VRを活用した競技タスクによるロボットの対人誘導スキルの評価

Evaluating the skill of robots guiding humans via a competitive task utilizing VR

水地良明*1 稲邑哲也*1*2 Yoshiaki Mizuchi Tetsunari Inamura

*1国立情報学研究所 National Institute of Informatics *2総合研究大学院大学

formatics SOKENDAI (The Graduate University for Advanced Studies)

We focus on evaluating an instruction ability of robots to appropriately guide human actions. We have proposed a simulation software to evaluate human-robot interaction in daily life environment based on immersive virtual reality (VR). However, difficulties still remain to evaluate the performance of human-robot communication which depends on subjective factors of humans. To develop evaluation criteria of the performance in such a human-robot communication, we designed and carried out a competitive task in which the robot has to generate comprehensible and unambiguous natural language expressions to guide test subjects to achieve several tasks in daily environments utilizing immersive VR. The evaluation results in a robot competition show the usefulness of the proposed approach. Additionally, we demonstrate that the evaluation based on human behavior itself such as changes of head directions is useful for evaluating the performance of instructions by the robot.

1. はじめに

より優れたロボットの対人コミュニケーション能力の構築に 向けた手立てを効率的に探る方法の一つは、RoboCup@Home などのロボット競技会の活用である.実機ロボットの開発に起 因する資金,時間,労力を削減し,効率的な評価プロセスを 実現する方法として, RoboCup Soccer や RoboCup Rescue などのロボット競技会ではシミュレータが活用されている. ロ ボットの対人コミュニケーション能力の評価においても,多様 な条件における試行や再現性の高い評価を比較的容易に実現 するための方法としてシミュレータの活用が望まれる.しか しながら、ロボットと人間のやり取りに重きが置かれている RoboCup@Home では、世界大会における正式なシミュレー ションリーグが存在しない. その要因の一つは, 対人コミュニ ケーションにおける人間の行動をシミュレートすることが困難 であることである. この問題については, 没入型 VR の活用 によって, 複数のロボットの評価を同時並列で実施可能とする ことで解決が図られている [Inamura 18]. 一方で,人間の主 観に依存する対人コミュニケーションの性能評価が困難である 点が課題として残っている.

本研究では、結果が人間の主観に依存する対人コミュニケー ションとして、ロボットによる対人誘導スキルに着目する. 我々 の日常生活では、状況に応じて場所や参照物、手順などを伝 え、相手の行動を誘導する場面は多く存在する. このような 対人誘導スキルは、知能ロボットが人間を相手として協働する 場面や手助けを求める場面における重要な要素の一つである. 対人誘導においてロボットがどれだけ効率的かつ気の利いた指 示を与えられているかを被誘導者となる人間の振る舞いから 評価する方法として、VR内のアバターに没入した被験者の行 動をロボットが誘導する競技タスクを設計した. 本研究の目的 は、競技タスクを通した実証実験により、評価方法の構築に向 けた前段階として、対人コミュニケーションの評価において、 従来の時間に基づく評価だけではなく、頭の向きの変化など人 間の振る舞いそのものに基づく評価の必要性を明らかにするこ とである. また、本研究で構築した競技システムの対人コミュ

連絡先: 水地 良明,国立情報学研究所,〒101-8430 東京都千 代田区-ツ橋 2-1-2, mizuchi@nii.ac.jp



図 1: 競技システムの構成



図 2: ソフトウェア構成

ニケーションの性能評価における有用性を明らかにすることで ある.同競技システムを用いたロボット競技会において実際の 人間の挙動に基づくロボットの対人コミュニケーション能力の 評価を行った結果から,所要時間以外の評価基準の必要性およ び提案したシステムの対人コミュニケーション能力の評価にお ける有用性について議論する.

2. 評価方法

被誘導者となる人間 (被験者) が, VR 内のアバターにログ インし,ロボットの指示に従ってタスクを遂行する競技タス クを設計した.同競技タスクにおけるロボットの役割は,人



図 3: 競技の様子

間の行動を的確に誘導するために,状況に応じた自然言語表 現やジェスチャを生成することである.図1,2に競技システ ムとソフトウェアの構成を示す.競技システムは,SIGVerse Ver.3[Mizuchi 17] プラットフォームの上に構成されている. SIGVerse Ver.3は Unityと ROS を統合した設計となってい るため,競技者は実機ロボットのソフトウェアを流用して仮想 ロボットのソフトウェアを構築できる.

構築した競技システムを用いて,経済産業省および NEDO 主 催の World Robot Summit 2018 の Service Category, Partner Robot Challenge (Virtual Space)*1 を実施した.提案す る競技タスクは、同競技会のルール*2 における "Human Navigation"に該当する. 従来のロボット競技では人間からの要求 をロボットが実行することに重きが置かれていたが、Human Navigation では両者の役割が入れ替わっている.具体的には, 被験者が対象物を掴んで、それを目的地に置くタスクとなって おり,競技者は達成時間に基づいて計算される得点を競う競技 システムは事前に GitHub*3 で公開し,競技者には仮想ロボッ トの制御ソフトウェアのみを実装してもらった. 被験者には, 環境中にどのような物体が存在するか、どのような間取りなの か、対象物が何であるか、目的地がどこであるかといった情報 は与えられない. ロボットには、各試行の開始にそれらの情報 が通知される. また, アバターの視線や何を持っているかなど の振る舞いを随時取得できるようになっている.

図3に競技中のシステムと被験者の様子を示す.競技全体の様子についても、YouTube^{*4}にて公開されている.

3. 評価結果

7 チームを対象としてそれぞれ 28 条件,計 196 回の試行を 行った.16人の被験者を対象として,7.9 時間 (196 試行*1 試 行あたり最大3分間)の振る舞いデータを収録した.この収録 データには、ロボットによる発話だけでなく,曖昧な指示が与 えられた際の迷う素振りや勘違いによる無意味な行動も含め, 被験者の身体的な振る舞いや物体とのインタラクションが含ま れている.また,それらの挙動は VR 空間上で再現可能なレ ベルで記録されているため,競技の採点では対象となっていな い挙動を振り返って分析することも可能となっている.

図 4-6 に点数,達成時間,対象物を把持するまでの時間を



図 5: チームごとの被験者が対象物を目的地に置くまで時間

それぞれ示す. 黒丸が平均, エラーバーは 95%信頼区間を表 す. ここでは, タスクを達成できなかった場合の時間を 180 秒 として平均時間を算出している. ロボットによる指示が分かり にくいほど, 対象物の把持またはタスクの完了に時間を要する 傾向が得られた.

頭の向きの変化は、対象物を見つけるためにどれだけ周囲 を見渡していた度合いと相関があると考えられる.つまり、的 確に対象物の場所を指示できているほど変化が少なくなること が予想される.各試行における頭の向きの変化量 v を、収録 されたデータに基づいて式 (1)を用いて計算した.

$$v = \frac{1}{T_c} \sum_{t=1}^{T_c} \|\mathbf{R}_t \boldsymbol{z} - \mathbf{R}_{t-1} \boldsymbol{z}\|$$
(1)

ここで、tは時刻、Nはデータのサンプル数、 \mathbf{R}_i は頭の姿勢 を表す回転行列, $z = (0,0,1)^{\mathrm{T}}$ は視線方向の単位ベクトル, T_c はタスクの完了までに要した時間である.図7に被験者の 頭の向きの変化を示す.所要時間を指標とした対人コミュニ ケーションの評価は既存研究(例えば, [Striegnitz 11]) にお いて利用されている.しかしながら、図 4-6 の所要時間を指 標とした結果では、上位4チーム(T1, T2, T3, T4)に有 意な差があるとは言えない (いずれの組み合わせにおいても p > 0.42). 一方で, 頭の向きの変化では, T1 と T2 に有意な 差がある (p < 0.01). 実際に, T1 や T5 は被験者の位置や視 線方向を基準とした対象の方向を指示していたのに対し,T2 などでは相対的な方向に関する指示が含まれない傾向があっ た. つまり, 所要時間での評価結果と頭の向きの変化での評価 結果は異なることが明らかとなった、これは、時間だけではな く,様々な身体的社会的な計測に基づいて評価する必要性を示 す結果である.

4. まとめ

本稿では、ロボットの対人コミュニケーションの性能を評価 するためのロボット競技タスクの実施手法について述べた.同 競技では、196 試行、7.9 時間のロボット・環境とのインタラ

^{*1} http://worldrobotsummit.org/en/wrc2018/service/

^{*2} http://worldrobotsummit.org/download/rulebook-en/ rulebook-simulation_league_partner_robot_challnege. pdf

^{*3} https://github.com/PartnerRobotChallengeVirtual/

^{*4} https://bit.ly/2QOjJAZ





図 7: チームごとの被験者の頭の向きの変化

クションを対象とした人間の振る舞いデータを収録した.これ は,提案した人間の挙動データの収集法の有用性を示してい る.また,頭の向きの変化など人間の振る舞いに着目すること で,従来の所要時間を指標とした評価では結果に有意な差が得 られない対象においても,評価結果に有意な差が得られるこ とを明らかにした. これらの結果は、ロボットの対人コミュニ ケーションの性能評価において,タスクの達成に必要な所要時 間だけではなく、様々な身体的社会的な計測に基づく評価の必 要性を示すものである. このような結論をエビデンスに基づい て示すことができる実験システムを提供したという点も本シス テムの寄与の一つである.

本研究では、人間による主観的評価は行っておらず、評価基 準の妥当性は明らかとなっていない. 今後は、VR 上で再現し た対人インタラクションにおける被験者の挙動を第三者に評価 してもらい、競技会のルールにおける評価の妥当性を検証して いく予定である.

謝辞

本研究の成果の一部は、国立研究開発法人新エネルギー・産業 技術総合開発機構 (NEDO) の委託業務,および JST CREST (グラント番号 JPMJCR15E3)の支援を受けたものです.

参考文献

- [Inamura 18] Inamura, T. and Mizuchi, Y.: Competition design to evaluate cognitive functions in human-robot interaction based on immersive VR, RoboCup 2017: Robot World Cup XXI, Lecture Notes in Artificial Intelligence 11175, pp. 84-94 (2018).
- [Mizuchi 17] Mizuchi, Y. and Inamura, T.: Cloud-based Multimodal Human-Robot Interaction Simulator Utilizing ROS and Unity Frameworks, in Proc. of

the IEEE/SICE Int. Symp. on System Integration, pp. 948-955 (2017).

[Striegnitz 11] Striegnitz, K., Denis, A., Gargett, A., Garoufi, K., Koller, A. and Theune, M.: Report on the First NLG Challenge on Generating Instructions in Virtual Environments (GIVE-2.5), in Proc. of the 13th European Workshop on Natural Language Generation, pp. 270–279 (2011).