

スマートフォン=ロボットとの継続的なインタラクションが人の行動に及ぼす影響

Does Continuous Interaction with a Smartphone=Robot Influence the Behavior of a User?

柴田 諒子*1
Ryoko Shibata

劉 よう*1
Liu Ye

深田 智*1
Chie Fukuda

尾関 基行*1
Motoyuki Ozeki

岡 夏樹*1
Natsuki Oka

*1 京都工芸繊維大学
Kyoto Institute of Technology

We studied the possibility of continuous human-robot interaction experimentally. Although robots are becoming increasingly common in social and home environments, the relationships between people and robots vary between individuals. If a user becomes attached to a robot, we can expect that he/she will be able to establish a long-term interaction with the robot. We developed a smartphone application which regards a smartphone as a robot to facilitate continuous human-robot interaction. We confirmed the adaptive change of the user behavior, according to the progress of learning by the robot.

1. はじめに

本研究の最終的な目標は、相互適応を通じてユーザにロボットを継続的に用いてもらうことである。

ロボットは現在、社会・家庭環境に徐々に浸透しつつあるが [1]、飽きなどの理由により長期的なインタラクションを持つことは未だに難しい [2]。しかしながら、ロボットに対してユーザが愛着を持つことで、ロボットとのより長期的なインタラクションを持つことが期待される。また、インタラクションが長期化することで、学習可能なロボットであればそのユーザに適應するための学習データを多く得ることも可能であると考えられる。

これまでに著者らが扱ってきた実験用ロボットを用いた印象評価の実験では、人数を確保した上で継続的に実験を行うことが困難であった。例えば実験室で実験を行うとき [3]、参加者への負担を考えると 1 回の実験時間はあまり長くなりすぎないようにし、継続的に行おうとしたとき、何度も同じ参加者を定期的に実験室に來させることもまた困難な負担となりうる。一方、一般のご家庭にロボットを貸し出すとき [4]、セッティングや教示、撤収などのために、ご家庭にあげてもらう必要が生じ、そこまでを許可して下さるご家庭を多く見つけることが困難であった。

これらの問題を解決するために、より多くのユーザにとって身近で気軽に用いることのできる端末で実験を行うために、本研究ではスマートフォンそのものをロボットとみなすアプリケーションを開発した。

現在主流であると考えられる前面がタッチパネルとなっているディスプレイを有するスマートフォンそのものを、聴覚（マイク入力）や触覚（タッチパネルへのタッピング）を備えた*1ロボットの身体とみなし、ディスプレイにロボットの顔と手を描画することでロボット側もまた身体表現を可能とし、インタラクションが可能なスマートフォン=ロボットを開発した。

また普及したハードウェアを用いることで、多人数からのインタラクションデータが取得できることが期待される。これ

らロボットの学習度合と人のロボットへの適應のデータを用いて、ユーザにロボットが適應するだけでなく、ユーザがロボットに適應することが理想的には可能となる。そのことによってロボットの学習が促進され、ロボットの更なるユーザへの適應、さらに将来的には、何を学習すべきかを学習するロボットの開発も期待される。

長井ら [5] が実装した、教示者とのインタラクションによって語彙を獲得するロボットでは、幼児の語彙学習における教示者と幼児のインタラクションダイナミクスに似た振る舞いを、ロボットへの教示者が見せたとされている。長井らの実験では、実験参加者にはロボットに単語を教えるように指示し、20個または 125 個の物体を用意した。これに対して本研究では、後述するように、ロボットに対する発話は、操作手順書で言葉を教えることに言及はするが、それだけには限定せず、より自由に行われた。

2. 実験

2.1 目的

本実験では、参加者自身が所有している、前面にタッチパネルディスプレイを搭載した Android OS のスマートフォンにロボットの顔と手を描画するアプリケーションをインストールし、端末そのものを、強化学習及び教示によって言葉を学習してゆくロボットとみなし、ロボット側の学習の進展に応じて人の行動が適応的に変化するか、また、それがロボットの学習を助けるかについて調べた。

2.2 実験方法

実験参加者 男性 2 人（参加者 A 24 歳、参加者 B 38 歳）と女性 1 人（参加者 C 48 歳）であった。

装置と刺激 実験には、実験参加者自身の持つ所有する Android を搭載したスマートフォン端末を用いてもらった。各自の端末に今回開発したロボットの顔と手を描画するアプリケーションをインストールさせた。実験に用いたアプリケーションによって描画したロボットを図 1 に示す。このロボットに対して、ユーザはマイク入力による音声対話と、タッチパネルへのタッピングやドラッグの操作によるロボットへの指さしやタッピング操作が可能となっている。また、ロボット側からユーザへの働きかけとして、ロボット自身の手を動かして指をさす、ユーザに指さされた（タップされた）パーツを動かすことで

連絡先: 連絡先: 京都工芸繊維大学, 〒 606-8585 京都府京都市左京区松ヶ崎橋上町, E-mail: liu@ii.is.kit.ac.jp

*1 脚注: ハードウェアとしては視覚（カメラ入力）や平衡感覚（加速度センサ等）等も備えているが、今回のアプリケーションでは利用していない。

こであったと認識されたかを通知する，ダイアログに文字列を表示することによって発話することが可能となっている．ロボットの持つ機能の詳細を以下に示す．

色々なところを，ダブルタップして，言葉を教えてあげてください

色々なところを，シングルタップすると，覚えた言葉をお話します

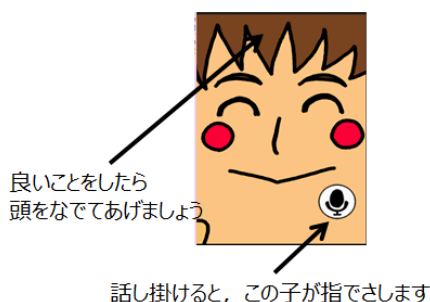


図 1: 実験参加者へ配布したアプリケーションの操作手順

● 顔のパーツをダブルタップされた時

顔のパーツは、「髪の毛」「右の眉毛」「左の眉毛」「右の目」「左の目」「右の頬」「左の頬」「鼻」「口」を離散的に定めており，それ以外の場所をさされた時は「どこもさされていない」とした．これらのパーツについてダブルタップされると，android.speech パッケージの RecognizerIntent クラスの音声認識が立ち上がる．このときのユーザの発話の音声認識結果の文字列をそのパーツの名前であると教示されたとして，ロボットは対応付ける．

● 顔のパーツをシングルタップされた時

ダブルタップのときと同様，顔のパーツは「髪の毛」「右の眉毛」「左の眉毛」「右の目」「左の目」「右の頬」「左の頬」「鼻」「口」を離散的に定めており，それ以外の場所をさされた時は「どこもさされていない」とした．ダブルタップの際，および後述のマイクのボタンをタップした際に入力された文字列のうち，タップ対象と結びつけて記憶されている文字列を，ダイアログにてロボットの発話として出力する．

また，ユーザがロボットの発話の後に「髪の毛」部分を指でなでる（ドラッグ）すると，ロボットに対して報酬を与えることが出来，ユーザのタップ位置とロボットの発話内容を対応付けて学習する．報酬はユーザの任意のタイミングで与えることが出来る．

● マイクのボタンをシングルタップされた時

マイクのマークのボタンをシングルタップすると，音声認識が立ち上がる．ユーザの発話の音声認識結果の文字列を入力とし，ロボットは入力と関連づけて記憶されている顔のパーツ「髪の毛」「右の眉毛」「左の眉毛」「右の目」「左の目」「右の頬」「左の頬」「鼻」「口」のいずれかを自身の指でさす．

ユーザはロボットの指さし行動の後に「髪の毛」部分を指でドラッグすると，ロボットに対して報酬を与えることが出来，ユーザの発話内容とロボットの指さし位置を対応付けて学習する．報酬はユーザの任意のタイミングで与えることが出来る．

本アプリケーションでは上記 3 種類のインタラクションが可能となっており，各インタラクションで学習された結果は他のインタラクションでも共有される．学習結果は強化学習の行動価値テーブル（図 2）の形で共有されるが，共有しない設計もありえる．チンパンジーは対称性推論が苦手であり [8]，色を色調として見て，その色名を漢字で選ぶという操作を学習しても，逆に漢字から色調を選択することが出来ないとされており，チンパンジーを模したものであれば，上述の強化学習結果は操作ごとに分けるべきであると考えられる．本研究の学習モデルは人を模しており，人の特徴は対称性推論をすることである [9] と言われているため，上述のように設計した．

	髪	左眉	右眉	左目	右目	鼻	左頬	右頬	口
花	0	0	-0.1	0	0	0.1	0	0	0
目だよ	0	0	0	0	0.19	0	0	0	0
髪の毛	0.1	0	0	0	0	0	0	0	0
...									

図 2: 強化学習の行動価値テーブル

手続き 実験参加者に対しては，図 1 と同様の内容のものを PDF ファイルでアプリケーション操作手順として配布して教示した．

この操作手順書では最低限の操作方法のみを教示しており，触れてみて最初は何をしてよいか分からないであろう状態から実験を始めた．ロボットの振る舞いによって人がロボットに適應してゆくか，また，適應した時ロボットの学習が促進するかについて実証的に調べる．

また，実験終了後に，著者らの研究 [4] にてロボットに対しての印象評価に用いたものと同じ質問をした．更に自由記述で，具体的にどのように扱ったか，どんな存在だと思ったか，今後どうなるとおもしろい，嬉しいか，その他思ったこと，などを回答させた．

2.3 結果と考察

図 3 に各実験参加者のアプリケーションの操作時間を示す．操作時間については，アプリケーションが停止しても参加者自身がすぐに再起動したときは参加者の操作時間として連続していると考えられるので，これらの試行は同じ行動として扱った．逆に起動後すぐに停止してしまってから再起動を試みなかった場合については，インタラクション継続時間が 0 であると扱っている．

参加者 A は 1 回あたりのインタラクション時間が長かった．参加者 C はインタラクション回数が多かった．参加者 B はあまりインタラクションをしていなかった．

このときの，各実験参加者のアプリケーションへの印象のプロフィール曲線を図 4 に示す．

全ての参加者が，可愛いについては高く評定した．全体的に高めに評価した参加者 B に比べて，参加者 A，C は低めに評価している．考えられる原因の一つとして，参加者 B は，参加者 A と C に比べインタラクションが短時間であった．改善の余地がある点として，3 者ともに，ロボット側の動作や表情の乏しさを挙げ，特に今回はロボットから人への自発的な働きかけがなかった点もまた指摘された．また，音声認識精度の低さについても不満点として挙げられていることから，継続的にインタラクションを持つことによって，ユーザのロボットの誤動作・誤認識の経験が増加し，飽きなどの理由により印象が下がる可能性が考えられる．

ロボットの頭をなでた時の，表情の変化が良かった，とも回答されたことから，より表情を豊かにすることはロボットへの

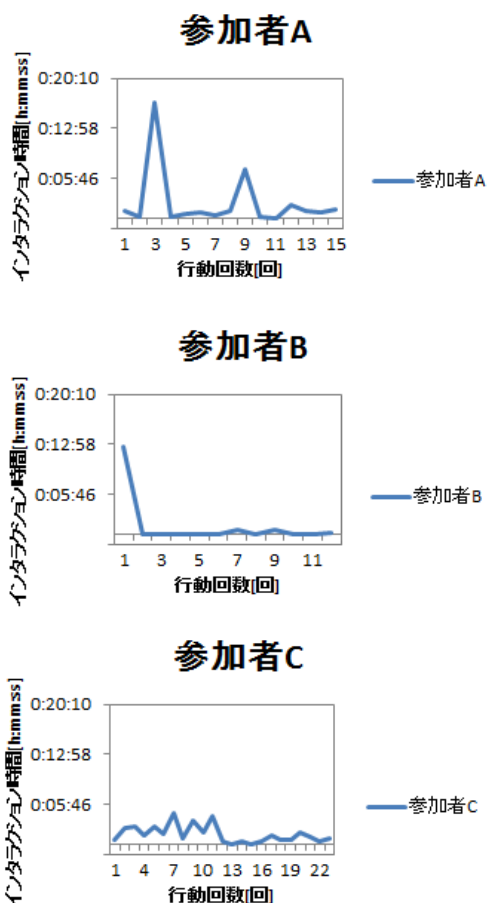


図 3: 各実験参加者のアプリケーションの操作時間

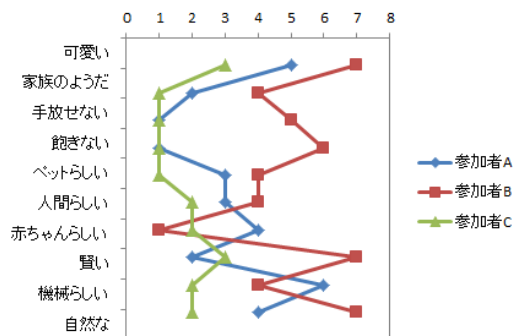


図 4: 各実験参加者のアプリケーションへの印象

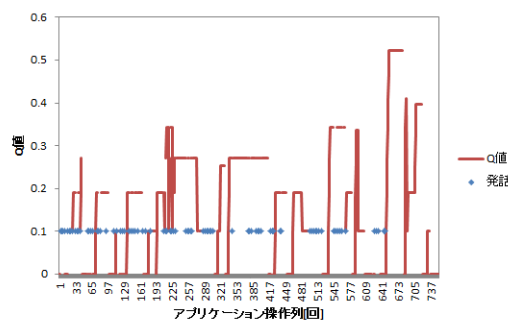


図 5: Q 値の変化と発話タイミング

印象の変化を向上させようと考えられる。

ここで、実験参加者がロボットに対して、どのように振る舞ったかについて、ロボット側の Q 値の変化と参加者の発話入力タイミングを図 5 に示す。ここで Q 値とは行動価値であり、報酬を得られる (=頭をなでられる) 行動が強化され増加し、その行動が選ばれやすくなる値であり値が大きいくほどその行動が選ばれやすくなる。

実線が Q 値の変化の様子を示してあり、点が参加者がロボットに対して発話入力したタイミングである。Q 値は、ダブルタップによってタップされた顔パーツの位置と音声入力された文字列を対応付けした際と、ロボットの行動や発言に対して頭をなでるといった報酬を与えた際に大きくなる。

この図 5 では、Q 値などのログが正常に残っていた実験参加者 B のインタラクションの様子について Q 値と発話タイミングのデータを用いている。

このグラフを参照すると、ユーザの発話 (青い点) が多く生じていると考えられるアプリケーション操作列 (横軸) の 200 回程度くらいまでは Q 値の上昇は小さいが、発話数が減少してきた 300 回程度くらいから、ロボットのパーツをタップし頭をなで報酬を与えるような操作が増えたのではないかと推測される。言葉を教える、という場を想定すると、言葉を覚える対象に量をできるだけたくさん話しかけていくフェーズから、対象が発話する内容に耳を傾け、正しいことを言えば褒めてあげるようなフェーズに移行しているような状況ではないかと推測される。

Q 値の更新が終盤も続き、さらに大きくなっていったことから、参加者側が、褒めると喜びその単語を覚える、というロボットの動作を理解した結果として話しかけることと報酬を与えることとを区別するようになった、つまり適応したと考えられる。

3. まとめと今後の展望

スマートフォン=ロボットとの継続的なインタラクションによって、人はロボットに適応し、ロボットが学習しやすいように行動が変化する可能性が示唆された。特に、ロボットがある程度単語の数を覚えてきたら褒めてあげることで報酬を与え、より学習を促進させるように、段階を追って人の振る舞いが変化した可能性があると考えられる。

現時点では 3 名に対する予備実験しか行っていない上に、アプリケーションが停止してしまう不具合があり、さらに、ログをとるプログラムの不備により十分なログがとれず、インタラクションの十分な解析ができなかった。まずは、停止の可能性を減らし、ログが完全にとれるように修正したアプリケーションを用いてインタラクション実験をやり直し、ログを詳細に分析して、どのようなインタラクションが起きているかを明らかにする予定である。短期の実験室実験や、少人数のロボット貸し出し実験では得られなかった、興味深いデータが入手できることを期待している。その上で、スマートフォン=ロボットの仕様を検討し、本格的な継続的インタラクション実験を行いたい。現時点で判明している課題は以下の通りである。

今回は、ロボットは共起情報と報酬に基づき言葉を覚えるが、実験協力者への教示は意図的に不十分にして、何をすればよいか分からない状態から、ロボットに言葉を教えてもらった。結果として、ロボットの振る舞いに適応して、人の教え方

が変わった可能性が示唆された。長井ら [5] の研究と異なり、ロボットの単語の学習が進みロボットの知っている語数が増えると、実験参加者の発話は少し減少し、ロボットの発話を確認して褒めることで報酬を与えることが多くなったと考えられる。しかしながら、ロボット側からの働きかけが乏しく、ユーザは飽きてしまい、ロボットへの印象は良くならなかった。

各実験参加者が挙げた、ロボット側からの働きかけは今後必ず実装すべき点であり、さらに、この働きかけによって、ロボット自身の学習度合をユーザ側に知らせることが出来れば、ユーザはそれにも気づくことが出来、更に行動が適応変化する可能性があると考えられる。また、ロボットの振る舞いが増えることから、飽きも生じにくくなり、ロボット自身への印象も低下しにくくなるのではないかと考えられる。

今後の展望として、学習するだけでなく、行動予測についてもロボットに実装したいと考えている。ロボット側が行動予測し、ロボットから自発的な行動が生じることが期待される。つまり、ロボット側の動作について、人からロボットへの働きかけだけでなく、ロボットから人への働きかけについても実装することで、人が何をすればよいかわからない時間を減らし、より人からロボットへの働きかけが増加することが期待される。

現在はユーザがロボットとのインタラクションに際して得られるものが達成感しかなく、継続的なインタラクションを実現する動力としては弱くなっている。ユーザにとってのモチベーションに繋がる、ロボットからユーザへ提示可能な報酬についても、今後検討したいと考えている。

また、現在は、ユーザからロボットへの働きかけとしてはタッチセンサとマイクによる音声入力、ロボットからユーザへの働きかけとしてダイアログによるロボットの発言及びロボットの笑顔しか用いておらず、スマートフォンを用いて追加できる機能は、加速度センサによる「高い高い」遊びや、マルチタッチによる顔の変形、音声合成や録音音声の再生などが存在する。これらの機能に関しても有効に活用し、ロボットの入出力を拡充させることでさらなるインタラクションの広がりが期待される。

謝辞

本研究は科研費 (25330260) の助成を受けたものである。

参考文献

- [1] R.A. Brooks, C.B. (Ferrell), R. Irie, C.C. Kemp, M. Marjanovic, B. Scasselati, and M.M. Williamson, "Alternative essences of intelligence," *Proceedings of the American Association of artificial Intelligence*, pp.961-968, 1998.
- [2] 神田崇行, 平野貴幸, ダニエルイートン, 石黒 浩, "日常生活の場で長期相互作用する人間型対話ロボット," *日本ロボット学会誌*, vol.22, no.5, pp.636-647, 2004.
- [3] R. Shibata, T. Kojima, C. Fukada, K. Sato, Y. Hashikura, M. Ozeki, and N. Oka, "Does talking to a robot in a high-pitched voice create a good impression of the robot?," in *Proceedings 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing(SNPD 2012)*, pp.19-24, Kyoto, Japan, 2012.
- [4] 柴田諒子, 小島隆次, 尾関基行, 岡 夏樹, "継続的に対乳児発話でロボットに語りかけることによる印象への影響," 第 27 回人工知能学会全国大会, no.1M3-OS-18a-1, 2 pages, 2013.
- [5] 長井隆行, 中村友昭, "ロボットの語彙学習におけるインタラクションのダイナミクス," *HAI シンポジウム 2013*, pp.237-243, 2013.
- [6] 板垣祐作, 小川浩平, 小野哲雄, "Itaco on the room: アンビエントな情報提供を行う生物感のあるエージェントの提案," *HAI シンポジウム 2008*, no.1B-4, 6 pages, 2008.
- [7] NTT docomo, シャベってコンシェル — サービス・機能 — NTT ドコモ, (オンライン), 入手先 https://www.nttdocomo.co.jp/service/information/shabette_concier/ (参照 2014-2-8).
- [8] 友永雅己, "チンパンジーにおける対称性の(不)成立," *認知科学 = Cognitive studies : bulletin of the Japanese Cognitive Science Society*, vol.15, pp.347-357, 2008.
- [9] 服部雅史・山崎由美子(編), "誌上討論 対称性:思考・言語・コミュニケーションの基盤を求めて," *認知科学*, vol.16, pp.120-184, 2009.