

身体的社会的対話経験の収集と共有のためのクラウド型 VR プラットフォーム SIGVerse

稲邑 哲也^{1,2,a)} 水地 良明^{1,b)}

概要： 知能ロボットが環境内で適切に自律行動を行うために、膨大な量の行動を行いながらその経験に基づいた機械学習を活用する研究が多くなされてきている。しかしながらロボットが人間からスキルを学んだり、人間と協働してタスクを実行する場合、膨大な量の人間との行動履歴を収集することが困難となる。そこで、人間が VR デバイスでクラウド上の仮想環境に設置された実験フィールドにログインし、アバターを経由してロボットと対話を行うことのできるインタラクションシステムを提案する。このシステムにより、被験者は場所・時間を問わず実験に参加することが可能となり、研究者は全世界を対象として被験者実験を展開することが可能となる。本論文ではそのシステムの構成、および応用事例について概説する。

A Cloud-based VR platform for collecting and sharing social and embodied interaction experience

TETSUNARI INAMURA^{1,2,a)} YOSHIKI MIZUCHI^{1,b)}

Abstract: Recently, intelligent robot systems often use machine learning techniques to acquire autonomous behavior through an immense amount of interaction experience in real world. However, it is difficult to acquire social interaction skill or strategy to collaborate with humans, through such huge interaction experience due to difficulties in asking a human to interact with robots. We thus propose a cloud-based virtual reality system named SIGVerse, which enables general people to take part in experiments in cloud cyber space. Subjects can log in to the system anytime and anywhere; researchers can carry out experiments with global scale. This paper describes the system architecture of the SIGVerse and explains an overview of several applications.

1. はじめに

近年の人工知能の理論や技術の発展とともに、知能ロボットの研究開発は加速度的にスピードアップしている。ロボットが様々な環境内で行動をする戦略を自律的に学習する研究は特に注目を浴びており、例えばマニピュレーターによる物体把持経験を並列に蓄積することで、把持技能の自律学習などが実現されている [1]。ただし、2 ヶ月間に渡る 80 万回という膨大な量の把持をロボットが行う必

要がある。物体を把持するという環境とのインタラクションでは問題にならないとしても、人間とのインタラクションが必要となるタスク・アプリケーションの場合、このような膨大な量のインタラクション経験をどのように収集することは極めて困難である。

この問題を解決するために、従来までに稲邑らは VR 空間上での Human-Robot Interaction の実験・研究を効率的に行うためのシミュレータシステム SIGVerse[2] を開発し、様々な応用展開を行ってきた。SIGVerse は、人間が VR 空間上のアバターにログインし、仮想ロボットとの対話を VR 空間上で効率的に行うことが可能なシミュレータであり、ユーザの身体的動作、ロボットとの社会的対話の経験を蓄積しクラウドを介して研究者間で共有することが可能となっている。

¹ 国立情報学研究所
National Institute of Informatics, Japan

² 総合研究大学院大学
SOKENDAI (The Graduate University for Advanced Studies), Japan

^{a)} inamura@nii.ac.jp

^{b)} mizuchi@nii.ac.jp

従来までの SIGVerse (version 2) は Linux 上にサーバを立ち上げ、Windows 上のクライアントアプリケーションからサーバに接続する形で動作するシステム構成であった。基本的なフレームワークは独自のソフトウェア構成となっており、VR 映像の生成および三次元ヘッドマウントディスプレイ (HMD) への出力、仮想アバターやロボットの制御、動作計測デバイスからのデータ読み込みなど、ほとんどの部分をオリジナルのアプリケーションとして作成する必要があった。近年の VR 技術の発展と廉価版デバイスの安定供給に伴い、非常に数多くの VR 機器が市場に存在する状況となった現在において、各 VR 機器に対応するソフトウェアモジュールを逐次構築するのは非常に効率が悪いものとなっているのが現状である。

そこで本稿では、ユーザ人口が非常に多いことで知られる VR アプリケーション開発のプラットフォームである Unity に着目し、Unity を基盤フレームワークとした新しい SIGVerse (Version 3) を提案する。これにより、デバイスごとのインタフェース作成のコストを低減させ、さらなる実験開発の効率化を図る。

2. SIGVerse のシステム構成

図 1 に、SIGVerse (ver.3) の構成を示す [5]。VR 環境の構築には、Unity^{*1}を採用している。Unity では、Asset として、仮想物体の 3D モデルや VR インタフェースのモジュールなどが共有されてる。また、操作性の高いユーザインタフェースを用いた VR 環境の構築が可能である。これらの豊富な開発資源を流用することで、少ないコストで多様な実験環境を構築することが可能となっている。

被験者は、ヘッドマウントディスプレイ (HMD) などの VR インタフェースを介して VR 環境内のアバターにログインすることができる。仮想ロボットには、ROS 環境とのブリッジ機能により、実機ロボットを制御するためのソフトウェアを再利用することができる。また、独自のブリッジ機能により、データサイズの大きな仮想センサ情報のリアルタイム通信もサポートしている。

以下に SIGVerse ver.3 の基本コンポーネントとその機能をまとめる。

2.1 仮想環境の構築

人間とロボットのインタラクション実験が行われる最も典型的な場面はリビングやキッチンなどの日常生活空間である。Unity では、その構成要素となる家具や日用品などの 3D 形状モデルや、アバターを制御するスクリプトが Assets として豊富に共有されている。さらに、Unity は操作性の高いユーザインタフェースを備えており、それらを活用することで容易に大規模かつ複雑な仮想環境が構築で

きる。

2.2 VR デバイスの利用

VR 空間内でマルチモーダルな対話を行うため、被験者は没入型の VR デバイスを利用してアバターにログインする。近年では、VR ヘッドセット (HTC Vive, Oculus Rift and Touch, FOVE, etc.) や簡易運動計測装置 (Kinect V2, Leap Motion, Perception Neuron, etc.) などの多種の VR インタフェースが開発されている。ほとんどの VR デバイスには Unity 用の Asset が無償で提供されているため、VR デバイスのハードウェア・ソフトウェアのアップデートに対してほとんど労力を割く必要がない。

2.3 サーバ/クライアントシステムの構築

従来までに実現していたサーバクライアント形式によるマルチエージェントシステムは、Unity の基盤機能を用いることで容易に実現可能となる。サーバとクライアントは、アバターやロボット、家具などのオブジェクトから構成される同一のシーンを持つ。ネットワークを経由して、環境内の登録されたオブジェクトの情報を通信することで、それぞれのシーンが同期される。クライアントがログインすると、アバターやロボットなどの動的なオブジェクトがサーバによって VR 環境中に生成される。各クライアントの所有するアバターやロボットの状態は逐次サーバに送信され、それらの情報がブロードキャストされることで、それぞれのアバターやロボットによる対話が実現される。

2.4 Unity と ROS 間のブリッジ機能

近年、多くのロボットがロボットミドルウェアを利用して開発されており、ROS^{*2}は最もスタンダードなミドルウェアとして利用されている。実機ロボットと仮想ロボットの制御プログラムにおける互換性を保つためには、Unity と ROS をブリッジし、制御コマンドやセンサ情報を通信するための機能が必要となる。ROS では、JSON 形式データを利用した WebSocket 通信により、外部プログラムとのデータ通信を可能とするブリッジ機能が提供されている。この JSON を用いて Unity と ROS の通信を可能とする提案 [3][4] もいくつか提案されてきているものの、データサイズの大きい RGB 画像や深度画像などのセンサ情報をリアルタイムに通信することは困難であった。そこで、BSON 形式データを利用した TCP/IP 通信によるブリッジ機能を提案することで、ロボットの制御に必要なリアルタイムな仮想センサフィードバックを実現した。

特にその効果はカメラ画像のリアルタイムな取得において顕著に表れる。640x480 ピクセルの RGB 24bit の画像を Unity 側で生成し、ROS 側に送信する速度を計測した

^{*1} <https://unity3d.com>

^{*2} <http://wiki.ros.org/>

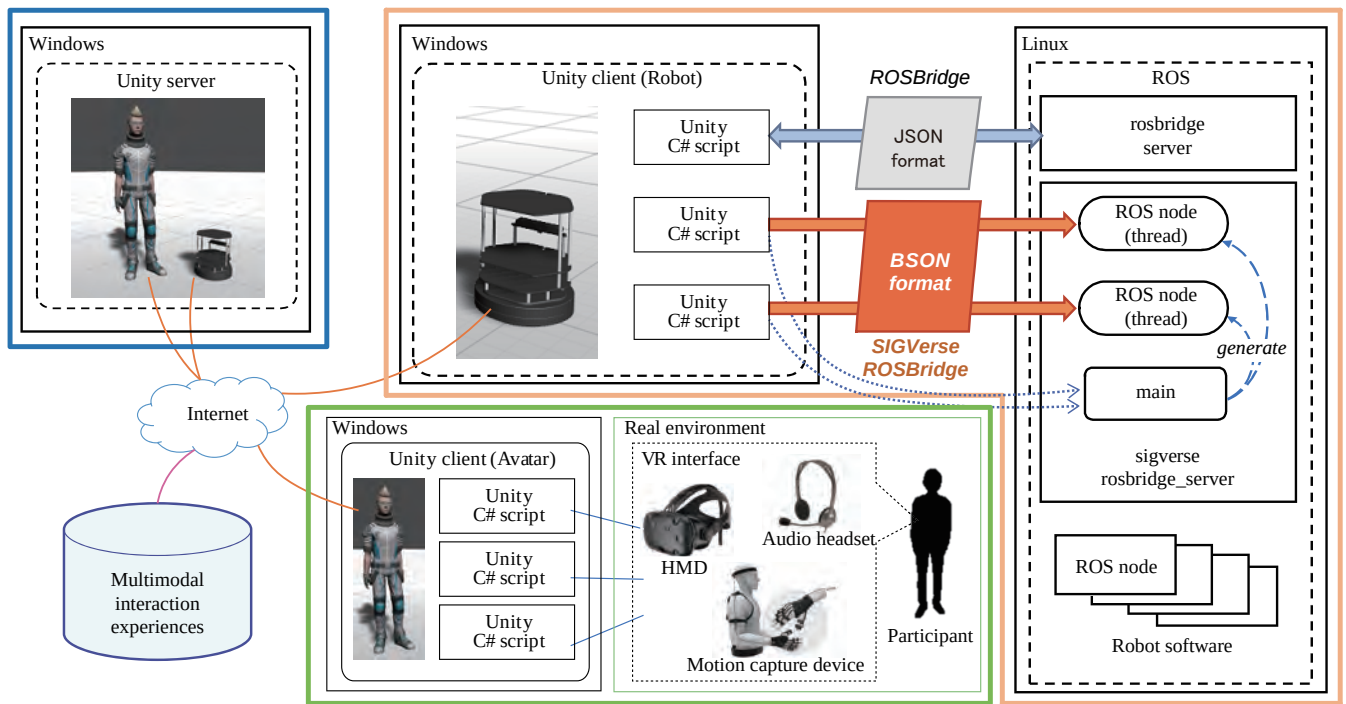


図 1 SIGVerse (ver.3) のシステム構成

ところ、GPU として GeForce GTX TITAN X を搭載する PC での結果は、従来の JSON 形式だと 0.55[fps] であったところ、提案する BSON 形式による通信だと 57.60[fps] の速度が出た [6]。

2.5 マルチモーダルな対話経験の蓄積と共有

マルチモーダルな対話経験データを共有するためには、仮想空間内での対話経験の実演のみでなく、それらを記録・蓄積する機能が重要となる。その実現のためには、エージェントのセンシング情報を記録するのみで不十分であり、マルチモーダルな対話経験を再現できるレベルで記録する必要がある。SIGVerse ver.3 では ver.2 から引き続き MySQL による対話経験の蓄積と共有を行う。これらをネットワーク経由でデータベースに蓄積し、共有することで、クラウドソーシングによる大規模な対話経験の収集と収集データの相互利用が可能となる。

以上の機能を統合することで図 2 や図 3 のような VR 空間上でのロボットとのインタラクション実験が可能となる。

3. ロボット競技会への展開

本章では、SIGVerse システムの応用例の一つとして、知能ロボットの競技会 (RoboCup@Home) の事例について紹介する。

RoboCup @Home 競技会は日常生活空間での Human-Robot Interaction を智能化することを目的とした競技会で、実際のモデルルームの中で活動するサービスロボットと人間が対話を行うことで競技を行ってきた。RoboCup



(a) Collection of behavior in washing task in VR environment



(b) Interactive clean-up task in VR environment

図 2 VR 空間上での人間ロボット間のインタラクション実験の例

には@Home 以外にもサッカーやレスキューなどのリーグが存在し、シミュレータによる競技も実施されてきているが、@Home リーグだけはシミュレータが存在していなかった。その根本的な要因は人間がロボットシミュレータに介入する必要がある、そのような機能を提供するロボットシミュレータが存在していなかったためである。そこで SIGVerse を用いた Human-Robot Interaction の競技会の実施を行うこととなった。

SIGVerse を用いた競技として、ロボットが自然言語発話によって人間に適切な行動を誘導するタスクを導入した [7]。被験者となる人間は未知の仮想環境にアバターを

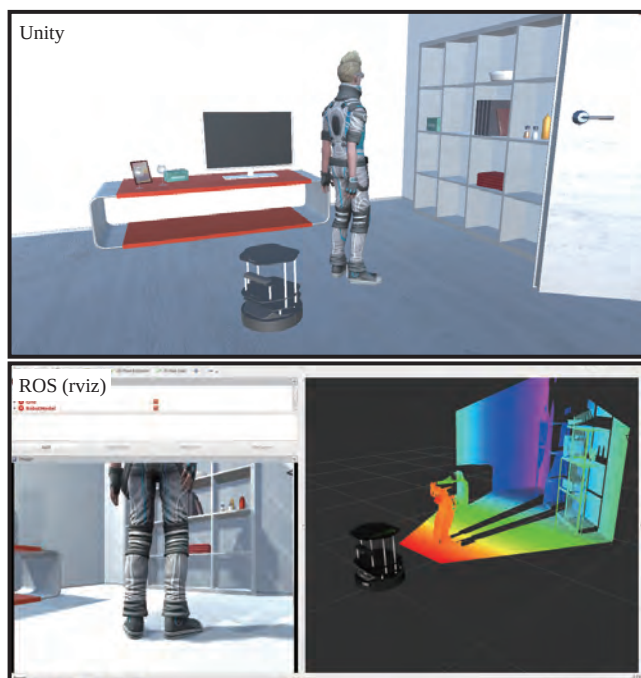


図 3 ROS-Unity 間のブリッジ機能を用いた距離センサ情報のリアルタイム通信機能および人間追従行動への応用。(実際の人間の動きをモーションキャプチャで計測し、VR 上のアバターに動きを反映させた上で VR 上の移動ロボットがそのアバターを追従している。)

介してログインし、ある物体を探し出すことをロボットからリクエストされる。ロボットはターゲットとなる物体が存在する位置情報元に、その場所を人間に説明するための自然言語を生成する。被験者はその自然言語表現を受け取り、仮想環境の中を歩き回りながら、ターゲットの物体を探し出すことになる。物体位置の座標を自然言語に変換するアルゴリズムは、競技に参加するチームが開発し、そのアルゴリズムの善し悪しの評価を被験者の行動で判断する流れとなる。対象物を発見するまでの時間が短ければ、自然言語表現への変換アルゴリズムは優秀であり、逆に長く時間がかかるようであれば、アルゴリズムの質はあまり良く無いと判断できる。

複数回行われる評価において一般の観客に参加してもらうことで、ロボットの発話に対する人間の反応を記録することが可能となり、例えば、自然言語による記述とそれに対応する行動・動作パターンの対データが得られることで、実世界における自然言語処理の研究にデータを提供可能となる。単なる競技会のソフトウェアシステムとして使われるだけでなく、蓄積された経験データが研究開発に有効活用される仕組みを容易に構築できるというメリットがある。

4. まとめ

人間が VR 環境にログインし、ロボットとの社会的身体的対話経験を蓄積・共有するための VR シミュレーションシステム SIGVerse のシステム構成を概説し、いくつか

の応用例について紹介した。この他にも、VR を活用したニューロリハビリテーションへの応用 [8] や、キッチン環境における料理や片付け行動を対象とした人間の動作の観察とモデル化に関する研究 [9],[10] など進めており、ロボットやリハビリに限らず、人間の日常生活空間での行動の分析やモデル化などにも各種の応用を展開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、JST CREST(グラント番号 JPMJCR15E3)、および科研費 26120007 の支援を受けた。

参考文献

- [1] S. Levine, P. Pastor, A. Krizhevsky, D. Quillen: "Learning Hand-Eye Coordination for Robotic Grasping with Deep Learning and Large-Scale Data Collection," arXiv, 1603.02199, 2016.
- [2] T. Inamura *et al.*: "Simulator platform that enables social interaction simulation - SIGVerse: Sociointelligence simulator," in the IEEE/SICE Int'l Symp. on System Integration, pp. 212-217, 2010.
- [3] Y. Hu and W. Meng, ROSUnitySim: development and experimentation of a real-time simulator for multi-unmanned aerial vehicle local planning, SIMULATION: Trans. Society for Modeling and Simulation Int., vol. 92, Iss. 10, pp. 931-944, 2016.
- [4] R. Codd-Downey, P. M. Forooshani, A. Speers, H. Wang, and M. Jenkin, From ROS to unity: Leveraging robot and virtual environment middleware for immersive teleoperation, in IEEE Int. Conf. on Information and Automation, pp. 932-936, 2014.
- [5] 水地, 稲邑: "Unity と ROS を統合したクラウド型マルチモーダル対話経験蓄積プラットフォーム", 第 31 回人工知能学会全国大会論文集, 4G1-OS-14a-4, 2017.
- [6] Y. Mizuchi and T. Inamura: "Cloud-based Multimodal Human-Robot Interaction Simulator Utilizing ROS and Unity Frameworks", in the IEEE/SICE Int'l Symp. on System Integration, 2017.
- [7] T. Inamura and Y. Mizuchi: "Competition design to evaluate cognitive functions in human-robot interaction based on immersive VR," in the Proc. of the Int'l Symp. on RoboCup, 2017.
- [8] T. Inamura, S. Unenaka, S. Shibuya, Y. Ohki, Y. Oouchida, and S. Izumi: "Development of VR platform for cloud-based neurorehabilitation and its application to research on sense of agency and ownership," Advanced Robotics, Vol.31, No.1-2, pp.97-106, 2017.
- [9] T. Bates, K. Amirez-Amaro, T. Inamura and G. Cheng: "On-Line Simultaneous Learning and Recognition of Everyday Activities from Virtual Reality Performances," in the Proc. of Int'l Conf. on Intelligent Robots and Systems, pp. 3510-3515, 2017.
- [10] 横田, 水地, 崔, 稲邑: "クラウドソーシングと没入型 VR を活用した身体動作パターン収集システムの設計," 計測自動制御学会 SI 部門講演会, 2017.