

ユーザの視線・発声に対するクロスモーダルウェアネス ～ユーザ状況把握を表すロボット～

Crossmodal Awareness to User's Gaze and Utterance for Expressing Anthropomorphic Attentiveness

米澤朋子^{1*} 山添大丈¹ 内海章¹ 安部伸治¹

Tomoko Yonezawa¹ Hirotake Yamazoe¹ Akira Utsumi¹ Shinji Abe¹

¹ ATR 知能ロボティクス研究所

¹ ATR Intelligent Robotics and Communication Laboratories

Abstract: In this paper, we introduce a daily-partner robot, that is aware of the user's situation or behavior by using gaze and utterance detection. Not only the crossmodal-awareness for the appropriate timing of communication but also the robot's expressions of the awareness (*attentiveness*) without any utterance are used for gradual notice of need for their communication. The system is based on both utterance detection and gaze detection of the user. The results of the experiments show the expressions of the robot's intention to speak are effectively recognized during the user's various situations of her/his tasks.

1 はじめに

人と人が過ごす時と同じように、ロボットが日常的に人間と共に過ごすためには、ロボットはユーザの状況や様子に気遣いながらメッセージを伝えたり仕事をしたりする必要がある。全世界的に加速しつつある高齢化社会や核家族化の傾向によって、家庭内での介護・看護やコミュニケーション相手といった人材は、不足してきており、結果として個人の生活を孤立させる傾向が強まりつつある。このような社会やコミュニティでは、独居の高齢者や障害者にとっての頼り場所のない状態であり、介護者の数も不足する。介護者やコミュニケーション相手の役割をある程度代理できるようなパートナー型ロボットが実現すれば、その需要は今後増加すると考えられる。ロボットには、あたかも実空間内で時間を共に過ごしているかのような感覚を与えながら人と人の遠隔コミュニケーションを実現することも期待される。実体を持つロボットの擬人的表現は、人をひきつけ自然なコミュニケーションを実現する上で必要不可欠であり、Picard [1] の Affective Computing の概念をさらに強力に実現する手法であるとも考えられる。

パートナーロボットとの日常生活を仮定すると、ユーザは常にロボットとコミュニケーションしているわけではなく、家事、他者とのコミュニケーションや、テレ

ビの視聴など、何らかの用事をしていることが多いことが想像できる。ロボットが何らかの注意や他者からのメッセージをユーザに伝える必要がある時（これを本稿では Active State と呼ぶ）、もしユーザの状況を考慮せず単純に音声でメッセージを伝えるなどすると、ロボットは邪魔な存在になったり、本来の目的であるメッセージを十分に聞いてもらえなくなる可能性もある。また一方で、例えばロボットがユーザのすべての発話に反応してしまうと、他者のコミュニケーションにおいて混乱をきたす可能性もある。もしくは、ユーザのロボットに対する視線にロボットがいちいち反応し、何らかの声を出すと、視線の通過にすら反応してしまうロボットとなる。

このように、単一のモダリティにロボットが反応する設計では、ユーザの状況を把握できず、ロボットが煩わしい存在となってしまう可能性があり、日常を共にするパートナーとしては不適切である。本稿では、ユーザの状況に応じて話しかけるロボットの行動モデルを提案する [3]。ユーザの視線方向と発話状況を組み合わせユーザの発話対象を推定するクロスモーダル適応型の行動指針を導入し、ユーザの注意がロボットに向けられていない時は、ロボットは、擬人的表現を通じ a) ロボットが何か物言いたげな行動を何も発声せずに示し、b) そのメッセージを伝えるタイミングをはかっているかのように見せかける、**話しかけ意図行動**を控えめに示しながら、ユーザに話しかけるべきタイミングを待つ。つまり、ユーザがロボットに話しかけの許可を出

*連絡先：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@atr.jp

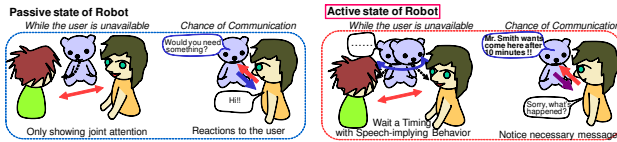


図 1: ロボットの行動設計指針

すまでは、ロボットからは声をかけないことで、ユーザの日常的生活を邪魔しない。ここでは、コミュニケーションにおける視線の役割 [2, 4, 5, etc.] と、実体型もしくは仮想エージェントにおける視線行動の効果 [7, 10, 8, 9, 6, etc.] などに基づいて、アイコンタクト・共同注視を交互に繰り返すことをモデルとし、ロボットがユーザの方向とユーザの注視対象を繰り返し見るという行動を取り入れた。

図 1 に、本稿で提案するクロスモーダル適応型のロボット行動設計を示す。ここでは、ロボットの状態を、I) 待機状態 (ロボット自身から発するメッセージのない状態, *passive state*) と II) ロボットがユーザに対するメッセージを伝えるべき状態 (*active state*) の二種類に設定した。I) の *passive state* では、ロボットはただ同じ空間を共有して存在している擬人的な媒体である。そのため、ユーザの視線方向に受動的に反応しユーザと同じ場所を見るという行動のみを示す。II) *active state* では、ユーザの視線と発話から発話対象や状況を推定し、ほかのタスク中など、ロボットに対するユーザの発話が見られない間は、ロボットは声を出さずに話しかけ意図行動を表し、ユーザの邪魔をせず、かつロボットの暗黙の意図に気づかせるチャンスを作る。一旦ユーザがロボットに対して話しかけたら、ロボットはメッセージをユーザに発声することが許される。

このようなロボットを構成するため、擬人的な媒体としてぬいぐるみロボットを採用し、ユーザの自由な動きに対応するため、遠隔視線方向推定手法 [11] を導入した。そして、ロボットの行動設計において、クロスモーダル適応型のユーザ状況推定を導入すること自体の効果や、その状況推定の結果に基づきユーザの状況をうかがっているかのような行動を見せる話しかけ意図行動の効果について、それぞれユーザの日常タスクと組み合わせた実験を行い、検証した。

2 関連研究

ユーザのマルチモーダル情報を用いてユーザの意図表出を汲み取り、直感的で、自然なインタラクションをもたらすシステムとして、これまでに提案され議論されてきている [12, etc.]. 例えば Schilit ら [13] や Selker ら [14] はヒューマンコンピュータインタラクションにおける文脈適応の手法について言及しており、それに応じ、ユーザの状況を様々なモダリティを通じ認識し、その時々適切な動作をするよう設計された様々なシス

テムも紹介している。Maglio ら [15, 16] は “Attentive User Interface” としてユーザの視線と発話に焦点を当てユーザの状況に敏感なインタフェースについて提案している。Selker [17] は特に、ユーザの注意をユーザの視線を検出することで推し量れるとしている。このような考え方や方法論は、知的な HCI システムのためのモジュールとして検討されてきている。しかし、擬人的な表現における、人と人のコミュニケーションにより近いロボットの行動設計までには導入されていない。Karl ら [18] や Sibert ら [19] はこのようなユーザの視線や発話に基づいたインタフェースの効果について議論をしており、Castellini ら [20] や Minotas ら [21] もデスクトップ環境のインタフェースにユーザの視線と発話を導入している。これらのシステムはユーザの視線と発話を「制御方法」の一環として捉えており、ユーザに対する気遣いといった擬人化した Attentiveness として捉えた例は少ない。

一方、バーチャルエージェント [7] や実体型ロボット [22, 10, 23, etc.] などの擬人的媒体の視線行動は、その擬人的媒体の内部状態をユーザが推定するためのマルチモーダル表現のチャンネルの一つとして着目されてきている。これは人と人の間のコミュニケーションにおいて視線が一つの重要な表現チャンネルとして機能していること [2] に基づいている。実際に、共同注視やその他の視線行動は、表現力のあるモダリティとしてロボットなどの擬人的媒体のデザインに適用されており [8, 6, 24, etc.], 文脈上での視線の社会的機能についての議論されてきた [9, 25] ように、ロボットの視線行動はユーザに気づきをもたらす何らかの情報を伝える手法として重要なキーとなると考えられる。

これらの文脈適応型インタフェースシステムのほとんどは、ユーザに対するシステムの注意自体を表現に用いるものではない。さらに多くのシステムでは、視線を計測するための様々な装置をユーザの体に装着させたり、ユーザが固定位置にいる必要などがあり、視線計測をユーザに意識させないことは難しい。我々は独自の遠隔視線計測技術により、ユーザの行動にあわせた擬人的媒体の意識・意図表現を実現することを狙う。クロスモーダル適応型のコミュニケーションの中でも特にユーザの発話対象を推定することで、ロボットへの話しかけにのみ積極的なリアクションを返すことで、ユーザにとって適切なタイミングで自然なコミュニケーションの段階を踏むことが可能になると考えた。

3 システム構成

図 2 にシステム構成を示す。広範囲のユーザの動きに対応した視線検出を行うための 2448×2048 ピクセルの高解像度カメラ、ユーザの発話を検出するためのショットガン型マイク、擬人的表現の媒体としてのぬいぐる

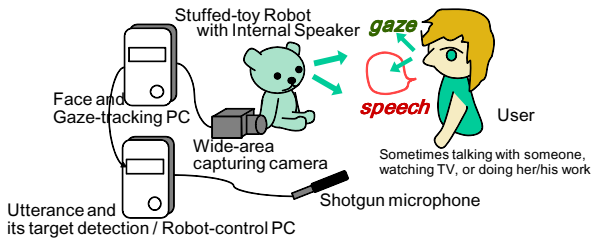


図 2: システム構成図

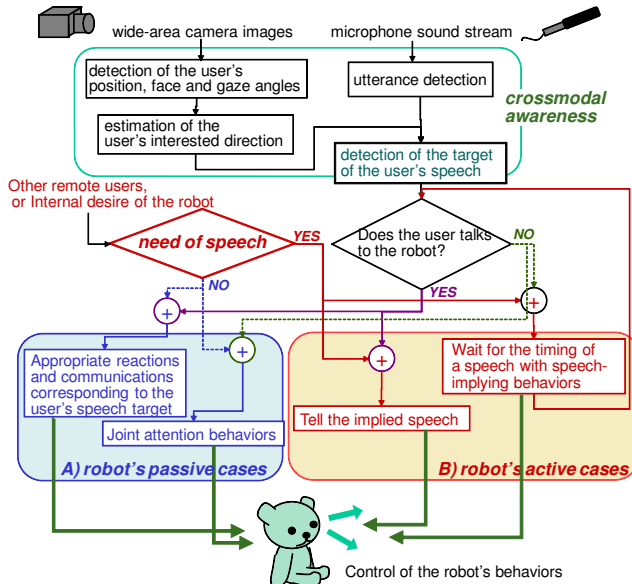


図 3: クロスモーダル適応型システムのフロー

みロボット [26] によりシステムが構成される。

本システム構成においては、話しかけ意図を示す行動として、ロボットの頭部方向がロボットの視線を表すとし、ロボットの頭部方向を、ユーザの方向 (EC) とユーザの注視対象の方向 (JA) へそれぞれ 2.0 秒～0.5 秒毎に交互に向けることとした。この話しかけ意図行動は、およそ 5～6 秒程度の間行う。ユーザが話しかけるまでの時間と、メッセージの緊急性・重要性に応じ、5 秒から 1 秒へ話しかけ意図行動の間隔を短くしていくことができる。

図 3 に提案するロボットの行動モデルのためのフローチャートを示す。カメラ画像に後述する視線検出手法を適用すると同時に、マイクからの音声信号から発話部分を判定する。倍音による判定 [27] も可能であるが、独居状態などのある程度静かな状況であれば発話の判定は、音量からでもある程度可能である。これらの入力信号から、ユーザの発話中の視線対象の比率を計算することで、ユーザの発話対象を判定する。i) 視線対象がロボットである時間の比率が 50% より高いとき、ユーザの発話対象がロボットであると推定し、ii) 10% 以上のときはユーザの発話開始時と発話終了時の視線位置によって判定する。これによりロボットはユーザの忙しさなどの状況に応じた行動を示したり、ロボットへ

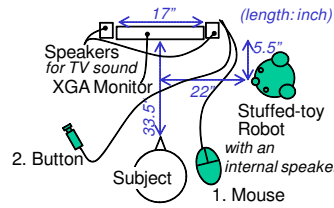


図 4: 実験状況



図 5: 実験時の様子

の話しかけに応じてメッセージを伝えることができる。

図 3 の下部には *passive state* と *active state* の 2 種類のモードについて記述している。これはロボット自身のユーザに対するメッセージ発信の必要性によって決定される。A) *passive state* では、ロボットはユーザと共同注視している行動を示し、ユーザの話しかけには受動的だがリアクションを返す。B) *active state* では、何らかのメッセージを発信するための適切なタイミングとして、ユーザがロボットに対し話しかけた 0.5 秒後を想定し、待ち状態を示すため話しかけ意図動作を見せる。メッセージを伝えるときのロボットは、いきなり音声を伝えないう、まず動作を見せた後、1 秒後に発声するよう設計した。

空間を共有しながら日常を共に過ごすロボットは、ユーザが部屋などの空間内を動き回ることを前提にユーザの視線方向を推定することが必要である。これまでの視線推定システムは様々な装着具 [28, etc.] やユーザに負担をかけるキャリブレーション作業 [29], 推定システムからの距離 [31, 30, etc.] など様々な制約があった。我々の視線推定手法 [11] では、明示的キャリブレーションを必要としない。また、高解像度単眼カメラを、視線推定角度の精度を向上させる [32] のではなく、人間の広範囲での頭部動作に適用できる。

4 システムの評価

本節では、提案するクロスモーダル適応型のロボット行動設計の有効性を示すため、まず、i) システム設計の前提としているロボットの発声のない話しかけ意図表示の効果を確認し、それに基づき、ii) ロボットに対するユーザの話しかけに応じた行動の有効性を試験する。各実験は、ロボットの異なる行動刺激により、ユーザ評価をそれぞれ行い、検証するものとした。

実験状況の設定: 被験者は 17 インチの画面前に座り、マウスを用いてタスクを行う。ぬいぐるみロボットはユーザと画面の間、右位置に置かれた (図 5)。

4.1 話しかけ意図行動に対するユーザの感覚

本稿で述べるロボットの話しかけ意図行動に対する人間の敏感さを検証するため、被験者による主観的なボタン押下を用いて以下の実験を行った。

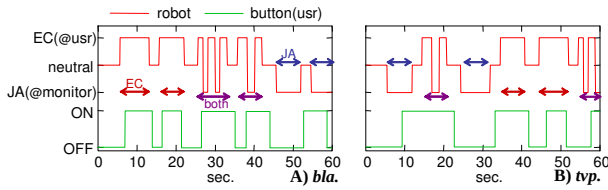


図 6: ボタンの信号の例

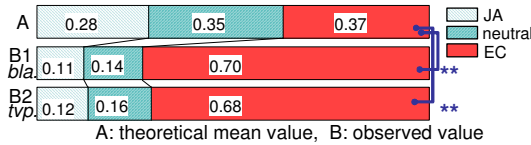


図 7: ボタン押下時間の理論値と実測値

仮説: i) ロボットの被験者への視線 (頭部方向) により, 被験者はロボットの物言いたげな様子を感じる. ii) 被験者は自身のタスクの状況に関わらずロボットの話しかけ意図行動を知覚する.

被験者: 男性 13 名女性 12 名の計 25 名 (21~40 歳).

実験条件: ロボットの視線方向は, 被験者側, モニタ側, もしくはその両者の間での頻繁な視線転換, の 3 パターンを示し, それぞれインターバルをはさみながら 5 秒程度提示される. 結果として, 被験者側は約 22 秒, モニタ側は約 17 秒, 約 21 秒は被験者とモニタの中間的な方向を向く.

実験は全体で 1 分で, その間被験者は a) 黒いモニタ (bla) もしくは b) テレビ番組を見ながら, 「ロボットが物言いたげだ」と感じたときにのみボタンを押下し続ける.

実験結果: ボタン押下信号の例を図 6 に示す.

[A] 押下時間の分析:] ロボットの視線 (頭部) 方向と被験者のタスク状況により押下時間の総合を比較するため 2 要因分散分析を行った結果, タスク状況による有意差がない一方 ($F_{(1,24)} = < .01, p > .99$), ロボットの視線方向による有意差が認められた ($F_{(2,48)} = 211.3, p < .01$). 事後検定として分散分析を行った結果 (図 7 および表 1), 両タスクにおいて, 時間に比例した理論値とは有意に異なる結果を得た.

[B] 押下タイミングの分析:] ボタン押下タイミングがどのような状況下であったかを分析するため, それぞれのロボットの視線方向ごとにタイミングを集計し χ^2 検定を実施した (表 2 参照). その結果, 押下タイミング分析においても bla 条件および tvp 条件ともに理論値とは有意に異なる結果であった. ユーザに対する視線は物言いたげと感じられたという結果であった. このように, ユーザの視聴コンテンツの有無に関わらず, ロボットの話しかけ意図行動が近くされることが示され, 仮説 i) および ii) が確認された.

表 1: ボタン押下時間の検定

cond.	EC		JA		neutral	
	a.(pd.)	T ₍₂₄₎	a.(pd.)	T	a.(pd.)	T
bla.	.70(.37)	8.87*	.11(.28)	-13.1*	.14(.35)	-11.1*
tvp.	.68(.37)	8.22*	.12(.28)	-15.6*	.16(.35)	-8.55*

a.=average, pd.= predicted value, *: $< .01$

表 2: ボタン押下タイミングの χ^2 検定

cond.	total	EC(prd.)	JA(prd.)	neu(prd.)	χ^2	p
bla.	125	94(46.4)	23(35.5)	8(43.1)	82.0	$< .001$
tvp.	139	92(51.5)	28(39.5)	19(48.0)	52.6	$< .001$

4.2 クロスモーダル適応型の行動の効果

実験仮説: I) ロボットの発声を伴わない話しかけ意図行動は, ユーザのタスクを邪魔せず話しかけ意図を知らせることができる. II) ユーザの話しかけを待つロボットの行動は, ユーザのタスクを邪魔しないことでユーザに好印象を与える.

被験者: 男性 13 名女性 13 名の計 26 名 (21~40 歳)

実験条件と提示刺激: 図 8 に 3 種類の実験条件を示す. 条件 *Thru.* (T) では, ロボットは, ユーザの話しかけを待たずにメッセージを伝えてしまう. 条件 *Wait* (W) では, ユーザが話しかけるまで待ち, メッセージを伝えるだけで, 待っている間は話しかけ意図行動は見せない. 条件 *ImPLY* (I) では, ユーザが話しかけるまでの間は話しかけ意図行動を一度見せて待ち, ユーザが話しかけたらメッセージを伝える.

ユーザの状況を変化させて効果を検証するため, 異なるタスクとして, a) 計算問題を解く (*cal*, 能動的タスク) and b) テレビを見る (*tv*, 受動的タスク) の 2 種類を第二刺激として準備した.

実験手順: 被験者は実験毎に各タスクを 2 セット行い, タスクの合間と最後にロボットに対し視線を投げかけて話しかけを行うよう指示された. 各実験の後, 被験者は下の質問項目について 5 段階評価で評価した.

- Q1: ロボットにタスクを邪魔された
- Q2: ロボットは話したげな様子だった
- Q3: ロボットとのコミュニケーションは快適だった
- Q4: ロボットに話しかけるのは好きではない
- Q5: ロボットに愛情を感じた
- Q6: ロボットは活発だった

実験は順序効果を相殺した反復測定により行われた.

実験結果: 各質問項目における主観評価の結果 (mean opinion score, MOS) の平均と標準偏差を, 各条件毎に図 9 に示す. 条件間の効果を測定するため, 2 要因分散分析を実施した (表 3) 後, 3 条件の多重検定を実施した (Sheffe, 表 4). この事後検定では, T-W 間比較はクロスモーダル適応型の効果, W-I 間比較は話しかけ意図行動の効果を示す. なお, T-I 間比較は両者の効果を示している. これらの結果のように, クロス

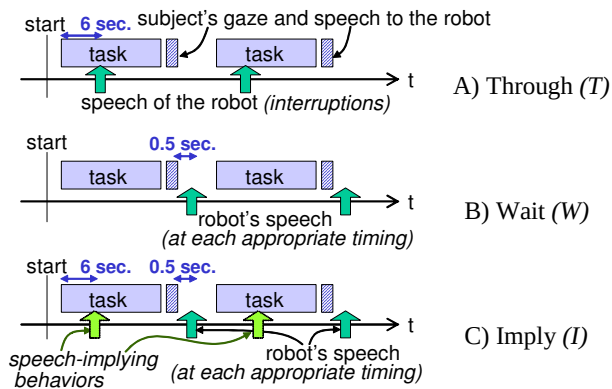


図 8: 実験条件：ロボットの行動

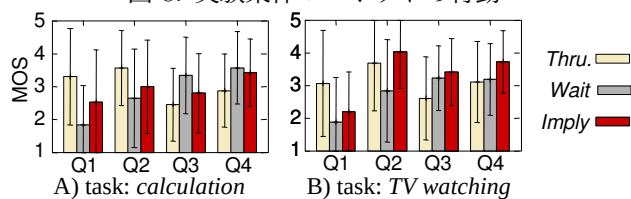


図 9: 主観評価の結果

モーダル適応型行動モデルの導入と、話しかけ意図行動の有意な効果が様々な項目において得られた。

Q1とQ6の事後検定によると、 $T-W$ 間比較においては、タスクの違いに関係なく有意差があった一方、 $W-I$ 間比較には有意差が認められなかった。このことから、話しかけ意図行動はユーザの邪魔をされた印象に影響を与えなかったことがわかる。

Q2とQ3のANOVAの結果より、タスク内容とロボットの行動それぞれの要因から有意差が導かれた。Q5においては、両要因の相互作用の有意傾向が認められた。Q3やQ5などロボットに対するポジティブな印象はタスクの内容により影響を与えられる可能性があることが示唆される。事後検定から、Q2は $T-W$ 、 $W-I$ 間の比較で有意差が認められた。これは、発言したこと自体が物言いたげであると解釈された可能性がある一方、話しかけ意図行動の表現が目的通り作用することも示している。また、Q3とQ5の事後検定では、 $T-W$ 、 $T-I$ 間比較において有意差や有意傾向が認められたものの、 $W-I$ 間比較においては有意差はなかった。これは、 T 条件がユーザの状況を顧みない行動であるため、クロスモーダル適応型の行動設計がポジティブな印象に影響すると考えられる。その一方で、擬人的媒体に対し話しかけることに抵抗を感じるか否かについては、どの検定においても有意差が認められなかった。

上記の分析より、仮説 I) ロボットによるユーザの状況推定に基づいた行動の効果が確認された。また、Q1, Q2の結果より、仮説 II) ロボットの話しかけ意図表示がユーザを邪魔することなく認識される可能性も確認された。また、ユーザのタスクの状況に応じたロボット

表 3: Two-factor ANOVA of the evaluation

		Q1	Q2	Q3	Q4	Q5	Q6
Bhb	$F_{(2,50)}$	13.53	6.45	14.69	0.96	6.01	5.57
	p	<.01	<.01	<.01	0.39	<.01	<.01
Tsk	$F_{(1,25)}$	1.59	7.56	3.57	2.28	0.14	0.48
	p	0.22	0.01	0.07*	0.14	0.71	0.50
Intr.	$F_{(2,50)}$	0.18	1.55	1.32	1.82	2.65	0.09
	p	0.84	0.22	0.28	0.17	0.08*	0.91

underlines: significance, *: significant tendency

Bhb: robot's behaviors, Tsk: user's tasks,

Intr.: interaction of Bhb and Tsk

表 4: Post-hoc tests for robot's behaviors

	Q1	Q2	Q3	Q4	Q5	Q6
$T-W$	<.01	0.01	<.01	0.49	0.08*	<.01
$W-I$	0.25	0.03	0.66	0.75	0.52	0.11
$(T-I)$	0.02	0.92	0.01	0.91	<.01	0.69

の行動に対する印象が変わる傾向がいくつかの項目で確認された。

5 考察

まず、ロボットの話しかけ意図行動に対するユーザの知覚は、視界に入っている状況においては敏感であることが実験から確認された。

次に、ユーザが話しかけるまでロボットが発話を待つことの効果 (仮説 1) や、ユーザに対しロボットが話しかける意図を示すことの効果 (仮説 2) のそれぞれが確認された。本節では、我々のパートナーロボットとして望ましい行動モデル設計について、実験結果に基づき考察する。

まず、2 要因分散分析の結果について考察する。ユーザのタスク内容により有意差が認められた項目もあった。このことより、ユーザにかかっている負荷に応じてロボットの行動に対する印象が変化しと考えられる。これは、ユーザの発話対象以外にも、ユーザの負荷状況に応じてその時々適切なロボットの行動を構成することの必要性を示唆している。

次に、各評価項目の多重比較分析結果について考察する。結果を観察すると、Q1とQ6の結果が同調しており、ロボットが活発である印象と邪魔になる印象において同様の結果が得られたことから、ロボットの活発さや積極性はユーザとの共存のためにある程度抑えられるべきであることが示唆されたと言える。

また、Q3やQ5の $T-W$ 間比較において有意差が認められたことより、ロボットにおけるクロスモーダル適応型の行動設計がロボットに対するポジティブな印象に影響を与えることがわかった。しかしこれらの項目では、ANOVAの結果より、タスク内容によっても影響されることが導かれた。このことや、Q1やQ2の結果から、ロボットの話しかけ意図行動はさほどユーザを邪魔しないという結果ではあるものの、ユーザに話しかけ意図を気付かせる効果はあることが示された。

本稿で提案する話しかけ意図行動は、ユーザの視野にロボットが入っていることが前提である。本手法をスムーズに導入するためには、ロボットがユーザの頭

部方向に応じ視野に入る段階も検討していくべきだと考えられる。

6 おわりに

本稿では、人間の快適な日常生活を保障するロボットの行動設計として、クロスモーダルなユーザ状況推定として、ユーザの発声対象を遠隔視線推定技術と音声発話区間検出の組み合わせにより実現し、検出されたユーザの状況に応じて適切な行動を示すロボットの行動モデルを提案した。この行動モデルでは、i) ロボットが何らかのメッセージを持つとき、ロボットは無言でユーザーユーザの視線対象間で視線をスイッチしながら話し掛け意図を示し、ユーザのロボットに対する話し掛けを待つ。ii) ユーザがロボットに対し話し掛けるとき、ロボットはようやくメッセージを音声で伝えることができる。被験者実験の結果より、話しかけ意図表示を用いることにより周辺視にあるロボットの意図を感じさせるという効果、および無言での意図表示によりユーザはタスクを邪魔されたと感じないという効果がそれぞれ確認された。

今後は、他者との関係やユーザの興味対象などのユーザの状況を正確に把握することで、話しかけまでのステップをより適切に表すロボットを実現していきたい。

謝辞

本研究の一部は科研費 20700106 の助成を受け推進した。研究に協力して下さった ATR の皆様に感謝する。

参考文献

- [1] R.W. Picard, Affective computing, *MIT Press*, 1997.
- [2] A. Kendon. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22–63, 1967.
- [3] Tomoko Yonezawa, Hirotake Yamazoe, Akira Utsumi, and Shinji Abe, “Evaluating Crossmodal Awareness of Daily-partner Robot to User’s Behaviors with Gaze and Utterance Detection,” *CASEMANS2009*, pp. 1–8, 2009.
- [4] C. Moore, P. J. Dunham, and P. Dunham. *Joint Attention: Its Origins and Role in Development*. Lawrence Erlbaum, 1995.
- [5] C. Peters, C. Pelachaud, E. Bevacqua, M. Mancini, and I. Poggi, A Model of Attention and Interest Using Gaze Behavior, *IVA2005*, pp. 229–240, 2005.
- [6] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe. Gaze-communicative behavior of stuffed-toy robot with joint attention and eye contact based on ambient gaze-tracking, *ICMI2007*, pp. 140–145, 2007.
- [7] 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*, vol. 1, pp. 41–49, 2002.
- [8] 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*, pp. 157–164, 2002.
- [9] B. Mutlu. The design of gaze behavior for embodied social interfaces, *CHI’08 extended abstracts*, pp. 2661–2664, 2008.
- [10] 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*, pp. 4564–4569, 2006.
- [11] H. Yamazoe, A. Utsumi, T. Yonezawa, and S. Abe. Remote Gaze Estimation with a Single Camera Based on Facial-Feature Tracking without Special Calibration Actions, *ETRA2008*, pp.245–250, 2008.
- [12] D. B. Koons, C.J. Sparrell, and K.R. Thorisson. Integrating Simultaneous Input from Speech, Gaze, and Hand Gestures, *Intelligent multimedia interfaces, MIT Press*, pp. 257–276, 1993.
- [13] B. Schilit, N. Adams, and R. Want. Context-aware computing applications, *WMCSA’94*, pp. 89–101, 1994.
- [14] T. Selker and W. Burleson. Context-Aware Design and Interaction in Computer Systems, *Workshop on Software Engineering for Wearable and Pervasive Computing*, 2000.
- [15] P.P. Maglio and C.S. Campbell, Gaze and Speech in Attentive User Interfaces, *ICMI2000*, pp.1–7, 2000.
- [16] P.P. Maglio, T. Matlock, C.S. Campbell, S. Zhai, and B.A. Smith. Attentive agents, *Communications of the ACM*, 46(3), pp. 47–51, 2003.
- [17] T. Selker. Visual attentive interfaces, *BT Technology Journal*, 22(4), pp. 146–150, 2004.
- [18] L. Karl, M. Pettet, and B. Shneiderman, Speech-Activated versus Mouse-Activated Commands for Word Processing Applications: An Empirical Evaluation *International Journal on Man-Machine Studies*, 1993.
- [19] L.E. Sibert and R. Jacob. Evaluation of eye gaze interaction, *CHI2000*, pp. 281–288, 2000.
- [20] E. Castellina, F. Corno and P. Pellegrino. Integrated Speech and Gaze Control for Realistic Desktop Environments, *ETRA2008*, pp.79–82, 2008.
- [21] D. Miniotas, O. Spakov, I. Tugoy, and I.S. MacKenzie. Speech-Augmented Eye Gaze Interaction with Small Closely Spaced Targets, *ETRA2006*, pp. 67–72, 2006.
- [22] 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*, pp. 512–517, 2001.
- [23] I. Haritaoglu, A. Cozzi, D. Kons, M. Flikner, D. Zotkin, R. Duraiswami, and Y. Yacoob. Attentive toys, *ICME2001*, pp. 1124–1127, 2001.
- [24] T. Yonezawa, H. Yamazoe, A. Utsumi, and S. Abe. GazeRoboard: Gaze-communicative Guide System in Daily Life on Stuffed-toy Robot with Interactive Display Board, *IROS2008*, pp. 1204–1209, 2008.
- [25] C.L. Bethel and R.R. Murphy Affective expression in appearance constrained robots *Proc. of ACM SIGCHI/SIGART conference on Human-robot interaction*, pp. 327–328, 2006.
- [26] D. Sekiguchi, M. Inami, and S. Tachi. RobotPHONE: RUI for interpersonal communication. In *CHI2001 Extended Abstracts*, pp. 277–278, 2001.
- [27] S. Basu, B. Clarkson, and A. Pentland. Smart Headphones: Enhancing Auditory Awareness through Robust Speech Detection and Source Localization, *ICASSP01*, Vol.5, pp. 3361–3364, 2001.
- [28] Applied Science Laboratories. Mobile Eye: Lightweight Tetherless Eye Tracking, <http://www.a-s-1.com/products/mobileeye.htm>
- [29] R. Newman, Y. Matsumoto, S. Rougeaux, and A. Zelinsky. Real-time stereo tracking for head pose and gaze estimation In *Proc. Int. Conf. Automatic Face and Gesture Recognition*, (FG2000), pp. 122–128, 2000.
- [30] T. Ohno, N. Mukawa, and S. Kawato. Just blink your eyes: A head-free gaze tracking system. In *Proc. CHI2003*, pp. 950–951, 2003.
- [31] C.H. Morimoto and M.R.M. Mimica. Eye gaze tracking techniques for interactive applications, *Computer Vision and Image Understanding*, 98(1), pp. 4–24, 2005.
- [32] J. Wang, E. Sung, and R. Venkateswarlu. Estimating the eye gaze from one eye, *Computer Vision and Image Understanding*, 98(1), Pages 83–103, 2005.