



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方式の検討

宍戸海¹⁾, 謝淳²⁾, 井上雅彦³⁾, 北原格²⁾

1) 筑波大学大学院システム情報工学研究群 (〒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カメラの視野角が限定され、通信から得られる情報は限定的である。そのため、通信相手の状況や意図、関心を理解することが困難である。そこで、複数台のカメラとヘッドマウントディスプレイ (HMD) を用いて、視野角が限定されない映像の生成、および相手の視点映像を取得することが可能な没入型テレコミュニケーションシステムの構築に取り組んだ。このシステムは遠隔での共同作業や指導活動に応用可能であると考えられる。

キーワード: VR, HMD, 没入型, テレコミュニケーション。

1. はじめに

情報通信技術の発展と普及により、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が注目されている。それらのビデオコミュニケーションは、図 1 左に示すように、パソコンに接続したカメラとマイクで取得した映像音声情報を介して行われるため、ユーザが受け取る情報は、音声と単一視点で観察される限定的な視野角の映像となり、相手の身振りや手振り、周囲の状況を把握するのは困難である。その結果、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。

我々はビデオコミュニケーションにおいて、情報送信側を主体、情報受信側を客体と定義した時、以下の問題への対応が必要と考えている。

問題 1: 限定的な視野角では主体が客体の意図や感情を理解することが困難。

問題 2: 単一視点では主体が客体の関心を理解することが困難。

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。

これら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発展を目指し、その実現法を提案する。

具体的には、図 1 に示すように、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、本手法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、相互理解が可能になることが期待される。ここで、360 度映像の観察方法としては、モニターを用いる方法とヘッドマウントディスプレイ (HMD) を用いる方法がある。本研究では空間共有感覚や没入感を得られるという利点から HMD を採用する。



図 1 従来のビデオコミュニケーション (左図) と本研究のテレコミュニケーション (右図) の比較:

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が表示されたディスプレイを代理人が装着することで、社会的な存在感と身体的な存在感の提示を実現することを目指した。実験を通じて、代理人はリモートユーザ本人とみなされる傾向にあることがわかり、一つのテレコミュニケーションシステムとして機能することが示された。

Cai ら[5]は、遠隔地に向くリモートユーザの肩に乗せた 360 度カメラの映像を屋内のワークスペースにいる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リモートユーザが Augmented Reality (AR) グラスによって確認するシステムを構築した。結果として二人のユーザは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が示された。笠原ら[6]は、指示役 (Ghost) が AR グラスとカメラを装着した人 (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を差し示す様子を Body の AR グラスに可視化した。その結果、遠隔ナビゲーションとしての有効性が示された。

VR 空間でユーザが互いにコミュニケーションすることが可能な Social VR が普及しつつある。代表例として VRChat[7]があり、ユーザはアバターを使用して音声チャットやボディランゲージによって他のユーザとコミュニケーションをとることが可能である。Social VR の社会的相互作用についての研究も行われ、VR 環境を利用したコミュニケーション方法の改善が探索されている。

このように、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における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者を好意的に評価すると同時に、集団間のバイアスを減少させることを明らかにした。実験参加者に年配の男性の写真を提示し、その男性の 1 日の生活についてエッセイを書くように指示した。参加者を A, B, C の三つのグループに分け、A グループには「あなたがその人になったつもりで書くように」と指示し、B, C グループには A グループに指示したような他者視点取得に関する指示とは別の指示をした。その結果、A グループは B, C グループと比較して、年配の人に対するステレオタイプなエッセイを書くことが少なく、その男性へ肯定的な評価をすることが示された。また、集団に関する実験では、実験参加者にコンピュータを使用した単純作業をさせた後、そのコンピュータが作業結果を過大評価していたことを伝えた。次に、参加者を A, B, C, D の四つのグループに分け、A グループには「まるで自分が過小評価されている人になったつもりで、過小評価されている人の 1 日の生活についてのエッセイを書くように」と指示した。B, C, D グループには A グループに指示したような他者視点取得に関する指示とは別の指示をした。その後、A, B, C, D グループに自分が属するグループ (内

集団) と、自分以外のグループ (外集団) について評価をするように指示した。その結果、A グループは外集団を内集団と同程度に評価した一方、B, C, D グループは内集団を外集団よりも明らかに好意的に評価した。以上の結果から、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た。

このように、他者視点取得は社会的相互作用において、相手との関係を改善し、相手を深く理解することに役立ち、より効果的なコミュニケーションを取ることを可能とする。

2.3 身体所有感と共感

Peak ら[9]は、白色人種の被験者に HMD を装着させ、彼らの動きと同期する白色人種と黒色人種のアバターをそれぞれ体験させる実験をした。その結果、黒色人種のアバターを使用した場合に、黒色人種への人種の偏見が減少することがわかった。また西田ら[10]は、子どもの認識を理解する機会を得るために、子どもの体験や経験を仮想的に再現することを目指した。被験者は、着用した HMD に自身の腰に付けたカメラの映像が映し出されることで、子どもの視点を取得し、手には子どもサイズの手の外骨格を着用して実験をした。結果として、被験者は周囲の人や物を大きく感じたり、圧迫感を感じたりする傾向にあり、「子供時代に戻った気がした」、「大人が目の前に立つと恐怖を感じる」という心理的效果も見られた。このように、視点位置や身体の変化によって人や物への知覚変化が確認されている。

Collaço Oliveira ら[11]は、目線の位置にカメラを付けた人 (パフォーマ) と HMD を装着した人 (ユーザ) において、パフォーマの映像をユーザが HMD で確認する状況を作った。この状況下で、パフォーマはユーザの動きを再現した。結果として、ユーザはパフォーマの体を自身の体であるという感覚が生じることが示された。またユーザはパフォーマに対して親近感を持つようになった。

このように、他者の視点を取得することは他者への理解や共感につながることを示されている。

3. 視点取得可能な没入型テレコミュニケーション

視点取得可能な没入型テレコミュニケーションシステムを図 2 に示す。図 2 の (a) では、客体の空間に置かれている 360 度カメラの映像を取得、(b) では客体の HMD の側面にマウントされたカメラの映像を取得し、主体の HMD で確認する。(c) では客体のマウントカメラの映像と HMD によって計測されるアイトラッキングデータを送受信し、主体の視点取得時に客体の注目領域を可視化する。また、主体と客体の立場を入れ替え同様の環境を構築することで双方向のコミュニケーションを実現する。

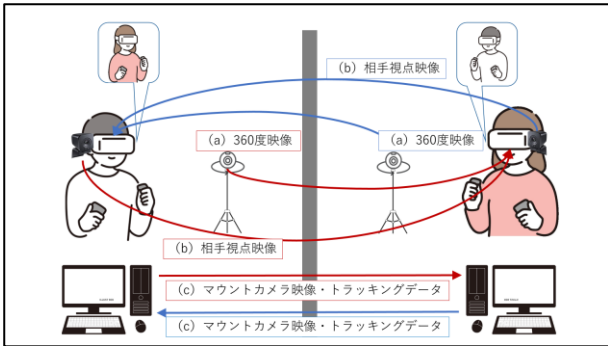


図 2 視点取得可能な実写映像に基づく没入型テレコミュニケーションシステムの概要図

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 マウントカメラ映像の送信

1節で述べた問題2の解決を目的として、相手の注目領域を取得する方法について述べる。主体が客体の視点を取得したときの様子を図3に示す。主体が使用しているデスクトップパソコンをPC-A、客体が使用しているデスクトップパソコンをPC-B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PC-BからPC-Aに送信され、主体のHMDで確認される。ユーザ間で主体と客体の役割を入れ替え、同様の処理を施すことにより、双方向での視点取得を実現する。

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TCP/IPソケット通信によりPC-BからPC-A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。これも主体と客体の立場を入れ替えて同様の処理を行うことにより、双方向での通信が可能となる。

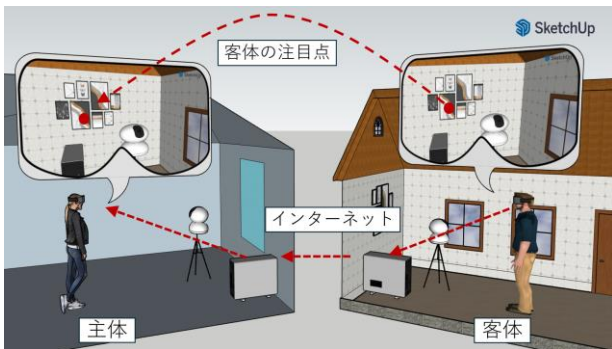


図 3 主体が客体の視点を取得したときの様子：客体のマウントカメラ映像と注目位置が主体に送信される。

5. 没入型テレコミュニケーション方式の実装

5.1 実装環境・デバイス

本研究ではHMDとしてMeta社のMeta Quest Proを用いる。解像度は片目あたり1800画素×1920画素、最大リフレッシュレートは90Hz、視野角は水平106度、垂直96度である。また、5個の赤外線センサーが内蔵されており、リアルタイムのフェイストラッキングが可能であるほか、フルカラーパススルー機能も備えている。映像切り替えに使用するコントローラーにはTouch Proを用いる。360度カメラはFun Tech社のInnex Cubeを用いる。4台のカメラユニットから構成されており、最大解像度は3840画素×964画素、最大フレームレートは30fps、最大視野角は水平360度、垂直90度、遅延は約250msである。マウントカメラにはサンワサプライ社のWebカメラCMS-V65BKを用いる。解像度は1920画素×1080画素、最大フレームレートは30fps、最大視野角は水平65度、垂直50度、遅延は約200msである。

5.2 VR空間の構築

Unityを用いてVR空間を構築する。360度映像を全天球に投影し、全天球の中心から映像を観察する。マウントカメラの映像は、平面に投影され、常にHMDの正面に配置されている。コントローラーによって、全天球と平面の表示/非表示を切り替えることで、映像の切り替えが可能である。

5.3 マウントカメラ映像の送受信

4節で述べたマウントカメラ映像の送受信はUnity Render Streaming[12]を利用したWebRTCによって行う。PC-Bはシグナリングサーバに接続要求を送信し、シグナリングサーバは接続要求を受け取り、PC-AにSDP(Session Description Protocol)情報を生成して返す。PC-AはSDPを受け取り、自身のSDPを生成し、シグナリングサーバを通じてPC-Bに返送する。PC-BとPC-Aは互いにICE(Interactive Connection Establishment)候補を交換し、最適な接続経路を確立する。SDPとICEの交換が完了すると、PC-BとPC-AはP2P接続が確立し、映像の送信が開始される。PC-AからPC-Bに映像を送信する場合においても同様の処理を施す。

5.4 視線情報の送受信

4.3節で述べたように視線情報の送受信はTCP/IPソケット通信によって行う。PC-BはTCPサーバを作成し、接続の受け入れ待機をする。PC-AはTCP/IPクライアントを作成し接続する。同様に、今度はPC-AがTCP/IPサーバを作成し、接続の受け入れ待機をする。PC-BはTCP/IPクライアントを作成し接続する。この手順によってデータの送受信を実現する。タイムスタンプを利用してデータ送受信の遅延を計測したところ、PC-BからPC-Aへの遅延平均は約23ms、標準偏差は約14ms、PC-AからPC-Bへの遅延平均は約28ms、標準偏差は29msであった。

5.5 視線情報の可視化

図4に示すように、主体が客体の視点取得時、客体の視

線情報を可視化して主体の映像に提示する。PC-BのUnity上でマウントカメラ映像が投影されている平面と、客体の視線の交点を平面のローカル座標で取得する。その座標をPC-BからPC-Aに送信し、PC-AのUnity上のPC-Bから送信された視点取得映像が投影された平面に表示する。PC-AからPC-Bに視線情報を送信する場合においても同様の処理を施した。

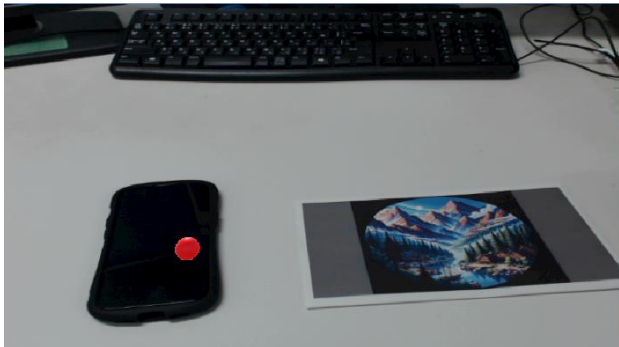


図 4 主体が客体の視点取得時に見ている映像：客体の注目位置を赤い丸で可視化している。

6. おわりに

本研究では、視野角が限定され、単一視点となる従来のビデオコミュニケーションで発生する問題、具体的には相手の意図や感情、関心を理解することが難しい状況を解決するために、HMDと360度カメラ、およびマウントカメラを用いて「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方式」を検討した。本システムの実用先として、遠隔での共同作業や指導活動に利用することが可能である。

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," 2023年4月18日。[オンライン]。Available: <https://hello.vrchat.com/>. [アクセス日:2024年7月3日]。
- [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] Unity Render Streaming [ソフトウェア]. (2024) Unity Technologies. 3.1.0-exp.7 Available: <https://docs.unity3d.com/ja/Packages/com.unity.renderstreaming@3.1/manual/index.html>.