TECNIACÚSTICA
CUENCA 2023

DESIGN OF A NOISE-SELECTIVE SOUND SEAT COMBINING AUDITORY SCENE ANALYSIS AND ACTIVE SOUND FIELD TECHNIQUES

Alberto González Salvador[1], Manuel Rosa Zurera[2], María De Diego Antón[1], Roberto Gil Pita[2], Miguel Ferrer Contreras[1], Manuel Utrilla Manso[2]*

[1] Universitat Politècnica de València, Valencia, España
[2] Universidad de Alcalá, Alcalá De Henares, España

RESUMEN

Se ha demostrado que la exposición prolongada a niveles elevados de ruido contribuye de forma significativa a diversos problemas de salud, como la discapacidad auditiva, la hipertensión y los trastornos del sueño. Además, los niveles elevados de ruido pueden exacerbar los niveles de estrés, aumentar la incidencia de accidentes laborales y fomentar comportamientos antisociales y agresivos.

Afortunadamente, estos problemas pueden aliviarse mediante la aplicación de técnicas de control activo del sonido, que consisten en manipular el entorno acústico utilizando altavoces para crear un campo sonoro deseado dentro de un área específica. Sin embargo, la integración de estas técnicas con el análisis de escenas auditivas basado en arrays plantea retos debido a las limitaciones físicas impuestas por la colocación de micrófonos y altavoces, que afectan al retardo máximo de procesamiento disponible para el sistema.

En este artículo se analiza el diseño de un asiento sonoro activo con selección de ruido, prestando especial atención a las limitaciones de diseño asociadas. El asiento propuesto combina el control activo del sonido con un amplio conjunto de técnicas derivadas del marco del Análisis Auditivo de Escenas. Estas técnicas permiten analizar mezclas de sonidos para separar fuentes sonoras individuales. Al analizar espacialmente la posición y las características de las distintas fuentes sonoras dentro de una sala, el asiento permite al usuario enmascarar selectivamente las fuentes sonoras no deseadas mediante técnicas de campo sonoro activo.

ABSTRACT

Prolonged exposure to elevated noise levels has been established as a significant contributing factor to various health problems, such as hearing impairment, hypertension, and sleep disturbance. Moreover, elevated noise levels can exacerbate stress levels, increase the incidence of workplace accidents, and foster anti-social and aggressive behaviors.

Thankfully, these issues can be alleviated through the implementation of active sound control techniques, which involves manipulating the acoustic environment by utilizing loudspeakers to create a desired sound field within a specific area. However, the integration of these techniques with array-based auditory scene analysis poses challenges due to physical limitations imposed by microphone and loudspeaker placement, which affect the maximum processing delay available for the system.

This paper analyzes the design of a noise-selective active sound seat, with a particular focus on the associated design constraints. The proposed seat combines active sound control with a comprehensive set of techniques derived from the Auditory Scene Analysis framework. These techniques enable the analysis of sound mixtures to separate individual sound sources. By spatially analyzing the position and characteristics of different sound sources within a room, the seat empowers the user to selectively mask undesirable sound sources using active sound field techniques.

Keywords/Palabras Clave— Active Noise Cancellation, Least Mean Squares Beamforming, Sound Seat.

* ***Corresponding author / Autor de contacto***: *roberto.gil@uah.es*
Acknowledgements: *This work has been funded by the Spanish Ministry of Science and Innovation with projects PDC2022-133651-C21 and PDC2022-133651-C22 (both funded by MCIN/AEI/10.13039/501100011033 and European Union NextGenerationEU/PRTR); by MICIN under grant PID2021-125736OB-I00 (MCIN/AEI/10.13039/501100011033/, "ERDF A way of making Europe"); by GVA under grant CIPROM/2022/20; by the Spanish Ministry of Science and Innovation with project PID2021-129043OB-I00 (MSI/FEDER); and by the Community of Madrid and the University of Alcalá under project EPU-INV/2020/003.*

1. INTRODUCTION

Active sound control (ASC) is the process of manipulating a sound field using electroacoustic transducers, typically loudspeakers, for the purpose of creating a desired sound field into a spatial zone. It can be used for canceling noise (ANC) [1], reshaping sound field [2], spatial audio reproduction [3], sound quality improvement [4], and others. Adaptive algorithms and methods are commonly used [1], although even machine learning techniques have been tested [5]. Currently, most of the efforts in ANC are focused on improving the performance of adaptive strategies customized to ANC, considering both linear controllers [6] and non-linear ones [7]. Thus, many algorithms to improve the convergence speed and tracking capability [8], the robustness [9], or the computational cost [10] have been proposed. In addition, control strategies particularized to the noise signal characteristics have also been extensively studied, for tonal or periodic noises [11] (typical in engines or revolution machines). Active Noise Equalization techniques (ANE) are used to keep a desired residue of the noise signal, instead of ANC. ANE can be used for narrowband noise (the most usual cases) [11, 12] or broadband noise [13, 14].

Most of these strategies require multichannel versions to extend control over a wider area of space, so most algorithms are proposed in their multichannel extension. These multichannel versions are high-computationally demanding when run on a centralized system, so decentralized strategies [15,16] and collaborative distributed algorithms [17,18] are being proposed. There are a variety of practical noise control problems currently addressed using active technologies, such as noise canceling headphones, and the control of: propeller induced noise in aircraft, noise in helicopters, and both engine and road noise in the automotive industry.

ANC can be complemented with the use of Auditory Scene Analysis (ASA) techniques. ASA consists in the analysis of sound mixtures to obtain objects corresponding to individual sound sources [19]. So, using information received by a set of microphones or array of microphones, ASA techniques aim at locating the different sound sources [20], classifying the typology of the sound source [21], and separating the sound received by this source from the possible interferent sounds present in the scene [22].

With this idea in mind, this paper studies the combination of these two big research fields: ANC and sound source separation, to propose a noise-selective sound seat. An array of microphones will select and separate the noisy sources, and two control speakers will apply a noise cancellation process, creating two spatial bubbles ideally placed on the ears of the subject, so that the person placed in the seat does not perceive the noise source, while maintaining the integrity of the desired sound sources.

The research is structured as follows: first, the basic concepts of ANC and array beamforming are presented. Then, the particularities of the proposed design are exposed. The paper then presents the results and conclusions derived from the experiments.

2. MATERIALS AND METHODS

The main sound seat system proposed in this paper is composed of two main subsystems, developed respectively at the Technical University of Valencia, and at the University of Alcalá. The first subsystem includes the development of algorithms for ANC, while the second one focuses on the development of algorithms for sound detection, location, and separation. The collaboration of both subsystems leads to the proposal of a system that combines an adaptive noise control subsystem with another system to generate knowledge of the acoustic scene, whose proof of concept is the objective of this paper.

2.1. Active Noise Cancellation

The basic idea of ANC originated in 1936. In summary, an ANC system consists in measuring the noise sound field with microphones, manipulating the resulting signal, and then feeding it to an electroacoustic secondary source, i.e., a loudspeaker, to produce an acoustic wave that attenuates the noise sound field at low frequencies. During the last four decades, much research effort has been devoted to understanding and solving most of the complexities that arise when implementing such an apparent system. These research efforts have made possible the implementation of useful ANC systems with application in different sectors of industry.

Possibly, the first practical application of an ANC system was aimed at attenuating the level of noise that traumatized pilots of fighter planes during military conflicts. Due to a lack of technology, the implementation of this ANC system could only be undertaken for luxury aircrafts and cars. However, the rapid development of inexpensive processors with advanced digital signal processing (DSP) capabilities has recently allowed large-scale commercialization of these systems for a vast number of applications. For the applications where ambient noise needs to be suppressed, the use of ANC systems for headset has become very popular. This is for instance the ANC system used by several flight companies to ensure the auditory comfort of passengers, who are generally seated for significant lengths of time and subjected to high ambient noise levels.

More recently, to remove the need of using headsets, several companies have already developed a cutting-edge technology that employs a set of loudspeakers and microphones when creating quiet bubbles where the ambient noise is attenuated. This technology can operate in a multi-user setting where multiple stand-alone and non-interacting ANCs can be simultaneously operating to create multiple quiet bubbles in the same room or hall, one per user. However, it suffers strong performance degradation if ANCs of different users are acoustically coupled, i.e., the signals

Figure 1. Car seats surrounded by sound transducers and microphones.

emitted by the loudspeaker of one local ANC are recorded by the microphones of another ANC. Moreover, due its non-cooperative nature, this technology does not leverage the spatial diversity of the signals to be acquired with microphones of the different ANCs to obtain a better cancellation of the noise within the different quiet bubbles of a multi-user setting. Finally, the solutions available in the market usually pursue maximal attenuation of all surrounding unwanted noise.

However, in some applications, instead of an ANC it is desirable to implement an ANE that retains a noise signal with a low level of power and a specified spectral shape. For instance, in a car or a truck, unlike the passengers, the driver needs some audible information about the engine speed to be able to control the vehicle safely. Therefore, adding a modern smart audition system to an ANC system will give an innovative noise selective ANE, whose noise target can be defined by each user depending on his needs or preferences.

2.1. Auditory Scene Analysis using a microphone array: sound source separation

As was stated above, the implementation of ANC solutions in a sound seat has already been carried out in the literature, but the problem of combining it with array-based sound source separation techniques has not been properly addressed.

Sound separation is a fundamental challenge in audio signal processing, aiming to extract specific sound sources from a mixture of acoustic signals. In recent years, beamforming techniques, especially when coupled with linear microphone arrays, have emerged as powerful tools for enhancing the separation of sound sources in various applications. This paper explores the principles and applications of sound separation using beamforming and linear microphone arrays.

Beamforming is a signal processing technique used to spatially filter acoustic signals received by an array of microphones, focusing on a particular direction while attenuating sounds from other directions. In essence, it creates a "beam" of enhanced sensitivity towards the desired source location. Beamforming algorithms exploit the spatial differences in arrival times and phases of sound waves at different microphones to enhance or attenuate signals, resulting in improved signal quality.

Linear microphone arrays consist of multiple microphones arranged in a linear fashion. The spatial arrangement of the microphones allows for precise control over the beamforming process. Each microphone captures a slightly different version of the incoming sound due to its location, enabling the system to differentiate between various sources based on their spatial characteristics.

3. DESCRIPTION OF THE PROPOSAL

The system proposed in this paper combines an ANC algorithm with a microphone linear array based beamformer technique, aiming at developing a highly innovative active sound seat, which can be eventually replicated and connected within a network. This sound seat simultaneously ensures the auditory comfort of multiple users located in different positions of a room by creating acoustically comfortable bubbles where the noise perceived by each user is selected and/or equalized according to some user-defined preferences.

This proof of concept is aimed at integrating and testing a novel knowledge-based listening and reproduction system for located listeners without the need of using headsets. This system, which we call active sound seat, is formed by two main parts: a knowledge-based listening module (KLM) and an active noise controller (ANC). Both subsystems interact to provide a personalized and pleasant acoustic environment to the listener, reducing user-selected undesired noises while keeping desired sounds (music, alarms, information, etc.). Furthermore, the system will provide information about the ambient sound and can be used within a collaborative network of active sound seats. By doing so, the proposed system can simultaneously ensure the auditory comfort of multiple users that are regularly exposed to high noise levels. The system can work in dynamic environments where the sound sources change or move.

Figure 1 shows the distribution of the microphones and speakers in the room, displaying the position of the noise and the desired source (red asterisk, 1 and 2 respectively), the position of the microphones in the linear array used to implement the KLM (black circles, 1 to 6), the position of the control mics where the space bubbles will be placed (blue circles, 7 and 8), and the position of the control speakers of the ANC (blue asterisk, 3 and 4). Please note that the separation between the 6 mics in the array (0.2 m between microphones) has been selected so that no aliasing is going to be present with signals under 850 Hz. Table 1 shows the coordinates of the different elements.

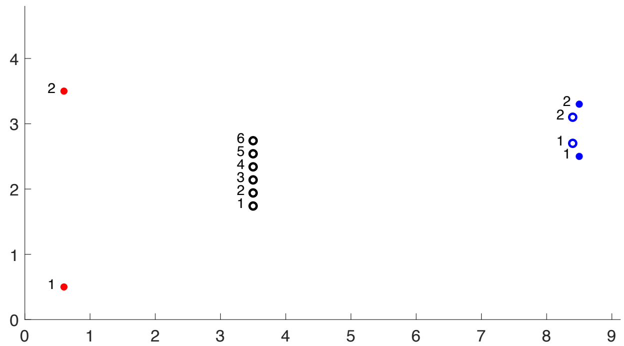

Figure 1. Room configuration displaying the position of the noise and the desired source (red asterisk, 1 and 2 respectively), the position of the microphone array (black circles, 1 to 6), the position of the control mics (blue circles, 7 and 8), and the position of the control speakers (blue asterisk, 3 and 4)**.**

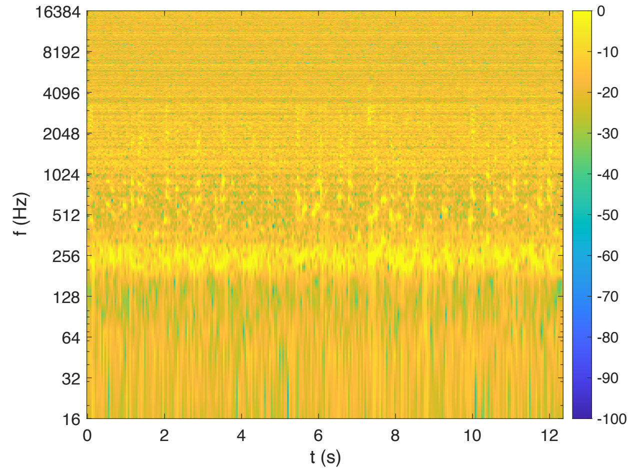

Figure 2. Reference mix in control microphone 1 without ANC**.**

Table 1. Spatial coordinates of the microphones and speakers.

	x (m)	*y* (m)	*z* (m)
Source 1 (noise source)	0.60	0.50	1.50
Source 2 (speech source)	0.60	3.50	1.50
Control Speaker 1	8.50	2.50	1.50
Control Speaker 2	8.50	3.30	1.50
Control Mic 1	8.40	2.70	1.50
Control Mic 2	8.40	3.10	1.50
Array Mic 1	3.50	1.74	1.50
Array Mic 2	3.50	1.94	1.50
Array Mic 3	3.50	2.14	1.50
Array Mic 4	3.50	2.34	1.50
Array Mic 5	3.50	2.54	1.50
Array Mic 6	3.50	2.74	1.50

The ANC algorithm used in this paper is based on the Frequency Domain Adaptive Filter algorithm (FDFA) [23]. This algorithm is designed to implement an adaptive finite impulse response (FIR) filter within the frequency domain using the fast block least mean squares (LMS) algorithm. This implementation allows for greater efficiency and control over filtering operations. Key parameters for this block include filter length and block length, which define the filter's length and the block size utilized by the algorithm.

The frequency-domain adaptive filtering process entails three essential steps: filtering, error estimation, and tap-weight adaptation. This algorithm performs FIR filtering within the frequency domain employing either the overlap-save or overlap-add method. Both error estimation and tap-weight adaptation are executed using the LMS algorithm, contributing to the algorithm's efficiency and adaptability.

Concerning the KLM, in this paper, a Filter-And-Sum beamformer (FASB) [24] is applied to separate the "desired" noise source from the rest of the sound sources present in the acoustic scene. In this algorithm, the signals received in each of the microphones in the array are filtered and then summed to obtain the separated noise reference, that will posteriorly be used by the ANC subsystem. In a design stage, the coefficients of the filter can be obtained by a LMS process, where the objective/target is the isolated noise signal received in the control microphones that are placed in the positions where the ANC want to be effective, typically close to the ears of the subject.

One important fact to take into consideration when designing an ANC system that includes real-time estimation of the desired noise source is the delay and available computing time. Due to the physicality of the proposal, the time difference that takes between the noise sound arriving at the microphone array and the same noise arriving at the control microphones must be larger than the processing time. That is, if we want to transmit a cancellation signal from our control speakers, we must ensure that this signal arrives at the same time as the noise to be removed. This fact strongly conditions the time requirements of both the ANC subsystem and the KLM subsystem, so that the total processing time (including buffers, signal acquisition and processing) must be lower or equal to the time it takes the sound to travel from the array to the chair. This is not an easy issue, and it forces the time processing frames to be shorter.

4. EXPERIMENTS AND RESULTS

To carry out the experiments, two sound sources have been artificially mixed. First, the impulse response of each microphone-speaker pair has been estimated using an ODEON software with a sampling frequency of f_s=48000 Hz, a room of 9.13m x 4.80m x 2.64m, and a reverberation

Table 2. Results of the source separation in terms of SNR (dB), in the case of introducing two sequences of white Gaussian noise in the two sources, plus an uncorrelated ambient noise of about 20 dB of SNR in the microphones.

	Ctrl Mic 1	Ctrl Mic 2
SNR (dB) of the KLM	14.97 dB	14.94 dB

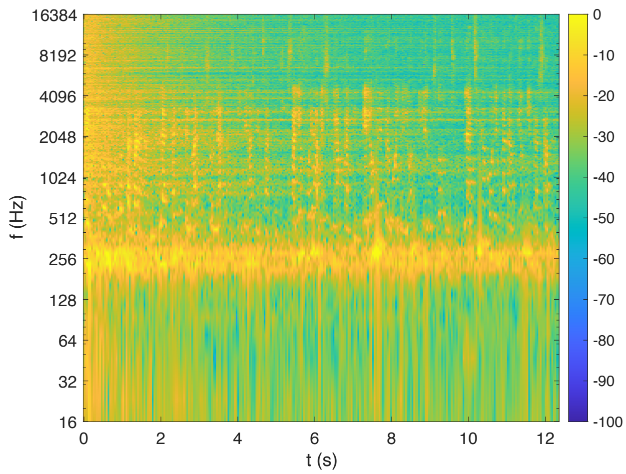

Figure 3. ANC using the isolated noise in array mic 1.

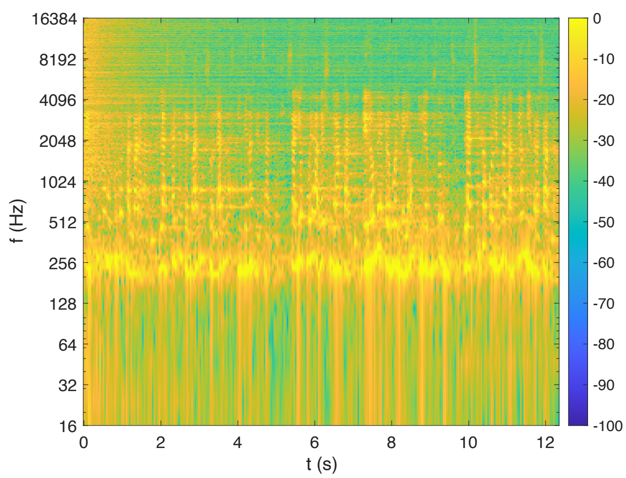

Figure 4. Combining ANC with sound source separation.

time of RT60=0.2s. Then, these impulse responses have been used to generate a mixture of two sound sources:

- The first source is a gaussian noise source, low pass filtered to contain all the energy below 850 Hz. This will be the "desired" noise source that must be removed in the spatial sound bubbles. This source will be placed in the position of the first source speaker.
- The second source is a speech from a female speaker. It will be placed in the position of the second source speaker.

Figure 2 shows the spectrogram of the mixed signals received by the first control microphone.

Once the two sources are artificially mixed, the KLM subsystem is implemented, to obtain the estimation of the noise sourced that should be received in each of the control microphones. Please note here that the information from these control microphones is not used in this process, they are only used to estimate the coefficients of the filters of the FASB.

So, the KLM subsystem obtains an estimation of the signals in the control microphones (those placed in the desired spatial bubbles). The obtained signal is estimated with an earliness of 380 samples. These 380 samples would be the margin of samples to be able to implement both the ANC and the acquisition delay of the system. As a reference, a standard nowadays Firewire low-delay audio card has an in-to-out delay of 70 samples, so the remaining 310 samples can be used to implement the ANC subsystem.

Table 2 shows the results of the sound source separation algorithm in terms of Signal-to-Noise Ratio (SNR). This SNR values have been measured at the control microphones 1 and 2. As we can see, we obtain an SNR of about 15 dB in each of the two estimates (estimate for control microphone 1 and for control microphone 2), in the case of introducing two sequences of white Gaussian noise in the two sources, plus an uncorrelated ambient noise of about 20 dB of SNR in the microphones.

These two estimated noise signals are then used to feed the ANC algorithm, and to implement the FDFA, obtaining the signals that must be reproduced by the control speakers, so that the effect of the noise is cancelled in the position of the sound bubbles (that is, in the position of the control microphones). Figure 4 shows the spectrogram of the signal received in the first control microphone, where the effect of the low frequency noise source has been mitigated.

For comparison purposes, Figure 3 shows the spectrogram of the results that could be obtained by the ANC in case a perfect noise signal were available. Comparing figures 3 and 4 we can appreciate the similarity of the ANC performance when the noise is estimated perfectly (ideal KLM), and in a real case noise estimation (real KLM).

5. CONCLUSIONS

Combining sound separation using beamforming and active noise cancellation has revolutionized our ability to isolate and extract specific sound sources from complex acoustic environments. The proposed seat combines active sound control with a comprehensive set of techniques derived from the Auditory Scene Analysis framework. These techniques enable the analysis of sound mixtures to separate individual sound sources. By spatially analyzing the position and characteristics of different sound sources within a room, the seat empowers the user to selectively mask undesirable sound sources using active sound field techniques.

6. BIBLIOGRAPHY

[1] P.A. Nelson, S, J. Elliott, “Active control of sound”. Academic press, 1991.

[2] F. Rumsey, “Sound field control”, Journal of the AES, vol. 61, no. 12, pp. 1046–1050, 2013.

[3] J. Francombe, T. Brookes, R. Mason, J. Woodcock, “Evaluation of Spatial Audio Reproduction Methods (Part 2): Analysis of Listener Preference”, Journal of the AES, vol. 65, pp. 212-225, 2017.

[4] J. Jiguang, L. Yun, “Review of active noise control techniques with emphasis on sound quality enhancement”, Applied Acoustics, vol. 136, pp. 139-148, 2018.

[5] Z. Hao, W. DeLiang, “Deep ANC: A deep learning approach to active noise control”, Neural Networks, vol. 141, pp. 1-10, 2021.

[6] L. Lu, K. L. Yin, R. C. de Lamare, Z. Zheng, Y. Yu, X. Yang, B. Chen, “A survey on active noise control in the past decade–Part I: Linear systems”, Signal Processing, vol. 181, 2021.

[7] L. Lu, K. L. Yin, R. C. de Lamare, Z. Zheng, Y. Yu, X. Yang, B. Chen, “A survey on active noise control in the past decade–Part II: Nonlinear systems”, Signal Processing, vol. 181, 2021.

[8] D. Shi, W. S. Gan, B. Lam and K. Ooi, "Fast Adaptive Active Noise Control Based on Modified Model-Agnostic Meta-Learning Algorithm", IEEE Signal Processing Letters, vol. 28, pp. 593-597, 2021.

[9] A. Mirza, A. Zeb, S.A. Sheikh, “Robust adaptive algorithm for active control of impulsive noise”, EURASIP Journal on Advances in Signal Processing, vol. 44, 2016.

[10] J. Fabry, S. Liebich, P. Vary and P. Jax, "Active Noise Control with Reduced-Complexity Kalman Filter" 16th International Workshop on Acoustic Signal Enhancement (IWAENC), pp. 166-170, 2018.

[11] S.M. Kuo, D.R. Morgan, “Active Noise Control Systems: Algorithms and DSP Implementation”, New York, NY, USA: Wiley, 1996.

[12] J. Estreder, G. Piñero, M. Ferrer, M. de Diego, A. Gonzalez, “Perceptual Active Equalization of Multi-Frequency Noise”, 18th International Conference on Signal Processing and Multimedia Applications (SIGMAP), “Best Paper Award”, 6-8 July 2021.

[13] S.M. Kuo, Y. Yang, "Broadband adaptive noise equalizer", IEEE Signal Processing Letters, vol. 3, no. 8, pp. 234-235, Aug. 1996.

[14] S.M. Kuo, J. Tsai. Residual noise shaping technique for active noise control systems, The Journal of the Acoustical Society of America, 95(3): 1665-1668, 1994.

[15] S. Pradhan, G. Zhang, X. Qiu, "A time domain decentralized algorithm for two channel active noise control", J. of the Acoustical Society of America, vol. 147, pp. 3808-3813, 2020.

[16] M. Ferrer, M. de Diego, G. Piñero and A. Gonzalez, "Affine Projection Algorithm Over Acoustic Sensor Networks for Active Noise Control," IEEE/ACM Trans. on Audio, Speech, and Language Processing, vol. 29, pp. 448-461, 2021.

[17] C.G. Lopes and A. H. Sayed, “Distributed processing over adaptive networks”, Proc. Adaptive Sensor Array Processing Workshop, MIT Lincoln Lab., Lexington, MA, June 2006.

[18] V. Molés-Cases, S. Elliott, J. Cheer, G. Piñero, and A. Gonzalez, “Weighted pressure matching with windowed targets for personal sound zones”, J. of the Acoustical Society of America 151, Online, 2022.

[19] A.S. Bregman, “Auditory scene analysis: the perceptual organization of sound.” Cambridge, MA: MIT Press, 1990.

[20] M. Rosa-Zurera, M. Utrilla-Manso, R. Gil-Pita, J. García-Gómez, “Analysis of error in sound sources position estimation with microphone arrays”, 2021 Signal Processing Symposium (SPSympo), pp. 232-237, Warsaw, September 2021.

[21] M. Bautista-Durán, J. García-Gómez, R. Gil-Pita, I. Mohíno-Herranz, M. Rosa-Zurera, “Energy-Efficient Acoustic Violence Detector for Smart Cities”. Int. Journal of Computational Intelligence Systems, Vol. 10, No. 1, pp. 49–56, 2017.

[22] D. Ayllón, R. Gil-Pita, and M. Rosa-Zurera. “Rate-constrained source separation for speech enhancement in wireless-communicated binaural hearing aids”, EURASIP Journal on Advances in Signal Processing,1, pp. 1-14, 2013.

[23] Farhang-Boroujeny, B., Adaptive Filters: Theory and Applications, Chichester, England, Wiley, 1998.

[24] M. Kajala and M. Hamalainen, "Filter-and-sum beamformer with adjustable filter characteristics," IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221), Salt Lake City, UT, USA, pp. 2917-2920 vol.5, 2001.