

Detecção Automática da Área Foliar Danificada da Soja através de Imagens Digitais

Antonio Carlos de Nazaré Júnior*, David Menotti*, José Maria Ribeiro Neves* e Tuneo Sediya[†]

*Departamento de Computação - Universidade Federal de Ouro Preto

Campus Universitário, Morro do Cruzeiro, CEP 35400-000, Ouro Preto, Minas Gerais, Brasil

[†]Departamento de Fitotecnia - Universidade Federal de Viçosa

Campus Universitário, CEP 36570-000, Viçosa, Minas Gerais, Brasil

juniorcristais@gmail.com, menottid@gmail.com, jmneves@iceb.br, tuneo@ufv.br

Resumo—Na agricultura, são muitas as dificuldades envolvidas no manejo de pragas na lavoura. Um dos tipos de ataque das pragas é feito diretamente na folha da planta. A detecção precisa da área foliar danificada é essencial para a determinação de medidas de controle. Essa análise muitas vezes é feita de forma demorada e imprecisa. Neste trabalho, é proposto um método para a detecção/quantificação automática da área foliar danificada por pragas usando imagens digitais. Inicialmente a imagem da amostra foliar é pré-processada através de filtragem e remoção de sombras e então segmentada. Em um segundo momento, uma reconstrução das estruturas danificadas da folha é realizada, e em seguida a área danificada é estimada. Experimentos mostram que o método proposto melhora em 40% o erro absoluto gerado por um método da literatura e que os resultados computacionais são muito mais precisos que a avaliação visual realizada por seres humanos.

Keywords—Segmentação de Imagens, Geometria Computacional, Soja, Área foliar danificada.

I. INTRODUÇÃO

A determinação da área foliar é de grande importância para a aplicação de técnicas como poda, adubação e densidade de plantio [1]. Ela também é um indicativo da produtividade, pois o processo fotossintético depende da interceptação da energia luminosa pela folha [2]. Ou seja, a superfície foliar de uma planta é a base do rendimento potencial da cultura.

Outra característica que pode ser extraída através da determinação da área foliar é a quantificação dos danos causados por pragas e doenças [3]. Para o controle das principais pragas é necessário, entre outros requisitos, conhecer os danos já causados na planta. Tais danos podem ser detectados através do estudo da área foliar atingida pelas pragas [4]. A detecção da quantidade precisa da área foliar danificada é essencial para a determinação de ações de controle como aplicação de defensivos, visto que uma pequena área foliar danificada pode dispensar medidas de controle.

Existem vários métodos para medição da área foliar, entretanto, na prática são utilizados principalmente três: a avaliação humana, o método de dimensões foliares e o método de utilização de aparelhos como planímetro e integrador de área [5].

A avaliação humana geralmente é realizada por um especialista (agrônomo). Além de ser um processo impreciso, nem sempre há um profissional disponível para análise da lavoura e dessa forma a demora no diagnóstico pode implicar em futuros prejuízos ao agricultor.

O método de dimensões foliares baseia-se nas dimensões do limbo foliar. Por exemplo, em [6] são utilizados o comprimento da nervura principal e a largura da folha para estimar a área foliar da erva-do-touro (*Tridax procumbens*) usando uma equação de regressão. Esse método é amplamente empregado no campo, entretanto, a sua aplicação demanda tempo e sua precisão não é satisfatória [5].

O planímetro é um aparelho que permite medir a área de uma figura plana, por meio de métodos mecânicos. Sua precisão é satisfatória, porém o seu uso torna-se difícil em folhas com bordas irregulares, além de demandar muito trabalho e tempo [7]. Os integradores de área foliar (ex.: *LI-COR 3100, Area Meter*) são precisos apenas na medição de área foliar de amostras com o limbo inteiro, portanto não conseguem estimar a área danificada. Outra desvantagem é que eles não são produzidos nacionalmente e por isso são caros e de difícil manutenção e acesso. Os integradores de área são encontrados apenas em grandes laboratórios de pesquisa, o que dificulta o acesso do agricultor a tal tecnologia.

Observa-se assim que mesmo existindo diversas técnicas tradicionais para determinação da área foliar, elas demandam trabalho exaustivo e tempo, além de possuírem certo grau de imprecisão [5]. Além disso, as técnicas de medição não são realizadas na maioria das vezes por um agricultor, o que dificulta a agilidade do diagnóstico.

Com os avanços da computação, principalmente na área de Processamento Gráfico (Processamento Digital de Imagens e Visão Computacional), é possível desenvolver métodos de determinação da área foliar alternativos aos métodos clássicos citados. Wilcken *et al.* [5] demonstram a viabilidade da utilização de digitalizadores de imagens (*scanner*) como método alternativo na detecção de área foliar consumida por insetos. Einhardt & Ferreira [8] concluíram que um *software* pode ser utilizado na determinação de área foliar, uma vez que mostrou resultados equivalentes aos métodos tradicionais. Mura *et al.* [9] mostraram que o uso de um

método de visão computacional, também utilizando imagens adquiridas por um *scanner*, é melhor na determinação do percentual da folha de soja afetada pela lagarta da soja, em relação à avaliação de um agrônomo.

Porém o trabalho desenvolvido em [9] apresenta algumas desvantagens: Para que o método funcione corretamente, é necessário que a imagem da amostra foliar esteja livre de impurezas, como ciscos e pequenos pedaços que se desintegram da folha durante o manejo da mesma sobre o *scanner*. Ou seja, as imagens devem ser compostas apenas pela amostra foliar e um fundo branco, o que é impraticável em um ambiente agrícola; Durante o pré-processamento da imagem não é realizado nenhum tratamento para remoção da sombra que é incorporada à folha durante o processo de varredura ótica para a aquisição da imagem, o que ocasiona uma detecção imprecisa da área foliar; E por último, o sistema não possibilita o preenchimento automático de extremidades comprometidas da folha, sendo assim, a área danificada de folhas com grandes danos presentes na borda é detectada de forma insatisfatória.

O presente trabalho propõe um novo método automático de detecção e quantificação automática da área foliar danificada da soja utilizando técnicas de processamento digital de imagens e geometria computacional que trata dos problemas apresentados na literatura como remoção de sombras, eliminação de objetos indesejáveis provenientes da aquisição da imagem e reconstrução de extremidades comprometidas da folha¹. O método proposto também é comparado com avaliações de especialistas humanos e o método de Mura *et al.* [9].

O restante deste artigo está organizado da seguinte forma. Na Seção II, são apresentadas as técnicas que constituem o método proposto. Os experimentos computacionais são descritos na Seção III e discutidos na Seção IV. Finalmente, na Seção V, são apresentadas conclusões e propostas de trabalhos futuros.

II. METODOLOGIA

A. Classificação

Para classificar os danos como danos de coleópteros e lagartos foi utilizado o classificador linear LDA (Linear Discriminant Analysis). O LDA procura a melhor reta de separação dos dados de forma que aumente a distância entre classes e diminua a distância intra classe.

O LDA calcula os centróides das classes (μ_i) e o centróide global (μ). O cálculo dos centroides é feito calculando a média sobre os vetores de características. Os vetores de características são os graus máximo e médio de conectividade da rede dado um limiar. Tendo os centróides locais, os dados são normalizados subtraindo-os. Em seguida a matriz

de covariância de cada classe é calculada para os dados normalizados. Depois a matriz de covariância conjunta é calculada utilizando as matrizes de covariância das classes e as respectivas probabilidades a priori. Tendo esta matriz é calculada a inversa da matriz de covariância (C^{-1}). Com isso podemos calcular a função discriminante para cada classe. Essa função é calculada pela fórmula:

$$f_i = \mu_i C^{-1} x_k^T - \frac{1}{2} \mu_i C^{-1} \mu_k^T - \ln(p_i) \quad (1)$$

A classe do dano que será classificado é aquela que a função discriminante é maximizada dado o vetor de características (x_i) do dano.

III. EXPERIMENTOS COMPUTACIONAIS

Para execução de testes foi utilizada uma base de 185 imagens de folíolos de soja. Essas imagens foram obtidas escaneando os folíolos com auxílio de uma scanner HP ScanJet G2410. As imagens foram pré-processadas utilizando os métodos mostrados na seção ???. Todos os danos extraídos foram salvos em imagens gerando uma base de 1700 danos.

Os danos extraídos foram classificados manualmente pelo autor Thiago para o aprendizado supervisionado como sendo danos de coleópteros ou lagartas. Essa classificação gerou uma base de 1000 causados por lagartas e 700 danos causados por coleópteros.

Para validação dos testes foi utilizada Validação Cruzada (“K-Fold Cross Validation”), onde nossa base foi separada em 10 partições, cada partição continha 170 imagens de danos, sendo 100 imagens de danos de lagartas e 70 de danos de coleópteros. foram feitas 10 iterações, a cada iteração uma partição diferente era separada. A partição separada foi utilizada em testes e as demais foram utilizadas no treinamento do sistema.

Os testes executados obtiveram precisão de classificação de 76,82%. A taxa elevada de erro de 23,18% pode ser explicada pois a supervisão do sistema não foi feita pelos especialistas na área, já que muitos danos podem ter sido erroneamente classificados como danos de lagartas ou coleópteros.

IV. RESULTADOS E DISCUSSÃO

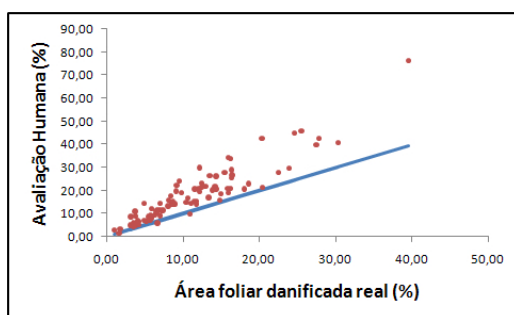
Na Tabela I são apresentados os valores de erro absoluto médio (EAM) e erro relativo médio (ERM) dos métodos avaliados. Os valores estão dispostos no seguinte formato: $\mu \pm \sigma$, onde μ representa a média e σ o desvio padrão. Os valores foram separados em grupos, divididos em faixas de porcentagem ideal de área danificada, para permitir uma análise mais minuciosa das avaliações

Quanto menor o valor de EAM, maior é a exatidão da estimativa da área foliar danificada. Com relação a esse critério, observou-se que o método *Proposto* apresentou o melhor resultado, pois o EAM gerado por ele foi um pouco

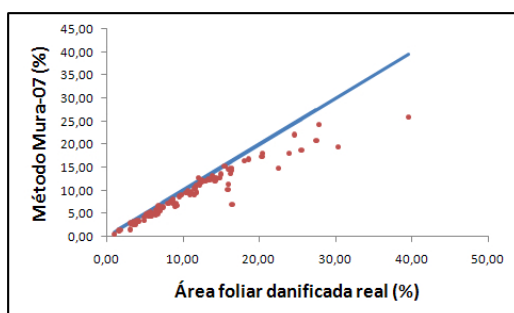
¹Uma versão preliminar deste trabalho que não contempla a remoção das sombras e reconstrução de danos nas bordas das folhas é apresentada em [10].

Tabela I
EAM E ERM DOS MÉTODOS AVALIADOS PARA CADA INTERVALO DE PORCENTAGEM DE ÁREA FOLIAR DANIFICADA.

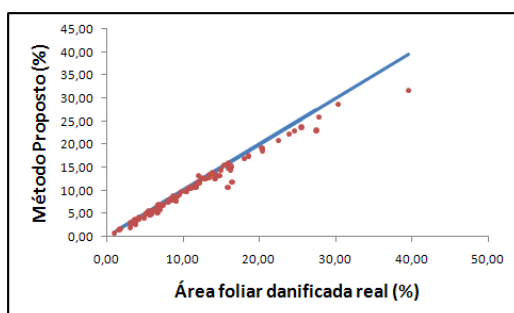
Grupo %	F	Humano		Mura-07		Proposto	
		EAM (%)	ERM (%)	EAM (%)	ERM (%)	EAM (%)	ERM (%)
0 — 5	54	2,90 ± 2,48	86,43 ± 93,43	0,53 ± 0,50	17,25 ± 20,65	0,37 ± 0,56	12,26 ± 25,19
5 — 10	40	5,54 ± 3,73	73,71 ± 46,72	1,13 ± 0,75	15,25 ± 9,17	0,65 ± 0,50	8,86 ± 6,39
10 — 15	31	6,35 ± 4,34	50,91 ± 34,36	1,93 ± 1,88	15,64 ± 15,82	1,08 ± 1,20	8,74 ± 9,88
15 — 20	13	9,61 ± 5,26	59,64 ± 33,46	3,77 ± 3,81	23,28 ± 23,87	2,05 ± 2,13	12,63 ± 13,36
20 — 99	10	14,93 ± 10,42	54,95 ± 34,36	6,32 ± 3,76	22,99 ± 10,07	2,58 ± 2,07	9,18 ± 4,86
TOTAL	148	5,74 ± 5,36	71,07 ± 65,84	1,66 ± 2,33	17,29 ± 16,94	0,89 ± 1,24	10,43 ± 16,69



(a)



(b)



(c)

Figura 1. Relação entre a área foliar danificada real e a estimada pelos métodos: (a) Humano, (b) Mura-07 e (c) Proposto.

mais da metade do EAM gerado pelo método *Mura-07* e oito vezes menor que o EAM do método *Humano*.

Os gráficos da Figura 1 estabelecem a relação entre a área foliar danificada real (referência) com os métodos avaliados. O método proposto apresentou boa estimativa de área foliar danificada, como pode ser observado pela

proximidade dos pontos à reta 1:1, na Figura 1(c). Já para o gráfico da Figura 1(b) pode-se observar que existem mais pontos afastados da reta 1:1, principalmente quando a área foliar ideal ultrapassa 20%. Note ainda que como apresenta o gráfico da Figura 1(c) os valores estimados pelos seres humanos estão quase que em sua maioria absoluta bem acima da área foliar danificada ideal. Ou seja, o ser humano tende a super avaliar a área foliar danificada.

A seguir, é apresentado um teste de significância estatística que tem como objetivo confirmar a superioridade do método *Proposto* em relação aos métodos *Mura-07* e *Humano*. Para tanto, foi aplicado o teste *t* de *Student* [16], com o nível de significância de 0,0005 tendo como hipótese alternativa, H_1 , que o EAM do método *Proposto* é menor que o EAM gerado pelos os outros dois métodos e como hipótese nula, H_0 , o contrário; isto é, que o EAM do método *Proposto* é maior ou igual ao EAM dos demais.

Foram calculados os valores do teste *t* de *Student* entre o método *Proposto* e os métodos *Humano* ($t_1 = -11,987$) e *Mura-07* ($t_2 = -6,799$). Para aceitar a hipótese nula, ao nível de significância de 0,0005, t_1 e t_2 devem ser maiores que $-3,291$. Como esta afirmação é negativa, a hipótese nula é rejeitada. Assim, podemos concluir que o método *Proposto* gera estimativas de área foliar danificada estatisticamente melhores que as dos outros dois métodos.

A Tabela II apresenta o erro máximo gerado por cada método, que foram observados na mesma amostra. Os valores mostram que o erro máximo do método *Proposto* é inferior ao gerado pelos dois outros métodos, o que significa um maior grau de confiabilidade.

Tabela II
ERRO MÁXIMO COMETIDO POR CADA MÉTODO.

Método	Valor real	Valor estimado	Erro máximo
<i>Humano</i>	39,54%	76,25%	36,71%
<i>Mura-07</i>	39,54%	25,84%	13,70%
<i>Proposto</i>	39,54%	31,64%	7,90%

Também foi verificada a capacidade de auxílio ao diagnóstico de cada método. A Tabela III mostra o número de indicações de ação de controle corretas realizadas por cada método. Os resultados mostram que o método *Proposto* apresentou uma taxa de acerto total de 95,95% $((17+125)/(23+125))$, valor superior ao método *Mura-07* e *Humano* que

apresentaram uma taxa de 92,57% $((12 + 125)/(23 + 125))$ e 75,78% $((23 + 89)/(23 + 125))$, respectivamente.

Tabela III
INDICAÇÕES CORRETAS DE CADA MÉTODO EM RELAÇÃO À
NECESSIDADE DE MEDIDAS DE CONTROLE NA LAVOURA.

Indicação	Referência	Humano	Mura-07	Proposto
Sim	23	23	12	17
Não	125	89	125	125

A alta taxa de acerto do método *Humano* para situações em que a indicação é positiva, se deve ao fato dos resultados serem superestimados como é ilustrado no gráfico da Figura 1(a).

Na Tabela IV é apresentada a matriz confusão para as imagens finais dos métodos *Mura-07* e *Proposto*. Valores altos de *TP* e *TN* significam uma semelhança grande entre as imagens analisadas. Com relação a esse critério é possível verificar que o método *Proposto* obteve um maior sucesso na segmentação e reconstrução das imagens das amostras, do que o método *Mura-07*.

Tabela IV
MATRIZ CONFUSÃO DAS IMAGENS FINAIS DOS MÉTODOS *Mura-07* E
Proposto EM RELAÇÃO AS IMAGENS DE REFERÊNCIA.

Imagem(Método)	TP	FN	FP	TN
Segmentada(<i>Mura-07</i>)	98,51%	1,49%	0,28%	99,72%
Segmentada(<i>Proposto</i>)	99,08%	0,92%	0,34%	99,66%
Reconstruída(<i>Mura-07</i>)	98,51%	1,49%	0,28%	99,21%
Reconstruída(<i>Proposto</i>)	99,60%	0,40%	0,39%	99,61%

V. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foi proposto um novo método para a detecção de área foliar danificada, que leva em consideração aspectos de remoção de ruído (*e.g.*, sombra, objetos indesejáveis, *etc.*) e reconstrução do contorno da folha usando conceitos de geometria computacional. Experimentos mostraram que o método proposto quantifica a área foliar danificada com uma precisão que é na média sete vezes maior que a avaliação humana e duas vezes maior que o método proposto por [9].

Como proposta de trabalhos futuros, pode-se aperfeiçoar este método com um algoritmo de interpolação de curvas para reconstrução das bordas ao invés da utilização de segmentos de reta. Acredita-se que com esta extensão o método teria uma maior precisão no auxílio ao diagnóstico do manejo da lavoura.

AGRADECIMENTOS

Agradecimentos à FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) pelo incentivo financeiro ao desenvolvimento deste trabalho. Os autores agradecem ao auxílio na segmentação manual das imagens de referência à Thiago Luiz Guimarães, Pedro Ribeiro

Mendes Júnior e Kayran dos Santos e também agradecem à Éder Matsuo pelo acompanhamento durante o processo da avaliação humana das amostras foliares.

REFERÊNCIAS

- [1] F. R. Gassen and D. N. Gassen, *Doenças Foliares em Soja*, 1st ed. Aldeia Norte, 2005.
- [2] J. L. Favarin, D. Dourado-Neto, A. G. García, N. A. V. Nova, and M. G. G. V. Favarin, "Equações para a estimativa do índice de área foliar do cafeeiro," *Pesquisa Agropecuária Brasileira*, vol. 37, no. 6, pp. 769–773, junho 2002.
- [3] J. E. Tavares-Jr, J. L. Favarin, D. Dourado-Neto, A. H. N. Maia, L. C. Fazuoli, and M. S. Bernardes, "Análise comparativa de métodos de estimativa de área foliar em cafeeiro," *Bragantia*, vol. 61, no. 2, pp. 199–203, maio/agosto 2002.
- [4] C. M. Borkert, J. T. Yorinori, B. S. Corrêa-Ferreira, Á. M. R. Almeida, L. P. Ferreira, and G. J. Sfredo, "Seja o doutor da sua soja - pragas: diagnóstico e controle," *Informações Agrônomicas*, no. 66, pp. 7–8, junho 1994.
- [5] C. F. Wilcken, R. C. B. Moraes, M. L. Haddad, and J. R. P. Parra, "Técnica de medição de área foliar consumida por insetos através de digitalizador de imagens," *Scientia Agrícola*, vol. 55, no. 2, pp. 218–221, maio/agosto 1998.
- [6] S. Bianco, R. A. Pitelli, and L. B. Carvalho, "Estimativa da área foliar de tridax procumbens usando dimensões lineares do limbo foliar," *Planta Daninha*, vol. 22, no. 2, pp. 247–250, abril/junho 2004.
- [7] C. A. A. Gonçalves, N. N. J. Chalfun, M. A. Regina, A. A. Alvarenga, M. T. Souza, and E. Abrahão, "Estimativa de Área foliar da videira sobre diferentes porta-enxertos," *Ciência e Agrotecnologia*, vol. 26, no. 3, pp. 500–504, maio/junho 2002.
- [8] P. M. Einhardt and O. G. L. Ferreira, "Validação de um método digital para determinação de área foliar," in *XVII Congresso de Iniciação Científica e X Encontro de Pós-Graduação*, Pelotas, novembro 2008, pp. 1–4.
- [9] W. D. Mura, A. L. Oliveira, E. M. Sgarbi, and L. G. Sachs, "Detecção automática da área foliar da soja danificada pela lagarta utilizando processamento digital de imagens," in *WUW-SIBGRABI 2007*, 2007, pp. 1–4.
- [10] A. C. Nazaré-Jr, D. Menotti, T. Sedyama, and J. M. R. Neves, "Detecção da área foliar danificada de soja," in *7º Congresso Brasileiro de Agroinformática - SBI Agro*, 2009, aceito para publicação.
- [11] H. Pedrini and W. R. Schwartz, *Análise de Imagens Digitais*, 1st ed. São Paulo: Thomson, 2008.
- [12] W. Shugen, G. Zejin, and L. Deren, "Shadow detection and compensation for color aerial images," *Geo-Spatial Information Science*, vol. 6, no. 3, pp. 20–24, setembro 2003.
- [13] R. C. G. e R. E. Woods, *Digital Image Processing*, 3rd ed. Prentice Hall, 2008.

- [14] X. Chen-He and N. H. C. Yung, "Corner detector based on global and local curvature properties," *Optical Engineering*, vol. 47, no. 5, pp. 057 008 1–12, maio 2008.
- [15] E. Azevedo and A. Conci, *Computação Gráfica: Teoria e Prática*, 1st ed. Rio de Janeiro: Campus, 2003.
- [16] M. F. Triola, *Introdução à Estatística*, 7th ed. Rio de Janeiro: LTC, 1999.