

新しい社会システムデザインに向けた情報基盤技術の創出
2018 年度採択研究者

2020 年度
年次報告書

武田 龍

大阪大学 産業科学研究所
准教授

音声対話系における言語・音響モデル自動適応

§ 1. 研究成果の概要

本研究では、音声対話を通じて、音声認識における言語モデル(単語)や音響モデル(発声特徴)の自動適応を目指す。システムがユーザに適応することで、次第に使いやすくなる音声対話システムの構築を目指している。本年度の主な研究成果は以下の 3 点である: ①音声認識中における音響モデルの教師なし適応手法の成果発表、②音声信号からのユーザ年齢推定、③音声対話システムの基盤構築。

①音響モデルの教師なし適応では、昨年度の研究成果を査読付き国際会議で発表した。ユーザ発話の認識と音響モデル教師なし適応をフレーム単位で交互に行う。これにより、ユーザ音声を聞くだけで、リアルタイムにユーザの音声へ適応できる。

②音声信号からのユーザ年齢推定では、年齢分布や話者数の異なる複数のデータセットからの年齢予測モデルの構築に取り組んだ。各データセットでの音響特性や分布の異なりが原因で、オープンな音声信号に対する年齢予測に失敗する。特徴量正規化に加え、ベイズモデルに基づきラベル・声年齢の曖昧性を表す確率モデルと深層学習を統合することで、予測性能が向上することを明らかにした。予測されたユーザ年齢は、システム発話や説明の表現をユーザへ適応することに用いる。

③音声対話システムの基盤構築では、実システム実装に必要な各種作業に取り組んだ。まず、再利用可能な対話システムモジュールの設計・実装を行った。同モジュールを用いて、対話ロボットコンペに参加し、その動作検証も進めている。次に、複数コーパスを用いて音響モデル学習を行い、可能な限りベースとなるモデルの性能向上に努めた。最後に、未知語・聞き誤りに対するユーザ応答のデータ収集・分析を行った。ユーザがどのような教え方をするかを分析するため、未知語の属性や認識誤りを想定した対話データを収集した。今後、このデータはシステム挙動の設計・ユーザ応答のモデル化に利用する。

【代表的な原著論文情報】(最大 5 件)

- 1) Ryu Takeda and Kazunori Komatani: "Frame-wise Online Unsupervised Adaptation of DNN-HMM Acoustic Model from Perspective of Robust Adaptive Filtering," Proceedings of Interspeech, pp.1291-1295, 2020.
- 2) Ryu Takeda and Kazunori Komatani: "Age Estimation with Speech-age Model for Heterogeneous Speech Datasets", Proceedings of Interspeech, pp.4164-4168, 2021.