陪席者として人間—人間コミュニケーションへ参加するアンドロイドの動作生成

Android's Nonverbal Participation for a Human-human Communication

○学 横山 晋 (大阪大) 学 力石 武信 (大阪大) 准 中村 泰 (大阪大)
正 松本 吉央 (産総研) 正 石黒 浩 (大阪大)

Shin Yokoyama, Osaka University
Takenobu Chikaraishi, Osaka University
Yutaka Nakamura, Osaka University
Yoshio Matumoto, National Institute of Advanced Industrial Science and Technology
Hiroshi Ishiguro, Osaka University

Appearance and motion of the robot have to be anthropomorphic, because when humans socially interact with robot, the robot must pretend to have intention. However, when the appearances of the robots closely resemble to real ones, humans expect to closely resemble motions from them. Such motions are very difficult to create, consequently the appearance inhibit natural interaction. Especially for an android, it is difficult to make motions, as motions are required to present sufficient intelligence equivalent to human. Then, we use following two approaches. (1) Equipped social interaction model for generating motion based on with sensing ability, (2) Participate human-human communication as a bystander avoiding gap of sensing ability between human and robot. In this paper, we propose a social interaction model for a bystanding android by observation of human. Finally, we report a result of an impression evaluation to verify the validity of our approaches.

Key word: Words: Android, Humanoid Robot, Human-Robot interaction, trilateral communication

1. 緒論

近年、日常生活で人間とコミュニケーションを行うロボットが期待され、様々なコミュニケーションロボットの研究開発が行われている[1],[2]. これらの多くは、人間や動物がデフォルメされた外観をもつ. 人間はロボットと社会的相互作用を行う場合、ロボットを擬人化するため、意志を持ったものとしての外観および動作が必要であるからである[3]. 我々は、可能な限り外観を人間に似せたヒューマノイドロボットであるアンドロイド(Fig.1)を用い、外観と動作に関する人間の社会的反応に関する研究を行って来た[4],[5],[6]. Shimada et al.は、アンドロイドと、機械的外観を持ったヒューマノイドロボットとの相互作用において、被験者の無自覚的な視線行動を比較し、アンドロイドに対しては、対人状況の視線行動と似ることを実験により示した[5]. すなわち、コミュニケーションロボットが人間に酷似した外観を持つ事は、人間に対する社会的反応と同等の反応を引き出し、人間とのコミュニケーション相手として重要な要素となることが示唆されている.

しかしながら、多くのコミュニケーションロボットは、実際の人間や動物に酷似した外観を持っていない. なぜなら、実際の動物や人間に酷似した動作を作ることが困難だからである. とりわけ、人間は外観と動作の落差に対して非常に敏感であり、人間に近い外観を持つロボットの場合、人間に近い知能を感じさせる動作を呈示しなければ、相互作用に大きな障害となる[7],[8]. すなわち、人間に酷似した外観をもつアンドロイドは、人間と同等の能力を感じさせる動作を行わせなければ、人間との自然な相互作用を行えない. 例えば、アンドロイドと人間が会話をを行った場合、人間と

Fig. 1 Android Robot ReplieeQ2.

同等の自然言語処理が行われる必要があり、既存の技術での実現は難しい.

上のような問題を回避しつつ、人間社会への影響を持つロボットの開発のために、本研究では、人間による3者コミュニケーションのうち、1者をアンドロイドに置き換えた実験を行う. Shimada et al.は、3者コミュニケーションにおいて、視線行動の違いによりアンドロイドへの好悪感情を変化させることができることを示した[9]. また、我々は人間と人間のコミュニケーションに第三者としてアンドロイドを陪席させた

実験を行い、被験者の表情などの非言語的動作を操作者が認識し、アンドロイドの動作を選択する事により、アンドロイドが直接発話することなしにコミュニケーションへの印象を操作できる事を示した[10].

一方、我々は、センサ群を用いて人間の行動を認識した結果からアンドロイドの対人行動モデルを作成し、これに基づき動作をさせることにより、アンドロイドの人間らしさが向上することを示した[11]. これらの研究は、アンドロイドが次の2点を備える事により、人間-人間コミュニケーションに対し影響を及ぼす自律システムを構築できることを示唆する.

1. センサを用いた人間行動認識能力を備えること.
2. 認識結果から動作を生成する陪席者としての対人行動モデルを持つこと.

これらのうち、1)人間行動認識は、既存の技術で実現可能である([12],[13]). 本稿では、2)対人行動モデルとして、アンドロイドが陪席者として人間-人間コミュニケーションに参加している、最も基本的な印象を形成する動作、すなわち、人間と人間の会話を陪席者として自然に聞いている印象を与える動作の生成に注目した. 実際の人間が陪席者として、人間と人間の会話に参加している場面を観察することにより生成したアンドロイドの動作に対して、被験者実験によりその印象評価を行った.

2. 陪席者としての動作生成

2.1 視線遷移モデルの作成

陪席するアンドロイドが他の2者の会話を聞いて理解している印象形成可能な動作を生成するモデルの構築を行う. 本研究では、実際の人間を他の2者の会話に陪席する状況に置き、非言語コミュニケーションである視線の動きに注目する. 現在の音声認識技術では、正確な認識が困難であること、また、文脈に応じた発話が困難(チューリングテストに合格する事は実現されていない)であることから、言語的な応答を自動で生成する事が困難である. しかし、非言語的な振る舞いにおいては、明確な一対一対応が存在しないため、偶然、不適切な応答を一時的に行ったとしても、人間が違和感を感じにくいと考えられるからである.

被験者を2人の実験者による会話に陪席させて、観察を行う. 被験者は、人間と人間の会話に陪席するアンドロイドと同様の位置に置かれる. 実験者Aおよび実験者Bは、被験者の前で会話を続け、被験者の視線行動を観察する(Fig.2). 実験者は多くの技術用語を用いることにより、被験者が会話の内容を完全には理解できないようにし、発言をしにくい状況におく. しかしながら、会話の話題は被験者についてであり、被験者が会話に関係している状況を維持する. これは、会話に対して発話することなしに、会話に参加し続けるアンドロイドの動作を想定しているからである. この時の被験者の動作を、被験者から気づきにくい場所に配置

Fig. 2 Gaze direction of human bystander.

されたカメラにより約5分間撮影する.

アンドロイドの待機動作を生成する際、多数の観察結果から動作を生成する必要はないため、実験者Aと実験者Bの会話に興味を持ち、参加をしている印象を最も強く与える被験者を解析対象とする. まず、被験者の視線方向を、実験者Aの頭部方向と身体方向、実験者Bの頭部方向と身体方向、実験者Aと実験者Bの中間地点の頭と同じ高さ方向、身体と同じ高さ方向、身体より下方向、および、それ以外の方向の8つに分類する(Fig. 2). 観察の結果、被験者の視線の方向、視線の滞留時間が、発話者に依存して変化する事がわかった. 実験者A及び実験者Bの顔方向に視線を向ける時間が最も長く(Fig. 3(a), 4(a)), 全体の約60%の時間、両者の顔に視線を向けていることが分かる. また、実験者A及び実験者Bの顔方向に視線を向ける回数が最も多い(Fig. 3(b), 4(b)). よって陪席者は、時間、回数ともに、会話者の顔方向に最もよく視線を向ける傾向にあることが分かる. すなわち、被験者の視線は、頻度、滞留時間共に、発話者の顔方向に向く傾向が大きい、必ずしも常に発話者の方を見る訳ではなく、その他の方向を見ている事もある. また、まばたき、呼吸などの生理的な動作は、陪席者として会話に参加している印象を形成させるためには大きな意味を持たないと考えられるため、陪席者として重要な意味を持つと考えられる視線を遷移させる動作を生成するのに必要な、各方向に視線を向ける頻度と、その視線の1回あたりの停留時間を確率的に選択するモデルを作成する. 現在の話者、選択された視線方向、その方向への視線の滞留時間、および、次に向ける視線遷移先を用いて、視線遷移マトリクスを作成する(Fig. 5). この視線遷移マトリクスからランダムに行を選択することにより、視線の遷移席を決定し、動作を生成する.

2.2 全身動作の作成

視線方向が提案モデルにより決定されると、アンドロイドの各自由度に対する指令値を計算する. ここでアンドロイドが行う動作は、物理的なタスクをこなす

Fig. 3 Gaze retention and frequency (when utterer is experimenter A).

Fig. 4 Gaze retention and frequency (when utterer is experimenter B).

utterer	Present posture	Retention time	Following posture
A	A face	10	B face
A	A face	5	middle high
A	A face	7	A low
B	A low	22	B low
B	A low	13	B face

Fig. 5 Gaze transition matrix.

ような物ではなく、単に振り返る動作であるため、便宜的に各姿勢を3次補間で近似する。ただし、実際の人間の眼球および頭部の動作は、前庭動眼反射という生理的反射により支配されることが知られており、CGエージェントにおいても、その動作が実現されている[14]。本稿においても同様の方法で、眼球頭部動作を生成した。また、視線方向に変化がない期間は、完全に静止し不自然な印象を与える事を回避するため、下に述べるPerlinノイズ[15]を指令値に加え瞬き呼吸といった生理的動作を実現している。

3. 被験者による印象評価実験

3.1 実験手順および条件

ここでは、提案手法により生成された動作と従来手法により生成された動作の比較を行う。生成された動作を行うアンドロイドの前で、被験者は実験者と会話を数分間行い、アンドロイドによる自分たちの会話への理解度を評価する。これは、被験者にとって、陪席者が自分たちの会話に参加しているかどうかの判断は、最低限、自分たちの会話を理解しているかどうかだと

考えられるからである。

被験者には「アンドロイドはある音声認識技術により、人間と人間の会話を理解して動作をします。この実験は、アンドロイドがどの程度会話を理解できるかの評価試験です」と教示をする。次に、実験者は被験者と約2分間会話を行う。その会話の隣でアンドロイドは次の3種類の動作を行う。被験者は、下のいずれかの動作を行うアンドロイドの隣で、実験者と会話をした後、別室に移動し、アンドロイドがどの程度会話を理解していたかを5段階に評価する。この手順を3回繰り返す。評価する。なお、呈示される順序による印象の変化を打ち消すために、被験者は6グループに分けられ、それぞれのグループはそれぞれ別の順序で、アンドロイドの動作を呈示される。被験者の年齢は19~24歳(平均:22.1, 分散:0.81), 被験者数22人。全員、アンドロイドを見た経験はない。

提案手法動作:

提案手法により作成された視線遷移モデルを元に生成された動作。操作者は、実験者と被験者のいずれかが発話しているかだけを判断し、コンピュータに入力する。コンピュータはモデルに基づき視線の遷移先と決定し、動作を生成する。

ハンドメイド動作:

事前にハンドメイドにより作成された動作。これまでのアンドロイドの動作の多くは、ハンドメイドにより作成されている。これは、アンドロイドの動作の自然さが、人間の主観によるものであり、何が自然動作で何が自然でないかといった事に対する調査や説明が、十分になされていない事から、現時点での最も自然な動作生成手法だと考えられている。

Perlinノイズ動作:

CGによる擬人化エージェントに、自然な存在感をあたえる手法としてPerlinノイズを用いる方法が知られている[15]。これは、人間動作のモデル化や観察といった過程を経ずに、簡単に人間の生理的動作を実現する方法である。アンドロイドの動作においても、瞬き、呼吸といった小さな動作は、Perlinノイズを用いて生成されている。

3.2 結果

それぞれの動作に対して得られた点数の平均値に、分散分析、および、多重比較を行った。その結果を(Fig. 6)に示す。全ての動作間で得た点数に有意差があり、話を聞いている印象は、提案手法、ハンドメイド手法、Perlinノイズ手法の順に高い事がわかった。よって、提案手法による動作生成は、他の従来手法に比べて、最も陪席者の動作として優れていると考えられる。

4. 結論

本稿では、陪席者として、人間-人間コミュニケーションに参加するアンドロイドの動作生成手法について述べた。実際に、人間観察から視線遷移モデルを作成した。そのモデルに基づいた動作の有効性を、会話の理

Fig. 6 Average score for the android's comprehension.

解度という尺度により印象評価を行い示した。

一方、本稿では、動作モデルの有用性より前に、視線を遷移させる動作そのものが、会話を聞いている印象形成する可能性を十分排除していない。Perlinノイズ動作が視線遷移動作を含まず、他の動作より聞いている印象が少ない事からも、この可能性はある。しかしながら、ハンドメイド動作が、視線を遷移させる動作を含んでいること、提案動作がハンドメイド動作より高得点を得ている事、自由回答では、会話と視線のタイミングやアイコンタクトに関する不満が多くあったことから、視線を遷移させる動作そのものが会話に関する印象を向上させる可能性はあるが、タイミングや頻度といった要素が、更に会話を聞いている印象を向上させるものであると考えている。よって、更なる比較対象として、WOZによる話者特定を行わず、ランダムなタイミングにより視線を動作させる場合と比較することも考えられる。また、本稿では、陪席者が人間に酷似した外観を持たない場合を考えていない。すなわち、機械的外観を持ったロボットが提案動作を用い動作をした場合、どのような印象形成がなされるかについては、考慮していない。しかしながら、人間にとって最も重要なコミュニケーション対象は人間であり、コミュニケーション対象として必要な要素を全て備えていると考えられる。よって、アンドロイドの外観的要素と他のロボットとの外観的要素の差分と、コミュニケーションへ参加している印象形成への関係は、将来の課題である。

参考文献

- [1] T. Kanda et al. Development of robovie as a platform for everyday-robot research. *Electronics and Communications in Japan*, 87(4), 2004.
- [2] S. Kajita. Research of biped robot and humanoid robotics project (HRP) in Japan. In *The Fourth International Conference on Machine Automation (ICMA'02)*, pages 1–8, 2002.
- [3] B. R. Duffy. Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3-4):177–190, 2003.
- [4] H. Ishiguro. Android science - toward a new cross-interdisciplinary framework. In *Proceedings of the International Symposium of Robotics Research*, 2005.
- [5] M. Shimada et al. Evaluation of android using unconscious recognition. In *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, pages 157–162, 2006.
- [6] M. Noma et al. Composition and evaluation of the humanlike motions of an android. In *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, pages 163–168, 2006.
- [7] S. Yamada et al. Mutual adaptation and adaptation gap between a human and an agent (special issue HAI: New trends in human agent interaction). *Journal of Japanese Society for Artificial Intelligence*, 21(6):648–653, 20061101.
- [8] M. Mori. On the uncanny valley. In *Proc. of the Humanoids-2005 Workshop: Views on the Uncanny Valley*, 2005.
- [9] M. Shimada et al. Social balancing effect of eye contact. In *18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 1119–1124, Sep.-Oct. 2009.
- [10] E. Takano et al. Psychological effects on interpersonal communication by bystander android using motions based on humanlike needs. In *the Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, pages 3721–3726, October 11-15 2009.
- [11] T. Chikaraishi et al. Development of an Android System Integrated with Sensor Networks. In *the Proceedings of the 2008 IEEE International Conference on Intelligent Robots and Systems (IROS 2008)*, pp.326-333, September, 2008.
- [12] T. Nishida et al. Minimally privacy-violative human location sensor by ultrasonic radar embedded on ceiling. In *Proceedings of the Third IEEE International Conference on Sensors (SENSORS2004)*, pages 433–436, 2004.
- [13] H. Ishiguro. Distributed vision system: A perceptual information infrastructure for robot navigation. *International Joint Conference on Artificial Intelligence (IJCAI-97)*, pages 36–41, 1997.
- [14] S. Masuko et al. Generating head-eye movement for virtual actor. *Systems and Computers in Japan*, 37(12):33–44, 2006.
- [15] K. Perlin. Real time responsive animation with personality. *IEEE trans. Visualization and Computer Graphics*, 1(1):5–15, 1995.