Pitch Extraction Using Half-wave Rectification in Narrow-Band Noise

Md. Saifur Rahman[1†] and Miss. Nargis Parvin[2††],

[1]Department of Information and Communication Technology, Comilla University, Cumilla, Bangladesh
[2]Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology (BAIUST) Cumilla Cantonment, Cumilla, Bangladesh

Summary
In this paper, we propose a pitch extraction method effective in narrow-band noise. We first use the fourth-root spectrum instead of the logarithm spectrum by reducing the effect of noise. The noise-free fourth-root spectrum is then flattened inspired by the lifter approach for reducing the effect of vocal tract characteristics. In our proposed method, we use the half-wave rectification approach instead of clipping operation on liftered spectrum to suppress the noise effect and maintain the periodicity in between the harmonics. Therefore, we can extract the more accurate pitch. Through experiments, the excellent pitch extraction performance of the proposed method is validated. It is also shown that the computation of the proposed method is simple and fast.
Key words:
Pitch, fourth-root, log spectrum, lifter, narrow-band, vocal tract, half-wave rectification.

1. Introduction

For voiced speech, pitch period is defined as the inverse of fundamental frequency of the excitation signal. Pitch period extraction is a key technique to understand most acoustical phenomena in speech communication and plays an important role in speech processing applications such as speech coding, speech recognition, speech enhancement, speech synthesis and so on. In the above systems, the system performance is significantly influenced by the accuracy of pitch extraction. Most of the pitch extraction methods in the literature are classified into three categories, as methods that use time domain properties, frequency domain properties, and both time and frequency domain properties of the speech signals [1][2].

Pitch extraction has proven to be a difficult task, even for speech in a noise-free environment. The clean speech waveform is not really periodic; it is quasi-periodic and highly non-stationary. On the other hand, when the speech signal is corrupted by noise, the reliability, and accuracy of pitch extraction algorithms face real challenges. Under noisy conditions, the periodic structure of the speech signal is destroyed so that the pitch extraction becomes an extremely complicated task. Among the conventional pitch extraction methods, the autocorrelation function (ACF) [3] is a straightforward computation one in the time domain and shows effectiveness against wide-band random noise such as white noise. The ACF corresponds to a correlation calculation between the input speech signal and its delayed version in the time domain, but it is also obtained by the inverse Fourier transform of the power spectrum of the speech signal. The ACF is, however, affected by the characteristics of the vocal tract. For reducing the vocal tract affect, many algorithms have been developed based on the ACF [4]-[12]. For example, YIN [4] focused on the conventional ACF, normalization, and interpolation to reduce the error rates in pitch extraction.

The average magnitude difference function (AMDF) [5] is a simplified version of the ACF, but provides a similar performance with the ACF. The AMDF treats a difference between the speech signal and its delayed version. It shows similar properties with the ACF. In [6], the AMDF was combined with the linear predictive analysis to eliminate the affect of vocal tract. Correntropy [7] also provides the similar properties to the ACF, which has a kernel function to transform the original signal into a high dimensional reproducing kernel Hilbert space (RKHS) in a nonlinear way. This transformation preserves the characteristics of the periodic signal. Higher order statistics [8] are also used to enhance the resolution of pitch extraction. However, the performance of correntropy in noisy environment is not so good. In [9], the harmonic sinusoidal autocorrelation (HSAC) was proposed. The pitch harmonic was utilized from the discrete cosine transform (DCT), and applied to the symmetric average magnitude sum function (SAMSF) for generating the periodic impulse-train to extract the pitch. In [9], many ideas and algorithms were embedded and the resulting pitch extractor is so complicated to implement. In [10], another pitch extraction method was proposed based on the time-frequency sparsity of speech signal by using an auditory filterbank. The auditory filterbank decomposes the speech signal into subbands. Then, the normalized autocorrelation function (NACF) [11] is applied to the subband signals, which are encoded for extracting the pitch.

Manuscript received October 5, 2020
Manuscript revised October 20, 2020
https://doi.org/10.22937/IJCSNS.2020.20.10.20

The NACF reduces more the variations in the signal amplitude than the ACF does. The filter bank approach in [10] is very effective, but inherently relies on a sophisticated post-processing technique to compensate for the pitch extraction errors.

In highly noisy environments, the two correlation-based methods; ACF and AMDF, are not good enough compared with the weighted autocorrelation function (WAF) [12]. The WAF also focuses on the ACF, but it is weighted by the inverse of the AMDF, resulting in an excellent pitch extractor in noisy environments. Most of the ACF based pitch extraction methods are effective in white noise. In general, the pitch extraction performance of the ACF based methods is degraded when the clean speech is corrupted by color noise.

In the frequency domain, one of the most widely accepted technique is based on the use of the cepstrum (CEP), which is proposed in [13][14]. In the CEP method, the pitch is extracted by applying the inverse Fourier transform of the log-amplitude spectrum, which is also effective. The logarithmic function involved in the CEP has the effect of shifting the vocal tract characteristics to low-quefrency parts. Utilizing high quefrency parts, we can extract the pitch without being affected by the characteristics of vocal tract.
Modified CEP (MCEP) in [15] further involves the liftering and clipping operations on the log spectrum, which is used to remove the characteristics of vocal tract as well as to eliminate the unnecessary notches of spectral valleys which correspond the noise phenomenon on the log spectrum. The MCEP also removes the high-frequency components for increasing the pitch extraction accuracy.
The ACF of the log spectrum (ACLOS) [16] utilizes the liftering and clipping operation on the log spectrum again. Then, the ACF is applied to the resulting log spectrum. The ACLOS emphasizes the periodicity of harmonics in the spectrum.

The CEP based methods are clearly expressed the harmonic structure of the speech signal under no noise conditions. However, in noisy environments, the CEP based methods does not perform well totally because the speech peaks are influenced by the noise peaks.

From the above point of view, the fourth-root (FROOT) method is introduced in [17]. The FROOT method is utilized in the amplitude spectrum instead of the logarithm spectrum for reducing the effect of noise. Therefore, the FROOT method is expected to get the clear harmonics structure by removing the unnecessary components by using the lifter and clipping operations. However, in noisy environments, the FROOT method is affected due to the noise characteristics in between the harmonics. Resulting in that, the pitch extraction accuracy of the FROOT method is severely degraded at the narrow-band noise.

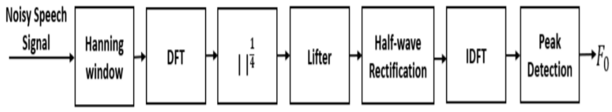


Fig. 1. Block diagram of proposed method.

A spectral harmonic technique has addressed in [18]. In this method, the bank of bandpass lifter is used to flatten the spectrum. Then, the autocorrelation function is applied to the spectrum domain for extracting the pitch periodicity by reducing the effect of vocal tract characteristics. This approach may be effective but the overall procedure is complicated to implement.

Recently, two sophisticated approaches have been addressed [19][20]. The pitch estimation filter with amplitude compression (PEFAC) [19] is a frequency domain pitch extraction method, which utilizes its sub-harmonic summation [21] in the log frequency domain. The PEFAC also attempts an amplitude compression technique for enhancing its noise robustness.
On the other hand, BaNa [20] considers the noisy speech peaks and results in a hybrid pitch extraction method that selects first five spectral peaks in the amplitude spectrum of the speech signal. BaNa calculates the ratios of the frequencies of the spectral peaks with tolerance ranges and accurately extracts the pitch of the speech signal.

In this paper, we propose a method for the narrow-band noise that utilizes the half-wave rectification (HWR) on the lifter spectrum for pitch extraction. In the narrow-band noise, the noise energy is distributed over a shorter range of frequencies. In this noise case, after lifter and clipping operations, the FROOT method fails to preserve the periodicity by the effect of the noise characteristics which is effective to extract the more accurate pitch. Therefore, the proposed method utilizing the HWR on the lifter output to present the clear harmonics and maintain the periodicity among the harmonics for increasing the extraction accuracy. Resulting in that, the proposed method emphasizes strongly the pitch peak in its waveform and simultaneously suppresses the noise components included in the noisy speech.

The remainder of this paper is organized as follows. Section 2 describes the principle of the proposed method. In Section 3, we first compare the performance of the proposed method with that of the conventional methods and then discuss the processing time. Finally, we conclude this paper in Section 4.

2. Proposed Method

Let us assume that the clean speech signal, $x(n)$, is corrupted by noise, $w(n)$. The noisy speech signal, $y(n)$, is expressed as

$$y(n) = x(n) + w(n) \qquad (1)$$

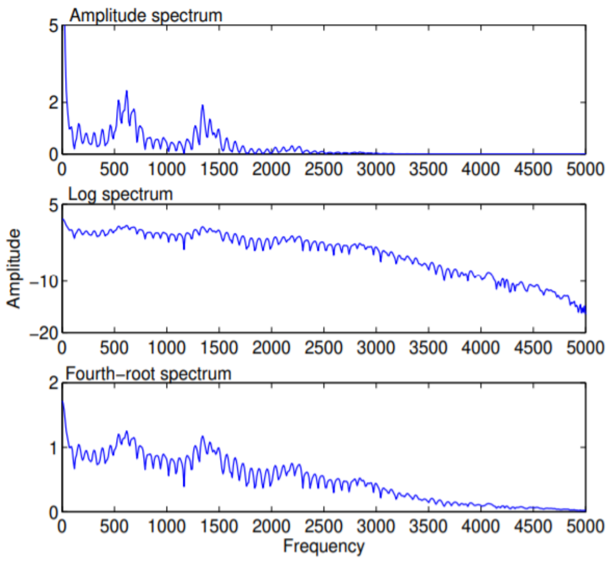


Fig. 2. Different spectral shapes of speech signal at SNR=0 [dB] (car interior noise).

Figure 1 shows a block diagram of the proposed method. In the proposed method, firstly we apply a low pass filter (LPF) to the noisy speech signal framed by a window function because the LPF can eliminate the noise characteristics for increasing the accuracy of pitch extraction. The LPF is often applied before the analysis of speech signals and filters out the high-frequency components of the noisy speech signals. We use an LPF with the telephone line cutoff frequency.

After windowing, we have considered the different spectral shapes of a speech signal as shown in Fig. 2. From figure 2, we have observed that the periodicity of the log spectrum is destroyed by the effect of the noise in the valley region. On the other hand, the fourth-root spectrum is emphasized to present the pitch harmonics in the low-frequency domain as well as to reduce the noise effect at the valley region in the high-frequency domain. Therefore, in this paper, we have selected the fourth-root spectrum for reducing the noise effect than that of the log spectrum.

However, the fourth-root spectrum sometimes faces the effect of vocal tract characteristics. To overcome this problem, the flattening procedure is more effective.

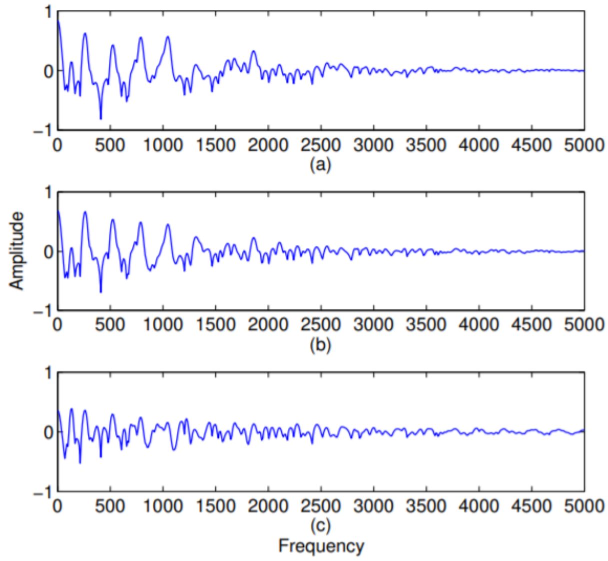


Fig. 3. waveform at different liftering output (a) using 10th lifter order (b) using 25th lifter order (c) using 40th lifter order at SNR=0 [dB] (car interior noise).

Therefore, we apply a lifter in the fourth-root spectrum. The lifter is implemented by multiplying by a lifter order in the quefrency domain and then converting back to the frequency domain resulting in a smoother signal. Basically, the vocal tract information is present at the lower part in the quefrency domain. At the higher part in the quefrency domain, the pitch information is present. Resulting in that, lifter order should be shorter to reduce the influence of the vocal tract characteristics. From figure 3, we have investigated that the lifter of 25th order preserve the high periodicity than that of the higher lifter order (ex. 40th order) at the lifter output. Therefore, the proposed method is used the short lifter of 25th order to eliminate the vocal tract information in the quefrency domain and to reduce the influence of noise. Resulting in that, the lifter is used to get the clear harmonics which maintain the periodicity for pitch extraction.

After lifter operations, we have observed that the noise components are present in between the harmonics and it can not maintain their periodicity in the noisy environments. In these cases, we have considered the harmonics from the lower frequency domain of the lifter output which is effective to extract the more accurate pitch by using the proposed method. Therefore, the HWR is applied to the lifter output which emphasizes the accurate pitch period by reducing the effect of noise and maintain the periodicity than that of the clipping operation in the noisy environments as shown in Fig. 4.

Figure 4 illustrates how to extract the pitch period by using the proposed method in the clean speech signal and the

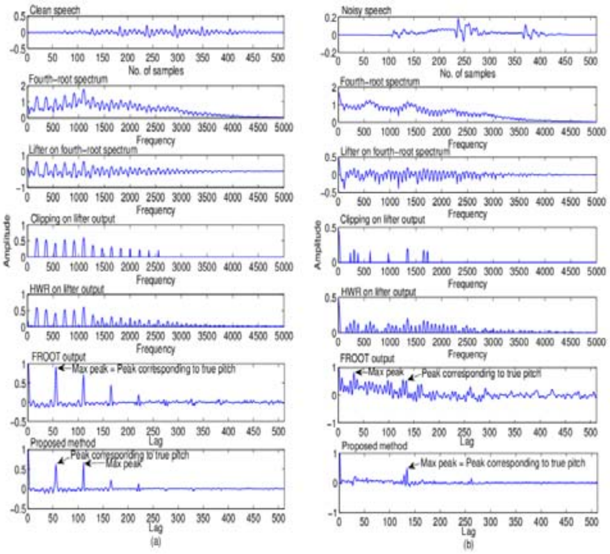


Fig. 4. Processing in step-by-step for FROOT method and the proposed method, (a) at clean speech (b) at SNR=0 [dB] (car interior noise).

noisy speech signal, respectively. We have observed that the energy level of the first three harmonics provides almost same amplitude in the lower frequency domain of the fourth-root spectrum. The pitch information also presents within these harmonics. After that, the clipping operation is applied to the lifter output where the pitch harmonics are maintained the periodicity. On the other hand, the HWR on lifter output fails to overcome the periodicity in between the harmonics in the clean speech signal. Therefore, the FROOT method extracts the more accurate pitch period in the clean speech signal as shown in Fig. 4 (a). In contrast, in the noisy speech case (such as car interior noise), we have investigated that the harmonics are not maintained their periodicity at the clipping output. Resulting in that, sometimes undesired peaks arise in the FROOT method by the affect of formant characteristics of the vocal tract.
Therefore, the HWR is applied to the lifter output to maintain the periodicity in between the harmonics. For this reason, the proposed method extracts the accurate pitch period by reducing the effect of the vocal tract characteristics as well as suppressing the noise effect in the noisy environments as shown in Fig. 4 (b).

3. Experiments

To investigate the performance of the proposed method, we conducted experiments on speech signals.

3.1. Experimental Condition

Speech signals are taken from three databases; NTT [22] and KEELE [23]. In the NTT database, which was developed by NTT Advanced Technology Corporation, the speech materials are 11 [s] long and spoken by four Japanese male and female speakers, which were sampled at a rate of 10 [kHz]. From the KEELE database, we utilize five male and five female speech signals spoken in English. The total length for the ten speakers' speeches are about 5.5 [m]. These speech signals were sampled at a rate of 16 [kHz]. To generate noisy speech signals, we added car interior noise, and military vehicle noise to the speech signals. The car interior noise, and military vehicle noise were taken from the NOISEX-92 database [24] with the sampling frequency of 19.98 [kHz]. These noises were resampled with the sampling frequency of 10 [kHz] and 16 [kHz], respectively, for using the speech data in [22]-[23]. The SNR was set to -5, 0, 5, 10, 20, ∞ [dB] and the other experimental conditions for pitch extraction were

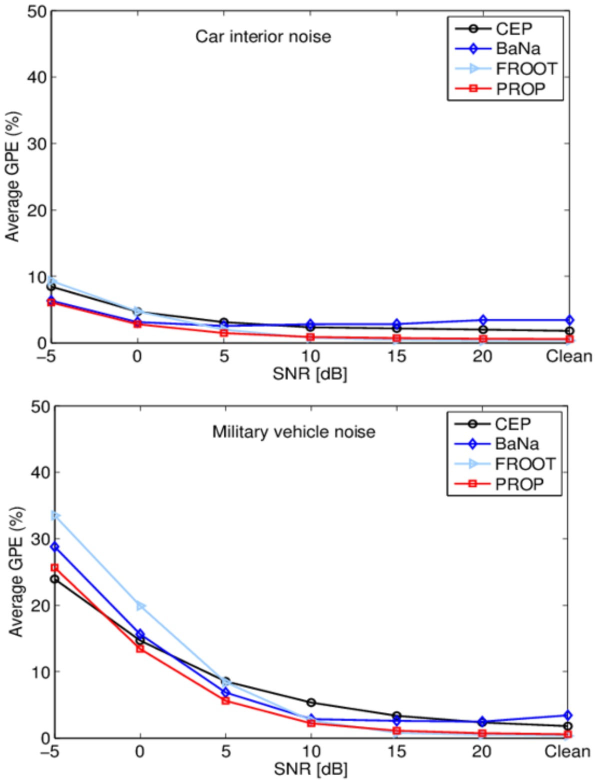
- frame length: 51.2 [ms], except for BaNa;
- frame shift: 10.0 [ms];
- window function: Hanning;
- band limitation of LPF: 3.4 [kHz];
- DFT (IDFT) length : 1024 points for the NTT database and 2048 points for the KEELE database except for BaNa.

The following pitch extraction error $e(l)$ was used for the evaluation of pitch extraction accuracy based on Rabiner's rule [2];

$$e(l) = F_{est}(l) - F_{true}(l) \tag{2}$$

where l corresponds to the frame number, $F_{est}(l)$ and $F_{true}(l)$ are the fundamental frequency extraction from the noisy speech signal, and the true fundamental frequency at the $l-th$ frame, respectively. If $|e(l)| > 10$ [%] from the ground truth fundamental frequency, we recognized the error as gross pitch error (GPE) and calculated the GPE rate (in percentage) over the total frames included in the speech data. Otherwise, we recognized the error as fine pitch error (FPE) and calculated the mean value of the absolute errors. We detected and accessed only voiced parts in sentences for pitch detection. For extracting the pitch, we used the search range of $f_{max} = 50\ [Hz]$ and $f_{min} = 400\ [Hz]$, which corresponds to the fundamental frequency range most of people have.

The ground truth information for the fundamental



frequency at each frame is included in the KEELE database , while the true fundamental frequencies at each frame in the NTT database were measured in [16] by observing the waveforms carefully, which are used here. Therefore, the $F_{true}(l)$ values in (2) are known a priori to evaluate.

Fig. 5. GPE under different SNR levels in NTT database.

3.2. Performance Comparison

The pitch extraction performance of the conventional (such as CEP [13], FROOT [17], and BaNa [20]) and proposed methods was investigated in noisy environments. Here, we consider the car interior noise and the military vehicle noise, which are the narrow-band noise. These are addressed in Fig. 9.

All parameters of the conventional methods are the same as those in the proposed method, except for the frame length and the DFT(IDFT) points for BaNa. Specifically, for BaNa the frame length was set as 60 [ms] and the DFT (IDFT) points were 2^{16} according to the suggestion in [20]. The source code to implement BaNa was collected from [26].

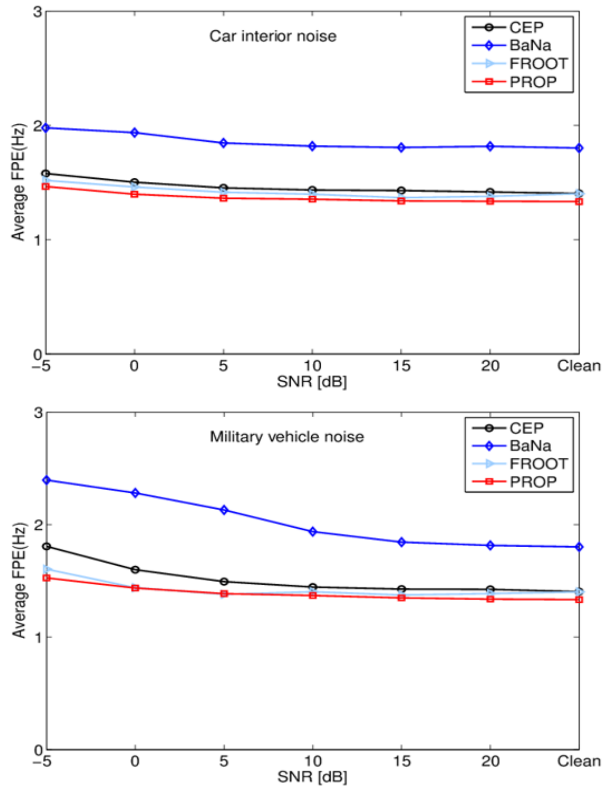


Figure 5 shows the average GPE rate in the NTT database with narrow-band noise. When the SNR is changed from -5 [dB] to ∞ [dB] (clean speech case), each plot has been obtained under each SNR condition. From Fig. 5, it is observed that the average GPE rate of the proposed method is better than that of the conventional methods for car interior noise at low SNRs. At high SNRs (>10 [dB]), the proposed method and the FROOT method is competitive with each other.

Fig. 6. FPE under different SNR levels in NTT database.

On the other hand, in the case of military vehicle noise, the proposed method provides better GPE rate at almost all SNRs levels than that of the other conventional methods except CEP at low SNR (-5 [dB]). At the low SNR (-5 [dB]) in the military vehicle noise case, the CEP provides slightly better GPE rate than that of the other methods. This could be due to a strong periodical nature of the two noises. See Fig. 9, we can observe that the car interior noise produces a sharp narrow band peak and the military vehicle noise also shows narrow band peaks, respectively. These give a strong periodical nature of the noise at low SNRs. The CEP is known to behave robustly against periodical noises.
In the clean speech case, BaNa shows the worse average GPE rate than that of the other methods due to the selection of spectral peaks.

However, it should noted here that the computational complexity of BaNa is extremely high.

Figure 6 shows the average FPE in the NTT database. The FPE represents the variations in detecting the pitch. Average FPE range for all methods is approximately from 1.5 [Hz] to 5.2 [Hz]. From Fig. 6, we see that the FPE of the proposed method is comparatively better than the other methods.

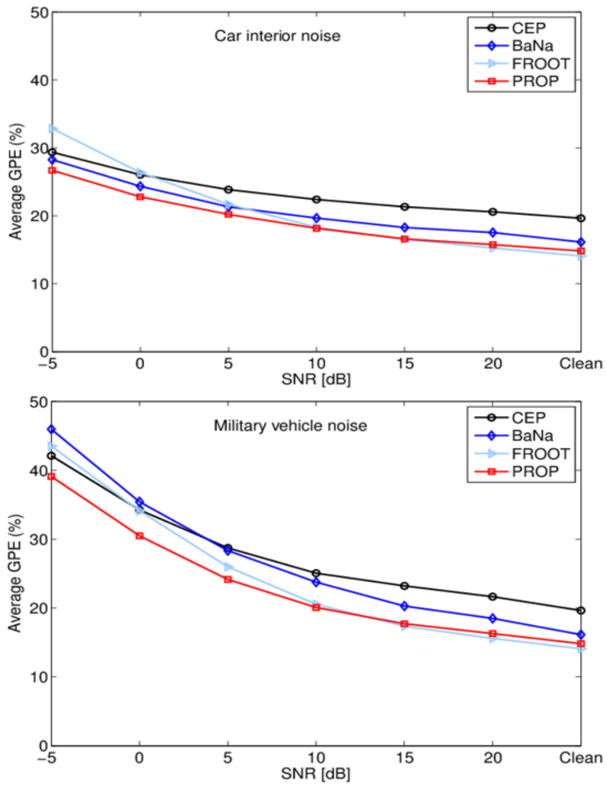


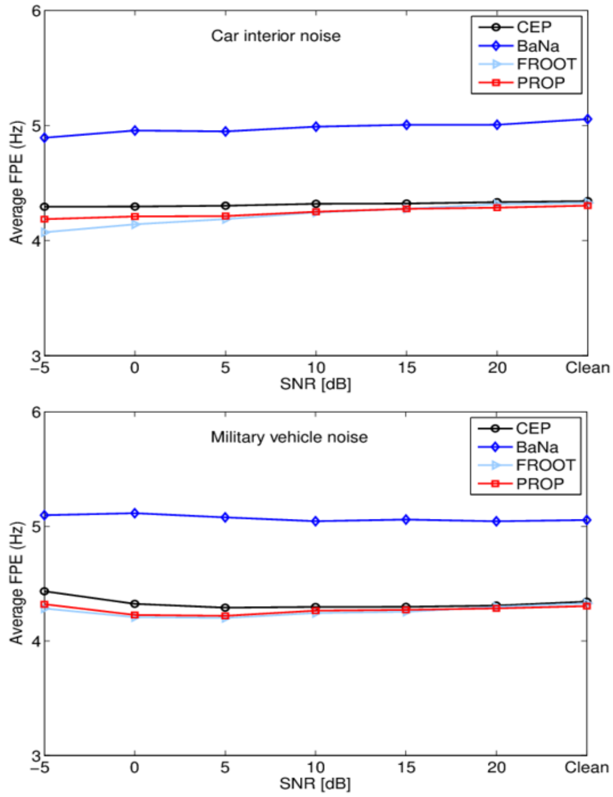
Fig. 7. GPE under different SNR levels in KEELE database.

To validate the proposed method, we have employed additionally the KEELE database. Fig. 7 shows the average GPE rate in the KEELE database with the narrow-band noise. The KEELE database provides the ground truth values of the pitch which are obtained from laryngograph signals. We analyzed the ground truth values of the pitch and found that some discontinuities are present. Therefore, the ground truth of the pitch is not so accurate. This appears in the resulting GPE rates. From Fig. 7, it is obvious that the GPE rates of the clean speech are higher than those of the clean speech in Fig. 5. This is reflected from the accuracy of the ground truth of the fundamental frequency in the KEELE database.

Fig. 7 indicates that the proposed method provides the smallest GPE rate at narrow-band noise except at high SNRs such as 20 [dB] and ∞ [dB]. At high SNRs, the proposed method is competitive to the FROOT method because FROOT method maintain their periodicity by using the clipping operation.

The average FPE performance in the KEELE database as shown in Fig. 8. From Fig. 8, it is obvious that the proposed

Fig. 8. FPE under different SNR levels in KEELE database.



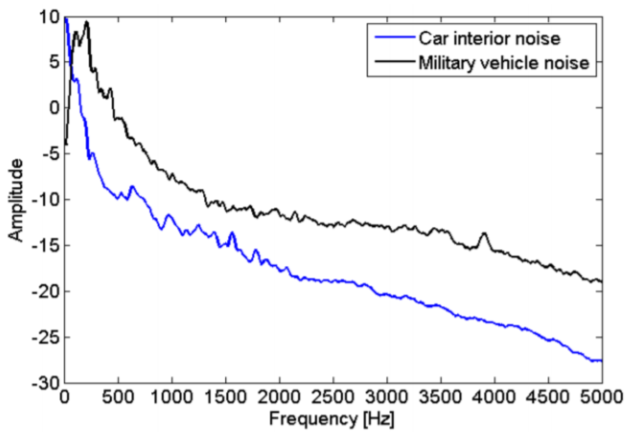


Fig. 9. Long term average spectra of noises.

method and the FROOT method are competitive with each other which provides the lower error rates than the other methods in the narrow-band noise.

3.4. Processing Time

In Table 1, we have compared the processing time per one-second data for each method in the NTT database. We have tested all methods on the PC with Intel (R) Core(TM) i5-6400K, 4 [GHz] clock speed of CPU and 8 [Gigabytes] of memory. For the evaluation, we have used five trials for each method, then calculated the average processing time to obtain reliable measurements. From Table 1, we notice that the processing time of the proposed method and the FROOT method are almost similar, which are commonly short than that of the other methods because the HWR and clipping operations are directly applied to the lifter spectrum. The computational time of the CEP method is shorter because it can consider the logarithm spectrum to identify the pitch candidates. On the other hand, BaNa faces the longest processing time because of the large FFT size, which is used for keeping high frequency resolution.

Table 1: Processing time per second of speech

CEP	**BaNa**	**FROOT**	**PROP**
0.232	29.427	0.146	0.176

4. Conclusion

In this paper, a new noise robust method has been presented to deal with the problem of pitch extraction from noise corrupted speech signals. In noisy environments, the proposed method has been derived by using HWR on the lifter spectrum. The proposed method behaves so as to reduce the effect of vocal tract characteristics as well as suppressing the non-pitch peaks, resulting in enhancing the pitch peak in the narrow-band noise. Through experiments, we have confirmed that the proposed method is an efficient and effective method to extract the pitch in the narrow-band noise.

References

[1] W. Hess, "Pitch determination of speech signals," Springer-Verlag, 1983.

[2] L. R. Rabiner, M. J. Cheng, A. E. Rosenberg and C. A. McGonegal, "A comparative performance study of several pitch detection algorithms," IEEE Trans. Acoust. Speech, and Signal Processing, vol. ASSP-24, no. 5, pp.339-417, Oct. 1976.

[3] L. R. Rabiner, "On the use of autocorrelation analysis for pitch detection," IEEE Trans. Acoust. Speech Signal Processing, vol. ASSP-25, no.1, pp. 24-33, Feb. 1977.

[4] A. Cheveigne and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," J. Acoust. Soc. Am., vol. 111, no. 4, pp. 1917-1930, 2002.

[5] M.J.Ross, "Average magnitude difference function pitch extractor," IEEE Trans, Acoust., Speech Signal Processing, vol. 22, no. 5, pp. 353-362, Oct. 1974.

[6] C. K. Un and S. Yang, "A pitch extraction algorithm based on LPC inverse filtering and AMDF," IEEE Trans. Acoust. Speech Signal Processing, vol. ASSP-25, no.6, pp. 565-572, Feb. 1977.

[7] J. W. Xu and J. C. Principle, "A pitch detector based on a generalized correlation function," IEEE Trans. on Audio, Speech and language Processing, vol. 16, no. 8, pp. 1420-1432, Nov. 2008.

[8] A. Moreno and J. A. R. Fonollosa, "Pitch determination of noisy speech using higher order statistics," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 1992.

[9] C. Shahnaz, W. Zhu and M. O. Ahmad, "Pitch estimation based on a harmonic sinusoidal autocorrelation model and a time-domain matching scheme," IEEE Trans. Audio, Speech and Language Processing, vol. 20, no. 1, pp. 322-335, Jan. 2012.

[10] S. Lin, "Robust pitch estimation and tracking for speakers based on subband encoding and the generalized labeled multi-Bernoulli filter," IEEE Trans. on Speech and Language Processing, vol. 27, no. 4, pp. 827-841, 2019.

[11] D. Talkin, "A robust algorithm for pitch tracking (RAPT)," in speech coding and synthesis, W. B. Kleijn and K. K. Paliwal, Eds. Amsterdam: Elsevier pp. 496-518, 1995.

[12] T. Shimamura and H. Kobayashi, "Weighted autocorrelation for pitch extraction of noisy speech," IEEE Trans. on Speech and Audio Processing, vol. 9, no. 7, pp. 727-730, 2001.

[13] A. M. Noll, "Cepstrum pitch determination," Journal on Acoust. Soc. Am., vol. 41, no. 2, pp. 293-309, 1967.

[14] S. Ahmadi and A. S. Spanias, "Cepstrum-based pitch detection using a new statistical V/UV classication algorithm," IEEE Trans. on Speech and Audio Processing, vol. 7, no. 3, pp. 333-338, 1999.

[15] H. Kobayashi and T. Shimamura, "A modified cepstrum method for pitch extraction," Proc. on IEEE Asia-Pacific International Conference on Circuits and Systems Microelectronics and Integrating Systems (APCCAS), 1998.

[16] N. Kunieda, T. Shimamura and J. Suzuki, "Pitch extraction by using autocorrelation function on the log spectrum," Electroni. Communi. Jpn, Part 3, vol. 83, no. 1, pp. 90-98, 2000.

[17] S. Motegi and T. Shimamura, "Fundamental frequency extraction in narrow band noise," Proc. of the Acoustical Society of Japan, pp. 277- 278, 2011.

[18] M. Lahat and R. J. Niederjohn, and D. A. Krubsack, "A spectral autocorrelation method for measurement of the fundamental frequency of noise-corrupted speech," IEEE Trans. on Acoust. Speech, and Signal Processing, vol. ASSP. 35, no. 6, pp. 741-750, 1987.

[19] S. Gonzalez and M. Brookes, "PEFAC - A pitch estimation algorithm robust to high levels of noise," IEEE/ACM Trans. on Audio, Speech, and Language Processing, vol. 22, no.2, pp. 518-530, Feb. 2014.

[20] N. Yang, H. Ba, W. Cai, I. Demirkol and W. Heinzelman, "BaNa : A noise resilient fundamental frequency detection

algorithm for speech and music," IEEE/ACM Trans. Audio, Speech, and Language Processing, vol. 22, no. 12, 2014.
[21] D. J. Hermes, "Measurement of pitch by subharmonic summation," Journal on Accoust. Soc. Am., vol. 83, no. 1, pp. 257-264, 1988.
[22] "20 Countries Language Database," NTT Advanced Technology Corp., Japan, 1988.
[23] F. Plante, G. Meyer, and W. Ainsworth, "A fundamental frequency extraction reference database," in Proc. Eurospeech, pp. 837-840, 1995.
[24] P. C. Bagshaw, S. M. Hiller, and M. A. Jack, "Enhanced pitch tracking and the processing of F0 contours for computer aided intonation teaching," in Proc. Eurospeech, pp. 1003–1006, 1993.
[25] A. Varga and H. J Steeneken, "Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems," Speech Communication, vol. 12, no. 3, pp. 247-251, 1993.
[26] Wcng, wireless communication networking group, [Online]. Available: http://www.ece.rochester.edu/projects/wcng/code.html

Md. Saifur Rahman received his B.Sc.(Honours) and M.Sc. degrees in information and communication engineering from University of Rajshahi, Rajshahi, Bangladesh, in 2006 and 2007, respectively. In 2012, he joined as a Lecturer in the Department of Information and Communication Technology, Comilla University, Comilla, Bangladesh, where he is currently serving as an Assistant Professor. He is received his PhD degree in 2020 from the Graduate School of Science and Engineering, Saitama University, Saitama, Japan. His research interests include speech analysis and digital signal processing.

Miss. Nargis Parvin received her B.Sc.(Honours) and M.Sc. degrees in information and communication engineering from University of Rajshahi, Rajshahi, Bangladesh, in 2006 and 2007, respectively. In 2015, she joined as a Lecturer in the Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Cumilla Cantonment, Cumilla, Bangladesh, where she is currently serving as an Assistant Professor. She is received her D. Engg. degree in 2019 from the Graduate School of Science and Engineering, Saitama University, Saitama, Japan. Her research interests include adaptive signal processing, blind equalization, distributed signal estimation and its applications.