



# Sprache Erkennen

Derzeitige Spracherkennungssysteme leisten relativ gute Dienste außerhalb von Gesprächssituationen, wie zum Beispiel beim Diktieren oder Anfordern von Telefonauskünften. Diese Anwendungen mögen nicht sehr beeindruckend erscheinen, jedoch erfordern Akzente, Tonfall und Pausen schon bei solch einfachen Situationen komplizierte Methoden, um Sprachfrequenzen akkurat in Worte umzuwandeln. Eine der geläufigsten mathematischen Methoden ist das sogenannte “versteckte Markov Modell”, dem bedingte Wahrscheinlichkeiten zugrunde liegen. Trainiert an Sprachproben einer Vielzahl von Versuchspersonen, findet das versteckte Markov Modell die Worte mit der besten Übereinstimmung für gegebene Eingangssignale.

Einem Computer Anweisungen zu diktieren, im Moment noch ein Luxus, wird zukünftig vielleicht eine Notwendigkeit sein bei immer kleiner werdenden Eingabegeräten. Forscher suchen nach neuen mathematischen Modellen und Algorithmen (die wahrscheinlich auf Statistik und künstlicher Intelligenz basieren werden), die Rauschanteile ausfiltern, Umgangssprache verstehen und sich auf unterschiedliche Sprecher einstellen können. Das sind schwierige Probleme. Aber einmal gelöst, wird es nicht lange dauern bis unsere Stimme unsere Tastatur, Maus und—am allerbesten—unsere vielen Fernbedienungen ersetzen wird.

**Für mehr Informationen:** *Speech Processing: A Dynamic and Optimization-Oriented Approach*, Li Deng und Douglas O’Shaughnessy, 2003.

Translation courtesy of Ina Lindemann, American Mathematical Society.



Abbildung mit freundlicher Genehmigung von ACM Crossroads.



Die **Mathematical Moments** sollen die Würdigung und das Verständnis der Rolle der Mathematik in Wissenschaft, Natur, Technologie und in der menschlichen Kultur fördern.

[www.ams.org/mathmoments](http://www.ams.org/mathmoments)