

DEVELOPMENT OF A REAL-TIME EOG-BASED ACOUSTICAL BEAMFORMER ALGORITHM FOR BINAURAL HEARING DEVICES

Olivier Valentin ^{*1}, Saumya Vij ^{†1} and Jérémie Voix ^{‡1}
¹École de technologie supérieure, Montréal, Québec, Canada

Résumé

L'électro-oculographie (EOG) est une technique permettant, notamment, de mesurer les fonctions de la motricité oculaire grâce à des électrodes périorbitales enregistrant la différence de potentiel entre la cornée (positif) et la rétine (négatif). Cet article présente la preuve de concept d'un algorithme de *beamforming* acoustique utilisant l'angle de déplacement oculaire obtenu à partir d'enregistrements EOG pour optimiser la localisation et la perception des sons. Un tel algorithme permettrait d'optimiser l'expérience utilisateur des personnes utilisant des dispositifs auditifs binauraux tels que des audioprothèses ou des protecteurs auditifs numériques, en améliorant, par exemple, la reconnaissance de la parole en présence de bruit.

Mots clés : beamforming acoustique, audioprothèse, électro-oculographie (EOG), simulink, binaural

Abstract

Electro-oculography (EOG), a technique that can be used to evaluate ocular motility, uses periorbital electrodes to record the difference in potential between the cornea (positive) and the retina (negative). This paper presents a proof of concept for an acoustical beamforming algorithm using the gaze angle obtained from EOG recordings to optimize sound localization and perception. This algorithm could enhance the user experience for binaural hearing devices such as hearing aids or digital hearing protectors by improving, for example, speech recognition in noise.

Keywords: beamforming, hearing aids, electrooculography (EOG), simulink, binaural

1 Introduction

Hearing aid users often complain about the difficulty of listening to a given sound source, for instance, speech from their interlocutor, in the presence of disturbances, such as concurrent babble, the so-called “cocktail party effect”. To help hearing aid wearers, the devices are now designed to simultaneously reduce background noise and increase speech intelligibility without adding artefacts or distortions to the signal. Two main features are considered to achieve this goal, signal processing algorithms, and directional microphones [1].

Directional hearing aids currently on the market are based on the assumption that people listen to what is in front of them. Such devices usually include more than one microphone to increase the signal-to-noise ratio through spatial consideration: the signal of interest from the frontal hemispheres is considerably amplified while the signal coming from rear azimuths is less amplified [1, 2]. Still, Thorpe *et al.* [3] and Srinivasan *et al.* [4] have concluded that head and eye orientations are the most obvious indicators of attentional orientation. Srinivasan *et al.* [4] mention that listeners may also pay attention to a direction that they are not actually facing, which is called, covert attention. Numerous characteristics associated to the aids, as

well as to the environment, must be taken into consideration in the evaluation of directional hearing aid benefits. Chung *et al.* relate some of these characteristics [1]. Notably, the directivity index of the microphone, characteristics of the sound sources (quantity, spatial location and type), characteristics of the environment (room, environmental acoustics), relative distance between the sound source of interest and the user as well as the location of the background noise, relative to the user. For example, Chung *et al.* [1] and Killion *et al.* [5] demonstrate that reverberation reduces the advantages of the hearing aids' directionality; the sounds are reflected from different surfaces in every direction, making it impossible to discriminate the sound source of interest from its spatial origin [1, 5]. Directional hearing aids also present disadvantages when the signal of interest is behind the user or when a surrounding wind noise is present [6, 7].

On the other hand, omnidirectional hearing aids amplify sounds coming from all directions equally. Gnewikow *et al.* [5] show that directional hearing aids give better speech-intelligibility performances than omnidirectional aids, but also that the difference decreases as the degree of hearing loss increases. Gnewikow *et al.* also reported that people with mild hearing loss mostly prefer directional hearing aids whereas people with moderate and severe hearing loss seem to prefer omnidirectional hearing aids. Generally, the omnidirectional mode is preferred in quiet environments, when the sound source of interest is not located in front of the user, or when the sound source is moving, while the directional mode is preferred when the sound source is

* m.olivier.valentin@gmail.com

† saumya.vij@gmail.com

‡ jeremie.voix@etsmtl.ca

located in front of the user, or close enough to the user, and if there is a consistent presence of background noise [8]. As a result, hearing aids that are both omnidirectional and directional have proven more successful, and are more commonly used. Also, it has been demonstrated in clinical studies that such aids can lead to good results when the user is able to switch properly between both functionalities [6].

Algorithms embedded in hearing aids can localise sound sources surrounding the hearing aid users. However, such algorithms remain unable to identify which is the source of interest.

Bulling *et al.* argue that eye movements can supply information on visual tasks and on cognitive processes of visual perception such as attention [9]. According to them, electrooculography (EOG) is a method well suited for mobile eye tracking, and is relatively inexpensive. They also qualify this method as reliable, easy to use, and unobtrusive, when compared to common eye tracking systems that use cameras. In their study, Bulling *et al.* use EOG to detect three specific eye movement patterns, namely saccades, fixations and blinks. The algorithms used to detect fixations use the fact that the gaze remains stable during a fixation [9]. Their results show an average precision in pattern detection of 69-93% for six out of eight participants, and less than 50% for the two last participants.

Joyce *et al.* developed a method based on EOG signals to track where an individual's gaze is directed on the surface of a flat screen (x, y coordinates) [10]. They obtained a mean error of 1-2 degrees on a 30-degree amplitude position (15 degrees on each side of the screen centre). However, according to them, the relationship between the EOG output and the angle of gaze stays linear within a limited range of up to +/- 70 degrees. Unfortunately, this method needs a calibration step before being usable, which means the participants had to sit at a precise distance from the screen and perform a few specific eye movements. If an electrode is displaced during the insertion of such an instrumented hearing aid device, the system would need to be recalibrated before being once again usable. Consequently, EOG-based eye tracking might not be well suited for real-world situations. However, the proposed system described in this paper might not necessarily require such recalibration since it can accommodate a certain range of error without altering the measurement of the eye gaze.

The paper is structured as follows: the design of the EOG-based beamforming is described in section 2. Section 3 contains preliminary results used to validate the proposed model. Conclusions and future works are presented in section 4.

2 Design of the EOG-based beamformer

The present study aims to develop a proof-of-concept of an EOG-based acoustical beamforming algorithm for improving speech discrimination in noisy environments and optimizing sound localization for users wearing binaural hearing aids.

2.1 Overview of fixed beamforming techniques

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 [11]. Different types of fixed beamformers exist, e.g. delay-and-sum beamforming, superdirective beamforming, differential microphone arrays and frequency invariant beamforming [12-17]. Fixed beamformers have mainly been used for monaural hearing aids [18-20]. Fixed beamforming techniques have also been proposed for binaural hearing aids combining spatial selectivity and noise reduction with the preservation of the speech source's binaural cues [21-23].

2.2 Design and implementation

Figure 1 shows a schematic of the proposed EOG-based beamformer prototype. The EOG signals recorded using electrodes F7 and F8 of the Emotiv (San Francisco, CA, USA) EPOC® headset are supplied to the EOG block, which calculates the corresponding angle. The angle is supplied to the beamformer for Interaural Time Difference (ITD) and Interaural Level Difference (ILD) calculations. The model takes stereo sound input from the left and right microphones, processes it in real-time and sends it back to the stereo headphones to create the binaural effect.

Figure 2 presents the block diagram of the model implemented on Simulink [24]. Four inputs are needed to perform the beamforming: the left and right in-ear microphone inputs, which are given by the *Audio processing* group of blocks, and the left and right delays, which are computed by the *EOG processing* group of blocks. Several steps are required to obtain the left and right delays.

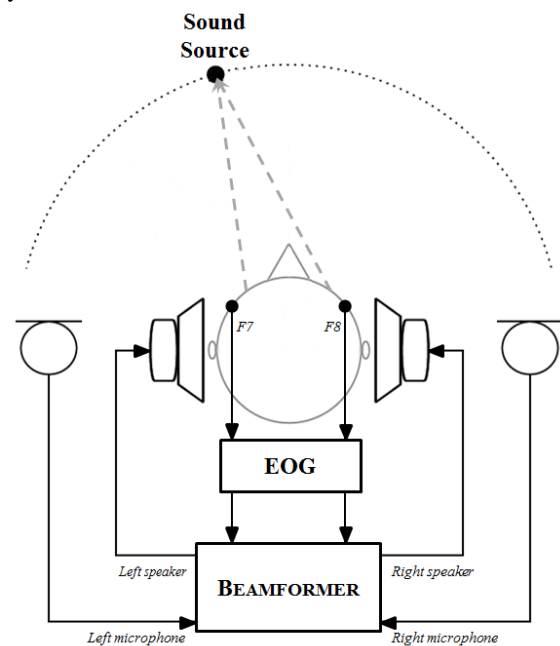


Figure 1: Schematic of the proposed EOG-based beamformer

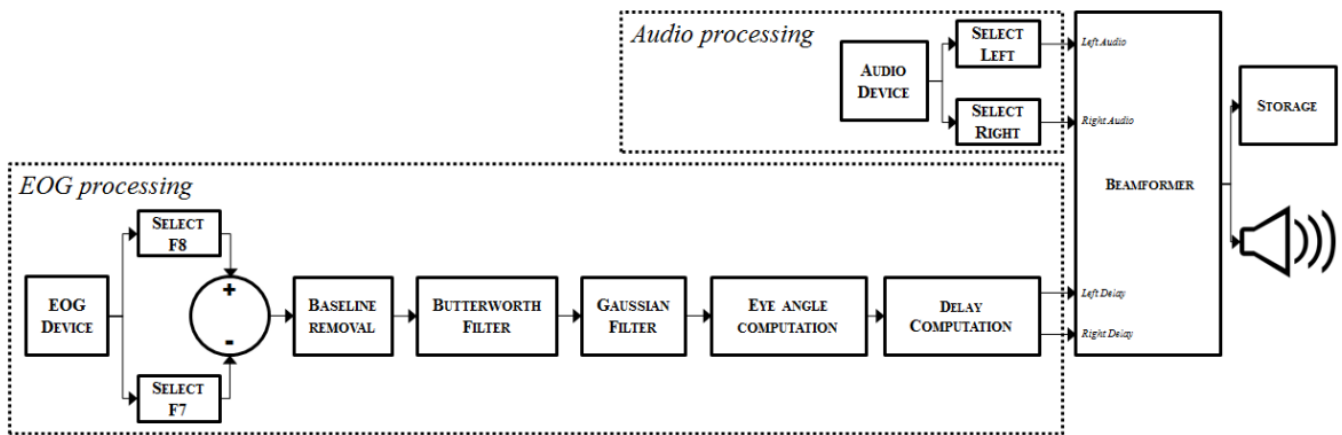


Figure 2: Simulink Block Diagrams of the proposed EOG-based beamformer.

First, EOG raw data obtained at F7 and F8 positions are selected from the output data given by the Emotiv headset and are formatted into a one-column format per electrode. These data are then subtracted to get the horizontal raw signal before being weight-averaged to remove the baseline from the horizontal raw signal.

Then, the signal is filtered and the corresponding horizontal angle is estimated mathematically. A MATLAB [25] function finally computes the delays corresponding to the given angle required to produce the binaural effect and stores the output into a multimedia file, for validation purposes.

The algorithm was first implemented completely in an offline mode using MATLAB and post-processing scripts. This offline mode includes scripts for EOG processing and angle calculation using windowed treatment.

Once the algorithm had been validated offline, a real-time model, with real-time input EOG signals and real-time inputs from the microphone, was developed using Simulink. The major advantage of Simulink was to facilitate the implementation of sample-by-sample real-time processing while MATLAB scripts used windowed treatment.

EOG recordings

EOG recordings, a technique used to evaluate ocular motility, can be used to assess the horizontal eye movements that are the most obvious indicators of where people are trying to direct their hearing attention. To obtain these horizontal eye movements, EOG signals were recorded on one subject using two electrodes of the Emotiv EPOC® headset placed on the external canthi (the bone on the side of the eye).

Minimal skin preparation was deliberate, to reproduce the limitations of acceptability as are to be expected of any device that is to be worn in social settings. Figure 3 shows the electrode placements used to measure EOG and Figure 4 presents a section of the raw EOG signal recorded on one subject.

Baseline removal

During electrophysiological recordings, drifts and/or direct current offsets due to sweating and skin conductance

variations or other noise sources can compromise the quality of the recorded signals. Therefore, a common procedure is used to relativize the signal of interest with respect to a control (baseline) signal, shortly recorded before a stimulus event. Such procedures require offline data processing. Consequently, two methods were developed to perform the baseline removal on the data recorded on one participant, without doing offline data processing.

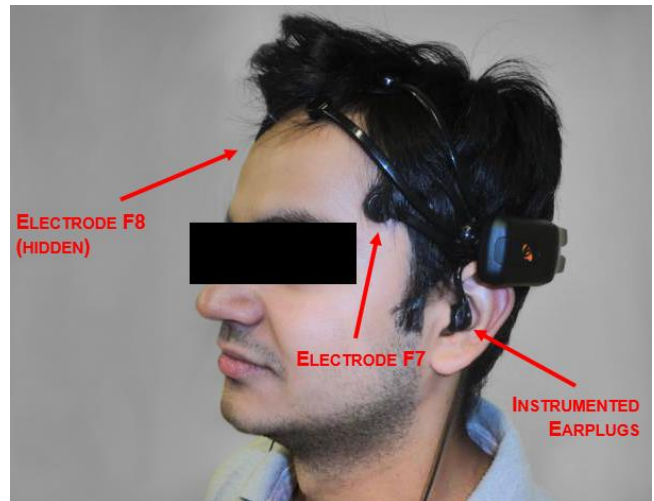


Figure 3: Electrode placement used for EOG recordings with the Emotiv EPOC® headset.

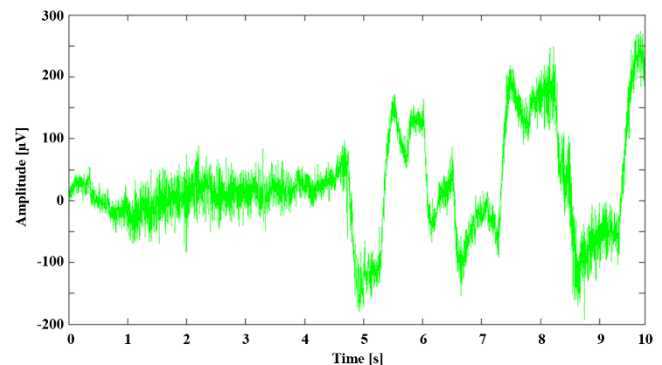


Figure 4: Raw EOG signal of F7-F8 signals, as recorded on one subject using electrode placement illustrated in Figure 3.

The first proposed method subtracts the mean amplitude value of the signal computed on three consecutive data windows of 10 ms to the data window being processed. The second proposed method subtracts the mean amplitude value of the signal computed from the first point of the first data window to the first point of the data window being processed. This method requires more data to compute the arithmetic mean since the number of data points being considered in the calculation increases as the data process moves forward.

Figure 5 presents the results obtained with the conventional baseline removal method (top) and the results obtained with the two proposed methods for baseline removal (middle and bottom). Figure 5 indicates that the results obtained with the second proposed method (bottom) are closer to those obtained with the conventional method (top) than those obtained with the first proposed method (middle).

Filtering

After the baseline removal, data were low-pass filtered using a fifth order Butterworth filter (-3dB cut-off at 10 Hz) before being filtered with a Gaussian filter to remove high frequency components since this type of filtering method requires fewer calculations than a frequency cut performed in the frequency domain. Figure 6 shows the filtered data from a subsection of the EOG signal presented in Figure 5.

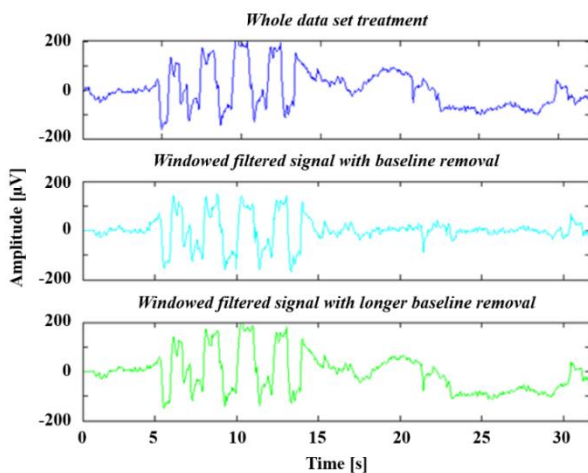


Figure 5: Baseline removal results obtained with a conventional method (top), with the first proposed method (middle) and with the second proposed method (bottom).

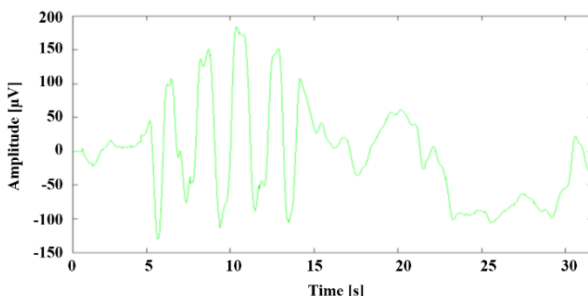


Figure 6: EOG signal filtered with a Gaussian filter, after baseline removal.

Horizontal eye angle computation

The EOG filtered results were used to establish the relationship (1) between the horizontal eye angle and the amplitude of the EOG signal, assuming that the angular movement of the eyes is linearly related to the amplitude of the signal, within a given angle interval.

$$\theta(i) = \frac{A(i)}{4.44\mu V} \quad (1)$$

Where θ is the angle in degrees, A is the EOG signal amplitude in μV and i is the index of the sample.

Delay-and-sum beamforming

The beamforming method presented here, relies on spatial coherence. Signals that reach the two binaural microphones are delayed in time (in number of samples), following the desired beamforming angle, and are summed, so their amplitude is doubled if they are perfectly in phase, and diminished, to a different extent, if they are not in phase [26].

Figure 7 shows schematically the different steps used to execute the delays and summations. In the example illustrated in Figure 7, the imposed delay consists of three signal samples received on the right microphone. Consequently, the signal that is to be amplified is the one coming from the left side of the user and which reaches the left microphone some time before the right microphone (with a delay of approximately three samples). The last step illustrated at the bottom of Figure 7, corresponds to the shift of the matrices' sum, formatted as the original data.

In this step, data that were originally present in the right and left matrices are replaced by the ones that have been summed: signals that were not in phase (such as random noise) are therefore subtracted. Once all these steps are executed, the two matrices are presented to the test subject, through the miniaturized speakers embedded inside the earpieces.

However, the presence of numerous discontinuities in the signal creates important artefacts in the signal that are sent to the user's ears.

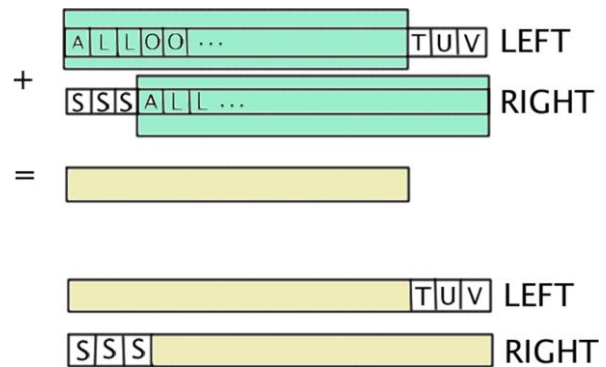


Figure 7: Steps in delay and summation, for a signal delayed by three samples (S) on the right microphone, while a signal starting with “ALLOO” is being played. “TUV” indicates the trailing audio message received on the left microphone but not yet picked-up by the right microphone.

At the beginning of a data window, when there is a difference between the amplitude of the last window's data and the amplitude of the first data of the processed window, an artefact may be generated. Also, at the limit between the section of the data window that has been summed and the rest of the data window, there can be a discontinuity that generates artefacts.

Figure 8 shows three successive data windows as well as the places where the discontinuities can be found.

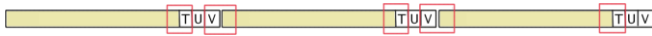


Figure 8: Discontinuities.

With this method, which consists of using only one window of data at a time, the main problem is not related to the discontinuities between summed values and original values, but to the discontinuities between two different windows of data, where an interpolation is necessary because it is impossible to filtrate (i.e., filtrated data that have already been sent to the user's ears). With interpolations, the results generally give better results, but the artifacts are persistent. Also, an interpolation between every single window implies a heavy calculation task. Another method was thus explored, and consists of using two data windows at a time instead of only one. This method is presented in Figure 9.

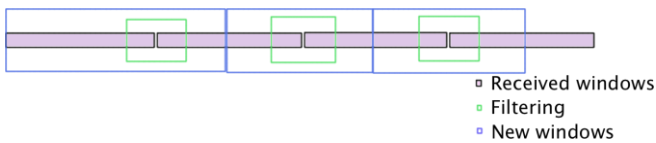


Figure 9: Reducing discontinuities.

3 Preliminary results

Gaze detection

Figure 10 presents the plot of the gaze angle versus time. As can be seen on this figure, the proposed EOG-driven beamforming algorithm was able to assess the movements generated by the eyes of the participant.

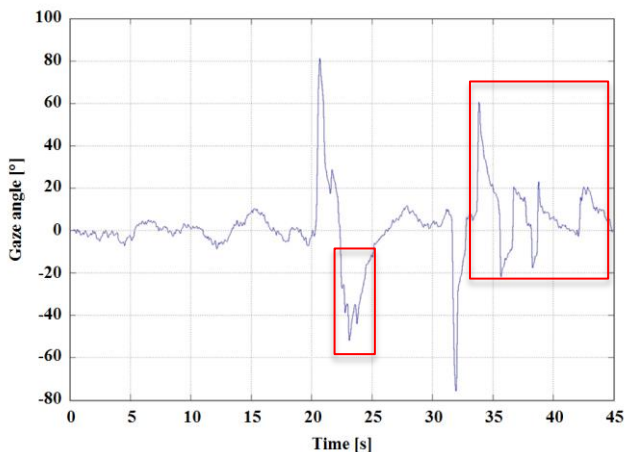


Figure 10: Gaze angle vs time plot in a Simulink scope.

Acoustical beamforming

With this proposed delay-and-sum method, a single filtering operation comprises one filtering operation and one interpolation. The filtering that is executed between two windows includes 50 samples. If a delay of 14 samples corresponds to a 45° angle, 25 samples on each side, included in the filtering, means that all the discontinuities are treated by this same filtering step. In other words, this filtering reduces the number of discontinuities present between two windows of data as well as the discontinuities present between the summed data and the original data in a window. This method reduces the complexity of calculation and at the same time, the time required for the calculation, and the interpolation generated artefacts. A drawback of this method is that two windows are required for processing instead of only one, which means that the delay between the time when the signals are sampled and the time when the signals are sent to the user's ears is increased. For 1500 samples per window and a sampling frequency of 44,100 Hz, the delay related to the proposed processing method is 0.068 seconds, excluding computation time.

Figures 11 and 12 show the results of a simulation that point out this phenomenon. In this simulation, the sound sources are localized all around the user, spaced at 1-degree intervals, and at a distance of one meter from the hearing aid user. For this, the beamformer is steered at 90° on the figures, which corresponds to the front of the user.

The reason why the lines are not smooth on Figures 11 and 12 is that the imposed delays are calculated in terms of samples instead of exact time. For a certain distance interval from the source, the imposed delay stays the same, which creates a "step" on the plot.

As can be seen from Figures 11 and 12, the proposed binaural beamforming approach only uses spatial coherence and does not provide the same enhancement across all the frequency range.

4 Conclusion and future work

This study presents an EOG-driven beamformer proof-of-concept. This real-time model is able to correct the audio signals recorded from the left and right in-ear microphones using EOG signals recorded from two electrodes, and present the optimized audio signals to the left and right in-ear speakers.

Further research is needed to overcome several limitations of this proposed EOG driven beamformer. Regarding the beamformer algorithm, further improvements could be achieved by using the resulting filtered signal to calculate the spectral weight to be applied on the head-related transfer function (HRTF) rather than being directly used as playback signal to be sent to the in-ear headphones. For the EOG gaze-detection algorithm, several filtering techniques that were applied in the post-processing mode are not implementable in a real-time embedded system. Therefore, alternate solutions must be applied to smooth the fast temporal transitions after each eye movement. Another limitation of this model remains the baseline removal method.

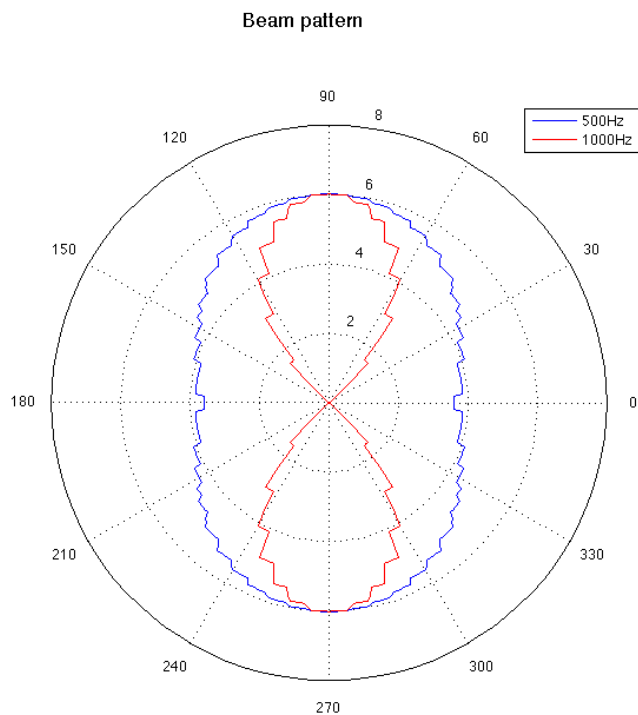


Figure 11: Simulation results of proposed beamformer in low frequency, at 500 and 1000 Hz showing clear improvement on front and back grains.

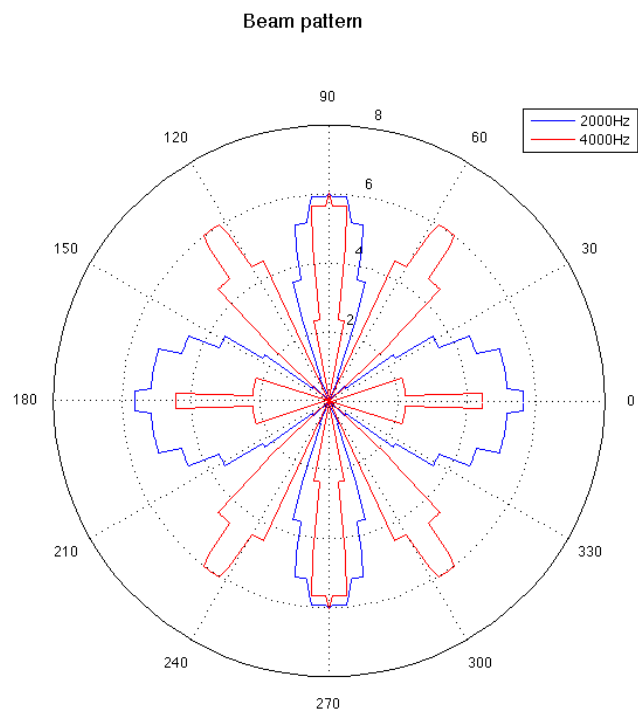


Figure 12: Simulation results of binaural beamformer in medium frequency, at 2000 and 4000 Hz showing the limits of the proposed delay-and-sum approach, with multiple side-lobe amplifications.

As can be seen in Figure 8, there is an immediate opposite spike whenever there is a large angle on either side. Also, the angles are not as stable as those derived from the

MATLAB post-processing code. The most probable source of error seems to be the baseline removal method. The beamformer requires validation through an ‘8-Plot diagram’ as well. EOG processing could also be further validated by a camera-based eye tracker.

Furthermore, eye angle estimates rely on two essential values, which are maximal amplitude of the signal and the maximum angle corresponding to that amplitude. These two values have been approximated here with the help of the complete set of data. A method needs to be found that would allow us to measure these two fundamental values first, within the initial sampled data, and second, with a method that ensures valid values that are user representative. To do so, one option is to perform a calibration at the beginning of every recording session, when the user puts the hearing device on. It could be in the form of two successively emitted sounds, one coming from the right, and the other from the left, to which the user would answer by looking towards the presented sound sources. Thus, the user’s maximum possible reachable angle can be obtained, as well as the related maximum and minimum amplitude of the EOG signal.

Finally, as the focus of listening is not given by eye movements or by head movements alone, but by some optimal combination of the two, the proposed EOG-based beamformer might also be improved with the adjunction of a head-tracking feature.

Acknowledgments

The authors would like to thank Marc Schöenwiesner, from the International Laboratory for BRAIn, Music and Sound Research (BRAMS) for the original idea and Marie-Hélène Faille for the initial work on the binaural beamformer. The authors wish to express their appreciation to EERS Global Technologies Inc. (Montréal, Québec, Canada) for providing the custom in-ear devices used in this study.

References

- [1] K. Chung. Challenges and recent developments in hearing aids: Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. *Trends in amplification*. 8(3):83-124, 2004.
- [2] D. Gnewikow, T. Ricketts, G.W. Bratt and L.C. Mutchler. Real-world benefit from directional microphone hearing aids. *The Journal of Rehabilitation Research and Development*. 46(5) 603-618, 2009.
- [3] S. Thorpe, M. D’Zmura and R. Srinivasan. Lateralization of frequency-specific networks for covert spatial attention to auditory stimuli. *Brain topography*. 25(1): 39-54, 2012.
- [4] R. Srinivasan, S. Thorpe, S. Deng, T. Lappas and M.D. Zmura. Decoding Attentional Orientation from EEG Spectra. *Human-Computer Interaction*. Part I, 5610, 176-183, 2009.
- [5] M. Killion, R. Schulein, L. Christensen, D. Fabry, L. Revit, P. Niquette and K. Chung. Real-world performance of an ITE directional microphone. *The Hearing Journal*. 51(4):1-6, 1998.
- [6] T.A. Ricketts. Directional hearing aids: Then and now. *The Journal of Rehabilitation Research and Development*. 42(4 Suppl 2), 113-144, 2005.

- [7] T. Ricketts, P. Henry and D. Gnewikow. Full Time Directional versus User Selectable Microphone Modes in Hearing Aids. *Ear & Hearing*. 24:424–439, 2003.
- [8] K.W. Grant, M. Elhilali, S.A Shamma, B.E. Walden, R.K. Surr, M.T. Cord and V. Summers. An Objective Measure for Selecting Microphone Modes in OMNI/DIR Hearing Aid Circuits. *Ear & Hearing*. 29:199–213, 2008.
- [9] A. Bulling, J.A. Ward, H. Gellersen and G. Tröster. Eye movement analysis for activity recognition using electrooculography. *IEEE transactions on pattern analysis and machine intelligence*. 33(4), 741–53, 2011.
- [10] C.A. Joyce, I.F. Gorodnitsky, J.W. King and M. Kutas. Tracking eye fixations with electroocular and electroencephalographic recordings. *Psychophysiology*. 39(5): 607-618, 2002.
- [11] S. Doclo, S. Gannot, M. Moonen and A. Spriet. Acoustic beamforming for hearing aid applications. Handbook on Array Processing and Sensor Networks. Chapter 10. John Wiley & Sons Inc., 2008.
- [12] J. Bitzer and K.U. Simmer. Superdirective Microphone Arrays. Microphone Arrays: Signal Processing Techniques and Applications. Chapter 2, pages 19-38. Springer-Verlag, 2001.
- [13] 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, 2003.
- [14] S. Doclo and M. Moonen. Design of far-field and near-field broadband beamformers using eigenfilters. *Signal Processing*. 83(12):2641-2673, 2003.
- [15] G. Elko. Superdirectional Microphone Arrays. Acoustic Signal Processing for Telecommunication. Chapter, pages 181-237. Kluwer Academic Publishers, 2000.
- [16] E.E. Jan and J. Flanagan. Sound capture from spatial volumes: Matched-filter processing of microphone arrays having randomly-distributed sensors. In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, pages 917-920, Atlanta, GA, USA, May 1996.
- [17] M. Kajala and M. H'am'al'ainen. Broadband beamforming optimization for speech enhancement in noisy environments. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 19-22, New Paltz, NY, USA, October 1999.
- [18] J.M. Kates. Superdirective arrays for hearing aids. *Journal of the Acoustical Society of America*. 94(4):1930-1933, 1993.
- [19] 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, 1996.
- [20] W. Soede, A.J. Berkhout, and F.A. Bilsen. Development of a directional hearing instrument based on array technology. *Journal of the Acoustical Society of America*. 94(2):785-798, 1993.
- [21] M.R. Bai and C. Lin. Microphone array signal processing with application in three-dimensional hearing. *Journal of the Acoustical Society of America*. 117(4):2112-2121, 2005.
- [22] 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, 1997.
- [23] I.L.D.M. Merks, M.M. Boone, and A.J. Berkhout. Design of a broadside array for a binaural hearing aid. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, October 1997.
- [24] Simulink 8.2, The MathWorks, Inc., Natick, Massachusetts, United States.
- [25] MATLAB, release R2013b, The MathWorks, Inc., Natick, Massachusetts, United States.
- [26] Y. Zeng and R.C. Hendriks. Distributed Delay and Sum Beamformer for Speech Enhancement via Randomized Gossip. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 22(1):260-273, 2014.

For
Digital Recorders

Introducing

For
USB A/D Systems

PHANTOM POWER

7052PH

Measurement Mic System

7052H Type 1.5™

Titanium Diaphragm

3Hz to >20 kHz

<20 dBA > 140 dB SPL

MK224 (<14 dBA to >134 dB SPL) Optional

4048 Preamp

Superior
IEC 1094 Type 1
Long-term Stability
Temperature and Humidity
Performance

Now in Stock

**Phantom
to IEPE/ICP
Adaptor
Supplies 3-4 mA
Power
Accelerometers
Microphones**

ICP1248

A B
C e
O g
u i
S n
w
S i
t t
h
i
A
C
C
S O



**MATT™
Family**

Mic Attenuator

Handle Higher Sound Pressure Levels

ACO Pacific, Inc., 2604 Read Ave., Belmont, CA 94002

Tel: (650) 595-8588 FAX: (650) 591-2891 E-Mail: sales@acopacific.com

Web Site: www.acopacific.com

TM