TRN
The Latest Technology Research News

PDA translates speech

By Kimberly Patch, Technology Research News

As speech recognition technology gets better, and as handheld computers get more powerful, audio translators are becoming a more practical proposition.

Researchers from Carnegie Mellon University, Cepstral, LLC, Multimodal Technologies Inc. and Mobile Technologies Inc. have put together a two-way speech-to-speech system that translates medical information from Arabic to English and English to Arabic and runs on an iPaq handheld computer.

The prototype falls short of Star Trek's fictional universal translator in several ways. The system is not transparent -- it must be switched between Arabic-to-English and English-to-Arabic modes. It also works only when the speakers are talking about medical information, and it's only about 80 percent accurate in the lab.

The device shows that it's becoming possible, however, to provide automatic translation using a portable device. "It's good enough to make yourself understood," said Alex Waibel, a professor of computer science at Carnegie Mellon University and a founder of Mobile Technologies Inc.

The effort is one of a series of projects aimed at providing the armed forces with automatic translation for medical and force protection situations and making automatic translation in a wider set of subject areas available for tourists during the 2008 Olympics in Beijing, said Waibel.

The Speechalator prototype uses a built-in microphone and a language-selection button. "You push on the button on the iPaq and speak a sentence and then the translation comes out... in the other language," said Waibel. "You can switch it into the opposite mode when the other person answers and it translates back into your own language."

The software consists of three components: a speech recognizer, a translator, and a speech synthesis engine. "Each one of these components have slight twists to them... in order to work properly for speech translation," said Waibel.

The researchers modified the speech recognition engine to optimize it for handling spontaneous speech.

The translation system has the biggest twist. It extracts the key meaning from the input sentence and translates it to an interlingual, or intermediate representation, and the process depends on the speech being contained in a certain domain, or context, like medical information. "It's just certain nuggets in the phrase that... you need to extract," said Waibel.

The process is akin to constructing a medical-context template that fits the key information, then filling in the template, said Waibel. This process makes it possible for the system to handle spontaneous speech. "We go fishing for the nuggets," he said. But it is also a limitation -- the system must know what domain a speaker is talking about.

The researchers are working on a system that can handle multiple contexts and automatically switch between them, said Waibel. "It can, for example, recognize 'now you're in the hotel reservation domain', or 'now you're in the conference registration mode', or 'now you're talking about medical problem'," he said.

To come up with templates that handle different domains, the researchers collect a lot

of data from people talking in those domains, said Waibel. "The more data we collect the better coverage of all the possible ways you could be saying [these things] becomes," he said.

The difficult part was fitting the software required to do two-way translation in the 64 megabytes of memory contained in the handheld computer, said Waibel. "You need two recognizers, two synthesizers and two translators to make [it] happen in both directions," he said.

The prototype also has a camera attachment that translates text like that on street signs, said Waibel. Snap a picture of a sign with the camera and it automatically extracts the text region, puts the text through a character recognition program, then translates it, he said. "What you then see on the screen is the picture of the scene with a sign and then underneath an English subtitle," he said.

The Speechalator is a practical proof of concept, said Bernard Suhm, a senior scientist at BBN Technologies. "They have engineered the recognizers and other algorithms sufficiently to make them work in real-time on the very limited computational resources of a consumer PDA," he said.

The device carries the promise of being useful not only for medical translation, but also situations such as travel or business, said Suhm. "This work could facilitate the transition of speech-to-speech translation research from the technology side of research, which focuses on algorithms and engineering, to the human factors side of research, which focuses on how people interact with devices, and how useful devices are to tasks from real-life," he said.

The device hasn't yet been run through its paces in a field test, however, Suhm said. "Until then we don't know whether the additional challenges in the field, [like] high levels of noise... or usability issues make it unusable," he said.

The researchers' next steps are to increase the accuracy of the device so that it can deal with ambient noise, and expand the coverage by collecting more data about how people communicate in different domains, said Waibel. The researchers are also working on building learning algorithms that automatically sort out different ways to say the same things.

The researchers' next prototype is scheduled to be finished in the summer of 2004, and will initially have two domains: hotel reservations and medical situations. "Then it [it will] gradually expand towards other domains as are necessary for tourists," he said.

The device can eventually be used to provide translation services for soldiers and relief workers in foreign countries and for travelers, said Waibel.

It could also address a medical problem in the U.S., he said. "There are a number of people in the U.S. who don't speak English and then when going to doctors... feel embarrassed to explain their health problems in front of somebody else who translates," he said.

The researchers are also working on a multilingual speech recognizer that can recognize speech in any of a set of languages, said Waibel. "In that case you might not have to switch the system between the two languages -- you just talk in any language and it will come out in any other language you choose," he said.

And they are aiming to develop a system that combines speech translation with human-to-machine translation, said Waibel. "There are certain situations as a traveler... where you want to communicate with a person in another language, but then there are certain other things which you could just as well do communicating with [a computer]," he said. You would want to talk to another person when ordering food, but communicate with a machine to get directions to a railway station, for example.

Longer-term the researchers are looking for ways to deal with spontaneous speech that is not limited to a certain domain, said Waibel.

Waibel's research colleagues were Ahmed Badran, Robert Frederking, Donna Gates, Alon Lavie, Lori Levin, Tanja Schultz and Dorcas Wallace from Carnegie Mellon University, Alan W. Black from Carnegie Mellon University and Cepstral, LLC, Kevin Lenzo from Cepstral, Monika Woszczyna from Multimodal Technologies Inc., and Jürgen Reichart and Jing Zhang from Mobile Technologies Inc. The researchers presented the results at Eurospeech 2003 in Geneva, Switzerland, September 1 to 4.

The research was funded by the Defense Advanced Research Projects Agency (DARPA).

Timeline: Now, 4 years
Funding: Government
TRN Categories: Applied Technology; Human-Computer Interaction
Story Type: News
Related Elements: Technical paper, "Speechalator: Two-Way Speech-To-Speech Translation on a Consumer PDA" posted at cmu.edu/~awb/papers/...speechalator.pdf," Eurospeech 2003 Geneva, Switzerland September 1-4

Discuss this story or anything else to do with science and technology on the TRN Forum.

TRN needs your help. Please click here for details.

Home   Archive   Resources   Feeds   Offline Publications   Glossary
TRN Finder   Research Dir.   Events Dir.   Researchers   Bookshelf
Contribute   Under Development   T-shirts etc.   Classifieds
Forum   Comments   Feedback   About TRN
TRN Newswire and Headline Feeds for Web sites