京都大学大学院情報学研究科 知能情報学専攻

音声メディア研究室

教授　河原 達也　　准教授　吉井 和佳　　助教　井上 昂治

志望区分：知-9

概要

人間の知の創造・伝達の多くは、音声によるコミュニケーションによって行われています。本分野では、人間どうしがやりとりを行う音声メディアを分析し、自動認識・理解した上で、インタラクションを行えるシステムの実現をめざします。具体的には、会議や講演のような自然な話し言葉音声をテキスト化するシステム、複数の話者や音源などからなる音環境や音楽を解析するシステム、非言語情報も統合しながら自然な対話が行える人間型ロボットなどの研究に取り組みます。

研究テーマ

1. 話し言葉の音声認識

講演・講義や会議のような実世界の話し言葉音声を自動認識し、情報・構造を抽出し、さらに講演録・会議録や字幕・要約などを生成する方法について研究しています。

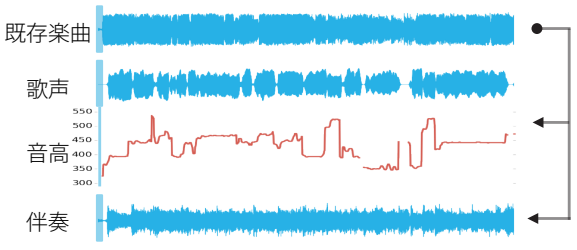
音声認識による講演への字幕付与

2. 音環境理解

複数の話者が存在したり、音声以外に様々な音が存在している環境の認識・理解を、マルチモーダルなセンシングと統計的な音響信号処理に基づいて行っています。

3. 音楽情報処理

音楽音響信号に対して音源分離や自動採譜を行う方法を研究しています。これに基づいて、歌声と伴奏音を分離して加工・編集を行うシステムを開発しています。



モノラル音楽信号に対する歌声と伴奏の分離

4. ロボットとの音声対話・インタラクション

人間のように振る舞い、インタラクションを行えるロボット（アンドロイド）の実現をめざして、言語情報と非言語情報を統合した音声対話のモデル・システムを研究しています。

望まれる素養

このような研究を行うに際しては、パターン認識、機械学習、統計的モデリング、デジタル信号処理、情報理論、人工知能、ヒューマンインタフェースに関する知識と音声情報処理に対する基礎的素養が望まれます。研究室Webサイト上の「研究プロジェクト」のページにPDFファイルをおいてある文献も参考にしてください。

問合せ先
総合研究7号館4階
河原達也(kawahara@i.kyoto-u.ac.jp)

研究室Webサイト
http://sap.ist.i.kyoto-u.ac.jp/