International Journal of Advance Research in Science and Engineering
Volume No.06, Special Issue No.(01), December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

Soft Computing Based Pattern Recognition Approach Using Mel-Frequency Cepstral Coefficients

Yadaiah V[1], Dr. R P. Singh[2]

[1.] *Research Scholar, Department of Computer Science and Engineering, Sri Satya Sai University of Technology and Medical Science, Sehore,Bhopal,Madhya Pradesh ,(India)*

[2.] *Professor, Department of Computer Science and Engineering,Sri Satya Sai University of Technology and Medical Science,Sehore,Bhopal, Madhya Pradesh , (India)*

ABSTRACT

Pattern recognition is one of the four cornerstones of Computer Science. It involves finding the similarities or patterns among small, decomposed problems that can help us solve more complex problems more efficiently. During the past twenty years, there has been a considerable growth of interest in problems of pattern recognition and image processing. Pattern recognition using soft computing deals with imprecision, uncertainty, partial truth, and approximation to achieve practicability, robustness and low solution cost. As such it forms the basis of a considerable amount of machine learning techniques. The pattern recognition techniques are widely used in many other applications such as data mining, face recognition, handwriting recognition etc. and in much more applications. A pattern can be anything. It can be human face, signals of speech, handwritten word etc. The problem of pattern recognition is classifies as classes which are defined by the system designer. The system designer has the full right to specify the needed constraints on the classes. With the advancement of technology, research works are going on inventing new techniques to make the process of data analysis less complex.

Keywords—Soft Computing, Machine learning, Imprecision, Uncertainty, Robustness.

I. INTRODUCTION

There are many traditional methods for pattern recognition which are used intensively. These days, the methodology of artificial neural is very popular. There are many things which needed to be cared in order to design a recognition system. These points are: definition of pattern classes, environment of sensors, representation of patterns, extraction of features, analysis of selection, test samples and their training and finally, the evaluation of performance.

In spite of a lot of research work done, the significant problem of complex patterns and their recognition with accurate location and scale have been unsolved. The pattern recognition techniques are widely used in many other applications such as data mining, face recognition, handwriting recognition etc. and in much more applications. A pattern can be anything. It can be human face, signals of speech, handwritten word etc. The problem of pattern recognition is classifies as classes which are defined by the system designer. The system designer has the full right to specify the needed constraints on the classes. With the advancement of technology, research works are going on inventing new techniques to make the process of data analysis less complex. Since,

most of the companies have large databases, so the need of an automatic pattern recognition system is there and engineers are working in that direction. There are primarily three aspects to design a pattern recognition system. These are: pre-processing and acquisition of data, representation of data and decision making. These three components are essential for designing a pattern recognition system. Artificial neural networks are also used for the purpose of pattern recognition. The reason behind popularity of these networks is their capability to learn complex relationships easily and procedural algorithm used by these networks.

Feed-forward network is the best type of neural networks which is used quite regularly for pattern recognition. The reason behind the most usage of this feed-forward network is the presence of multi-layer perceptron in it. The architecture of whole network is updated in order to track it by using artificial neural networks. The biggest advantage of using artificial neural networks is that they don't depend much on the domain-specific knowledge and efficient algorithms used in it for the task of pattern recognition.

Pattern recognition is concerned primarily with the description and classification of measurements taken from physical or mental processes. Many definitions of pattern recognition have been proposed. Pattern recognition is one of the four cornerstones of Computer Science. It involves finding the similarities or patterns among small, decomposed problems that can help us solve more complex problems more efficiently. During the past twenty years, there has been a considerable growth of interest in problems of pattern recognition and image processing. This interest has created an increasing need for theoretical methods and experimental software and hardware for use in the design of pattern recognition and image processing systems.

A number of special languages have been proposed for the description of patterns such as English and Chinese characters, chromosome images, spark chamber pictures, two-dimensional mathematics, chemical structures, spoken words, and fingerprint patterns. For the purpose of effectively describing high dimensional patterns, high dimensional grammars such as web grammars, graph grammars, tree grammars, and shape grammars have been used for syntactic pattern recognition.

II.RELATED WORK

Mohammed Algabri (2017) A method that uses fuzzy logic to classify two simple speech features for the automatic classification of voiced and unvoiced phonemes is proposed. In addition, two variants, in which soft computing techniques are used to enhance the performance of fuzzy logic by tuning the parameters of the membership functions, are also presented. The three methods, manually constructed fuzzy logic (VUFL), fuzzy logic optimized with genetic algorithm (VUFL-GA), and fuzzy logic with optimized pstudy swarm optimization (VUFL-PSO), are implemented and then evaluated using the TIMIT speech corpus. Performance is evaluated using the TIMIT database in both clean and noisy environments. Four different noise types from the AURORA database-babble, white, restaurant, and car noise-at six different signal-to-noise ratios (SNRs) are used. In all cases, the optimized fuzzy logic methods (VUFL-GA and VUFL-PSO) outperformed manual fuzzy logic (VUFL). The proposed method and variants are suitable for applications featuring the presence of highly noisy environments. In addition, classification accuracy by gender is also studied.

Mehdi Sotudeh Chafi (2010) We propose a novel soft computing (SC) based approach to design fault detection and isolation (FDI) systems for industrial plants, in particular a highly nonlinear CNC X-axis drive system's

component fault detection. The aim of this study is twofold. One is to present a general description of various concepts such as the novel fuzzy-neuro architecture that uses fuzzy clustering to build a nominal model, fuzzy decision-making subsystems, a central processing unit for estimation of fault location, and finally RBF neural networks to estimate fault size. The other aim is to apply proposed method to diagnosis of component faults of a CNC X-axis drive system amid significant noise levels. Simulation results demonstrate the significance of the proposed approach.

Kreangsak Tamee (2015) We present a new result of pattern recognition generation scheme using a small-scale optical muscle sensing system, which consisted of an optical add-drop filter incorporating two nonlinear optical side ring resonators. When light from laser source enters into the system, the device is stimulated by an external physical parameter that introduces a change in the phase of light propagation within the sensing device, which can be formed by the interference fringe patterns. Results obtained have shown that the fringe patterns can be used to form the relationship between signal patterns and fringe pattern recognitions.

Ludovic Seifert (2014) The aim of this study was to propose a method for full-body movement pattern recognition in climbing, by computing the 3D unitary vector of the four limbs and pelvis during performance. One climber with an intermediate skill level traversed two easy routes of similar rates of difficulty (5c difficulty on French scale), 10m in height under top-rope conditions. The first route was simply designed to allow horizontal edge-hold grasping, while the second route was designed with more complexity to allow both horizontal and vertical edge-hold grasping. Five inertial measurement units (IMUs) were attached to the pelvis, both feet and forearms to analyse the 3D unitary vector of each limb and pelvis. Cluster analysis was performed to detect the number of clusters that emerged from coordination of the four limbs and pelvis during climbing performance. Analysis revealed 22 clusters with 11 clusters unique across the two routes. Six clusters were unique to the simple hold design route and five clusters emerged only in the complex hold design route. We conclude that clustering supported identification of full-body orientations during traversal, representing a level of analysis that can provide useful information for performance monitoring in climbing.

Barry K. Lavine (2006) Pattern recognition methods have been applied to a wide variety of chemical problems. In a typical pattern recognition study, samples are classified according to a specific property using measurements that are indirectly related to the property of interest. An empirical relationship or classification rule is developed from a set of samples for which the property of interest and the measurements are known. The classification rule can then be used to predict the property in samples that are not part of the original training set. In this review, the three major subdivisions of pattern recognition methodology are discussed and the analytical literature is surveyed. Much of the literature on pattern recognition focuses on novel and not so novel applications.

A. Chamkalani (2016) Asphaltene precipitation is one of the most common problems in both oil recovery and refinery processes. In oil recovery, formation of asphaltene aggregation followed by deposition causes blocking in the reservoir. This makes the remedial process costly and sometimes uneconomical. Unfortunately, there is no predictive model for asphaltene problem treatment. The problem is very difficult mainly because of the fuzzy

nature of asphaltene and the large number of parameters affecting precipitation. In this study, a novel and intelligent method is presented to screen asphaltene stability in oil by using a pattern recognition tool, an intelligent system that utilizes SARA analysis, to determine how severe a problem is.

III. PROPOSED PATTERN RECOGNITION TECHNIQUE

Pattern recognition using soft computing (PRSC) solutions are unpredictable, uncertain and between 0 and 1. Pattern recognition using soft computing became a formal area of study in Computer Science in the early 2000s. Earlier computational approaches could model and precisely analyze only relatively simple systems. More complex systems arising in biology, medicine, humanities, management science and similar fields often remained intractable to conventional mathematical and analytical methods. However, it should be pointed out that simplicity and complexity of systems are relative, and many conventional mathematical models have been both challenging and very productive. Pattern recognition using soft computing deals with imprecision, uncertainty, partial truth, and approximation to achieve practicability, robustness and low solution cost. As such it forms the basis of a considerable amount of machine learning techniques.

Speech is a non-stationary signal so to extract spectral features of sub-phones we analyse the spectrum in successive narrow time windows of about 20-25 ms width. For reliable frequency analysis, the human speech is considered to be fairly stationary over 20-25msec time windows. The analysis is carried out using the Fast Fourier transform algorithm (FFT) of each window. This gives us the intensity of several bands on the frequency scale. After digitization and quantization of the wave form our goal is to transform the input waveform into a sequence of acoustic feature vectors, such that each feature vector represent the information in a small time window of the signal. Mel Scale Cepstrum Coefficients (MFCC) are the most widely used features extracted by Cepstrum analysis of the signal. These MFCC are Human Listening perception based features. As the human ear is not equally sensitive to all frequency bands in MFCC also the features are extracted by attenuating the high frequency components using the mel scale. The overall extraction process can be represented in sequence of steps as defined below.

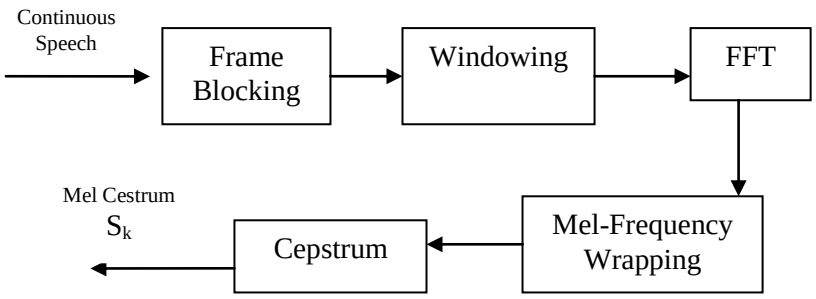

Figure: Block diagram of an MFCC processor

For obtaining MFCC features for the samples of the database words were divided into window of 25ms with frame rate of 10ms.The window taken during the extraction process is the Hamming window .Fast Fourier Transform was allowed on the windowed data and this was followed by bank of filters spaced logarithmically above 1000 Hz to obtain the Cepstrum coefficients. For training of the system one set of input was prepared for 12 coefficients from every frame of each word and the other set of input was taken as 12 MFCC coefficients from every frame along with one energy value.

IV. RESULTS AND DISCUSSION

Multi-layer Feed Forward Networks **[MLFFW]** are one of many different types of existing neural networks. They comprise of number of neurons connected together to form a network. The strengths or weights of the links between the neurons is where the functionality of the network resides. Neural networks are useful to model the behaviours of real-world phenomena. Being able to model the behaviours of certain phenomena, a neural network is able subsequently to classify the different aspects of those behaviours, recognize what is going on at the moment, diagnose whether this is correct or faulty, predict what it will do next, and if necessary respond to what it will do next. This study uses a multilayer feed forward neural networks with one hidden layer .The activation function at hidden and output layer is sigmoid and the network is trained with scaled conjugate gradient back propagation with momentum. The model can be extended to include more MFCC features for analysis purpose. These extracted features are fed as input to the network. These inputs are processed by hidden layers and fed to the output layer. Each neuron at the output layer corresponds to one input digit. Only one neuron is activated at one time. The overall training of the network is done in multiple epochs. The input for each frame is kept in the input file in required format. The training target is shown in the table where each of the digits will activate a different output neuron.

The analysis has been carried on the basis of number of neurons in the hidden layer as well as the number of speech features as input to the network is changed during its training. In the first phase the Input to the network is the 12 MFCC features of each frame of the word .In the second phase along with these energy of each frame is included as the 13^{th} input to the network. The network is trained with 75% of the total size of training database containing samples of digits recorded in neutral emotion.10% of the data is used for validation to check the generalization of the network where as rest 15% is used to test the network. Analysis is done on the basis of change of neurons in the hidden layer.

Table: Neural Network Configuration

Network Properties	**Information**
Training Method	Scaled Conjugate Gradient descent with momentum
Input Layer	12 MFCC and energy Features from short time duration frames

Neuron Transfer Function for hidden layer	Log sigmoid Transfer Function
Neuron Transfer Function for output layer	Log sigmoid Transfer Function
Epochs	1500
Learning Rate	0.01
Momentum Constant	0.9
Maximum performance parameter	0.01

Speech Signal

Features of an extracted speech signal include the words spoken by the person, pitch, pitch range and mean etc. In this research, we use a word extract method as the most general autocorrelation approach. We extracted a pitch value every 0.1 seconds and we calculated the average of the values defined by pitch mean and variance value was acquired the equal data. We canceled out the noise from our data because noise is a big factor which decreases the efficiency of our system. We obtained the selection number by looking for a concave extreme point of energy after finding a search starting point and ending point of a sentence from pitch contour.

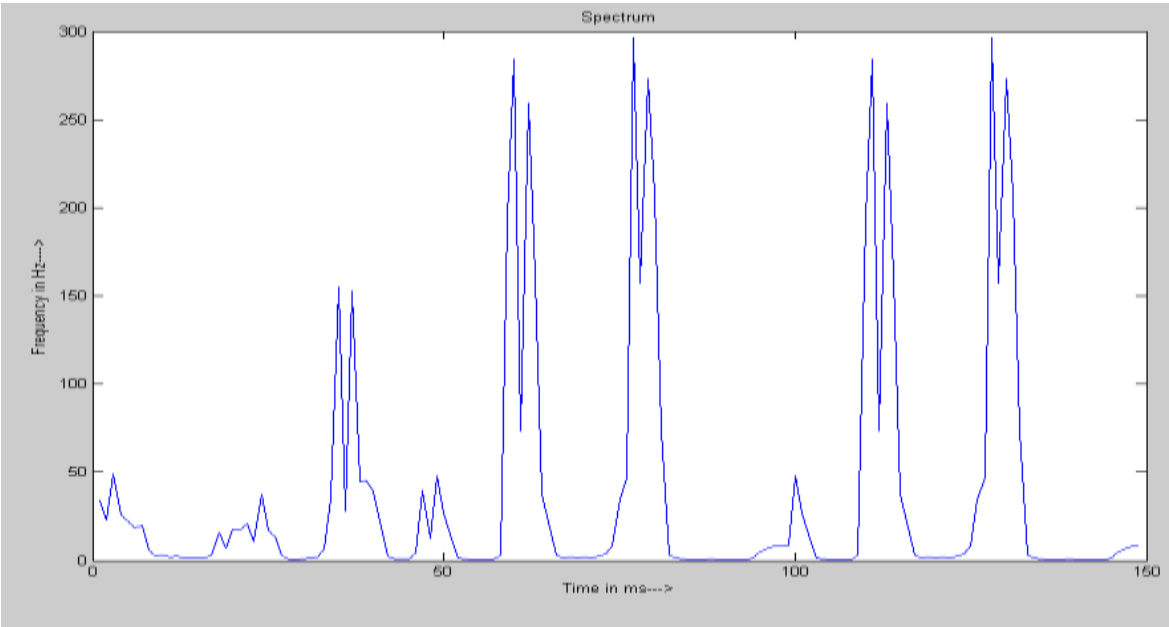

Figure : Power spectrum of the speech.

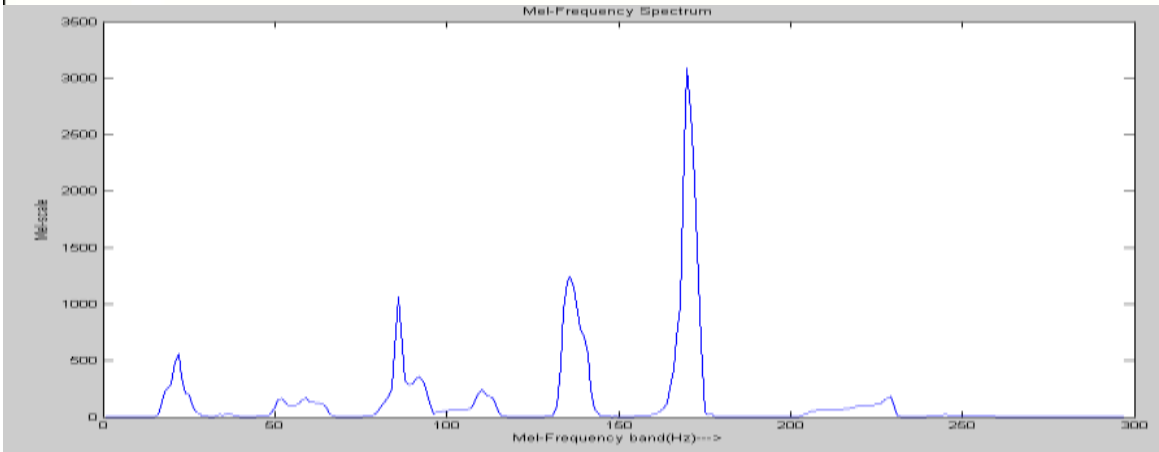

Figure: Mel-frequency spectrum of speech

V.CONCLUSION

Soft Computing is a consortium of methodologies which works synergistically and provides, in one form or another, flexible information processing capability for handling real life ambiguous situations. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning and partial truth in order to achieve tractability, robustness and low-cost solutions. Speech is the primary form of communication used by humans. Along with the linguistic content of speech, the way a word is said is equally important. The tone of the speech contains cues to the emotional state of the person speaking, and we as humans naturally recognize these emotions. When translating using speech recognition systems, emotions are only noise that degrades the performance of the speech recognition systems. Results show that the performance of the proposed work is better than the previous work.

REFERENCES

[1] Sucharita Mitra et al., A Simple Online Histogram and Pattern Recognition Based ECG Analyzer, IETE Journal of Research, 2008, Vol. 54, Issue 3, pp. 241-248.

[2] Lokendra Shastri et al., Character Recognition Digit Recognition Pattern Recognition Spatiotemporal Neural Networks Modular Networks Segmentation Problem, Connection Science, 1995, Vol. 7, Issue 3-4, pp. 211-246.

[3] Barry K. Lavine et al., Pattern Recognition, Critical Reviews in Analytical Chemistry, 2006, Vol. 36, Issue 3-4, pp. 153-161.

[4] Kreangsak Tamee et al., Fringe patterns generated by micro-optical sensors for pattern recognition, Artificial Cells, Nanomedicine, and Biotechnology, 2015, Vol. 43, Issue 4, pp. 252-257.

[5] Mohammed Algabri et al, Soft Computing Techniques for Classification of Voiced/Unvoiced Phonemes, Intelligent Automation & Soft Computing, 2017 , pp 1-7.

[6] Daniel L. Dinsmore et al., The Role of Initial Learning, Problem Features, Prior Knowledge, and Pattern Recognition on Transfer Success, The Journal of Experimental Education, 2014, Vol. 82, Issue 1, pp. 121-141.

[7] S. Sharma et al., A Novel Soft Computing-Based Retrieval System For Medical Applications, Applied Artificial Intelligence, 2012, Vol. 26, Issue 7, pp. 645-661.

[8] Konstantinos Demertzis et al., An innovative soft computing system for smart energy grids cyber-security, Advances in Building Energy Research, 2017, pp. 1-22.

[9] Cheng Zhiqiang et al., Control Chart Pattern Recognition Based on PSO and LS-SVM, HKIE Transactions, 2010, Volume 17, Issue 1, pp. 07-10.