



STEREOPHONIC PERSONAL AUDIO REPRODUCTION USING PLANARITY CONTROL OPTIMIZATION

Philip Coleman, Philip J. B. Jackson, Marek Olik

*Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, Surrey, GU2 7XH, United Kingdom,
e-mail: p.d.coleman@surrey.ac.uk*

Jan Abildgaard Pedersen

Bang & Olufsen A/S, Peter Bangs Vej 15, DK7600, Struer, Denmark

Sound field control to create multiple personal audio spaces (sound zones) in a shared listening environment is an active research topic. Typically, sound zones in the literature have aimed to reproduce monophonic audio programme material. The planarity control optimization approach can reproduce sound zones with high levels of acoustic contrast, while constraining the energy flux distribution in the target zone to impinge from a certain range of azimuths. Such a constraint has been shown to reduce problematic self-cancellation artefacts such as uneven sound pressure levels and complex phase patterns within the target zone. Furthermore, multichannel reproduction systems have the potential to reproduce spatial audio content at arbitrary listening positions (although most exclusively target a ‘sweet spot’). By designing the planarity control to constrain the impinging energy rather tightly, a sound field approximating a plane-wave can be reproduced for a listener in an arbitrarily-placed target zone. In this study, the application of planarity control for stereo reproduction in the context of a personal audio system was investigated. Four solutions, to provide virtual left and right channels for two audio programmes, were calculated and superposed to achieve the stereo effect in two separate sound zones. The performance was measured in an acoustically treated studio using a 60 channel circular array, and compared against a least-squares pressure matching solution whereby each channel was reproduced as a plane wave field. Results demonstrate that planarity control achieved 6 dB greater mean contrast than the least-squares case over the range 250-2000 Hz. Based on the principal directions of arrival across frequency, planarity control produced azimuthal RMSE of 4.2/4.5 degrees for the left/right channels respectively (least-squares 2.8/3.6 degrees). Future work should investigate the perceived spatial quality of the implemented system with respect to a reference stereophonic setup.

1. Introduction

Approaches to sound zone reproduction may be broadly categorized as energy cancellation approaches or sound field synthesis (SFS) approaches [1]. Energy cancellation techniques [2, 3] optimize the loudspeaker weights based on some function of the squared pressure in the zones, meaning that the phase is uncontrolled. This means that although excellent levels of contrast between zones can be obtained, there is no opportunity to intentionally create spatial effects. SFS approaches, on

the other hand, strictly specify the sound field and it is therefore straightforward to specify interesting spatial sound fields. Often, complex sound scenes are rendered in this way with massive multichannel sound systems [4]. Classical analytical SFS approaches such as mode matching and wave field synthesis may be applied to sound zone reproduction by means of coefficient translation between the local (zone) sound fields and the global sound field [5]. Alternatively, multi-point SFS approaches such as least-squares pressure matching (PM) [6] may minimize the error between a desired sound field at a number of microphone positions and the reproduced pressures. Least-squares approaches may also be weighted to achieve improved contrast [7]. It is convenient to use these approaches as the filters can be calculated based on measured transfer functions, lifting the tight constraints on loudspeaker and microphone array geometries and allowing room reflections to be considered in the optimization. Similarly, least-squares optimization may be adopted for the bright zone, with alternate constraints on the dark zone weights to create contrast [8, 9]. Energy cancellation and multi-point SFS approaches to sound zones have been compared in [1] for circular arrays, and in [10] for a line array. Although the SFS approaches to sound zone reproduction give the potential for spatial audio, such a system (synthesizing multiple plane wave directions per zone) has not previously been realized (although a superdirective array was used to create two-channel audio in [11]).

An alternative optimization approach planarity control (PC) was proposed in [12]. This approach is similar in concept to the wavenumber domain point focusing (WDPF) [13] in that the bright zone energy is focused towards being a plane-wave field by projecting the energy in to a spatial domain. However, in WDPF there is no cancellation region. PC was shown in [12] to reduce self-cancellation artefacts sometimes arising when energy cancellation methods are adopted, and it was suggested that by suitably narrowing the angular pass-range, plane wave energy impinging from a certain direction could be produced.

As with the SFS approaches, in theory any number of plane wave components can be approximated using PC. Here, two components are superposed to approximate stereo reproduction. As such, the aim is to reproduce two virtual sources placed at $\pm 30^\circ$, corresponding to the left and right loudspeakers in a conventional stereo setup. This situation is illustrated in Fig. 1. The extension to stereo represents a significant advance in demonstrating the potential of personal audio systems, as stereophonic reproduction has been used for decades in consumer audio systems. It greatly enhances the width of the reproduced audio, and depending on the zone geometry and virtual source locations, may reduce binaural unmasking and improve the perceived level difference over the monaural case.

To realize stereo, four sets of sound zone filters are required (target zone A, left and right channels; target zone B, left and right channels). The two most important properties for stereo sound zone reproduction are the acoustic contrast between the zones and the preservation of interaural level and phase differences (ILD, IPD). Acoustic contrast is strongly linked to the listening experience in a sound zone as it aims maximally to suppress the interference of the alternate zone's programme [14, 15], and for multi-channel audio it is especially important as any residual sound pressure in the dark zone will be summed. To preserve the spatial sound scene, the ILD and IPD should be accurately reproduced. Here, a physical approach is used whereby the microphone array-based spatial filtering is used to observe the direction from which the virtual source energy impinges on the target zone.

Using PC for stereo reproduction is advantageous for a number of reasons. As with PM, it is applicable to arbitrary (including irregular) loudspeaker arrays, and incorporates room compensation in to the cost function. However, the freedom in optimizing energy rather than explicitly defined phase means that PC can reproduce contrast over a wider bandwidth and with lower effort than PM. Furthermore, the listening regions may be arbitrarily placed within the array, and the concepts could be extended for moving listeners. In this paper, measured results are presented, comparing PM and PC for stereo personal audio reproduction applied to two static zones surrounded by a 60 channel circular loudspeaker array.

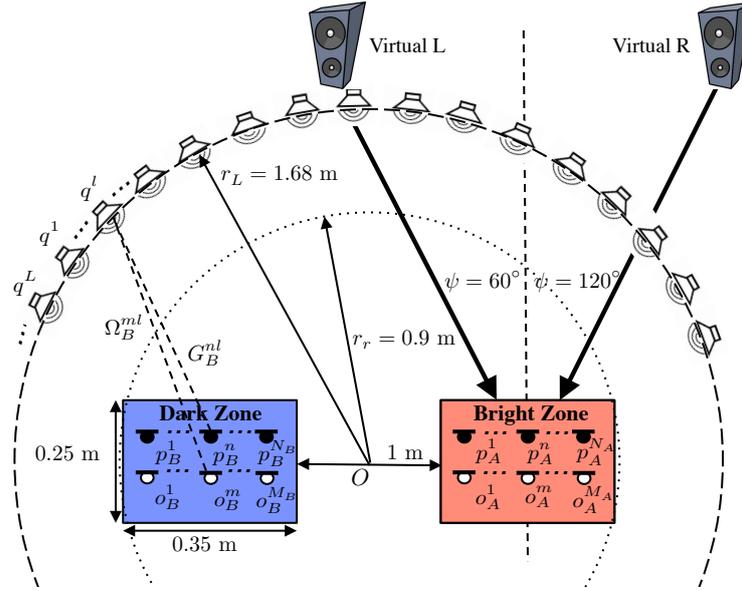


Figure 1: Notation and system geometry with L loudspeakers, zones A and B comprising N_A and N_B control microphones (black) and M_A and M_B monitor microphones (white) respectively. The transfer functions G and Ω are also shown. The concept of stereo reproduction is indicated with the plane wave directions ψ corresponding to the loudspeaker positions.

2. Background

Figure 1 shows an example sound zone system layout. Two audio programs A and B are to be reproduced in zones A and B, respectively. The rest of the room is uncontrolled. The zones (defined by the control microphone positions) and loudspeakers may be placed arbitrarily in the room. For each frequency, the source weights can be written in vector notation as $\mathbf{q} = [q^1, q^2, \dots, q^L]^T$, where there are L loudspeakers and q^l is the complex source weight of the l th loudspeaker. Similarly, the complex pressures at the control microphone positions in zones A and B are written as $\mathbf{p}_A = [p_A^1, p_A^2, \dots, p_A^{N_A}]^T$ and $\mathbf{p}_B = [p_B^1, p_B^2, \dots, p_B^{N_B}]^T$ respectively, where there are N_A control microphones in zone A and N_B in zone B, and the complex pressures at the n th microphones in each zone are p_A^n and p_B^n . The observed pressures at the monitor microphones in each zone are denoted as $\mathbf{o}_A = [o_A^1, o_A^2, \dots, o_A^{M_A}]^T$ and $\mathbf{o}_B = [o_B^1, o_B^2, \dots, o_B^{M_B}]^T$ respectively, where there are M_A monitor microphones in zone A and M_B in zone B, and the complex pressures at the m th microphones in each zone are o_A^m and o_B^m . Spatially distinct microphones are used in order to reduce possible bias due to measurement of performance at the exact control positions.

The plant matrices contain the transfer functions between each loudspeaker and microphone, and are considered with respect to the control and monitor microphones in each zone. For zone A they are defined as

$$\mathbf{G}_A = \begin{pmatrix} G_A^{11} & \cdots & G_A^{1L} \\ \vdots & \ddots & \vdots \\ G_A^{N_A 1} & \cdots & G_A^{N_A L} \end{pmatrix}, \quad \mathbf{\Omega}_A = \begin{pmatrix} \Omega_A^{11} & \cdots & \Omega_A^{1L} \\ \vdots & \ddots & \vdots \\ \Omega_A^{M_A 1} & \cdots & \Omega_A^{M_A L} \end{pmatrix}, \quad (1)$$

where G_A^{nl} and Ω_A^{ml} are the transfer functions between the n th control microphone and the m th monitor microphone in zone A, respectively, and the l th loudspeaker. The equivalent notation is used for \mathbf{G}_B and $\mathbf{\Omega}_B$. The pressures at the microphone positions may be written as $\mathbf{p}_A = \mathbf{G}_A \mathbf{q}$, $\mathbf{o}_A = \mathbf{\Omega}_A \mathbf{q}$, $\mathbf{p}_B = \mathbf{G}_B \mathbf{q}$ and $\mathbf{o}_B = \mathbf{\Omega}_B \mathbf{q}$.

3. Theory

In the following, PM and PC optimizations are introduced [6, 12]. These utilize constraints on the sum of squared pressures in zone A and the sum of squared source weights. The former can be expressed as $A = N_A |p_r|^2 \times 10^{T/10}$, where T is the target spatially averaged level in decibels relative to the threshold of hearing $p_r = 20 \mu\text{Pa}$.

3.1 Pressure matching

The desired virtual sources can be synthesized as plane wave sound fields, $\mathbf{d}_A = D_A e^{jk\mathbf{r}_n \cdot \mathbf{u}_\varphi}$, for $n = 1, 2, \dots, N_A$, where D_A gives the pressure amplitude, \mathbf{r}_n is the position of the n th control microphone in zone A, \cdot denotes the inner product, and \mathbf{u}_φ is the unit vector in the direction of the incoming plane wave. The desired zone B sound field is given by a vector of length N_B populated with zeros, $\mathbf{d}_B = \mathbf{0}$. The cost function, with a constraint to fix the effort to a certain E , is [6]:

$$J_{PM} = (\mathbf{p}_A - \mathbf{d}_A)^H (\mathbf{p}_A - \mathbf{d}_A) + \mathbf{p}_B^H \mathbf{p}_B + \lambda (\mathbf{q}^H \mathbf{q} - E). \quad (2)$$

Using the method of Lagrange multipliers the solution can be found by taking the derivatives with respect to \mathbf{q} and λ :

$$\mathbf{q} = (\mathbf{G}_A^H \mathbf{G}_A + \mathbf{G}_B^H \mathbf{G}_B + \lambda \mathbf{I})^{-1} \mathbf{G}_A^H \mathbf{d}_A; \quad \mathbf{q}^H \mathbf{q} = E. \quad (3)$$

The Lagrange multiplier λ is numerically chosen to satisfy the control effort constraint, and it is assumed that the solution is appropriately scaled by setting $\mathbf{d}_A^H \mathbf{d}_A = A$.

3.2 Planarity control

The PC optimization cost function can be introduced as a minimization of the dark zone pressures, with the bright zone energy constraint enforced via the spatial domain, and with an effort constraint [12]:

$$J_{PC} = \mathbf{p}_B^H \mathbf{p}_B + \mu (\mathbf{p}_A^H \mathbf{Y}_A^H \mathbf{\Gamma} \mathbf{Y}_A \mathbf{p}_A - A) + \lambda (\mathbf{q}^H \mathbf{q} - E), \quad (4)$$

where μ and λ are Lagrange multipliers, \mathbf{Y}_A is an $I \times M_A$ steering matrix populated by superdirective beamforming as in [16], and $\mathbf{\Gamma} = \text{diag}[\gamma_1, \gamma_2, \dots, \gamma_I]$, with $0 \leq \gamma_i \leq 1$ the weighting corresponding to the i th steering angle. By setting the weightings appropriately, we can attempt to place the virtual source at a certain angle $i = \varphi$. The solution is found by taking the derivatives with respect to \mathbf{q} and each of the Lagrange multipliers, and setting to zero:

$$\mu \mathbf{q} = (\mathbf{G}_A^H \mathbf{Y}_A^H \mathbf{\Gamma} \mathbf{Y}_A \mathbf{G}_A)^{-1} (\mathbf{G}_B^H \mathbf{G}_B + \lambda \mathbf{I}) \mathbf{q}; \quad \mathbf{p}_A^H \mathbf{Y}_A^H \mathbf{\Gamma} \mathbf{Y}_A \mathbf{p}_A = A; \quad \mathbf{q}^H \mathbf{q} = E. \quad (5)$$

The optimal source weights are proportional to the eigenvector corresponding to the maximum eigenvalue of $(\mathbf{G}_B^H \mathbf{G}_B + \lambda \mathbf{I})^{-1} (\mathbf{G}_A^H \mathbf{Y}_A^H \mathbf{\Gamma} \mathbf{Y}_A \mathbf{G}_A)$. The values of the Lagrange multipliers are determined iteratively, where the sum of squared pressures (projected via the spatial domain) is fixed to satisfy the constraint $A = \mathbf{p}_A^H \mathbf{Y}_A^H \mathbf{\Gamma} \mathbf{Y}_A \mathbf{p}_A$, with $\lambda = 0$. Then, λ is chosen such that the constraint on $\mathbf{q}^H \mathbf{q}$ is satisfied. If $E > \mathbf{q}^H \mathbf{q}$ when $\lambda = 0$, the constraint is not active. Otherwise, λ is determined numerically using a gradient descent search such that $\mathbf{q}^H \mathbf{q} \leq E$, with A being fixed at each step.

4. Reproduction system realization

A reproduction and measurement system was designed and mounted on a bespoke spherical structure, the ‘‘Surrey Sound Sphere’’, placed in an acoustically treated room of dimensions $6.55 \times 8.78 \times 4.02$ m (RT60 235 ms averaged over 0.5 kHz, 1 kHz and 2 kHz octave bands). The loudspeakers (Genelec 8020b) were clamped to the equator of the sphere to form a 60 channel circular array (radius of 1.68 m, as Fig. 1), and 48 microphones (Countryman B3 omni) were attached to a

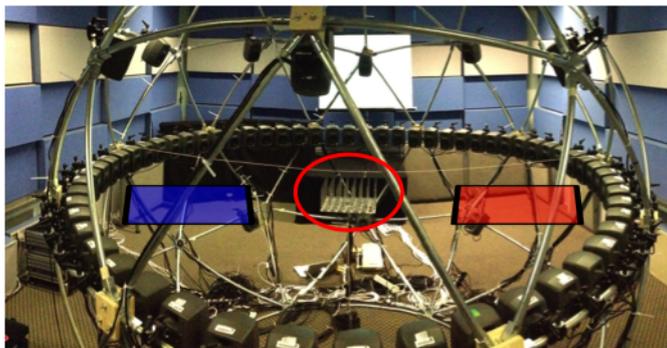


Figure 2: Photograph of the experimental system showing the circular loudspeaker array and microphone grid (circled), with the approximate zone positions overlaid (not to scale).

grid mounted on a microphone stand. In order to achieve the required sampling density of microphone locations, 8 positions of the microphone stand were measured per zone. A photograph of the equipment is shown in Fig. 2. A Mac Pro computer running Matlab was used to play the audio and also to record the signals from the microphones, via the ‘playrec’ utility. A 72 channel MOTU PCIe 424 sound card was used for the analogue to digital interface, with the microphone inputs first passed through a pre-amplifier stage (PreSonus Digimax D8). Level differences between the input and output signal channels were compensated through calibration. Room impulse responses (RIRs) between each microphone position and each loudspeaker were measured using the maximum length sequence (MLS) approach (15th order) and cropped at 150 ms. Finite impulse response (FIR) filters were populated and measured by considering a bin-by-bin approach. The RIRs were first down-sampled to the simulation sample rate of 20 kHz, and a 8192 point fast Fourier transform (FFT) was taken. The source weights were collated for each frequency bin, the negative frequency bins populated by complex conjugation, and the inverse FFT taken to obtain a time-domain filter. Regularization was applied by initializing λ (Eqs. (3) and (5)) such that the condition number of the matrix to be inverted did not exceed 10^{10} , before enforcing a control effort limit of 0 dB relative to a single loudspeaker equidistant from both zones reproducing the same sound pressure level (76 dB) in the bright zone [1]. A 4096 sample modelling delay was applied to ensure causality. Measurements of objective performance were made by convolving an MLS sequence with each of the FIR control filters, simultaneously replaying them through the loudspeakers, and sampling the reproduced sound pressures with the microphone array. Finally, the FFT was taken of the recorded system responses, and the evaluation metrics were calculated in the frequency domain.

5. Performance

The sound pressure level difference will be evaluated using the metric of acoustic contrast, which describes the attenuation achieved between the bright zone and the dark zone, and is therefore of paramount importance for assessing sound zone algorithms. It is defined as the ratio of spatially averaged pressures in each zone due to the reproduction of program A, expressed in decibels:

$$\text{Contrast} = 10 \log_{10} \left(\frac{M_B \mathbf{o}_A^H \mathbf{o}_A}{M_A \mathbf{o}_B^H \mathbf{o}_B} \right). \quad (6)$$

To quantify the accuracy of the virtual source placement, the root mean square error (RMSE) in degrees between the desired location and the measured location, based on the principal direction of the energy impinging on the bright zone, was used. The measured sound pressures are related to the energy flux distribution as $\mathbf{w} = \frac{1}{2} |\mathbf{H}_A \mathbf{o}_A|^2$, where \mathbf{H}_A is a $I \times M_A$ steering matrix [16].

The main advantage of using PC over PM for stereophonic personal audio reproduction is in terms of the cancellation achieved for each channel. The measured acoustic contrast values of the

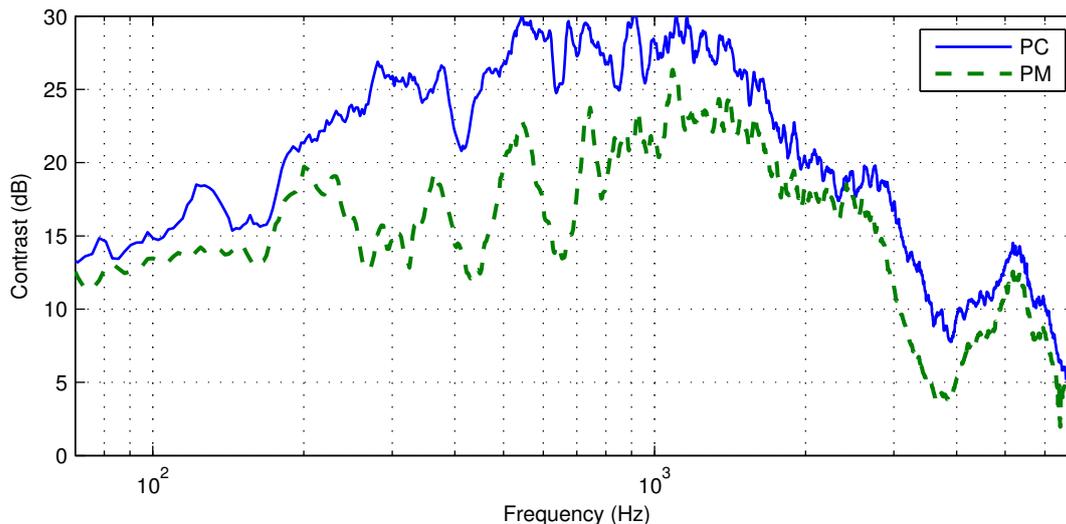


Figure 3: Combined acoustic contrast for the stereo scenario, for PC (solid) and PM (dashed), based on individual measurements of the left and right channel sound pressure levels. The plot shows data smoothed (after combining the channels) using a 15-bin wide moving average filter.

combined left and right channels are shown in Fig. 3. At all frequencies above 70 Hz, PC produced a greater acoustic contrast than PM, with an improvement of at least 3 dB between 200–2000 Hz, and greater than 10 dB improvement at some frequencies below 1 kHz. PC contrast was above 15 dB over 100–3100 Hz, whereas for PM the range was much narrower (650–2800 Hz). The upper frequency of contrast performance is given by $f_{max} = cL/4\pi r_r$, i.e. the projected half-wavelength spacing around the reproduction radius. For this configuration, accurate reproduction is expected up to approximately 1800 Hz, and it is indeed evident that the measured contrast of both methods drops above this frequency. However, PC produces up to 5 dB more contrast than PM even in this range.

Figure 4 shows the distribution of the energy impinging on the zones as a function of azimuth, showing PC and PM at 500, 1000 and 2000 Hz. The target window for PC is also indicated. The fundamental result shown in Fig. 4 is that the normalized energy peak was correctly located for both PC and PM. This result generalized across significant portions of the frequency range tested. In comparing PM and PC, additional energy sidelobes were present for PC. It is not clear what kind of perceptual impact these sidelobes have on the quality of the stereo image achieved.

The mean placement RMSEs over various frequency ranges are shown in Table 1. The energy distribution over azimuth for PC was compromised at some frequencies above the array aliasing limit by the additional cancellation it achieved, whereas PM tended to still produce the desired target field at the cost of contrast. This effect was especially noticeable for the PC left channel, where in the range 100–7000 Hz the RMSE was 40.0° for PC and 4.1° for PM. For the same filter set between 100–1800 Hz (i.e. below the spatial aliasing limit for the array), the errors were 11.4° (PC) and 12.2° (PM), which are more comparable. For the right channel, the RMSEs over 100–7000 Hz were 6.3° and 5.2° for PC and PM, respectively. Some inflation of the RMSE across this frequency range may also be attributed to the beamformer resolution, and the RMSEs in the range 500–1800 Hz were 4.6° (left) and 3.7° (right) for PC, and 2.0° (left) and 1.9° (right) for PM. These values are, for both methods, in the same order of magnitude as human localization, reported in [17] to have a mean accuracy of 5° at 30° azimuth (standard deviation of 2°).

The large RMSE for the PC left channel can be explained in terms of the angle between the left channel beam and the dark zone, which is larger than the corresponding angle with the right channel beam. Spatial aliasing for PC was problematic due to grating lobes (across the dark zone) emerging in the frequency range 2.3–3 kHz. At these frequencies, the priority cancellation in PC led to sound

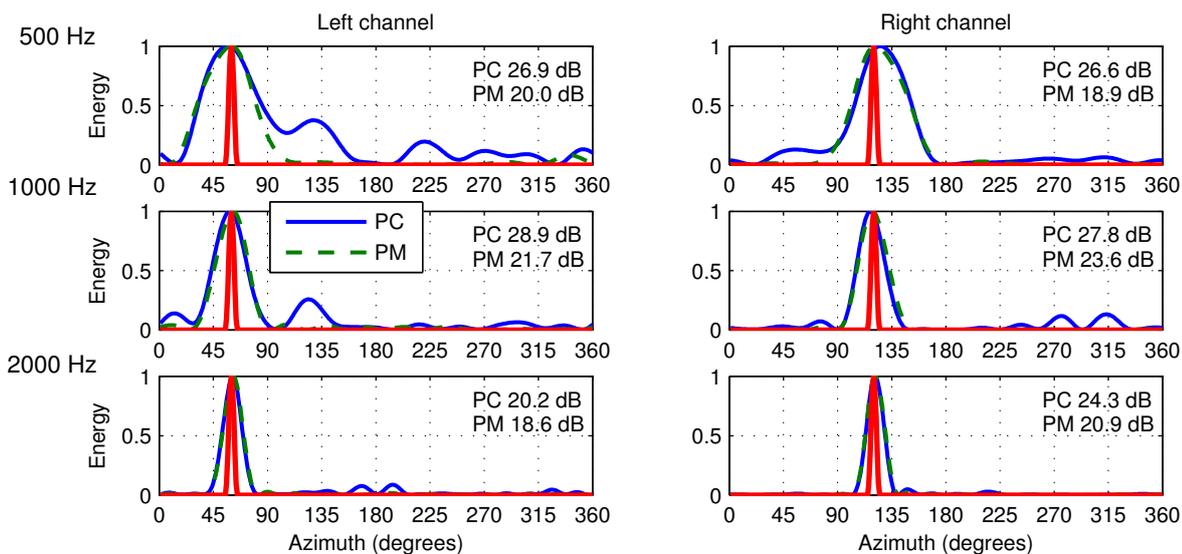


Figure 4: Performance of PC and PM for placing plane wave energy at angles corresponding to the left and right channels of a stereo loudspeaker setup, at 500 Hz, 1 kHz and 2 kHz. The target window for PC is indicated by the thick red line. The corresponding (measured, single channel) acoustic contrast values are given in each case.

Table 1: Mean RMSE (left and right channel) and combined contrast values for PC and PM for the stereophonic application, showing the effect of the frequency band on performance.

	RMSE (L, deg)		RMSE (R, deg)		Contrast (dB)	
	PC	PM	PC	PM	PC	PM
100–7000 Hz	40.0	4.1	6.3	5.2	15.7	12.1
100–1800 Hz	11.4	8.1	12.2	10.2	26.0	19.9
500–1800 Hz	4.6	2.0	3.7	1.9	27.0	21.3

fields reminiscent of energy cancellation techniques [1, 12] being reproduced. The principal energy directions therefore switched from the desired 60° placement towards a mode of operation whereby significant energy components impinged on the zone from around 180° . However, in order to satisfy the bright zone energy constraint (Eq. (4)), a significant (but not the principal) component of energy was placed in the desired location. The effect of such energy distributions in a minority of frequency bands on source localization has yet to be investigated. On the other hand, the right channel beam was able to reproduce the direction constraint while also steering the grating lobes away from the dark zone, thereby giving improved accuracy over the left channel.

6. Summary

Reproduction of stereophonic programme material for personal audio by rendering two virtual loudspeakers, while creating a cancellation region, was investigated. Measured performance results comparing PC and PM were presented based on a 60 channel circular array in a reflective room. At frequencies up to the array aliasing limit, PC and PM produced comparable RMSEs in terms of the principal energy direction impinging on the bright zone, with PC producing 6.1 dB better mean contrast over 100–1800 Hz. At higher frequencies for PC, some energy was placed at the desired location, but this was not always the principal direction, however PM still produced the specified plane wave relatively accurately. The contrast was limited by the physical distribution of the loudspeakers for both methods, with PC producing slightly greater contrast than PM. The perceptual properties of

stereophonic reproduction conducted in this way, considering both localization and interference, are an interesting and necessary topic of further work.

7. Acknowledgements

This research has arisen from the POSZ project, www.posz.org, supported by Bang & Olufsen A/S, and has also been supported by EPSRC Grant Ref. EP/L000539/1 (S3A).

REFERENCES

- ¹ Coleman, P., Jackson, P. J. B., Olik, M., Møller, M., Olsen, M., and Pedersen, J. A., “Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array,” *J. Acoust. Soc. Am.* **135**(4) (2014). <http://dx.doi.org/10.1121/1.4866442>.
- ² Choi, J. and Kim, Y., “Generation of an acoustically bright zone with an illuminated region using multiple sources,” *J. Acoust. Soc. Am.* **111**, 1695–1700 (2002).
- ³ Shin, M., Lee, S. Q., Fazi, F. M., Nelson, P. A., Kim, D., Wang, S., Park, K. H., and Seo, J., “Maximization of acoustic energy difference between two spaces,” *J. Acoust. Soc. Am.* **128**, 121–131 (2010).
- ⁴ Spors, S., Wierstorf, H., Raake, A., Melchior, F., Frank, M., and Zotter, F., “Spatial sound with loudspeakers and its perception: A review of the current state,” *Proc. IEEE* **101**(9), 1920–1938 (2013).
- ⁵ Wu, Y. and Abhayapala, T., “Spatial multizone soundfield reproduction: Theory and design,” *IEEE Trans. Audio Speech Lang. Proc.* **19**(6), 1711–1720 (2011).
- ⁶ Poletti, M., “An investigation of 2-d multizone surround sound systems,” in *Proc. 125th AES Conv., San Francisco, CA, 2-5 October 2008*, (2008).
- ⁷ Chang, J.-H. and Jacobsen, F., “Sound field control with a circular double-layer array of loudspeakers,” *J. Acoust. Soc. Am.* **131**(6), 4518–4525 (2012).
- ⁸ Betlehem, T. and Teal, P., “A constrained optimization approach for multi-zone surround sound,” in *IEEE Int. Conf. Acoust., Speech Signal Proc. (ICASSP), Prague, 22-27 May 2011*, pp. 437–440 (2011).
- ⁹ Cai, Y., Wu, M., and Yang, J., “Sound reproduction in personal audio systems using the least-squares approach with acoustic contrast control constraint,” *J. Acoust. Soc. Am.* **134**(2), 734–741 (2014).
- ¹⁰ Simón Gálvez, M. F., Elliott, S. J., and Cheer, J., “A superdirective array of phase shift sources,” *J. Acoust. Soc. Am.* **132**(2), 746–756 (2012).
- ¹¹ Park, J.-Y., Chang, J.-H., and Kim, Y.-H., “Generation of independent bright zones for a two-channel private audio system,” *J. Audio Eng. Soc.* **58**(5), 382–393 (2010).
- ¹² Coleman, P., Jackson, P. J. B., Olik, M., and Pedersen, J. A., “Optimizing the planarity of sound zones,” in *Proc. 52nd AES Int. Conf., Guildford, UK, 2-4 September 2013*, (2013).
- ¹³ Chang, J.-H., Choi, J.-W., and Kim, Y.-H., “A plane wave generation method by wave number domain point focusing,” *J. Acoust. Soc. Am.* **128**(5), 2758–2767 (2010).
- ¹⁴ Baykaner, K., Hummersone, C., Mason, R., and Bech, S., “The prediction of the acceptability of auditory interference based on audibility,” in *Proc. 52nd AES Int. Conf., Guildford, UK, 2-4 September 2013*, (2013).
- ¹⁵ Francombe, J., Mason, R., Dewhirst, M., and Bech, S., “Determining the threshold of acceptability for an interfering audio programme,” in *Proc. 132nd AES Conv., Budapest, Hungary, 26-29 April 2012*, (2012).
- ¹⁶ Jackson, P. J. B., Jacobsen, F., Coleman, P., and Pedersen, J. A., “Sound field planarity characterized by superdirective beamforming,” in *Proceedings of Meetings on Acoustics*, **19**, 055056 (2013).
- ¹⁷ Makous, J. C. and Middlebrooks, J. C., “Two-dimensional sound localization by human listeners,” *J. Acoust. Soc. Am.* **87**(5), 2188–2200 (1990).