

Voice synthesis software delivers high-quality singing automatically

May 27 2015



A song synthesis software that brings out the natural beauty in off-key singing or speaking was introduced to Singaporeans in 2013 through ‘Sing for Singapore’, part of the National Day Parade 2013 mobile application. Credit: International

Speech Communication Association

Whether you give it your best—or worst—effort, I2R Speech2Singing technology will make you sound like the melodious singer you've always wanted to be. The voice synthesis software developed by A*STAR researchers is the first to deliver high-quality singing automatically, while still preserving the original character of your natural voice.

"Many people like [singing](#) but they lack the skills to do so," says Minghui Dong who led the research at the A*STAR Institute for Infocomm Research. "We want to use our [technology](#) to help the average person sing well."

Speech consists of three key elements—content, prosody and timbre. The content is conveyed using words, the prosody—or melody in the case of singing—is expressed through rhythm and pitch, but the timbre is that distinctive quality that makes a banjo sound different from a trumpet, and one singer's voice from another's. I2R Speech2Singing works by polishing melody, while retaining the original content and timbre of a sound.

Existing technologies that focus on correcting melody try to align off-tune sounds either to the closest note on the musical scale or to the exact note in the original score. The former works well for professional singers who may only be slightly out of tune, but cannot fix drastically off-key singing or simply reading out loud. The latter is better at correcting discordant tunes, but ignores many other aspects of melody such as vibrato and vowel-stretching.

Instead, I2R Speech2Singing uses recordings by professional singers as templates against which to correct the [melody](#) of a [singing voice](#) or

convert a speaking voice into a singing one. The software detects the timing of each phonetic sound using [speech recognition technology](#), and then stretches and compresses the duration of the signal using voice conversion technology to match the rhythm to a professional singer's. A speech synthesizer then combines the time-corrected voice with pitch data and background music to produce a beautiful solo.

"When we compared the output with other applications in the market and in research, we realized that our software generated much better [voice](#) quality," says Dong.

Singaporeans were first introduced to the software in 2013 through 'Sing for Singapore', part of the National Day Parade 2013 mobile application (see image). And in 2014, I2R Speech2Singing won the award for best Show & Tell contribution at INTERSPEECH, a major global venue for research on the science and technology of speech communication.

Dong and his team are now working to improve the accessibility of the [software](#) and to add a feature that allows users to tune their singing as they wish.

More information: "I2R Speech2Singing Perfects Everyone's Singing." INTERSPEECH 2014, 2148-2149 (2014). [www.isca-speech.org/archive/in ... h_2014/i14_2148.html](http://www.isca-speech.org/archive/in...h_2014/i14_2148.html)

Provided by Agency for Science, Technology and Research (A*STAR), Singapore

Citation: Voice synthesis software delivers high-quality singing automatically (2015, May 27) retrieved 21 May 2024 from <https://phys.org/news/2015-05-voice-synthesis-software-high-quality-automatically.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.