Modulbezeichnung:	**Sprach- und Audiosignalverarbeitung (SAV)** (Speech and Audio Signal Processing)	**5 ECTS**
Modulverantwortliche/r:	Walter Kellermann	
Lehrende:	Walter Kellermann,	

Startsemester: SS 2022	Dauer: 1 Semester	Turnus: jährlich (SS)
Präsenzzeit: 60 Std.	Eigenstudium: 90 Std.	Sprache: Englisch

Lehrveranstaltungen:

Sprach- und Audiosignalverarbeitung (SS 2022, Vorlesung, 3 SWS, Walter Kellermann)
Übung zur Sprach- und Audiosignalverarbeitung (SS 2022, Übung, 1 SWS, N.N.)

Empfohlene Voraussetzungen:

Vorlesung Signale und Systeme I & II

Inhalt:

It concentrates on algorithms for speech and audio signal processing with applications in telecommunications and multimedia, especially

- physiology and models for human speech production and hearing: source-filter model, filterbank model of the cochlea, masking effects,
- representation of speech and audio signals: estimation and representation of short-term and long-term statistics in the time and frequency domain as well as the cepstral domain; typical examples and visualizations
- source coding for speech and audio signals: criteria, scalar and vector quantization, linear prediction, prediction of the pitch frequency; waveform coding, parametric coding, hybrid coding, codec standards (ITU, GSM, ISO-MPEG)
- basic concepts of automatic speech recognition (ASR): feature extraction, dynamic time warping, Hidden Markov Models (HMMs)
- basic concepts of speech synthesis: text-to-speech systems, model-based and data-driven synthesis, PSOLA synthesis system
- signal enhancement for acquisition and reproduction: noise reduction, acoustic echo cancellation, dereverberation using single-channel and multichannel algorithms.

Es werden Grundlagen und Algorithmen der Verarbeitung von Sprach- und Audiosignalen mit Anwendungen in Telekommunikation und Multimedia behandelt, insbesondere:

- Physiologie und Modelle der Spracherzeugung und des Hörens: Quelle-Filter-Modell, Filterbank-Modell der Cochlea; Maskierungseffekte;
- Darstellung von Sprach- und Audiosignalen: Schätzung und Darstellung der Kurzzeit- und Langzeitstatistik in Zeit-, Frequenz- und Cepstralbereich; typische Beispiele, Visualisierungen;
- Quellencodierung für Sprache und Audiosignale: Kriterien; skalare und vektorielle Codierung; lineare Prädiktion; Pitchprädiktion; Wellenform-/Parameter-/Hybrid-Codierung; Standards (ITU, GSM, ISO-MPEG)
- Spracherkennung: Merkmalextraktion, Dynamic Time Warping, Hidden Markov Models
- Grundprinzipien der Sprachsynthese: Text-to-Speech Systeme, modellbasierte und datenbasierte Synthese, PSOLA-Synthese
- Signalverbesserung bei Signalaufnahme und - wiedergabe: Geräuschbefreiung, Echokompensation, Enthallung mittels ein- und mehrkanaliger Verfahren;

Lernziele und Kompetenzen:

The students

- understand basic physiological mechanisms of human speech production and hearing and can apply them for the analysis of speech and audio signals
- apply basic methods for the estimation and representation of the short-term and long-term statistics of speech and audio signals and can analyze such signals by means of these methods
- understand current methods for source coding of speech and audio signals and can analyze current coding standards

- verstehen die Grundbausteine von Spracherkennungssystemen und können deren Funktion mittels Rechnersimulation analysieren
- understand the basic principle of text-to-speech systems and can apply fundamental methods for speech synthesis
- can apply basic algorithms for speech enhancement and understand their functionality for real-world data.

Die Studierenden

- verstehen die grundlegenden physiologischen Mechanismen der Spracherzeugung und des Hörens beim Menschen und können diese zur Analyse von Sprach- und Audiosignalen anwenden
- wenden die grundlegenden Methoden zur Schätzung und Darstellung der Kurzzeit- und Langzeitstatistik von Sprach- und Audiosignalen an und können diese damit analysieren
- verstehen die aktuellen Methoden zur Quellencodierung von Sprache-und Audiosignalen und können aktuelle Codierstandards analysieren
- verstehen die Grundbausteine von Spracherkennungssystemen und können deren Funktion mittels Rechnersimulation analysieren
- verstehen die Grundprinzipien von Text-to-Speech Systemen und können elementare Algorithmen zur Sprachsynthese anwenden
- können elementare Algorithmen zur Signalverbesserung anwenden und für reale Daten analysieren

Literatur:

Gemäß themenbezogenen Angaben in der Lehrveranstaltung

Verwendbarkeit des Moduls / Einpassung in den Musterstudienplan:

Das Modul ist im Kontext der folgenden Studienfächer/Vertiefungsrichtungen verwendbar:

[1] **Computational Engineering (Rechnergestütztes Ingenieurwesen) (Master of Science)**
(Po-Vers. 2013 | TechFak | Computational Engineering (Rechnergestütztes Ingenieurwesen) (Master of Science) | Gesamtkonto | Wahlpflichtbereich Technisches Anwendungsfach | Information Technology - DSP | Speech and Audio Signal Processing)

Dieses Modul ist daneben auch in den Studienfächern "123#67#H", "Advanced Signal Processing & Communications Engineering (Master of Science)", "Berufspädagogik Technik (Bachelor of Science)", "Berufspädagogik Technik (Master of Education)", "Communications and Multimedia Engineering (Master of Science)", "Computational Engineering (Master of Science)", "Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science)", "Elektrotechnik, Elektronik und Informationstechnik (Master of Science)", "Information and Communication Technology (Master of Science)", "Informations- und Kommunikationstechnik (Bachelor of Science)", "Informations- und Kommunikationstechnik (Master of Science)", "Mathematik (Bachelor of Science)", "Mechatronik (Master of Science)", "Medizintechnik (Master of Science)", "Wirtschaftsingenieurwesen (Master of Science)" verwendbar.

Studien-/Prüfungsleistungen:

Speech and Audio Signal Processing (Prüfungsnummer: 64601)
(englische Bezeichnung: Speech and Audio Signal Processing)

Prüfungsleistung, Klausur, Dauer (in Minuten): 90
Anteil an der Berechnung der Modulnote: 100%

Erstablegung: SS 2022, 1. Wdh.: WS 2022/2023
1. Prüfer: Walter Kellermann
