

### Editorial

Liebe Leserinnen und Leser, gute Nachrichten! Wir freuen uns mit zwei unserer ITG-Mitglieder über den ITG-Preis 2019, mit einigen von uns über eine weitere Förderung der DFG-Forschungsgruppe „Akustische Sensornetze“ und nehmen darüber hinaus eine sehr spannende Neuerung zur Kenntnis: Subjektive Tests zur Bewertung der Sprachqualität müssen in Zukunft nicht mehr im stillen Stübchen durchgeführt werden. Lesen Sie von neuen ITU-T-Empfehlungen, die dies nun mittels *Crowdsourcing* ermöglichen, und nehmen Sie ggfs. an der [Interspeech 2020 Deep Noise Suppression Challenge](#) von Microsoft teil, die diesen Standard gleich auch mal zur Evaluierung anwendet. So kommt man hinein in die Thematik. Ein Meilenstein.

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 18.11.2019 wurde das neue Akustiklabor des Arbeitsbereichs [Signal Processing \(SP\) von Prof. Timo Gerkmann an der Universität Hamburg](#) mit einem Gastvortrag von Prof. Simon Doclo feierlich eröffnet. Der neue Laborraum ermöglicht es flexibel und



reproduzierbar, den akustischen Nachhall im Bereich zwischen 0,2 und 1,6 Sekunden einzustellen. Durch klappbare Paneele an den Wänden und Decken des Raums ist dieses Umstellen in wenigen Minuten erledigt. Im Anschluss wurden Echtzeitdemonstrationen vorgeführt. Einen Eindruck von Raum und Demos bekommen Sie auf dem [SP-Youtube-Kanal](#).

### Persönliches

● Dr.-Ing. Ina Kodrasi (IDIAP Research Institute) und Prof. Dr. Simon Doclo (Universität Oldenburg) wurden für ihre Veröffentlichung [„Analysis of Eigenvalue Decomposition-Based Late Reverberation Power Spectral Density Estimation“](#) mit dem ITG Preis 2019 ausgezeichnet. Die Preisverleihung fand im Rahmen des VDE Hauptstadtforums in Berlin statt. In der ausgezeichneten Veröffentlichung wird ein neuartiges mehrkanaliges Verfahren zur Schätzung der spektralen Leistungsdichte der Nachhallkomponente vorgeschlagen, wobei keine Schätzung der relativen Übertragungsfunktion des Zielsprechers erforderlich ist.



### Projekte und Aktivitäten

● In der Studiengruppe 12 der International Telecommunication Union (ITU-T) sind wieder einige neue Empfehlungen entstanden, die sich mit der Messung und Vorhersage der Sprachqualität beschäftigen. So gibt es die neue [ITU-T Rec. G.107.2](#) mit einem vollbandigen E-Modell zur Netzwerkplanung, und [ITU-T Rec. P.808](#) mit *Crowdsourcing*-basierten Methoden zur subjektiven Messung der Sprachqualität. Nähere Informationen bei [Sebastian Möller](#) von der TU Berlin.

● Die Deutsche Forschungsgemeinschaft setzt die Förderung der Forschungsgruppe „Akustische Sensornetze“ um weitere drei Jahre fort. In dem Projekt arbeiten Kollegen der Universitäten Paderborn (Häb-Umbach, Schmalenströer, Karl), Bochum (Martin, Enzner) und Erlangen-Nürnberg (Kellermann) an neuartigen Signalverarbeitungsverfahren, die auf räumlich verteilten Mikrofonen beruhen. Dies umfasst Fragestellungen, wie etwa die Verteilung von Teilaufgaben auf die Verarbeitungsknoten unter Berücksichtigung der Randbedingungen der akustischen Signalausbreitung und des Funknetzes, die Lokalisierung von Sensorknoten, Synchronisation verteilter Mikrofone und maschinelle Lernverfahren zur Signalverbesserung, Erkennung und zum Schutz der Privatsphäre.

- Auf der Interspeech 2019 hielten T. Nakatani (NTT, Japan) und R. Häb-Umbach (Univ. Paderborn) ein Tutorial mit dem Titel „[Microphone Array Signal Processing and Deep Learning for Speech Enhancement - Strong Together](#)“.

- Auf dem 47. Erlanger Kolloquium for Audiological Research and Development hielt T. Fingscheidt (TU Braunschweig) am 21.02.2020 eine Keynote zum Thema: „[Data Driven Speech Enhancement: History, Networks, and Loss Functions](#)“.

## Bücher, Dissertationen

- Benjamin Weiss (2020): "[Talker Quality in Human and Machine Interaction – Modeling the Listener’s Perspective in Passive and Interactive Scenarios](#)". In this book, published in the T-Labs Series in Telecommunication Services with Springer Nature, Cham, the background, state of research, and own contributions to the assessment and prediction of talker quality that is constituted in voice perception and in dialog are presented. Starting from theories and empirical findings from human interaction, major results and approaches are transferred to the domain of human-computer interaction. The main subject of this book is to contribute to the evaluation of spoken interaction in both humans and between human and computer, and in particular to the quality subsequently attributed to the speaking system or person, based on the listening and interactive experience.

- T. Schmitz: [Continuous-Amplitude Error Protection of Discrete-Time Signals](#), RWTH Aachen University (Betreuer: P. Vary, zweite Gutachterin: A. Schmeink). A concept of continuous-amplitude, discrete-time



signal transmission is investigated, which uses Analog Modulo Block Codes (AMB codes) for error protection. A major advantage of this concept is that there is no quality limitation with increasing quality of the transmission channel. The main contributions are the thorough design, analysis, and evaluation of AMB codes, as well as a novel method to semi-analytically estimate the transmission quality. For same channel bandwidth and transmission power, discrete-time transmission with AMB codes may outperform traditional digital transmission in terms of quality, latency and complexity. AMB codes are implemented by digital signal processing and are especially well suited for dedicated applications, such as hearing aids, wireless microphones, and wireless sensor networks.

- Christopher Schymura wurde am Institut für Kommunikationsakustik (Ruhr-Uni Bochum) am 12. November 2019 zum Dr.-Ing. promoviert. Seine Arbeit "[Cognitive models for acoustic and audiovisual source localization](#)" zeigt, wie auf Basis von multimodalen Informationen, von optimalen Bewegungsstrategien und kausalen Modellen Quellen in schwierigen akustischen und in dynamischen Szenarien besser lokalisiert werden können. Die Arbeit entstand im Team von Dorothea Kolossa, in Kooperation mit Boaz Rafaely und den NTT Communication Science Labs (Kyoto).

## Journalartikel

- M. Zohourian and R. Martin: [Binaural Direct-to-Reverberant Energy Ratio and Speaker Distance Estimation](#)

This article addresses the problem of distance estimation using binaural hearing aid microphones in reverberant rooms. The proposed distance estimators are based on the direct-to-reverberant energy ratio (DRR) and a calibration procedure. We present two novel approaches to estimate the DRR of binaural signals. Furthermore, we investigate informed calibration using the critical distance of the reverberant room and semi-blind calibration using the listener's own voice.

- C. Sohn, H. Krüger, and P. Vary: [Logarithmic pyramid vector quantization-design and theoretical analysis](#)

Vector quantization is an integral part of modern speech and audio codecs. This paper proposes the logarithmic pyramid gain-shape vector quantizer (LPVQ) and its optimization for memoryless Laplace distributed sources. We formulate analytically the distortion of the shape quantizer with respect to rate and vector dimension and compare with rate distortion theory. Furthermore, we proposed a strategy for the gain-shape bit allocation and verify the results by simulations.

- W. Mack, U. Bharadwaj, S. Chakrabarty, E.A.P. Habets

- [Signal-Aware Broadband DOA Estimation Using Attention Mechanisms](#)

Broadband multi-source direction-of-arrival (DOA) estimation is often based on the multi-microphone phase information observed in different frequency bands. To estimate only the DOAs of specific sound sources, such as human speakers, we propose to use attention mechanisms in combination with a deep neural network (DNN). The attention mechanisms allow the DOA estimator to focus on frequency bands dominated by the sound sources of interest. The attention can be obtained with a secondary DNN. Demo videos are available [here](#).

- N. K. Desiraju, S. Doclo, M. Buck and T. Wolff:  
[Online Estimation of Reverberation Parameters For Late Residual Echo Suppression](#)

This paper proposes two signal-based methods (i.e. output error and equation error) to jointly estimate frequency-dependent reverberation scaling and decay parameters in online mode. The estimated parameters are then used to generate an estimate for the late residual echo PSD. Extensive simulation results using both artificial as well as measured room impulse responses show that the proposed output error method with mean squared log error minimization outperforms state-of-the-art methods for late residual echo suppression.

- W. Mack and E.A.P. Habets  
[Deep Filtering: Signal Extraction and Reconstruction Using Complex Time-Frequency Filters](#)

Single-channel signal extraction of a desired source signal from a mixture with additional other signals and degradations is typically performed using time-frequency masks estimated by deep neural networks (DNNs). These masks are, by design, not capable of reconstructing severe destructive degradations, e.g., caused by packet-loss. Complex time-frequency filters are capable of reconstructing such degradations by incorporating neighboring time and frequency bins. In this paper, we elaborate on how to estimate such filters with DNNs and compare them to time-frequency masking techniques. A follow-up paper in which the proposed technique was used for de-clipping was presented at WASPAA 2020.

### Tagungen (nach Paper Deadline sortiert)

- [EUSIPCO](#), 24.-28.08.2020, Amsterdam, Niederlande,  
Paper Deadline (extended): 02.03.2020 [\[CfP\]](#)
- [ESSV](#), 04.-06.03.2020, Magdeburg,  
[keine Einreichungen mehr]
- [DAGA](#) 16.-19.03.2020 in Hannover,  
[keine Einreichungen mehr]
- [Interspeech](#), 02.-06.09.2020, Shanghai, China,  
Paper Deadline: 30.03.2020 [\[CfP\]](#)
- [IWAENC](#), 31.08.-03.09.2020, Bamberg,  
Paper Deadline: 24.04.2020 [\[CfP\]](#)
- [ICASSP](#) 04.-09.05.2020 in Barcelona, Spain  
[keine Einreichungen mehr]
- [ITG Conference on Speech Communication](#)  
9.-11.09.2020 in Aachen,  
Paper Deadline: 18.05.2020 [\[CfP\]](#)
- [SLT](#) 13.-16.12.2020 in Shenzhen, China  
Paper Deadline: 03.07.2020 [\[CfP\]](#)

### Stellenanzeigen

- Im Arbeitsbereich Signal Processing (SP) an der Universität Hamburg wird ein/e wissenschaftliche/r Mitarbeiter/in im Bereich „KI-basierte Sprachverbesserung für Hörhilfen“ gesucht. [\[Link\]](#)
- In der Abteilung Signalverarbeitung des Departments für medizinische Physik und Akustik an der Universität Oldenburg wird ein Postdoctoral Research Associate im Bereich „Individualized signal processing for hearing devices“ gesucht. [\[Link\]](#)
- Das Institut für Nachrichtentechnik der TU Braunschweig sucht einen Mitarbeiter/in (TV-L E13) für Erforschung und Entwicklung einer audiovisuellen Spracherkennung im Rahmen des Förderprojekts [SPEAKER](#). [\[Kontakt\]](#)