

Spracherkennungssysteme

Inhaltsverzeichnis

1. Was ist „Spracherkennung“?
2. Entwicklung
3. Probleme
4. Arbeitsschritte
5. Aktueller Stand
6. Übung

1. Was ist „Spracherkennung“?

- ▶ Hauptfrage: „Was wurde gesagt?“
- ▶ Teilgebiet der Informatik mit stark phonetischen Ansätzen
- ▶ *„Im engeren Sinn versteht man unter Spracherkennung die Aufgabe, aus einer gesprochenen Äußerung die Wörter richtig zu rekonstruieren.“ (Euler 2006)*
- ▶ Nicht zu verwechseln mit „Sprechererkennung“

2. Entwicklung

- ▶ Erste Systeme in den 1960er Jahren
- ▶ 1984: Spracherkennungssystem von IBM erkennt ca 5.000 englische Einzelworte auf Großrechnern
- ▶ Ab 1993 käuflich zu erwerben für tragbare Computern
- ▶ Als Open Source seit 2004 erhältlich

3. Probleme

- ▶ Unterscheidung zwischen diskreter und kontinuierlicher Sprache
 - Diskret: durch kleine Pausen unterbrochene Sprache
 - Kontinuierlich: flüssiger normaler Sprachgebrauch (keine unnatürlichen Pausen)

Beispiel:

Bald – kommt – der – Frühling.

– diskret

Bald kommt der Frühling.

– kontinuierlich

3. Probleme

- ▶ Größe des Wortschatzes ist begrenzt
- ▶ Homophone:
 - gleich klingende Wörter, mit unterschiedlichen Bedeutungen
 - Hierunter fällt auch die Groß- und Kleinschreibung
- ▶ Stopplaute(Plosive) weisen im spektralen Muster keine charakteristischen Merkmale auf
 - Z. B. der Laut p

3. Probleme

- Benutzerkreise: je mehr Benutzer, desto schwieriger wird die Erkennung für das System

	Einfach	Schwierig
Sprechweise	Einzelne Wörter	Sätze
Wortschatz	Kleiner Wortschatz	Großer Wortschatz
Benutzerkreis	Sprecherabhängig	Sprecherunabhängig

4. Arbeitsschritte

- ▶ Schritt 1: Steigt die Lautstärke, beginnt die Aufnahme.
- ▶ Schritt 2: Die Aufnahme wird digitalisiert.
- ▶ Schritt 3: Filterung von Geräuschen.
- ▶ Schritt 4: Transformation in ein Spektrum.
- ▶ Schritt 5: Abgleich mit gespeicherten Phonemen.
- ▶ Schritt 6: Trigrammmethode zur Verifikation.
 - „Tag am mehr“

4. Arbeitsschritte

- ▶ Textcorpora werden in Trigramme zerlegt
- ▶ „ConTexte“ werden extra für Speech Recognition Systeme entworfen.
- ▶ Die Qualität eines Systems wird in Geschwindigkeit und Akkuratheit gemessen.

5. Aktueller Stand

- ▶ Bei geringem Wortschatz sehr effizient:
 - Bandansagen: „Sagen Sie ja!“
- ▶ Niedersächsischer Landtag testet automatisches Stenographieren: erfolglos.
 - Stenomasken bieten optimale Vorraussetzungen
- ▶ Neuer Ansatz: Kombination mit kameraunterstütztem Lippenlesen
- ▶ Aktuelle Systeme sind meist kommerziell.
 - Mit wenigen Ausnahmen ;)

6. Übung

Testet mithilfe eures Headsets die Anwendung auf der folgenden Seite!

Link: <http://www.midomi.com/>

Fragen: Welche Probleme treten auf?
Was wird gut erkannt?

6.1 Unsere Lösung

- ▶ Probleme bei „Query-by-humming-Systemen“:
 - Ausgelassene und hinzugemogelte Töne
 - Tonartwechsel
 - Verschiedene Tempi
 - Oft werden deutsche Titel nicht erkannt
 - Kleine Datenbank

Quellen

- ▶ Stock, W.G. (2007): Information Retrieval. Informationen suchen und finden. – München: Oldenbourg.
- ▶ Euler, S. (2006): Grundkurs Spracherkennung. – Wiesbaden: Friedr. Vieweg & Sohn Verlag.
- ▶ <http://de.wikipedia.org/wiki/Spracherkennung>

**Vielen Dank für Eure
Aufmerksamkeit!!
Viel Spaß noch ;)**