ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

A Methodology for Efficient Gender Dependent Speaker Age and Emotion Identification System

Shivaji J Chaudhari[1], Ramesh M Kagalkar[2]

M.E Final Year Student, Dept. Computer Engineering, Dr. D Y Patil School of Engineering and Technology, Lohegaon, Pune[1]

Asst. Professor Dept. Computer Engineering, Dr. D Y Patil School of Engineering and Technology, Lohegaon, Pune[2]

Abstract: In this paper human unique voice characteristics used to increase the human to computer interaction (HCI) and automation in the machine uses. This system is a combination of training phase and testing phase, training phase includes the training of the system and testing phase work on the to identify the new or existed speaker information based on training. Noise elimination algorithm applied on the new audio file to eliminate the noise from voice signal and help to extract the features easily. The Mel Frequency Cepstral Coefficient (MFCC) features extraction technique used to extract unique features from voice, on the extracted feature Gaussian Mixture Model (GMM) and Dimension reduction technique applied to increase the efficiency of performance, A GMM is a parametric probability density function. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum a Posteriori (MAP) estimation from a well-trained prior model. GMM creates the super vectors as features for a Support Vector Machine (SVM) model for classification of a speaker voice according classified age group like child, young, adult, senior, age group classification help identify the age group and precise age of speaker, reduce the complexity of pattern matching using the SVM classification. Proposed techniques increase the performance and accuracy of system.

Index Terms: Mel Frequency Cepstral Coefficient (MFCC), Gaussian Mixture Model (GMM), support vector machine (SVM), Expectation-Maximization (EM), Maximum a Posteriori (MAP), Hidden Markov Models (HMMs), Suprasegmental Hidden Markov Models (SPHMMs).

I. INTRODUCTION

In HCI many characteristics like a human face, fingerprint, iris, voice used to identify the human and human information, this system focused on the unique human voice each human has a unique voice. The speaker-specific characteristics of the signal can be exploited by listeners and technological applications to describe and classify speakers, based on age, gender, accent, language, emotion or health, gender, age, accent and emotion are some of speaker characteristics being investigated in voice-based speaker classification systems. A major motivation comes from the desire to develop human machine interfaces that are more adaptive and responsive to a user's behavior. There is an increasing need to know not only what an information user conveys but also how it is being conveyed. Given the very importance of emotions in human communication and decision making, for instance, automatic dialog systems with the ability to recognize emotions can respond to callers according to the detected emotional state or they can pass control over to human operators. Interactive voice response (IVR) systems are one of the most mature applications of automatic speech recognition (ASR) today and are widely deployed for customer care and service applications, also In call centers, classification of speakers into age categories is used to perform user-profiling, which is a basis for important applications like market research, targeted advertising, and service customization.

Several speech based age and gender estimation systems were proposed, using and combining different kinds of acoustic features and classification algorithms.

A first noise removal process used to remove the noise from the speaker audio file. The noise removal process is generally called signal denoising. The noise removal process removes the background noise and help to extract the only clear feature from the signal. This system investigates new methods for feature extraction from the speaker's voice and classification to meet this challenge and find the exact outcome. Extracted speech features range from traditional features in speech recognition such as Mel-Frequency Cepstral Coefficients (MFCCs). Many types of feature are extracted using the MFCC extraction like spectral and prosodic features include energy, frequency, duration, pitch of the signal, range, etc. Feature selection, then performed to find a more suitable feature set for building speaker models using the GMM/SVM super vector system for speaker age and gender recognition, GMM create the Model of each speaker for training purpose. SVM classification technique is adopted from state-of-the-art speaker classification in gender and age group and recognition of precise age [1]. Radial Basis Function (RBF) kernel is used, the accuracy is improved compared to using the linear kernel; however, the computation complexity is more sensitive to the feature dimension. Classic

Copyright to IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

dimension reduction methods like Principal Component Analysis (PCA) and linear discriminant analysis (LDA) tend to eliminate the relevant feature information and cannot always be applied without damaging the model's accuracy. Hidden Markov Models (HMMs) and Suprasegmental Hidden Markov Models (SPHMMs) have been used as classifiers in the two-stage recognizer, to identify the hidden state means feature from spectral and prosodic feature to identify the exact characteristics of the speaker.

II. RELATED WORK

Gil Dobry, Ron M. Hecht [1] presents a novel dimension reduction method Principal Component Analysis (PCA) which reduce large dimensions of feature into small dimension and improves the accuracy and the efficiency of speakers age estimation systems based on speech signal. Two different gender male & female based age estimation approaches were implemented, the first age group (Senior, Adult, and Young) classification, and the second, exact age identification using the SVM classification technique.

Hugo Meinedo1, Isabel Trancoso [2] present gender detection is a very useful task for a wide range of usage and application. In the Spoken Language Systems lab of INESC-ID, the Gender Identification module is one of the basic components of our Voice processing system, where it is mainly used for speaker separation, in order to avoid mixing speakers from different genders in the same cluster. Gender information (male, female & children) is also used for building gender-dependent acoustic models for speech recognition. Here the third group classification is children because sometime difficult distinguish the child's voice in two group male and female.

Mohamad Hasan Bahari, Hugo Van Hamme [3] introduces new gender detection and an age estimation approach. To create this strategy, after determining an acoustic model for all speakers of the database, Gaussian mixture weights are extracted and concatenated to create a super vector for each speaker. Then, hybrid architecture of GRNN and WSNMF is developed using the super vectors of the training data set.

Ismail Mohd, Adnan Shahin [4] focused on improving emotion identification performance and accuracy based on a two-stage recognizer that is composed of gender recognizer followed by an emotion recognizer. This work is a gender dependent and speaker-independent emotion recognizer. Both HMMs and SPHMMs have been used as classifiers in the two-stage architecture. Two types of databases are used first is collected database using different emotions sample and second standard Emotional Prosody Speech and Transcripts database.

Chul Min Lee and Shrikanth Narayanan [5] explore the detection of domain-specific emotions using language and discourse information in conjunction with acoustic correlates of emotion in speech signals. The main focus is on detecting emotions (happy or unhappy, angry, sad, neutral, etc) using spoken language data obtained from a call center application and from another source.

Tetsuya Takiguchi and Yasuo Ariki [6] investigate robust feature extraction using kernel PCA instead of DCT, where kernel PCA is applied to the mel scale filter bank output, because the expectation is that kernel PCA will project the main speech element onto low-order features, while noise element onto high order ones. The use of kernel PCA provides high performance and high accuracy.

Michael Feld, Felix Burkhardt and Christian Muller [7] present a GMM/SVM supervector system (GaussianMixture Model combined with Support Vector Machine) for speaker age and gender recognition, a technique that is adopted from state-of-the-art speaker recognition. This emphasizes the need of providing flexible in car dialog that take into account the specific needs and preferences of the respective user (group).

Afzal Hossan, Sheeraz Memon and Mark A Gregory [8] present MFCC feature extraction method is a leading approach for speech feature extraction and current research aims to identify performance enhancements. One of the recent MFCC implementations is the Delta-Delta MFCC, which improves speaker verification. In this paper, a new MFCC feature extraction method based on distributed Discrete Cosine Transform (DCT-II) is presented. Speaker verification tests are proposed based on three different feature extraction methods including: conventional MFCC, Delta-Delta MFCC and distributed DCT-II based Delta-Delta MFCC with a Gaussian Mixture Model (GMM) classifier.

III. PROPOSED SYSTEM

A. System Architecture

System Architecture divided into two phase that is
1. Training phase
2. Testing phase

Most of the operations are same in the training phase and testing phase

B. Feature Extraction

Prior to the MFCCs, Linear Prediction Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) used for feature extraction and were the main feature type for automatic speech recognition (ASR). Step by step MFCC procedure.

1. Divide the signal into short frames.
2. For each frame calculate the periodogram estimate of the power spectrum.
3. Apply the mel filter bank to the power spectra, sum the energy in each filter.
4. Take the logarithm of all filter bank energies. Take the log of each of the 26 energies from step
5. Take the Discrete Cosine Transform (DCT) of the 26 log filter bank energies to give 26 cepstral coefficients. For ASR, only the lower 12-13 of the 26 coefficients are kept.
6. Keep DCT coefficients 2-13, discard the rest.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

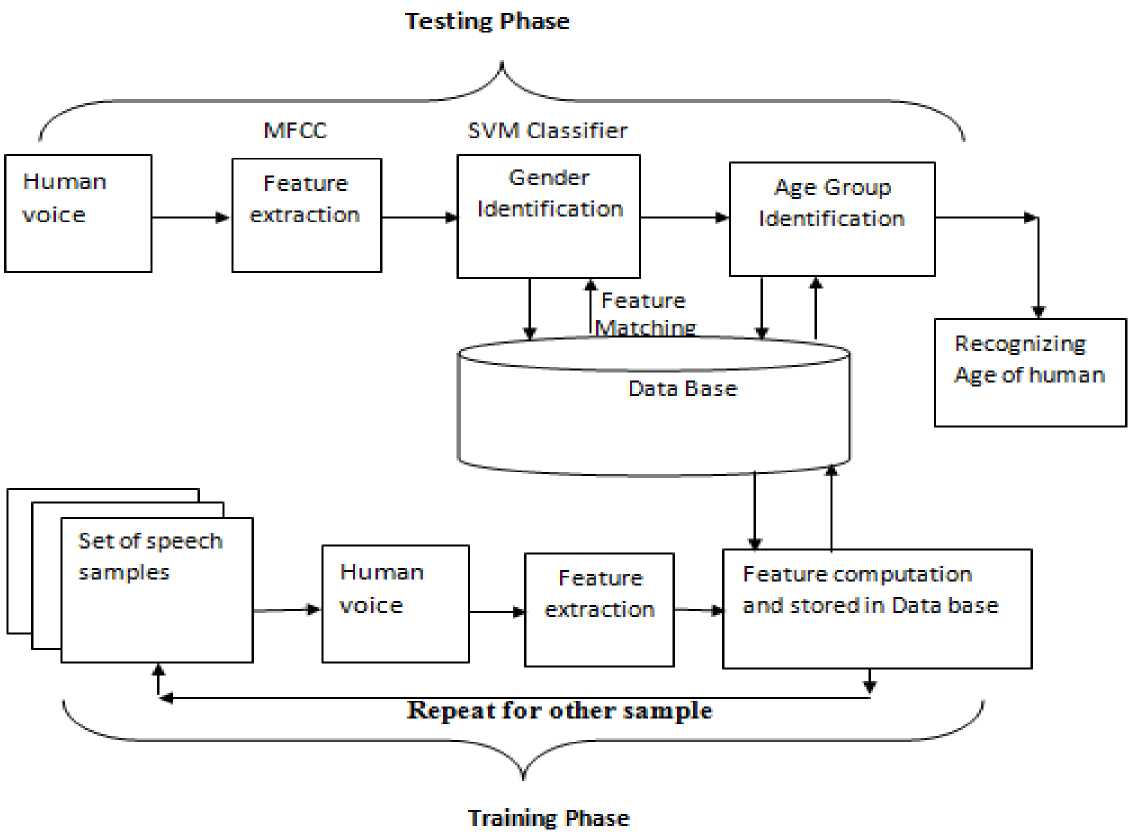

Fig 1. System Architecture Block Diagram

An audio signal is constantly changing nature. Assume that on short time scales the audio signal doesn't change much. Frame the signal into 20-40ms frames. If the frame is much shorter we don't have enough samples to get a reliable spectral estimate [6][14]. The next step is calculating the power spectrum of each frame, identifying which frequencies are present in which frame. Each frame has to be multiplied with a hamming window in order to keep the continuity of the first and the last points in the frame [14].

The final step is to compute the DCT of the log filter bank energies. There are 2 main reasons this is performed. Because our filter banks are all overlapping, the filter bank energies are quite correlated with each other. The DCT decorrelates the energies which mean diagonal covariance matrices can be used to model the features.

The Mel scale relates perceived frequency, energy or pitch, of a pure tone to its actual measured frequency. The formula for converting from frequency to Mel scale is,

$$M(f)=1125\ \ln(1+f/500) \qquad \text{.......(3.2.1)}$$

To go from Mels back to frequency,

$$M^{-1(m)}=700(\exp(m/1125)-1) \qquad \text{........(3.2.2)}$$

Delta cepstrum take the time derivatives of (energy + MFCC) as new features, shows the velocity and acceleration (energy + MFCC). MFCC work at both testing side and training, on extracted feature Gaussian Mixture Model (GMM) applied to create the model for each type of speaker.

C. Gaussian Mixture Modeling

GMM is a parametric probability density function used as a weighted sum of Gaussian component densities. GMM are commonly used as a parametric model of the probability distribution of continuous measurements or features. A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by the equation,

$$P\ (X|\lambda) = \sum^{m}{}_{i=1} W_i\ g(x|u_{i,}\ \Sigma_i) \qquad \text{............ (3.3.1)}$$

Where x is a D-dimensional continuous-valued data vector (i.e. measurement or features),

W_i, i = 1. . . M, are the mixture weights, and

g (x|ui, ∑i) , i = 1, . . . ,M, are the component Gaussian densities with mean vector μi and covariance matrix i. The mixture weights satisfy the constraint that $\sum^{m}{}_{i=1.}$ The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. Using the features of training vectors and a GMM configuration, we wish to estimate the parameters of the GMM, λ, which in some sense best matches the distribution of the training feature vectors. There are several techniques available for estimating the parameters of a GMM. By far the most popular and well-established method is maximum likelihood (ML) estimation. The aim of ML estimation is to find the model parameters which maximize the likelihood of the GMM given the training data. For a sequence of T training vectors X = {x1. . . xT }, the GMM likelihood, assuming independence between the vectors1, can be written as,

$$P(X|\lambda) = \pi^{T}{}_{t=1}\ P(x_t|\lambda) \qquad \text{....... (3.3.2)}$$

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

ML parameter estimates can be obtained iteratively using a special case of the expectation-maximization (EM) algorithm. The basic idea of the EM algorithm is, beginning with an initial model _, to estimate a new model λ, such that p (X| λ) >= p(X| λ). The new model then becomes the initial model for the next iteration and the process is repeated until some convergence threshold is reached.

Maximum Posteriori (MAP) estimation. MAP estimation is used, for example, in speaker recognition applications to derive speaker model by adapting from a universal background model (UBM). It is also used in other pattern recognition tasks where limited labeled training data is used to adapt a prior, general model.

D. Feature Classification

A support vector machine is a powerful technique for pattern classification according to the each group speaker feature. SVM map feature vector into a high dimensional space and then separate feature classes with a hyper plane. A critical aspect of using SVMs successfully is the design of the inner product, the kernel, induced by the high dimensional mapping. We consider the application of SVM to speaker and language recognition and classification. An SVM is a discriminative classifier, it models the boundary between different speaker feature using hyper plane, for example, a speaker and a set of impostors. This approach contrasts to traditional methods for speaker recognition which separately model the probability distributions of the speaker and the general population, by exploring SVM methods.

$$f(x) = \sum^{n}_{k=}1a_it_iK(x;x_i) + d\ldots\ldots\ldots \quad \ldots\ldots..(3.4.1)$$

The vectors xi is support vectors and obtained from the training set by an optimization process. The ideal outputs are either 1 or -1, depending upon whether the corresponding support vector is in class 0 or class 1, respectively. For classification, a class decision is based upon whether the value f(x), is above or below a threshold.

$$D = \{(x_i,y_i)|xi \in R^p\, y_i(1,-1)\}^n_{i=1}\ldots\ldots\ldots\ldots(3.4.1)$$

Where the y_i is either 1 or −1, indicating the class to which the point X_i belongs. Each X_i is a - dimensional vector. We want to find the maximum-margin hyper plane that divides the points having y_i = 1 from those having y_i = -1.

Any hyper plane can be written as the set of points X satisfying maximum margin hyper plane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors [14].

$$W \,.\, x - b = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots..(3.4.2)$$

And

$$W \,.\, x - b = -1 \quad \ldots\ldots\ldots\ldots\ldots\ldots..(3.4.3)$$

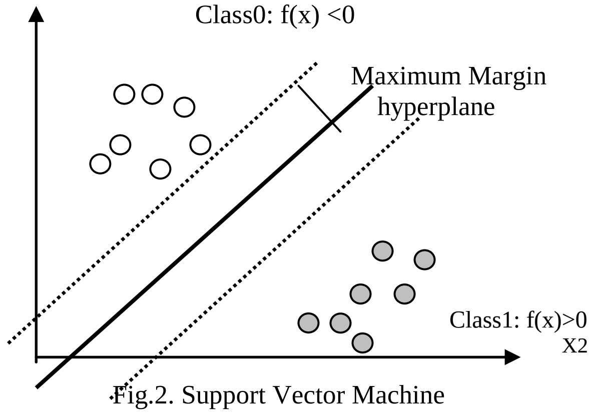

Fig.2. Support Vector Machine

Equation 3.4.2 and 3.4.3 satisfy the condition and classify the feature into two classes class0 and class 1. SVM classifer used the a hyperplane or set of hyperplanes in a high or infinite dimensional space, which can be used for classification, regression, or other tasks, Maximum Margin hyper plane increase the feature or record classification grater as compare to other classifier.

IV. DATA TABLES AND RESULT ANALYSIS

In the [1] previous system used the voice or speech data used to train the system model was taken from the LDC's Switchboard corpus annotated with age and gender labels. Proposed system use audio/voice data from two sources one is using the internet and second is the collected audio/voice dataset in .wav format because of quality issue. Proposed system used the classification technique for Data classification, according to the age group and emotion. While training the voice processing system Training dataset is divided into seven group notations are C, T, M_Y, M_A , M_S ,F_Y , F_A , and F_S in the following Table .I.

TABLE I Classification of Speaker's Dataset

Classified Dataset Name	**Age Range (Year)**	**Notation**
Child	0-08	C
Teenage	09-17	T
Male young	18-30	M_Y
Male Adult	30-60	M_A
Male Senior	>60	M_S
Female young	18-30	F_Y
Female Adult	30-60	F_A
Female Senior	>60	F_S

Classifications of training dataset's, shown in Table I, were selected such that there is no speaker overlap between them. Collect the audio file from different above mentioned age group type and in each emotion and train the system on well classified database to increase the performance and accuracy.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

Table II Result for Speaker Dependent sample

Notation	Classified Type	Training	Testing	Overall Result
C	Child	8	8	80
TA	Teen Age	8	8	88
MY	Male Young	10	10	100
MA	Male Adult	10	10	100
MS	Male Senior	10	10	100
FY	Female young	10	10	100
FA	Female Adult	10	10	100
FS	Female Senior	10	10	100

The database is collected from different source like an audio clip from movie, news channel, some dialog, and recorded audio clip[2][4] [3].

Emotion identification and age identification are gender dependent; two recognizers are used that is composed of gender recognizer followed by an emotion recognizer. And two databases, one is the collected database and second is emotional prosody speech and Transcripts database. For each type of emotion training dataset is collected from different source and also from the emotional prosody speech dataset [4].

MFCC plays important role in proposed system; feature of training data is extracted using the feature extraction technique. MFCC is a highly efficient algorithm for feature extraction, here provide the noise free small voice file to feature extraction and for GMM modeling with dimension reduction, because large file increase the processing time and reduce the system performance Figure show file size effect on feature extraction and other processes.

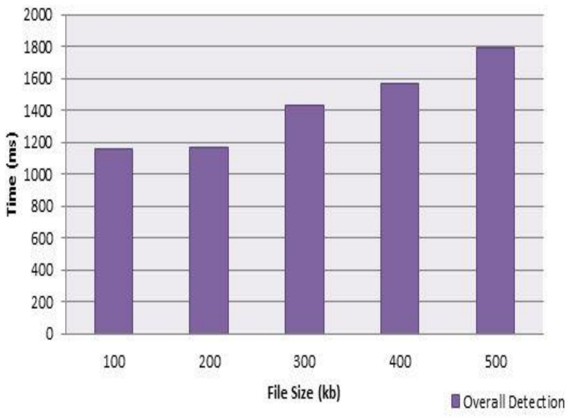

Fig.3. Effect of the file in overall system.

Fig.3. Shows the effect of the file size on overall system, file size impact at each steps feature extraction, GMM model and classification, because unnecessary extra audio file decrease the performance and increase the complexity of system. Table.2. Show the overall system results for speaker dependent for each group 8 for child and teenage 10 for other group samples audio files are trained the system, stored to database. In testing phase test the audio sample to identify the age group and precise age and emotion using train database.

V. CONCLUSION

The techniques implemented on different kind's data type as discussed data table, an age-group classifier and a precise age estimator by SVM classifier and regression. The results taken on multiple data sets of speech and different languages and different and the number of voice files. Gender dependency increases the classification result of male and female speaker. SVM classifier classifies the age group speaker and identifies the age group and precise age of speaker using regression technique. System gives the greater result accuracy and efficiency of system performance, average result accuracy of system is 96%.

ACKNOWLEDGMENT

The authors would like to thank the researchers as well as publishers for making their resources available and teachers for their guidance. We are thankful to the **Ramesh M Kagalkar** for his valuable guidance and constant guidelines also thank full the computer department staff of DYPSOET, lohegoan, Pune and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

REFERENCES

[1] Gil Dobry, Ron M. Hecht, Mireille Avigal, and Yaniv Zigel, "Supervector Dimension Reduction for Efficient Speaker Age Estimation Based on the Acoustic Speech Signal," IEEE transction VOL. 19, NO. 7, SEPTEMBER 2011.
[2] Hugo Meinedo1, Isabel Trancoso, "Age and Gender Classification using Fusion of Acoustic and Prosodic Features," Spoken Language Systems Lab, INESC-ID Lisboa, Portugal, Instituto Superior Tecnico, Lisboa, Portugal,2008.
[3] Ismail Mohd Adnan Shahin, "Gender-dependent emotion recognition based on HMMs and SPHMMs",Int J Speech Technol, Springer 16:133141, 2013.
[4] Mohamad Hasan Bahari, Hugo Van hamme, "Speaker Age Estimation and Gender Detection Based on Supervised Non-Negative Matrix Factorization,"Centre for Processing Speech and Images Belgium, ITN-2008.
[5] Chul Min Lee, and Shrikanth S. Narayanan, "Toward Detecting Emotions in Spoken Dialogs," IEEE transaction 1063-6676, 2005.
[6] Tetsuya Takiguchi and Yasuo Ariki, "Robust Feature Extraction Using Kernel PCA," Department of Computer and System Engineering Kobe University, Japan, ICASSP 1-4244-0469 2006.
[7] Michael Feld, Felix Burkhardt, Christian Muller, "Automatic Speaker Age and Gender Recognition in the Car for Tailoring Dialog and Mobile Services," German Research Center for Artificial Intelligence, INTERSPEECH 2010.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

[8] Md Afzal Hossan, Sheeraz Memon, Mark A Gregory, "A Novel Approach for MFCC Feature Extraction," RMIT University, Melbourne, Australia, IEEE, 2010.
[9] Shivaji Chaudhari and Ramesh Kagalkar "A Review of Automatic Speaker recognization and Identifying Speaker Emotion Using Voice Signal" International Journal of Science and Research (IJSR) Volume 3, Issue 11 November 2014.
[10] Shivaji Chaudhari and Ramesh Kagalkar " Automatic Speaker Age Estimation and Gender Dependent Emotion Recognition", International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 17, May 2015.
[11] Ruben Solera-Ure na, "Real-time Robust Automatic Speech Recognition Using Compact Support Vector Machines," TEC 2008-06382 and TEC 2008-02473, 2008.
[12] Marc Ferras, Cheung-Chi Leung, Claude Barras, and Jean-Luc Gauvain, "Comparison of Speaker Adaptation Methods as Feature Extraction for SVM-Based Speaker Recognition," IEEE Transaction 1558-7916, 2010
[13] Wei HAN, Cheong-Fat CHAN, "An Efficient MFCC Extraction Method in Speech Recognition," Department of Electronic Engineering, The Chinese University of Hong Kong Hong Kong, 7803-9390-06/IEEE ISCAS-2006.
[14] Arif Ullah Khan, L. P. Bhaiya "Text Dependent Method for Person Identification through Voice Segment",ISSN- 2277-1956 IJECSE 2008.
[15] Felix Burkhardt, Martin Eckert, Wiebke Johannsen, Joachim Stegmann, "A Database of Age and Gender Annotated Telephone Speech," Deutsche Telekom AG Laboratories, Ernst-Reuter-Platz 7, 10587 Berlin, Germany, 2010.
[16] Lingli Yu, Kaijun Zhou, "A Comparative Study on Support Vector Machines Classifiers for Emotional Speech Recognition," Immune Computation (IC) Volume2, Number1, March 2014.
[17] Rui Martins, Isabel Trancoso, Alberto Abad, Hugo Meinedo, "Detection of Childrens Voices," Intituto Superior Tecnico, Lisboa, Portugal INESC-ID Lisboa, Portugal 2009.
[18] Chao Gao, Guruprasad Saikumar, Amit Srivastava, Premkumar Natarajan, "Open-Set Speaker Identification in Broadcast News," IEEE 978-1-4577-0539, 2011.
[19] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955.
[20] Manan Vyas "A Gaussian Mixture Model Based Speech Recgnition System Using Matlab", Signal & Image Processing : An International Journal (SIPIJ) Vol.4, No.4, August 2013
[21] Parwinder Pal Singh, Pushpa Rani "An Approach to Extract Feature using MFCC", IOSR Journal of Engineering (IOSRJEN) Vol. 04, Issue 08.2014.
[22] Marc Ferras, Cheung-Chi Leung, Claude Barras, and Jean-Luc Gauvain, 2010 "Comparison of Speaker Adaptation Methods as Feature Extraction for SVMBased Speaker Recognition",IEEE Transaction 1558- 7916.
[23] Wei HAN and Cheong fat CHAN, 2006 "An Efficient MFCC Extraction Method in Speech Recognition",Department of Electronic Engineering, The Chinese University of Hong Kong Hong Kong, 7803- 9390-06/IEEE ISCAS.
[24] Arif Ullah Khan and L. P. Bhaiya, 2008 "Text Dependent Method for Person Identification through Voice Segment",ISSN- 2277-1956 IJECSE. International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 17, May 2015 10
[25] Felix Burkhardt, Martin Eckert, Wiebke Johannsen and Joachim Stegmann, 2010"A Database of Age and Gender Annotated Telephone Speech", Deutsche Telekom AG Laboratories, Ernst-Reuter-Platz 7, 10587 Berlin, Germany.
[26] Lingli Yu and Kaijun Zhou, March 2014, "A Comparative Study on Support Vector Machines Classifiers for Emotional Speech Recognition", Immune Computation (IC) Volume:2, Number:1.
[27] Rui Martins, Isabel Trancoso, Alberto Abad and Hugo Meinedo, 2009, "Detection of Childrens Voices", Intituto Superior Tecnico, Lisboa, Portugal INESC-ID Lisboa, Portugal.
[28] Chao Gao, Guruprasad Saikumar, Amit Srivastava and Premkumar Natarajan, 2011, "OpenSet Speaker Identification in Broadcast News", IEEE 978-1-4577- 0539.

BIOGRAPHIES

Shivaji J. Chaudhari is a PG Scholar of ME Computer Network at Dr. D Y Patil School of Engineering and Technology, Lohegaon, Pune; He received B.E. in Information Technology from SVPM COE Malegaon, Baramati, and Pune University. Currently He is pursuing M.E. in Computer Network from Dr. D. Y. Patil School of Engineering & Technology, Pune, and University of Pune. He has published 3 research papers in International Journals and presented one paper in national conferences.

Ramesh M. Kagalkar was born on June 1st, 1979 in Karnataka, India and presently working as an Assistant. Professor, Department of Computer Engineering, Dr. D.Y. Patil School Of Engineering and Technology, Charoli, B.K.Via –Lohegaon, Pune, Maharashtra, India. He has 13.5 years of teaching experience at various institutions. He is a Research Scholar in Visveswaraiah Technological University , Belgaum, He had obtained M. Tech (CSE) Degree in 2006 from VTU Belgaum and He received BE (CSE) Degree in 2001 from Gulbarga University, Gulbarga. He is the author of text book Advance Computer Architecture which cover the syllabus of final year computer science and engineering, Visveswaraiah Technological University, Belgaum. He is waiting for submission of two research articles for patent right. He has published more than 18 research papers in International Journals and presented few of there in international conferences. His main research interest includes Image processing, Gesture recognition, speech processing, voice to sign language and CBIR. Under his guidance four ME students awarded degree in SPPU, Pune, five students at the edge of completion their ME final dissertation reports and two students started are started new research work and they have publish their research papers on International Journals and International conference.