Fundamental Frequency Analysis of Bodo Vowels

Dr. Jogen Boro

Abstract— A successful speech recognition system can help in many applications and environments in our daily life. It can help in non critical operations such as presenting the driving route to the driver, dialing a phone number, light switch turn on/off, the coffee machine on/off etc. apart from speaker verification-caste wise, community wise and locality wise including *identification of sex*. In this paper an attempt has been made to find out the Fundamental Frequency of Bodo (Vowels) Phonemes and observe the spectrogram and bifurcation of the same through the utterances of native Bodo speakers. These studies will help to extract the features need for speech/speaker recognition of Bodo Language. Bodo is a local language of the North-East India.

Index Terms— Bodo Language, Bifurcation, Fundamental Frequency, Pitch, Speech Recognition, Spectrogram

I. INTRODUCTION

The opening and closing of the vocal folds that occur during speaking break the air steam into chains of pulses. The rate of repetition of these pulses is the **pitch** and it defines the fundamental frequency of the speech signal [1]. In other words, the rate of vibrations of the vocal folds is the fundamental frequency of the voice. The frequency increases when the vocal folds are made taut. Relative differences in the fundamental frequency of the voice are utilized in all languages to study the various aspects of linguistic information conveyed by it [2].

The general problem of fundamental frequency estimation is to take a portion of signal and to find the dominant frequency of repetition. Thus, the difficulties that arises in the estimation of fundamental frequency are (i) all signals are not periodic, (ii) those are periodic may be changing in fundamental frequency over the time of interest, (iii) signals may be contaminated with noise, even with periodic signals of other fundamental frequencies, (iv) signals which are periodic with interval T are also periodic with interval 2T, 3T etc., so we need to find the smallest periodic interval or the highest fundamental frequency, and (v) even signals of constant fundamental frequency may be changing in other ways over the interval of interest.

In general, the fundamental frequency of the speech wave is estimated using autocorrelation. The mathematical model used for estimating the fundamental frequency is given below [3]:

Dr. Jogen Boro, Guest Faculty, Department of Bodo, Gauhati University, Assam, India, 781014

A discrete short-time sequence is given by

$$s_n[m] = s[m]w[n-m] \qquad (1)$$

Where w[n] is an analysis window of duration N_W. The short-time autocorrelation function $r_n[\tau]$ is defined by

$$s_n[m] = s[m]w[n-m]$$

$$= \sum_{m=-\infty}^{\infty} s_n[m]s_n[m+\tau] \qquad (2)$$

Where s[m] is periodic with period p, $r_n[\tau]$ contains peak at or near the pitch period p. For unvoiced sound no clear peak occurs near an expected pitch period. Location of the peak in the pitch period range provides a measure of pitch estimation and voicing decision. The above correlation pitch estimator can be obtained, more formally, by minimizing over possible pitch periods (p>0), and the error criterion is given by:

$$E[p]$$
$$= \sum_{m=-\infty}^{\infty} (s_n[m] - s_n[m+\tau])^2 \qquad (3)$$

Minimizing E[p] with respect to p yields

$$\hat{p} = \frac{max}{p}\left(\sum_{m=-\infty}^{\infty} s_n[m]s_n[m+p]\right) \quad (4)$$

Where p >ε, i.e., p is sufficiently far from zero.

This alternative view of autocorrelation pitch estimation is used for detecting the pitch of Bodo vowels. The speech waveform and corresponding pitch spectra of six Bodo vowels have been depicted in *Figure (1), Figure (2)* for both male and female informants. The estimated values of the fundamental frequency or pitch have been given in *Table (1)* for Bodo vowels.

Table 1: Pitch of six Bodo vowels

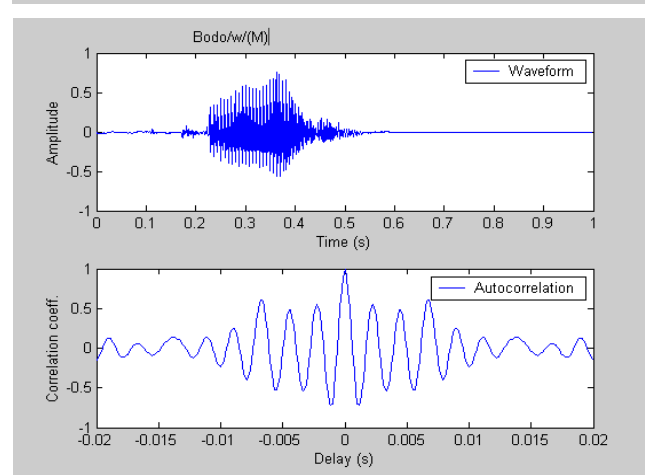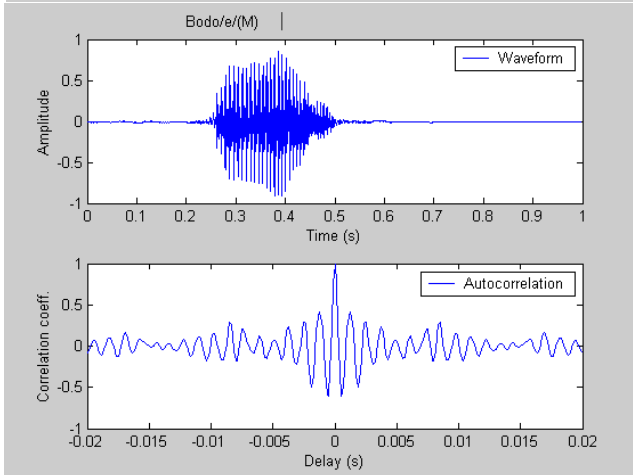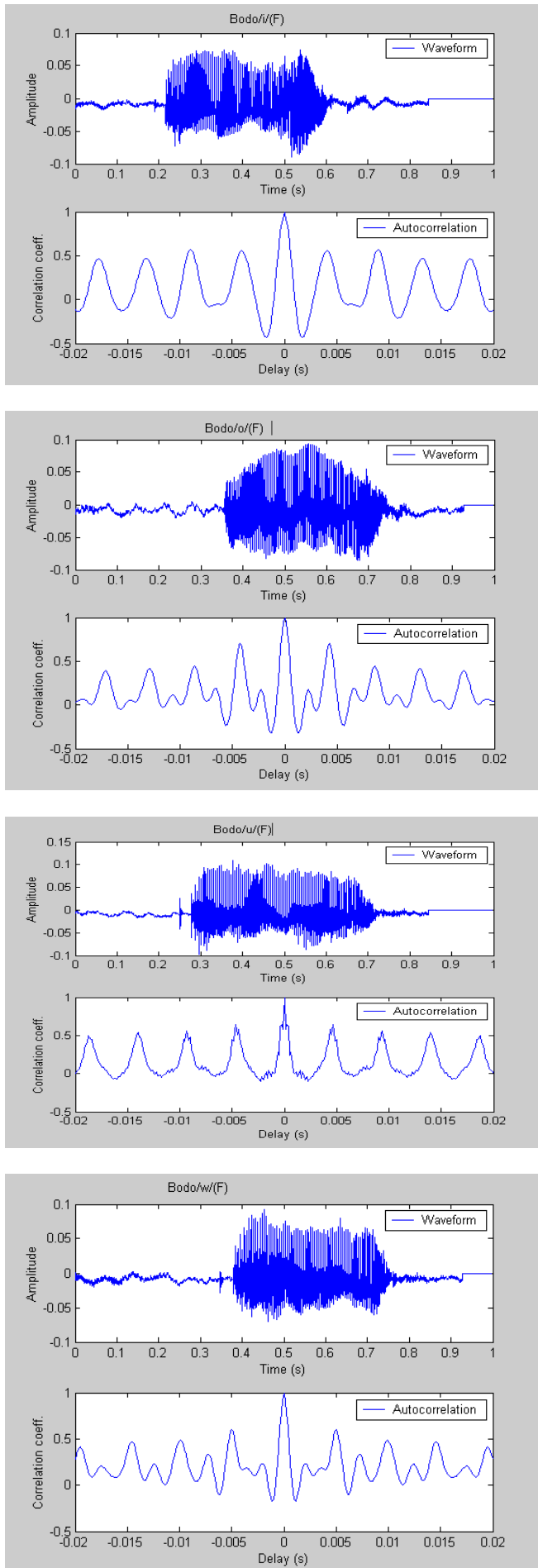| SPECIMEN	FUDAMENTAL FREQUENCY (in Hz.)					
	/a/	/e/	/i/	/o/	/u/	/ɯ/
MALE	130.001	242.049	148.200	154.004	126.003	146.030
FEMALE	229.000	381.001	268.250	250.900	118.500	251.002
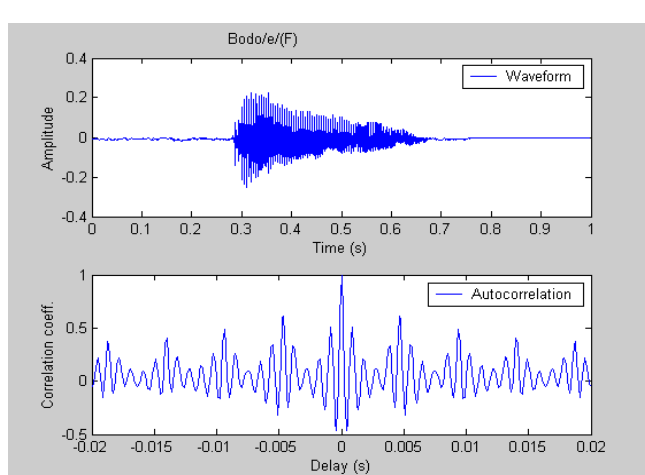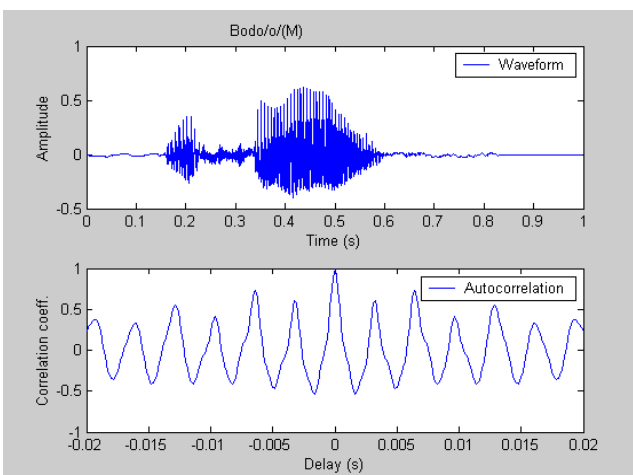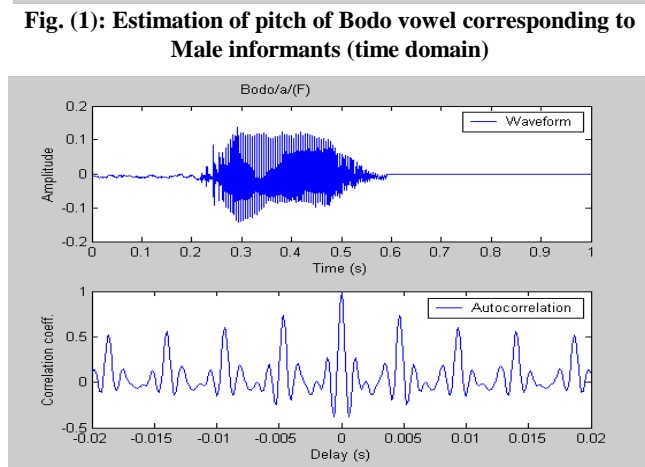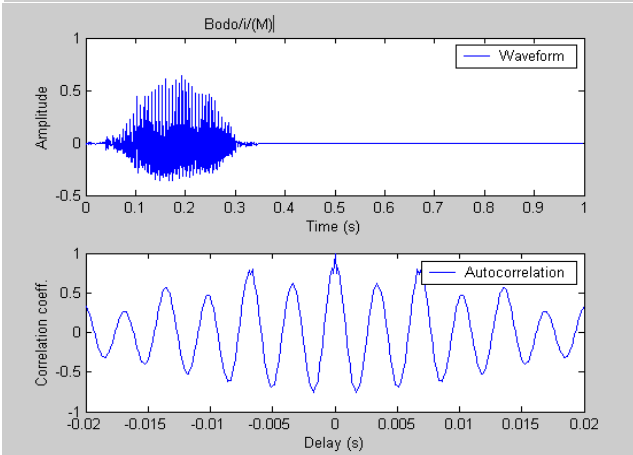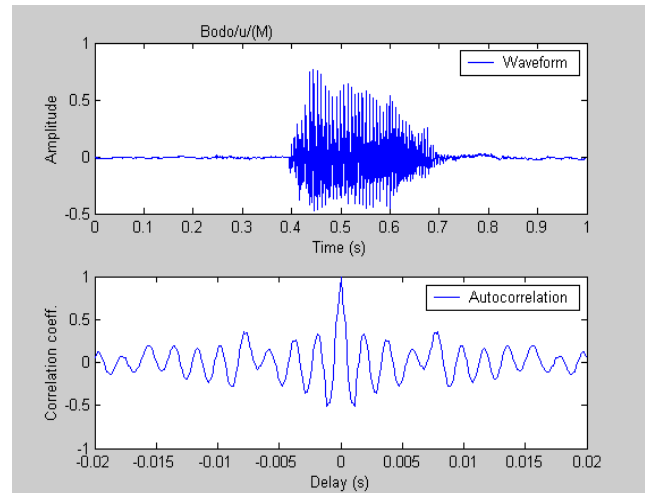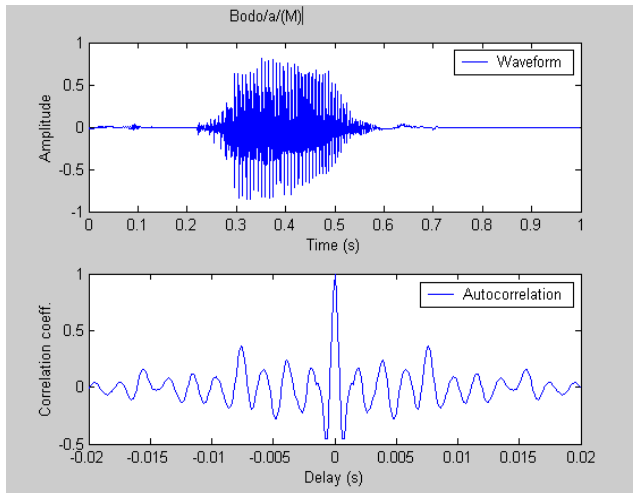
Fig. (1): Estimation of pitch of Bodo vowel corresponding to Male informants (time domain)
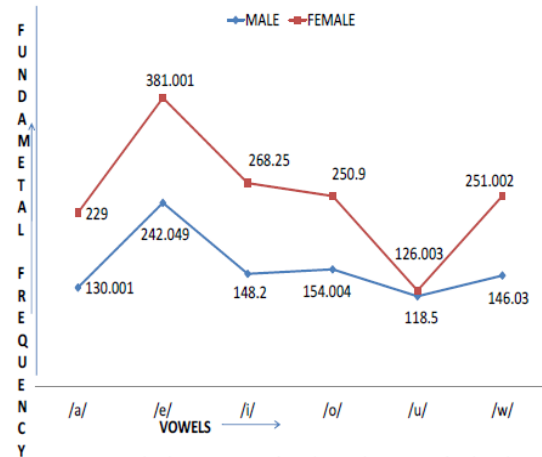
Fig. (3): Graphical representation of Fundamental Frequency of Bodo Male and Female for vowel utterances

II. RESULTS AND DISCUSSION

Typically, the pitch or fundamental frequency ranges from 80Hz to 160Hz for male speakers and from 140Hz to 400Hz for female speakers **[4].** The formant frequencies are usually greater than the pitch frequency. So, in speech encoding, synthesis and recognition, the estimation of pitch and formant frequencies finds extensive use. In adult, generally the length of vocal folds in male is more than that of female counterpart. The more is the vocal fold length, less is the pitch frequency. **Thus the pitch differs in male and female informants.**

In the present study, from the *Table (1) and Figure (3)*, it is observed that the values of pitch or fundamental frequency for female informant are higher than that of male informants as proposed by Pinto-et-al **[5].** Thus, it can be concluded that the pitch or fundamental frequency can be affectively used for the verification of sex of the Bodo speakers through the vowel utterances.

REFERENCES
[1] Wanda, W.Zaho and T.Ogunfinm: "Formant and Pitch detection using Time-Frequency distribution", International Journal of Speech Technology, 1999, pp.35-49
[2] Choudhury, S.: "A Statical Analysis of Speech Dependent/Independent Pattern Congruity of Assamese and Bodo Phonemes" a Ph.D. Thesis
[3] Thomas, F. Quqtieri: "Discrete- Time Speech Signal Processing: Principle and Practice", Pearson Education
[4] Wanda, W.Zaho and T.Ogunfinm: "Formant and Pitch detection using Time-Frequency distribution", International Journal of Speech Technology, 1999, pp.35-49
[5] Pinto, N.B. and Childers, D.G.: "Formant Speech Analysis".

Fig. (2): Estimation of pitch of Bodo vowel corresponding to Female informants (time domain)