身体性遠隔コミュニケーションにおけるユーザとアバタの視点の一致

石井 健太郎[†1] 谷口 祐司[†2] 大澤 博隆[†3]
中臺 一博[†4] 今井 倫太[†3]

本論文では，仮想的な身体を持つアバタを用いた遠隔コミュニケーションシステム PROT AVATAR におけるアバタ操作手法に関する実験をもとに，得られた知見について議論する．PROT AVATAR によるコミュニケーションでは，アバタの操作者の映像を遠隔地に投影するため，表情により感情を伝えることができる．さらに，アバタの操作者にとっては明確ではない，アバタの投影に適切な位置をシステムが自動で計算するため，アバタの操作者は投影位置を考えることなく遠隔の環境内を指し示すことができる．しかし，アバタの操作者のとる視点はアバタの視点とは異なることがあるため，アバタの操作者の発話がアバタとの対話者にとっては自然ではない場合がある．本論文では，アバタの操作手法として，自動操作手法と半自動操作手法の 2 つの手法を設計・実装し，比較実験を行った．実験の結果，半自動操作手法のほうが自動操作手法よりも，アバタとの対話者にとって自然な発話を行うことが示された．また，実験を通して得られた遠隔コミュニケーションシステム関する知見をまとめる．

Merging Viewpoints of User and Avatar in Embodied Telecommunication

Kentaro Ishii,[†1] Yuji Taniguchi,[†2] Hirotaka Osawa,[†3]
Kazuhiro Nakadai[†4] and Michita Imai[†3]

This paper discusses the findings of the viewpoint of an avatar-controlling user on the basis of experimentation with an implemented embodied telecommunication system named PROT AVATAR. Communication using an avatar with facial expressions is useful when a user wants to express emotions. On top of this feature, our system supports automatic avatar movement toward nearest visible location to the target, which is not obvious for the avatar controller. With our system, the avatar controller can easily refer to something remotely. However, sometimes, the words of an avatar controller may not be intuitive for an avatar viewer, because the avatar controller does not necessarily share the viewpoint of the avatar. We designed automatic and semi-automatic methods for controlling the avatar, and we conducted an experiment to compare the two methods. The results show that semi-automatic control is more intuitive than fully automatic control for an avatar viewer, and they have design implications for embodied telecommunication systems.

1. はじめに

遠隔コミュニケーションのための技術が発達し，テレビ電話やビデオチャットを利用して離れた場所にいる相手と対面対話を行うことができる．この場合，通常の電話では困難な，相手の表情から感情を読みとることや，相手の様子をうかがうことができる．一方，レーザポインタを利用した Gesture Laser[1] や Visual SLAM を利用した遠隔アノテーションシステム[2] を用いることで，遠隔から相手の環境中の物体を指し示すことができる．例えば，海外の工場で機械が故障して修理が必要な場合，修理方法を知る社員がその場にいなくても，これらのシステムを利用して現地に赴くことなく，修理を行うべき正確な位置を現地社員に伝えるといったことができる．

本論文では，上記の 2 つの特徴をあわせもつ遠隔コミュニケーションシステムとして開発した PROT AVATAR の操作手法について議論する．PROT AVATAR は，アバタ投影システム Remy[3] をもとに

†1 東京大学大学院 情報学環
Interfaculty Initiative in Information Studies, the University of Tokyo
†2 慶應義塾大学大学院 理工学研究科
Graduate School of Science and Technology, Keio University
†3 慶應義塾大学 理工学部
Faculty of Science and Technology, Keio University
†4 （株）ホンダ・リサーチ・インスティチュート・ジャパン
Honda Research Institute Japan Co., Ltd.

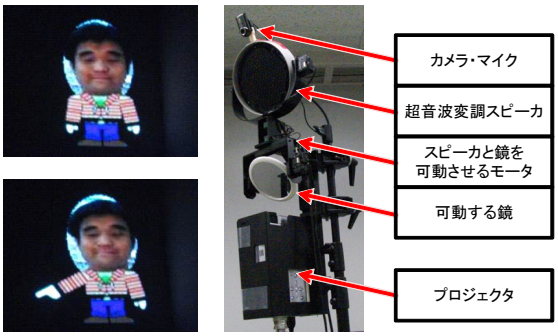


図 1 PROT AVATAR. 可動する鏡と超音波変調スピーカを用いた身体性遠隔コミュニケーションシステム
Fig. 1 PROT AVATAR. telecommunication system utilizing actuated mirror and ultrasonic speaker for embodied interaction

したシステムであり，カメラでキャプチャした自分の顔に，身体アニメーションを付与して作成したアバタを投影することができる（**図 1**）．また，音声は超音波変調スピーカから射出され，投影されているアバタの位置から聞こえてくるように感じることができる[4),5)]．PROT AVATAR は，上記の Remy の特徴に加えて，アバタとの対話者（以後，単に対話者と呼ぶ.）にとって好ましい位置を自動的に計算する機能を備えている．3.2 節の予備実験で示す通り，アバタがある物体を指し示す際に，対象物体から離れたとしても平坦で単色の背景に投影するほうが，対象物体の近くの平坦で単色ではない背景に投影するよりも，対話者には好まれる．PROT AVATAR は，あらかじめ定められたアバタの投影に適した背景の範囲のうち，対象物体に最も近い位置を計算して投影位置を決定する．アバタの投影に適した背景はアバタの操作者（以後，単に操作者と呼ぶ.）にとっては明確ではなく，アバタの投影位置をシステムが自動計算することは操作者の負担を軽減し，操作しながらの対話を支援する．

しかし，本システムの利用時には，対話者にとって自然ではない発話を操作者が行うことが起こりうる．なぜならば，対話者はアバタの位置に操作者がいるものとして知覚するのに対して，操作者はアバタから見た視点で遠隔地を見ているのではなく，システムに付属するカメラからの映像を見て対話するためである（**図 2**）．典型的な例として，指示語の発話が挙げられる．実空間上の物体を指し示すことができる状況では，対象物体を指定するために指示語が用いられることがある．これは，指示語の利用によって効率よく対話を行うことができるためだと考えられる[6)]．指示語は，発話者とその発話相手の位置関係や状況により解釈が決まる言葉であるため，発話者・発話相手の位置関係や状況に関する共通の理解がないと指示語は正しい意味を成さない．

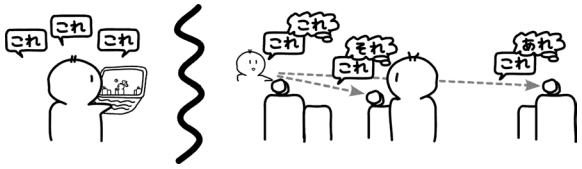


図 2 操作者のとる視点と対話者が考える視点の不一致
Fig. 2 Mismatch between Viewpoints of an Avatar Controlling User and an Interaction Partner

本論文では，操作者のアバタ操作インタフェースを工夫することで，操作者がとる視点をアバタの視点と一致させることができることを示す．Imai らの実験では，ロボットを物理的に動かすことによって，人間のとる視点がロボットの視点に変わる場合があることが示されており[7)]，同様に操作者のとる視点をアバタの視点に変えることができると考えられる．適切な投影位置を自動的に計算する PROT AVATAR の操作手法として，自動操作と半自動操作の 2 種類を設計し，プロトタイプを実装して比較実験を行った．自動操作手法は，操作者が指し示したい画面上の物体をクリックすると，自動的にアバタが適切な位置まで移動する手法である．半自動操作手法は，操作者が指し示したい画面上の物体をクリックすると，適切な位置を画面上に提示する手法で，操作者が提示された位置をもう 1 度クリックすることによってアバタは移動する．比較実験の結果にもとづき，実空間を移動可能な遠隔コミュニケーションシステムの操作手法について議論する．本論文の貢献として，まず実空間を介した遠隔コミュニケーションにおいては，単純な方法では視点の不一致が起こりコミュニケーションがうまくいかないという事例を示す．また，実験を通して得られた知見をもとに，遠隔コミュニケーションシステムの設計の参考となる情報を提供する．

2. 関 連 研 究

これまでに，対面コミュニケーション用のアバタを用いる遠隔コミュニケーションシステムが作成されており，例えば Xbox LIVE[8)] や Second Life[9)] において，利用者はソーシャルネットワーク上において自身のアバタを作成・操作することができる．一方，Gesture Man[10)]・Geminoid[11)]・駅における実験[12)] のような，遠隔から操作されるロボットを用いることにより，遠隔地における方向を指し示したり対話相手に物理的な存在感を感じさせるといったことが行われてい

る．また，Nakanishi らは，遠隔地にいる対話者の移動に応じて，カメラが動くこと[13] やディスプレイが動くこと[14] によって，対話者の存在感が増すことを示した．これらの物理的な移動を伴うシステムを用いた研究では，遠隔地にいる対話者は主に，移動するロボットや物体の視点をとって対話する．したがって，ロボットや物体の移動にともなって遠隔地にいる対話者が知覚する映像も変化していくが，本研究で用いるシステムでは，アバタをプロジェクタで投影することにより生じさせているため，移動に対する自由度が高く，対話者が知覚する映像は変わらないままで，アバタが移動するといったことが起こる．そのため，視点の移動に関して考慮することが必要となる．

指示語の意味は発話者の視点に大きく影響を受ける．これらは psycholinguistics と呼ばれる分野で，人間がどのように視点を選択しているかを中心に研究されている[15]．McNeill によると，発話者のとる視点は Origo と呼ばれる．Ullmer-Ehrich の実験では，発話者が自分の部屋の構成を説明するときに，ドアから部屋の内側に向かう視点をとることが示された[16]．Klein の実験では，発話者が対話相手に目的地までの道順を教える際に，その経路にそって視点を変化させていくことを示した[17]．この実験では，対話相手も一緒に視点を変化させていくことが確認されており，このとき変化していく視点のことを，Klein は joint viewpoint と呼んだ．また，Imai らの実験では，人間がロボットを持ち上げて移動させる際に，人間がロボットの視点をとるように視点を変化させていることを示した[7]．Imai らは，ロボットを持ち上げて物理的に操作したことが joint viewpoint を作り出したと結論づけている．本論文では，これらの視点の変化をアバタの操作者に対しても起こすことができることを示す．

3. PROT AVATAR

3.1 システム概要

本研究で開発した PROT AVATAR は，操作者側と対話者側のシステムからなる（**図 3**）．操作者側のハードウェアは，スピーカ付きコンピュータ・マイク付きカメラによって構成されている（**図 4**）．アバタ操作ソフトウェアはこのコンピュータで稼働する．操作者側のシステムはカメラ画像から操作者の顔を認識して切り出し，対話者側に投影されるアバタを作成する（図 1）．また，マイクにより操作者の音声をキャプチャし，スピーカにより対話者の音声を出力する．一方，対話者側のハードウェアは，PROT・マイク付きカメラによって構成される．PROT は，可動式の鏡を

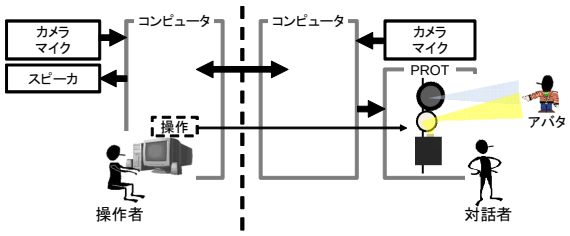


図 3 システム構成．操作者側と対話者側のシステムはネットワーク接続されている．操作者と対話者は互いに対話することができ，操作者は遠隔地に投影されている自分のアバタを操作することができる．

Fig. 3 System Configuration. Both avatar controller and viewer side components are network connected. The avatar controller can control the remote avatar, while both the avatar controller and viewer can talk to each other.

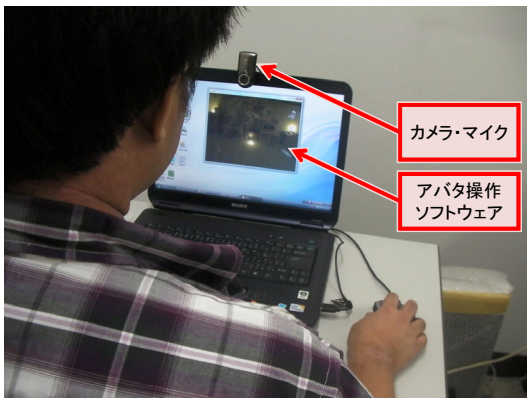


図 4 対話者側のハードウェア

Fig. 4 Avatar Controller Side Components

取り付けたプロジェクタと可動式の超音波変調スピーカからなる映像・音声の投影システムである[4]．超音波変調スピーカは離れた位置を音源として音声を生成することができるスピーカであり[5]，モータで可動させることにより任意の位置に音声を投影することが可能である．この機能により，対話者にアバタの位置から音声が聞こえるように知覚させることができる．カメラは対話者側の部屋を撮影するものであり，その映像を操作者は図 4・**図 5** の通り画面上で見ることができる．また，対話者の音声をマイクでキャプチャし，操作者側のシステムに送信して出力する．

操作者側システムからモータ制御情報が対話者側システムに送られると，対話者側システムは PROT のモータを動かし，アバタ画像と音声の投影位置を変更する．また，同様に操作者側システムから，アバタの体部分の画像を変化させ，指差しのジェスチャを行うことができる．本研究では，基本的な操作の方法として，操作者が部屋の映像上の物体をクリックすることによりアバタを操作することを考える．これは，操作

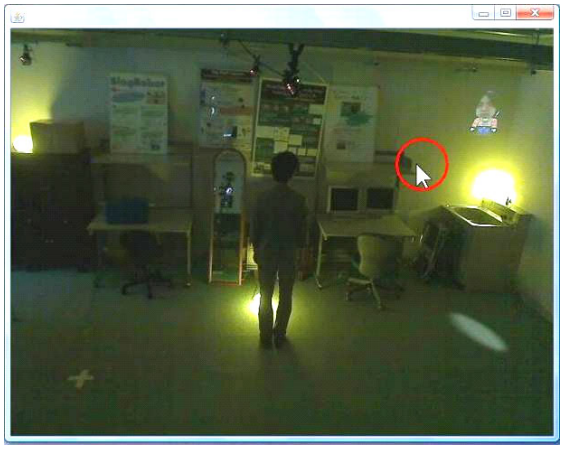


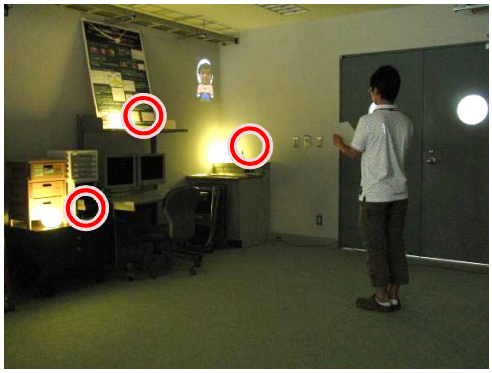
図 5 アバタ操作ソフトウェアのスクリーンショット
Fig. 5 Screenshot of Avatar Control Software

者が物体を指し示す場合に通常先に意識するのは，アバタをどう移動させるかではなく，どの物体を指し示すかであるためである．

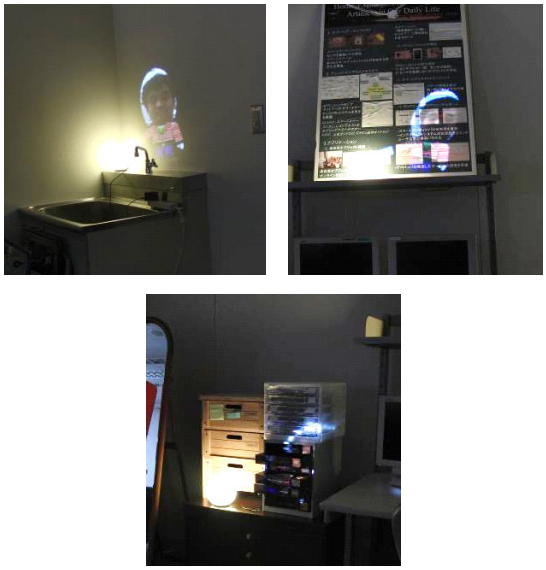
3.2 予備実験

アバタを投影する場所を検討するための予備実験を行う．PROT AVATAR はアバタの投影にプロジェクタを用いているため，投影場所がアバタの見えやすさに影響を及ぼす可能性があるためである．**図 6** に予備実験の実験環境を示す．対話者側の部屋に 3 つのライトを用意し，1 つ目は白い壁の前，2 つ目はポスターの前，3 つ目は棚の前に配置した．白い壁は平坦で単色の背景，ポスターは平坦だが単色でない背景，棚は平坦でない背景にそれぞれ対応するものとして，どの場所に投影されるのを好むかを評価する．

予備実験は 5 名の評価者に対して行った．評価者は，配置したライトの近くにアバタが移動して指差す場面を観察し，「アバタの見えやすさ」と「アバタの指し示す対象のわかりやすさ」に関して 7 段階のリッカート尺度によって点数を付けるというタスクをそれぞれ 3 つのライトについて行った．その結果，両方の質問に対して，5 名全員が平坦で単色な背景に対して最も高い評価をつけ，平坦だが単色でない背景，平坦でない背景の順で評価が低くなった．評価者に対しインタビューしたところ，指し示す対象からアバタが離れたとしても平坦で単色の背景にアバタを投影するほうが，指し示す対象の近くの平坦で単色ではない背景に投影するよりも好まれることがわかった．したがって，アバタの操作手法を設計するうえで，必ずしも物体のそばに移動すればよいわけではなく，平坦で単色の背景に移動し，その場所から物体を指し示すほうがよいことがわかった．



アバタの投影場所



各場所おける投影の様子

図 6 予備実験の実験環境．実験参加者は PROT AVATAR により投影されたアバタを見て，「アバタの見えやすさ」と「アバタの指し示す対象のわかりやすさ」を評価した．
Fig. 6 Preliminary study setting. the participant saw the projection of PROT AVATAR and determined how much of the avatar the participant could see and how well the participant could identify the location pointed by the avatar.

3.3 2 種類の操作手法の設計

予備実験の結果にもとづき，自動操作手法と半自動操作手法の 2 種類の操作手法を設計した（**図 7**）．どちらの手法においても，指し示す物体の位置が与えられると，システムは投影に適した背景の範囲のうち，対象物体に最も近い位置をアバタが移動すべき位置として求める．自動操作手法では，操作者が指し示したい画面上の物体をクリックすると，システムは自動的にアバタが移動すべき位置までアバタを移動させる．半自動操作手法では，操作者が指し示したい画面上の物体をクリックすると，システムはアバタが移動すべき位置を画面上に赤丸で提示し，操作者が提示された

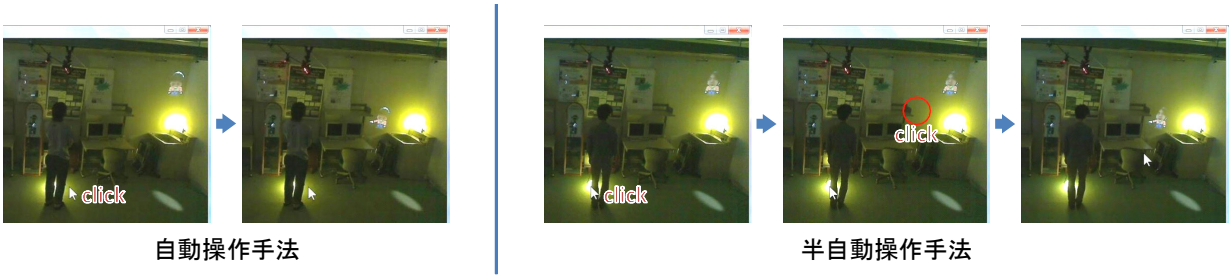


図 7 アバタ操作手法．自動操作手法では，操作者が物体をクリックするとアバタは移動すべき位置に移動する．半自動操作手法では，操作者が物体をクリックするとシステムはアバタが移動すべき位置を提示し，操作者が提示された位置をさらにクリックするとアバタはその場所に移動する．

Fig. 7 Avatar Control Methods. With automatic control, the avatar just moves toward the appropriate location when the avatar controller specifies a pointing location. With semi-automatic control, the appropriate location is first indicated when the avatar controller specifies a pointing location, and the avatar moves after the avatar controller clicks the indicated location.

赤丸内をもう 1 度クリックするとアバタはその場所に移動する．移動と同時に，アバタの指差しの向きが対象物体に向くよう，システムはアバタの身体を回転させる．

4. 実　　験

自動操作手法と半自動操作手法における，操作者の指示語の利用を比較する実験を行う．指示語の利用は，発話者がとる視点に密接に関連しているため[7]，指示語の利用を計測することで発話者がどの視点をとっていたかを測ることができる．対話者にとって自然なのは，操作者がアバタの視点をとることである．本実験では，どちらの手法が対話者にとって自然な指示語を用いたかを比較する．

4.1 実 験 手 順

図 8 に，対話者側の実験環境を示す．実験参加者は，はじめにこの部屋に案内され，実験者が自身のアバタを操作して行われる PROT AVATAR の紹介デモを体験する．紹介デモの際に，実験者からアバタを介して実験参加者へいくつか質問をたずね，対話者として実験参加者は操作者である実験者との対話を体験する．その後，実験参加者は別の実験者によって操作者側の部屋に案内される．

操作者側の部屋への案内後，実験参加者に自動・半自動のどちらかのアバタ操作手法についての説明がなされる．実験参加者は対話者側の部屋に 3 つのライトが設置してある様子（図 8）を画面上で観察できるようになっており，操作手法の説明終了後，しばらくアバタ操作の練習を行う．十分な練習のあと，実験者は実験参加者に，以下の実験で行うタスクを説明する．

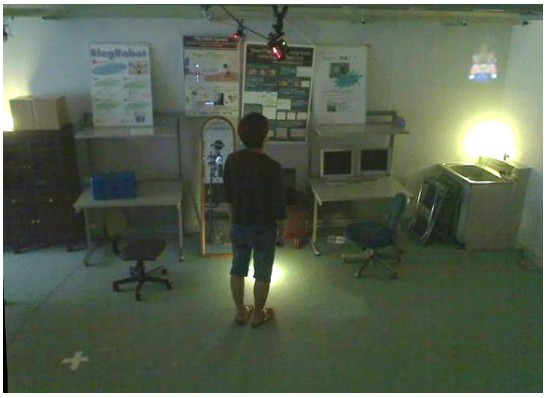


図 8 実験における対話者側の実験環境

Fig. 8 Avatar Viewer Side Environment in the Experiment

実験で行うタスクは，「これ」・「それ」・「あれ」のいずれかの指示語を用いながら説明を受けた手法でアバタを操作して，対話者側の部屋にいる実験者にライトの色の変更指示を出すというものである．対話者側の実験者は実験参加者の指示にしたがい，リモコンでライトの色を変更する．この一連のタスクを 3 つのライトすべてに対して行う．タスク終了後には，実験参加者の主観的な感想を調べるため，タスクに対するアンケートを行う．なお，この実験では，予備実験で議論したアバタの投影に適した背景の範囲は，あらかじめ手動でキャリブレーションされており，このアバタの投影に適した背景の範囲とクリック点から，システムはオンラインでアバタの移動位置を求める．

日本語では，発話者の近くにあるものを示すときに「これ」，発話相手の近くにあるものを示すときに「それ」，両者から遠くにあるものを示すときに「あ

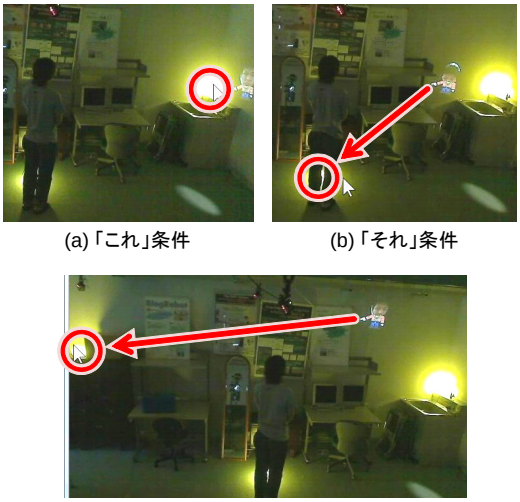


図 9 各条件におけるアバタと対象物体の位置
Fig. 9 Avatar and Target Location for Each Condition

れ」という指示語を用いる．**図 9** は，3 つのライトに対してシステムが計算した投影に適切な位置である．図 9(a),(b),(c) はそれぞれ，対話者にとっては「これ」・「それ」・「あれ」を用いるのが自然と考えられる位置関係になっており，以降それぞれを「これ」条件・「それ」条件・「あれ」条件と呼ぶ．

操作者の本来の視点をとると，3 つのどの条件においても「これ」という指示語を使うのが最も自然である．なぜならば，画面上のどの場所も操作者からは十分に近い位置であるためである．そのため，実験参加者が「これ」という指示語を使うか，各条件に応じて対話者にとって自然な指示語を選んで使うかを調べることによって，どの視点をとっているかを解析することを意図している．

4.2 実 験 結 果

日本人の 36 名を実験参加者として実験を行った．29 名が男性・7 名が女性であり，年齢は 19 歳から 43 歳であった．36 名のうち 18 名ずつを自動操作手法・半自動操作手法の 2 グループに分けて実験を行った．指示対象となるライトや色の変更順は，学習効果を避けるためにカウンターバランスをとって実験を行った．

表 1 に，実験の結果として指示語の発話人数を示す．それぞれの列は条件ごとに 18 名中何名がその指示語を発話したかを表している．影付きのセルは対話者にとって最も自然な発話を示している．「それ」条件・「あれ」条件において，自動操作手法よりも半自動操作手法を用いた場合に，対話者にとって自然な指示語を用いた実験参加者の人数が多くなった．対話者にとって自然な指示語を用いた人数の割合を，自動操作手法と半自動操作手法の比較においてフィッシャーの正確確率検定により検証したところ，「それ」条件において有意差（$p < 0.05$）・「あれ」条件において有意傾向（$p < 0.1$）が見受けられた．「これ」条件においては，統計学的差異は見受けられなかった．この結果は，半自動操作手法を用いた実験参加者のほうが対話者にとって自然な指示語を用いる傾向を示している．

表 1 実験結果．「それ」条件で有意差（$p < 0.05$）・「あれ」条件で優位傾向（$p < 0.1$）が示された．影付きのセルは各条件における最も自然な発話を示している．
Table 1 Experimental Result. Significantly different ($p < 0.05$) in condition “SORE” and difference trend ($p < 0.1$) in condition “ARE.” The shaded cells indicate the most natural utterance for the avatar viewer.

	「これ」条件		「それ」条件		「あれ」条件	
発話	自動	半自動	自動	半自動	自動	半自動
これ	18	17	14	4	13	6
それ	0	1	4	11	1	2
あれ	0	0	0	3	4	10

タスク終了後に行ったアンケートでは，なぜその指示語を用いたかをたずねる自由記述の問いに対して，自動操作手法を用いた複数の実験参加者が，「画面上のクリックした点は 3 つとも自分から近い位置だから」といった，自分の視点をとっているとみられる回答をしたのに対して，半自動操作手法を用いた複数の実験参加者が，「アバタから近い/遠い位置だから」といった，アバタの視点をとっているとみられる回答をした．本実験では，事前に統制をとって指示語を用いた主観的な理由をたずねる準備をしていなかったため参考にしかならないが，今回の実験では自動操作手法でアバタの視点・半自動操作手法で自分の視点をとったとみられる回答をした実験参加者はいなかった．

5. 考察と議論

実験の結果は，アバタ操作の完全な自動化は身体を持つアバタを用いた遠隔コミュニケーションシステムにおいては必ずしも適切ではなく，半自動操作手法を用いたほうが自然な発話が行われるということを示唆している．なぜ半自動操作手法で，操作者がアバタの視点をとるようになるかについては，今後実験を重ねて明らかにしなければならないが，半自動操作手法でアバタの移動位置をクリックさせることにより，明示的にアバタの位置を意識することが最も大きな理由であると考えている．物理的なロボットを物理的に持ち上げて移動させることで，人間がロボットの視点をとりうることが先行研究では示されているが[7]，仮想的な身体を持つアバタを画面のクリックにより仮想的に

操作することで，物理的な操作と同様に人間がアバタの視点をとりうるという本研究の結果は，ヒューマンエージェントインタラクションの研究を進める大きなきっかけとなると考えている．

一方，操作者にとっては，半自動操作手法は自動操作手法に比べて 1 クリック多く操作を必要とする手法である．しかし，その差は十分に少ないと考えており，半自動操作手法においても投影に適した位置はシステムによって求められ，操作者は画面内のどの位置が投影に適した場所であるかを考える必要はないため，手動でアバタを移動させて物体を指し示す操作に比べると，大きく手間が軽減されている．

Sugiyama らは，日本語の「それ」と「あれ」という言葉の示す範囲が人間ごとに異なっていることを指摘している[6]．実験結果における「それ」条件下の「あれ」という発話と「あれ」条件下の「それ」という発話はこの知見により説明可能である．「あれ」と「それ」は対話者によって互いに解釈可能であるととらえると，「これ」と発話したかそれ以外の指示語を発話したかにより，対話者にとっての理解可能性を判断することができる．この前提で自動操作手法と半自動操作手法を比較すると，「あれ」条件においても「それ」条件と同様に有意差（$p < 0.05$）が見受けられた．

本研究では，プロジェクタにより操作者の表情と仮想的な身体を投影するシステムを用いた．このことが，レーザポインタの点で位置を指示する Gesture Laser[1] や矢印で位置を指示する遠隔アノテーションシステム[2] とは異なり，平坦で単色の背景に投影してほしいという対話者の要求を生み出していると考える．点や矢印であれば，対象物体の位置に情報を損なうことなく直接投影できる．しかし，相手の表情を見ながら対話できるテレビ電話やビデオチャットの特徴を保ったまま動作する PROT AVATAR にとっては，対象物体から離れても投影に適切な位置を選択することは必要な機能である．

一方で，物理的な身体を持つ人間型ロボットを用いて同様のコミュニケーションを行う場合は，対話者が知覚する操作者はロボットとなると考えられるが，そのロボットにカメラを搭載することができるため，操作者はロボットからの視野を確認することができる．ただし，PROT AVATAR のようなプロジェクタで投影してできるアバタは，物理的なロボットより素早く自由に移動できるという特徴を持つ．

本論文で示したアバタ操作手法の知見は，人間や自律動作するロボットやエージェントがインタラクションに関与するようなシステムに活かすことができると考える．なぜならば，人間はインタラクションの相手がシステムを介した人間であることや自律動作する人工物である場合には，本研究で示した「操作者は視点をアバタの位置にとる」といった期待を自然と持つためである．このような人間や自律動作する人工物との対話においては，注意深く人間の特性を考慮してインタラクションをデザインする必要がある．

PROT AVATAR の対話者側システムで用いられているカメラは 1 つだけであり，アバタの視点を直接的に操作者に提供しているものではない．複数のカメラが対話者側の部屋に設置されているような環境では，本論文で示したのとは異なる手法で，問題を直接的に解決できる可能性がある．しかし，本論文で示した手法は，通常の部屋に 1 つにパッケージ化された PROT AVATAR を配置すれば実現できるという点で，より制約の少ない理にかなった手法であると考える．

Ullmer-Ehrich[16] や Klein[17] が示したように，対話中の視点の変化は日本語での対話に限らず起こるものと考えているが，本論文において視点の変化を計測するために用いた手段は，日本語特有の使い回しに依存する．そのため，別の言語での対話において視点の変化を計測するためには，異なる手法が求められる．

6. ま と め

本論文では身体を持つアバタを用いた遠隔コミュニケーションシステムにおいて，アバタの操作者がとる視点を調べるための実験を行った．自動操作手法・半自動操作手法の 2 つの手法をプロトタイプに実装し比較したところ，半自動操作手法を用いた実験参加者のほうがアバタとの対話者にとって自然な発話をすることが示され，アバタの操作者がアバタの視点をとっていることが示唆された．仮想的な身体を持つアバタの操作の際に，明示的にアバタの位置を意識させるようなインタラクションをデザインすることによって，コミュニケーションがより自然となるという知見が示唆され，人間や自律動作する人工物との対話の際には，注意深く人間の特性を考慮してインタラクションをデザインすることが必要になるといえる．

今後は，アバタの投影に適した背景の範囲を画像処理により自動取得する仕組みの開発に取り組むとともに，アバタを用いた遠隔コミュニケーションにおけるさらなる改善と知見の獲得のための追加実験を進めていく予定である．

参 考 文 献

1) Yamazaki, K., Yamazaki, A., Kuzuoka, H., Oyama, S., Miki, H.: Development of an Embodied Space to Support Remote Instruction, European Conference on Computer Supported Cooperative Work, pp.239–258, (1999).
2) Reitmayr, G., Eade, E., Drummond, T. W.: Semi-automatic Annotations in Unknown Environments, IEEE and ACM International Symposium on Mixed and Augmented Reality, pp.1–4, (2007).
3) 藤村亮太, 郭斌, 大村廉, 中臺一博, 今井倫太: 実物体を扱う遠隔協調作業を支援する壁面投影移動型アバタシステム Remy の提案, 知能と情報, Vol.21, No.5, pp.701–712, (2009).
4) Fujimura, R., Nakadai, K., Imai, M., and Ohmura, R.: PROT – An Embodied Agent for Intelligible and User-Friendly Human-Robot Interaction, International Conference on Intelligent Robots and Systems, pp.3860–3867, (2010).
5) Ishii, K., Yamamoto, Y., Imai, M., Nakadai, K.: A Navigation System Using Ultrasonic Directional Speaker with Rotating Base, International Conference on Human-Computer Interaction, Lecture Notes in Computer Science, Vol.4558, pp.526–535, (2007).
6) Sugiyama, O., Kanda, T., Imai, M., Ishiguro, H., and Hagita, N.: Human-Like Conversation with Gestures and Verbal Cues based on a Three-Layer Attention-Drawing Model, Connection Science, Vol.18, No.4, pp.379–402, (2006).
7) Imai, M., Hiraki, K., Miyasato, T., Nakatsu, R., and Anzai, Y.: Interaction With Robots: Physical Constraints on the Interpretation of Demonstrative Pronouns, International Journal of Human-Computer Interaction, Vol.16, No.2, pp.367–384, (2003).
8) Xbox avatars. http://www.xbox.com/en-US/live/avatars/
9) Second Life. http://secondlife.com
10) Kuzuoka, H., Oyama, S., Yamazaki, K., Suzuki, K., and Mitsuishi, M.: GestureMan: A Mobile Robot that Embodies a Remote Instructor's Actions, ACM Conference on Computer Supported Cooperative Work, pp.155–162, (2000).
11) Ishiguro, H., and Nishio, S.: Building artificial humans to understand humans, Journal of Artificial Organs, Vol.10, No.3, pp.133–142, (2007).
12) Shiomi, M., Sakamoto, D., Kanda, T., Ishi, C.T., Ishiguro, H., and Hagita, N.: A Semi-autonomous Communication Robot: A Field Trial at a Train Station, ACM/IEEE International Conference on Human-Robot Interaction, pp.303–310, (2008).
13) Nakanishi, H., Murakami, Y., Kato, K.: Movable Cameras Enhance Social Telepresence in Media Spaces. ACM SIGCHI Conference on Human Factors in Computing Systems, pp.433–442, (2009).
14) Nakanishi, H., Kato, K., Ishiguro, H.: Zoom Cameras and Movable Displays Enhance Social Telepresence, ACM SIGCHI Conference on Human Factors in Computing Systems, pp.63–72, (2011).
15) McNeill, D.: Psycholinguistics: A New Approach, Harper & Row Press, (1987).
16) Ullmer-Ehrich, V.: The structure of linving space descriptions, Speech, Place and Action, pp.219–250, (1982).
17) Klein, W.: Local Deixis in Route Directions, Speech, Place and Action, pp.161–182, (1982).