Optical Diagnosis of Cervical Cancer by Intrinsic Mode Functions

Sabyasachi Mukhopadhyay[1*], Sawon Pratiher[2], Souvik Pratiher[3], Asima Pradhan[2], Nirmalya Ghosh[1], Prasanta K. Panigrahi[1]

[1]Indian Institute of Science Education and Research Kolkata, Nadia-741246, West Bengal, India
[2]Indian Institute Technology Kanpur, Kanpur- 208016, Uttar Pradesh, India
[3]KIIT University, Bhubaneswar-751024, Odisha, India

*mukhopadhyaysabyasachi@gmail.com

ABSTRACT

In this paper, we make use of the empirical mode decomposition (EMD) to discriminate the cervical cancer tissues from normal ones based on elastic scattering spectroscopy. The phase space has been reconstructed through decomposing the optical signal into a finite set of bandlimited signals known as intrinsic mode functions (IMFs). It has been shown that the area measure of the analytic IMFs provides a good discrimination performance. Simulation results validate the efficacy of the IMFs followed by SVM based classification.

Keywords: intrinsic mode functions, empirical mode decomposition, Analytic representation, Hilbert transform, cervical cancer.

1. INTRODUCTION

In order to diagnose properly, detection of cancer at earlier stages has become major concern for researchers. During the progress of precancer, tissues undergo subtle morphological changes [1]. The potential biomarkers like wavelets [2-3], multifractal detrended fluctuation analysis (MFDFA) [2-3], s-transform [8] played a significant role in feature extraction in diagnosis of cancer. The machine learning methodologies like principal component analysis (PCA) [4], artificial neural network (ANN) [6], support vector machine (SVM) [7,10] demonstrated their efficacy in production of optimum classification accuracies. The signal processing technique like empirical mode decomposition (EMD) proved to be very useful for discrimination of seizure and seizure-free EEG signals [11]. In this manuscript, we demonstrate the efficacy of IMFs as a potential biomarker followed by SVM based classification in optical diagnosis of cancer.

2. THEORY

Intrinsic Mode Functions (IMFs): Intrinsic mode functions (IMFs) denotes simple oscillatory mode as a counterpart to the simple harmonic function. An IMF function should have the same number of extrema and zero crossings, whose mean value of envelopes defined by local maxima and local minima are zero [12].

Empirical Mode Decomposition (EMD): The core part of the Hilbert-Huang transform (HHT) is the empirical mode decomposition (EMD) method. By breaking signals into various components, EMD can be compared with other analysis methods like wavelets. The EMD method can be deployed to decompose any intricate data set into a finite and small number of components which form an orthogonal basis for the original signal. Moreover, the decomposition does not require any conditions about the stationarity and linearity of the signal. For more detail, please see reference [12-13].

Support Vector Machine (SVM): The support vector machine (SVM) is a supervised machine learning methodology which can be interpreted as an extended version of perceptron. SVM takes benefit of the structural risk minimization (SRM) scheme of statistical machine learning by forming an optimum separating hyperplane (OSH) in order to maximize the margin among different classes. The procedure of maximization of geometric margin reduces the empirical classification errors which lead SVMs to perform as maximum margin classifiers. The prediction over a function $f(x)$ by SRM scheme can be validated as: $f(x) = \sum_{i=1}^{N} w_i k(x, x_i) + w_0$,where, $k(x, x_i)$ is the kernel function and $\{w_i\}$ is the corresponding model weights.The distant training data points from OSH plane receive weight zero and do not participate under the classification scheme, whereas, the nearer training data points to decision boundary receive non-zero weights and these category of training data points are known as 'Support vectors' [14] .

Dynamics and Fluctuations in Biomedical Photonics XIV, edited by Valery V. Tuchin, Kirill V. Larin, Martin J. Leahy, Ruikang K. Wang, Proc. of SPIE Vol. 10063, 100630V · © 2017 SPIE
CCC code: 1605-7422/17/$18 · doi: 10.1117/12.2251225

Selection of Kernel: The 'mercer's theorem' plays the key role behind the choice of kernel function which actually controls the success of SVM [15]. A linear classifier is a dot product of support vector x_i and the data set x in the input space and generates the linear OSH. Mathematically, $k(x_i, x) = x_i, x + 1$, where, the feature space and N-dimensional input space are same.

The higher order non-linear kernels like d th order kernel can be expressed mathematically as: $k(x_i, x) = (x_i, x + 1)^d$.

The Gaussian RBF kernel is expressed as: $k(x_i, x) = exp\left(-\|x_i - x\|^2 / 2\sigma^2\right)$, where σ is the width of Gaussian.

3. DATA ACQUISITION AND PROPOSED METHODOLOGY

There are 57 observations in the dataset of cervical tissue regions, which have been marked in the following order, such as, 1st to 15th observations have been marked as healthy, 16th to 31st observations have been marked as grade I, 32nd to 40th observations have been marked as grade II and 41st to 57th observations have been marked as grade III.

3.1 Data Processing:

Fig. 1, highlights the workflow of the proposed methodology.

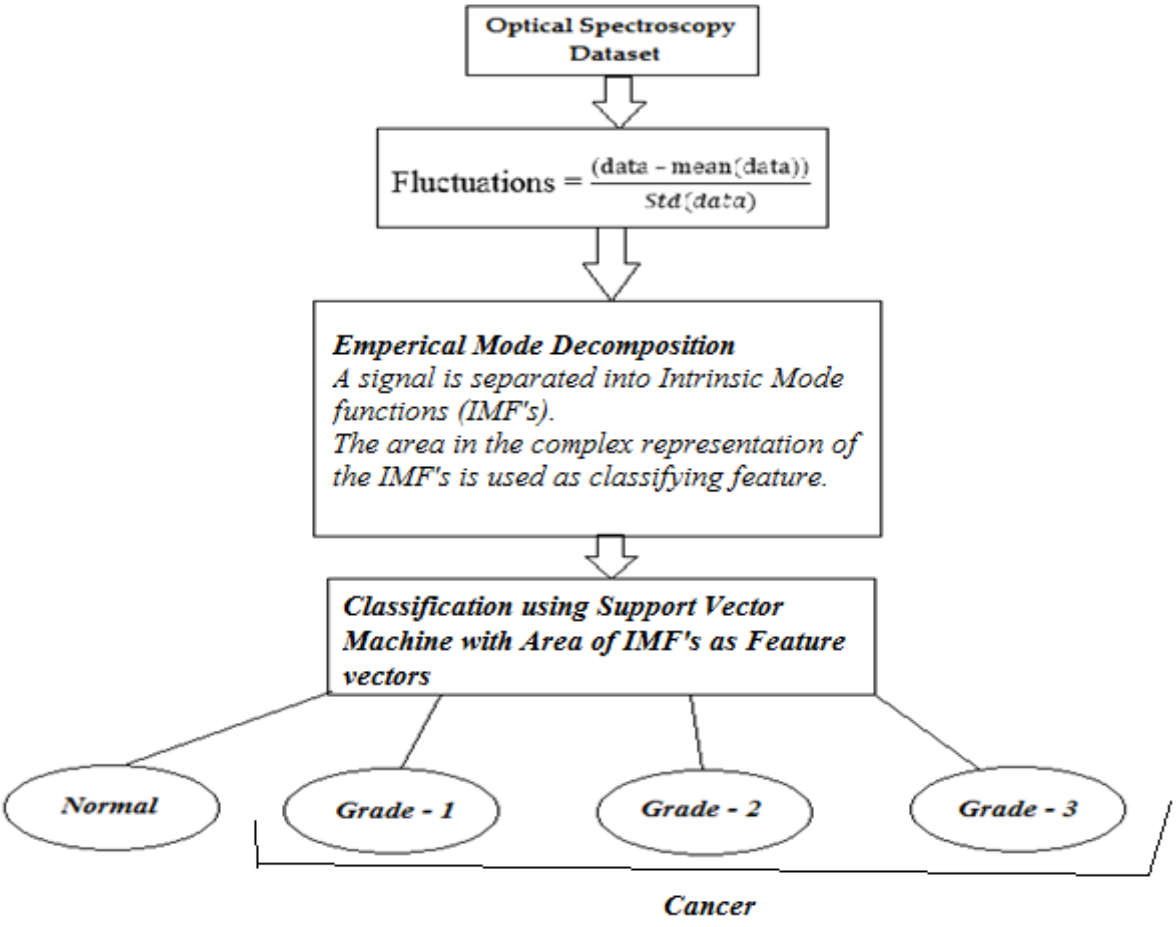

Fig.1 Workflow of the proposed methodology

4. RESULTS AND DISCUSSIONS

The efficacy of intrinsic mode functions by their analytic signal representation to discriminate spectroscopy signals containing normal and cancerous scattering data has been shown. The use of EMD to decompose spectral data into IMFs

provides estimation of the set of proper rotations for classifying the phase spectrum centers and surface areas in the complex plane.

A. Emperical Mode Decompostion (EMD) of Intrinsic Mode Functions (IMF's)

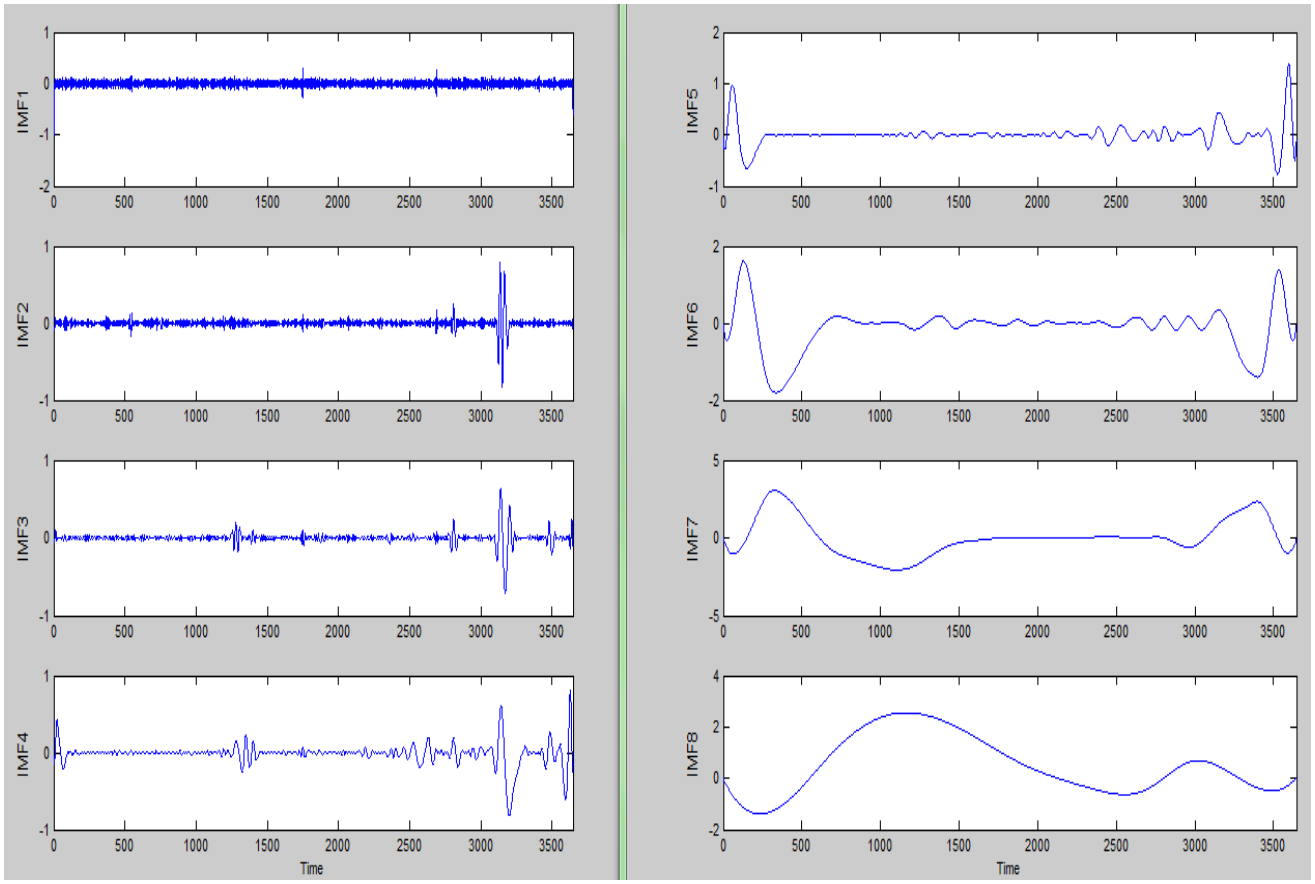

Fig.2 Empirical Mode Decompsotion (EMD) of a normal scattering signal (A representative demonstration of EMD for healthy cervical tissues)

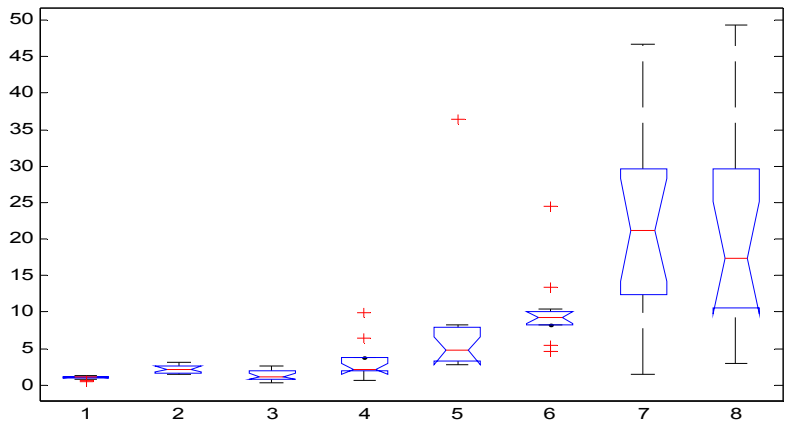

Fig.3 Area parameter of different intrinsic mode functions (IMFs) for scattering spectra of healthy cervical tissues

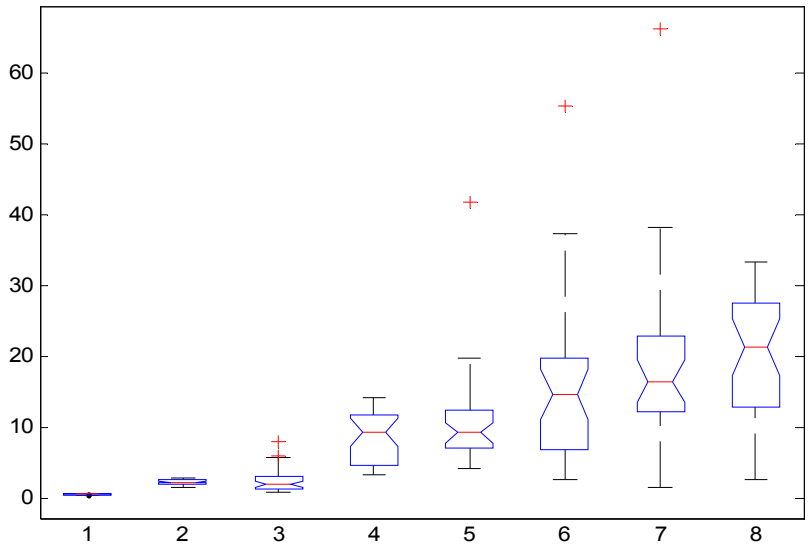

Fig.4 Area parameter of different intrinsic mode functions (IMFs) for scattering spectra of different cancerous cervical tissues (Grade III)

The area parameter of the analytic IMFs provides statistically significant difference between normal and cancerous spectra. The normal tissue spectra had significant greater surface area than compare to the cancerous tissues. The increased surface area in the complex plane for IMFs of the cancerous tissues is due to the large perturbations of the cancerous tissues.

B. Area Feature Variation of Analytic Representation of IMF's

Table I – Mean Area of analytic signal representation of intrinsic mode functions

Grade	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
Normal	0.99563	2.1632	1.3839	3.2050	7.4508	9.9325	22.222	21.707
Grade-I	0.67298	2.5627	1.6115	6.7552	11.7600	21.245	21.160	19.397
Grade-II	0.63081	2.2559	2.7633	9.3851	9.9975	11.739	17.955	17.943
Grade-III	0.64484	2.0555	4.1887	9.8397	9.6416	10.995	15.059	19.981

Table II – Standard Deviation of the Areas of analytic signal representation of intrinsic mode functions

Grade	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
Normal	0.211360	0.51268	0.76445	2.2948	8.2985	4.5202	13.9650	14.3740
Grade I	0.040888	0.28107	0.65467	3.6696	8.7716	13.1500	14.7540	10.1280
Grade II	0.033800	0.33357	1.52060	3.2325	4.2245	7.2485	9.1433	8.1713
Grade III	0.045213	0.23233	2.24340	3.6124	4.7113	7.3024	6.3471	7.1131

The phase space has been reconstructed through decomposing the optical signal into a intrinsic mode functions (IMFs) in Fig.5.

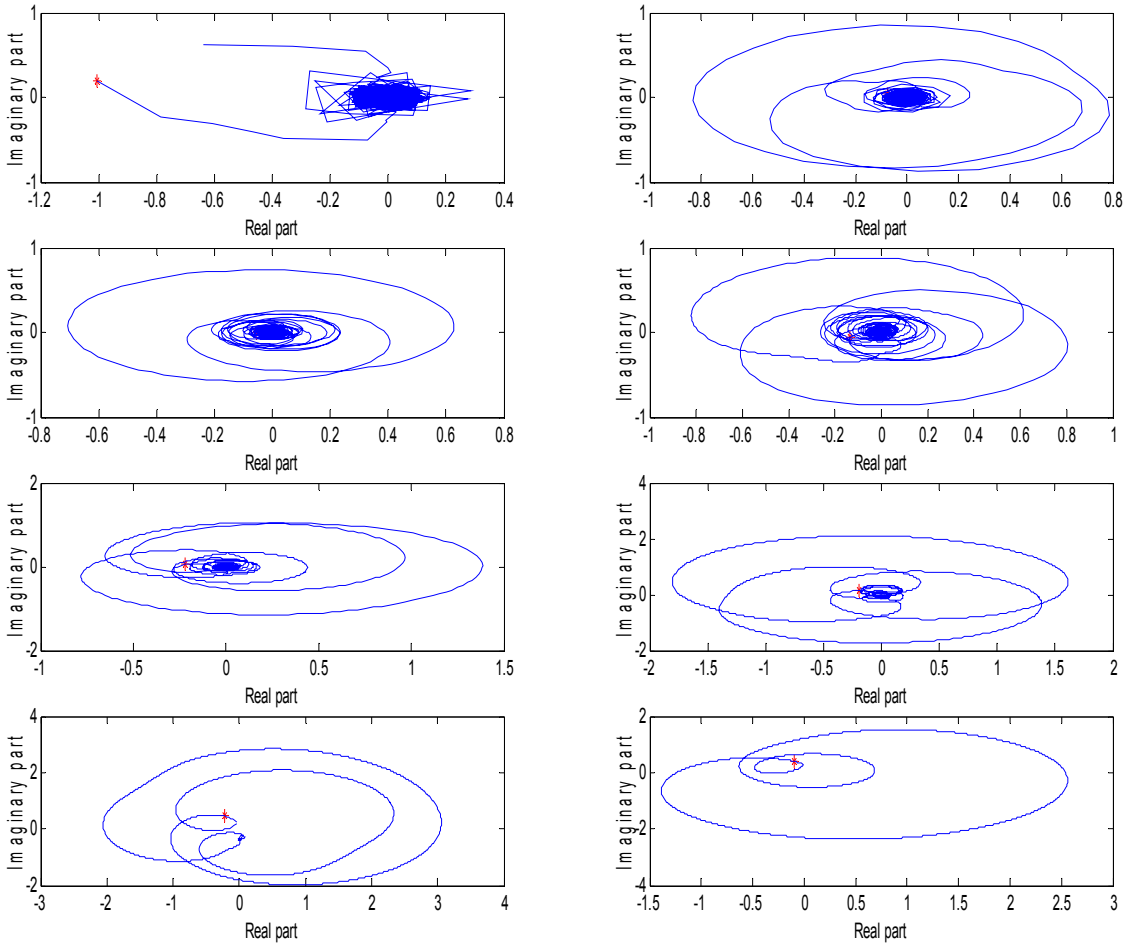

Fig.5 Analytic signal representations in the complex plane of Normal Scattering signal (a) IMF1, (b) IMF2, (c) IMF3, (d) IMF4, (e) IMF5 (f) IMF6, (g) IMF7, and (h) IMF8. (A representative demonstration of phase space reconstruction for healthy cervical tissues)

The simulated results in below Table II validate the efficacy of the IMFs as a potential biomarker followed by the SVM based classifications for different kernels. It is quite evident that SVM with linear kernel produces the optimum results.

Table II – SVM Classifier Performance for different kernels

Classifier	Accuracy	Sensitivity	Specificity	Precision	Recall
SVM with Linear Kernel	0.95652	1.00	0.875	0.9375	1.00
SVM with Quadratic Kernel	0.86957	0.93	0.750	0.8750	0.93
SVM with Polynomial Kernel	0.82609	0.80	0.875	0.92308	0.80
SVM with RBF Kernel	0.91304	1.00	0.750	0.88235	1.00

5. CONCLUSION

In this contribution, these primary results show considerable promise of IMF as a novel biomarker for optical diagnosis of cervical cancer. To establish the clinical use for this proposed technique for pre-cancer detection, it is necessary to test on *in vivo* datasets. Future lines of research include application of area measure to analyze the

segregation of different grades of cancer and to determine the feasibility of prediction of cancer using area measures of analytic IMFs.

REFERENCES

[1] Prasad,P.N., *Introduction to biophotonics,* Wiley-Interscience (2003).

[2] Mukhopadhyay,S., Das,N.K., Pradhan,A., Ghosh,N., Panigrahi,P.K.," Wavelet and multi-fractal based analysis on DIC images in epithelium region to detect and diagnose the cancer progress among different grades of tissues*", Proceeding of SPIE*, Belgium, SPIE Photonics Europe (2014).

[3] Mukhopadhyay,S., Das,N.K., Pradhan,A., Ghosh,N., Panigrahi,P.K., "Pre-cancer Detection by Wavelet Transform and Multi-fractality in various grades of DIC Stromal Images", *Proceeding of SPIE*, USA, SPIE BIOS (2014).

[4] Devi,S., Panigrahi,P.K., and Pradhan,A., "Detecting cervical cancer progression through extracted intrinsic fluorescence and principal component analysis", *J.Biomed.Opt.***19** (2014).

[5] Mukhopadhyay,S., Mandal,S., Das,N.K., Dey,S., Mitra,A., Ghosh,N., Panigrahi,P.K., "Diagnosing Heterogeneous Dynamics for CT Scan Images of Human Brain in Wavelet and MFDFA domain", *Proceeding of Springer*, IEM Optronix-2014, India (2014).

[6]Rovithakis,G.A.,Maniadakis,A.N.,Zervakis,M.,Fillipidis,G.,Zakarakis,G.,Katsamouris,A.N.,Papazoglou,T.G., ''Artificial neural networks for discriminating pathologic from normal peripheral vascular tissue", *IEEE Trans.BME*.**48** (2001).

[7] Majumder,S.K., Ghosh,N., and Gupta,P.K., ''Support vector machine for optical diagnosis of cancer", *J.Biomed.Opt.***10** (2005).

[8] Mukhopadhyay,S., Ghosh,N., Barman,R., Panigrahi,P.K., Pratiher,S., Pradhan,A., "S-TRANSFORM BASED FLUCTUATION ANALYSIS- A METHOD FOR PRE-CANCER DETECTION", *IEEE conference proceeding*, Microcom-2016, India (2016).

[9] Das,N.K., Mukhopadhyay,S., Ghosh,N., Chhablani,J., Richhariya,A., Rao,K.D., Sahooe, N. K., "Investigation of alterations in multifractality in optical coherence tomographic images of in vivo human retina", *Journal of Biomedical Optics* (2016).

[10] Mukhopadhyay,S., Kurmi,I., Dey,R., Das,N.K., Pradhan,S., Pradhan,A., Ghosh,N., Panigrahi,P.K.,Mohanty, S.," Optical diagnosis of colon and cervical cancer by support vector machine", *Proceeding of SPIE*, Belgium, SPIE Photonics Europe (2016).

[11] Pachori, R.B., "Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition","*Research Letters in Signal Processing* (2008).

[12] Huang, N. E.; Attoh-Okine, N. O., *The Hilbert-Huang Transform in Engineering*, CRC Taylor & Francis, ISBN 978-0849334221 (2005).

[13] Huang, N. E. et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. A, 454: 903–995 (1998).

[14] Vapnik, V.N., *Statistical Learning Theory*, Wiley, New York (1998).

[15] Burges, C. J. C., ''A tutorial on support vector machines for pattern recognition,'' *Data Min. Knowl. Discov.* **2**(2), 121–167 (1998).