

USO DE REDES NEURAI ARTIFICIAIS E O MODELO DE SPOT NOISE PARA A CLASSIFICAÇÃO DE IMAGENS DE CEREAIS

ANTONIO C.G. MARTINS, ALEXANDRE S. SIMÕES, GUSTAVO I. PRADO

LAPI – Laboratório de Automação e Processamento de Imagens. Universidade Estadual Paulista “Júlio de Mesquita Filho” – Campus de Sorocaba.

Avenida Três de Março, 511, Alto da Boa Vista – 18087-180 - Sorocaba – SP – Brasil.

E-mails: amartins@sorocaba.unesp.br, assimoes@sorocaba.unesp.br, gustavoprado@grad.sorocaba.unesp.br

Abstract— This work presents a cereal grain image classification method based on the analysis of the frequency domain by an artificial neural network. The Spot Noise model is used as a guideline for the interpretation of the power spectrum and the definition of texture characteristic parameters. Results are presented for the classification of cereals according to the grains presented in the images. This method can be used for automatic evaluation of grains in loading belts.

Keywords— Artificial neural networks, texture, Spot Noise, cereal grain classification.

Resumo— Este artigo apresenta uma proposta de classificação de imagens de grãos de cereais a partir da análise de texturas no domínio da frequência com uma rede neural artificial. O modelo de Spot Noise é utilizado como guia para a interpretação do espectro de potência das imagens e definição dos parâmetros característicos da textura. São apresentados resultados para a classificação de diversas imagens. Este método pode ser utilizado na avaliação automática de grãos sendo carregados por esteiras em portos ou postos de distribuição.

Palavras-chave— Redes neurais artificiais, texturas, Spot Noise, classificação de grãos de cereais.

1 Introdução

O uso de visão computacional em um sistema de controle de qualidade busca a identificação e reconhecimento de objetos ou a classificação de uma cena. No caso de imagens que contenham objetos individualizados, como, por exemplo, os dispositivos em uma placa de circuito integrado, estes podem ser classificados de acordo com o formato e a cor. No entanto, existem situações em que se está interessado em analisar o conjunto das informações presentes na cena, como no caso de tecidos em uma linha de produção onde não existem objetos que possam ser segmentados. Em outros casos, mesmo que a cena seja formada por elementos individualizados, a identificação e classificação de cada elemento pode não ser uma abordagem viável, uma vez que os posicionamentos relativos acarretariam dificuldades na interpretação, como ocorre no caso de imagens de grãos obtidas de uma esteira de carregamento. Nestas situações, o conceito de textura pode ser utilizado para se analisar e classificar estas imagens.

São várias as aplicações que utilizam texturas, tais como: análise e classificação de carne bovina [Basset 2000], análise de superfícies de alimentos [Quevedo 2002], segmentação automática de terrenos através de fotos de satélites [Paget, 1996] e diagnósticos médicos [Christodoulou, 2003; Madablushi, 2003].

Além disto, para aplicações de visão computacional em tempo real é importante que o processo de análise de textura a ser utilizado seja rápido e robusto. Uma solução que preenche estes requisitos é a

utilização de redes neurais artificiais, que uma vez treinadas com exemplos característicos, produz resultados a partir de operações simples. Além disto, a capacidade de generalização, a partir do exemplo, leva a robustez requerida.

Devido às características de imagens com texturas aleatórias, o problema passa a ser de como utilizar redes neurais para realizar a análise.

Este artigo apresenta uma proposta de classificação de grãos de cereais a partir da análise de texturas aleatórias com uma rede neural artificial para a avaliação de imagens no domínio de frequência.

Uma aplicação para o método proposto é a avaliação automática de grãos sendo carregados por esteiras em portos ou postos de distribuição.

2 Fundamentação

2.1 Análise de texturas

A caracterização de uma textura pode ser realizada em diversos domínios, como o espacial ou de frequência. Sendo uma textura associada a qualidade visual global de uma imagem, o domínio da frequência apresenta uma característica bastante apropriada que é a de que cada ponto contém informação de todos os pixels da imagem.

Neste sentido, Weszka et al. (1976) apresentaram um método para gerar parâmetros característicos de uma textura a partir das somas das energias presentes em regiões do espectro de potência. Ao analisar estes parâmetros, os autores mostraram ser possível a discriminação de diferentes texturas.

Informações adicionais podem ser ainda extraídas do espectro de potência de uma textura calculando-se a entropia dos valores de energia ao invés da soma como sugerido por Jernigan e D'Astous (1984).

Existem também na literatura trabalhos que discutem o desempenho de diferentes métodos de caracterização de textura, tais como: a soma das energias, filtros espaciais, filtro de Gabor, estatísticas de tons de cinza e wavelet [Christodoulou, 2003]. Chen e Chen (1999), por exemplo, mostram que a análise de frequência consome menos tempo de processamento do que a baseada em filtros de Gabor.

Outras técnicas procuram utilizar redes neurais artificiais associadas ao resultado da aplicação de filtros de Gabor [Kachouie, 2003] ou na implementação de um banco de filtros [Jain, 1996a] no processo de classificação de texturas.

2.2 Redes Neurais Artificiais

O processo de análise de padrões pode utilizar técnicas de Inteligência Artificial (I.A.), tais como: redes neurais artificiais [Kovacs, 2002], lógica fuzzy [Shaw, 1999], algoritmos genéticos e sistemas especialistas [Russel, 1995]. Na área de classificação de imagens, as técnicas baseadas na teoria de redes neurais artificiais (RNAs) têm apresentado bons resultados [Jain, 1996a; Simões, 2003].

Entre as principais características que tornam uma RNA apropriada para a classificação de imagens estão [Mitchell, 1997]:

- Capacidade de aproximar qualquer função;
- Capacidade de aprender através de exemplos;
- Capacidade de generalização para situações que a rede não foi treinada.

Uma RNA baseia-se no modelo de funcionamento dos neurônios biológicos e mapeia os elementos de um conjunto de entrada em elementos de um conjunto de saída através da aproximação de uma função. A forma desta função é determinada a partir de um conjunto de exemplos que contém entradas e saídas pré-definidas, sendo, no entanto, que a arquitetura da RNA (número de neurônios e de camadas) e o processo de treinamento utilizados devem ser tais que permitam a representação desta função.

Assim, ao se decidir utilizar uma RNA, deve-se definir a arquitetura da rede, a forma da função de ativação, o conjunto de exemplos e o algoritmo de treinamento.

No processo de análise de imagens por uma RNA, pode-se reduzir o número de entradas ao se utilizar parâmetros que caracterizem esta imagem [Markopoulos, 2000]. A correta obtenção destes parâmetros é crucial para o sucesso da classificação, sendo assim, é importante que haja alguma fundamentação que leve a escolha destes. Neste trabalho, o modelo de *Spot Noise* é utilizado para se definir tais parâmetros.

2.3 Modelo de Spot Noise

Inicialmente utilizado por Wijk (1991) para sintetizar diferentes texturas, este modelo parte do princípio de que, dadas duas funções $f(x,y)$ e $g(x,y)$, a transformada de Fourier (TF) da convolução destas funções é igual a multiplicação das TFs de $f(x,y)$ e $g(x,y)$:

$$\mathfrak{F}(f(x,y) * g(x,y)) = \mathfrak{F}(f(x,y)) \cdot \mathfrak{F}(g(x,y)) \quad (1)$$

onde $\mathfrak{F}(\)$ denomina a operação da TF.

Se $f(x,y)$ representa uma mancha (*spot*) e $g(x,y)$ é um campo aleatório de impulsos de diversas intensidades, o espectro de potência da imagem obtida pela convolução de f e g é uma versão com ruídos do espectro de potência de $f(x,y)$, uma vez que o espectro de potência de um campo aleatório é essencialmente constante no caso de uma imagem digital. Deste modo, uma textura aleatória pode ser modelada como uma versão filtrada de um ruído, onde o filtro é representado por um *spot* de certa forma e tamanho, geralmente muito menor que o da imagem.

Em um processo de análise de texturas, parte-se do pressuposto de que cada textura está associada a um *spot* e que, portanto, o espectro de potência traz informações das características deste *spot*. Assim, deve-se buscar obter parâmetros deste espectro que permitam caracterizar o *spot*, e desta forma, classificar a textura.

Mais detalhes sobre o uso do modelo de *Spot Noise* na análise de texturas pode ser obtido em Martins (2007).

3 O método de classificação

Propõe-se a classificação de imagens com texturas a partir da análise de parâmetros obtidos do espectro de potência que permitam caracterizar o tamanho e formato dos *spots*. Para isto, o espectro de potência é dividido em 8 setores circulares, com ângulos internos de 45° cada um, e calculados os primeiros momentos destas sub-regiões dados por:

$$m_i = \frac{1}{N} \sum_{(u,v) \in R_i} r_{u,v} \cdot P(u,v) \quad (2)$$

onde $P(u,v)$ é o valor do espectro de potência na posição (u,v) , R_i é um setor circular com $i=1$ correspondendo a sub-região entre -22,5° e 22,5° e N é um fator de normalização. Pode-se notar que, a partir da análise da distribuição dos valores dos momentos, têm-se informações a respeito da simetria do *spot* (quanto mais próximos os valores de m_i , mais próximo de um círculo será o *spot*) e de seu tamanho (quanto maiores os valores de m_i , menores tendem a ser os *spots*).

Este método de classificação de textura necessita de um processo que permita a subdivisão e agrupamento dos pontos no espaço \mathbb{R}^8 dos m_i s de forma que cada uma destas regiões seja associada a uma textura em particular. Para se realizar esta tarefa, propõe-se a

utilização de uma RNA com 8 neurônios na camada de entrada, um para cada m_i , duas camadas ocultas e um neurônio na camada de saída. Levando-se em consideração que as regiões podem ter formatos genéricos, escolheu-se utilizar duas camadas ocultas conforme sugerido por Jain (1996b).

Nos testes realizados, foram utilizados 16 neurônios em cada uma das camadas ocultas e a seguinte função de ativação:

$$G = a. \tanh(bv) \quad (3)$$

onde v é a soma ponderada das entradas em um dado neurônio e os parâmetros a e b assumem, respectivamente, os valores 1,7159 e 2/3, como sugerido por Haykin (1999). Os números de neurônios nas camadas ocultas foram escolhidos de forma a gerarem um processo de treinamento convergente.

Para o processo de treinamento desta rede foi utilizado o algoritmo de retropropagação, adotando-se o modo *batch* para a atualização dos pesos sinápticos, descartando-se as atualizações que levassem a um aumento do erro total. Desta forma, se um peso sináptico é atualizado e o erro total aumenta, mantêm-se o valor anterior. Com isso, melhora-se o padrão de convergência da rede, não permitindo que o valor do erro total fique oscilando.

4 Resultados obtidos

Avaliou-se a capacidade do método proposto de classificar uma imagem de acordo com o tipo de grão de cereal presente, considerando-se existir apenas um tipo em cada imagem. Desta forma, o sistema de análise destina-se a verificar automaticamente se uma dada esteira está fazendo o carregamento do cereal correto.

Para isto, foram utilizadas imagens de três tipos de cereais: lentilha, soja e grão de bico. As imagens foram adquiridas com uma câmara digital comum (Pentax Optio S50), com 2560 pixels x 1920 pixels representando uma área total de 19,1 cm x 14,2 cm, não tendo sido empregado nenhum sistema de iluminação especial.

Em cada imagem foi selecionada uma área central de 512 x 512 pixels que continham os grãos. Em seguida, estas imagens foram transformadas para tons de cinza. A figura 1 apresenta 3 imagens características dos tipos de cereais utilizados.

Adquiriu-se 7 imagens para cada tipo de cereal sendo que os grãos foram despejados de forma aleatória em um suporte plástico. A partir de cada uma das imagens com os grãos, foram obtidas 16 imagens de 256 x 256 pixels utilizando-se uma máscara móvel com deslocamentos horizontais e verticais de 50 pixels. Assim, para cada tipo de grão obteve-se 112 imagens diferentes, sendo que 32 foram utilizadas para o treinamento da RNA e 80 para a validação.

Os valores de saída da rede associados a cada textura na fase de treinamento são apresentados na tabela 1. Estes valores procuram ordenar as imagens

de acordo com o tamanho dos grãos individuais dos cereais.

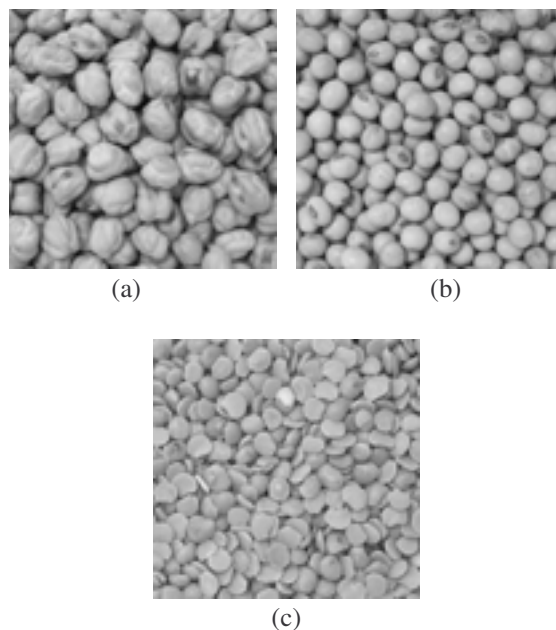


Figura 1 – Imagens de 512 x 512 pixels com grãos de: (a) grão de bico, (b) soja e (c) lentilha.

Na fase de treinamento, o processo convergiu utilizando-se uma taxa de aprendizado de 0,001 e como critério de parada a condição de que o erro quadrático para cada imagem fosse menor que 0,001.

Tabela 1. Valores de treinamento.

Tipo de grão	Valor de saída
Grão de bico	1
Soja	0
Lentilha	-1

Obtidos os pesos sinápticos, utilizou-se a RNA para classificar as 240 imagens da fase de validação. Os resultados obtidos são apresentados na figura 2 em forma de histograma.

Deve-se ressaltar que não foi aplicado nenhum filtro de pré-processamento nas imagens.

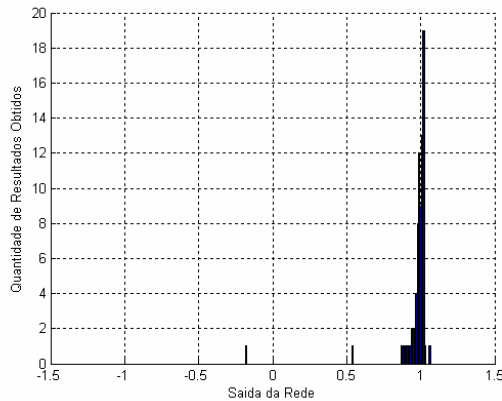
5 Conclusão

Este trabalho apresentou um processo de classificação de imagens com grãos de cereais a partir da análise do espectro de potência por uma rede neural artificial.

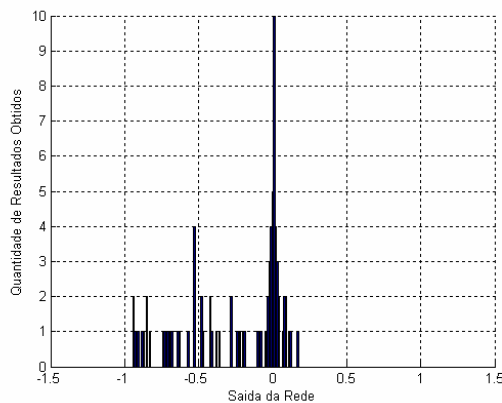
Analisando-se os histogramas com os valores de saída e definindo-se o intervalo [-1,5; -0,5] como sendo os referentes as imagens de lentilha, [-0,5; 0,5] referentes as imagens de soja e]0,5; 1,5] referentes as imagens de grão de bico, o método proposto, de acordo com os testes realizados, classificou corretamente 100% das imagens de lentilha, 73% das de soja e 98% das de grão de bico. Utilizando-se intervalos menos simétricos, [-1,5; -0,8] para lentilha, [-

0,8; 0,8] para soja e]0,8; 1,5] para grão de bico, as taxas de acerto passam para 100%, 89% e 97% respectivamente.

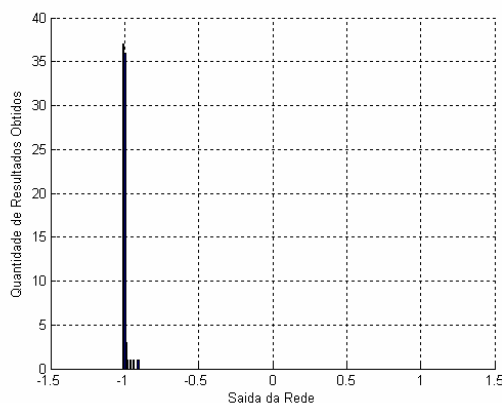
Assim, os resultados obtidos mostram a adequação do método para a avaliação automática de imagens de grãos sendo carregados por esteiras em portos ou postos de distribuição.



(a)



(b)



(c)

Figura 2 – Valores de saída da RNA para as imagens de: (a) grão de bico, (b) soja e (c) lentilha.

Referências Bibliográficas

- Basset, O., Buquet, B., Abouelkaram, S., Delachartre, P. e Culioli, J. (2000) "Application of texture image analysis for the classification of bovine meat", *Food Chemistry*, 69, p. 437-445.
- Chen, C. C. e Chen, C. C. (1999) "Filtering methods for Texture Discrimination", *Pattern Recognition Letters*", v. 20, p. 783-790.
- Christodoulou, C. I., Pattichis, C. S., Pantziaris, M. e Nicolaides, A. (2003) "Texture-based classification of atherosclerotic carotid plaques", *IEEE Transactions on Medical Imaging*, v. 22, no. 7, p. 902-912.
- Haykin, S. (1999). "Neural Networks: a comprehensive foundation" 2 ed., Upper Saddle River: Prentice-Hall.
- Jain, A. K. e Karu, K. (1996a) "Learning Texture Discrimination Masks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, no. 2.
- Jain, A. K., Mao, J. e Mohiuddin, K.M. (1996b) "Artificial Neural Networks: A Tutorial", *Computer*, v. 9, no. 3, p. 31-44.
- Jernigan, M. E. e D'Astous, F. (1984) "Entropy-Based Texture Analysis in the Spatial Frequency Domain", *IEEE Transactions on Pattern Recognition and Machine Intelligence*, v. PAMI-6, no. 2.
- Kachouie, N. N. e Alirezaie, J. (2003) "Texture Segmentation using Gabor Filter and Multi-Layer Perceptron", *Proceedings of the IEE International Conference on Systems, Man and Cybernetics*, v.3, p. 2897-2902.
- Kovacs, Z. L. (2002) "Redes Neurais Artificiais - Fundamentos E Aplicações". 4ed. São Paulo: Livraria da Física.
- Madablushi, A. E. e Metaxas, D. N. (2003) "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions", *IEEE Transactions on Medical Imaging*, v. 22, n. 2, p.155-169.
- Markopoulos, C.; Kouskos, E.; Koufopoulos, K.; Kyriakou, V.; Gogas; J. (2001) "Use of artificial neural networks (computer analysis) in the diagnosis of microcalcifications on mammography". *European Journal of Radiology*, v.39, p. 60-65.
- Martins, A. C. G.; Simões, A. S.; Prado, G. I. "Classificação de texturas usando redes neurais artificiais e o modelo de Spot Noise". *Anais do VI Encontro Nacional de Inteligência Artificial*.
- Mitchell, T.M. (1997) "Machine Learning". Boston: WCB/McGraw-Hill.
- Paget, R. D. e Longstaff, D. (1996) "Terrain mapping of radar satellite images", *Electronic Imaging*, 6(2), p. 6-7.
- Quevedo, R., Carlos, L., Aguilera, J. M. e Cadoche, L. (2002) "Description of food surfaces and microstructural changes using fractal image

- texture analysis”, *Journal of Food Engineering*, 53, p. 361-371.
- Russel, S.J. e Norvig, P. (1995). “Artificial Intelligence – A Modern Approach.” 2 ed. London: PrenticeHall International.
- Shaw, I. e Simões, M.G. (1999). “Controle e Modelagem Fuzzy”. São Paulo: Edgard Blucher.
- Simões, A. S. ; Costa, A.R. (2003). “Classificação de laranjas baseada em padrões visuais”. *Anais do 6o Simpósio Brasileiro de Automação Inteligente*.
- Weszka, J. S., Dyer, C. R. e Rosenfeld, A. (1976) “A Comparative Study of Texture Measures for Terrain Classification”, *IEEE Transactions on Systems, Man, and Cybernetics*, v. SMC-6, no. 4, p. 269-285.
- Wijk, J. J. (1991) “Spot Noise”, *Computer Graphics*, 25(4), p. 309-318.