

擬人的媒体によるユーザの視線に応じた視線行動の検討

米澤 朋子[†] 山添 大丈[†] 内海 章[†] 安部 伸治[†]

[†] ATR 知能ロボティクス研究所 〒619-0288 京都府相楽郡精華町光台 2-2-2

E-mail: †{yone,yamazoe,utsumi,sabe}@atr.jp

あらまし 本研究は、擬人的媒体の視線行動によるユーザのコミュニケーション欲の誘発や継続を狙いとし、画像処理による非装着型視線推定から検出されたユーザの視線方向に応じ、共同注視とアイコンタクトを段階的に適用しぬいぐるみ型ロボットが視線コミュニケーションを行うシステムを紹介する。擬人的媒体の視線行動に対するユーザの印象や行動への影響を評価した結果、i) ロボットの注視対象に対する被験者の共同注視行動、ii) ロボットの共同注視行動に喚起される、被験者のロボットへの興味、iii) 被験者とのアイコンタクトに応じたロボットの反応に喚起される被験者のロボットへの好意、iv) 被験者の視線に応じたロボットの共同注視行動による自然性の評価の上昇、を確認した。

キーワード ぬいぐるみロボット、視線行動、共同注視、アイコンタクト、視線コミュニケーション、

Effectiveness of Gaze-communicative Interaction between Human and Stuffed-toy Robot Corresponding to User's Gaze

Tomoko YONEZAWA[†], Hirotake YAMAZOE[†], Akira UTSUMI[†], and Shinji ABE[†]

[†] ATR Intelligent Robotics and Communication Labs. Hikaridai 2-2-2, Seika-cho, Kyoto 619-0288 Japan

E-mail: †{yone,yamazoe,utsumi,sabe}@atr.jp

Abstract This paper introduces a gaze-communicative system for a human and a stuffed-toy robot with both behaviors of joint attention and eye-contact reactions. These behaviors are corresponding to the user's gaze which is estimated by ambient gaze-tracking. For free and natural interaction, we adopted our remote gaze-tracking method based on 3D eye-balls models. From demonstration experiments, we confirmed that i) joint attention draws the user's interest along with the user-guessed interest of the robot, ii) "eye contact" brings the user a favorable feeling for the robot, and iii) this feeling is enhanced when "eye contact" is used in combination with "joint attention."

Key words Stuffed-toy Robot, Gazing behavior, Joint Attention, Eye-contact, Gaze-communication

1. はじめに

実体型や仮想型のロボットやエージェントが人とマルチモーダルな表現を伴いながら社会的にコミュニケーションする状況を想定した様々な研究が行われている [4, 23, etc.]. このような擬人的システムは、擬人的表現や外観による自然で親しみやすいやりとりによって、感情を込めたコミュニケーションを喚起する可能性を期待するものである。Kendon [8] は人間同士の視線におけるふるまい（以後**視線行動**と呼ぶ）について検討しており、マルチモーダル表現の1つのチャンネルとして視線行動の重要性が注目される。特に、擬人的媒体、たとえば仮想エージェント [5] やヒューマノイドロボット [6, 25] などによる視線行動は、彼ら擬人的媒体の内部状態を人間が推測する時の1つの材料となっている可能性がある。多種多様な視線行動のうち、共同注視は、母親と幼い子供の間のコミュニケーションで見られる視線行動で、子供の学習に重要な役割を果たしている [11]。ロボットのふるまいとして上記の共同注視を取り入れることで、子供のようなふるまい

を見せる Infanoid [9] も提案されている。

擬人的媒体の中でもぬいぐるみは、ままごと遊びや腹話術などにおいて、アバターや（擬人的）媒体として、もしくはパートナーとして遊びの中で適応的に用いられている。この多様性を利用して、大地震などによってトラウマを負った子供たち [16] や認知症の患者 [12] の心理面のケアでは、親しみやすく強制的ではないコミュニケーションを提供する動物や人形のぬいぐるみが用いられている。

本稿では、ぬいぐるみロボットが人間に対しより自然なマルチモーダル表現を行うために、特に視線行動について検証する。動物を模したぬいぐるみでは、白目が表されず黒目のみにより構成されたものが多いため、実際の人間の視線ほど視点がはっきりせず独特の視線の表現性がある。よって、効果的にぬいぐるみの視線行動を示すためには、ふるまいが単純で分かりやすいことが必要とされる。

本研究は、様々な要因からコミュニケーション欲が低い状態にあるユーザに対し、コミュニケーションへのきっかけとなる仕組

みを明らかにし、擬人的媒体による自然なコミュニケーションの喚起を実現させることを狙い、ぬいぐるみロボットとユーザの間にコミュニケーションを誘発・成立させることを目指す。コミュニケーションの動機としてユーザの“興味”をひきつけ、その結果ユーザの“好意”を得ることで更なるコミュニケーションの継続がもたらされる、といったように、コミュニケーションの進行度に応じてロボットの視線行動が段階的に変化すれば、ユーザの様々な身体的／精神的な状況に適応しつつユーザの意識的もしくは潜在的な感覚を刺激することができると予想した。そこで、1) ユーザの視線に対する共同注視のふるまい（**共同注視行動**）によりユーザの興味を間接的にひきつけ、2) ユーザのぬいぐるみへの視線に対するふるまい（**アイコンタクト反応**）により直接働きかけることでユーザの好意をひきつける、という段階性をもたせた視線コミュニケーションのモデルを仮定し、ユーザの視線位置に応じてコミュニケーションをしているような視線行動を示すぬいぐるみロボットを IPRobotPHONE [18] を用いて実装した [24, 30]。

ぬいぐるみとの視線コミュニケーションをより自然で親しみやすいものにするためには、ユーザに意識させない視線検出装置であることがのぞましい。しかし従来の視線検出手法は特定の装具が必要であったり [1, 13]、画像処理による視線検出手法であってもステレオカメラが必要、特殊なキャリブレーションによってユーザに負担が生じる [10, 14, 15]、もしくは高解像度画像が必要である [21, 27] など、様々な制限が存在している。これらに対し、単眼カメラにより遠隔からユーザの視線角度を推定する我々の視線推定手法 [22] は、遠隔にある隠れたカメラなどから何もユーザに装着させずに視線を検出できるため、ユーザの無意識的な視線を検出することが可能である。よって提案システムの自然な視線コミュニケーションを実現するためにこの視線推定手法を導入した。

2. 関連研究

ユーザの視線と関連付けられた、実体を持つエージェントによる身体動作に関する多くの研究がある。Castellini らはユーザの視線を入力として導入したロボットの遠隔操作手法を提案している [3]。しかしこの研究は、制御のための入力チャンネルとしての視線に焦点を当てており、ロボットをコミュニケーションに用いておらず、視線を用いることの社会的・心理的影響力には言及していない。これに対し小川原ら [29] は、視線により迷いの状態を検出することで効率的な協調作業を進めるシステムを提案している。

一方で、人間とロボットの間の社会的インタラクションについて検討しているいくつかの研究もある。Scassellati はヒューマノイドロボットに対する幼児の心理に関する理論とインタラクションの設計について述べている [17]。また、中野ら [28] はユーザの視線に応じた観光案内エージェントの挙動を扱っている。しかしこれらは実験などによる十分な検証はなされていない。Breazeal らは、他者のまねをするようなロボットのふるまいが社会的コミュニケーションにおいて効果を発揮する可能性を示唆している [2]。また、Sidner ら [19] はロボットによって共同注視行動や様々なジェスチャを含めた身体動作を実現し、ユーザの注意を引き付けるのにこのようなジェスチャが有効であることを示している。さらに、Thomaz ら [20] はロボット Leonardo を用い、指差しなどの社会的参照行為のモデルに基づいて感情に関して議論している。これらの研究は、コミュニケーションにおける人間の行動のモデルをそれぞれ実現しているが、これらは共同注視を取り上げているにもかかわらず、視線コミュニケーションを段階的に構成するように複数種類の視線行動を組み合わせたか、その構成を十分に検

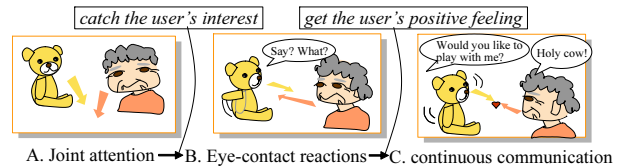


図 1 段階的な視線コミュニケーションモデル
Fig. 1 Gradual Gaze-communication Model

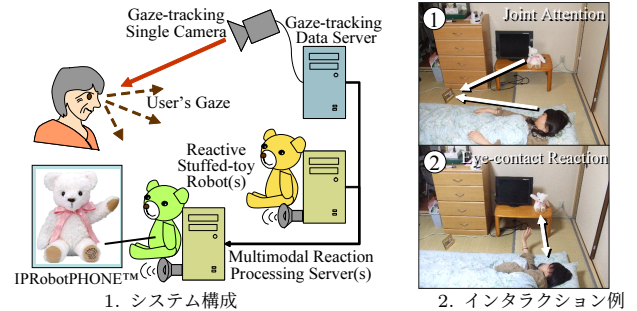


図 2 提案システム
Fig. 2 System Implementation

証するなどの議論はされていない。

3. システム構成

3.1 視線コミュニケーションのモデル化とデザイン

ここでは、ユーザのコミュニケーション欲を喚起するため、ユーザの意識的もしくは無意識的な視線に応じて視線行動を表現して視線コミュニケーションを実現するぬいぐるみロボットを提案する。

まず、本システムの基盤として視線コミュニケーションのモデルを仮定した（図 1）。最初の段階では、ぬいぐるみロボットの共同注視行動が間接的にユーザの注意を喚起し、コミュニケーションのきっかけとなる（図 1-A）。この段階ではまだ視線は合っており、ロボットは積極的に話しかけないためユーザにとってもロボットにとっても互いに距離感のある状態であるが、同時にユーザの潜在的な注意がロボットにひきつけられた状態とも考えられるため、共同注視を介した無意識的なコミュニケーションが成り立ったものとする。これによって、以後ロボットが動きや音声によりユーザへ話しかける段階に進んだ時に、コミュニケーションをいきなり求められたかのようにユーザが捉えたり負担を感じる可能性が低くなると考える。

その後、ユーザがロボットに興味を持ちロボットへ視線を向けると、アイコンタクトが成立する。このときロボットが動きや音声などによってアイコンタクトの成立を表し、ユーザへの好意を示したりユーザの好意をひきつける（図 1-B）。これにより継続的なコミュニケーション（図 1-C）がもたらされる。

3.2 視線コミュニケーションシステムの構成

ぬいぐるみロボットは腕や頭によるジェスチャ動作や音声出力ができるものを用いるため IPRobot-PHONE [7] を採用した。このロボットは頭部と腕がそれぞれ 2 自由度、計 6 自由度あり、専用ライブラリにより動作制御プログラムを構成した。

本稿で提案するぬいぐるみロボットによる視線コミュニケーションシステムの構成を、図 2-1 に示す。まずユーザの視線角度を検出するために、ユーザの顔画像を取り込む単眼カメラを設置する。視線検出サーバは、カメラから取り込まれた顔画像から画像処理によってユーザの縦横の視線角度を検出し、マルチモーダル反応生成サーバに各視線角度の情報を送信する。このマルチモーダル反応生成サーバは、ユーザとぬいぐるみロボットの 3 次元上

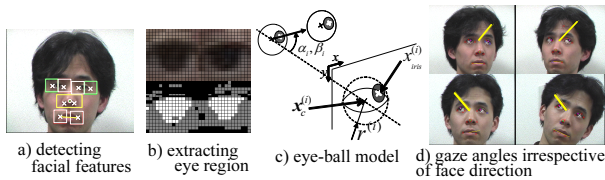


図 3 視線検出結果

Fig. 3 Results of Gaze-tracking

の位置情報とリアルタイムに得たユーザの視線角度から導出されるユーザ視線の相対位置に応じて、1 体もしくは複数のぬいぐるみロボットの視線動作やその他のジェスチャ動作および音声表現を制御する。複数のぬいぐるみロボットを用いる際には、他のロボットの動きと連動しながらユーザの視線に反応する。

提案システムにおける段階的な視線コミュニケーションを伴うインタラクションの例として、

0. ユーザがぬいぐるみの周囲から離れた位置を見ている時、ぬいぐるみはユーザの視線に応じた動きを伴わない。

1. ユーザがぬいぐるみの周囲（ぬいぐるみにとっての視界範囲内）を見ているとき、ぬいぐるみは声は出さずに、ユーザが見ている位置へ顔を向ける。ユーザの視線が範囲内で動いたときはその動きに追隨して顔方向を変化させる（共同注視行動、図 2-2-①）。

2. ユーザがぬいぐるみを見たとき、アイコンタクトが成立したことを示すために、音声と頭・腕を動かすジェスチャ動作により、「何?」「遊んでくれるの?」などの反応を見せる（アイコンタクト反応、図 2-2-②）。

3. ユーザのアイコンタクトの継続により、継続的なコミュニケーションの反応として、あらかじめ準備された音声とジェスチャのマルチモーダル反応を示す。

といった段階を伴いつつコミュニケーションを成り立たせる。

3.3 非装着型遠隔視線検出手法

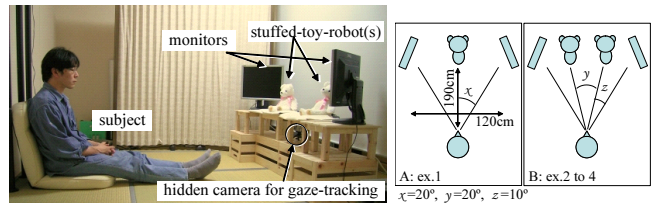
上記のような擬人的媒体との自然な視線コミュニケーションを実現するためには、ユーザに意識させることなくユーザの視線が検出できることが望ましい。そこで本システムでは、ユーザがカメラ等の装置を装着する必要がなく遠隔からの視線検出が可能な手法 [26] を用いる。手法 [26] では単眼カメラから得られた画像を分析することにより視線推定を行う。

手法 [26] では、眼球中心と虹彩中心を結ぶ 3 次元ベクトルを視線方向として検出している。画像からは実際の眼球中心は計測できないが、あらかじめ眼球中心と顔特徴点の相対関係をモデル化しておくことで、顔特徴点から眼球中心位置を求めることが可能となる。眼球中心と顔特徴点のモデル化にあたって、従来手法と異なり、特定位置を注視する必要はない。

視線検出においては、まず画像から目・鼻・口などの顔特徴点を抽出し（図 3-a）、目周辺領域画像から虹彩領域を検出する（図 3-b）。次に、虹彩領域から楕円近似により虹彩中心を推定するとともに、顔特徴点から眼球中心位置を推定し、3 次元眼球モデル（図 3-c）に基づき視線方向を推定する（図 3-d）。視線推定精度は水平方向 $\pm 5^\circ$ 、垂直方向 $\pm 7^\circ$ となっており、遠隔にあるぬいぐるみロボットとユーザの視線が合っているかどうかなど、ぬいぐるみの位置とユーザの視線との関係を測定するには十分な精度が得られている。

4. 実験

本節では、擬人的媒体とユーザの間における段階的な視線コミュニケーションのモデルについて、それぞれの視線行動の効果を検証する。まず、ぬいぐるみロボットの顔方向に応じたユーザの共同注視行動とその効果を検証するため、1) ロボットの顔方向に応じたロボットの興味の推測、および 2) ロボットの興味と同



1. 実験風景

図 4 実験環境

Fig. 4 Experimental Environment

じ対象へのユーザの興味の喚起への影響をそれぞれ調べる。次に、ユーザの視線に応じて変化するロボットの視線行動が与える印象を検証するため、3) ロボットの共同注視行動（ユーザ主体）とアイコンタクト反応による視線行動がロボットに対する好意や自然性の印象へ与える影響を調べる。

共通実験手法・実験設定、および実験教示: 各実験は 1 分ずつ行った。実験中に被験者がぬいぐるみロボットの様子を見るよう、「実験後にロボットの考えていたことを推定し 10 文字以内で答える」というタスクをあらかじめ説明した。

図 4-1、4-2 に実験環境と実験装置の配置を示す。ロボットの視線行動の対象として、17 インチの液晶ディスプレイを 2 個準備し、それぞれの液晶ディスプレイ上に表示領域の 4 分の 1（縦 2 分の 1 横 2 分の 1）の大きさで準備したアニメーション対を表示した。これらのアニメーション対は、野菜などの単純なコマ送りアニメーションであり、6 秒毎に更新される。この視覚的刺激対は、実験の第一刺激である一体もしくは二体のぬいぐるみロボットの挙動（第一刺激）を評価する際、共同注視対象としてもしくは被験者の視線の対象を絞るための刺激（第二刺激）として用いる。

実験に共通して参加した被験者: 21 歳から 40 歳の男性 12 名、女性 10 名の計 22 名

4.1 ロボットの共同注視行動による興味誘引の評価

実験仮説: I) 被験者は、ぬいぐるみロボットが長い間顔を向けていたほうのアニメーションを、ロボットの“興味”がある対象として推測する。II) 被験者のアニメーションへの“興味”は、ロボットの視線行動に影響を受ける。III) 被験者はロボットがより長く顔を向けたアニメーションのほうを、もう一方のアニメーションより長く見つめる。

実験刺激: ぬいぐるみロボットは図 4-2-A のように 2 つのディスプレイの中心位置に配置される。約 1 分間の実験中、ロボットの顔方向は、A) 長く見つめるアニメーションの映される液晶ディスプレイ (LM) の方向には、合計約 30 秒間（最短 1 秒間、最長 6 秒間）、B) あまり見つめないアニメーションが映されるディスプレイ (OM) の方向には、合計約 20 秒間（最短 1 秒間、最長 5 秒間）、C) 被験者の方向には、合計約 10 秒間（最短 0.5 秒間、最長 1 秒間）へ、それぞれ向けられる。これらの視線方向の転換のタイミングはあらかじめ乱数により決定され、被験者間で共通のシナリオとして用いられる。左右のディスプレイはアニメーション対をそれぞれ表示する。左右ディスプレイと各アニメーション対の表示間の組み合わせは、被験者毎にランダムに決定する。

実験手法と教示: 被験者はぬいぐるみロボットの正面に座り（図 4-2-A）、楽な姿勢をとる。実験の前に、実験者から『実験後、ぬいぐるみの考えていたことを推し量って 10 文字以内で記述してもらうので、実験中はぬいぐるみの様子を見ていてください』と教示される。被験者は 1 分間の実験の後、次の各質問項目

Q1: ぬいぐるみは左右どちらのアニメーションに興味を持っていたか

Q2: 被験者は左右どちらのアニメーションに興味を持ったか

表 1 コミュニケーションにおける興味の主観評価分析
Table 1 Analyses of Interest

比較対	興味	ave.	stdev.	$t_{(20)}$	p(t)
LM & OM	robot(Q1)	5.19	0.98	5.56	< .01
	subject(Q2)	4.43	1.33	1.48	= .154

についてそれぞれ、7: 絶対左側, 6: おそらく左側, 5: たぶん左側, 4: どちらともいえない, 3: たぶん右側, 2: おそらく右側, 1: 絶対右側として、7段階評価を決定する。実験後、右が LM だった場合の評価を反転し集計する。

実験結果: ロボットの挙動を失敗した 1 人の被験者を除き、残りの 21 名の被験者の実験データを取得した。以下に 3 種類の分析結果を示す。

ロボットと被験者の興味に関する主観評価の分析: Q1 および Q2 に対する被験者の主観評価をまとめ、評価値の傾向を調べるため 1 標本の T 検定 (予測値=4) を行った。分析結果を表 1 に示す。これによると Q1 の平均が予測値と有意に異なることがわかった ($T_{(20)} = 5.56, p < .01$)。仮説 I で述べたように、被験者は、ロボットが長時間顔方向を向けていた LM 側が、ロボットの興味のあるアニメーションであると推測したことが示された。一方仮説 II については有意な結果は確認されず、被験者が興味を持つアニメーションはロボットが顔を向けている方向に影響されることはなかった。

モニターやロボットへの注視時間の分析: 被験者の視線の方向を、LM, OM, およびぬいぐるみロボット (SR) の 3 方向に分類し、それぞれの注視時間を被験者毎に合計した (表 2)。1 要因分散分析 (ANOVA, $\alpha = .05, df = 20, 2$) により、平均の差が有意であるという結果を得た ($F_{(20,2)} = 8.34, p < .01$)。Scheffé の事後検定を行った結果、LM-SR 間 ($p < .01$)、および LM-OM 間 ($p = .01$) に有意差を確認した。これらの結果は仮説 III を支持しており、被験者は、ロボットが顔方向を長時間向けていたディスプレイ側を長時間見ていたことが確認された。

被験者の視線の転換とロボットの顔方向の関係: 主観評価では確認されなかった、被験者の LM への無意識の興味を計測するため、ロボットの顔方向と被験者の視線方向の関係性を分析することとした。その中で特に、被験者がロボット (SR) から LM または OM へ視線方向を転換させるときの、ロボットの顔方向に着目した。

まず、被験者の視線方向およびロボットの顔方向の例を図 5 に示す。多くの被験者は、図 5-subject A のように、ロボットの顔方向を追いかけるような視線の動きを見せたが、一方で、図 5-subject B のように、ロボットが見ている方向とは無関係に両方のディスプレイをすばやく交互に見るような動きも観察された。

次に、被験者の視線方向の転換点について分析する。図中の“▼”記号は SR から LM への転換点、“▲”記号は SR から OM への転換点を示す。これらの転換点についての分析を表 3 に示す。SR から LM への転換点は被験者間合計で 191 個あり、そのうち 127 個は、ロボットが LM へ顔を向けていた時だった。SR から OM への転換点は、被験者間合計 191 個のうち 87 個がロボットが OM へ顔を向けていた時だった。これらのついて、ロボットの顔方向の時間配分から予測値を計算した結果を表中に prd. として示している。これらの値を元にそれぞれ χ^2 検定を行った結果、予測値と実測値の差は有意であることが分かった。つまり、Q2 の主観評価では被験者の興味は LM にひきつけられていないという結果であったが、この分析により、被験者は無意識のうちにロボットの顔方向に応じてアニメーションを見ていたことが示された。このようにロボットの視線行動が被験者の視線行動に影響を

表 2 被験者の方向別注視時間
Table 2 Total Gazing Duration of Subjects by Each Direction

	LM	OM	SR
average (stdev)	25.45 (7.255)	16.05 (7.06)	17.49 (5.41)

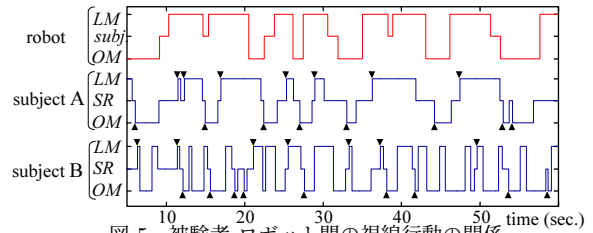


図 5 被験者-ロボット間の視線行動の関係

Fig. 5 Relation of Robot's and Subject's Gaze

表 3 視線の方向転換についての χ^2 検定
Table 3 Summaries of Gaze-switching and χ^2 Test

switch to	total	same (prd.)	else (prd.)	χ^2	p
LM “▼”	191	127(95.5)	64(95.5)	20.8	< .01
OM “▲”	157	87(52.3)	70(104.7)	34.4	< .01

表 4 アイコンタクト反応による好意性と自然性の主観評価
Table 4 Analyses of Liking and Naturalness with Eye-contact Reaction

比較対	評価項目	ave.	stdev.	$t_{(20)}$	p(t)
B1 & B2	好意 (robot,Q3)	5.190	1.29	4.232	< .01
	好意 (subject,Q4)	4.714	1.65	1.987	= .06
	自然性 (Q5)	4.286	1.271	1.031	= .315

及ぼすことが確認された。

4.2 ロボットの視線行動に対する印象評価

4.2.1 アイコンタクト反応による好意性と自然性

実験仮説: I) 被験者は、ぬいぐるみロボットがアイコンタクト反応を示すことにより、ロボットの被験者への好意を推測する。II) 被験者は、ロボットのアイコンタクト反応により、ロボットへの好意を感じる。III) 被験者は、ロボットのアイコンタクト反応に自然性を感じる。

実験刺激: ぬいぐるみロボットは、図 4-2-B に示すように 2 体を並べて配置され、それぞれが以下の 2 つの異なる視線行動をとるものとした。

–B1: 被験者の視線がロボットに注がれた時に、ロボットは音声および動作によりアイコンタクト反応を示す (アイコンタクト反応あり)。

–B2: 被験者の視線に関係なく、ロボットは音声および動作による反応を示す (アイコンタクト反応なし)。

実験中、ロボット以外の被験者の視線を分散させないため、両端のディスプレイに第二刺激であるアニメーション対をそれぞれ表示する。

実験手法と教示: 被験者は、2 体のぬいぐるみロボットの正面に座り (図 4-2-B)、前節で述べた実験と同様に、実験前に「実験中ぬいぐるみの考えていたことを想像し、実験後単純に記述する」ことを教示される。1 分間の実験の後、被験者は次の各質問項目 Q3: 左右どちらのロボットが被験者に対してより好意を持っていたか Q4: 被験者は左右どちらのロボットに対し好意を感じたか Q5: 被験者は左右どちらのロボットが自然だと感じたか についてそれぞれ 7 段階で主観評価する。

実験結果: ロボットの挙動を失敗した 1 人の被験者を除き、残りの 21 名の被験者による Q3 から Q5 の主観評価の結果を集計し、評価値の傾向を調べるため 1 標本の T 検定 (予測値=4) を行った (表 4)。分析結果より、Q3 の平均値と予測値は有意に異なり ($T = 5.19, p < .01$)、Q4 では有意傾向が認められた ($T = 4.714, p = .06$)。一方 Q5 については有意性は認められなかった。このように実験仮説 I は支持され、仮説 II の傾向もわずかに認められたが、仮説 III は証明されなかった。これらの結果より、被験者

表 5 共同注視行動による好意性と自然性の主観評価
Table 5 Analyses of Liking and Naturalness with Joint Attention

比較対	好意性	ave.	stdev.	$t_{(21)}$	$p(t)$
B3 & B4	好意 (robot,Q3)	3.86	1.25	-.513	= .613
	好意 (subject,Q4)	4.09	1.41	.302	= .766
	自然性 (Q5)	4.682	1.323	2.417	= .025

の視線に応じたアイコンタクト反応をロボットが示すことは、コミュニケーションにおけるユーザの好意をもたらす可能性があるといえる。

4.2.2 ロボットの共同注視行動による好意性と自然性

実験仮説: I) 被験者は、ぬいぐるみロボットの共同注視行動によりロボットの好意を推測する。II) 被験者は、ロボットの共同注視行動によりロボットへの好意を感じる。III) 被験者は、ロボットの共同注視行動を自然と感じる。

実験刺激: ぬいぐるみロボットは、図 4-2-B のように 2 体を並べて配置され、それぞれ以下の視線行動をとるものとした。

–B3: ロボットは、被験者の視線の位置に応じて、顔方向を向け続ける (共同注視行動あり)。

–B4: ロボットは、被験者の視線の位置に関係なく、ランダムな速度・ベクトルで、顔方向を変化させ続ける (共同注視行動なし)。

実験中、ロボットと被験者の共同注視行動の対象として、両端のディスプレイに第二刺激であるアニメーション対をそれぞれ表示する。

実験手法と教示: 4.2.1 節と同様に、被験者はまず教示をうけ、1 分間の実験に臨んだ。また実験後に、評価項目 Q3 から Q5 について 7 段階評価を求めた。

実験結果: 被験者 22 名全員において実験が手順どおり遂行された。Q3 および Q4 の主観評価について、予測値を 4 として 1 標本の T 検定を行った。主観評価の平均はそれぞれ予測値と比べても有意な差は見出されなかったが、その一方で Q5 において有意性が示された ($T = 2.417$, $p = .025$, 表 5)。このように仮説 I, II については有意な結果は得られなかったが、III は確認された。つまり、共同注視行動単独では、コミュニケーションにおけるユーザのロボットに対する好意やロボットの好意を感じさせることができるとはいえないが、ユーザに自然さを感じさせることはできるといえる。

4.2.3 複合的な視線行動による好意性と自然性

実験仮説: I) 被験者は、ロボットがアイコンタクト反応だけを行うときに比べ、アイコンタクト反応と共同注視行動の両方を行うとき、ロボットの被験者への好意をより強く感じたと推測する。II) 被験者は、ロボットがアイコンタクト反応だけを行うときに比べ、アイコンタクト反応と共同注視行動の両方を行うとき、好意をより強く感じる。III) 被験者は、ロボットが共同注視行動だけを行うときに比べ、共同注視行動とアイコンタクト反応の両方を行うとき、ロボットの被験者への好意をより強く感じたと推測する。IV) 被験者は、ロボットが共同注視行動だけを行うときに比べ、共同注視行動とアイコンタクト反応の両方を行うとき、好意をより強く感じる。V) 被験者は、ロボットの複合的な視線行動と単一の視線行動を比べ、複合的な視線行動に自然性を感じる。

実験刺激: ぬいぐるみロボットは、図 4-2-B に示すように 2 体を並べて配置された。視線行動は、

–B5: 被験者の視線がロボットに注がれた時に、ロボットは音声および動作によりアイコンタクト反応を示すとともに、それ以外のときは、被験者の視線位置にロボットの顔方向を向ける (アイコンタクト反応および共同注視行動あり、つまり B1 と B3 を同時に満たす)。

–B6: 被験者の視線に関係なく、ロボットは音声および動作による反応を示すとともに、被験者の視線位置にロボットの顔方向を向ける (アイコン

表 6 複合的な視線行動による好意性と自然性の主観評価
Table 6 Analyses of Liking and Naturalness with Combined Gazing-reactions

比較対	好意性	ave.	stdev.	$t_{(20)}$	$p(t)$
B5 & B6	好意 (robot,Q3)	3.81	1.29	-.68	= .74
	好意 (subject,Q4)	4.10	1.37	.32	= .74
	自然性 (Q5)	4.238	0.995	1.096	= .286
B5 & B7	好意 (robot,Q3)	4.62	1.07	2.648	= .02
	好意 (subject,Q4)	4.62	1.28	2.210	= .04
	自然性 (Q5)	4.286	0.784	1.671	= .110

タクト反応なし、共同注視行動あり、つまり B2 と B3 を同時に満たす)。

–B7: 被験者の視線がロボットに注がれた時に、ロボットは音声および動作によりアイコンタクト反応を示すとともに、それ以外のときは、被験者の視線位置に関係なく、ランダムな速度でランダムな方向へロボットの顔方向を向ける (アイコンタクト反応あり、共同注視行動なし、つまり B1 と B4 を同時に満たす)。

の 3 つの異なる視線行動を準備し、B5 と B6、B5 と B7 をそれぞれ比較対として、2 体のロボットそれぞれによって、各比較対 (B5–B6 または B5–B7) の視線行動を提示する。

実験中、ロボット以外の被験者の視線を分散させないため、また、ロボットの共同注視行動の対象として、両端のディスプレイに第二刺激であるアニメーション対をそれぞれ表示する。

実験手法と教示: 視線行動 B5 と B6 を比較する実験と、視線行動 B5 と B7 を比較する実験を、それぞれ前節の実験と同様に 2 体のロボットにより行う。同様の手順で実験前に教示を行い、実験後に、評価項目 Q3 から Q5 によって、左右のロボットについての主観評価を 7 段階評価で求める。

実験結果: ロボットの挙動を失敗した 1 人の被験者を除き、残りの 21 名の被験者による主観評価データを取得した。評価項目 Q3 から Q5 について、B5 と B6 の間および B5 と B7 の間における比較評価結果に対し、予測値を 4 として 1 標本の T 検定を行った結果を表 6 に示す。B5 と B7 の比較評価結果と予測値とは有意な差が認められた (ロボットの好意についての評価: $T = 2.648$, $p = .02$, 被験者の好意についての評価: $T = 2.210$, $p = .04$)。それに対し、B5 と B6 の比較評価結果と予測値には有意差は認められなかった。一方、自然性の比較評価では有意差は、B5–B6 間、B5–B7 間ともなかった。

これらの結果から、仮説 III と IV は確認されたが、仮説 I と II および V は確認されなかった。つまり、共同注視行動を常に伴うとき、アイコンタクト反応の有無が視線コミュニケーションにおける好意性に与える影響があるかは確認されなかったが、アイコンタクト反応を常に伴うとき、共同注視行動の有無が視線コミュニケーションにおける好意性に影響した可能性がある。

5. 考 察

4.1 節に述べた実験結果より、ぬいぐるみロボットの顔方向が、被験者に「ロボット自身の興味」を推定させる要素であることが示唆された。LM:OM の時間の比率は 3:2 であり大差はなかったのにもかかわらず、Q1 については顕著な結果となった。被験者自身のアニメーションへの興味は、主観評価 (Q2) では LM–LM 間で有意な差は観察されなかったが、被験者の視線について客観評価を導入した結果、被験者も有意に LM を注視しており、また、視線の方向転換はその時々ロボットの顔方向に伴っていることが分かった。つまりこれらの結果から、被験者の無意識的な視線行動はロボットの視線行動に影響されるといえる。よって、視線コミュニケーションを伴うぬいぐるみロボットが、ユーザの興味を無意識的なレベルでひきつける可能性が示唆された。

4.2 節では、ユーザの視線に伴ったロボットの共同注視行動およびアイコンタクト反応が、ユーザの好意や自然性の評価へそれ

それ異なる影響を及ぼすことが示された。アイコンタクト反応が無いときに比べて、アイコンタクト反応が単独で用いられる時は、被験者のロボットに対する好意性とロボットが被験者に対し持っていると推測される好意性はともに強まるという結果が得られた。ロボットが共同注視行動のみを示した場合と共同注視行動を示さなかった場合を比較したときは好意性の有意な差は見られなかったものの、自然性において有意差が確認された。また、アイコンタクト反応を伴いながらの共同注視行動の有無を比較評価した場合では、共同注視行動をした場合のほうが視線コミュニケーションの好意性を有意にひきつけていた。これらの結果より、コミュニケーションの自然性を高めるために共同注視行動を積極的に用いたり、コミュニケーションにおける好意性を強めるために、共同注視行動とアイコンタクト反応をコミュニケーションの状態に応じ両立させたり使い分けたりするといった、ぬいぐるみロボットの効果的な視線行動を構成することが重要であることを示唆している。

視線コミュニケーションシステムの構成においてあらかじめ予測していたモデルのとおり、アイコンタクト反応は直接的に好意性を高めることが確認された。その一方で、共同注視行動が単独で用いられるだけでは好意性は高まらなかった。実際のシステムでは、この2種類の視線行動にとどまらず、ぬいぐるみロボットの様々な反応を適切に組み合わせたり使い分けられるよう準備することで、より効果的に感情表現を含む豊かなコミュニケーションを可能にすることができると考えられる。

従来の視線追跡手法では非装着および遠隔での視線追跡が難しかったため、現実の空間内で制限や心理的な負担をかけない自然な視線コミュニケーションを実現することができなかったが、提案する視線コミュニケーションシステムでは、視線検出手法 [26] により、ユーザは物理的にも心理的にも自由にふるまいながら、ぬいぐるみロボットと視線コミュニケーションを交わすことが可能になった。そして、視線コミュニケーションのモデルの仮説のとおり、A) ぬいぐるみロボットの共同注視行動が自然性をもたらしつつユーザの無意識的な興味をひきつけ、B) アイコンタクト反応が好意性を強めることでユーザのコミュニケーション欲を喚起させることが、実験により示唆された。さらに、C) コミュニケーションが成立した後にアイコンタクト反応と両立してさらに共同注視行動を行うことで、ユーザの好意性をより強め、継続的な視線コミュニケーションを可能にすることができると考えられる。本稿では、様々なモダリティを伴う擬人的な表現のうちの、視線行動という1チャンネルの効果的な表現を議論した。今後の実体を伴う擬人的媒体の表現では、視線行動とあわせて多チャンネルの表現モダリティを複合的に構成することを検討されるべきである。

6. おわりに

本稿では、ユーザとロボットの視線コミュニケーションを段階的にモデル化し組み込んだ、ユーザの視線に応じ視線行動を行うぬいぐるみロボットを紹介した。本システムでは、ユーザの視線を計測するために遠隔視線検出を用い、ユーザの視線に応じたロボットの視線行動を、共同注視行動およびアイコンタクト反応の2種類の視線行動より導出している。被験者実験により、i) ユーザの視線位置に応じたロボットの共同注視行動は、ロボットの『興味』を推測させたり、ユーザの無意識的な興味をひきつけること、ii) ロボットのユーザの注視に対するアイコンタクト反応は、ユーザに対するロボットの『好意』を推測させたり、ユーザのロボットへの行為をひきつけること、iii) ロボットの共同注視行動は自然

性の評価を向上させること、iv) ii) で述べたコミュニケーションにおける好意性は、ロボットが、ユーザの注視に対するアイコンタクト反応に加えて、ユーザの視線位置に応じた共同注視行動を示すことにより、さらに強まること、がそれぞれ示され、実体を伴うぬいぐるみロボットにおける視線コミュニケーションのモデルの有効性を確認した。これらの結果に基づき、ユーザの継続的なコミュニケーションを可能にするためには、視線コミュニケーションの最終段階として、ロボットの視線行動では共同注視行動とアイコンタクト反応を複合的に表現させることが重要だと結論付ける。視線コミュニケーションは、人と人のマルチモーダルなコミュニケーションの重要なチャンネルのひとつであり、より表情豊かなロボットのマルチモーダル表現として、今後も十分検討されるべきである。

謝 辞

本研究は情報通信研究機構 (NICT) の研究委託により実施したものである。実験に関する議論でアドバイスを頂いた鈴木紀子氏、また実験にご協力をいただいた ATR-IRC の皆様に感謝する。

文 献

- [1] Arrington Research. Viewpoint eyetracker. <http://arringtonresearch.com/index.html>.
- [2] C. Breazeal, D. Buchsbaum, J. Gray, D. Gatenby, and B. Blumberg. Learning from and about others: Towards using imitation to bootstrap the social understanding of others by robots. *Autonomous Robots*, 11, 2005. Issues 1-2.
- [3] C. Castellini and G. Sndini. Gaze tracking for robotic control in intelligent teleoperation and prosthetics. In *COGAIN 2006*, pages 73-77, 2006.
- [4] B. R. Duffy. Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3):177-190, 2003.
- [5] A. Fukayama, T. Ohno, N. Mukawa, M. Sawaki, and N. Hagita. Messages embedded in gaze of interface agents - impression management with agent's gaze-. In *Proc. ACM SIGCHI2002*, volume 1, pages 41-49, 2002.
- [6] M. Imai, T. Ono, and H. Ishiguro. Physical relation and expression: Joint attention for human-robot interaction. In *IEEE Int. Workshop on Robot and Human Communication*, pages 512-517, 2001.
- [7] IWAYA-Corporation: "IP RobotPHONE", <http://www.iwaya.co.jp/Fgoods/ip.html>, 2002.
- [8] A. Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22-63, 1967.
- [9] H. Kojima. Infanoid: A babybot that explores the social environment. *Socially Intelligent Agents, K. Dautenhahn and A. H. Bond and L. Canamero and B. Edmonds, Kluwer Academic Publishers*, pages 157-164, 2002.
- [10] Y. Matsumoto and A. Zelinsky. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *Proc. Int. Conf. Automatic Face and Gesture Recognition*, pages 499-504, 2000.
- [11] C. Moore, P. J. Dunham, and P. Dunham. *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum, 1995.
- [12] D. Moore. 'it's like a gold medal and it's mine' - dolls in dementia care. *Journal of Dementia Care*, Vol 9 No.6 2001, 9(6):20-22, 2001.
- [13] nac image technology Inc. eye mark recorderEMR-8B. http://www.ceatec.com/en/2005/news/ne_web_detail.html?volume=15.
- [14] T. Ohno, N. Mukawa, and S. Kawato. Just blink your eyes: A head-free gaze tracking system. In *Proc. CHI2003*, pages 950-951, 2003.
- [15] E. Poggini, A. Redert, I. Patras, and E. Hendriks. Gaze tracking by using factorized likelihoods particle filtering and stereo vision. *Proc. 3DPVT'06*, pages 57-64, 2006.
- [16] S. Ratjatanawan. Tsunami children foundation (tcf) psychological support services, 2005. <http://tsunamifcf.org/psychological.htm>.
- [17] B. Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12, 2002.
- [18] D. Sekiguchi, M. Inami, and S. Tachi. RobotPHONE: RUI for interpersonal communication. In *CHI2001 Extended Abstracts*, pages 277-278, 2001.
- [19] C. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich. Explorations in engagement for humans and robots. *Artificial Intelligence*, 12, 2005. Issues 1-2.
- [20] A. L. Thomaz, M. Berlin, and C. Breazeal. An embodied computational model of social referencing. *ROMAN 2005*, 2005.
- [21] J. Wang, E. Sung, and R. Venkateswarlu. Estimating the eye gaze from one eye. *Computer Vision and Image Understanding*, 98(1):83-103, 2004.
- [22] H. Yamazoe, A. Utsumi, and S. Abe. Remote gaze direction estimation with a single camera based on facial-feature tracking. *Asian Conference on Computer Vision (ACCV2007) Demo Session*, 2007, to appear.
- [23] T. Yonezawa, N. Suzuki, S. Abe, K. Mase, and K. Kogure. Crossmodal coordination of expressive strength between voice and gesture for personalized media. *Proc. of ICM106*, pages 43-50, 2006.
- [24] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe. Gazecoppet: Hierarchical gaze-communication in ambient space. In *SIGGRAPH Proceedings DVD-ROM*, 2007.
- [25] Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, and T. Miyamoto. The effects of responsive eye movement and blinking behavior in a communication robot. In *Proc. IROS2006*, pages 4564-4569, 2006.
- [26] 山添大丈, 内海章, 米澤朋子, 安部伸治. 単眼カメラによる視線検出のための眼球モデルパラメータの推定. *信学技報 IE2007-11*, pages 1-6, 2007.
- [27] T. Ishikawa, S. Baker, I. Matthews and T. Kanade. "Passive driver gaze tracking with active appearance models", *Proc. World Congress on Intell. Transportation Systems* (2004).
- [28] 中野, 岡, 佐藤, 西田: "ユーザの視線に気づく会話エージェント-アテンションの知覚と制御を利用した会話の円滑化-", 第 19 回人工知能学会全国大会, pp. 3B2-08 (2005).
- [29] 小川原, 崎田, 池内: "視線運動からの意図推定に基づいたロボットによる行動支援", *インタラクシオン 2005 情報処理学会シンポジウム*, pp. 103-110 (2005).
- [30] 米澤, 山添, 内海, 安部: "視線検出環境による擬人的媒体の階層的視線コミュニケーション", *映像メディア学会技報*, ME2007-83, pp.1-4, 2007.