

Introduction

Noise reduction using multi-microphone arrays with beamforming algorithms is a powerful means for the enhancement of speech in ambient noise. Current automatic speech recogniser (ASR) still need high SNR to perform accurately. Hence, a close distance between user and microphone of the ASR is usually required. In the ambient assistive living (AAL) project "Design of Environments for Ageing", an acoustical interface for the interaction between users and assistive systems in their home environment is developed, including an ASR system for user input. In order to allow the users to interact with the system at any position in their home, a special, ambient system for the acquisition of acoustical signals is being developed. It consists of two spherical microphone arrays and algorithms for localisation and beamforming for SNR improvement. The noise-reduced signal is provided to the ASR system.

Background: The AAL project „GAL“

The Lower Saxony research network "Design of Environments for Ageing" ("Gestaltung Altersgerechter Lebenswelten" – GAL) deals with information and communication technologies for promoting and sustaining quality of life, health and self-sufficiency in the second half of life.

Objective: quality of life in the ageing society

- independence within one's own residence
- development of systems for assisting older people, relatives and caregivers
- identification of threats
- support of care structures



Approach: interdisciplinary research

- synergy of geriatrics, gerontology, economics, computer science, engineering, medicine, nursing science and special needs education
- survey of requirements and resources
- development and evaluation of exemplary assisting systems

Expected Outcome: Four exemplary assisting systems

Scenario 1: Personal activity and household assistant

Assisting system for everyday planning of activities and housekeeping



Scenario 2: Monitoring of sports activities in prevention & rehabilitation

Monitoring of relevant vital parameters



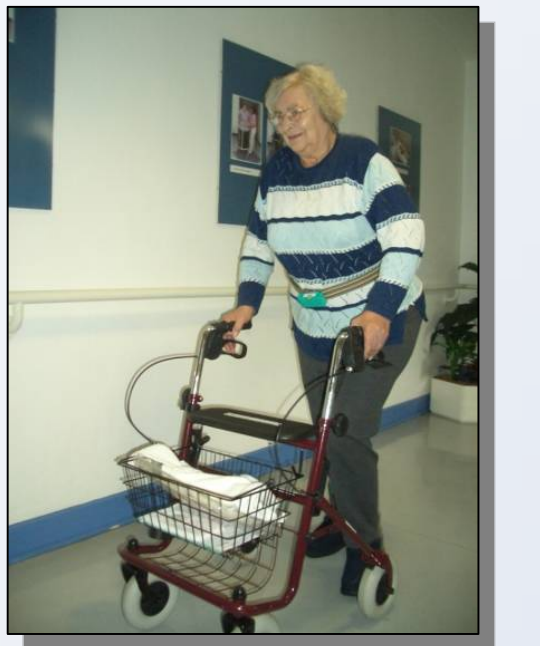
Scenario 3: Sensor-based activity determination

Comprehensive, automatic and continuous determination of activities at home



Scenario 4: Sensor-based fall prevention and fall recognition

Automatic recognition of falls and risk of falling



Acoustical front-end of the assisting systems

For the acoustical interaction between user and assisting systems, an acoustical front-end is developed.

Components of the front-end

- microphone array for signal acquisition
- automatic source localisation
- noise reduction by beamforming
- signal classification for automatic event detection
- automatic speech recognition

Demands for application in home environments

- ambient, non-intrusive integration possible (mounted microphones should be invisible / barely visible)
- low-priced equipment
- no microphone calibration required

Realisation of signal acquisition technology

Two spherical microphone arrays

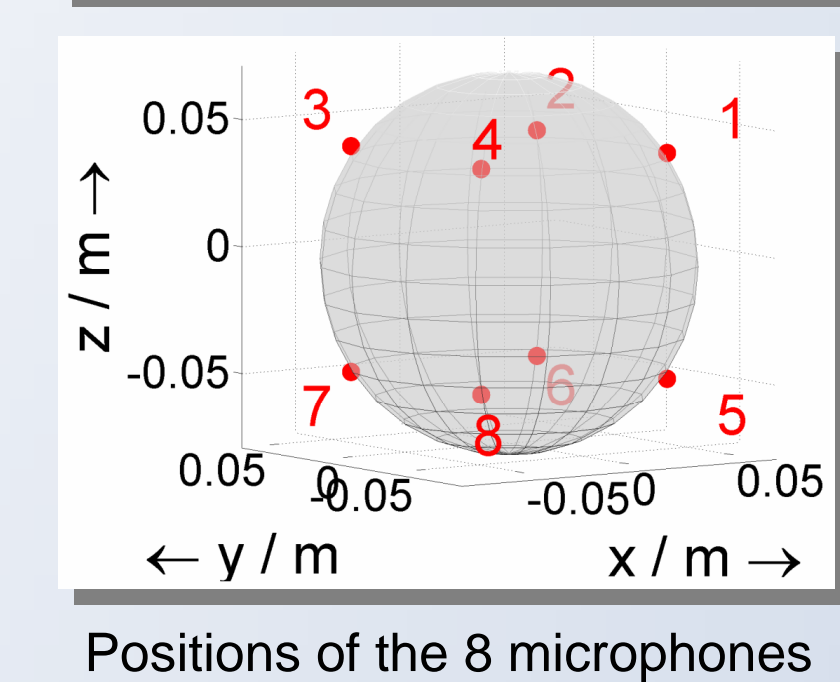
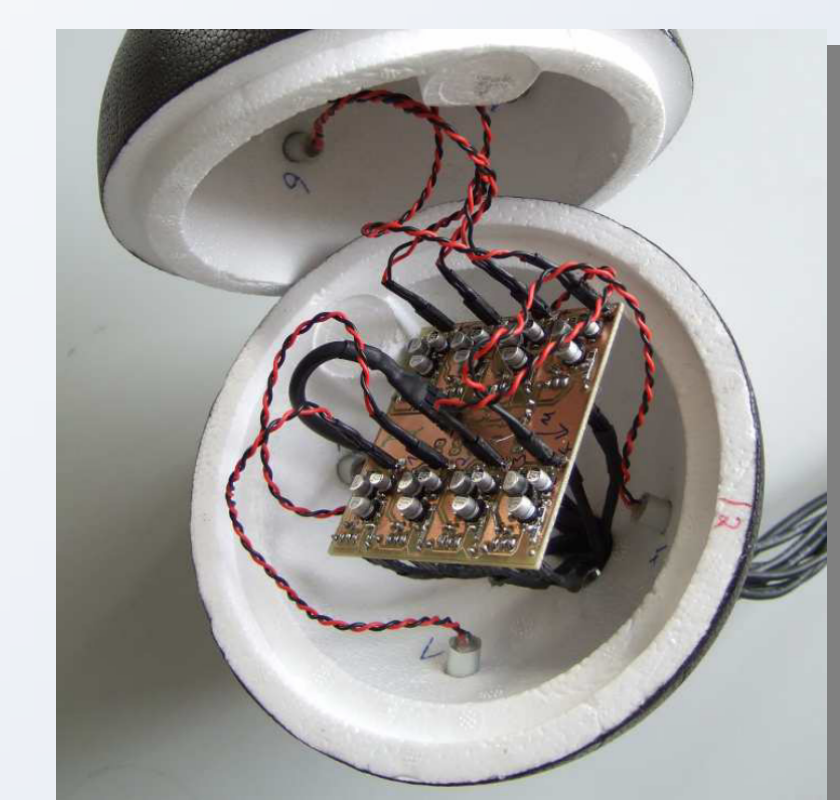
- $\varnothing = 15$ cm, 8 microphones each
- mimics head shadow effect
- cheap; no calibration needed
- ambient integration possible (e.g. in lamp)

Algorithms:

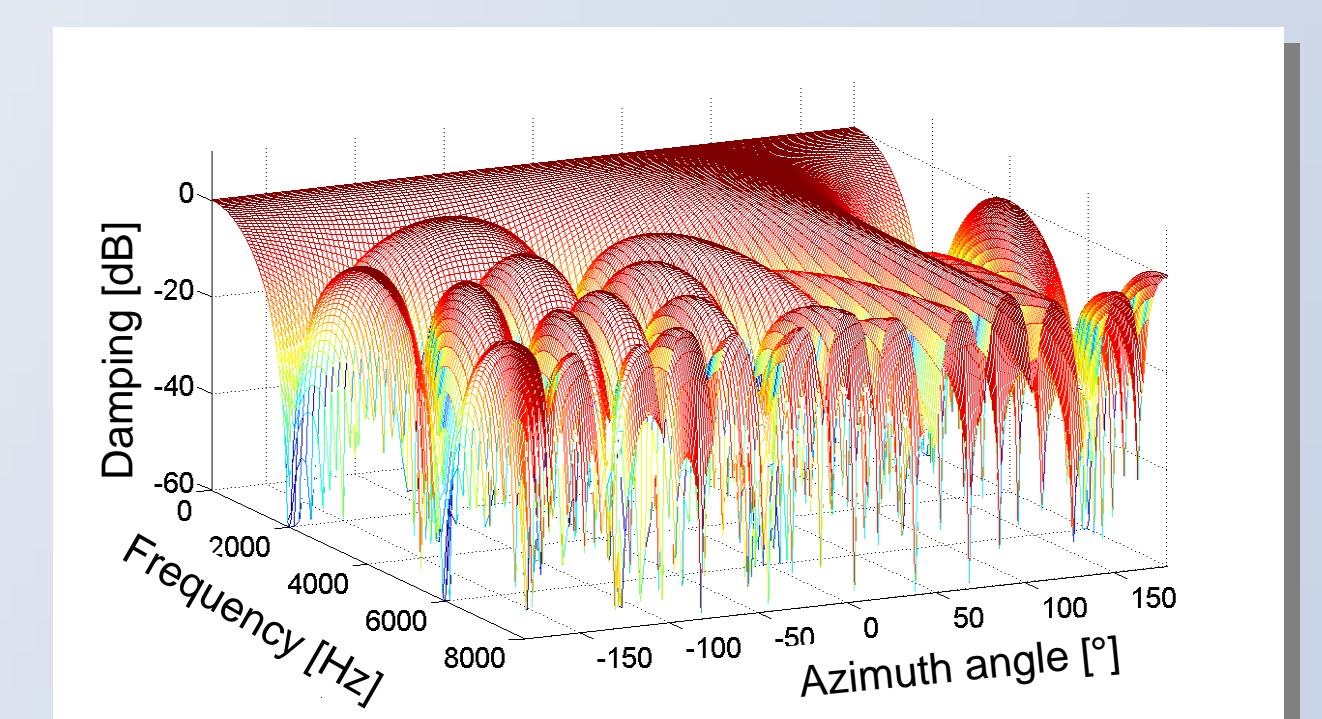
- Time Delay of Arrival estimation by
 - Generalized Cross Correlation¹ (GCC)
 - Phase-Transform spectral weighting¹ of GCC
- localisation: Global Coherence Field algorithm²
- beamforming: Minimum Variance Distortion Response algorithm³
- Voice Activity Detector⁴

First results

- localisation with both spheres within a radius of 30 cm at a hit rate of 90%
- enhanced speech intelligibility (compared to binaural listening in original sound field)



Exemplary ambient integration in lamp



Theoretical directivity pattern of array directed to 90°

Conclusions

- A cheap, ambient solution for signal acquisition including noise reduction in home environments could be found
- First results show an enhancement of speech intelligibility à benefit for ASR and hence for the AAL system is expected

Acknowledgement

This research was (partly) funded by grant VWZN2420 ("Lower Saxony Research Network Design of Environments for Ageing") from the Ministry for Science and Culture of Lower Saxony, Germany.

References

- 1) Knapp, C.H., and Carter, G.C. (1976) *The Generalized Correlation Method for Estimation of Time Delay*. IEEE Transactions on Acoustics, Speech and Signal Processing 24(4), pp. 320–327.
- 2) Brutti, A., Omologo, M., and Svaizer, P. (2008) *Localization of multiple speakers based on a two step acoustic map analysis*. Proc. ICASSP '08, Las Vegas, USA
- 3) Bitzer, J. and Simmer, K. U. (2001) *Superdirective microphone arrays*, in Brandstein, Ward (Editors), "Microphone Arrays", Springer Verlag
- 4) Marzinek, M., and Kallmeier, B. (2002) *Speech pause detection for noise spectrum estimation by tracking power envelope dynamics*. IEEE Transactions on Speech and Audio Processing, 10(2), p. 109-118.