

Gender Recognition Using Speech Features

Trupti Vishe
PG Student

Electronics & Telecommunication Engineering
PHCET, Rasayani, India

Email: trupti.vishe@gmail.com

Prof. Kiran Inamdar
Assistant Professor

Dept. of Electronics & Telecommunication Engineering
PHCET, Rasayani, India

Email: kinamdar@mes.ac.in

Abstract - The paper describes the novel approach to gender identification using speech features. The speech signal is in the form of recorded utterances from number of speakers. The speech features are analyzed using Praat Software. Utterances are recorded with different emotions to calculate comparative study of features between male and female speakers. The method is experimented using speech samples in English. Adaptive Network-based Fuzzy Inference System (ANFIS) will be adopted as one of the fuzzy neural network classifiers identifying gender from speech. Formant frequency, Pitch & Energy are features used to identify speaker gender. Gender recognition is a task of recognizing the gender from his or her voice. With the current concern of security worldwide speaker identification has received great deal of attention among of the speech researchers.

Keywords – Gender Recognition; ANFIS; F_0 ; Feature Extraction

I. INTRODUCTION

Speech signal conveys data for correspondence among individuals. It additionally passes on the data with respect to Speaker feelings, expressions and so forth. Speaker data is conveyed as components like force range thickness, vitality, pitch, power. Applications like human PC collaboration, biometrics, manmade brainpower need discourse signal preparing. The extremely essential element of any communicating individual is his or her sex. The discourse flag likewise convey the data of the specific speaker including social elements, full of feeling element and the properties of the physical voices creation mechanical assembly for which person can perceive whether the speaker is a male or a female effortlessly, amid phone discussion or any shrouded state of the speaker[1][2]. The overall security worries with the majority of the consideration towards sexual orientation recognition among discourse research. Individuals can normally separate amongst male and female speakers however for machines required in various specialized field applications it's a non paltry undertaking in view of the variability in volume and pitch levels of male and female voices. The most vital procedure in discourse and speaker acknowledgment framework is to order sexual orientation of speaker. For discourse signal based sex recognizable proof, the most ordinarily utilized components are pitch period and Mel-Frequency Cepstral Coefficients (MFCC) [3]. The primary instinct for utilizing the pitch period originates from the way that the normal crucial recurrence (equal of pitch period) for men is commonly in the scope of 100-146 Hz, though for ladies it is 188-221 Hz [4]. Speakers with various age and sex, demonstrates a considerable measure of information on the recurrence of the voice crucial (F_0) in the discourse. The work talked about in this paper for the most part focuses on crucial recurrence for sexual orientation order.

II. Related Work

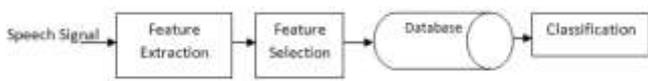
Jerzy SAS1, Aleksander SAS2 tentatively assessed sex acknowledgment utilizing discourse tests as a part of English and in Polish. The examination portrayed that MFCC highlights utilized as a part of the simulated neural system classifier guarantee high exactness of sex acknowledgment in light of short expressions. The most effective related work is "Speaker acknowledgment in a multispeaker domain" was submitted in Proc. seventh European Conference on Speech Communication and Technology (Eurospeech 2001) (Aalborg, Denmark, 2001), pp. 787–90[5]. Md. Sadek Ali,

Md. Shariful Islam and Md. Alamgir Hossain built up a framework forfor discourse encoding, investigation, amalgamation and sexual orientation recognizable proof. The framework was tried with 10 speakers(5 male individuals and 5 female people).The acknowledgment precision of this framework is 80% on average.The acknowledgment process included the measurable investigation and edge technique.The exploratory result, had demonstrated that acknowledgment rate diminishes as the quantity of speaker increments.

In Gender Classification in Speech Recognition utilizing Fuzzy Logic and Neural Network by Kunjithapatham Meena, Kulumani Subramaniam and Muthusamy Gomathy, the new technique with fluffy rationale and neural system is utilized to recognize the sexual orientation of the speaker.The strategy talked about utilizations edge worth to distinguish that speaker has a place with which sex. This quality is ascertained as a mean worth from results got from fluffy rationale and neural system. The testing was done utilizing Matlab Platform. The database utilized was Harvard-Haskins database [6]. Sachin Lakra, Juhi Singh and Arun Kumar Singh, displayed results on Results on ordering a speaker on the premise of sexual orientation by handling discourse and investigating the voice tests. A computerized sexual orientation order was effectively performed by ANFIS. A programmed sexual orientation based grouping was actualized utilizing discourse handling of a discourse signal by characterizing it into quiet/unvoiced/voiced. In the two cases, i.e., male and female sex acknowledgment, the framework is 92% and 100% effective, separately. The general proficiency of the framework was observed to be 96.55 %. In International Journal of Engineering Science and Innovative Technology (IJESIT),S.Sravan Kumar, T.RangaBabu introduced deal with Emotion and Gender Recognition of Speech Signals Using SVM.The framework is created for feeling and sexual orientation acknowledgment. The framework comprises of two subsystems, to be specific Emotion acknowledgment (ER)&Gender acknowledgment (GR). For this two bolster vector machines (SVM'S)are utilized for the male and female speaker's feeling acknowledgment. The trial examination demonstrates the execution as far as exactness of the proposed ER framework. The test results highlight that the Gender Recognition (GR) subsystem permits expanding the general feeling acknowledgment exactness from 80.4% to 84.5% because of the from the earlier information of the speaker sexual orientation [7].

III. System Implementation

The general block diagram for gender recognition can be shown as follows:



A. Feature Extraction

Features selected for emotion Recognition are pitch, Energy and MFCC. Feature extraction is performed using PRAAT. Pitch and energy/intensity information are known to be useful in emotion recognition and are thus very commonly used. MFCCs are powerful speech features and are commonly used in automatic speech recognition and speaker and language recognition technologies. we have compared standard autoscale of each feature with the use of min-max normalization.

B. Feature Selection

Feature selection is the important step in gender classification. The selected features decide the efficiency for a particular gender classification system. Feature selected for this methods are energy, Intensity, fundamental frequency and pitch. Out of these features most important features for gender identification are pitch and fundamental frequency. Using the Template.

C. Database

For database preparation voice samples from number of speakers are collected using PRAAT software. The performance of a classifier that can distinguish different genders ultimately depends on the completeness of the samples and how similar it is compared to the real world data. Data capturing method employs speakers to utter various words. The speech signal contains a large number of information which reflects the behavioural characteristics. So in the research of speech based gender recognition, the most important thing is that how to extract and select better speech. features with which gender can be classified. PRAAT is a very flexible tool to do speech analysis. It offers a wide range of standard and non-standard procedures, including spectrographic analysis, articulatory synthesis, and neural networks.

D. Classification

A versatile system based fluffly induction framework (ANFIS) is used to arrange sexual orientation of speaker. The acronym ANFIS gets its name from versatile neuro-fluffly surmising framework. Utilizing a given information/yield information set, the tool stash capacity Anfis develops a fluffly surmising framework (FIS) whose enrollment capacity parameters are tuned (balanced) utilizing either a back proliferation calculation alone, or in blend with a minimum squares sort of technique. This permits your fluffly frameworks to gain from the information they are demonstrating. The parameters connected with the enrollment capacities will change through the learning process. The calculation of these parameters (or their change) is encouraged by an angle vector, which gives a measure of how well the fluffly deduction framework is displaying the info/yield information for a given arrangement of parameters. Once the angle vector is acquired, any of a few improvement schedules could be connected keeping in mind the end goal to conform the parameters in order to decrease some mistake measure (normally characterized by the aggregate of the squared contrast amongst genuine and sought outputs)[8].

Anfis utilizes either back proliferation or a mix of minimum squares estimation and back spread for participation capacity parameter estimation. ANFIS advances the utilization of fluffly deduction framework as the motor to make a tenet based methodology from the given input-output relationship. It comprises of five layers with every layer performing diverse procedures; to be specific: fuzzification, augmentation, standardization, execution and de-fuzzification separately. To handle this variability, pitch has been arranged into the fluffly arrangements of male and female persons. The inspiration for ordering sexual orientation into fluffly sets is the presence of high variability and also certain similitudes in the voices of male and female speakers. Fluffly sets permit the production of a scientific model consolidating high variability and in the meantime obliging similitudes as a fluffly limit. This permits exceptions to be arranged with higher exactness than on account of a fresh classification.[9].

The procedure of sex acknowledgment can be clarified as takes after:

Information the discourse test. Extricate the voiced part of the chose discourse test for pitch extraction on the premise of an abundancy limit. Ascertain the pitch from the voiced bit. Train an ANFIS system to group pitch into two fluffly sets, in particular, male and female. Utilize the prepared ANFIS system for pitch-based sexual orientation acknowledgment.

IV. EXPERIMENTAL RESULTS

The detailed recognition results can be shown in the following table:
Table 1: Gender Recognition Results

No. of speaker	Gender	% Recognition
1	Female	92.45
2	Male	85.69
3	Female	90.21
4	Male	85
5	Female	94.38
6.	Male	86

V. Conclusion

The main aim of this work was to develop a gender recognition system using speech signal. The feature selection is one of the most important factors in designing a gender recognition system. From the study of different previous research works it was found that out of the different features the Fundamental frequency gives best classification rate. The average recognition accuracy is 89 %. The proposed approach is based on the analysis of the fundamental frequency F0, which is typically higher in females than in males. In future some more combinations of speech features can be considered to have better recognition rate.

REFERENCES

- [1] Jerzy SAS¹, Aleksander SAS², "Gender Recognition using neural networks and ASR techniques"
- [2] Huang X., Acero, A., and Hon, H.-W. "Spoken Language Processing: a Guideto Theory, Algorithm and System Development" prentice-Hall, New Jersey, 2001.
- [3] H. Harb and L. Chen, "Voice-based gender identification in multimedia applications," Journal of Intelligent Information Systems, vol. 24, no. 2, pp. 179-198, 2005.
- [4] M. Gelfer and V. Mikos, "The relative contributions of speaking fundamental frequency and formant frequencies to gender

-
- identification based on isolated vowels,” *Journal of Voice*, vol. 19, no. 4, pp. 544–554, 2005
- [5] Martin, A., and Przybocki, M. Speaker recognition in a multi-speaker environment. In *Proc. European Conference on Speech Communication and Technology (Eurospeech 2001)* (Aalborg, Denmark, 2001), pp. 787–90.
- [6] Kunjithapatham Meena, Kulumani Subramaniam & Muthusamy Gomathy, “Gender Classification in Speech Recognition using Fuzzy Logic and Neural Network”, *The International Arab Journal of Information Technology*, Vol. 10, No. 5, September 2013, 475-485.
- [7] S.Sravan Kumar, T.RangaBabu, “Emotion and Gender Recognition of Speech Signals Using SVM ”, *International Journal of Engineering Science and Innovative Technology (IJESIT)*, Volume 4, Issue 3, May 2015, ISSN: 2319-5967.
- [8] J. Vieira, F.M. Dias, A. Mota, Neuro-fuzzy systems: A survey, *5th WSEAS International Conference on Neural Networks and Applications*, Udine, Italy, 25-27 March, 2004.
- [9] Sachin Lakra, Juhi Singh and Arun Kumar Singh, “Automated Pitch-Based Gender Recognition using an Adaptive Neuro-Fuzzy Inference System”, *2013 International Conference on Intelligent Systems and Signal Processing (ISSP)*, 82-86.
- [10] Devi M., Kasthuri N., and Natarajan A., “Performance Comparison of Noise Classification using Intelligent Networks,” *International Journal of Electronics Engineering*, vol. 2, no. 1, pp. 49-54, 2010.
- [11] Sigmund M., “Gender Distinction using Short Segments of Speech Signal,” *International Journal of Computer Science and Network Security*, vol. 8, no. 10, pp. 159-162, 2008.