![PHYS ORG]
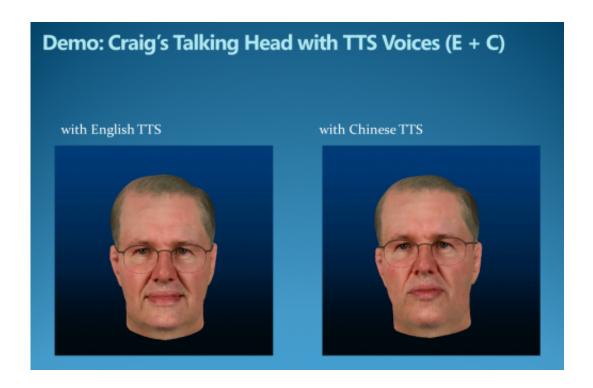
Bilingual avatar speaks Mundie language

March 10 2012, by Nancy Owano

(PhysOrg.com) -- This week's Microsoft Big Idea event, TechFest 2012, presented the latest advances on the part of researchers at Microsoft. A bilingual talking head received much of the attention. Called "Monolingual TTS," the Microsoft research effort involves software that can translate the user's speech into another language and in a voice that sounds like the original user's. As Microsoft explains, with the use of a speaker's monolingual recording, the system's algorithm can render speech sentences in different languages for building "mixed coded

bilingual text to speech (TTS) systems."

According to the team, "We have recordings of 26 languages which are used to build our TTS of corresponding languages. By using the new approach, we can synthesize any mixed language pair out of the 26 languages."

The software does this by first "learning" what the user's voice sounds like. The tool works by using speech recognition, followed by translation, followed by the final text to speech output in a different language. The demo at Microsoft this week used an avatar of Craig Mundie, Microsoft's chief research and strategy officer, to illustrate the system in action.

A synthetic version of Mundie's voice, in English, welcomed the audience to Microsoft Research. Then the voice shifted to the same phrase in Mandarin. The words in Mandarin were reported to be recognizably Mundie's voice.

Some obvious applications might be in a wide range of service-related activities, from the hospitality and tourism market sectors to government workers making use of the software with communities at home and in their international travels.

"We will be able to do quite a few scenario applications," said Frank Soong, who is a principal researcher in Microsoft's speech group. Soong helped create the system with his colleagues at Microsoft's research lab in Beijing.

Microsoft, meanwhile, has had a vision for a while about virtual avatars being used along with this kind of technology. The vision is one where avatars not only look like their users, with photo-realistic effects, but can also successfully mimic their users' voices and approximate their lip

movements to put speech translation into instant, and personalized, action.

Last year, [Mundie](#) was on hand at the Microsoft Research Asia facility in Beijing, where he said that the coming-together of touch, vision, [speech](#) synthesis and recognition, will be an important advancement.

"Another dream we have is that I should be able to sit in my office, send my avatar to meet somebody in Beijing, and I can speak in English and the avatar speaks in Mandarin in real-time," he said. "We want the computer to be a simultaneous translator."

More information: [research.microsoft.com/en-us/p … o-real_talking_head/](#)
via [Technology Review](#)

© 2011 PhysOrg.com

Citation: Bilingual avatar speaks Mundie language (2012, March 10) retrieved 2 May 2024 from https://phys.org/news/2012-03-bilingual-avatar-mundie-language.html