New Feature Vector Based Gender Identification System Using GMM –

Dr.B.Chandrasekaram
Associate Professor
Department of Computer Science
National Sanskrit University
Tirupati

Abstract:

In this paper a new feature vector for Gender Identification system using Guassian mixture model(GMM) is explored. From the literature it is found that In first using MFCC new type of feature vectors are created using a GMM probability density function. The conventional GMM based gender identification system require large amount of training data to capture all the gender discrimination information present in the speech signal. Where as the new feature vectors based gender Identification system require less amount training data. The performance of new feature vector based gender identification system is compared with the conventional MFCC based Gender identification system. From the analysis it is found that the performance of new feature vector based Gender identification system is outperform than the MFCC GMM based Gender identification system.

Keywords: Gender identification system, GMM, MFCC.

DOI Number: 10.14704/nq.2022.20.11.NQ66107 **NeuroQuantology 2022; 20(11): 1117-1121**

1.Introduction

The language identification, speaker identification, emotion recognition, gender identification applications has increased recent years using speech signal properties. The task of gender identification system is identifying the person is male or female using short duration of speech. Humans identify the gender of persons easily by listening few seconds of speech, but it is tough task for the machines. The speech contains the information about the speaker such as speaker identity,gender, age, accent, emotional state and health state. The extraction of this information from speech signal have lot of useful applications such as person identification in banking system, voice bots such as health assistance voice bots, customer care applications in call centers, human computer interaction and intelligent voice assistants.

The researchers explored the different gender identification cues are physiological (vocal tract length), phonetic, pitch and voice quality differences. From this study, they found that the human perception of pitch location and fundamental frequency (f0) are most important features for gender identification. From the recent studies it is observed that some attempts are made to explore the represent the alternative from of feature vectors. In this work we also made an attempt to explore the new form of feature vectors based on the GMM feature extraction. The new form feature vectors are robust against the noise [1].

The paper is organized as follows, first section describes the transformation of MFCC feature vectors into new feature vectors. The second section explains the creation gender GMM model using this new feature vector to training the Gender identification system. Followed by the thirdsection discussion about the performance analysis of GMM gender identification system.

eISSN 1303-5150

2.New feature Vector for Gender Identification

In this paper the robust new form of feature vectors are extracted from the speech data using GMM. The gender discrimination information present in the speech signal is represented as a Guassians. From the speech signal 12 dimensional MFCC feature vectors are extracted with window size of 20ms and window shift of 4ms. These extracted MFCC features are distributed into R clustered as shown in the fig.1. Here we considered R=15 and each clustered is represented as a one Gaussian[2][3].

Fig.1: R Gaussians for speaker S_i.

New each training feature vector X is passed through this Gaussians G_1(clusters) by producing the probability P_1 using the Gaussian probability density function. Now this probability P_1form the first coefficient in the new feature vector [4]. In this way the feature vector X pass through 15 Gaussians to produce 15 dimensional feature vector as shown in the fig.2.

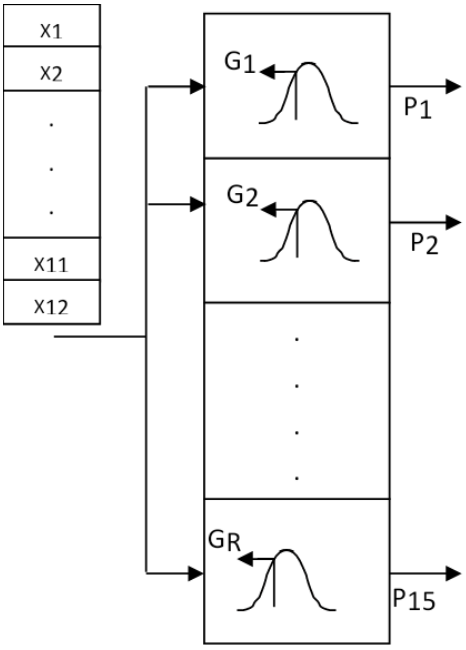

Fig.2: Each coefficients estimation for new feature vector P of dimension 15.

The all the feature vectors are passed through 15 Gaussians (G_1, G_2,……., G_{15}) by producing new 15 dimensional feature vectors. In other way, the training 12 dimensional MFCC feature vector of size N is transformed into new form of 15 dimensional feature vector of size Z. The 12 dimensional MFCC feature vector is transformed into 15 dimensional feature vector as shown in fig.3.

eISSN 1303-5150

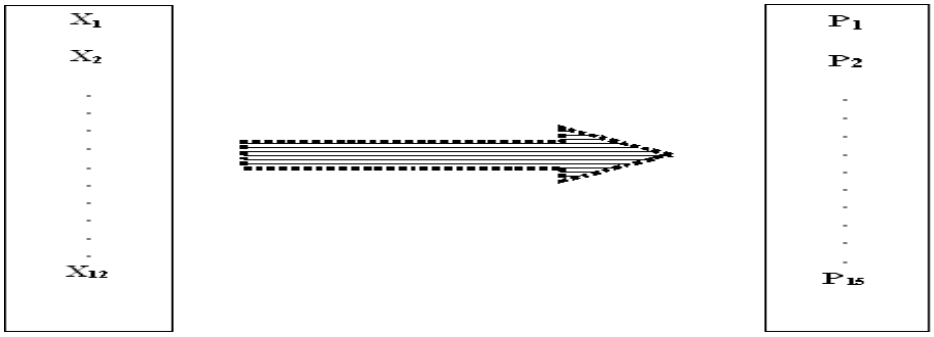

Fig.3: Converting 12dimensional MFCC feature vector into 15 dimensional feature vector.

3.Gender Identification Using GMM

Gaussian mixture model (GMM) is a best parametric method to model the gender identification feature densities effectively, because Gaussians have the capability of representing gender specific information effectively. From the literature it is observed that the GMM classifier is effective used in many gender identification applications. The GMM is probability distribution of the observed data is given by $p(\bar{x} \mid \lambda) = \sum_{i=1}^{M} p_i b_i(\bar{x})$

Where M represent number of mixture components, $\bar{x}$ observed data of D dimensional p_i indicate the mixture weight for M mixture components and $b_i(\bar{x})$ represents component density given by

$$b_i(\bar{x}) = \frac{1}{(2\pi)^{D/2}|\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\bar{x}-\bar{\mu}_i)'\Sigma_i^{-1}(\bar{x}-\bar{\mu}_i)\right\}$$

Here each component density $b_i(\bar{x})$ is a D dimensional normal distribution having with mean vector $\bar{\mu}_i$ and covariance matrix $\sum_i$. This parameter is represented as model $\lambda = \{p_i, \bar{\mu}_i, \Sigma_i\}$ for i = 1, …M . The each speaker is represented by a unique GMM and referred as speaker model for gender identification system [5][6].

3.1 Training the GMM model Using New feature Vectors

Using the new feature vector the clusters are formatted. Each cluster represented as a multiple probability distribution function (pdf) and combination of such many Gaussia probability distribution functions indicate one GMM model. Now we have to estimate the parameters of this GMM model λ using a iterative approach called expected maximization (EM) algorithm. The EM algorithm start with the initial model λ and tend to calculate the new model $\bar{\lambda}$ such that $p(X \mid \bar{\lambda}) \geq p(X \mid \lambda)$. In the next iteration, the new model is the initial model and entire process is repeated until the certain threshold is reached [7] [8].

The training stage of new feature vector based gender identification system is shown in fig.4. For this study the TIMIT speech corpus is used. The TIMIT speech corpus contain total of 6300 sentences and 10 sentences are spoken by each of 630 speakers. For this study we are considering 100 male speaker and 100 female speakers. In this first the 13 dimensional MFCC feature vectors are extracted with frame size 25ms and frame shift 15ms from training speech data. Next this MFCC feature vectors are converted into 15 dimensional new feature vectors as explained previously. Next this speaker new feature vectors are modeled

eISSN 1303-5150

using GMM to create the gender identification system. For every speaker one GMM model is created and is represented S_i. In this way for each speaker under consideration one GMM model is created.

In the testing stage, from the speaker test sample, MFCC feature vectors are extracted. Using this MFCC feature vectors new feature vectors are extracted. Next with this new feature vectors, the likelihood values are calculated against the each speaker GMM model for gender identification purpose. Which model yields the maximum likelihood value is identified the type of gender of that speaker.

3.2 Experimentation and Discussion

The new feature vector based gender identification system is developed using Python and trained using 10sec , 20sec and 30 sec speech data for the varying Gaussian mixtures such as 8, 16, 32 and 64. For testing purpose 1sec, 2sec and 3sec speech samples are used.

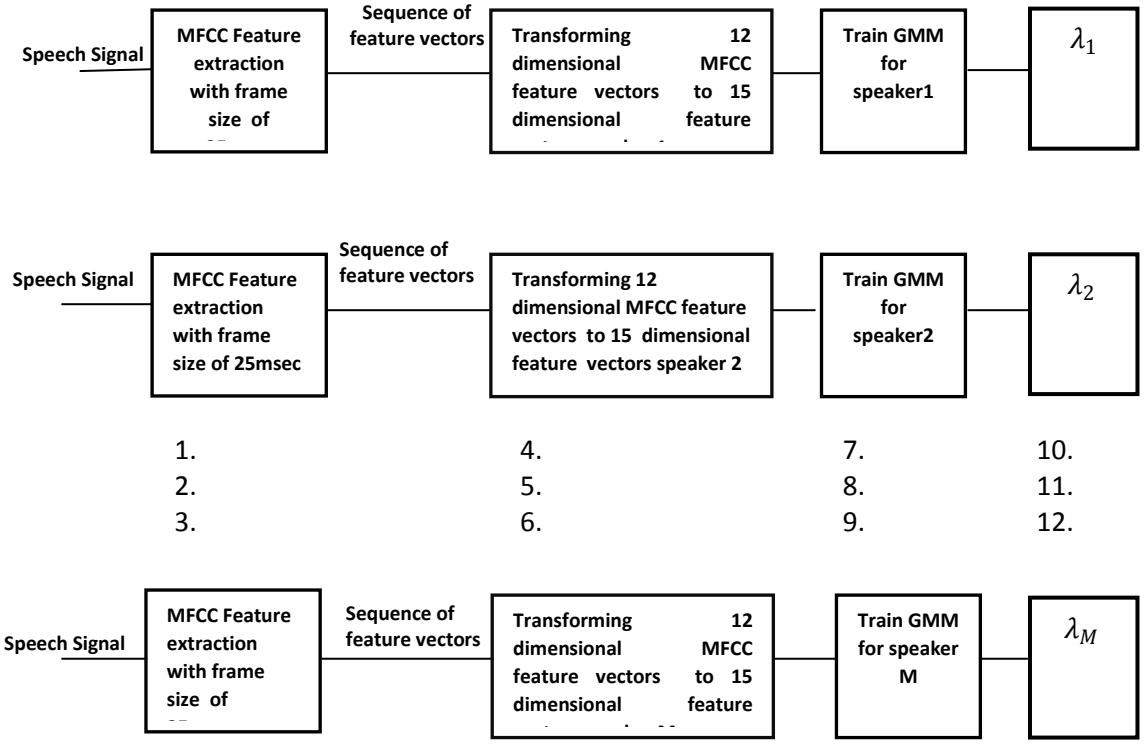

Fig. 4: GMM training for gender identification system.

From the gender identification performance, it is found that new feature vector based GMM gender identificationsystem outperformed conventional GMM identification system. It is also observed that for small training data also, theperformance very good.

eISSN 1303-5150

Table.1 The gender identification for varying the number of Gaussian components

S.No	Training speech Duration	No. of Gaussian Components	Gender Identification based on new feature vectors			Gender Identification based on MFCC feature vectors		
			1 Sec	2Sec	3Sec	1Sec	2Sec	3Sec
1	10 Sec	8	72	75	79	57	61	65
		16	74	76	82	58	63	67
		32	75	78	83	60	64	68
		64	76	79	86	61	66	71
2	20 Sec	8	78	81	87	64	69	72
		16	79	84	89	65	71	74
		32	80	86	90	67	74	76
		64	82	88	91	69	75	77
3	30Sec	8	8 2	84	88	67	72	75
		16	84	85	90	69	73	77
		32	85	87	92	70	74	78
		64	87	89	95	72	75	80

Conclusion

In this paper we have demonstrated new feature vector based gender identification using Gaussian mixture model. The GMM effectively capture the gender specific information effectively. First the new feature vectors are created MFCC features as explained. Using these new feature vectors, GMM based gender identification system is created. The performance of new feature vector based GMM gender identification model is evaluated by varying the number of mixture components from 8 to 64 and varying the test speech duration also. It is found that, the performance is increasing when the no. of mixture components are increased.

References

1.H.Meinedo, Age and gender classification using late fusion of acoustic and prosodic features, in proceedings of interspeech 2010, Makuhari, Japan, pp 2818-2821, 2010.

2. R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor, “Emotion recognition in human-computer interaction,” IEEE Signal Processing Magazine, vol. 18 (1), no. 1, pp. 32 – 80, January 2001.

3. Y. Zeng, Z. Wu, T. Falk, and W. Y. Chan. “Robust GMM based gender classification using pitch and RASTA-PLP parameters of speech”. In Proc.5th.IEEE Int. Conf. Machine Learning and Cybernetics, pages 3376–3379.Dalian, China, 2006.

4. Rakesh K., Dutta S., and Shama K., “Gender Recognition using Speech Processing Techniques” in LABVIEW,” International Journal of Advances in Engineering & Technology, vol. 1, no. 2, pp. 51-63, 2011.

eISSN 1303-5150

5. JakubKarwan, Khalid Saeed, "A New Algorithm for Speech and Gender Recognition on the Basis of Voiced Parts of Speech" Communications in Computer and Information Science Volume 245, 2011, pp 113-120.

6.Reynolds D.A and Rose R.C, Robust text-independent speaker identification using Gaussian mixtures models, IEEE transactions on speech and audio processing, Vol.3 no.1, pp.72-83, 1995.

7.Djemili R, Bourouba H, A speech based gender identification system using four classifiers, In proceedings of the 2012 International conference on multimedia computing and systems, Morocco, May 2012, pp.184-187.

8. Mahamoodi D, Marvi H, Taghizadeh M, Soleimani, Age estimation based on speech features and support vector machine, In proceedings of the 2011 3rd computer science and electronic engineering conference (CEEC), Colchester UK, 13-14 Jul 2011, pp.60-64.

Mahmoodi, D.; Marvi, H.; Taghizadeh, M.; Soleimani, A.; Razzazi, F.; Mahmoodi, M. Age Estimation Based on Speech Features and Support Vector Machine. In Proceedings of the 2011 3rd Computer Science and Electronic Engineering Conference (CEEC), Colchester, UK, 13–14 July 2011; pp. 60–64.