

Editorial

Liebe Leserinnen und Leser, die TU Berlin freut sich, eine Professorin aus unseren Reihen gewonnen zu haben, Deutschlands Grundschulkindern freuen sich auf eine neue App zum Lesenlernen aus Braunschweig und München, und ansonsten geht ein Medientechnologiepreis nach Oldenburg. Und Dissertationen, die teils zu neuen ITU-T-Empfehlungen geführt haben! Lesen Sie selbst.

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

• Das BMBF fördert mit über 1,3 Mio. Euro das Projekt [BesserLesen](#) an der TU Braunschweig, TU München und LMU München. Ziel des Verbundprojekts ist die Bereitstellung einer demonstrativen App, die in ein didaktisches Gesamtkonzept zur kooperativen Leseförderung für Kinder im Grundschulalter eingebettet ist. Dabei kommt eine KI-basierte automatische Aussprachebewertung zum Einsatz, die Lernen im individuellen Lerntempo ermöglicht und durch gezieltes Feedback zur Verbesserung der Lesekompetenz von Kindern beitragen soll. Verbund-Projektleiter ist Prof. Tim Fingscheidt vom Institut für Nachrichtentechnik der TU Braunschweig. Die an der TU Braunschweig bereits u.a. im Projekt SPEAKER entwickelte automatische Spracherkennung soll im Rahmen des Projekts [BesserLesen](#) mit Algorithmen zur Bewertung der Aussprachequalität beim Lesen kombiniert werden. Mehr Details zum Projekt [hier](#).



Persönliches

• An der TU Berlin wurde zum 01.10.2022 das Fachgebiet für [Elektronische Systeme der Medizintechnik](#) gegründet. Es wird von Prof. Dr.-Ing. Dorothea Kolossa geleitet. Im Fachgebiet werden aktuelle Ansätze des maschinellen Lernens mit den klassischen Verfahren der Systemtheorie verbunden, im Sinne des interpretierbaren und zuverlässigen Lernens aus Daten und (bspw. kausalen und physikalischen) Strukturen.



• Wiebke Middelberg, Doktorandin am Dept. für medizinische Physik und Akustik der Universität Oldenburg, ist mit dem 3. Platz beim [ARD/ZDF-Förderpreis „Frauen + Medientechnologie“](#) 2022 für ihre Masterarbeit ausgezeichnet worden. In ihrer Masterarbeit erforschte Wiebke Middelberg die Möglichkeiten zur Verbesserung der Sprachverständlichkeit in Hörgeräten, indem externe Mikrophone zur Reduzierung von Störgeräuschen und konkurrierenden Sprechern mitgenutzt werden. Die Arbeit entstand im von der DFG geförderten Sonderforschungsbereich Hörakustik. Aktuell promoviert Wiebke Middelberg in der Abteilung Signalverarbeitung von Prof. Simon Doclo.

Bücher, Dissertationen

• Alicia Flores Requardt: [Driver Affect Recognition from Real-World Speech Data in In-Vehicle Driving Environments](#), Otto-von-Guericke-Universität Magdeburg (Betreuer: Prof. Dr. Andreas Wendemuth) Die Arbeit untersucht, wie tatsächliche intrinsische Zustände des Fahrers aus indirekten Beobachtungen der Emotionalität des Fahrers anhand seiner Sprache zuverlässig erhalten werden können. Verfahren des maschinellen Lernens zur Bewertung von sprachlichen Emotionen werden auf reale Szenen des Fahrens im Fahrzeug angewendet. Die Wirksamkeit solcher Verfahren in der gesamten Prozesskette von Datenerhebung, Datenvoranalyse, Klassifizierung und Bewertung wird im Umfeld von Horizon-2020 [ADAS&Me](#) nachgewiesen.

• Thilo Michael: [Simulating Conversations for the Prediction of Speech Quality](#), Quality and Usability Lab der Technischen Universität Berlin (Betreuer: Prof. Dr.-Ing. Sebastian Möller)

In der Arbeit von Thilo Michael geht es um die Vorhersage von Konversationsqualität einer VoIP-Verbindung durch die Simulation von Konversationen. In der Dissertation wird ein Simulations-Framework beschrieben, welches standardisierte Konversationszenarien auf der Sprachsignalebene simulieren kann. Die Simulation kann durch die Modellierung von Sprecherwechseln (Turn-Taking) und von Missverständnissen das

Verhalten von Nutzern bei Konversationen, die durch Übertragungsverzögerung und Paketverluste gestört sind, nachbilden. Mit Hilfe von aus den simulierten Konversationen extrahierten Parametern wird dann die Konversationsqualität vorhergesagt. Der in der Dissertation beschriebene Simulationsansatz wurde in der International Telecommunication Union ITU-T Recommendation P.836 standardisiert.

- Naveen Kumar Desiraju: [Low-Complexity Acoustic Echo Cancellation and Model-Based Residual Echo Suppression](#), Universität Oldenburg (Betreuer: Prof. Dr. Simon Doclo, Dr.-Ing. Tobias Wolff)
The main objective of this PhD thesis is to achieve low-complexity acoustic echo cancellation (AEC) for multichannel systems. First, novel tap selection schemes are proposed which exploit input signal sparsity across the dimensions of frequency, channels and time, leading to efficient partial updates of multichannel AEC filters in the subband domain. Second, novel signal-based methods are proposed to estimate the early and late residual echo PSD in online mode. Simulation results using both artificially generated as well as measured impulse responses show that the proposed output error method with the recursive prediction error algorithm outperforms state-of-the-art offline and online methods in terms of residual echo PSD estimation accuracy.

Journalartikel

- P. Didier, T. van Waterschoot, S. Doclo and M. Moonen
[Sampling Rate Offset Estimation and Compensation for Distributed Adaptive Node-Specific Signal Estimation in Wireless Acoustic Sensor Networks](#)
In this paper, we present a sampling rate offset (SRO) estimation and compensation method to allow the deployment of the distributed adaptive node-specific signal estimation (DANSE) algorithm in a wireless acoustic sensor network. The signals available at each node are first utilized in a coherence-drift-based method to blindly estimate SROs, which are then compensated for via phase shifts in the frequency domain. A modification of the weighted overlap-add (WOLA) implementation of DANSE is introduced to account for full-sample drifts, permitting per-sample signal transmission via an approximation of the WOLA process as a time-domain convolution. The performance of the proposed algorithm is evaluated in the context of distributed noise reduction.
- I. Siegert, N. Weißkirche, A. Wendemuth
[Acoustic-Based Automatic Addressee Detection for Technical Systems – A Review](#)
Acoustic addressee detection is a challenge that arises in human group interactions, as well as in interactions with technical systems like voice

assistants. The research domain is relatively new, and no structured review is available. To allow a natural interaction on the same level as human interactions, many studies focused on the acoustic analyses of speech. This survey gives an overview on the different studies and compares them in terms of utilized features, datasets, as well as classification architectures.

Tagungen (nach Paper Deadline sortiert)

- [Interspeech](#), 20.-24.08.2023, Dublin, Irland
Paper Deadline: 01.03.2023 [CfP]
- [ESSV 2023](#), 01.03.-03.03.2023, München,
[keine Einreichungen mehr]
- [DAGA 2023](#), 06.-09.03.2023, Hamburg,
[keine Einreichungen mehr]
- [WASPAA 2023](#), 22.-25.10.2023, New Paltz, NY, USA,
Paper Deadline: 28.04.2023 [CfP]
- [ITG Conference on Speech Communication 2023](#)
20.-22.09.2023, Aachen
Paper Deadline: 26.05.2023 [CfP]
- [ICASSP 2023](#), 04.-09.06.2023, Rhodos, Griechenland
[keine Einreichungen mehr]
- [EUSIPCO](#), 04.09.-08.09.2023, Helsinki, Finnland
[keine Einreichungen mehr]
- [ASRU](#), 16.-20.12.2023, Taipei, Taiwan,
Paper Deadline: 03.07.2023 [CfP tbd]

Stellenanzeigen

- Das neu gegründete Fachgebiet für Elektronische Systeme der Medizintechnik der TU Berlin sucht eine/n wissenschaftliche Mitarbeiter/in (TV-L E13, befristet auf 5 Jahre) für die Entwicklung neuer Methoden der multimodalen Sprachverarbeitung. [Weitere Infos via Mail \(Prof. Dorothea Kolossa\)](#)
- Das Quality and Usability Lab der TU Berlin hat mehrere offene Stellen im Bereich der Sprachqualitätsanalyse, insbesondere durch Crowdsourcing-Methoden und Konversationsimulation, sowie zur Anonymisierung von Sprechereigenschaften. [Kontakt: [Sebastian Möller](#)]
- Das Speech and Language Technology Lab des Deutschen Forschungszentrums für Künstliche Intelligenz (DFKI) in Berlin hat offene Stellen im Bereich der Sprecheranonymisierung sowie der multimodalen Falschnachrichtenerkennung. [Kontakt: [Sebastian Möller](#)]
- Das Institut für Nachrichtentechnik der TU Braunschweig sucht wiss. Mitarbeiter*innen im Bereich der automatischen Spracherkennung und des akustischen Few Shot Learning. Bewerbungen bitte an [Tim Fingscheidt](#).