Mathematicians Solve the 'Cocktail Party Problem'

August 22 2006

Officials at the CIA and scientists around the world have pondered the "cocktail party problem" for decades. How could they separate one sound - perhaps a voice - from a group of other recorded sounds, perhaps a multitude of voices at a cocktail party? Now, two researchers at the University of Missouri-Columbia have found a mathematical solution to this problem.

"Theoretically, our solution says you should be able to pick up voices on a squeaky old microphone and then separate them all out so that you can hear what each person is saying in his or her own voice," said Peter Casazza, professor of mathematics in MU's College of Arts and Science. "This is a very old problem, and we have the first mathematical solution to it."

Casazza and Dan Edidin, also a professor of mathematics at MU, worked with Radu Balan of Siemens Corporate Research to solve the problem. Their solution shows that it is possible to separate voices and still retain vocal characteristics. Researchers had previously found a solution for separating and reconstructing voices, but they were only able to reconstruct the words spoken, not the characteristics of the voice itself.

"Our solution is called 'signal reconstruction without noisy phase,'" Edidin said. "In speech recognition technology, a 'signal' could be a recording of 25 people in a room talking at the same time. Our solution shows that we can pull out each voice individually, not just with the

words, but with the voice characteristics of each individual. We showed that this 'cocktail party problem' is mathematically solvable."

Although Casazza, Edidin and Balan do not have a computer program that can do this automatically, they hope to find a way to develop one. Currently, their solution runs on a computer, but the process cannot be easily replicated or distributed.

"The computer we use is doing the work without an algorithmic program. It uses a system called a neural net, which is designed for the computer to teach itself. Basically, it works on trial and error," Casazza said. "This isn't consistent and cannot be duplicated easily. We need to find a way to design an implementable algorithm that could do this consistently and quickly."

Casazza said that there are already programs that can separate and reconstruct voices, but they are not completely reliable. For example, such programs have difficulty separating voices with similar pitch characteristics. A program using the researchers' solution would be more exact.

Source: University of Missouri