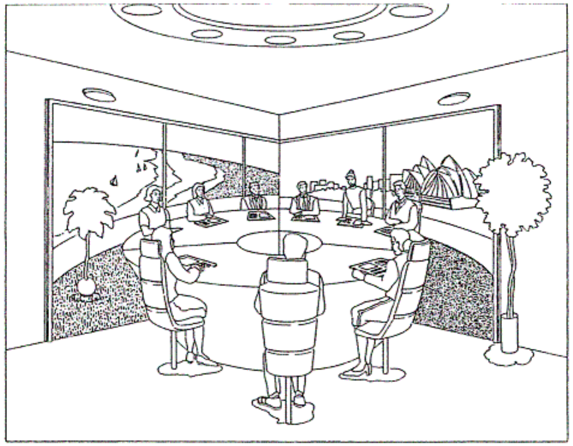
Theoretical Study and Numerical Analysis of 3D Sound Field Reproduction System Based on Directional Microphones and Boundary Surface Control

Toshiyuki Kimura

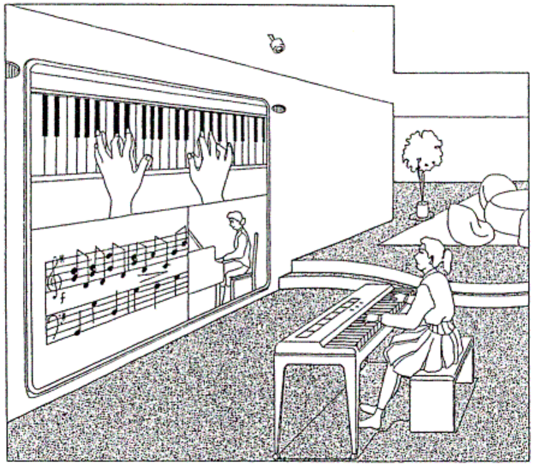
Universal Media Research Center, National Institute of Information and Communications Technology, Japan

3D Sound Field Reproduction System

- Facilitate more realistic experience
- Tele-conference
  - People in different places feel **as if they have a meeting in the same room**

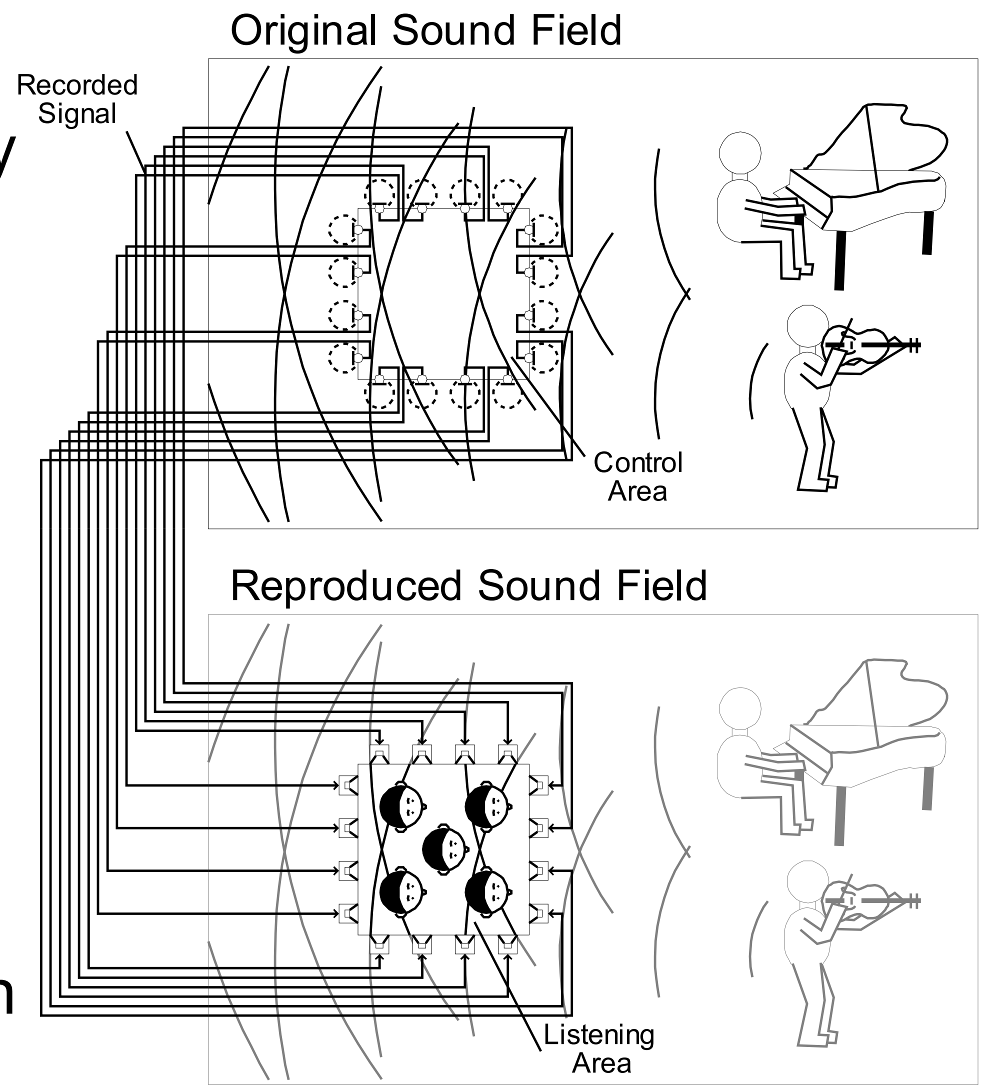


- Tele-ensemble
  - People in different places feel **as if they play a music in the same concert hall**



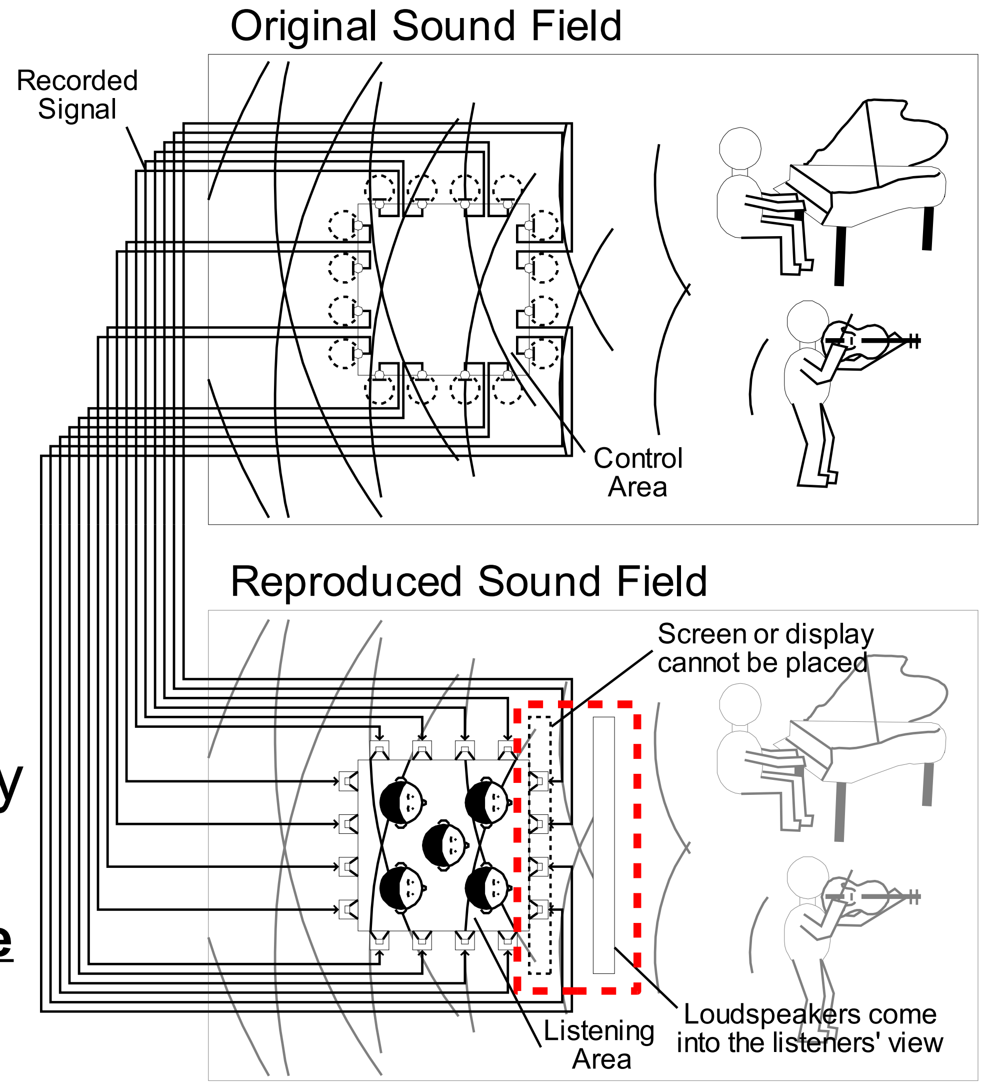
Wave Field Synthesis System

- Original sound field
  - Sound is recorded by directional microphones
- Reproduced sound field
  - Recorded sound is replayed by loudspeakers
  - Wave fronts are accurately reproduced based on Huygens' principle



Problem of Conventional System

- Audio-visual system
  - Screen or display should be placed on or outside the boundary surface
  - On the boundary surface
    - **Screen or display cannot be placed**
  - Outside the boundary surface
    - **Loudspeakers come into the listeners' field of view**



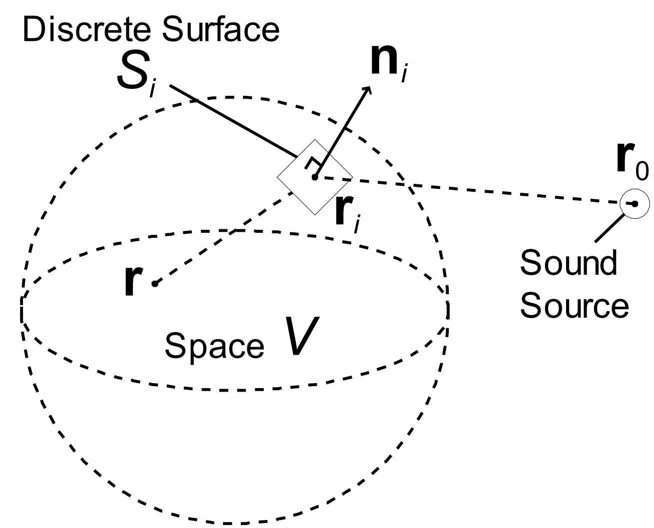
Aim of Study

- A novel 3D sound field reproduction system is proposed
  - Based on directional microphones and boundary surface control
  - Reproduce the 3D sound field without requiring loudspeakers to be placed at the boundary surface
- Theoretical Study
  - The 3D sound field can be accurately reproduced by carrying out inverse filtering based on acoustic transfer functions
- Numerical Analysis
  - A computer simulation is performed

Theory of Conventional System

- 3D sound field can be reproduced in V

$$P(\mathbf{r},\omega)=jk\sum_{i=1}^{M} D_m(\mathbf{r}_0|\mathbf{r}_i)P(\mathbf{r}_i,\omega)$$
$$G(\mathbf{r}_i|\mathbf{r},\omega)\Delta S_i \quad (\mathbf{r}\in V)$$



  - $k(=\omega/c)$: Wave number
  - c: Sound velocity
  - ΔS_i: Area of S_i
  - M: Number of S_i
  - $\mathbf{r}_0$, $\mathbf{r}_i$, $\mathbf{r}$: Position vector of sound source, S_i and V
  - $P(\mathbf{r}_i,\omega)$, $P(\mathbf{r},\omega)$: Sound pressure of $\mathbf{r}_i$ and $\mathbf{r}$
  - $D_m(\mathbf{r}_o|\mathbf{r}_i)$: Directivity of microphones
  - $G(\mathbf{r}_i|\mathbf{r},\omega)$: Acoustic transfer function from $\mathbf{r}_i$ to $\mathbf{r}$

Theory of Proposed System (1)

- 3D sound field is reproduced in V'
  - $\Delta S'_l$, N: Area and number of S'_l

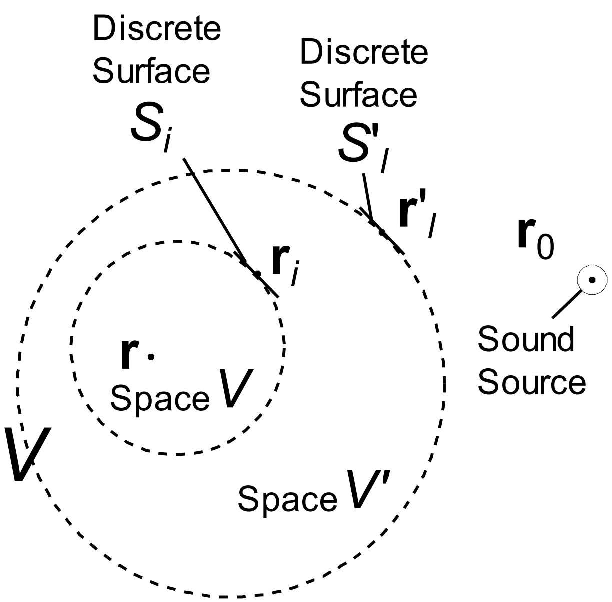
$$P(\mathbf{r},\omega)=jk\sum_{l=1}^{N} D_m(\mathbf{r}_0|\mathbf{r}'_l)\,P(\mathbf{r}'_l,\omega)\,G(\mathbf{r}'_l|\mathbf{r},\omega)\,\Delta S'_l \quad (\mathbf{r}\in V')$$

- $\mathbf{r}_i$ is always in V'

$$P(\mathbf{r}_i,\omega)=jk\sum_{l=1}^{N} D_m(\mathbf{r}_0|\mathbf{r}'_l)\,P(\mathbf{r}'_l,\omega)\,G(\mathbf{r}'_l|\mathbf{r}_i,\omega)\,\Delta S'_l \quad (\mathbf{r}_i\in V')$$

- 3D sound field is reproduced in V
  - There are sound sources in $\mathbf{r}'_l$

$$P(\mathbf{r},\omega)=jk\sum_{i=1}^{M} D_m(\mathbf{r}'_l|\mathbf{r}_i)\,P(\mathbf{r}_i,\omega)\,G(\mathbf{r}_i|\mathbf{r},\omega)\,\Delta S_i \quad (\mathbf{r}\in V)$$



Theory of Proposed System (2)

- 3D sound field is reproduced in V'
  - $\Delta S'_l$, N: Area and number of S'_l

$$P(\mathbf{r},\omega)=jk\sum_{l=1}^{N} D_m(\mathbf{r}_0|\mathbf{r}'_l)P(\mathbf{r}'_l,\omega)\underline{G(\mathbf{r}'_l|\mathbf{r},\omega)}\Delta S'_l \quad (\mathbf{r}\in V')$$

$$P(\mathbf{r},\omega)=jk\sum_{l=1}^{N} D_m(\mathbf{r}_0|\mathbf{r}'_l)P(\mathbf{r}'_l,\omega)\left\{\underline{jk\sum_{i=1}^{M} D_m(\mathbf{r}'_l|\mathbf{r}_i)G(\mathbf{r}'_l|\mathbf{r}_i,\omega)G(\mathbf{r}_i|\mathbf{r},\omega)\Delta S_i}\right\}\Delta S'_l \quad (\mathbf{r}\in V,\mathbf{r}_i\in V')$$

$$G(\mathbf{r}'_l|\mathbf{r},\omega)=jk\sum_{i=1}^{M} D_m(\mathbf{r}'_l|\mathbf{r}_i)G(\mathbf{r}'_l|\mathbf{r}_i,\omega)G(\mathbf{r}_i|\mathbf{r},\omega)\Delta S_i \quad (\mathbf{r}\in V,\mathbf{r}_i\in V')$$

Theory of Proposed System (3)

- System via *M*-input *N*-output filters $H_{li}(\omega)$

$$\underbrace{P'(\mathbf{r}'_l,\omega)}_{\text{Filtered signal}}=\sum_{i=1}^{M} H_{li}(\omega)\underbrace{D_m(\mathbf{r}_0|\mathbf{r}_i)P(\mathbf{r}_i,\omega)}_{\text{Recorded signal}}$$

- Reproduced 3D sound field $P'(\mathbf{r},\omega)$

$$P'(\mathbf{r},\omega)=\sum_{l=1}^{N} P'(\mathbf{r}'_l,\omega)G(\mathbf{r}'_l|\mathbf{r},\omega)\Delta S'_l$$

$$=\sum_{i=1}^{M} D_m(\mathbf{r}_0|\mathbf{r}_i)P(\mathbf{r}_i,\omega)\left\{\sum_{l=1}^{N} H_{li}(\omega)\boxed{G(\mathbf{r}'_l|\mathbf{r},\omega)}\Delta S'_l\right\}$$

$$\boxed{G(\mathbf{r}'_l|\mathbf{r},\omega)}=jk\sum_{i=1}^{M} D_m(\mathbf{r}'_l|\mathbf{r}_i)G(\mathbf{r}'_l|\mathbf{r}_i,\omega)G(\mathbf{r}_i|\mathbf{r},\omega)\Delta S_i$$

$$(\mathbf{r}\in V,\mathbf{r}_i\in V')$$

Theory of Proposed System (4)

- $H_{li}(\omega)$ is defined as inverse filters

$$\sum_{l=1}^{N} H_{li}(\omega) D_m(\mathbf{r}'_l|\mathbf{r}_n) G(\mathbf{r}'_l|\mathbf{r}_n,\omega) \Delta S'_l = \begin{cases} 1 & (n=i) \\ 0 & (n \neq i) \end{cases}$$

- Reproduced 3D sound field $P'(\mathbf{r},\omega)$

$$P'(\mathbf{r},\omega) = jk \sum_{i=1}^{M} D_m(\mathbf{r}_0|\mathbf{r}_i) P(\mathbf{r}_i,\omega) \left[\sum_{n=1}^{M} G(\mathbf{r}_n|\mathbf{r},\omega) \left\{ \sum_{l=1}^{N} H_{li}(\omega) D_m(\mathbf{r}'_l|\mathbf{r}_n) G(\mathbf{r}'_l|\mathbf{r}_n,\omega) \Delta S'_l \right\} \Delta S_n \right]$$

$$= jk \sum_{i=1}^{M} D_m(\mathbf{r}_0|\mathbf{r}_i) P(\mathbf{r}_i,\omega) G(\mathbf{r}_i|\mathbf{r},\omega) \Delta S_i = P(\mathbf{r},\omega)$$

3D sound field can be reproduced in space *V*

$$(\mathbf{r} \in V, \mathbf{r}_i \in V')$$

Diagram of Proposed System (1)

- Directional microphones are placed and sound is recorded
  - Directional microphones are directed toward the outside of the control area

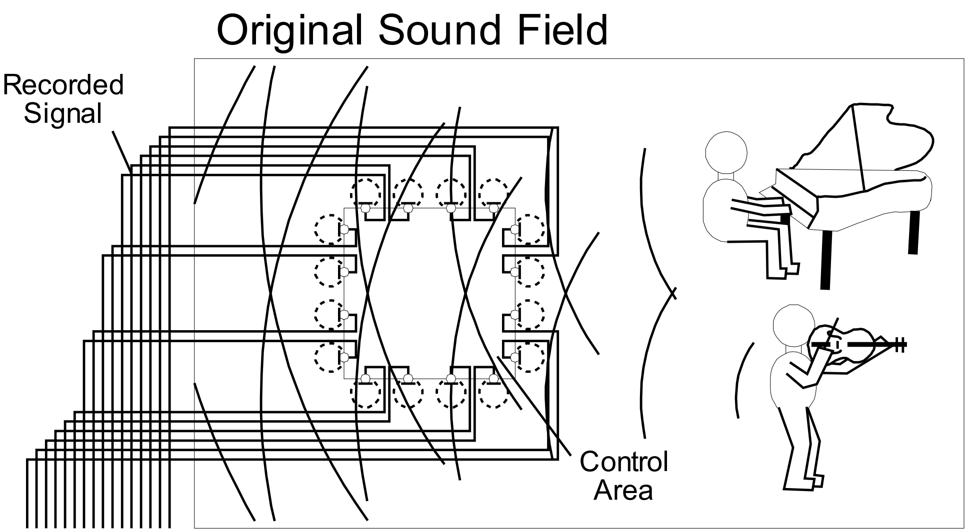


Diagram of Proposed System (2)

- Directional microphones and loudspeakers are placed
  - The position and direction of directional microphones are the same as in the recording
  - Loudspeaker array envelops microphone array

Reproduced Sound Field

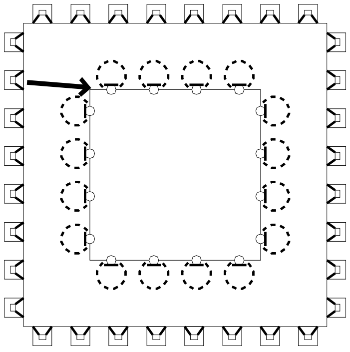


Diagram of Proposed System (3)

- Acoustic transfer functions are measured and the inverse filters are calculated

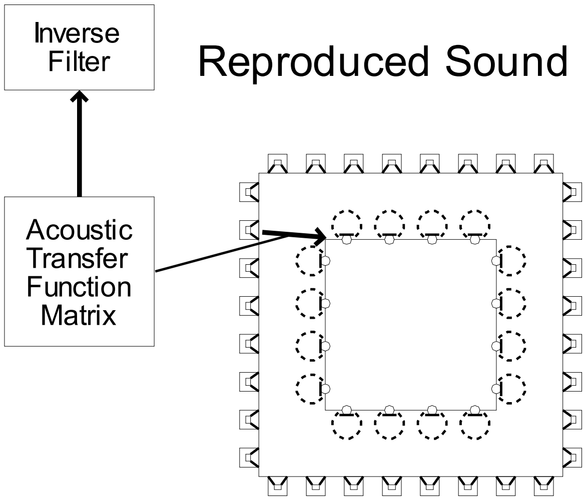


Diagram of Proposed System (4)

- Recorded signals are filtered by the inverse filters, and the filtered signals are played by loudspeakers
  - Listeners feel as if they are listening to the sound in the original sound field

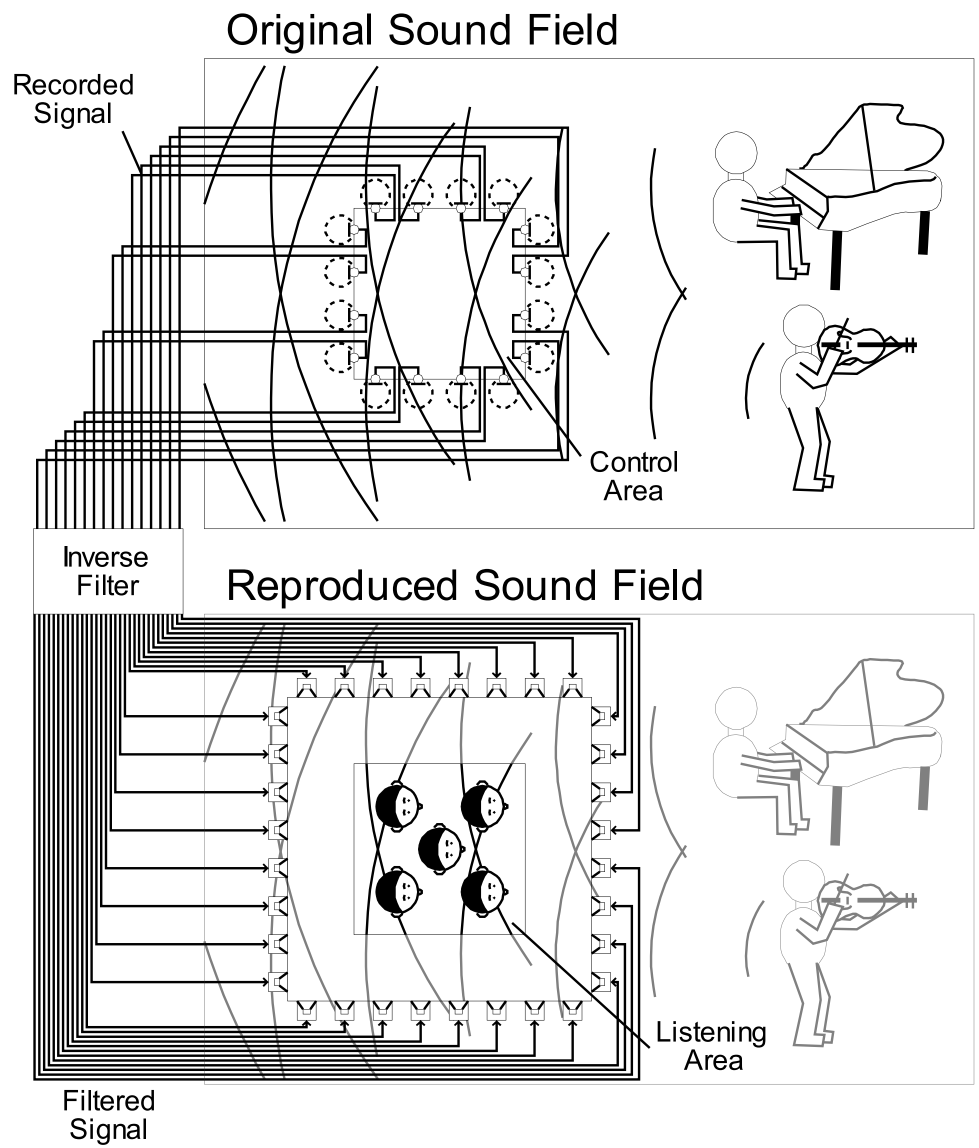
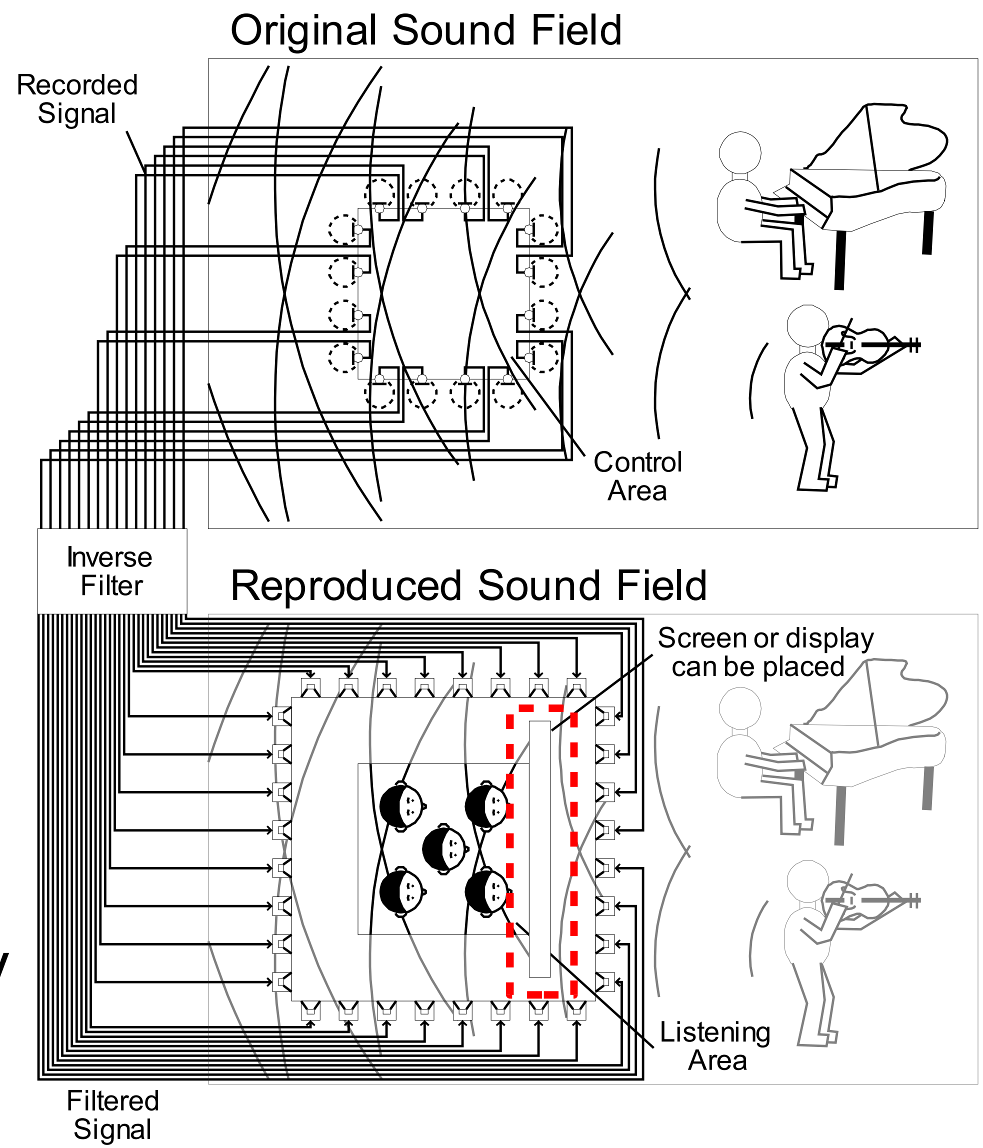


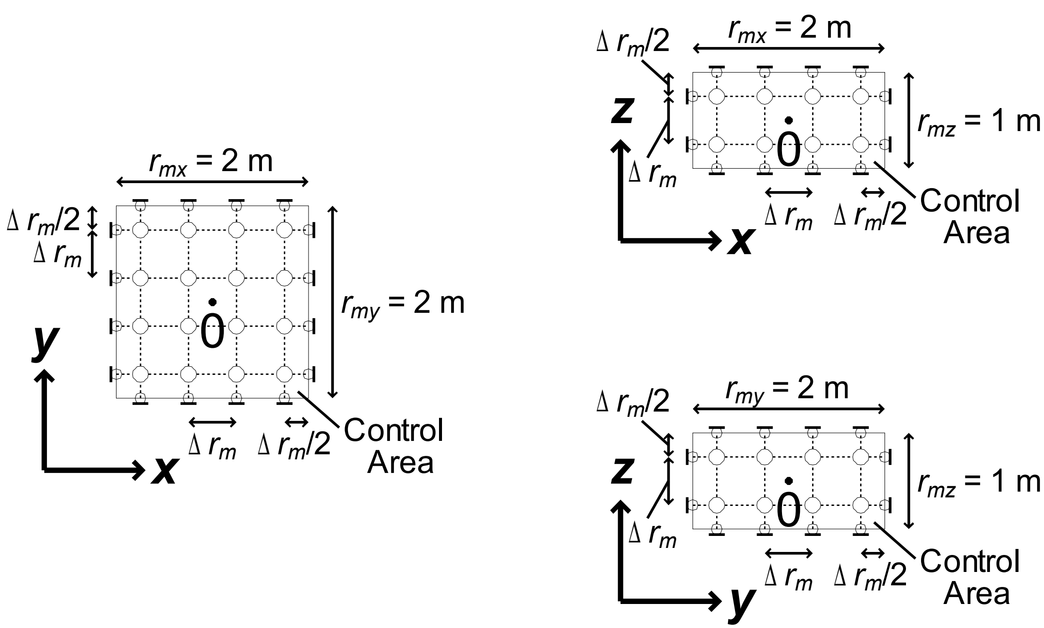
Diagram of Proposed System (5)

- Directional microphones do not need to be placed during playing
- It is possible to construct an audio-visual system
  - Screen or display of the visual system can be placed on or outside the boundary surface



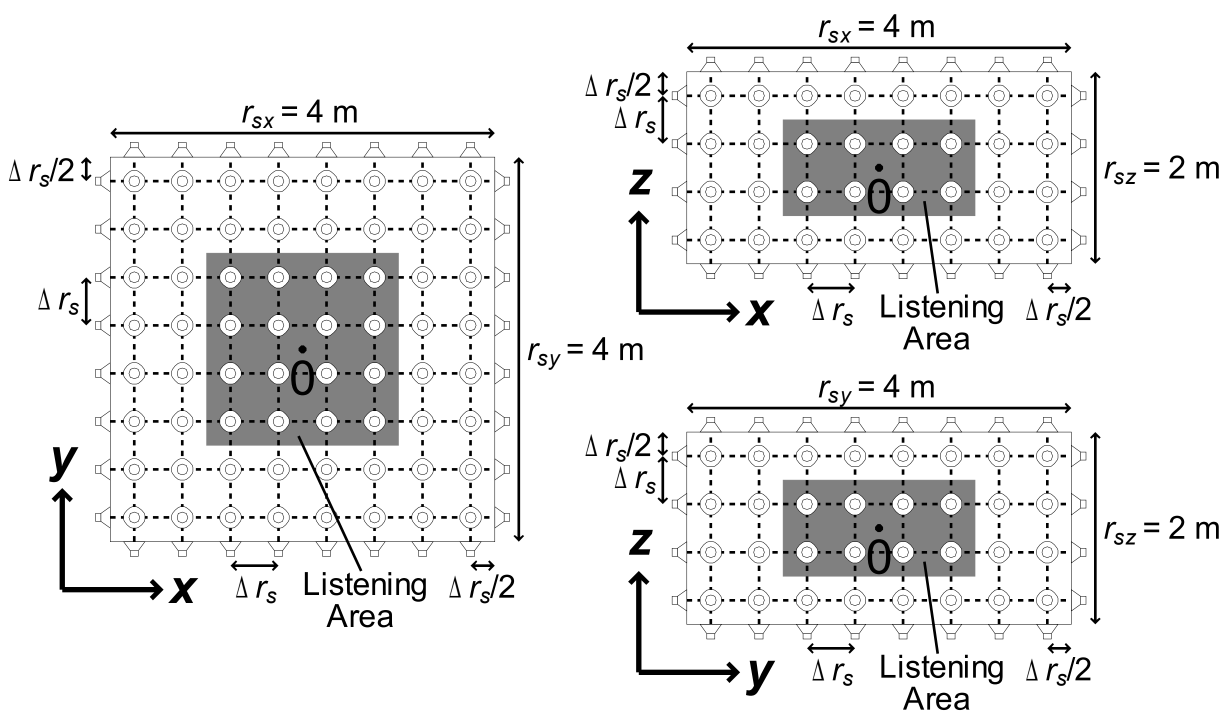
Original Sound Field

- Control area
  - Size...2 m width, 2 m depth, 1 m height
- Microphone array (placed on 6 planes)
  - Size...2 m width, 2 m depth, 1 m height



Reproduced Sound Field

- Listening area (same size as control area)
  - Size…2 m width, 2 m depth, 1 m height
- Loudspeaker array (placed on 6 planes)
  - Size…4 m width, 4 m depth, 2 m height



Original Sound Field Synthesis

- Sound source signal $s(t)$
  - Sinusoidal signal (frequency f, amplitude A)

$$s(t) = A \sin 2\pi f t$$

- Sound pressure in the original sound field
  - $\mathbf{r}_0$: Position vector of sound source
  - $\mathbf{r}$: Position vector of point in control area
  - c: Sound velocity

$$p_o(\mathbf{r}, t) = \frac{1}{|\mathbf{r} - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{r} - \mathbf{r}_0|}{c}\right)$$

$$= \frac{A}{|\mathbf{r} - \mathbf{r}_0|} \sin\left\{2\pi f \left(t - \frac{|\mathbf{r} - \mathbf{r}_0|}{c}\right)\right\}$$

Reproduced Sound Field Synthesis (1)

- Recorded signal $x_i(t)$ (i=1...M)
  - $\mathbf{r}_i$: Position vector of ith directional microphone
  - $D_m(\mathbf{r}_o|\mathbf{r}_i)$: Directivity of ith directional microphone
  - M: Number of directional microphones

$$x_i(t) = \frac{D_m(\mathbf{r}_0|\mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{r}_i - \mathbf{r}_0|}{c}\right)$$

$$= \frac{D_m(\mathbf{r}_0|\mathbf{r}_i) A}{|\mathbf{r}_i - \mathbf{r}_0|} \sin\left\{2\pi f\left(t - \frac{|\mathbf{r}_i - \mathbf{r}_0|}{c}\right)\right\}$$

Reproduced Sound Field Synthesis (2)

- Filtered signal $y_l(t)$ (l=1...N)
  - Ξ_{li}(=$|H_{li}(\omega)|$): Absolute value of inverse filters
  - Θ_{li}(=arg$H_{li}(\omega)$): Angle of inverse filters
  - M: Number of directional microphones
  - N: Number of loudspeakers

$$y_l(t) = \sum_{i=1}^{M} \Xi_{li} x_i \left(t - \frac{\Theta_{li}}{2\pi f} \right)$$

$$= \sum_{i=1}^{M} \frac{\Xi_{li} D_m(\mathbf{r}_0 | \mathbf{r}_i) A}{|\mathbf{r}_i - \mathbf{r}_0|} \sin\left\{ 2\pi f \left(t - \frac{|\mathbf{r}_i - \mathbf{r}_0|}{c} \right) - \Theta_{li} \right\}$$

Reproduced Sound Field Synthesis (3)

- Sound pressure in reproduced sound field
  - $\mathbf{r}'_l$: Position vector of lth loudspeaker
  - N: Number of loudspeakers

$$p_p(\mathbf{r},t)=\sum_{l=1}^{N}\frac{1}{|\mathbf{r}-\mathbf{r}'_l|}y_l\left(t-\frac{|\mathbf{r}-\mathbf{r}'_l|}{c}\right)$$

$$=\sum_{l=1}^{N}\sum_{i=1}^{M}\frac{\Xi_{li}D_m(\mathbf{r}_0|\mathbf{r}_i)A}{|\mathbf{r}-\mathbf{r}'_l||\mathbf{r}_i-\mathbf{r}_0|}\sin\left\{2\pi f\left(t-\frac{|\mathbf{r}-\mathbf{r}'_l|+|\mathbf{r}_i-\mathbf{r}_0|}{c}\right)-\Theta_{li}\right\}$$

Parametric Condition

Amplitude (A)	1
Frequency (f)	63, 125, 250, 500, 1000 Hz
Distance (d)	2, 10, 50 m
Direction vector ($\mathbf{u}$)	$(1,0,0)^T(1/\sqrt{2},1/\sqrt{2},0)^T(2/3,2/3,1/3)^T$
Sound velocity (c)	340 m/s
Microphone number (M)	576
Microphone interval (Δr_m)	0.1667 m
Loudspeaker number (N)	2304
Loudspeaker interval (Δr_s)	0.1667 m

$$D_m(\mathbf{r}_0|\mathbf{r}_i)=\begin{cases}\cos\theta_{im} & (\theta_{im}\le 90^\circ)\\ 0 & (\theta_{im}>90^\circ)\end{cases}\qquad \cos\theta_{im}=\frac{\mathbf{n}_{im}\cdot(\mathbf{r}_0-\mathbf{r}_i)}{|\mathbf{n}_{im}||\mathbf{r}_0-\mathbf{r}_i|}$$

– $\mathbf{n}_{im}$: Directional vector of ith directional microphone

Synthesis of Conventional System

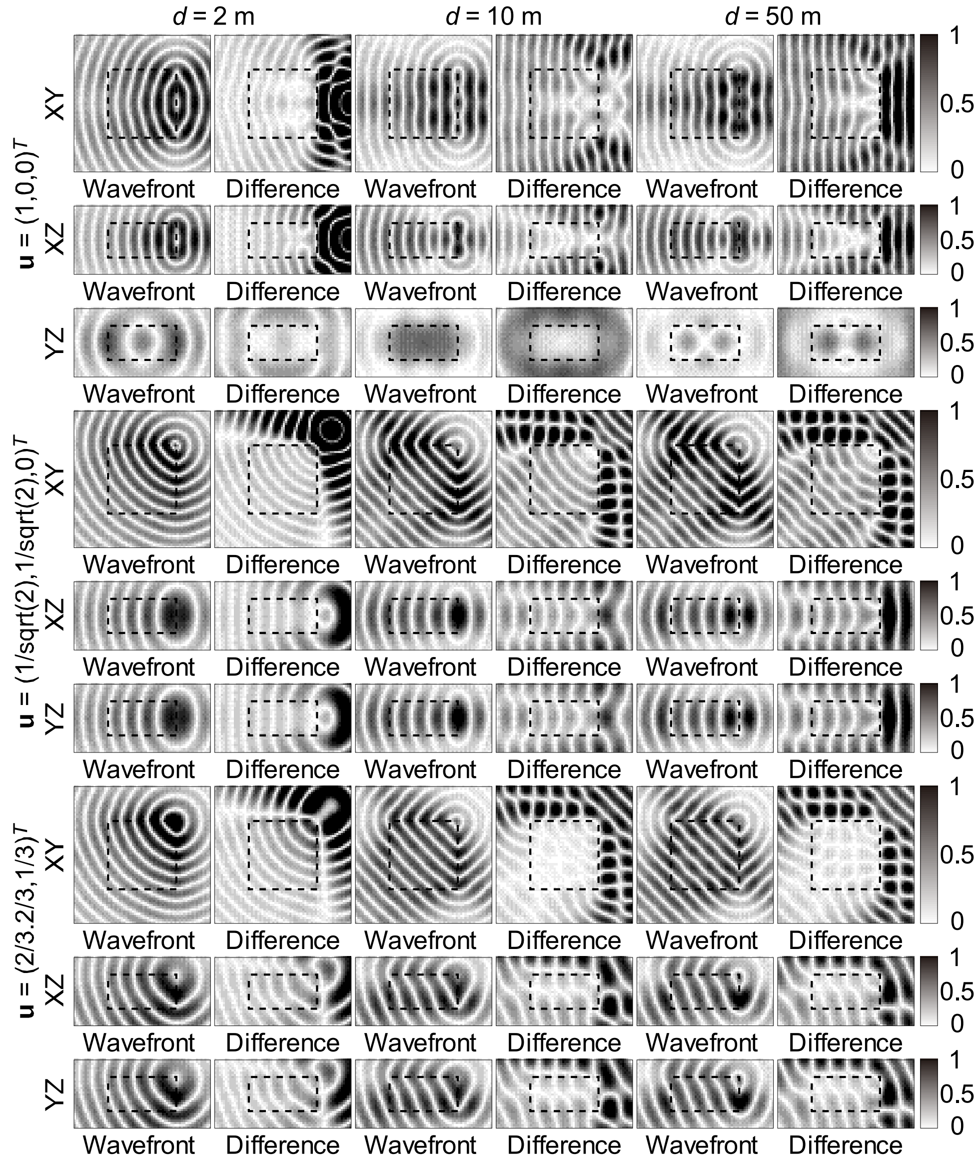
- Listening area (same size as control area)
  - Size...2 m width, 2 m depth, 1 m height
    - Same size as proposed system
- Loudspeaker array (placed on 6 planes)
  - Size...2 m width, 2 m depth, 1 m height
    - Half scale size as proposed system

$$p_c(\mathbf{r},t)=\sum_{i=1}^{M}\frac{1}{|\mathbf{r}-\mathbf{r}_i|}x_i\left(t-\frac{|\mathbf{r}-\mathbf{r}_i|}{c}\right)$$

$$=\sum_{i=1}^{M}\frac{D_m(\mathbf{r}_0|\mathbf{r}_i)A}{|\mathbf{r}-\mathbf{r}_i||\mathbf{r}_i-\mathbf{r}_0|}\sin\left\{2\pi f\left(t-\frac{|\mathbf{r}-\mathbf{r}_i|+|\mathbf{r}_i-\mathbf{r}_0|}{c}\right)\right\}$$

Results of Conventional System

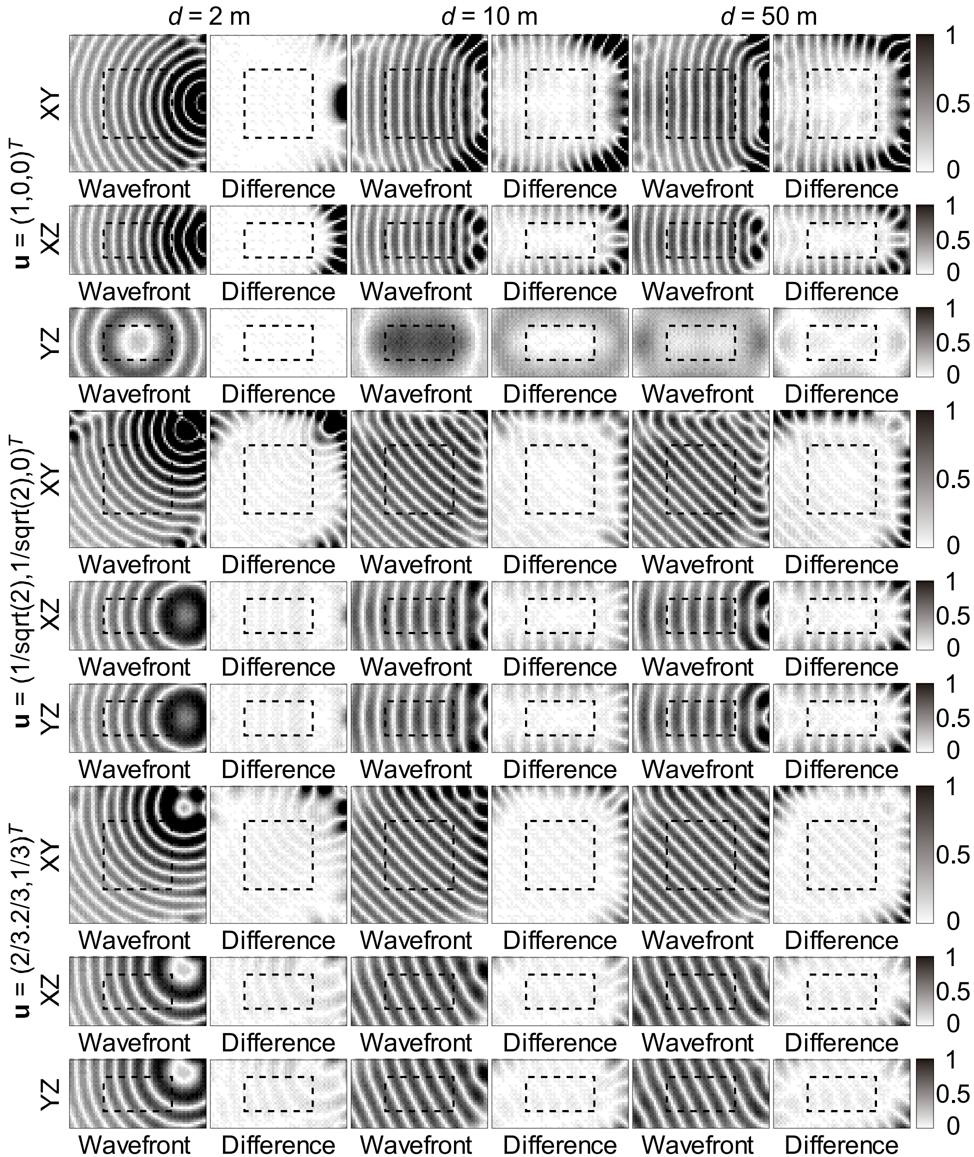
- Wave fronts are not accurately reproduced
  - Differences are not white



Results of Proposed System

- Wave fronts are accurately reproduced
  - Differences are white

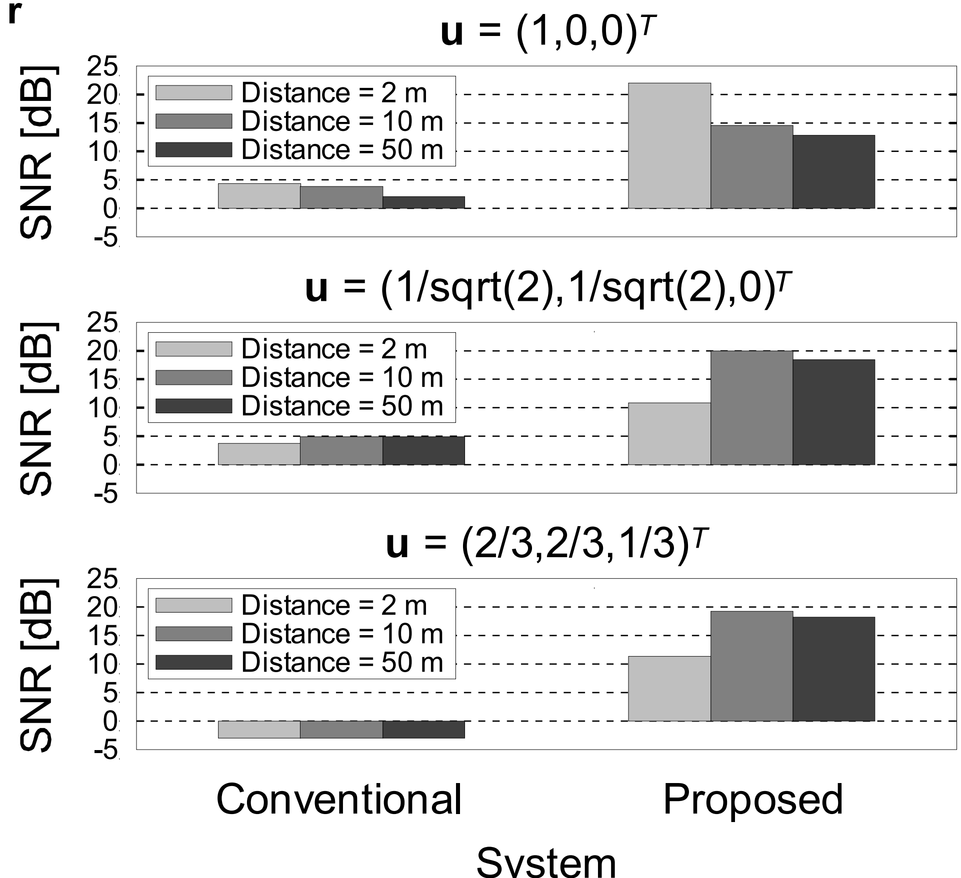
Proposed system can reproduce wave fronts more accurately than conventional system



Results of Signal-to-Noise Ratios

$$\mathrm{SNR} = 10\log_{10}\frac{\sum_{f}\sum_{\mathbf{r}}\left\{p_o(\mathbf{r},0)\right\}^2}{\sum_{f}\sum_{\mathbf{r}}\left\{p_{c(p)}(\mathbf{r},0)-p_o(\mathbf{r},0)\right\}^2}$$

- Conventional system
  - Lesser than 5 dB
- Proposed system
  - Greater than 10 dB
- Wave fronts can be accurately reproduced in the proposed system



Conclusion

- 3D sound field reproduction system based on directional microphones and boundary surface control is proposed
  - Constructed by using the inverse filters in the conventional system
  - Reproduce a 3D sound field in a listening area even if loudspeakers are not placed on the boundary surface
- Computer simulation
  - The proposed system can reproduce wave fronts more accurately than the conventional system

Future Works

- Further study for an actual room
  - The effect of reflected and reverberant sounds in a room
  - The variation of the acoustic transfer functions by the placement of the screen or display
- The reduction of the number of microphones and loudspeakers
  - The evaluation of the performance of the constructed real system by listening tests