Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 3, N. 1, 2015.
Trabalho apresentado no XXXV CNMAC, Natal-RN, 2014.

Wavelet Packet Energy-Entropy Feature Extraction and Principal Component Analysis for Signal Classification

Marcus Varanis, **Robson Pederiva**,
Faculty of Mechanical Engineering, Mechanical Design Department, University of Campinas, Brazil
Campinas, SP
E-mail: mvaranis@fem.unicamp.br

Abstract: *This paper has the usage of energy and entropy parameters associated with Wavelet Packet Transform (WPT) as the target to the automatic signal classification as well as the detection of voltage disturbances in electric signals. One can apply Wavelet Packet Transform to remove noise presented in the signals by means of decomposition to obtain the energy and entropy characteristics. Principal component analysis (PCA) is used to reduce the dimensions of the parameters vector and to classify the kinds of signal and disturbances presented using k-nearest neighbor (kNN). Numerical simulations showed the effectiveness of the proposed method.*

1 INTRODUCTION

The poor quality of electric power signals is attributed due to various disorders, such as sags, elevations, interruptions, switching transients, impulses, flicker, harmonics, and notches [3]. Such disturbances may be highly prejudicial to the power grid users. [2]. This paper aims to classify voltage disturbances in electric signals, through the extraction parameters using techniques based on wavelet transform.

2 Basic Concepts

2.1 Wavelet Packet Transform (WPT)

Wavelet packet transform (WPT) is the generalization of the classic wavelet transform. In WPT, the coefficients of detail are decomposed at the first decomposition level, generating what is known in the literature as wavelet packet tree [9]. This process can be viewed in Figure 1. In this way, results are obtained with better time and frequency domain resolution [4]. The wavelet packet transform of a signal x(t) is defined in equation (1).

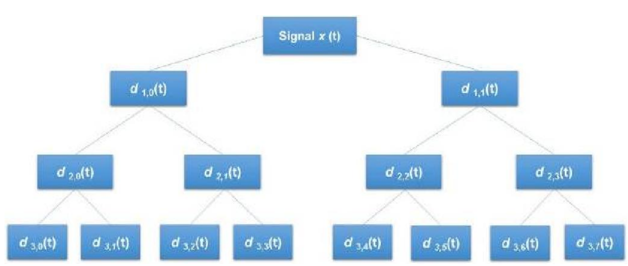


Figure 1: Wavelet Packet decomposition Tree (third Level decomposition).

$$x_p^{n,j} = 2^{-j/2} \int_R x(t) \overline{\mu_n (2^{-j} t - p)} dt \qquad (1)$$

$\mu_n(t)$ is the function wavelet packet; j represents the number of decomposition levels, also known as scale parameters; the symbol p represents the position parameter; n is the number of packets due to decomposition process. To the signal x(t), decomposed by WPT, it is used the expressed by [10]:

$$W_{2n}(t) = \sqrt{2}\sum_{l} h_l W_n(2t-1) \quad (2)$$
$$W_{2n+1}(t) = \sqrt{2}\sum_{l} g_l W_n(2t-1)$$

The operators g and h are known as Quadrature Mirror Filters (QMF) and must satisfy the following orthogonality conditions [12]:

$$HG^* = GH^* \quad (3)$$

$$HH^* = GG^* = I \quad (4)$$

2.2 Wavelet Energy

The calculation energy method of each packet from decomposition WPT is a more robust signal representation than using directly the decomposition coefficients. The energy associated with the wavelet packet decomposition is given by:

$$E_i = \int_{-\infty}^{+\infty} x_j^i(t)dt \quad (5)$$

The total signal energy is expressed by:

$$E_{tot} = \sum_{i=1}^{2^j} E_i \quad (6)$$

The energy of each sub-band is defined as E_i. The normalized energy value, which corresponds to the energy of each wavelet packed is given by:

$$P_l = \frac{E_i}{E_{tot}} \quad (7)$$

where P_l is the probability distribution of each sub-band.

2.3 Wavelet Entropy

The concept of entropy has been widely used as a measure of the system disorder. In this paper, the entropy is obtained by the WPT. The energy probability distribution for each sub-band was given by Equation 7 [8].

Using the definition proposed by Shannon , the entropy is expressed by :

$$S = -\sum P_l \ln(P_l) \quad (8)$$

2.4 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a statistical technique applied in several fields and it is widely used to find patterns in high dimension data, reducing the redundancy of informations presented in a data block through the projection of the original data group onto a subspace which has reduced dimensions and defined some orthogonal variables that contain the most part of the original data variance [5].

Considering $X = (x_1, x_2, ..., x_m)$ a set of data which has the dimension . X can be decoupled in the following way showed by Equation 9:

DOI: 10.5540/03.2015.003.01.0471

$$X = TP^T = \sum_{i=1}^{m} t_i p_i^T \qquad (9)$$

which is defined as principal component vector and p_i is the score vector. The least important components, which generally describe the noise presented in data, can be discarded without losing significant information [7]. This wau the matrix X can be rebuilt as the summation that involves an estimation value $\hat{X}$ and the residual effect E, so:

$$\hat{X} = \sum_{i=1}^{k} t_i p_i^T \qquad (10)$$

$$X = \hat{X} + E \qquad (11)$$

where k represents the number $t_1 .. t_k$ of principal components. The score vectors form a reduced dimension subspace to be used for subsequent analysis [5].

3 Proposed Methodology

The proposed method to classify the signals consists in the application of WPT until the fourth decomposition level, resulting 16 packets. It was used Daubechies-10 filter (Daub-10).

The last decomposition level energy is calculated as well as each packet energy and the relative energy of each packet. The characteristics vector is assembled considering the relative energy values and entropy, resulting a vector with 32 characteristics or dimension 32.

Principal Component Analysis method (PCA) is applied to the data set in order to reduce the size of the vector characteristics to the dimension [1], because the first two principal components corresponds to a value greater than 90% of the system variance.

To perform the classification, the classifier k-nearest neighbor (KNN) type was used.

3.1 Numerical Simulation

In order to test the method´s efficiency, one have created two signal groups. In the first group, called simulation 1, there are four classes of analytical signals. In the second one, named as simulation 2, there are five classes of voltage disturbance in electric signals, which were simulated through parametric equations.

3.1.1 Simulation 1

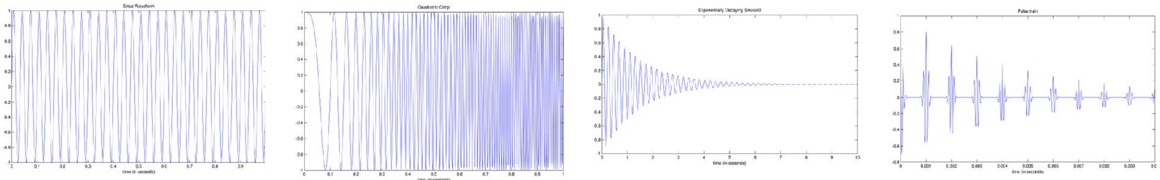
All the signal classes have the addition of white noise. For each signal class, there are 100 training signals and 900 test signals in which the white noise is generated for each time instant. In figure 3, it is shown a sample of each signal class with 1024 points (with noise and without noise, respectively), which allows the signal decomposition until the 10th level (log(1024) / log(2)=10, ou, 2^{10}=1024), using the WPT.

In equation 12, it is shown analytical signals that will be used for processing and classification.

$$\begin{cases} x^{(1)}(t) = \sin(\omega t) + noise(t) \\ \mathrm{x}^{(2)}(\mathrm{t}) = \sin(\omega \mathrm{t}^2) + noise(t) \\ x^{(3)}(t) = \mathrm{pulstran(t)} + noise(t) \\ \mathrm{x}^{(4)}(\mathrm{t}) = \mathrm{e}^{(-0.7*\mathrm{t})} .* \sin(\omega * \mathrm{t}) + noise(t) \end{cases} \qquad (12)$$

Which,

$x^{(1)}(t)$ Sinus Waveform , $x^{(2)}(t)$ Quadratic chirp, $x^{(3)}(t)$ Pulse train, $x^{(4)}(t)$ Exponentially Decaying Sinusoid.



a : Sinus Waveform b : Quadratic chirp c : Pulse train d : Exponentially Decaying Sinusoid

Figure 2: analytical signals that will be used for processing and classification.

3.1.2 Simulation 2

As simulated in case 1, all the signal classes have the addition of white noise. For each class, there are 100 training signals and 900 test signals whereas the white noise is generated for each time instant. All the signals have 1024 points. In figure 3, it is shown a sample of each signal class (with noise and without it, respectively). In equation 12, the parametric equations used to simulate the disturbed signals, which will be used in processing and classification, are presented.

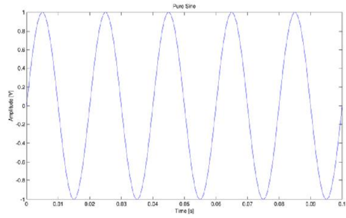
$$\begin{cases} v^{(1)}(t) = \sin(\omega t) \\ \mathrm{v}^{(2)}(\mathrm{t}) = (1 - \alpha_{ss}(1(t-t_b) - 1(t-t_e)))\sin(\omega t) \\ v^{(3)}(t) = (1 + \alpha_{sw}(1(t-t_b) - 1(t-t_e)))\sin(\omega t) \\ \mathrm{v}^{(4)}(\mathrm{t}) = (1 + \alpha_f \sin(\beta_f \omega t))\sin(\omega t) \\ \mathrm{v}^{(5)}(\mathrm{t}) = (\sin(\omega t) + \alpha_{osc}\exp(-(t-t_b)/\tau_{osc})\sin(\omega_{nosc}(t-t_b))) \end{cases} \tag{12}$$

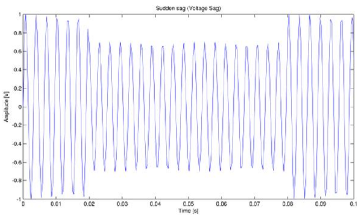
In Table 1 are presented the parameters for generating the analytical signals from the parametric equations.

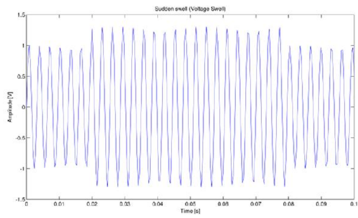
Class (or disturbance)	**Parameters variation**
Sinus Waveform	Amplitude: $1.$ Frequency: $50Hz$
sudden Sag	Duration: $(t_1 - t_2) = (0-9)T$ amplitude: $\alpha_{ss} = 0.3 - 0.9$
Sudden swell	Duration: $(t_2 - t_1) = (0-8)T$ amplitude: $\alpha_{sw} = 0.3 - 0.7$
Flicker	Frequency: $(5-10)Hz$ Amplitude: $\alpha_f = 0.1 - 0.2$
Oscillatory transient	Time const: $(0.008 - 0.04)s$ Frequency: $(100 - 400)Hz$

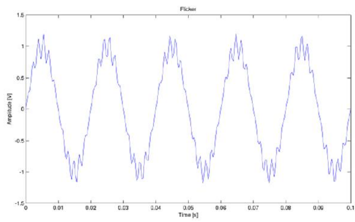
Table 1: Parameters of parametric equations [3].

In Figure 3 presents a sample of each signal of the five simulated signals of of voltage disturbances.


a: Sinus Waveform -Free disturbance


b: Sudden Sag


c: sudden Swell


d: Flicker

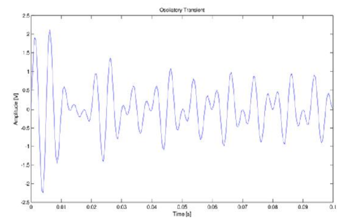
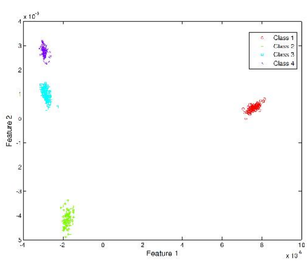

e: Oscillatory transient

Figure 3: Five simulated signals. (a) Pure sinusoid, (b) sudden sag (c) sudden swell (d) Flicker (e) oscillatory transient. (Parameters used to generate the signals shown in Table 1).

4 Results and Discussions

The parameters of energy and entropy are then used to form variables present for classification. For training set generated by the two features of reducing the size of original features vector, produced the clustering result as show in Figure 4-a.

Figure 5-b indicates the scatter plot of the testing set by the two features (feature dimension reduced through the technique of principal component analysis). It is seen that, using the proposed method associated with the KNN classifier, the entire set of test data are classified successfully. To simulate one hit rate ranking of the signals was 100%. The results are presented in Table 2.


a: Training Set

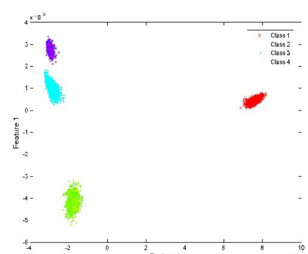

b: Testing Set

Figure 4: Training and Testing Set (Simulation1)

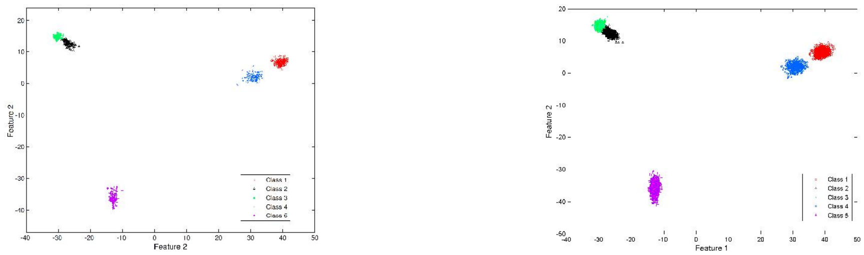
Signal Class	Number of signals per class	Number of Signals correctly classified	Correct classification rate (%)
Class 1	900	900	
Class 2	900	900	
Class 3	900	900	
Class 4	900	900	
Sum	3600	3600	100%

Table 2: Classification rate based on WPT Energy-Entropy and KNN classifier for testing set (simulation 1).

DOI: 10.5540/03.2015.003.01.0471

In simulation 2, the parameters of energy and entropy are then used to form variables present for classification. For training set generated by the two features of reducing the size of original features vector, for five class of signals, produced the clustering result as shown in Figure 5-a.

Figure 5-b indicates the scatter plot of the testing set by the two features (feature dimension reduced through the technique of principal component analysis). It is seen that, using the proposed method associated with the KNN classifier, the entire set of test data, retained high levels of accuracy and were classified successfully.



a: Training Set b: Testing Set

Figure 5: Training and Testing Set) (Simulation2)

The results of simulation 2 are presented in Tables 3 and 4.

Class	Sinus waveform	Sudden sag	Sudden swell	Flicker	Oscillatory transient
Sinus waveform	900	0	0	0	0
Sudden sag	0	889	11	0	0
Sudden swell	0	0	900	0	0
Flicker	0	0	0	900	0
Transient	0	0	0	0	900

Table 3: Signal Classification results for testing set (simulation 2).

Type of disturbance	Number of disturbance	Number of cases correctely identified	Correct classification rate (%)
Sinus waveform	900	900	
Sudden sag	900	888	
Sudden swell	900	900	
Flicker	900	900	
Transient	900	900	
Sum	4500	4488	98,88%

Table 4: Classification rate based on WPT Energy-Entropy and KNN for testing set (simulation 2).

5 Conclusions

This paper proposed the use of energy and entropy parameters for classification of signals and showed that when combined with the technique of dimensionality reduction (PCA), become quite effective. Such efficiency is observed with the use of a very simple algorithm is widely known, k-nearest neighbor, kNN [11].

This paper has the usage of energy and entropy parameters associated with Wavelet Packet Transform (WPT) as the target to the automatic signal classification as well as the detection of voltage disturbances in electric signals. One can apply WPT to remove noise presented in the signals by means of decomposition to obtain the energy and entropy characteristics. Principal component analysis (PCA) is used to reduce the dimensions of the parameters vector and to classify the kinds of signal and disturbances presented using k-nearest neighbor (KNN). Numerical simulations showed the effectiveness of the proposed method.

The accuracy rate for classifying signals (voltage disturbances) were higher than those found in the literature, which use methods based on wavelet transform, but without the use of dimensionality reduction techniques.

ACKNOWLEDGEMENTS

The authors want to express their thanks to CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for the financial support.

REFERENCES

[1] D. Zhao, J. Yan. Performance prediction methodology based on pattern recognition, Signal Processing 912194–2203. 2011.

[2] J. Arrillaga, N. R. Watson, S Chen .Power System Quality Assessment, New York, Wiley, p. 23. 2000.

[3] Janik, P., & Lobos, T. Automated classification of power-quality disturbances using SVM and RBF networks. IEEE Transactions on Power Delivery, 21(3), 1663–1669. 2006.

[4] S Mallat. A Wavelet Tour of Signal Pro-cessing, Third Edition: The Sparse Way. Elsevier. 2009.

[5] Jolliffe, I. T. Principal Component Analysis; Springer-Verlag: New York. 1986.

[6] Yen G, Lin K. Wavelet packet feature extraction for vibration monitoring. IEEE Trans IndElectron 47(3):650–667. 2000.

[7] S.S. Haykin. Neural Networks: A Comprehensive Foundation, 2nd ed., Prentice Hall, New Jersey. 1999.

[8] Z.K. Peng, F.L. Chu. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, Machine Systems and Signal Processing 18 199–22. 2004.

[9]Daubechies I. Ten lectures on wavelets. SIAM, Philadelphia, PA. 1992.

[10] B. Liu. Selection of wavelet packet basis for rotating machinery fault diagnosis, Journal of Sound and Vibration 284 567–582. 2005.

[11] Duda R, Hart P, Stork D. Pattern classification. Wiley-Interscience, New York. 2000.

[12] G. Strang, T. Nguyen. Wavelets and Filters Banks Wellesley-Cambridge Press, Wellesley, USA. 1997.

