Method of Pitch Measurement Based on Packet Speech

Xulai Zhu
Anhui Jianzhu University, Hefei 230039, China
zhuxulai@126.com

Abstract. In view of the characteristics of speech signals, dullness in speech has obvious periodicity characteristics, and dullness is only quasi periodic. So we use short time averaging to estimate pitch period. In view of these characteristics, the speech signal processed in this paper is grouped according to the frequency, and then the pitch period is extracted after grouping, which reduces the complexity and accuracy of the processing.

1. Introduction
In speech signal processing, pitch period is one of the most important parameters. [1], it is very important to extract this parameter, such as speech coding, speech synthesis and speech recognition.

2. The difficult of the pitch measurement

2.1. The pitch does not have complete periodicity
The speech signal is a nonlinear time-varying signal. In order to facilitate the study, a short time window is used to deal with it. It is considered as a linear time invariant signal. The speech signal is produced by binary incentive way, namely the voiceless and voiced the voiced glottal excitation waveform is not a complete cycle sequence, and thus will not be able to estimate the pitch [2][3] accurately.

2.2. The influence of the channel
Because of the influence of the channel, the periodicity of pitch and formant will overlap. The peak value of the waveform may be very different from the original peak value, so we need to remove the influence of the vocal tract from the speech signal. The common methods are the inverse filter method [2], homomorphic analysis and so on. It is not easy to take out the sound information directly related to the vibration of the vocal cord and take out the voice signal directly in the vocal band.

2.3. Veracity of endpoint detection
The endpoint detection of a speech signal, an input speech signal, which is a silent segment, and which is a sound segment. In the audio segment, which is voiceless, which is voiced, it is difficult to accurately judgment, so it cannot accurately determine the start and end position of pitch [3].

2.4. Pitch cycle change
How to accurately detect the pitch period of different frequency segments, from the 80Hz of the bass male voice to the girl's 500Hz, the frequency span is large, the complexity and the difficulty are great.
In addition, for the pitch period of voiced sound, there may be many times of harmonic components. Compared with the fundamental components, harmonic components become the strongest component, which brings difficulty to pitch detection. [3]

3. Common pitch measurement method

3.1. Waveform estimation method
The pitch period is estimated according to the periodic peak of the speech signal waveform. A parallel processing method [3] [4], the simple waveform mixed together, pitch extraction is decided by the peak detector; data reduction method [3], according to some algorithm, to remove the redundant information from the waveform, subject only to retain information, namely the pitch signal audio signal the zero crossing rate; [3] [4], using the signal waveform by horizontal frequency, namely zero rate, repeating pattern, pitch estimation.

3.2. Correlation processing method
The pitch period is obtained by the autocorrelation function of the speech signal. No need for voiceless autocorrelation function, because it will not have periodic waveform for voiced, its autocorrelation function in time domain, the peak just appeared in the integer pitch period, according to the voiced autocorrelation function and voiced has the same periodicity, we can calculate the pitch.

3.3. Transformation method
The speech signal is time-varying in the time domain, can transform, Fu Liye transform, transform to the frequency domain, can also pitch estimation by transformation to the cepstral domain, and then eliminate the channel effect, belongs to incentive part information, consistent cycle cycle and speech signal excitation sequences, so according to this a feature of staterepresentation pitch [3].

4. Pitch detection based on voice packet

4.1. Introduction
The range of the pitch cycle is large, and the difference of the pitch cycle between the old and the children, the male voice and the female voice is quite obvious. Even the same person has a slightly different pitch cycle in a different environment. So we first group the frequency of the sound to reduce the scope of the gene cycle detection, reduce the complexity and increase the accuracy.

4.2. Simulation process
The sampling rate of the speech signal is 8kH, the window sequence uses 300 points Hamming window, and the frame is stacked 200 points. The concrete steps are as follows:

(1) Using 900Hz low pass filter to filter the speech signal \{x (n)\}, remove the first 15 output values and get \{X ' (n)\};
(2) The signal \{X ' (n)\} is clipped. The clipper uses three level central clipping, and the signal \{y(n)\} is obtained;
(3) The band pass filter of 40Hz-290Hz is used to filter the signal \{X ' (n)\}, and the signal \{y' (n)\} is obtained;
(4) The band pass filter is used to filter the signal \{X ' (n)\} with the 290Hz-520Hz band pass filter, and the signal \{y' (n)\} is obtained;
(5) The signal \{y' (n)\} and \{y' (n)\} are used to calculate the pitch period by autocorrelation.

4.3. Simulation result
In the simulation environment, we use Microsoft Windows XP Professional version 2002, Service Pack 3, Intel (R) Core (TM) 2 Duo, Core, the memory.
Figure 1. Is a time wave, with respect to the number of sampling points and the normalized sound pressure.

Figure 2. Is a base period value without frequency grouping.

Figure 3. Is the gene cycle value derived from the frequency from 60-290 Hz.
Figure 4. is the frequency of the frequency of the gene from 290-520 Hz.

It can be seen from the simulation diagram of pitch period value that the pitch period value after frequency grouping is more accurate and clear, which shows that the practice of this paper has certain effect and practical significance.

5. Concluding remarks
With the gradual maturity of voice technology, the pitch detection technology in the voice signal is constantly developing and improving. The method proposed in this paper improves the accuracy of pitch period to a certain extent, but increases some additional expenses in time. In addition, whether it has the same effect for different environments and different people to detect, it needs to be verified in the future.

Acknowledgments
This paper is funded by the project of scientific research of the Anhui provincial education department. The projects are KJ2016A821 and KJ2016JD13.

References
[1] Sung M o-kung, Elements of low power design for integrated systems[M], Low Power Electronics and Design, 2003.
[2] Andrei Grebennikov. RF and microwave power amplifier design[M]. McGraw-Hill Companies, 2005.
[3] Katsumi Sakakibara, et al. Backoff Algorithm with Release Stages for Slotted ALOHA Systems [J]. ECTI Transactions On Electrical Engineering, 2005, 3 (1): 59-70.
[4] K.Yu, I.B. Collings. Performance of low-complexity code acquisition for direct-sequence spread spectrum systems. IEE Proc-Commun, 2003, 150 (6):453–460.