

# 人と協調するAI

Cooperative AI for Human - Machine Work

松永沙織\*  
Saori Matsunaga  
森本卓爾\*  
Takuji Morimoto  
三塚由浩†  
Yoshihiro Mitsuoka

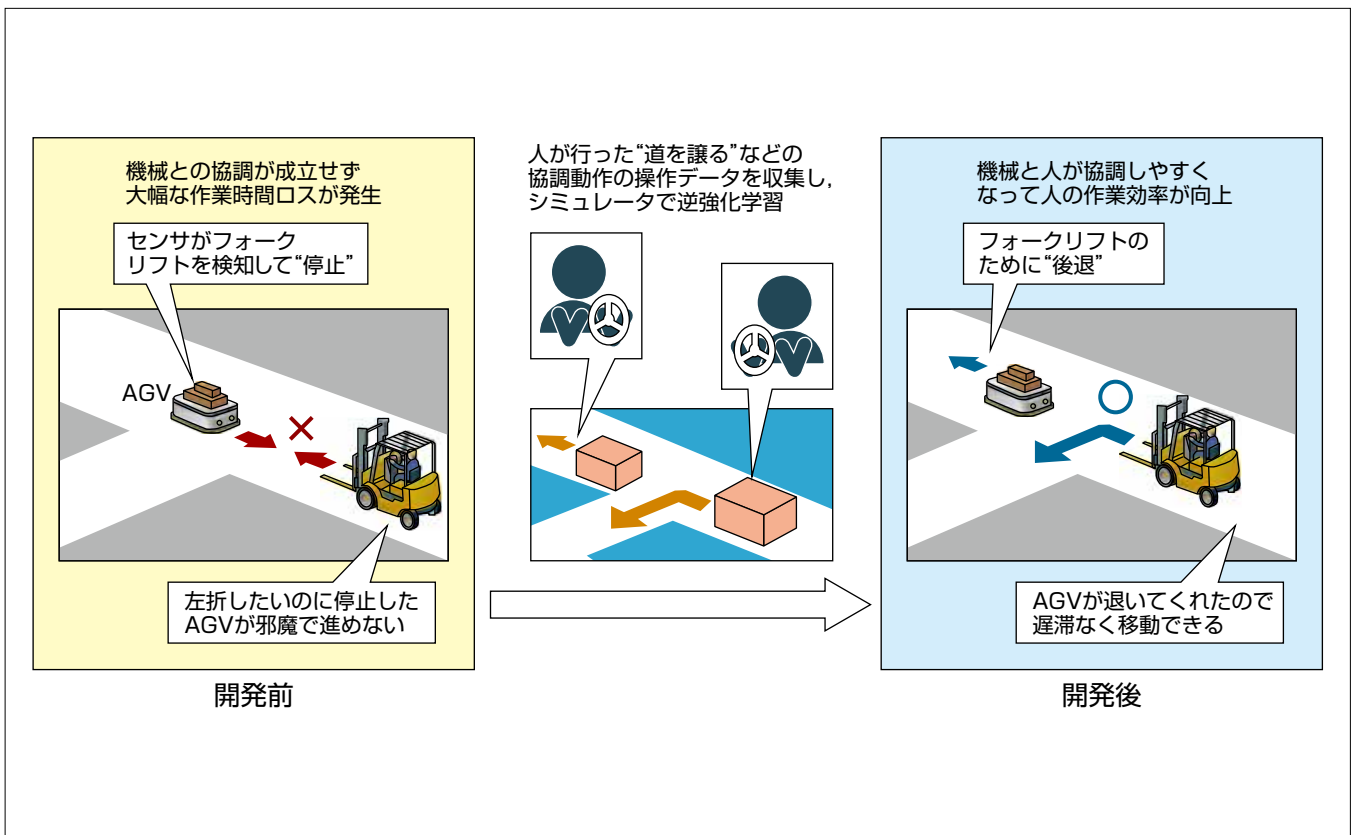
佐藤 匠\*  
Takumi Sato  
穂山利貞‡  
Toshisada Mariyama

## 要 旨

人手不足や作業者間の距離確保ニーズによって、人と機械が混在する作業環境がより一層普及すると考えられる中、人と機械が互いに連携して安全かつ効率的な動作を行うことが求められる。AIを用いた機械の制御手法の代表例として深層強化学習が挙げられるが、深層強化学習に必要な報酬関数は、複雑なタスクになるほど設計が困難になる。一方で、人同士であれば、相手の動きに合わせて自身の行動を決定することで相手の行動を妨害しないような協調動作を実現できる。そこで“人と協調するAI”では人の動作や操作をAIによって模倣することで、人のような自然な

協調動作を可能にすることを目的とする。模倣学習の中でも逆強化学習と呼ばれる手法を使用し、人手による動作の設計なしに、少ない教師データ数で効率的な学習を実現している。

無人搬送車(Automated Guided Vehicle : AGV)を模した小型の自律移動体にこの技術を適用した結果、教師あり学習に対して1/10以下のデータ量で学習可能であることを確認した。また、入力画像の工夫、制御モジュールとの組合せによって、従来困難であった逆強化学習による実機でのリアルタイムな協調動作を実現した。



## “人と協調するAI”を適用した場合の動作例

AGVへの適用事例を示す。単純なルールで動作させる場合、他車両(フォークリフト)が危険領域に進入した時点でAGVが緊急停止し、お見合いが発生する。一方“人と協調するAI”を適用した場合、AGV側がフォークリフトへ進路を譲る動作を見せて、遅滞のないスムーズな移動を実現する。

## 1. ま え が き

現在、少子高齢化による人手不足などに対応するために、人に代わる機械の導入や、AIによる機械の知能化が進んでいる。例えば生産や物流の現場では、AGVの導入が加速している。省人化や作業員間の距離を確保することが望まれるニューノーマル社会では、人と機械の混在する環境がより一層増加すると考えられる。現在は安全確保のために人と機械の作業空間が分かれていることがほとんどであるが、今後は更なる効率化のため、人と機械がより連携することが求められる。そこで、当社のAI技術“Maisart”を活用し、人のような自然な動作を模倣学習することで人との協調を容易にする“人と協調するAI”を開発した。逆強化学習と呼ばれる手法を使用し、少ないデータ数で効率的な学習を実現している。

本稿では、AGVを模した小型の自律移動体に適用した例を基に、“人と協調するAI”について述べる。

## 2. 人と協調するAI技術

### 2.1 全体の構成

人と協調するAI技術の全体構成を図1に示す。今回、人が行う協調動作の操作データを“逆強化学習”(2.2節)で模倣することで、自然な協調動作を実現している。逆強化学習

学習では繰り返しの試行錯誤が必要になるため、適用する実環境に対応したシミュレーション環境を用意する必要がある。そして、シミュレータや実機から出力されたセンサーデータは適切な前処理を行って学習しやすい形へ変換する必要がある。また実機の制御に当たっては、機器ごとに適した制御周期や制御手法を採用しなければならない。通常、自律移動体の大まかな経路の決定に関しては高周期の推論は不要であり、百ミリ～数百ミリ秒周期でも問題ない。一方で自律移動体の制御に当たっては数ミリ～数十ミリ秒周期でのリアルタイム制御が不可欠になる。このような背景から、入力データを変換する“俯瞰(ふかん)図生成器”“特徴量抽出ネットワーク”，大まかな目標動作を決定する“学習器”“学習済みモデル”，きめ細かな制御を行う“制御モジュール”といった各モジュールに機能を分割している。制御モジュールでの具体的な制御手法は機器ごとに最適なものを選択すればよいため、AIの適用以前から培ってきた制御技術を生かすことも可能である。高速で滑らかな制御と難しい判断を実環境で両立させることが可能な構成になっている。

次に、学習器及びその前段の俯瞰図生成器と特徴量抽出ネットワークについて述べる。

### 2.2 学習器(逆強化学習)

複数の人(又は人の操作する機械)が同じ環境で作業をするとき、人は相手の様子を見て進路や順番を譲ったり、作業スピードを調整したりしながら協調して作業を行うこと

で、互いの安全性と効率性を両立させている。そのためAGVなどの機械と人が共存する環境下でも、機械側が人に進路を譲ったり、周囲の人に注意して徐行したりすることが望ましい。しかしながら、現状では“前に進む”“障害物を検知すると止まる”といったルールで動作するため、場合によっては互いに身動きが取れなくなってお見合い状態になるなど、双方の作業効率が低下する可能性がある。起こり得る事象一つ一つに対応する行動のルールを

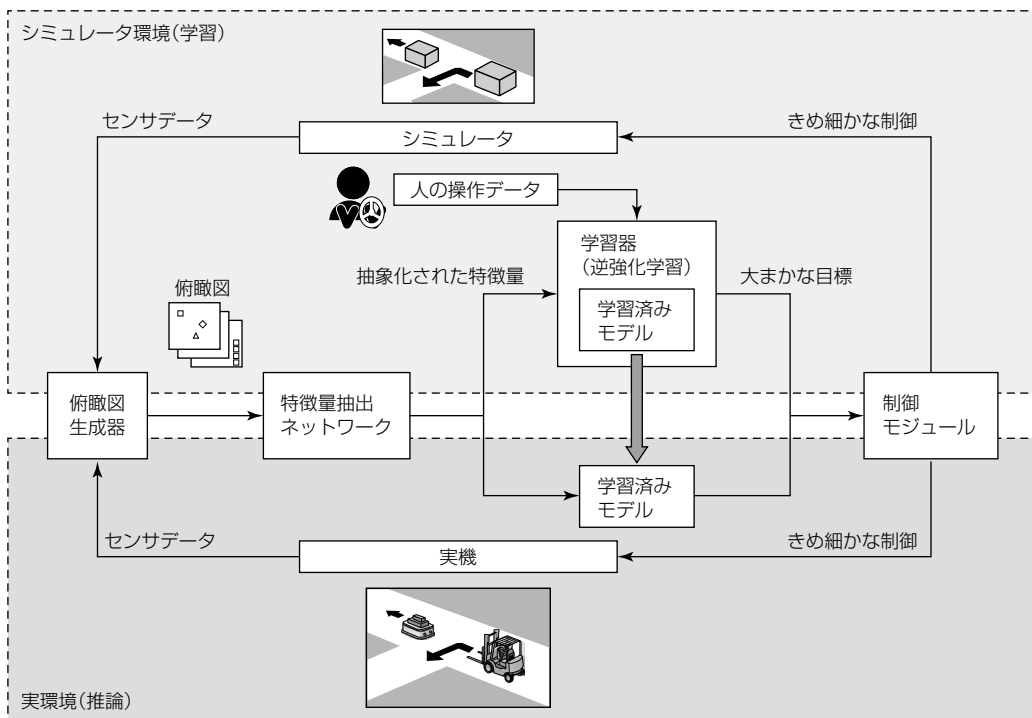


図1. 人と協調するAI技術の全体構成

全て書き下して協調作業を実現することも考えられるが、全てのルールを網羅することは困難である。そこで、AIを活用した人との協調の実現を考える。AIを用いた手法として強化学習を適用する例が挙げられる。強化学習では、報酬関数と呼ばれる、ある状態に対する行動の良しあしを表す関数を設計する必要がある。しかし、周辺環境が複雑に変化する自律移動体ではその報酬関数の設計自体が困難になる。一方で、報酬関数を人手で設計するのではなく、見本になる行動を与えてそれに近づくように振る舞い方を学習する“模倣学習”という手法がある。このうち、見本のデータ(以下“デモンストレーションデータ”という。)を教師データとして教師あり学習で行動を学習する場合、大量のデモンストレーションデータを必要とするという側面がある。これは、学習データに含まれない状態では適切な行動を決定できないため、制御中に発生する誤差や誤差の蓄積も考慮しながらあらゆるデータを準備する必要があるからである。一方、デモンストレーションデータから報酬関数を推定し、推定した報酬関数を用いて強化学習を行って最適な行動を獲得する手法は、逆強化学習と呼ばれる。与えられたデモンストレーションデータの入力情報と行動のマッピングを得るのではなく、デモンストレーションデータを基にシミュレーションで自ら試行錯誤を繰り返しながら報酬関数を推定する。逆強化学習の手法の一つである敵対的模倣学習(Generative Adversarial Imitation Learning: GAIL)<sup>①</sup>は画像生成分野で注目されている敵対的生成ネットワーク(Generative Adversarial Network: GAN)<sup>②</sup>の要領で最適な行動を学習する方法であり、教師あり学習に対して少ないデモンストレーションデータ数で学習できることが報告されている。こうした背景から、この技術ではGAILをベースにした手法を使用している。

### 2.3 俯瞰図生成器と特徴量抽出ネットワーク

AIを用いた移動体制御では画像入力が用いられることが多い。自車の状態も周辺物体の状態も時々刻々変化する走行環境では、自車近傍の物体を選択して数値ベクトルを生成するよりも、自車周辺の画像として表現した方が複数物体の位置関係を容易に表現できることが理由の一つである。しかし通常、得られた画像には自車の行動決定に無関係な情報が含まれている。またシミュレータを活用した試行錯誤による学習を前提とする場合、学習時に得られる画像と運用時に得られる実画像とのギャップが課題になる。そこで“俯瞰図生成器”によってシミュレータや実機から得られる情報をバーチャル画像に変換することで、不要な情報をそぎ落としながら実環境とシミュレーション環境のギャップを抑えるという方法を取る。このとき、進行方向前方の画像ではなく上から見下ろした俯瞰図にすることで、

移動体自身と周辺物体との位置関係を簡潔に表現している。俯瞰図導入によって周辺情報はシンプルな表現になったが、画像データの次元数は大きく、学習の際に学習速度や安定性といった面でのネックになり得る。そこで“特徴量抽出ネットワーク”を導入し、得られた俯瞰図を更に圧縮して低次元の特徴量を抽出する。ここでは、VAE(Variational AutoEncoder)を使用して低次元ベクトルへと変換している。VAEについても学習が必要になるが、周辺物体との位置関係を様々に変えた俯瞰図を人工的に大量に作り出すことで、逆強化学習による自車行動の学習とは別に、特徴抽出器としての学習をあらかじめ完了させておくことが可能である。

最終的に、VAEのエンコーダによって抽出された俯瞰図の特徴量情報に、俯瞰図に含まれない情報(自車速度等)を結合したものが、学習器又は学習済みモデルに入力される。

## 3. 小型自律移動体への適用

### 3.1 シナリオ

最後に、AGVを模した小型自律移動体へこの技術を適用した際の実験結果について述べる。今回はAGVの譲り合いのシーンを想定した図2に示すシナリオで実験を行う。図中の実線のように、自車AGVは図の右側に向かって直進走行している。このとき、フォークリフトと衝突したりフォークリフトの作業を邪魔したりすることのない範囲で、できるだけ早く右端まで到達することが目的である。一方人が運転するフォークリフトは、図中の点線で示すとおり、AGVの走行経路に重なるように後退して切り返した後、同じく右側へ向かって走行する。そのため、“前に進む”“障害物を検知すると停止する”といったルールでAGVを走行させた場合、フォークリフトの割り込みタイミングによっては図3のようにお見合い状態が発生して互いの作業効率が低下するシナリオになっている。今回は、フォークリフトがAGVの進路に割り込むタイミングを、①AGVの後退が必須になる急な割り込み、②AGVの後退が不要な余裕を持った割り込みの2シナリオ用意し、①②の二つのシナリオを混合して学習・評価する。①②のどちらのシナリオでも、車両の出発タイミングや出発位置はその都度微小に変化するものとする。

### 3.2 シミュレータでの学習

独自に構築したシミュレータを使用し、3.1節で述べたシナリオで協調動作を学習した。今回は、自車位置を画像中心とした図(実環境で2m四方に相当)の中に、他車だけを描画した二値画像を俯瞰図として使用している。特徴量

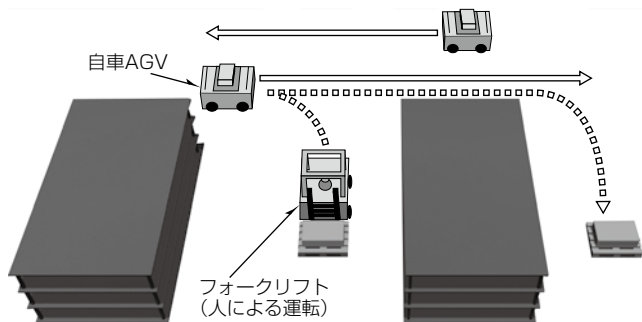


図2. 実験シナリオ

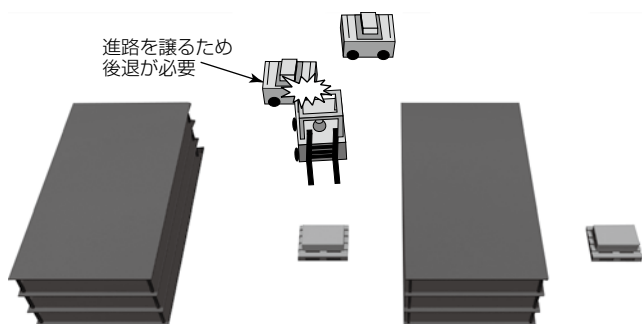


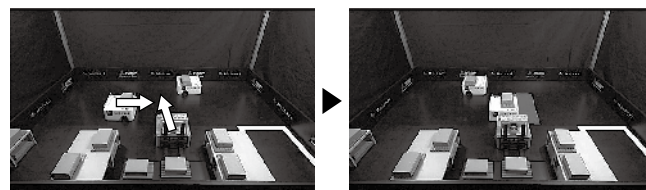
図3. 協調動作の必要なシーン

抽出ネットワークで俯瞰図を16次元のベクトルへと変換し、自車速度情報と合わせた計17次元のベクトルが学習器へと入力される。

逆強化学習に使用するデモンストレーションデータは、3.1節の①②の各シナリオで人がシミュレータ上のAGVを操作して収集している。これらのデモンストレーションデータを用いて逆強化学習と教師あり学習を行い、衝突の有無や作業完了に要した時間から算出したスコアを比較した。その結果、今回のシナリオに限定すれば、逆強化学習では①②の各4個以上のデモンストレーションデータがあれば人と同等以上のスコアを獲得できる一方、教師あり学習ではその10倍以上のデモンストレーションデータを使用しても逆強化学習使用時のスコアには届かず、かつスコアのばらつきも大きいことが分かった。逆強化学習を用いたこの手法は必要なデータ数、動作の安全性や効率性、安定性という点で教師あり学習よりも優位であることが確認できた。

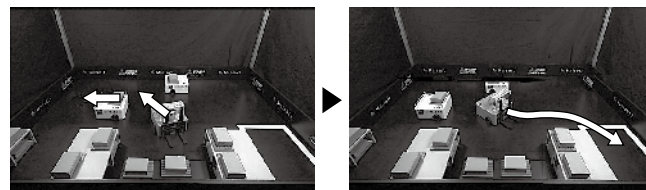
### 3.3 実機での実験

最後に学習結果を実機へ適用した際の実験結果を示す。図4に示すとおり、図2のシミュレーション環境と同等のAGV、フォークリフト、周辺物体を再現している。自車



フォークリフトが緊急領域に進入したためAGVが緊急停止

(a) “人と協調するAI”非適用



AGVが道を譲ることでフォークリフトが遅滞なく移動できる

(b) “人と協調するAI”適用

図4. 実機実験の様子

AGVには、デモンストレーションデータ20個を使用して学習したモデルを、追加学習や調整なしにそのまま適用した。

この技術を適用せず、“前に進む”“障害物を危険領域内に検知すると停止する”というルールで動作する場合、フォークリフトが危険領域に進入した時点でAGVが緊急停止し、作業時間ロスが発生している(図4(a))。一方、この技術を適用した場合、AGVが後退してフォークリフトに進路を譲ることでフォークリフトが遅滞なく移動でき、スムーズな作業の実現に貢献していることが分かる(図4(b))。

## 4. むすび

今後ますます人と機械の混在が増加すると考えられる中、人の模倣によって協調作業を実現する“人と協調するAI”を開発した。逆強化学習によって教師あり学習より少ないデモンストレーションデータで自然な動作を獲得し、俯瞰画像生成器、特徴量抽出ネットワーク、制御モジュールと組み合わせることで実機での協調動作を実現した。この技術は、作業の効率化又は作業員同士の一定の作業空間の確保が求められる場合に有効な技術になり得る。今後は人と機械の混在が想定される生産・物流現場、自動運転などでの実運用に向けて開発に取り組んでいく。

### 参考文献

- (1) Ho, J., et al.: Generative adversarial imitation learning, In Advances in Neural Information Processing Systems, 4565~4573 (2016)
- (2) Goodfellow, I., et al.: Generative adversarial nets, Advances in neural information processing systems 27, 2672~2680 (2014)