

Reconhecimento de Padrões de Texturas em Imagens Digitais Usando uma Rede Neural Artificial Híbrida

Wonder Alexandre Luz Alves, Sidnei Alves de Araújo, André Felipe H. Librantz

Departamento de Ciências Exatas – Centro Universitário Nove de Julho (UNINOVE)
São Paulo – SP – Brasil

wonder@woniz.com, {saraujo,librantz}@uninove.br

***Abstract.** This work presents a mechanism of images indexation based on textural features applying artificial neural network. Attributes originated from Haralick descriptors based on co-ocurrence matrix, are utilized 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 applying an unsupervised model (self-organization map) and secondly by using a supervised model (Learning Vector Quantization) aiming to enhance the patterns classes segmentation previously gathered since the application of the unsupervised model.*

***Resumo.** Este trabalho apresenta um mecanismo de indexação de imagens baseado em características texturais utilizando redes neurais artificiais. Os atributos que descrevem as texturas e que são utilizados para classificá-las advêm dos descritores de Haralick os quais são baseados em matrizes de co-ocorrência. Uma rede neural híbrida é utilizada para reconhecer os diferentes padrões de texturas. Primeiramente, é realizado um agrupamento inicial dos padrões por um modelo não supervisionado (Mapas Auto-organizáveis) e numa segunda fase, utiliza-se o modelo supervisionado (Quantização Vetorial por Aprendizagem) para melhorar a segmentação das classes de padrões previamente agrupados pelo modelo não supervisionado.*

1. Introdução

Desde os anos 70 diversas técnicas de recuperação de imagens vêm sendo desenvolvidas, principalmente pelas áreas de gerenciamento de banco de dados (indexação textual) e visão computacional (Rui et al., 1997). No entanto, o aumento considerável de bancos de dados de imagens, devido principalmente às inovações tecnológicas, exigiu algum tipo de mecanismo de indexação para posterior recuperação das imagens armazenadas. Entretanto, executar esta indexação manualmente é uma tarefa sensível a interpretações subjetivas, é laboriosa e sujeita a erros (Antani et al., 2002; Sebe & Lew, 2001; Brunelli & Mich, 2000; Rui et al., 1997; Smith & Chang, 1996). Métodos automáticos de indexação e recuperação baseados em algum tipo de característica (cor, forma, textura ou região de uma imagem) são interessantes neste contexto uma vez que podem reduzir a intervenção humana possibilitando, desta forma, maior eficácia e, em muitos casos, uma significativa diminuição da margem de erro (Liu & Zhou, 2003).

Os métodos de classificação de imagens tem sido largamente explorados nos últimos anos, dada a sua aplicabilidade em sistemas automáticos de indexação e recuperação de imagens digitais, principalmente na área médica, em sistemas denominados PACS (*Picture Archiving and Communication System*) que permitem integrar num só sistema, informações textuais do paciente além de informações gráficas, provenientes de exames como raio-X, tomografia, ultra-sonografia, entre outras (Felipe & Traina, 2002).

Em sistemas automáticos de indexação de imagens, as mesmas podem ser recuperadas a partir de um banco de dados por meio de algum elemento gráfico como cor, textura, forma e outros. Dentre estes elementos, os atributos de textura têm sido uma das principais características utilizadas na tarefa de classificação de imagens digitais (Aksoy & Haralick, 1998). Na literatura, diferentes métodos de classificação de imagens, principalmente utilizando atributos de texturas, podem ser encontrados.

Felipe & Traina (2002) desenvolveram um software para indexação e recuperação de imagens médicas por meio de suas texturas, as quais foram caracterizadas pelos descritores de Haralick (Haralick, 1979). Em (Martins et al., 2002) foi proposto um método baseado em uma rede neural híbrida SOM (*Self-Organizing Maps*) e LVQ (*Learning Vector Quantization*) para classificação de imagens de sensoriamento remoto a partir dos atributos de textura extraídos com uso de filtros de Gabor. Shiba et al. (2005) propuseram um método baseado em árvores de decisão para classificação de imagens de sensoriamento remoto descritas pelas três bandas do sistema de cores RGB. Alves & Araújo (2005) fizeram uma análise comparativa do desempenho de dois diferentes modelos de redes neurais, 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 rede neural híbrida (modelo de rede neural supervisionado com um modelo de rede não supervisionado) no reconhecimento de imagens baseado nas suas características texturais empregando-se os descritores de Haralick. Os resultados obtidos pelo modelo de rede neural híbrido são comparados aos resultados 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

2.1. Imagens Digitais

Uma imagem digital pode ser entendida como uma matriz onde os índices de linha e coluna identificam um ponto na imagem e o valor do elemento da matriz identifica 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 $i, j \in Z^2$. Uma imagem em níveis de cinza G^y pode ser descrita como uma função (Equação 1) cujo contra-domínio representa diferentes níveis de cinza, por exemplo, o intervalo real de 0 a 1. Esse intervalo é geralmente discretizado computacionalmente resultando em um intervalo inteiro de 0 a 255.

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

2.2. Texturas

Apesar de não haver consenso sobre a definição de textura, pode-se referenciá-la como sendo um conjunto de determinados padrões existentes nas superfícies físicas, perceptíveis ao olho humano e que trazem grande quantidade de informações sobre a natureza da superfície, como suavidade, rugosidade e regularidade. Para 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 através de regras estatísticas que governam a distribuição e a relação entre os diferentes níveis de cinza. Inúmeras medidas podem ser extraídas da distribuição dos níveis de cinza em imagens, entre as quais podemos citar: o cálculo de média, variância, desvio padrão, energia ou entropia, entre outras. A abordagem estatística pode ser dividida em duas importantes vertentes: a abordagem de primeira ordem, na qual se extrai características a partir de histogramas de primeira ordem e a abordagem de segunda ordem que leva em conta o posicionamento relativo da ocorrência dos níveis de cinza.

A principal diferença entre as duas 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, apresentarão um mesmo histograma enquanto na abordagem de segunda ordem, os histogramas serão diferentes.

Em (Haralick, 1979), foi proposta uma metodologia para descrição de texturas baseada em estatística de segunda ordem, onde são definidas 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.

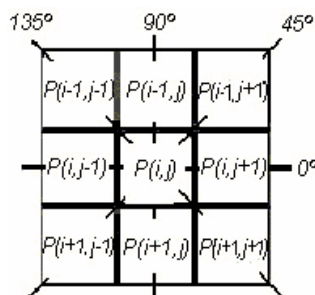


FIGURA 1 – Janela das variações angulares (θ) utilizadas no cálculo das matrizes de co-ocorrência, considerando $d=1$.

As matrizes de co-ocorrência formam a base para elaboração de diversas medidas estatísticas conhecidas como descritores de Haralick. Utilizou-se nesse trabalho, 8 descritores: segundo momento angular (SMA), contraste (Cont), entropia (Ent), momento da diferença inversa (MDI), correlação (Cor), variância (Var), dissimilaridade

(Dis) e média (Med). As equações desses descritores são dadas, respectivamente, pelas equações (2), (3), (4), (5), (6), (7), (8) e (9) a seguir.

Para cada pixel $P(i, j)$ processado na imagem, é considerado 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.

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

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

$$Ent(d, \theta) = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j) \log_2[P(i, j)] \quad (4) \quad MDI(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1+(i-j)^2} P(i, j) \quad (5)$$

A entropia fornece o grau de dispersão de níveis de cinza de uma imagem. O momento da diferença inversa denota a homogeneidade da matriz de co-ocorrência.

$$Cor(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{ijP(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (6) \quad Var(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-j)^2 * P(i, j) \quad (7)$$

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.

$$Dis(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |(i-j)| * P(i, j) \quad (8) \quad Med(d, \theta) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i * P(i, j) \quad (9)$$

A dissimilaridade mede o desvio dos valores da combinação de pares de pixels diagonais, onde apenas a contribuição do desvio é considerada. Média da matriz de co-ocorrência.

3. Redes Neurais Artificiais

As redes neurais artificiais têm sido amplamente utilizadas em aplicações relativas à classificação e recuperação de imagens. O uso dos modelos neurais para estas atividades está diretamente relacionado à suas principais características: a capacidade de aprender e de generalizar a informação aprendida. Outra característica importante das redes neurais é a capacidade de auto-organização, em que a rede atribui um padrão de entrada a uma classe dentre um conjunto de classes conhecidas. O procedimento empregado no aprendizado de uma rede neural é denominado de algoritmo de treinamento e sua função principal é a modificação dos pesos sinápticos visando atingir o objetivo (Haykin, 2001).

3.1. Mapas Auto-Organizáveis

Os Mapas Auto-Organizáveis também conhecidos como redes de Kohonen são modelos de redes neurais 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 através da auto-organização inspirada na neurobiologia, no qual os neurônios da camada de saída representam 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, que também é chamado de neurônio vencedor. Ou seja, a rede funciona como um mapeador de um conjunto de atributos de entrada representando-os por neurônios que são ativados na camada de saída (Kohonen, 1990).

3.2. Quantização Vetorial por Aprendizagem

O algoritmo SOM proporciona um agrupamento inicial dos padrões de entradas, todavia, é comum haver problemas de classificação nas fronteiras da decisão entre as classes. A quantização vetorial por aprendizagem (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 maneira: 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 denotado 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 8 elementos (os oito descritores estatísticos).

O treinamento da rede neural híbrida inicia-se rotulando o conjunto P através do algoritmo não supervisionado SOM, o qual 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 então 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 a fase de testes para verificar o desempenho do mesmo. Para os conjuntos de treinamento (P) e de teste (Q) foram utilizadas imagens de tecidos biológicos extraídas de (Kuhnel, 1991; Schneider, 1998; U.N.S., 2005). Foram selecionadas 50 imagens para compor o conjunto P e outras 50 imagens diferentes para compor o conjunto Q , totalizando 100 imagens. Tanto no conjunto de treinamento quanto no conjunto de teste, as imagens foram divididas em 5 classes, cada uma contendo 10 imagens.

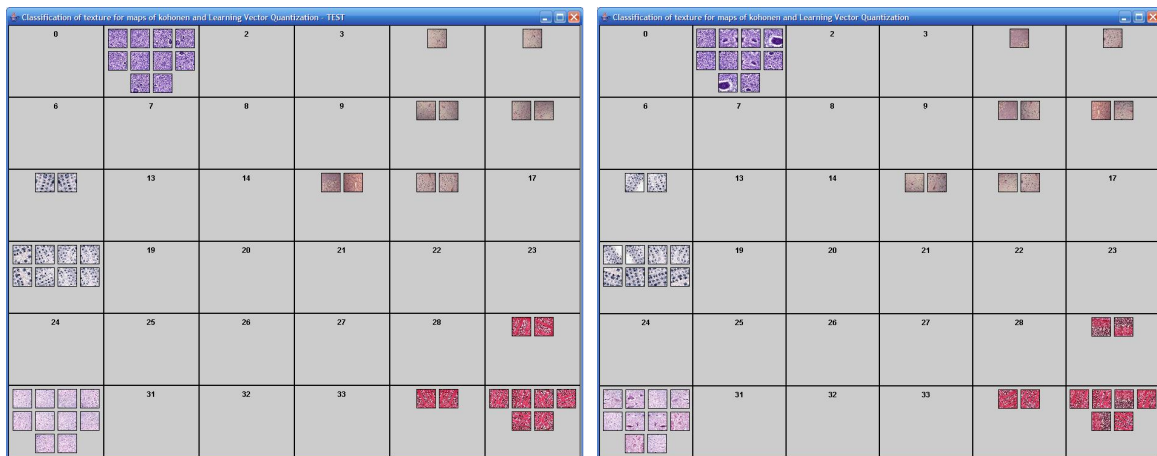
Para a fase de rotulação dos padrões, foi proposta uma arquitetura para rede SOM com 8 neurônios na camada de entrada e um mapa de saída com 36 neurônios (matriz quadrada 6X6). Para averiguar o desempenho da rede, o algoritmo de treinamento foi executado por 700 épocas utilizando uma taxa de aprendizagem de 5×10^{-3} . Após o

agrupamento, o mapa de saída da rede é apresentado em uma janela (Figura 2), onde as imagens com texturas similares estão organizadas em uma determinada região do mapa.



FIGURA 2 - Mapa de saída após treinamento com algoritmo SOM.

Observando o mapa de saída ilustrado na 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 que não pertencem à sua classe de padrões. Desta forma, para melhor desempenho no reconhecimento, fez-se necessário melhorar a divisão do espaço dos padrões agrupados pelo algoritmo SOM através do algoritmo supervisionado LVQ. Para isto, 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 usadas 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 ilustra o mapa de saída com as imagens do conjunto Q (imagens-teste). A Tabela 1 mostra os resultados dos testes efetuados com a rede treinada apenas com o algoritmo SOM. Já a Tabela 2, mostra os resultados relativos aos testes com a rede híbrida proposta.



3a. Mapa de saída da rede após treinamento com LVQ com conjunto P .

3b. Mapa de saída da rede treinada com LVQ após fase de testes com conjunto Q .

FIGURA 3 - Mapas de saída da rede híbrida proposta

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

Classe	Qtd. Acertos	Qtd. Erros	% Acertos
1	3	7	30
2	10	0	100
3	10	0	100
4	6	4	60
5	7	3	70
% Médio de Acertos			72

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

Classe	Qtd. Acertos	Qtd. Erros	% Acertos
1	10	0	100
2	10	0	100
3	10	0	100
4	10	0	100
5	10	0	100
% Médio de Acertos			100

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 a partir da descrição de suas texturas, enquanto que o resultado obtido com o uso do algoritmo SOM sem o LVQ mostrou uma eficiência média de 72%. Isto comprova que a aplicação do algoritmo LVQ ao SOM permitiu uma melhora significativa da classificação dos padrões.

5. Conclusão

Neste trabalho foi demonstrada a aplicação de um modelo de rede neural híbrido (SOM com LVQ) no reconhecimento de imagens caracterizadas pelos descritores de Haralick. Os resultados obtidos na classificação de imagens com a aplicação da técnica mostraram-se superiores quando comparados àqueles obtidos a partir do uso de um único modelo de rede neural não supervisionado (SOM). Isto sinaliza que o método proposto pode ser uma alternativa promissora para aplicações que envolvam a indexação e a recuperação automáticas de imagens digitais, principalmente em bancos de imagens médicas.

Agradecimentos

Os autores agradecem a Uninove pelo apoio financeiro.

Referências

Aksoy, S. & Haralick, R. M. (1998) "Textural Features for Image Database Retrieval", In: Proceedings of the IEEE Workshop on Content - Based Access of Image and Video Libraries, CBAIVL. IEEE Computer Society, Washington.

- Alves W. A. L. & Araújo S. A. (2005) “Uma Avaliação da Utilização de Dois Diferentes Modelos de Redes Neurais na Classificação de Imagens Baseada em Características Texturais”, In: Anais do XVIII Brazilian Symposium on Computer Graphics and Image Processing, Natal.
- Antani, S. et al. (2002) “A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video”, In: Pattern Recognition, Amsterdã, Elsevier Science B.V., n. 35, p. 945-965.
- Brunelli, R. & Mich, O. (2000) “Image Retrieval by Examples”, In: IEEE Trans. Multimedia, vol. 2, n.3, p. 164-171.
- Felipe, J. C. & Traina, A. J. M., “Utilizando Características de Textura para Identificação de Tecidos em Imagens Médicas”. In: Anais do II Workshop de Informática Médica – WIM’2002. Gramado.
- Gonzalez, R. C. & Woods, R. E. (2002) “Digital Image Processing”. 2. nd., Prentice Hall, New Jersey.
- Haralick, R. M. (1979) “Statistical and Structural Approaches to Texture”. In: Proceedings of The IEEE, vol. 67, n. 5, p. 786-804.
- Haykin, S. (2001) Redes Neurais Princípio e Prática”, 2. ed., Porto Alegre, Bookman.
- Kohonen, T. (1990) “The self-organizing map”. In: Proceedings of the Institute of Electrical and Electronics Engineers. Vol. 78, p. 1464-1480.
- Kuhnel, W. (1991) “Atlas de citologia, histologia e anatomia microscópica: teoria e prática”, São Paulo, Guanabara Koogan.
- Liu, Y. & Zhou, X. (2003) “A Simple Texture Descriptor for Texture Retrieval”, In: Proceedings of ICTT, p. 1662-1665.
- Martins M. P. et al. (2002) “Classificador de Texturas por Redes Neurais”, In: Anais do II Congresso Brasileiro de Computação, Itajaí-SC.
- Rui, Y. et al. (1997) “Image Retrieval: Past, Present and Future”, International Symposium on Multimedia Information Processing, Taiwan.
- Schneider, M. L. (1998) “Atlas de diagnóstico diferencial em citologia ginecológica”, Rio de Janeiro, Revinter.
- Sebe, N. & Lew, M. S. (2001) “Color-Based Retrieval”, In: Pattern Recognition Letters, Amsterdã, Elsevier Science B.V., n. 22, p. 223-230.
- Shiba, M. H. et al. (2005) “Classificação de imagens de sensoriamento remoto pela aprendizagem por árvore de decisão: uma avaliação de desempenho”, In: Anais do XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia-GO, p. 4319-4326.
- Smith, J. R. & Chang, S. (1996) “Tools and Techniques for Color Image Retrieval”, In: Symposium on Electronic Imaging: Science and Technology – Storage & Retrieval for Image and Video Databases IV, San Jose, CA. vol. 2670, p. 426-437.
- U.N.S. Dto. (2005) “Biología Bioquímica y Farmacia”, Universidad Nacional del Sur, disponível em: <http://www.anatomohistologia.uns.edu.ar>.