ENHANCE LOW-FREQUENCY SUPPRESSION OF GSC BEAMFORMING

Zhaorong Hou, Ying Jia

Intel China Research Center, Beijing 100020, China
zhao.rong.hou@intel.com, ying.jia@intel.com

ABSTRACT

Usually the generalized sidelobe canceller (GSC) beamformer requires additional highpass pre-filtering due to insufficient suppression of low frequency directional interference, and it deteriorates the bandwidth quality of speech enhancement, especially for small size microphone array. This paper proposes a new GSC beamformer with multiple frequency dependent norm-constrained adaptive filters (FD-NCAF), which combine bin-wise constraint in low frequency band and norm constraint in high frequency band, to improve the performance of the adaptive interference canceller (AIC) for low frequency interference. Simulation on five testing signals shows that directional response of the proposed beamformer is less sensitive to the spectrum of interference than the full-band GSC. In the experiments based on real recordings, when the cut-off frequency of highpass pre-filtering extended to lower frequency, the residual directional interference of the full-band GSC will be increased by 3.37dB, while that of the proposed beamformer will be increased by only 0.95dB. Therefore, the passband of the proposed GSC beamformer can be extended without loss of performance.

1. INTRODUCTION

The generalized sidelobe canceller (GSC) beamforming is an effective approach for microphone array based speech enhancement to achieve high interference reduction with a small number of microphones [1]-[4]. Since classical GSC structure, such as Griffiths-Jim beamformer, is sensitive to target signal leakage and cancellation in the presence of steering-vector error, Hoshuyama *et al.* proposed a robust GSC structure with an adaptive blocking matrix (ABM) consists of coefficient-constrained adaptive filters (CCAF), and sidelobe canceller with norm-constrained adaptive filters (NCAF) to avoid target signal cancellation [5]. However, the interference rejection of this robust GSC deteriorates in the presence of colored lowpass directional interference [6]. One simple solution is to preprocess the input signal with additional highpass filters, which will deteriorate the bandwidth quality of speech enhancement. Since the required cut-off frequency is inverse proportional to the microphone spacing, the bandwidth limitation will become more serious for small size microphone array. Subband processing can also be combined with the robust GSC to enhance the suppression of low frequency directional interference [6].

On the other hand, real-time microphone array applications require as less computing complexity as possible. Compared with sample-by-sample adaptive GSC, frequency domain implementation of robust GSC based on block adaptation is more computational efficient without loss of performance [7]. Furthermore, it is easy to realize frequency dependent constraint of adaptive filter on frequency domain [8].

In this paper, first we present the formulations and analysis showing the highpass effect of the blocking matrix of GSC, and then based on frequency domain implementation of robust GSC, we present an adaptive interference canceller (AIC) with multiple frequency dependent norm-constrained adaptive filters (FD-NCAF) to improve sidelobe cancellation at low frequency band. The proposed beamformer has two advantages: first, it requires less processor consumption; and second, the frequency resolution of the bin-wise norm constraint can be better than subband norm constraint.

2. THE PROPOSED FREQUENCY DEPENDENT NORM CONSTRAINED ADAPTIVE FILTER

2.1. The highpass effect of blocking matrix

As shown in Fig.1, the frequency domain implementation of GSC beamformer consists of highpass filters, steering filters, fixed beamforming, blocking matrix and AIC. To investigate the low frequency issue, we will analyze the highpass effect of the blocking matrix in frequency domain, and present the formulations below.

In the case of the target signal comes from straight ahead, the fixed and adaptive BMs have similar transfer functions for interference. To analyze the ABM proposed by Hoshuyama [5], assume the sensor spacing of a M-sensor equi-spacing array is d, and the equivalent direction of arrival after steering compensation is θ, so the total frequency response of acoustic delay and steering compensation in channel m is

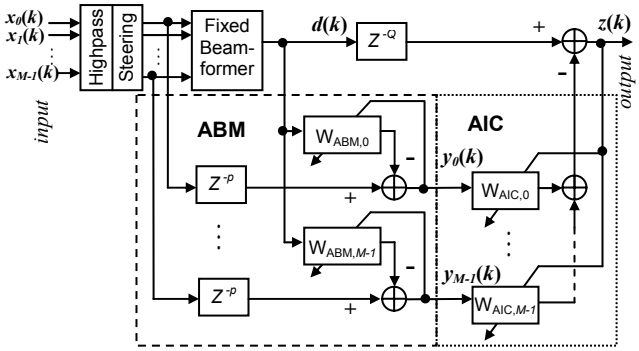

Fig.1 Diagram of the frequency domain implementation of robust GSC

$$H_m(\theta,\omega) = e^{-jm\omega\frac{d\cdot\sin\theta}{c}} \tag{1}$$

where c is the sound speed. The output of delay-and-sum beamformer is

$$H_{FBF}(\theta,\omega) = \frac{1}{M}\sum_{m=0}^{M-1} H_m(\theta,\omega), \tag{2}$$

and the output of blocking matrix in channel m is

$$H_{ABM,m}(\theta,\omega) = H_m(\theta,\omega)e^{-j\omega\tau_p} - H_{FBF}(\theta,\omega)\cdot W_{ABM,m}(\theta_T,\omega) \tag{3}$$

where τ_p is the delay compensation for ABM, and θ_T is the direction of target after steering. In the case of $\theta_T = 0°$, for completely target blocking, the ideal ABM filter response in channel m is

$$W_{ABM,m}(\theta_T,\omega)\big|_{\theta_T=0} = \frac{H_m(\theta_T,\omega)}{H_{FBF}(\theta_T,\omega)}e^{-j\omega\tau_p}\Big|_{\theta_T=0} = e^{-j\omega\tau_p} \tag{4}$$

Thus

$$H_{ABM,m}(\theta,\omega)\big|_{\theta_T=0} = e^{-j\omega\tau_p}\left(H_m(\theta,\omega) - H_{FBF}(\theta,\omega)\right) \tag{5}$$

is close to highpass filter response as shown in Fig. 2.

2.2. The desired frequency response of AIC filters for directional interference

The highpass effect of BM leads to the requirement of higher response at low frequency in the adaptive interference canceller (AIC) filters. Assuming $\theta_T = 0°$, for completely suppression of single directional interference comes from θ, the desired AIC filter responses follow

$$\sum_{m=0}^{M-1} W_{AIC,m}(\theta,\omega)H_{ABM,m}(\theta,\omega) = H_{FBF}(\theta,\omega)\cdot e^{-j\omega\tau_Q} \tag{6}$$

where τ_Q is the delay compensation for AIC. There are infinite solutions for (6), but within these solutions we can investigate the minimum sum of the filter responses

$$\underline{\Omega}(\theta,\omega) = \min\left\{\sum_{m=0}^{M-1}\left|W_{AIC,m}(\theta,\omega)\right|^2\right\} \tag{7}$$

In practice, based on the following observation,

$$\underline{\Omega}(\theta,\omega) \le \left|\frac{H_{FBF}(\theta,\omega)}{H_{ABM,m}(\theta,\omega)}\right|^2, \forall m \in [0,M) \tag{8}$$

we define the rough estimation of $\underline{\Omega}(\theta, \omega)$ as

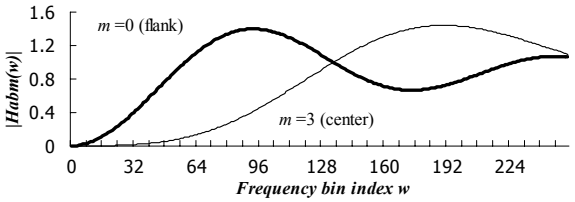

Fig. 2 Output frequency response of BM (M=8, d=4cm, θ=20°, θ_T =0, f_s = 12kHz, analyzed by 512-point DFT)

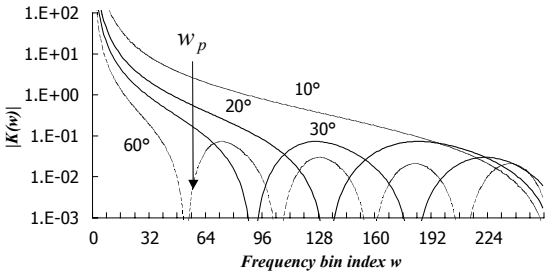

Fig. 3 Estimation of desired AIC filter response at different θ (M=8, m=0, d=4cm, f_s = 12kHz, analyzed by 512-point DFT)

$$K(\theta,\omega) = \alpha\cdot\min_{m\in[0,M)}\left\{\left|\frac{H_{FBF}(\theta,\omega)}{H_{ABM,m}(\theta,\omega)}\right|^2\right\} \tag{9}$$

where parameter α is optimized in experiments.

After analysis with 2L-point DFT as shown in Fig. 3, some conclusions can be drawn:

1. In the lower partition (the frequency bins with $w < w_p$)
   a. $K(w)$ increases rapidly as the frequency decreases, and tends to be infinite at the lowest end. So bin-wise constraint can achieve better resolution than subband constraint approach.
   b. On the other hand, $K(\theta)$ is a monotonic decreasing function of θ at each frequency bin. That is to say, to form a mainlobe with width θ_A, $K(\theta_A/2, w)$ can be used as a threshold for bin-wise constraint at bin index w.
2. In the higher partition (the frequency bins with index $w \ge w_p$), $K(\theta)$ is no longer monotonic function of θ at each individual bin w, but the following sum
   $$K_H(\theta) = \sum_{w=w_p}^{L-1} K(\theta,w) \tag{10}$$
   is still a monotonic decreasing function of θ. Therefore, $K_H(\theta_A/2)$ can be used as the threshold for group constraint in this partition.

2.3. The frequency dependent norm-constrained AIC

The block-adaptive interference canceller in Fig.1 utilizes FFT to implement fast FIR convolution, and the bin-wise NLMS adaptation [7] of filter coefficients is

$$W'_{m,w}(k+1) = W_{m,w}(k) + \beta\frac{Z_w(k)}{P_w(k)}Y_{m,w}(k) \tag{11}$$

where w is the frequency bin index, $P_w(k)$ is the recursive

average power of $Y_{m,w}$ in the bin w at the kth block, and β is the step size parameter.

In the frequency domain robust GSC [7], the full-band NCAF proposed by [5] is implemented in the following approach:

$$\Omega = \sum_{m=0}^{M-1}\sum_{w=0}^{L-1}\left|W'_{m,w}(k+1)\right|^2 \tag{12}$$

$$W_{m,w}(k+1) = \begin{cases} \sqrt{\frac{K}{\Omega}}W'_{m,w}(k+1), & for:\Omega > K \\ W'_{m,w}(k+1) & otherwise \end{cases} \tag{13}$$

where K is a parameter determined by experiment.

Due to the highpass effect of blocking matrix, the beam width of the full-band NCAF with one fixed threshold will be enlarged when the interference signal contains more low frequency components.

Following the analysis in 2.1-2.2 and Fig. 3, to achieve constant beamwidth, the optimum constraint threshold depends on the frequency. Therefore, we present a frequency dependent norm constraint method for $W_{m,w}(k)$, which combines bin-wise constraint in the lower band and group constraint in higher band to extend (12) and (13).

In the lower partition ($w \le w_p$) of a 2L-point FFT, bin-wise constraint is performed as

$$\Omega_w = \sum_{m=0}^{M-1}\left|W'_{m,w}(k+1)\right|^2 \tag{14}$$

$$W_{m,w}(k+1) = \begin{cases} \sqrt{\frac{K_w}{\Omega_w}}W'_{m,w}(k+1), & for:\Omega_w > K_w \\ W'_{m,w}(k+1) & otherwise \end{cases} \tag{15}$$

where $K_w = K\left(\frac{\theta_A}{2}, w\right)$ is a threshold at bin w, and θ_A is the desired mainlobe width.

In the higher partition, group constraint is performed

$$\Omega_H = \sum_{m=0}^{M-1}\sum_{w=w_p}^{L-1}\left|W'_{m,w}(k+1)\right|^2 \tag{16}$$

$$W_{m,w}(k+1) = \begin{cases} \sqrt{\frac{K_H}{\Omega_H}}W'_{m,w}(k+1), & for:\Omega_H > K_H \\ W'_{m,w}(k+1) & otherwise \end{cases} \tag{17}$$

where $K_H = \sum_{w=w_p}^{L-1} K\left(\frac{\theta_A}{2}, w\right)$.

3. SIMULATION AND EXPERIMENT RESULTS

Simulations were conducted to examine the directional responses of GSC algorithms with NCAF and FD-NCAF in the following setup: M=8, d=4cm, f_s=12kHz. Highpass pre-filtering was skipped in this simulation. Five typical signals with different spectrums were used as input: white Gaussian noise, lowpass Gaussian noise (cutoff at 500Hz), highpass Gaussian noise (cutoff at 1kHz), female voice and male voice from NIST Speech Disc. The threshold of NCAF was optimized for highpass Gaussian noise, and the thresholds set of FD-NCAF was designed for a mainlobe of [-7°, +7°]. During simulation, the ABM was fixed for direction 0°, and the adaptation of AIC was active except at 0°.

The results are shown in Fig. 4 and 5. NCAF with one fixed threshold provides quite different directional response for each testing signal, while the proposed FD-NCAF achieves coherent beamwidth for all testing signals with the same thresholds set. The improvement on suppression of lowpass Gaussian noise outside the mainlobe (7°) is up to 18 dB over NCAF.

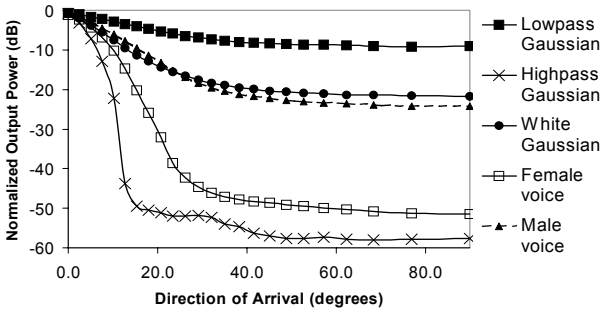

Fig. 4 Directional response of robust GSC with NCAF (ABM is fixed to 0°)

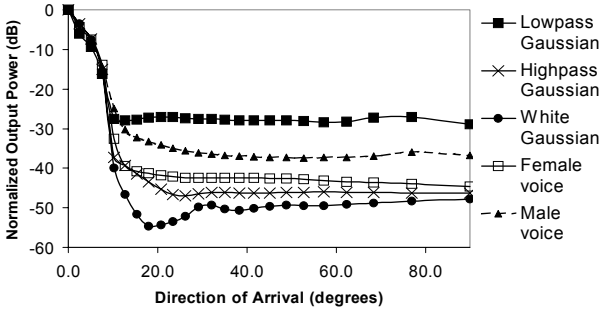

Fig. 5 Directional response of robust GSC with FD-NCAF (ABM is fixed to 0°)

Experiment based on real recordings was also conducted to validate the quality of GSC beamforming with FD-NCAF in typical conference room environment. The parameters of microphone array and the threshold settings of the two beamformers were the same as those in the simulation. They were tested twice with two different highpass cut-off frequencies 400Hz and 100Hz. The output powers of the single microphone, the GSC with full-band NCAF and the GSC with FD-NCAF are drawn in Fig. 6. The target talker was active during the period of [0.88, 2.04], and there were two interference talkers speaking from 45° and -45° during the period of [0, 0.5], and [3.1, 6.2]. During the active period of target talker, the adaptation of ABM was active while the adaptation of AIC was frozen. In other periods, the blocking matrix was fixed to 0° while the adaptation of AIC was active.

The target cancellation, and the residual directional interference were measured on the average of corresponding periods of power curve in Fig. 6, while the residual diffuse field noise was measured during the period of [6.2, 7.1]. They are shown in Table 1.

Based on the following observations from Table 1:

1. Basically, the target cancellation of FD-NCAF is at the same level of full-band NCAF.

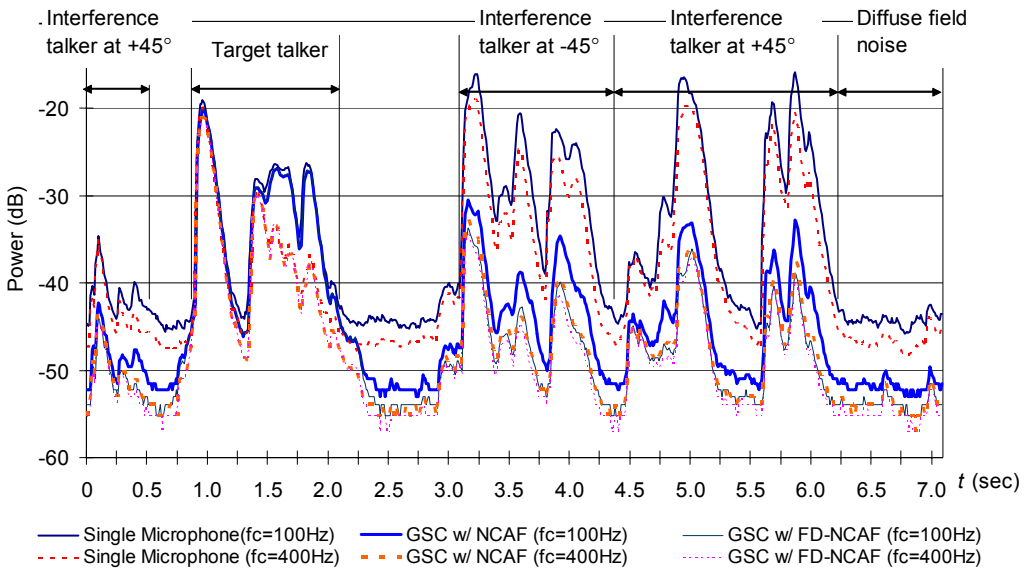

Fig. 6 Output power of single microphone, the GSC with NCAF and the GSC with FD-NCAF (M=8, d=4cm, f_s=12kHz)

2. For the cut-off frequency of 400Hz, the residual directional interferences of the two beamformers are basically at the same level.

We can draw the conclusions below:

i. When the cut-off frequency of highpass filter is extended to 100Hz, the residual directional interference of full-band NCAF will be increased by 3.37dB, while that of FD-NCAF will be increased by only 0.95dB.

ii. The proposed FD-NCAF is also effective to control the increase of residual diffuse field noise when the cut-off frequency of highpass filter is extended to lower frequency.

Table 1. Experiment result of FD-NACF

Algorithm	NCAF		FD-NCAF	
Cut-off frequency (Hz)	400	100	400	100
Target Cancellation (dB)	-0.80	-0.89	-0.79	-0.80
Residual directional interference (dB)	-42.83	-39.46	-44.49	-43.54
- Increased by		3.37		0.95
Residual diffuse field noise (dB)	-53.90	-51.72	-54.84	-53.79
- Increased by		2.19		1.05

4. CONCLUSIONS

In this paper, multiple frequency dependent norm constrained adaptive filters are proposed to improve the low frequency suppression of GSC beamformer. Simulation and experiment results show that directional response of the proposed beamformer is less sensitive to the spectrum of interference than the full-band GSC, and it is effective to control the residual directional interference when the passband is extended to lower frequency or the microphone spacing is reduced.

5. ACKNOWLEDGEMENT

The authors would like to thank Wolfgang Herbordt and Walter Kellermann, University Erlangen-Nuremberg, Germany, for making their real-time GSC code available and for fruitful discussions about robust GSC.

6. REFERENCES

[1] L. J. Griffiths and C. W. Jim, "An alternative approach to linear constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27-34, Jan. 1982.

[2] B.D. Van Veen, K.M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, pp. 4-24, April 1988.

[3] J. E. Greenberg, and P. M. Zurek, "Evaluation of an adaptive beamforming method for hearing aids," *J. Acoust. Soc. Amer.*, vol. 91, no. 3, pp. 1662-1676, Mar. 1992.

[4] M. S. Brandstein, D. B. Ward. *Microphone arrays: Signal processing techniques and applications*, Springer, Berlin, 2001.

[5] O. Hoshuyama, A. Sugiyama, A. Hirano, "A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters," *IEEE Trans. on Signal Processing*, vol. 47, no. 10, pp. 2677-2684, October 1999

[6] W. H. Neo and B. Farhang-Boroujeny, "Robust microphone arrays using subband adaptive filters," *IEE Proc.-Vis. Image Signal Process*., Vol. 149, No. 1, February 2002, pp.17-25

[7] W. Herbordt and W. Kellermann, "Efficient frequency-domain realization of robust generalized sidelobe cancellers," *IEEE Workshop on Multimedia Signal Processing*, October 2001, pp. 377-382

[8] B. Rafaely and S. J. Elliott, "A computationally efficient frequency-domain LMS algorithm with constraints on the adaptive filter," *IEEE Trans. on Signal Processing*, vol. 48, no. 6, pp. 1649-1655, June 2000