NEWSLETTER

AALBORG UNIVERSITY

CENTRE FOR ACOUSTIC SIGNAL PROCESSING RESEARCH (CASPR)

January 2025

The Centre for Acoustic Signal Processing Research (CASPR) is a research centre at the Section for Artificial Intelligence & Sound, Department of Electronic Systems, Aalborg University, Denmark. CASPR is supported by the Demant Foundation, Oticon A/S, and Aalborg University.

If you are interested in learning more about the research and teaching taking place in CASPR:

Check our webpage at:

http://caspr.es.aau.dk

Reach out to:

- Professor Jan Østergaard (jo@es.aau.dk)
- Professor Zheng-Hua Tan (zt@es.aau.dk)
- Professor Jesper Jensen (jje@es.aau.dk).

Contents

CASPR

Centre for Acoustic Signal Processing Research

RESEARCH IN CASPR

ABOUT CASPR

CASPR is conducting research related to advanced statistical signal processing solutions for assistive listening. The research finds direct use in communication devices such as hearing aids, helmets, headphones, cochlear implants, ear monitor, ear pieces, diagnostic equipment, etc. However, the envisioned research also finds use in related areas such as robust speech/speaker recognition, brain computer interfaces, acoustic event detection, etc.

CASPR will navigate in a rapidly changing technological landscape: we envision a near future, where the technological landscape allows very different, and better, hearing assistive devices than are known today.

The scientific scope of CASPR encompasses:

- Signal processing for hearing assistive devices that use multiple modalities such as sound, vision, EEG, and other body signals.
- Advanced signal processing concepts in closer symbiosis with the user in order to e.g., automatically adapt to the user's needs.
- Beyond audibility. Restore audibility, increase intelligibility, decrease listening effort.

SCIENTIFIC DISCIPLINES OF CASPR

Machine Learning

- Deep Neural Networks
- Automatic Speech Recognition
- Multimodal Biometric
- Brain-Computer Interfaces
- Big Data

Information and Communication Theory

- Information in the brain and the auditory system
- Steaming of sound
- Compression of sound
- Wireless acoustic sensor networks

Statistical Signal Processing

- Speech processing
- Signals in noise
- Spatial signal processing
- Multimodal signal processing
- EEG signal processing
- Processing of body worn sensor data

Signal Processing and Perception

- Signal processing that compensates auditory system
- Signal processing that complements auditory system

TEACHING IN CASPR

CURRENT COURSES RELATED TO CASPR

CASPR is heavily involved in teaching and education at B.Sc., M.Sc., PhD., and Postdoc level in disciplines that are relevant to the scientific scope of CASPR:

- Machine learning, Master and PhD courses.
- Deep learning, Master and PhD courses.
- Self-Supervised Learning, PhD course.
- CASPR Seasonal Schools on Signal Processing for Hearing Assistive Devices.
- Information theory, Master course.

DURING AUTUMN 2024, CASPR WAS INVOLVED IN THE FOLLOWING STUDENT PROJECTS

- Quantifying speech comprehension using Large Language Models. Mathematical Engineering – joint 9+10th semesters long-thesis and 4+4 PhD – Sif B. Lindby.

- Speech Enhancement using Sequence Modelling Neural Architectures and Spoken Large Language Models. Mathematical Engineering – joint 9+10th semesters long-thesis and 4+4 PhD - Nikolai Lund Kühne.

- Active/Adaptive Noise Cancellation. Electronic Systems, 9+10th semester long thesis – Jacob El-Omar.

- ML based classification of audio and. speech codecs. Computer Engineering, 9th semester – Simon Loi Baks.

- Shaping the distribution of the quantization noise of lattice vector quantizers. Mathematical Engineering, 9th semester – Alexander Topholm Bertelsen.

- Thinking, Understanding, and Navigating Environments with Spatial Sound using Large Language Models. Mathematical Engineering, 7th semester – Jeppe Roden Münster, Søren Pilegaard Rasmussen, Rasmus Erik Villadsen, Simon Bock Segaard, Marcus Grøn, Andreas Peter Juhl Hansen.

CASPR RESEARCH IN FOCUS

PHD PROJECT: BRAIN-INFORMED ACOUSTIC SIGNAL PROCESSING FOR HEARING ASSISTIVE DEVICES

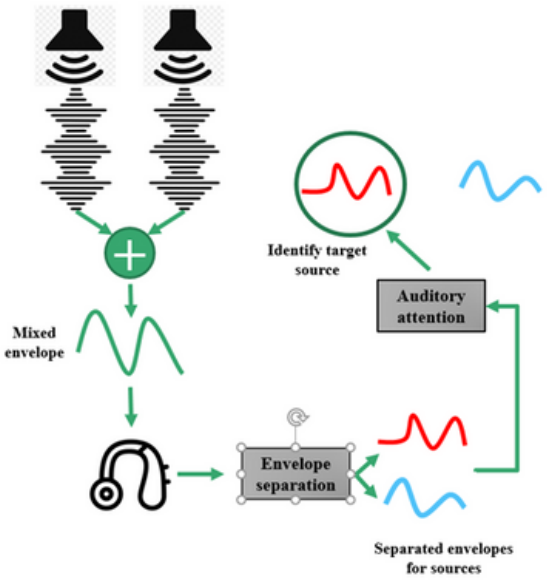
PhD researcher: Asjid Tanveer
Start date: 15.10.2023
Supervisors: Prof. Jan Østergaard, Prof. Zheng-Hua Tan, Prof. Jesper Jensen

PROJECT SUMMARY

Hearing aids play a pivotal role in improving communication for individuals with hearing impairments. Their primary function is to enhance and amplify speech signals, making conversations more accessible. Advances in EEG and hearing aids technology hold the promise of significantly improving hearing aids by integrating brain-response monitoring and feedback. For instance, hearing aids could use brain-controlled beamformers to identify a specific speaker in environments with multiple competing voices.

In particular, imagine a setting where several people are speaking at once, or where background music competes with a speaker's voice. The microphones in hearing aids capture a complex, mixed signal composed of overlapping voices, noise, and music. The challenge lies in enabling the hearing aid to isolate the target sound source—whom the listener is focusing on—while suppressing irrelevant sounds.

Using brain signals alongside the mixed acoustic inputs, this project aims at developing deep learning-based methods to identify, extract, and enhance the correct target sound source. This involves separating the mixed audio into distinct sources and determine which one aligns with the listener's auditory attention as is illustrated in the figure below. This is particularly challenging, when the competing sources originate from the same spatial direction or when the directions of the sources are changing. Solving these problems could pave the way for hearing aids that not only amplify sound but also intelligently adapt to the listener's focus, thereby improving the auditory experience.



NEWS

PhD students

Nikolai Lund Kühne started as a PhD student in CASPR on September 1st, 2024. The research topic is on Speech Enhancement using Sequence Modelling Neural Architectures and Spoken Large Language Models. The supervisors are Profs. Zheng-Hua Tan, Jan Østergaard, and Jesper Jensen.

Sif Bjerre Lindby started as a PhD student in CASPR on September 1st, 2024. The research topic is on Quantifying Speech Comprehension using Large Language Models and Brain Feedback. The supervisors are Profs. Jan Østergaard, Zheng-Hua Tan, and Jesper Jensen.

On Friday, January 17, Andreas Jonas Fuglsig successfully defended his PhD thesis: “Joint Far- and Near-end Speech and Listening Enhancement with a Minimum Processing Perspective”.

Background noise can make communication difficult, affecting both the speaker and the listener. Andreas Jonas Fuglsig’s research focuses on how to tackle these challenges by combining different methods to improve both the speaker’s and listener’s experience. His approach focuses on reducing noise, enhancing clarity, and maintaining the natural quality of speech.

Assessment Committee:
Associate Professor Jan Dimon Bendtsen (Chair), Aalborg University, Denmark
Associate Professor Richard Christian Hendriks, Delft University of Technology, The Netherlands
Assistant Professor Aki Härmä, Maastricht University, The Netherlands

Supervisors:
Professor Zheng-Hua Tan, Aalborg University, Denmark
Professor Jan Ostergaard, Aalborg University, Denmark
Project Engineer Lars Søndergaard Bertelsen, RTX
CTO Jens Christian Lindof, RTX

Advisor:
Professor Jesper Jensen, Aalborg University and Oticon A/S

NEWS

Courses

The following two courses on machine learning and information theory will be available in 2025. Participation is free of charge for PhD students and costs DKK 2000,- per ECTS for industry participants. The courses will take place physically at Aalborg University, Department of Electronic Systems.

COURSE 1: MACHINE LEARNING

ECTS: 3.0

Date/Time: 17, 19, 21, 24, 26 March 2025

Description: This course will give a comprehensive introduction to machine learning both by presenting technologies proven valuable and by addressing specific problems such as pattern recognition, prediction, clustering, generative modeling and anomaly detection. This course covers both theory and practices for machine learning, but with an emphasis on the practical side namely how to effectively apply machine learning to a variety of problems. Topics will include: Supervised learning methods: logistic regression, support vector machines, neural networks, K-nearest neighbors, decision trees, boosting. Unsupervised learning and clustering methods: K-means, Gaussian mixture models, Expectation Maximization algorithm, principal component analysis. Deep learning methods: deep neural networks, long short-term memory recurrent neural networks, convolutional neural networks, generative adversarial networks. Probabilistic graphical models. Reinforcement learning.

Prerequisites: Basic probability and statistics theory, linear algebra and basic programming skills.

Link: https://phd.moodle.aau.dk/course/index.php?categoryid=308

Lecturer(s): Prof. Zheng-Hua Tan (zt@es.aau.dk)

COURSE 2: INTRODUCTION TO INFORMATION THEORY IN NEUROSCIENCE

ECTS: 2.0

Date/Time: 7, 8, 9 April 2025

Description: In this course, we introduce information theoretic notions that are applicable to several neuroscience systems. Our focus will be on directed information measures, which are useful for establishing statistical relationships between time series data such as EEG. We will also introduce non-directed measures such as phase synchrony. You will learn about concepts such as mutual information, transfer entropy, redundant and synergistic information, connectivity matrices and coupling strengths between time series. These concepts will be demonstrated on EEG data and you will be able to apply the tools on your own real-world physiological data.

Prerequisites: Basic courses on statistics and probability theory.

Link: https://phd.moodle.aau.dk/course/view.php?id=2550

Lecturer(s): Prof. Jan Østergaard (jo@es.aau.dk) and Postdoc Payam Shahsavari Baboukani.

PUBLICATIONS

Since August 2024

CONFERENCE PAPERS

1. Detecting and Defending Against Adversarial Attacks on Automatic Speech Recognition via Diffusion Models. N. L. Kühne, A. H. F. Kitchen, M. S. Jensen, M. S. L. Brøndt, M. Gonzalez, C. Biscio, Z.-H. Tan. Accepted for IEEE ICASSP, April 2025.
2. Rate-Distortion under Neural Tracking of Speech: A Directed Redundancy Approach. J. Østergaard, S. G. Jayaprakash, R. Ordonez. Accepted for the IEEE Data Compression Conference, March 2025.
3. Near-End Listening Enhancement Using a Noise-Robust Linear Time-Invariant Filter. F. Villani, W.-Y. Chan, Z.-H. Tan, J. Østergaard, J. Jensen., Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), September 9 – 12, 2024.
4. Bayesian Sound Field Estimation Using Uncertain Data. J. Brunnstrom, M. B. Møller, J. Østergaard, and M. Moonen. Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), September 9 – 12, 2024.
5. Audio Mamba: Selective State Spaces for Self-Supervised Audio Representations. S. Yadav and Z.-H. Tan, Interspeech 2024, Kos Island, Greece, September 1-5, 2024.
6. No-Reference Speech Intelligibility Prediction Leveraging a Noisy-Speech ASR Pre-Trained Model. H. Wang, J. Jensen, I. Lopez-Espejo, W.-Y. Chan, Proc. Interspeech, September 1-5, 2024.
7. Deep Digital Joint Source-Channel Based Wireless Speech Transmission. M. Bokaei, J. Jensen, S. Doclo, J. Østergaard. Proc. 32nd European Signal Processing Conference (EUSIPCO 2024), Lyon, France, August 26–30, 2024.
8. Speaker and Style Disentanglement of Speech Based on Contrastive Predictive Coding Supported Factorized Variational Autoencoder. Y. Xie, M. Kuhlmann, F. Rautenberg, Z.-H. Tan, and R. Haeb-Umbach, The 32nd European Signal Processing Conference (EUSIPCO 2024), Lyon, France, August 26–30, 2024.
9. Envelope Based Deep Source Separation and EEG Auditory Decoding for Speech and Music, M. A. Tanveer, J. Jensen, Z.-H. Tan, J. Østergaard, Proc. 32nd European Signal Processing Conference (EUSIPCO 2024), Lyon, France, August 26–30, 2024.

JOURNAL PAPERS

1. Low-latency Deep Analog Speech Transmission using Joint Source Channel Coding. Mohammad Bokaei, Jesper Jensen, Simon Doclo, Jan Østergaard, Journal of Selected Topics in Signal Processing, Accepted 2025.
2. Noise-Robust Hearing Aid Voice Control, Ivan Lopez-Espejo, Eros Rosello, Amin Edraki, and Jesper Jensen, IEEE Signal Processing Letters, Accepted 2025
3. Hearing Loss Compensation Using Deep Neural Networks: A Framework and Results from a Listening Test. Peter Asbjørn Leer Bysted, Jesper Jensen, Laurel Carney, Zheng-Hua Tan, Jan Østergaard, Lars Bramsløw, IEEE/ACM Transactions on Audio, Speech, and Language Processing, Accepted 2025.
4. Investigating the design space of diffusion models for speech enhancement, Philippe Gonzalez, Zheng-Hua Tan, Jan Østergaard, Jesper Jensen, Tommy Sonne Alstrøm, Tobias May, IEEE/ACM Transactions on Audio, Speech, and Language Processing, p. 4486-4500. Vol.32, October 2024.
5. Identifying principal attributes for evaluating audio quality of reproduction systems with spatially dynamic program material. P. N. P. Moreta, S. Bech, J. Francombe, J. Østergaard, and S. van de Par. Journal of the Audio Engineering Society, September 16, 2024.
6. The Effect of Training Dataset Size on Discriminative and Diffusion-Based Speech Enhancement Systems. P. Gonzalez, Z.-H. Tan, J. Østergaard, J. Jensen, T. S. Alstrøm, and T. May, Vol. 31, pp. 2225-2229, IEEE Signal Processing Letters, August 2024.
7. Joint Far- and Near-end Speech and Listening Enhancement with Minimum Processing. A. J. Fuglsig, Z.-H. Tan, L. S. Bertelsen, J. Jensen, J. C. Lindof, and J. Østergaard, IEEE Access, August 2024.