Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 16 Issue 2 Version 1.0 Year 2016
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Accuracy Comparison using Different Modeling Techniques under Limited Speech Data of Speaker Recognition Systems

By Satyanand Singh & Ajeet Singh
JNTU Hyderabad, India

Abstract- Pointing towards programmed machine learning by human, a technique for speaker recognition with speaker identity in light of man machine interface is an interest of science. Motivated by the same, we propose a philosophy to recognize speakers. Inside of our investigation, obtaining speech signal, analysis of spectrogram, neutralization, extraction of speaker specific features for recognition, mapping of speech using Novel Vector Quantization (NFVQ) is presented. NFVQ is particularly suitable for colossal arrangement of information and yield discourse mapping. Furthermore Speaker Recognition by utilizing NFVQ Model additionally will be exhibited in this paper. During feature extraction, traditional triangular shaped bins have been replaced by Gaussian shaped filter (GF) and Tukey filter (TF) to calculate Mel Frequency Cepstral Coefficients (MFCC). This work performs an experimental evaluation of three simple modelling techniques namely, Fuzzy c-means, FVQ2 and NFVQ. Among these NFVQ shows significant improved performance compared to Fuzzy c-means and FVQ2. For about 10 s of training and testing speech data of speakers the efficiency for NFVQ, FVQ2 and Fuzzy c-means are 98.8%, 73.33, and 8, respectively, for a set of 630 speakers taken from the TIMIT database. We additionally got 5% outright EER change for both-sex trials on the 10 s-10 s condition contrasted with the FVQ2 approach.

Index-terms: *gaussian filter, triangular filter, tukey filter, subbands, MFCC, vector quantization, novel fuzzy vector quantization.*

GJSFR-F Classification : *MSC 2010: 97N80*

Strictly as per the compliance and regulations of :

Accuracy Comparison using Different Modeling Techniques under Limited Speech Data of Speaker Recognition Systems

Satyanand Singh [α] & Ajeet Singh [σ]

Abstract- Pointing towards programmed machine learning by human, a technique for speaker recognition with speaker identity in light of man machine interface is an interest of science. Motivated by the same, we propose a philosophy to recognize speakers. Inside of our investigation, obtaining speech signal, analysis of spectrogram, neutralization, extraction of speaker specific features for recognition, mapping of speech using Novel Vector Quantization (NFVQ) is presented. NFVQ is particularly suitable for colossal arrangement of information and yield discourse mapping. Furthermore Speaker Recognition by utilizing NFVQ Model additionally will be exhibited in this paper. During feature extraction, traditional triangular shaped bins have been replaced by Gaussian shaped filter (GF) and Tukey filter (TF) to calculate Mel Frequency Cepstral Coefficients (MFCC). This work performs an experimental evaluation of three simple modelling techniques namely, Fuzzy c-means, FVQ2 and NFVQ. Among these NFVQ shows significant improved performance compared to Fuzzy c-means and FVQ2. For about 10 s of training and testing speech data of speakers the efficiency for NFVQ, FVQ2 and Fuzzy c-means are 98.8%, 73.33, and 8, respectively, for a set of 630 speakers taken from the TIMIT database. We additionally got 5% outright EER change for both-sex trials on the 10 s-10 s condition contrasted with the FVQ2 approach.

Index-terms: *gaussian filter, triangular filter, tukey filter, subbands, MFCC, vector quantization, novel fuzzy vector quantization.*

I. Introduction

A speaker recognition system mainly consists of two main modules, speaker specific feature extractor as a front end followed by a speaker modelling technique for generalized representation of extracted features [1, 2]. Since long time MFCC is considered as a reliable front end for a speaker recognition application because it has coefficients that represents audio, based on perception [3, 4]. In MFCC the frequency bands are positioned logarithmically which approximated the human auditory systems response more closely than the linear spaced frequency bands of FFT or DCT. The main speaker specific information is pitch [5], residual phase [6], prosody [7], dialectical features [8] etc. These features are related with vocal chord vibration and it is very difficult to extract speaker specific information [9]. The MFCC modeled by Fuzzy c-Means, FVQ2 and NFVQ [10] technique.

In this paper a NFVQ is proposed for speaker recognition modelling. All the above vector quantization routines perform the codebook outline by utilizing crisp choice making systems [11], the feeling that every preparation vector is allocated to one

Author α σ: JNTU Hyderabad. e-mails: yogitechno@gmail.com, ajeetsingh534@gmail.com

and only group. Inevitably, these routines overlook the likelihood that a particular preparing vector might likewise have a place with another group. Fuzzy set theory created by Zadeh has been seen as a distinct option for more conventional contemplations keeping in mind the end goal to manage perplexing, poorly characterized and less scientifically justifiable frameworks [12].

The fundamental issue in Fuzzy logic is that a particular item can be allocated to more than one bunch with specific degrees of support [13]. The use of Fuzzy systems in speaker recognition gives two fundamental advantages. Firstly, Fuzzy set hypothesis can show the vulnerability included in the information set of the preparation vectors [14]. Furthermore, it offers a computational system, which is algorithmically furnished with a strong and all around organized scientific foundation [15]. The Fuzzy logic strategies, which can be productively utilized as a part of speaker recognition, are principally in view of Fuzzy bunching investigation [16]. The most illustrative Fuzzy bunching calculation is the understood Fuzzy c-means system, which was produced by Bezdek in [17]. The Fuzzy c-means system regards every group as a Fuzzy set and along these lines, the codebook configuration is a delicate choice making procedure [18]. Since Fuzzy grouping can demonstrate the vulnerability included in the segment of the preparation vector space, it can be utilized to take out or if nothing else fundamentally diminish the reliance of the codebook outline on the introduction [19].

In this paper we propose a NFVQ algorithm for speaker recognition. In the first step, we introduce a simple modification of the fuzzy c-means objective function and reformulate this objective function. In the next step, we extract analytical learning conditions for the codebook design by minimizing the reformulated function. In many real situations, fuzzy clustering is more natural than hard clustering, as objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial memberships. So the present study was undertaken with the objective of to find out the speaker recognition efficiency improving components with the help of a novel algorithms.

II. Fuzzy Clustering for Vector Quantization

In SR Vector quantization is fretful with the demonstration of a set of unlabeled data vectors $X = \{x_1, x_2, x_3 \ldots., x_n\} \in \mathbb{R}^p$ by a set $V = \{v_1, v_2, v_3 \ldots., v_c\} \in \mathbb{R}^p$ with $c \ll n$. Here x_k is called training vector and the set $\mathbf{X}$ is referred as training set, while each v_i is called codebook vector and the set V is referred as codebook. The key issue in vector quantization is the codebook outline. The codebook can be composed by utilizing hard or crisp choice making systems. In both cases, the nature of the last codebook is generally assessed by the accompanying normal distortion measure,

$$D = \frac{1}{n}\sum_{k=1}^{n} \min_{1 \le i \le c} \{\|x_k - v_i\|^2\} \quad (1)$$

We depict two Fuzzy grouping based vector quantization calculations in next segment, which are understood Fuzzy c-means and the FVQ2 created by Karayiannis and Pai in [20].

a) The Fuzzy c-Means Algorithm

The Fuzzy c-means is the most generally utilized calculation to deliver constrained fuzzy c-partitions in speaker recognition [21]. $u_{i\,k} = \{u_i(x_k), 1 \le i \le c, 1 \le$

Ref

16. Li-juan Ma, Ming hu Ha., "Support Vector Machines Based on Sectional set Fuzzy K-Means Clustering," Fuzzy Info. And Engineering, vol. 54, Jan 2009, pp. 420-425.

$k \leq n\}$ represents membership degree of the k^{th} training vector to i^{th} cluster. There are constrained in cluster if the next three conditions are satisfied,

$$0 \leq u_{i\,k} \leq 1\,, \quad \forall\, i,k$$

$$0 < \sum_{k=1}^{n} u_{i\,k} < n, \quad \forall\, i$$

$$\sum_{i=1}^{c} u_{i\,k} = 1, \quad \forall k \tag{2}$$

At whatever point the last situation is not fulfilled the Fuzzy c-means is said to be unconstrained. The usage of the Fuzzy c-means depends on the minimization, under the fairness imperative given in eq. (2),

$$J_m = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{i\,k})^m \|x_k - v_i\|^2 \tag{3}$$

Where $m \in (1,\infty)$ is a component to adjust the membership degree weighting effect. The cluster centers (codebook vectors) and the membership degrees that take care of the above compelled improvement issue are separately given by the accompanying mathematical statements [13],

$$v_i = \frac{\sum_{k=1}^{n}(u_{i\,k})^m\ x_k}{\sum_{k=1}^{n}(u_{i\,k})^m}\ , \qquad 1 \leq i \leq c \tag{4}$$

And

$$u_{i,k} = \frac{1}{\sum_{j=1}^{c}\left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|}\right)^{\frac{2}{m-1}}}\ , 1 \leq i \leq c\,, 1 \leq k \leq n \tag{5}$$

Mathematical statements in eq. (4) and (5) speak to an iterative enhancement methodology, where m is the Fuzziness controlling parameter. If m takes extensive values then the participation degrees of every preparation vector tend to approach 1/c.

In what follows, results due to the case studies are presented to minimize the objective function to enhance the percentage of speaker recognition accuracy. For example, let us consider a case of 64 clusters and 35 iteration. Fuzzy c-means clustering and its minimum objective function J_m= 0.1878. Fig. 1 shows the plot of objective function by fuzzy c-means clustering.

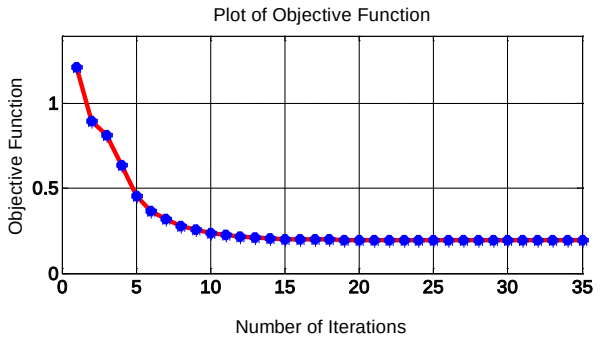

Fig. 1 : Plot of objective function J_m of Fuzzy c-means clustering

Ref
13. Jayanna, H.S., Prasanna, S.R. Mahadeva., "Fuzzy Vector Quantization for Speaker Recognition under Limited Data Conditions," IEEE Region 10 Conference, TENCON 2008, Hyderabad India, Nov 2008, pp.1-4.

Fig. 2 shows the plot of distortion of the spea kers using fuzzy c-means clustering.

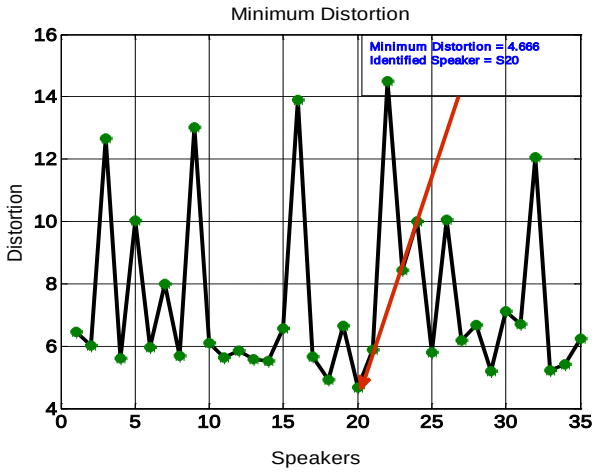

Fig. 2 : Plot of distortion measurement by fuzzy c-means clustering

Distortion of the voice data of the speakers during testing phase using fuzzy c-means clustering. D = 6.4711, 6.0261, 12.6531, 5.6056, 10.0395, 5.968, 8.0071, 5.6889, 13.0149, 6.0971, 5.6337, 5.8435, 5.5733, 5.5329, 6.5637, 13.8927, 5.6554, 4.9273, 6.6624, **4.6669**, 5.8925, 14.5138, 8.4431, 9.9943, 5.7945, 10.0552, 6.195, 6.6755, 5.192, 7.1233, 6.7116, 12.066, 5.209, 5.4083, 6.2515.

b) Fuzzy Vector Quantization2

Vector quantization is completed by relating every preparation vector to a solitary codebook vector. In this manner, the utilization of the Fuzzy c-means to vector quantization ought to be founded on relegating every preparation vector to the codebook vector. In any case, such a fresh understanding of the Fuzzy c-means amid the codebook configuration may affect the nature of the last codebook, since this methodology shrouds the presence of anomalies and replaces them by their nearest codebook vectors.

The arrangement of the codebook vectors that fit in with the hyper circle focused at the kth preparing vector is meant as T_k. At that point, the move from Fuzzy to crisp mode is proficient by steadily contracting the covering hyper circles amid the grouping procedure. In addition, it was found that the move speed straightforwardly influences the nature of the subsequent codebook [20]. As the configuration procedure continues, the set T_k is upgraded by taking after method:

In the ν-th iteration the set $\mathrm{T_k^{(v)}}$ contains $\aleph \mathrm{T_k^{(v)}}$ codebook vectors. The average distance is defined as,

$$\check{d}_k^{(v)} = \frac{1}{\aleph(T_k^{(v)})} \sum_{v_i \in T_K^{(v)}} \|x_k - v_i\|^2 \tag{6}$$

The T_k is updated in the $(v+1)th$ iteration as follows,

$$T_k^{(v+1)} = \left\{ v_i \in T_k^{(v)} : \|x_k - v_i\|^2 \leq \check{d}_k^{(v)} \right\} \tag{7}$$

The above upgrading guideline requires that the enrollment degrees of $\mathrm{x_k}$ to the codebook vectors, which during the $(i+1)^{th}$ iteration are removed from the set $\mathrm{T_k^{(v)}}$, are set equal to zero. In this manner, at first, every preparation vector is relegated to

Ref
21. James C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms." Plenum Press, New York, 1987.

the majority of the codebook vectors. These sureties the interest of all the codebook vectors in the codebook outline process. As this improvement continues, the cardinality $\aleph(T_k^{(v)})$ of the set T_k diminishes, until T_k will incorporate stand out component. For this situation the k-th preparing vector is exchanged from Fuzzy to crisp mode. For assessment reasons, from the methodologies created in [13], we utilize the FVQ2 calculation, in light of the fact that just this calculation is specifically identified with the compelled minimization of the target capacity given in eq. (3). The usage of the FVQ2 requires that the codebook vectors are redesigned by utilizing eq. (4)· Additionally, in the v-th emphasis, if the preparation vector x_k is in fuzzy mode its participation degrees are figured as,

$$u_{i\,k} = \frac{1}{\sum_{v_j \in T_k^{(v)}} \left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}}}, with\ v_i \in T_k^{(v)} \tag{8}$$

while in the event that x_k is in fresh mode then the participation degrees are given by the following closest neighbor condition,

$$u_{i\,k} = \begin{cases} 1, if\ \|x_k - v_i\|^2 = \min_{1 \le i \le c}\{\|x_k - v_i\|^2\} \\ 0, otherwise \end{cases} \tag{9}$$

Alluding to the last comparison, the utilization of eq. (4) is still legitimate, subsequent to for this situation it holds that $(u_{i\,k})^{\,m} = u_{i\,k}$ notwithstanding the estimation of m. To this end, the FVQ2 calculation comprises on iteratively utilizing the eqs (4), (8) and (9) to figure the codebook vectors and the participation degrees, in blend with the beforehand dissected methodology for the move from fuzzy to crisp decisions. FVQ2 clustering and its minimum objective function $\check{d}_k^{(v)}$=0.1887. Fig. 3 shows the plot of objective function by FVQ2.

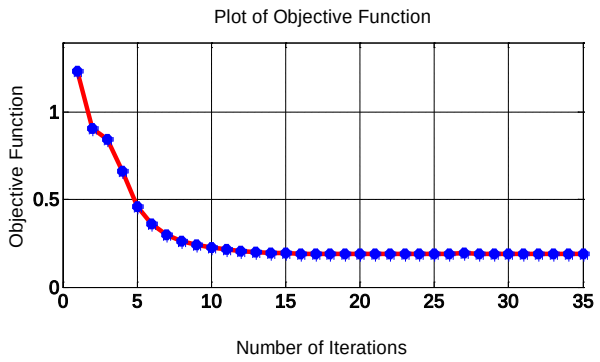

Fig. 3 : Plot of objective function of Fuzzy Vector Quantization2 clustering

Ref
13. Jayanna, H.S., Prasanna, S.R. Mahadeva., "Fuzzy Vector Quantization for Speaker Recognition under Limited Data Conditions," IEEE Region 10 Conference, TENCON 2008, Hyderabad India, Nov 2008, pp.1-4.

Fig. 4 shows the plot of distortion of the speakers using FVQ2 clustering.

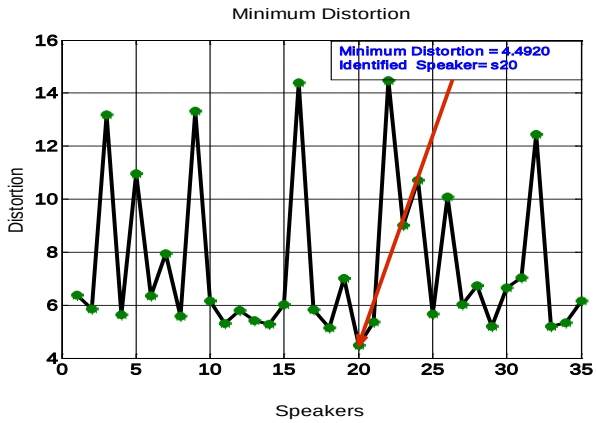

Fig. 4 : Plot of distortion measurement of Fuzzy Vector Quantization2

Distortion of the voice data of the speakers during testing phase using FVQ 2 clustering. D =6.3721, 5.8452, 13.1923, 5.6251, 10.963, 6.3408, 7.938, 5.582, 13.319, 6.163, 5.296, 5.785, 5.412, 5.264, 6.011, 14.384, 5.815, 5.131, 7.005, **4.492**, 5.362, 14.482, 9.007, 10.723, 5.671, 10.07, 6.0151, 6.7238, 5.2056, 6.6534, 7.0366, 12.4455, 5.1927, 5.3277, 6.1519.

III. The Novel Fuzzy Vector Quantization Algorithm

In this section we present a detailed analysis of the NFVQ algorithm. The algorithm is based on the following novel objective function of the fuzzy *c*-means method,

$$J = \sum_{k=1}^{n}\sum_{i=1}^{c} f(u_{i\,k})\|x_k - v_i\|^2 \tag{10}$$

With

$$f(u_{i\,k}) = \frac{1}{2}u_{ik} + \frac{1}{2}(u_{ik})^2 \tag{11}$$

Where u_{ik} is the membership degree of the *k*-th training vector to the *i*-th codebook vector. The objective is to minimize the above function under the following equality constraint,

$$\sum_{i=1}^{c} u_{ik} = 1, \quad \forall\, k \tag{12}$$

The membership degrees and the codebook vector values that solve the above minimization problem are given by the following theorems,

Theorem 1

If v_i are settled then u_{ik} that minimize J in eq (10), under the imperative in eq. (12), is presented as follows,

$$u_{ik} = \frac{c+2}{2}.\frac{1}{\sum_{j=1}^{c}\left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|}\right)^2} - \frac{1}{2} \tag{13}$$

Proof

By eq. (12), the Lagranian of eq.(10) for a only one training vectorx_k is,

$$F(u_{ik},\lambda_k)=\sum_{i=1}^{c} f(u_{ik})\|x_k-v_i\|^2-\lambda_k\left(\sum_{i=1}^{c}u_{ik}-1\right)$$

Notes

The partial derivative of the Lagranian with respect to λ_k is,

$$\frac{\partial F(u_{ik},\lambda_k)}{\partial \lambda_k}=-\left(\sum_{i=1}^{c}u_{ik}-1\right)$$

Equating the above derivative equal to zero and get the eq. (12).

$$\frac{\partial F(u_{ik},\lambda_k)}{\partial u_{ik}}=\frac{\partial f(u_{ik})}{\partial u_{ik}}\|x_k-v_i\|^2-\lambda_k \tag{14}$$

Considering (u_{ik}), in eq. (11), the eq. (14) can be written as,

$$\frac{\partial F(u_{ik},\lambda_k)}{\partial u_{ik}}=\left(\frac{1}{2}+u_{ik}\right)\|x_k-v_i\|^2-\lambda_k \tag{15}$$

Setting the above derivative equal to zero and illuminating as for u_{ik} we get the following equation,

$$u_{ik}=\frac{\lambda_k}{\|x_k-v_i\|^2}-\frac{1}{2} \tag{16}$$

Combining eqs (16) and (12) it follows that,

$$\sum_{j=1}^{c}\left(\frac{\lambda_k}{\left\|x_k-v_j\right\|^2}-\frac{1}{2}\right)=1 \tag{17}$$

Fathoming the last mathematical statement concerning λ_k and substituting into eq. (16) we can without much of a stretch determine the eq. (13). This finishes the confirmation of the theorem 1

Theorem 2

In the event that the u_{ik} are settled, then the cluster centers v_i that minimize J in eq. (11) is given by the following mathematical equation.

$$v_i=\frac{\sum_{k=1}^{n} f(u_{ik})x_k}{\sum_{k=1}^{n} f(u_{ik})} \tag{18}$$

Proof

In perspective of eq. (10), setting the partial derivative $\partial J/\partial v_i$ equal to zero and tackling regarding v_i we can undoubtedly get the eq. (18). This finishes the verification of of theorem 2.

Substituting eq. (13) into the objective function in (10) we can easily obtain the following reformulating function,

$$R_J=R_{J1}+R_{J2} \tag{19}$$

$$R_{J1} = \frac{2+c}{4}\sum_{k=1}^{n}\sum_{i=1}^{c}\left\{\|x_k - v_i\|^2\left[\sum_{j=1}^{c}\frac{\|x_k - v_i\|^2}{\|x_k - v_j\|^2}\right]^{-1}\right\}$$

$$-\frac{1}{4}\sum_{k=1}^{n}\sum_{i=1}^{c}\|x_k - v_i\|^2 \Rightarrow$$

$$R_{J1} = c\frac{2+c}{4}\sum_{k=1}^{n}\left[\sum_{j=1}^{c}\frac{1}{\|x_k - v_j\|^2}\right]^{-1}$$

$$-\frac{1}{4}\sum_{k=1}^{n}\sum_{i=1}^{c}\|x_k - v_i\|^2 \quad (20)$$

Relationally,

$$R_{J2} = \frac{4-c^2}{8}\sum_{k=1}^{n}\left[\sum_{i=1}^{c}\frac{1}{\|x_k - v_i\|^2}\right]^{-1}$$

$$+\frac{1}{8}\sum_{k=1}^{n}\sum_{i=1}^{c}\|x_k - v_i\|^2 \quad (21)$$

Substituting (20) and (21) into (19), the reformulating function is novel as follows,

$$R_J = K_1\sum_{k=1}^{n}\left[\sum_{i=1}^{c}\frac{1}{\|x_k - v_i\|^2}\right]^{-1} - K_2\sum_{k=1}^{n}\sum_{i=1}^{c}\|x_k - v_i\|^2 \quad (22)$$

Where

$$K_1 = \frac{(2+c)^2}{8} \text{ and } K_2 = \frac{1}{8} \quad (23)$$

By minimizing the reformulating function in (22) with respect to the codebook vectors, the gradient-descent based learning rule for the *i*-th codebook vector is given as,

$$v_i(t+1) = v_i(t) - a(t)\sum_{k=1}^{n} f\left(u_{ik}(t)\right)x_k \quad (24)$$

Where $f(u_{ik})$ is given in eqn. (11), and $a(t)$ is the learning rate parameter, which can be calculated as follows,

$$a(t) = a_0\left(1 - \frac{t}{t_{max}}\right) \quad (25)$$

Where a_0 the initial is value for the learning parameter, and t_{max} is the maximum number of iteration. Based on the above analysis, the proposed fuzzy learning vector quantization algorithm for speaker recognition given as follows,

The Novel vector quantization for speaker recognition

Randomly select initial values for the v_i.

Set values for the design parameters t_{max} and a_0.

For $t = 1\ to$ t_{max}

Using eqn (13) calculate the u_{ik}

Using eqn. (11) calculate $f(u_{ik})(1 \le i \le c, 1 \le k \le n)$.

For $i = 1\ to\ c$

Using eqn. (24) to update the codebook vectors.

Endfor

Endfor

End

NFVQ clustering and its minimum objective function J=0.073. Fig. 5 shows the plot of objective function by novel Fuzzy vector quantization.

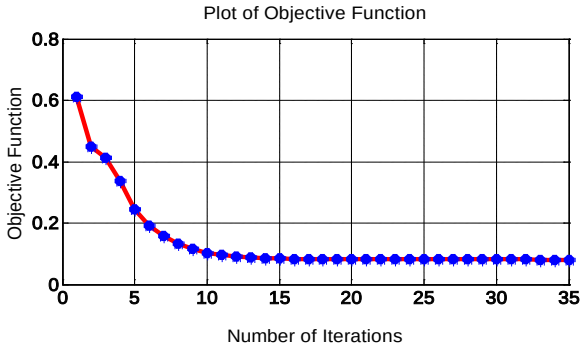

Fig. 5 : Plot of objective function of Novel Fuzzy Vector Quantization

Fig. 6 shows the plot of distortion of the speakers using NFVQ clustering.

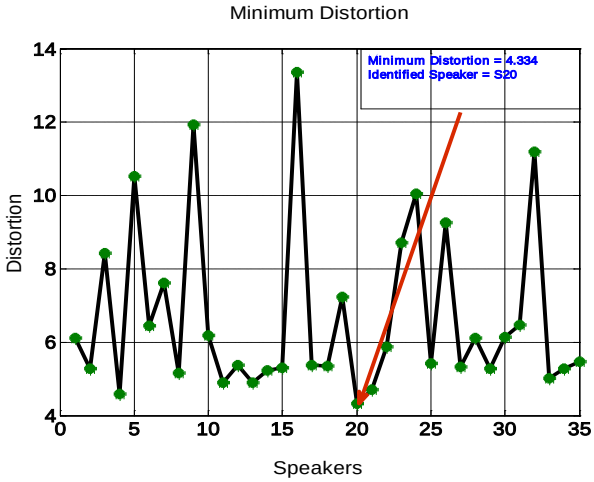

Fig. 6 : Plot of distortion measurement of NFVQ

Distortion of the voice data of the speakers during testing phase using NFVQ. D = 6.1162, 5.2875, 8.4223, 4.5901, 10.5284, 6.4421, 7.6078, 5.1621, 11.9275, 6.1786, 4.8972, 5.3811, 4.8904, 5.2311, 5.3065, 13.3695, 5.3833, 5.3433, 7.2431, **4.334**, 4.7034, 5.8735, 8.7089, 10.0453, 5.4298, 9.2664, 5.3304, 6.1129, 5.2765, 6.1359, 6.4721, 11.205, 5.0118, 5.2669, 5.4778.

In the following section we present the experimental results.

IV. Experimental Results

The proposed algorithm was implemented in MATLAB and results were compared with those of the Fuzzy c-Means and FVQ2 algorithms.

The proposed speaker recognition system efficiency is evaluated with the following design parameters t_{max} = 35 and a_o= 0.5 codebook size = 256. The experiment uses two sets of databases TIMIT and self-collected database.

a) Experimental Result

The performance of MFCC based classifier has been evaluated where each feature set was tested using TF, GF and Tukey Filter. A total 1000 utterances were put to test for 100 speakers. For the above cases, recognition accuracy has been calculated using the expression:

Percentage of Identification Accuracy=No of utterance correctly identified/Total No of utterance under test.

Table I shows the identification accuracies of TIMIT database for TF, GF and Tukey based filters and Fuzzy c-means, FVQ2 and NFVQ techniques respectively. It can be observed from this table that use of GF and NFVQ show significant improvement.

Table 1 : Speaker Recognition of TIMIT Database

Filters	Fuzzy c-Means Accuracy (%)	FVQ 2 Accuracy (%)	NFVQ Accuracy (%)
Triangular Filter	96.9	97	**97.2**
Gaussian Filter	98.1	98.3	**98.8**
Tukey Filter	97.3	97.5	**97.9**

It can be observed from this table that the combination of GF and NFVQ algorithms shows significant improvement up to 98.8%. In a real life situation, a biometric security system, which is usually imperfect, the characteristic curves of FRR and FAR intersect at a certain point called 'Equal Error Rate (EER). If one fixes a very low threshold value, then the system would exhibit very low FRR and very high FAR and accept all identity claims. Alternatively, if one fixes a very high threshold value, then the system would exhibit very high FRR and very low FAR and reject all identity claims. In this context, one could plot a curve called 'Receiver Operating Characteristic (ROC)', which involves FRR and FAR. ROC curve is a graphical indication of the system performance. Fig. 7 shows a typical EER curve.

Notes

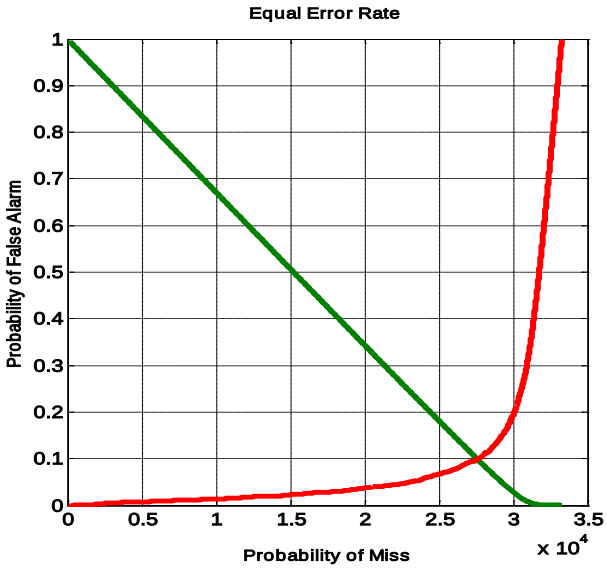

Fig. 7 : Plot of Equql Error Rate

Fig. 8 shows a typical DET curve showing the optimum detection cost for Fuzzy c-means clustering based speaker recognition system. In Fuzzy c-means clustering the EER is 6.5.

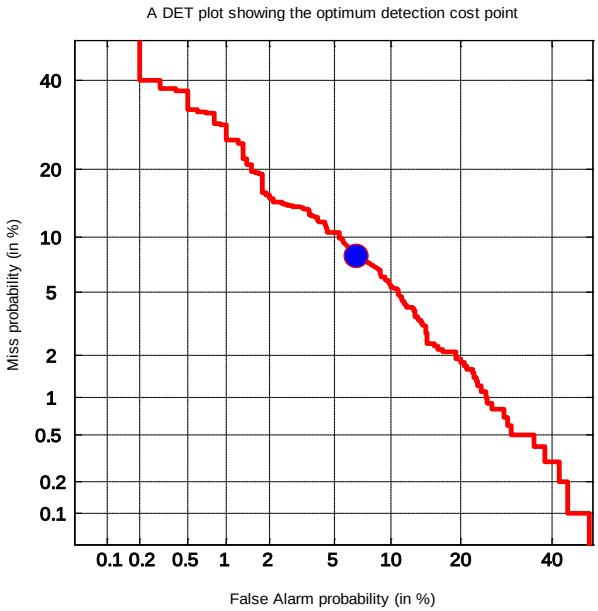

Fig. 8 : Plot of DET showing the optimum detection cost for Fuzzy c-means

Fig. 9 shows a typical DET Curves showing the optimum detection cost for FVQ2 clustering based speaker recognition system and EER is 6.1.

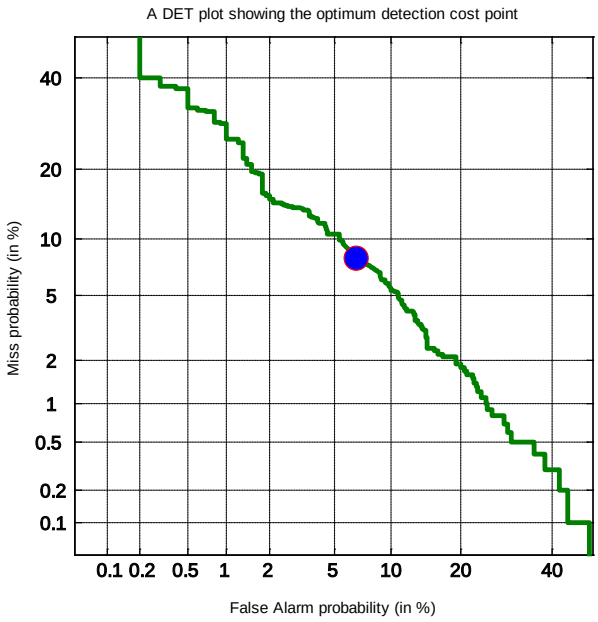

Fig. 9 : Plot of DET showing the optimum detection cost for FVQ2

Fig. 10 shows typical DET curve of optimum detection cost for NFVQ clustering based speaker recognition system and EER is 5.5. In this case the EER is minimum compare to fuzzy c-means and FVQ2. Proposed NFVQ algorithm gives the lower EER that is 5.6, FVQ2 algorithms gives medium EER performance that is 5.9. Finally, the Fuzzy c-means provided the highest EER of 6.1.

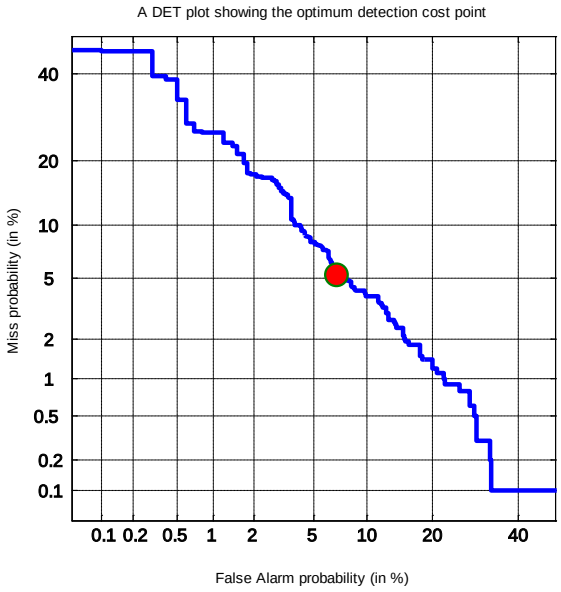

Fig. 10 : Plot of DET showing the optimum detection cost for NFVQ

Fig. 11 shows a typical DET Curves showing the speaker recognition evaluation for NFVQ, FVQ2 and Fuzzy c-means clustering techniques.

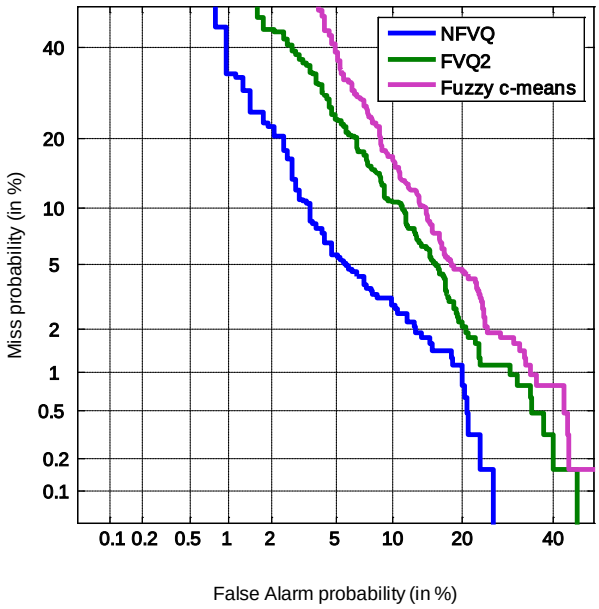

Fig. 11 : Plot of DET Curves for a speaker recognition evaluation

V. Conclusion and Future Work

This paper presented the evaluation of a NFVQ algorithm for speaker recognition. The calculation was intended to catch the favorable circumstances gave by Fuzzy choice making procedures, while keeping up the computational abilities accomplished by fresh crisp making procedures. This was accomplished by developing and reformulating a novel objective function for the well known fuzzy c-means. Several simulations were performed, in which the proposed algorithm was compared to other techniques found in the literature. The objective function is minimized and distortion of the new NFVQ approach is reduced when compared with the objective function and distortion of Fuzzy c-means, and FVQ2. The NFVQ clustering algorithm for speaker recognition is promising as it shows improvement compared to other methods. Equal Error Rate (EER) due to NFVQ is very small when compared to the EER due to Fuzzy c-means clustering and FVQ2 hence NFVQ algorithm for speaker recognition is better than the others. The aftereffect of this examination demonstrates that the calculation can be utilized as a solid instrument as a part of speaker recognition applications. The system performance and speaker recognition efficiency can be further improved by using systematic hierarchical database.

VI. Acknowledgment

The author would like to thank to Dr. B. Yegnanarayana, Director IIIT, Hyderabad for his support in providing TIMIT database for research have resulted in bringing out this original research. The authors acknowledge the support and encouragement received from. Dr. Pragya Singh and Meehika Singh for her invaluable support and encouragement.

Notes

References Références Referencias

1. Premakanthan, P. Mikhael, W.B., "Speaker verification/recognition and the importance of selective feature extraction," Circuits and Systems Proceedings of the 44th IEEE. Vol. 1. Midwest Symposium 2001.
2. Ming, J., Hazen, T.J, Glass, J.R. and Reynolds, D.A., "Robust Speaker Recognition in Noisy Conditions," IEEE Tran. on Audio Speech Lang. Proc., Vol. 15 (5), 2007, pp. 1711 - 1723.
3. Kumerasan, R. and Rao, A., "Model-based approach to envelope and positive instantaneous frequency estimation of signals with speech applications," Journal of Acoustical Society of America, Vol. 105, no 3, Mar. 1999, pp. 1912-1924.
4. Athineos, M. and Ellis, D., "Autoregressive modelling of temporal envelopes," IEEE Tran. Signal Proc., Vol. 55, 2007, pp. 5237-5245.
5. Zhu Jian-wei, Sun Shui-fa, Liu Xiao-li, Lei Bang-Jun, "Pitch in Speaker Recognition." Hybrid Intelligent Systems, 2009, pp. 33-36.
6. Zhu Jian-wei, Sun Shui-fa, Liu Xiao-li, Lei Bang-jun, "Pitch in Speaker Recognition," Hybrid Intelligent Systems, vol. 1, 2009. HIS '09. Ninth International Conference.
7. W. Han, C. Chan., "An efficient MFCC extraction method in speech recognition," IEEE International Symposium on Circuits and Systems, 2006, pp. 145-148.
8. Chakroborty, S., Roy, A. and Saha, G, "Improved Closed set Text- Independent Speaker Identification by Combining MFCC with Evidence from Flipped Filter Bank," International Journal of Signal Processing, 4(2), 2007, pp. 114-122.
9. S. Singh and Dr. E. G. Rajan, "Application Of Different Filters In Mel Frequency Cepstral Coefficients Feature Extraction And Fuzzy Vector Quantization Approach In Speaker Recognition," International Journal of Engineering Research & Technology, Vol. 2 Issue 6, June – 2013.
10. Angtao Cao, Kubota, N. Honghai Liu, "A two-stage pattern matching method for speaker recognition of partner robots," Fuzzy Systems (FUZZ), IEEE International Conference on 18-23 July 2010, PP 1-6.
11. Melek, W.W., Emami, M.R., Goldenberg, A.A., "An improved robust fuzzy clustering algorithm," Fuzzy Systems Conference Proceedings, FUZZ-IEEE '99. IEEE International vol.3, 1999, PP 261 – 1265.
12. Lin Lin, Chen Jian, Sun Xiaoying, "A Novel Fuzzy Kernel Vector Quantization for Speaker Recognition with Short Utterances," Electrical and Control Engineering (ICECE), 25-27 June 2010, PP 194 – 197.
13. Jayanna, H.S., Prasanna, S.R. Mahadeva., "Fuzzy Vector Quantization for Speaker Recognition under Limited Data Conditions," IEEE Region 10 Conference, TENCON 2008, Hyderabad India, Nov 2008, pp.1-4.
14. Daniel Graves, Witold Pedrycz, "Kernel-Based Fuzzy Clustering and Fuzzy Clustering A Comparative Experimental Study," Fuzzy Sets and Systems. vol. 6, no. 6, June 2010, pp. 522-543.
15. Jacek M. Leski, Marian Kotas, "Generalized fuzzy c-means clustering strategies using L p norm distances," Journal Fuzzy Sets and Systems., Vol 279 Issue C, Nov 2015, pp. 112-129.

Notes

16. Li-juan Ma, Ming hu Ha., "Support Vector Machines Based on Sectional set Fuzzy K-Means Clustering," Fuzzy Info. And Engineering, vol. 54, Jan 2009, pp. 420-425.
17. Pinheiro, H.N.B., Tsang Ing Ren, Cavalcanti, G.D.C, Tsang Ing Jyh, "Type-2 Fuzzy GMM-UBM for Text-Independent Speaker Verification," Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on 13-16 Oct. 2013, pp. 4328-4331.
18. Roy. K, Shelton. J, O'Connor. B, Kamel. M. S. "Multibiometric system using fuzzy level set, and genetic and evolutionary feature extraction," IET, Biometrics, Aug 2015, Vol. 4 , Issue. 3, pp. 151-161.
19. I. R. Tsang, D. Gabriel, H. N. B. Pinheiro, and G. D. C. Cavalcanti, "Speaker verification using type-2 fuzzy gaussian mixture models," International Conference on Systems, Man, and Cybernetics, IEEE, oct. 2012, pp.2336-2340.
20. Bhattacharyya, P. Mitra, A. Chatterjee, "Vector quantization based image compression using generalized improved fuzzy clustering." IEEE Conference on Control, Instrumentation, Energy and Communication (CIEC), Jan 2014, pp 662-666.
21. James C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms." Plenum Press, New York, 1987.

Notes

This page is intentionally left blank