다수의 청취자를 위한 가상음향 시스템의 크로스토크 제거

김영태*, O. Deille**, P.A. Nelson**

*: 삼성종합기술원 (yt65.kim@samsumg.com), **: ISVR, University of Southampton, U.K.

Crosstalk cancellation of virtual acoustic imaging systems for multiple listeners

Youngtae Kim*, O. Deille**, P.A. Nelson**

*: Samsung Advanced Institute of Technology, **: ISVR, University of Southampton, U.K.

Abstract

The central aspect of the research into virtual acoustic imaging for multiple listeners involves a detailed study of the relative orientation of both sources and listeners. It is vital in any multiple listener system to first establish the conditioning of the potential geometrical arrangements of transducers and listeners. In this paper, this made clear the important link between the conditioning of the electro-acoustic transfer function matrix between transducers and ears and the design of inverse filters for crosstalk cancellation. This work has been undertaken by using simple free field models of the electro-acoustic transfer functions. Therefore, optimal transducer arrangements have been identified and these have been proved by time domain solutions.

1. Introduction

Perfect crosstalk cancellation of virtual acoustic imaging systems requires that a signal is desired at one ear of each listener and silence is sought at the other ear. The technical developments undertaken to date for a single listener, for example 2 source-2 receiver systems are described in full in references [1,2]. Meanwhile, in many practical applications, virtual acoustic imaging systems are required to design for multiple users. In such cases, the systems have to be able to create virtual environments for multiple listeners at the same time. Therefore, it is clearly necessary that researches be initiated to examine the developed techniques for the delivery of high quality virtual acoustic images to multiple listeners. In this paper, a virtual imaging system having four sources and four receivers (or equivalently two listeners) is focused. However, since this is just the start of the observation of the dependence of the condition number on geometrical factors for multiple listeners, it will be more helpful to keep matters simple in the first instance

2. Crosstalk cancellation with a free field model for multiple listeners

For the crosstalk cancellation, inverse filters are often deduced by using the inversion of the plant matrix [1] which is defined as acoustic transmission paths between the loudspeakers and listener's ears. In dealing with this kind of the problem, the condition number has been introduced to represent the control performance and robustness of the system [2]. The condition number of the plant matrix **C**, $\kappa(\mathbf{C})$, is simply defined by $\kappa(\mathbf{C}) = \|\mathbf{C}\| \|\mathbf{C}^{-}\|$ (where $\| \|$ denotes the two-norm of the matrix)[3]. For example, in the 4 source-4 receiver system depicted in Fig.1, the receivers (w_1, w_2, w_3 and w_4) are assumed to locate in a horizontal plane. In the same plane, four monopole sources (v_1, v_2, v_3 and v_4) are aligned in front of the listeners. For this case, the condition numbers in dB ($=10\log_{10}\kappa(\mathbf{C})$) as a function of frequency are shown in Fig.2.

In order to perform time domain simulations with the geometry depicted in Fig.1, assume that each source signals are emitted from a Gaussian pulse specified by $d_{\text{pulse}}(t) = e^{-\pi(at)^2}\cos(\omega_o t)$. Here a is an arbitrary constant and $\omega_o (= 2\pi f_o)$ is the angular frequency. The source inputs are designed to achieve crosstalk cancellation at the listeners' right ears. The sampling frequency is 20480 Hz. The total sound field is represented by nine frames which are listed sequentially in a reading sequence from top left to bottom right. The time increment between each frame is 1.22×10^{-3} second. In each frame, the sound field is calculated at 131-by-131 discrete points over the whole area. For each point, the transfer function between each source and the position is calculated, and then convolved with the source signal to obtain the acoustic pressure at the position as a function of time. In the frames shown in Fig. 3 & 4, the results show the sound wave interference pattern in the same horizontal plane of the sources and the receivers, and their values greater than 1 are plotted as white, values smaller than -1 as black. Firstly in Fig.3, the solution with the pulse in the well-conditioned frequency (f_o=1275 Hz, see Fig.2) shows better result compared with that illustrated in Fig.4 for the pulse in the ill-conditioned frequency (f_o=3740 Hz, see Fig.2). As shown in Fig.3, the emitted signals from the pulse with f_o =1275 Hz vanish rapidly before reaching receivers. However, in the case presented in Figure 4, 'ring phenomenon' can be found, and thus the listeners must not move at all in order to avoid any disturbance.

3. Conclusions

It has shown that, in achieving crosstalk cancellation, the control performance of virtual acoustic imaging systems is crucially determined by the conditioning of the plant matrix. These examples investigated have also shown promise of designing a specific arrangement of sources and receivers for different frequency bands.

References

1. P.A. Nelson, O. Kirkeby, T. Takeuchi, H. Hamada, Sound fields for the production of virtual acoustic images, *Journal of Sound and Vibration*, 204, 386-396, 1997

2. T. Takeuchi, P.A. Nelson, Optimal source distribution for binaural synthesis over loudspeakers, *Journal of the Acoustical Society of America*, 112(6), 2786-2797, 2002.

3. Y. Kim, Spatial resolution limits for reconstruction of acoustic source distribution by inverse techniques, PhD Thesis, University of Southampton, 2002

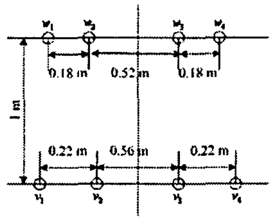

Fig.1 An example of a configuration for the 4 source-4 receiver system

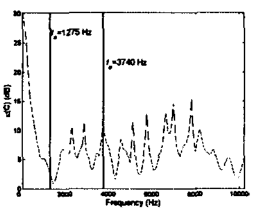

Fig.2 Condition numbers (in dB) of the plant matrix for the configuration of the four aligned sources considered here

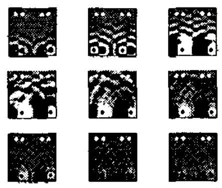

Fig.3 Sound field reproduced by the aligned sources at the well-conditioned frequency band (f_o =1275 Hz).

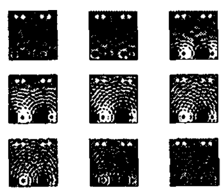

Fig.4 Sound field reproduced by the aligned sources at the well-conditioned frequency band (f_o =3740 Hz).