

Editorial

Liebe Leserinnen und Leser,

gerade erst haben wir eine tolle ITG Fachtagung Sprachkommunikation im virtuellen Kiel erlebt. Prof. Gerhard Schmidt und sein Team haben uns unter all den schwierigen Bedingungen wunderbar zusammen gebracht. Und: Es hagelt Preise. Neben dem Best Paper Award der ITG Fachtagung gibt es dreimal einen ersten Platz in internationalen Challenges. Außerdem freuen wir uns über eine Reihe abgeschlossener Dissertationen. Fragen Sie die doch mal an!

Ihr Tim Fingscheidt & Reinhold Häb-Umbach

Sie wünschen ein Abo oder haben einen Beitrag? Sehr gerne! Bitte melden Sie sich einfach per Email unter Hinweis darauf, ob Sie nur [Abonnent](#), oder [Abonnent und auch möglicher Autor](#) sein möchten! Wir weisen aus datenschutzrechtlichen Gründen darauf hin, dass Sie unter gleicher Emailadresse jederzeit Auskunft über Ihre gespeicherten Daten erfragen können, sowie die Löschung Ihrer Kontaktdaten erwirken können.

Latest News

- Am Freitag, den 15.10.2021, fand die 72. Sitzung der [ITG-Fachgruppe "Signalverarbeitung und Maschinelles Lernen"](#) statt, leider weiterhin im Online-Format. Der Gastgeber war diesmal Bin Yang von der Universität Stuttgart. Nach einer kurzen Einführung von ihm („Die einzige Konstante im Leben ist die Veränderung“) über die Entwicklung der Forschung und Lehre in Stuttgart wurden zehn Fachvorträge aus seinem Institut über Methoden und Architekturen von statistischer Signalverarbeitung über Deep Learning bis hin zum kausalen Lernen sowie deren Anwendungen in der Medizintechnik, Sprachverarbeitung, beim autonomen Fahren und Halbleiter-Chiptest gehalten. Das Programm aus Stuttgart wurde komplettiert durch einen Beitrag über Videokomprimierung aus der Fachgruppe. Über 30 Teilnehmer nahmen an der Sitzung teil.

- Vom 29.09. bis 01.10. fand die 14. ITG-Fachtagung Sprachkommunikation statt. Pandemiebedingt wurde die Konferenz vollständig online durchgeführt – auf einer virtuellen Karte von Kiel konnten sich die Teilnehmenden durch zahlreiche Orte (Auditorium, Posterräume, einen Strand, ein Gym für aktive Pausen mit einer Sportwissenschaftlerin und eine Einkehrmöglichkeit für einen „Kieler Abend“) bewegen (siehe Foto). Großes Lob gab es für drei exzellente Keynotes von Timo Gerkmann, Juan Rafael Orozco-Arroyave und Christophe Couvreur, welche die Themenbereiche pathologische Sprache, maschinelles Lernen in der Sprachsignalverbesserung



und Sprachtechnologie in Fahrzeugen adressierten. Damit die Konferenz nun auch in Zukunft einen festen Platz im Internet hat, wurde folgende [Webseite](#) dauerhaft eingerichtet.

- Ein Team der Ruhr-Universität Bochum (Doktorand Benedikt Bönninghoff und Prof. Dorothea Kolossa) und der Bucknell University (Prof. Robert M. Nickel, USA) hat zum zweiten Mal in Folge den internationalen PAN-Wettbewerb „Authorship Verification“ gewonnen. PAN steht für Plagiarism Analysis, Authorship Identification, and Near-Duplicate Detection. Mit dem entwickelten Verfahren konnten 2020 und 2021 in allen Metriken signifikant bessere Ergebnisse erzielt werden als von den konkurrierenden Systemen. Außerdem wurde der Beitrag mit dem "Best of Lab Papers Award" ausgezeichnet.

Resultate 2020: <https://pan.webis.de/clef20/pan20-web/author-identification.html>

Resultate 2021: <https://pan.webis.de/clef21/pan21-web/author-identification.html>

Persönliches

- Auf der 14. ITG-Fachtagung Sprachkommunikation wurde Ragini Sinha mit dem Best-Paper-Award für ihren Beitrag „Speaker-Conditioned Target Speaker Extraction Based on Customized LSTM Cells“ ausgezeichnet. Das Paper wird ab Dezember auf IEEExplore verfügbar sein. Ragini Sinha studierte am IIT Mandi in Indien. Seit 2019 arbeitet sie am Fraunhofer Institutteil Hör-, Sprach- und Audiotechnologie in Oldenburg und promoviert zugleich an der Universität Oldenburg (Betreuer: S. Doclo) im Bereich Quellentrennung und audiovisuelle Sprachsignalverbesserung mittels tiefer neuronaler Netze.



- Jose A. Lopez vom Human and AI Systems Research (HAR) Lab in den Intel Labs hat den 1. Platz bei der diesjährigen IEEE Audio and Acoustic Signal Processing Society Challenge on Detection and Classification of Acoustic Scenes and Events ([DCASE 2021](#)), Task 2 erreicht [[Paper](#)]. In der internationalen Challenge werden Anomalien in Maschinengeräuschen unüberwacht detektiert.

Projekte und Aktivitäten

- Das Institut für Akustik und Sprachkommunikation der TU Dresden (Prof. P. Birkholz) stellt in einem [Video auf YouTube](#) seine Technik zur exakten Vermessung der akustischen Transferfunktion von physischen Modellen des Vokaltrakts vor.

Bücher, Dissertationen

- Simon Stone: [A silent speech interface using electro-optical stomatography](#), TU Dresden (Betreuer: P. Birkholz)

In der Arbeit wurde eine neue intra-orale Messtechnik für die Erfassung von Sprechbewegungen entwickelt und Experimente zur stillen Spracherkennung damit durchgeführt.

- Ali Aroudi: [Cognitive-driven speech enhancement using EEG-based auditory attention decoding for hearing aid applications](#), Universität Oldenburg (Betreuer: S. Doclo)

This thesis deals with the problem of identifying and enhancing the target speaker in realistic acoustic environments based on decoding the auditory attention of the listener using single-trial EEG recordings. First, the impact of different acoustic conditions on the performance of a least-squares-based AAD method is analyzed. Second, several cognitive-driven speech enhancement systems are proposed. The first system is an open-loop cognitive-driven binaural beamformer, aiming at enhancing the target speaker and suppressing the interfering speaker and background noise. The second system is a closed-loop cognitive-driven gain controller, where real-time AAD enables the listener to directly interact with an adaptive gain controller. In addition, a novel AAD method is proposed based on a state-space model, which improves the decoding performance of linear and non-linear methods using small correlation windows.

- Benjamin Cauchi: [Non-intrusive quality evaluation of speech processed in noisy and reverberant environments](#), Universität Oldenburg (Betreuer: S. Doclo)

This thesis deals with speech enhancement in noisy and reverberant environments, mainly focusing on non-intrusive measures to evaluate noise reduction

and dereverberation algorithms. First, a single-channel algorithm for joint noise and reverberation reduction is proposed, where a spectro-temporal gain is computed using a combination of a statistical room acoustics model, minimum statistics and temporal cepstrum smoothing. Second, two non-intrusive speech quality measures are proposed that combine perceptually motivated features and predicting functions based on machine learning. The results show that the speech quality measure using a recurrent neural network as predicting function outperforms existing non-intrusive measures and yields a similar performance as intrusive measures when trained and evaluated for a single category of algorithms. When trained and evaluated for several categories of algorithms, it even outperforms the intrusive benchmark measures, making it suitable for the selection of algorithms or algorithm parameters.

- Stefan Liebich: [Active Noise and Occlusion Effect Cancellation in Headphones and Hearing Aids](#), Institut für Kommunikationssysteme der RWTH Aachen (Betreuer: P. Vary)

This PhD thesis is dealing with the distorted perception of one's own voice when wearing hearing aids, hearables or headsets that occlude the ear canal. The occlusion effect which creates a hollow perception of one's own voice but also amplifies any kind of body conducted sounds, e.g., during eating, drinking, or walking, is still one of the biggest problems in devices nowadays. Dr.-Ing. Liebich investigated the effect and developed an active solution, the so-called Occlusion Effect Cancellation, which combines methods of active noise cancellation with personalized design. Main contributions are novel design concepts w.r.t. variations of ear canals and earpieces, novel combined feedforward-feedback filters with adaptive stability control, real-time implementation as well as objective instrumental and subjective auditive evaluation of the concepts. Based on these works and together with a strong team, Dr.-Ing. Liebich has founded the Elevear GmbH. Here the concepts are further developed into a licensing product.

Journalartikel

- A. Chinaev, P. Thüne, and G. Enzner [Double-Cross-Correlation Processing for Blind Sampling-Rate and Time-Offset Estimation](#)

Coherent processing of signals captured by a wireless acoustic sensor network (WASN) requires an estimation of unknown parameters such as the sampling-rate and sampling-time offset (SRO and STO) of asynchronous sensor clocks. We therefore introduce a novel approach for blind SRO and STO estimation in the spirit of both a recently available double-cross-correlation processor (DXCP) and the

generalized cross-correlation with phase transform (GCC-PhaT). For the proposed approach, called DXCP-PhaT, we specifically introduce secondary cross-quantities to restore ergodicity of the otherwise drifting cross-correlation function or cross-spectrum of asynchronous input signals. Further, we derive the DXCP-PhaT algorithm, which brings along a number of improvements over state-of-the-art: advanced estimation accuracy, environmental robustness to larger microphone distances, real-time-applicability, and reduced complexity. Those claims are confirmed by comprehensive experiments, both offline and online, on simulated data and real recordings from a Raspberry-Pi-based WASN.

- A. Nelus and R. Martin

[Privacy-Preserving Audio Classification Using Variational Information Feature Extraction](#)

In this article we investigate and tackle the privacy risks of deep-neural-network-based feature extraction for sound classification in acoustic sensor networks. To this end, we analyze a single-label domestic activity monitoring and a multi-label urban sound tagging scenario. We show that in both cases, the feature representations designed for sound classification also carry a significant amount of speaker-dependent data, thus posing serious privacy risks for speaker recognition attacks based on feature interception. We then propose to mitigate the aforementioned privacy risks by introducing a variational information feature extraction scheme that allows sound classification while, concurrently, minimizing the feature representation's level of information. Thus, it exhibits robustness against x-vector-based state-of-the-art speaker recognition attacks.

- T. Kabzinski and P. Jax

[A Causality-Constrained Frequency-Domain Least-Squares Filter Design Method for Crosstalk Cancellation](#)

Crosstalk cancellation filters are crucial for binaural reproduction over loudspeakers. While the time-domain least-squares filter design problem is too complex to be solved in a dynamic real-time setup, the frequency-domain problem is computationally cheap. However, it yields circular convolution artifacts, which degrade the performance. We derive a causality-constrained frequency-domain least-squares problem, exploiting the fact that the real part and the imaginary part of a causal signal's spectrum are related via Hilbert transform. By approximating the Hilbert transform, we reduce the proposed method's computational complexity by two to three orders of magnitude while the performance remains close to that of the time-domain solution. The causality-constrained method thus enables a flexible

tradeoff between computational complexity and performance figures.

Tagungen (nach Paper Deadline sortiert)

- [DAGA 2022](#), 21.-24.03.2022, Stuttgart, Paper Deadline: 01.11.2021 [[CfP](#)]
- [ESSV 2022](#), 02.-04.03. 2022, Sonderborg, Dänemark. Extended Abstract Deadline: 12.12.2021
- [ASRU](#), 13.-17.12.2021, online, [keine Einreichungen mehr]
- [EUSIPCO](#), 29.-02.09.2022, Belgrad, Serbien. Paper Deadline 20.02.2022 [[CfP](#)]
- [Interspeech](#), 18.-22.09.2022, Incheon, Korea Paper Deadline: noch unbekannt
- [IWAENC 2022](#), 05.-08.09.2022, Bamberg, Paper Deadline: 22.04.2022 [[CfP](#)].
- [ICASSP 2022](#), 22.-27.05.2022, Singapur [keine Einreichungen mehr]
- [ITG Conference on Speech Communication 2023 Herbst 2023](#) in Aachen. Die Konferenz findet nun in allen **ungeraden** Jahren statt.

Stellenanzeigen

- Am Institut für Kommunikationsakustik der Ruhr-Universität Bochum ist die Stelle eines Akademischen Rats/Rätin (m/w/d) zum nächstmöglichen Zeitpunkt zu besetzen. Die Ausschreibung zu dieser Stelle findet sich [hier](#).
- Das Institut für Nachrichtentechnik der TU Braunschweig sucht laufend wiss. Mitarbeiter*innen (TV-L E13) in den Forschungsbereichen Speech und Vision. [[Kontakt](#)]