

# マルチモーダル対話システムのための話者間の親しさの識別

藤井 更紗<sup>†</sup> 山崎 善啓<sup>††</sup>  
<sup>†</sup> 東北大学工学部電気情報物理工学科

千葉 祐弥<sup>††</sup> 伊藤 彰則<sup>††</sup>  
<sup>††</sup> 東北大学大学院工学研究科

## 1. はじめに

近年、雑談対話システムの研究が盛んに行われている。雑談対話システムにおいては、ユーザの主観評価が重要な指標の一つとなる。そのため、ユーザ評価の向上を目標として様々な検討が行われてきた [1]。特に、文献[2]ではユーザとシステムの関係性を考慮することでユーザの評価を改善できることが示されている [2]。しかしながら、対話システムが対話相手との親しさに基づいてふるまいを変えるためには、対話中の情報から対話相手との親しさを推定できる必要がある。本研究ではユーザとの親しさの推定手法の確立を目指して、まずは人間同士の対話を用いて話者間の親しさを推定する手法を検討する。

## 2. 親しさを考慮したマルチモーダル雑談コーパス

実験データには、文献 [3]で収集されたマルチモーダル雑談コーパスを用いる。当該のコーパスには、大学生・大学院生 19 名 (男性 15 名, 女性 4 名)の計 95 対話 (約 10 時間)のデータが収録されている。対話はビデオカメラによってそれぞれの話者の正面から撮影されている。全 19 ペアのうち 8 ペアは初対面, 11 ペアは面識のある話者同士の会話である。本研究では前者を親しさの度合いが低いグループ, 後者を高いグループとして扱う。各グループの総発話数は 7,219 発話及び 12,087 発話である。それぞれの発話には人手で付与された対話行為のラベルと, Convolutional Neural Network (CNN)によって自動で付与された笑顔またはその他の 2 値の表情ラベルが付与されている。

## 3. マルチモーダル情報に基づく親しさの識別

本稿では、各発話における話者間の親しさを高低の 2 値で分類する識別モデルを学習する。識別器には隠れ層 3 層の Multilayer Perceptron (MLP)を利用した。文献 [3]では、対話行為ラベルの遷移, 音声パワー及び対数 F0 の同調傾向, 笑顔の出現頻度が話者間の親しさによって異なることが確かめられている。そこで、本稿においてもこれらの特徴量を識別に用いる。表 1 に実験で用いた音声・画像・言語特徴量をまとめる。識別器には、当該発話の特徴量と先行発話の特徴量を結合して入力した。ここで、音声パワーと対数 F0 に関しては、当該発話の先頭及び先行発話の末尾 500 [ms]区間の平均と最大値を利用した。

## 4. 識別実験の条件

実験では、特徴量の組み合わせを変えることでそれぞれの特徴量の効果を比較した。実験は 4 分割交差検証に

よって行った。この時、各フォールドは話者がオープンになるように分割された。また、学習時にはオーバーサンプリングによって各クラスのサンプル数を均等にした。MLP の隠れ層のノード数は事前の実験により 32 とした。

表 1 推定に用いた特徴量

音声特徴量(A)	交替潜時, パワー, 対数 F0
画像特徴量(V)	表情ラベル
言語特徴量(L)	対話行為タグ

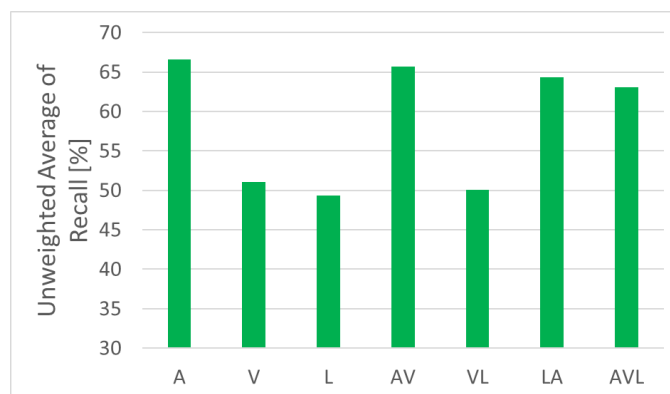


図1. 各特徴量における識別精度

## 5. 実験結果

図 1 にそれぞれの特徴量を用いた場合の各クラスの再現率の平均を示す。結果より、音声特徴量を用いた場合がその他の特徴量を用いた場合よりも性能が高く、66.6%の識別精度が得られた。しかしながら、音声・画像・言語特徴量を組み合わせることで性能は低下した。したがって、今回の条件では音声特徴量のみを用いることで話者間の親しさがある程度分類できることが示唆される。

しかしながら、性能自体はまだ十分とは言えない。したがって今後は音声・画像・言語特徴量の統合方法に関して検討を行う予定である。

## 6. まとめ

本研究では、対話中の話者間の親しさの度合いを識別する手法を検討した。実験では、音声特徴量を用いた場合が最も性能が高く、66.6%の精度が得られた。今後はより効果的に特徴量を統合するため、識別器のネットワークの構造に関しても検討を行う予定である。

## 参考文献

- [1] Kobori et al., <https://doi.org/10.18653/v1/W16-3646>, 2016
- [2] Kageyama et al., <https://doi.org/10.18653/v1/W18-5026>, 2018
- [3] 山崎他, <http://id.nii.ac.jp/1004/00010461/>, 2019