

Modulbezeichnung: Sprach- und Audiosignalverarbeitung (SAV) **5 ECTS**
(Speech and Audio Signal Processing)

Modulverantwortliche/r: Walter Kellermann

Lehrende: Mhd Modar Halimeh, Walter Kellermann

Startsemester: SS 2020

Dauer: 1 Semester

Turnus: jährlich (SS)

Präsenzzeit: 60 Std.

Eigenstudium: 90 Std.

Sprache: Englisch

Lehrveranstaltungen:

Sprach- und Audiosignalverarbeitung (SS 2020, Vorlesung, 3 SWS, Walter Kellermann)

Übung zur Sprach- und Audiosignalverarbeitung (SS 2020, Übung, 1 SWS, Mhd Modar Halimeh)

Empfohlene Voraussetzungen:

Vorlesung Signale und Systeme I & II

Inhalt:

The course 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.

Die Vorlesung behandelt Grundlagen und Algorithmen der Verarbeitung von Sprach- und Audiosignalen mit Anwendungen in Telekommunikation und Multimedia, 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, Enthaltung 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] **123#67#H**

(Po-Vers. 2008 | TechFak | Computational Engineering (Master of Science with Honours) | Gesamtkonto | Wahlpflichtbereich Technisches Anwendungsfach | Speech and Audio Signal Processing)

[2] **123#67#H**

(Po-Vers. 2013 | TechFak | Computational Engineering (Master of Science with Honours) | Gesamtkonto | Wahlpflichtbereich Technisches Anwendungsfach | Information Technology - DSP | Speech and Audio Signal Processing)

[3] **Advanced Signal Processing & Communications Engineering (Master of Science)**

(Po-Vers. 2016w | TechFak | Advanced Signal Processing & Communications Engineering (Master of Science) | Gesamtkonto | Wahlpflichtmodule | Technical Mandatory Electives | Speech and Audio Signal Processing)

[4] **Berufspädagogik Technik (Bachelor of Science)**

(Po-Vers. 2011 | TechFak | Berufspädagogik Technik (Bachelor of Science) | Studienrichtung Elektrotechnik und Informationstechnik | Gesamtkonto | Wahlpflichtmodule Fachwissenschaft | Speech and Audio Signal Processing)

[5] **Berufspädagogik Technik (Master of Education)**

(Po-Vers. 2010 | TechFak | Berufspädagogik Technik (Master of Education) | Studienrichtung Elektro- und Informationstechnik (Masterprüfungen) | Wahlpflichtmodule Fachwissenschaft | Speech and Audio Signal Processing)

[6] **Berufspädagogik Technik (Master of Education)**

(Po-Vers. 2018w | TechFak | Berufspädagogik Technik (Master of Education) | Gesamtkonto | Wahlpflichtmodule Fachwissenschaft | Speech and Audio Signal Processing)

[7] **Communications and Multimedia Engineering (Master of Science): 2. Semester**

(Po-Vers. 2011 | TechFak | Communications and Multimedia Engineering (Master of Science) | Gesamtkonto | Pflichtmodule | Speech and Audio Signal Processing)

[8] **Computational Engineering (Rechnergestütztes Ingenieurwesen) (Master of Science)**

(Po-Vers. 2008 | TechFak | Computational Engineering (Rechnergestütztes Ingenieurwesen) (Master of Science) | Gesamtkonto | Wahlpflichtbereich Technisches Anwendungsfach | Speech and Audio Signal Processing)

[9] **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)

[10] **Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science): 5-6. Semester**

(Po-Vers. 2007 | TechFak | Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science) | Studienrichtungen (Wahlpflichtmodule) | Studienrichtung Informationstechnik | Vertiefungsmodule Informationstechnik |

Speech and Audio Signal Processing)

- [11] **Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science): 5-6. Semester**
(Po-Vers. 2009 | TechFak | Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science) | Studienrichtungen | Studienrichtung Informationstechnik | Vertiefungsmodule (Wahlpflichtmodule) Informationstechnik | Speech and Audio Signal Processing)
- [12] **Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science)**
(Po-Vers. 2017w | TechFak | Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science) | Studienrichtung Informationstechnik | Vertiefungsmodule Informationstechnik | Speech and Audio Signal Processing)
- [13] **Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science)**
(Po-Vers. 2019w | TechFak | Elektrotechnik, Elektronik und Informationstechnik (Bachelor of Science) | Gesamtkonto | Studienrichtung Informationstechnik | Kern- und Vertiefungsmodule Informationstechnik | Vertiefungsmodule Informationstechnik | Speech and Audio Signal Processing)
- [14] **Elektrotechnik, Elektronik und Informationstechnik (Master of Science): 1-4. Semester**
(Po-Vers. 2010 | TechFak | Elektrotechnik, Elektronik und Informationstechnik (Master of Science) | Studienrichtung Informationstechnik | Vertiefungsmodule Informationstechnik | Speech and Audio Signal Processing)
- [15] **Elektrotechnik, Elektronik und Informationstechnik (Master of Science)**
(Po-Vers. 2015s | TechFak | Elektrotechnik, Elektronik und Informationstechnik (Master of Science) | Gesamtkonto | Studienrichtung Informationstechnik | Vertiefungsmodule Informationstechnik | Speech and Audio Signal Processing)
- [16] **Information and Communication Technology (Master of Science)**
(Po-Vers. 2019s | TechFak | Information and Communication Technology (Master of Science) | Gesamtkonto | Pflicht- und Wahlpflichtmodule der Studienschwerpunkte | Schwerpunkt Embedded Systems | Wahlpflichtmodul aus EEI im Schwerpunkt Embedded Systems | Speech and Audio Signal Processing)
- [17] **Information and Communication Technology (Master of Science)**
(Po-Vers. 2019s | TechFak | Information and Communication Technology (Master of Science) | Gesamtkonto | Pflicht- und Wahlpflichtmodule der Studienschwerpunkte | Schwerpunkt Media Processing Systems | Wahlpflichtmodul aus EEI im Schwerpunkt Media Processing Systems | Speech and Audio Signal Processing)
- [18] **Informations- und Kommunikationstechnik (Bachelor of Science)**
(Po-Vers. 2009 | TechFak | Informations- und Kommunikationstechnik (Bachelor of Science) | Gesamtkonto | Wahlpflichtmodule, Wahlmodule, Seminar, Praktika, Bachelorarbeit | Wahlpflichtmodule aus Katalog für IuK | Speech and Audio Signal Processing)
- [19] **Informations- und Kommunikationstechnik (Master of Science)**
(Po-Vers. 2010 | TechFak | Informations- und Kommunikationstechnik (Master of Science) | Gesamtkonto | Schwerpunkte im Masterstudium | Schwerpunkt Eingebettete Systeme | Wahlpflichtmodule | Wahlpflichtmodul aus EEI im Schwerpunkt Eingebettete Systeme | Speech and Audio Signal Processing)
- [20] **Informations- und Kommunikationstechnik (Master of Science)**
(Po-Vers. 2010 | TechFak | Informations- und Kommunikationstechnik (Master of Science) | Gesamtkonto | Schwerpunkte im Masterstudium | Schwerpunkt Kommunikationsnetze | Wahlpflichtmodule | Wahlpflichtmodul aus EEI im Schwerpunkt Kommunikationsnetze | Speech and Audio Signal Processing)
- [21] **Informations- und Kommunikationstechnik (Master of Science)**
(Po-Vers. 2010 | TechFak | Informations- und Kommunikationstechnik (Master of Science) | Gesamtkonto | Schwerpunkte im Masterstudium | Schwerpunkt Multimediasysteme | Wahlpflichtmodule | Wahlpflichtmodul aus EEI im Schwerpunkt Multimediasysteme | Speech and Audio Signal Processing)
- [22] **Informations- und Kommunikationstechnik (Master of Science)**
(Po-Vers. 2010 | TechFak | Informations- und Kommunikationstechnik (Master of Science) | Gesamtkonto | Schwerpunkte im Masterstudium | Schwerpunkt Realisierung von Informations- und Kommunikationssystemen | Wahlpflichtmodule | Wahlpflichtmodul aus EEI im Schwerpunkt Realisierung von Informations- und Kommunikationssystemen | Speech and Audio Signal Processing)
- [23] **Informations- und Kommunikationstechnik (Master of Science)**
(Po-Vers. 2016s | TechFak | Informations- und Kommunikationstechnik (Master of Science) | Gesamtkonto | Schwerpunkte im Masterstudium | Schwerpunkt Eingebettete Systeme | Wahlpflichtmodule | Wahlpflichtmodul aus EEI im Schwerpunkt Eingebettete Systeme | Speech and Audio Signal Processing)
- [24] **Informations- und Kommunikationstechnik (Master of Science)**
(Po-Vers. 2016s | TechFak | Informations- und Kommunikationstechnik (Master of Science) | Gesamtkonto | Schwerpunkte im Masterstudium | Schwerpunkt Multimediasysteme | Wahlpflichtmodule | Wahlpflichtmodul aus EEI im

Schwerpunkt Multimediasysteme | Speech and Audio Signal Processing)

- [25] **Mathematik (Bachelor of Science)**
(Po-Vers. 2015w | NatFak | Mathematik (Bachelor of Science) | Module des Nebenfachs | Nebenfach Informations- und Kommunikationstechnik | Wahlpflichtmodule aus Katalog für IuK | Speech and Audio Signal Processing)
- [26] **Mathematik (Bachelor of Science)**
(Po-Vers. 2019w | NatFak | Mathematik (Bachelor of Science) | weitere Module der Bachelorprüfung | Module des Nebenfachs | Nebenfach Informations- und Kommunikationstechnik | Wahlpflichtmodule aus Katalog für IuK | Speech and Audio Signal Processing)
- [27] **Mechatronik (Master of Science)**
(Po-Vers. 2012 | TechFak | Mechatronik (Master of Science) | Gesamtkonto | M3 Technische Wahlmodule | M3 Technische Wahlmodule | Speech and Audio Signal Processing)
- [28] **Mechatronik (Master of Science)**
(Po-Vers. 2020w | TechFak | Mechatronik (Master of Science) | Gesamtkonto | M3 Technische Wahlmodule | Speech and Audio Signal Processing)
- [29] **Medizintechnik (Master of Science)**
(Po-Vers. 2013 | TechFak | Medizintechnik (Master of Science) | Studienrichtung Medizinische Bild- und Datenverarbeitung | M2 Ingenieurwissenschaftliche Kernmodule (BDV) | Speech and Audio Signal Processing)
- [30] **Medizintechnik (Master of Science)**
(Po-Vers. 2013 | TechFak | Medizintechnik (Master of Science) | Studienrichtung Medizinelektronik | M2 Ingenieurwissenschaftliche Kernmodule (MEL) | Speech and Audio Signal Processing)
- [31] **Medizintechnik (Master of Science)**
(Po-Vers. 2018w | TechFak | Medizintechnik (Master of Science) | Studienrichtung Medizinische Bild- und Datenverarbeitung | M2 Ingenieurwissenschaftliche Kernmodule (BDV) | Speech and Audio Signal Processing)
- [32] **Medizintechnik (Master of Science)**
(Po-Vers. 2018w | TechFak | Medizintechnik (Master of Science) | Studienrichtung Medizinelektronik | M2 Ingenieurwissenschaftliche Kernmodule (MEL) | Speech and Audio Signal Processing)
- [33] **Medizintechnik (Master of Science)**
(Po-Vers. 2019w | TechFak | Medizintechnik (Master of Science) | Modulgruppen M1, M2, M3, M5, M7 nach Studienrichtungen | Studienrichtung Medizinische Bild- und Datenverarbeitung | M2 Ingenieurwissenschaftliche Kernmodule (BDV) | Speech and Audio Signal Processing)
- [34] **Medizintechnik (Master of Science)**
(Po-Vers. 2019w | TechFak | Medizintechnik (Master of Science) | Modulgruppen M1, M2, M3, M5, M7 nach Studienrichtungen | Studienrichtung Medizinelektronik | M2 Ingenieurwissenschaftliche Kernmodule (MEL) | Speech and Audio Signal Processing)
- [35] **Medizintechnik (Master of Science)**
(Po-Vers. 2019w | TechFak | Medizintechnik (Master of Science) | Modulgruppen M1, M2, M3, M5, M7 nach Studienrichtungen | Study Field Health and Medical Data Analytics | M2 Engineering Core Modules (HMDA) | Speech and Audio Signal Processing)
- [36] **Wirtschaftsingenieurwesen (Master of Science): 1-2. Semester**
(Po-Vers. 2009 | TechFak | Wirtschaftsingenieurwesen (Master of Science) | Masterstudiengang Wirtschaftsingenieurwesen (bis 30.09.2018) | Gesamtkonto | Ingenieurwissenschaftliche Studienrichtungen | Technische Wahlmodule | Technische Wahlmodule | Speech and Audio Signal Processing)
- [37] **Wirtschaftsingenieurwesen (Master of Science): 1-2. Semester**
(Po-Vers. 2009 | TechFak | Wirtschaftsingenieurwesen (Master of Science) | Masterstudiengang Wirtschaftsingenieurwesen (bis 30.09.2018) | Gesamtkonto | Ingenieurwissenschaftliche Studienrichtungen | Studienrichtung Informations- und Kommunikationssysteme | Wahlpflicht- und Vertiefungsmodul Modulgruppe 2 | Vertiefungsmodul Modulgruppe 2 | Speech and Audio Signal Processing)
- [38] **Wirtschaftsingenieurwesen (Master of Science)**
(Po-Vers. 2018w | TechFak | Wirtschaftsingenieurwesen (Master of Science) | Masterstudiengang Wirtschaftsingenieurwesen (Studienbeginn ab 01.10.2018) | Gesamtkonto | Studienrichtung Maschinenbau | Technische Wahlmodule und Hochschulpraktikum | Technische Wahlmodule | Speech and Audio Signal Processing)
- [39] **Wirtschaftsingenieurwesen (Master of Science)**
(Po-Vers. 2018w | TechFak | Wirtschaftsingenieurwesen (Master of Science) | Masterstudiengang Wirtschaftsingenieurwesen (Studienbeginn ab 01.10.2018) | Gesamtkonto | Studienrichtung Elektrotechnik | Schwerpunkt Informa-

tionstechnik | Wahlpflichtmodul 2.2 | Vertiefungsmodul 2.2 | Speech and Audio Signal Processing)

[40] **Wirtschaftsingenieurwesen (Master of Science)**

(Po-Vers. 2018w | TechFak | Wirtschaftsingenieurwesen (Master of Science) | Masterstudiengang Wirtschaftsingenieurwesen (Studienbeginn ab 01.10.2018) | Gesamtkonto | Studienrichtung Elektrotechnik | Technische Wahlmodule und Hochschulpraktikum | Technische Wahlmodule | Speech and Audio Signal Processing)

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%

Erstablingung: SS 2020, 1. Wdh.: WS 2020/2021

1. Prüfer: Walter Kellermann
