International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

Pitch Estimation using Time Domain Algorithm

Madhukar S. Chavan

Abstract: ***Speech is classified into voice, unvoiced and silence. The voice speech is the periodic vibration of vocal folds. Background noise affects the speech signals. In many speech applications calculation of pitch plays a major role. The paper proposes a pitch detection algorithm based on the short-time average magnitude difference function (AMDF) and the short-term autocorrelation function (ACF). Detecting the Pitch within the speech signal is important in most of all the speech related applications. Detection of Pitch is useful in identification of speaker. One solution to get detect with the pitch is by using the time domain algorithms. This paper gives idea about estimation and detection of pitch in time domain algorithm for different voice samples.***

Keywords: Pitch, Energy of clipped signal, Autocorrelation, Average magnitude difference function.

I. INTRODUCTION

Human's primary mode of communication is speech. Through the vocal cords the air is passed which also passes through the vocal tract and the production of speech takes place. While the creation of voice takes place due to the periodical vibrations of vocal cords which gives the glottal wave. Pitch period (T0) is referred as the time duration of one glottal cycle and the reciprocal is referred to as the fundamental frequency (F0). It has a wide applications areas like the digital cellular communication, conferencing of video and coding of speech signals etc. Estimation of speech is also important in many applications like synthesis of speech signals, speech recognition, identification of a speakers, coding of speech and its verification.

In many speech processing algorithm detection of pitch is very important. Recognition of speech system of tongue uses pitch chase for tone recognition. Pitch is additionally crucial for speech variation in language systems and text-to-speech systems. The main cue of pitch is its fundamental frequency (F0). However it is tough to make a mathematical model which helps to calculate the pitch of signal. Therefore it is much more important to design a pitch detection algorithm (PDA) in many of the speech processing applications systems. A convolution of a voice signal of time-varying stimulus is delivered by a vocal tract and a glottal flow.

The vocal cords is a popular model for human vocal system which acts as "repeated pulse source" which also acts as a filter and through which the sound is shaped is provided by the train of pulses within the vocal tract. The pitch is the period between the pulses emitted by the vocal cords. Fig 1.1 shows the signal processing of a voiced signal. The ear of a human respond at a frequency of 20 Hz-20 KHz. This frequency is obtained with the help of vibrations which are caused due to the continuous flow of pressure. The human sound produced could be categorized into unvoiced signal and voice signal. During the process of utterance of unvoiced sounds the vocal cords do not vibrate and stay open whereas during the process of utterance of voiced sounds, the vocal cords vibrate and produce glottal pulses. A pulse is a summation of a sinusoidal wave of fundamental frequency (F0) and its harmonics. Pitch is important in identification of the characteristics of voice within the human being, the information of speaker, like the classification of gender that is whether male, female and age whether adults or children or age old and healthy or deceased

Manuscript published on 30 September 2019.
***Correspondence Author(s)**
Dr. Madhukar S. Chavan, Dept. of Electronics and Telecomm. Engineering, P. V. P. Institute of Technology Budhgaon, Sangli, India. Email: mschavan67@gmail

II. LITERATURE REVIEW

This system is used for identifying the Pitch within the wave. The purpose of the paper is to implement algorithms which helps in modifications in the existing speech coding techniques [1].

K. Inbanila et.al [3] have described a method which is used to improve substitution quality of voice signals produced by artificial larynx transducer.

Geliang Zhang et.al [5] have described about the technique which uses the particle filter approach to track pitch period.

Yuan Zong et.al [6] have described a new modified technique for pitch detection. This technique is termed as Empirical Mode Decomposition.

III. PITCH DETECTION ALGORITHMS

Speech signal always varies with time that is it always varies with respect to time. To get the exact feature of speech we need to identify the smaller portion of speech signal. The signal is recorded and converted to wave file so as to process the signal in the Matlab. Time domain pitch analysis includes the Autocorrelation method and AMDF techniques.

IV. TIME DOMAIN ALGORITHM

There are many different ways for detection of pitch signals. Out of this the Autocorrelation is the most reliable pitch detector technique. The requirement for higher processing is needed for this approach. When a signal is transmitted over a telephone line the autocorrelation is a very good method for getting the expected results. It does displays prominent peaks.

A. Autocorrelation Method

The Cross-correlation of signal with itself is defined as autocorrelation. The repeated patterns are fined using the autocorrelation function. The pitch periods are estimated for a given signals. The short term analysis technique is used for pitch detection.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

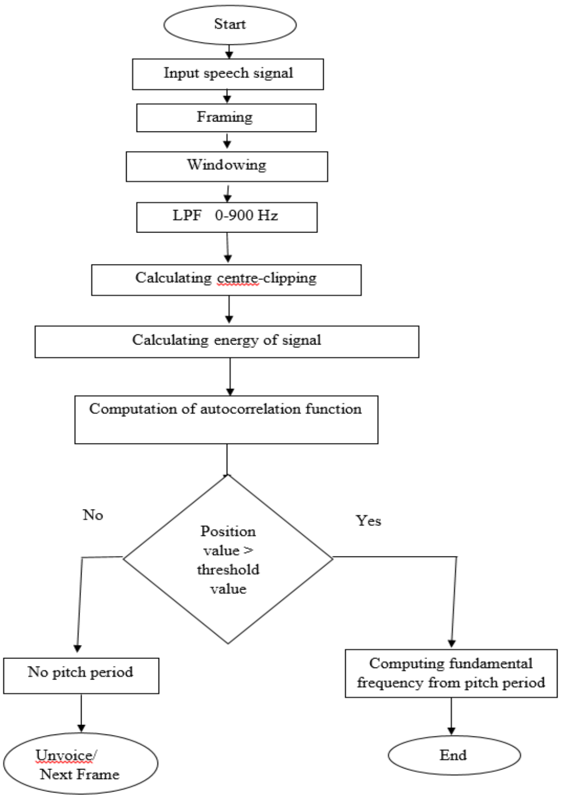

Fig. (1) Flowchart of Pitch Detection Algorithm using Autocorrelation Method

The detail of the algorithm are depicted in fig. (1) And a summary of algorithm is given below.

1. The signal is considered and it is filtered at a 900Hz and applied with a sampling frequency of 10 KHz.
2. 10ms intervals are selected from a length of segments of 30msec. Thus the segments overlaps by 20msec.
3. The Autocorrelation, Equation (1), samples using a hamming window. Overlapping is allowed within the frames. The peak signal level for 30msec determines background noise. If the signal goes above the threshold value then it is termed as a voice otherwise it is termed as silence and no further action is taken.
4. Fixed % is determined by the clipping level.
5. Using clipping level, processing of speech signal takes place.
6. The largest value from the signal is computed and compared to a fixed threshold. If the peak value is below threshold it is termed as unvoiced signal and when it is above the threshold then it is defined as the largest peak location.

For appropriate normalization process the autocorrelation function at zero delay is computed. It always search for maximum peaks. The values are always computed and compared with the threshold. Within the pitch fundamental frequency is computed by,

$$Rx(\mathrm{m}) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x(\mathrm{n})\,x(\mathrm{n+m}) \qquad (1)$$

Where m is the of sample points time lag. Autocorrelation algorithm is relatively impervious to noise.

B. Average Magnitude Difference Function

It is similar to Autocorrelation it estimates the gap rather than the similarities within the frames. It is termed as the different form of autocorrelation analysis. The signals are formed by a delay difference between the input signal and it is given by relation,

$$D_{\mathrm{m}} = \frac{1}{L}\sum_{n=1}^{L} |(X(n) - X(\mathrm{n-m})| \qquad (2)$$

Where x (n) are input speech samples and x (n-m) time shifted samples. The 'Dm' is the difference signal which is formed by delaying the input signal for certain amount of time. The pitch period is identified as the value of the lag at which the minimum AMDF occurs.

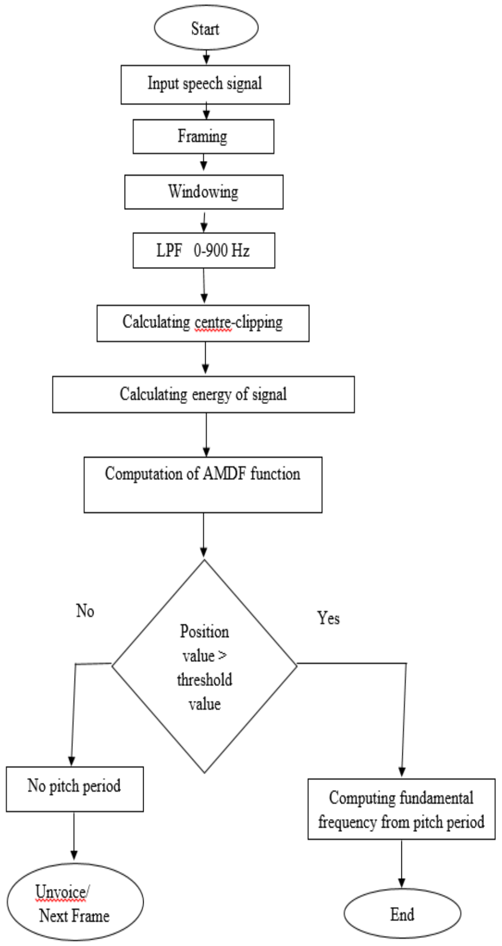

Fig. (2) Flowchart of Pitch Detection Algorithm using AMDF Method

The detail of the algorithm are depicted in fig. (2) And a summary of algorithm is given below.

1. The signal is considered and it is filtered at a 900Hz and applied with a sampling frequency of 10 KHz.
2. 10ms intervals are selected from a length of segments of 30msec. Thus the segments overlaps by 20msec.
3. The Autocorrelation, Equation (2), samples using a hamming window. Overlapping is allowed within the frames. The peak signal level for 30msec determines background noise. If the signal goes above the threshold value then it is termed as a voice otherwise it is termed as silence and no further action is taken.
4. Fixed % is determined by the clipping level.
5. Using clipping level, processing of speech signal takes place.

Retrieval Number: J97330881019/19©BEIESP
DOI: 10.35940/ijitee.J9733.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

6. The largest value from the signal is computed and compared to a fixed threshold. If the peak value is below threshold it is termed as unvoiced signal and when it is above the threshold then it is defined as the largest peak location.
7. The largest peak of the Average magnitude difference function is located and the peak value is compared to a fixed threshold (e.g., 30% of Rn(0)). If the peak value falls below the threshold, the segments is classified as unvoiced and if it is above, the pitch period is defined as the location of the largest peak.

- AMDF tapers off with the lag m. the AMDF is divided by the length of overlap just to avoid tapping-off:

V. RESULTS AND DISCUSSION

- The range for children lies between 200Hz to 500Hz i.e., the pitch which is in this range are classified as voice of a child.
- The range for female lies between 150Hz to 350Hz i.e., the pitch which is in this range are classified as voice of a female.
- The range for children lies between 80Hz to 200Hz i.e., the pitch which is in this range are classified as voice of a men.

A. For Children

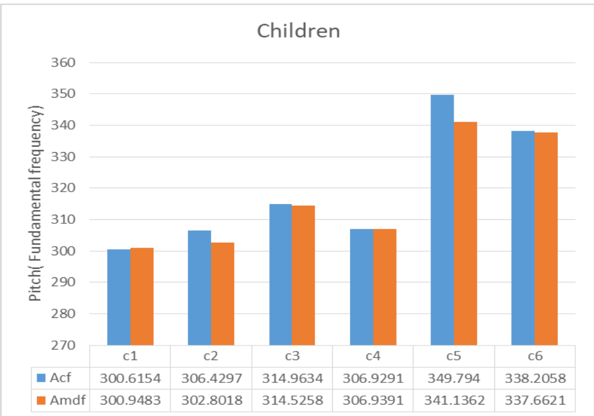

Fig. (a) Graph of ACF and AMDF for Children

Fig. (a) Shows results of autocorrelation and average magnitude difference function for 6 samples of children where the pitch or frequency range for children matches with the range 200Hz to 500Hz hence the values of pitch are classified under children categories.

B. For Female

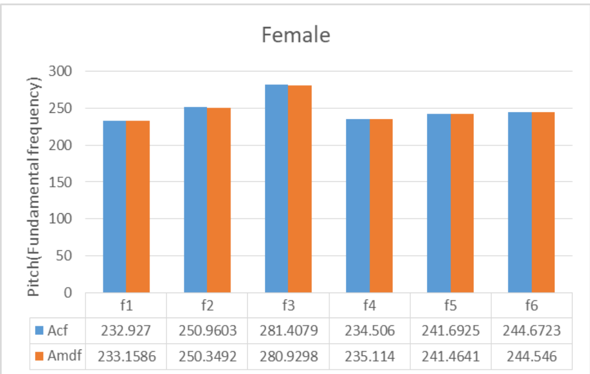

Fig. (b) Graph for ACF and AMDF for Female

Fig. (b) Shows results of autocorrelation and average magnitude difference function for 6 samples of female where the pitch or frequency range for female matches with the range 150Hz to 350Hz hence the values are classified under female categories.

C. For Male

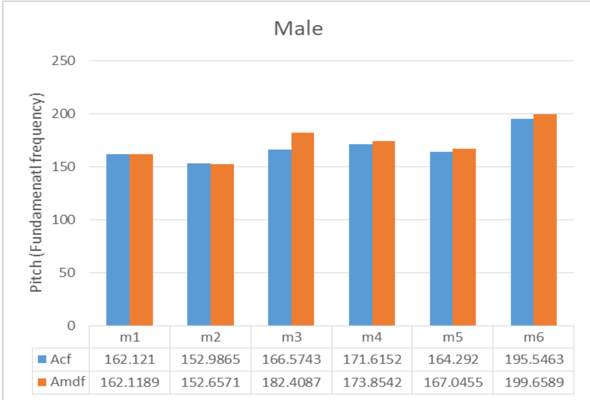

Fig. (c) Graph for ACF and AMDF for Male

Fig. (c) Shows results of autocorrelation and average magnitude difference function for 6 samples of male where the pitch or frequency range for male matches with the range 80Hz to 200Hz hence the values of pitch are classified under male categories.

- The autocorrelation and average magnitude difference function methods are used for calculation of pitch signal from the different collected samples.
- If the pitch values are in the range of 80Hz to 200Hz. Then the pitch signal is classified under male categories
- If the pitch values are in the range of 150Hz to 350Hz. Then the pitch signal is classified under female categories
- If the pitch values are in the range of 200Hz to 500Hz. Then the pitch signal is classified under children categories
- The autocorrelation and average magnitude difference function methods gives pitch value and by comparing the pitch value with the standard values we can easily classify the sample according to different categories i.e., children, female and male.

VI. CONCLUSION

The time domain method is used for calculation of pitch in which the values are compared to the standard format which is used to classify the gender of a person. In future we may extend this for noise reduction and multiple speaker source separation.

REFERENCES

1. L. R.Rabiner and R. W.Schafer ,Digital Processing of Speech Signals, Prentice-Hall, Englewood Cliffs, NJ, 1978.
2. Pratibha and Chandrashekar H. M, “Estimation and Tracking of Pitch for Noisy Speech Signals using EMD based Autocorrelation Function Algorithm),” 2nd IEEE International Conference on Recent Trends in Electronics Information & Communication Technology (RTEICT), May 19-20,2017.
3. K. Inbanila and E. Krishnakumar, ‘‘Enhancement of Substitution Voices Using F1 Formant Deviation Analysis and DTW Based Template Matching” International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) 2017.

4. Muhammad Navid Anjum Aadit, Sharadindu Gopal Kirtania and Mehnaz Tabassum Mahin, "Pitch and Formant Estimation of Bangla Speech Signal Using Autocorrelation, Cepstrum and LPC Algorithm" 19th International Conference on Computer and Information Technology, North South University, Dhaka, Bangladesh, December 18-20, 2016.
5. Geliang Zhang, Simon Godsill, ''Tracking Pitch period using particle filters" IEEE Workshop on Applications of Signal Processing to Audio
6. Yuan Zong, Yumin Zeng, Mengchao Li, Rui Zheng, ''Pitch Detection Using EMD- Based AMDF" Fourth International Conference on Intelligent Control and Information Processing (ICICIP), Beijing, China June 9, 2013.
7. Bageshree V. Sathe, Pathak and Ashish R. Panat, "Extraction of Pitch and Formants and its Analysis to identify 3 different emotional states of a person" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012.
8. zhijun Cui, ''Pitch Extraction Based on Weighted Autocorrelation Function in Speech Signal Processing", 2nd International Conference on Computer Science and Network Technology 2012.
9. Savitha S Upadhya, ''Pitch Detection in Time and Frequency Domain" International Conference on Communication, Information & Computing Technology (ICCICT), Oct 1920, Mumbai, India 2012.
10. Li Hui, Bei-qian Dai, Lu Wei, ''A pitch detection algorithm based on AMDF and ACF" 2006 IEEE.

AUTHORS PROFILE

Dr. Madhukar S. Chavan, M.Tech & M.E. degree in E&TC Engg. from Rajasthan Vidyapeeth and Shivaji University. M.B.A.(Marketing) from Yashwantrao Chavan Open University and Ph.D. degree in Electronics Engg. from Shivaji University. Currently working as Professor and PG Co-ordinator in E&TC Engg. Department of P.V.P. Institute of Technology Budhgaon, Sangli

Retrieval Number: J97330881019/19©BEIESP
DOI: 10.35940/ijitee.J9733.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)