Aalto University
School of Science
and Technology

Broadband Direction Estimation Method utilizing Combined Pressure and Energy Gradients from Optimized Microphone Array

Jukka Ahonen and Ville Pulkki

Department of Signal Processing and Acoustics, Aalto University

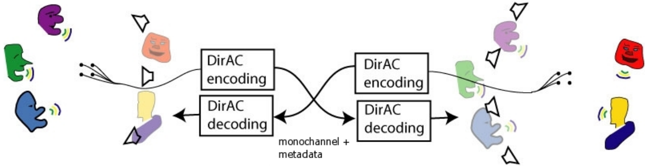
27.5.2011

Outline

- Introduction
- Background
  - Direction estimation in energetic analysis
  - Microphone array to provide required signals for analysis
- Direction estimation with PEG (pressure and energy gradients)
  - Microphone array for PEG method
- Microphone array optimization for PEG method
- Evaluations
- Summary

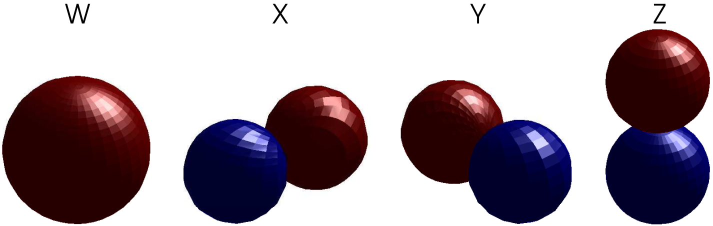
Introduction

- Direction of arrival (DOA) of sound
  - Source localization, beamforming/spatial filtering
  - Parametric spatial audio coding
- Large scale of methods for estimating DOA
  - MUSIC, ESPRIT, time differences of microphone signals, ...
- DOA based on sound intensity vector (energetic analysis)
  - Applied here with Directional Audio Coding (DirAC) for recording and reproducing spatial sound



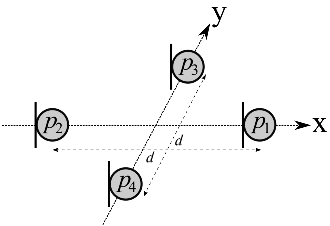
Direction estimation in energetic analysis

- Based on active sound intensity vector $\boldsymbol{I}_a$
  - Direction and magnitude of the net flow of sound energy
  - $p \times \vec{\mathbf{u}}$ (pressure times particle velocity)
- DOA$(t, f) = \frac{-\mathbf{I}_a(t,f)}{\|\mathbf{I}_a(t,f)\|}$ (opposite dir. to intensity vector)
- Pressure and velocity with B-format microphone signals



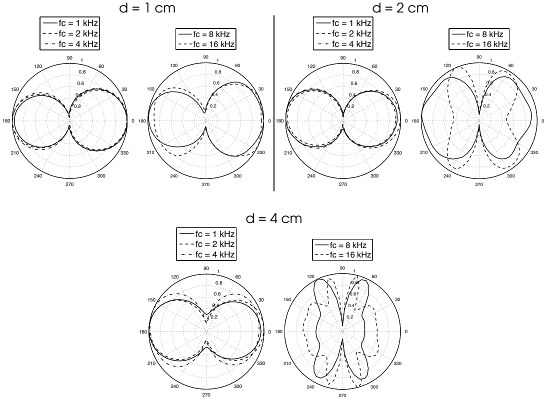
Square microphone array

- Four omnidirectional microphones in square
  - Miniature microphones
  - Mics. placed a few centimeters apart from one another
- Horizontal B-format signals (omni W, dipoles X and Y)
  - Pressure gradients: X = p_1 - p_2, Y = p_3 - p_4
  - W = $\mathbf{mean}(p_1, p_2, p_3, p_4)$



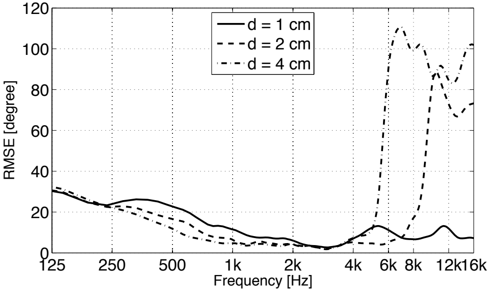
Issues

- Deformed dipoles above spatial-aliasing frequency f_{sa}
  - Depend on spacing, $f_{\mathrm{sa}} = \frac{c}{2d}$ (c = speed of sound)



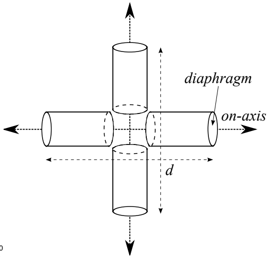
Issues

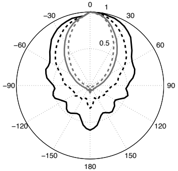
- Spatial aliasing
  $\Rightarrow$ erroneous DOA estimation at high frequencies
- Self-noise of microphones
  $\Rightarrow$ erroneous DOA estimation at low frequencies
- Error for DOA estimation expressed as RMSE

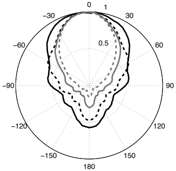


Microphone array providing acoustic shadowing

- Array of omnidirectional microphones with large housing
  - ⇒ Shadowing at high frequencies
  - ⇒ Inter-microphone level differences
- Could shadowing be utilized in energetic analysis to estimate DOA?







Left: dir. patterns for AKG mic. (diaph 2.1cm) at 5, 8, 12.5, and 16 kHz

Right: dir. patterns for G.R.A.S mic. (diaph 1.27cm) at 8, 10, 12.5, and 16 kHz

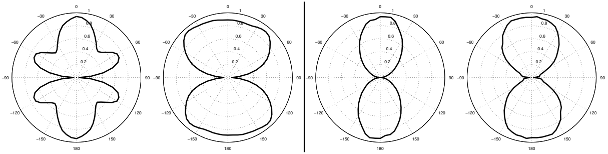
Direction estimation using energy gradients

- Sound intensity approximated using energy gradients

$$\widetilde{I}_x(t,f) = |P_1(t,f)|^2 - |P_2(t,f)|^2$$
$$\widetilde{I}_y(t,f) = |P_3(t,f)|^2 - |P_4(t,f)|^2$$

- Dir. patterns at 12 kHz (1/3-oct. band)

Pressure gradient | Energy gradient



AKG | G.R.A.S | AKG | G.R.A.S

- PEG method:

  Intensity vectors with pressure and energy gradients at low and high frequencies, respectively ⇒ DOA from intensity vectors

Spacing optimization of microphone array

- Idea: match spatial-aliasing frequency f_{sa} with the frequency-limit f_{lim} for using the energy gradients
  - f_{sa} depends on inter-microphone distance
  - f_{lim} depends on diaphragm size of the microphone
- Directivity index (DI): on-axis pickup energy vs. total pickup energy (integrated over all directions)

$$DI(f, ka) = 10 \log_{10}(\Delta L(f, ka))$$

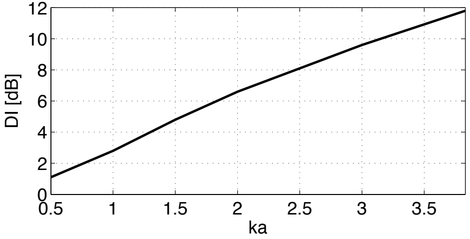
- DI depends on ratio of diaphragm circumference and wavelength:

$$ka = \frac{2\pi r}{\lambda},$$

where r is the radius of the diaphragm and $\lambda = c/f_{\mathrm{lim}}$

Spacing optimization of microphone array

- DI as a function of ka for omnidirectional microphone

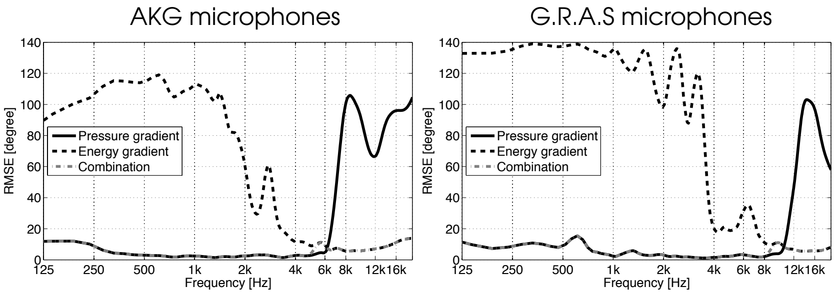


- Optimized spacing between microphones in array:

$$d_{\mathbf{opt}} = \frac{\pi r}{ka}$$

Evaluation

- Evaluation of PEG method and spacing optimization
  - Array of AKG mics. (diaph 2.1 cm) ⇒ spacing d_{opt} = 3.3 cm
  - Array of G.R.A.S mics. (diaph 1.27 cm) ⇒ spacing d_{opt} = 2 cm
- Error for DOA estimation expressed as RMSE



Summary

- DOA of sound estimated from sound intensity vectors (p and $\vec{\mathrm{u}}$ from B-format signals)
- Array of four omni microphones providing B-format signals
  - X- and Y-dipoles as pressure gradients
  - Spatial aliasing $\Rightarrow$ erroneous DOA estimation at high freq.
- Broadband DOA estimation method (PEG) proposed
  - Energy gradients $\Rightarrow$ overcomes spatial aliasing issue
  - Inter-mic. level differences (caused by shadowing of mics)
- Method to optimize mic. array for PEG
  - Inter-mic. distance based on the size of diaphragm
- Evaluation: PEG produces reliable DOA estimation at the entire audio freq. range using optimized mic. array

Acknowledgment

The Academy of Finland (Projects #105780 and #119092) and Fraunhofer Gesellshaft IIS have supported this work. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [240453].