Zürcher Hochschule
für Angewandte Wissenschaften

School of
Engineering

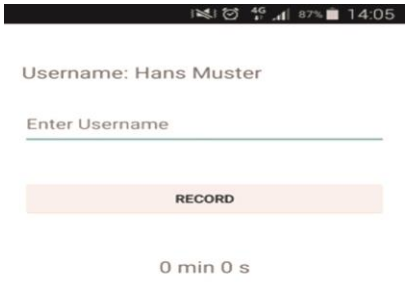
InIT Institut für angewandte
Informationstechnologie

Talkalyzer: Entwicklung einer Android-App zur Erkennung von Redeanteilen im Unterricht

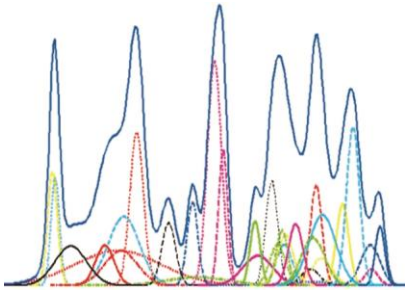
Eine Person kann automatisch anhand ihrer Stimme erkannt werden. Dieses Verfahren wird auch automatische Sprechererkennung genannt und kann mit dem heutigen Stand der Technik in einer Smartphone-Applikation implementiert werden. Der Benutzer kann mit einer Smartphone-App ausserdem seinen Redeanteil für ein aufgenommenes Gespräch berechnen lassen, was eine Vielzahl von Anwendungsmöglichkeiten ermöglichen würde. So lässt sich dann beispielsweise herausfinden, wie gross der Redeanteil eines Dozenten während einer Vorlesung ist beziehungsweise wie stark sich die Studenten am Unterricht beteiligen. Jedoch findet sich zurzeit keine entsprechende Smartphone-App auf dem Markt, die so ein Verfahren ermöglichen würde. In dieser Bachelorarbeit wurde zur Berechnung des Redeanteils eine Android-App entwickelt.
Für die Implementierung der Android-App wurden verschiedene Bibliotheken getestet, wobei die Kombination von drei Bibliotheken zum Erfolgt geführt hat. Anhand der Funktionen der Sphinx4 Bibliothek wird das Sprachsignal der Stimme zuerst im Preprocessing-Vorgang verstärkt und für die Weiterverwendung in viele kleine Blöcke unterteilt. Aus diesen werden dann die Features extrahiert, welche die Merkmale der Stimme repräsentieren. Mit dem k-Means-Algorithmus aus der Apache Commons Math Bibliothek werden im nächsten Schritt die Features in mehreren Gruppen unterteilt, woraus sich dann die Parameter bestimmen lassen. Diese Parameter dienen dann zum Initialisieren des Gaussian Mixture Models (GMM) aus der sclib Bibliothek, welches ein Sprechermodell der Stimme darstellt. Mit der Log-Likelihood-Funktion eines GMM kann schlussendlich die Wahrscheinlichkeit einer Stimme zum Sprechermodell bestimmt werden. Zum Testen der Android-App wurden die TIMIT-Daten verwendet, die hochqualitative Aufnahmen von 630 Sprechern enthalten. Die Testergebnisse zeigten bezüglich der 630 Sprechern eine 100%-Sprechererkennungsrate. In weiteren Experimenten wurde die Android-App noch anhand von selbst aufgezeichneten Gesprächen getestet, die mit dem Smartphone-Mikrophone aufgenommen wurden. Diese Experimente führten zur Erkenntnis, dass für eine hohe Sprechererkennungsrate ein Gespräch am besten alle 2 Sekunden analysiert werden sollte. Im Allgemeinen zeigen die Ergebnisse, dass die Android-App für Aufnahmen mit hoher Qualität eine sehr gute Sprechererkennungsrate erreicht.

Diplomand
Arash Besadi

Dozierende
Mark Cieliebak
Thilo Stadelmann



Android-App zur Berechnung des Redeanteils



GMM: ein Sprechmodell der Stimme