

# 音声情報処理特論

2

Speech Processing, Adv.

【配当学年】 後期

【担当者】 奥乃・河原・尾形・北村・平原

## 【内 容】

本講義では、音について基礎理論から、最先端の音の認識技術とその応用について学ぶ。

第1部では、音の物理モデルについて述べ、生理学および心理物理学の立場から人の聴覚機能について学び、聴覚による情景分析・理解についての理解を深める。次に、人の発声機能について述べ、ソース・フィルタ理論や声道内の共鳴現象について解説する。

第2部では、音声の基本的な性質と特徴量について述べ、代表的な音声分析手法を解説する。音声符号化、音声合成、及び音声認識についてその概要を述べる。音声認識においては、代表的な音響モデル・言語モデルについて解説する。

第3部では、音声に限らず楽音や環境音を含む音環境理解について述べ、両耳聴の原理から空間的な知覚の原理を探り、視覚と聴覚の情報統合による認知機能について説明する。また、具体的な応用としてロボット聴覚の課題と現状についても取り上げる。

## 【授業計画】

項 目	回 数	内 容 説 明
聴覚 (平原)	2	・ 音の物理モデル ・ 人の聴覚機能 (聴覚生理学・心理物理学)
音声生成理論 (北村)	1	・ ソース・フィルタ理論 ・ 声道内の共鳴現象
音声の分析・認識・合成 (河原)	4	・ 音声に含まれる情報、線形分離等価モデル ・ 音声分析 ・ 音声合成・音声符号化・音声認識の概要 ・ 音声認識のための音響モデル・言語モデル
音環境理解とロボット聴覚 (奥乃)	4	・ 音環境理解とは、音源定位 ・ 音源分離と分離音認識 ・ ロボット聴覚 ・ 音楽情報処理、擬音語認識
マルチモーダル情報処理とロボットによる情報生成 (尾形)	2	・ ダイナミクスに基づくマルチモーダル情報処理 ・ ロボットによるマルチモーダル情報生成

【教科書】 なし。

## 【参考書】

- ・ 板橋秀一編: 『音声工学』 (森北出版, 2005)
- ・ Moore, B.C.J.: “*An Introduction to the Psychology of Hearing*”, 4th Ed., Academic Press, 1997. (大串監訳 『聴覚心理学概論 (第3版)』 (誠信書房), 1994)
- ・ Bregman, A.: “*Auditory Scene Analysis*” (MIT Press, 1990)
- ・ Rosenthal, D. and Okuno, H.G. (eds.): “*Computational Auditory Scene Analysis*” (Lawrence Erlbaum Associates, 1998)

## 【その他】

- ・ 成績評価は、5項目のレポートの成績の平均に毎回の議論への参加度を加味して行う。
- ・ 講師の都合により、一部順序の変更がありうる。