International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in; editorijoer@gmail.com

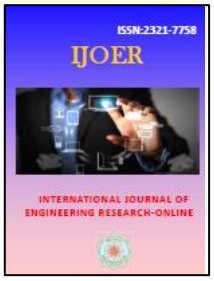
Vol.5., Issue.2, 2017
March-April

RESEARCH ARTICLE

ISSN: 2321-7758

A NOVEL MULTIPLE FEATURES EXTRACTION BASED FUZZY APPROACH FOR GENDER IDENTIFICATION IN CLASSICAL MUSIC

RAJBEER KAUR[1], GAURAV[2]
[1]M.Tech Scholar, [2]Assistant Professor
Electronics and Communication Engineering Department, PTU, SUSCET, Tangori, Mohali
[1]rajbeer.aquarious@gmail.com;[2]gauravece@sus.edu.in



ABSTRACT
Gender classification on the basis of audio or voice samples is one of the prominent topics for implementing in research work. As per the traditional work many techniques had been developed such as ANN, SVM etc., but these methods used earlier for classification were not able to achieve the accurate results. Hence the paper proposed a new method for gender classification uses the technique of feature extraction along with the fuzzy system for the purpose of classification. The motivation behind the implementation of the technique is to increase the accuracy level of the system and to reduce the number of errors. The efficiency of the technique is calculated by performing the simulation on the audio samples in two modules i.e. Direct Classification and Matching analysis. The result section depicts the results obtained after implementation.

Keywords-, Gender Classification System, Fuzzy System, Direct Classification, Matching Analysis.

1. INTRODUCTION

The study of Gender identification is one of the most rarely touched applications of Music Information retrieval research. The gender detection is a challenging task in terms of special information that we infer from audio file i.e. audio description. In any music information retrieval research, various routine audio feature extractors and classifiers are used to solve typical problems like singer identification, music melody finder, and audio data indexing. Usually, the interest in voice-gender conversion was of a more theoretical nature rather than founded in real–life applications. With increase in biometric security applications, mobile and automated telephonic communication and resulting limitation in transmission bandwidth, practical applications of gender recognition have increased many folds. Gender-based differences in human speech are partly due to physical differences such as vocal fold thickness or vocal tract length and partly due to the variations in speaking style. Since these changes are reflected in the speech signal, we hope to exploit these properties to automatically classify a speaker as male or female. For gender identification problem the pitch information was used which considerably depends on the speech quality. This disadvantage makes this approach unsuitable for video indexing.

Audio content extraction its mgt can be achieved by different means, sometimes overlapping, and anyway non-exclusive means. There are different types of audio content extraction techniques: computational auditory scene analysis techniques, statistical techniques, manual keyword labeling techniques. Those techniques can work synergistically with automatic

techniques in order to help describe and organize more thoroughly the content of audio material. Describing sound content involves the use of procedures, techniques and data that have been originated and developed in poles apart research areas i.e. signal processing, music cognition, artificial intelligence, etc. for solving related problems. Automatic acquisition of attributes characterizing a person interaction with computer and makes it possible to conduct the communication in the way that is best suited to predict the needs of the user. One of key features of a interacting user is her/his gender. Therefore Automatic gender recognition can be considered as a method from the domain of biometry. In medical applications automatic gender recognition is a source of important information of the interacting patient, especially in voice-controlled network applications e.g. emotional state of the patient may be important feature necessary to more precisely diagnose the patient or provide necessary aid. In typical applications based on automatic speech recognition, speaker gender can be used to select appropriate acoustic model in order to improve automatic speech recognition accuracy. In most practical applications, the gender needs to be known at the beginning of the automatically controlled dialog. Hence, it should be possible to differentiate the gender using very short sample of speech, usually the single sound whose duration is in the order of 2-4 seconds.

In gender detection, a range of audio features of the male singer and the female singer fluctuate with respect to the aspects of the physical distinctiveness of singer. In some of the research done, till now, on the gender identification of a narrator, the major focus is for eternity on audio information that is extracted and manipulated. For a singer singing the North Indian classical music becomes much intricate when add-on the musical instruments since the instruments generate frequencies that are merged in the singing voice that it becomes impossible to extract exact voice and its features from a homogenous music. Typical accompanying instruments such as 'Tanpura' or violin or harmonium being continuously played in the same scale and almost at the same note of the frequency where the singer has taken the long stay on the musical note, makes it impracticable at some times to know that which sound relates to the singer and which is of musical instrument such as Tanpura or violin [1].

2. PROBLEM FORMULATION

Till date we do not have an audio player which can classify music based upon gender for audios in their database or libraries. Gender identification is one of the most rarely touched applications of Music Information Retrieval Research. In gender detection, various audio features of male singer and female singer differ with respect to the aspects of physical and psychological characteristics of the singer. For a singer, singing North Indian classical music, the process becomes complex when accompanying musical instruments generate frequencies that are merged in the singing voice so that it becomes impossible to extract exact voice and its features. Which are typical accompanying instruments such as 'Tanpura' or violin/harmonium being continuously played in the same scale and are almost at the same note of frequency where the singer has taken a long stay on musical note, makes it impossible to differentiate the sound of singer and that of Tanpura or violin [1].

Also in Indian music all the supplementary instruments are toned or pitched according to the voice of singer, so unless we detect that whether this is male or female, we cannot have different set of classified options for the music processing. In digital recording environments and wherever we need technical analysis it is important that each voice should be separated and classified for further processing. It is an important part of digital sound processing. Merely applying filters to this background voice may help for polyphonic music recordings, such as, western classical or popular music, but in case of North Indian classical music, which is homophonic, the continuous interference of a supportive musical instrument such as 'Tanpura' or 'Harmonium' cannot be neglected or filtered out [1]. Fuzzy schemes are good classification techniques that can be easily corporate to deal with above type of problem.

3. PROPOSED SYSTEM

In the presented work, the gender recordings are firstly collected and a database is formed. Then the voice is differentiated from instrumental voice

by training a classifier for instrument. Then the gender is identified from retrieved vocal voice. Features of the gender are then analyzed and compared with database by applying classification method namely FUZZY INFERENCE SYSTEM (FIS). This method helps in identifying gender. The result of Fuzzy is carried out then and as a result the accuracy of the system is then checked.

- **Gender Detection**

The presented work consists of four stages. First stage is input audio database in which we obtain the north Indian classical music dataset in proper sampled and registered form. As input we use the audio of male, female and a combination of male and female. Second stage is feature extraction in which input data is transformed into a set of features. Audio features like pitch, pitch variation, zero crossing rate (ZCR), Energy, Spectral Roll Off, Spectral Centroids will be analyzed for the proposed work.

Third stage is discriminating vocal and instrumental voice in which vocal voice will be differentiating from instrumental voice (Tanpura, Tabla, Harmonium, and Violin) by training the classifier for instrument. Fourth stage is gender classification in which the obtained vocal voice will be further classified based upon gender. The male or female classification of the singer can be obtained by applying the algorithm fuzzy inference system (FIS) which is a rule based classifier.

4. METHODOLOGY

The proposed system is implemented in order to perform gender classification on the basis of audio files. This comprises of various process such as feature extraction, classification etc. Fuzzy System is used to perform the classification on the basis of various set of rules. The flow of the proposed work is as follows in figure 1. The present technique works on the basis of two modules i.e. direct classification and matching analysis.

The step wise methodology of the present work on the basis of modules is as follows:

4.1 DIRECT CLASSIFICATION

This module performs the direct gender classification on the basis of selected particular audio by the user. The steps covered under this module are as follows:

1. Take an input in the form of the sample audio from the database. The audio file contains the voice of male or female or both singers.
2. After providing an input audio to the system, next step is to extract the features from the audio sample to perform the further processing on it. The features that are extracted from the audio are as follows:
   - ZCR (Zero Crossing Rate)
   - Pitch
   - Energy
   - Spectral Roll Off
   - Spectral Centroids
3. Next Step is to generate Fuzzy interface or system. The creation of the fuzzy interface is done on the basis of membership functions and rules that are pre-defined.
4. After generation of fuzzy interface, next step is to perform the gender classification. The classification will be performing on the basis of defined rules of the fuzzy system.
5. After this, the final result will be generated that whether the audio file consist of male voice, female voice or both.
6. Then next step is to perform result evaluation to check the proficiency of the proposed system.

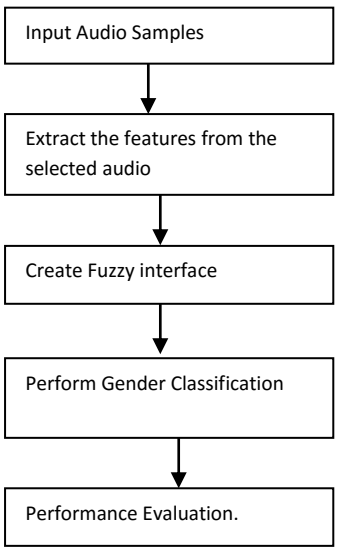


Figure 1. Block Diagram of proposed work

4.2 MATCHING ANALYSIS

Matching analysis is another module of gender classification on the basis of audio. It takes more than one audio files and then evaluate the

gender of the audios in continuity. The steps for this module are as follows:

1. First of all in this module the creation of fuzzy system will take place.
2. Next step is to perform the evaluation of the audio files that are stored in the database. The evaluation will be performed on the audios individually.
3. After evaluation last step is to evaluate the performance parameters.

5. RESULTS AND DISCUSSIONS

This section depicts the results that are obtained after implementing the present work. In this, feature extraction from the audio is done on the basis of five parameters and then classification is performed by using fuzzy interface or system. The results of the proposed work in proper sequence order are as follows:

Select a module for gender classification: as shown in the figure2, user have click on a button i.e. whether Direct classification or matching analysis. The purpose behind using these two modules is defined in previous section along with the processing sequence of them. Let's consider the user clicks on "Direct Classification."

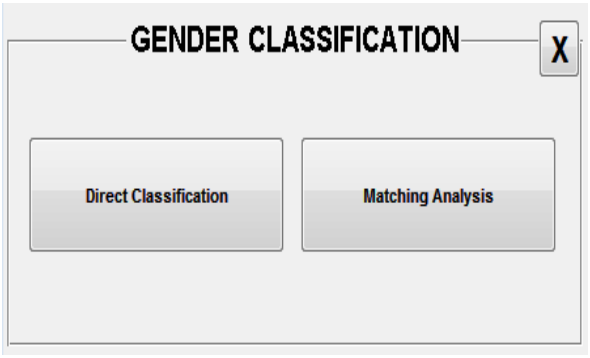


Figure 2 Graphical user interface of the proposed work

DIRECT CLASSIFICATION: The image in figure3 depicts the layout of this module; first step is to click on "Browse Sample to Process" button. After then a dialogue box will be open from which user have t select an audio file. After then next step is to click on "Extract Feature" button. After this user have to follow the steps that are defined in the methodology section.

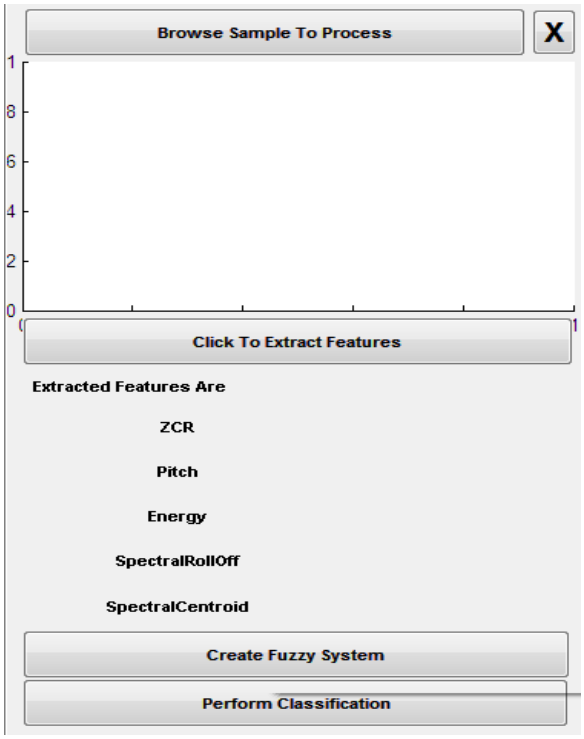


Figure 3 GUI for direct classification

The figure depicts the graph of wave signals with respect to the audio samples and amplitude of the signals.

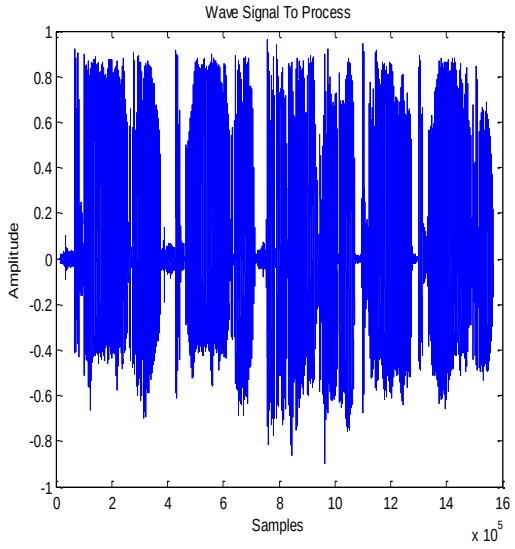


Figure 4 Graph of wave signals to process

The figure 5 below defines the fuzzy inference system of proposed work which takes 5 inputs and generates 1 output. The Mamdani model for fuzzy interference is used for proposed system. The output produced is in the form of gender classification on the basis of audios, it has three possibilities that either the output is male, Female or neutral which means no result.

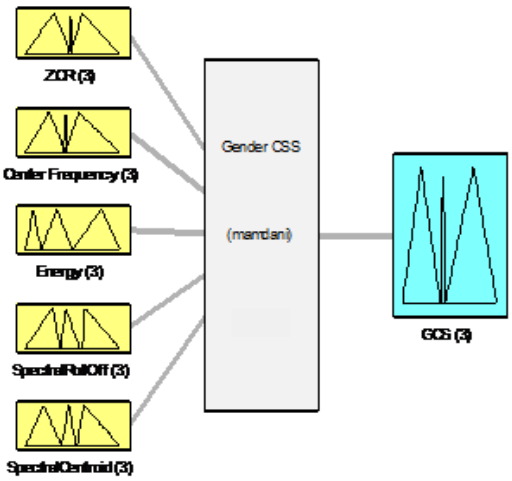


Figure 5 Fuzzy Interface of the proposed work

The figure below represents the output of the direct matching module. In this figure6 the window depicts the list of extracted features along with their obtained values i.e. ZCR (Zero Crossing Rate), Pitch, Energy, Spectral Roll Off, Spectral Centroids. The classification depicts that the selected audio sample contains of male voice.

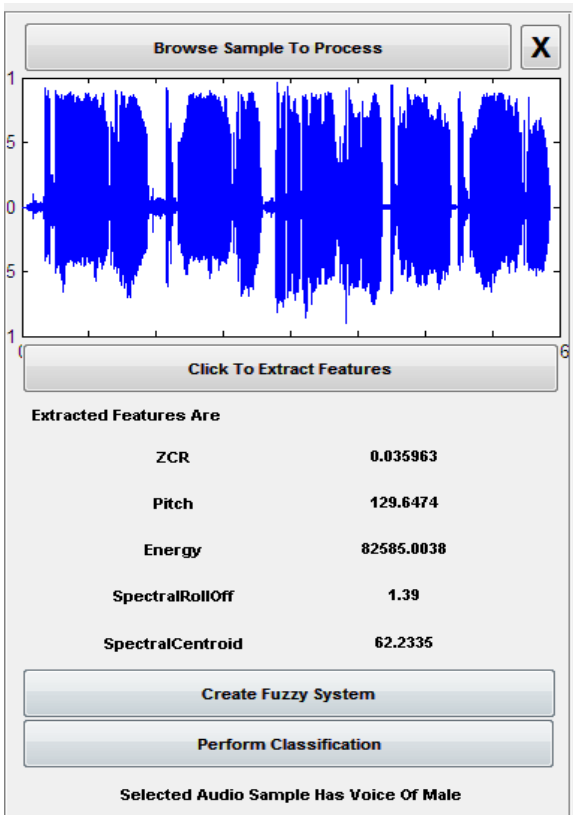


Figure 6: Results obtained after Direct Classification

The table below depicts the values of extracted parameters from the audio samples that is given to the system as an input sample.

Table1 Features Extracted Parameters

S.No	Features Extracted	Values
1	ZCR	0.035963
2	Pitch	129.6474
3	Energy	82585.0038
4	Spectral Roll Off	1.39
5	Spectral Centroids	62.2336

MATCHING ANALYSIS: Matching analysis is another module of gender classification on the basis of audio. It takes more than one audio files and then evaluate the gender of the audios in continuity. The figure7 shows the graphical framework corresponding to this module. In this first of all user have to click on "Create Fuzzy System" and then evaluation will be performed on all of the audio samples that are saved in the database. Then on the basis of evaluated results the performance of the system will be measured or evaluated.

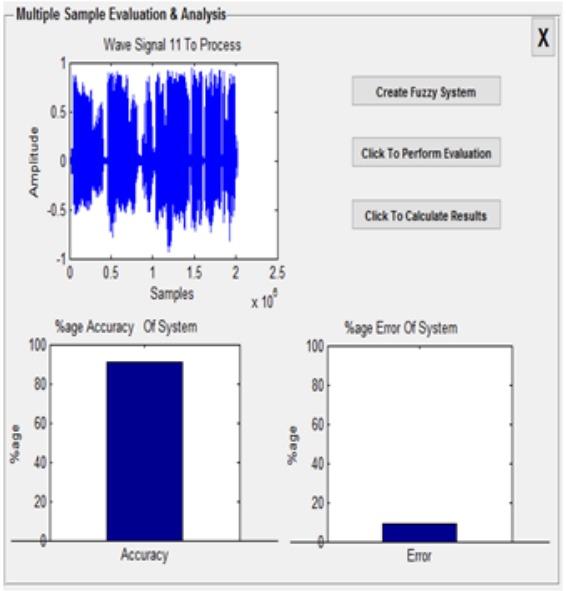


Figure 7: GUI framework of "Matching Analysis"

The figure8 portrays the comparison graph on the basis of accuracy achieved after implementing the proposed work and results or accuracy that were achieved in traditional work. The comparison is done between 5 techniques which are as follows:

1. ANN
2. SVM
3. ZMRI

4. ZRMB
5. PFuzzy (present work)

On the basis of values that are obtained from the graph it is concluded that the proposed technique has high level of accuracy as compare to the traditional techniques. ANN has less accuracy out of all the techniques. The accuracy level is calculated in percentage.

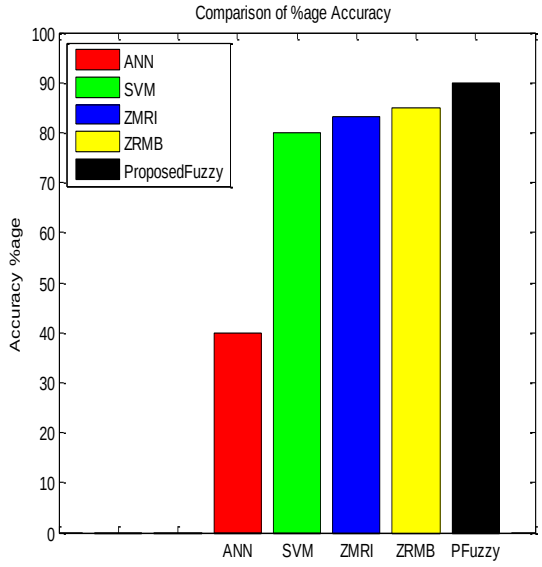


Figure 8: Comparison graph illustrating the level of accuracy

The Graph (figure9) depicts the comparison between proposed and traditional work on the basis of percentage Error that is generated while implementing the techniques. The obtained value proves that the ANN has the higher rate of the error whereas PFuzzy has least amount of errors with respect to all other techniques.

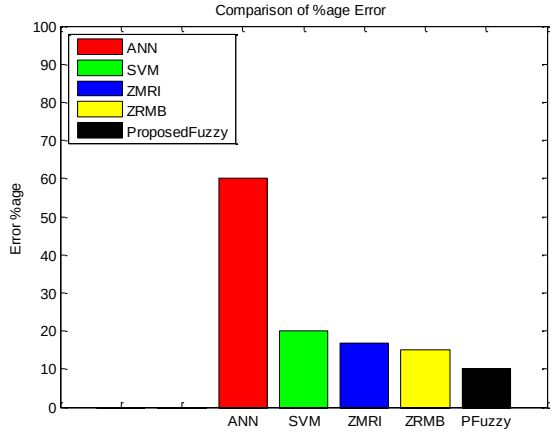


Figure 9 Comparison graph illustrating the Error (%age)

6. CONCLUSION AND FUTURE SCOPE

At last it is concluded that the present work implements the Fuzzy System in order to evaluate the gender classification on the basis of audio samples. The present work is quite efficient and has the high level of accuracy as compare to traditional methods as can be seen from the calculated results in the previous section. The objective that was considered is totally accomplished after the implementation process that is the accuracy achieved by the proposed work is more than 90%.

In Future the method can be enhanced by increasing the number parameters for feature extraction and by implementing the more advanced classification technique in order to obtain more proficient results.

REFERENCES

[1] S.H. Deshmukh, S.G. Bhirud,"A Novel Method to Identify Audio Descriptors, Useful in Gender Identification from North Indian Classical Music Vocal", International Journal of Computer Science and Information Technologies, Vol. 5, NO. 2, 2014.

[2] S. Gulati, A. Bellur, J. Salamon, R. H.G, V. Ishwar, H. A. Murthy, X. Serra, "Automatic Tonic Identification in Indian Art Music: Approaches and Evaluation", Journal of New Music Research, Vol. 43, Issue 1, 2014.

[3] F. Weninger, C. Wagner, M. Wollmer, Bjorn Schuller, L.P. Morency, "Speaker Trait Characterization in Web Videos: Uniting Speech, Language, and Facial Features", IEEE International Conference on Acoustics, Speech and Signal Processing, 2013.

[4] J. SAS, A. SAS, "Gender Recognition Using Neural Networks and ASR Techniques", Journal of Medical Informatics & Technologies, Vol. 22, 2013.

[5] O. Singh, G. Bommagani, S. R. Ravula , V. K. Gunjan, "Pattern Based Gender Classification", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue 10, 2013.

[6] Z. Ma, E.Fokoue, "Speaker Gender Recognition via MFCCs and SVMs", Rochester Institute of Technology, 2013.

[7] S.A. Khan, M.Nazir, N.Riaz, "Optimized features selection for gender classification using optimization algorithms", Journal of

Electrical Engineering & Computer Sciences, 2013.

[8] K.Meena, K.R.Subramaniam, M.Gomathy, "Performance Analysis of Gender Clustering and Classification Algorithms", International Journal on Computer Science and Engineering, Vol. 4, NO. 3, 2012.

[9] Y. Hu, D. Wu, A.Nucci, "Pitch-based Gender Identification with Two-stage Classification", Journal Security and Communication Networks, Vol. 5, Issue 2, 2012.

[10] K. Rakesh, S. Dutta, K. Shama, "Gender Recognition Using Speech Processing Techniques in LABVIEW", International Journal of Advances in Engineering & Technology, 2011.

[11] M.H. Sedaaghi,"A Comparative Study of Gender and Age Classification in Speech Signals", Iranian Journal of Electrical & Electronic Engineering, Vol. 5, NO. 1, 2009.

[12] H. Harb, L. Chen,"A General Audio Classifier based on human perception motivated model", Journal Multimedia Tools and Applications, Vol. 34, Issue 3, 2007.

[13] G.Tzanetakis, "Audio-Based Gender Identification Using Bootstrapping", IEEE Pacific Rim Conference on Communications, Computers and signal Processing, 2005.

[14] H. Harb, L. Chen, "Gender Identification Using a General Audio Classifier", International Conference on Multimedia and Expo, Vol. 2, 2003.

[15] A. Divakaran, K.A. Peker, R. Radhakrishnan, Z. Xiong, R.Cabasson, "Video Summarization Using MPEG-7 Motion Activity And Audio Descriptors", A Mitsubishi Electric Research Laboratory, 2003.

[16] P. Herrera, X. Serra,G.Peeters, "Audio Descriptors and Descriptor Schemes in the Context of MPEG-7", International Computer Music Conference, 1999.

[17] IMarc A. Zissman, "Comparison of :Four Approaches to Automatic Language Identification of Telephone Speech", IEEE Transactions on speech and audio processing, Vol. 4, No. 1, 1996.

[18] Saurabh H. Deshmukh, Dr. S.G.Bhirud, "Analysis and application of audio features extraction and classification method to be used for North Indian Classical Music's singer identification problem", international Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 2, 2014.

[19] Lakshmikanth.S, Natraj.K.R, Rekha.K.R, "Noise Cancellation in Speech Signal Processing-A Review", International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 1, 2014.

[20] Lie Lu, Rui Cai, and Alan Hanjalic, "Towards a unified Framework For Content-Based Audio Analysis", IEEE, 2005.

[21] Md. Sadek Ali, Md. Shariful Islam and Md. Alamgir Hossain, "Gender Recognition System Using Speech Signal", International Journal of Computer Science, Engineering and Information Technology, Vol.2, No.1, 2012.

[22] Kunjithapatham Meena, Kulumani Subramaniam, and Muthusamy Gomathy, "Gender Classification in Speech Recognition using Fuzzy Logic and Neural Network", The International Arab Journal of Information Technology, Vol. 10, No. 5, 2013.

[23] Alireza Zourmand, Hua-Nong Ting, and Seyed Mostafa Mirhassani, "Gender Classification in Children Based on Speech Characteristics: Using Fundamental and Formant Frequencies of Malay Vowels", The Voice Foundation, 2013.

[24] Milan Sigmund, "Gender Distinction Using Short Segments of Speech Signal", International Journal of Computer Science and Network Security, Vol.8, No.10, 2008.

[25] Gang Chen, Xue Feng, Yen-Liang Shue, and Abeer Alwan, "On Using Voice Source Measures in Automatic Gender Classification of Children's Speech", Interspeech, 2010.

[26] L.P.Bhaiya, Arif Ullah Khan, "Hindi Speaking Person Identification using Zero Crossing rate and Short-Term Energy", International Journal of Soft Computing and Engineering ISSN: 2231-2307, Vol.2, Issue-4, 2012.

[27] Rashmi C R, "Review of Algorithms and Applications in Speech Recognition System",

International Journal of Computer Science and Information Technologies, Vol. 5, 2014.

[28] Samiksha Sharma, Anupam Shukla and Pankaj Mishra,"Speaker and Gender Identification on Indian Languages using Multilingual Speech",International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 4, 2014.

[29] Mr. Kashyap Patel, Dr. R.K. Prasad, " International Journal of Advanced Research in Computer Science and Software Engineering", Volume 3, Issue 5, 2013.

[30] R Rajeswara Rao," Robust Feature Based Automatic Text-Independent Gender Identification System Using Gmm", Int. Journal of Engineering Research and Applications, Vol. 4, Issue 1, 2014.