

体内エージェントの心理的効果検証とワイヤレスデバイスの開発

Psychological Evaluation of an In-Body Virtual Agent and Development of a Wireless Speaker Device

熊木 拓也^{1*} 田中 文英²
Takuya Kumaki¹ Fumihide Tanaka²

¹ 筑波大学 知能機能システム学位プログラム

¹ Master Programs in Intelligent and Mechanical Interaction Systems, University of Tsukuba

² 筑波大学 システム情報系 知能機能工学域

² Institute of Systems and Information Engineering, University of Tsukuba

Abstract: With advancing AI, various companion agents have been developed. This study provides an interactive experience in which a virtual agent is spatially localized inside the user's body. We enhanced the sense of in-body localization by expanding the speaker array channels and by modifying the type of the agent's voice. Experimental results showed that the 8-channel configuration improved the perception of front-back movement and the feeling of closeness, while the self-similar voice increased the sense of body ownership but also emotional loneliness. To enhance immersion and practical utility, our future work focuses on developing a wireless system based on these findings.

1 はじめに

音声エージェントとのインタラクションは、AI技術の発展に伴い高度化している。これらの音声エージェントはスマートフォンやスピーカーなどから出力されるため、ユーザーは基本的に外部のデバイスと対話を行っている。しかし、本研究ではベルト型のスピーカーアレイデバイスにより身体内部の腹部付近にバーチャルエージェントを仮想的に定位させる体内エージェント [1] を用いることにより、ユーザーの身体の中という最も近接した位置に存在するエージェントとの対話を可能にしている。体内エージェントのコンセプト図を図 1 に示す。

Hall の近接学では、人と人との物理的距離は対人関係における心理的親密さに影響を与えることが示されており、人は相手との距離に応じて安心や信頼を感じると述べられている [2]。この考え方はロボットやエージェントとのインタラクションにも応用されており、好感度や信頼性の低いロボットに対して、人間は物理的に距離をとり、個人的な情報を開示することを避ける、あるいはロボットとの交流を完全に停止する傾向にあることが報告されている [3]。しかし、これらの研究はいずれも身体外部のロボットやエージェントに対する



図 1: 体内エージェントのコンセプト図

ものであり、身体内部という自分と一体化したエージェントを対象とした研究は存在しない。体内エージェントを用いることでこのような負の距離のエージェントとのインタラクションが可能となるが、そのためにはユーザーにエージェントが身体内部にいるように強く知覚される必要がある。

本論文では、エージェントが身体内部に存在している感覚を向上させることを目的として実施した施策とその検証結果を示し、現在行っているワイヤレスデバイスの開発について述べる。

*連絡先: 筑波大学知能機能システム学位プログラム
〒 305-8577 茨城県つくば市天王台 1-1-1
E-mail: takuya.kumaki@ftl.iit.tsukuba.ac.jp

2 関連研究

エージェントの定位感を向上させるための関連研究としては、スピーカーアレイ構成の拡張が挙げられる。例えば、体外における音像定位について、音像の位置を変化させることで方向知覚の精度が向上することが報告されている [4, 5, 6]。河原らはスピーカーから出力される音量を調整することで音像を移動させ、体内の音像をユーザーが知覚しやすいようにしている [1]。しかし、既存デバイスは前後左右の4方向に配置されたスピーカーアレイで構成されており、特に前後方向の音像移動を正確に知覚することが難しい。

また、自己とエージェントの一体感や身体所有感などの心理的要因も定位感に影響を与えられられる。VRの研究では、自己類似の声を持つ仮想空間上のアバターに対して好感度や信頼性が高まる傾向にあり [7]、アバターの性別がユーザーと一致している場合に身体所有感が高まることが報告されている [8]。これらを含む研究の多くは体外または画面上のアバターを対象としたものであり、体内に提示される音声の知覚や心理的効果を調査した研究は存在しない。

本研究では、スピーカーアレイの構成を4chから8chに拡張し、体内での定位感を向上させる試みと、エージェントの音声を変更した際の心理的影響を調査する。

3 実験

本研究では、体内での音像の定位感を向上させることを目的として、以下の2つの試みを実施した。

1. スピーカーアレイの構成を4chから8chへ拡張
2. エージェントの音声の種類を変更

また、体内エージェントの潜在的なユースケースを探るため、評価実験の終了後にユーザーインタビューを実施した。

まずスピーカーアレイ構成の拡張について述べる。音像の定位位置の計算には Vector Base Amplitude Panning (VBAP) 法を採用した [9]。VBAP法では、任意の位置に配置されたスピーカーを用いて仮想音源を定位させている。したがって、ベクトルを構成するスピーカーの数を増やすことで仮想音源を定位させる領域を細かく分割し、定位の精度を上げることが可能である。既存のデバイス [1] は前後左右の4方向にスピーカーが配置されており、振幅パンニングに基づいた音像の移動を可能にしている。本研究ではこれに加えて斜め方向にスピーカーを追加し8ch構成にすることで、音像移動の知覚を容易にするとともに、音像定位の精度向上を図った。



図 2: 8ch スピーカーベルト

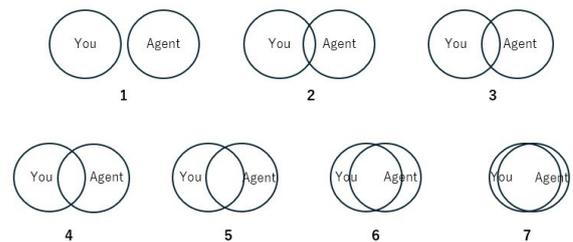


図 3: 主観的な親近感の測定に用いた IOS スケール

次にエージェントの音声の変更について述べる。音声としては男性音声、女性音声、およびユーザーの音声をもとに生成した自己類似音声の3種類を用意している。自己類似音声に関しては、VR分野の研究において好感度や信頼性の向上に寄与することが報告されている [7]。本研究では、体内に定位する自己類似音声ユーザーに与える心理的影響を調査するとともに、音声の特性が体内定位感に及ぼす影響を明らかにすることを目的とした。

3.1 実験方法

実験には、8chのスピーカーベルト(図2)を使用した。実験参加者は、普段から体内エージェントを使用していると仮定し、その日にあった楽しかったことや悲しかったことについて話すように求められた。音声認識には Google Cloud Speech-to-Text API v1, 応答生成には Google Gemini 2.0 Flash, 音声合成には VOICEVOX または Elevenlabs Instant Voice Clone を使用した。

3.2 実験参加者

本実験には、31名の大学生・大学院生(男性21名、女性10名、M=22.72, SD=1.26)が参加した。本研究は筑波大学システム情報系の研究倫理審査委員会の承

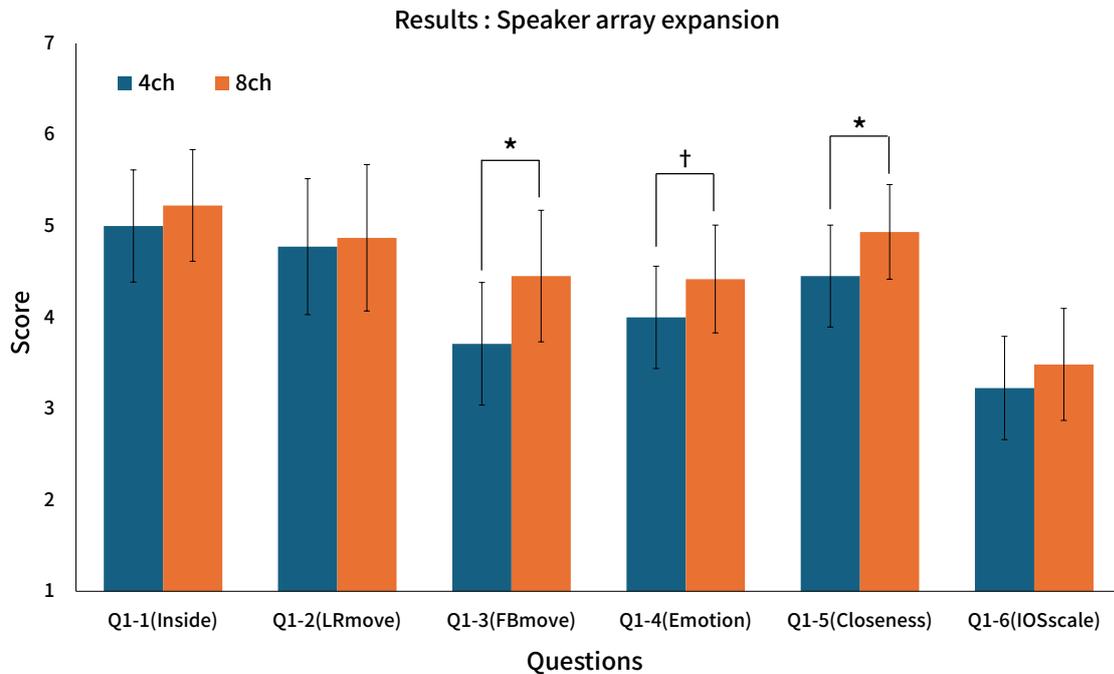


図 4: 実験 1 の結果 (平均値と標準誤差を示す * $p < 0.050$, † $p = 0.062$)

認を得ている (2023R723)。実験参加前に参加者全員からインフォームドコンセントを得た。また、体内エージェントの効果が不明確であるため、中枢性脳障害や精神疾患の既往歴のある参加者は対象から外した。

3.3 実験結果

3.3.1 スピーカーアレイ構成の拡張

参加者はまずスピーカーアレイ構成の拡張に関する実験として、4ch スピーカーベルトの条件と 8ch スピーカーベルトの条件での対話を体験した。この 2 つの条件はランダムな順序で実施された。それぞれの条件終了後、印象を調査する以下のアンケートに回答した。これらの項目はすべて 7 段階の評価尺度で採点された。また、2 条件を比較するための自由記述欄を設けた。

- Q1-1 どの程度エージェントの声はあなたの身体の中から聞こえてきましたか
- Q1-2 どの程度エージェントが左右に移動しているのを感じましたか
- Q1-3 どの程度エージェントが前後に移動しているのを感じましたか
- Q1-4 どの程度エージェントから感情を感じましたか
- Q1-5 どの程度エージェントに親近感を感じましたか
- Q1-6 あなたとエージェントの関係を最も良く表しているものを図 3 から 1 つ選択してください

実験結果を図 4 に示す。分析を行う前に正規性の確認を行った。対応のある t 検定を行った結果、エージェントの前後移動の知覚に関する Q1-3 ($t(30) = -2.178$, $p < 0.05$, $dz = -0.391$) およびエージェントへの親近感に関する Q1-5 ($t(30) = -2.136$, $p < 0.05$, $dz = -0.384$) で有意差が見られ、エージェントから感じる感情に関する Q1-4 ($t(30) = -1.938$, $p = 0.062$, $dz = -0.348$) で有意傾向が見られた。

3.3.2 エージェントの音声の変更

10 分間の休憩をはさんだ後、エージェントの音声の変更に関する実験として、3 種類の音声のエージェントとの対話を体験した。3 つの条件はランダムな順序で実施された。それぞれの条件終了後、印象を調査する以下のアンケートに回答した。これらの項目はすべて 7 段階の評価尺度で採点された。また、3 条件を比較するための自由記述欄を設けた。

- Q2-1 このエージェントはどの程度信頼できると感じましたか
- Q2-2 このエージェントに対してどの程度親近感を感じましたか
- Q2-3 このエージェントが自分の一部であると感じましたか
- Q2-4 このエージェントの声は身体の中から聞こえてくると感じましたか

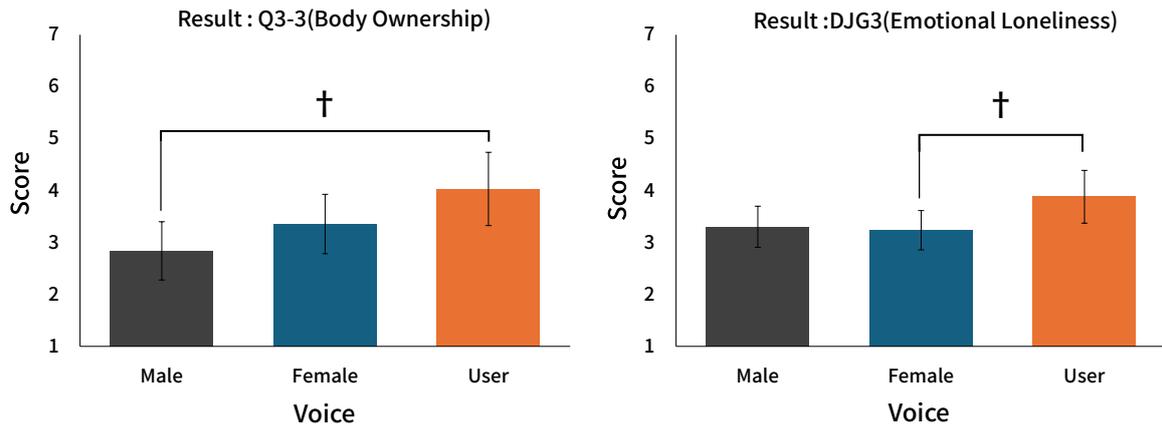


図 5: 実験 2 の結果 (平均値と標準誤差を示す † $p < 0.060$)

これに加えて、ユーザーの孤独感を測定するアンケートを行った。質問紙には、De Jong Gierveld Loneliness Scale[10]における情緒的孤独感に関する3つの項目を採用した。

実験結果を図5に示す。分析を行う前に正規性の確認を行った。一元配置反復測定分散分析を行った結果、エージェントに対する身体所有感に関するQ2-3 ($F[2, 60] = 3.738, p < 0.05, \eta_p^2 = 0.111$) および情緒的孤独感に関するDJG孤独感尺度 ($F[1, 30] = 4.500, p < 0.05, \eta_p^2 = 0.130$) で有意な主効果が認められた。事後検定では、自己類似音声において男性音声と女性音声と比較して身体所有感と情緒的孤独感がそれぞれ高い傾向を示した。

3.4 考察

スピーカーの数を増やした結果、音像の前後移動の知覚性能が向上した。これは、両耳間時間差が小さく、人間が知覚しにくい正中面の音像定位において新たに斜め方向の手がかりを追加することにより、前後の音声知覚の混乱が軽減されたためと考えられる。また、エージェントに対する親近感なども向上した。このことから、スピーカーアレイの構成を拡張することにより、体内エージェントの機能は向上したと考えられる。

エージェントの音声を変更した結果、自己類似音声において男性音声と比較してエージェントに対する身体所有感が向上した。これは、体内から自分の声に似たエージェントの音声聞こえることにより、エージェントを自分の一部としてとらえられているためと考えられる。一方で、自己類似音声を使用すると孤独感が高まる結果となった。これは、自分に似た音声と会話することで、一人で話している感覚が高まるためと考えられる。したがって、孤独感の軽減効果は、親近感や自己同一性とは比例しない可能性が示唆された。

また、評価実験の終了後に実施したユーザーインタビューでは、8chデバイスにより定位精度が向上しエー

ジェントに対する親近感が高まるという有効性が確認された一方で、現状の有線デバイスでは身動きがとりづらく、装着した状態での移動が困難であるという実用面での課題が指摘された。体内エージェントのユースケースを日常活動へと広げるためには、こうした物理的制約を解決する必要があると考えられる。

4 ワイヤレスデバイスの開発

ユーザーインタビューで指摘された物理的制約に関する課題に基づき、実用性の向上を目的とした小型ワイヤレススピーカーのシステムの設計および試作機の開発を行った。そのデバイスを図6に示す。本システムは、マイクロコントローラにESP32を採用し、モバイルバッテリー給電による無線でのウェアラブル化を実現している。これにより、有線のデバイスでは困難であった、歩行や家事などといった移動を伴う環境下での体内エージェント体験が可能となった。

本システムでは、Wi-Fiによる無線通信を利用して各スピーカーユニットを制御する構成をとっている。空間的な音像定位を実現するために、複数のスピーカーユニット間でマイクロ秒単位での再生同期が求められる。そこで本研究では、IEEE 802.11 TSF[11]に基づいたハードウェアベースでの時刻同期手法に加え、UDPブロードキャスト通信とバッファリングを用いた時刻指定による再生手法を実装し、音声出力のタイミングをマイクロ秒の精度で制御した。本手法により、ユーザーの姿勢変化や移動に依存しない安定した音像提示を実現している。現在は試作したシステムの実装および動作検証を進めており、今後はデバイスのさらなる装着性の改善を図る。これにより、体内エージェントがユーザーの生活に寄り添うパートナーとしての役割を果たすことが期待される。

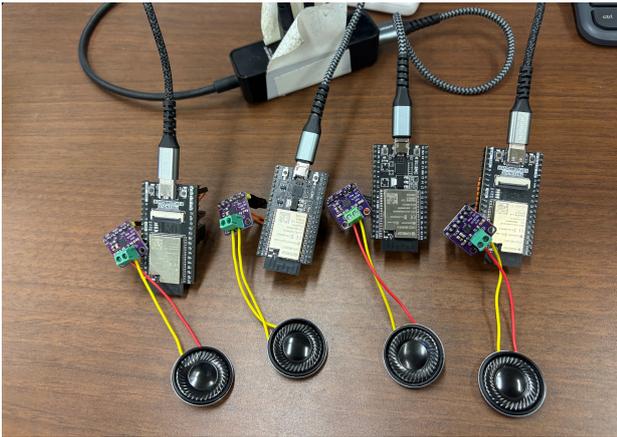


図 6: 開発したワイヤレスデバイス

5 おわりに

本研究では、体内エージェントの定位感向上を目的とした取り組みとその検証結果を報告し、現在実施しているワイヤレスデバイスの開発について報告した。今後はさらに会話への没入感を高める機能の開発に取り組んでいきながら、応用可能性を探索していく。

謝辞

本研究は JSPS 科研費 23H00484 の助成を受けて行われたものです。

参考文献

- [1] 河原歩夢, 田中文英. 体内エージェントの提案とユーザの情緒的孤独感抑制の検証. HAI シンポジウム 2024, March 2024.
- [2] E.T.Hall. *The Hidden Dimension*. Anchor, 1966.
- [3] Jonathan Mumm and Bilge Mutlu. Human-robot proxemics: Physical and psychological distancing in human-robot interaction. In *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 331–338, 2011.
- [4] Hans Wallach. The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, Vol. 27, No. 4, p. 339, 1940.
- [5] Frederic L. Wightman and Doris J. Kistler. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *The Journal of the Acoustical Society of America*, Vol. 105, No. 5, pp. 2841–2853, 05 1999.
- [6] Toshiyuki Kimura and Hiroshi Ando. [paper] 3d audio system using multiple vertical panning for large-screen multiview 3d video display. *ITE Transactions on Media Technology and Applications*, Vol. 2, No. 1, pp. 33–45, 2014.
- [7] Siqi Guo, Minsoo Choi, Dominic Kao, and Christos Mousas. Collaborating with my doppehgänger: The effects of self-similar appearance and voice of a virtual character during a jigsaw puzzle co-solving task. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, Vol. 7, pp. 1–23, 05 2024.
- [8] Tiffany D. Do, Camille Isabella Protko, and Ryan P. McMahan. Stepping into the right shoes: The effects of user-matched avatar ethnicity and gender on sense of embodiment in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, Vol. 30, No. 5, p. 2434–2443, May 2024.
- [9] Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of The Audio Engineering Society*, Vol. 45, pp. 456–466, 1997.
- [10] Jenny Gierveld and Theo van Tilburg. A 6-item scale for overall, emotional, and social loneliness confirmatory tests on survey data. *Research on Aging*, Vol. 28, , 09 2006.
- [11] Ieee standard for information technology–telecommunications and information exchange between systems - local and metropolitan area networks–specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications. *IEEE Std 802.11-2020 (Revision of IEEE Std 802.11-2016)*, pp. 1–4379, 2021.