

Speech Based Gender Identification 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: In these times one of the most important processes in speech processing is the classification of gender. Typically classification of gender is based on considering pitch as feature. The pitch value of male is lower than the female. In most of the topical research works the process of the classification of gender is performed using the aforementioned condition. In some cases the pitch value of female is lower and also pitch of some male is high, in that case the exact required result is not being produced by this classification method. By considering the aforesaid problem we have here suggested a new method for gender classification which considers three features. In this new method the gender of the speaker is identified by using fuzzy logic. Now by using the above three features training data set is generated to train fuzzy logic. The performance of the proposed technique in gender classification is shown by the implementation result.

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

I. INTRODUCTION

A Speech is one of the common methods for communication between humans in modern civilized societies [1]. By applying some proper grammatical rules, the speaker's different ideas are communicated by speech in the form of phrases, words, and sentences [2]. The sound segmentation of speech can be considered by classifying it with voiced, unvoiced and silence.

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

In previous researches gender classification is done on the basis of pitch. But this feature has certain limitations. Here a new method for gender classification is proposed by considering three features using fuzzy logic. Speech signals can be considered as a non-stationary and high redundant signal. To work with speech signal is very challenging because of these properties. Speech recognition system can be classified into two categories namely: 1. Speaker dependent systems that often only one person can use and train. 2. Speaker independent systems that anyone can use. We can classify speaker recognition into two categories: identification and verification. Speaker identification determines who is speaking and is based on the speaker's speech information. And (Is the speaker who we think she or he is?) is determined by the speaker verification process.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

The tools which we used in the above mentioned work are the MATLAB and fuzzy logic toolbox. The MATLAB tool is used to process speech signals and many systems was designed based on this tool as in [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

Recently some of the research works related to classification of speech are discussed as follows. Rakesh et al. [9] by using several speech processing techniques have proposed two different models. One of the two models is used to produce pitch value of the voice sample and the other one is used to produce formant values of the voice sample. These two models are employed to extract the gender based features and speaker's pitch value. By employing a model having loops and counters which perform a mean of formant 1 and pitch value of the speaker, the mean of pitch and formants of the entire speaker's sample were calculated. By computing Euclidean distance from the mean value of females and males of the generated mean values of pitch and formant 1 by using a nearest neighbor technique the speaker is classified between female and male. The algorithm is performed in real time, by practicing NI Lab VIEW.

Devi et al. [10], have proposed a fuzzy ARTMAP network and a modified fuzzy ARTMAP network to categorize different background noise signals. In their study have discussed that the performance of speech processing systems like speech recognition and speech coding have been reduced due to background noise from noisy environment. For example street noise, factory, helicopter, car, bus etc. Therefore to improve the performance of the speech recognition system the classification of noise is need to be improved. In the design of a signal classification system the key process is to efficiently separate the signals on the basis of the selection of fine set of features. Noise classification is a fundamental process in order to lessen the effect of environmental noises on speech processing tasks.

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

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

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

The ratio of the average energy in the low-frequency wavelet sub bands has been evaluated 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-determined threshold. Now to obtain the classification decision, an experimentally confirmed criterion which depends upon the result of comparison process has been used.

Silovsky and Nouza [13], for automatic recording of broadcast programs in a system have proposed a set of methods in his research to classify 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 obtained 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 accelerates the recognition process in a significant way due to large number of database speakers. While reviewing these recent researches which have discussed the same

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

problem, in most of the researches gender is classified on the basis of pitch and in some other researches gender is recognized 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 also considered.

III. GENDER CLASSIFICATION USING FUZZY LOGIC

In speech processing gender classification plays a key role. The gender of the speaker is identified by means of this technique. There are different-different methods used for gender recognition. But the main problem is that most of these works depends on pitch value. The pitch generally depends on the frequency of sound. In some of the cases the pitch of female is high and pitch value of male is low. While in some cases the pitch value of male is higher than female and pitch of female is lower than male. In these situations gender recognition using pitch as feature will not produce proper results. Due to this drawback here we have proposed a new method for gender classification using three features namely; power amplitude, spectrum power, and total harmonic distortion. Firstly these three features are computed and given as an input to the fuzzy logic and in the output it gives the percentage of female and male feature. In this way the process of gender classification is done. In the below sections the process which takes place in the proposed method is explained briefly. Now at first, the features which are used in our method have discussed.

3.1. Feature Analysis for Speech Signals

The gender classification depends on the features that we have selected in our method. Features selection plays a vital role in speaker recognition. The three 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.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

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

Recorded speech	Power amplitude (Watt)	Spectrum power (db)	Total harmonic distortion
F1	0.00071947	-7838900	19.3628
F2	0.00032606	-18587000	8.9833
F3	0.00062043	-19575000	12.6998
F4	0.00050457	-18177000	38.8943
F5	0.00060963	-26534000	13.7608
F6	0.00083228	-24146000	18.979
F7	0.00076215	-22702000	11.382
F8	0.000905	-8082000	7.2447
F9	0.00058168	-8375100	10.4355
F10	0.00091456	-7998000	17.681
F11	0.00069989	-20428000	37.9648
F12	0.00065209	-25679000	20.3866
F13	0.00099794	-5945700	9.4794
F14	0.00087153	-7743700	6.2683
F15	0.00062134	-5705000	6.0127
F16	0.00075488	-8278100	10.8918
F17	0.0007131	-8414400	17.2326
F18	0.001	-8153200	5.0871
F19	0.00087716	-7413200	7.7817
F20	0.00057489	-19348000	12.211
F21	0.00028207	-8978900	11.8113
M1	0.0031	-21331000	1.3889
M2	0.0022	-32573000	1.3882
M3	0.0011	-29870000	2.3883
M4	0.0014	-32155000	1.3113
M5	0.0014	-31000000	1.6174
M6	0.0016	-30233000	11.1365
M7	0.0027	-33994000	5.1485
M8	0.0015	-23466000	1.8453
M9	0.0038	-28983000	1.0136
M10	0.00094268	-19716000	1.708
M11	0.00086686	-32117000	1.382
M12	0.0011	-33450000	9.588
M13	0.0022	-9052000	2.9548
M14	0.0021	-20665000	0.8916
M15	0.0023	-9084200	4.5345

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

M16	0.0022	-23475000	1.8146
M17	0.0026	-21682000	1.9205
M18	0.0017	-21696000	1.4566
M19	0.00089398	-24407000	1.1613
M20	0.0015	-20093000	3.3368
M21	0.00099569	-19721000	2.8312

IV. IDENTIFYING FEMALE AND MALE FEATURE USING FUZZY LOGIC

In many control problems fuzzy logic offers a number of unique parameters which alternatively produce better results [14]. Here the percentage of various female and male features present in the given speech signals is evaluated by fuzzy logic. Normally there are three important steps in fuzzy logic. These are fuzzification, generating fuzzy rules and defuzzification. Fuzzification is the process of converting system data into fuzzy data. In the fuzzification process we have used triangular and trapezoidal membership functions. After this the next step is to generate fuzzy rules. The structure of fuzzy logic used in the proposed method with three inputs and one output are as shown in figure 2 and The Membership function Editor of the speech based gender identification system is shown in figure 3.

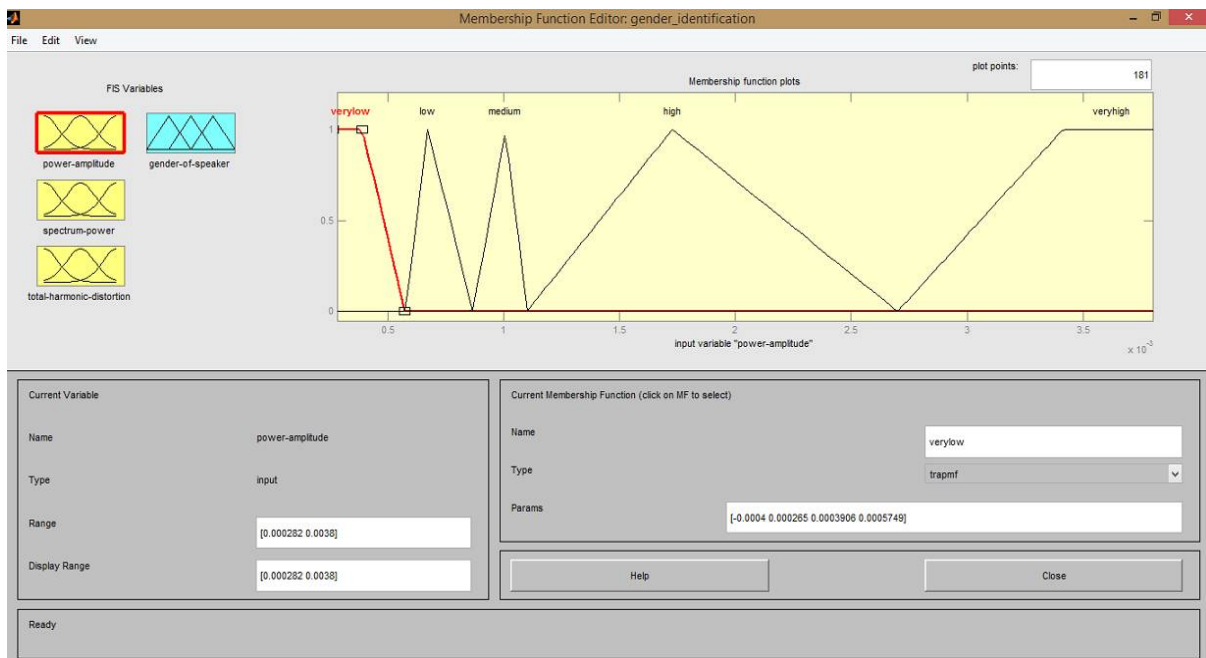


Figure 3. The Membership function Editor of the speech based gender identification system.

The above figure shows the membership function editor for the speech based gender identification system. In which the five membership functions for the power amplitude is shown.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

V. FUZZY RULES GENERATION

The input to the fuzzy logic is power amplitude, spectrum power, total harmonic distortion and the output obtained from fuzzy logic is the percentage of various female and male features present in the given speech signal. The five different sets in which inputs are fuzzified are named as; very small, small, medium, large, very large where as the three sets in which output variable is fuzzified are named as; female, male, male/female. In male/female the speaker is either female or male. The generated fuzzy rules are as shown in figure 4 and the Rule Viewer of the speech based gender identification system is shown in figure 5. The next process after this is to train fuzzy logic. Fuzzy logic is trained by using these rules. Training datasets will have to be generated to train fuzzy logic. Now this trained fuzzy logic is used for practical operation. During testing if we will provide power amplitude, spectrum power and total harmonic distortion as input to the fuzzy logic then it will give the output as the feature belong to female or male.

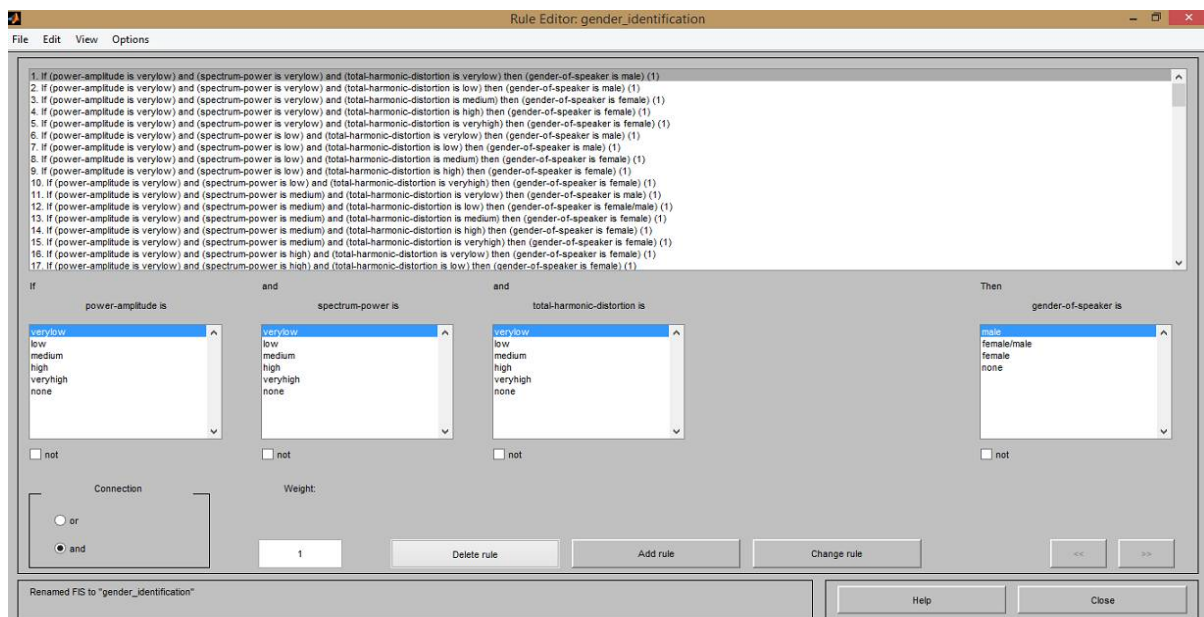


Figure 4. The Rule Editor of the speech based gender identification system.

The above figure shows the rule Editor of the speech based gender identification system in which we have generated 125 rules on the bases of which we can calculate the gender of the speaker.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

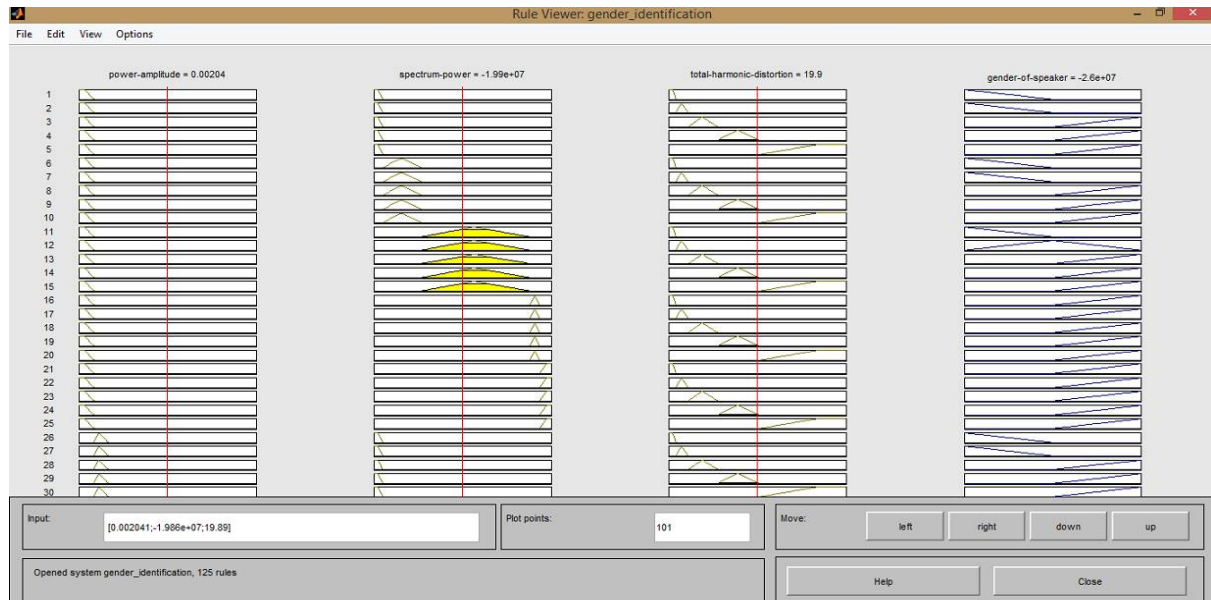


Figure 5. The Rule Viewer of the speech based gender identification system.

The rule viewer of the speech based gender identification system 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

This proposed technique is tested for various speech signals based on the existing MATLAB fuzzy logic toolbox demo in MATLAB 8.1.0.604 (R2013a). Here 21 speech signals are taken as an input. Primarily the fuzzy logic is trained by using some speech signal and to recognize the gender of the speaker a set of speech signal as input to the proposed method is used to perform the testing. Three rules were used in the system to identify the speaker as mentioned above.

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].

There is an error value may result like every system. As we see in the result section fuzzy logic was turned out to be a successful system in recognizing male and female speakers. Fuzzy logic is an adjustable system where the membership functions are used to take true and false decisions. The Surface Viewer of the speech based gender identification system is shown in figure 6.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

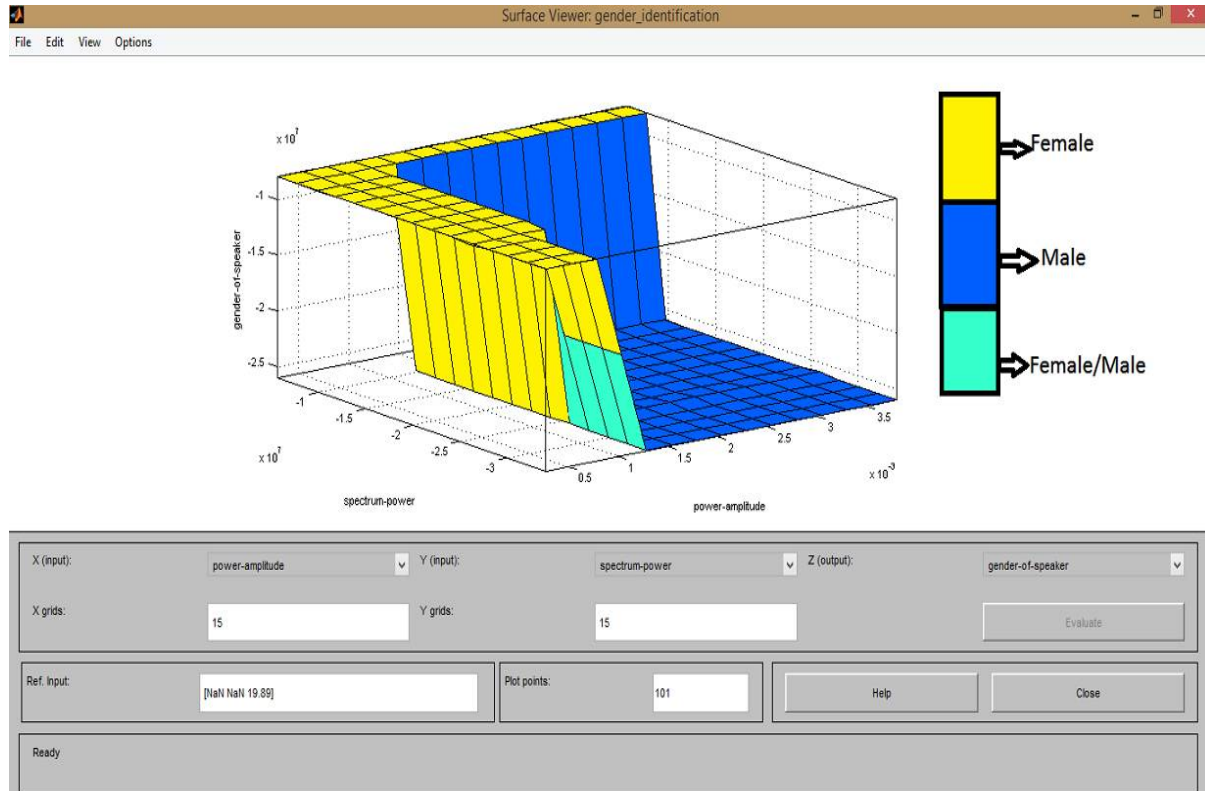


Figure 6. The Surface Viewer of the speech based gender identification system

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.