EFFECTIVENESS OF HORIZONTAL PERSONAL SOUND SYSTEMS FOR LISTENERS OF VARIABLE HEIGHTS

Nasim Radmanesh, Ian S Burnett

RMIT University, Melbourne, VIC, Australia
nasim.radmanesh@rmit.edu.au, ian.burnett@rmit.edu.au

ABSTRACT

Standard surround systems for generation of isolated wideband soundfields employ uniformly-spaced array of speakers in the horizontal plane. For these systems, the evaluation of sound reproduction with height is important due to listener's variable heights. Previous work demonstrated that controlling both the speakers' location and their complex weights using two-stage Lasso-LS pressure matching optimization allows isolated sound reproduction with limited number of speakers within the speakers' plane. This work demonstrates that deployment of this technique can also give up to 24dB in suppression of sound at heights between zero and one meter from speakers' plane over single-stage LS using e.g. 90 speakers in a semicircular array.

Index Terms— Isolated sound fields, Lasso-LS, listeners' height, personal sound system, wideband.

1. INTRODUCTION

Isolated sound reproduction has attracted significant research attention in the last two decades due to its broad range of applications. Examples are the provision of private sound spaces in communal areas such as dance areas, museums and planes. In such scenarios, the personal sound spaces are desired to be flexible to people's variable heights. In addition, in some cases e.g. in plane scenario it is desirable to provide isolated sound zones for both choices of sitting and lying down. Thus, an array design with limited number of speakers which provides high performance multizone sound system with height is of interest here.

Several sound reproduction methods have been recently proposed based on active sound control strategies [1]. Using this technique, the speaker weights are calculated to control the acoustic energy within bright and dark zones [2] [3] or to least squares (LS) pressure match the desired and reproduced soundfields within control zones [4], [5]. Similar scenarios are investigated in [6]-[8] for the generation of personal soundfields. Moreover, broadband beamforming techniques are used in [9] to provide a robust solution to manufacturing tolerances for focusing the sound to the user. Finally, the number of speakers required for wideband isolated sound generation was minimized in [10], [11] by the present authors through the control of both the speaker locations and their complex weights in a two-stage Lasso-LS pressure matching approach. As the personal spaces should be flexible to users of different heights, the evaluation of sound system at those heights is of paramount importance in real world scenarios. In this paper, the performance of the horizontal non-uniform array design using Lasso-LS algorithm is evaluated in terms of sound suppression at various heights in comparison to a regularly-spaced array. The robustness to limited height variation makes the horizontal sound systems usable for all listeners without the addition of more speakers at other heights.

This paper is structured as follows: In section 2 the isolated soundfields reproduction is presented and then the Lasso-LS algorithm for the speakers' positions and weightings optimization is outlined. Section 3 provides the simulation results for evaluation of the multizone system performance at variable heights using Lasso-LS in comparison to the LS approach. Section 4 concludes on the work with a discussion of the results.

2. ISOLATED SOUNDFIELDS REPRODUCTION

The aim of this section is to generate a desired field for every source in one active zone and to suppress it effectively in the other *N*-1 zones (silent zones) while all control zones are located in the same plane as virtual sources and speakers. It is assumed here that the sound field propagates under free field conditions and the virtual sources and speakers are considered to be point sources. In the following analysis the aim is to generate S isolated sound fields ($s = 1,\ldots,S$) for wideband sources (with constituent frequencies fq, $q = 1,\ldots,Q$) in N zones. The radius and angle of the sth source is r_s and θ_s, respectively and L speakers are located on a semi-circular array.

Fig.1 illustrates the task scenario with the reproduction zones located at radius R_z from the origin and the nth zone's angle given by ψ_{zn}. All zones are located within a circle of radius R_c surrounded by an array of L speakers placed on a semicircle of radius *R*. Each zone is of radius r_z with a covering of M matching points distributed uniformly

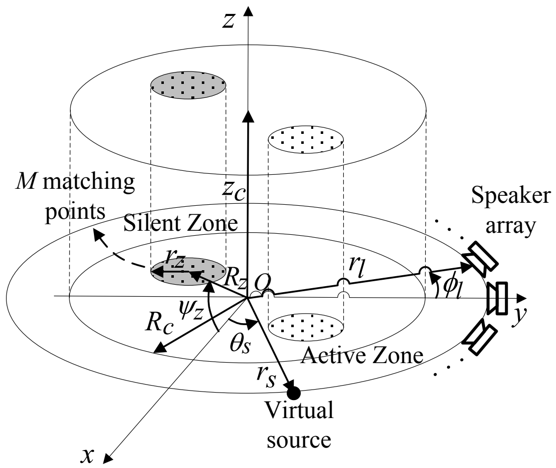


Fig.1. Diagram of reproduction of isolated sound fields with height in a multizone system using an arc of speakers.

over a Euclidean grid. For each source frequency, fq, a pressure matching approach is performed to control the complex sound pressure at the MN matching points within the zones. While the pressure amplitude is directly controlled within the N zones, the pressure outside of the control zones is limited by control of the total speakers' power. Assuming a time dependency $e^{j\omega t}$, the pressure $p_{s,q}(r_m,\phi_m,z_m)$ produced by the speakers at a given matching point m is given by [4]:

$$p_{s,q}(r_m,\phi_m,z_m)=\sum_{l=1}^{L}W_{s,q}(\ell).h_q(m,\ell) \quad (1)$$

where $W_{s,q}(\ell)$ is the ℓth speaker weight for reproduction of the sth source at frequency fq and $hq(m,\ell)$ is the three-dimensional Green's function which relates the pressure amplitude of the ℓth speaker and the pressure at the matching point m according to:

$$h_q(m,\ell)=\frac{e^{-jk_q\left|\vec{r_l}-\vec{r_m}\right|}}{\left|\vec{r_l}-\vec{r_m}\right|} \quad (2)$$

where $\left|\vec{r_l}-\vec{r_m}\right|=\sqrt{(x_l-x_m)^2+(y_l-y_m)^2+(z_l-z_m)^2}$, $k_q=2\pi fq/c$ is the acoustic wave number and c the speed of sound propagation in air. $\vec{r_l}=(r_l,\varphi_l,z_l)$ and $\vec{r_m}=(r_m,\phi_m,z_m)$ are respectively the vector positions of the speakers and matching points in cylindrical coordinates. The desired sound field $D_{s,q}(r_m,\phi_m,z_m)$ of a virtual source s located at $\vec{r_s}=(r_s,\theta_s,z_s)$ to be reproduced e.g. in the first zone is then given by:

$$D_{s,q}(r_m,\phi_m,z_m)=\begin{cases}h_q(m,s), m=1,2,\ldots,M\\ \alpha.h_q(m,s), m=M+1,\ldots,NM\end{cases} \quad (3)$$

where $hq(m,s)$ relates the pressure amplitude of the sth source and the pressure at the matching point m, the first zone is covered by the first M matching points and $\alpha=0.001$ is the sound field attenuation in inactive zones.

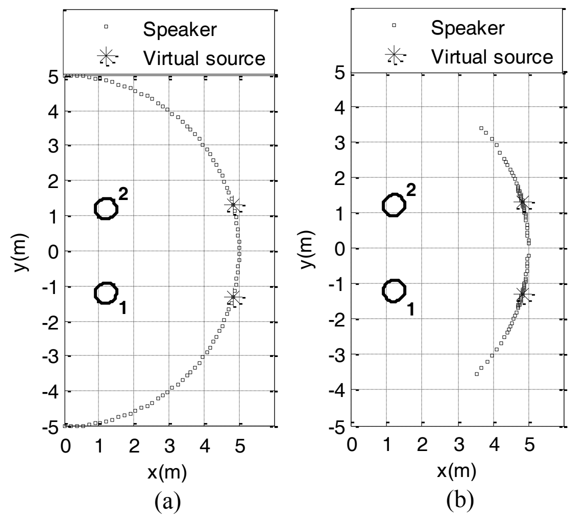


Fig.2. Speaker locations for (a) single-stage LS and (b) two-stage Lasso-LS, sources located at $\theta_1=-15°$ and $\theta_2=15°$. The number of speakers used in the reproduction of two wideband sources is identical, e.g. $L=La=90$.

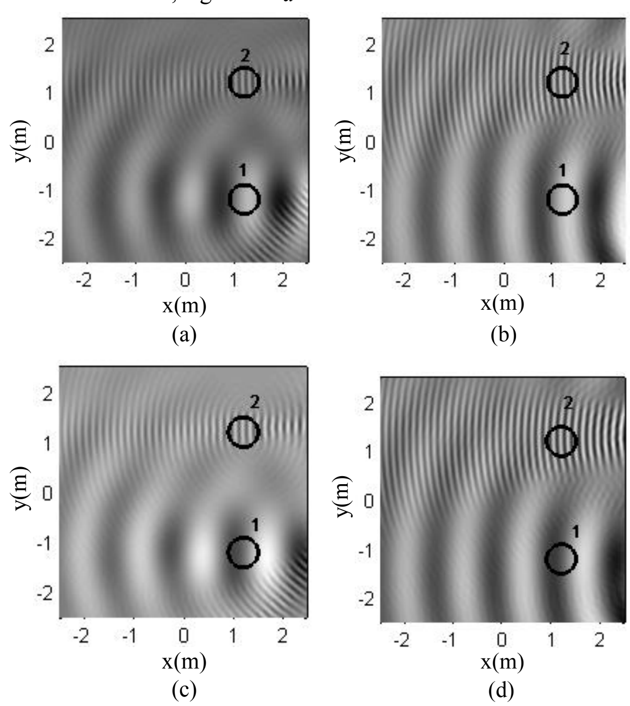
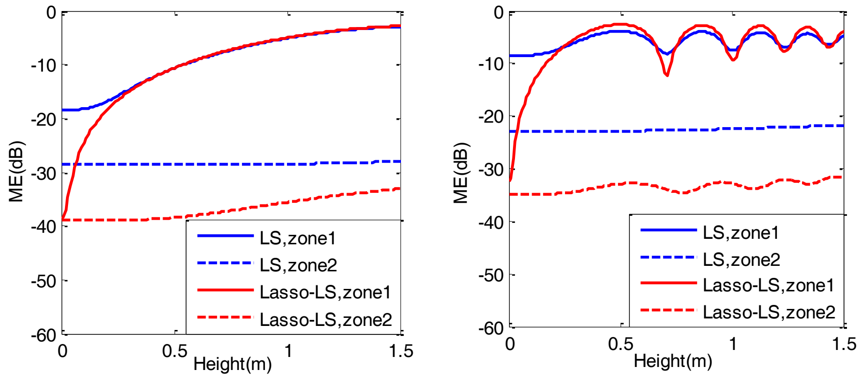


Fig.3. Sound field visualization using (a) single-stage LS at speakers' plane, (b) two-stage Lasso-LS at speakers' plane (c) single-stage LS at height $z_c=1$m from speakers' plane and (d) two-stage Lasso-LS at height $z_c=1$m from speakers' plane. Two wideband sources are located at $\theta_1=-15°$ and $\theta_2=15°$. (Single frequencies of $f_1=300Hz$ and $f_2=2.5kHz$ are shown for clarity). In both methods the number of speakers is identical, i.e. $L=La=90$.



(a) (b)

Fig.4. (a) The mean error vs height at low frequency $f_1 = 500Hz$ and (b) the mean error vs height at high frequency $f_2 = 5kHz$ for source1 at $\theta_1 = -15°$ with zone1 as the target zone for this source and zone2 as the corresponding silent zone. In the Lasso-LS algorithm, the speaker locations at horizontal plane are selected considering two wideband sources located at $\theta_1 = -15°$ and $\theta_2 = 15°$. The number of active speakers in both methods is identical i.e. $L = L_a = 90$.

The speaker weights, $Ws,q(\ell)$ are estimated using a two-stage Lasso-LS algorithm for speaker location and weight estimation in wideband sound reproduction [10].

2.1. Two stage, Combined Lasso-LS Optimization

In this paper, a two-stage Lasso-LS algorithm [10] is employed to optimize the speakers' positions and weightings for the generation of wideband soundfields. A first-stage Lasso [12] optimization was employed to select the most efficient locations from L_c potential positions in terms of achieving the lowest reproduction MSE for a limited number of speakers. A second stage regularized LS estimation was then employed for complex weighting optimization as it is theoretically guaranteed to result in the lowest MSE for the selected set of speakers

In the first stage of the Lasso-LS algorithm, *SQ* Lasso problems are solved to determine all active speakers used for reproduction of *S* wideband sound fields ($s = 1,\ldots,S$) with constituent frequencies f_q, $q = 1,\ldots,Q$. In this paper the center band frequencies of one-third octave bands [13] from 100Hz to 16kHz were used to select active speakers in the first stage Lasso algorithm. The first-stage Lasso penalty parameter λ_1 determines the number of selected active speakers. The larger the first-stage penalty parameter, λ_1, is made, the lower the number of speakers selected. The solutions of the Lasso problems are used to calculate the total speaker weights vector $\mathbf{W}_\Sigma$ as explained in [10]. The locations of the active speakers to be used in the second stage are then extracted on the basis of the nonzero entries of $\mathbf{W}_\Sigma$. The number of those nonzero entries in $\mathbf{W}_\Sigma$ thus determines the number of active speakers, L_a, to be used in the second stage. In the second stage, the non-uniformly spaced arc of L_a active speakers are utilized for sound reproduction of all constituent frequencies, f_u, $u = 1,\ldots,U$ of S wideband sources using LS optimization. The number of LS problems to be solved for generation of *S* isolated audio signals is thus *SU*. The penalty parameter δ_2 limits the power of the second-stage LS solution.

2.2. Reproduction Error

In order to compare the error performance of different methods, a reproduction mean error (ME), $E_{s,q}$, generated by every source *s* at frequency f_q in each zone at height z_c from speakers' plane is calculated as:

$$E_{s,q} = \frac{1}{A}\iint_A \left| D_{s,q}(r,\phi,z{=}0) - p_{s,q}(r,\phi,z{=}z_c) \right| d_A \qquad (4)$$

where A is the area of each zone and $D_{s,q}(r,\phi,z{=}0)$ and $p_{s,q}(r,\phi,z{=}z_c)$, $0 \le r \le R_c$, $0 \le \phi \le 2\pi$ are respectively the desired soundfield in the area A at speakers' plane and the reproduced soundfields in the area A at height z_c for the *s*th source at frequency f_q.

3. SIMULATION RESULTS

The speakers' complex weights are calculated using pressure matching approach to generate *S=2* isolated wideband sound fields within *N=2* zones located in the same plane as virtual sources and speakers. The sound reproduction performance is then evaluated in *N=2* zones at heights $z_c = 0-1.5m$ from the speakers' plane using the single-stage LS and two-stage Lasso-LS algorithms. The zones are located at $R_z = 1.7m$ from the origin and zone angles are $\psi_{z1} = -45°$ and $\psi_{z2} = 45°$. The speakers and sources radii are considered to be $r_s = R = 5m$ and *S=2* sources are located at $\theta_1 = -15°$ and $\theta_2 = 15°$. The number of matching points used in each zone of radius $r_z = 0.3m$ is $M_z = 111$. The sound field attenuation of 60dB is considered in the silent zone ($\alpha = 0.001$). One-third octave bands with centre band frequencies from 100Hz to 16kHz were used in the first

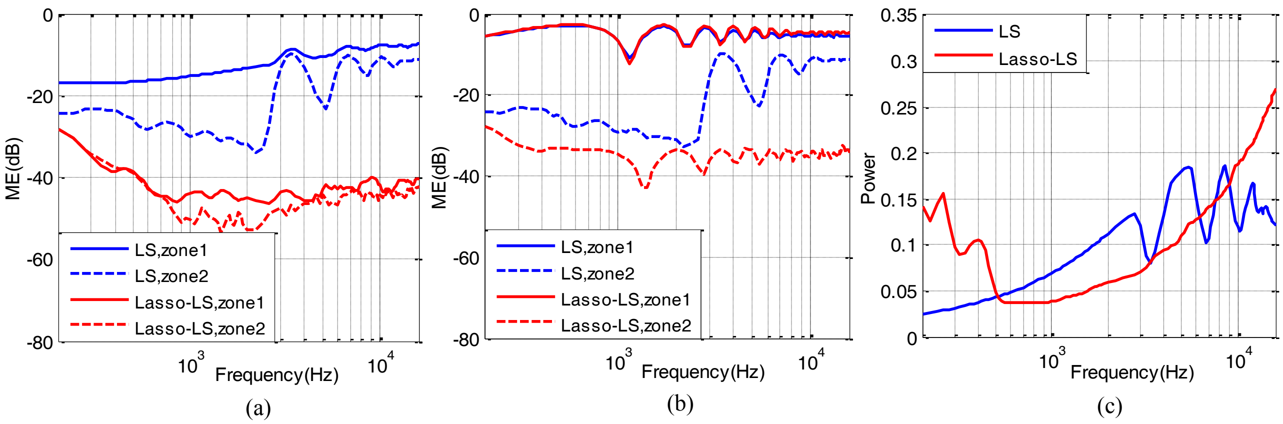


Fig.5. (a) The mean error vs frequency within control zones at speakers' plane (b) the mean error vs frequency within control zones at height $z_c = 1\text{m}$ from speakers' plane and (c) total speaker weight power vs frequency for source1 at $\theta_1 = -15^\circ$ with zone1 as the target zone for this source and zone2 as the corresponding silent zone. In the Lasso-LS algorithm, the speaker locations at horizontal plane are selected considering two wideband sources located at $\theta_1 = -15^\circ$ and $\theta_2 = 15^\circ$. The number of active speakers in both methods is identical i.e. $L = L_a = 90$.

stage Lasso algorithm and *SQ=46* sets of active speakers were selected out of $L_c = 580$ candidate positions. The unified 46 sets of speakers give a total of $L_a = 90$ active speakers which their locations are shown in Fig. 2(b). For a fair comparison between the single-stage LS and two-stage Lasso-LS algorithm in terms of wideband sound reproduction, the same number of active speakers selected in the Lasso-LS algorithm was employed in the LS method ($L = 90$) at a comparable total speakers' power. Speakers used in the single-stage LS were arranged in a uniformly-spaced array as demonstrated in Fig. 2(a). The Lasso-LS soundfield illustrated in Fig. 3(b) is 37dB more accurate than the LS soundfield in Fig.3(a) within control zones on the speakers' plane. At a height of $z_c = 1m$ from the speakers' plane, the Lasso-LS reproduced soundfield in Fig. 3(d) is still 10dB more accurate than the LS soundfield in Fig. 3(c).

Fig. 4 illustrates that the Lasso-LS ME for reproduction of a low frequency $f_1 = 500Hz$ (Fig. 4(a)) and high frequency $f_2 = 5kHz$ (Fig. 4(b)) within the control zones on the speakers' plane is up to 20dB and 24dB less than the LS error. The superior performance of Lasso-LS over LS in the active zone (zone 1) reduces with increasing height and becomes comparable at $z_c = 0.5m$ and $z_c = 0.25m$ respectively for the frequencies $f_2 = 500Hz$ and $f_2 = 5kHz$. However, the Lasso-LS approach remains up to 10dB and 12dB more successful than LS in sound suppression of frequencies $f_2 = 500Hz$ and $f_2 = 5kHz$ respectively within the silent zone (zone 2) at heights between $z_c = 0-1.5m$ from the speakers' plane. Fig. 5(a) illustrates that the Lasso-LS ME within the control zones on the speakers' plane is up to 37dB less than the LS ME across frequency. Fig 5(b) demonstrates that at a height of $z_c = 1m$, although the two methods are becoming competitive in terms of sound reproduction within the active zone, the Lasso–LS is still up to24dB more successful than LS in suppression of sound in the silent zone. This is due to the capability of the Lasso-LS algorithm to suppress sound within the silent zone on the speakers' plane directly and suppression its leakage to the non-optimized area at height $z_c = 0-1.5m$ indirectly with locating a limited number of speakers on a horizontal plane at LS-optimal locations. Fig 5(c) shows the total speaker weight power resulting from the LS and Lasso-LS techniques across frequency.

4. CONCLUSION

To implement a multizone sound system, listening height should be considered in the reproduction of isolated soundfields to make the personal sound system suitable to listeners at different heights. In this paper, a semicircular non-uniformly spaced array was employed with its speakers' locations and weightings calculated on the bases of a two-stage Lasso-LS optimization.

The results show that locating the speakers on a horizontal plane at the LS-optimal locations not only improves the performance of the multizone system by up to 37dB over a single-stage LS optimization on the speakers's plane but also can give up to 24dB in sound suppression at heights between zero and one meter from the speakers' plane using e.g. 90 speakers. This makes the technique appropriate for realistic soundfield installations without the addition of more speakers away from the speaker plane.

5. ACKNOWLEDGMENTS

This work has been supported by the Australian Research Council (ARC) through the grant DP1094053.

6. REFERENCES

[1] P. A. Nelson and S. J. Elliott, *Active Control of Sound*, Academic Press, 1993.

[2] J. Choi and Y. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources," *J. Acoust. Soc. Am.,* vol. 111, no. 4, pp. 1695–1700, Apr. 2002.

[3] M. Shin, S. Lee, F. Fazi, P. Nelson, D. Kim, S. Wang, K. Park, and J. Seo, "Maximization of acoustic energy difference between two spaces," *J. Acoust. Soc. Am*., vol. 128, no. 1, pp. 121-131, Jul. 2010.

[4] M. Poletti, "An Investigation of 2D Multizone Surround Sound Systems," *in Proc. AES 125th Convention. Audio Eng. Society*, San Francisco, CA, Oct. 2008.

[5] N. Radmanesh, I.S. Burnett, "Reproduction of Independent Narrowband Soundfields in a Multizone Surround System and its Extension to Speech Signal Sources," *in Proc. IEEE Int. Conf. Acoust, Speech, Signal Process., ICASSP'2011*, Prague, Czech Republic, May 22-27, 2011.

[6] W. F. Druyvesteyn and J. Garas, "Personal sound," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 685–701, Sep. 1997.

[7] S. J. Elliott and M. Jones, "Active headrest for personal audio," *J. Acoust. Soc. Am.,* vol. 119, no. 5, pp. 2702–2709, May 2006.

[8] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," *J. Acoust. Soc. Am.* , vol. 125, no. 4, pp. 2091–2097, Apr. 2009.

[9] I. Tashev, J. Droppo, M. Seltzer, and A. Acero, "Robust design of wideband loudspeaker arrays," *in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, *ICASSP'2008,* Las Vegas, Nevada, Mar 30 - Apr 4, 2008.

[10] N. Radmanesh and I.S. Burnett, "Generation of Isolated Wideband Sound Fields Using a Combined Two-stage Lasso-LS Algorithm," *IEEE Audio, Speech, Language Process.*, vol. 21, no. 2, pp. 378-387, Feb 2013.

[11] N. Radmanesh, I.S. Burnett, "Wideband Sound Reproduction in a 2D Multizone System Using a Combined Two-stage Lasso-LS Algorithm" *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop, SAM 2012,* Hoboken, NJ, Jun 17-20, 2012.

[12] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society*, Series B, 58(1):267–288, 1996.

[13] *Specification for Octave-Band and Fractional-Octave-Band Analog and Digital Filters*, ANSI s1.11-2004, Feb. 2004, Standards Secretariat Acoustical Society of America.