Disorder Speech Clustering For Clinical Data Using Fuzzy C-Means Clustering And Comparison With SVM Classification

C.R.Bharathi[1],Dr.V. Shanthi[2]
[1.] Research Scholar, Department of ECE, Sathyabama University,
AP, Vel Tech University, Avadi,Chennai, India.
bharathi_dham@yahoo.com
[2.] Professor, Department of MCA, St. Joseph's College of Engineering,
Chennai, India.

Abstract

Speech is the most vital skill of communication. Stammering is speech which is hesitant, stumbling, tense or jerky to the extent that it causes anxiety to the speaker. In the existing system, there are many effective treatments for the problem of stammering. Speech-language therapy is the treatment for most kids with speech and/or language disorders. In this work, mild level of mental retardation (MR) children speech samples were taken for consideration. The proposed work is, the acute spot must be identified for affording speech training to the speech disordered children. To begin with the proposed work, initially Clustering of speech is done using Fuzzy C-means Clustering Algorithm. Feature Extraction is implemented using Mel Frequency Cepstrum Coefficients (MFCC) and dimensionality reduction of features extracted is implemented using Principal Component Analysis (PCA). Finally the features were clustered using Fuzzy C-Means algorithm and compared with SVM classifier output[13].

***Keywords**: speech, stammering, Mel Frequency Cepstrum Coefficients (MFCC), Principal Component Analysis (PCA), clustering*

1. Introduction

A speech disorder refers to a problem with the actual production of sounds, whereas a language disorder refers to a difficulty understanding or putting words together to communicate ideas. A speech pathologist has been trained to assess and treat people who have a communication disability. Speech synthesis among the disabled children is the first step for making speech corrective tool kit.

In this paper, the MR pathological subject's speeches are analysed by comparing it with the normal children. This work is very useful for the speech practitioners to know in which position they have to improve the speech of the abnormal child.

The samples of the normal as well as the abnormal children speech have been obtained and with the aid of these samples, the further process has carried out. Features are extracted using the MFCC of both (normal and abnormal) speeches and the PCA is applied to the MFCC to reduce the dimensionality of the speeches. The parameters extracted from MFCC of both speeches are used for clustering using fuzzy C-means clustering algorithm.

Feature extraction is a key issue for efficient speaker recognition. Additionally, a reduced feature set would allow more robust estimates of the model parameters, and less computational resources would be required.

MFCC convey not only the frequency distribution identifying sounds, but also the glottal source and the vocal tract shape and length, which are speaker specific features, so MFCC are commonly used as features.

Clustering of numerical data forms the basis of many classification and system modelling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behavior.

Fuzzy c-means (FCM) is a data clustering technique in which a dataset is grouped into n clusters with every datapoint in the dataset belonging to every cluster to a certain degree. For example, a certain datapoint that lies close to the centre of a cluster will have a high degree of belonging or membership to that cluster and another datapoint that lies far away from the centre of a cluster will have a low degree of belonging or membership to that cluster.

The rest of the paper is organized as follows. The Feature Extraction, Clustering using Fuzzy C-Means Clustering are described in Section 2. The experimental setup is described in Section 3, including the comparision between Fuzzy C-means clustering and SVM classification and Experiments Results are presented and discussed in Section 4: (1) Feature Extraction using MFCC and PCA, (2) Fuzzy C-Means Clustering

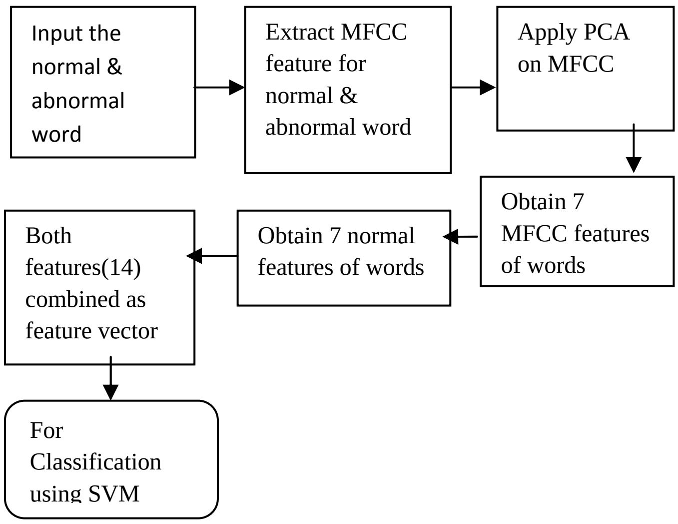
(3)comparison between Fuzzy C-means clustering and SVM classifier result. Finally, Section 5 summarizes our approach.

2. Methodology

2.1 Feature Extraction

Normal and pathological subject's speech dataset is obtained and then the MFCC is extracted from both the speeches. This speech dataset is developed through in which both the databases are same set of scripts. Subsequently, with the aid of the PCA the dimensionality is reduced for MFCC features extracted. The step by step processes are explained in detail in the following subsections. [13]

Fig.1 Data Flow for Feature Extraction



The speech samples are extracted from the normal and pathological subjects with the aid of the audio synthesizer. Let AB D_a and N D_b are the abnormal children, normal children speech datasets respectively and from these datasets the MFCC feature is extracted.

$$ABD_a = \{w_1, w_2, w_3 ... w_{N_w - 1}\} \quad (1)$$

$$ND_b = \{\omega_1, \omega_2, \omega_3 ... \omega_{N_w - 1}\} \quad (2)$$

MFCC's are based on the known variation of the human ear's critical bandwidths with frequency. The MFCC technique makes use of two types of filter, namely, linearly spaced filters and logarithmically spaced filters. The MFCC features are obtained from the normal as well as abnormal datasets which is referred as N M_a and AB M_b

PCA is used abundantly in all forms of analysis - from neuroscience to computer graphics - because it is a simple, non-parametric method of extracting relevant information from confusing data sets. With minimal additional effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified dynamics that often underlie it. PCA was used to reduce the dimensionality. After this process completed, the following parameters are obtained from the MFCC featured vectors N M_a, ABM_b. The parameters are mean, standard deviation, maximum amplitude value and its id, minimum amplitude value and its id, MFCC length are extracted for the MFCC featured word and as well as for the original word also extracted and hence for each word we have 14 inputs. [13]

2.2 Fuzzy C-Means Clustering

Clustering is a mathematical tool that attempts to discover structures or certain patterns in a data set, where the objects inside each cluster show a certain degree of similarity.

Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition. In fuzzy clustering, each point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to just one cluster. Thus, points on the edge of a cluster may be *in the cluster* to a lesser degree than points in the center of cluster.

Any point x has a set of coefficients giving the degree of being in the kth cluster $w_k(x)$. With fuzzy c-means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

$$c_k = \frac{\sum_x w_k(x)\,x}{\sum_x w_k(x)}.$$

The degree of belonging, $w_k(x)$, is related inversely to the distance from x to the cluster center as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest center. The fuzzy c-means algorithm is very similar to the k-means algorithm.

- Choose a number of clusters.
- Assign randomly to each point coefficients for being in the clusters.
- Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than ε, the given sensitivity threshold) :
    - Compute the centroid for each cluster, using the formula above.
    - For each point, compute its coefficients of being in the clusters, using the formula above.

3. Experimental Setup

3.1 The speech database

In this with the aid of the **Free Audio Editor** we generate the dataset with the normal and abnormal female children within the age limit 6-10 yrs. For normal data, 2 female children were utilized and for abnormal data, a female child was utilized for the system and the normal frequency range is from 20 - 4 kHz. The proposed technique is tested with the database of 100 words with two normal children and an abnormal child each.

3.2 Fuzzy C-Means Clustering

Fuzzy Cluster Analysis presents advanced and powerful fuzzy clustering techniques.It is based on minimization of the following objective function:

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

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-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|}\right)^{\frac{2}{m-1}}} \quad , \quad c_j = \frac{\sum_{i=1}^{N} u_{ij}^m \cdot x_i}{\sum_{i=1}^{N} u_{ij}^m}$$

This iteration will stop when $\max_{ij}\left\{\left| u_{ij}^{(k+1)} - u_{ij}^{(k)} \right|\right\} < \varepsilon$, where ε is a termination criterion between 0 and 1, whereas k are the iteration steps. This procedure converges to a local minimum or a saddle point of J_m.

The algorithm is composed of the following steps:

1. *Initialize $U=[u_{ij}]$ matrix, $U^{(0)}$*
2. *At k-step: calculate the centers vectors $C^{(k)}=[c_j]$ with $U^{(k)}$*

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

3. *Update $U^{(k)}$, $U^{(k+1)}$*

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

4. *If $\| U^{(k+1)} - U^{(k)} \| < \varepsilon$ then STOP; otherwise return to step 2.*

According to the dataset of normal and abnormal children speech, the clustering is done by tuning the 3 optional parameters for the FCM algorithm (exponent, maximum number of iterations and minimum amount of improvement) into two clusters(for normal and abnormal children).

4. Experimental Results

4.1 Feature Extraction

Initially, the words are extracted from the both normal and abnormal children and then the MFCC feature has been extracted from it. Subsequently, the PCA is applied to reduce the dimensionality of the words. Few speech samples which are sent as input to MFCC Feature Extraction.[13]

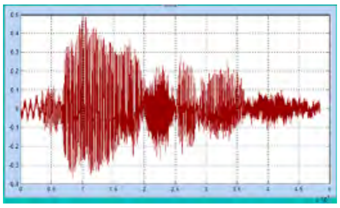


Fig. 2 Normal speech 1 “dinosaur”

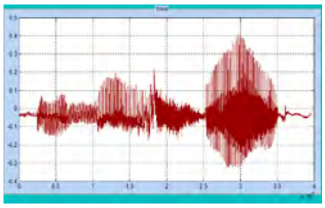


Fig. 3 Normal speech 2 “dinosaur”

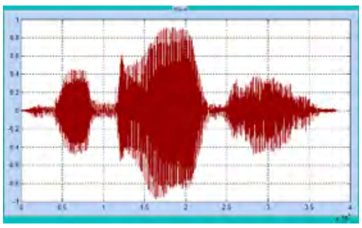


Fig. 4 Abnormal speech – “dinosaur”

4.2 Features extracted using MFCC

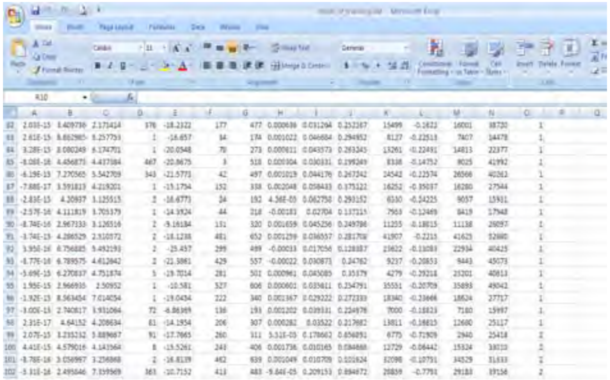
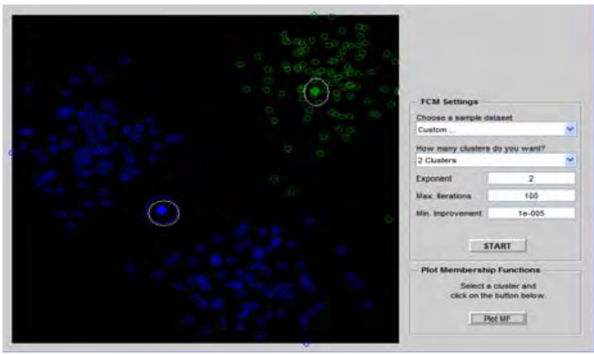


fig. 5 Parameter values of MFCC

4.3 Fuzzy C-Means Clustering and its Membership Function (MF)

The function fcm takes a data set and a desired number of clusters and returns optimal cluster centres and membership grades for each data point. This information helps to build a fuzzy inference system by creating membership functions that represent the fuzzy qualities of each cluster. According to the dataset of normal and abnormal children speech, the clustering is done by tuning the 3 optional parameters for the FCM algorithm (exponent, maximum number of iterations and minimum amount of improvement)

Fig. 6 Two Clusters output



A *membership function* (MF) is a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. The input space is sometimes referred to as the *universe of discourse*, a fancy name for a simple concept. The output-axis is a number known as the membership value between 0 and 1. The curve is known as a *membership function.*

Fig. 7 Cluster 1 (blue) Membership Function

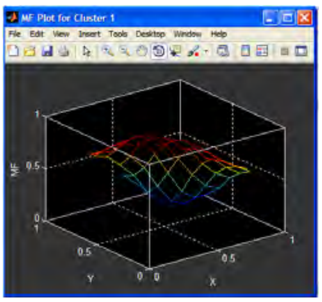
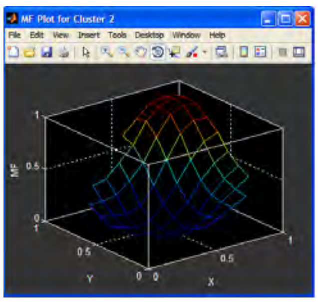


Fig 8. Cluster 2 (green) Membership Function



Certain datapoint that lies close to the centre of a cluster will have a high degree of belonging or membership to that cluster and another datapoint that lies far away from the centre of a cluster will have a low degree of belonging or membership to that cluster. Therefore from cluster 1, MF has low degree of belonging.

4.4 Comparison between Fuzzy C-means Clustering and SVM Classification

Classification is a type of supervised learning (Background knowledge is known) and Clustering is a type of unsupervised learning(No such knowledge is known).

4.4.1 SVM Classification [13]

The PCA is applied to MFCC features to reduce the dimensionality of the words and then they are inputted to the SVM to identify the abnormal and the normal word. The result is shown as [13],

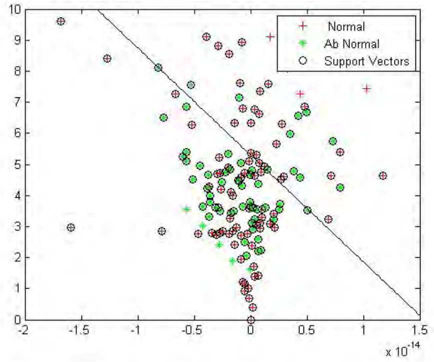


Figure 9. SVM classification Regression result to identify the abnormal word

Table 1 Some Output for the SVM classifier for the identification of abnormal word

Original	**Proposed (SVM)**
apple.wav	abnormal word
mango.wav	abnormal word
analysis	abnormal word
dinosaur	abnormal word
birds	abnormal word
ink pot	abnormal word
umbrella	abnormal word
zebra	abnormal word
vehicle	normal word
domestic animals	abnormal word

The MF of cluster 1 and cluster 2 got by Fuzzy C-means clustering is compared with the SVM classification[13] result. The MF has low degree of belonging when compared to MF of cluster2. This result is compared with SVM classifier output[13], the conclusion we get is the performance of the SVM classifier using MFCCs as features was very high, with up to 90% accuracy.

5. CONCLUSIONS AND FUTURE WORK

The proposed system was implemented in the working platform of MATLAB (version 7.11). In this with the aid of the Free Audio Editor we generate the dataset with the normal and abnormal female children within the age limit 6-10 yrs. For 100 normal data (words) we utilized 2 female children each and for 100 abnormal data we utilized a female child for our system and their normal frequency range is from 20 – 4khz. The speech input is recorded at a sampling rate of 44.1kHz. To develop an effective system to identify the abnormal word and spotting aberration in speech of pathological subject, where the speech has to be improved, initially the MFCC is obtained from both the normal and abnormal words and then PCA dimensionality reduction is done. Then clustering is done using FCM in this work and compared with SVM classifier result[13] of normal and abnormal words. The performance of the SVM classifier using MFCCs as features was very high, with up to 90% accuracy. After that, the work will be extended by (i) acute Spotting aberration in speech of pathological subject.

References

[1] Thomas Lotter, Christian Benien and Peter Vary, "Multichannel Direction-Independent Speech Enhancement Using Spectral Amplitude Estimation", EURASIP Journal on Applied Signal Processing, Vol. 2003, No. 11, pp. 1147-1156, 2003

[2] Sven Nordholm, Thushara Abhayapala, Simon Doclo, Sharon Gannot, Patrick Naylor and Ivan Tashev, "Microphone Array Speech Processing", EURASIP Journal on Advances in Signal Processing, Vol. 2010, pp. 1-3, 2010

[3] Md. Tariquzzaman*†, Jin Young Kim*††, Seung You Na*†† "A "Correction of Missing Reliability for Robust Bimodal Speaker Identification", 2009 International Conference on Information and Multimedia Technology

[4] Aida–Zade, Ardil and Rustamov, "Investigation of Combined use of MFCC and LPC Features in Speech Recognition Systems", World Academy of Science, Engineering and Technology, Vol. 3, No. 2, pp. 74-80, Spring 2007

[5] Rashad, Hazem M. El-Bakry and Islam R. Ismail, "Diphone Speech Synthesis System for Arabic Using MARY TTS ", International journal of computer science & information Technology (IJCSIT), Vol. 2, No. 4, pp. 18-26, August 2010

[6] Fusheng Yu, Ruiqiong Cai "Optimized Fuzzy Information Granulation of Temporal Data ", 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2010)

[7] NIU Dong-xiao, WANG Yong-li, MA Xiao-yong, "Optimization of support vector machine power load forecasting model based on data mining and Lyapunov exponents,"Journal of Pattern Recognition Research, Vol. 17, Pp. 406−412, 2010.

[8] R. Sangeetha, B. Kalpana, "Performance Evaluation of Kernels in Multiclass Support Vector Machines,"International Journal of Soft Computing and Engineering, Vol. 1, Pp. 138-145No. 5, 2011

[9] Giorgos Mountrakis, Jungho Im, Caesar Ogole, "Support vector machines in remote sensing, "Journal of Photogrammetry and Remote Sensing, Vol. 66, Pp. 247-259, 2011.

[10] Reginaldo K. Fukuchi, Bjoern M. Eskofier, Marcos Duarte, Reed Ferber, "Support vector machines for detecting age-related changes in running kinematics, Journal of Biomechanics, Vol. 44, Pp. 540-542, 2011.

[11] Ahmad Ghodselahi, "A Hybrid Support Vector Machine Ensemble Model for Credit Scoring, International Journal of Computer Applications, Vol. 17, No.5, Pp. 1-5, 2011.

[12] Sebastián Maldonado, Richard Weber, Jayanta Basak, "Simultaneous feature selection and classification using kernel-penalized support vector machines, Journal of Information Sciences, Vol. 181, Pp. 115-128, 2011.

[13] C.R. Bharathi,Dr. V. Shanthi,"Disorder speech Classification for Clinical Data using SVM" , National Conference by IETE, EXCITE, 2012

C.R.Bharathi received AMIE (ECE) and M.E.(Applied Electronics) in 2001 and 2005. Since June 2002, she has been a Assistant Professor in an Engineering College and at present working as Assistant Professor in Vel Tech University, Avadi, Chennai.She is an Research scholar at Sathyabama University.

Dr. V. Shanthi working as Professor in St.Joseph's College of Engineering,Chennai. Her research interest includes Artificial Intelligence, cloud computing. Dr. V. Shanthi is co-author.