

Sprachverarbeitung: Musterlösung zur Übung 22

Ziffernerkenner mit CDHMM-Wortmodellen

Aufgabe 1: Viterbi-Algorithmus für den kontinuierlichen Fall

Die Musterlösung der Funktion `[logP,optQ] = log_cont_viterbi_alg(a,mu,sigma,c,X)` ist im Directory `Uebung22/Loesung/` zu finden.

Aufgabe 2: Ermitteln der Erkennungsrate

Mit dem Viterbi-Algorithmus erkennt der Ziffernerkenner 94.5 % der Wörter des Testsets korrekt.

Diese Erkennungsrate entspricht etwa der Erwartung, weil die Wort-HMM nur 8 emittierende Zustände und nur eine Mischkomponente aufweisen.

Aufgabe 3: Verwechslungsmatrix und Erkennungsrate pro Sprecher

Die Musterlösung des Matlab-Skripts `ueb22_3` ist im Directory `Uebung22/Loesung/` zu finden. Sie zeigt folgendes:

a) Verwechslungsmatrix für die Ziffern

Die Ziffern werden nicht gleich gut erkannt. Verwechselt werden insbesondere Ziffern, die ähnlich sind, z.B. “zwei” und “drei” jedoch kaum “zwo” und “dreiß”.

b) Erkennungsrate pro Sprecher

Der Worterkenner ist mit Daten von 61 Sprecherinnen und Sprechern trainiert worden und ist daher in einem gewissen Mass sprecherunabhängig.

Für die einzelnen Sprecher des Testsets (die Mengen der Trainings- und Testsprecher sind disjunkt) ergeben sich trotzdem recht unterschiedliche Erkennungsraten, je nachdem, wie gut die Stimme und die Sprechweise eines Sprechers in den Trainingsdaten repräsentiert ist. Gut repräsentiert ist offensichtlich Sprecher 2, jedoch nicht Sprecher 5, der einen hoch-deutschen Akzent hat.