Extended Average Magnitude Difference Function Based Pitch Detection

Ghulam Muhammad
Department of Computer Engineering, King Saud University, Saudi Arabia

Abstract: *This paper presents a new extended average magnitude difference function for noise robust pitch detection. Average magnitude difference function based algorithms are suitable for real time operations, but suffer from incorrect pitch detection in noisy conditions. The proposed new extended average magnitude difference function involves in sufficient number of averaging for all lag values compared to the original average magnitude difference function, and thereby eliminates the falling tendency of the average magnitude difference function without emphasizing pitch harmonics at higher lags, which is a severe limitation of other existing improvements of the average magnitude difference function. A noise robust post processing that explores the contribution of each frequency channel is also presented. Experimental results on Keele pitch database in different noise level, both with white and color noise, shows the superiority of the proposed extended average magnitude difference function based pitch detection method over other methods based on average magnitude difference function.*

Keywords: *Pitch detection, AMDF, EAMDF, and noise robust.*

Received May 13, 2009; accepted January 3, 2010

1. Introduction

Precise calculation of pitch in speech signal has demonstrated to be a basic task in almost all areas of speech research including speech/music recognition, speaker recognition/verification, speech forensics, voice-enabled security, etc., a wide range of perceptual models and algorithms using a variety of techniques and a varying degree of accuracy to extract pitch exist. However, the pitch detection algorithms face a real challenge in presence of noise [11]. Noise often buries real harmonic peaks and creates false peaks, which cause problems to find pitch [8].

There are many pitch detection algorithms such as Average Magnitude Difference Function (AMDF) [14], Short-Term Autocorrelation Function (ACF) [10], cepstrum [16], and different combinations and modifications of them [1, 2, 3, 6, 7, 12, 17]. Among the algorithms, AMDF is used for real time application as it involves less computation, and it is the focus of this paper. However, AMDF in [14] has 'falling tendency' at the later half of a frame and that makes it erroneous even in less noisy condition. Also, noise and intensity variation produce false pitch in AMDF. To overcome this situation, several improvements of AMDF are proposed in the literature. For example, High Resolution AMDF (HRAMDF) [5] and circular AMDF (CAMDF) [18] conquer the falling tendency of AMDF by well averaging all the lag values. But in doing so, these methods generate new double pitch errors by emphasizing pitch harmonics at higher lags.

In this paper, we introduce a new Extended AMDF (EAMDF) that eliminates the falling tendency of AMDF without emphasizing pitch harmonics at higher lags. The proposed EAMDF is similar to AMDF except it spreads over second half of the previous frame, the current frame, and the first half of the next frame. With this length, the EAMDF averages a reasonable amount for all the lag values, and provides better smoothing than the other improvements.

We also explore channel contribution to detect true pitch. The EAMDF is applied on a bank of band pass filters, and each channel EAMDF is verified for a proper pitch candidate. Pitch candidates are assigned weight with respect to presence or absence of peaks at their harmonics. The highest weighted candidate is then selected as pitch.

All the experiments are carried out using Keele pitch reference database [13]. There are a quite a number of literatures that present with the result on Keele database. We choose the same database, which is publicly available, to compare with other methods.

The paper is organized as follows: section 2 reviews AMDF, HRAMDF, and CAMDF; section 3 introduces the proposed EAMDF; section 4 describes the proposed EAMDF based pitch detection method, and section 5 gives experimental results with discussion. Section 6 presents results under restaurant noise. Finally, section 7 draws some conclusion.

2. Review of AMDF and its Variations

As this paper focuses on the improvement of AMDF-based pitch detection, we describe some major and well known AMDF-based pitch detection algorithms.

2.1 AMDF

The original AMDF was proposed in [14] and it is defined as

$$D(\tau)=\frac{1}{N-\tau-1}\sum_{n=0}^{N-\tau-1}|x(n)-x(n+\tau)| \quad (1)$$

where $x(n)$ is the speech sample sequence multiplied by a rectangular window of length N, and τ is the lag number. The range of τ is between 0 and N-1, and the constant term outside summation is for normalization. For a periodic or quasi periodic signal with a period of T_P, eq should exhibit minimum at lag T_P and minimum peaks with lower degree at its multiple. In general, a rough estimation of pitch is derived by

$$T_P=\arg_{\tau}\underset{\tau=\tau_{\min}}{\overset{\tau_{\max}}{MIN}}(D(\tau)) \quad (2)$$

where τ_{min} and τ_{max} correspond to possible minimum and maximum pitch periods in samples. In equation 1 less data is involved to calculate D at higher lags, because speech signal is weighted by a rectangular window and outside the window the values are zero. Therefore, AMDF cannot show periodic nature at the later half of a frame and it is often called as the 'falling trend' in literature. Furthermore, D is sensitive to noise and intensity, which makes it difficult to produce minimum at T_P. In noisy conditions, it can output minimum at $T_P/2$ or $2T_P$, which is normally termed as 'half pitch error' or 'double pitch error', respectively.

Figure 1 (b) shows an example of double pitch error using AMDF. In the figure, speech is a female voiced frame as shown in Figure 1(a) contaminated with white noise at signal-to-noise ratio (SNR) = -5 dB. The true pitch period is at lag 85, but AMDF falsely detect it at its double (at lag 170).

2.2. HRAMDF

To avoid the falling tendency of AMDF, HRAMDF was proposed in the speech coding standard LPC-10 [5]. HRAMDF is defined as

$$D_H(\tau)=\sum_{n=(N/2-\tau)/2+1}^{(N/2-\tau)/2+N/2}|x(n)-x(n+\tau)| \quad (3)$$

There are two major differences between equations 1 and 3. Unlike AMDF, HRAMDF involves two speech frames and all the lags are well averaged resulting in the elimination of the falling trend of AMDF. However, as pitch multiples are emphasized, this type of modification introduces mostly double pitch errors. From Figure 1(c), we can see that though HRAMDF successfully eliminates falling tendency of AMDF, double pitch error prevails.

2.3. CAMDF

CAMDF was proposed in [18] and is defined by

$$D_C(\tau)=\sum_{n=0}^{N-1}|x(\mathrm{mod}(n+\tau,N))-x(n)| \quad (4)$$

where mod(n+τ, N) represents the modulo operation. This function is symmetric around τ = N/2, meaning that it can produce pitch only within N/2. To determine pitch within full-length (N), it needs double sized frame. The lags are equally averaged and also there is no falling tendency. It has better performance than HRAMDF, however, magnitudes at all the pitch multiples are enhanced introducing new errors. Figure 1(d) shows CAMDF calculating over a double sized frame. Double pitch error is not eliminated, but the difference of magnitudes between true pitch and double pitch is reduced.

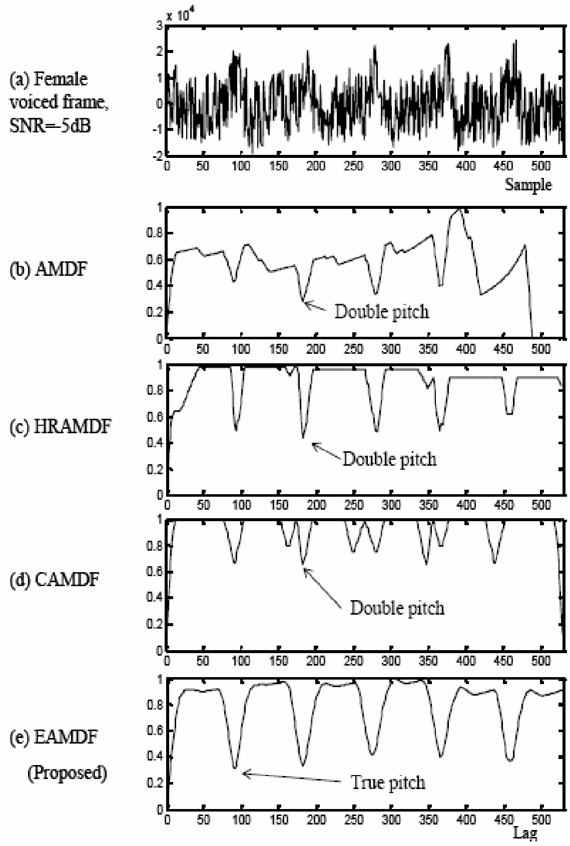

Figure 1. Comparison between (b) AMDF, (c) HRAMDF, (d) CAMDF, and (e) the proposed EAMDF on (a) a female voiced frame with SNR = -5 dB (white noise). The proposed EAMDF extracts the true pitch, while the others produce double pitch error.

3. The Proposed Extended AMDF

HRAMDF and CAMDF almost successfully avoid the falling trend of AMDF by well averaging for all lag values. However, giving equal emphasis on all lags enhances peaks (minimums) at multiple pitch periods and these results in frequent double pitch errors.

We propose a new EAMDF that avoids enhancing peaks at pitch multiples and at the same time eliminates the falling trend of AMDF. The proposed EAMDF is defined as:

$$D_E(\tau) = \frac{1}{N-\tau} \sum_{n=-N/2}^{N+N/2-\tau} | x(n) - x(n+\tau) | \quad (5)$$

In contrast of the original AMDF, the calculation of EAMDF is spread over three frames: second half of the previous frame, the current frame, and first half of the next frame (so a total of 2*N* samples). It does not give equal weights to all lags, but still at some extent eliminates the falling trend by averaging double number of samples compared to the original AMDF. It can be seen in Figure 1 that only EAMDF as shown in Figure 1(e) can correctly determine the pitch period at lag 85. The proposed function has also greater smoothing power than the other improvements of AMDF, by taking into account some samples of the previous and the next frames. The EAMDF is applied to detect pitch in noisy speech signal and the procedure is described in the next section.

4. The Proposed EAMDF-Based Pitch Detection

Figure 2 shows a block diagram of the proposed EAMDF-based pitch detection method. First, input speech is passed through a bank of four band-pass filters. The bands of the filters are in the range of 50-200 Hz, 150-300 Hz, 250-400 Hz, and 350-500 Hz, respectively. High frequency components are blocked, because they do not contain significant pitch information. We could have used a low-pass filter with a cut off frequency of 500 Hz instead of using four individual band-pass filters encourages us to use several overlapping band-pass filters instead of one low-pass filter. Figure 3 illustrates such an example. Figure 3 (a) shows a female voiced frame with SNR = -5 dB.

Figure 3 (b), (c), (d) and (e) output the results of EAMDF on the four band-pass filters, and 3 (f) gives that on the low-pass filter. The minimum peaks that determine the pitch are encircled in the figures. The actual pitch is 89, which is wrongly detected with double pitch error (lag = 178) on the low-pass filter. However, first and second band-pass filters can correctly determine the pitch. This finding concludes that if we can select the appropriate channels, there is a high possibility to reduce half pitch or double pitch errors. After band-pass filtering, each filter output is half-wave rectified and center clipped in a fairly conventional way. The half-wave rectification is used to mimic phase-lock property of human and center clipping is used to remove unwanted noise up to some limit. These two blocks can also be used as a way to reduce computation.

The proposed EAMDF, described in section 3, is applied on center clipped speech. EAMDF values are normalized by the maximum value in a frame. In theory, EAMDF should exhibit minimum at pitch period, however, due to noise effect and intensity variation, there are some instances where it shows minimum at half or double the pitch period. To eliminate these errors, we apply some post processing that includes candidate refinement and weight assignment, channel selection, and final pitch detection. In the experiments, we consider 50- 450 Hz to be the pitch range, and calculation and selection are made inside the corresponding lag of the EAMDF.

For candidate refinement, we first mirror EAMDF values about lag axis by '1 – EAMDF' to convert the notches into peaks (Fig. 4). Then, all the peaks that are above 0.2 in amplitude and greater than the values at ±2 lags, are retained. We call these peaks as pitch candidates. After extracting pitch candidates, they are assigned some weights (initially weights are one) by exploring periodicity of pitch.

For a pitch candidate at t_1, its weight will be increased by 1/2 if there is a candidate at $2t_1$, by another 1/3 if there is a candidate at $3t_1$, and so on. Similarly, its weight will be decreased by 1/2 or 1/3 in case of absence of candidate at $2t_1$ or $3t_1$, respectively.

The presence / absence of candidates (at $2t_1$ or $3t_1$) is determined within ±0.5 ms of $2t_1$ or $3t_1$. This kind of weight assignment enhances the possibility to detect true pitch, as well as suppress a false pitch. For example, in Figure 4, candidate at true pitch will have more weight than the candidate at double pitch.

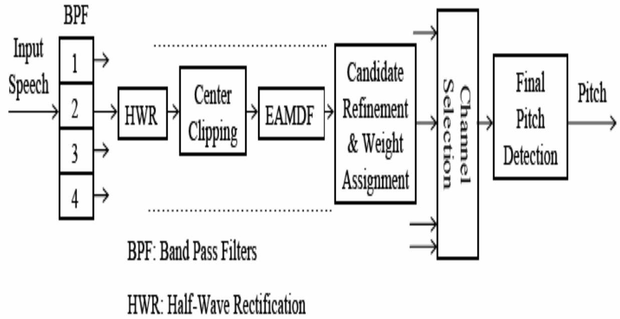

Figure 2. Block diagram of the EAMDF-based pitch detection.

After weight assignment, the channel(s) with highest weighted candidate(s) will be selected. If the selected channels contain the same highest weighted candidate, then this candidate is detected as the final pitch. Otherwise the candidates of the channels are added lag by lag, and the same weight assignment procedure is applied on the added candidates. Then the candidate with highest weight is selected as the final pitch. If several candidates with same highest weight exist, the candidate with the maximum height is detected as pitch.

5. Experiments

5.1. Database

The performance of the proposed EAMDF-based pitch detection algorithm is evaluated using the *Keele* pitch extraction reference database [13]. There are five male (M) and five female (F) speakers in the database. The data consists of a phonetically balanced English text of around 35 second. Speech data is sampled at 20 kHz with 16-bit resolution. The pitch values are provided at 100 Hz frame rate with 26.5 ms window. We added white Gaussian noise to the clean speech at different SNR (SNR = 10 dB, 5 dB, 0 dB, -5dB, -10 dB).

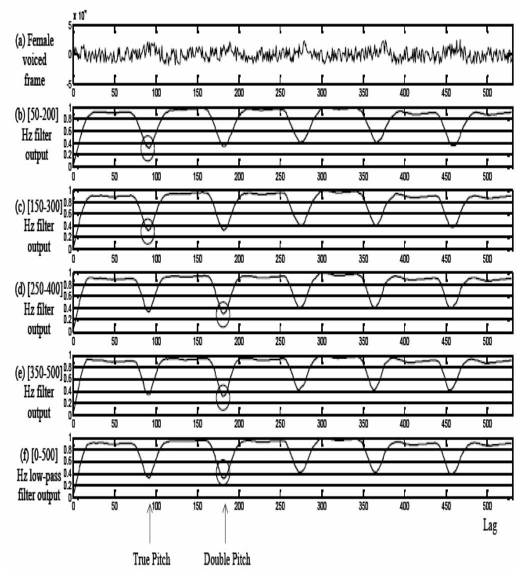

Figure 3. Illustration of the use of several band-pass filters [(b) to (e)] instead of a single low-pass filter (f) for (a) female voiced frame. The first two filters can correctly determine pitch from minimum peak (encircled) obtained by the EAMDF, while the low-pass filter produces double pitch error.

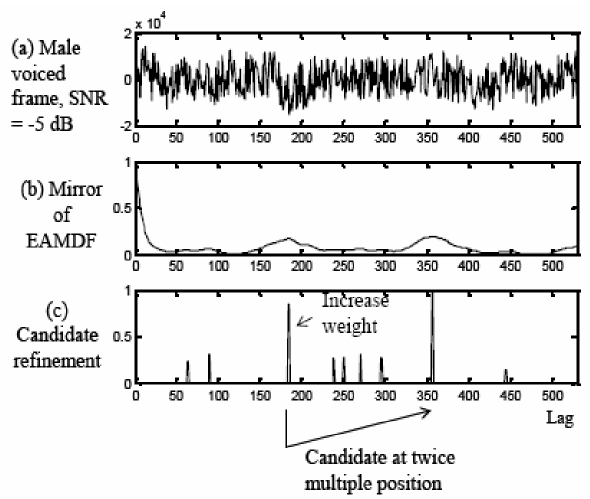

Figure 4. (a) A male voiced frame with SNR = -5 dB (white noise), (b) mirroring of EAMDF and (c) candidate refinement and weight assignment of the noisy speech. Mirroring is used to convert notches of EAMDF into peaks. Higher peaks (within pitch range) are retained and assigned weights corresponding to presence / absence of harmonics.

In our experiment, we use all the female (F1, F2, F3, F4, and F5) and male voices (M1, M2, M3, M4, and M5) from *Keele* database. The results of individual female speaker are averaged to give average result for female speakers, and similarly the results of individual male speakers are averaged to give average result for male speakers.

5.2. Experimental Setup

We evaluate AMDF, HRAMDF, CAMDF, and the proposed EAMDF-based pitch detection algorithms on Keele pitch database. AMDF, HRAMDF, and CAMDF involve pitch detection using only minimum peak. For a fair comparison, we also present the result using EAMDF without any post processing.

The window length is set to 26.5 ms and frame rate is 100 Hz. The errors are reported in terms of average percentage gross pitch error (%GPE). Gross pitch error is measured if the measured pitch defers 1 ms from the actual pitch. The true pitch values are obtained from the original database. Some instances of pitch values from that database correspond to half / double pitch errors or '-1' that are manually corrected. Results are shown based on these manually corrected values. Segmental SNR is used for evaluation.

Table 1. Comparison of different methods in terms of %GPE. The results are on male speech contaminated with white noise at different SNR.

Method	Clean	10 dB	5 dB	0 dB	-5 dB	-10 dB
AMDF	9.69	13.44	17.15	20.36	34.89	71.46
HRAMDF	8.01	11.32	16.13	18.84	31.43	68.79
CAMDF	7.26	10.69	15.46	17.79	29.45	66.39
EAMDF	5.36	8.43	13.57	15.72	24.50	61.87
EAMDF + post processing	0.52	2.47	4.10	6.19	12.43	38.12

Table 2. Comparison of different methods in terms of %GPE. The results are on female speech contaminated with white noise at different SNR.

Method	Clean	10 dB	5 dB	0 dB	-5 dB	-10 dB
AMDF	11.58	15.64	19.12	23.39	39.03	73.29
HRAMDF	10.32	13.52	18.32	21.31	36.39	70.63
CAMDF	9.47	12.09	17.58	19.84	34.61	68.78
EAMDF	6.80	10.19	16.89	17.10	29.47	63.54
EAMDF + post processing	0.83	2.96	3.15	4.98	10.99	41.69

Table 3. Comparison between the proposed EAMDF and WAC [8] at SNR = 0 dB for five males (M1, M2, M3, M4, and M5) speech.

Method	M1	M2	M3	M4	M5
WAC	16.01	43.67	33.46	16.38	31.42
EAMDF	15.58	41.80	31.04	15.80	30.14
EAMDF + post processing	6.03	16.31	12.86	6.31	11.71

Table 4. Comparison between the proposed EAMDF and WAC [8] at SNR = 0 dB for five females (F1, F2, F3, F4, and F5) speech.

Method	F1	F2	F3	F4	F5
WAC	44.41	46.37	42.81	54.76	44.13
EAMDF	16.04	16.78	15.54	18.12	15.89
EAMDF + post processing	4.86	4.91	4.88	5.04	4.79

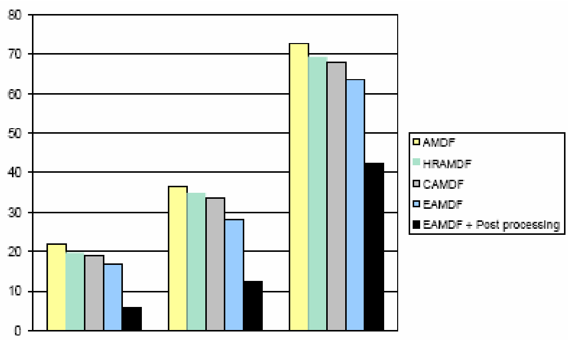

Figure 5. The average performance in terms of %GPE of different methods at low SNRs.

5.3. Results and Discussion

Tables 1 and 2 show %GPE of the different methods for male and female speech, respectively, at different SNR. From these tables we can see that the proposed EAMDF based pitch detection has the least %GPE for both male and female speech at different SNR. Even without post processing and only by selecting minimum as pitch period, the proposed EAMDF performs better than CAMDF. For example, in SNR = 0 dB, the original AMDF has 20.36% GPE, HRAMDF has 18.84%, CAMDF gives 17.79%, while EAMDF reduces it to 15.72% for male speech.

However, post processing that include candidate refinement and weight assignment, has a great impact on reducing error. For instance, applying post processing to EAMDF further reduces GPE to 6.19% from 15.72% in the above example.

Figure 5 shows the average performance (male and female) in terms of %GPE of the different methods at low SNRs (SNR = 0 dB, -5 dB, -10 dB). From the plots, the superiority of the proposed EAMDF based pitch detection method can easily be seen in very noisy condition.

A comparative evaluation in terms of half and double pitch errors for a female voiced speech segment consisting of 112 frames at SNR = 10 dB is illustrated in Figure 6. AMDF produces many half and double pitch errors, which are marked as deviation in dashed line. HRAMDF and CAMDF reduce the errors, however, at the cost of new double pitch error (encircled in the figure). EAMDF further reduces the error without introducing any new error, and the proposed EAMDF based pitch detection matches with the true pitches throughout the segment.

It can be noted that the proposed EAMDF is evaluated and compared with other AMDF-based family of pitch detection. However, we also compare the proposed method with one of the ACF-based pitch extraction algorithm called weighted autocorrelation (WAC) [15]. The WAC reported here includes no post-processing steps. The results are given in Table 3 for male speech and in table 4 for female speech. From these tables, we can see that proposed EAMDF has comparable performance with WAC for male speech and far better performance for female speech even without the post-processing. It can be mentioned that WAC involves many multiplication, whereas EAMDF involves only addition / subtraction.

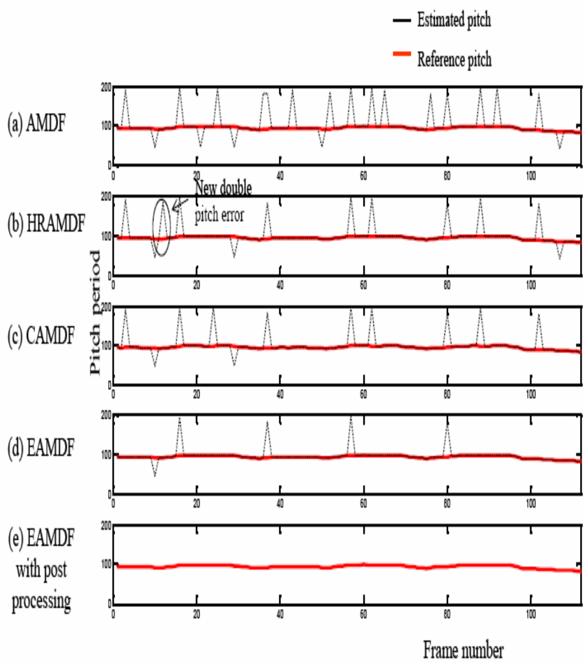

Figure 6. Pitch error in different methods for a female voiced segment of 112 frames at SNR = 10 dB. (a) AMDF produces many half and double pitch errors, while (b) HRAMDF and (c) CAMDF reduce error at the cost of new double pitch error (encircled). (d) EAMDF has less error, and (e) the proposed EAMDF based pitch detection has no error.

6. Evaluation in Restaurant Noise

We evaluate the proposed EAMDF-based PDA with AMDF, HRAMDF, and CAMDF in color noise also. The noise we consider is restaurant noise, which is taken from Freesound [4]. The noise was taken at a hotel during breakfast, and the filename was chosen as 22529_LG_Breakfast04.wav. The recording was done with 44.1 kHz, and we downsample it to 20 kHz. The waveform of the restaurant noise, its spectrogram, and pitch contour are shown in Fig. 7. This kind of noise is chosen in our experiment to evaluate the pitch detection algorithms in a real-world scenario. Restaurant noise is one of the difficult noises to deal with, because there are many human voices in this noise that introduces new pitches in addition to the pitch of main speaker. The noise is added artificially to the clean speech from the *Keele* pitch database at different SNR. Experimental setup for this evaluation is similar to as described in section 5.

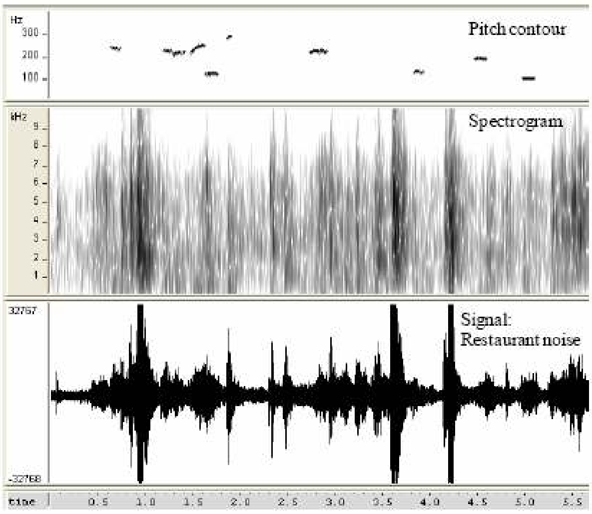

Figure 7. Wave signal of a restaurant noise, and its spectrogram and pitch contour.

The experimental results are shown in Table 5 and Table 6 for male and female speakers, respectively. From these tables, we can easily see the superiority of the proposed EAMDF-based PDA. If we compare these tables with Tables 1 and 2, we can find that %GPE in restaurant noise is much higher than in white noise. For example, at SNR = 0 dB and male speakers, for white noise EAMDF without post processing has 15.72% GPE, while for restaurant noise it has 32.47% GPE, which is almost double comparing to white noise. This is quite understandable with the fact that the background speakers present in restaurant noise contribute some pitch as shown in Figure 7 and this causes false detection of pitch of the main speaker. The false pitch is very difficult to remove even with the post-processing, unless we apply some multi-pitch tracking procedures.

7. Conclusions

A noise robust pitch detection method based on EAMDF was presented. First, EAMDF was introduced to overcome the shortcomings of the original AMDF and its existing improvements. The EAMDF spread over previous and next frames along with current frame, and thereby possesses greater smoothing power. Then post processing was applied on the EAMDF. An experiment shows efficient noise robustness of the proposed method both in white and color (restaurant) noise. This method can significantly contribute to speech/music discrimination, voice-enabled security, among others.

Table 5. Comparison of different methods in terms of %GPE. The results are on male speech contaminated with restaurant noise at different SNR.

Method	Clean	10 dB	5 dB	0 dB	-5 dB	-10 dB
AMDF	9.69	21.52	28.65	40.45	57.89	86.12
HRAMDF	8.01	19.22	24.76	37.23	50.41	76.72
CAMDF	7.26	18.93	22.98	36.12	45.73	71.33
EAMDF	5.36	15.27	20.28	32.47	41.33	69.40
EAMDF + post processing	0.52	5.67	12.54	20.49	31.17	52.67

Table 6. Comparison of different methods in terms of %GPE. The results are on female speech contaminated with restaurant noise at different SNR.

Method	Clean	10 dB	5 dB	0 dB	-5 dB	-10 dB
AMDF	11.58	25.32	33.24	41.25	57.68	86.78
HRAMDF	10.32	19.77	25.63	38.42	51.52	78.55
CAMDF	9.47	19.47	23.91	37.98	45.22	73.23
EAMDF	6.80	15.62	21.05	33.56	42.71	70.20
EAMDF + post processing	0.83	5.81	13.04	21.05	32.10	54.11

References

[1] Amado G. and Filho V., "Pitch Detection Algorithms Based on Zero-Cross Rate and Autocorrelation Function For Musical Notes," *in Proceedings of IEEE ICASSP'08,* pp. 449-454, 2008.

[2] Boril H. and Pollak P., "Direct Time Domain Fundamental Frequency Estimation of Speech in Noisy Conditions," *in Proceedings of European Signal Processing Conference*, pp. 1003-1006, 2004.

[3] Dong J., Wei X., Zhang Q., and Zhao Y., "Pitch Detection Using Circular Average Magnitude Difference Function Based on Wavelet Transform," *International Journal of Innovative Computing, Information and Control*, vol. 5, no. 9, pp. 2717-2724, 2009.

[4] Freesound:http://www.freesound.org/samplesViewSingle.php?id=22529, Last Visited 2009.

[5] Gu L. and Liu R., "The Government Standard Linear Predictive Coding Algorithm," *in Proceedings of Speech Technology*, pp. 40-49, 1982.

[6] Heckmann M., Joublin F., and Goerick C., "Combining Rate and Place Information For Robust Pitch Extraction," *in Proceedings of Interspeech*, pp. 2765-2768, 2007.

[7] Lindh J. and Eriksson A., "Robustness of Long Time Measures of Fundamental Frequency," *in Proceedings of Interspeech*, pp. 2025-2028, 2007.

[8] Mahadeva R. and Yegnanarayana B., "Extraction of Pitch in Adverse Conditions," *in Proceedings of IEEE ICASSP'04*, pp. 109-112, 2004.

[9] Mamun K., Sarker F., and Ghulam M., "A High Resolution Pitch Detection Algorithm Based on AMDF and ACF," *Computer Journal of Science Research*, vol. 1, no. 3, pp. 508-515, 2009.

[10] Mei X. D., Pan J., and Sun H., "Efficient Algorithms for Speech Pitch Estimation," *in Proceedings of Intelligent Multimedia, Video and Speech Processing*, pp. 421-424, 2001.

[11] Nakatani T. and Irino T., "Robust and Accurate Fundamental Frequency Estimation Based on Dominant Harmonic Components," *Computer Journal Acoustic Society America*, vol. 3, no. 3. pp. 3690-3700, 2004.

[12] Pelle P., "A Robust Pitch Extraction System Based on Phase Locked System," *in Proceedings of IEEE ICASSP*, pp. 249-252, 2006.

[13] Plante F., Meyer G., and Ainsworth A., "A Pitch Extraction Reference Database," *in Proceedings of Eurospeech'95*, pp. 837-840, 1995.

[14] Ross M., Shaffer H., Cohen A., Freudberg R., Manley H., "Average Magnitude Difference Function Pitch Extractor," *in Proceedings of IEEE Transactions Lon Speech and Audio*, pp. 353-362, 1974.

[15] Shimamura T. and Kobayashi H., "Weighted Autocorrelation for Pitch Extraction of Noisy Speech," *in Proceedings of IEEE Transactions Speech and Audio*, pp. 727-730, 2001.

[16] Wang F. and Yip P., "Cepstrum Analysis Using Discrete Trigonometric Transforms," *Computer Journal of IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 39, no. 2, pp. 538-541, 1991.

[17] Ying G., Jamieson H., and Mitchell C., "A Probabilistic Approach to AMDF Pitch Detection," *in Proceedings of ICSLP*, pp. 1201-1204, 1996.

[18] Zhang W., Xu G., and Wang Y., "Pitch Estimation Based on Circular AMDF," *in Proceedings of IEEE ICASSP*, pp. 341-344, 2002.

Ghulam Muhammad received his BSC in computer science and engineering in 1997, from Bangladesh University of Engineering and Technology, and MSc and PhD degrees in 2003 and 2006, respectively, from Toyohashi University of Technology, Japan. After serving as a JSPS (Japan Society for the Promotion of Science) fellow, he joined as a faculty member in the College of Computer and Information Sciences at King Saud University, Saudi Arabia. His research interests include automatic speech recognition, signal processing, and multimedia forensics.