

[研究助成 (A)]

複合現実空間における漫符を活用したロボットによる 自己表現生成 AI に関する研究



2241023

研究代表者

産業技術総合研究所

研究員

藤 井 綺 香

[研究の目的]

近年，社会における様々な場所でロボットを含めたエージェントが活躍を始めている。同じ空間で活動する仲間としてエージェントと人の親近感を高めるためには，情報伝達を行うだけでなく，感情表現を伴うインタラクションが重要であると考えられる。またロボットが，モータの不調や認識機能のエラーなどの内部状態を，人間に対してわかりやすく表現することも有用である。本研究では，自己表現を提示するハードウェアを持たないロボットを含めた様々なエージェントに適用可能で，汎用的な感情表現のインタラクション技術を構築することで，人間とエージェントのより良い協調・共生関係を構築することを目的とする。具体的には，漫符表現（漫画でキャラクターの状態を表現するためにシンボリックに用いられる手法）を活用して，豊かな感情表現を実現することを目指す。

ロボットによる感情表現に関する研究において，表情や動作，音声など，多様な手法の有効性が示されている。しかし，ロボットの形態は，ヒューマノイド，動物型，車輪移動型，アーム型など多岐にわたるため，多様な外観に幅広く適用可能な感情表現の手法を構築する必要があると考えられる。そこで本研究では，漫画において登場人物やキャラクターだけでなく，動物や物体にも豊かな感情表現を与えることのできる漫符に着目する（図1）。過去にも，ロボットにおいて漫符表現を用いることで，感情が伝

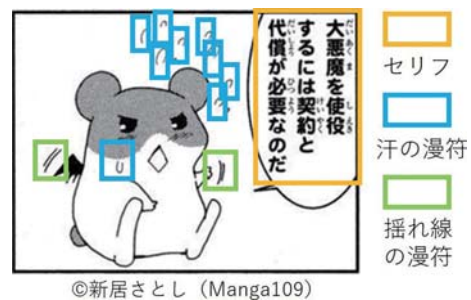


図1 漫画内で使用されている漫符の例

わりやすくなるという基礎実験が行われてきている^[1]。しかし，漫符表現に関する体系化されたデータベースや，対話の文脈に合わせて漫符を生成するモデルの開発は行われていないため，本研究ではまずこれらに取り組む。また，漫符を用いた従来研究では，ロボットにLEDディスプレイなどの専用の表現デバイスを組み込むことで感情表現を実現している。しかし，全てのロボットに表現デバイスを組み込むのは現実的ではないと考えられるため，複合現実（MR）技術を活用した提示手法の検討を行う。

[研究の内容，成果]

漫符を用いたヒューマンエージェントインタラクションの実現に向け，本年度は，①漫画のデータセットを用いた漫符のアノテーション，②漫符を表示しながら対話インタラクションを行うエージェントのプロトタイプ開発，③複合現実内でロボットに合わせて漫符を提示するシステムの基礎検討，を実施した。

1. 漫画のデータセットを用いたアノテーション

1.1 漫符のアノテーション方法の検討

まずは、人物や人型のキャラクターに着目して、漫符のアノテーション方法の検討を実施した。漫符をエージェントに合わせて表示するためには、漫符の種類と提示位置が大事であると考え、それぞれに関して、分類を行った。

漫符に関する文献^[2]等を参考に、図2に示すような55種類の漫符を抽出した。オノマトペや、背景や風景に関連するもの、表情やジェスチャーに分類されるようなものは除外した。付与位置に関しては、2次元での位置に関して検討した従来研究^[3]を参考にするとともに、最終的に3次元空間に存在するロボットにも適用することを想定し、人間の各身体部位の前、後ろ、外側、内側、などを含めた分類を行った。

また、実際に漫符を提示する上では、表示する個数も大事となる。例えば、汗を表現するときには、水しぶきの漫符が複数用いられる場合が多い。また、眠りのマークや気づき線などの、典型的に複数個の組み合わせで用いられるものに関しては、まとめて1つとみなして取り扱うこととした。



図2 分類を行った漫符一覧

1.2 アノテーションツールの開発及びアノテーション

漫符に関するアノテーションを行うデータセットとして、Manga109^[4,5]を利用した。これは、109冊の日本の漫画作品からなるデータセットであり、収録されている作品のジャンルは多岐にわたる。コマや、セリフなどのテキスト、キャラクターの顔や身体に関するアノテーションが付与されている。さらに、Manga109



(アノテーション対象例としている漫画：
© よしまさこ (Manga109 より))

図3 開発したアノテーションツール

内のセリフと発話者を対応づけたデータセットも存在する。しかし、漫符に関するアノテーションは提供されていない。

本研究ではまず図3に示すように、漫画内で描かれている漫符と、関連する登場人物の発話に対応づけてアノテーションするためのツールを開発した。開発は、Pythonでグラフィカルユーザーインターフェース (GUI) を作成するためのツールキットであるTkinterを用いて行った。GUIには、アノテーションの対象とするページとその前のページの漫画の画像を表示した。漫画の画像内の対象となる漫符をバウンディングボックスでアノテーションできるようにするとともに、その種類と位置をクリックするボタンや、個数を入力する欄を配置した。また、漫符が付与されているキャラクターの名前とセリフの内容を選択するプルダウンボックス及び、その前の文脈に関連するキャラクターの名前とセリフの内容を選択するプルダウンボックスを設けた。そして、これらの漫符と対話内容に関するアノテーション結果をcsv形式で保存できるようにした。

開発したアノテーションツールを用いて、Manga109データセットのアノテーション作業を実施した。本年度はまず人物に関連する漫符に焦点を当て、データセットの一部に対してアノテーションを行った。今後は、動物や物体などに付与されている漫符等も含めてアノテーションを拡張していきたいと考えている。

2. 漫符を用いた対話インタラクションシステム

2.1 文脈に合わせた漫符の生成

会話の文脈に応じて適切な漫符を推定・出力するモデルを構築するために、大規模言語モデルのファインチューニングを行った。大規模言語モデルには、OpenAI の GPT4o (gpt-4o-2024-08-06) を採用した。また、漫符と発話内容を関連づけた csv 形式のアノテーションデータを、GPT-4o に対応する JSONL 形式に変換することで、ファインチューニング用のデータを作成した。そして、アノテーションの方針に沿い、会話内容に関連した漫符の種類、場所、個数を、OpenAI API の function calling 機能を用いて取得できるようにした。

2.2 漫符を提示可能な対話エージェントの開発

図4に示すように、3DCG エージェントが対話の文脈に合わせた漫符を提示することで、豊かな感情表現を行うことを目指した対話インタラクションシステムのプロトタイプ開発を行った。

システム構築基盤として、キャラクター制御やユーザーインターフェースの設計に優れ、MR アプリケーションも開発可能なゲームエンジンである Unity を採用した。また、エージェントとして表示される 3D キャラクターモデルは、VRoidHub という 3D キャラクターの共有プラットフォーム内で公開されているものを利用した。

ユーザーインターフェースはシンプルな構成



図4 漫符を用いた対話インタラクションシステムの
プロトタイプ

であり、ユーザーが任意の会話内容をテキストで入力すると、それに対するエージェントからの応答がテキストで出力される。同時に、その対話内容に応じた漫符がエージェントの 3D モデルに合わせて表示されるというものである。エージェントのテキストによる応答内容及び、表示される漫符の種類や表示位置、個数に関しては、前述のファインチューニングを行った大規模言語モデルを用いて決定している。

漫符の提示に関する具体的な実装としては、まず Unity のシステム内にあらかじめ漫符の画像データを準備しておくとともに、3D エージェントモデルの既存のボーンに対応付けた漫符表示ポイントを設定した。そして、大規模言語モデルからの出力に応じて、対応する表示ポイントに画像データを Unity の GameObject として追加したり削除したりすることで、インタラクティブに漫符を提示するシステムとしている。

3. ロボットへの応用に関するシステム検討

実世界のロボットに漫符を用いた感情表現を拡張するために、MR 空間でロボットの身体に合わせて漫符を提示する手法の検討を実施した。

ロボットのモデル情報や、各関節角度の情報、Simultaneous Localization and Mapping (SLAM) 等による位置推定情報が取得可能な場合は、ヘッドマウントディスプレイ (HMD) 装着者の座標系とロボット座標系を対応付けることで、ロボットの骨格に応じて漫符を配置して提示することができる。また位置情報に関しては、AR マーカーを利用することでも、HMD 装着者とロボットとの相対的な位置関係を取得可能である。

一方で、ロボットの位置情報や関節角度情報などを取得できない場合は、HMD のカメラなどで取得した画像からのロボットの骨格推定が必要と考えられる。初期検証として、人型に近い身体デザインを有するロボットの画像に対して、MediaPipe や MMPose などの従来は人間

の骨格推定に用いられているモデルを適用し、検討を行った。その結果、HRP-4 や NAO などの二足歩行ロボットでは主要関節を推定でき、既存モデルを適用可能であった一方で、Miroki などの車輪移動ロボットなどは上半身も含めてうまく推定できない場合が多かった。今後は、ロボット実機のデータやシミュレータ上でのデータを用いて既存モデルの転移学習などを行うことで、ロボットのための骨格推定器を構築することを検討している。

上記の事柄を考慮して、Unity を通じて Robot Operating System 2 (ROS2) と MR ヘッドセットを組み合わせたシステムの開発に着手した。

4. おわりに

4.1 本研究のまとめ

本研究では、エージェントの感情表現を拡張する技術として、漫画におけるシンボル表現である漫符を活用する新たなアプローチを提案した。

まず、55 種類の漫符を抽出するとともに、人型のキャラクターにおける 3 次元空間への拡張を視野に入れて付与位置の整理を行った。そして、漫画のデータセット内の漫符と発話内容に対応づけたアノテーションデータと大規模言語モデルを用いて、会話の文脈に基づいて適切な漫符を提示するためのシステムを構築した。

また、漫符を用いたエージェントとのインタラクションのプロトタイプとして、3DCG エージェントに合わせて漫符を表示しながら対話を行うアプリケーションを実装した。ロボット実機への応用に向けて、ロボットの骨格推定や位置推定情報を基に、MR の HMD を用いて漫符の表示を行う仕組みの検討を進めた。

4.2 今後の展望

今後はまず、現在開発中の MR 空間でロボットの身体に合わせて漫符を提示するシステムの開発をすすめるとともに、対人実験を通じた評価を実施したいと考えている。

また現在は、人間に近い身体デザインのキャラクターを対象に、漫符のアノテーションを実施しているが、動物や物体に付与されている漫符に関しても分析を行いたい。そして、アーム型ロボットやロボット掃除機のような人型にも動物型にも該当しないエージェントの場合は、どのように漫符を提示するのが適切かということなども検討していきたい。

提示する漫符に関しても、現在の画像データから 3D モデルに変更し、動的エフェクトやアニメーションなどをつけることで、より豊かな表現を行うことができると考えている。

汎用的な感情表現手法である漫符と、比較的低コストで実装可能な複合現実技術を組み合わせることで、家電などの身近なものも含めたインタラクション技術への拡張も可能であると考えられる。本研究成果を発展させ、エージェントとのより良い共生社会の実現に貢献したい。

[参考文献]

- [1] T. Kishi, H. Futaki, G. Trovato, N. Endo, M. Destephe, S. Cosentino, K. Hashimoto, A. Takanishi: Development of a Comic Mark Based Expressive Robotic Head Adapted to Japanese Cultural Background, 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.2608-2613, 2014.
- [2] こうの史代: ギガタウン漫符図譜, 朝日新聞出版, 2018.
- [3] 赤井友紀, 盛山将広, 松下光範: 漫符を利用したコミック内の心情・行動描写検索システム, 情報処理学会研究報告, Vol.2017-DCC-15, No.42, pp.1-7, 2017.
- [4] Y. Matsui, K. Ito, Y. Aramaki, A. Fujimoto, T. Ogawa, T. Yamasaki, K. Aizawa: Sketch-based Manga Retrieval using Manga109 Dataset, Multimedia Tools and Applications, vol.76, no.20, pp.21811-21838, 2017.
- [5] K. Aizawa, A. Fujimoto, A. Otsubo, T. Ogawa, Y. Matsui, K. Tsubota, H. Ikuta: Building a Manga Dataset "Manga109" with Annotations for Multimedia Applications, IEEE MultiMedia, vol.27, no.2, pp.8-18, 2020.

[成果の発表, 論文など]

- [1] Ayaka Fujii, Ken Fukuda: Development of a Human-Agent Dialogue Interaction System Using Comic Symbols for Enriching Agent Expression, HCI International 2025, 2025. (発表予定)