



実環境とバーチャル環境間での アバタを用いた接客体験の違いに関する探索的調査

畑田 裕二¹⁾, 武内 一晃²⁾, 加藤 寛聡²⁾, 山崎 洋一³⁾, 鳴海 拓志¹⁾

Yuji Hatada, Kazuaki Takeuchi, Hiroaki Kato, Yoichi Yamazaki, Takuji Narumi

1) 東京大学 (〒 113-8656 東京都文京区本郷 7-3-1, {hatada, narumi}@cyber.t.u-tokyo.ac.jp)

2) 株式会社オリー研究所 (〒 103-0023 東京都中央区日本橋本町 3-8-3, {k.takeuchi, h.kato}@orylab.com)

3) 神奈川工科大学 (〒 243-0292 厚木市下荻野 1030, {yamazaki}@he.kanagawa-it.ac.jp)

概要: サービス産業では、バーチャル環境におけるバーチャルアバタのみならず、実環境におけるロボットアバタの活用も進んでいる。本研究では、ロボットアバタとバーチャルアバタの体験の質を比較するために、「分身ロボットカフェ DAWN ver. β」にてロボットアバタを通じた遠隔就労に従事しているユーザ 11 名を対象として、可能な限り実環境を模倣したバーチャル環境において接客実験を実施し、インタビュー調査を行った。

キーワード: テレプレゼンスロボット, アバタ, 接客, オンラインコミュニケーション

1. はじめに

ユーザが自己の身体だと認識し、ある環境とインタラクションするために用いられる表象はアバタと呼ばれる。オペレータが遠隔地からアバタを操作する遠隔接客は、サービス産業の現場を中心に盛んに実証実験が進んでいる。COVID-19 の世界的流行に伴うリモートワークの推進は、こうしたアバタ活用を後押ししている。3DCG で描画されるバーチャルアバタは、実環境に設置された 2D モニタ¹²⁾や、客側もアクセス可能なバーチャル環境に構築された店舗³⁴⁾を通じて提示される。バーチャルアバタのみならず、物質的実体を持ち実環境で利用される遠隔操作ロボットもアバタの一種だと捉えることができる [1]。ロボットアバタは、テレプレゼンス技術の応用として、近年ではカフェの接客 [2] や商品販促を行う店員などに活用されている。例えば、東京都日本橋に 2021 年 6 月にオープンした「分身ロボットカフェ DAWN ver. β」(以下、分身ロボットカフェ)では、パイロットが遠隔地から卓上のロボット「OriHime」を操作して注文を取り、配膳ロボット「OriHime-D」を操作して料理を提供するという実証実験が行われている。

バーチャルアバタとロボットアバタという、メディアの異なる二つのアバタ体験の間には多くの共通点がある [1]。両者を同一のフレームワークで扱うことができれば、実環境とバーチャル環境を自在に行き来し、時空間に縛られない社会参画を可能にする応用が可能になる。バーチャルリアリティの意味が「みかけや形は原物そのものではないが、本質的あるいは効果としては現実であり原物であること」で

あるならば、バーチャルアバタを通じた接客は、理想的には対面におけるそれと等価な体験になるはずである。しかし技術的制約から、これらの体験が実際に等価になるとは限らない。本研究では、実環境とバーチャル環境それぞれにおける接客体験の質を比較することで、現行のテレプレゼンスおよび VR 技術の限界を明らかにし、両者を同一のフレームワークで扱うための指針を検討する。

2. 方法

実験は、2022 年 5 月 23 日から 6 月 2 日にかけて、4 日間に分けて行われた。各実験日において午前中に 1 セッション、午後に 2 セッションが実施された。各セッションには、分身ロボットカフェにてロボットアバタを通じた遠隔就労に従事しているユーザ (パイロット) が 1 名ずつ、合計 11 名が参加した。各パイロットに対して 1 名または 2 名の非パイロット参加者が客役として割り当てられた。非没入型のバーチャル環境において接客実験が実施された後、普段の業務体験との違いに関するインタビュー調査を行った。

2.1 実験環境

本実験は、分身ロボットカフェの 2021 年 6 月当時の内装 (図 1b を再現したバーチャル環境「バーチャル分身ロボットカフェ」(図 1c,d) を用いて行われた。また「アバタ」と明示されていない OriHime は、実環境におけるロボットアバタを表す。以下では、日本橋に実在する場所としてのカフェを「分身ロボットカフェ」、バーチャル分身ロボットカフェを「V カフェ」と表記する。V カフェは、オープンから本実験が行われるまでの一ヶ月程度の期間、控室としてパイロット同士の業務外交流に用いられていた。V カフェは、

¹<https://prtimes.jp/main/html/rd/p/000000651.000009215.html>

²<https://prtimes.jp/main/html/rd/p/000001071.000016751.html>

³<https://prtimes.jp/main/html/rd/p/000000032.000033579.html>

⁴<https://group.ntt.jp/newsrelease/2022/02/02/220202a.html>



図 1: (a) OriHime の操作画面. (b) 実際の分身ロボットカフェ. (c) 実験に用いた V カフェ. (d) 接客の様子.

表 1: 店員役の参加者 (パイロット)

	性別	年齢	OriHime を 使い始めた時期	V カフェ 利用経験
P1	女	40 代	2021 年夏	日常利用
P2	男	50 代	2021 年夏	日常利用
P3	男	50 代	2021 年夏	日常利用
P4	男	20 代	2018 年秋	日常利用
P5	女	50 代	2021 年夏	日常利用
P6	女	40 代	2021 年夏	日常利用
P7	女	30 代	2018 年秋	数回
P8	女	20 代	2019 年秋	なし
P9	女	50 代	2019 年秋	なし
P10	女	40 代	2021 年夏	日常利用
P11	女	40 代	2021 年春	日常利用

monoAI technology 株式会社が提供する「XR CLOUD⁵」上に構築されている。XR CLOUD は、大規模なオンラインイベント開催を実現するプラットフォームで、双方向での音声コミュニケーション機能を備えている。参加者はスマートフォン、タブレット、PC でログインすることができる。ユーザは人型 4 種類、OriHime-D (図 1c)、OriHime の合計 6 種類をアバタとして使用することができ、参加中にいつでも変更が可能であった。人型アバタはジャンプやアクション (感情表現などに用いるアバタのアニメーション) が可能だが、OriHime-D 及び OriHime のアバタではいずれも不可能 (ボタンを押しても何も起きない) であった。なお、OriHime アバタは OriHime と同様、空間内を移動することができない。V カフェ内に設置されているスクリーンには、パイロットが普段の接客で使用しているのと同じ内容の PDF ファイル (パイロットの自己紹介やフードメニューなど含まれる) が投影された。

2.2 参加者

分身ロボットカフェに勤務している 11 人のパイロット (表 1) が店員役として、22 名 (女性 2 名) の大学生または大学院生が客役として本実験に一度または二度参加した。3 名の客役参加者は分身ロボットカフェに一度だけ行ったことがあると回答し、残り 19 名は行ったことがないと回答した。

2.3 手順

店員役のパイロットは、30 分間の接客セッションを 2 度連続して行った後、V カフェ内に残って 30 分程度の半構造化インタビューに答えた。インタビューでは、「バーチャル環境と実環境の間で接客体験の違いを生み出す要因は何か」というリサーチクエスション (RQ) のもとで、普段のロボッ

トアバタを使用した就業との比較に基づく本実験の感想についての質問がなされた。実験は直前の打ち合わせや 10 分間の休憩を合わせて、約 110 分間に渡って行われた。実験の様子は、参加者の同意のもとで、実験環境に同席した実験実施者が録画・録音した。客役の参加者 (以下、客) は、10 分間の実験説明を受けた後、約 30 分間の接客セッションに参加した。客は実験開始直前に、パイロットの案内に従いながら、自由に振る舞うよう実験実施者から指示された。

2.4 解析

半構造化インタビューの録音データから作成された簡易的な逐語録に対して、テーマティック・アナリシス [3] が行われた。まず筆頭著者が逐語録を読み込み、帰納的アプローチで予備的なテーマを作成した。その後、RQ に照らし合わせて演繹的アプローチでテーマをまとめた。各段階において、全ての著者を交えて結果の妥当性について議論した。

3. 結果と考察

以下では、インタビュー中の語りを括弧付で引用しながら、得られた四つのテーマを順に述べる。

3.1 受け取る情報の質と量

3.1.1 視線・表情の欠如

V カフェには、バーチャルアバタの視線や表情を制御する機能が実装されていなかった。多くのパイロットは、視線や表情といった社会的手がかりが欠如したコミュニケーションにおいて、声の調子のみから客の心理状態を推測することを強いられており、客が話し始めるタイミングや、現在の話題に対する興味関心の度合いを把握することが難しかったと述べた。人はコミュニケーションに際して知覚と行動に関する様々な予測を無意識下で行っており、その予測から外れる結果に晒されることはストレスにつながると考えられる。例えば、実環境で実践されているのは知覚的に異なるコミュニケーションが疲労や認知負荷をもたらすことは、「ZOOM 疲れ [4]」に関する研究でも指摘されている。他方、P4 が「OriHime だと相手の目を見るように気をつけているけど、今日はそれがなかったので話に集中できなかった」と語ったように、参加者全員の視線情報が欠如している環境では、「相手の目を見る」という行為に認知資源を割く必要がなく、かえって会話に集中できるという側面も見出された。

3.1.2 高品質な音声情報

バーチャル環境での接客体験の優位点として、会話の音質の高さと遅延の少なさが挙げられた。分身ロボットカフェ

⁵<https://xrcloud.jp/>

での接客では、隣のテーブルの声がマイクに拾われることによって、接客に集中できない場面もある。P4はVカフェにおける高品質な音声情報について「お客さんとの会話がメインなこの仕事において、聞こえづらいというのはかなりストレス。それがないだけでバーチャルカフェの評価は高くなる」と語った。本研究ではバーチャル環境の中に一組の客しかいない状況が設定されたため、賑やかな時間帯の実環境との単純な比較はできないが、バーチャル環境ではユーザごとの声の大きさを調節できるため、ノイズが少ないという利点は保たれると考えられる。

3.1.3 素性の知れなさ

パイロットの1名(P5)が、バーチャルアバタと対面する体験を「相手の本当の姿がわからないので不安」と評したことから、現在のアバタには、その人の過去の経験やその結果構築された価値観を推定できる外見的特徴が実身体に比べて少ないことが示唆された。P5は分身ロボットカフェでの普段の接客において、客の外見から様々な推測をしていた。肉体、髪型や服、所持品といった外見的情報は、その人が生きてきた文化や社会、その結果構築された価値観を彷彿とさせる。「どんな姿で、どんな話し方をするのか、どういう髪型をしているのか、どういう服を着ているのか、どういう小物を持っているのか、そういうので人となりが見えてくる。多分こういう雑誌読んでるんだろうな、とか。(P5)」。

これに対して、本実験で用いたアバタはいずれもサービスによって画一的に提供されているものであった。ソーシャルVRの中には、アバタの細やかなカスタマイズが可能なものもあるが、しかし、アバタの外見の表現はまだ、実社会ほどの多様性がない場合や、外見が持つ文脈（歴史的意味や流行、ブランド価値など）が乏しい場合も多い。

3.1.4 情報不足を補うコツ

社会的手がかりのなさが円滑なコミュニケーションを阻害したという訴えがあった一方で、何名かのパイロットはそれを問題としなかった。そうしたパイロットは「相手の状態を想像で補った(P2)」と語り、「情報がこないものは考えても仕方ない(P9)」と潔く割り切る姿勢が見られた。またP10はVカフェでのコミュニケーションを「電話している感覚」「音声だけの会話をしている感じ」と表現し、音声以上の情報を期待していないために「情報がなくても気にならない」と語った。さらに、Vカフェでのコミュニケーションにあまり支障を感じなかったパイロットは、自ら相手の状態を探るために積極的な質問を心がけていた。「相手のことは気にするけど、情報がこないものは考えても仕方ない。だからせめて、返事で状況がわかるように質問を投げかけたりした。(P8)」。

3.2 自己表現

3.2.1 能動的な移動

分身ロボットカフェでの接客は、主に卓上に置かれたOriHimeを通して行われる。標準的なOriHimeは車輪などの移動機構を持たないため、接客はOriHimeが設置されている定点で完結するよう行われる。一方で本実験では、OriHime-

Dアバタや人型アバタの持つ移動能力が積極的に活用された。例えば、Vカフェの入り口からテーブルまで客と共に歩いていくことに始まり、Vカフェ内の他エリア（トイレやグッズ販売エリアなど）を巡るツアーや、分身ロボットカフェでの配膳の模擬的な実演などが行われていた。パイロットは、移動を活用して（現地のスタッフではなく）自ら客の案内ができることを喜んでいて、「現実だと物販コーナーを説明する時は（中略）一応顔の向きを変えてあっちですと言ったりするけど、本当だったらそこまで連れて行けたらいいと思う。それができたのはよかった。(P4)」

3.2.2 モーションの役割

OriHimeを用いた接客では、OriHimeが標準的に備えているジェスチャーモーション（「バタバタ」「困った」といったラベルで呼ばれる）が活用されている。これらのモーションは、感情を表現するだけでなく、会話の随所に挿入されることで会話の間を持たせ、コミュニケーションを円滑にする役割も持っている。しかし本実験において、アバタのモーションはあまり活用されなかった。これは、Vカフェのモーションシステムにパイロットが不慣れであった（「普段使っているモーションが使えず(P11)」）のみならず、Vカフェのアバタが有するモーションがいずれも「大袈裟すぎる(P8)」のものであり、汎用性に欠けているためだと考えられる。Vカフェで採用されているアバタのモーションは「はっきりとしたメッセージを出す時しか使えない(P6)」一方で、OriHimeが有する「バタバタ」や「困った」といったジェスチャーモーションは「可愛い動作(P2)」で「どんな時でも使えるような曖昧(P8)」さを持っていることが指摘された。会話中のモーション活用が乏しかったことで、パイロットは明示的な言語表現のみでコミュニケーションすることを強いられ、不便さを訴えていた。

アバタモーションの曖昧さは、モーションの内容ではなく、そのモーションを表現するアバタの自由度のいずれによってもたらされている可能性がある。人型アバタがOriHimeと異なり多くの自由度を持ち、動きのディティールを豊かに表現できることは、Vカフェアバタのモーションが大袈裟に感じられたことと関係しているかもしれない。

3.2.3 操作の難しさ

モーション以外に、操作上の難しさが主に三つ挙げられた。それは、客と視線を合わせること、カメラの位置と角度を適切に調節すること、アバタを適切に移動させることである。こうした操作が難しく感じられる理由には、パイロットが操作インタフェースに慣れていないことだけでなく、本来緊密に連携した複合的動作として行われている行為が、インタフェースによって個々に分離され、独立した意思決定や操作が必要になっていることも考えられる。例えば、発話すること、ジェスチャーをすること、相手の目を見ることは、実環境においてはそれぞれ独立した意思決定に基づいて表出されるというより、ある一つの目的（例えば相手の話を聴いているというメッセージを伝えるなど）に基づいて、同時連携的に表出されるといった方が適切である。

これに対してバーチャル環境では、こうした一つ一つの動作がボタンの押下などの個別要素へと還元され、逐一独立した意思決定や操作が必要になる。

3.3 身体と環境のフレーミング

3.3.1 自己と他者の対等さ

分身ロボットカフェでの接客では、パイロット以外の人は現地に居合わせている場合が多い。オンライン参加が少数派となるコミュニケーションの場では、パイロットは通信環境の問題によって「自分だけ視界が悪くなる (P9)」、音声品質の問題によって「自分だけ会話を聞き取れない (P9)」といった問題に直面するため、「かえて孤独を感じる (P11)」こともある。さらに OriHime の機能的制約ゆえに、パイロットは接客を遂行するために現地のスタッフのサポートを必要とする。V カフェでの接客は、OriHime で現地の客と接するのと異なり、全員がオンラインという同一の場に集まるために対等なコミュニケーションが取れ、精神的に楽だったという感想があった。例えば P9 は「お客さん側にも (マイクがミュートになっていて聞こえないなどの) トラブルが起きていて、逆に自分が大丈夫だったのが新鮮」だと語った。

3.3.2 プレゼンスとリアリズム

多くのパイロットが、OriHime を用いた接客では「家について喋っていることを忘れ (P9)」、(自宅ではなく) あたかもその場所にいるかのような感覚 (プレゼンス) が強いと述べた。その一方で V カフェでの接客は、「自分は自宅のペットにいて、パソコン越しの世界をのぞいている感覚 (P9)」や「ネットで働いている感覚 (P8)」だとされた。さらに、V カフェは実環境とは地続きにはない、(見かけは似ているが) 分身ロボットカフェとは質的に異なる空間だと指摘するパイロットもいた。こうしたプレゼンスやリアリズムの差には、ノイズと捉えられがちな微細な環境刺激の有無が寄与していることが示唆された。例えば P7 は「光や音はもちろん、誰かの悪口を聞いちゃうとか、誰かが顔を顰めるとか、OriHime はそういうノイズが入ってくるけど、外出ができない時はそういうリアリティを求めている自分がいる」と語った。

3.4 アバタの自己同一化

3.4.1 インタラクションを通じた自己知覚

OriHime を自分そのものだと感じる自己同一化には、他者や環境との相互作用が重要な働きをしていると思われる。特に「お客さんが自分をまなざしてくれる (目線が合う) (P10)」こと、そこにモノではなく人が実在するのと同じやり方で扱われる (「自分みたく扱ってくれる (P7)») ことは、プレゼンスのみならずアバタの自己同一化を強く促進することが示唆された。先行研究では、テレプレゼンスロボットを人として捉えているか、単なる道具として捉えているかによって、人はそのロボットの扱い方が変わることが指摘されている [5]。こうした行動変容は、行為の受け手が他者反応を通じて自己同一化するプロセスにも寄与すると考えられる。OriHime はタブレット端末とは異なり人型の表象であるため、周囲の人から人として扱われやすい。

他方で V カフェでは、人型アバタを使用したパイロットも多くいたものの、そこに人が実在するかのようインタラククションが不足していたと考えられる。

他にも、環境に働きかけようと立てた行動計画が预期通りに実行されること、即ち行為主体感が担保されることに言及したパイロットもいた。特に主体感の伴った行為が他者に観測され、受容された際に自己同一化が起こることが示唆された。例えば P3 は「普段の OriHime だと会話の中でちゃんと頷いているけれど、バーチャルだとそれができないので、もう少し手を加えないと自分になりきれていないという感じがする」と語った。

3.4.2 自己同一化しやすいアバタ

自己同一化を促すインタラククション以外の要因として、アバタの長期的な利用とカスタマイズが挙げられた。またアバタの自己同一化に寄与する要因の重みづけは、身体や性格などの個人特性によって変化する可能性がある (「体が動かない方にとっては OriHime は分身だと思うけど、私は肢体不自由ではないので、フォーカスしている部分が違うかもしれない (P7)»)。

4. おわりに

本研究では、実環境に置かれたテレプレゼンスロボットを通じて遠隔で接客に従事するパイロットが、職場によく似たバーチャル環境においてバーチャルアバタを通じた接客体験を行った。インタビューを通じて、実環境 (ロボットアバタ) とバーチャル環境 (バーチャルアバタ) それぞれの体験がどのように異なるのかを明らかにした。

謝辞 本研究の一部は、JST ムーンショット型研究開発事業 (JPMJMS2013) および科研費特別研究員奨励費 (22J14271) から支援を受けて行われた。

参考文献

- [1] L. Aymerich-Franch. Towards a common framework for mediated embodiment. *Digital Psychology*, Vol. 1, No. 2, pp. 3–12, Nov. 2020.
- [2] K. Takeuchi, Y. Yamazaki, and K. Yoshifuji. Avatar work: Telerwork for disabled people unable to go outside by using avatar robots. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '20, pp. 53–60, 2020.
- [3] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, Vol. 3, No. 2, pp. 77–101, 2006.
- [4] J. N. Bailenson. Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue. *Technology, Mind, and Behavior*, Vol. 2, No. 1, 2021.
- [5] L. Takayama and J. Go. Mixing metaphors in mobile remote presence. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pp. 495–504, 2012.