A PITCH DETECTION ALGORITHM BASED ON AMDF AND ACF

Li Hui, Bei-qian Dai, Lu Wei

MOE-Microsoft Key Laboratory of Multimedia Computing and Communication,
University of Science and Technology of China
(hli@ustc.edu.cn, bqdai@ustc.edu.cn, luwei@ustc.edu.cn)

ABSTRACT

The paper proposes a pitch detection algorithm based on the short-time average magnitude difference function (AMDF) and the short-term autocorrelation function (ACF). At first, AMDF values are computed by AMDF algorithm for a frame of speech signal. And then ACF values are computed by ACF algorithm for the AMDF values. In order to decreases computational expense and complexity, the AMDF values of the frame of speech signal are then transformed into one bit signals. The method can also decrease the effects of amplitude and formants the speech signal for pitch detection. The pitch period is calculated by ACF algorithm for the one bit signals. The multiplication operation for short-time autocorrelation function of the one bit signals is replaced by simple addition operation. The algorithm can largely reduce its computational cost compared with calculating ACF directly. The experiments show that the algorithm can decreases computational expense and complexity compared with ACF algorithm. The algorithm can decrease the error rate of pitch detection compared with AMDF algorithm. The proposed algorithm is easy to implement in real-time signal processing.

1. INTRODUCTION

The problem of pitch estimation is a topic of research during all the evolution of digital speech processing and the pitch period is an important speech parameter. The pitch estimation plays a very important role in speech compression, speech coding, speech recognition and synthesis, as well as in voice translation. A good estimation of the pitch period is crucial to improving the performance of speech analysis and synthesis systems. There are many pitch detection algorithms such as the short-time average magnitude difference function (AMDF), short-term autocorrelation function (ACF) [1], spectrum, cepstrum [2], neural networks and wavelet transform [3] [4] algorithms. Although many pitch detection methods have been proposed, few of them have been built in special-purpose digital hardware capable of real-time operation [5].

The short-time AMDF and short-term ACF are often used in pitch detector. AMDF only needs subtraction, add and absolute operations for the pitch detection. AMDF algorithm has relatively low computational cost and is easy to implement. AMDF algorithm is used in the fields for the real-time processing. For example, LPC10e uses AMDF for pitch extraction. But one major disadvantage of the short-term AMDF approach for pitch detection is that the magnitude of the minimum in each frame is highly influenced by the intensity variation and the background noise of the speech signal.

In order to decrease the error rate of pitch detection of AMDF algorithm, this paper proposes a low complexity pitch detection algorithm based on AMDF and ACF. The proposed algorithm only needs subtraction, add and absolute operations. So the complexity computation of the proposed algorithm is low. At first, an AMDF values are computed by AMDF algorithm for a frame of speech signal and then the pitch period is computed by ACF algorithm for the AMDF values. But the ACF algorithm needs multiplication operation. To void use multiplication operation and decrease the computational expense, AMDF values of the frame of speech signal are transformed into one bit signals. The transform step can also decrease the effects of amplitude and formants the speech signal for pitch detection. The pitch period is then calculated by ACF algorithm for the one bit signals. The multiplication operations of autocorrelation function of the one bit signals are transformed into simple addition operations. The experiments show that the algorithm can decreases computational expense and complexity compared with ACF algorithm. The error rate of pitch detection is decreased compared with AMDF algorithm. The effects of the speech signal amplitude and formants for pitch detection can also be decreased.

2. AMDF PITCH DETECTION ALGORITHM

For periodic signal s(n) of period T, the difference function x(m)=s(n+m)-s(n) will be approximately zero for m=0, ±T, ±2T, ….Because the voiced speech signals are quasi-periodic, the difference function x(m)=s(n+m)-s(n) will be small if m=0, ±T, ±2T, …. Based on the reason, the

short-time average magnitude difference function of a frame of speech signal is defined as:

$$x_w(m) = \frac{1}{N-m-1} \sum_{n=0}^{N-m-1} \left| s_w(n+m) - s_w(n) \right| \qquad (1)$$

where $s_w(n)$ is speech signal. N is the length of a frame of speech signal. The range of m is between 0 and N. When a frame of speech signal is quasi-periodic, the values of the short-time average magnitude difference function of a frame of speech signal are quasi-periodic.

$x_w(m)$ has the minimum value in the position of the integer times of the pitch period. The position of the minimum value is found and the distance between the position and origin point is calculated. And then the pitch period can be obtain. The position of the minimum value may not match the pitch period because of the effect of the amplitude, noise and formant of speech signal. A frame of speech signal （sampling frequency is 8kHz，180 samples / a frame of speech signal） is shown in Figure 1.a. The AMDF values are shown in Figure 1.b. The pitch period T2 using AMDF algorithm is 15.5ms or 124 samples. The pitch error is made by the intensity variation, formants and the background noise of the speech signal. But the correct pitch period T1 is 3.25ms or 26 samples.

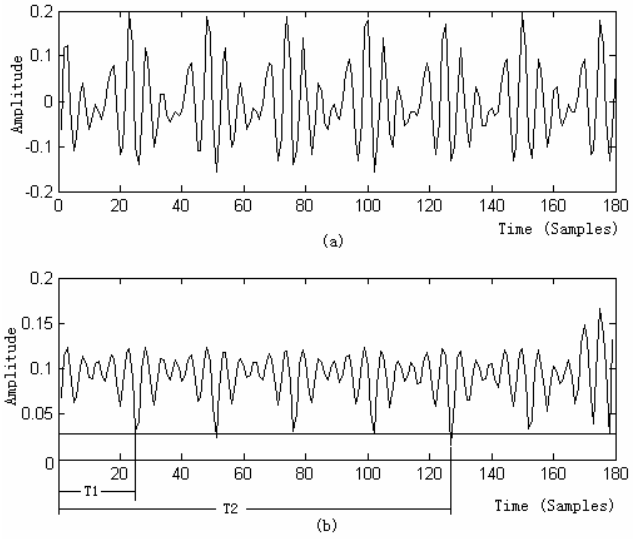

Figure 1 (a)A frame of speech signal; (b) AMDF values of a frame of speech signal

3. A PITCH DETECTION ALGORITHM BASED ON AMDF AND ACF

In order to decrease the effect of the amplitude and formants of speech signal for pitch detection, this paper proposes a pitch detection algorithm based on AMDF and ACF. The function block diagram of the pitch detection algorithm based on AMDF and ACF is shown in Figure 2.

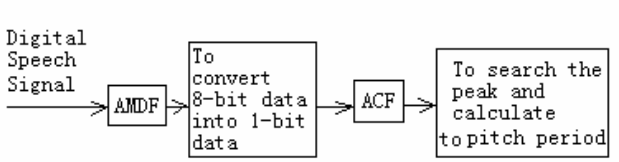

Figure 2 Function block diagram of the pitch detection algorithm based on AMDF and ACF

At first, the AMDF values are obtained for a frame of speech signal. And then the pitch period is obtained by ACF for the AMDF value. The ACF algorithm needs multiplication operation. In order to void to use multiplication operation, AMDF values of the frame of speech signal are transformed into one bit signals. According to the maximum $V_{\max}$ and minimum value $V_{\max}$ of the AMDF values of the frame of speech signal, the clipping level θ is chosen. The clipping level θ equals $\alpha(V_{\max} + V_{\min})$. The factor α is set to 0.4. When the value of that the short-time AMDF value is less than θ the value is set to 1. When the value of that the short-time AMDF value is greater than θ the value is set to 0. The short-time AMDF signal is converted into the one bit signal. And then the pitch period is calculated using the short-term autocorrelation function algorithm. The short-term autocorrelation function is defined as：

$$R(k) = \sum_{n=0}^{N-k-1} x(n)x(n+k) \qquad (2)$$

where $x(n)$ is one bit signal. The signal $x(n)$ is 1 or 0. $x(n)x(n+k)$ equals 1 when both of $x(n)$ and $x(n+k)$ equal 1. $x(n)x(n+k)$ equals 0 when one of $x(n)$ and $x(n+k)$ equals 0. The multiplication operation of autocorrelation function for the one bit signals is transformed into simple addition operation. The reduction in computational expense and complexity is very promising. The method can also decrease the effects of amplitude and formants the speech signal for pitch detection.

According to the variable range of the pitch period of speech, the variable range is between 16 samples and 140 samples (8kHz sampling frequency) [6]. The position of the peak of the autocorrelation value R(k) is searched in the range between k=16 and k=160. The distance between the position and k=0 is calculated, and then the pitch period of a frame of speech signal is obtained. The accuracy of the pitch detection is increased compared with AMDF algorithm. Figure 3.a shows the AMDF value of a frame of speech signal as same as Figure 1.a. Figure 3.b shows the one bit signal that has been transformed. Figure 3.c shows the autocorrelation value R(k). R(k) has the peak in the position of integrate times of pitch period. The position of peak is found and the distance between the position of peak

and the position of $k = 0$ is calculated. The correct pitch period is 3.25ms（26sampls，the sampling frequency is 8kHz）.

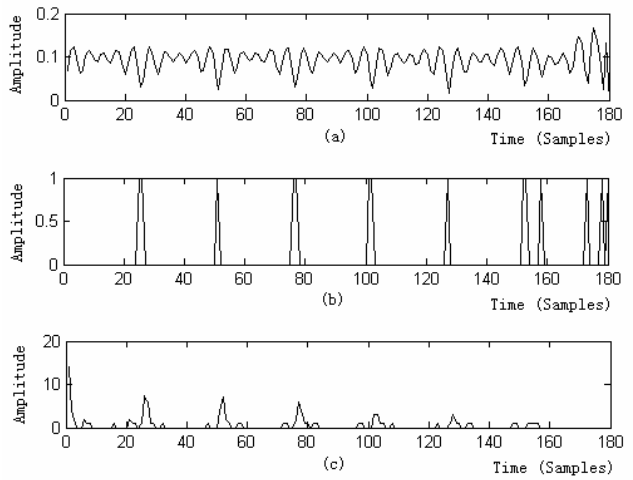

Figure 3 (a) AMDF values of a frame of speech signal; (b) The one bit signal of AMDF values; (c) ACF values of the one bit signal

The error rate is decreased using the pitch detection algorithm based on AMDF and ACF compared with AMDF. But the AMDF values above are not very good enough for pitch detection. Every AMDF value $x_w(m)$ has N-m-1 magnitude difference values $|s_w(n+m)-s_w(n)|$. The more great the parameter m in the equation (1) is, the less the number of the magnitude difference value $|s_w(n+m)-s_w(n)|$ is. When $(N-m-1)\times T_s$ (T_s is sampling period) is less than the pitch period, AMDF value $x_w(m)$ can not show the periodic feature of the frame of the speech signal.

A frame of speech signal is shown in Figure 4.a. The AMDF signals of the frame of speech signal are shown in Figure 4.b. If the parameter m is greater than 126, the AMDF value $x_w(m)$ can not show the periodic feature of the frame of the speech signal. It will increase the error rate of pitch detection.

In order to improve the performance of the pitch detection algorithm based on AMDF and ACF, the short-time average magnitude difference function is redefined as:

$$x_w(m) = \frac{1}{N-1}\sum_{n=2N-1}^{N} |s_w(n-m)-s_w(n)| \qquad (3)$$

where $s_w(n)$ has two frames of speech signal including current and previous frame of speech signal. N is the length of a frame of speech signal. Every AMDF value $x_w(m)$ has N magnitude difference values. The shot-time average magnitude difference values made by using equation (3) have the more periodic property than the shot-time average magnitude difference values made by using equation (1).

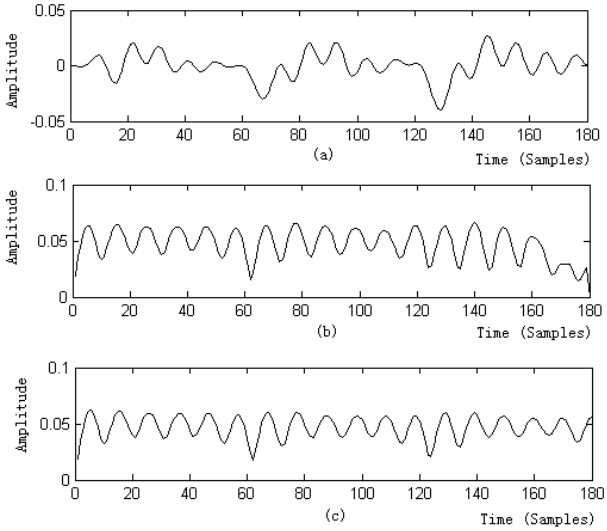

Figure 4 (a) A frame of speech signal; (b) AMDF values of a frame of speech signal made by equation (1); (c) AMDF values of a frame of speech signal made by equation (3)

4. EXPERIMENTS AND RESULTS

Figure 5 shows the performance comparison with three kinds of different pitch detection algorithm, which are the pitch detection based on AMDF, average magnitude difference function and autocorrelation function. The speech signal is a sentence "What about your family?" （ 8kHz sampling frequency, total 54 frames ）. The original speech signal is shown in Figure 5.a.

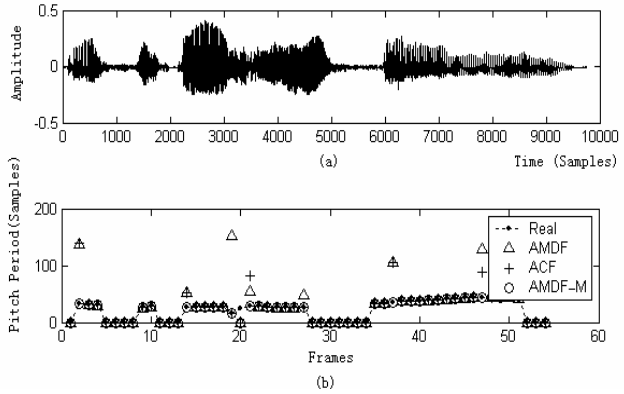

Figure 5 (a) A sentence of original speech signal; (b) Comparison of three kinds of pitch detection algorithms

Figure 5.b shows the comparison results of using three kinds of different pitch detection algorithms for the same sentence. All of three kinds of different pitch detection algorithms do not use error correlation or non-linear smoothing and correct technology. The correct pitch period

(-·-·- Real), the results of using the pitch detection based on AMDF (AMDF_M), average magnitude difference function (AMDF) and autocorrelation function (ACF) for the same sentence are shown in Figure 5.b. The error rate of using AMDF algorithm is 0.13. The error rate of using ACF algorithm is 0.1. The error rate of using he pitch detection based on AMDF algorithm is 0.02.

Table 1 shows the comparison of error rates of pitch detection of three kinds of different pitch detection algorithms for the 1280 frames of speech signal. The experiments show that the error rate of pitch detection of the proposed algorithm is approximately equal to the error rate of pitch detection of the ACF algorithm. The proposed algorithm can decrease the error rate of pitch detection compared with AMDF algorithm.

Table 1 Comparison of error rate of pitch detection

AMDF	ACF	AMDF_M
15%	9 %	8%

5. CONCLUSIONS

This paper proposes a pitch detection algorithm based on AMDF and ACF. Because the proposed algorithm does not need multiplication operations the algorithm can decrease computational expense and complexity compared with ACF algorithm. The error rate of pitch detection is decreased compared with AMDF algorithm. The effects of the speech signal magnitude and formants for pitch detection can also be decreased. Experiment results show that the proposed algorithms are easy to implement in real-time signal processing.

6. ACKNOWLEDGEMENTS

This work has been supported by the Science Research Fund of MOE-Microsoft Key Laboratory of Multimedia Computing and Communication (No.05071810) and the National Science Foundation of China (No.60272039).

7. REFERENCES

[1] Xiao-Dan Mei, Jengshyang Pan, Sheng-He Sun, Efficient algorithms for speech pitch estimation, Intelligent Multimedia, Video and Speech Processing, 2001, pp. 421 – 424, 2001

[2] Fangming Wang; Yip, P, Cepstrum analysis using discrete trigonometric transforms, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on Volume 39, Issue 2, pp 538 – 541, 1991

[3] Li Ji Chun, Du Li Ming, A Pitch Detection Algorithm Based on Edges Feature Extraction，Acta Electronic Sinica, Vol.31, No.10, pp. 1500~1502, 2003

[4]Du Limin,Hou ZiQiang, Determination of the Instants of Glottal Closure from Speech Wave Using Wavelet Transfor，ICSP'96 Beijing: IEEE & PHEI, pp. 473~476，1996

[5] J.J.Dubnowski, R.Schafer and L.R.Rabiner， "Real-time Digital Hardware Pitch Detector", IEEE Trans. Acoust., Speech and Signal Processing, Vol. ACSSP-24, No.1, 1976.

[6] Hu Jian Ling, Xu Sheng, Chen Jian, Spectrum Based Voiced/Unvoiced Decision, Journal of Data Acquisition & Processing, Vol. 17, No.1, pp 20~24, 2002.