

# 3차원 라우드스피커 어레이를 이용한 음장재현기술

## Sound Field Reconstruction Technology Using a Three Dimensional Loudspeaker Array

서정일\*, 강경옥\*, Filippo M. Fazi\*\*, Philip A. Nelson\*\*  
(Jeongil Seo\*, Kyeongok Kang\*, Filippo M. Fazi\*\*, Philip A. Nelson\*\*)

\*한국전자통신연구원 방통융합미디어연구부, \*\*ISVR, Southampton University, U.K.  
(접수일자: 2009년 11월 5일; 채택일자: 2009년 11월 16일)

본 논문에서는 3차원 라우드스피커 어레이를 이용하여 다수의 청취자에게 동일하게 실감 있는 음장을 제공하기 위한 음장 재현 알고리즘을 제안한다. 제안된 알고리즘은 제어하고자 하는 3차원 공간 영역을 정의하고, 정의된 영역에서 원음장과 라우드스피커 어레이를 이용하여 재현되는 음장간의 자승오차가 최소가 되도록 라우드스피커 구동신호를 제어하는데 기초한다. 제안된 알고리즘의 성능을 검증하기 위하여 40개의 라우드스피커를 이용한 구형 3차원 어레이를 구성하였으며, 이를 이용한 실험결과와 고찰결과를 제시한다.

**핵심용어:** 3차원 음장, 음장 재현, 라우드스피커 어레이, 최소자승오차 방법

**투고분야:** 음향 신호처리 분야 (1,2)

In this paper, we propose a novel sound field reconstruction algorithm using a three dimensional loudspeaker array for providing realistic sound field to multiple listeners. The proposed algorithm is based on minimization of the squared error between the original sound field and the reconstructed sound field by the loudspeaker array over a predefined three dimensional region of the space using a loudspeaker array surrounding the listening area. For evaluating the proposed algorithm, we constructed the three dimensional array composed of 40 loudspeakers and discuss the relevant experiment results.

**Keywords:** Three Dimensional Sound Field, Sound Field Reconstruction, Loudspeaker Array, Least Squared Method

**ASK subject classification:** Acoustic Signal Processing (1,2)

### I. 서론

20세기를 대표하던 스테레오 음향재생 시스템은 DVD (Digital Versatile Disc)와 HDTV (High Definition Television)의 등장으로 인하여 지난 10여 년간 5.1채널 서라운드 음향재생 시스템으로 진화되어 갔다 [1]. 5.1채널 서라운드 음향재생 시스템은 좌우 스테레오 라우드스피커를 기본으로 하여, 중앙 라우드스피커는 배우나 가수의 목소리를 화면과 동일한 위치에 놓이도록 하고 좌우 후방 라우드스피커를 이용하여 배경음을 충실히 재현하는데 목적을 두고 있다. 그러나 궁극의 현장감을 제공하기 위한 3DTV와 UHDTV (Ultra High Definition

Television)의 등장으로 인하여 5.1채널 서라운드 음향재생 시스템만이 고현장감 영상 신호에 부합하는 실감음향 음 제공하는 것에는 한계가 있음이 제기되어, 5.1채널 이상의 라우드스피커 어레이를 이용하여 3차원 실감 음장을 제공할 수 있는 시스템들이 제안되고 있다.

물론 지난 30년간 라우드스피커 어레이를 이용하여 원하는 음장을 재현하기 위한 기술들이 제안되어 왔다. Kirchhoff Helmholtz 적분식에 기반을 두는 파면음장합성 (Wave Field Synthesis) 기술이 대표적인 기술이다. [2-3] 멀티채널 오디오에 광범위하게 적용 중인 엠비소닉스 (Ambisonics) 기술도 이러한 기술 중의 하나이며, Gerzon에 의해 처음으로 제안된 이후 [4] 최근 들어 음장을 재현하는데 보다 효과적인 다차원 엠비소닉스 (High Order Ambisonics)로 [5] 알고리즘이 발전되어 가고 있다. 또한, 최근에는 구형 하모닉스 (spherical harmonics)를 이용하여 원하는 음장을 합성하는 기술들과 [6], 음향

학적 역함수의 해를 풀어가는데 기반 하거나 새로운 원리를 이용하는 기술들이 제안되고 있다 [7-8].

본 논문에서는 3차원 라우드스피커 어레이를 이용하여 음장을 재현하는 새로운 알고리즘을 제안한다. 제안되는 알고리즘은 먼저 라우드스피커 어레이가 2차음원 (secondary sound source)으로 3차원 공간상에 연속적으로 채적을 이루는 경계면에 분포된다고 가정한다. 이러한 2차음원의 크기는 Kirchhoff Helmholtz 적분식으로 잘 알려진 적분식의 해를 구하는 과정으로 계산될 수 있다. 이러한 배경 논리는 본 논문의 저자들에 의해 처음 제안되었으며 [9-10], 본 논문에서는 실제 구현을 통해 제안된 알고리즘의 유효성을 검증하는데 그 목적이 있다.

본 논문은 다음과 같이 구성된다. 2장에서는 제안된 알고리즘의 이론적 기초를 설명하고, 3장에서는 라우드스피커 어레이로 구성되는 2차음원이 구면으로 존재하는 환경에서의 제안된 알고리즘의 적용방법에 대해 설명한다. 4장에서는 실제 제한된 라우드스피커를 이용하였을 경우의 계산적 방법을 설명하고, 5장에서는 실제 구현을 통해 얻어진 실험결과를 제시하고 분석한다. 마지막으로 6장에서 결론을 맺는다.

## II. 제안된 알고리즘의 이론적 기초

그림 1에서와 같은 3차원 공간상에서 라우드스피커에 의해서 재현되는 영역을  $V$ 라고 할 때 그 경계면은  $\partial V$ 라고 할 때, 음원이  $V$ 내에 존재하지 않는다고 가정하면 복원하고자 하는 음장  $p(x)$ 은 아래와 같은 균등 파면 방정식 (homogeneous wave equation)을 따른다.

$$\nabla^2 p(x,t) - \frac{1}{c^2} \frac{\partial p(x,t)}{\partial t^2} = 0 \quad (1)$$

여기서  $c$ 는 소리의 속도를 나타낸다.

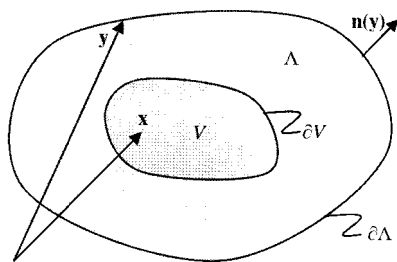


그림 1. 임의의 3차원 공간에서의 영역  $V$ 와  $A$ .  
Fig. 1. Three dimensional regions  $V$  and  $A$ .

단일 주파수  $\omega$ 를 가정하고  $p(x,t) = Re\{p(x,t)e^{j\omega t}\}$ 라 면 수식 1은 Helmholtz 방정식으로 아래와 같이 변형 된다.

$$\nabla^2 p(x) + k^2 p(x) = 0, x \in V \quad (2)$$

여기서  $k = \omega/c$ 는 파수 (wave number)를 의미하며, 시간축 성분인  $e^{j\omega t}$ 는 생략되어 있다.

라우드스피커가 배치되는 이차음원 (monopole 음원) 들로 이루어지는 경계  $A$ 가  $V$ 를 완전히 둘러싼다고 가정한다.  $V$ 의 경계에서의 음장  $P_y(x)$ 는 그린함수 (green function)로 표현할 수 있으며, 따라서 아래 수식과 같은 비균질 (inhomogeneous) Helmholtz 방정식으로 표현된다.

$$\Delta^2 p_y(x) + k^2 p_y(x) = -a(y)\delta(x-y), x \in \bar{V}, y \in \partial A \quad (3)$$

여기서  $a(y)$ 는 이차음원의 크기를 나타낸다.

자유공간 (free space)에서  $P_y(x)$ 는 자유공간 그린함수로 다음 수식과 같이 표현할 수 있다.

$$p_y(x) = a(y)g(x,y) = a(y) \frac{e^{-jk|y-x|}}{4\pi|y-x|}, x \in R^3, y \in \partial A \quad (4)$$

잔향환경에서는 그린함수가 보다 복잡한 형태로 표현되며 이는 잔향을 발생시키는 구조물의 형태와 벽면의 특성에 의존하게 된다. 균일하게 분포된 이차음원들에 의하여 발생하는 음장  $P_y(x)$ 는 단일음원에 의해 발생된 음장의 중첩 (superposition)으로 다음 수식과 같이 표현할 수 있다.

$$\hat{p}(x) = (Sa)(x) = \int_{\partial A} G(x,y)a(y)dS(y), x \in V \quad (5)$$

식 5와 같이 음장을 표현할 경우 아래와 같은 두 가지 문제점이 발생한다.

- 만약 복원하고자 하는 음장  $p(x)$ 가 복원하고자 하는 영역  $V$ 의 경계  $\partial V$ 에서 복원할 수 있을 때, 영역  $V$  내에서도 원하는 음장을 복원할 수 있는가?
- 만약 복원하고자 하는 음장  $p(x)$ 를 영역  $V$ 의 경계  $\partial V$ 에서 알 수 있을 때,  $p(x)$ 가 이차음원으로부터 재현되는 음상  $\hat{p}(x)$ 와 같아지도록 하는 이차음원의 크기 (라우드스피커 구동신호의 크기)  $a(y)$ 를 계산할 수 있는가?

첫 번째 문제점은 Interior Dirichlet 문제에 대한 해를 구할 수 있는가에 대한 문제와 관련이 있으며, 이러한 문

제는 아래와 같은 두 개의 미분방정식의 단일해가 존재하는가에 대한 문제로 귀결된다.

$$\begin{aligned} \nabla^2 p(x) + k^2 p(x) &= 0, \quad x \in V \\ p(x) &= f(x), \quad x \in \partial V \end{aligned} \quad (6)$$

여기서 함수  $f(x)$ 는 복원하고자 하는 음장  $p(x)$ 의 복원하고자 하는 영역의 경계  $\partial V$ 에서의 값을 나타낸다.

이러한 미분방정식은 Dirichlet 문제로 알려져 있으며, 관련된 경계조건 (boundary condition)도 같은 이름으로 사용되고 있다. Kirchhoff-Helmholtz 적분방정식에서 제시하는 것과 같이 음장에 대한 지식과 음장에 대한 일반 도함수 (normal derivative)에 대한 지식을 이용하면 복원하고자 하는 영역  $V$ 내에서 해를 구할 수 있다. 또한 함수  $f(x)$ 에 대한 적절한 정규성 (regularity) 가정을 이용하면 Dirichlet 문제는 단일해를 가진다. 단 파수  $k$ 가 Dirichlet 고유치 (eigenvalue)  $k_n$  중 하나가 되어서는 안되며, Dirichlet 고유치: 음장을 복원하고자 하는 영역  $V$ 의 모양과 경계에서의 압력에 의해 정의되는 내부 공진 주파수  $\omega_n = k_n c$ 로 결정된다. 따라서 내부 공진 주파수에서는 식 6의 해를 구할 수 없게 되므로 원하는 음장을 재현할 수 없게 된다.

두 번째 문제점은 복원하고자 하는 영역의 경계면  $\partial V$ 가 알려졌을 때 연속적인 모노폴 이차음원을 이용하여 원하는 음장을 재현할 수 있는가와 연관된다. 이는 식 5로 표현되는 복원음장  $\hat{p}(x)$ 가 원음장  $p(x)$ 와 같아지는 이차음원들의 크기  $a(y)$ 를 찾아가는 과정과 연관된다. 앞서 설명한 Dirichlet 문제를 따라가서 만약 동작 주파수가 Dirichlet 고유치에 의해 정의되는 공진주파수가 아니고, 복원하고자 하는 영역의 경계면  $\partial V$ 에서 정확하게 복원되었다면, 영역  $V$  내에서도 동일하게 재현되게 된다. 이러한 재현문제는 다음과 같이 수식화 할 수 있다.

$$p(x) = (Sa)(x) = \int_{\partial V} G(x|y)a(y)dS(y), \quad x \in \partial V \quad (7)$$

식 7에서 원음장  $p(x)$ 가 주어진다면 이차음원들의 크기  $a(y)$ 를 찾아야 한다. 식 (7)은 Volterra 적분방정식의 첫 번째 해가 되고, 경계면  $\partial V$ 에서의 원음장  $p(x)$ 로 알 수 있는  $a(y)$ 의 determination은 역함수 문제로 풀어낼 수 있게 된다. 그러나 역함수를 계산하는 과정을 포함하기 때문에 식 (7)은 illposed problem이며, 해가 존재하더라도 유일한 해가 아니거나 원음장  $p(x)$  데이터에 연속적으로 의존되지 않는 해로 존재할 가능성이 있다. 다만 유일한 해가 존재하지 않거나 안정적인 해가 존재하지

않더라도 적절한 가정이나 근사화를 통해 안정적인 해를 구하는 방법들이 제시되어 있다. 본 논문에서는 식 (7)의 해를 구하기 위하여 적분식을 SVD (singular value decomposition)하여 안정된 해를 구하는 방식을 적용하였다 [5].

같은 도메인의 함수  $f(x)$ 와  $g(x)$ 의 내적은 다음과 같이 표현된다.

$$\langle fg \rangle = \int_V f(x)^* g(x) dS(x) \quad (8)$$

매트릭스  $S$ 의 수반행렬(adjoint matrix)을  $S^+$ 라고 하면 수반연산 과정은 다음과 같이 표현된다.

$$\langle g|Sa \rangle = \langle S^+g|a \rangle \quad (9)$$

수반연산자(adjoint operator)에 의한 식 7에서의 적분 연산자  $S$ 는 아래와 같이 변형된다.

$$(S^+g)(y) = \int_{\partial V} G(y|x)^* g(x) dS(x), \quad y \in \partial V \quad (10)$$

단일계층전위 (single layer potential)  $(Sa)(x)$ 의 물리학적 의미는 경계면  $\partial V$ 에 연속적으로 위치하는 이차음원들에 의해 생성되는 음장을 나타낸다. 따라서 이의 수반연산자  $(S^+a)(y)$ 는 경계면  $\partial V$ 에 위치하는 이차음원들에 의해 생성되는 음장신호가 시간축으로 뒤집힌 것으로 나타난다. 이러한 시간축상의 반전은 식 (10)의 커널 (kernel)이 그린함수  $G(y|x)$ 의 켈레복소수 (complex conjugate)이기 때문이다. 따라서 식 4에서의 자유공간 그린함수  $g(x|y)$ 의 경우에  $g(y|x)$ 와  $g(y|x)^*$ 가 다르다는 것은 지수부의 부호가 달라진다는 것에 따라 쉽게 알 수 있다. 이러한 사실은  $g(\cdot|x)$ 가  $x$  위치에서 밖으로 발산하는 구면파(spherical wave)로 여길 수 있다면,  $g(\cdot|x)^*$ 는 안으로 들어오는 구면파로 여길 수 있다는 것을 의미한다. 이와 같이  $g(\cdot|x)^*$ 로 표현되는 음장은  $g(\cdot|x)$  음장의 시간축상의 반전된 형태로 인표현할 수 있다. 따라서

$$\begin{aligned} Re\{e^{-j\omega t} g(x|y)^*\} &= Re\left\{ \frac{e^{-j(\omega(x-y) + \omega(t))}}{4\pi|x-y|} \right\} \\ &= Re\{e^{-j\omega(t)} g(y|x)\} \end{aligned} \quad (11)$$

이러한 해석과정을 따르면, 생성된 구면파의 파면은  $x$ 를 향하여 수렴된다. 결론적으로 수반연산자  $S^+$ 은 경계면  $\partial V$ 에 입력되는 파면을 이루는 음원들의 연속된 분

포로 해석된다. 주파수축 상에서의 분해 (decomposition) 를 수행하기 위하여 자기 수반연산자 (self adjoint operator) 특성을 이용하면 이차음원을 표현하는 함수  $a(y)$ 는 아래와 같이 간략화 할 수 있다.

$$(S^+ S a_n)(y) = \lambda_n a_n(y), y \in \partial A, n = 1, 2, 3, \dots, \infty \quad (12)$$

여기서 eigenvalue  $\lambda_n$ 은 실수이며 양의 값을 가진다.

$S$ 와  $S^+$ 에 대한 해석을 식 (12)와 같은 개념으로 이해 하면, 합성연산자 (composite operator)  $(S^+ S a)(y)$ 는 다음과 같이 이해될 수 있다.

음장 (sound field)은 경계면  $\partial A$ 에 있는 연속적으로 분포된 모노폴 이차음원들로부터 생성되며, 이들의 크기는  $a(y)$ 이다.

이러한 경계면  $\partial V$ 는 음장은 함수  $\hat{p}(x) = (S a)(x)$ 로 표현된다.

실제적인 문제로 전환하여 이차음원들은 라우드스피커 어레이로 치환되고, 함수  $a_n(y)$ 는 어레이 모드로 해석된다. 함수  $a_n(y)$ 의 집합은 경계면  $\partial A$ 에서는 함수들의 직교 집합 (orthogonal set)으로 구성된다. 이는  $\partial A$ 에서 정의 되는  $a(y)$ 는 아래와 같이 표현될 수 있음을 의미한다.

$$a(y) = \sum_{n=1}^N \langle a_n | a \rangle a_n(y) + (Qa)(y) \quad (13)$$

여기서 정수  $N$ 은 영역  $S$ 의 차수에 의존하고 무한대까지 확장될 수 있으며,  $(Qa)(y)$ 는 영공간 (null space)  $S$ 로의 직교투영 (orthogonal projection)이다. 함수  $\tilde{a}(y)$ 로 이루어지는 집합은,

$$N(S) = \{ \tilde{a}(y) : (S \tilde{a})(x) = 0 \} \quad (14)$$

로 정의된다.

반약 식 (14)로 표현되는 집합의 원소가 존재하지 않으면,  $a_n(y)$  함수들로 구성되는 집합은 완전해지고, 식 14는 일반적인 푸리에 역급수 (Fourier Series)로 표현할 수 있다. 또한 적분연산자  $S$ 를  $a_n(y)$ 에 적용하면,  $\partial V$ 에서의 직교함수  $p_n(x)$ 들은 다음과 같이 표현된다.

$$\mu_n p_n(x) = (S a_n) \quad n = 1, 2, 3, \dots, \infty \quad (15)$$

여기서 양수이면서 실수인  $\sigma_n = \sqrt{\lambda_n}$ 는 적분연산자  $S$ 의 eigenvalue를 의미한다. 따라서 식 15는 다음과 같이 표현될 수 있다.

$$\delta_n a_n(x) = (S^+ p_n) \quad n = 1, 2, 3, \dots, \infty \quad (16)$$

경계면  $\partial V$ 에서의 입의의 평방적분함수 (square integral function)  $p(x)$ 는 다음과 같은 무한 역급수로 표현할 수 있다.

$$p(x) = \sum_{n=1}^N \langle p_n | p \rangle p_n(x) + (Rp)(x) \quad (17)$$

여기서  $(Rp)(x)$ 는 영공간  $S^+$  영역으로 직교투영이다.

식 13과 15를 결합하고, 영공간의 정의에 따라  $(S(Qa))(x) = 0$ 이라면,  $a(y)$ 에 대한  $S$ 의 영향은 다음 수식으로 표현할 수 있다.

$$(Sa)(x) = \sum_{n=1}^N \delta_n \langle a_n | a \rangle p_n(x) = \hat{p}(x), x \in \partial V \quad (18)$$

따라서 식 7에 대한 근사해는 다음과 같이 계산할 수 있다.

$$a(y) = \sum_{n=1}^N \frac{1}{\delta_n} \langle p_n | p \rangle a_n(y) \quad (19)$$

식 (15), (17), (19)를 고려하고, 식 (18)에 의해 계산되는 단일계층전위는 다음과 같이 표현된다.

$$\begin{aligned} & \int_{\partial A} G(x|y) a(y) dS(y) \\ &= \int_{\partial A} \left( \sum_{n=1}^N \frac{1}{\delta_n} \langle p_n | p \rangle a_n(y) \right) dS(y) \\ &= \sum_{n=1}^N \frac{1}{\delta_n} \langle p_n | p \rangle (S a_n)(x) = \sum_{n=1}^N \langle p_n | p \rangle p_n(x) \\ & \quad - p(x) - (Rp)(x) = \hat{p}(x) \end{aligned} \quad (20)$$

식 (20)은 복원음장  $\hat{p}(x)$ 는 복원을 복표로 하는 음장 중 부분일 뿐 수반연산자  $S^+$ 로 정의되는 영공간에는 포함되지 않는다는 것을 의미한다.  $1/\delta_n$  팩터는 원음장  $p(x)$ 에 포함되어 있는 에러성분에 의한 증폭치와 관계가 있으며, 이는 복원음장의 안정성 (stability)와 관련된다.  $1/\delta_n$  팩터에 의하여 발생하는 ill-condition 문제는 주파수 영역에서의 cut-off 주파수의 조정으로 정규화 (regularization) 하거나 다른 정규화 기술을 이용하여 극복될 수 있다. 원음장  $p(x)$ 를 복원하기 위하여  $a(y)$ 를 계산할 수 있는 가능성은 전술한 내용으로 설명되며,  $a(y)$ 를 정확하게 계산할 수 없다고 하더라도 근사해를 도출할 수 있으므로 완벽한 음장을 복원할 수 없다는 것을 의미하지는 않는다.

### III. 구형 음장환경에서의 해석적 방법

식 19가 식 7의 해를 계산하기 위한 효율적인 방법이라고 하더라도 함수  $a_n(y)$ 와  $p_n(x)$ 의 determination은 일반적인 해를 가진 뿐 명백한 해를 가지지는 않는다. 따라서 해를 구하는데 필요한 과정은 일반적으로 수식을 풀어나가며 진행된다. 그러나  $V$ 와  $A$ 가 구 (spherical)와 같은 특별한 모양을 가질 때는 해석상의 근사화를 통해 간단하게 계산될 수 있다. 만약  $V$ 와  $A$ 가 동심원 형태의 구 (concentric sphere)이고 각각  $R_V$ 와  $R_A$ 의 반지름을 가진다면, 자유공간에 대한 가정 (assumption)을 적용할 수 있다. 따라서 아래와 같은 해석상의 전개가 가능하다.

$$\begin{aligned} a_n(y) &= \frac{1}{R_A} Y_n^m(\theta_y, \phi_y) \\ p_n(x) &= \frac{\gamma_n}{R_V} Y_n^m(\theta_x, \phi_x) \\ \delta_n &= R_A R_V k |h_n^{(1)}(kR_A) j_n(kR_V)| \end{aligned} \tag{21}$$

여기서  $j_n(\cdot)$ 은  $n$ 차의 구형 베셀함수 (spherical Bessel function)를,  $h_n^{(1)}(\cdot)$ 은  $n$ 차의 구형 행켄함수 (spherical Hankel function)를 의미하며, 구형 하모닉스 (spherical harmonics)  $Y_n^m(\theta, \phi)$ 는 다음과 같이 정의된다.

$$Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos \theta) e^{im\phi} \tag{22}$$

여기에서  $P_n^m$ 은 레젠드르 함수 (Legendre function)와 관련되며, 일원 norm을 가지는  $\gamma_n$ 은 다음 수식으로 표현할 수 있으며  $S$ 연산에 의한  $n$ 차수의 구형 하모닉스들의 위상 이동으로 해석된다.

$$\gamma_n = j \frac{h_n^{(1)}(kR_A) j_n(kR_V)}{[h_n^{(1)}(kR_A) j_n(kR_V)]} \tag{23}$$

상기 수식들에서 고유치  $\delta_n$ 은 하나의 인덱스를 가지고, 구형 하모닉스  $Y_n^m(\theta, \phi)$ 는 두 개의 인덱스를 가진다. 이는  $(2n+1)$ 차수의 하나의 고유영역 (eigenspace)에서는 고유치  $\delta_n$ 만 관련됨을 의미하며, 따라서 각  $n$ 차수마다  $(2n+1)$ 개의 직교 구형 하모닉스들의 집합을 가질 수 있음을 의미한다. 즉  $n$ 차수의  $m$  계수의 구형 하모닉스는 같은 고유치  $\delta_n$ 과 연관된다. 이러한 일반화는 구와 같은 대칭 (symmetric) 구조에서만 가능하다. 구형 하모닉스의 직교성 (orthogonality)을 바탕으로 식 (21)에서의 정의된  $p_n(x)$ 와  $a_n(y)$  함수들의 직교성이 유도된다. 또한 자유

공간 그린함수의 확장을 이용하여 함수  $p_n(x)$ 와  $a_n(y)$ 와 고유치  $\delta_n$ 은 식 (12), (15), (16)을 만족함을 유도할 수 있다.

전술한 근거를 바탕으로 복원을 목표로 하는 음장  $p(x)$ 의 구형파 스펙트럼  $S_{nm}(r)$ 은 다음 수식과 같이 정의된다. 단  $r$ 은  $R_V$ 를 의미한다.

$$S_{nm}(R_V) = \int_0^{2\pi} d\theta \int_0^\pi p(R_V, \theta, \phi) Y_n^m(\theta, \phi) \sin(\theta) d\theta \tag{24}$$

따라서  $a(y)$ 는 해석적으로 다음과 같이 계산될 수 있다.

$$a(y) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{Y_n^m(\theta_y, \phi_y)}{R_A^2 j k h_n^{(1)}(kR_A) j_n(kR_V)} S_{nm}(R_V) \tag{25}$$

또한 상기 수식은 단일 점음원일 경우와 평면파 (plane wave)일 경우에는 다음과 같이 간략화 된다.

$$\begin{aligned} a^{ps} &= \sum_{n=0}^{\infty} \frac{h_n^{(1)}(ky_0)}{R_A^2 h_n^{(1)}(kR_A)} \frac{2n+1}{4\pi} P_n(\cos(\zeta)) \\ a^{pw} &= \sum_{n=0}^{\infty} \frac{4\pi j^{n-1}}{kR_A^2 h_n^{(1)}(kR_A)} \frac{2n+1}{4\pi} P_n(\cos(\zeta)) \end{aligned} \tag{26}$$

여기서  $P_n(\cdot)$ 은  $n$ 차수의 레젠드르 다항식 (polynomial)이며,  $\cos(\zeta)$ 는 다음과 같이 계산된다.

$$\cos(\zeta) = \sin(\theta_y) \sin(\theta) \cos(\theta_y - \phi) + \cos(\theta_y) \cos(\theta) \tag{27}$$

여기서 단일 점음원일 경우에는  $\theta'$ 는  $\theta_m$ 이고  $\phi'$ 은  $\phi_m$ 이며, 평면파이면  $\theta'$ 는  $\theta_{k_0}$ 이고  $\phi'$ 은  $\phi_{k_0}$ 이다.

### IV. 실제 제한된 라우드스피커 환경에서의 계산적 방법

이상적인 가정들을 근거로 논의된 이론적인 접근은 실제 환경에서는 적절하지 못할 수 있다. 만약 음장복원에 기초하는 서라운드 시스템을 실현하기 위해서는 이러한 가정들 몇몇은 제거되어야 한다. 먼저 경계면  $\partial A$  위에 모노폴로 이루어지는 이차음원들이 연속적으로 분포한다는 가정은 라우드스피커 어레이로 변경되어야 한다. 이는 이차음원이 더 이상 모노폴 음원으로 가정할 수 없음을 의미한다. 더군다나 라우드스피커의 개수 또한 한정되게 된다. 따라서 연속적인 이차음원들의 분포는 이

차음원들이 표본화된 이산분포 (discrete distribution)로 변경되어야 한다.

복원하고자 하는 음장을 표현하는 기법도 실제적인 방법으로 변경되어야 한다. 이는 복원하고자 하는 음장이 지향성 (directivity)과 위치 등과 같은 정보가 알려진 음원으로부터 생성되는 환경과 복원하고자 하는 음장에 대한 사전정보 없이 음장을 표현하는 방법이 있을 수 있다. 후자의 경우에는 복원하고자 하는 음장에 대한 정보는 마이크로폰 어레이로 측정되어야 하며, 이러한 과정은 복원음장이 공간적으로 표본화 (sampling) 된다는 것을 의미한다. 따라서 이러한 과정은 공간적 위신평 (aliasing)와 같은 오차를 발생하게 된다.

따라서  $\partial A$  위에 존재하는 이차음원들에 대한 연속분포 (continuous distribution)가 이산분포 (discrete distribution)로 변경되고 보노폴 음원을 전방향 점음원 (omnidirectional point source)으로 변경하면, 2장에서 도출한 이론을 복잡한 지향성을 가지는 라우드스피커 어레이의 경우도 확장하는 것이 이론적으로 가능한 것 인지를 이장에서 논의하고자 한다.

라우드스피커의 개수를  $L$  이라고 하고, 라우드스피커가  $\partial A$  위에 위치한다는 것을 제외하고 라우드스피커들의 배치에 대한 어떠한 가정도 두지 않는다면, 델타함수  $\delta(\mathbf{y}-\mathbf{y}_0)$ 는  $\partial A$ 에서 정의되는 함수  $f(\mathbf{x})$ 로 아래와 같이 정의된다.

$$f(\mathbf{y}_0) = \int_{\partial A} f(\mathbf{y})\delta(\mathbf{y}-\mathbf{y}_0)dS(\mathbf{y}), \mathbf{y}_0 \in \partial A \quad (28)$$

표본화된 밀도함수  $a_s(\mathbf{y})$ 는 다음과 같이 표현할 수 있다.

$$a_s(\mathbf{y}) = \sum_{l=1}^L a_l(\mathbf{y})\delta(\mathbf{y}-\mathbf{y}_l)\nabla S_l \quad (29)$$

여기서 실수이며 양수인 계수  $\nabla S_l$ 는 영역의 차수와 적용된 표본화 기법에 의해 결정되며, 등간격 표본화 (regular sampling)인 경우에는  $\nabla S_l$ 이 상수가 되고 인덱스  $l$ 과도 독립적인 계수가 된다.

따라서 재현음장 (reconstructed sound filed)은 다음과 같이 표현할 수 있다.

$$\hat{p}(\mathbf{x}) = \int_{\partial A} G(\mathbf{y}|\mathbf{x})a_s(\mathbf{y})dS(\mathbf{y}), \mathbf{x} \in V \quad (30)$$

이러한 표본화 과정으로 인해 발생하는 음장복원 성능을 살펴보기 위하여  $\partial V$  경계면에서의 복원음장은 다음

과 같이 주어진다.

$$\hat{p}(\mathbf{x}) - p(\mathbf{x}) = e(\mathbf{x}), \mathbf{x} \in \partial V \quad (31)$$

여기서  $e(\mathbf{x})$ 는 복원과정에서 발생하는 오차를 나타내며,  $e(\mathbf{x})$ 는 다음과 같이 정의할 수 있다.

$$e(\mathbf{x}) = \sum_{n=0}^{\infty} \sum_{v=0}^{\infty} e_{nv} \frac{\delta_n}{\delta_v} < p_v | p > p_n(\mathbf{x}) \quad (32)$$

$$e_{nv} = \sum_{l=1}^L a_n(\mathbf{y}_l)\Delta S_l - \delta_{nv} \quad (33)$$

여기서  $e_{nv}$ 의 직교특성은  $a_n(\mathbf{y})$ 와  $a_v(\mathbf{y})$  신호의 표본화 과정으로 인하여 라우드스피커 어레이의 배열환경과 밀접한 연관을 가진다.

#### 4.1. 이차음원들이 구면에 정규적으로 분포할 때

$\partial A$ 와  $\partial V$ 가 동심원 형태의 구라면,  $p_n(\mathbf{x})$ ,  $a_n(\mathbf{y})$ ,  $\delta_n$ 을 식 21과 같이 간략화하여 표현할 수 있으며, 복원음장은 식 24와 같이 주파수영역에서 표현할 수 있다. 따라서 복원과정에서 발생하는 오차신호는 다음과 같이 정의된다.

$$e_{nv}^{mv} = \frac{4\pi}{L} \sum_{l=1}^L Y_n^m(\theta_l, \phi_l)^* Y_n^m(\theta_v, \phi_l) - \delta_{nv}\delta_{mv} \quad (34)$$

$$e(\mathbf{x}) = \sum_{n,v=0}^{\infty} \sum_{m=-n}^n e_{nv}^{mv} \frac{h_n^{(1)}(kR_A)j_n(kR_V)}{k_n^{(1)}(kR_A)j_n(kR_V)} Y_n^m(\theta_v, \phi_v) S_{nm}(R_V), \mathbf{x} \in \partial V \quad (35)$$

여기서  $\Delta S_l = 4\pi R_A^2/L$  관계는 반지름  $R_A$ 를 가지는 구에서 이상적으로 균일하게  $L$ 개의 이차음원을 표본화 하였을 경우이다.

구면기하학 (spherical geometry)에서 Helmholtz 방정식은 분리될 수 있으므로 식 35의 구면베셀함수 (spherical Bessel function)  $j_n(kR_V)$ 를  $j_n(k|\mathbf{x}|)$ ,  $\mathbf{x} \in V$ 로 변경하면 영역  $V$  내부에서의 원하는 음장을 정의할 수 있다. 식 34와 35에서 확인할 수 있는바와 같이  $e_{nv}^{mv}$ 는 이차음원들의 배열환경에 의해 결정되고,  $S_{nm}(R_V)/j_n(kR_V)$ 는 복원음장에 의해서 결정된다.

상기 오차신호  $e_{nv}^{mv}$ 는  $n, v < (L+1)^2$ 일 때 작아진다고 알려져 있고 [5], 이것은 라우드스피커의 개수  $L$ 이 주어진  $n$  또는  $v$  차수의 구면 하모닉스 (spherical harmonics)의 전체 개수보다 커야 한다는 것을 의미한다. 큰 차수들 사용하는 하모닉스에 대해서는  $e_{nv}^{mv}$ 가 증가하는 것은 공간적인 표본화에 의한 영향이라고 할 수 있

다. 음장복원 과정에서 발생하는 오차를 줄이기 위해서는 공간 하모닉스들  $S_{nm}(R_V)$  중에서 차수가  $(L+1)^2$ 인 것들을 제외하는 방법이 있을 수 있으나, 이러한 과정은 공간 하모닉스를 제거함에 따른 새로운 오차를 야기할 수 있다. 다음 수식은 공간 하모닉스의 차수를  $N_T$ 로 한정할 경우에 오차신호에 대한 것이다.

$$e_r(x) = \sum_{n > N_T} \sum_{m = -n}^n Y_n^m(\theta_x, \phi_x) S_{nm}(R_V), x \in V \quad (36)$$

### 4.2. 이차음원들의 위치와 무관한 계산적 음장재현 방법

식 7로 표현된 적분방정식을 표본화된 환경으로 변경하여 임의의 이차음원들의 위치에도 적용하는 방법은 다음과 같다. 먼저 함수  $a(y)$ 와  $p(x)$ 를 표본화 하여 행렬로 표현하면 다음과 같다.

$$\begin{aligned} p &= Sa \\ p &= [p(x_1)\Delta S_1, p(x_2)\Delta S_2, \dots, p(x_K)\Delta S_K] \\ a &= [a(y_1)\Delta S_1, a(y_2)\Delta S_2, \dots, a(y_L)\Delta S_L] \end{aligned} \quad (37)$$

$$S = \begin{bmatrix} G(x_1|y_1), G(x_1|y_2), \dots, G(x_1|y_L) \\ G(x_2|y_1), G(x_2|y_2), \dots, G(x_2|y_L) \\ \dots \\ G(x_K|y_1), G(x_K|y_2), \dots, G(x_K|y_L) \end{bmatrix}$$

여기서  $K$ 는  $p(x)$ 의 표본화 개수이다.

이차음원의 크기를 나타내는  $a$  매트릭스는 다음과 같이 행렬  $S$ 와 이의 의사역행렬 (pseudo inverse matrix)을 이용하여 계산할 수 있다.

$$S^+ = (S^H S)^{-1} S^H \quad (38)$$

여기서  $[ ]^H$ 는 행렬의 에르미트 변환 (Hermitian transform) 의미한다. 따라서  $a$  벡터는 다음과 같이 계산된다.

$$a = S^+ p \quad (39)$$

2장에서 설명한바와 같이 행렬  $S$ 를 SVD 과정을 통해 분해하고 정규화 (regularization) 과정을 추가하면 안정적인 값을 계산할 수 있다.

## V. Experiments

본 논문에서 제안하는 음장재현 알고리즘의 성능을 분석하기 위하여 40개의 라우드스피커로 이루어지는 음장재현 시스템을 무향실내에 설치하여 실험하였다. 모든 라우드스피커들은 2미터의 반지름을 가지는 구면에 일정한 간격으로 배치하였으며, 라우드스피커의 음향학적 중심이 연결점으로부터 0.2미터에 위치하도록 라우드스피커를 설치하였다. 따라서 구면 라우드스피커 어레이의 반지름은 1.8미터 정도가 된다. 그림 2는 전술한 바와 같이 구현되어 영국 사우스햄튼 대학의 무향실에 설치된 라우드스피커 어레이를 나타낸다.

본 실험에 사용된 라우드스피커는 KEF<sup>®</sup> HTS3001이며, 멀티채널 오디오 입출력을 위하여 RME<sup>®</sup>사의 MADI (Multichannel Audio Digital Interface) 장비들을 이용하였다. 또한 그림 3과 같이 29개의 전방향성 (omnidirectional) 마이크로폰들을 7cm 간격으로 배치한 마이크로폰 어레이를 구성하여 라우드스피커 어레이 중심으로 한 4m<sup>2</sup> 영역에서의 복원음장을 실측하였다.

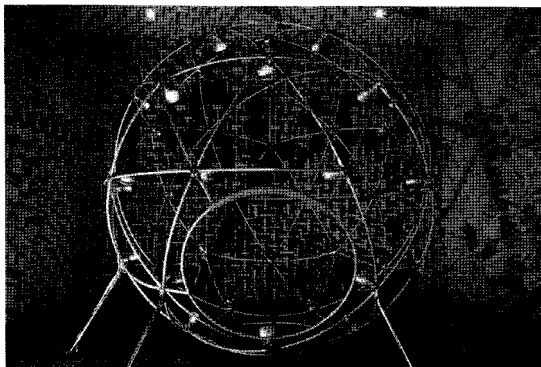


그림 2. 무향실내 설치된 40채널 라우드스피커 어레이  
Fig. 2. Picture of the loudspeaker array with 40 channels in the anechoic chamber.

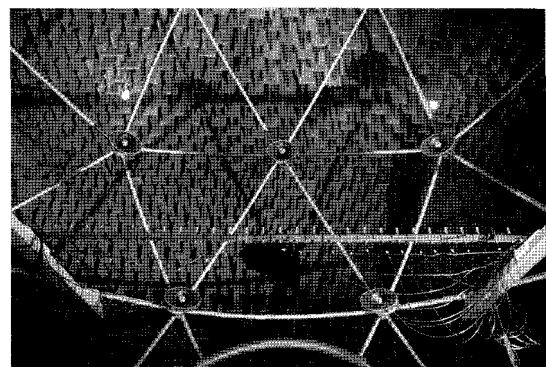


그림 3. 복원음장 실측용 마이크로폰 어레이  
Fig. 3. Picture of the microphone array for measuring the reconstructed sound field.

그림 4는 500Hz 사인파 가상음원을  $r_{FSM} = 3.23m$ ,  $\theta_{FSM} = 77.3^\circ$ ,  $\phi_{FS} = 72^\circ$  에 위치시켰을 때 원음장 (그림 4 (a)), 해석적인 방법으로 복원된 음장 (그림 4 (b)), 계산적 방법으로 복원된 음장 (그림 4 (c))을 나타낸다. 그림 4에서 폐곡선은 다음과 같이 정의된 NRE (normalized reconstruction error)가 20 % 미만인 경우를 나타낸다.

$$NRE(x) = 100 \frac{|p(x) - \hat{p}(x)|^2}{|p(x)|^2} \quad (40)$$

그림 5는 동일한 가상음원의 위치를 이용하면서 900Hz로 주파수를 높였을 때의 결과를 나타낸다. 두 그림에서 확인할 수 있듯이 해석적 방법과 계산적 방법은 유사한 성능을 발휘하며 주파수가 높아질수록 공간적 위신호 (spatial aliasing)에 의하여 성능이 열화 됨을 확인할 수 있다. 또한 음장을 완벽하게 복원하는 최적영역 (sweet spot)도 입력신호의 주파수에 관련됨을 알 수 있었다. 제안된 음장재현 시스템의 경우에는 약 2 m<sup>2</sup> 가량의 영역에서 원음장과 유사한 복원음장을 재현함을 확인하였고, 알고리즘 연구에서 목표로 하였던 다중 청취자

에게 최적의 음장을 제공할 수 있음을 확인하였다.

## VI. Conclusions

본 논문에서는 두 명이상의 다중 청취자에게 최적의 음장을 제공하기 위한 다채널 라우드스피커 어레이를 이용하는 음장재현 알고리즘과 실제 구현 결과를 제시하였다. 실제 40채널 라우드스피커 어레이를 이용한 실험결과 제안된 알고리즘은 1 kHz 이하의 저주파수 대역에서는 만족할만한 성능을 보였으나 주파수가 증가할수록 그 성능이 열화 됨을 확인하였다. 이러한 결과는 라우드스피커 어레이를 구성하는 라우드스피커의 개수가 제한됨으로 인한 공간적 위신호 때문이며 현재 본 실험에 사용된 것보다 더 적은 개수의 라우드스피커를 이용할 경우에는 성능열화가 심화될 것으로 예상된다. 그러나 이러한 현상은 WFS나 HOA 에서도 동일하게 나타나는 현상이며 알고리즘의 개선을 통한 극복보다는 대역별로 상이한 음장재현 기법을 이용하는 것과 같은 새로운 접근방법이

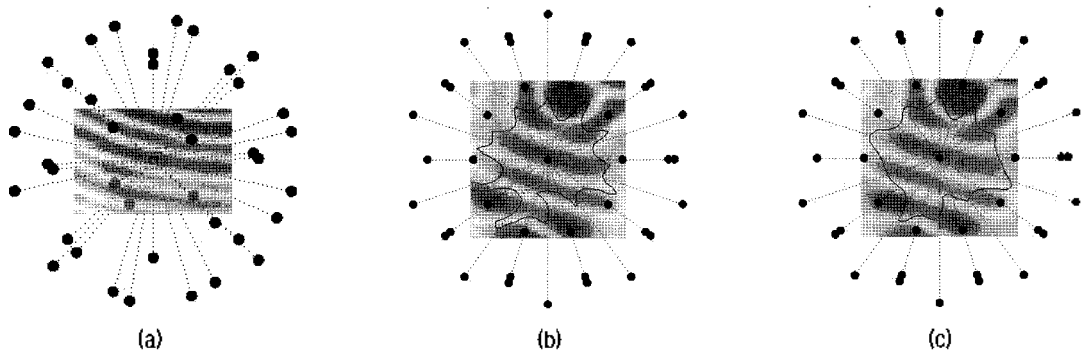


그림 4. 500Hz 가상음원에 대한 원음장 (a), 해석적 방법으로 복원된 음장(b), 계산적 방법으로 복원된 음장(c)들의 측정치  
Fig. 4. Measurement of the original sound field (a) and analytical (b) and numerical (c) reconstructed sound field for 500Hz virtual source.

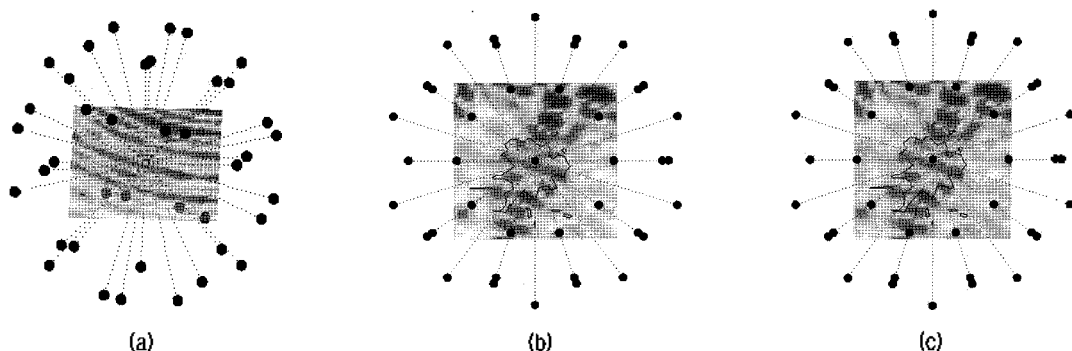


그림 5. 900Hz 가상음원에 대한 원음장 (a), 해석적 방법으로 복원된 음장(b), 계산적 방법으로 복원된 음장(c)들의 측정치  
Fig. 5. Measurement of the original sound field (a) and analytical (b) and numerical (c) reconstructed sound field for 900Hz virtual source.



필요할 것으로 예측된다. 또한, 실제 가정이나 극장 등에서도 적용 가능한 라우드스피커 개수를 이용한 제안된 알고리즘의 성능평가 등이 요구된다.

## 감사의 글

본 연구는 방송통신위원회, 지식경제부 및 한국산입기술평가관리원의 IT 원천기술개발사업의 일환으로 수행한 연구로부터 도출된 것이다 [과제관리번호: 2008-F 011, 과제명: 차세대 DTV 핵심기술 개발].

## 참고 문헌

1. F. Okano, et al., "Ultrahigh-Definition Television System with 4000 Scanning Lines", *2004 NAB BEC Proceedings*, April 2004.
2. A. J. Berkhout, "A holographic approach to acoustic control," *Journal of the Audio Engineering Society*, vol. 36, pp. 977-995, 1988.
3. D. de Vries, E. W. Start, V. G. Valstar, "The wave field synthesis concept applied to sound reinforcement: restrictions and solutions," *The 96th Convention of the Audio Engineering Society*, 1994.
4. M. A. Gerzon, "Periphony: with-height sound reproduction," *Journal of the Audio Engineering Society*, vol. 21, no. 1, pp. 2-10, 1973.
5. J. Daniel, R. Nicol, S. Moureau, "Further investigation of high-order ambisonics and wavefield synthesis for holophonic sound imaging," *The 114th Convention of the Audio Engineering Society*, 2003.
6. M. Poletti, "Three dimensional surround system based on spherical harmonics," *Journal of the Audio Engineering Society*, vol. 53, no. 11, 2005.
7. O. Kirkeby and P. A. Nelson, "Reproduction of plane wave sound fields," *Journal of the Acoustical Society of America* 94, 2992-3000, 1993.
8. V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456-466, 1997.
9. F. M. Fazi and P. A. Nelson, "Application of functional analysis to the sound field reconstruction," *Proceedings of the Institute of Acoustics, the 23rd conference on reproduction sound*, 2007.
10. F. M. Fazi, P. A. Nelson, Jens E. N. Christensen and J. Seo, "Surround system based on three dimensional sound field reconstruction," *The 125th Convention of the Audio Engineering Society*, 2008.

## 저자 약력

### •서정일 (Jeongil Seo)

1994년: 경북대학교 전자공학과 (학사)  
 1996년: 경북대학교 전자공학과 (석사)  
 2005년: 경북대학교 전자공학과 (박사)  
 1998년~2000년: LG반도체 주임연구원  
 2000년~현재: 한국전자통신연구원 방통융합미디어연구부 선임연구원선임연구원  
 ※ 관심분야: 오디오 부호화, 다채널 음장재현 시스템, 3차원 오디오, 디지털방송 시스템, 객체기반 오디오

### •강경옥 (Kyeongok Kang)

1985년: 부산대학교 물리학과 (학사)  
 1988년: 부산대학교 물리학과 (석사)  
 2004년: 한국항공대학교 전자공학과 (박사)  
 2006년: 영국 University of Southampton (방문 연구원)  
 1991년~현재: 한국전자통신연구원 (책임연구원, 미디어응용연구팀장)  
 ※ 주 관심 분야: 오디오 신호처리, 객체 기반 오디오, 3D 오디오, 음성 및 오디오 코덱

### •Filippo M. Fazi

Ph.D. Candidate  
 Institute of Sound and Vibration Research, University of Southampton, U.K.

### •Philip A. Nelson

Professor  
 Institute of Sound and Vibration Research, University of Southampton, U.K.