University of Western Sydney
Bringing knowledge to life

RESEARCH DIRECTIONS

Sound selection above the din

Associate Professor Jeesun Kim and Professor Christopher Davis from the MARCS Institute along with Professor Martin Cooke from Basque Foundation for Science are undertaking an investigation into speech recognition in noisy situations. The interdisciplinary project, which is supported by the Australian Research Council Discovery Project, will involve collaboration between human and machine speech recognition researchers, combining signal processing and cognition.

'What characterises human communication is its dynamic nature.' says Associate Professor Kim. Determining how speakers control variations in sounds and rhythm in auditory/visual speech production, and the impact this has on speech intelligibility, is an important foundation for understanding how people communicate in difficult circumstances, such as background noise or with hearing loss'.

Noise is one of the big issues in automatic speech recognition not simply because it interferes with the signal, but because people alter the way they speak in noise. An alternative approach to tackling this problem is based on the idea that speech recognition is still possible with only partial information. In other words, there is no need to reconstitute a clear speech signal because fragments of speech can yield sufficient clues and information to make the overall gist comprehensible.

While speech recognition in noise would seem to provide the perfect topic for collaboration between human and machine speech recognition researchers, little has been done to date. Associate Professor Kim's project uses a detailed computational model ("glimpsing", developed by computer scientist Professor Martin Cooke) and a set of human recognition experiments. The glimpsing model draws on human perception and cognition research that shows comprehension of speech can still be achieved where there is missing data, perceptual grouping and preconceived ideas. The ability to make sense of speech depends upon the manner in which information from these glimpses is used and integrated.

This research project combines expertise in human speech recognition (with an emphasis on auditory-visual processing) and expertise in automatic speech recognition (with an emphasis on missing data techniques). The findings will contribute to theories of speech pattern recognition and the development of a better intelligibility model. The applications of the latter include the potential for speech generation technology that boosts intelligibility by adapting to context, and the design of hearing aids that are more robust to the noisy quirks of everyday living.

Project Title: Understanding speech in noise: linking perception and computation

Funding has been set at: $438,000

Contact Details: j.kim@uws.edu.au
URL http://marcs.uws.edu.au/
March 2014