Comparative Performance Evaluation of Near 3D Sound Field Reproduction System with Directional Loudspeakers and Wave Field Synthesis

Toshiyuki Kimura[1], Yoko Yamakata[2], Michiaki Katsumoto[1], Takuma Okamoto[3], Satoshi Yairi[4], Yukio Iwaya[3], Yo-iti Suzuki[3]

[1]NICT, [2]Kyoto Univ., [3]Tohoku Univ., [4]Sendai NCT

Outline

- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

Outline

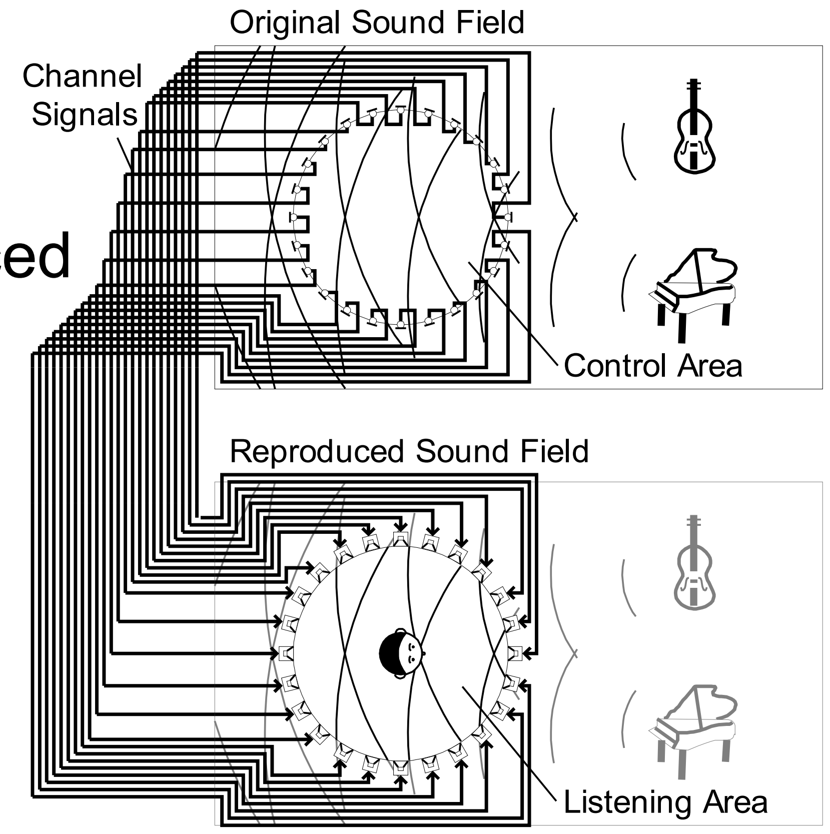
- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

Ultra-Realistic Communication

- Future 3D television
  - 3D video and audio appear in a 3D space
  - People view an object anywhere in its vicinity
    - Without glasses
  - 3D sound field reproduction systems without headphones must be developed

Wave Field Synthesis (WFS) System

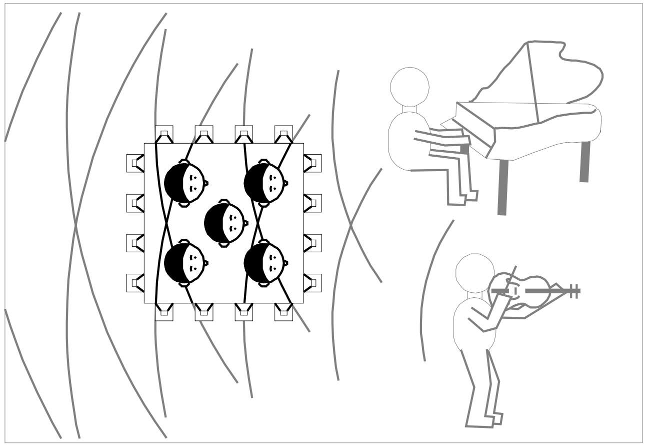
- Original sound field
  - Sound is recorded by the microphone array
- Reproduced sound field
  - Recorded sound is played by the loudspeaker array
  - 3D sound field is reproduced by Kirchhoff-Helmholtz integral equation
- Multiple listeners can listen to a sound without headphones



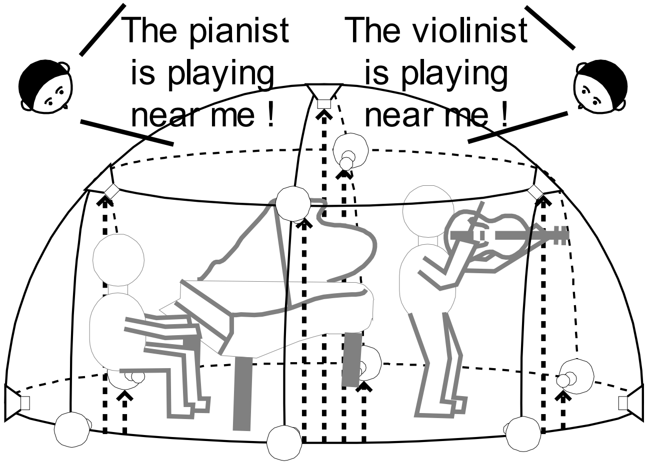
WFS System for Future 3DTV

- Conventional system
  - Loudspeakers are placed around the listeners
  - Sound scene is reproduced
- Proposed system
  - Loudspeakers are placed around sound sources
  - Radiated sound field is reproduced

Conventional System



Proposed System



Aim of Study

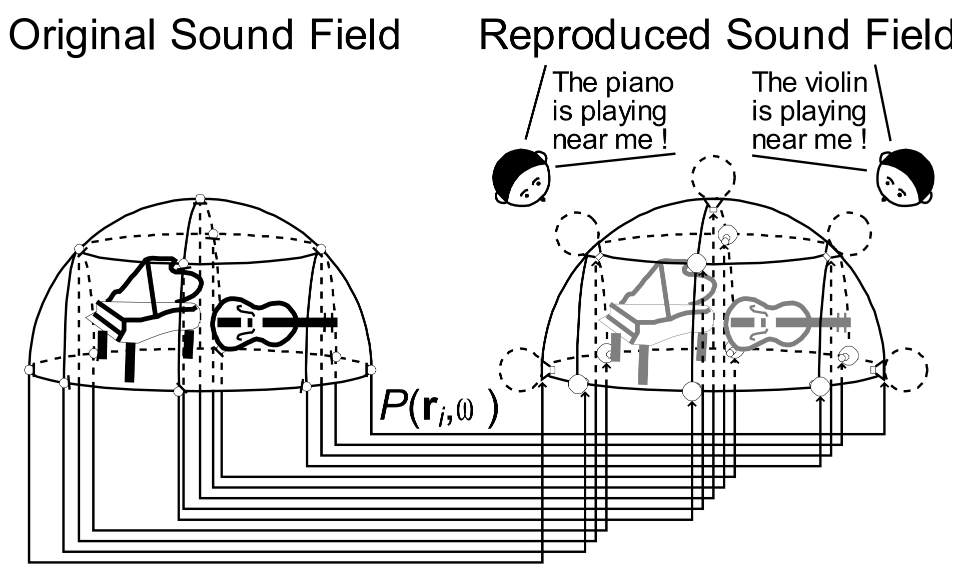
- We have proposed and developed the WFS system for future 3DTV
  - Near 3D sound field reproduction system using directional loudspeakers and wave field synthesis
- Evaluation of the performance of the sound image localization in the developed system
  - Estimation of the position of sound images
    - Computer simulation
    - Acoustical measurement

Outline

- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

Diagram of Proposed System

- Original sound field
  - Sound is recorded by microphones placed around sound sources
- Reproduced sound field
  - Recorded sound is played by directional loudspeakers
  - 3D sound field is reproduced
- Listeners feel that sound sources are playing in the array



Outline

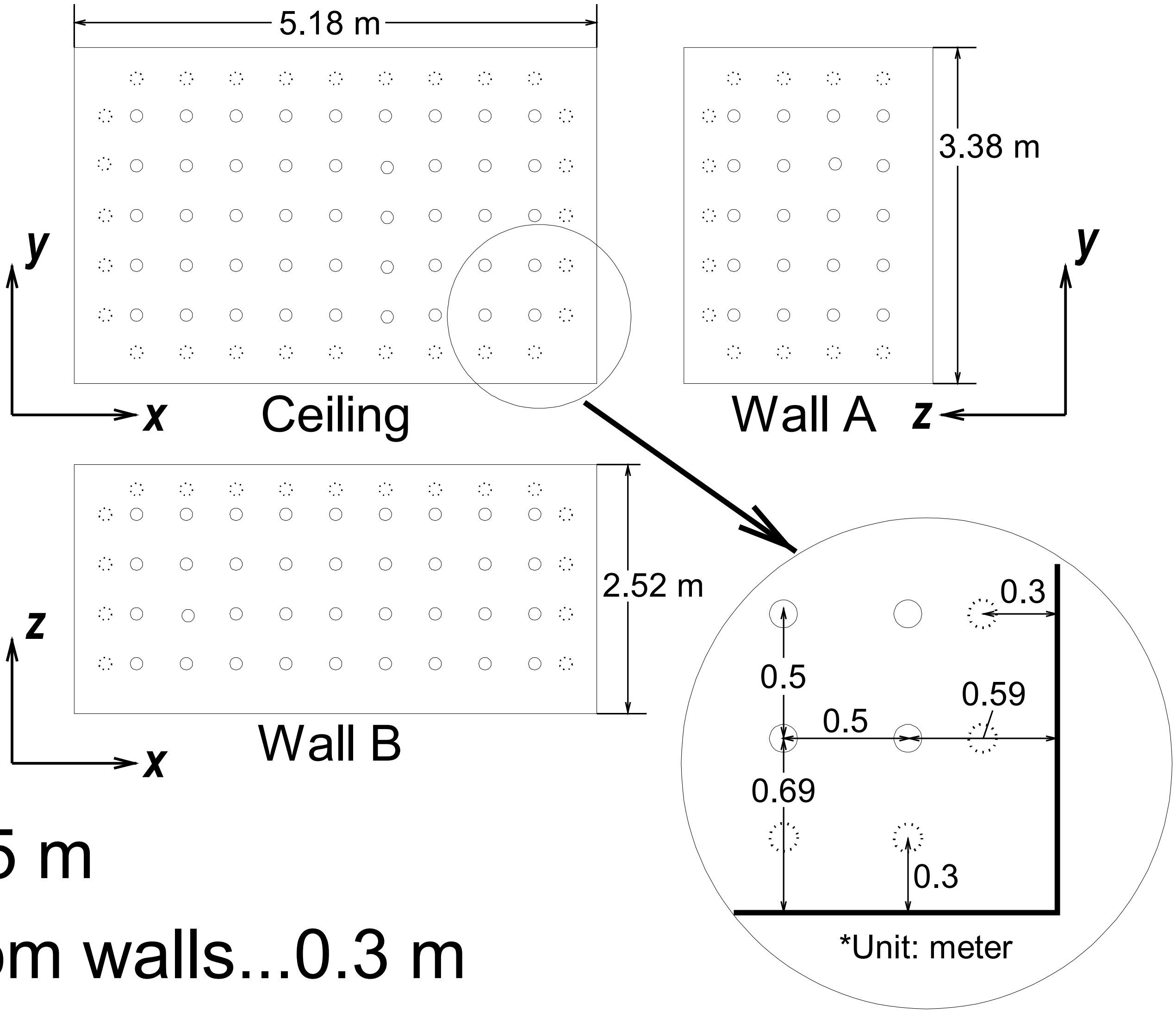
- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

Surrounding Microphone Array

- Surrounding microphone array room
  - Equipped in RIEC, Tohoku Univ.
  - Reverberation time...about 150 ms
- 157 omnidirectional microphones
  - B&K: Type 4951
  - Placed on 5 planes
  - Mainly record the direct sound from sound sources
- 10 Amplifiers
  - B&K: Type 2694

Arrangement of Microphones

- Wall A
  - 2 planes
  - 20 (=5×4)
- Wall B
  - 2 planes
  - 36 (=9×4)
- Ceiling
  - 1 plane
  - 45 (=9×5)
- Interval...0.5 m
- Distance from walls...0.3 m

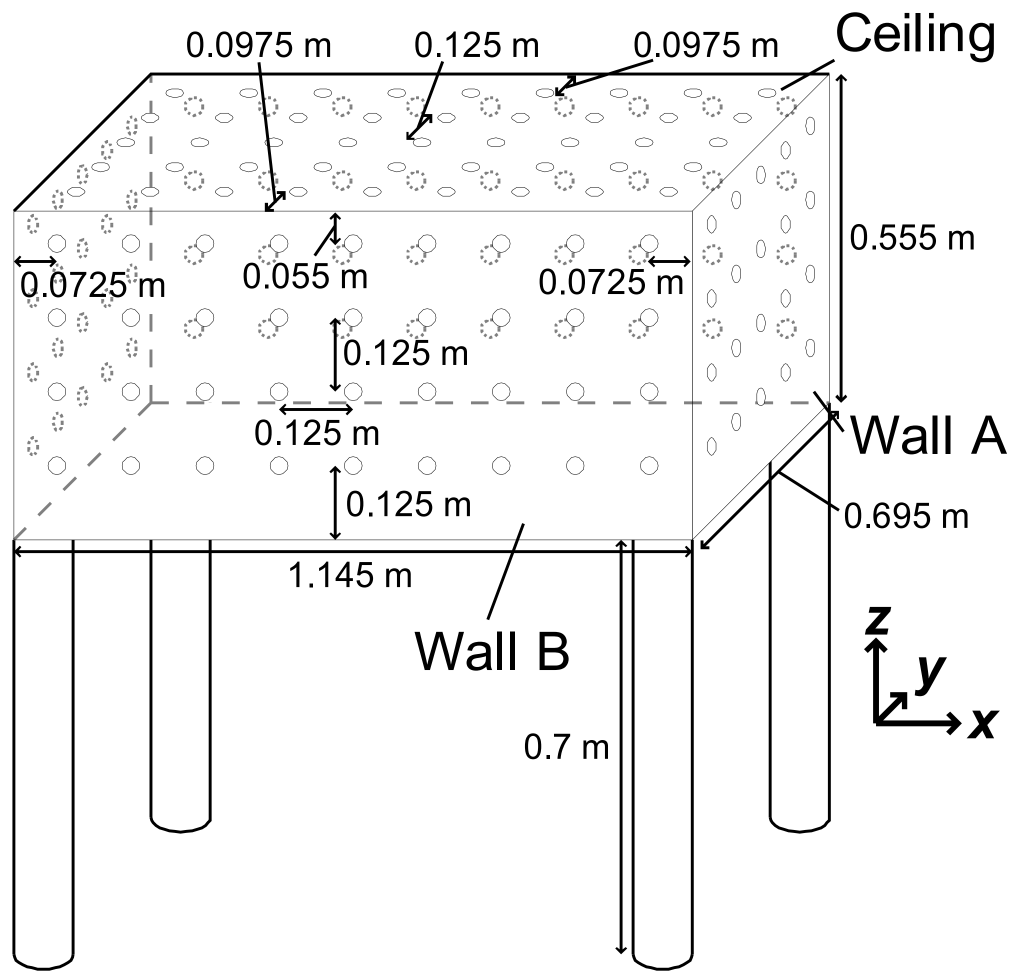


Radiated Loudspeaker Array

- Rectangular enclosure
  - Size...1/4 of surrounding microphone array
  - Material...Plywood and aluminum panels
- 157 loudspeaker units
  - AURASOUND:
    - NSW1-205-8A suitable
  - Attached to 5 planes
  - Size...1 inch
  - Directivity
    - Towards outside
- Amplifier
  - Custom-made (157ch)

Arrangement of Loudspeaker Units

- Interval...0.125 m
- Wall A
  - 2 planes
  - 20 (=5×4)
- Wall B
  - 2 planes
  - 36 (=9×4)
- Ceiling
  - 1 plane
  - 45 (=9×5)
- Elevated by 0.7 m



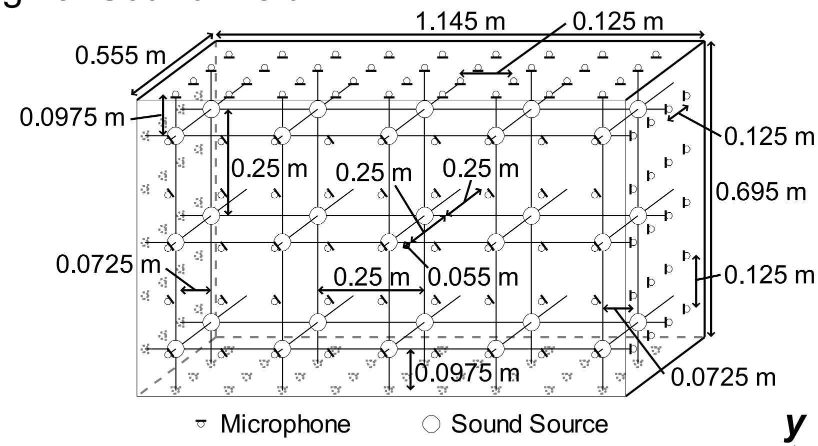
Outline

- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

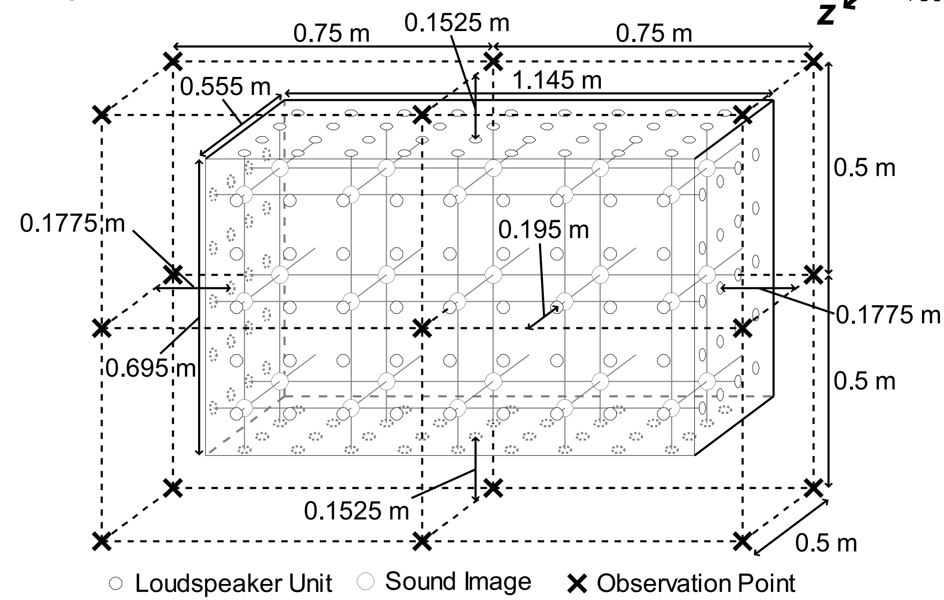
Simulation Environment

- Original sound field
  - Free space
  - 157 microphones
  - 30 sound sources
- Reproduced sound field
  - Free space
  - 157 loudspeaker units
  - 17 Observation points

Original Sound Field



Reproduced Sound Field



Synthesis of Recorded Signals

- Sound source signal $s(t)$
  - Octave-band noise (central frequency f_{cent})
- Signal of ith microphone $x_i(t)$

$$x_i(t) = g_i(t) * s(t) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{r}_i - \mathbf{r}_0|}{c}\right)$$

  - *: Convolution operation
  - $\mathbf{r}_i$: Position vector of ith microphone
  - $\mathbf{r}_0$: Position vector of sound sources
  - c: Sound velocity

Synthesis of Sound Pressure

- Sound pressure in observation points $p(\mathbf{R}_j, f_{\mathrm{cent}}, t)$

$$p(\mathbf{R}_j, f_{\mathrm{cent}}, t) = \sum_{i=1}^{M} \frac{D_{si}}{|\mathbf{R}_j - \mathbf{r}_i|} x_i\left(t - \frac{|\mathbf{R}_j - \mathbf{r}_i|}{c}\right)$$

$$= \sum_{i=1}^{M} \frac{D_{si}}{|\mathbf{R}_j - \mathbf{r}_i||\mathbf{r}_i - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{R}_j - \mathbf{r}_i| + |\mathbf{r}_i - \mathbf{r}_0|}{c}\right)$$

  - $\mathbf{R}_j$: Position vector of *j*th observation point
  - *M*: Total number of loudspeaker units
  - D_{si}: Radiation directivity of *i*th loudspeaker unit

Calculation of Sound Intensity

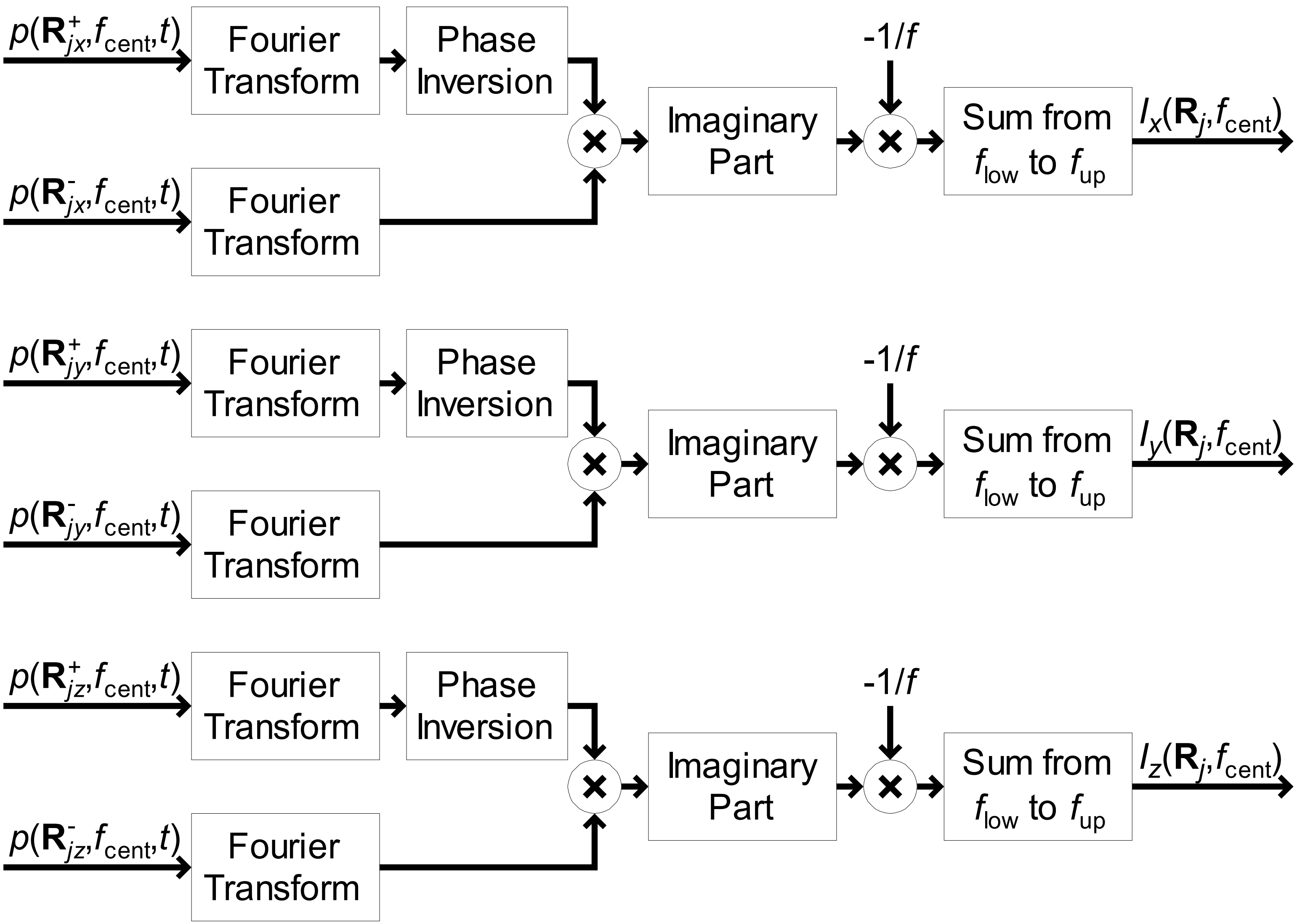
- Direction of sound intensity vector
  - Corresponds to arrival direction of sound sources

$$\mathbf{I}(\mathbf{R}_j, f_{\mathrm{cent}}) = \left\{ I_x(\mathbf{R}_j, f_{\mathrm{cent}}), I_y(\mathbf{R}_j, f_{\mathrm{cent}}), I_z(\mathbf{R}_j, f_{\mathrm{cent}}) \right\}^T$$

- Calculation by cross-spectral method
  - 3 directions are calculated from sound pressures at 6 points
    - Sound pressure: $p(\mathbf{R}^{+}_{jx}, f_{\mathrm{cent}}, t)$, $p(\mathbf{R}^{-}_{jx}, f_{\mathrm{cent}}, t)$, $p(\mathbf{R}^{+}_{jy}, f_{\mathrm{cent}}, t)$, $p(\mathbf{R}^{-}_{jy}, f_{\mathrm{cent}}, t)$, $p(\mathbf{R}^{+}_{jz}, f_{\mathrm{cent}}, t)$, $p(\mathbf{R}^{-}_{jz}, f_{\mathrm{cent}}, t)$,

$$\mathbf{R}^{\pm}_{jx} = \mathbf{R}_j \pm (\Delta, 0, 0)^T$$
$$\mathbf{R}^{\pm}_{jy} = \mathbf{R}_j \pm (0, \Delta, 0)^T \quad \Delta = 0.001\ \mathrm{m}$$
$$\mathbf{R}^{\pm}_{jz} = \mathbf{R}_j \pm (0, 0, \Delta)^T$$

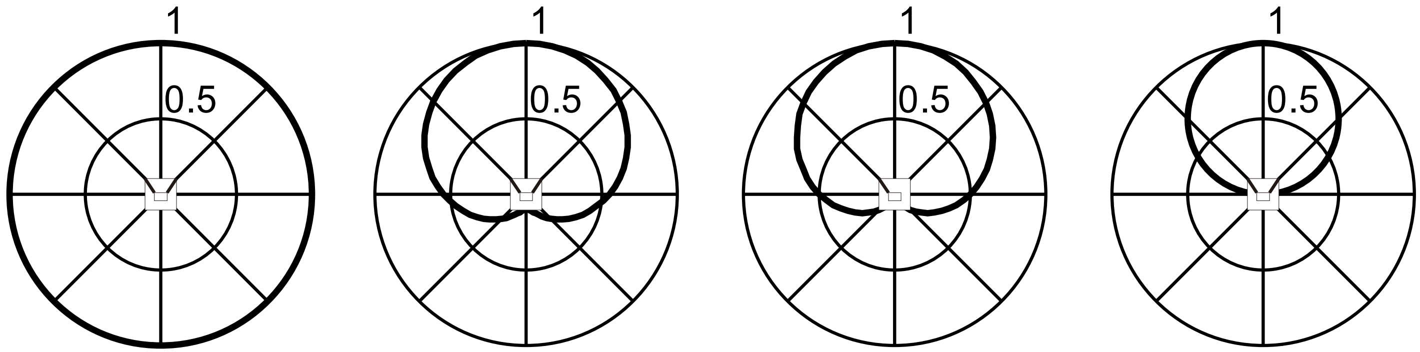
Block Diagram of SI Calculation



Parametric Conditions

Central frequency (f_{cent})	250, 500, 1000 Hz
Lower frequency (f_{low})	f_{cent}÷sqrt(2)
Upper frequency (f_{up})	f_{cent}×sqrt(2)
Sound velocity(c)	340 m/s
Number of loudspeaker units (M)	157
Radiation directivity of loudspeaker units (D_{si})	Omnidirectional, Decay20dB, Unidirectional, Shotgun

Radiation Directivity of Loudspeakers



(Omnidirectional) $D_{si} = 1$

(Decay 20dB) $D_{si} = 0.55 + 0.45\cos\theta_{si}$

(Unidirectional) $D_{si} = \dfrac{1 + \cos\theta_{si}}{2}$ $\qquad \cos\theta_{si} = \dfrac{\mathbf{n}_{si} \cdot (\mathbf{R}_j - \mathbf{r}_i)}{|\mathbf{n}_{si}||\mathbf{R}_j - \mathbf{r}_i|}$

(Shotgun) $$D_{si} = \begin{cases} \cos\theta_{si} & (\theta_{si} \leq 90^\circ) \\ 0 & (\theta_{si} > 90^\circ) \end{cases}$$

Estimation of Sound Image Position

- Estimated position of sound images $\mathbf{r}_E$

$$\mathbf{r}_E = \frac{1}{FN} \sum_{f_{cent}}^{250,500,1000} \sum_{j=1}^{N} \left\{ \mathbf{R}_j - \frac{\mathbf{I}(\mathbf{R}_j, f_{cent})}{p(\mathbf{R}_j, f_{cent})} \right\}$$

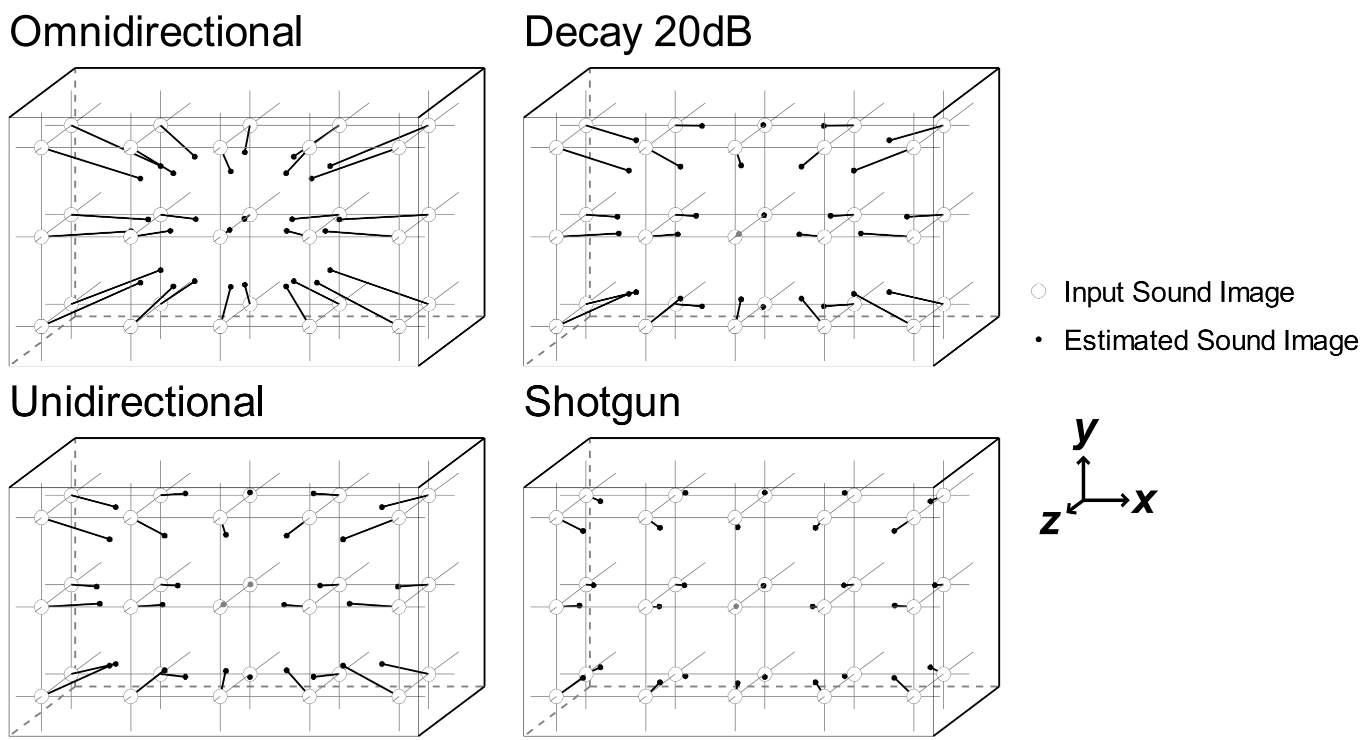
  - $\mathbf{R}_j$: Position vector of jth observation point
  - $\mathbf{I}(\mathbf{R}_j, f_{cent})$: Sound intensity at $\mathbf{R}_j$
  - $p(\mathbf{R}_j, f_{cent})$: RMS of sound pressure at $\mathbf{R}_j$

$$p(\mathbf{R}_j, f_{cent}) = \sqrt{\frac{1}{T} \int_0^T \{p(\mathbf{R}_j, f_{cent}, t)\}^2 dt}$$

  - T: Period
  - F(=3): Number of octave-band noises
  - N: Number of observation points
    - All observation points, 4 observation points near source

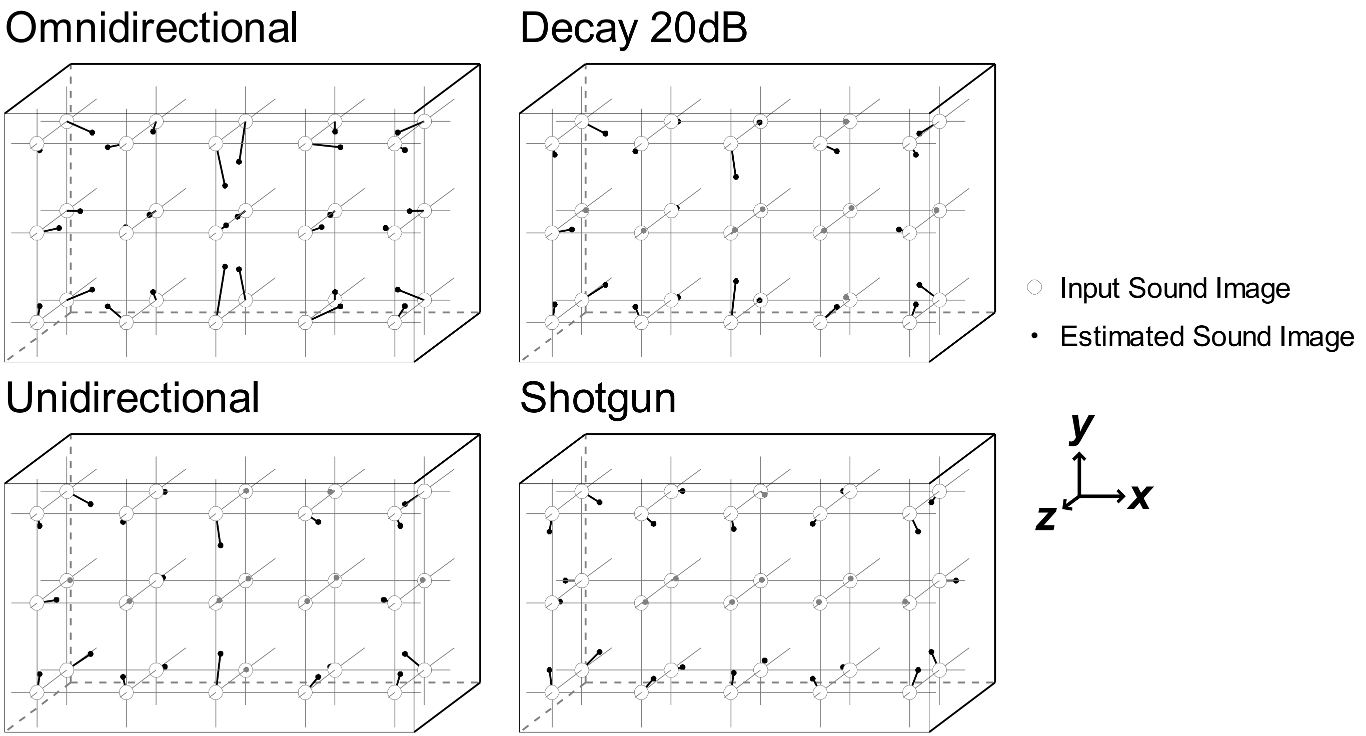
Estimated Position (All Points)

- If radiation directivity is sharpened
  - Sound images are accurately estimated
  - Listeners can accurately localize sound images at any listening position around loudspeaker array



Estimated Position (4 Points Near Source)

- Even if radiation directivity is not sharpened
  - Sound images are approximately estimated
  - Listeners can accurately localize sound image close to listening position

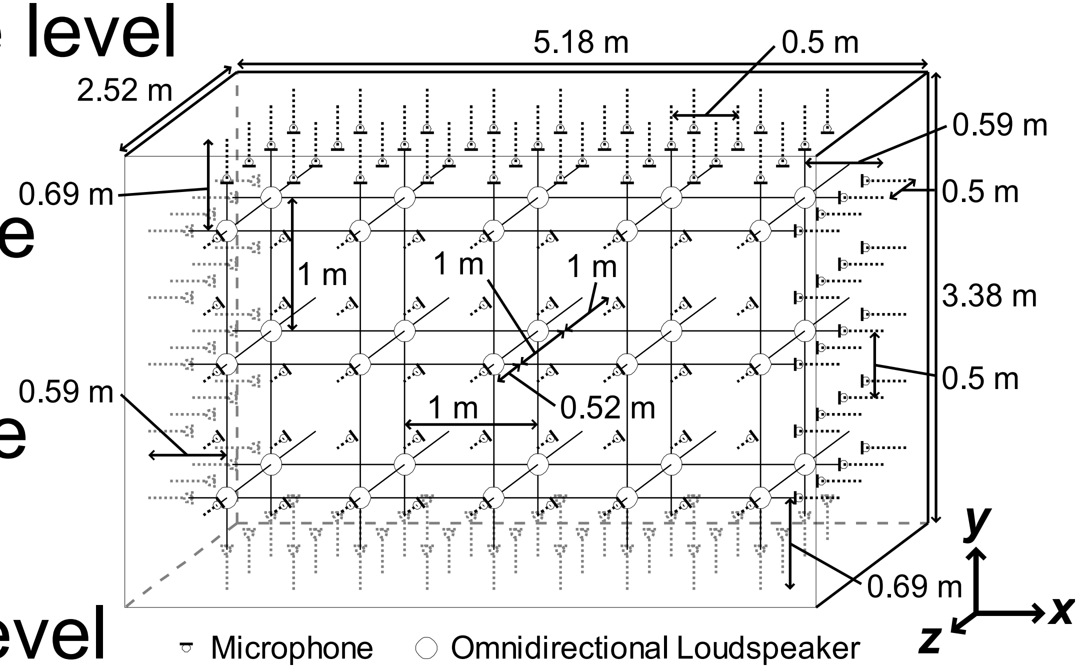


Outline

- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

Room Impulse Response Measurement

- Omnidirectional loudspeakers are placed at 30 positions in the surrounding microphone array room
- Background noise level
  - 18.4 dB(A)
- Reverberation time
  - 150 ms
- Room temperature
  - 20°C
- Sound pressure level
  - 85.6 dB(A) at 1m distance from loudspeakers



RIR Measurement Conditions

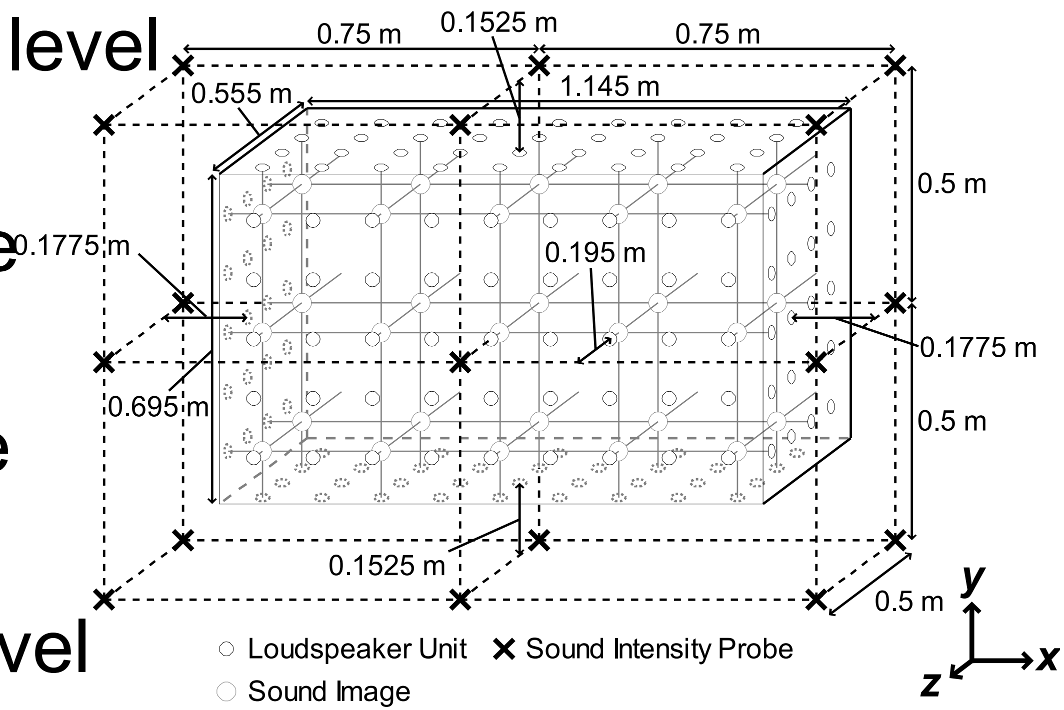
- TSP signal
  - Sampling frequency...48 kHz
  - Quantization bits...16 bits
  - Length...65536 samples
- FIR filter
  - Length...14400 taps
  - Number of synchronous repetitions...16
  - Sampling frequency is regarded as 192 kHz
    - match the size of surrounding microphone array to that of radiated loudspeaker array
    - Reverberation time...about 37.5 ms (1/4 of 150 ms)

Synthesis of 157 Channel Signals

- Convolve measured room impulse responses to sound source signals
- Sound source signal
  - Octave-band noise
    - Central frequency...250, 500, 1000 Hz
  - Sampling frequency...48 kHz
  - Quantization bits...16 bits
  - Length...10 s
  - Time of fade-in and fade-out...1 ms
- Room Impulse response
  - Transformed to 48 kHz before synthesis

Measurement of Sound Intensity

- Synthesized 157ch signals are played by the radiated loudspeaker array and sound intensity probes are placed at 17 positions
- Background noise level
  - 22 dB(A)
- Reverberation time
  - 180 ms
- Room temperature
  - 22°C
- Sound pressure level
  - 71 dB(A) at 1 m from central sound image



Estimation of Sound Image Position

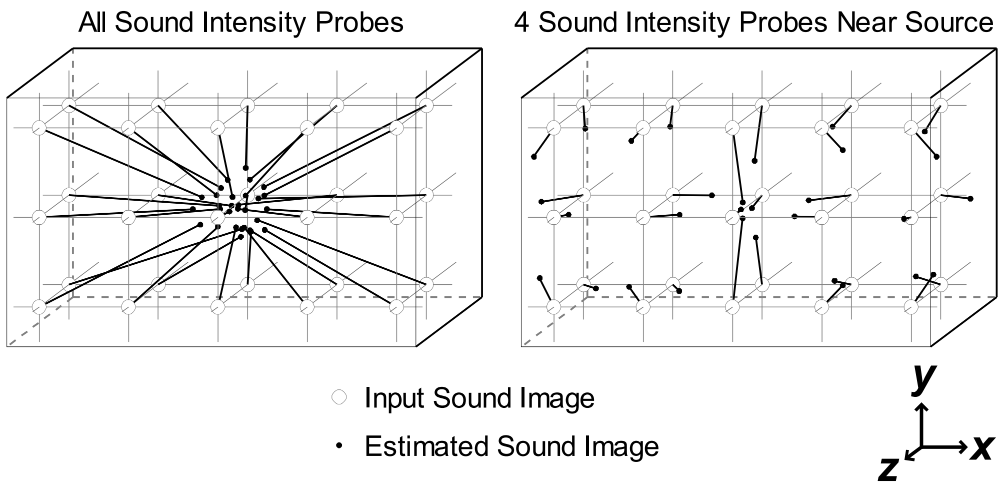
- Estimated sound image position $\mathbf{r}_E$

$$\mathbf{r}_E = \frac{1}{FN} \sum_{f_{\text{cent}}}^{250,500,1000} \sum_{j=1}^{N} \left\{ \mathbf{R}_j - \frac{\mathbf{I}(\mathbf{R}_j, f_{\text{cent}})}{p(\mathbf{R}_j, f_{\text{cent}})} \right\}$$

  - $\mathbf{R}_j$: Position vector of jth sound intensity probe
  - $\mathbf{I}(\mathbf{R}_j, f_{\text{cent}})$: Sound intensity at $\mathbf{R}_j$
  - $p(\mathbf{R}_j, f_{\text{cent}})$: Sound pressure at $\mathbf{R}_j$
  - $F(=3)$: Number of octave-band noises
  - N: Number of sound intensity probes
    - All sound intensity probes
    - 4 sound intensity probes near source

Estimated Sound Image Position

- All sound intensity probes
  - Estimated positions are biased to the center
- 4 sound intensity probes near source
  - Sound images are approximately estimated
- Listeners can accurately localize sound image close to listening position



Outline

- Introduction
  - Ultra-realistic communication
  - Wave field synthesis system
- Diagram of proposed system
- Developed system
- Evaluation of developed system
  - Computer simulation
  - Acoustical measurement
- Conclusion

Conclusion

- Evaluation of the performance of the sound image localization in the developed system
  - The positions of sound images were estimated by computer simulation and acoustical measurement
    - If the radiation directivity is sharpened, listeners can accurately localize sound images in any listening position
    - Even if the radiation directivity is not sharpened, listeners can accurately localize the sound image close to the listening position
- Future works
  - Development of a system with better performance in a real environment by sharpening the radiation directivity of the loudspeaker units