THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA VOL.25, NO.1E 2006. 3 pp. 8~13

Optimized Ambisonic Panning Algorithm Using Directional Psychoacoustic Criteria

Sin-Lyul Lee*, Seung-Rae Lee*, Koeng-Mo Sung*
*School of Electrical Engineering and Computer Science, Seoul National University
(Received January 13 2006; revised March 8; accepted March 22 2006)

Abstract

In this paper, an Optimized Ambisonic Panning Algorithm (OAPA) which reduces sound localization error, is proposed. In the conventional Ambisonic Panning Algorithm (APA), sound localization is usually different from the panning angle, especially when listeners are not in an ideal listening position, because of low signal separation among other channels. To overcome this problem, an OAPA using window functions is proposed. A proper window function can be verified, comprising of higher harmonic components than 2M+1 and improved DPC and channel separation. Analysis results demonstrate that the proposed method results in higher signal separation among other channels and lower sound localization errors than the conventional APA.

Keywords: *Ambisonic, Panning, Optimized Ambisonic Panning Algorithm, Directional Psychoacoustic Criteria*

I. Introduction

In audio systems, the use of a parametric potentiometer (pan-pot) is essential for sound recording and reproduction. Recently, a study on panning laws for multi channel sound recording and reproduction has made active progress because home theater systems using DVD is provided in homes for receiving HDTV signals[1].

Nowadays, the most frequently used multi channel panning algorithm is the multi channel extended pair wise constant power panning algorithm, which is widely used in stereo panning law and has easy implementation and acceptable sound quality. As a consequence, in a pair wise constant power panning algorithm, two effects occur, that is, a detent effect (the perceived sound image is not coincident with the intended panning angle and has a tendency to be pulled closer to the nearest loudspeaker) and an elevation effect (the perceived location of the phantom sound image is elevated above the horizontal plane due to the difference of HRTFs between real and phantom sound images) [2-3]. The APA is a virtual sound localization method, where the wave fed from a real sound source coincides with the sum of an incoming wave fed from the loudspeakers to a listener located an equal distance from all loudspeakers. This method is a kind of global panning algorithm, where all loudspeakers feed signals wherever the position of the virtual sound source, is widely applied to the multi channel panning algorithm and three dimensional sound reproduction systems[4]. For non-sweet spot listeners (listener is not at an equal distance from all loudspeakers), the APA, however, has a serious problem where a phantom sound image is inclined to the nearest loudspeaker due to the Hass effect[5]. In order to solve this problem, a new panning algorithm is devised, combining the APA reducing sound localization error, with the local panning algorithm that extends the sweet spot and introduces a Directional Psychoacoustic Criteria (DPC), which is devised by M. A. Gerzon, to assess the panning algorithm objectively[6]. Using DPC, the conventional APA and proposed OAPA were analyzed, combining the merit of global panning that has better DPC and local panning that has a better signal separation among other channels to obtain a stable phantom sound image for a non-sweet spot listener.

Corresponding author: Seung-Rae Lee(srlee@acoustics.snu.ac.kr)
School of Electrical Engineering and Computer Science, Seoul National University, Shillim dong Kwanak Ku, Seoul 151 742, Korea

II. Analysis of conventional ambisonic panning algotithm

2.1. Directional Psychoacoustic Criteria

Directional Psychoacoustic Criteria (DPC) describes sound localization in terms of a hierarchy of some relatively primitive theories, as one approximates a complex function into the sum of simple polynomials in applied mathematics. The purpose of these criteria is to describe the physical requirements using primitive theories, for accurate localization[6]. The most important primitive theories comprising the lowest components in the hierarchy are the velocity vector direction (θ_V), energy vector direction (θ_E), velocity vector magnitude (r_V) and energy vector magnitude (r_E). In the case of n-loudspeaker arrangement, θ_V, θ_E, r_V and r_E are expressed as follows:

$$\theta_V = \tan^{-1}\left(\frac{\sum_{i=1}^{n} g_i \sin\theta_i}{\sum_{i=1}^{n} g_i \cos\theta_i}\right) \tag{1}$$

$$\theta_E = \tan^{-1}\left(\frac{\sum_{i=1}^{n} g_i^2 \sin\theta_i}{\sum_{i=1}^{n} g_i^2 \cos\theta_i}\right) \tag{2}$$

$$r_V = \sum_{i=1}^{n} g_i \tag{3}$$

$$r_E = \sum_{i=1}^{n} g_i^2 \tag{4}$$

Here, g_i and θ_i are the gain and the angular location of i loudspeaker, respectively. θ_V is the apparent direction of the sound source in a low frequency range (below 700 Hz) according to the inter-aural phase localization theory[4]. θ_E determines the apparent direction of the sound source in the frequency range from 700 Hz to 3.5 kHz. θ_E is considered as more useful criterion than $\theta_V$for the estimation of the angular position of the phantom sound source, in both cases of central listening and offsignificant frequency region for human perception. The equality between θ_V and θ_E should be retained for all panning angles, because θ_V is equal to θ_E for real sound sources. r_V is a quality index of the virtual images at low frequencies and must equal 1 for optimum quality. r_E is a quality criterion of the virtual images at high frequencies and must be as close to 1 as possible for optimum quality.

2.2. Gerzon' s method

Gerzon defined the feed for a second order APA, Vanderkooy and Lipshitz also used a similar feed[7]. According to the Gerzon, the feed for a second order N-loudspeaker APA is:

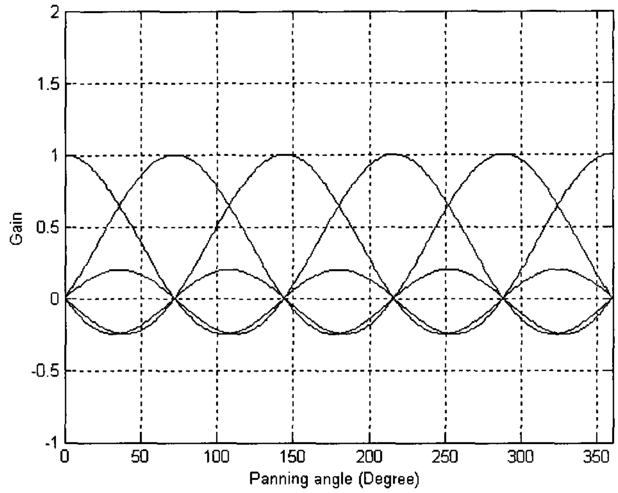

Fig. 1. Loudspeaker gains of a regular pentagon array: Gerzon' s 2nd order.

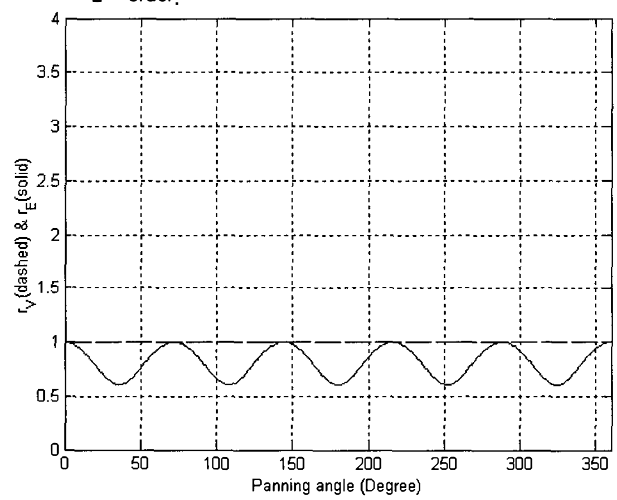

Fig. 2. r_V (dashed line) and r_E (solid line): Gerzon' s 2nd order.

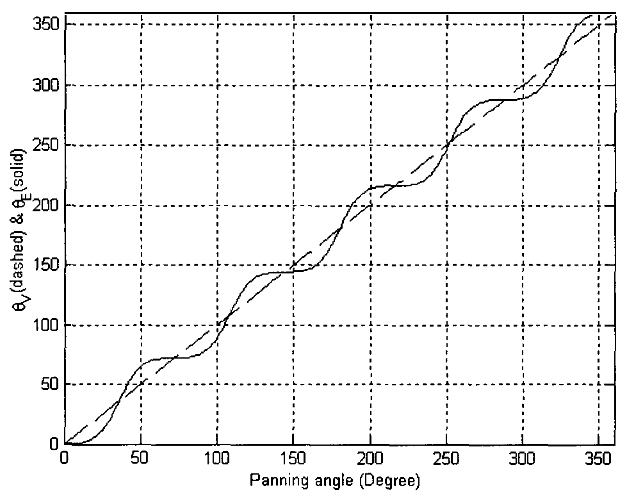

Fig. 3. θ_V (dashed line) and θ_E (solid line): Gerzon' s 2nd order.

$$P_n = \frac{1}{N}(W + 2X\cos\phi_n + 2Y\sin\phi_n + 2U\cos 2\phi_n + 2V\sin 2\phi_n) \tag{5}$$

where the W is 0^{th} order, X and Y is 1^{st} order, U and V is 2^{nd} order Ambisonic signal and P_n is the output signal of the n^{th} loudspeaker.

Fig. 1 shows the loudspeaker gains of a regular pentagon array according to the Eq. (1). Contrary to Gerzon's 1st order APA, which reproduces sound source by all loudspeakers, in order to localize in loudspeaker direction, the sound is being reproduced only by a single loudspeaker. Therefore, Gerzon's 2nd order panning algorithm performs considerably better in terms of channel separation and stability of the phantom sound image for a non-sweet spot listener, especially in the loudspeaker direction.

Fig. 2 shows the r_V and r_E according to the Gerzon's 2nd order APA. This has better performance in terms of sound localization, because of a larger and smoother value of r_E for all direction, when compared to Gerzon's 1st order. However, this method still has problems relating to the phantom sound image, which has minima of r_E in the center of both loudspeakers and maxima in the loudspeaker direction.

Fig. 3 shows the θ_V and $\theta_E$according to the Gerzon's 2nd order APA. This shows that θ_V is equal to panning angle but θ_E, representing a more important sound localization factor for human perception, as opposed to θ_V, which has a large detent effect.

2.3. Daniel' s method

Daniel indicated the problems of Gerzon's APA and subsequently proposed an APA using DPC[8]. According to psychoacoustic phenomenon that a phantom sound direction depends on its frequency components, Daniel uses correcting gains g_m in order to retain the property that θ_V is equal to panning angle for low frequency region (<700Hz) and that θ_E is equal to the panning angle for high frequency regions (>700Hz).

Equation (6) shows Daniel's Ambisonic panning algorithm using correcting gains.

$$\begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{pmatrix} = M_d \cdot \begin{pmatrix} g^0 W \\ g^1 X \\ g^1 Y \\ g^2 U \\ g^2 V \end{pmatrix} \tag{6}$$

where M_d is the decoding matrix and P_N is the output signal of the nth loudspeaker.

Fig. 4 shows the loudspeaker gains of a regular pentagon array

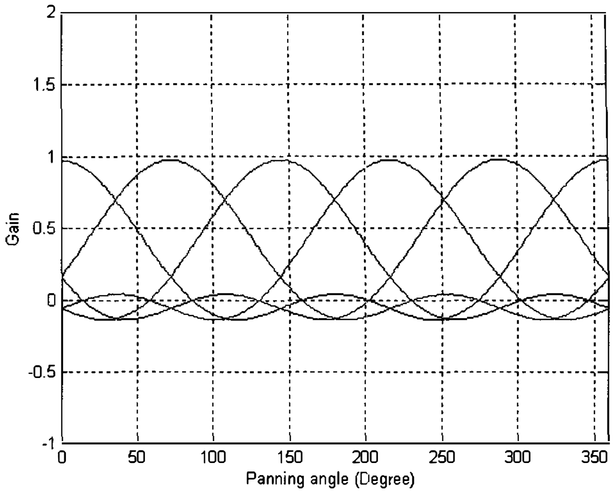

Fig. 4. Loudspeaker gains of a regular pentagon array: Daniel' s 2nd order.

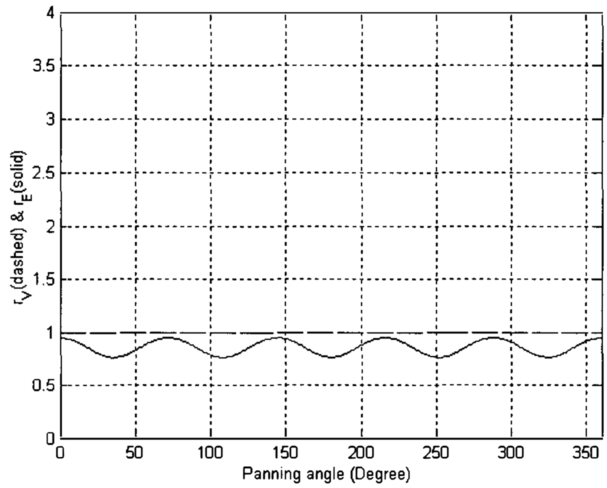

Fig. 5. r_V (dashed line) and r_E (solid line): Daniel' s 2nd order.

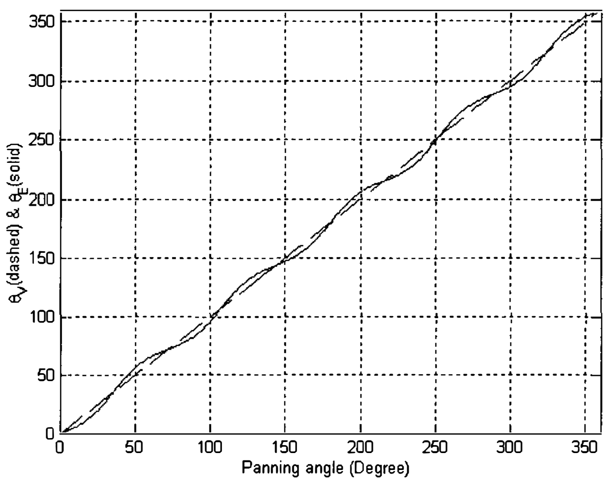

Fig. 6. θ_V (dashed line) and θ_E (solid line): Daniel' s 2nd order.

using Daniel's 2nd order APA, meeting the requirements to maximize r_E for high frequencies, and $r_V = 1$ for low frequencies. However, when compared to Gerzon's 2nd order APA, it has a low channel separation at loudspeaker directions and a high channel separation at other directions.

Fig. 5 shows the r_V and r_E according to the Daniel's 2nd order panning algorithm. This method has better performance of the sound localization, because of a larger and smoother value of r_E for all direction than Gerzon's 2nd order.

Fig. 6 shows the θ_V and $\theta_E$according to Daniel's 2nd order APA. It shows that θ_V is equal to panning angle and θ_E has smaller detent effect than Gerzon's 2nd order. Thus, sound localization error can be considerably reduced. To optimize DPC, Daniel's algorithm uses correcting gains to reduce the detent effect and maximize r_E to improve the quality of the phantom sound image.

For a non-sweet spot listener who isn't located in the central position of all loudspeakers, however, the phantom sound image with Daniel's APA has a tendency to be pulled closer to the nearest loudspeaker, not exactly matched with the intended panning angle, because every loudspeaker reproduces same sound wherever the position of the virtual sound source is located.

To eliminate this localization error, a local panning algorithm is required: for each position of the panning angle, 2 or 3 nearby loudspeakers are selected, these loudspeakers are fed with the sound source signal scaled by the panning algorithm and the gains of all the other loudspeakers are set to zero. This panning algorithm also must satisfy the DPC in order to minimize the sound localization error.

III. Optimized Ambisonic Panning Algorithm

To reduce the sound localization error and sound quality degradation of the conventional APA, an OAPA is derived using DPC. The requirements of the DPC in the design of an OAPA are as follows:

- The total power should be independent of the panning angle.
- The difference between θ_V , θ_E and the panning angle should be minimized.
- r_V, r_E should approximate to 1 and deviations according to the panning angle must be minimized.

To satisfy the requirements, the feed for an n-th loudspeaker (P_n) can be expressed as follows:

$$P_n = \frac{1}{N}(W + 2[f(n)(\sum_{m=1}^{M} P_\varphi \cos(m\varphi)\cos m\phi_n + P_\varphi \sin(m\varphi)\sin m\phi_n)]) \qquad (7)$$

where the f_n is a window function. Equation (7) is mathematically identical to the frequency response of an FIR filter designed by Fourier approximation and the stop band suppression, which corresponds here with the crosstalk to adjacent loudspeakers, and can be improved by applying an appropriate window function f_n, at the expense of an increased band pass width[9]. In the convention APA, the number of loudspeakers required to reproduce the two-dimensional Ambisonic panning signal is limited to 2M+1, where M is the order of the system, since prior research argues that higher order components higher than 2M+1 cause aliasing, increasing sound localization error[10].

In this paper, however, a proper window function can be

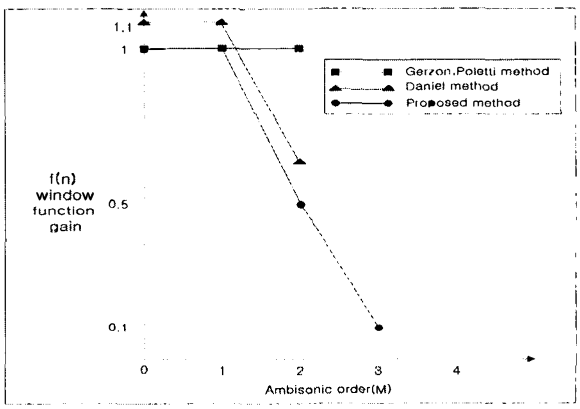

Fig. 7. Analysis of an Ambisonic panning algorithm using window function.

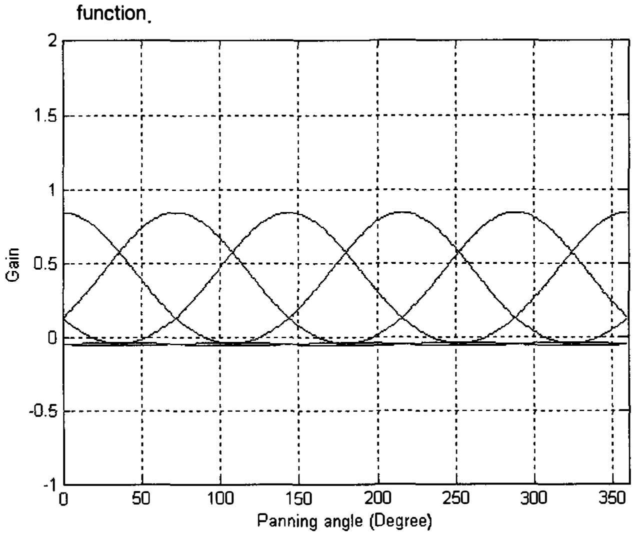

Fig. 8. Loudspeaker gains of a regular pentagon array: OAPA.

verified, comprising of higher harmonic components than 2M+1 and improved DPC and channel separation. A window function is chosen using the least square operation, showing the iterative cycle for finding an appropriate window function of the gains of the mth order Ambisonic components.

Fig. 7 shows an analysis of APA using different window functions, and the gains of window functions versus Ambisonic orders are drawn. Using window functions, existing APAs can be analyzed thoroughly. Gerzon and Poletti's method uses a brick wall window function, which has large side-lobe causing poor channel separation and large sound localization error. Daniel's method uses a simple window function, which has some small 2nd order Ambisonic component gain, and has superior DPC over Gerzon and Poletti's method. Since prior research limits the number of loudspeakers (N=2M+1), all methods use only 0th, 1st and 2nd Ambisonic components in a pentagon loudspeaker array. This cannot improve DPC and cannot reduce channel separation simultaneously. To resolve this problem, OAPA applies a window function, which uses higher Ambisonic components than 2nd order, chosen the gains of Ambisonic components using a least square operation.

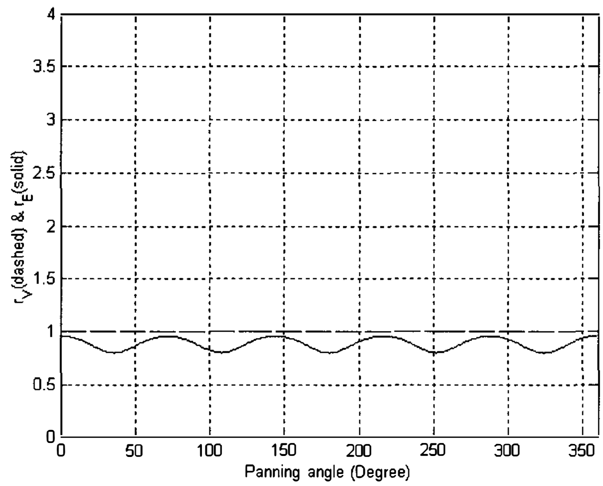

Fig. 9. r_V (dashed line) and r_E (solid line): OAPA.

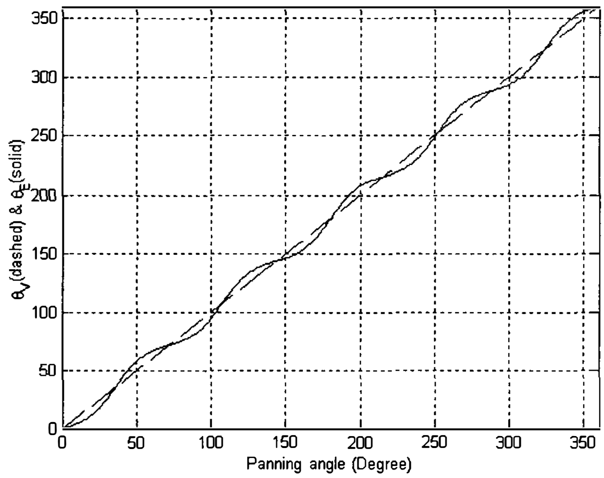

Fig. 10. θ_V (dashed line) and θ_E (solid line): OAPA.

Fig. 8 shows the loudspeaker gains of a regular pentagon array according to OAPA. Contrary to the conventional APA, every loudspeaker reproduces same sound wherever the position of the virtual sound source is, the sound is being reproduced by only 3 nearby loudspeakers and the gains of all the other loudspeakers are set to zero (e.g., triple-wise panning algorithm).

Hence, OAPA performs considerably better in terms of channel separation and stability of the phantom sound image for non-sweet spot listener. For a non-sweet spot listener, this triple-wise panning can significantly reduce sound localization error with which the phantom sound image has a tendency to be pulled closer to the nearest loudspeaker and is not exactly matched with the intended panning angle in the conventional APA.

Fig. 9 shows the r_V and r_E according to OAPA. OAPA has superior performance with regard to sound localization, because it has a larger and smoother value of r_E for all directions, in contrast to Gerzon's 2nd order method and Daniel's method.

Fig. 10 shows the θ_V and $\theta_E$according to OAPA. This means that θ_V is equal to the panning angle and θ_E has a smaller detent effect than Gerzon's 2nd order, almost the same as Daniel's 2nd order. The OAPA using a window function is a more appropriate APA than the conventional APA, because of the properties of triple-wise panning and DPC.

IV. Conclusion

Using both DPC analysis and signal separation among other channels, the conventional APA is confirmed as an inappropriate panning algorithm for surround sound systems. The sound localization is usually different from the panning angle, especially when listeners are not in an ideal listening position, because of low signal separation among other channels. To overcome this problem, the OAPA, using window functions, is proposed. An analysis shows that the proposed method results in higher signal separation among other channels and lower sound localization errors than the conventional APA. The OAPA can be used for various kinds of audio systems, for example, mixing consoles, sound reproduction systems and sound field processors.

References

1. S.-L. Lee, *Optimized multi-channel panning algorithm using directional psychoacoustic criteria and head related transfer function*, (Ph.D. thesis, Seoul National University, 2005)
2. H. D. Harwood, *Stereophonic image sharpness*, (Wireless World, **74**, 207-211, 1968)
3. C. Avendano, V. R. Algazi and R. O. Duda, "A head and torso model for low-frequency binaural elevation effects" , Proc.1999 IEEE Workshop on Application of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. **17** (20) 1999.
4. M. A. Gerzon, "Panpot laws for multispeaker stereo" , 92nd Convention of the Audio Engineering Society, J. Audio Eng. Soc., preprint 3309, 1992.
5. M. A. Gerzon, "Psychoacoustic decoders for multispeaker stereo and surround sound" , 93rd Convention of the Audio Engineering Society, J. Audio Eng. Soc., preprint 3406, 1992.
6. M. A. Gerzon, "General metatheory of auditory localization" , 92nd Convention of the Audio Engineering Society, J. Audio Eng. Soc., preprint 3306, 1992.
7. J. S. Bamford, *An analysis of Ambisonics of first and second order*, (M.Sc. thesis, University of Waterloo, Waterloo, Ont., Canada, 1995)
8. J. Daniel, J.-B. Rault and J.-D. Polack "Ambisonics encoding of other formats for multiple listening conditions" , 105th Convention of the Audio Engineering Society, preprint 4795, 1998.
9. U. Horbach "New techniques for the production of multichannel sound" , 103rd Convention of the Audio Engineering Society, preprint 4624, 1997.
10. J. S. Bamford and J. Vanderkooy. "Ambisonic Sound for us" , 99th Convention of the Audio Engineering Society, preprint 4138, 1995.

Acknowledgment

This work was supported by grant No. (R-01-2006-000-10717-0) from the Basic Research Program of the Korea Science & Engineering Foundation.

【Profile】

● Sin-Lyul Lee

was born in Gosong, Korea in 1971. He received the B.S. degree from the School of Computer Engineering, Kwangwoon University, Korea, in 1999 and the M.S. and Ph. D. degree from the School of Electrical Engineering and Computer Science, Seoul National University, Korea, in 2002 and 2005, respectively. Currently he is a postdoctoral researcher with the Institute of New Media and Communications at Seoul National University. His research interests are in the fields of audio signal processing, spatial audio, psychoacoustics and electro-acoustics. He is now a member of Acoustical Society of Korea, Audio Engineering Society of Korea and Society for 3D Broadcasting and Imaging.

●Seung-Rae Lee

received the B.S. degree in 1991, M.S. degree in 1993, and doctoral degree in 1997, all in mathematics from Aachen University of Technology, Germany. From 1995 to 1996 he was an assistant in mathematics at the Gerhard-Mercator University of Duisburg, Germany. From 1998 to 2000, he was a postdoctoral researcher with the Automation and Systems Research Institute at Seoul National University. In 2000, he was a research professor with the School of Electrical and Computer Engineering at Sungkyunkwan University. From 2001 to 2003, he was a senior researcher with the Institute of New Media and Communication at Seoul National University. Currently he is an assistant professor with the Electrical Engineering and Computer Science at Seoul National University. His current research interests are biometrics, 3-D sound system, differential games and optimal control.

●Koeng-Mo Sung

was born in Incheon, Korea in 1947. He was in the Department of Electronics Engineering at Seoul National University from 1965 to 1971. He received the Dipl.-Ing. in communication engineering in 1977 and the Dr.- Ing. degree in acoustics in 1982 from Technische Hochschule Aachen, Germany. He was a research engineer at RWTH Aachen from 1977 to 1983. Since 1983 he has been with Seoul National University, where he is a Professor in the School of Electrical Engineering. He was the president of Directors of the Acoustical Society of Korea. His research interests are ultrasonics, musical acoustics and audio signal processing.