

人—計算機インタラクションでの視覚的手がかりの提示による 不自然さへの気づき

池ヶ谷 啓伍, 清水 悠生, 長原 瑛吾, 金野 武司

Keigo Ikegaya, Yusei Shimizu, Eigo Nagahara, Takeshi Konno

金沢工業大学 工学部 電気電子工学科

Kanazawa Institute of Technology, College of Engineering,

Electrical and Electronic Engineering

{b1914269, b1833115, b1813289}@planet.kanazawa-it.ac.jp, konno-tks@neptune.kanazawa-it.ac.jp

概要

人と機械のインタラクションで実現できないことの1つに、ターンテイキングと呼ばれる対話やジェスチャーにおける主従関係の自然な入れ替わりがある。視覚的なインタラクション環境で実施された先行研究における人と計算機での実験では、構築された計算モデルが相手に応じて主従関係を切り替えるようになっていなかったにも関わらず、その計算モデルに相対した人間は、相手が人であるか計算機であるかを識別できなかったことが報告されている。我々は、主従の状態をより意識できるような手掛かりを示せば、動作アルゴリズム自体には何ら変更がなくとも、その計算モデルを識別できるようになるのではないかと予想した。結果、我々の実験では人間と計算機を識別する正答率がチャンスレベルを有意に上回ることが確認された。ただし、その正答率と主従の切り替えに対する意識の変化を調べたアンケートや、インタラクションパートナーとの対応関係を見出すことはできなかった。

キーワード: Turn-taking, Nonverbal communication, Human agent interaction

1. はじめに

近年、一般家庭に進出するコミュニケーションを主目的としたロボットや、モニターに映し出される AI エージェントは、発話だけでなく身振り手振りを交えて人とインタラクションする。しかし、そのようなロボットや AI エージェントは急に動き出したり、不意に話し出したりするなど、人どうしのインタラクションに比べて自然にコミュニケーションできないことが多い。これは、人間と AI の間で自然なターンテイキング [1] が成立しないことが原因であると考えられる。

飯塚 [2] は触覚を通じた人どうし、あるいは人と計算機によるインタラクション実験の中で、ターンテイキングは相手が人であることを判断する重要な手がかりであることを示した。この実験を元にして、入江ら

[3] は視覚に限定したインタラクション環境を構築し、人どうしと同様の主従関係（以下、リーダー／フォロワー）の入れ替わりを再現する計算モデルを用いて人と計算機でのインタラクション実験を行なった。結果、その計算モデルとインタラクションした人間は、相手が人間であるか計算機であるかの判別ができなくなるが示された。しかし、そこで用いられた計算モデルは、リーダー／フォロワーの入れ替わるタイミングがあらかじめ設定された確率分布（パレート分布）に従って決められるようになっており、相手の動きに応じてリーダー／フォロワーを切り替えるようになっていない点で、人が行なう自然なインタラクションを再現するものではなかった。

入江らの実験において人と計算機の判別ができなかった原因は、実験参加者が用意された環境においてリーダーやフォロワーといった役割に注意を向けることがなかったからではないかと我々は考えた。そこで本研究では、その役割に意識を向けるような視覚的手掛かりを加えたとき、インタラクションする相手が計算機であることに気付くようになるか検証することを目的とした。

2. 実験方法

2.1 実験室実験

実験では、二人の参加者が向かい合うように設置されたコンピュータ端末の前に座り、目の前に設置された衝立により互いの姿は見えないようになっていた。両者の端末の画面上には中央に直径 18.7[mm] の円形オブジェクトが横並びに 2 つ配置された（図 1）。それぞれの丸は一方を参加者 A が操作し、他方を参加者 B が操作した。お互いに操作する丸はオレンジ色で表され、相手側の丸は緑色で表示された。丸はテンキーの 2 つのボタンによって直線上を左右に動かすことができ、自身が動かす丸の位置の変化は相手の画面に映る丸（緑色）に直ちに反映され、丸は互いにすり抜ける

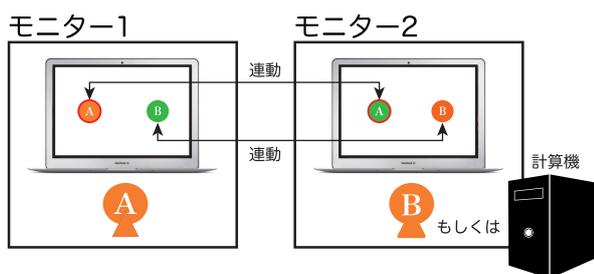


図1 実験環境

ことができなかった。この環境において、本研究では参加者がリーダー／フォロワーの関心に注意を向けるための2種類の手がかりを用意し、それぞれを実験1、および2として参加者間計画の実験を実施した。

いずれの実験でも、視覚的手がかりには1つの円形オブジェクトの上に描く赤色のリングを用いた(図1)。1つ目の実験では、そのリングは互いの円の接触によって遷移させた(実験1:接触遷移)。この方法では明示的な接触というイベントによってリングの受け渡しが行なわれるため、何らかの役割に意識的になる可能性が高いと予想した。また、2つ目の実験では、二人が動かす円の軌道を10秒間の時間窓で逐次分析し、動きが先行する側にリングを表示させた(実験2:リーダー表示)。先行する動きの判別には入江ら[3]の研究と同様に、2つの軌道の相互相関係数を計算する方法¹を用いた。実験2ではリングが遷移する仕組みが明示されないため、参加者がリングから何らかの役割を意識することは困難であると予想した。いずれの実験でも、参加者にはリングの表出条件は提示しなかった。

2.2 実験手続き

実験参加者は、コンピュータ端末越しに簡単なゲームに取り組むことが説明され、最初に操作に慣れるための練習を行なった。参加者は90秒を1セッションとし計15セッションに取り組んだ。前半の5セッションは自由に動かすように指示され、人どうしで行った(人-人課題)。後半の10セッションでは相手が人と計算機で5回ずつランダムに入れ替わった(人計算機混在課題)。人計算機混在課題で参加者は、1セッションが終了するごとに、相手が人であったか計

¹2つの軌道に対して、一方の軌道を時間的に遅らせた場合に相関係数が上がるならば、遅らせた側の動きが先行したリーダー側であると判断し、逆に進めた場合に相関係数が上がるならばフォロワーであると判断する方法。同調度合いに閾値を設けるようなことはせず、必ずリーダーもしくはフォロワーのどちらかの状態を判定した。

算機であったかについてのアンケートに答えた。その回答の正誤についてのフィードバックはしなかった。前半の5セッションが終わった後で、参加者には後半の10セッションは相手側の操作が人もしくは計算機になることが伝えられた。そこでは、相手が人であるか計算機であるかを判定するとともに、自身が人であることを相手に伝えようとするように指示された。

このセッション中、参加者は相手側のキーの操作音が聞こえないように、ノイズキャンセリングヘッドホン(Sony製WH-1000XM2)を装着した。セッション終了後、参加者は相手の動かし方やリングの表出条件についての簡単なアンケートに答えた。

2.3 実験設備

モニターはiiyama製ProLiteXUB2790HSを用い、画面サイズは597.9×336.3[mm](1920×1080[pixel])だった。実験システムはMathworks社のMATLABにPsychtoolbox[4]を組み合わせて構築した。丸の直径は18.7[mm]とし、リーダー状態を表示する赤色のリングの内径を丸の直径と等しく18.7[mm]に、そして外形を21.8[mm]とした。描画更新のサンプリング時間は33[Hz](30[ms])とし、1サンプリング時間あたりの丸の移動距離を10[pixel]とした。このため、人が操作する丸の移動は79.7[mm/sec]の等速度運動であった。

2.4 実験参加者

それぞれの実験には金沢工業大学の男子学生10ペア20名ずつ(実験1:平均=21.7, SD=0.61, 実験2:平均年齢=21.6, SD=0.66)が参加した。また実験はすべて同大学の研究室で行なわれた。

3. 結果

リングの表出という手がかりの提示によって、参加者は人または計算機を判別することができたのだろうか。これを確認するために人計算機混在課題の10セッションで行なった、人であるか計算機であるかを答えたアンケートの正答率を、リングありとリングなしの状態と比較した(図2。リングなしの結果は入江ら[3]の実験データを使用した)。

いずれの実験においても、後半10セッションにおいて相手が人であるか計算機であるかの正答率はチャンスレベル(1/2)を有意に超えた。先行研究では人を相手にした場合と計算機を相手にした場合の正答率

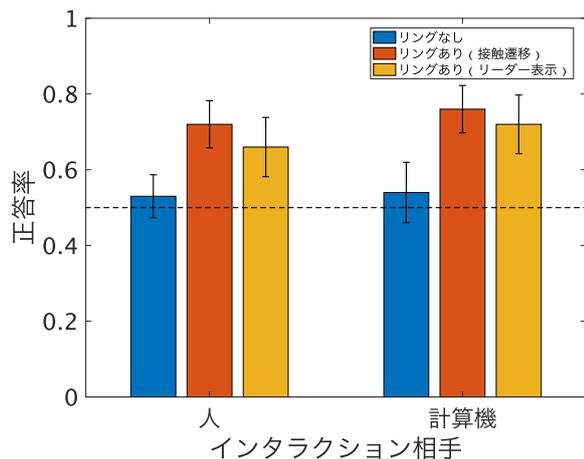


図2 人計算機混在課題（後半 10 セッション）での正答率

の平均がそれぞれ 0.53 と 0.54 であったのに対して、接触を条件とした実験 1 では 0.72 と 0.76, さらに実験 2 では 0.66 と 0.72 であった。チャンスレベルとペア相手（人／計算機）について行なった二要因分散分析では、いずれの実験も交互作用は有意ではなく、リングを表示した場合の有意差が見られた（実験 1: $F(1,76)=29.58$, $MSE=0.039$, $p < .001$, $\eta^2=0.279$, 実験 2: $F(1,76)=11.93$, $MSE=0.061$, $p < .001$, $\eta^2=0.135$ ）。

それでは、リングのあり／なしによって参加者の丸の動かし方に違いは生じていたのだろうか。これを調べるため、入江らの先行研究に倣い、2つの丸の同調度（図3）とリーダー比均等率（図4）を計算した。同調度は2つの丸の軌道の相関係数が計算される。またリーダー比均等率は先述した相互相関係数を算出する方法により、1セッション内でリーダーをしていたと判定された二者の時間が均等（1/2）であれば1に、どちらか一方が全てリーダーをしていたら0と計算される。

同調度について、リングが表示されない状態では、計算機を相手にした場合の方が高い傾向があり、リーダー比均等率においては人を相手にした場合の方が高い傾向があった。これは、人を相手にした場合の方が相手との動きの同調度が下がり、リーダーをより均等に行っていたことを示唆している。これに対して、リングを表示させた場合の重要な結果は、そのような違いがなくなったということであり、それにも関わらず正答率がチャンスレベルよりも有意に高かったということではないかと思われる。この結果は、リングの表示によって上がった正答率を説明するのは、全体的な動作特徴ではない可能性を示唆する。

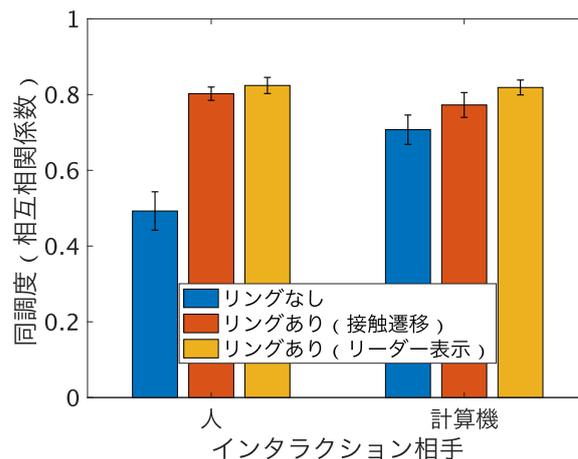


図3 人計算機混在課題での同調度の比較

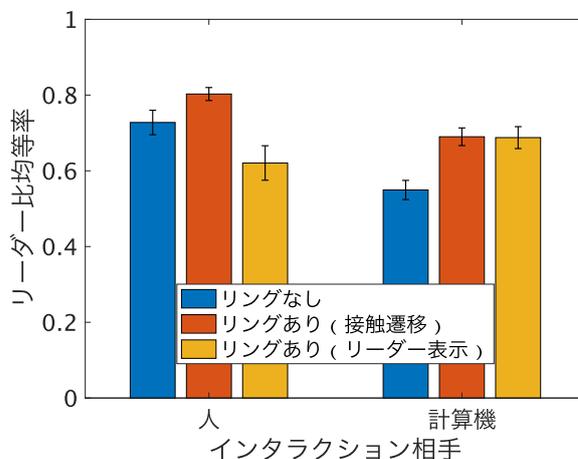


図4 人計算機混在課題でのリーダー比均等率の比較

加えて、同調度と正答率の間の相関関係を調べると、人を相手にした場合には有意な相関関係が見られる ($r = .266$, $p = .040$)。しかし、個別の条件ごとの相関分析には有意な関係はなく、同調度が正答率の違いを説明する要因とは考えにくい。これはリーダー比均等率についても同様で、条件全体での分析および個別の分析においても有意な相関関係は見られなかった。これらの結果は、リングが表示された事による人および計算機の識別率の向上が、同調度やリーダー比均等率といった全体的な動作特徴によって説明できるものではないことを重ねて示していると思われる。

4. 議論

接触によってリングが遷移した実験 1 において、有意に識別率が上がったことは我々の予想どおりであったが、リングが遷移する仕組みが明示されなかった実験 2 においても識別率が実験 1 とほぼ同程度になった

ことは興味深い結果であると我々は考えている。実験1において、その判断理由を自由記述で尋ねた実験後のアンケートでは、相手と鬼ごっこをしたことや、リングを渡すふりをしたことなど、通時的な動作とは異なるレベルに出現する逐次的な動作パターンを判断理由としたことが伺える記述が見られた。他方、実験2において、リングの表出を人または計算機の判断に利用したかを「はい/いいえ」で尋ねた質問で、「はい」と回答したのは2名だけだった。また、その2名の判断理由についての自由記述も、背後に用意された仕組みを正しく答えたものではなかった。この結果は、リングの表出によって相手が人であるか計算機であるかの識別率が上がったにも関わらず、その判断は無意識的に行なわれたことを示唆している。

このような結果となった原因を特定するため、我々は丸の移動軌道の分析を行なったが、現時点では有効な特徴の抽出には至っていない。アンケートの回答からはその認識のプロセスが無意識下において行なわれたらうことが示唆されるため、二人の丸の動かし方からそれを推定できるような分析を行う必要があると考えられる。また、リングを表出させる規則が本当に識別率に貢献したかどうか併せて検証する必要がある。この検証には、リングの遷移になんの規則性も持たせないようにするなどの条件を新たに設定し、実験を実施することが考えられる。

5. 結論

我々は、視覚的な二者間のインタラクションにおいて、二者の動きの特徴を明示的（接触遷移）/非明示的（リーダー表示）に反映させたリングの提示により、自身の役割を意識することで人/計算機の区別が付きやすくなると予想した。特に我々が注目したのはターンテイキングにおいてリーダーとフォロワーを相手に応じて切り替える仕組みがないことによる違和感に気づくかどうかであった。

結果として、手掛かりの提示により正答率は向上したが、非明示的なリングの提示条件ではその規則は理解されず、手掛かりの提示とその違和感を露出させることの因果関係は確認できなかった。また、我々の実験では人を相手にした場合と計算機を相手にした場合の動作特徴に、入江ら [3] の実験で見られたような差が生じなくなることが確認された。この結果は、リングがあることによって与えられる違和感が、全体的な行動特徴によってもたらされるのではなく、より短期的な動きの特徴によって生じる可能性を示唆している。

文献

- [1] Levinson, S. (2016). Turn-taking in Human Communication—Origins and Implications for Language Processing, *Trends in cognitive sciences*, 20(1), pp.6–14.
- [2] 飯塚 博幸 (2014). 複雑系アプローチによる他者の認識とコミュニケーションの創発, *計測と制御*, 53(9), pp.821–827.
- [3] Irie, R. and Konno, T. (2019). Effects of turn-taking dynamics without contingency: A visual interaction experiment. In *proceedings of the 7th International Conference on Human-Agent Interaction (HAI2019)*, pp.239–240.
- [4] Kleiner, M., Brainard, D., Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36 ECVF abstract supplement.