais (RNAs). Os atributos que descrevem as texturas e que são utilizados para classificá-las advêm dos descritores de Haralick, que se fundamentam em matrizes de co-ocorrência. Uma RNA híbrida é utilizada para reconhecer os diferentes padrões de texturas. Na primeira fase, é realizado um agrupamento inicial dos padrões por um modelo não-supervisionado (mapas auto-organizáveis) e, na segunda, utiliza-se o modelo supervisionado (quantização vetorial por aprendizagem) para melhorar a segmentação das classes de padrões previamente agrupados pelo modelo não-supervisionado.

Palavras-chave: Descritores de Haralick. Padrões de textura. Redes Neurais artificiais.



1 Introdução

Desde os anos 1970, diversas técnicas de recuperação de imagens vêm sendo desenvolvidas, principalmente, pelas áreas de gerenciamento de banco de dados (indexação textual) e visão computacional (RUI; HUANG; CHANG, 1997). No entanto, o aumento considerável de bancos de dados de imagens, devido às inovações tecnológicas, exigiu a utilização de algum tipo de mecanismo de indexação para posterior recuperação das imagens armazenadas. Executar manualmente esse processo é, entretanto, uma tarefa sensível a interpretações subjetivas, é laboriosa e sujeita a erros (ANTANI; KASTURI; JAIN, 2002; SEBE; LEW, 2001; BRUNELLI; MICH, 2000; RUI; HUANG; CHANG, 1997; SMITH; CHANG, 1996). Métodos automáticos de indexação e de recuperação baseados em algum tipo de característica (cor, forma, textura ou região de uma imagem) são interessantes nesse contexto, uma vez que podem reduzir a intervenção humana, possibilitando, dessa forma, maior eficácia e, em muitos casos, uma significativa diminuição da margem de erro (LIU; ZHOU, 2003).

Os métodos de classificação de imagens têm sido largamente explorados nos últimos anos, dada a sua aplicabilidade em sistemas automáticos de indexação e recuperação de imagens digitais, principalmente na área médica. É o caso dos sistemas denominados *picture archiving and communication system* (Pacs), que permitem integrar, num só sistema, informações textuais do paciente, além de informações gráficas provenientes de exames como raios X, tomografia e ultra-sonografia (FELIPE; TRAINA, 2002).

Em sistemas automáticos de indexação de imagens, elas podem ser recuperadas de um banco de dados, por meio de algum elemento gráfico

como cor, textura e forma. Entre esses elementos, destacam-se os atributos de textura, que são muito utilizados na tarefa de classificação de imagens digitais (AKSOY; HARALICK, 1998).

Felipe e Traina (2002) desenvolveram um *software* para indexação e recuperação de imagens médicas por meio de suas texturas, que foram caracterizadas pelos descritores de Haralick (1979). Em Martins, Guimarães e Fonseca (2002), foi proposto um método baseado em uma rede neural artificial (RNA) híbrida, com modelos não-supervisionados (mapas auto-organizáveis, em inglês *self-organizing maps* [SOM]) e supervisionados (quantização vetorial por aprendizagem, em inglês *learning vector quantization* [LVQ]), para classificar imagens de sensoriamento remoto, com base em atributos de textura extraídos com uso de filtros de Gabor. Shiba e colaboradores (2005) propuseram um método fundamentado em árvores de decisão para classificação de imagens de sensoriamento remoto descritas pelas três bandas do sistema de cores RBG (do inglês *red, green, blue*; em português, vermelho, verde, azul). Alves e Araújo (2006) fizeram uma análise comparativa do desempenho de dois diferentes modelos de RNA, na classificação de imagens digitais, a partir de suas texturas, caracterizadas por descritores estatísticos de primeira e de segunda ordem.

Neste trabalho, aborda-se o uso de uma RNA híbrida (modelo de RNA supervisionado com um modelo de rede não-supervisionado) no reconhecimento de imagens com base nas suas características texturais, empregando-se os descritores de Haralick. Os resultados obtidos pelo modelo de RNA híbrido são comparados aos do modelo que emprega apenas o algoritmo de treinamento não-supervisionado, no intuito de averiguar a validade do modelo híbrido proposto.

2 Texturas em imagens digitais

Conforme enumerações a seguir.

2.1 Imagens digitais

Uma imagem digital pode ser entendida como uma matriz em que os índices de linha e coluna identificam um ponto na imagem, e o valor do elemento da matriz, o nível de cinza naquele ponto. Cada elemento dessa matriz é chamado de *pixel* (GONZALEZ; WOODS, 2002). Dessa forma, uma imagem digital é normalmente definida como uma função bidimensional $f(i, j)$, com $i, j \in \mathbb{Z}^2$. Uma imagem em níveis de cinza G^y pode ser descrita como uma função (Equação 1), cujo contradomínio representa diferentes níveis de cinza, por exemplo o intervalo real de 0 a 1 que, geralmente, é discretizado computacionalmente, resultando em um intervalo inteiro de 0 a 255.

$$G^y: \mathbb{Z}^2 \rightarrow [0, \dots, 255] \quad (1)$$

2.2 Texturas

Apesar de não haver consenso sobre a definição de textura, pode-se referenciá-la como um conjunto de determinados padrões existentes nas superfícies físicas, perceptíveis ao olho humano, e que trazem grande quantidade de informações sobre a sua natureza, tais como suavidade, rugosidade e regularidade. Na descrição de padrões de texturas em imagens, normalmente são utilizadas três abordagens principais: a estatística, a estrutural e a espectral (GONZALEZ; WOODS, 2002). Neste trabalho, considerou-se apenas a abordagem estatística para descrição das texturas das imagens.

2.3 Descrição de texturas por matriz de co-ocorrência

A extração de características usando abordagem estatística propicia a descrição de imagens por meio de regras estatísticas que governam tanto a distribuição quanto a relação entre os diferentes níveis de cinza. Inúmeras medidas podem ser extraídas da distribuição dos níveis de cinza em imagens, entre as quais o cálculo de média, variância, desvio-padrão, energia ou entropia. A abordagem estatística pode ser dividida em duas vertentes: a de primeira ordem, na qual se extraem características a partir de histogramas de primeira ordem, e a de segunda ordem, que leva em conta o posicionamento relativo da ocorrência dos níveis de cinza.

A principal diferença entre as abordagens é que, dadas duas imagens que contenham a mesma quantidade de níveis de cinza, diferenciando-se apenas em termos de posicionamento, na abordagem de primeira ordem, apresenta-se um mesmo histograma, enquanto na de segunda ordem, os histogramas serão diferentes.

Haralick (1979) propõe uma metodologia para descrição de texturas com base em estatística de segunda ordem, em que são definidas as características provenientes do cálculo de matrizes denominadas “matrizes de co-ocorrência”, que consistem de uma contagem de quantas combinações diferentes de níveis de cinza ocorrem em uma imagem, em uma determinada direção. Para obtenção de tais matrizes, considera-se a variação da distância e direção (d, θ) entre pixels vizinhos. Normalmente, são utilizados quatro direcionamentos: 0° , 45° , 90° e 135° , conforme ilustra a Figura 1.

As matrizes de co-ocorrência formam a base para elaboração de diversas medidas estatísticas conhecidas como descritores de Haralick. Utilizaram-se, neste trabalho, oito descritores: segundo momento angular (SMA), contraste (Cont), entro-

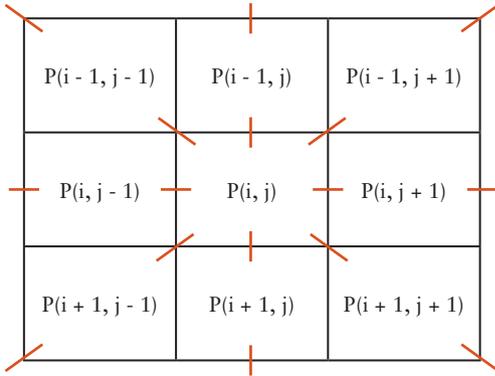


Ilustração 1: Janela das variações angulares (θ) utilizadas no cálculo das matrizes de co-ocorrência, considerando $d = 1$

Fonte: Os autores.

pia (Ent), momento da diferença inversa (MDI), correlação (Cor), variância (Var), dissimilaridade (Dis) e média (Med). Esses descritores são dados, respectivamente, pelas equações (2), (3), (4), (5), (6), (7), (8) e (9) a seguir.

Para cada *pixel* $P(i, j)$ processado na imagem, há uma janela em torno dele, com distância $d = 1$ nas quatro direções θ . N_g denota a quantidade de níveis de cinza em uma imagem.

$$SMA(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [P(i, j)]^2 \quad (2)$$

O segundo momento angular fornece a medida da homogeneidade local dos níveis de cinza em uma imagem.

$$Cont(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 P(i, j) \quad (3)$$

O contraste reflete a quantidade de variação local de níveis de cinza em uma imagem.

$$ENT(d, \theta) = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j) \log_2 [P(i, j)] \quad (4)$$

A entropia fornece o grau de dispersão de níveis de cinza de uma imagem.

$$MDI(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i - j)^2} P(i, j) \quad (5)$$

O momento da diferença inversa denota a homogeneidade da matriz de co-ocorrência.

$$COR(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{ijP(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (6)$$

A correlação mede a dependência linear dos níveis de cinza nas combinações dos pares de *pixels* em uma determinada direção. μ_x e μ_y denotam as médias nas direções x e y , enquanto σ_x e σ_y representam os desvios-padrão.

$$VAR(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 \cdot P(i, j) \quad (7)$$

A variância diz respeito à heterogeneidade da matriz de co-ocorrência em forma de desvio dos valores P da matriz.

$$DIS(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j| \cdot P(i, j) \quad (8)$$

A dissimilaridade mede o desvio dos valores da combinação de pares de *pixels* diagonais, em que apenas a contribuição do desvio é considerada.

$$MED(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i \cdot P(i, j) \quad (9)$$

Média da matriz de co-ocorrência.

3 RNA

As RNAs têm sido amplamente utilizadas em aplicações relativas tanto à classificação quanto à recuperação de imagens. O uso dos modelos neurais para essas atividades está diretamente relacionado às suas principais características: a capacidade de aprender e de generalizar a informação aprendida. Outra característica importante dessas redes é a capacidade de auto-organização, em que a rede atribui um padrão de entrada a uma classe entre um conjunto de classes conheci-

das. O procedimento empregado no aprendizado de uma RNA é denominado de algoritmo de treinamento, e sua função principal é a modificação dos pesos sinápticos, visando a atingir o objetivo (HAYKIN, 2001).

3.1 SOM

Os SOMs, também conhecidos como redes de Kohonen (1990), são modelos de RNA não-supervisionados, ou seja, o ajuste dos pesos sinápticos é feito automaticamente sem a necessidade de um “professor”. Trata-se de um mapa topográfico artificial que aprende, por meio da auto-organização inspirada na neurobiologia, qual dos neurônios da camada de saída representa uma classe de padrões de entrada, de acordo com suas características ou atributos (HAYKIN, 2001).

A idéia básica da rede é a seguinte: cada vetor de entrada, que representa um padrão, pode ser visto como um estímulo aplicado à rede e corresponde a um só neurônio na camada de saída, também chamado de neurônio vencedor, ou seja, a rede funciona como mapeador de um conjunto de atributos de entrada, representando-os por neurônios, que são ativados na camada de saída (KOHONEN, 1990).

3.2 LVQ

O algoritmo SOM proporciona um agrupamento inicial dos padrões de entrada; no entanto, é comum haver problemas de classificação nas fronteiras da decisão entre as classes. A LVQ, de natureza supervisionada, possibilita melhorar a divisão do espaço dos padrões previamente agrupados. O algoritmo LVQ usa a informação sobre as classes de padrões para melhorar a qualidade das regiões de decisão do classificador. O funcionamento do LVQ pode ser descrito da seguinte forma: toma-se um vetor de entrada x aleatoriamente do espaço de entrada. Se houver concordância entre os rótulos de classe do vetor de entrada

x e de um vetor de Voronoi w , move-se o vetor de Voronoi w em direção ao vetor de entrada x . Caso haja discordância entre os rótulos de classe do vetor de entrada x e do vetor de Voronoi w , o vetor de Voronoi w é afastado do vetor de entrada x (HAYKIN, 2001).

4 Classificação de texturas pela aprendizagem

O problema de classificação de texturas pela aprendizagem envolveu dois conjuntos de padrões baseados nos descritores de Haralick, um para o treinamento da rede, denotado por $P = \{p_1, p_2, \dots, p_n\}$, e outro para a fase de testes da rede, indicado por $Q = \{q_1, q_2, \dots, q_n\}$. Cada padrão $p_i \in P$ e $q_i \in Q$ é representado por um vetor de oito elementos (os oito descritores estatísticos).

O treinamento da RNA híbrida inicia-se rotulando o conjunto P por meio do algoritmo não-supervisionado SOM, que proporciona um agrupamento inicial dos padrões de treinamento. Entretanto, nas regiões de separação entre as classes, pode haver problemas de classificação (sobreposição de classes). A aprendizagem por LVQ, de natureza supervisionada, é, portanto utilizada para melhorar a divisão do espaço das classes estabelecidas pelo algoritmo SOM.

5 Resultados experimentais

Após a fase de treinamento do modelo híbrido, procedeu-se à etapa de testes para verificação de seu desempenho. Para os conjuntos de treinamento (P) e de teste (Q), foram utilizadas imagens de tecidos biológicos extraídas de vários autores (KUHNEL, 1991; SCHNEIDER, 1998; WOLF; SCARBROUGH, 2005). Selecionaram-se 50 imagens para compor o conjunto P e outras 50



diferentes para formar o conjunto Q, totalizando 100 imagens. Tanto no conjunto de treinamento quanto no de teste, as imagens foram divididas em cinco classes, cada uma contendo dez imagens.

Para a fase de rotulação dos padrões, foi proposta uma arquitetura para rede SOM com oito neurônios na camada de entrada e um mapa de saída com 36 neurônios (matriz quadrada 6 x 6). Para averiguar o desempenho da rede, o algoritmo de treinamento foi executado por 700 épocas¹, utilizando uma taxa de aprendizagem de 0,005. Após o agrupamento, o mapa de saída da rede é apresentado em uma janela (Figura 2), em que as imagens com texturas similares estão organizadas em uma determinada região do mapa.

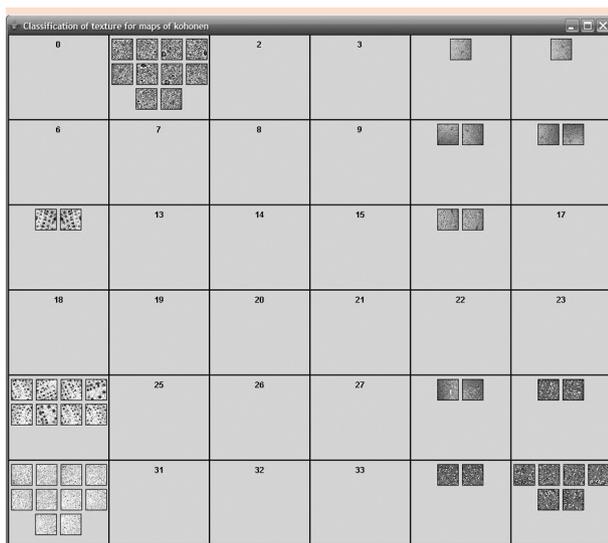
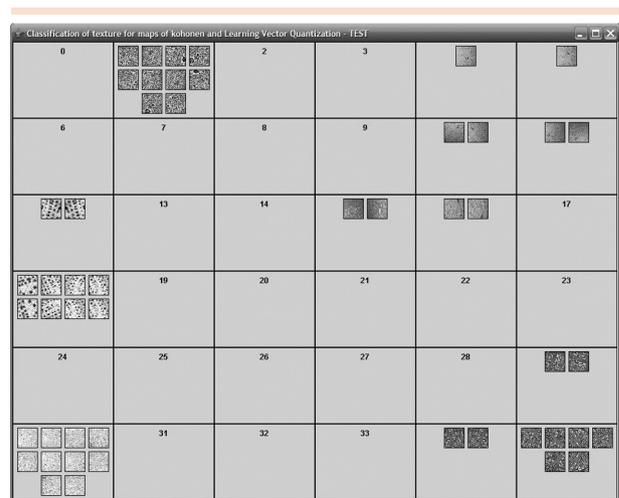


Figura 2: Mapa de saída após treinamento com algoritmo SOM

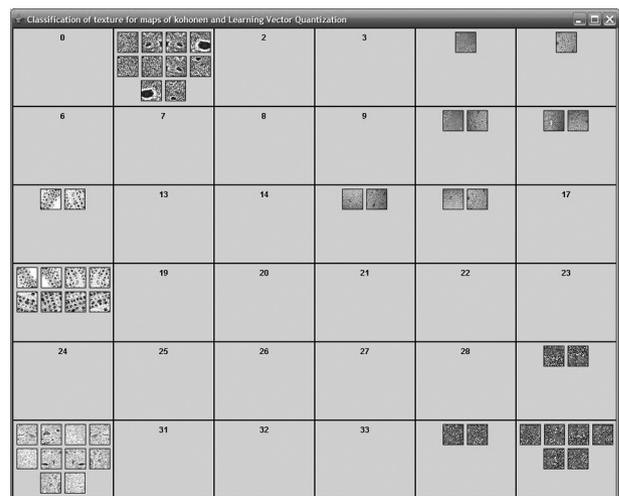
Fonte: Os autores.

Observando o mapa de saída (Figura 2), percebe-se que a rede treinada com algoritmo SOM conseguiu agrupar todas as imagens do conjunto de treinamento. Entretanto, em alguns pontos houve imagens classificadas em regiões a que não pertenciam sua classe de padrões. Dessa forma, para obter um melhor desempenho no reconhecimento, foi preciso melhorar a divisão do espaço dos padrões agrupados pelo algoritmo SOM por

meio do algoritmo supervisionado LVQ. Para isso, a partir dos dados de saída do algoritmo SOM, foi realizado o treinamento com LVQ, utilizando a mesma taxa de aprendizagem e a mesma quantidade de épocas adotadas no treinamento de SOM. A Figura 3a ilustra o mapa de saída após o treinamento com LVQ, usando o conjunto P, enquanto a Figura 3b mostra o mapa de saída com as imagens do conjunto Q (imagens-teste). A Tabela 1 traz os resultados dos testes efetuados com a rede treinada apenas com o algoritmo SOM. Já a Tabela 2



(a)



(b)

Figura 3: Mapas de saída da rede híbrida proposta

Obs.: a) Mapa de saída da rede após treinamento com LVQ com conjunto P; b) Mapa de saída da rede treinada com LVQ após fase de testes com conjunto Q.

Fonte: Os autores.

demonstra os resultados relativos aos testes com a rede híbrida treinada e proposta.

Tabela 1: Resultados dos experimentos com o conjunto de testes Q usando a rede treinada apenas com SOM

Classe	Quantidade de acertos	Quantidade de erros	Acertos (%)
1	3	7	30
2	10	0	100
3	10	0	100
4	6	4	60
5	7	3	70
Média de acertos			72

Fonte: Os autores.

Tabela 2: Resultados dos experimentos com o conjunto de testes Q usando a rede híbrida treinada com SOM + LVQ

Classe	Quantidade de acertos	Quantidade de erros	Acertos (%)
1	10	0	100
2	10	0	100
3	10	0	100
4	10	0	100
5	10	0	100
Média de acertos			100

Fonte: Os autores.

A partir dos resultados apresentados nas Tabelas 1 e 2, verificou-se que o emprego do algoritmo híbrido utilizado no agrupamento dos padrões de treinamento proporcionou uma eficiência média de 100% no reconhecimento das imagens, tendo como base a descrição de suas texturas, enquanto o resultado obtido com o uso do algoritmo SOM sem o LVQ mostrou uma eficiência média de 72%. Isso comprova que a aplicação do algoritmo LVQ ao SOM permitiu uma melhora significativa da classificação dos padrões.

6 Considerações finais

Neste trabalho, foi demonstrada a aplicação de um modelo híbrido de RNA (SOM com LVQ)

no reconhecimento de imagens caracterizadas pelos descritores de Haralick. Os resultados obtidos na classificação de imagens com a aplicação dessa técnica mostraram-se superiores na comparação com aqueles obtidos a partir do uso de um único modelo de RNA não-supervisionado (SOM). Isso sinaliza que o método proposto pode ser uma alternativa promissora para aplicações que envolvam tanto a indexação quanto a recuperação automática de imagens digitais, principalmente em bancos de imagens médicas.

7 Agradecimentos

Os autores agradecem à Uninove pelo apoio financeiro.

Recognition texture patterns in digital images using an hybrid artificial network

This work presents a mechanism of images indexation based on textural features applying artificial neural network. Attributes originated from Haralick descriptors based on co-occurrence matrix, are applied in order to classify the textures. An hybrid implementation of neural network is used in the recognition of different texture patterns, which firstly propitiates a initial gathering of them by using an unsupervised model (self-organization map [SOM]) and secondly by using a supervised model (learning vector quantization [LVQ]) aiming to enhance the patterns classes segmentation previously gathered from the application of the unsupervised model.

Key words: Haralick descriptors. Texture patterns. Artificial neural network.

Notas

- 1 N. Ed.: cada época significa um ciclo de treinamento.



Referências

- AKSOY, S.; HARALICK, R. M. Textural features for image database retrieval. In: IEEE WORKSHOP ON CONTENT-BASED ACCESS OF IMAGE AND VIDEO LIBRARIES, 1998, Washington. *Proceedings...* Washington: IEEE, 1998.
- ALVES, W. A. L.; ARAÚJO, S. A. de. Avaliação da utilização de dois diferentes modelos de redes neurais na classificação de imagens baseada em características texturais. *Exacta*, São Paulo, v. 4, n. 1, p. 77-86, 2006.
- ANTANI, S.; KASTURI, R.; JAIN, R. A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. *Pattern Recognition*, Amsterdã, v. 4, n. 35, p. 945-965, 2002.
- BRUNELLI, R.; MICH, O. Image retrieval by examples. *IEEE Transactions on Multimedia*, Nova Jersey, v. 2, n. 3, p. 164-171, 2000.
- FELIPE, J. C.; TRAINA, A. J. M. Utilizando características de textura para identificação de tecidos em imagens médicas. In: WORKSHOP DE INFORMÁTICA MÉDICA, 2., 2002, Gramado. *Anais...* Gramado: WIM, 2002.
- GONZALEZ, R. C.; WOODS, R. E. *Digital image processing*. 2. ed. Nova Jersey: Prentice Hall, 2002.
- HARALICK, R. M. Statistical and structural approaches to texture. *Proceedings of the IEEE*, v. 67, n. 5, p. 786-804, 1979.
- HAYKIN, S. *Redes neurais: princípio e prática*. 2. ed. Porto Alegre: Bookman, 2001.
- KOHONEN, T. The self-organizing map. *Proceedings of the IEEE*, v. 78, p. 1.464-1.480, 1990.
- KUHNEL, W. *Atlas de citologia, histologia e anatomia microscópica: teoria e prática*. 7. ed. Rio de Janeiro: Guanabara Koogan, 1991.
- LIU, Y.; ZHOU, X. A simple texture descriptor for texture retrieval. In: INTERNATIONAL CONFERENCE ON COMMUNICATION TECHNOLOGY, 2003, Beijing. *Proceedings...* Beijing: ICTT, 2003. p. 1.662-1.665.
- MARTINS, M. P.; GUIMARÃES, L. N. F.; FONSECA, L. M. G. Classificador de texturas por redes neurais. In: CONGRESSO BRASILEIRO DE COMPUTAÇÃO, 2., 2002, Itajaí. *Anais...* Itajaí: Univali, 2002. v. 1, p. 1-12.
- RUI, Y.; HUANG, T. S.; CHANG, S-F. Image retrieval: past, present and future. In: INTERNATIONAL SYMPOSIUM ON MULTIMEDIA INFORMATION PROCESSING, 1997, Taiwan. *Proceedings...* Taiwan, 1997.
- SCHNEIDER, M. L. *Atlas de diagnóstico diferencial em citologia ginecológica*. 1. ed. Rio de Janeiro: Revinter, 1998.
- SEBE, N.; LEW, M. S. Color-Based Retrieval. *Pattern Recognition Letters*, Amsterdã, v. 22, n. 2, p. 223-230, 2001.
- SHIBA, M. H. et al. Classificação de imagens de sensoriamento remoto pela aprendizagem por árvore de decisão: uma AVALIAÇÃO de desempenho. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12., 2005, Goiânia. *Anais...* Goiânia: Inpe, 2005. p. 4.319-4.326.
- SMITH, J. R.; CHANG, S-F. Tools and techniques for color image retrieval. In: SYMPOSIUM ON ELECTRONIC IMAGING: SCIENCE AND TECHNOLOGY – STORAGE & RETRIEVAL FOR IMAGE AND VIDEO DATABASES, 4., 1996, San Jose. *Proceedings...* San Jose: IS&T/Spie, 1996. v. 2.670, p. 426-437.
- WOLF, M.; SCARBROUGH, M. The JayDoc HistoWeb. *Site The University of Kansas*, 2005. Disponível em: <<http://www.kumc.edu/instruction/medicine/anatomy/histoweb/index.htm>>. Acesso em: 5 set. 2006.

Recebido em 5 set. 2006 / aprovado em 5 nov. 2006

Para referenciar este texto

ALVES, W. A. L.; ARAÚJO, S. A. de; LIBRANTZ, A. F. H. Reconhecimento de padrões de texturas em imagens digitais usando uma rede neural artificial híbrida. *Exacta*, São Paulo, v. 4, n. 2, p. 325-332, jul./dez. 2006.