

令和 6 年 6 月 25 日現在

機関番号：55201

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K12147

研究課題名（和文）実環境における深層強化学習の実現と群ロボットへの展開

研究課題名（英文）Realization of deep reinforcement learning in real environment and its application to swarm robots

研究代表者

堀内 匡（HORIUCHI, Tadashi）

松江工業高等専門学校・電子制御工学科・教授

研究者番号：50294129

交付決定額（研究期間全体）：（直接経費） 3,400,000円

研究成果の概要（和文）：本研究では、実環境での深層強化学習を実現し、群ロボットの協調行動の獲得に応用した。まず、シミュレーションと実機の差異の低減手法を検討し、単一の移動ロボットの視覚情報に基づく行動獲得を深層強化学習により実現した。次に、複数のロボットが存在する環境を想定し、実機の群ロボットを対象とした協調行動の獲得を実現した。具体的には、群ロボットの環境における協調行動として、追従行動および追い抜き行動の獲得を深層強化学習により実現した。さらに、深層強化学習を用いたロボットの行動選択の説明性の向上に向けて、ロボットがカメラ画像内のどの領域に注目して行動を選択したかという注視領域の可視化について、有効性を確認した。

研究成果の学術的意義や社会的意義

本研究は、人工知能の中核である機械学習の技術を実機ロボットに応用したものである。深層学習と強化学習を組み合わせた深層強化学習という手法を用いて、実機ロボットが視覚情報に基づいて行動を自分自身で学習することを実現した。また、複数のロボットが存在する環境を想定し、実機の群ロボットを対象とした協調行動の獲得を実現した。具体的には、群ロボットの環境において、追従行動および追い抜き行動の獲得を実現し、深層強化学習による協調行動の学習が可能であることを明らかにした。これらの成果は、物流倉庫や福祉施設などにおける搬送ロボット群の行動制御などにつながると期待できる。

研究成果の概要（英文）：In this study, we realized deep reinforcement learning in real environment and applied it to acquiring cooperative behavior of swarm robots. First, we investigated the method to reduce the difference between the simulation environment and the real robot environment. Using this method, we achieved behavior acquisition based on visual information for a single mobile robot using deep reinforcement learning. Next, assuming the environment in which multiple robots exist, we achieved the acquisition of cooperative behavior for real swarm robots. More concretely, we realized to acquire following and overtaking behaviors as cooperative behaviors in swarm robot environment by deep reinforcement learning to enable robot. Finally, in order to improve the explainability of robot action selection using deep reinforcement learning, we confirmed the effectiveness of visualizing the attention area, that indicates which area in the camera image the robot focused on when selecting an action.

研究分野：知能システム

キーワード：深層強化学習 知能ロボティクス 群ロボット 行動獲得

1. 研究開始当初の背景

強化学習は、報酬を手掛かりにしてタスク達成のための行動を「自己学習」する方法である。近年、強化学習は深層学習と結びつき、深層強化学習として大きな注目を集めている。Deep Q-network (DQN) を用いた人間レベルのゲーム AI の開発とトッププロ棋士以上の強さを誇る AlphaGo がそのきっかけである。両者ともに深層学習と強化学習を組み合わせた深層強化学習が成功の鍵となっている。

しかし、深層強化学習は、多数の試行が容易にできるゲーム AI やシミュレーション環境での研究がほとんどであり、実環境への応用は進んでいない。深層強化学習を実機ロボットなどの実環境に適用するには、「実機での試行回数を削減すること」が重要な課題として存在する。そのため、シミュレーションの活用が重要となるが、「シミュレーションと実機の差異をどのように低減するか」という課題が存在する。

上記の課題に対して、我々は実環境に限りなく近いシミュレーション環境を構築するのではなく、実機のカメラ画像に対してセグメンテーション(領域分割)と低解像度化をすることで実環境での見え方をシミュレーション環境での見え方に近づける手法を導入した。この手法により、深層強化学習の代表的な手法である DQN を用いて、単一の実機の移動ロボットが壁や障害物を避けながら進む行動を獲得することを実現した(図1参照)。そこで、本研究では「シミュレーションと実機の差異の低減手法に着目した実環境での深層強化学習の実現」を図るとともに、実社会での応用において重要となる「深層強化学習における行動選択の説明性の向上」に取り組む。

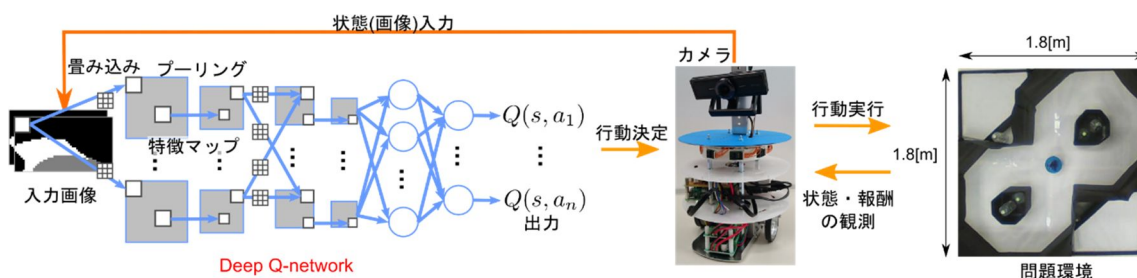


図1. DQN を用いた実機の移動ロボットの視覚ベースの行動獲得

2. 研究の目的

本研究では、車輪型移動ロボットを対象として、以下の三つを研究の目的とする。

第一目的として、「複雑な実環境での単一のロボットの行動獲得」を実現する。家具などが存在する室内環境における単一の移動ロボットの視覚情報に基づく行動獲得を深層強化学習により実現する。

第二目的として、複数のロボットが存在する環境を想定し、「実機の群ロボットを対象とした協調行動の獲得」を実現する。実機の群ロボットの環境における「追い抜き行動」や「すれ違い行動」などの協調行動の獲得を深層強化学習により実現する。

第三目的として、「深層強化学習における行動選択の説明性の向上」を実現する。深層強化学習において、ロボットがカメラ画像内のどの領域に注目して行動を選択したかを明らかにする。

3. 研究の方法

まず、本研究の第一目的である「複雑な実環境での単一のロボットの行動獲得」の実現に関して、以下の方法により研究を実施した。

- (1) 室内環境を模したシミュレーション環境および実環境において、車輪型移動ロボットを対象として行動獲得を深層強化学習により実現した。深層強化学習のアルゴリズムとして、DQN (Deep Q-network) およびその発展手法である Rainbow を適用し、単一の移動ロボットの行動獲得における学習性能を比較した。さらに、ベイズ最適化を用いて深層強化学習のハイパーパラメータと報酬関数を最適化することを実現した。

次に、本研究の第二目的である「実機の群ロボットを対象とした協調行動の獲得」について、以下の方法により研究を実施した。

- (2) 3台の車輪型移動ロボットが存在する群ロボット環境における協調行動を獲得するタスクとして「追い抜き行動」の獲得および「追従行動」の獲得を目指した。シミュレーション環境および実環境の両方において、それぞれのタスクに対して DQN および Soft Actor-Critic を用いた実現方法を検討し、学習性能を比較した。

さらに、本研究の第三目的である「深層強化学習における行動選択の説明性の向上」について、以下の方法により研究を実施した。

- (3) 深層学習における注視領域の可視化手法である Grad-CAM および Grad-CAM++を深層強化学習に適用する手法を提案し、ロボットがカメラ画像内のどの領域に注目して行動を選択したかを検証した。さらに、Vision Transformer (ViT) における注視領域の可視化手法について実験的に特徴を明らかにして、深層強化学習における行動選択の説明性の向上について検討した。

4. 研究成果

- (1) 複雑な実環境での単一のロボットの行動獲得

室内環境を模したシミュレーション環境において深層強化学習 DQN の発展手法である Rainbow アルゴリズムを用いることでロボットのゴール到達率の改善を実現した。また、より複雑な室内環境で実験を行うことにより、Rainbow を使用することによる適応能力の向上を確認することができた。図 2 に示す 2 箇所のゴールと障害物となる家具等を設置した室内環境を

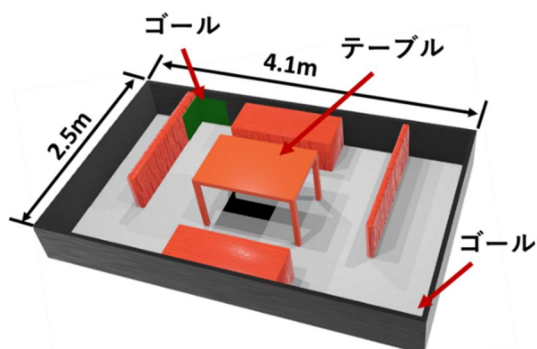


図 2. 問題環境 (シミュレーション環境)

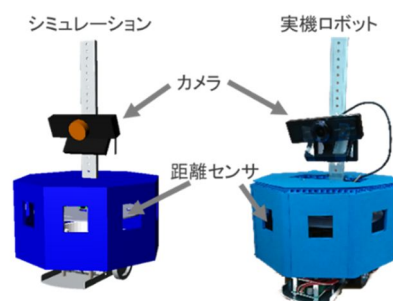


図 3. 車輪型移動ロボット

想定し，図 3 に示す車輪型移動ロボットの 8 方向の距離センサの値と Web カメラの画像を Rainbow の入力とする．5 箇所のスタート位置から壁や障害物に衝突せずにゴールに到達することを学習する．ゴールに到達できれば成功とし正の報酬を与え，壁や障害物に衝突または行動が 1000 ステップ経過した場合は負の報酬を与える．

各ステップにおける成功率（ゴール到達率）の平均を図 4 に示す．比較対象の手法として，DQN の発展手法である Dueling DDQN (Dueling Double DQN) を用いた．Rainbow を適用することによって，平均成功率を Dueling DDQN の 67.5% から 97.5% に大きく向上させることを実現した．

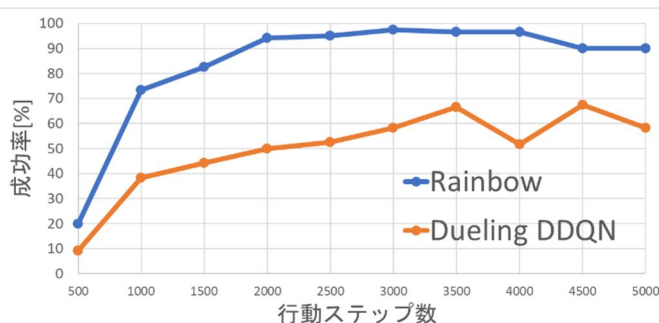


図 4. 成功率（ゴール到達率）の比較：5 セット平均

(2) 実機の群ロボットを対象とした協調行動の獲得

深層強化学習 DQN の発展手法である Dueling DDQN を群ロボットの行動獲得に適用した．3 台のロボットが同じ方向に周回する問題環境において，各ロボットが LiDAR の値と Web カメラを用いて，他のロボットと通信せず行動を選択することを想定した．タスク 1 として他のロボットに追従する行動の獲得，タスク 2 として他のロボットを追い抜く行動の獲得を実現するためにシミュレーション実験と実機実験を行った．シミュレーション環境を図 5 に示し，実機環境を図 6 に示す．

タスク 1 の追従行動では，遅いロボットを先頭に，他の速いロボットが壁や障害物，前のロボットに衝突せずに，直進（高速）と直進（低速）を繰り返しながら追従する行動が獲得できた．シミュレーション環境と実機環境における学習時の 1 エピソード毎のロボット 1 台あたりの衝突回数の推移を図 7 に示す．実機実験ではシミュレーション実験で得た DQN の重みを初期値として用いた．

タスク 2 の追い抜き行動では，遅いロボットに衝突せずに障害物と遅いロボットの間を走行して追い抜く行動が獲得できた．シミュレーション環境と実機環境における学習時の 1 エピソード毎のロボット 1 台あたりの衝突回数の推移を図 8 に示す．こちらも実機実験ではシミュレーション実験で得た DQN の重みを初期値として用いた．

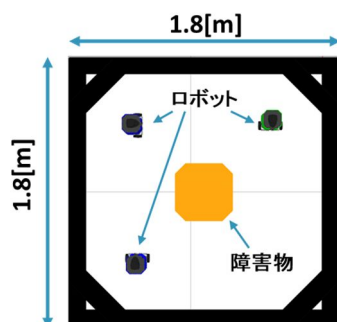


図 5. シミュレーション環境

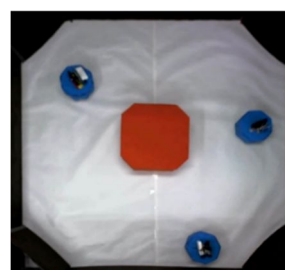


図 6. 実機環境

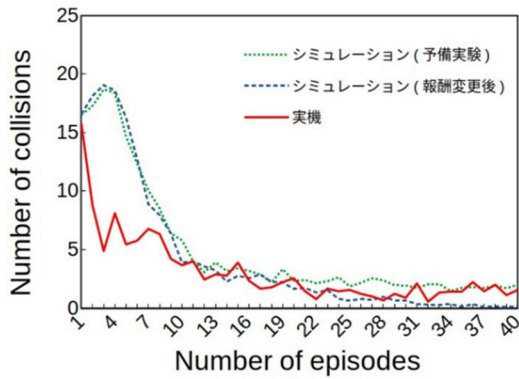


図 7. タスク 1 : 学習中の衝突回数の推移 (シミュレーション実験 10 セット, 実機実験 3 セットの平均)

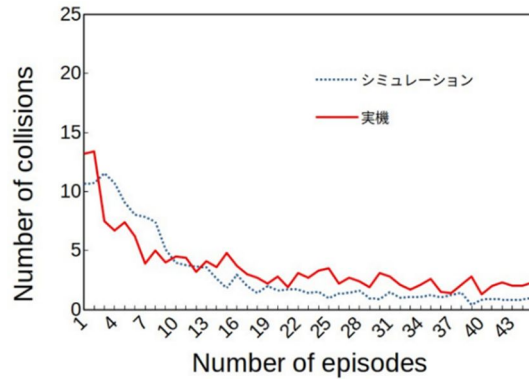


図 8. タスク 2 : 学習中の衝突回数の推移 (シミュレーション実験 10 セット, 実機実験 3 セットの平均)

(3) 深層強化学習における行動選択の説明性の向上

深層強化学習 DQN を用いた群ロボットの行動学習におけるロボットの注視領域の可視化を実現した . 図 9 に示すシミュレーション環境において, 3 台のロボットを対象に CNN の注視領域の可視化手法である Grad-CAM++ を適用することでロボットの注視領域の可視化をした .

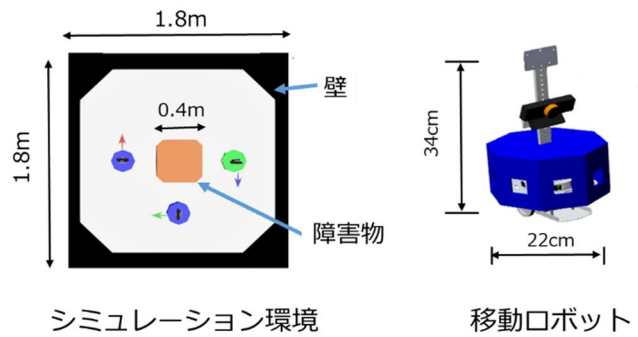


図 9. 実験環境

2 台の青色の筐体の速いロボットと緑色の筐体の遅いロボットを対象に DQN の行動学習を行った後, DQN に対して Grad-CAM++ を適用した . そのときの入力画像と結果例を図 10 に示す . 図 10 左の入力画像に対して, 遅いロボットを追い抜くための行動として速いロボットが右折行動を選択しているとき (図 10 中央), Grad-CAM++ により得られた注視領域を図 10 右に示す . 紙面の都合上 Grad-CAM の適用結果は省略するが, Grada-CAM に比べて Grad-CAM++ の方が注視領域は広く Grad-CAM では見逃していた特徴量を捉えていると考えられる . Grad-CAM++ では, 前方のロボットに特に注目していることを定量的に明らかにした .

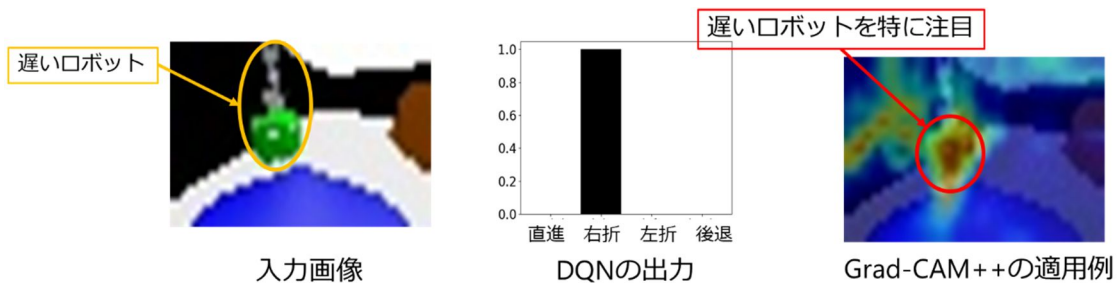


図 10. 入力画像 (ロボットのカメラ画像) と Grad-CAM++ の適用例

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 R. Sota, T. Nishimura and T. Horiuchi	4. 巻 15
2. 論文標題 A Study on Optimization of Hyper-Parameters in Deep Reinforcement Learning by Bayesian Optimization	5. 発行年 2024年
3. 雑誌名 ICIC Express Letters, Part B: Applications	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 R. Watanuki, T. Horiuchi and T. Aodai	4. 巻 11
2. 論文標題 Vision-based Behavior Acquisition by Deep Reinforcement Learning in Multi-Robot Environment	5. 発行年 2020年
3. 雑誌名 ICIC Express Letters, Part B: Applications	6. 最初と最後の頁 237-244
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計29件（うち招待講演 1件／うち国際学会 5件）

1. 発表者名 R. Sota, T. Nishimura and T. Horiuchi
2. 発表標題 A Study on Bayesian Optimization of Hyper-Parameters and Reward Settings in Deep Reinforcement Learning
3. 学会等名 The 2nd International Symposium on Data-Driven Intelligent Optimization for Decision Making（国際学会）
4. 発表年 2024年

1. 発表者名 R. Sota, T. Nishimura and T. Horiuchi
2. 発表標題 A Study on Optimization of Hyper-Parameters in Deep Reinforcement Learning by Bayesian Optimization
3. 学会等名 The 17th International Conference on Innovative Computing, Information and Control（国際学会）
4. 発表年 2023年

1. 発表者名 曾田涼介, 西村拓人, 堀内 匡
2. 発表標題 ベイズ最適化による深層強化学習のハイパーパラメータと報酬関数の最適化
3. 学会等名 第68回システム制御情報学会研究発表講演会
4. 発表年 2024年

1. 発表者名 原 歩陸, 堀内 匡
2. 発表標題 Vision Transformerによる画像認識と判断根拠の可視化
3. 学会等名 第28回日本知能情報ファジィ学会中国・四国支部大会
4. 発表年 2024年

1. 発表者名 生和直央, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの協調行動の獲得
3. 学会等名 第28回日本知能情報ファジィ学会中国・四国支部大会
4. 発表年 2024年

1. 発表者名 曾田涼介, 西村拓人, 堀内 匡
2. 発表標題 ベイズ最適化による深層強化学習のハイパーパラメータと報酬関数の最適化の試み
3. 学会等名 電気学会 システム/制御 合同研究会
4. 発表年 2023年

1. 発表者名 曾田涼介, 西村拓人, 堀内 匡
2. 発表標題 ベイズ最適化を用いた深層強化学習のハイパーパラメータの最適化に関する検討
3. 学会等名 2023年電気学会電子・情報・システム部門大会
4. 発表年 2023年

1. 発表者名 福島 英, 曾田涼介, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの協調行動の獲得に関する検討
3. 学会等名 2022年電気学会電子・情報・システム部門大会
4. 発表年 2022年

1. 発表者名 曾田涼介, 福島 英, 堀内 匡
2. 発表標題 深層強化学習を用いた移動ロボットの行動獲得における改良手法の検討
3. 学会等名 2022年電気学会電子・情報・システム部門大会
4. 発表年 2022年

1. 発表者名 生和直央, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの追い抜き行動の獲得
3. 学会等名 第27回日本知能情報ファジィ学会中国・四国支部大会
4. 発表年 2023年

1. 発表者名 曾田涼介, 西村拓人, 堀内 匡
2. 発表標題 ベイズ最適化を用いた深層強化学習のハイパーパラメータの最適化
3. 学会等名 第27回日本知能情報ファジィ学会中国・四国支部大会
4. 発表年 2023年

1. 発表者名 福島 英, 綿貫零真, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの行動獲得に関する実験的考察
3. 学会等名 2021年電気学会電子・情報・システム部門大会
4. 発表年 2021年

1. 発表者名 福島 英, 綿貫零真, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの協調行動の獲得
3. 学会等名 第26回日本知能情報ファジィ学会中国・四国支部大会
4. 発表年 2022年

1. 発表者名 綿貫零真, カン サリユー, 福島 英, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの実機環境での行動獲得の試み
3. 学会等名 計測自動制御学会システム・情報部門学術講演会2020
4. 発表年 2020年

1. 発表者名 福田隼也, 綿貫零真, 堀内 匡
2. 発表標題 深層強化学習における移動ロボットの注視領域の可視化手法に関する検討
3. 学会等名 電気学会システム研究会
4. 発表年 2020年

1. 発表者名 福島 英, 山田航平, 堀内 匡
2. 発表標題 深層強化学習を用いた群ロボットの行動獲得に関する検討
3. 学会等名 第25回日本知能情報ファジィ学会中国・四国支部大会
4. 発表年 2021年

1. 発表者名 R. Watanuki, T. Horiuchi and T. Aodai
2. 発表標題 A Study on Vision-based Behavior Acquisition of Multi-Robot System by Deep Reinforcement Learning
3. 学会等名 SICE Annual Conference 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 R. Watanuki, T. Horiuchi and T. Aodai
2. 発表標題 Vision-based Behavior Acquisition by Deep Reinforcement Learning in Multi-Robot Environment
3. 学会等名 The 14th International Conference on Innovative Computing, Information and Control (国際学会)
4. 発表年 2019年

1. 発表者名 栗山海渡, 綿貫零真, 堀内 匡, 青代敏行
2. 発表標題 深層強化学習を用いた実機ロボットの室内環境での行動獲得の試み
3. 学会等名 2019年電気学会電子・情報・システム部門大会
4. 発表年 2019年

1. 発表者名 R. Watanuki, T. Horiuchi and T. Aodai
2. 発表標題 A Study on Behavior Acquisition by Deep Reinforcement Learning in Multi-Robot Environment
3. 学会等名 2019年電気学会電子・情報・システム部門大会
4. 発表年 2019年

1. 発表者名 西田吉克, 福田隼也, 綿貫零真, 堀内 匡, 青代敏行
2. 発表標題 深層強化学習における移動ロボットの注視領域の可視化の試み
3. 学会等名 計測自動制御学会システム・情報部門学術講演会2019
4. 発表年 2019年

1. 発表者名 堀内 匡
2. 発表標題 人工知能・機械学習の活用とロボティクス
3. 学会等名 技術コミュニティラボ 第9回ミーティング (招待講演)
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	青代 敏行 (AODAI Toshiyuki) (40571849)	東京都立産業技術高等専門学校・ものづくり工学科・准教授 (52605)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------