Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Signal Processing in Acoustics
Session 1aSP: Array Signal Processing for Three-Dimensional Audio Applications I

1aSP9. Design and implementation of a personal audio system in a car cabin

Jordan Cheer* and Stephen Elliott

***Corresponding author's address: Institute of Sound and Vibration Research, University of Southampton, Southampton, SO17 1BJ, Hampshire, United Kingdom, j.cheer@soton.ac.uk**

The generation of personal listening zones in a car cabin would allow the different occupants to listen to different audio programmes without the use of headphones. This would allow, for example, the driver to listen to a navigation system whilst the rear passengers watched a film. Personal audio systems have previously been implemented in mobile devices and monitors, for example, however, the investigation of the effects of an enclosure on the generation of personal listening zones has been limited. This paper presents an investigation of the effects of a car cabin sized enclosure on the generation of independent listening zones in the front and rear seats. The standard car audio loudspeaker array is used to produce independent listening zones at low frequencies, while a second array of small loudspeakers positioned at the four headrest positions is used to provide control over the rest of the audio bandwidth. The proposed arrays are implemented in a real car and the results of a real-time implementation are presented.

Published by the Acoustical Society of America through the American Institute of Physics

Received 11 Jan 2013; published 2 Jun 2013

INTRODUCTION

The development of personal audio systems that aim to reproduce sound over a specified region of space, whilst minimising the sound reproduced in other regions, have seen a significant amount of interest over recent years. This is due to the rapid increase in both the number of methods via which media, including video, audio and telecommunications, may be received. Due to the desire to create personalised listening zones, there has been a number of personal audio systems proposed for different applications such as in mobile devices [1, 2] and video monitors [3].

In the car cabin environment there is also a desire to implement personalised rear seat entertainment, in part due to the introduction of flat panel displays [4]. Based on previous work on the generation of personal listening zones, this paper describes a loudspeaker array system which is designed to achieve personal listening zones in the front and rear seats of a car cabin. This system might allow the driver and front seat passenger to listen to the radio, whilst the rear seat passengers watch a film, for example. The least squares array optimisation strategy is first presented with a constraint on the individual loudspeaker drive signals. The proposed loudspeaker array geometry is then detailed along with the microphone array that is used to define the two listening zones. A method of designing practical filters is detailed and the performance of the proposed system is presented.

LEAST SQUARES ARRAY OPTIMISATION

The performance of personal audio systems can be quantified by the acoustic contrast [5], which can be defined as the ratio between the sum of the squared pressures in the bright, or listening zone, to that in the dark, or quiet zone. The vectors of complex pressures at a given frequency in the bright and dark zones are given by

$$\boldsymbol{p}_B = \boldsymbol{Z}_B \boldsymbol{q} \qquad \text{and} \qquad \boldsymbol{p}_D = \boldsymbol{Z}_D \boldsymbol{q}, \tag{1}$$

respectively, where $\boldsymbol{q}$ is the vector of M complex signals driving the source array and $\boldsymbol{Z}_B$ and $\boldsymbol{Z}_D$ are the $(L_B \times M)$ and $(L_D \times M)$ matrices of transfer impedances from the input to each source to each of the L_B bright zone locations and L_D dark zone locations respectively. The acoustic contrast can then be expressed at a single frequency as

$$C = \frac{L_D \boldsymbol{p}_B^H \boldsymbol{p}_B}{L_B \boldsymbol{p}_D^H \boldsymbol{p}_D} = \frac{L_D \boldsymbol{q}^H \boldsymbol{Z}_B^H \boldsymbol{Z}_B \boldsymbol{q}}{L_B \boldsymbol{q}^H \boldsymbol{Z}_D^H \boldsymbol{Z}_D \boldsymbol{q}}. \tag{2}$$

The acoustic contrast can be used directly to optimise the signals driving the loudspeaker array in order to maximise the difference in level between the bright and dark zones [5], however, it does not constrain the phase of the pressures in the listening zone and, therefore, may degrade the audio quality [6]. Recent studies of personal audio systems [6, 7] have therefore used the least squares optimisation method, which has been widely used in the context of sound reproduction [8]. This optimisation method allows the phase as well as the magnitude of the pressures in the bright zones to be controlled.

In the least squares optimisation, the transfer impedances to the bright and dark zones are combined to form a single $(L \times M)$ matrix

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_B \\ \boldsymbol{Z}_D \end{bmatrix}, \tag{3}$$

where $L = L_B + L_D$. The vector of pressures at the L control points is then given by

$$\boldsymbol{p} = \boldsymbol{Z}\boldsymbol{q}. \tag{4}$$

The least squares optimisation is formulated by defining a vector of target pressures, $\boldsymbol{p}_T$, which the system aims to produce at the control points. Using this vector of target pressures, an error vector may be defined as

$$\boldsymbol{e} = \boldsymbol{p}_T - \boldsymbol{p}, \tag{5}$$

and by minimising these errors in a least squares sense it is possible to optimise the driving signals, $\boldsymbol{q}$, such that the array produces an approximation of the targeted sound field.

It is also possible to introduce constraints on the individual driving signals and the cost function to minimise can then be expressed as

$$J_{LS} = \boldsymbol{e}^H \boldsymbol{e} + \sum_{m=1}^{M} \lambda_m \left(|q_m|^2 - E_m \right), \tag{6}$$

where λ_m is the Lagrange multiplier governing the constraint on the m-th driving signal, q_m is the signal driving the m-th source and E_m is the defined limit on the individual modulus squared driving signal. Substituting equations 4 and 5 into equation 6, differentiating J_{LS} with respect to the real and imaginary parts of $\boldsymbol{q}$, equating this to zero and rearranging gives the vector of optimal driving signals as

$$\boldsymbol{q} = \left[\boldsymbol{Z}^H \boldsymbol{Z} + \boldsymbol{\lambda}_M \right]^{-1} \boldsymbol{Z}^H \boldsymbol{p}_T, \tag{7}$$

where $\boldsymbol{\lambda}_M$ is a diagonal matrix of M Lagrange multipliers that can be independently set to fulfil specified constraints on each loudspeaker. The physical limits of control have been explored in computer simulations [7] and the results of a practical implementation are described here.

Specification of the Car Cabin Personal Audio System

The vehicle within which the personal audio system has been implemented was a right-hand drive, small people carrier which has a capacity for seven passengers as can be seen from the plan view presented in Figure 1. The interior dimensions of the car are approximately 3 m × 1.8 m × 1.2m, with an internal volume of approximately 6.48 m^3. The aim of the personal audio system was to produce independent listening zones at the two front seats and at the three rear seats – the sound field produced at the two extreme rear seats was not considered.

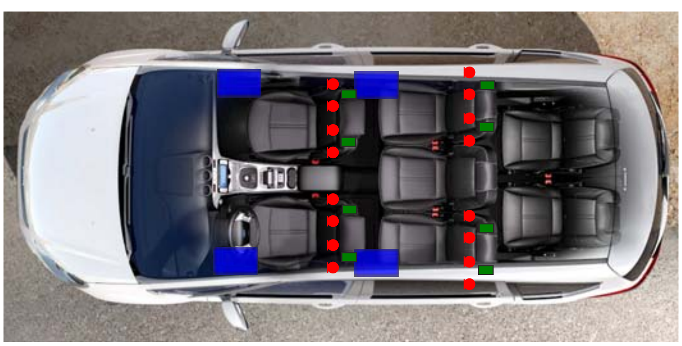


Figure 1: Plan view of the small people carrier showing the positions of the KEF loudspeakers positioned close to the standard car audio loudspeakers' positions as blue rectangels, the phase-shift headrest loudspeakers as green rectangles and the microphones are shown in red (Original image from [9]).

Based on simulations of loudspeaker arrays in a rectangular walled enclosure [7] two loudspeaker arrays have been implemented in the car cabin. The first array is the standard car audio loudspeaker array which aims to achieve sound field control at low frequencies and the second is a headrest array employing directional loudspeakers, which aims to achieve control at higher frequencies. This combined array avoids the potential safety, cost and weight issues of mounting full-range loudspeakers at the headrest positions.

The car audio loudspeaker array consists of four KEF B200G drivers which have a 183 mm diameter cone and were mounted in closed-back cabinets with an internal volume of approximately 0.01 m^3. These loudspeakers were positioned adjacent to the standard car audio loudspeakers in the front and rear of the car as shown by the blue rectangles in Figure 1. These loudspeakers were used instead of the standard car audio loudspeakers to avoid unnecessary difficulty in accessing the standard car audio loudspeakers' connectors.

The headrest array consists of eight phase-shift loudspeakers, as described in [10], with one loudspeaker mounted to each side of each headrest as shown in the plan presented in Figure 1 and in the photo in Figure 2. A photo of an individual phase-shift loudspeaker is presented in Figure 3. As can be seen from Figure 3, the phase-shift loudspeaker enclosure consists of a rear opening in which a resistive material is placed. By specifying the dimensions of the enclosure and the rear opening, and the resistive and compliant properties of the material covering the rear port, the directivity response of the phase-shift loudspeaker can be altered [10, 11, 12]. The phase-shift loudspeakers employed here were designed to produce a hypercardioid directivity pattern, which minimises the radiated sound power in a freefield environment [1]. The resistive material positioned covering the rear opening was a fine metal gauze and the size of the opening was empirically determined using a prototype phase-shift loudspeaker with a variable size rear opening. The directivity index of the implemented phase-shift loudspeaker measured in an anechoic chamber is shown in Figure 3b along with the directivity index of a theoretical hypercardioid source. The directivity index was measured as the ratio of the squared pressure produced on-axis to the average squared pressures produced at 24 additional positions evenly distributed on a circle surrounding the source in the horizontal plane. From this plot it can be seen that the directivity index of the phase-shift loudspeaker is close to that of a hypercardioid source at frequencies between around 200 Hz, where the loudspeaker begins to operate effectively, and 1 kHz. At frequencies between around 1 kHz and 3.5 kHz the directivity index is negative indicating that the phase-shift loudspeaker radiates more efficiently to the rear of the device. Despite the limited directivity index of the phase-shift loudspeakers over this frequency range, it is expected that at these frequencies the diffraction effects introduced by the positioning of the loudspeakers in close proximity to the car seats and headrests will significantly alter their directivities. These effects would require a significant amount of further work to fully understand. The low directivity index over the mid frequency range has been solved in subsequent work by introducing high frequency absorption at the rear opening of the phase-shift loudspeaker [6].

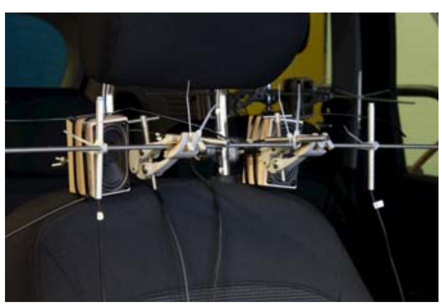
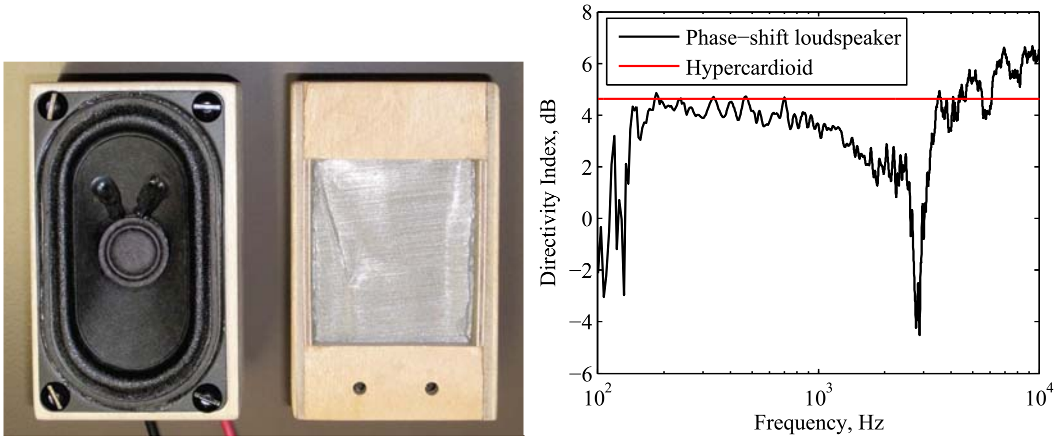


FIGURE 2: The loudspeaker array and four microphones positioned at one of the headrests.

To define the bright and dark zones an array of microphones was used. To avoid the generation of small listening zones around individual microphones it is necessary to employ a sufficient number and distribution of microphones. Four microphones were positioned at each headrest, as shown in Figures 1 and 2, and, therefore, the two listening zones were each defined



(A) Front and rear view of an individual phase-shift loudspeaker.

(B) Directivity index of the phase-shift loudspeaker and a theoretical hypercardioid source.

FIGURE 3: Details of the phase-shift loudspeaker.

by 8 microphones. The inner two microphones were separated by around 12 cm and the outer two microphones were spaced a further 8 cm from these microphones. Additionally, to provide some spread in the other two dimensions, the outer microphones were positioned around 4 cm above and 4 cm behind the inner two microphones.

Using the least squares optimisation method defined by equation 7, it is necessary for each of the loudspeaker arrays to define a vector of target pressures. The target pressures in the bright zone for the car audio loudspeaker array have been defined as the pressures produced in the bright zone when the four car audio loudspeakers are driven in-phase. The target pressures in the bright zone for the headrest loudspeaker array have been defined as the pressures produced in the bright zone when the four headrest sources in the nearfield of the bright zone are driven in-phase. For both arrays the target pressures in the dark zone have been defined as zero.

FILTER DESIGN

To calculate the optimal filter frequency responses the transfer responses have first been measured between the voltage input to each of the 12 loudspeakers and the resulting pressures measured at the 16 microphones. These transfer responses have then been used as the appropriate elements of the $\boldsymbol{Z}_B$ and $\boldsymbol{Z}_D$ matrices and the optimal filter responses have been calculated in the frequency domain according to equation 7. A number of constraints have been defined and are summarised in Table 1. These constraints are enforced by specifying the matrix of Lagrange multipliers, $\boldsymbol{\lambda}_M$. The two constraints detailed in columns f_{min} and f_{max} of Table 1 ensure that the loudspeakers are not driven at high levels at frequencies outside of their operating frequency ranges and the third constraint, which acts on the maximum acoustic contrast, has been introduced to ensure that the arrays do not attempt to achieve unnecessarily high levels of contrast. High levels of contrast tend to require high driving voltages and to produce small bright and dark zones concentrated around the microphones. An acoustic contrast of 15 dB has been chosen as an initial target level, based on the results of the subjective tests presented in [13], which indicate that the minimum required level difference between a desired audio programme and an interfering audio programme is around 11 dB.

Using the optimal filters calculated at each frequency using equation 7, practical time-domain filter responses have been calculated according to the following method, which

TABLE 1: Filter optimisation constraints

Array	f_{min}	f_{max}	Maximum Contrast
Car audio	20 Hz	200 Hz	15 dB
Headrest	200 Hz	20kHz	15 dB

follows that previously been presented in [2]:

1. The optimal frequency responses of each filter have been calculated at each frequency using equation 7 and the responses have beeen windowed to ensure zero level at 0 Hz and the Nyquist frequency (24 kHz in this case).

2. The frequency domain windowed responses have been Inverse Fast Fourier Transformed (IFFT) to give impulse responses of length

$$I = \left(\frac{Fs}{\Delta f} + 1\right), \tag{8}$$

   where Fs is the sample rate and Δf is the separation between adjacent frequency points in the optimal filter frequency responses. Δf is dependent on the length of the Fast Fourier Transform (FFT) used in the calculation of the transfer responses $\boldsymbol{Z}_B$ and $\boldsymbol{Z}_D$ from the measured response data. Δf has been set to 1.4648 Hz, which ensures the IFFT has largely decayed to zero over its duration, i.e. 682 ms.

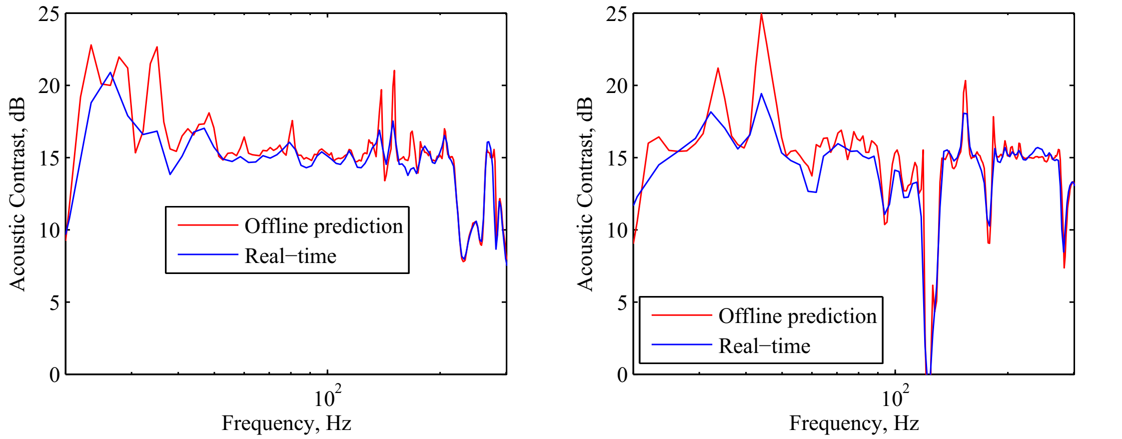
3. The filter impulse responses have been shifted to accommodate for the periodicity of the FFT.

4. At this stage a practicable set of filters is available, however, since this method of filter design imposes no constraint on causality, the filters have a response of length $(I-1)/2$ before zero time. This response can be implemented using a modelling delay and, for the employed sample rate and frequency point separation, a modelling delay of 341 ms is required. However, this leads to practical issues in applications such as two-way telecommunications and such a significant pre-echo leads to subjectively poor audio quality [14]. Therefore, it is desirable to truncate the length of the filters in the time domain to reduce the required modelling delay. For both arrays the filters have been truncated to the shortest length possible without significant reductions in the predicted acoustic contrast or significant enhancements in the array effort. For the car audio array the filter impulse responses have been truncated to $I = 25000$ which requires a modelling delay of 260 ms and for the headrest array the filters have been truncated to $I = 2000$ which requires a modelling delay of 21 ms.

5. The truncated filter impulse responses have been windowed using a Hanning window to avoid artefacts due to non-zero responses at the start and end of the filters.

SYSTEM PERFORMANCE

Using the time-domain filters designed, as detailed above, the performance of the car audio loudspeaker array has been predicted for the case when a bright zone is produced in either the front or the rear seating region. The resulting acoustic contrast is shown by the red lines in Figures 4a and 4b respectively. From these plots it can be seen that the car audio loudspeaker array largely achieves the 15 dB acoustic contrast target over the 20 – 200 Hz bandwidth, although there is a dip in the contrast for the rear right zone at around 120 Hz due to the optimisation of λ_M at this frequency not reaching the optimum value within the specified limits of the optimisation routine. This problem could be avoided by allowing the optimisation routine to run for a larger number of iterations.



(A) Acoustic contrast between *front bright* and dark rear zones.

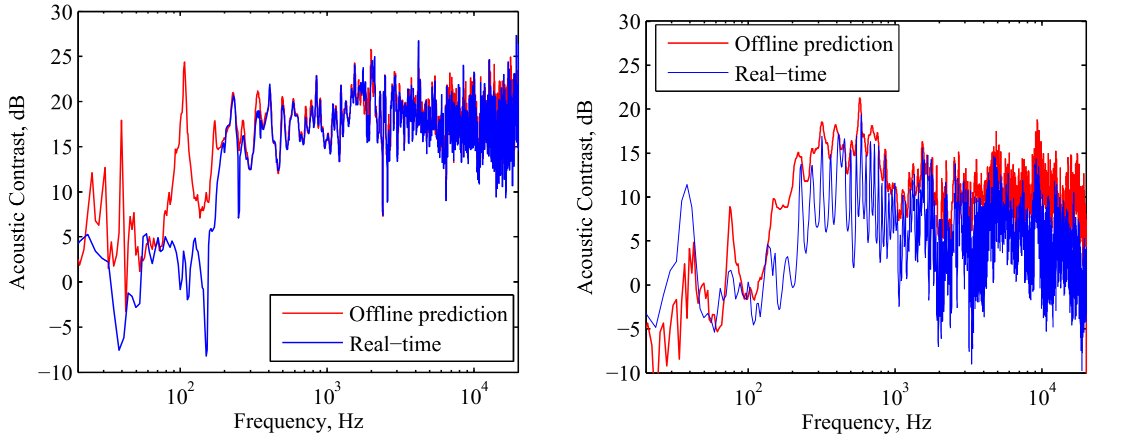
(B) Acoustic contrast between *rear bright* and dark front zones.

FIGURE 4: The acoustic contrast and array effort plotted as a function of frequency for the *car audio loudspeaker array* predicted using offline simulations (red) and measured in real-time (blue).

To confirm the practical performance of the car audio loudspeaker array using the designed filters the acoustic contrast performance has been measured in real-time by driving the four loudspeakers simultaneously with a pink noise signal filtered using the associated truncated filters and measuring the resulting pressures at the 16 microphones. The measured performance of the car audio array when producing a bright zone in either the front or rear seating regions is shown by the blue lines in Figures 4a and 4b respectively. From these results it can be seen that the real-time performance achieved by the array is close to the predicted levels. The level of acoustic contrast achieved for both bright zones is practically useful and produces subjectively impressive level differences in informal listening tests.

The performance of the headrest loudspeaker array using the time-domain truncated filters has also been predicted for the two bright zone scenarios and the results are shown by the red lines in Figures 5a and 5b. From Figure 5a it can be seen that when producing a bright zone in the front seats the proposed headrest loudspeaker array employing the filters designed as detailed above mostly exceeds the 15 dB contrast level between 200 Hz and 20 kHz. However, when a bright zone is produced in the rear seating region the predicted level of contrast shown in Figure 5b struggles to meet the 15 dB target level. The limited performance of the headrest array to produce a rear bright zone is related to the forward-direction of the phase-shift loudspeakers so that it is difficult to control forward radiation from the rear loudspeakers using the front loudspeakers.

Despite the limited predicted performance of the headrest array when producing a rear bright zone, the real-time performance of the array has been measured as for the car audio loudspeaker array and the results are shown in Figure 5 by the blue lines. From Figure 5a it can be seen that when producing a front bright zone the real-time performance is close to the predicted performance at frequencies above around 200 Hz, where the loudspeakers begin to operate effectively. The blue line in Figure 5b shows the measured performance of the array producing a rear bright zone and it can be seen that there are more significant differences between the predictions and measured results. This can be related to variations in the positions of the loudspeakers between the transfer response measurements and the real-time



(A) Acoustic contrast between *front bright* and dark rear zones.

(B) Acoustic contrast between *rear bright* and dark front zones.

FIGURE 5: The acoustic contrast and array effort plotted as a function of frequency for the *headrest loudspeaker array* predicted using offline simulations (red) and measured in real-time (blue).

measurements. These variations were not introduced systematically and, therefore, their magnitude is unknown, however, it does highlight the practical issue of robustness in personal audio systems, as discussed in a theoretical context in [15].

CONCLUSIONS

This paper has presented a description of the design and implementation of a personal audio system in a car cabin. The system attempts to produce two independent listening zones – one in the front seats and one in the rear seats – and thus enable the front seat occupants to listen to one audio programme while the rear seat passengers listen to another.

The least squares, frequency domain array optimisation strategy has first been detailed with the capability of constraining the magnitude of the individual driving signals. The least squares optimisation method has been employed since it has been reported to achieve improved audio quality [6] and, with suitably selected target pressures, improved numerical conditioning [7] compared to alternative personal audio optimisation strategies.

To achieve two independent listening zones over the full audio bandwidth two loudspeaker arrays have been proposed. At frequencies below 200 Hz an array of four low-frequency loudspeakers, positioned adjacent to the standard car audio loudspeakers, has been employed. At higher frequencies this array is not able to achieve sound field control due to the increasing number of dominant acoustic enclosure modes and, therefore, a second array of 8 small directional loudspeakers have been positioned at the four headrest positions. By combining the standard car audio loudspeakers with a small array of headrest loudspeakers avoids potential safety, cost and weight issues of mounting full-range loudspeakers at the headrests.

The optimisation of the driving signals using the least squares method is performed in the frequency domain and, therefore, a method of calculating practical time-domain filters has been presented. The car audio loudspeaker array has been shown to be capable of achieving significant levels of acoustic contrast at frequencies between 20 and 200 Hz, however, this requires a long filter response with a long modelling delay. This may lead to issues in employing the proposed array in two-way telecommunications applications and, more importantly, may lead to poor audio quality due to audible pre-echos. The headrest loudspeaker array has also been shown to achieve significant levels of acoustic contrast with a significantly shorter filter response, although when producing a rear bright zone, the performance is limited due to the directivity of the phase-shift loudspeakers and their orientation relative to the control zones. It has also been highlighted that the performance of the headrest loudspeaker is susceptible to variations in the loudspeaker positions, although this requires further investigation.

REFERENCES

[1] S. J. Elliott, J. Cheer, H. Murfet, and K. R. Holland, "Minimally radiating arrays for mobile devices", J. Acoust. Soc. Am. **128**, 1721–1728 (2010).

[2] J. Cheer, S. J. Elliott, Y.-T. Kim, and J.-W. Choi, "Practical implementation of personal audio in a mobile device", J. Audio Eng. Soc. (provisionally accepted September 2012).

[3] 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. **125**, 2091–2097 (2009).

[4] B. Crockett, E. Benjamin, and M. Smithers, "Next generation automotive sound research and technologies", in *120th Audio Engineering Society Convention* (Paris, France) (2006), URL `http://www.aes.org/e-lib/browse.cfm?elib=13453`.

[5] J.-W. Choi and Y.-H. Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources", J. Acoust. Soc. Am. **111**, 1695–1700 (2002).

[6] M. F. Simón Gálvez, S. J. Elliott, and J. Cheer, "A superdirective array of phase-shift sources", J. Acoust. Soc. Am. **132**, 746–756 (2012).

[7] J. Cheer, "Active control of the acoustic environment in an automobile cabin", Ph.D. thesis, University of Southampton, Southampton, UK (2012).

[8] P. Nelson, F. Orduna Bustamante, and H. Hamada, "Multichannel signal processing techniques in the reproduction of sound", J. Audio Eng. Soc. **44**, 973–989 (1996).

[9] Ford, "Ford S-Max – spacious and versatile interior design - ford UK", `http://www.ford.co.uk/Cars/S-MAX/Interiordesign` (2012 (last accessed 8/10/2012)).

[10] M. F. Simón Gálvez, "Loudspeaker arrays for family television", Master's thesis, University of Southampton, UK (2011).

[11] J. Cheer, "Designing loudspeaker directivity for mobile devices", Master's thesis, University of Southampton, UK (2009).

[12] T. J. Holmes, "The acoustic resistance box - a fresh look at an old principle", J. Audio Eng. Soc. **34**, 981–989 (1986).

[13] W. F. Druyvesteyn, R. M. Aarts, A. J. Asbury, P. Gelat, and A. Ruxton, "Personal sound", in *Proceedings of the Institute of Acoustics*, volume 16, 571 (1994).

[14] P. Marins, F. Rumsey, and S. Zielinski, "The relationship between basic audio quality and selected artefacts in perceptual audio codexs - part II: validation experiment", in *122nd Audio Engineering Society Convention* (Vienna, Austria) (2007).

[15] S. J. Elliott, J. Cheer, Y.-T. Kim, and J.-W. Choi, "Robustness and regularisation of personal audio systems", IEEE Trans. Audio, Speech, Lang. Process. **20**, 2123–2133 (2012).