

数理・情報のフロンティア
2021 年度採択研究者

2021 年度 年次報告書

井上 昂治

京都大学 大学院情報学研究科
助教

マルチモーダルなふるまいに基づく音声対話の人間目標型評価

§ 1. 研究成果の概要

本研究課題では、音声対話システムの「人間らしさ」について、ユーザのマルチモーダルなふるまい(反応)から間接的に評価をする方法を確立することを目指している。そのために、人間どうしの対話とシステムとの対話において、ユーザのふるまいの違いを分析し、それをもとに評価モデルを学習する。このように人間どうしの対話を基準とした評価を「人間目標型評価」として提案している。

2021年度は、本研究のベースとなる対話データの収集およびアノテーション作業を実施した。はじめに、複数のマイクおよびカメラを同期させながら対話データを収録するシステムを構築した。このシステムを用いて傾聴対話70対話を収録した。このうち20対話はシステムによる対話、50対話は人間が背後で話した対話(Wizard-of-OZ)である。傾聴のシステムはこれまでに構築してきたものを活用した。システムのインターフェイスはアンドロイドロボットを用いた。

続いて、収録した対話データに対して、人間らしさのラベルのアノテーションを行うためのツールを作成し、試験的なアノテーションを実施した。上記の対話データのうち、ユーザ役のみが映された映像と、ユーザ役の音声のみを強調(システムの音声を抑圧)した音声動画像を作成し、これを2分間毎に区切り評価サンプルとした。各サンプルをアノテータに視聴してもらい、その対話相手(システム)が人間かシステムのどちらかと思うかを、二値で判定してもらった。次年度はこの二値判定の結果とユーザのふるまいとの関係性を調査し、さらに大規模な対話データ収録およびアノテーションを実施していく予定である。また、上記と並行して、ユーザのふるまいのアノテーションも実施した。発話区間、発話内容、相槌、笑い、フィラー、視線についてのデータセットを構築した。