

Survey of Spatial Audio Enhancement for Digital Hearing Aid

Pradeep M N¹ and Udayashankara V²

¹Research Scholar, Dept. of Electronics and communication Engineering,
Sri Siddhartha Institute of Technolgy, Tumakuru, India
pradeepmn81@gmail.com

²Prof. and Head, Dept. of Electronics and Instrumentation Engineering,
Sri Jayachamarajendra College of Engineering, Mysuru, India
v_udayashankara@sjce.ac.in

Abstract— A multi microphone array technique for hearing aid noise reduction is described. This study evaluates a spatial-filtering algorithm as a method to improve speech reception for monaural and binaural hearing impaired users in reverberant environments with multiple noise sources. The system has to filter sounds using phase differences between two microphones situated some distance apart in a behind-the-ear hearing-aid capsule. The approach is to linearly combine microphone signals after amplitude scaling and phase shifting. This approach is linear and time-invariant and can be used to generate known directional responses. Such approaches are generally referred to as beamforming algorithms. Beamformers can be implemented in analog, multiport, microphone circuitry to produce directional microphones. These techniques help to filter sounds using phase differences between microphones and provide improvements in speech signal across listening conditions compared with the omnidirectional response also spatial filtering can improve speech reception for hearing impaired users even in highly reverberant conditions with multiple noise sources.

Index Terms— Spatial filtering, Adaptive Beamformer, Multi-Channel Wiener Filtering (MWF), Blind Source Separation (BSS), Computational Auditory Scene Analysis (CASA), Minimum Variance Distortionless Response (MVDR), Generalized Side-lobe Canceler (GSC), Adaptive Noise Canceler (ANC).

I. INTRODUCTION

Noise jammer techniques in hearing aids are essential for hearing impaired persons to enhance speech quality in background noise (e.g. traffic, cocktail party situation). Many hearing aids today have more than one microphone, enabling the use of multi-microphone speech enhancement techniques [1]. When compared with single-microphone techniques, which can only use spectral and temporal information, multi-microphone techniques can additionally exploit the spatial information [beamforming] of the sound sources. This actually results in a better performance, especially when the speech and the noise sources are spatially separated. Most hearing impaired patients have a hearing loss at both ears; they are fitted with a hearing aid at each ear. In a so-called bilateral system, no cooperation between the hearing aids takes place. Current noise reduction techniques in bilateral hearing aids are not designed to preserve the binaural localization cues, i.e. the

Interaural Time Difference (ITD) and the Interaural Level Difference (ILD) [2]. These binaural cues play an important role in sound localisation and speech segregation in noisy environments [3, 4, 5]. In order to achieve true binaural processing, both hearing aids need to cooperate with each other, e.g. through a wireless link, such that a binaural hearing aid can be considered a simple acoustic sensor network. The objective of a binaural signal enhancement technique then is not only to selectively extract the useful speech signal and suppress background noise, but also to preserve the binaural cues of the sound sources, so as to preserve the auditory impression of the acoustic scene and exploit the binaural hearing advantage.

Section II gives an overview of several multi-microphone noise reduction techniques for hearing aids.

II. OVERVIEW OF NOISE REDUCTION TECHNIQUES

This paper gives the survey of some of the multi-microphone speech enhancement techniques and its advantages and limitations for monaural and binaural hearing aids, based on Fixed and Adaptive Beamforming, Multi-channel Wiener Filtering (MWF), Blind Source Separation (BSS) and Computational Auditory Scene Analysis (CASA). Each class of techniques has its own advantages and limitations.

A. Fixed Beamforming

Fixed Beamformers combine the microphone signals using a time-invariant filter-and-sum operation and are hence data-independent. The objective of a Fixed Beamformer is to obtain spatial focusing on the desired speech source, thereby reducing background noise not coming from the direction of the speech source. Different types of Fixed Beamformers exist, e.g. delay-and-sumbeamforming, superdirective Beamforming, differential microphone arrays and frequency-invariant Beamforming [6, 7]. For the design of Fixed Beamformers, the direction of the speech source and the complete microphone configuration need to be known. Hence, Fixed Beamformers have mainly been used for monaural hearing aids [8, 9], although for binaural hearing aids Fixed Beamforming techniques have also been proposed that aim to combine spatial selectivity and noise reduction with the preservation of the binaural cues of the speech source [10, 11].

B. Adaptive Beamforming

In practice, since the background noise is unknown and can change both spectrally and spatially, information about the noise field has to be adaptively estimated. Adaptive Beamformer will combine the spatial focusing of Fixed Beamformer with adaptive noise suppression [12]. Hence, they generally exhibit a higher noise reduction performance than Fixed Beamformers.

C. In a Minimum Variance Distortionless Response (MVDR) Beamformer

[13, 14], the energy of the output signal is minimised under the constraint that signals arriving from the assumed direction of the desired speech source are processed without distortion. A widely studied adaptive implementation of this Beamformer is the Generalised Sidelobe Canceler (GSC) [15]. The standard GSC consists of a spatial pre-processor, i.e. a Fixed Beamformer and a blocking matrix, combined with a (multichannel) adaptive noise canceler (ANC). The Fixed Beamformer provides a spatial focus on the speech source, creating a so-called speech reference; the blocking matrix steers nulls in the direction of the speech source, creating so-called noise references; and the ANC eliminates the noise components in the speech reference that are correlated with the noise references. Due to room reverberation, microphone mismatch and look direction error, speech components may however leak into the noise references of the standard GSC, giving rise to speech distortion and possibly signal cancellation. Several techniques have been proposed to limit the speech distortion resulting from this speech leakage, by reducing the speech leakage components in the noise references [16, 17] and by limiting the distorting effect of the remaining speech leakage components [18, 19]. The GSC or one of its more robust variants is a widely used multi-microphone noise reduction technique for monaural hearing aids with an endfire microphone array configuration. In an effort to combine adaptive noise reduction with binaural processing, adaptive Beamforming techniques producing a binaural output signal have also been proposed [20, 21].

In [22] the frequency spectrum is divided into a low-pass and a high-pass portion, and the low-pass portion is passed through unaltered in order to preserve the ITD cues of the speech source, while adaptive noise reduction is performed only for the high-pass portion. Other techniques restrict the preservation of the binaural cues to an angular region around the frontal direction, while reducing background noise for the other angles [23].

Multi-channel Wiener filtering (MWF) In [24] an MWF technique has been proposed that produces a minimum mean square error (MMSE) estimate of the desired speech component in one of the microphone signals, hence simultaneously performing noise reduction and limiting speech distortion. In addition, the MWF is able to take speech distortion into account in its optimisation criterion, resulting in the speech distortion weighted MWF (SDW-MWF). The SDW-MWF is uniquely based on estimates of the second-order statistics of the recorded speech signal and the noise signal, and hence requires a method that determines time-frequency regions where the desired source is dominant and time-frequency regions where the interference is dominant. The SDW-MWF has been successfully applied as a speech enhancement technique in monaural multi-microphone hearing aids [25, 26].

Since the SDW-MWF produces an estimate of the speech component in the microphone signals and does not make any assumptions regarding the microphone configuration and the room impulse responses relating the speech source and the microphones, it is obviously well suited to combine noise reduction with binaural processing. It was shown in [27, 28] that the binaural MWF perfectly preserves the binaural cues of the speech component, but undesirably changes the noise cues to those of the speech component. Several extensions have been proposed to preserve the binaural cues of both the speech and the noise component, either by partial noise estimation or by extending the MWF cost function with terms related to the ITD, ILD or interaural transfer function [29]. In addition, recently a distributed version of the binaural SDW-MWF has been presented [30].

D. Blind Source Separation (BSS)

The signals received at the microphones can essentially be considered a mixture of all sound sources, filtered by the respective room impulse responses between the sound source and the microphones. The goal of BSS is to recover all original signals. Many BSS algorithms exploit the independence and the non-Gaussianity of the sources, enabling the use of e.g. independent component analysis (ICA) techniques [31]. While time-domain ICA-based techniques are well suited to solve the instantaneous mixing problem, they are not able to address the convolutive mixture problem encountered in typical reverberant environments. By considering the BSS problem in the frequency-domain, the convolutive mixing problem can be transformed into an instantaneous mixing problem for each frequency bin. An inherent permutation and amplitude/phase scaling problem however occurs in frequency-domain BSS approaches, for which several solutions have been proposed [31, 32]. Nevertheless, due to its computational complexity the use of BSS techniques in hearing aids has found only limited practical interest.

E. Computational Auditory Scene Analysis

Algorithms based on computational auditory scene analysis (CASA) aim to perform sound segregation by modeling the human auditory perceptual processing [05, 33]. A typical CASA model involves two stages. In the first stage, a time-frequency representation of the incoming mixture of signals is generated, e.g. using short-time Fourier transform (STFT) processing or by incorporating a more advanced cochlear model. In the second stage, the resulting time-frequency elements are grouped into separate perceptual streams based on distinctive perceptual cues. As summarised in [34], some cues characterise the monaural acoustic properties of the sources, such as common pitch, amplitude modulation, onset [35]. In addition, for a binaural system sound sources can also be distinguished based on their spatial direction information, e.g. ITD, ILD and interaural envelope difference [36]. A gain factor is applied to each time-frequency element, such that regions dominated by the desired sound stream receive a high gain and regions dominated by other streams receive a low gain.

III. CONCLUSION

This paper reviewed several multi-microphone beamforming techniques that can be used for noise reduction in monaural and binaural hearing aids. For monaural noise reduction technique the transfer-function GSC and the multi-channel Wiener filter are focused, which are derived from different cost functions but are in fact closely related. For binaural noise reduction several extensions of the standard binaural MWF that aim to preserve the binaural cues of the speech and the noise sources are focused.

REFERENCES

- [1] V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, and U. Rass. Signal processing in high-end hearing aids: State of the art, challenges, and future trends. *EURASIP Journal on Applied Signal Processing*, 2005(18):2915–2929, 2005.
- [2] T. Van den Bogaert, T. J. Klasen, M. Moonen, L. Van Deun, and J. Wouters. Horizontal localisation with bilateral hearing aids: without is better than with. *Journal of the Acoustical Society of America*, 119(1):515–526, January 2006.
- [3] R. Beutelmann and T. Brand. Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 120(1):331–342, July 2006.
- [4] J. Blauert. *Spatial Hearing: The Psychophysics of Human Sound Localisation*. MIT Press, Cambridge MA, USA, 1983.
- [5] A. S. Bregman. *Auditory Scene Analysis*. MIT Press, Cambridge MA, USA, 1990.
- [6] J. Bitzer and K. U. Simmer. Superdirective Microphone Arrays, chapter 2 in “Microphone Arrays: Signal Processing Techniques and Applications” (Brandstein, M. S. and Ward, D. B., Eds.), pages 19–38. Springer-Verlag, May 2001.
- [7] S. Doclo and M. Moonen. Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics. *IEEE Trans. Signal Processing*, 51(10):2511–2526, October 2003.
- [8] J. M. Kates. Superdirective arrays for hearing aids. *Journal of the Acoustical Society of America*, 94(4):1930–1933, October 1993.
- [9] J. M. Kates and M. R. Weiss. A comparison of hearing-aid array-processing techniques. *Journal of the Acoustical Society of America*, 99(5):3138–3148, May 1996.
- [10] M. R. Bai and C. Lin. Microphone array signal processing with application in threedimensional hearing. *Journal of the Acoustical Society of America*, 117(4):2112–2121, April 2005.
- [11] J.G. Desloge, W.M. Rabinowitz, and P.M. Zurek. Microphone-array hearing aids with binaural output—Part I: Fixed-processing systems. *IEEE Trans. Speech and Audio Processing*, 5(6):529–542, November 1997.
- [12] F. L. Luo, J. Y. Yang, C. Pavlovic, and A. Nehorai. Adaptive null-forming scheme in digital hearing aids. *IEEE Trans. Signal Processing*, 50(7):1583–1590, 2002.
- [13] J. Capon. High resolution frequency-wavenumber spectrum analysis. *Proc. IEEE*, 57(7):1408–1418, August 1969.
- [14] O. L. Frost III. An Algorithm for Linearly Constrained Adaptive Array Processing. *Proc. IEEE*, 60(8):926–935, August 1972.
- [15] L. J. Griffiths and C.W. Jim. An alternative approach to linearly constrained adaptive beamforming. *IEEE Trans. Antennas Propagat.*, 30(1):27–34, January 1982.
- [16] I. Claesson and S. Nordholm. A Spatial Filtering Approach to Robust Adaptive Beaming. *IEEE Trans. Antennas Propagat.*, 40(9):1093–1096, September 1992.
- [17] I. Cohen, S. Gannot, and B. Berdugo. An integrated real-time beamforming and postfiltering system for non-stationary noise environments. *EURASIP Journal on Applied Signal Processing*, (11):1064–1073, October 2003.
- [18] H. Cox, R. M. Zeskind, and M. M. Owen. Robust Adaptive Beamforming. *IEEE Trans. Acoust., Speech, Signal Processing*, 35(10):1365–1376, October 1987.
- [19] S. Doclo, A. Spriet, J. Wouters, and M. Moonen. Frequency-Domain Criterion for Speech Distortion Weighted Multichannel Wiener Filter for Robust Noise Reduction. *Speech Communication*, 49(7–8):636–656, Jul.-Aug. 2007.
- [20] J. E. Greenberg and P. M. Zurek. Evaluation of an adaptive beamforming method for hearing aids. *Journal of the Acoustical Society of America*, 91(3):1662–1676, March 1992.
- [21] M. Kompis and N. Dillier. Performance of an adaptive beamforming noise reduction scheme for hearing aid applications. *Journal of the Acoustical Society of America*, 109(3):1123–1143, 2001.
- [22] D.P. Welker, J.E. Greenberg, J.G. Desloge, and P.M. Zurek. Microphone-array hearing aids with binaural output—Part II: A two-microphone adaptive system. *IEEE Trans. Speech and Audio Processing*, 5(6):543–551, November 1997.
- [23] R. Nishimura, Y. Suzuki, and F. Asano. A new adaptive binaural microphone array system using a weighted least squares algorithm. In *Proc. IEEE Int. Conf. Acoustics, REFERENCES 33 Speech, and Signal Processing (ICASSP)*, pages 1925–1928, Orlando FL, USA, May 2002.
- [24] S. Doclo and M. Moonen. GSVD-based optimal filtering for single and multimicrophone speech enhancement. *IEEE Trans. Signal Processing*, 50(9):2230–2244, September 2002.
- [25] J.-B. Maj, M. Moonen, and J. Wouters. SVD-based optimal filtering technique for noise reduction in hearing aids using two microphones. *EURASIP Journal on Applied Signal Processing*, 2002(4):432–443, April 2002.
- [26] A. Spriet, M. Moonen, and J. Wouters. Robustness Analysis of Multi-channel Wiener Filtering and Generalized Sidelobe Cancellation for Multi-microphone Noise Reduction in Hearing Aid Applications. *IEEE Trans. Speech and Audio Processing*, 13(4):487–503, July 2005.
- [27] S. Doclo, T. J. Klasen, T. Van den Bogaert, J. Wouters, and M. Moonen. Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions. In *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Paris, France, September 2006.

- [28] T.J. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters. Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues. *IEEE Trans. Signal Processing*, 55(4):1579–1585, April 2007.
- [29] S. Doclo, R. Dong, T. J. Klasen, J. Wouters, S. Haykin, and M. Moonen. Extension of the multi-channel Wiener filter with localisation cues for noise reduction in binaural hearing aids. In *Proc. Int. Workshop on Acoustic Echo and Noise Control (IWAENC)*, pages 221–224, Eindhoven, The Netherlands, September 2005.
- [30] S. Doclo, T. Van den Bogaert, J. Wouters, and M. Moonen. Comparison of reducedbandwidth MWF-based noise reduction algorithms for binaural hearing aids. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WAS-PAA)*, pages 223–226, New Paltz NY, USA, October 2007.
- [31] A. J. Bell and T. J. Sejnowski. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1004–1034, 1995.
- [32] H. Buchner, R. Aichner, and W. Kellermann. A generalization of Blind Source Separation algorithms for convolutive mixtures based on second-order statistics. *IEEE Trans. Speech and Audio Processing*, 13(1):120–134, January 2005.
- [33] S. Makino. *Adaptive Signal Processing: Applications to Real-World Problems*, chapter Blind source separation of convolutive mixtures of speech. Springer-Verlag, 2003.
- [34] G. J. Brown and D. L. Wang. *Speech Enhancement*, chapter Separation of speech by computational auditory scene analysis, pages 371–402. Springer-Verlag, 2005.
- [35] G. N. Hu and D. L. Wang. Monaural speech segregation based on pitch tracking and amplitude modulation. *IEEE Trans. on Neural Networks*, 15(5):1135–1150, September 2004.
- [36] B. Kollmeier and R. Koch. Speech enhancement based on physiological and psychoacoustical modals of modulation perception and binaural interaction. *Journal of the Acoustical Society of America*, 95(3):1593–1602, March 1994.