

Researchers developing new systems to improve voice recognition

September 4 2013

Graduate students and researchers at UT Dallas have developed novel systems that can identify speaking voices despite conditions that can make it harder to make out a voice, such as whispering, speaking through various emotions, or talking with a stuffy nose.

By improving this ability to detect voice through changing conditions, the research could be used in voice recognition applications such as signing into a bank, getting into locked rooms, logging onto a computer, or verifying purchases online.

The researchers are working in the Center for Robust Speech Systems (CRSS) under the direction of Dr. John Hansen, associate dean for research in the Erik Jonsson School of Engineering and Computer Science.

Using algorithms and modeling techniques, the group's solutions are being sought after by other researchers in the signal processing field. The team has been recognized in international competitions sponsored by the federal government, as well as by the largest professional engineering organization in the world – the Institute of Electrical and Electronics Engineers (IEEE).

Last fall, CRSS lab work was recognized with high rankings in the National Institute of Standards and Technology Speaker Recognition Evaluation. Earlier this summer, the team earned the Best Paper Award given at the IEEE International 38th International Conference on

Acoustics, Speech and Signal Processing.

Speaker verification works by either accepting or rejecting a signal that a sound matches the speech of a person. The process can be complicated by background noise or by the type of microphone used. A speaker's voice can also change if the person hears background noise, is ill or ages.

Oftentimes, speaker verification systems are developed during ideal conditions – when background noise is controlled, when the person is prepared for the recording or reading a prepared text.

In a recent National Institute of Standards and Technology Speaker Recognition Evaluation challenge, officials sent about 80 million voice verification trials with added noise – natural background sounds or artificial computer-generated sounds – to more than 50 universities, research labs and companies throughout the world. Teams had to determine whether speech recordings were from certain speakers or not.

UT Dallas' CRSS lab members had already been refining this process based on earlier research and participation in similar competitions. They had created algorithms that more efficiently converted acoustic waveforms into computer processing for pattern analysis. Their process also eliminated silences and background noise to allow computers to spend more resources on the important speech sounds that reveal speaker identity traits.

Team members had implemented models that incorporated all of these algorithms. They further refined their techniques with regular online input from other teams in The Netherlands, Singapore, Finland, Australia, the United Kingdom, France and Switzerland.

"In our PhD research, we were already identifying these problems for application in the real world without considering the competition," said

Omid Sadjadi, an electrical engineering doctoral candidate at UT Dallas. "Because we were already seeing results on some of the data used for our development, we expected to do well in the challenge."

The team was one of a few that were asked to give oral presentations of their approach to other challenge participants during a two-day workshop held in Orlando, Fla., in December.

"It was a great team effort," said Taufiq Hasan, who was a lead student in the team. "Everyone contributed and was passionate about it. We worked day and night throughout the process, and the team was very humbled by their success."

Earlier this summer, the paper written about their systems earned the a student paper award during the IEEE International Conference – the top conference in the signal processing field that was attended by more than 2,000 people. IBM is a sponsor of the award, and the winning paper received \$500. In addition to Hasan and Sadjadi, other students contributing to the paper were Gang Liu and Navid Shokouhi, doctoral students in electrical engineering. Hynek Boril, an assistant research professor in electrical engineering, was a coauthor on the paper and collaborator in the competition.

"We were surprised because a team of mostly graduate students earning this award is rare, since companies and government research labs with much more resources also compete," Sadjadi said. "That we took a systematic approach to the challenge of recognizing speakers made the difference."

Since the competitions, team members said more companies and researchers have pursued them for collaborations.

"The goal of the competitions such as these has been to inspire research

and start discussions about new and different ways to process speech for the real world, as well as give students the opportunity to work on real-world problems," said Hansen, holder of the Distinguished Chair in Telecommunications. "I'm proud that our students and staff within CRSS have made significant contributions to this aim."

More information: www.icassp2013.com/

Provided by University of Texas at Dallas

Citation: Researchers developing new systems to improve voice recognition (2013, September 4) retrieved 12 September 2024 from <https://phys.org/news/2013-09-voice-recognition.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.