Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 46 (2015) 372 – 380

Procedia Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

Performance comparison of Variational Mode Decomposition over Empirical Wavelet Transform for the classification of power quality disturbances using Support Vector Machine

Aneesh C*, Sachin Kumar, Hisham P M, K P Soman

Centre for Excellence in Computational Engineering and Networking
Amrita Vishwa Vidyapeetham,
Coimbatore Campus, Tamilnadu-641112.

Abstract

This work considers the classification of power quality disturbances based on VMD (Variational Mode Decomposition) and EWT (Empirical Wavelet Transform) using SVM (Support Vector Machine). Performance comparison of VMD over EWT is done for producing feature vectors that can extract salient and unique nature of these disturbances. In this paper, these two adaptive signal processing methods are used to produce three Intrinsic Mode Function (IMF) components of power quality signals. Feature vectors produced by finding sines and cosines of statistical parameter vector of three different IMF candidates are used for training SVM. Validation for six different classes of power qualities including normal sinusoidal signal, sag, swell, harmonics, sag with harmonics, swell with harmonics is performed using synthetic data in MATLAB. Classification results using SVM shows that VMD outperforms over EWT for feature extraction process and the classification accuracy is tabled.

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: Intrinsic mode function; Empirical wavelet transform; Empirical mode decomposition;Alternate direction method of multipliers; Variational mode decomposition; AM-FM signal

* Corresponding author. Tel.: 8754699468
*E-mail address:*a4aneeshc@gmail.com

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)
doi:10.1016/j.procs.2015.02.033

1. Introduction

Power quality signal detection and its classification is of great importance in industrial and commercial power system. The disturbances in power quality signal are threats for electronic equipments which are sensitive to these disturbances[1]. Commonly these types of disturbances like swell, sag, harmonics are referred to as are referred to as the pollutions in power grid. The main sources include SMPS in computers, rectification circuits, non-linear resistive loads, inverters and verities of switching circuits, rotating machineries etc. Various techniques for detection and classification of power quality disturbances have been proposed recently. Mainly it includes hidden Markov models[2], artificial neural networks[3], Fourier and Wavelet transform for the analysis of harmonic pollution[4] etc. Wavelet transform based processing of power quality disturbances is explained in[5]. There are rule based system for disturbance classification[6] in which decisions for classifications are performed using a set of rules designed by human experts. As the dimensionality increases, this method has the drawback of extending the rules else poor classification may be the result.

The remaining section of this paper is organized as follows: Section 2 describes about Empirical Wavelet Transform (EWT) and Empirical Mode Decomposition (EMD). Variational Mode Decomposition (VMD) is described in Section 3. Section 4 explains the Support Vector Machine (SVM) for classification. Section 5 describes the proposed methodology. Section 6 discusses the result and conclusion of the paper with the scope of future work are described in section 7.

2. Empirical Wavelet Transform

EWT was introduced in 2013 by Jerome Gilles[7]. Aim is to extract a series of (Amplitude Modulated- Frequency Modulated) AM-FM signals from the given signal. This method of decomposition outperforms over EMD introduced in 1998 by Huang *et al*[8]. EMD can represent the signal in terms of fast and slow oscillations and the splitting of signal is local and fully data driven. in many cases, EMD algorithm fails to extracts the harmonics accurately if the signal is affected with noises hence Ensemble EMD (EEMD) method is introduced[9]. Moreover it gives too many number of modes in case of ECG like signals. In many situations, this many numbers of modes are irrelevant. Even though, this method doesn't have any strong mathematical background, it is widely used in different kinds of applications like in biomedical signal processing[10], signal processing and communication[11]. The purpose of the decomposition is to extract the principal modes of signal given by the equation,

$$m(t) = \sum_{k=1}^{N} A_k(t) \cos \varphi_k(t) \tag{1}$$

Amplitude A_K and time derivative of phase change φ'_k are assumed to be positive and slowly varying component compared to φ_k . These modes have compact support in Fourier domain. So any IMF has a short frequency support or in other words frequency vary in a small range. Analytic signal can be obtained using Hilbert transform applied to IMF components (EMD+Hilbert transform) which helps to determine the instantaneous frequency and instantaneous amplitude of the signal. EWT performs adaptive decomposition of Fourier spectrum by appropriately choosing the boundaries so that the information is relevant within the considered boundary regions[7]. These support boundaries are set by identifying local maximas and local minimas in the normalized Fourier domain. Wavelet filters are provided for each compactly supported Fourier regions within the boundary regions to extract IMF components. Wavelet filters have perfect reconstruction property which means that the original input signal can be exactly reconstructed without any losses by summing the filter output signals. EWT algorithm uses Mayer's filter bank for decomposition.

3. Variational Mode Decomposition

Variational Mode decomposition[12] decomposes the signal into various modes or intrinsic mode functions using calculus of variation. Each mode of the signal is assumed to have compact frequency support around a central frequency.VMD tries to find out these central frequencies and intrinsic mode functions centered on those frequencies concurrently using an optimization methodology called (Alternate Direction Method of Multipliers) ADMM[13]. The original formulation of the optimization problem is continuous in time domain. The constrained formulation is given in[12] as,

$$\min_{u_k,\omega_k}\left\{\sum_k \left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\}$$

$$\text{s.t.}\quad \sum_k u_k = f$$

3.1. Final algorithm for VMD[12]

Initialize $\hat{u}_k^1,\ \hat{\omega}_k^1,\ \hat{\lambda}^1,\ \mathrm{n} \leftarrow 0$

Repeat,

$$n \leftarrow n+1$$

For $k=1:K$
Update $\hat{u}_k$ for all $\omega \geq 0$

$$\hat{u}_k^{n+1} \leftarrow \frac{\hat{f}-\sum_{i<k}\hat{u}_i^{n+1}-\sum_{i>k}\hat{u}_i^{n}+\frac{\hat{\lambda}^n}{2}}{1+2\alpha(\omega-\omega_k^n)^2} \tag{2}$$

Update ω_k:

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega\left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_k^{n+1}(\omega)\right|^2 d\omega} \tag{3}$$

End for
Dual ascent for all $\omega \geq 0$:

$$\hat{\lambda}^{n+1} \leftarrow \hat{\lambda}^n + \tau(\hat{f}-\sum_k \hat{u}_k^{n+1}) \tag{4}$$

Until convergence: $\sum_k \left\|\hat{u}_k^{n+1}-\hat{u}_k^n\right\|_2^2 \Big/ \left\|\hat{u}_k^n\right\|_2^2 < \varepsilon$

The formulation is worded is as follows:
Minimize the sum of the bandwidths of k modes subject to the condition that sum of the k modes is equal to the original signal. So the unknowns are k central frequencies and k functions centered at those frequencies. Since parts of the unknowns are functions, calculus of variation is applied to derive the optimal functions. Bandwidth of an AM-FM signal primarily depends on both, with the maximum deviation of the instantaneous frequency $\Delta f \sim \max\left(\left|\omega_k(t)-\omega_k\right|\right)$ and the rate of change of instantaneous frequency. Dragomiretskiy and Zosso proposed a

function that can measure the bandwidth of a Intrinsic mode function $u_k(t)$. At first they computed Hilbert transform of the $u_k(t)$. Let it be $u_k^H(t)$. Then formed an analytic function $\left(u_k(t)+ju_k^H(t)\right)$. The frequency spectrum of this function is one sided(exist only for positive frequency) and assumed to be centered on ω_k. By multiplying this analytical signal with $e^{-j\omega_k t}$, the signal is frequency translated to be centered at origin. The integral of the square of the time derivative of this frequency-translated-signal is a measure of bandwidth of the Intrinsic mode function $u_k(t)$.

Let

$$u_k^M(t)=\left(u_k(t)+ju_k^H(t)\right)e^{-j\omega_k t} \tag{5}$$

It is a function whose spectrum is around origin (baseband). Magnitude of time derivative of this function when integrated over time is a measure of bandwidth. So

$$\Delta\omega_k=\int\left(\partial_t\left(u_k^M(t)\right)\right)\overline{\left(\partial_t\left(u_k^M(t)\right)\right)}dt \tag{6}$$

Where,

$$\partial_t\left(u_k^M(t)\right)=\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right] \quad . \tag{7}$$

The integral can also expressed as a norm.

$$\Delta\omega_k=\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]\right\|_2^2 \tag{8}$$

The Sum of bandwidths of k modes is given by $\sum_{k=1}^{K}\Delta\omega_k$. The resulting variational formulation is as follows.

$$\min_{u_k,\omega_k}\left\{\sum_k\left\|\partial_t\left[\left(\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right)e^{-j\omega_k t}\right]\right\|_2^2\right.$$

$$s.t.\ \sum_k u_k=f$$

Where f is the original signal. Using augmented Lagrangian Multiplier method the objective function can be casted as an unconstrained optimization problem as follows.

$$L\left(u_k,w_k,\lambda\right)=\alpha\sum_k\left\|\partial_t\left[\left(\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right)e^{-j\omega_k t}\right]\right\|_2^2+\left\|f-\sum_k u_k\right\|_2^2+\left\langle\lambda,f-\sum_k u_k\right\rangle \tag{9}$$

k ranges from 1 to K. In ADMM philosophy, we solve for one variable at a time assuming all others are known. So the formula for updating u_k at the n+1 the iteration is as follows. The convergence of the algorithm is given in[12].

The final updated equations are given by,

$$\hat{u}_k^{n+1} = \left(\hat{f} - \sum_{i \neq k} \hat{u}_i + \frac{\hat{\lambda}}{2} \right) \frac{1}{\left(1 + 2(\omega - \omega_k)^2\right)},\ \omega \geq 0 \tag{34}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k(\omega) \right|^2 d\omega} \tag{35}$$

$$\lambda^{n+1} \leftarrow \lambda^n + \tau\left(f - u_k^{n+1}(t) \right) \tag{36}$$

4. Support Vector Machine

Support Vector Machines (SVM) are widely using supervised classification tool in machine learning. Based on the training data $\{x_1 \ldots x_n\}$, where $x_i \in R^n$, together with class labels $\{y_1 \ldots y_n\}$ where $y_i \in \{-1,1\}$, SVM can be trained to create a model[14]. Using this model, it predicts the class of a new testing sample. The idea in SVM is to create a hyperplane as shown in Fig. 1 with a maximum margin(separation for adjacent classes). This helps to reduce the generalization error for classifying a new data point. Fig. 1 shows a two class problem where a linear separation is achieved using a straight line. But in N-dimensional space, a hyperplane can be used for separation of different classes. In cases where data points are clustered so that linear separation is not possible, the data points can be mapped into feature space (higher dimensional space) where a linear separation is possible. This hyperplane which is linear in feature space will be non linear in its corresponding input space [16].

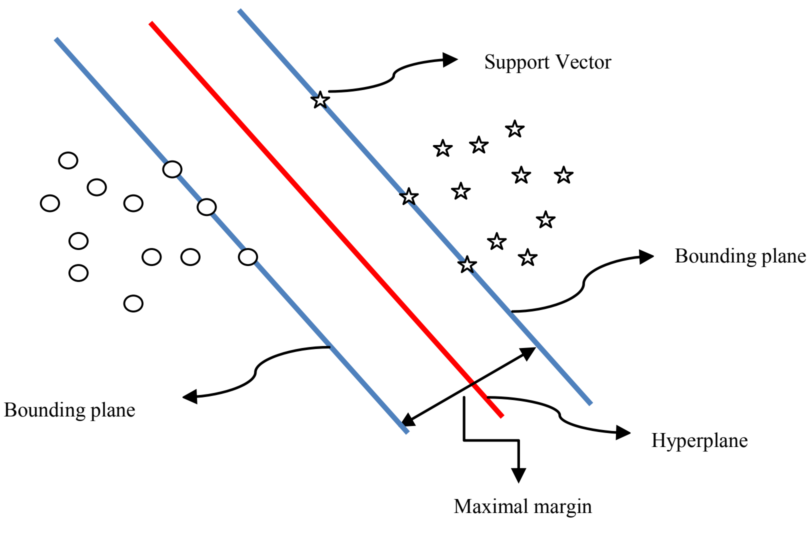

Fig. 1. Linear separation of data points using SVM.

Consider a two case scenario. Let x_1 and x_2 be the two variables, the objective of SVM is to find a hyperplane of the form $w_1x_1 + w_2x_2 - \gamma = 0$ with the bounding planes $w_1x_1 + w_2x_2 - \gamma \geq 1$ and $w_1x_1 + w_2x_2 - \gamma \geq -1$. Here γ is the bias term which is scalar. During the training stage, SVM finds the appropriate $w_{i's}$ and γ. Ones these parameters are found, decision boundary is obtained as $w^T x - \gamma = 0$, where w is a vector. When new data point comes, the decision for the label is found using the function $f(x) = sign(w^T x - \gamma)$

5. Proposed Methodology

Five different classes of power quality disturbances with their corresponding classes considered are swell-K1, sag-K2, harmonics-K3, sag with harmonics-K4 and swell with harmonics-K5. Also a class K6 with pure sinusoidal signal (50Hz frequency) is included. All signal has 50 Hz normal frequency. Number of samples per cycle is 256 and each signal is ten cycle duration, so every signal consists of 2560 number of total samples. Each class contains 120 cases out of which 100 signals feature vector is used for training the SVM and remaining for testing purpose. Fig. 1 shows one such generated signal from all the six classes considered. Each signal is decomposed in to three IMF candidates. Choosing higher the number of candidates may results in good classification accuracy at the cost of increase in the dimension of feature vector computational cost. The total size of training matrix is 600×25. Where 600 includes 100 cases per class multiplied by 6 classes and 25 includes sines and cosines statistical parameters of three IMF candidates together with a label vector. Similarly 120×25 is the dimension of testing data matrix. The statistical parameters considered are variance, kurtosis, median and range.

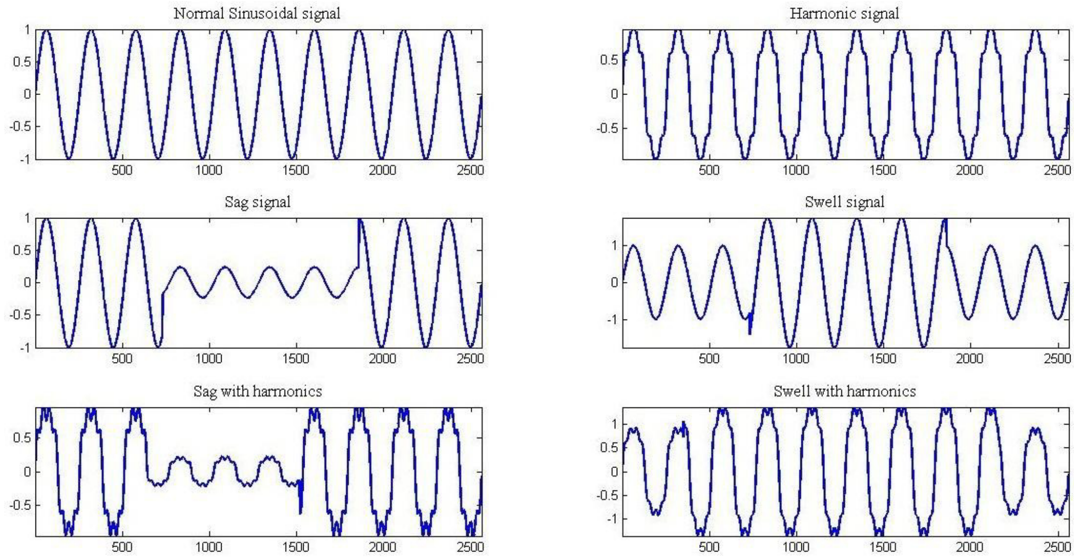

Fig. 1. PQ disturbances Waveform.

6. Results and discussion

In the present work, VMD and EWT algorithm is applied for all cases of power signal disturbances which are shown in Fig. 1. The estimated three components of sag with harmonic signal is Fig. 2. The α value is set to 2000. Very low value of α produces high amount of noises in the decomposed modes. It is noticed that the time complexity of VMD is much higher than EWT. VMD requires around 1918.37 seconds for feature extraction, at the same time EWT requires around 943.14 seconds. The experiments are performed on Windows system having 2.50Ghz Core i5 with 4GB RAM.

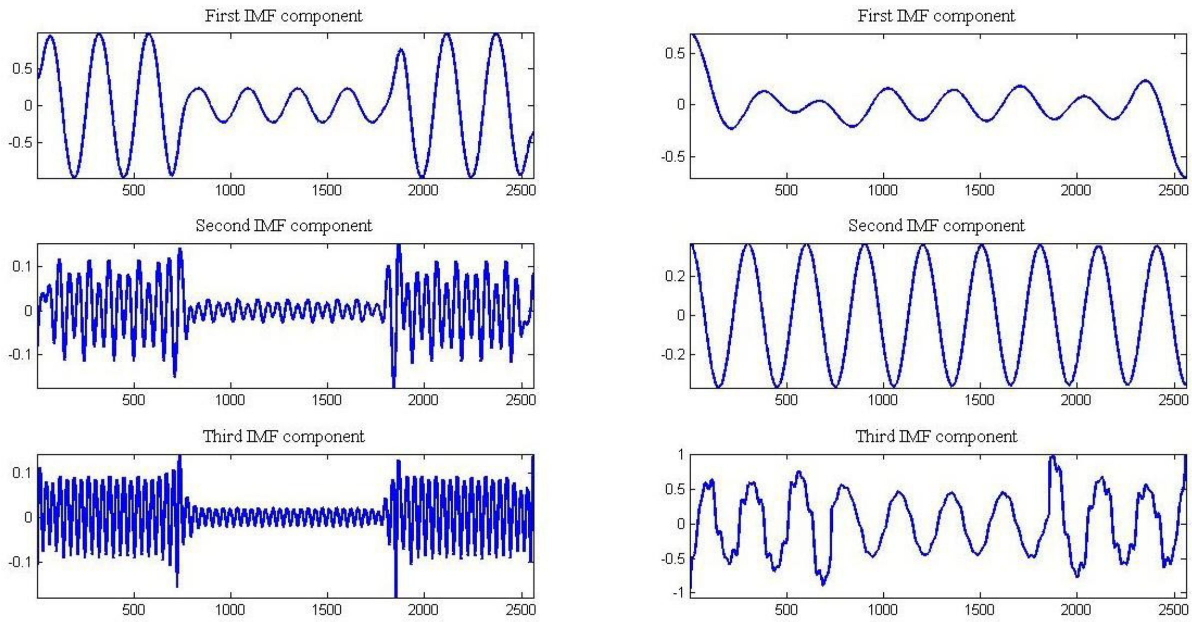

Fig. 2. (a) Three modes extracted using VMD with $\alpha = 2000$; (b) Three modes extracted using EWT.

The classification results using SVM for EWT based feature extraction is shown in Table 1. A 6×6 confusion matrix is created in which diagonal elements shows the correctly classified PQ disturbances and off-diagonal elements represents the misclassified PQ disturbances. The classification accuracy is different for different types of kernels. An optimum selection of kernel function and parameters gives good classification accuracy[16,17,18]. From the confusion matrix obtained, it is noticed that classification results for VMD based feature extraction outperforms the EWT based feature extraction. Table 2 shows the classification results for VMD based feature extraction. Taking more number of statistical parameters with higher number of IMFs may improve the classification accuracy but it will results more time SVM to train.

Table 1. Classification results for EWT using SVM.

	SVM						
Linear Kernel Function	**Class**	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
	K1	19	0	1	0	0	0
	K2	0	18	1	1	0	0
	K3	0	0	20	0	0	0
	K4	0	5	4	11	0	0
	K5	12	0	2	0	6	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**78.33%**					
Polynomial Kernel Function	**Class**	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
	K1	18	0	0	0	2	0
	K2	0	20	0	0	0	0
	K3	0	0	14	1	5	0
	K4	0	1	0	19	0	0
	K5	3	0	0	0	17	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**90%**					

	Class	K1	K2	K3	K4	K5	K6
RBF Kernel Function	**K1**	13	0	6	0	1	0
	K2	1	14	1	0	4	0
	K3	0	0	12	6	2	0
	K4	0	1	0	19	0	0
	K5	1	0	1	0	18	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**80%**					
Sigmoid Kernel Function	**Class**	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
	K1	10	0	4	5	0	1
	K2	1	6	3	10	0	0
	K3	0	0	5	0	15	0
	K4	5	2	1	6	5	1
	K5	0	0	11	0	5	4
	K6	0	0	0	0	0	20
	Overall Accuracy	**43.33%**					

Table 2.Classification results for VMD using SVM

SVM							
	Class	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
Linear Kernel Function	**K1**	17	3	0	0	0	0
	K2	0	16	0	4	0	0
	K3	0	0	20	0	0	0
	K4	0	1	9	10	0	0
	K5	0	0	4	0	16	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**82.50%**					
Polynomial Kernel Function	**Class**	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
	K1	20	0	0	0	0	0
	K2	0	20	0	0	0	0
	K3	0	0	20	0	0	0
	K4	0	0	0	20	0	0
	K5	0	0	0	0	20	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**100%**					
RBF Kernel Function	**Class**	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
	K1	19	0	0	1	0	0
	K2	0	20	0	0	0	0
	K3	0	2	17	1	0	0
	K4	0	0	2	17	1	0
	K5	1	0	0	0	19	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**93.33%**					
Sigmoid Kernel Function	**Class**	**K1**	**K2**	**K3**	**K4**	**K5**	**K6**
	K1	13	0	0	1	6	0
	K2	5	10	0	5	0	0
	K3	0	0	11	6	0	3
	K4	0	3	2	14	0	1
	K5	4	0	0	0	16	0
	K6	0	0	0	0	0	20
	Overall Accuracy	**70%**					

7. Conclusion

This paper considers the classification of power quality disturbances using SVM applied on feature extracted from the signals using EWT and VMD. It is the purpose of this work to introduce recent adaptive signal processing method called VMD which can accurately separate the harmonic components of non-stationary signals regardless of how close their frequency components are. Unlike EWT, without any wavelet filter bank, VMD model is able to generate IMF components concurrently using ADMM optimization method. It can be concluded that the effectiveness of VMD can replace ordinary method like EMD which doesn't have a strong mathematical foundation and EWT which builds adaptive wavelets for signal decomposition. Instead of directly taking the statistical parameters of different modes as feature vector, it is found that the feature vector derived by taking sines and cosines of statistical parameters of modes gives better classification accuracy and less time SVM to converge during training. A good automated system for maintaining the power quality requires detection, identification and classification of power quality disturbances.

References

1. S. Taskin, H. Gokozan. Determination of the spectral properties and harmonic levels for driving an induction motor by an inverter driver under the different load conditions. *Elektronika ir Elektrotechnika (Electronics and Electrical Engineering)*. no. 2. p.75–80. 2011.
2. J. Chung, E. J. Powers, W. M. Grady, and S. C. Bhatt. Power disturbance classifier using a rule-based method and wavelet packet-based hidden markov model. *IEEE Trans. Power Delivery*. vol. 17. p.233–241. Jan. 2002.
3. A. K. Ghosh and D. L. Lubkeman. The classification of power system disturbance waveforms using a neural network approach. *IEEE Trans. Power Delivery*. vol. 10. p. 109–115. Jan. 1995.
4. G. Rata, M. Rata, C. Filote, Theoretical and experimental aspects concerning Fourier and wavelet analysis for deforming consumers in power network. *ElektronikairElektrotechnika (Electronics and Electrical Engineering)*. no. 1. p. 62–66, 2010.
5. S. Chen, H. Y. Zhu. Wavelet Transform for Processing Power Quality Disturbances. *EURASIP Journal on Advances in Signal Processing*. 2007.
6. S. Santoso, J. Lamoree, W. M. Grady, E. J. Powers, and S. C. Bhatt. A scalable PQ event identification system. *IEEE Trans. Power Delivery*. vol. 15. p.738–743. Apr 2000.
7. Jerome Gilles. Empirical wavelet transform. *IEEE transactions on signal processing*. vol. 61, no.*16*. August 15, 2013.
8. N.E. Huang, Z. Shen, S.R. Long, M.L. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu. The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc. LondonSeriesA: Mathematical, Physical and Engineering Sciences*. vol. 454, no.1971. p.903–995. 1998.
9. M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin. A complete ensemble empirical mode decomposition with adaptive noise. *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.(ICASSP)*. p. 4144–4147. May 2011.
10. S. Liu, Q. He, R. X. Gao, and P. Freedson. Empirical mode decomposition applied to tissue artifact removal from respiratory signal. *IEEE Engineering in Medicine and Biology Conference (EMBC)*. p.3624– 3627. Jan. 2008.
11. I. Mostafanezhad, O. Boric-Lubecke, V. Lubecke, and D. P. Mandic. Application of empirical mode decomposition in removing fidgeting interference in doppler radar life signs monitoring devices. *IEEE Engineering in Medicine and Biology Conference (EMBC)*. p.340– 343. Jan. 2009.
12. Konstantin Dragomiretskiy and Dominique Zosso. Variational Mode Decomposition. *IEEE Transactions on Signal Processing*. vol. 62, no. 3. p.531–544. Feb.*2014*.
13. Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato and Jonathan Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*. vol. 3, no.1. p.1–122. 2010.
14. Dr. K. P. Soman, Ajav. V, Loganathan R. Machine Learning with SVM and other Kernel Methods. *PHI Learning Pvt. Ltd*. 2009.
15. Thorsten Joachims. Transductive Inference for Text Classification using Support Vector Machines. *In Proceedings of the Sixteenth International Conference on Machine Learning (ICML '99), Ivan Bratko and SasoDzeroski (Eds.). Morgan Kaufmann Publishers Inc, San Francisco, CA,USA*, p.200-209. 1999.
16. B. Scholkopf, A. Smola, R. C.Williamson, and P. L. Bartlett. New support vector algorithms. *Neural Computation*. vol. 12. p.1207-1245. May 2000.
17. V. Vapnik. Statistical Learning Theory. *Wiley*. 1998.
18. C. Cortes and V. Vapnik. Support-vector network. *Mach. Learning*. vol. 20. p.273–297. 1995.