Application of Classifier Integration Model to Disturbance Classification in Electric Signals

Dong-Chul Park

Abstract—An efficient classifier scheme for classifying disturbances in electric signals using Classifier Integration Model (CIM) is proposed in this paper. The classifier integration model does not use the entire feature vectors extracted from the original data in a concatenated form to classify each datum, but rather uses groups of features related to each feature vector separately. The previous Partitioned Feature-based Classifier (PFC) model is compared with the CIM in terms of classification accuracy. In order to compare the performances of the PFCM and CIM, Fourier transform and wavelet transform are utilized for extracting features in disturbance signals. Centroid Neural Network is also adopted as the local classifiers for PFCM and CIM. In order to evaluate the classification accuracies of different classifier models, experiments are performed with several types of disturbances in signals. The results show that the proposed CIM can improve its classification accuracy over conventional classification algorithms.

Keywords—classifier design, disturbance signal, feature, neural networks.

I. INTRODUCTION

The signal classification task is one of the ongoing important topics in various pattern recognition tasks. Especially, many signal classification methods have been proposed and traditional signal classification methods consist of a feature extraction procedure and a classifier design process. One very important characteristic of conventional classifier schemes is that their classification accuracy heavily depends on the feature extraction method they adopt and finding a specific best feature extractor for a specific object becomes a critical task for the object classification problem. Several techniques including STFT (Short Time Fourier Transform) [1] and wavelet transform [2] have been successfully used. Fourier transform can analyze signals in frequency domain. However, it is not easy to find time information from the Fourier transformed signal. The STFT was introduced to solve this problem by applying time windows to time domain signals and finding their Fourier transforms inside of each windowed signal separately. In addition to STFT, wavelet transform has been widely used for retrieving time and frequency information from time domain signals. Wavelet transform is used to divide a time domain signal into wavelets and can construct a time-frequency representation of a signal that offers excellent time and frequency information. In wavelet transform, we need to choose a mother wavelet with scaling function and high vanishing moments. Instead of covering all the spectrum with an infinite number of levels, a finite combination of the scaling function to cover the spectrum is utilized and the number of wavelets required to cover the entire spectrum can be effectively reduced. With various efforts, different features are extracted for different reasons: nevertheless, all the features can help to describe signals more precisely. In most cases, more features facilitate more accurate classification. When using different features for describing a signal, different magnitudes may become a problem, because each feature is independent of the others. With various efforts, different features are extracted for different reasons: nevertheless, all the features can help to describe signals more precisely. In most cases, more features facilitate more accurate classification. When using different features for describing a signal, different magnitudes may become a problem, because each feature is independent of the others [3]. Centroid Neural Network (CNN) [4] is utilized as the local classifier for PFC model [5] and the PFC model is utilized for detecting disturbance signals from normal signals [3]. For real time detection and classification of disturbance signals in electrical signals, neural network has been successfully utilized [6]-[8]. Even though PFC model is successfully used for real time detection and classification of disturbance signals in electrical signals, an improved version of PFC model, CIM, can yield more accurate classification results because CIM utilizes the information on how the local classifiers makes mistakes in classifying disturbance signals in addition to the correct classifications[9]. That is, CIM can utilizes both of incorrect and correct classification tendencies of local classifiers while PFC model utilizes only the correct classification results of local classifiers. This additional information of incorrect classification tendencies of local classifiers can play an important role in improving the classifier accuracy of disturbance signal classification problem and this improvement can be witnessed in this paper.

The rest of this paper is organized as follows: Section II and Section III provide the brief summaries of feature extraction methods and classifier models including CNN, respectively. Experiments on several disturbance signals with various classifier models are reported in Section IV. Section V concludes this paper.
II. FEATURE EXTRACTION METHODS AND CENTROID NEURAL NETWORK

Fourier transform and Wavelet transform is adopted as feature extraction methods in this paper [3]. Fourier transform is a widely used tool for analyzing frequency information in various signals. In this paper, original signal is assumed to have its fundamental frequency as 50 Hz and various kinds of disturbances are added to the original signal. Figure 1 shows various disturbance signals and their Fourier transform results.

In attempting to solve the time series data problem, many researchers have thought to represent the original complex signal in many simpler signals in a cascade. Therefore, wavelet transform analysis is used to decompose the original signal at many different frequencies with different resolutions. On the other hand, it produces high-quality local representations of original signal in both time domain and frequency domain. In multiresolution model, the discrete wavelet transform (DWT) plays low-pass filters. First, a low-pass filter is used to suppress the high frequency components of a signal while allowing the low frequency components to pass through. The scaling function associated with the low-pass filter is then used to calculate the average of elements, which results in a smoother signal. The smooth data \( c_j(t) \) at given resolution \( j \) can be obtained by performing successive convolutions with the discrete low-pass filter \( h \). Next, the received signal is to become an input signal of another low-pass filter. The synthetic process to obtain the signal by summing the signal at each scale.

\[
c_j(t) = \sum_k h(k)c_{j-1}(t + 2^{j-1} k)
\]

where \( h \) is a discrete low-pass filter and \( c_0(t) \) is the original signal.

From the sequence of smoothing signals, wavelet transforms are obtained by smoothing signals.

\[
w_j(t) = c_{j-1}(t) - c_j(t)
\]

Consequently, the original signal is represented into many wavelets from coarse resolution level to the finest resolution level. Figure 2 shows examples of wavelet transform results on two disturbance signals.

The CNN [4] is utilized as the local classifier in this paper. The CNN is an unsupervised competitive learning algorithm based on the classical k-means clustering algorithm. It finds the centroids of clusters at each presentation of the data vector. The CNN first introduces definitions of the winner neuron and the loser neuron. When a data \( x_i \) is given to the network at the epoch \((k)\), the winner neuron at the epoch \( (k) \) is the neuron with the minimum distance to \( x_i \). The loser neuron at the epoch \( (k) \) to \( x_i \) is the neuron that was the winner of \( x_i \) at the epoch \( (k-1) \) but is not the winner of \( x_i \) at the epoch \( (k) \). The CNN updates its weights only when the status of the output neuron for the presenting data has changed when compared to the status from the previous epoch.

When an input vector \( x \) is presented to the network at epoch \( n \), the weight update equations for winner neuron \( j \) and loser neuron \( i \) in CNN can be summarized as follows [4]:

\[
w_{j}(n + 1) = w_{j}(n) + \frac{1}{N_{j+1}}[x(n) - w_{j}(n)]
\]

\[
w_{i}(n + 1) = w_{i}(n) - \frac{1}{N_{i+1}}[x(n) - w_{i}(n)]
\]

Where \( w_{j}(n) \) and \( w_{i}(n) \) represent the weight vectors of the winner neuron and the loser neuron, iteration, respectively.
The CNN has several advantages over conventional algorithms such as SOM or k-means algorithm when used for clustering and unsupervised competitive learning. The CNN requires neither a predetermined schedule for learning gain nor the total number of iterations for clustering. It always converges to sub-optimal solutions while conventional algorithms such as SOM may give unstable results depending on the initial learning gains and the total number of iterations. Note that the CNN was designed for deterministic data because the distance measure used in the CNN is the quadric (Euclidean) distance. More detailed description on the CNN can be found in [4].

### III. Classifier Models

With a given set of training data in conventional classifier shown in figure 3, the trainer finds a set of code vectors called a codebook by using various clustering algorithms when the classifier is based on unsupervised learning algorithms. Once the codebook is found and given test data, the trained classifier finds the closest code vector from the codebook and the test data is labeled by the class that represents the closest code vector. In conventional classifier, N classes of features describing the data are together used as input to the trainer and classifier. In this case, each feature usually has different properties including its magnitude. When using different features for describing input data, different magnitudes may present a problem, because each feature is independent of the others. In using N sets of combined features that consist of different individual features for a classification problem, the different individual features are typically assigned the same minimum and maximum magnitude (usually zero and one). However, owing to this normalization process on features with different characteristics and the use of unnatural treatment of these different features for training, the characteristics of individual features are not fully exploited. As a consequence, the classification results may not be accurate.

Another method to use features with different characteristics is shown in figure 4. Here, each group of features with the same characteristics is treated independently from other features and used separately in the local classifier. Various approaches to utilize different feature vectors are proposed including Ensemble Classifier. A classifier model, called Partitioned Feature-based Classifier (PFC), that efficiently uses various available features extracted through various tools and enhances the classification performance was proposed [3]. In PFC, the entire available features are grouped into several groups, where each group has homogeneous features and forms a feature vector. Each feature vector in PFC is separately used in the independent classifier and can preserve the properties of individual features in the same group. Each local classifier is independently trained with a specific group of features. Since each feature group is used only for a local classifier in the training stage independently, features from different groups will not interfere with each other. In PFC, each local classifier produces a specific accuracy result on the classification during the training stage. The accuracy for each local classifier is then used as a weight for the local classifier when making decisions in the test stage. The PFC model demonstrates that conventional clustering algorithms can significantly improve their classification accuracy when the PFC model is used with them. For given data, however, the PFC model draws the classification results based on each local classifier's performance regardless of how each classifier performs on each class. If we know from the training data set that a certain trained local classifier classifies very well on a certain class of data while it gives very inaccurate classification results on another class of data, then it may be reasonable to give a different credit according to the classification results from the local classifier.

In order to improve the classification accuracy, the classifier integration model (CIM) is introduced as a fusion method for multiple classifier [9]. As shown in figure 5, individual local
Comparision of classification accuracy

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Accuracy Average</th>
<th>Std. Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid Neural Network</td>
<td>0.961</td>
<td>1.328</td>
</tr>
<tr>
<td>PFC Model</td>
<td>0.972</td>
<td>0.890</td>
</tr>
<tr>
<td>Classifier Integration Model</td>
<td>0.994</td>
<td>0.782</td>
</tr>
</tbody>
</table>

Comparision of training time

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Training time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid Neural Network</td>
<td>1.22</td>
</tr>
<tr>
<td>PFC Model</td>
<td>0.33</td>
</tr>
<tr>
<td>Classifier Integration Model</td>
<td>0.31</td>
</tr>
</tbody>
</table>

IV. Experiments and Results

In order to evaluate the proposed CIM for the classification of disturbance signals, basic sinusoidal signals with 50 Hz are generated. 7 different disturbance classes including sag, swell, interruption, harmonics, flicker, sag harmonics, swell harmonics are also generated. For each of 7 disturbance classes, 200 disturbance signal data with random amplitudes and random starting times are also collected. From the available data sets, 90% data in each class are randomly chosen for training classifiers while the remaining data are used for evaluating classifiers. This training data and test data combination is collected 10 times. That is, there exist 10 different combinations of randomly chosen training data sets and test data sets. The accuracy of different algorithms used in experiments is reported by mean and variance from these data sets. Through experiments, CNN and PFCM are compared with the proposed CIM-based classifier. For CNN, the number of clustered is set to seven and the final class for a test data is determined by calculating the distance to the code vector for each cluster. For PFC model classifier and CIM-based classifier, two different local classifiers are utilized; one with low-frequency features and another one with wavelet features.

The results are summarized in Table I. As shown in Table I, the CIM-based classifier outperforms the conventional CNN and PFCM-based classifier in most cases. Unlike CIM model, classifier of CNN shows some misclassification results on sag, sag harmonic, and swell harmonic signals. Experiments with various numbers of code vectors are performed. Local classifiers used in CIM model can separate different disturbances in different ways. In order to compare the training speed for different classifiers, training times are measured under the following computational environment: CPU: Intel Core2 Quad Q8200 2.33GHz, RAM: 6 GByte and 64bit Window 7. As shown in Table II, the CIM and PFCM require 25.4% of training time when compared with conventional CNN algorithm.

IV. Conclusion

A classification model for disturbance signals is proposed by using Classifier Integration Model that efficiently utilized all the available features extracted through various feature extraction methods from training data sets and enhances the classification accuracy. Instead of using the entire feature vectors extracted from the original data at once, the CIM groups the entire available features into several groups, where each group has homogeneous features and forms a feature vector. Each feature vector is separately used in the independent local classifier and preserves the properties of individual features in the same group. The results demonstrate that the CIM-based classifier improves its classification accuracy significantly while the conventional classifier utilizing CNN gives misclassifications on disturbance signals such as sag, sag harmonic, and swell harmonic signals. The CIM-based classifier utilizes each local classifier’s classification tendency while the classifier based on PFC-model based does not. When compared with the classifier based on PFC-model, the CIM-based classifier improves the
classification accuracy slightly. The CIM-based classifier also saves its training time significantly when compared with CNN based classifier. With further experiments on larger training data sets, the training speed and classification accuracy can be measured and its advantage will be witnessed further.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of The MKE/KEIT (10040191, The development of Automotive Synchronous Ethernet combined IVN/OVN and Safety control system for 1Gbps class).

REFERENCES


Dong-Chul Park received the B.S. degree in electronics engineering from Sogang University, Seoul, Korea, in 1980, the M.S. degree in electrical and electronics engineering from the Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1982, and the Ph.D. degree in electrical engineering, with a dissertation on system identifications using artificial neural networks, from the University of Washington (UW), Seattle, in 1990. From 1990 to 1994, he was with the Department of Electrical and Computer Engineering, Florida International University, The State University of Florida, Miami. Since 1994, he has been with the Department of Electronics Engineering, MyongJi University, Korea, where he is a Professor. From 2000 to 2001, he was a Visiting Professor at UW. He is a pioneer in the area of electrical load forecasting using artificial neural networks. He has published more than 140 papers, including 40 archival journals in the area of neural network algorithms and their applications to various engineering problems including financial engineering, image compression, speech recognition, time-series prediction, and pattern recognition. Dr. Park was a member of the Editorial Board for the IEEE TRANSACTIONS ON NEURAL NETWORKS from 2000 to 2002. He is a Senior Member of IEEE.