

エージェント対話システムのための 高速カメラを用いた表情認識

秋山 俊貴[†][†] 埼玉大学小林 貴訓^{†,††}^{††} 科学技術振興機構さきがけ久野 義徳[†]

1. はじめに

現在、日本では認知症高齢者の数が増加し、認知症の進行抑止や介護支援の必要性が高まっている。認知症の予防や進行の抑制には、家族との会話など、認知活動を行うことが重要である[1]。しかし、会話には必ず相手が必要であり、いつでも行うことは難しい。そこで、我々の研究グループでは、認知症高齢者の表情を認識し、感情を推定することで行動を変化させるエージェント対話システムを検討している。しかし、認知症高齢者の表情は乏しい傾向があり、従来の表情認識手法で感情を推定することは難しい。そこで、Yan ら[2]などによって研究が進められている微表情を手がかりとすることを検討している。微表情は、感情を抑えようとしている時やあまり強くない感情の時に見られる表情であるが、表情の持続時間が 0.04 秒程度と非常に短いため、一般的なカメラで検出することは難しい。そこで、本稿では、高速カメラを用いた表情認識を検討する。

2. 高速カメラを用いた表情認識

高速カメラは 90fps で撮影が可能なものを利用する。撮影した 0.1 秒間のフレーム画像列を図 1 に示す。



図 1. 0.1 秒間のフレーム画像列

一般的なカメラを用いて 0.1 秒間のうちに生起する表情変化とらえることは難しい。そこで、本システムでは、平常時は 9 フレームごと(1 秒間に 10 回)に認識処理を行い、同時に過去 2 秒間(180 フレーム)の画像をバッファに保存しておく。本システムはエージェント対話システムに適用することを想定しているため、エージェントの発言に対する感情変化を捉えることを目的とし、エージェントが特定のキーワードを発言するタイミングでバッファに保存されている過去 2 秒間の全ての画像に対して認識処理を行い、細かな感情変化の観測を試みる(図 2)。

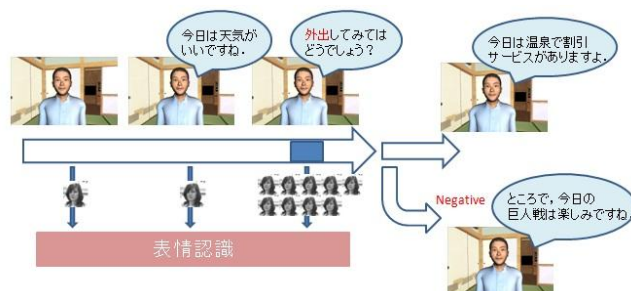


図 2. 表情認識を行うタイミング

表情認識の手順としては、まずアピランススペース手法の一つである Active Shape Model を用いることで得られる顔の特徴点の座標データを得る。座標値を正規化した値を特徴量として用い、Support Vector Machine を用いて表情の分類を行う。そしてあらかじめ作成しておいたデータベースをもちいて学習した Support Vector Machine を用いて認識を行う。分類するクラスとしては、一般的に用いられることの多い無表情、怒り、悲しみ、驚き、喜びの 5 種類としており、表情の認識の結果をエージェントの行動に反映させ、システムの対応力を向上させる予定である。

3. まとめ

本稿では、エージェント対話システムのための表情認識手法について検討した。微表情の認識を行うため、高速カメラを用いてエージェントの発言に合わせて適切なタイミングで細かな表情の認識を行うシステムを提案した。今後の展望としては、微表情を獲得するような状況で観測を行い、データベースを作成し、実験を行っていく。

本研究の一部は埼玉県先端産業実用化開発事業補助金および JSPS 科研費 26240038, 26540131 の助成を受けた。

参考文献

- [1] K. Ball, et al, "Effects of Cognitive Training Interventions With Older Adults," Journal of American Medical, Vol.288, pp.2271-2281, 2002.
- [2] W. J. Yan, et al, "CASMEII: An Improved Spontaneous Micro-Expression Database and the Baseline Evaluation," PLOS ONE, Vol.9, e86041, 2014.