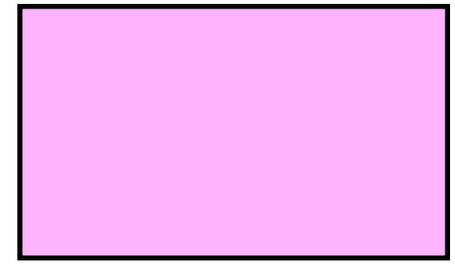
Implementation of higher order Ambisonics recording array with 121 microphones

T. Okamoto[(1)(2)], Y. Iwaya[(1)(3)], S. Sakamto [(1)(3)] and Y. Suzuki[(1)(3)]

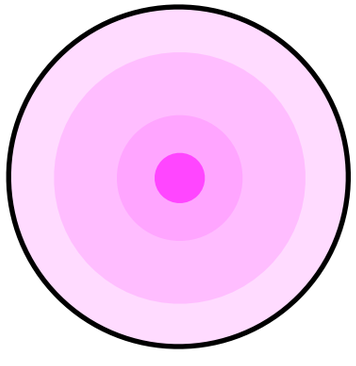
(1) Research Institute of Electrical Communication
(2) Graduate School of Engineering
(3) Graduate School of Information Sciences
Tohoku University, Japan

Introduction and research aim

- Implementation of 3D sound field recording and reproduction system
  - Realizing communications with high sense-of-presence
- Sound field recording and reproduction based on higher order Ambisonics (HOA)
  - Highly sound field reproduction around a center of array than WFS (Wave Field Synthesis) and BoSC (Boundary Surface Control)
    - We have already implemented 5th order Ambisonics reproduction system using 157 loud-speaker array
    - High definition Ambisonics recording system is required



WFS or BoSC



Ambisonics

Relationship between reproduction accuracy and frequency

- Today's topics
  - Implementation of 3D sound field recording system
    - Overview of HOA
    - Designing 121 spherical microphone array
    - Implementation with 121 Digital Electric Condenser Microphones (ECMs) and FPGA board

Concept of HOA

- Simple example of orthogonal expansion
  - Taylor expansion

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

- Orthogonal expansion of sound field information based on spherical harmonics in HOA

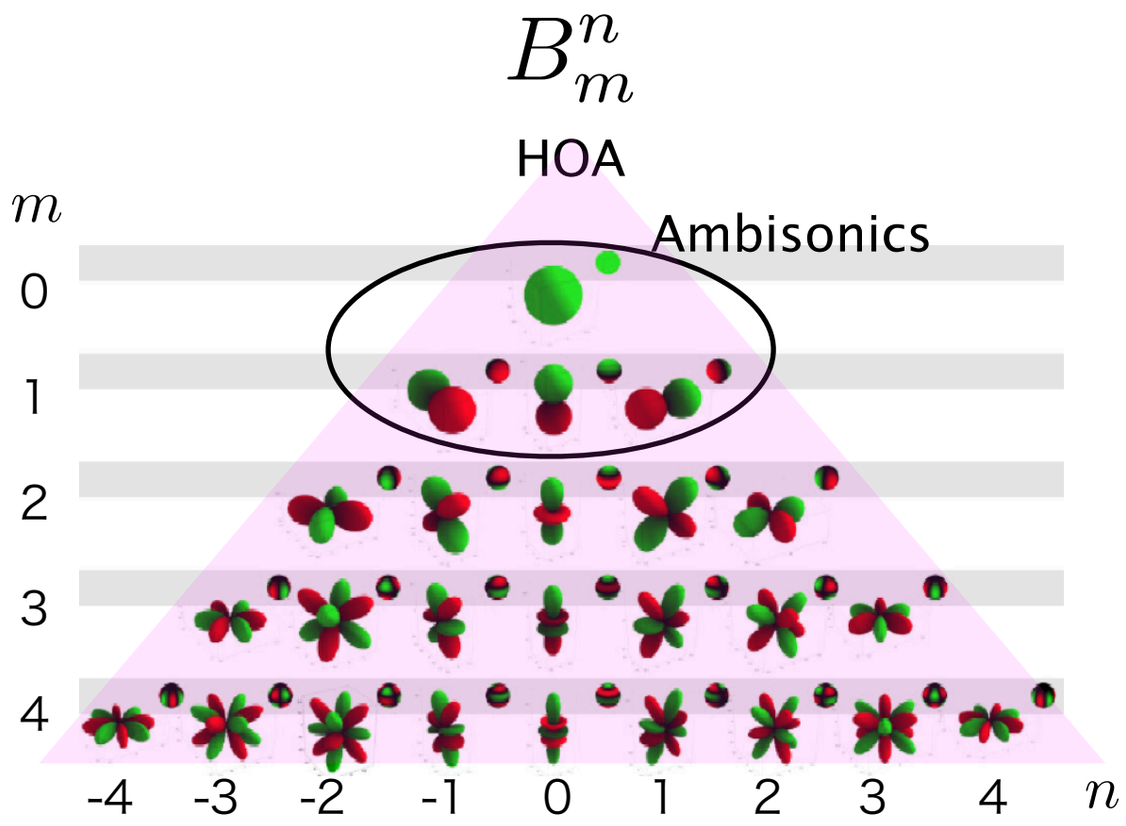
$$p(kr, \theta, \phi) = \sum_{m=0}^{+\infty} i^m j_m(kr) \sum_{n=-m}^{m} B_m^n Y_m^n(\theta, \phi)$$

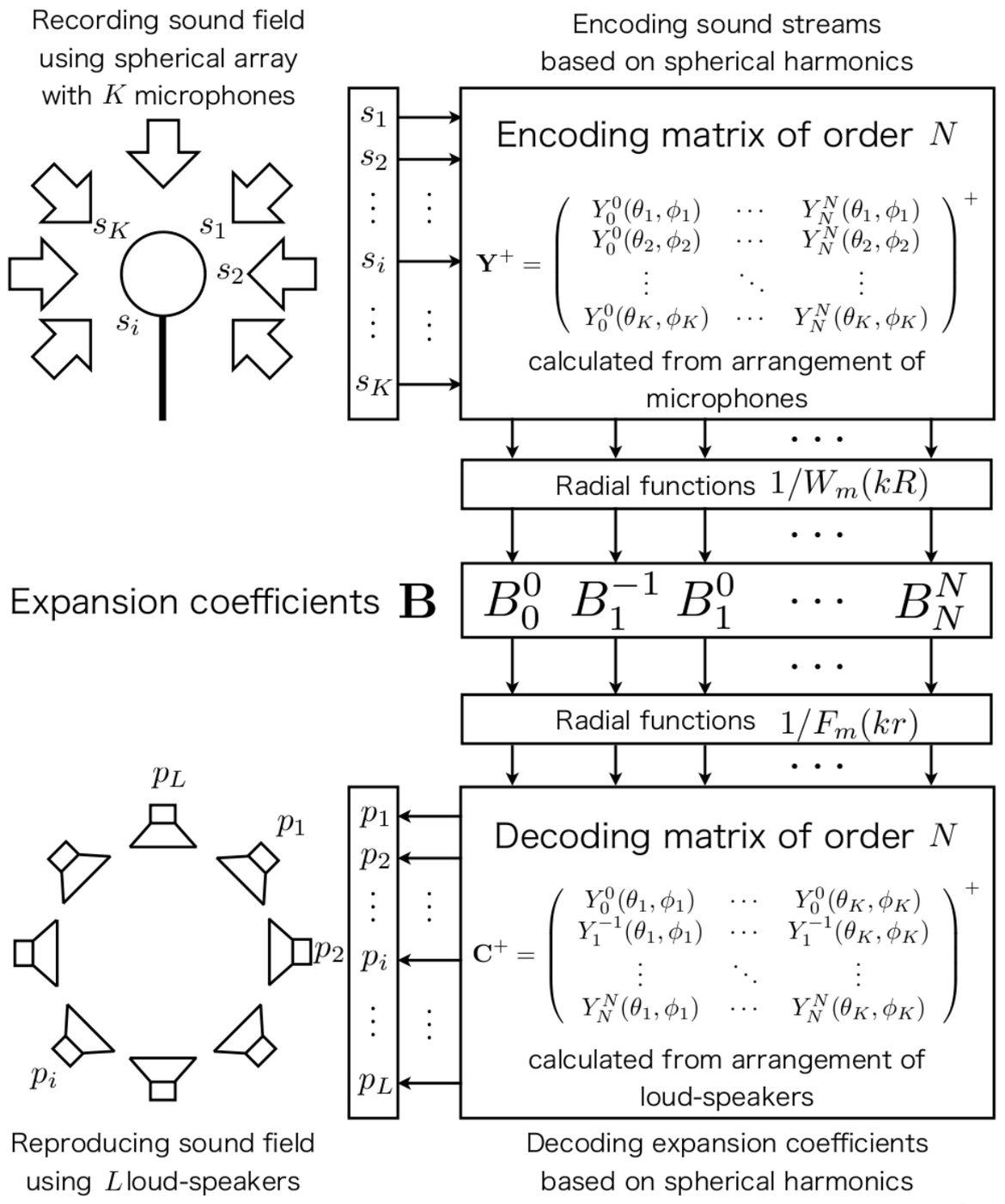
  - Sound information is decomposed to expansion coefficients B_m^n using spherical harmonics $Y_m^n(\theta, \phi)$ of each direction (θ, ϕ)
  - Using more higher order m, more accurate sound field is reproduced

$$\mathrm{Y}_{mn}(\theta, \varphi) = \begin{cases} N_m^0 P_m^0(\cos\varphi) & \text{if } n = 0, \\ \sqrt{2} N_m^n P_m^n(\cos\varphi)\cos n\theta & \text{if } n > 0, \\ \sqrt{2} N_m^n P_m^{-n}(\cos\varphi)\sin n\theta & \text{if } n < 0, \end{cases} \quad N_m^n = \sqrt{\frac{(2m+1)}{4\pi}\frac{(m-|n|)!}{(m+|n|)!}}$$

Sound field recording and reproduction based on HOA system

- Aspects of HOA system
  - Using more higher order m, more accurate sound field is reproduced
  - More than $(m+1)^2$ channels are needed to encode or decode order m
  - Recording system and reproduction system are independent each other

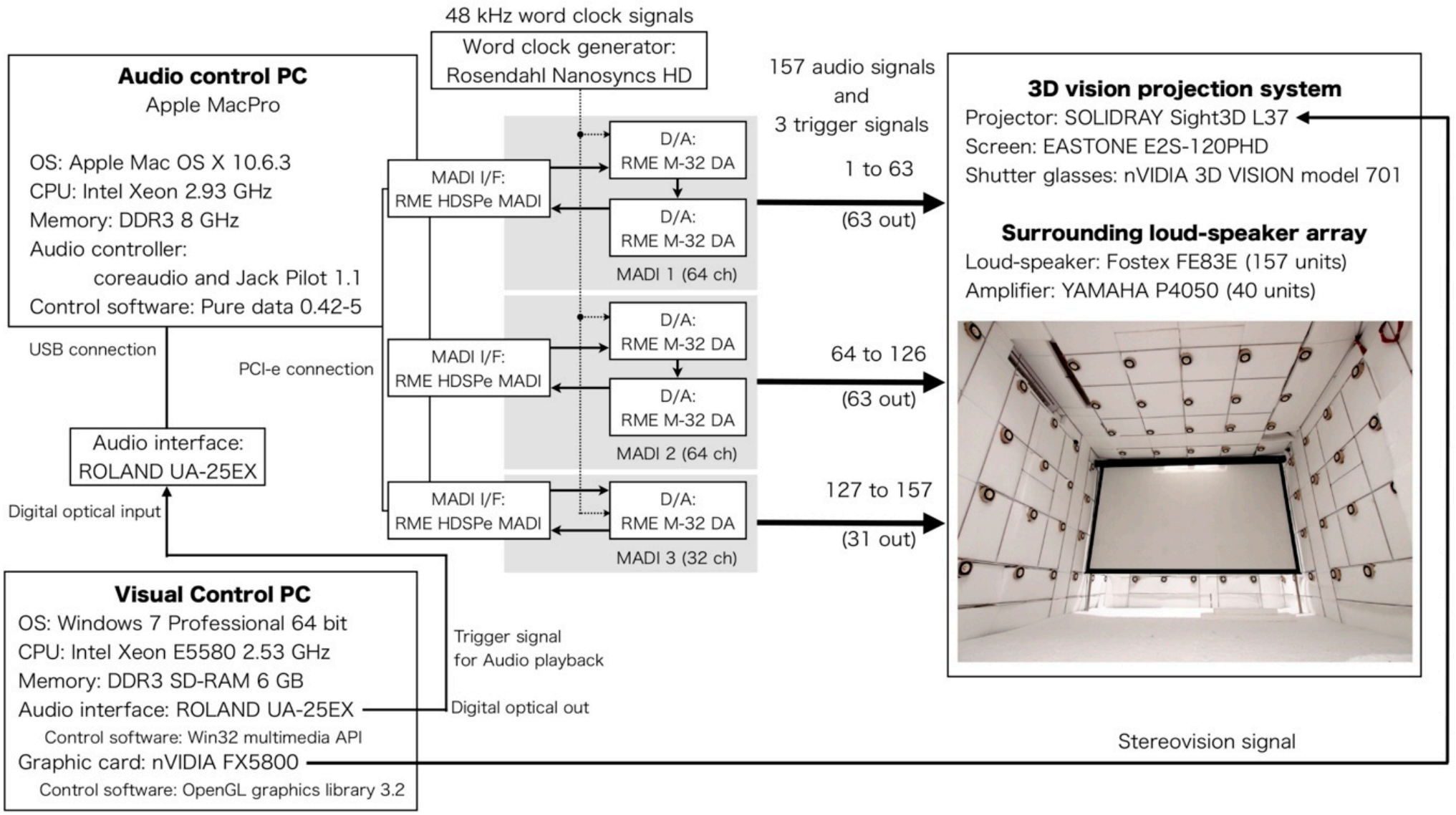




Implementation of 3D audio-visual display

T. Okamoto *et al.*, *Proc. IEEE IC-NIDC 2010*

- Combining HOA system with 3D projection display
  - Audio system: Completely synchronous 157-loudspeaker array system
  - Visual system: Stereo shutter technique with acoustic transparent screen
    - 5th order decoding, which is highest order in the world, was realized



Implementation of HOA recording system with more than 100 microphones

- Current HOA recording system
  - Up to 64 channels with up to 6th order encoding
  - These systems use analog microphones and A/Ds

S. Moreau *et al.* (32 ch)
@Orange Labs

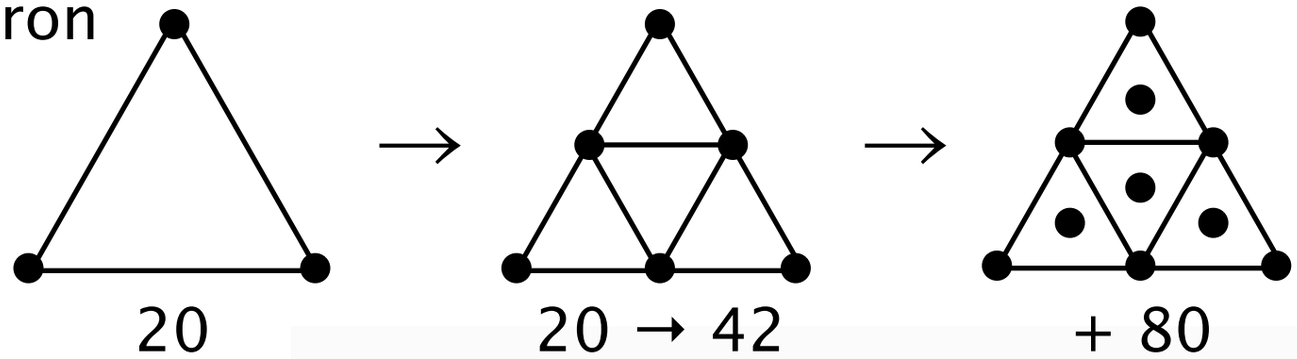
D. N. Zotkin *et al.* (64 ch)
@University of Maryland

A. Parthy *et al.* (64 ch)
@University of Sydney

- Our research aim
  - Implementation of HOA recording system with more than 100 microphones
    - More than 100 ch system cannot realize when using analog microphones and A/Ds
      - Problems: Size of array must be large! Interconnection lines must be huge!
    - Digital Electric Condenser microphones (ECMs) and FPGA board are introduced

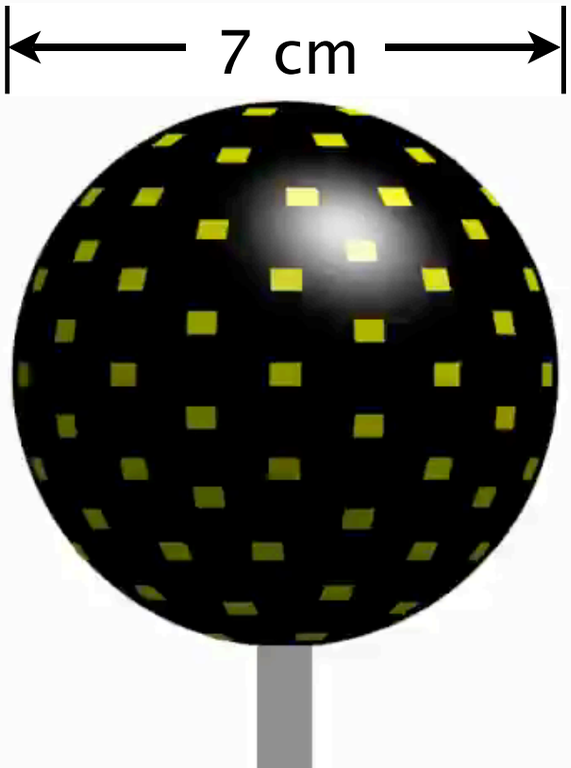
Designing 121 spherical microphone array

- Arrangement of microphones
  - In HOA system, regular arrangement (equal density) is important
  - Nearly equal density arrangement is calculated from regular icosahedron based on polyhedral subdivision method
  - Total vertex number is 122



- Radius of sphere
  - 3.5 cm
    - Designed from size of Digital ECMs, interconnection lines and control board in the sphere
    - 1 channel at the bottom is used for interconnection output
    - Spatial aliasing frequency: 8.9 kHz

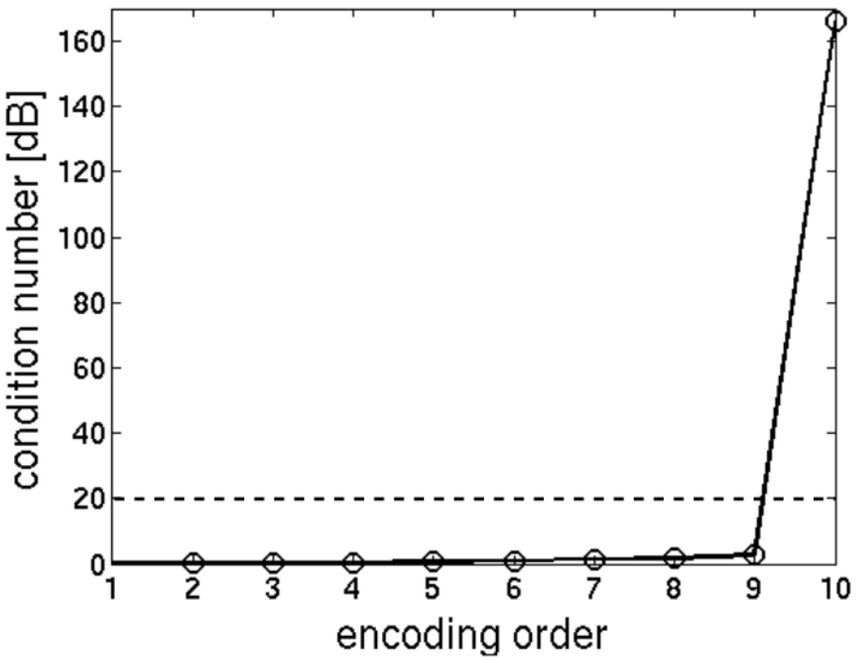
Designing of spherical arrangement with 121 channels



Spec of 121 spherical microphone array

- Estimation of encoding order
  - In 121 ch system, up to 10th order can be calculated $(10+1)^2 = 121$
  - Condition number of encoding matrix
    - Maximum eigenvalue / minimum eigenvalue
  - If condition number is less than 20 dB, stable encoding can be realized

M. Noisternig *et al.* 2010



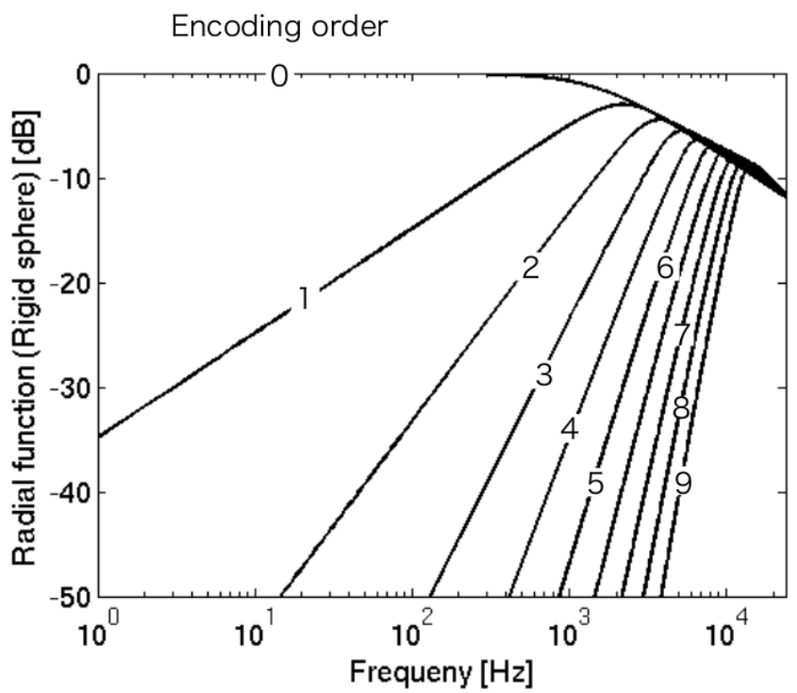
- Up to 9th order encoding
  → highest order in the world

- Radial functions
  - Relationship between radial function of rigid sphere and encoding order

$$W_m(kR) = i^{-m+1}(kR)^2 h_m^{-\prime}(kR)$$

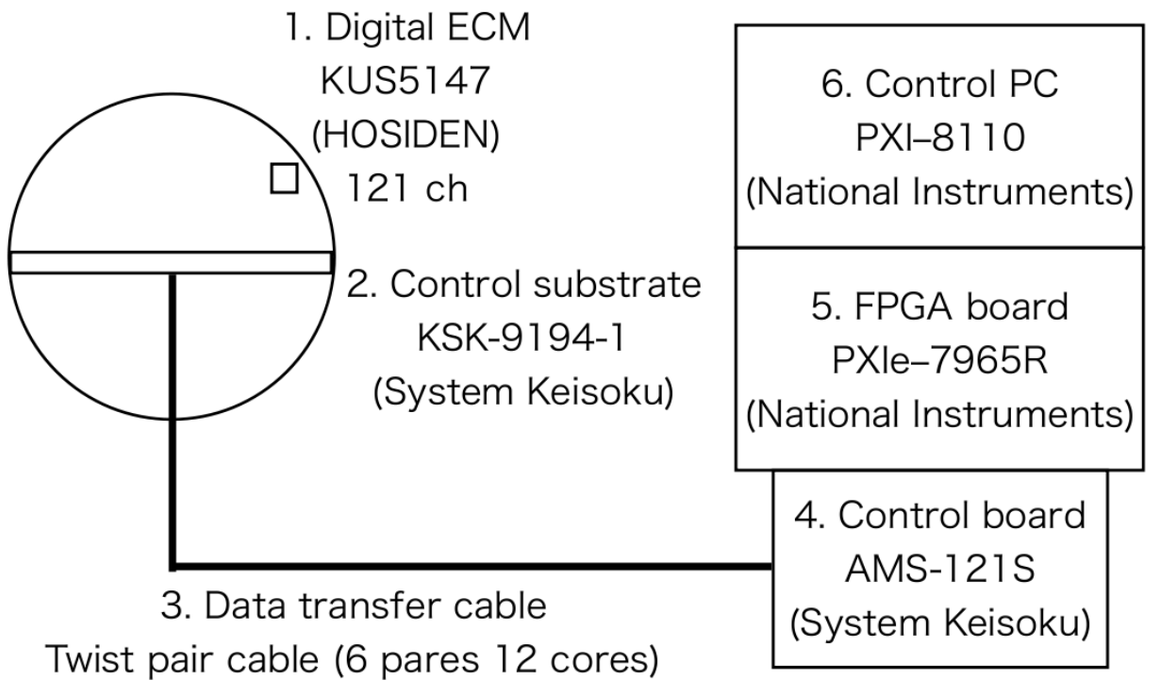
$h_m^{-\prime}(kR)$: Hankel function

- In higher encoding order, sound information in higher frequency components are much



Implementation using 121 Digital ECMs and FPGA

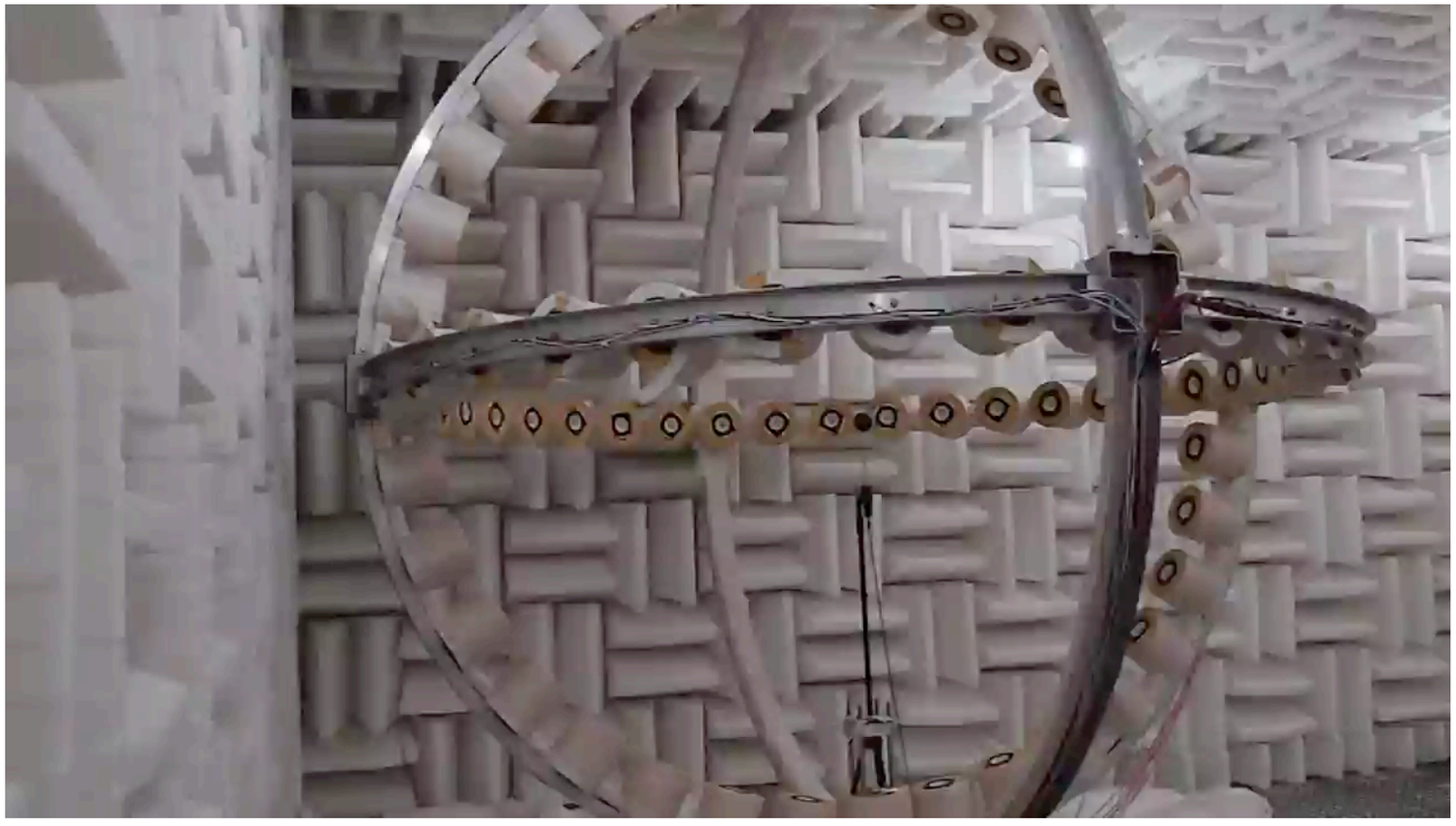
- 121 ch recording system
  - Controlled by LabVIEW (National Instruments) in Windows XP (SP3)
  - Completely synchronous 121-audio recording system using Digital Electric Condenser Microphone (Digital ECM) and FPGA board
    - Sampling frequency is 48 kHz
    - 4th Delta-Sigma modulation is used for 1-bit signals



Concluding remarks

- Conclusions
  - Implementation of 3D sound field recording system based on HOA
    - 9th order encoding, which is highest order in the world, was realized
    - Completely synchronous 121-audio recording system using Digital Electric Condenser Microphone (Digital ECM) and FPGA board

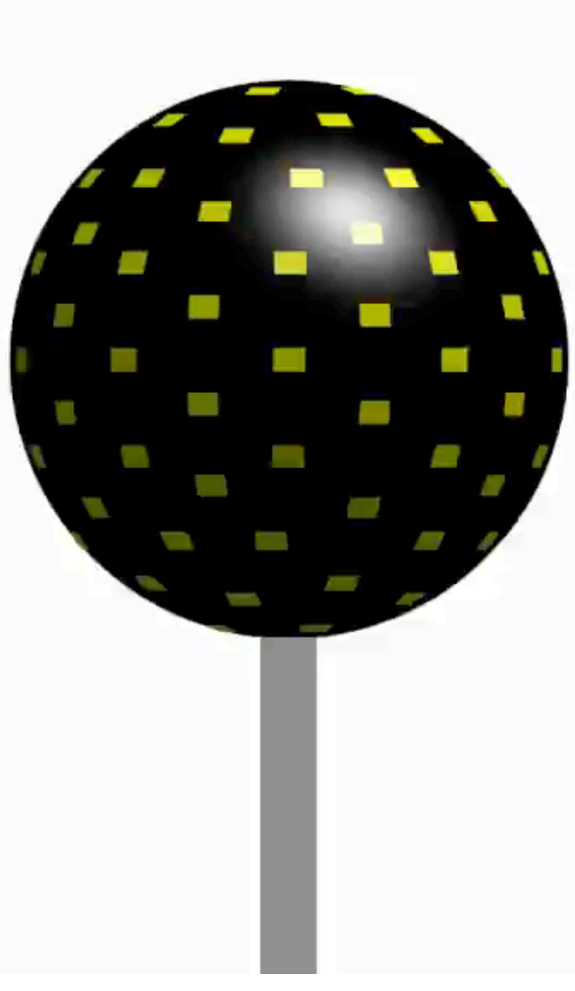
- Future works
  - System calibration
    - Acoustical measurement in anechoic chamber using spherical loudspeaker array
  - Performance evaluation
    - Recording sound field in actual environment and reproducing in our 157-loudspeaker array



Spherical loudspeaker array in R.I.E.C

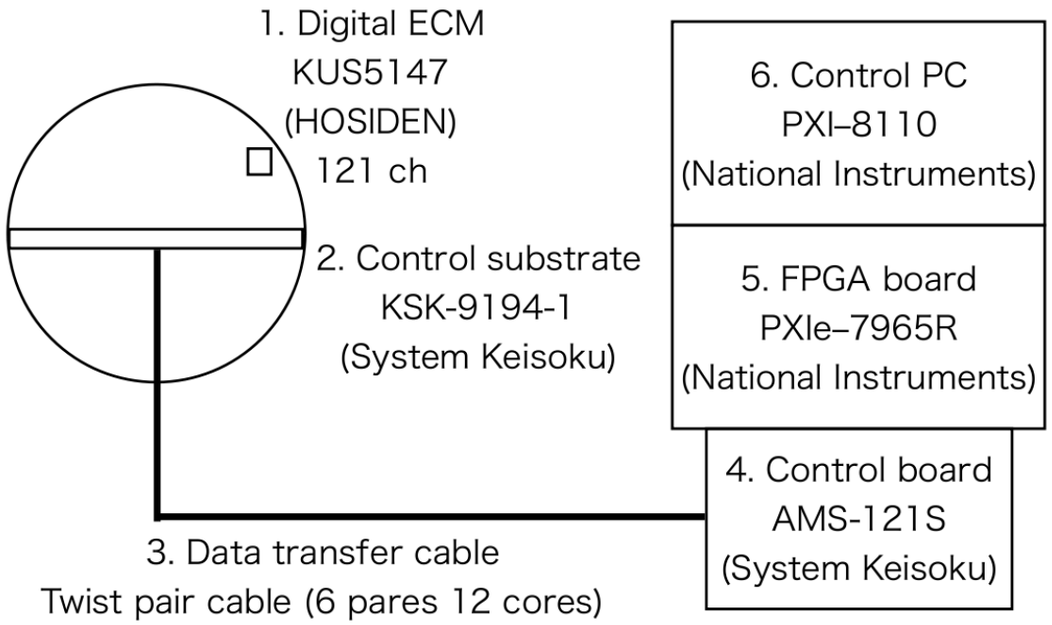
Acknowledgements

- Mr. Y. Shigeno (HOSIDEN)
  - for contribution of Digital ECMs
- Mr. J. Trevino (in our lab., MC2)
  - for making the animation movie of rotative 121 spherical microphone array
- Mr. J. Matsunaga (in our lab., MC1)
  - for cooperation to take acoustical measurements for system calibration
- Mr. Y. Sato (in our lab., MC1)
  - for cooperation to record several movies
- GCOE program (CERIES)
- Kakenhi (no.19001004) from the JSPS



Signal format at each device

1. 121 Digital ECMs
   - 48 kHz x 64 = 3.072 MHz, 4th ΔΣ modulation, 16 bit
2. Control substrate
   - Transforming time-division serial data using Ser/Des communication
   - 3 pairs of Ser/Des transmission
     - 3.072 MHz × 48 = 147.457 bps (1 pair)
3. Data transfer cable
   - 3 pairs of Ser/Des transmission = 442.368 bps
4. Control board
   - Decomposing time-division serial data to multipoint parallel data
   - Input signals to FPGA board
   - Sampling frequency: 3.072 MHz, 1 bit
5. FPGA board
   - Transforming signals using Cascaded Integrator Comb (CIC) filter
   - Sampling frequency: 48 kHz, 16 bit
6. Control PC
   - Recording 48 kHz x 121 ch data to HDD



※SerDes
SERializer/DESerializer