

## Editorial

Liebe Leserinnen und Leser, aufgrund der Corona-Krise befinden wir uns alle in einer seltsamen Zeit des Lockdown, oder auch der langsamen Rückkehr in „normale“ Verhältnisse. Home Office, digitale Lehre – das waren alles mal progressive Schlagworte, viele von uns wünschen sich vielleicht einfach nur mal wieder direkte analoge Begegnung – privat oder im Büro – und vielleicht auch mal wieder eine Vorlesung mit Tafel und Kreide, zumindest aber mit einem Blick in aufmerksame (oder ratlose) Gesichter. Auch in diesen Zeiten gibt es erfreuliche News: Wir haben drei neue IEEE Fellows aus Deutschland im Bereich Sprache/Audio, die ITG-Fachgruppe Signalverarbeitung verabschiedet ihren langjährigen Leiter Prof. Dr.-Ing. Bin Yang und begrüßt einen neuen aus unseren Reihen. Schauen Sie selbst, um welche Personen es sich handelt. Bleiben Sie gesund!

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

- Die ITG-Vorstandswahlen für die Amtsperiode 2021 bis 2023 stehen in Kürze an. Aus unseren Reihen kandidieren in der Gruppe „Wissenschaft und Lehre“ Prof. Dr.-Ing. Dorothea Kolossa (Wiederwahl) und Prof. Dr.-Ing. Tim Fingscheidt. Schauen Sie sich in Ruhe die bald zur Verfügung gestellte komplette Kandidatenliste an. Das Wichtigste ist: Nehmen Sie an der Wahl teil.

## Persönliches

- Drei Kolleginnen und Kollegen aus dem Sprach- und Audio-Bereich aus Deutschland wurden zum 1.1.2020 zu IEEE Fellows ernannt: Reinhold Häb-Umbach, *for contributions to robustness of automatic speech recognition*; Meinard Mueller, *for contributions to music signal processing*; und Tanja Schultz, *for contributions to multilingual speech recognition and biosignal processing*. Herzlichen Glückwunsch!

## Projekte und Aktivitäten

- Die mittlerweile 69. Sitzung der ITG-Fachgruppe [Signalverarbeitung und maschinelles Lernen](#) fand am 06. März 2020 am [Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS](#) in Sankt



Augustin statt. Das Fraunhofer IAIS ist eines der führenden und anwendungsorientierten Forschungsinstitute für Machine Learning und Künstliche Intelligenz in Europa und somit ein geeigneter Veranstaltungsort für die ITG-Fachgruppe, die vermehrt Themen zum Machine Learning in Kombination mit Signalverarbeitung erforscht.

Das abwechslungsreiche Tagungsprogramm wurde von der Abteilung [NetMedia](#) des Fraunhofer IAIS gestaltet. Nach der Vorstellung des Institutes durch Dr. Joachim Köhler wurden zahlreiche Fachvorträge zu den Themen Spracherkennung (Dr. Christoph Schmidt), Sprachassistenzsysteme (David Laqua), KI-basierte Dokumentenanalyse (Marcin Namysl), Audio- und Sprachsignalverarbeitung für Verteidigung und Sicherheit (Prof. Dr. Frank Kurth, Fraunhofer-Institut für Kommunikation, Informationsverarbeitung und Ergonomie FKIE) gehalten. Diese von Fraunhofer entwickelten Technologien wurden im Anschluss anhand von praktischen Demonstratoren (u.a. Live-Subtitling, Audio-Mining und Sprachassistentenanwendung) den 35 Teilnehmerinnen und Teilnehmern vorgestellt. Abschließend präsentierten die Mitglieder der ITG-Fachgruppe neue Forschungsergebnisse aus ihren Bereichen.

Höhepunkt der Veranstaltung war jedoch die Verabschiedung von Prof.



Dr.-Ing. Bin Yang (Universität Stuttgart), der über 11 Jahre die Veranstaltungen der Fachgruppe eindrucksvoll organisiert hat. In einer sehr persönlichen Dankesrede durch seinen Nachfolger Prof. Dr. Gerhard Schmidt (Universität Kiel) wurden

die umfangreichen Verdienste von Prof. Dr.-Ing. Bin Yang gewürdigt und eine reibungslose Staffelübergabe zu Prof. Dr. Gerhard Schmidt vollzogen.

- Im Fachgebiet Quality & Usability der TU Berlin ist das DFG-Projekt [Simulation von Konversationsverhalten bei gestörter Telefonübertragung](#) erfolgreich beendet worden. In diesem Projekt wurden Konversationen und insbesondere der Sprecherwechsel bei durch Übertragungsverzögerungen und Paketverluste gestörten Telefonübertragungen simuliert. Die so simulierten Konversationen können dazu verwendet werden, die Konversationsqualität zu schätzen.

### Bücher, Dissertationen

- Andreas Schwarz: "[Dereverberation and Robust Speech Recognition Using Spatial Coherence Models](#)" (Betreuer: W. Kellermann, FAU Erlangen-Nürnberg) Spatial coherence models allow a simple and robust modeling of noisy and reverberant speech signals. In this thesis, the application of such models for dereverberation and noise suppression is investigated, and a real-time, low-complexity and direction-of-arrival-independent enhancement system is proposed. For the application to automatic speech recognition, spatial features which capture the short-time diffuseness of the sound field are proposed.

### Journalartikel

- A. Brendel and W. Kellermann [Distributed Source Localization in Acoustic Sensor Networks Using the Coherent-to-Diffuse Power Ratio](#) Acoustic source localization in an acoustic sensor network (ASN) based on estimates of the coherent-to-diffuse power ratio (CDR) is investigated. It is shown that the CDR reflects the distance between an acoustic source and an observing microphone pair in real-world acoustic enclosures. As this relation is dependent on various physical parameters of the given acoustic environment, which are unknown in practice, we estimate it by Gaussian process regression. Subsequently, the obtained distance estimates are used by an ASN with known relative node positions for estimating the position of the acoustic source. The algorithm is designed in a completely distributed fashion, i.e., no fusion center is needed.

- A. Schmidt, H. W. Löllmann, and W. Kellermann [Acoustic Self-Awareness of Autonomous Systems in a World of Sounds](#) Autonomous systems (ASs), such as self-driving cars, robots and drones, operating in real-world environments, are exposed to a plurality and diversity of airborne sounds that carry a wealth of information for perception in cognitive dynamic systems. It is investigated in this overview article to what extent

current technical ASs exploit the acoustic modality to support self-awareness. First, relevant generic techniques and methods for acoustic scene analysis are reviewed. This is followed by an overview of current techniques for ego-noise suppression, as a specific additional challenge for ASs. Finally, active sensing is considered with its unique potential for acoustic scene analysis and, thus, for supporting self-awareness of ASs. Underlining the multimodal nature of self-awareness, links to other modalities and non-acoustic reference information are pointed out, where appropriate.

- C. Evers, H. W. Löllmann, H. Mellmann, A. Schmidt, H. Barfuss, P. A. Naylor, and W. Kellermann [The LOCATA Challenge: Acoustic Source Localization and Tracking](#)

Algorithms for acoustic source localization and tracking are essential for a wide range of applications such as smart homes, tele-conferencing systems, hearing aids, or autonomous systems. Numerous algorithms have been proposed for this purpose which, however, are not evaluated and compared against each other by a common database so far. The aim of the IEEE-AASP challenge on acoustic source LOCALization and TrACKing (LOCATA) is to provide an open-access framework for the objective evaluation and benchmarking of broad classes of algorithms for sound source localization and tracking. This paper provides a review of relevant localization and tracking algorithms and, within the context of the existing literature, a detailed evaluation and dissemination of the LOCATA challenge submissions. The presented evaluation highlights achievements in the field, open challenges, and identifies potential future directions.

- A. Aroudi, S. Doclo: [Cognitive-Driven Binaural Beamforming Using EEG-Based Auditory Attention Decoding](#)

Although several speech enhancement algorithms are available to reduce background noise or to perform source separation in multi-talker scenarios, their performance strongly depends on correctly identifying the target speaker to be enhanced. In this paper we propose a cognitive-driven binaural beamforming system, consisting of an MVDR or an LCMV beamformer, which is steered based on EEG-based auditory attention decoding. The performance of the proposed system is evaluated both in anechoic as well as reverberant noisy conditions.

- H. Schepker, S. Nordholm, S. Doclo: [Acoustic Feedback Suppression for Multi-Microphone Hearing Devices Using a Soft-Constrained Null-Steering Beamformer](#)

For an earpiece with multiple microphones, in this paper we propose an algorithm for acoustic feedback suppression, where the usage of a soft constraint allows to trade off between feedback suppression

and incoming signal preservation. We formulate two versions of the algorithm, either minimizing the residual feedback power or directly maximizing the maximum stable gain. Experimental results using an earpiece with 3 microphones show that an added stable gain of more than 50 dB can be obtained, while limiting the distortions of the incoming signal.

#### Tagungen (nach Paper Deadline sortiert) (Fettdruck: Verschiebungen durch Corona-Pandemie)

[ETSI STQ Workshop](#), 27.-28.08.2020 in Bratislava, Slowakei, **Paper Deadline: 30.06.2020** (siehe unten!)

[SLT](#) **19.-22.01.2021** in Shenzhen, China

**Paper Deadline: 14.08.2020** [Cfp]

[ICASSP](#) 06.-11.06.2021 in Toronto, Kanada

**Paper Deadline: 19.10.2020** [Cfp]

[Interspeech](#), **25.-29.10.2020**, Shanghai, China,

[keine Einreichungen mehr]

[IWAENC](#), **17.-20.05.2021**, Bamberg,

**Paper Deadline: 08.01.2021** [Cfp]

Der Workshop, der ursprünglich vom 31.08. bis 03.09.2020 stattfinden sollte, wurde aufgrund der Corona-Krise verschoben.

[EUSIPCO](#), **18.-22.01.2021**, Amsterdam, Niederlande,

[keine Einreichungen mehr]

[ITG Conference on Speech Communication 2020](#)

Die Tagung, die ursprünglich vom 09.09. bis 11.09.2020 stattfinden sollte, wird verschoben. Neues Datum noch unbekannt.



ETSI [TC STQ](#) is pleased to invite you to a [Workshop on Emerging Services for Speech and Audio](#) taking place in Bratislava in Slovakia on 27-28 October 2020.

Currently, the Programme Committee is calling for presentations on Emerging Services for Speech and Audio. Should you want to submit a presentation proposal, please [fill in this form](#) by 30 June 2020.

Speech based services are gaining increasing attention. While in the past, voice calls were the primary way of communicating, today the number of devices, of communication scenarios, and acoustical environments have exploded. New service types

enabled by automated speech recognition have emerged. The variety of Use Cases to qualify and optimize the performance of all the components that have to interwork has grown exponentially as well, requiring an emphasis on efficient testing.

Within this context, this event intends to address implementation topics, efficient testing and optimization and will bring together researchers, equipment manufacturers, test houses, operators, service/platform providers and regulators to exchange the latest information, to discuss about new applications and solutions and to jointly work on how to provide the best possible services for speech and audio applications.

#### Stellenanzeigen

- Das Institut für Nachrichtentechnik der TU Braunschweig sucht eine/n [Mitarbeiter/in \(TV-L E13\) für Erforschung und Entwicklung einer audiovisuellen Spracherkennung](#) im Rahmen des Förderprojekts SPEAKER ([Homepage des Projekts SPEAKER auf den FhG-Seiten](#)) [Kontakt]

- Bei der Technischen Universität Berlin ist/sind folgende Stelle/n zu besetzen: Wiss. Mitarbeiter\*in (d/m/w) - Entgeltgruppe 13 TV-L.

Das Quality and Usability Lab ist als Fachgebiet der Fakultät IV mit der Evaluierung und Gestaltung von Mensch-Maschine-Interaktionen befasst, wobei sowohl Aspekte der menschlichen Wahrnehmung, der technischen Systeme, sowie der Gestaltung der Interaktion Gegenstand unserer Forschung sind.

Fakultät IV - Institut für Softwaretechnik und Theoretische Informatik / Quality and Usability Lab  
Kennziffer: IV-265/20 (besetzbar ab 01.07.2020 / befristet für 34 Monate / Bewerbungsfristende 05.06.2020)

Aufgabenbeschreibung: Die Stelle soll Forschungsaufgaben im Bereich der Bewertung der Qualität von Sprachdiensten unter Verwendung eines Crowdsourcing-Ansatzes übernehmen. Diese Forschungsaufgaben sind von der Deutschen Forschungsgemeinschaft gefördert. [Kontakt]