

Talking Heads

von
Jan Paller

1.2.2001
Seminar "Multimodale Mensch-Maschine-Kommunikation"
WS 00/01
Universität Bielefeld

1 Was sind Talking Heads?

Talking Heads sind synthetische Agenten, die die Mensch-Maschine-Kommunikation vereinfachen, da sie die beidseitige Kommunikation möglichst natürlich gestalten. Sie bestehen aus einem Gesicht, dessen wichtigster Teil die Lippen sind, da diese die akustische Kommunikation am meisten unterstützen.

Zur automatischen Modellierung des Gesichts gibt es mehrere Ansätze:

1. Reproduktion von durch Motion Capturing erhaltenen Daten
2. Kontrolle über einen Puppenspieler, wie es gerne in Filmen gemacht wird
3. Animationsdaten, die sie aus dem gesprochenen Text selbst extrahieren (text-to-speech)
4. der semiautonome Conversational Agent, der aus mehreren Modulen besteht und unter anderem eine Persönlichkeit darstellen und auf den Nutzer eingehen kann

Auch um die Lippenbewegungen darzustellen gibt es mehrere Wege:

1. cartoonartig nur eine kleine Auswahl an Lippenstellungen zur Verfügung stellen
2. die Bewegungen durch eine relativ kleine Anzahl von Parametern zu beschreiben
3. die aktuelle Position aus dem vorherigen und dem kommenden Frame zu berechnen
4. eine Reihe von vorgefertigten Lippenstellungs-Bildern zur Verfügung zu stellen, aus denen dann jeweils das richtige Bild ausgesucht wird
5. die durch EMG-Elektroden gemessenen Muskelkontraktionen als Parameter zu verwenden

Besondere Schwierigkeiten ergeben sich aus Synchronisationsfehlern zwischen Audio- und Video-Kanal. Ein Beispiel dafür ist der McGurk Effekt, bei dem sich aus zwei unterschiedlichen Lauten aus diesen beiden Quellen der Mischlaut ergibt.

2 Welche Talking Head Systeme gibt es und wie werden sie eingesetzt?

In meinem Referat zeige ich drei Talking Head Systeme:

1. Baldi
2. Olga und
3. August

Alle bis auf Baldi sind sind Conversational Agents. Baldi baut auf der text-to-speech-Methode auf.

1. August war sechs Monate in der Öffentlichkeit im Einsatz. Man hat aus den dabei aufgezeichneten Gesprächen mit unerfahrenen Benutzern folgende Erkenntnisse gewonnen:
 - Die Nutzer verwandten hauptsächlich kürzere Äusserungen mit geringer syntaktischer Komplexität.
 - Scheiterte die Kommunikation versuchten, sie es mit Wiederholungen, bei denen teilweise die syntaktische und lexikalische Struktur geändert wurde.
 - Es wurden sehr viele allgemein gebräuchliche Wörter verwendet. Dies trifft besonders auf den sozialen Teil der Äusserungen zu, der relativ gross war.
 - Die Grösse dieses Teils mag unter anderem daran liegen, dass die Entwickler sehr viel Wert auf natürliche Kommunikation legten.
2. Olga ist ein Conversational Agent, der bei der Bearbeitung von Datenbanken hilft. Im Gegensatz zu den beiden anderen hat Olga nicht nur einen Kopf, sondern einen kompletten Körper.
3. Baldi wurde sehr erfolgreich zum Lehren von Sprache bei taubstummen Kinder eingesetzt. Dies liegt unter anderem an der hohen Motivation, die dieses System bei den Kindern erzeugte.

3 Fazit

Talking Heads unterstützen die Kommunikation zwischen Mensch und Maschine, indem sie diese auf ein natürliches (zwischenmenschliches) Level heben. Des Weiteren erhöhen sie die Robustheit dieser Kommunikation durch ihren multimodalen Charakter. Darüberhinaus eignen sie sich sehr gut zum Lernen und Lehren von Sprache.

4 Literaturliste

- Massarro, D.W., et al.: Developing and Evaluating Conversational Agents. In Cassel, J., et al.: Embodied Conversational Agents (S. 287-318). Cambridge
- Internet http://www.nada.kth.se/~osu/olga/e_index.html letzter Zugriff: 26.2.01
- Internet <http://www.speech.kth.se/august/> letzter erfolgreicher Zugriff: 2.2.01 - **nicht mehr zugänglich**