

映像から身体が出現する実体重畳型ビデオ会議による 同室感の強化

Emergence of Physical Embodiment from an Overlay-Type Videoconferencing Enhances Social Telepresence

大西 裕也¹ 田中 一晶² 中西 英之¹

Yuya Onishi¹, Kazuaki Tanaka² and Hideyuki Nakanishi¹

¹大阪大学大学院工学研究科 知能・機能創成工学専攻

¹Department of Adaptive Machine Systems, Osaka University

²京都工芸繊維大学 情報工学・人間科学系

²Information and Human Sciences, Kyoto Institute of Technology

Abstract: There is a problem that the videoconference had a window in which the surface of a display forms a separate boundary as the remote space and the local space. In addition this problem reduces social telepresence (i.e. the sense that a participant feels as if he/she meets with the conversation partner in the same place). In this paper, as a method to reduce the visual difference between these two spaces, we tackled to embody a part of a remote partner's body and object in a video conference and make them appear in the middle of conversation. Through the experiment, we found that appearance of the embodied partner and object improved those presence and enhanced the sense of being together in the same room.

1. はじめに

ビデオ会議は、実際に対面して会話する時と比べ対話相手の存在感が低下する[13]. そのため、ソーシャルテレプレゼンス（遠隔地にいる人と対面している感覚[2]）が欠落する恐れがある. 我々は、遠隔地側の空間とユーザの空間の境界であるディスプレイが窓のような役割を持ち、二つの空間を分離させているのではないかと考えた. この問題を解決するために、ビデオ会議に握手用ロボットアームを組み合わせ、身体接触を負荷することによってソーシャルテレプレゼンスの強化を行った[8]. さらに、対話相手の映像とその身体の代替であるそのロボットアームがつながって見えるように提示することで、ソーシャルテレプレゼンスを強化した[9]. これらの実験で使用したデバイスは、遠隔地の空間とユーザ側の空間の視覚的な差異を少なくし、実体として提示ことで相手の存在感を強化してきた. つまり、相手空間の人・物を実体化することによって存在感が向上し、ソーシャルテレプレゼンスを強化することができる可能性がある. 本研究の目的は、物・人さらにそれらが映像から出現する時の存在感を調査することで、ソーシャルテレプレゼンスを強化する要因を

明らかにすることである.

これまでにビデオ会議を用いた遠隔地の空間を視覚的につながっている様に提示する様々な方法が提案されてきた. ClearBoard [4]や Video Whiteboard [14]ではガラスの板やホワイトボードを挟んで向かい合っている状況を設定し、それを描画面として視線などを遠隔地で共有する方法が提案された. また、机上に映像を表示する方法が提案されている[5,11,15,17]. DIGITABLEでは、従来のビデオ会議に加えて机上に対話相手の腕の映像を表示することで水平面の空間を共有した[11]. VideoArmsでは、スタイラスペンと映像を組み合わせることで接触跡を追加しスケッチを可能にした[16]. また、Remote Lag という手法によりジェスチャ映像が実物や人の手などに隠れて見えなくなってしまう状況を緩和し[16]、高さの表現を付加することや [3]、指示対象を色や形で視覚的に提示することで[10]、ジェスチャの解釈を改善できることを示された. また、指示棒の影を投影することで遠隔地間の机上で指示を共有する投影映像を用いる方法が提案された[18]. しかし、これらの方法はリモートの空間とローカルの空間の連続性を向上することはできても、いずれも実体として提示されていない. また、物理的に遠隔地にジェスチ

ャなどを伝えるため、ロボットを使用する方法も提案されている。遠隔地を自由に動き回ることが可能なロボットを遠隔操作し、そのロボットに搭載したレーザポインタで指示する方法[6]が提案されている。しかし、ロボットによる身体動作の提示は、操作者の姿を提示しないため、ソーシャルテレプレゼンスが低下することが知られている[13]。

また、ビデオ会議に握手用のロボットハンドを組み合わせ、身体接触の機能を付加する方法 [9]や対話相手の映像から身体の代替であるそのロボットアームがつながって見えるように提示する方法[9]、さらに、テーブルに鉛直方向に可動式のピンを格子状に配置し、そのピンの個々の高低差によって遠隔地側の腕先の形状を描写する方法が提案されている [7]。これら実験で使用した装置は、あらかじめ対話相手を表示するディスプレイの前にデバイスが置かれたままであった。

本研究では、映像内で提示されていた物や相手の身体の一部が会話の途中で実体として出現する実体重量型ビデオ会議システムを開発した (図 1)。このデバイスは、遠隔地側の空間の顔映像と手元の映像を 2 枚のディスプレイによって表示し、普段はディスプレイに遮られて見えないロボットアームが映像内の人の動きに同期してディスプレイの裏から出現するシステムである。また、遠隔地にいる対話相手が把持する物であれば実体として出現することも可能である。本研究では、物・身体さらにそれらが映像から出現する時の存在感を調査することで、同室感を強化する要因を検証する。人はロボットアームに、物は遠隔地にいる対話相手及びロボットアームが把持するぬいぐるみに対応させ実験を行った。

2. 構造

我々が開発したデバイスは、ビデオ会議とロボットアームを組み合わせたシステムである。遠隔地にいる対話相手の空間を映す 2 つのディスプレイは顔映像と手元映像を表示し、遠隔対話相手の上半身を等身大で映す。さらに遠隔対話相手とユーザの間には机を設置し、手元映像を映すディスプレイにもそれを表示している。ロボットアームは手元を映すディスプレイに遮られて見えない位置に設置している。図 2 にロボットアーム及びぬいぐるみが出現する様子を示す。ロボットアームのサイズは成人男性の平均的な大きさと設計しており、映像内に映る対話相手の腕や手の大きさと一致する。また実験者と同じ部屋にいる感覚を増すために、実験者の服とロボットアームの袖を同じ服で統一した。さらに、ロボットアームが人工物であるという印象を

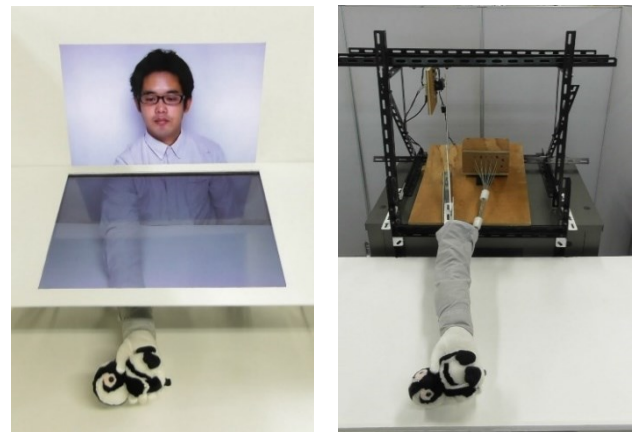


図 1. システムの外観

軽減させるために白い手袋をかぶせた。対話相手の手首には赤いリストバンド状のマーカを付け、手元を映すディスプレイの腕の位置をピクセル単位でリアルタイムに取得し、映像の腕とロボットアームが同期して動く。相手の映像の表示とロボットの制御を並列に処理しており動作遅延は目立たないようにしている。また、このロボットアームはスライドレールを含むリンク機構で構成されており、サーボモータの回転により手前や奥方向に伸縮する。さらに出現したロボットアームに機械的な印象を与えないようにするため、会話中に提示するぬいぐるみを軽く握るような動きを追加した (図 3)。これは手を開く・閉じる動きに相当するが、把持していたぬいぐるみを落とさない程度の開閉である。ロボットアームの手の開閉機構は、強度の高いワイヤを各指に通し、サーボモータでワイヤを引くことにより実現している。ぬいぐるみのみを実体として提示する場合は、ロボットアームがユーザに見えないようにアームの伸縮する距離を短くし、ぬいぐるみをディスプレイの裏からユーザが見える位置まで押し出すようにした。このときぬいぐるみの転がる方向を安定化させるため先端が平面である円筒をロボットアームにかぶせた。

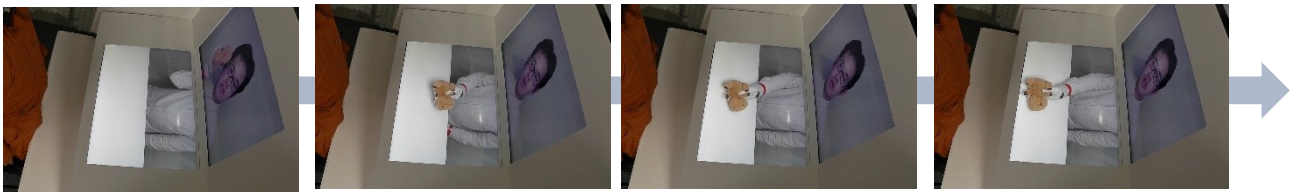


図 2.1 映像条件



図 2.2 物実体条件

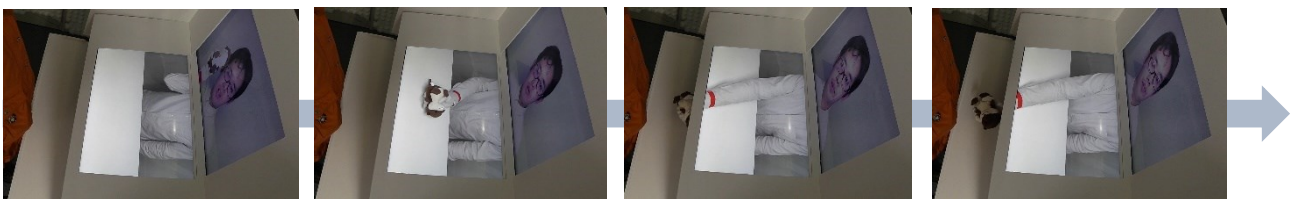


図 2.3 物+人実体条件

図 2. 実験の様子



図 3. ぬいぐるみを握る動作（上図：手が閉じている状態，下図：手が開いている状態）

3. 実験

本研究で開発したデバイスは、遠隔地にいる対話相手の腕や物を実体化し会話の途中で出現させるシステムである。そこで、遠隔地にいる対話相手の体の一部を実体として提示することで物・身体さらにそれらが映像から出現する時の存在感を調査することで、同室感を強化する要因を検証する。遠隔地に

いる対話相手の身体動作を映像で表示し、それと同期してロボットアームやぬいぐるみが実体として提示されると、映像内に表示されていた人・物がこちら側の空間に出てきたように感じられ存在感が向上すると考えられる。そして、ロボットによるジェスチャが本物の人間によって行われたというリアリティを生み、同室感が向上することが考えられる。これより以下の仮説が考えられる。

仮説 1: 対話相手の映像と同期して身体や物体が実体として出現することで、それらの存在感が向上される。

仮説 2: 対話相手の映像と同期して身体や物体が実体として出現することで、同室感が強化される。

これらの仮説を明らかにするため、以下の 3 条件を被験者内実験で比較を行った（図 2）。

映像条件 : 対話相手の腕とぬいぐるみの両方を映像で表示する。このとき、実験者の手が映像から見切れない程度にできるだけ被験者側に差し出す。

物実体条件 : 対話相手の腕は映像で表示し、ぬいぐるみは実体として会話の途中で出現させる。このとき、映像と出現するぬいぐるみのタイミングや向きを一致するようにした。

物+人実体条件：対話相手の腕とぬいぐるみを実体として会話の途中で出現させる。
このとき、映像と出現するロボットアームのタイミングを一致するようにした。

3.1 実験内容

全ての条件において、実験者が説明するぬいぐるみを差し出し、それを見ながらぬいぐるみに関してクイズを含む簡単な会話を行うタスクにした。会話開始時、把持するぬいぐるみは顔映像を表示するディスプレイに映像として表示する。そして、会話に途中で実験者が「こちらをご覧ください」と発言するタイミングで手元を表示するディスプレイにぬいぐるみを映し、被験者の手元までぬいぐるみを差し出す。統制された実験を行うためには、全ての実験で、会話時間を等しくする必要がある。会話が長いことや、会話中の質問が多くなると、そばにいる感じで高いスコアがつけられやすくなり、天井効果が発生しやすくなるため、会話は短く、会話中の質問の数は少なくした。

実験後にアンケートを実施し、それを実験の評価として用いた。アンケートで使用した各条件の写真を図5に示す。これらの写真は実験者がぬいぐるみを被験者の前に提示した後の様子を抜き出した写真となる。アンケートでは、条件間での映像や音声、会話内容のクオリティの差が実験結果に影響を与えなかったことを確かめるため、下記の質問項目を設定した。

- Q1 映像は十分にきれいだと感じた。
- Q2 音声は十分にきれいだと感じた。
- Q3 対話相手の説明は分かりやすかった。
また仮説を検証するため、物・人・差し出した動作の存在感に関して尋ねる質問と同室感に関して尋ねる質問として、以下の質問項目を設定した。
- Q4 あたかも同じ部屋の中で対話相手の差し出したぬいぐるみがあるように感じた。
- Q5 あたかも同じ部屋の中で対話相手の身体があるように感じた。
- Q6 あたかも同じ部屋の中で対話相手がぬいぐるみを差し出したように感じた。
- Q7 あたかも同じ部屋の中で対話相手が目の前にいるように感じた。

質問項目では7段階のリッカート尺度を用いた。アンケートには自由解答欄を用意し、被験者にスコアを付けた理由を記入してもらった。また、アンケート終了後に点数を付けた理由についてインタビューで尋ねた。

3.2 実験環境

被験者側の空間は、ディスプレイの裏にロボットアームを設置する。実験者側と被験者側の両方に、マイクとスピーカがあり、音声通話ソフトを用いて遠隔地間で会話を行うことができる。被験者側のスピーカは画面の方向から音声が聞こえるように、ディスプレイの後ろに設置した。ビデオ会議において、遠隔地の人の映像を等身大で提示する方法や[10]、アイコンタクトが成立するようにカメラとディスプレイを設置する方法[1]によって、同室感が強化されることがわかっている。そのため、実験前に被験者の目線の高さを測定し、実験者と被験者の目線の高さが一致するようにした。また、ディスプレイに表示される実験者を等身大の映像にするため、顔の縦の長さが22cmでとなるように調節した。ディスプレイの枠は実験者側の背景の色と合わせて白くした。顔を表示するディスプレイと手元を表示するディスプレイは同じ性能のウェブカメラを用意し、ディスプレイ間で対話相手が繋がっているようにみせるためにウェブカメラの位置を調節した。

実験開始前である被験者が実験室に入室する時、ディスプレイの裏にあるロボットアームやぬいぐるみを被験者から見えないようにするため、装置全体を布で覆った。実験開始時に実体であるロボットアームやぬいぐるみが見えてはいけなかったため被験者の立つ位置を固定した。また、デバイスの裏に人が隠れており、その人がぬいぐるみを差し出したという疑念を払拭するために、事前説明では、会話中実験室には1人しかいないことを被験者に説明した。事前説明の後、被験者に3つの条件を体験してもらった。順序効果による影響を避けるために、被験者によって条件の順番は変えて行った。

3.3 被験者

我々の大学キャンパスの近くに住む18歳から24歳の大学生16名（男性8名、女性8名）に実験に参加してもらった。このとき実験条件の順序による影響が起こらないようにカウンタバランスをとった。

3.4 実験結果

実験結果を図4に示す。3つの条件は、一要因分散分析を用いて比較した。棒グラフは各項目のスコアの平均値を表し、エラーバーは標準誤差を表す。

実験の映像、音声、会話内容のクオリティに関する質問では各条件で違いが無かったことを確認した。

存在感に関する質問項目では、ぬいぐるみの存在感に関する項目($F(2,15)=34.2, p<.01$)で有意な差が見られた。多重比較の結果、映像条件が他条件よりも

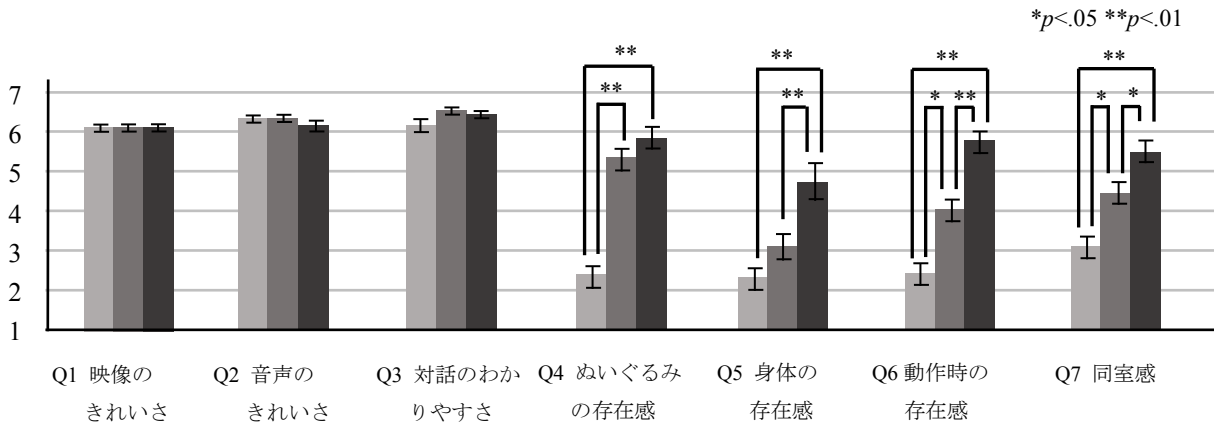
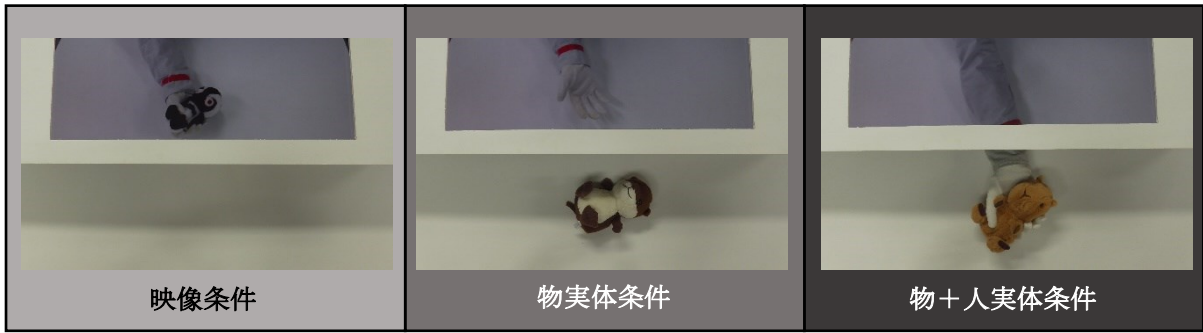


図 5. 実験のアンケート結果

低いことが分かった（それぞれ、 $p < .01$, $p < .01$ ）。また、身体の実在感に関する項目($F(2,15) = 12.5, p < .01$)においても有意な差が見られた。多重比較の結果、物+人実体条件が他条件よりも高いことが分かった（それぞれ、 $p < .01$, $p < .01$ ）。また、動作時の存在感に関する項目($F(2,15) = 31.9, p < .01$)においても有意な差が見られた。多重比較の結果、物+人実体条件が他条件よりも高く（それぞれ、 $p < .01$, $p < .01$ ）さらに物実体条件が映像条件よりも高いことがわかった（ $p < .05$ ）。以上のことから、物や身体が映像と同期して実体として出現ことでその物体の存在感が向上し仮説 1 を支持する結果となった。

同室感の項目($F(2,15) = 18.8, p < .01$)においても有意な差が見られ、多重比較の結果、物+人実体条件が他条件より高いことが分かった（それぞれ、 $p < .01$, $p < .05$ ）。さらに物実体条件が映像条件よりも高いことがわかった（ $p < .05$ ）。物実体条件及び物+人実体条件が映像条件よりもスコアが高くなっていることから仮説 2 を支持する結果となった。また、対話相手の身体が実体として出現する場合、物体のみが出現するよりもさらに同室感が向上することもわかった。

4. 考察

ぬいぐるみの存在感の質問項目では、インタビューにおいて、いずれの被験者も映像内で実験者が把

持していたぬいぐるみと実体として出現したぬいぐるみは同一の物と解釈して回答していた。したがって、映像と同期して物体の実体が提示されれば、対話相手が持っていた物体そのものが差出されたと感じていることが分かる。また、物実体条件と物+人実体条件において、ぬいぐるみが出現するタイミングで映像と実体の動機のずれを感じる被験者はおらず、むしろぬいぐるみが画面から消えるタイミングと、画面の下から出てくるタイミングが一致していたことを理由に存在感が向上したと回答した被験者が多かったことから、実体の提示の際には映像と同期していることが重要となる。相手の身体動作の映像を提示するだけでは物体の存在感は得られないが、何らかの物体が映像と同期して実体として出現することにより物体の存在感が創出されることが示された。また、身体の実在感においては、映像条件と物体条件の間に有意な差が見られなかった。これは実験者の腕が映像でしか提示されていないためであると考えられる。すなわち、映像内の空間とこちらの空間がつながっていると感じていても、実体が示されないと身体の実在感はあまり向上しないということである。差し出した動作の存在感については、ぬいぐるみと身体の実在感の両方を理由に挙げた被験者が多かった。以上をまとめると、何らかの物体が映像と同期して実体として提示されることでその物

体の存在感が創出されることがわかった。また、相手の身体が直接実体として提示しなくても相手の動作が想像される場合には、相手の動作の存在感が向上すると考えられる。

同室感の質問項目では、差し出した動作の存在感と同じ傾向を示している。物実体条件が映像条件よりもスコアが高くなっていることから、相手の身体が実体として提示することがなくても、相手の動作によって物が実体として提示されたことが想像することができれば、あたかも対話相手が同じ部屋にいるように感じられる。したがって、本システムのように相手の身体動作を提示する映像を用いる場合は、何らかの実体が出現することによって相手の動作の存在感を創出することが同室感の強化において重要となる。また、実体である身体を提示することにより更に同室感が強化されることが分かった。

被験者ごとに実験条件の順序を変えていたが、初めて実体の出現を体験した物実体条件及び物+人実体条件において、ぬいぐるみが映像から実体に出現したタイミングで驚きのリアクションをとった被験者が3人いた。インタビューにおいて、「物が出てきた瞬間はどのようにして差出されたか分からなかった」と回答していた。ぬいぐるみが映像から実体に出現することを想像しておらず、気が付いたら実体のぬいぐるみが目の前にあったためその瞬間は戸惑ってしまったが、会話の流れから対話相手が差出したぬいぐるみであることはすぐに理解できたようであった。しかし、この印象はアンケートの質問項目にマイナスの印象を与えていることはなかった。このように、ぬいぐるみを差し出す動作を映像内で表示していても、実体として物や身体が出現することを予期できず驚く被験者が一定数いた。

また、初めて実体の出現を体験した物実体条件及び物+人実体条件において、ぬいぐるみに触れようとする被験者が8人（物実体条件5名、物+人実体条件3名）いた。実験者が「こちらをご覧ください」と発言していても、手元に実体のぬいぐるみがあると触って確かめたい被験者が多かった。ぬいぐるみを触ったことが各存在感や同室感のスコアに影響しなかったことをインタビューで確認した。実体が出現することにより存在感や同室感が強化される他に、映像のみでは発生しない心理が働く可能性がある。

本研究で開発したデバイスの応用例について議論する。実験で使用したデバイスは実験者が被験者に物を差し出す行為に着目したが、腕の動きの機能などを拡張することにより様々な応用が可能となる。例えば、家庭教師などの遠隔教育におけるリアルタイムでの指導や採点、カードゲームやボードゲーム

などの遠隔テーブルゲームが挙げられる。遠隔教育では、教師の腕が出現することによりどこを指導しているのか分かりやすくなる。さらに、用紙の読み込みと印刷技術により、生徒の答案用紙を教師がロボットハンドで受け取り、遠隔地で再現された答案用紙を教師が採点を行い、同様の方法で生徒に返すことでリアルタイムの教育が可能となる。遠隔テーブルゲームでは、使用する道具が常に卓上もしくはプレイヤーの手元にある。プレイヤーの腕を代替するロボットアームによって、手札を卓上に出すことや相手にカードやコマなどを渡すことが可能となる。さらに、本実験により実体で差し出された物体に触れようとした被験者がいたことから、上記とは異なる応用が可能となる。例えば、乾杯や名刺交換である。映像内に映った相手が液体の入ったコップを実体として差し出し、コップをあてることによって乾杯を行うことができる。また、相手が差し出した名刺を受け取り、自分の名刺を渡すことで名刺交換も実現することができる。これらの方法は、ユーザは目の前に出現した物体に接触する・手に取るといった行動をとることが可能となる。また、これらの方法は遠隔地にいる相手とユーザの間である卓上の空間だけで成立するだけではなく、それぞれの行為を行った後、相手の行動を映像で確認することができ、更なる同室感の向上が期待できる。本研究では物体や身体が出現した瞬間の効果を検証したが、その後の効果について調査することは今後の課題である。

5. おわりに

本研究では、映像内で提示されていた物や相手の身体の一部が会話の途中で実体として出現する実体重型ビデオ会議システムを開発した。実験より、何らかの物体が映像と同期して実体として出現することでその物体の存在感が創出されることがわかった。さらに、相手の身体が実体として出現しなくても、何らかの物体の実体として提示するによって相手の動作が想像される場合には、同室感が強化されることが分かった。また実体のある物体の出現が、視覚的な効果に加え、物体に触れようと試みるなど、これまで確認されなかった行動を新たに創出できる可能性がある。この効果について調査することは今後の課題である。

謝辞

本研究で使用したデバイスの開発及び実験に協力していただきました山口隆浩氏に心より感謝いたします。本研究は、JSPS 科研費 JP26280076, JP15K12081, JP16K16140, KDDI 財団, 科学技術融合振興財団, 電気通信普及財団, 立石科学技術振興

財団からの支援を受けました。

参考文献

- [1] Bondareva, Y. and Bouwhuis, D.: Determinants of Social Presence in Videoconferencing, Proc. AVI 2004, pp.1-9, (2004)
- [2] De Greef, P. and Ijsselsteijn, W.: Social Presence in a Home Tele-Application, In Proceedings of Cyberpsychology, Behavior, and Social Networking 2001, pp. 307-315, (2001)
- [3] Genest, A., and Gutwin, C.: Evaluating the effectiveness of height visualizations for improving gestural communication at distributed tabletops, Proc. CSCW 2012, pp. 519-528, (2012)
- [4] Ishii, H. and Kobayashi, M.: ClearBoard: a Seamless Medium for Shared Drawing and Conversation with Eye Contact, Proc. CHI 1992, pp.525-532, (1992)
- [5] Izadi, S., Agarwal, A., Criminisi, A., Winn, J., Blake, A. and Fitzgibbon, A.: C-Slate: a Multi-touch and Object Recognition System for Remote Collaboration using Horizontal Surfaces, Proc. Tabletop 2007, pp.3-10, (2007)
- [6] Kuzuoka, H., Oyama, S., Yamazaki, K., Suzuki, K. and Mitsuishi, M.: GestureMan: a Mobile Robot that Embodies a Remote Instructor's Actions, Proc. CSCW 2000, pp.155-162, (2000)
- [7] Leithinger, D., Follmer, S., Olwal, A. and Ishii, H.: Physical telepresence: shape capture and display for embodied, computer-mediated remote collaboration, Proc UIST 2014, pp. 461-470, (2014)
- [8] Nakanishi, H., Tanaka, K. and Wada, Y.: Remote Handshaking: Touch Enhances Video-Mediated Social Telepresence. Proc. CHI 2014, pp.2143-2152, (2014)
- [9] Onishi, Y., Tanaka, K. and Nakanishi, H.: Embodiment of Video-mediated Communication Enhances Social Telepresence, Proc. HAI2016, pp.171-178, (2016)
- [10] Ou, J., Chen, X., Fussell, S. and Yang, J.: DOVE: Drawing over Video Environment, Proc. Multimedia 2003, pp.100-101, (2003)
- [11] Pauchet, A., Coldefy, F., Lefebvre, S., Louis, S., Perron, L., Bouguet, A., Collober, and t M., Guerin, J. and Corvaisier, D.: TableTops: Worthwhile Experiences of Collocated and Remote Collaboration, Proc. TABLETOP 2007, pp.27-34, (2007)
- [12] Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H. and Hagita, N.: Android as a Telecommunication Medium with a Human-like Presence, Proc. HRI2007, pp. 193-200, (2007)
- [13] Tanaka, T., Nakanishi, H. and Ishiguro, H.: Comparing Video, Avatar, and Robot Mediated Communication: Pros and Cons of Embodiment, Proc. CollabTech2014, pp. 96-110, (2014)
- [14] Tang, J. and Minneman, S.: VideoWhiteboard: Video Shadows to Support Remote Collaboration, Proc. CHI 1991, pp. 315-322, (1991)
- [15] Tang, A., Pahud, M., Inkpen, K., Benko, H., Tang, C.J and Buxton B.: Three's Company: Understanding Communication Channels in Three-way Distributed Collaboration, Proc.CSCW2010, pp. 338-348, (2010)
- [16] Yamashita, N., Kaji, K., Kuzuoka, H. and Hirata, K.: Improving Visibility of Remote Gestures in Distributed Tabletop Collaboration, Proc. CSCW 2011, pp. 95-104, (2011)
- [17] 村田雄一, 志築文太郎, 田中二郎: SyncSurface:デジタルワークスペースとフィジカルワークスペースを統合し遠隔地間で共有するシステム. 日本ソフトウェア学会第26回大会講演論文集, (2009)
- [18] 渡辺 貴文, 上杉 繁, 三輪 敬之: 異なる空間への指示行為が可能な仮想の影による道具インタフェースの開発, 情報処理学会論文誌, Vol.48, No.12, pp.3919-3930, (2007)