



Lautdauerelastizität und Sprechtempo

Abschlussarbeit (Bachelor, evtl. Master)

Motivation: In anspruchsvollen Sprachsyntheseanwendungen (z. B. Vorlesen) sind variable Sprechtempi erforderlich; entsprechend flexible Lautdauermodelle sind bisher nicht verfügbar.

Kurzfassung: Bei langsamerem oder schnellerem Sprechen werden die einzelnen Laute entsprechend länger (bzw. kürzer) gesprochen. Allerdings äußert sich der Einfluss der Sprechgeschwindigkeit nicht gleichmäßig in allen unterschiedlichen Lauten: Einige Laute sind elastischer, d. h. ändern ihre Dauer stärker als andere. Analysen hierfür liegen bisher nur für Japanisch vor [1], das sich durch seine einfache Silbenstruktur vom Deutschen abgrenzt. Donnellan et al. [2] klassifizieren Laute um für die Klassen jeweils separate (lineare oder exponentielle) Dehnbarkeitsmodelle zu erstellen.

Ziel der Abschlussarbeit ist die Erhebung von Sprachdaten mit unterschiedlichen Sprechgeschwindigkeiten, sowie die Analyse der Dehnbarkeiten. Idealerweise soll ein Lautdauer vorhersagendes Modell auf Basis der Datenanalyse erstellt werden, welches variable Sprechtempi unterstützt. Das Vorgehen gliedert sich im einzelnen in folgende Schritte:

- Aneignung von Wissen über spontane, gesprochene Sprache und Sprachsynthese,
- Experimentdesign für die Korpuserstellung unterschiedlich schnell gesprochener Äußerungen,
- automatisches Lautalignment und weitere Erschließung des erstellten Korpus (z. B. mit [3]),
- Analyse der gesammelten Daten bezüglich der Dehnbarkeit von Lautklassen und/oder einzelnen Lauten, ggfs. nach Kontextklassen [2],
- Training und Evaluation eines Vorhersagemodells für Lautdauern auf Basis der gesammelten Daten (für Masterarbeit),
- Evaluation des Modells im Sprachsynthesekontext [4] (für Masterarbeit).

Die erfolgreiche Bearbeitung des Themas erfordert solide Fähigkeiten auf der Kommandozeile, Sorgfalt bei der Datenhaltung, sowie Organisationstalent für die Datensammlung. Für die Erstellung eines Vorhersagemodells sind Kenntnisse maschinellen Lernens hilfreich. Wissen aus den Lehrveranstaltungen Sprachsignalverarbeitung und/oder Computerlinguistik ist von Vorteil, aber nicht notwendig.

Die Abschlussarbeit kann auf Deutsch oder Englisch erarbeitet und verfasst werden; relevante Literatur ist fast ausschließlich auf Englisch verfügbar. Bei jeder Abschlussarbeit streben wir die Veröffentlichung relevanter Forschungsergebnisse auf einer internationalen oder nationalen Konferenz an.

Stichworte: Verarbeitung gesprochener Sprache, Dauermodelle, Maschinelles Lernen, Regression, Psychoakustik, Sprachsynthese

Literatur

- [1] Hisao Kuwabara. „Acoustic properties of phonemes in continuous speech for different speaking rate“. In: *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*. Bd. 4. IEEE. 1996, S. 2435–2438.
- [2] Olivia Donnellan, Elmar Jung und Eugene Coyle. „Speech-adaptive time-scale modification for computer assisted language-learning“. In: *Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference on*. IEEE. 2003, S. 165–169.
- [3] Florian Schiel. „MAUS goes iterative“. In: *Proceedings of the LREC*. 2004.
- [4] Timo Baumann und David Schlangen. „INPRO_iSS: A Component for Just-In-Time incremental Speech Synthesis“. In: *Procs. of ACL System Demonstrations*. Jeju, Korea, 2012.

Kontakt

Timo Baumann (baumann@inf...), Prof. Wolfgang Menzel

URL dieses Dokuments:

