

EINLADUNG

- Zeit: Montag, 06. März 2006, 15.30 Uhr
- Ort: Raum 5056
- Referent: Dipl.-Ing. Hauke Schramm
- Thema: Modellierung von Sprachvariabilität in Spontansprache für Spracherkennungssysteme mit großem Vokabular

Abstract:

Automatische Spracherkennung erlaubt heute die Transkription von gelesener Sprache mit sehr großer Genauigkeit. Die Leistung der Systeme sinkt jedoch drastisch bei unvorbereiteter, *spontaner* Sprache, wie sie z.B. bei natürlicher Mensch-zu-Mensch Kommunikation vorkommt. Zahlreiche Studien haben gezeigt, dass eine der wichtigsten Ursachen hierfür die ausserordentlich große Variabilität von Spontansprache, insbesondere bezüglich alternativer Aussprachen und Sprechgeschwindigkeit, ist. Die sehr unterschiedliche Natur verschiedener Sprachvariabilitäten erfordert individuelle Modellierungsansätze, welche geeignet kombiniert werden müssen.

Im Rahmen dieses Vortrags wird ein neues Verfahren vorgestellt, das eine *spezifische* Behandlung *mehrerer unterschiedlicher* Sprachvariabilitäten im akustischen Modell und Aussprachemodell eines Spracherkennungssystems ermöglicht. Diese Technik basiert auf dem Einsatz klassenspezifischer akustischer Modelle und Aussprachemodelle und deren Kombination auf Wortebene mit Hilfe einer neuartigen lexikalischen Modellkombination. Dazu wird eine Theorie zur effizienten Integration der klassenspezifischen Alternativmodelle in ein 1-Pass Suchverfahren präsentiert, in der die gewichteten Beiträge der Alternativen aufsummiert werden.

Verschiedene Verfahren zur klassenabhängigen Modellierung wichtiger Sprachvariabilitäten (insbesondere Sprechgeschwindigkeit, Aussprache und so genannte "gefüllte Pausen") im akustischen Modell und Aussprachemodell werden vorgestellt, welche vor allem auf der Trennung der Trainingsdaten und systematischen Veränderung der verwendeten Aussprachemodelle basieren. Darüber hinaus wird ein neues diskriminatives Trainingsverfahren zur Optimierung der Gewichte von Aussprachealternativen präsentiert, welches zur weiteren Verbesserung des Aussprachemodells verwendet werden kann.

Evaluationsergebnisse auf hochgradig spontanen medizinischen Diktaten und natürlicher Konversationsprache werden gezeigt und diskutiert. Danach erlaubt die Kombination der genannten Verfahren eine Senkung der Wortfehlerrate von bis zu 14% relativ.

Es laden ein: Die Dozenten der Informatik