A Review of Sound Field Control

Jun Yang 1,2,*, Ming Wu 1,2 and Lu Han 1,2

1 Key Laboratory of Noise and Vibration Research, Institute of Acoustics, Chinese Academy of Sciences, No. 21 North 4th Ring Road, Haidian District, Beijing 100190, China; mingwu@mail.ioa.ac.cn (M.W.); hanlu@mail.ioa.ac.cn (L.H.)
2 School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, No. 19(A) Yuquan Road, Shijingshan District, Beijing 100049, China
* Correspondence: jyang@mail.ioa.ac.cn

Abstract: Sound field control (SFC) technology enables the active management of audio delivered within an acoustical environment. It includes three research directions: sound field reproduction, personal audio systems, and active noise control. Sound field reproduction uses loudspeaker arrays to replicate a sound field in a target region; personal audio systems extend sound field reproduction over multiple regions so that different listeners can hear personalized audio in a shared space; and active noise control aims to cancel the original sound field in the target area by generating a secondary sound field. In this paper, we briefly review the advances of the three different types of techniques with a discussion of their algorithms and applications.

Keywords: sound field reproduction; personal audio system; active noise control

1. Introduction

Sound field control (SFC) technology aims to generate the desired sound field over a target area through secondary sound sources or actuators. It has a wide range of applications and generally includes sound field reproduction, personal audio systems, and active noise control.

Sound field reproduction technology reconstructs the original sound field using secondary sound sources, so that the listener can obtain an immersive listening experience. Sound field reproduction can be divided into binaural sound field reproduction and large region sound field reproduction according to the size of the control area. Binaural sound field reproduction uses headphones or loudspeakers to reproduce the desired sound field in both ears [1–4] which can find applications in augmented reality (AR) and virtual reality (VR) products. It is essentially a virtual sound field reproduction method based on the perception characteristics of human ears. Large region sound field reproduction creates the desired sound field in an extended region [5–9]. This reproduction method has a larger listening area compared with the binaural one but requires a larger number of loudspeakers. Therefore, its practical application in entertainment is limited, and this method is more suitable for laboratory use and demonstration.

The control area of the sound field reproduction technology is a single listening area, and the personal audio systems require that the original sound field is reproduced in the listening area while unwanted interferences are reduced in other areas [10–12]. Therefore, the control area for personal audio systems is divided into two kinds of zones: the acoustic bright zone, a listening area where the desired sound field is reproduced; and the acoustic dark zone, a quiet area where the sound pressure level needs to be suppressed. Personal audio systems can be used on TVs and tablet computers to produce directional sound to avoid interference with people around and are used in museums and other occasions to avoid the influence of sound from other sources (e.g., when the audio guides of multiple exhibits operate simultaneously).
Different from sound field reproduction and personal audio systems which reproduce the desired sound field in a certain area, active noise control technology can suppress the original sound field in the target area to create a quiet environment for the listeners through generating a secondary sound field with the same amplitude as and the opposite phase to the original noise field according to the interference principle of sound waves [13–15]. Compared with passive control technologies, the active ones are more suitable for low frequency noise. Active noise control can be divided into binaural active noise control, local area active noise control, and global active noise control from the size of the control area. A typical application of binaural active noise control is the active noise control headphones, where the control area is at the eardrum. Local area active noise control generally refers to the reduction in noise in the user’s head area, such as an active noise control headrest. Global active noise control is to achieve noise reduction in an entire space, such as the active control of cabin interior noise and aircraft radiated sound.

This article introduces some of the research progress in sound field reproduction technology, personal audio systems, and active noise control technology in recent years.

2. Sound Field Reproduction

A sound field reproduction system generates the desired sound field over a target region using loudspeaker arrays to provide an immersive listening experience for listeners. In this section, some technologies of sound field reproduction are discussed, during which sound field recording methods for acquiring the desired sound field to be reproduced are introduced first.

2.1. Sound Field Recording

Sound field recording is the foundation for reproducing the desired sound field precisely. This technology aims to accurately estimate the distribution of the original sound field in an entire target area through limited acoustical data measured by microphone arrays.

Expressing a 3D sound field into spherical harmonics reveals the underlying characteristics of the sound field, and thus enables a high accuracy recording of the sound field [16–19]. The spherical harmonics expansion for sound pressure at any point in a source-free region is of the form:

\[
p(r, \theta, \phi, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \alpha_n^m(k) j_n(kr) Y_n^m(\theta, \phi)
\]

where \(r\), \(\theta\), and \(\phi\) are radius, elevation, and azimuth angles, respectively; \(k = 2\pi f / c\) is the wave number; \(f\) is the frequency; \(c\) is the speed of the sound propagation; \(j_n(kr)\) is the \(n\)th order spherical Bessel function; \(Y_n^m(\theta, \phi)\) is the spherical harmonic of order \(n\) and degree \(m\); and \(\alpha_n^m(k)\) is the spherical harmonic coefficients. The sound field in the target area can be represented by these spherical harmonic coefficients.

Because the sound pressure received by a spherical array is suitable for the spherical harmonic transform [18], the spherical microphone array is considered to calculate the above-mentioned coefficients. However, the placement of the microphones in a spherical array must follow a strict rule. A great number of microphones are required especially when the target area is large and the frequency is high. Therefore, the practical use of spherical arrays is limited. An array of higher order microphones (HOMs) has been proposed to decrease the number of microphones [20,21]. The array of HOMs is distributed enclosing the large area of interest. The local spatial recordings by the HOMs are translated to the global sound field coefficients of interest based on the Graf’s addition theorem. This method replaces the omnidirectional microphone with the HOM, which reduces the microphone units of an array at the expense of the additional complexity of each unit. Microphone arrays with 2D geometry are easier to implement than a spherical array and were therefore designed to record 3D sound fields [22]. The planar array utilizes both omnidirectional and first-order microphones to calculate the even and odd coefficients, respectively. The
proposed 2D microphone array in [22] has good accuracy of sound field recording, and its geometry is more feasible than traditional spherical microphone arrays in real-world applications.

2.2. Typical Techniques for Sound Field Reproduction

Sound field reproduction includes binaural sound field reproduction and large region sound field reproduction. Binaural sound field reproduction is a virtual sound field reproduction method based on the perception characteristics of human ears, which does not generate the original sound field in the spatial region. The best listening area of this technology is generally small. Sound field reproduction over a large-scale region aims to provide better listening experience to more listeners. There exist a number of techniques for the reproduction: pressure matching (PM), higher order Ambisonics (HOAs), and wave field synthesis (WFS).

The PM method based on the least-square (LS) principle aims to minimize the sum of squared errors between the desired and reproduced sound pressure at the control points [9]. The conventional PM method usually activates all the candidate loudspeakers to minimize the reproduction error. However, in many cases, the desired sound field can be accurately reproduced using only a few loudspeakers. The least absolute shrinkage and selection operator (Lasso) can be adopted to limit the number of active loudspeakers and optimize the positions of the loudspeakers. The two-stage Lasso-LS optimization further reduces the reconstruction error [23]. The first stage using Lasso can effectively reduce the number of activated speakers, and then the reproduction error is reduced by the second stage using LS. However, these algorithms are designed in the frequency domain, which means that the loudspeaker weights are independently optimized at each control frequency. The PM method in the time domain combined with the framework of the group Lasso can optimize the number and position distribution of loudspeakers at all control frequencies simultaneously [24]. In recent years, some studies introduce the velocity matching [25,26]. When the loudspeaker array is non-uniformly spaced, the velocity matching method can improve the accuracy of the reproduced sound pressure and intensity direction.

The HOA approach expands the sound field in orthogonal modes and minimizes the modal coefficients error between the original sound field and the reproduced sound field. It was first proposed by Gerzon in 1973 [27] based on first-order spherical harmonic decomposition, which uses four loudspeakers located at the vertices of a regular tetrahedron. This approach is only applicable to small-scale sound field reproduction. The HOA based on higher order spherical harmonic decomposition was then proposed to enlarge the reproduction area [7,28,29]. Ward and Abhayapala [8] analyzed the reproduction error based on the HOA method and presented the quantitative relationship between the radius and frequency of the reproduction area, the order of the spherical harmonic decomposition, and the number of loudspeakers required. The performance of the HOA reproduction in terms of the size of reproduction area and the upper frequency limit improves with the increased order of the spherical harmonic decomposition. An Nth-order 3D Ambisonics requires at least \((N + 1)^2\) loudspeakers.

The concept of WFS was introduced by Berkhout in 1993 [5] based on the Kirchhoff–Helmholtz integral. The sound field in a bounded and source-free region is determined by the pressure and normal velocity on the boundary of the region. This boundary condition is generated by a continuous distribution of secondary sources. Then the transformation from continuous sources to discrete loudspeakers leads to spatial aliasing effects, which were analyzed in [30]. Early WFS research focused on the sound field reproduction in a 2D plane using linear and planar loudspeaker arrays [31,32]. The 2.5D WFS was proposed to synthesize the field of 3D sources using 2D array configurations by deriving the 2.5D operator [33,34]. Local Sound Field Synthesis (LSFS) features a more accurate reproduction in a small region. The region is smaller than the area surrounded by the loudspeaker distribution. Winter et al. [35] proposed two LWFS techniques to improve the reproduction accuracy compared with conventional WFS.
The HOA and WFS methods are both able to generate the desired sound field in a single and large region, but they require a large number of loudspeakers. Moreover, the HOA method imposes additional requirements on the array structure, i.e., the use of spherical loudspeaker arrays. The PM method has no specific requirements for the number and structure of sound sources. Compared with the WFS and HOA methods, the PM method only reproduces the sound pressure at the control points but improves the operability in the application.

2.3. Applications of Sound Field Reproduction

Sound field reproduction technology can be applied to entertainment places such as cinemas and concert halls. Representative companies are IOSONO in Germany and Sonic Emotion in Switzerland. They have many successful engineering cases. Sound field reproduction for virtual reality (VR) provides users with the feeling of being presented in natural auditory environments. Sonke [36] tried to combine wavefield analysis and the WFS method into virtual reality, and Bruijin [37] applied the WFS system to the video conference scene. There are also some products for HOA systems. For example, BlueRippleSound developed the VST2 series software using the third-order Ambisonics, which can encode and decode the sound field with the third-order Ambisonics. In addition, it also provides a platform for acoustic measurement. By simulating a real noise environment, it is used for the subjective and objective evaluations of acoustic equipment such as active noise cancellation headphones and speech recognition systems.

3. Personal Audio Systems

3.1. Basic Theory

Personal audio systems aim to focus sound to a listening zone without disturbing others, which can be realized by multi-zone sound field reproduction technologies. Multi-zone sound field reproduction creates two different kinds of zones: the bright zone where a desired sound field is reproduced, and the dark zone where the sound pressure level needs to be attenuated. Figure 1 is the schematic diagram of the multi-zone sound reproduction. Assuming that there are M discrete control points in the target area, the sound pressure generated by the loudspeaker array composed of L units in the target area can be expressed as:

\[ p = Gq \]  

(2)

where \( p = [p(r_1), p(r_2), \cdots, p(r_M)]^T \) is the sound pressure vector at all control points, \( G \) is the acoustic transfer function matrix, and \( q = [q_1, q_2, \cdots, q_L]^T \) is the weight vector. The matrix \( G \) reads:

\[
G = \begin{pmatrix}
g(r_1|s_1) & \cdots & g(r_1|s_L) \\
\vdots & \ddots & \vdots \\
g(r_M|s_1) & \cdots & g(r_M|s_L)
\end{pmatrix}
\]  

(3)

where \( g(r_m|s_l) \) represents the transfer function from the \( l \)th loudspeaker to the \( m \)th control point.

![Figure 1. Schematic diagram of multi-zone sound reproduction.](image-url)
The primary evaluation metric for a personal audio system is the acoustic contrast, which is defined as the ratio of acoustic potential energy density between the bright and the dark zone:

$$\max \frac{q^H R_b q}{q^H R_d q}$$

where $R_b = G_b^H G_b$ and $R_d = G_d^H G_d$ are the spatial correlation matrices of the bright and dark areas, respectively.

The acoustic contrast control (ACC) method obtains the weight vector by maximizing the acoustic contrast. However, the main drawbacks of the ACC method are the lack of phase control and the uneven distribution of energy in the bright zone. To deal with these problems, pressure matching (PM) [38] and planarity control (PC) [39] were proposed. As mentioned above, the PM method minimizes the errors between the desired and reproduced sound pressure at the control points. In personal audio systems, the desired sound pressure in the bright zone is set according to the target sound field, and the desired sound pressure in the dark zone is equal to zero. The PC method uses superdirective beamforming for spatial filtering of the sound pressure at the control points and then maximizes the bright zone energy after spatial filtering with the dark zone energy constraints.

Both the PM and PC methods can reproduce a directional sound field in the bright zone, but their acoustic contrasts are less than that of the ACC method. Considering that the main purpose of personal audio system is the higher acoustic contrast, the ACC method is more appropriate. For the lack of phase control and other issues, many studies have proposed improvements on robustness and sound quality based on the ACC method. These improvements are described in detail in the following subsection.

### 3.2. Algorithm Improvements

The ACC method focuses more on acoustic contrast performance, but the phase control and sound pressure distribution in the bright area are neglected, which affects the subjective perception of the listener. The PM method is concerned with the accurate reproduction of sound fields. If the two types of methods are combined, the system can take the advantages of both. The ACC-PM method [40] obtains the compromise effect between the reproduction error in the bright zone and the acoustic contrast by adjusting the weight factor. The SFR-ACC method [41] is to minimize the reproduction error with the acoustic contrast to not be lower than a given value. The latter can precisely adjust the acoustic contrast performance as needed.

In practical applications, the acoustic transfer functions can be perturbed by such factors as inconsistencies in the sensitivities of loudspeaker, mismatch in positions of loudspeaker, and the scattering from the obstacles. Moreover, the performance of reproduction can be sensitive to errors in the acoustic transfer functions. Therefore, the robustness of reproduction systems is concerned. Elliott et al. [42] studied the robustness and regularization problem of the ACC method and increased the robustness by limiting the array effort. The regularization method is commonly used for robustness improvement, but it is difficult to select a good regularization parameter. The best regularization parameter can be found by Monte-Carlo simulations using the information about errors in the system, but the computational complexity is high. Zhu et al. [43] proposed a design framework for robust reproduction as shown in Figure 2 and provided a method to quickly estimate the regularization parameters. They modeled the acoustic transfer function as the superposition of the deterministic part ($G$) and the perturbation part ($\Delta G$):

$$\tilde{G} = G + \Delta G.$$  

The perturbation part describes the transfer function variations due to possible errors. Based on the model, two strategies are provided. One is to estimate the bound of $\Delta G$ and achieve good performance in the worst acoustic situation. The other is to estimate the probability distribution of $\Delta G$ and establish the mean cost function to improve the average
performance over all possible cases. Through such robust control, the ability of the system to resist disturbance is improved.

The methods mentioned above are all designed to optimize the weight vectors at a set of discrete frequencies. However, the acoustic contrast can decrease significantly at non-control frequencies. Moreover, the designed method has the causality problem when transforming the weight vectors from the frequency domain to the time domain. To tackle these problems, Elliot and Cheer [44] presented a broadband acoustic contrast control (BACC) method by calculating the coefficients of the filters directly. However, the BACC method cannot control the consistency of frequency responses, which may not ensure a flat frequency response at a point in the bright zone and may lead to intolerable distortions. The BACC-RV method introduced a response variation (RV) term into the BACC method [45], which can effectively solve the frequency response consistency problem and improve the sound quality in the bright zone. In [45], a linear loudspeaker array consisting of eight loudspeaker units was used to generate the multi-zone sound field. The bright zone and dark zone were located at $45^\circ$ and $-45^\circ$, respectively, deviating from the loudspeaker array center. The sampling rate and the length of filter are set to 8000 Hz and 200, respectively, and thus the control frequency interval is 40 Hz. The weight vectors were computed according to traditional acoustic contrast control (TACC), BACC, and BACC-RV. The performance of the acoustic contrast at each frequency was compared as provided in Figure 3. To demonstrate the acoustic contrast at non-control points, the observed frequency interval is 10 Hz. Note that the acoustic contrast of TACC is high at control frequencies but rapidly decreases when the frequency deviates from the control frequency. BACC and BACC-RV both alleviate this problem, reducing the difference in contrast between uncontrolled and controlled frequencies. However, compared with BACC, BACC-RV has a larger acoustic contrast. Figure 4 shows the frequency response at the center of the microphone array in the bright zone. The frequency response of TACC has obvious fluctuations over the entire frequency range. Moreover, the frequency response of BACC is similar to line spectrum,
which performs worst. To maximize the acoustic contrast in the time domain, the BACC method only retains the frequency component with the highest acoustic contrast in the frequency domain. With the RV constraint, BACC-RV achieves the best frequency response consistency among all three methods.

The above methods implemented in the frequency domain or in the time domain utilize the cost function at discrete spatial control points. Thus, the performance at the spatial non-control points is degraded. The method designed in the modal domain can solve this problem [46]. The energy in the bright and dark zones is transformed into the modal domain. Modal coefficients contain complete sound field information at any point in a space. Therefore, the energy is controlled in the entire region, rather than at discrete control points. Moreover, the reproduced sound field has a more uniform energy distribution.

3.3. Applications of Personal Audio Systems

Personal audio systems, designed to meet the different listening needs of different listeners, improve the privacy of audio transmission and reduce interference to others significantly. Such applications potentially include directional audio guides in museum and galleries, private communication devices in banks and hospitals, and personalized entertainment devices in cabins and other public places.

Figure 3 shows several typical microphone and loudspeaker arrays for personal audio systems. Figure 5a shows a microphone array for sound field recording, and Figure 5b-d...
show several different personal audio systems. In Figure 5b, a linear loudspeaker array is used to control the sound field. The bright and dark zones are generated at the two microphone arrays in this figure, respectively. Figure 5c shows a circular array, and the bright and dark zones are inside the circular area. Another circular array provided in Figure 5d is able to achieve sound field radiation in a certain direction.

Elliott and Jones investigated the application of the ACC method in a headrest [47]. Cheer et al. implemented a two-source personal audio system on a mobile phone [48]. The personal audio systems in a car cabin were also implemented [49] using two loudspeaker arrays to achieve independent listening zones over the wide frequency ranges. A significant level of acoustic contrast was achieved when a bright zone was produced in the front seats. The bright zone of this system was expanded to either the front or rear seats by Liao et al. [50]. In this study, a loudspeaker array was mounted on the ceiling to generate independent listening zones at higher frequencies, and the array locations are selected from some of the configurations under study. Compared with the system reported in [49], the improved one achieved a significant boost in contrast performance between 1 and 4 kHz, and more consistent contrast when generating either a front or a rear bright zone.

Figure 5 shows several typical microphone and loudspeaker arrays for personal audio systems [51](c,d) circular loudspeaker arrays for personal audio systems.

4. Active Noise Control

Active noise control (ANC) can suppress the original sound field in a target area to create a quiet environment for the listeners. Since its first introduction in 1936 [52], this approach has been greatly developed. The ANC system can efficiently attenuate low frequency noise. Due to excellent performance, ANC systems have been widely used to control the vibration or noise of headphones, cabins, and boards. In this section, we introduce ANC in aspects including the system model, algorithm design, and application.

4.1. ANC System Model

An ANC system mainly consists of reference microphones, controllers, and error microphones. The reference signal vector $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-L+1)]$ is
processed by the controller \( w(n) = [w_0(n), w_1(n), \ldots, w_L(n)] \) to produce an output signal:

\[
y(n) = w^T(n)x(n),
\]

where \( L \) is the length of the control filter. The error signal \( e(n) \) measured by the error sensor represents the acoustical combination of the desired signal \( d(n) \) and the output signal [13–15, 53, 54]. The control filter is updated as:

\[
w(n + 1) = w(n) + \mu x'(n)e(n),
\]

where \( x'(n) \) is filtered-x signal obtained by convolving the reference signal and the secondary path \( S(z) \).

ANC systems can be divided into narrowband systems and wideband systems according to the noise bandwidth to be eliminated. Narrowband systems are mainly used to handle periodic (harmonic) noise. In a narrowband system, the output of the control filter is the harmonic frequency signal

\[
y(n) = w_1 x_\omega(n) + w_2 x_\theta(n),
\]

\[
x_\omega(n) = \cos(\omega n), \quad x_\theta(n) = \sin(\omega n),
\]

where \( \omega \) is the frequency of narrowband signal, and \( w_1 \) and \( w_2 \) are the control filter coefficients. Each periodic noise requires independently a two-order control filter. Frequency mismatch refers to the inconsistency of the frequency of the reference signal and the noise signal, which destroys the noise reduction performance. Solving the frequency mismatch is an important issue for narrowband ANC systems [55–58]. By contrast, broadband systems use a high-order filter to deal with broadband noise and do not need to control each frequency point. In the presence of wideband and narrowband noise components, it is necessary to separate the narrowband noise first and then process the narrowband noise and the wideband noise independently.

Alternatively, ANC systems can be divided into feedforward and feedback systems depending on the control structure [59–62]. Feedforward systems are a classic type of ANC systems, which obtain a reference noise signal through a reference sensor and then performs noise reduction at the error point. Feedforward systems have good stability and have thus been widely used. The block diagram of a single channel feedforward ANC system is shown in Figure 6. However, this system requires high coherence between the reference signal and the primary noise signal, which often requires experienced selection of the reference point. Feedback systems, on the other hand, require no reference microphone and are suitable for occasions where the reference signal is not easy to obtain. Therefore, feedback systems perform well in consumer electronics with low cost and complexity such as a noise-canceling headset. However, the feedback algorithm needs to make a trade-off among water bed effect, stability, and noise amplification. Moreover, hybrid feedforward and feedback systems can improve the noise reduction performance with the poor coherence between the feedforward reference signal and the noise signal.

![Figure 6. Block diagram of a single channel feedforward ANC system.](image-url)
4.2. Controller Algorithm

The algorithm of the controller is the most important part of an ANC system, which directly determines the noise reduction performance. Convergence speed, computational complexity, and robustness are important metrics for algorithm evaluation.

Faster convergence speed and lower steady-state error result in better noise control. The speed of convergence of the algorithm depends on the condition number of the covariance matrix of the filtered-x signal, which is the ratio of the largest eigenvalue to the smallest eigenvalue. The classical filtered-x least mean square (FxLMS) algorithm has been widely used due to its simplicity and efficiency. Because the condition number of the covariance matrix may be very large, the FxLMS algorithm has a slow convergence rate. Therefore, the filtered-x recursive least square (FxRLS) and filtered-x affine projection (FxAP) algorithms [63] use information from past moments to whiten the signal and reduce the condition number. Moreover, an algorithm for whitening the filtered-x signal with an inverse secondary path was also proposed. Since the inverse secondary path is a part of the Hessian matrix, the algorithm is called the filtered-x Newton (Fx-Newton) algorithm [64].

The speed of convergence can also be increased by changing the step size instead of changing the matrix condition number. The variable step-size algorithm uses a large step size to quickly converge in the initial stage and uses a small step size to obtain a smaller residual error in the steady-state stage [65].

The reduction in algorithm complexity is mainly achieved by computation optimization and control strategy. The calculation amount of the ANC algorithm is mainly concentrated in three aspects: filtered signal generation, output signal generation, and controller update. The frequency-domain algorithm uses the transform domain [66,67], and the subband algorithm [68] uses the multi-rate method to reduce the computational cost of the calculation of the filtered reference and output signal [69,70]. The partial update algorithm updates only part of the filter coefficients per sample period, instead of all coefficients, which reduces the cost of controller update [71].

Multi-channel systems have more secondary sources and error microphones [72,73], and thus they may control noise over a larger area. However, as the number of secondary sources and error points increases, the computational complexity of multi-channel systems increases dramatically. The filtered error algorithm convolves secondary path and the error signal instead of the reference signal, which can significantly reduce the complexity of the filtered error signal generation, especially in multiple reference systems. Multi-channel algorithms can reduce the complexity through control strategies, mainly turning centralized algorithms into decentralized or distributed algorithms. The centralized algorithm has only one computing core in the center, and the computational burden is heavy. The decentralized algorithm splits the multi-channel system into multiple single-channel systems. The split system does not need to consider the cross secondary path, thereby reducing the filtered reference signal and controller update. The distributed method introduces multiple computable nodes through the network, thus reducing the computational cost of each computational node [74,75].

Robustness refers to the ability of an algorithm to cope with changes in the environment. The changes in the environment mainly include variations in noise and secondary paths that require different optimal controllers. In headrest, all secondary paths are first identified offline. When the head moves, the controller switches to the optimal filter according to the head position determined by other sensors [76,77]. We can also design parallel filters that match the noise to remove different types of noise [78]. Online modeling is a way to deal with a time-varying secondary path and is mainly divided into two types: additional noise modeling and noiseless modeling.

Nonlinearity is introduced due to the nonlinear effects of the secondary path, secondary source, and primary path [60]. It limits adaptive algorithm performance. Methods such as the Volterra filter [79], optimization algorithm, and neural network [80] can predict and eliminate nonlinear terms.
Theoretical analysis can provide a better understanding of the algorithm and provide a guide to its practical applications. Through theoretical analysis, we can improve the algorithm in a goal-oriented and effective way. In recent years, the re-analysis of some classic LMS-based algorithms remove the assumption of slow convergence and is suitable for arbitrary input distributions. Such analysis provided us a more accurate prediction of the transient behavior and steady-state performance [81–85].

4.3. ANC Applications

ANC facilitates the protection of instruments or human hearing from damage by controlling noise or vibration. It has been widely used in noise control and vibration elimination of industrial and mechanical equipment, such as pipeline noise, aircraft noise, car noise, magnetic resonance imaging (MRI) noise, etc.

Noise-canceling headphones are a typical one-dimensional noise reduction scenario and may be the most successful application of ANC, as shown in Figure 7a. Because of low computational cost and flexibility, fixed and digital controllers are often used to implement both feedforward and feedback control [62,86]. The feedforward controller may be designed by approximating the optimal Wiener solution regardless of stability. In feedback ANC headphones, the noise reduction performance, noise amplification, and robustness to secondary path variations should be considered for the design of the controller [15]. Most ANC feedback controller design methods have employed the ad-hoc techniques named loop shaping. Rafely assumed a multiplicative uncertainty model and optimized the controller with $H_2$ performance criteria and $H_{\infty}$ constraints of secondary path uncertainty [87]. However, these methods require some guess work and personal experience. In Figure 7a, the controller is implemented by a codec ADAU1772 with 44.1 kHz sampling rate. This codec has several biquad filters to approximate target controller response [88]. Recently, chips with higher sampling frequencies such as 768 kHz are developed and applied to commercial headphones, expecting lower latency and noise reduction bandwidth.

![Figure 7. ANC applications. (a) is the ANC headphone in [88], (b) is the pipe muffler in [55], (c) is the adjustable ANC headrest for electric vehicles [77], (d) is the local quiet zone in [76].](image)

Pipe mufflers in Figure 7b are another one-dimensional noise reduction application which involves acoustic feedback and effects of airflow and temperature. Because the primary noise is usually periodic, non-acoustic reference signals such as revolving speed, acceleration signal, and generated tonal signal are widely used to remove the secondary...
path feedback effect. In the case of narrowband active noise control, frequency mismatch (FM) between estimated frequencies and true frequencies may decrease the noise reduction and even make the system unable to work properly in some extreme circumstances. Some improvements have been made in the narrow-band ANC system involving variable step-size algorithms [89] and frequency estimation algorithms [90]. In Figure 7b, a circular pipe system with an active-passive hybrid compound muffler was established. The primary source is set at the front-end of the pipe system while the secondary source at the after-end of the muffler. Signal processing is performed online by a controller with a development board Omap-L137 based on TMS320C6747. This system achieves good reduction performance of multiple tones with time-varying frequencies.

For spatial noise reduction, active headrests in Figure 7c are also very promising designs. They do not need to be worn on the ears. By placing speakers around the seats or windows, a good noise reduction can be obtained near the ears. In Figure 7d, active headrests are used to control the noise around passengers’ ears in cabins, improving the robustness when the head is rotated. In order to expand the noise reduction area, existing ANC systems are often improved from perspectives of combining optical sensors, virtual sensing technology, and modal domain algorithms. Cabin and aircraft noise control is interior noise control. In this regard, piezoelectric materials are utilized to actively reduce the structural sound radiation, through which enclosed sound fields are controlled [91–94].

Combined with psychoacoustics, ANC systems tend to achieve a comfortable residual sound field rather than simply a quiet one. Specifically, a noise source with relatively high acoustic power may somehow sound “pleasant”, as a result of its loudness, sharpness, roughness, tonality, pleasantness, and other indicators. The algorithm combined with psychoacoustics is mainly to perform weighted filtering in the error signal to obtain the residual sound field in the desired frequency band. This makes the sound field sound more comfortable instead of just making the error signal smaller. Moreover, using music or natural sounds to mask noise is also an interesting method.

5. Future Works

SFC achieves satisfactory results in local and stationary acoustic environments. In the future, in a complex and varying environment, the system can focus on human hearing and achieve effective and robust performance. In the following, combined with the remaining problems in the development process, some future directions are suggested:

- Designing adaptive systems to deal with the variable acoustic environment;
- Combining with psychoacoustics to improve the user’s subjective sense of hearing;
- For personal audio systems, using highly directional loudspeakers to focus more sound energy in the listening area;
- Developing algorithms with low complexity, fast convergence, and good robustness;
- Combining ANC systems and multiple sensors (vision, position, and motion) to control noise more flexibly and achieve noise reduction in a larger area;
- Applying deep learning and communication network to ANC systems.

6. Conclusions

This article presented an overview of the research progress of sound field control technology and separately outlined the algorithms and applications of the three research directions of sound field control, namely, sound field reproduction, personal audio systems, and active noise control. In response to the critical issues in sound field control, researchers have made many improvements. The application of sound field reproduction is limited by the large number of loudspeakers required. Thus, a major improvement direction is to reduce the number of sensors and loudspeakers. For sound field reproduction and personal audio systems, the improvements of reproduction accuracy and sound quality in the listening area are important problems. The performance of ANC systems can be improved by optimizing the sensor and secondary source positions, increasing the robustness, and
reducing the complexity of the algorithm. These improvements are conducive to the realization and application of these technologies.

The practical application of SFC systems must solve the problems of system nonlinearity and adaptability to complex environments such as high reverberation. In order to realize the sound field control of a larger area, the control algorithm of the multi-sensor and multi-speaker system must be optimized, and the setup must be simplified, e.g., the distributed algorithm.

Funding: This research was funded by Frontier Exploration Project Independently Deployed by Institute of Acoustics, Chinese Academy of Sciences under Grant QYTS202009, Natural Science Foundation for Scientists of Shanxi Province under Grant 20210302123107 and 20210302123188.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.


Conflicts of Interest: The authors declare no conflict of interest.

References


44. Elliott, S.J.; Cheer, J. Regularisation and Robustness of Personal Audio Systems; University of Southampton: Southampton, UK, 2011.


56. Han, R.; Wu, M.; Li, F.; Sun, H.; Yang, J. A narrowband active noise control system with a frequency estimator based on Bayesian inference. *J. Sound Vib.* **2019**, *455*, 299–311. [CrossRef]

57. Gong, C.; Wu, M.; Guo, J.; Zhang, Z.; Cao, Y.; Yang, J. Multichannel narrowband active noise control system with a frequency estimator based on DFT coefficients. *J. Sound Vib.* **2021**, *521*, 116660. [CrossRef]


91. Fuller, C.; von Flotow, A. Active control of sound and vibration. *IEEE Control Syst.* 1995, 15, 9–19. [CrossRef]