TIME-DOMAIN ACOUSTIC CONTRAST CONTROL WITH A SPATIAL UNIFORMITY CONSTRAINT FOR PERSONAL AUDIO SYSTEMS

Sipei Zhao and Ian S. Burnett

Centre for Audio, Acoustics and Vibration, Faculty of Engineering and IT, University of Technology Sydney, NSW, Australia

Sipei.Zhao@uts.edu.au, Ian.Burnett@uts.edu.au

ABSTRACT

Personal audio systems with multiple sound zones for listeners to enjoy different music/audio contents privately in a shared physical space have attracted great research interest in the past two decades. Acoustic Contrast Control (ACC) is one of the most popular methods for generating multiple personal sound zones because it produces the minimum inter-zone interference. However, the ACC method has been found to be inferior to the pressure matching method in terms of sound quality due to an uneven frequency response and nonuniform spatial sound field distribution in the bright zone. This paper proposes a spatial uniformity constraint on time-domain broadband ACC in addition to the frequency response trend estimation constraint with the aim of ensuring a uniform sound field distribution in the bright zone. Simulation results with measured room impulse responses demonstrate that the proposed algorithm reduces the magnitude variations in the bright zone to be less than 1 dB higher than the just noticeable level difference at a cost of a perceptually negligible degradation in acoustic contrast.

Index Terms— Personal audio systems, acoustic contrast control, spatial uniformity, personal sound zone

1. INTRODUCTION

Personal audio systems, which are also described as multizone sound field reproduction, aim to generate multiple personal sound zones where people can listen to their own music/audio content without disturbing each other in a public space. The approach eliminates the inconvenience and discomfort of wearing headphones for long periods and has wide applications in mobile devices [1]–[3], personal computers [4], car cabins [5]–[7], and outdoor concerts [8] *etc*.

Various methods have been developed to generate personal sound zones, such as Acoustic Contrast Control (ACC) [9], Pressure Matching (PM) [10], Mode Matching (MM) [11], and weighted Pressure Matching (wPM) which is a combination of the ACC and PM methods [12]. Among these methods, ACC is the most commonly used because it minimizes the inter-zone interference by generating the highest acoustic potential energy contrast between the bright and dark zones. This has been recently proved theoretically under the variable span linear filtering framework [13].

The ACC method was originally formulated in the frequency domain [9] and its robustness has been thoroughly investigated [14]. However, although the ACC method generates a high acoustic contrast between zones, it does not provide a uniform sound field distribution over the bright zone, resulting in potentially unsatisfactory sound quality [15], [16]. Therefore, planarity control [17] and a sound differences constraint [18] have been applied to the original ACC methods to reduce the sound pressure variation in the bright zone. These frequency-domain methods were optimized at each frequency and the optimal filter coefficients obtained by the Inverse Fourier Transform (IFT) which results in two disadvantages. One is the non-causal nature and need for a modelling delay of half of the filter length [19] and the other is the poor acoustic contrast performance at non-control frequencies, which is especially serious when the filter length is short [20].

To overcome these problems, the broadband ACC (BACC) method has been developed in the time domain directly [21]. Although BACC ensures causality, it suffers from uneven frequency response due to its tendency to extract a specific frequency component with the highest contrast [22]. To alleviate this problem, various methods have been explored to increase the flatness of the frequency response, among which the BACC-RV method [20] appears to be the first. In the BACC-RV method, a Response Variation (VR) term with respect to the response at a reference frequency was introduced as a constraint on the BACC cost function [20]. However, the performance of BACC-RV depends on the choice of the reference frequency, hence the RV term is replaced with a Response Differential (RD) term, which is a measure of the summed sound pressure differences between consecutive frequencies, in the BACC-RD method [23]. Similarly, the BACC-RTE method [24] modifies the RD term to a Response Trend Estimation (RTE) term to sum the sound pressure differences between further frequencies.

The abovementioned time-domain BACC algorithms have improved the frequency responses dramatically, but the

spatial sound field distribution in the bright zone was not controlled. This paper proposes the BACC-RTE-SUC method, which applies a Spatial Uniformity Constraint (SUC) in addition to the RTE term to the BACC cost function so as to improve both the frequency response and the spatial sound field uniformity in the bright zone. Simulations with measured Room Impulse Responses (RIRs) are performed to validate the proposed approach.

2. THEORY

A personal audio system, as illustrated in Fig. 1, utilizes an array of loudspeakers to generate an acoustic bright zone and a dark zone. The input signal $x[n]$ (n denotes the time index) is filtered with a Finite Impulse Response (FIR) filter $\mathbf{w}_l$ (l = 1, 2, …, L, L is the total number of loudspeakers) with a length of I before being reproduced through each loudspeaker. The RIR from the l-th loudspeaker to the m-th microphone in the control zones is modelled as a FIR filter $\mathbf{h}_{ml}$ with a length of K.

To investigate the frequency response of the system, the input signal $x[n]$ is assumed to be a Dirac delta function [23], hence the sound pressure at the m-th microphone due to the l-th loudspeaker is written as

$$p_{\mathrm{B},ml}[n] = \sum_{i=0}^{I-1} h_{\mathrm{B},ml}[n-i]\, w_l[i], \tag{1}$$

where the subscript $_\mathrm{B}$ denotes the bright zone. Arranging the sound pressure of all the time indices into a vector, *i.e.*, $\mathbf{p}_{\mathrm{B},ml} = [p_{\mathrm{B},ml}[0], p_{\mathrm{B},ml}[1], \ldots, p_{\mathrm{B},ml}[I+K-2]]^\mathrm{T}$, (1) can be expressed concisely as [22]

$$\mathbf{p}_{\mathrm{B},ml} = \mathbf{H}_{\mathrm{B},ml}\mathbf{w}_l, \tag{2}$$

where $\mathbf{w}_l = [w_l[0], w_l[1], \ldots, w_l[I-1]]^\mathrm{T}$ and

$$\mathbf{H}_{\mathrm{B},ml} = \begin{bmatrix} h_{\mathrm{B},ml}[0] & 0 & 0 \\ \vdots & \ddots & 0 \\ h_{\mathrm{B},ml}[K-1] & \ddots & h_{\mathrm{B},ml}[0] \\ 0 & \ddots & \vdots \\ 0 & 0 & h_{\mathrm{B},ml}[K-1] \end{bmatrix} \tag{3}$$

is a matrix of dimension $(I + K - 1) \times I$ [22]. It should be noted that $P_{\mathrm{B},ml}$ in (1) is in fact the global impulse response with a length of $(I + K - 1)$.

Summing up the contribution from L loudspeakers to the m-th microphone, one obtains [22]

$$\mathbf{p}_{\mathrm{B},m} = \sum_{l=1}^{L} \mathbf{p}_{\mathrm{B},ml} = \sum_{l=1}^{L} \mathbf{H}_{\mathrm{B},ml}\mathbf{w}_l = \mathbf{H}_{\mathrm{B},m}\mathbf{w}, \tag{4}$$

where $\mathbf{w} = [\mathbf{w}_1^\mathrm{T}, \mathbf{w}_2^\mathrm{T}, \ldots, \mathbf{w}_L^\mathrm{T}]^\mathrm{T}$ and $\mathbf{H}_{\mathrm{B},m} = [\mathbf{H}_{\mathrm{B},m1}, \mathbf{H}_{\mathrm{B},m2}, \ldots, \mathbf{H}_{\mathrm{B},mL}]$. Denoting the sound pressure at all the M_B microphones in the bright zone in a vector form, *i.e.*, $\mathbf{p}_\mathrm{B} = [\mathbf{p}_{\mathrm{B},1}^\mathrm{T}, \mathbf{p}_{\mathrm{B},2}^\mathrm{T}, \ldots, \mathbf{p}_{\mathrm{B},M_\mathrm{B}}^\mathrm{T}]^\mathrm{T}$, one obtains $\mathbf{p}_\mathrm{B} = \mathbf{H}_\mathrm{B}\mathbf{w}$, where $\mathbf{H}_\mathrm{B} = [\mathbf{H}_{\mathrm{B},1}^\mathrm{T}, \mathbf{H}_{\mathrm{B},2}^\mathrm{T}, \ldots, \mathbf{H}_{\mathrm{B},M_\mathrm{B}}^\mathrm{T}]^\mathrm{T}$ is an $M_\mathrm{B}(I + K - 1) \times IL$ matrix of the impulse responses from the loudspeaker array to the bright zone [22].

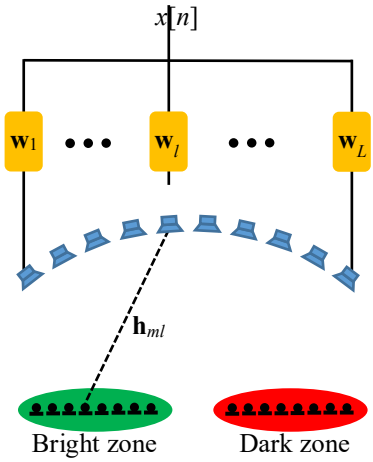

Fig. 1. Block diagram of the personal audio systems.

Similarly, the sound pressure in the dark zone can be obtained as $\mathbf{p}_\mathrm{D} = \mathbf{H}_\mathrm{D}\mathbf{w}$ with $\mathbf{H}_\mathrm{D}$ being the $M_\mathrm{D}(I + K - 1) \times IL$ impulse response matrix from the loudspeaker array to the dark zone.

BACC maximizes the ratio of the average acoustic potential energy in the bright zone to that in the dark zone, but suffers from an uneven frequency response and non-uniform spatial sound field distribution in the bright zone. To increase the flatness of the frequency response, BACC-RTE introduces an extra constraint on the response differences between frequencies, i.e., [24]

$$\mathbf{w}_\mathrm{RTE} = \arg\max_{\mathbf{w}} \frac{\mathbf{w}^\mathrm{T}\mathbf{R}_\mathrm{B}\mathbf{w}}{\mathbf{w}^\mathrm{T}\left[(1-\alpha)\mathbf{R}_\mathrm{D} + \alpha\mathbf{C}_\mathrm{RTE}\right]\mathbf{w} + \lambda\mathbf{w}^\mathrm{T}\mathbf{w}}, \tag{5}$$

where $\mathbf{R}_\mathrm{B} = \mathbf{H}_\mathrm{B}^\mathrm{T}\mathbf{H}_\mathrm{B}/M_\mathrm{B}$, $\mathbf{R}_\mathrm{D} = \mathbf{H}_\mathrm{D}^\mathrm{T}\mathbf{H}_\mathrm{D}/M_\mathrm{D}$, λ is the regularization parameter, α is the weighting factor, and $\mathbf{C}_\mathrm{RTE}$ is an $IL \times IL$ matrix to account for the mean frequency response difference (refer to [24] for details).

The solution to Eq. (5) is proportional to the eigenvector corresponding to the largest eigenvalue of the matrix $((1 - \alpha)\mathbf{R}_\mathrm{D} + \alpha\mathbf{C}_\mathrm{RTE} + \lambda\mathbf{I})^{-1}\mathbf{R}_\mathrm{B}$. It is clear that when $\alpha = 0$, BACC-RTE degenerates to the original BACC method.

2.1. BACC-RTE-SUC

The BACC-RTE method improves the frequency response but does not control the spatial sound field distribution. This paper proposes to apply a Spatial Uniformity Constraint (SUC) on the BACC-RTE algorithm to improve the sound field distribution in the bright zone.

The sound pressure at the m-th microphone in the bright zone is [23]

$$P_{\mathrm{B},m}(f) = \sum_{n=0}^{I+K-2} \mathbf{p}_{\mathrm{B},m}[n]\mathrm{e}^{-\mathrm{j}2\pi nfT_\mathrm{s}} = \mathbf{w}^\mathrm{T}\mathbf{H}_{\mathrm{B},m}{}^\mathrm{T}\underline{\boldsymbol{s}}(f), \tag{6}$$

where $\underline{\boldsymbol{s}}(f) = [1, \mathrm{e}^{-\mathrm{j}2\pi fT\mathrm{s}}, \ldots, \mathrm{e}^{-\mathrm{j}2\pi f(J-1)Ts}]^\mathrm{T}$ ($J = I + K - 1$). The SUC term is defined as

$$SUC = \frac{1}{C}\sum_{j=1}^{J}\sum_{m=1}^{M_B-2b+1}\left|P_{B,m+b}(f_j) - P_{B,m}(f_j) + \ldots + P_{B,m+2b-1}(f_j) - P_{B,m+b-1}(f_j)\right|^2 , \quad (7)$$

where $C = J(M_B - 2b + 1)$ and b is a positive integer.

Substituting (6) into (7), one obtains

$$\begin{aligned} SUC &= \frac{1}{C}\mathbf{w}^T\mathcal{R}\left\{\sum_{j=1}^{J}\mathbf{H}_B{}^T\underline{\boldsymbol{u}}(f_j)\mathbf{V}_b\mathbf{V}_b{}^T\underline{\boldsymbol{u}}(f_j)^H\mathbf{H}_B\right\}\mathbf{w} \\ &= \mathbf{w}^T\left\{\frac{1}{C}\mathbf{H}_B{}^T\mathcal{R}\left\{\sum_{j=1}^{J}\underline{\boldsymbol{u}}(f_j)\mathbf{V}_b\mathbf{V}_b{}^T\underline{\boldsymbol{u}}(f_j)^H\right\}\mathbf{H}_B\right\}\mathbf{w} \\ &= \mathbf{w}^T\left\{\frac{1}{C}\mathbf{H}_B{}^T\mathcal{R}\left\{\underline{\boldsymbol{U}}\mathbf{V}\underline{\boldsymbol{U}}^H\right\}\mathbf{H}_B\right\}\mathbf{w} \\ &= \mathbf{w}^T\mathbf{C}_{SUC}\mathbf{w} \end{aligned} , \quad (8)$$

where $\underline{\boldsymbol{u}}(f_j) = \mathrm{diag}\{\underline{\boldsymbol{s}}(f_j)\}$ is a block diagonal matrix with a dimension of $M_B J \times M_B$, $\underline{\boldsymbol{U}} = [\underline{\boldsymbol{u}}(f_1), \underline{\boldsymbol{u}}(f_2), \ldots, \underline{\boldsymbol{u}}(f_J)]$ with a dimension of $M_B J \times M_B J$,

$$\mathbf{V}_b = \frac{1}{b^2}\begin{bmatrix} -1_1 & 0 & 0 \\ \vdots & \ddots & 0 \\ -1_b & \ddots & -1_1 \\ 1_1 & \ddots & \vdots \\ \vdots & \ddots & -1_b \\ 1_b & \ddots & 1_1 \\ 0 & \ddots & \vdots \\ 0 & 0 & 1_b \end{bmatrix} \quad (9)$$

is a $M_B \times (M_B - 1)$ matrix, and $\mathbf{V} = \mathrm{diag}\{\mathbf{V}_b\mathbf{V}_b{}^T\}$ is a diagonal block matrix with a dimension of $M_B J \times M_B J$.

A uniform sound field distribution in the bright zone requires a small SUC term, thus the proposed BACC-RTE-SUC algorithm is described as:

$$\mathbf{w}_{SUC} = \arg\max_{\mathbf{w}} \frac{\mathbf{w}^T\mathbf{R}_B\mathbf{w}}{\mathbf{w}^T\left[\gamma\mathbf{R}_D + \alpha\mathbf{C}_{RTE} + \beta\mathbf{C}_{SUC}\right]\mathbf{w} + \lambda\mathbf{w}^T\mathbf{w}} , \quad (10)$$

where $\gamma = (1 - \alpha - \beta)$. The solution to Eq. (10) is proportional to the eigenvector corresponding to the largest eigenvalue of the matrix $(\gamma\mathbf{R}_D + \alpha\mathbf{C}_{RTE} + \beta\mathbf{C}_{SUC} + \lambda\mathbf{I})^{-1}\mathbf{R}_B$. It is clear that when $\beta = 0$, the proposed approach in Eq. (10) is the same as the BACC-RTE method in Eq. (5).

3. RESULTS

Simulations with measured RIRs were performed to demonstrate the efficacy of the proposed BACC-RTE-SUC algorithm in improving the spatial sound field uniformity in the bright zone.

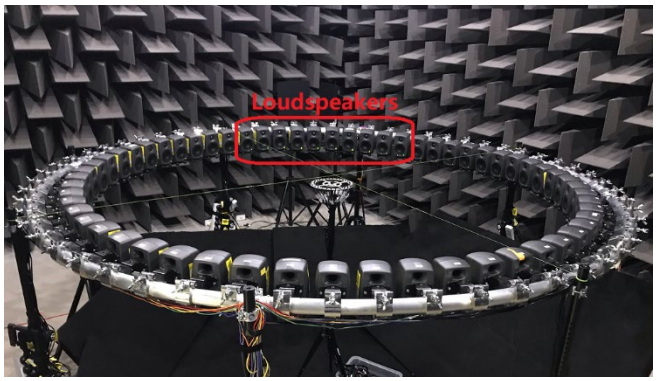

Fig. 2. Photograph of the experimental setup.

The RIRs were measured in a semi-anechoic chamber, as depicted in Fig. 2. Sixty Genelec 8010A loudspeakers were uniformly mounted on a circle truss with a diameter of 1.5 m at a height of 1.5 m above the ground. The distance between the loudspeaker acoustic centers was approximate 15.7 cm. The ground was covered with sound absorbing material to reduce ground reflections. Eight DPA 4060 Miniature Omnidirectional microphones were placed in the bright and dark zones to form a linear microphone array with an interval of 4 cm, as illustrated in Fig. 1. The bright and dark zones were separated 0.8 m from each other. A Log Sweep Sine signal sent to each loudspeaker through a YAMAHA Rio1608-D2 I/O interface. The RIRs were measured at a sampling rate of 48 kHz and down sampled to 8 kHz.

In the simulations, only ten loudspeakers (in the red rectangle in Fig. 2) were used to reduce the computational burden and each RIR was modelled with a 256-sample FIR filter ($K = 256$). The length of the control filters $\mathbf{w}_l$ was set as $I = 128$. The regularization parameter λ was chosen as 10^{-10} times the largest eigenvalue of the matrix $\mathbf{R}_D$. The weighting factors α and β were 0.6 and 0.3, respectively.

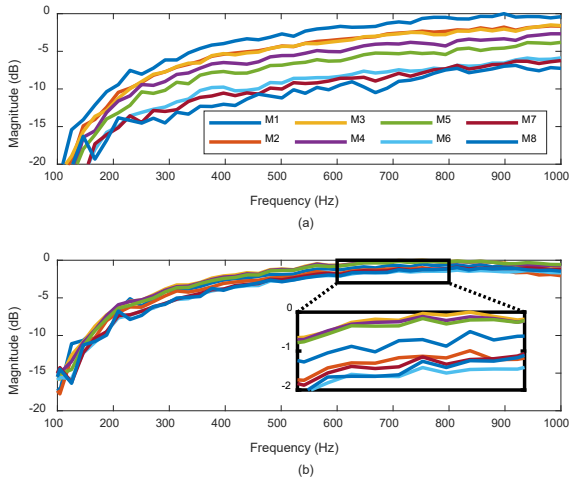

Fig. 3. Frequency responses at different microphones (M1-8) in the bright zone, (a) BACC-RTE and (b) BACC-RTE-SUC.

The frequency responses at the eight microphones in the bright zone are shown in Fig. 3(a) and (b) for the BACC-RTE and BACC-RTE-SUC algorithms, respectively. It can be observed that the sound pressure magnitude generated by the BACC-RTE algorithm varies with spatial locations in the bright zone while the BACC-RTE-SUC algorithm mitigates the spatial variation and produces a more uniform sound field distribution in the bright zone. Fig. 3(a) shows that sound pressure magnitudes between microphones deviate by up to 8 dB for the BACC-RTE method. By contrast, the deviations in magnitudes in the bright zone are less than 2 dB for BACC-RTE-SUC, as denoted by the subset in Fig. 3(b) for the detailed frequency responses between 600 Hz and 800 Hz.

The Just Noticeable Difference (JND) level of human ears depends on many factors, including but not limited to the frequency, bandwidth, duration, sound pressure level *etc*, but the order of magnitude of the JND is approximate 1 dB [25]. Therefore, for the sake of simplicity without loss of generality, the spatial sound pressure distribution in the bright zone is supposed to be perceptually uniform if the difference in the sound pressure magnitudes is less than 1 dB.

To more clearly illustrate improvement of the BACC-RTE-SUC approach, the magnitudes at different microphones were compared in Fig. 4 for 500 Hz and 1000 Hz signals. Fig. 4(a) shows that, at 500 Hz the sound pressure magnitudes produced by the BACC-RTE algorithm at Microphone M1 is 8.5 dB higher than that at Microphone M8. By contrast, the maximum difference in magnitudes generated by the BACC-RTE-SUC algorithm is 1.6 dB, thus providing a more uniform spatial sound field distribution in the bright zone. Similarly, the maximum difference in SPL between the eight microphones is reduced from 6.9 dB by BACC-RTE to 1.5 dB by BACC-RTE-SUC at 1000 Hz in Fig. 4(b). Although the maximum differences in sound pressure magnitudes in the bright zone achieved by the proposed BACC-RTE-SUC algorithm are slightly higher than the JND, it has been significantly improved compared to the BACC-RTE method.

To quantify the performance of the algorithms at different frequencies, magnitude difference (ΔL_p) in the bright zone and the Acoustic Contrast (AC) [23] between the bright and dark zones in Eqs. (11) and (12), respectively, were calculated and compared.

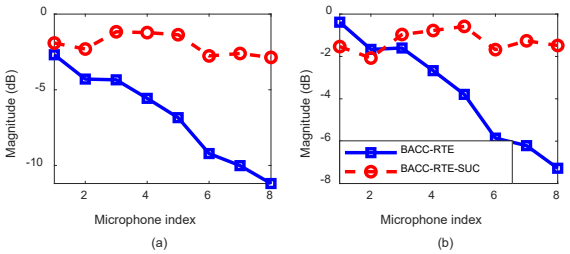

Fig. 4. Sound pressure magnitude at different microphones in the bright zone at (a) f = 500 Hz and (b) f = 1000 Hz.

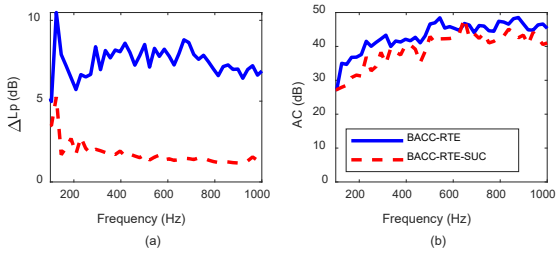

Fig. 5. Comparison of the (a) magnitude difference (ΔL_p) and (b) acoustic contrast (AC) in the bright zone.

$$\Delta L_p(f) = L_{p,\max}(f) - L_{p,\min}(f). \tag{11}$$

$$AC(f) = 10\log_{10}\frac{M_D\sum_{m=1}^{M_B}\left|P_{B,m}(f)\right|^2}{M_B\sum_{m=1}^{M_D}\left|P_{D,m}(f)\right|^2}. \tag{12}$$

In Eq. (11), $L_{p,\max}(f)$ and $L_{p,\min}(f)$ are the maximum and minimum sound pressure levels at frequency f at the different microphones in the bright zone, respectively. In Eq. (12), $P_{B,m}(f)$ is the sound pressure at the m-th microphone defined as in Eq. (6).

The ΔL_p and AC produced by the BACC-RTE and BACC-RTE-SUC are compared in Fig. 5(a) and (b), respectively. The average ΔL_p between 100 Hz and 1000 Hz is reduced from 7.5 dB of BACC-RTE to 1.7 dB of BACC-RTE-SUC. That comes at a cost of an average degradation of ~4 dB in the acoustic contrast, as shown in Fig. 5(b). Subjective evaluation of audio-on-audio interference shows that an level difference of 27 – 29 dB is needed to achieve a non-distracting personal audio system [26]. The average AC by BACC-RTE-SUC is 39.4 dB, thus the degradation in AC is supposed to be perceptually negligible.

4. CONCLUSIONS

This paper has presented a time-domain broadband acoustic contrast control method which improves the spatial sound field distribution. A spatial uniformity constraint is applied on the BACC-RTE algorithm to form a new BACC-RTE-SUC algorithm. Simulations based on measured room impulse responses were performed and the results demonstrate that the proposed algorithm improves the spatial sound field distribution in the bright zone at the cost of an insignificant decrease in the acoustic contrast. Future work includes subjective evaluation of the perceptual improvement of the proposed method.

5. REFERENCES

[1] J. Cheer, S. J. Elliott, Y. Kim, and J. W. Choi, "Practical implementation of personal audio in a mobile device," *AES J. Audio Eng. Soc.*, vol. 61, no. 5, pp. 290–300, 2013.

[2] Z. Tu, J. Lu, and X. Qiu, "Robustness of a compact endfire personal audio system against scattering effects (L)," *J.*

Acoust. Soc. Am., vol. 140, no. 4, pp. 2720–2724, 2016, doi: 10.1121/1.4964752.

[3] S. Jeon and J. Choi, "Personal audio system for neckband headset with low computational complexity," *J. Acoust. Soc. Am.*, vol. 148, no. 6, pp. 3913–3927, 2020, doi: 10.1121/10.0002941.

[4] 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, 2009, doi: 10.1121/1.3082114.

[5] J. Cheer, S. J. Elliott, M. F. S. Gálvez, and M. F. S. I. M. On, "Design and implementation of a car cabin personal audio system," *J. Audio Eng. Soc.*, vol. 61, no. 6, pp. 412–424, 2013.

[6] J.-H. Chang and W.-H. Cho, "Evaluation of independent sound zones in a car," in *Proceedings of the 23rd International Congress on Acoustics*, 2019, pp. 5196–5203.

[7] H. So and J.-W. Choi, "Subband optimization and filtering technique for practical personal audio systems," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2019, pp. 8494–8498.

[8] J. Brunskog *et al.*, "Full-scale outdoor concert adaptive sound field control," in *Proceedings of the International Congress on Acoustics*, 2019, pp. 1170–1177, doi: 10.18154/RWTH-CONV-239890.

[9] J.-W. Choi and Y.-H. 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, 2002, doi: 10.1121/1.1456926.

[10] M. Poletti, "An investigation of 2D multizone surround sound systems," in *Audio Engineering Society - 125th Audio Engineering Society Convention 2008*, 2008, pp. 167–175.

[11] Y. J. Wu and T. D. Abhayapala, "Spatial multizone soundfield reproduction: Theory and design," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 19, no. 6, pp. 1711–1720, 2011, doi: 10.1109/TASL.2010.2097249.

[12] J.-H. Chang and F. Jacobsen, "Sound field control with a circular double-layer array of loudspeakers," *J. Acoust. Soc. Am.*, vol. 131, no. 6, pp. 4518–4525, 2012, doi: 10.1121/1.4714349.

[13] T. Lee, J. K. Nielsen, and M. G. Christensen, "Signal-Adaptive and Perceptually Optimized Sound Zones with Variable Span Trade-Off Filters," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 28, pp. 2412–2426, 2020, doi: 10.1109/TASLP.2020.3013397.

[14] S. J. Elliott, J. Cheer, J. W. Choi, and Y. Kim, "Robustness and regularization of personal audio systems," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 20, no. 7, pp. 2123–2133, 2012, doi: 10.1109/TASL.2012.2197613.

[15] P. Coleman, P. J. B. Jackson, M. Olik, M. Møller, M. Olsen, and J. Abildgaard Pedersen, "Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array," *J. Acoust. Soc. Am.*, vol. 135, no. 4, pp. 1929–1940, 2014, doi: 10.1121/1.4866442.

[16] S. Zhao, X. Qiu, and I. S. Burnett, "Acoustic contrast control in an arc-shaped area using a linear loudspeaker array," *J. Acoust. Soc. Am.*, vol. 137, no. 2, pp. 1036–1039, 2015, doi: 10.1121/1.4906184.

[17] P. Coleman, P. J. B. Jackson, M. Olik, and J. Abildgaard Pedersen, "Personal audio with a planar bright zone," *J. Acoust. Soc. Am.*, vol. 136, no. 4, pp. 1725–1735, 2014, doi: 10.1121/1.4893909.

[18] N. Yanagidate, J. Cheer, S. Elliott, and T. Toi, "Car cabin personal audio: Acoustic contrast with limited sound differences," in *Proceedings of the AES International Conference*, 2014, pp. 1–8.

[19] M. F. Simon Galvez, S. J. Elliott, and J. Cheer, "Time Domain Optimization of Filters Used in a Loudspeaker Array for Personal Audio," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 23, no. 11, pp. 1869–1878, 2015, doi: 10.1109/TASLP.2015.2456428.

[20] Y. Cai, M. Wu, and J. Yang, "Design of a time-domain acoustic contrast control for broadband input signals in personal audio systems," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2013, pp. 341–345, doi: 10.1109/ICASSP.2013.6637665.

[21] S. J. Elliott and J. Cheer, "Regularisation and Robustness of Personal Audio Systems," 2011.

[22] M. Hu and J. Lu, "Theoretical explanation of uneven frequency response of time-domain acoustic contrast control method," *J. Acoust. Soc. Am.*, vol. 149, no. 6, pp. 4292–4297, 2021, doi: 10.1121/10.0005270.

[23] Y. Cai, M. Wu, L. Liu, and J. Yang, "Time-domain acoustic contrast control design with response differential constraint in personal audio systems," *J. Acoust. Soc. Am.*, vol. 135, no. 6, pp. EL252–EL257, 2014, doi: 10.1121/1.4874236.

[24] D. H. M. Schellekens, M. B. Moller, and M. Olsen, "Time domain acoustic contrast control implementation of sound zones for low-frequency input signals," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2016, pp. 365–369, doi: 10.1109/ICASSP.2016.7471698.

[25] H. Fastl and E. Zwicker, *Psychoacoustics: facts and models*. Springer, 2007.

[26] J. Rämö, S. Bech, and S. H. Jensen, "Validating a real-time perceptual model predicting distraction caused by audio-on-audio interference," vol. 153, 2018, doi: 10.1121/1.5045321.