

SPEECH SYNTHESIS RESEARCH BENEFITS INDUSTRY AND PATIENTS



Synthesised voices are becoming more natural-sounding

In the 2015 film *The Theory of Everything*, Stephen Hawking is presented with his new voice, via a text-to-speech device. The stilted American tones are a shock to his wife. This scenario, no doubt played out in hundreds of cases since text-to-speech became available, is on its way to becoming a thing of the past, thanks to research conducted by the University's Centre for Speech Technology Research (CSTR), part of the University's world-renowned School of Informatics.

Creating a more natural voice

Unlike existing aids, which provide a small range of inappropriate voices, the technology developed by the Centre's researchers is unique in having the ability to create normal-sounding personalised voices from recordings even of disordered speech, therefore enabling people to communicate while retaining personal identity and dignity. The 2010 pilot study used a three minute sample of a Motor Neurone Disease sufferer's voice, and this synthetic voice is now in daily use. A more extensive trial involved the voice banking of 600 people (including the then Scottish First Minister) to gather the data needed to train the underlying statistical model, which has now been used to provide several more reconstructed voices to patients via smartphone or tablet.



CREATING THE VOICE

Complementing the CSTR's research is an ambitious programme of voice recording, carried out by the Euan MacDonald Centre for MND Research. Ideally, a person's voice is recorded soon after diagnosis, and before speech has become affected. The 400 sentences that are read have been chosen to capture all the speech sounds of English in all the different possible combinations.

This voice recording is then "banked" and stored ready to create a synthetic voice for a communication aid if, and when, that person needs one. Using software developed by CSTR scientists, all the parameters of that unique voice can be automatically analysed and synthetically reproduced in a process called "voice cloning".



Voice bank recording



Fighting back against motor neurone disease



The collaborative project is giving people with MND a better voice



The research in Natural Speech Technology is based on a common statistical modelling framework for synthesis and recognition, and is organised into four tracks:

Learning and Adaptation: Models and algorithms for synthesis and recognition that can learn from continuous streams of data, can compactly represent and adapt to new scenarios and speaking styles, and seamlessly adapt to new situations and contexts almost instantaneously.

Natural Transcription: Speech recognisers that can detect “who spoke what, when, and how” in any acoustic environment and for any task domain.

Natural Synthesis: Controllable speech synthesisers that automatically learn from data, and are capable of generating the full expressive diversity of natural speech.

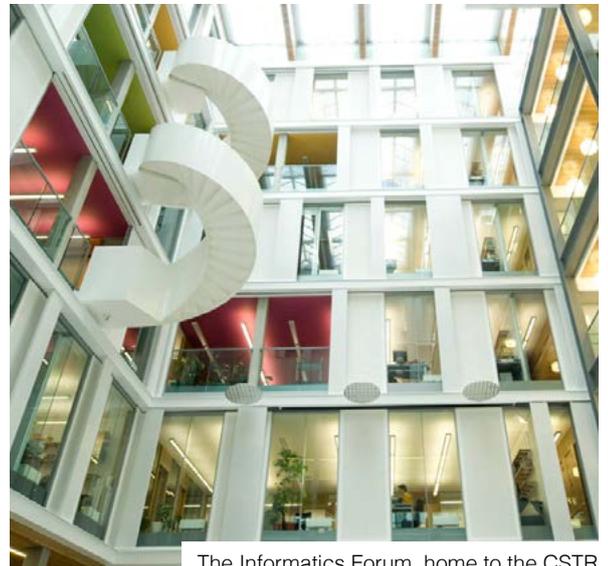
Exemplar Applications: Deployment of these advances in novel applications, with an emphasis on the health/social domain, media archives, and personal listeners.

The results of this research at the School of Informatics have been incorporated into software tools which are freely available, leading to widespread use in commercial products and in research and development, as well as direct commercial spinouts. These systems have become the benchmarks by which others are judged, placing Edinburgh firmly at the centre of speech synthesis development.



Sometimes it is only possible to gain a short recording from the patient. In this case, during the voice cloning process, the synthetically reproduced parameters of a patient’s voice are combined with those of healthy donor voices.

Features of donor voices with the same age, sex and regional accent as the patient are pooled together to form an “average voice model” (AVM), which acts as a base on which to generate the synthetic voice.



The Informatics Forum, home to the CSTR

CONTACT

Centre for Speech Technology Research
Informatics Forum

T: +44 (0)131 650 4434
E: admin@cstr.ed.ac.uk

www.cstr.ed.ac.uk

If you require this document in an alternative format, such as large print, please contact:

Communications and Marketing

Tel: +44 (0)131 650 2252

Email: communications.office@ed.ac.uk



THE UNIVERSITY of EDINBURGH