International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Improved the Efficiency of Generalized Sidelobe Canceller Algorithm by using Speech Presence Probability

Quan Trong The

Abstract*: Speech signal processing application always encounter certain difficulties in real complex environment. The captured signal on microphones often interfered by coherent, incoherent, stationary, non-stationary noise and self acoustic mismatch. To solve this problem, the necessary requirement is speech enhancement to extract target speaker from observed signals in condition minimum speech distortion, while removing background noise. The author proposed a speech enhancement generalized sidelobe canceller based on an estimation of speech presence probability. Main ideal of the algorithm is accuracy estimation of auto and cross power spectral densities of main and reference signal, which used in process of filtering. The experimental result ensures the effectiveness of the proposal algorithm, the background noise is suppressed while the quality of speech is improved in compared with the conventional generalized sidelobe canceller. The proposed algorithm can be evaluated as a frontend for automatic speech application.

Keywords*: microphone array, dual-microphone, generalized sidelobe canceller, noise reduction, speech presence probability, Wiener filter.

I. INTRODUCTION

Mainly speech signal receiving systems work in noisy environments, where the target speaker or useful signal desired speech signal is corrupted by interfering signals, nonstationary, stationary or diffuse noise signal. The problem is need to suppress background noise and extract desired signal of interest. Speech enhancement ensures the intelligibility of the output signal is kept or improved after processing noisy signal, which degraded by additive noise. For solving that problem, single-channel approach extensively studied, but results is not always achievable due to musical noise self-algorithm; especially in such application: communications, teleconferencing, hands-free mobile. The limitation of single-channel speech enhancement is use only spectral information. To overcome this disadvantage; multi-microphone [1-3] has the ability of exploiting the spatial information: direction of arrival (DOA), a priori information of coherence noise field, coherence between noisy signals.

Revised Manuscript Received on January 30, 2020.
***** Correspondence Author
Quan Trong The*, Department of Information Technologies and Programming Faculty, University ITMO, Saint-Petersburg, Russian Federation. Email: quantrongthe1984@gmail.com

The generalized sidelobe canceller (GSC) has been an effective solution in many applications using two or more microphones to enhance speech quality in noisy environments [4-6]. The basic idea of this approach is forming an additional signal toward to direction of desired speaker and a reference signal. This allows for using adaptive processing signal Wiener filter to separate target and remove noise signal. In this paper, the author proposes a simple and efficient method, which based an estimation of speech presence probability [7-8], to control the updating of auto, cross power spectral densities in GSC system. The proposed algorithm is performed in various condition: in an anechoic chamber, in diffuse noise field. The objective measure estimated proved that capability of suggested algorithm.

II. GENERALIZED SIDELOBE CANCELLER SYSTEM

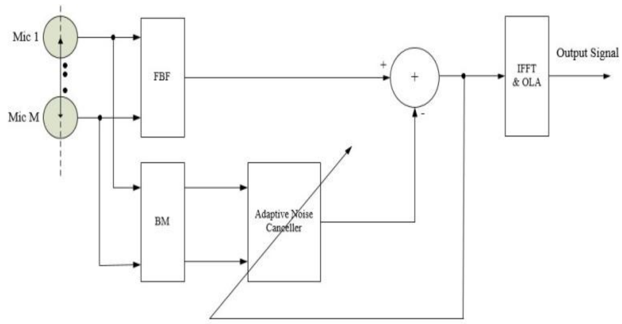


Fig. 1. The scheme of SGC Structure.

As shown in Fig. 1, the GSC process is composed of three parts: a fixed beamformer (FBF) forms a beam in the look direction so that a desired speech signal is passed and all other signals are attenuated. A blocking matrix (BM) is applied to the input microphone signal to compute an estimate for a noise reference signal by blocking the components of a desired speech signal. The noise references at its output drive an adaptive noise canceller (ANC) whose coefficients are adapted to suppress the noise in the FBF output.

In real speech application, BM requires more correctly information of direction of arrival of target speaker for remove speech in lower branch. Estimation errors in the direction of arrival and reflections of signals by objects and walls cause leakage of the desired speech signal into noise references, resulting in signal cancellation in the beamformer output.

III. THE PROPOSED ALGORITHM

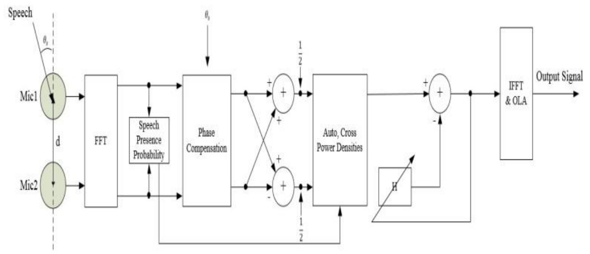


Fig. 2. The scheme of the proposed algorithm.

In this section, the author proposed a new method for estimating the coherence of noise with an estimation of speech presence probability [7-8]. Let $s(t)$ denotes desired target speech signal, $n_1(t)$ and $n_2(t)$ denote two-channel noise signals, which captured at the input two spatially separated microphones.

In the time domain, observed noisy signal at the m-th microphone can be modelled as:

$x_1(t) = s_1(t) + n_1(t)$ (1)

$x_2(t) = s_2(t) + n_2(t)$ (2)

Where $s_1(t)$, $s_2(t)$ is the speech signal at microphones, and $n_1(t)$, $n_2(t)$ are additive noises. Using short-time Fourier transform (STFT), we have got:

$X_1(f,k) = S_1(f,k)e^{j\Phi_s} + N_1(f,k)$ (3)

$X_2(f,k) = S_2(f,k)e^{-j\Phi_s} + N_2(f,k)$ (4)

Where $X_1(f,k), X_2(f,k)$ is the STFT of the $x_1(t), x_2(t)$ for the $k-th$ frame and the $f-th$ frequency bin, and $S(f,k), N_1(f,k), N_2(f,k)$are the STFT of $s(t), n_1(t), n_2(t)$ respectively. The direction of arrival of the plane wave target signal $S(\omega,k)$ is characterized by the angle θ_s relative to the axis passing through the microphones, and $\Phi_s = \pi f \tau_0 cos(\theta_s)$, $\omega = \pi f$ is the angular frequency,$\tau_0 = d/c$ is the sound delay between the microphones, d is the distance between the microphones, c is the sound speed (340*m/s*).

After step "Phase Compensation", the main signal $Y_s(f,k)$, and reference signal $Y_r(f,k)$ as follows:

$$Y_s(f,k) = \frac{X_1(f,k)e^{-j\Phi_s}+X_2(f,k)e^{j\Phi_s}}{2} \quad (5)$$

And

$$Y_r(f,k) = \frac{X_1(f,k)e^{-j\Phi_s}-X_2(f,k)e^{j\Phi_s}}{2} \quad (6)$$

$P_{Y_rY_r}(f,k), P_{Y_sY_r}(f,k)$ are the Power Spectrum Density (PSD) of $Y_r(f,k)$ and Cross Power Spectrum Density (CPSD) of $Y_s(f,k)$ and $Y_r(f,k)$, respectively. The PSD can be estimated as:

$$P_{Y_rY_r}(f,k) = \alpha P_{Y_rY_r}(f,k-1) + (1-\alpha)Y_r(f,k)Y_r^*(f,k) \quad (7)$$

$$P_{Y_sY_r}(f,k) = \alpha P_{Y_sY_r}(f,k-1) + (1-\alpha)Y_s(f,k)Y_r^*(f,k) \quad (8)$$

Where α is a smoothing factor in the range $\{0 \div 1\}$.

Formulation of Wiener filter for ANC is calculated as follows:

$$H(f,k) = \frac{P_{Y_sY_r}(f,k)}{P_{Y_rY_r}(f,k)} \quad (9)$$

In the real environment, the target speaker may not stay precisely; furthermore, the position and frequency response of the microphones may not be as precise as expected, leading to imperfect cancellation of the desired speech. For avoiding the adversities mentioned above, it is necessary to update the coherence noise according a recursive formulation of SPP. The author proposed a new method, which uses SPP to estimate auto and cross power densities as:

$$P_{Y_rY_r}(f,k) = SPP(f,k)P_{Y_rY_r}(f,k-1) + (1- SPPf,kYrf,kYr*f,k \quad (10)$$

$$P_{Y_sY_r}(f,k) = SPP(f,k)P_{Y_sY_r}(f,k-1) + (1- SPPf,kYsf,kYr*f,k \quad (11)$$

From (10) - (11), auto and cross PSD are combined from the previously value and current estimation. This procedure ensures updating and accurate estimation ANC filter.

IV. EXPERIMENTS AND RESULTS

The sampling rate was set 16 kHz for all two received noisy signals. The author use a Hamming window, number samples for Short-Fast Fourier Transform is 512 point; and level of overlap is 50%. The objective measures of speech quality NIST SNR, WADA SNR [9] were used to estimate and evaluate the improvement of suggested algorithm (GSC-SPP) compared to conventional GSC algorithm (GSC-CONV). The purpose of the experiments was to prove the ability of filtering noisy signal in coherent and diffuse noise using the suggested GSC algorithm, and compared to others dual-microphone speech processing algorithms.

A. Experiment in anechoic chamber

The purpose of the experiment was to test the GSC algorithm on real signals and to assess the possibilities of adaptation to changes in the direction of noise arrival. The scheme of the experiment is shown in Fig. 3. Dual microphone was placed on a table at the center of chamber.

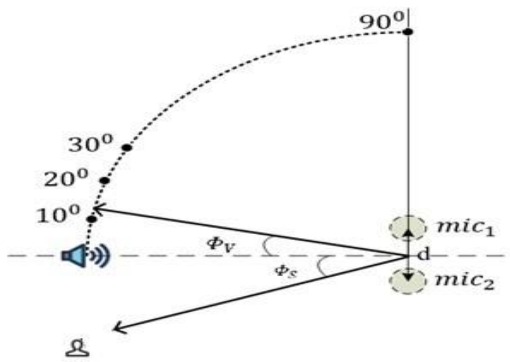
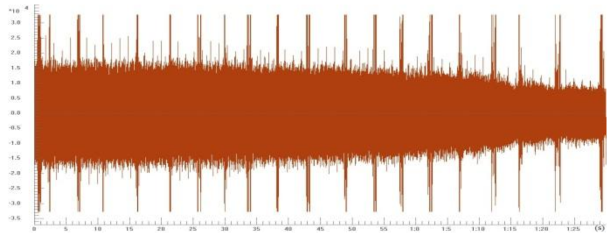


Fig. 3. The scheme of experiments

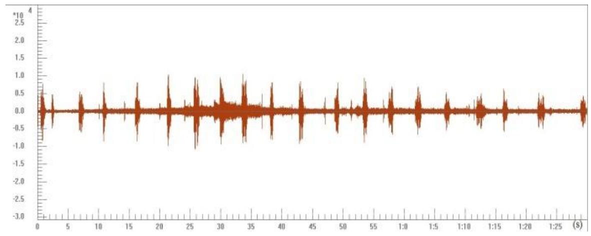
The target direction was set in the direction of the speaker ($\Phi_s = -30^0$), the distance between the microphones $d = 5cm$.

Table i. Snr estimation (db)

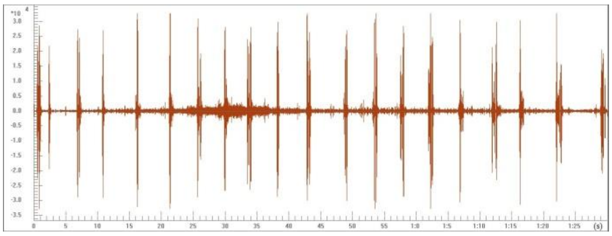
Estimation Method	Original signal	GSC-CONV	GSC-SPP
NIST SNR	4	16.8	38.3
WADA SNR	-0.1	13	43.6



(*a*)



(*b*)



(*c*)

Fig. 4. Amplitude of the (*a*) original signal; (*b*) processed signal by conventional GSC algorithm; (*c*) GSC-SPP algorithm.

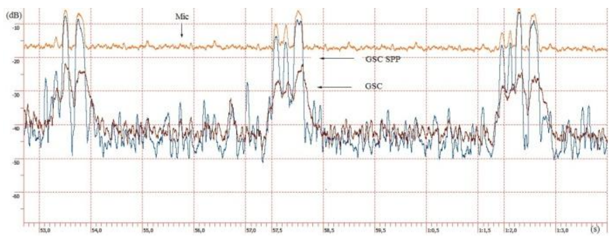
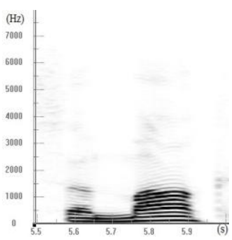


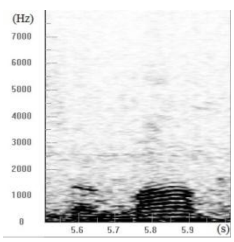
Fig. 5. RMS of original signal and processed signals by GSC SPP, conventional GSC algorithm

As we can see from Fig 4, Fig 5; the proposed GSC-SPP algorithm can save target speech and suppress efficiently background noise. The suggested algorithm GSC SPP can save sufficient target speaker while suppressing background noise. The noise reduction is about 15-30 (dB). When compared with conventional GSC (GSC-CONV), GSC-SPP have shown the advantage of reducing the speech distortion up to 15dB; and the quality of speech is increased about 21.5 ÷ 30 (dB) compared to GSC-CONV. The results of performance shown in Table 1.

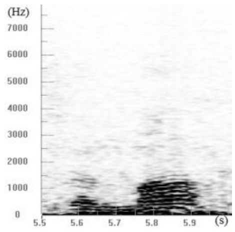
B. Experiments With Mixtures Of Diffuse Noise



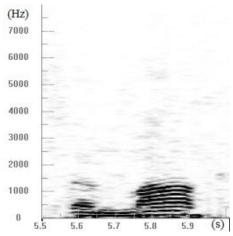
(*a*)



(*b*)



(*c*)



(*d*)

Fig. 6. Fragment of original, mixture and processed signals by GSC-CONV, GSC-SPP

The author used the two-channel mixtures of speech and real-world background in SiSEC 2010 noisy speech dataset [10]. Background noise signals were recorded via a pair of omnidirectional microphones spaced by 8.6 cm in public environments. As from Figure 6, we can realize the improvement in comparison of spectrogram between mixture and processed signal. The noise components were obviously removed, and objective measure SNR, which represented in Table II, provided the effectiveness of GSC-SPP. From Table II; the suggested algorithm gives us the advantage of noise reduction while remaining the target signal; the quality of speech is improved about 2.8 ÷ 8.9 (dB) compared to with GSC-CONV.

Table ii. Snr estimation (db)

Estimation Method	Original signal	GSC-CONV	GSC-SPP
NIST SNR	4.5	5.5	8.3
WADA SNR	-3.4	1.3	10.2

V. Conclusions

Two-microphone system is well-known in speech applications, due to the low calculation, processing time; easily implementation. The challenge are more robustness of position of desired signal, a priori information about noise field, different sensitivity of microphones, complex environment require appropriate approaches to enhance noisy signals. In both conditions of coherent noise, diffuse noise; the proposal algorithm ensures save target speaker and suppressing background noise. An estimation of speech presence probability is necessary to control the updating rate, and the effectiveness of the proposal algorithm is proven.

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. Griffiths L.J., Jim C.W. An alternative approach to linearly constrained adaptive beamforming. IEEE Trans. Antennas Propag.,vol.ASSP-30, no.1, pp.27–34, Jan. 1982.
5. Hoshuyama O., Sugiyama A., Hirano A. A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters. IEEE Trans. Signal Process., Vol.47, No.10, pp.2677–2684, Oct. 1999.
6. Herbordt W., Kellermann W. Analysis of blocking matrices for generalized sidelobe cancellers for non-stationary broadband signals. IEEE International Conference on Acoustics, Speech, and Sig-nal Processing, vol.4, pp.IV-4187, May 2002.
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 measures of speech quality. https://labrosa.ee.columbia.edu/projects/snreval/.
10. SiSEC 2010 noisy speech database. http://sisec2011.wiki.irisa.fr/tikiindex79f3.html?page=Two-channel+mixtures+of+speech+and+realworld+background+noise

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.