

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 17 日現在

機関番号：32641

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24700190

研究課題名(和文) ヒューマンロボットインタラクションのための注意誘導のモデル化

研究課題名(英文) Modeling of Attention Drawing for Human-Robot Interaction

研究代表者

田村 雄介 (Tamura, Yusuke)

中央大学・理工学部・助教

研究者番号：40515798

交付決定額(研究期間全体)：(直接経費) 3,500,000円、(間接経費) 1,050,000円

研究成果の概要(和文)：ロボットが人間と円滑なインタラクションを行うためには、人間の注意と行動を誘導することのできる能力が望まれる。これに対して本研究では、手品の解析に基づいて人間の注意を誘導するための注意モデルを開発した。

手品鑑賞者の注視点と手品師が注意を誘導しようと意図していた領域の関係を調べた結果、手品師の手と顔の位置、および視線の方向の関係性が注意誘導に重要な役割を果たしていることがわかった。

この結果に基づき、これらの関係性を表現する誘導マップと、ボトムアップの視覚刺激を表現する顕著性マップを統合した注意マップを提案した。実験の結果、提案モデルは人間の視覚的注意をよく説明できることが示された。

研究成果の概要(英文)：In order for robots to smoothly interact with humans, it is desirable that robots can manipulate human attention and behaviors. In this study, we developed a model of human visual attention for the manipulation of human attention based on an analysis of magic. We measured human gaze points while watching magic videos and compared them with the area where the magician intended to draw spectators' attention. The analysis showed that the relationship among the positions of the magician's face and hands and his gaze direction plays an important role in manipulating spectators' attention. On the basis of the analysis, we integrated a saliency map and a manipulation map that describes the relationship to develop a novel human attention model. The evaluation using the observed gaze points demonstrated that the proposed model can better explain human visual attention than the original saliency map.

研究分野：総合領域

科研費の分科・細目：情報学 知覚情報処理・知能ロボティクス

キーワード：注意誘導 ヒューマンロボットインタラクション 手品

1. 研究開始当初の背景

人間と共存するロボットには、人間と円滑なインタラクションをすることが求められる。円滑なインタラクションを行うためには、相手の意図を推定することが重要であり、これまでにロボットによる人間の意図推定に関する研究は数多く行われてきた。しかしながら、一般に意図推定は相互に行われるものであり、人間の行動を受動的に観察することによる意図推定には限界がある。これに対して、我々人間は相手に積極的に働きかけることによって相手の行動を誘導するということをしばしば行っており、これにより円滑なインタラクションを実現している。

相手の行動を誘導するためには、相手の内部モデルを推定し、自らの行動が相手の認知や動作にどのような影響を与えるかを予測することが不可欠である。また、他者の行動を誘導するためには、その前に注意を誘導することが必要となる。

2. 研究の目的

本研究では、

- (1) 人間が行っている巧みな注意誘導の方法を計測に基づいて理解すること
- (2) 他者の注意誘導をロボットに実現させるためのモデル化を行うこと

を目的とする。

3. 研究の方法

(1) 注意誘導の解析

本研究では、手品という、人間の行う注意誘導の中でも非常に極端な例について解析を行う。これは、我々人間の日常生活における些細かつ無意識的な行動を解析するのではなく、手品師が意識的かつ巧みに行っている行動を解析の対象とすることで、注意誘導の本質に確実に迫ることができると考えたためである。

手品師の振る舞いが鑑賞者の注意にどのような影響を与えるのかを調べるため、手品の技法の一つである「フェイクトランスファー」を題材とし、この動画を鑑賞している鑑賞者の注視点を計測することとした。フェイクトランスファーとは、手から手へ物品を渡したように見せる技法であり、通常は手品の一連の流れの中で用いられる。

実験においては、手品の演技における文脈の影響を排除するため、フェイクトランスファーの部分のみの動画を用いるものとした。実験参加者には、(a) 通常のフェイクトランスファー、(b) 正面を向いた状態でのフェイクトランスファー、(c) 斜め上方を見ながらのフェイクトランスファー、の3条件の動画(図1)を鑑賞してもらい、その時の注視点を非接触型の頭部眼球運動追跡システム(faceLAB4.2, Seeing Machines)によって60Hzで計測した。

条件(a)において、手品師は鑑賞者の注意をどこに向けたいか、という意図をもって振る



図1 手品動画鑑賞実験

舞っている。ここでは、動画を9つのシーンに分け、各シーンで手品師が意図していた領域(意図領域)と鑑賞者の注視点との関係を解析することとした。

(2) 人間の注意のモデル化

(1)での注意誘導の解析結果に基づいて、ロボットに他者注意の誘導を行わせるためのモデル化を行った。

本研究では、人間の視覚的注意には、入力される画像における顕著性のようなボトムアップの要素だけではなく、他者の顔や手の位置、視線の方向が影響するという考えに基づき、これらを統合した視覚的注意の計算モデルを提案した。

提案モデルが、どの程度人間の注意を説明できるかを検証するため、(1)で用いたフェイクトランスファーの動画を鑑賞している時の注視点と提案モデルによる出力の比較を行った。

4. 研究成果

(1) 注意誘導の解析

まず、実験で用いた手品動画を、手品師に対するインタビューに基づいて、以下の9シーンに分けた(括弧内に手品師の意図領域を示す)。

- シーン1: ピンク色のボールを持った右手を上げる。(顔)
- シーン2: ボールを観客に見せる。(右手)
- シーン3: 左手を上げる。通常条件の場合、左手の方を見る。斜め上条件の場合、ここで斜め上方に視線を移す。(顔, 左手)
- シーン4: ボールを右手から左手に渡す動作をする(実際には、右手に残す)。(右手, 左手)
- シーン5: 右手を下ろすと同時に左手を上げる。(左手)
- シーン6: 左手の指をこすり、ボールを消す動作をする。(左手)
- シーン7: 左手の向きを変える。通常条件の場合、視線を左手から正面に戻す。(顔, 左手)

- シーン8：左手を開く。(左手)
 - シーン9：左手を下ろす。(特に無し)
- 手品師に対するインタビューから、本実験における意図領域は、手品師の顔、右手、左手であった。そこで、まず実験に用いた3条件の動画の各フレームにおける意図領域の

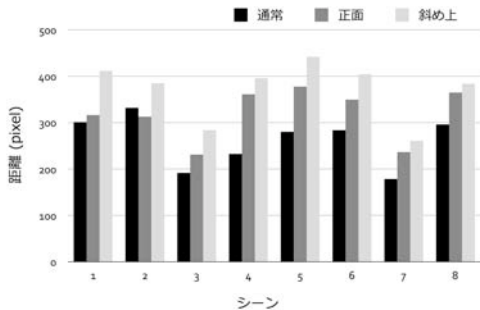


図2 意図領域重心と注視点との距離

位置を検出した。具体的には、HSV色空間において閾値処理を行うことで、動画中の肌色領域を抽出し、得られた領域の相対的な位置関係から顔、右手、左手を判別した。また、各領域の重心を顔、右手、左手の位置とした。これらの処理は、Windows PC上でOpenCV 2.4を用いて行った。

次に、手品師が観客の注意をどの程度コントロールできているのかを評価するため、検出した意図領域の重心と実験参加者の注視点位置との距離を測定した。図2にシーン1から8における各条件での距離を示す。

手品師の動作が条件間で異なるシーン3以降に注目してみると、通常条件の場合の意図領域重心と注視点との距離が他の2条件に比べて短くなっていることがわかる。

顔、右手、左手のうち最も実験参加者の注視点が引きつけられていた領域の時間変化を図3, 4, 5に示す。

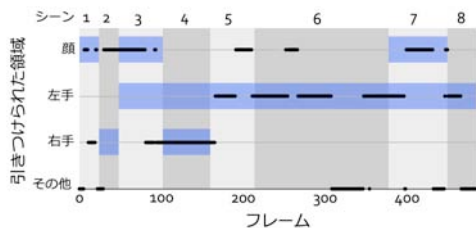


図3 引きつけられた領域の変化 (通常)

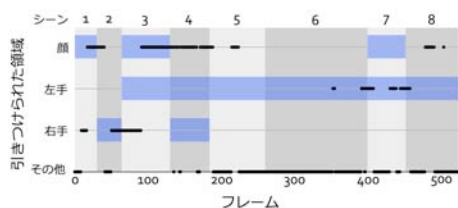


図4 引きつけられた領域の変化 (正面)

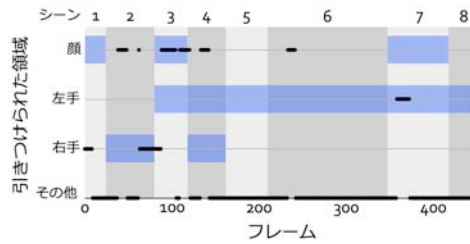


図5 引きつけられた領域の変化 (斜め上)

図3, 4, 5において、青い帯は手品師の意図領域を表し、黒いプロットが、実験参加者が平均的に引きつけられた領域を表す。つまり、手品師が完璧に観客の注意をコントロールできた場合、プロットは全て青い領域の中に含まれることになる。

図3からわかるように、通常条件の場合、手品師は観客の注意をうまくコントロールできている。

一方、図4からは、手品師が正面を向いている条件では、シーン4において、本来注意を向けてほしい手の領域ではなく、顔に注意が向いているということがわかる。また、その後のシーンでも左手にあまり注意をコントロールできていない。

図5から、手品師が斜め上を見る条件においても、シーン4以降、ほとんど意図領域に注意を誘導できていないことがわかる。

もし、手品師の視線だけが観客の注意に影響を与えるのであれば、手品師が正面を向いている条件では観客の注意は手品師の顔に固定されるはずであるが、実際にはそうではない。一方、通常条件のシーン2では手品師の意図領域である右手にうまく注意を誘導できていないことがわかる。これは手品師の視線が意図領域に向いていないからと考えられる。

以上のことから、手や顔、視線単独ではなく、それらの関係が他者の注意の誘導に寄与するということが示唆された。

(2) 人間の注意のモデル化

人間の注意には、(1)での解析結果から得られたような他者の存在や振る舞いだけでなく、ボトムアップの刺激が影響を与えることが知られている。

そこで本研究では、人間の視覚的注意には以下の3要素が影響を与えると仮定した。

- 入力画像の視覚的顕著性
- 他者の手や顔の位置
- 他者の視線

視覚的顕著性については、Itti らによる Saliency Map によって表現することとし、他者からの影響については Manipulation Map というものを提案した。また、これら2つのマップを統合した Attention Map によって人間の視覚的注意をモデル化した。

a. Saliency Map

本研究では, Itti らの手法に基づき, 入力画像に対して, 輝度, 色, 回転, 動きの4つの要素に関する特徴マップを生成し, それらを正規化したものの線形和をとることで視覚的顕著性を表現する Saliency Map を生成した.

b. Manipulation Map

Manipulation Map は, 顔と手の位置を表す Face-hand Map と視線の方向を表す Gaze Map の線形和によって生成することとした.

Face-hand Map は顔の位置と両手の位置まわりの2次元正規分布を重ね合わせたものとして, 以下の式で表現した.

$$F(\mathbf{x}) = \sum_i \frac{1}{2\pi\sqrt{|S_i|}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mathbf{x}_i)^T S_i^{-1}(\mathbf{x} - \mathbf{x}_i)\right\}$$

ここで, $i=F, R, L$ であり, それぞれ顔, 右手, 左手を表す.

Face-hand Map の例を図6に示す.

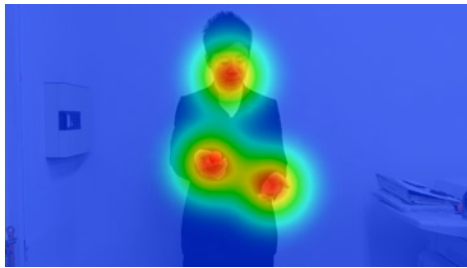


図6 Face-hand Map の例

また, Gaze Map は視線方向を表す半直線との距離 d に応じた以下の式で表現することとした.

$$G(\mathbf{x}) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp\left(-\frac{d(\mathbf{x})^2}{2\sigma_d^2}\right)$$

Gaze Map の例を図7に示す.

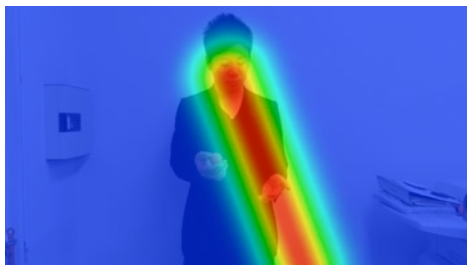


図7 Gaze Map の例

これら Face-hand Map と Gaze Map の線形和で Manipulation Map を生成する.

c. Attention Map

上記のようにして生成された Saliency Map

と Manipulation Map の線形和をとることで, Attention Map を生成する.

図8に, Attention Map 生成の概略図を示す.

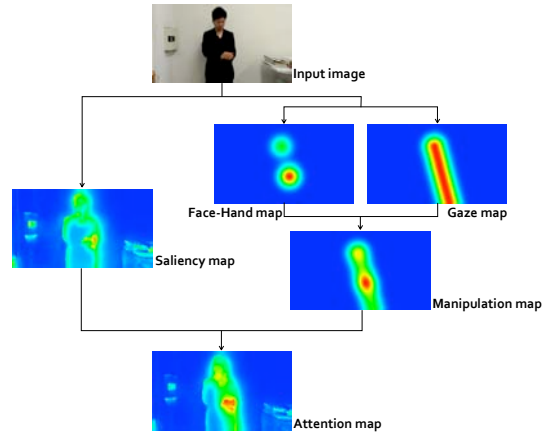


図8 Attention Map の生成

提案したモデルが前述の仮定に基づいた設計意図通りの挙動を示すかを確認するため, まず, (1)で使用した3つの手品動画に対して, モデルを適用した. 図9に各シーンの代表画像におけるモデルの適用結果の例を示す.

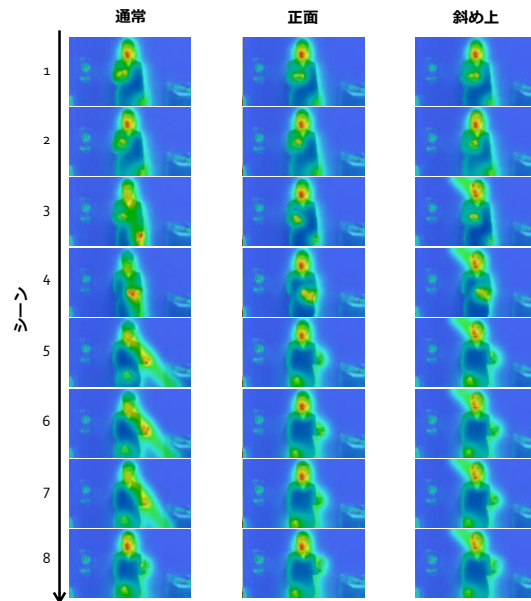


図9 手品動画への Attention Map の適用

図9から, 定性的にはあるが, 設計意図通りの挙動を示すことが確認できた.

次に, 提案モデルの有用性を確認するため, モデルがどの程度実際の注視点の分布を説明できているかを, Saliency Map と比較することで確認した. 図10に, 通常条件での動画における Saliency Map, 提案モデル (Attention Map), 注視点分布を示す.

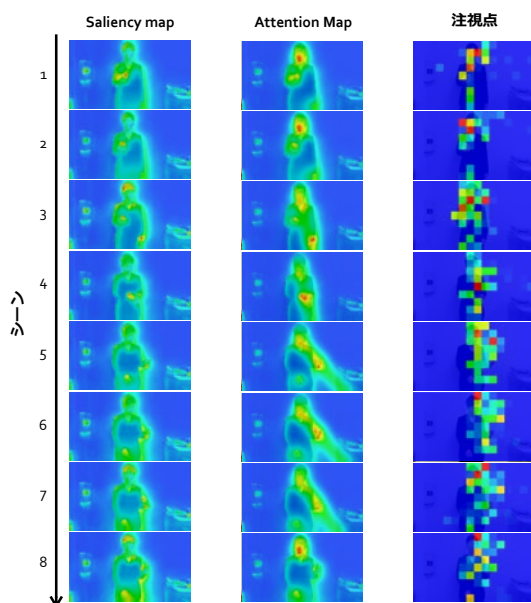


図 10 注視点と提案モデル, Saliency Map の比較

図 10 から, 提案モデルの方が, Saliency Map よりも注視点分布によく似ていることがわかる。

注視点の位置における Attention Map の出力値および Saliency Map の出力値の, それぞれの画像全体での平均値との比を比較すると, 図 11 のようになった。

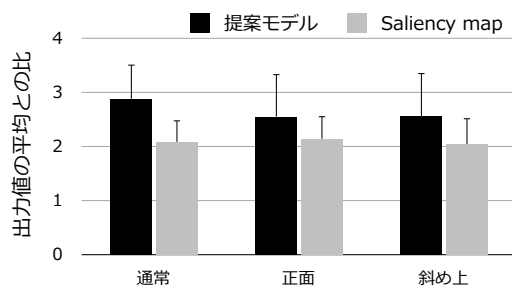


図 11 提案モデルと Saliency Map の注視点における出力値/画像全体の平均値の比較

この図からわかるように, 3 条件全てにおいて, 提案モデルが Saliency Map よりも高い値を示した. この比が高いということは, モデルが注視点分布をよく説明できているということを意味している. したがって, このことから, 本研究で提案した注意モデルは人間の注意をよく説明できると言うことが示された。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計 0 件)

〔学会発表〕 (計 7 件)

- ① Yusuke Tamura, Shiro Yano, Hisashi Osumi, Visual Attention Model for

Manipulating Human Attention by a Robot, Proceedings of the 2014 IEEE International Conference on Robotics and Automation, pp.3507-3512, 査読有り, Hong Kong, China, 2014 年 6 月 4 日

- ② Yusuke Tamura, Shiro Yano, Hisashi Osumi, Modeling of Human Attention Based on Analysis of Magic, Proceedings of the 9th ACM/IEEE International Conference on Human-Robot Interaction, pp.302-303, 査読有り, Bielefeld, Germany, 2014 年 3 月 4 日
- ③ Takafumi Akashi, Yusuke Tamura, Shiro Yano, Hisashi Osumi, Analysis of Manipulating Other's Attention for Smooth Interaction between Human and Robot, Proceedings of the 2013 IEEE/SICE International Symposium on System Integration, pp.340-345, 査読有り, Kobe, Japan, 2013 年 12 月 15 日
- ④ 田村雄介, 矢野史朗, 大隅久, 注意誘導のための視覚的注意のモデル化と手品鑑賞時の注視点計測に基づくモデルの評価, 第 31 回日本ロボット学会学術講演会, 2H1-04, 首都大学東京, 2013 年 9 月 5 日
- ⑤ 明石貴文, 田村雄介, 矢野史朗, 大隅久, 人間とロボットのインタラクション円滑化に向けた注意誘導の解析, 日本機械学会ロボティクス・メカトロニクス講演会 2013, 2P1-P14, つくば, 2013 年 5 月 24 日
- ⑥ 田村雄介, 明石貴文, 矢野史朗, 大隅久, 手と視線の関係が注意誘導に与える影響, HAI シンポジウム 2012, 2A-3, 京都工芸繊維大学, 2012 年 12 月 8 日
- ⑦ 田村雄介, 矢野史朗, 大隅久, 注意誘導動作のモデル化のための手品鑑賞時の視線計測, 第 30 回日本ロボット学会学術講演会, 2N1-4, 札幌, 2012 年 9 月 18 日

〔その他〕

ホームページ等

http://www.tamlab.jp/?page_id=145

6. 研究組織

(1) 研究代表者

田村 雄介 (TAMURA YUSUKE)

中央大学・理工学部・助教

研究者番号: 40515798