Expert Systems with Applications 38 (2011) 12491–12496

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A novel analytic method of power quality using extension genetic algorithm and wavelet transform

Meng-Hui Wang *, Yi-Feng Tseng

Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung, Taiwan

ARTICLE INFO

Keywords:
Power quality analysis
Wavelet transform
Extension theory

ABSTRACT

The power quality affects the power stability of power company and customers. In order to avoid economic losses caused by the power disturbances, it is necessary to monitor power parameters. This paper aimed at power quality analyses by wavelet transform and proposed a novel algorithm called extension genetic algorithm (EGA). The paper introduced the fundamental theory of wavelet transform, current applications and the theoretical framework of EGA. Then, it described the definition of power quality problems and the characteristics of power waves. Finally, this paper compared the analysis results of EGA and other methods. As the results of simulation, this paper mentioned of methods has a very high accuracy. It can also provide an application tool on power quality and data classification for future researchers.

1. Introduction

The power quality analysis is an important factor in the modern power systems. The electrical engineer must understand a certain statistical data and system when they analyze the electric power quality issues. Various kinds of electric power quality event used long-term monitoring of voltage to calculate the value rms. And then, we observing the values in a certain unit time of value change value to judgment the voltage were lower, rise or the electricity cut off (Chen, 2007). By the peak voltage and frequency changes, determine whether such problems as voltage flicker or harmonic power. In order to accurate analyze of the various electric power quality problems, usually measure many kinds of the electric power characteristics. Therefore, this paper utilized the wavelet transform to deal with the disturbance waveform of electric power quality and expect to reduce the number of characteristics of pick waveform to maintain the accuracy of the identification (Galli & Heydt, 1997).

In recent years, the research pointed out the wavelet transform had Multi-Resolution Analysis (MIRA) of characteristics and in accordance with different frequencies of analyzing degree to provide different frame width that the transient change will leave obvious signal.

Therefore, some scholars used wavelet transform to monitoring the location of the power system transient and classify different power system accident (Santoso, 1996). When the electric power quality perturbation signal continuous wavelet transform because different input signal, it will show the corresponding wavelet coefficient in the low frequency. If the wave of the signal has high frequency ingredient, it will also produce a bigger wavelet coefficient in the part of high frequency. Therefore, according to the different high and low frequency coefficient, it can be used as recognize of electric power quality perturbation wave.

The extension of matter-element theory established the failure of the matter-element model can save a lot of modeling space and the advantages of rapid classification. Some of diagnosis case has been successfully (Wang & Ho, 2005; Wang, 2002). This method needs to adjust the weight and set up the matter-element model by experience rule, which can achieve the highest accuracy rate.

Therefore, if we can reduce the extension method to the experience rule, this method will increase the universality and the practicality. In view of this, this paper puts forth the EGA and wavelet transform to extract the characteristic value and carry on a classification and identification to the electric power quality of disturbance wave. In this paper, the methods to propose its very innovation and the highest accuracy rate have 97%. So it has a very high practical value to provide a new reference for scholarly research.

2. Synopsis of wavelet theory

The wavelet transform utilize wavelet conversion of the zoom and translation two characteristic to sample time-shifting and scaling as show by Eq. (1). $W(j,k)$ is the discrete wavelet transform in $x(k)$, and $\psi(t)$ is a finite energy and quickly constringency of time-function (Wang & Wang, 2007)

* Corresponding author. Tel.: +886 4 23954505 2200; fax: +886 4 23924419.
E-mail address: wangmh@ncut.edu.tw (M.-H. Wang).

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.04.032

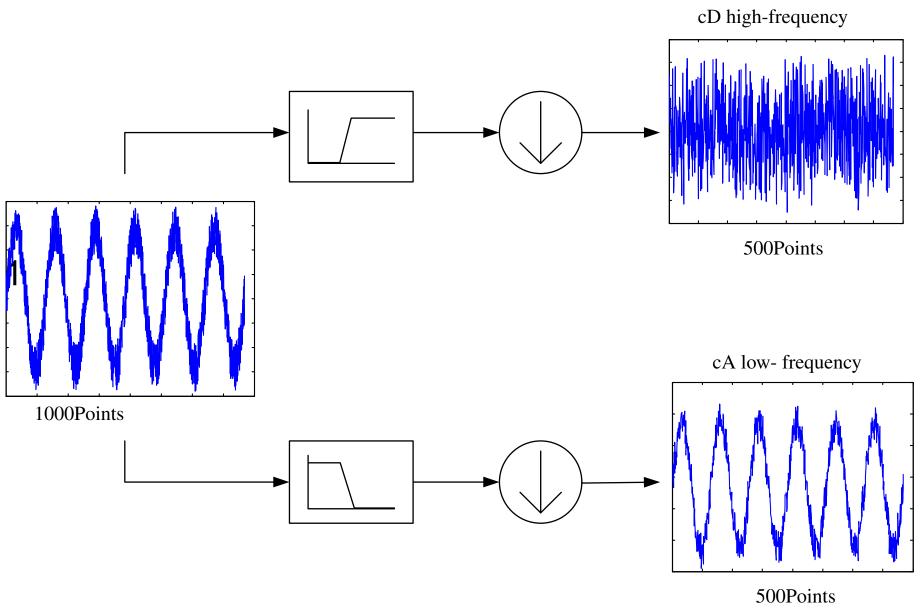

Fig. 1. The signal filters and reduces the sampling points.

Table 1
Three different sorts of mathematical sets.

Compared item	Cantor set	Fuzzy set	Extension set
Objects	Data variables	Linguistic variables	Contradictory problems
Model	Mathematics model	Fuzzy mathematics model	Matter-element model
Descriptive function	Transfer function	Membership function	Correlation function
Descriptive property	Precision	Ambiguity	Extension
Range of set	{0, 1}	[0, 1]	$[-\infty, \infty]$

$$W(j,k) = \sum_j \sum_k x(k) 2^{-j/2} \psi(2^{-j}n - k) \tag{1}$$

The angle of observation is picked from the characteristic value therefore the continuous wavelet functions can carry the wavelet function on all times and scales. However, the obtained scale is huge. If only a part of the scale is used, a great deal of current factor can be reduced without losing accuracy. The wavelet transform can also use a similar carrier signal. For most signals the low-frequency is usually an approximate original signal. The high-frequency signal carries the detailed changes or partial perturbation. A signal filter is based on the wavelet transform concept. The signal after filtering is divided into two parts: high-frequency and low-frequency. If the original signal is described as having 1000 points, there will be 1000 signals of 2 after the high, low-frequency filter. This will make the signal data become longer, so we usually reduce the data to a sampling point as shown by Fig. 1.

3. Summary of extension theory

The extension theory was first introduced in 1983 by a China scholar Cai, W. There are two main points in extension theory that are matter-element model and extension set (Cai, 1998). The matter-element model can describe the data that can analyze the quantitative change and the qualitative change. The extension set is built by matter-element model. It can solve contradictory problems which cannot solve by classical methods and fuzzy methods. The extension sets extent the fuzzy set from [0, 1] to $[-\infty, \infty]$ (Das, 2006). The distance of extension describes the value from the interval in the region. The interval is called classical field, and the region is called joint field. This value could calculate by correlation function that can describe the element to be positive field, negative field or zero boundary (Wang, Chung, & Sung, 2011). Table 1 shows three different sorts of mathematical sets.

3.1. Definition of matter-element

In the extension theory, the element is R, and N is the name of element. The characteristic is c, and v is characteristic of value. The matter-element in ET can be described as follows (Cai, 1999):

$$R = (N, c, v) \tag{2}$$

Assuming $R = (N, c, v)$ a multidimensional matter-element, $C = [c_1\ c_2 \cdots c_n]^T$ a characteristic vector and $V = [v_1\ v_2 \cdots v_n]^T$, a value vector of then a multidimensional matter-element is defined as

$$R = (N, C, V) = \begin{bmatrix} N, & c_1, & v_1 \\ & c_2, & v_2 \\ & \vdots & \vdots \\ & c_n, & v_n \end{bmatrix} \tag{3}$$

3.2. Definition of extension set

U is universe of discourse and u is a generic element of U. The u belongs $U(u \in U)$, then an extension set A in U is defined as a set of ordered pairs (4):

$$A = \{(u, y) | u \in U, y = K(u) \in (-\infty, \infty)\} \tag{4}$$

$y = K(u)$ is correlation function of extension set A. The extension set A and the universe of discourse U are defined as follows:

$$A = A^+ \cup A^0 \cup A^- \tag{5}$$

$$A^+ = \{(u, y) | u \in U, y = K(u) > 0\} \tag{6}$$

$$A^0 = \{(u, y) | u \in U, y = K(u) = 0\} \tag{7}$$

$$A^- = \{(u, y) | u \in U, y = K(u) < 0\} \tag{8}$$

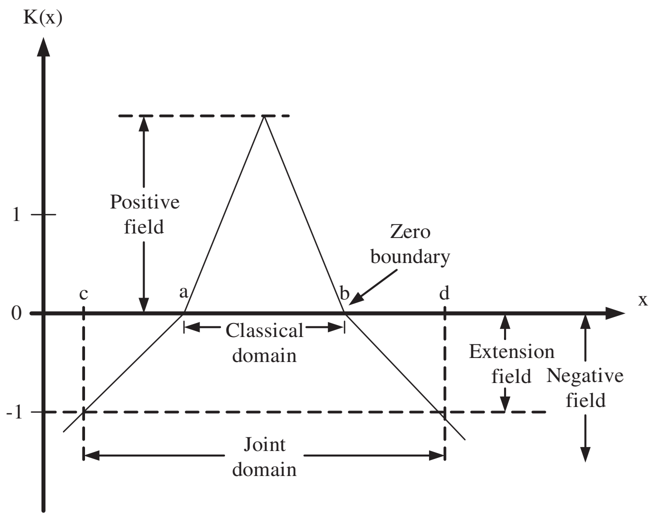

Fig. 2. The graphic relationship of extension sets.

A^+ is called positive field, and A^- is called negative field. A^0 is called zero boundary. Fig. 2 is the graphic relationship of extension.

4. The basic theory of genetic algorithm

The best-known evolutionary algorithm (EA) is the genetic algorithm (GA), which transposed the notion of evolution in nature to computers and imitates natural evolution and selection. Basically, they find solution to a problem by maintaining a population of possible solutions according to the "survival of the fittest" principle. The genetic algorithm constitutes a class of search algorithms especially suited to solving complex optimization problems. In addition to parameter optimization, the genetic algorithm is also suggested for solving problems in creative design, such as combining components in a novel creative way. In general, the major advantage of using the GA is that the optimal solution is obtained globally (Hwang & He, 2006). The genetic algorithm generally includes the following five parts:

(1) *Gene coding:* Combining all genes into a chromosome of sequence 0 and 1.
(2) *Fitness function:* It describes the capability of a certain individual gene to reproduce and is usually equal to the proportion of the individual's genes in all genes of the next generation.
(3) *Selection mechanism:* It is the intentional manipulation by chromosome of the fitness of individuals in a population to produce a desired evolutionary response.
(4) *Crossover*: A process in which chromosomes exchange genes through the breakage and reunion of two chromosomes.
(5) *Mutation*: A change in a gene resulting in new or rearranged hereditary determinants. Mutations are rare, random events in which the base sequence of the gene is changed.

5. The proposed extension genetic algorithm

In this paper, the proposed clustering method involves a combination of the extension theory and genetic algorithm. The extension theory provides a means for distance measurement in the classification process. The genetic algorithm has the ability to search for an optimal solution within a wide space. The EGA is a kind of supervised learning that finds the best classical domain and gets better accuracy without adjusting the weight (Wang, Tseng, Chen, & Chao, 2009).

This section will present a mathematical description of the EGA. The extension method can be found in out at the paper by (Wang, 2004), so it is not necessary to explain here. We define several variables before using the algorithm.

5.1. The training stage

The chromosomes propagate next generation of chromosomes to combine the matter-element models in the EGA. Setting *Patterns* = $\{p_1, p_2, \ldots, p_n\}$ with *i*th as follows: $p_i^j = \{c_1, c_2, \ldots, c_k\}$. In the patterns, *i* is the total number of genes, and *j* is the type of pattern. Using the proposed EGA can be simply described as follows:

Step 1. Set the epoch, the crossover rate C_r, the mutation rate m_u, the tolerance of error rate E_r, and the chromosome rate R_a.
Step 2. Find the gene of lower limit and upper limit value

$$v_a^j = \min(c_{kn}^j) \tag{9}$$

$$v_b^j = \min(c_{kn}^j) \tag{10}$$

$$v^j = \langle v_a^j, v_b^j \rangle \tag{11}$$

k is number of characteristic. v_a is the upper limit, and v_b is lower limit.

Step 3. Produce new gene of lower limit and upper limit value with chromosome rate. The chromosome rate is produced with random generator.

$$v_a^j - R_a \leqslant G_L^j \leqslant v_a^j + R_a \tag{12}$$

$$v_b^j - R_a \leqslant G_L^j \leqslant v_b^j + R_a \tag{13}$$

Step 4. The genes make up the chromosome.

$$chrom = \{G_L^{11}, G_L^{11}, G_L^{12}, G_L^{12}, \ldots, G_L^{jk}\} \tag{14}$$

The amount of gene in a chromosome is calculated by the function $2 * k * j$.

Step 5. Building the matter-element model from gene.

$$R_j = \begin{bmatrix} N, & c_1, & \langle G_L^1, G_U^1 \rangle \\ & c_2, & \langle G_L^2, G_U^2 \rangle \\ & \cdots & \cdots \\ & c_n, & \langle G_L^k, G_U^k \rangle \end{bmatrix} \quad j = 1, 2, \ldots, m \tag{15}$$

Step 6. Input the training of data that is the value of gene.

$$x^j = \{c_1, c_2, \ldots, c_k\} \tag{16}$$

Step 7. Calculate the correlation function.

$$z^k = (G_L^k + G_U^k) \tag{17}$$

$$K_{nk} = \sum_{i=1}^{n} \left[\frac{|x_{nk}^j - z_{jk}| - (G_U^{jk} - G_L^{jk})/2}{|(G_U^{jk} - G_L^{jk})/2|} + 1 \right] \tag{18}$$

Step 8. Normalizing the value of correlation function for the matter-element model to be between 1 and -1.
Step 9. Input the next training of data to repeat Step 6 to Step 8.
Step 10. Input the next matter-element model, and repeat Step 5 to Step 9.
Step 11. Calculate the fitness function.

$$Fitness = \frac{N_r}{N_a} \tag{19}$$

N_r is the right amounts, and N_a is the total mounts.

Step 12. The selection of the parental chromosomes put into the mating pool, and the genes implement crossover mechanism.

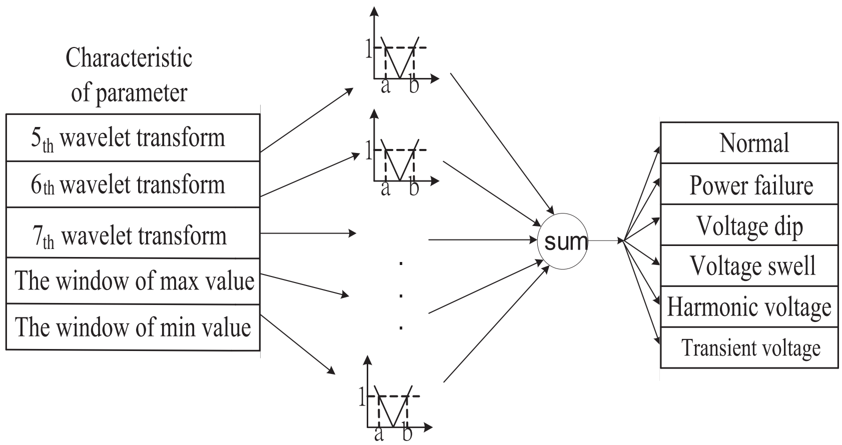

Fig. 3. The EGA of analysis pattern.

Table 2
The extension with experience rule result.

The type of parameter adjustment		1st	2nd	3rd	4th	5th
The ratio of training	Normal	34/40	33/40	35/40	35/40	34/40
	Power failure	37/40	33/40	34/40	30/40	37/40
	voltage swell	36/40	32/40	33/40	32/40	35/40
	Voltage dip	30/40	36/40	31/40	35/40	34/40
	Harmonic voltage	40/40	38/40	40/40	38/40	40/40
	Transient voltage	37/40	36/40	39/40	33/40	38/40
The ratio of accuracy		214/240	208/240	212/240	203/240	218/240
Accuracy rate (%)		89	87	88	85	91

Step 13. Let the next generation of chromosomes to replace the chromosomes, and implement mutation mechanism.

Step 14. Calculate the correct rate.

$$E_r = (1 - Fitness) \times 100\% \tag{20}$$

Step 15. Until the training is finished. If training process is not finished; otherwise go to Step 3.

5.2. The recognizing stage

Step 1. Build the matter-element model by optimization solution.

$$R_j = \begin{bmatrix} N, & c_1, & \langle G_L^1, G_U^1 \rangle \\ & c_2, & \langle G_L^2, G_U^2 \rangle \\ & \dots & \dots \\ & c_n, & \langle G_L^k, G_U^k \rangle \end{bmatrix} \quad j = 1, 2, \dots, m \tag{21}$$

Step 2. Input the data that is recognize.

$$x^j = \{c_1, c_2, \dots, c_k\} \tag{22}$$

Step 3. Calculate the correlation function.

$$z^k = (G_L^k + G_U^k)/2 \tag{23}$$

$$K_{nk} = \sum_{i=1}^{n} \left[\frac{|x_{nk}^j - z_{jk}| - (G_U^{jk} - G_L^{jk})/2}{|(G_U^{jk} - G_L^{jk})/2|} + 1 \right] \tag{24}$$

Step 4. Find min(K_{nk}). If the K_{nt} is bigger than k, then the data does not belong to any.

Step 5. Until recognizing is finished; otherwise go to Step 2.

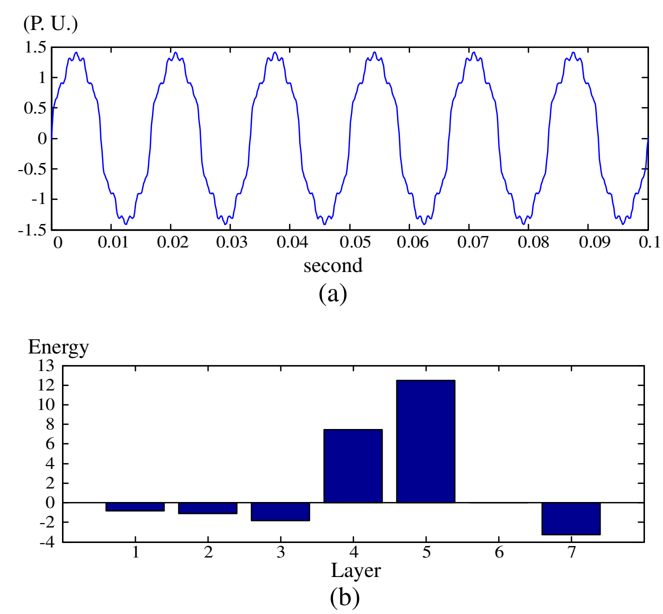

Fig. 4. The power harmonic wave and the energy layers with the fundamental frequency removed.

6. Result and analysis

This text compliant IEEE STD 1159-1995 (IEEE recommended practice for monitoring electric power quality) to build the information for simulating a power failure, voltage dip, voltage swell, power harmonic, and transient voltage for the power quality problem (Bhattacharyya, 2007; Dash, 2003; Kwang, 2008). The simulation cycle is set to six cycles, describing 128 points for a cycle. This paper presents a method that utilizes the wavelet transform to deal with the disturbance waveform in electric power quality. The EGA needs the parameter characteristics for system training including the max value and min value windows, 5th, 6th and 7th wavelet transform. Fig. 3 shows the EGA analysis pattern.

6.1. Testing the ability to training

There are numerous problems maintaining power quality. Power quality maintenance problems include vibration amplitude issues, random changes in frequency and timing. In the training stage 40 samples of every type of power quality problem are

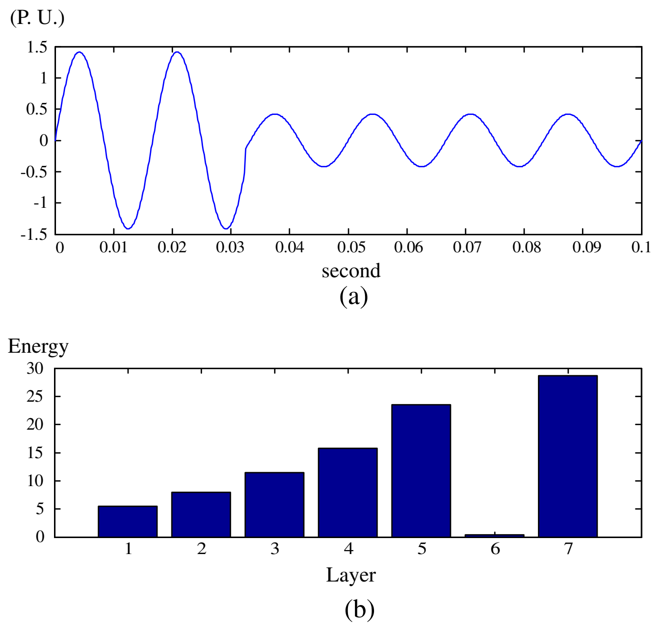

Fig. 5. The voltage dip wave and the energy layers with the fundamental frequency removed.

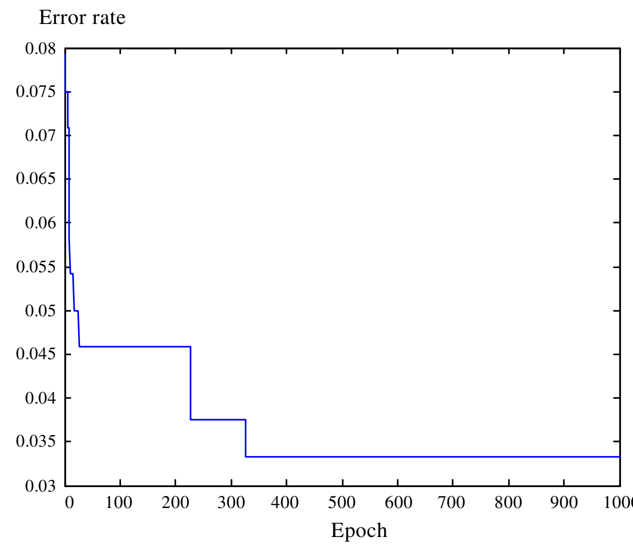

Fig. 6. The highest accuracy rate of convergence curve.

Table 3
The training times of accuracy rate.

training times (epoch 1000)	The ratio of accuracy	Accuracy rate (%)	Average accuracy rate (%)
1	230/240	96	96
2	228/240	95	
3	231/240	96	
4	229/240	95	
5	231/240	96	
6	229/240	95	
1	230/240	96	
2	228/240	95	
3	231/240	96	

Table 4
The matter-element model.

$$R_1 = \begin{cases} Har, c_1, \langle 0.000003, 39.846 \rangle \\ c_2, \langle -0.23328, 0.3924 \rangle \\ c_3, \langle -7.5126, -2.4913 \rangle \\ c_4, \langle 0.077703, 1.6189 \rangle \\ c_5, \langle 1.1198, 1.5632 \rangle \end{cases} \quad R_2 = \begin{cases} Intt, c_1, \langle 0.0000002, 37.667 \rangle \\ c_2, \langle -4.3416, 0.54634 \rangle \\ c_3, \langle 3.0461, 41.274 \rangle \\ c_4, \langle 0.000003, 1.343 \rangle \\ c_5 \langle 1.3119, 1.6361 \rangle \end{cases}$$

$$R_3 = \begin{cases} Sag, c_1, \langle 0.00144, 31.787 \rangle \\ c_2, \langle -4.6065, 0.7138 \rangle \\ c_3, \langle 1.3537, 40.614 \rangle \\ c_4, \langle 0.24499, 1.2107 \rangle \\ c_5, \langle 1.3524, 1.5198 \rangle \end{cases} \quad R_4 = \begin{cases} Swell, c_1, \langle -19.05, -1.1995 \rangle \\ c_2, \langle -1.6369, -0.0588 \rangle \\ c_3, \langle -27.01, -2.1059 \rangle \\ c_4, \langle 1.328, 1.7938 \rangle \\ c_5, \langle 0.66809, 4.3545 \rangle \end{cases}$$

$$R_5 = \begin{cases} Norm, c_1 \langle -2.3726, 2.9789 \rangle \\ c_2, \langle -0.25736, 0.1746 \rangle \\ c_3, \langle -2.2624, 4.1231 \rangle \\ c_4, \langle 1.1768, 1.5236 \rangle \\ c_5, \langle 1.3085, 1.7008 \rangle \end{cases} \quad R_6 = \begin{cases} Osci, c_1, \langle 0.004125, 77.752 \rangle \\ c_2, \langle -3.0257, 0.24934 \rangle \\ c_3, \langle -3.8565, 1.4996 \rangle \\ c_4, \langle 0.44524, 1.5328 \rangle \\ c_5, \langle 2.6703, 7.466 \rangle \end{cases}$$

randomly produced. Two hundred and forty pieces of training data are produced in total. The training is divided into two parts. The first part is the extension rule of thumb. The second part is the training EGA. Table 2 shows the extension with experience rule result. In the test stage a sample of every type of acceptable power quality is produced. The total training data is 240 pieces. Fig. 4(a) shows the power harmonic wave, and Fig. 4(b) shows the layers of power harmonic energy difference without the fundamental frequency. Fig. 5(a) shows the voltage dip wave, and Fig. 5(b) shows the layers of voltage dip energy difference without the fundamental frequency. When the voltage is normal, the energy layers are under the zero form 1st to 7th layers. Fig. 4(b) shows when the fault is harmonic voltage, the 4th layer and the 5th layer of energy is over zero. Fig. 5(b) shows the layers of energy are over zero. Therefore, the wavelet transform can be reasoned to explain the power efficiency problem in many cases.

The EGA parameter sets the error rate tolerance to 0.1, the crossover rate at 0.2, and the mutation rate at 0.005. The number of epochs is 1000. The EGA promotes non-progressive convergence. Two hundred and forty pieces of data produce 1000 epochs at a training iteration of 10. The training convergence result is recorded. Table 3 shows the training accuracy rate. The average accuracy rate is 96% after 10 iterations. The highest training accuracy rate is 97%. Fig. 6 shows the highest convergence curve accuracy rate for each epoch. The error rate is 0.033. Table 4 shows the matter-element model. This model is obtained after training.

6.2. Testing the ability to analysis

By the result, the matter-element model is built by the highest accuracy rate of chromosome in the testing stage. The data of testing are 240 pieces. In the text, there are 6 types of fault. Type 1 is meaning the problem of harmonic voltage. Type 2 is power failure. Type 3 is power dip. Type 4 is voltage swell. Type 5 is the voltage is normal. And type 6 is transient voltage (Dugan, 2002). Table 5 is the value of correlation function that is normalizing between -1 and 1. The value of correlation function is 1. That is belonged to which type. If the value of correlation function is close to -1. That means the data is not belonged to this type. At the result, it shows which type is. In Table 5, when the result is 1, the diagnosis is problem of harmonic voltage. And the result of value is conforming to each type in Table 5. Table 6 is comparing the different sorts of clustering with 240 pieces of testing data. The K-means clustering of accuracy rate is 85%. The fuzzy c-means clustering of accuracy rate is 70%. And the EGA of accuracy rate is 92%. By the testing, Table 6 is showing that the EGA of accuracy is batter then the extension, K-mean, and fuzzy c-means of accuracy rate. And the EGA has an advantage that is sloughing off the extension of experience rule.

Table 5
The value of correlation function.

The data of number	The value of correlation function type						Result
	Harmonic voltage	Power failure	Voltage dip	voltage swell	Normal	Transient voltage	
11	1	0.45	0.17	−0.05	−1	0.49	1
45	−0.21	1	0.96	0.07	−1	0.11	2
93	−0.88	0.89	1	−1	−0.66	0.14	3
130	−0.37	0.2	−0.56	1	−1	0.78	4
182	−0.48	−0.07	−1	0.58	1	−0.19	5
211	−0.38	−0.02	−1	0.75	−0.83	1	6

Table 6
The comparing the different sorts of clustering.

Compared methods	The training of accuracy rate (%)	The testing of accuracy rate (%)
K-means clustering	63	59
Fuzzy c-means clustering	64	61
The extension theory	91	85
The extension genetic algorithm	96	92

7. Conclusion

This paper supports a method which can advantage accuracy and conveniently with power quality analysis. It uses the wavelet transform to extract characteristic value and to build various types of electric power quality problems with EGA. In the matter-element model of space, the method of using notion of genetic algorithm's space in search is implemented to enhance the accuracy and to select a best matter-element. From the result, the simulate system showed the accuracy rate that is better than before. In this research, EGA is compared extension theory, K-means and c-means. This theory has proved a 96% accuracy rate with proper training. If it is not trained, the accuracy rate would be 92%. This fact has proved that the wavelet transform and EGA is an outstanding system for power quality analysis. In the future, the EGA to power quality analysis can be the foundation of further research that might lead to other perspective or other problems.

Acknowledgement

The author would like to acknowledge the financial support of the National Science Council of Taiwan, ROC through its Grant No. NSC-97-2221-E-167-020.

References

Bhattacharyya, S. (2007). Consequences of poor power quality – An overview. In *The 42nd international universities power engineering conference* (pp. 651–656).
Chen, L. J. (2007). Study of partial discharge measurement in power equipment using acoustic technique and wavelet transform. *IEEE Transactions on Power Delivery, 22*(3), 1575–1580.
Cai, W. (1998). The extension set and incompatibility problem. *Journal of Scientific Exploration, 1*, 81–93.
Cai, W. (1999). Extension theory and its application. *Journal of Scientific Exploration, 44*(7), 673–682.
Das, B. (2006). Fuzzy logic-based fault-type identification in unbalanced radial power distribution system. *IEEE Transactions on Power Delivery, 21*(1), 278–285.
Dash, P. K. (2003). Power quality analysis using S-transform. *IEEE Transactions on Power Delivery, 18*(2), 406–411.
Dugan, R. C. (2002). *Electrical power systems quality 2E*. McGraw-Hill, Inc.
Galli, A. W., & Heydt, G. T. (1997). Transient power quality problems analyzed using wavelets. *IEEE Transactions on Power Delivery, 12*(2), 908–915.
Wang, M. H., & Ho, C. Y. (2005). Application of extension theory to PD pattern recognition in high-voltage current transformers. *IEEE Transactions on Power Delivery, 20*(3), 1939–1946.
Kwang, T. C. (2008). Power Quality event characterization. In *The 4th IET conference on power electronics, machines and drives* (pp. 543–547).
Santoso, S. (1996). Power quality assessment via wavelet transform analysis. *IEEE Transactions on Power Delivery, 11*(2), 924–930.
Wang, M. H. (2002). A novel extension method for transformer fault diagnosis. *IEEE Transaction on Power Delivery, 22*(8), 63–64.
Wang, M. H., Chung, Y. K., & Sung, W. T. (2011). Using thermal image matter-element to design a circuit board fault diagnosis system. *Expert Systems with Applications, 38*(5), 6164–6169.
Wang, J. D., & Wang, C. G. (2007). Detection of power quality disturbance based on binary wavelet transform. *TENCON 2007–2007 IEEE region 10 conference* (pp. 1–3).
Wang, M. H., Tseng, Y. F., Chen, H. C., & Chao, K. H. (2009). A novel clustering algorithm based on the extension theory and genetic algorithm. *Expert Systems with Applications, 36*(4), 8269–8276.