The CSI Journal on
Computer Science and Engineering
Vol. 3, No. 2&4, 2005
Pages 39-45
Regular Paper

Subband Beamforming of Miniature Microphone Arrays for Audio Applications*

S.M. Makki M.T. Manzuri-Shalmani

Department of Computer Engineering, Sharif University of Technology, Tehran, Iran, and
IPM School of Computer Science, Tehran, Iran

Abstract

Microphone array processing techniques have generated a tremendous interest in both theoretical and applied areas, especially, over the past decade. So long as an array performs a type of spatial filtering, the dimension of array will determine the acceptable frequency-range of the incoming waves. Consequently, it is well known that, miniature microphone arrays are not appropriate for audio applications. In this paper, we introduce a novel technique to utilize miniature microphone arrays for detecting audio signals. Besides, the problem of low sensitivity of such arrays in low frequency beamforming, due to dealing with wideband audio signals, is resolved. We start our discussion with a Delay-and-Sum beamformer, and extend it to the GSC-based beamformer, as an example of popular beamformers. Constant directivity is an important result of the proposed approach.

Keywords: Beamforming, microphone array, Filter banks, delay-and-sum beamformer, generalized sidelobe canceller.

1. Introduction

Microphone arrays are very popular in nowadays electronic supplies. Tape recorders in automobiles, laptops, hearing aid instruments and many other electronic instruments use microphone arrays. We can say everywhere that sound directionality is an issue, microphone array is one of the first choices. By beamforming techniques, not only the sound will be grasped directionally, but also a natural enhancement especially for voice is obtained. A microphone array with a beamforming algorithm, create an automatic beamsteering, instead of changing the direction of the microphone, mechanically. This is what makes microphone arrays popular for sound applications.

Beamsteering is accomplished by choosing suitable dimensions for the applied array. It is well known that in audio applications, the dimension of a microphone array cannot be very small. For example, the inter-spacing between the microphones of a normal array is typically not less than 4cm. From another side, recent developments in the field of small electronic appliances, allow us to fabricate miniature microphones arrays, in the range of a few millimeters. However, there is no well-behaved algorithm to be able to achieve the beamsteering job for these kinds of arrays.

There is a long list of various researches about microphone arrays, which are mostly about beamsteering techniques. For example, in some works, subband techniques have been used for speech recognition [2,3] or speech enhancement [13].

The idea of spectral subtraction is another issue that can be mixed with subband techniques [4]. In abovementioned techniques a filter bank is usually used [5,10]. Adaptive filtering is also a well-known tool in beamforming, and especially when subband techniques are the choice. Research in this area is not new but is still in progress [6-9,11,12]. In [21], an algorithm has been presented to improve the beamforming in low frequencies, using Linear Prediction Coding. There is another algorithm which utilize Gradient Flow concept; however it is more a localization algorithm than a beamsteering tool, and is very sensitive to noise [20]. In this paper we introduce a novel technique that works in low frequencies and for small arrays, as well as, in high frequencies and for typical arrays.

We start our discussion with a Delay-and-Sum beamformer. Section two is denoted to the problem formulation. In this section we discuss the problem from mathematical point of view. Section three presents the theoretical solution of the problem. In this section we introduce the block diagram of the proposed system, and study the Constant-Directivity issue. In section four, the proposed algorithm is applied to a GSC-based system, and necessary modifications are explained. In section five we evaluate the performance of the proposed method by experimental results. Finally, the subjects are concluded in section six.

2. Problem Formulation

Before introducing the technique, the problem must be well understood. To understand the problem, consider Figure 1 which shows a typical beam-former. For seek of simplicity, assume that the beam-former is a Delay-and-Sum one. In this figure, θ is the direction of arrival (DOA), and x_is are input signals. Besides,

$$Z\{y(x_i;\theta)\} = Y(z) \tag{1}$$

Here, $Y(z)$ indicates the Z transform of $y(x_i, \theta)$.

Definition 1: In system shown in Figure 1, if for a specific direction of arrival, $Y(z_p)=1$, we say the system has a pole in z_p.

Definition 2: In system shown in Figure 1, if for a specific direction of arrival, $Y(z_i)=0$, we say the system has a zero in z_i.

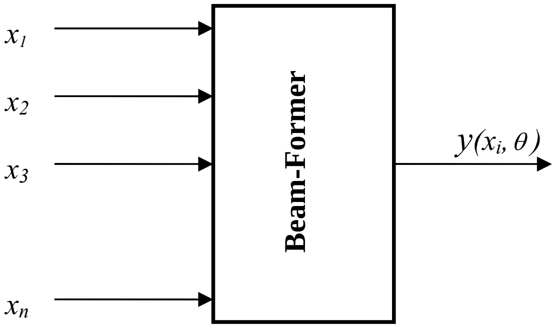


Figure1. A typical beamformer

Theorem 1: Without lose of generality, assume that the system in Figure 1 has a pole in $\theta = 90^\circ$; then, it will be impossible to have a zero by adding delays for $d < \lambda / n$. Note that, d is inter-spacing between the microphones of array, n is the number of microphones, and λ is the wavelength of signal.

Proof: We define the fundamental frequency of an array as follows:

$$f_a \overset{\Delta}{=} \frac{c_s}{nd} \tag{2}$$

Here, c_s is the speed of sound. Then the distance between two microphones will be equal to:

$$d = \frac{c_s}{nf_a} = \frac{f\lambda}{nf_a} \tag{3}$$

Note that f denotes the frequency of sound. Now the arrival wave can be defined as:

$$y(x_i;\theta) = \sum_{i=1}^{n} x_i(t - t_i) \tag{4}$$

where in Z domain can be presented as

$$Y(z) = \sum_{i=1}^{n} z^{-t_i}.X(z) \tag{5}$$

Finally, the transfer function of beam-former can be derived as:

$$F(z) = \sum_{i=1}^{n} e^{-j\omega t_i} = \sum_{i=1}^{n} e^{\frac{-j2\pi f d}{ndf_a}} = \sum_{i=1}^{n} e^{\frac{-j2\pi}{nK}} \tag{6}$$

where, K presents the normalized frequency, and is defined by:

$$K \overset{\Delta}{=} \frac{f_a}{f} \tag{7}$$

Note that if $\lambda > d.n$, then we will have $K>1$, and vice versa. From another side, it can easily be shown that for $K>1$:

$$\sum_{i=1}^{n} e^{-j\frac{2\pi}{nK}} \neq 0 \tag{8}$$

This means that the transfer function of system in Figure 1 will not have any zero. In other words, beamforming by such a system is not possible for small-size arrays. For example if we assume that d=5cm and n=4, then for any sound-wave with a wavelength above 20 cm (frequencies below 16 KHz), the value of K will be greater than one, and consequently it is impossible to force the transfer function of array to be zero in a desired direction. As a result, the inter-space of microphones in array can not be chosen less than 4cm, for audio applications.

3. Proposed Method

To solve this problem, remember that the sound hits different microphones with the same speed but with different wavelengths. For a specific wavelength, if the inter-spacing of microphones is too small, as in the case of miniature microphone arrays, beamsteering cannot be carried out, properly. From another point of view, by looking over the transfer function of array, it is well know that each delay in the system is related to a pole of that system [8]. Consequently, poles are in different locations but are spaced in low frequency section of a zero-pole plot.

Referring to Figure 2, if we transfer the poles from low frequency part to high frequency section, beamsteering for low-frequency waves become possible. In fact, by using this method, we temporarily increase the frequency components of the coming waves and decrease the apparent wavelengths. This can be done by shifting in frequency domain

$$F_1(e^{j\omega}) = F(e^{j\omega_0}.e^{j\omega}) = F(e^{j(\omega+\omega_0)}) \tag{9}$$

Here, ω_0 indicates the quantity of rotation and upshifting.

Regarding Figure 2, another key point in the proposed approach can be outlined as pole-alignment in a row. As it has shown in Figure 2, the poles have been rotated in such a way that after rotation, they have been aligned in a row. This guaranties that the beamwidths of all poles to be equal.

Figure 3 shows the block diagram of the proposed system. Obviously, audio signals, especially speech waveforms are wideband signals. Consequently, we use a filter bank to be able to assume the frequency components are not changing in each subband. The outputs of the analysis filter bank are downsampled to resume the number of samples. After downsampling and before beamforming, frequency upshifting is carried out. Therefore, each beamformer will be working in each subband and we can use narrowband beamformers.

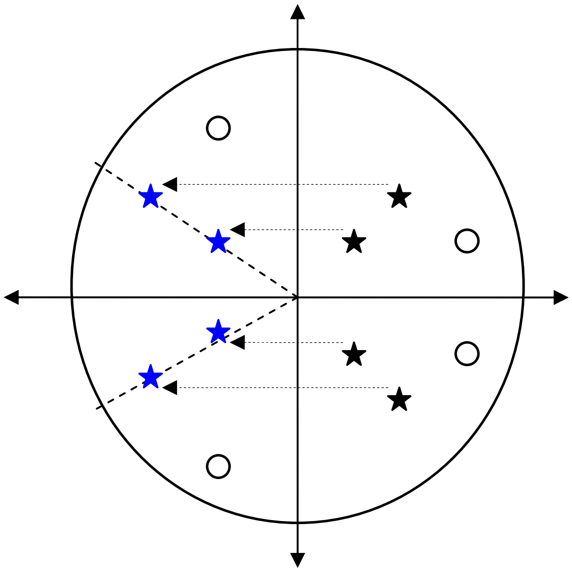


Figure 2. Rotation of poles in z-plane, due to proposed method

After beamforming, frequency is down-shifted and upsampling is carried. Finally, the upsampled signals pass through the synthesis filter bank to obtain the beamformed signal.

In appendix it is shown that the relationship between input signals from array-microphones, and output signal of beamformer is as follows:

$$Y(z) = \sum_{k=1}^{n} A_k(z).X_k(z.W_D^d) \tag{10}$$

where,

$$A_k(z) = \sum_{d=0}^{D-1}\sum_{m=0}^{M-1} A_{m,d,k}(z), \text{ and } \quad W_D = e^{-j\frac{2\pi}{D}} \tag{11}$$

Note that:

$$A_{m,d,k}(z) = \frac{1}{D} F_{m,k}(z^D).H_k(z.W_M^m.W_D^d).G_k(z.W_M^m) \tag{12}$$

presents the transfer function between input signal X_k and output signal of beamformer.

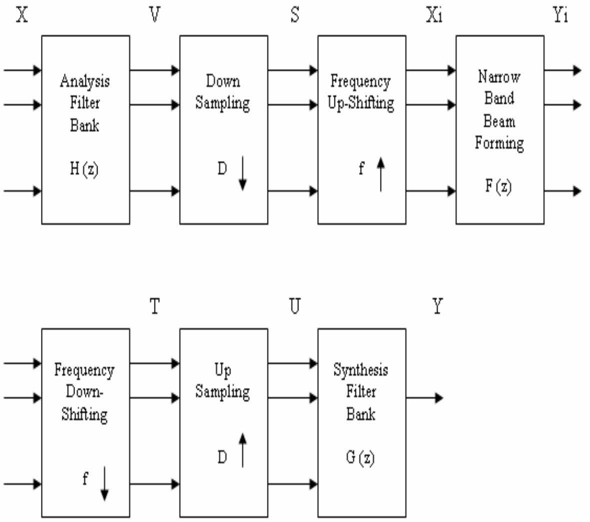


Figure 3. Block diagram of the proposed system

***Theorem 2**:* The system in Figure 3 with the relations (10) to (12) shows a beamformer which have at least one pole and one zero, regardless of the value of *d*.

Proof: Consider equation (3). For a specified value of *n*, increasing *f* will decrease the proper value of *d*. Now, the system defined in Figure 3 permit us to perform frequency up-shifting operation, and consequently, the proper value of *d* can be defined and adjusted. Hence, based on theorem 1 this system has at least one pole and one zero.

Theorem 3: The system in Figure 3 with the relations (10) to (12) is a constant beamwidth system.

Proof: We are changing the transfer function of beamformer from *F(z)* to *F(az)*. On the other hand, the beamwidth is dependent on frequency of poles. But as it has shown in Figure 2, the frequency upshifter aligns all poles of system on a row. Therefore, the beamwidth for all frequency components of sound will be the same.

4. GSC-based Beamformers

The problems and solutions mentioned above can be extended to GSC-based beamformers in a similar manner. Figure 4 shows a typical generalized side-lobe canceller (GSC) beamformer. Here, the fixed beamformer (FBF) provides a pole for the system, while the blocking matrix (BM) provides a zero. Multiple input canceller (MIC) is the adaptive part of the system. Many researches have been reported about using subband techniques for GSC beamformers [14-19].

For seek of simplicity, we assume that the BM part of the system is fixed. Normally, the sampling rate for all parts of the beamformer is the same. Hence, the system is combined of two fixed sections, and one adaptive part. But by using the frequency shifting technique, adaptive filters cannot follow the stringent time dictated by this approach, because we have to use a very long FIR filter, which can be quite impractical. In order to prevent such problems we can apply the frequency shifting technique as shown in Figure 5. In this figure which is an overall view of the system, frequency down-shifting is carried out before the adaptive filter used in MIC. Moreover, *K* presents the ratio of high to low sampling frequencies.

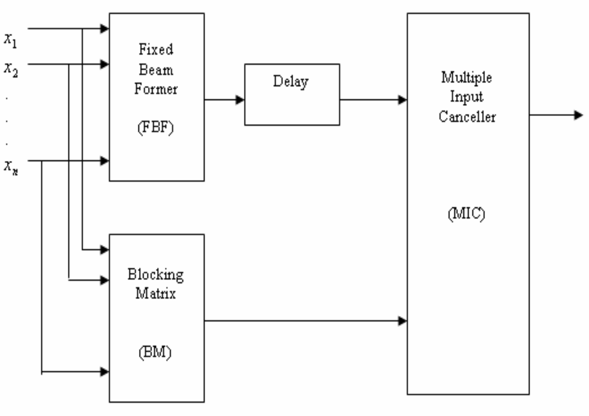


Figure 4. Block diagram of a GSC beamformer

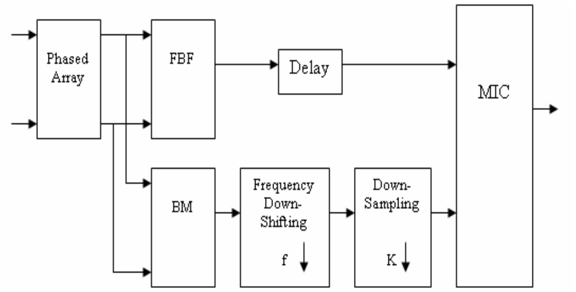
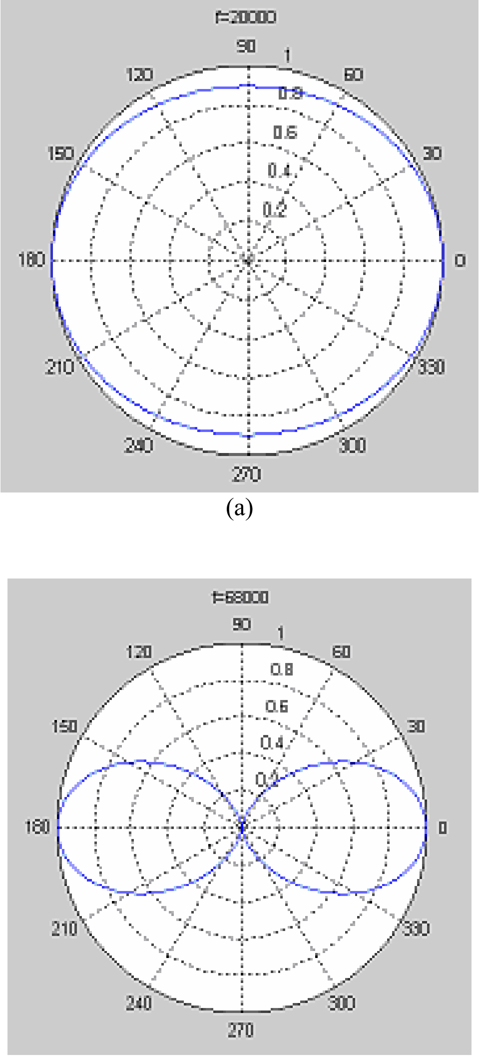


Figure 5. Block diagram of the modified GSC beamformer

5. Performance Evaluation

Let us look over the performance of the presented approach. We assume n=2, d=2.5mm and $\tau=d/c_s$=7.35µsec. Here, τ is the maximum delay of array. Fig. 6a shows the polar plot for $f=20$ KHz and Figure 6b shows the polar plot for $f=68$ KHz, which is equal to the fundamental frequency array.

It is clear that beamforming for the highest possible sound cannot be done by any array with any delay value. But the same array can be used for much lower frequency by the frequency shifting technique described before. In Figure 7a, sounds coming from different angles pass the array. On the other hand, as shown in Figure 7b, sounds coming from the front are passed and sounds coming from other angles are attenuated. If we want to have a narrower lobe, we can use more microphones. Figure 8 shows the polar plot having $n=6$ microphones.



(a)

(b)

Figure 6. Polar plots, (a)-before and (b)-after applying frequency shifting

Definition 3: $\gamma = \rho_{max}(\theta)/\rho_{min}(\theta)$ is a criterion for effective beamforming. Here, ρ presents the magnitude of the beamformer for a specific angle. Note that $\gamma \rightarrow 1$ means inefficient beamforming and $\gamma \rightarrow \infty$ means a fully efficient beamforming.

Definition 4:

$$\psi = \frac{\int_{\varphi_1}^{\varphi_2} y(\theta)d\theta}{\int_{-\pi/2}^{\varphi_1} y(\theta)d\theta + \int_{\varphi_2}^{\pi/2} y(\theta)d\theta}$$

is another criterion for effective beamforming. Here, y is the magnitude of the output of the system (Figure 1). The more the value of ψ, the better the beamforming is.

Table 1 shows the effects of using frequency shifting technique, as well as increasing the number of microphones. As it is clear from this table, without performing the frequency shifting method, γ is very close to unity, while using this technique causes the value of γ approaches large values. The same results is obtained for ψ too, except that for this performance metrics the changes are smoother.

Table 1. Performance metrics evaluation for using frequency shifting technique

	n=2, No frequency Shifting	n=2, proposed Method	n=6, proposed Method
γ	1.1171	25.2219	19.5036
ψ	1.0581	2.2322	4.9263

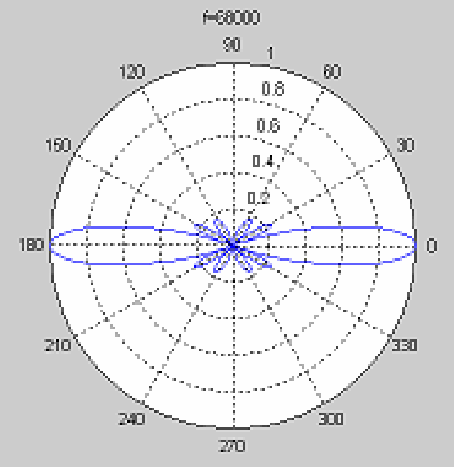


Figure 8. Polar plot after frequency shifting for 6 microphones

To obtain a better imagination of what really happens, look at Figure 9. For this figure it is assumed that a 1 KHz acoustic wave is hitting the array with a direction of arrival equal to 90 degrees, and another wave with frequency of 3 KHz, is hitting the array, while its direction has 40 degrees difference with array direction. Obviously, the system passes the first signal while attenuates the second one.

Figure 10 shows the changes in angle domain. The continuous curve is for a 500 Hz signal while the dotted curve is for a 5 KHz signal. Using frequency shifting technique makes the two curves to be identical, drastically.

6. Conclusion

Beamforming for small-size arrays in the range of audio frequencies fails to eliminate sounds coming from side angles. This problem can not be resolved by changing delay values in a phased array. But the frequency shifting technique described in this paper solves the problem and opens the doors of beamforming world to small-size arrays for audio applications. Performance evaluation shows promising results for beamsteering, using this method. Constant directivity is another gift of this technique.

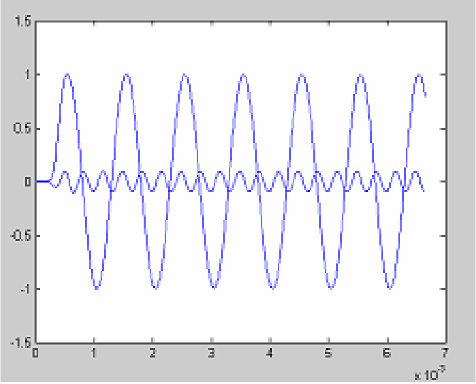


Figure 9. Attenuation of signals outside the zooming direction

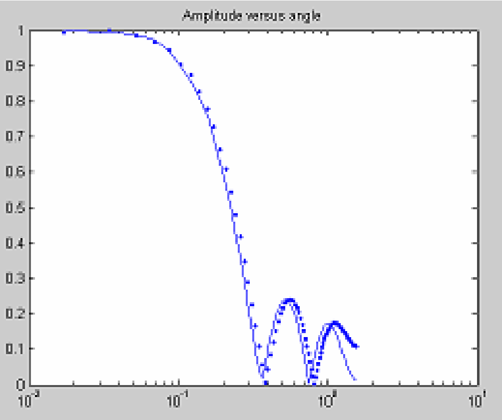


Figure 10. Amplitude versus angle for 500 Hz (continuous curve) and 5 KHz (dotted curve) frequencies

References

[1] B. A. Goh, D.C. McLernon, A.G. Orozco-Lugo and M.Lara, "A New Subband Structure for an Acoustic Beamformer with Leaky Adaptive Filters in the Blocking matrices," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 4, pp. 801-804, 2005.

[2] M.L. Seltzer and R.M. Stern, "Parameter Sharing in Subband Likelihood-Maximizing Beamforming for Speech Recognition Using Microphone Arrays," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 1, pp. 881-884, 2004.

[3] M.L. Seltzer and R.M. Stern, "Subband Parameter Optimization of Microphone Arrays for Speech Recognition in Reverberant Environments," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 1, pp. 408-411, 2003.

[4] L. Yong, N. Grbic and S. Nordholm, "Robust Microphone Array Using Subband Adaptive Beamformer and Spectral Subtraction," *Proc. of the 8th International*

Conference on Communication Systems, Vol. 2, pp. 1020-1024, 2002.

[5] J. M. Haan, N. Grbic, and S. Nordholm, "Design and Evaluation of Nonuniform DFT Filter Banks in Subband Microphone Arrays," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 2, pp. 1173-1176, 2002.

[6] W. H. Neo and B. Boroujeny, "Robust Microphone Arrays Using Subband Adaptive Filters," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 6, pp. 3721-3724, 2001.

[7] F. Lorenzelli, A. Wang and K. Yao, "Broadband Array Processing Using Subband Techniques," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 5, pp. 2876-2879, 1996.

[8] S. Nordholm, J.M. Haan I. Claesson and N. Grbic, "Performance Limits in Subband Beamforming," *IEEE Transactions on Speech and Audio Processing,* Vol. 11, No. 3, pp. 193-203, 2003.

[9] Y. Denda, T. Nishiura, and H. Kawahara, "Study of Talker Localization Based on Subband CSP Analysis in Real Noisy Environments," *Proc. of IEEE-Eurasip Nonlinear Signal and Image Processing,* pp. 27-31, 2005.

[10] J. M. Haan, N. Grbic, I. Claesson and S. Nordholm, "Filter Bank Design for Subband Adaptive Microphone Arrays," *IEEE Transactions on Speech and Audio Processing*, Vol. 11, No. 1, pp. 14-23, 2003.

[11] Y. R. Zheng, R.A. Goubran and M. El-Tanany, "Experimental Evaluation of a Nested Microphone Array with Adaptive Noise Cancellers," *IEEE Transactions on Instrumentation and Measurement,* Vol. 53, No. 3, pp. 777-786, 2004.

[12] W. H. Neo and B. Farhang-Boroujeny, "Robust Microphone Arrays Using Subband Adaptive Filters," *IEEE Proceedings 0n Vision, Image and Signal Processing*, Vol. 149, No. 1, pp. 17-25, 2002.

[13] L. Yong, N. Grbic and S. Nordholm, "Speech Enhancement Using Multiple Soft Constrained Subband Beamformers and Non-Coherent Technique," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 5, pp. 489-492, 2003.

[14] L. Yong, S. Nordholm and N. Grbic, "Subband Generalized Sidelobe Canceller - a Constrained Region Approach," *Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 41-44, 2003.

[15] L. Wei, S. Weiss and L. Hanzo, "Subband Adaptive Generalized Sidelobe Canceller for Broadband Beamforming," *Proc. of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, pp. 591-594, 2001.

[16] L. Wei, S. Weiss and L. Hanzo, "A Subband-Selective Broadband GSC with Cosine-Modulated Blocking Matrix," *IEEE Transactions on Antennas and Propagation,* Vol. 52, No. 3, pp. 813-820, 2004.

[17] L. Wei, S. Weiss and L. Hanzo, "Subband-Selective Partially Adaptive Broadband Beamforming with Cosine-Modulated Blocking Matrix," *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing,* Vol. 3, pp. 2913-2916, 2002.

[18] S. Weiss, R.W. Stewart, M. Schabert, I.K. Proudler, and M.W. Hoffman, "An Efficient Scheme for Broadband Adaptive Beamforming," *Proc. of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*, Vol. 1, pp. 496-500, 1999.

[19] S. Gannot and I. Cohen, "Speech Enhancement Based on the General Transfer Function GSC and Postfiltering," *Proc. of IEEE Transactions on Speech and Audio Processing,* Vol. 12, No. 6, pp. 561-571, 2004.

[20] M. Stanacevic, G. Cauwenberghs, and G. Zweig, "Gradient Flow Adaptive Beamforming and Signal Separation in a Miniature Microphone Array," *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 4, pp. 4016-4019, 2002.

[21] C. Yuchang, and S. Sridharan, "Improving the Performance of a Small Microphone Array at Low Frequencies Using Critical Band and LPC Codebooks," *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 2, pp. 1033-1036, 2000.

Appendix: Calculating the Transfer Function of Proposed System

Referring to Figure 2, one can calculate the transfer function of system as follows. Assume that h(n) and g(n) present the impulse responses of two low pass filters. Then the impulse response of each filter in analysis and synthesis filter banks, and their Z transforms can be shown as:

$$h_m(n) = h(n).W_M^{-mn} \qquad \leftrightarrow \qquad H_m(z) = H(z.W_M^m)$$

$$g_m(n) = g(n).W_M^{-mn} \qquad \leftrightarrow \qquad G_m(z) = G(z.W_M^m)$$

where, $W_M = e^{-j\frac{2\pi}{M}}$.

Now the output signals of analysis filter bank can be defined as:

$$V_m(z) = H_m(z).X(z) = H(z.W_M^m).X(z)$$

Next, in the output of sampling block we will have:

$$S_m(z) = \frac{1}{D}\sum_{d=0}^{D-1} V_m(z^{\frac{1}{D}}.W_D^d) = \frac{1}{D}\sum_{d=0}^{D-1} H(z^{\frac{1}{D}}.W_M^m.W_D^d).X(z^{\frac{1}{D}}.W_D^d)$$

The relationship between the outputs and inputs of upshifting block can be expressed as:

$$X_m(z) = S_m(a_m z) = \frac{1}{D}\sum_{d=0}^{D-1} H[(a_m z)^{\frac{1}{D}}.W_M^m.W_D^d].X[(a_m z)^{\frac{1}{D}}.W_D^d] \cdot$$

Similarly, the outputs of narrow band beamforming section will be equal to:

$$Y_m(z) = F_m(z).X_m(z).$$

After beamforming, the main job has been done, and we are ready to demodulate the signal. For this purpose, first perform downshifting operation:

$$T_m(z) = Y_m(z / a_m) = \frac{1}{D}\sum_{d=0}^{D-1} F_m(z).H(z^{\frac{1}{D}}.W_M^m.W_D^d).X(z^{\frac{1}{D}}.W_D^d) \cdot$$

Next, the resulting signals will be upsampled. That is:

$$U_m(z) = T_m(z^D) = \frac{1}{D} F_m(z^D) \sum_{d=0}^{D-1} H(z.W_M^m.W_D^d).X(z.W_D^d) \cdot$$

Finally, the output of beamformer can be determined:

$$Y(z) = \sum_{m=0}^{M-1} U_m(z).G_m(z) \cdot$$

Equivalently, we will have:

$$Y(z) = \frac{1}{D} \sum_{d=0}^{D-1} X(z.W_D^d) \sum_{m=0}^{M-1} F_m(z^D).H(z.W_M^m.W_D^d).G(z.W_M^m) \cdot$$

Assuming, we have an array with *n* microphones, it can be concluded that:

$$Y(z) = \frac{1}{D} \sum_{k=0}^{n-1} \sum_{d=0}^{D-1} X_k(z.W_D^d) \sum_{m=0}^{M-1} F_{m,k}(z^D).H_k(z.W_M^m.W_D^d).G_k(z.W_M^m) \cdot$$

Finally, we can formulate the relationship between the output and input signals of the proposed system as:

$$Y(z) = \sum_{k=1}^{n} A_k(z).X_k(z.W_D^d)$$

where,

$$A_{m,d} = \frac{1}{D} F_{m,k}(z^D).H_k(z.W_M^m.W_D^d).G_k(z.W_M^m)$$

and

$$A_k(z) = \sum_{d=0}^{D-1} \sum_{m=0}^{M-1} A_{m,d,k}(z) \cdot$$

Seyed Mohammad Makki received his B.S. in Electrical Engineering from Amirkabir University of Technology in Iran in 1997, and his M.S. in Electrical Engineering from Sharif University of Technology in Iran in 1999. His research interests are signal processing and array structures.
Email: Makki72@yahoo.com

Mohammad T. Manzuri Shalmani received his B.S. and M.S. in Electrical Engineering from Sharif University of Technology (SUT), Iran, in 1984 and 1988, respectively; and PhD in Electrical and Computer Engineering from Vienna University of Technology, Austria, in 1995. Currently, he is an assistant professor in Computer Engineering Department of SUT, Tehran, Iran. His main research interests include digital signal processing, robotics, image processing, and data communications.
Email: manzuri@sharif.com

*****This research was in part supported by a grant from IPM School of Computer Science (No. CS1385-4-03).