

Gedächtnisprotokoll mündliche Prüfung Sprachkommunikation

Prüfungsdatum: 04.03.2015

Anmerkung:

Sämtliche Fragen und Antworten werden lediglich sinngemäß wiedergegeben. Insbesondere die Richtigkeit der Antworten kann nicht gewährleistet werden.

Frage 1 (aus Fragenkatalog Sprache/Sprachsignaleigenschaften):

Was ist ein LTI System? Was sind die Eigenschaften eines LTI Systems?

Antwort:

LTI: Linear Time Invariant

Time Invariant: Wird ein Eingangssignal verzögert, so ergibt sich die gleiche aber verzögerte Systemantwort.

Linear: $S[a * f(x) + b * g(x)] = a * S[f(x)] + b * S[g(x)]$

Frage 2 (aus Fragenkatalog Sprachsignalanalyse):

Was ist ein Spektrogramm? Was ist in diesem Spektrogramm zu sehen?
(Es wurde ein Bild eines Spektrogramms mit Farbkodierung vorgelegt.)

Antwort:

Drei Dimensionen:

Frequenz (Y-Achse linear), Zeit (X-Achse), Betrag (Farbkodierung)

Stimmhafte Laute:

Grundfrequenz von ungefähr 120Hz und Spektrallinien bei vielfachen der Grundfrequenz. Die Grundfrequenz variierte an manchen Stellen.

Formanten:

Spektrallinien variieren in der Intensität. Stellen größerer Intensität entsprechen den Formanten.

Rauschartige Laute:

Durchgängige Frequenzbereiche höherer Intensität. Zum Beispiel höhere Frequenzen könnten S-Laute sein.

Plosive:

Sehr kurze Pause gefolgt von einem Impuls (hohe Intensität über alle Frequenzen)

Frage 3 (aus Fragenkatalog Auditive Wahrnehmung):

Erklären Sie zeitliche und spektrale Maskierung. Wofür kann man Maskierung nutzen?

Antwort:

Spektrale Maskierung: Maskierer einer bestimmten Frequenz verhindert, dass andere leisere Töne in der Nähe dieser Frequenz wahrgenommen werden können. Die Maskierung ist nicht symmetrisch zur Frequenz. Frequenzen größer als die Frequenz des Maskierers werden deutlicher maskiert als tiefere Frequenzen.

Zeitliche Maskierung: Das menschliche Gehör ist träge. Ein Ton einer bestimmten Frequenz benötigt daher eine gewisse Zeit um stabil "anzuliegen". Verschwindet dieser Ton wieder, so wird genauso eine bestimmte Zeit benötigt, bis dieser Ton nicht mehr wahrgenommen wird.

Ein anderer leiserer Ton kann also auch dann maskiert werden, wenn er für eine kurze Zeit kurz nach dem der Maskierer ertönt. Gleichzeitig kann ein leiserer Ton auch dann maskiert werden, falls er für eine kurze Zeit kurz vor dem Maskierer ertönt. Dies ist deshalb möglich, da der Maskierer bei Erscheinen durch die viel größere Amplitude bereits beim Einschwingvorgang zur Maskierung führt.

Nutzen: Kodierung im Frequenzbereich

Frage 4 (aus Fragenkatalog Spracherkennung):

Nach welchen Annahmen kann man Spracherkennung klassifizieren? Wie wirken sich diese Annahmen (Klassen) auf die Leistung eines Spracherkenners aus?

Antwort:

- Sprache (einfache Sprache → bessere Performance)
- sprecherabhängig (bessere Performance) ↔ sprachunabhängig
- Zielgruppe (Jugendliche → spontane Sprache, Fachpersonal → formale Sprache)
- Allgemeingültig ↔ Spezialisiert auf Anwendung (Komplexität/Größe des Vokabulars)
- Umgebung (viele Umgebungsgeräusche → schlechtere Performance)

Frage 5 (aus Fragenkatalog Sprachdialogsysteme):

Erklären Sie den Aufbau eines Sprachdialogsystems in Hub-Struktur. Welche Vorteile bietet ein solcher Aufbau?

Antwort:

Grundstruktur eines Sprachdialogsystems:

Eingabe → Spracherkennung → Semantische Analyse → Sprachdialogmanager
→ Antwortgenerator → Sprachsynthese → Ausgabe

Hub-Struktur:

Alle Module sind mit dem Hub verbunden (Eingabe sowie Ausgabe). Damit können alle Module auf die Ergebnisse aller anderen Module zugreifen. Es gibt zusätzlich einen History-Tracker. Zum Beispiel kann der Antwortgenerator auf die Semantische Analyse oder History zugreifen und somit Wörter verwenden, die dem Benutzer offensichtlich bekannt sind. Außerdem können die Module bereits anfangen zu arbeiten, noch bevor alle notwendigen Eingaben vollständig sind.