

アバタコミュニケーションのための ポインティング操作インタフェースの設計

Designing Pointing Control Interface for Avatar Communication

谷口祐司¹ 石井健太郎² 大澤博隆³ 中臺一博⁴ 今井倫太³

Yuji Taniguchi¹, Kentaro Ishii², Hirotaka Osawa³, Kazuhiro Nakadai⁴, and Michita Imai³

¹ 慶應義塾大学大学院 理工学研究科

¹ Graduate School of Science and Technology, Keio University

² 東京大学大学院 情報学環

² Interfaculty Initiative in Information Studies, the University of Tokyo

³ 慶應義塾大学 理工学部

³ Faculty of Science and Technology, Keio University

⁴ (株) ホンダ・リサーチ・インスティテュート・ジャパン

⁴ Honda Research Institute Japan Co., Ltd.

Abstract: Communication using an avatar with body expressions is useful when a user wants to refer to something remotely. In this paper, we discuss the design of pointing control interface for such an avatar communication system named PROT AVATAR. We start from the existing prototype and describe an iterative design process on the basis of user study with an implemented prototype to show insights and lesson learned. In particular, the user studies showed that people preferred that the avatar was projected on a solid background even if the avatar's location was far from the pointing target, and semi-automatic control is more intuitive than fully automatic control for a remote user. We believe that our report provides design implications for avatar communication systems and will contribute HAI community when designing related communication systems.

はじめに

アバタ技術の活用により、離れた場所にいる相手とのコミュニケーションを円滑に行うことが可能となる。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いでも、現地に赴くことなくアバタを利用して、修理を行うべき正確な位置とその方法を現地社員に伝えるといったことができる。

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 [1]や Second Life [2]において、利用者はソーシャルネットワーク上で自身のアバタを作成・操作することができる。その一方で、Gesture Man [3]や Geminoid [4]のようなロボットを用いることにより、遠隔地の対話相手に物理的な存在感を感じさせることができる。

人間の姿をかたどったアバタを用いることの1つ

の特徴は、アバタの指差し動作で、実空間上の物体をポインティングできることである。本論文では、プロジェクタによりアバタを任意の場所に投影可能なシステムである PROT [5]をもとに設計した PROT AVATAR というシステム（図 1）の、ポインティング動作の操作インタフェースの設計について議論する。既存の操作インタフェースでは、マウスドラッグによってアバタを移動し、マウスクリックによってポインティングジェスチャを行うことができる [6]。しかし、ユーザが物体をポインティングする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、どの物体をポインティングするかである。また、移動してポインティングを行うのに既存の手法では2操作必要とする。そこで、基本方針として、ユーザが物体をクリックすることでアバタを移動させることを考える。

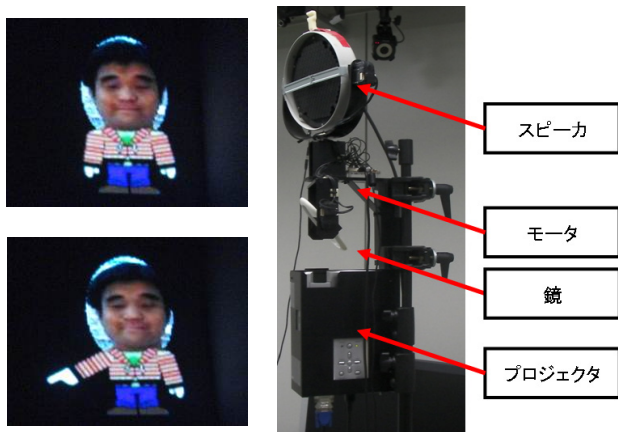


図 1: PROT AVATAR

本研究では、プロトタイプの実装・評価を目的として、そのプロトタイプを用いてユーザ評価や比較実験を行い、その結果をもとに段階的に次のインタフェースを設計・実装するというアプローチをとる。本論文では、その段階的な設計プロセスについて、どのようにインタフェースを検討したのかを述べるとともに、ユーザ評価から得られた知見をまとめることによって、アバタを介した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。

PROT AVATAR

システム概要

本論文では、システムプラットフォームとして、アバタシステム Remy [6]で用いられているものと同様の遠隔アバタコミュニケーションシステム PROT AVATAR を利用した。PROT AVATAR は、図 2 のようにローカルとリモートのシステムからなる。

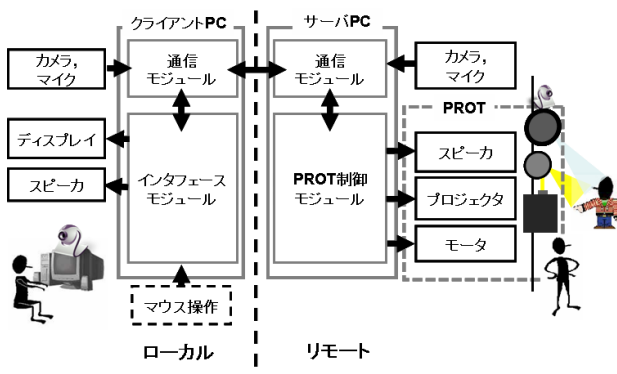


図 2: システム構成

ローカルシステムのハードウェアは、図 3 にあるようにスピーカ付きコンピュータ・マイク付きカメラ・スピーカによって構成されている。本論文において述べるアバタ操作インタフェースはこのコンピュータ上で起動される。ローカルシステムはカメラにより操作者の顔を撮影し、図 1 にあるようなアバタを作成する。また、マイクにより操作者の音声をキャプチャし、スピーカによりリモートの対話相手の音声を出力する。

一方、リモートシステムのハードウェアは、PROT・マイク付きカメラによって構成される。PROT は、可動式の鏡を取り付けたプロジェクタと可動式の超音波スピーカからなる画像・音声の投影システムである[5]。超音波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、モータを利用するより任意の位置に画像や音声を投影することが可能である。この機能により、対話相手にアバタの位置から音声聞こえるように感じさせることが可能である。カメラはリモートの様子を撮影するものであり、その映像をローカルの操作者は図 3・図 4 のように画面上で見ることができる。マイクによりリモートの対話相手の音声をキャプチャし、ローカルシステムに送信して操作者に対して出力する。

ローカルシステムからモータ制御情報がリモートシステムに送られると、リモートシステムは 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、同様にローカルシステムから、アバタの体部分の画像が変更して、指差しのジェスチャを行うことができる。

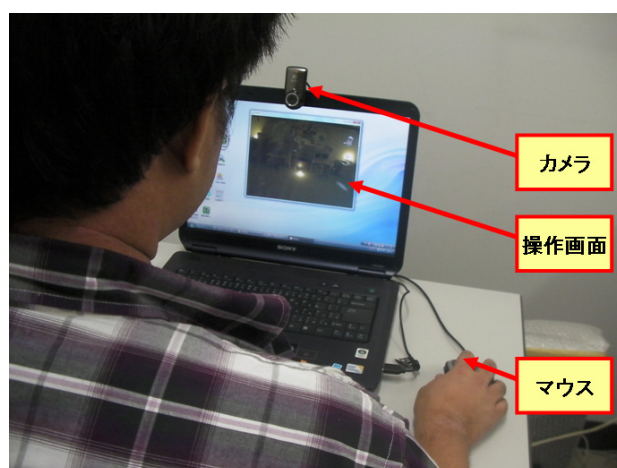


図 3: ローカルシステム

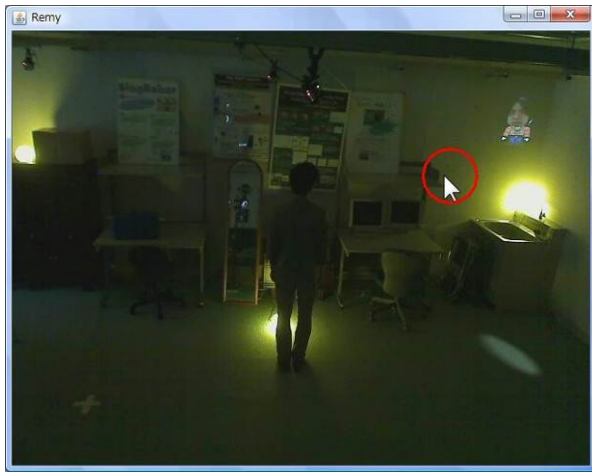


図 4: アバタ操作のスクリーンキャプチャ

物体クリックによる操作

ローカルの操作者が物体をポインティングする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、どの物体をポインティングするかである。既存の方法で、対象物体のそばにアバタが移動するようにするためには、アバタの現在の位置と目標位置を考慮してドラッグする量を考えなければならない。これは、必ずしも容易な方法とはいえず、少なくとも操作者の慣れを必要とする。

そこで、操作者は物体をクリックするだけで、システムが現在の位置から移動量を計算して、物体のそばまで自動的にアバタが移動する手法を用いることを考える。この手法を用いると、ユーザの操作は非常に容易なものとなる。一方、自由な移動ができなくなるが、ポインティング動作の目的のためには、自由な移動が必要になることは少ないと考える。また、既存のドラッグによる移動とは排他的なものではないため、必要があればドラッグによる微調整ができるよう、既存の方法を残しておくこともできる。

問題点

しかし、この手法を操作インタフェースに実装し、実際に動作させたところ、アバタの移動先をシステムが自動的に決めるため、アバタが見えにくい場所を避けられないという問題点が明らかになった。PROT AVATAR は、プロジェクタによりアバタを投影しているため、背景によりアバタが見えにくいことがある。1 つは投影先に物体があり表面が平面ではない場合で、アバタの人間をかたどった姿が提示されない。もう 1 つは投影先が平面であっても一樣ではないカラフルな背景である場合で、表情やポインティング点がわかりにくい。

ユーザ評価

上記の問題点の重要性と改善の方向を検討するために、投影場所に関するユーザ評価を行った。図 5 にユーザ評価の設定について示す。環境内に 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設定した。

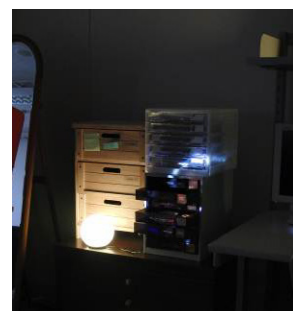
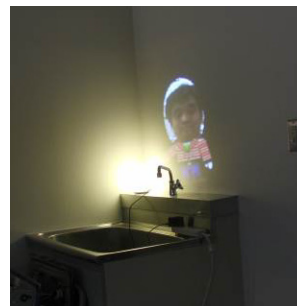


図 5: 投影場所に関するユーザ評価の設定

5 名の評価者に対してユーザ評価を行った。評価者は、設置した物体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つの物体について行う。

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、平坦で単色な背景でなくても指し示す対象の近くに投影するよりも好まれることがわかった。

投影に適した場所への自動移動

投影場所に関するユーザ評価により、アバタ投影に適切な位置は指し示す対象物体の近くであるとは限らず、そのままでは投影が困難な場所を操作者が選択した際にはシステムが自動でアバタを適切な位置に移動する必要があることがわかった。

そこで、クリックした物体のそばが投影に適した位置である場合は従来のままで、そうでない場合は最寄りの投影に適した位置にアバタを移動させ、そこからクリックされた物体に向けて指差しを行う手法を考える。現在のところ、投影に適した位置はリモートシステムの設置時に手動で設定することとしている。将来的には、カメラ画像を処理することにより、投影に適した位置を自動取得する手法を導入することを検討している。

この手法では、ローカルの操作者は物体をクリックするだけでよく、従来の手法で必要であった、投影場所が適切ではない場合の微調整が必要ではなくなる。実際には、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もない。

問題点

しかし、この手法を操作インタフェースに実装し、実際に動作させたところ、ローカルの操作者はどの物体に対しても「これ」という指示語を使い、リモートの対話相手にとってわかりにくいという問題点が明らかになった。物体を指し示す際に効果的な指示を行うために、指示語を利用することは人間にとって自然なふるまいである[7]。しかし、ローカルの操作者とアバタを介して対話しているリモートの対話相手にとっては、アバタから遠くて自分に近い物体は「あれ」、アバタと自分の双方から遠い物体は「それ」を用いて物体を参照するのが自然である。一方、ローカルの操作者は、物体をクリックするだけで操作が可能となったため、画面上の物体の位置を強く意識するようになったためか、上記のような場合でも「これ」を多用する。

解決策

上記の問題を解決策として、ローカルの操作インタフェースを少し変更したものを実装した。変更したインタフェースでは、操作者が物体をクリックすると、自動で投影に適した位置に移動するのではなく、まず投影に適した位置を画面上に赤丸で提示し（図4）、その位置を操作者が改めてクリックすることにより、その位置にアバタが移動するという2ステップの操作手法を採用した。この手法は、ローカルの操作者がアバタの移動先を意識することを促し、アバタの位置を考慮に入れて発話することを狙ったものである。

指示語は発話者のとる視点に強い影響を受ける[8]。また、発話者のとる視点は、条件によって変化することが示されている[9,10,11]。これらのことより、元の手法の問題点は、リモートの対話者はローカルの操作者がアバタの視点をとっていることを期待しているのに対して、ローカルの操作者は目の前にあるコンピュータを操作するため自分の視点で発話をしていることだと考えられる。変更した操作手法では、赤丸を提示してローカルの操作者がアバタの移動先を意識することで、アバタの位置に視点をとることを狙っている。

比較実験

元の手法を自動操作手法・変更した操作手法を半自動操作手法として、2つの手法における指示語の利用を比較する実験を行った。図6に実験の様子を示す。被験者は最初にリモートの部屋に案内され、ローカルの実験者の操作によって行われるPROT AVATARの紹介デモを体験する。体験後、被験者は別の実験者にローカルの部屋に案内される。

ローカルの部屋への案内後、被験者に自動・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。図6にあるように、被験者はリモートの部屋に3つのライトが設置してある様子を画面上で観察できるようになっており、被験者にしばらくアバタ操作の練習してもらった。その後、以下のとおり実験に関する教示を行った。指示語を用いながらリモートの部屋にいる実験者に自動・半自動のいずれかの手法でアバタを操作して、ライトの1つの色の変更指示を出すというタスクを行う。リモートの実験者は被験者の指示に従い、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。

日本語では、発話者の近くにあるものを示すときに「これ」、対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「あれ」という指示語を用いる。図7に各ライトに対してシステムが計算した投影に適切な位置を示す。図7(a),(b),(c)はそれぞれ、リモートの対話相手にとっては「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。

日本人の36名を被験者として実験を行った。7名が女性、29名が男性であり、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、学習効果为了避免の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回数を示す。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。影付きのセルはリモートの対話相手にとって最も自然な発話を示している。自動操作手法よりも半自動操作手法を用いた場合に、対話相手にとって自然に感じられる発話を行った被験者の数が多くなった。自動操作手法と半自動操作手法において自然な発話の割合を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被験者のほうがリモートの対話相手にとって自然に感じられる指示語を用いることを示している。

杉山らは日本語の「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[7]。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」はリモートの対話相手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、対話相手にとってのわかりやすさ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。

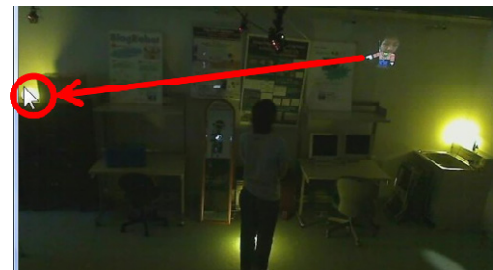


図6: 指示語に関する比較実験の様子



(a) 「これ」条件

(b) 「それ」条件



(c) 「あれ」条件

図7: 各ライトに対するアバタの位置

表1: 指示語に関する比較実験の結果

	「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

最終的な手法と考察

アバタ操作の完全な自動化は身体を持つアバタによる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対話が行われるという実験結果が得られた。ローカルの操作者にとっては、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く、元の手動操作手法に比べるとはるかに手間が少ないといえる。

本論文で示した操作インタフェースの設計プロセスから得られた知見は、アバタ操作の設計において利用することができる。特に、投影位置を適切に選択すること・操作者にアバタの投影位置を意識させるインタラクションデザインを設計することは重要である。

本論文において議論した指示語の問題は、日本語特有の使い回しに起因する。そのため、別の言語でコミュニケーションを行う場合には異なる結果が得られる可能性がある。

まとめ

本論文では身体を持つアバタを利用した遠隔コミュニケーションシステムにおいて、アバタのポインティング動作の操作インタフェースの設計について議論した。プロトタイプの実装した操作インタフェースを実装し、そのプロトタイプを用いたユーザ評価や比較実験の結果をもとに段階的に次のインタフェースを設計・実装したプロセスを記述したことで、関連するコミュニケーションシステムの設計の参考となると考える。特に、指し示す対象からアバタが離れたとしても、アバタが見えやすい背景に投影することをリモートの対話相手が好むため、自動でアバタを適切な位置に移動することが望ましいことと、アバタ操作の完全な自動化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対話が行われるということがわかった。

今後は、アバタの投影に適した場所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評価実験を進めていく予定である。

参考文献

[1] Xbox avatars, <http://www.xbox.com/en-US/live/avatars/>

- [2] Second Life, <http://secondlife.com>
- [3] Kuzuoka, H., Oyama, S., Yamazaki, K., Suzuki, K., and Mitsuishi, M.: GestureMan: a mobile robot that embodies a remote instructor's actions, In Proc. CSCW 2000, pp. 354-354, (2000)
- [4] Ishiguro, H. and Nishio, S.: Building artificial humans to understand humans, JOURNAL OF ARTIFICIAL ORGANS 10, pp. 133-142, (2007)
- [5] Fujimura, R., Nakadai, K., Imai, M., and Ohmura, R.: PROT - An embodied agent for intelligible and user-friendly human-robot interaction, IROS 2010, pp. 3860-3867, (2010)
- [6] 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)
- [7] Sugiyama, O., Kanda, T., Imai, M., Ishiguro, H., and Hagita, N.: Humanlike conversation with gestures and verbal cues based on a three-layer attention-drawing model, Connection science 18, pp. 379-402, (2006)
- [8] McNeill, D.: Psycholinguistics: A New Approach, Harper & Row Press, (1987)
- [9] Ullmer-Ehrich, V.: Speech, place, and action, Wiley Press, pp. 219-250, (1982)
- [10] Klein, W.: Speech, place, and action, Wiley Press, pp. 161-182, (1982)
- [11] 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, pp. 367-384, (2003)