

Editorial

Liebe Leserinnen und Leser, der ehemalige Fachbereich 4 in der ITG heißt mittlerweile Fachbereich Audiotechnik (AT), und die Herausgeber dieses Newsletters sind entsprechend die Fachausschüsse FA 3 (Sprachakustik) und FA 4 (Sprachverarbeitung). Nach kleiner Erinnerung an potentielle Autoren kam plötzlich eine Flut von Beiträgen – schön, dass es in der *Speech Community* eigentlich immer eine Menge zu berichten gibt. Besonders erinnern möchten wir an die 13. ITG-Fachtagung Sprachkommunikation, die im Oktober in Oldenburg stattfinden wird. Die Paper Submission Deadline findet sich am Ende des Newsletters.

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!

Latest News

● Am 2. Februar 2018 hat das [Institut für Kommunikationssysteme](#) der RWTH Aachen University ein Festkolloquium zum 70. Geburtstag von Prof. Dr.-Ing. Peter Vary veranstaltet. Dieser runde Geburtstag war eine schöne Gelegenheit, auf die drei Jahrzehnte zurückzublicken, in denen Peter Vary engagiert, prägend und mit großem Ideenreichtum das Institut geleitet hat. Durch die Anwesenheit zahlreicher Kollegen und ehemaliger Doktoranden wurden seine wissenschaftlichen Beiträge und die Ausstrahlung seiner Forschungsarbeiten in den Bereichen der digitalen Sprachverarbeitung und der Mobilfunktechnik gewürdigt.



● Das NRW.BANK.Venture Center lädt Start-Ups, Forschungsgruppen und Unternehmen am 2. März 2018 zu einem „Innovationsforum Sprache & Audio“ nach Düsseldorf ein. Es werden Szenarien für Sprach- & Audiotechnologien und Chancen, die sich für Neugründungen ergeben, diskutiert. Für die Anmeldung (**bis zum 26. Februar 2018**) und für weitere Informationen wenden Sie sich bitte an E-Mail: venture-center@nrwbank.de oder Tel: 0211 / 91 741 1026.

Persönliches

● Simon Receveur (ehemals TU Braunschweig) wurde in der Berlin-Brandenburgischen Akademie der Wissenschaften vom Vorsitzenden der Informationstechnischen Gesellschaft (ITG) Prof. Rüdiger Kays im November einer der [ITG-Preise](#) 2017 überreicht.



Ausgezeichnet wurde die Journalpublikation in den IEEE/ACM Transactions on Audio, Speech, and Language Processing: [Turbo Automatic Speech Recognition](#). Die Arbeit beschreibt den Transfer des Turbo-Prinzips aus der Kommunikationstechnik in das Feld der automatischen Spracherkennung und zeigt geeignete Fusionsmethoden im Zuge des Erkennungsprozesses auf.

● Herr Dr.-Ing. Nilesh Madhu hat einen Ruf auf die Universitätsprofessur für [Audio- und Sprachtechnologie](#) an der Universität Gent in Belgien erhalten. Die Professur hat er zum 1. Dez. 2017 angetreten.



Projekte und Aktivitäten

● Vom 5. bis 7. Dezember 2017 wurde von der Professur Kognitive Systeme der TU Dresden (Jun.-Prof. Peter Birkholz) ein Workshop zum Thema „Modeling the Larynx and the Voice“ im Rahmen des von der DFG-geförderten SimPhon-Projekts durchgeführt.

● Im Rahmen von [ahoi.digital](#) fördert die Hansestadt Hamburg das Projekt "[Adaptive crossmodale Sensordatenerfassung](#)". In dem Projekt entwickeln Forscher der Universität Hamburg und der Technischen Universität Hamburg gemeinsam neuartige Algorithmen zur robusten, beschleunigten und crossmodalen Sensordaten-Erfassung und -Verarbeitung. Anwendungsgebiete sind Sprachkommunikationsgeräte (Hörhilfen, Smartphones), Robotik und Medizintechnik. Sprecher ist [Timo Gerkmann](#), Universität Hamburg.

● An der TU Braunschweig wird mit Unterstützung des Bundesministeriums für Bildung und Forschung (BMBF) ein zweijähriges Projekt zur Förderung von Qualifizierungsmaßnahmen und Forschungsvorhaben im Bereich Maschinelles Lernen im Rahmen des Förderprogramms "IKT 2020 – Forschung für Innovationen" gestartet. Es handelt sich um das sog. *Deep Learning Lab* der TU Braunschweig, kurz

TUBS.dll, das in Ergänzung zur Lehrveranstaltung „Mustererkennung“ von Tim Fingscheidt eine Praxisphase der Studierenden u.a. mit einem Wettbewerb der Teilnehmer/innen untereinander unter Nutzung realer Daten von Unternehmen vorsieht. Das Projekt wird durch sechs kleine und große Unternehmen der Region Braunschweig/Wolfsburg/Hannover unterstützt und nimmt zurzeit ein großes Rechencluster aus Grafikprozessoren (GPUs) am Rechenzentrum der TU Braunschweig in Betrieb.

- Das Fraunhofer IAIS erforscht und entwickelt in dem vom BMBF geförderten Projekt [KA3](#) (Kölner Zentrum für Analyse und Archivierung audiovisueller Daten) Methoden und Anwendungen zur automatischen Analyse von Interviews. In dem Anwendungsfall *Oral History* werden Zeitzeugeninterviews der Fernuniversität Hagen mittels des Fraunhofer IAIS Audio Mining Systems segmentiert, transkribiert und indiziert. In einem weiteren Anwendungsfall zur Erforschung von Interaktionsmustern von gesprochener Sprache werden neue Verfahren zur Sprachsegmentierung untersucht.

- Bei der [voice INTER connect GmbH](#) gibt es Neuigkeiten rund um Sprachsteuerung und Sprachdatenerfassung. Mit dem vicCONTROL Industrial SDK stellen wir pünktlich zur [Embedded World 2018](#) unsere verbesserte embedded Sprachsteuerung auch für ARMv7-Plattformen vor. Außerdem zeigen wir mit unserem DSP-Modul vicFVAS, wie Farfield Voice Acquisition mittels Mikrofonarray funktioniert.

- Im Institut für Nachrichtentechnik an der TU Braunschweig wurde im Rahmen einer BMWi-Förderung das Projekt “HiFi-AEC” in Kooperation mit der Firma Linguwerk GmbH in Dresden gestartet. Die Idee ist die Aufwertung der Sprachbedienung bei HiFi- und Infotainment-Systemen durch eine am Institut zu entwickelnde akustische Echokompensation für Stereo-Signale.

- Die Leiter des DEGA-Fachausschusses Sprachakustik laden Interessenten herzlich zur FA-Sitzung während der DAGA 2018 in München ein: Dienstag, der 20.03.2018, 12:40-13:20h, Raum MW 1050. Während der Sitzung sollen insbesondere Möglichkeiten einer engeren Zusammenarbeit mit den ITG-Fachausschüssen FA 4.3 und 4.4 (AT3 und AT4) und der ESSV (Elektronische Sprachsignalverarbeitung) erörtert werden. Anregungen und Diskussionsbeiträge in Vorbereitung der FA-Sitzung – auch von Interessenten, die nicht an der Sitzung teilnehmen können – sind gerne willkommen (bitte Email an [Alexander Raake](#) und [Janto Skowronek](#)).

Dissertationen

- Philipp Thüne: [Advances in Blind Multichannel Wiener Filtering of Noisy Speech](#), Ruhr-Universität Bochum (Berichter: G. Enzner, P. Jax). The main contribution is a novel two-stage postfilter estimation technique for multichannel Wiener filters. First, ML estimators for speech and noise PSDs are derived to form an ML-based postfilter. In a second step, this ML-based postfilter itself is considered as a statistical quantity, allowing for a Bayesian refinement based on an *a priori* model of an ideal Wiener postfilter. The Bayes refinement can be implemented as a lightweight extension to reduce artifacts in blind multichannel speech processing.

- Ante Jukić: [Sparse Multi-Channel Linear Prediction for Blind Speech Dereverberation](#), Universität Oldenburg (S. Doclo)

Key contributions: general framework for blind speech dereverberation based on multi-channel linear prediction and using a sparse prior that promotes sparsity of the speech signal in the time-frequency domain. Batch as well as adaptive algorithms are derived, which are suitable for online processing and dynamic scenarios. In addition, extensions to sparsity-based joint dereverberation and noise reduction are presented.

Journalartikel

- R. Rehr, T. Gerkmann: [On the Importance of Super-Gaussian Speech Priors for Machine-Learning Based Speech Enhancement](#)

Einige Ansätze zur Sprachverbesserung, die auf maschinellem Lernen basieren, nutzen lediglich grobe spektrale Einhüllende für die Schätzung der Sprachleistung. Wir zeigen und begründen, dass die Annahme super-Gauß'scher Sprachkoeffizienten dann zu sehr deutlichen Qualitätsgewinnen gegenüber der Gauß-Annahme führt.

- C. Huemmer, C. Hofmann, R. Maas und W. Kellermann:

- [Estimating Parameters of Nonlinear Systems Using the Elitist Particle Filter Based on Evolutionary Strategies](#)

An ‘elitist’ particle filter based on evolutionary strategies (EPFES) is presented as an efficient parameter estimation approach for nonlinear systems, generalizing the widely-used Gaussian particle filter. Its application to nonlinear acoustic echo cancellation confirms algorithmic efficacy and computational efficiency.

- J. Stahl, P. Mowlae: [A Pitch-Synchronous Simultaneous Detection-Estimation Framework for Speech Enhancement](#)

This paper presents a simultaneous detection-estimation framework for speech enhancement under the assumption of a harmonic plus noise model

for speech. In order to achieve better estimates of the harmonic components, the detection-estimation procedure is combined with a pitch-synchronous signal analysis stage that increases the achievable accuracy of the harmonic parameter estimates.

• M. Taseska and E.A.P. Habets:

[Blind source separation of moving sources using sparsity-based source detection and tracking](#)

A blind source separation algorithm is proposed for an unknown number of moving sources. In this paper, an approximate Bayesian tracker is derived and the association of each time-frequency bin to the dominant source is performed using the tracker's measurement-to-source association probabilities. These probabilities are used to estimate the statistics of the sources, which are required to estimate the source signals. Some audio examples can be found [here](#).

• C. Zheng, A. Deleforge, X. Li, W. Kellermann:
[Statistical Analysis of the Multichannel Wiener Filter Using a Bivariate Normal Distribution for Sample Covariance Matrices](#).

The statistical performance of the multichannel Wiener filter (MWF) is studied when the weights are computed using estimates of the sample covariance matrices of the noisy and the noise signals. As the key contribution, a bivariate normal distribution is introduced to approximate the joint p.d.f. of the noisy and the noise sample covariance matrices and allows to quantitatively predict the performance of the MWF and provide explanations for hitherto unexplained phenomena.

• M. Zohourian, G. Enzner, R. Martin:

[Binaural Speaker Localization Integrated Into an Adaptive Beamformer for Hearing Aids](#)

In diesem Beitrag stellen wir Verfahren für die kombinierte Sprecherlokalisierung und Sprechertrennung für Hörgeräte vor. Es werden mehrere Lösungen für die Quellenlokalisierung im Zusammenspiel mit einem adaptiven Beamformer verglichen und zudem eine verbesserte Methode für die Berechnung einer "Target Presence Probability (TPP)" vorgeschlagen.

• A. Kumar, C. Schmidt, J. Köhler:

[A Knowledge Graph Based Speech Interface for Question Answering Systems](#)

Das Journal-Paper stellt verschiedene Ansätze vor, um Wissensgraphen zur Verbesserung der Spracherkennung und des *Retrievals* zu verwenden. Speziell werden Ansätze zum Thema *Question/Answering* vorgestellt und bewertet.

• N. K. Desiraju, S. Doclo, and T. Wolff:

[Efficient multichannel acoustic echo cancellation using constrained tap selection schemes in the subband domain](#)

This paper proposes a novel tap selection scheme for partially updating the adaptive filters in multichannel acoustic echo cancellation, by exploiting sparsity across time, frequency, and channels. Simulation

results show that the proposed tap selection scheme outperforms state-of-the-art tap selection schemes in terms of echo cancellation performance and enables to significantly reduce computational cost.

Tagungen (nach Paper Deadline sortiert)

[EUSIPCO](#) 03.-07.09.2018 in Rom, Italien

Extended Paper Deadline: 28.02.2018 [\[CfP\]](#)

[ESSV](#) 07.-09.03.2018, Ulm, Deutschland

[keine Einreichungen mehr]

[DAGA](#) 19.-22.03.2018 in München, Deutschland

[keine Einreichungen mehr]

[Interspeech](#) 02.-06.09.2018 in Hyderabad, Indien

Paper Deadline: 23.03.2018 [\[CfP\]](#)

[ECDA](#) (European Conference on Data Analysis) 04.-

06.7.2018 in Paderborn: Die Tagung bietet eine sehr gute Gelegenheit, die europäische und insbesondere die deutsche *Machine-Learning*-Community kennenzulernen.

Deadline für Abstracts: 1.4.2018 [\[CfP\]](#)

[ICASSP](#) 15.-20.04.2018 in Calgary, Alberta, Kanada

[keine Einreichungen mehr]

[IWAENC](#) 17.-20.09.2018 in Tokyo, Japan

Paper Deadline: 20.04.2018 [\[CfP\]](#)

[13. ITG Fachtagung Sprachkommunikation](#)

10.-12.10.2018 in Oldenburg, Deutschland

Paper Deadline: 18.05.2018 [\[CfP\]](#)

[ICPR](#) 20.-24.08.2018 in Beijing, China

[keine Einreichungen mehr]

ITG Fachtagung Sprachkommunikation

Die Vorbereitungen für die 13. ITG Fachtagung Sprachkommunikation, die vom 10.-12.10.2018 an der Universität Oldenburg stattfindet, laufen auf Hochtouren.



Der Call for Papers ist jetzt [online](#) und wir möchten Sie alle herzlich einladen, einen Beitrag bis zum 18.05.2018 einzureichen. Auf Sie warten spannende Sitzungen aus den Bereichen Machine Learning for Speech Enhancement, Speech and Diagnostics, Automatic Speech Recognition und Multimodal Speech Processing, sowie Keynotes u.a. von Prof. Sharon Gannot. Wir freuen uns darauf, Sie in Oldenburg willkommen zu heißen und hoffen auf interessante Beiträge und Diskussionen.