Phase Differences Based Coherence Filter for Dual-Microphone System

Quan Trong The

Abstract*: This paper proposes a dual-microphone coherence filter, which based on phase difference of two noisy signals. This method passes, without dirtortion target speaker. In real noisy environment, it's difficult to get exactly information of condition noise field, especially in complex environment. Therefore a post-filtering, which is a function depends on speech presence probability, is suitable for compact dual-microphone and is adapted in the author's approach. The performance of suggested algorithm proved the stability and efficiency of the algorithm. This approach allows increasing performance of speech enhancement in term of signal-to-noise ratio and the amount of noise reduction. The author intends to continue using phase differences as conditions for different algorithms speech quality improvement under various noise conditions.

Keywords*: phase differences, coherence function, microphone array, dual-microphone, filter, post-filtering, speech presence probability.

I. INTRODUCTION

Environment noises are always major disturbing sources to sound capturing equipment, especially in complex condition of noise filed. Almost single – channel, which based on spectral domain, can solve problem of separating the target sound source, causes missing a priori information of geometry situation of recordings. Microphone arrays [1-3], which is a modern of processing signal technique, allow exploiting the spatial information to pre-processing, or reduce non-stationary, interference noise. Dual-microphone, which is one of the most simple type microphone array, used very popular due to their easily implementation, low cost. Coherence function [4-6] based speech enhancement algorithms have been developed for many years. A knowledge of target speech is very necessary for almost speech acquisition system. A accuracy speech presence probability (SPP) estimator obviously effects on performance of dual-microphone algorithms. In this paper, the author proposed a new coherence based algorithm, which use an estimation of speech presence probability (SPP) [7-8] to obtain target speaker, while ensuring suppress diffuse noise field. The results provide that, the suggested algorithm is suitable for diffuse noise conditions.

Revised Manuscript Received on January 15, 2020
***** Correspondence Author
Quan Trong The*, Information Technologies and Programming Faculty, ITMO University, Saint-Petersburg, Russian-Federation. Email: quantrongthe1984@gmail.com.

II. COHERENCE FUNCTION

An overview of coherence function, which based on dual-microphone system is studied in this section. At first, two noisy signal were sampled and transformed into Fourier Transform. In frequency domain, an adaptive beamforming, which exploits the spatial information (the direction of arrival signal (DOA), a priori of knowledge background noise, coherence of noise field) or digital signal processing applied to enhanced noisy signal. Finally the output signal obtained by inverse Fourier transform and the overlap-and-add (OLA) algorithm.

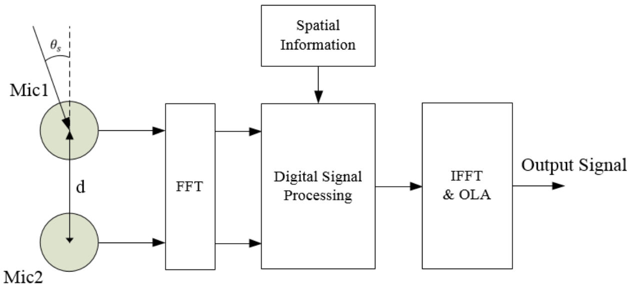

Fig. 1. Digital signal processing of dual-microphone system.

In noisy environment, dual-microphone receive speech signal, that can be interfered by background noise. The two noisy signals $x_1(t), x_2(t)$ at the microphones modeled as:

$$x_1(t) = s_1(t) + n_1(t) \tag{1}$$

$$x_2(t) = s_2(t) + n_2(t) \tag{2}$$

where $s_1(t), s_2(t)$ is the speech signal at microphones, and $n_1(t), n_2(t)$ are additive noises. The representation in the short-time Fourier transform (STFT) domain as:

$$X_1(f,k) = S_1(f,k) + N_1(f,k) \tag{3}$$

$$X_2(f,k) = S_2(f,k) + N_2(f,k) \tag{4}$$

Where $X_1(f,k), X_2(f,k)$ is the STFT of the $x_1(t), x_2(t)$, $S_1(f,k), S_2(f,k), N_1(f,k), N_2(f,k)$ are the STFT of $s_1(t), s_2(t), n_1(t), n_2(t)$respectively for the k −th frame and f −th frequency bin.
The magnitude of the coherence function of two signals $X_1(f,k)$ and $X_2(f,k)$ can be computed as follow and is expressed as follow:

$$\Gamma_{X_1X_2}(f,k) = \frac{P_{X_1X_2}(f,k)}{\sqrt{P_{X_1X_1}(f,k)P_{2X_2}(f,k)}} \quad (5)$$

In the equation (5), $P_{X_1X_1}(f,k)$, $P_{X_2X_2}(f,k)$ and $P_{X_1X_2}(f,k)$ are the Power Spectrum Density (PSD) of $X_1(f,k)$, $X_2(f,k)$ and Cross Power Spectrum Density (CPSD) of $X_1(f,k)$ and $X_2(f,k)$,respectively. The CPSD and PSD can be estimated recursively as:

$$P_{X_iX_j}(f,k) = \gamma P_{X_iX_j}(f,k-1) + (1-\gamma)X_i(f,k)X_j^*(f,k) \quad (6)$$

With $i,j \in \{1,2\}$. Where γ is a smoothing factor in the range [0,1].

Yousefian N and Loizou P. C [6] proposed a coherence filter in condition of coherent noise. The filter, which uses an estimation of signal-to-noise ratio for calculating the enhanced signal, can be presented as:

$$G(f,k) = 1 - \left|\Gamma_{X_1X_2}(f,k)\right|^{F(f,k)} \quad (7)$$

where function $F(f,k)$ depends on ratio signal-to-noise.

III. THE PROPSED ALGORITHM

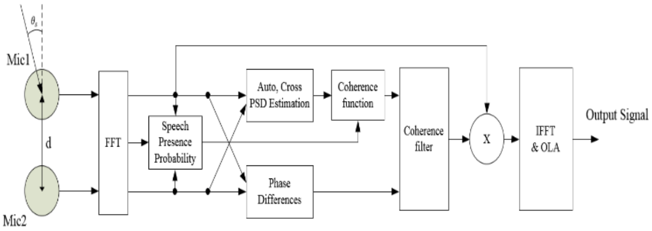

Fig. 2. The proposal algorithm.

One of the characteristic factors that can be exploited and utilized in dual-microphone system is the phase difference, which changes rapidly corresponding with the presence or absence speech component. The phase difference $\Delta arg_X(f,k)$ is defined as:

$$\Delta arg_X(f,k) = arg(X_1(f,k)) - arg(X_2(f,k)) \quad (8)$$

Due to the different sensitivity of microphones, or failure of sampling; which affected on the magnitude of the captured signal, caused attenuation. The attenuation can be considered as a function of phase difference as:

$$\Delta\Phi_X(f,k) = 1 + tan\left(\frac{\Delta arg_X(f,k)}{2}\right)^{-2} \quad (9)$$

The author proposed an exploiting of phase difference, speech presence probability to decrease non-stationary, stationary noise and save speech component by using coherence filter. The coherence function is estimated recursively respect to speech presence probability in (10):

$\Gamma_{X_1X_2}(f,k) = SPP(f,k)\Gamma_{X_1X_2}(f,k-1) + (1 -$ $SPPf,kPX1X2f,kPX1X1f,kPX2X2f,k$ (10)

And suggested algorithm can be presented with following equation:

$$G(f,k) = 1 - \left|\Gamma_{X_1X_2}(f,k)\right|^{\Delta\Phi_X(f,k)} \quad (10)$$

The output signal obtained by:

$$Y(f,k) = G(f,k)X_1(f,k) \quad (11)$$

IV. EXPERIMENTS AND RESULTS

In this section, the author apply the suggested method to the speech enhancement problem and evaluate its performance in a anechoic chamber; dual microphone was placed on a table at the center of chamber. All two signals were sampled with frequency 16 kHz. The two noisy signal segmented with 512 samples, overlap is 50%. For short-time Fourier transform (STFT) a Hamming window used. Smoothing parameter γ was set to 0.5. The purpose of the experiment was to test the proposed algorithm on real signals. The scheme of the experiment shown in Fig. 3. The objective measure NIST STNR, WADA SNR [9] used to estimate the quality of original and processed signal by suggested algorithm.

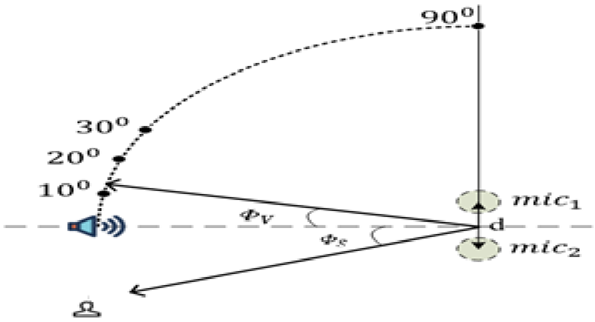

Fig. 3. The scheme of experiments

The distance between the microphones $d = 5cm$, and direction of arrival of target speaker is $\Phi_s = -30^0$.

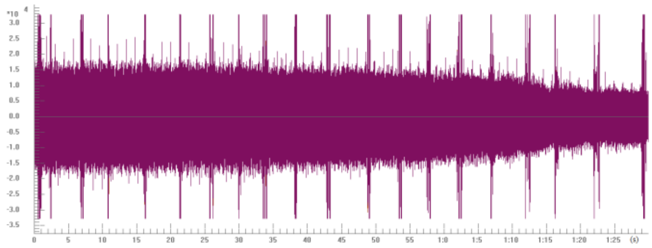

Fig. 4. Amplitude of original signal.

Due to objective reasons, the speaker didn't stand still, and move his head during the experiment; microphone mismatch, the sensitivity of the microphones is different, etc. So the most common speech enhancement algorithms encounter is the speech distortion at the output signal. Estimation of coherence function between two noisy signal is an optimal solution to overcome the above problem.

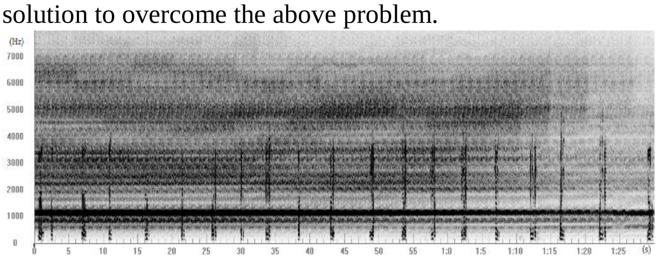

Fig. 5. Spectrogram of original signal.

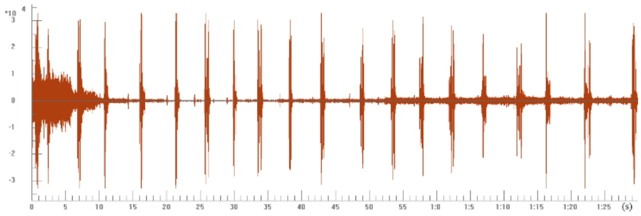

Fig. 6. Amplitude of processed signal.

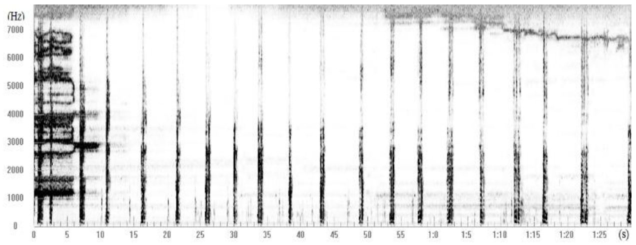

Fig. 7. Spectrogram of processed signal.

Figure 4, 5, 6, 7 demonstrate that the suggested algorithm, which is a combination of coherence function and speech presence probability, can save target speaker while reducing background coherent noise. The ratio signal-to-noise SNR is improved to 28.5 ÷ 43 (dB) as in the Table I. The characteristic of phase difference proves the effectiveness of speech enhancement. At the speech frames, $\Delta arg_X(f,k)$ tends to 0, $\Delta\Phi_X(f,k) \approx \infty$, due to $\Gamma_{X_1X_2}(f,k) < 1$, $H(f,k) \approx 1$, so the speech component can be remain. In contrast, $\Delta arg_X(f,k)$ tends to $\frac{\pi}{2}$ or $-\frac{\pi}{2}$, $\Delta\Phi_X(f,k) \approx 1$, so $1 - \left|\Gamma_{X_1X_2}(f,k)\right|^{\Delta\Phi_X(f,k)} \approx 0$, noise component suppressed at the noise frames.

Table-I: The signal-to-noise ratio SNR (dB)

Estimation method	**Signal**	
	Original	***The proposed algorithm***
NIST STNR	4.0	32.5
WADA SNR	-0.1	42.9

V. CONCLUSION

In this paper, the author introduced an useful, effective algorithm, which based phase differences, can be incorporated in multi-microphone system to improve the performance and robustness. The results provides the author's approach to use and implement one of observed signal's feature in dual-microphone system. Furthermore, the suggested algorithm can combine with a post-filtering to decrease background noise.

REFERENCES

1. Brandstein M. and Ward D. (Eds.). Microphone Arrays: Signal Processing Techniques and Applications, Springer, 2001.
2. Benesty J. and Chen J. Study and Design of Differential Microphone Arrays, Springer, 2013.
3. Benesty J., Chen J., Pan C. Fundamentals of Differential Microphone Arrays, Springer, 2016.
4. Yousefian N., Loizou P. C. A Dual-Microphone Algorithm That Can Cope With Competing-Talker Scenarios. IEEE Transactions on Audio, Speech, and Language Processing, Vol. 21, No. 1, January 2013.
5. Yousefian N., Kokkinakis K., Loizou P.C. A coherence-based algorithm for noise reduction in dual-microphone applications. Proc. Eur. Signal Process. Conf. (EUSIPCO10), Alborg, Denmark, Aug. 2010, pp. 1904–1908.
6. Yousefian N., Loizou P. C. A dual-microphone speech enhancement algorithm based on the coherence function. IEEE Trans. Speech Audio Process., vol. 20, no. 2, pp. 599–609, Feb. 2012.
7. Gerkmann T. Unbiased MMSE-Based Noise Power Estimation with Low Complexity and Low Tracking Delay, IEEE TASL, 2012.
8. Gerkmann T., Hendriks R.Noise Power Estimation Based on the Probability of Speech Presence, WASPAA 2011.
9. The objective measure: Available: https://labrosa.ee.columbia.edu/projects/snreval/.

AUTHORS PROFILE

Quan Trong The, is a graduate student within the Computer Science program at Le Quy Don Technical University in 2011. He also graduated with an MS in Information Systems at the Posts and Telecommunications Institute of Technology in 2015. Now he is Ph.D student at National Research ITMO University, Russian Federation. His research interests include Speech Enhancement, Microphone Array.