

Title	視野による制約を前提としたゲームAIの強化学習
Author(s)	勝又, 翼; シュエ, ジュウシュエン; 池田, 心
Citation	ゲームプログラミングワークショップ2024論文集, 2024: 111-116
Issue Date	2024-11-15
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/20011
Rights	社団法人 情報処理学会, 勝又 翼, シュエ ジュウシュエン, 池田 心, ゲームプログラミングワークショップ2024論文集, 2024, 111-116. ここに掲載した著作物の利用に関する注意 本著作物の著作権は情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	第29回ゲームプログラミングワークショップ (GPW-24), 箱根セミナーハウス, 2024年11月15日-17日



視野による制約を前提としたゲーム AI の強化学習

勝又 翼^{1,a)} シュエ ジュウ シュエン^{1,b)} 池田 心^{1,c)}

概要：ゲーム AI は主に強さの向上を目的として研究されてきたが、近年人間らしい挙動の獲得も研究が進められている。本研究では、人間の視野特性に着目し、強化学習エージェントに動的な視野を実装することで、「視野の制約を前提とした行動」の再現を目指す。提案手法では、注視点の動的な決定と注視点からの距離に応じた情報のぼかしを導入し、エージェントがゲーム内行動と視野移動を同時に学習する。Atari Breakout に簡単な視野の制約を導入した実験では、視野の移動を含めた学習が可能であることを確認した。さらに視野の制約がある場合に、序盤の成績がかえって向上する傾向を確認した。今後は、格闘ゲームなどのより複雑なゲーム環境での学習の確認や、「視野の制約を前提とした行動」の再現を目指す。

キーワード：深層強化学習、人間らしい AI、生物学的制約

Reinforcement Learning of Game AI Assuming Constraints on Visual Fields

TSUBASA KATSUMATA ^{1,a)} CHU-HSUAN HSUEH ^{1,b)} KOKOLO IKEDA ^{1,c)}

Abstract: While game AI research has primarily focused on improving playing skills, recent studies have also explored the acquisition of human-like behavior. This research aims to reproduce “actions based on visual field constraints” by implementing a dynamic visual field in reinforcement learning agents, focusing on human visual characteristics. The proposed method introduces dynamic determination of the gaze point and information blurring based on distance from the gaze point. In the system, the agents simultaneously learn in-game actions and visual field movement. Experiments with simple visual field constraints in Atari Breakout confirmed the feasibility of learning even with visual field movement. Furthermore, we observed a tendency for improved scores in the early stages when visual field constraints were present. Future challenges include verifying learning in more complex game environments, such as fighting games, and reproducing actions based on visual field constraints.

Keywords: Deep Reinforcement Learning, Human-like AI, Biological Constraints

1. はじめに

これまで、ゲーム AI は主に強さを向上させることを目的として研究されてきた。ビデオゲームの Dota2 では人間のプロプレイヤーに対してチーム戦で勝利を収め [1]、完全情報のボードゲームである将棋や囲碁では既に人間のプロ棋士を圧倒する強さを獲得し、強くなるためにゲーム AI を用いて学習することは当たり前になっている。

強さを追求する研究は現在も継続されているが、強さに最適化されたゲーム AI の挙動はしばしば人間らしいとされるプレイとはかけ離れることがある。特にリアルタイムで進行するビデオゲームでは、人間よりはるかに速い反応を見せたり、人間には同時に処理できない量の情報を 1 度に処理してしまったり、1 フレーム単位で全くミスなく操作を行ったりと、人間プレイヤーの対戦相手やお手本として不適切な場合がある。このような状況を背景に、ゲーム AI の人間らしさに対する需要が高まってきている。例えば、2D 格闘ゲームのストリートファイター 6 では、CPU (AI) に対して人間らしい行動を実装する試みがなされている [2]。

本研究では、「視野」という観点からゲーム AI に人間ら

¹ 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology,
Nomi, Ishikawa 923-1211, Japan

a) t-katsumata@jaist.ac.jp

b) hsuehch@jaist.ac.jp

c) kokolo@jaist.ac.jp

しい挙動をさせることを試みる。人間の視覚による認知には制限があり、周囲全てを視野に入れることができないだけでなく、視野内でも認識の精度に差がある。人間は、自然とそういった制限を前提とした行動を身に付けている場合があり、それをゲーム AI にも獲得させることが本研究の目的である。例えば自動車のドライバーが左折する際、同時に複数箇所を見なくても済むように、あらかじめ左側に車を寄せることで、自転車横断が横をすり抜けるリスクを減らしている。また、2D 格闘ゲームにおける「置き」や「仕込み」（詳しくは付録を参照）といった戦略は、視野による制約を考慮した行動の典型例である。視野による制約を強化学習に組み込むことで、ゲーム AI に制約を前提とした行動を獲得させることが、この研究の目的である。

2. 関連研究

本章では、人間らしいゲーム AI に関する関連研究として、実際のゲーム開発での事例と先行研究を紹介する。また、人間の視覚特性についても取り上げ、これらがゲーム AI の開発にどのように関連するかを探る。

2.1 ストリートファイター 6 における人間らしいゲーム AI の事例紹介

安原は対戦格闘ゲームにおける CPU の役割を二つの側面から捉えている [2]。一つは、一人用アクションゲームとしての CPU であり、プレイヤーに操作に慣れてもらうことや攻略する楽しさを提供する。もう一つは、対人戦としての CPU であり、システムや駆け引きの理解を促し、オンライン対戦の予行演習となる役割を果たす。この、対人戦前の練習としての役割を効果的に果たすためには、CPU がある程度人間らしい行動をとることが重要である。ストリートファイター 6 では、AI の行動を「戦術」「作戦」「実行能力」の 3 つの観点で分解し、CPU のレベルを段階的に表現している。各レベルの CPU は、3 つの観点全てで単調に強くなっていくのではなく、人間プレイヤーが対人戦に習熟していく過程での能力の伸びに合うように、各ランクの人間プレイヤーにありそうなミスを行っている。

2.2 人間らしいゲーム AI に関する先行研究

藤井らは、認知・行動のゆらぎや遅れといった人間プレイヤーに共通して生じる現象を、生物学的制約として定義し、強化学習に導入した [3]。マリオのようなゲームを題材として、位置情報にゆらぎや遅れを加えて学習させた結果、エージェントは敵を余裕を持って避けたり、危険そうな箇所では待つから進んだりするなど、制約を前提とした人間らしいゲーム行動を獲得することに成功した。この研究は、制約を前提とした行動獲得という点で、本研究の直接の先行研究として位置づけられる。しかし、この手法では画面全体に均一にガウスノイズを付与しているため、人間

の視覚特性における中心視と周辺視の違いを再現できていない。人間の視覚特性の詳細については次節で取り上げるが、実際の人間の視覚は視野の中心部と周辺部で情報の精度が異なることが知られている。この特性を考慮することで、より人間らしい行動を AI に再現できる可能性がある。

平井らの研究では、人間の視覚特性である周辺視の特徴をシューティングゲームに取り入れている [4]。注視点から離れるほどゆらぎをかけ、オブジェクトの位置情報を不正確にすることで、人間らしいゲーム AI の作成を目指した。視野の中心部と周辺部で情報精度の差を再現したことは重要だが、平井らの研究では注視点に操作キャラクタに固定されている。人間プレイヤーは必ずしも操作キャラクタだけを見ているわけではなく、また敵キャラクタや敵の撃つ弾のどれかだけを注視しているわけでもない。実際の人間プレイヤーは、ゲーム画面内の様々な場所に視線を移動させ、状況に応じて注視する箇所を変えているはずである。

2.3 視野特性に関する研究

本節では、人間の視野特性に関する研究を紹介する。人間の視野は中心視野と周辺視野に分けることができ、それぞれ異なる特性を持つことが知られている。中心視野は比較的高い空間解像度を有し、詳細な情報を捉えることができるのに対し、周辺視野は主に動きの検出に特化している。

小松原は、人間の視界からの認知について、視野の限界や弁別能力を詳細にまとめている [5]。また、単に中心視野と周辺視野の二分だけでなく、視野内での視力の変化や、弁別能力の段階的な違いも指摘している。図 1 は視線からのズレに応じた視力の変化を示すグラフであり、図 2 は視野角度ごとの弁別能力を小松原が体系的にまとめたものである。

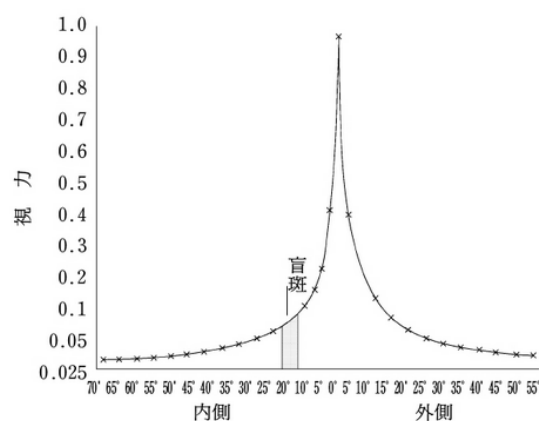


図 1 視線からのズレと視力

真島英信, “生理学”, p241, 文光堂 (1978) より引用

さらに、ある物体から別の物体に視線を跳躍的に移動させる際に生じるサッカード中には、視覚感度が低下する現象（サッカード抑制）が生じることが知られている。人間

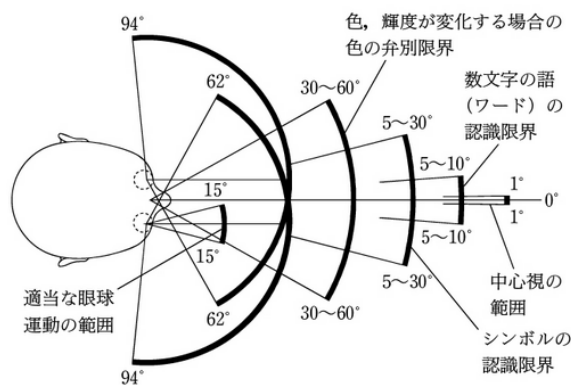


図 2 視野と弁別能力 [5] より引用

プレイヤーがゲームをプレイしている際には、こういった特性から生じる制約を無意識に前提とした行動を取っている。

3. 提案手法

本章では、強化学習エージェントに動的な視野を実装し、視野の制約を前提とした行動を再現させる手法について詳細に説明する。提案手法の核となる 2 つの要素は、(1) 注視点の動的な制御と (2) 注視点からの距離に応じた情報のぼかしである。

まず、画面内のある座標を注視点として定義する。注視点の座標は各フレームごとに、エージェントの行動の一部として決定される。ゲーム内の行動だけでなく、注視点の移動もエージェントの学習対象とし、エージェントはゲーム内の行動と注視点移動の両方を同時に出力する。これにより、例えばゲーム内で操作キャラクターを左に動かしながら、同時に注視点を右に移動させるといった行動が可能となる。

次に、注視点からの距離に応じて、エージェントが受け取る情報にぼかしをかけ、人間プレイヤーに見られる視野内の視力差を簡易的に再現する。この処理の結果、エージェントはゲーム画面の情報全てを均等に受け取るのではなく、注視点から離れるほど解像度が低い情報を受け取り、行動を決定することになる。

エージェントは、注視点によってぼかされた画像を入力として受け取り、ゲーム内の行動と注視点の移動という 2 つの行動を出力する。その後、ゲーム内行動の結果により画面が更新され、新たな注視点の座標に基づいて画像がぼかされる。このぼかされた画像が再びエージェントの入力となり、次の行動決定が行われる。

4. Breakout を題材とした実験

この研究では視野の影響によって生まれる行動を学習によって獲得することが目的だが、そもそも視野の制約下で強化学習が行えるかどうかはまだ確認されていない。本章では、まず Atari Breakout を対象問題として、画面を複

数のエリアに分割する簡単な方法で視野の制約を導入した強化学習を行い、ぼかしの有無や強度による性能の比較を行う。

実験は次の 3 つの段階で構成され、各段階で視野の制約方法を変更しながら学習の可能性を検証した。

- (1) ブラー効果による視野制約の影響評価
- (2) 3×3 エリアによる視野の実装
- (3) 7×7 エリアによる視野の実装

エージェントの行動を環境内行動と視野移動行動（上下左右とその場の 5 種類）の組み合わせで定義したため、行動の種類が通常に比べてかなり増加した。そのため、すべての実験はこうした行動数の増加に強い DuelingNetwork を用いている。このネットワークに対して、グレースケール化した 84×84 の画像を入力した。視野を含めた学習が可能かどうかを確認することが目的であるため、ある程度のプレイができるようになる 200000 エピソードまでの実験を行った。

4.1 ブラーによる性能低下の確認

2 章で述べたように、人間の視野は場所により視力・分解能が異なる。このような現象はブラー等を使用することで近似し、ネットワークの入力として与えることはできるが、それが強化学習エージェントにとって制約になり、性能の低下を引き起こすのかについては自明ではない。まずは人間の分解能の差を再現するため画面全体に対して、図 3 に示すようなボールが認識できる程度のブラー（以下弱いブラー）と、それよりも強くボールが認識できないほどのブラー（以下強いブラー）の 2 種類を用意し、ブラーをかけていないものと性能を比較する。

図 4 は各設定における性能（壊したブロックの数）の推移を示したものである。各設定 3 回の試行を行い、3 試行の平均を実線で、その 95%信頼区間を帯で表した。強いブラーをかけたものは試行ごとのばらつきが激しいが、ブラーをかけ学習したものは標準のものに比べて性能が落ちていることがわかる。ボールが認識できない強いブラーの 1 つの試行が弱いブラーの平均に迫る結果を出しているが、大まかに認識できる壊れたブロックを利用した挙動を学習しようとしているものと推測している。

今回画面全体にかけた弱いブラーと強いブラーがエージェントに対して制約になっていることが確認できたため、次節からそれぞれを中心に近い周辺視野に対応する加工、中心から離れた周辺視野に対応する加工として実験を行う。

4.2 3×3 のエリアに分割した実験

ここからは、画面に対して注視エリアからの距離に応じて強さの異なるブラーをかけて実験を行う。この節では画面を 3×3 エリア分割して学習させた実験について記載す

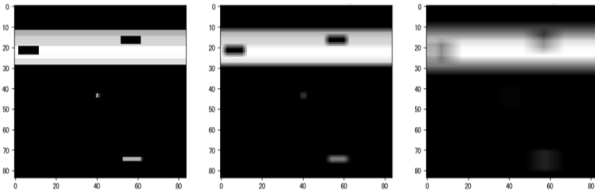


図3 ブラーの強弱による入力の違い

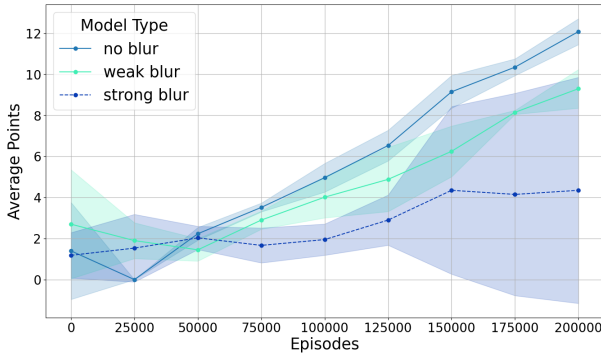


図4 ブラーの強弱による性能差

る。注視エリアは1つで、その周囲8近傍に弱いブラーを、さらにその外側には強いブラーをかけて学習を行った。

注視エリアが常にボールの位置になるよう設定したモデル（以下 auto-ball-fix）を用意し、これと注視エリアを上下左右に移動できるモデル（以下 manual-move-area）を比較したグラフを図6に示す。緑で示した manual-move-area は、この段階ではあまりブロックを壊すことができておらず、学習初期のランダム行動による性能を超えられていない。一方橙色で示した auto-ball-fix は、視野による画像の加工があるにもかかわらず、標準の設定よりも多くのブロックを壊すことに成功している。

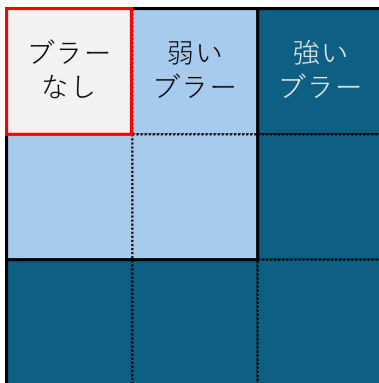


図5 実験2でのブラーのかけ方

最終的な性能については同等か、視野を設定していないものが上回ることが予想されるが、少なくとも学習序盤時点では視野によって画像がぼやけている auto-ball-fix の方が多くのブロックを破壊できることが確認できた。ボールの位置に固定された注視点からの距離に応じた画像のぼかしによって、ボールの位置情報が強調された入力となり、

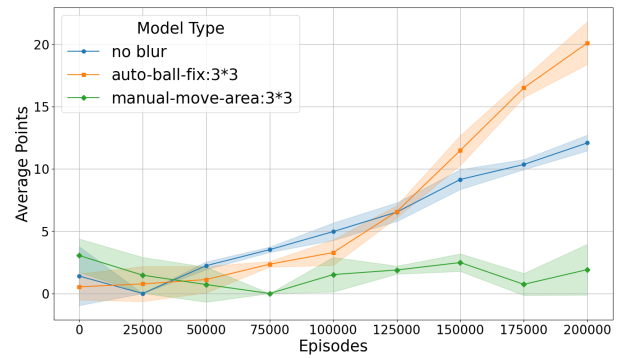


図6 3×3 エリアに分割して行った実験

ボール周りの重点的な学習をむしろ促進した可能性もあると考えている。

4.3 7×7のエリアに分割した実験

最後に、視野エリアの分割を先ほどより少し細かくする代わりに、ブラーの範囲が重なるよう実装した実験を行った。具体的には、視野エリアを7×7マスの分割としたうえで、注視エリアを中心とした3×3マスにはブラーをかけず、弱いブラーをその周囲2マス外側の7×7マスに、それよりも外側に強いブラーをかけた（参考図7）。これにより、注視エリアを動かしたときブラーのかかっていたエリアのうち再度ブラーのかからないエリアが3つ存在するため、ボールが注視エリアの境界線付近を動く際にも、モデルがボールを追いやすく、認識しやすくなっている。

先ほどの実験の結果も含めて全てプロットしたグラフを図8に示す。濃い緑で示したものは重なった視野エリアを動かすことができるもので、この点数も標準の設定を超える結果となった。このモデルがプレイしている様子を観察すると、ボールが視野の範囲に収まるようにエリアを移動し、跳ね返す場面を確認することができた。エリア分割の実装を変更したことにより、「ボールが各エリアの境界線付近を動く際、突然強いブラーのかかるエリアに入ってしまう、エージェントの入力から消えてしまう」現象が抑制されたことが、性能向上の一因ではないかと推測している。

この結果から、エージェントに視野の制約を課したうえで、ゲーム内行動と視野移動行動の両方を学習させることが可能であると考えられる。

5. FightingICE を題材とした実験

前章では、Breakout を題材に視野の移動を含めた学習が可能であることを示した。本章では、これが Breakout に特有の現象かどうかを確かめるため、他ジャンルのゲームを題材として行った実験について記載する。本実験では、格闘ゲームプラットフォームである FightingICE に対して、前章と同じような実装で視野を導入し強化学習させた。先述した通り、格闘ゲームには視野による制約を前提とした

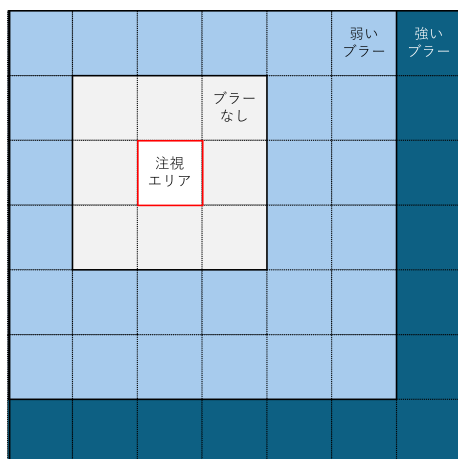


図 7 実験 3 でのブラーのかけ方

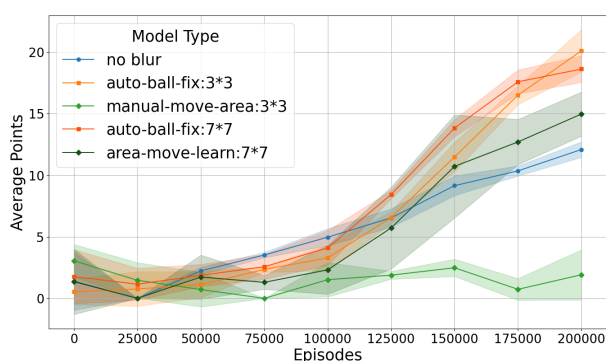


図 8 画面を複数エリアに分割して行った実験

行動がいくつか知られている。これを再現することが本研究の目標であるが、まずは視野を含めた学習の確認を本実験の目的とした。

5.1 実験設定

本実験は、音声データのみを入力とする強化学習サンプルである、TeamFightingICE の BlindAI [6] をベースに実装を行った。BlindAI は画像情報を使用せず、音声データのみを入力として学習を行う AI である。視覚障害のある人間プレイヤーにとって強すぎない程度の強さに制限された MctsAi を固定の対戦相手とし、Proximal Policy Optimization(PPO) を用いて訓練される。いくつかのエンコーダを試し、弱体化された MctsAi に対して音声情報だけをを入力として勝率 0.63、平均体力差 37.07 を記録した。

これをベースに、入力形式を音声データから画像データに変更して実験を行った。画像を入力とする深層強化学習のため、15 フレーム遅れた画面情報が渡される設定を無視することができるが、生物学的制約の中に「遅れ」が存在するため、この実験では実際よりも 15 フレーム遅れた画像を取得する。各フレーム時点で受け取るグレースケールの画像は、128×192 にリサイズしてからブラー処理を行っている。FightingICE から取得できるディスプレイ画像の解像度は本来 64×96 であるが、キャラクターがブラーをかける

前からつぶれて表示されてしまうため、元の 640×960 で画像データを取得できるよう書き換え、128×192 にリサイズした。エリア分割は画面の縦横比に合わせて 6×9 とし、ブラーの強さや範囲は Breakout と同じ設定を適用した。図 9 は実際に入力となるぼかし済みの画像に対して、注視エリアをあらわす白い枠と、弱いブラーがかかっている範囲を表す黒い枠を付けたものである。画面左に白く映っているものが操作キャラクターで、画面右は相手キャラクターである。ネットワーク構造も Breakout 時と同じ、CNN による畳み込み層と全結合層を採用している。対戦相手を含むほとんどのパラメータは BlindAI の実装 [7] を参照した。ただし、報酬は各フレームごとに、相手に与えたダメージと自分の減った体力の差を、1 回の最大攻撃力で割り正規化したものに変更した。

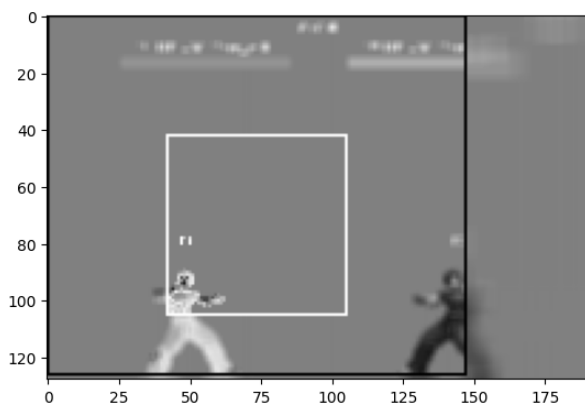


図 9 画面左の操作キャラクター付近を注視していた際の入力画像

5.2 現時点での結果

RTX4090 で約 1 日学習させたものを、BlindAI と同じく制限された MctsAI と 30 試合させたところ、ラウンド勝率は 0.6、平均体力差は 27.8 となった。現時点では、制限された MctsAI に勝ち越すことができる AI は学習させることができたものの、いずれの設定においても視野の移動を含めた学習ができているようには判断できなかった。エージェントは隙の少ない攻撃のみ連続で選択する行動を取っており、視野の移動や状況に応じた戦略的な行動は見られなかった。弱い MctsAI に対しては画面から得られる情報が極端に少なくとも勝ち越すことができる可能性や、反対に強いブラーを Breakout 時にボールが消えるよう設定したため、今回はキャラクターの位置が判別できてしまい、十分に情報を制限できなかった可能性が残されている。まだ各種パラメータの設定が不適切である可能性を考慮し、FightingICE での視野を含めた学習の確認を目的とした実験を発表までに行う予定である。

6. おわりに

本研究では、視野による制約を強化学習に組み込むことで、ゲーム AI に制約を前提とした行動を獲得させることを目的に、視野の移動を含めた学習が可能かどうかを検証した。Breakout を題材とした実験では、実際に視野を含めた学習が確認できただけでなく、視野制限がある場合の方が、序盤の学習成績が良いという結果を得た。一方で、FightingICE を用いた実験では、まだ視野を含めた学習が確認できていない。今後は他ジャンルでの視野を含めた学習の確認や、制約を前提とした行動の獲得を目指したい。

されることは少ないと考えるが、人間は、注視すべき場所を減らすことができる合理的な戦略として採用することがある。

参考文献

- [1] Berner, Christopher, et al., “Dota 2 with large scale deep reinforcement learning.”, arXiv preprint arXiv:1912.06680 (2019).
- [2] 安原直宏, “『ストリートファイター 6』初心者から上級者まで対応した人間らしい有機的な行動を行う CPU (AI) のしくみ”, CEDEC2024(2024), https://cedil.cesa.or.jp/cedil_sessions/view/2946
- [3] 藤井叙人, et al., “生物学的制約の導入によるビデオゲームエージェントの「人間らしい」振舞いの自動獲得”, 情報処理学会論文誌 55.7, pp.1655-1664, 2014.
- [4] 平井弘一, Grimbergen Reijer, “弾幕の認識に人間の視覚特性を取り入れたシューティングゲーム AI の研究”, ゲームプログラミングワークショップ 2016 論文集, No. 2016, pp. 158–161, (2016).
- [5] 小松原 (2011), 「ヒューマンエラー 第 2 版」, 丸善出版株式会社
- [6] Van Nguyen, Thai, et al., “A deep reinforcement learning blind AI in DareFightingICE.”, 2022 IEEE Conference on Games (CoG). IEEE, 2022.
- [7] <https://github.com/TeamFightingICE/BlindAI>, (アクセス: 2024/10/15)

付録：2D 格闘ゲームにおける行動「置き」

2D 格闘ゲームにおいて、相手の目の前で、当たることを前提としないで通常技を振る行為を「置き」と呼び、これがヒットしていた時に必殺技などが連続でヒットするタイミングでコマンドを入力する行為を「仕込み」と呼ぶことがある。2D 格闘ゲームにおけるほとんどの行動は、見ながら対処することが原理的には可能であるが、人間が相手の行動すべての行動に対して準備し、適切に対応することは現実的ではない。例えば地上からの急な接近やリーチの長い通常技、飛び道具での攻撃やジャンプからの攻撃に対して、適切な対応のために画面上で注視すべき場所はそれぞれ異なる。視野の制約により、それら全てを警戒し全てに対して最適な行動を取ることは困難なため、これらの攻撃のいくつかに対抗できるよう技を置きつつ、画面上方に注視点移動させ、相手のジャンプ攻撃に備える戦略が存在する。対戦相手に読まれた時には置きの際に反撃を受ける可能性があるため、視野による制約がなければ採用