

Speaker's Gender Identification Depends on Speech using Fuzzy Logic

Atif Khan¹, Vikas Kumar², Santosh Kumar³

P.G. Student, Department of Electronics and Communication Engineering, Invertis University, Bareilly,
UP, India¹

Associate Professor, Department of Electronics and Communication Engineering, Invertis University, Bareilly,
UP, India²

Associate Professor, Department of Electronics and Communication Engineering, Invertis University, Bareilly,
UP, India²

ABSTRACT: Identification of gender is one of the most important process in speech processing nowadays. Generally identification of gender is based on pitch value of the speech signal. Mostly pitch value of female is higher than male. But in some cases potch value of male is higher than female. In that case this gender classification method does not produce the exact required result. Due to this above mentioned problem here we have proposed a new method of gender identification considering four features. This new method is performed by using fuzzy logic on the working platform of MATLAB. Now fuzzy logic is trained by using the training dataset generated by these four features. The implementation results have shown the proposed technique performance in gender identification.

KEYWORDS: Gender identification, fuzzy logic, power amplitude, spectrum power, total harmonic distortion, Pitch.

I. INTRODUCTION

For communication between humans speech is one of the common methods in modern civilized societies [1]. The speaker's different ideas are communicated in the form of words, phrases and sentences by speech applying some proper grammatical rules [2]. We can consider the sound segmentation of speech by dividing it in voiced, unvoiced and silence.

Speech processing is known as the study of speech signals and the various methods used to process these speech signals. By applying electronic circuits and computers the linguistic information is determined and extracted by a speech wave [3]. Gender classification is applied in many fields. It is applied in several applications such as in speech indexing, synthesis, recognition and retrieval of multimedia database etc. This process is performed for various applications such as for security devices, cellular phones, ATM machines, household appliances and computers [4]. It turns out to be a complex and challenging problem [5]. The differences in speaking style of humans, vocal fold thickness or vocal tract length are the reason of gender based differences in human speech and these physiological differences can be identified [6,7].

Here we have proposed a new method for gender identification due to certain drawbacks of the previous research method. The new method considering four features where as the previous method considers pitch as feature. The new method is implemented on the working platform of MATLAB using fuzzy logic. To work with speech signal is very challenging because it is highly redundant and non-stationary signal. Speech recognition system can be divided into two categories namely: 1. Speaker independent systems that can be used by any person. 2. Speaker dependent systems that can be often used and trained by one person only. Speaker recognition can be classified into two categories namely: 1. Identification. 2. Verification. The first one determines who is speaking and the second one determines (Is the speaker who we think he or she is?).

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

The tools we used in the proposed method are the MATLAB and fuzzy logic toolbox. Many systems were designed based on this MATLAB tool and it is also used to process speech signals [8].

Rest of the paper is structured as: In section 2 the related works are reviewed in brief. In section 3 the proposed technique with satisfactory mathematical models and examples are elaborated. In section 4 the obtained implementation results are discussed. The decision is made in the section 5 of the paper.

II. RELATED WORK

In recent times some of the research works associated with classification of speech are discussed as follows. Rakesh et al. [9] by using numerous speech processing techniques have considered two different models. One of the two models is used to generate pitch value of the voice sample and the second one is used to generate formant values of the voice sample. These models are employed to extract the speaker's pitch value and gender based features. We can calculate the mean of pitch and formants of the entire speaker's sample by employing a model having loops and counters which perform a mean of formant 1 and pitch value of the speaker. By calculating Euclidean distance from the mean value of males and females of the generated mean values of formant1 and pitch by using a nearest neighbor technique the speaker is categorized between female and male. The algorithm is carried out in real time, by practicing NI Lab VIEW. Devi et al. [10] have suggested a fuzzy ARTMAP network and a modified fuzzy ARTMAP network to classify different background noise signals. In their study have discussed that due to background noise from noisy environment, the performance of speech processing systems like speech coding and speech recognition have been reduced. For example street noise, factory, helicopter, car, bus etc. Therefore the performance of the speech recognition system has been improved by improving the classification of noise. The key process in the design of a signal classification system is to effectively separate the signals on the basis of the selection of fine set of features. In order to reduce the effect of environmental noises on speech processing tasks, noise classification is a fundamental process.

Sigmund [11], in a short segment of normally spoken continuous speech have suggested an approach for automatic identification of gender. To observe which phonemes are useful for classification of gender they have studied all the vowels individually. And have also computed two different simple identifiers which are based on the selected Mel-Frequency Cepstral Coefficients (MFCCs). Now for gender identification more than 90% of efficiency has been achieved in a very short-time analysis (20 msec) by using the vowel phonemes. particularly there was no error for vowel "a". For text-independent analysis the duration of speech of 500 msec is enough to identify female/male speakers with the precision of more than 93%. For achieving high-quality dialogue systems automatic prediction of gender by her/his voice has been an important aspect.

Mahdi and Jafer [12], for unvoiced and voiced classification of speech segments have suggested wavelet based algorithm. This classification process is categorized into two steps:

1. The energy-frequency distribution of various speech signals has been statistically analysed by using wavelet transform.
2. The short-time zero-crossing rate of the signal has been calculated.

The ratio of the average energy in the low-frequency wavelet sub bands has been calculated for each time segment of the pre-emphasized speech in comparison to that of highest-frequency wavelet sub band by using a 4-level dyadic wavelet transform. And then it has been compared to a pre-decided threshold. Now the classification decision has been obtained by using an experimentally confirmed criterion which depends upon the result of comparison process.

Silovsky and Nouza [13], for automatic recording of broadcast programs in a system have considered a set of methods in his research to categorize various audio segments in a system. Their task is to conclude:

1. Whether the segment have to be marked as voiced or as unvoiced and also in the preceding case.
2. Whether the speaker is one of the speakers in the data base.
3. If not, the talking person belongs to which gender.

The information achieved from record system has been extended and the performance of speech recognition system has been improved by using the classification result. Their proposed method is based on GMM like all the other modern speaker recognition systems. They have developed a method that speed up the recognition process in a major way due

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

to large number of database speakers. While reviewing these recent researches which have discussed the same problem, in most of the researches gender is identified on the basis of pitch and in some other researches gender is classified on the basis of other statistical features. To test their methods, in some researches emotional speech dataset is used, while in some other continuous speech dataset is used and in some of the researches any words speech dataset was used.

III. GENDER CLASSIFICATION USING FUZZY LOGIC

Gender identification plays a vital role in speech processing. Speaker's gender is identified by using this technique. Gender identification is done by various techniques. But the problem is that most of these techniques depend on pitch value. The pitch typically depends on the sound's frequency. Commonly pitch of female is higher than pitch of male but in some cases pitch of male is higher than female. This is the reason gender identification using pitch as feature does not produce proper results. To overcome this problem here we have suggested a new method for gender identification using four features namely; power amplitude, spectrum power, total harmonic distortion and pitch. Firstly these four features are given as an input to the fuzzy logic and then in the output it gives the percentage of male and female feature. The process of gender identification is done in this way. In below sections the process which takes place in the Suggested method is explained briefly. Now, the features which are used in our method have discussed firstly.

3.1. Feature Analysis for Speech Signals

In gender identification feature selection plays an important. The identification of gender depends on these features. The four features used in our method are as discussed below:

3.1.1. Power amplitude

Power is the amount of energy consumed per unit time. It is the rate of doing work. Since it has no direction it is a scalar quantity. The SI unit of power is joule per second (J/s), known as watt. From the testing results we observed that the value of power amplitude is high for males and for females it is low.

3.1.2. Power spectrum

Power spectrum of a signal exhibits the role of every frequency of the spectrum to the power of the overall signal. From the testing results we have observed that the spectrum power for male is low where as for female it is high.

3.1.3. Total harmonic distortion

Total harmonic distortion, or THD, of a signal is defined as the ratio of distortion power to the signal power. From the testing results we observed that the THD for female speech is higher than that of the male speech. In the above sections the features we used in our method are explained. The next process in the proposed method is to identify the percentage of female and male feature present in the given speech signal by using fuzzy logic. Table below shows the values of the three features for male and female samples.

3.1.3. Pitch

In the sense associated with musical melodies it is possible to judge sounds as "higher" and "lower" because of the quality of pitch. We can determine pitch only in sounds that have a frequency that is clear and fixed enough to distinguish from noise. Pitch is a major sonic property of musical tones, along with duration, loudness and rhythm.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

Table (1): Values of the inputs for male and female recorded samples

Recorded speech	Power amplitude (Watt)	Spectrum power (db)	Total harmonic distortion	Pitch (Hz)
F1	0.00071947	-7838900	19.3628	347.8261
F2	0.00032606	-18587000	8.9833	2000
F3	0.00062043	-19575000	12.6998	1600
F4	0.00050457	-18177000	38.8943	2666.7
F5	0.00060963	-26534000	13.7608	1000
F6	0.00083228	-24146000	18.979	250
F7	0.00076215	-22702000	11.382	533.3333
F8	0.000905	-8082000	7.2447	888.8889
F9	0.00058168	-8375100	10.4355	320
F10	0.00091456	-7998000	17.681	1600
F11	0.00069989	-20428000	37.9648	333.3333
F12	0.00065209	-25679000	20.3866	571.4286
F13	0.00099794	-5945700	9.4794	347.8261
F14	0.00087153	-7743700	6.2683	380.9524
F15	0.00062134	-5705000	6.0127	320
F16	0.00075488	-8278100	10.8918	347.8261
F17	0.0007131	-8414400	17.2326	666.6667
F18	0.001	-8153200	5.0871	938.2979
F19	0.00087716	-7413200	7.7817	
F20	0.00057489	-19348000	12.211	800
F21	0.00028207	-8978900	11.8113	1000
M1	0.0031	-21331000	1.3889	210.5263
M2	0.0022	-32573000	1.3882	666.6667
M3	0.0011	-29870000	2.3883	500
M4	0.0014	-32155000	1.3113	444.4444
M5	0.0014	-31000000	1.6174	242.4242
M6	0.0016	-30233000	11.1365	235.2941
M7	0.0027	-33994000	5.1485	363.6364
M8	0.0015	-23466000	1.8453	228.5714
M9	0.0038	-28983000	1.0136	135.5932
M10	0.00094268	-19716000	1.708	200
M11	0.00086686	-32117000	1.382	250
M12	0.0011	-33450000	9.588	137.931
M13	0.0022	-9052000	2.9548	156.8627
M14	0.0021	-20665000	0.8916	242.4242
M15	0.0023	-9084200	4.5345	250
M16	0.0022	-23475000	1.8146	163.2653
M17	0.0026	-21682000	1.9205	195.122
M18	0.0017	-21696000	1.4566	320
M19	0.00089398	-24407000	1.1613	580.2632
M20	0.0015	-20093000	3.3368	615.3846
M21	0.00099569	-19721000	2.8312	228.5714

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

IV. IDENTIFYING FEMALE AND MALE FEATURE USING FUZZY LOGIC

A number of unique parameters offer by fuzzy logic in many control problems which alternatively produce better results in many control problems [14]. Here the percentage of a variety of female and male features which is present in the given speech signals is evaluated by fuzzy logic. In general there are three key steps in fuzzy logic. These are fuzzification, generating fuzzy rules and defuzzification. The process of converting system data into fuzzy data is known as fuzzification. We have used triangular and trapezoidal membership functions in the fuzzification process. The next step is to generate fuzzy rules. The structure of fuzzy logic used in the suggested method with four inputs and one output are as shown in figure 2 and The Membership function Editor of the speaker's gender identification system depends on speech is shown in figure 3.

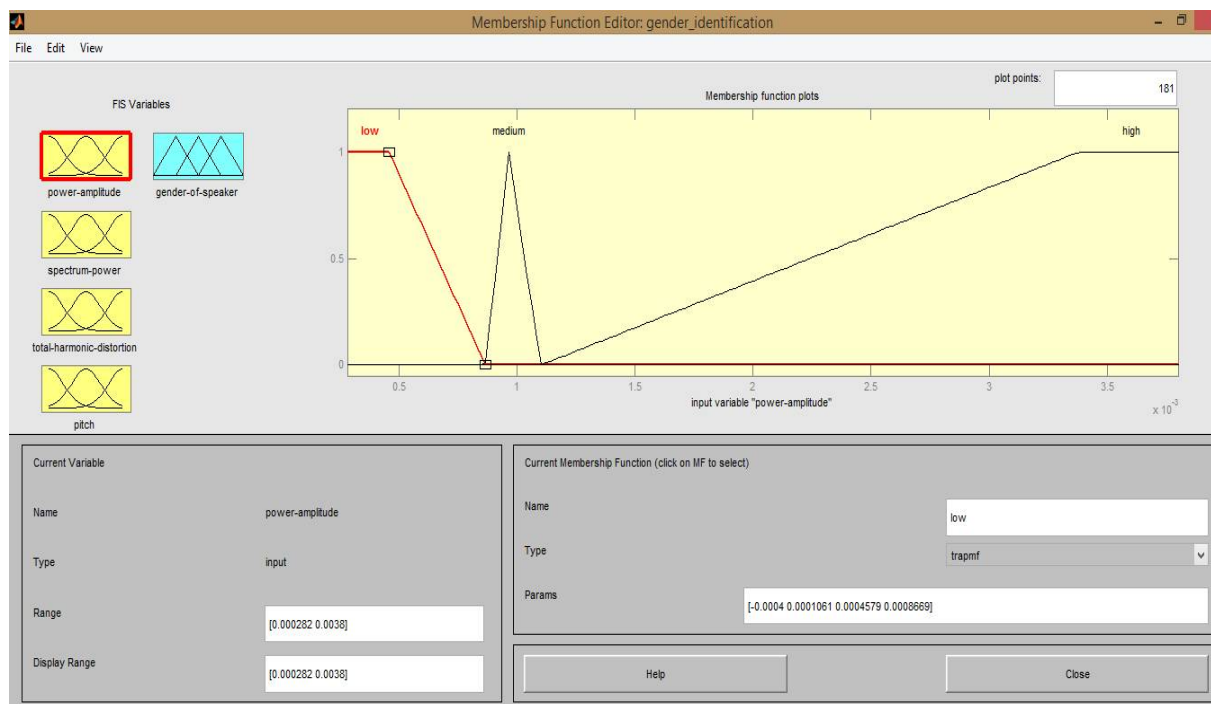


Figure 3. The Membership function Editor of the speaker's gender identification system depends on speech.

The above figure shows the membership function editor for the speaker's gender identification system depends on speech. In which the five membership functions for the power amplitude is shown.

V. FUZZY RULES GENERATION

The four inputs to the fuzzy logic are power amplitude, spectrum power, total harmonic distortion, pitch and the output obtained is the percentage of various female and male features present in the given speech signal. The inputs are fuzzified into three sets namely: small, medium large where as the output is fuzzified into three sets namely male, female/male, female. The fuzzy rules generated are as shown in figure 4 and the Rule Viewer of the speaker's gender identification system depends on speech is shown in figure 5. The next operation after this is to train fuzzy logic. These rules helps to generate fuzzy rules. To train fuzzy logic Training datasets will have to be generated. Now the trained fuzzy logic is used for practical operation. During testing if we will provide power amplitude, spectrum power, total

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

harmonic distortion and pitch as input to the fuzzy logic then it will give the output as the gender belongs to male or female.

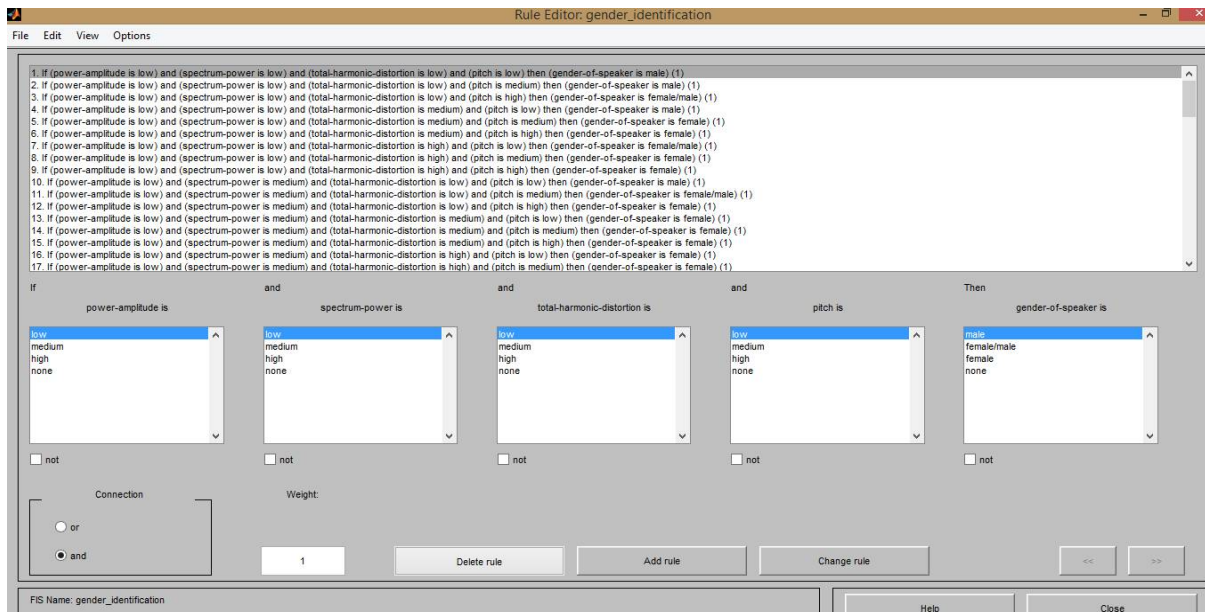


Figure 4. The Rule Editor of the speaker's gender identification system depends on speech.

The above figure shows the rule Editor of the speaker's gender identification system depends on speech in which we have generated 81 rules on the bases of which we can calculate the gender of the speaker.

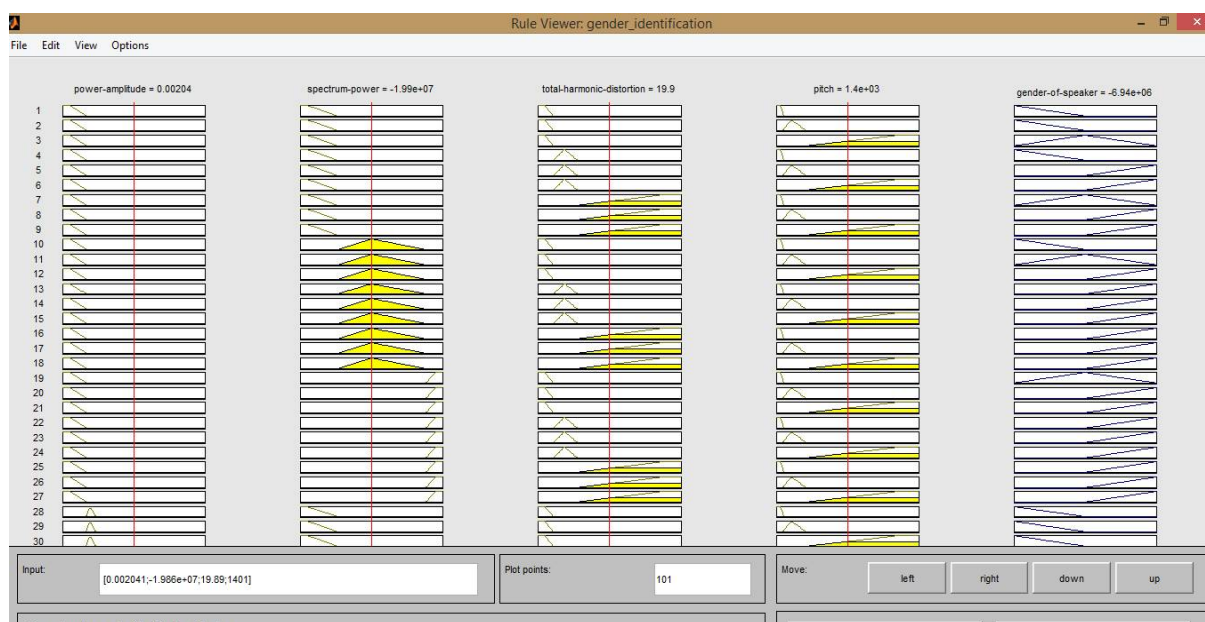


Figure 5. The Rule Viewer of the speaker's gender identification system depends on speech.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

The rule viewer of the speaker's gender identification system depends on speech gives the output value on the bases of these three input values. And then with the help of fuzzy rules we can calculate the gender of the speaker.

VI. RESULT AND DISCUSSIONS

The suggested technique is tested for a range of speech signals based on the existing MATLAB fuzzy logic toolbox demo in MATLAB 8.1.0.604 (R2013a). Here 42 speech signals are taken as an input. First and foremost the fuzzy logic is trained by using some speech signal and to identify the gender of the speaker a set of speech signal as input to the suggested method is used to perform the testing. The set of laws that were used in the system to identify the gender as mentioned above are:

1. If power of the speech signal is high then male speaks where as if power is low then female speaks. With range specified [0.000282 0.0038].
2. If spectrum power of the speech signal is high then female speaks and if low then male speaks. With range specified [-3.4e+07 -5.71e+06].
3. If total harmonic distortion of speech signal is low then male speaks and if THD of speech signal is high then female speaks. With range specified [0.8916 38.89].
4. If the pitch value of the speech signal is high then female speaks and if the pitch value of the speech signal is low then male speaks.

There is an error value may result like every system. The result section has shown that has turned out to be a successful system in identifying both male and female speakers. Fuzzy logic is a flexible system where the true and false decisions are taken by using membership functions. The Surface Viewer of the speaker's gender identification system depends on speech is shown in figure 6.

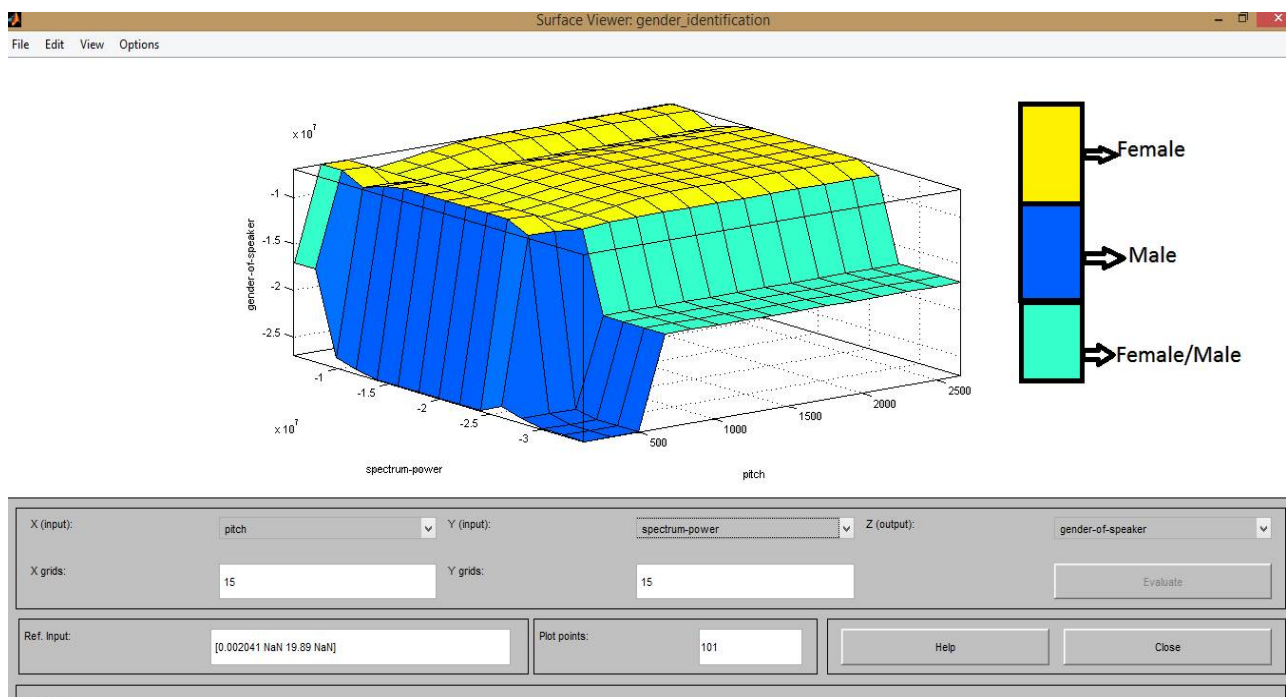


Figure 6. The Surface Viewer of the speaker's gender identification system depends on speech.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 8, August 2017

REFERENCES

- [1] Gudi A., Shreedhar H., and Nagaraj H., "Signal Processing Techniques to Estimate the Speech Disability in Children," IACSIT International Journal of Engineering and Technology, vol. 2, no. 2, pp. 169-176, 2010
- [2] Gudi A. and Nagaraj H., "Optimal Curve Fitting of Speech Signal for Disabled Children," International Journal of Computer Science & Information Technology, vol. 1, no. 2, pp. 99107, 2009.
- [3] Haraty R. and Ariss O., "CASRA+: A Colloquial Arabic Speech Recognition Application," American Journal of Applied Sciences, vol. 4, no. 1, pp. 23-32, 2007.
- [4] Patel I. and Rao S., "Speech Recognition using HMM with MFCC- an Analysis using Frequency Spectral Decomposition Technique," Signal & Image Processing : An International Journal, vol. 1, no. 2, pp. 101-110, 2010.
- [5] Kotti M. and Kotropoulos C., "Gender Classification in Two Emotional Speech Databases," in Proceedings of the 19th International Conference on Pattern Recognition, Tampa, pp. 1-4, 2008.
- [6] Shue Y. and Iseli M., "The Role of Voice Source Measures on Automatic Gender Classification," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, pp. 4493-4496, 2008.
- [7] Zengi Y., Wu Z., Falk T., and Chan W., "Robust GMM Based Gender Classification using Pitch and Rasta-PLP Parameters of Speech," in Proceedings of the 5th International Conference on Machine Learning and Cybernetics, Dalian, pp. 13-16, 2006.
- [8] Abdul-Bary, R., S., Soad, T. A., "MATLAB-Based Design and Implementation of Time-Frequency Analyzer", College of Electrical Eng., Technical Institute.
- [9] 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.
- [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.
- [12] Mahdi A. and Jafer E., "Two-Feature Voiced/Unvoiced Classifier Using Wavelet Transform," The Open Electrical and Electronic Engineering Journal, vol. 2, no. 1874-3005, pp. 8-13, 2008.
- [13] Silovsky J. and Nouza J., "Speech, Speaker and Speaker's Gender Identification in Automatically Processed Broadcast Stream," Radio Engineering Journal, vol. 15, no. 3, pp. 42-48, 2006.
- [14] Haider T. and Yusuf M., "A Fuzzy Approach to Energy Optimized Routing for Wireless Sensor Network," The International Arab Journal of Information Technology, vol. 6, no. 2, pp. 179188, 2009.