

行動と状況の認識に基づく人間とロボットのコミュニケーション

Human-Robot Communication Based on the Recognition of Actions and Situations

プロジェクト代表者： 久野義徳 (大学院理工学研究科・教授)

Yoshinori Kuno

Professor, Graduate School of Science and Engineering

1. はじめに

人間同士の会話では視覚で相手の行動や周囲状況を認識していることにより、簡単な発話で相互の意図が理解できる。例えば、テーブルの上の本を取ってもらいたいとき、顔をその本の方に向けて「それ取って」といきなり言っても、相手に通じることが多い。これは、顔を本の方に向けて、本を見ているという行動が、相手に視覚で認識され、「それ」がテーブルの上の本を指すと相手に理解されるからである。本プロジェクトではロボットにも人間とのコミュニケーションの際に、この能力を持たせる方法を実現することを目指して研究を行った。また、人間とロボットのコミュニケーションでは人間もロボットの行動を見て、そこから情報を得ると考えられる。したがって、人間の視覚に適切な情報を与える行動をロボットが行うことについても検討を行った。ただし、どちらについても人間のあらゆるコミュニケーションの場面に対して対応できるように始めから考えるのは困難である。そこで、ロボットが人を見ることについては、頼んだものを取ってきてくれるなどの軽作業をしてくれる介護ロボットへの利用を想定して、人間の自然な簡略化された依頼を理解することに問題を限定した。また、人間に対して適切な行動を見せることについては、博物館や美術館でのガイドロボットの頭部の動きについて検討した。人にもものを説明するとき、ずっと聞き手の方を見ていたり、ずっとものの方を見ていたりしない。説明者はものの方を見て説明する場合でも、ときどき聞き手の方を振り返って、聞き手の様子を確認しながら、そして聞き手を引き付けるように説明を行う。今回は、このように効果的なロボットの頭部動作を検討した。

今回のプロジェクトは教養学部の社会学が専門の山崎教授との共同研究である。人間とロボットのコミュニケーションを考える前に、人間同士ではどうなっているか、社会学のエスノメソドロジ（人間の行動の方法を調べる分野）の会話分析・相互行為分析の手法で調べる。そして、その知見に基づきロボットを開発する。ロボットが動作するようになったら、そのロボットと人間がコミュニケーションを行う場面を会話分析・相互行為分析で調べ、ロボットの評価を行う。そして、そこで得られた知見からロボットをさらに改良する。プロジェクトの研究は、基本的にはこのように、社会学と工学の研究を循環的に行うことにより進めるアプローチをとった。

2. 簡略化発話の理解

2.1 行動認識に基づく簡略化発話の理解

先に述べたように、人間同士の会話においては、それまでに会話で触れられていない物体に対しても「あれ取って」と言うだけで意味が通じることがある。このような簡略化された発話の理解を一般的に扱うのは困難なので、介護ロボットへの依頼に限って検討した[1]。ロボットへの依頼は「何を」(目的語、対象物体)「どうして欲しい」(述語、動詞)という2つの要素からなると考えられる。そこで、発話の中の対象物体と動詞のそれぞれが以下の3つのどれであるかを判断する。(i) 明確に言われている、例:「赤い本」、「取って」、(ii) 指示語(対象物体の場合)、代動詞(動詞の場合)が使われている、例:「あれ」、「して」、(iii) 省略されている。そして、対象物体

に関して(ii)か(iii)の場合、人間が明確に言わなくてもわかると判断して発話したのは、それが会話の当事者(人間とロボット)の行動に関連しているからだと考える。関連する行動にどのようなものがあるかは、今後、さらに検討する必要はあるが、まず以下の4つの行動を考えた。(1) その物体を指差している、(2) 手で扱っている、(3) その物体の近くにいる、(4) その物体を見ている(視線が向いている、実際にロボットシステムを実現するときは、視線の概略値として顔の向きを用いた)。以上の4つの行動に対して、それに関わっている物体を視覚で検出し、検出された物体を対象物体であるとした。動詞が(ii)、(iii)の場合は、対象物によって、それに対して人間がして欲しい行為は決まっているとして、解釈するようにした。

以上の検討に基づき、簡略化発話を理解できるロボットを開発した。図1にロボットの外観を示す。このロボットは2組のステレオカメラを持っている。このロボットでは下段のステレオカメラは常に人間の方を向き、その視線や指差しの方向を求めている。上段のステレオカメラは下段のステレオカメラで得られた視線あるいは指差しの方向上の物体を検出するのに用いられる。人間の眼球や首は速く動くので、相手の指差している方向などをすぐに見ることができるが、それを機械で実現するのは困難なので、ここでは人間を見るカメラと、物体を検出するカメラの2組のステレオカメラを使うことにした。

このロボットを実際に用いて、対象物体の名前を言わなかったり、「あれ」や「これ」で済ませても、ロボットが人間の依頼を理解できることを確認した。発話理解の例を図2に示す。



図1 簡略化発話を理解するロボット

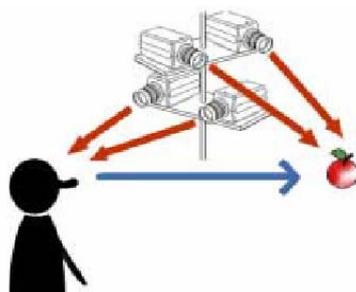


図2 簡略化発話の理解の例
「それ取って」 ロボット近くの物体、その物体を見てもいる

2.2 介護場面における人間の依頼行動の調査

以上のように想定した簡略化発話の理解が視覚を用いて行えることを確認した。ただし、このシステムでは4つの行動に関連する物体を検出すれば簡略化された発話が理解できると仮定している。4つの行動に着目するのは有効であることは示されたが、相手の発話の理解には他にも多くの情報を利用している可能性がある。そこで、実験室に普通の家庭内のような環境を作り、車椅子使用者と介護者を想定した被験者に過ごしてもらい、そこに現れる依頼と、それがどのように理解されるかを調べた。また、老人介護福祉施設にビデオカメラを設置して、被介護者の意図を介護者がどのように理解しているかも調べた。調査は教養学部の山崎教授と共同して行った。「はじめに」で述べた研究の順序と逆になったが、社会学の会話分析・相互行為分析の手法による検討の結果、人間同士の会話では、発話にかなりの簡略化があるのが確認された。特に、介護施設の場合はその傾向が大きいことがわかった。仮定した4つの行動に関する簡略化が実際に出現することが確認されたが、それ以外にも手や他の身体各部の多様な動きを見ることにより、相手の意図を理解していることが観察された。ここでは詳細は省略するが、検討結果は論文として報告してある[2]。今後は、この知見に基づき、ロボットを改良し、人間同士の場合

のように、自然な簡略化されたことばを使って指示のできるロボットを開発していく予定である。

3. ガイドロボットの頭部動作

3.1 人間の説明場面の分析

人間が他人に展示物を説明する様子を2つの場合について調べた。一つは埼玉大学内にある古代朝鮮半島の瓦についての展示の説明である。展示内容の研究者が説明者(ガイド)になって、訪問者(すべて埼玉大学の学生)に説明する様子をビデオカメラで撮影した。15分程度の説明を4回(別の4人)、30分程度の説明を2人ペアの相手に2回(2回は別人)行った。もう一つは、共同研究をしている、はこだて未来大学の山崎晶子講師のところで実施したもので、タイの写真の展示を、写真を撮影した研究者がはこだて未来大学の学生に説明しているところを記録した。30分程度の説明を一人に対して行うものを3回(別の3人)、2人ペアに対して行うものを2回(2回は別人)実施した。

撮影されたビデオから、ガイドの頭部の動きを社会学で使われている会話分析の方法を用いて調べた。特に、頭部を展示の方から訪問者の方に向ける部分を中心に検討した。詳細は[3]に報告したので、ここでは割愛する。分析のまとめとして、ガイドが訪問者の方へ頭部を向けたときの発話等の内容を計数した結果を表1に示す。話の切れ目になる場所は、Sacksたち[4]が発話交代部分と呼んだところと考えられる。この部分で、ガイドは顔を訪問者に向けているが、これは相手が自分の話についてきているか、何か質問がないかを確認するためだと考えられる。その他の場合では、重要な語を言うときに振り向くのは、重要な語であることを認識してもらいたい、それがわかっているか確認したいためだと考えられる。難しい語や数字を言うときも、それが伝わっているか確認したいためだと考えられる。「これ」などの指示語を使うときにも訪問者の方を向くことが多くあったが、これは、指示した方を訪問者が見ているか確認するためだと考えられる。また、このような指示語と同時に進行することが多いが、指差しなどの手のジェスチャをしたときに、振り向くことが多くあった。さらに、これは当然と思われるが、訪問者が質問をしたときには、訪問者の方を向いて質問を聞いた。

3.2 簡易ロボットによる科学技術館での実験

	回数
話の切れ目	61
重要な語を言うとき	14
難しい語や数字を言うとき	6
「これ」などの指示語を使うとき	26
手のジェスチャといっしょに	41
訪問者が質問したとき	12

表1 ガイドが訪問者の方へ頭部を向けた場合とその回数：全部で136回、複数の場合があてはまる場合は、重複して計数

人間の場合の分析結果に基づいて頭部を動かす簡単なガイドロボットを開発して、頭部を動かすことの効果を調べる実験を行った。このロボットは展示品の近くの人間がロボットの方を見ると、その人の顔をビデオカメラの画像から見つけて、その人に近づく。そして、展示品の説明をするようになっている。

このロボットを使って、科学技術館(東京北の丸公園)で実験を行った。図3に実験の様子を示す。磁性流体という磁石の性質をもった液体を用いた芸術作品の展示会の際に、その作品の一つをロボットに説明させた。ロボットは、人間についての分析結果に基づき、話の切れ目や重要な語や指示語を言うときに、訪問者の方へ頭部を向ける(首を振る)ようにした。この頭部動作の効果を調べるために、16人の被験者により実験を行った。16人を8人ずつの2つのグループに分けた。一つをAグループ、もう一つをBグループと呼ぶことにする。Aグループの人には、はじ

めに頭部を訪問者の方へずっと向けたままのロボットにより展示品の説明を受けてもらった。そして、次に頭部を動かしたロボットの説明を受けてもらった。Bグループの人には、この逆の順序で説明を受けてもらった。なお、2回の実験でのロボットの動作の違いなどは被験者には告げなかった。

図4は実験結果をまとめたものである。この図は、実験に参加した個人ごとに、首を振らずに固定したロボットに対して、説明の間にロボットの方に頭部を向けた数を横軸に、説明の適当なところで首を振るロボットに対して、同様にロボットの方を向いた数を縦軸に示したグラフである。○はAグループの人、△はBグループの人に対する結果であることを示す。頭部を動かすロボットに対しては、人間の方も頭部を動かす回数が有意に増えていることがわかる($p < 0.01$)。

回数が増えれば良いとは必ずしもいえないが、回数が多いほうがロボットとのかかわりが強かったということで、ロボットの説明がより訪問者をとらえていたと考えられる。ただし、この実験では頭部を動かさないロボットと比較したので、動かし方が良かったかどうかは断定できない。そこで、現在、ヒューマノイドロボットを用いて、頭部を提案のように動かす場合と、ランダムに動かす場合の比較実験を行っている。

4. おわりに

会話をするときにはことばを使うが、対面のコミュニケーションではことばで表される以外の多くの情報を使うことにより、円滑なコミュニケーションが可能になっている。その情報のうちのおもなものは視覚により得られる情報である。また、このように視覚情報が重要だということは、相手の視覚に適切な情報を与えるような行動を人間はしているということにもなる。このプロジェクトではまだ限られた場面であるが、人間と円滑にコミュニケーションできるロボットを実現するために、ロボットが人間を理解するために視覚情報を利用することと、行動で人間に適切な視覚情報を与えることを検討した。

本研究の一部は総務省戦略的情報通信研究開発推進制度、日本学術振興会人文・社会科学振興プロジェクト研究「日本の文化政策とミュージアムの未来」、科学研究費補助金(14350127)による。

文献

- [1] ザリヤナ モハマド・ハナフィア, 山崎千寿, 中村明生, 久野義徳, 視覚によるサービスロボットのための簡略化発話の理解, 電子情報通信学会論文誌, Vol.J88-D-II, No.3, pp.605-618, 2005.
- [2] 秋谷直矩, 丹羽仁史, 坪田寿夫, 鶴田幸恵, 葛岡英明, 久野義徳, 山崎敬一, 介護ロボット開発に向けた高齢者介護施設における相互行為の社会的分析, 電子情報通信学会論文誌, 採録決定.
- [3] 森山正太, 関口博之, 坪田寿夫, 山崎敬一, 久野義徳, 山崎晶子, 解説時の視線のエスノメソドロジ的分析に基づくガイドロボット, 電子情報通信学会技術研究報告 人工知能と知識処理, Vol.105, No.639, pp.29-34, 2006.



図3 科学技術館での実験

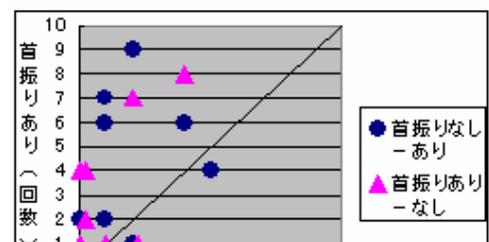


図4 ロボットの首振りがある場合とない場合に対する人間の頭部動作の回数

[4] H. Sacks, E. Schegloff, and G. Jefferson, A simplest systematics for the organization of turn-taking in conversation, *Language*, Vol.50, pp.696-735, 1974.