# 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

田中 一晶<sup>†1, †2</sup> 中西 英之<sup>†1</sup> 石黒 浩<sup>†3</sup>

近年,新たな遠隔会議の方法としてロボットを介して遠隔地にいる相手と対話するロボット会議が研究され始めた. ロボット会議の特徴は,対話相手の身体動作を実体のあるロボットで提示することである. しかし,実体のあるメディアを介して対話することが遠隔会議においてどのような効果をもたらすかは明らかになっていない. 本研究の目的は,遠隔会議における実体の効果を明らかにし,ロボット会議の有効性を検証することである. 我々は,ロボットや様々な既存の遠隔対話メディアを介して対話する被験者実験を実施した. その結果,実体は対話相手のソーシャルテレプレゼンス(相手と対面で対話している感覚の度合)を強化することが分かった. さらに,実体は対面会議で感じるような緊張感を生み出すことも分かった. これらの結果からロボット会議の有効性が示唆された.

## Robot Conferencing that Presents Physically Embodied Body Motions Enhances Social Telepresence

## KAZUAKI TANAKA<sup>†1, †2</sup> HIDEYUKI NAKANISHI<sup>†1</sup> HIROSHI ISHIGURO<sup>†3</sup>

In recent years, studies have begun on robot conferencing in which people talk with a remote conversation partner through a robot as a new teleconferencing method. The main feature of robot conferencing is that the robot presents the bodily motions of the conversation partner with a physical embodiment. However, the effects of physical embodiment on teleconferencing had not yet been demonstrated. The purpose of this study is to clarify the effects of physical embodiment on teleconferencing to investigate the effectiveness of robot conferencing. We conducted an experiment in which subjects talked with a partner through a robot and various existing telecommunication media. As a result, we observed that the physical embodiment enhanced social telepresence, i.e., the sense of resembling face-to-face interaction. Furthermore, we also observed that physical embodiment enhances the sense of tension as in the case of a face-to-face meeting. These results indicated the effectiveness of robot conferencing.

#### 1. はじめに

電話のような音声のみの対話では、対話相手のソーシャルテレプレゼンスが希薄になるという問題がある. ソーシャルテレプレゼンスとは、相手と対面で対話している感覚の度合である[9]. ソーシャルテレプレゼンスを強化する最も一般的な方法はビデオ会議であり、対話相手の映像はソーシャルテレプレゼンスを強化することが知られている[8][11]. しかし、ビデオ会議におけるソーシャルテレプレゼンスは、対面会議と比較するとまだ十分とは言えない.また、ビデオ会議では、対話相手の服装や部屋の様子などの余計な情報も一緒に伝えてしてしまうため、プライバシーの問題がある.

このビデオ会議に代わる新たな遠隔会議の方法として、ロボット会議が研究され始めた[18][19][21]. ロボット会議の特徴は、実体を持つロボットで対話相手の身体動作を提示することである. 本研究の目的は、遠隔会議における実体の効果を明らかにし、ロボット会議の有用性を示すこと

である. 我々の先行研究では、対話相手の映像を表示したディスプレイを前後に移動させる方法で対話相手の前後移動を強調して提示すると、ソーシャルテレプレゼンスが強化されることが分かっている[17]. したがって、実体のあるロボットで対話相手の身体動作を提示することは、ソーシャルテレプレゼンスを強化する可能性がある.

遠隔会議における実体の効果として、対面会議で感じるような緊張感を生み出すことも期待される. 先行研究では、テキストチャットで感じる緊張感は対面会議に劣ることが報告されている[6]. 一方、ボイスチャット、ビデオ会議、ロボット会議など他の遠隔会議で感じる緊張感については明らかになっていない. 本研究では、ボイスチャットやビデオ会議など実体の無い対話方法とロボット会議や対面会議のような実体のある対話方法を比較し、実体があることによる緊張感への影響についても調査する.

心理学研究の分野では、情緒的な緊張が吃りの原因の一つであることが知られている[13]. 我々は、緊張することで途切れがちな発話になると考え、緊張感を測る方法として発話に含まれる途切れの頻度に着目した。図1に途切れがちな発話の例を示す。図中の四角は発言を表しており、それらの間隔が途切れである。本研究では、この途切れの頻度を観察することで、遠隔対話メディアを通して対話す

<sup>†1</sup> 大阪大学大学院工学研究科 知能・機能創成工学専攻

Department of Adaptive Machine Systems, Osaka University

<sup>†2</sup> 独立行政法人科学技術振興機構,CREST

CREST, Japan Science and Technology Agency †3 大阪大学大学院基礎工学研究科 システム創成専攻

Department of Systems Innovation, Osaka University



図1 途切れがちな発話の例

Figure 1. Example of speech including a lot of pauses.

ることによる緊張感の度合の測定を試みる.

近年, avatar Kinect[a]に代表されるようなアバタチャット が容易に利用できるようになり, アバタを介して対話する ことによる遠隔会議への効果を調査した研究が多く行われ ている[2][4][10][12][22]. アバタチャットはロボット会議と 同様に、対話相手の外見は伝えず、顔や身体の動きのみを 伝えるものである. したがって, 我々は, アバタチャット に実体を加えることでロボット会議になると仮定した. 同 様に、ボイスチャットに実体を加えると、静止したロボッ トを介して対話するメディアになり、ビデオチャットに実 体を加えると対面会議になると仮定することができる. ま た,これらの対話方法が相手に伝える情報に関しては,ボ イスチャットと静止ロボットは音声のみ, アバタとロボッ トは音声と身体動作, ビデオと対面は音声, 身体動作, 外 見であると仮定した. これらの仮定は, 6 つの対話方法を 実体の有無と伝達情報の違い(音声のみ,音声+身体動作, 音声+身体動作+外見) に要因を分けて比較することを可 能にした.

#### 2. 関連研究

ロボット会議に関する研究は多く行われており、様々な遠隔操作ロボットが提案されている[14][16][18][19][21][23]. それらのロボットは、コミュニケーションに必要と思われる操作者の頭部の動きを実体で提示することができる. 先行研究では、ロボット会議のソーシャルテレプレゼンスは、音声のみの対話やビデオ会議よりも高いことが示されている[19]. しかし、この研究で用いられたロボットは、非常に人に酷似した外見を持っており、実体で身体動作を提示したことと、人に近いリアルな外見を持っていること、どちらの要因がソーシャルテレプレゼンスを強化したのかは明らかにされていない. 本研究では、特定の人物の外見を持たない遠隔対話用ロボットである Telenoid[18]を用い、実体で身体動作を提示することによるソーシャルテレプレゼンスへの効果を明らかにする.

人とインタラクションを行う自律ロボットの研究分野では、ロボットの持つ実体の効果が検証されており、実体のあるロボットは映像で提示されたエージェントよりも高い存在感があることが分かっている[3][15]. 遠隔対話用のロボットにおいても、実体があることで対話相手のソーシャルテレプレゼンスを強化できる可能性がある.

さらに, 本研究では, 遠隔対話メディアに実体を追加す

a) http://www.xbox.com/en-us/kinect/avatar-kinect

ることによる発話への影響を調査する.遠隔対話メディアの違いによる発話への影響を調べた研究では,会話構造(話者交代や同時発話など) [1][7][20]が主に用いられている.本研究では,吃りの度合を調べるため,発話に含まれる途切れの頻度を用いる.緊張感は吃りの原因とされていることから[13],途切れの頻度を調べることで,緊張感の測定を試みる.

上述の先行研究では、遠隔対話メディアを単一の要因として扱っている。これに対し、実体の有無、伝達する情報の種類に要因に分けて遠隔対話メディアを比較することも本研究の特徴である。本研究では、実体要因と伝達情報要因によるソーシャルテレプレゼンスと緊張感への影響を明らかにする。

### 3. 実験

#### 3.1 仮説

先行研究では、人に近いリアルな外見を持ったロボットを通して対話すると、対話相手のソーシャルテレプレゼンスが向上することが分かっている[19]. また、我々の先行研究において、対話相手の移動をディスプレイの移動で強調して提示することでソーシャルテレプレゼンスが強化されることが分かっている[17]. さらに、人とロボットとのインタラクションの研究分野では、映像上のエージェントより、実体のあるロボットの方が高い存在感があることが分かっている[3][15]. そこで、我々は、実体のあるロボットで身体動作を提示すると、対話相手のソーシャルテレプレゼンスが強化されると考えた. 本研究の1つ目の仮説は次の通りである.

# 仮説 1: 実体は対話相手のソーシャルテレプレゼンスを強化する.

また、対話相手の身体動作を再現するアバタを介して対話するアバタチャットで感じるソーシャルテレプレゼンスは、音声のみで対話するよりも高いが、ビデオ会議には劣ることが分かっている[10][12]. したがって、我々は、伝達情報(音声、身体動作、外見)が増加するほど対話相手のソーシャルテレプレゼンスが増すと考えた。本研究の2つ目の仮説は次の通りである.

# 仮説 2: 伝達情報の増加は対話相手のソーシャルテレプレゼンスを強化する.

対面会議で感じる緊張感はテキストチャットよりも高いことが知られているが[6],ボイスチャット,アバタチャット,ビデオ会議等で感じる緊張感がテキストチャットよりも高いという報告は無い.対面会議と異なり,これらの遠隔対話メディアは実体を持たない.そこで,我々は,実体があることが緊張感を生むのではないかと考えた.

本研究では、緊張感は吃りの原因とされていることから [13]、発話に含まれる途切れの頻度を緊張の度合と仮定す る. つまり、話者が感じる緊張感の度合を、その発話に含

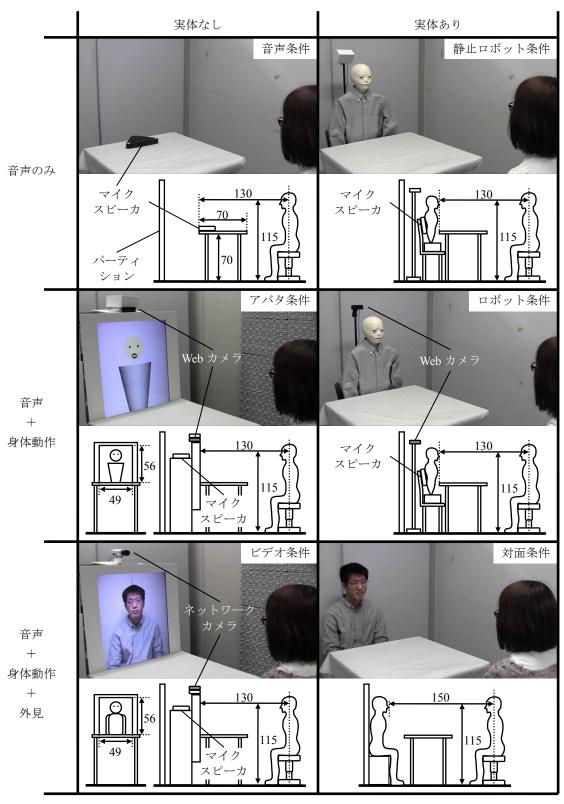


図2 実験条件と実験環境(単位:cm)

Figure 2. Experimental conditions and setups (length unit: centimeters).

まれる途切れの頻度で測定する.本研究の3つ目の仮説は次の通りである.

## 仮説 3: 実体は途切れの頻度 (緊張感) を増加させる. 3.2 実験条件

前節で述べた本研究の仮説は、実体要因と伝達情報要因

の2つの要因から構成される. 実体要因は有無の2水準, 伝達情報要因は, 音声のみ, 音声+身体動作, 音声+身体動作+外見の3水準である. したがって, この仮説を検証するため, 我々は図2に示す2×3の6つの実験条件を設定した. 各実験条件の詳細は以下で述べる.





Figure 3. Teleoperated robot used in the experiment and its modifications.

図3 実験で使用したロボットとその変更点

**音声条件(実体なし、音声のみ)**:一般的なボイスチャットと同様であり、被験者は目の前の机に置かれたマイクスピーカを通して実験者と対話する.

静止ロボット条件(実体あり、音声のみ):被験者は目の前に置かれた動かないロボットを見ながら実験者と対話する。マイクスピーカはロボットの背中に設置している。本条件ではTelenoidを使用するが、図3に示すように2つの変更を施した。Telenoidは人のようなリアルな目を備えているが、見た目の抽象度を揃えるために、目を黒の単色に変更した。さらに、後述のビデオ条件や対面条件で実験者が着用したものと同じ灰色のシャツを着せた。後述のロボット条件で用いるロボットの背後に設置したフェイストラッキング用のWebカメラは本条件では使用しないため、カバーで隠した。被験者にはロボットが動かないことを予め伝えてから実験を行った。

アバタ条件(実体なし、音声+身体動作):被験者は目の前のディスプレイに表示された3DCGのアバタを見ながら実験者と対話する.マイクスピーカはディスプレイの後ろに設置した.アバタは、黒い目と唇が配置された肌色の円柱の頭部と、実験者が着用したシャツとほぼ同じ灰色の円錐の胴体で構成されている.頭部の直径は、ロボットの頭部の横幅と同じ13.5cmに設定した.被験者と実験者の顔の動きは、フェイストラッキングツールであるFaceAPIを用いて追従した.頭部の向き(pitch/roll/yawの角度)、口の形(3 次元座標の14 頂点)のトラッキングデータを30fpsで取得し、アバタで再現した.アバタは縦置きにした40 イン

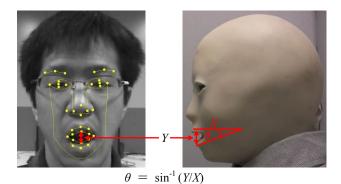


図4 口のトラッキングデータの角度への変換

Figure 4. Conversion from mouth tracking data to degrees.

チのディスプレイに表示した.ディスプレイの枠は、被験者の位置から見えないように、背景のパーティションと同じ白色の板で隠した.また、他の条件と同じく、前方に机を設置したため、ディスプレイの実際の表示域は56×49cmになった.アバタを表示したディスプレイの上部には、フェイストラッキング用のWebカメラを設置しており、後述のビデオ条件で用いるネットワークカメラは本条件では使用しないため、カバーで隠した.

ロボット条件(実体あり、音声+身体動作):被験者は目の前のロボットを見ながら実験者と対話する. ロボットは、静止ロボット条件と同様に、図 3 の変更を施した Telenoidを使用した. ロボットの首は3自由度(pitch/roll/yaw)、口は開閉のみの1自由度である. アバタ条件と同じ方法で取得した顔のトラッキングデータをロボットで再現する. ただし、ロのトラッキングデータは、図 4 中に示した式でロボットの口の角度  $\theta$  に変換した. Yは上下の唇中央のy座標の差、つまり口の開き具合である. また、X はロボットの唇先端から顎までの長さ(5.8cm)である.

ビデオ条件(実体なし、音声+身体動作+外見):被験者は目の前のディスプレイに表示された実験者の映像を見ながら実験者と対話する.実験者は灰色のシャツを毎回着用した.実験者の映像は、実験者側のネットワークカメラ(解像度1280×720、フレームレート30fps)で撮影し、縦置きにした40インチのディスプレイに表示した.前述のアバタ条件と同じく、実際の表示範囲は56×49cmであり、その解像度は720×800になった.また、実験者の頭部がロボットの頭部の横幅と同じ13.5cmになるように、実験者を撮影するネットワークカメラの画角を調整した(水平画角87度).実験者を表示するディスプレイの上部には被験者を撮影するネットワークカメラを設置しており、前述のアバタ条件で用いたフェイストラッキング用のWebカメラは本条件では使用しないため、カバーで隠した.

対面条件(実体あり,音声+身体動作+外見):被験者は 目の前の実験者と直接対話する. ビデオ条件と同様に,実 験者は灰色のシャツを毎回着用した. 被験者から見て,実 験者の顔の大きさがロボットの顔とほぼ同じ大きさに見えるように、被験者の座る位置から実験者までの距離を調整した(150cm).この距離は、被験者からディスプレイやロボットまでの距離(130cm)よりも20cm遠いが、実験では、遠過ぎると感じた被験者はいなかった.

実験では、被験者側の対話環境と実験者側の対話環境が同じであることを予め被験者に伝えた.これを被験者に信じさせるために、アバタ条件、ロボット条件、ビデオ条件では、実験の前に実験者から自分がどのように見えているか(アバタ、ロボット、ビデオ映像)を7インチのディスプレイで提示し、対話環境が双方向であることを説明した.その際、アバタ条件とロボット条件は、実験者と被験者自身の顔の動きがアバタやロボットで実際に再現されていることを確認してもらった.この確認作業を行うのは、予備実験において、本当に実験環境が双方向であるのか、アバタやロボットが自動で動いているのではないかと疑った被験者がいたためである.確認用に使用した7インチのディスプレイは実験の前に撤去した.

先行研究では、ビデオ会議において、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と、ソーシャルテレプレゼンスが向上することが分かっている[5].本研究でも、全ての実験条件でアイコンタクトが成立するように、被験者の椅子の高さを調整した。この目線の高さの調整で、実際にアイコンタクトがあったように被験者が感じることを予備実験で確認した。

#### 3.3 実験タスク

本研究では、ソーシャルテレプレゼンスの評価だけでな く、緊張感の度合を測るため、発話に含まれる途切れの頻 度を分析する. この分析を行うために、出来るだけ長い 1 ターンの発話データを得る必要があった. そこで, 特定の 電子機器の問題点とその改善案、および追加してほしい機 能について述べるタスクを設定した. 実験は, 1 人の被験 者が6つ全ての条件を体験する被験者内計画で実施するた め、異なる6つの電子機器(電子ブックリーダー、携帯ゲ ーム機, スマートフォン, ロボット掃除機, 携帯音楽プレ イヤー, 3D テレビ) を話題として設定した. これらの話題 は、被験者が問題点等の意見を述べやすいように、最近話 題になっている電子機器から選定した. 話題は各条件で実 験を開始する際に実験者から伝えた.被験者が話している 間の実験者の行動は、被験者の発言を遮らないように、頷 きながら相槌を打つのみとした. 実験では, 実験者が途中 で発話を促すことは無かったが、ほとんどの被験者が1分 以上話すことができた.

6 つの条件での実験の順番はラテン方格法に基づいて設定した.また、被験者に実験タスクに慣れてもらうため、練習として対面条件で練習を行ってから6回の実験を行った.練習での話題は毎回 IC カード型乗車券とした.

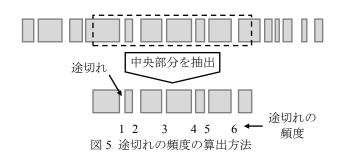


Figure 5. Method to calculate the frequency of pauses.

### 3.4 実験データの収集

実験には我々の大学の近辺に住む学部生36名(男:19名, 女:17名)が被験者として参加した.

#### 3.4.1 アンケートによるソーシャルテレプレゼンスの評価

6 つ条件の実験を行った後、被験者にアンケートを実施した.アンケートでは、被験者が感じた対話相手(実験者)のソーシャルテレプレゼンスを測定するため、「同じ部屋の中で相手に話しかけている感覚があった」という質問について、9 段階のリッカード尺度で 6 つの条件を相対的に評価してもらった. 1、3、5、7、9 は、それぞれ「まったくあてはまらない」、「あてはまらない」、「どちらともいえない」、「あてはまる」、「非常にあてはまる」に対応させた.アンケートの後に、点数を付けた理由についてインタビュで尋ねた.

各条件に対応するアンケート項目の順番は、実験を行った順番に並べ替えており、各条件の実験装置の写真共にアンケートに記した。この並べ替えと写真は、実験の想起を助ける手がかりである。

#### 3.4.2 途切れの頻度の測定

本研究では、遠隔対話メディアを用いて対話することによる緊張感への影響を明らかにするため、1 ターンの発話中に含まれる途切れの頻度を測定する.

途切れの頻度を測定するため、実験での被験者の発話を全て録音し、アノテーションツールである ELAN[b]を用いて書き起こした。その際、発話の途切れ(無音区間)を得るために、発話区間の開始時間と終了時間を可能な限り正確に決定した。ただし、促音「つ」による無音区間は吃りとは無関係であるため無視した。また、50ms よりも短い区間はホワイトノイズと被験者の発話を区別することができなかったため、50ms 以上の無音区間のみを途切れとした。ホワイトノイズはノイズフィルタによって除去することが可能であるが、非常に小さい音量の発言も除去されてしまうため、ノイズフィルタは使用しなかった。

図5に途切れの頻度の算出方法を示す.まず,全体の発 話から一定量の発話を抽出する必要がある.発話の開始部 分は,被験者がその話題に興味を持っている場合,流暢に 話す傾向があり,話題による影響が大きい.一方,発話の

b) http://www.lat-mpi.eu/tools/elan/

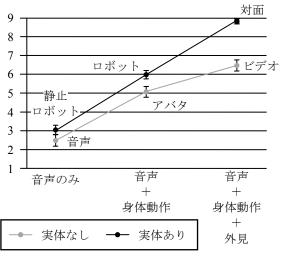


図 6 同じ部屋で相手に話しかけている感覚 についてのアンケート結果

Figure 6. Results of the questionnaire on the feeling of speaking to the partner in the same room.

終了部分は、話すことがほとんど無くなり、途切れの多い発話になる場合が多い。そこで、全体の発話の中央部分を抽出することとした。この中央部分に含まれる途切れの数が途切れの頻度である。発話の分量を量る単位は、発音の最小単位である音節を採用した。前述の通り、ほとんどの被験者が1分以上話すことができており、1分に相当する音節数は約200であったため、中央から抽出する発話の量は200音節とした。ELANで作成した書き起こしデータから、音節数および途切れの頻度を算出する作業は、表計算ソフトを用いて行った。

#### 4. 実験結果

#### 4.1 ソーシャルテレプレゼンス

図6に、同じ部屋の中で相手に話しかけている感覚についてのアンケート結果を示す。図中の点は、9 段階のリッカート尺度による各実験条件の評価の平均値である。

このアンケート結果について実体要因×伝達情報要因の2要因分散分析を行った.その結果,実体要因(F(1,35)=36.955,p<.001)と伝達情報要因(F(2,70)=279.603,p<.001)の主効果,および交互作用(F(2,70)=14.794,p<.001)が有意であった.交互作用が有意であったため,Bonferroni補正法を用いて単純主効果の検定を行った.伝達情報が音声+身体動作の場合,および音声+身体動作+外見の場合,実体要因に有意差があり(それぞれ,F(1,105)=8.857,p<.01,F(1,105)=65.470,p<.001),音声のみの場合は有意傾向であった(F(1,105)=3.460,p=.086).つまり,ロボットを介した対話でのソーシャルテレプレゼンスはアバタよりも高いこと,対面で感じるソーシャルテレプレゼンスはビデオよりも高いことが分かった.これらの結果は仮説 1(実体は対話相手のソーシャルテレプレゼンスを強化する)を支持す

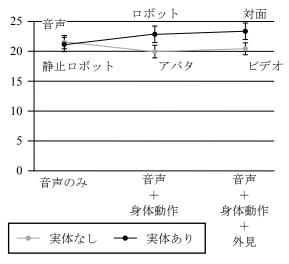


図7 途切れの頻度の平均値

Figure 7. Average of the frequency of pauses.

るが、伝達情報が音声のみの場合には、実体がソーシャル テレプレゼンスを強化する効果は低い.

さらに、実体がある場合と無い場合、両方において伝達情報要因に有意差が認められた(それぞれ、F(2, 140)=107.141,p<.001,F(2, 140)=223.095,p<.001). 多重比較の結果、対面で感じるソーシャルテレプレゼンスはロボットや静止ロボットよりも高いこと(共にp<.001)、ロボットを介した対話でのソーシャルテレプレゼンスは静止ロボットよりも高いこと(p<.001)、ビデオを介した対話でのソーシャルテレプレゼンスはアバタや音声のみでの対話よりも高いこと(共にp<.001)、アバタを介した対話でのソーシャルテレプレゼンスは音声のみでの対話よりも高いこと(p<.001)が分かった。これらの結果は仮説 2(伝達情報の増加は対話相手のソーシャルテレプレゼンスを強化する)を支持する.

#### 4.2 緊張感

実験では、ほとんどの被験者が全ての条件において 1分以上話すことができた.しかし、10名の被験者は、発話時間が 1分に満たない実験条件があった.3.4.2節で述べた通り、途切れの頻度を算出する上で、発話の開始・終了部分を削除し、中央部分を抽出する必要があるが、発話時間が 1分に満たない場合、分析可能な量の発話を抽出することができなかった.そこで、途切れの頻度の分析は全ての条件で 1分以上の発話データが揃っている 26名 (男性: 15名、女性: 11名)で行った.

図 7 に途切れの頻度の平均値を示す. この途切れの頻度について,アンケート結果と同様に,実体要因×伝達情報要因の2 要因分散分析を行った. その結果,実体要因の主効果が有意であり (F(1,25)=8.004,p<.01),交互作用が有意傾向であった (F(2,50)=2.947,p=.062). 交互作用が有意傾向であったため,Bonferroni 補正法を用いて単純主効果の検定を行った. 伝達情報が音声+身体動作の場合,および

音声+身体動作+外見の場合,実体要因に有意差が認められた(それぞれ,F(1,75)=6.799,p<.05,F(1,75)=6.981,p<.05). つまり,ロボット条件の途切れ頻度は,アバタ条件よりも高いこと,対面条件での途切れ頻度はビデオ条件よりも高いことが示された.これらの結果は仮説 3(実体は途切れの頻度(緊張感)を増加させる)を支持するが,伝達情報が音声のみの場合には実体による途切れの頻度への影響は見られなかった.

#### 5. 考察

実験結果から, 実体は対話相手のソーシャルテレプレゼ ンスを強化することが分かった. 被験者へのインタビュで は、目の前にロボットが物理的に存在するために、ロボッ ト条件はアバタ条件よりも実験者が目の前にいるように感 じたという意見が得られた(36 名中 7 名).一方,静止口 ボット条件も物理的な実体があるが、音声条件との間に有 意な差は見られなかった、インタビュでは、実験者の反応 を見ることができないため静止ロボット条件は音声条件と 変わらないという意見が得られており(36名中3名),ア ンケートでは36名中8名の被験者が音声条件と静止ロボッ ト条件に同じ点数を付けていた. さらに, ロボットが動い たときに実験者がいるように感じたと述べた被験者もいた (36 名中 5 名). したがって、実体がソーシャルテレプレ ゼンスを強化する効果は、対話相手の身体動作を伝えない 音声のみでの対話の際には低いと考えられる. これらの結 果は、対話相手の身体動作を実体で再現するロボット会議 の有用性を示している.

また、伝達情報の増加(身体動作や外見の伝達)もソーシャルテレプレゼンスを強化することが分かった.しかし、図6を見ると、ロボット条件でのソーシャルテレプレゼンスは、対話相手の外見を伝えていないにも関わらず、ビデオ条件と同等であることが分かる.実際、アンケートでは、半数近くの被験者がロボット条件にビデオ条件以上の点数を付けていた(36名中16名). 我々は、実体によるソーシャルテレプレゼンスの強化が、ロボットが外見を伝えないことを補ったのではないかと考えている. この結果は、プライバシーの問題があるビデオ会議をロボット会議が代替する可能性を示唆している.さらに、身体動作だけでなく、外見も伝えるロボットが実現できれば、ロボット会議でのソーシャルテレプレゼンスはビデオ会議を上回る可能性がある.

本研究では、遠隔会議において感じる緊張の度合を測定するため、発話に含まれる途切れの頻度を観察した。その結果、実体は途切れの頻度(緊張感)を増加させることが分かった。しかし、伝達情報が音声のみである静止ロボット条件は、途切れの頻度を増加させなかった。これは、上述の通り、被験者が音声条件と静止ロボット条件が同じであると感じていたことが原因かもしれない。インタビュで

は、36人中16人の被験者が緊張感について言及していた. その内、緊張した対話方法として、15人の被験者が対面条件を,5人の被験者がロボット条件を挙げていた.一方、ビデオ条件とアバタ条件を挙げた被験者は2人のみであり、静止ロボット条件と音声条件を挙げた被験者はいなかった.したがって、伝達情報が音声のみの場合には、実体による緊張感への影響は低いと考えられる.これらの結果は、ロボット会議が対面会議で感じるような緊張感を生み出す可能性を示唆している.

図7を見ると、アバタ条件およびビデオ条件の途切れの 頻度は、統計的有意差は無いが音声条件および静止ロボット条件よりも低い、被験者へのインタビュによると、実験 者の身体動作が見えると意見を聞いてくれていることが分 かるため、話しやすいと感じるようである。この感覚がア バタ条件やビデオ条件の途切れの頻度を減少させたのかも しれない、本実験と並行して行われた実験[22]においても、 アバタやビデオによる身体動作の伝達は途切れの頻度を減 少させ、発話を滑らかにするという結果が出ており、本実 験でも同様の効果が出たものと思われる。統計的有意差が 見られなかった原因は現在調査中であるが、本実験ではア バタやビデオを等身大で表示しているなど、実験環境にい くつかの差異があるためだと思われる。

#### 6. まとめ

本研究では、ロボット会議の有用性を示すため、遠隔会議における実体の効果を検証した。その結果、実体は対話相手のソーシャルテレプレゼンスを強化することが分かった。この実体の効果により、ロボット会議でのソーシャルテレプレゼンスは、対話相手の外見を提示しなくても、ビデオ会議と同等であった。この結果から、プライバシーの問題があるビデオ会議をロボット会議が代替する可能性が示唆された。さらに、対話相手の外見を伝えることもソーシャルテレプレゼンスを強化することが分かったため、ロボットが対話相手の外見を提示することができれば、ロボット会議でのソーシャルテレプレゼンスはビデオ会議を上回る可能性がある。

また、本研究では、発話に含まれる途切れの頻度を観察することで緊張感の測定を試みた.その結果、実体は緊張感を与えることが分かった.したがって、ロボット会議は、対面会議で感じるような緊張感を生み出す可能性があり、面接や講義などの用途で使用できるかもしれない.

謝辞 実験に協力していただいた山本健太氏,尾上聡氏に感謝する.本研究は,若手研究(A)「テレロボティックメディアによる社会的テレプレゼンスの支援」,基盤研究(S)「遠隔操作アンドロイドによるソーシャルテレプレゼンスの研究」,JST CREST「人の存在を伝達する携帯型遠隔操作アンドロイドの研究開発(研究領域:共生社会に向け

た人間調和型情報技術の構築)」,グローバル COE プログラム「認知脳理解に基づく未来工学創成」からの支援を受けた。

#### 参考文献

- Anderson, A.H., Newlands, A., Mullin, J., Fleming, A., Doherty-Sneddon, G. and Van Der Velden, J.M. Impact of Video-Mediated Communication on Simulated Service Encounters. Interacting with Computers, Vol.8, No.2, pp.193-206 (1996).
- [2] Bailenson, J.N., Yee, N., Merget, D. and Schroeder, R. The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. Presence: Teleoperators & Virtual Environments, Vol.15, No.4, pp.359-372 (2006).
- [3] Bainbridge, W.A., Hart, J., Kim, E.S. and Scassellati, B. The benefits of interactions with physically present robots over video-displayed agents. International Journal of Social Robotics, Vol.3, No.1, pp.41-52 (2011).
- [4] Bente, G., Ruggenberg, S., Kramer, N.C. and Eschenburg, F. Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations. Human Communication Research, Vol.34, No.2, pp.287-318 (2008).
- [5] Bondareva, Y. and Bouwhuis, D. Determinants of Social Presence in Videoconferencing. Proc. AVI2004 Workshop on Environments for Personalized Information Access, pp.1-9 (2004).
- [6] Bordia, P. Face-to-face versus computer-mediated communication: A synthesis of the experimental literature. Journal of Business Communication, Vol.34, No.1, pp.99-120 (1997).
- [7] Daly-Jones, O., Monk, A.F. and Watts, L. Some Advantages of Video Conferencing over High-quality Audio Conferencing: Fluency and Awareness of Attentional Focus. International Journal of Humancomputer Studies, Vol.49, No.1, pp.21-58 (1998).
- [8] de Greef, P. and Ijsselsteijn, W. Social Presence in a Home Tele-Application. CyberPsychology & Behavior, Vol.4, No.2, pp.307-315 (2001).
- [9] Finn, K.E., Sellen, A.J. and Wilbur, S.B. Video-Mediated Communication. Lawrence Erlbaum Associates, (1997).
- [10] Garau, M., Slater, M., Bee, S. and Sasse, M.A. The Impact of Eye Gaze on Communication Using Humanoid Avatars. Proc. CHI 2001, pp.309-316 (2001).
- [11] Isaacs, E.A. and Tang, J.C. What Video Can and Can't Do for Collaboration: A Case Study. Multimedia Systems, Vol.2, No.2, pp.63-73 (1994).
- [12] Kang, S., Watt, J.H. and Ala, S.K. Communicators' Perceptions of Social Presence as a Function of Avatar Realism in Small Display Mobile Communication Devices. Proc. HICSS 2008, (2008).
- [13] Kraaimaat, F.W., Vanryckeghem, M. and Van Dam-Baggen, R. Stuttering and social anxiety. Journal of Fluency Disorders, Vol.27, No.4, pp.319-331 (2002).
- [14] Kuzuoka, H., Yamazaki, K., Yamazaki, A., Kosaka, J., Suga, Y. and Heath, C. Dual Ecologies of Robot as Communication Media: Thoughts on Coordinating Orientations and Projectability. Proc. CHI 2004, pp.183-190, (2004).
- [15] Lee, K.M., Jung, Y., Kim, J. and Kim, S.R. 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, Vol.64, No.10, pp.962-973 (2006).

- [16] Morita, T., Mase, K., Hirano, Y. and Kajita, S. Reciprocal Attentive Communication in Remote Meeting with a Humanoid Robot. Proc. ICMI 2007, pp.228-235 (2007).
- [17] Nakanishi, H., Kato, K. and Ishiguro, H. Zoom Cameras and Movable Displays Enhance Social Telepresence. Proc. CHI 2011, pp.63-72 (2011).
- [18] Ogawa, K., Nishio, S., Koda, K., Balistreri, G., Watanabe, T. and Ishiguro, H. Exploring the Natural Reaction of Young and Aged Person with Telenoid in a Real World. Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.15, No.5, pp.592-597 (2011).
- [19] Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H. and Hagita, N. Android as a Telecommunication Medium with a Human-like Presence. Proc. HRI 2007, pp.193-200 (2007).
- [20] Sellen, A.J. Remote Conversations: The Effects of Mediating Talk with Technology. Human-Computer Interaction, Vol.10, No.4, pp.401-444 (1995).
- [21] Sirkin, D. and Ju, W. Consistency in physical and on-screen action improves perceptions of telepresence robots. Proc. HRI 2012, pp.57-64 (2012).
- [22] 尾上聡, 山本健太, 田中一晶, 中西英之: 身体動作の提示による遠隔対話の円滑化, インタラクション2012, pp.57-64 (2012).
- [23] 三澤 加奈, 石黒 祥生, 暦本 純一, LiveMask: 立体顔形状ディスプレイを用いたテレプレゼンスシステムにおけるコミュニケーションの評価, インタラクション2012, pp.41-48, (2012).