

ロボット操作者の偽存在感による ソーシャルテレプレゼンスの生成

田中 一晶^{†1, †2} 宇野 弘晃^{†1} 山下 直美^{†3} 中西 英之^{†1} 石黒 浩^{†4}

概要: 遠隔操作ロボットと自律ロボットの本質的な違いは遠隔地にいる操作者の存在の有無である。この存在の有無をユーザがどのように判断しているのかは未だ良く分かっていない。その判断のメカニズムを明らかにすることによって、自律ロボットとの対話を人との対話のように感じさせることが本研究の目的である。我々は、被験者が遠隔操作状態と自律状態のロボットとそれぞれ対話する実験をソーシャルテレプレゼンステストに基づいて行った。ソーシャルテレプレゼンステストとは、自律システムが人と同等の存在感を生み出すかどうかを評価する我々が考案したテストである。実験の結果、自律状態のロボットとの対話における遠隔操作者の存在感の有無は、遠隔操作状態の同じロボットとの対話を事前に行ったかどうかによって判断されることが分かった。これは、被験者が遠隔操作状態のロボットと事前に対話した場合、その際に感じた遠隔操作者の存在感が、自律状態のロボットとの対話でも継続して感じられたと考えられる。また、事前の対話において自律システムが遠隔操作者を装った場合であっても、自律状態のロボットとの対話で遠隔操作者の存在感が生み出されることも分かった。

Pseudo Presence of Robot Operator Produces Social Telepresence

KAZUAKI TANAKA^{†1, †2} HIROAKI UNO^{†1} NAOMI YAMASHITA^{†3}
HIDEYUKI NAKANISHI^{†1} HIROSHI ISHIGURO^{†4}

Abstract: The essential difference between a teleoperated robot and an autonomous robot is the presence or absence of a remote operator. It has remained unclear how the user who is talking with a robot judges this difference. The purpose of this study is to produce the sense of talking with a remote operator when talking with an autonomous humanoid robot by finding the user's judgment mechanism. We conducted experiments based on the social telepresence test, in which subjects talked with autonomous and teleoperated robots. We proposed the social telepresence test that evaluates whether an autonomous system can produce a presence. As a result, we found that subjects tended to judge the presence/absence of a remote operator by the presence/absence of their prior experience in which they talked with the remote operator through the robot. We considered that if the subjects previously talked with a teleoperated robot, they continued to feel the presence of remote operator even while talking with the autonomous robot that has a same appearance with the teleoperated robot. Furthermore, even if subjects had talked with an autonomous robot under the guise of a remote operator, such a pseudo presence also seemed to produce the presence of remote operator when they later on talked with an autonomous robot.

1. はじめに

ヒューマノイドロボットの種類には大きく分けて遠隔操作ロボットと自律ロボットがある。遠隔操作ロボットは遠隔操作者の存在感を伝えるものであるのに対し[18][22]、自律ロボットは人間のような存在感を生み出すものである[3]。したがって、両者の本質的な違いは遠隔操作者の存在の有無と考えることができる。この違いをロボットと対話するユーザがどのように判断しているかは明らかになっていない。本研究の目的は、その判断のメカニズムを明らかにし、自律ロボットとの対話でも人と話しているようにユーザに感じさせることである。

遠隔操作ロボットの特徴は、物理的実体のあるロボット

で遠隔操作者の身体動作をユーザに提示できる点である。先行研究では、他の遠隔コミュニケーションメディアに対する遠隔操作ロボットの優位性がいくつか報告されている[14][18]。存在感伝達もその1つであり、人の外見に酷似した遠隔操作ロボットはビデオチャットやボイスチャットよりも高いソーシャルテレプレゼンス（遠隔地にいる人と対面している感覚[5]）を生み出すことが分かっている[18]。リアルな外見を持ったロボットをユーザに応じて製作することはコストの問題から現実的ではないため、人と分かる顔を持つが特定の人物の外見は持たない遠隔操作ロボットも提案されている[16]。そのような匿名のロボットであっても、アバターチャットやボイスチャットよりも高いソーシャルテレプレゼンスを生み出すことができる[22]。アバターチャット[1][2][7][10][21]はコンピュータグラフィックスのアニメーションで相手の身体動作を提示するメディアであり、遠隔操作ロボットと類似した特長を持つが、物理的実体を持たない点において異なる。したがって、物理的実体で身体動作を提示することが遠隔操作者の存在を感じさせ

†1 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

†2 独立行政法人科学技術振興機構, CREST
CREST, Japan Science and Technology Agency

†3 NTTコミュニケーション科学基礎研究所
NTT Communication Science Laboratories

†4 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

る要因かもしれない。

遠隔操作者の身体動作を伝達する代わりに、その動作を自律化する方法が提案されている。具体的には、人の音声情報から発話動作を自動生成するアルゴリズムである[4][11][12][13][19][26]。ロボットとの対話では、ビデオチャットとは異なり相手の外見や振る舞いを直接見ることができないため、自動生成されたロボットの発話動作が人間らしければ、その動作が遠隔操作者の振る舞いに基づいているとユーザは信じるかもしれない。これが可能であれば、録音音声から発話動作を自動生成することで、遠隔操作者が存在しなくてもユーザはその存在を感じながら自律ロボットと対話できる可能性がある。しかし、その動作が自動生成であるとユーザが知っていた場合には、その感覚は低下する恐れがある。発話動作を生成するアルゴリズムによって動作する自律ロボットが遠隔操作者の存在感を生み出すかどうかは先行研究では明らかにされていない。本研究では、遠隔操作モードから自律モードにロボットの対話モードを切り替える方法によって、たとえユーザが自律ロボットと対話していることを知っていても、遠隔操作者の存在が感じられることを実験で示した。

まず、遠隔操作者の発話動作に同期して動く遠隔操作モードのロボットを介してユーザは遠隔操作者と対話する。次に、自動生成された発話動作を提示する自律モードにロボットの対話モードを切り替え、ユーザはその自律ロボットと対話する。自律モードのロボットは、事前に録音された音声から唇の動きなどの発話動作を生成し、ユーザの発話から頷き動作を生成する。我々は、遠隔操作モードでの対話で感じた遠隔操作者の存在感が、自律モードでの対話で想起され、その存在感が継続して得られるのではないかと考えた。例えば、ユーザの話し相手になるコミュニケーションロボットの場合[17][23]、一人暮らしのユーザが遠隔介護者と遠隔操作モードで対話すると、自律モードに切り替えた後もユーザは遠隔介護者の存在が感じられ、効果的に孤独感が低減されるかもしれない。また、遠隔講義ロボットの場合[8]、最初に遠隔講師が遠隔操作モードで生徒と挨拶を交わすと、その後に講義が自律モードで行われたとしても、生徒は講師の存在が感じられ、真面目に講義を受けるかもしれない。遠隔操作モードと自律モードの両方を備えたヒューマノイドロボットは既に提案されているが[17]、それらのモードを切り替えることによる存在感への影響は明らかになっていない。

本研究では、上述の予想をチューリングテストに類似した実験で検証した。一般的なチューリングテストでは、被験者は自律システムに人間同様の知的能力があるかどうかを判断するのにに対し、我々の実験では、被験者は自律システムに人間同様の存在感があるかどうかを判断する。我々はこれをソーシャルテレプレゼンステストと名づけた。

自律ロボットが遠隔操作者の存在感を生み出す要因を明

らかにするため、ソーシャルテレプレゼンステストに基づいた3つの実験を行った。まず実験1では、ロボットの持つ物理的実体の効果を検証した。ボイスチャットのような音声のみの対話も自律化することが可能であり、音声のみで対話する自律システムでも人の存在感を生み出すことができれば後述の実験2,3でロボットを使用する必要は無い。したがって、物理的実体の効果は最初に検証する必要があった。

実験1では、遠隔操作モードでの事前の対話において、自律システムが遠隔操作者を代替していた。つまり、実際には自律モードだが、被験者には遠隔操作モードと偽って実験を行った。実験の結果、このような遠隔操作者の偽の存在感であっても、その後の自律ロボットとの対話において遠隔操作者の存在感を生み出すことができた。実験2では、実際に遠隔操作者と対話した場合の実際の存在感が、自律ロボットとの対話で遠隔操作者の存在感をより効果的に生み出すかを、偽の存在感と比較することで検証した。

実験1,2では、被験者が話し手であり、ロボットは聞き手として相槌を返した。実験3では、両者の役割を反対にし、ロボットが被験者に一方的に話しかける対話で実験を行った。このようなインタラクティブ性を欠いた対話では、ビデオメッセージを見ているのと同様であるため、遠隔操作者の存在を感じにくいことが予想される。実験3では、インタラクティブ性を欠いた対話においても対話モードを切り替える我々のアプローチが遠隔操作者の存在感を生み出すか検証した。

2. 対話モード

発話動作を生成するアルゴリズムを提案した研究のほとんどは顔の動作(例えば頷きや唇の動作)に注目していた。それらは対話における最も基本的な動作であるため、本研究においてもそれらの動作を扱う。既存の遠隔操作ロボットの多くはそれらの動作を提示できる頭部を備えている[8][16][18][26]。本研究では、人のような顔を持つが特定の人物の外見を持たないヒューマノイドロボットであるTelenoidを使用した[16]。このロボットは3自由度の首と1自由度の口を備えている。我々は以下で説明する遠隔操作モードと自律モードでこのロボットを操作した。また、後述の3つの実験では、被験者の事前知識を統制するため、遠隔操作モードと自律モードについて説明する紙を実験の前に被験者に提示し、口頭でも同様の説明を行った。

2.1 遠隔操作モード

このモードでは、フェイストラッキングソフトであるfaceAPIで取得した遠隔操作者の顔のトラッキングデータに基づいて30fpsでロボットの首と口の角度をコントロールする。このソフトは遠隔地にある端末で動作し、Webカメラによって遠隔操作者の首と口の動きを取得する。

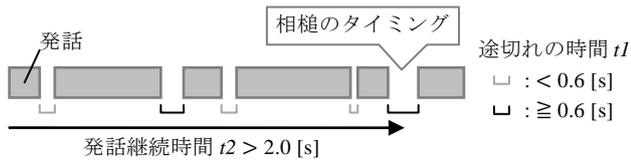


図 1 相槌のタイミングを検出するルール

Figure 1. Method to detect timing of back-channel response.

2.2 自律モード

ユーザとの対話における遠隔操作者の役割は聞き手と話し手に分けることができ、それらの役割における主な動作はそれぞれ頷きと発話である。したがって、我々は相槌のタイミングをユーザの発話から検出する相槌システムと、遠隔操作者の発話とロボットの口の動きを同期させるリップシンクシステムを構築した。

対話モードを切り替える我々のアプローチによって、非常に単純な自律システムであっても遠隔操作者の存在感を生み出すことができた場合、より自然で多様な発話動作を生成する技術を用いた場合においても我々のアプローチが有効に働くと考え、我々は相槌システムとリップシンクシステムを可能な限り単純化した。

2.2.1 相槌システム

相槌のタイミングを検出する多くの方法が存在し、そのタイミングの判断材料として発話の途切れ[15][20][24][26]、基本周波数[6][15][24][25]が主に用いられている。発話の途切れは、相槌のタイミングとして適切と思われる文章の切れ目や末尾を判断する有効な手がかりであるため、我々の方法は発話の途切れのみを使用した。

図 1 に相槌のタイミングを検出するルールを示す。四角は一続きの発話を示しており、それらの間隔が途切れの時間 $t1$ である。発話部分と途切れ部分は音圧の高/低にそれぞれ対応する。相槌システムは 0.6 秒以上の途切れのみを相槌の候補とし、それより短い途切れは全て無視する。発話継続時間 $t2$ は、発話の開始から相槌の候補となる途切れまでの時間である。 $t2$ が 2.0 秒以上のとき、システムは途切れを相槌のタイミングと判断する。つまり、このルールは 2.0 秒以上発話が継続した後の途切れを相槌のタイミングと判断するものである。 $t2$ は相槌のタイミングが決定した場合と $t1$ が 1.2 秒以上になった場合にゼロにリセットされる。

相槌のタイミングにおいて、ロボットは頷き動作と予め録音した相槌の音声「はい」を再生する。予備実験では、1 種類の頷き動作と音声を用いたため、被験者からロボットの相槌が一定であることを指摘された。そこで、ピッチ角度と速度の異なる 3 種類の頷き動作と、音程がわずかに異なる 2 種類の音声を用意し、それらを相槌のタイミングでランダムに選んで再生した。この方法により、相槌が一定であることを被験者に指摘されることは無くなった。

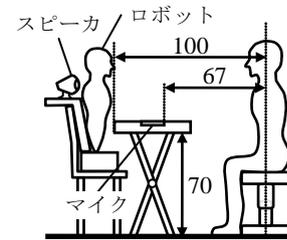


図 2 実験環境 (単位: cm)

Figure 2. Experimental setup (length unit: centimeters).

2.2.2 リップシンクシステム

人の発話から唇の動きを生成し、ロボット[9][26]やコンピュータグラフィックスのアバタ[4][19][26]の口を操作する様々なリップシンクの方法が提案されている。本研究で使用するロボットの口は 1 自由度であり、高精度のリップシンクは必要無いため、単純化することができた。

我々のリップシンクシステムは、人の発話の音圧を測定し、その大きさに合わせてロボットの顎の角度を変化させるものである。つまり、ロボットの口の開閉度合は発話音声の波形に基づいている。実験では、このシステムを用いて予め録音した遠隔操作者の発話からロボットの口の動きを生成した。

3. 実験環境

図 2 に実験環境を示す。実験では、ロボットを机の前に設置し、被験者とその反対側に座らせた。被験者の発話を取得する指向性マイクはその机の上面に埋め込み、被験者から直接見えないように布で覆った。遠隔操作者の発話を再生するスピーカはロボットの背後に設置した。

4. 実験 1

この実験では、遠隔操作モードでの対話で遠隔操作者の存在を感じることが、自律モードでの対話において遠隔操作者の存在感を生み出すのか調査する。また、遠隔操作者の存在感を生み出す上で、物理的実体のあるロボットが必要であるかについても調査する。

遠隔操作ロボットと同様に、ボイスチャットのような音声のみの対話も遠隔地にいる相手の姿を提示しないため、自律化することが可能である。音声のみの対話でも自律的に応答するシステムとの対話が相手の存在感を生み出すことができるならば、物理的実体は不要である。

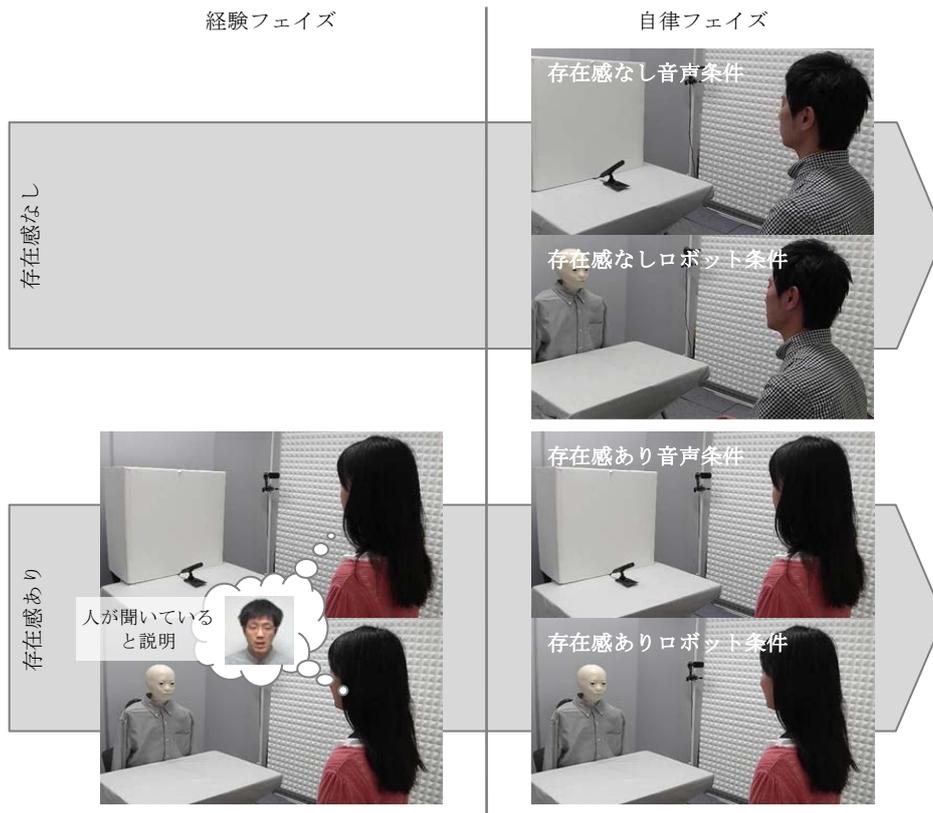


図 3 実験 1 の条件
Figure 3. Conditions of first experiment.

先行研究では、音声のみの対話と比較して遠隔操作ロボットを介した対話は、物理的実体の効果によって高いソーシャルプレゼンスを生み出すことが報告されている[22]。我々は、自律ロボットとの対話においても物理的実体が遠隔操作者の存在感を生み出す上で有効に働くと考えた。実験 1 の仮説は以下の通りである。

仮説 1: ユーザが遠隔操作状態のロボットと対話した場合、自律状態の同じロボットとの対話においてもユーザは遠隔操作者と対話しているように感じる。

4.1 実験条件

仮説 1 を調査するため、図 3 に示す 4 つの実験条件を設定した。実験は、遠隔操作状態のロボットと対話する段階（経験フェイズ）と自律状態のロボットと対話する段階（自律フェイズ）に分けて行った。経験フェイズは存在感あり条件にのみ含まれている。経験フェイズの前に、我々は、遠隔地にいる人とロボット／システムを介して対話することを被験者に説明した。しかし、被験者の発話に対して実験者が相槌を返した場合、相槌のタイミングを統制することが困難であったため、実際には経験フェイズでも自律モードを使用して相槌を返した。全ての被験者が我々の説明を信じたことは 4.4 節で述べるアンケートで確認した。

一方、自律フェイズは全ての条件に含まれている。自律フェイズの前に、我々は、遠隔地にいる人の代わりに自律

モードのロボット／システムが相槌を返し、被験者の発話は録音されていることを被験者に説明した。自律フェイズでは、被験者への説明の通りに自律モードで実験を行った。ロボット条件では、被験者の発話に対しロボットが録音音声の相槌と頷き動作を再生した。これに対し、音声条件では、録音音声の相槌のみを再生した。この条件では、ロボットは使用しないため白い箱で隠した。

ロボットの対話モードがどちらであるかという実験設定は、各フェイズの実験の前に被験者に説明した。

4.2 被験者

実験 1 には 16 人の学部生が被験者として参加した。8 人（女性 5 名、男性 3 名）は存在感あり音声／ロボット条件に参加した。この条件では、被験者は経験フェイズと自律フェイズにおいて音声のみの対話とロボットを介した対話の両方を行った。残りの 8 人（女性 4 人、男性 4 人）は存在感なし音声／ロボット条件に参加し、自律フェイズにおいて音声のみの対話とロボットを介した対話を行った。音声のみの対話とロボットを介した対話の順番はカウンターバランスを取った。

4.3 タスク

この実験では、被験者が聞き手であった場合、音声条件は一方的に録音音声を再生するだけであるため、話し手の身体動作を再生できるロボットと比較して不利になる可能性があった。この理由から、実験 1 では、被験者が話し手

になり、ロボット／システムは聞き手として被験者の発話に相槌を返すタスクとした。

実験の開始時、スピーカを通して電子機器について意見を述べるように被験者に録音音声で指示を与えた。この指示や相槌は当研究室のメンバーの音声を録音したものである。経験フェイズにおける話題は携帯音楽プレイヤーとロボット掃除機、自律フェイズにおける話題はスマートフォンと3DTVとした。各フェイズにおいて話題が2つあるのは音声条件とロボット条件があるからであり、それらの話題と条件の組み合わせはカウンターバランスを取った。

4.4 アンケート

1つの話題について話した後、被験者が正しく実験設定を理解していることを確認するため、はい／いいえで回答する下記のアンケートを実施した。

- ・ 先ほどの実験では、別の部屋に人がいて、あなたの話を聞いていた。
- ・ 先ほどの実験では、別の部屋に人はおらず、あなたの話は録音されていた。

自律フェイズの後、被験者が自律フェイズで感じた対話相手の存在感を計測するアンケートを実施した。アンケートの項目は以下の通りである。

- ・ 同じ部屋の中で人に話を聞いてもらっているように感じた。

この項目には、被験者は7段階のリッカート尺度で回答した。1～7は、「全くあてはまらない」、「あてはまらない」、「ややあてはまらない」、「どちらともいえない」、「ややあてはまる」、「あてはまる」、「非常によくあてはまる」に対応させた。また、被験者は回答の理由を自由記述欄に記述した。

4.5 結果

アンケートの結果、全ての被験者が正しく実験設定を理解していた。実験1の結果を図4に示す。グラフは「人に話しを聞いてもらっている感覚」の平均値であり、エラーバーは標準誤差を示す。また、存在感なし条件と存在感あり条件を対応なしt検定で比較した結果を図中に示す。

音声のみの対話では、存在感なし条件と存在感あり条件の間に有意な差は無かった ($t(14)=0.664, n.s.$)。一方、ロボットを介した対話では、存在感あり条件は存在感なし条件よりも有意に高いことが示された ($t(14)=2.575, p<.05$)。これは、ロボットを通して対話相手の顔き動作を見ることができる場合、経験フェイズにおける対話相手の存在感が自律フェイズでも生み出されたことを意味している。しかし、音声のみの対話の場合、対話相手の存在感は自律フェイズでは生み出されなかった。これらの結果は4節で述べた仮説1を支持する。

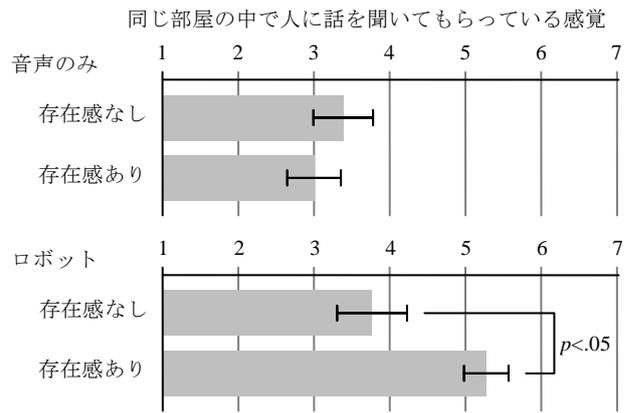


図4 実験1の結果

Figure 4. Result of first experiment.

存在感あり条件の被験者は対話を2回行ったが、存在感なし条件の被験者は対話を1回のみ行った。この回数の差が対話相手と話している感覚に影響を与えた恐れがある。しかし、音声のみの対話においても対話の回数に差があったが、存在感なし条件と存在感あり条件の差は有意ではなかった。したがって、自律フェイズにおいて対話相手の存在感を生み出したのは、対話回数ではなく、ロボットの持つ物理的実体の効果であったと思われる。さらに、次節で述べる実験2では、各条件の対話回数を統制して遠隔操作者の存在感の有無を比較した。

5. 実験2

実験1の結果、経験フェイズにおいて遠隔操作者を装った自律システムが相槌を返した場合（偽存在感）でも、自律フェイズにおいて遠隔操作者の存在感を生み出すことができた。これに対し、実際に遠隔操作者がユーザの発話に応える場合（実存在感）には、自然で多様な反応を返すことにより遠隔操作者の存在をより強く印象付けられるため、偽存在感よりも効果的に働く可能性がある。実験2の仮説は以下の通りである。

仮説2: 遠隔操作者を装った自律ロボットと対話した場合と比較して、実際に遠隔操作ロボットと対話した場合には、自律ロボットとの対話において遠隔操作者と対話している感覚が向上する。

さらに、全ての条件において対話回数を統制して仮説1を検証した。

5.1 実験条件

仮説1, 2を検証するため、図5に示す3つの条件を設定した。存在感なし条件は、実験1の存在感なしロボット条件に相当するが、被験者は自律モードのロボットとの対話を2回行う。偽存在感条件は、実験1の存在感ありロボット条件に相当する。これらの条件では、両方のフェイズにおいてロボットは全て自律モードでコントロールされる。

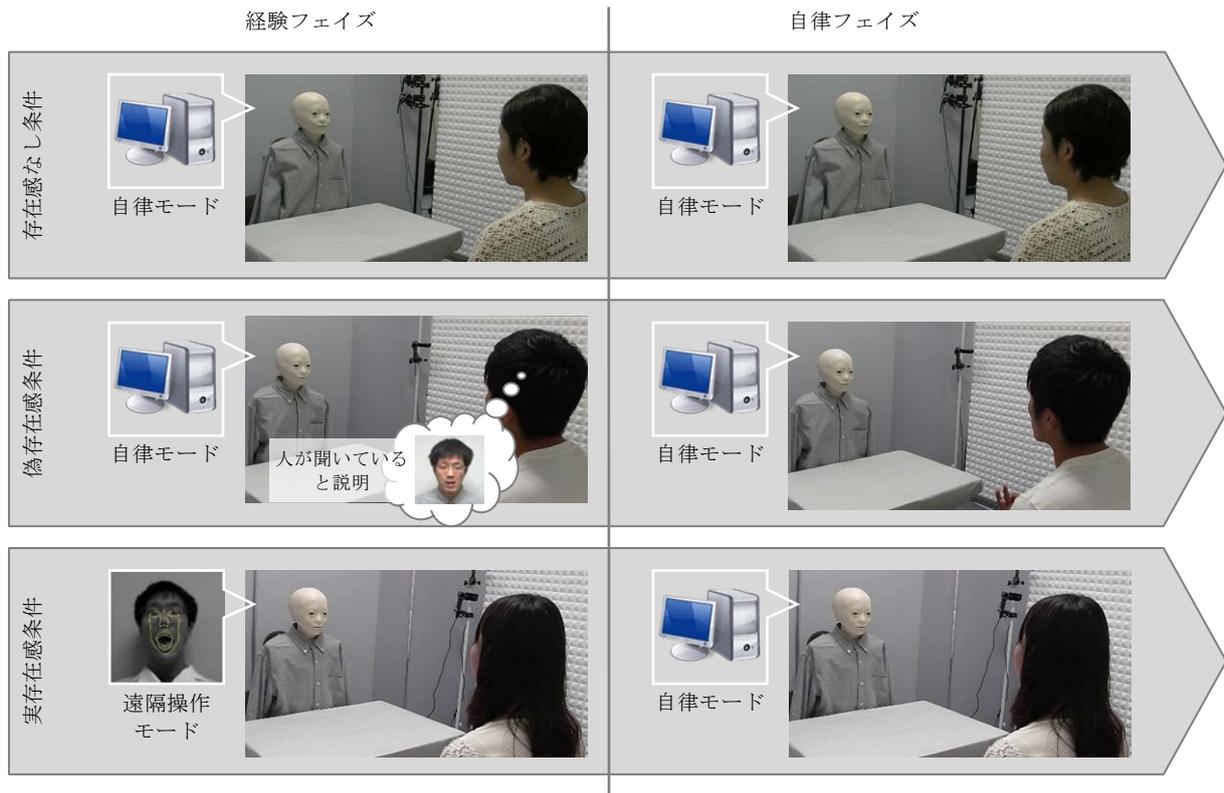


図 5 実験 2, 3 の条件
Figure 5. Conditions of second and third experiments.

したがって、これらの条件の差は経験フェイズにおける被験者への説明のみであり、存在感なし条件では自律モードであると正しい説明を行うが、偽存在感条件では遠隔操作モードであると偽った説明を行う。

仮説 2 の検証のために新たに追加した条件が実存在感条件である。この条件では、遠隔操作者が被験者の発話に対して相槌を打つだけでなく、被験者の意見を繰り返したり、言い換えて確認したり、多様な反応を返す。このような反応を自律化することは相槌のタイミングの検出よりも困難であるため、経験フェイズでは遠隔操作モードを用いてロボットをコントロールした。

ロボットの対話モードがどちらであるかという実験設定は、各フェイズの実験の前に被験者に説明した。

5.2 被験者

実験 2 には 30 人の学部生が被験者として参加した。10 人（女性 6 人、男性 4 人）は存在感なし条件に参加した。10 人（女性 5 人、男性 5 人）は偽存在感条件に参加した。残りの 10 人（女性 5 名、男性 5 名）は実存在感条件に参加した。

5.3 タスク

タスクは基本的に実験 1 と同様である（4.3 節）。存在感なし条件と偽存在感条件では、ロボットは当研究室のメンバーの録音音声を再生した。実存在感条件では、同じメンバーが遠隔操作者を務め、被験者への話題の指示や相槌を

実際に行った。経験フェイズと自律フェイズにおける話題はそれぞれ 3DTV とスマートフォンとした。

5.4 アンケート

実験を終えた後、被験者が正しく各フェイズの実験設定を理解していることを確認するため、はい/いいえで回答する下記のアンケートを実施した。

- ・ 1 回目の実験では（自律フェイズに対しては「2 回目の実験では」）、ロボットは遠隔操作モードで動いていた。
- ・ 1 回目の実験では、ロボットは自律モードで動いていた。

実験 1 のアンケートの自由記述欄では、自律的にロボットが相槌を返しているように感じた点と、遠隔操作者が相槌を返しているように感じた点を両方記述する場合がいくつか見られた。そこで、これらの感覚を別々に評価できるように以下の 2 つの項目を用意した。

- ・ ロボットは自動で相槌を打っている感じがした。
- ・ ロボットは質問者の相槌を伝えている感じがした。

被験者は 7 段階のリッカート尺度で回答し、その理由を自由記述欄に記述した。2 つ目の項目のスコアから 1 つ目の項目のスコアを引いた値を「遠隔操作者と対話している感覚」とし、この値を条件間で比較した。

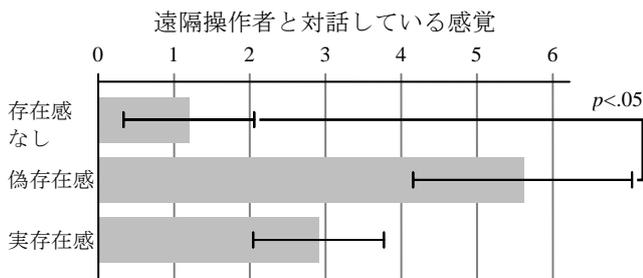


図 6 実験 2 の結果
Figure 6. Result of second experiment.

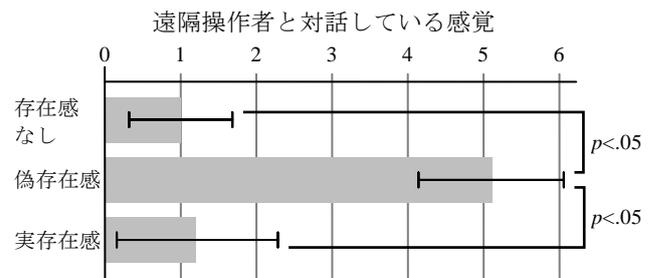


図 7 実験 3 の結果
Figure 7. Result of third experiment.

5.5 結果

アンケートの結果、全ての被験者が正しく実験設定を理解していた。実験 2 の結果を図 6 に示す。グラフは「遠隔操作者と対話している感覚」の平均値であり、全ての値が正になるように 4 を足した。エラーバーは標準誤差を示す。また、3 つの条件の比較を対応なし一元配置分散分析で行い、ボンフェローニ補正法を用いて多重比較を行った結果を図中に示す。

分散分析の結果、3 つの条件間の差が有意であった ($F(2, 27)=4.881, p<.05$)。多重比較の結果、偽存在感条件は存在感なし条件よりも有意に高いことが示された ($p<.05$)。これは、経験フェイズにおいて自律システムが遠隔操作者を装って被験者と対話した場合 (偽存在感) であっても、自律フェイズにおいて遠隔操作者の存在感が生み出されたことを意味している。この結果は実験 1 と同様に仮説 1 (4 節) を支持する。

一方、実存在感条件と存在感なし条件および偽存在感条件の差は有意ではなかった。したがって、仮説 2 (5 節) は示されなかった。

6. 実験 3

実験 1, 2 では、被験者が話し手でありロボットが聞き手であった。実験 3 では、これらの役割を反対にして仮説 1, 2 を検証する。この状況では、自律ロボットは録音音声を再生しながら発話動作を一方向的に提示するため、ビデオメッセージを再生している状況のように、対話相手の存在感を生み出しにくいことが予想される。この実験の目的はこのような非インタラクティブな対話においても我々のアプローチが有効に働くかを確認することである。

6.1 実験条件

実験条件は実験 2 と同様である (図 5)。この実験では、ロボットは被験者の発話に回答する必要が無いため、自律モードではリップシンクシステムのみ使用した。実存在感条件の経験フェイズでは、遠隔操作者は被験者に質問を行うため遠隔操作モードを使用した。被験者とインタラクションを行う実存在感条件と一方向的に録音音声と発話動作を再生する偽存在感条件を比較することで仮説 2 を検証した。

6.2 被験者

実験 3 には 30 人の学部生が被験者として参加した。各 10 人 (女性 5 人, 男性 5 人) はそれぞれ存在感なし条件, 偽存在感条件, 実存在感条件に参加した。

6.3 タスク

存在感なし条件, 偽存在感条件では、ロボットは電子機器について意見を述べた録音音声再生し、発話動作を提示した。この意見は当研究室のメンバーの音声を録音したものである。これらの条件では、被験者はロボットから再生される音声を聞くのみである。実存在感条件では、同じメンバーが遠隔操作者を務め、電子機器について意見を述べる中で 3 つの質問 (例えば「3D 映画を見たことはありませんか」など) を行い、被験者はそれに答えた。

6.4 アンケート

実験を終えた後、被験者が正しく各フェイズの実験設定を理解していること 5.4 節に示したアンケートで確認した。被験者が自律フェイズで感じた「遠隔操作者と対話している感覚」を評価するため、以下の 2 つの項目を用意した。

- ・ロボットは自動で発話動作を実行している感じがした。
- ・ロボットは遠隔操作で話者の発話動作を伝えている感じがした。

被験者は 7 段階のリッカート尺度で回答し、その理由を自由記述欄に記述した。2 つ目の項目のスコアから 1 つ目の項目のスコアを引いた値を「遠隔操作者と対話している感覚」とし、この値を条件間で比較した。

6.5 結果

アンケートの結果、全ての被験者が正しく実験設定を理解していた。実験 3 の結果を図 7 に示す。グラフは「遠隔操作者と対話している感覚」の平均値であり、全ての値が正になるように 4 を足した。エラーバーは標準誤差を示す。また、3 つの条件の比較を対応なし一元配置分散分析で行い、ボンフェローニ補正法を用いて多重比較を行った結果を図中に示す。

分散分析の結果、3 つの条件間の差が有意であった ($F(2, 27)= 5.806, p<.01$)。多重比較の結果、偽存在感条件は存在感なし条件よりも有意に高いことが示された ($p<.05$)。こ

れは、経験フェイズでの遠隔操作者の偽存在感によって、自律フェイズでも遠隔操作者の存在感が生み出されたことを意味している。したがって、ロボットが一方的に発話動作を提示する非インタラクティブな対話でも仮説 1 (4 節) が示された。

また、偽存在感条件は実存在感条件よりも有意に高く ($p < .05$)、実存在感条件と存在感なし条件との差は有意ではなかった。つまり、実存在感には自律フェイズにおいて遠隔操作者の存在感を生み出す効果は認められなかった。これは仮説 2 (5 節) に反する結果である。

7. インタラクション分析

実験 2, 3 の結果を観察データから確認するため、インタラクション分析を行った。これらの実験では、ロボット(または遠隔操作者)は対話の最後に「ありがとうございます」という音声で被験者に挨拶をした。我々は、この挨拶に被験者が応じるかどうかを観察した。被験者が遠隔操作者と対話していると感じていた場合には、挨拶に応じるかもしれないが、そうでなければ挨拶を無視するかもしれない。

分析の結果、多くの被験者がロボットの挨拶に対して「ありがとうございます」と言いながら会釈を行ったが、声のみまたは会釈のみで挨拶に応じた被験者もいたため、これらの反応を別々に数えることとした。その結果を以下の表に示す。

表 1 実験 2 においてロボットの挨拶に応じた被験者数
Table 1. Subjects who replied to robot's greeting in second experiment.

実験 2	経験フェイズ		自律フェイズ	
	声	会釈	声	会釈
存在感なし	7	7	6	7
偽存在感	10	10	9	10
実存在感	8	10	4	8

表 2 実験 3 においてロボットの挨拶に応じた被験者数
Table 2. Subjects who replied to robot's greeting in third experiment.

実験 3	経験フェイズ		自律フェイズ	
	声	会釈	声	会釈
存在感なし	2	5	1	3
偽存在感	5	8	5	7
実存在感	7	10	2	6

挨拶に応じた被験者の数は全体的に実験 2 よりも実験 3 の方が少ないことが見受けられる。これは実験タスクによる影響と思われる。実験 3 では、実存在感条件の経験フェ

イズを除いて、被験者はロボットの発話を聞くだけであり、その発話に応答する必要が無い。これにより、遠隔操作者の存在を感じにくかったため、挨拶に応じた被験者が少なかったと考えられる。実際、遠隔操作者とのやり取りがあった実存在感条件の経験フェイズでは、実験 2, 3 共にほぼ全ての被験者が挨拶に応じていた。

しかし、実存在感条件の自律フェイズでは、他の条件と比較し、挨拶に応じた被験者数が経験フェイズから大きく減少していることが見受けられる。これは自律フェイズにおいて遠隔操作者の存在感が低下したことを示しているのかもしれない。これに対し、存在感なし条件、偽存在感条件ではこのような大きな減少は見られない。しかし、存在感なし条件では、経験フェイズ、自律フェイズ共に低く、遠隔操作者の存在感の欠如が挨拶に応じた被験者数を減少させたと考えられる。したがって、挨拶に応じた被験者数が多く、自律フェイズにおいてその被験者数の減少が見られない偽存在感条件のみが両方のフェイズを通して高い存在感を保つことができたと思われる。

以上の通り、インタラクション分析において実験 2, 3 のアンケート結果を支持する結果が得られた。実存在感条件の自律フェイズにおいて遠隔操作者の存在感が減少した理由は次節で考察する。

8. 考察

我々の予想に反し、自律フェイズで遠隔操作者の存在感を生み出す上で、実存在感は効果的ではなく、偽存在感が最も効果的であった。アンケートの自由記述によると、実存在感条件に参加した実験 2, 3 の被験者(各 10 人)の半数が、ロボットとのインタラクション(様々な反応や問いかけ)が自律フェイズで減少したことに言及していた。したがって、実存在感条件では、経験フェイズと自律フェイズの間に、インタラクションの度にギャップがあったことが遠隔操作者の存在感を減少させたと考えられる。

実験 2 の結果(図 6)を見ると、差は有意では無かったが、実存在感条件の平均値は非存在感条件よりも高く、偽存在感条件よりも低かった。この結果は、遠隔操作者と実際に対話したことによるプラス効果が、インタラクションのギャップによるマイナス効果によって減少したものと考えられる。実験 2 の実存在感条件に参加した 10 人中 4 人の被験者は、自律フェイズでの相槌のタイミングは経験フェイズと大きな差が無かったと述べており、自律フェイズにおける適切な相槌がインタラクションのギャップをある程度軽減していたと思われる。一方、実験 3 の結果(図 7)を見ると、実存在感条件の平均値は非存在感条件とほぼ同じであり、偽存在感条件よりも有意に低かった。この実験では、自律フェイズにおいて被験者はロボットの話を聞くだけであったため、実存在感条件でのインタラクションのギャップは実験 2 よりも大きく、そのマイナス効果の増加

により、遠隔操作者と実際に対話したことによるプラス効果が完全に打ち消されたものと考えられる。7 節で述べたインタラクション分析においても、ロボットの挨拶に応じた被験者数の減少としてこのギャップの効果が表れている。

偽存在感条件では、ほとんどの被験者がロボットの挨拶に応じており、その被験者数の減少は見られなかった。偽存在感条件に参加した被験者は、経験フェイズにおいて自律ロボットとの対話を遠隔操作者との対話であると信じていたため、経験フェイズと自律フェイズのギャップを感じることはない。実際、実験 2 と 3 の偽存在感条件にそれぞれ参加した 10 人中 5 人と 6 人の被験者は、自律モードは遠隔操作モードとほとんど変わらないと述べており、これらのモードのギャップを曖昧にしたことが遠隔操作者の存在感を生み出す上で有効に働いたと考えられる。

自律フェイズでは、ロボットは 2.2 節で述べた非常に単純な方法で発話動作を生成した。より自然で多様な発話動作を生成できる技術を自律ロボットとのインタラクションに利用した場合、実際に遠隔操作者と対話するインタラクションとのギャップを埋められる可能性がある。したがって、発話動作生成に関する技術の進歩によって、実存在感も遠隔操作者の存在感を生み出す上で有効に働くかもしれない。この仮説の検証は、Wizard of Oz 法を用いて、経験フェイズと自律フェイズの両方で遠隔操作者が被験者と対話するが、自律フェイズでは自律システムでロボットが動作すると被験者に説明する条件を追加して実験を行うことで可能だと考えられる。この検証は今後の課題である。

実験 1 の結果から、物理的実体のあるロボットで身体動作を提示することが遠隔操作者の存在感を生み出す要因であることが分かった。しかし、物理的実体は無いが身体動作を提示することができるコンピュータグラフィックスのアバタでも同様の効果が得られるかは明らかになっていない。物理的実体と身体動作を切り分けてそれらの効果を検証することも今後の課題である。

9. まとめ

本研究では、ロボットが自律的に動いているのか遠隔操作者によってコントロールされているのかを人が判断するメカニズムの解明を目指し、我々が提案するソーシャルテレプレゼンステストに基づいた実験を行った。ソーシャルテレプレゼンステストは、自律システムが人間と同様の存在感を生み出すかどうかを評価するものである。実験の結果、物理的実体のあるロボットで身体動作を提示すること、ユーザがロボットを介して遠隔操作者と事前に対話することがソーシャルテレプレゼンステストに合格する要因であることが分かった。これは、ロボットを介した対話で感じた遠隔操作者の存在感を、ロボットが自律システムで動作する状態に切り替わった後でも継続して感じられたためだと推測される。しかし、遠隔操作者との事前の対話と自律

ロボットとの対話においてインタラクションの度合いにギャップがあると、自律ロボットとの対話で感じられる遠隔操作者の存在感が減少することも分かった。このギャップは、事前の対話において遠隔操作者を装って自律システムがユーザと対話することによって埋めることができ、自律ロボットとの対話で遠隔操作者の存在感を効果的に生み出すことができる。さらに、発話動作を生成する技術の進歩によって、ロボットの振る舞いの自然さや多様さを向上させることでもこのギャップを埋められる可能性ある。我々は、本研究によってテレロボティクスと知能ロボティクスの研究が相互に促進されることを期待している。

謝辞 本研究は、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」、基盤研究 (B)「ソーシャルテレプレゼンスのためのロボットエンハンスドディスプレイ」、SCOPE「遠隔身体インタラクションインタフェースの研究開発」、大阪大学と NTT との共同研究「遠隔地間で同じ部屋にいる状態を作り出す次世代コミュニケーション環境の研究」からの支援を受けた。

参考文献

- [1] 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*, 15(4), (2006), 359-372.
- [2] 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*, 34(2), (2008), 287-318.
- [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*, 3(1), (2011), 41-52.
- [4] Cao, Y., Tien, W.C., Faloutsos, P. and Pighin, F. Expressive Speech-Driven Facial Animation. *ACM Transactions on Graphics*, 24(4), (2005), 1283-1302.
- [5] Finn, K.E., Sellen, A.J. and Wilbur, S.B. Video-Mediated Communication. *Lawrence Erl-baum Associates*, (1997).
- [6] Fujie, S., Fukushima, K. and Kobayashi, T. A Conversation Robot with Back-Channel Feedback Function based on Linguistic and Nonlinguistic Information. *Proc. ICARA2004*, (2004), 379-384.
- [7] Garau, M., Slater, M., Bee, S. and Sasse, M.A. The Impact of Eye Gaze on Communication Using Humanoid Avatars. *Proc. CHI 2001*, (2001), 309-316.
- [8] Hashimoto, T., Kato, N. and Kobayashi, H. Development of Educational System with the Android Robot SAYA and Evaluation. *International Journal of Advanced Robotic Systems*, 8(3), (2011), 51-61.
- [9] Ishi, C., Liu, C., Ishiguro, H. and Hagita, N. Evaluation of formant-based lip motion generation in tele-operated humanoid robots. *Proc. IROS2012*, (2012).
- [10] 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. HICSS2008*, (2008).
- [11] Le, B.H., Ma, X. and Deng, Z. Live Speech Driven Head-and-Eye Motion Generators. *IEEE Transactions on Visualization and*

- Computer Graphics*, 18(11), (2012), 1902-1914.
- [12] Lee, J., Wang, Z. and Marsella, S., Evaluating Models of Speaker Head Nods for Virtual Agents, *Proc. AAMAS2010*, (2010), 1257-1264.
 - [13] Liu, C., Ishi, C.T., Ishiguro, H. and Hagita, N. Generation of Nodding, Head Tilting and Eye Gazing for Human-Robot Dialogue Interaction. *Proc. HRI2012*, (2012), 285-292.
 - [14] Morita, T., Mase, K., Hirano, Y. and Kajita, S. Reciprocal Attentive Communication in Remote Meeting with a Humanoid Robot. *Proc. ICM2007*, (2007), 228-235.
 - [15] Noguchi, H. and Den, Y. Prosody-Based Detection of the Context of Backchannel Responses, *Proc. ICSLP1998*, (1998).
 - [16] 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*, 15(5), (2011), 592-597.
 - [17] Ranatunga, I., Torres, N.A., Patterson, R.M., Bugnariu, N., Stevenson, M. and Popa, D.O. RoDiCA: a Human-Robot Interaction System for Treatment of Childhood Autism Spectrum Disorders. *Proc. PETRA2012*, (2012).
 - [18] Sakamoto, D., Kanda, T., Ono, T., Ishiguro, H. and Hagita, N. Android as a Telecommunication Medium with a Human-like Presence. *Proc. HRI2007*, (2007), 193-200.
 - [19] Salvi, G., Beskow, J., Moubayed, S.A. and Granstrom, B. SynFace: Speech-driven Facial Animation for Virtual Speech-reading Support. *EURASIP Journal on Audio, Speech, and Music Processing*, 2009(3), (2009).
 - [20] Takeuchi, M., Kitaoka, N. and Nakagawa, S. Generation of Natural Response Timing Using Decision Tree Based on Prosodic and Linguistic Information. *Proc. Interspeech2003*, (2003).
 - [21] Tanaka, K., Onoue, S., Nakanishi, H. and Ishiguro, H. Motion is Enough: How Real-Time Avatars Improve Distant Communication. *Proc. CTS2013*, (2013), 465-472.
 - [22] Tanaka, K., Nakanishi, H. and Ishiguro, H. Comparing Video, Avatar, and Robot Mediated Communication: Pros and Cons of Embodiment. *Proc. CollabTech2014*, CCIS460, (2014), 96-110.
 - [23] Tanaka, M., Ishii, A., Yamano, E., Ogikubo, H., Okazaki, M., Kamimura, K., Konishi, Y., Emoto, S. and Watanabe, Y. Effect of a human-type communication robot on cognitive function in elderly women living alone. *Medical Science Monitor*, 18(9), (2012), CR550-CR557.
 - [24] Truong, K.P. and Poppe, R. and Heylen, D. A rule-based backchannel prediction model using pitch and pause information. *Proc. Interspeech2010*, (2010), 26-30.
 - [25] Ward, N. and Tsukahara, W. Prosodic Features which Cue Back-channel Responses in English and Japanese. *Journal of Pragmatics*, 32(8), (2000), 1177-1207.
 - [26] Watanabe, T., Okubo, M., Nakashige, M. and Danbara, R. InterActor: Speech-Driven Embodied Interactive Actor. *International Journal of Human-Computer Interaction*, (2010), 17(1), 43-60.