

课程详述

COURSE SPECIFICATION

以下课程信息可能根据实际授课需要或在课程检讨之后产生变动。如对课程有任何疑问，请联系授课教师。

The course information as follows may be subject to change, either during the session because of unforeseen circumstances, or following review of the course at the end of the session. Queries about the course should be directed to the course instructor.

1.	课程名称 Course Title	语音信号处理 Speech Signal Processing
2.	授课院系 Originating Department	电子与电气工程系 Department of Electrical and Electronic Engineering
3.	课程编号 Course Code	EE328
4.	课程学分 Credit Value	3
5.	课程类别 Course Type	专业基础课 Major Foundational Courses
6.	授课学期 Semester	春季 Spring
7.	授课语言 Teaching Language	中英双语 English & Chinese (English with Occasional Explanations in Chinese)
8.	授课教师、所属学系、联系方式 (如属团队授课, 请列明其他授课教师) Instructor(s), Affiliation & Contact (For team teaching, please list all instructors)	陈霏副教授, 电子与电气工程系 Associate Professor CHEN Fei, Department of Electrical and Electronic Engineering 南山智园 A7 栋 1104 室 Rm 1104, Building A7, iPark Email: fchen@sustc.edu.cn Tel: 0755-88018554
9.	实验员/助教、所属学系、联系方式 Tutor/TA(s), Contact	王蕾, 助教, 电子与电气工程系 WANG Lei, TA, Department of Electrical and Electronic Engineering 南山智园 A7 栋 1111 室 Rm 1111, Building A7, iPark Email: wangl6@mail.sustc.edu.cn Tel: 0755-88018582
10.	选课人数限额(可不填) Maximum Enrolment (Optional)	30

11. 授课方式 Delivery Method	讲授 Lectures	习题/辅导/讨论 Tutorials	实验/实习 Lab/Practical	其它(请具体注明) Other (Please specify)	总学时 Total
学时数 Credit Hours	32		32		64

12. 先修课程、其它学习要求 Pre-requisites or Other Academic Requirements	EE 323 数字信号处理 Digital Signal Processing
13. 后续课程、其它学习规划 Courses for which this course is a pre-requisite	
14. 其它要求修读本课程的学系 Cross-listing Dept.	

教学大纲及教学日历 SYLLABUS

15. 教学目标 Course Objectives

本课程涵盖语音信号处理的基本内容，介绍以下 5 大方面的知识：

1. 语音信号产生的过程，听觉的过程，语音链的构成
2. 语音信号处理的基本方法，短时能量、幅度、过零率分析，短时傅里叶分析，同态分析方法，和线性预测分析
3. 语音参数估计常用的方法和算法，包括语音区间的检测，声调估计，谐波共振峰
4. 语音信号处理的基本应用原理，包括语音编码，语音合成，语音识别
5. 语音信号处理相关的重要工具程序和 Matlab 程序

The subject covers the fundamentals of speech signal processing. The main objective is to gain knowledge/skill of five main topics of speech signal processing:

1. The process of speech production, the process of hearing, and speech chain
2. The fundamentals of speech signal processing, including short-time energy, magnitude and zero-crossing analysis, short-time Fourier analysis, homomorphic methods, and linear predicative methods
3. The commonly used methods and algorithms for detecting speech/non-speech segments, voiced/unvoiced/non-speech, pitch and formants
4. The principles of the applications of speech signal processing in speech coding, speech synthesis, and speech recognition

The skills to utilize open-source toolkits and Matlab functionality for speech processing

16. 预达学习成果 Learning Outcomes

通过这门课程的学习，学生能够：

1. 掌握语音和听觉产生的基本过程，综合声学、生理学、语言学、信号处理等知识来描述语音信号的特征
2. 掌握语音信号处理的常用方法，了解各个短时分析方法的原理和运用场合
3. 掌握提取语音区间、声调、谐波共振峰的技术和算法
4. 了解倒谱的概念和算法，运用倒谱技术提取语音信号的重要参数
5. 了解语音技术应用的历史、常用方法和未来发展，熟悉语音编码、合成和识别技术的基本原理
6. 运用 Matlab 软件工具对语音信号进行分析和特征提取，使用当前常用的语音分析软件

After completing this course, the students will be able to

1. Understand the processes of speech production and speech perception, and make use of acoustic, physiological, linguistic, and signal processing knowledge to characterize speech signals
2. Know the commonly-used techniques and methods for speech signal processing, and understand the principles and applications of short-term speech analysis methods
3. Understand the fundamentals of techniques and algorithms for detecting speech/non-speech segments, voiced/unvoiced/non-speech, pitch, formant, etc.
4. Understand the concept and implementation of cepstrum, and use cepstrum to extract important speech features
5. Know the history of speech technology application, present technologies and future development trend, and understand the principle of speech coding, synthesis and recognition
6. Analyze speech signals with Matlab functionality and commonly-used open-source toolkits

17. 课程内容及教学日历（如授课语言以英文为主，则课程内容介绍可以用英文；如团队教学或模块教学，教学日历须注明主讲人）

Course Contents (in Parts/Chapters/Sections/Weeks. Please notify name of instructor for course section(s), if this is a team teaching or module course.)

第一周：语音信号处理介绍：语音信号简介，课程安排介绍，离散时间信号与系统，常用的变换方法、数字滤波器设计等的复习

第二周：语音产生的基础知识：语音产生的过程和生理基础，语音产生的声学 and 语音学知识，音素发音的特征

第三周：听觉及其模型，语音识别：语音链，耳的解剖结构和功能，声音的感知，听觉模型，语音识别实验设计，测量语音音质和清晰度

第四周：语音处理的时域方法：语音信号的短时分析，短时能量、幅度和过零率，短时自相关函数

第五周：频域表示：离散时间傅里叶分析，短时傅里叶分析，重叠相加合成法，滤波器相加合成法

第六~七周：倒谱和同态语音处理：卷积的同态系统，语音模型的同态分析，语音信号短时倒谱和复倒谱的计算，自然语音的同态滤波，零极点模型的倒谱分析，倒谱距离测量

第八~九周：语音信号的线性预测分析：线性预测分析的基本原理，模型增益的计算，线性预测分析的频域表述，求解线性预测分析方程

第十周：估计语音参数的算法：语音信号中值平滑处理，语音背景/静默期区分，声调周期估计，谐波共振峰估计

第十一~十二周：语音信号的数字编码：语音信号的统计模型，幅度的瞬时自适应量化，Delta 调制，差分 PCM，声码器的分析与合成

第十三~十四周：语音和音频信号的频域编码：子带编码，自适应变换编码，MPEG-1 音频编码标准，其它音频编码标准

第十五周：文本语音合成的方法：文本分析，语音合成方法的进展，语音合成方法，单位选择方法，文本语音合成的未来

发展

第十六周：自动语音识别和自然语言理解：自动语音识别的基本构成，语音识别过程，自动语音识别中的检测过程，口语理解

Week 1: Introduction to Speech Signal Processing

The Speech Signal, Applications of Digital Speech Processing, Discrete-Time Signals and Systems, Transform Representation of Signals and Systems, Fundamentals of Digital Filters

Week 2: Fundamentals of Human Speech Production

The Process of Speech Production, Acoustic Phonetics

Week 3: Hearing, Auditory Models, and Speech Perception

The Speech Chain, Anatomy and Function of the Ear, The Perception of Sound, Auditory Models, Human Speech Perception Experiments, Measurement of Speech Quality and Intelligibility

Week 4: Time-Domain Methods for Speech Processing

Short-Time Analysis of Speech, Short-Time Energy and Short-Time Magnitude, Short-Time Zero-Crossing Rate, The Short-Time Autocorrelation Function

Week 5: Frequency-Domain Representations

Discrete-Time Fourier Analysis, Short-Time Fourier Analysis, Overlap Addition Method of Synthesis, Filter Band Summation Method of Synthesis

Weeks 6-7: The Cepstrum and Homomorphic Speech Processing

Homomorphic Systems for Convolution, Homomorphic Analysis of the Speech Model, Computing the Short-Time Cepstrum and Complex Cepstrum of Speech, Homomorphic Filtering of Natural Speech, Cepstrum Analysis of All-Pole Models, Cepstrum Distance Measures

Weeks 8-9: Linear Prediction Analysis of Speech Signals

Basic Principles of Linear Predictive Analysis, Computation of the Gain for the Model, Frequency Domain Interpretation of Linear Predictive Analysis, Solution of the LPC Equations

Week 10: Algorithms for Estimating Speech Parameters

Median Smoothing and Speech Processing, Speech-Background/Silence Discrimination, Pitch Period Estimation, Formant Estimation

Weeks 11-12: Digital Coding of Speech Signals

A Statistical Model for Speech, Instantaneous and Adaptive Quantization, Delta Modulation, Differential PCM, Analysis-by-Synthesis Speech Coders

Weeks 13-14: Frequency-Domain Coding of Speech and Audio

Sub-band Coding, Adaptive Transform Coding, MPEG-1 Audio Coding Standard, Other Audio Coding Standards

Week 15: Text-to-Speech Synthesis Methods

Text Analysis, Evolution of Speech Synthesis Methods, Early Speech Synthesis Approaches, Unit Selection Methods, TTS Future Needs

Week 16: Automatic Speech Recognition and Natural Language Understanding

Basic ASR Formulation, Overall Speech Recognition Process, The Detection Processes in ASR, Spoken Language Understanding

18. 教材及其它参考资料 **Textbook and Supplementary Readings**

Lawrence R. Rabiner, and Ronald W. Schafer, <Theory and Applications of Digital Speech Processing>, Pearson, 2011.

课程评估 **ASSESSMENT**

19. 评估形式 Type of Assessment	评估时间 Time	占考试总成绩百分比 % of final score	违纪处罚 Penalty	备注 Notes
出勤 Attendance		10		
课堂表现 Class Performance				
小测验 Quiz				
课程项目 Projects		20		
平时作业 Assignments		50		
期中考试				

Mid-Term Test
期末考试
Final Exam
期末报告
Final Presentation
其它（可根据需要
改写以上评估方
式）
**Others (The
above may be
modified as
necessary)**

	20		

20. 记分方式 **GRADING SYSTEM**

- A. 十三级等级制 **Letter Grading**
 B. 二级记分制（通过/不通过） **Pass/Fail Grading**

课程审批 REVIEW AND APPROVAL

21. 本课程设置已经过以下责任人/委员会审议通过
This Course has been approved by the following person or committee of authority

