

# テレプレゼンスロボットの自動相槌による遠隔対話経験の想起

宇野弘晃<sup>†1</sup> 田中一晶<sup>†1†2</sup> 中西英之<sup>†1</sup>

テレプレゼンスロボットは身体動作の伝達によって対話相手の存在感を伝えるものである。先行研究では、この身体動作を自動的に生成する様々な方法が提案されているが、自動生成であるとユーザが分かっていると相手の存在が感じられない恐れがある。我々は、身体動作を伝達するテレプレゼンスモードでの対話をユーザが事前に経験していると、身体動作を自動生成する自律ロボットとの対話で相手の存在を錯覚すると考えた。本研究では、会話中の身体動作として相槌に着目し、人が遠隔制御して相槌を行うテレプレゼンスモードと、人の発話に対して自動で相槌を行う自律モードでヒューマノイドロボットと対話する実験を行った。その結果、テレプレゼンスモードで対話した経験によって自律モードでも人と話している感覚が得られること、テレプレゼンスモードでの対話において自律モードでは実行されない多様な相槌を反すと対話経験の効果を弱めることが分かった。

## Remembrance of the remote conversation experience by the automatic nodding of the telepresence robot

HIROAKI UNO<sup>†1</sup> KAZUAKI TANAKA<sup>†1†2</sup>  
HIDEYUKI NAKANISHI<sup>†1</sup>

Telepresence robot is intended to convey the presence of a remote conversation partner by the transfer of body motion. In previous studies, various methods of automatically generating the body motion have been proposed. However, there has a risk that the presence of the partner is not felt when the user know that it is automatically generated. We thought the experience of a conversation with the robot of the telepresence state lets the user get an illusion of the presence of the partner in the conversation with the robot of the autonomous state. In this study, we focused on the nodding as body motion in a conversation. And we conducted an experiment in which subjects talked with the humanoid robot which has telepresence mode controlled by a remote person or autonomous mode which controlled by automatic-nodding system. As a result, we found that the experience of remote conversation with a telepresence robot enhance the illusion of the presence of the partner in the conversation with the autonomous robot. And also, a variety of reaction in a conversation with the telepresence robot can't enhance this illusion.

### 1. はじめに

テレプレゼンスロボットは身体動作の伝達によって遠隔地にいる対話相手の存在感を伝えるものである。近年では遠隔会議において、そのテレプレゼンス向上効果を期待して音声通話やビデオ会議に代わるものとしてテレプレゼンスロボットを用いたロボット会議の研究が盛んに行われている[8][9]。また教育や介護などの分野においても実際に人が移動することなく遠隔地で講演や講義[3][11]、介護対象との対話[12]などが行えるため、そのようなロボットの研究や活用がなされ始めている。

一方、テレプレゼンスロボットは操作者がロボットを操作する必要があるが、一部の動作を自動化する、または操作者の操作に自動生成した動作を重畳することで、労力の削減やインタラクションの活性化を行う動き[14]や自動で人間とインタラクションを行うコミュニケーションロボットの研究・開発も行われている[10]。

人間らしい動作を自動で生成する方法は多く研究されており、発話音声からアバタやロボットの頭部動作を生成

するもの[1][4][5]や、発話音声に対して相槌などの身体動作を提示しコミュニケーションを円滑化するもの[6]などがある。

しかしテレプレゼンスロボットの自動化を行う際に、ユーザがロボットの身体動作を自動で生成していると分かっていると、ユーザが感じる操作者のソーシャルテレプレゼンスが低下するおそれがある。我々の先行研究[15]では身体性とテレプレゼンスロボットとの遠隔対話の経験が自動で動いていると分かっているロボットとの対話でも人と対話しているような感覚を強化することが分かっている。本研究ではテレプレゼンスロボットとの遠隔対話の経験に着目し、どのような対話が人と対話している感覚を強化するのか、より詳しく調査する。

### 2. テレプレゼンスロボット

#### 2.1 ヒューマノイドロボット

本研究ではテレプレゼンスモードおよび自律モードを持つロボットとして特定の人物の外見を持たないテレプレゼンスロボットである Telenoid R2[7]を使用した。駆動部は首（Pitch 方向, Roll 方向, Yaw 方向の 3 回転軸）と口（開閉の 1 回転軸）の計 4 自由度である。

<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> 科学技術振興機構 CREST

CREST, Japan Science and Technology Agency

## 2.2 テレプレゼンスモード

テレプレゼンスモードでは実際に対話相手の身体動作を伝達するため、顔の傾き具合や口の開き具合などを検出する顔追従システム Face API を用いて対話相手の顔をトラッキングしたデータをロボットの動きに反映させた。また、机の内部に設置したマイクおよびロボットの背後に設置したスピーカを用いて、ユーザと対話相手が音声の相互伝達を行えるようにした。

## 2.3 自律モード

自律モードでは我々の先行研究で構築した自動相槌システムを用いて、ユーザの発話の途切れで音声と頷き動作による相槌を返した。

頷きのタイミングに関しては多く研究されているが[2][13]、この自動相槌システムはユーザの発話を音圧で検出し、一定値以上の音圧を“発話”、2.0秒以上の発話の後の0.6秒以上の途切れを“発話の途切れ”とみなして、発話の途切れに相槌を行うシステムである。また、発話中の詰まりや促音などから発話の途切れを誤検出しないために0.6秒以内に発話が再開された場合は発話を継続しているとみなした。

相槌を行うには、頷き動作は人の相槌のトラッキングデータを参考に予め作成した頷きの速さとピッチの異なる3種類のモーションから発話の途切れごとにランダムで選択し、相槌の音声はテレプレゼンスモードの際に対話相手を務める実験者の2種類の「ハイ」という音声の録音から発話の途切れごとにランダムに選択し、頷き動作と音声を同時に再生した。

## 3. 実験

### 3.1 仮説

先行研究では、身体性のある自律ロボットとの対話で、ロボットの相槌によってテレプレゼンスロボットを用いた操作者との遠隔対話の経験が想起され、ロボットを通して人と話している感覚が強化されることが分かっている。このことから以下の仮説を立てた。

**仮説1** テレプレゼンスロボットを通して遠隔地の操作者と対話した経験（遠隔対話経験）があれば、自律ロボットとの対話でもロボットが人の動きを伝達しているような感覚が強まる。

また、多様な反応が得られるインタラクションは相手の存在をより強く感じられるため、テレプレゼンスロボットを通して遠隔地の操作者と対話の際に自律ロボットが行わない「ハイ」以外の多様な相槌を行うことで操作者の存在をより強く感じ、自律ロボットとの対話でその経験が想起され、相手の存在がより強く感じられるのではないかと考えた。このことから以下の仮説を立てた。

**仮説2** テレプレゼンスロボットとの遠隔対話で、自律ロボットでは行わない多様な反応を返す対話を行うと、ロボットが人の動きを伝達しているような感覚がより強化



図1 実験の様子

Fig 1. Snapshot of the experiment

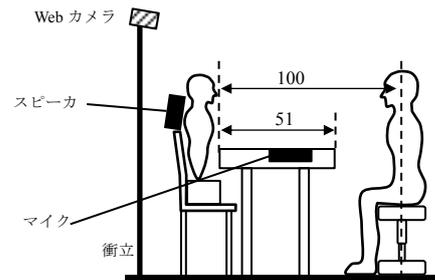


図2 実験環境（単位：cm）

Fig 2. Experimental setups (length unit: centimeters).

される。

### 3.2 実験方法

仮説を検証するため、テレプレゼンスロボットとの遠隔対話の経験の種類を要因とし、以下の3条件を設定した。

**自動経験条件：**テレプレゼンスロボットとの遠隔対話の経験を与えない統制条件である。各条件間でロボットとの対話の回数を統制するため、この条件では被験者は自律モードのロボットと2回対話を行った。

**単調経験条件：**自律モードと同程度のインタラクションを行う遠隔対話の経験を与える条件である。先行研究同様、被験者間で相槌のクオリティを統制するため、また、テレプレゼンスモードのロボットと遜色ない相槌を提示するために、被験者にはテレプレゼンスモードのロボットであると教示をした上で、自律モードのロボットと対話させテレプレゼンスロボットとの遠隔対話の経験を与え、その後、自律モードのロボットと対話させた。

**自然経験条件：**自律モードのロボットでは行わない多様な反応を返す遠隔対話の経験を与える条件である。この条件では被験者はテレプレゼンスモードのロボットを通して遠隔地の実験者と対話を行った後、自律モードのロボットと対話を行った。

この3条件で、大学生30人（各条件男性5人、女性5人）に対して被験者間計画で実験を実施した。

本実験では、被験者が事前に与えられたテーマについての意見を述べ、ロボットがそれに対して相槌を返すというタスクを設定した。まず、被験者にロボットのそれぞれの

モードに関する教示およびテーマの告知を行い、その後 2 回のタスクを行った。実験の様子を図 1、実験環境の模式図を図 2 に示す。被験者は椅子に座り、机を挟んでロボットと対話を行う。テレプレゼンスモードのとき、実験者は机の内部に設置したマイクを用いて被験者の音声を聞き、衝立の上部に設置した Web カメラを用いて被験者の様子を見ることができる。また、ロボットの背後に設置したスピーカから実験者の音声を流すことで被験者と対話を行う。自律モードのときは、マイクから取得した音声をを用いて自動相槌システムで被験者の発話を検出し、相槌の音声をスピーカから流す。また、自律モードのときは Web カメラを撤去した。1 回目のタスクは被験者に遠隔対話の経験を与える経験フェイズである。自動経験条件と単調経験条件では自律モード、自然経験条件ではテレプレゼンスモードのロボットでタスクを行った。単調経験条件では被験者にテレプレゼンスモードであると教示を行っているため、Web カメラを設置した。また、自然経験条件では、実験者は「ハイ」という相槌の他に、「ええ」「なるほど」といった相槌や、被験者の発言の一部を繰り返す発言など、被験者が意見を述べることを阻害せず、対話の主導権を握らないような相槌を行った。1 回目のタスクの後、2 回目のタスクでのテーマの告知を行った。2 回目のタスクは、どの条件とも自律モードのロボットでタスクを行う実験フェイズである。2 回のタスクの後、被験者にアンケートによって 2 回目のタスクに関する評価を行わせた。アンケートの項目は次節で述べる。

被験者が意見を述べるテーマは 1 回目のタスクを 3D テレビ、2 回目のタスクをスマートフォンと設定した。

### 3.3 アンケート

アンケートでは 2 回目のタスクに関して被験者が抱いた印象の度合いについて質問した。全 7 段階のリッカート尺度を用い、1: 全くあてはまらない、2: あてはまらない、3: ややあてはまらない、4: どちらともいえない、5: ややあてはまる、6: あてはまる、7: 非常によくあてはまる、に対応させた。アンケートの項目は以下の通りである。

- ロボットは自動で相槌を打っている感じがした。  
 この質問では、ロボットが自律ロボットであるように感じたかどうかを調べた。
  - ロボットは質問者の相槌を伝えている感じがした。  
 この質問では、ロボットがテレプレゼンスロボットであるように感じたかどうかを調べた。
- これらの質問は、被験者がロボットはどちらのモードで動いていたと感じたかを詳細に調べるために設定した。自律らしさ、テレプレゼンスらしさの片方のみを感じた被験者は一方のスコアを高く他方のスコアを低く付け、両方を感じた被験者は双方の評価をそれぞれ行うため、被験者の印象を詳しく尋ねることができる。また、どちらの被験者でも、スコアの差分を取ることで被験者がロボットを自律

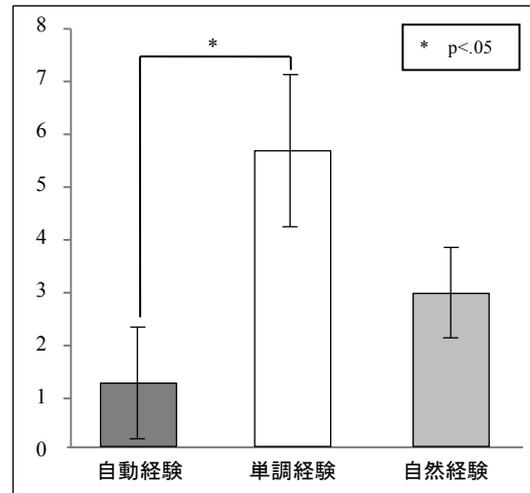


図 3 アンケート結果

Fig 3. Results of the questionnaire

ロボットまたはテレプレゼンスロボットのどちらだと思ったかを測ることができる。

## 4. 結果および考察

アンケートの結果を図 3 に示す。グラフのスコアはそれぞれの条件において「ロボットは質問者の相槌を伝えている感じがした。」のスコアから「ロボットは自動で相槌を打っている感じがした。」のスコアを引いた値の平均値に 4 を加えたものおよびその標準誤差である。スコアが高いほど被験者はロボットがテレプレゼンスを感じており、スコアが低いほど被験者はロボットが自動で動いているように感じていることになる。

このデータで一要因分散分析を行ったところ、要因間で有意な差が認められた( $F(2,27)=4.881, p<.05$ )。ボンフェローニ補正法で多重比較を行うと、自動経験条件と単調経験条件の間に有意差が見られた( $p=.013$ )。また、単調経験条件と自然経験条件の間、経験なし条件と自然経験条件の間には差がなかった。

この結果から、自動経験条件と単調経験条件の間では仮説 1 が成り立つが、自動経験条件と自然経験条件の間では仮説 1 が成り立たないこと、および仮説 2 が成り立たないことが分かる。

仮説 1, 2 に反して自然経験条件のスコアが低くなった理由について考察する。自然経験条件では操作者の存在をより強く感じさせるため、自律モードのロボットが行う「ハイ」という相槌以外の相槌を行った。しかし、多様な相槌を行うロボットを見た後で、「ハイ」のみの相槌を行うロボットを見たことで、対話のクオリティが低下したと感じ、スコアを下げる原因になったと考えられる。自然経験条件で低いスコア (スコアの差が 0 以下) をつけた被験者 7 人のうち 4 人は経験フェイズとの差を指摘しており、特に相槌

のバリエーションを指摘した被験者は3人いた。また、高いスコアをつけた被験者は「1回目とさほど違いがない」「人と同じような相槌」などとコメントしており、経験フェイズでの多様な相槌を重視していなかったことが分かった。このため、自律モードとテレプレゼンスモードの相槌のクオリティが同程度であれば、自律モードでの相槌からテレプレゼンスを感じられる可能性があり、自律モードの相槌クオリティの向上によってテレプレゼンスモードの相槌と遜色ない相槌を返すことで対話相手の存在を錯覚させる。

## 5. おわりに

本研究では、テレプレゼンスロボットとの遠隔対話の経験が同じデザインの自律ロボットとの対話において人と対話している感覚に与える影響を調べる実験を行った。その結果、テレプレゼンスロボットとの対話経験は、その対話と遜色ない自律ロボットとの対話においてロボットが人の動きを伝達している感覚が強化されることが分かった。また、自律ロボットよりも多様な反応を返す対話を含むテレプレゼンスロボットとの対話経験は、その対話のクオリティのギャップによって、自律ロボットとの対話においてロボットが人の動きを伝達しているような感覚を強化しないことが分かった。

**謝辞** 実験に協力していただいた大西裕也氏、加藤良治氏、塩崎恭平氏に深く感謝する。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（研究領域：共生社会に向けた人間調和型情報技術の構築）」、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」からの支援を受けた。

## 参考文献

- 1) BUSSO, Carlos, et al. Rigid head motion in expressive speech animation: Analysis and synthesis. *Audio, Speech, and Language Processing, IEEE Transactions on*, 2007, 15.3: 1075-1086.
- 2) FUJIE, Shinya; FUKUSHIMA, Kenta; KOBAYASHI, Tetsunori. A conversation robot with back-channel feedback function based on linguistic and nonlinguistic information. In: *Proc. ICARA Int. Conference on Autonomous Robots and Agents*. 2004. p. 379-384.
- 3) HASHIMOTO, Takuya; KOBAYASHI, Hiroshi; KATO, Naoki. Educational system with the android robot SAYA and field trial. In: *Fuzzy Systems (FUZZ), 2011 IEEE International Conference on*. IEEE, 2011. p. 766-771.
- 4) LE, Binh H.; MA, Xiaohan; DENG, Zhigang. Live speech driven head-and-eye motion generators. *IEEE transactions on visualization and computer graphics*, 2012, 18.11: 1902-1914.
- 5) LIU, Chaoran, et al. Generation of nodding, head tilting and eye gazing for human-robot dialogue interaction. In: *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on*. IEEE, 2012. p. 285-292.
- 6) OGAWA, Hiroki; WATANABE, Tomio. InterRobot: a speech

- driven embodied interaction robot. In: *Robot and Human Interactive Communication, 2000. RO-MAN 2000. Proceedings. 9th IEEE International Workshop on*. IEEE, 2000. p. 322-327.
- 7) 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.
- 8) SAKAMOTO, Daisuke, et al. Android as a telecommunication medium with a human-like presence. In: *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*. IEEE, 2007. p. 193-200.
- 9) TANAKA, Kazuaki; NAKANISHI, Hideyuki; ISHIGURO, Hiroshi. Robot conferencing: physically embodied motions enhance social telepresence. In: *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2014. p. 1591-1596.
- 10) TANAKA, Masaaki, et al. Effect of a human-type communication robot on cognitive function in elderly women living alone. *Medical science monitor: international medical journal of experimental and clinical research*, 2012, 18.9: CR550.
- 11) YAMAZAKI, Ryuji, et al. How does telenoid affect the communication between children in classroom setting?. In: *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2012. p. 351-366.
- 12) YAMAZAKI, Ryuji, et al. Teleoperated android as an embodied communication medium: A case study with demented elderlies in a care facility. In: *RO-MAN, 2012 IEEE*. IEEE, 2012. p. 1066-1071.
- 13) WATANABE, Tomio, et al. Interactor: Speech-driven embodied interactive actor. *International Journal of Human-Computer Interaction*, 2004, 17.1: 43-60.
- 14) 瀬島吉裕; 渡辺富夫; 山本倫也. うなずき反応モデルを重畳した VirtualActor を介する身体的コミュニケーションの合成的解析 (機械力学, 計測, 自動制御). *日本機械学会論文集. C 編*, 2009, 75.758: 2773-2782.
- 15) 宇野弘晃; 田中一品; 中西英之. チューリングテストに基づく自動相槌システムの研究. *2013 年度人工知能学会全国大会 (第27回)*, 2013.