Binaural Beamformer: An early Proof of Concept for Wearables Audio Devices

Stéphane Dedieu*[1], Thomas Padois†[2], and Jérémie Voix‡[2]
[1]Bloo Audio Inc, Ottawa (ON), Canada
[2]ÉTS, Université du Québec, Montréal (QC), Canada

1 Introduction

Binaural Beamforming, using signals from microphone arrays from both ears and respecting 3D localization, is a promising approach to improve speech intelligibility in the presence of interfering sound sources for all wearable audio devices. Despite the imposed small size of the microphone arrays, it offers strong interfering noise reduction potential due to the large inter-array distance and the natural attenuation provided by the head. This study is a proof of concept conducted in 2015, for a technology that several commercial products now casually feature (hearing aids, digital hearing protectors, hearables, etc.)

2 Method

A pair of small identical 3-microphone arrays is designed for a "Behind The Ear (BTE)" headset. The array geometry is visible in Figure 1. In one simple implementation of binaural beamforming, where left and right arrays are connected (through wired or wireless link), the monaural beamformer output is filtered with Head-Related Transfer Functions (HRTFs) to preserve binaural localization cues. The monaural beamforming is based on a fixed Minimum Variance Distortionless Response (MVDR) algorithm for robustness.

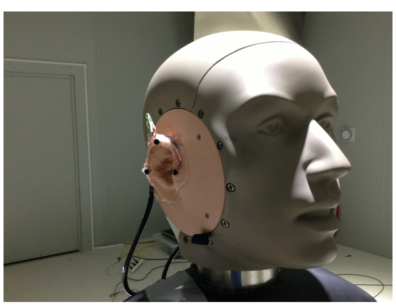

Figure 1: Microphone array on artificial head's right ear

2.1 MVDR beamformer

MVDR minimum-variance distortion-less response beamformer is built in the frequency range of [150-7000] Hz and a set of direction of interests: "look directions" from -90 to 90 degrees, with 30 degree increments. Assuming that $d_0(\omega)$ is the look direction defined by an acoustic monopole at 1 m in the simulation, filters $w(\omega)$ are found by solving the following optimization problem (dependency on the circular frequency ω is omitted for clarity) [1]:

$$min_w \; w^H \Gamma_{\nu\nu} w \;\text{ subject to }\; w^H d_0 = 1 \tag{1}$$

$\Gamma_{\nu\nu}$ is the noise covariance matrix that is generally ill-conditioned at low frequencies and requires a regularization. The constraint guarantees a constant gain in the look direction d_0. The optimum weight vector w_{opt} is defined by:

$$w_{opt} = \frac{\Gamma_{\nu\nu}^{-1} d_0}{d_0^H \Gamma_{\nu\nu}^{-1} d_0} \tag{2}$$

$\Gamma_{\nu\nu}$ can easily be built in free-field, assuming isotropic 3D or back-to-front noise fields. Here, the arrays are adjacent to a large rigid obstacle made from the head structure. Therefore the computation of optimal beamforming filters requires the use of Boundary Element Method (BEM) implemented with FEMAP Finite Element Modeling and Postprocessing software (Simcenter Femap/NASTRAN XaaS) for characterizing the acoustic field surrounding the head for noise and speech sources. Such approach is more effective than traditional binaural beamforming implementations [2]. In this investigation, the beamformer is designed and validated with the 45CB artificial head (GRAS, Holte, Denmark). The BEM mesh model is illustrated in Figure 2.

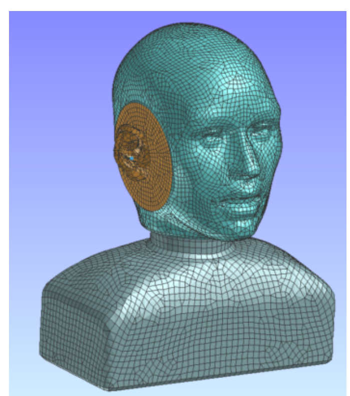

Figure 2: BEM mesh model of GRAS 45CB artificial head

2.2 Simulations

Simulations consists of determining optimal complex filters $w(\omega)$, Directivity Index (DI), White Noise Gain (WNG), and producing 3D beam patterns. DI varies from 5 dB around 200 Hz to a maximum of 9 dB around 3 - 4 kHz. Maximum WNG degradation is limited, around 10-15 dB at low frequencies.

*stephane.dedieu@bloo-audio.com
†thomas.padois@etsmtl.ca
‡jeremie.voix@etsmtl.ca

2.3 Validations

Validation took place in a 100 m^3 anechoic room, with source at 1 m from the center of the artificial head, at all azimuths, in the 150-7,000 Hz frequency range.

Figure 3: Test set-up in the anechoic room

In the horizontal plane, the agreement is excellent for all frequencies of interest, at all azimuths. An example is shown in Figure 4.

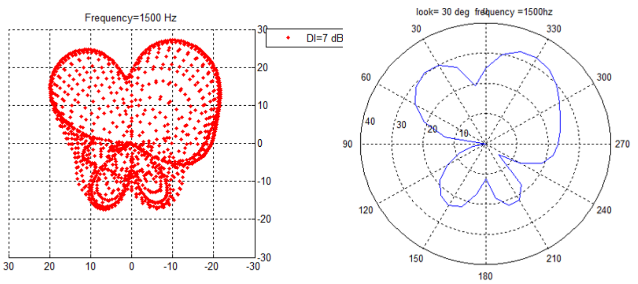

Figure 4: Validation horizontal plane - 30 deg. azimuth - 1500 Hz

If a larger array is an advantage when rejecting ambient noise at low frequencies, 1,500 Hz is an interesting frequency because with an inter-array distance of $\sim$ 16 cm, spatial aliasing starts to take place for all "look directions" angles.

2.4 Binaural GSC implementation

There are several approaches to handle such edge case. One can work with individual Left & Right small arrays, but this would make steering at 30 deg. and 60 deg. relatively inefficient. In this study, we implement (1) a Binaural strategy based on the monaural beamformer output, filtered with HRTFs and (2) a simple realization of a General Sidelobe Canceller (GSC) [1], by adding a "noise channel". The "noise channel", equivalent to the usual GSC blocking matrix, is built by filtering all microphones' outputs. Optimal filters are determined by steering a null in the look direction $d_0(\omega)$ using equation 1 and a proper set of constraints. Their definition is not unique. The noise channel signal is then used for implementing a Noise Reduction algorithm at the output of the main beamformer, based on a modified Wiener gain. The method is tested with recordings in the next section.

3 Results

Tests are performed in a large reverberant room. Four large loudspeakers at each corner generate the interfering noise field. Speech samples taken from the HINT database [3] are played from a smaller loudspeaker placed 1 m away from the head, in the plane of the arrays. Recordings and array processing are performed in 4 different conditions: in Silence, and with different background noises (White noise, Industrial noise, Cocktail party noise). In the present demo, the sound source is located at 60 deg. on the right. The binaural beamformer properly positions the speech source at 60 deg., but the residual noise is also located at 60 deg. which sounds a bit unnatural. The waveform of the right channel signal of the binaural beamformer is shown in Figure 5. The MVDR beamformer improves reverberations and reduces the white noise, overall by about 4 to 8 dB. The General Sidelobe Canceler (GSC) strategy with a modified Wiener gain further attenuates reverberations, and reduces white noise by about 9 to 12 dB. Speech sounds a bit "dry" and distorted, but it is clear and intelligible.

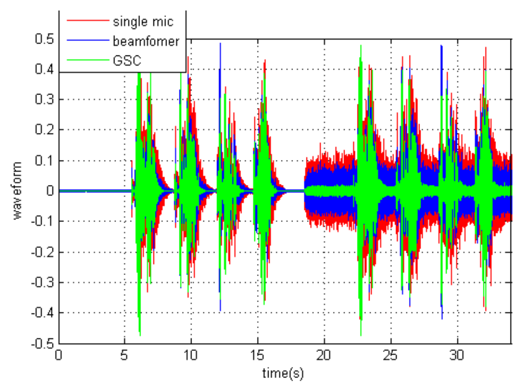

Figure 5: Right channel audio waveforms in various conditions

Acknowledgments

The authors acknowledge the technical support received from the NSERC-EERS Industrial Research Chair in In-Ear Technologies (CRITIAS) and from ICAR laboratories (Infrastructure commune en acoustique pour la recherche ÉTS-IRSST).

References

[1] Michael Brandstein, Darren Ward: "Microphone Array", Springer-Verlag, 2001

[2] Zhengwei Luo, "Beamforming for binaural hearing aids", Master Thesis, Univ. of Ottawa, 2009

[3] Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc. Am. 1994 Feb;95(2):1085–1099