Theories and signal processing techniques for the implementation of sound ball in space using loudspeaker array

Jung-Woo Choi* and Yang-Hann Kim†
*Korea Advanced Institute of Science and Technology, Daejeon, Korea
E-mail: khepera@kaist.ac.kr Tel: +82-42-350-3244
† Korea Advanced Institute of Science and Technology, Daejeon, Korea
E-mail: yanghannkim@kaist.edu Tel: +82-42-350-3025

Abstract— **It has been known that we can make a certain region of more acoustic energy than others. This can be achieved by utilizing loudspeaker arrays and designing a multichannel filter that effectively controls interferences of sound waves in space and time. The concept of a bright and dark zone[Choi and Kim, "Generation of an acoustically bright zone with an illuminated region using multiple sources,"J. Acoust. Soc. Am., Vol. 111(4), 1695-1700, Apr. 2002] showed that this idea can be realized in practice. Then we attempted to make a "sound ball" that utilizes the concept of bright and dark zone for generating a small spatial region of concentrated sound energy inside. The 32 speaker system which surrounds the zone of interest tried to implement the ball that can be positioned and also allowed to be moved. However, it was also found that the solution based on the bright and dark zones control does not, in strict sense, guarantee an effective radiation of sound from the ball. Understanding this inherent limitation motivated us to design a novel mean to have a sound ball that can radiate effectively. This has to solve the well-known Kirchhoff-Helmholtz equation for the case of which the source or sources are surrounded by an array of speakers. The sound ball is implemented by using a 50-channel spherical loudspeaker array, in which the loudspeakers are positioned on the Lebedev quadrature grid.**

I. INTRODUCTION

A loudspeaker array can generate a desired sound field in a region of interest. Generating a sound field in a region of interest, denoted as *sound manipulation* in this paper, can be used to control sound fields for various purposes. We can focus sound energy on a region, suppress sound energy on another region, and make a virtual sound source.

Mathematically, the fundamental theory of sound manipulation can be found from Kirchhoff-Helmholtz(KH) integral. It says that a sound field at the selected region is determined by the pressure and velocity distribution of the closed contour of the selected region. Therefore, in principle, it is possible to synthesize a desired sound field using transducers or loudspeaker arrays that control pressure and velocity distribution on the surface. The key problem is, however, how we define the desired or target sound field that best describes what we want.

As an example, let us consider making an acoustic source that looks like a sphere, called *sound ball*, in 3D space. The expression 'looks like a sphere' can be described in many ways depending on the objective of manipulation. For example, it can be an acoustic sphere of focused energy or a radiating acoustic sphere that would likely provide the desired impression (Fig. 1).

The focused sphere is to make an acoustic sphere that can make sound energy as much as possible in the volume that we desire. The second type of sound ball refers to a sphere that generates a wave front in the desired region. Since the acoustic sphere can be constructed by using many loudspeakers in space, the problem definition is to find the best possible means; the best method to control the gains and phases of the array speakers, which can satisfy the dissimilar objectives. In this paper, we introduce two representative methods for synthesizing two different kinds of sound balls.

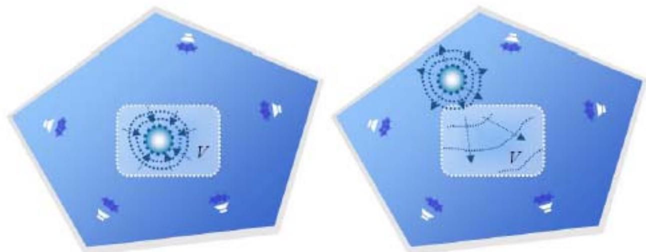

Fig. 1 Two different concepts and definitions of a sound ball

II. FOCUSED SOUND BALL

The sound focusing methods aim to focus the sound energy within a selected region using transducer arrays, and can be regarded as a manipulation counterpart of the microphone beamforming technology. However, the generation of a sound ball requires more general control strategy other than the conventional beamforming developed for the simple loudspeaker and listener geometries. The concept of acoustic brightness and contrast control [1] was proposed to implement

This work was supported in part by the Ministry of Knowledge Economy grant funded by the Korea government (No. 10037244), the BK21 project initiated by the Ministry of Education, Science and Technology, the Converging Research Center Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2011K000665), the project of Agency for Defense Development (ADD-07-07-01), and UTRC(Unmanned technology Research Center) at KAIST(Korea Advanced Institute of Science and Technology), originally funded by DAPA, ADD

zone-selective or regional focusing with enhanced flexibility (Fig. 2).

Fig. 2 The concept of regional sound focusing

There are numerous applications that can utilize the regional focusing concept. The highlights are all related with the zone that we defined. By introducing the zone or field in the formulation, we can get some degree of freedom when we attempt to implement the sound focusing in space. The applications depends, therefore, totally how we define the zone for field. We can have multiple bright and dark zones. The following is one of the examples.

A. Personal audio system

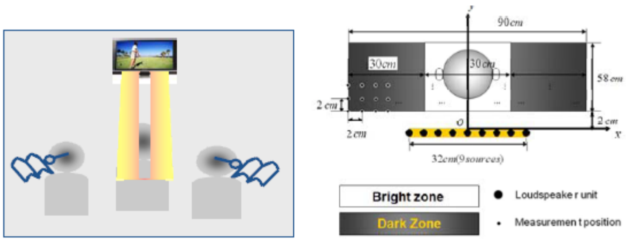

Fig. 3 A personal audio system[2]

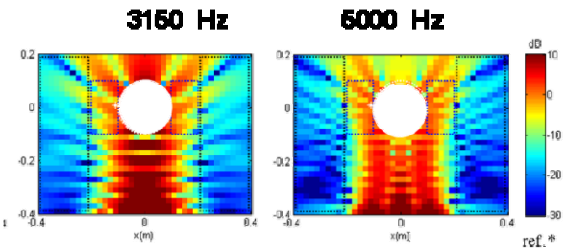

Fig. 4 Focused field generated by acoustic contrast control[3] (with head scattering model)

Consider a situation in which we want to enjoy audio without uncomfortable earphones or headsets and without bothering the user's neighbors (Fig. 3). To realize this situation, the energy of sound waves must be focused on a specific region called the 'listening zone'. The system that can create this kind of listening zone is mentioned as 'sound focused personal audio system' [2]. For this purpose, nine loudspeakers were embedded on a 17 inch monitor and the 'listening zone' and 'un-listening zone' were configured as shown in the Fig. 3. In this experiment, the frequency region of interest was between 800Hz and 5 kHz. Using acoustic contrast control, we were be able to generate the sound field that has over 20dB SPL sound pressure level difference between the 'listening zone(bright zone)' and 'un-listening zone(dark zone)' for 1.5-5 kHz frequency band. Sound fields measured at 3.15 kHz and 5 kHz are depicted in Fig. 4. Even though the SPL difference is noticeable in both frequencies, there are areas to be improved to get practically acceptable personal audio system. For example, we need to reduce the array size for mobile devices applications, reverberation, and etc.

III. RADIATING SOUND BALL

The radiating sound ball refers to the virtual sound source that emits a desired wave front (Fig. 1(b)). The practical application of this sound ball can be the synthesis of a virtual sound source in front of the loudspeaker array (Fig. 5). By definition, this type of sound source is not related to the focused sound ball discussed in Sec. II, because the focusing of sound energy on a point is not directly associated with the wave front emanating from the source.

However, through informal subjective test performed in a listening room (Fig. 15), a certain sense of localization has been reported, when 32 loudspeakers surrounding the listener were manipulated by the focusing solution. This has prompted the question on the relation of the focused and radiating sound ball.

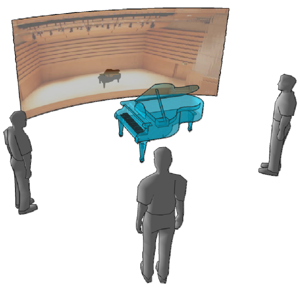

Fig. 5 The concept of a radiating sound ball

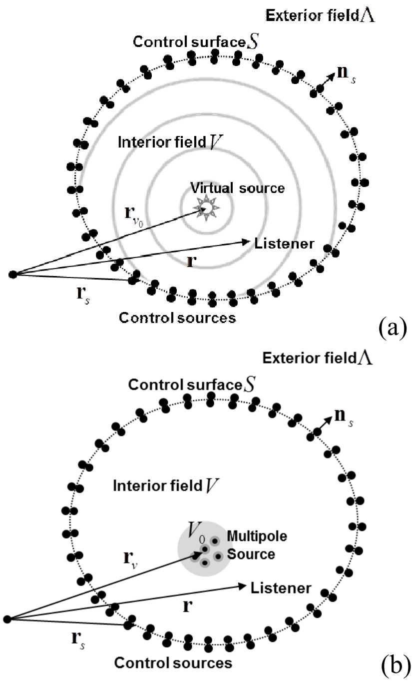

Fig. 6 Schematic illustrating the radiating sound ball problem (a) point virtual source (b) multipole virtual source

A. Integral equation for a focused sound ball

The point focusing of a sound field can be implemented by time-reversing the wavefield impinging on an array(e.g., [4][5]). It also has been demonstrated that the regional focusing solution converges to the time-reversal solution in case of the point focusing [11]. The known form of integral that is mostly related for this problem is the *generalized holographic imaging integral* [6], which is given by

$$\int_S \left[G(\mathbf{r}_{v_0} \mid \mathbf{r}_s)^* \frac{\partial G(\mathbf{r} \mid \mathbf{r}_s)}{\partial n_s} - \frac{\partial G(\mathbf{r}_{v_0} \mid \mathbf{r}_s)^*}{\partial n_s} G(\mathbf{r} \mid \mathbf{r}_s) \right] dS(\mathbf{r}_s) \tag{1}$$
$$= G(\mathbf{r} \mid \mathbf{r}_{v_0}) - G(\mathbf{r} \mid \mathbf{r}_{v_0})^*, \quad \mathbf{r}_{v_0}, \mathbf{r} \in V \setminus S$$

where $\mathbf{n}_s$ is the surface normal vector pointing at the exterior region of the boundary surface S, and $G(\mathbf{r} \mid \mathbf{r}_s)$ represents the Green's function between the control source location $\mathbf{r}_s$ and the virtual source location $\mathbf{r}_{v_0} \in V \setminus S$. The asterisk denotes the complex conjugate operator in frequency domain, which is equivalent to the time reversal in time domain.

Equation (1) implies that the time reversal of the surface pressure and its derivative at the surface S generate two types of sound fields: a converging wave field ($-G(\mathbf{r} \mid \mathbf{r}_{v_0})^*$) and a diverging wave field ($G(\mathbf{r} \mid \mathbf{r}_{v_0})$). Since the time-reversed field can exist for $t < 0$, the alternative field first converges towards to the focal point $\mathbf{r}_v$ and then begins to diverge afterwards. From the viewpoint of sound field reproduction, the diverging wavefront is what we want to reproduce, and the converging wavefront is the artifact to be removed or suppressed. Since the converging wavefront is the time-reverse of the omni-directional radiation and as such, incidents from every direction; its sound is audible in every location between the control surface and the virtual source.

In addition, the integral of (1) requires monopole and its spatial derivative(dipole sources), which can be difficult to implement in practice. A single layer formula that only incorporates monopole sources can be found by applying far-field approximation to (1)[7]. That is,

$$G(\mathbf{r} \mid \mathbf{r}_{v_0}) - G(\mathbf{r} \mid \mathbf{r}_{v_0})^*$$
$$\approx 2ik \int_S \left[G(\mathbf{r}_s \mid \mathbf{r}_{v_0})^* G(\mathbf{r} \mid \mathbf{r}_s) \right] dS(\mathbf{r}_s), \quad \mathbf{r}_{v_0}, \mathbf{r} \in V \setminus S \tag{2}$$

for large spherical surface S of which radius is much greater than the wavelength($\lambda = 2\pi / k$, k : wavenumber). Therefore, the focusing solution for a point virtual source can be related to the integral equation for the reproduction of a sound field. The drawback of this approach is, however, that the time-reversed wave front converging to the location of a virtual source always exists.

B. Solution for the radiating sound ball

Having found that the point focusing can be expressed by the integral equation with the time-reversed driving function, we can find its relation to the optimal solution of the radiating sound ball problem.

The optimal solution for the radiating sound ball should be the one that can suppress the converging wave front G^* of the integral (1). In order to reduce the converging wave front, we have proposed a formula that can separate the converging and diverging wavefronts. It is derived by replacing the point source of Fig. 6(a) with a compact multipole source shown in Fig. 6(b).

Then the propagating direction of the multipole source is reversed in time(TRR; time-reversed radiation). If we describe the original radiation from the multipole source as

$$p_t(\mathbf{r}) = \int_{V_0} G(\mathbf{r} \mid \mathbf{r}_v) q(\mathbf{r}_v) dV(\mathbf{r}_v), \tag{3}$$

then its time-reversed radiation can be written as

$$p_{tr}^*(\mathbf{r}) = \int_{V_0} G(\mathbf{r} \mid \mathbf{r}_v)^* q(\mathbf{r}_v) dV(\mathbf{r}_v). \tag{4}$$

The sound field of TRR on the surface S can be used as a driving function of monopole and dipole sources. The integral equation for the radiating sound ball thus can be derived as

$$p_a(\mathbf{r}) = p_t(\mathbf{r}) - p_{tr}(\mathbf{r})^*$$
$$= \int_S \left[p_{tr}(\mathbf{r}_s)^* \frac{\partial G(\mathbf{r} \mid \mathbf{r}_s)}{\partial n_s} - \frac{\partial p_{tr}(\mathbf{r}_s)^*}{\partial n_s} G(\mathbf{r} \mid \mathbf{r}_s) \right] dS(\mathbf{r}_s).$$
$$(\mathbf{r} \in V \setminus S) \tag{5}$$

where the reproduced field p_a, called *alternative field*, includes two sound fields: the original multipole field p_t and the TRR field p_{tr}^*. This equation is similar to (1), in that both the diverging and converging wave fronts are reproduced. However, for a compact multipole source, it can be shown that the radiating pattern of the TRR field is a mirror image of that of the original multipole field [7]. That is,

$$p_t(\mathbf{r}) \approx b(\mathbf{e}) \frac{e^{ikR}}{4\pi R}, \quad p_{tr}(\mathbf{r})^* \approx b(-\mathbf{e}) \frac{e^{-ikR}}{4\pi R}. \tag{6}$$

where the directional pattern $b(\mathbf{e})$ represents the radiation of a multipole source in the direction $\mathbf{e}$, and R denotes the distance between the listener position $\mathbf{r}$ and the virtual source location $\mathbf{r}_{v_0}$ (Fig. 7).

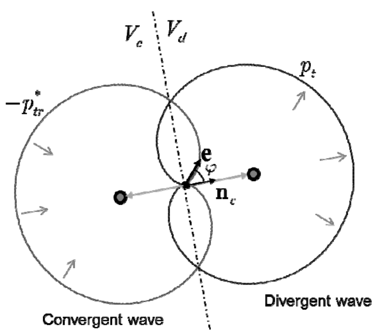

Fig. 7 Opposite directivities of the target field and time-reversed radiations: (V_c : region where the converging wavefront dominates, V_d : region where the diverging wavefront dominates, , $\mathbf{n}_v$: surface normal of the boundary separating V_c and V_d , $\mathbf{e}$: unit directional vector to the listener position, and φ : angle between $\mathbf{n}_v$ and $\mathbf{e}$).

The practical importance of this behavior is that the converging and diverging wave fronts are separated in space, so we can produce a *sweet spot* in which the contribution of the converging wavefront is minimal. Using (6), we can obtain the following approximation:

$$\begin{aligned} p_a(\mathbf{r}) &\approx G(\mathbf{r}\,|\,\mathbf{r}_{v_0})b(\mathbf{e}) - G(\mathbf{r}\,|\,\mathbf{r}_{v_0})^* b(-\mathbf{e}) \\ &\approx \frac{1}{4\pi}\int_S \Big[\left(b(-\mathbf{e})G(\mathbf{r}\,|\,\mathbf{r}_{v_0})^*\right)\partial G(\mathbf{r}\,|\,\mathbf{r}_s)/\partial n_s \\ &\quad -\partial\left(b(-\mathbf{e})G(\mathbf{r}\,|\,\mathbf{r}_{v_0})^*\right)/\partial n_s G(\mathbf{r}\,|\,\mathbf{r}_s)\Big] dS(\mathbf{r}_s) \end{aligned} \tag{7}$$

A single layer version of this integral [7] also has been addressed, which has a form of

$$\begin{aligned} p_a(\mathbf{r}) &\approx \frac{ik}{8\pi^2}\int_{V_0}\left[\int_S\left(\frac{\partial R_v}{\partial n_s}\frac{e^{ik(R_s-R_v)}}{R_v R_s}\right)dS(\mathbf{r}_s)q(\mathbf{r}_v)\right]dV(\mathbf{r}_v) \\ &\approx \frac{ik}{2\pi}\int_S\left(\frac{\partial R_v}{\partial n_s}p_{tr}^*(\mathbf{r}_s)\right)\frac{e^{ikR_s}}{R_s}dS(\mathbf{r}_s), \qquad \mathbf{r}\in V\setminus S \end{aligned} \tag{8}$$

with R_v and R_s being the distances of the virtual source and the listener from a loudspeaker, respectively.

Therefore, the focused and radiating sound balls can be generated by similar integral equations ((2) and (8)). The only difference is that the point multipole is used for the radiating sound balls, instead of the point omni-derectional source for the focused sound ball.

IV. VERIFICATIONS

A. Simulations

To validate the formulas for focused and radiating balls, simulations have been conducted with circular and spherical arrays. Fig. 8 shows the example of the sound field reproduction in the 2D propagation model. Multipoles consisting of up to the 10th order circular harmonics with the half-angular radiation pattern were considered as the target sound field. For the control sources, 64 monopoles and dipoles were arranged on a circle of radius 4λ (λ : wavelength), in order to reduce the artifacts from the discrete sampling of the control surface [8]. The virtual multipole source was located at $\mathbf{r}_{v_0} = (r_{v_0}, \phi_{v_0}) = (2\lambda, \pi/3)$ from the center of the sphere. Figure 8 depicts the real parts of the target cardioid field from the virtual multipole and the reproduced sound field, respectively. The main difference between the target and reproduced fields is the rear-half of the reproduced field. The backward radiation behind the virtual source represents the converging wave front from the loudspeaker array

Next, a three-dimensional case was tested with a spherical loudspeaker array (Fig. 9). For this example, a sphere with radius 2λ was sampled according to the rule of Lebedev quadrature [9]. In all, 170 loudspeakers were employed such that the effect of spatial sampling could be avoided ($N_{spk} \geq (kr_s+1)^2$, [8]) over the entire internal volume V . The resultant field of the 3D case (Fig. 9) also shows the target field reproduced within the finite angular region, with the external field identical to the target field.

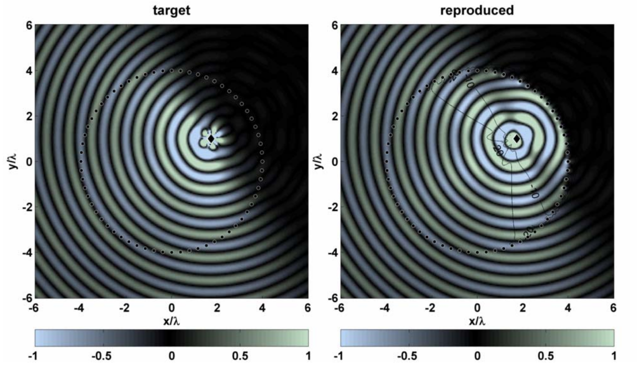

Fig. 8 2D simulation showing the target multipole field (left) and the reproduced sound field (right)

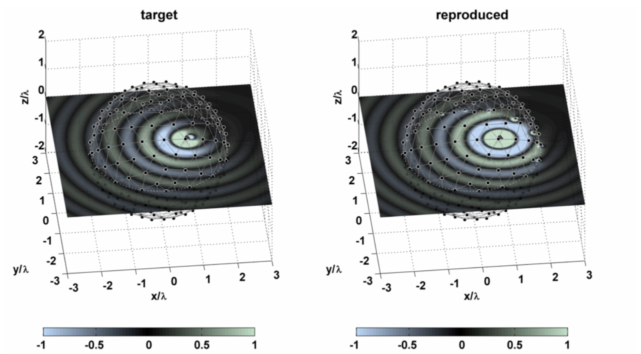

Fig. 9 3D simulation showing the target multipole field (left) and the reproduced sound field (right)

B. *Construction of 32 and 50 channel loudspeaker arrays*

To implement the proposed algorithm, 32 channel loudspeaker array system has been developed(Fig. 10). The specification of the system is available at following webpage: (http://soundmasters.kaist.ac.kr/sound_ball.htm). The 32 channel system has irregular sampling grids, but the transfer function of each loudspeaker was measured and calibrated such that the acoustic contrast control can be applied. The focused sound ball implementaed by the array shows that more than 15dB SPL difference can be achieved(Fig. 11).

As a more idealistic system having regular sampling grids, the 50 channel loudspeaker system also has been constructed (Fig. 12). This array has a complete spherical shape and designed according to the Lebedev quadrature rule, hence it is expected to show more idealistic response close to the theoretical prediction.

The operating program of the system consists of four main units. The communication unit receives raw signals and transmits to signal processing unit. The signal processing unit receives raw signals from the playback unit and control parameters from control region unit and loudspeaker position information from the source control unit. Control region unit and source control unit are developed based on the Max/MSP monitoring object developed by P. Kocher [10]. Fig.13 shows the actual program developed, in which the user can select the manipulation algorithm: Ambisonics, Brightness/Contrast control or Multipole rendering. The position of the point is updated periodically with interval of 10ms, location of the focused source can be renewed and new solution is given to signal processing unit. Multiple points can be generated with the program.

To realize the feedback control, simple remote controller using smart phone is implemented. Using Open Sound Control (OSC) protocol, the controller and the host (PC) transmit control parameter values. Fig. 14 shows the layout of the smart phone application that controls the Spatial Equalizer. Two points are controlled separately using separate layout. The horizontal position of the focused source is controlled with X-Y slider, and elevation is controlled with vertical slider. These objects are typical control object used in midi controls. The smart phone and the host (multichannel receiver) can be connected with Wi-Fi or Bluetooth. Then a listener inside the array system can respond in real-time, by sending a feedback from the smart phone (Fig. 15).

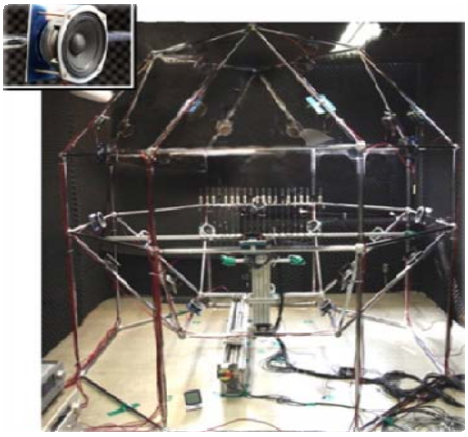

Fig. 10 A 32channel loudspeaker array system installed on a listening room

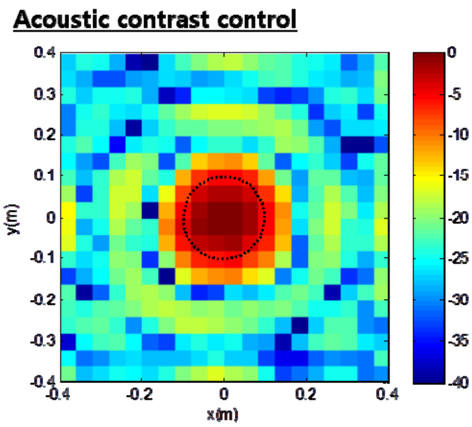

Fig. 11. The focused sound ball implemented by the 32 channel loudspeaker system

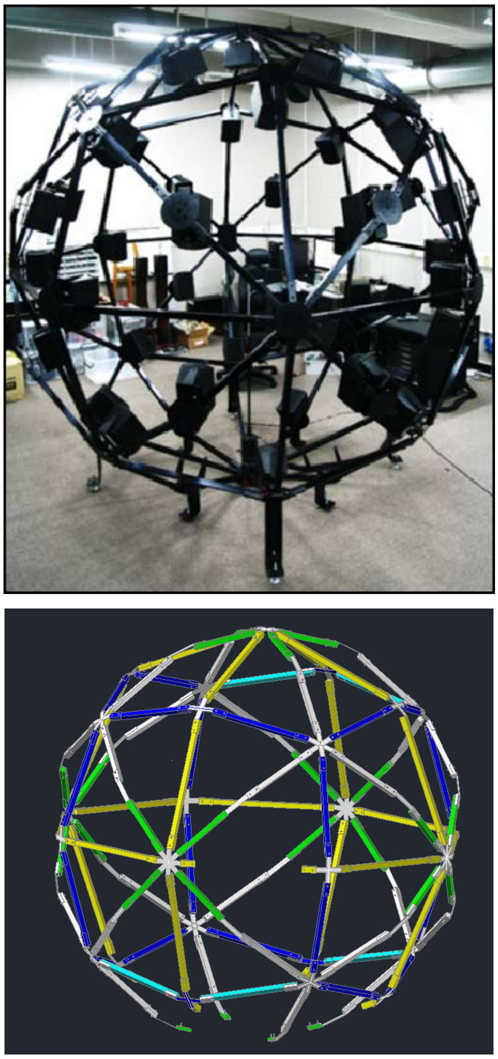

Fig. 12 A 50 channel loudspeaker array designed to realize the Lebedev quadrature grids

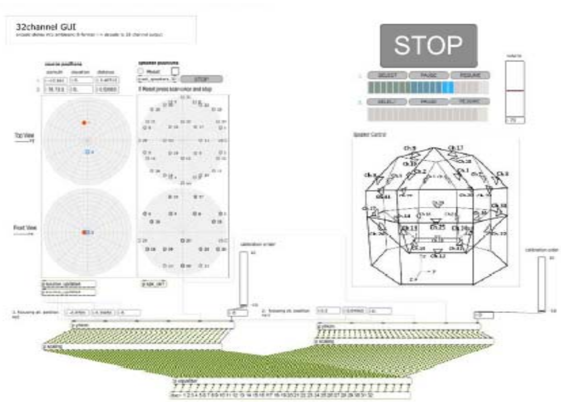

Fig. 13 Max/MSP interface programmed in host PC

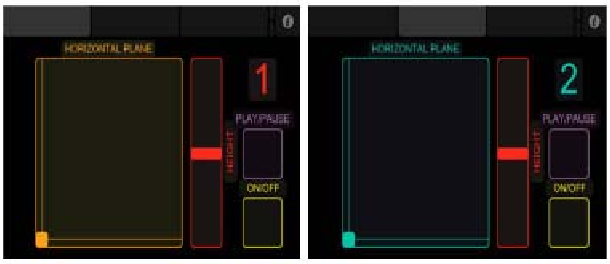

Fig. 14 Mobile controller for remote manipulation of a sound ball

Fig. 15 A listener manipulating the 32 channel loudspeaker array system

V. SUMMARY

We have investigated the relation of the sound focusing and virtual source synthesis techniques. Two methods of sound ball synthesis deal with different physical quantities: acoustic energy and wave front radiation. However, it was shown that both methods can be explained in terms of integral equations, which describe the behavior of the internal sound field in terms of surface source distributions surrounding the internal volume. The main difference is that the synthesis of a virtual sound source requires directive multipole excitation, as compared to the point focusing solution utilizing the time-reversal of the sound field from an omni-directional virtual source. The simulation results and experimental configurations for implementing the proposed methods were presented.

REFERENCES

[1] J.-W. Choi, Y.-H. Kim, “Generation of an acoustically bright zone within an illuminated region,” *J. Acoust. Soc. Am.,* 111(4), pp. 1695-1700 (2002)

[2] J.-H. Chang, C.-H. Lee, J.-Y. Park, and Y.-H. Kim, "A realization of sound focused personal audio system using acoustic contrast control," *J. Acout. Soc. Am.*, 125(4), pp.2091-2097, 2009.

[3] J.-H. Chang, J.-Y. Park, and Y.-H. Kim, "Scattering effect on the sound focused personal audio system," *J. Acout. Soc. Am.*, 125(5), pp.3060-3066, 2009.

[4] D. R. Jackson and D. R. Dowling, “Phase conjugation in underwater acoustics,” *J. Acout. Soc. Am,* 89(1), pp. 171-181, 1991.

[5] M. Fink, “Time Reversal of Ultrasonic Fields – Part I: Basic Principles,” *IEEE Transactions on Ultrasonics,Ferroelectric and Frequency Control,* 39(5), pp. 555-566, 1992.

[6] A. J. Devaney, R. P. Porter, “Holography and the inverse source problem part II: Inhomogeneous media,” *J. Opt. Soc. Am. A*, vol. 2, no. 11, pp. 2006-2011, Nov. 1985.

[7] J. -W. Choi, and Y.-H. Kim, "Integral approach for the reproduction of a virtual sound source surrounded by loudspeaker array" IEEE Trans. Audio Speech Lang. Process., vol. 20, no. 7, pp. 1976-1989, Sept. 2012.

[8] D. B. Ward, T. D. Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Trans. Speech Audio Proc.*, vol.9, pp.697-707, Sep. 2001.

[9] V. I. Lebedev, “Values of the nodes and weights of ninth to seventeenth order gauss-markov quadrature formulae invariant under the octahedron group with inversion,” *USSR Comp. Mathematics and Mathematical Phys.*, vol. 15, no. 1, pp. 44-51, 1975.

[10] J. C Schacher, P. Kocher, "Ambisonics Spatialization Tools for Max/MSP" in Proc. of the 2006 Int. Comp. Music Conf., New Orleans, USA, 2006

[11] J.-W. Choi, *Spatial Manipulation and Implementation of Sound*, Doctoral Thesis, KAIST, pp. 36-99, 2005.