

計器情報と視覚情報に基づく 熟練者ノウハウの学習による航空機操縦支援

Aircraft Operation Support by Learning Expert Know-how
Based on Instrument Information and Visual Information



筈井 祐介*1
Yusuke Hazui

別府 啓史*2
Hirofumi Beppu

小野 仁意*3
Hitoi Ono

土屋 武司*4
Takeshi Tsuchiya

航空機の操縦士育成には多大なコストと時間が必要であり、訓練効率化の重要性が高い。効率的な技量向上には熟練者による適切な指導が必要となる一方で、操縦士の高齢化に伴う熟練者不足が課題となっている。そこで、訓練に利用することを目的に、熟練者が操縦に際して着目している計器情報や視覚情報に基づき、操縦ノウハウを学習する逆強化学習手法を開発した。本技術を航空機の着陸問題に適用し、熟練者の着陸技術を習得できることを確認した。さらに、得られた操縦モデルを活用したパイロットの教育システムに関しても言及する。

1. はじめに

近年の日本における少子高齢化問題は、様々な業界に影響を与えている。中でも育成に時間と費用を要する航空機パイロットの不足は深刻で、効率的な訓練によるパイロットの早期育成は航空業界の課題の一つである。通常、座学を終えた訓練生は、実機での飛行訓練の前にシミュレータを用いて離着陸や高度維持、旋回等の基本的な操縦操作を体感、習得する。この段階において、熟練者である教官による適切な指導が早期育成の鍵となるが、指導に十分な人材と時間を確保することが難しくなっている。そこで、教官に代わり適切なアドバイスを自動生成でき、訓練生による自学可能な訓練支援システムの開発・適用が課題解決に繋がると考え、熟練者の操縦を模倣する熟練者 AI を生成し、熟練者と訓練生との差異を明示し訓練生に気づきを与える評価指標を構築した。本報では、着陸操作を一例とし、熟練者の操縦により得られた計器情報と視覚情報を使い、強化学習の一手法である敵対的逆強化学習手法を用いて、熟練者 AI と評価指標となる報酬関数を同時に学習⁽¹⁾した。図1に、本技術のフレームワークを示す。

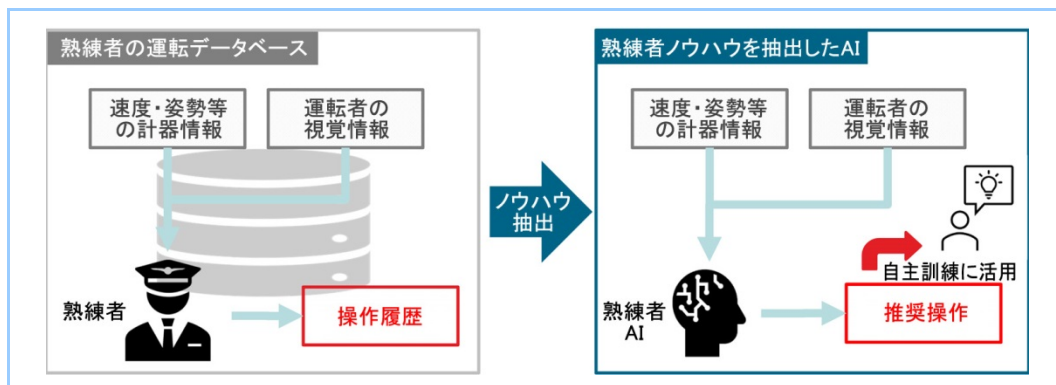


図1 本技術のフレームワーク

*1 デジタルイノベーション本部 CIS 部

*3 デジタルイノベーション本部 CIS 部 首席技師 工博

*2 デジタルイノベーション本部 CIS 部 工博

*4 東京大学 教授 工博

以降、2章では開発手法の概要を示し、3章では航空機の着陸問題に開発手法を適用した結果について述べる。4章にて航空機の訓練支援への開発手法の適用可能性について言及し、5章でまとめを述べる。

2. マルチモーダル情報に基づく逆強化学習技術

2.1 強化学習と逆強化学習

動的に変化する環境下で状況に応じた操作方法を自律的に学習する手法として、強化学習が知られている。強化学習は機械学習手法の一つであり、報酬関数と呼ばれる操作を評価する指標を事前に設計し、得られる報酬が最大となるような操作方法を学習する。

一方で、強化学習を航空機操縦の学習に適用するにあたり、熟練者の評価基準を定式化した報酬設計が容易ではないこと、熟練者が想定し得ない操縦方法は現場では受け入れられ難いこと等の課題がある。

これらの課題に対し、熟練者の操縦データを用いることで、操縦方法を習得すると同時に報酬関数の推定も可能な技術として、逆強化学習がある。ある状況下における熟練者操作を覚える手法としては教師あり学習が有名であるが、運用したい全ての状況を含む熟練者操作データを用意する必要がある。一方で、逆強化学習は熟練者操作データ取得時と異なる状況においても、推定した報酬関数を用いて新たに学習を行うことにより、報酬関数を設計することなく、熟練者操作データが存在しない状況に対しても学習が可能となる。その結果、少ないデータや試行回数で広範な状況に適用可能な操作方法を学習できる。図2に強化学習と逆強化学習の構成の違いを示す。

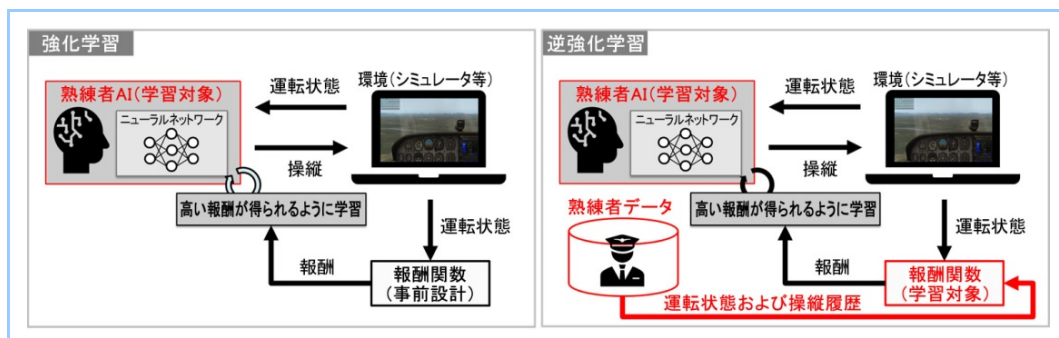


図2 強化学習及び逆強化学習の概要

2.2 マルチモーダル情報に基づく逆強化学習

通常、人は計器による計測値や視覚情報などの複数種類の情報に基づいて、意思決定に必要な状況認識を行う。例えば航空機操縦において熟練者は、計器の数値、コックピット窓外の視覚情報、エンジン音、体感重力などに基づいて意思決定を行っている。そのため、熟練者AIを学習するためには、これらマルチモーダルな情報を扱える逆強化学習手法が必要である。2.2節では、開発したマルチモーダル情報に基づく逆強化学習技術について述べる。

図3に開発した逆強化学習手法の構成図を示す。本報ではベースとなる逆強化学習手法として、深層学習技術の一つである敵対的生成ネットワークを取り入れることで、より複雑な行動を学習できるAIRL(Adversarial Inverse Reinforcement Learning⁽²⁾)を採用した。

敵対的逆強化学習は、生成器と識別器の2つのニューラルネットワークで構成される。生成器では、強化学習により熟練者の操縦を学習するとともに熟練者AIによる操縦データを生成する。一方で識別器では、熟練者AIが生成した操縦データ(贋)か、学習データとして用意した熟練者の操縦データ(真)かを識別する。生成器は、識別器を騙そうとより学習データに近い操縦を学習するのに対し、識別器は生成器に騙されないように真贋をより見分けられるように学習する。生成器が熟練者の操縦を学習することができれば、識別器は生成器による操縦と熟練者の操縦を分

類できなくなり、結果として生成器が熟練者の行動を模倣する学習ができたことになる。

またマルチモーダル情報は、シミュレータ環境から取得されるセンサ値（計器による計測値）と画像情報で構成される。画像情報にCNN (Convolutional Neural Network) や幾何学変換等の前処理を行うことで、センサ値と画像情報を合成することを考えた。熟練者は、視覚情報として得られる滑走路の形状を利用して、航空機の高度や現在位置から滑走路までの距離を認識する。これに対して、本報では図4に示す滑走路幅と滑走路辺のなす角度の視覚情報（ディスプレイ上での見え方）への幾何学的変換⁽³⁾により、学習に必要な視覚情報への変換を行った。

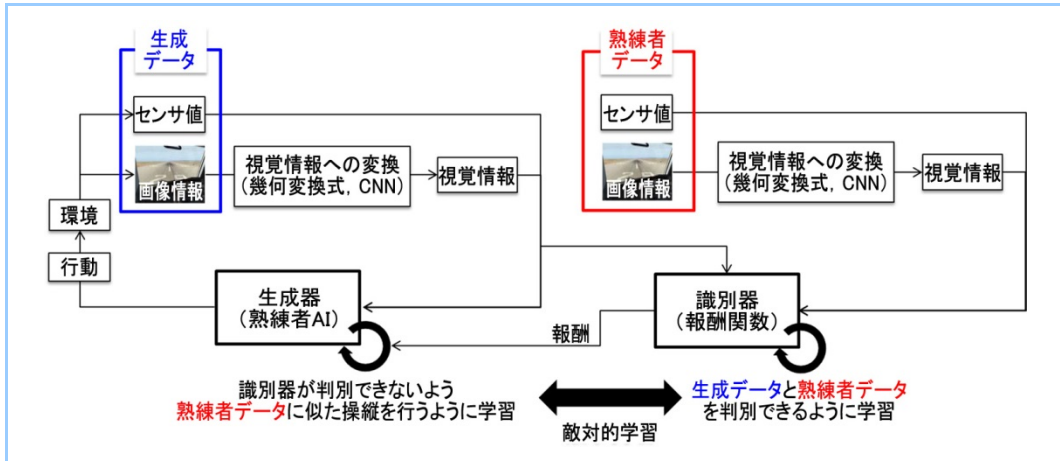


図3 開発手法の構成

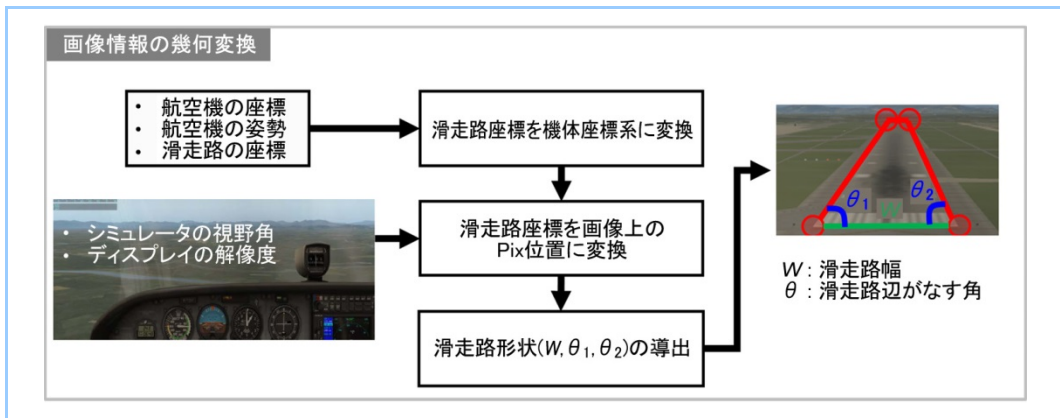


図4 視覚情報への変換例

3. 航空機着陸問題への適用

3.1 問題設定

市販フライトシミュレータ X-PLANE11⁽⁴⁾上において、滑走路が視野に入る航空機位置及び姿勢を初期状態とし、高度や対気速度に関する条件を満たしながら、航空機を滑走路に着陸させることを考える。初期状態におけるコックピットからの視界を図5に示す。また、飛行中には向かい風による外乱が発生し、所望の状態に着陸するために適宜航空機を制御する必要がある。ここでは、エレベータとスロットルによる高度制御及び速度制御を学習対象とした。

なお、2.2 節に記載の開発手法にて学習を行うため必要な熟練者による着陸操縦データは、図6に示すような市販のヨーク（操縦桿）及びスロットルを用い、実際に熟練者が X-PLANE11 上で着陸操縦を行うことで収集した。



図5 X-PLANE11 におけるコックピットからの視界



図6 熟練者データ収集に用いた機器の構成

3.2 学習結果

開発手法の学習プロセスを監視するために、許容高度や対気速度等の着陸成功のために満たすべき要件を満足する飛行により高いスコアを与える評価関数を設計した。本評価関数値と学習エピソードの関係を図7に示す。図7から試行回数が増加するにしたがい、開発手法による学習がうまく進み、207回の学習を経ることで、最終的に着陸操縦を学習できていることがわかる。なお207回の学習にかかった時間は17時間強であった。

学習した熟練者 AI(方策関数)によって着陸操縦を行った結果を図8に示す。学習に使用した熟練者データは赤線、開発手法での結果を青線で示しており、許容高度及び対気速度に関する要件を満たすことが確認できた。

また、エレベータとスロットルの操作履歴は熟練者と開発手法で類似している。例えば、風外乱の発生直後に、航空機のピッチ角や対気速度を維持するためエレベータ及びスロットルを低減する動作は両データで共通である。さらに、風外乱により不足する推進力を補うため、スロットルを初期設定値と比較して増加させる動作についても共通している。特にスロットル操作は単純に対気速度を増減させるだけでなく、航空機の姿勢にも影響があり、高度変化やさらなる対気速度変化を引き起こすことから、初心者にとって習得が難しいことで知られている。本検証の前に行った従来の強化学習手法による試行では、熟練者と類似したスロットル操作を獲得することはできなかった。以上より、開発手法によって熟練者の着陸操縦を模擬できる熟練者 AI を学習可能であることが確認できた。

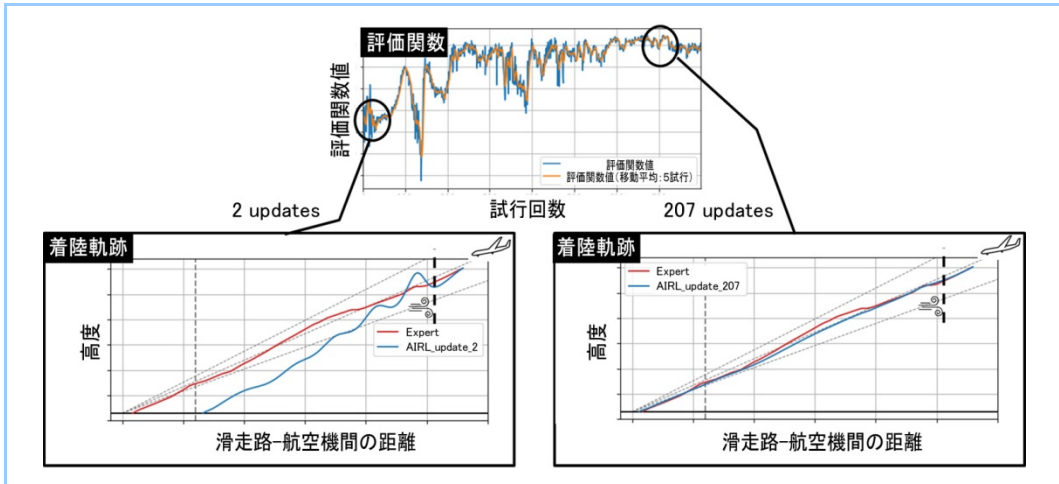


図7 評価関数と試行回数の関係

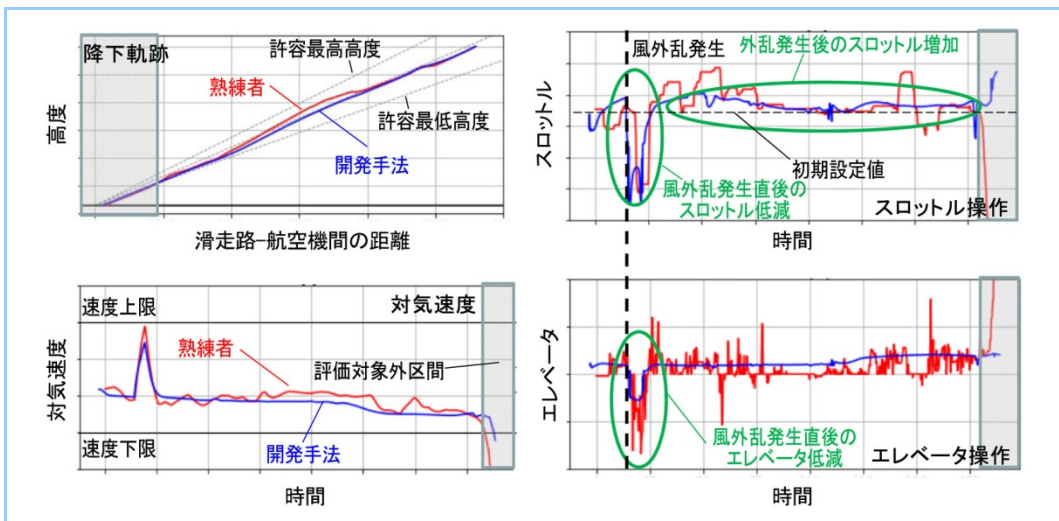


図8 熟練者 AI による着陸操縦データ

3.3 熟練者 AI のロバスト性検証

3.2 節で評価した熟練者 AI は、風外乱の発生高度及び風速を一定とした条件で学習を行った結果得られたものであり、このとき学習に用いた熟練者データは同条件における着陸操縦1回分のデータである。一方で、熟練者 AI はパイロット教育に使用されることを想定しており、様々な航空機状態や風外乱条件に対して適切な操作を提供できることが求められる。つまり、熟練者 AI は環境変化に対してロバストである必要がある。そこで得られた熟練者 AI のロバスト性を検証するため、風外乱の条件を変更した場合における挙動を評価した。

具体的には、風速と風外乱の発生タイミングを変更した。風速を変更したときの挙動を [図9](#) に、風外乱の発生タイミングを変更したときの挙動を [図10](#) に示しており、両方のケースにおいて必要な要件を満たしつつ着陸に成功することを確認した。この結果により、想定する全ての条件における熟練者データを事前収集しなくとも、環境変化にロバストな熟練者 AI を生成できることが確認できた。

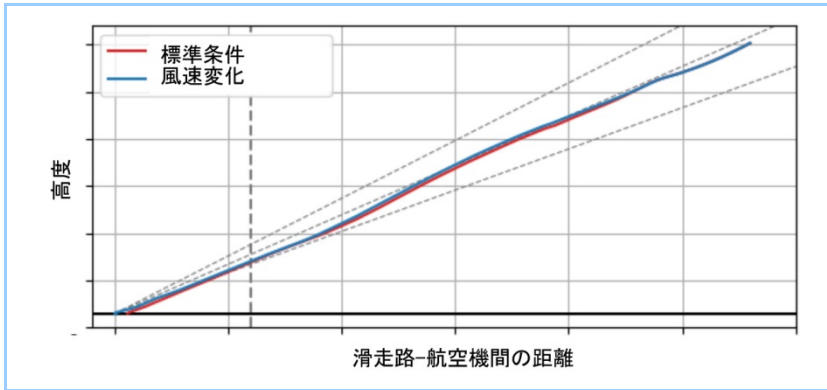


図9 風速変更ケースにおける熟練者 AI による着陸軌跡

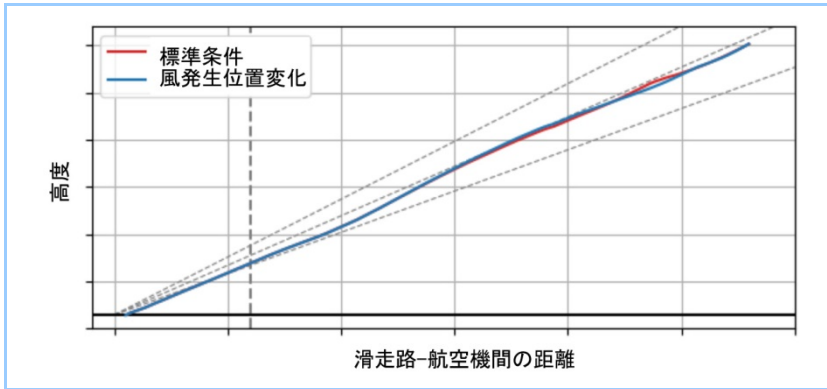


図 10 風発生タイミング変更ケースにおける熟練者 AI による着陸軌跡

4. 将来の飛行教育システムへの適用に向けて

将来的には、教官パイロットが訓練生の練習をサポートするかのようにより、熟練者の推奨操作や意図をフィードバックしながら学ぶことができる自己訓練システムの実現を目指している。ここでは、熟練者 AI による推奨操作と飛行状態価値の評価を利用した試作システムを紹介する。

図 11、図 12 は訓練生の航空機シミュレータ上での飛行操縦に対するフィードバック動画のキャプチャである。図 11 左のcockpit視点の画像は航空機シミュレータ画面を、右下の図は訓練生と熟練者の軌跡の比較を示しており、赤い点は訓練生の現在の飛行位置である。また、右上の図の緑色のバーは訓練生の現在のエレベータ（操縦桿）とスロットル（出力）の制御量を示し、赤色の線は熟練者 AI が出力する推奨操作を表している。例えば図 11 の状態において、熟練者 AI は操縦桿を押すべきと推奨しているが、訓練生は操縦桿をニュートラル位置にしており、好ましくない操縦である。また、スロットルについて、熟練者 AI は出力低減を推奨しているが、訓練生は低減操作できていない。図 12 左上図と左下図は図 11 の図と同様であり、中央下の状態価値ヒートマップは熟練者 AI にて算出した飛行状態の価値を示している。この状態ヒートマップでは、価値の高い領域が熟練者 AI の目指す状態を意味している。また、赤い丸とバツ印はそれぞれ熟練者の飛行状態と訓練生の飛行状態を示しており、訓練生は時々刻々と変化する状況に対して、どのような飛行状態を目指せば良いかを学ぶことができる。例えば図 12 の状態において熟練者 AI は、現状の飛行状態よりも対気速度及びピッチ角を低減した状態が好ましいと算出しており、その状態へ遷移するために必要な操作を推奨操作として出力している。以上のような情報を訓練生にフィードバックすることで、訓練生の操縦矯正に繋がれる可能性があると考えており、今後検討を進めていく。



図 11 訓練生の飛行操縦に対するフィードバック動画のキャプチャ①

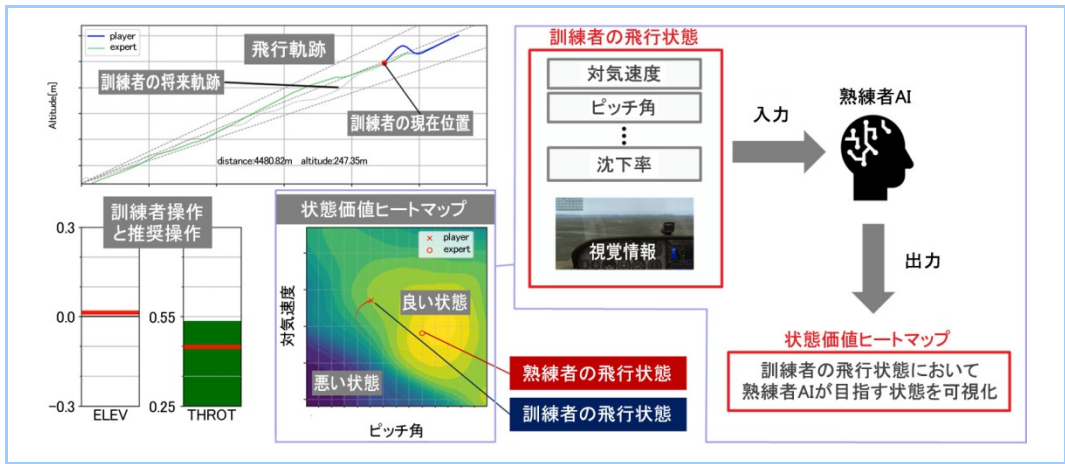


図 12 訓練生の飛行操縦に対するフィードバック動画のキャプチャ②

5. まとめ

本報では、計器情報と視覚情報からなるマルチモーダルな情報を組み込むことにより、熟練者のノウハウを学習するための逆強化学習手法を開発した。また、開発手法を航空機着陸問題に適用し、熟練者と同様の操縦で航空機を制御できるロバストな熟練者 AI が生成できることを確認した。さらに、飛行教育を支援する試作システムを紹介し、熟練者 AI の推奨操作と状態価値ヒートマップを利用した支援のユースケースを示した。

今後は、実際の操縦で求められるより複雑な問題に対応するため、階層的な逆強化学習技術や、より高い説明性を有した手法の開発を進めるとともに、飛行教育システムへの本技術適用を検討していく。

参考文献

- (1) H. Beppu, et al., Learning of Piloting Skills for Aircraft by Adversarial Inverse Reinforcement Learning with Multimodal Information, Proceedings of 2023 62nd Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 2023
- (2) J. Fu, et al., Learning Robust Rewards with Adversarial Inverse Reinforcement Learning, International Conference on Learning Representations, 2018
- (3) 森亮太ら, パイロットの着陸操縦分析手法に関する研究, 東京大学 博士論文, 2009
- (4) Laminar Research, X-PLANE 11 Professional, 2017, <https://www.x-plane.com/product/x-plane-11-professional-use-digital-download/>