# Pitch and formant estimation of bangla speech signal using autocorrelation, cepstrum and LPC algorithm

**Conference Paper** · December 2016

DOI: 10.1109/ICCITECHN.2016.7860226



**3 authors:**


Muhammad Navid Anjum Aadit
University of California, Santa Barbara



Sharadindu Gopal Kirtania
Washington State University



Mehnaz Tabassum Mahin
University of California, Riverside

# Pitch and Formant Estimation of Bangla Speech Signal Using Autocorrelation, Cepstrum and LPC Algorithm

Muhammad Navid Anjum Aadit[1], Sharadindu Gopal Kirtania[2], and Mehnaz Tabassum Mahin[3]
[1,2]Department of Electrical and Electronic Engineering
[3]Department of Computer Science and Engineering
Bangladesh University of Engineering and Technology
Email: [1]navidanjumaadit@gmail.com, [2]sharadindugopal@gmail.com, and [3]mahin.cse.09@gmail.com

***Abstract*—In this paper, we present comparative study of digital speech processing on Bangla speech signal. We represent oral characteristics of Bangla alphabet in terms of pitch and formant. We worked with both vowels and consonants to show their difference in practical use. We take oral speech signals as voice record and extract phonemes to analyze in both time and frequency domains. Both male and female voices are included in the analysis to find the effect of gender on Bangla voice speech. We verify our work by showing similarity of the results in time and frequency domains. We calculate first three formants of Bangla phonemes to show their relative impacts on Bangla speech. This extensive study of Bangla speech would provide listed information and data for further research to extract some other important features in this field.**

***Keywords*—Bangla Speech, Cepstrum, Formant, LPC, Pitch**

## I. Introduction

Speech signal processing is the study of speech signals to extract different important features [1]. Speech spectrum analysis, pitch and formant estimation, noise characteristics and its elimination, speech recognition, speech synthesis are some useful tools for speech processing. It is important due to digitization, transformation, transportation, storage, manipulation of speech signals in our modern life. Pitch estimation of voice signal is one of the most preeminent tasks in speech processing [2]. It is used in phonetics [3], music information retrieval [4], speech coding [5]. Pitch, a subjective psycho-physical property of a sound refers to the auditory characteristics according to which sounds can be ordered on a scale from low to high [6]. Pitch estimation can be done in time and frequency domains [7]. Pitch estimation or detection is useful in sound transformations, musical instrument recognition melodies [8], converting a signal captured by a microphone into a midi number [9], determining music notation number. It also contains information about human emotions [10]. So it is an important analysis in signal processing. Formant is a spectral peak of a sound spectrum [11]. A formant, acoustic resonance of human vocal tract refers to a range of frequencies in which there is a relative or absolute peak in the sound spectrum [12]. These frequencies are used for characterization of speech signal. It is a useful analytical model for speech sound production and the independence of the source of sound in the vocal tract from the filter that shapes that sound [13]. Many researches have been done on pitch and formant estimation of speech signal. Chenghui *et al.* proposed an algorithm of formant estimation of whispered speech [14]. Wang and Quateri proposed high pitch frequency estimation by exploiting temporal change of pitch [15]. Khulage extracted pitch, duration and formant frequencies of emotion recognition system [16]. Chen and Loizu addressed the problem of formant frequency estimation of speech signals corrupted by colored noise [17].

Bangla (or Bengali) is a very popular spoken language but digital speech processing is quite a new concept for this language. However, there are still some promising research works found in literature that introduce Bangla speech processing. Kamal *et al.* presented a fuzzy based Bangla vowel sign recognition system [18]. Paul *et al.* developed Bangla speech recognition using Linear Predictive Coding (LPC) and Artificial Neural Network (ANN) [19]. Moreover, Kotwal *et al.* proposed a technique of gender effect suppression that is composed of two Hidden Markov Model (HMM) based classifiers [20]. Hossain *et al.* further synthesized vowel space illustration by spectral analysis of Bangla vowel sounds to locate them in both acoustic and articulatory spaces of the vocal tract tube [21]. Rahaman *et al.* also extracted some important features which include pitch and first two formants for few vowels [22]. But to the best of our knowledge, estimation of pitch and first three formants of Bangla alphabet including both vowels and consonants is not found in literature. Unlike other languages, these features are not recorded so far in tabular form for Bangla phonemes that limits further speech processing study on Bangla language.

In this paper, we report pitch and formant estimation of Bangla vowels and consonants, for the first time. We present our results in both graphical and tabular forms. The results are comparable in time and frequency domains that validates the accuracy of our findings. We present first three formants of vowels as list and of consonants as groups. Effect of gender on Bangla voice speech is also studied by taking both male and female voice samples. We use autocorrelation and cepstrum algorithms for pitch detection and Linear Predictive Coding

(LPC) algorithm for formant detection. Estimation of pitch and formant for Bangla speech would open up many new scopes of research on this language based on digital speech processing when listed data of these important features are available in tabular form.

## II. Overview of Database Used

Bangla alphabet can be classified as vowels and consonants, like most other languages. There are mainly 11 vowels and 39 consonants in Bangla. However, if we consider phonemic inventory, there are 7 vowel phonemes and 29 consonant phonemes [23]. In this work, we work on all 36 phonemes and find their pitch and first three formants. However, Bangla consonants can be grouped based on their pronunciation style. In this paper, pitch and first three formants of consonants are recorded based on these groups so that general effects of each group can be realized and redundancy in data is removed. Bangla vowels and consonants are recorded using voice recorder as oral speech by native Bangla speakers in a noise-free silent environment. The database consists of voice records by 25 males and 25 females, whose age range is from 20 to 25 years.

## III. Simulation Technique

Speech signal varies with time and so for extracting features properly, we need to identify a smaller portion or window segment of speech. Then the recorded speech signal is downsampled and read using 'audioread' function in MATLAB. This function reads a small window of the downsampled signal and returns the sampled values as well as the sampling frequency. For pitch detection, both time and frequency domain analyses are done in this work. Formants are estimated using well known Linear Predictive Coding (LPC) method [24].

### A. Pitch Detection

*1) Time Domain:* Pitch refers to the fundamental frequency of a voice signal. It evaluates how much stress is applied while generating the voice. To detect pitch in time domain, autocorrelation is the most popular and effective technique developed so far [25]. Autocorrelation is the cross-correlation of a speech signal with itself in time domain. At first, the downsampled Bangla speech is split into a 40 ms window segment and the corresponding samples are read using 'audioread' function in MATLAB. Then we invoke the built-in function 'xcorr' which returns cross-correlation sequence of the window segment of the sampled signal. The pitch period is then found by calculating time lag between the central peak and the second highest peak of the autocorrelation sequence. Finally, pitch frequency is calculated simply by taking inverse of the pitch period.

*2) Frequency Domain:* To detect pitch in frequency domain, cepstrum method is a widely used algorithm. Cepstrum of a signal is obtained by taking the Inverse Fourier transform (IFT) of the logarithm of the spectrum of that signal [26]. Mathematically,

$$c[n] = F^{-1}\{\log(|F\{x[n]\}|)\} \ , \qquad (1)$$

where, $x[n]$ is the sampled speech signal,
$F$ indicates its Fourier transform,
and $c[n]$ are the cepstrum coefficients.

Like autocorrelation method, cepstrum method also reads a 40 ms window segment of the downsampled Bangla voice signal using 'audioread' function in MATLAB. Then the signal is multiplied by a hamming window. Fast Fourier Transformation (FFT) of this windowed frame gives the spectrum of the speech signal in frequency domain. Taking Inverse Fourier Transformation (IFT) of logarithm of the spectrum gives cepstrum in quefrency domain. Once in the quefrency domain, the pitch can be estimated by determining the peak of the cepstrum which represents pitch lag. The lag at which there is the most energy represents the dominant frequency in the log spectrum and thereby it gives the pitch frequency. A flow diagram of cepstrum algorithm is shown in Fig. 1.

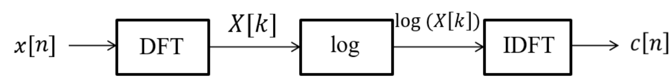


Fig. 1. Flow diagram of cepstrum algorithm for pitch detection

### B. Formant Detection

Linear Predictive Coding (LPC) is one of the most potential speech analysis tools used in audio and speech signal processing which provides precise prediction of formants of a speech signal [27]. The formant frequencies are obtained by finding the roots of the prediction polynomial. As in pitch detection, a small segment of 40 ms is extracted from the recorded Bangla voice signal and multiplied by a hamming window. Then a pre-emphasis filter is applied that is a high-pass all-pole filter. LPC coefficients are obtained by using the MATLAB built-in function 'lpc' and then the roots of the prediction polynomial returned by 'lpc' are calculated. Only imaginary parts of the roots with one sign are retained and the corresponding angles are computed. The distances of the zeros of the prediction polynomial from the unit circle represents the bandwidths of the formants. To determine first three formants, the criteria that the upper limit of the bandwidths is 400 Hz and the lower limit of the formant frequencies is 90 Hz are used.

Formant frequencies can also be determined from the LPC frequency domain spectrum. Local maxima of this spectrum which lie in a small limited bandwidth approximately represent the formants [28]. A flow diagram of LPC algorithm is shown in Fig. 2.

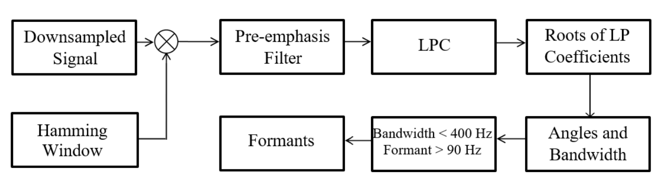


Fig. 2. Flow diagram of LPC algorithm for formant detection

## IV. Results and Discussions

At first, we collect both male and female vowel voice samples from native Bangla speakers, downsample them and read small segments of 40 ms in MATLAB. Then we calculate pitch for each voice segment using autocorrelation method as described in Section III-A1. Fig. 3 shows the results of autocorrelation for the vowel 'আ' (pronounced as /a/) uttered by a male speaker. From the time axis, we see that the voice signal is clipped into a 40 ms windows segment where the amplitude of the voice is significant. We can also see that the zeroth lag of the output correlation is in the middle of the sequence. We consider it as the delay for central peak. Then we find that the second highest peak is at about 0.0095 s delay. Thus, the pitch period for vowel 'আ' is about 0.0095 s and the corresponding pitch frequency is 105.3 Hz.

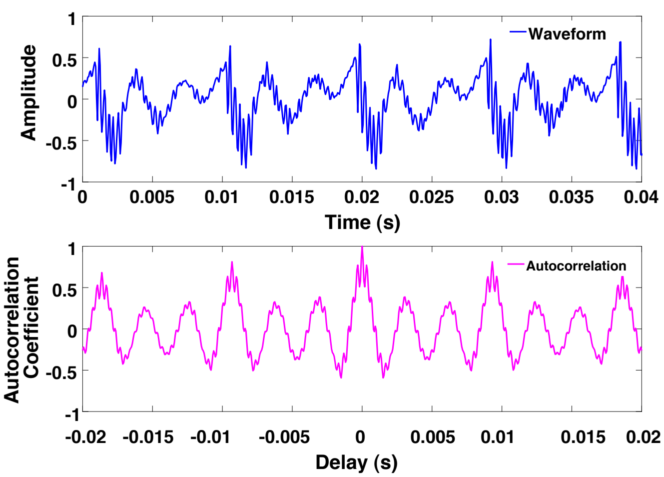


Fig. 3. Results of autocorrelation for pitch estimation of vowel 'আ' uttered by a male

Now, we investigate pitch of the same vowel 'আ' uttered by the same human using cepstrum algorithm which is implemented in frequency domain as described in Section III-A2. Fig. 4 shows the results of cepstrum method for vowel 'আ'. The upper part shows the clipped speech of 40 ms where the middle part shows the log spectrum of its FFT. Taking IFT of the log spectrum gives the cepstrum which is shown in the bottom part. The x axis of the cepstrum is quefrency which is measured in second. As the spike with highest peak in cepstrum determines the fundamental frequency, we note from figure that the highest peak occurs at quefrency 0.0094 s. Therefore, the pitch period is 0.0094 s and the corresponding pitch frequency is 106.7 Hz. This value is very close to our findings using autocorrelation method, which proves the accuracy of both methods applied for Bangla speech.

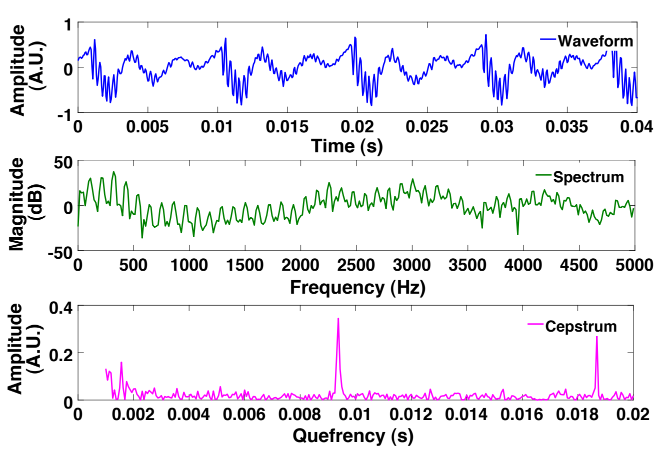


Fig. 4. Results of cepstrum for pitch estimation of vowel 'আ' uttered by a male

Fig. 5 and Fig. 6 show pitch calculation of vowel 'আ' uttered by a female using autocorrelation and cepstrum method respectively. Using the similar approaches as those for male voice, pitch frequency of vowel 'আ' for a female is found to be about 250.0 Hz (pitch period 0.004 s) in both the methods. Clearly, pitch frequency for female is much higher, even higher than double of that of male. Females have smaller vocal folds and so voices are uttered with higher pitch frequencies. Thus, gender has an important effect on pitch frequencies of Bangla speech. This is quite similar to the results found for pitch frequencies of English speech [29].

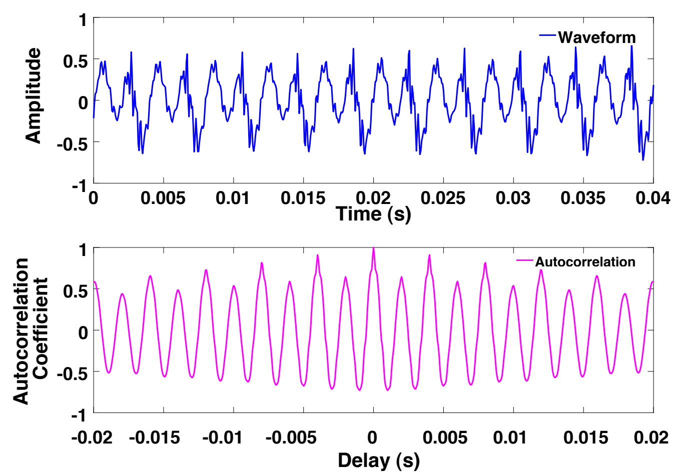


Fig. 5. Results of autocorrelation for pitch estimation of vowel 'আ' uttered by a female

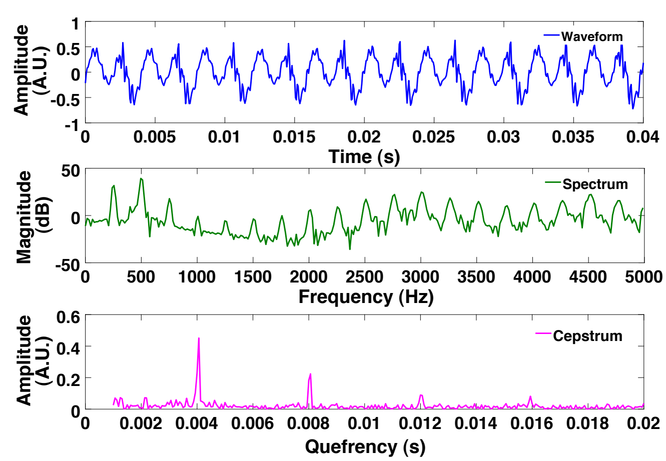


Fig. 6. Results of cepstrum for pitch estimation of vowel 'আ' uttered by a female

Now, we present formant estimation using LPC method as described in Section III-B. Fig. 7 shows the first three formants estimation for vowel 'উ' (pronounced as /u/) uttered by a native male. The upper part is the clipped and downsampled voice signal and the lower part represents the frequency response of LPC spectrum. From the local maxima of this spectrum, we can estimate that the first three formants are at about 400, 800 and 1500 Hz. Moreover, solving for roots of the LPC polynomial we find that the first three formants are accurately at 410.7, 843.4 and 1593.2 Hz. Both the frequency spectrum observation and root solving techniques provide close values that proofs the accuracy of the simulations

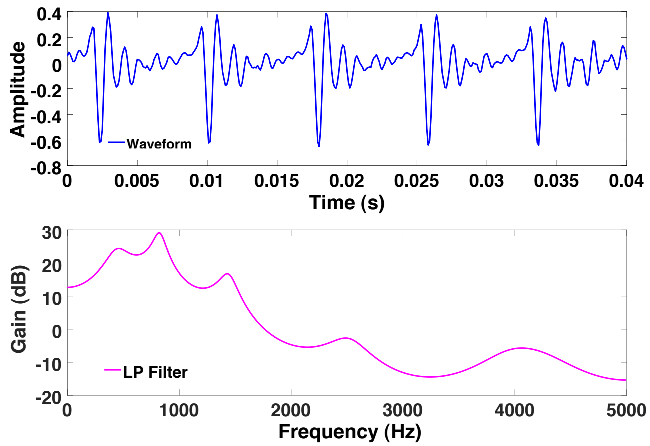


Fig. 7. Results of LPC for formant estimation of vowel 'উ' uttered by a male

Fig. 8 shows the first three formants estimation for vowel 'উ' uttered by a native female. In this case, the first three formants are found at 547.5, 1698.5 and 3190.8 Hz. Clearly, all the three formants are at higher values than the male counterparts. However, we see that the difference between formant frequency in male and female voice increases from first formant to third formant, as we increase the order of the formant. Thus, effect of gender plays an eminent role in both pitch and formant frequencies of Bangla speech.

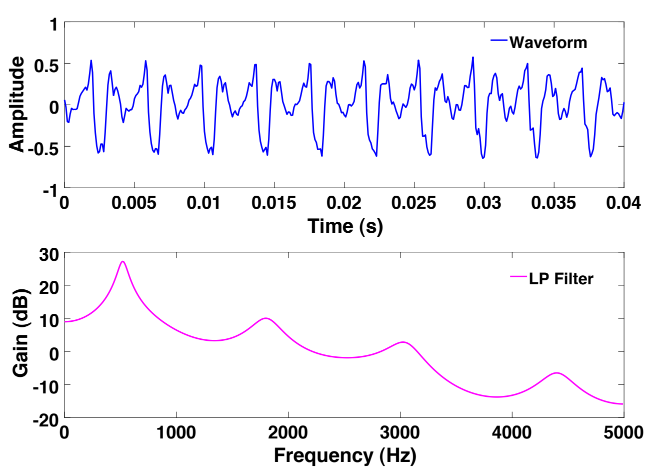


Fig. 8. Results of LPC for formant estimation of vowel 'উ' uttered by a female

Finally, we present pitch and formant estimation for Bangla consonants in Fig. 9. The results are shown for Bangla consonant 'প' (pronounced as /pɔ/) uttered by a native female speaker. From the figure, we find that the pitch period is about 0.0033 s in both auto-correlation and cepstrum methods. The corresponding pitch frequency is 301.8 Hz. The first three formants are estimated from the local maxima of the LPC spectrum shown at the bottom part of the figure and also calculated solving roots of LPC polynomial as 523.3, 1800.1 and 2706.6 Hz.

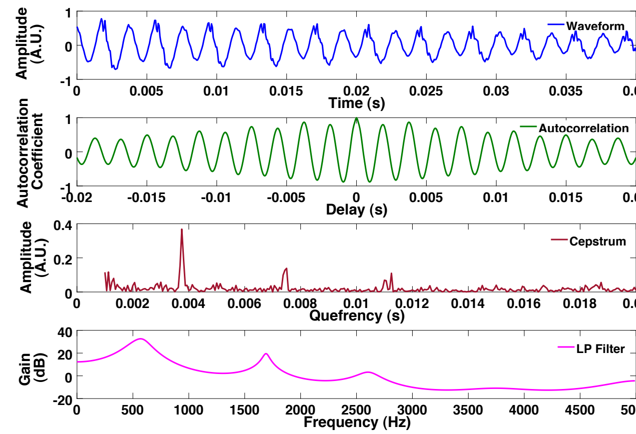


Fig. 9. Results of pitch and formant estimation of consonant 'প' uttered by a female

We summarize our work in this paper in Table I, Table II and Table III so that data can be easily collected from the tables for further research use. In the tables, male indicates average from 25 males and female indicates average from 25 females. The data are further averaged taking different time frames. Table I records pitch estimation of 7 important Bangla vowels. The table contains data collected from both autocorrelation and cepstrum methods. Table II records first three formants estimation of those 7 Bangla vowels using LPC method. Finally, Table III summarizes estimation of pitch and first three formants of Bangla consonants. As mentioned earlier, we prepare the table by considering consonants as groups so that data redundancy can be avoided. Moreover, as autocorrelation and cepstrum methods provide almost similar data without significant difference, we take average for pitch calculation of consonants collecting data from these two methods.

We further investigate effects of gender on pitch and formant frequencies of Bangla consonants. From Table III, it is found that both pitch and formant frequencies are higher for female speakers. This effect is quite similar to that of vowels. However, it is to be noted that the highest pitch and formant frequencies do not appear in the same consonant groups. Moreover, the consonant group that has the highest frequencies uttered by the male speakers may not contain the highest frequencies uttered by the female ones. Thus, the data recorded in tabular forms contain information of gender effects that may be helpful for future studies on Bangla alphabet and speech.

TABLE I. ESTIMATION OF PITCH FREQUENCIES OF BANGLA VOWELS

| Bangla Vowel | Pronunciation | Pitch Frequency Using Autocorrelation (Hz) | | Pitch Frequency Using FFT- Cepstrum (Hz) | |
|---|---|---|---|---|---|
| | | *Male* | *Female* | *Male* | *Female* |
| অ | /ɔ/ | 100.0 | 248.5 | 100.0 | 247.9 |
| আ | /a/ | 105.3 | 250.0 | 106.7 | 250.0 |
| ই | /i/ | 145.7 | 275.3 | 146.1 | 274.4 |
| উ | /u/ | 119.6 | 282.5 | 120.0 | 285.1 |
| এ | /e/ | 125.0 | 255.5 | 125.0 | 253.9 |
| ও | /o/ | 109.4 | 258.7 | 111.1 | 257.8 |
| অ্যা | /æ/ | 108.0 | 262.1 | 107.5 | 260.8 |

TABLE II. ESTIMATION OF FORMANT FREQUENCIES OF BANGLA VOWELS

| Bangla Vowel | Pronunciation | Formant Frequency (Hz) | | | | | |
|---|---|---|---|---|---|---|---|
| | | *Male* | | | *Female* | | |
| | | *F1* | *F2* | *F3* | *F1* | *F2* | *F3* |
| অ | /ɔ/ | 550.4 | 798.1 | 980.1 | 985.5 | 1492.8 | 3003.1 |
| আ | /a/ | 680.7 | 1250.2 | 1560.4 | 999.1 | 1748.2 | 3406.7 |
| ই | /i/ | 440.3 | 997.8 | 1372.2 | 946.3 | 1796.6 | 3250.6 |
| উ | /u/ | 410.7 | 843.4 | 1593.2 | 547.5 | 1698.5 | 3190.8 |
| এ | /e/ | 490.9 | 1176.7 | 1790.6 | 529.3 | 2100.0 | 3250.5 |
| ও | /o/ | 640.6 | 779.3 | 1650.1 | 780.7 | 800.0 | 3400.1 |
| অ্যা | /æ/ | 695.5 | 923.7 | 1602.3 | 835.5 | 2436.4 | 2730.4 |

TABLE III. ESTIMATION OF PITCH AND FORMANT FREQUENCIES OF BANGLA CONSONANT GROUPS

| Bangla Consonant Group | Pronunciation | Pitch Frequency Taking Average from Autocorrelation and FFT-Cepstrum (Hz) | | Formant Frequency (Hz) | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *Male* | *Female* | *Male* | | | *Female* | | |
| | | | | *F1* | *F2* | *F3* | *F1* | *F2* | *F3* |
| ক | /kɔ/ | 108.5 | 273.1 | 647.2 | 1025.6 | 1750.1 | 960.5 | 1200.0 | 3370.5 |
| চ | /tɕɔ~tsɔ~sɔ/ | 100.0 | 290.9 | 480.8 | 936.4 | 1500.0 | 740.2 | 2100.4 | 3350.5 |
| ট | /t̠ɔ/ | 100.0 | 270.3 | 695.5 | 1426.4 | 1842.3 | 1680.5 | 3239.9 | 3600.0 |
| ত | /t̪ɔ/ | 105.5 | 270.4 | 486.3 | 720.7 | 925.5 | 998.8 | 3030.3 | 3725.5 |
| প | /pɔ/ | 99.5 | 301.8 | 825.0 | 1480.9 | 1750.5 | 523.3 | 1800.1 | 2706.6 |

## V. CONCLUSION

We have presented estimation of pitch and formants for Bangla alphabet including both vowels and consonants. We have calculated pitch frequencies in both time and frequency domains and shown that the results are comparable. To estimate formants, we have used Linear Predictive Coding (LPC) algorithm in frequency domain. We have calculated first three formants using spectral observation as well as polynomial solution that proves the accuracy of our simulation. Both male and female voice samples have been recorded and effects of gender on Bangla speech signal have been discussed. Finally, we have presented our results in tabular forms that would be a useful resource to aid research works on Bangla speech in future.

## REFERENCES


[1] M. Saundade and P. Kurle, "Speech Recognition using Digital Signal Processing," *International Journal of Electronics, Communication and Soft Computing Science & Engineering (IJECSCSE)*, vol. 2, no. 6, p. 31, 2013.

[2] R. Pindak, C. Huang, and J. Ho, "The intrinsic pitch of cholesteryl nonanoate," *Solid State Communications*, vol. 14, no. 9, pp. 821–825, 1974.

[3] C. Gussenhoven, "Intonation and interpretation: phonetics and phonology," in *Speech Prosody 2002, International Conference*, 2002.

[4] G. Tzanetakis, A. Ermolinskyi, and P. Cook, "Pitch histograms in audio and symbolic music information retrieval," *Journal of New Music Research*, vol. 32, no. 2, pp. 143–152, 2003.

[5] R. P. Ramachandran and P. Kabal, "Pitch prediction filters in speech coding," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 37, no. 4, pp. 467–478, 1989.

[6] H. Fastl, "Psycho-acoustics and sound quality," in *Communication acoustics*. Springer, 2005, pp. 139–162.

[7] H. Kawahara, I. Masuda-Katsuse, and A. De Cheveigne, "Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds," *Speech communication*, vol. 27, no. 3, pp. 187–207, 1999.

[8] A. Eronen and A. Klapuri, "Musical instrument recognition using cepstral coefficients and temporal features," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*, vol. 2, 2000, pp. II753–II756 vol.2.

[9] R. J. McNab, L. A. Smith, and I. H. Witten, "Signal processing for melody transcription," *Australian computer science communications*, vol. 18, pp. 301–307, 1996.

[10] J. H. Hansen and S. Patil, "Speech under stress: Analysis, modeling and recognition," in *Speaker Classification I*. Springer, 2007, pp. 108–137.

[11] D. H. Klatt, "Prediction of perceived phonetic distance from critical band spectra: A first step," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'82.*, vol. 7. IEEE, 1982, pp. 1278–1281.

[12] W. Fitch and M. D. Hauser, "Vocal production in nonhuman primates: acoustics, physiology, and functional constraints on honest advertisement," *American Journal of Primatology*, vol. 37, no. 3, pp. 191–219, 1995.

[13] I. R. Titze, "Nonlinear source–filter coupling in phonation: Theory," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 2733–2749, 2008.

[14] G. Chenghui, Z. Heming, L. Gang, and L. Jianxin, "Formant estimation of whispered speech based on spectral segmentation," in *2006 IEEE International Symposium on Signal Processing and Information Technology*, Aug 2006, pp. 562–566.

[15] T. T. Wang and T. F. Quatieri, "High-Pitch Formant Estimation by Exploiting Temporal Change of Pitch," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 1, pp. 171–186, Jan 2010.

[16] A. A. Khulage, "Extraction of pitch, duration and formant frequencies for emotion recognition system," in *Communication and Computing (ARTCom2012), Fourth International Conference on Advances in Recent Technologies in*, Oct 2012, pp. 7–9.

[17] B. Chen and P. C. Loizou, "Formant frequency estimation in noise," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 1, May 2004, pp. I–581–4 vol.1.

[18] M. S. Kamal, M. M. Hoque, M. M. U. Hasan, and M. S. Arefin, "Bangla vowel sign recognition by extracting the fuzzy features," in *Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference on*, Dec 2008, pp. 306–311.

[19] A. K. Paul, D. Das, and M. M. Kamal, "Bangla Speech Recognition System Using LPC and ANN," in *Advances in Pattern Recognition, 2009. ICAPR '09. Seventh International Conference on*, Feb 2009, pp. 171–174.

[20] M. R. A. Kotwal, F. Hassan, M. S. Alam, S. I. Daud, F. Ahmed, and M. N. Huda, "Gender Effects Suppression in Bangla ASR by Designing Multiple HMM-Based Classifiers," in *Computational Intelligence and Communication Networks (CICN), 2011 International Conference on*, Oct 2011, pp. 390–394.

[21] S. A. Hossain, M. L. Rahman, and F. Ahmed, "Spectral Analysis of Bangla Vowels," in *2005 Pakistan Section Multitopic Conference*, Dec 2005, pp. 1–5.

[22] M. A. N. R. Rahaman, A. Das, M. Z. Nayen, and M. S. Rahman, "Special feature extraction techniques for Bangla speech," in *Computer and Information Technology (ICCIT), 2010 13th International Conference on*, Dec 2010, pp. 114–119.

[23] Wikipedia. (2016) Bengali Phonology. [Online]. Available: https: //en.wikipedia.org/wiki/Bengali phonology

[24] V. Jain, "Formant estimation by LPC with a new error criterion," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '82.*, vol. 7, May 1982, pp. 1968–1971.

[25] Wikipedia. (2016) Autocorrelation. [Online]. Available: https://en.wikipedia.org/wiki/Autocorrelation

[26] Wikipedia. (2016) Cepstrum. [Online]. Available: https://en.wikipedia. org/wiki/Cepstrum

[27] Wikipedia. (2016) Linear Predictive Coding. [Online]. Available: https://en.wikipedia.org/wiki/Linear predictive coding

[28] D. Gargouri, M. A. Kammoun, and A. B. Hamida, "A comparative study of formant frequencies estimation techniques," in *Proceedings of the 5th WSEAS international conference on Signal processing, Istanbul, Turkey*. Citeseer, 2006.

[29] M. Iseli, Y.-L. Shue, and A. Alwan, "Age-and gender-dependent analysis of voice source characteristics," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 1, May 2006, pp. I–I.


978-1-5090-4090-2/16/$31.00 ©2016 IEEE  ISBN 978-1-5090-4089-6