Power Quality Disturbance Classification Based on Wavelet Transform and Support Vector Machine

[1]J. A. Bosnic, [2]G. Petrovic, [3]A. Putnik, [4]P. Mostarac

[1,2,3]Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, Split, Croatia

[4] Faculty of Electrical Engineering and Computing University of Zagreb, Zagreb, Croatia

Email: jbosnic@fesb.hr

Abstract. *This paper presents an effective approach for classification of power quality (PQ) disturbances based on wavelet transform (WT) and support vector machine (SVM). Wavelet transform was applied to disturbance signal in order to obtain decomposition coefficients at six levels that represents signal in time and frequency domain. Eight statistical methods were used to extract features that characterize each disturbance signal. Forward sequential feature selection was then applied to the feature vector to identify the most discriminative features. SVM method was used for designing the classifier which is trained with the data simulated in MATLAB. High classification accuracy, reliability and robustness of the proposed classifier were confirmed on the testing data in noisy and noiseless environment.*

Keywords: Wavelet Transform, SVM, Power Quality (PQ), Feature Extraction

1. Introduction

Power quality (PQ) has become of a great interest in recent decades due to increased number of sensitive electrical and electronic devices which require reliable and high quality power supply. PQ deterioration has been caused by extensive use of power electronic devices and non linear loads as well as renewable energy production. Effective and accurate detection, localization and identification of PQ disturbances is ste first and the most important step in power disturbance mitigation process. The identification of PQ disturbances can be divided into two stages: signal processing and feature extraction and classification of PQ disturbances. Since voltage disturbances are generally non-stationary signals, various time-frequency signal transformation techniques have been used to represent disturbance signal simultaneously in time and frequency domain, such as WT [1,2], S-transform [3], TT-transform [4] and Hilbert Huang transform [5]. This kind of transformed data is not suitable to be used as a input of the classifier because of the large memory usage that would significantly increase computational time and reduce classification accuracy. Several different feature extraction methods have been utilized to obtain features that could characterize distortion and deviations in the voltage waveforms. In [2,3] various statistical techniques have been applied to signals in time and frequency domain to obtain distinctive features. Second part of PQ disturbance identification process is designing a classifier that will diferentiate among disturbance classes. Probabilistic neural network [6], fuzzy decision tree [7] and support vector machine (SVM) [1,2,3] are just one of the artifical intelligence tools employed for designing the classifier. SVM has emerged as one of the most popular techniques for classification of PQ disturbances due to its ability to obtain high classification performance for high-dimensional, non-linearly separable data.

In this paper, wavelet transform, eight feature extraction methods and feature selection technique were used to obtain distinctive features of voltage disturbances. SVM classifier was trained using distinctive features and classification performance in noisy and noiseless environment were verified on the independent data set.

2. Subject and Methods

Wavelet Transform

Power quality (PQ) disturbances that will be detected and classified in this paper are mostly nonstationary signals (time-variant frequency domain characteristics). Thus Fourier transform isn't adequate for analysis of this kind of signals because it assumes stationary signals. Wavelet transform is answer to this problem because it provides us with time-frequency representation of the nonstationary signal. While FT uses sine waves of different frequencies and infinite length, WT represents the signal as a sum of wavelets of finite and variable length at different positions. Different wavelets are obtained by applying scaling and translation on mother wavelet. Mathematical theory behind WT is explained in detail in [8].

Feature Extraction and Feature Selection

Feature extraction and feature selection are the most significant procedures in event recognition system. Classifiers that are build using poorly selected features will almost never yield good classification performance. The detail and approximation coefficients obtained from WT are not suitable to be used as classifier input because of their large data size. The aim of the feature extraction process is to identify the most relevant features, features that can discriminate between different events. The forward feature selection (FFS) is applied after feature extraction to further reduce feature number by discarding features with low discrimination ability. Consequences of applying this dimensionality reduction procedures is reduced data size of the classifier input (lower computational time) and higher classification accuracy (decreased chance of overfitting).

Distinctive features are obtained by applying different feature extraction methods to detail coefficients at every decomposition level and to approximation coefficients at last decomposition level [2]. Mean, standard deviation, skewness, kurtosis, RMS, Shannon entropy, log energy entropy and norm entropy are feature extraction methods used in this paper. This feature extraction methods are just some of many possible signal analysis methods that can be utilized to characterize original or transformed signal.

Feature selection process was realized in the next step to further decrease number of the features. The method used in this paper was forward sequential feature selection (FFS). FFS selects a subset of features by sequentially adding features starting from the empty set until the criterion is met. The criterion is to minimize 10-fold cross-validation (CV) error which is computed for every feature subset. This feature selection method is model-dependent (depends on the algorithm used for computing CV error). The classification algorithm used in this paper is support vector machine (SVM) which main principles will be presented in the next chapter.

Support Vector Machine

SVM is supervised learning algorithm that has gain high popularity in recent years for solving classification problem across different areas, from face and voice recognition to classification of PQ disturbances. SVM is particularly useful when dealing with high-dimensional, nonlinearly separable data. SVM's statistical foundations and mathematical formulation are well-known and well-documented so it will not be the topic of consideration. Refer to [9] for complete review of SVM's theory and application.

SVM is fundamentally a binary (two-class) classifier. Multi error-correcting output code (ECOC) technique was used in this paper for dealing with the multiclass classification problems. Multi ECOC technique is based on a reduction of multiclass classification problems to a set of binary SVM classifiers where certain decoding scheme and coding design are used

Table 1. The results of FFS algorithm for noiseless signals

Indices of the training data set	Classification accuracy (%)	Number of selected features	The first best 10 features
1	99.25	16	39, 46, 22, 15, 40, 8, 4, 44, 47, 19
2	99.50	25	36, 42, 32, 31, 23, 7, 43, 11, 45, 15
3	99.12	25	28, 46, 32, 31, 40, 39, 15, 1, 33, 7
4	99.69	23	36, 48, 20, 47, 23, 8, 4, 26, 44, 14
5	99.62	24	36, 42, 32, 31, 39, 7, 23, 15, 9, 1
6	99.38	21	36, 46, 32, 31, 7, 23, 8, 42, 47, 4
7	**99.50**	**15**	**36, 48, 20, 47, 7, 8, 4, 30, 14, 44**
8	99.50	24	36, 46, 32, 31, 7, 15, 23, 41, 35, 17
9	99.19	25	28, 42, 32, 31, 40, 7, 33, 9, 15, 25
10	99.50	25	36, 46, 32, 31, 7, 23, 27, 28, 15, 17

for the prediction of classification results according to binary SVM classifiers predictions. Refer to [10] for more information about ECOC multiclass classification.

3. Results

PQ disturbances were simulated in MATLAB based on equations given in [3]. 200 instances of each class were randomly generated for both training and testing process. 5-level WMRA was applied to training and testing data set to obtain detail and approximation coefficients. Feature extraction methods were then applied to decomposition coefficients to obtain feature vector for every signal. Features were then standardized to avoid classifier sensitivity to different feature ranges.

Thus, total size of the training and testing data set is 1600×48. FFS algorithm was applied to the 10 randomly generated training data sets to analyse relationship between training data sets and the selected features. Classification accuracy that is evaluated by the FFS algorithm, number of features selected by FFS algorithm and the first best 10 features (for practical reasons) are shown in Table 1. If only the first best 5 features are considered, certain pattern of best features can be recognized. But because of the high classification accuracy, size of the feature vector that is related to the computational time should be the decisive factor at selecting appropriate feature vector. Data set used for training the classifier is displayed in bold.

Also, to study the performance of the proposed classification algorithm in a noisy environment, same signals were added by Gaussian white noise with the SNR 30 dB. Same signal processing methods were also applied to a noisy signals and Table 2 displays results of FFS algorithm for noisy signals and selected data set.

Power system disturbances classifiers for noisy and noiseless signals are constructed using multiclass ECOC SVM algorithm. Coding design, kernel scale parameter σ of Gaussian kernel function and penalty parameter C were determined using Bayesian optimization algorithm [11]. Following values were obtained: one-versus-one coding design, $\sigma = 57.56$, $C = 9885.4$ for noisy signal classifier and one-versus-all coding design, $\sigma = 106.55$, $C = 9523.9$ for noiseless signal classifier. Classification results for the testing data are given in Table 3 and Table 4 where diagonal elements represent correctly classified PQ disturbances and others represent the misclassified PQ disturbances.

Table 2. The results of FFS algorithm for the signals with the 30 dB noise

Indices of the training data set	Classification accuracy (%)	Number of selected features	The first best 10 features
1	96.25	25	48, 22, 47, 31, 1, 19, 39, 42, 26, 36
2	96.38	20	47, 22, 48, 31, 36, 40, 42, 39, 38, 45
3	96.50	25	47, 48, 30, 31, 27, 43, 9, 1, 26, 23
4	95.75	24	48, 30, 47, 31, 9, 33, 11, 41, 3, 29
5	96.19	24	29, 46, 36, 47, 10, 30, 45, 31, 4, 41
6	96.06	17	45, 29, 47, 39, 31, 16, 42, 32, 41, 33
7	96.38	22	36, 45, 26, 48, 1, 9, 11, 19, 17, 32
8	96.19	23	32, 46, 47, 36, 16, 31, 39, 26, 1, 23
9	96.69	21	47, 22, 48, 31, 36, 8, 41, 40, 45, 44
10	**96.56**	**16**	**32, 46, 47, 36, 16, 31, 29, 30, 23, 40**

Table 3. Classification results for noiseless signals

True classes	Predicted classes							
	C1	C2	C3	C4	C5	C6	C7	C8
C1	198	0	0	0	0	0	0	2
C2	0	193	7	0	0	0	0	0
C3	0	0	199	0	0	0	0	1
C4	0	0	0	200	0	0	0	0
C5	0	0	0	0	200	0	0	0
C6	0	0	0	0	0	200	0	0
C7	0	0	0	0	0	0	199	1
C8	0	0	0	0	0	0	0	200
Overall accuracy (%): 99.31								

Table 4. Classification results for noisy signals

True classes	Predicted classes							
	C1	C2	C3	C4	C5	C6	C7	C8
C1	191	0	0	0	0	0	0	9
C2	0	192	7	0	0	0	0	1
C3	0	0	200	0	0	0	0	0
C4	0	0	0	200	0	0	0	0
C5	0	0	0	0	200	0	0	0
C6	0	0	0	1	0	199	0	0
C7	0	0	0	0	0	0	198	2
C8	1	0	0	0	0	0	0	199
Overall accuracy (%): 98.69								

As can be seen from Table 3 and Table 4, classifier shows good classification performance in both noiseless and noisy environment. Misclassification at noiseless and noisy signals occurs for sag signals because of the similarity between sag and interruption amplitude ($a = 0.9$). Also, it can be noticed that swell signals with added noise can be misclassified as flickers. Other classes have been classified correctly.

4. Conclusions

PQ classification system based on wavelet transform and SVM is presented in this paper. Wavelet transform was used for presenting the signal as a set of decomposition coefficients at 6 different levels. Eight different feature extraction methods were then applied to decomposition coefficients to obtain feature vector for every disturbance signal. The most distinctive features were then determined by applying FFS to the feature vector. SVM ECOC technique was utilized for designing the classifier, separately for noisy and noiseless signals. Classification results show that proposed classification system has good performances in both noisy and noiseless environment.

Further improvement of this study can be done if other signal transformation methods and/or feature extraction methods are utilized. Also, classification performance of the classifier trained on simulated data will be investigated on data generated from Simulink and on real power system data.

Acknowledgements

This paper is fully supported by Croatian Science Foundation under the project Metrological infrastructure for smart grid IP-2014-09-8826.

References

[1] Borras MD, Bravo JC, Montano JC. Disturbance ratio for optimal multi-event classification in power distribution networks. *IEEE Trans. Ind. Electron.*, 63(5): 3117-3124, 2016.

[2] Eristi H, Ucar A, Demir Y. Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines. *Electric Power Systems Research*, 80(7): 743-752, 2010.

[3] Li J, Teng Z, Tang Q, Song J. Detection and classification of power quality disturbances using double resolution S-transform and DAG-SVMs. *IEEE Trans. Instr. Meas.*, 65(10): 2302-2312, 2016.

[4] Biswal B, Biswal MK, Dash PK, Mishra S. Power quality event characterization using support vector machine and optimization using advanced immune algorithm. *Neurocomputing*, 103: 75-86, 2013.

[5] Ozgonenel O, Yalcin T, Guney I, Kurt U. A new classification for power quality events in distribution systems. *Electric Power Systems Research,* 95: 192-199, 2013

[6] Mishra S, Bhende CN, Panigrahi BK. Detection and classification of power quality disturbances using S-transform and probabilistic neural network. *IEEE Trans. Power Del.*, 23(1): 280-287, 2007.

[7] Biswal M, Dash PK. Measurement and classification of simultaneous power signal patterns with an S-transform variant and fuzzy decision tree. *IEEE Trans. Ind. Informatics*, 9(4): 1819-1827, 2012.

[8] Daubechies I. Ten Lectures on Wavelets. SIAM, Philadelphia, 1992.

[9] Vapnik VN. Statistical Learning Theory. John Wiley & Sons, New York, 1998.

[10] Allwein EL, Schapire RE, Singer Y. Reducing multiclass to binary: A unifying approach for margin classifiers. Journal of Machine Learning Research, 1: 113-141, 2000.

[11] Mockus J. Bayesian Approach to Global Optimization. Kluwer Academic Publishers, Dordrecht, 1989.