Head tracking for local active sound control

S J Elliott, M Simón-Gálvez, J Cheer and W Jung
Institute of Sound and Vibration Research, University of Southampton, Southampton, UK

Paper Number: O11000135

Abstract

The spatial extent of the cancellation that can be achieved with a local active sound control system is limited, particularly at frequencies above about 300 Hz, so that control of the pressure at fixed points in space does not give satisfactory performance as the listener's head moves about, unless the position of the head can be tracked. The availability of low-cost head tracking systems for gaming applications opens up the possibility of such a head tracking active control system at a modest cost. In the experiments reported here, such a system was used to track the head position in a local active sound controller, implemented in a headrest. The positions of the listener's ears were calculated from the head position. This was used in lookup tables to estimate both the responses from the secondary sources to the remote microphone positions at the ear, and also the weightings on an array of monitoring microphones mounted on the headrest, used to estimate the pressures at these remote microphone positions, using an observer. These estimates were then used to adapt the two control filters driving the secondary sources.

1. Introduction

Active sound control is subject to fundamental limits on its performance due to the physics of acoustic cancellation [1]. Global active control can only be achieved in an enclosure if there are not too many acoustic modes contributing to the response, which in enclosures such as those inside cars limits its upper frequency range to below about 300 Hz. The performance of local active control systems depends more on the detailed arrangement of secondary source and error microphone, but a useful rule of thumb is that the zone of quiet, within which the sound can be reduced by 10dB, has a diameter of no more than about 1/10 of an acoustic wavelength [2]. The maximum diameter of this zone of quiet is thus about 10 cm at 300 Hz but only about 3.5 cm at 1 kHz.

With a fixed arrangement of secondary sources and error microphones, as in the active headrest arrangement shown in figure 1, a local active control system will work well provided the listener's ears are within such a zone of quiet. In practice, the listeners head will typically move around, however, and for head movements of about 5 cm, this local system will thus only be effective up to about 300 Hz. This upper frequency limit could be improved if the error microphone moved around with the head location, as it does inside an active headphone, but the use of headphones is often undesirable for comfort and safety reasons. It is possible, with a fixed array of microphones, to estimate the pressure at some "virtual" microphone positions remote from this array [3,4], and thus avoid the need for physical microphones close to the listener's ears. The performance of such a system at different frequencies is illustrated in figure 2 [5]. The position of these virtual microphones must be known for such a system to be effective, however, and so such techniques are also limited by the movement of the listener's head, unless the head position can be tracked [4].

In this paper we will initially discuss two separate problems associated with implementing a local active sound control system when the listener's head is moving: the tracking of the plant response, which is required to maintain stability, and the tracking of the

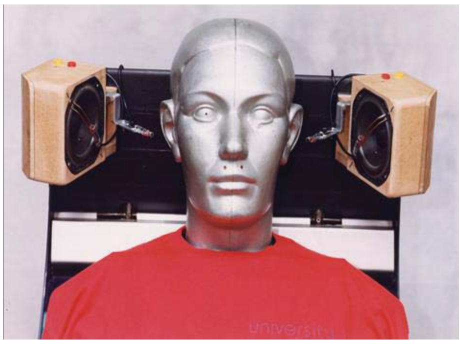

Figure 1. Physical arrangement of local active sound control in a headrest

ISBN: 978-981-09-7961-4 — doi: 10.3850/978-981-09-7961-4_O11000135

error signal, which is required to maintain performance. Both of these problems will be considered in the frequency domain, which allows a straightforward analysis of their performance limits. An experimental system will then be described in which a Kinect system is used to track the head position in a local active headrest control system, initially to schedule the change in plant response and maintain stability.

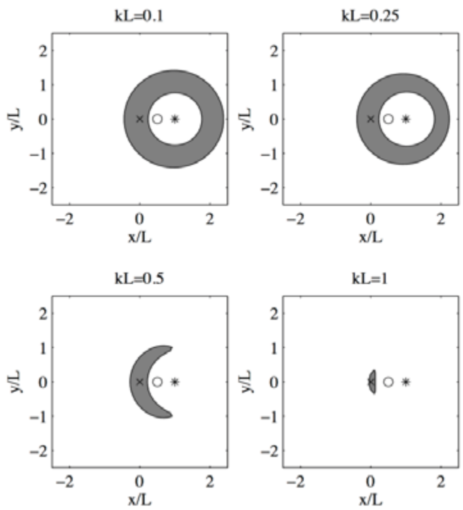

*Figure 2. The shaded regions show the zone of quiet within which the average pressure is reduced by 10 dB from simulations of cancellation of a diffuse field using the virtual microphone method at a single virtual error microphone position, ×, using a single monitoring microphone, ○, and a single monopole secondary source, *, at four different normalized frequencies.*

2. Tracking of plant response

We initially assume that a local active sound control system is implemented having two loudspeakers and two microphones, at the ears of the listener, as shown in figure 3. If the controller is operated at a single frequency, the vector of error signals, $\boldsymbol{e}(n)$ can be written, at the n-th iteration, as

$$\boldsymbol{e}(n) = \boldsymbol{d}(n) + \boldsymbol{G}\ \boldsymbol{u}(n)\,, \qquad (1)$$

where $\boldsymbol{d}(n)$ and $\boldsymbol{u}(n)$ are the disturbance and control signals at the n-th iteration and $\boldsymbol{G}$ is the

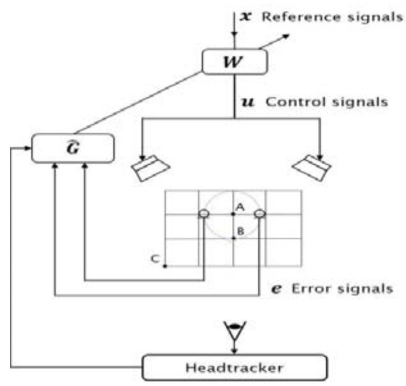

Figure 3. The use of a tracker to determine the position of the head in a local active sound control system and hence update the estimate of the plant response, $\hat{\boldsymbol{G}}$, between the loudspeakers and microphones.

matrix of physical plant responses between the loudspeakers and the microphones. The control signals can be adapted using a leaky steepest-descent algorithm, which can be written as

$$\boldsymbol{u}(n+1) = (1-\alpha\beta)\boldsymbol{u}(n) - \alpha\hat{\boldsymbol{G}}^{H}\boldsymbol{e}(n)\,, \qquad (2)$$

where α is the convergence coefficient, β is a leakage factor that improves robustness, and $\hat{\boldsymbol{G}}^{H}$ is the Hermitian transpose of the estimated plant response matrix. For slow adaptation, it is known that the adaptive algorithm is stable provided

$$\mathrm{Re}\left(eig[\hat{\boldsymbol{G}}^{H}\boldsymbol{G} + \beta\boldsymbol{I}]\right) > 0\,, \qquad (3)$$

where $\mathrm{Re}\left(eig[\ ]\right)$ denotes the real parts of the eigenvalues of the matrix in square brackets. The center of the listener's head in figure 3 is assumed to be located on a grid of 5 x 4 points spaced 5 cm apart. The nominal location of the head, is shown as "A" in figure 3, together with two other locations shown as "B", which is about 5 cm away, and "C", which is about 14 cm away. The 2 x 2 plant matrices have been measured for the physical arrangement shown in figure 1 with the dummy head at all of these locations. It is assumed that the estimated plant response is $\hat{\boldsymbol{G}}$, but when the head is moved the physical plant response is $\boldsymbol{G}$. Figure 4 shows the real part of the eigenvalues in equation 3, with no leakage, plotted at different excitation frequencies, for the head in the physical positions, "B" and "C", after the estimated plant response has been measured at position "A".

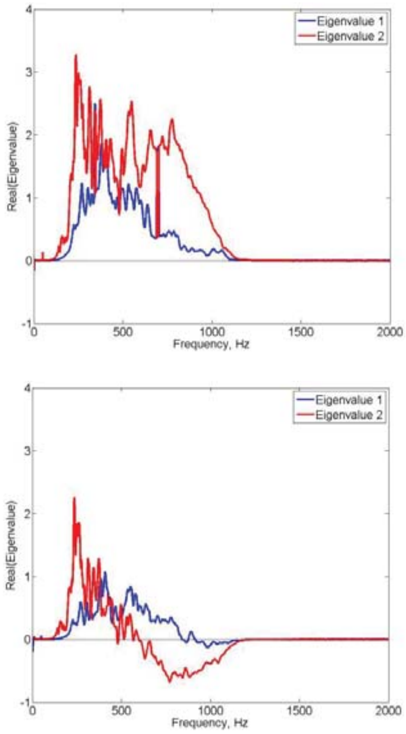

Figure 4. Real parts of the eigenvalues of the matrix, $\hat{G}^H G$ when $\hat{G}$ is measured with the head in position "A", but G is measured in either position "B", 5cm from position "A" (top) or position "C", 14 cm from position "A" (bottom). The real part of the eigenvalue becomes negative in the latter case above about 600 Hz, indicating that the system will be unstable.

Although the real parts of both eigenvalues remain positive up to 1 kHz at position "B", it is clear that the real part of one of the eigenvalues becomes negative above 600 Hz when the head is in position "C", indicating that an untracked system will be unstable above this frequency. Head tracking is thus required to adapt $\hat{G}$, depending on the head position, and thus maintain stability.

3. Tracking of error signals

It is often not possible to position the error microphones at the ears of the listener, as above. In this case the physical error signals, at the listener's ears, can be estimated from the outputs of a stationary array of monitoring microphones, using a virtual or remote microphone arrangement [3,4], as shown in figure 5. The estimates of the error signals, $\hat{e}$, at the remote microphones can be calculated using the outputs of the monitoring microphones, m, and the control signals, u, as

$$\hat{e} = \hat{G}_e u + \hat{O}\left[m - \hat{G}_m u \right], \qquad (4)$$

where $\hat{G}_e$ and $\hat{G}_m$ are estimates of the plant responses from the control signals to either the remote microphone signals or the monitoring microphones signals, and $\hat{O}$ is an observation matrix, used to estimate the disturbance signals at the remote microphones from those at the monitoring microphones [4,5]. Figure 6 shows the block diagram of a system in which the head is tracked and used to update a remote microphone system with the appropriate values of $\hat{G}_e$, $\hat{G}_m$ and $\hat{O}$ at the different head positions.

The form of this observation matrix, if it is designed to minimise the mean squared difference between the estimated disturbance at the remote microphones and the true disturbance there, can be written as

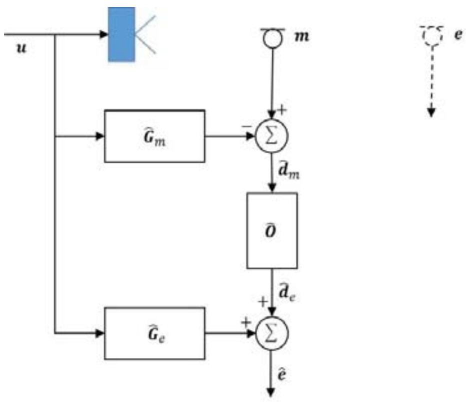

Figure 5 The remote microphone arrangement, in which the error signals at a remote microphone location, e, are estimated from the signals at an array of monitoring microphones, m, by subtracting an estimate of the control signals at the monitoring microphones, via the estimated monitoring plant response, $\hat{G}_m$ to estimate the disturbance signals at these microphones, $\hat{d}_m$, using an observation matrix, $\hat{O}$, to estimate the disturbance at the remote microphones, $\hat{d}_e$ and then adding back the component of the control signals at these microphones, via $\hat{G}_e$.

$$\boldsymbol{O}_{opt} = \boldsymbol{S}_{me}\left[\boldsymbol{S}_{mm} + \beta_0 \boldsymbol{I}\right]^{-1}, \qquad (5)$$

where the spectral density matrices are defined such that $\boldsymbol{S}_{me} = E[\boldsymbol{e}\,\boldsymbol{m}^H]$, $\boldsymbol{S}_{mm} = E[\boldsymbol{m}\,\boldsymbol{m}^H]$, E is the expectation operator calculated over all the positions of the primary sources, which can be calculated from the measured responses from an array of assumed primary sources [5]. $\boldsymbol{O}_{opt}$ can be termed the spatial Wiener filter. A regularisation factor, β_0, is introduced to reflect the uncertainty in $\boldsymbol{S}_{mm}$ [6]. A similar formulation for this observation matrix has been used by Moreau et al [7] to analyse the form of this matrix for an idealised, diffuse, primary field. In this case, the elements of $\boldsymbol{O}_{opt}$ are entirely real and can be calculated from the theoretical form of the spatial correlation functions in a diffuse field.

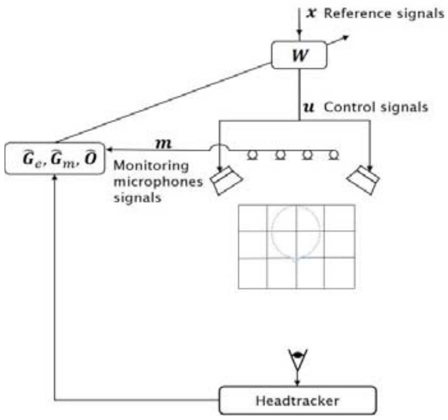

Figure 6. The use of a tracker to determine the position of the head in a local active control system and hence update the parameters required to estimate the error signals at the listener's ears from those at an array of static monitoring microphones using the remote microphone method.

4. Experimental results

A Microsoft Kinect system has been used in a small listening room to track, in the first instance, the position of a dummy head, as shown in figure 7. For the Kinect system to recognise the head, it was necessary to fix a mask representing a face to the dummy head.

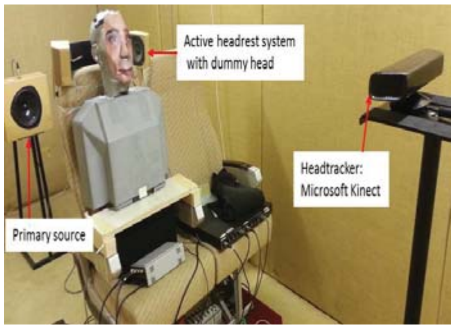

Figure 7. Installation of the active headrest with the head tracker.

Figure 8 shows the time history of an error signal, measured in one ear of in the dummy head, during the course of a tracking experiment. Initially, with the active noise control, ANC, and the head tracker off, the error signal represents the disturbance from the primary source, which is a loudspeaker behind a headrest generating a single frequency disturbance at 650 Hz. With the dummy head in a fixed position, the ANC is then turned on and control of the error signal down to the level of the background noise is observed. The dummy head is then moved with the head tracking off, until the control systems becomes unstable, approximately when it reaches position "C" in figure 3, as predicted in section 2. The head tracker is then turned on and rapidly acquires the current head position, corrects the estimate of the plant response, and brings the error under control again.

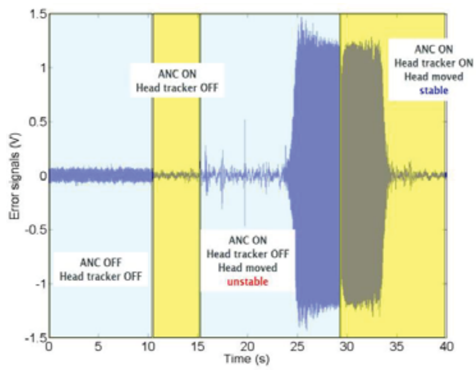

Figure 8 The time history of the error signal during the course of a tracking experiment.

5. Conclusions

This paper describes the use of head tracking to improve the performance of a local active noise control system. In practice a headrest system is used, with two secondary loudspeakers and, initially, two microphones in the ears of a dummy head. Different situations are analysed where head tracking is required. In the first, tracking of the plant responses is required to maintain the systems stability. Predictions are made of the distance that the dummy head can be moved before the control system becomes unstable, and these agree reasonably well with what is observed in experiments. In the experiments a Kinect system is employed to track the position of the dummy head and this is used to update the estimates of the plant response used in the control system, thus bringing it back into stability. The second situation analysed is where the error signals at the ears of a listener are estimated from a fixed array of monitoring microphones, using a remote microphone method. The practical implementation of this system is the subject of current work.

Acknowledgments

The work of Marcos Simón Gálvez is supported by the EPSRC under the S3A Programme Grant, EP/L0005391/1, and the Ph.D. work of Woomin Jung is supported by Jaguar Land Rover.

References

[1] P. A. Nelson, and S.J. Elliott "Active control of sound.", Academic Press, London, (1992).
[2] S. J. Elliott, P. Joseph, A. J. Bullmore and P. A. Nelson "Active cancellation at a point in a pure tone diffuse sound field.", Journal of sound and vibration, **120**, 183-189 (1988).
[3] S. J. Elliott, and A. David. "A virtual microphone arrangement for local active sound control." Proceedings of the 1st International Conference on Motion and Vibration Control, Yokohama, Japan, 1027-1031, (1992).
[4] D. Moreau, B. Cazzolato, B. Zander and C. Petersen "A review of virtual sensing algorithms for active noise control.", Algorithms, **1**, 69-99 (2008).
[5] S. J. Elliott, and J. Cheer. "Modelling local active sound control with remote sensors in spatially random pressure fields.", The Journal of the Acoustical Society of America, **137**, 1936-1946 (2015).
[6] S. J. Elliott, J. Cheer, J. W. Choi, and Y. Kim "Robustness and regularization of personal audio systems.", Audio, Speech, and Language Processing, IEEE Transactions on, **20**, 2123-2133 (2012).
[7] D. Moreau, J. Ghan, B. S. Cazzolato, and A. C. "Active noise control in a pure tone diffuse sound field using virtual sensing.", The Journal of the Acoustical Society of America, **125**, 3742-3755 (2009).