

仮想空間への没入とハンドジェスチャを利用した遠隔会議システム

根岸 匠^{*1} 田中 二郎^{*2}

Remote Meeting System Using Hand Gesture and immersion in the virtual space

Takumi Negishi^{*1}, Jiro Tanaka^{*2}

Abstract – In most studies of remote meeting with virtual space, participants are identified by displaying their real names or their faces. In this study, we focus on the anonymity feature of the meeting. We developed a remote meeting system which maintains anonymity of the user, implemented using immersive head-mounted display and hand tracking. Our system offers the user a sense of immersion in the conference room that exists in the virtual space and it can apply the movement of the user's head and finger to the character avatar. Furthermore, using the network, we can share each participant's movement and voice in virtual space, and make it possible to perform a remote meeting where participants remain anonymous. Our paper shows the significance of performing an anonymity-based remote meeting system and discusses our contribution.

Keywords : Virtual Reality, Remote Meeting, Hand Tracking, Anonymity.

1. はじめに

遠隔会議技術は Egido ら^[1] の提唱したもの以来 25 年で大きく進化をしている。背景として、高速なインターネット回線やビデオチャットをするのに十分な性能を持ちながらも低コストであるコンピュータが一般に普及したこと、そして技術そのものの普及によるものである。インターネットを介して遠隔で会議をすることで、相手と意思疎通ができるほか、移動に伴う経費や時間の削減に有効であることから実際に会議への代替として用いる事例が増えてきた。

こうして遠隔会議の需要が大きくなってきたことに伴いビデオチャットベースの遠隔会議の研究やサービスが生まれてきたが、近年では遠隔会議に 3 次元の仮想空間や参加者をアバター表現して用いる研究がなされている。河口ら^[2] は展示会や学会などの相互コミュニケーションを伴うイベントに関して仮想空間上に同様のものを再現し、実空間と仮想空間の間で映像と音声によって相互のコミュニケーションが可能なシステムを開発している。Demeulemeester ら^[3] は仮想空間内に円形のテーブルと椅子を設置し、参加者を椅子に座る人型のアバターで表現し、アバターの頭上に参加者の名前を表示することで特定することを可能にしている。この時アバターの操作としてカメラ画像を用い

て実際のユーザの手と腕の動きを検出しアバターに適用している。これにより挙手やジェスチャーといった簡単な操作が可能であるシステムを開発した。こうした仮想空間やアバターを用いた遠隔会議システムの研究では、参加者の個人としての判断が求められるよう、参加者の情報に重点が置かれた状況を想定している。参加者が顔を表示したりアバターに実名を添えるなどして誰が誰であるかを特定できるシステムとなっており、実名を出さず匿名で行いたい遠隔会議においての有効な手段ではない。

インターネットを介した参加者が匿名の遠隔会議は、実際に対面している会議と比べ個人の情報が失われるため課題志向的な会議になると Sproull ら^[4] らは述べている。また、Walther ら^[5] は実際に対面した場合のコミュニケーション (FTF) とコンピュータを介した場合のコミュニケーション (CMC) を比較した。その結果、匿名的な CMC では、相手との親密なコミュニケーション活動が促進され、その結果コミュニケーション相手に対して過度な親密感を感じるという結果を得ている。これは自らの外見や非言語的行動に心的資源を割り当てる必要がないためその分発言内容に多くの資源を割くことができ親密で話しやすい環境を構築できるからであると主張した。これらから匿名の遠隔会議では例えば、ビジネスの場における会議で会社内で上司や部下などの社会的な関係に関わらず自己の意見を主張し意見を出しあいたい時に有効である。また匿名のいじめ相談窓口において、心を開きながら自分の情報は出さずにいじめに関する情報を提供し意見

*1: 筑波大学大学院 システム情報工学研究科

*2: 筑波大学 システム情報系

*1: Graduate School of Systems and Information Engineering, University of Tsukuba

*2: Faculty of Engineering, Information and Systems, University of Tsukuba

を仰ぐ時など、自己の情報は伝えず状況のみを相手に伝えたい場合も、匿名であるほうが有効である。

しかしこうした匿名の対話のみでは、非言語的な情報は限定されたものになってしまう問題がある。喜多^[6]は対話におけるジェスチャーの役割について論じており、ジェスチャーは相手にある内容を伝えるという行為の達成を助けるために存在すると論じていた。その役割において具体的に、体の動きそのものが伝えるべき内容を表現する場合であること、体の動きによって会話の流れを調節することによって発話による情報のやりとりをスムーズにすると指摘しているため、こうした非言語的な情報を表現できる手法が必要がある。

それらをふまえた上で本研究では、匿名と実名それぞれの遠隔会議の特徴を活かしどちらにも使用できるシステムを構築する。また非言語的な情報として、ハンドトラッキングシステムを用いてジェスチャーを行うことができる、さらに会議としての側面を考慮し、資料の提示や遠隔協調作業も可能であるシステムを構築する。本稿では匿名で数人で行う遠隔会議を想定し、システムの技術的詳細や手法の検討、本システムの貢献を論じる。

2. システムの構成

本システムの構成を説明する。本システムは没入感の高いヘッドマウントディスプレイである Oculus-RiftDK2 と、両手や指の位置や角度をセンシングできる LeapMotion を用いる。それぞれを PC に接続してシステムを起動する。

参加者は、図 1 のように OculusRift を装着し、机の上に置いた LeapMotion に対して手をかざすことで、仮想空間の会議室にいるキャラクターを一人称視点で操作することができる。現在、本システムが対象と

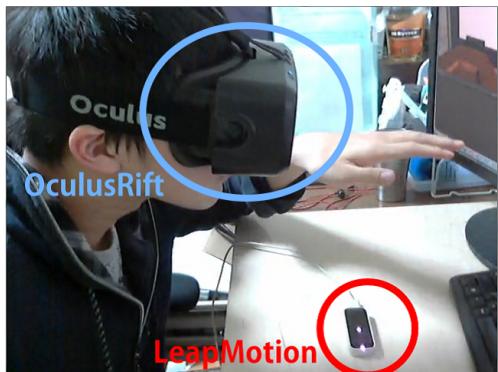


図 1 ユーザの利用想定図

する会議の規模は参加者が 3,4 人の小規模なものを想定したシステムとなっている。

3. システムの概要

システムの動作を説明する。現在の実装ではサーバサイドのシステムは作られていない。そのためクライアントサイドでどのように動作するのかに焦点を当て説明する。システムを起動後、クラウドサーバへと接続することで仮想空間の会議室へログインする。

3.1 アバターの操作

システム起動後、仮想空間に設置された会議室へと映像が切り替わる。まず頭部の操作について、OculusRift から得られるユーザの頭のヨー・ピッチ・ロール角からアバターの頭の方向を決定している(図 2)。その際アバターの自然な振る舞いを目指すため、一定以上の角度が検出されたらそれ以上頭が回転しないようとする処理を施し、代わりに胴体を捻るような処理を施している。ヘッドマウントディスプレイの映像に関して、アバターの頭部の位置からの映像を提供することでアバターを一人称視点で操作できるようにしている。



図 2 左:アバターが左を向いている様子 右:ユーザが左を向いている様子

手の操作については、LeapMotion によって得られた手の位置と角度の値から、腕の関節角度と手の角度を算出する。まず腕の関節角度については、得られた手の位置を終端部分として腕に対して逆運動学の処理を行いアバターの腕の適切な関節角度を算出後、アバターの腕に対して適用している。手の角度については、これは得られた手の角度をそのままアバターに適用している。

一方指の操作は順運動学に従い、親指以外の指については得られた指のそれぞれの DIP 関節、PIP 関節、MP 関節の角度の値をアバターの指へ適用する。親指は IP 関節、MP 関節の角度の値のみをアバターの指に対して適用している。

こうした処理を行い図 3 のような操作を達成する。これによりユーザは実世界で行うのと同様に任意のジェスチャーを行うことができる。さらに仮想空間内のメニューを操作することができるほか、任意のボディランゲージや拍手などをすることができます。

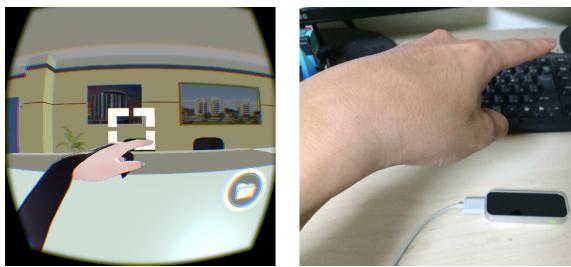


図3 左:ユーザの手と指のアバターへの適用 右:
その際のユーザの手と指

3.2 資料の操作・遠隔協調作業

次に仮想空間内の資料の操作方法について説明する。本システムに資料を用いるには実行ファイル以下のフォルダに ppt,pptx/png ファイルを置くとシステムが自動的に認識して仮想空間で使うことができる。ユーザの前方にはフォルダボタンが存在し、それをクリックすることで資料を選択するダイアログを開く。図4のようにダイアログには認識された資料のファイルが列挙されており、右のボタンのクリックにより任意のファイルを選択することができる。

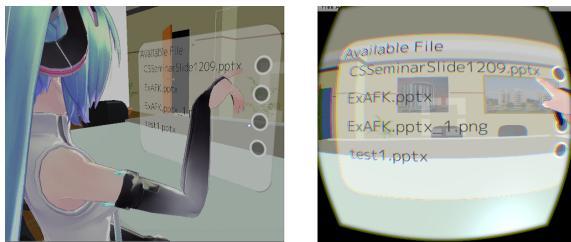


図4 左:ダイアログを選択する様子 右:この時ヘッドマウントディスプレイに表示される視界

例えば pptx ファイルを選択すると図5のような画面が目の前に生成される。この時資料の画像の周りにその資料を操作するためのボタンも同時に生成される。これらのボタンは資料を消す、資料を移動させる、資料を前方の大画面へ共有する、資料の次のページへ進む、前のページに戻る、資料に書き込みをする、などの機能を持つ。

資料への書き込みについて、図5の左上のペンのボタンを押すと資料が図6左のような書き込み状態に遷移する。書き込み状態では、図6右のようにキャラクターの指によってスライドに書き込みを行うことができる。このとき3つのボタンが同時に生成されており、書き込み状態を解除する、書き込みを前方の大画面と参加者に対して共有する、書き込みした資料を画像ファイルとして保存するなどの機能を持つ。またこの時書き込みを共有するボタンによってお互いに同じ資料に書き込みをすることができ、これによって遠隔協調作業を達成する。

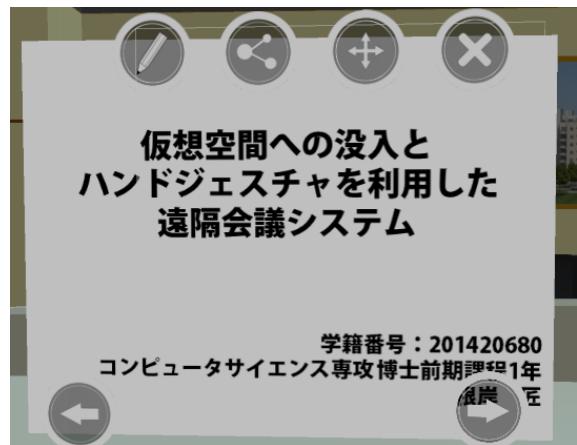


図5 仮想空間内における資料とボタンの生成

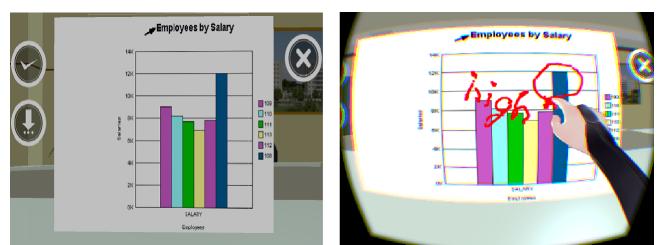


図6 左:書き込み状態に遷移した資料とボタン
右:実際に指によって書き込みを行っている様子

本システムにおける全てのボタンは指で一定時間触れた時、クリックとみなす。この理由としては、本システムにおいて手と指は資料の操作だけではなく拳手や拍手などのジェスチャとしての役割も持つため触れた時に即時にクリックとみなすと誤操作を起こしてしまうためである。

3.3 音声によるコミュニケーション

本システムではネットワークを介した音声によるコミュニケーションを可能とする。ただしそのままの声では匿名性が失われてしまう可能性があるためボイスチャットをするにあたり声をリアルタイムに加工する必要があると考えており、そのような実装を行う予定である。

またユーザの発話に伴いアバターの口をリップシンクさせることで発話していることを他ユーザに知らせる。アバターの頭部操作とこれを合わせることで誰に向かって話しているかを表現することを予定している。

4. 実装方法

本システムのソフトウェアの実装について PC の WindowsOS 用アプリケーションとして開発した。開発は、Unity5.0.1f と C# を用いて行った。

各デバイスのソフトウェアへの組み込みには各自の Unity 用 SDK を利用して実装しており、OculusRift の

SDKにはOVR SDK ver0.5.0.1を使用し、LeapMotionのSDKにはLeapMotion SDK ver2.2.7+30199を用いた。

5. 操作手法に関する予備実験と考察

5.1 予備実験に用いた3種の手法と実験の説明

本章では本システムで用いた資料の操作手法が有用であるかと使用感を確かめるため行った予備実験について述べる。実験にあたり図7に示した3種類の操作手法を用意した。本システムは没入型ヘッドマウントディスプレイを用いるため視認して操作する必要のあるキーボードやマウスなどを用いることは不適当だと考え、またジェスチャを行うにあたりデバイスを持つことはそれを阻害してしまうと考え3種類の手法に絞り込んだ。手法1は、OculusRiftをタップすることによってクリックとみなすものである。映像では正面に照準がありその位置にあるボタンにクリック処理を施す。手法2は机の上にLeapMotionを置きそれに手をかざすことによって仮想空間内に手を出現させる。この手の人差し指には棒状のオブジェクトがついておりこれをカーソルとする。図7の下段中央のようにボタンにそれを一定時間重ねているとクリックとみなす。手法2は現時点で本システムで用いられている手法である。手法3はOculusRiftの前面にLeapMotionを取りつけて正面に手をかざすことによって仮想空間内に手を出現させる。手の役割やクリック方法は手法2と同じである。

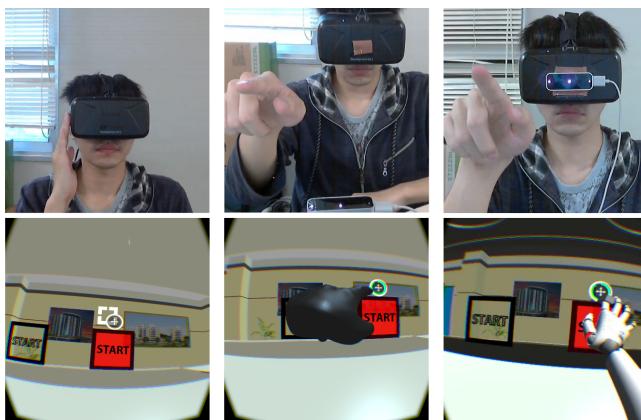


図7 上段左:手法1 上段中央:手法2 上段右:手法3
下段はそれぞれの手法に対応した映像の図

以上3種類の方法を用いてそれぞれその操作の正確性や使用感を確かめるため実験を行った。被験者9人(男性7名女性2名、いずれもコンピュータに慣れ親しんだ人物)にタスクを行ってもらった。被験者には一度これら3種の手法をそれぞれ練習してから仮想空間内に存在する10個の仮想空間のメニューを選択し

規定の位置へと操作するというタスクを行ってもらった。そのタスク完了時間を計測し、操作に正確性を求める場合の早さを確かめる。またタスクの完了後被験者には使用感に関する感想をアンケートに書いてもらい、実際の使用感を確かめた。

5.2 実験の結果

図8はそのタスクの完了時間をグラフ化したものである。被験者9人から得られた実験結果として、タ

タスク完了にかかった時間

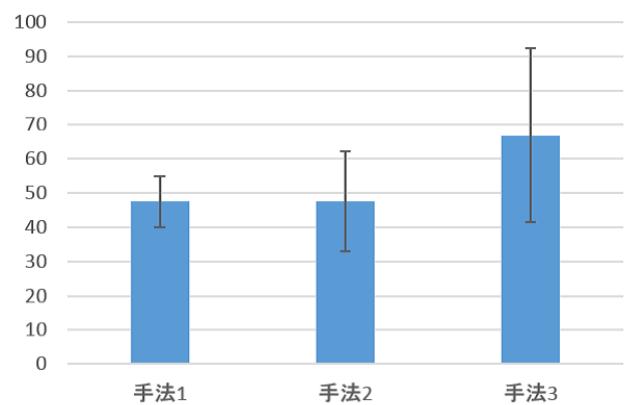


図8 タスク完了時間(秒)の各操作手法における平均値

スクの完了時間の平均は手法1が47.454秒、手法2が47.589秒、手法3が66.907秒であった。またタスクを完遂できなかった被験者はいなかった。平均の結果を見ると手法1と手法2の早さが等しく手法3が遅いように思えるが、得られたデータに対して二群の平均値に差があるかt検定を行った所、手法1と3では有意傾向ではあるが有意差は認められず($t(16)=2.07$, $p=0.054 > 0.05$)、手法2と3に関しても同様であった($t(16)=1.867$, $p=.085 > 0.05$)。

使用感の感想については、まず手法1では力加減がわからない、頭ごと目標に向けなければいけないため操作が難しかった、などのネガティブな意見が多く見られた。実際に練習の際、私が想定していたより弱くタップしていた被験者が多かった。手法2では、ポジティブな意見として手を用いた操作は直感的であった、精度がよく操作がしやすかった、奥行きを操作できるのは良かったという意見を得ることができた。ただし一方で操作をする際に腕が疲れるというネガティブな意見も得た。手法3についてはポジティブな意見として手法2よりも操作が直感的であったという意見を得た。一方で精度が悪かった、操作に腕が疲れるといったネガティブな意見も得た。

5.3 結果からの考察

平均達成時間が遅いが有意差が認められなかったことについて、手法3ではトラッキングの精度に問題があつたためと考える。LeapMotionによる手のトラッキングは赤外線画像から指を認識して行うものである。しかしこの時正面に対して指を向けるという設計上、図7の上段右のようにLeapMotionに対して指の方向が平行になり、指が手の甲によって隠れやすくなってしまう。これにより結果としてトラッキング精度が落ちてしまう問題があった。実験において指が手の甲に隠れてしまう問題に気づいた4人の被験者が手をLeapMotionに対して平行にならぬようにすることで操作をスムーズに終えており、手の方向が平行になっているまま操作した被験者は精度の低下により操作の完了までに時間がかかってしまっていた。これが平均値には差があるが、有意差が認められなかつた原因であると考える。手法3のこの問題を解決するにはアバターの手のひらにカーソルを手に垂直になるよう設置することで、手のひらで押し出すようにクリックするという操作手法に変えることによりLeapMotionに対して手の向きを垂直にして精度をなるべく高めながら操作を行う方法が考えられる。

使用感の感想について、手法1では操作のしにくさの感想が目立った。操作がしにくいうがこのことタイムに影響していない理由としては、タップという操作は再度行う操作をしても時間がかかるので多少のタップミスがあっても影響しなかったと考える。また手法2,3について共通して直感的な操作ができた、前後の奥行きの操作ができるることはよかった、などのポジティブな意見が見られた一方で腕が疲れたという意見が多数見られた。これはシステムで一定時間重ねているとクリックとみなすという手法を用いているからであると考えられる。中村ら^[7]は現実の大画面環境における研究で、ポインタを一定時間静止する操作をウェイティングと呼称しており、認識のブレにより位置がずれてしまう問題や決まった時間だけ静止していかなければならぬため選択に時間がかかるなどの問題が起きる可能性を指摘していた。また、これを解決するためにダブルクロッシングという対象を短時間に2度横切ることでクリックとみなす手法を提案していた。本研究でも腕の疲れという問題を解決するためにダブルクロッシングが有効である可能性がある。しかし中村らの研究と違い本研究では手の操作の役割として資料の操作だけではなくジェスチャを行うことも想定しているのでジェスチャの際に資料の誤操作が起きてしまう可能性もあり、取り入れるには慎重な考察が必要である。

加えて手法3では見ている方向と手を利用した入力

が噛み合つていて操作が手法2より直感的であったという意見を得た。これは手法2と比べ、トラッキング範囲が異なるためであると考える。手法3ではOculusRiftの正面に対してLeapMotionが設置されているためトラッキング範囲がOculusRiftの位置と向きに依存し、どの方向に対してもトラッキングできる、一方で手法2では机の上に設置するためトラッキング範囲は限られてしまう。このため例えば手法3真横に对しても手の操作を行うことができ、頭の動きと手の位置が合っているため更に直感的であったという意見を得たと考えている。手法3に関しては腕が疲れることと現状のシステムでは精度が低いことを除けばポジティブな意見が多くを占めた。

6. 関連研究

本研究と同じく仮想空間内の遠隔会議システムを本題とした研究にDeanら^[8]の研究がある。これはディスプレイの前に置かれたKinectを用いて実世界のユーザの動きをアバターに適用し、それをインターネットを介した仮想空間での会議で、頭部を実写画像にした棒人間のモデルに対して適用する。またユーザはその仮想空間での会議をディスプレイで視聴する。本研究では、ディスプレイを用いることに伴う頭部の可動範囲の制約を解決している。この制約は、例えば仮想空間内で真横にいる人に対して話したい場合、ディスプレイを見ないと仮想空間の様子がわからないため必然的にディスプレイをみることができる角度までしか頭を動かせないという制約がある。本研究ではヘッドマウントディスプレイを用い、装着したユーザがどこを向いていても映像が追従するため、この問題を解決する。

福井ら^[9]は、既存のビデオチャットにおいて誰が誰に注目して発話しているか認識できない問題に焦点を当て、カメラ画像からユーザの顔の方向を判別し、その方向に存在する仮想空間中のアバターの表示を拡大することで誰に注目して発話しているかを認識可能にするシステムを開発した。本研究も同様の問題に対して、ユーザの頭の方向をヘッドマウントディスプレイから取得することで取り組んでいる。

本システムと同様に没入型ヘッドマウントディスプレイを用いての遠隔協調作業研究にAmoresら^[10]の研究がある。作業者と支援者の2人に分かれて使うシステムであり、両者はカメラが設置されたヘッドマウントディスプレイを装着する。映像を共有することで協調作業を行うものであり、作業者は自らの視界と手の動きを、支援者は手の動きのみをお互いの映像に共有することで作業を支援する。本研究ではスライドへの書き込みによる協調作業を行う。

また本研究と同様のシステム構成を持つものに Lee ら [11] の研究が存在する。これは OculusRift に Leap-Motion を貼り付け、それを活かしたゲームを開発した研究である。この手法について、本研究の操作手法の実験における 1 つの選択肢として参考とした。

7. 結論

本稿では、匿名で行う遠隔会議を想定し、既存の遠隔会議システムを利用した時に発生する問題点を解決するシステムを提案・実装した。

今後はシステムを完成させた後、提案したアプローチが実際に有効であるか本システムを用いた被験者実験を行う。実験ではシステムの使用感に関するアンケートによって定性評価を行い、有用性を検証する。また本研究は匿名の会議を対象としシステムを開発しているが、実名においても資料の提示や遠隔協調作業のシステムは有効である可能性があるためそれを活かした新たな利用方法を模索する。

8. 参考文献

- [1] Egido, C.: Video Conferencing As a Technology to Support Group Work: A Review of Its Failures; Proceedings of the 1988 ACM Conference on Computer-supported Cooperative Work (CSCW '88), pp.13-24 (1988).
- [2] 河口, 梶, 中山, 浅井: SAMTK-3D を用いた多地点間仮想空間コミュニケーションとその応用. マルチメディア、分散、協調とモバイル (DICOMO) シンポジウム論文集, pp. 1664-1671, (2011).
- [3] Demeulemeester, A., Kilpi K., Elprama S. A., Lievens, S., et al.: The ICOCOON Virtual Meeting Room: A Virtual Environment As a Support Tool for Multipoint Teleconference Systems, Proceedings of the 11th International Conference on Entertainment Computing, pp.158-171, (2012).
- [4] Sproull, L., Kiesler, S.: Reducing Social Context Cues: Electronic Mailin Organizational Communication. Management Science, pp.1492-1512, (1986).
- [5] Walther J. B.: Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. Communication Research, pp. 3-43, (1996).
- [6] 喜多: ジェスチャーの認知科学 ひとはなぜジェスチャーをするのか. Japanese Cognitive Science Society, pp.9-21, (2000).
- [7] 中村, 高橋, 田中: 大画面環境におけるハンドジェスチャーの選択手法：ダブルクロッシングの提案と他の選択手法との比較（ヒューマンコンピュータインタラクション）. 電子情報通信学会論文誌. D, 情報・システム, pp.978-988, (2013).
- [8] Dean, J., Apperley, M., Rogers,B.: Refining Personal and Social Presence in Virtual Meetings. Proceedings of the Fifteenth Australasian User Interface Conference (AUIC'14), pp. 67-75 , (2014).
- [9] 福井, 喜多野, 岡田: 仮想空間を使った多地点遠隔会議システム:e-MulCS. 情報処理学会論文誌, pp. 3375-3384 , (2002).
- [10] Amores, J., Benavides, X., Maes, P.: ShowMe: A Remote Collaboration System That Supports Immersive Gestural Communication. Proceedings

of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems(CHI EA '15), pp.1343-1348 ,(2015).

- [11] Lee, P., Wang, H., Tung, Y., et al. TranSection: Hand-Based Interaction for Playing a Game Within a Virtual Reality Game. Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, pp.73-76, (2015).