

dradio.de

URL: <http://www.dradio.de/dlf/sendungen/computer/557777/>

COMPUTER & KOMMUNIKATION

28.10.2006 · 16:30 Uhr



Antje Düsterhöft entwickelt Verfahren zur Suche in Audiomaterial. (Bild: hs-wismar.de)

Suche im Geräusch-Dickicht

Audio-Clipping durchforstet Ton-Archive auf Inhalte

Von Mirco Smiljanic

Ohne Suchmaschinen wäre das Internet wohl nahezu nutzlos. Doch Google und Co behandeln vor allem Textdokumente, nicht aber die enormen Mengen an Audiomaterialien. Das so genannte "Audio-Clipping" soll das zukünftig ändern.

Die Diskussion im Radio war spannend, ein CDU-Politiker griff die aktuelle Sozialpolitik der großen Koalition mit Argumenten an, die so noch niemand öffentlich geäußert hatte. Nur, wann war die Sendung? Letzte Woche, oder vorletzte Woche, oder doch gestern? Was genau hat er gesagt? Wer es wissen will, muss viele Stunden Tonmaterial hören, das erfordert viel Zeit, die nicht jeder investiert. Einfacher wäre da das so genannte Audio Clipping: ein oder zwei Stichworte eingeben und schon sucht die Software in gespeicherten Audiodokumenten. Das Ergebnis wird wie bei Suchmaschinen üblich in Trefferlisten angezeigt, allerdings auch mit den bei Suchmaschinen üblichen Ungenauigkeiten:

Die Probleme liegen da, dass man eventuell ganz allgemeine Suchbegriffe eingibt und dann diese allgemeinen Suchbegriffe in ganz verschiedenen Gebieten findet. Also, man sucht den Ball, findet ihn in Fußball und natürlich auch in Tanzball.

Die Lösung, sagt Antje Düsterhöft, [Professorin für Multimediasysteme und Datenbanken an der Hochschule Wismar](#), sind eingrenzende Suchbegriffe. Womit klar ist, Recherchen in Audiodaten unterscheiden sich kaum von Recherchen in Texten. Selbst der technische Vorgang ist fast identisch:

Stellen Sie sich das so vor, dass man aus dem Schall ein Signal erzeugt und das in Einheiten unterteilt, wobei die Einheiten so genannte Phoneme beschreiben, und dass diese dann zusammengesetzt werden zu Wörtern beziehungsweise zu Äußerungen, und dass man versucht, diese Äußerungen statistisch basiert in einem Lexikon oder in einer Datenbank zu finden, und dann die eigentliche Textfolge abzuleiten.

Spracherkennung wandeln das Audiomaterial zunächst in eine Lautschrift und anschließend über Rechtschreib- und Grammatikkorrekturprogramme in normale, geschriebene Texte um. In diesen Texten findet die Suche statt. Eine häufig angewandte Methode, die aber ihre Grenzen hat. So lassen sich etwa Sprecher nur identifizieren, wenn jemand ihren Namen nennt. Wer direkt nach Forschungsministerin Schavan sucht, muss im Audiomaterial selbst suchen. "Patternmatching" heißt die Methode, bei der die Software charakteristische Muster im Audiomaterial analysiert:

Da wird das Frequenzspektrum aufgesplittet, so dass ich sehen kann: für diesen Moment bildet sich beispielsweise gerade ein Vokal, ein A, ab. Und jetzt kann ich, wenn ich das Signal weiterhin, also kontinuierlich, aufsplittle, diese Lautschrift sozusagen aus dem Audiosignal live herausfiltern und dann dadurch Worte bilden...

...sagt Lars Mainka, Entwicklungschef der Firma [COM VISION](#) in Wismar. Die Spracherkennung basiert auf Methoden, die Ende der 80er Jahre, Anfang der 90er Jahre des letzten Jahrhunderts entwickelt worden sind. Fortschritte haben vor allem leistungsstärkere Prozessoren geliefert, grundsätzlich neue

Methoden sind zurzeit nicht zu erkennen. So leidet beispielsweise das Audio Clipping immer noch darunter, dass die Trefferquote sinkt, sobald mehrere Menschen gleichzeitig reden:

Mein Computer hat keine Konzentrationsfähigkeit, er kann sich nicht auf eine Stimme unter vielen konzentrieren, wenn alle mit der gleichen Intensität sprechen.

Je nach verfahren liegt die Trefferquote zwischen 50 und 90 Prozent. Ab 80 Prozent, sagen Fachleute, wird das Audio Clipping auch für private Nutzer interessant, die endlich einmal in noch unerschlossenen Tonarchiven recherchieren wollen oder einfach wissen möchten, was wo läuft. Lars Mainka:

Ich denke, es wird am ehesten in jenen Bereichen zu sehen sein, wo, nennen wir es einmal diese Video- und Music-on-demand-Generation jetzt immer mehr zum Tragen kommt. Das heißt, das Programm der Öffentlich-rechtlichen und der Privatsender wird immer mehr beispielsweise im Internet angeboten und wenn ich in dieser Programmviefalt etwas finden möchte, was mich interessiert, dann werde ich immer mehr auf audiovisuelle Suchmechanismen zurückgreifen müssen, einfach weil nicht mehr alles mitgeschrieben wird.

© 2006 Deutschlandradio