

Editorial

Vor nunmehr einem Jahr haben wir seitens der ITG-Fachausschüsse 4.3 und 4.4 unseren ersten Newsletter namens *Voice Message* herausgegeben. Laufend kommen neue Abonnenten hinzu. Damit kommt auch zunehmend die Frage nach älteren Ausgaben des Newsletters auf. Diese finden sich im [Archiv auf den Webseiten der ITG](#).

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 Department für medizinische Physik und Akustik der Universität Oldenburg wurde ein neues Labor mit variabler Akustik eingerichtet. Mit Hilfe von ausklappbaren akustisch absorbierenden und reflektierenden Paneelen, die an den Wänden und an der Decke angebracht sind, kann die Nachhallzeit innerhalb von Minuten beliebig zwischen 0,2 und 1,2 Sekunden variiert werden. Das Labor wird für Messzwecke, für Forschung an Echtzeitalgorithmen zur Verbesserung der Sprachqualität und für Hörversuche eingesetzt.



Persönliches

• Janek Ebberts, Lukas Drude, Thomas Glarner (auf dem Bild von rechts nach links) und Jahn Heymann wurden auf der Interspeech 2017 mit einem Best Student Paper Award ausgezeichnet. Ihre Veröffentlichung mit dem Titel "[Hidden Markov Model Variational Autoencoder for Acoustic Unit Discovery](#)" und den weiteren Koautoren Bhiksha Raj (Carnegie Mellon University, USA) und Reinhold Häb-Umbach beschreibt einen auf Zeitsignale angepassten *variational autoencoder*, um die akustischen Grundbausteine einer Sprache auf unüberwachte Weise zu lernen. Das Verfahren verknüpft probabilistische grafische Modelle (hier: HMMs) mit neuronalen Netzen zur Inferenz.



Projekte und Aktivitäten

- Spielerisches Lernen korrekter Artikulation steht im Mittelpunkt des neuen EFRE-Projekts [AVATAR](#). Koordiniert von der FH Dortmund, gemeinsam mit dem Institut für Kommunikationsakustik (Prof. Kolossa), der Hochschule für Gesundheit und Partnern aus der Wirtschaft werden hier seit August 2017 die kommenden drei Jahre logopädisch fundierte Lerntools für Kinder mit phonetisch-phonologischen Artikulationsstörungen auf Basis audiovisueller Spracherkennung entwickelt.
- Der Sonderforschungsbereich SBF-TRR 62 (Ulm / Magdeburg) beendet im Dezember 2017 seine neunjährige Forschung über [Companion-Technologie für Kognitive Technische Systeme](#). Ergebnisse sind insbesondere die Erkennung und dialogische Nutzung von Emotionen, Dispositionen und Intentionen aus realen Sprachdaten und weiteren bis zu zehn Modalitäten. [Kontakt: [Andreas Wendemuth, Universität Magdeburg](#)]
- Gemeinsam mit der Firma [viasto GmbH](#) aus Berlin arbeitet das Institut für Nachrichtentechnik der Technischen Universität Braunschweig an der Entwicklung eines selbstlernenden Expertensystems (Voice Analytics Suite) zur objektiven Bewertung von zeitversetzten Jobinterviews basierend auf akustischer Keyworderkennung und psychologischen Modellen. Die TU Braunschweig entwickelt in diesem Zusammenhang einen robusten multilingualen Keywordspotter.
- Am Institut für Mustererkennung an der Hochschule Niederrhein werden im Rahmen eines ZIM-Projekts kleine Sprachdialogkomponenten zur Hausautomation entwickelt, die z.B. in einer Schalterdose Platz finden sollen. Zur Verbesserung der hausintern stattfindenden Spracherkennung wird der Einsatz mehrerer im Raum verteilter Mikrofone, z.B. in Leuchtmitteln oder in Rauchmeldern, untersucht.
- In einem einmonatigen gemeinsamen Forschungsprojekt zwischen dem MARCS Institute der University of Western Sydney und dem Quality and Usability Lab der TU Berlin werden die neurophysiologischen Auswirkungen von Verzögerungen bei der Audio-Kommunikation untersucht. Dazu werden die Gehirnströme zweier Probanden simultan gemessen (Dual-EEG), während diese Konversationstests durchführen.
- In der ITU-T SG12 Q4 wurde eine Vielzahl neuer Standards bzw. Updates von Standards erstellt. Dies betrifft Freisprechen im KFZ ([P.1100](#) – narrowband, [P.1110](#) – wideband, [P.1120](#) – super-wideband und fullband, sowie [P.1140](#) – jetzt auch inkl. wideband). Die Rapporteur-Gruppe der [Study Group 12 \(SG12\), Question 4](#), unter der Leitung von Rapporteur H. W. Gierlich, HEAD acoustics konzentriert sich derzeit auf

[P.ICC](#) – In-Car Communication Audio Specification. Ein Rapporteurtreffen hierzu fand in Genf am 02.08.2017 statt. Weitere Arbeiten hierzu erfolgten beim regulären SG12-Treffen vom 19.-28.9.2017. Auch bei der ETSI TC STQ wurde eine Vielzahl neuer Standards für mobile und VoIP-Endgeräte und Gateways verabschiedet: [ES 202 737](#) – 40; [TS 103 737](#) – 40; [ES 202 718](#). Darüber hinaus gibt es jetzt standardisierte Testverfahren für Sprachqualität bei Störschall in super-wideband und fullband: [TS 103 281](#). Zwei perzeptiv motivierte Modelle, die beide eine hohe Korrelation zu auditiven Tests aufweisen, sind hier beschrieben.

• Beim letzten Treffen der Study Group 12 der ITU-T wurden neue Empfehlungen zur Evaluierung der Qualität von Sprache und Multimedia diskutiert und beschlossen. So wurde eine neue Methode zur subjektiven Bestimmung der Sprachqualität in einer Konversationssituation beschlossen, welche auf die Dissertation von Köster (2017) zurückgeht. Ferner wurden unterschiedliche Erweiterungen des E-Modells zur Vorhersage von Sprachqualität für Super-Breitband-Übertragung diskutiert. Desweiteren wurde ein Update der Empfehlung ITU-T [P.1301](#) zur perzeptiven Beurteilung von Konferenz- und Telemeeting-Systemen beschlossen. Dieses Update basiert im Wesentlichen auf der Dissertation von [Skowronek](#) (2017). Als erfreuliche Entwicklung macht die ITU-T nun auch ihre Software Tools Library nach ITU-T Rec. G.191 online über [github](#) verfügbar und lädt zur Mitarbeit ein. [Nähere Infos gibt es bei [Sebastian Möller](#), [Alexander Raake](#), [Janto Skowronek](#), TU Berlin und TU Ilmenau]

Dissertationen

• Patrick Bauer: [Artificial Bandwidth Extension of Telephone Speech Signals Using Phonetic A Priori Knowledge](#), Technische Universität Braunschweig (T. Fingscheidt)

Key contributions: HMM approach to artificial speech bandwidth extension including side information. Use of phonetic knowledge as side information: (a) offline with known phonetic transcription, (b) online for telephony using neural-network-based phonetic estimators for fricative sounds to provide the side information. In total 1.0 points of CMOS improvement vs. coded narrowband. Hearing-impaired listeners have more than 20% relative intelligibility improvement with fricative sounds over the telephone.

Journalartikel

• A. Jukić, T. van Waterschoot, S. Doclo: [Adaptive Speech Dereverberation Using Constrained Sparse Multichannel Linear Prediction](#)
This paper presents a blind adaptive dereverberation method using multiple microphones, showing that

the proposed constraint increases the usability for dynamic scenarios, e.g., a moving speech source.

• M. Azarpour und G. Enzner:

[Binaural Noise Reduction via Cue-Preserving MMSE Filter and Adaptive-Blocking-Based Noise PSD Estimation](#)

Diese Lösung bietet eine zweikanalige Signalverbesserung unter den vielschichtigen Randbedingungen niedriger Komplexität, niedriger Laufzeit, und strikter binauraler Merkmalsbewahrung. Die im Haus verfügbare Echtzeitdemonstration profitiert außerdem vom Charme der Einfachheit und Modularität.

• J. Abel and T. Fingscheidt:

[Artificial Speech Bandwidth Extension Using Deep Neural Networks for Wideband Spectral Envelope Estimation](#)

Die vorgestellte Methode zur künstlichen Sprachbandbreitenerweiterung mit tiefen neuronalen Netzen liefert detailreiche Spektren im oberen Band und erzielt hochsignifikante 1,37 CMOS-Punkte Sprachqualitätsverbesserung gegenüber dem AMR-codierten Eingangs-Schmalband-Sprachsignal.

Tagungen (nach Paper Deadline sortiert)

[ISSP](#) (International Seminar on Speech Production)

16.-19.10.2017 in Tianjin, China

[keine Einreichungen mehr]

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

[keine Einreichungen mehr]

[ICASSP](#) – Achtung! Statt Seoul, Südkorea, nun vom

15.-20.04.2018 in Calgary, Alberta, Kanada

Paper Deadline: 27.10.2017 [CfP]

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

Paper Deadline (Abstract): 01.11.2017 [CfP]

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

Paper Deadline: 05.01.2018 [CfP]

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

Paper Deadline (Abstract): 01.12.2017

[Kontakt]

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

Paper Deadline: 17.02.2018 [CfP]

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

Paper Deadline: 23.03.2018 [CfP]

[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]

Stellenangebote

• Die [HEAD acoustics GmbH](#) bietet Stellen für Ingenieure in den Bereichen Forschung, Entwicklung, Applikation, Messtechnik und Support an. [\[Link\]](#)