Pitch Estimation Using Pulse Modelling Approach and Its Comparison with Existing Methods

[1]Sunil Kumar, [2]Ameya K. Naik, [3]Swapna M. Patil
[1]M.E Student, [2]Assistant Professor, [3]Assistant Professor
[1]Electronics and Telecommunication,
[1]Shree L.R. Tiwari College of Engineering, Thane, India

***Abstract* - Pitch can be considered to be an important tool for much speech processing application and its estimation can be considered vital to many systems. Many methods exist for estimation of pitch with its advantages and disadvantages. In this paper, we have given a comparison between pulse modelling approach for estimation of pitch along with the traditional method and found that estimation of pitch with pulse modelling approach is more accurate with respect to the noise in the signal.**

***Index Terms* – Pitch estimation, Real-time processing, Speech processing.**

I. Introduction

In Speech processing, pitch defines the degree of the highness and lowness of a tone. Speaker identification and verification is done by pitch estimation and it is used as a very essential tool, it has various applications, for example speech synthesis, speech coding and speech instructions. Pitch Estimation is nothing but accurate estimation of the fundamental frequency of the signal. It is a very challenging task to estimate the pitch of a speech signal because of the mixed nature of excitation and quasi-periodic nature of speech. There are different methods of pitch estimation developed by researchers and all those methods are categorized in time domain and frequency domain [1].

The methods that operate directly on the given speech waveform are in time domain behavior. Auto-correlation, zero-crossing, valley and peak measurements are the algorithms used in pitch estimation. As speech signal is a quasi-periodic waveform so time domain measurement provides better estimation of the period [2].

There are many methods in time domain to calculate pitch estimation such as Parallel processing method, average magnitude difference function (AMDF), auto-correlation and threshold-crossing analysis. All these methods use complex mathematical calculation. This paper presents a better way of pitch estimation of the given speech signal in time-domain as compared to existing methods in a straightforward way.

II. Background

A. Pitch Estimation using Auto-correlation

Autocorrelation will give lots of information of the speech signal. Information is the pitch period (the fundamental frequency). To make the speech signal Closely approximate a periodic impulse train, we must use some kind of spectrum flattening. “Center clipping spectrum flattener” is used to do this [2]. After the this the fundamental frequency is extracted and the autocorrelation is calculated. An overview of the system is shown in Fig. 1.

B. Pitch Estimation using Cepstrum

The Fourier transform of a pitched signal usually have a number of regularly peaks, and it is representing the harmonic spectrum. These peaks are reduced when log magnitude of a spectrum is taken [1]. This gives a result which is a periodic waveform in the frequency domain, where the period of pitched signal is related to the fundamental frequency of the original signal. It means that a Fourier transformation of this waveform has a peak representative of the fundamental frequency [2].

III. Pulse Modelling Approach

Block diagram of the pulse Modelling system system is shown in Fig. 2. To estimate pitch, the input must be a speech signal. The main blocks are explained as follows.

A. Pulse Converter

Block diagram of the pulse converter is shown in Fig. 3. The switch requires a threshold, which is a speech signal passed through the low pass filter [6].

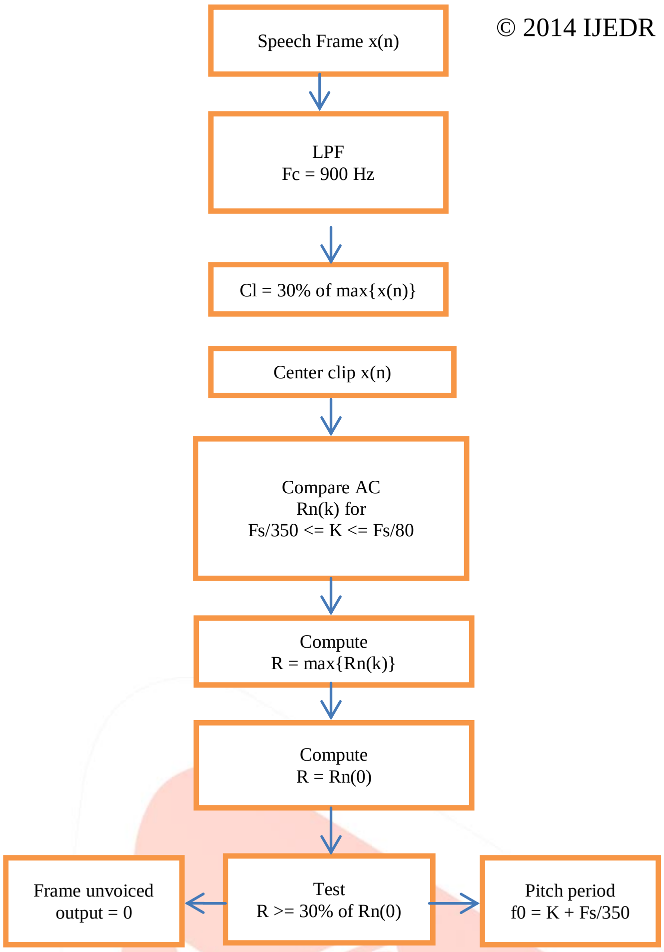

Fig.1 Overview of the autocorrelation pitch detector

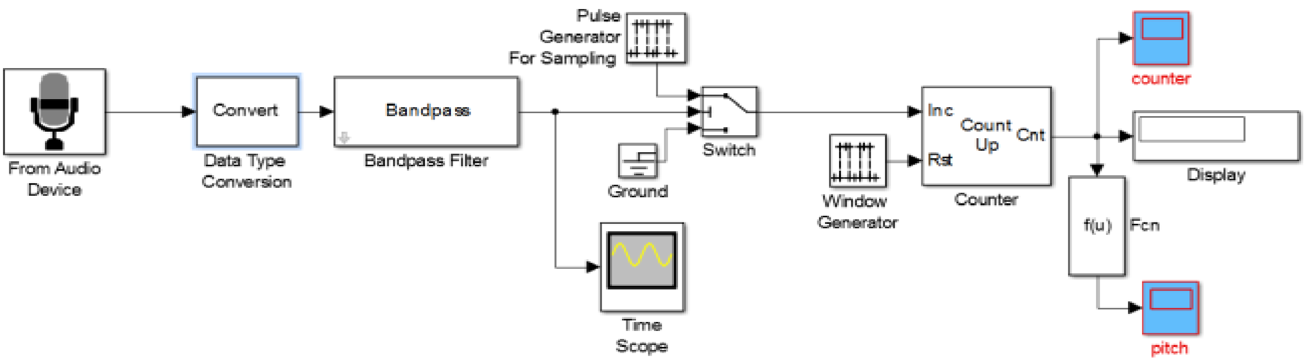

Fig. 2 Overview of Pulse modeling Approach

The threshold is detected by the switch and it changes its value between ground potential and the pulse generator. X (n) is the number of pulses and these pulses are passed through the switch. The frequency of the speech signal is directly proportional to the number of pulses. Switching will be less if the speech signal frequency is low and hence the number of pulses are allowed to be passed[1].

Switching time is directly proportional to frequency, hence if switching time increases the frequency of the signal will also increase and vice versa. Switching time is not directly proportional to the number of pulses. If switching time is 0.001 seconds a number of pulses will be generated that helps to eliminate the possibility of unwanted high frequency signals [5]. The system will sample only formant frequencies. Higher frequencies are not converted to the required pulse train [1].

Threshold in the Switch controls the switching action. 0.03 is set to be the threshold value and it is the normalized amplitude other than interference. Pickup of low energy signals and noise will also be eliminated. To get the pitch of speech signals the threshold can be varied. It helps in detecting the pitch of an attenuated speech signal [2].

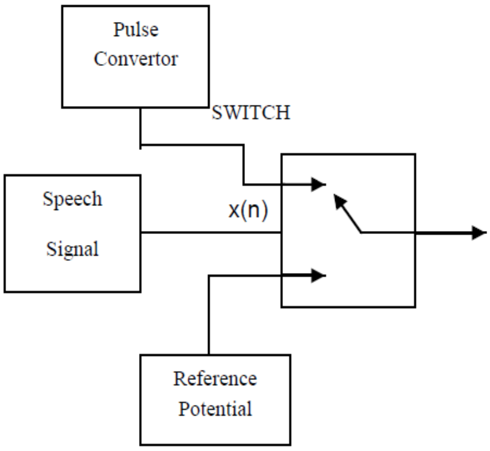

Fig. 3 Block Diagram of Pulse converter

B. Filter

A low pass filter is used and its stop-band frequency is 1 KHz. High frequency signals and noise are filtered, because High frequency signals and noise interfere the speech signal.

C. Window Function

After multiplying with a Rectangular window frame of time period W(n)=0.05ms, the pulse stream is extracted. It allows pitch estimation over a fixed time duration, and it also helps to extract short time pitch effectively .

D. Counter

Up counter is used to count the pulses within specified window length and counter will be reset at the end of the window time period. The pitch of the waveform can be plotted in that time instant[4]. We can calculate the pitch by the formula shown below:

$$Pitch = \frac{Number\ of\ pulses\ Counted\ \times Time\ period\ of\ window}{Time\ Period\ of\ pulse}$$

IV. RESULT

The signal whose pitch has to be estimated is the vowel /e/. Fig. 4 below shows the waveform of the speech signal with respect to time. The auto correlation waveform is shown in Fig. 5. The pitch is then calculated using by taking its peak and the time at which the peak occurs. The Fig. 6 shows the cepstrum waveform which is the frequency conversion of the speech signal in time domain [1]. The waveform in the Fig. 7 is the waveform of the pitch by using pulse modelling approach. As you can see the pitch is the peak that occurs at 138 Hz.

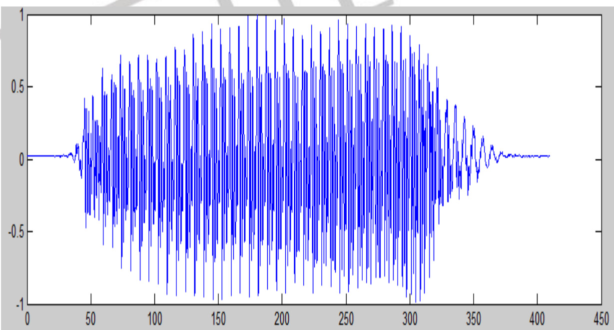

Fig. 4 Waveform of Speech Signal /e/

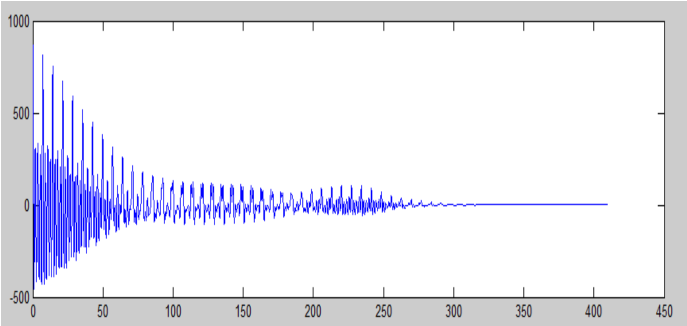

Fig. 5. Auto – Correlation of Speeh waveform

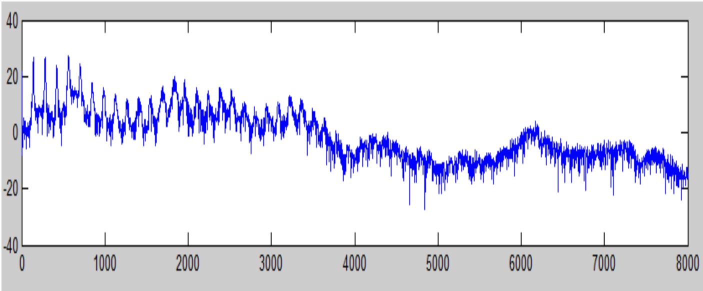

Fig. 6. Cepstrum Frequency Waveform

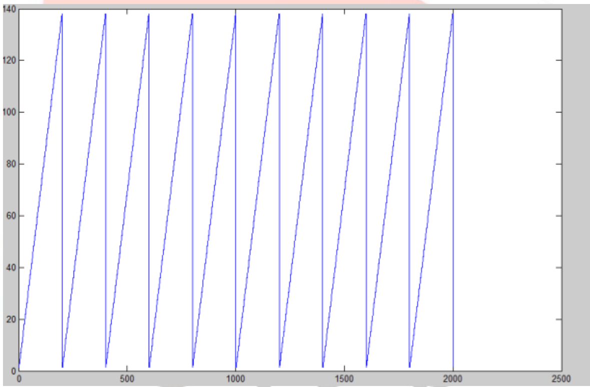

Fig. 7 Pitch using Pulse modeling

The table below shows the calculated pitch with noise added to the speech waveform. The table shows that the effect of noise is relatively low in the pulse modelled approach [3].

Table 1. Effect of noise in PMA

Method	Pitch (/e/) Without noise	Pitch(/e/) with noise
Auto Correlation	139.13	145.45
Cepstrum	137.93	130
Pulse Modelling	138	136

V. Conclusion

The Pulse Modelling approach depends on the time period for which the fundamental frequency in the time domain signal is high, hence it is immune to white noise. Also, due to its approach it can be used in real time application where as auto correlation cannot be used and the speed of the cepstrum depends on the hardware available.

REFERENCES

[1] Kelvin D'souza, Jason Fernandes,"Simplistic Implementation of Realtime Pitch Estimation using Pulse Modelling Approach," International Journal of Computer Applications, Volume 33– No.10, November 2011.

[2] Dominik Joho, MarenBennewitz, and Sven Behnke"Pitch Estimation Using Models of voice, speech on three levels", Proceedings of 2007 IEEE International Conferen Ce on Acoustics, Speech and Signal Processing (ICASSP).

[3] Marek4Szczerba&AndrzejCzy¿ewski "Pitch Estimation Enhancement Employing Neural Network- Based Music Prediction", Sound & VisionEngineeringDepartmentTechnicalUniversity of Gdañskul.Narutowicza..

[4] Hess W,"Pitch Determination of Speech signals: Algorithms and Devices,"Springer Verlag, Berlin Heidelberg, New York, Tokyo, 1983.

[5] Dushyant Sharma, Patrick. A. Naylor,"Evaluation of Pitch Estimation in Noisy Speech for Application in Non-Intrusive Speech Quality Assessment," Department of Electrical and Electronic Engineering, Imperial College, London,UK.

[6] Lawrence R. Rabiner, Michael J. Cheng, Aaron E. Rosenberg, Carol A. McGonegal "AComparative Performance Study of Several PitchDetection Algorithms" IEEE Transactions on Acoustics Speech and Signal Processin.