

知能メディア講座 音声メディア分野

教授 奥乃 博 准教授 尾形 哲也

特定助教 糸山 克寿 GCOE助教 高橋 徹 特定助教 西出 俊

志望区分：知-7

概要

本研究室では、音声だけでなく、楽音や環境音などのさまざま音、さらにそれらの混合音に含まれる情報の知覚・理解を通じた音環境理解の研究を、知能情報学の立場から行っている。研究のキーワードは、「音を聞き分ける」、「能動的音楽鑑賞」、「認知発達ロボティクス」である。とくに、実環境でのロボット聴覚を実現するために、ノンパラメトリックベイズ推定によるマルチチャンネル音響信号処理、実時間組込みシステム、同時発話認識、擬音語認識の課題に取り組み、音楽共演ロボットや、神経力学による人とロボットとのインタラクション手法などに展開をしている。

研究テーマ

1. 音環境理解とロボット聴覚:

ロボットを実環境に配備し、人との共生を進めるには、必要な事前知識が極力少ないロボット聴覚システムの開発が不可欠である。これまでに同時発話を『聞き分ける』ために開発したロボット聴覚ソフトウェア HARK の高機能化のために、ノンパラメトリックベイズ手法による音源定位・音源分離・音源追跡、自己生成音抑制機能の高速化などに取り組んでいる。ここで、同時発話認識や擬音語認識が可能となった場合のロボットヒューマンインタラクションの設計が重要な課題である。

2. 能動的音楽鑑賞・音楽共演ロボット:

膨大なデジタル音楽を自由に加工し、音楽の新しい楽しみ方を支援するために、多重奏音楽音響信号から最小限の事前知識だけで楽器音を分離し、残響抑制を行う機能をノンパラメトリックベイズ手法により取り組んでいる。また、ハードウェアに依存しない電子楽器テルミン演奏ロボットの開発を行い、人のリードによる音楽共演ロボットの開発にも取り組んでいる。ここで、人の演奏の揺らぎに対応するためには、ノンパラメトリックベイズ手法による楽譜追跡が重要な課題である。

3. ロボットのマルチモーダルインタラクションによるシンボル発達:

認知発達ロボティクスの観点から、声道モデルと人工神経回路モデルを有したロボットを用い、幼児の音声模倣、音素獲得過程を、脳科学知見を規範としつつモデル化する研究を行っている。また音声やジェスチャなどの原始記号の創発、人間型ロボットを用いた物体能動知覚、道具身体化などに関して、神経力学系を規範とした構成論的研究を展開している。

分野基礎問題出題範囲

上記のような研究を行うに際しては、人工知能、コミュニケーションモデル、パターン認識と機械学習、聴覚心理学、音響・デジタル信号処理、認知発達ロボティクス、発達認知神経科学等（すべてを要求しない）に関する基礎的な素養とともに、人間が音を知覚し、理解する過程に対する深い洞察と旺盛な好奇心が望まれる。具体的には以下の書籍から出題する。

1. 日本音響学会編 『音のなんでも小事典』（講談社ブルーバック）、
2. 長尾他著 『文字と音の情報処理』（岩波講座マルチメディア情報学第4巻）の音に関する章。
3. 【参考】 Al Bregman: “Auditory Scene Analysis” (1991, MIT Press) の第1章。

問合せ先

京都大学工学部 10号館 335号室 奥乃 博 (tel: 075-753-5376)

電子メール: okuno@i.kyoto-u.ac.jp

研究室ホームページ: <http://winnie.kuis.kyoto-u.ac.jp/>

情報学研究科 知能情報学専攻 音声メディア分野 奥乃・尾形研究室

メンバ

- ・ 教授 奥乃 博 (東大→NTT→JST→東理大→京大)
- ・ 准教授 尾形 哲也 (早大→理研→京大)
- ・ 特定助教 糸山 克寿 (京大)
- ・ GCOE助教 高橋 徹 (名工大→和大→京大)
- ・ 特定助教 西出 俊 (阪大・京大)
- ・ 秘書 岡崎 裕美
- ・ 博士課程 4名 (学振特別研究員2名, 留学生2名)
- ・ 修士課程 12名 (留学生2名, 京大他学部2名・他大学1名(英))
- ・ 卒研究生 5名程度
- ・ ロボット Actoroid, HIRO, HRP-2, Robovie-R2, NAO, Robovie-lls, Robovie-X, SIG-2, Keepon, Palo, ...

主要研究テーマのキーワード

- 聞き分ける・音を見せる**
 - ・ ロボット聴覚オープンソースソフトウェア **HARK**
 - ・ Non-parametric Bayesian 音源定位・分離・追跡
- 音楽共演ロボット・能動的音楽鑑賞**
 - ・ 視聴覚情報統合による音楽演奏共演モデル
 - ・ 調波非調波統合モデル上のBayes音源分離
- 認知発達ロボティクス**
 - ・ インタラクション能力の神経力学モデル
 - ・ 経験学習による身体自己モデルの獲得



参加プロジェクト

- ・ 科研費(S): **音環境理解からのロボット聴覚の構築**
- ・ JST日仏交流: **ヒューマノイドのための能動的両耳聴**
- ・ JSTさきがけ: **長期インタラクション創発を可能とする知能化空間の設計論**

発表論文・受賞

- ・ 重要国際会議を中心に発表(平均年1回以上)
- ・ 特許出願・ソフトウェアライセンス
- ・ 数多くの受賞・海外発表助成
- ・ **ほぼ全員が奨学金返済免除・学振特別研究員**
- ・ IEEE/RSJ IROS NTF Award for Entertainment Robots

修士論文の一部

- ・ 共演音楽ロボットのためのPFによる実時間楽譜追従
- ・ 神経力学モデルを用いた言語と感覚運動系の統合
- ・ 発語行為情報による発話意図解釈による音声対話
- ・ NP-Bayes 音源分離・残響除去による楽器音置換
- ・ 楽譜を援用したクラシック音楽解釈の推論
- ・ 楽器音イコライザによる類似楽曲検索

就職先: 博士進学, 京大, 名大, 産総研, 名工大, 日大, 任天堂, キヤノン, パナソニック, ソニー, 日立, ホンダ, NTT, DoCoMo, 東芝, NEC, デンソー, 朝日放送, 毎日放送等

URL: <http://winnie.kuis.kyoto-u.ac.jp/>

音環境理解 (CASA)

奥乃・尾形研究室

音声メディア分野



あかちゃんの挙動模倣

音による Human Robot Interaction

能動的音楽鑑賞

認知発達
ロボティクス

常時発話可能対話

音楽共演ロボット

音高, 音長,
音色操作

テンポによる
音楽解釈理解

バフリング
による模倣

多様な音を聞き分ける

自分の音は抑制

音を見せる

身体性

ロボット聴覚
ソフト**HARK**

HARK
on Kinect



SpeakBySinging

道具身体化

能動的両耳聴

複数話者同時
発話認識

実時間視聴
覚情報統合

音楽情報処理

楽譜追跡

神経力学

移動ロボット
AV-SLAM

擬音語認識

音源定位・分離

調波非調波統合モデル

Recurrent
Neural
Network

マルチチャンネル
音響信号処理

Non-Parametric
ベイズ推定

音によるField 観測

カエルホタル

発達認知神経科学

人工知能 (AI)

計算機科学 (CS)

合唱の数理モデル

