

# 近未来のワークプレイスを創り出す ソーシャルテレプレゼンス技術

中西 英之

(大阪大学大学院准教授)

ビデオ会議システムなどを用いつつ地理的に離れたワークプレイス間で共同作業を行う場合、遠隔にいる共同作業者のプレゼンス（存在感）が希薄化する。この問題はビデオ会議システムのディスプレイをロボット技術で拡張することによって緩和可能である。なぜならロボットはディスプレイ上のカメラ映像とは異なり物理的な実体を有しており、その実体を動かすことで遠隔にいる他者の身体の動きを表現することができ、それによってプレゼンスを伝えることができるからである。本稿ではディスプレイをロボット技術で拡張することでソーシャルテレプレゼンス（実際には離れた場所にいる他者とあたかも同じ場所にいるかのような感覚になることができる現象）の創出能力を向上させる3つの方法を紹介する。1つ目はディスプレイやそこに取り付けられているカメラをロボット化して移動できるようにすることで、他者や自分が同じ部屋の中において歩いているかのような感覚を創出する方法である。2つ目は身振り手振りなどの身体動作や握手などの身体接触に用いられる身体の一部をロボットで再現してカメラ映像に結合することで、他者が自分の目の前にいて動いているかのような感覚を創出する方法である。3つ目はテーブルや椅子などの体が触れる物体をロボット化して動くようにすることで、自分のそばにある物体を他者が触って動かしているかのような感覚を創出する方法である。これらの方法は全て、カメラ映像がディスプレイという平面の上に描かれた単なる絵に過ぎないという印象を、ロボットの使用によって打ち消すものであると言える。

## 目次

- I はじめに
- II ロボットとソーシャルテレプレゼンス
- III テレプレゼンスロボット
- IV ロボットハンド
- V ロボット家具
- VI おわりに

## I はじめに

地理的に離れたワークプレイス間での共同作業をビデオ会議（テレビ会議、ビデオチャット）などの音声映像通信技術によって支援する研究が古く

から行われてきている（Gaver et al. 1993; Heath and Luff 1991）、これらの研究では、遠隔地のワークプレイスの様子、特に机の上などの作業領域の様子が見えるようになることは作業効率の向上にとって重要である一方、遠隔地にいる共同作業者の様子が見えるようになることは、それが作業内容に深く関わる場合を除いて、作業効率の向上には結び付きにくいと結論付けられている。作業効率の向上に結び付くことを示した研究もあるが（Olson, Olson and Meader 1995）、その効果はあまり明確ではないと言える。そこで、互いの様子の共有ではなく、共同作業における資料の共有やその資料への書き込みの共有を目的としたビデオ会

議システムとして、カメラで撮影した互いの机の上の様子を他方の机の上にプロジェクタで重畳的に投影する技術が盛んに開発された (Tang and Minneman 1991)。例えば ClearBoard (Ishii and Kobayashi 1992) ではホワイトボードのように利用できるガラスの板を挟んで2人で議論をする状況を遠隔地間で再現できるようにしており、Agora (Luff et al. 2006) では遠隔地にいる議論相手の映像を正面に表示するとともに、互いの書類やその書類への指差し・書き込みを机の上で共有できるようにしている。これらのシステムにおいて共同作業者の様子を映し出す目的は、資料のどこを共同作業者が見ているのか (Gaze Awareness)、資料のどこを共同作業者が指差しているのか (Pointing Gestures)、といった非言語的の手掛かり (Nonverbal Cues) の取得を可能にすることである。

上記のような作業効率に着目した資料共有を主たる機能とするビデオ会議システムの研究と並行して、離れた場所にいる作業者同士があたかも同じ部屋で一緒に作業しているかのような感覚を生むための研究、すなわち資料ではなくプレゼンス (存在感) の共有に着目した研究も盛んに行われてきた。この感覚の創出には、他者の様子を映し出すカメラ映像が大きな役割を果たすことが知られており (Isaacs and Tang 1994; de Greef and Ijsselstein 2001)、遠隔地との間で視線を一致させること (アイコンタクト) が可能な時や (Bondareva and Bouwhois 2004)、映像が実物大 (等身大) である時や (Prussog, Mülbach and Böcker 1994)、顔だけではなく体 (上半身) も含めて映し出せる場合に (Nguyen and Canny 2009)、この感覚が強まることが分かっており、それらを可能にするシステムが盛んに研究されてきた。これらは基本的に、映像を通してしか見ることのできない遠隔地にいる他者を、同じ部屋に物理的に存在するのと同じ状態に可能な限り近付ける VR (バーチャルリアリティ) 技術の研究であるとみなすことができるが、双方向で使用しても問題の無いユーザーインタフェースが求められる点が通常の VR 技術とは異なる。例えば、リアリティの向上を目的としてヘッドマウントディスプレイのようなウエ

アラブル機器を使用すると、それらの機器の装着がカメラ映像を通して他者に伝わる見た目を変化させてしまい、そのような見た目の不自然さがリアリティを損なうという本末転倒な結果となってしまう。このような制約から単に既存の VR 技術を利用するだけでは済まず、アイコンタクトを実現するだけでも大掛かりな装置の開発が必要となる (Nguyen and Canny 2007)。

本稿では近未来においてワークプレイスを変容させる様々な技術の中でも特にプレゼンスの共有を可能にする技術に焦点を当て、ソーシャルテレプレゼンス (メディアを使用することによって実際には離れた場所にいる他者とあたかも同じ場所にいるかのような感覚になることができる現象) を生み出すための技術に関して議論を行う。その理由は、遠隔会議による対面会議の代替や、それによるワークプレイスの分散化においてボトルネックとなっているものは、単に会議や共同作業の効率が低下することではなく、場所を共有していないことから生じる様々な心理的影響であるからである。このような心理的影響に関する古典的な研究に、同じ場所にいる人間同士の間で自然に発生するような雑談 (インフォーマルコミュニケーション) を、異なる部屋や異なる建物の間などの遠隔地間において発生させることを試みた研究がある。これらの研究ではメディアスペースと呼ばれる常時接続型のビデオ会議システム (Bly, Harrison and Irwin 1993) が用いられたが、そのような単なる音声映像通信によって部屋と部屋をつなげるだけではインフォーマルコミュニケーションを発生させることは困難であり、他者と同じ場所にいる環境 (対面環境) よりも発生頻度が減少する結果となった (Fish, Kraut and Chalfonte 1990)。メディアスペースの問題点として、画面の向こう側にいる他者の身体動作や行動に気付にくいことや (Heath and Luff 1992)、画面の向こう側にいる他者よりも同じ部屋の中にいる他者に対してより話しかける傾向が生じてしまうこと (Mantei 1991) が挙げられている。

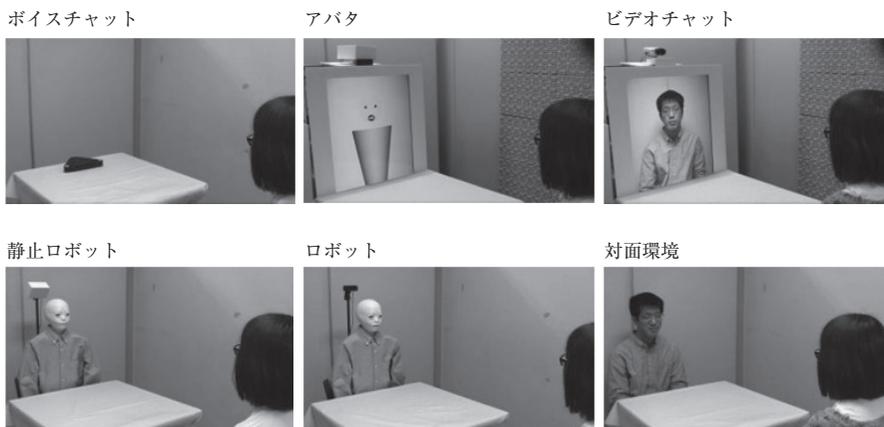
## II ロボットとソーシャルテレプレゼンス

上で議論したように代表的なソーシャルテレプレゼンスの創出手段は、遠隔地にいる他者のカメラ映像をディスプレイに表示する手法であるが、これ以外に創出手段として研究されている代表的なものに遠隔操作人型ロボットがある。遠隔操作人型ロボットは将来的には人間を遠隔地に仮想的に転送すると表現できる程にまでその見た目や動きのクオリティが向上する可能性のある技術と考えられているため有望視されており、盛んに研究されている。ヒューマンロボットインタラクションの研究分野ではロボットが持つ物理的身体の効果が検証されており、物理的に現実空間に存在するロボットの動きは映像上の変化にのみ基づくコンピュータグラフィックスキャラクタの動きよりも高いプレゼンスを生むことが判明している (Lee et al. 2006; Bainbridge et al. 2011)。ロボットはディスプレイ上の映像とは異なり、物理的な物体の運動によって身体動作を表現できる上に、それによって握手、ハイタッチ、肩をたたく、などの相手の体に触れる身体接触をも擬似的に可能にするという利点を持つ。しかしその一方で、ロボットとして表現されている他者、すなわち遠隔地にいるロボット操作者の見た目を提示することが困難であり、そのためカメラ映像と同様にはプレゼンスを伝達することができないという問題がある。この問題に対してロボットの見た目のクオリティを高めるアプローチで挑んだものに、実在

の人間に酷似した見た目を持つロボット (アンドロイド) の研究がある (Sakamoto et al. 2007)。従来のいかにも機械っぽいロボットに比べると非常に高度な迫真性を与えるものであり、そのようなロボットで操作者の見た目を再現することは有効かもしれない。しかしながら、製作に多大なコストと時間が必要である上に、あくまである時点における姿形を静的に模して作られるものであるため、カメラ映像よりも写真に近く、言わば動く蠟人形であり、見た目の「再現」と「伝達」は根本的に異なることを考えると、抜本的な解決策にはならないと思われる。

このようにソーシャルテレプレゼンスにおけるロボットの有効性を考えた場合、物理的身体を持つことに効果があるのかが重要な疑問となる。そして仮に効果があるとして、その原因は純粹に物理的に存在することにあるのか、それともその物理的な身体を動作の伝達に利用することにあるのか、どちらなのかという疑問が湧く。さらに、見た目を伝えることができないことはどれくらいデメリットとなるのか、ということもロボットの有効性を考える上で重要である。これらの疑問に答えるために、物理的身体・身体動作・見た目の有無の影響を調べるための図1に示すような6条件の比較実験を実施した (Tanaka, Nakanishi and Ishiguro 2014)。物理的身体の有無を第一の要因とし、身体動作と見た目の両方を伝える場合、身体動作のみ伝える場合、どちらも伝えない場合の3段階の情報伝達方式を第二の要因

図1 身体・動作・見た目の影響を調べる実験



とした。なお全条件において音声は伝えるようにし、簡単な音声対話を実験タスクとした。

図1左上の条件（物理的身体なし，身体動作なし，見た目なし）はボイスチャット（電話）に対応しており，右上（物理的身体なし，身体動作あり，見た目あり）はビデオチャット（テレビ電話）に対応している。下段中央（物理的身体あり，身体動作あり，見た目なし）が調査の焦点となるロボット条件であり，そこから物理的身体を引いたものが上段中央（物理的身体なし，身体動作あり，見た目なし）のアバタ条件である。これは，ビデオチャットツールにしばしば付属している顔の部分をキャラクタに変える画像処理機能を使用している状況に近い。VTuberとボイスチャットをしている状況とも言える。このようなアバタを使用することによって，単純なボイスチャットよりもプレゼンスが伝わりやすくなることが確認されている（Garau et al. 2001; Kang, Watt and Ala 2008）。ロボットとアバタは同じフェイストラッキング機構で動くようになっており，被験者の対話相手を務めた実験者の首の傾き（ロール・ピッチ・ヨーの3自由度）と唇の動き（開閉の1自由度）を反映させた。右下（物理的身体あり，身体動作あり，見た目あり）は，ビデオチャットに物理的身体を加えたものと解釈して，実験者そのものが登場する条件，すなわち対面環境とした。ロボットに見た目を加えたアンドロイド条件と解釈して実験を行うことも考えられるが，ソーシャルテレプレゼンス創出能力についてはアンドロイドよりも対面環境のほうが明らかに上であるので，まずは対面環境との比較を行った。左下（物理的身体あり，身体動作なし，見た目なし）は，ただ単に物理的身体が目前にあることに意味があるのかどうかを試すための条件である。フェイストラッキングによって動かすことを一切行わない点以外はロボット条件と同一であり，伝達する情報はボイスチャット条件と同一である。

以上の6条件を比較した結果，2要因ともにソーシャルテレプレゼンスに対して有意な影響を与えた。具体的には，物理的身体を持っているほうが，身体動作を伝えるほうが，見た目を伝えるほうが，よりプレゼンスが伝わりやすくなった。これ

は予想していた結果であるが，意外だったのはロボット条件とビデオチャット条件がほぼ同レベルのプレゼンス伝達能力を示したことである。ロボットは見た目の情報を伝達しないにもかかわらず，物理的身体を持つというアドバンテージによって，カメラ映像に匹敵するソーシャルテレプレゼンス創出手段になりうるということである。

### III テレプレゼンスロボット

ロボットはカメラ映像と同等のソーシャルテレプレゼンス創出手段になりうる。なぜならロボットはその物理的身体によって，目の前に実際に人間（のようなもの）が存在するという感覚を与えることができるからである。しかしながら，そのロボットが伝えようとしているプレゼンスの出所となる人間（遠隔地にいる他者＝ロボット操作者）の今現在の見た目をリアルタイムで反映することができないため，カメラ映像を完全に代替する手段にはなり得ない。ディスプレイ上のカメラ映像には，どこか別の場所にいる実在の人間の今現在の見た目が映し出されていることによる存在感があるが，ロボットのような物理的実体に基づいた存在感はない。そこで，ロボットとカメラ映像を融合することによって，それぞれの利点を併せ持つ新しいメディアを作り出すことが望ましいと考ええる。

ロボットとカメラ映像を融合させたメディアの代表的な例としては，既に多数の製品が販売されているテレプレゼンスロボットがある。テレプレゼンスロボットと呼ばれる製品の中にはカメラ映像を用いておらず，IIで述べたロボット条件と同様にロボット自身の身体動作が視覚情報の全てであるものもあるが，基本的にはロボットとカメラ映像が組み合わさった構成をしている。まず重要な構成要素として遠隔操作可能な台車ロボットがあり，その上にロボット操作者の顔を映し出すためのディスプレイが搭載されており，それらに加えて音声対話のためのマイクロフォンおよびスピーカと，ロボットが動き回る遠隔地の様子をロボット操作者が眺めるためのカメラが備わっている。つまり，テレプレゼンスロボットとは通常の

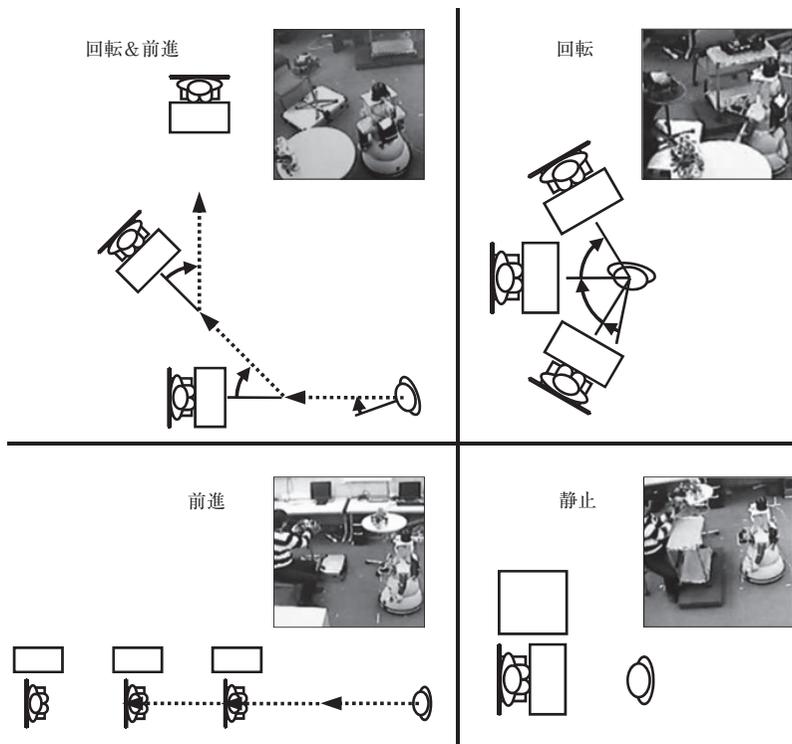
ビデオ会議システムに対して遠隔地の物理空間を動き回る機動性を与えたものである。

カメラ映像の中の人間は現地にいる人間のように部屋の中を歩き回ることにはできないので、例えば歩いて誰かに近付いて話しかけたりすることはできない。遠隔地にいる人間も現地にいる人間と同じように振る舞うことができるようにしないと対面環境を再現したとは言えず、「I はじめに」で述べたインフォーマルコミュニケーションの再現も難しいであろう。現地の部屋を動き回ることのできる機動性をビデオ会議システムに付与したテレプレゼンスロボットの開発機はおそらくこのようなものであり、実際にそのような機動性の有効性が過去の研究で示されている (Paulos and Canny 2001)。しかしながら、そのような有効性は機動性が有用な特定の状況において機動性が問題無く使用されたという自明に近いものであり、機動性が遠隔地間のコミュニケーションに及ぼす具体的な影響はほとんど明らかになっていない。例えば、廊下でばったり会った相手とその場で雑談をする、という偶発的出会いの再現には機動性

が必須のように思われるが、そのような状況をオンラインで再現するシステムがいくつも考案されており (Root 1988; Nakanishi et al. 1999), 偶発的の出会いを再現するためだけにテレプレゼンスロボットを使用する意義は薄い。仮にテレプレゼンスロボットの機動性が、廊下をうろついて誰かと会うためだけに必要であり、会った後に立ち止まって (停止して) 雑談を始めた後は通常の動かないビデオ会議システムを使っているのと全く同じなのであれば、その機動性を上記の偶発的出会いを再現するシステムに置き換えても良いことになる。つまり、機動性に着目してロボットとカメラ映像の融合の有効性を考えた場合、機動性そのものの有用性のみならず、機動性がコミュニケーションにどのような影響を及ぼすのかが問題となる。

図2に示す実験は、テレプレゼンスロボットの機動性がカメラ映像のソーシャルテレプレゼンス創出能力を強化するのかどうかを検証したものである (Nakanishi et al. 2008)。この実験ではテレプレゼンスロボットの遠隔操作における左右回転の

図2 テレプレゼンスロボット実験



有無と前後移動の有無の2つを要因とする4条件の比較を行い、ロボットに搭載されたカメラを通して見る相手のプレゼンスの強さを調査した。多くのテレプレゼンスロボットは左右回転と前後移動の組み合わせで動き回るようになっていて。このうち、前後移動にソーシャルテレプレゼンス強化の効果があるという仮説を立てた。なぜなら過去の研究では両眼視差を発生する立体映像に効果のあることが示唆されており (Prussog, Mülbach and Böcker 1994), 同様に奥行き情報を伝える視差である運動視差を発生させることになるカメラの位置の変化にも効果があるのではないかと考えたからである。被験者に与えられたタスクは3つのテーブルそれぞれに置かれたロボットトイの説明を聞くというもので、テーブルの配置を変えることによって被験者が操作するロボットの移動軌跡が4種類発生するようにした。回転&前進条件では右回転と前進を組み合わせることでテーブル間を移動する必要があったが、回転条件ではテーブルが円環状に並んでいたため右回転だけで済み、前進条件ではテーブルが一直線に並んでいたため前進だけで済んだ。静止条件では同じテーブルで3つのロボットトイの説明全てが行われたので、遠隔操作の必要は無かった。前進条件と回転&前進条件は軌跡に前進を含んでおりカメラの位置が変化

するので、ソーシャルテレプレゼンスが強化されると予想し、結果はその通りになった。映像上の変化は小さいが運動視差を生じさせる「前進」には効果が見られ、映像変化は大きい運動視差を生じさせない「回転」には効果が見られなかった。この結果に加えて、ロボットが自動的に移動して被験者が操作しない場合には、たとえ軌跡に前進が含まれていても効果が発揮されないことを確認した。この原因は、自分で操作していないために移動先を予測できず、それによってロボット(に搭載されたカメラ)の移動を自分自身(の視点)の移動と感じられなくなるためと思われる (Vogeley and Fink 2003)。

上述の実験ではロボットの移動がソーシャルテレプレゼンスを強化した。しかしながら、同様の効果を発揮させるために果たしてロボットは必要なのだろうか。もっと機能が限定された単純なデバイスでも同じ効果を出せるかもしれない。図3に示す実験はこの問いに答えたものであり、メディアスペースに対して運動視差を発生させるための最小限の機能拡張を施したものによってソーシャルテレプレゼンスの強化を試みた (Nakanishi, Murakami and Kato 2009)。具体的には、2つの部屋を繋いでいるメディアスペースにおいて、一方の部屋にいる利用者のディスプレイへの接近に連

図3 ロボットカメラ実験



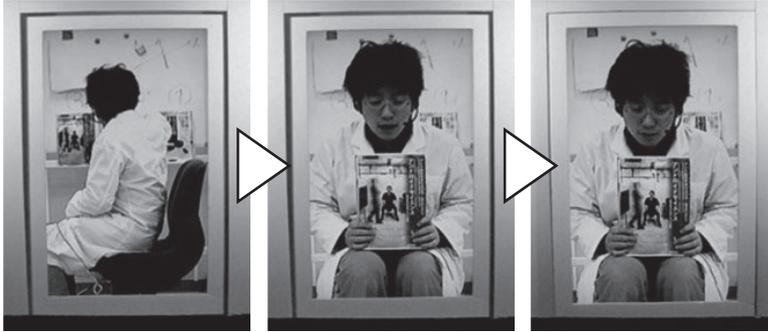
動して、もう一方の部屋に置いてあるカメラが前進するようにした。この仕組みによって、ある利用者がディスプレイを眺めながら近づくと、その利用者は自分の歩行に連動した運動視差を伴う遠隔の部屋の映像を見ることになる。図2の実験で用いたシステムとの違いはカメラを前進させる方法のみであり、ロボットを操作するのか、それとも自分で歩くのか、という点である。通常のビデオ会議システムは着席して会議に参加している状態で使用するものであり、従って利用者がディスプレイに近づくとという状況は基本的に発生しない。一方、メディアスペースではたいがい大型のディスプレイが用いられ、そこに遠隔の部屋の風景が常時映し出されているので、それを良く見るために利用者が歩いてディスプレイに近づくと、という状況はしばしば発生する。実験ではこのような状況を想定し、被験者のディスプレイへの接近を距離センサで検出し、その距離の変化に同期して3種類の映像変化を発生させ、それらと映像変化が無い条件の4条件を比較した。被験者に与えられたタスクは別の部屋にいる説明者から3つのロボットトイの説明を聞くというもので、各説明の前の準備の間、ディスプレイから離れたところにある椅子で休憩してもらい、説明を開始する度にディスプレイの手前まで歩いて来てもらった。よって、被験者は映像変化を計3回見るようになった。映像変化の内容については一切事前に説明しなかった。図3の前進条件ではカメラが位置決めテーブルに取り付けられており、ディスプレイへの接近に同期してカメラが前進し、説明者の映像が2倍に拡大した。ズーム条件ではカメラを前進させる代わりに、運動視差を生じさせない光学ズームを用いて映像を2倍に拡大した。無変化条件の映像は常に2倍に拡大された大きさの状態であった。映像変化は前進条件と全く同じであるが、カメラは説明者のリモコンによって操作され、被験者の歩行と同期することなく前進するリモコン条件というものも試した。

クオリティの観点からは映像が常に拡大された状態にある無変化条件がもっとも優れており、従ってプレゼンスの伝達においてもっとも有利である。それにもかかわらず前進条件では無変化条

件に比べてプレゼンスが伝わりやすくなった。一方、前進条件と全く同様の映像変化を与えたりリモコン条件や、同様に映像が拡大したズーム条件においては、無変化条件との間にプレゼンスの差は見られず、前進条件よりもプレゼンスが伝わりにくいという結果となった。このように前進条件では、カメラの移動を被験者が操作したわけではないにもかかわらず、ソーシャルテレプレゼンスを強化する効果を確認できた。おそらくは、被験者自身の歩行が運動視差を生じさせている感覚、すなわちディスプレイに映し出されている向こう側の部屋の中を歩いている感覚にさせることができたためと思われる。同様に運動視差が生じて、それが歩行に起因するものではなかつたりリモコン条件では効果を確認できず、また同様に映像が拡大するが運動視差は生じなかつたズーム条件でも効果を確認できなかった。

上述の実験ではカメラに前進機能を付与するという最小限の機能拡張によって、テレプレゼンスロボットと同様のプレゼンス伝達能力をビデオ会議システムに与えることができた。これはテレプレゼンスロボットを操作して話しかけたい相手などに接近させる操作者の体験をビデオ会議システムに移植できたと言えるであろう。では逆方向の体験、すなわちテレプレゼンスロボットと同じ部屋の中において接近される側の人間の体験を移植することは可能であろうか。図4に示す実験はこの問いに答えたものであり、ビデオ会議システムに対してディスプレイが迫り出す機能を追加し、テレプレゼンスロボットの接近による対人距離の変化をディスプレイの接近によって模擬できるようにしたものによってソーシャルテレプレゼンスの強化を試みた(Nakanishi, Kato and Ishiguro 2011)。実験では図に示すようにディスプレイが僅かに前にせりだしてくる可動条件と、常に迫り出した位置にある固定条件とを比較した。被験者に与えられたタスクは別の部屋にいる説明者からビデオ会議システムを通して3冊の本の説明を聞くというもので、説明者の背後に設置されたホワイトボードのトレイに3冊の本が立て掛けられており、説明者は3回、後ろを振り返って本を取り、前向き直って説明を始める、という行動を繰り返し

図4 ロボットディスプレイ実験



行った。その際、可動条件では後ろを振り返る直前にディスプレイが6cm後退し、前に向き直った直後に6cm前進した。固定条件ではディスプレイは常に手前の位置にあって動かなかった。ディスプレイのサイズは30インチであり、被験者と1.2m離れた位置に設置されていたので、6cmという移動距離はかなり目立たないものであった。それにもかかわらず、可動条件では固定条件よりもプレゼンスが伝わりやすくなった。インタビューにおける被験者のコメントから、せりだすディスプレイはそこに映し出されている説明者が被験者に近付いて来ているかのような印象を与え、それによって実際に目の前に説明者がいるかのような感覚を与えたことがうかがえた。このようにカメラの前進機能に引き続き、ディスプレイに前進機能を付与するという機能拡張によってもテレプレゼンスロボットと同様のプレゼンス伝達能力をビデオ会議システムに与えることができた。これらの実験結果は、ロボットを操作する側とロボットに対峙する側の両側においてロボットの機動性がカメラ映像のソーシャルテレプレゼンス創出能力を強化できることを示すとともに、ロボットとカメラ映像を通常のテレプレゼンスロボットよりも直接的な方法で融合可能であることを示している。

#### IV ロボットハンド

テレプレゼンスロボットにおけるロボットの役割はビデオ会議システムのカメラ映像に対して機動性を与えることだけであり、ゆえにロボットと

カメラ映像は完全に分離している。すなわち、遠隔にいる人間の身体の動きや見た目の表現の全てをカメラ映像が担っており、ロボットは身体表現には利用されていない。テレプレゼンスロボットはロボットとカメラ映像の融合における価値ある第一歩ではあるが、身振り手振りなどの身体動作が映像でしか表現できない点や、ロボットの重要な利用法である身体接触の再現が行えない点など、本格的な融合からは程遠いと言える。今現在の見た目を伝えることが重要と思われる顔などの提示にはカメラ映像を用いつつも、物理的実体として再現したほうが良いと考えられる身体の一部を映像の代わりにロボットで表現するなど、より密接なロボットとカメラ映像の融合が試みられるべきであると考えられる。

図5に示す実験では指差しに用いる手の部分だけをロボットハンドを用いて物理的に実体化する方法を試した(Onishi, Tanaka and Nakanishi 2016)。このロボットハンドはディスプレイに映っている人間の上半身と繋がって見えるように設置されており、その人間の手の動きに合わせて動くようになっている。実験ではこの方法と、通常のビデオ会議システムと同様にカメラ映像のみで身体を表現する方法、さらには過去の分散ワークプレイスの研究でよく用いられてきたような机に手を表示する方法(Tang and Minneman 1991; Luff et al 2006)を用いて、動物のぬいぐるみを指差す説明者の姿が被験者に対してどのような印象を与えるか調査したところ、ロボットハンドを用いる方法では他の2つの方法に比べてプレゼンスが伝わりやすくなった。

図5 指差し用ロボットハンド実験



図6に示す実験ではこのようなカメラ映像内の人間と連動するロボットハンドを用いて、遠隔地間での握手の実現を試みた (Nakanishi, Tanaka and Wada 2014)。ロボットハンドを握手のような身体接触に用いる場合は図5の実験のように身振り手振りに用いる場合とは異なり、見た目や動きなどの視覚情報に加えて触覚情報をもリアルに再現する必要が生じる。既存のロボットハンドはそのような目的では作られていないため、実験に先立って握手用のロボットハンドを自作した。

触覚情報の伝達が可能なロボットハンドは国内外で研究されているが、それらは一般的に、ロボットハンドに取り付けられたセンサーが検出した圧力を、ロボットハンドを遠隔操作する側のインタフェースデバイスを通して再現する仕組みになっている。このようなロボットハンドを握手に利用した場合、ロボットハンドを握る側の人間は圧力検出装置の付いた機械を触ることになり、人の手を握っているかのような触覚は得られない。また、ロボットハンドを操作する側の人間は特殊なデバイスを手に装着することになり、自然な状態で握手を行うことができない。そこで、人間の

手にそっくりなロボットハンドを両者の側で用いるというアプローチによって、特殊なデバイスを装着することなく、目の前にいる人間の手を握ると同様の状態で遠隔の人間の手を擬似的に握ることを可能にした。

これまでに多種多様なロボットハンドが開発されてきたが、その多くは物体を掴むことに焦点を当てているためにいかにも機械っぽいものであり、人間の手にそっくりなものは数少ない。人間の手の見た目や動きを模しているロボットハンドがいくつか開発されているものの、触覚まで再現しているものはほとんどない。自作した握手用ロボットハンドは十分な握力を発揮するためのワイヤー駆動機構、皮膚の柔らかさを出すためのゲルやスポンジ、体温のような温かさを出すためのヒーターが備わっており、人間の手の握力・人間の皮膚の柔らかさ・人肌の温度を再現可能である。

単に触覚を遠隔地間でのコミュニケーションに用いただけであれば、人間の手にそっくりなロボットハンドよりもずっと単純なデバイスで済み、実際にそのようなデバイスを開発して実験を

図6 握手用ロボットハンド実験



行った既存研究が多数存在する (Haans and Ijsselstein 2006)。しかしながら、そのような単純な触覚デバイスによって、等身大映像を映すような高品位のビデオ会議システムのプレゼンス伝達能力をさらに向上させることは困難である。低品位なシステムの能力を向上させたという報告はあるが、高品位なシステムが既に有する高い能力をさらに向上させたという報告はこれまでに無い。実験ではそのような高品位システムのプレゼンス伝達能力を、自作の握手用ロボットハンドが向上させようことを確認した。

2 地点を繋ぐビデオ会議システムに対して、自作の握手用ロボットハンドを両地点のディスプレイに取り付け、各地点のディスプレイの前に立っている利用者が自分の目の前のロボットハンドを握るという形式で遠隔握手を行う実験を行った。その結果、遠隔握手を行わない場合に比べてプレゼンスが伝わりやすくなり、さらに相手への親しみが向上することが確認された (図6の遠隔握手条件)。従来型のテレプレゼンスシステムのようにロボットハンドを片方にだけ設置し、もう片方の利用者はロボットハンドを握らずに遠隔地に設置されているロボットハンドを遠隔操作するだけの形式もテストしたが、プレゼンスは伝わりやすくなるはならなかった (図6の遠隔操作条件)。プレゼンス伝達の効果が見られなかった理由は、触覚情報が得られなかったことで握手をしている感覚が薄れてしまったからと思われる。また、両地点においてロボットハンドを握るという行為が発生していることを強調するために、相手がロボットハンドを握っている様子をディスプレイに映して可視化する形式もテストしたが、これにもプレゼンスを伝わりやすくする効果はなかった (図6の可視化条件)。効果が見られなかった理由は、カメラ映像内に相手の手が映ってしまったことで、相手の体とロボットハンドが繋がっているという感覚が消えてしまったためと思われる。

## V ロボット家具

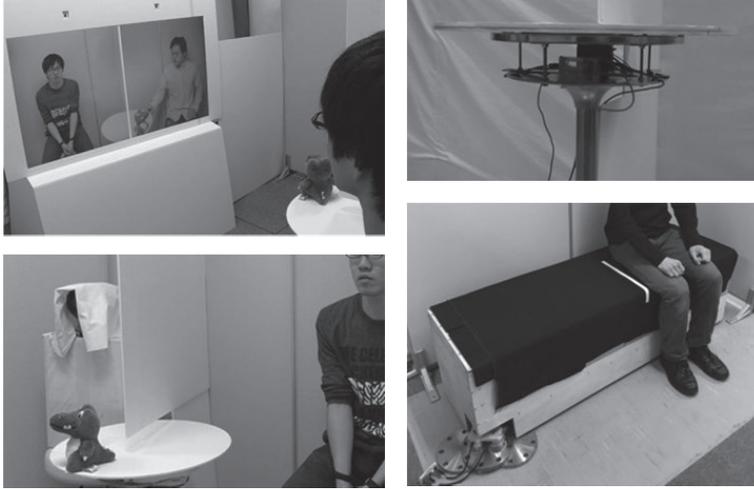
ロボットによる身体部位の表現はメリットばかりをもたらすものではない。まず、ロボット化し

た身体部位とそれ以外のカメラ映像で表現されている身体部位との間の継ぎ目に残る不自然さを完全に払拭することはできない。そして、ロボット化の必要性が薄い状況ではむしろカメラ映像のほうが自然でプレゼンス伝達に有利である。例えばIVで挙げたロボットハンドの場合、指差しや握手を行っていない間はカメラ映像の手に戻るべきであろう。このような問題を避けつつ、ロボットが有する物理的実体やそれによる触覚情報の再現をプレゼンスの伝達に利用するにはどうすれば良いのであろうか。一つの方法は、他者の身体をロボット化するという直接的なアプローチではなく、他者の身体動作の結果が反映される環境中の物体をロボット化するという間接的なアプローチである。

図7に示す実験ではテーブルや椅子といった家具をロボット化し、それを介して遠隔地にいる他者の身体動作を感じ取ることができるようにした (Nakanishi et al. 2017)。実験ではそれらの家具を互いに遠隔にいる2人の人間が共有している状況を表現するために、通常とは異なるビデオ会議システムを用いた。通常のビデオ会議システムのディスプレイは、遠隔地を眺めるために覗き込む窓の役割を果たす。それに対して実験用システムのディスプレイは図7左上のように、遠隔地と自分のいる部屋が融合したと仮定して、遠隔地にいる他者と自分が同じ部屋の中で隣り合っている状況を映し出す大型のウォールミラーとして機能する。このようなミラー型のビデオ会議システムは一般には普及していないが、これまでに数多くのシステムが研究開発されてきている (Morikawa and Maesako 1998)。実験用システムでは図7左下のように実際には隣には誰も座っていないが、その不在の他者があたかも隣に座っているかのように遠隔地のカメラ映像が自分の部屋のカメラ映像の隣に表示された。2人が同じテーブルや同じ椅子を共有している様子がディスプレイに映るように、同じ物をそれぞれの部屋に置き、ディスプレイ上でうまく繋がって見える位置に設置した。

実験では右に映っている遠隔の説明者がテーブルを手で回転させ、それによってテーブルの上に乗っている動物のぬいぐるみがあたかも説明者か

図7 ロボット家具実験



ら送られてきたかのような錯覚を与えることで、プレゼンスが伝わりやすくなるのかどうかを調べた。テーブル同様にぬいぐるみも同じ物がそれぞれの部屋にあり、ディスプレイ上ではテーブルの上に一つだけ乗っているように見えるように設置位置が調整されていた。図7右上のようにテーブルにはサーボモーターが取り付けられており、説明者の部屋のテーブルの回転角度を読み取って、それと同じ角度だけ被験者の部屋のテーブルを回転させる仕組みにし、一つのテーブルであるかのように見せた。実験では、最初からぬいぐるみがテーブルの被験者側に置いてあり説明者はテーブルを回転させない条件を基準にして、手で回転させる条件と、テーブルに手を触れずにリモコン操作によって回転させる条件を比較した。すると、テーブルを回転させる2条件では回転させない条件よりもプレゼンスが伝わりやすくなり、さらに手で回転させる条件の方がリモコン操作の条件よりも伝わりやすくなるという結果になった。この結果はテーブルを回転させるという説明者の身体動作がテーブルの回転を通して間接的に被験者に伝わった効果を示していると考えられることができる。

上記の実験に加えて、テーブルの代わりに図右下に示すような2人が同時に座ることのできる長椅子を設置し、遠隔の説明者の動作が椅子にもたらず振動を被験者側で再現することでプレゼンスを伝わりやすくなる実験も行った。説明者の部屋

の椅子の座面に圧力センサを取り付け、座ったり立ったりする動作を検出した。そして、そのような動作が検出されると、被験者の部屋の椅子の一本の足に取り付けられた電磁石を動かし、あたかも隣に説明者がいるかのような振動を発生させた。実験では振動を発生させる場合と発生させない場合を比較し、振動がプレゼンス伝達に有効であることを確認した。この結果もテーブル実験と同様に、説明者の身体動作が間接的に被験者に伝わった効果を示していると考えられる。

両実験とも開始前にスピーカしか置かれていない隣の無人の空間を被験者に見せ、実際には隣に誰も居ないことを分かってもらった上で実施している。よって、プレゼンス伝達の効果が発揮された理由は、隣に人が座っていると思いつつも騙されてしまったからではない。実際には無人であることを分かった上で、ロボット家具の動作によって、あくまで印象の上で隣に人がいるような気分にさせられたということである。

## VI おわりに

ワークスペースの分散化は共同作業者のプレゼンスの希薄化という問題を生む。プレゼンスが希薄化しても問題の無い共同作業の方法を編み出すというアプローチはもちろん有効であるが、失われたプレゼンスを技術の力によって取り戻すこと

も同時に検討されるべきであろう。そして、それはロボット技術を導入すれば可能である。しかしながら、ロボットは見た目の情報を伝えるのが苦手であり、この問題はビデオ会議システムとの融合によって解決できる。ロボットはその物理的な実体によって身体の動きを伝えることができることにより、カメラ映像とは異なった方法でプレゼンスを伝えることができる。この特徴をビデオ会議システムのディスプレイと融合させる方法として3種類のシステムを本稿では紹介した。1つ目は対人距離の変化などを生み出す身体全体の移動を、ディスプレイやそこに取り付けられているカメラのロボット化によって再現する方法である。ディスプレイやカメラを移動させることで、他者や自分が歩いている感覚を与えることができる。2つ目は身振り手振りなどを生み出す身体の一部をロボットで再現して、ディスプレイ上のカメラ映像に結合させる方法である。ロボット化した身体部位の動きが迫真性を与えたり身体接触を可能にしたりするとともに、その身体部位以外はカメラ映像で表現することによって、見た目の情報の損失を最小限に抑えることができる。3つ目は環境中の物体を押し回したりする身体動作の結果を、その物体のロボット化によって再現する方法である。ミラー型ビデオ会議システムなどによって物体を他者と共有している状況をカメラ映像として作り出し、その上で物体を動かすことで、他者の身体動作が物体を通して伝わって来るような感覚を与えることができる。これらのロボットの利用法は全て、カメラ映像がディスプレイという平面の上に描かれた単なる絵に過ぎないという感覚を打ち消すことを狙ったものであると言える。

現在の技術レベルでは本稿で述べたような仕掛けを駆使したとしても、対面環境に匹敵するプレゼンスの共有を遠隔地間で実現することは非常に困難であるが、ディスプレイやロボットの性能はまだまだ発展途上であり、実際に会うことと仮想的に会うこととの差はますます縮小していくと期待される。

#### 参考文献

- Bainbridge, W. A., Hart, J., Kim, E. S. and Scassellati, B. (2011) "The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents," *International Journal of Social Robotics*, 3 (1), 41-52.
- Bly, S. A., Harrison, S. R. and Irwin, S. (1993) "Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment," *Communications of the ACM*, 36 (1), 28-46.
- Bondareva, Y. and Bouwhuis, D. (2004) "Determinants of Social Presence in Videoconferencing," *Proc. AVI 2004 Workshop on Environments for Personalized Information Access*, 1-9.
- de Greef, P. and Ijsselstein, W. A. (2001) "Social Presence in a Home Tele-Application," *CyberPsychology & Behavior*, 4 (2), 307-315.
- Fish, R. S., Kraut, R. E. and Chalfonte, B. L. (1990) "The VideoWindow System in Informal Communication," *Proc. CSCW 1990*, 1-11.
- Garau, M., Slater, M., Bee, S. and Sasse, M. A. (2001) "The Impact of Eye Gaze on Communication Using Humanoid Avatars," *Proc. CHI 2001*, 309-316.
- Gaver, W. W., Sellen, A., Heath, C. and Luff, P. (1993) "One is not Enough: Multiple Views in a Media Space," *Proc. CHI 1993*, 335-341.
- Haans, A. and Ijsselstein, W. (2006) "Mediated Social Touch: a Review of Current Research and Future Directions," *Virtual Reality*, 9 (2=3), 149-159.
- Heath, C. and Luff, P. (1991) "Disembodied Conduct: Communication through Video in a Multi-media Office Environment," *Proc. CHI 1991*, 99-103.
- Heath, C. and Luff, P. (1992) "Media Space and Communicative Asymmetries: Preliminary Observations of Video-Mediated Interaction," *Human-Computer Interaction*, 7 (3), 315-346.
- Isaacs, E. A. and Tang, J. C. (1994) "What Video Can and Cannot Do for Collaboration: A Case Study," *Multimedia Systems*, 2 (2), 63-73.
- Ishii, H. and Kobayashi, M. (1992) "ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact," *Proc. CHI 1992*, 525-532.
- Kang, S., Watt, J. H. and Ala, S. K. (2008) "Communicators' Perceptions of Social Presence as a Function of Avatar Realism in Small Display Mobile Communication Devices," *Proc. HICSS 2008*.
- Lee, K. M., Jung, Y., Kim, J. and Kim, S. R. (2006) "Are Physically Embodied Social Agents Better than Disembodied Social Agents?: The Effects of Physical Embodiment, Tactile Interaction, and People's Loneliness in Human-Robot Interaction," *International Journal of Human-Computer Studies*, 64 (10), 962-973.
- Luff, P., Heath, C., Kuzuoka, H., Yamazaki, K. and Yamashita, J. (2006) "Handling Documents and Discriminating Objects in Hybrid Spaces," *Proc. CHI 2006*, 561-570.
- Mantei, M. M., Baecker, R. M., Sellen, A. J., Buxton, W. A. S., Milligan, T. and Wellman, B. (1991) "Experiences in the Use of a Media Space," *Proc. CHI 1991*, 203-208.
- Morikawa, O. and Maesako, T. (1998) "HyperMirror: Toward Pleasant-to-Use Video Mediated Communication System," *Proc. CSCW 1998*, 149-158.
- Nakanishi, H., Kato, K. and Ishiguro, H. (2011) "Zoom Cameras and Movable Displays Enhance Social Telepresence," *Proc.*

- CHI 2011, 63-72.
- Nakanishi, H., Murakami Y. and Kato K. (2009) "Movable Cameras Enhance Social Telepresence in Media Spaces," *Proc. CHI 2009*, 433-442.
- Nakanishi, H., Murakami, Y., Nogami D. and Ishiguro, H. (2008) "Minimum Movement Matters: Impact of Robot-Mounted Cameras on Social Telepresence," *Proc. CSCW 2008*, 303-312.
- Nakanishi, H., Tanaka, K., Kato, R., Geng, X. and Yamashita, N. (2017) "Robotic Table and Bench Enhance Mirror Type Social Telepresence," *Proc. DIS 2017*, 779-790.
- Nakanishi, H., Tanaka, K. and Wada, Y. (2014) "Remote Handshaking: Touch Enhances Video-Mediated Social Telepresence," *Proc. CHI 2014*, 2143-2152.
- Nakanishi, H., Yoshida, C., Nishimura, T. and Ishida, T. (1999) "FreeWalk: A 3D Virtual Space for Casual Meetings," *IEEE MultiMedia*, 6 (2), 20-28.
- Nguyen, D. T. and Canny, J. (2007) "Multiview: Improving Trust in Group Video Conferencing through Spatial Faithfulness," *Proc. CHI 2007*, 1465-1474.
- (2009) "More than Face-to-Face: Empathy Effects of Video Framing," *Proc. CHI 2009*, 423-432.
- Olson, J. S., Olson, G. M. and Meader, D. K. (1995) "What Mix of Video and Audio Is Useful for Small Groups Doing Remote Real-time Design Work?" *Proc. CHI 1995*, 362-368.
- Onishi, Y., Tanaka K. and Nakanishi, H. (2016) "Embodiment of Video-Mediated Communication Enhances Social Telepresence," *Proc. HAI 2016*, 171-178.
- Paulos, E. and Canny, J. (2001) "Social Tele-Embodiment: Understanding Presence," *Autonomous Robots*, 11 (1), 87-95.
- Prussog, A., Mühlbach, L. and Böcker, M. (1994) "Telepresence in Videocommunications," *Proc. Annual Meeting of Human Factors and Ergonomics Society*, 25-38.
- Root, R. W. (1988) "Design of a Multimedia Vehicle for Social Browsing," *Proc. CSCW 1988*, 25-38.
- Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H. and Hagita, N. (2007) "Android as a Telecommunication Medium with a Human-like Presence," *Proc. HRI 2007*, 193-200.
- Tanaka, K., Nakanishi, H. and Ishiguro, H. (2014) "Comparing Video, Avatar, and Robot Mediated Communication: Pros and Cons of Embodiment," *Proc. CollabTech 2014*, CCIS 460, 96-110.
- Tang, J. C. and Minneman, S. L. (1991) "VideoDraw: a Video Interface for Collaborative Drawing," *ACM Transactions on Information Systems*, 9 (2), 170-184.
- Vogele, K. and Fink, G. R. (2003) "Neural Correlates of the First-Person-Perspective," *Trends in Cognitive Sciences*, 7 (1), 38-42.

なかにし・ひでゆき 大阪大学大学院工学研究科知能・機能創成工学専攻准教授。情報学専攻。