



音声を認識する

電話番号案内といった（会話を含まない）単純な状況では、現在の音声認識システムは十分機能しています。こんな応用では大したことはないという気がするかもしれませんが、アクセントやなまりや間といった要素を考えると、単純な状況下とはいえ、音声の波形を正確に単語に変換するために高度な技術が使われています。もっともよく使われる技術に、条件付き確率を用いる数学的方法「隠れマルコフモデル」があります。入力された音声にもっともふさわしい単語を特定するため、このモデルによって候補となる音声で学習していくのです。

機械に対し声で命令をだすことは現時点ではまだ高価な技術ですが、入力装置が非常に小さくなると絶対必要な技術となることでしょう。雑音をカットし、不完全な日常会話の文章を理解し、発語者の違いに対応できる数学的モデルとアルゴリズム（それにはたぶん統計学や機械学習といった分野も用いることになるでしょう）を研究者たちは探しています。難しい課題ではありますが、これが解決すれば、キーボードやマウスや、たくさんあり過ぎるリモコン類はいらなくなるはずです。

より詳しくは:*Speech Processing: A Dynamic and Optimization-Oriented Approach*, Li Deng and Douglas O'Shaughnessy, 2003 を参照してください。

Translation courtesy of volunteer mathematicians of the University of Tokyo.



Image courtesy of ACM Crossroads.



AMS

AMERICAN MATHEMATICAL SOCIETY

Mathematical Moments プログラムは、
科学・自然・科学技術・人類の文化における
数学の役割への評価と理解を推し進めています。

www.ams.org/mathmoments