International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 12, December 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

FEATURE CLUSTERING AND LABELLING TECHNIQUE FOR CONTINUOUS SPEECH RECOGNITION

M.Kalamani[1], M.Krishnamoorthi[2] ,R.S.Valarmathi[3]

[1,3]Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, (India)

[2]Department of CSE, Bannari Amman Institute of Technology, Sathyamangalam, Tamilnadu, (India)

ABSTRACT

In this research work, Feature Clustering and labelling technique is proposed through analysis in order to cluster and label the extracted feature vectors of continuous speech signal into discrete symbols and label them. Through various literature surveys, the Fuzzy C-Means is the most powerful unsupervised feature clustering and labelling technique and it employs the fuzzy partitioning, such that each data point belongs to all groups with different membership grades between 0 and 1. Simulation is carried out for different speech recognition algorithms with proposed feature clustering and labelling technique under various number of clusters for clean speech signal. From the results, it is observed that the optimal number of clusters as found as 5 and it is used for continuous speech recognition.

Keywords: Centroid, Continuous Speech Recognition, FCM, Feature Clustering, K-Means, Labelling.

I. INTRODUCTION

Clustering is one of the most useful data mining processes for discovering groups by identifying interesting distributions and patterns in the underlying data. It is an unsupervised process which groups similar objects. The resulting groups are called clusters and they ignore the labels while clustering. In this, the samples are interpreted as points in a d-dimensional euclidean space, and clustering is made according to the distances between points. Most clustering benchmarks deal with low dimensional real world or synthetic data sets, unlike speech data. A frequent problem that many clustering algorithms encounter is the choice of the number of clusters. Most of the clustering methods require the user to specify a value, though some provide a means to estimate the number of clusters inherent within the data [1,2].

K-means clustering is a very popular approximate method that can be used to simplify and accelerate convergence. In general, K-means does not achieve a global minimum over the assignments due to discrete rather than a set of continuous parameters. In addition, the results can greatly depend on initialization [3]. Further, the unsupervised K-means clustering algorithm is introduced and it is one of the usual clustering algorithms. This has a good sense of geometry and statistics significance. It is used to label the feature vectors before modelling. In this, speech features are clustered around centroid locations. The resulting cluster centers constitute a codebook which is used for training phase. It is widely used in many applications and it is more popular due to its simplicity, efficiency, and low cost of computation [4].

Fuzzy clustering is a process of assigning these membership levels and then using them to assign data elements to one or more clusters. One of the most widely used fuzzy clustering algorithms is the Fuzzy C-Means Algorithm. This algorithm minimizes the intra-cluster variance as well, but has the same problems as K-Means; the minimum is a local minimum, and the results depend on the initial choice of weights [5]. It is an unsupervised clustering technique, which is more suitable to normalize the extracted features. Hence, this is used for clustering the data vectors in order to reduce the word error rate in the CSR system [6,7].

This paper is organized as follows: Section II provides the K-means clustering technique. Section III presents the proposed Feature clustering and labelling technique for CSR system. Section IV provides the evaluation results and Section V provides conclusions.

II. K-MEANS CLUSTERING TECHNIQUE

Clustering is used to convert continuous-valued observation vectors into quantized discrete observation symbols before the classification or modelling process. It will generate the code book for training sequence which contains the acoustic information of the signal. For this purpose, K-means clustering is used for speech recognition. The K-means clustering is an algorithm to cluster ‘n’ objects based on attributes into ‘K’ partitions (where K < n). Given K, the number of clusters and training data are clustered iteratively to minimize the objective function as shown here under:

$$J = \sum_{j=1}^{K} \sum_{i=1}^{n} \left\| x_i^j - c_j \right\|^2 \quad (1)$$

where x_i^j is the ith data point in cluster j and c_j centroid of cluster j.

Then, the feature vector in the training matrix is clustered based on the squared Euclidean distance between a training vector and each of the code vectors in the codebook. The total distortion is computed once when all the training vectors have been labelled with the corresponding code vectors. Total quantization error is the summation of all these distances [8, 9].

III. PROPOSED FEATURE CLUSTERING AND LABELLING TECHNIQUE

Fuzzy C-means (FCM) clustering is the most powerful fuzzy clustering technique. This is an unsupervised method for the analysis of data and construction of models. It is converged to the local minima and this process is more natural than hard clustering. FCM clustering employs fuzzy partitioning such that a data point can belong to all groups with different membership grades between 0 and 1. In this, membership values are assigned based on the distance between the cluster center and the data points. In this approach, the membership values and cluster centers are updated iteratively [5, 9,10]. The FCM clustering algorithm has the following steps:

1. Let the N_T data points represented by x_i, i = 1,2, … , N_T are to be clustered.
2. Assume the number of clusters to made, i.e., CL, where $2 \leq CL \leq N_T$.
3. Choose an appropriate level of cluster fuzziness, f > 1.
4. Initialize the $N_T \times CL$ size membership matrix at random, such that U_{ij}€ [0 ,1] and $\sum_{j=1}^{CL} U_{ij} = 1$ for each 'i'
5. Determine the cluster centers CC_j , for j^{th} cluster using the expression given below:

$$cc_j = \frac{\sum_{i=1}^{N_T} U_{ij}{}^f x_i}{\sum_{i=1}^{N_T} U_{ij}{}^f} \tag{2}$$

6. Calculate the Euclidean distance, D_{ij} between i^{th} data point and j^{th} cluster center with the following formula:

$$D_{ij} = \left\| x_{im} - CC_j \right\| \tag{3}$$

7. Update fuzzy membership matrix **U** according to D_{ij}. If D_{ij}>0, then

$$U_{jij} = \frac{1}{\sum_{C=1}^{CL} \left(\frac{D_{ij}}{D_{iC}} \right)^{\frac{2}{f-1}}} \tag{4}$$

8. If D_{ij}=0, then the data point coincides with the corresponding data point of j^{th} cluster center CC_j and it has the full membership value, that is, U_{ij}=1. Repeat from Step 5 to Step 7 until the changes in$U<\chi$, where χ is a pre-specified termination criterion.

IV. RESULTS AND DISCUSSIONS

Throughout the evaluation process, the different input speech signals from the database of IIIT used. For these database signal, simulation is carried out for various number of clusters for different speech recognition algorithms under clean speech environment. Figure 1 shows the recognition accuracy for different Speech

Recognition Algorithms under various Number of Clusters for Clean Speech. From the results, it is observed that the significant improvement in recognition accuracy is found at number of cluster as 5 and this is considered as optimal feature cluster value for continuous speech recognition algorithms.

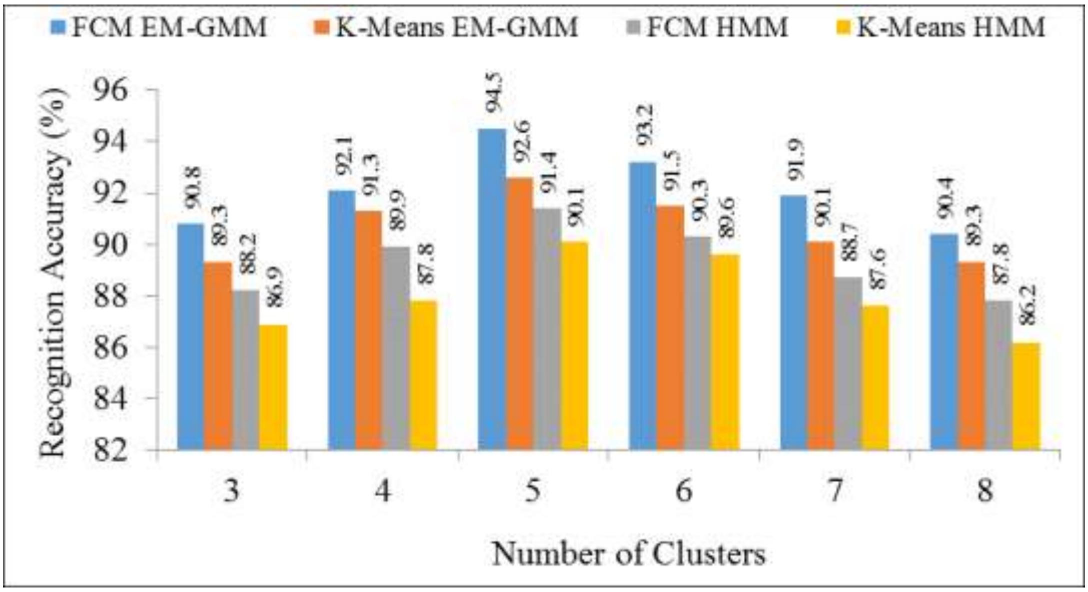

Fig.1. Recognition Accuracy (%) for the different Speech Recognition Algorithms under various Number of Clusters for Clean Speech

V. CONCLUSION

Clustering is used to convert extracted continuous-valued feature vectors into quantized discrete observation symbols before the classification or modelling process of continuous speech recognition. In this research work, FCM feature clustering and labelling technique is proposed for continuous speech recognition. From the simulation results, it is observed that the significant recognition accuracy improvement is obtained for number of cluster as 5 and it is used as optimal value for continuous speech recognition. Hence, it is concluded that this optimal FCM clustering technique will provides best clustering and labeling of the extracted feature vectors for the real time Continuous Speech Recognition.

REFERENCES

[1]. T.Kinnunen, T.Kilpeläinen, and P.FrÄnti, Comparison of Clustering Algorithms in Speaker Identification, Proceedings of the international conference on signal processing and communications, 2000.

[2]. J.R.Deller, J.G. Proakis, and J.H. Hansen, Discrete-Time Processing of Speech Signals, IEEE Press, New York, 2000.

[3]. Y.Linde, A. Buzo, and R.M. Gray, An algorithm for vector quantizer design, IEEE Transactions on Communications, 28(1), 1980, pp. 84-95.

[4]. X.G.Li, M.F.Yao, and W.T.Huang, Speech recognition based on k-means clustering and neural network ensembles, Proceedings of the IEEE seventh international conference on natural computation, 2, 2011, pp. 614-617.

[5]. S.Chattopadhyay, D.K. Pratihar, and S.C.D. Sarkar, A Comparative study of Fuzzy C-Means Algorithm and Entropy-based Fuzzy Clustering Algorithms, Computing and Informatics, vol. 30, no. 4, 2011, pp. 701-720.

[6]. M.Vyas, A Gaussian Mixture Model based Speech Recognition system using MATLAB, Signal & Image Processing: An International Journal (SIPIJ), 4, 2013, pp. 109-118.

[7]. Q.Zhu, and A.Alwan, Non-linear feature extraction for robust speech recognition in stationary and non-stationary noise, Computer speech & language, 17(4), 2003, pp. 381-402.

[8]. R.S.Kurcan, Isolated word recognition from in-ear microphone data using Hidden Markov Models. Ph.D. thesis, Naval Postgraduate School Monterey CA, 2006.

[9]. G.Li, J. Zhang, and W.J. Freeman, Mandarin digital speech recognition based on a chaotic neural network and fuzzy c-means clustering, Proceedings of the IEEE international conference on fuzzy systems, 2007, pp. 1-5.

[10]. J.C.Bezdek, R. Ehrlich, and W.Full, FCM: The Fuzzy c-means clustering algorithm’, Computers & Geosciences, 10(2), 1984, pp. 191-203.