



視点取得可能な実写映像に基づく 没入型テレコミュニケーションに関する検証

宋戸 海¹⁾, 謝 淳²⁾, 井上 雅彦³⁾, 北原 格²⁾

1) 筑波大学大学院エンパワーメント情報学学位プログラム（〒305-8573 茨城県つくば市天王台 1-1-1）

2) 筑波大学計算科学研究センター（〒305-8577 茨城県つくば市天王台 1-1-1）

3) 鳥取大学医学系研究科（〒683-8503 鳥取県米子市西町 86 番地）

概要：一般的なビデオコミュニケーションでは、対話相手を外側から撮影した有限視野角の映像が提示されるため、通信相手の身体的状況や視線・意図・関心といった非言語的情報を的確に把握することが困難である。本研究では、複数のカメラとヘッドマウントディスプレイ（HMD）を組み合わせ、“視野制限のない映像”と“相手視点の映像”を双方向提示する没入型テレコミュニケーションシステムを提案する。実証実験を通じて、従来手法と比べて相手の感情や意図の理解が促進されることが示され、遠隔共同作業や指導場面への応用可能性を確認する。

キーワード：他者視点取得、テレコミュニケーション、没入型映像提示、ユーザインタフェース

1. はじめに

情報通信技術の発展により、テレコミュニケーションの需要が増加している[1]。中でも Zoom や Microsoft Teams, Google Meet などのビデオ会議システムを利用したビデオコミュニケーションは、我々の生活に欠かせないツールとなりつつある。ビデオコミュニケーションでユーザが受け取る情報は、パソコンに接続したカメラとマイクで取得した音声と単一視点で観察される、視野角が限定された映像である。その結果、相手の身振りや手振り、周囲の状況を把握することは困難であり、また、コミュニケーションにとって重要な相手の意図や感情、関心を理解することが難しい状況が発生する。

本稿では、ビデオコミュニケーションの情報送信側を主体、情報受信側を客体と定義し、以下の問題について検討する。なお、主体、客体の立場を入れ替えた場合も同様の問題が起こり得る。

問題 1：視野角が有限な映像では、主体が客体の意図や感情を理解することが困難である。

問題 2：単一視点映像では、主体が客体の関心を理解することが困難である。

上述した問題の解消を目的として、360 度映像をリモートユーザに共有する Jackin Head[2]や、二人のリモートユーザが二人のローカルユーザの一人称視点や三人称視点からの観察を可能とする Jackin Space[3]など、周辺状況の共通理解支援に関する研究が行われている。しかし、これらの研究では、ビデオコミュニケーションのような双方向

通信を対象としていない。本稿では、主体と客体が互いの様子を観察可能な双方向通信への導入を想定し、その実現法を提案するとともに、有効性を検証する。具体的には、360 度映像とコミュニケーション相手の視点からの映像の観察が可能なテレコミュニケーション方式を提案し、その実現法について述べる。相手の視点を取得するアプローチは、心理学における他者の思考感情・視点を理解すること、または、その能力の意を表す「他者視点取得」から着想を得ている。他者の体験を疑似的に得ることによって、相手に対する理解を深め、相互理解の実現が期待される。

2. 関連研究

2.1 VR テレコミュニケーション

Misawa ら[4]は、リモートユーザの顔を表示するディスプレイを代理人が装着することで社会的・身体的存在感を提示する方式を提案し、代理人が本人とみなされやすく、テレコミュニケーションシステムとして有効であることを確認している。Cai ら[5]は、リモートユーザが肩に装着した 360 度カメラの映像をローカルユーザが取得し、ローカルユーザのハンドジェスチャーを AR グラスを介してリモートユーザが確認するシステムを提案した。両ユーザの同位置感覚が向上し、円滑な遠隔コミュニケーションが実現されていることを示している。笠原ら[6]は、指示役（Ghost）が AR グラスとカメラを装着した人物（Body）からの一人称視点映像を観察しながら映像の任意の一部

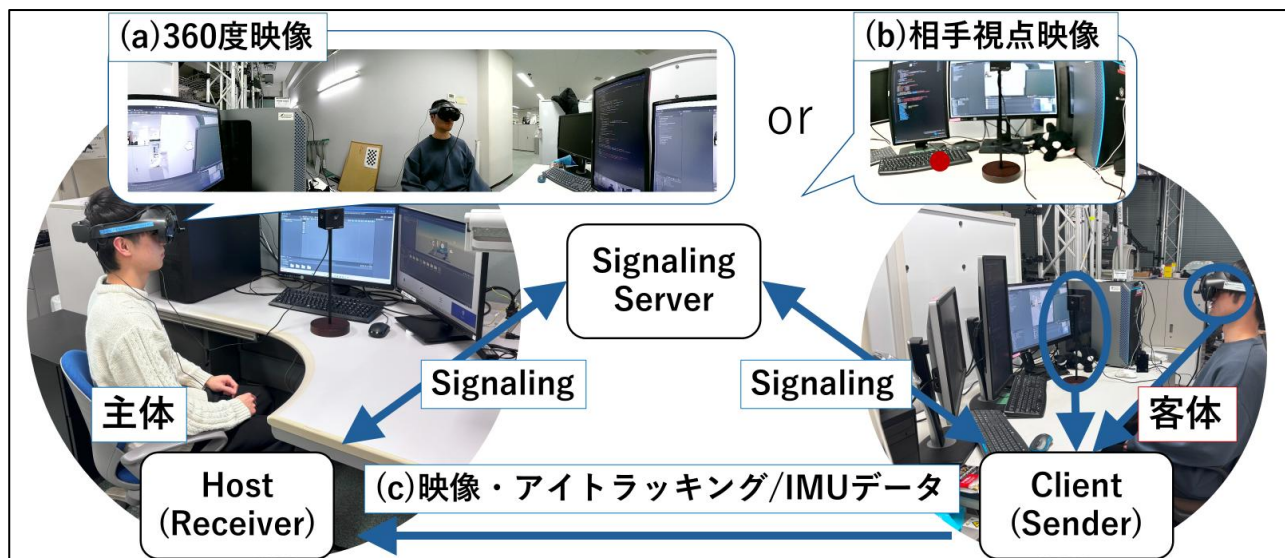


図 1 視点取得可能な実写映像に基づく没入型テレコミュニケーションシステムの概要図

を差し示す様子を Body の AR グラスに可視化することが、遠隔ナビゲーションとして有効であることを確認した。

ユーザ同士が交流可能な Social VR が普及しつつある。代表的な例の VRChat[7]では、ユーザはアバターを介して音声やボディランゲージで他者とコミュニケーションを取ることができる。Social VR における社会的相互作用に関する研究も活発に行われており、VR 環境を活用した新たなコミュニケーション手法が模索されている。

このように、VR 技術を用いて物理的な距離による制約を克服し、人々の効果的な相互作用が可能な手法の構築は遠隔コミュニケーションの重要な研究課題である。

2.2 他者視点取得

Galinsky ら[8]は、他者視点取得が対象への好意的評価や集団間バイアスの低減に効果的であることを示した。年配男性の生活についてエッセイを書く課題では、「その人物になったつもりで書くように」と指示されたグループが、他のグループに比べてステレオタイプの記述が少なく、より肯定的な評価を下した。また別の実験では、「過小評価された人物の視点に立って書く」よう指示されたグループは、内集団と外集団を同等に評価したが、他のグループでは内集団を明確に優遇する傾向が見られた。これらの結果から、他者視点取得は、対象者を好意的に評価すると同時に、集団間のバイアスも減少させることが実証されている。

このことから、他者視点取得は社会的相互作用における相手との関係の改善・理解に役立ち、より効果的なコミュニケーションを実現することが期待できる。

2.3 身体所有感と共感

Peak ら[9]は、白人の参加者が HMD を用いて、自身の動きと同期する白人および黒人アバターを体験した際、黒人アバターを用いた場合に人種の偏見が有意に低下することを示した。

西田ら[10]は、子どもの認識を理解する手段として、腰のカメラ映像を HMD に映し出し、子どもサイズの外骨格

を装着する体験を設計した。これにより、参加者は周囲を大きく捉え、圧迫感や恐怖といった心理的变化を示した。

また、Collaço Oliveira ら[11]は、パフォーマの視点映像をユーザが HMD で視聴し、パフォーマがユーザの動きを模倣する状況を通じて、ユーザに身体所有感や親近感が生じることを示した。

このように、他者の視点や身体感覚を共有することは、他者理解や共感の促進に寄与する。

3. 視点取得可能な没入型テレコミュニケーション

図 1 に、視点取得が可能な没入型テレコミュニケーションシステムの構成を示す。(a) は客体空間に設置した 360 度カメラ映像、(b) は客体の HMD 側面に取り付けたカメラ映像を主体が HMD で確認する構成であり、主体はコントローラーによってこれらを切り替え可能である。(c) では、カメラ映像とアイトラッキング、IMU データを送受信し、注目領域の可視化や VR 酔いの低減を目的にホモグラフィ変換を適用する。さらに、主体と客体の立場を入れ替えることで、双方向のコミュニケーションが可能となる。

4. 没入型テレコミュニケーション

4.1 360 度カメラとマウントカメラ映像の取得

VR 空間に 360 度カメラとマウントカメラで撮影した映像を取り込み、HMD を用いて観察する。HMD の装着者は、頭を上下左右に回転させることで 360 度映像を観察する。

4.2 カメラ映像の送信

1 節で述べた問題に対処するために、二つの映像撮影・送信方法を用いる。一つは問題 1 に対応するため視野角に制限のない映像を撮影・送信する方法であり、もう一つは問題 2 に対応するため、客体の注目領域を取得する方法である。図 1 は映像とデータの送信の様子を示したものである。主体が客体空間の 360 度映像を取得したときの様子を図 1 (a)、主体が客体の視点を取得したときの様子を図 1 (b) に示す。主体が使用しているデスクトップパソコン

を Host、客体が使用しているデスクトップパソコンを Client と定義すると、客体のマウントカメラの映像は WebRTC によって Client から Host に送信され、主体の HMD に提示される。

4.3 視線情報の可視化

主体が客体の視点取得時において、客体がマウントカメラ映像中のどこに注目しているかを可視化するため、客体の視線情報をソケット通信によって Client から Host に送信する。このとき、主体が確認する映像において客体の注目位置を可視化する。

4.4 視点取得時のホモグラフィ変換

視点取得時には、客体の頭部の動きが主体の映像に直接反映されるため、主体が VR 酔いを引き起こす可能性がある。この問題に対応するため、本システムでは主体が視聴する映像の視覚的な安定性を確保し VR 酔いの軽減を図る、客体が装着する HMD に内蔵された IMU から得られるデータをソケット通信を通じて Client から Host に送信する。Host 側ではこの IMU データに基づいて、主体に提示される視点取得映像にホモグラフィ変換を適用する。

5. 評価実験

5.1 実験 1：相手の意図や感情の理解度に関する比較実験

実験 1 の目的は、視野角の制限解除が主体による客体の意図や感情の理解に対して有効であるかの検証である。

5.1.1 手順

客体が、表 1 に示す特定の感情と関連があるとされる 5 種類のボディランゲージ[12][13][14]を行う様子を、客体の肩から上が映った映像（従来のビデオコミュニケーションを想定したもの）として撮影する。次に、本システムを用いて、客体が 5 種類のボディランゲージを行う様子を撮影する。主体は両映像群を観察し客体がどのような感情であったかを推測する。また、主体は両映像に対して、客体のボディランゲージの観察のしやすさについて評価する。

5.1.2 評価方法

客体の各ボディランゲージにおいて、主体は客体において各感情がどの程度含まれているか 7 段階（-3=「とてもそう思わない」~3=「とてもそう思う」）で評価する。また、従来のビデオコミュニケーションと比較して、本システムでは客体のボディランゲージが見えやすくなったかについて 7 段階で評価する。

表 1 ボディランゲージと感情の対応

ボディランゲージ	感情
腕を組む	警戒・拒絶
手のひらを見せる	正直・誠実
袖口を触る	動揺・不安
足を小刻みに震わせる	苛立・焦燥
猫背・動作が遅い	悲しみ

5.1.3 結果

実験参加者 15 名に対し、実験 1 を行った結果を図 2 に示す。従来のビデオコミュニケーションと本システムを比

較したところ、従来方式では客体が意図した感情が主体に十分に伝わらず、回答にばらつきが見られた、一方で本システムを用いた場合では、客体の意図する感情がより正確に主体に認識された。さらに、ボディランゲージの視認性に関する評価においても、従来方式よりも提案システムの方が明確に観察しやすいと認識された。

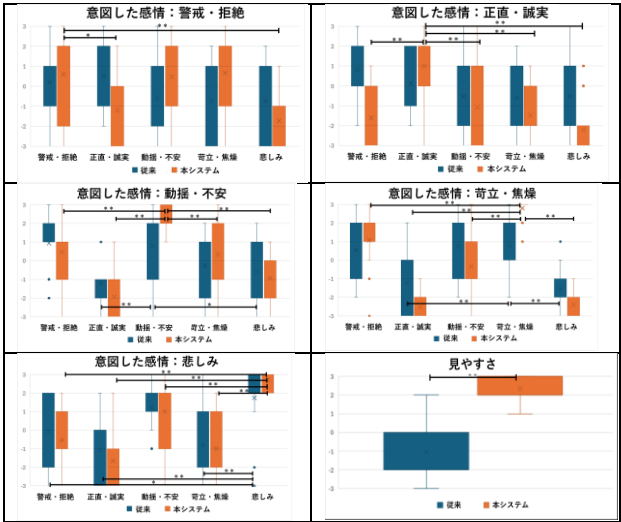


図 2 主体が意図した感情と客体の回答した感情：意図した感情は、警戒／拒絶（左上）、正直／誠実（右上）、動揺／不安（中央左）、苛立／焦燥（中央右）、悲しみ（左下）である。右下の図はボディランゲージの読み取りやすさを比較したものである。（*）は有意水準 5%、（**）は有意水準 1%で統計的に有意であることを示す。

5.1.4 考察

実験 1 の結果に基づき、従来のビデオコミュニケーションと本システムの両条件において、各意図された感情と他の感情群との比較を行うため、分散分析（ANOVA）を実施した。図 2 に示すように、有意な差が認められた場合には、*（ $p < 0.05$ ）、**（ $p < 0.01$ ）で示している。全体として、本システムは従来のビデオコミュニケーションに比べて、感情の伝達においてより効果的であることが確認された。ただし、「警戒／拒否」と「動揺／不安」および「警戒／拒否」と「苛立／焦燥」の間には有意差が認められなかった。これらの感情が類似していることに加え、HMD の装着により目元周辺の表情が遮られたことが、識別の困難さに影響を与えた可能性があると考えられる。

5.2 実験 2：相手の関心の対象を理解可能か確かめる実験

実験 2 の目的は、主体による客体の視点取得が、客体の注目領域の理解に対して有効であるかの検証である。

5.2.1 手順

客体はデスクに置かれた五つの注目物体を順に注視する。その際、主体は客体の視点取得をし、客体の視線情報が可視化された状況下で、客体は何を見ていたかをその都度回答する。一人の実験参加者（主体）に対し、注視物体を変えながら同様の質問を 5 回実施する。

5.2.2 評価方法

実験2では、主体が客体の注視対象を正しく回答できたかどうかを基準に正誤を判定し、正答数をもって評価指標とした。

5.2.3 結果

実験2では、全75件の回答のうち、すべてが正答であった（正答率100%）。この結果は、すべての実験参加者が客体の注視対象を正確に特定できたことを示している。

5.2.4 考察

本実験では、主体が全75問に正答したことから、視点取得によって客体の注目対象を高精度で把握できることが示された。特に、本システムでは客体の注視点が可視化されており、これにより主体は対象物との空間的関係を直感的に理解できたと考えられる。また、視点取得のみで赤点を表示しなかった条件については、本研究の先行実験[15]において、全48試行中46問の正答（正答率約95.8%）が得られている。この結果は、視線の可視化によって主体の認知的負荷が軽減され、理解精度が向上する可能性を示唆している。

6. おわりに

本稿では、HMD・360度カメラ・マウントカメラを用いた没入型テレコミュニケーション手法を提案し、従来のビデオコミュニケーションが抱える視野制限や視点の単一性といった課題に対処した。従来のシステムでは、相手の意図・感情・関心の焦点を把握することが困難であったが、提案システムでは、被験者が対象者のボディランゲージを有意に観察しやすいと感じ、感情の解釈精度も向上した。さらに、視点取得に関する実験では、実験参加者が客体の注視領域を正確に特定できることが確認された。これらの結果は、本システムが相手の感情状態および注意の向きを理解する能力を高めることを示唆している。本システムは遠隔協働や教育支援などへの応用が期待される。

参考文献

- [1]. DeFilippis, E., Impink, S. M., Singell, M., Polzer, J. T., and Sadun, R., "The impact of COVID-19 on digital communication patterns," *Humanities and Social Sciences Communications*, vol.9, Article 180. May 2022.
- [2]. Kasahara, S., and Rekimoto, J. "JackIn head: Immersive visual telepresence system with omnidirectional wearable camera for remote collaboration," *VRST '15: Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pp.217-225, November 2015.
- [3]. komiyama, R., Miyaki, T., Rekimoto, J., JackIn space: designing a seamless transition between first and third person view for effective telepresence collaborations," *AH '17: Proceedings of the 8th Augmented Human International Conference*, pp.1-9, March 2017.
- [4]. Misawa, K., Rekimoto, J., "ChameleonMask: Embodied Physical and Social Telepresence using Human Surrogates," *CHI EA '15: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pp.401-411, April 2015.
- [5]. Cai, M., Tanaka, J., "Mixed-Reality Communication System Providing Shoulder-to-shoulder Collaboration," *International Journal On Advances in Software*, vol.12, no.3-4, December 2019.
- [6]. Kasahara, S., Rekimoto, J., "JackIn: integrating first-person view with out-of-body vision generation for human-human augmentation," *AH '14: Proceedings of the 5th Augmented Human International Conference*, Article no.46, pp.1-8, March 2014.
- [7]. "VRChat," <https://hello.vrchat.com/>, last accessed 2025/07/08.
- [8]. Galinsky, A. D., and Moskowitz, G. B. "Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism," *Journal of Personality and Social Psychology*, vol.78, issue 4, pp.708-724, May 2000.
- [9]. Tabitha, C. Peck, Sofia Seinfeld, Salvatore M. Aglioti, Mel Slater, "Putting yourself in the skin of a black avatar reduces implicit racial bias," *Consciousness and Cognition*, vol.22, Issue3, pp.779-787. September 2013.
- [10]. Nishida, J., Takatori, H., Sato, K., Suzuki, K., "CHILDHOOD: Wearable Suit for Augmented Child Experience," *VRIC '15: Proceedings of the 2015 Virtual Reality International Conference*, no.22. pp.1-4, April 2015.
- [11]. Oliveira, E. C., Bertrand, P., Lesur, M. R., Palomo, P., Demarzo, M., Cebolla, A., Baños, R. M., & Tori, R. "Virtual Body Swap: A New Feasible Tool to Be Explored in Health and Education," *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*, pp.81-89, June 2016.
- [12]. "Science of People: 23 Essential Body Language Examples and Their Meanings." Archived at <https://web.archive.org/web/20250125185533/https://www.scienceofpeople.com/body-language-examples/>, last accessed 2025/07/08.
- [13]. Dael, N., Mortillaro, M., Scherer, K. R.: Emotion expression in body action and posture. *Emotion* 12(5), 1085–1101 (2012).
- [14]. Bindemann, M., Burton, A. M., Langton, S. R. H.: How do eye gaze and facial expression interact? *Visual Cognition* 16(6), 708–733 (2008)
- [15]. 穴戸 海, 謝 淳, 井上 雅彦, 北原 格: 相手視点を取得可能な実写ベースの没入型テレコミュニケーション方式. *メディアエクスペリエンス・バーチャル環境基礎研究会 (MVE 2024)*.