

遠隔対話用ロボットの頷きの自動化と モーションチューリングテストによる検証

Evaluating Automatic Nods of Teleconferencing Robots with Motion Turing Test

大嶋 悠司*¹ 田中 一晶*² 中西 英之*¹ 石黒 浩*³
Yuji Oshima Kazuaki Tanaka Hideyuki Nakanishi Hiroshi Ishiguro

*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

This research suggests Motion Turing test which measure the credibility of that a robot is synchronized with human. In this test, a participant talks to robot which is synchronized with its operator or move automatically. Then, the participant judges the robot's motion in conversation the robot is synchronized with its operator and move automatically. We discover elements, which make believe motion of robot is synchronized with human by Motion Turing test experiment.

1. はじめに

近年、人の動きを捉えるモーショントラッキング装置が安価で手に入るようになり、ユーザの身体動作を模倣して遠隔地に伝えることのできる分身エージェントの研究が盛んになりつつある [Watanabe 99]. また人型ロボットの研究開発も盛んに行われており、Breazealらが開発した感情を表出する顔ロボット [Breazeal 99] や、石黒の開発した人間に酷似した外見をもつアンドロイド [Sakamoto 07] 等がある。これらのロボットは人の動きを再現することが可能であり、モーショントラッキング装置を用いて人の身体動作をリアルタイムに遠隔地に伝達する遠隔対話用ロボットが新たなコミュニケーションメディアとして研究され始めた [Kidd 04, Shinozawa 05].

これまでにも、身体動作を伝達する手段としてビデオ会議システムが通信メディアとして存在するが、自分のライブカメラ映像を送ると身体動作以外にも必要以上の情報が伝達され、プライバシーの問題が発生する。さらにビデオ会議システムには物理空間が共有できない [Kraut 03]、視線が合わない [morikawa 98] などの問題があり、これらビデオ会議システムの問題をすべて解決する手段として遠隔対話用ロボットは非常に有用なコミュニケーションメディアである。しかし、そのような遠隔対話用ロボットを用いた対話では、操作者の実際の姿を提示しないため、操作者の動きをリアルタイムに反映しているのではなく、何らかの規則で自動的に動いているのではないかと対話相手に疑われてしまうという問題がある。

本研究では遠隔対話用ロボットの身体動作を自動化することによってロボットの動きが人の動きに由来していると信じられる要素を解明することを目的とし、日常的な対話において行われる代表的な身体動作として頷きに注目して実験を行った。本論文では、ロボットの動きが人の動きに由来しており、人の動きをそのまま伝達していると信じられる度合いを動きの信憑性という言葉で表現する。本研究では、遠隔対話用ロボットの動きの信憑性を測る方法として、モーションチューリングテストを提案する。これは操作者の身体動作を再現するロボットを通して音声と身体動作による対話を行い、そのロボットの動き

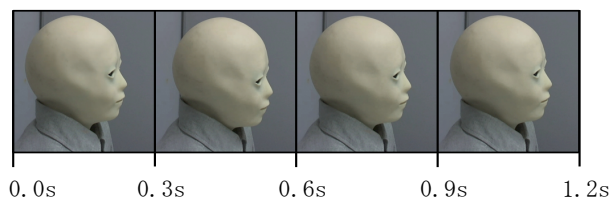


図 1: 作成したロボットの頷き

が操作者のものか、自動的に生成された作り物であるかを判断する方法である。

2. 仮説

会話中に行う大きな身体動作として、代表的なものは頷きであり、短い会話であれば頷き以外の大きな身体動作を行わない場合でも、自然なインタラクションが行える可能性が高い。我々は会話中のロボットの動きの信憑性には、頷き動作が自然であれば頷きのタイミングが重要であると考えた。

頷きと音声と同期していないと、信憑性を大きく損なう可能性があるため、「はい」という録音された音声を毎回の頷きに同期して再生させた。「はい」という録音音声は種類のみ用意した。これは、音声にバリエーションがあることが動きの信憑性に影響することを防ぐためである。

今回遠隔対話用ロボットとして、テレノイドを使用した。頷きに着目したため、首部のみ可動するようにし、動いていない口部が見えることが悪印象を与えないようにロボットにはマスクを着用した。

ロボットの頷きモーションは実際の頷きのトラッキングデータに近くなるように関数により再現した。頷きを再現する関数に従ってロボットが頷く様子が図 1 である。この頷きの動作は経験的なものであるが実験において違和感があると答えた被験者はいなかった。また頷いていない間に全く動かないということにロボットらしさを感じ、信憑性を損なう恐れがあったので会話中は常にランダムな微小動作を行っていた。

連絡先: 中西 英之, 大阪大学大学院工学研究科 知能・機能創成工学専攻, nakanishi@ams.eng.osaka-u.ac.jp

3. モーションチューリングテスト

本研究ではロボットの動きの信憑性を測る方法としてモーションチューリングテストという方法を提案する。モーションチューリングテストとは操作者の身体動作を再現するロボットを通して対話を行い、そのロボットの動きが操作者のものか自動的な作り物であるかを判断する方法である。通常のチューリングテストのコミュニケーションメディアはテキストチャットであり、変動することはない。しかしモーションチューリングテストの場合、使用する遠隔対話用ロボットによりコミュニケーションチャンネルのクオリティが変動する。モーションチューリングテストでは、ロボットの動きが作り物である可能性があることを被験者に伝えた上で、そのロボットを介して遠隔地の対話相手との会話をを行う。会話終了後にそのロボットの動きが本物なのか作り物なのか判断してもらう。判断結果の取得方法としては、本物か作り物かの判断が付くまで会話を続けてもらったうえでどちらなのか答えさせる方法と、一定時間の会話をを行った後に本物であると信じる度合いをアンケートで尋ねる方法がある。前者の場合は本物と判断する時間が短い程、後者の場合はアンケートの得点が高いほどテストの成績が高いということになる。

本研究における実験では後者の方法を選択した。

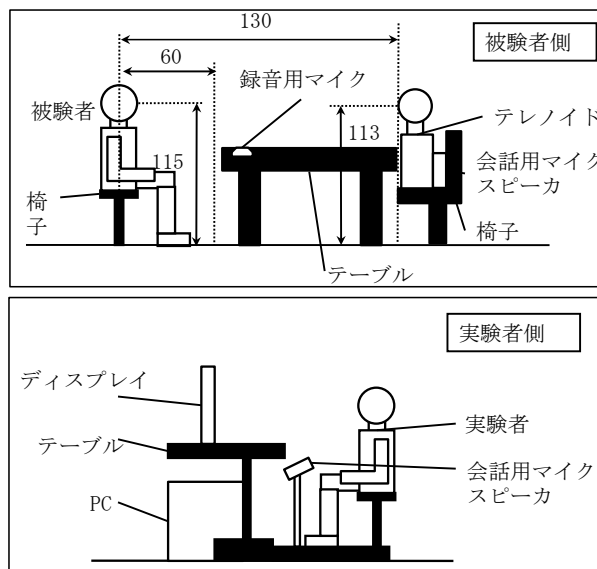
4. 実験

本研究で行ったモーションチューリングテストの実験環境の様子を図2に示す。実験者と被験者はそれぞれ別室に居り、被験者からは実験者の様子は見えない。被験者にはあらかじめ資料を配布し、実験に用いるロボットが首のみが動くこと、実験者の身体動作を再現している場合と自動で動く場合があることを説明する。また、実験の直前にもロボットの操作者ではない説明者が、同様の内容を説明する。その後、ロボットが相手の身体動作を再現できることを被験者に理解させるため、ロボットの同期確認を行う。その方法は、同期確認の間のみ実験者の様子が見えるディスプレイを被験者に見せ、ロボットが実験者の動きを再現できることを確認させる。同期確認の様子を図3に示す。そして被験者を一旦実験室から出し、実験者の様子が見えるディスプレイを撤去した後、実験者と被験者はロボットを通して90秒から120秒間会話し、会話終了後アンケートにてロボットの動きが、実験者の動きの再現と感じたか自動で動いていると感じたかを質問する。本研究では、頷きに注目しているのでロボットは常に会話の受け手であり、頷きのみ行うような実験を設定した。

4.1 実験条件

本研究の仮説は、頷きを行うタイミングが信憑性に影響するというものであり、そのなかでも頷きの間隔の影響を与えるかを調べるために3の条件を比較した。ロボットが7.7秒という一定間隔で頷く条件、6.0秒一定間隔で頷く条件、頷きの間隔をランダムで決定し頷く条件である。ランダムで頷く条件の会話全体での頷きの回数は7.7秒間隔で頷くものと同程度になるよう設定しており、頷きの間隔は $\mu = 6.5$, $\sigma^2 = 16$ の正規分布である。ただし頷きの間隔は最低2秒、最高18秒の間で計算される。

どの条件でも再生する頷きのモーションは同じであり、頷きに同期して、録音された「はい」という種類の音声を再生した。



単位:cm

図2: 実験環境の様子

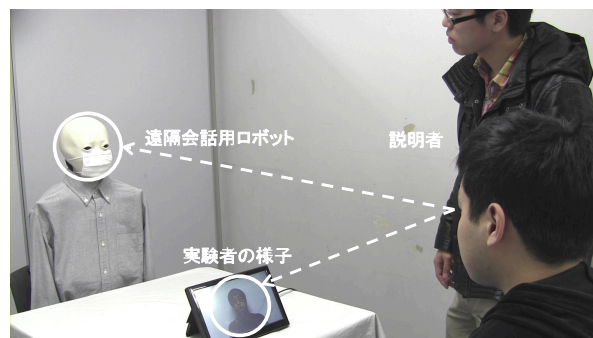


図3: 同期確認の様子

4.2 実験結果

ランダム条件と7.7秒一定間隔条件で各6人、6.0秒一定間隔条件で5人の被験者にロボットと会話してもらい、アンケートによる主観評価を用いて比較した。アンケートは動きの信憑性、頷きのタイミングに関しては11段階、話しやすさ、同じ部屋の中で話しかけている感覚に関しては9段階のリッカート尺度を用いて評価した。図4に実験結果を示す。アンケート結果に対し一要因分散分析(3水準)を行った。

信憑性について有意な主効果 ($F(2,14)=4.33$ $p<0.05$) が見られた。下位検定の結果7.7秒一定間隔条件が6.0秒一定間隔条件より有意 ($p<0.05$) に高かった。特に7.7秒一定間隔条件では半数の被験者が信憑性において最も高いスコアである11を付けていた。これら最高のスコアを付けた被験者の全員がスコアの理由に頷きが文脈に沿っていた、音声と同時に頷いたという理由のどちらか、または両方であると答えた。これから、頷きが文脈に依存しない場合でも、頷きの間隔が適切であれば信憑性を確保できることが分かる。

タイミングに関しては有意傾向 ($F(2,14)=3.28$ $p<0.1$) であり、下位検定の結果7.7秒一定間隔が6.0秒一定間隔に対して

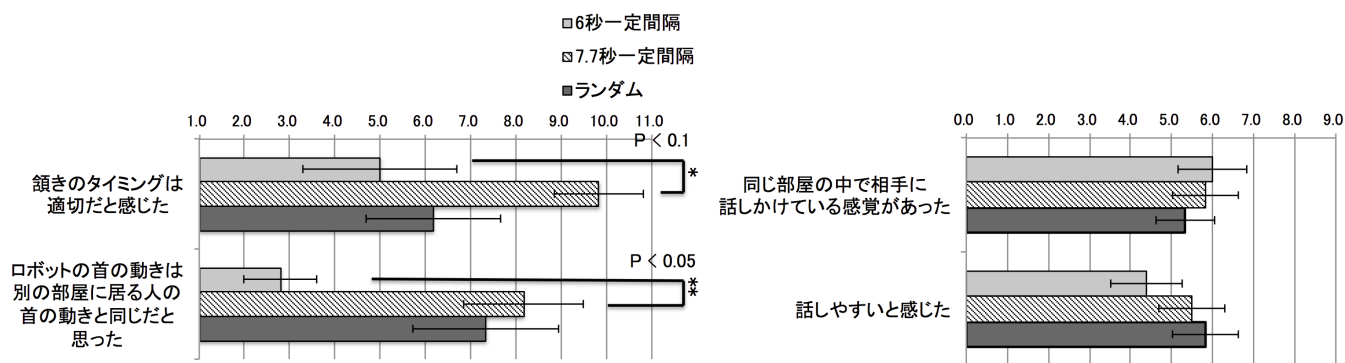


図 4: 実験結果

タイミングは適切であると判断された ($p < 0.1$).

話しやすさ、同じ部屋で相手に話しかけている間隔については主効果は見られなかった。

5. 考察

7.7 秒一定間隔条件と 6.0 秒一定間隔条件に有意な差が見られたことから、信憑性に大きく影響するのは傾きの頻度もしくは間隔である。今回の実験においてタイミングの項目に低いスコアを付けた被験者は比較的途切れがちな話し方であり、それは特に 6.0 秒一定間隔条件で見られた。これはロボットに短い間隔で相づちを打たれることによって被験者の発言が途切れてしまうのではないかと考えられる。発言中に傾きを行う場合は傾くタイミングに関してあまり敏感ではなく、自然だと感じる割合が多いが、被験者が途切れがちな話し方の場合、ロボットが不自然な途切れの場所で頷いたりすることで傾きのタイミングが不適切であると感じる確率が高くなるのではないかと考えられる。ランダム条件でもロボットが短い間隔で相づちを行う可能性がある。そのため被験者が途切れがちな話し方になることがあり、タイミングの項目のスコアが 7.7 秒間隔と比較し低くなっている。

今回多くの被験者が、信憑性に高い評価を付ける場合に傾きが音声と同期していたことをあげる場合が多かったことから傾きと音声との同期も信憑性に影響する要因の一つであると考えられる。遠隔対話用ロボットの操作にフェイストラッキングを用いた場合には必ずしも音声と傾きが同期するとは限らないが、そのような音声と同期していない傾きは動きの信憑性を低下させる可能性がある。

予備実験において非接触なフェイストラッキング装置を用いて実験を行った場合、特に傾きのような縦方向の大きな動きではフェイストラッキング装置のデータのノイズにより動きがぎこちなくなってしまう。その際多くの被験者に、実際にフェイストラッキングのデータを用いているにもかかわらず動きのぎこちなさを理由に自動で動いていると判断されてしまった。それと比較し今回の 7.7 秒一定間隔では信憑性の平均値は 8.2 と高く、タイミングが十分自然であれば傾き動作があらかじめ作り込まれたモーションの再生である方がノイズの混じったフェイストラッキングのデータを用いるよりも高い信憑性を得られる可能性が高い。

話しやすさの項目に関しては有意な差は見られなかったが、スコアの理由に傾きのタイミングをあげる被験者が多かった。

そのため 6.0 秒一定間隔条件のスコアが最も低くなっている。これは上で述べたように短い間隔に相づちを打たれることで途切れがちな話し方になり、話し難さを感じたためであると考えられる。

同じ部屋の中で話しかけている間隔に関しては条件間で差はなかった。これは今回のロボットははいと言いながら頷くだけであり、被験者とロボットの間にインタラクションが殆どなかったためであると考えられる。ロボット側が常に受け手ではなく質問をするなどより多くのインタラクションを行うことで動きの信憑性との関係が出る可能性もある。

本実験では、音声録音であることに気づいた被験者はほとんどいなかった。そのため音声録音であることによる、結果への影響はなかったものと思われる。今回のタスクが被験者が一方的に話し、ロボットは常に受け手であるというものであったため、会話は 90~120 秒と短いものであった。実験時間をより長くすることで、今回高いスコアであった 7.7 秒一定間隔条件でも違和感が蓄積し、結果が変わる可能性がある。

6. おわりに

本論文では、遠隔対話用ロボットの傾きを自動化する際に、そのロボットの動きが遠隔地にいる人の動きを伝達していると感じさせるためには作り込まれた傾きモーション、傾きに同期して再生される録音音声で十分であり、信憑性に大きく影響するのは傾きの間隔であることを示した。

これを我々は、適切な一定の間隔で頷く条件と短い一定の間隔で頷く条件を比較することで示した。ロボットが短い間隔で相づちを行うと、被験者の発言が途切れがちになり、被験者が不自然なタイミングで頷いたと感じやすくなるためであると考えられる。

謝辞

実験に協力していただいた和田侑也氏に感謝する。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存在感の研究」、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

- [Breazeal 99] C. Breazeal and B. Scassellati: “A context-dependent attention system for a social robot”, Proceedings Int. Joint Conference on Artificial Intelligence, pp. 1146-1151, (1999).
- [Kidd 04] C. Kidd and C. Breazeal: “Effect of a Robot on User Perceptions.”, IEEE/RSJ International Conference on Intelligent Robots and Systems, (IROS '04), (2004).
- [Sakamoto 07] D.Sakamoto, T. Kanda, T. Ono, H. Ishiguro and N. Hagita, : “Android as a Telecommunication Medium with a Human-like Presence”, Proceedings of the ACM/IEEE international conference on Human-robot interaction, HRI '07, pp. 193–200, (2007).
- [Shinozawa 05] K. Shinozawa, F. Naya, J. Yamato, and K. Kogure: “Differences in Effect of Robot and Screen Agent Recommendations on Human Decision-Making”, International Journal of Human-Computer Studies, Vol 62, pp 267-279, (2005).
- [morikawa 98] O. Morikawa and T. Maesako, HyperMirror: “Toward Pleasant-to-use Video Mediated Communication System”, Proceedings Conference on Computer supported cooperative work, pp. 149-158, (1998).
- [Kraut 03] R. E. Kraut, S. R. Fussell, and J. Siegel: “Visual information as a conversational resource in collaborative physical tasks.”, Human-Computer Interaction., 18, 13-49. (2003).
- [Watanabe 99] 渡辺 富夫:身体的コミュニケーション解析のためのバーチャルコミュニケーションシステム 情報処理学会論文誌 40(2), 670-676, 1999-02-15.