Tutorial

Session: Tutorial 7
Time: Wednesday December 16, 14:00-17:20
Place: Room Y304

Spoofing and Anti-Spoofing: A Shared View of Speaker Verification, Speech Synthesis and Voice Conversion

Speakers: Zhizheng Wu, University of Edinburgh, UK
Tomi Kinnunen, University of Eastern Finland, Finland
Nicholas Evans, EURECOM, France
Junichi Yamagishi, University of Edinburgh, UK

Abstract

Automatic speaker verification (ASV) offers a low-cost and flexible biometric solution to person authentication. While the reliability of ASV systems is now considered sufficient to support mass-market adoption, there are concerns that the technology is vulnerable to spoofing, also referred to as presentation attacks. Spoofing refers to an attack whereby a fraudster attempts to manipulate a biometric system by masquerading as another, enrolled person. On the other hand, speaker adaptation in speech synthesis and voice conversion techniques attempt to mimic a target speaker's voice automatically, and hence present a genuine threat to ASV systems. The research community has responded to speech synthesis and voice conversion spoofing attacks with dedicated countermeasures which aim to detect and deflect such attacks. Even if the literature shows that they can be effective, the problem is far from being solved; ASV systems remain vulnerable to spoofing, and a deeper understanding of speaker verification, speech synthesis and voice conversion will be fundamental to the pursuit of spoofing-robust speaker verification. While the level of interest is growing, the level of effort to develop spoofing countermeasures for ASV is lagging behind that for other biometric modalities. What's more, the vulnerabilities of ASV to spoofing are now well acknowledged. A tutorial on spoofing and anti-spoofing from the combined perspective of speaker verification, speech synthesis and voice conversion is much needed. The tutorial will attract, not only members of the growing anti-spoofing research community, but also the broader community of general practitioners in speaker verification, speech synthesis and voice conversion. The speakers have led the research community in anti-spoofing for ASV since 2013, have jointly authored a growing number of conference papers, book chapters and the latest survey paper published in Speech Communications in 2015. Between them they have organised two special sessions and one evaluation/challenge (http://www.spoofingchallenge.org/) on the same topic. The experience gained through these activities is be the foundation of this tutorial proposal for APSIPA ASC 2015.

Biographies
Zhizheng Wu (University of Edinburgh, UK, zhizheng.wu@ed.ac.uk) is a research fellow in the Centre for Speech Technology Research (CSTR) at the University of Edinburgh since 2014, and he received the Ph.D. degree from Nanyang Technological University (NTU), Singapore. From 2007 to 2009, he was with Microsoft Research Asia as an intern researcher. He received the best paper award in APSIPA ASC 2012. His research interests includes speech synthesis, voice conversion, spoofing and anti-spoofing, and speaker verification.

Tomi Kinnunen (University of Eastern Finland, Finland, tomi.kinnunen@uef.fi) received the Ph.D. degree in computer science from the University of Eastern Finland (UEF, formerly Univ. of Joensuu) in 2005. From 2005 to 2007, he was an associate scientist at the Institute for Infocomm Research (I2R) in Singapore. Since 2007, he has been with UEF. In 2010-2012, his research was funded by the Academy of Finland in a post-doctoral project focusing on speaker recognition. He is the principal investigator of a 4-year Academy of Finland project focusing on speaker recognition, voice conversion and anti-spoofing techniques. Dr. Kinnunen's team is a regular participant to the NIST speaker recognition evaluations. He was the chair of Odyssey 2014: The Speaker and Language Recognition workshop. He is a partner in a recently kicked-off, large Horizon 2020 funded "OCTAVE" project (octave-project.eu) that trials technology transfer of speaker verification technology to both logical and physical access control scenarios, involving integration of practical spoofing countermeasures.

Nicholas Evans (EURECOM, France, evans@eurecom.fr) is an Assistant Professor at EURECOM where he heads research in Speech and Audio Processing. In addition to other interests in speaker diarization, speech signal processing and multimodal biometrics, and in the scope of the EU FP7 ICT TABULA RASA project, he has studied the threat of spoofing to automatic speaker verification systems and developed new spoofing countermeasures. He serves as Lead Guest Editor for the IEEE T-IFS special issue on Biometric Spoofing and Countermeasures and the IEEE SPM special issue on Biometric Security and Privacy and is an Associate Editor of the EURASIP Journal on Audio, Speech, and Music Processing. He was general co-chair for IWAENC 2014 and will be technical co-chair for EUSIPCO 2015. He also contributed to the organisation of the TABULA RASA Spoofing Challenge held in conjunction with ICB 2013.

Junichi Yamagishi (University of Edinburgh, UK, jyamagis@inf.ed.ac.uk) is a senior research follow and holds an EPSRC Career Acceleration Fellowship in the Centre for Speech Technology Research (CSTR) at the University of Edinburgh. He was awarded a Ph.D. by Tokyo Institute of Technology in 2006 for a thesis that pioneered speaker-adaptive speech synthesis and was awarded the Tejima Prize as the best Ph.D. thesis of Tokyo Institute of Technology in 2007. Since 2006, he has been in CSTR and has authored and co-authored about 100 refereed papers in international journals and conferences. His recent important work includes spoofing against speaker-verification systems using the adaptive speech synthesis and the development of its countermeasures. He is a scientific committee and area coordinator for Interspeech 2012.