KLEBER MARIANO RIBEIRO

ABORDAGENS ESTATÍSTICAS NA ANÁLISE DO SPECKLE LASER DINÂMICO

LAVRAS - MG

2014

KLEBER MARIANO RIBEIRO

ABORDAGENS ESTATÍSTICAS NA ANÁLISE DO SPECKLE LASER DINÂMICO

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, área de concentração em Instrumentação, para a obtenção do título de Doutor.

Orientador

Dr. Roberto Alves Braga Júnior

Coorientadores

Dra. Thelma Sáfadi

Dr. Danton Diego Ferreira

Dr. Graham William Horgan

LAVRAS - MG

2014

Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e Serviços da Biblioteca Universitária da UFLA

Ribeiro, Kleber Mariano.
Abordagens estatísticas na análise do speckle laser dinâmico / Kleber Mariano Ribeiro. – Lavras: UFLA, 2014.
122 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2014.
Orientador: Roberto Alves Braga Júnior.
Bibliografia.

1. Biospeckle laser. 2. Fourier. 3. Wavelets (Matemática). 4. PCA. 5. ICA. I. Universidade Federal de Lavras. II. Título.

CDD – 621.367

KLEBER MARIANO RIBEIRO

ABORDAGENS ESTATÍSTICAS NA ANÁLISE DO SPECKLE LASER DINÂMICO

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, área de concentração em Instrumentação, para a obtenção do título de Doutor.

APROVADA em 28 de janeiro de 2014.

Dr. Roberto Alves Braga Júnior	UFLA
Dr. Giovanni Francisco Rabelo	UFLA
Dra. Thelma Sáfadi	UFLA
Dra. Ellem Waleska Nascimento da Fonseca Contado	UFLA
Dr. Adilson Machado Enes	UFS

Dr. Roberto Alves Braga Júnior

Orientador

LAVRAS - MG

2014

A Luiz, meu pai, e a Ivone, minha mãe,
pessoas maravilhosas e exemplos para mim.

A Kátia e Kássio, meus irmãos, pelo companheirismo e carinho.

A Daniele e Camila, minhas sobrinhas, pelas brincadeiras e diversões.

E a minha namorada, Ana Carolina, pelo incentivo
e companheirismo durante esta caminhada.

DEDICO.

AGRADECIMENTOS

A Deus, por tudo o que sou e tudo o que tenho, e por estar sempre presente em cada passo dessa caminhada.

Ao meu pai, Luiz Franz Ribeiro, pelo apoio, carinho e conselhos.

À minha mãe, Ivone Aparecida Ribeiro, pela dedicação, amor, pela compreensão e incentivo nos momentos difíceis.

Aos meus irmãos, Kátia Daniela e Kássio Mariano, pela convivência, apoio e amizade.

À minha companheira, Ana Carolina Lopes, que esteve sempre do meu lado, paciente, compreensiva, desejando o melhor para mim.

Às minhas sobrinhas, Daniele Ribeiro e Camila Ribeiro, pelas brincadeiras e momentos de descontração.

A todos os meus familiares, pelo exemplo de luta e vida.

A todos os meus colegas de Carmo da Cachoeira, pela amizade, convívio, prontidão e atenção.

À Universidade Federal de Lavras (UFLA), em especial ao Programa de Pós-Graduação em Engenharia Agrícola, pela oportunidade de cursar o doutorado e ao CNPq, pela concessão da bolsa, que tornou possível a realização deste trabalho.

Ao professor Roberto Braga, pelo apoio, confiança e incentivo.

Ao pesquisador Graham Horgan e aos professores Thelma Sáfadi e Danton Diego, pelas colaborações, sugestões e auxílios nos artigos desenvolvidos.

A Lionel Dupuy e Laurence Ducreux, que tão bem me receberam em Dundee, e também aos amigos Dimitris, Michel, Cris, Antônio, Eva e Laura, pela ajuda e momentos de descontração nos quatro meses de trabalho no Instituto James Hutton.

Ao Sr. Antônio, Juninho, Diego, Éberson, Bianca, Baiano, Ísis e demais amigos da pós-graduação e da graduação da UFLA, pela ajuda nos trabalhos desenvolvidos, paciência, momentos de descontração, etc. Valeu!

A Greice e a Hellem, secretárias da Pós-Graduação em Engenharia Agrícola, pela ajuda e paciência.

Agradeço a todos os professores e técnicos do Programa de Pós-Graduação em Engenharia Agrícola, que me proporcionaram conhecimento no decorrer desse período.

Enfim, a todos aqueles que colaboraram, de alguma forma, para o cumprimento desta etapa da minha vida.

Obrigado!!!

O trabalho não é a sua vida, apenas uma parte dela.

(Autor desconhecido)

RESUMO

A técnica do *speckle* laser dinâmico processa os padrões de interferência óptica formados quando uma luz coerente incide sobre uma superfície com rugosidade igual ou superior ao comprimento de onda da luz incidente. Quando a superfície em questão tem origem biológica, a técnica tem sido chamada simplesmente de *biospeckle* laser e tem sido utilizada cientificamente para a quantificação da atividade biológica de materiais em diferentes áreas do conhecimento. O elevado número de situações em que vêm sendo aplicado, associado à complexidade do sinal *speckle* laser, demanda técnicas de análise de sinais que possam auxiliar a extração das informações, o que pode ser realizado por meio de técnicas gráficas, como os métodos de Fujii e de diferenças generalizadas (DG), ou por meio de interpretações numéricas, como, por exemplo, a técnica da diferença dos valores absolutos (AVD). Ainda assim, a análise dos sinais no domínio da frequência aparece como uma alternativa que pode fornecer informações adicionais sobre os padrões de interferência do *biospeckle* laser. Neste contexto, as transformadas de Fourier e de wavelets têm sido as ferramentas mais utilizadas para a análise espectral do *speckle* laser dinâmico e, embora haja trabalhos na literatura com resultados favoráveis utilizando ambas as técnicas de análise espectral, não há relatos indicando qual ferramenta é a mais indicada para os dados do *biospeckle* laser. Por outro lado, as características de evolução no tempo e aleatoriedade dos dados do *speckle* laser dinâmico favorecem a utilização de ferramentas estatísticas, tal como a análise dos componentes principais (PCA) e a análise dos componentes independentes (ICA), as quais apresentam a vantagem das diferentes vertentes em que podem ser aplicadas e por se adequarem às características dos sinais em estudo. Dessa forma, o presente trabalho foi realizado com o objetivo de avaliar as transformadas de Fourier e wavelet, de forma a indicar a ferramenta mais apropriada para a análise espectral do *biospeckle* laser, propor a filtragem dos dados do *speckle* laser dinâmico por meio da técnica estatística PCA e, por fim, o pré-processamento das saídas gráficas de análise do biospeckle laser via ICA,visando à melhoria da qualidade visual das imagens digitais finais. O trabalho está dividido em três partes, sendo a primeira uma revisão de literatura sobre os principais temas abordados na tese; na segunda seção apresentam-se três artigos científicos desenvolvidos para atender aos objetivos propostos e, na última divisão, estão as considerações finais.

Palavras-chave: *Biospeckle* laser. Fourier. Wavelet. PCA. ICA.

ABSTRACT

The technique of the dynamic laser speckle processes the optics interference patterns formed when a coherent light focuses on a surface with roughness equal or higher than the wavelength of the incident light. When the surface in question is biological, the technique has been called simply of biospeckle laser and it has been used scientifically to quantify the biological activity of materials in different knowledge areas. The high number of situations in which has been applied, associated to the complexity of the speckle signal, demand by signals analysis techniques that can assist the extraction of information, which can be accomplished by graphical methods, such as Fujii and generalized differences (GD), or through numerical interpretations, for example, the technique of the absolute values of the differences (AVD). Even so, the signal analysis in the frequency domain is an alternative that can provide additional information about the interference patterns of the biospeckle laser. In this context, Fourier and wavelet transforms have been the most used tools in the spectral analysis of the dynamic laser speckle and, although there are reports in the literature with favorable results using both techniques of spectral analysis, there are no reports indicating which tool is most indicated for the biospeckle laser data. Moreover, the characteristics of time evolution and randomness of the dynamic laser speckle encourage the use of statistical tools, such as principal component analysis (PCA) and the independent component analysis (ICA), which have as advantage the different aspects that can be applied and fit the characteristics of the signals under study. Thus, the present study aimed to evaluate the Fourier and wavelet transforms to indicate the most appropriate tool for the spectral analysis of the biospeckle laser, to propose the filtering of the dynamic laser speckle data using the statistics technique PCA and, finally, pre-processing the graphical output of biospeckle laser analysis via ICA, to improve the visual quality of final digital images. The work was divided into three parts, in which the first is a literature review of the main topics discussed in the thesis; the second section presents three scientific papers developed to perform the proposed approaches and, the last division, the final considerations of the thesis.

Keywords: Biospeckle laser. Fourier. Wavelet. PCA. ICA.

LISTA DE FIGURAS

PRIMEIRA PARTE

SEGUNDA PARTE

ARTIGO 1

ARTIGO 2

ARTIGO 3

LISTA DE TABELAS

SEGUNDA PARTE

ARTIGO 2

ARTIGO 3

SUMÁRIO

PRIMEIRA PARTE

1 INTRODUÇÃO GERAL

A técnica óptica do *speckle* laser dinâmico ou do *biospeckle* laser analisa padrões de interferência formados quando uma luz coerente se espalha sobre uma superfície rugosa. Consiste de uma ferramenta não destrutiva que vem sendo utilizada em processos industriais, na medicina, na agropecuária e em outras áreas do conhecimento, como instrumento de quantificação e monitoramento da atividade biológica do material em estudo.

A atividade biológica consiste de um conjunto de processos relacionados ao movimento de partículas do objeto estudado e que pode ser interpretado como o somatório de movimentos moleculares e estruturais do material em análise, bem como pelo movimento browniano dos líquidos presentes nos materiais e pelas variações dos índices de refração na região de iluminação, contribuindo para o batimento Doppler que é observado pelas câmeras.

A análise do *biospeckle* laser é realizada pela correlação dos padrões de interferência óptica no tempo, o que resulta em saídas gráficas ou numéricas. Nas análises gráficas, processamentos são realizados nas imagens *speckle* e o resultado são mapas com a variabilidade espacial da atividade biológica; por outro lado, as interpretações numéricas atribuem valores às atividades biológicas, o que elimina a subjetividade dos métodos gráficos.

Todavia, a adoção de outras técnicas que possam auxiliar e melhorar a análise e a interpretação dos dados, de forma a oferecer informações adicionais sobre os padrões de interferência óptica que compõem o *biospeckle* laser, é bem-vinda e a análise dos dados no domínio da frequência tem sido uma interessante opção.

Todavia, a adoção de outras técnicas que possam auxiliar e melhorar a análise e interpretação dos dados, de forma a oferecer informações adicionais sobre os padrões de interferência óptica que compõe o *biospeckle* laser é bem vinda, e a análise dos dados no domínio da frequência tem sido uma interessante opção.

Nesse sentido, diversos trabalhos vêm sendo conduzidos utilizando a análise espectral do sinal do *biospeckle* laser, de forma complementar a classificação gráfica e numérica, e, dentre as ferramentas disponíveis para as análises, as transformadas de Fourier e de wavelets são as mais utilizadas.

A transformada de Fourier apresenta o conteúdo espectral do sinal sem o intervalo de tempo em que as componentes espectrais ocorreram, enquanto a técnica de wavelets permite analisar os dados em diferentes escalas e em cada instante.

Embora ambas as técnicas venham apresentando resultados satisfatórios na análise do sinal do *biospeckle* laser, não há trabalhos que investiguem qual delas é a mais apropriada e se a transformada de Fourier, que é mais simples, do ponto de vista computacional, é suficiente para a análise espectral do *speckle* laser dinâmico.

Por outro lado, a natureza aleatória e a evolução no tempo dos dados do *biospeckle* laser favorecem a utilização de ferramentas estatísticas na análise dos dados, o que abre espaço para as técnicas de análise dos componentes principais (PCA) e para a análise dos componentes independentes (ICA).

PCA e ICA são técnicas estatísticas multivariadas clássicas que transformam os dados iniciais em um novo conjunto de observações descorrelacionados e estatisticamente independentes, respectivamente.

Ambas as ferramentas têm chamado a atenção da comunidade acadêmica pelo número de abordagens em que podem ser aplicadas, seja na redução do volume de dados com o mínimo de perdas, na extração, na

identificação e na classificação de características, bem como na filtragem de dados e como técnica de pré-processamento para a melhoria da qualidade visual de imagens.

Partindo do exposto, este trabalho foi realizado com a finalidade de aplicar técnicas de análise no domínio da frequência e de ferramentas estatísticas multivariadas em dados do *biospeckle* laser. De forma específica, objetivou-se:

a) avaliar as ferramentas de análise no domínio da frequência, Fourier e wavelet, nos dados do *biospeckle* laser;
b) propor a filtragem alternativa dos dados do *speckle* laser dinâmico, por meio da análise dos componentes principais;
c) explorar a análise dos componentes independentes no pré-processamento dos métodos gráficos de análise do *biospeckle*, buscando melhoria na qualidade visual das imagens finais.

O trabalho foi dividido em três partes. A primeira parte consiste de uma revisão de literatura em que se abordam os principais temas discutidos na tese. Já na segunda seção apresentam-se, de forma estruturada, três artigos científicos que procuraram responder aos objetivos específicos descritos anteriormente e, por último, na terceira divisão do trabalho, estão as considerações finais.

2 REFERENCIAL TEÓRICO

2.1 *Speckle* laser dinâmico

O *speckle* laser é um fenômeno de interferência óptica que ocorre quando um material é iluminado por uma luz coerente que se dispersa sobre uma superfície qualquer (RABAL; BRAGA, 2008). Quando aplicado a superfícies dinâmicas, observa-se uma contínua formação de novos e diferentes padrões de interferência óptica e esse padrão aleatório e dinâmico de interferência denomina-se *speckle* laser dinâmico ou *biospeckle* laser, se a superfície em questão for de origem biológica (RABAL; BRAGA, 2008; ZDUNEK et al., 2013).

A análise desses padrões de interferência no tempo vem sendo validada como instrumento de quantificação e monitoramento da atividade biológica na agropecuária e trabalhos, como os apresentados por Carvalho et al. (2009), na avaliação dos parâmetros de motilidade de sêmem congelado bovino; por Kurenda et al. (2012), na análise do efeito de diferentes níveis de temperatura na atividade biológica de maçãs e por Zdunek e Herppich (2012), no estudo da relação entre pigmentos de clorofila presentes em maçãs e sua respectiva atividade biológica são alguns exemplos recentes de aplicação do *biospeckle* laser.

A atividade biológica expressa no contexto do *speckle* laser dinâmico pode ser entendida como a combinação de movimentos moleculares e estruturais ocorridos no material em análise, variações do índice de refração na área iluminada, movimento browniano nos líquidos presentes nas amostras e efeito Doppler observado nas câmeras (ZDUNEK et al., 2013), dentre outros.

A observação direta dos padrões de interferência óptica não permite sua quantificação e, por ser um fenômeno dinâmico, o *biospeckle* laser deve ser analisado com técnicas de processamento de imagem e tratamento estatístico (RABAL; BRAGA, 2008), o que pode ser realizado por meio de técnicas gráficas ou numéricas.

Os métodos gráficos são processamentos das imagens digitais *speckle* que geram mapas indicando a variabilidade espacial da atividade biológica, à medida que as interpretações numéricas conferem valores para a atividade biológica.

2.1.1 Técnicas gráficas de análise do *biospeckle* laser

Há disponíveis, na literatura, diversos métodos gráficos para a análise dos dados do *biospeckle* laser e as técnicas de Fujii e de diferenças generalizadas vêm sendo as mais utilizadas nos trabalhos científicos atuais, merecendo destaque.

Método de Fujii

O método de Fujii (FUJII et al., 1987) tem como princípio trabalhar as imagens obtidas de um corpo iluminado, identificando a intensidade luminosa de cada pixel que a compõe. A técnica consiste no somatório das diferenças de intensidades luminosas entre uma imagem e outra, sobre a soma das intensidades entre uma imagem e sua subsequente (fator de ponderação), conforme descrito matematicamente na Equação 1.

$$\mathbf{Fujii}\ (x,y) = \sum_{k=1}^{N} \left| \frac{I_k(x,y) - I_{k+1}(x,y)}{I_k(x,y) + I_{k+1}(x,y)} \right| \quad (1)$$

em que $I_k(x,y)$ é a intensidade do pixel de coordenadas; (x,y) da $k^{ésima}$ imagem.

O resultado será uma nova imagem, em que as regiões de alta atividade são representadas em tons de cinza-claro e as áreas escuras ilustram regiões de baixas atividades.

A amplificação dos movimentos em áreas mais escuras é uma característica do método de Fujii, resultando em imagens mais claras de forma geral, quando comparadas com as do método de diferenças generalizadas (DG) (BRAGA et al., 2009).

Método de diferenças generalizadas

A técnica gráfica de diferenças generalizadas, por sua vez, foi apresentada por Arizaga et al. (1999) como uma alternativa ao método de Fujii. A expressão matemática (Equação 2) que descreve o método de diferenças generalizadas não apresenta o fator de ponderação no denominador e as diferenças entre as intensidades dos pixels foram generalizadas a todo o conjunto de imagens capturadas.

$$DG(x,y) = \sum_{k} \left[\sum_{l} \left[\left| I_k(x,y) - I_{k+1}(x,y) \right| \right] \right]$$

(2)

em que DG (x,y) é a imagem resultante da aplicação da técnica sobre um conjunto de imagens; I(x,y) corresponde à intensidade luminosa presente na coordenada x e y de cada imagem k.

2.1.2 Análises numéricas do *biospeckle* laser

As análises numéricas vieram para suprir as interpretações subjetivas dos métodos gráficos. A medição dos níveis de atividade abre espaço para a utilização do *speckle* laser dinâmico como ferramenta de metrologia e complementa as informações visuais advindas das técnicas gráficas, melhorando a qualidade das análises.

A técnica numérica adotada no presente trabalho para a análise do *biospeckle* laser foi a diferença dos valores absolutos (AVD), cujo cálculo tem início com a história temporal dos padrões *speckle*.

História temporal do padrão de *speckle*

A história temporal do padrão de *speckle* (*time history speckle patterns*, THSP) foi proposta por Oulomara, Tribillon e Duvernoy (1989) e avaliada por Xu, Joenathan e Khorana (1995) e consiste de uma imagem que oferece informações sobre a evolução temporal dos padrões de *speckle*.

A técnica consiste em coletar uma mesma linha na imagem *speckle* em instantes sucessivos e organizá-los verticalmente, lado a lado, em uma imagem bidimensional intermediária. As imagens THSP são matrizes m x n com pixels em níveis de cinza, em que o eixo das abscissas apresenta informações sobre a evolução temporal dos pixels selecionados e, no eixo das ordenadas, há o registro da distribuição espacial dos padrões de interferência.

Com a história temporal do padrão de *speckle* construída, o próximo passo para o cálculo dos valores AVD é caracterizar a matriz de coocorrência.

Matriz de coocorrência

A matriz de coocorrência (MCO) na análise numérica do *biospeckle* laser é utilizada como passo intermediário no cálculo dos valores AVD e

expressa o número de transições de cada pixel da imagem THSP com respeito ao seu vizinho imediato (ARIZAGA et al., 1999).

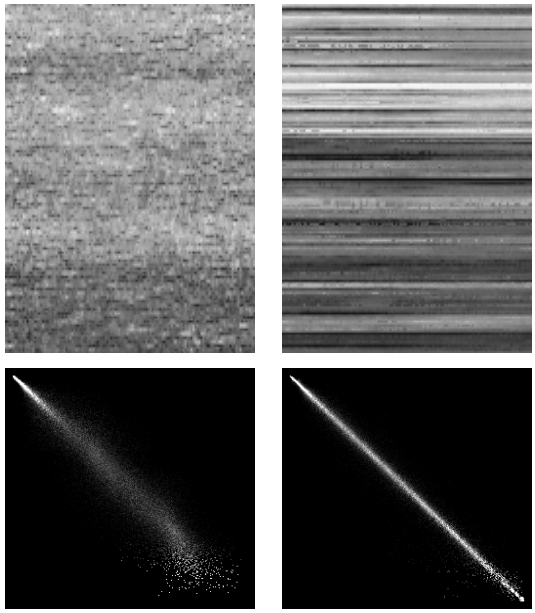
A MCO é uma técnica muito utilizada em processamento de imagens para caracterizar a textura de imagens digitais (PARTIO; CRAMARIUC; GABBOUJ, 2007) e é descrita matematicamente pela Equação 3.

$$\mathbf{MCO} = [\, N_{ij} \,] \tag{3}$$

em que MCO consiste na matriz de coocorrência; i e j são as intensidades sucessivas e N_{ij} corresponde ao número de ocorrências de um valor de intensidade i, seguido por um valor de intensidade j, ao se deslocar pelas linhas ou colunas da história temporal.

Se um material apresentar baixa atividade, o THSP praticamente não apresenta variações de intensidade ao longo do tempo e sua matriz de coocorrência é caracterizada por pixels que apresentam tons de cinza com pouca variação, ou seja, a mudança de intensidade do pixel i para o pixel j será pequena.

Entretanto, se o material a ser analisado apresentar alta atividade, será possível observar THSP com grande atividade e a matriz de coocorrência apresenta os elementos não nulos próximos da diagonal principal. Na Figura 5 apresenta-se, graficamente, a THSP e suas respectivas matrizes de coocorrência para materiais com alta e baixa atividade biológica.



(A) (B)

Figura 1 História temporal e matriz de coocorrência. (A) Material com alta atividade e (B) material com baixa atividade
Fonte: Silva et al. (2011)

A matriz de coocorrências é modificada para a quantificação das variações de intensidades do *speckle*. A modificação é realizada dividindo-se os números de transições de cada linha da MCO pelo somatório dos valores de ocorrência de cada linha (RABAL; BRAGA, 2008), conforme Equação 4.

$$\mathbf{M_{ij}} = \frac{\mathbf{N_{ij}}}{\sum \mathbf{N_{ij}}} \tag{4}$$

Dessa forma, o somatório das componentes em cada linha da matriz de coocorrência modificada é igual a 1.

Diferença dos valores absolutos

A partir da matriz de coocorrência modificada, Braga et al. (2011) propuseram um método para quantificar a atividade biológica, denominado diferença dos valores absolutos (AVD), que indica com que frequência ocorreram mudanças bruscas de intensidade na matriz THSP. A Equação 5 descreve matematicamente a diferença dos valores absolutos.

$$\mathbf{AVD} = \sum_{\mathbf{ij}} \mathbf{M_{ij}}\,(\mathbf{i} - \mathbf{j}) \tag{5}$$

Esta técnica exibe altos valores quando a amostra apresentar alta atividade e baixos valores em situação contrária. Este cálculo é uma ferramenta importante para estimar a atividade global em diversas aplicações biológicas e não biológicas.

A combinação da análise no domínio da frequência com métodos gráficos e numéricos tem sido uma alternativa para a análise dos dados do *biospeckle* laser, permitindo abordagens que proporcionam meios para o isolamento e a marcação de frequência de vários fenômenos observados.

2.2 Análises no domínio da frequência

Filtragem dos dados, melhor contraste das imagens *speckle* e definição de marcadores de frequência associados a fenômenos biológicos são algumas áreas de trabalho referidas com a análise dos dados do *biospeckle* laser no domínio da frequência. As transformadas de Fourier e wavelets têm sido as ferramentas mais utilizadas para a análise espectral desses dados.

2.2.1 Transformada de Fourier

A transformada de Fourier pode ser entendida como uma ferramenta matemática que transforma um conjunto de dados do domínio do tempo para o domínio da frequência por meio de funções senos e cossenos de período 2π (MORETTIN, 1999). A Equação 6 descreve matematicamente a transformada de Fourier.

$$f(\omega) = \int_{-\infty}^{\infty} f(t) \cdot e^{-i\omega t}\, dt \qquad (6)$$

em que

$\omega = 2 \pi f$

$| f(\omega) |$ = amplitude de cada componente ω do sinal.

Uma forma útil de entender a transformada de Fourier é pensar que o sinal f(t) é projetado em um conjunto de funções senos e cossenos representadas pelas exponenciais complexas (Equação 6).

Nas ciências agrárias, a transformada de Fourier apresenta inúmeras aplicações. Por exemplo, Ferrão et al. (2003) propuseram uma nova metodologia para a quantificação do tanino em café cru, por meio da espectroscopia por reflexão difusa no infravermelho; Dick et al. (2008) utilizaram a teoria de Fourier para avaliar o efeito residual de queimadas periódicas nos atributos químicos e na matéria orgânica de um Latossolo Vermelho e Pontelli et al. (2010) a empregaram como ferramenta na avaliação de dois tipos de suspensão passiva de barra em condições simuladas de pista, além de vários outros trabalhos.

Existe também a transformada inversa de Fourier, técnica que retorna os dados do domínio da frequência para o domínio do tempo e que é representada pela Equação 7.

$$f(t) = \int_{-\infty}^{\infty} f(\omega) \cdot e^{-i\omega\tau} \, d\omega \tag{7}$$

A transformada de Fourier apresenta o conteúdo espectral do sinal sem fornecer o intervalo de tempo em que as componentes espectrais aparecem e são indicadas para a análise de sinais estacionários. Entende-se por séries estacionárias séries cujos momentos estatísticos, como a média, a variância e outros, não variam ao longo do tempo.

Dessa forma, a análise de sinais não estacionários, que corresponde à maioria dos casos presentes na natureza, ou em situações em que o instante em que as frequências ocorreram é importante, inviabiliza ou limita a aplicação da transformada de Fourier (SIFUZZAMAN; ISLAM; ALI, 2009), e abre espaço para a adoção da transformada de wavelet para a análise espectral dos dados.

2.2.2 Transformada de wavelet

A transformada de wavelet é uma ferramenta matemática que tem despertado o interesse da comunidade acadêmica pelos resultados satisfatórios que tem apresentado na análise de problemas em diversas áreas do conhecimento, como, por exemplo, no processamento de sinais, em áreas de saúde, hidrologia, geofísica espacial e outras (BOLZAN, 2006; LIU et al., 2010).

Definida como uma decomposição multirresolução para a análise de sinais e imagens, a transformada de wavelet caracteriza os dados por meio da

energia de cada escala e translação (XU et al., 1994), conforme descrito matematicamente na Equação 8.

$$< f, \psi_{a,b} > = \int_{-\infty}^{\infty} f(t)\psi \left(\frac{t-b}{a}\right) dt \quad (8)$$

em que < f, $\psi_{a,b}$> corresponde ao espectro de wavelet, f(t) é o sinal estudado no domínio do tempo, a variável *a* representa a escala e *b*, a translação da função wavelet mãe $\psi_{a,b}(t)$.

A transformada de wavelet é matematicamente descrita como uma convolução entre o sinal de interesse e a função wavelet mãe na escala *a* e translação *b*.

A escala é a compressão ou a dilatação de uma função e está relacionada à frequência em que altas escalas equivalem a baixas frequências e as baixas escalas equivalem a altas frequências. Já a translação corresponde ao deslocamento de uma função $\psi(t)$ por um valor k, o que é representado matematicamente por $\psi(t - k)$ (KARIMI; PAWLUS; ROBBERSMYR, 2012).

Existem, disponíveis na literatura, diversas funções mãe de wavelet, cada qual com suas particularidades, e a função mãe de Morlet foi a escolhida para as análises dos dados, no presente trabalho, pela próxima relação entre sua escala e a frequência de Fourier, como discutido por Polansky et al. (2010).

Wavelet de Morlet

Apresentada por Grossmann e Morlet (1984), em conjunto com pesquisadores da equipe de Alex Grossman, do Centro de Física Teórica de Marseille, na França, a wavelet mãe de Morlet é um complexo de amortecimento

exponencial com um conjunto de parâmetros de oscilação conforme descrito matematicamente na Equação 9 e ilustrado graficamente na Figura 2.

$$\psi_o(\eta) = \pi^{-0,25} \times e^{-i\eta w_o} \times e^{\frac{-\eta^2}{2}} \quad (9)$$

em que w_o é frequência (adimensional) com um valor que satisfaça à condição de admissibilidade.

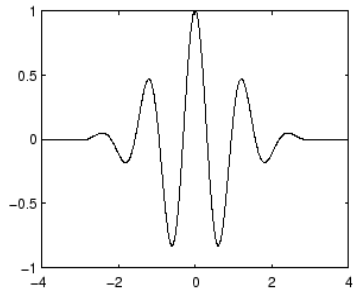


Figura 2 A wavelet mãe de Morlet
Fonte: Liu et al. (2010)

É possível observar, na Figura 2, a regularidade e a simetria da função mãe de Morlet, duas importantes características dessa ondaleta.

O regresso dos dados do domínio da frequência para o domínio do tempo também é possível por meio da transformada inversa de wavelet e a Equação 10 descreve o processo de reconstrução dos dados para a onda mãe de Morlet (TORRENCE; COMPO, 1998).

(10)

em que $s_j^{0,50}$ é um fator que converte a transformada de wavelet em densidade de energia e δ_j, $\delta_t^{0,50}$, C_δ e $\psi_0(0)$ são constantes específicas da função base de Morlet.

Por outro lado, a característica de evolução no tempo e a aleatoriedade dos dados do *speckle* laser dinâmico abrem espaço para a adoção de técnicas estatísticas na análise dos dados (RABAL et al., 2012), buscando facilitar as interpretações.

2.3 Técnicas estatísticas

Dentre as diversas ferramentas estatísticas disponíveis na literatura, a análise de componentes principais (PCA) e a análise de componentes independentes (ICA) têm chamado a atenção da comunidade acadêmica, pelas diferentes abordagens em que podem ser utilizadas na análise de sinais.

2.3.1 Análise de componentes principais

A análise de componentes principais é uma técnica clássica de análise estatística multivariada dos dados que consiste em transformar um conjunto de variáveis originais correlacionadas em outro conjunto de variáveis sintéticas descorrelacionadas, as chamadas componentes principais (HYVÄRINEN; KARHUNEN; OJA, 2001).

A transformação dos dados para o domínio PCA é realizada por meio da decomposição da matriz de covariância em autovalores e autovetores, e as componentes principais calculadas correspondem a uma combinação linear de

todas as variáveis originais e ortogonais entre si, o que não permite a redundância de informações (STONE, 2004).

O ponto de partida da análise de componentes principais é organizar os dados em uma matriz **X** de dimensões *M* x *N*, conforme descrito na Equação 11.

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots & \mathbf{x}_{1N} \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots & \mathbf{x}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{M1} & \mathbf{x}_{M2} & \cdots & \mathbf{x}_{MN} \end{bmatrix} \quad (11)$$

em que *M* representa o número de observações e *N*, as variáveis.

Ao estudar dados de diferentes unidades, é comum que os pontos mais distantes do centro dos dados tenham maior influência que os pontos mais próximos e, para evitar que essa influência ofusque a análise, a média dos dados é extraída utilizando-se a Equação 12.

$$\mathbf{y}_N = \mathbf{x}_N - E(\mathbf{x}_N) \quad (12)$$

em que $\mathbf{y}_N$ são os dados centralizados em torno da média, $\mathbf{x}_N$ são as observações referentes à variável *N* (Equação 11) e $E(\mathbf{x}_N)$ é a esperança estatística do vetor amostra $\mathbf{x}_N$.

Entende-se por variáveis ou vetores amostra cada coluna da matriz **X** ilustrada na Equação 11 e, seguindo os procedimentos da análise PCA, tem-se que o próximo passo é o cálculo da matriz de covariância, que é igual à esperança do produto da matriz de dados centralizados em torno da média (**Y**) e sua transposta (Equação 13).

$$C_Y = E(Y \cdot Y^T) \quad (13)$$

em que $\mathbf{C_y}$ é a matriz de covariância, **Y** e $\mathbf{Y^T}$ são, respectivamente, a matriz de dados centralizados em torno da média e sua matriz transposta e E corresponde à esperança estatística.

Os elementos da diagonal principal de $\mathbf{C_y}$ correspondem às variâncias dos vetores amostras e os elementos fora da diagonal principal representam a covariância entre as variáveis. Valores nulos para a covariância significam que as variáveis aleatórias em estudo são descorrelacionadas (HADFIELD, 2010).

A matriz de covariância é real e simétrica e essas características permitem sua decomposição em um conjunto de autovalores e autovetores ortogonais (JUNG; SEN; MARRON, 2012), conforme Equação 14.

$$(C_y - \lambda . I) . V = 0 \quad (14)$$

em que **λ** são os autovalores, **V** é a matriz de autovetores e **I**, a matriz identidade.

Os autovetores, também denominados de *loadings*, na terminologia PCA, representam a contribuição com que cada um dos eixos originais entra na composição dos novos eixos, as componentes principais. Já os autovalores expressam a quantidade de variância original descrita pelos respectivos autovetores (SILVA; MINIM; RIBEIRO, 2005).

O produto entre a matriz de autovetores (**V**) e a matriz de dados centralizados na média (**Y)** (Equação 15) traz como resultado um novo conjunto de dados descorrelacionados, as chamadas componentes principais ou escores (SILVA; MINIM; RIBEIRO, 2005), encerrando a sequência de cálculos da PCA.

$$\mathbf{CP} = \mathbf{V}^{\mathbf{T}} . \mathbf{Y} \tag{15}$$

em que **CP** são as componentes principais.

Conforme descrito por Cheriyadat e Bruce (2003), Jung, Sen e Marron (2012) e Zhang et al. (2010), a análise dos dados no domínio PCA permite identificar, extrair e classificar características do sinal, reduzir o volume de dados com menor perda possível de informação, bem como filtrar sinais. Essa versatilidade de aplicação da PCA, aliada à sua baixa complexidade computacional, tem despertado o interesse da comunidade acadêmica em diferentes áreas do conhecimento.

Neste contexto, trabalhos como os conduzidos por Silva, Minim e Ribeiro (2005), na análise sensorial de diferentes marcas comerciais de café orgânico; por Mélem Júnior et al. (2008), na avaliação de resultados analíticos de fertilidade de solos do Amapá e por Salgado e Nääs (2010), no estudo do risco de produção de frango de corte no estado de São Paulo, são exemplos da aplicação da PCA na agropecuária.

A reconstrução dos dados originais também é pertinente e permite visualizar a contribuição de cada uma das componentes principais para o conjunto de observações iniciais, possibilitando a compactação de dados e a eliminação de informações indesejadas. A Equação16 descreve matematicamente a transformada inversa de PCA.

$$\mathbf{X} = (\mathbf{V} . \mathbf{CP}) + \mu(X) \tag{16}$$

A ferramenta PCA busca a descorrelação dos dados originais, entretanto, pode-se optar por uma propriedade mais restritiva, tal como a independência estatística, o que abre espaço para a aplicação da técnica análise de componentes independentes para o processamento das observações iniciais.

2.3.2 Análise de componentes independentes

Se duas variáveis aleatórias *a* e *b* são estatisticamente independentes, as observações contidas em uma das variáveis não fornecem nenhuma informação sobre a outra (MORETO, 2008), e essa afirmativa não é válida para variáveis descorrelacionadas.

Duas variáveis descorrelacionadas linearmente, por exemplo, significa que, utilizando um modelo linear, não se podem combinar esses sinais. Todavia, nada impede uma relação entre essas variáveis por meio de modelos matemáticos de ordem superior, pensamento comum no impulso das interpretações.

Dessa forma, a independência estatística é uma propriedade mais restrita que a correlação. Duas variáveis aleatórias estatisticamente independentes são, necessariamente, duas variáveis descorrelacionadas; entretanto, se duas variáveis são não correlacionadas, nada se pode afirmar quanto à sua independência estatística (MORETO, 2008).

Neste contexto, a análise dos componentes independentes (ICA) é uma técnica que visa identificar e separar as fontes originais de dados estatísticos multidimensionais (as misturas), por meio de uma busca por componentes estatisticamente independentes e não gaussianos (AHMAD; GHANBARI, 2011).

Conforme descrito por Faier (2011), as variáveis transformadas são componentes implícitos que descrevem a estrutura essencial dos dados, em que se espera que essas componentes correspondam a alguma causa física envolvida no processo de geração dos dados.

Para melhor entendimento da análise ICA, suponha um vetor aleatório **s** com N fontes originais de interesse e não conhecido, tal como

$\boldsymbol{s} = [s_1 \quad s_2 \quad \cdots \quad s_N]^T$. As fontes **s** foram misturadas por meio da matriz **A**, denominada matriz de misturas, e resultou nos sinais observados **x** (OJA; YUAN, 2006; STONE, 2004), como descrito matematicamente na Equação 17.

$$\mathbf{x} = \mathbf{A}.\boldsymbol{s} \tag{17}$$

em que **x** corresponde à matriz com as observações misturadas; **A** é a matriz de misturas e **s,** as fontes originais.

A única variável conhecida na Equação 17 é o vetor **x** com as observações misturadas e o objetivo da técnica ICA é encontrar uma matriz de separação **W** que realize o processo inverso da matriz de misturas **A**, recuperando as fontes originais **s**.

A identificação e a separação das componentes independentes são realizadas por meio da maximização da não gaussianidade e a sequência de cálculos para a estimativa das fontes originais tem início com a centralização dos dados em torno da média, semelhante ao procedimento realizado na análise PCA e cuja operação matemática está descrita na Equação 12.

A fim de simplificar o problema ICA e auxiliar as rotinas computacionais de estimativa das componentes independentes a convergir mais rápido, é recomendado o branqueamento dos dados centralizados em torno da média (**y**), conforme relatado por Ahamad e Ghanbari (2011), Bell e Sejnowski (1995) e Karhunen (1996).

O branqueamento é uma transformação linear que visa transformar os dados centralizados na média (**y**) em um novo vetor (**k**) em que as observações sejam descorrelacionadas e tenham variância unitária (COSTA, 2006). A Equação 18 reproduz matematicamente a operação de branqueamento.

$$\mathbf{m} = \mathbf{E}.\mathbf{D}^{\frac{1}{2}}.\mathbf{E}^{\mathbf{T}}.\mathbf{y} \quad (18)$$

em que **m** são os dados branqueados, **E** é a matriz de autovetores e **D** corresponde à matriz diagonal de autovalores da matriz de covariância $\mathbf{C_m}$ = ($y \cdot y^T$).

Uma vez os dados branqueados, o desafio da ICA é encontrar uma matriz **W**, denominada matriz de separação, que permita recuperar as fontes originais (AHAMAD; GHANBARI, 2011; OJA; YUAN, 2006), conforme Equação19.

$$\mathbf{n} = \mathbf{W} \cdot \mathbf{m} \quad (19)$$

em que **n** são as fontes originais estimadas, **W** é a matriz de separação ($\mathbf{W} = \mathbf{A}^{-1}$) e **m** são as fontes branqueadas.

Existem, na literatura, diversos algoritmos para a estimativa das componentes independentes de um conjunto de observações e a base de cálculo de cada um está na maximização da não gaussianidade, o que é realizado minimizando a informação mútua, maximizando os valores absolutos da curtose ou da negentropia, dentre outros.

Caso as fontes originais **s** sejam estatisticamente independentes e não gaussianas, restrições que devem ser consideradas para o não comprometimento da ICA, as fontes estimadas (**n**) são as mais próximas possíveis das observações de **s**.

REFERÊNCIAS

AHMAD, T.; GHANBARI, M. A review of independent component analysis (ICA) based on kurtosis contrast function. **Australian Journal of Basic and Applied Sciences**, Amman, v. 5, n. 9, p. 1747-1755, 2011.

ARIZAGA, R. et al. Speckle time evolution characterization by the co-occurrence matrix analysis. **Optics and Laser Technology**, Surrey, v. 31, n. 2, p. 163-169, 1999.

BELL, A. J.; SEJNOWSKI, T. J. An information-maximization approach to blind separation and blind deconvolution. **Neural Computation**, Cambridge, v. 7, n. 6, p. 1129-1159, 1995.

BOLZAN, M. J. A. Transformada em ondeleta: uma necessidade. **Revista Brasileira de Ensino de Física**, São Paulo, v. 28, n. 4, p. 563-567, 2006.

BRAGA, R. A. et al. Evaluation of activity through dynamic laser speckle using the absolute value of the diferences. **Optics Communications**, Amsterdam, v. 284, n. 2, p. 646-650, 2011.

BRAGA, R. A. et al. Live biospeckle laser imaging of root tissue. **European Biophysics Journal**, London, v. 38, n. 5, p. 679-686, 2009.

CARVALHO, P. H. A. et al. Motility parameters assessment of bovine frozen semen by biospeckle laser (BSL) system. **Biosystems Engineering**, London, v. 102, n. 1, p. 31-35, 2009.

CHERIYADAT, A.; BRUCE, L. M. Why principal component analysis is not an appropriate feature extraction method for hyperspectral data. In: GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 3., 2003, Toulouse. **Proceedings…** Toulouse: IEEE International, 2003. p. 3420-3422.

COSTA, I. M. **Projeto e implementação em ambiente foundation fieldbus de filtragem estocástica baseada em análise de componentes independentes**. 2006. 68 p. Dissertação (Mestrado em Ciências) - Universidade Federal do Rio Grande do Norte, Natal, 2006.

DICK, D. P. et al. Impacto da queima nos atributos químicos do solo, na composição da matéria orgânica e na vegetação. **Pesquisa Agropecuária Brasileira**, Brasília, v. 43, n. 5, p. 633-640, maio 2008.

FAIER, J. M. **Análise dos componentes independentes para monitoração da qualidade de dados em séries temporais**. 2011. 153 p. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

FERRÃO, M. F. et al. Técnica não destrutiva de análise de tanino em café empregando espectroscopia no infravermelho e algoritmo genético. **Tecnológ**, Santa Cruz do Sul, v. 7, n. 1, p. 9-26, 2003.

FUJII, H. et al. Evaluation of blood flow by laser speckle image sensing. **Applied Optics**, New York, v. 26, n. 24, p. 5321-5325, 1987.

GROSSMANN, A.; MORLET, J. Decomposition of Hardy functions into square integrable wavelets of constant shape. **SIAM Journal on Mathematical Analysis**, Philadelphia, v. 15, n. 4, p. 723-736, Dec. 1984.

HADFIELD, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. **Journal of Statistical Software**, Los Angeles, v. 33, n. 2, p. 1-22, 2010.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. **Independent component analysis**. New York: J. Wiley, 2001. 504 p.

JUNG, S.; SEN, A.; MARRON, J. S. Boundary behavior in high dimension, low sample size asymptotics of PCA. **Journal of Multivariate Analysis**, New York, v. 109, p. 190-203, Aug. 2012.

KARHUNEN, J. Neural approaches to independent component analysis and source separation. In: EUROPEAN SYMP. ON ARTIFICIAL NEURAL NETWORKS, 4., 1996, Bruges. **Proceedings…** Bruges: ESANN, 1996. p. 249-266.

KARIMI, H. R.; PAWLUS, W.; ROBBERSMYR, K. G. Signal reconstruction, modeling and simulation of a vehicle full-scale crash test based Morlet wavelets. **Neurocomputing**, New York, v. 93, p. 88-99, Sept. 2012.

KURENDA, A. et al. Temperature effect on apple biospeckle activity evaluated with different indices. **Postharvest Biology and Technology**, Amsterdam, v. 67, p. 118-123, May 2012.

LIU, Y. et al. Daubechies wavelet-based method for elastic problems. **Engineering Analysis with Boundary Elements**, Southampton, v. 34, n. 2, p. 114-121, 2010.

MELÉM JÚNIOR, N. J. et al. Análise de componentes principais para avaliação de resultados analíticos da fertilidade de solos do Amapá. **Semina. Ciências Agrárias**, Passo Fundo, v. 29, n. 3, p. 499-506, 2008.

MORETO, F. A. L. **Análise dos componentes independentes aplicada à separação de sinais de áudio**. 2008. 83 p. Dissertação (Mestrado em Engenharia Elétrica) - Escola Politécnica da Universidade de São Paulo, São Paulo, 2008.

MORETTIN, P. A. **Ondas e ondeletas**. São Paulo: EDUSP, 1999. 276 p.

OJA, E.; YUAN, Z. The FastICA algorithm revisited: convergence analysis. **IEEE Transactions on Neural Networks**, Chicago, v. 17, n. 6, p. 1370-1381, 2006.

OULOMARA, G.; TRIBILLON, J.; DUVERNOY, J. Biological activity measurements on botanical specimen surfaces using a temporal decorrelation effect of laser speckle. **Journal of Moderns Optics**, London, v. 36, n. 2, p. 136-179, 1989.

PARTIO, M.; CRAMARIUC, B.; GABBOUJ, M. An ordinal co-occurrence matrix framework for texture retrieval. **Journal on Image and Video Processing**, Heidelberg, v. 2007, n. 1, p. 1-5, 2007.

POLANSKY, L. et al. From moonlight to movement and synchronized randomness: fourier and wavelet analyses of animal location time series data. **Ecology**, Durham, v. 91, n. 5, p. 1506-1518, May 2010.

PONTELLI, C. O. et al. Comparação entre dois tipos de suspensão passiva de barra em condições simuladas de pista de prova normalizada. **Engenharia Agrícola**, Jaboticabal, v. 30, n. 4, p. 761-775, ago. 2010.

RABAL, H. J.; BRAGA, R. A. **Dynamic laser speckle and applications**. Boca Raton: CRC, 2008. 304 p.

RABAL, H. J. et al. Q-statistics in dynamic speckle pattern analysis. **Optics and Lasers in Engineering**, London, v. 50, n. 6, p. 855-861, 2012.

SALGADO, D. D.; NÄÄS, I. A. Avaliação de risco à produção de frango de corte do estado de São Paulo em função da temperatura ambiente. **Engenharia Agrícola**, Jaboticabal, v. 30, n. 3, p. 367-376, 2010.

SIFUZZAMAN, M.; ISLAM, M. R.; ALI, M. Z. Application of wavelet transform and its advantages compared to Fourier transform. **Journal of Physical Sciences**, Kelantan, v. 13, n. 1, p. 121-134, 2009.

SILVA, A. F.; MINIM, V. P. R.; RIBEIRO, M. M. Análise sensorial de diferentes marcas comerciais de café (*Coffea arabica L.*) orgânico. **Ciência e Agrotecnologia**, Lavras, v. 29, n. 6, p. 1224-1230, nov./dez. 2005.

SILVA, M. M. et al. Optical mouse acting as biospeckle sensor. **Optics Communications**, Amsterdam, v. 284, n. 7, p. 1798-1802, 2011.

STONE, J. V. **Independent component analysis:** a tutorial introduction. Cambridge: MIT, 2004. 193 p.

TORRENCE, C.; COMPO, G. P. A practical guide to wavelet analysis. **Bulletin of the American Meteorological Society**, Boston, v. 79, n. 1, p. 61-78, 1998.

XU, Y. et al. Wavelet transform domain filters: a spatially selective noise filtration technique. **IEEE Transactions on Image Processing**, New York, v. 3, n. 6, p. 747-758, 1994.

XU, Z.; JOENATHAN, C.; KHORANA, B. M. Temporal and spatial proprieties of the time varying speckles of botanical. **Journal of Optical Enginnering**, Bellingham, v. 34, n. 5, p. 1487-1502, 1995.

ZDUNEK, A. et al. The biospeckle method for the investigation of agricultural crops: a review. **Optics and Lasers in Engineering**, London, v. 52, p. 276-285, Jan. 2013.

ZDUNEK, A.; HERPPICH, W. B. Relation of biospeckle activity with chlorophyll content in apples. **Postharvest Biology and Technology**, Amsterdam, v. 64, n. 1, p. 58-63, 2012.

ZHANG, L. et al. Two-stage image denoising by principal component analysis with local pixel grouping. **Pattern Recognition**, Ezmsford, v. 43, n. 4, p. 1531-1549, 2010.

SEGUNDA PARTE - ARTIGOS

ARTIGO 1

O primeiro artigo está apresentado na íntegra conforme foi publicado pela revista Applied Mathematics em sua edição especial Harmonic Analysis and Wavelets.

Comparison Between Fourier and Wavelets Transforms in Biospeckle Signals

Kleber Mariano Ribeiro[1] *, Roberto Alves Braga Júnior[1],
Thelma Sáfadi[2], Graham Horgan[3]

[1] Engineering Department, Federal University of Lavras,
Lavras, Minas Gerais, Brazil.
Email addresses: klebermariano@gmail.com robbraga@deg.ufla.br
[2] Exact Science Department, Federal University of Lavras,
Lavras, Minas Gerais, Brazil.
Email address: safadi@dex.ufla.br
[3] Biomathematics and Statistics Scotland,
Rowett Institute of Nutrition and Health, Aberdeen, Scotland.
Email address: g.horgan@abdn.ac.uk

* Corresponding author: Tel. + 55 35 3829 1210

Abstract

The dynamic speckle is a non-destructive optical technique that has been used as a tool for the characterization of the biological activity and several studies are conducted to obtain for more information about the correspondence of the observed phenomena and their expressions in the interference images. Analysis in the frequency domain has been considered as powerful alternative, and although there are works using Fourier transform in the frequency analysis of the biospeckle signals, the majority present the wavelet transform as tool for spectral analysis. In turn, there are still doubts if the Fourier transform is not enough for the analysis of the biospeckle, which would enable the reduction of processing time since an operation is computationally simpler. In this context, the present study aims to compare the constituents' parts of the speckle signal according to Fourier and wavelet transforms for numerical analysis. The comparative analysis based on the absolute values of the differences technique (AVD), were carried out for performance evaluation of the Fourier and wavelet transforms, in which the speckle signals were decomposed spectrally and subsequently reconstructed with the elimination of specific frequency bands. Results showed that the wavelet transform allowed more information about signals constituents of the dynamic speckle, emphasizing its use instead of the Fourier transform, which in turn was restricted the situations in which the only interest is to know the spectral content of the data.

Keywords: Spectral analysis; dynamic speckle; biological activity.

Comparação entre as transformadas de Fourier e de wavelets no sinal do *biospeckle*

Resumo

O *speckle* dinâmico é uma técnica óptica não destrutiva que vem sendo validada como uma ferramenta para a caracterização da atividade biológica e vários estudos estão sendo conduzidos para se obter maiores informações sobre as correspondentes dos fenômenos observados e suas expressões nas imagens interferências. A análise no domínio da frequência tem sido considerada uma potencial alternativa e, apesar de existirem trabalhos que utilizam a transformada de Fourier nas análises em frequência do sinal do *biospeckle*, a maioria apresenta a transformada de wavelet como ferramenta para análise espectral. Dessa forma, ainda há dúvidas se a transformada de Fourier não é suficiente para as análises espectrais do *biospeckle*, o que permitiria a redução do tempo de processamento uma vez que a operação é computacionalmente mais simples. Neste contexto, o presente estudo foi realizado com o objetivo de comparar as partes constituintes do sinal do *speckle* de acordo com as transformadas de Fourier e de wavelet para análises numéricas. As análises comparativas, baseada na técnica da diferença dos valores absolutos (DVA), foram realizadas para a avaliação do desempenho das transformadas de Fourier e de wavelet, em que os sinais do *speckle* foram decompostos espectralmente e, em seguida, reconstruídos com a eliminação de bandas de frequências específicas. Os resultados mostraram que a transformada de wavelet apresentou maiores informações sobre os sinais constituintes do *speckle* dinâmico, enfatizando sua utilização em vez da transformada de Fourier que, por sua vez, ficou restrita a situações em que o único interesse é conhecer o conteúdo espectral dos dados.

Palavras-chave: Análise Espectral; *Speckle* Dinâmico; Atividade Biológica.

1 Introduction

When a coherent light, such as laser, illuminates a rough surface, compared the wavelength of laser, occurs a phenomenon of optical interference with the formation of light and dark regions, called speckle (ZHAO et al., 1997).

After applying the dynamic surface, there is a continuous formation of new and different speckles, and these random and dynamic interference patterns is called dynamic speckle or biospeckle, if the area concerned is biological. This technique allows extracting information about the structures movement of the illuminated material, making it an interesting tool in several knowledge areas (RABAL; BRAGA, 2008).

The biospeckle has been used as a technique to measure detailed extensions of pine roots (RATHANAYAKE et al., 2008) or even the biological activity of roots in tissue culture (BRAGA et al., 2009), in assessing the water activity in maize and beans seeds (CARDOSO et al., 2011), to studies of the relationship between chlorophyll pigments present in apples and their respective biological activity (ZDUNEK; HERPPICH, 2012), and several other papers.

The biological activity expressed in the context of speckle does not present a clear definition of what phenomenon is creating, however can be understood as structural and molecular motions occurring in the material analysis (BRAGA et al., 2009), Doppler effect, Brownian motion, variations of the refractive index (PASSONI et al., 2005), among others. It is a complex signal and with causes still investigated (COSTA et al., 2010), which is a challenge and at the same time a motivation.

In this context, the use of image processing techniques and signal analysis tools can be used in the biospeckle signal to understand better this optical phenomenon.

The interference patterns analysis can use graphical methods, that generate maps indicating the spatial variability of the biological activity, or a numerical

interpretation of the temporal variation of patterns formed. An alternative the graphical and numerical classifications are signal analysis in the time domain or in the frequency domain (CARDOSO et al., 2011).

The analysis of biospeckle signals in the frequency domain has been an alternative for many applications, allowing the filter and images contrast, beyond search of frequency markers of phenomena that contribute to the formation of the interference patterns in time, as described by (CARDOSO et al., 2011). Thereby, Fourier and wavelet transforms can be a good choice to make such analysis in the frequency domain.

Several studies have been conducted using the spectral analysis in the biospeckle signal, such as Rabelo et al. (2011) that used the Fourier transform to analysis bean seeds contaminated by two kinds of fungi and managed to differentiate them using the harmonics amplitude; Sendra et al. (2005) assessed damage in apples and seed germination using wavelet transform and defined frequency markers for biological phenomena, as well as Cardoso et al. (2011) who studied maize and beans, and cancer isolation and others.

Although there are many papers applying spectral analysis in the biospeckle signal the most journals use wavelet transform and there is not works evaluating if Fourier transform, that is simpler than wavelet transform, is enough in the frequency analysis of the dynamic speckle. In this context, the present study aims to compare the Fourier and wavelet transform in the spectral analysis of biospeckle signal.

2 Theory

2.1 Time history of the speckle patterns (THSP)

The biospeckle is a nondestructive optical technique based on the analysis of the variations of the laser light scattered from material, and the biological activity presented reflects the state of the investigated object (ANSARI; NIRALA, 2013).

Follow a set of pixels of the images speckles in the time is a method of monitoring their time variations and consequently the biological activity of the studied object, and, in this context, Oulamara, Tribillon e Duvernoy (1989) proposed the Time History of the Speckle Patterns (THSP).

The THSP is a two dimensional image that record a certain line or column of pixels in successive moments and arrange them vertically side by side. The x axis show information about the time evolution of the selected pixels and the y axis is the spatial distribution of the interference patterns (OULAMARA; TRIBILLON; DUVERNOY, 1989).

2.2 Co-occurrence matrix

The co-occurrence matrix was presented by Arizaga, Trivi e Rabal (1999), and expresses the number of the transitions of each THSP pixel with respect to its immediate neighbor. Equation 1 describes mathematically the co-occurrence matrix.

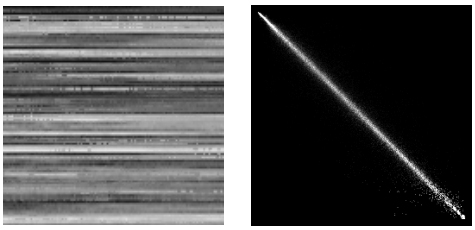
$$\mathbf{M_{CO}} = [\mathbf{N_{ij}}] \tag{1}$$

which:

M_{CO} is the co-occurrence matrix, N_{ij} correspond the number of occurrences of an intensity value i, followed by an intensity value j to move through rows or

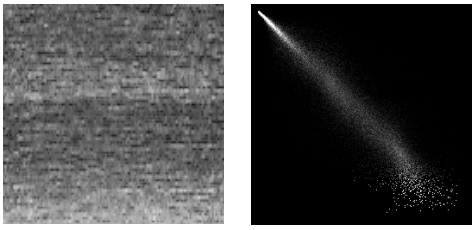
columns of the time history.

Phenomenon that show low biological activities, their time variations of the speckle patterns are slow and present a THSP horizontally in the elongated shape and the co-occurrence matrix is characterized by small changes of the pixels intensity to i and j, as illustrated in the Figure 1A. However, materials that exhibit high biological activity shows fast intensity variations in the THSP that resemble an ordinary spatial speckle patterns and their co-occurrence matrix has nonzero elements near the main diagonal (Figure 1B) (BRAGA et al., 2003).



THSP Moc

(A)



THSP Moc

(B)

Figure 1 Time history of the speckle patterns and their respective co-occurrence matrix. Materials with low (A) and high (B) biological activity

2.3 Absolute values of the differences (AVD)

One of the methods for analyzing of the speckle patterns is the technique of the absolute values of the differences (AVD), proposed by Braga et al. (2011) as an alternative the inertial moment technique.

The AVD method is a statistics moment of first order which it is applied on the co-occurrence matrix and generates a number (ANSARI; NIRALA, 2013) which allow quantify the biological activity of the studied material. Equation 2 presents mathematically the AVD technique.

$$AVD = \sum_i \sum_j \{M_{ij} \cdot |(i - j)|\}$$

(2)

which:

AVD is a dimensionless value, i and j are coordinates of the row and column respectively, and M_{ij} is called of modified co-occurrence matrix and that is presented in the Equation 3.

$$M_{ij} = \frac{N_{ij}}{\sum N_{ij}} \quad (3)$$

According Braga et al. (2011), the inertial moment showed be more sensitive than AVD on analyzing processes that involve high biological activities, although when this variation is not so intense, this method is less efficient.

2.4 Fourier transform

Information of the biospeckle data in the frequency domain has been an alternative to the interpretation of the interference patterns (CARDOSO et al., 2011), with the possibility of improve the visualization of some phenomena of the studied material and to know their spectral signatures. In this context, the Fourier transform is one of the tools that can be used to spectral analysis of the

biospeckle.

Fourier transform can be understood as the mathematical technique that transforms a signal from the time domain to the frequency domain, and it is formed by a set $w_n(t) = e^{int}$, n = 0, 1, ... of orthogonal functions, of period 2π (MORETTIN, 1999). Equation 4 described mathematically the Fourier transform.

$$f(\omega) = \int_{-\infty}^{\infty} f(t) \, . \, e^{-i\omega t} \, dt \tag{4}$$

which:

$\omega = 2 \pi f$

$| f(\omega) |$ = amplitude of each component ω of the signal.

There is also the inverse Fourier transform, which is used to transform the signal from frequency domain to time domain with the reconstruction of the original function. Equation 5 presents the mathematical expression of the inverse Fourier transform.

$$f(t) = \int_{-\infty}^{\infty} f(\omega) \, . \, e^{-i\omega t} \, d\omega \tag{5}$$

The Fourier transform indicates the spectral information of the signal without providing the instant which these components happen, and in situations that to know when the frequencies occur are interesting precludes the use of Fourier transforms, unless if the series is stationary (SIFUZZAMAN; ISLAM; ALI, 2009). In this context, the wavelet transform is an alternative that provides

the instant the frequency components occur.

2.5 Wavelets transform

The wavelets are simply waves of duration adjusted with energy concentrated in variables intervals (GRAPS, 1995), which makes it a great useful method for time series analysis, that exhibit characteristics that can change in the time and in frequency.

The continuous wavelet transform is defined as the convolution of f(t) with a scaled and translated version of ψ (TORRENCE; COMPO, 1998), called wavelet mother. Equation 6 describes mathematically the continuous wavelet transform.

$$< f, \psi_{a,b} > = \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt \qquad (6)$$

which:

f(t) is the studied signal

a scale parameter

b translation value

$\psi_{a,b}(t)$ is the mother function of wavelets

$< f, \psi_{a,b} >$ is the spectrum wavelets.

The scale is related to the frequency, in which high scales correspond to low frequencies and low scales correspond to high frequencies, whereas the translation is the displacement of the mother function about the studied signal (KARIMI; PAWLUS; ROBBERSMYR, 2012).

The return of the signal from frequency domain to time domain, inverse wavelets transform, allows observe the behavior of the signal in specifics

frequencies bands and also the reconstruction of the original function f(t). According Torrence and Compo (1998), the inverse wavelet transform can be realized by the sum of real part of wavelet spectrum on all scales (Equation 7).

(7)

which:

$s_j^{0.50}$ is a factor that convert the wavelets transform in energy density,

δ_j; $\delta_t^{0.50}$; C_δ; $\psi_0(0)$ are specific constants of the base function used.

One of the major difficulties in wavelet analysis is the identification of the scales set used in the wavelet transform. Orthogonal wavelet, there is a limit and a discrete set of scales, as given by Farge (1992), however, for analysis of non-orthogonal wavelet, can use an arbitrary scales set to build a more complete signal (TORRENCE; COMPO, 1998).

In this context, Torrence and Compo (1998) suggested the Equations 8 and 9 to calculate the scales interval to be used in the wavelet transform, in which s_j is the lowest and J is the highest scale.

$$s_j = s_0 . 2^{j.\delta j} \qquad j = 0, 1, 2, \ldots, J \tag{8}$$

$$J = \delta j^{-1} . LOG_2 \left(\frac{N . \delta t}{s_0} \right)$$

(9)

The s_0 should be chosen so that the Fourier period is 2.δt, and to the Morlet wavelet the largest value that can adjust the scale is δj of 0.5. For other wavelet functions can be used a larger value.

2.6 Sampling theorem

The sampling theorem describes the relationship between sampling frequency of a signal and the frequency maximum of the reconstructed signal. Below is transcript the sampling theorem as presented by Shannon (1949).

"Theorem 1: If a function f (t) contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced 1/2 seconds W apart".

According to the theorem, the number of samples per unit time of a signal is called rate or frequency sampling (W), and half the sampling frequency corresponds to the frequency maximum of the signal which can be reproduced in full without aliasing error.

The sampling theorem is used in this work to define the highest frequently during the decomposition of signals.

3 Materials and methods

It was conducted a comparison between Fourier and wavelet transforms using the time history of speckle patterns (THSP) relative to a paint drying process and presented by Silva et al. (2011).

The database was formed by 8 THSP's collected each 20 minutes during the paint drying using the back-scattering experimental setup. Each time history was made by a set of 128 images, resolution of 512 by 640 pixels, whose time acquisition between images was of 0.08 seconds (sampling frequency of 12.5 Hz).

The lines of the THSPs were concatenated creating a new signal that was decomposed into frequency spectra using Fourier and wavelet transforms with application posterior of the inverse transform. Some frequency bands were eliminated before the reconstruction of the signal in order to analyze the results of the speckle signal using a numerical method to measure the speckle activity. The selective filtering was conducted as well in order to create some frequency markers linked to the physical phenomena under monitoring.

According the sampling theorem the highest frequency that can be seen in the reconstruction process is 6.25 Hz, and using the Equations 8 and 9 were calculated the number of frequency bands used in the transform. In addition, in the continuous wavelet transform was used mother function of Morlet, a damped complex exponential with a set of oscillation parameter that preserves an approximate relationship between the scale of the wavelet analysis and the frequency in a Fourier analysis, as described by Polansky et al. (2010).

The signal resulting of the inverse transform was converted to THSP format again and numerically analyzed using the technique of the absolute values of the differences (AVD) (BRAGA et al., 2011), and their values compared to the gravimetrical measurement.

Figure 2 illustrated all the methodology used.

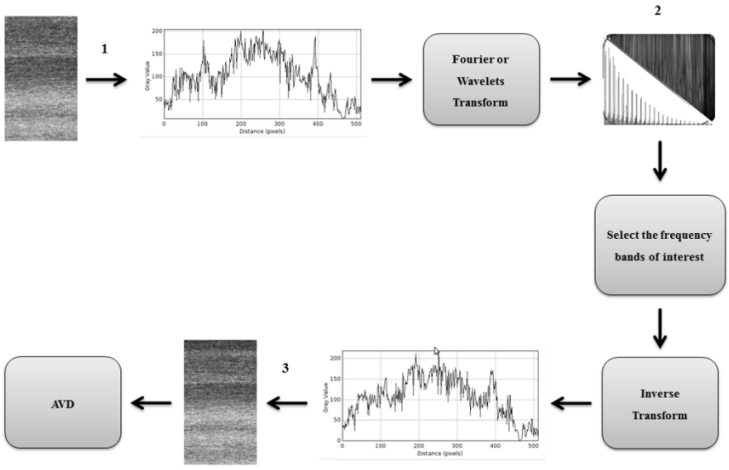


Figure 2 Methodology used to the data analysis, in which 1 represent the concatenation, 2 is the Fourier or wavelets spectrum and 3 correspond to the inverse process of the concatenation

4 Results and discussion

The Figure 3 present the absolute value of the differences for the THSP's of the paint drying process with decomposition and reconstruction of some frequency bands using Fourier transform.

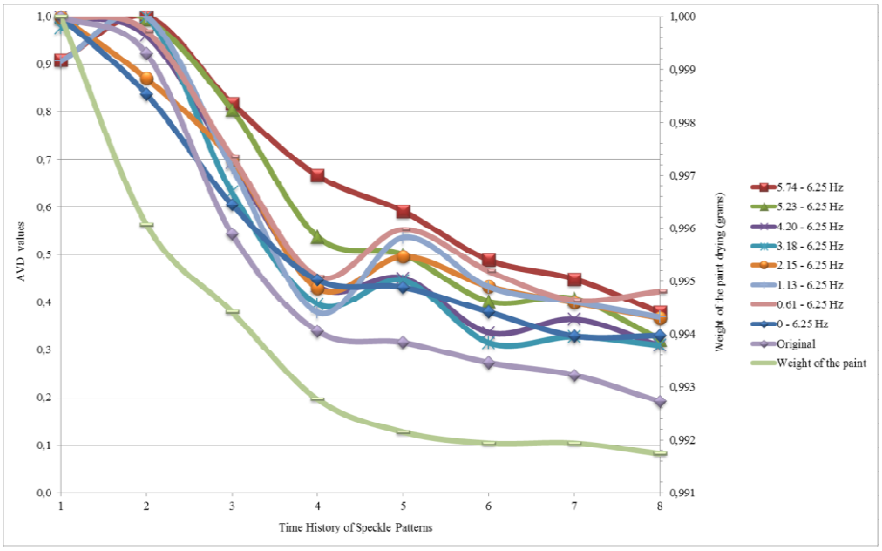


Figure 3 AVD values of the THSP's of the paint drying using Fourier transform for spectral analysis

The data reconstructed in the frequency band from 0 to 6.25 Hz (Figure 3), here called of total reconstruction, correspond to highest reconstruction possible in accordance with the sampling theorem, and the original data are the AVD values of the time history of the paint drying without filtering.

The dynamic speckle signals reconstructed in frequencies bands 5.74-6.25 and 5.23-6.25 Hz presented gradual reduction of AVD values along of the paint drying process, closer to the behavior of the original data, total reconstruction and to the weight of the paint drying. However, the addition of low frequencies components in the reconstruction process resulted in the oscillation of the AVD values in the fifth time history, as observed in the Figure 3. We attribute those oscillations to the influence of the atmospheric conditions that occur in experiment of paint drying as the realized by Silva et al. (2011), which did not interfered in the first moments since the paint volatility was higher and thus undermined the presence of the modulation of the signal in low

frequencies linked to the atmospheric conditions such as temperature and humidity, and the Figure 4 illustrates these information's.

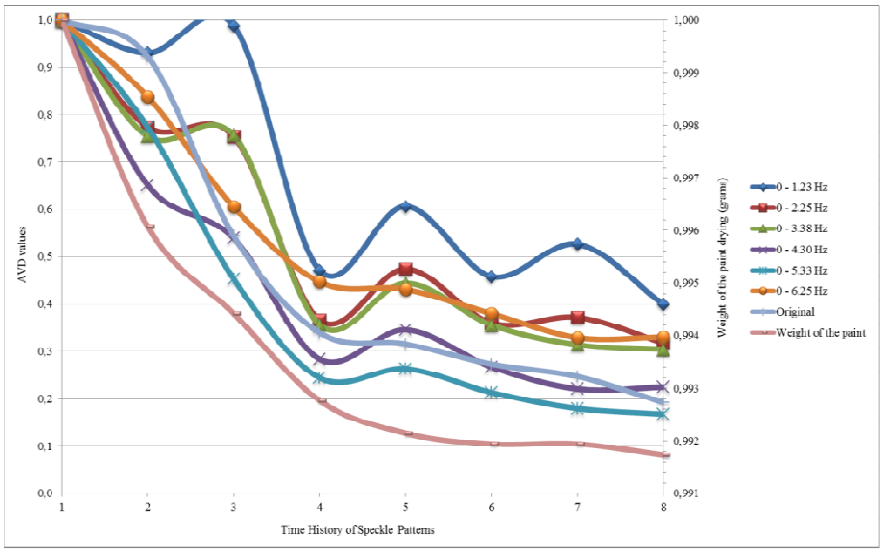


Figure 4 AVD values of the time history of the paint drying reconstructed with some frequencies band using Fourier transform

The Figure 4 presents that the signals reconstructed with components of low frequencies (associated the temperature and humidity variables) are mixed with the data reconstructed using components of high frequencies (linked the volatility) in the first moments, and which along the paint drying process, the paint volatility stabilize and make possible to observe the high oscillations of the data reconstructed using low frequencies component.

In order to clarify the spectral information found in the fifth THSP, a mathematical model was adjusted in the AVD values of the original data, using the least squares method, to describe the process of paint drying. The Equation 10 present the mathematical model adjusted and the Figure 5 illustrates the regression curve.

$$y = -0.003624\,x^4 + 0.062957\,x^3 - 0.340004\,x^2 + 0.450417\,x + 0.889444 \quad (10)$$

which:

x is the TSHP number,

y is the normalized AVD value.

The mathematical model adjusted showed a correlation index of 0.98 and mean square error of 0.0048 with respect to the original.

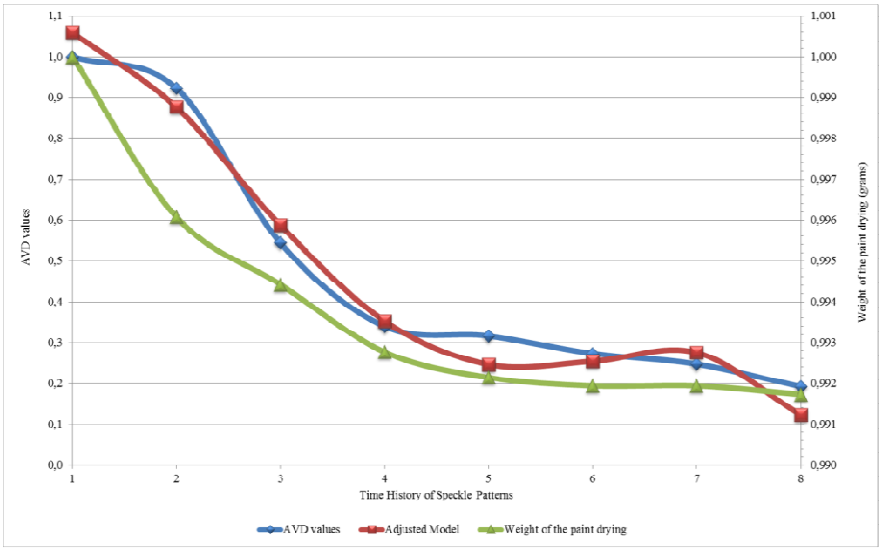
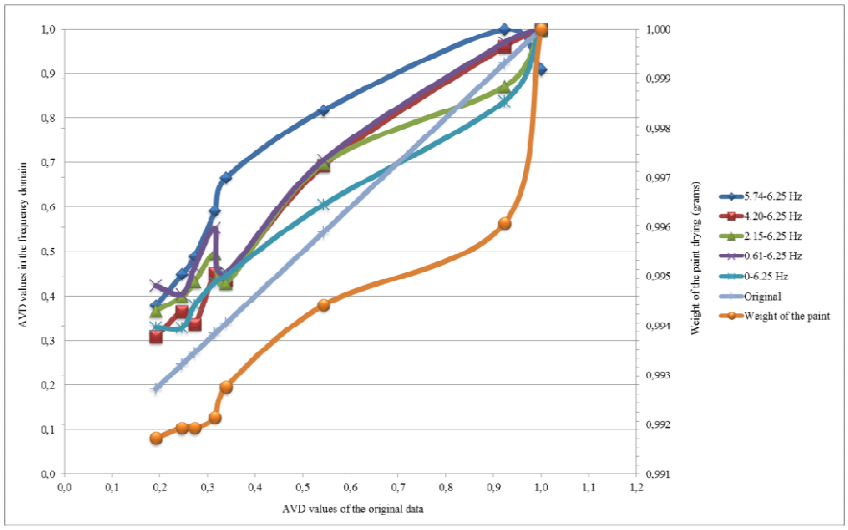


Figure 5 Regression adjusted to describe the paint drying process

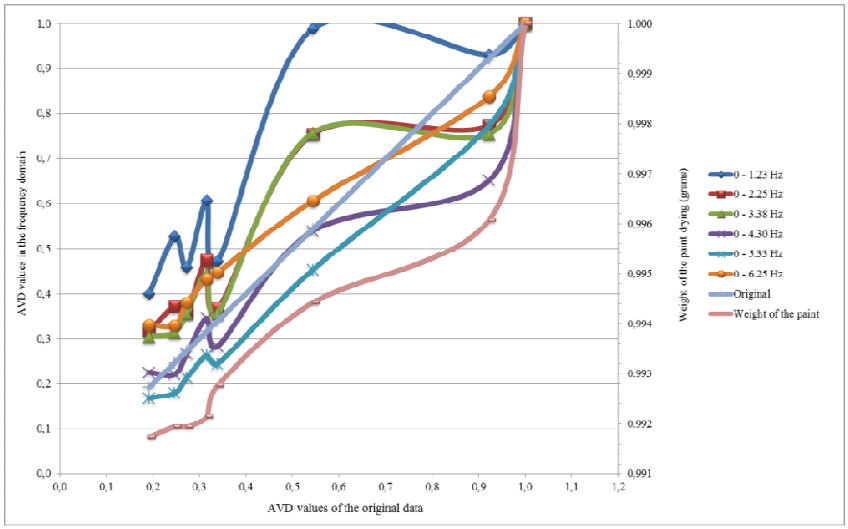
The first derivative of the adjusted model ($y' = -0.014496\,x^3 + 0.188869\,x^2 - 0.680007\,x + 0.450417$) when equaled to zero, showed that the region the fifth THSP is a minimum local, and this means that the AVD values reaches a minimum value and then initiates an oscillations related to the variations of the temperature and humidity.

These results are similar to the behavior of the paint weight at time presented by Silva et al. (2011) (Figure 5), in which is possible to observe a stabilization of the weight values after the fifth data acquisition, and leads us to assume that the drying time of this paint is one hour and twenty minutes, approximately. In this context, the dynamic speckle analyzed by Fourier transform allowed to observe this transition occurred in paint drying structure.

In addition, the distance between the original signals and reconstructed signals were evaluated and the results are illustrated in the Figure 6.



(A)



(B)

Figure 6 Signals reconstructed in specific frequency bands by inverse Fourier transform and the original signal. (A) Addition of components of low frequencies in the signals reconstruction and (B) increase high frequencies components in the inverse transform

The largest distances were observed in the signals reconstructed using few spectral components and these distances between original signals and reconstructed signals were reducing to use a larger number of components in the inverse Fourier transform, as waited.

In addition, the correlation index between the original signal and the signals reconstructed in the frequencies bands of 5.74-6.25, 5.23-6.25 and 4.20-6.25 Hz were of 0.86, 0.94 and 0.98, respectively. The high correlation of 0.86 using just a small portion of the frequency band can be explained by compact support of the Fourier basis functions in the frequency domain, which allows using in data compression with minimum loss of information, as discussed by Morettin (1999).

The speckle signals is not stationary (SENDRA; MURIALDO; PASSONI, 2007), unviable to use the Fourier transform. Thus, frequency analysis using the Fourier transform is restricted to situations in which are interesting to know the spectral information of the data.

Otherwise, the spectral analysis using the wavelet transform presented different behavior in the high frequencies and the signal reconstructed using more components of low frequencies made the results closer to the original signal, as illustrated in the Figure 7.

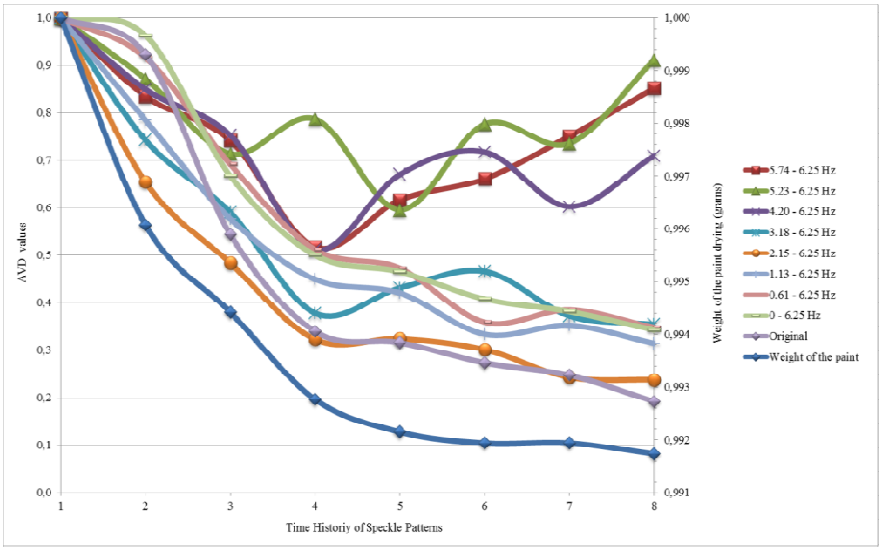


Figure 7 AVD values of the time history of speckle patterns of paint drying using wavelet transform for spectral analysis

The total reconstruction of the biospeckle signal, in the frequency band from 0 to 6.25 Hz, showed behavior similar to the original data and to the weight of paint drying, with the gradual reduction of the AVD values in the time.

Furthermore, removing the low frequency components in the reconstruction process resulted in oscillations in the AVD values, in special, in the signals reconstructed within the frequency range from 4.20 to 6.25 Hz, which the AVD values decreasing until the fourth time history and subsequently increasing.

In this context, the high AVD values of the last time history in the frequency bands 4.20-6.25 Hz are attributed the random oscillations and noise presents in the biospeckle signal, without significant information's about the paint volatility, since that the energy of the time history in the high frequencies showed was reducing along of the paint drying and presented low values after of the fourth THSP, as illustrated graphically in the Figure 8.

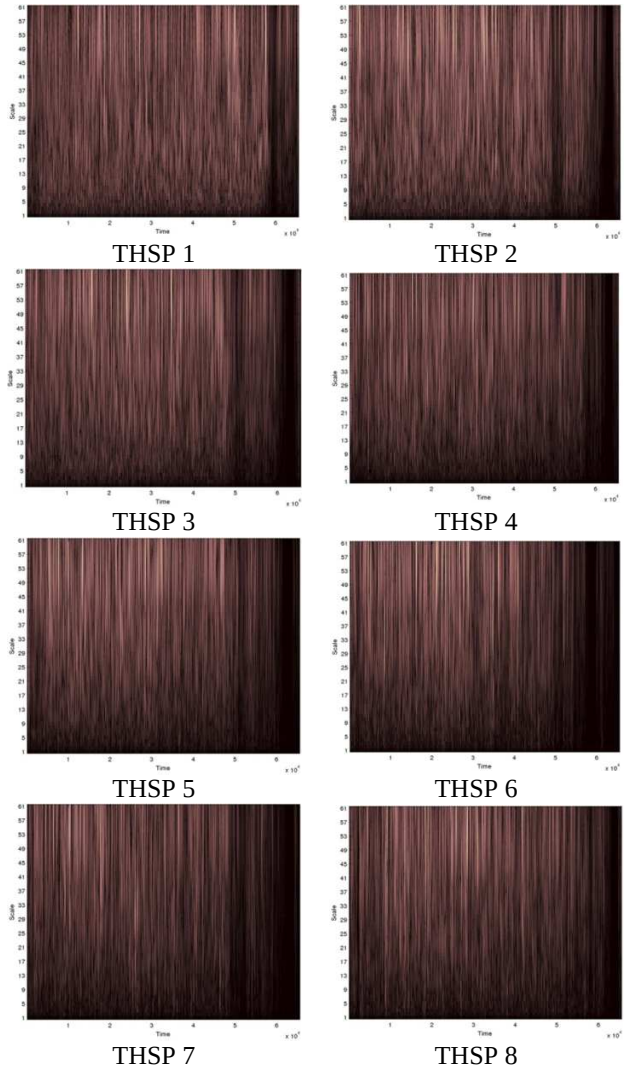


Figure 8 Energy of the 8 THSP’s for different frequencies

The energy of the THSP's (Figure 8) is represented in pseudo-colors, the ordinate axis correspond the scales and in the abscissas axis is the time. The light pseudo-colors indicate high energy while dark shades are associated the low energy, and the scales are inversely proportional to the frequencies, which

the low scales are attached with high frequencies and high scales with the low frequencies.

It is possible to observe in the Figure 8 that the pseudo-colors in the high frequencies (low scales) are darkening in the time, which means reduce of the energy in high frequency bands along of the paint drying and which does not justify the high AVD values in the last time histories.

The reconstruction of the signals using components of low frequencies was also analyzed, and the Figure 9 shows the results.

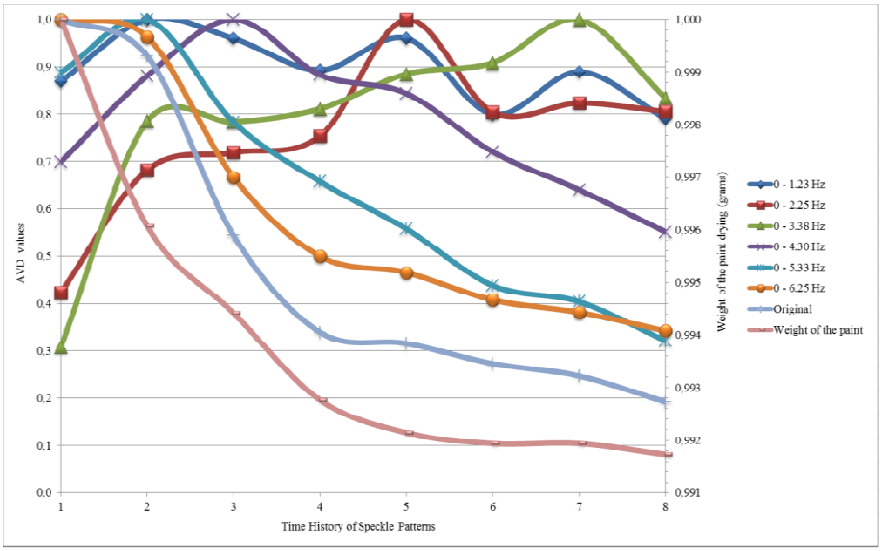
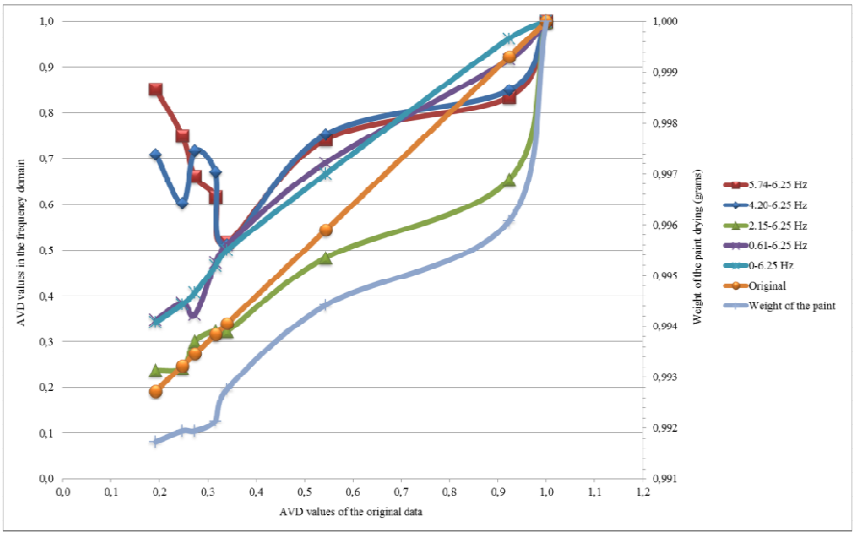


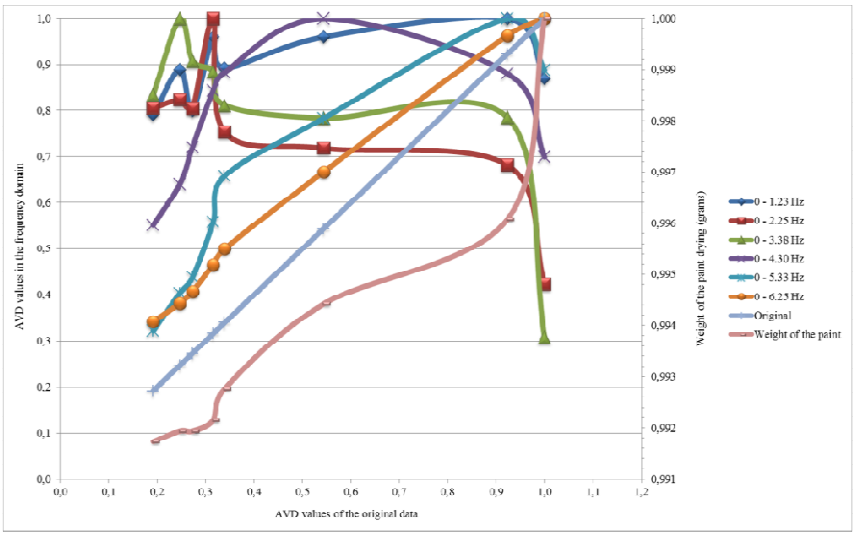
Figure 9 Absolute value of the difference of the THSP's of the paint drying using wavelet transform for spectral analysis

In the first moments we see that the signals reconstructed using components of low frequency and of high frequency mixed, instant in which the paint volatility was intense. Over time, the phenomena linked to the high frequencies stabilized, allow to observe the oscillations of the signals reconstructed with components of low frequency, as the humidity and the temperature.

The reconstruction of the signals showed correlation index higher than 0.90 with the original signal when used a wide frequency band, the opposite showed when we used the Fourier transform. The Figure 10 presents the signals reconstructed using wavelet transform against the original signal, and allow observation the distance between the curves when we added components of high and low frequencies in the reconstruction of the signals.



(A)



(B)

Figure 10 Signals reconstructed using wavelet transform and the original signal. (A) Addition of components of low frequencies in the inverse wavelet transform and (B) increase high frequencies components in the signal reconstruction

Signals reconstructed in frequency bands of 5.23-6.25, 4.20-6.25 and

3.18-6.25 Hz presented correlation index of 0.30, 0.69 and 0.92, respectively. The Morlet function does not have compact support, which explains the need for a large number of spectral components for great approximation of the original signal. Furthermore, the addition of components of low frequencies (Figure 10A) and of high frequencies (Figure 10B) in the signal reconstruction made the reconstructed data closer to the original signal.

In this context, the wavelet transform details spectral information's in time, which does not occur in the Fourier transform, and that in the analysis of the biospeckle signals provides further information of the studied process and facilitates to understand the signals that create this complex phenomenon of optical interference, being the adequate tool for studies in the frequency domain.

5 Conclusion

The Fourier transform allowed data analysis with compact support, while the wavelets provided definition of frequency markers and information's not presented in the Fourier analysis about the participants of the dynamic speckle signal and being the tool most adequate for frequency analysis.

Acknowledgements

This work was partially financed by CNPq, Fapemig, Capes, Finep in Brazil, and and partly supported by the Scottish Government Rural and Environment Science and Analytical Services division.

REFERENCES

ANSARI, M. D. Z.; NIRALA, A. K. Assessment of the bio-activity using the

methods of inertia moment and absolute value of the differences. **Optiks - International Journal for Light and Electron Optics**, Amsterdam, v. 124, n. 15, p. 2180-2186, 2013.

ARIZAGA, R.; TRIVI, M.; RABAL, H. Speckle time evolution characterization by the co-occurrence matrix analysis. **Optics and Laser Technology**, New York, v. 31, n. 2, p. 163-169, 1999.

BRAGA, R. A. et al. Assessment of seed viability by laser speckle techniques. **Biosystems Engineering**, London, v. 86, n. 3, p. 287-294, 2003.

BRAGA, R. A. et al. Evaluation of activity through dynamic laser speckle using the absolute value of the diferences. **Optics Communications**, Amsterdam, v. 284, n. 2, p. 646-650, 2011.

BRAGA, R. A. et al. Live biospeckle laser imaging of root tissue. **European Biophysics Journal**, London, v. 38, n. 5, p. 679-686, 2009.

CARDOSO, R. R. et al. Frequency signature of water activity by biospeckle laser. **Optics Communications**, Amsterdam, v. 284, n. 8, p. 2131-2136, 2011.

COSTA, R. M. et al. Técnicas estatísticas aplicadas em imagens do speckle dinâmico. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 2, p. 27-39, 2010.

FARGE, M. Wavelet transforms and their applications to turbulence. **Annual Review of Fluid Mechanics**, Evanston, v. 24, n. 1, p. 395-457, 1992.

GRAPS, A. An introduction to wavelets. **Computation Science and Engineering, IEEE**, Ottawa, v. 2, n. 2, p. 50-61, 1995.

KARIMI, H. R.; PAWLUS, W.; ROBBERSMYR, K. G. Signal reconstruction, modeling and simulation of a vehicle full-scale crash test based Morlet wavelets. **Neurocomputing**, New York, v. 93, p. 88-99, Sept. 2012.

MORETTIN, P. A. **Ondas e ondeletas**. São Paulo: EDUSP, 1999. 280 p.

OULOMARA, G.; TRIBILLON, J.; DUVERNOY, J. Biological activity measurements on botanical specimen surfaces using a temporal decorrelation effect of laser speckle. **Journal of Moderns Optics**, London, v. 36, n. 2, p. 136-179, 1989.

PASSONI, I. et al. Dynamic speckle processing using wavelets based entropy. **Optics Communications**, Amsterdam, v. 246, n. 1/3, p. 219-228, 2005.

POLANSKY, L. et al. From moonlight to movement and synchronized randomness: fourier and wavelet analyses of animal location time series data. **Ecology**, Durham, v. 91, n. 5, p. 1506-1518, May 2010.

RABAL, H.; BRAGA, R. A. **Dynamic laser speckle and applications**. New York: CRC, 2008. 304 p.

RABELO, G. F. et al. Frequency response of biospeckle laser images of bean seeds contaminated by fungi. **Biosystems Engineering**, London, v. 110, n. 3, p. 297-301, 2011.

RATHNAYAKE, A. P. et al. A novel optical interference technique to measure minute root elongations of Japanese red pine (Pinus densiflora Seibold & Zucc.) seedlings infected with ectomycorrhizal fungi. **Environmental and Experimental Botany**, Elmsford, v. 64, n. 3, p. 314-321, 2008.

SENDRA, G. H. et al. Decomposition of biospeckle images in temporary spectral bands. **Optics Letters**, Washington, v. 30, n. 13, p. 1641-1643, 2005.

SENDRA, H.; MURIALDO, S.; PASSONI, L. Dynamic laser speckle to detect motile bacterial response of Pseudomonas aeruginosa. **Journal of Physics: Conference Series**, Moscow, v. 90, n. 1, 2007. Disponível em: <http://iopscience.iop.org/1742-6596/90/1/012064>. Acesso em: 10 jan. 2014.

SHANNON, C. E. Communication in the presence of noise. **Proceedings of the IRE**, Ottawa, v. 37, n. 1, p. 10-21, 1949.

SIFUZZAMAN, M.; ISLAM, M. R.; ALI, M. Z. Application of wavelet transform and its advantages compared to Fourier transform. **Journal of Physical Sciences**, Kelantan, v. 13, n. 1, p. 121-134, 2009.

SILVA, M. M. et al. Optical mouse acting as biospeckle sensor. **Optics Communications**, Amsterdam, v. 284, n. 7, p. 1798-1802, 2011.

TORRENCE, C.; COMPO, G. P. A practical guide to wavelet analysis. **Bulletin of the American Meteorological Society**, Boston, v. 79, n. 1, p. 61-78, 1998.

ZDUNEK, A.; HERPPICH, W. B. Relation of biospeckle activity with chlorophyll content in apples. **Postharvest Biology and Technology**, Amsterdam, v. 64, n. 1, p. 58-63, 2012.

ZHAO, Y. et al. Point-wise and whole-field laser speckle intensity fluctuation measurements applied to botanical specimens. **Optics and Lasers in Engineering**, London, v. 28, n. 6, p. 443-456, 1997.

ARTIGO 2

O conteúdo do segundo artigo está apresentado na íntegra conforme foi aceito para publicação na revista Journal of the European Optical Society: Rapid Publication (JEOS:RP).

Principal Component Analysis in the Spectral Analysis of the Dynamic Laser Speckle Patterns

Kleber Mariano Ribeiro[1] *, Roberto Alves Braga Júnior[1],
Graham William Horgan[2], Danton Diego Ferreira[1], Thelma Sáfadi[3]

[1] Engineering Department, Federal University of Lavras,
Lavras, Postal Address 3037, 37200-000, MG, Brazil.
Email addresses: klebermariano@gmail.com robbraga@deg.ufla.br
danton@deg.ufla.br
[2] Biomathematics and Statistics Scotland, Aberdeen,
AB21 9SB, Aberdeen, Scotland.
Email address: g.horgan@abdn.ac.uk
[3] Exact Science Department, Federal University of Lavras,
Lavras, Postal Address 3037, 37200-000, MG, Brazil.
Email address: safadi@dex.ufla.br

* Corresponding author: Tel. + 55 35 3829 1210

Abstract

Dynamic laser speckle is a phenomenon that interprets an optical patterns formed by illuminating a surface under changes with coherent light. Therefore, the dynamic change of the speckle patterns caused by biological material is known as biospeckle laser. Usually, these patterns of optical interference evolving in time are analyzed by graphical or numerical methods, and the analysis in frequency domain has also been an option, however involving large computational requirements which demands new approaches to filter the images in time. Principal component analysis (PCA) works with the statistical decorrelation of data and it can be used as a data filtering. In this context, the present work evaluated the PCA technique to filter in time the data from the biospeckle images aiming the reduction of time computer consuming and improving the robustness of the filtering. It was used 64 images of biospeckle in time observed in a maize seed. The images were arranged in a data matrix and statistically uncorrelated by PCA technique, and the reconstructed signals were analyzed using the routine graphical and numerical methods to analyze the biospeckle. Results showed the potential of the PCA tool in filtering the dynamic laser speckle data, with the definition of markers of principal components related to the biological phenomena and with the advantage of fast computational processing.

Keywords: Biospeckle; principal components; filter.

Análise de componentes principais na análise espectral dos padrões do *speckle* laser dinâmico

Resumo

O *speckle* laser dinâmico é um fenômeno que interpreta padrões ópticos formados ao iluminar uma superfície em movimento com uma luz coerente. Assim, a movimentação dinâmica dos padrões *speckle* causadas por materiais biológicos são conhecido como *biospeckle* laser. Normalmente, esses padrões de interferência óptica que evolui no tempo são analisados por meio de métodos gráficos ou de interpretações numéricas e a análise no domínio da frequência também tem sido uma opção, entretanto, envolvendo grandes recursos computacionais que demandam novas abordagens para filtrar as imagens no tempo. A análise de componentes principais (ACP) trabalha a descorrelação estatística de um conjunto de observações e pode ser utilizada para filtrar os dados. Neste contexto, no presente trabalho, avaliou-se a ACP na filtragem das imagens do *biospeckle* laser no tempo, buscando a redução do tempo de processamento computacional e a melhoria na robustez da filtragem. Foram utilizadas 64 imagens do *biospeckle* laser de uma semente de milho. As imagens foram organizadas em uma matriz de dados e descorrelacionadas estatisticamente por meio da técnica ACP, e os sinais reconstruídos foram analisados utilizando-se métodos gráficos e numéricos de análise do *biospeckle* laser. Os resultados mostraram o potencial da ferramenta ACP na filtragem dos dados do *speckle* laser dinâmico, com a definição de marcadores de componentes principais associados a fenômenos biológicos e com a vantagem do rápido processamento computacional.

Palavras-chave: *Biospeckle*; componentes principais; filtro.

1 Introduction

Dynamic laser speckle, also known as biospeckle when applied to biological materials, is an optical technique that processes the interference patterns formed when a material is illuminated by coherent light. It is a non-destructive technique and that has been validated as a tool for analysis and quantification of biological activity in the material under study (RABAL; BRAGA, 2008).

The term ‘biological activity’ expressed in the context of speckle does not present a precise definition and it is understood as the result of phenomena such as the Doppler effect, Brownian motion, variations of the refractive index, structural and molecular motions occurring in the material analyzed, among others (BRAGA et al., 2009; PASSONI et al., 2005).

Dynamic laser speckle technique has been used in several areas of research, such as in medicine, industrial processes and agriculture. Some examples of recent application of this tool are the works of Zakharov et al. (2009) imaging blood flow in rodent brain, Mavilio et al. (2010) studying the process of paint drying, Ansari and Nirala (2012) monitoring the maturation of Indian fruits, among others. In addition, the high number of applications of biospeckle brings with themselves the need for techniques of image and signal processing that can help in the interpretation, and offer additional information derived from these optical interference patterns.

The analysis of the data from optical interference patterns can be accomplished using graphical and numerical approaches (RABAL; BRAGA, 2008), in turn, Cardoso et al. (2011) associated graphical and numerical analysis using the frequency domain to create signatures and isolate some phenomena.

There are many studies analyzing the spectral information of the biospeckle data in different types of material and most use either Fourier or

wavelet transforms as tools to analyze the data in the frequency domain. Each method has distinct characteristics and properties. The Fourier transform is suited for stationary signals, which is not the case of the dynamic laser speckle as reported by Sendra, Murialdo e Passoni (2007), and this can compromise or limit the use of the technique.

Moreover, wavelet transforms have shown useful results in the segmentation of tissues, definition of frequency markers, and data filtering, as demonstrated by Sendra et al. (2005) in the assessment of apple damage and seed germination, as well as by Cardoso et al. (2011) studying seeds of maize and bean and animal cancer. However, the wavelets transform demands complex computational operations, as well as requiring some subjective choices such as that of a mother wavelet. Argoud, Azevedo e Mariano Neto (2004) claimed that the methodology for selection of the base function is not clear yet.

Despite the success of using Fourier and wavelet transform in frequency analysis, there are other filtering techniques in the literature which can be considered as alternative, overcoming the limitations of the methods used currently and providing information about this complex pattern of optical interference. Additionally, even though existing methods that have presented important contributions to dynamic speckle analysis, it may still be considered a complex problem and, therefore, alternative methods should be examined in order to undertake a thorough analysis.

In this context, statistical tools, such as principal component analysis, stand out as an option to analyze biospeckle data. As described by Rabal et al. (2012), the statistical techniques are indicated for data with random nature and with time evolution, which is the case with dynamic laser speckle. In the basis of the dynamic laser speckle phenomenon, the laser light scattering in a dynamic way can be related to a multiple range of physical and chemical phenomena that can be the considered the key factor to understand and correlate the dynamic

scattered output with the analyzed phenomenon itself (BERNE; PECORA, 1976).

Principal component analysis - PCA - is a classic technique for multivariate statistical analysis of data, which consists essentially in transforming orthogonally a set of correlated observed variables into a new set of uncorrelated variables, called the principal components. The transformation is accomplished by calculation of the eigenvalues and eigenvectors of the data covariance matrix (JUNG; SEN; MARRON, 2012; SILVA; MINIM; RIBEIRO, 2005; ZHANG et al., 2010).

PCA has been used in many applications as a tool to reduce the data volume with the least possible loss of information, classification and clustering of data, extraction and identification of patterns and also filtering of signals (NIELSEN et al., 2002; RINGNÉR, 2008). Papers presented by Chen and Qian (2011) and Souza Filho and Dinniss (1997) confirm the potential of principal component analysis as filtering technique.

In this context, the present work aims at proposing the usage of this multivariate statistical tool as an alternative for the spectral analysis of the dynamic laser speckle signal. The proposed method consists in applying the PCA technique as a preprocessing tool for biospeckle signal analysis. The combination of PCA and existing methods like Fujji and GD is shown and promising results have been achieved for real data.

The next section reviews the background theory of the methods used in this work. The first subsection describes the technique of principal component analysis, and sections 2.2 and 2.3 relates the Fujii and GD methods of graphical analysis of biospeckle patterns while the last part presents the use of the logarithm unit to carry out numerical interpretation of the data.

2 Theory

2.1 Principal component analysis (PCA)

Principal component analysis (PCA) is a multivariate statistical technique that describes a set of correlated observations in terms of a new set of orthogonal and uncorrelated variables, called principal components, which are linear combinations of the original variables (ABDI; WILLIAMS, 2010).

The transformation of the data to the PCA domain is performed by the decomposition of the covariance matrix into eigenvalues and eigenvectors, and this technique has been used in several application areas under different approaches, such as use as a denoising method, and with the advantage of being a convenient tool from a computational viewpoint (JUNG; SEN; MARRON, 2012; XANTHOPOULOS; PARDALOS; TRAFALIS, 2013; ZHANG et al., 2010).

Principal components analysis begins with the organization of the data in a matrix **X** of dimension *M* x *N*, which *M* represents the number of observations and *N* the number of variables, as illustrated in the Equation 1.

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{11} & \mathbf{x}_{12} & \cdots & \mathbf{x}_{1N} \\ \mathbf{x}_{21} & \mathbf{x}_{22} & \cdots & \mathbf{x}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{M1} & \mathbf{x}_{M2} & \cdots & \mathbf{x}_{MN} \end{bmatrix} \quad (1)$$

In order to avoid points distant from the data center having a greater influence than nearby points, (as would arbitrarily occur when data are in different units), the mean of each variable is removed from data. This process is called centralization of data and it is represented by Equation 2.

$$\mathbf{y}_i = \mathbf{x}_i - \mu(\mathbf{x}_i) \quad (2)$$

where $\mathbf{y_i}$ correspond the data vectors centralized around of the mean, $\mathbf{x_i}$ are the *N* sample vectors studied and $\mu(\mathbf{x_i})$ consists of the mean of the sample vectors, which can be calculated by the Equation 3.

$$\mu(\mathbf{x_i}) = \frac{1}{m}\sum_{i=1}^{m}\mathbf{x_i}(j) \tag{3}$$

The variables or sample vectors, as it is called by Zhang et al. (2010), are each column of **X** and are expressed mathematically in the Equation 4.

$$\mathbf{x_i} = [\mathbf{x}_{1i} \quad \mathbf{x}_{2i} \quad \cdots \quad \mathbf{x}_{mi}]^{\mathbf{T}} \tag{4}$$

The data matrix organized and centralized on the mean is used to compute the covariance matrix as shown in the Equation 5.

$$\mathbf{C_Y} = \mathbf{E}(\mathbf{Y} \cdot \mathbf{Y}^{\uparrow}\mathbf{T}) \tag{5}$$

which **Y** and $\mathbf{Y^T}$ are, in order, the data matrix centered on the mean and its transpose, and $\mathbf{C_Y}$ is the covariance matrix.

The diagonal elements of $\mathbf{C_Y}$ represent the statistical variance while the off-diagonal elements characterize the covariance between variables. Null diagonal covariance means that the random variables are uncorrelated (HADFIELD, 2010), though we cannot affirm about the statistical independence for the biospeckle, since the speckle patterns in time cannot be represented by a

Gaussian behavior. Furthermore, the covariance matrix is real and symmetric, which permits us to decompose $\mathbf{C_Y}$ into a set of eigenvalues and orthogonal eigenvectors (JUNG; SEN; MARRON, 2012) using the Equation 6.

$$\mathbf{C_Y = V . \Lambda . V^T} \quad (6)$$

where $\mathbf{V} = [\emptyset_1\ \emptyset_2\ \ldots\ \emptyset_m]$ is a M x M orthonormal eigenvectors matrix and $\mathbf{\Lambda} =$

diag $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$ is the diagonal matrix of eigenvalues, which $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$.

The eigenvectors represent the contribution to each of the original axes to the composition of the new axes, the principal components. The eigenvalues, in turn, are associated with the original amount of the variance described by each of the eigenvectors (BATINA; HOGENBOOM; WOUDENBERG, 2012; SILVA; MINIM; RIBEIRO, 2005).

The last step of the analysis is the construction of the uncorrelated data matrix that is also known as the principal component scores, and which is formed by the product of the orthonormal eigenvector matrix **V** and the data matrix organized and centralized on the mean **Y**, as expressed in the Equation 7.

$$\mathbf{PC = V^T . Y} \quad (7)$$

which **PC** is the matrix of uncorrelated principal component scores.

From the data in the PCA domain, it is possible to extract signal characteristics, and according to Zhang et al. (2010), the signal and the noise of a data set can be better distinguished in the PCA domain, since the signal energy and noise energy will concentrate in different subsets of the uncorrelated data. Because of this ability, PCA is referred to as a statistical data filtering method.

We can also consider the inverse PCA transform, which is used to back transform the principal component scores (uncorrelated data), thereby reconstructing the original dataset. Equation 8 presents the mathematical expression of the inverse PCA transform.

$$\mathbf{X} = (\mathbf{V} \,.\, \mathbf{PC}) + \mu(\mathbf{X}) \tag{8}$$

The inverse PCA transformation is a useful operation since reconstruction of original data with only some specific PCs, discarding the rest of them, can enhance important features not previously easily seen in the data and/or remove the contribution of undesirable features such as noise. Such an operation is also widely used for data compaction.

2.2 Fujii method for biospeckle

One way to analyze the interference patterns of the dynamic laser speckle is the use of graphical methods, which display maps of the spatial variability of the biological activity of the material studied, and the Fujii method is a tool that fits this classification.

Fujii et al. (1987) presented this technique in the analysis of a sequence of dynamic laser speckle images. The method consists of the summation of the weighted differences between each image and the subsequent image (Equation 9).

$$\mathbf{Fujii}\ (\mathbf{x},\mathbf{y}) = \sum_{k=1}^{N} \left| \frac{\mathbf{I_k}(\mathbf{x},\mathbf{y}) - \mathbf{I_{k+1}}(\mathbf{x},\mathbf{y})}{\mathbf{I_k}\ (\mathbf{x},\mathbf{y}) + \mathbf{I_{k+1}}\ (\mathbf{x},\mathbf{y})} \right|$$

(9)

where Fujii(x,y) is the resulting image and $I_k(x,y)$ is the gray level in the coordinates x and y of the k^{th} image.

The result is a new image, in which it is possible to visualize the spatial variability of biological activity. Regions of high activity are represented in light tonalities while dark areas illustrate regions of low biological activity.

In addition, a feature of the Fujii method is the amplification of movements in darkest areas, making the images clearer when compared with other approaches such as the generalized difference method (BRAGA et al., 2009).

2.3 Generalized difference method (GD)

The generalized difference approach was introduced by Arizaga et al. (2002) as an alternative to the Fujii technique. The method generalized the summation of the differences of the intensities along the whole sequence of images and the weighting factor was eliminated (Equation 10).

$$\mathbf{DG}\ (\mathbf{x},\mathbf{y}) = \sum_{k} \left[\sum_{l} \left[\left| \mathbf{I}_{k}\ (\mathbf{x},\mathbf{y}) - \mathbf{I_{k+l}}(\mathbf{x},\mathbf{y}) \right| \right] \right]$$

(10)

where GD(x,y) is the resulting image, and $I_k(x,y)$ is the pixel intensity located in the coordinates x and y of the k^{th} image.

2.4 Logarithm unit

Comparison between the results before and after the adoption of the filtering promoted by PCA of the biospeckle data were carried out by means of the logarithm scale, in particular by using the decibel scale.

The decibel (dB) is defined by a logarithmic relationship that expresses the ratio of a value being measured with a reference (SPECHT et al., 2009). Equation 11 describes mathematically the logarithmic unit in decibels.

$$dB = 10\ LOG_{10}\left(\frac{W_1}{W_2}\right) \quad (11)$$

where dB is the result of the logarithmic relationship expressed in dB, and W_1 and W_2 are the energies of the signal studied and the reference signal, respectively.

Negative dB results indicate that the data processing promoted attenuation of the signal energy, whereas positive values express energy gain after application of the analysis.

The energy of a discrete signal k[n] is the summation of squares over time as shown in Equation 12.

$$W_k = \sum_{n=1}^{N} [\,|k[n]|\,]^2 \quad (12)$$

3 Materials and methods

In order to evaluate the proposed method, a database from a maize fruit illuminated by laser was used (BRAGA et al., 2001), and the approach adopted

was the back-scattering. In the back-scattering approach adopted, the laser beam reached the object in a plane and the scattered light that returned from the sample was collected by a CCD camera in the same side of the plane where the laser was positioned. The images in time were acquired in the CCD were processed by image analysis and by statistical procedures in order to quantify or qualify the biospeckle phenomenon. In this work, the database from the illuminated maize had 64 gray level images, each with a resolution of 490 by 256 pixels, and they were collected using the experimental setup with a time rate of 0.08 seconds. The time rate adopted was enough to acquire all the relevant frequencies in the signal, since the biological activity of the maize seed is below 6 Hz (CARDOSO et al., 2011; SENDRA et al., 2005). The images were collected in order to get a sufficient focus of the maize, as well as with a clear definition of the speckle grains, avoiding the saturation of the light or the sub-exposition on the whole sample.

Each image of the database was concatenated and the signals formed were vertically arranged side by side following the sequence of the images. Figure 1 illustrates the construction of the concatenated images in the data matrix **X**.

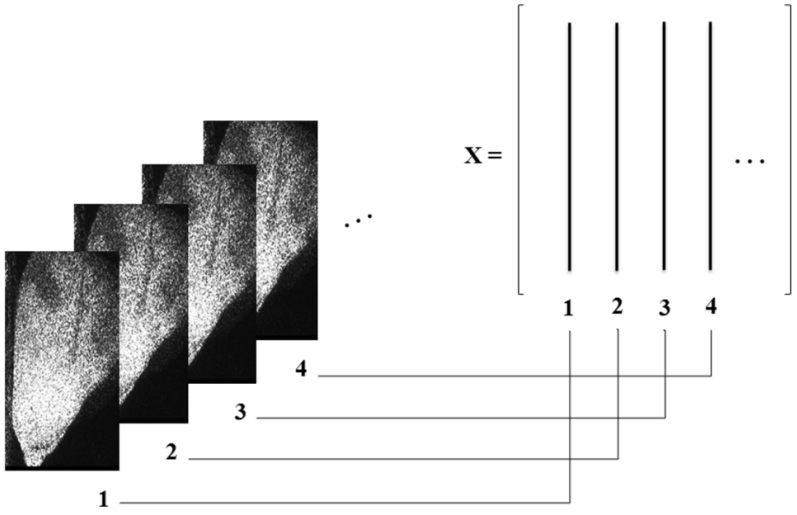


Figure 1 Organization of the concatenated images in a new data matrix

The data matrix **X** was transformed to a set of statistically uncorrelated coordinates by the PCA technique, converting the original data to the PCA score domain. In order to study the contribution of each principal component to the composition of the original signal, some principal components were eliminated before application of the inverse PCA transform, and this selection process of the PC's was performed using three approaches:

a) Emphasis on the first *g* principal components;
b) Using only the last *h* PC's;
c) A random choice of some PC's.

After selecting PC's, the inverse PCA transform was obtained. Then, the inverse process of concatenating image was done. Afterward, the reconstructed data were analyzed graphically by the Fujii and GD methods. Figure 2 summarizes the proposed methodology in a flow chart.

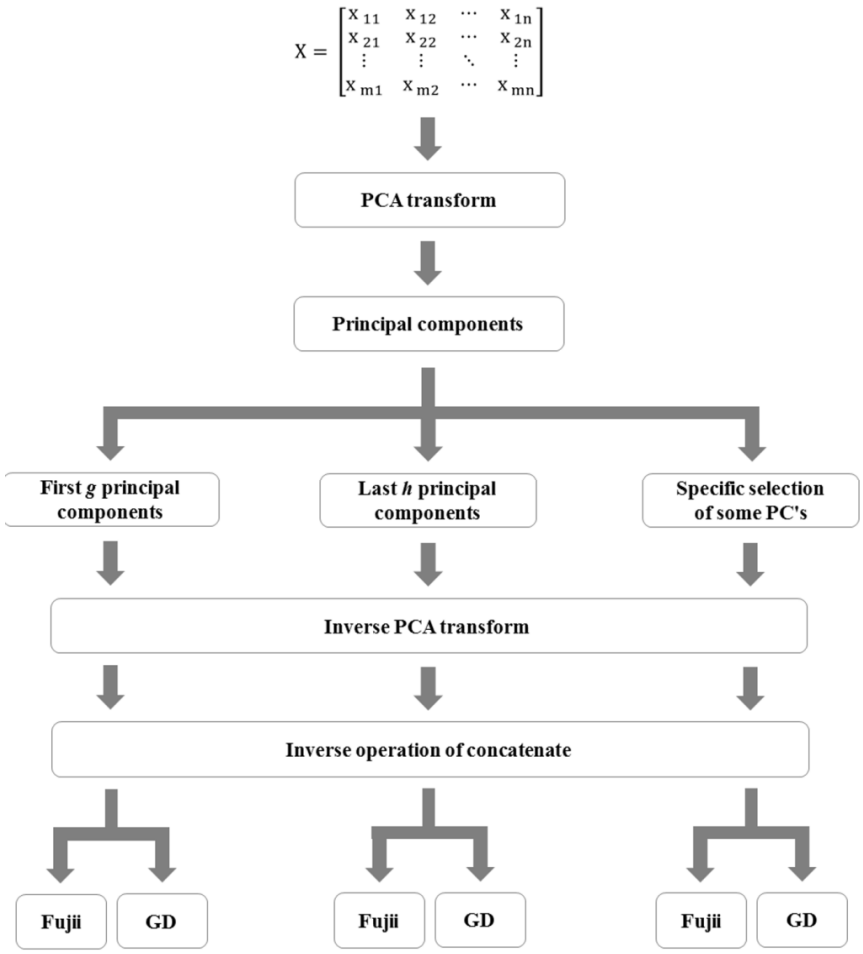


Figure 2 Methodology used

In order to carry out a numerical analysis and to assist the interpretation of the processed data, one line each from the Fujii and GD images resulting of the graphical methods was selected, as illustrated in Figure 3. Each line was shown in the same figure to compare its behavior in terms of amplitude. In addition, quantitative analyses were also carried out by calculating the energy of the chosen lines on the dB scale.

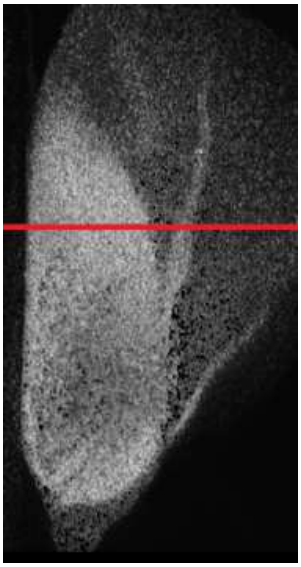
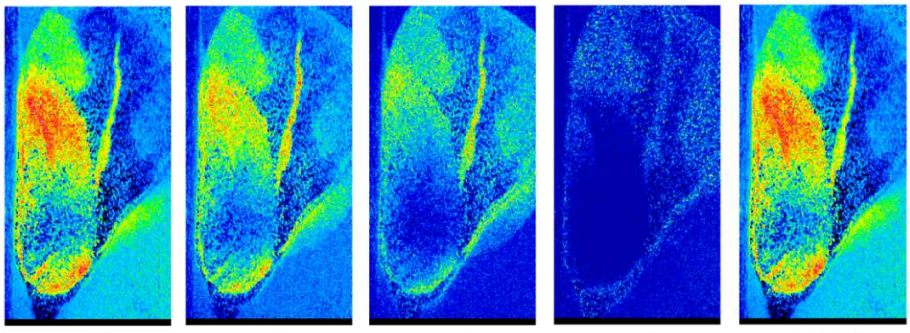


Figure 3 Position of the line selected in the resulting images

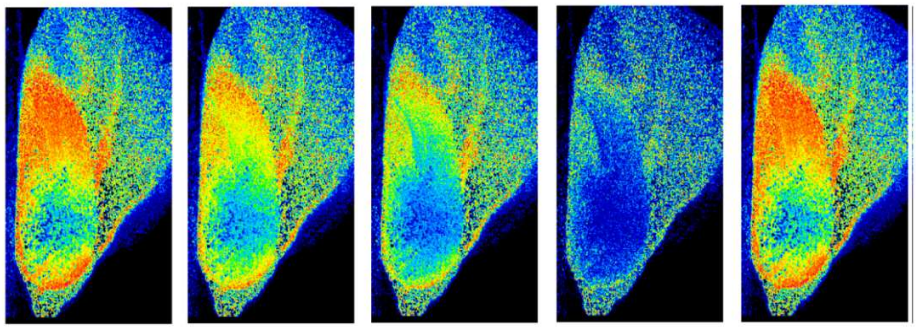
4 Results and discussion

4.1 Signal reconstruction using the first *g* principal components

Figure 4 illustrates the biospeckle activity maps of the maize fruit analyzed using the PCA technique, in which we used the first *g* principal components in the reconstruction process of the signal.



(A)



(B)

Figure 4 Fujii (A) and GD (B) images performed by PCA analysis with the signal reconstruction using the first g PC's and the correspondent original images

The areas of high biological activity are illustrated by the light gray in the images whereas the dark shades are linked to low activity (in pseudocolors red means light gray and blue means dark gray). Furthermore, the images named as 'Original' presented in the Figures 4A and 4B are, respectively, the Fujii and GD graphics of the biospeckle of the maize fruit unprocessed with the PCA technique, and they are the reference images for the data analysis.

The total reconstruction of the data, using all 64 principal components in the inverse transform, presented images visually identical to the reference in both graphical methods, in Figure 4A and Figure 4B as expected. Moreover, decreasing of the number of the first PC's used in the inverse PCA transform attenuated the embryo information and kept the endosperm separation, so filtering the data and segmenting the tissues.

Figure 5 shows the selected rows in the GD images where it is possible to observe the filtering effect in the tissues of the embryo and endosperm for different values of g, and Table 1 presents the results of the numerical analysis,

based on the data from Figure 5. In the embryo it was expected the highest activity since there are live tissues and water movement contributing to the Doppler beating of the scattered light, though in the endosperm the expected activity should be lower than in the embryo since there is no presence of live tissues in there, but only a reserve of nutrients (BRAGA et al., 2001). Therefore, the outputs presented the ability to tag that difference with different levels depending on the g values of PCA adopted.

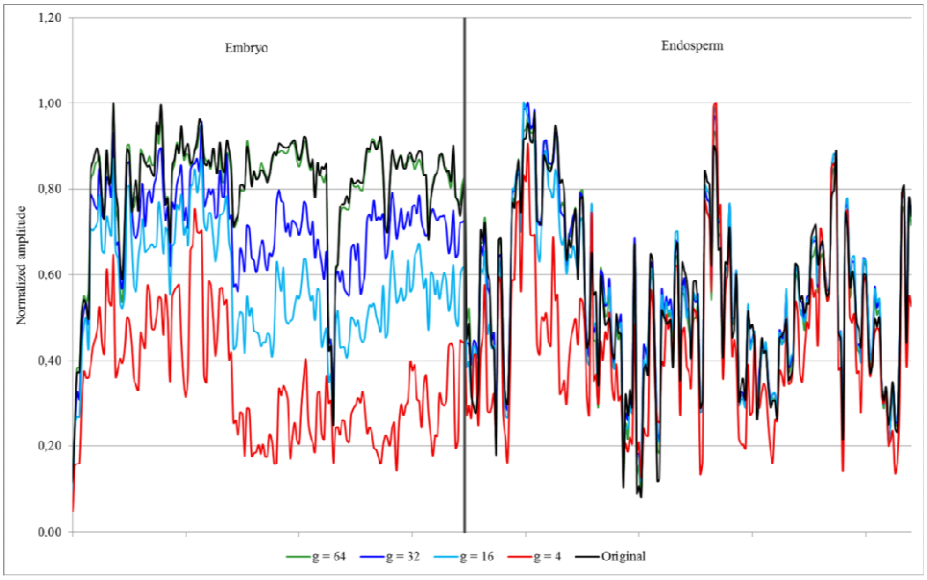


Figure 5 Filtering effect for different values of *g* used in the inverse PCA transform

Table 1 Decibels and correlation index of the signals reconstructed using the first *g* principal components and of the original signal

---	Maize fruit		Embryo tissue		Endosperm tissue	
	dB	R^2	dB	R^2	dB	R^2
g = 64	-0.05	0.99	-0.05	1.00	-0.05	0.96

g = 32	-0.72	0.91	-1.29	0.94	0.15	0.96
g = 16	-1.66	0.70	-2.87	0.80	-0.07	0.94
g = 4	-4.28	0.25	-6.86	0.41	-1.70	0.67

Negative decibel values in Table 1 indicate attenuation of the energy and positive values denote gain of energy in the acquired line. Null values of decibels mean that the two signals compared have the same energy.

The dB values (Table 1) oscillated between 0.05 and 6.86 dB for embryonic tissue whereas for endosperm tissue they kept close to zero, except for g equal to 4, which presented an attenuation of 1.70 dB. These results show numerically a higher attenuation of the embryo data relative to endosperm. Such attenuation is shown in the Figure 5, where the embryo signal exhibits large changes for the different g values, decreasing the normalized amplitude with the decrease of g, whilst the endosperm signal remained near the original curve. In addition, the correlation index presented lower fluctuations in the values for the endosperm tissue, which also demonstrates preservation of the characteristics of the endosperm signal and modifications of the embryo signal. The better estimation of the level of those noise and variations in the signal can be addressed by some techniques (SKIPETROV et al., 2010) which can validate the filtering outputs at each case.

Kaiser (1960) proposed a statistical criterion to define the optimal number of principal components to represent a dataset. Applying this criterion to the database used here, the optimal number of principal components was 16, which explain 94% of the variance of the data. Therefore, setting g equal to 4, which describes 89% of the data variance, is considered too low to represent the dataset by the criterion of Kaiser (1960). It explains the achieved attenuation in

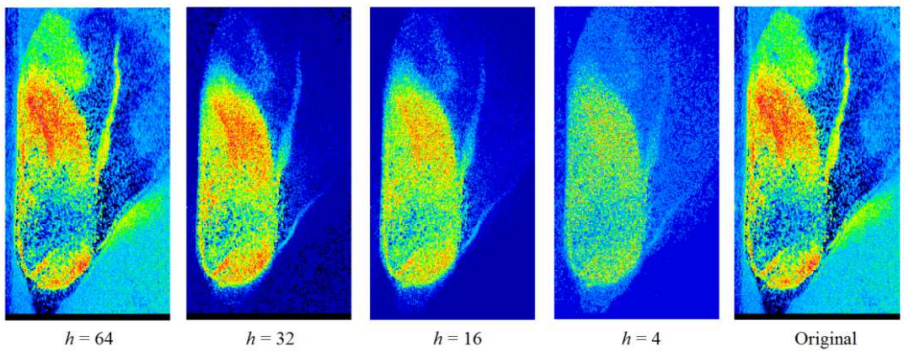
the amplitude of the endosperm representation signal and the low correlation index.

In this context, the inverse PCA transform using only the first *g* PC's implements a low pass filter in the time expression of the images, attenuating amplitudes associated to the high frequencies (embryo) and preservation of the low frequencies, which represents the endosperm activity.

According to Scalassara, Barin e Maciel (2004), the first principal components contain information of a large proportion of the signal variance and the last contain basically the noise variance (high frequency signal). Consequently, the use of the first *g* PC's produces a data filtering with elimination of high frequency activities, related to the images varying in time domain, which means concerning to the temporal Fourier transform

4.2 Signal reconstruction using the last *h* principal components

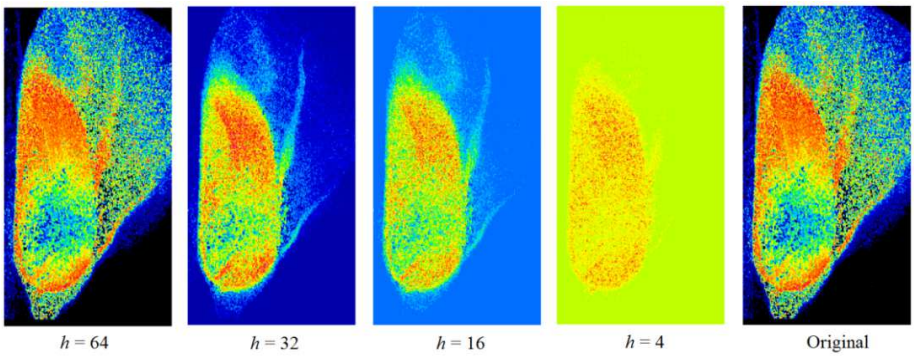
Figure 6 presents the results of the PCA analysis based on signal reconstruction using the last *h* PC's, in which are illustrated the Fujii and GD maps, and the selected line in the graphics output with the behavior of the signals for different values of *h*, respectively.



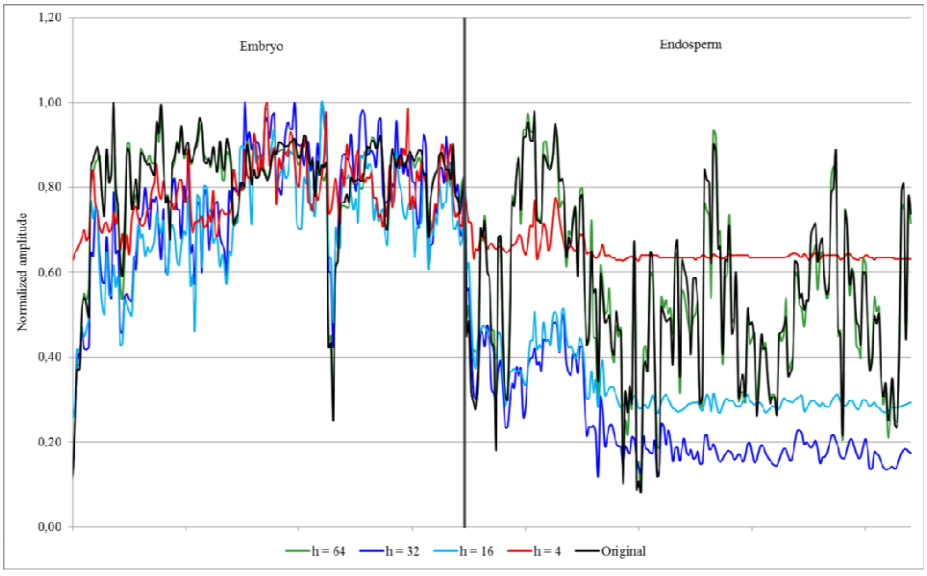
(A)

Figure 6 Biological activity according to Fujii (A) and GD (B) techniques and the filtering effect in the embryo and endosperm tissues for different numbers of PC's used in the signal reconstruction (C)

(…continue…)



(B)



(C)

Figures 6A and 6B show similar graphical results to Fujii and GD methods, with maps visually identical to original pictures for $h = 64$. In addition, we note the emphasis in the embryonic part by decreasing the number of PC's.

These results show that the preprocessing with PCA using the last h principal components served as a high pass filter, highlighting the high frequencies, such as in the embryonic portion, and filtering of the lowest frequencies, which are linked to the biological activity of the endosperm, as discussed by Cardoso et al. (2011).

Quantitative results point out higher attenuation in the endosperm activities for low values, achieving -6.97 dB ($h = 32$) and a correlation index of 0.19 ($h = 16$), summarized in the Table 2. Figure 6C allows us to visualize the filtering effect in the endosperm signal, where the reconstructed lines present amplitudes considerably different from those of the reference signal, except for $h = 64$, which corresponds to the total reconstruction of the original signal.

Table 2 Numerical analysis for signals reconstructed using the last h principal components

---	Maize fruit		Embryo tissue		Endosperm tissue	
	dB	R^2	dB	R^2	dB	R^2
$h = 64$	-0.05	0.99	-0.05	1.00	-0.05	0.96
$h = 32$	-1.89	0.58	-0.49	0.83	-6.97	0.28
$h = 16$	-1.97	0.52	-0.99	0.74	-4.65	0.19
$h = 4$	0.42	0.42	0.06	0.46	1.02	0.21

Otherwise, the results for the embryo signal presented low oscillation of the decibel values, where the highest attenuation achieved was -0.99 dB for h = 16. The correlation indices (Table 2) also kept high values for different h, except

for the last 4 PC's. These results show the preservation of the information retained in the high frequencies, thereby performing a high pass filter by PCA.

4.3 Random selection of some principal components to application of the inverse PCA transform

Figure 7 illustrates four GD images in which the signals were reconstructed using a small and random number of principal components.

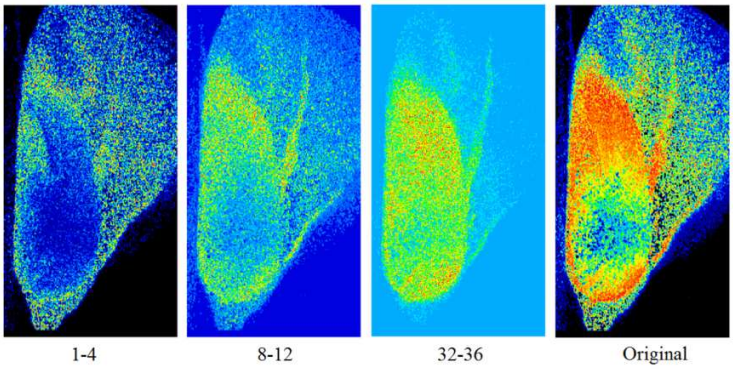


Figure 7 GD images resulting of the signal reconstruction using a short and random number of principal components

The goal of using this specific and random number of principal components is to combine both high and low pass filters obtained from PCA in order to improve the results of Fujii and GD methods. Use of high pass filters, low pass filters or band pass filters allows us to define small spectral ranges in which the characteristics of biological, physical or chemical phenomena are concentrated and occurring more intensely, the frequency markers as it is called by Cardoso et al. (2011) and Sendra et al. (2005).

In principal component analysis, the terminology used are based on the principal component scores and loadings, and not frequency, but the signal

reconstruction using random and specific number of PC's opens an option to define markers of principal components and associate them to biological phenomena, as presented in the Figure 7. The characteristic of the biospeckle signal allowed the use of the PCA as a filtering tool, based in the advantage of performing as a non-parametric and adaptive method, which is desirable for practical implementations. In addition, the PCA filtering presents the advantage of the reducing of the computational time consuming which is relevant in the quasi-online applications.

The first image presented in the Figure 7 is the signal reconstruction using the PC's from 1 to 4 followed by the GD processing, which highlighted information from the endosperm and filtered the embryo signals. Thus, the PC's interval 1-4 can be considered as a marker of principal component for biological activity of the endosperm tissue of the maize fruit.

The same perception occurs in the third image of the Figure 7, in which the PC's from 32 to 36 also are markers of principal components but for biological activity of the embryonic tissue. The result of the analysis emphasized the embryo and attenuated information from endosperm tissue in the GD image.

Finally, the GD image constructed using the signals reconstructed from 8 to 12 PC's (second image of the Figure 7), improved the quality of the output however without any mark.

5 Conclusion

Principal component analysis was proposed as tool to spectral analysis of dynamic laser speckle data and showed to be a powerful tool to analyze biospeckle data, allowing the implementation of filters with different frequency pass band ranges for data analysis concerning to the temporal Fourier transform.

The proposed PCA based method allowed the decomposition of biological activity in the endosperm and embryo of the maize seed example

used, with advantage of a blind source separation technique with fast computational processing, in which the orthogonal basis functions used for data decomposition are statistically optimum fitted. In addition, in comparison to conventional low-pass and high pass filters, the PCA based filtering has the advantage of performing as a non-parametric and adaptive method, which is desirable for practical implementations.

The proposed method also provided tissues segmentation of the biological materials, improving the visual quality of the final images and the definition of markers of principal components of the biological phenomena, which supports its potential for biospeckle data analysis.

Acknowledgements

This work was partially financed by CNPq, Fapemig, Capes, Finep in Brazil, and and partly supported by the Scottish Government Rural and Environment Science and Analytical Services division.

REFERENCES

ABDI, H.; WILLIAMS, L. J. Principal component analysis. **Wiley Interdisciplinary Reviews: Computational Statistics**, New York, v. 2, n. 4, p. 433-459, 2010.

ANSARI, M. D.; NIRALA, A. K. Biospeckle activity measurement of Indian fruits using the methods of cross-correlation and inertia moments. **Optik-International Journal for Light and Electron Optics**, Amsterdam, v. 124, n. 15, p. 2180-2186, Aug. 2012.

ARGOUD, F. I. M.; AZEVEDO, F. M. de; MARIANO NETO, J. Comparative study concerning to wavelet functions and its different applicabilities to pattern recognition in electroencephalogram. **Revista Brasileira de Engenharia Biomédica**, Campinas, v. 20, n. 2/3, p. 49-59, 2004.

ARIZAGA, R. et al. Display of local activity using dynamical speckle patterns. **Optical Engineering**, Redondo Beach, v. 41, n. 2, p. 287-294, 2002.

BATINA, L.; HOGENBOOM, J.; WOUDENBERG, J. G. J. van. Getting more from PCA: first results of using principal component analysis for extensive power analysis. In: _____. **Topics in cryptology - CT-RSA 2012**. Berlin: Springer, 2012. p. 383-397.

BERNE, B. J.; PECORA, R. **Dynamic light scattering:** with applications to chemistry, biology, and physics. New York: J. Wiley, 1976. 384 p.

BRAGA, R. A. et al. Live biospeckle laser imaging of root tissue. **European Biophysics Journal**, London, v. 38, n. 5, p. 679-686, 2009.

BRAGA, R. A. et al. Potencial do bio-speckle laser para avaliação da viabilidade de sementes. **Ciência e Agrotecnologia**, Lavras, v. 25, n. 3, p. 645-649, maio/jun. 2001.

CARDOSO, R. R. et al. Frequency signature of water activity by biospeckle laser. **Optics Communications**, Amsterdam, v. 284, n. 8, p. 2131-2136, 2011.

CHEN, G.; QIAN, S. E. Denoising of hyperspectral imagery using principal component analysis and wavelet shrinkage. **Geoscience and Remote Sensing, IEEE Transactions**, New York, v. 49, n. 3, p. 973-980, 2011.

FUJII, H. et al. Evaluation of blood flow by laser speckle image sensing. **Applied Optics**, New York, v. 26, n. 24, p. 5321-5325, 1987.

HADFIELD, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. **Journal of Statistical Software**, Los Angeles, v. 33, n. 2, p. 1-22, 2010.

JUNG, S.; SEN, A.; MARRON, J. S. Boundary behavior in high dimension, low sample size asymptotics of PCA. **Journal of Multivariate Analysis**, New York, v. 109, p. 190-203, Aug. 2012.

KAISER, H. F. The application of electronic computers to factor analysis. **Educational and Psychological Measurement**, Thousand Oaks, v. 20, p. 111-117, 1960.

MAVILIO, A. et al. Characterization of a paint drying process through granulometric analysis of speckle dynamic patterns. **Signal Processing**, Amsterdam, v. 90, n. 5, p. 1623-1630, 2010.

NIELSEN, T. O. et al. Molecular characterisation of soft tissue tumours: a gene expression study. **The Lancet**, London, v. 359, n. 9314, p. 1301-1307, 2002.

PASSONI, I. et al. Dynamic speckle processing using wavelets based entropy. **Optics Communications**, Amsterdam, v. 246, n. 1, p. 219-228, 2005.

RABAL, H. J.; BRAGA, R. A. **Dynamic laser speckle and applications**. Boca Raton: CRC, 2008. 304 p.

RABAL, H. J. et al. Q-statistics in dynamic speckle pattern analysis. **Optics and Lasers in Engineering**, London, v. 50, n. 6, p. 855-861, 2012.

RINGNÉR, M. What is principal component analysis? **Nature Biotechnology**, New York, v. 26, n. 3, p. 303-304, 2008.

SCALASSARA, P. R.; BARIN, C. S.; MACIEL, C. D. Minimização de ruído eletroquímico usando processamento digital de sinais. **Semina: Ciências Exatas e Tecnológicas**, Passo Fundo, v. 25, n. 2, p. 135-144, 2004.

SENDRA, G. H. et al. Decomposition of biospeckle images in temporary spectral bands. **Optics Letters**, Washington, v. 30, n. 13, p. 1641-1643, 2005.

SENDRA, H.; MURIALDO, S.; PASSONI, L. Dynamic laser speckle to detect motile bacterial response of Pseudomonas aeruginosa. **Journal of Physics: Conference Series**, Moscow, v. 90, n. 1, 2007. Disponível em: <http://iopscience.iop.org/1742-6596/90/1/012064>. Acesso em: 10 jan. 2014.

SILVA, A. F.; MINIM, V. P. R.; RIBEIRO, M. M. Análise sensorial de diferentes marcas comerciais de café (*Coffea arabica L.*) orgânico. **Ciência e Agrotecnologia**, Lavras, v. 29, n. 6, p. 1224-1230, nov./dez. 2005.

SKIPETROV, S. E. et al. Noise in laser speckle correlation and imaging techniques. **Optics Express**, Washington, v. 18, n. 14, p. 14519-14534, 2010.

SOUZA FILHO, C. R.; DINNISS, A. Periodic noise suppression techniques applied to remote sensing images. **Boletim IG-USP. Série Científica**, São Paulo, v. 28, p. 23-62, 1997.

SPECHT, L. P. et al. Noise evaluation using the SPBI (Statistical Pass-By Index) for different pavements. **REM: Revista Escola de Minas**, Ouro Preto, v. 62, n. 4, p. 439-445, 2009.

XANTHOPOULOS, P.; PARDALOS, P. P. M.; TRAFALIS, T. B. **Robust data mining**. New York: Springer, 2013. 71 p.

ZAKHAROV, P. et al. Dynamic laser speckle imaging of cerebral blood flow. **Optics Express**, Washington, v. 17, n. 16, p. 13904-13917, 2009.

ZHANG, L. et al. Two-stage image denoising by principal component analysis with local pixel grouping. **Pattern Recognition**, Ezmsford, v. 43, n. 4, p. 1531-1549, 2010.

ARTIGO 3

O terceiro artigo está submetido para avaliação na Revista Brasileira de Biometria e segue formatado nas instruções da Associação Brasileira de Normas Técnicas (ABNT), NBR 6022.

Independent Component Analysis Preprocessing the Graphical Output of the Dynamic Laser Speckle Data

Kleber Mariano Ribeiro[1], Roberto Alves Braga Júnior[1],
Danton Diego Ferreira[1], Thelma Sáfadi[2],
Graham Horgan[3], Diego Eduardo Costa Coelho[1] *

[1] Engineering Department, Federal University of Lavras,
Lavras, Postal Address 3037, 37200-000, MG, Brazil.
Email addresses: klebermariano@gmail.com robbraga@deg.ufla.br
danton@deg.ufla.br diegocoelho.ufla@gmail.com
[2] Exact Science Department, Federal University of Lavras,
Lavras, Postal Address 3037, 37200-000, MG, Brazil.
Email address: safadi@dex.ufla.br
[3] Biomathematics and Statistics Scotland, James Hutton Institute,
University of Aberdeen, AB21 9SB, Aberdeen, Scotland.
Email address: g.horgan@abdn.ac.uk

* Corresponding author: Tel. + 55 35 3829 1210

Abstract

This paper proposes the use of independent component analysis (ICA) as a preprocessing tool of the graphical outputs from the analysis of the dynamic laser speckle data. It was used 64 images in gray levels from the dynamic laser speckle of a maize fruit, which have been rearranged in a data matrix and transformed into independent statistically components by FastICA algorithm. The independent components were graphically analyzed by the traditional methods designed to create maps of activity. The results showed that the combination of the ICA with the generalized difference method provided quality improvement of the images that enhanced the interpretation of the activity maps. Otherwise, the combination of the ICA with the Fujii method didn't present improvements since the Fujii method is vulnerable to the changes of the order of the images caused by the independent components.

Key-words: Biospeckle laser; Fujii; generalized differences; blind sources separation.

Pré-processamento das saídas gráficas dos dados do *speckle* laser dinâmico por meio da análise de componentes independentes

Resumo

Neste trabalho propõe-se o uso da análise de componentes independentes (ACI) como uma ferramenta de pré-processamento das saídas gráficas das análises dos dados do *speckle* laser dinâmico. Foram utilizadas 64 imagens em níveis de cinza do *speckle* laser dinâmico de um fruto de milho, que foram reorganizados em uma matriz de dados e transformadas em componentes estatisticamente independentes por meio do algoritmo FastICA. As componentes independentes foram graficamente analisadas pelos métodos tradicionais destinados a criar mapas de atividade. Os resultados mostraram que a combinação do ACI com o método de diferenças generalizadas proporcionou melhoria na qualidade das imagens e facilitou a interpretação dos mapas de atividade. Por outro lado, a combinação da ACI com o método de Fujii não apresentou melhorias, uma vez que o método de Fujii é vulnerável às mudanças na ordem das imagens componentes independentes.

Palavras-chave: *Biospeckle* laser; Fujii; diferenças generalizadas; separação cega de fontes.

1 Introduction

The dynamic laser speckle is a technique that processes the optical interference patterns formed when a coherent light illuminates a surface with roughness higher than its wavelength (RABAL; BRAGA, 2008). This technique has been validated as a tool of analysis and quantification of structural movements and/or molecular occurring in the analysis material.

When applied to biological surfaces the dynamic laser speckle can be called simply of biospeckle laser, and works such as Li, Tai e Nie (2011) in the process of sedimentation of silver chloride, Murialdo et al. (2012) with the detection and differentiation of bacteria and fungi and Amaral et al. (2013) in the study of biological maturation of meats, illustrate the potential of application of the technique of the dynamic laser speckle in different areas of the knowledge.

The result of the application of biospeckle laser in a given surface consists of a complex signal (COSTA et al., 2010), result of phenomena such as Brownian motion, Doppler effect, changes in the refractive index, molecular and structural motions (BRAGA et al., 2009; PASSONI et al., 2005) among others.

The biospeckle laser is analyzed with techniques of image processing and statistical treatment, since the visual observation does not allow quantifying it (RABAL; BRAGA, 2008), and the best methodology for the analysis of data from dynamic laser speckle is determined by the nature of the signals collected (BRAGA et al., 2009) and can be graphic or numeric.

The graphic analysis are digital processing of the images containing the interference patterns that result in maps with the spatial variability of biological activity, whereas the numerical analysis assign values to biological activities allowing evaluating the activity of the studied material at different instants. Numerical analyses have been used in order to provide more objective results in

comparison to graphical analysis. On the other hand, both graphical and numerical analysis should be preferred to make the results more consistent.

There are several graphical methods of analysis of the biospeckle laser available in the literature, each with its own characteristics, and Fujii and generalized differences (GD) techniques are two important approaches used in a large number of scientific papers.

However, works such as Braga et al. (2005) in the analysis of fungi on bean seeds and Braga et al. (2009) in the study of the roots of *Coffea arabica* and *Eucayptus grandis* show that there are situations in which both graphical methods of analysis present final images with lots of unwanted information and high heterogeneity of the activity in the same biological tissue.

The variability of the biological activity in the same tissue may be caused by several factors, for example, an experimental setup not appropriate or deficient, the complexity involved in the studied material and even the inexperience of the researcher, and amplifying the differences used by the Fujii and GD methods can guide areas of high and low activity for different levels, allowing the segmentation of the phenomena.

The maximization of differences can be performed by signal analysis techniques, through digital image processing or using statistical tools, which Rabal et al. (2012) stands out as an interesting method for analysis of dynamic laser speckle seen their random nature and evolution in time.

Costa et al. (2010), in the segmentation of biospeckle data from bovine semen and Rabal et al. (2012) optimizing methods that measure the biological activity of dynamic laser speckle, are examples of work available in the literature in which biospeckle data were analyzed by statistical tools, and the results were favorable to the quality improvement in the results.

Among the existing statistical techniques, the independent component analysis (ICA) has drawn the attention of the academic community by the

number of applications which offers (BEDOYA; BERMEJO; CABESTANY, 2003). ICA is a method of blind source separation in which a set of original data is decomposed into components non-Gaussian and statistically independent (DAMMERS et al., 2010). A key concept that constitutes the foundation of independent component analysis is statistical independence. Two random variables are considered statistically independent if the value of one gives no information on the values of the other. The measure of statistical independence takes into account higher-order statistics and, therefore, it is a measure stronger than correlation (HYVÄRINEN; KARHUNEN; OJA, 2001). In this sense, the transformation performed by ICA over the dynamic speckle data may amplify the differences between the data on the view of the Fujii and GD methods providing improvements on the results.

In this context, the present study proposed the use of independent component analysis as preprocessing to graphical methods of dynamic speckle analysis, in special for Fujii and generalized difference techniques, aiming at improving the visual quality of the final images, assisting the analysis and interpretation of results.

The next section presents the theory of the methods used in this work, starting with the independent component analysis description followed by the characterization of the Fujii and GD methods.

2 Theory

2.1 Independent component analysis

The independent components analysis - ICA - is a technique of blind source separation which aims to recover a set of sources signals of multidimensional statistical data by searching components non-Gaussian and

statistically independent (AHMAD; GHANBARI, 2011; KOLDOVSKÝ; TICHAVSKÝ; OJA, 2006; TICHAVSKÝ; KOLDOVSKÝ; OJA, 2006).

Currently, the ICA has become a major tool in the analysis of digital signals due their applicability in different knowledge areas, such as telecommunications systems, in analysis of biomedical signals and stock market prices (STONE, 2004).

The start point of the ICA analysis is to consider a vector **s** with *n* independent sources unknown, such that $\mathbf{s} = [s_1 \quad s_2 \quad \dots \quad s_n]^T$. The vector **s** was mixed using a matrix **A** and resulted in a vector **x** of observed signals (OJA; YUAN, 2006), as described mathematically in the Equation 1.

$$\mathbf{x} = \mathbf{A}.\mathbf{s} \tag{1}$$

where **x** is the matrix with the observations mixed, **A** consists of the matrix of mixtures and **s** is the vector with the original independent sources.

In Equation 1 only the vector **x** with the observations mixed is known and the ICA technique aims to find a matrix **W** that performs the inverse process of **A**, with the separation of the independent sources present in **x**.

The identification and separation of the independent statistically components is possible by maximizing of the non-gaussianity and the data analysis begins with the centralization of the vector **x** around the average (Equation 2).

$$\mathbf{x}'_n = \mathbf{x}_n - \mu(\mathbf{x}_n) \tag{2}$$

in which $\mathbf{x}'_n$ is the vector with the observations mixed and average zero and $\mu(x_n)$ is the average of each column of **x**, which can be calculated using Equation 3.

$$\mu(\mathbf{x_n}) = \frac{1}{\mathbf{m}} \sum_{\mathbf{j=1}}^{\mathbf{m}} \mathbf{x_n(j)} \tag{3}$$

After of the centralized data around the average is recommended the data whitening (decorrelation), an operation that simplifies the ICA problem and assists the algorithms of estimating of the independent components to converge faster (AHAMAD; GHANBARI, 2011; BELL; SEJNOWSKI, 1995; KARHUNEN, 1996).

The whitening operation transforms linearly **x'** in a new vector (**z**) with components uncorrelated, unit variance (KUMAR et al., 2013) and orthogonal. Ahamad and Ghanbari (2011) emphasize that the principal component analysis is one of the ways to realize the data whitening, which is described mathematically in the Equation 4.

$$\mathbf{z = E.D^{\frac{1}{2}}.E^{T}.x'} \tag{4}$$

where **z** are the data whited, **E** = (e_1, e_2, ..., e_n) is the eigenvectors matrix and $\mathbf{D}^{1/2}$ = diag($\lambda_1^{1/2}$, $\lambda_2^{1/2}$, ..., $\lambda_n^{1/2}$) consists of the diagonal matrix of eigenvalues of the covariance matrix $\mathbf{C_x}$ = E [x'.x'^{T}].

The covariance matrix of the data whited $\mathbf{C_z}$ is equal to identity matrix **I** (Equation 5), and the ICA problem now consists in estimating the matrix of separation **W** which allows to recover the independent sources (AHAMAD; GHANBARI, 2011; OJA; YUAN, 2006) , according to Equation 6.

$$\mathbf{E[z.z^{T}] = I} \tag{5}$$

$$\mathbf{y_n} = \mathbf{W}.\mathbf{z} \quad (6)$$

in which $\mathbf{y_n}$ are the independent components estimated and **W** corresponds the matrix of separation ($\mathbf{W} = \mathbf{A}^{-1}$).

The vector $\mathbf{y_n}$ consists of the recovered signal and the closer possible of the original sources **s** if the sources are statistically independent and non-Gaussian (HYVÄRINEN; OJA, 2000). It is worth noting also that the ICA model estimated the order of components is random.

There are available in the literature various computational codes for the estimation of independent components, and calculation basis on these algorithms is the maximization of non-gaussianity, which can be realized by minimizing of the mutual information, maximization of the absolute values of the kurtosis or negentropy, among other.

2.2 FastICA algorithm

FastICA or fixed-point algorithm was proposed by Hyvärinen (1999) and Hyvärinen and Oja (1997) to solve the ICA problem and the blind source separation. It is one of the most successful algorithms to estimate the separation matrix **W** and is distinguished by accuracy, robustness, low computational complexity and fast convergence (KOLDOVSKÝ; TICHAVSKÝ; OJA, 2006; TICHAVSKÝ; KOLDOVSKÝ; OJA, 2006).

There are two types of FastICA algorithm available in the literature: the deflation and the symmetric. The first estimates the independent components successively under orthogonality conditions and in the symmetric algorithm the independent components are calculated in parallel (DELFOSSE; LOUBATON, 1995), and both varieties are based on the optimization of the contrast function.

Among the several variables used to measure the non-gaussianity of the data, the kurtosis gains prominence, being widely used as contrast function in

algorithms of estimation of the independent components, such as FastICA (DELFOSSE; LOUBATON, 1995; PAPADIAS, 2000). Equation 7 presents the mathematical expression of the kurtosis.

$$\mathbf{kurtosis}(X) = E\{X^4\} - 3[E\{X^2\}]^2 \quad (7)$$

Random variables with Gauss distribution have kurtosis null and high kurtosis values are linked to the signals with distribution non-Gaussian. In addition, the fourth-order statistical moment at which the kurtosis is based makes it very sensitive to variations in the data.

2.3 Fujii method

The visual observation does not allow to quantify the patterns of optical interference of the dynamic laser speckle (RABAL; BRAGA, 2008), and an alternative to the analysis and interpretation is to use graphical methods.

In this context, the technique proposed by Fujii et al. (1987) consists of a graphic method of analysis of the dynamic laser speckle which is based on work images from a body illuminated, identifying the luminous intensity of each pixel that is composed.

The technique consists of the summation of the differences in light intensities between an image and its follower, divided by the sum of intensities between an image and its subsequent (weighting factor), as described mathematically in the Equation 8.

$$\mathbf{Fujii}(x,y) = \sum_{k=1}^{N} \frac{I_k(x,y) - I_{k+1}(x,y)}{I_k(x,y) + I_{k+1}(x,y)} \quad (8)$$

where I_k (x,y) is the pixel intensity of coordinated (x,y) of the k^{th} image.

The result is a new image in which regions of high activity are represented in light gray levels and dark areas illustrate regions of low activity.

The amplification of the movement in darker areas is a characteristic of the Fujii method, resulting in clearer images when compared with the generalized differences method (BRAGA et al., 2009).

2.4 Generalized differences method (GD)

Presented by Arizaga, Trivi e Rabal (1999) as an alternative to the Fujii method, the generalized differences (GD) does not have the weighting factor in the denominator of the mathematics expression (Equation 9), and the differences between the intensities of the pixels have been generalized to the whole the captured images.

$$GD(x,y) = \sum_{k}\sum_{l} I_k(x,y) - I_{k+1}(x,y) \tag{9}$$

in which GD(x,y) is the final image of the application of the graphical technique in a set of images.

3 Materials and methods

Actual data of a maize fruit illuminated with a laser light, obtained from Braga et al. (2001) were used for the realization of the proposed approaches.

The database is composed by 64 images in gray levels with resolution of 490 x 256 pixels, which have been captured using the experimental configuration known as back-scattering and acquisition rate between images of 0.08 seconds.

Each image of the data set was concatenated and the signal formed

ordered vertically side by side in a new matrix (**X**), as illustrated in Figure 1.

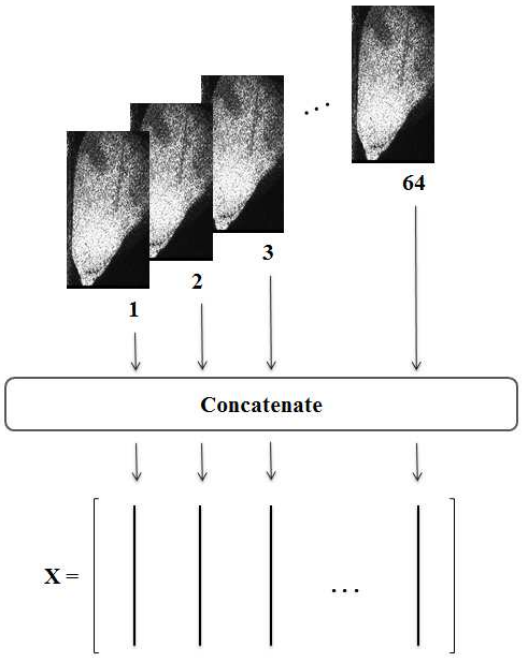


Figure 1 Order of the images concatenated

The matrix **X** (Figure 1) was considered as the matrix with the mixed observations and the data were transformed to the domain of the independent components using the deflation FastICA algorithm, developed by Hyvärinen (1999) and Hyvärinen and Oja (1997).

It is noteworthy that during ICA processing occurred without reduction of the data volume and the choice of the computational routine FastICA was due to its accuracy, robustness, low computational complexity and fast convergence, as highlighted by Koldovský, Tichavský and Oja (2006) and Tichavský, Koldovský and Oja (2006).

The inverse process of the concatenation was performed on the independent components, in which the data is regrouped to form the reconstructed images. After this, the reconstructed images are analyzed by using

the Fujii and GD techniques aiming at presenting the spatial variability of the biological activity of the maize fruit. Figure 2 summarizes the methodology used.

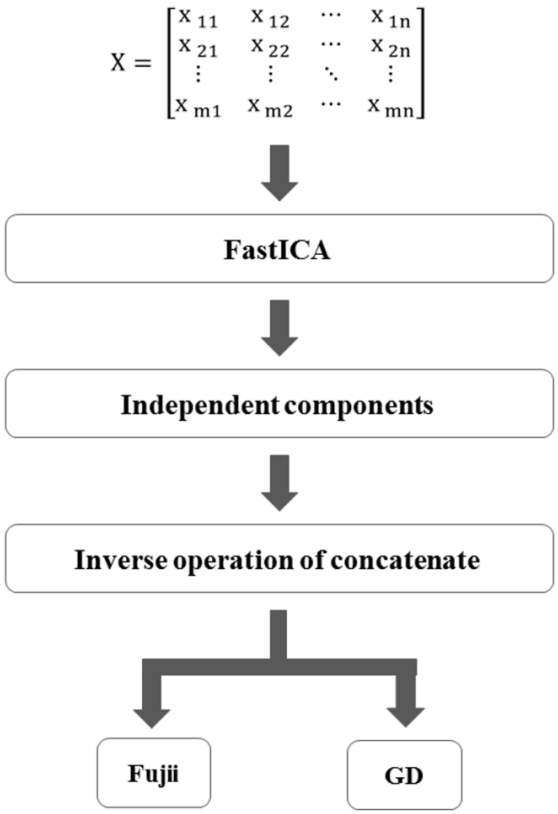


Figure 2 Block diagram of the proposed methodology

Ten repetitions of the proposed methodology were carried out, and the histograms of the reconstructed images by Fujii and GD were used to evaluate the effect of the randomness in the independent components order on the graphical methods of analysis of the dynamic laser speckle.

Finally, the average of five rows in the final images (Fujii, GD) with and without ICA preprocessing were selected in the position illustrated in the Figure

3 and used for comparatively performance evaluation. The variation coefficients, ratio between standard deviation and the data average, in the embryo and endosperm tissues were calculated and the lines concerning the same graphical method were shown in the same figure to observe the behavior of the biological activity in the tissues of the maize fruit in terms of amplitude.

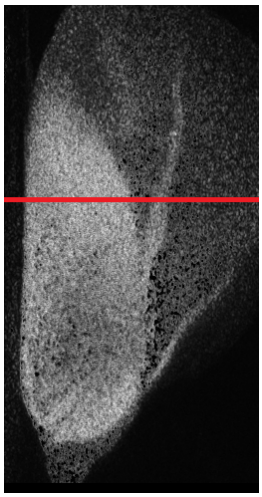
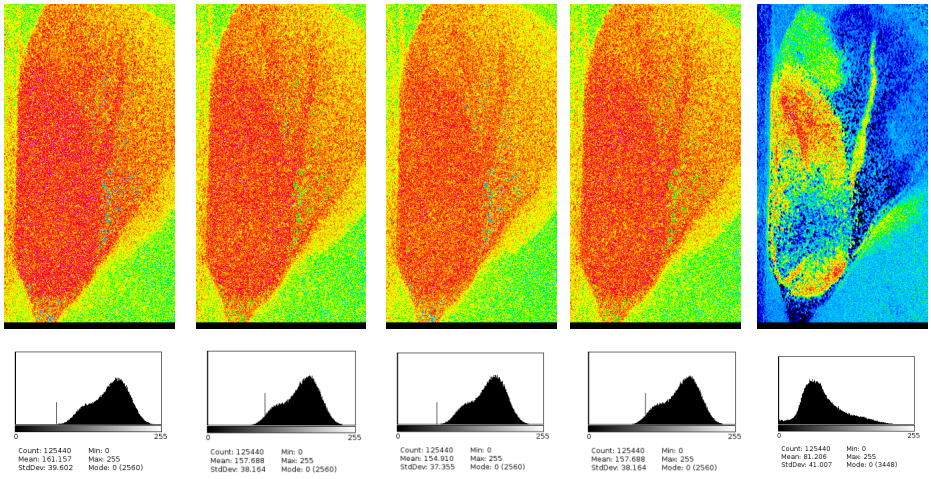


Figure 3 Line selected in the images

4 Results and discussions

4.1 Fujii method

Figure 4 shows the final images Fujii preprocessed with the independent component analysis and their respective histograms.



Fujii 1 Fujii 2 Fujii 3 Fujii 4 Original

Figure 4 Final Fujii images preprocessed with the ICA technique

The images titled Fujii 1, Fujii 2, Fujii 3 and Fujii 4 are four of the ten results obtained with the proposed methodology, while the image named Original corresponds to the Fujii image without preprocessing with ICA, which was considered as reference for the analysis.

The histograms illustrated in the Figure 4 show that the usage of ICA as data preprocessing technique conducted the gray levels from dark area to the clear region of the interval of gray levels. This behavior was observed in the ten repetitions performed and resulted in final images with clearer gray levels.

A second view of the images lightening is illustrated in the Figure 5, where the average of the five lines of the Fujii 1 and Original images (Figure 4) was presented in the same window.

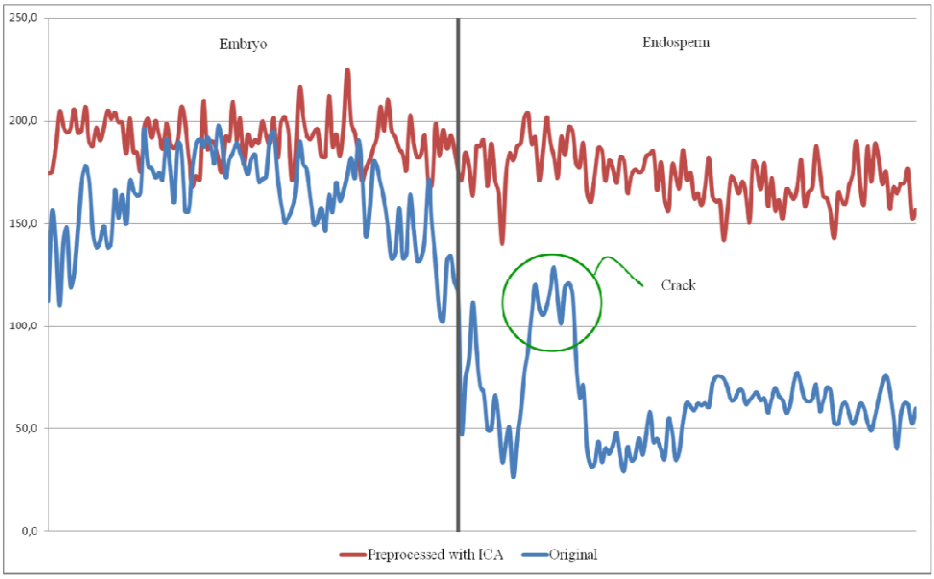


Figure 5 Line selected in the Fujii images preprocessed with ICA and in the reference Fujii image

The preprocessing of the data using ICA increased the signal amplitude to gray levels close to white (gray level 256) compared with the reference signal, in special to the endosperm tissue, which confirms the whitening of the final images.

The intense whitening is called saturation of the image, and the images shown in the Figure 4 is the result of the maximization of the differences produced by the ICA technique combined with the characteristic of the Fujii method, which basically amplifies movements in darker areas as reported by Braga et al. (2009).

The proposed methodology produced Fujii images with gray levels of the histogram concentrated in a narrow region, which characterizes low contrast images as specified by Gonzalez and Woods (2000) and explains the complex differentiation of the biological tissues of the maize fruit, in which the biological activities in embryo and endosperm tissues presented close.

However, despite the low visual quality of the final images, a reduction in the variation coefficients of the biological activity is verified when the data are preprocessed with independent component analysis. Table 1 shows the variation coefficients for biological tissues of the maize fruit estimated on the lines selected of the Fujii's images presented in the Figure 5.

Table 1 Coefficients of variation for the embryo and endosperm tissues with and without preprocessing using ICA

	Original	Preprocessed with ICA
Embryo	0.132	0.056
Endosperm	0.353	0.072

The variation coefficient was lower in the embryonic tissue than in the endosperm tissue in both statistical treatments. Note that the selected row in the reference Fujii image displays information of a crack located in the endosperm (highlighted in the Figure 5), which has high and atypical biological activity, which contributes significantly to the high value of the variation coefficient (0.353) of the reference endosperm.

The preprocessing using ICA reduced the variation coefficient in both biological tissues, and the explanation are in the small oscillations of the resulting signal amplitude, which is translated into low value of standard deviation and, consequently, reduced coefficient of variation.

The repetitions showed that the Fujii method is vulnerable to randomness in the order of the independent components, producing different results for each test performed, as summarized in Table 2.

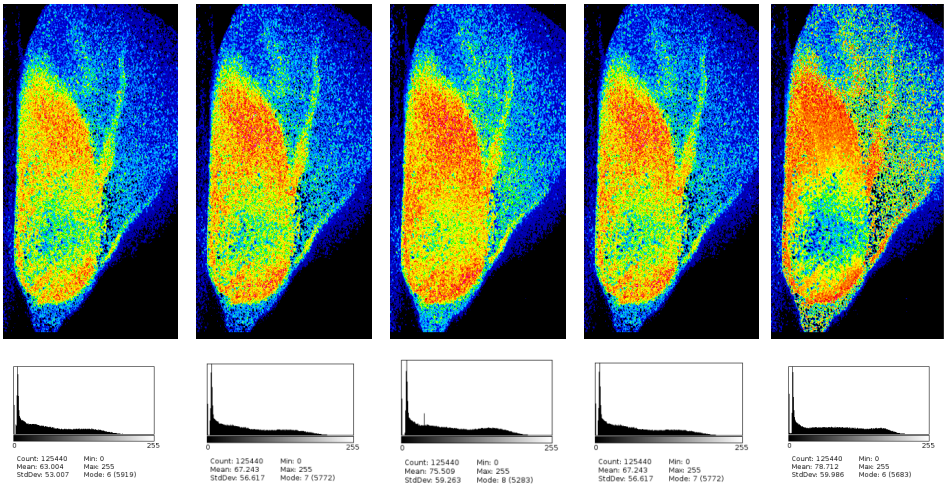
Table 2 Mean values of the histograms gray levels of the ten repetitions

	Mean values
Repetition 1:	161.2
Repetition 2:	154.9
Repetition 3:	157.7
Repetition 4:	159.5
Repetition 5:	154.9
Repetition 6:	154.9
Repetition 7:	157.7
Repetition 8:	159.5
Repetition 9:	163.2
Repetition 10:	161.0

Although the mean values of the gray levels for the repetitions were not constant, the differences in the values were small, with the data standard deviation presented in Table 2 of 2.76.

4.2 Generalized diferences method

The usage of ICA as preprocessing technique for the GD method leads to final images with better quality compared to the results provided by Fujii method, and the Figure 6 presents these graphical results.



GD 1 GD 2 GD 3 GD 4 Original

Figure 6 GD images preprocessed with ICA and their respective histograms

Addition to the best visual quality, the preprocessing of the biospeckle data by the ICA provided higher homogeneity of the biological activity in the maize fruit tissues as well as a better view of the crack present in the endosperm, as can be seen in the Figure 6.

The histograms of the final GD images (Figure 6) showed significant scattering of the gray levels, which is associated with digital images with high contrast as discussed by Gonzalez and Woods (2000), and this feature facilitates the analysis and visual interpretation of the data.

The expressive distribution of the gray levels of the histogram brings an increase in the standard deviation values, thus, increasing the level of the variation coefficients (Table 3).

Table 3 Coefficients of variation of the maize fruit tissues with and without ICA preprocessing

	Original	Preprocessed with ICA
Embryo:	0.112	0.148
Endosperm:	0.311	0.378

Similar to the results presented in Fujii method, endosperm had higher variation coefficient than the embryonic tissue in both treatments, noting that the high biological activity in the crack located in the endosperm tissue significantly influence these numbers.

In the opposite direction of the Fujii results there was an increase in the variation coefficients when the biospeckle data were analyzed with ICA and GD. However, these results show that only the variation coefficients are not enough to qualify the data since the numbers shown themselves inversely proportional to the visual quality of the final images.

Figure 7 graphically displays the selected rows in the reference GD image and in the GD 1 image previously processed with ICA, which allows another view of the behavior of biological activity in the embryo and endosperm.

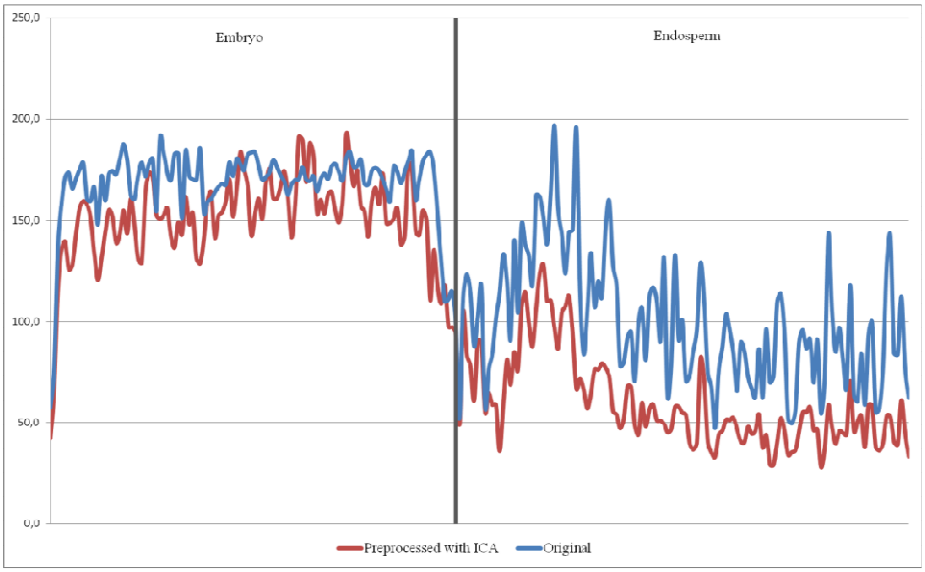


Figure 7 Behavior of the biological activity in the maize fruit tissues with and without preprocessing via ICA

The signal resulting from proposed methodology showed behavior close to the reference signal, with the biological activities in the embryo region and of the endosperm at different levels, and these results facilitate the application of digital images segmentation techniques.

It is noteworthy the abrupt transition of the biological activity between embryo and endosperm in the data analyzed by ICA, which come against the common sense in which there should be abrupt transition of the activity to be two distinct tissues.

In terms of sensitivity the order in which the independent components are shown, the generalized differences technique, as well as the Fujii method proved to be vulnerable to the randomness of the IC's, as the data in Table 4.

Table 4 Mean values of the gray levels of the histograms

	Mean values
Repetition 1	63.0
Repetition 2	75.5
Repetition 3	67.2
Repetition 4	66.5
Repetition 5	75.5
Repetition 6	75.5
Repetition 7	67.2
Repetition 8	66.5
Repetition 9	74.2
Repetition 10	78.4

The standard deviation of the mean values of the gray levels (Table 4) was of 5.66, higher than the value found in the Fujii method (2.76). The values indicate higher susceptibility to randomness in the order of the independent components by graphical method of GD, which was not expected since all the images are compared with each other in this processing unlike of the Fujii analysis in which the image is compared with its subsequent.

5 Conclusion

The preprocessing of the dynamic speckle data through independent component analysis presented improvement in the visual quality of the final images when associated with graphical method of generalized differences.

Both graphical methods of analysis of the biospeckle tested showed vulnerable to randomness in the order of the independent components from the FastICA algorithm, which must be overcome for authentication of the proposed methodology.

Acknowledgements

This work was partially financed by CNPq, Fapemig, Capes, Finep in Brazil, and partly supported by the Scottish Government Rural and Environment Science and Analytical Services division.

REFERENCES

AHMAD, T.; GHANBARI, M. A review of independent component analysis (ICA) based on kurtosis contrast function. **Australian Journal of Basic and Applied Sciences**, Amman, v. 5, n. 9, p. 1747-1755, 2011.

AMARAL, I. C. et al. Application of biospeckle laser technique for determining biological phenomena related to beef aging. **Journal of Food Engineering**, Essex, v. 119, n. 1, p. 135-139, 2013.

ARIZAGA, R.; TRIVI, M.; RABAL, H. Speckle time evolution characterization by the co-occurrence matrix analysis. **Optics and Laser Technology**, New York, v. 31, n. 2, p. 163-169, 1999.

BEDOYA, G.; BERMEJO, S.; CABESTANY, J. Comparison of neural algorithms for blind source separation in sensor array applications. In: EUROPEAN SYMPOSIUM ON ARTIFICIAL NEURAL NETWORKS, 11., 2003, Bruges. **Proceedings…** Bruges: Université Catholique de Louvain, 2003. p. 131-136.

BELL, A. J.; SEJNOWSKI, T. J. An information-maximization approach to blind separation and blind deconvolution. **Neural Computation**, Cambridge, v. 7, n. 6, p. 1129-1159, 1995.

BRAGA, R. A. et al. Detection of fungi in beans by the laser biospeckle technique. **Biosystems Engineering**, London, v. 91, n. 4, p. 465-469, 2005.

BRAGA, R. A. et al. Live biospeckle laser imaging of root tissue. **European Biophysics Journal**, London, v. 38, n. 5, p. 679-686, 2009.

BRAGA, R. A. et al. Potencial do bio-speckle laser para avaliação da viabilidade de sementes. **Ciência e Agrotecnologia**, Lavras, v. 25, n. 3, p. 645-649, maio/jun. 2001.

COSTA, R. M. et al. Técnicas estatísticas aplicadas em imagens do speckle dinâmico. **Revista Brasileira de Biometria**, São Paulo, v. 28, n. 2, p. 27-39, 2010.

DAMMERS, J. et al. Signal enhancement in polarized light imaging by means of independent component analysis. **NeuroImage**, Orlando, v. 49, n. 2, p. 1241-1248, 2010.

DELFOSSE, N.; LOUBATON, P. Adaptive blind separation of independent sources: a deflation approach. **Signal Processing**, Amsterdam, v. 45, n. 1, p. 59-83, 1995.

FUJII, H. et al. Evaluation of blood flow by laser speckle image sensing. **Applied Optics**, New York, v. 26, n. 24, p. 5321-5325, 1987.

GONZALEZ, R. C.; WOODS, R. E. **Processamento de imagens digitais**. São Paulo: Blucher, 2000. 509 p.

HYVÄRINEN, A. Fast and robust fixed point algorithm for independent component analysis. **IEEE Transations on Neural Networks**, New York, v. 10, n. 3, p. 626-634, 1999.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. **Independent component analysis**. New York: J. Wiley, 2001. 504 p.

HYVÄRINEN, A.; OJA, E. A fast fixed-point algorithm for independent component analysis. **Neural Computation**, Cambridge, v. 9, n. 7, p. 1483-1492, 1997.

HYVÄRINEN, A.; OJA, E. Independent component analysis: algorithm and applications. **Neural Networks**, New York, v. 13, n. 4/5, p. 411-430, 2000.

KARHUNEN, J. Neural approaches to independent component analysis and source separation. In: EUROPEAN SYMP. ON ARTIFICIAL NEURAL NETWORKS, 4., 1996, Bruges. **Proceedings…** Bruges: ESANN, 1996. p. 249-266.

KOLDOVSKÝ, Z.; TICHAVSKÝ, P.; OJA, E. Efficient variant of algorithm FastICA for independent component analysis attaining the Cramér-Rao lower bound. **IEEE Transactions on Neural Networks**, New York, v. 17, n. 5, p. 1265-1277, 2006.

KUMAR, S. et al. Suitability of independent component analysis in digital image forgery detection. **International Journal of Engineering and Technology**, Bremen, v. 5, n. 1, p. 226-231, 2013.

LI, X.; TAI, Y.; NIE, Z. Application of dynamic speckle method using in sedimentation process of silver chloride. **Optik - International Journal for Light and Electron Optics**, Amsterdam, v. 122, n. 23, p. 2155-2157, 2011.

MURIALDO, S. E. et al. Discrimination of motile bacteria from filamentous fungi using dynamic speckle. **Journal of Biomedical Optics**, New York, v. 17, n. 5, p. 56011-1-56011-5, 2012.

OJA, E.; YUAN, Z. The FastICA algorithm revisited: convergence analysis. **IEEE Transactions on Neural Networks**, Chicago, v. 17, n. 6, p. 1370-1381, 2006.

PAPADIAS, C. B. Globally convergent blind source separation based on a multiuser kurtosis maximization criterion. **IEEE Transactions on Signal Processing**, Amsterdam, v. 48, n. 12, p. 3508-3519, 2000.

PASSONI, I. et al. Dynamic speckle processing using wavelets based entropy. **Optics Communications**, Amsterdam, v. 246, n. 1/3, p. 219-228, 2005.

RABAL, H. J.; BRAGA, R. A. **Dynamic laser speckle and applications**. Boca Raton: CRC, 2008. 304 p.

RABAL, H. J. et al. Q-statistics in dynamic speckle pattern analysis. **Optics and Lasers in Engineering**, London, v. 50, n. 6, p. 855-861, 2012.

STONE, J. V. **Independent component analysis:** a tutorial introduction. New York: J. Wiley, 2004. 193 p.

TICHAVSKÝ, P.; KOLDOVSKÝ, Z.; OJA, E. Performance analysis of the FastICA algorithm and Cramér-Rao bounds for linear independent component analysis. **IEEE Transactions on Signal Processing**, Amsterdam, v. 54, n. 4, p. 1189-1203, 2006.

CONSIDERAÇÕES FINAIS

A demanda por técnicas de análise de sinais que possam auxiliar nas análises e interpretações dos dados do *biospeckle* laser tem aumentado com a utilização cada vez maior da técnica do *speckle* laser dinâmico em estudos de movimentos moleculares e estruturais de superfícies. As técnicas de Fourier, wavelet, PCA e ICA são algumas interessantes ferramentas.

Observando-se os resultados encontrados, pode-se inferir que a transformada de wavelet mostrou ser mais apropriada que a transformada de Fourier na análise no domínio da frequência dos dados do *biospeckle* laser, com informações adicionais sobre as partes constituintes da base de dados e permitindo a definição de bandas de frequências associadas a fenômenos biológicos. A característica não estacionária do sinal do *speckle* laser dinâmico limita a aplicação da técnica de Fourier e restringe sua aplicação a trabalhos que buscam conhecer apenas o conteúdo espectral dos dados.

Se, por um lado, a transformada de wavelet mostra-se mais adequada aos estudos espectrais do *biospeckle* laser, com bons resultados nos trabalhos conduzidos, a sua utilização despende elevado tempo de processamento e também a escolha de uma função base para a análise. E os critérios para essa seleção ainda não são claros.

Neste contexto, a filtragem dos dados do *speckle* laser dinâmico por meio da análise dos componentes principais é uma alternativa que permite implementar filtros de diferentes bandas passantes de frequência e também definir de marcadores de componentes principais relacionado a fenômenos biológicos. Além disso, na filtragem PCA, as funções bases utilizadas na decomposição dos dados são estatisticamente ótimas e seu rápido processamento computacional abre espaço para a sua aplicação em sistemas online de monitoramento e análise da atividade biológica de materiais.

O pré-processamento dos dados do *biospeckle* laser meio da análise de componentes independentes, associado ao método gráfico de diferenças generalizadas, melhorou a qualidade das imagens finais, não acontecendo o mesmo para o método de Fujii. Ambos os métodos gráficos utilizados mostraram-se sensíveis às saídas aleatórias das componentes independentes, obstáculo que deve ser superado para a validação da metodologia proposta.

Como propostas para trabalhos futuros destacam-se:

a) avaliar diferentes funções base de wavelets na análise espectral do *speckle* laser dinâmico, de forma a identificar qual a onda mãe mais apropriada;
b) propor a filtragem estatística do *biospeckle* laser por meio da análise dos componentes independentes e comparar os resultados com os obtidos na filtragem PCA;
c) aperfeiçoar o algoritmo FastICA, de forma a solucionar a ordem aleatória das componentes independentes estimadas, barreira para a confirmação da metodologia de pré-processamento das saídas gráficas com ICA.