

# 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 助教 駒谷 和範 助教 高橋 徹

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な音、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、「聞き分ける」、「音響信号のシンボルへのグラウンディング」、「身体性を利用した他メディアとの統合」である。とくに、ロボットを対象とする組み込みシステムへの応用を視野に入れ、実環境下で複数のセンサー入力を階層的に統合化することにより、実時間で処理可能な技術の研究開発を行っている。

## 研究テーマ

### 1. 音環境理解とロボット聴覚:

混合音から各々の音響ストリームを『聞き分ける』ために、ビームフォーマや独立成分解析による音源分離とミッシングフィーチャマスク自動生成に基づいた音声認識に取り組んでいる。ここでは、音源分離と分離音認識の統合や画像情報との統合による高性能化も重要な課題である。さらに、開発した聴覚システムをヒューマノイドロボットに搭載し、評価するとともに、知覚と拳動との相互作用を介したロボットヒューマンインタラクションの研究も行っている。

### 2. 楽器音分離による音楽情報処理・環境音の擬音語認識:

膨大なデジタル音楽や音データに対する検索を効率化するために、楽器音同定、歌手認識、ドラム音認識などの音楽情報認識の研究、および、環境音の擬音語自動認識の研究を行っている。ドラム音イコライザ、パート抽出による楽器カラオケ、楽器構成のリミックスによる類似楽検索などの直接的な応用だけでなく、音楽に合わせて踊り歌うロボット、楽曲推薦、など多様な分野に応用している。

### 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、力学系を規範とした構成論的研究を展開している。

### 4. 音声を用いた対話インタフェース:

音声認識誤りに頑健な音声対話の実現に向けて、バージン（発話への割り込み）処理から発話行為理論に基づく対話文脈のモデル化まで、階層的に取り組んでいる。具体的な課題は、ユーザモデリングに基づく適応的音声インタラクション、初心者への動的ヘルプ生成、実環境ロボットとの音声対話、音声対話システムのラピッドプロトタイピング、マルチドメイン音声対話システムなどである。

## 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、聴覚心理学、音響・デジタル信号処理、脳科学等（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/>