

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7367226号
(P7367226)

(45)発行日 令和5年10月23日(2023.10.23)

(24)登録日 令和5年10月13日(2023.10.13)

(51)国際特許分類		F I	
G 0 5 B	19/18 (2006.01)	G 0 5 B	19/18 W
G 0 5 B	23/02 (2006.01)	G 0 5 B	23/02 T
G 0 5 B	19/4155(2006.01)	G 0 5 B	23/02 3 0 1 N
B 2 3 Q	17/09 (2006.01)	G 0 5 B	19/4155 U
		B 2 3 Q	17/09 Z
請求項の数 19 (全23頁)			
(21)出願番号	特願2022-543267(P2022-543267)	(73)特許権者	000006013
(86)(22)出願日	令和2年10月14日(2020.10.14)		三菱電機株式会社
(65)公表番号	特表2022-549531(P2022-549531 A)	(74)代理人	110001195
(43)公表日	令和4年11月25日(2022.11.25)		弁理士法人深見特許事務所
(86)国際出願番号	PCT/JP2020/039473	(72)発明者	ウィヘルン, ゴードン
(87)国際公開番号	WO2021/075583		アメリカ合衆国、0 2 1 3 9 - 1 9 5 5
(87)国際公開日	令和3年4月22日(2021.4.22)		マサチューセッツ州、ケンブリッジ、ブ
審査請求日	令和4年3月18日(2022.3.18)		ロードウェイ 2 0 1、エイス・フロア
(31)優先権主張番号	16/655,272		、ケアオブ・ミツビシ・エレクトリック
(32)優先日	令和1年10月17日(2019.10.17)		・リサーチ・ラボラトリーズ・インコー
(33)優先権主張国・地域又は機関	米国(US)	(72)発明者	ポレイテッド
			ル・ルー, ジョナタン
			アメリカ合衆国、0 2 1 3 9 - 1 9 5 5
			マサチューセッツ州、ケンブリッジ、ブ
			最終頁に続く

(54)【発明の名称】 音波分離ニューラルネットワークを用いた製造自動化

(57)【特許請求の範囲】

【請求項1】

1つ以上のツールを補助して1つ以上のタスクを実行する複数のアクチュエータを含む機械の動作を制御するためのシステムであって、

前記システムの動作中に、タスクを実行するツールによって生成された信号と前記ツールを作動させる前記複数のアクチュエータによって生成された信号との音波混合物を受信するように構成された入力インターフェイスと、

前記ツールを作動させる前記アクチュエータによって生成された信号から前記タスクを実行する前記ツールによって生成された信号を、音波混合物信号から分離するように訓練されたニューラルネットワークを記憶するように構成されたメモリと、

前記音波混合物信号を前記ニューラルネットワークに投入することによって、前記音波混合物信号から前記タスクを実行する前記ツールによって生成された前記信号を抽出し、前記抽出された信号を分析することによって、前記タスクの実行状態を生成し、前記タスクの前記実行状態に応じて選択された制御動作を実行するように構成されたプロセッサとを含み、

前記機械の動作中に、複数のツールは、第1のタスクを実行する第1のツールと第2のタスクを実行する第2のツールとを含み、複数のタスクを実行し、

前記ニューラルネットワークは、前記音波混合物から、前記第1のタスクを実行する前記第1のツールによって生成された第1の信号と前記第2のタスクを実行する前記第2のツールによって生成された第2の信号とを分離するように訓練され、

前記システムの動作中に、前記プロセッサは、前記ニューラルネットワークを用いて前記音波混合物から前記第1の信号および前記第2の信号を抽出し、前記第1の信号と前記第2の信号とを独立して分析することによって前記第1のタスクの第1の実行状態と前記第2のタスクの第2の実行状態とを推定し、前記第1の実行状態に応じて選択された第1の制御動作を実行し、前記第2の実行状態に応じて選択された第2の制御動作を実行するように構成される、システム。

【請求項2】

前記機械は、ワークピースを機械加工するための一連の操作を実行するように構成され、
前記ツールは、機械加工ツールであり、

前記プロセッサは、ツール経路に沿って前記機械加工ツールを作動させるコンピュータ数値制御（CNC）を実行し、

前記機械加工ツールによって生成された前記信号は、前記機械加工ツールによる機械加工中に前記ワークピースの変形によって生成された振動信号である、請求項1に記載のシステム。

【請求項3】

前記機械は、前記ワークピースの機械加工、はんだ付け、および組み立てのうちの1つまたは組み合わせを含む前記ワークピースを製造するための一連の操作を実行するように構成され、

前記ツールによって生成された前記信号は、前記ワークピースの製造中に前記ワークピースの変更によって生成された振動信号である、請求項2に記載のシステム。

【請求項4】

前記機械は、潜在的な異常を検出するために監視されている変速装置であり、

前記変速装置は、モータからの振動、カップリング、または可動部からの他の振動が存在する場合のみ記録される、請求項1に記載のシステム。

【請求項5】

前記ニューラルネットワークは、各分離済み信号が前記システムの動作中に存在する1種類の信号であり、前記分離済み信号が前記音波混合物の合計を構成するように、前記音波混合物から異なる信号を分離するように訓練される、請求項1に記載のシステム。

【請求項6】

前記システムの動作中に存在する信号の種類は、弱ラベルによって識別され、

各弱ラベルは、動作中のある時点に存在する信号の種類を特定する、請求項5に記載のシステム。

【請求項7】

前記ニューラルネットワークは、前記弱ラベルによって識別された前記信号の種類を分類するように構成された分類器と共同で訓練される、請求項6に記載のシステム。

【請求項8】

前記ニューラルネットワークは、前記分類器と共同で訓練されて、前記弱ラベルと前記分類器を通る分離ニューラルネットワークとの間のクロスエントロピー項を含む損失関数を最小化する、請求項7に記載のシステム。

【請求項9】

前記ニューラルネットワークは、前記弱ラベルによって識別された1種類の分離済み信号が入力として前記ニューラルネットワークに投入されると、前記ニューラルネットワークが前記分離済み信号を出力として生成するように訓練される、請求項6に記載のシステム。

【請求項10】

前記プロセッサは、状態推定器を実行することによって、前記タスクの前記実行状態を推定し、

前記状態推定器は、前記タスクを実行する前記ツールによって生成された信号、および、前記ニューラルネットワークによって前記システムの異なる反復動作の異なる音波混合物から抽出された信号に基づいて訓練される、請求項1に記載のシステム。

10

20

30

40

50

【請求項 1 1】

前記タスクを実行する前記ツールによって生成された抽出信号の各サンプルは、前記タスクの実行状態を定義し、前記機械の動作状態を定義するには不十分である、請求項 1 に記載のシステム。

【請求項 1 2】

前記タスクを実行する前記ツールによって生成された前記信号は、前記ツールを作動させる前記アクチュエータによって生成された信号と混合され、前記音波混合物において同じ時間および周波数スペクトルを占有する、請求項 1 に記載のシステム。

【請求項 1 3】

前記音波混合物は、単一のマイクロフォンの単一の出力チャンネルから得られる、請求項 1 に記載のシステム。

【請求項 1 4】

少なくとも一部の前記アクチュエータは、前記機械内の物理位置に基づいて空間的に重なる、請求項 1 に記載のシステム。

【請求項 1 5】

1 つ以上のツールを補助して 1 つ以上のタスクを実行する複数のアクチュエータを含む機械の動作を制御するための方法であって、前記方法は、前記方法を実行するために格納された命令と結合されたプロセッサを使用し、前記命令は、前記プロセッサによって実行されると、前記方法の以下のステップを実行し、前記ステップは、

タスクを実行するツールによって生成された信号とツールを作動させる複数のアクチュエータによって生成された信号との音波混合物を受信するステップと、

前記ツールを作動させる前記アクチュエータによって生成された信号から前記タスクを実行する前記ツールによって生成された信号を、音波混合物信号から分離するように訓練されたニューラルネットワークに前記音波混合物信号を投入することによって、前記音波混合物信号から前記タスクを実行する前記ツールによって生成された前記信号を抽出するステップと、

前記抽出された信号を分析することによって、前記タスクの実行状態を生成するステップと、

前記タスクの前記実行状態に応じて選択された制御動作を実行するステップとを含み、前記機械の動作中に、前記 1 つ以上のツールは、第 1 のタスクを実行する第 1 のツールと第 2 のタスクを実行する第 2 のツールとを含み、複数のタスクを実行し、

前記ニューラルネットワークは、前記音波混合物から、前記第 1 のタスクを実行する前記第 1 のツールによって生成された第 1 の信号と前記第 2 のタスクを実行する前記第 2 のツールによって生成された第 2 の信号とを分離するように訓練されており、

前記ステップは、

前記機械の動作中に、前記ニューラルネットワークを用いて前記音波混合物から前記第 1 の信号および前記第 2 の信号を抽出するステップと、

前記第 1 の信号と前記第 2 の信号とを独立して分析することによって前記第 1 のタスクの第 1 の実行状態と前記第 2 のタスクの第 2 の実行状態とを推定するステップと、

前記第 1 の実行状態に応じて選択された第 1 の制御動作を実行するステップと、

前記第 2 の実行状態に応じて選択された第 2 の制御動作を実行するステップとをさらに含む、方法。

【請求項 1 6】

前記タスクを実行する前記ツールによって生成された前記信号は、前記ツールを作動させる前記アクチュエータによって生成された信号と混合され、前記音波混合物において同じ時間および周波数スペクトルを占有する、請求項 1 5 に記載の方法。

【請求項 1 7】

前記音波混合物は、単一のマイクロフォンの単一の出力チャンネルから得られる、請求項 1 5 に記載の方法。

【請求項 1 8】

少なくとも一部の前記アクチュエータは、前記機械内の物理位置に基づいて空間的に重なる、請求項 15 に記載の方法。

【請求項 19】

機械の動作を制御するための方法を実行するために、プロセッサによって実行可能なプログラムを記憶する非一時的なコンピュータ可読記憶媒体であって、前記方法は、
1 つ以上のタスクを実行する 1 つ以上のツールによって生成された信号と前記 1 つ以上のツールを作動させる複数のアクチュエータによって生成された信号との音波混合物を受信するステップと、

前記ツールを作動させる前記アクチュエータによって生成された信号から前記タスクを実行する前記ツールによって生成された信号を、音波混合物信号から分離するように訓練されたニューラルネットワークに前記音波混合物信号を投入することによって、前記音波混合物信号から前記タスクを実行する前記ツールによって生成された前記信号を抽出するステップと、

前記抽出された信号を分析することによって、前記タスクの実行状態を生成するステップと、

前記タスクの前記実行状態に応じて選択された制御動作を実行するステップとを含み、
前記機械の動作中に、前記 1 つ以上のツールは、第 1 のタスクを実行する第 1 のツールと第 2 のタスクを実行する第 2 のツールとを含み、前記複数のタスクを実行し、
前記ニューラルネットワークは、前記音波混合物から、前記第 1 のタスクを実行する前記第 1 のツールによって生成された第 1 の信号と前記第 2 のタスクを実行する前記第 2 のツールによって生成された第 2 の信号とを分離するように訓練されており、
前記方法は、

前記機械の動作中に、前記ニューラルネットワークを用いて前記音波混合物から前記第 1 の信号および前記第 2 の信号を抽出するステップと、

前記第 1 の信号と前記第 2 の信号とを独立して分析することによって前記第 1 のタスクの第 1 の実行状態と前記第 2 のタスクの第 2 の実行状態とを推定するステップと、

前記第 1 の実行状態に応じて選択された第 1 の制御動作を実行するステップと、

前記第 2 の実行状態に応じて選択された第 2 の制御動作を実行するステップとをさらに含む、非一時的なコンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般に、機械学習技術を用いた製造自動化に関し、特に、音波混合物から信号を分離するように訓練されたニューラルネットワークを用いた製造自動化に関する。

【背景技術】

【0002】

安全性および品質の監視および制御は、高速および強力な機械が高速で一連の複雑な操作を実行することができる製造において非常に重要である。意図した一連の動作またはタイミングからの逸脱は、品質の低下、原材料の浪費、故障時間および設備の破壊を引き起こし、産出を低下する可能性がある。作業への危険性は、主な懸念事項である。このため、予期しない事象を最小限に抑えるために製造プロセスを細心に設計するように厳重な注意を払う必要があり、様々なセンサおよび緊急スイッチを用いて、生産ラインに安全保護装置を設計する必要もある。

【0003】

製造の種類は、プロセス製造および個別製造を含む。プロセス製造において、製品、例えば油、天然ガスおよび塩などの製品は、一般的に分離されない。個別製造は、例えば自動車、家具、玩具、および飛行機のような個別の製品を生産する。

【0004】

安全性を増大させ、材料および産出の損失を最小限に抑えるための 1 つの実用的なアプローチは、いつ生産ラインが異常状態で運転しているかを検出し、生産ラインの異常を検

10

20

30

40

50

出した場合に、必要に応じて生産ラインを停止することである。このアプローチを実現する1つの方法は、許容可能な動作領域を規定し、許容可能な動作領域から逸脱した動作点を検出するための、例えば、温度および圧力などの測定可能な変数範囲についての、で、生産ラインの正常動作の記述を使用することである。この方法は、通常、物理変数の許容範囲がよく理解されており、製品の品質メトリックがしばしばこれらの変数で直接に定義されているプロセス製造業、例えば油の精製においてよく使用される。

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、個別製造の加工プロセスの本質は、プロセス製造の加工プロセスの本質とは異なり、通常の加工プロセスからの逸脱は、非常に異なる特性を有する。個別製造は、作業ユニットに対して行われた一連の作業、例えば、機械加工、はんだ付けおよび組み立てを含む。異常は、1つ以上のタスクの不正確な実行、またはタスクの不正確な順序を含むことができる。異常があっても、温度または圧力などの物理変数がしばしば範囲から外れていないため、これらの変数を直接に監視しても、異常を確実に検出することができない。

【0006】

さらに、複雑な製造システムは、プロセス製造および個別製造の組み合わせを含むことができる。プロセス製造と個別製造とを1つの生産ラインに混在して設けた場合、異なる種類の製造のために設計された異常検出方法は、不正確になる可能性がある。例えば、プロセス製造のための異常検出方法は、データの外れ値を検出することを目的とする一方、個別製造のための異常検出方法は、動作の実行順序の異常を検出することを目的とする。したがって、異なる種類の製造作業に対して異なる異常検出方法を設計することが必要である。しかしながら、複雑な製造システムにおいて異なる検出技術を使用することは、複雑になり過ぎる可能性がある。

【0007】

したがって、異なる種類の製造システムの異常検出に適したシステムおよび方法を開発する必要がある。例えば、US 2019/0302707 A1に記載された方法は、機械学習技術を適用することによって、プロセス製造および個別製造のうちの1つまたはその組み合わせの異常を検出する。機械学習を用いることによって収集されたデータは自動学習システムに利用され、訓練を通してデータの特徴が学習され得る。訓練済みモデルは、リアルタイムデータの異常を検出することができ、予防保全および故障時間の削減を実現することができる。しかしながら、機械学習による支援を用いても、いくつかの製造作業を表すために収集される必要があるデータは、正確な異常検出を不可能にする。

【課題を解決するための手段】

【0008】

いくつかの実施形態の目的は、1つ以上のタスクを実行する1つ以上のツールを動作させる複数のアクチュエータを含む複雑な産業システムの製造自動化に適したシステムおよび方法を提供することである。追加的にまたは代替的に、いくつかの実施形態の目的は、機械学習を用いて、1つ以上のタスクの実行状態を推定し、それに応じてシステムを制御することである。

【0009】

いくつかの実施形態は、機械学習を用いて、未知の基本的な動的システムに従って状態が経時的に変化する物理システムについてのデータ駆動型の時系列予測推論のために用いることができるという認識に基づいている。これらのシステムについて、システムの状態に関連する観測結果のみが測定される。これは、モデル化することが難しい複雑な産業システムの制御に有利であり得る。

【0010】

しかしながら、いくつかの実施形態は、観測結果がシステムの状態を明瞭に表す場合に機械学習技術を使用することができるという別の認識に基づいており、これは、いくつか

10

20

30

40

50

の状況では問題となり得る。実際に、一組の観測結果がシステムの状態に一意に対応する場合、機械学習法を用いて様々な予測変数を設計することができる。しかしながら、各時点の検出データは、システムの実際の状態に関する情報を十分に提供できない場合がある。必要な観測結果の数は、システムの次元 d に依存し、線形システムでは d に等しく、非線形システムでは $2d + 1$ に等しい。収集された測定値が十分な観測結果を含まない場合、機械学習方法は、失敗する可能性がある。

【0011】

いくつかの実施形態は、タスクを実行するシステムの状態を検討する代わりに、タスク自体の実行状態を検討することができるという認識に基づいている。例えば、システムが1つ以上のタスクを実行するための複数のアクチュエータを含む場合、システムの状態は、これらのアクチュエータの全ての状態を含む。しかしながら、場合によって、アクチュエータの状態は、主な制御対象ではない。実際に、性能、すなわち、タスクの実行を導くために、アクチュエータの状態が必要である。したがって、実行の状態は、主な目的であり、タスクを実行するアクチュエータの状態は、二次的な目的にすぎない。

【0012】

いくつかの実施形態は、システムの状態のみを非常に頻繁に測定または観測することができ、十分な観測結果を収集した場合に、システムの状態が事実上実行状態を表すことができるため、システムの状態をタスクの実行状態と等しくすることができるという理解に基づいている。しかしながら、場合によって、システムの状態を測定することが困難であるため、タスクの実行状態を定義することが困難である。

【0013】

例えば、切削ツールでワークピースを機械加工するコンピュータ数値制御（CNC）を検討する。システムの状態は、ツール経路に沿って切削ツールを移動させるアクチュエータの状態を含む。機械加工の実行状態は、実際の切削状態である。産業用CNCシステムは、多くの異なるアクチュエータ、場合によっては冗長なアクチュエータを含み、これらのアクチュエータは、互いに複雑な非線形関係にある多数の状態変数を有するため、CNCシステムの状態を観測することが困難となる。したがって、ワークピースの機械加工状態を測定することも困難であり得る。

【0014】

いくつかの実施形態は、タスクの実行状態が、タスクの実行によって生成された音波信号で表すことができるという認識に基づいている。例えば、ワークピースのCNC機械加工の実行状態は、機械加工中のワークピースの変形によって生成された振動信号で表すことができる。したがって、このような振動信号を測定できれば、機械学習法を含む様々な分類技術を用いてこのような振動信号を分析することによって、タスクの実行状態を推定することができ、実行を制御するための適切な制御動作を選択することができる。

【0015】

しかしながら、このアプローチの問題は、このような振動信号が単独で存在していないことである。例えば、1つ以上のタスクを実行する1つ以上のツールを作動させる複数のアクチュエータを含むシステムにおいて、タスクを実行するツールによって生成された信号には、常に、ツールを作動させる複数のアクチュエータによって生成された信号が混合されている。例えば、ワークピースの変形によって生成された振動信号には、常に、切削ツールを移動させるモータからの信号が混合されている。何らかの方法でこのような振動信号を単独で生成することまたは単独で捕捉することができれば、このような信号を抽出するように、ニューラルネットワークなどの機械学習システムを訓練することができる。しかしながら、CNC機械加工を含む多くの場合に、タスクの実行を表す孤立信号の生成は、非現実である。同様に、複数のマイクロフォンで異なる信号を個別に記録することも、非現実である。

【0016】

したがって、1つ以上のツールを作動させて1つ以上の作業を実行する複数のアクチュエータを含むシステムの製造自動化を効率化するために、作業を実行するツールによって

10

20

30

40

50

生成された信号と、ツールを作動させる複数のアクチュエータによって生成された信号とを音波混合物から音源信号を分離する必要がある。したがって、いくつかの実施形態の目的は、単独の音源が存在していない場合に、混合信号の音源を分離するためのニューラルネットワークを訓練することである。本明細書に使用された場合、混合信号中の少なくとも一部の音源は、音波混合物において同じ時間、空間および周波数スペクトルを占有する。

【0017】

いくつかの実施形態は、例えば、音波混合物中の音源を単独で容易に記録できる場合または人間が音波混合物中の音源を容易に識別できる場合に、教師付き学習で訓練を行うことができるという認識に基づいている。このような分類は、本明細書において強ラベルと呼ばれる。理論的には、人間は、このようなラベルの近似値を提供することができるが、時間および周波数の両方でアクティブ音を正確に分類することによって、強ラベルを付けることは、殆ど非現実である。しかしながら、一部の時間範囲内でアクティブ音を限定的に分類すること（弱ラベル）は、現実的である。通常の場合、このような弱ラベルは、指定範囲の全体にわたって音がアクティブである必要がなく、指定範囲内の短時間のみにアクティブであればよい。

【0018】

いくつかの実施形態は、通常、製造プロセスの仕組みが既知であれば、音波信号を生成する音源も既知であるという認識に基づいている。したがって、製造システムの動作を表す音波混合物の音源に弱ラベルを付けることができる。よって、いくつかの実施形態は、弱ラベルのみを含む音波混合物の訓練データが利用可能である場合に、この音波混合物中の音波を分離するように学習することができる方法を開発する。

【0019】

したがって、いくつかの実施形態は、音波混合物から、タスクを実行するツールによって生成された信号とツールを作動させるアクチュエータによって生成された信号とを分離するようにニューラルネットワークを訓練する。例えば、ニューラルネットワークは、各分離済み信号がシステムの動作中に存在する信号の1つのクラスに属しており、分離済み信号が足し合わされることによって音波混合物が構成されるように、音波混合物から異なる信号を分離するように訓練される。弱ラベルは、システムの動作中に存在する信号のクラスを識別する。各弱ラベルは、動作中のある時点に存在する信号のクラスを特定する。

【0020】

一実施形態において、ニューラルネットワークは、弱ラベルによって識別された信号のクラスを分類するように構成された分類器と共同で訓練される。例えば、ニューラルネットワークは、分類器の出力とニューラルネットワークの出力との間のクロスエントロピー項を含む損失関数を最小化するように、分類器と共同で訓練されてもよい。このような共同訓練は、単独で存在しない信号を分類することが困難であるという事実を考慮し、分離ニューラルネットワークおよび分類ニューラルネットワークのエンドツーエンド訓練を可能にする。

【0021】

いくつかの実装形態において、品質を保証するために、ニューラルネットワークの訓練に付加制約が追加される。例えば、一実施形態において、ニューラルネットワークは、弱ラベルによって識別されたあるクラスの分離済み信号が入力としてニューラルネットワークに投入されると、ニューラルネットワークはこの分離済み信号を出力として生成するように訓練される。追加的にまたは代替的に、一実施形態において、ニューラルネットワークは、弱ラベルによって識別された2つ以上のクラスからの分離済み信号を再結合し、ネットワークにフィードバックして再分離するように訓練される一方で、敵対的損失（Adversarial Loss）が実際の音波混合物と再結合された合成混合物とを区別するように訓練される。

【0022】

信号分離を行うようにニューラルネットワークを訓練する場合、いくつかの実施形態は、このネットワークの出力を使用して、タスクの実行状態を推定するための状態推定器を

10

20

30

40

50

訓練する。例えば、一実施形態において、状態推定器は、タスクを実行するツールによって生成され、ニューラルネットワークによってシステムの異なる反復動作の異なる音波混合物から抽出された信号に基づいて訓練される。特に、いくつかの実施形態において、タスクを実行するツールによって生成された抽出信号の各サンプルは、タスクの実行状態を定義し、システムの動作状態を定義するには不十分である。しかしながら、タスクの実行状態は、適切な制御動作を選択するのに十分である。このように、実施形態は、製造自動化用途の次元を削減することができる。

【0023】

次元削減に加えてまたはその代わりに、いくつかの実施形態は、異なる作業タスクを実行する異なるツールを個別に制御することができる。例えば、システム全体の状態を示す信号を分析すれば、システム全体の制御を行うことができる。したがって、異なるタスクを実行する異なるツールを個別に分析する場合、異なるツールを個別に制御することができる。

10

【0024】

したがって、システムの動作中に複数のツールが複数のタスクを実行する場合、いくつかの実施形態は、複数のタスクを個別に制御する。例えば、一実施形態は、第1のタスクを実行する第1のツールと第2のタスクを実行する第2のツールとを含むシステムを制御する。ニューラルネットワークは、音波混合物から、第1のタスクを実行する第1のツールによって生成された第1の信号と第2のタスクを実行する第2のツールによって生成された第2の信号とを分離するように訓練される。システムの動作中に、ニューラルネットワークを用いて音波混合物から第1の信号および第2の信号を抽出し、互いに独立して分析することによって、第1のタスクの第1の実行状態と第2のタスクの第2の実行状態とを推定する。この実施形態は、第1の状態に応じて選択された第1の制御動作を実行し、第2の状態に応じて選択された第2の制御動作を実行するように構成される。

20

【0025】

したがって、一実施形態は、1つ以上のツールを補助して1つ以上のタスクを実行する複数のアクチュエータを含む機械の動作を制御するためのシステムを開示する。システムは、システムの動作中に、タスクを実行するツールによって生成された信号とツールを作動させる複数のアクチュエータによって生成された信号との音波混合物を受信するように構成された入力インターフェイスと、ツールを作動させるアクチュエータによって生成された信号からタスクを実行するツールによって生成された信号を、音波混合物信号から分離するように訓練されたニューラルネットワークを記憶するように構成されたメモリと、音波混合物信号をニューラルネットワークに投入することによって、音波混合物信号からタスクを実行するツールによって生成された信号を抽出し、抽出された信号を分析することによってタスクの実行状態を生成し、タスクの実行状態に応じて選択された制御動作を実行するように構成されたプロセッサとを含む。

30

【0026】

別の実施形態は、1つ以上のツールを補助して1つ以上のタスクを実行する複数のアクチュエータを含む機械の動作を制御するための方法を開示する。この方法は、方法を実装する格納された命令と結合されたプロセッサを使用し、命令は、プロセッサによって実行されると、方法のステップを実行する。この方法は、タスクを実行するツールによって生成された信号とツールを作動させる複数のアクチュエータによって生成された信号との音波混合物を受信するステップと、ツールを作動させるアクチュエータによって生成された信号からタスクを実行するツールによって生成された信号を、音波混合物信号から分離するように訓練されたニューラルネットワークに音波混合物信号を投入することによって、音波混合物信号からタスクを実行するツールによって生成された信号を抽出するステップと、抽出された信号を分析することによって、タスクの実行状態を生成するステップと、タスクの実行状態に応じて選択された制御動作を実行するステップとを含む。

40

【0027】

さらに別の実施形態は、方法を実行するために、プロセッサによって実行可能なプログ

50

ラムを記憶する非一時的なコンピュータ可読記憶媒体を開示する。この方法は、タスクを実行するツールによって生成された信号とツールを作動させる複数のアクチュエータによって生成された信号との音波混合物を受信するステップと、ツールを作動させるアクチュエータによって生成された信号からタスクを実行するツールによって生成された信号を、音波混合物信号から分離するように訓練されたニューラルネットワークに音波混合物信号を投入することによって、音波混合物信号からタスクを実行するツールによって生成された信号を抽出するステップと、抽出された信号を分析することによって、タスクの実行状態を生成するステップと、タスクの実行状態に応じて選択された制御動作を実行するステップを含む。

【図面の簡単な説明】

【0028】

【図1A】いくつかの実施形態に従う、機械の動作中に振動を生成する異なるアクチュエータに対応する信号を分離し、これらの分離済み信号を用いて後続のタスクを監視するためのシステムを示すブロック図である。

【図1B】いくつかの実施形態に従う、機械の動作中に振動を生成する異なるアクチュエータに対応する信号を分離し、これらの分離済み信号を用いて後続のタスクを監視するためのシステムを示すブロック図である。

【図1C】いくつかの実施形態に従う、タスクを実行するツールによって生成された信号とツールを作動させる複数のアクチュエータによって生成された信号との音波混合物のスペクトログラムを示す図である。

【図2】いくつかの実施形態に従う、音源を分離するための音波信号処理システムを訓練することを示す流れ図である。

【図3A】いくつかの実施形態に従う、単一チャネルマスク推論音源分離ネットワークのアーキテクチャを示すブロック図である。

【図3B】いくつかの実施形態に従う、音事象の分類を行うための畳み込みリカレント型ネットワークのアーキテクチャを示すブロック図である。

【図4】いくつかの実施形態に従う、弱ラベルを用いて音源分離ネットワークを訓練するためのいくつかの方法ステップを示す流れ図である。

【図5】いくつかの実施形態に従う、分離済み信号を強調するために必要な分類損失関数およびターゲットを示す概略図である。

【図6】いくつかの実施形態に従う、分離された信号に基づいて実行状態を分析することを示すグラフである。

【図7】いくつかの実施形態に従う、機械加工動作の制御を示す概略図である。

【図8】いくつかの実施形態に従う、製造異常検知システムのアクチュエータを示す概略図である。

【発明を実施するための形態】

【0029】

図1Aおよび図1Bは、いくつかの実施形態に従う、実行を分析して機械102の動作を制御するためのシステム100を示すブロック図である。機械102は、1つ以上のアクチュエータ（コンポーネント）103を含むことができる。各アクチュエータ103は、特定のタスクを実行し、調節装置104に接続される。アクチュエータ103によって実行されるタスクの例は、機械加工、はんだ付け、または組み立てを含むことができる。いくつかの実施形態において、機械アクチュエータ103は、同時に動作することができるが、調節装置104は、各アクチュエータを個別に制御する必要がある。被調節装置104の例は、タスクを実行するためのツールを含む。例えば、いくつかの実施形態において、システム100は、1つ以上のツールを補助して1つ以上のタスクを実行する複数のアクチュエータを含む機械の動作を制御する。

【0030】

いくつかの実施形態において、マイクロフォンまたは複数のマイクロフォンからなるアレイであり得るセンサ101は、機械102の動作中に各アクチュエータ103によって

10

20

30

40

50

生成された振動を捕捉する。また、一部の機械アクチュエータ１０３は、同じ空間領域の同じ場所に配置されることがある。これによって、センサ１０１は、マルチマイクロフォンアレイであっても、アクチュエータ１０３によって生成された振動を個別に捕捉することができない。したがって、センサ１０１によって捕捉された振動信号は、各機械アクチュエータ１０３によって生成された振動信号の和からなる音波混合物信号１９５である。いくつかの実施形態において、音波混合物信号のスペクトログラム内の少なくとも一部の音源は、音波混合物において同じ時間、空間および周波数スペクトルを占有する。

【００３１】

いくつかの実施形態において、音が大きいアクチュエータが静かなアクチュエータを覆い隠す可能性があるため、音波混合物信号１９５から各機械アクチュエータ１０３の実行状態を推定できない可能性がある。したがって、システム１００は、音波混合物信号１９５から各機械アクチュエータ１０３によって生成された振動を分離できる分離ニューラルネットワーク１３１を含む。各機械アクチュエータ１０３の信号が音波混合物信号１９５から分離される。それらは、タスクの実行状態の推定１３５およびタスクの実行１３７のために用いられ得る。

【００３２】

したがって、システム１００は、機械の動作を制御するように、プロセッサ１２０によって実行される様々なモジュールを含む。プロセッサは、音波混合物信号をニューラルネットワーク１３１に投入することによって、タスクを実行するツールによって生成された信号を音波混合物信号から抽出し、推定器１３５を用いて抽出した信号を分析することによってタスクの実行状態を生成し、機械１０２の障害回避および円滑な動作維持などの用途のために、タスクの実行状態に応じてコントローラ１３７によって選択され、制御インターフェイス１７０を介して通信された制御動作を実行する。

【００３３】

システム１００は、システム１００を他のシステムおよび装置に接続するための複数の入力インターフェイス１０８および出力インターフェイス１１６を含むことができる。例えば、ネットワークインターフェイスコントローラ１５０は、バス１０６を介して、システム１００をネットワーク１９０に接続するように構成される。システム１００は、無線または有線ネットワーク１９０を介して、音波混合物入力信号１９５を受信することができる。いくつかの実装形態において、システム１００内のヒューマンマシンインターフェイス１１０は、システムをキーボード１１１およびポインティング装置１１２に接続する。ポインティング装置１１２は、とりわけ、マウス、トラックボール、タッチパッド、ジョイスティック、ポインティングスティック、スタイラス、またはタッチスクリーンを含むことができる。システム１００は、インターフェイス１１０またはNIC１５０を介して、機械１０２の動作中に生成された音波混合物信号１９５などのデータを受信することができる。

【００３４】

システム１００は、機械１０２の動作中に各アクチュエータ１０３によって生成された振動に対応する分離済み音波信号、または分離済み音波信号に基づいて動作する実行状態推定システム１３５の出力を出力するように構成された出力インターフェイス１１６を含む。例えば、出力インターフェイスは、分離済み音波信号または状態推定結果をレンダリングするためのメモリを含むことができる。例えば、システム１００は、バス１０６を介して、ディスプレイインターフェイス１８０に接続されてもよい。ディスプレイインターフェイス１８０は、システム１００を、例えば、スピーカ、ヘッドフォン、コンピュータモニタ、カメラ、テレビ、プロジェクタ、またはモバイル装置のようなディスプレイ装置１８５に接続するように構成される。また、システム１００は、アプリケーションインターフェイス１６０に接続されてもよい。アプリケーションインターフェイス１６０は、システムを様々な動作を実行するための設備１６５に接続するように構成される。

【００３５】

システム１００は、格納された命令を実行するように構成されたプロセッサ１２０と、

プロセッサによって実行可能な命令を格納するメモリ 140 とを含む。プロセッサ 120 は、シングルコアプロセッサ、マルチコアプロセッサ、コンピューティングクラスタ、または任意の数の他の構成を含むことができる。メモリ 140 は、ランダムアクセスメモリ (RAM)、読み取り専用メモリ (ROM)、フラッシュメモリ、または任意の他の適切なメモリシステムを含むことができる。プロセッサ 120 は、バス 106 を介して、1つ以上の入力装置および出力装置に接続される。これらの命令は、性能の推定および将来の制御を行うために、機械 102 の動作中に生成された振動信号を分離するための方法を実装する。

【0036】

いくつかの実施形態は、タスクを実行するシステムの状態を検討する代わりに、タスク自体の実行状態を検討することができるという認識に基づいている。例えば、システムが1つ以上のタスクを実行するための複数のアクチュエータを含む場合、システムの状態は、これらのアクチュエータの全ての状態を含む。しかしながら、場合によって、アクチュエータの状態は、主な制御対象ではない。実際に、性能、すなわち、タスクの実行を導くために、アクチュエータの状態が必要である。したがって、実行の状態は、主な目的であり、タスクを実行するアクチュエータの状態は、二次的な目的に過ぎない。

【0037】

しかしながら、場合によって、システムの状態を測定することが困難であり、タスクの実行状態を定義することが困難である。例えば、切削ツールでワークピースを機械加工するコンピュータ数値制御 (CNC) を検討する。システムの状態は、ツール経路に沿って切削ツールを移動させるアクチュエータの状態を含む。機械加工の実行状態は、実際の切削状態である。産業用 CNC システムは、多くの異なるアクチュエータ、場合によっては冗長なアクチュエータを含み、これらのアクチュエータは、互いに複雑な非線形関係にある多数の状態変数を有するため、CNC システムの状態を観測することが困難となる。したがって、ワークピースの機械加工状態を測定することも困難であり得る。

【0038】

いくつかの実施形態は、タスクの実行状態が、タスクの実行によって生成された音波信号で表すことができるという認識に基づいている。例えば、ワークピースの CNC 機械加工の実行状態は、機械加工中のワークピースの変形によって生成された振動信号で表すことができる。したがって、このような振動信号を測定できれば、機械学習法を含む様々な分類技術を用いてこのような振動信号を分析することによって、タスクの実行状態を推定することができる、実行を制御するための適切な制御動作を選択することができる。

【0039】

しかしながら、このアプローチの問題は、このような振動信号が単独で存在しないことである。例えば、1つ以上のタスクを実行する1つ以上のツールを作動させる複数のアクチュエータを含むシステムにおいて、タスクを実行するツールによって生成された信号は、常に、ツールを作動させる複数のアクチュエータによって生成された信号が混合されている。例えば、ワークピースの変形によって生成された振動信号には、常に、切削ツールを移動させるモータからの信号が混合されている。何らかの方法でこのような振動信号を単独で生成することまたは単独で捕捉することができれば、このような信号を抽出するように、ニューラルネットワークなどの機械学習システムを訓練することができる。しかしながら、CNC 機械加工を含む多くの場合に、タスクの実行を表す孤立信号の生成は、非現実である。同様に、複数のマイクロフォンで異なる信号を個別に記録することも、非現実である。

【0040】

図 1C は、いくつかの実施形態に従う、タスク 102 を実行するツールによって生成された信号とツール 103 を作動させる複数のアクチュエータによって生成された信号との音波混合物信号のスペクトログラム 196 を示す。センサ 101 が単一のマイクロフォンである場合に、全ての機械アクチュエータ 103 によって生成された信号は、時間および周波数 196 の両方で重なることがある。この場合、様々な音波信号処理技術、例えば周

10

20

30

40

50

波数または時間の選択フィルタリングは、機械アクチュエータによって生成された個々の音波信号を効果的に分離することができない。センサ 101 がマイクロフォンアレイである場合に、アレイ内の異なるマイクロフォン間の遅延の差を考慮することによって、異なる空間位置で音波信号を生成する音源を分離することができる。例えば、図 1C において、機械 102 は、C1、...、C5 によって示されるアクチュエータ 103 を含み、アクチュエータ C3 が一意的な空間位置に配置されているため、マイクロホンアレイは、アクチュエータ C3 によって生成された信号を分離することができる。しかしながら、信号対 (C1, C2) および (C4, C5) が、機械 102 における物理位置によって空間的に重なり、音波信号スペクトログラム 196 において時間および周波数の両方で重なるため、従来の技術は、これらを分離することができない。

10

【0041】

したがって、1つ以上のツールを作動させて1つ以上の作業を実行する複数のアクチュエータを含むシステムの製造自動化を効率化するために、作業を実行するツールによって生成された信号と、ツールを作動させる複数のアクチュエータによって生成された信号とを音波混合物から音源信号を分離する必要がある。したがって、いくつかの実施形態の目的は、単独の音源が存在していない場合に、混合信号の音源を分離するためのニューラルネットワークを訓練することである。本明細書に使用された場合、混合信号中の少なくとも一部の音源は、音波混合物において同じ時間、空間および周波数スペクトルを占有する。

【0042】

場合によって、少なくとも一部の音源は、音波混合物のスペクトログラムにおいて同じ時間および／または周波数スペクトルを占有する。例えば、一実施形態において、混合物中の音源は、室内の同じ領域を占有するため、マイクロフォンアレイ技術の使用は、非現実である。追加的にまたは代替的に、いくつかの実施形態の音波混合物は、単一のマイクロフォンの単一の出力チャンネルのみから得られる。

20

【0043】

いくつかの実施形態は、例えば、音波混合物中の音源を単独で容易に記録できる場合または人間が音波混合物中の音源を容易に識別できる場合に、教師付き学習で訓練を行うことができるという認識に基づいている。このような分類は、本明細書において強ラベルと呼ばれる。理論的には、人間は、このようなラベルの近似値を提供することができるが、時間および周波数の両方でアクティブ音を正確に分類することによって、強ラベルを付けることは、殆ど非現実である。しかしながら、一部の時間範囲内でアクティブ音を限定的に分類すること（弱ラベル）は、現実的である。通常の場合、このような弱ラベルは、指定範囲の全体にわたって音がアクティブである必要がなく、指定範囲内の短時間のみにアクティブであればよい。

30

【0044】

いくつかの実施形態は、通常、製造プロセスの仕組みが既知であれば、音波信号を生成する音源も既知であるという認識に基づいている。したがって、製造システムの動作を表す音波混合物の音源に弱ラベルを付けることができる。よって、いくつかの実施形態は、弱ラベルのみを含む音波混合物の訓練データが利用可能である場合に、この音波混合物中の音波を分離するように学習することができる方法を開発する。

40

【0045】

したがって、いくつかの実施形態は、音波混合物から、タスクを実行するツールによって生成された信号とツールを作動させるアクチュエータによって生成された信号とを分離するようにニューラルネットワーク 131 を訓練する。例えば、ニューラルネットワークは、各分離済み信号がシステムの動作中に存在する 1 種類の信号であり、分離済み信号が音波混合物の合計を構成するように、音波混合物から異なる信号を分離するように訓練される。弱ラベルは、システムの動作中に存在する信号の種類を識別する。各弱ラベルは、動作中のある時点に存在する信号の種類を特定する。

【0046】

図 2 は、本開示のいくつかの実施形態に従って、音波混合物信号を分離するための音波

50

信号処理システム 200 を訓練することを示す流れ図である。図 2 は、一般的な音源分離シナリオを示す。このシナリオにおいて、システムは、ターゲット音波信号とノイズなどの他の可能な非ターゲット音波信号との混合物から、複数のターゲット音波信号を推定する。ターゲット音波信号が機械のアクチュエータの振動によって生成され、これらの振動が単独で存在していない例において、音源分離システムの訓練ターゲットは、弱ラベルによって識別される。すなわち、特定の時間ブロックにおいて、孤立の音源ではなく、音源の有無のみが訓練に必要とされる。音波混合物入力信号 204 は、複数の重なる音源の合計を含み、機械の動作中に記録された音波混合物信号および対応する弱ラベル 222 を含む訓練セット 202 からサンプリングされる。

【0047】

スペクトログラム推定器 206 は、混合入力信号 204 を処理することによって、音波混合物の時間一周波数表現を計算する。その後、スペクトログラムは、記憶されたネットワークパラメータ 215 を使用するマスク推論ネットワーク 230 に入力される。マスク推論ネットワーク 230 は、スペクトログラムの各時間周波数ビンにおいて各クラスの音源の有無を判断し、一組の振幅マスク 232 を推定する。各音源は、1つの振幅マスクを有する。各マスクに音波混合物の複雑な時間一周波数表現を乗算することによって、一組の拡張スペクトログラム 234 を計算する。拡張スペクトログラム 234 の各々を信号再構成プロセス 236 に渡し、スペクトログラム推定器 206 によって計算された時間一周波数表現を反転させることによって、一組の推定音波信号波形 216 を取得する。

【0048】

拡張スペクトログラム 234 を、格納されたネットワークパラメータ 215 を使用する分類器ネットワーク 214 に渡すことができる。分類器ネットワークは、各拡張スペクトログラムの各時間フレームに所定の音源クラスが存在するか否かに関する確率を提供する。分類器は、スペクトログラムの各時間フレームに対して 1 回動作するが、弱ラベル 222 は、スペクトログラムよりも遥かに低い時間分解能を有する。したがって、分類器ネットワーク 214 の出力は、時間プーリングモジュール 217 に渡される。いくつかの実施形態において、時間プーリングモジュール 217 は、1つの弱ラベル時間フレームに対応する全てのフレームレベル判定値の最大値をとることができ、フレームレベル判定値の平均値をとることができ、または他のプーリング操作を使用して、フレームレベル判定値を結合することができる。ネットワーク訓練モジュール 220 は、目的関数を用いて、ネットワークパラメータ 215 を更新することができる。

【0049】

図 3 A は、本開示の実施形態に従って、単一チャネルマスク推論ネットワークのアーキテクチャ 300 A を示すブロック図である。入力混合物から得られた特徴ベクトルのシーケンス、例えば、入力混合物の短時間フーリエ変換の対数振幅は、混合物エンコーダ 310 に入力される。例えば、シーケンス内の入力ベクトルの次元は、F で表すことができる。混合物エンコーダ 310 は、第 1 の BLSTM 層 330 から最後の BLSTM 層 335 までの双方向長短期メモリ (BLSTM) ニューラルネットワークの複数の層から構成される。各 BLSTM 層は、順方向長短期メモリ (LSTM) 層および逆方向 LSTM 層からなる。各 BLSTM 層の出力は、結合され、次の層に入力される。例えば、第 1 の BLSTM 層 330 内の各 LSTM の出力の次元を N で表すことができ、最後の BLSTM 層 335 を含む他の全ての BLSTM 層内の各 LSTM の入力と出力の両方の次元を N で表すことができる。最後の BLSTM 層 335 の出力は、線形ニューラルネットワーク層 340 および非線形性 345 を含むマスク推論モジュール 312 に入力される。時間一周波数領域、例えば短時間フーリエ変換領域における各時間フレームおよび各周波数に対して、線形層 340 は、最後の BLSTM 層 335 の出力を使用して、分離すべきターゲット音源の数である C 個の数値を出力する。非線形性 345 を各時間フレームおよび各周波数の C 個の数値に適用することによって、各時間フレーム、各周波数、各ターゲット音源、およびその時間フレームおよびその周波数における入力混合物の優勢ターゲット音源を示すマスク値を生成する。マスクモジュール 313 による分離符号化推定は、マスクを推定

10

20

30

40

50

した時間一周波数領域、例えば振幅短時間フーリエ変換領域における入力混合物の表現と共に、これらのマスクを使用して、各ターゲット音源についての分離符号化を出力する。例えば、マスクモジュール313による分離符号化推定は、ターゲット音源のマスクに入力混合物の振幅短時間フーリエ変換を乗算することによって、単独で観測された場合に、そのターゲット音源の分離符号化として使用される分離済み信号の振幅短時間フーリエ変換の推定値を得ることができる。

【0050】

図3Bは、本開示の実施形態に従って、単一チャネル畳み込みリカレント型分類ネットワークのアーキテクチャ300Bを示すブロック図である。入力混合物から得られた特徴ベクトルのシーケンス、例えば、入力混合物の短時間フーリエ変換の対数強度は、混合物エンコーダ320に入力される。例えば、シーケンス内の入力ベクトルの次元は、Fで表すことができる。混合物エンコーダ320は、第1の畳み込みブロック301から最後の畳み込みブロック302までの複数の畳み込みブロック、およびそれに続くリカレントBLSTM層303から構成される。各畳み込みブロックは、学習済み重みおよびバイアスを有する畳み込み層を含み、時間次元および周波数次元の両方に対するプーリング操作を行うことによって、後続層への入力次元を低減する。混合物エンコーダ320の末端のBLSTM層303は、順方向長短期メモリ(LSTM)層および逆方向LSTM層からなる。BLSTM層303の出力は、結合され、分類器モジュール322への入力として用いられる。分類器モジュール322は、線形ニューラルネットワーク層305と、シグモイド非線形性307を実装するモジュールとを含む。分類器モジュール322は、入力信号の各時間フレームに対してC個の数値を出力する。これらの数値は、当該時間フレームにおいて特定の種類の音源がアクティブである確率を表す。

【0051】

図4は、データの収集および単独で存在することまたは互いに空間的に離れる必要のない音波混合物信号を分離するためのニューラルネットワークの訓練を行うための方法を示す流れ図である。機械または分離すべき信号を含む音環境からの一連の音波記録を収集することによって、一組の訓練データ記録400を形成する。これらの記録に注釈を付けることによって、全ての訓練データ信号410に弱ラベルを生成する。この注釈手順は、訓練データ記録400を収集した時の音源がアクティブおよび非アクティブである期間中に取られたメモを含むことができる。追加的にまたは代替的に、訓練データ記録400を聞き、分離すべき音源がアクティブおよび非アクティブである期間中に注釈を付けることによって、全ての訓練データ信号410の弱ラベルを法的(forensically)に収集することができる。追加的にまたは代替的に、弱ラベルは、被制御機械の動作の仕様に基づいて決定される。

【0052】

次のステップは、全ての訓練データ信号410の弱ラベルをターゲットとして用いて、各訓練データ信号400の弱ラベル420を予測するように分類器を訓練することである。この分類器は、音事象検出システムとも呼ばれる。続いて、分類器を監視信号430として用いて、分離器ネットワークを訓練する。これは、以下でさらに詳細に説明する。次いで、分離器ネットワークは、一組の信号である分離された音源440を抽出する。各分離済み信号は、単一のクラスまたは信号種類に属する。これらの分離済み信号を用いて分類器420をさらに訓練することによって、分離済み信号のクラスのみを予測する一方で、他の全てのクラスの出力はゼロである。さらに、以前に分離された信号を入力として用いて分離器ネットワーク430を訓練することができ、正しく分離された信号が分離器ネットワークを不変に通過するようにネットワークの重みを更新することができる。

【0053】

図5は、いくつかの実施形態に従う、弱ラベルと信号分離を促進する分類器出力との間の関係を示す概略図である。これらの実施形態において、ニューラルネットワークは、弱ラベルと分類器を通る分離ニューラルネットワークとの間のクロスエントロピー項を含む損失関数を最小化するように、分類器と共同で訓練される。図5の例示は、単一の時間フ

10

20

30

40

50

レーンに対するものであるが、弱ラベルを利用できる時間分解能で繰り返される。

【 0 0 5 4 】

【数 1】

分離ネットワークによって抽出されたC個の信号の各々を分類器に通過させることによって、分類器出力5 1 0が取得される。すなわち、分類器出力行列5 1 0の各行は、1つの分離済み信号に対する分類器出力に対応する。弱ラベルターゲット行列5 3 0は、分離システムの訓練に使用できるような方式で、与えられた弱ラベルを配置する。C個の信号種類の各々に対して、種類*i*の信号が存在しない場合 $t_i=0$ であり、種類*i*の信号が存在する場合 $t_i=1$ である。分類器を訓練する場合、クロスエントロピー損失関数5 2 0は、音波混合物から得られた分類器出力、および弱ラベルターゲット行列5 3 0の対角要素のみを必要とし、以下のように数学的に表すことができる。

$$H(t_i, p_i) = - \sum_{i=1}^C [t_i \log(p_i) + (1 - t_i) \log(1 - p_i)]$$

式中、

$i=1, \dots, C$ 、

p_i は、種類*i*の音波混合物信号に対して動作するときの分類器出力である。

【 0 0 5 5 】

【数 2】

しかしながら、分離済み信号を必要とするため、実施形態は、分類器出力行列5 1 0の非対角項を、弱ラベルターゲット行列5 3 0においてゼロに等しくなるように強制する。これによって、各分離済み信号は、単一の音源クラスのみに属する。したがって、分離のためのクロスエントロピー損失関数5 2 0は、以下のように数学的に表すことができる。

$$L_{class} = \sum_{i=1}^C H(t_i, p_{ii}) + \sum_{i=1}^C \sum_{j \neq i}^C H(0, p_{ij})$$

式中、

$H(t_i, p_{ii})$ は、上記に定義されたクロスエントロピー損失である。

【数 3】

10

20

30

40

50

分類損失を用いて分離システムを訓練することに面する 1 つの問題は、分類器がしばしば利用可能な周波数スペクトルのごく一部に基づいて判断を行うため、スペクトルの一部のみを分離するように分離器を学習させる場合、分離信号を抽出する目的を達成できないことである。これを避けるために、他の損失項を追加することによって、全てのアクティブ音源の抽出信号の合計が音波混合物信号を形成し、非アクティブ音源のエネルギーにペナルティーを科す。この関数は、以下のように数学的に表すことができる。

$$L_{Mag} = \sum_f \left| X(f) - \sum_{i=1}^C t_i * \hat{S}_i(f) \right| + \sum_f \sum_{i=1}^C (1 - t_i) * \hat{S}_i(f) \quad 10$$

式中、

f は、周波数であり、

$X(f)$ は、音波混合物の振幅であり、

$\hat{S}_i(f)$ は、離れた音源の振幅である。

最後に、これらの 2 つのアクチュエータを組み合わせると、全体損失関数は、以下のよう
に実行される：

$$L_{overall} = L_{class} + \alpha L_{Mag} \quad 20$$

式中、

α は、各損失アクチュエータの相対重要度を重み付けすることを可能にする項である。

【 0 0 5 6 】

このように、本開示は、音波分離ニューラルネットワークを用いて製造自動化を行うためのシステムおよび方法を提供する。本開示によって提案された機械学習アルゴリズムは、ニューラルネットワークを用いて、製造時に異なる部品の振動によって生成された個々の信号を分離する。また、音波混合物を構成する振動信号は、単独でまたは空間的に互いに分離できる形で存在していないため、弱ラベルを用いて、提案されたシステムを訓練することができる。この場合、弱ラベルは、音波混合物を構成する分離済み信号ではなく、音波混合物の異なる音源がアクティブである期間のみを利用できることを意味する。いったん音波混合物を構成する信号が分離されると、分離済み信号を用いて、個々の機械部品の実行状態を監視することができ、または分離済み信号を制御信号として用いて、異なるアクチュエータをそれぞれ制御することができる。

30

【 0 0 5 7 】

図 6 は、いくつかの実施形態に従って分離された信号に基づいて、実行状態を分析することを示すグラフである。信号を分離するためにニューラルネットワーク 1 3 1 を訓練した後、いくつかの実施形態は、このネットワークの出力を用いて、タスクの実行状態を推定するための状態推定器 1 3 5 を訓練する。例えば、一実施形態において、状態推定器は、タスクを実行するツールによって生成された信号 6 0 1 およびニューラルネットワークによってシステムの異なる反復動作の異なる音波混合物から抽出された信号 6 0 2 に基づいて訓練される。

40

【 0 0 5 8 】

例えば、いくつかの実施形態において、状態推定器 1 3 5 は、実行状態を示す予測パターン 6 1 5 を検出するように構成される。例えば、期間 6 1 7 にわたって収集された孤立信号の実数値時系列は、正常領域 6 1 8 と、故障点 6 2 1 をもたらす異常領域 T 6 1 9 とを含むことができる。状態推定器 1 3 5 は、異常領域 6 1 9 を検出することによって、障害 6 2 1 を回避するように構成されてもよい。例えば、いくつかの実装形態において、状態推定器 1 3 5 は、Shapelet 発見法を用いて、最良の予測パターンを発見するまで予測パターンを検索する。Shapelet 発見アルゴリズムの少なくとも 1 つの利点は、異なる長さの

50

予測パターンを効率的に検索できることである。内部では、Shapelet発見法は、所定の測定基準に従って予測パターンを最適化している。予測パターンは、例えば、異常領域内の1つのパターンとは可能な限り類似し、正常領域内の全てのパターンとは可能な限り異なる。

【0059】

追加的にまたは代替的に、別の実施形態において、状態推定器135は、抽出ニューラルネットワーク131の出力からタスクの実行状態を推定するように訓練されたニューラルネットワークとして実装される。有利なことに、本実施形態の状態推定器は、ニューラルネットワーク131およびニューラルネットワーク131を訓練するために使用される分類器の両方と共同で訓練されてもよい。したがって、本実施形態において、複雑な機械を制御するためのエンドツーエンド訓練解決策を提供する。

10

【0060】

特に、いくつかの実施形態において、タスクを実行するツールによって生成された抽出信号の各サンプルは、タスクの実行状態を定義し、システムの動作状態を定義するには不十分である。しかしながら、タスクの実行状態は、適切な制御動作を選択するのに十分である。このように、実施形態は、製造自動化用途において次元を削減することができる。

【0061】

例えば、いくつかの実施形態において、システムは、ワークピースを機械加工するための一連の操作を実行するように構成される。これらの実施形態において、ツールは、機械加工ツールであり、プロセッサは、ツール経路に沿って機械加工ツールを作動させるコンピュータ数値制御(CNC)を実行し、機械加工ツールによって生成された信号は、機械加工ツールによる機械加工中にワークピースの変形によって生成された振動信号である。多くの機械加工システムは、ツールを位置決めするための複数のアクチュエータを含む。また、いくつかの機械加工システムは、ツールを各自由度に沿って位置決めするための冗長なアクチュエータを含むことができる。さらに、ツールの種類も、システムの性能に影響を与える可能性がある。しかしながら、これらの全ての種類は、タスクの実行を示す信号の分離および分類に基づく実施形態によって捕捉することができる。

20

【0062】

図7は、いくつかの実施形態に従う、機械加工動作の制御を示す概略図である。一組の機械加工命令701は、例えば、ネットワークを介してファイルとしてNC機械加工コントローラ702に提供される。コントローラ702は、プロセッサ703と、メモリ704と、機械の動作を表示するためのディスプレイ705とを含む。いくつかの実施形態に従って、プロセッサは、抽出ニューラルネットワーク131、状態機械加工の状態推定135、およびコントローラ動作137を実行する。いくつかの実装形態において、ニューラルネットワーク131、状態推定器135およびコントローラ137は、ワークピース710の異なる種類の機械加工712、714、716および718を実行する機械加工の異なるツール702、704、706および708に適合される。例えば、被制御機械は、ワークピースの機械加工、はんだ付けおよび組み立てのうちの1つまたは組み合わせを含むワークピースを製造するための一連の操作を実行することができ、ツールによって生成された信号は、ワークピースの製造中にワークピースの変更によって生成された振動信号である。

30

40

【0063】

要するに、抽出された信号に基づく状態推定は、制御を異なる種類の複雑な製造に適合させる。しかしながら、いくつかの実施形態は、工場の自動化に限定されない。例えば、一実施形態において、被制御機械は、潜在的な異常を検出するために監視されている変速装置であり、変速装置は、モータまたはカップリングからの振動、あるいは可動部からの他の振動が存在する場合のみ記録されてもよい。

【0064】

次元削減に加えてまたはそれに代えて、いくつかの実施形態は、異なる作業タスクを実行する異なるツールを個別に制御することができる。例えば、システム全体の状態を示す

50

信号を分析すれば、システム全体の制御を行うことができる。したがって、異なるタスクを実行する異なるツールを個別に分析する場合、異なるツールを個別に制御することができる。

【0065】

したがって、システムの動作中に複数のツールが複数のタスクを実行する場合、いくつかの実施形態は、複数のタスクを個別に制御する。例えば、一実施形態は、第1のタスクを実行する第1のツールと第2のタスクを実行する第2のツールとを含むシステムを制御する。ニューラルネットワークは、音波混合物から、第1のタスクを実行する第1のツールによって生成された第1の信号と第2のタスクを実行する第2のツールによって生成された第2の信号とを分離するように訓練される。システムの動作中に、ニューラルネットワークを用いて音波混合物から第1の信号および第2の信号を抽出し、互いに独立して分析することによって、第1のタスクの第1の実行状態と第2のタスクの第2の実行状態とを推定する。この実施形態は、第1の状態に応じて選択された第1の制御動作を実行し、第2の状態に応じて選択された第2の制御動作を実行するように構成される。

10

【0066】

図8は、いくつかの実施形態に従って、製造異常検知システム800のアクチュエータを示す概略図である。システム800は、製造生産ライン810と、訓練データプール820と、機械学習モデル830と、異常検出モデル840とを含む。生産ライン810は、センサを用いてデータを収集する。センサは、デジタルセンサ、アナログセンサ、およびこれらの組み合わせであってもよい。収集されたデータは、2つの目的に使用される。一部のデータは、訓練データプール820に格納され、機械学習モデル830を訓練するための訓練データとして使用され、一部のデータは、異常を検出するように、動作時間データとして異常検出モデル840によって使用される。機械学習モデル830と異常検出モデル840の両方は、同じデータを使用してもよい。

20

【0067】

製造生産ライン810の異常を検出するために、まず訓練データを収集する。機械学習モデル830は、訓練データプール820内の訓練データを用いて、抽出ニューラルネットワーク131を訓練する。訓練データプール820は、ラベル付きデータまたはラベルなしデータのいずれかを含むことができる。ラベル付きデータは、例えば、異常または正常を示すラベルを有する。ラベルなしデータは、ラベルを有していない。機械学習モデル830は、訓練データの種に基づいて、異なる訓練アプローチを適用する。ラベル付き訓練データの場合、通常、教師付き学習が使用される。ラベルなし訓練データの場合、通常、教師なし学習が使用される。このように、異なる実施形態は、異なる種類のデータを取り扱うことができる。

30

【0068】

機械学習モデル830は、訓練データから、正常データパターンおよび異常データパターンを含む特徴およびパターンを学習する。異常検出モデル840は、訓練済み機械学習モデル850および収集された運転時間データ860を用いて、異常を検出する。動作時間データ860は、正常データであってもよく、異常データであってもよい。例えば、訓練済み機械学習モデル850は、正常データパターン855および858を用いて、動作時間データを正常データ870および異常データ880に分類することができる。例えば、操作時間データX1 863およびX2 866は、正常データとして分類され、操作時間データX3 869は、異常データとして分類される。異常を検出すると、必要な行動890を取る。

40

【0069】

いくつかの実施形態において、抽出ニューラルネットワークは、各監視プロセスX1 863およびX2 866およびX3 869に対して訓練される。コントローラは、1つのプロセス、例えばプロセスX1 863を、他のプロセス、例えばプロセスX2 866およびX3 869とは独立して制御するための行動890を取ることができる。信号抽出に基づいたプロセス制御分離は、複雑な製造プロセスの制御を単純化することによって

50

、より正確且つ実用的に制御を行うことができる。

【0070】

上述した本発明の実施形態は、多くの方法で実装されてもよい。例えば、実施形態は、ハードウェア、ソフトウェア、またはそれらの組み合わせで実装されてもよい。ソフトウェアで実装される場合、ソフトウェアコードは、単一のコンピュータに設けられたまたは複数のコンピュータに分散されたことにも拘らず、任意の適切なプロセッサまたは一群のプロセッサで実行されてもよい。このようなプロセッサは、集積回路として実装されてもよい。1つの集積回路要素は、1つ以上のプロセッサを含むことができる。しかしながら、プロセッサは、任意の適切な回路で実装されてもよい。

【0071】

また、本発明の実施形態は、一例として提供された方法として具現化されてもよい。本方法の一部として実行される動作は、任意の適切な方法で順序付けられてもよい。したがって、例示的な実施形態において順次に実行される動作とは異なる順序で動作を実行すること、いくつかの動作を同時に実行することを含み得る実施形態を構築することができる。

【0072】

請求項において請求項要素を修飾するための順序用語、例えば第1、第2などの使用は、別の請求項要素に対する1つの請求項要素の優先順位、前後順位もしくは順序、または方法の動作を実行する時間順序を意味しておらず、単に請求項要素を区別するためのラベルとして使用され、（順序用語を使用することによって）特定の名前を有する1つの請求項要素と同じ名前を有する別の要素とを区別させる。

【0073】

いくつかの好ましい実施形態を参照して本発明を説明したが、理解すべきことは、本発明の精神および範囲内で、様々な他の改造および修正を行うことができることである。

【0074】

したがって、添付の特許請求の範囲は、本開示の真の精神および範囲内にある全ての変形および修正を網羅する。

10

20

30

40

50

【図面】

【図 1 A】

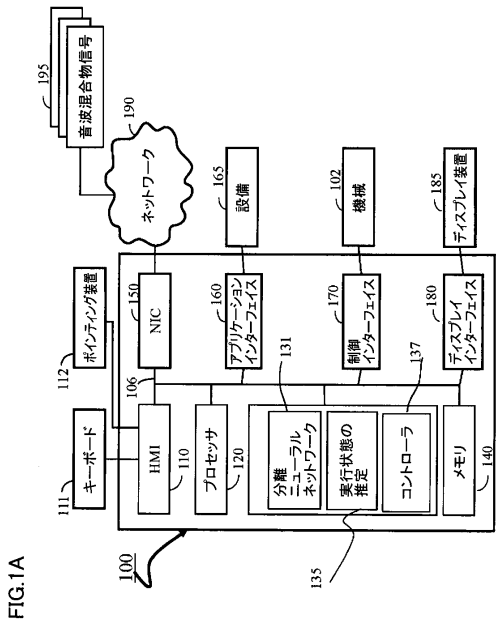


FIG. 1A

【図 1 B】

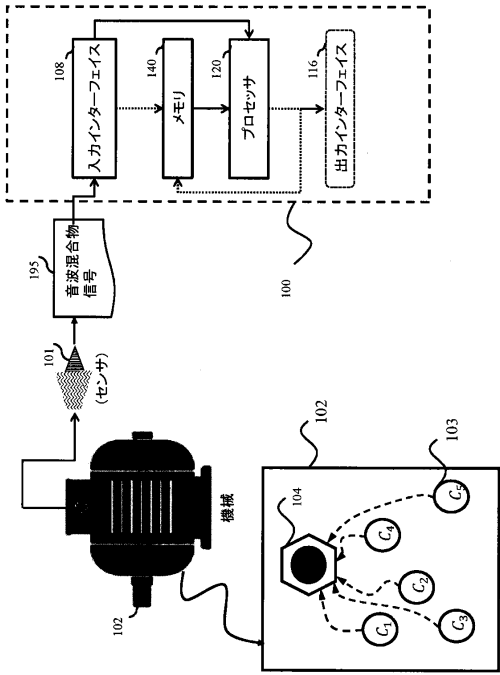


FIG. 1B

【図 1 C】

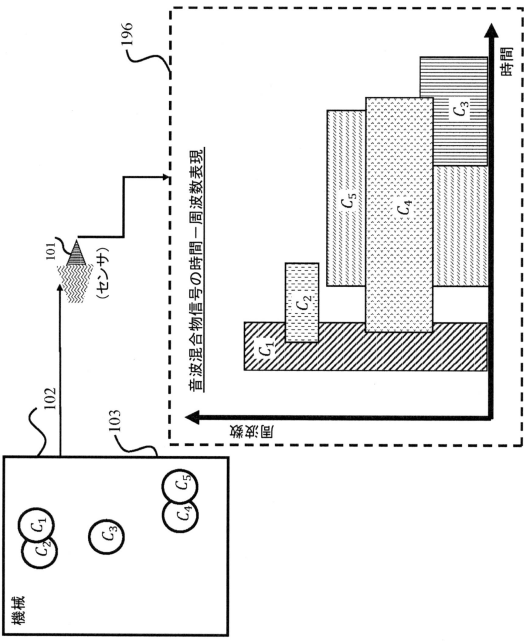


FIG. 1C

【図 2】

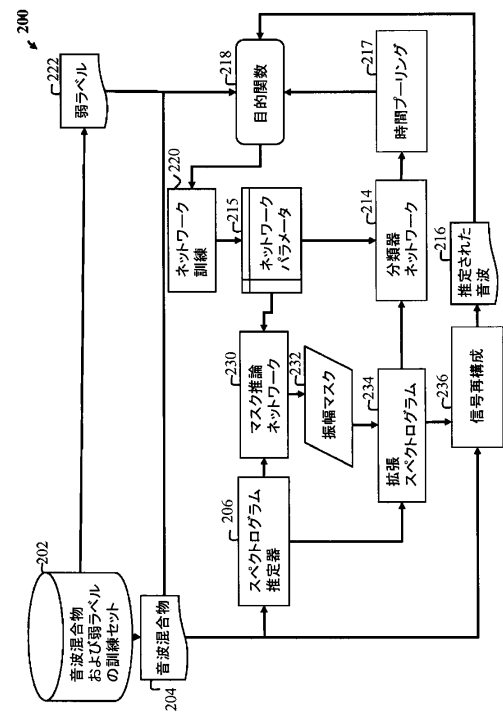


FIG. 2

10

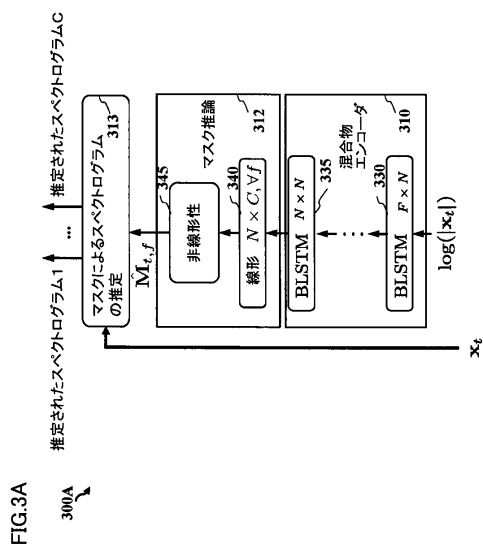
20

30

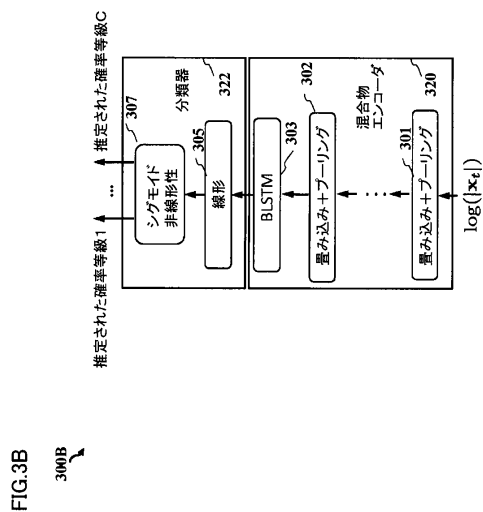
40

50

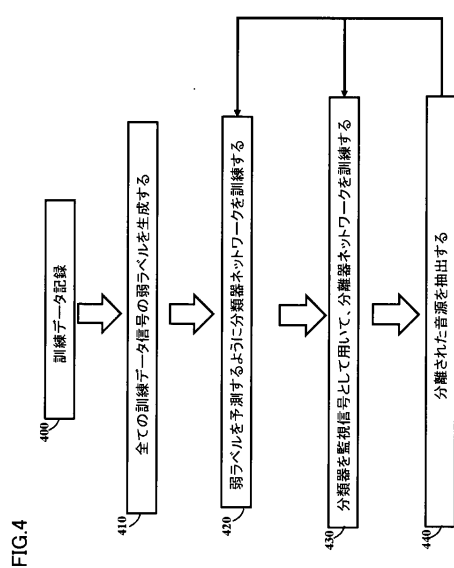
【図 3 A】



【図 3 B】



【図 4】



【図 5】

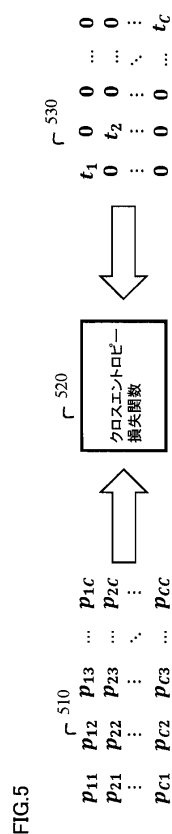
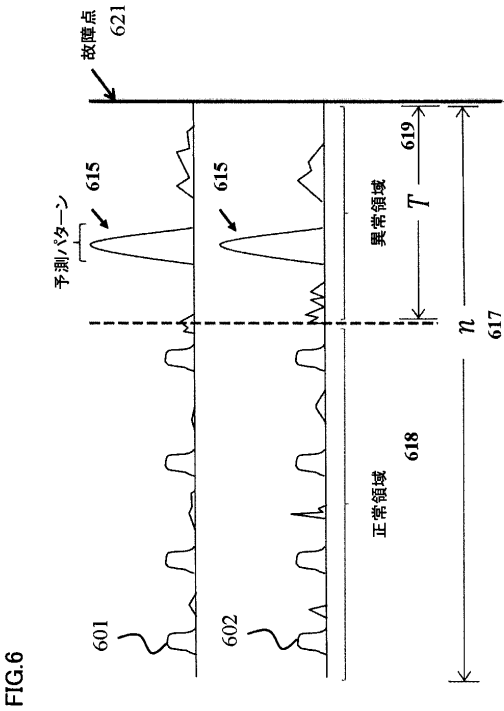
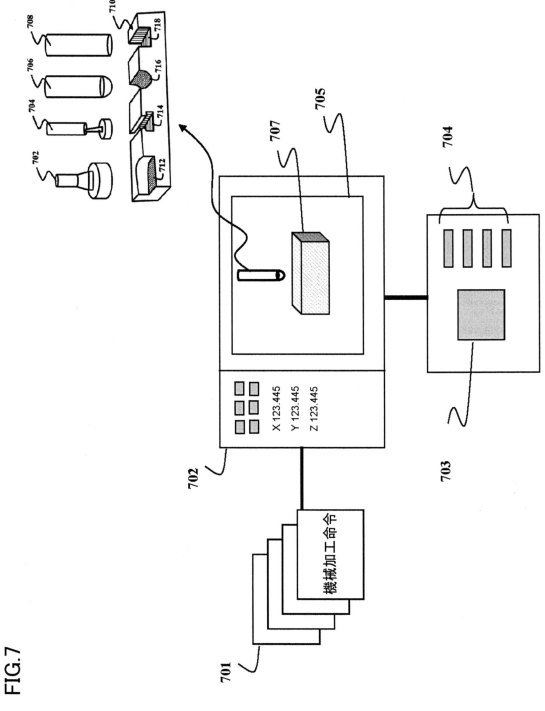


FIG. 5

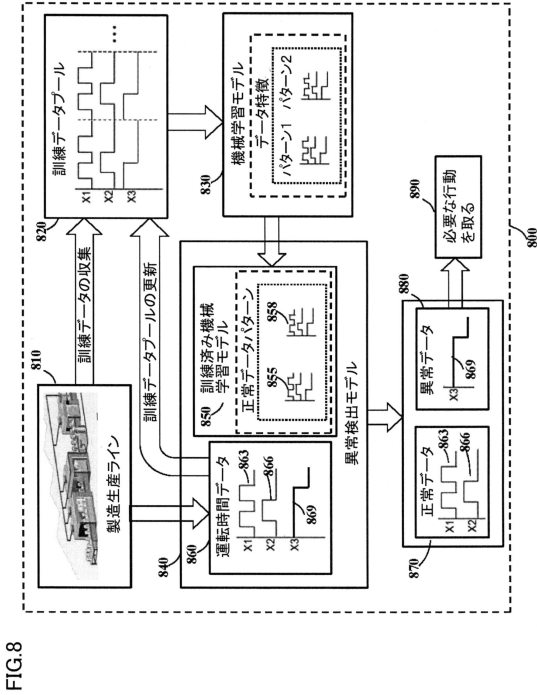
【図 6】



【図 7】



【図 8】



10

20

30

40

50

フロントページの続き

ロードウェイ 201、エイス・フロア、ケアオブ・ミツビシ・エレクトリック・リサーチ・ラボ
ラトリーズ・インコーポレイテッド

(72)発明者 ピシュダディアン, ファテメ

アメリカ合衆国、02139-1955 マサチューセッツ州、ケンブリッジ、ブロードウェイ 2
01、エイス・フロア、ケアオブ・ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・イン
コーポレイテッド

審査官 小川 真

(56)参考文献 特開2014-213412 (JP, A)

中国特許出願公開第110070183 (CN, A)

特開2008-137102 (JP, A)

国際公開第2018/105194 (WO, A1)

国際公開第2016/125256 (WO, A1)

Gubin, M.V., " Using Convolutional Neural Networks to Classify Audio Signal in Noisy Sound
Scenes ", 2018 Global Smart Industry Conference (GloSIC), 2018年, p.1-6, DOI: <http://dx.doi.org/10.1109/GloSIC.2018.8570117>

(58)調査した分野 (Int.Cl., DB名)

G05B 19/18

G05B 23/02

G05B 19/4155

B23Q 17/00、17/09

G06N 20/00

G01H 5/00