



Recognizing Speech

Current speech recognition systems perform fairly well in non-conversational settings such as dictation or requests for directory assistance. Applications like this may not appear impressive, but because of accents, inflections, and pauses, even such simple situations require sophisticated techniques to transform speech waveforms into words accurately. One of the most common techniques is a mathematical tool known as a *hidden Markov model*, involving conditional probabilities, which trains on candidate sounds so as to locate the best match for a given input.

Dictating directions to machines, a luxury now, may become a necessity as input devices become too small. Researchers are looking for new mathematical models and algorithms (which will probably use subjects like statistics and machine learning) that can filter out noise, understand casual speech, and adjust to different speakers. Those are difficult problems, but once solved, it won't be long before your voice replaces your keyboard, mouse, and—best of all—your many remotes.

For More Information: *Speech Processing: A Dynamic and Optimization-Oriented Approach*, Li Deng and Douglas O'Shaughnessy, 2003.



Image courtesy of ACM Crossroads.



The **Mathematical Moments** program promotes appreciation and understanding of the role mathematics plays in science, nature, technology, and human culture.

www.ams.org/mathmoments