

## Editorial

Liebe Leserinnen und Leser,

Highlight dieses Newsletters ist sicherlich, dass wir Prof. Reinhold Hüb-Umbach (Universität Paderborn) zum Distinguished Lecturer der ISCA gratulieren können, eine hohe Auszeichnung und Anerkennung seiner Forschungsleistungen. Dann gibt es z.B. eine Menge verschiedenster (guter) Nachrichten aus Oldenburg und die Meldung, dass unsere ITG-Fachtagung „Sprachkommunikation“ statt vor Ort in Kiel diesmal virtuell stattfinden wird, Gerhard Schmidt aber Überraschungen bereithalten wird. Wir sind gespannt und freuen uns auf die Begegnung ...

Ihr Tim Fingscheidt

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

- Für die ITG-Fachtagung „Sprachkommunikation“ wurde die Entscheidung getroffen, die vom 29.09.-01.10.2021 geplante Konferenz nicht vor Ort in Kiel, sondern rein virtuell stattfinden zu lassen. Dieser Beschluss wurde von der Tagungsleitung gemeinsam mit den Mitgliedern des Programmausschusses insbesondere vor dem Hintergrund gefasst, dass bei fortbestehenden Abstandsregelungen eine hybride Tagungsform nicht zufriedenstellend realisierbar wäre. Durch diese Entscheidung zum jetzigen Zeitpunkt kann die verbleibende Zeit bis zur Tagung nun dafür genutzt werden, eine ansprechende Onlinepräsenz der Tagung zu entwickeln. Worauf man sich freuen kann: ca. 50 thematisch breit gestreute Beiträge, drei spannende Keynotes und natürlich ein paar kleine Überraschungen. Nähere Informationen zum Programm und zur Registrierung werden voraussichtlich Mitte Juli auf [www.itg-speech-communication.de](http://www.itg-speech-communication.de) veröffentlicht.

- Acoustic Echo Cancellation Challenge at Interspeech 2021: Acoustic echo cancellation (AEC) plays an essential part in today's VoIP speech communication and video conferencing systems. Due to the vast variety of the acoustic setup, interferences such as non-linear echoes or background noise can be quite prominent. To compensate for these effects, Evolve.tech and Graz University of Technology proposed a cross-domain echo controller (CDEC) model. It consists of a time-delay compensator, a frequency-domain block-based AEC, and a small

neural network which operates in the time domain. The hybrid structure achieves an overall system delay of 32ms, and an average MOS score of 4.04. This solution was awarded third place at the Interspeech 2021 AEC-Challenge.

## Persönliches

- *Reinhold Hüb-Umbach* (Universität Paderborn) wurde von der ISCA zum Distinguished Lecturer (DL) für die Periode 2021/2022 ernannt. Ein DL gilt als Botschafter der ISCA zur Förderung des Interesses und des Engagements für Forschungsthemen rund um Sprachtechnologien. Dazu fördert die ISCA Vortragsreisen insbesondere in Länder, die auf der internationalen Bühne der Forschung an gesprochener Sprache noch unterrepräsentiert sind.

Reinhold Hüb-Umbach hofft natürlich, dass solche Reisen bald wieder möglich sein werden.

- Ina Kodrasi wurde für ihre Dissertation [“Dereverberation and noise reduction techniques based on acoustic multi-channel equalization”](#) (Universität Oldenburg, Betreuer: Prof. Dr. Simon Doclo) mit dem *EURASIP Best PhD Award* ausgezeichnet. Dr.-Ing. Kodrasi leitet zurzeit die Arbeitsgruppe [Signal Processing for Communication](#) am IDIAP Forschungsinstitut in der Schweiz.

## Projekte und Aktivitäten

- Since January 2021, the MSCA European Training Network [“Service-Oriented, Ubiquitous, Network-Driven Sound” \(SOUNDS\)](#) started its activities (partners: KU Leuven, Imperial College, Universität Aalborg, Universität Oldenburg, Fraunhofer IDMT, Bang & Olufsen, Oticon, Nuance, Cerence). Inspired by the ubiquity of mobile and wearable devices capable of capturing, processing, and reproducing sound, the SOUNDS network aims to bring audio technology to a new level by exploiting network-enabled cooperation between devices. Envisaged applications are hearing assistance, immersive audio reproduction and advanced voice control.

- Am 26. Februar wurde die 71. Sitzung der ITG-Fachgruppe [„Signalverarbeitung und maschinelles Lernen“](#) gemeinsam von der Wehrtechnischen Dienststelle für Schiffe und Marinewaffen, Maritime Technologie und Forschung (WTD 71) durch Dr. Uwe Kretschmer und Dr. Jan Abshagen und der Christian-Albrechts-Universität zu Kiel ausgerichtet. Fokus der Sitzung waren Anwendungen aus dem maritimen Bereich der Signalverarbeitung, der Messtechnik und des maschinellen Lernens. Neben sechs Vorträgen zu SONAR-Themen kam auch die Unterwasser-

Sprachsignalverarbeitung durch einen Beitrag von Owe Wisch (Uni Kiel) nicht zu kurz. Pandemiebedingt wurde die Sitzung online durchgeführt. Alles in allem hat alles gut funktioniert und es wurden zur 71. Sitzung erstmals auch Studierende zugelassen. Eine der studentischen Teilnehmerinnen, Karoline Seidel, sagte im Anschluss an die Sitzung: „Ich habe am ITG-Fachgruppentreffen teilgenommen, da ich gerade meine Abschlussarbeit im Bereich der Unterwasser-Signalverarbeitung schreibe und sehr an den Themen interessiert bin. Vor allem die Vielfältigkeit der in den Vorträgen behandelten Themen und wie an diesen theoretisch und praktisch geforscht wird, fand ich wirklich spannend.“ Die nächste Sitzung wird am 15. Oktober in Stuttgart, ausgerichtet von Prof. Bin Yang, stattfinden. Auch hier sind wieder neben den Mitgliedern der Fachgruppe auch interessierte Doktoranden und Studierende willkommen.

### Bücher, Dissertationen

- In der Arbeitsgruppe für Kognitive Signalverarbeitung gab es in den letzten zwei Monaten gleich zwei neue Dissertationen zu feiern: [Lea Schönherr](#) entwickelte im [CASA-Cluster](#) Methoden zur Evaluierung und Verbesserung der Robustheit von Sprach- und Sprechererkennung: “Adversarially Robust Speech and Speaker Recognition”, (Betreuer: D. Kolossa und T. Holz) und ...
- [Dennis Orth](#) promovierte in einem gemeinsamen Projekt mit dem Honda Research Institute Europe zu intelligenter [Fahrerassistenz](#), mit dem Titel “Gap After the Next Vehicle – Investigation and Realization of Speech-Based, Personalized Intersection Assistance” (Betreuer: D. Kolossa und M. Heckmann).

- Samy Elshamy: [Speech Enhancement Exploiting the Source-Filter Model](#), Technische Universität Braunschweig (Betreuer: T. Fingscheidt)  
This PhD thesis presents speech enhancement methods based on the source-filter model. The core novelty are simple manipulation methods of the residual signal in the cepstral domain, called cepstral excitation manipulation (CEM), robustly allowing to suppress (non-stationary) noise between harmonics. Another quite surprising finding is that already a simple F0 estimator w/o additional algorithmic latency provides robust control of CEM in non-stationary environments even at SNRs below 0 dB. Moreover, spectral envelope enhancement estimators are proposed. Some of the methods including CEM are used in millions of smartphones today.

- Nico Gößling: [Binaural Beamforming Algorithms and Parameter Estimation Methods Exploiting External Microphones](#), Universität Oldenburg (Betreuer: S. Doclo)

This PhD thesis presents advanced binaural beamforming algorithms, aiming at simultaneously performing noise reduction and preserving binaural cues by exploiting external microphones. For the binaural MVDR beamformer with partial noise estimation, it is shown that incorporating an external microphone allows to significantly increase the noise reduction performance compared to using only the head-mounted microphones, while preserving the spatial impression of the background noise. In addition, computationally efficient methods to estimate the relative transfer function (RTF) vectors of the desired source are proposed, exploiting one or more external microphones that are spatially separated from the head-mounted microphones. Experimental results for a moving desired source in a reverberant environment show that the proposed methods are applicable in realistic and highly dynamic acoustic scenarios, and outperform state-of-the-art RTF vector estimation methods at a much lower computational complexity.

### Journalartikel

- R. Haeb-Umbach, J. Heymann, L. Drude, S. Watanabe, M. Delcroix, T. Nakatani: [Far-Field Automatic Speech Recognition](#)

Übersichtsartikel zur automatischen Spracherkennung, wenn der Abstand zwischen Sprecher und Mikrofon(en) groß ist.

- N. Gößling, E. Hadad, S. Gannot, S. Doclo [Binaural LCMV Beamforming With Partial Noise Estimation](#)

Besides reducing undesired sources, i.e., interfering sources and background noise, another important objective of a binaural beamforming algorithm is to preserve the spatial impression of the acoustic scene, which can be achieved by preserving the binaural cues of all sound sources. Aiming at merging the advantages of the binaural linearly constrained minimum variance (BLCMV) beamformer and the binaural MVDR beamformer with partial noise estimation (BMVDR-N), in this paper we propose the BLCMV beamformer with partial noise estimation (BLCMV-N). Experimental results using recorded signals as well as the results of a perceptual listening test show that the BLCMV-N is able to preserve the binaural cues of an interfering source (like the BLCMV), while enabling to trade off between noise reduction performance and binaural cue preservation of the background noise (like the BMVDR-N).

- K. Tesch, T. Gerkmann: [Nonlinear Spatial Filtering in Multichannel Speech Enhancement](#)

The paper shows via theoretical and experimental analysis that in non-Gaussian noise, the traditional concatenation of a linear beamformer and a postfilter is not optimal but outperformed by a nonlinear joint spatial-spectral filter. This is an exciting motivation for research into DNNs for nonlinear spatial processing as opposed to the mere DNN-based parameter estimation often applied in multichannel speech enhancement. Do not miss the impressive sound examples [here](#).

- R. Rehr, T. Gerkmann: [SNR-Based Features and Diverse Training Data for Robust DNN-Based Speech Enhancement](#)

The paper shows how normalizing the input of a neural network using conventional PSD estimators reduces the amount of required training data and increases the generalizability of DNN-based speech enhancement.

- D. Fischer, S. Doclo: [Robust Constrained MFMVDR Filters for Single-Channel Speech Enhancement Based on Spherical Uncertainty Set](#)

Aiming at exploiting speech correlation across consecutive time frames in the STFT domain, the multi-frame minimum variance distortionless response (MFMVDR) filter for single-channel speech enhancement has been proposed. The MFMVDR filter requires an accurate estimate of the speech correlation vector in order to avoid speech distortion and artifacts. In this paper we investigate the potential of using robust MVDR filtering techniques to estimate the speech correlation vector as the vector maximizing the signal output power within a spherical uncertainty set. To set the upper bound of the spherical uncertainty set, we propose to use a trained non-linear mapping function that depends on the a-priori SNR. Experimental results for different speech signals, noise types and SNRs show that the proposed constrained approaches yield a more accurate estimate of the speech correlation vector than a state-of-the-art maximum-likelihood (ML) estimator. An instrumental and a perceptual evaluation show that both constrained MFMVDR filters lead to less speech and noise distortion but a lower noise reduction than the ML-MFMVDR filter.

## Tagungen (nach Paper Deadline sortiert)

- [HSCMA](#), June 2021 gecancelt,  
für neues Datum Webseite checken bitte.
- [DAGA 2021](#), 15.-18.08.2021, Wien, Österreich,  
**hybrid**,  
[keine Einreichungen mehr]
- [EUSIPCO](#), 23.-27.08.2021, Dublin, Irland, **virtuell**  
[keine Einreichungen mehr]
- [Interspeech](#), 30.08.-03.09.2021, Brno,  
Tschechien, **hybrid**  
[keine Einreichungen mehr]
- [ITG Conference on Speech Communication 2021](#)  
29.09.-01.10.2021 in Kiel, **virtuell**,  
[keine Einreichungen mehr]
- [ICASSP 2022](#), 23.-27.05.2022, Singapur  
Paper Deadline: 01.10.2021 [\[CfP\]](#)
- [ASRU](#), 13.-17.12.2021, Cartagena, Kolumbien,  
[keine Einreichungen mehr]
- [IWAENC 2022](#), **05.-08.09.2022**, Bamberg,  
Paper Deadline: 22.04.2022 [\[CfP\]](#).

## Stellenanzeigen

- Der Arbeitsbereich [Signal Processing \(SP\)](#) der Universität Hamburg sucht einen Postdoc (w/m/d) auf einer 100% TV-L E14 Position. Die Stelle ist zunächst auf 3 Jahre befristet. [\[Kontakt\]](#)