FURTHER READING
www3.ntu.edu.sg/home/mcloughlin/bionicvoice

Page **
'Linear predictive analysis for ultrasonic speech', F. Ahmadi, I.V. McLoughlin, H.R. Sharifzadeh

The relatively new field of ultrasonic speech is looking set to provide socially important solutions to both medical and non-medical applications

talking ultrasound*

Near-realtime speech regeneration for voice-loss patients using ultrasonic speech for phoneme detection and classification is under development by researchers at Nanyang Technological University in Singapore. Their research is important for a range of applications including: prosthesis for voice-loss patients; improving the performance of artificial speech recognition systems, especially in the presence of audible acoustic noise; and for the more private use of mobile telephones in public.

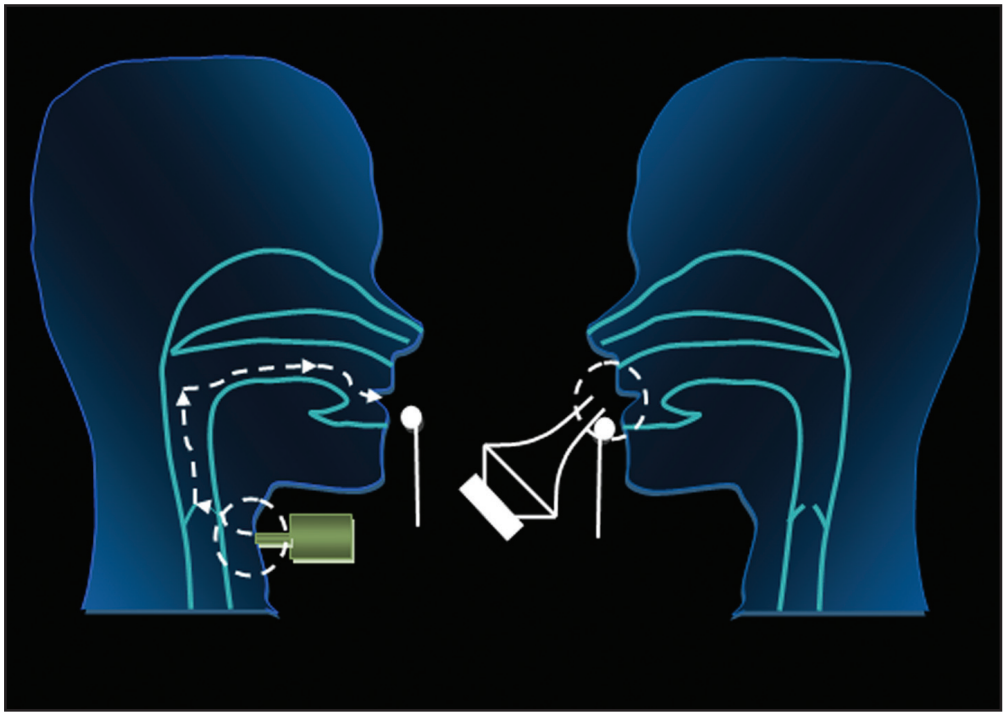
A new voice

Speech relies upon lung exhalation passing through a taut glottis to create audible pitch impulses which resonate through the vocal tract: throat, nasal cavity, mouth and out through the lips. In ultrasonic speech, the resonating pitch is artificially injected at a frequency slightly above the threshold of human hearing. The signal may be applied simultaneously to normal pitch excitation, or in replacement of normal pitch - for example in post-laryngectomised patients who have lost the ability to generate pitch with the surgical removal of their glottis.

Although the term 'ultrasonic speech' was coined 20 years ago, the progress made so far globally revolves primarily around the improvement of automatic speech recognition engines, and Professor Ian McLoughlin and his team at Nanyang Technological University were the first to apply the idea of ultrasonic speech for voice-loss patients. In their *Electronics Letters* paper in this issue they have shown how the 3D complexities of the ultrasonically resonated cavities can be reduced to a tube model which is the most common tool for analysis of audible speech. "This opens up the field of ultrasonic speech research as it can be based upon a very well established body of knowledge, and not require difficult 3D wave equation solutions for every processing step," said McLoughlin.

Practical challenges

This ultrasonic speech analysis is part of a larger project which aims to return the power of natural speech to post-laryngectomised patients: the Bionic Voice project funded by the National Medical Research Council, Singapore. At this stage, however, there are several more research questions to tackle before a viable medical prosthesis can be clinically trialled. One is that there is not necessarily a one-to-one mapping between ultrasonic speech phonemes and audible phonemes. Another difficulty is the practical challenge of efficiently coupling an ultrasonic signal into the vocal tract: bone conduction can not be used since bones are highly nonlinear at ultrasonic frequencies, and so a headset-style microphone, or neck-mounted contact transducer is probably required. Regulatory issues are another major challenge: low-frequency ultrasonics falls outside of the standards governing both medical and industrial ultrasound and McLoughlin and his team are currently working on a new standards proposal for human exposure to low-frequency ultrasound in medical applications.



ABOVE: Speech can be generated by the artificial injection of a low frequency ultrasonic signal into the vocal tract, for example, in the throat or in front of the mouth. A sensing transducer is placed in front of the mouth.
RIGHT: An electrolarynx is one of the current options for post-laryngectomised patients.

Application

The team have patented part of the basic technology behind the voice prosthesis and they are now working toward pre-clinical trials around mid-year. "If the results of these trials on patients proceed to plan, we expect to contract with the medical device industry toward the end of the year and to subsequently hand off commercialisation to them, although the first mass-market product may still require another 3 or 4 years development from that point," said McLoughlin. "The driving force behind this work is to help to improve the quality of life of laryngectomy, and other voice-loss, patients. However we also recognise a quality-of-life argument for a mobile phone version: anyone who has sat next to an unrepentant mobile phone addict on a long train journey, or received an urgent call in a public library or board meeting (with the ring set to vibration mode, I hope), would appreciate the ability for users to communicate at a whisper without actually speaking out loud, and yet for the call recipient to hear a near-normal voice. The phone would be equipped to extract speech parameters from an active ultrasonic mapping of the vocal tract, and this is an exciting application which we may choose to develop later."