

## 多人数動作解析に基づく相互関係の理解

李 春軒<sup>†</sup> 福田 悠人<sup>†</sup> 久野 義徳<sup>†</sup> 小林 貴訓<sup>†</sup>  
<sup>†</sup>埼玉大学

### 1. はじめに

少子高齢化による労働者不足を背景に、人の代わりにサービスを提供できるサービスロボットに期待が寄せられている。このようなロボットを実現するためには、ロボットがサービスを求めているユーザを特定できなければならない。この問題に対して、本研究では、カメラ画像を用いた行動認識により解決を試みる。従来の行動認識手法は、比較的単純な環境を想定した手法が多く、複数人環境では適用が難しい。そこで、「手を上げている」などの特定の行動認識ではなく、行動の種類を識別しないで、周りとは異なる行動を取っている人を認識して仮定ユーザと見なし、ある程度近づいてからジェスチャー認識などを用いてユーザを特定することにした。向かい合っている人の間には、ある程度の同期現象が発生することが知られている[1]。特に、会話するときには、自発的なジェスチャーの同期がある[2]。よって、本稿では、人の相互関係を理解し、会話など相互行為をしている人とそうでない者（仮定ユーザ）の分類をめざす。具体的には、会話している二人とそうでない人を同じ場所に座らせ、人数分のビデオカメラで別々に動画を取り、得られたデータから、相互の同期性を分析する。他の被験者との同期性が一番低い人が会話に参加していない仮定ユーザである。

### 2. 提案手法

まず、動画から Openpose を用いて人の骨格情報をそれぞれ取得・保存する。次に、関節の点を数フレーム前と比較し、その移動量や相対移動量を計算する。最後に、区間を分けて移動量或は移動量の内積を特徴量として一組二人のデータの動的タイムワーピング(DTW)距離を計算する。移動量の場合、距離が小さい区間数が多い二人が会話者で、もう一人が仮定ユーザである。内積の場合は、距離が大きい区間が多いほど二人が相互行為をしている可能性が高い。

### 3. 実験

図1で示したように、三人に同じテーブルに座ってもらい、テーブルに三人それぞれに向けたビデオカメラを三台設置し、一組1分30秒撮影した。三人のう

ち二人には話題を与えて会話してもらい、もう一人はスマートフォンの操作してもらった。また、スマートフォンを持つ被験者には遠くにいる人とスマートフォンを通じて会話してもらった。また、個人差を考慮して、被験者を変えて、複数回実験を行った。動的タイムワーピング距離の計算に使う実験データは30fpsの動画で、計算区間は300フレーム(10s)とした。



図1. 実験の様子

移動量を特徴量とした場合、三人のDTW距離が小さい区間数に差は見られなかった。内積の場合、会話者のDTW距離が大きい区間が80%以上を占めていた。よって、内積を特徴量として計算することで、期待する結果が得られた。

### 4. おわりに

今回の実験では、移動量の内積で計算した動的タイムワーピング距離は期待通りの結果となったが、今回の実験では、仮定ユーザとなる被験者はあまり動いていないため、実際に会話していないにも関わらず、動きの多い二人の結果値が大きくなった可能性も考えられる。今後は、仮定ユーザ役に人が座っているだけでなく、何かの行動をしている場合でも正しく分類できるようにする予定である。また、特徴量を定め、機械学習などの手法での分類を試みるなど、より有効な手法を継続して検討していく予定である。

### 参考文献

- [1] Shuntaro Okazaka, et al, "Unintentional Interpersonal Synchronization Represented as a Reciprocal Visuo-Postural Feedback System: A Multivariate Autoregressive Modeling Approach," PLoS one, 10(9), 2015.
- [2] 城綾実, 細馬宏通, "多人数会話における自発的ジェスチャーの同期," 認知科学 16(1), pp.103-119, 2009.