



VII Workshop de Visão Computacional

Anais do Evento

www.wvc2011.ufpr.br

22 a 25 de maio de 2011 - UFPR - Curitiba

ISSN - 2175-6120

WVC 2011

Anais do VII Workshop de Visão Computacional

Universidade Federal do Paraná

ISSN 2175-6120

Editores:

Luiz Antônio Pereira Neves

Lucas Ferrari de Oliveira

Hugo Vieira Neto

Curitiba, 22 a 25 de Maio de 2011

Geração de um Classificador para Detecção do HLB em Citros Utilizando Técnica *Box-Counting Multifractal*

Diego Carlos Pereira da Silva¹; Lúcio André de Castro Jorge²; Adolfo Posadas³; Maria Stela Veludo de Paiva⁴

1- Mestrando, Escola de Engenharia de São Carlos - Universidade de São Paulo - USP, Depto. de Engenharia Elétrica, São Carlos-SP, dicapessilva@usp.br.

2- Pesquisador, Embrapa Instrumentação Agropecuária, São Carlos – SP, lucio@cnpdia.embrapa.br.

3- Pesquisador, International Potato Center (CIP)- CGIAR, São Carlos-SP, a.posadas@cgiar.org

4- Professora, Escola de Engenharia de São Carlos - Universidade de São Paulo - USP, Depto. de Engenharia Elétrica, São Carlos-SP, mstela@sc.usp.br.

Abstract

The Huanglongbing (HLB) is a serious disease that attacks citrus. It is caused by insect Diaphorina citri. Contamination occurs when the insect feeds on infected plants, from that moment he becomes the transmission vector of the same. HLB infected leaves with symptoms that can be mistaken with symptoms of some diseases and nutritional deficiencies. HLB to date is an incurable disease, and difficult to diagnose. In this way, this work developed a tool based on multifractal theory in order to characterize the self-similarities present in the leaves infected with HLB. Were analyzed 80 samples of leaves, 20 leaves infected with CVC, 20 with deficiency of manganese, 20 with deficiency of zinc and 20 infected with HLB. The results show that the method is quite effective for detection of HLB.

1. Introdução

O Huanglongbing (HLB, ex-greening) é uma doença destrutiva que ataca o citros. Essa doença tem como vetor transmissor o inseto *Diaphorina citri* na Ásia e na América, e o inseto *Trioza erytreae* na África. A contaminação se dá quando o inseto se alimenta de plantas doentes. A partir desse momento o inseto passa a ser o transmissor da doença [1]. As folhas das plantas contaminadas com o HLB apresentam manchas com formatos irregulares, de cor verde clara ou amarela, mescladas com o verde normal, sem apresentar uma nítida divisão entre elas. Os sintomas do HLB presentes nas folhas das plantas podem ser confundidos facilmente com sintomas de deficiências nutricionais como deficiência de Manganês e de Zinco e doenças como a CVC (Clorose Variegada dos Citros). A deficiência de Manganês faz com que as folhas apresentem clorose mais pálida e menos acentuada entre as nervuras. Na deficiência de

Zinco as folhas apresentam clorose acentuada do limbo entre as nervuras. Na CVC surgem pequenas manchas amareladas e irregulares, espalhadas na frente da folha [2].

Para diagnosticar o HLB existem atualmente dois métodos, o método de inspeção visual e a análise do PCR (*Polymerase Chain Reaction*) [3]. O primeiro método apresenta a desvantagem de permitir diagnosticar a doença, somente vários meses após a contaminação da planta. Além disso, devido às condições de campo a que são acometidos os inspetores, os sintomas podem ser confundidos com outras doenças e deficiências nutricionais, proporcionando taxas de acerto inferiores a 50%. O método de análise do PCR é o mais eficaz, porém, o alto custo da análise impossibilita a sua realização em uma escala que possa permitir o controle da propagação da doença.

A prevenção do HLB se dá por meio do controle do vetor transmissor com a utilização de tratamentos químicos e pela erradicação das plantas contaminadas após o diagnóstico positivo para a doença.

Por se tratar de uma doença incurável até o presente momento, e de difícil diagnóstico, esforços vem sendo realizados na tentativa de se desenvolver ferramentas eficazes que auxiliem no diagnóstico dessa doença.

O presente artigo apresenta uma ferramenta baseada na teoria *multifractal* que caracteriza as auto-similaridades das folhas das plantas contaminadas pelo HLB, demais doenças e deficiências nutricionais de forma a permitir o reconhecimento de padrões e a identificação do HLB.

A teoria *multifractal* permite a caracterização de fenômenos complexos de maneira quantitativa, mesmo em meio a variações tanto temporais como espaciais. Esta teoria vem sendo cada vez mais utilizada para a análise e simulação de fenômenos complexos em uma grande variedade de escalas [4], devido aos bons resultados obtidos [5], [6], [7], [8], [9] e [10]. A análise *multifractal* é um método robusto utilizado para descrever eventos,

estruturas ou fenômenos, e possui como característica fundamental a auto-similaridade. Esta característica permite que independente da complexidade da forma de um objeto, observando-se mais detalhadamente a sua estrutura, um observador pode encontrar as mesmas formas (ou semelhantes) em escalas diferentes [11].

2. Materiais e métodos

2.1 Aquisição de Dados

O Grupo Fischer [12] disponibilizou à EMBRAPA Instrumentação Agropecuária, 80 amostras de folhas de citros de Valência com o porta-enxerto *Swingle*. Essas folhas já vieram com um pré-diagnóstico realizado por meio de inspeção visual realizada em laboratórios de maneira controlada longe das adversidades que dificultam o diagnóstico em campo, por técnicos especializados (pragueiros), não sendo realizado nenhum teste químico que pudesse confirmar a autenticidade do diagnóstico. Do total de 80 folhas, 20 delas foram diagnosticadas contendo a doença CVC, 20 apresentando deficiência de Manganês, 20 apresentando deficiência de Zinco e 20 delas contendo o HLB.

Foi realizada a aquisição das imagens dessas 80 folhas, utilizando um *scanner* de mesa da marca HP modelo *Scanjet G4050*. A resolução do mesmo foi ajustada para 100 DPI e a dimensão das imagens obtidas foi 400 x 200 *pixels*.

Na Figura 1 pode ser observada a imagem de folhas utilizadas neste trabalho, sendo a imagem (a) de uma folha com CVC, (b) de uma folha com deficiência de Manganês, (c) de uma folha com deficiência de Zinco e (d) de uma folha com HLB.

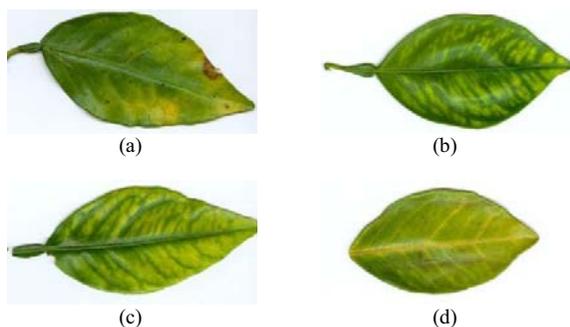


Figura 1 – (a) Imagem de uma folha com CVC, (b) de uma folha com deficiência de Manganês, (c) de uma folha com deficiência de Zinco e (d) de uma folha com HLB.

2.1. Segmentação por cor e Binarização

Como pode ser percebido nas imagens mostradas na Figura 1, em todas as doenças e deficiências nutricionais apresentadas neste artigo, os sintomas presentes nas folhas evidenciam a cor amarelo, e devido a essa característica presente em todas as folhas foi realizada a segmentação por cor [13].

Para a segmentação por cor da folha, foi desenvolvida uma ferramenta baseada em uma Rede Neural Artificial do tipo MLP (*MultiLayer Perceptron*) [14] com algoritmo *backpropagation*, de forma a separar a cor amarelo das demais cores presentes na folha. Foram utilizados como parâmetros de entrada os valores R, G e B do pixel, sendo selecionadas manualmente as áreas de interesse na imagem que pudessem representar as diversas tonalidades existentes para a cor amarela. Para o treinamento da rede foram utilizadas 64 amostras das folhas de citros e para os testes foram utilizadas 16 amostras. Dessa forma foram criadas as classes Amarelo e Fundo. Para os parâmetros referentes à taxa de aprendizagem, *momentum* e quantidade de neurônios na camada neural escondida, foram definidos, respectivamente, os valores 0,3, 0,2 e 3, com base em testes previamente realizados.

Na Figura 2 pode ser observada a imagem das folhas da Figura 1 após a segmentação, sendo a imagem (a) de uma folha com CVC, (b) de uma folha com deficiência de Manganês, (c) de uma folha com deficiência de Zinco e (d) de uma folha com HLB. O preto representa a região amarela da folha.

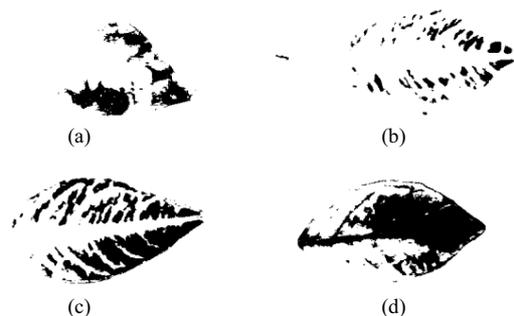


Figura 2 – (a) Imagem de uma folha após a segmentação com CVC, (b) de uma folha com deficiência de Manganês, (c) de uma folha com deficiência de Zinco e (d) de uma folha com HLB.

Em seguida foi aplicado o método *Box-Counting*, para a análise *multifractal* [15].

2.3. Análise multifractal

Como já mencionado na introdução, a análise *multifractal* permite a caracterização das auto-similaridades presentes nas diversas estruturas do objeto analisado [11].

Neste artigo foi utilizada a teoria *multifractal* na tentativa de se caracterizar as estruturas nas folhas amostradas que pudessem evidenciar características inerentes ao CVC, deficiências nutricionais e HLB.

Para a análise *multifractal* das imagens utilizadas neste artigo, foi utilizado o método *Box-Counting* [16], sendo este método eficiente para a análise de sinais binários.

No método *Box-Counting* a imagem é dividida em caixas de tamanho L e a probabilidade de se encontrar uma estrutura naquela região é calculada de acordo com a equação (1):

$$P_i(N) = \frac{N_i(L)}{Nt} \quad (1)$$

Na equação (1) $P_i(N)$ representa o resultado do cálculo da probabilidade de determinada estrutura em cada uma das caixas em que foi dividida a imagem, $N_i(L)$ é a soma total dos valores de determinada estrutura de uma das caixas da imagem e N_i corresponde o valor da soma dessa estrutura para a imagem inteira.

Os valores obtidos na equação (1) são utilizados para determinar o valor de $\mu_i(q, l)$. Esse valor corresponde à normalização das probabilidades em relação ao intervalo de valores q , chamado de momento estatístico, esse valor permite analisar as estruturas explorando as diferentes regiões de medidas singulares. Para valores de $q > 1$ a região de medida mais singular é amplificada, para $q < 1$ a região com menor singularidade é acentuada e para $q = 1$ a medida original é replicada. A função $\mu_i(q, l)$ é dada pela equação (2):

$$\mu_i(q, l) = \frac{[P_i(l)]^q}{\sum_j [P_j(l)]^q} \quad (2)$$

Os valores das funções $f(q)$ e $\alpha(q)$ determinam a dimensão fractal, e são descritos pelas equações (3) e (4):

$$f(q) = \lim_{l \rightarrow 0} \frac{\sum_i \mu_i(q, l) \times \ln(\mu_i(q, l))}{\ln(l)} \quad (3)$$

$$\alpha(q) = \lim_{l \rightarrow 0} \frac{\sum_i \mu_i(q, l) \times \ln(P_i(l))}{\ln(l)} \quad (4)$$

A Figura 3 mostra o gráfico da relação $\alpha(q)$ por $f(q)$ de uma folha, sendo a imagem (a) de uma folha com CVC, (b) de uma folha com deficiência de Manganês, (c) de uma folha com deficiência de Zinco e (d) de uma folha com HLB.

Na figura 3 o eixo das abscissas é constituído pelos valores de $\alpha(q)$ e o eixo das ordenadas é constituído pelos valores de $f(q)$.

O formato do gráfico representa uma assinatura para cada um dos padrões estudados.

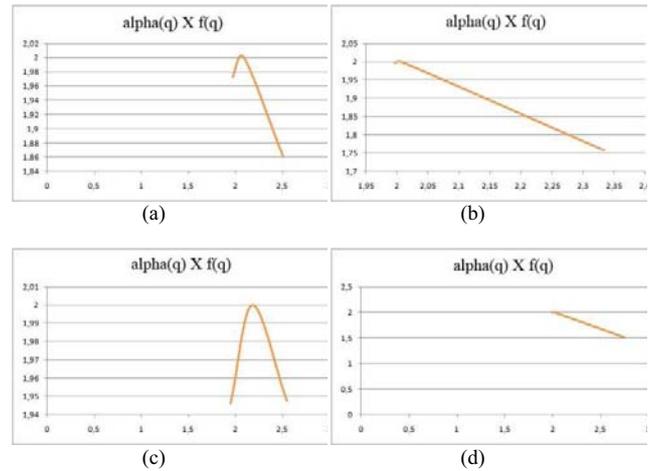


Figura 3 – (a) Imagem de uma folha com CVC, (b) de uma folha com deficiência de Manganês, (c) de uma folha com deficiência de Zinco e (d) de uma folha com HLB.

3. Classificadores

Na tentativa de se tentar identificar padrões para a realização do diagnóstico das folhas, alguns valores obtidos com análise a *multifractal* foram utilizados para compor o vetor de características.

3.1. Vetor de características

O vetor de características foi composto pelos valores das funções $\alpha(q)$, $f(q)$ e pelos valores das coordenadas de $X1, X2, X3, Y1, Y2$ e $Y3$ que representam os lados do triângulo formado pelo comportamento *fractal*.

Além dos valores das funções supracitadas foram utilizados os valores da dimensão fractal $D0, D1$ e $D2$ dados pela equação (5):

$$FD = \lim_{r \rightarrow 0} \frac{\log(N(r))}{\log(r)} \quad (5)$$

Na equação (5) $N(r)$ corresponde ao número de caixas necessárias para cobrir toda a imagem.

3.2. Classificadores

Neste artigo para a geração do classificador foi utilizada uma rede neural artificial do tipo MLP (*MultiLayer Perceptron*) [14] presente na API (*Application Programming Interface*) do *toolbox* WEKA [17]. Essa rede neural é baseada no algoritmo *backpropagation*.

Com base nos testes realizados durante a classificação, foram definidos os parâmetros referentes à taxa de aprendizagem, *momentum* e quantidade de neurônios na

camada neural escondida, respectivamente, como 0,3, 0,2, e 20. A função de ativação utilizada foi a sigmóide.

4. Resultados e Discussão

Foi realizado o experimento utilizando o vetor de características com os valores obtidos por meio da análise *multifractal* conforme mencionado na seção 3.1.

Durante a classificação foram utilizados 10 *folds* com validação cruzada, aplicados a todas as imagens das folhas.

Do total das 80 amostras analisadas pelo classificador, ele conseguiu classificar corretamente 60 delas, representando um percentual de 75%.

O valor do erro absoluto médio foi de 0,1628 e o valor do erro quadrático médio foi de 0,3175.

A tabela 1 mostra a matriz de confusão contendo os resultados da classificação de acordo com cada classe.

Tabela 1 – (a) CVC, (b) deficiência de Manganês, (c) deficiência de Zinco e (d) de uma folha com HLB.

A	b	c	d
15	0	1	4
1	18	1	0
2	2	15	1
5	1	2	12

A interpretação dos valores presentes na tabela 1 é dada da seguinte forma:

- Do total de 20 instâncias contendo CVC, 1 instância foi classificada com deficiência de Zinco e 4 instâncias com HLB;
- Do total de 20 instâncias contendo deficiência de Manganês, 1 instância foi classificada com CVC e 1 instância foi classificada com deficiência de Zinco;
- Do total de 20 instâncias contendo deficiência de Zinco, 2 instâncias foram classificadas como CVC, 2 instâncias foram classificadas com deficiência de Manganês e 1 instância foi classificada como HLB;
- Do total de 20 instâncias contendo HLB, 5 instâncias foram classificadas como CVC, 1 instância foi classificada como deficiência de Manganês e 2 instâncias foram classificadas como deficiência de Zinco.

Para o CVC a precisão dada pelo classificador foi de 65,2%, para a deficiência de Manganês foi de 85,7%, para a deficiência de Zinco foi de 78,9% e para o HLB foi de 70,6%.

5. Conclusões

Este trabalho apresenta resultados da aplicação de um classificador baseado em uma rede neural artificial, tendo como parâmetros de entrada os valores obtidos pela análise *multifractal*, aplicados a uma área de extrema importância. Atualmente os valores obtidos por meio da inspeção visual para a detecção do HLB são inferiores a 50%, sendo os sintomas presentes nas folhas contaminadas com HLB confundido muitas vezes com os sintomas de outras doenças e deficiências nutricionais.

Dessa forma, a ferramenta desenvolvida nesse artigo baseada na teoria *multifractal* apresenta resultados satisfatórios se comparada ao método de inspeção visual.

Este trabalho pode ser utilizado como referência para o desenvolvimento de um sistema que possa ser ter aplicação direta no campo.

10. Referências

- [1] BOVÉ, J. M. Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. **Journal of Plant Pathology**, Pisa, v. 88, p.7-37, 2006.
- [2] FUNDECITRUS. Manual Técnico do Greening 2009. Disponível em: <<http://www.fundecitrus.com.br/ImageBank/PageFlip/pagflip.aspx?idPage=143>>. Acesso em: 29 abr. 2010.
- [3] INNIS, M. A.; GELFAND, D. H.; SNINSKY, J. J.; WHITE, T. J. (eds.), PCR Protocols: A guide to methods and applications, Academic Press: San Diego, 1990.
- [4] SCHERTZER, D., S. LOVEJOY, 1994: Multifractal Generation of Self-Organized Criticality, in Fractals In the natural and applied sciences Ed. M.M. Novak, Elsevier, North-Holland, 325-339.
- [5] MCCARTNEY, H.A. & FITT, B.D.L. Dispersal of foliar fungal plant pathogens: mechanisms, gradients and spatial patterns. In Jones, D.G. (ed.). **The Epidemiology of Plant Diseases**. Dordrecht, Kluwer. 1998. p. 139-160.
- [6] MCATEER J.R.T., YOUNG C.A., IRELAND J., and GALLAGHER P.T. 2007. The bursty nature of solar flare x-ray emission. The American Astronomical Society..The Astrophysical Journal, 662:691-700.
- [7] POSADAS A., GIMENEZ D., QUIROZ R., and PROTZ R. 2003. Multifractal Characterization Of Soil Pore Systems. **Soil Sci. Soc. Am. J.** 67: 1361-1369.
- [8] POSADAS A., ZOROGASTÚA P., QUIROZ R. and LÉON-VELARDE C. 2005. Multifractal characterization of the spatial distribution of ulexite in a Bolivian salt flat.

International Journal of Remote Sensing. Vol. 26 (3): 615–627.

[9] LATKA M., GLAUBIC-LATKA M., LATKA D., WEST B. 2002. The loss of multifractality in migraines. <http://arxiv.org/PS_cache/physics/pdf/0204/0204010v1.pdf>.

[10] IVANOV P. C., AMARAL L. A. N., GOLDBERGER A. L., HAVLIN S., ROSENBLUM M. G., STRUZIK Z. R., and STANLEY H. E. 1999. Multifractality in healthy heartbeat dynamics. *Nature* 399: 461-465.

[11] MANDELBROT, B. 1983. *The Fractal Geometry of Nature*. Copyrighted material, USA.

[12] Grupo Fischer. Disponível em: <http://www.citrosuco.com.br/fischer/fischer/sites/fischer/portal_grupo/home/index.html>. Acesso em: 20 fev. 2011.

[13] SIMÕES, A. S. Segmentação de imagens por classificação de cores: uma abordagem neural. DISSERTAÇÃO DE MESTRADO. ESCOLA POLITÉCNICA DA USP. São Paulo. 2000.

[14] WITTEN, I. H.; FRANK, E. *Data mining : practical machine learning tools and techniques*. 2 ed. San Francisco, 2005.

[15] LOPES, R.; BETROUNI, N. Fractal and multifractal analysis. **Medical Image Analysis**. 13: 634-649.

[16] RUSSEL, D.; HANSON, J.; OTT, E. Dimension of strange attractors. *Physical Review Letters* 45 (14), 1175-1178.

[17] WEKA. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 29 abr. 2010.

11. Agradecimentos

Agradecemos o apoio financeiro do CNPq (processo 578627/2008-6), fornecimento das folhas de citros da empresa Fischer S.A.