



人との対話を拡張するアンドロイドアバター共生社会



大阪大学
大学院基礎工学研究科
博士前期課程
株式会社国際電気
通信基礎技術研究所
石黒浩特別研究所
学外実習生

たにがわ とも き
谷川 智樹



株式会社国際電気
通信基礎技術研究所
石黒浩特別研究所
研究技術員

ふなやま とも き
船山 智



国立研究開発法人
理化学研究所
ガーディアンロボット
プロジェクト
チームディレクター
株式会社国際電気
通信基礎技術研究所
石黒浩特別研究所
客員研究員

みなと たか し
港 隆史

1. はじめに

遠隔地から操作して様々な作業等を行うことができるサイバネティック・アバター（CA）は、人の日常活動を身体的、空間的、時間的制約から解放することができるため、人の活動能力を拡張させている（作業量や作業効率の向上）と見ることができる。このような能力拡張を、コミュニケーション能力においても実現することができる。例えば、ホスピタリティを求められるサービスタスクでは、丁寧な振る舞いが求められるが、そのような振る舞いを習熟するのは容易ではない。サービス業務を行う者が操作するCAが、操作者本人よりも丁寧な振る舞いを行うことができれば、操作者本人のコミュニケーション能力を拡張できていると考えられる。さらに、CAがその身体だけでなく、様々な表出デバイスをも制御することによって、多様な情報を表出することができれば、生身の人間ではできないような意図伝達もCAによって可能になる。

本研究では、これまでに存在感CA（人のような存在を感じさせるアンドロイドロボットアバター）を用いて人のコミュニケーション能力を拡張させる技術開発に取り組んできた。本研究において主に2つの課題があると考えている。1つ目は、コミュニケーションの性能を拡張するための存在感CAの表出方法である。本稿では、これに関する取組みとして、Shosaと呼ぶインタラクションの開発について紹介する。2つ目は、コミュニケーションにおいて操作者が相手に伝えたい、あるいは表出したい意図と、CAが表出する意図とのアライメントの課題である。本研究では、2者間でのコミュニケーションにおいて双方がCAを用いるシステムの実現に取り組んでいるが、このシステムの上でも上記の問題が発生する。本稿では、この取組みについても紹介する。

2. マルチモーダル表出を用いたコミュニケーション能力拡張の試み—Shosa—

2.1 Shosa

人のコミュニケーションにおいて、心がつながる感覚は極めて重要であり、それには身体表現が重要である。我々が他者との関わりの中で安心感や親近感を得られるのは、言葉の正確さだけでなく、表情や視線、身体の動きといった非言語的な情報を通じて互いの心を感じ取れるからである。人は生得的に身体を持ち、他者の所作から微細なシグナルを読み取るように知覚が最適化されている。それゆえ、ロボットにおいても身体性を伴った非言語情報は、意図や情動の伝達、さらには良好な関係性の形成において大きな役割を果たし得る（例えば[1]）。

本節で紹介する Shosa は、ロボットの身体動作に音と光を重ね合わせることで、ロボットと対面する人がロボットと「心が通うように感じられる」インタラクションの実現を目指した試みである（図1）。Shosaは、人とロボットが互いの動きや存在を感じ合う中で、情動的なつながりを生み出すことを目的として設計されている。CAとなるロボットにこのような振る舞い機能を持たせることができれば、操作者はCAを用いることで他者と「心が通うような」コミュニケーション能力を有することができる。



図1. Shosaの体験の様子



2.2 概念設計：所作の拡張としてのShosa

「所作」は身体の微細な動きによって心情を伝える日本的な身体文化であり、そこに内在する特徴の1つが曖昧性である。明確な意味を指定せず、見る側の解釈に一定の余白を残すことで受け手の想像力や感情的な参与を促す。曖昧性は現代アートデザインの実例からも、ヒューマンコンピュータインタラクションにおいてユーザに個人的で感情的な体験を生み出すことに有用なデザインリソースであることが指摘されている^[2]。Shosaではこのような動作表現をアンドロイドロボットに実装しつつ、この概念を拡張し、音声なし（動きのみ）のインタラクションにおいてアンドロイドの動作に音響・照明演出を連動させることで“複合的な情動表現”を生成することを試みた。

2.3 システム構成

Shosaは、RGBDカメラと光学式モーションキャプチャを併用し、アンドロイドと対面する人の姿勢・動きのリアルタイム認識を行う。それらの情報に基づいて、アンドロイドの動作・音・光がリアルタイムに制御される。印象に大きく影響する視線の制御には、反射的な注視行動と意図的な動作を組み合わせる確率モデルを導入し、「自律性」と「相互反応性」のバランス調整をしている。また、音と光は人とアンドロイドの手先位置と速度の相関度を定量指標とした調和度に基づいて変化させ、アンドロイドとの一体感が高まるよう設計した（詳細は[3] 参照）。

2.4 体験者の評価

ATRオープンハウス2023内で開催した日本国際芸術祭特別企画の展示では、71名の一般来客にShosaを体験してもらい、体験後のアンケートでは6割を超える体験者がアンドロイドとの心のつながりを感じたと回答した（図2）。また多くの体験者から「表情が豊か」「生きているようだった」「人間と対話している感覺」といった肯定的な回答を得た。

この取組みにより、動き・音・光の複合表現を通じて、機械的な応答を超えた“心的なつながり”を実感させる可能性が示された。心と心がつながるロボット体験の実現は、人とロボットの共生における新しい方向性を提示し、より



豊かな社会的・情動的関係性の構築に寄与すると期待される。本展示では、アンドロイドは自律的に動作しており、CAとしての役割を持っていないが、このような表出が可能なアンドロイドを存在感CAとして用いながら他者と対話するシステムを構築できれば、人のコミュニケーション能力を拡張した対話を実現することができる。

3. 双方向CAシステムの開発

3.1 CAコミュニケーションの拡張

存在感CAを用いた遠隔コミュニケーションの形態は、これまで非対称であった。すなわち、一方はCAの操作者で、遠隔地に置かれたCAや周囲の環境を映像等で観察しながらPCなどでCAを操作する。他方は、CAと対面して、操作者の身代わりとなったCAとインタラクションを行う。この形態では、CAと対面する人は操作者の存在を感じながら対話ができるが、操作する者は同様の効果を得ることができない。そこで、本研究では、双方向が存在感CAを用いて対話できる形態（図3）の実現を目指して開発に取り組んでいる。本システムは、操作者の身体動作をリアルタイムに計測し、相手との対話において伝えたい「意図」を損なわずに、計測した情報に基づいて遠隔地のCAの動作を表出させる。

3.2 双方向CAコミュニケーションの課題

上述したシステムにおいて、計測した対話者の姿勢（関節角度や手先位置）と相同的の姿勢を遠隔地のCAにとらせると、対話者A（B）とアバターA（B）の身体的特徴の違い、対話者とアバターの相対的位置の両地点で違いにより、不自然なインタラクションが発生する。例えば、対話者Aが相手のアバターBと目を合わせて話し掛けたとする。このときアバターBと対話者Bの身長が異なっていると、対話者Aの首の角度と同じ角度だけアバターAの首の角度を実現したとしても、アバターAは対話者Bの目を見るとは限らない。すなわち対話者Aが視線を合わせようとしている意図が対話者Bに伝わらない。したがって、対話者の動作から対話に必要な意図を抽出し、対話者の動作を意図から外れない動作とそれ以外の動作に分類し、意図から外れない動作の意図を伝えるための動作をCAに行わせるような機能を実現する必要がある。

3.3 双方向CAコミュニケーションシステムの開発

開発したシステムは、大きく分けると、表情を制御する表情制御モジュール、視線と頭部動作を制御する視線頭部制御モジュール、手の動きを制御する身体動作制御モ

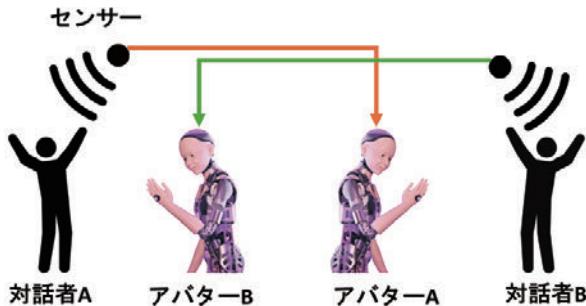


図3. 双方向CAコミュニケーションシステム

ジュー、対話音声を遠隔地に伝送する対話音声モジュールの4つのモジュールで構成されている。

表情制御モジュールは、対話者の表情をCAに反映させる。このモジュールでは意図を伝える動作として、対話の際に有効となる笑顔が相手に伝わることを目指した。そこで、対話者の顔画像から対話者の表情を6基本感情表情に分類し、「喜び」と判定された場合に、CAのアクションユニット(AU: 表情動作の最小単位)のうち、笑顔に関するAUの指令値を修正することで、CAの笑顔表出を強調する。それ以外の表情と判定された場合は、顔画像から計測された対話者のAUの動きをそのままCAに反映する。こうすることで、対話者の喜んでいる状態が効果的に相手に伝わるようにした。

視線頭部制御モジュールは、対話者の視線と頭部の動きをCAに反映させるモジュールである。意図を伝える動作として、対話者がCAを見たという動作が、遠隔地のCAでも再現されるようにした。顔画像とRGBD情報から対話者の視線及び頭部の仰角・方位角を計測する。この情報から、対話者が対面しているCAを見ているかどうかを判定し、見ているときは遠隔地のCAの視線を対話者に向けるように頭部と視線の姿勢を制御する。見ていないときは、遠隔地のCAの頭部・視線の姿勢を計測した角度と一致させるようにした。

身体制御モジュールは、対話者の腕と胴体の動きをCAに反映させるモジュールである。Azure Kinectによって対話者の手先位置が体の右、左、上、下、前、中央のどの領域にあるかを計測し、手先位置に応じて、対応する座標の手先位置にCAの手先を移動させる。

対話音声モジュールは、マイクで取得した対話者の発話音声を遠隔地に伝送し、CAの背後に設置したスピーカで再生する。このとき、音声に合わせて口唇動作を自動生成し、CAに反映している。

3.4 体験者の評価

上述したシステムをATRオープンハウス2025の参加者に

体験してもらい、体験者の評価を集めた。2名の参加者がそれぞれのCAを介して対面し、自由な内容で数分間対話してもらった。本展示では、握手やハイタッチといった日常的な対話では生じにくい身体接触が、参加者間で自発的かつ積極的に行われる様子が確認された。

また、体験者からは、「顔、声、存在感などいろいろなモダリティで実際に人と話している感覚が強かった」、「実際に対面で話している感じがして面白かった」、「遠隔で握手できるのは想像以上に新しい体験だった」、「生身の人間だと手を合わせたりつないだりしないけれど、それが不思議と自然にできて、それをきっかけに少し親しくなれた感じがあった」というような感想が得られた。

身体接触は、対人間の親密度や関係性を高める要因として知られている。本展示での観察結果を踏まえると、本コミュニケーション形式は、遠隔でありながらも相互の心理的距離を縮め、親密度を向上させる効果を持つ可能性が示唆される。

4. おわりに

Shosaのシステムで構築した表出機能を、双方向CAシステムに持たせることにより、双方向でコミュニケーション能力を拡張しつつ対話ができるアバターの実現が今後の課題である。本稿で紹介したシステムの体験者の評価から分かるように、存在感CAは、より親密な対話や、発話の持つ意味の交換以上の対話者間のつながりをもたらす可能性がある。このようなCAが日常生活の中に人と共存して、いつでも容易にCAを使えるようになると、人々の関係の維持・発展に貢献できるのではないかと考えている。

参考文献

- [1] G. Laban, A. Kappas, V. Morrison, and E. S. Cross. Building long-term human-robot relationships: Examining disclosure, perception and wellbeing across time. International Journal of Social Robotics, 16(5): 1-27, 2024.
- [2] W. W. Gaver, J. Beaver, and S. Benford. Ambiguity as a resource for design, Conf. on Human factors in computing systems, pp. 233-240, 2003.
- [3] M. Sato, T. Minato, T. Funayama, H. Sumioka, K. Sakai, R. Mikata, H. Ishiguro, K. Horibe, A. Kikuchi, and K. Sakuma, Analysis of heart-to-heart communication with robot using transfer entropy, Int. Conf. on Robot and Human Interactive Communication, pp.75-82, 2024.