

Improvement of Microphone Array Characteristics for Speech Capturing

Stolbov Mikhail Borisovich¹ & Aleinik Sergei Vladimirovich¹

¹ St. Petersburg National Research University of Information Technologies, Mechanics and Optics, Russia

Correspondence: Stolbov Mikhail Borisovich, St. Petersburg National Research University of Information Technologies, Mechanics & Optics, 197101, St.Petersburg, Kronverkskiy pr, 49, Speech Technology Center, Russia, 196084, St.Petersburg, Krasutskogo street, 4, Russia. E-mail: stolbov@speechpro.com/aleinik@speechpro.com

Received: December 18, 2014

Accepted: January 25, 2015

Online Published: May 30, 2015

doi:10.5539/mas.v9n6p310

URL: <http://dx.doi.org/10.5539/mas.v9n6p310>

Abstract

This paper presents a new adaptive technique for speech capturing in adverse conditions using microphone arrays. The proposed technique is based on frequency-domain alignment of microphone signals with the output of the fixed beamformer directed to the target speaker. This alignment procedure improves pattern directivity and reduces sidelobes. The low complexity of the technique is achieved by means of a frequency-domain implementation of the algorithm. This makes it possible to implement this technique in real-time applications with a large number of microphones. The technique was evaluated on speech data corrupted by varying levels and directions of noise and interference. The proposed technique improves the directivity pattern of a traditional Dealy & Sum beamformer as well as provides additional suppression of spatially incoherent noise, diffuse noise and interference, with minimal loss of the target signal quality.

Keywords: microphone array, speech processing, noise reduction, adaptive beamforming

1. Introduction

Microphone arrays (MAs) are an effective tool for many applications (speech transcription, voice/speaker identification, etc.). MAs also enable long-distance speech capture in adverse conditions (Zang *et al.*, 2010). For speech enhancement, the main advantage of using microphone arrays rather than a single microphone is that a microphone-array-based beamformer can spatially suppress multiple interfering signals and different background noises (diffuse noises, independent noises, etc.) while maintaining minimum distortion of the target signal from the look direction (Gannot, *et al.* 2001). A great number of well-known fixed or adaptive algorithms for signal processing in MA are described in (Widrow & Stearns, 1985) and (Brandstein & Ward, 2010). Comparison of different beamforming techniques may be found, for example, in (VanVeen & Buckley 1988). Different approaches can be classified into three main categories (Fischer & Simmer, 1995):

- Conventional beamforming (Kellermann 1991), (Mahieux, *et al.* 1993).
- Adaptive beamforming (Frost, 1972), (Hoshuyama, *et al.* 1999).
- Adaptive postfiltering (Zelinski, 1988), (Simmer, *et al.* 2001).

A conventional fixed beamformer (FBF) uses known “Delay & Sum” or “Filter & Sum” algorithms (Flanagan, *et al.* 1985), enables the suppression of both coherent interference and statistically independent noise. However, not always with enough good efficiency. Different adaptive algorithms (Frost, 1972), (Griffiths & Jim, 1982), etc. can be used for more efficient coherent interference suppression, while adaptive postfiltering (Löllmann & Vary, 2007), (Cao, *et al.* 1994) can be applied for incoherent noise cancellation. Several different integrated techniques were developed for the suppression of both coherent and incoherent noise, see, for example, (Ortega-García & González-Rodríguez, 1996).

However, practical application of these algorithms presents a number of difficulties: audible music noise, target signal suppression, sensitivity to microphone mismatch (Doclo & Moonen 2007), etc. In our previous paper (Stolbov & Aleinik, 2014) we proposed a new adaptive beam-forming algorithm. It was developed for an 8-element MA and its efficiency was proved in practice by experimental studies on artificial model and real signals. This paper presents a modification of that algorithm, as well as a detailed study of the algorithm and its

characteristics, advantages and disadvantages.

2. The Proposed Technique

Here we briefly describe the modification of our method proposed in (Stolbov & Aleinik, 2014).

2.1 The Description of the Modified Method

The method is based on frequency domain alignment of microphone signals with the output of a fixed beamformer directed to the target speaker. The block diagram of the method (example for a 4-microphone array) is shown in Figure 1. We implement this procedure in frame-based frequency domain (Simmer & Wasiljeff, 1992). The three initial steps are well-known as a frequency-domain Delay & Sum FBF. At the first step, the input signals of each microphone $x_m(t)$ are transformed into spectra $X_m(f, k)$ using short-time Fourier (STFT),

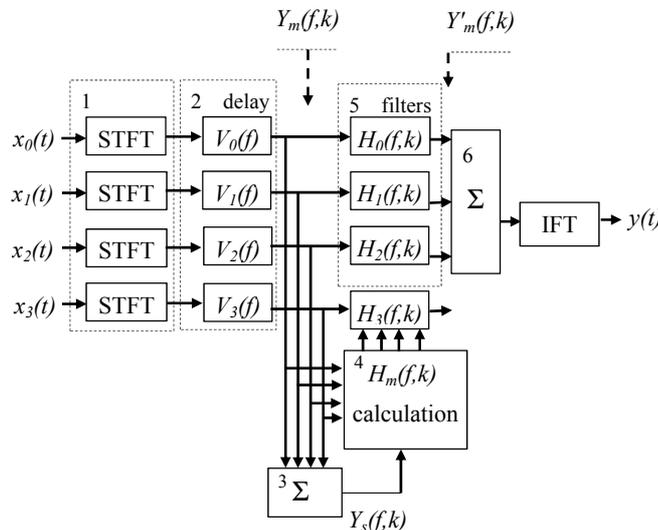


Figure 1. Block diagram of the proposed technique.

where $m = 1, M$ is the microphone index, f is the frequency bin index and k is the frame index (Figure 1, block 1). Then each signal $X_m(f, k)$ is delayed in block 2 by multiplication with the complex steering vector $V_m(f, \theta)$:

$$Y_m(f, k) = V_m(f, \theta) X_m(f, k), \tag{1}$$

where θ is the desired source direction. At the third step the frequency-domain FBF output is calculated (block 3) using the following equation:

$$Y_s(f, k) = \frac{1}{M} \sum_{m=1}^M Y_m(f, k). \tag{2}$$

We point here that the signals (1) and (2) represent an ideal situation where microphone characteristics are equal to each other. In practice there are fluctuations in microphone sensitivity and phase (both are generally frequency dependent). To compensate for these fluctuations, we first estimate the transfer function $H_m(f, k)$:

$$H_m(f, k) = \frac{\hat{\Phi}_{Y_m Y_s}(f, k)}{\hat{\Phi}_{Y_m Y_m}(f, k)}, \tag{3}$$

where $\hat{\Phi}_{Y_m Y_s}(f, k)$ and $\hat{\Phi}_{Y_m Y_m}(f, k)$ are the estimated cross spectra and the power spectra of the $Y_m(f, k)$ and $Y_s(f, k)$ signals. These estimates are obtained using exponential frame-by-frame averaging:

$$\begin{aligned} \hat{\Phi}_{Y_m Y_s}(f, k) &= (1 - \alpha) \hat{\Phi}_{Y_m Y_s}(f, k - 1) + \alpha Y_m(f, k) Y_s^*(f, k) \\ \hat{\Phi}_{Y_m Y_m}(f, k) &= (1 - \alpha) \hat{\Phi}_{Y_m Y_m}(f, k - 1) + \alpha Y_m(f, k) Y_m^*(f, k) \end{aligned} \tag{4}$$

where $\alpha = T/\tau$ is the time constant of the decaying, T and τ are the frame duration and adaptation time constant, respectively. The calculation of (3-4) is performed in block 4, Figure 1. Then the matched transfer function $H_m(f, k)$ is transferred to block 5 where the microphone signals are transformed as:

$$Y_m'(f, k) = H_m(f, k)Y_m(f, k). \quad (5)$$

After modification the new output spectrum is calculated in block 6 as follows:

$$Y_{out}(f, k) = \frac{1}{M} \sum_{m=1}^M Y_m'(f, k). \quad (6)$$

2.2 Analysis of the Proposed Algorithm (Magnitude and Phase Influence)

We should point out here that further on we omit the frame index k for the sake of simplicity. Consider again the MA directed perpendicularly to the line of microphones. Naturally the transfer function (3) is a complex function, i.e. we can write:

$$H_m(f) = |H_m(f)| \exp(j \arg(H_m(f))) \quad (7)$$

where $|H_m(f)|$ and $\arg(H_m(f))$ are the magnitude and the phase of the function, respectively. It is clear that generally both components (magnitude and phase) affect the results (5) and (6). Transforming $H_m(f)$ in (5) to “only the phase of $H_m(f)$ ” or “only the magnitude of $H_m(f)$ ” allows us to investigate these influences separately. Results these studies are presented below.

2.2.1 Phase Influence

The physical meaning of the influence of the phase can be easily explained as follows. Consider a 3-element equidistant linear MA. In that case the middle (second) microphone represents the phase center of the MA. If our MA is directed perpendicularly to the line of microphones, then every element of the steering vector is real valued and equal to 1. Accordingly, the MA output $Y_s(f)$ is a simple sum of three microphone output signals $X_m(f), m=1, 2, 3$ divided by 3. Now let the input signal be a harmonical plane wave from some direction θ . It is clear that in this case phase differences between $Y_s(f)$ and $X_m(f)$ are equal to $\varphi, 0, -\varphi$ for the first, second and third microphone, respectively. Moreover, exactly the same phase differences are obtained when we calculate cross-spectra (4). So for our 3-element MA the phase of $H_m(f)$ can be written as: $\arg(H_1(f)) = \varphi$; $\arg(H_2(f)) = 0$; $\arg(H_3(f)) = -\varphi$. Thus when we calculate (5) we provide additional rotation of the $X_1(f)$ and $X_3(f)$ signals to angles φ and $-\varphi$, i.e. phase differences are doubling. As a result, for the input signal frequency f we have the MA directivity pattern corresponding to the frequency $2f$: the width of the mainlobe and sidelobes are two times narrower, the sidelobes are smaller, etc. Of course, this conclusion is valid for any number of microphones. Figure 2 shows the resulting directivity patterns of an 8-microphone MA for a Conventional Delay & Sum FBF and the proposed method when only $H_m(f)$ phase was used in (5). The distance between the microphones was 5 cm, the input signal was a 2000 Hz harmonical plane wave.

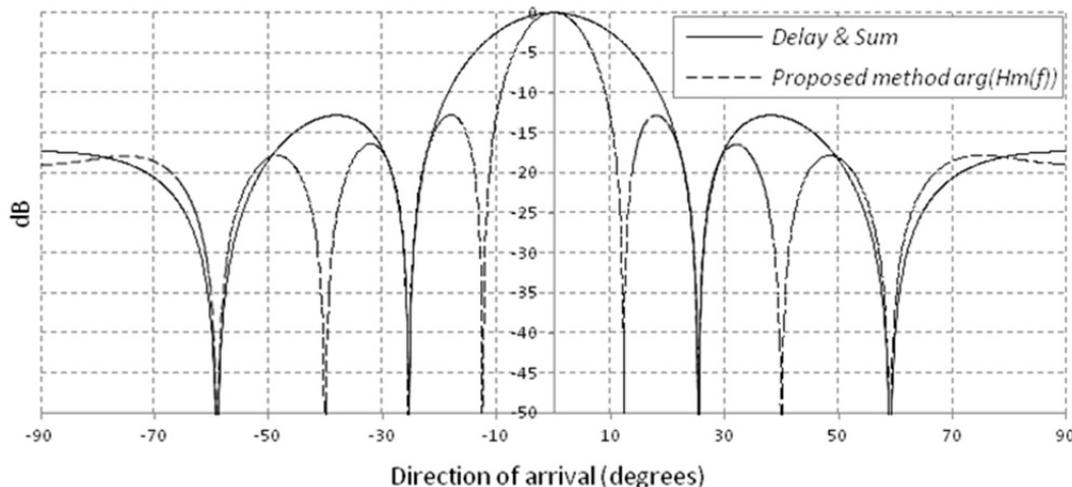


Figure 2. Directivity patterns for Delay & Sum and the proposed method for $\arg(H_m(f))$

It is clear that the dashed curve corresponds to the input signal with the frequency equal to 4000 Hz, while in reality the input signal frequency was 2000 Hz.

2.2.2 Magnitude Influence

If we use only the magnitude of $H_m(f)$ in (5), the result will be quite different. If the input signal is a plane wave from an angle θ , then the amplitude A_{out} of the output signal Y_{out} can be written as:

$$A_{out} = A_{in}D(f, \theta), \tag{8}$$

where A_{in} is the input signal amplitude and $D(f, \theta)$ is the magnitude directivity pattern of the MA. A close inspection of (5) and (6) shows that in this case

$$E(|H(f)|) = D(f, \theta), \tag{9}$$

where $E(\cdot)$ denotes the mathematical expectation. Rewriting (8) and (9) in terms of expectations we get:

$$E(A_{out}) = E(|H(f)|A_{in}D(f, \theta)) = E(A_{in}D^2(f, \theta)) \tag{10}$$

This means that the resulting directivity pattern is equal to the square of the initial one. These conclusions are confirmed by Figure 3 which shows the resulting directivity patterns for the same parameters of the microphone array and the input signal as in Figure 2 when only the magnitude of $H_m(f)$ is used in (5).

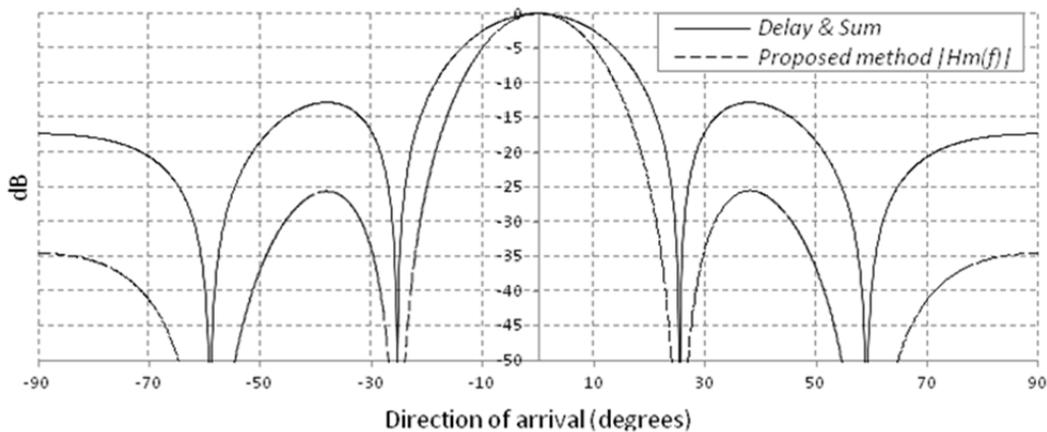


Figure 3. Directivity patterns for Delay & Sum and the proposed method for $|H_m(f)|$.

We can see that main lobe of the dashed curve is slightly narrower, and the level of sidelobes for the dashed curve on the dB scale is two times lower than that for the solid curve.

If we compare Figs. 2 and 3, we can conclude that the phase mainly reduces the width of the main lobe, while the magnitude has a greater effect on the amplitude of the sidelobes. Of course mainlobe width reduction looks very useful. Moreover our investigation have shown that the use of the complex $H_m(f)$ allows to get both mainlobe width and sidelobes reduction. However, the data in Figs. 2 and 3 were obtained using harmonic signals without interfering noise, i.e. in ideal conditions. On the other hand, our experiments showed that, first, the phase method works poorly in the presence of high level additive noise. Second, the use of the phase increases parasitic musical noise in the output speech signal. In the end we left the phase method for further research and focused on the investigation of (3) with magnitude in the numerator, i.e. everywhere below we assume that for our proposed method:

$$H_m(f, k) = \frac{|\hat{\Phi}_{Y_m Y_s}(f, k)|}{\hat{\Phi}_{Y_m Y_m}(f, k)}, \tag{11}$$

2.3 Analysis of the Proposed Algorithm for Coherent Noise

Consider coherent noise suppression with the presence of a target signal. The microphone output signal in this case equals interference plus speech signal:

$$Y_m(f) = S(f) + N(f, \theta). \quad (12)$$

where $Y_m(f)$ is the output microphone signal; $S(f)$ is target signal coming from the look direction; $N(f, \theta)$ is interference coming from the direction θ . Conventional FBF output can be written as:

$$Y_s(f) = \frac{1}{M} \sum_{m=1}^M Y_m(f) = S(f) + N(f, \theta)D(f, \theta). \quad (13)$$

In this case the transfer function (11) can be written as:

$$H_m(f, \theta) = \frac{\langle Y_s(f)Y_m^*(f) \rangle}{\langle Y_m^2(f) \rangle} = \frac{\langle (S(f) + N(f, \theta)D(f, \theta))(S(f) + N^*(f, \theta)) \rangle}{\langle (S(f) + N(f, \theta))^2 \rangle}. \quad (14)$$

where $\langle \cdot \rangle$ is the time averaging symbol. Hence we get:

$$H_m(f, \theta) = \frac{SIR(f, \theta) + D(f, \theta)}{SIR(f) + 1}, \quad (15)$$

where: $SIR(f, \theta) = \frac{\langle |S(f)|^2 \rangle}{\langle |N(f, \theta)|^2 \rangle}$ is the Signal-to-Interference ratio. If the target speech signal is stronger than the interference, we get $SIR(f, \theta) \gg 1 \Rightarrow H_m(f, \theta) \approx 1$, which means that the target signal and the noise come from the FBF output unchanged, and the SIR relation remains the same compared with FBF. In case there is no target signal, $SIR(f, \theta) = 0 \Rightarrow H_m(f, \theta) = D(f, \theta)$ and according to (5) we can draw the same conclusion as in (10), i.e. noise reduction $NR(f, \theta)$ is equal to:

$$NR(f, \theta) = H_m(f, \theta)D(f, \theta) = D^2(f, \theta) \quad (16)$$

Therefore, the proposed method doubles noise suppression in the dB scale. For an 8-element MA with 2 kHz interference, the level of the first sidelobe is equal to -13 dB. After the application of the proposed algorithm, the level is reduced to -26 dB.

2.4 Analysis of the Proposed Algorithm for Incoherent Noise

For independent microphone noise, we get:

$$Y_m(f) = S(f) + N_m(f), \quad (17)$$

where $N_m(f)$ is the noise on the m -th microphone such that $E(N_m(f)N_k(f)) = 0, \forall m \neq k$. In this case the FBF output can be expressed as:

$$Y_s(f) = \frac{1}{M} \sum_{m=1}^M Y_m(f) = S(f) + \frac{1}{M} \sum_{m=1}^M N_m(f) \quad (18)$$

In this case: $\langle Y_s(f)Y_m^*(f) \rangle = S(f) + \frac{1}{M} \langle N_m^2(f) \rangle$ and $\langle Y_m^2(f) \rangle = \langle S^2(f) \rangle + \langle N_m^2(f) \rangle$, hence:

$$H_m(f) = \frac{\langle Y_s(f)Y_m^*(f) \rangle}{\langle Y_m^2(f) \rangle} = \frac{S^2(f) + \frac{1}{M} \langle N_m^2(f) \rangle}{\langle S^2(f) \rangle + \langle N_m^2(f) \rangle} = \frac{SNR_m(f) + \frac{1}{M}}{SNR_m(f) + 1}. \quad (19)$$

Thus, for different SNR we get: if $SNR(f) \gg 1 \Rightarrow H_m(f) \approx 1$; if $SNR(f) \ll 1 \Rightarrow H_m(f) \approx 1/M$.

Therefore, in the presence of a speech signal, the simple FBF signal appears at the output. On the other hand, white noise in speech pauses (i.e. when $SNR(f) \ll 1$) is additionally suppressed by the factor of M . If $M=8$, suppression amounts to 9 dB (with total suppression of ≈ 18 dB).

3. Experiments and Results

3.1 Equipment and Parameters

In our experiments we investigated the proposed method using 8 equally spaced microphone array with 5 cm inter-microphone spacing and 35 cm total aperture length. Omni-directional microphones (Knowles Electronics) were used. The signals of the microphones were sampled with the frequency $F_s=16$ kHz. We used a standard Overlap-and-Add (OLA) technique with the frame length of 512 samples, 50% overlapping and a Hann window. The time constant α in (4) was about 0.5...0.25.

In practice, one of the detected issues was audible music noise caused by estimated spectrum random fluctuations (4). It is known that it is mostly low-frequency estimated spectrum fluctuations that cause this type of noise. To reduce the music noise, we propose the following transfer function limitations:

$$\begin{aligned} \text{if } (H(f) > H_{\max}) \text{ then } H(f) &= H_{\max} \\ \text{if } (H(f) < H_{\min}) \text{ then } H(f) &= H_{\min} \end{aligned} \quad (20)$$

We used fixed $H_{\max} = 1$, H_{\min} is set by user.

Estimated directivity patterns, as well as values of noise reduction (NR) and interference reduction (IR) were used for evaluating the performance of the proposed method. White Gaussian noise (WGN) and Wide-Band Gaussian noise (WBGN) with a 300-5000 Hz band were used as input signals. We investigated the method for three scenarios: incoherent, coherent and diffuse noises. Testing was conducted with both model signals and real signals recorded in an office. For comparison we also estimated the characteristics of a conventional FBF.

3.2 Results

3.2.1 Directivity Patterns

Directivity patterns (DPs) shown in Figures 2 and 3 were obtained using an artificial single-frequency harmonic signal and therefore do not provide complete information about the performance of the proposed method. For clarification, we conducted a series of experiments using the model and real WBGN as input signal for MA. Figure 4 shows the DPs when the input signal is artificial WBGN from the direction $\theta=0^\circ$. It can be seen that for our WBGN oscillating sidelobes are absent and the DPs gradually decrease as the MA angle moves away from the signal direction $\theta=0^\circ$. We can also see that the width of the mainlobe is large enough. This fact can be explained by the effect of low-frequency components of WBGN (as our antenna has a length of 35 cm, which provides poor directivity at frequencies below 1000 Hz). Using our method we obtain a similar behavior of the DP curve (dashed curve), but the mainlobe is narrower and the sidelobes are smaller.

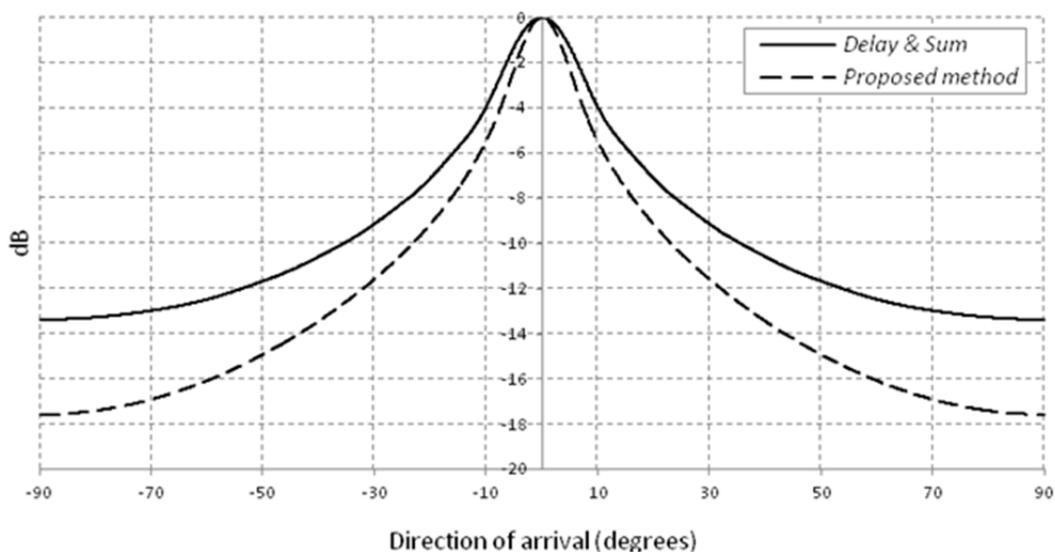


Figure 4. Directivity patterns for Delay & Sum and the proposed method for artificial wide-band Gaussian noise (bandwidth is 300 – 5000 Hz)

Figure 5 shows the DPs for the signal obtained in a real experiment: WBGN noise emitted by high-quality speakers located at a distance of 4 meters from the MA in an office with the reverberation constant equal to 860 ms.

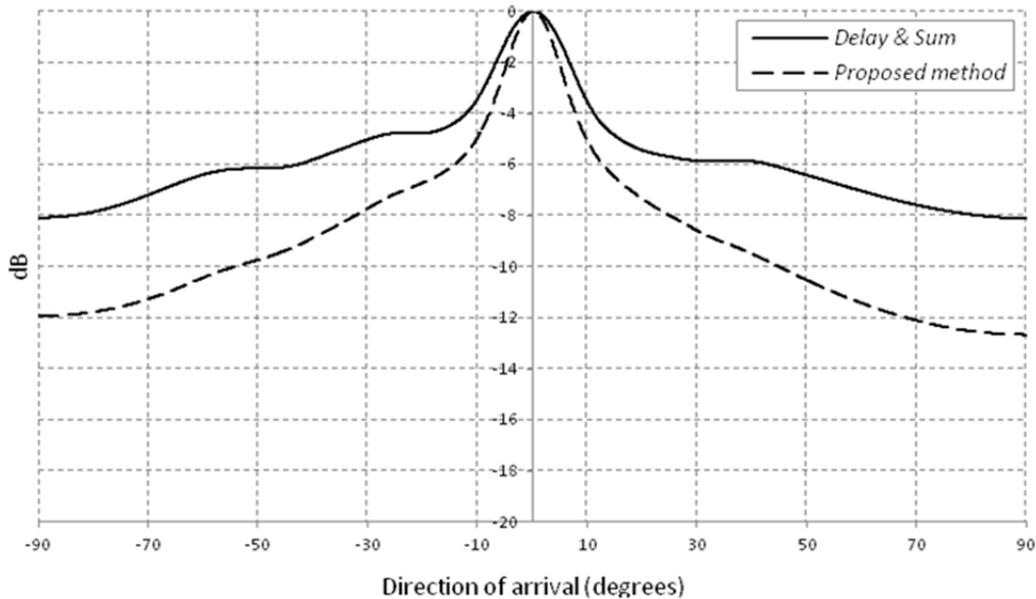


Figure 5. Directivity patterns for Delay & Sum and the proposed method for wide-band Gaussian noise recorded in a real office

We can see that despite the fact that the overall MA performance deteriorates, the proposed method has better characteristics than the conventional Delay & Sum FBF.

3.2.2 Noise Reduction as a Function of Frequency

We already mentioned the well-known fact that noise reduction in a MA depends on noise frequency. Consider Figure 6 which shows three power spectra. In this experiment, WGN emitted by high-quality speakers located 6 m from the MA and hidden from the MA by a screen simulated diffuse noise in the office.

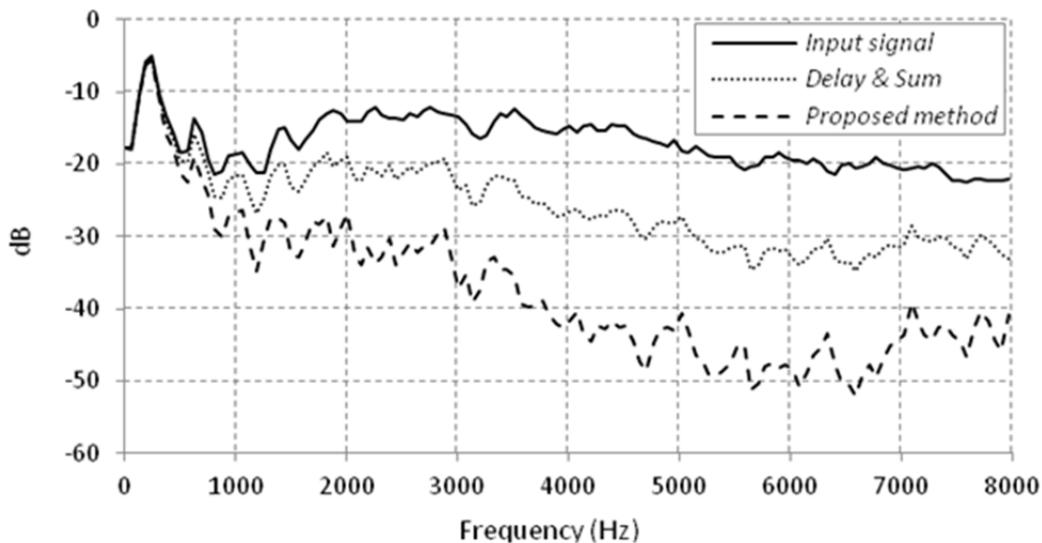


Figure 6. Power spectra of input and output signals of the MA for diffuse noise

The solid curve is the power spectrum of the single microphone signal; the dotted and dashed curves are the power spectra of the MA output signal when Delay & Sum and the proposed methods of BF are used,

respectively. The figure shows that all three curves are almost equal to each other in the low-frequency domain (0 – 500 Hz). Then the dotted and the dashed curves begin to go down, indicating that both (Delay & Sum and the proposed) algorithms start to suppress the diffuse noise. The lower the curves are in comparison with the solid one, the stronger the noise suppression. So Figure 6 clearly demonstrates the advantages of the proposed method compared to the Delay & Sum FBF. A slight convergence of the curves in the region above 6000 Hz can be explained, in our opinion, by the appearance of parasitic sidelobes with a high amplitude. We should point out that the data shown in Figure 6 are very useful as they allow us to calculate the level of diffuse noise reduction in the frequency bands. For example, let us denote the power spectrum of the single microphone signal (solid curve) as $P_{in}(f)$ and the power spectrum of the MA output signal for the proposed method as $P_{out}(f)$. Then noise reduction in the discrete frequency band (f_1, f_2) can be written as:

$$NR(\Delta f) = 10 \log_{10} \left(\sum_{f=f_1}^{f_2} P_{in}(f) \right) - 10 \log_{10} \left(\sum_{f=f_1}^{f_2} P_{out}(f) \right) \quad (20)$$

3.2.3 Noise Reduction for Diffuse Noise and for Interference

We calculated noise reduction for three types of noises: model incoherent noise, real diffuse noise and real interference. Model incoherent noise is a set of 8 independent artificial random sequences with Gaussian distribution, generated using the “WELL1024a” algorithm described by (Panneton *et al.* 2006). Diffuse noise is a real WGN recorded by the MA in an office and generated as explained in the description of Figure 6. Interference is also a real recorded WGN in an office coming to the MA from the angle of 40 degrees when the MA is oriented at $\theta = 0^\circ$. Estimations of NR and IR were provided in the frequency band 1000 – 8000 Hz.

Table 1. Estimated noise and interference reduction

Noise type	FBF	Proposed method
Incoherent noise	9 dB	18 dB
Coherent noise (40°)	10 dB	17 dB
Diffuse noise	6.9 dB	12.9 dB

The results presented in Table 1 show that the proposed method has considerable advantages compared to FBF.

4. Discussion

In this paper we proposed a novel frequency-domain alignment technique of speech enhancement for directional microphone array systems. The proposed technique improves the MA directivity pattern compared to the conventional Delay & Sum FBF: we show both analytically and experimentally that the proposed method narrows the mainlobe and reduces the sidelobes of the DP. Two modifications of the proposed method (“phase” and “magnitude” modifications) were considered and studied in detail. We performed a number of experiments using real (emitted and recorded in a real office) and simulated acoustic signals. Various types of signals and noises were used in the experiments: harmonic signals, white Gaussian noise, wide-band Gaussian noise (300 – 5000 Hz), etc. We achieved a good similarity of theoretical and experimental results. Our experiments showed that for an 8-microphone array it is possible to achieve noise reduction of up to 12.9 dB for diffuse noise and up to 18 dB for incoherent noise.

5. Conclusion

The proposed technique showed good results in the processing of acoustic data. Moreover, the technique, which involves a low-complexity algorithm, can be used in systems with a large number of microphones (for example, a two-dimensional microphone array) in real-time applications. We think that the technique can be combined with several types of microphone array processing algorithms, e.g. null-steering, post-filtering, etc. In a limited number of experiments we found that the combined use of the null-steering algorithm and the proposed method enhances the suppression in the null direction. We also found that the system consisting of our algorithm and a sequentially connected Zelinsk post-filter performs better than either of these components separately. However, we did not conduct detailed studies in this area. We should point out that musical noise is one of the major problems limiting the performance of the proposed method. Combination with other adaptive frequency-domain algorithms as well as musical noise suppression and detailed study of the performance of the phase and complex transfer function are a task for future work.

Acknowledgments

The work was financially supported by the Ministry of Education and Science of the Russian Federation

(Contract 14.575.21.0033, ID RFMEFI57514X0033) and by the Government of the Russian Federation, Grant 074-U01.

References

- Cao, Y., & Sridharan, S. (1994). Moody, Post-Microphone-Array Speech Enhancement with Adaptive Filters for Forensic Applications. *Proceedings of International Symposium on Speech, Image Processing and Neural Networks* (pp. 2543-255).
- Doclo, S., & Moonen, M. (2007). Superdirective Beamforming Robust Against Microphone Mismatch. *IEEE Trans. on Audio, Speech, and Signal Processing*, 15(2), 617-631.
- Brandstein, M., & Ward, D. B. (2001). *Microphone Arrays: Signal Processing Techniques and Applications*. Springer-Verlag.
- Fischer, S., & Simmer, K. U. (1995). An adaptive Microphone Array for Hands-Free Communication. *Proceedings of International Workshop on Acoustic Signal Enhancement* (pp. 1-4).
- Flanagan, J. L., Johnston, J. D., Zahn, R., & Elko, G. W. (1985). Computer-steered microphone arrays for sound transduction in large rooms. *J. Acoust. Soc. Amer.*, 78(5), 1508-1518.
- Frost, O. L. (1972). An algorithm for linearly constrained adaptive array processing. *Proceedings of IEEE*, 60, 926-935.
- Gannot, S., Burshtein, D., & Weinstein, E. (2001). Signal enhancement using beamforming and nonstationarity with applications to speech. *IEEE Trans. Signal Processing*, 49, 1614-1626.
- Griffiths, L. J., & Jim, C. W. (1982). Alternative approach to linearly constrained adaptive beamforming. *IEEE Transaction on Antennas Propagation*, AP-30, 27-34.
- Hoshuyama, O., Sugiyama, A., & Hirano, A. (1999). A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters. *IEEE Trans. Signal Processing*, 47, 2677-2684.
- Kellermann, W. (1991). A self-steering digital microphone array. *Proceeding of the Internat. Conference on Acoustics, Speech and Signal Processing* (pp. 3581-3584).
- Löllmann, H. W., & Vary, P. (2007). Post-filter design for superdirective beamformers with closely spaced microphones. *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (pp. 291-294).
- Mahieux, Y., Le Tourneur, G., Gilloire, A., Saliou, A., & Jullien, J. P. (1993). A microphone array for multimedia workstations. *Proceeding of the 3rd International Workshop on Acoustic Echo Control* (pp. 145-149).
- Ortega-García, J., & González-Rodríguez, J. (1996). Overview of speech enhancement techniques for automatic speaker recognition. *Proceedings of International Conference on Spoken Language Processing* (pp. 929-932).
- Panneton, F., L'Ecuyer, P., & Matsumoto, M. (2006). Improved Long-Period Generators Based on Linear Recurrences Modulo 2. *ACM Transactions on Mathematical Software*, 32(1), 1-16.
- Simmer, K. U., Bitzer, J., & Marro, C. (2001). Post-filtering techniques. In *Microphone arrays*, M. Brandstein, and D. Ward (Eds.), Springer-Verlag, Berlin Heidelberg New York, (ch. 3, pp. 39-60).
- Simmer, K. U., & Wasiljeff, A. (1992). Adaptive microphone arrays for noise suppression in the frequency domain. *Proceedings of Second Cost 229 Workshop on Adaptive Algorithms in Communications* (pp. 185-194).
- Stolbov, M., & Aleinik, S. (2014). Speech Enhancement with Microphone Array Using Frequency-Domain Alignment Technique. *Proceedings of 54th Conference on Audio Forensics Techniques, Technologies, and Practice* (pp. 1-5).
- VanVeen B. D., & Buckley, K. M. (1988). Beamforming: A versatile approach to spatial filtering. *IEEE Acoust., Speech, Signal Processing* (pp. 4-24).
- Widrow, B., & Stearns. S. D. (1985). *Adaptive Signal Processing*. New Jersey: Prentice-Hall, Inc.
- Zang, C., Tao, Y., & Hansen, J. H. L. (2010). Microphone array processing for distance speech capture: a probe study on whisper speech detection. *Proceedings of Asilomar Conference on Signals, Systems & Computers*, pp. 1707-1710.
- Zelinski, R. (1988). A microphone array with adaptive postfiltering for noise reduction in reverberant rooms.

Proceeding of the Internat. Conference on Acoustics, Speech and Signal Processing (pp. 2578-2581).

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).