

知能メディア講座 音声メディア分野

教授 河原 達也 准教授 井本 桂右

助教 井上 昂治 特定助教 越智 景子

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、講演や会話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・理解

最先端の深層学習モデルに基づいて、講演・講義や会議・会話のような実世界の話し言葉音声を自動認識するとともに、話者や感情などの情報を抽出し、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。アイヌ語のような少数言語も対象としています。

2. 音環境解析

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、マルチチャネル・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. ロボットとの音声対話

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、言語情報だけでなく、相槌・笑い・感情などの非言語情報も考慮した音声対話のモデル・システムを研究しています。複数人との会話に対象を広げています。

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、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ファイルが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

E-mail: kawahara@i.kyoto-u.ac.jp

研究室ホームページ: <http://sap.ist.i.kyoto-u.ac.jp>