CHAPTER 8

AUTOMATIC DISTURBANCE RECOGNITION USING
HILBERT-HUANG TRANSFORM

8.1 INTRODUCTION

In the previous chapter, feature extraction based on Hilbert Transform has been discussed. In the Hilbert Transform based approach, the features are extracted by considering only the top envelope of the signal. But the bottom envelope of the disturbance signal also contains useful information. The performance of the ANN can be improved, if both the envelopes are taken into consideration for extracting the features. In this chapter, Hilbert-Huang Transform (HHT) is proposed in which both the envelopes are considered for extracting the features of the disturbance waveforms.

HHT is based on the instantaneous frequencies resulting from Intrinsic Mode Functions of the signal being analyzed (Huang 2000). These functions are obtained from both the top and bottom envelopes of the disturbance signal. Further, RBF network based on Fuzzy-C-Means (FCM) clustering algorithm is employed for the automatic disturbance recognition. Computer simulation is carried out to evaluate the performance of HHT based RBF network for power quality disturbance recognition. Simulation results demonstrate the effectiveness of HHT based RBF network for power quality disturbance recognition.
8.2 HILBERT-HUANG TRANSFORM

The Hilbert-Huang Transform (HHT) is an adaptive data analysis method designed for analyzing non-stationary signals. In HHT, the signal is decomposed into a finite small number of components, called Intrinsic Mode Functions (IMF). This process of decomposition is called Empirical Mode Decomposition (EMD). An IMF is defined as any function having the same number of zero-crossing and extrema and also having symmetric envelopes defined by the local maxima and minima respectively. This decomposition method operating in time domain is adaptive and therefore highly efficient. Since the decomposition is based on the characteristics of the time scale data, it is suitable for the analysis of non-stationary signal.

8.2.1 Empirical Mode Decomposition (EMD)

The EMD decomposes the signal in terms of IMFs, each of which is a mono-component function (Huang et al 2003). Then the Hilbert Transform is applied to calculate the instantaneous frequencies of the original signal. The procedure for extracting the IMFs from a signal \( x(t) \) is given below:

1. Identify the local extrema of the signal. A signal has at least two extrema (maximum and minimum). The extrema are found by determining the change of sign of the derivative of the signal \( x(t) \).

2. Two smooth splines are constructed which connects all the maxima and minima of \( x(t) \) to obtain its upper and lower envelope. Connect all the local extrema through curve fitting to...
generate the upper and lower envelopes. Cubic spline line is employed as a curve fitting function. The upper and the lower envelopes should cover all the data between them.

Let \( m \) be the mean of the two envelopes. The difference between the input signal \( x(t) \) and \( m \) is the first component, \( h \)

\[
x(t) - m = h
\]  
(8.1)

This is the first round of shifting. In this round, \( h \) is treated as proto-IMF. In the next round, it is treated as the signal, then

\[
h - m = h_1
\]  
(8.2)

After repeated shifting, upto \( k \) times, \( h_k \) is obtained as

\[
h_{k-1} - m = h_k
\]  
(8.3)

The function \( h_k \) is defined as the first IMF component and expressed as

\[c = h_k
\]  
(8.4)

The stopping criteria of the decomposition process is determined by the Cauchy type of convergence test, given by the equation

\[
SD_k = \sum_{t=0}^{T} \left| h_{k-1}(t) - h_k(t) \right|^2 / \sum_{t=0}^{T} h_{k-1}^2(t)
\]  
(8.5)
The shifting process is stopped, when $SD_k$ becomes smaller than a pre
determined value. Once the shifting process is stopped, the first IMF $c_i$
can be obtained, which contains the finest scale or the shortest period
component of the signal.

After separating $c_i$ from the original signal $x(t)$, the residue of the signal
is obtained

$$x(t) - c_i = r_i \quad (8.6)$$

This process can be repeated for all $r_j$ and the result is

$$r_1 - c_2 = r_2 \quad (8.7)$$

$$r_{n-1} - c_n = r_n \quad (8.8)$$

The shifting process stops finally when the residue $r_n$ becomes a
monotonic function, from which no more IMFs can be extracted. By
summing up the equations it follows that

$$x(t) = \sum_{j=1}^{n} c_j + r_n \quad (8.9)$$

Thus the decomposition of a signal in $n$-empirical modes is achieved.
The components of the EMD are physically meaningful, as the
characteristic scales are defined by the physical data. The instantaneous
frequency can be computed by finding the Hilbert Transform of the IMF
components.
After performing the Hilbert Transform to each IMF component, the original data can be expressed as the real part $\text{RP}$, in the following form:

$$ x(t) = \text{RP}\left\{ \sum_{j=1}^{n} a_j(t)e^{iw_j(t)t} \right\} $$

(8.10)

The equation (8.10) gives both amplitude and frequency of each component as function of time. The same signal $x(t)$ is represented in Fourier series as

$$ x(t) = \text{RP}\left\{ \sum_{j=1}^{\infty} a_j e^{i\omega_j t} \right\} $$

(8.11)

The time-frequency representation of the amplitude is designated as the Hilbert amplitude spectrum or Hilbert spectrum $H(w,t)$.

The IMF components for various power quality disturbances are shown in figure 8.1 to 8.6. The disturbance signals are generated using the same method adopted in section 3.4. The IMF components of these disturbances are found out using EMD technique. The IMF components for voltage sag are shown in figure 8.1.
8.1 Voltage sag and its IMF components

As shown in figure 8.1, the voltage sag signal is decomposed into four IMF components c1, c2, c3 and c4. The maximum amplitude of the IMF decreases gradually from c1 to c4. Each IMF component shows variations in magnitude and its value depends on the amplitude of both top and bottom envelopes.

Figure 8.2 shows the voltage swell and its IMF components

Figure 8.2 Voltage swell and its IMFs
As shown in figure 8.2, the voltage swell signal is decomposed into six IMF components. The total number of IMF varies depending on the maximum and minimum amplitude of the signal and the duration of the disturbance present in the signal.

Figure 8.3 shows the transient signal and its IMF components.

![Voltage transient and IMF Components](image)

**Figure 8.3 Voltage transient and its IMF components**

As shown in figure 8.3, the voltage transient is decomposed into four IMF components. The maximum amplitude of the IMF depends on the amplitude of the transient present in the signal.
Figure 8.4 Harmonics and its IMF components

The harmonic signal is decomposed into six IMFs as shown in figure 8.4. The first four IMFs show amplitude variations in the entire duration, whereas the IMFs c5 and c6 show slight variation only. Figure 8.5 shows flicker and its IMF components.

Figure 8.5 Flicker and its IMF components
The flicker signal contains three IMFs as shown in figure 8.5.

8.3 HILBERT SPECTRUM BASED POWER QUALITY DISTURBANCE ANALYSIS

Hilbert spectrum is a statistical tool for analyzing the non-stationary signals. The spectrum itself is decomposed into component sources (Ruqiang 2005).

The Hilbert spectrum is computed by performing the following steps.

1. Preprocessing the disturbance signal into Intrinsic Mode Functions (IMF) using Empirical Mode Decomposition (EMD) approach as explained in section 8.2.1.

2. Apply Hilbert Transform to the IMF components to obtain the instantaneous frequency spectrum of each of the components.

The energy-frequency-time representation of a signal is known as Hilbert spectrum.

The Hilbert spectrum for various power quality disturbances are shown in figures 8.6 to 8.11.
Figure 8.6 Voltage sag and its Hilbert spectrum

Figure 8.6 shows the voltage sag and its Hilbert spectrum. The spectrum shows slight variations in frequency during the period of the disturbance.

Figure 8.7 shows the Hilbert spectrum of the voltage swell signal.

Figure 8.7 Voltage swell and its Hilbert spectrum
The Hilbert spectrum of the voltage swell signal shows increase in frequency as shown in the figure.

Figure 8.8 shows the Hilbert spectrum of the transient signal.

![Voltage transient and its Hilbert spectrum](image)

**Figure 8.8 Voltage transient and its Hilbert spectrum**

The frequency content of the transient signal is high compared to the frequency content of the voltage sag and swell signals.

Figure 8.9 shows the Hilbert spectrum of the harmonic signal.
Figure 8.9 Hilbert spectrum of the harmonic signal

As shown in figure 8.9, the frequency content of the harmonic signal varies for the entire duration of the signal.

Figure 8.10 shows the Hilbert spectrum of the flicker signal.

Figure 8.10 Voltage flicker and its Hilbert spectrum
The amplitude of the Hilbert spectrum is very less as shown in figure 8.10 and the frequency varies within 100 Hz. Thus from the Hilbert spectrum of the disturbance signals, it is inferred that the shape of the spectrum as well as the maximum amplitude and power of the Hilbert spectrum depends on the type of power quality disturbance present in the signal.

8.4 FEATURE EXTRACTION USING HILBERT-HUANG TRANSFORM

The features of the disturbance signals are extracted by finding the energy of the IMFs which are derived from each of the disturbance waveforms. Let $c_1$, $c_2$, $c_3$ be the first three IMF components and $E_1$, $E_2$ and $E_3$ be their corresponding energies. Energy of the IMF is calculated using the following equations.

$$E_1 = ||c_1||^2$$  \hspace{1cm} (8.12)

$$E_2 = ||c_2||^2$$  \hspace{1cm} (8.13)

$$E_3 = ||c_3||^2$$  \hspace{1cm} (8.14)

where $||\cdot||$ represents norm.

The energy of the IMF for some of the power quality disturbances are given in Table 8.1
Table 8.1 Energy of IMF components for power quality disturbances

<table>
<thead>
<tr>
<th>Disturbance type</th>
<th>Energy of IMF</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_1$</td>
<td>$E_2$</td>
</tr>
<tr>
<td>Voltage sag</td>
<td>0.424</td>
<td>0.698</td>
</tr>
<tr>
<td>Voltage swell</td>
<td>2.412</td>
<td>4.012</td>
</tr>
<tr>
<td>Transient</td>
<td>7.521</td>
<td>8.012</td>
</tr>
<tr>
<td>Harmonics</td>
<td>5.2219</td>
<td>7.364</td>
</tr>
<tr>
<td>Flicker</td>
<td>4.327</td>
<td>5.304</td>
</tr>
</tbody>
</table>

The energy $E_1$, $E_2$ and $E_3$ of the IMF component varies for different types of disturbances as shown in Table 8.1. These features are given as input to the RBF network for further classification. The classification results are given in Table 8.2.

8.5 DEVELOPMENT OF RBF NETWORK USING FUZZY C-MEANS ALGORITHM

In the previous chapters, the RBFNN was developed using k-means algorithm. Because of the crisp nature of the k-means algorithm, there exists difficulty in assigning a cluster membership to data points. Depending on the minimum distance, a data point always
becomes a member of one of the clusters. Further, the k-means algorithm is not capable of dealing with overlapping clusters and outlier points, since it has to include a data point into one of the existing clusters. Because of this, even extreme outlier points will be included into a cluster based on the minimum distance.

In this chapter, Fuzzy C-Means algorithm (Yegnanarayana 1999) is proposed for determining the unit cluster centers of the RBF network. The fuzzy clustering provides a flexible and robust method for handling the data with vagueness and uncertainty. In fuzzy clustering, each data point will have an associated degree of membership for each cluster. This membership value has the range zero to one and indicates the strength of its association in that cluster.

The Fuzzy C-Means algorithm consists of two processes: the calculation of cluster centers and the assignment of points to these centers using a form of Euclidian distance. This process is repeated until the cluster centers stabilize. The details of FCM algorithm is given below:

The Fuzzy C-Means clustering algorithm is based on the minimization of an objective function called C-Means functional. It partitions the sample data into a number of clusters. These clusters have ‘fuzzy’ boundaries, in the sense that each data value belongs to each cluster to some degree or other.

FCM is an iterative algorithm to find cluster centers that minimize dissimilarity function. The algorithm starts with selecting the number of clusters as defined in the problem and initializing the
membership matrix $U$. This matrix contains the membership value for all points for each cluster. The initialization is done randomly and the cluster centers are computed using the membership matrix $U$. The FCM algorithm is as follows:

1. The membership matrix ($U$) is randomly initialized according to the equation

$$\sum_{i=1}^{n} U_{ij} = 1, i = 1, 2, \cdots n$$

(8.15)

Where $c$ is the number of clusters

2. Calculate the dissimilarity function or objective function given by the equation

$$J(U, c_1, c_2, \cdots c_c) = \sum_{i=1}^{c} \sum_{j=1}^{n} U_{ij}^m d_{ij}^2$$

(8.16)

Where $U_{ij}$ is between 0 and 1 and.

$c_{ij}$ is the centroid of cluster 1.

$d_{ij}$ is the Euclidian distance between $i^{th}$ centroid and $j^{th}$ data point. To reach the minimum of dissimilarity function there are two conditions given by the equations

$$c_i = \frac{\sum_{j=1}^{n} U_{ij}^m x_j}{\sum_{j=1}^{n} U_{ij}^m} \quad \text{and} \quad d_{ij}$$

(8.17)
By iteratively updating the cluster centers and the membership grades for each data point, FCM iteratively moves the cluster centers to the correct location within the data set.

### 8.6 SIMULATION RESULTS (synthetic data)

The disturbance waveforms are generated using the MATLAB simulation in section 3.4.1. Totally, 3000 signals are collected, out of which, 500 disturbance-free waveforms, and the remaining are the disturbance signals (500 signals for each disturbance type). 300 set of signals are used for training and 200 set of signals are used for testing. These signals are processed through HHT.

The relevant features are extracted by finding the energy of the IMFs. These features are given as input to the RBF Neural Network with the unit centers selected using FCM algorithm for training. Then the network is tested with the test data. The performance of the trained ANN during testing is given in Table 8.2. For comparison, the performance measures of the RBF network based on k-means algorithm are evaluated and the results are given in Table 8.2. Similarly, HHT and HT based performance measures are compared.
Table 8.2  Comparison of performance measures of HHT and HT based RBFNN using FCM and k-means clustering

**Algorithm (Simulated data)**

<table>
<thead>
<tr>
<th>Performance Measures</th>
<th>Hilbert-Huang Transform</th>
<th>Hilbert Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without noise</td>
<td>With noise</td>
</tr>
<tr>
<td><strong>Type of clustering algorithm</strong></td>
<td>FCM</td>
<td>k-Means</td>
</tr>
<tr>
<td>Classification Accuracy (%)</td>
<td>99.7</td>
<td>98.6</td>
</tr>
<tr>
<td>Mean Square Error (MSE)</td>
<td>0.0002</td>
<td>0.0083</td>
</tr>
<tr>
<td>Secure Misclassification (SMC) (%)</td>
<td>1.216</td>
<td>3.184</td>
</tr>
<tr>
<td>Insecure Misclassification (IMSC) (%)</td>
<td>0.016</td>
<td>1.0124</td>
</tr>
</tbody>
</table>
From Table 8.2, it is found that, the HHT based RBFNN produces high classification rate and results in low MSE, SMC and IMSC. Also, RBFNN based on FCM algorithm produces best performance results.

8.6.1 Performance results for real-time power quality disturbances

Next, the performance of the RBF network is evaluated with the real time power quality data generated using the experimental setup explained in section 3.4.3. Totally 500 samples are generated for each disturbance type. Out of this, 300 samples are used for training and 200 for testing. The relevant features are extracted by finding the energy of the IMFs as explained in section 8.1.2. These features are given as input to the RBF Neural Network with the unit centers being selected using FCM algorithm for training. Then the network is tested with the test data. The performance of the trained ANN during testing is given in Table 8.4. Here also, it is found that RBF based on FCM clustering algorithm produces high accuracy rate.

Next, the performance of the network is tested under the noisy environment. The disturbance signals are added with Gaussian noise of 30 dB. The relevant features are extracted from the noise corrupted signals based on the same method explained in section 8.1.2. These features are given as input to the RBFNN for training and it is further tested with the test data to assess the generalization ability of the network. The performance of the network during testing is given Table 8.4. From the results, it is found that the performance of the RBF is improved using FCM based clustering.
Table 8.3 Performance measures of HHT based RBFNN
(Real-time data)

<table>
<thead>
<tr>
<th>Performance measures</th>
<th>Hilbert-Huang Transform</th>
<th>Without noise</th>
<th>With noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of clustering algorithm</td>
<td>FCM</td>
<td>FCM</td>
<td>k-Means</td>
</tr>
<tr>
<td>Classification Accuracy, CA(%)</td>
<td>98.3</td>
<td>96.7</td>
<td>95.0</td>
</tr>
<tr>
<td>Mean Square Error (MSE)</td>
<td>0.0058</td>
<td>0.033</td>
<td>0.068</td>
</tr>
<tr>
<td>Secure Misclassification SMC (%)</td>
<td>2.800</td>
<td>3.658</td>
<td>5.014</td>
</tr>
<tr>
<td>Insecure Misclassification ISMC (%)</td>
<td>0.071</td>
<td>4.783</td>
<td>5.395</td>
</tr>
</tbody>
</table>

Table 8.3 shows that the HHT based RBFNN using FCM clustering algorithm produces higher percentage of Classification Accuracy, less MSE, less SMC and ISMC values.
A new feature extraction method based on Hilbert-Huang Transform is presented in this chapter for automatic classification of power quality disturbances. An RBF network developed using FCM is used for the classification of disturbance waveforms. The simulation results show that the Hilbert-Huang Transform based RBF can categorize power quality disturbances efficiently and accurately.