ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 2, February 2013

Non-Stationary Power Quality Signals Classification using Fuzzy C-means Algorithm

M.Venkata Subbarao[1], N.Sayedu.Khasim [2],Jagadeesh Thati[3] and M.H.H.Sastry [4]

***Abstract*— This paper presents a new approach for classification of various non-stationary power signal waveforms using modified Gaussian windowing technique. In comparison to the earlier Gaussian window, the modified Gaussian window is found to provide excellent normalized frequency contours of the power signal disturbances suitable for accurate visual localization, detection and classification. Various non-stationary power signals are processed through S-Transform with modified Gaussian window to generate time-frequency contours for extracting optimal feature vectors for automatic pattern classification. The extracted features are clustered using Fuzzy C-means algorithm. From simulation results, it is shown that the proposed Fuzzy C-means algorithm reveal very encouraging results in terms of the quality of solution found, the average no of function evaluations and processing time required. The average classification accuracy of the disturbances is 94.37%**

Index Terms*— *Non-Stationary power quality (PQ) signals; Time-Frequency Analysis; S-Transform (ST); Fuzzy Clustering.

I. INTRODUCTION

Data clustering is a process of dividing a given data set into classes or clusters so that items belonging to same class are similar to each other as possible, and items in different classes are dissimilar enough to be differentiated. Depending on the nature of the data and the purpose for which clustering is being used , different measures of similarity can be used to place items into classes, where the similarity measures controls how the clusters are being formed.

In electrical power networks, the voltage and current signals exhibit fluctuations in amplitude, phase, and frequency due to the operation of solid-state devices that are prolifically used for power control. The sudden increase and decrease in voltage signal is known as swell and sag, respectively. Similarly, the introduction of the multiples of fundamental frequency of the signal in the signal itself is known as harmonic distortion. Apart from these steady state disturbances, transients are seen in power networks[1-6].

In addition to these, disturbances like multiple voltage notches, due to solid-state converter switching, are observed in the electric power networks. To classify the non-stationary disturbance signal patterns like voltage sag, voltage swell, transients, notches, harmonics, chirps, momentary interruptions etc, in the normal sinusoidal signal frequencies, windowing techniques play a vital role. Consequently solutions to electrical network disturbance problems involve continuous monitoring of electrical network voltage and current waveforms by passing them through a range of frequencies, called window, in order to localize the different disturbance patterns. This is done by expressing the amount of overlap the window function has as it is shifted over the power signal. In order to do that, short time discrete Fourier transform (STFT) is most often used. But for non-stationary signals, the STFT does not track the signal dynamics properly due to the limitations of a fixed window width chosen a priori. Thus, a different approach for power quality analysis using a modified wavelet transform, known as S–transform and the analysis of the several power quality problems using an S–transform is incorporated. Because power signals are usually non–stationary, this on-line detection process implements the accurate classification of power disturbances, which is of paramount importance in signal processing.

Manuscript received Feb, 2013.

M.Venkata Subbarao, *Assistant Professor, Department of ECE, Tirumala Engineering College, Guntur, A.P., India.*

N.Sayedu Khasim, *Assistant Professor, Department of ECE, Tirumala Engineering College, Guntur, A.P., India.*

Jagadeesh Thati, *Sr. Assistant Professor, Department of ECE, Tirumala Engineering College, Guntur, A.P., India.*

M.H.H.Sastry, *Associate Professor, Head of the Department, Department of ECE, Tirumala Engineering College, Guntur, A.P., India.*

II. FEATURE EXTRACTIOIN

In this work we have used two types of features i.e. time-domain and frequency-domain features. Out of the seven features used, spectral entropy is a frequency-domain feature and the rest, which includes energy, mean, standard deviation, variance, autocorrelation and maximum deviation, are time-domain features. From these features 7 statistics $f1,\ldots,f7$ are calculated which are fed to the network for classification task. All the time domain feature statistics calculated are normalized between the ranges zero to one. This is done by dividing the individual features obtained for all the training patterns or (disturbance signals) with the maximum value of the respective feature. This ensures better stability of the network weights during training. The detailed description of the feature statistics used is provided below[1].

Energy values for different types of power system disturbance signals are different. Hence short time energy is an effective feature in classifying power system disturbances. The short time energy of a infinite signal sequence is given by:

$$E_j = \sum_{n=0}^{N-1} x_j(n)^2 \quad (1)$$

Mean of a signal is the property that tracks the mean values in a sequence of inputs over a period of time. It computes the arithmetic mean of all data in the series. The mean is calculated as:

$$\mu_j = \frac{1}{N}\sum_{n=0}^{N-1} x_j(n) \quad (2)$$

All Rights Reserved © 2013 IJARCET

Standard deviation of a signal is defined as the square root of the variance. Hence standard deviation of a probability distribution, random variable, or population or multi set of values is a measure of the spread of its values. Power system disturbances have different values of spread, which qualifies standard deviation as a good feature. The standard deviation is calculated as,

$$\sigma_j = \sqrt{\left(\frac{1}{N-1}\sum_{i=0}^{N-1}(x_j(n)-\mu_j)^2\right)} \quad (3)$$

Variance is the measure of the spread. When we are dealing with a *sample* (that is, a subset of the complete population), we cannot ofcourse compute the mean and variance exactly, but rather estimate them. The variance is calculated as,

$$r_j xx = \sum_{n=0}^{N-1} x_j(n) x_j(n-1) \quad (4)$$

Autocorrelation is the degree of matching between two signals and thus to extract some information that depends to a large extent on the application. It is computed by the sum of the cross products between the data set and its shifted versions. The level of similarity for steady state disturbances like voltage sag, voltage swell etc. is more than that of short time disturbances like oscillatory transients and flicker. So the autocorrelation sequence is used as a distinguishing feature in this work. Autocorrelation closely resembles convolution and is given by,

$$r_j xx = \sum_{n=0}^{N-1} x_j(n) x_j(n-1) \quad (5)$$

To obtain the statistical feature, standard deviation of the autocorrelation sequence is used.

Maximum deviation gives a measure of the extent of variation of a signal from its mean. The mean value of uniform signals like steady state disturbances is close to zero. Hence maximum deviation provides an indication about the level of the disturbance occurred. Maximum deviation is particularly effective in identifying amplitude disturbances like voltage sag, voltage swell, sag with harmonics and swell with harmonics. It is also effective in separating spikes from oscillatory transients.

$$\Delta j = \max(x_j) - \mu_j \quad (6)$$

III. Data Clustering

Clustering is a process of classification of objects into different groups or more precisely the partitioning of a data set into midgets: so that the data in each subset share some common trait often proximity according to some standard defined distance measure. Data clustering is a common technique for statistical data analysis which is used in many fields such as machine learning, data mining, pattern recognition, bioinformatics and image processing.

We have taken the help of data clustering methodology in order to distinguish patterns in power signal features and hence leading to their efficient classification.

An important step in any clustering technique is to select a distance measure, which will determine how effectively the two elements bear resemblance to each other. This will influence the shape of the clusters, as some elements close to one another according to one distance and may be farther away according to another.

For example in a two dimensional space the distance between the point (x=1,y=0) and the origin (x=0,y=0) is always 1 according to usual norms, but the distance between the point (x=1,y=1) and the origin can be 2, or 1 if you take the 1-norm, 2-norm or infinity norm respectively.

***Common Distance Functions*:**

- The Euclidean distance(also called the distance as the crow flies or 2-norm distance). A review of cluster analysis in health psychology research found that the most common distance measure used in research area is the Squared Euclidean distance.
- The Manhattan distance (also called the taxicab norm or 1-norm).
- The Infinity Norm
- The Mahalanobis corrects data for different scales and correlations in the variables.

A. Fuzzy Clustering:

Data clustering is a process of dividing a given data set into classes or clusters so that items belonging to same class are similar to each other as possible, and items in different classes are dissimilar enough to be differentiated. Depending on the nature of the data and the purpose for which clustering is being used , different measures of similarity can be used to place items into classes, where the similarity measures controls how the clusters are being formed.

In hard clustering data is divided into distinct clusters where each data element belongs exactly to one cluster as in K-Means. In fuzzy clustering data elements can belong to more than one cluster and associated with each element is a set of membership levels, and then use them to assign data elements to one or more clusters.

The K-Means algorithm is an algorithm used to cluster objects based on attributes into k partitions. It is a variant of the expectation maximization algorithm in which the goal is to determine the k means of data generated from Gaussian distributions. It assumes that the Object attributes form a vector space. The objective it tries to achieve is to minimize total intra- cluster variance or the squared error function.

$$V = \sum_{i=1}^{k} \sum_{x_j \in S_i} |x_j - \mu_i|\,2 \quad (7)$$

Where there are k clusters S_i, i=1, 2,3……..k and μ_i is the centroid or mean points of all these points for $x_j \in Si$.

IV. Fuzzy C-Means Algorithm

Fuzzy C-means is a method of clustering which allows one piece of data to belong to two or more clusters. This method developed by Dunn in 1971 and further improved by Bezdek in 1981 serves as a wonderful tool in pattern classification problems. It's based on minimization of the objective function.

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}\,||x_i - c_j||^2 \qquad \text{For } 1 \le m < \infty$$

Where m is any real number greater than 1 , U_{ij} Is the degree of membership of X_i in the cluster j , X_i is the ith of d-dimensional measured data , C_j is the d-dimensional centre of the cluster , and ||*|| is any norm expressing the similarity between any measured data and the center .Fuzzy partitioning

All Rights Reserved © 2013 IJARCET

is carried throughout an iterative optimization of the objective function made above with an update of membership function U_{ij} and the cluster C_j centers by :

$$U_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i-c_j\|}{\|x_i-c_k\|}\right)} 2/m - 1$$

$$C_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

This iteration will stop when the value of $\max_{ij}\{|U^{k+1}_{jj}-U^{k}_{jj}|\}<\delta$ where δ is a termination criterion between 0 and 1 where k are the iteration steps. This procedure converges to a local minimum of saddle point of J_m.

1) swell with flicker
2) sag with flicker
3) swell with harmonic
4) flicker with harmonic
5) momentary with flicker
6) momentary with harmonic
7) momentary with transient
8) sag with harmonic

These eight disturbances are classified according to the following fuzzy tree.

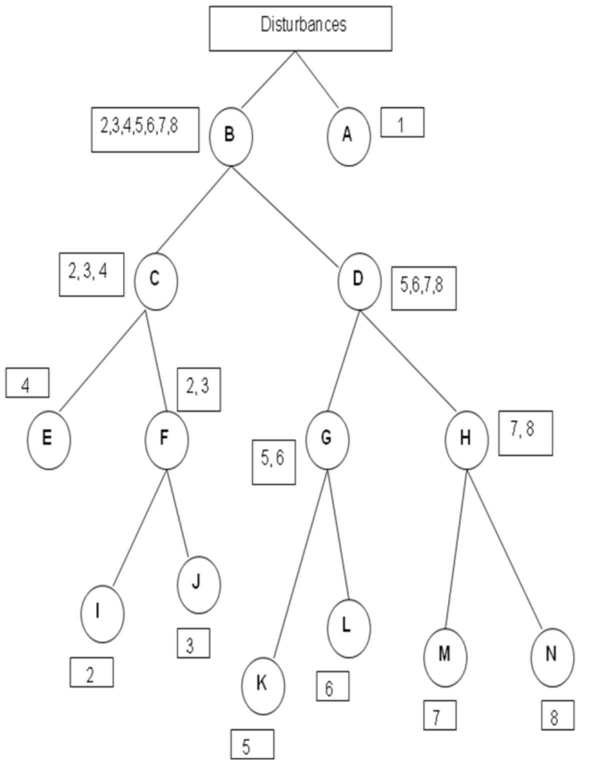

Figure 1. Fuzzy C-means Tree

V. Simulation Results

For clustering, the desired feature vectors (input space) will be divided into two clusters namely A and B. Further cluster B is sub-clustered into clusters C and D. Cluster C is then sub-clustered into cluster E and F and cluster D represents the remaining power signal disturbances. This process is continued for all other power signal disturbances until we get 8 clusters each representing one disturbance. The order and clustering process of the power signal disturbances is same for genetic based reformulated Fuzzy C-means tree.

The non-stationary power signal pattern recognition using reformulated Fuzzy C-means algorithm is shown where the classification of entire dataset into two clusters are depicted by star and triangle data points as shown in Figure 2. Similarly the remaining disturbances are sorted into different clusters by dividing each cluster into two clusters as shown in the figures below.

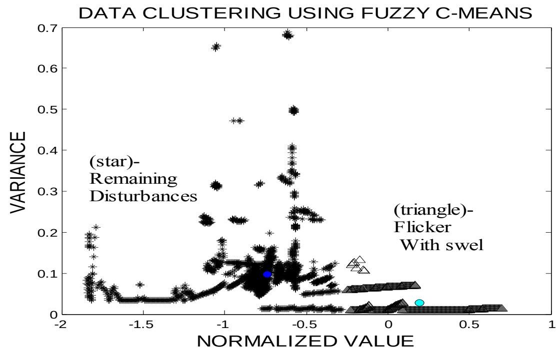

Figure 2. Classification of whole data into cluster A and B

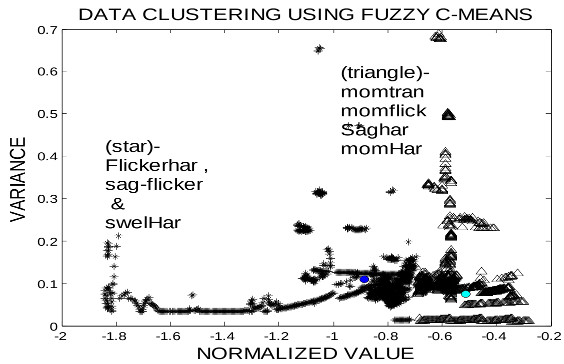

Figure 3. Classification of cluster B into C and D

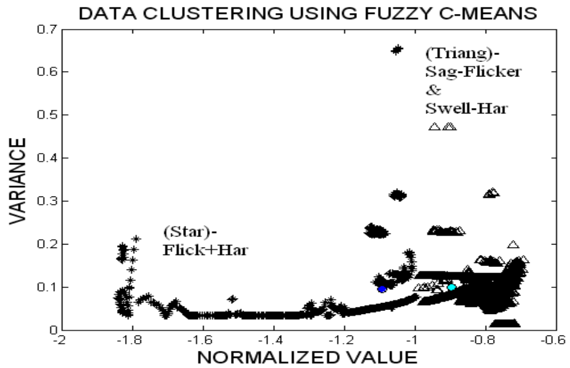

Figure 4. Classification of cluster C into E and F

All Rights Reserved © 2013 IJARCET

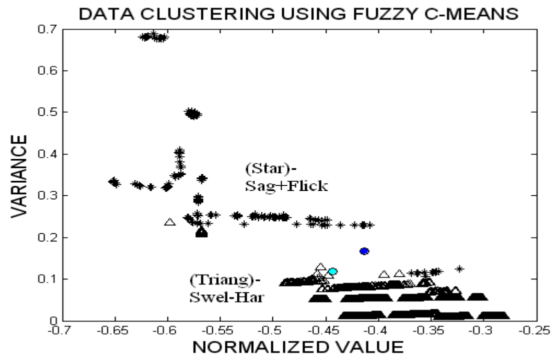

Figure 5. Classification of cluster F into I and J

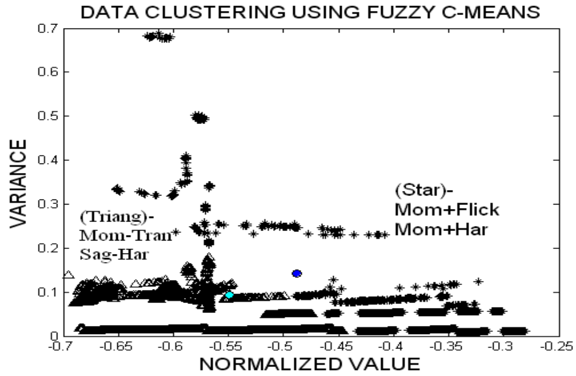

Figure 6. Classification of cluster D into G and H

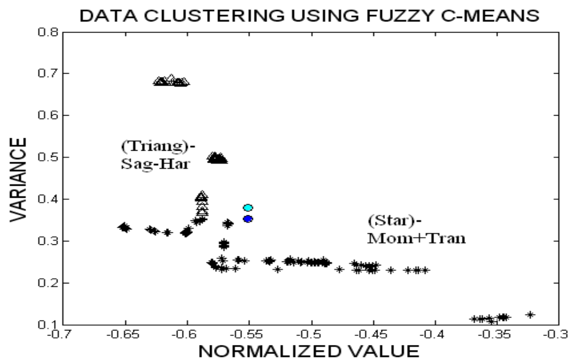

Figure 7. Classification of cluster G into K and L

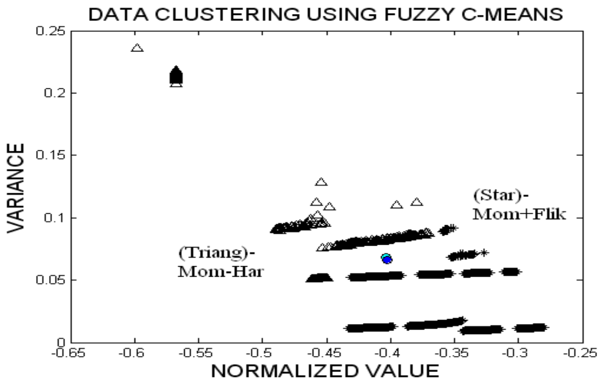

Figure 8. Classification of cluster H into M and N

These are the simulation results by using traditional Fuzzy C-means clustering. The accuracy values using is tabulated as shown in Table 1.

TABLE 1. ACCURACY IN PERCENTAGE USING FUZZY C-MEANS CLUSTERING

S.No.	DISTURBANCE	ACCURACY (C-Mean)
1	Swell with Flicker	92
2	Sag with Flicker	94
3	Swell with Harmonic	93
4	Flicker with Harmonic	94
5	Momentary & Flicker	92
6	Momentary&Harmonic	93
7	Momentary &Transient	94
8	Sag with Harmonics	96

VI. Conclusion

This work has proposed a new approach for classification of power quality disturbances in a power distribution system called Fuzzy C-Means algorithm. Extracted relevant features from generated normalized time-frequency contours of different power signal disturbances for pattern classification. In future this method can be widely applied to image processing, Radar signal classification and seismic signal processing. Automatic classification can be done by extracting feature vectors from the ST frequency contours and finally passing those pertinent feature vectors through an intelligent classifier for pattern classification.

REFERENCES

[1] Birendra Biswal, P.K.Dash, B.K.Panigrahi, ”Power quality disturbance classification using Fuzzy C means Algorithm and Adaptive particle Swarm Optimization” IEEE Trans, Ind.Electron., Vol 56,No.1 Jan 2009.

[2] A. M. Gaouda, S. H. Kanoun, M. M. A. Salama, and A. Y. Chikhani, “Wavelet base signal processing for disturbance classification and measurement,” *Proc. Inst. Electr. Eng.—Generation Transmission Distrib.* vol. 149, no. 3, pp. 310–318, May 2002.

[3] P. K. Dash, M. Nayak, M. R. Senapati, and I. W. C. Lee, “Mining for similarities in time series data using wavelet-based feature vectors and neural networks,” *Eng. Appl. Artif. Intell.*, vol. 20, no. 2, pp. 185–201, Mar. 2007.

[4] K.Deb, A. Anand and D. Joshi (2002). A Computationally Efficient Evolutionary Algorithm for Real-Parameter Optimization, Evolutionary computation, 10(4), pp. 371 – 395.

[5] S.R.Shaw, C.R.Laughman, S.B.Leeb, and R.F.Lepard, “A power quality prediction system”, IEEE Transaction on Industrial Electronics, vol.47, No.3, Jun 2000, pp.511-517.

[6] R.G. Stockwell, L.Mansinha, and R.P. Lowe, Localization of the complex Spectrum: The S-Transform”, IEEE Transaction on Signal Processing, vol.44, no.4, 1996, pp.998-1001. [5]

All Rights Reserved © 2013 IJARCET

M. Venkata Subbarao received his B.Tech degree from JNTU, Hyderabad, India in 2008 and M.Tech from JNTU Kakinada, India in 2011. Presently he is working as a Assistant Professor, in Department of Electronics and Communication Engineering,Tirumala Engineering College, Guntur, A.P, India.He is an Associate Member in IETE. So far he published more than 15 research papers in various international journals and conferences. His current areas of research interests include Signal Processing, Communication systems and Pattern Classification.

N.Sayedu khasim received his M.Tech degree from ANU, A,P, India in 2010. Presently he is working as a Assistant professor, in Department of Electronics and Communication Engineering, Tirumala Engineering College, Guntur, A.P, India. He is an Associate Member in IETE and Member in ISTE. So far he published more than 5 research papers in various international journals and conferences. His current areas of research interests include Communication and Radar systems, Digital signal processing.

T.Jagadeesh received his M.Tech degree from JNTUH, A,P, India in 2009 and M.S. degree from BTH, Sweedan in 2010. Presently he is working as a Sr. Assistant professor, in Department of Electronics and Communication Engineering, Tirumala Engineering College, Guntur, A.P, India. He is an Associate Member in IETE. So far he published more than 20 research papers in various international journals and conferences. His current areas of research interests include Signal processing, Image Processing and Nurual Networks.

M.H.H.Sastry received his M.Tech degree from JNTUK, A,P, India in 2006. Presently he is working as a Associate professor, in Department of Electronics and Communication Engineering, Tirumala Engineering College, Guntur, A.P, India. He is an Associate Member in IETE . So far he published more than 5 research papers in various international journals and conferences.His current areas of research interests include Communication Systems and Signal Processing.

.

All Rights Reserved © 2013 IJARCET