

プロジェクションマッピングを用いた対話エージェントによる内部状態表出

Expressing Internal State of Listening Agent with Projection Mapping based Technology

石原 義久 *1 小林 一樹 *2
Yoshihisa Ishihara Kazuki Kobayashi

*1信州大学大学院 総合理工学研究科 *2信州大学 学術研究院
Graduate School of Science and Technology, Shinshu University Academic Assembly, Shinshu University

This paper proposes a method for manufactured objects such as anime figures to exhibit highly realistic behavioral expressions to improve speech interaction between a user and an artifact. Using a projection mapping technique, an anime figure provides back-channel feedback to a user by appearing to nod or shake its head. We developed a simple and highly feasible method to express internal states of the artifact by translating a projected solid graphic on the artifact.

1. はじめに

近年, Apple 社の Siri や Google 検索などにおいて自然言語を用いるユーザインタフェースがスマートフォンをはじめとする身近なデバイスに搭載されている。これらのインタフェースは, ユーザがデバイスに話しかけることによって電話をかけたり Web 検索といった機能を提供する。しかし, ユーザは音声入力を公の場所で使用することをためらう傾向がある。音声入力時には, 音声通話に比べて趣味や趣向に関する言葉を使用する可能性が高く, プライバシーを保つために使用していない可能性がある。しかし, 音声通話と音声インタフェースに対する受容の差は, デバイスの人間らしさや知性に起因している可能性もある。たとえば, コンピュータを相手に画面上でしとりゲームを行うとき, 人間が相手であると知らされた場合には, 相手がコンピュータプログラムであると知らされた場合よりも長い時間ゲームを続けることが報告されている [山本 94]。また, 程度の差はあれど, 誰もがコンピュータに対し社会性を見出し, コンピュータに対し礼儀正しく接したり社会の一員として接することが報告されている [Byron 01]。

本研究ではエージェントのコミュニケーションのリアリティを向上させることを目的とし, 既存の人工物に動きを中心とする非言語情報を付加してユーザと対話する手法を提案する。対話対象として, 市販のアニメフィギュアを用い, うなずき動作と首振り動作によるあいづち表現をプロジェクションマッピング [Aliaga 12][Amano 13] により実現する。

2. 関連研究

これまで, 機械の人間らしさやユーザと機械との対話体験の向上に関する研究が行われてきた。Breazeal ら [Breazeal 05] は, 非言語情報を利用する人型ロボットが人間とのチームワークを向上させることを示している。Riek ら [Riek 10] は, チンパンジーに似せた頭部ロボットを用い, ユーザのうなずきなどの非言語情報を模倣し, ユーザと調和した関係の構築を試みている。Goetz ら [Goetz 03] は, 社会的なタスクを実行するロボットがより人々に受け入れられることを示している。また, ロボットとソフトウェアエージェントとを比較する研究と

表 1: 先行研究の実験における各水準の内容

	PV	V	P	C3
プロジェクションマッピング 音声	(1) ↑ +はい/うん (2) ⇄ +ううん	なし	なし	(1) ↑ +はい/うん (2) ⇄ +ううん
音声	なし	(1) はい/うん (2) ううん	なし	(1) はい/うん (2) ううん
プロジェクションマッピング	なし	なし	(1) ↑ (2) ⇄	(1) ↑ (2) ⇄

↑ プロジェクションマッピングによるうなずき動作
⇄ プロジェクションマッピングによる首振り動作

して, Powers ら [Powers 07] は画面に表示されたエージェントと実ロボットとのインタラクションを比較する実験を行い, ロボットが画面に表示されたエージェントよりも, 実ロボットのほうがユーザに対して好印象を与えることを示した。さらに, 原田 [原田 06] は第三者がいる環境において, 画面に話しかけるよりも, ロボットに話しかける方が, ユーザの違和感を和らげると報告している。

上記の先行研究では, 非言語情報を用いたり, 社会的タスクを実行するロボットが人間らしさを向上させること, ソフトウェアエージェントよりも実ロボットのほうがユーザに対して好印象を与えることが示されている。しかし, ロボットの場合, アクチュエータを用いるために, 動作にぎこちなさが生じるため, ロボット制御の問題点として指摘されている [石黒 01]。

このような問題に対し, 著者らは先行研究として, プロジェクションマッピングを用いた投影型行動表出エージェントを開発した [Ishihara 16][石原 16a][石原 16b]。エージェントに道案内を行うタスクを実行する参加者実験を表 1 に示す水準を設定し行なった。その結果プロジェクションマッピングを用いた行動表出エージェントは, ユーザに対しエージェントに意思があるように感じさせたり, エージェントの振る舞いが人間であるかのような振る舞いに感じさせたり, エージェントが話を真剣に聞いているかのように感じさせる効果を出し, かつ会話のテンポを円滑にするという効果があり, エージェントのリアリティを高めることができることを示唆する結果を得ている。本手法の利点は, 対象物の形状に依存せず, モータのようなアクチュエータを用いないため, ぎこちない振動を生じさせることなく自然な表現を可能にし, 非言語情報を用いて人工物の人間

連絡先: 石原義久, 信州大学大学院総合理工学研究科

〒 380-8553 長野県長野市若里 4-17-1

E-mail: 16w2003f@shinshu-u.ac.jp



図 1: 投影型行動表出エージェントの外観

らしさを向上させる。会話におけるうなずきなどの非言語情報は合意形成や、会話内容、会話の長さに影響を与える重要な要素である [Clark 94][Schuller 07]。これまで、MMDAgent をはじめとして人間らしいソフトウェアエージェントが提案されているが [Lee 13][André 98][Bickmore 01]，実体を持つロボットの方がリアリティが高い [Kiesler 08]。

また、最近ではプロジェクションマッピングを用いた表現として注目されている変幻灯という技術がある [Kawabe 16]。これは動作の差分を投影することで静的な対象が動いているかのような微細な表現を可能にする技術である。変幻灯と同じ原理を用いた手法を用いることでより高いリアリティを実現することができる可能性がある。著者らの先行研究では、コミュニケーションのリアリティの向上を目的として、人工的でシンプルな表現を用いていた。それに対し、変幻灯は対象の視覚的なリアリティを向上させることができる。これらの表現を比較することにより、よりコミュニケーションのリアリティを向上させるエージェントの行動表現が明らかになることが考えられる。

以降では、著者らの既存の表現手法と変幻灯と同じ原理を用いた手法を用いた表現手法に関する説明、両表現手法の比較を行う実験計画について述べる。

3. 投影型行動表出エージェント

開発した投影型行動表出エージェントを図 1 に示す。エージェントはアニメフィギュアと台座 (310 × 240 × 110mm)、プロジェクタで構成され、台座の中に小さなプロジェクタを配置し、その上に人型のアニメフィギュアを置く。プロジェクタはフィギュアの顔にアニメーションを投影できるように固定した。本エージェントは既存の人工物にうなずき動作や首振り動作などの動的な行動表出を付加してユーザと対話を行うことが可能になり、身近な愛着物に動作を加えてエージェント化することが可能になり、ユーザとエージェントの円滑な対話を実現することが期待できる。

3.1 エージェントの振り舞い

エージェントはユーザの発話に対して肯定的表現としてうなずき動作、否定的表現として首振り動作を行い、ユーザの

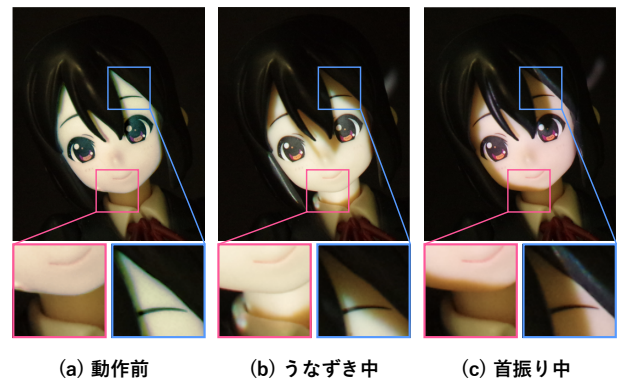


図 2: 平行移動表現の動作

発話に対して応答を行う。また、肯定的表現時には「はい」と「うん」、否定的表現時には「ううん」という音声出力を行う。音声は女性の声を録音し、ピッチを変更したものを使用した。エージェントはユーザの発話を常に監視しており、一定以上の音量の発話が 200ms 以上継続した場合にあいづち動作を行う。

3.2 平行移動表現

平行移動表現はプロジェクションマッピングを用いて、エージェントの顔の輪郭に合わせた白黒の図形を投影する表現である。本手法は対象のテクスチャに依存せず、画像の加工が容易であるためシンプルで適用範囲が広いと考えられる。図 2 に平行移動表現の動作の様子を示す。うなずき動作は投影画像を 150ms で 20px 下に動かし、150ms で 20px 上に動かす。首振り動作は投影画像を 150ms で 10px 右に動かし、150ms で 20px 左に、150ms で 10px 右に動かす。投影画像は Web ブラウザで動作し、操作画面と投影画面を有する。投影画面に表示された表現アニメーションは PC に接続したプロジェクタから出力している。投影ソフトウェアは JavaScript と PHP を用いて開発した。

3.3 画素値差分表現

画素値差分表現は変幻灯と同じ原理を用いた手法を用いて、動作前画像と動作後画像の差分画像を投影する表現である。図 3 に画素値差分表現の動作の様子を示す。エージェントの動作表現は、エージェントの動作前のグレースケール画像から動作後のグレースケール画像を作成し、動作後の画像から動作前の画像を引いた差分画像を生成して投影することで表現する。うなずき動作は縦方向に 5px 縮小して、動作前と同じこめかみの位置に合わせた差分画像を投影する。首振り動作は横方向に 5px 縮小して、5px 右に移動させた画像と元画像の差分画像を投影する。

4. 実験計画

平行移動表現と画素値差分表現の有効範囲及び特徴の比較を行うための実験計画について述べる。

4.1 実験タスク

参加者には、エージェントに信州大学工学部正門から JR 長野駅までの道案内を行うように指示する。また、会話時間が 3 分間であること、エージェントの反応を見ながら会話すること、エージェントに伝わるようにわかりやすくはっきりとした声で発話をすることが指示される。対話終了後、参加者に対してアンケート調査を実施する。

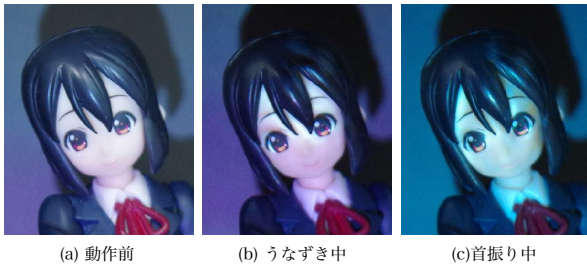


図 3: 画素値差分表現の動作

4.2 水準

実験には以下の 2 水準を設ける。平行移動表現のプロジェクトンマッピングを用いた表現によるあいづち動作を行う水準、画素値差分表現のプロジェクトンマッピングを用いた表現によるあいづち動作を行う水準とする。参加者間配置とし、参加者一人あたり 1 水準を経験する。これらの水準間はエージェントのあいづち動作の表現以外の条件は全て同じとする。

4.3 評価指標

評価指標は、アンケート調査による主観評価、エージェントとの対話における参加者の発話回数、参加者の平均発話時間、参加者の平均無音時間、参加者が発したフィラーの数とする。会話中のフィラーは話者の発話権の保持や深く考えていることを表すと報告されており [Candea 05] [水上 07][Clark 02], 集中の度合いを比較するために採用した。

5. まとめと今後の展望

本稿では、プロジェクトンマッピングによって行動表出を行うエージェントの表現手法の比較を行うための実験計画について述べた。参加者実験を実施することで、エージェントの顔の輪郭に合わせた白黒の図形を投影する平行移動表現と動作前画像と動作後画像の差分画像を投影する画素値差分表現のそれぞれの特徴と有効範囲が明らかになると考えられる。

提案手法は既存の静的な人工物に対して反応性を付与することができる。そのため、たとえば思い入れのある人工物や、アクチュエータを組み込む改造が困難な人工物に提案手法を適用することで、そのような人工物とのコミュニケーションのリアリティを高められる可能性がある。

謝辞

本研究の一部は、JSPS 科研費新学術領域研究「認知的デザイン学」(No.JP26118005) の助成を受けたものである。

参考文献

- [Aliaga 12] Aliaga, D. G., Yeung, Y. H., Law, A., Sajadi, B., and Majumder, A.: Fast high-resolution appearance editing using superimposed projections, *ACM Transactions on Graphics (TOG)*, Vol. 31, No. 2, p. 13 (2012)
- [Amano 13] Amano, T.: Projection based real-time material appearance manipulation, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 918–923 (2013)

- [André 98] André, E., Rist, T., and Müller, J.: Integrating reactive and scripted behaviors in a life-like presentation agent, in *Proceedings of the second international conference on Autonomous agents*, pp. 261–268 ACM (1998)
- [Bickmore 01] Bickmore, T. and Cassell, J.: Relational agents: a model and implementation of building user trust, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 396–403 ACM (2001)
- [Breazeal 05] Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., and Berlin, M.: Effects of nonverbal communication on efficiency and robustness in human-robot teamwork, in *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)*, pp. 708–713 (2005)
- [Byron 01] Byron, R. and Clifford, N.: 人はなぜコンピュータを人間として扱うか「メディアの等式」の心理学, 翔泳社 (2001)
- [Candea 05] Candea, M., Vasilescu, I., and Adda-Decker, M.: Inter-and intra-language acoustic analysis of autonomous fillers, in *DISS 05, Disfluency in Spontaneous Speech Workshop*, pp. 47–52 (2005)
- [Clark 94] Clark, H. H.: Managing problems in speaking, *Speech communication*, Vol. 15, No. 3, pp. 243–250 (1994)
- [Clark 02] Clark, H. H. and Tree, J. E. F.: Using uh and um in spontaneous speaking, *Cognition*, Vol. 84, No. 1, pp. 73–111 (2002)
- [Goetz 03] Goetz, J., Kiesler, S., and Powers, A.: Matching robot appearance and behavior to tasks to improve human-robot cooperation, in *Proceedings of the International Workshop on Robot and Human Interactive Communication (ROMAN)*, pp. 55–60 (2003)
- [Ishihara 16] Ishihara, Y., Kobayashi, K., and Yamada, S.: Behavioral Expression Design Onto Manufactured Figures, in *Proceedings of the Fourth International Conference on Human Agent Interaction, HAI '16*, pp. 243–244, New York, NY, USA (2016), ACM
- [Kawabe 16] Kawabe, T., Fukiage, T., Sawayama, M., and Nishida, S.: Deformation Lamps: A Projection Technique to Make Static Objects Perceptually Dynamic, *ACM Trans. Appl. Percept.*, Vol. 13, No. 2, pp. 10:1–10:17 (2016)
- [Kiesler 08] Kiesler, S., Powers, A., Fussell, S. R., and Torrey, C.: Anthropomorphic interactions with a robot and robot-like agent, *Social Cognition*, Vol. 26, No. 2, p. 169 (2008)
- [Lee 13] Lee, A., Oura, K., and Tokuda, K.: Mmdagent—a fully open-source toolkit for voice interaction systems, in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8382–8385 (2013)

-
- [Powers 07] Powers, A., Kiesler, S., Fussell, S., and Torrey, C.: Comparing a computer agent with a humanoid robot, in *Proceedings of the International Conference on Human-Robot Interaction (HRI)*, pp. 145–152 (2007)
- [Riek 10] Riek, L. D., Paul, P. C., and Robinson, P.: When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry, *Journal on Multimodal User Interfaces*, Vol. 3, No. 1-2, pp. 99–108 (2010)
- [Schuller 07] Schuller, B., Seppi, D., Batliner, A., Maier, A., and Steidl, S.: Towards more reality in the recognition of emotional speech, in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 4, pp. IV–941IEEE (2007)
- [原田 06] 原田悦子: 感性心理とロボティクス ロボットの「使いやすい」?: インタフェイスとしてのコミュニケーション型ロボットを考える, *日本ロボット学会誌*, Vol. 24, No. 6, pp. 720–724 (2006)
- [山本 94] 山本吉伸, 安西祐一郎: 特集 人間とシステムのインタラクションインタラクションの楽しさ, *人間工学*, Vol. 30, No. 1, pp. 15–25 (1994)
- [水上 07] 水上悦雄, 山下耕二: 対話におけるフィラーの発話権保持機能の検証, *認知科学*, Vol. 14, No. 4, pp. 588–603 (2007)
- [石原 16a] 石原義久, 小林一樹: 表出プロジェクションを用いた対話エージェントの開発, *Human-Agent Interaction Symposium 2016 P-4* (2016)
- [石原 16b] 石原義久, 小林一樹, 山田誠二: 投影型行動表出エージェントによる対話体験の向上, *HCG Symposium 2016 C-4-3* (2016)
- [石黒 01] 石黒章夫: きびきびと動くロボットを目指して「心身一元論」に立脚した人工知能システムの構築に向けて, *電気学会論文誌. C, 電子・情報・システム部門誌* = The transactions of the Institute of Electrical Engineers of Japan. C, A publication of Electronics, Information and System Society, Vol. 121, No. 1, pp. 48–49 (2001)