EFFICIENT FEATURE EXTRACTION AND CLASSIFICATION OF POWER QUALITY DISTURBANCES

Marija Markovska, Dimitar Taškovski

Faculty of Electrical Engineering and Information Technologies,
"Ss. Cyril and Methodius" University in Skopje,
Rugjer Bošković bb, P.O. box 574, 1001 Skopje, Republic of Macedonia
marijam@feit.ukim.edu.mk

A b s t r a c t: The increasing number of renewable energy sources, together with the installation of modern control equipment requires higher power quality (PQ) in the generation, transmission and distribution systems. In order to maintain and improve the power quality, power disturbances should be monitored continuously. Power quality monitoring and analysis must be able to detect and classify the disturbances present in voltage or current waveforms. In this paper a method for feature selection and classification of power quality disturbances using wavelet transform (WT) and random forest (RF) algorithm is proposed. The classification results for seven and ten PQ classes are compared with results obtained by applying some previously published methods, proposed by different authors. Moreover, a comparison of the classification accuracies obtained in noisy environment, by using our method and the other proposed wavelet based methods, is made. The investigation has shown that the proposed method represents an efficient method for feature extraction and classification with high accuracy.

Key words: power quality; wavelet analysis; classification; feature extraction

ЕФИКАСНО ИЗВЛЕКУВАЊЕ НА КАРАКТЕРИСТИКИ И КЛАСИФИКАЦИЈА НА НАРУШУВАЊАТА НА ЕЛЕКТРИЧНАТА ЕНЕРГИЈА

А п с т р а к т: Зголемениот број на обновливи извори на електрична енергија, заедно со инсталацијата на модерна опрема за управување, има потреба од висок квалитет на енергијата во системите за нејзино производство, пренос и дистрибуција. Со цел да се одржува и подобрува квалитетот на енергијата, потребно е тој да биде континуирано мониториран. Мониторирањето и анализата на квалитетот на електричната енергија треба да овозможат детекција и класификација на нарушувањата присутни во напонот или струјата. Во овој труд е предложен метод за избор на карактеристики и класификација на нарушувањата на електричната енергија со употреба на вејвлет-трансформација и "random forest" (RF) алгоритам за класификација. Резултатите од класификацијата на седум и десет класи се споредени со резултати добиени при примена на методи порано предложени од разни автори. Дополнително е направена споредба на точноста на класификациите добиени при присуство на шум со примена на нашиот метод и на порано предложените методи базирани на вејвлет. Истражувањето покажа дека предложениот метод претставува ефикасен метод за извлекување на карактеристики и класификација и дека се одликува со висока точност.

Клучни зборови: квалитет на електрична енергија; вејвлет-анализи; класификација; извлекување на карактеристики

1. INTRODUCTION

Nowadays, when power consumption continuously increases, the world is facing potential energy shortages as oil and gas supplies run down. Together with the requirements for reducing carbon emission in the environment leads to increased proliferation of renewable energy sources. The integration of this sources as well as smart transmission systems, well equipped with modern control equipment, increases the application of nonlinear devices in distributions systems. These devices produce one of the major power quality (PQ) disturbances. Contrary to that, utilities and consumers of electric power are expected to acquire an ideal voltage and current waveform at rated frequency. In order to ensure high quality power supply of the power grid, it is very important to analyze and recognize these PQ

disturbances, where the key problem is how to extract the feature vectors automatically from large amount of PQ data and how to classify the PQ disturbances with high accuracy [1].

Several methods for automated detection and classification of PQ disturbances have been published recently. Some frequently used artificial intelligence (AI) based classifiers are rule-based expert systems (ES) [2], fuzzy logic (FL) classification systems [3], artificial neural networks (ANN) [4], support vector machines (SVM) [5], but none of them is robust as the data mining based classifiers [6]. All these methods use feature vectors derived by the process of feature selection from disturbance waveforms to classify power quality events.

The researchers have introduced different digital signal processing techniques that can be used in the process of extraction features that characterize PQ disturbances [7], such as the fast working Fourier transform (FT), the short-time Fourier transform (STFT), the S-transform (ST), the Neural Network, the Fuzzy logic and the wavelet transform (WT) [8, 9]. Among them, wavelet transform has been used extensively in the last years. Wavelet transform analysis approach gives time and frequency information accurately, by convolving the dilated and translated wavelet with signal. This feature makes the wavelet transform well suited for the analysis of the power system transients caused by various PQ disturbances [10].

Large number of feature extraction and classification methods has been proposed in the literature. Here, we will consider several of them used for comparison purposes later in this paper. In [11] a learning-based method is used for classification of seven different types of PQ disturbances, including the case with no power disturbance. A decision tree (DT) is created, using wavelet analysis in the feature extraction process. In [12] the same feature vector is used, as in [11], however the learning-based method is different. A type of neural network, called SOLAR (self-organizing learning array), is used for classification of the power disturbances. A comparative study using SVM is also done, by obtaining accuracy, which in some cases is very close to the accuracy obtained using SOLAR. In [13], wavelet norm entropy-based effective feature extraction method for PQ disturbances classification is presented, where also seven types of PQ disturbances are considered. The disturbance classification schema is performed with the wavelet neural network (WNN). It performs a feature extraction and a classification algorithm composed of a wavelet feature extractor based on norm entropy and a classifier based on a multi-layer perception.

In this paper, an efficient wavelet based method for feature selection and classification of PQ disturbances is presented. The proposed method is experimentally tested and compared with the different feature extraction approaches used in [11–16]. It is shown that the wavelet feature extraction based on energy and log-energy entropy together with the random forest (RF) algorithm gives more accurate classification results, compared to the other methods, for both number of seven and ten PQ classes and in case of 20 dB, 30 dB, 40 dB and 50 dB white Gaussian noise (WGN). RF is an effective data mining based classifier, especially when dealing with massive amounts of data. It has more advantages, compared to the other type of classifiers, such as good anti-noise performance, less parameters and less influence of the over-fitting problem. Furthermore, RF has better generalization ability than DT [17]. In the verification of multiple public data sets it is shown that the classification accuracy of RF is the highest among all classification methods [18].

2. DISCRETE WAVELET TRANSFORM

The continuous wavelet transform of a signal $f(t)$ is given with the following expression:

$$CWT(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{ab}(t)\,dt \qquad (1)$$

where the function $\psi(t)$ is the base function or the mother wavelet and is defined as:

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad a,b \in \mathfrak{R}; a \neq 0. \qquad (2)$$

The parameters a and b are the dilation and translation parameters, respectively. Since the transformation is achieved by dilating and translating the mother wavelet continuously, it generates substantial redundant information. Therefore, instead of continuous dilation and translation, the mother wavelet can be dilated and translated discretely by selecting $a = a_0^m$ and $b = nb_0 a_0^m$, where a_0 and b_0 are fixed constants with $a_0 > 1$ and $b_0 > 0$, $m, n \in \mathbb{Z}$ and $\mathbb{Z}$ is the set of integers. The resulting expression is the discrete wavelet transform (DWT) which is defined as:

$$DWT[m,n] = \frac{1}{\sqrt{a_0^m}} \sum_{k=-\infty}^{\infty} f[k]\psi\left[\frac{k - nb_0 a_0^m}{a_0^m}\right]. \qquad (3)$$

With the choice $a_0 = 2$ and $b_0 = 1$, a dyadic orthonormal wavelet transform is obtained [19, 20],

and can be easily and quickly implemented by filter bank techniques known as Multi-Resolution Analysis (MRA). The filter bank is used to decompose the signal into various levels using a low-pass filter with a transfer function $H_0(z)$ and a high-pass filter with a transfer function $H_1(z)$, as shown in Figure 1.
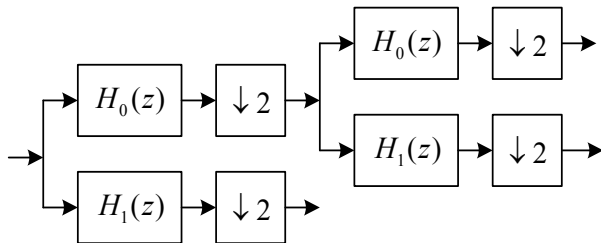

Fig. 1. Wavelet decomposition over 2 levels

The basic idea of the MRA is that of the successive approximation, together with that of "added detail". The low frequency part (approximation signal) is split again into two parts of high and low frequencies. Depending on the application and on the size of the input signal, the process could be repeated several times. As result, logarithmic decomposition of the frequency spectra of the input signal is obtained.

3. FEATURE EXTRACTION METHOD

The PQ disturbance signals in the feature extraction process are decomposed using DWT. Compared to the discrete Fourier transform (DFT), DWT is a better choice for the analysis of the PQ disturbances, since it provides not only frequency information, but also information about time localization of the components. Thus, the DWT has been used intensively in the last years. The wavelet analysis is in fact a measure of similarity between the basis function, so called mother wavelet, and the signal itself. Therefore, the selection of the wavelet mother function is one of the key factors for designing a successful wavelet application. As general rule, for detection of fast transient disturbances, shorter filters are proposed as better, while for slow transient disturbances long filters are presented as particularly good [21, 22]. Thus, selection of best filter length for detection and classification is not an easy task. In this work Daubechies wavelet filter is used, as one of widely used wavelets in power quality problems.

The number of decomposition levels, is also of significant importance. Firstly, higher number of decompositions levels means more calculations which results in increased computational cost. Secondly, choosing higher l will, generally, bring more information in the system, i.e. extended feature vector, and in that way higher accuracy. Experimental results show that the increasing the number of levels after $l = 6$ do not considerably affect the accuracy of algorithm [14]. Thus, in this work we have chosen $l = 6$.

With the resulting detail coefficients at each decomposition level and approximation coefficients at the last level feature vector is extracted. Different feature extraction techniques have been proposed using the wavelet based features presented in [16]. In [11–14] a feature extraction approaches are presented using one of the wavelet based features, i.e. in [13] wavelet based norm entropy is used, while in the others wavelet based energy is used. In [15] three of the features are used: energy, standard deviation and Shannon entropy. In [16] all nine features are used, which are: mean, standard deviation, skewness, kurtosis, RMS, energy, Shannon entropy, log-energy entropy, norm entropy. These approaches are considered in Section 5, for comparison purposes of the accuracy of the feature extraction method proposed in this paper.

The proposed DWT based feature extraction process is performed by using two wavelet based features, energy and log-energy entropy, given with the following equations:

$$SE_i = -\sum_{j=1}^{N} C_{ij}^{\,2} \log\left(C_{ij}^{\,2}\right) \qquad (4)$$

$$LOE_i = \sum_{j=1}^{N} \log(C_{ij}^2). \qquad (5)$$

In the equations, $i = 1,2,3,\dots,l$ represents the level of decomposition, C_{ij} stands for detail coefficients at each multiresolution level and approximation coefficients from the last level, while N is the number of those coefficients. The feature vector, constructed with the proposed feature extraction method is an input of the random forest classifier.

4. RANDOM FOREST CLASSIFIER

Random forest represents a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The RF classifier is defined as:

$$\left\{ f(x,\delta_k), k = 1,...,\right\}, \qquad (6)$$

where $f(x,\delta_k)$ is a meta classifier and it is a tree construct classifier that can be formed by several algorithms, x is the input vector, while δ_k are k random vectors and each of them determines the growth of a single decision tree. The RF summarizes the classification results of different decision trees in order to achieve the optimal classification results. The classification ability of RF can be measured by analyzing its generalization error, an important index for measuring the extrapolation ability of the classifier [17]. Given a set of classifiers $f_1(x), f_2(x), ..., f_k(x)$, and with the training set drawn by random sampling of the original data set (X, Y), define the margin function as:

$$mg(X,Y) =$$
$$= av_k I(f_k(X) = Y) - \max_{j \neq Y} av_k I(f_k(X) = j), \quad (7)$$

where $I(\cdot)$ is an indicator function, $av_k(\cdot)$ is the average value, Y is correct classification of the vector, j is an incorrect classification of the vector. The margin function measures the extent to which the average number of votes for the correct class exceeds the average number of votes for any other class. The larger the margin, the better the performance of the classification. The generalization error is given with following equation:

$$PE^* = P_{X,Y}(mg(X,Y) < 0), \qquad (8)$$

where the index X, Y defines space. In RF algorithm $f_k(x) = f(x, \delta_k)$. For a large number of trees, from the Strong Law of Large Numbers and the tree structure, it can be known that:

$$\lim_{k \to \infty} PE^* = P_{X,Y}(P_\delta(f(X,\delta) = Y) -$$
$$\max_{j \neq Y} P_\delta(f(X,\delta) = j) < 0). \qquad (9)$$

Here, $P_\delta((X, \delta) = Y)$ denotes the probability of the classification results as the correct class, while $P_\delta((X, \delta) = j)$ denotes the maximum probability of the classification results as any other class. Hence, this results explains why random forest do not overfit as the number of trees increases, but produce a constant value of PE^*.

RF has simple structure, good generalization ability and anti-noise performance [23]. Compared to other classification techniques, it can achieve higher classification accuracy and has lower time

complexity. Thus, it can satisfy the demands of situations when real-time performance is highly needed. Although RF can meet application needs of massive PQ data for signal classification.

The steps of the classification process of RF classifier used in this paper and referring to [24] are presented with a flow diagram given in the following figure.
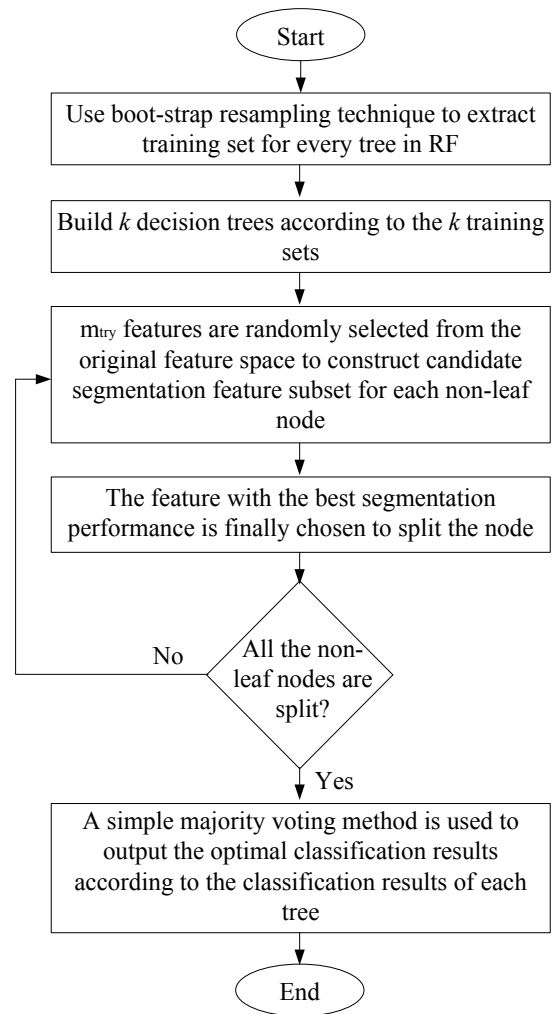

Fig. 2. Random forest flow diagram

5. EXPERIMENTAL AND COMPARATIVE CLASSIFICATION RESULTS

Different types of PQ disturbances are generated in Matlab in order to test the proposed feature extraction and classification method. The experimental and comparative testing was made using two separate groups of PQ signals. In the first group seven different classes are considered, including the case with no power disturbances, with class labels given in Table 1.

T a b l e 1

Class labels of the PQ disturbances

PQ disturbances	Class labels
Normal	C1
Swell	C2
Sag	C3
Harmonic	C4
Interruption	C5
Sag with harmonic	C6
Swell with harmonic	C7

T a b l e 2

Class labels of the PQ disturbances

PQ disturbances	Class labels
Sag	V1
Swell	V2
Interruption	V3
Transient	V4
Harmonics	V5
Sag with harmonic	V6
Swell with harmonic	V7
Notch	V8
Spike	V9
Flicker	V10

Ten cycles are included in every signal with a sampling frequency of 256 samples/cycle (12.82 kHz), so that every signal has 2560 samples. The fundamental frequency is assumed to be 50 Hz. The learning of the classifier was done using training data set with size of 1400 examples. In the training and testing data sets each of the training and testing classes has size of 200 data examples.

The second group is consisted from increased number of PQ disturbance classes, compared to the first one. Namely, in this group 10 different classes are considered for classification, denoted with the class labels that are given in Table 2.

Ten cycles are included in every PQ signal with a sampling frequency of 64 samples/cycle (3.2 KHz), so that every signal is consisted of 640 samples. The fundamental frequency is also 50 Hz.

In this group, 20100 PQ signals are generated for feature extraction. The training and testing data sets have size of 10050 signals, and correspondingly each class has size of 1005 examples.

For the comparison purposes of this paper the power disturbances for the both groups are modelled in the same way as they are modelled in [11–16]. That includes the same disturbance types and the same mathematical definitions, which are given in Table 3. Also for the aim of the comparison, the generated signals are synthesized in noisy environment by adding 20 dB, 30 dB, 40 dB and 50 dB evenly distributed WGN..

T a b l e 3

Mathematical definitions for the used PQ disturbances

Disturbance	Model	Parameters
Normal	$x(t) = \sin(\omega t)$	$f = 50$ Hz, $\quad \omega = 2\pi f$
Sag	$x(t) = [1 - \alpha(u(t - t_1) - u(t - t_2))]\sin(\omega t)$	$0.1 \le \alpha \le 0.9, \quad T \le t_2 - t_1 \le 9T$
Swell	$x(t) = [1 + \alpha(u(t - t_1) - u(t - t_2))]\sin(\omega t)$	$0.1 \le \alpha \le 0.8, \quad T \le t_2 - t_1 \le 9T$
Interruption / outage	$x(t) = [1 - \alpha(u(t - t_1) - u(t - t_2))]\sin(\omega t)$	$0.9 \le \alpha \le 1, \quad T \le t_2 - t_1 \le 9T$
Flicker	$x(t) = [1 + \alpha \sin(2\pi\beta t)]\sin(\omega t)$	$0.1 \le \alpha \le 0.2, \quad 5$ Hz $\le \beta \le 20$ Hz
Oscillatory transient	$x(t) = \sin(\omega t) + \alpha \exp(-(t - t_1)\tau)(u(t - t_1) - u(t - t_2))\sin(2\pi f_n t)$	$0.1 \le \alpha \le 0.8, \quad 8$ ms $\le \tau \le 40$ ms $0.5T_n \le t_2 - t_1 \le 3T_n$, 300 Hz $\le f_n \le 900$ Hz
Harmonics	$x(t) = \alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t)$	$0.05 \le \alpha_3, \quad \alpha_5, \quad \alpha_7 \le 0.15, \quad \sum \alpha_i^2 = 1$
Notch	$x(t) = \sin(\omega t) - \mathrm{sgn}(\sin(\omega t))\left\{ \sum_{k=0}^{9} k[u(t - (t_1 + 0.2n)) - u(t - (t_2 + 0.2n))] \right\}$	$0.1 \le k \le 0.4, \quad 0 \le t_1, \quad t_2 \le 0.5T$ $0.01T \le t_2 - t_1 \le 0.05T$
Spike	$x(t) = \sin(\omega t) + \mathrm{sgn}(\sin(\omega t))\left\{ \sum_{k=0}^{9} k[u(t - (t_1 + 0.2n)) - u(t - (t_2 + 0.2n))] \right\}$	$0.1 \le k \le 0.4, \quad 0 \le t_1, \quad t_2 \le 0.5T$ $0.01T \le t_2 - t_1 \le 0.05T$

The first data group is used for two different experimental and comparative testings of the proposed method. The first testing is made only for the accuracy of the proposed feature extraction method. Classification results obtained using our method are compared with the results obtained using the methods proposed in [11, 12, 14–16], considering same training and testing sets and same RF classifier for all of the methods. The results obtained for seven types of pure PQ disturbances are presented in Table 4.

T a b l e 4

Comparison of the feature extraction results of the proposed method for 7 PQ classes

Class	Feature extraction method			
	Ref [11], [12], [14]	Ref [15]	Ref [16]	This paper
C1	100	100	100	100
C2	98	100	100	94.5
C3	80.5	82.5	94	100
C4	100	100	100	99
C5	85.5	88.5	97	100
C6	95	100	99	99
C7	98.5	99	100	100
Overall	93.93	95.71	98.57	98.93

As could be seen from the table, the proposed wavelet based feature extraction method gives better results compared to the other wavelet based methods. However, the overall accuracy obtained considering the method proposed in [16] is very close to the accuracy obtained with the method proposed in this paper. That is due to the fact that [16] uses nine wavelet based features at each decomposition level, which for 6 levels of decomposition results in 63 features. It is a large number of features, compared with our feature vector consisted from 14 features. That means more calculations which result in increased computational cost.

The same procedure is made for signals accompanied with WGN. Obtained results for 20 dB, 30 dB, 40 dB and 50 dB are graphically presented in Figures 3, 4, 5 and 6, respectively.

According to the figures, the conclusion for each noise level is more or less the same as that for pure PQ signals. Hence, the feature extraction method proposed in this paper represents an efficient feature extraction method with high accuracy.
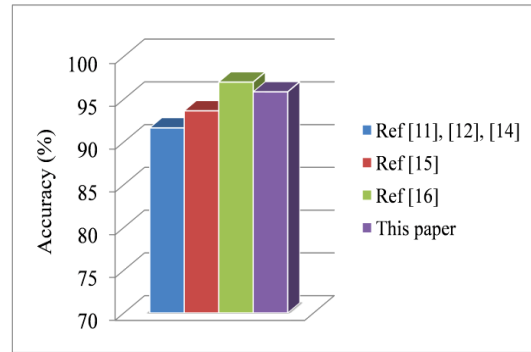

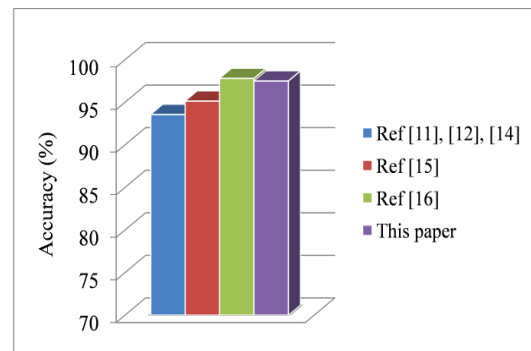**Fig. 3.** Comparison of the classification results for the proposed feature extraction method in case of 20 dB WGN

Fig. 4. Comparison of the classification results for the proposed feature extraction method in case of 30 dB WGN
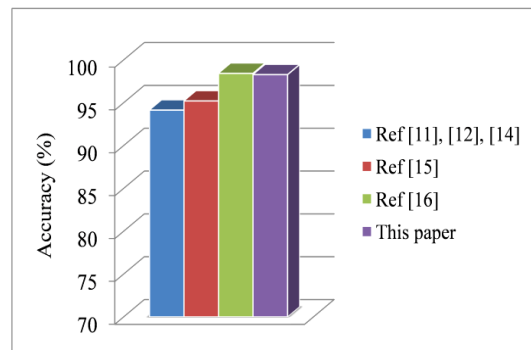

Fig. 5. Comparison of the classification results for the proposed feature extraction method in case of 40 dB WGN
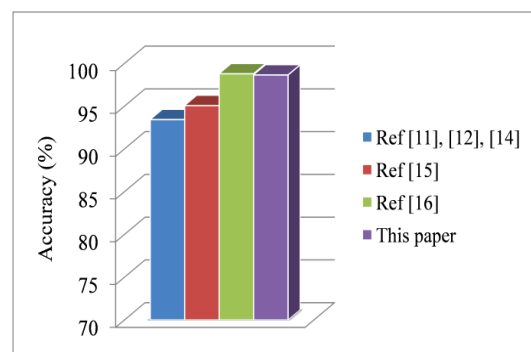

Fig. 6. Comparison of the classification results for the proposed feature extraction method in case of 50 dB WGN

The other testing using the first data group is about the performance of proposed feature extraction method together with the proposed classification method. The results obtained using this methods are comparatively presented in Table 5, together with some previously presented results in the referenced papers.

T a b l e 5

Comparison of the classification accuracy using 7 types of PQ disturbances

Class	Ref [11]	Ref [12]	Ref [13]	This paper
C1	100	100	100	100
C2	97	100	96.5	94.5
C3	76.5	87	88	100
C4	100	100	100	99
C5	90	80.5	85.5	100
C6	71.5	97	100	99
C7	98	100	100	100
Overall	**90.4**	**94.93**	**95.71**	**98.93**

From the given results, it is evident that the performance of our feature extraction and classification method exceeds the performance of the considered methods.

The second group of data was used for comparative testing of the classification accuracy of the proposed method in case of ten types of PQ disturbances. For that purpose, the computed values of the %CA (classification accuracy) and %MC (misclassification rate) for our method and for the methods proposed in [14] for pure PQ signals and PQ signals accompanied with 20 dB and 30 dB of WGN, are comparatively presented in Tables 6, 7 and 8, respectively.

T a b l e 6

Comparison of the classification accuracy using 10 types of pure PQ disturbances

	Class	Ref [15]		This paper
		SVM	DT	
MC (%)	V1	0	0	0.07
	V2	0	0	0.01
	V3	1.48	0	0.07
	V4	0	0	0.01
	V5	0	1.28	0
	V6	0.04	0	0.02
	V7	1.71	0.08	0.02
	V8	1.02	0.46	0.03
	V9	3.11	0	0.06
	V10	2.33	1.76	0
%CA	**Total**	**88.52**	**96.67**	**99.08**

T a b l e 7

Comparison of the classification accuracy using 10 types of PQ disturbances accompanied with 30dB WGN

	Class	Ref [15]		This paper
		SVM	DT	
MC (%)	V1	2.1	0	0.11
	V2	2.52	0.63	0.04
	V3	1.7	0	0.08
	V4	0.1	2.53	0.05
	V5	6.21	1.45	0
	V6	0	0	0.05
	V7	1.03	0.04	0.01
	V8	4.53	3.12	0.1
	V9	2.31	2.19	0.15
	V10	4.12	0	0.08
%CA	**Total**	**84.03**	**91.81**	**98.03**

T a b l e 8

Comparison of the classification accuracy using 10 types of PQ disturbances accompanied with 20dB WGN

	Class	Ref [15]		This paper
		SVM	DT	
MC (%)	V1	0	0	0.18
	V2	13.92	1.16	0.12
	V3	10.48	2.32	0.11
	V4	0	2.19	0.03
	V5	26.41	1.28	0.02
	V6	0.04	0	0.05
	V7	10.74	6.08	0.03
	V8	6.91	3.46	0.26
	V9	11.11	2.3	0.22
	V10	23.33	1.76	0.19
%CA	**Total**	**81.52**	**89.95**	**95.89**

In order to better perceive the difference between accuracies, the same results are graphically presented in Figure 7.

From the presented results it is obvious that the proposed feature extraction and classification method has lower misclassification value and higher classification accuracy.
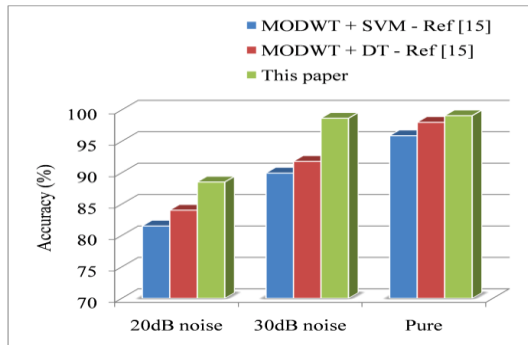
Fig. 7. Comparison of the classification accuracies obtained in case of 10 types of PQ disturbances

6. CONCLUSION

An efficient method for feature extraction and classification of PQ disturbances has been presented in this paper. The method is based on DWT, used in the feature extraction process, and a RF algorithm for classification. In order to test its accuracy, different experimental classifications were made, using seven and ten types of PQ disturbances. The obtained classification results were compared with some previously published results, obtained from different authors, using similar wavelet based feature extraction methods and different classification techniques. The comparisons have shown that the presented method has higher classification accuracy.

REFERENCES

[1] *IEEE Recommended Practice for Monitoring Electric Power Quality*, IEEE Std. 1159–1995, 1995.

[2] Styvaktakis, E., Bollen, M. H. J., Gu, I. Y. H.: Expert system for classification and analysis of power system events, *IEEE Trans. Power Del.*, vol. **17**, No. 2, pp. 423–428 (2002).

[3] Dash, P. K., Mishra, K. S., Salama, M. M. A.: Classification of power system disturbances using a fuzzy expert system and a Fourier linear combiner, *IEEE Trans. Power Del.*, vol. **15**, No. 2, pp. 472–477 (2000).

[4] Monedero, I., León, C., Ropero, J., García, A., Elena, J. C., Montaño, J. C.: Classification of electrical disturbances in real time using neural networks, *IEEE Trans. Power Del.*, vol. **18**, No. 2, pp. 406–1296 (2003).

[5] Janik, P., Lobos, T.: Automated classification of power-quality disturbances using SVM and RBF networks, *IEEE Trans. Power Del.*, vol. **21**, No. 3, pp. 1663–1669 (2006).

[6] Dehghani, H., Vahidi, B., Naghizadeh, R., Hosseinian, S.: Power quality disturbance classification using a statistical and wavelet-based hidden Markov model with dempster-shafer algorithm, *International Journal of Electrical Power & Energy Systems*, vol. **47**, pp. 368–377 (2013).

[7] Bollen, M.H.J. and Gu, I.Y.H.: *Signal Processing of Power Quality Disturbances*, Wiley – IEEE Press, 2006.

[8] Gaing, Z. L.: Wavelet-based neural network for power disturbance recognition and classification, *Power Delivery, IEEE Transaction*, vol. **19**, No. 4, pp. 1560–1568 (2004).

[9] Decanini, J. G., Tonelli-Neto, M. S., Malange, F. C. Minussi, C. R.: Detection and classification of voltage disturbances using a fuzzy-artmap-wavelet network, *Electric Power Systems Research*, vol. **81**, No. 12, pp. 2057–2065 (2011).

[10] Rată, G., Rată, M., Filote, C., Strugaru C.: Theoretical and experimental aspects concerning Fourier and wavelet analysis for deforming consumers in power network, *Elektronika ir Elektrotechnika (Electronics and Electrical Engineering)*, vol. **97**, No. 1, pp. 62–66, (2010).

[11] Abdel-Galil, T. K. A., Kamel, M., Youssef, A. M., El-Saadany, E. F. E., Salama, M. M. A.: Power quality disturbance classification using the inductive inference approach, *IEEE Trans. Power Delivery*, vol. **19**, No. 4, pp. 1812–1818 (2004).

[12] He, H., Starzyk, J. A.: A self-organizing learning array system for power quality classification based on wavelet transform, *IEEE Trans. Power Delivery*, vol. **21**, No. 1, pp. 286–295 (2006).

[13] Uyara, M., Yildirima, S., Gencoglub, M.T.: An effective wavelet-based feature extraction method for classification of power quality disturbance signals, *Journal of Elec. Power Syst. Res.*, vol. **78**, No. 10, pp. 1747–1755 (2008).

[14] Milchevski, A., Taskovski, D.: Classification of power quality disturbances using wavelet transform and SVM decision tree, *Proc. of IEEE EPQU 2011*, pp. 1–5, 2011.

[15] Upadhyaya, S., Mohanty, S.: Localization and Classification of Power Quality Disturbances using Maximal Overlap Discrete Wavelet Transform and Data Mining based Classifiers, *Elsevier, IFAC – Papers On Line*, vol. **49**, No. 1, pp. 437–442 (2016).

[16] Erişti, H., Yıldırım, Ö., Erişti, B., Demir, Y.: Optimal feature selection for classification of the power quality events using wavelet transform and least squares support vector machines, *Int. J. Electr. Power Energy Syst.*, vol. **49**, pp. 95–10 (2013).

[17] Breiman, L.: Random Forests, *Machine Learning*, vol. **45**, No. 1, pp. 5–32 (2001).

[18] Fernández-Delgado, M., Cernadas, M., Barro, S., Amorim, S.: Do we need hundreds of classifiers to solve real world classification problems?, *J. Mach. Learn. Res.*, vol. **15**, pp. 3133–3181, (2014).

[19] Vaidyanathan, P.: *Multirate Systems and Filter Banks*, Prentice-Hall, 1993.

[20] Strang, G., Nguyen, T.: *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Cambridge, MA, 1996.

[21] Santoso, S., Powers, E. J., Grady, W. M., Hofmann, P.: Power quality assessment via wavelet transform analysis, *IEEE Trans. Power Delivery*, vol. **11**, No. 2, pp. 924–930, (1996).

[22] Hamid, E. Y., Kawasaki, Z. I., Yoshida, H., Doi, H.: Wavelet analysis of voltage disturbances for power quality applications, *Trans. Inst. of Electrictrical Engineering of Japan*, vol. **122-B**, No. 2, pp. 323–330, (2002).

[23] Li, T., Ni, B., Wu, X., Gao, Q., Li, Q. and Sun, D.: On random hyper-class random forest for visual classification, *Neurocomputing*, vol. **172**, pp. 281–289 (2016).

[24] Huang, N., Lu, G., Cai, G., Xu, D., Xu, J., Li, F., Zhang, L.: Feature selection of power quality disturbance signals with an entropy-importance-based random forest, *Entropy*, vol. **18**, No. 2, pp. 1–21 (2016).