

AI 特許紹介(44)
AI 特許を学ぶ！究める！
～言語条件付き模倣学習特許～

2022 年 9 月 9 日

河野特許事務所
所長弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第 4 次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

1.概要

特許出願人 Google

出願日 2021 年 5 月 14 日

公開日 2021 年 11 月 18 日

公開番号 WO2021231895

発明の名称 自然言語ベースのロボット制御で使用するための機械学習モデルのトレーニングおよび/または利用

895 特許は、自由形式の自然言語命令をロボットに与え、目標条件付きポリシー ネットワークにより、視覚データのインスタンス、および、自由形式の自然言語命令の潜在的な目標表現の処理に基づいて、出力を生成する技術に関する。

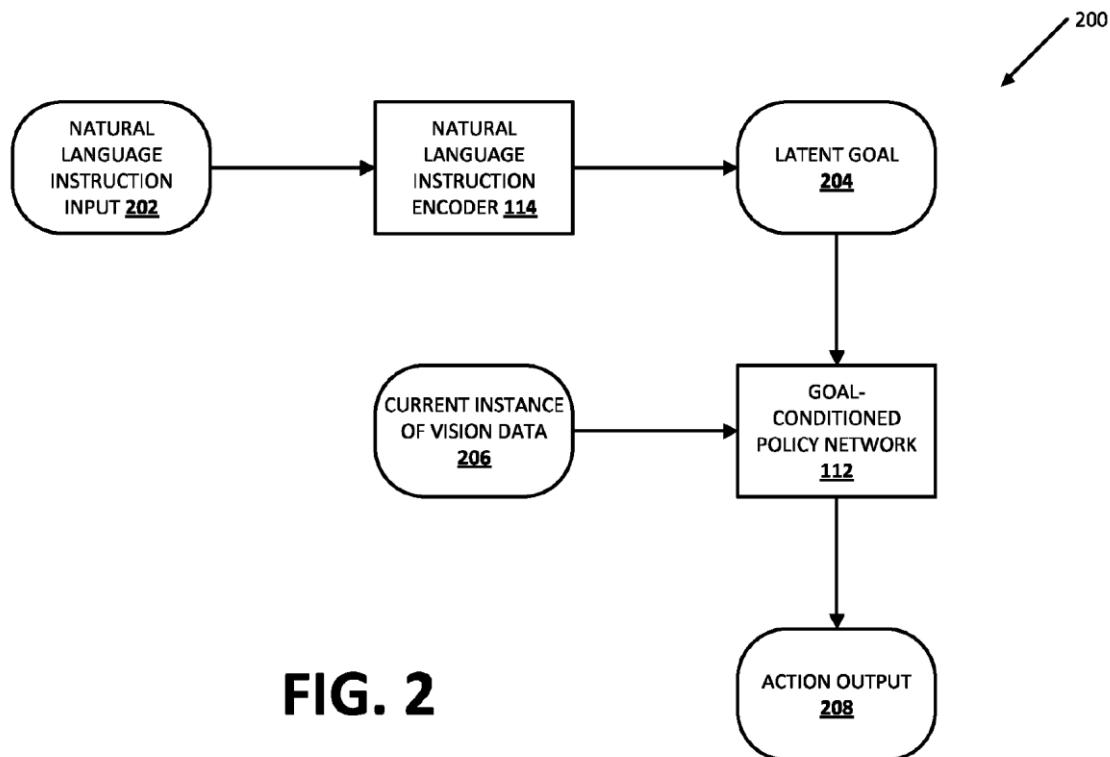
2.特許内容の説明

多くのロボットは、特定のタスクを実行するようにプログラムされている。たとえば、組み立てラインのロボットは、特定のオブジェクトを認識し、それらの特定のオブジェクトに対して特定の操作を実行するようにプログラムできる。

さらに、一部のロボットは、特定のタスクに対応する明示的なユーザインターフェース入力に応答して、特定のタスクを実行することができる。例えば、掃除機ロボットは、「ロボット、きれいに」という発話に応答して、一般的な掃除機掛けタスクを実行する。

しかしながら、通常、ロボットに特定のタスクを実行させるユーザインターフェース入力は、明示的にタスクをマップ化する必要がある。そのため、ロボットは、ロボットを制御しようとするユーザのさまざまな自由形式の自然言語入力に応答して、特定のタスクを実行できないという問題がある。例えば、ロボットは、「ドアを出て、左に曲がり・・」といった自由形式の自然言語入力では、目標の場所に移動できない。

895 特許は言語アノテーションのコストを低減しつつ、自由形式の自然言語条件付けを模倣学習に組み込むものである。下記図は、アクション出力 208 を生成する処理を示すフローチャートである。



システム 200 は、自由形式の自然言語命令入力 202 を受信する。自然言語命令エンコーダ 114 は、自然言語命令入力 202 を処理して、自然言語命令 204 の潜在目標空間表現を生成する。目標条件付きポリシーネットワーク 112 は、アクション出力 208 を生成するために、視覚データ 206 の現在のインスタンス（例えば、視覚コンポーネント 106 を介してキャプチャされた視覚データのインスタンス）と共に潜在的な目標 204 を処理する。

アクション出力 208 は、ロボットが自然言語命令入力 202 によって命令されたタスクを実行するための 1 つまたは複数のアクションを記述する。ロボットのアクチュエータは、ロボットが自然言語命令入力 202 によって示されるタスクを実行するためのアクション出力 208 に基づいて制御される。

下記図は、目標条件付きポリシーネットワーク、自然言語命令エンコーダ、および目標画像エンコーダをトレーニングするプロセス 600 を示すフローチャートである。

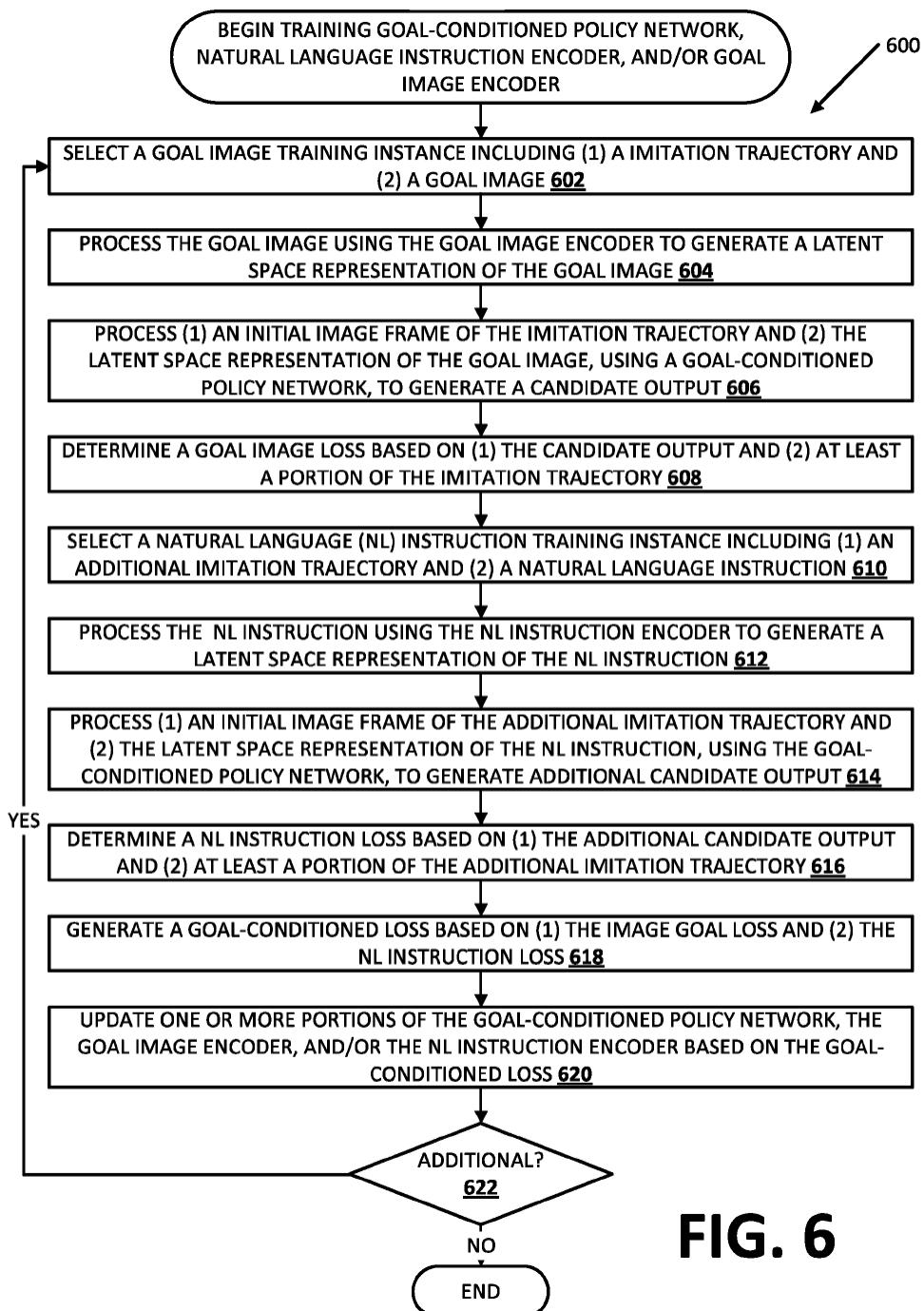


FIG. 6

ブロック 602 で、システムは、(1) 模倣軌道および (2) 目標画像を含む目標画像訓練インスタンスを選択する。

ブロック 604 で、システムは、目標画像エンコーダを使用して目標画像を処理し、目標画像の潜在目標空間表現を生成する。

ブロック 606 で、システムは、目標条件付きポリシーネットワークを使用して、少な

くとも（1）模倣軌道の初期画像フレームおよび（2）目標画像の潜在空間表現を処理して、候補出力を生成する。

ブロック 608 で、システムは、（1）候補出力および（2）模倣軌跡の少なくとも一部に基づいて、目標画像損失を決定する。

ブロック 610 で、システムは、（1）追加の模倣軌跡および（2）自然言語命令を含む自然言語命令トレーニングインスタンスを選択する。

ブロック 612 で、システムは、自然言語エンコーダを使用して自然言語命令訓練インスタンスの自然言語命令部分を処理し、自然言語命令の潜在空間表現を生成する。

ブロック 614 で、システムは、目標条件付きポリシーネットワークを使用して、（1）追加の模倣軌跡の初期画像フレームおよび（2）自然言語命令の潜在空間表現を処理して、追加の候補出力を生成する。

ブロック 616 で、システムは、（1）追加の候補出力および（2）追加の模倣軌跡の少なくとも一部に基づいて、自然言語損失を決定する。

ブロック 618 で、システムは、（1）画像目標損失および（2）自然言語命令損失に基づいて、目標条件付き損失を生成する。

ブロック 620 で、システムは、目標条件付き損失に基づいて、目標条件付きポリシーネットワーク、目標画像エンコーダ、および自然言語命令エンコーダの 1 つまたは複数の部分を更新する。

ブロック 622 で、システムは、目標条件付きポリシーネットワーク、目標画像エンコーダ、および自然言語命令エンコーダに対して追加のトレーニングを実行するかどうかを決定する。

3. クレーム

895 特許のクレーム 1 は以下の通りである。

1. 1 以上のプロセッサにより実装される方法において、

ロボットのタスクを記述する自由形式の自然言語命令を受信し、

自由形式の自然言語命令は、1 つまたは複数のユーザインターフェース入力デバイスを介してユーザにより提供されたユーザインターフェース入力に基づいて生成され、

自由形式の自然言語命令の潜在的な目標表現を生成するために、自然言語命令エンコーダを使用して自由形式の自然言語命令を処理し、

視覚データのインスタンス、ロボットの少なくとも 1 つの視覚コンポーネントによっ

て生成された視覚データのインスタンス、およびロボットの環境の少なくとも一部をキャプチャする視覚データのインスタンスを受信し、

目標条件付きポリシーネットワークを使用して、少なくとも (a) 視覚データのインスタンス、および (b) 自由形式の自然言語命令の潜在的な目標表現の処理に基づいて出力を生成し、

目標条件付きポリシーネットワークは、少なくとも (i) トレーニングタスクが目標画像を使用して記述されるトレーニングインスタンスの目標画像セット、および (ii) トレーニングタスクが自由形式の自然言語命令を使用して記述されるトレーニングインスタンスの自然言語命令セットに基づき、トレーニングされ、

生成された出力に基づいてロボットの 1 つまたは複数のアクチュエータを制御し、
ロボットの 1 つまたは複数のアクチュエータを制御すると、ロボットは、生成された出力によって示される少なくとも 1 つのアクションを実行する。

4. 本特許に関連する論文

本特許に関する論文“Language Conditioned Imitation Learning over Unstructured Data”¹が、Google の Corey Lynch 氏らにより公表されている。

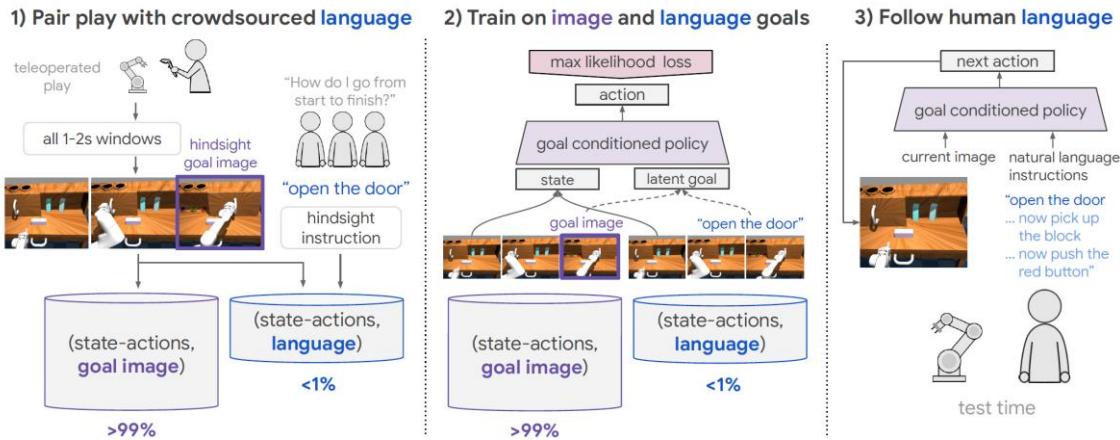
自然言語は、人間がタスクをロボットに伝えるための最も柔軟で直感的な方法である。模倣学習における従来の方法は、各タスクをタスク ID または目標画像で指定する必要があった。これは、オープンワールド環境では現実的ではない。

本論文は、自由形式の自然言語条件付けを模倣学習に組み込む方法を紹介しており、ピクセルからの知覚、自然言語の理解、マルチタスクの連続制御を単一のニューラルネットワークとしてエンドツーエンドで学習する。模倣学習における以前の研究とは異なり、ラベル付けされていない構造化されていないデモンストレーションデータを組み込むことができる（つまり、タスクや言語のラベルはない）。

これにより、言語アノテーションのコストを全データの 1%未満に削減しながら、言語条件付きのパフォーマンスを劇的に向上させる。テスト時には、トレーニングされた単一の言語条件付き視覚運動ポリシーは、各タスクの自然言語の説明のみで指定された 3D 環境で、さまざまなロボット操作スキルを実行することができる（例：「引き出しを開けて…今すぐブロックを手に取って…緑のボタンを押して…」）

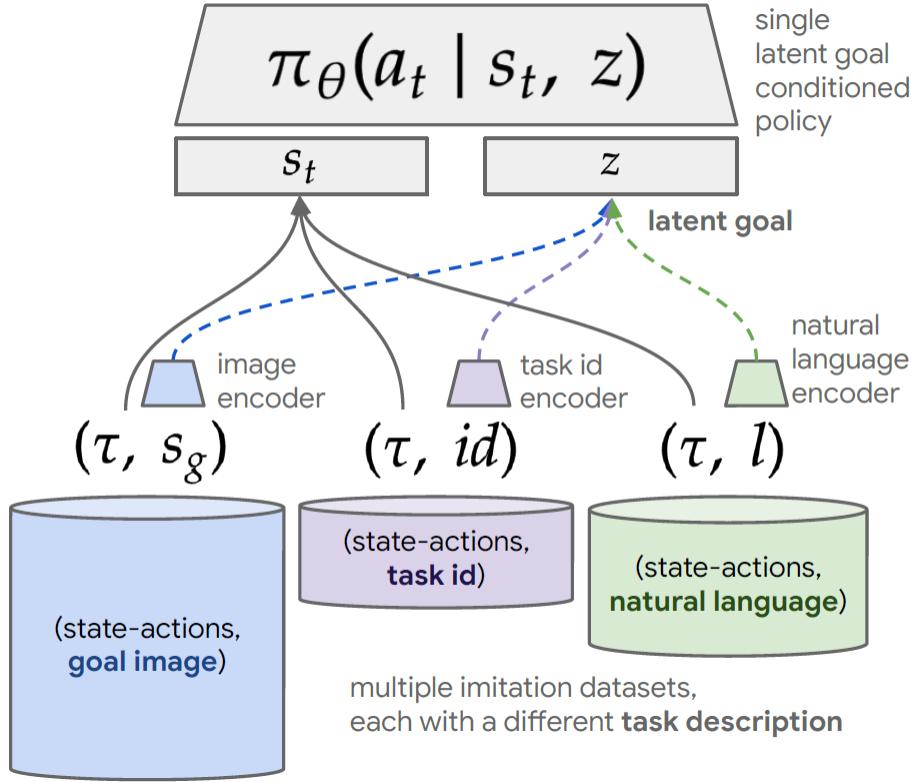
¹ Corey Lynch, Pierre Sermanet “Language Conditioned Imitation Learning over Unstructured Data” arXiv:2005.07648v2 [cs.RO] 7 Jul 2021

下記図は非構造化データに続く自由形式の自然言語命令のスケールアップを示す説明図である。



- 1) まず、構造化されていない遠隔操作の「プレイ」データ（タスクラベルなし）を後知恵の目標イメージの例に再ラベル付けする。次に、ちょっととした遊びと後知恵の指示を組み合わせる。
- 2) イメージまたは言語のいずれかの目標に従うように単一のポリシーをトレーニングする。
- 3) テスト時に言語条件付けのみを使用する。

下記図はマルチコンテキスト模倣学習 (MCIL : Multicontext Imitation Learning) を示す説明図である。



MCIL は、すべてのデータセットに対して単一の潜在目標条件付きポリシーを同時にトレーニングし、データセットごとに 1 つのエンコーダをトレーニングし、それが共有潜在目標空間にマッピングする。

これにより、ラベル付けされたデモンストレーションデータセットとラベル付けされていないデモンストレーションデータセットの両方で言語条件付きポリシーをトレーニングできる。

Method	Input	Training source	Task conditioning	Multi-18 Success (18 tasks)	Chain-4 Success (925 long-horizon tasks)
LangBC	pixels	predefined demos	text	20.0% ±3.0	7.1% ±1.5
Restricted LangLfP	pixels	unstructured demos	text	47.1% ±2.0	25.0% ±2.0
LfP	pixels	unstructured demos	image	66.4% ±2.2	53.0% ±5.0
LangLfP (ours)	pixels	unstructured demos	text	68.6% ±1.7	52.1% ±2.0
TransferLangLfP (ours)	pixels	unstructured demos	text	74.1% ±1.5	61.8% ±1.1
LangBC	states	predefined demos	text	38.5% ±6.3	13.9% ±1.4
Restricted LangLfP	states	unstructured demos	text	88.0% ±1.4	64.2% ±1.5
LangLfP (ours)	states	unstructured demos	text	88.5% ±2.9	63.2% ±0.9
TransferLangLfP (ours)	states	unstructured demos	text	90.5% ±0.8	71.8% ±1.6

TABLE I: Human language conditioned visual manipulation experiments

上記表は各モデルのスコアを示す説明図である。本手法(TransferLangLfP)の成功率

が高いことが理解できる。

自然言語によりロボットを動作させる動画は下記サイトで閲覧することができる。

<https://language-play.github.io>

以上

著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI研究所 AI コース修了。

[AI 特許コンサルティング](#)、[医療 AI 特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech 特許入門」、「[AI/IoT 特許入門 2.0](#)」、「[ブロックチェーン 3.0](#)(共著)」がある。