

全方位カメラを用いた実機ロボティック・スワーム を志向した Reality Gap の低減に関する研究

村上 力丸* 伍賀 正典**

Reality gap reduction in real robotic swarms using an omnidirectional camera and incomplete data

Rikimaru MURAKAMI*, Masanori GOKA**

ABSTRACT

In recent years, robots have been widely introduced in society, performing roles in disaster relief and rescue, household cleaning, manufacturing, and more. Computer simulation is commonly used for real robot development, and is particularly prevalent in the development of robotic swarms consisting of many robots. However, there are various differences between simulation and reality, which can compromise the efficacy control algorithms developed in simulated environments when applied to real robots. This is sometimes called the “reality gap”. In this paper, we investigate methods to reduce the reality gap by training robots in a simulated environment using incomplete information. In this way, we aim to make embedded machine learning methods that implement machine learning on real robots more efficient, effective, and practical in the future.

キーワード：ロボティック・スワーム, Reality Gap, 全方位カメラ, 画像処理

Keywords: Robotic Swarms, Reality Gap, Omnidirectional Camera, Image Processing

1. はじめに

近年、ロボットは産業を中心に様々な分野で活躍している。また、従来のアームロボットから警備巡回ロボットや分配・配送ロボットなど複雑なタスクを実行するロボットが登場し、今後ロボットに求められる能力は高くなっている。ロボットが難易度の高いタスクを達成する手法として、ロボットの高性能化が一般的であるが、これに代わる手法として複数台のロボットを用いるマルチロボットがある。ロボティック・スワームはマルチロボットシステムの制御に群知能を適応し、小型かつ低性能であっても台数で能力の低さを補うことができ、単体のロボットと比較して堅牢性の高いシステムになる。また、

コストの面においても、ロボットの高性能化よりも有利とされている[1]。

ロボティック・スワームは、比較的 low performance で単純な機能を持つ個々の自律ロボットがそれぞれ周囲環境を認識し行動決定し、複雑なタスクを実行する。そのため、センサを多数搭載することはこれらの利点を損なうため、少ないセンサで正確な情報を得ることが必須となる。これまで我々が開発してきた実機ロボットには環境認識のために赤外線測距センサと全方位カメラが搭載されており、このうち測距センサを主に用いて障害物等を検知していた[2]。周囲の環境をより詳細に検出するためには測距センサの数を増やすことで解決できるが、搭載数には限界があり、検出範囲も限定的なため死角が存在する。先

行研究では機体に搭載されている全方位カメラを用いて得られる全周画像を用いて壁面検出する実験を行っている[3].

これら実機ロボットを用いて、障害物回避、相互の衝突回避行動を達成させるための実験を行ってきたが、実機を用いた実験には時間、労力等のあらゆる面で大きなコストがかかるために、シミュレーションが多く用いられている。しかしながら、実際には現実の世界を完全に模倣することは不可能であり、様々な特性上の矛盾が生じる。また、シミュレーション環境で機械学習を用いた場合、機械学習アルゴリズムが非現実的な特性に対し適応するために実際には有効ではない結果となる場合があり、このようなシミュレーションと現実とのギャップは

“Reality Gap” と呼ばれ、克服の困難さが指摘されている[4]。本研究では、実機ロボットに機械学習を実装する組み込み機械学習の手法を高効率・高性能・実用的にするために、Reality Gap を低減する手法を検討する。

2. 実機ロボットのシステム

本章では実機ロボットのハードウェア構成と搭載している測距センサ・全方位カメラの仕様を紹介する。

2.1 ハードウェア

Fig.1 に我々が開発し実験に用いてきた実機の全体像を示す。外形 (W×D×H) 120×120×170[mm], 走行用モータとしてサーボモータSG90を改造したものを用いている。カメラ制御・処理用にRaspberry pi Zero W, モータ, センサ処理・制御用にESP32 Develop Moduleを搭載している。これらは無線LANを通じてデータの共有ができる。センサは前方6方向の赤外線測距センサと9軸測距センサが搭載されている。バッテリーを含めた総重量は320gで、3Dプリンタで出力した部品で構成されている[3]。機体上部には任意の色のバンパーを装着することができ、この色によってロボットの個体・グループを識別することができる。Fig.2 にIR セン

サの検出範囲を示す。前方90度(左右に45度)最大で50cm先の物体を検出できる。



Fig. 1 Overview of robot

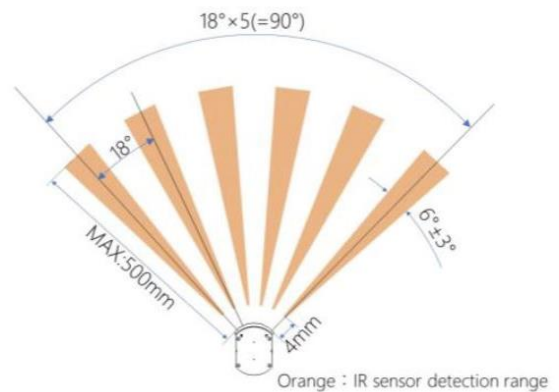


Fig. 2 IR sensor mounting direction and detection

2.2 全方位カメラ

スワーム・ロボティクスにおいて、協調行動の発現には周囲の環境認識が不可欠である。全方位カメラは単体で多くのデータが収集できるため、スワーム・ロボットでしばしば搭載されている。実機ロボットに搭載しているカメラは直径30mmのステンレス半球ミラーを直径28mmに切削加工したものを、長さ50mmの亚克力パイプに装着し全方位化している。

2.3 キャリブレーション

全方位カメラで得られた画像の歪みはロボットの距離推定に大きく影響を及ぼすため、キャリブレーションが必要である。Adriana等は、スワーム・ロ

ボットの相対位置特定システムを志向した全方位画像のキャリブレーションを最小二乗法による多項式回帰を用いて行っており [5], この手法を用いて本研究室が扱っている全方位カメラのキャリブレーションを行っている [6]. シミュレーション上のカメラでも同様のキャリブレーションを行う.

x をシミュレーション上の距離, y をピクセル距離とした場合, 半球ミラーによる変形量は以下の多項式近似によって求めることができる [5].

$$y = a_2x^2 + a_1x + a_0$$

シミュレーション上で計測した距離と全方位画像のピクセル距離を Table.1 に示す.

Table. 1 Simulator distance and pixel distance on omnidirectional images

simulation distance[cm]	pixel distance[pix]
10	59
15	63.16
20	66.6
30	73.6
40	77.5
50	79.68
60	81.2
70	81.7
80	83.2
90	83.7
100	83.72
110	84.2
120	84.25

これをもとに最小二乗による多項式近似をした. 式1の係数を Table. 2 に, グラフを Fig3 に示す.

Table. 2 Coefficients of polynomial

a_2	a_1	a_0
-0.0031	0.5883	56.8872

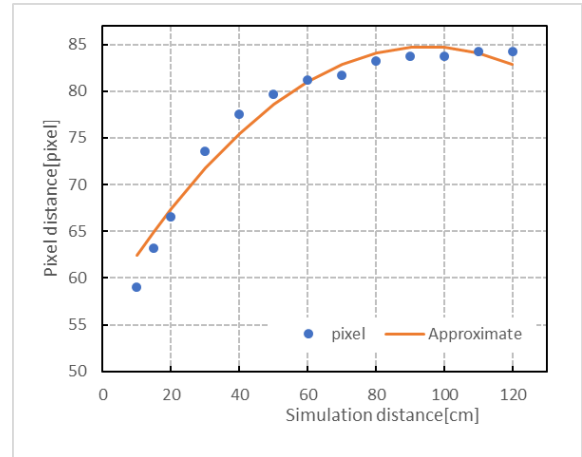


Fig. 3 Polynomial approximation of mirror deformation

3. 計算機シミュレーション

3.1 Reality Gap

近年, タスクの複雑化によってロボット制御に強化学習が導入されている [7]. 特にロボティック・スワームは制御器に強化学習をはじめとした自動化手法が多用されており, これらのトレーニングには安全性・迅速性の高いコンピュータシミュレーションがしばしば用いられる. しかし, コンピュータシミュレーションは離散的かつ形式的に施行され, 個々のロボットが環境から得られるデータは完全である場合が多い. しかし実世界では連続的かつ非形式的であり, 個々のロボットが得られるデータは環境からロボットに実装されたセンサなど, 身体を通して取得され本質的に不完全となる. このシミュレーションとリアルギャップは “Reality Gap” と呼ばれる. このギャップはシミュレーションによってトレーニングされた制御器が現実世界に適合せず, タスクが達成できない原因となりうる.

Reality Gap の低減には環境乱択化・敵対的強化学習・転移学習など様々な研究例が存在する [8]. しかしながら, Reality Gap の低減においては, 身体性に基づいた不完全な環境情報に対し適応させるという視点に重点を置く.

3.2 計算機シミュレーションの構成

本研究では不完全な環境情報としてロボットの位置情報に注目した。シミュレータ上ではロボットの位置は絶対座標で示すことができ、位置情報として完全な環境情報といえる。しかし、実機が取得できるのは多くの場合は相対座標あるいは相対的な位置関係である。不完全な環境情報に対し適応させるために、実機で用いている画像処理プログラムを入力として学習する。Fig.4に学習環境、Fig.5に学習環境のシステム構成を示す。物理シミュレーションはUnityを用い、ML-Agents v2.2によって学習した。また、Fig1で示した実機ロボットのIRセンサはRaycastPerception3Dコンポーネントを用い、全方位カメラによる視覚はFig.6に示すように曲面反射を用いてシミュレートした。

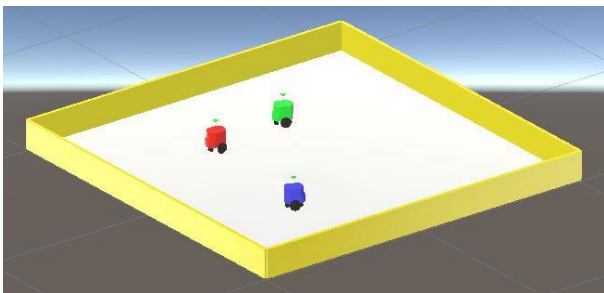


Fig. 4 Simulation environment

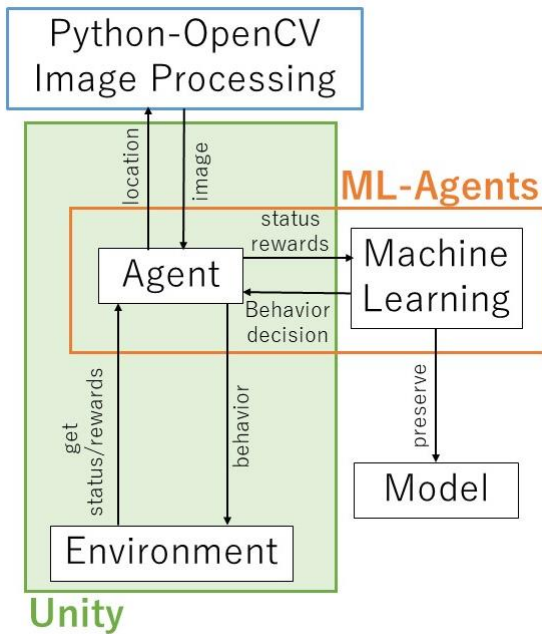
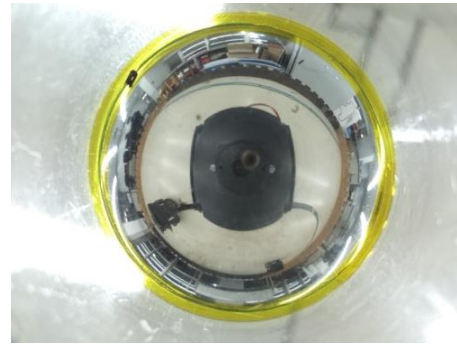
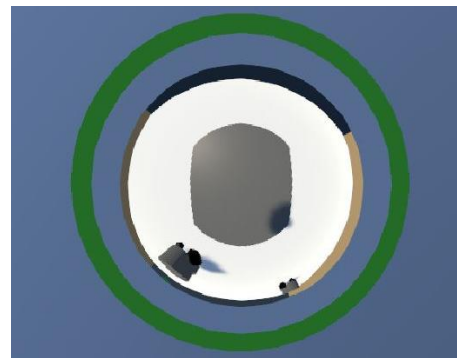


Fig. 5 Schematic figure of the simulation system



(a)



(b)

Fig. 6 real omnidirectional image (a),
Simulated omnidirectional image (b)

3.3 実験設定

与えるタスクはこのフィールド内をランダム移動するMaster Robot(青色)に対してSlave Robot(赤色)が追従行動を行うものである。また、追従行動に無関係なロボット(緑色)をランダム移動させ、対象のロボットのみを追従するようにした。制御器はそれぞれ入力に完全情報、不完全情報を用いた2つ用意し、性能比較する。制御器Aは絶対座標から得られたユークリッド距離・角度と距離センサ値・ロボット検出の有無・画像処理プログラムの応答の有無を入力した、制御器Bは画像処理によって推定されたロボットの距離・角度と距離センサ値・ロボット検出の有無・画像処理プログラムの応答の有無を入力としている。ただし、制御器Aについてはロボット検出の有無と画像処理プログラムの応答については常に1が入力される。Fig.7に示すように画像はPython-Open-CVによる色抽出アルゴリズムを用いた外部プロセスで処理している。

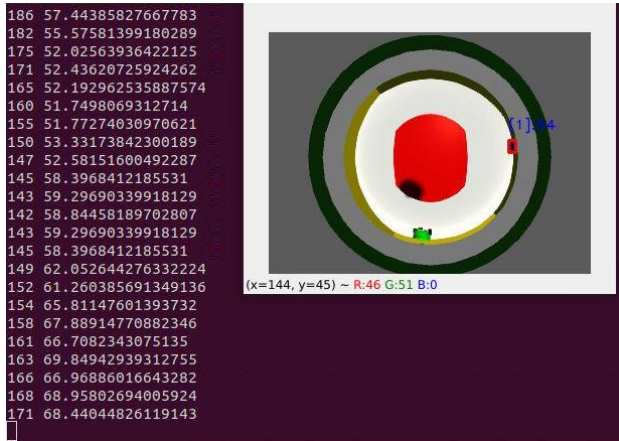


Fig. 7 Image processing process

入力ユニット数は 22, 出力ユニット数 2 で左右モータそれぞれの回転数とした. 隠れ層 2, 1 層のユニット数 20, 学習アルゴリズムは PPO を用いた.

4. 実験結果

4.1 振る舞い

Fig.8 に完全情報を用いた制御器 A の学習で得られた振る舞いの一例を示す. 最終ステップは 500000 であり, 学習時間は 6h である. ①では Slave Robot が IR センサの範囲外で Master Robot に方向転換する. ②では Slave Robot は Master Robot に接近する. ③で双方のロボットの IR センサ検出範囲に侵入し, 接近スピードは弱まる. ④では Slave Robot は停止し, Master Robot は回避行動を行い, 以後, 追従を継続する. Fig.8 は Slave Robot の IR センサ検出範囲に 2 つのロボットが侵入した場合の振る舞いを示している. ①では交互に機体をむける振る舞いをし, ②では正しいロボットを追従している.

制御器 B についてもほぼ同様の振る舞いがみられた. 最終ステップは 500000 であり, 学習時間は 20h である.

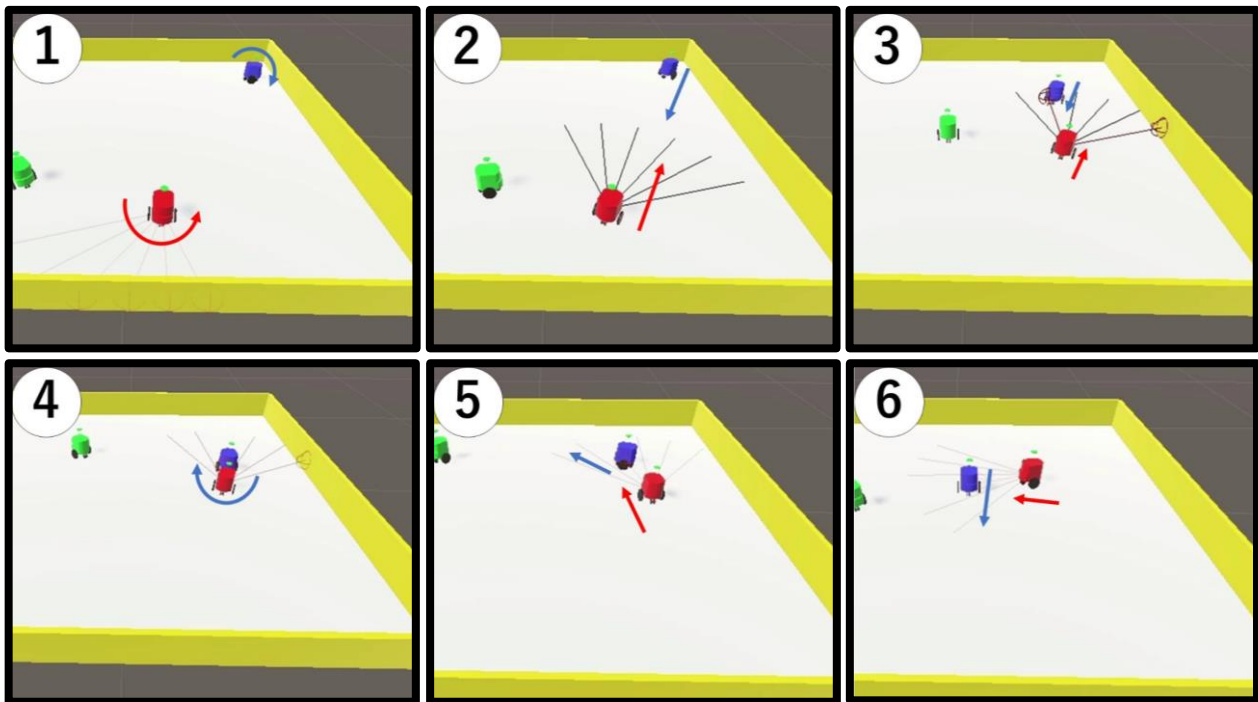


Fig. 8 Robot behavior

(visualizing the detection beam of the IR sensor)

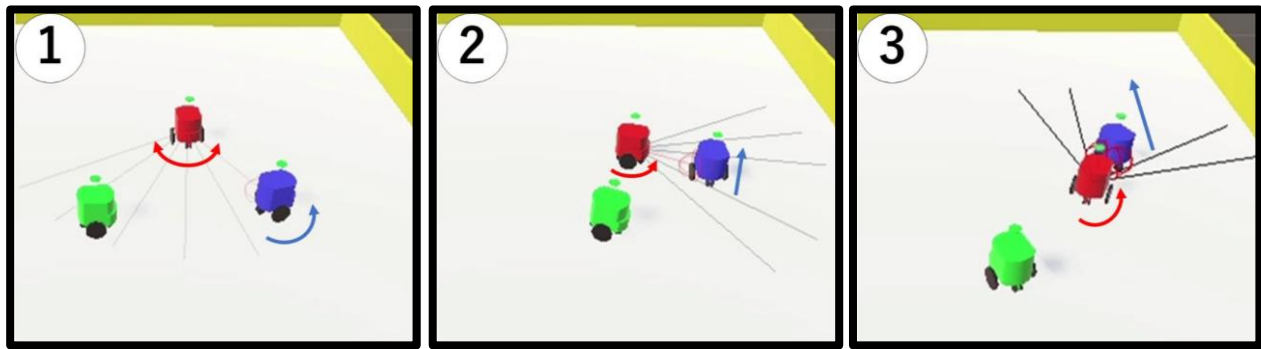


Fig. 9 Behavior when two robots are detected

4.2 性能評価

制御器のタスク達成率を比較するために、追従行動の定義を“機体方位がターゲットから ± 20 度の範囲で向いている、かつ距離がIRセンサの範囲(50cm以下)にとどまるように追従すること”とした。1エピソードを1000シミュレートステップとし、100エピソード実行する。そのうち上記の定義を満たしているステップの割合をタスク達成率とし、Table.3に示す。ただし、エピソード開始時は、双方のロボット位置はランダムに設定される。

完全情報を用いた制御器と不完全情報を用いた制御器では性能に10%の差異がみられた。振る舞いの差としては、制御器BのロボットはIRセンサが壁面を検出した時、スタックしてしまう振る舞いがみられた。また、IRセンサの入力が強く作用し、ダミーとして走行させたロボットに誤って追従してしまう場合もあった。制御器Aにも同様の事象がみられたが、すぐに正しい追従に戻る行動をとった。このような差が発生したのは学習の際、ダミーロボットがMaster Robotを遮り検出できない、推定された方角・距離を用いるなど不完全情報が大きな原因だと考えられる。

Table. 3 Task completion rate for controller

Controller A	controller B
44.973 %	33.954 %

5. おわりに

本研究では、全方位視覚をモデル化し外部プロセスによって処理することで不完全情報をシミュレータ上で再現し、不完全情報を用いた強化学習を実行した。その結果、不完全情報で学習した制御器は完全情報で学習した制御器より性能が落ち、シミュレーション内でReality Gapが確認できた。不完全情報で学習することで学習モデルを用いることで実機に移した時のReality Gapは軽減できる可能性がある。この結果を踏まえ今後は実機実験を通してReality Gap低減を目指す。

参考文献

- 1) Erol Sahin : Swarm Robotics: Inspiration to Domains of Application(2004)
- 2) 伍賀, 深坂, 大野 : 実機ロボティックスワームを志向したセンサ統合に関する研究 : 第21回計測自動制御学会システムインテグレーション部門講演会, 3B2-12(2020)
- 3) 村上, 伍賀 : 実機ロボティック・スワームを志向した小型全方位カメラの画像処理に関する研究, 計測自動制御学会システム・情報部門 学術講演会, GS2-2-4, 2021.
- 4) Kleeberger K: A Survey on Learning-Based Robotic Grasping ,Current Robotics Reports, 1, pp. 239-249, 2020, doi: 10.1007/s43154-020-00021-6
- 5) C. M. Adriana Milasan: Least Squares Estimation of Round

Convex Mirror Image Distortion for Indoor Localization in
Robot Swarms, MTA REVIEW, XXVI-1, 27/38,(2016)

- 6) 本田,村上,伍賀: スマート農業用環境計測デバイスの全方位カメラの改良に関する研究, 第 31 回計測自動制御学会中国支部学術講演会,2022, pp. 103 – 104
- 7) Kober, J., Bagnell, J.A., Peters, J.: Reinforcement learning in robotics: A survey.
- 8) The International Journal of Robotics Research 32(11), 1238–1274 (2013)
- 9) J. Kober, “Reinforcement learning in robotics: A survey”, The International Journal of Robotics Research 32(11), pp. 1238–1274 (2013)
- 10) Erica Salvato: Crossing the Reality Gap: A Survey on Sim-to-Real Transferability of Robot Controllers in Reinforcement Learning, IEEE Access, 2021, 9, doi: 153171-153187.