Journal of Sound and Vibration (1997) **204**(2), 386–396

SOUND FIELDS FOR THE PRODUCTION OF VIRTUAL ACOUSTIC IMAGES

P. A. NELSON, O. KIRKEBY AND T. TAKEUCHI
Institute of Sound and Vibration Research, *University of Southampton*, *Southampton SO*17 1 *BJ*, *England*

AND

H. HAMADA
Department of Information and Communication Engineering, *Tokyo Denki University*, 2-2 *Kanda-Nishiki-cho*, *Chiyoda-ku*, *Tokyo* 101, *Japan*

(*Received* 20 *February* 1997)

1. INTRODUCTION

In stereophonic sound reproduction one generally uses two sources of sound (loudspeakers) that are placed symmetrically in front of a listener in order to subtend an angle of about 60°. The inputs to the loudspeakers are determined in order to produce the illusion in the listener of the existence of a "virtual" source of sound. The classical approach to this [1, 2], which is still widely used to this day, relies on the application of a difference in gain (or time delay) applied to the virtual source signals prior to being input to the loudspeakers. This produces the illusion of a virtual source placed somewhere between the loudspeakers. A more sophisticated approach in which such a loudspeaker arrangement is also used, is that generally attributed to Atal and Schroeder [3], who showed how the virtual source signal could be processed by a pair of linear filters prior to being input to the loudspeakers. These (analogue) filters were designed to ensure that the acoustical signals produced in the region of the listeners' ears were a reasonable approximation to those that would be produced by the virtual source. Such an approach can produce the illusion in the listener that the virtual source is located at a much wider range of spatial positions. The basis of this method has been used more recently by a number of workers using digital signal processing and a variety of filter design techniques [4–7]. The purpose of this note is to point out that this approach can also be made to operate remarkably effectively, and arguably more effectively, by using a pair of loudspeakers that are placed very close together. Such a loudspeaker arrangement appears to have received little attention in the past, although it has been referred to by Bauck and Cooper [8] in a recent paper. These authors also referred to some very early experiments by Lauridsen (reported by Heegaard [9]), who used a combination of a conventional boxed loudspeaker together with an open-backed loudspeaker in an alternative scheme for stereophonic sound reproduction. Such an arrangement can be used to approximate the combination of a monopole and dipole source. In this note it is demonstrated that the use of two closely spaced loudspeakers also approximates such a source combination and that the form of sound field produced makes it particularly suitable for the generation of virtual acoustic images. Preliminary reports of the authors' findings have also been given in references [10–15]. In this preliminary work, such a sound reproduction system has been referred to as a "Stereo Dipole". It is shown in what follows that this is actually a slight misnomer.

0022–460X/97/270386 + 11 $25.00/0/sv970967

2. THEORY

An analysis of a two-loudspeaker virtual source imaging system can be undertaken in the frequency domain in order to describe its main physical characteristics. The geometry is illustrated in Figure 1. Two monopole sources of sound have strengths (volume accelerations) defined by the elements of the complex vector $\mathbf{v} = [v_1(\mathrm{j}\omega)\ \ v_2(\mathrm{j}\omega)]^{\mathrm{T}}$. These two sources produce acoustic pressure signals given by the elements of $\mathbf{w} = [w_1(\mathrm{j}\omega)\ \ w_2(\mathrm{j}\omega)]^{\mathrm{T}}$ at two spaced apart "ear" positions. This vector is given by $\mathbf{w} = \mathbf{C}\mathbf{v}$ where $\mathbf{C}$ is a matrix of frequency response functions that relates the source outputs to the acoustic pressures at the ears. The "desired" values of $\mathbf{w}$ are given by $\mathbf{d} = [d_1(\mathrm{j}\omega)\ \ d_2(\mathrm{j}\omega)]^{\mathrm{T}}$. The elements of $\mathbf{d}$ generally specify the signals that would be produced at the ear positions by a virtual source of strength proportional to $u(\mathrm{j}\omega)$ such that $\mathbf{d} = \mathbf{a}u(\mathrm{j}\omega)$, where $\mathbf{a}$ is a vector of frequency response functions. In this analysis, for the sake of simplicity, any effect of scattering by the head of the listener is omitted. Our studies have shown, however, that the essential physical mechanisms described here are not markedly changed by its inclusion.

The task is now to seek the linear filters given by the elements of $\mathbf{h} = [H_1(\mathrm{j}\omega)\ \ H_2(\mathrm{j}\omega)]^{\mathrm{T}}$ which operate on the virtual source signal $u(\mathrm{j}\omega)$ in order to give the signals $\mathbf{v}$ of the real sources that ensure that the signals at the ears are equal to those that would be produced by a virtual source. Thus one requires the condition $\mathbf{w} = \mathbf{d}$, which can be written as

$$\mathbf{w} = \mathbf{d} = \mathbf{C}\mathbf{v}. \tag{1}$$

The solution for the source strengths which produce the desired ear signals can thus be written as

$$\mathbf{v} = \mathbf{C}^{-1}\mathbf{d}, \tag{2}$$

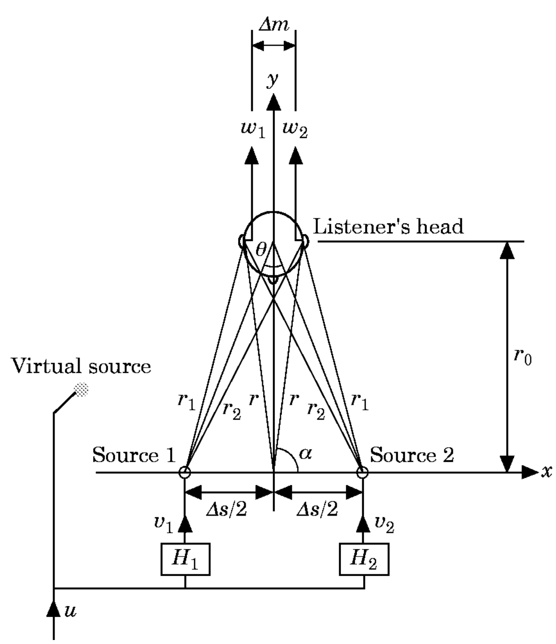

Figure 1. The geometry of the virtual source imaging system. The filters H_1 and H_2 are designed in order to ensure that the signals w_1 and w_2 produced at the listeners' ears are close approximations to those produced by a virtual source at a given spatial position.

and since $\mathbf{v} = \mathbf{h}u(\mathrm{j}\omega)$ and $\mathbf{d} = \mathbf{a}u(\mathrm{j}\omega)$ one can also write the solution for the optimal filters as

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{a}, \tag{3}$$

The matrix $\mathbf{C}$ is particularly simple if one assumes free field propagation and the geometry of Figure 1, in which case

$$\mathbf{C} = \frac{\rho_0}{4\pi}\begin{bmatrix} \mathrm{e}^{-\mathrm{j}kr_1}/r_1 & \mathrm{e}^{-\mathrm{j}kr_2}/r_2 \\ \mathrm{e}^{-\mathrm{j}kr_2}/r_2 & \mathrm{e}^{-\mathrm{j}kr_1}/r_1 \end{bmatrix} \tag{4}$$

where an $\mathrm{e}^{\mathrm{j}\omega t}$ time dependence is assumed with $k = \omega/c_0$, and where ρ_0 and c_0 are the density and sound speed. The inverse of this matrix is readily computed analytically.

The solution for the source strengths given by equation (2) is particularly illuminating when it is assumed that a given value of desired signal is required at one of the "ears", while a zero signal is required at the other ear. Consider the case $\mathbf{d} = [0 \;\; d_2(\mathrm{j}\omega)]^{\mathrm{T}}$: i.e., a signal $d_2(\mathrm{j}\omega)$ is desired at ear 2 while a zero signal is desired at ear 1. Furthermore, in order to ensure a causal solution for the source outputs one can assume that $d_2(\mathrm{j}\omega) = D(\mathrm{j}\omega)\rho_0\,\mathrm{e}^{-\mathrm{j}kr_1/4\pi r_1}$.This amounts to assuming that one wishes to simulate the signal produced at ear 2 by a source of strength $D(\mathrm{j}\omega)$ at a distance r_1 from ear 2. The source outputs given by equation (2) required to achieve this can be written as

$$v_1(\mathrm{j}\omega) = -g\,\mathrm{e}^{-\mathrm{j}\omega\tau}\frac{D(\mathrm{j}\omega)}{1 - g^2\,\mathrm{e}^{-2\mathrm{j}\omega\tau}}, \qquad v_2(\mathrm{j}\omega) = \frac{D(\mathrm{j}\omega)}{1 - g^2\,\mathrm{e}^{-2\mathrm{j}\omega\tau_1}}, \tag{5a, b}$$

where $g = r_1/r_2$ and $\tau = (r_1 - r_2)/c_0$. The denominator of these expressions can be written in series form by using the identity

$$(1 - x)^{-1} = \sum_{n=0}^{\infty} x^n \text{ (for } |x| < 1)$$

in order to show that

$$v_1(\mathrm{j}\omega) = -g\,\mathrm{e}^{-\mathrm{j}\omega\tau}D(\mathrm{j}\omega)\sum_{n=0}^{\infty} g^{2n}\,\mathrm{e}^{-2n\mathrm{j}\omega\tau}, \qquad v_2(\mathrm{j}\omega) = D(\mathrm{j}\omega)\sum_{n=0}^{\infty} g^{2n}\,\mathrm{e}^{-2n\mathrm{j}\omega\tau}. \tag{6a, b}$$

As recognized by Atal and Schroeder [3], the solution for the required source strengths is intrinsically recursive. In the time domain, the solution can be written in the form

$$v_1(t) = -gD(t - \tau) * [1 + g^2\delta(t - 2\tau) + g^4\delta(t - 4\tau) \ldots], \tag{7a}$$

$$v_2(t) = D(t) * [1 + g^2\delta(t - 2\tau) + g^4\delta(t - 4\tau) \ldots], \tag{7b}$$

where the asterisk denotes convolution. If, for example, $D(t)$ is a pulse the duration of which is short compared to the delay τ, source 2 first emits a pulse $D(t)$ that travels to ear 2 to give the desired signal $d_2(t)$. This pulse then arrives at ear 1 but is cancelled by the pulse $-gD(t - \tau)$ that has been emitted from source 1. This pulse however, causes an unwanted pulse at ear 2. This in turn is cancelled by the pulse $g^2D(t - 2\tau)$ emitted from source 2, and so on. This process is illustrated in Figure 2 which shows a sequence of

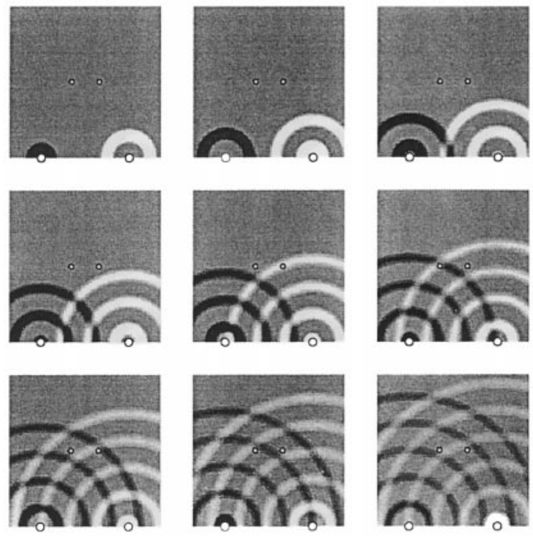

Figure 2. A series of illustrations of the sound field produced when two point monopole sources are used to generate a desired pulse at "ear 2" (on the right-hand side above) while producing zero pressure at "ear 1". Each illustration depicts the magnitude of the acoustic pressure on a greyscale, with lighter shading denoting positive pressures and darker shading denoting negative pressures. Values of 1 or greater are plotted as white and value of -1 or less are plotted as black. The right source emits the first pulse at $t = 0$ and the first illustration in the sequence is at $t = 0{\cdot}2/c_0$ with subsequent illustrations at intervals of $0{\cdot}1/c_0$ (with $c_0 = 344\ \mathrm{m\ s^{-1}}$). Each illustration is calculated at 101×101 points over an area measuring $1\ \mathrm{m} \times 1\ \mathrm{m}$. The separation of the "ears" $\Delta m = 0{\cdot}18$ m and the sources are spaced to give $\theta = 60^\circ$.

"snapshots" of the instantaneous pressure field produced by the two sources when the desired signal is a Hanning pulse specified by

$$d_2(t) = \begin{Bmatrix} 0 & t < 0,\, t > 2\pi/\omega_0 \\ (1 - \cos\omega_0 t)/2, & 0 \leqslant t \leqslant 2\pi/\omega_0 \end{Bmatrix}, \tag{8}$$

where ω_0 is $6{\cdot}400\pi$ (which thus implies that the first zero in the spectrum is at 6·4 kHz). A considerable reduction in the spatial complexity of the sound field generated in the region of the listener's head can be produced while still achieving the same objective. This is accomplished simply by placing the sources close together. In this case, the two sources can be used to synthesize a close approximation to the sound field that would be produced in the region of the listener's head by the superposition of a point monopole source and a point dipole source, these two equivalent point sources being placed at the mid-point between the two real sources. The strengths of these two equivalent sources are those necessary to produce the desired ear signals. This can be demonstrated by evaluating the monopole and dipole moments associated with the two sources in the limit $\omega\tau \to 0$, $g \to 1$.

The monopole moment is given by

$$v(\mathrm{j}\omega) = v_1(\mathrm{j}\omega) + v_2(\mathrm{j}\omega) = D(\mathrm{j}\omega)(1 - g\,\mathrm{e}^{-\mathrm{j}\omega\tau})/(1 - g^2\,\mathrm{e}^{-2\mathrm{j}\omega\tau}), \tag{9}$$

and since the denominator of this expression can be written as $(1 - g\,\mathrm{e}^{-\mathrm{j}\omega\tau})/(1 + g\,\mathrm{e}^{-\mathrm{j}\omega\tau})$, this expression becomes

$$v(\mathrm{j}\omega) = D(\mathrm{j}\omega)/(1 + g\,\mathrm{e}^{-\mathrm{j}\omega\tau}). \tag{10}$$

Using the series expansion $e^x = 1 + x + x^2/2 \ldots$ shows that in the limit $g \to 1$ and $\omega\tau \to 0$, one has

$$v(j\omega) = D(j\omega)/2. \tag{11}$$

Similarly, the dipole moment can be written as

$$\mathbf{f}(j\omega) = \frac{\rho_0 \mathbf{\Delta} s}{2}(v_2(j\omega) - v_1(j\omega)) = \frac{\rho_0 \mathbf{\Delta} s}{2} \frac{D(j\omega)(1 + g\, e^{-j\omega\tau})}{1 - g^2\, e^{-2j\omega\tau}}, \tag{12}$$

where $\mathbf{\Delta} s$ is the distance separating the two sources and $\mathbf{f}(j\omega)$ is the point force acting along the axis of the point sources that defines the strength of the equivalent point dipole. This expression can be written as

$$\mathbf{f}(j\omega) = (\rho_0 \mathbf{\Delta} s/2)\; D(j\omega)/(1 - g\, e^{-j\omega\tau}). \tag{13}$$

Again using a series expansion of the denominator, together with the approximation $|\mathbf{r}_2 - \mathbf{r}_1| = |\mathbf{\Delta} s| \cos\alpha$, where the angle α is defined in Figure 1, shows that in the limit $g \to 1$, $\omega\tau \to 0$,

$$\mathbf{f}(j\omega) \approx \frac{\rho_0 D(j\omega)}{2\cos\alpha}\left(\frac{1}{r} + \frac{j\omega}{c_0}\right)^{-1}. \tag{14}$$

In using this limiting case one assumes, of course, that the acoustic wavelength is much larger than the path length difference $(r_2 - r_1)$ between one of the sources and the ears.

Identical results are derived by assuming *a priori* that the two sources used to produce the same desired ear signals are a superposition of point monopole and dipole sources. In Figure 3 is shown the sound field produced when the requisite combination of point monopole and dipole sources is used to produce the same desired ear signals. Clearly, the sound field produced has a far less complex spatial behaviour than that produced by the widely spaced sources. The sources can acheive the desired objective with a single pulse emitted by the source combination.

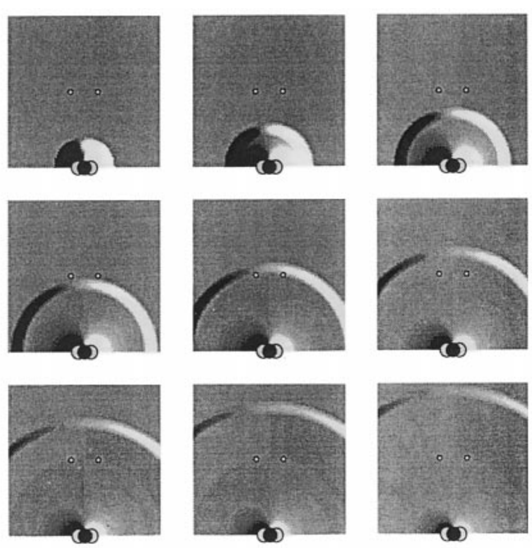

Figure 3. A series of illustrations that are equivalent to those shown in Figure 2, except that the desired field is produced by using the superposition of a point monopole and a point dipole.

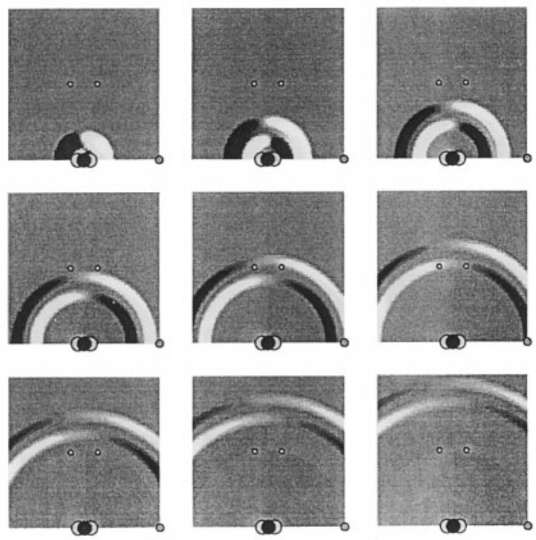

Figure 4. The sound field produced by the monopole/dipole combination when attempting to synthesize the signals at the "ears" that would be produced by the virtual source shown located at the bottom right of the illustrations.

3. THE PRODUCTION OF VIRTUAL IMAGES

In order to produce the illusion in a listener of the existence of a virtual source, the desired signals at the ears specified by the vector **d** must of course be those that would be produced by a real source placed at the intended virtual source position. The combination of a point monopole and a point dipole can in principle be used to produce these signals. The solutions for the source outputs can be deduced readily by using the above analysis. As an illustration, in Figure 4 is shown the sound field produced when the desired signal at ear 2 is again given by equation (8) while the desired signal at ear 1 is given by $g_v d_2(t - \tau_v)$. In this case, g_v and τ_v are the gain and time delay that account for the reduction in amplitude and the relative delay of the signal produced at ear 1 by the virtual source located at the position illustrated in Figure 4. The form of sound field produced now consists of a succession of two pulses that are emitted by the source combination. The first pulse produced the desired signal $d_2(t)$ at ear 2 and a zero signal at ear 1, while the second pulse produces the desired signal $d_1(t)$ at ear 1 and a zero signal at ear 2. Note that this form of sound field would be produced irrespective of the frequency content of the desired pulse; these idealized point sources could operate to produce such a sound field over an infinite bandwidth.

However, a real system operating with two individual sources will be capable of producing this form of sound field over only a limited frequency range. Nevertheless, we have found that this can produce remarkably good virtual source images. As an illustration, in Figure 5 is shown the sound field produced by two closely spaced monopole sources when used to synthesize the ear signals produced by a virtual source. The virtual source signal is again assumed to be a Hanning pulse which must arrive first at ear 2 and second at ear 1, the time delay and amplitude difference between these pulses replicating those that would be produced by the virtual source.

The results presented in Figure 5 clearly demonstrate the form of the sound field produced. It must be emplasized, however, that with a finite source separation such a sound field can be generated only over a limited bandwidth. Thus the time difference τ

will always remain finite and the recursive behaviour of the sources described by equations (7) is still present, except that the time duration between successive source output pulses is, in the example shown, very small compared to the duration of the pulse $D(t)$. This timescale is characterized by the frequencies at which the denominator of equations (5) becomes small, this being the case when the path length difference $r_2 - r_1$ is an integer number of half wavelengths. In the example shown in Figure 2, with the included angle $\theta = 60^\circ$, this frequency is 1·9 kHz, while for the closely spaced loudspeakers shown in Figure 5, with $\theta = 10^\circ$, the frequency becomes 10·8 kHz. The high frequency "ripples" in the sound field shown in Figure 5 correspond to this frequency. It is also worth emphasizing that the sound field produced by a pair of sources with finite separation is only a good approximation to the monopole/dipole field when $|\mathbf{r}_2 - \mathbf{r}_1|w|\boldsymbol{\Delta}s \cos\alpha|$ is small compared to the wavelength. Thus the idealized monopole/dipole field at higher frequencies is approximated only in the region of the listeners ears and not over the entire sound field. A comparison of the directivity patterns of the fields produced by the monopole/dipole and that produced by two sources shows that the two fields remain very closely matched in the region of the listeners ears up to the frequency at which $r_2 - r_1$ is equal to about one quarter of the acoustic wavelength.

4. PRACTICAL IMPLEMENTATION

In order to implement such a system effectively, one has to design digital realizations of the filters $H_1(j\omega)$ and $H_2(j\omega)$. It is also important in practice to account for the scattering of the listener's head. This can be achieved by ensuring that the filters are designed in order to ensure a good approximation to the desired signals at the ears of a dummy head. The transfer function matrix $\mathbf{C}$ and the vector $\mathbf{a}$ are thus of considerably greater complexity than in the case of free field propagation assumed above. Furthermore, it may also be necessary to account for the frequency response functions of the loudspeakers used, although it has been found in practice that provided the frequency response functions are

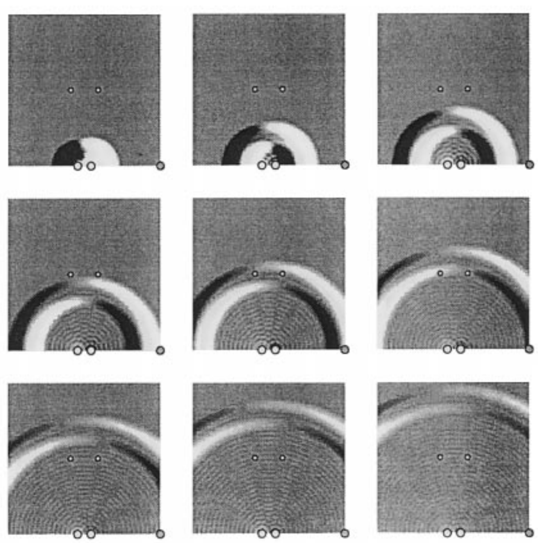

Figure 5. The sound field produced by two closely spaced monopole sources ($\theta = 10^\circ$) when attempting to synthesize the signals at the "ears" that would be produced by the virtual source shown located at the bottom right of the illustrations. Note that the field shown in Figure 4 of the monopole/dipole combination is well approximated, except for the high frequency "ripples" in the field.

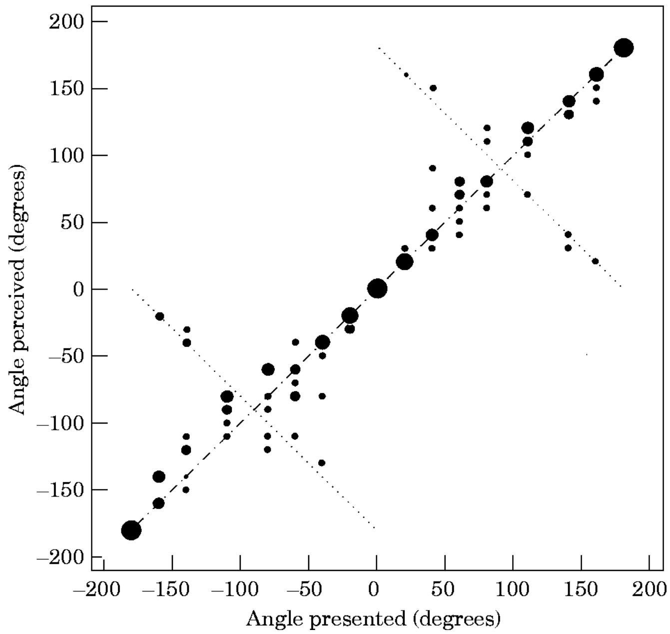

Figure 6. Results of subjective experiments with the virtual source imaging system using closely spaced loudspeakers ($\theta = 10^\circ$) for the seven subjects for whom the system worked well.

well matched (typically to within 0·5 dB in amplitude and a few degrees in phase) then this can be neglected in the filter design process. In any event, the filters have to be designed with considerable care, and a number of useful methods have been presented previously [7–19]. These methods generally make use of least squares techniques wherein one finds a stable causal realization of the filters that minimize the sum of the squared deviations (in time or frequency domains) between the desired signals and the reproduced signals at the ears of a dummy head.

The results of some preliminary subjective experiments with such a system are presented in Figures 6–9. In all these figures, the results are shown of the "angle perceived" by the subjects plotted against the "angle presented" by the virtual source imaging system. The number of subjects reporting a given angle perceived is indicated by the area of the circles in the figures. These experiments were undertaken by first designing the filters using the techniques described in reference [18] and by using a database [20] of head related transfer functions for the KEMAR dummy head in order to specify the elements of the matrix **C** and the vector **a**. The filters used were FIR filters each having 1600 coefficients, each pair of filters being designed to produce a required position of virtual source image. These filters were implemented on special purpose signal processing equipment at a sample rate 44·1 kHz. More details of the equipment used is specified in references [16, 17]. Eleven male subjects, all with normal hearing, were then seated in front of a pair of Celestion 1 loudspeakers under anechoic conditions, the listeners being surrounded by acoustically transparent black cloth behind which the loudspeakers were placed. The speakers were arranged to subtend an angle θ of either 10° or 60° at the ears of the subjects. The listeners were seated such that their ears and the centers of the loudspeakers were in the same horizontal plane. They were asked to look straight ahead, judge the apparent location of the source and then report this location by using a series of markers which were placed at 10° intervals in the horizontal plane. No attempt was made to generate virtual images outside this plane. The virtual source signals used in the case shown was

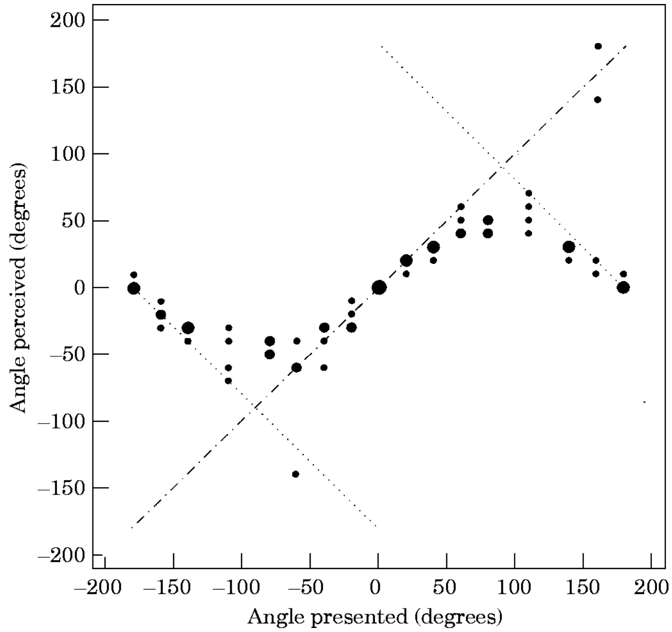

Figure 7. Results of subjective experiments with the virtual source imaging system using closely spaced loudspeakers ($\theta = 10^{\circ}$) for the four subjects for whom the system worked poorly.

filtered white noise according to EAIJ RC-7603 which has a spectrum that is flat between 200 Hz and 2 kHz and which rolls off at about 3 dB per octave at both low and high frequencies.

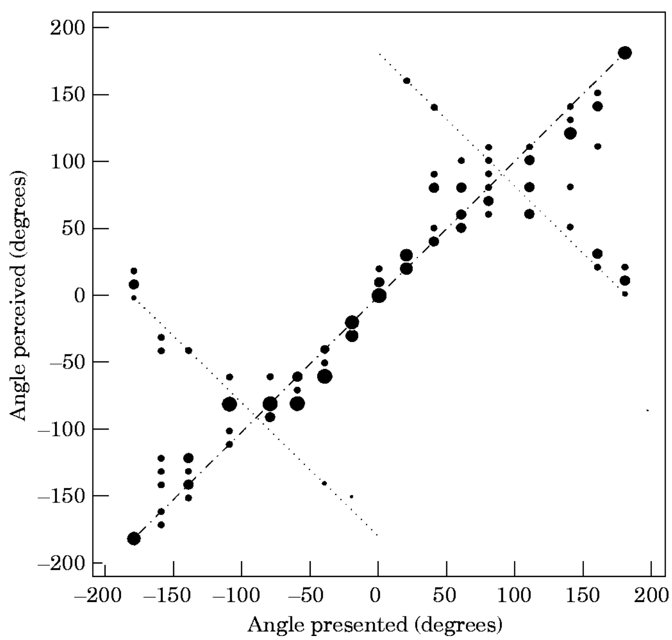

Figure 8. Results of subjective experiments with the virtual source imaging system using widely spaced loudspeakers ($\theta = 60^{\circ}$) when the listener's head was displaced laterally by 5 cm from the optimum listening position. Results are shown for only those seven subjects for whom the system worked well.

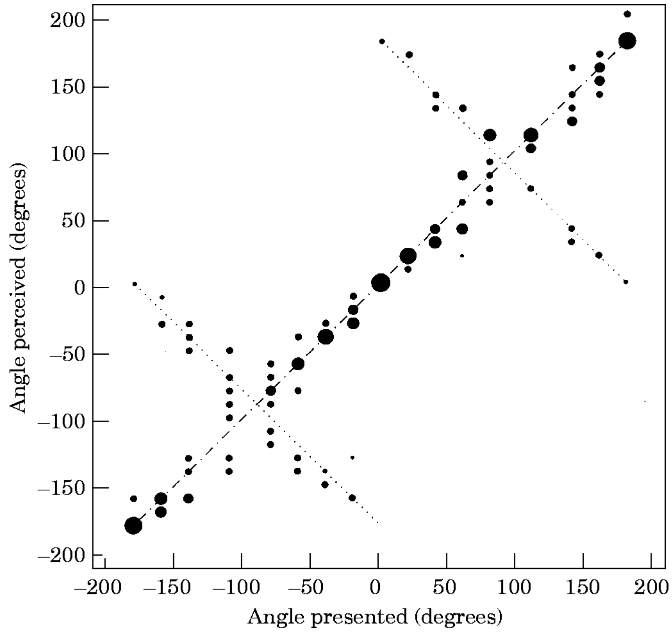

Figure 9. Results of subjective experiments with the virtual source imaging system using closely spaced loudspeakers ($\theta = 10^\circ$) when the listener's head was displaced laterally by 5 cm from the optimum listening position. Results are shown for only those seven subjects for whom the system worked well.

An interesting feature of the results is that both the systems ($\theta = 10^\circ$ and 60°) worked extremely well for seven of the eleven subjects whilst the system was not so effective for four of the eleven subjects. The results for the two classes of subjects are shown only for the $\theta = 10^\circ$ system in Figures 6 and 7. The difference between the two sets of results is that the group of four subjects did not localize images presented in the rear half of the horizontal plane, but tended to perceive their location to be at "mirror image" positions in the front half of the horizontal plane. The results for the $\theta = 60^\circ$ system were not dissimilar. Preliminary experiments have also been undertaken in order to establish the robustness of image location in respect to the movement of the listener's head. The results are shown in Figures 8 and 9, in which are compared the performance of the $\theta = 10^\circ$ and $\theta = 60^\circ$ systems when the listeners' heads were displaced by 5 cm to the right of the optimal listening position. The results presented are for the group of seven subjects for whom both systems worked best. These preliminary results suggest that the imaging capabilities of the system with more closely spaced loudspeakers are more robust in respect to head movement than that of the system with widely spaced loudspeakers. The extent to which the frequency content of the virtual image signal and the form of the dummy head HRTF each affect the performance of the system is a matter of current investigation. The simplification in the spatial complexity of the field suggests that the imaging system should be more tolerant to head movement, and these preliminary results suggest that this is indeed the case.

5. CONCLUSION

The intention of this brief report is to introduce a method for generating virtual images that makes use of a form of sound field whose spatial complexity is less than that produced by existing techniques. The method can be implemented in practice extremely effectively and preliminary results suggest that it may have many practical benefits, including an increased robustness in respect to head movement and, of course, the ability to site both transducers used for reproduction within one compact cabinet.

REFERENCES

1. A. D. Blumlein 1931 *British Patent No*. 394325. Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems.
2. J. Blauert 1983 *Spatial Hearing*. Cambridge, Massachusetts: MIT Press.
3. B. S. Atal and M. R. Schroeder 1962 *U.S. Patent* 3,236,949. Apparent sound source translator.
4. H. Hamada, N. Ikeshoji, Y. Ogura and T. Miura 1985 *Journal of the Acoustical Society of Japan* (*E*) **6**, 143–154. Relation between physical characteristics of orthostereophonic system and horizontal plane localisation.
5. G. Neu, E. Mommertz and A. Schmitz 1992 *Acustica* **76**, 183–192. Investigations of true directional sound reproduction by playing head referred recordings over two loudspeakers: part I
6. G. Urbach, E. Mommertz and A. Schmitz 1992 *Acustica* **77**, 153–161. Investigations on the directional scattering of sound reflections from the playback of head referred recordings over two loudspeakers: part II
7. P. A. Nelson, H. Hamada and S. J. Elliott 1992 *IEEE Transactions on Acoustics Speech and Signal Processing* **40**, 1621–1632. Adaptive inverse filters for stereophonic sound reproduction.
8. J. Bauck and D. H. Cooper 1996 *Journal of the Audio Engineering Society* **44**, 683–705. Generalized transaural stereo and applications.
9. Fr. Heegaard 1958 *EBU Rev. pt A—Technical*, *No*. 52, 2–6. (Reprinted in *Journal of the Audio Engineering Society* **40**, 692–705, 1992).
10. O. Kirkeby, P. A. Nelson and H. Hamada 1996 *British Patent Application No*. 9603236.2. Sound recording and reproduction systems (short title: Stereo Dipole).
11. H. Hamada, H. Tokunoh, O. Kirkeby and P. A. Nelson 1996 *IEICE Technical Report EA*96-2. A study of a new sound reproduction system (Stereo Dipole).
12. H. Tokuno, Y. Watanabe, H. Hamada, O. Kirkeby and P. A. Nelson 1996 *Proceedings of the Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*. Binaural reproduction in the stereo dipole.
13. Y. Watanabe, H. Tokuno, H. Hamada, O. Kirkeby and P. A. Nelson 1996 *Proceedings of the Third Joint Meeting of the Acoustical Society of America and the Acoustical Society of Japan*, 307–310. Subjective investigation of new sound reproduction system (Stereo Dipole).
14. O. Kirkeby, P. A. Nelson and H. Hamada 1997 Audio Engineering Society Preprint 4463 (16). The "Stereo Dipole" Binaural sound reproduction using two closely spaced loudspeakers. (Presented at the 102*nd AES Convention*, 22–25, *March*, *Munich*, *Germany*.)
15. T. Taleuchi, P. A. Nelson, O. Kirkeby and H. Hamada 1997 *Audio Engineering Society Preprint* 4464 (17). Robustness of the Performance of the "Stereo Dipole" to misalignment of head position. (Presented at the 102nd AES Convention, March 22–25, Munich, Germany.)
16. P. A. Nelson, F. Orduna-Bustamante and H. Hamada 1996 *Journal of the Audio Engineering Society* **44**, 973–989. Multi-channel signal processing techniques in the reproduction of sound.
17. P. A. Nelson, F. Orduna-Bustamante, D. Engler and H. Hamada 1996 *Journal of the Audio Engineering Society* **44**, 990–1007. Experiments on a system for the synthesis of virtual acoustic sources.
18. O. Kirkeby, P. A. Nelson, H. Hamada and F. Orduna-Bustamante 1996 *ISVR Technical Report* 255 Fast deconvolution of multi-channel systems using regularisation.
19. O. Kirkeby, P. A. Nelson, H. Hamada and F. Orduna-Bustamante 1996 *Journal of the Acoustical Society of America* **100**, 1584–1593. Local sound field reproduction using digital signal processing.
20. B. Gardner and K. Martin 1994 *MIT Media Lab Technical Report* 280. HRTF measurements of a KEMAR dummy-head microphone.