

Editorial

Was für eine lebendige Speech Community: Schauen Sie sich allein mal die Projekte in diesem zweiten Newsletter der ITG-Fachausschüsse 4.3 und 4.4 an! Die planmäßige Fertigstellung und das gute Interesse geben uns die Hoffnung, die dritte Ausgabe bereits im Juni 2017 verteilen zu können.

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 auch möglicher Autor sein möchten!

Latest News

● Im letzten Newsletter berichteten wir von der Gründung der japanischen Far Field Sound Corporation (Herr Ishikawa) – mittlerweile ist auch die Gründung der deutschen Schwesterfirma [Far Field Sound GmbH](#) in Braunschweig geschehen. „Unser Produktportfolio bietet u.a. automotive Freisprech-Technologie, sowie In-Car-Kommunikationssysteme“, sagt Tim Fingscheidt, einer der drei Mitbegründer.

**Persönliches**

● Gerald Enzner (Ruhr-Universität Bochum) wurde von den Mitgliedern des IEEE Germany Chapters zum Vice-Chair des IEEE Signal Processing Society Germany Chapters gewählt. Ein Ziel des Engagements ist der Ausbau der Zusammenarbeit von IEEE und VDE/ITG auf Fachgruppenebene. Initiativen und Beiträge sind sehr willkommen!

Projekte und Aktivitäten

● **12. ITG-Fachtagung Sprachkommunikation.** Zur 12. Fachtagung in dieser Reihe trafen sich vom 05.-07.10.2016 rund 150 Wissenschaftler und Anwender in Paderborn, um neueste Forschungsergebnisse zu präsentieren und zu diskutieren. Neben den rund 70 regulären Präsentationen fanden vor allem die drei eingeladenen Übersichtsvorträge von Bhiksha Raj, Hynek Hermansky und Richard Hendriks großen Anklang. [Die Vortragsfolien sind online verfügbar.](#)

● Die Deutsche Forschungsgemeinschaft richtet an den Universitäten Paderborn, Bochum und Erlangen-Nürnberg eine neue [Forschergruppe "Akustische Sensornetze"](#) ein, deren Sprecher Reinhold Häb-Umbach ist. Weitere beteiligte PI's sind H. Karl, J. Schmalenströr, R. Martin, G. Enzner und W. Kellermann. Verteilte Mikrofone bieten u.a. den Vorteil, dass stets ein Sensor der relevanten Schallquelle nahe ist, wodurch die Signalqualität der Aufnahme erhöht wird. Die Herausforderungen liegen neben der Kommunikation, Synchronisation, und Sicherung der Netze, in der Klassifikation und Filterung der relevanten Signale. Eine Stelle im

Projekt ist noch mit einem wiss. Mitarbeiter zu besetzen [[Ansprechpartner G. Enzner](#)].

● Das EU-Projekt [DREAMS \(Dereverberation and Reverberation of Audio, Music, and Speech\)](#), mit 12 Partnern aus der akademischen Welt und der Industrie (u.a. Universität Oldenburg, Fraunhofer IDMT, Nuance Communications), ist Ende 2016 erfolgreich abgeschlossen worden. In den 16 Teilprojekten sind neuartige Verfahren zur Nachhallunterdrückung, raumakustische Modelle und objektive Bewertungsmaße entwickelt worden. Anfang 2017 erscheint dazu eine Sonderausgabe im [Journal of the Audio Engineering Society](#).

● Das interdisziplinäre NRW-Fortschrittskolleg [SecHuman - Sicherheit für Menschen im Cyberspace](#) beginnt im Oktober seine Arbeit, unter anderem mit dem von Dorothea Kolossa und Karin Pittner geleiteten Projekt "Sprachliche Imitations- und Verschleierungsstrategien," in dem Erfahrungen aus der Linguistik mit maschinellem Lernen kombiniert werden, um Vortäuschungen von Identitäten im Internet zu erkennen.

● In Kooperation zwischen der CAU Kiel und der TU Berlin werden im DFG-Erkenntnis-Transferprojekt [Qualitätsattribute und Gesamtqualität übertragener Sprache](#), geleitet von Gerhard Schmidt und Sebastian Möller, Erkenntnisse zur Identifikation perceptiver Dimensionen weiterentwickelt, aus denen sich der Qualitätseindruck eines Hörers zusammensetzt. Die Schätzung der Gesamtqualität ergibt sich dabei durch Überlagerung der Dimensionsschätzungen. Die Projektergebnisse werden den beteiligten Industriepartnern, Deutsche Telekom und SwissQual, zur Entwicklung von neuartigen Qualitätsschätzern zugänglich gemacht.

● Das Institut für Nachrichtentechnik der Technischen Universität Braunschweig kooperiert im Rahmen einer zweijährigen Förderung des BMWi mit der Firma viasto GmbH in Berlin. Seit Januar 2017 wird an einem selbstlernenden Expertensystem zur objektiven Bewertung von zeitversetzten Jobinterviews geforscht und entwickelt. Der Braunschweiger Beitrag an der „Voice Analytics Suite“ wird ein robuster akustischer Keywordspotter sein.

● Das erste Meeting der neuen ITU-T-Studiengruppe 12 fand vom 10.-18.01.2017 in Genf statt. In der Study Group 12 (SG12), Question 4 arbeitet man am Standard P.carSFS „Super-WideBand (SWB) and FullBand (FB) stereo hands-free communication in motor vehicles“ sowie Updates bestehender Standards. Darüber hinaus erfolgen Revisionen der Standards [P.1100](#), [P.1110](#) sowie [P.1140](#).

● In zwei BMBF-Forschungsvorhaben entwickelt die [voice INTER connect GmbH](#) Assistenzsysteme der Zukunft mit. Akustische Quellenlokalisierung und -trennung mittels Mikrofonarray sind der Beitrag für MOVA3D (Multimodaler omnidirektionaler 3D-Sensor für die Verhaltensanalyse von Personen); die

Gestaltung von Sprachdialogen und Nutzermodellierung stehen im Fokus bei RadAR+ (Reiseassistenzsystem für dynamische Umgebungen auf Basis von Augmented Reality).

Produkte & Sprachtechnologie ready for transfer

- VoCAS (Voice Control Analysis System) ist eine neue leistungsfähige Software von HEAD acoustics zur Bewertung von Spracherkennungssystemen in den unterschiedlichsten Anwendungsfeldern. Mit VoCAS lässt sich die Sprachqualität der Systeme objektiv und schnell bewerten – unter realistischen und reproduzierbaren Testbedingungen.

- In Kooperation mit der R&D Gruppe der Firma NXP Software B.V. in Leuven, Belgien, hat die Technische Universität Braunschweig ein neuartiges Verfahren zur Störgeräuschreduktion entwickelt, das in der aktuellsten Produktsuite von NXP Software namens [LifeVibes VoiceExperience 10](#) für Smartphones angeboten wird.

Publikationen und Datenbanken

Buch: P. Mowlae, J. Kulmer, J. Stahl, F. Mayer: [Single Channel Phase-Aware Signal Processing in Speech Communication: Theory and Practice](#), John Wiley and Sons, December 2016.

The book provides a comprehensive guide to phase-aware speech signal processing and reviews the history, fundamentals of phase processing, phase estimation in noise. It exemplifies several applications to demonstrate the usefulness of phase-aware signal processing in speech technologies.

Dissertation:

- M. Krawczyk-Becker: [Phase-Aware Single-Channel Speech Enhancement](#), Uni Oldenburg (T. Gerkmann)

Key contributions: novel methods for phase estimation; derivation of novel MMSE-based clean speech estimators given a phase estimate and its uncertainty; the proposed phase-aware speech enhancement algorithms may significantly improve the perceived signal quality also in a blind setup.

- S. Han: [Contributions to Improved Hard- and Soft-Decision Decoding in Speech and Audio Codecs](#),

Technische Universität Braunschweig (T. Fingscheidt) Key contributions: AMR narrowband and AMR wideband soft decision decoding, HE AAC soft decision decoding (variable-length coding with soft-decision), improved hard-decision decoding using prediction-based adaptive receiver codebooks.

Journalartikel:

- P. Mowlae, J. Stahl, J. Kulmer: [Iterative Joint MAP Single-Channel Speech Enhancement Given Non-Uniform Phase Prior](#)

The article proposes an iterative joint maximum a posteriori (MAP) amplitude and phase estimator assuming a non-uniform phase distribution. Experimental results demonstrate the effectiveness of the proposed method in recovering both amplitude and phase in noise, justified by perceived quality and speech intelligibility.

- M. Krawczyk-Becker, T. Gerkmann:

[An evaluation of the perceptual quality of phase-aware single-channel speech enhancement](#)

The key contribution of this work are listening experiments showing that phase-aware speech enhancement algorithms may significantly improve signal quality as compared to traditional phase-blind approaches.

- J. Abel, M. Kaniewska, C. Guillaume, W. Tirry, T.

Fingscheidt: [An Instrumental Quality Measure for Artificially Bandwidth-Extended Speech Signals](#)

Wideband PESQ and POLQA are not reliably predicting speech quality for artificial speech bandwidth extension test conditions. We propose an SVM-based high-performance MOS predictor which is superior to both WB-PESQ and POLQA.

Tagungen (nach Paper Deadline sortiert)

[EUSIPCO](#) 28.08.-02.09.2017 auf Kos, Griechenland,

Paper Deadline 17.02.2017 [CfP]

[ETSI Workshop on Multimedia Quality in Virtual, Augmented or Other Realities](#) 10.-11.05.2017 in

Sophia Antipolis, Frankreich

Presentation Deadline 20.02.2017 [CfP]

[HSCMA](#) 01.-03.03.2017 in San Francisco, CA, USA

[keine Einreichungen mehr]

[ICASSP](#) 05.-09.03.2017 in New Orleans, LA, USA,

[keine Einreichungen mehr]

[DAGA](#) 06.-09.03.2017 in Kiel,

[keine Einreichungen mehr]

[Interspeech](#) 20.-24.08.2017 in Stockholm, Schweden,

Paper Deadline 14.03.2017 [CfP]

[ESSV](#) 15.-17.03.2017 in Saarbrücken

[keine Einreichungen mehr]

[WASPAA](#) 15.-18.10.2017 in New Paltz, NY, USA,

Paper Deadline 20.04.2017 [CfP]

[ASRU](#) 16.-20.12.2017 in Okinawa, Japan,

Paper Deadline 28.06.2017 [CfP]

SAVE THE DATE: Die 13. ITG-Fachtagung Sprachkommunikation wird vom 10.-12. Oktober 2018 an der Universität Oldenburg stattfinden.

Stellenangebote

- Die [voice INTER connect GmbH](#) sucht Verstärkung in den Bereichen Blind Source Separation, Beamforming, Speech Dialogue Development, Natural Language Understanding und Deep Neural Networks. [Link]

- Die [HEAD acoustics GmbH](#) bietet Stellen für Ingenieure in den Bereichen Forschung, Entwicklung, Applikation, Messtechnik und Support sowie für Praktikanten und studentische Hilfskräfte an. [Link]

- Die Sivantos Gruppe ist einer der weltweit führenden Anbieter von Hörgeräten und sucht für ihren Entwicklungsstandort in Erlangen motivierte Mitarbeiter in den Bereichen „Entwicklungsingenieur Audiosignalverarbeitung“ und „Softwareentwickler Signalverarbeitung“. [Link]