

Automatische Erkennung von Sprache

Stand der Technik und Einschränkungen

Einleitung

Es gibt zwei grosse Klassen von Spracherkennungssystemen, die heute im Einsatz sind. Zum einen sind das Diktiersysteme und zum anderen Sprachdialogsysteme. Die Systeme können sich dabei in verschiedenen Aspekten unterscheiden. Wichtige Unterscheidungsmerkmale sind sprecherabhängigkeit, bzw. sprecherunabhängigkeit, Grösse des Vokabulars und die Komplexität der Eingabe (komplexe Sätze oder nur Sätze nach bestimmten Mustern) [1]. Im folgenden werde ich einige neuere Spracherkennungssysteme vorstellen, über deren Funktionalität berichten und auf die eben genannten Aspekte eingehen. Am Ende des Berichtes werde ich noch einige Projekte erläutern, welche in Planung stehen und über Einschränkungen und eigene Erfahrungen mit solchen Systemen schreiben.

Diktiersysteme

Diktiersysteme sollen es dem Benutzer ermöglichen ohne Tastatur Texte zu schreiben. Er spricht den Text in ein Mikrofon, worauf der Computer diesen am Bildschirm anzeigt (speech-to-text). Das Vokabular eines Diktiersystemen muss sehr gross sein. Es sollte natürlich möglich sein komplexe Sätze „einzugeben“ und nicht nur solche nach einem bestimmten Muster. Alle guten Diktiersysteme, die es heute gibt sind sprecherabhängig und müssen somit vor der Benutzung zuerst trainiert werden. Im folgenden werde ich ein Produkt, welches in diese Kategorie fällt vorstellen.

VoiceOffice PREMIUM

Mit VoiceOffice Premium kann man in kontinuierlichem Sprechrhythmus bis zu 150 Wörter pro Minute diktieren. Das Programm greift dabei auf ein Basisvokabular von 1'000'000 Wörter zu. Da das Vokabular für einzelne Berufsgruppen sehr verschieden sein kann, ist es möglich spezielle Vokabularerweiterungen (z.B. für Ärzte) zu kaufen. Weiter gibt es auch eine Spracherweiterung für die englische Sprache und ein Programm Namens VoiceOffice Net, welches den Versand der Diktate über ein Netzwerk ermöglicht [2]. Ein komplettes System mit VoiceOffice Premium und VoiceOffice Net und den Erweiterungen würde zum

Beispiel in einem Spital sinnvoll sein [3]. Bei einer Korrektur des Benutzers lernt das Programm und passt das Benutzerprofil an. Bei Tests hat VoiceOffice vor allem durch seine erstaunlich hohe Erkennungsrate überzeugt [4]. Dieses Programm (VoiceOffice Premium) ist für 189 Euro erhältlich (exklusive Erweiterungen) [2].

Sprachdialogsysteme

Sprachdialogsysteme sollen Anfragen wie Telefonauskunft, Fahrplanauskunft, etc. übernehmen. Solche Systeme müssen natürlich vom Sprecher unabhängig sein. Denn müsste jeder, der eine Anfrage machen will, dieses System zuerst trainieren, so wäre die Akzeptanz und auch die Wirtschaftlichkeit solcher Systeme nicht gegeben. Die Betreiber solcher Systeme erhoffen sich einerseits eine Einsparung bei den Personalkosten und auch eine bessere Kundenzufriedenheit, wegen der 24-stündigen Verfügbarkeit [1]. Im folgenden stelle ich einige solche Systeme vor.

Voice Portale in Österreich

In Österreich wurde Anfang dieses Jahres (2002) zwei neuartige Voice Portale eingeführt. Das eine Portal dient Abfrage über aktuelle Informationen. Man kann das Sprachdialogsystem anrufen um sich zum Beispiel über das aktuelle Kinoprogramm oder Freizeittipps zu erkundigen. Das Andere dient als Chatroom [5].

Audiopoint

Audiopoint ist eine Anwendung, die es in den USA gibt. Sie ist speziell auf Reise- und Hotel-Industrie zugeschnitten. Ohne direkte menschliche Interaktion können Kunden Informationen über Flüge, Wetter, Unterhaltung (Kino, Bücher etc.) und Reservationsdienste erhalten [6].

Einige andere Systeme, welche in Europa schon im Einsatz sind, sind die Fahrplanauskunft der Deutschen Bahn, das Aktienkurssystem von Speechworks (Deutschland) und ein Restaurantführer für London [1].

Ankündigungen

Sympafly

Mit Sympafly soll es in Zukunft möglich sein Flüge bei einem Call-Center-Agenten zu buchen. Das System wurden an der CeBIT in diesem Jahr vorgestellt. Das System hat bei vorgängigen

Test eine sehr gute Performance gezeigt und soll in Zukunft von Flughäfen und Reisebüros eingesetzt werden können [7].

Erstes komplett sprachgesteuertes Mobiltelefon

Nuance hat am 13. März 2001 das erste komplett sprachgesteuerte Mobiltelefon angekündigt [8]. Ob dieses aber schon auf dem Markt ist oder wann es erhältlich sein wird habe ich bei meinen Recherchen nicht herausgefunden.

Bluetooth-Freisprecheinrichtung mit Sprachsteuerung

Parrot der französische Hersteller von Spracherkennungssystemen hat ein Headset angekündigt, welches sich mit einem Mobiltelefon kombinieren lässt. Mit diesem Handset soll es ermöglicht werden beim Autofahren sein Handy vollkommen sprachgesteuert zu bedienen [9].

Einschränkungen

Die neusten Diktiersysteme funktionieren gemäss Testberichten relativ gut [4]. Diese müssen jedoch wie oben erwähnt vom Benutzer konfiguriert bzw. trainiert werden, was eine Einschränkung ist. Von den Sprachdialogsystemen habe ich eines der Voice Portale von Österreich getestet. Die Sprache des Gesprächspartners habe ich dabei als angenehm empfunden. Bei diesem System hat man immer verschiedene Befehlsmöglichkeiten. Das System lenkt also die Kommunikation. Das System fragt einem zum Beispiel über welches Thema man etwas wissen möchte und gibt folgende vier Auswahlmöglichkeiten zur Auswahl: Kinotipps, Eventtipps, Witz des Tages und Gewinnspiel. Die Anzahl an Auswahlmöglichkeiten ist also relativ gering und somit auch die Befehle, welche das System erkennen muss. Dafür hat das Programm aber gut funktioniert. Weiter wollte ich den Fahrplan der Deutschen Bahn testen. Dieser ist aber im Moment nicht verfügbar, da er verbessert wird. Dies lässt darauf schliessen, dass solche Systeme noch nicht ausgereift sind. Bei einem Fahrplan hätte man viel mehr Auswahlmöglichkeiten als beim Voice Portal von Österreich, da es viele verschiedene Destinationen gibt. Die Anzahl der Wörter, die ein solches System in einem Schritt erkennen muss, ist also ein kritischer Faktor.

Literaturverzeichnis

- [1] Gregor Erbach, Sprachdialogsysteme für Telefondienste, November 2000.
<http://www.speech.ftw.at/~gor/pub/ivsw00/ivsw2000-erbach.pdf>
- [2] Mende Speech Solutions GmbH & Co. KG, Mende stellt auf der SYSTEMS das erste Diktiersystem mit voller Netzwerkimtegration vor: VoiceOffice.NET, Oktober 2001.
<http://www.speech.de/presse/011001.htm>
- [3] MEDICA Messeführer 2001, Spracherkennung für das Krankenhaus, 2001.
www.speech.de/presse/mende03.pdf
- [4] SoftGuide, VoiceOffice Premium, 2000.
http://www.softguide.de/prog_o/po_0509.htm - 10k
- [5] Sonorys Technology GmbH, Neuartige Voice Portale erstmalig in Österreich!, Januar 2002.
http://www.sonorys.at/press/service/releases/jan_2002/index.htm
- [6] Audiopoint, Inc, What is Audiopoint?, 2001.
<http://www.myaudiopoint.com/faq.jsp?reg1=MrfYV1H3RXosD5IMLf3zFQ®2=BsouP6cJwoIOc0%2BhCdXopA>
- [7] Sympalog Speech Technologies AG, CeBIT 2002: SympaFly-Flugbuchung einfach via Telefon: Anrufen, losreden, abheben, April 2002.
http://www.sympalog.de/aktuelles/pressemitteilungen/presse_02_03_18_2.htm
- [8] Nuance, Nuance Enables World's First Completely Voice-Driven Wireless Phone, März 2001.
http://www.nuance.com/press/2001/20010313_nuance_telefree_firstvoicedrivenwirelessphone.html
- [9] Dirk Delbrouck, Bluetooth-Freisprecheinrichtung mit Spracherkennung, Februar 2002.
<http://news.zdnet.de/story/0,,s2103860,00.html?020206175214>