



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: V Month of publication: May 2020

DOI: <http://doi.org/10.22214/ijraset.2020.5397>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Identification of Gender Speech Recognition by Implementation of Fuzzy Logic System

Manisha Verma

Department of Computer Science & Engineering, Rama University, Kanpur, Uttar Pradesh, India

Abstract: *Speech analysis is still a challenging area in recognition, classification and retrieval and it is one of the most interesting fields in Digital Signal Processing. Many researches already done on it that is based on different scientific programs and materials tools and that produce an analysis that start from speech production, processing, coding and recognition, Chester, F.J .Taylor and M.Doyle were the first to find out analysis of speech recognition [1]. In this research, identification of gender females and males speech recognition will take speech samples of the word, "Close", were used to build a system in Neural Network and Fuzzy Logic to recognize the male and female speech voice and compared between two systems, which are part of a class of universal approximations of continuous functions.*

The system of the fuzzy logic was developed based on three features of the speaker voice that are power spectrum of the signal ,energy value of the signal and vowel sound "O" in the word "close" in the speech samples to increase its ability in recognizing an individual speaker also system security can differentiate for intruders by making decision based on the recognition to the speech of a one person and giving a voice acceptance authority to that person and make an access denied to others to prevent accessing the system, using samples of one person against others female's and male's samples.

I. INTRODUCTION

The voice signal is a high redundant and non-stationary signal. This attributes of speech recognition for identification to be very challenging work [2]. The speech recognition systems fall into two categories according to Speaker dependent systems that are used and often trained by one person. Those independent systems that can be used by anyone. In general, voice recognition can be partitioned into voice identification (who is speaking?) and voice and gender verification.

In addition, voice identification can be closed-set (The voice of speaker is always one of a closed set that is used for training.) or open-set (The voice of speakers from outside the training set) .Also, each variant may be implemented as text-dependent or text-independent.

Human voice (speech) can be divided into two parts: sound production and sound shaping. Sound production is caused by air passing across the vocal cords (as in a, e, i, o, u) or from a construction in the vocal word (as in "sss", "p", or "sh"). Sound words using the vocal is called voice recognition and that used in tracking the vowel sound during speech in male and female samples in my research; unvoiced speech is produced by the tongue, lips, teeth, and mouth .In signal processing sound producing is a combination of the vocal tract, the placement of the tongue, lips, teeth and the nasal passages. For each fundamental vocal words or phoneme, of English, the shape of the vocal chords is somewhat different which may lead to a different speech recognition database. In that processing, sound formatting is called filtering [3]. MATLAB program was used in the recognition work which a very useful program in processing non- stationary signals such as speech signal and many systems was designed based on its tool as in [5] .

Voice Comparison process will first get the input in the form of recorded files, and then after whatever spoken words are there that will be converted into machine-readable format from which certain characteristics will be extracted. Those characteristics will be compared with databases in which certain characteristics are identified. Gender and Voice Identification technique is totally different than Gender and Voice Verification. Because here the word used authentication.

Speaker confirmation is normally utilized as a Speaker verification so as to give access to a safe framework. These frameworks work with the user's information and ordinarily require their collaboration. While Voice Identification is the process of determining an unknown speaker's identity. So both Gender and Voice Verification and Identification will be performed after getting pattern from comparison process. Voice identification is possible because of the reason every person has a unique set of voice characteristics and speech patterns like speech, tone, etc. This example reflect both size state of the throat and mouth just as different educated personal conduct standards like voice pitch, and talking style. Voice Identification requires to get several recordings .first, one or more recordings of the voice to can be identified by getting one or more recorded voice samples of one or more persons. From these recordings the one can make determination about the identity of the unknown voice. So first it needs to

get a recording without knowing person and it is better to work with random speech of a person. Sample recorded voices and gender needs to be compared with input. For result generation here role of fuzzy logic proposed. Fuzzy means uncertainty. If comparison is made with multiple recorded voices there may be chance that instead of exact match can have similarity up to certain extent means partly match; may be 80% , 90 % or more, then one can go for using fuzzification and defuzzification.

The transformation of fuzzy set to single fresh esteem is called defuzzification and is the converse procedure of fuzzification [3]. In this case there will be role of linguistic variables related to exact match, partial match or no match, etc. In the area of Voice Technology following type of work already done:

- 1) Various types of speech recognition software developed.
- 2) In the field of healthcare, speech recognition tool developed which recognizes spoken words and will be displayed immediately on the screen after they are spoken.
- 3) With limited vocabularies speech recognition used to reduce workload of pilot.
- 4) It is also used in telephonic world also like conversion of Voicemail into the form of text.

Voice Stress Analysis technology developed, which is based on hypothesis. The application of a VSA related to examination is to determine the truthfulness of responses made by an examinee regarding the subject under investigation. The application related to automatic speaker voice recognition developed, which is more particularly for verification of a speaker authorized to access a service application, whether independently of or depending on the content of the voice segment spoken by the speaker, such as a password.

A. About Work

The model should work for gender's speech recognition, comparison, and identification. In actual various thing stated in review of literature related to speech recognition. But it needs to state that basically speech recognition can be text dependent or text independent. In case of speech and gender, verification needs to perform text dependent recognition and in identification, it needs to perform text independent recognition. If it requires providing system access based on user's knowledge then it need to verify user first. And voice identification requires identifying unknown person. This model should identify unknown person by doing comparison between voice recordings of a particular person with the database of pre-recorded voices using fuzzy logic. The system provides list of people having voice match up to certain required level.

II. MATERIALS AND METHODS

A. Recording the Voice Samples

The initial phase in the examination was building the database of voice sample, that progression requires to gather the male and females sample and an individual voice speaker sample for a female the choose word for recording the speaker's voices was speaking, 'Close', word as the sound can be heard unmistakably. Computer system was utilized for recording the speech signals for the male and females sample and utilizing Sound Recorder programming of windows (XP) and a Microphone (Creative) model HS-350. In a perfect world in recording known voices the examiner should endeavor to copy the physical conditions related with the unknown call.

That endeavors ought to incorporate chronicle the predefined voice with a similar account gadget used to record the undefined voice. Correspondingly, if the inquiry discussion was recorded over a call, it is alluring to discover the known model via call system utilizing the comparative gadget. On the off chance that there is space on the inquiry recording tape, the realized model can be set on the comparative tape? In the event that the call system is to be discover the known model, pull back straightforwardly into recording device ought to be gotten all the while [1].

Various techniques may be used to get an account of voice of a suspect. The specialist may move toward the phone and record the conversation. The authority moreover may record the assume's voice without individual's data by using a covered mouthpiece; models gained with a concealed recipient every now and again wind up being lacking quality for relationship purposes. [1]. Here, the underlying advance will be to evaluate the record of the voice using fuzzy basis, by then it need to watch that the account has a sufficient proportion of speech with which to work and that the idea of the recording has satisfactory clarity or not in the repeat run required for examination.

If it is of low quality, by then no convincing motivation to do also getting ready. By then as inspected before first it need to prepare database and a short time later after requires to get test recording that will be differentiated and existing database in conclusion it will help with perceiving known individual and will in like manner license to get summary of individuals whose voice matches with genuine model up to generally $\geq 85\%$.

B. Low-Pass Filter

Low-pass Filter, It is critical to use for the recording speech signals sample have noise that produces during speaking and typically that noise lies at the higher frequencies while the speech information lies at the lower frequencies so this channel was utilized twice, first when reading the speech signal in MATLAB program environment and second while computing the intensity of the Discrete - Fourier Transform to isolate the information of the force amplitudes for the speech signal from the noise.

III. FEATURES EXTRACTION

A. Pitch

Pitch is the most distinctive difference among male and female speakers. An individual's contribute begins the vocal strings, and the rate at which the vocal lines/folds vibrate is the recurrence of the pitch. The explanation pitch varies between genders is the size, mass and the pressure of the laryngeal exchange which incorporates the vocal folds and the glottis (the spaces between and behind the vocal folds). The crucial recurrence or pitch of the human voice is around 250 Hz and the crease length is about 10.4mm. After adolescence the human body develops to its full grown-up size, changing the element of the larynx region. The vocal overlay length in males increments to around 15-25mm. While female's vocal crease length increments to around 13-15mm. The normal pitch falls somewhere in the range of 60 and 120 Hz, and the scope of female's pitch can be found somewhere in the range of 120 and 200 Hz. Females have a higher pitch go than males in light of the fact that the size of their larynx is littler [4]. This component was extremely useful in perceiving between the male and the female voice in numerous examines, one of them was demonstrated utilizing fuzzy logic system as in [6]. In view of the contribute include I found the examination that the amplitudes of the intensity of the speech signal is another element that can be utilized in causing acknowledgment between the male and the female voice By applying the Fast Fourier to change calculation.

B. Fast Fourier Transform (FFT)

The discrete Fourier change (DFT) with a million focuses are regular in numerous applications. Present day signal and picture handling applications would be outlandish without a productive technique for processing the discrete Fourier change which change time or space – based information into recurrence based information [7]. The DFT was utilized as a features extractor in light of the fact that the recurrence size contains data about the pitch and the formants. Alongside the unearthly size additionally holds a lot of other data next to the pitch and the formants greatness. The DFT of the vector of length n is another vector y of length n . Where, $X(k)$: The discrete information signal in time area $X(j)$: The discrete information signal in recurrence space N . The last condition is a N th foundation of solidarity. When utilizing FFT calculations, a qualification is made between the the transform length and window length that is the length of the input data vector. It is defined by, the size of an external buffer. The transform length is the length of the output, the computed DFT.

C. Spectrogram

A wide range of scopes of frequencies forms signals, for example, voice. Along these lines, recurrence frequency is vital in the understanding of a speech signal. Spectrogram is one of notable recurrence frequency of the first speech. The vertical pivot compares to time. The force of the example at any moment of time relates to the vitality level. Spectrogram permits user to know the measure of vitality a speech, may have regarding recurrence scale. This is valuable device to recognize voiced and unvoiced zones, and distinguishing the pertinent recurrence that is created in this speech. There are different assortments of explores that can be performed utilizing spectrogram. The measure of data a spectrogram can give is colossal and numerous speech investigates can recognize plain English content from spectrogram [3]. In the exploration spectrogram device was utilized in MATLAB language because the force range of a sign speaks to the commitment of each recurrence of the range to the intensity of the general sign. Adjacent to it is exceptionally useful in commotion undoing and system distinguishing proof.

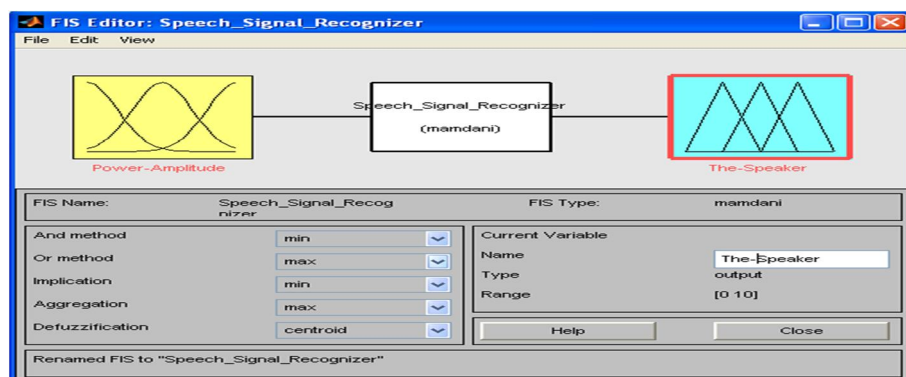
The procedure began Using MATLAB Signal Processing Block set, by stacking a speech signal from MATLAB workspace for the word, 'Close', for each example. At that point isolating the speech sign to number of sections called frames with given example time (1/8000) and 80 examples for every edge. The yield cradle size per channel (128) with (48) support cover. In light of proposals sitting parameters, the principal yield frame contains 48 beginning condition esteems followed by the main info frame. The subsequent yield frame contains the last 48 qualities from the past casing followed continuously 80 examples from the subsequent information frame, etc. The buffering of the information signal into a yield signal with 128 examples for every edge was to limit the evaluating commotion added to the speech signal. After this progression was finished, the Fourier change was taken for the sign utilizing Periodogram square, which compute a nonparametric estimation of the force range of the speech signal.

This activity is relying upon windowing process like Hamming and the reason for it is to make serious in some zone while unessential in different zones. It was applied intermittently to the speech sign and normal two spectra at once. The length of FFT was expected 128, which is the quantity of sample per frame.

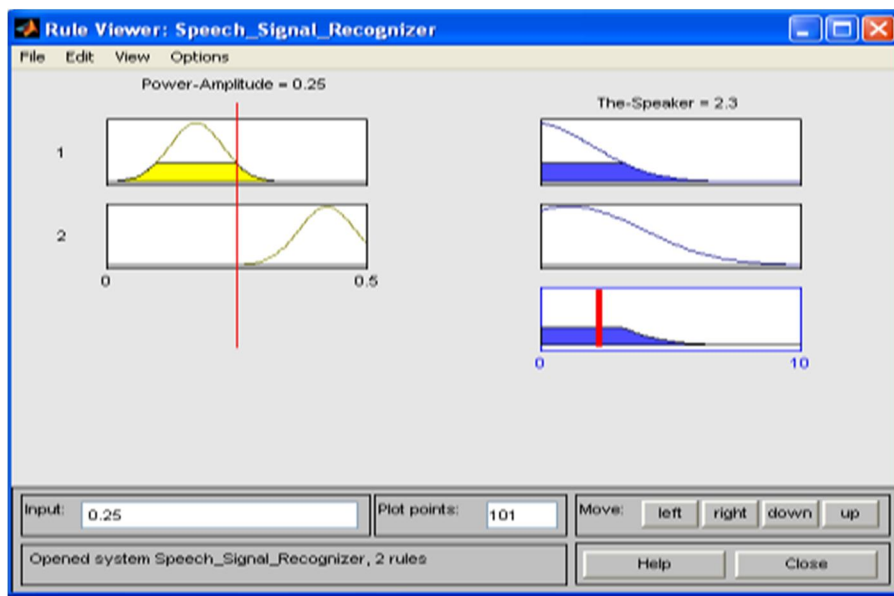
IV. TRAINING NEURAL NETWORKS FOR SPEECH RECOGNITION

An Artificial Neural Networks (ANNs) is a data handling worldview that is enlivened by the way organic sensory systems, for example, the cerebrum, process data. The main feature of the voice structure of the database framework. It is made out of an enormous number of exceptionally interconnected preparing components (neurons) working as one to take care of explicit issues. ANNs, similar to individuals, learn by model. An ANNs are designed for a particular application, for example, design acknowledgment or information arrangement, through a learning procedure [8]. Present day second-request calculations, for example, conjugate inclination descent and Levenberg-Marquardt (Bishop, 1995; Shepherd, 1997) are generously quicker (e.g., a request for extent quicker) for some issues, yet back engendering still has focal points in certain conditions, and is the least demanding calculation to comprehend. In back proliferation, the inclination vector of the blunder surface is determined. This vector focuses along the line of steepest plunge from the present point, so we realize that in the event that we move along it a "short" separation, we will diminish the mistake. A grouping of such moves (easing back as we close to the base) will inevitably locate at least some sort. The troublesome part is to choose how huge the means ought to be. Huge advances may unite all the more rapidly, yet may likewise violate the solution or (if the error surface is flighty) go off course. An example of this in neural system preparing is the place the calculation advances gradually along a precarious, thin, valley, skipping from one side across to the next. Interestingly, little advances may go in the right bearing, yet they likewise require countless cycles. Practically speaking, the progression size is relative to the incline (with the goal that the calculations settle down in a base) and to an uncommon consistent: the learning rate. The right setting for the learning rate is application-subordinate, and is commonly picked by explore; it might likewise be time-fluctuating, getting littler as the calculation advances. The calculation along these lines advances iteratively, through various ages. On every age, the preparation cases are each submitted thus to the system, and target and genuine yields thought about and the blunder determined. This blunder, along with the mistake surface inclination, is utilized to alter the loads, and afterward the procedure rehashes. The underlying system setup is irregular and preparing stops when a given number of ages pass, or when the blunder arrives at a worthy level, or when the mistake quits improving. In the examination, Back engendering Neural Network has been utilized to assemble a system that ready to perceive male from female's examples dependent on the force amplitudes reaction of the DFT for the speech signals sample where the outcomes shows that the force sufficiency reaction for male' examples were higher than the adequacy reaction for females' examples. The means to make a Neural Network based recognizer are [9]:

- 1) Specify the phonetic classifications that the system will perceive. In the exploration it was the force amplitudes for the DFT for the two male and females sample.
- 2) Find numerous examples of every one of these classes in the speech information. In the examination the best female and male examples relying upon its capacity amplitudes were utilized. The littlest force an incentive for the female example appointed objective [0 1 0 1] and the biggest force an incentive for the male example relegated focus as [0 0 1 1].
- 3) Train a system to perceive females and male sample. In the exploration the Back Propagation Neural Network (BPNN) was prepared for 462 ages, with 50-shrouded unit as a beginning however it was insufficient number to fulfill the acknowledgment procedure.



Fig(1) : The Speech Signal Recognizer System



Fig(2): Rule Viewer Speech Signal Recognizer System

V. FUZZY LOGIC

Lotfi Zadeh as a method of permitting uncertainly considered fuzzy set hypothesis and fuzzy rationale in 1965 or unclearness to be spoken to scientifically .Fuzzy sets are a super – set of old style sets .Each component in a fuzzy set is related with genuine number, which speaks to the level of enrollment of the component in the set. Fuzzy sets are generally communicated as a lot of components that having level of participation for reality esteems in shut unit interim [0, 1]. The thought behind a fuzzy set speaks to an idea and having a setting is a further extended by etymological factors. A semantic variable is allocated to a fuzzy area, a lot of fuzzy sets that speak to a total idea. Also, it is a super-arrangement of traditional rationale that manages relational words which required being either evident or bogus. Fuzzy rationale permits exceptionally nonlinear, ineffectively comprehended or scientifically complex frameworks to be displayed dependably and proficiently. What's more, it manages noise information. These qualities as in [10] recommended in the inquires about that fuzzy rationale may be a powerful device for speech acknowledgment. The fuzzy rationale tool kit for use with MATLAB is an apparatus for tackling issues. It assists with making and alter fuzzy deduction frameworks by utilizing graphical apparatuses or order line working [7]. In the examination the primary fuzzy system program was worked to perceive among male and females sample dependent on two standards:

- 1) If the force adequacy of the speech signal is little worth then female talks
- 2) If the power amplitude of the speech signal is huge worth then male speaks

The two rules helped the system to perceive the speaker as female and as a male; the system is appeared in fig (1) that was assembled utilizing mamdani MATLAB7.6.0 (R2008a) and fig (2) shows the standard watcher of the framework.

In view of the separated features, the force sufficiency, power range and vowel-sound, the above system was created to have the option to perceive an individual speaker a female example against different male and females sample. Testing the system through a few cases and changing the principles .A system was worked of 24-rules dependent on the estimations of the three extracted features, with 3-inputs speak to the three extracted features for the speech samples, utilizing the force adequacy of the DFT for the speech signal element, The force Spectrogram of the speech signal component and the sign vitality include that was gathered in the music of the vowel sound 'o' in, "Close". The system had the option to recover the necessary speech signal for the necessary speaker.

VI. CONCLUSION

Computer Science engineering and Technology is consistently changing and developing, with new innovations being identification of Gender speech Recognition utilizing Fuzzy Logic model will be an important apparatus in the zone of Voice innovation. This device will recognize rundown of those people from accessible database whose voice matches with test recording. What's more, will likewise assist with producing report that up to which level match will found.

From collecting the speech sample for both male's and female's samples, the back proliferation Neural Network (BPNN) was tried to perceive the speech signal, utilizing 50-concealed units, The objective was considered [0 0 1 1] for various male examples and females sample (F). Sexual orientation speech Recognition location for each word

by Using All Words Simultaneously with Imbalanced Corpus. Territory of all words for every speaker is determined. Twelve-dimensional component vector is utilized for every speaker where each measurement speaks to a region of the word. The exactness of 96% is accomplished, where acquired discovery. Twenty-seven times, out of 28, system identifies the sexual orientation accurately: 17 out of 18 for male and 10 out of 10 for females. By breaking down the finish of sexual orientation location by utilizing words separately and all words all the while, it very well may be induced that it is smarter to utilize more than single word to accomplish high discovery rate. It may be expected that the proposed include is one-sided when quantities of sample are distinctive for male and female speakers. To give the appropriate response about biasness of the proposed highlight, it is important to utilize the equivalent number of male and female speakers for the preparation and testing of the framework. The main accessible choice is to incorporate the female speakers of other tongue locales. It will make a harmony between the two sexual orientations and it will build the all out number of speakers. It is deduced in this segment utilization of all words at the same time gave great voice and sex discovery rate and watch the conduct of the proposed include for clean speech .

The fuzzy interface system was constructed utilizing MATLAB7.6 as a recognizer speech system for male and females sample. Two guidelines were utilized in the system to perceive the male from the female examples.

- 1) If the force sufficiency of the speech signal is little value then female talks. With extend indicated (0 – 0.5).
- 2) If the force adequacy of the speech signal is huge, esteem the male talks. With extend determined (0 – 0.5).

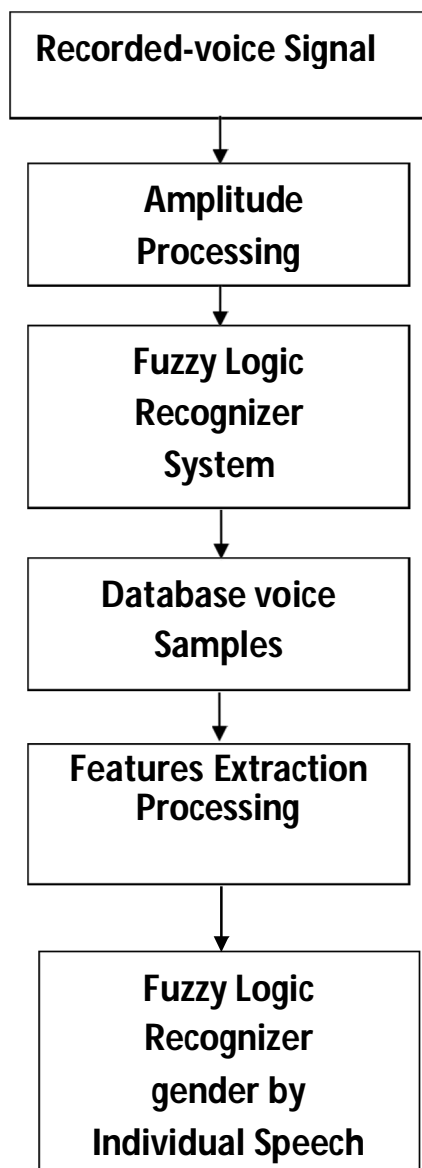
The output range for the speaker (0 – 10), female speaker brought about range (0 – 2.5), while male speaker brought about range (2.6 or more). Like each framework, there is a mistake worth may result, the outcomes are appeared in table (1) , The little power amplitude esteems came about for a females speaker , while the huge force abundancy esteems came about for male speaker ,then it will give the acknowledgment achievement pace of the framework.

Table (1): Speech Recognizer Fuzzy Interface System

Recorded voice	Amplitude (Watt)	Fuzzy output (Females Samples)	Recorded voice	Amplitude (Watt)	Fuzzy output (Males Samples)
F1	0.2826	2.27	M1	0.3368	4.07
F2	0.2647	2.07	M2	0.3209	3.47
F3	0.2816	2.26	M3	0.2673	2.1
F4	0.2573	2	M4	0.3126	3.06
F5	0.2056	1.72	M5	0.3935	2.86
F6	0.1132	1.83	M6	0.2645	2.09
F7	0.2337	1.84	M7	0.2911	2.41
F8	0.3033	4.01	M8	0.7108	3.74
F9	1.4565	2.72	M9	0.2709	2.13
F10	0.2929	3.74	M10	0.3420	3.94
F11	0.4648	2.45	M11	0.4163	2.68
F12	0.2875	2.97	M12	0.3235	3.61
F13	0.2424	2.35	M13	0.1728	1.69
F15	0.4980	3.69	M15	0.3262	3.77

VII. DISCUSSION

Identification of Gender Speech Recognition is as yet a difficult zone in acknowledgment, grouping and recovery .numerous examines have done in speech acknowledgment, .As found in the end area, Fuzzy rationale was end up being a fruitful system in both acknowledgment and recovery tasks. Fuzzy rationale is where valid and bogus choices are taken dependent on a recognizer framework. The primary system was utilized in perceiving among male and females sample and it shows a decent achievement rate however to expand the acknowledgment capacity of this system, more features from speech signal were required to be characterize ,so spectrogram investigation which utilized in speech examination inquires about as in [4] ,[11] was extremely useful examination in characterizing two other vigorous features that characterize extra standards to the fuzzy interface system utilizing MATLAB 7.6 tool and builds the system security and authority against unapproved speakers that are not permitted to enter the system or to indicated work zone. Fig (3) shows the Identification of gender speech recognition model by collection of speech signals samples.



Fig(3) Identification of gender speech recognition model

REFERENCES

- [1] Chester ,F.J ,Taylor ,and M.Doyle , "The WingerDistribution In Speech Processing Applications", J.Franklin Inst.,vol.318 ,pp. 415 – 420 ,1984.
- [2] Ho, K. Lai, and, Octavian, C., " Speech ProcessingWorkshop", Department of Electrical and Electronic Engineering, Part IV Project Report 2003.
- [3] Jere, B. , " Digital Signal Processing Application using the ADSP-2100 Family ", Application Engineering staff of Analog Devices DSP Division Edited ,PTR ,Prentice Hall ,New Jersey ,p.330-p.331.
- [4] Brain, J. Love, and Jennifer, V. and, Xuening, S., "Automatic Speaker Recognition Neural Networks", Electrical and Computer Engineering Department, the University of Texas at Austin, Spring 2004.
- [5] Abdul-Bary, R., S., Soad, T. A., "MATLAB-Based Design and Implementation of Time-Frequency Analyzer", College of Electrical Eng., Technical Institute.
- [6] M.Liu, C.Wan, L.Wang, " Content-Based Audio Classification and Retrieval using a Fuzzy Logic System towards Multimedia Search Engines".
- [7] The Math Work, Inc. "MATLAB the Language of Technical Computing ", Version .7.6.0(R2008a), USA. [8]. Christos, S. and, Dimitrios, S., "Neural Networks",
- [8] John-Paul, H., and Ron, C. and Mark, F. and John ,S. , and Yonghong ,Y. and Wei ,W. , "Training Neural Networks for Speech Recognition", Center for spoken language understanding , (CSLU), Organ Graduate Institute of Science and Technology.
- [9] Patrick.M. Mills, "Fuzzy Speech Recognition", Bachelor of Science in Engineering, Swathmore College, 1994, University of South Carolina, Thesis 1996.
- [10] Dunia, J.,Jamma, "Low Cost Wavel et Analyzer", Computer Engineering Department, Technical College/ Mosul, Foundation of Technical College.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)