**Gender Recognition using Speech Processing Techniques**

**Introduction:**

Gender Detection plays a key role in Human Computer Interaction but the ultimate focus is about on what basis the male and female can be discriminated and how it can be done. Traditionally, the interest in voice-gender conversion was of a more theoretical nature than founded in real-life applications. But, with the increase in biometric security applications, mobile and automated telephonic communication and the resulting limitation in transmission bandwidth, practical applications of gender recognition have increased many folds.

In contrast to speech recognition which recognizes the spoken words of individual but doesn’t emphasize on the voice characteristics of individuals, **Voice recognition** is playing an important role in many electronic devices to recognize the voice of individuals. Various speech processing techniques and algorithms are used for generating the Formant values of the voice sample and for the generation of pitch value for the voice sample. Physiological differences such as vocal fold thickness or vocal tract length and the speaking accents help us find gender-based differences.

**Benefits:**

* Facilitating automatic speaker recognition by cutting the search space in half, therefore reducing computations and enhancing the speed of the system.
* Enhance speaker adaptation as a part of an automatic speech recognition system.
* Sorting telephone calls by gender for gender sensitive surveys.

**Literature Review:**

We have used some of the speech processing techniques that have been developed over time to develop the gender recognition system.

**Framing:**

Framing is implemented on the speech signal. The recorded signal has a finite length. The signal again is framed into pieces of same length. The vocal tract is not able to change its shape faster than fifty times in a second, so we can assume the signal will be stationary for 20 milliseconds. Overlapping of individual frames is used so as to increase precision of recognition process.

**Windowing:**

The signals that are divided into equal sized frames are windowed to develop non-zero values of specific frequencies.

**Pitch Detection using Auto Correlation:**

The short speech segments, consisting of only N samples obtained from framing are operated with a short-time autocorrelation function to detect the highest value of Auto Correlation Function (ACF).

The auto correlation function between times s and t is generally defined as


R(s,t) = \frac{\operatorname{E}[(X_t - \mu_t)(X_s - \mu_s)]}{\sigma_t\sigma_s}\, ,


**Formant Detection:**

The resonances of the vocal tract are called formants, The frequencies at which the formants occur are primarily dependent upon the shape of the vocal tract, which is determined by the positions of the articulators.

In continuous speech the formant frequencies vary in time with the change in position of articulators.

1. **Methodology:**

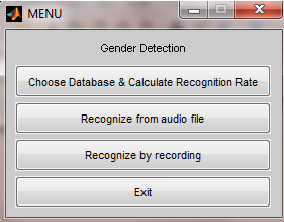
**Tools:** MATLAB 7.9.0

**Database:** CMU\_ARCTIC speech synthesis databases

1. **Experiments/Results:**

It is known that the fundamental frequency of an adult male is 85Hz to 180Hz and a typical adult female frequency varies from 165Hz to 255Hz. Thus based upon these values pitch detection and Formant analysis is performed on the voice signals and classification is made.

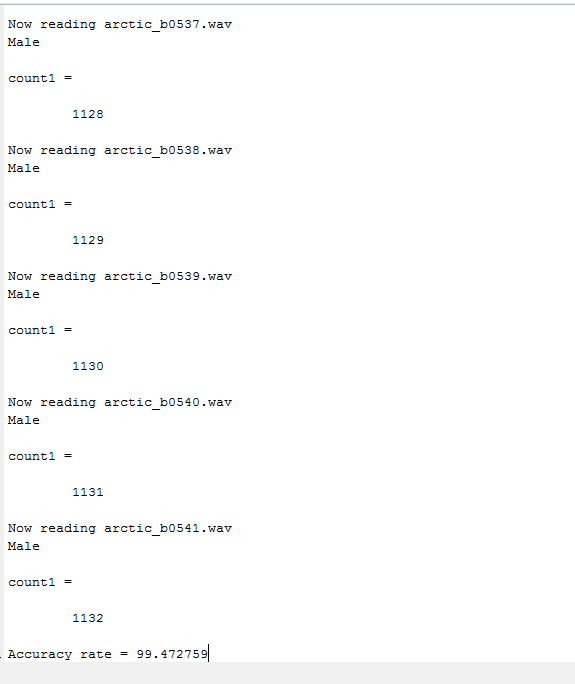
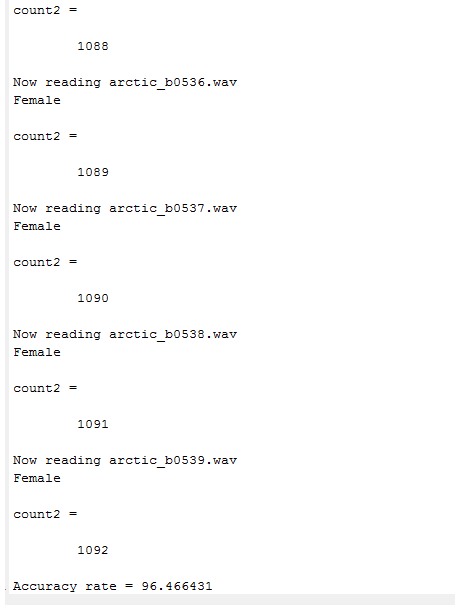
The GUI provides us with an option to choose whether to analyze the database, Recognize the gender on feeding an audio file or from recording the audio signals dynamically and analyze the gender of the speaker. It is shown in Fig 1

****

**Fig 1: GUI**

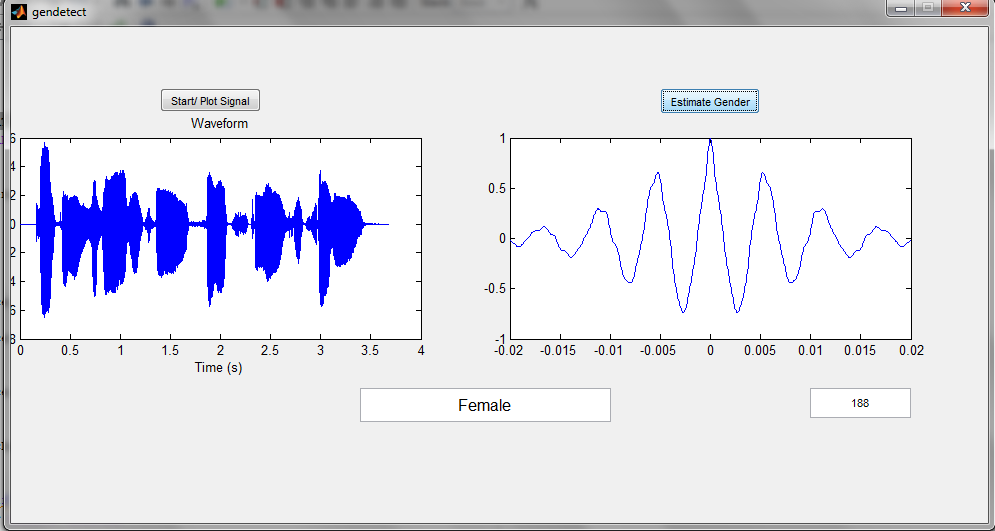
On choosing the ‘Choose Database & Calculate Recognition Rate’ a pop up appears to choose the directory of wav files on which the experiment has to be performed.

On successful selection the gender detection is performed on the particular wav files and the accuracy rate of gender classification is returned as shown below.

****

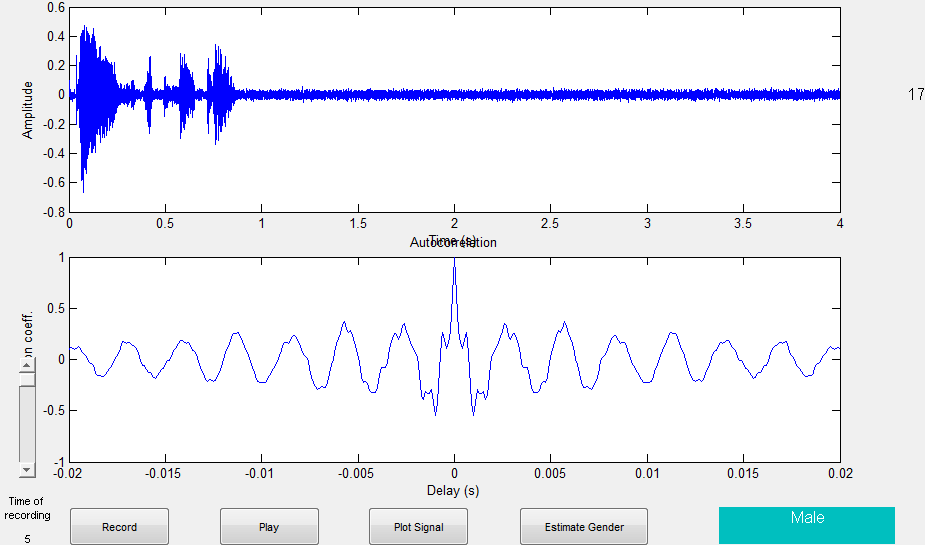
**Fig 2**

On choosing Recognize from audio file a GUI appears as shown below, Thus on selecting a specific wav file it interprets the wav file and returns the gender of the speaker in the audio file as shown in the figure.

****

**Fig 3**

Finally on selection of ‘recognize by recording’ a GUI is prompted which allows you to record the voice of the speaker and perform gender detection on the recorded file.

****

**Fig 4.**

**Conclusion:**

From the efficiency of the results obtained it is concluded that the algorithm implemented in MATLAB for gender detection is working successfully.

**References:**

1. <http://www-scf.usc.edu/~audhkhas/audhkhasi_fp_icassp09.pdf>
2. <http://www.ijcaonline.org/volume10/number3/pxc3871976.pdf>
3. <http://www.speech-recognition.de/matlab-examples.html>
4. <http://www.festvox.org/cmu_arctic/>