Design and Implementation of an Active Noise Control Headphone with Directional Hear-Through Capability

Vinal Patel, *Member, IEEE*, Jordan Cheer, and Simone Fontana

Abstract—This paper presents the design and implementation of an active noise control (ANC) headphone system with a directional hear-through capability and compares the performance of this system to that of a standard hear-through headphone system. The directional hear-through ANC headphones are a novel integration of microphone array beamforming and ANC technologies into a pair of headphones, which provide the consumer with additional functionality and new, digitally augmented ways to interact with their acoustic environment. As the microphone array is necessarily compact, superdirective beamforming is utilised to increase its low and mid frequency directional performance. In this unique integration of two current consumer technologies, first, the ANC subsystem attempts to maximise the attenuation and then the beamformer output is added to the control signal and reproduced by the headphones' loudspeakers, with the appropriate compensation to avoid self-cancellation. The experimental study demonstrates that the proposed spatially selective ANC headphones provide a hear-through capability in the look direction, whilst reducing ambient noise and enabling the wearer to experience reduced noise communication in a noisy environment. The proposed system thus offers the consumer the potential for an electronically enhanced acoustic experience, allowing a selective reduction in environmental noise whilst desired exterior noise remains audible.

Index Terms—Active noise control, Microphone array, Superdirective beamformer, Adaptive control, Hardware, Augmented reality, Feedforward systems.

I. INTRODUCTION

ACTIVE noise control (ANC) headphones have seen significant success in the consumer market over almost two decades, since they improve the listening experience in conditions that are usually hostile to audio reproduction, such as trains, aeroplanes, and busy urban areas. The earliest ANC headphones were demonstrated in the 1950s and utilised analogue feedback controllers, however, they could not readily adapt to changes in the unwanted sound or small changes in the plant responses due to the positioning of the headphones [1]. In the recent past, a large volume of research has been carried out to develop digital ANC headphones that can potentially provide higher levels of performance than analogue ANC due to the greater level of flexibility in the controller design and this has resulted in a significant level of adoption of

J. Cheer is with the Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, United Kingdom (e-mail: j.cheer@soton.ac.uk).

V. Patel was with the Institute of Sound and Vibration Research, but now with IIITM Gwalior, 474015, India (e-mail: vp@iiitm.ac.in).

S. Fontana is with HUAWEI European Research Center, Riesstrasse 25 C3.0G, 80992, Munich, Germany (e-mail: Simone.Fontana@huawei.com).

ANC technology into consumer headphones, rather than just in the original aerospace and military applications [2]. The basic single channel feed-forward ANC system consists of a reference microphone that senses the noise to be cancelled, a loudspeaker that generates the anti-noise and an error microphone that measures the level of noise cancellation achieved. The anti-noise, produced by the loudspeaker, is generated using an adaptive controller that is conventionally an adaptive filter trained using a suitable learning algorithm [3].

Although the idea of implementing noise cancellation using digital processing integrated into the ANC headphones seems somewhat straightforward, the realisation is a complicated process where many system components must be considered [3]. One such issue in digital ANC headphones is the inter-linked delays due to analogue-to-digital and digital-to-analogue conversion, antialiasing and reconstruction filtering, and controller processing. These delays limit the bandwidth over which control can be achieved and, therefore, a number of systems have been proposed that combine an analogue feedback controller with a digital feedforward controller [4], [5]. It has been reported that as the processing speed continues to increase, higher sampling rates could be readily used in the digital controller implementations to reduce the inherent delay in digital implementations of ANC headphones [3]. Regardless of the type of headset ANC is implemented in (in-ear, supra-aural or circumarual and open or closed back headsets), there are three basic control strategies to reduce ambient noise: feedforward, feedback and hybrid control [4]–[9].

In general, ANC headphones, irrelevant of the specific control strategy, have used independent controllers for the left and right ears. Recently, however, a multi-reference control strategy for ANC headphones has been proposed, which uses the reference signals from the reference microphones on each ear to improve the performance of the ANC headphones [10]. The main advantage of the multi-reference ANC system over existing ANC headphone strategies is the availability of additional time-advanced information, which results in an increased control bandwidth [10].

In more recent years, in addition to reducing the sound due to exterior noise sources, consumer audio systems have begun to focus on virtual and augmented reality [11]–[15]. One area of current interest focuses on designing headphones that allow hear-through functionality in addition to reducing the sound due to exterior noise sources [11], [12], however, this previous work has not actually demonstrated the combination of these technologies. These devices have generally been designed to

allow the consumer to hear the exterior acoustic environment without coloration, and there is an increasing desire to augment this listening experience through the addition of virtual sound sources [11]–[15]. The limitation of current augmented reality devices is that they do not allow the consumer to reduce unwanted environmental noise, by selecting the spatial region over which the environmental noise is fed-through, whilst reducing the noise in other directions. The need for greater and more flexible control of acoustic noise is extremely important for consumers, who are subjected to increasing levels of environmental noise in their daily lives, but still need to communicate and interact with the exterior acoustic environment.
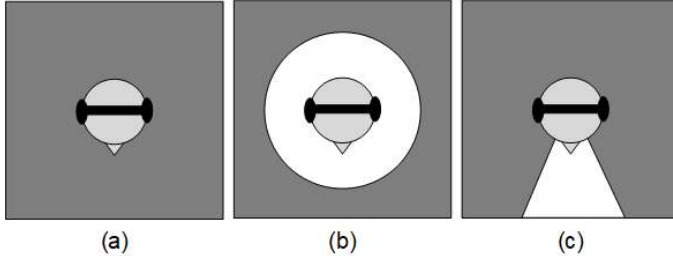

Fig. 1. Conceptual difference between (a) ANC headphones (b) ANC headphones with direct hear-through (c) Directional hear-through ANC headphones. Dark and bright regions illustrate the controlled and uncontrolled spatial directivities respectively.

In this paper, an ANC headphone system with a directionally selective hear-through capability is presented for the first time. In comparison to standard ANC headphones, which cancel the ambient noise in all directions, the proposed system cancels the ambient noise whilst aiming to leave the sound incident from a selected direction unaltered. This will provide the perception of hear-through from the target direction, as if the wearer was not wearing headphones for the sound coming from that particular direction, as depicted in Fig. 1. This is distinct from direct hear-through headphones, in which the signal output from the exterior reference microphone is reproduced by the loudspeakers inside the ear cup after equalization to compensate for the electro-acoustic response [11], [12]. The presented technology thus provides a functionality improvement over the current state-of-the-art systems. In addition, the proposed ANC headphones could provide the consumer with the ability to select the spatial regions over which both hear-through and active control are focused. The spatially selective hear-through functionality will be achieved in the proposed system by integrating a directional microphone array into an ANC system. These subsystems have both previously been implemented independently in consumer electronics devices and, therefore, this paper will focus on the theoretical framework of combining the directional microphone array with the ANC system to facilitate spatial selectivity. In addition, the advantage of the spatially selective ANC headphones over direct hear-through headphones with and without ANC is also investigated.

The rest of this paper is organised as follows. The proposed directional hear-through ANC system, with a brief overview of both the multi-reference ANC and the superdiretive beam-forming strategies, is discussed in Section II. A detailed study of the hardware design for the directional hear-through ANC system is presented in Section III. Experimental validation of the spatially selective ANC headphones under different noise fields is demonstrated in Section IV and concluding remarks are made in Section V.

II. PROPOSED METHODOLOGY

The directional hear-through ANC headphones comprise of two primary subsystems: an ANC headset and a microphone array, which will ultimately be interlinked. A Multi-reference ANC system is used due to its reduced sensitivity to the direction of arrival of the incident noise source compared to the conventional single-reference feedforward control system [10]. This ANC system is described in the following section and then the superdirective beamforming strategy is presented.

A. Multi-Reference Active Noise Control System

Figure 2 shows the block diagram of the multi-reference ANC system. In the multi-reference feedforward controller, the signal fed to the control loudspeaker can be expressed as the summation of the control signals generated by filtering both the left and right reference signals, x_L and x_R respectively. Using the left ear controller as an example, the control signal driving the left loudspeaker is given by

$$
\begin{aligned}
u_L(n) &= u_{LL}(n) + u_{LR}(n) \\
&= \mathbf{w}_{LL}^T(n)\mathbf{x}_L(n) + \mathbf{w}_{LR}^T(n)\mathbf{x}_R(n). \quad (1)
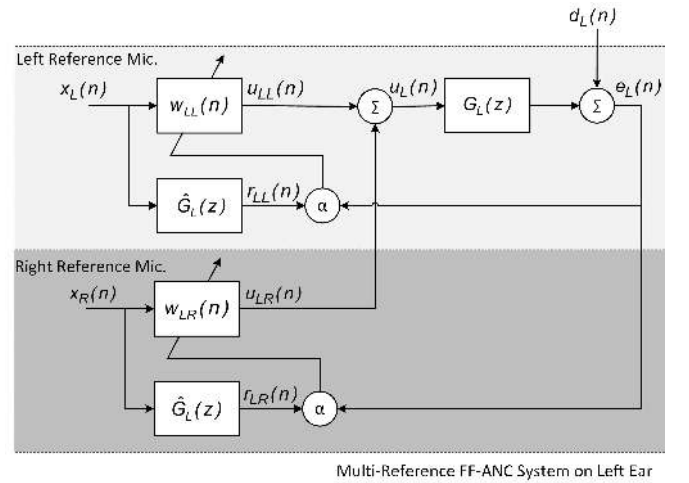\end{aligned}
$$

Fig. 2. Block diagram of the multi-reference feedforward active noise control system for the left ear of the headphones; the equivalent block diagram for the right ear controller can be obtained by interchanging the L and R subscripts.

In the above expression, $\mathbf{x_L}(n) = [x_L(n), x_L(n-1), ..., x_L(n-I+1)]^T$ and $\mathbf{x_R}(n) = [x_R(n), x_R(n-1), ..., x_R(n-I+1)]^T$ are the left and right reference signal vectors respectively, where I is the length of the vectors, which is determined by the control filter length. Moreover, $\mathbf{w}_{LL} \in \mathbb{R}^{I_{LL} \times 1}$ and $\mathbf{w}_{LR} \in \mathbb{R}^{I_{LR} \times 1}$ are the weights of the control filters, where $\mathbf{w}_{LL}$ corresponds to the control filter

whose input is from the left reference microphone and is used to control the left ear error sensor, whereas $\mathbf{w}_{LR}$ corresponds to the control filter whose input is from the right reference microphone and is used to control the left ear error sensor. The weights of the two control filters are updated using the filtered-reference leaky LMS algorithm [3], [7], given as

$$\mathbf{w}_{LL}(n+1) = (1-\alpha_{LL}\beta_{LL})\mathbf{w}_{LL}(n)-\alpha_{LL}e_L(n)\mathbf{r}_{LL}(n) \quad (2)$$

$$\mathbf{w}_{LR}(n+1) = (1-\alpha_{LR}\beta_{LR})\mathbf{w}_{LR}(n) - \alpha_{LR}e_L(n)\mathbf{r}_{LR}(n) \quad (3)$$

where α_{LL} and α_{LR} are the normalized convergence coefficient given as $\frac{\mu}{\mathbf{r}_{LL}^T(n)\mathbf{r}_{LL}(n)}$ and $\frac{\mu}{\mathbf{r}_{LR}^T(n)\mathbf{r}_{LR}(n)}$ respectively. β_{LL} and β_{LR} are the leakage parameters, $\mathbf{r}_{LL}(n) \in \mathbb{R}^{I_{LL}\times 1}$ and $\mathbf{r}_{LR}(n) \in \mathbb{R}^{I_{LR}\times 1}$ are the reference signals, $\mathbf{x}_L(n)$ and $\mathbf{x}_R(n)$ respectively, filtered by a model of the plant response between the left loudspeaker and left error microphone, $\hat{G}_L(z)$. The error signal at the left ear can be expressed as

$$e_L(n) = d_L(n) + \mathbf{g}_L^T\mathbf{u}_L(n), \quad (4)$$

where $d_L(n)$ is disturbance at the left error microphone, $\mathbf{u}_L(n)$ is the control signal vector that drives the left side control loudspeaker, and, $\mathbf{g}_L^T$ is the impulse response corresponding to the physical plant $G_L(z)$. The main advantage of the multi-reference ANC system over the single-reference ANC headphone strategy is the availability of additional time-advanced information when controlling primary disturbances incident from the left (right) at the right (left) ear [10]. This means that the ANC headphones implemented using this strategy provides a more uniform attenuation in the noise with respect to its angle of incidence.

B. Superdirective Beamformer System

The design and implementation of superdirective microphone arrays has seen significant research interest, due to the broad range of applications [16]–[18]. Here, the superdirective beamforming problem is formulated in terms of a bright region, where the consumer wishes to hear the exterior sound environment, and a dark region, where the consumer does not want to hear the exterior environment, as depicted in Fig. 1c. This follows the description presented in [19] for the reciprocal loudspeaker array problem, and this optimisation problem is reformulated for the microphone array system. To facilitate the design using measured responses, it has been assumed that we are interested in detecting the sound from a particular (desired) direction due to a discrete number of sources at a given distance around the microphone array. The response of the microphone array in the desired direction is given as

$$\mathbf{y}_D = \mathbf{X}_D\mathbf{w}, \quad (5)$$

where $\mathbf{w} = [w_1, w_2, \cdots, w_M]^T$ is a vector of complex weights by which the microphone signals are scaled and $\mathbf{X}_D$ is an $(L \times M)$ matrix in which L denotes the number of sources in the look direction and M denotes the number of microphones in the array, given as

$$\mathbf{X}_D = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{L1} & x_{L2} & \cdots & x_{LM} \end{bmatrix} \quad (6)$$

where, $x_{l,m}$ is the response between the lth source point and the mth microphone in the array. Similarly, the responses between the microphone array and the source points in the undesired direction are given as

$$\mathbf{y}_U = \mathbf{X}_U\mathbf{w}, \quad (7)$$

where $\mathbf{X}_U$ is an $(N \times M)$ matrix in which N denotes the number of source points in the undesired direction, and can be written as

$$\mathbf{X}_U = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NM} \end{bmatrix}. \quad (8)$$

The superdirective microphone array weights are obtained by maximizing the ratio of the signal power $(\mathbf{y}_D^H\mathbf{y}_D)$ coming from the look-direction (desired) to the signal power $(\mathbf{y}_U^H\mathbf{y}_U)$ coming from the other directions (undesired) at each frequency and can be written as [19], [20]

$$R = \frac{\mathbf{y}_D^H\mathbf{y}_D}{\mathbf{y}_U^H\mathbf{y}_U} = \frac{\mathbf{w}^H\mathbf{X}_D^H\mathbf{X}_D\mathbf{w}}{\mathbf{w}^H\mathbf{X}_U^H\mathbf{X}_U\mathbf{w}}, \quad (9)$$

where, ^{H} denotes the Hermitian complex conjugate transpose. The maximum value of the cost function R at a single frequency can be obtained using constrained optimization [20] as

$$\text{minimize:} \quad \mathbf{y}_U^H\mathbf{y}_U \quad (10)$$
$$\text{subject to:} \quad \mathbf{y}_D^H\mathbf{y}_D = \text{constant } (c). \quad (11)$$

Using the method of Lagrange multipliers [20], the function to be minimized can be written in a generalized form as

$$\begin{aligned} J &= \mathbf{y}_U^H\mathbf{y}_U + \lambda(\mathbf{y}_D^H\mathbf{y}_D - c) \\ &= \mathbf{w}^H\mathbf{X}_D^H\mathbf{X}_D\mathbf{w} + \lambda(\mathbf{w}^H\mathbf{X}_U^H\mathbf{X}_U\mathbf{w} - c), \end{aligned} \quad (12)$$

where λ is the Lagrange multiplier. Taking the derivatives of (12) with respect to $\mathbf{w}$ and equating to zero, the following expression can be obtained

$$\mathbf{X}_D^H\mathbf{X}_D\mathbf{w} + \lambda\mathbf{X}_U^H\mathbf{X}_U\mathbf{w} = 0. \quad (13)$$

On rearranging (13), we have

$$\mathbf{w} = -\lambda[\mathbf{X}_D^H\mathbf{X}_D]^{-1}[\mathbf{X}_U^H\mathbf{X}_U\mathbf{w}]. \quad (14)$$

Thus, the optimal weight vector $\mathbf{w}$ is proportional to the eigenvector corresponding to the smallest eigenvalue value of the matrix $[\mathbf{X}_D^H\mathbf{X}_D]^{-1}[\mathbf{X}_U^H\mathbf{X}_U]$. To increase the robustness of the array, the optimal filter weights can alternatively be

calculated as $[\mathbf{X}_D^H \mathbf{X}_D + \gamma \mathbf{I}]^{-1}[\mathbf{X}_U^H \mathbf{X}_U]$, where γ is a regularization parameter, whose value can be increased to improve the robustness of the array at the expense of array performance [16]–[20].

C. Directional Hear-Through ANC System

Figure 3 shows the block diagram of the proposed directional hear-through ANC system. In this novel integration approach, conceptually, the ANC subsystem first attempts to achieve as much attenuation as possible and then the beamformer output is added to the control signal and reproduced by the headphone loudspeakers. The beamformer output is thus able to feed-through, or potentially enhance, the signal in the look-direction (desired direction) for the consumer. In Fig. 3, the beamformer subsystem processes the signals from the M microphones and generates the beamformer output, $s(n)$. As shown in Fig. 3, the beamformer output, $s(n)$, will be detected by the error microphones inside the ear cups and, therefore, an estimate of the beamformer signal measured at the error microphones must also be subtracted from the error signals to avoid the ANC subsystem attempting to cancel this signal. This means that the error signal fed to the ANC system can be expressed as

$$\hat{e}_L(n) = e_L(n) - \hat{s}_L(n), \tag{15}$$

where $\hat{e}_L(n)$ is the pseudo-error which is used to update the weights of the controller; this is similar to the process required when combining ANC with audio reproduction to avoid the ANC system cancelling the audio programme [21]. In (15), $\hat{s}_L(n)$ is the estimate of the beamformer signal at the error microphones, which is estimated by filtering the beamformer signal, $s(n)$, via a model of the plant response, between the loudspeaker and error microphone for the left ear, $\hat{G}_L(z)$, as shown in Fig. 3.
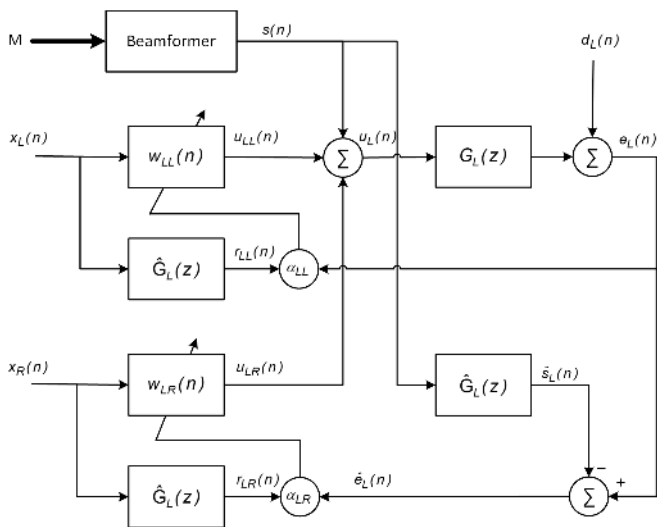

Fig. 3. Block diagram of the directional hear-through ANC system for the left ear; the equivalent block diagram for the right ear can be obtained by interchanging the L and R subscripts.

The directivity of the array over frequency inherently determines the level of spatial selectivity that the system can achieve; however, with the proposed integration approach, the delay introduced by the beamformer does not limit the performance of the ANC subsystem. For the system shown in Fig. 3, the allowable delay is instead limited by the acceptability of delay to the consumer. Many studies have considered this effect, particularly for hearing aids, but a recent study has investigated the tolerable levels of delay in speech signals for both normal and hearing-impaired listeners [22]. Listeners with normal hearing are less tolerant to delays in the system and the results presented in [22] suggest that delays should be kept below 10 ms.

III. MICROPHONE ARRAY PERFORMANCE

An overview of the proposed system is shown in Fig. 4(a), which is composed of a microphone array integrated with the ANC headphones. A detailed study on the design of the beamformer using a plus-shaped microphone array including the delays introduced by the microphone array, its beam patterns and directivity is presented in this section.

The microphone arrangement for four microphones in a plus-shaped configuration is shown in Fig. 4(b). The separation between the microphones on the plus-shaped array is 30 mm, as shown in Fig. 4(b). For such a small separation, a superdirective beamformer generally achieves greater directivity than the delay and sum beamformer, especially at low frequencies [23]. Therefore, a superdirective beamformer has been considered. The schematic of the measurement setup used to compute the filter weights for the superdirective beamformer is shown in Fig. 5. From this diagram it can be seen that the loudspeakers were positioned at angular increments of 30 degrees at a radial distance of 1.3 m around the dummy head. In order to include the effect of head reflections, the measurements have been taken with the dummy head wearing the plus-shaped microphone array mounted on the ANC headphones, as shown in Fig. 4(c). The responses between each of these loudspeakers and the microphones on the prototype headphones have been measured. Using (14), the complex weights at each frequency have been computed and these weights have then been transformed to the time domain to obtain the impulse responses or time domain filter weights for the superdirective beamformer. The parameter γ, described at the end of Section II-B controls the trade-off between the beamformer robustness or speech quality and the directivity, as suggested in [24] for the reciprocal loudspeaker array system. In the following, $\gamma = 1 \times 10^{-5}$ has been chosen when computing the weights of the plus-shaped beamformer.

The performance of the plus-shaped microphone array with the optimal weights calculated according to the formulation presented in Section II is summarised by the results presented in Fig. 6. Fig. 6(a) shows the impulse responses of the filters used by the four microphones forming the plus-shaped superdirective beamformer. It is evident from Fig. 6(a) that these beamforming filters introduce a delay of 5 ms, which is less than the tolerable delay for a normal listener according to [22]. The sensitivity of the beamformer to a plane wave impinging from various directions is given by the directivity pattern, which is shown in Fig. 6(b) for the plus-shaped
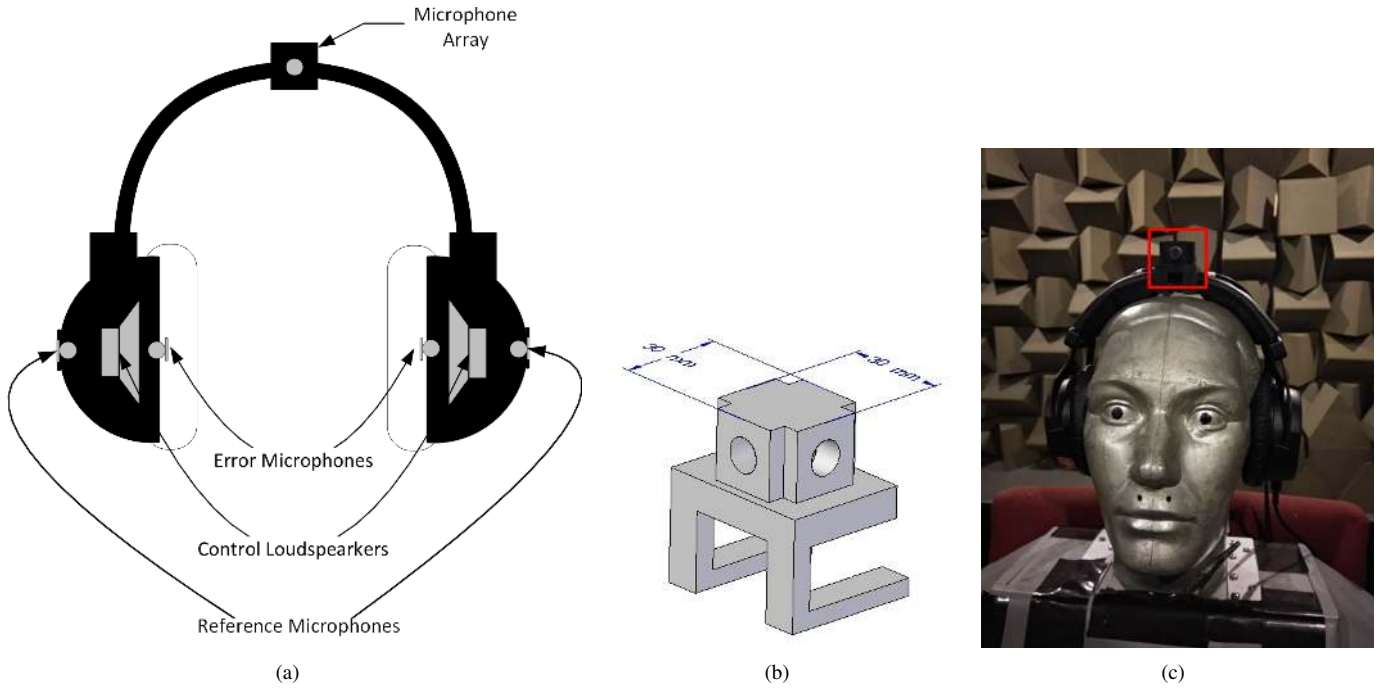
(a) (b) (c)

Fig. 4. (a) The practical realisation of the directionally selective ANC headphones system, which incorporates a directional microphone array in addition to the error and reference microphones. (b) Model of the plus-shaped microphone array mounting arrangement having a separation of 30 mm between microphones. (c) A four microphone plus shaped array integrated on the ANC headphones.
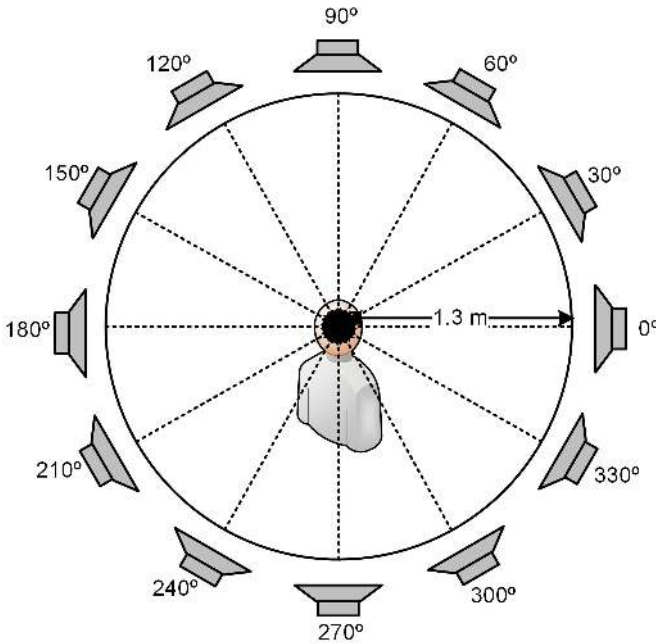

Fig. 5. Schematic of the measurement setup used to compute the weights for the plus-shaped superdirective beamformer.

beamformer at four frequencies (100, 500, 1000 and 5000 Hz), spanning the typical speech bandwidth. From these directivity responses, it can be seen that the plus-shaped array achieves significant levels of directivity over the speech bandwidth, with a front-to-back ratio of at least 10 dB. Apart from observing the diectivity pattern of the beamformer, it is also beneficial to quantify its directional response over frequency. This can

be achieved using the directivity index (DI), which is defined as the ratio between the on-axis pressure, i.e. the pressure in the look-direction and the average pressure measured by the beamformer in all other directions. The mathematical expression for the DI of a microphone array obtained for a set of discrete measurement points is given by [25]

$$DI(f) = \tag{16}$$

$$10\log_{10} \left[\frac{4\pi |p_{ax}|^2}{2 \sum\limits_{m=1}^{2\pi/\Delta_\theta} \sum\limits_{n=1}^{\pi/\Delta_\phi} |p(\phi_n, \theta_m, r)|^2 |cos\theta_n| \Delta_\theta \Delta_\phi} \right]$$

where p_{ax} is the on-axis sound pressure, Δ_θ and Δ_ϕ are the number of horizontal and vertical source points receptively. If the measurement is performed only for the horizontal plane, then (16) can be simplified as

$$DI(f) = 10\log_{10} \left[\frac{4\pi |p_{ax}|^2}{2 \sum\limits_{m=1}^{2\pi/\Delta_\theta} |p(\theta_m)|^2 \Delta_\theta} \right]. \tag{17}$$

Using (17) the DI of the plus-shaped superdirective beamformer for a look-direction to the front ($0°$) of the consumer is shown in Fig. 6(c). From this plot it can be seen that the plus-shaped beamformer achieves an almost flat DI over the entire frequency range considered, with an average DI of 6 dB between 100-5000 Hz. To allow steerability of the array, the beamformer weights for other look directions can be determined using the same process outlined above.
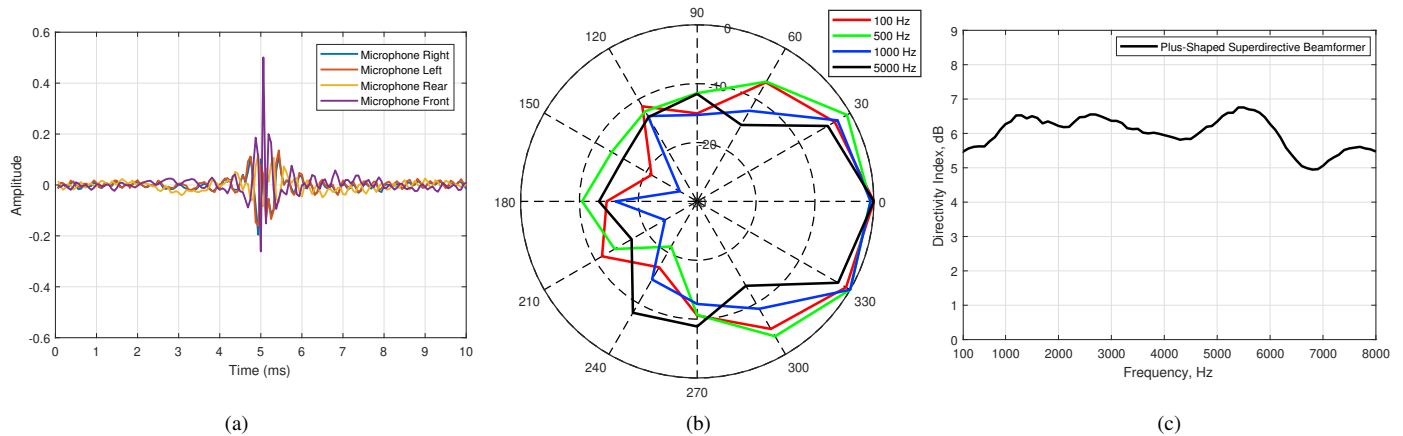
Fig. 6. (a) The optimal beamformer impulse responses corresponding to the four microphones (b) The directivity pattern at discrete frequencies over the typical speech bandwidth (c) The directivity index of the plus-shaped array including the effects of head scattering.

IV. EXPERIMENTAL VALIDATION OF THE INTEGRATED SYSTEM

In this section, experimental evaluation of the directional hear-through ANC headphones (ANC+Beamformer system) under different acoustic field scenarios will be presented. This study aims to demonstrate the benefits of the ANC+Beamformer system over a direct hear-through (HT) system integrated with the ANC system; that is, where the signals from the microphones on the exterior of the ear cups are fed, via equalization filters, to the loudspeakers directly. In addition, the following analysis highlights the performance difference between the beamformer system and the direct hear-through system with the ANC system on and off. In this investigation, the beamformer signal, as well as the signal from the direct hear-through system are passed through a highpass filter with a cut-on frequency of $Fc = 300Hz$ and then through an equalization filter, which is designed to compensate for the electro-acoustic response of each ear cup, as described for example in [11].

Fig. 7. Experimental setup used in the Institute of Sound and Vibration Research audio lab.

The experiments were conducted at a sampling rate of 16 kHz and the control algorithms were implemented on a digital signal processing board. The control filters were implemented using FIR filters with tap-lengths of $I_{LL} = I_{RR} = 160$ and

$I_{LR} = I_{RL} = 320$, and the plant response models were implemented using 128-tap FIR filters. The equalizers designed for both the beamformer and the direct hear-through system were 700-tap FIR filters, while the beamformer weights were implemented using an 80-tap FIR filter for each microphone. The learning rates $\alpha_{LL} = \alpha_{RR} = 2.5 \times 10^{-2}$, $\alpha_{LL} = \alpha_{RR} = 1 \times 10^{-3}$ and leakage parameters $\beta_{LL} = \beta_{RR} = 2 \times 10^{-5}$, $\beta_{LR} = \beta_{LR} = 2 \times 10^{-6}$ have all been chosen to maximize the convergence speed for all configurations.

Figure 7 shows the experimental setup used in the audio laboratory at the Institute of Sound and Vibration Research; the lab has a volume of around 36 m^3 and a reverberation time of 0.11 s. The audio lab comprises of loudspeakers positioned around the dummy head at a separation angle of $30°$. The loudspeaker at the front of the dummy head is used to generate a speech signal, while pink noise is generated by the loudspeakers between 30 and 330 degrees. To evaluate the performance of the directional hear-through ANC headphones, the speech signal, played from the loudspeaker at $0°$, and the noise, played via the surrounding loudspeakers from $30°$-$330°$, are considered separately.

Figure 8 shows the comparison between the ANC system, direct hear-through (HT) system, ANC+Direct hear-through (ANC+HT) system, beamformer (BF) system and ANC+Beamformer (ANC+BF) system for the case when the speech is being played from the loudspeaker at $0°$. The objective in this case is for the signal at the error microphone to match that at the reference microphone, so that the effect of the cup is effectively removed; i.e. the attenuation between the reference and the error signals should be 0 dB. It can be seen from the spectra presented in Fig. 8 that the error signal for both the ANC+HT system (cyan line) and ANC+BF system (pink line) match with the reference signal (blue line) over the speech bandwidth. The average attenuation in the speech bandwidth (500 Hz to 4kHz octave bands) is 0.95 dB for the left ear, and 0.89 dB for the right ear for the ANC+Beamformer system, and the ANC+Direct hear-through system gives a similar attenuation of 0.75 dB and 0.98 dB for left and right ears respectively. The two systems are thus considered effective at providing hear-through of the speech
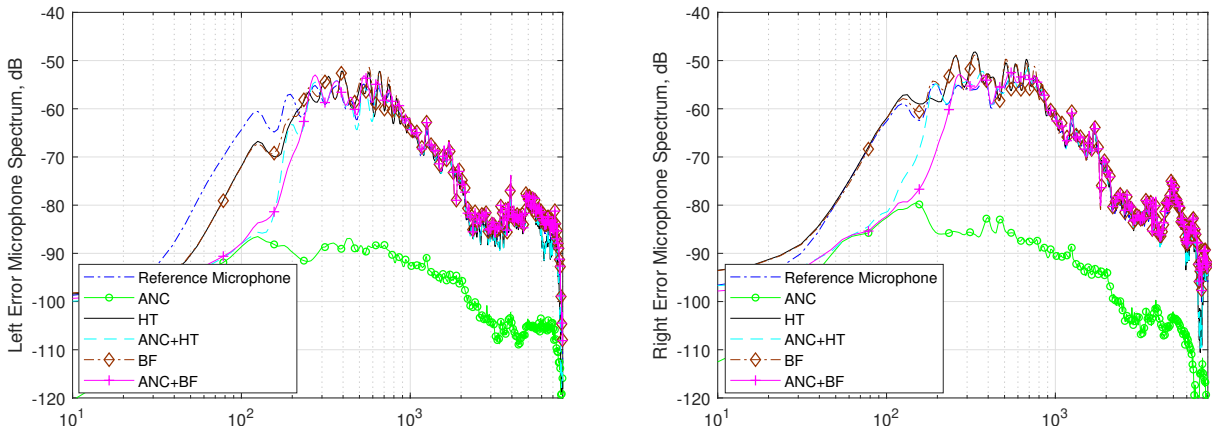
Fig. 8. Performance comparison between ANC system, direct hear-through (HT) system, ANC+Direct hear-through system, beamformer (BF) system and ANC+Beamformer system when the speech signal is played from $0°$.
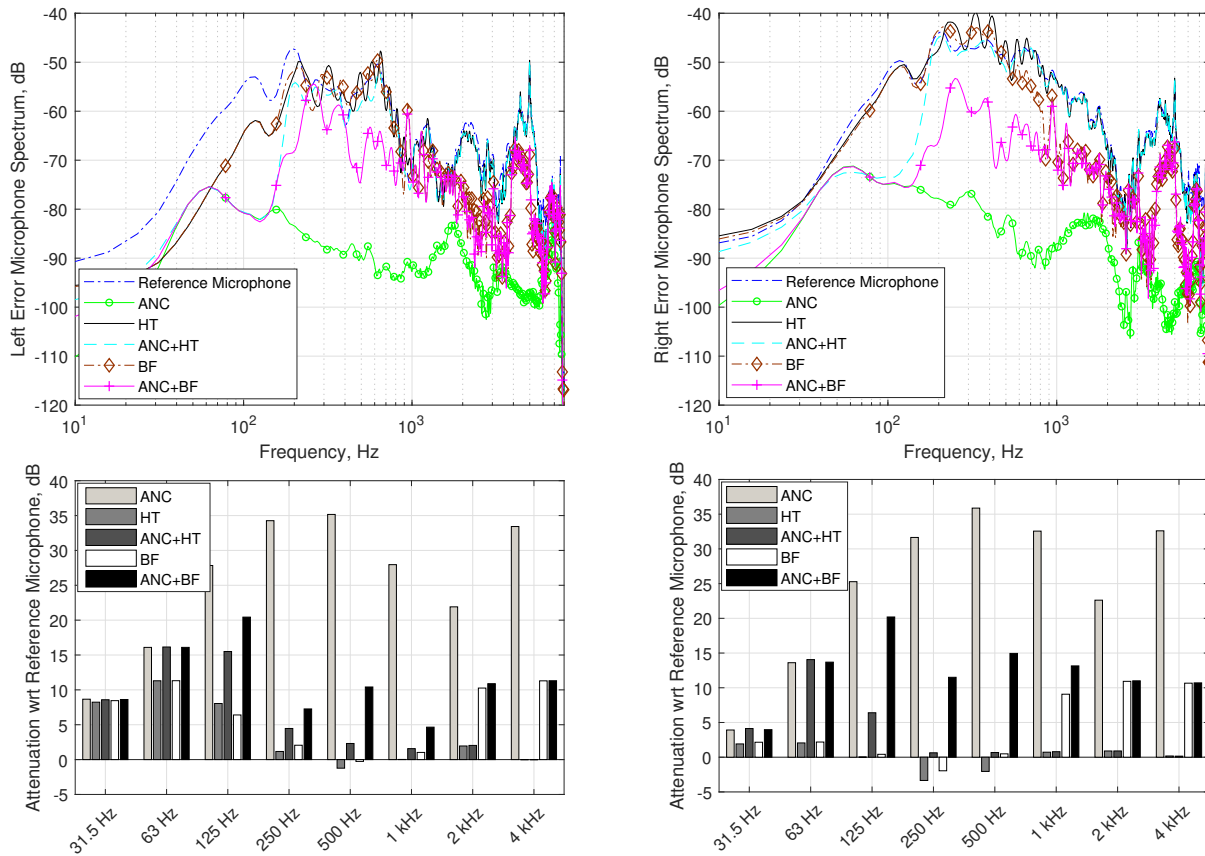

Fig. 9. Performance comparison between ANC system, direct hear-through system, ANC+Direct hear-through system, beamformer system and ANC+Beamformer system (directional hear-through ANC system) on the left and right ear when the noise signal is played from 30-330 degrees.

signal located at $0°$. It is also interesting to note from Fig. 8 that the hear-through system (black line) and beamformer system (yellow line) operating without ANC match the reference microphone signal over a wider bandwidth, because the ANC system is not active and thus the attenuation below 200 Hz is removed.

Figure 9 shows the performance comparison between the ANC system, HT system, ANC+HT system, BF system and ANC+BF system for the scenario when pink noise has been generated by each of the loudspeakers between 30 and 330 degrees. Since there is no signal being generated in the look direction, the objective in this case is for the reduction between the reference and error signals to be maximized. The details of this reduction in each case can be assessed from the spectral plots in Fig. 9, but a potentially clearer presentation is provided by the octave band plots, which show the attenuation at the

TABLE I
ATTENUATION ACHEIVED BY THE ANC SYSTEM, DIRECT HEAR-THROUGH (HT) SYSTEM, ANC+HT SYSTEM, BEAMFORMER (BF) SYSTEM AND
ANC+BF SYSTEM WITH RESPECT TO THE REFERENCE MICROPHONE, FOR THE CONDITION WHEN PINK NOISE IS PLAYED FROM THE LOUDSPEAKERS
POSITIONED BETWEEN 30-330 DEGREES AND THE BEAM IS STEERED AT $0°$ (SPEECH BANDWIDTH OCTAVE BANDS ARE SHOWN IN GREY).

		Left Ear (dB)					Right Ear (dB)				
		ANC	HT	ANC+HT	BF	ANC+BF	ANC	HT	ANC+HT	BF	ANC+BF
	31.25 Hz	9	14	9	14	9	4	2	4	2	4
	63 Hz	16	11	16	11	16	14	2	14	2	14
	125 Hz	28	8	16	6	20	25	0	6	0	20
	250 Hz	34	1	4	2	7	32	-3	1	-2	12
Octave bands	500 Hz	35	-1	2	0	10	36	-2	1	0	15
	1 kHz	28	0	2	1	5	33	1	1	9	13
	2 kHz	22	2	2	10	11	23	1	1	11	11
	4 kHz	33	0	0	11	11	33	0	0	11	11

error microphones compared to the reference microphones in each case; these are also included in Fig. 9. From these results, it can be observed that the ANC system achieves maximum attenuation of the noise, but this system also attenuates the speech signal from the look direction as shown in Fig. 8. That is, the ANC system attenuates the incident noise irrespective of its direction.

In the HT system, the signal detected by the reference microphones is reproduced by the loudspeakers inside each ear cup, which results in a lower level of attenuation being provided by the HT and ANC+HT systems as shown in Fig. 9. It should be noted, however, that in the case of the ANC+HT system, the attenuation is enhanced by the ANC subsystem in the lower octave bands compared to the HT system. These results are summarised in Table I, and from this summary it can be observed that the ANC+HT system does not provide significant noise attenuation in the speech bandwidth, as indicated by the attenuations achieved in the 500 Hz, 1 kHz, 2 kHz and 5 kHz octave bands. The average attenuation achieved by the ANC+HT system is 6 dB for the left ear and 4 dB for right ear.

In comparison to the HT system, the BF system focuses on the sound field incident from the desired direction, as shown in Section III, whilst limiting the sound reproduced by the headphones due to sources incident from the undesired directions. This means that the reproduction of sounds from the undesired directions will be lower than in the case of the HT system. This means that the noise attenuation performance achieved by the ANC+BF system (pink line) is greater than that achieved by the ANC+HT system (cyan line) for broadband noise, as shown by the results presented in Fig. 9. The advantage of the ANC+BF system over the BF system is that the ANC+BF system provides noise attenuation in the low-frequency octave bands, as can be seen from Table I and Fig. 9. Additionally, it can be seen that the ANC+BF system (pink line) provides significant attenuation for the broadband noise in the speech bands. From Table I it can be evaluated that the average broadband attenuation achieved by the ANC+BF system is 11 dB and 12 dB at the left and right ears respectively. In particular, it should be noted from Table I that the ANC+BF system achieves around 10-11 dB attenuation in the 1 kHz-4

kHz octave bands, while the ANC+HT only achieves between 1 and 2 dB of attenuation.

V. CONCLUSIONS

This paper presents a directional hear-through active noise control headphone system that provides a hear-through capability to the consumer in the look-direction of a microphone array, whilst cancelling the ambient noise emanating from other directions. This system thus presents an extension over the current state-of-the-art, in that it augments the consumer experience by both actively reducing unwanted ambient noise, whilst allowing sound in a particular direction to be fed-through. An experimental study has been conducted to demonstrate the performance of the proposed ANC headphones compared to conventional ANC, beamformer and direct hear-through based headphones. It has been demonstrated that both the directional and direct hear-through configurations are able to provide hear-through functionality for a wanted speech signal, but the directional hear-through ANC headphones (ANC+Beamformer) provide 6-7 dB more attenuation in the unwanted noise compared to the ANC+Direct hear-through headphones, especially in the higher octave bands. In addition, by using beamformer weights for different directions, the consumer can select the region over which the hear-through functionality will operate.

REFERENCES

[1] E. D. Simshauser and M. E. Hawley, "The noise-cancelling headset—an active ear defender," *J. Acoust. Soc. Amer.*, vol. 27, no. 1, pp. 207–207, 1955.
[2] C. Hansen, S. Snyder, X. Qiu, L. Brooks, and D. Moreau, *Active control of noise and vibration*. CRC press, 2012.
[3] S. J. Elliott, *Signal Processing for Active Control*. London: Academic Press, 2001.
[4] L. R. Ray, J. A. Solbeck, A. D. Streeter, and R. D. Collier, "Hybrid feedforward-feedback active noise reduction for hearing protection and communication," *J. Acoust. Soc. Amer.*, vol. 120, no. 4, pp. 2026–2036, 2006.

[5] C. Chang, A. Siswanto, C. Ho, T. Yeh, Y. Chen, and S. M. Kuo, "Listening in a noisy environment: Integration of active noise control in audio products," *IEEE Trans. Consum. Electron.*, vol. 5, no. 4, pp. 34–43, Oct 2016.

[6] W. S. Gan and S. M. Kuo, "An integrated audio and active noise control headset," *IEEE Trans. Consum. Electron.*, vol. 48, no. 2, pp. 242–247, 2002.

[7] D. A. Cartes, L. R. Ray, and R. D. Collier, "Experimental evaluation of leaky least-mean-square algorithms for active noise reduction in communication headsets," *J. Acoust. Soc. Amer.*, vol. 111, no. 4, pp. 1758–1771, 2002.

[8] W. S. Gan, S. Mitra, and S. M. Kuo, "Adaptive feedback active noise control headset: implementation, evaluation and its extensions," *IEEE Trans. Consum. Electron.*, vol. 51, no. 3, pp. 975–982, 2005.

[9] H. Bao and I. M. Panahi, "A novel feedforward active noise control structure with spectrum-tuning for residual noise," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, 2010.

[10] J. Cheer, V. Patel, and S. Fontana, "The application of a multi-reference control strategy to noise cancelling headphones," *J. Acoust. Soc. Amer.*, vol. 145, no. 5, pp. 3095–3103, 2019.

[11] R. Ranjan and W. Gan, "Natural listening over headphones in augmented reality using adaptive filtering techniques," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 11, pp. 1988–2002, Nov 2015.

[12] V. Valimaki, A. Franck, J. Ramo, H. Gamper, and L. Savioja, "Assisted listening using a headset: Enhancing audio perception in real, augmented, and virtual environments," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 92–99, March 2015.

[13] S. Yao, "Headphone-based immersive audio for virtual reality headsets," *IEEE Trans. Consum. Electron.*, vol. 63, no. 3, pp. 300–308, August 2017.

[14] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, and G. Lorho, "Augmented reality audio for mobile and wearable appliances," *J. Audio Eng. Soc.*, vol. 52, no. 6, pp. 618–639, 2004.

[15] S. Serafin, M. Geronazzo, C. Erkut, N. C. Nilsson, and R. Nordahl, "Sonic interactions in virtual reality: state of the art, current challenges, and future directions," *IEEE Comput. Graph. Appl.*, vol. 38, no. 2, pp. 31–43, 2018.

[16] J. Bitzer and K. U. Simmer, "Superdirective microphone arrays," in *Microphone arrays.* Springer, 2001, pp. 19–38.

[17] S. M. Kim, C. J. Chun, and H. K. Kim, "Multi-channel audio recording based on superdirective beamforming for portable multimedia recording devices," *IEEE Trans. Consum. Electron.*, vol. 60, no. 3, pp. 429–435, Aug 2014.

[18] R. Tong and Z. Ye, "Enhanced, blind and robust far-field audio acquisition for portable devices," *IEEE Trans. Consum. Electron.*, vol. 63, no. 1, pp. 62–67, February 2017.

[19] M. F. Simón Gálvez, S. J. Elliott, and J. Cheer, "A superdirective array of phase shift sources," *J. Acoust. Soc. Amer.*, vol. 132, no. 2, pp. 746–756, 2012.

[20] S. J. Elliott, J. Cheer, H. Murfet, and K. R. Holland, "Minimally radiating sources for personal audio," *J. Acoust. Soc. Amer.*, vol. 128, no. 4, pp. 1721–1728, 2010.

[21] M. Bouchard and Y. Feng, "Inverse structure for active noise control and combined active noise control/sound reproduction systems," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 9, no. 2, pp. 141–151, Feb 2001.

[22] T. Goehring, J. L. Chapman, S. Bleeck, and J. J. Monaghan*, "Tolerable delay for speech production and perception: effects of hearing ability and experience with hearing aids," *Int. J. Audiol.*, vol. 57, no. 1, pp. 61–68, 2018.

[23] G. W. Elko, "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication.* Springer, 2000, pp. 181–237.

[24] M. F. S. Galvez, "Design of an array-based aid for the hearing impaired," November 2014. [Online]. Available: https://eprints.soton.ac.uk/375092/

[25] L. L. Beranek and T. Mellow, *Acoustics: sound fields and transducers.* Academic Press, 2012.

Vinal Patel (M'19) received his Bachelor of Technology degree in Electronics and Telecommunication Engineering from the Chhattisgarh Swami Vivekananda Technical University, Bhilai, in 2010 and Master of Technology degree in Signal Processing from the Indian Institute of Technology, Guwahati, in 2013, and Ph.D. degree in Electrical Sciences from Indian Institute of Technology Gandhinagar, in 2018. Currently he is working as an Assistant Professor in the department of Information Technology at ABV-Indian Institute of Information Technology and Management Gwalior. His research interests include active noise control, adaptive signal processing and neural network.

Jordan Cheer received the B.Mus. (Tonmeister) degree in music and sound recording from the Institute of Sound Recording, University of Surrey, Guildford, U.K., in 2008, and the M.Sc. and Ph.D. degrees in sound and vibration studies from the Institute of Sound and Vibration Research (ISVR), University of Southampton, Southampton, U.K., in 2009 and 2012, respectively.

From 2012 to 2014, he was a Research Fellow in Active Control at the ISVR. He was promoted to a Senior Research Fellow in 2015, Lecturer in 2017 and currently holds an Associate Professorship at the same institution. His research has focused on the active control of noise and vibration and array signal processing for audio reproduction. He has published 80 papers in journals and conference proceedings.

Dr. Cheer is a Subject Editor for the Journal of Sound and Vibration and an Associate Editor for the Journal of the Acoustical Society of America. He is a Member of the International Institute of Acoustics and Vibration, a Fellow of the Higher Education Academy, a Member of the Institute of Mechanical Engineers and a Chartered Engineer. He received the P. E. Doak Award for academic performance in the taught part of the M.Sc., the E. J. Richards Prize for the best masters dissertation, and the Sir James Lighthill Award for the best student paper at the 19th International Congress on Sound and Vibration. He has served on the scientific committees for a number of international conferences including the International Congress on Sound and Vibration and the 2016 Audio Engineering Society Conference on Sound Field Control.

Simone Fontana received the M.Sc. degree in telecommunication engineering from Politecnico di Torino, Torino, Italy, and Telecom ParisTech, Paris, France, and the Ph.D. degree in audio signal processing and technical physics from Telecom ParisTech, Paris, France, and University of Parma, Parma, Italy. Since 2013, he works in Huawei European Research Center as Senior Research Fellow and Technical Team Leader. His current research interests include spatial audio processing, speech enhancement, and new transducers for communication devices.