International Journal of Computer Applications (0975 – 8887)
Volume 129 – No.11, November2015

Dual Microphone Beamforming Algorithm for Acoustic Signals

Pallavi Agrawal
Department of ECE,
Maulana Azad National Institute of Technology,
Bhopal (M. P.)-462051

Madhu Shandilya
Department of ECE,
Maulana Azad National Institute of Technology,
Bhopal (M. P.)-462051

ABSTRACT
This paper covers delay-and-sum beamformer and Filter and sum beamformer - Minimum Variance Distortion-less Response (MVDR) beamformer. Both the beamformers were simulated and tested in terms of noise source separation at various frequencies and computational complexity using MATLAB. Even though Generalized Side-Lobe Cancellers (GSC), Superdirectivity and Post-Filtering are also available. It actually covers two-sensor array beamforming which can be extended to multisensory array. MVDR beamformer gave better results as compared to delay-and-sum beamformer, as it adopts to noise condition and also improves the beamformer output, but has a higher computational complexity. Seeing the simulation results, MVDR proves to be a better option for implementing on smart phone applications.

Keywords
Beamformer, MVDR, array, noise

1. INTRODUCTION
Noise reduction speech de-reverberation and echo cancellation are some of the problems which arise whenever loudspeakers and microphones are kept in the same enclosure as in the case of the smart phone applications. Earlier the solution of this problem was to place the microphone closed to the desired speaker. But this also had several drawbacks as even though microphones were closed to vicinity of the speaker [1]. There used to be variations in the sound as the user used to move his head and mouth here and there relative to the microphone in which he was talking. So the most appropriate solution is to use beamforming microphone arrays [2], [3].

Array processing involves the use of multiple sensor arrays to transmit or receive a signal carried by propagating waves. Microphone arrays are set of microphones positioned to capture spatial information. They are used for acquisition and de-noising of speech signals [4]. Here two sensor array has been considered which uses beamforming for improving the quality of the speech signals. This concept can be further enhanced for multiple microphone array for speech signal enhancement and noise reduction. Microphone arrays utilizes spatial-temporal filtering methods [5], [6] which is very powerful as compared to conventional temporal filtering techniques in their ability to get rid of unwanted noise. Spatial diversity is represented by the acoustic impulse response from radiating source to sensors and these acoustic channels are modelled by FIR filters.

1.1 Beamforming-
Speech acquisition in adverse acoustic environment is corrupted due to background noise, room reverberation and far-end echo signals. Signal processing techniques used for reducing the background noise i.e. the noise from computer, fans, audio equipments etc. are known as acoustic noise reduction techniques [7-9]. Also the desired speech signals may be reverberated due to room acoustics. The main task of beamformer is to pick up signals coming from a particular predefined direction, called as steering direction while nullifying signals from other directions. The microphones can be placed linearly in broadside array type or end-fire array pattern, in a line, or an arc or 3-D manner. This influences the performance of the multi-microphone signal enhancement algorithms [10-12]. In order to avoid spatial aliasing the microphone spacing d_{min} which in turn depends on the maximum frequency f_{max} and the speed of sound c.

$$d_{\min} \langle \frac{\lambda_{\min}}{2} = \frac{c}{2 f_{\max}} \tag{1.1}$$

Now delay-and-sum beamformer and MVDR beamformer has been covered in context of noise reduction and signal enhancement.

1.2 Problem Description
For sensor array, the model assumes that each channel introduces delay and attenuation. Consider a scenario, wherein there are N sensors then output at time k can be written as:

$$y_n(k) = \alpha_n S[\mathrm{k-t-F_n}(\tau)] + \mathrm{V_n(k)} = x_n(k) + V_n(k), \tag{1.2}$$

Where α_n is the attenuation factor, n = 1, 2, …….., N, S(k) is the unknown source signal (It can be narrowband or wideband signal), $V_n(k)$ is the noise signal at the nth receiver, τ is the relative delay between sensor 1 and 2, and $F_n(\tau)$ is the delay between sensor 1 and n. It assumes here that τ and $F_n(\tau)$ are either known or can be estimated. The main aim is to now reduce the noise $V_n(k)$, which impinges on the desired source signal thereby improving signal-to-noise ratio (SNR).

2. Delay and Sum Beamformer-
The delay and sum beamformer [13] as shown in Fig. 1 consists of two basic processing steps. Firstly to calculate time-difference-of-arrival between the reference signal and the next arrived signal at the sensor. After time-shifting equation (1.2) becomes

$$\begin{aligned} y_{a,n}(k) &= y_n[k + F_n(\tau)] = \alpha_n \mathrm{S(k-t)} + \mathrm{V_{a,n}(k)} \\ &= \mathrm{x_{a,n}(k)} + \mathrm{V_{a,n}(k)},\ \textit{Where}\ \mathrm{n} = 1,2,......\mathrm{N} \end{aligned} \tag{1.3}$$

Where a is the aligned copy of sensor signal.

Second step is adding up all these time shifted signals. Thus the output Z(k) of delay and sum beamformer is:

$$Z(k) = \frac{1}{N}\sum_{n=1}^{N} y_{a,n}(k) = \alpha_s S(k-t) + \frac{1}{N} V_s(k) \tag{1.4}$$

$$\alpha_s = \frac{1}{N}\sum_{n=1}^{N}\alpha_n$$

$$V_s(k) = \sum_{n=1}^{N} V_{a,n}(k) = \sum_{n=1}^{N} V_n[k + F_n(\tau)] \tag{1.5}$$

Input SNR is given by:

$$SNR_i = \frac{\sigma_{x_1}^2}{\sigma_{v_1}^2} = \alpha_1^2 \frac{\sigma_s^2}{\sigma_{v_1}^2} \tag{1.6}$$

Where $\sigma_{x1}^2, \sigma_{v_1}^2$ and σ_s^2 are the variances of the signal $x_1(k), v_1(k)$ and $s(k)$ respectively.

Output SNR is given by-

$$SNR_o = N^2\alpha_s^2 \frac{E[s^2(k-t)]}{E[v_s^2(k)]} = N^2\alpha_s^2 \frac{\sigma_s^2}{\sigma_{V_s}^2} = \left(\sum_{n=1}^{N}\alpha_s\right)^2 \frac{\sigma_s^2}{\sigma_{V_s}^2} \tag{1.7}$$

$$\sigma_{Vs}^2 = E\left[\sum_{n=1}^{N} V_n[k + F_n(\tau)]^2\right] = \sum_{n=1}^{N}\sigma_{V_n}^2 V_n + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\rho_{v_i v_j} \tag{1.8}$$

In above equation first term is the variance of the noise signal, and second term is the cross-correlation between $V_i(k)$ and $V_j(k)$.

The delay and sum beamformer is of interest when output SNR is more than input SNR. Here N=2 has been used in microphones which are separated by a distance d.

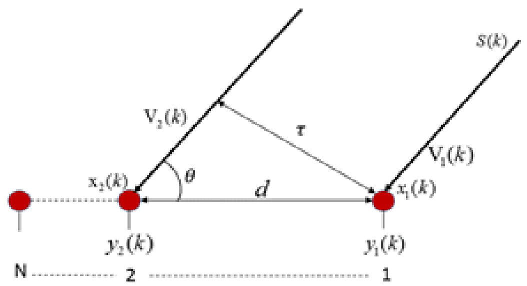


Fig. 1: Delay and sum beamformer

2.1 Minimum Variance Distortion-less Response Filter (MVDR)

MVDR [14], [15] refers to identifying a linear filter which minimizes variance at its output and at the same time the filter maintains a distortion-less response towards a specific input vector direction of interest. It is due to Capon and most widely used adaptive beamformer. It also minimizes the total output power compared to DAS. Mathematically, if r is random, zero mean, complex input vector of dimension L, $r \in C^L$, which is processed by a L-tap filter h, $h \in C^L$, then the filter output variance is $h^H Rh$, $R = E\{rr^H\}$ is the input autocorrelation matrix. $E\{\bullet\}$ = Expectation operator and H denote Hermitian. The MVDR filter minimizes $h^H Rh$ and at the same time satisfies $h^H\alpha = 1$, $\alpha =$ input signal vector direction to be protected. i.e. unity gain should be maintained at the target direction. Also minimum average output power of the beamformer per time frame i and per frequency bin k can be given by-

$$\min_h h^H Rh \text{ Subject to } h^H\alpha = \alpha_1 \tag{1.9}$$

This equation can be solved by Lagrange multipliers as

$$h_c = \alpha_1 \frac{R^{-1}\alpha}{\alpha^T R^{-1}\alpha}, C = \text{Capon} \tag{1.10}$$

Output of beamformer with MVDR filter is:

$$Z_c(k) = h_c^T y_n(k) = \alpha_1 \frac{\alpha^T R^{-1} y(k)}{\alpha^T R^{-1}\alpha} = x_1(k) + r_n(k), \tag{1.11}$$

Where $r_n(k) = \alpha_1 \frac{\alpha^T R^{-1} \mathrm{V_n}(k)}{\alpha^T R^{-1}\alpha}$ is the residual noise.

The output SNR with the capon filter can be evaluated as-

$$SNR(\mathrm{h}) = \alpha_1^2 \frac{\sigma_s^2}{\sigma_{rn}^2} = \frac{\sigma_n^2}{\sigma_{rn}^2} SNR, \sigma_{rn}^2 = E\left\{r_n^2(k)\right\} \tag{1.12}$$

The residual noise power is: $\sigma_{rn}^2 = \left(\alpha^T R^{-1}\alpha\right)^{-1}$ (1.13)

Output SNR can also be written as:

$$SNR(h) = \sigma_s^2\left(\alpha^T R^{-1}\alpha\right) \tag{1.14}$$

Comparing from DAS: $SNR(h) = N \cdot SNR$ which complies that capon filter degenerates to DAS when noise is uncorrelated having the same power. The main advantage of Capon beamformer is that it being adaptive in nature can adapt to changes in noisy environment and thus can be employed for maximum noise reduction. MVDR beamformer is a special case of the more generalized form of linearly constrained minimum variance filter (LCMV).

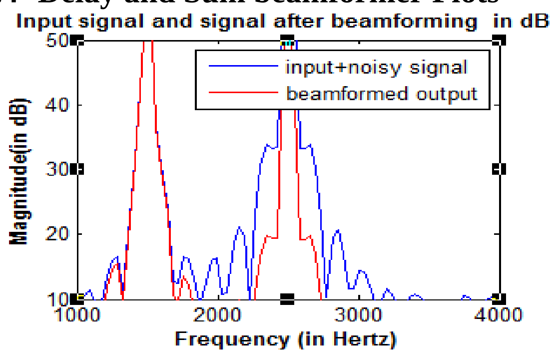
2.2 Generalised Side lobe Canceller (GSC)-

GSC [16-21] transforms the Linearly Constraint (LCMV) Beamformer from a constraint optimization problem into an unconstraint form. It can be used in adaptive operation for time-varying environments. GSC comprises of three blocks, a fixed beamformer which aligns the desired signal component, a blocking matrix to block the desired part of the signal so that only noise and interference remains and an adaptive noise canceller which removes the noise that leaks through side-lobes of fixed beamformer. In case the blocking matrix is not chosen properly, it can lead to signal leakage which does not block the speech signal properly.

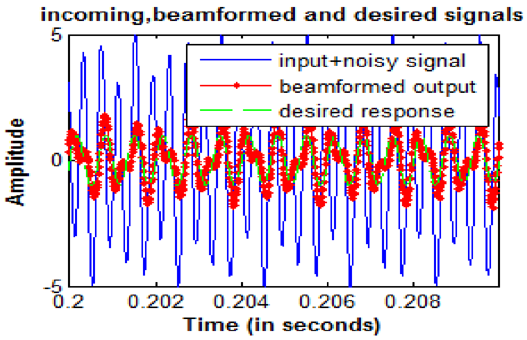
2.3 Performance Evaluation

MATLAB simulations were carried out for delay and sum beamformer and for MVDR beamformers. GSC beamformers were not chosen as they were expected to have similar response. These beamformers are tested for single noise and target source when noise was present. For the first test 1500 Hz frequency was chosen for signal coming from desired direction and 2500 Hz frequency was chosen with four times bigger amplitude for signal coming from undesired direction. Again the same was carried out for 300 Hz for desired direction and 600 Hz for undesired direction. Distance between two microphones was chosen to be 0.12 m, with total sample time of 1 sec and time frames of 20 msec. Both spectrum and time plots are shown.

2.4 Delay and Sum beamformer Plots-



(a)

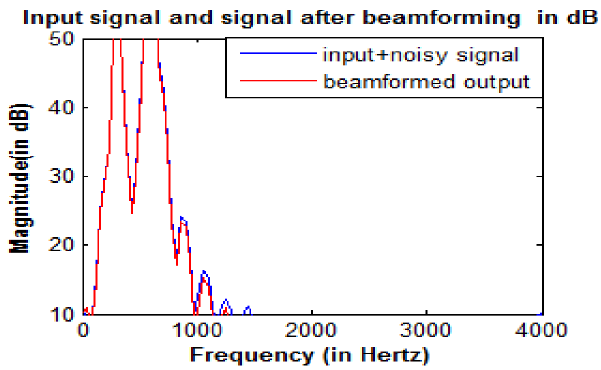


(b)

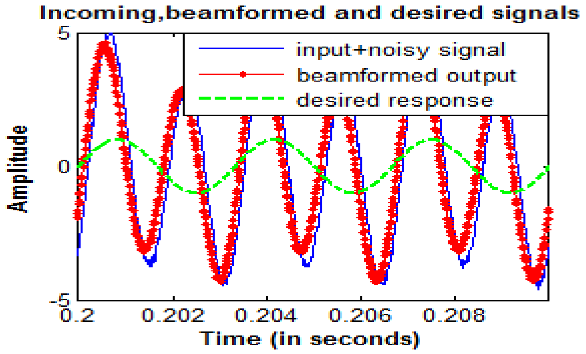
Fig. 2: Spectrum plot at 1500 Hz and 2500 Hz frequency for (a) Input signal and signal after beamforming in dB (b) Incoming, beamformed and desired signal

Fig. 2(a) shows the Spectrum plot for Input signal and signal after beamforming in dB and Fig. 2(b) shows the spectrum plot for Incoming, beamformed and desired signal at 1500 Hz and 2500 Hz frequency respectively. Frequency domain plot Fig. 2 shows that undesired sine wave has 4 times more amplitude. At frequency of 2500 Hz, the undesired sine wave is reduced by almost 20 dB. In time domain beamformed signal follows the desired signal.

Fig. 3(a) shows the Spectrum plot for Input signal and signal after beamforming in dB and Fig. 3(b) shows the spectrum plot for Incoming, beamformed and desired signal at 300 Hz and 600 Hz frequency respectively. As shown in Fig. 3, the plots at 300 Hz and 600 Hz, it is seen that at low frequencies, results are not satisfying. Beamformed signal did not suppress the unwanted signal significantly. Thus DAS performs well for some frequencies and only at some angles but overall its performance is quite poor.



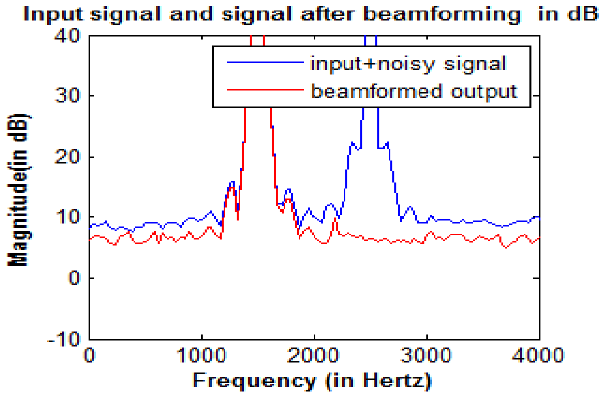
(a)



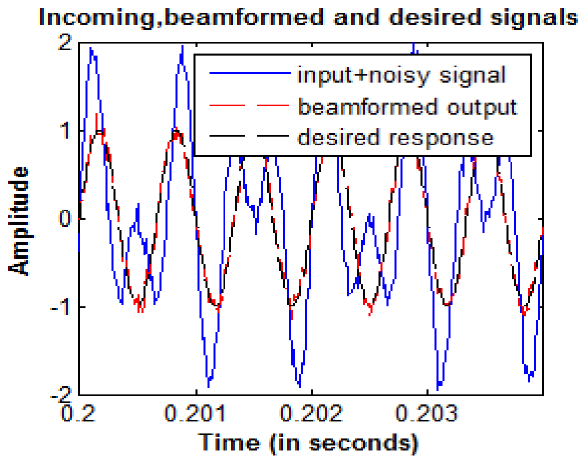
(b)

Fig. 3: Spectrum plot at 300 Hz and 600 Hz frequency for (a) Input signal and signal after beamforming in dB (b) Incoming, beamformed and desired signal

2.5 MVDR Beamformer Plots-



(a)

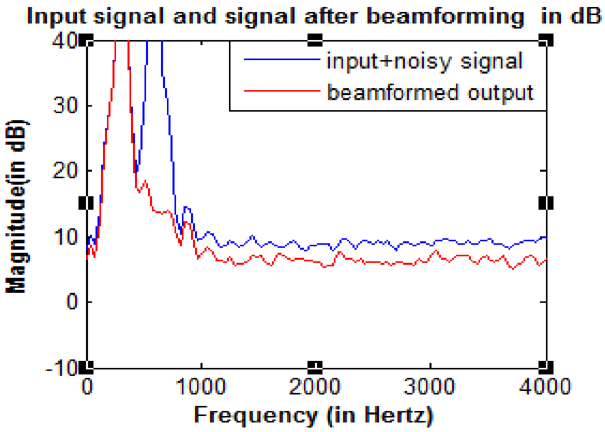


(b)

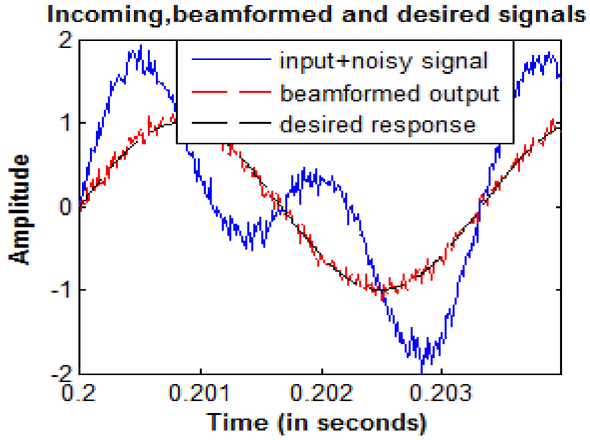
Fig. 4: Spectrum plot at 1500 Hz and 2500 Hz frequency for (a) Input signal and signal after beamforming in dB (b) Incoming, beamformed and desired signal

Fig. 4(a) shows the Spectrum plot for Input signal and signal after beamforming in dB and Fig. 4(b) shows the spectrum plot for Incoming, beamformed and desired signal at 1500 Hz and 2500 Hz frequency respectively. Spectrum plot as shown in Fig. 4 clearly depicts that undesired signal at frequency 2500 Hz, even though had 4 times more amplitude and coming from 400 direction, is suppressed as expected as compared to frequency 1500 Hz. It filtered the white noise by

3dB and removed noise source perfectly. Also in time domain the reconstructed signal follows the desired signal.



(a)



(b)

Fig. 5: Spectrum plot at 300 Hz and 600 Hz frequency for (a) Input signal and signal after beamforming in dB (b) Incoming, beamformed and desired signal

Similar results were seen at Fig. 5, at 300 Hz and 600 Hz frequencies as against DAS, which did not suppress the unwanted frequencies at low frequency bands.

3. CONCLUSION

In this paper a detailed description of two microphone array beamformer was described and in particular DAS and MVDR beamformer. This can be easily extended for evaluating the other existing beamforming techniques. As seen from the MATLAB simulations, it is very clear that DAS beamformer operated in a limited frequency range whereas MVDR operated satisfactorily in low and high frequencies and suppressed signals coming from unwanted directions. But MVDR requires high degree of complexity, still it outperforms DAS, as it steers properly at look direction and blocks noise from all other directions. Also MVDR covers full audio range from 300-3400 Hz.

4. REFERENCES

[1] Carroll and Heiser, G. 2010. An analysis of power consumption in a smartphone, in Proceedings of the 2010 USENIX conference on USENIX annual technical conference, USENIXATC'10. Berkeley, CA, USA: USENIX Association, 2010, pp. 21. [Online] Available: http://dl.acm.org/citation.cfm?id=1855840.1855861.

[2] Brandstein, M. and Ward, D. 2001. Microphone Arrays: Signal Processing Techniques and Applications, Springer-Verlag (Chapter 2), 2001.

[3] Elko, G. W. 2000. Superdirectional microphone arrays, In Acoustic Signal Processing for Telecommunication, Kluwer Academic Publishers, 181–235 (Chapter 10), 2000.

[4] Cohen, I., Benesty, J. and Gannot, S. 2010. Speech processing in modern communication: challenges and perspectives, Springer topics in signal processing, Springer-Verlag Berlin Heidelberg, 2010.

[5] Veen, B. V. and Buckley, K. 1988. Beamforming: a versatile approach to spatial filtering, ASSP Magazine (Apr. 1988), IEEE, vol. 5, no. 2, 4–24.

[6] Gannot S. and Dvorkind, T. G. 2006. Microphone Array Speaker Localizers Using Spatial-Temporal Information, EURASIP Journal on Applied Signal Processing, Hindawi Publishing Corporation, vol. 2006, Article ID 59625, 1–17, 2006.

[7] Puder, H. 2009. Acoustic noise control: An overview of several methods based on applications in hearing aids, in IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PacRim 2009), 871-876.

[8] Hansler, E. and Schimdt, G. 2006. Topics in acoustic echo and noise control, Springer-Verlag Berlin Heidelberg, 2006.

[9] Levitt, H. 2001. Noise reduction in hearing aids: a review, Journal of Rehabilitation Research and Development (2001), vol. 38, no. 1, 111–121.

[10] Luo F-L. and Uvacek, B. 2002. The Combination Of Binaural Processing With Adaptive Beamforming for Dual Microphones in Speech Communications, 1st IEEE International Conference on Circuits and Systems for Communications (ICCSC 2002), 436 – 439.

[11] Chen, Y., Gong, Q. and Zhang, T. 2013. Research on the Amplitude Mismatch in Dual-microphone Device, in Proceedings of International Conference on Complex Medical Engineering (ICME) (May 25 - 28, 2013), Beijing, China, 433-438.

[12] Chen, J., Phua, K., Shue, L., and Sun, H. 2009. Performance evaluation of adaptive dual microphone systems, Speech Communication (Dec. 2009), vol. 51, no. 12, 1180–1193.

[13] Greensted. Delay sum beamforming. [Online]. Available: http://www.labbookpages.co.uk/audio/beamforming/delaySum.html.

[14] Demba E. B., Florencio D. and Zhang, C. 2007. Enhanced MVDR Beamforming for Arrays of Directional Microphones, in proceedings ICME (July 2007), IEEE.

[15] Richmond, C. D. 2006. Adaptive Array Detection, Estimation and Beamforming, Workshop on Stochastic Eigen-Analysis and its Applications (July 2006), MIT Lincoln Laboratory.

[16] Cohen, I. 2003. Analysis of Two-Channel Generalized Sidelobe Canceller (GSC) With Post-Filtering, IEEE Transactions on Speech And Audio Processing (Nov. 2003), vol. 11, no. 6, 684-699.

[17] Lee, Y. and Wu, W-R. 2005. A Robust Adaptive Generalized Sidelobe Canceller with Decision Feedback, IEEE Transactions on Antennas and Propagation (Nov. 2005), vol. 53, no. 11, 3822-3832.

[18] Glentis, G-O. 2003. Implementation of Adaptive Generalized Sidelobe Cancellers Using Efficient Complex Valued Arithmetic, Int. J. Appl. Math. Comput. Sci. (2003), vol. 13, no. 4, 549–566.

[19] Yu, S-J. and Ueng, F-B. 2000. Blind adaptive beamforming based on generalized sidelobe canceller, Signal Processing (Dec. 2000), vol. 80, Issue 12, 2497–2506.

[20] Townsend, P. 2009. Enhancements to the Generalized Sidelobe Canceller for Audio Beamforming in an Immersive Environment. University of Kentucky Master's Theses, Paper 645. http://uknowledge.uky.edu/gradschool_theses/645.

[21] Talmon, R., Cohen, I. and Gannot, S. 2009. Convolutive Transfer Function Generalized Sidelobe Canceler, IEEE Transactions on Audio, Speech, and Language Processing (Sept. 2009), vol. 17, no. 7, 1420-1434.