17th May 2012, Acoustics 2012 @ Hong Kong Convention and Exhibition Centre

Acoustic privacy area generation based on simple summation of numerous loudspeaker signals

Takuma OKAMOTO[(1)], Yukio Iwaya[(2)] and Yôiti Suzuki[(3)]

(1) National Institute of Information and Communications Technology (NICT), Japan

(2) Tohoku Gakuin University, Japan

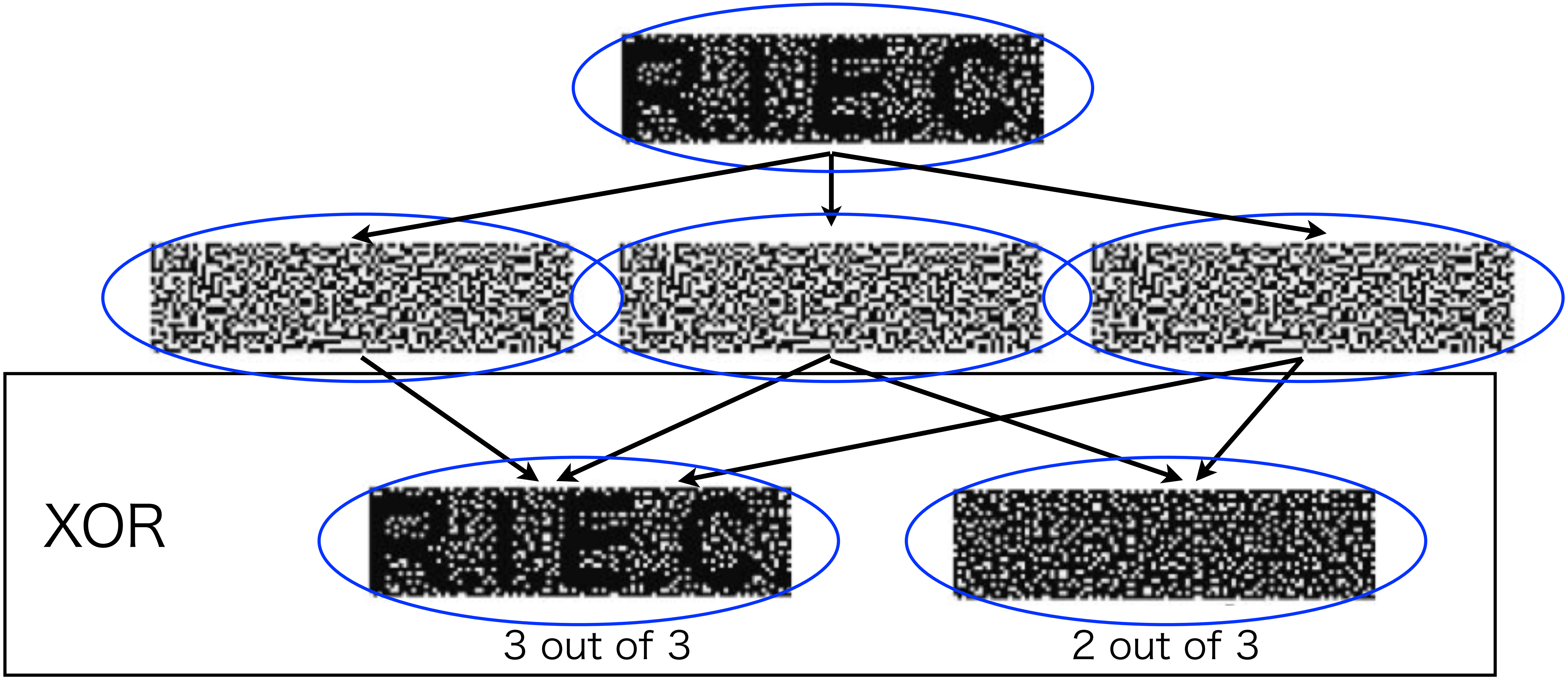
(3) Tohoku University, Japan

Motivation

- Audio secret sharing for 1-bit audio

  N. Fujita, R. Nishimura and Y. Suzuki, *Acoust. Sci. & Tech.* **27**, 3 (2006)

  - based on Secret sharing schema (A. Shamir 1979)



Motivation

- Audio secret sharing for 1-bit audio

  N. Fujita, R. Nishimura and Y. Suzuki, *Acoust. Sci. & Tech.* **27**, 3 (2006)

  - based on Secret sharing schema (A. Shamir 1979)
    - Original speech can only be restored when all streams can be got
    - Privacy speech communication on network

- Realizing these communications on actual sound field!!
  - Enhancing a speech signal at a point
    - Speech privacy technique on actual environment
  - Problem
    - XOR cannot be calculated on actual sound field and hearing

Previous methods

- Speech enhancement techniques at a point
  - based on multiple points control
    - J.-W. Choi *et al.* (2002)
    - S. Enomoto *et al.* (2005)
  - based on beamforming
    - Y. Tamai *et al.* (2003)

Techniques for enhancing objective signal

Objective signal can be listened at outside of sweet spot

- Speech privacy techniques
  - based on masking

Proposed method

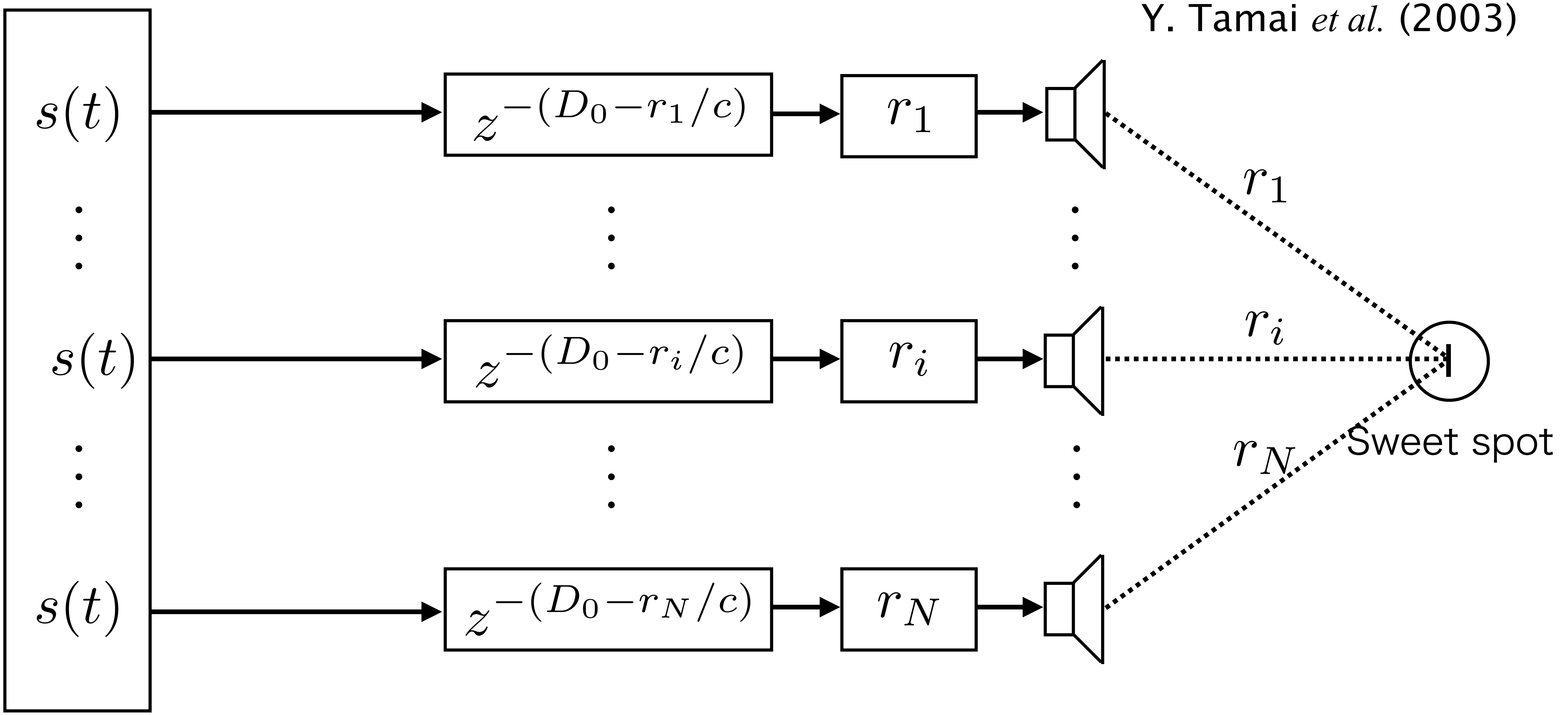
Techniques for masking objective signal

Spacial information is not taken into account

Previous method

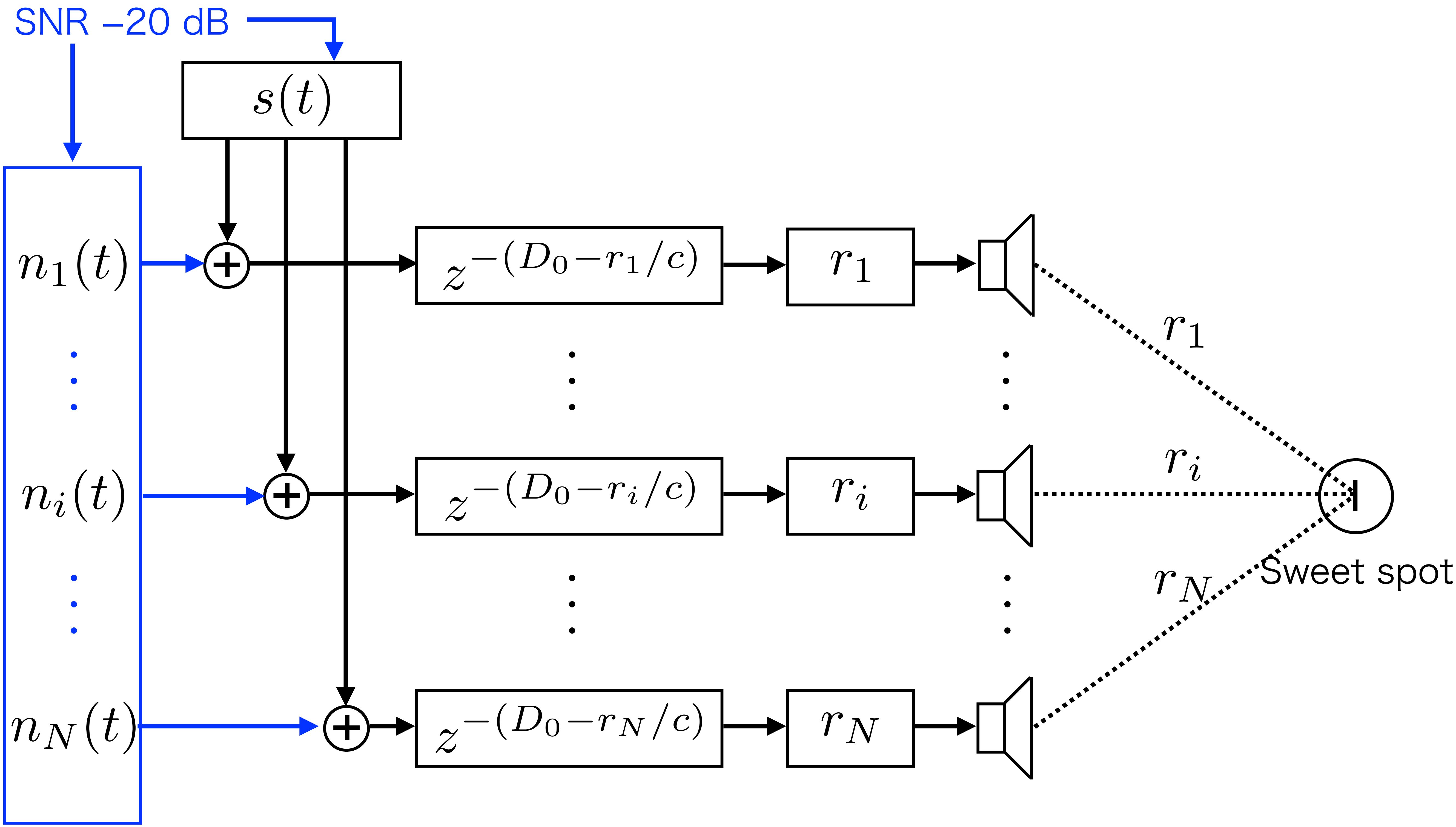
- Sound spot generation based on delay-and-sum

Y. Tamai *et al.* (2003)



- Amplitude of signal is amplified N times at a sweet spot
- Problem : Speech signal can be listened outside of a sweet spot

Proposed method 0



Basis of proposed method 0

- Sound spot generation based on simple summation with random masking signals

$$x_i(t) = s(t) + n_i(t)$$ Loudspeaker signal at i–th channel

$$y_i(t) = \frac{1}{r_i} x_i(t - r_i/c)$$ Observed signal from i–th channel

$$y(t) = \sum_{i=1}^{N} r_i \cdot \frac{1}{r_i} x_i(t - D_0) = Ns(t - D_0) + \sum_{i=1}^{N} n_i(t - D_0)$$ Total observed signal

  - Masking signals are mutually uncorrelated each other

$$n(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{t^2}{\sigma^2}\right)$$

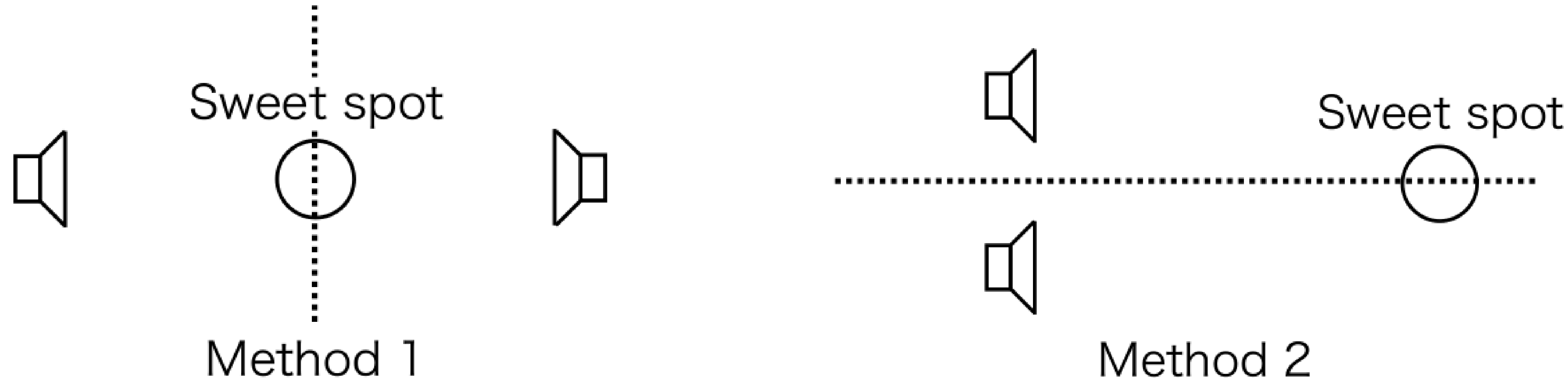
$$y(t) = Ns(t) + \frac{N}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{1}{2}\frac{t^2}{n\sigma^2}\right) = Ns(t) + \sqrt{N} n(t)$$

  - We can only masking signal at each loudspeaker position
  - SNR increases proportionally to $\sqrt{N}$ and the objective signal can be heard only at the sweet spot

Introduction of reverse phase pairs

- Problem in Proposed method 0
  - Masking signals can be listened at a sweet spot

- Improved method : introduction of reverse phase pairs of masking signals
  - Masking signal components can completely be 0!!
    - Proposed method 1 : reverse phase pairs composed of set to each other
    - Proposed method 2 : reverse phase pairs composed of next to each other



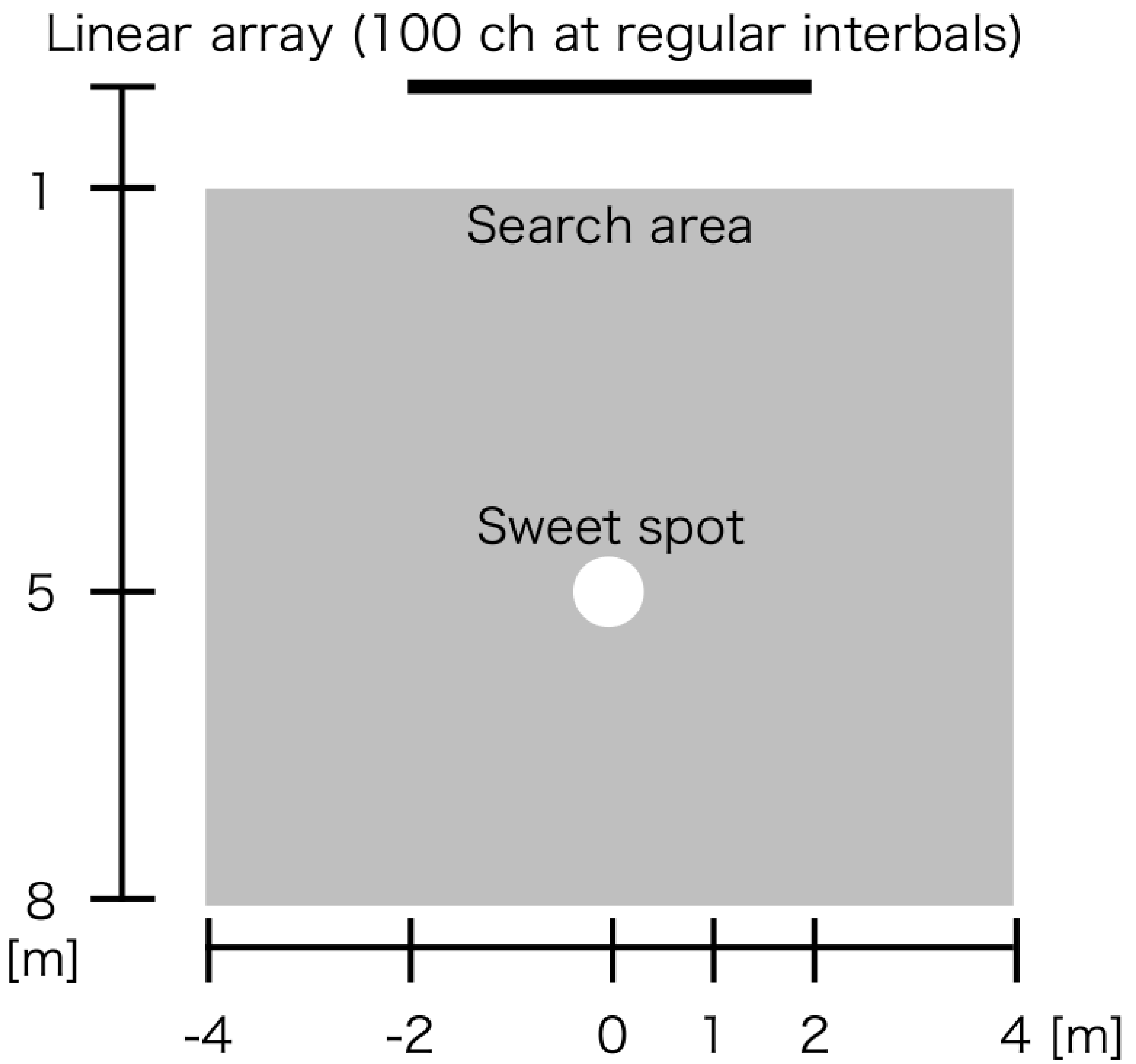
Performance evaluation

- Computer simulations
  - Signal to Noise Ratio (SNR)
  - Speech Intelligibility Index (SII) ANSI S3.5-1997
    - Evaluate speech intelligibility as a score 0 to 1
    - Calculated from each SNR between speech and noise at each frequency band

  - Simulation conditions
    - Sampling frequency : 48 kHz
    - Objective signal : a male voice from Speech Database by NTT-AT
    - Masking signals : Pink noise
    - SII conditions : 1/3 octave band (18 bands)

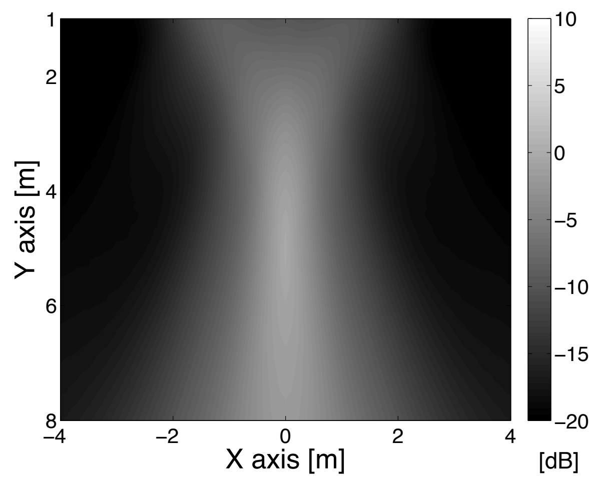
      (center frequency:160 Hz to 8000 Hz)
    - Analysis bandpass filter : FIR filter (128 samples)

Linear array

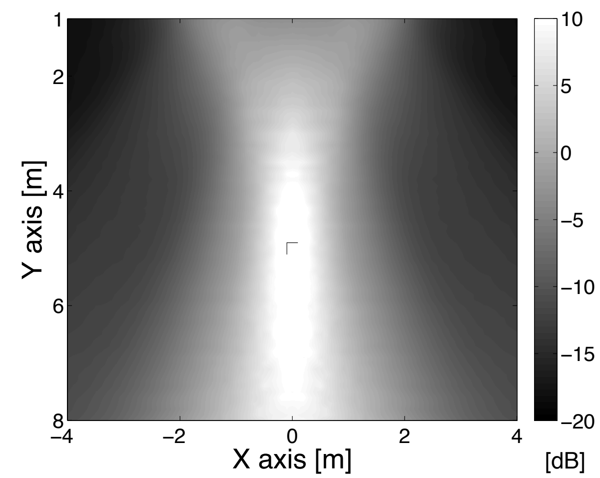


Linear array : Results

- SNR

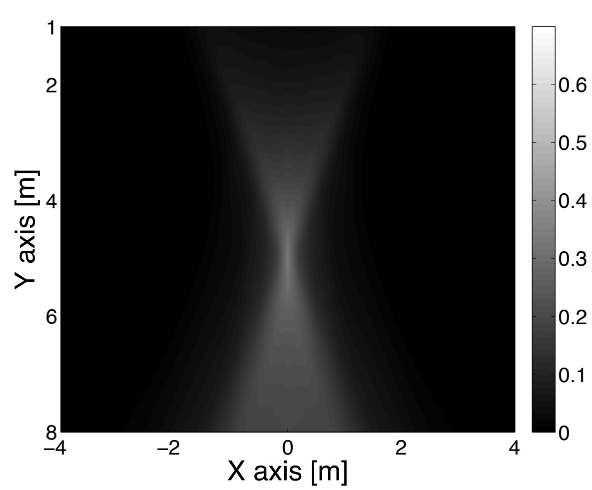


Proposed 0

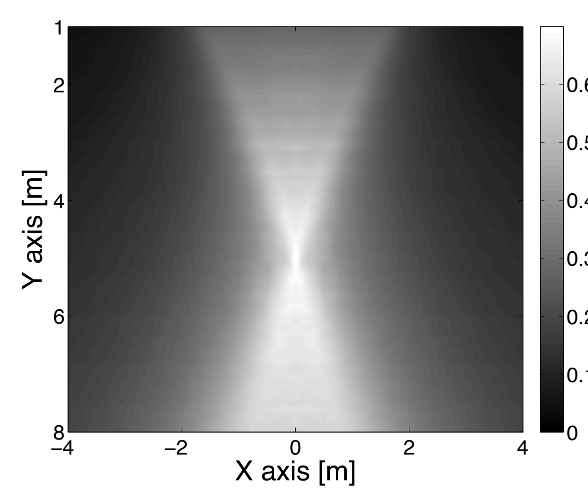


Proposed 2

- SII



Proposed 0

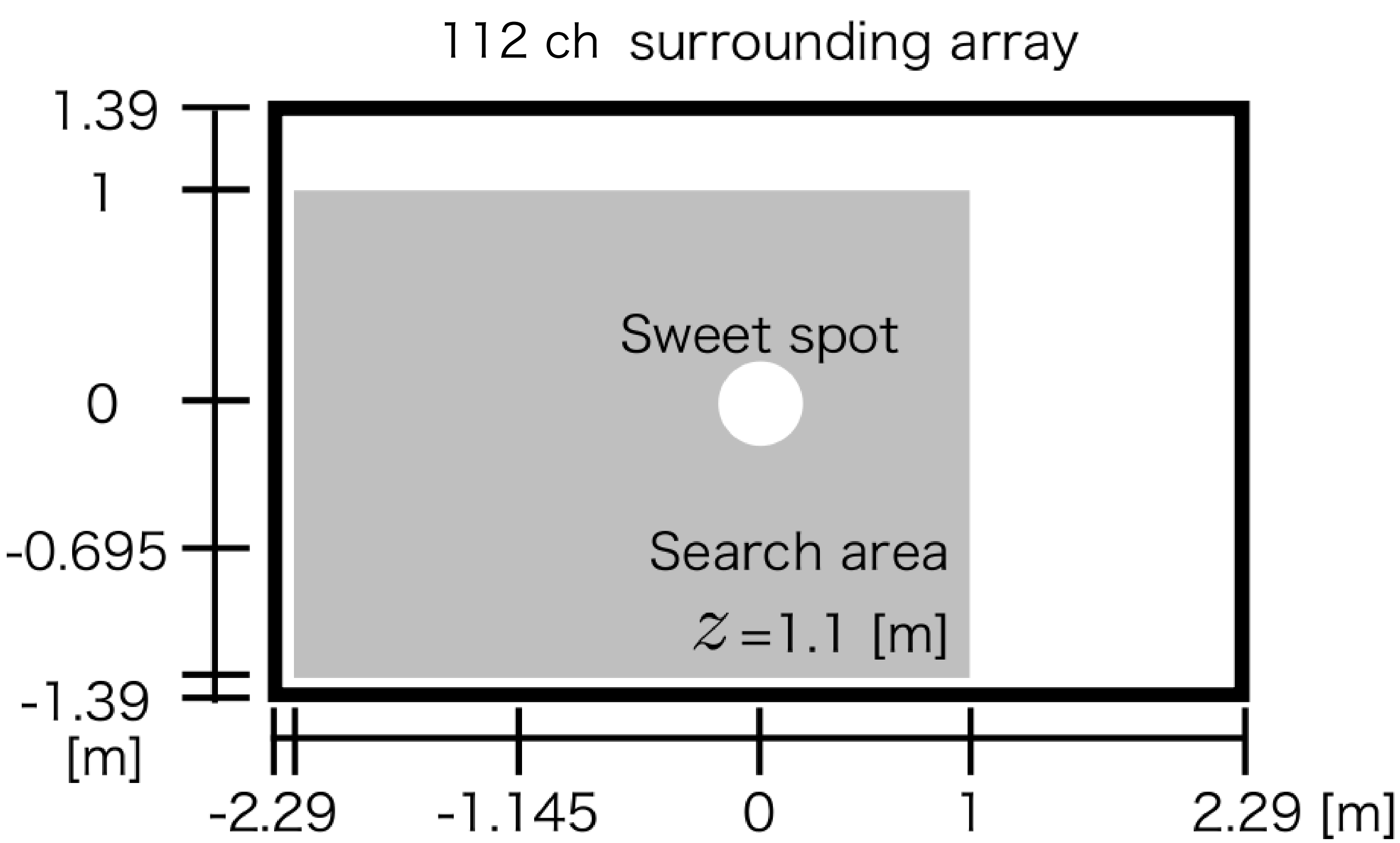


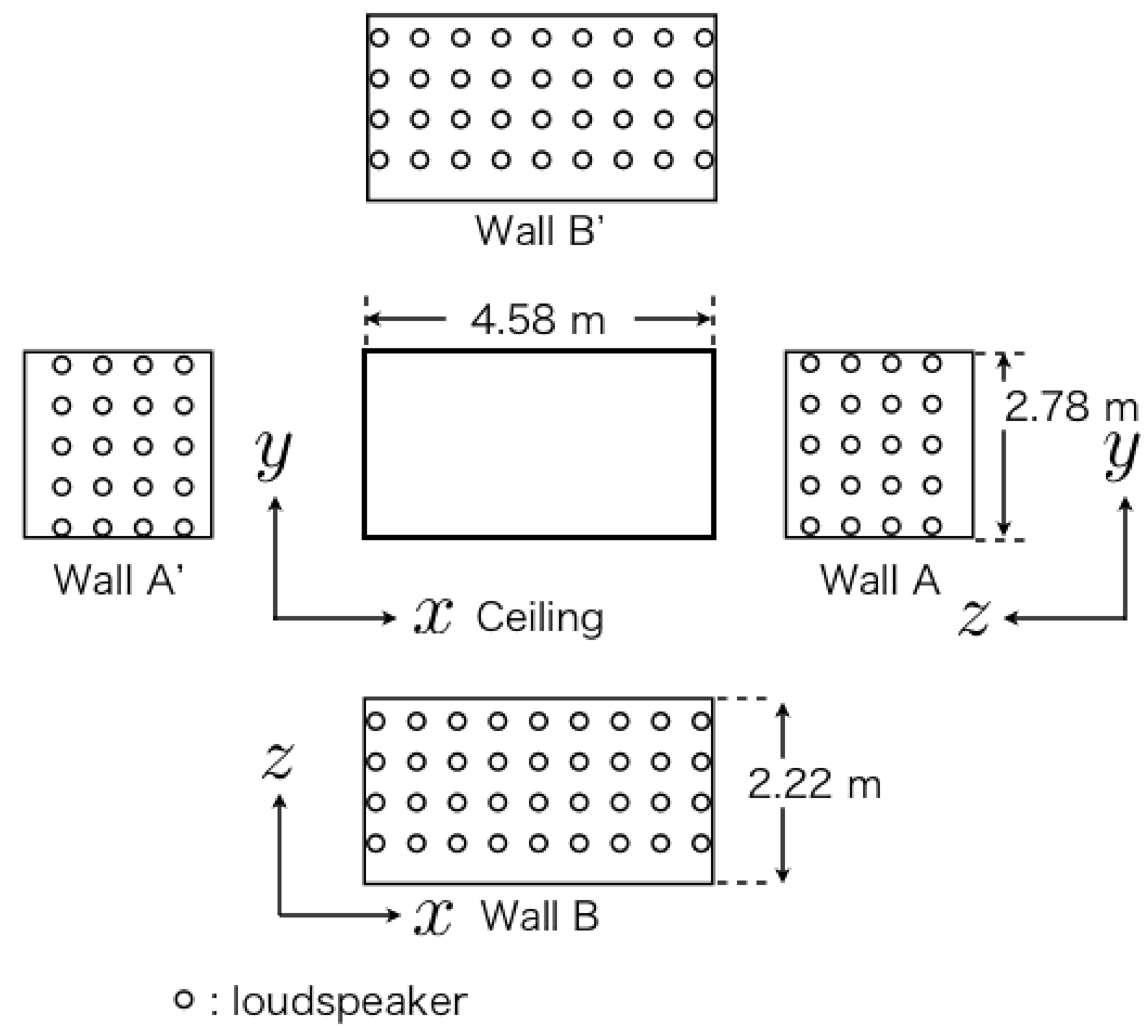
Proposed 2

Surrounding 157 array

Surrounding 157 loudspeaker array
@R.I.E.C., Tohoku Univ.

Surrounding 112 array

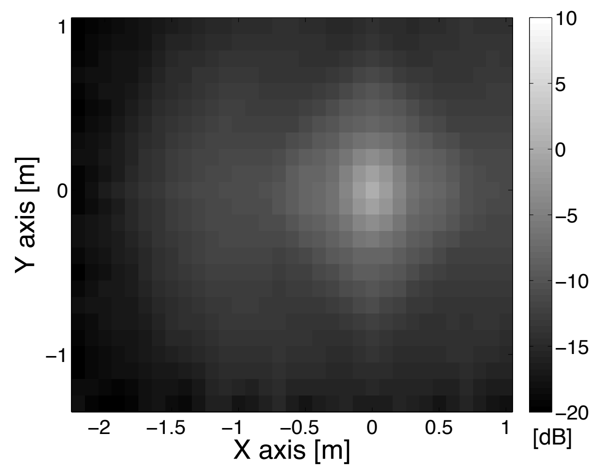




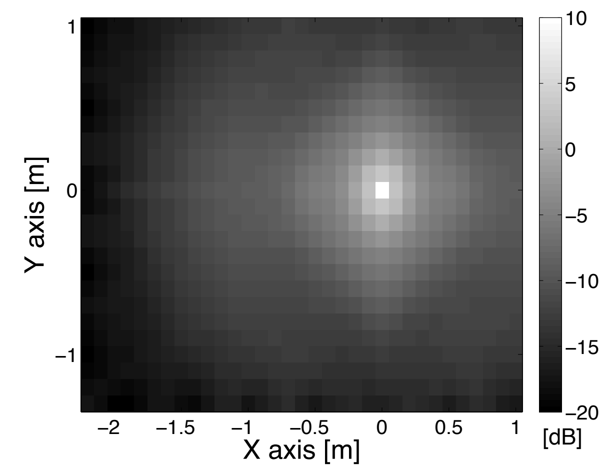
Arrangement of loudspeakers

Surrounding 112 array : Results

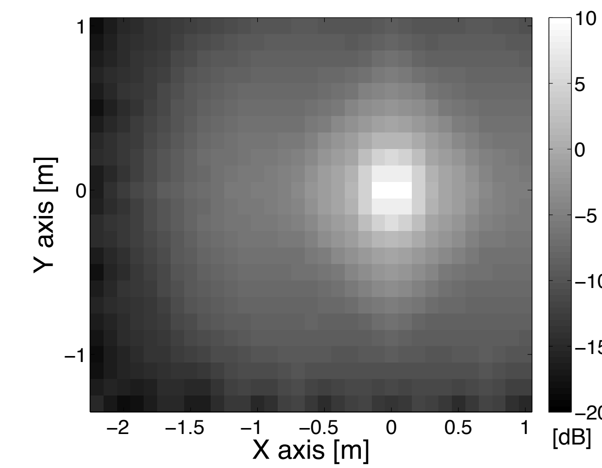
- SNR



Proposed 0

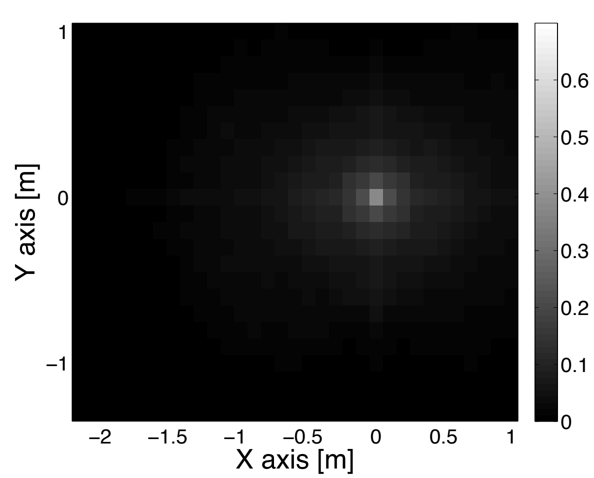


Proposed 1

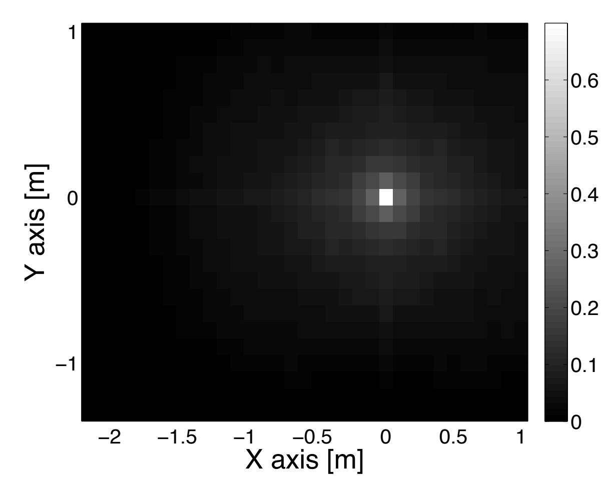


Proposed 2

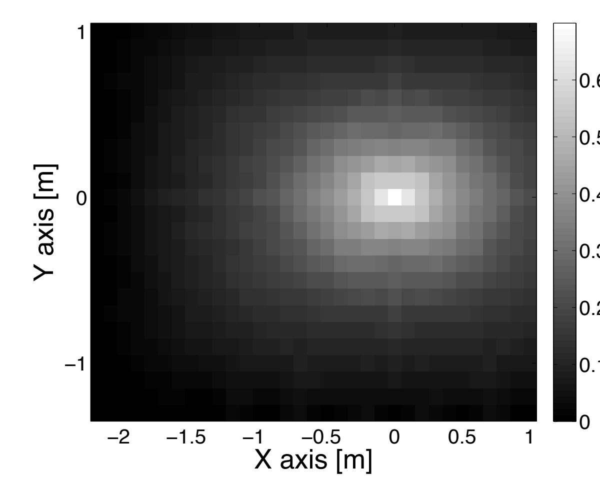
- SII



Proposed 0



Proposed 1



Proposed 2

Concluding remarks

- Acoustic privacy area generation based on simple summation
  - Using mutually uncorrelated masking signals
  - Introduction of reverse phase pairs for expansion of sweet spot
    - Reverse phase pairs composed of next to each other can increase SNR around a sweet spot efficiently

- Acknowledgement
  - This study was partly supported by the GCOE program (CERIES) of the School of Engineering, Tohoku University