



## AFCEA Bonn e.V. Studienpreis 2018 Kernthesen der Arbeit

<b>Titel der Arbeit:</b>	Neural Networks and Speaker Identification
<b>Tag der Einreichung:</b>	13.05.2017
<b>Hochschule:</b>	Rheinische Friedrich-Wilhelms-Universität Bonn
<b>Name des Verfassers:</b>	Kevin Wilkinghoff
<b>Betreuender Professor:</b>	Prof. Dr. Frank Kurth

*Kurze Beschreibung (1 Seite !) der Kernthesen.*

*Was ist die Quintessenz der Arbeit?*

Bei Fraunhofer FKIE werden im Kontext des Projektbereichs EvA (Erschließung von Audiomassendaten) Methoden zur robusten Detektion und Klassifikation von Audioereignissen in digitalen Signalen untersucht. Sprachsignale stellen dabei eine wesentliche Signalklasse dar. Die automatische Sprechererkennung bzw. -klassifikation ist ein wichtiger Schritt zur Informationserschließung solcher Sprachsignale. Aus Anwendungssicht sind dabei oftmals schwierige Randbedingungen zu beachten, die Methoden nach dem Stand der Technik vor Probleme stellen. Zu diesen Randbedingungen zählen eine schlechte Signalqualität, unzureichendes Trainingsmaterial zur geeigneten Parametrierung der Sprechererkenner und eine Abweichung in der Signalqualität zwischen Trainingsdaten und Anwendungsdaten (Mismatched Conditions).

Herrn Wilkinghoffs Aufgabenstellung bestand vor diesem Hintergrund einerseits darin, verschiedene Möglichkeiten zur Anpassung bekannter Methoden zur Sprecherklassifikation auf ein spezielles Anwendungsszenario zu untersuchen. Andererseits sollte er Kombinationsmöglichkeiten dieser Methoden mit neuen Ansätzen aus dem Bereich des Deep Learnings, sowie den Einsatz komplementärer Sprachmerkmale untersuchen. Das hierbei zugrunde gelegte spezielle Anwendungsszenario war die closed-set Sprecherklassifikation für eine relativ geringe Sprechermenge (hier: ca. 10 bekannte Sprecher), eine sehr geringe Menge von Trainingsmaterial (in der ungefähren Größenordnung von einigen 10 Sekunden) und Telefonie-Sprache (Analogtelefonie aus dem NTIMIT-Datensatz). Als Benchmark wird ein am FKIE entwickelter Sprecherklassifikator herangezogen, der eine Klassifikationsfehlerrate von etwas über 5% erreicht.

Herrn Wilkinghoffs Arbeit stellt vor dem Hintergrund der Aufgabenstellung eine informatisch und mathematisch ausgereifte Leistung dar. Er hat die Arbeit sehr zügig und selbständig bearbeitet. Dabei hat er sich insbesondere selbständig in zugrundeliegende Primär- und Sekundärliteratur, sowie eine Reihe von Softwarebibliotheken eingearbeitet. Die vorgelegte Arbeit ist von exzellenter Qualität. Sie verknüpft die fundierte theoretische Darstellung der verwendeten Methoden mit einem komplexen, systematisch strukturierten Experimentalteil. Die im Theorieteil durchgeführten Herleitungen zusammen mit dem Schließen gewisser in der Literatur offen gelassener Lücken belegen Herrn Wilkinghoffs große Leistungsfähigkeit in dieser Hinsicht. Die sehr anspruchsvolle Aufgabe der geeigneten Parametrierung und Kombination der verschiedensten Methoden der Mustererkennung im praktischen (siehe S. 78 der Arbeit für eine Übersicht) Teil meistert Herr Wilkinghoff ebenfalls sehr gut. Die als Resultat von ihm erzielten Praxisergebnisse - eine Verbesserung von über 5 % auf ca. 1,4 % Klassifikationsfehler - ist als herausragend zu bezeichnen.

Die Ergebnisse der Arbeit sind bereits als Konferenzbeitrag (EUSIPCO'17) und als FKIE-Bericht publiziert. Weitere Konferenzbeiträge werden eingereicht (u.a. EUSIPCO'18). Eine Journal-Publikation soll in Folge vorbereitet werden. Auf praktischer Seite sind einige der aufgeworfenen Fragestellungen Grundlage aktueller Forschungsaktivitäten bei Fraunhofer FKIE. Die resultierenden Softwaremodule werden aktuell schon im realen Anwendungskontext erprobt.