

Text Independent Speaker Identification using Integrated Independent Component Analysis with Generalized Gaussian Mixture Model

N M Ramaligeswararao

Department of Electronics &
Communication Engineering,
GIET affiliated to JNTUK,
Rajahmundry, India

Dr.V Sailaja

Department of Electronics &
Communication Engineering,
GIET affiliated to JNTUK,
Rajahmundry, India

Dr.K. Srinivasa Rao

Department of statistics,
Andhra University
Visakhapatnam, India

Abstract - Recently much work has been reported in literature regarding Text Independent speaker identification models. Sailaja et al (2010)[34] has developed a Text Independent speaker identification model assuming that the speech spectra of each individual speaker can be modeled by Mel frequency cepstral coefficient and Generalized Gaussian mixture model. The limitation of this model is the feature vectors (Mel frequency cepstral coefficients) are high in dimension and assumed to be independent. But feature represented by MFCC's are dependent and chopping some of the MFCC's will bring falsification in the model. Hence, in this paper a new and novel Text Independent speaker identification model is developed by integrating MFCC's with Independent component analysis(ICA) for obtaining independency and to achieve low dimensionality in feature vector extraction. Assuming that the new feature vectors follows a Generalized Gaussian Mixture Model (GGMM), the model parameters are estimated by using EM algorithm. A Bayesian classifier is used to identify each speaker. The experimental result with 50 speaker's data base reveals that the proposed procedure outperforms the existing methods.

Keywords - Independent component analysis; Generalized Gaussian Mixture Model; Mel frequency cepstral coefficients; Bayesian classifier; EM algorithm.

I. INTRODUCTION

The growing need for automation in complex work environments and the increased need for voice operated services in many commercial areas have motivated the present work. While many existing systems for speaker identification achieve good performance in relatively constrained environments, performance invariably deteriorates in noisier environment. Speaker identification system is the process of selecting the best matched speaker among the enrolled speakers, with features extracted from speech signals.

Many techniques involving statistical or probabilistic approaches have been applied to speaker specific speech patterns (Leena Mary and Yegnanarayana (2008), Jyoti et al (2011)) [22] [18]. Several methods were employed to separate mixed signals known as 'Blind Source Signals' (BSS) [13]. The term blind refers to the fact that the method of

combination and source signal characteristics are unknown, so BSS permits a wide range of signals as input.

Text independent speaker identification system has many potential applications like security control, telephone banking, information retrieval systems, speech and gender recognition systems, etc. Speaker identification system involves two parts: front-end (feature extractions) and back-end (actual recognition). These system use processed form of speech signals instead of using raw speech signals as it is obtained. This is to reduce the time consumed in identifying the speaker and to make the process easy, by reducing the data stream and exploiting its advantage of being redundant. Computation of cepstral coefficients using preprocessing and feature extraction phases plays a major role in text independent speaker identification systems Ning Wang et al (2010) [24].

Various studies made by Zhu, (1994) and Furui, (1982)[38][28] have also shown that computing cepstral coefficient is the best among all the parameters for any type of speaker recognition. It was proved that the performance of the speech recognizers can improve using cepstral representation of the signals for both clean and noisy speech (Erell, (1993)) [11]. Recently Reynold (1994)[5] have used Mel frequency cepstral coefficients as base line acoustic features for text independent speaker identification and assumed that Mel frequency cepstral coefficients associated with speakers speech spectra follows a Gaussian Mixture Model.

Sailaja, Sriniva Rao and Reddy (2010a, 2010b, 2010c, 2011)[35][33][32][34] have developed and analyzed text independent speaker identification models with Mel frequency coefficients as feature vectors and follow either Doubly truncated multivariate Gaussian mixture model or Generalized Gaussian mixture model. The Generalized Gaussian mixture model will also include Gaussian mixture model as particular case. In all these papers the authors considered only a first few Mel frequency cepstral coefficients and the remaining coefficients are dropped as insignificant due to high dimensionality problems. They have also assumed that the Mel frequency cepstral coefficients of each speaker speech spectra are independent.

But in reality the adjacent Mel frequency cepstral coefficient are correlated. Ignoring the dependences and dropping some of the Mel cepstral coefficients may lead to a serious falsification of model. So, to have a robust model for the speakers speech spectra, it is need a to reduce the dimensionality and avoid dependence among the features this can be done by further processing the Mel frequency cepstral coefficients of each speakers speech spectra with Independent component analysis.(ICA). The integration of ICA with Mel frequency cepstral coefficients will make the feature vector more robust in characterizing the speech spectra of the speaker.

The ICA, tries to express a set of random variables with some noise as linear combination of components that are statistically independent. ICA is the powerful tool available for high dimensional multivariate analysis. Application of this tool to speech analysis results in computational and conceptual simplicity. Independent component analysis is a technique used for text independent speaker identification system. ICA helps in capturing some of the essential features of speech data in many Speaker identification Systems (Hyvarinen, (2001))[10]. ICA techniques can be used to separate mixed signals. The required assumption in applying ICA is that the observed signals are linearly mixed. As general feature vectors for patter recognition are nonlinearly mixed, distinguishing among characteristics is intractable. However, the normomorphic process of cepstrum parameter extraction, which transforms convolutively mixed filters functions into additive ones, ICA is applied to this problem domain and transformed feature vectors are used for the training and testing of a speaker identification system.

Hence, in this paper a text independent speaker identification method is developed and analyzed integrating independent component analysis (ICA) with Mel frequency cepstral coefficient and using Generalized Gaussian mixture model. The procedure for extracting the feature vectors of the speaker speech spectra using Mel frequency cepstral coefficients and independent component analysis is presented. Assuming that the feature vector of each individual speech spectra follows a Generalized Gaussian mixture model, the estimation of the model parameters is done by using the updated equation of EM algorithm. The speaker identification algorithm under Bayesian frame work is also presented. The efficiency of the developed procedure is studied by conducting experimentation with 50 speakers with 12 utterances of locally recorded data base. A comparative study of different earlier models with the proposed (ICA + MFC coefficients) hybridized method is also discussed.

II. FEATURE VECTORS EXTRACTION

In this section we briefly describe the feature vector extraction of each speaker speech spectra. Reliable and efficient smoothing of the frequency response of a human vocal tract is to be obtained as feature vector. Mel frequency cepstral coefficients have gained importance in speaker identification to describe the signal characteristics. According to psychophysical studies human perception of the frequency content of sounds follows a subjectively defined nonlinear scale called the Mel scale [4] [10], . This is defined as

$$f_{mel} = 2595 \log [1 + (f/700)]$$

Where, f_{mel} is the subjective pitch in Mel's corresponding to f , the actual frequency in Hz. This leads to the definition of MFCC, a base acoustic feature for speech and speaker recognition applications. ICA is the liner and a supervised dimensional reduction algorithm. ICA is not necessarily orthogonal and it extracts independent components even with smaller magnitudes. This section gives brief introduction and analysis of ICA algorithms. Text independent speaker identification system is a highly complex model associated with a huge number of feature vectors. Analysis of such models a challenging task. Under such circumstance dimensional reduction of the data becomes a major requirement for obtaining better identification results. This can be achieved by using ICA algorithms.

Each person's voice has distinguishing properties and features which makes them unique. Air stream pumped by the lungs modifies itself to generate desired sequence of sounds every time a person tries to speak. This implies that there exist some differences in the characteristics of speech depending on the changes in the shape of the vocal tract, vibration of the vocal chords and the nasal cavity. Vocal tract can then be considered as a set of filters that change or alter a set of excitation signals. ICA aims at extracting a set of statistically independent vectors from the matrix of training data, the Mel-frequency Cepstral feature vectors derived from the original signal. It tends to find directions of minimum mutual information. It aims at capturing certain Correlations among the frequencies present in the spectral based representation of a speech signal. This is achieved by ICA in the form of linear combinations of basic filter functions specific to each person. Specific sounds are then generated by combining these functions in a statistically independent nature.

The signal \mathbf{X} is used as a proper Mel-Cepstral based representation of the original signal and the data can be observed as a set of multivariate time series resulting from a hidden linear mixing process \mathbf{A} of independent functions \mathbf{S} (Potamitis, (2000))[25], (Hyvarinen, (2001))[14]. Linear combination of such sources or functions can be summarized as (Cardoso, (1996))[4]

$$\mathbf{A}\mathbf{x} = \mathbf{S} \quad (1)$$

The problem of ICA is to determine both the excitation signal \mathbf{s} and the scalars \mathbf{A} and the only known component is the matrix of the MFCC coefficients of the input speech signal. \mathbf{s} can be computed as follows (Hyvarinen, (1997))[16]

$$\mathbf{S} = \mathbf{A} \mathbf{x}^{-1} \quad (2)$$

\mathbf{A} can be computed by consider \mathbf{x} as a vector of observations, where each observation is expressed as a linear combination of independent components. In order to estimate one of the independent components, a linear combination of x_i is chosen such that (Hyvarinen, (1997))[16], (Michael, (2002))[19]

$$\mathbf{y} = \mathbf{W}^T \mathbf{x} \quad (3)$$

with respected to the conditions stated in equation (1) and equation (2), the linear combination represents in equation (3) is a true combination of independent components if w were one of the columns of the inverse of \mathbf{A} . After preprocessing and whitening, the final equation is given as (Michael, (1999)) [24]

$$S \approx \mathbf{y} = \mathbf{W}^T \mathbf{X} = \tilde{\mathbf{W}} \mathbf{P} \mathbf{x} \quad (4)$$

Fast ICA algorithm (Hyvarinen, (1999))[11] is used to estimate w_i which constitutes the rows of \mathbf{W} . Since the components are considered to be statistically independent, the variance between them is high. The following steps are used to estimate w

1. Choose an initial random guess for \mathbf{W}
2. Iterate: $\mathbf{W} \leftarrow E\{xg(W^T X)\} - E\{g'(W^T X)\}W$
3. Normalize: $\mathbf{W} \leftarrow \frac{w}{\|w\|}$
4. If not converged, go back to step 2.

Once w is estimated the final step is to project the signal into the space created by ICA.

New dataset = $\mathbf{W}_{ica} * \text{Mean Adjusted original Data}$,

where, \mathbf{W}_{ica} is the transformation matrix obtained from Fast ICA algorithm.

The multivariate dimensionality reduction techniques (ICA) can be applied to Mel spectral energies (Ding, (2001))[9], or the Mel frequency cepstral coefficient after the feature vectors obtained through MFCC's makes the dimension reduction possible and more efficient (Wanfeng, 2003)[36]. This is because the Mel frequency feature vector characteristics agree with the assumptions made in ICA algorithms (Somervuo, (2003))[30].

The extraction of Feature vectors for speaker identification is done in two steps. :

1. Compute MFCC's, and 2. Apply ICA to transform them to get new feature vectors.

The computation steps for extracting the new feature vectors are as follows.

Step 1.

- (a) Take the Fourier transform of a signal
- (b) Map the powers of the spectrum obtained above on to the Mel scale, using triangular overlapping windows
- (c) Take the logs of the powers at each of the Mel frequencies

- (d) Take the discrete cosine transform of the list of the Mel log powers, as it were a signal
- (e) The MFCCs are the amplitudes of the resulting spectrum.

Step 2.

- (f) Apply ICA transform to Mel frequency cepstral coefficients to get the new feature vectors of each speaker speech spectra

III. SPEAKER IDENTIFICATION MODEL WITH GENERALIZED GAUSSIAN DISTRIBUTION:

In this section we describe the speaker identification process. Fig.1 represents the block diagram of the proposed text independent speaker identification system with Generalized Gaussian distribution using integrating ICA in the system after feature extraction.

Here it is assumed that the feature vector (after processing the MFCC with ICA) follows a multivariate Generalized Gaussian mixture model. The motivation for considering the Generalized Gaussian mixture models is that the individual component densities of a multi model density like the mixture model may model some underlying set of acoustic processes.

It is reasonable to assume the acoustic space corresponding to a speaker voice can be characterized by acoustic classes representing some broad phonetic events such as vowels, nasals or fricatives. These acoustic classes reflect some general speaker dependent vocal tract configurations that are useful for characterizing speaker identity. The spectral shape of its acoustic class can in turn be represented by the mean of its component density and the variation of the average spectral shape can be represented by the co-variance matrix. Therefore the entire speech spectra of the each individual speaker can be characterized as a M component Finite Multivariate Generalized Gaussian mixture distribution.

The probability density function of the each individual speaker speech spectra is

$$p(\vec{x}_t | \lambda) = \sum_{i=1}^M \alpha_i b_i(\vec{x}_t | \lambda) \quad (2.1)$$

where, $\vec{x}_t = (x_{tj})$ $j=1,2,\dots,D$; $i=1,2,3,\dots,M$;
 $t=1,2,3,\dots,T$ is a D dimensional random vector representing the MFCC vector

λ is the parametric set such $\lambda = (\mu, \rho, \Sigma)$

α_i is the component weight such that $\sum_{i=1}^M \alpha_i = 1$

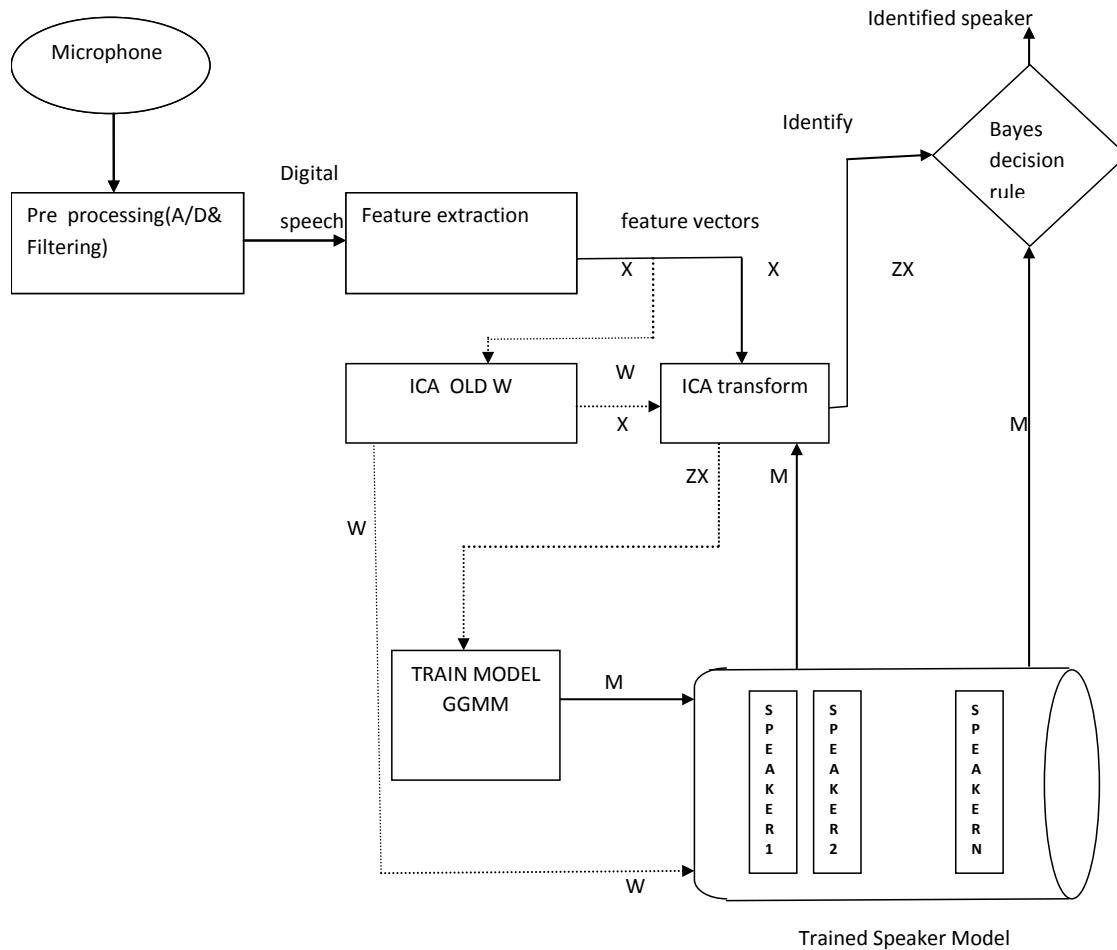


Fig 1. Schematic diagram representing the Speaker Identification process

$b_i(\vec{x}_t|\lambda)$ is the probability density of i^{th} acoustic class represented by new vectors of the speech data and the D -dimensional Generalized Gaussian (GG) distribution (M. Bicego et al (2008))(23) and is of the form

$$b_i(\vec{x}_t|\mu, \rho, \Sigma) = \frac{[\det(\Sigma)]^{-1/2}}{[z(\rho)A(\rho, \sigma)]^D} \exp\left(-\left\|\frac{\Sigma^{-1/2}(\vec{x}_t - \vec{\mu}_i)}{A(\rho, \sigma)}\right\|_{\rho}\right) \quad (2.2)$$

Where, $z(\rho) = \frac{2}{\rho} \Gamma\left(\frac{1}{\rho}\right)$ and

$$A(\rho, \sigma) = \sqrt{\frac{\Gamma(1/\rho)}{\Gamma(3/\rho)}} \quad (2.3)$$

And $\|x\|_{\rho} = \sum_{i=1}^D |x_i|^{\rho}$ stands for the l_{ρ} norm of vector x , Σ is a symmetric positive definite matrix. The parameter $\vec{\mu}_i$ is the mean vector, the function $A(\rho)$ is a scaling factor which allows the $\text{var}(x) = \sigma^2$ and ρ is the shape parameter when $\rho=1$, the Generalized Gaussian corresponds to a laplacian or double exponential Distribution. When $\rho=2$, the Generalized Gaussian corresponds to a Gaussian distribution. In limiting case $\rho \rightarrow +\infty$ Equation (2.2) Converges to a

uniform distribution in $(\mu - \sqrt{3}\sigma, \mu + \sqrt{3}\sigma)$ and when $\rho \rightarrow 0+$, the distribution becomes a degenerate one when $x = \mu$.

The model can have one covariance matrix per a Generalized Gaussian density of the acoustic class of each speaker. Based on the previous studies, the diagonal covariance matrix is used for speaker model. As a result of diagonal covariance matrix for the feature vector, the features are independent and the probability density function of the feature vector is

$$b_i(\vec{x}_t|\lambda) = \prod_{j=1}^D \frac{\exp\left(-\left|\frac{(x_{tj} - \mu_{ij})}{A(\rho_{ij}, \sigma_{ij})}\right|^{\rho_{ij}}\right)}{\frac{2}{\rho_{ij}} \Gamma\left(1 + \frac{1}{\rho_{ij}}\right) A(\rho_{ij}, \sigma_{ij})} = \prod_{j=1}^D f_{ij}(x_{tj})$$

To find the estimate of the model parameters μ_{ij} and σ_{ij} for $i=1,2,3 \dots, M, j=1,2, \dots, D$, we maximize the expected value likelihood (or) log likelihood function. Here the shape parameters ' ρ_{ij} ' is estimated by the procedure given by Armando.J el at (2003) [1] for each acoustic class of each speech spectra.

The updated equations of the parameters for EM algorithm are as given by Sailaja et al (2010) [34] are

The updated equation for estimating α_i is

$$\alpha_i^{(l+1)} = \frac{1}{T} \sum_{t=1}^T \left[\frac{\alpha_i^{(l)} b_i(\vec{x}_t, \lambda^{(l)})}{\sum_{i=1}^M \alpha_i^{(l)} b_i(\vec{x}_t, \lambda^{(l)})} \right]$$

Where $\lambda^{(l)} = (\mu_{ij}^{(l)}, \sigma_{ij}^{(l)})$ are the estimates obtained

The updated equation for estimating μ_{ij} is

$$\mu_{ij}^{(l+1)} = \frac{\sum_{t=1}^T t_1(\vec{x}_t, \lambda^{(l)}) A(N, \rho_{ij}) (x_{tj} - \mu_{ij})}{\sum_{t=1}^T t_1(\vec{x}_t, \lambda^{(l)}) A(N, \rho_{ij})}$$

where, $A(N, \rho_{ij})$ is some function which must be equal to unity for $\rho_i = 2$ and must be equal to $\frac{1}{\rho_{ij} - 1}$ for $\rho_i \neq 1$, in the case of $N=2$, we have also observed that $A(N, \rho_{ij})$ must be an increasing function of ρ_{ij} .

The updated equation for estimating σ_{ij} is

$$\sigma_{ij}^{(l+1)} = \left[\frac{\sum_{t=1}^N t_1(\vec{x}_t, \lambda^{(l)}) \left(\frac{\Gamma(\frac{3}{\rho_{ij}})}{\rho_{ij} \Gamma(\frac{1}{\rho_{ij}})} \right) |x_{tj} - \mu_{ij}^{(l)}|^{\frac{1}{\rho_{ij}}}}{\sum_{t=1}^T t_1(\vec{x}_t, \lambda^{(l)})} \right]^{\frac{1}{\rho_{ij}}}$$

IV. SPEAKER IDENTIFICATION USING BAYES' DECISION RULE

The extracted feature vectors from each test speaker were applied to the function "ICA Transform" (Fig. 1) and were estimated into the space of ICA created by the associated speaker with unique speaker ID. This uses the stored ICA W from the trained speaker model. The new feature vectors of the test utterances and the trained models were fed to a Byes classifier for identification applications which employ large group of data sets and the corresponding test speaker was identified (Domingo's, (1997))[8][9].

$p(i/x_t, \lambda)$ is called a posteriori probability for an acoustic class i and is defined by the following equation

$$p(i/x_t, \lambda) = \frac{p_i b_i(x_t^{\wedge})}{\sum_{k=1}^M p_k b_k(x_t^{\wedge})}$$

For a given observation sequence the main goal is to find the speaker model that has the maximum posteriori probability represented as (Reynolds,(1995))[7]

$$\begin{aligned} \hat{s} &= \max_{1 < k < S} p_l(\lambda_k | X) \\ &= \arg \max_{1 < k < S} [p(\lambda_k | X) p_r(\lambda_k)] \end{aligned}$$

The speaker identification system finally computes S using the logarithms and the independence between the observations.

V. EXPERIMENTAL RESULTS

The performance of the developed model is evaluated by using a database of 50 speakers. For each speaker there are 12 conversations of approximately 4sec.each recorded in 12 separate sessions. Out of which three, four & five sessions are used for training data and the remaining sessions used for testing data. The speaker's speech data was recorded locally by using high quality Microphone. The test speech was first processed by front end analysis explained in section 2 to produce a new transformed feature vectors which are obtained for test sequence length 4 seconds, ith the procedure given by Smitha Gangisetty (2005)[29].

Using the classified data for each speaker the updated equations of the parameters are calculated by the procedure explained in section 3, the refined estimates of the parameters are obtained. The global model for each speaker density is estimated by using the derived parameters. With the test data set, the efficiency of the developed model is studied by identifying the speaker with the Speaker identification algorithm given in section (4). The average percentage of identification was computed and the results are tabulated. Two sets of experiments were carried out for evaluating the performance of the develop model.

EXPERIMENT -1. This experiment involves clean train and test signals.

TABLE 1. AVERAGE PERCENTAGE OF CORRECT IDENTIFICATION VERSUS FOR VARIOUS SPEAKER IDENTIFICATION MODELS

| Model | Percentage of accuracy |
|-----------------------------------|------------------------|
| GMM-nv | 94.6±1.8 |
| GMM-gv | 89.7±2.4 |
| TGMM | 80.2±3.1 |
| GC | 67.3±3.76 |
| FDTMGMM (K-means) | 96.4±1.7 |
| FDTMGMM (Hierarchical clustering) | 97.1±1. |
| Embedded ICA with GMM | 94.12% |
| Embedded ICA with GGMM | 97.8±12. |

A comparative study of the Performance of the developed model is carried with reference to existing speaker modeling techniques. Specially the other techniques are the unimodel Gaussian classifier given by H.Gish (1985)[12], Tied Gaussian Mixture model given by J. Oglesby and J. Mason,(1991)[20] and the Gaussian Mixture Model using nodal variance(GMMnv) and Gaussian Mixture Model using global variance (GMMgv) by Douglas A Reynolds (1994)[5], and Finite Doubly Truncated Gaussian Mixture Model .

From, Table 1, it is observed that the average percentage of correct identification for the developed model is

97.8±12 The percentage correctness for the Gaussian Mixture Model with integrated ICA is 94.12%. This clearly shows that the speaker identification model with multivariate Generalized Gaussian Mixture Model is having higher average percentage of correct identification than the other models.

EXPERIMENT -2. Additive white Gaussian noise (AWGN) affected train signals at a particular SNR of 30dB and AWGN affected test signals with varying SNR's. From the table we found that as the signals to noise ratio increases the identification rate also increases. The performance of speaker identification is improved using ICA with Generalized Gaussian Mixer Model even in noisy conditions.

TABLE 2. PERFORMANCE OF PCA AND ICA WITH VARIATION IN SNR OF THE TEST SIGNALS

| MODEL TRAIN : 30dB SNR | TEST VALUES OF SNR dB | | |
|------------------------|-----------------------|--------|--------|
| | 0 | 10 | 20 |
| ICA with GMM | 51.00% | 70.00% | 85.50% |
| PCA with GMM | 40.33% | 62.41% | 76.70% |
| ICA with GGMM | 55.01% | 72.34% | 88.45% |

VI. CONCLUSION

In this paper a novel Text Independent Speaker Identification model is developed with the assumption that the feature vector associated with the speech spectra of each individual speaker follows a Generalized Gaussian Mixture Model. The Generalized Gaussian Mixture Model also includes the leptokurtic or platykurtic nature of the feature vector associated with each vocal class of individual speaker's speech spectrum. The new feature vectors are derived for the each speaker's speech data through the mel frequency cepstral coefficients + ICA.

This procedure avoids the losing of information on speech spectra, by dropping some of the MFC Coefficients and avoids the dependencies among them. The model parameters are obtained by deriving the updated equations from the EM Algorithm associated with Finite Multivariate Generalized Gaussian Mixture Model. An experimentation with 50 speakers speech data revealed that this Text Independent Speaker Identification with integrating ICA using Finite Multivariate Generalized Gaussian Mixture Model outperforms the earlier existing Text Independent speaker identification models. This method works well even with large population with small time speech data sets.

REFERENCES

- [1] Armando. J et al (2003), "A practical procedure to estimate the shape Parameters in the Generalized Gaussian distribution."
- [2] Ben Gold and Nelson Morgan (2002), "speech and audio processing" part 4 chapter 14 pp.189-203 John willis and sons.
- [3] Cardoso(1996) Automated speech/speaker recognition over Digital wireless 29, pp. 641-662.
- [4] Cardoso, (1996), "Equivariant adaptive source separation" IEEE Transaction on signal processing, Volume.44, no.12, pp.3017-3030.
- [5] D.A. Reynolds (1994), "Experimental evaluation of features for robust speaker identification", IEEE Trans. Speech Audio Process, pp. 639-643.
- [6] D.O. Shoaughnessy (1987), "Speech Communication Human and Machine, wisely publication. NEWYORK.
- [7] Douglas A. Reynolds, and Richard C. Rose (1995), "Robust Text Independent Speaker Identification using Gaussian Mixture Speaker Model," IEEE trans. Speech and Audio Processing, vol.3, pp. 72-83.
- [8] Domingos, (1997), "On the optimality of the simple Bayesian classifier under zero-one loss," Machine Learning, vol. 29, pp. 103-130, 1997
- [9] Domingo(1997), "Speaker Recognition", A Tutorial", Proceedings of the IEEE, Vol. 85, no. 9,
- [10] Ding(2001), "Personal Recognition Using ICA," 8th International Conference on Neural Information Processing,
- [11] Erell and MBH Juang, LR Rabiner, and JG Wilpon, "On the use of bandpass filtering in speech recognition," IEEE Trans. Acoust., Speech, Signal Processing, vol.573-576.
- [12] Furui, (1992), "toward robust speech recognition and adverse conditions" proceeding of the ESCA workshop in speech processing under adverse conditions, PP.31-41.
- [13] H. Gish et al (1985), "Investigation Of Text-dependent Speaker Identification Over Telephone Channels," in Proc. IEEE ICASSP, pp. 379-382.
- [14] Hyvarinen, (2001), "Independent component analysis, John Wiley and sons.
- [15] Hyvarinen,(1999), "Fast and Robust Fixed Point Algorithm for ICA" proceedings of the IEEE transactions on neural networks, Volume.10, no.3, pp.626-634.
- [16] Hyvarinen.(1997) "A family of fixed point algorithms for independent component analysis" proceedings of the IEEE international conference on acoustics speech and signal processing (ICASSP), VOLUME5, pp.3917-3920.
- [17] J.F. Cardoso (1997), "Informax and maximum likelihood for Blind Source Separation", IEEE Signal Processing Letters, Vol.4.
- [18] Jyothi et al (2011), "Text independent speaker identification with finite multivariate generalized Gaussian mixture model with distant microphone speech" proceeding of the international journal of computer applications(IJCA)14(4):5-9,
- [19] Heidelberg (2003), "ICANN/ICONIP'03 of the 2003 joint international conference on Artificial neural networks/neural information processing".
- [20] J.Oglesby and J. Mason (1991), "Radial basis function networks for Speaker Recognition," in Proceedings of IEEE ICASSP, pp. 393-396.
- [21] JP Campbell (1997), "Speaker Recognition: A Tutorial", Proceedings of the IEEE, Vol. 85, no.9.
- [22] Leena Mary and Yegnanarayana(2008), "Extraction and representation of prosodic feature for language and speaker recognition" SPEECH COMMUNICATION 50(10):782-796. Michael Charles (1999), "Orthogonal GMM in Speaker Recognition," Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing", pp. 845-848.,
- [23] Md M. Bicego, D Gonzalez, E Grosso and Alba Castro(2008) "Generalized Gaussian distribution for sequential Data Classification" IEEE Trans. 978-1-4244-2175-6.
- [24] Ning Wang, P. C. Ching(2011), "Nengheng Zheng, and Tan Lee, "Robust Speaker Recognition Using Denoised Vocal Source and Vocal Tract Features speaker verification," IEEE Transaction on Audio Speech and Language processing, Vol. 1, No. 2, pp. 25-35.
- [25] Potamitis, N. Fakotakis and G. Kokkinakis (2000), "Independent component analysis applied to feature extraction for robust automatic speech recognition", IEE Electronic Letters, Vol.36, No.23, pp.1977-1978. Pierre C (1994). "Independent component analysis", A new concept? Signal Processing, 36:pp. 287-314.
- [26] Sandipan Chakraborty, Anindya Roy and Goutam Saha(2006) "Improved Closed Set text-Independent Speaker identification by combining MFCC with Evidence from Flipped Filter banks " IJSp vol 4, no 2
- [27] S. Furui (1986), "Cepstral analysis technique for automatic speaker verification", IEEE Trans. Acoust. Speech Signal Process., Vol. 29, No. 2, pp. 254-272.

- [28] S. Furui(1982), "Speaker recognition by statistical features of Speech Communication S. Furui / Individual features in speech waves /97 cepstral parameters", Trans. IECE, Vol. J65-A, pp. 183-190.
- [29] Smitha Gangisetty (2005), "Text-Independent Speaker Recognition", MS Thesis, College of Engineering and Mineral Resources at Morgantown, West Virginia University.
- [30] Somervuo, chen B and Zhu Q, ,(2003) "Feature Transformations and Combinations for Improving ASR Performance" proceedings of the 8th European conference on speech communication and technology (Euro speech),PP.477-480.
- [31] Somervuo,(2003) " Experiments with Linear and Non Linear Feature Transformations in HHM based phone Recognition" proceedings of the IEEE International Conference on Acoustics, Speech and signal processing (ICASSP), volume.1,PP.52.55.
- [32] V Sailaja, K Srinivasa Rao & K V V S Reddy(2010), "Text Independent Speaker Identification Using Finite Doubly Truncated Gaussian Mixture Model " International Journal of Information Technology and Knowledge Management July-December 2010, Volume 2, No. 2, pp. 475-480
- [33] V Sailaja, K Srinivasa Rao & K V V S Reddy(2010), "Text Independent Speaker Identification Model Using Finite Doubly Truncated Gaussian Distribution And Hierarchical Clustering" International Journal of Computer Science & Communication Vol. 1, No. 2, , pp. 333-338.
- [34] V Sailaja, K Srinivasa Rao & K V V S Reddy(2010), "Text Independent Speaker Identification Model with Finite Multivariate Generalized Gaussian Mixture Model and Hierarchical Clustering Algorithm " International Journal of Computer applications Vol. 11, No. 11, , pp. 25-31.
- [35] V.Sailaja (2010), "Some Studies on Text Independent Speaker Identification Models with Generalizations of Finite Gaussian Mixture Models", unpublished Thesis notes Department of Electronics and Communication Engineering, Andhra University, Visakhapatnam.
- [36] Wanfeng,(2003) "experimental evaluations of a new speaker identification frame work using PCA" IEEE international conference on systems ,man and cybernetics, volume.4147-4152,PP.5-8.
- [37] Z. Wanfeng et al (2003), "Experimental Evaluation of a New Speaker Identification Framework using PCA", IEEE Inter- national Conference on Systems, Man and Cybernetics, Vol. 5, pp.4147-4152.
- [38] Zhu(1994), "Text-independent speaker recognition using VQ, mixture Gaussian VQ Andergodic HMMs", in ASRIV,pp.55-58.

AUTHORS PROFILE

*Mr N.Murali .Ramalingeswararao Completed M.tech in Digital Electronics and Communication Engineering in GIET in Jawaharlal Nehru Technological University, Kakinada, INDIA in 2011.

* Vemuri Sailaja is professor of electronics and communication engineering at GIET. She received his Ph.D instatistical signal processing from Andhra University. She published 10 research papers in referred International and National Journals and She guided 10 M.Tech students. She is a fellow of IETE. She also presents several papers in National and International conferences.

*Dr.K.S.Rao is professor of statistics at Andhra University. He Introduced four new and novel probability distributions namely bimodal distribution, generalized Laplace distribution, four parameters generalized Gaussian distribution and published. He received his Ph.D degrees and he published 88 research papers in refereed international and national journals. He completed 9 research projects sponsored by ministry of industries, Govt of India of AP, UGC and vizag airport. He was the chief editor of Journal of Indian Society for probability and statistics.vol12.