

Zusammenfassung der Dissertation:

**Übertragungsfehlerrobuste Spracherkennung
durch Verwendung von weichen Merkmalsvektoren
des Herrn Valentin Ion**

Die Benutzung des Mobiltelefons, die sich in den letzten Jahrzehnten rasant verbreitet hat, bietet ein bedeutendes Entwicklungspotenzial für sprachbasierte Dienste an. Dafür ist die Remote-Spracherkennung eine geeignete Technologie, wobei für die Erfüllung der Erkennungsaufgabe, statt des mobilen Gerätes ein entfernter Server eingesetzt wird. Trotz der Vorteile einer Client/Server-Architektur, ist die Verschlechterung der Erkennungsgenauigkeit aufgrund Übertragungsfehler eine inhärente Schwachstelle dieses Verfahrens.

Die Robustheit der Remote-Spracherkennung gegen Übertragungsfehler wurde durch viele Forschungsarbeiten angesprochen. Eine sehr verbreitete Fehlerbehandlungstechnik basiert auf der Ersetzung des fehlerbehafteten empfangenen Merkmalsvektors durch einen Schätzwert des fehlerfreien Vektors. Der Schätzwert wird anschließend für die Klassifikation verwendet. Die durch dieses Verfahren ermöglichte Qualitätsverbesserung ist jedoch begrenzt, denn der geschätzte Merkmalsvektor stimmt nicht genau mit dem gesendeten Merkmalsvektor überein, d.h., der Schätzwert ist unsicher.

Diese Arbeit konzentriert sich auf die Änderungen in dem Rahmenwerk der Spracherkennung, die notwendig sind, um die Unsicherheitsinformation auszuwerten. Die neue Darlegung des Klassifikationsproblems ergibt eine neuartige Decodierregel, die anstatt einen Schätzwert anzuwenden, die Posterior-Verteilungsdichtefunktion des gesendeten Merkmalsvektors ausnutzt. Die Annahme, die häufig in der Hidden Markov Model basierten Spracherkennung gemacht wird, dass die einzelnen Beobachtungen unabhängig voneinander sind, wird hier erleichtert. Somit hängt die Verteilungsdichtefunktion nicht nur von einer Beobachtung ab, sondern von der gesamten beobachteten Merkmalsvektorfolge. Dadurch wird die Aussagefähigkeit der Posterior-Verteilungsdichtefunktion erhöht.

Die neuartige Decodierregel ermöglicht die Realisierung eines gegen Übertragungsfehler robusten Remote-Spracherkennungssystems. Es wird aufgezeigt, wie die oben erwähnte Verteilungsdichtefunktion für Kommunikationsnetzwerke, die Bitfehlern oder Paketverluste aufweisen, ausgerechnet werden kann. Das zur Ausrechnung zugrunde gelegte wahrscheinlichkeitstheoretische Model fasst sowohl A-priori Kenntnisse über den Merkmalsvektor als auch die Bitzuverlässigkeitsinformation über die empfangenen Daten zusammen.

Die Verbesserung der Robustheit unter verschiedenen widrigen Übertragungsumständen wird für Erkennungsaufgaben mit kleiner und mittlerer Vokabulargröße experimentell beurteilt. Die Ergebnisse für die beiden Modelle der Remote-Spracherkennung: verteilte und Netzwerk-basierte Spracherkennung sind dargestellt. Die Letztere setzte für Voice-over-IP verbreitete Sprachcodierungsverfahren ein.