

腕部の提示モダリティがビデオ通話における社会的存在感に与える影響の調査

中村 理歩^{1,a)} 長谷川 実紀² 川口 一画³

概要: ビデオ通話はディスプレイを境界面とした窓越しのコミュニケーションツールである。そのため、遠隔地にいる人と対面している感覚である社会的存在感（ソーシャルテレプレゼンス）が十分であるとは言えない。これまでに社会的存在感を向上するために対話相手の提示方法の影響を調査した研究が報告されている。これらの研究では、実体で動作を提示するロボット会議は従来の遠隔会議より社会的存在感を強化することが示されている。また、ディスプレイから腕が「飛び出す」ように見せる指差し用ロボットアームを映像と同期させると、相手との距離感が近く感じられ、社会的存在感が向上することが報告されている。近年では AR（Augmented Reality）によるロボットの身体的部位の提示方法が、遠隔下でも他者との協働実感を促せる可能性が示唆されている。ただし AR により提示される腕部について、ビデオ通話に追加した場合の効果は不明である。そこで、腕部について物理的に提示される身体部位と AR により提示される仮想的な身体部位に切り替え可能なロボットを設計した。このロボットでは映像によるビデオ通話に加えて、物理的な腕部または AR を用いた腕部を追加提示することが可能である。本研究ではこのロボットを用いて、ビデオ通話における追加腕部の提示が対話相手の社会的存在感に与える影響を調査する。

1. はじめに

ビデオ通話は社会的に広く普及したがその構造上、相手の身体の向きや姿勢、微細な視線変化、空間的距離感といった対面での豊富な手がかりが希薄化しやすい。実世界の対話場面では、言語の意味情報以外のすべてを含む非言語情報が重要な役割を果たすとされている。このような非言語情報が不足した結果、相手と同じ場を共有しているという感覚や、互いに向き合っているという確信が弱まり、社会的存在感の十分な成立が難しくなる。

社会的存在感の伝達では、提示対象の実体性と空間的手がかりが鍵となる。田中らは、映像提示に比して実体の身体動作提示が相手のソーシャルテレプレゼンスを高めることを示し [13]、後続研究でも複数メディア比較によりこの傾向を再確認している [12]。また大西らは、映像枠を越えて腕が受け手側空間に迫る見せ方が、距離感の近接化と指示の伝達性向上につながることを示した [11]。さらに、VR 空間では全身 CG アバタにより高度な非言語表現が容易に実現され [4][2]、物理×AR のハイブリッド検討では、空間参照と社会的存在感には物理頭部が、正確な指差しには

物理腕部が有効であることが示されている [10]。総じて平面映像のみの提示よりも、実体性や空間の手がかりを付与することが社会的存在感を高めると示唆される。

加えて非言語的要素の意義は近年普及が進む仮想現実（VR）空間における CG アバタを介したコミュニケーションでも有効性が示されている。既存のテレプレゼンスロボットやアーム型ロボットに対して、人間の全身を再現した CG アバタを重畳表示する手法が提案されている [4][2]。これらのシステムでは、人間の全身を再現した CG アバタを用いて高度な非言語表現を容易に提示することが可能である。

また川口ら [10] は物理と AR のハイブリッド型ロボットで頭部と腕部の提示方式を切り替えて比較した結果、空間参照と社会的存在感の伝達には物理頭部が重要であり、指差しの正確な伝達には物理腕部が重要であることを示した。大西ら [11] は、ビデオ通話に物理腕を付加することで指示伝達や社会的存在感の向上を示した。また、CG アバタ（AR）によっても非言語表現の伝達が可能であることが報告されている。これらの知見から、腕を AR に置き換えても同様の効果が期待できると考えられる。そこで本研究では、この示唆を拡張し、AR（Augmented Reality）の利用可能性を前提に、ビデオ通話映像に重畳する AR 提示の腕と、現実空間に配置する物理提示の腕を比較検討する。具体的には、(a) 映像のみのベースライン、(b) 映像+物理提示の

¹ 筑波大学 情報学群 情報科学類
² 筑波大学 情報理工学位プログラム
³ 筑波大学 システム情報系
a) s2311869@u.tsukuba.ac.jp

腕, (c) 映像+ AR 提示の腕の三条件を設ける.

また本研究では効果測定は次の4指標で行う. 他者の視線キューに対する印象(知覚のしやすさ/肯定的・否定的影響), 会話体験の主観的評価, 対話相手に対する社会的認知, 共在感(co-presence). 事後確認では, AR 提示と物理実体を相対比較し評価する. また, 提示方法(物理/AR)の違いに関する自由記述も収集する. とくに非言語情報の知覚を各条件がどの程度補完・強調できるかに着目し, 遠隔参加者の存在感と会話への参与への影響を明らかにする.

2. 関連研究

2.1 題名1

実世界のコミュニケーションでは, 非言語情報が会話の開始や進行を左右する重要な役割を担う. ここでいう非言語情報とは, 言葉そのものの意味以外のあらゆる情報を指す. たとえば視線や身体配置は, 参加者間の関心や注意の向きを示すだけでなく, 相互の関係性や会話の進み具合を示すシグナルとしても機能する. ケンドン [3] は, 視線および身体の向きが会話の展開に決定的な影響を与え, 相手の注意喚起にも寄与することを示している. また, 視線や身体の向きは相手に対する積極性や関心の表明ともなり, その結果, 会話はより自然に進みやすくなる. さらにホール [1] は, 相手との身体的距離(パーソナルスペース)が社会的関係性を示す重要な要因であることを明らかにしており, 非言語的要素が会話の進行に深く関与していることが確認されている.

2.2 題名2

ロボット通話を用いた一連の研究は, 身体動作を物理実体として提示することの有効性を実証している. 田中ら [13] は, 遠隔者の身振りを物理的に再現して伝達することで, 映像のみの提示に比してソーシャルプレゼンスが強化されることを報告し, 後続の論文 [12] でも, 物理実体を介した提示が会話の自然さや参与感の向上に寄与することを詳細に論じている. 加えて, 大西ら [11] は, 腕を映像化することによって, 全身を仮想的に提示しなくても, 腕などの要部を強調するだけで存在感を効果的に高められることを示した. さらに, 川口ら [10] は, 物理的提示と AR 提示を併用するハイブリッド型ロボットにおいて, 頭部と腕部という部位ごとに提示方法の違いが知覚・行動・関与へ与える影響を比較し, 部位特性×提示モダリティ(物理/AR)の相互作用が設計上の要点であることを明確化した. これらの知見を総合すると, 腕部などの要部を選択的に強調する部分的実体化, 物理実体による奥行き・接触可能性・空間共有の手掛かり, 部位に応じた AR/物理のハイブリッド運用, という三つの設計原則が, 平面映像のみの提示に比して社会的存在感を有意に高める有望な方策であると位置づけられる. 以上より, 今後の遠隔コミュニケーション

設計では, 場面や制約に応じてどのモダリティで提示するかを最適化することが, 会話の開始・進行の円滑化と参与感の維持に不可欠である.

3. システム

本研究では, ビデオ通話に腕を付与したシステムにおいて, 腕の提示モダリティによる影響を調査する. そのため, 腕の物理的提示と AR 提示を切り替え可能なビデオ通話システムを実装した.

3.1 システム構成

川口ら [10] のシステムを参考に物理的な腕と AR 提示を切り替え可能なシステムを実装した. システム構成を図 1 に示す.

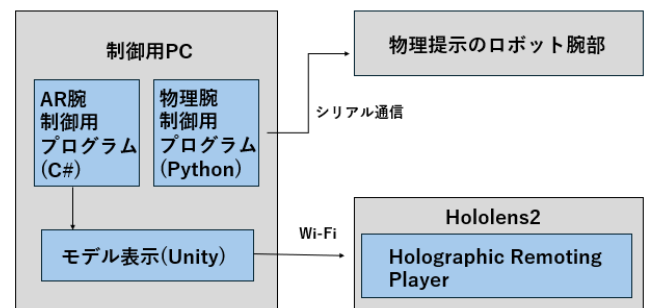


図 1 システム構成図

本システムは, 制御用 PC を中核に, AR 提示と物理提示の両モダリティを同一基盤で扱う構成である. 物理提示では, 制御用 PC からシリアル通信を介し, Python で実装した腕制御プログラムによりロボット腕を駆動する. AR 提示では, C # で実装した腕制御プログラムを Unity [8] 上で動作させ, Wi-Fi 経由で HoloLens2 [6] の Holographic Remoting Player [5] に送出し, 現実空間へ重畳表示する. これにより物理腕と AR 腕を切り替え, 両条件の比較実験が可能となる.

3.2 物理提示

物理提示の腕は, 本研究で開発した実験用システムの 1 つであり, 実世界における実体的の手がかりを付与することを狙いとしている. 腕部機構は 3D プリンタで設計・出力した骨組みにサーボを組み合わせて構成し, 遠隔参加者の指差しを再現可能とした. このロボットは片腕あたり肩 3 自由度と肘 1 自由度を有する. をこの物理提示の腕を図 4 に示す. これにより, 受け手は机上の実体として腕部の動きを知覚可能である.

3.3 AR 提示

AR 提示の腕はを用いて, ビデオ通話映像に 3 次元の腕部を重畳表示する手法である. Python によりロボットモデル

およびそのモーションを設計し、その動作データを Unity[8] に取り込み画面上へ描画する。受け手は HoloLens2[6] を装着し、現実空間に存在する通話画面に遠隔参加者の腕部を AR で重畳提示された状態で観察することができる。この物理提示の腕を図 4 に示す。

4. 実験

本研究では、物理提示の腕と AR 提示の腕を腕部提示のモダリティの違いが遠隔参加者の社会的存在感および会話の開始・進行に与える影響を明らかにする。通常の映像のみのビデオ通話に 2 手法を追加提示し、同一の対話タスク上で比較し実験を行った。本章では実験設計、および評価項目を述べる。

4.1 実験設計

本研究では、ビデオ通話映像に非言語的の手がかりを付加する腕部提示手法を比較するため、下記 3 条件を設定した。

4.1.1 タスク

参加者は遠隔対話ルームに参加し、遠隔操作者と 1 対 1 で対話する。参加者と遠隔操作者の様子を図 5 と図 6 で示す。はじめに遠隔操作者が特定地域のお土産 2 種について説明、その後参加者に「どちらを選ぶか」を質問する。参加者が回答した後、選択の理由を追加で質問する。遠隔操作者の発話内容はあらかじめスクリプトを用意し読み上げることで統制を図った。指さしの提示方法は以下の通りである。

- **提示タイミング** スクリプト中の指示語「こちら」または「右/左」に合わせ、対象側を指差しで示す。参加者の回答後は、必要に応じて短い再指差しや静止姿勢で間を取り、再選択プロンプトに接続する。
- **操作手段** 遠隔操作者がキーボードのショートカットで各モーションを行う。同一の台本に同期して、物理条件では実機腕のクリップ再生、AR 条件では Unity[8] 上の 3D 腕モーション再生を行い、動作内容は条件間で共通である。

4.1.2 実験条件

遠隔対話は Zoom[9] を用い、制御用 PC(Unity[8]/Visual Studio Code[7] 実行)、参加者用 PC、および HoloLens2[6] を同一 Wi-Fi ネットワークに接続して実施した。開始前に HoloLens2[6] の視線キャリブレーションを行い、表示の整合性を確保した。いずれの条件でも参加者は HoloLens2[6] を装着し、説明内容は遠隔操作者用スクリプトで統一した。このスクリプトはお土産の二択提示と理由の聴取を行う内容である。

- **C1. 映像のみ** 机上の iPad に Zoom[9] 映像を表示し、遠隔話者は音声のみで左右指示を行う。HoloLens2[6] は装着するが、腕提示は行わない。
- **C2. 物理条件** 参加者視点で、iPad 上の遠隔話者の身体に実体ロボット腕が接続されて見える位置に設

置する。制御用 PC から Python スクリプトを起動し、シリアル通信でサーボを駆動して指差しを再生する。HoloLens2[6] は装着するが、腕提示は行わない。

- **C3. AR 条件** 制御用 PC 上で Unity[8] を実行し、C # 実装の AR 腕制御プログラムにより 3D 腕モデルを描画、Wi-Fi 経由で HoloLens2[6] 内の Holographic Remoting Player[5] へ送出して現実空間に重畳表示する。腕モーションは C2 と同等の指差しジェスチャを提示する。

この構成により、対話タスクを共通としたまま、腕のモダリティだけを切り替えて比較できる。

4.1.3 参加者

参加者は大学内の学生・大学院生を対象に募集し、各参加者には 3 条件をすべて 1 回ずつ実施してもらった。事前に氏名・年齢・性別・VR 利用経験等を調査した。3 条件ともに合計 6 名。(男性 5 名、女性 1 名、平均 22.33 歳、SD = 1.37)

4.1.4 実験手順

(1) 目的説明と同意取得後、事前アンケートを iPad で実施した。(2) システム準備として Zoom[9] の接続、録画開始、ロボットの起動確認を行う。AR 条件では HoloLens2[6] の視線キャリブレーションと位置合わせを実施する。(3) 練習として音声のみで短い練習タスク(沖縄)を行い、通話品質と手順に慣れてもらう。(4) 本試行で割付順に各条件を実施する。セッション中、遠隔操作者はスクリプトに従って対象の左右を示し、参加者はお土産の選択と理由を述べる。各セッションの上限は約 5 分とした。(5) セッション終了ごとに条件別アンケートに回答してもらう。録画は 2 条件ごとに停止・新規開始してファイルを分割で保存した。(6) 全条件終了後、事後アンケートを実施した。

4.1.5 実験手順

実験手順は以下の通りである。

- **事前説明・同意・事前アンケート** 目的説明に同意後、iPad で事前アンケートに回答する。
- **HoloLens2[6] の装着** 参加者は全条件共通で HoloLens2[6] を装着する。
- **視線キャリブレーションと位置合わせ** 表示整合性を確保するため、HoloLens2[6] で視線のキャリブレーション実施する。
- **練習タスク** 音声のみの短い練習課題で対話手順に慣れる。
- **本試行(割付順に各条件を実施)** 各セッションで
 - (1) 遠隔操作者が特定地域のお土産 2 種を説明、
 - (2) 参加者は「どちらを選ぶか」を回答。
 - (3) 続けて選択理由を口頭で回答する。
 - (4) 遠隔操作者は上記の回答ともう一方の選択理由を説明する。
 - (5) 参加者は再び「どちらを選ぶか」を回答。

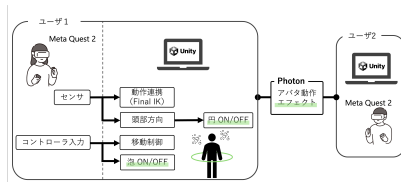


図 2 映像のみの写真

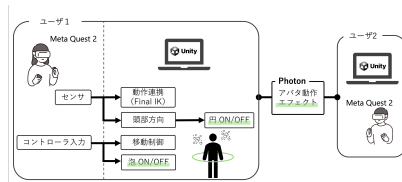


図 3 物理の写真

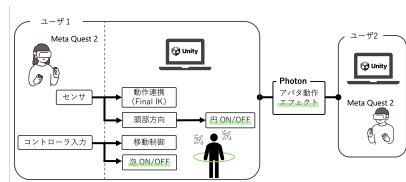


図 4 AR の写真



図 5 参加者



図 6 遠隔操作者

図 7 参加者と遠隔操作者の写真

1セッションの上限は約5分とする。

- **条件別アンケート** 各セッション終了後、該当条件に関するアンケートに回答する。
- **事後アンケート** 全条件終了後、事後アンケートに回答する。

4.2 評価項目

川口ら [10] の評価項目を参考に本研究において、以下の5つの評価項目を設定した。

● 指差し知覚

遠隔ユーザの指差し方向の変化に気づいた、指しているカードの同定が容易だったを7段階で評価。

● 社会的存在感・相互注意

互いに気づいた／存在が明らかだった、注意を引いた／引かれた、他事で気が散りやすい、対話中に集中し続けた、などを7段階で評価。

● 相互理解の主観評価

自分／相手の考えが明確だった、理解しやすかった、理解が難しかったを7段階で評価

● 知覚品質評価 (AR-物理の比較)

腕提示の画質 (解像度)・明るさ (輝度)・動作の滑らかさを、AR と物理を対比させた7段階の両極アンカーで評価。

● 自由記述

提示方式の違い (物理的実体 vs AR 提示) で気づいた点や改善提案を収集。

4.2.1 他者の視線キューに対する印象評価、これ以下は分析後かなと

本実験では、他者からの視線キューに対する参加者の印象を評価するために、Ichino?らの研究を参考に、3つの下位尺度 (知覚?, 肯定的?, 否定的な影響?) を設定した。Ichino らの文献?に基づくアンケートは、他者からの視線キューに対する印象に関して、4つの尺度を評価する計14個の質問から構成されている。本実験においては、本実験と無関係と判断した質問項目を除外し、3つの下位尺度の設定および計6個の質問を行い、それぞれ「非常に当てはまる」から「全く当てはまらない」の7段階のリッカート尺度により回答させた。

4.2.2 対話相手への社会的認知に対する主観評価分析後変更

本実験では、対話相手への社会的認知に対する印象を評価するため、アンケートを行った。Zhang?らの研究を参考に、計4項目の質問を用いた。これらの質問項目は、参加者が対話相手をどのように認識したかを直接評価することを目的とする。各項目に対する回答は、「全く当てはまらない」から「非常に当てはまる」の7段階のリッカート尺度を用いて収集した。

4.2.3 会話体験に対する主観的評価分析後変更

本実験では、参加者の会話体験に対する印象を評価するため、アンケートを行った。Zhang ら?の研究を参考に、本実験にて関係があると判断した計7項目の質問を用いた。1から5の質問に対する回答は、「全く当てはまらない」から「非常に当てはまる」の7段階のリッカート尺度を用いて収集した。また、6および7に対する回答は、「非常にネガティブに変化した」から「非常にポジティブに変化した」の7段階のリッカート尺度を用いて収集した。

4.2.4 共存在感 (Co-presence) に対する主観的評価分析後変更

本実験では、参加者が VR 空間内で対話相手とどの程度共に存在していると感じたかを測定するために、アンケー

トを行った。Bailenson 等の研究を参考にし、計 4 項目の質問を用いた。共存在感 (Co-presence) は、社会的存在感の一要素として、VR 空間における対話相手との相互作用のリアリティを測定する指標である。各質問項目に対する回答は、「全く当てはまらない」から「非常に当てはまる」の 7 段階のリッカート尺度を用いて収集した。

参考文献

- [1] Hall, E. T.: A System for the Notation of Proxemic Behavior, *American Anthropologist*, Vol. 65, No. 5, pp. 1003–1026 (online), DOI: 10.1525/aa.1963.65.5.02a00020 (1963).
- [2] Jones, B., Zhang, Y., Wong, P. N. and Rintel, S.: Belonging there: VROOM-ing into the uncanny valley of XR telepresence, *Proceedings of the ACM on Human-Computer Interaction*, Vol. 5, No. CSCW1, pp. 1–31 (2021).
- [3] Kendon, A.: Some functions of gaze-direction in social interaction, *Acta Psychologica*, Vol. 26, pp. 22–63 (online), DOI: 10.1016/0001-6918(67)90005-4 (1967).
- [4] Luo, L., Weng, D., Hao, J., Tu, Z. and Jiang, H.: controllable telepresence: A robotic-arm-based mixed-reality telecollaboration system, *Sensors*, Vol. 23, No. 8, p. 4113 (2023).
- [5] Microsoft: Holographic Remoting Player, <https://apps.microsoft.com/detail/9nblggh4sv40>. Accessed 2025-11-03.
- [6] Microsoft: HoloLens 2, <https://www.microsoft.com/hololens>. Accessed 2025-11-03.
- [7] Microsoft: Visual Studio Code, <https://code.visualstudio.com/> (2025). (Accessed on 11/01/2025).
- [8] Technologies, U.: Unity Real-Time Development Platform — 3D, 2D VR & AR Engine, <https://unity.com/> (2025). (Accessed on 11/01/2025).
- [9] Zoom Video Communications, Inc.: Zoom, <https://zoom.us>. Accessed 2025-11-03.
- [10] 川口一画, 井原圭一, 市川あゆみ, 佐方葵, 守新太郎ほか: 物理的提示と AR 提示を併用するハイブリッド型ロボットにおける頭部と腕部の提示方法の違いによる影響, *情報処理学会論文誌*, Vol. 65, No. 3, pp. 706–717 (2024).
- [11] 大西ほか: 身体映像の部分的実体化によるソーシャルテレプレゼンスの強化, *情報処理学会論文誌*, Vol. 57, No. 1, pp. 228–235 (2016).
- [12] 田中, 一晶, 中西, 石黒浩: ロボット会議: 物理的実体を介した身体動作の提示によるソーシャルテレプレゼンスの強化, *情報処理学会論文誌*, Vol. 57, No. 1, pp. 209–217 (2016).
- [13] 田中一晶, 中西英之, 石黒浩: 実体で身体動作を提示するロボット会議によるソーシャルテレプレゼンスの強化, *Interaction (13INT017)* (2013).