

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2023年11月16日(16.11.2023)



(10) 国際公開番号

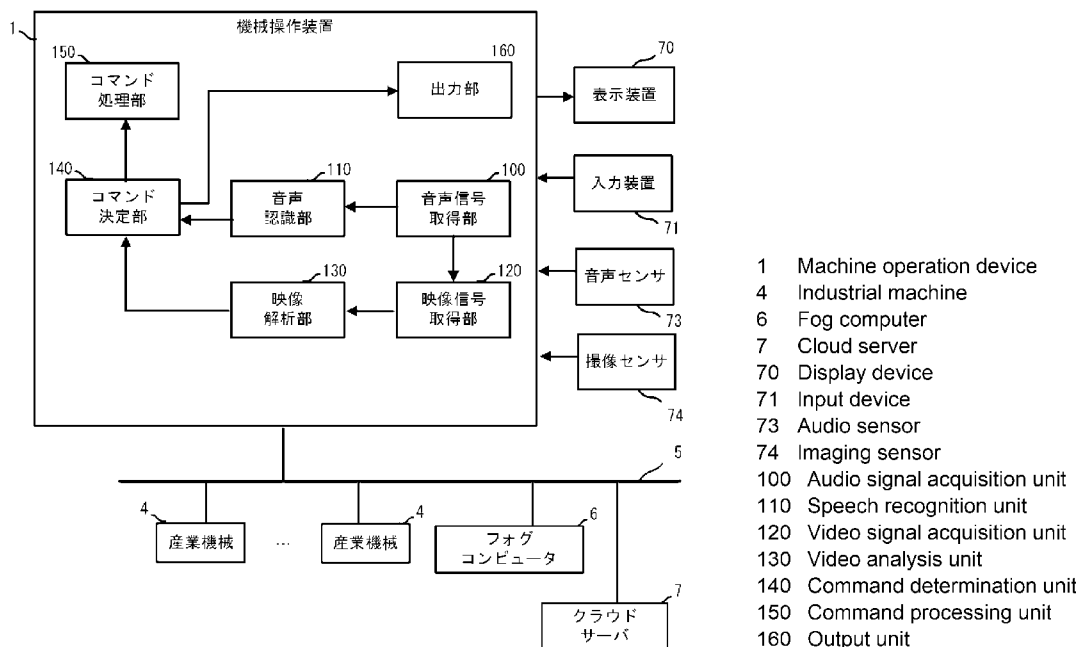
WO 2023/218522 A1

- (51) 国際特許分類:
G06F 3/16 (2006.01) *G06F 3/01* (2006.01)
- (21) 国際出願番号: PCT/JP2022/019792
- (22) 国際出願日: 2022年5月10日(10.05.2022)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: ファナック株式会社 (FANUC CORPORATION) [JP/JP]; 〒4010597 山梨県南都留郡忍野村忍草字古馬場 3 5 8 0 番地 Yamanashi (JP).
- (72) 発明者: 胡 蓮成(HU Liancheng); 〒4010597 山梨県南都留郡忍野村忍草字古馬場 3 5 8 0 番地 ファナック株式会社内 Yamanashi (JP).
- (74) 代理人: あいわ弁理士法人 (AIWA INTERNATIONAL PATENT AGENCY); 〒1040045 東京都中央区築地一丁目 1 2 番 2 2 号 コンビル4階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

(54) Title: MACHINE OPERATION DEVICE

(54) 発明の名称: 機械操作装置

[図2]



(57) Abstract: The machine operation device according to the present disclosure comprises: a speech recognition unit that recognizes first data indicating a predetermined operation on the basis of an audio signal into which speech uttered by a worker and detected by an audio sensor is converted; a video analysis unit that analyzes second data indicating a predetermined operation on the basis of a video signal obtained by imaging a motion of the worker using an imaging sensor; a command determination unit that determines an operation command for operating an industrial machine on

[続葉有]



MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

一 国際調査報告 (条約第21条(3))

the basis of the first data and the second data; and a command processing unit that processes the operation command determined by the command determination unit.

(57) 要約：本開示による機械操作装置は、音声センサにより検出された作業者が発話した音声を変換した音声信号に基づいて、所定の操作を示す第1のデータを認識する音声認識部と、撮像センサにより作業者の動きを撮像して得られた映像信号に基づいて、所定の操作を示す第2のデータを解析する映像解析部と、前記第1のデータ及び前記第2のデータに基づいて、産業機械を操作するための操作コマンドを決定するコマンド決定部と、コマンド決定部が決定した操作コマンドを処理するコマンド処理部と、を備える。

明 細 書

発明の名称：機械操作装置

技術分野

[0001] 本発明は、機械操作装置に関する。

背景技術

[0002] 現在、製造業などの産業分野では、ロボット、搬送機、工作機械、機械設備などの様々な産業機械が作動している。このような産業機械には、操作部を備えたものも多く、PLC (Programmable Logic Controller)、NC (Numerical Controller)、制御盤など、各産業機械を制御する機器も操作部を備える。

[0003] 産業機械の操作部は、ボタンや操作画面が多いが、操作が複雑で習熟に時間のかかることがある。音声入力インタフェースは、音声コマンドを発話するだけで目的の操作が実行できる。そのため、音声入力インタフェースを用いた操作性の向上が試みられている。音声入力インタフェースは、産業機械のドア開閉、産業機械の運転開始や一時停止、室内点消灯、切削液の使用停止、切粉の清掃など、あらゆる場面において当該産業機械に触れることなく指令することができる。また、ボタンや操作画面を直接操作する必要が無いため、例えば、試作の場合だと加工状況を見ながら操作したり、ロボットを安全柵外から操作したり、危険を感じる時にボタンに近寄るまでもなく速やかに音声で停止させることもできる。

[0004] 産業機械の操作に使用する音声コマンドは、音声コマンドを使用する産業機械の種類、産業機械を設置する現場、産業機械の操作内容などにより想定することができる。そのため、想定される音声コマンドを文法（構文及び単語）で作成することができる。例えば、特許文献1参照。

先行技術文献

特許文献

[0005] 特許文献1：特開平9－325787号公報

発明の概要

発明が解決しようとする課題

[0006] 産業機械が設置される工場などの製造現場では、通常、複数の産業機械が動作している。そして、製造現場では常に騒音が発生している。そのため、音声コマンドのみで産業機械を操作する場合、音声コマンドと同時に加工音などがマイクに入力される。これが原因で、音声コマンドの認識精度が低下することがある。騒音が発生している現場においては、音声よりもジェスチャなどを撮像し、撮像した映像からコマンドを認識する方法も考えられる。このようにすることで、騒音の影響を排除することができる。しかしながら、ジェスチャのみの場合、言語より指令の表現力に限界があり、使いにくい場面がある。

そのため、製造現場において生産性と精度とを両立させた機械操作の手法が望まれている。

課題を解決するための手段

[0007] 本発明による機械操作装置は、工作機械やロボット等の産業機械の操作に、マイク等で検出した音声に基づく音声認識と、カメラ等で検出した画像に基づく顔（口形）乃至ジェスチャ認識を融合させることで、手動操作より生産性を向上した機械操作を行えるようにし、上記課題を解決する。

[0008] そして、本開示の一態様は、製造現場において入力された音声信号を音声認識して産業機械を操作するための操作コマンドとして利用する機械操作装置であって、音声センサにより検出された作業者が発話した音声を変換した音声信号に基づいて、所定の操作を示す第1のデータを認識する音声認識部と、撮像センサにより作業者の動きを撮像して得られた映像信号に基づいて、所定の操作を示す第2のデータを解析する映像解析部と、前記第1のデータ及び前記第2のデータに基づいて、前記産業機械を操作するための操作コマンドを決定するコマンド決定部と、前記コマンド決定部が決定した操作コマンドを処理するコマンド処理部と、を備えた機械操作装置である。

発明の効果

[0009] 本開示の一態様により、作業の簡略化や効率化により生産性を向上するとともに、機械本体と非接触で機械操作が可能となるため、作業者の安全性を確保できるようになる。

図面の簡単な説明

[0010] [図1]本発明の一実施形態による機械操作装置の概略的なハードウェア構成図である。

[図2]本発明の一実施形態による機械操作装置の概略的な機能を示すブロック図である。

[図3]本発明の変形例による機械操作装置の概略的な機能を示すブロック図である。

[図4]本発明の他の実施形態による機械操作装置の概略的な機能を示すブロック図である。

[図5]複数の産業機械を用いて作業する例を示す図である。

発明を実施するための形態

[0011] 以下、本発明の実施形態を図面と共に説明する。

図1は本発明の一実施形態による機械操作装置の要部を示す概略的なハードウェア構成図である。本実施形態による機械操作装置1は、工場などの製造現場に設置された産業機械2を制御する制御装置の上に実装することができる。また、機械操作装置1は、制御装置に併設されたパソコンや、制御装置と有線乃至無線のネットワークを介して接続されたフォグコンピュータ6、クラウドサーバ7などのコンピュータ上に実装することができる。以下では、本実施形態による機械操作装置1を、産業機械2を制御する制御装置上に実装した例に基づいて説明する。

[0012] 本実施形態による機械操作装置1が備えるCPU11は、機械操作装置1を全体的に制御するプロセッサである。CPU11は、バス22を介してROM12に格納されたシステム・プログラムを読み出し、該システム・プログラムに従って機械操作装置1全体を制御する。RAM13には一時的な計

算データや表示データ、及び外部から入力された各種データ等が一時的に格納される。

[0013] 不揮発性メモリ 14 は、例えば図示しないバッテリーでバックアップされたメモリやSSD (Solid State Drive) 等で構成され、機械操作装置 1 の電源がオフされても記憶状態が保持される。不揮発性メモリ 14 には、産業機械 2 から取得されたデータ、インタフェース 15 を介して外部機器 72 から読み込まれた制御用プログラムやデータ、入力装置 71 を介して入力された制御用プログラムやデータ、ネットワーク 5 を介して他の装置から取得された制御用プログラムやデータ等が記憶される。不揮発性メモリ 14 に記憶された制御用プログラムやデータは、実行時／利用時にはRAM 13 に展開されても良い。また、ROM 12 には、公知の解析プログラムなどの各種システム・プログラムがあらかじめ書き込まれている。

[0014] インタフェース 15 は、機械操作装置 1 のCPU 11 とUSB 装置等の外部機器 72 と接続するためのインタフェースである。外部機器 72 側からは、例えば産業機械 2 の制御に用いられる制御用プログラムや設定データ等が読み込まれる。また、機械操作装置 1 内で編集した制御用プログラムや設定データ等は、外部機器 72 を介して外部記憶手段に記憶させることができる。PLC 16 は、ラダープログラムを実行して産業機械 2 及び産業機械 2 の周辺装置（例えば、工具交換装置や、ロボット等のアクチュエータ、産業機械 2 に取付けられている温度センサや湿度センサ等の複数のセンサ 3）に I/O ユニット 19 を介して信号を出力し制御する。また、産業機械 2 の本体に配備された操作盤の各種スイッチや周辺装置等の信号を受け、必要な信号処理をした後、CPU 11 に渡す。

[0015] インタフェース 20 は、機械操作装置 1 のCPU と有線乃至無線のネットワーク 5 とを接続するためのインタフェースである。ネットワーク 5 には、工作機械や放電加工機などの他の産業機械 4 やフォグコンピュータ 6、クラウドサーバ 7 等が接続され、機械操作装置 1 との間で相互にデータのやり取りを行っている。

- [0016] 表示装置 70 には、メモリ上に読み込まれた各データ、プログラム等が実行された結果として得られたデータ等がインタフェース 17 を介して出力されて表示される。表示装置 70 は、機械の状態を示す LED インジケ이터や警告ランプを備えていてもよい。また、キーボードやポインティングデバイス等から構成される入力装置 71 は、オペレータによる操作に基づく指令、データ等をインタフェース 18 を介して CPU 11 に渡す。
- [0017] インタフェース 21 は、機械操作装置 1 の CPU 11 と音声センサ 73 とを接続するためのインタフェースである。音声センサ 73 は、例えばマイクなどの集音機器であってよい。音声センサ 73 は、例えば入力装置 71 や図示しない機械操作盤、ペンダント（携帯型機械操作盤）などに取り付けられていてよい。音声センサ 73 が検出した作業者の発する音声は、音声信号として CPU 11 に渡される。
- [0018] インタフェース 23 は、機械操作装置 1 の CPU 11 と撮像センサ 74 とを接続するためのインタフェースである。撮像センサ 74 は、例えばカメラなどであってよい。撮像センサ 74 は、例えば表示装置 70 や図示しない機械操作盤、ペンダント（携帯型機械操作盤）などに取り付けられていてよい。また、機械の周辺を撮像できるように、機械が設置された位置の上方に取り付けられていてもよい。撮像センサ 74 が検出した作業者の顔や体の動きは、映像信号として CPU 11 に渡される。
- [0019] 産業機械 2 が備える軸を制御するための軸制御回路 30 は CPU 11 からの軸の移動指令量を受けて、軸の指令をサーボアンプ 40 に出力する。サーボアンプ 40 はこの指令を受けて、工作機械が備える軸を移動させるサーボモータ 50 を駆動する。軸のサーボモータ 50 は位置・速度検出器を内蔵し、この位置・速度検出器からの位置・速度フィードバック信号を軸制御回路 30 にフィードバックし、位置・速度のフィードバック制御を行う。なお、図 1 のハードウェア構成図では軸制御回路 30、サーボアンプ 40、サーボモータ 50 は 1 つずつしか示されていないが、実際には制御対象となる産業機械 2 に備えられた軸の数だけ用意される。

[0020] 図2は、本発明の一実施形態による機械操作装置1が備える機能を概略的なブロック図として示したものである。本実施形態による機械操作装置1が備える各機能は、図1に示した機械操作装置1が備えるCPU11がシステム・プログラムを実行し、機械操作装置1の各部の動作を制御することにより実現される。

[0021] 本実施形態の機械操作装置1は、音声信号取得部100、音声認識部110、映像信号取得部120、映像解析部130、コマンド決定部140、コマンド処理部150、出力部160を備える。

[0022] 音声信号取得部100は、音声センサ73が検出した音声信号を取得する。そして、取得した音声信号から1回の発話として認識される音声信号を抽出する。音声センサ73が検出する音声信号は、主として作業者が発話した音声に基づくものである。音声信号取得部100は、その中から作業者の1回の発話に対応する音声信号を切り出すようにしてよい。これは、例えば音声信号が予め定めた所定のレベル $L_{v_{th}}$ 以下となる状態が予め定めた所定の期間 $T_{s_{th}}$ 以上続いたところを音声の区切りとして、音声の区切りで挟まれた所定の期間 $T_{n_{th}}$ 以上の音声信号を1回の発話に対応する音声信号として切り出すようにすればよい。また、音声の切り出しにその他の公知の音声信号の解析技術を用いるようにしてもよい。音声信号取得部100が切り出した音声信号は、音声認識部110に出力される。

[0023] 音声認識部110は、音声信号取得部100から入力された作業者の1回の発話に対応する音声信号に対して音声認識処理を行う。そして、認識した結果をコマンド決定部140へと出力する。音声認識処理としては、動的時間伸縮法や隠れマルコフモデル、ニューラルネットワークを用いた手法などの公知の音声認識アルゴリズムを用いればよい。音声認識部110が出力する音声信号の認識結果は、例えば当該音声信号において発話される言葉を表す文字列、及び認識結果の信頼度であってよい。この時、複数の認識結果、それぞれの認識結果の信頼度を併せて出力するようにしてもよい。また、ニューラルネットワークを用いる場合、音声信号の認識結果は、音声信号の発

話の特徴を示す特徴量であってもよい。

[0024] 映像信号取得部 120 は、撮像センサ 74 が検出した音声信号を取得する。そして、取得した映像信号から作業者の発話に対応する部分の映像信号を抽出する。撮像センサ 74 が検出する映像信号は、作業者の顔、特に口元を撮像したものであってよい。または、撮像センサ 74 が検出する映像信号は、作業者の両手を含む上半身を撮像したものであってよい。映像信号取得部 120 は、音声信号取得部 100 が作業者の 1 回の発話に対応する音声信号を切り出した際に、その切り出した音声信号と同じタイミングで取得される映像信号を、作業者の 1 回の発話に対応する部分の映像信号として切り出すようにしてよい。映像信号取得部 120 が切り出した映像信号は、映像解析部 130 に出力される。

[0025] 映像解析部 130 は、映像信号取得部 120 から入力された作業者の 1 回の発話に対応する映像信号を解析する。そして、解析した結果をコマンド決定部 140 へと出力する。映像信号が作業者の顔、特に口元を撮像したものである場合は、映像解析処理として AAM (Active Appearance Model)、LipNet などの公知のアルゴリズムを用いることができる。また、映像信号が作業者の両手を含む上半身を撮像したものである場合、隠れマルコフモデルや多層ニューラルネットワークを用いた公知のジェスチャ認識アルゴリズムを用いればよい。映像解析部 130 が出力する映像信号の解析結果は、例えば当該映像信号から認識される口の動きを解析して得られる産業機械 2 の操作を表す文字列、及び解析結果の信頼度であってよい。また、作業者のジェスチャに対応する産業機械 2 の操作を表す文字列、及び解析結果の信頼度であってよい。いずれの場合においても、複数の解析結果、それぞれの認識結果の信頼度を併せて出力するようにしてもよい。更に、映像信号の解析結果は、映像信号から認識される口の動きに係る特徴を示す特徴量や、ジェスチャの特徴を示す特徴量であってもよい。

[0026] コマンド決定部 140 は、音声認識部 110 から入力された音声信号の認識結果と、映像解析部 130 から入力された映像信号の解析結果とに基づい

て、作業者による産業機械２に対する操作コマンドを決定する。そして、決定したコマンドをコマンド処理部１５０、及び出力部１６０へと出力する。コマンド決定部１４０は、例えば音声信号の認識結果に含まれる信頼度と、映像信号の解析結果に含まれる信頼度とを比較して、信頼度が高い方の結果に基づいて操作コマンドを決定してもよい。この時、音声信号の認識結果に含まれる信頼度と、映像信号の解析結果に含まれる信頼度とに、予め定めた所定の重みを掛け合わせてから比較するようにしてもよい。また、複数の認識結果、及び解析結果が入力された場合、同一の操作コマンドを示す認識結果、及び解析結果のそれぞれの信頼度が予め定めた所定の閾値以上である操作コマンドの中で、認識結果、及び解析結果のいずれかの信頼度が最も高い操作コマンドに決定するようにしてもよい。

[0027] より厳密に判定をする場合には、コマンド決定部１４０は、多層ニューラルネットワークのモデルを用いて操作コマンドを決定するようにしてもよい。この場合、例えば音声信号において発話される言葉を表す文字列、認識結果の信頼度、作業者のジェスチャに対応する産業機械２の操作を表す文字列、及び解析結果の信頼度を入力とし、操作コマンドを出力とする教師データを用いて学習を行ったモデルを作成しておく。そして、コマンド決定部１４０は、音声認識部１１０から入力された認識結果、及び映像解析部１３０から入力された解析結果を学習済みのニューラルネットワークへと入力し、出力された操作コマンドを作業者による産業機械２に対する操作コマンドとして決定する。

[0028] 多層ニューラルネットを用いた他の例としては、例えば予め音声信号の特徴量及び口の動きやジェスチャの特徴量を入力とし、操作コマンドを出力とする教師データを用いて学習を行ったモデルを作成しておく。そして、コマンド決定部１４０は、音声認識部１１０から入力された認識結果、及び映像解析部１３０から入力された解析結果を学習済みのニューラルネットワークへと入力し、出力された操作コマンドを作業者による産業機械２に対する操作コマンドとして決定する。

[0029] コマンド処理部 150 は、コマンド決定部 140 から出力された操作コマンドに従って、該操作コマンドに対応する所定の機能を実行する。所定の機能は、制御装置が備える一般的な機能であってよい。例えば機械操作装置 1 の所定の画面を呼び出す機能や、所定のパラメータの設定機能、産業機械 2 に対する制御に係る機能などが例示される。また、PLC 16 を介して産業機械 2 の周辺装置に対して信号を送ることで、産業機械 2 の周りに配置される安全ドアの開閉や、室内点消灯、切削液の使用停止、切粉の清掃などの機能を実行することも考えられる。

[0030] 出力部 160 は、コマンド決定部 140 から出力された操作コマンドを表示装置 70 に対して表示出力する。出力部 160 は、操作コマンドを表示装置 70 の画面上で実行されている所定の機能の表示の邪魔にならない位置（例えば、画面の最下段の状態表示領域など）に表示するようにしてもよい。また、ダイアログなどの形で表示出力するようにしてもよい。更に、図示しないスピーカーから操作コマンドを復唱するように音声出力するようにしてもよい。出力部 160 は、単に決定した操作コマンドを出力するだけでなく、操作コマンドの認識処理開始、処理中、処理完了などといったように、処理のそれぞれの過程を出力するようにしてもよい。また、表示装置 70 にLED や警告ランプなどの点灯を併用するようにしてもよい。出力部 160 は、操作コマンドをネットワーク 5 を介して他の産業機械 4 やフォグコンピュータ 6、クラウドサーバ 7 などの上位のコンピュータに送信出力するようにしてもよい。また、不揮発性メモリ 14 などの上に予め設けられたログ記録領域に出力するようにしてもよい。

[0031] 上記構成を備えた機械操作装置 1 は、取得した音声信号及び映像信号に基づいて、作業による産業機械 2 に対する操作コマンドを決定する。音声信号の認識結果のみに基づいて操作コマンドを決定する場合と比較して、製造現場における騒音などが原因での認識精度の低下を映像信号により補うことができるため、操作コマンドの認識精度の向上が見込まれる。また、映像信号の解析結果のみに基づいて操作コマンドを決定する場合と比較すると、指

令の表現力の限界を音声信号で補うことが期待される。

[0032] 本実施形態による機械操作装置 1 の一変形例として、機械操作装置 1 は、作業者を認証する機能を備えていてもよい。図 3 は、本変形例による機械操作装置 1 が備える機能を概略的なブロック図として示したものである。本変形例による機械操作装置 1 は、更に作業者認証部 170 を備える。

[0033] 作業者認証部 170 は、音声認識部 110 が認識した作業者が発話した声を検出した音声信号や、映像解析部 130 が解析した作業者の顔を撮像した映像信号に基づいて、作業者を認証する。機械操作装置 1 の RAM 13 乃至不揮発性メモリ 14 上には、作業者の認証に係る情報として、作業者の発話する声の特徴を示す特徴量や、作業者の顔の特徴を示す特徴量に対して、作業者を一意に識別する識別情報を関連付けて記憶している認証情報記憶部 210 が予め用意されている。作業者認証部 170 は、作業者が発話した際に、作業者が発話した声や、作業者の顔を撮像した映像信号から、作業者の発話する声の特徴を示す特徴量や、作業者の顔の特徴を示す特徴量を解析する。そして、認証情報記憶部 210 に記憶されている複数の作業者の声乃至顔の特徴量と照合して、発話している作業者がいずれの作業者であるのかを認証する。作業者認証部 170 は、認証した作業者の識別情報をコマンド処理部 150 に出力する。作業者の発話した声、作業者の顔は、一方のみを認証に用いてもよいし、双方を認証に用いるようにしてよい。

[0034] 機械操作装置 1 の RAM 13 乃至不揮発性メモリ 14 上には、更に作業者の識別情報に対して、それぞれの作業者が実行可能な操作コマンドを関連付けた認可情報を記憶する認可情報記憶部 220 が用意されている。そして、コマンド処理部 150 は、コマンド処理部 150 から決定された操作コマンドが入力された際に、作業者認証部 170 から入力された作業者の識別情報に関連付けられている実行可能な操作コマンドを認可情報記憶部 220 から取得する。そして、取得した実行可能な操作コマンドの中に、決定した操作コマンドが含まれている場合は、通常通りその操作コマンドを処理する。一方で、含まれていない場合には、当該操作コマンドは作業者に認可されてい

ない旨を出力部１６０に出力する。

[0035] 本変形例による機械操作装置１は、作業者の職種や権限などにより実行できない操作コマンドが実行できないように制限することができる。作業者の認証は、作業者が発話した声や作業者の顔に基づいて行われるため、なりすましによる操作コマンドの実行を防止することが期待できる。例えば、他の作業者の音声を録音して悪用することを防止することができる。

[0036] 以上、本発明の実施形態について説明したが、本発明は上述した実施の形態の例のみに限定されることなく、適宜の変更を加えることにより様々な態様で実施することができる。

例えば、上記した変形例において認証情報記憶部２１０には、作業者の識別情報と関連付けて、更に作業者の用いる言語に係る情報を記憶するようにしてもよい。そして、予め音声認識部１１０が音声認識に用いる言語モデルを、作業者の用いる言語に応じて切り替えるようにしてもよい。このように構成することで、言語が異なる様々な作業者に対して、操作コマンドの認識精度を向上させることが可能となる。

[0037] 更に、作業者が撮像センサ７４の撮像領域に入った場合に、認証情報記憶部２１０は、自動的に作業者を認証するようにしてよい。そして、作業者が認証されていない場合に、コマンド決定部１４０は、操作コマンドの決定に係る処理を停止するようにしてもよい。このような構成により、作業者が産業機械２の操作を行う範囲内にいない限り、操作コマンドの入力を受け付けないようにすることができるため、他の音声を検出することによる産業機械２の誤動作を防止することができる。また、複数台の産業機械２を用いて１人の作業者が作業している場合、作業者がいずれの産業機械２の操作範囲内にいるかに応じて、操作の対象とする産業機械２を切り替えるようにすることもできる。

[0038] また、作業者の動作として、作業者の口の動きやジェスチャ以外に、例えば作業者の作業動作や立ち位置を検出して、これを操作コマンドの認識に用いるようにしてもよい。この場合、映像解析部１３０は、作業者がどのよう

な作業動作をしているのか（例えば、ワークを運んでいるのか、機械の中に腕を入れているのか否かなど）といったことや、作業者が産業機械 2 の近傍のいずれの位置に立っているのかなどを解析する。そして、その解析結果をコマンド決定部 140 へと出力する。コマンド決定部 140 は、作業者が行っている作業動作や、立ち位置に応じて、作業者が行おうとしている操作コマンドの絞り込みを行うことで、操作コマンドの決定をする。例えば作業者が機械の安全扉の中に手を入れている状態で、安全扉を閉じる操作コマンドや、運転再開の操作コマンドを入力することは危険な行為である。このような状況において、音声認識部 110 がそのような音声信号を認識した場合であっても、コマンド決定部 140 は、危険な操作コマンドを処理対象として決定しないようにすることができる。また、工場などの製造現場には、複数の作業者が協働して作業を進めている場合もある。このような場合、作業者同士の会話を音声コマンドとして誤検出してしまう可能性がある。そのような場合に、作業者を認証した上で、作業者の動作内容から他の作業者と会話をしていると解析されている場合には、操作コマンドの入力を受け付けないようにすることも可能となる。これにより、産業機械の誤操作を抑制することが期待される。

[0039] 更に、上記に加えて産業機械 2 の動作状態を取得し、その動作状態から作業者が行おうとしている操作コマンドの絞り込みを行うことも可能である。例えば、ワークの加工が行われている状態で安全扉を開く操作コマンドが入力されることは通常あり得ない。そのため、ワークの加工中に安全扉を開く操作コマンドが認識乃至解析された場合、コマンド決定部 140 は、そのような操作コマンドについては決定対象として除外するようにしてもよい。

[0040] 近傍に設置された複数台の産業機械を用いて並列して作業している場合、作業者の立ち位置やジェスチャにより操作対象とする産業機械 2 を決定するようにしてもよい。このように構成する場合、図 4 に示すように、それぞれの産業機械 2 を制御する制御装置とネットワークを介して接続されたコンピュータ上に機械操作装置 1 を実装する。作業者が発話した音声は、例えば機

械操作装置 1 と有線／無線のネットワーク 5 を介して接続される音声センサ 7 3 を用いて収集すればよい。また、作業者の映像については、製造現場の天井や作業者の近傍に設置され、機械操作装置 1 と有線／無線のネットワーク 5 を介して接続される撮像センサ 7 4 や、作業者の管理対象となる産業機械 4 に取り付けられた撮像センサ 4 7 4 などを用いて収集する。

[0041] そして、図 5 に例示されるように、複数台の産業機械 4 を用いて作業をするに際して、撮像センサ 7 4 や撮像センサ 4 7 4 で撮像された作業者の立ち位置や、作業者の向き（体の向きや、目線など）、作業者のジェスチャ（どの産業機械 4 を指しているか、など）に基づいて複数の産業機械 4 の中から操作対象の産業機械 4 を特定し、その産業機械 4 に対する操作コマンドを認識及び解析するように構成すればよい。この時、作業者が発話した音声の中から産業機械 4 を特定する用語（特定の産業機械 4 でのみ実行される操作コマンド、産業機械 4 の名称など）も併せて検出して利用するようにしてもよい。

符号の説明

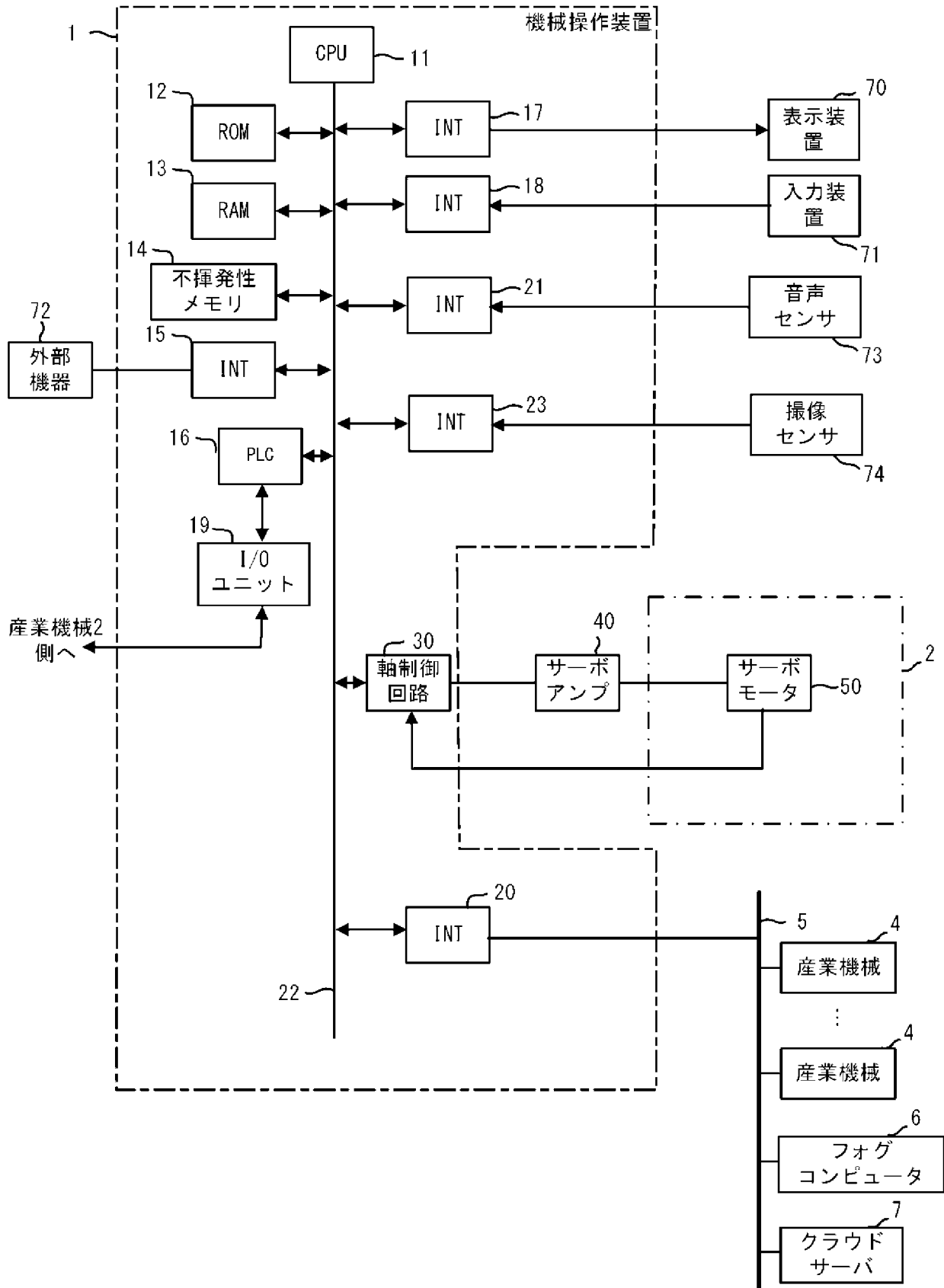
- [0042]
- 1 機械操作装置
 - 2 産業機械
 - 4 産業機械
 - 5 ネットワーク
 - 6 フォグコンピュータ
 - 7 クラウドサーバ
 - 1 1 C P U
 - 1 2 R O M
 - 1 3 R A M
 - 1 4 不揮発性メモリ
 - 1 5, 1 7, 1 8, 2 0, 2 1 インタフェース
 - 1 6 P L C
 - 1 9 I / O ユニット

- 2 2 バス
- 3 0 軸制御回路
- 4 0 サーボアンプ
- 5 0 サーボモータ
- 7 0 表示装置
- 7 1 入力装置
- 7 2 外部機器
- 7 3 音声センサ
- 7 4 撮像センサ
- 1 0 0 音声信号取得部
- 1 1 0 音声認識部
- 1 2 0 映像信号取得部
- 1 3 0 映像解析部
- 1 4 0 コマンド決定部
- 1 5 0 コマンド処理部
- 1 6 0 出力部
- 1 7 0 作業者認証部
- 2 1 0 認証情報記憶部
- 2 2 0 認可情報記憶部

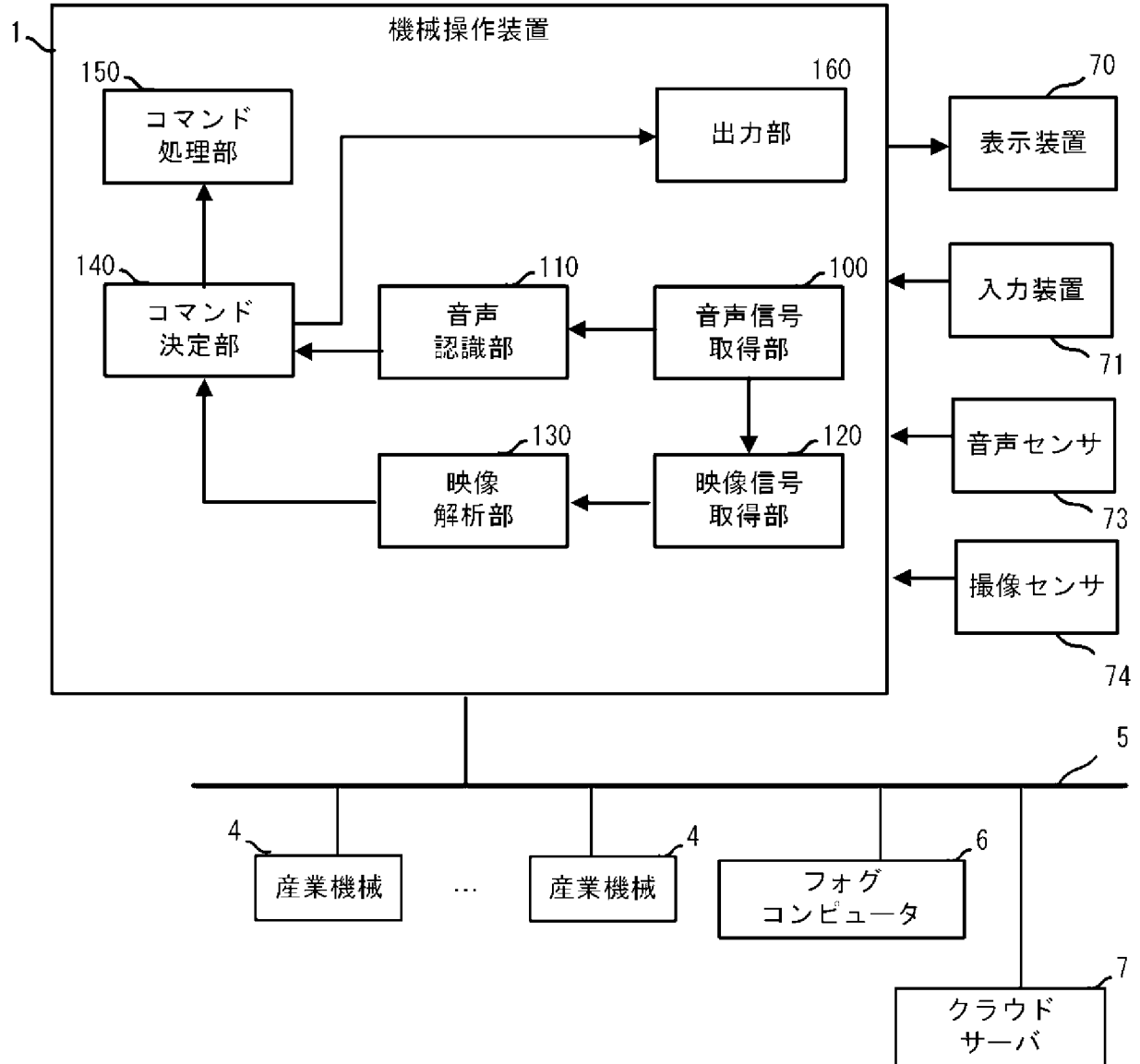
請求の範囲

- [請求項1] 製造現場において入力された音声信号を音声認識して産業機械を操作するための操作コマンドとして利用する機械操作装置であって、
- 音声センサにより検出された作業者が発話した音声を変換した音声信号に基づいて、所定の操作を示す第1のデータを認識する音声認識部と、
- 撮像センサにより作業者の動きを撮像して得られた映像信号に基づいて、所定の操作を示す第2のデータを解析する映像解析部と、
- 前記第1のデータ及び前記第2のデータに基づいて、前記産業機械を操作するための操作コマンドを決定するコマンド決定部と、
- 前記コマンド決定部が決定した操作コマンドを処理するコマンド処理部と、
- を備えた機械操作装置。
- [請求項2] 前記第1のデータは、作業者が発話した音声及び作業者の顔に基づいて該作業者を認証する作業者認証部を更に備え、
- 前記コマンド処理部は、前記作業者認証部により認証された作業者に認可されていない操作コマンドの実行を禁止する、
- 請求項1に記載の機械操作装置。
- [請求項3] 認証された作業者が用いる言語に応じて、前記音声認識部が用いる言語モデルを切り替える、
- 請求項2に記載の機械操作装置。
- [請求項4] 前記コマンド決定部は、前記映像解析部による作業者の検出が行われない場合に、前記操作コマンドの決定を停止する、
- 請求項1に記載の機械操作装置。
- [請求項5] 前記音声認識部による認識結果及び前記映像解析部による解析結果の少なくともいずれかに基づいて、複数の産業機械の中から操作対象となる産業機械を特定する、
- 請求項1に記載の機械操作装置。

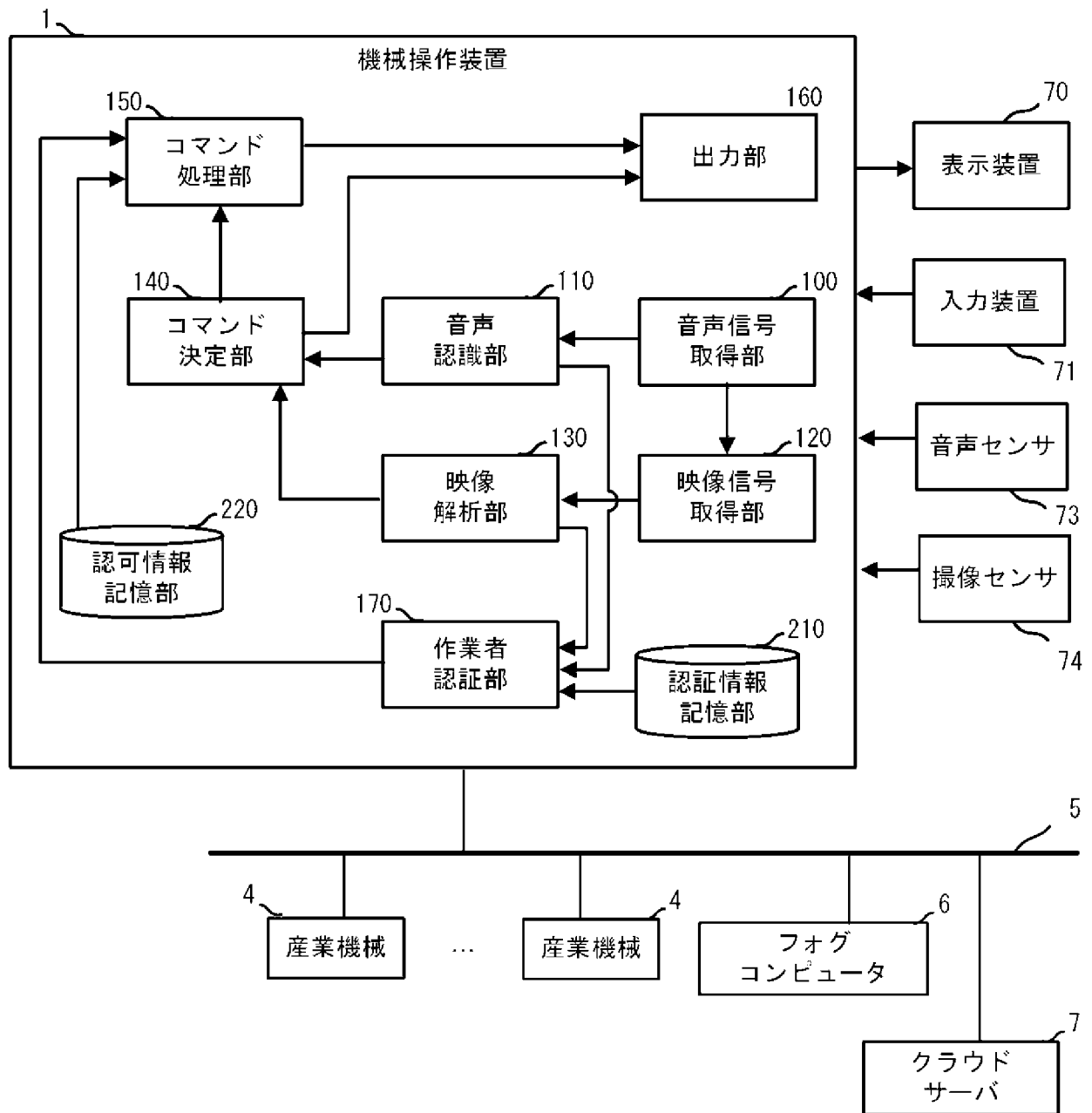
[図1]



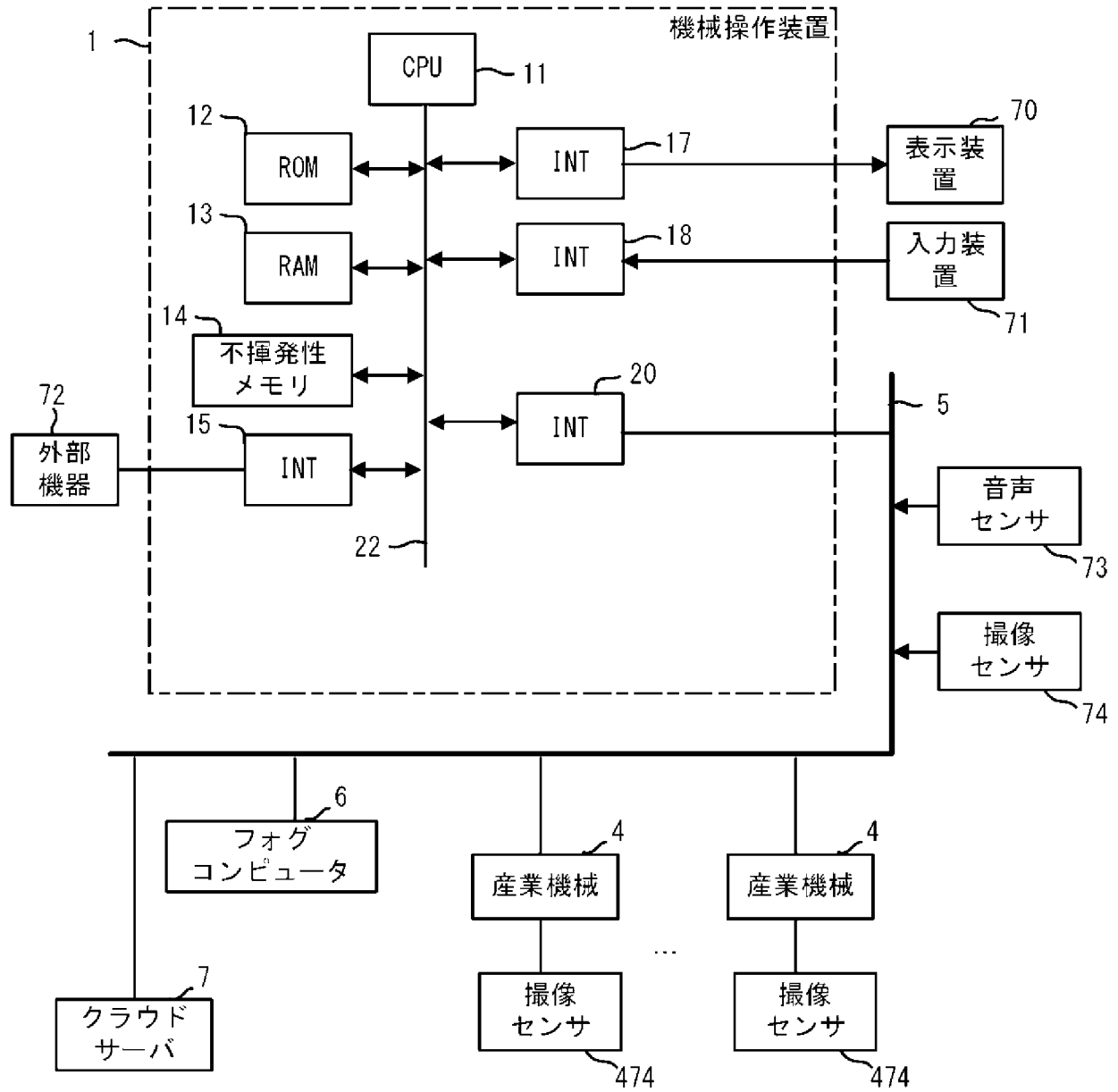
[図2]



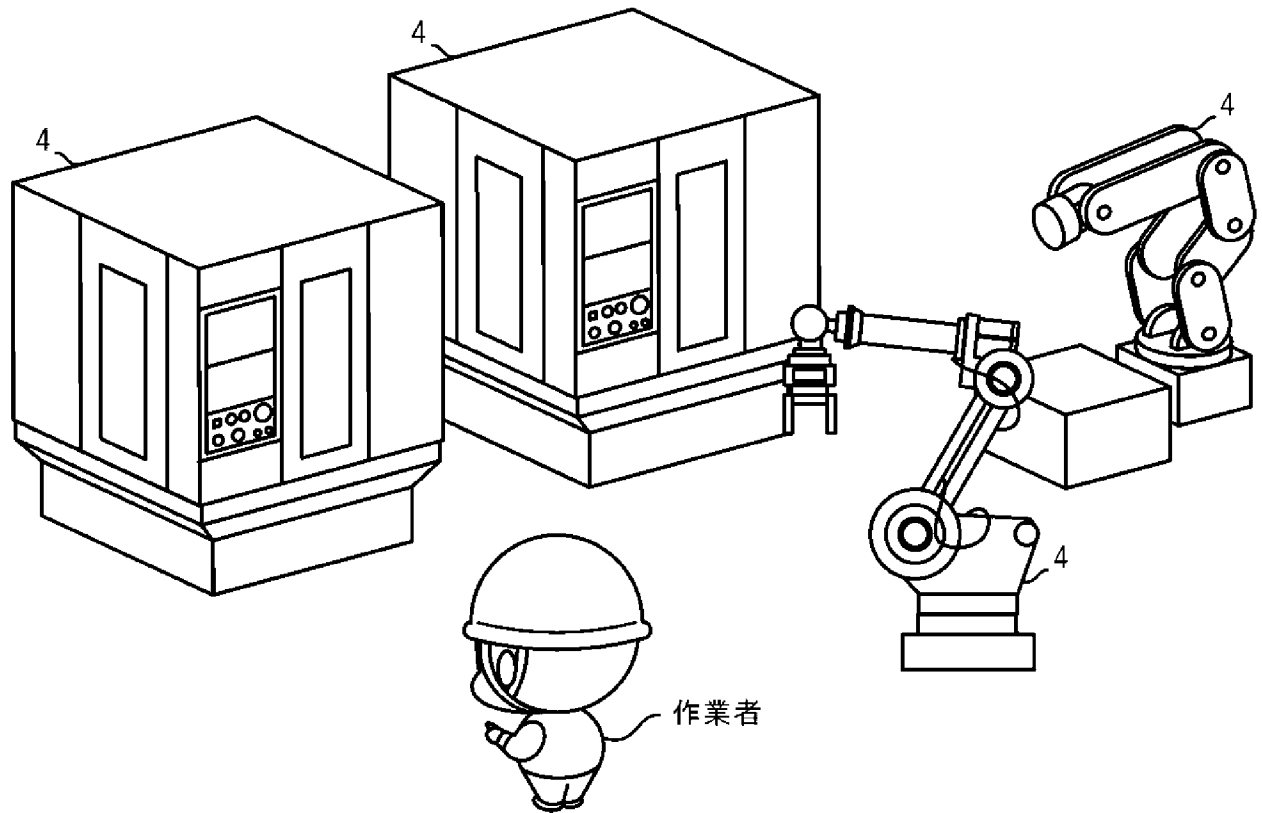
[図3]



[図4]



[図5]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/019792

A. CLASSIFICATION OF SUBJECT MATTER**G06F 3/16**(2006.01)i; **G06F 3/01**(2006.01)i

FI: G06F3/16 610; G06F3/01 570

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F3/01; G06F3/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2022

Registered utility model specifications of Japan 1996-2022

Published registered utility model applications of Japan 1994-2022

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2012-103840 A (SONY CORPORATION) 31 May 2012 (2012-05-31) paragraphs [0030]-[0169]	1-4
X	JP 2007-121576 A (MATSUSHITA ELECTRIC WORKS, LTD.) 17 May 2007 (2007-05-17) paragraphs [0017]-[0132]	1, 4-5
A	JP 2013-75005 A (GE MEDICAL SYSTEMS GLOBAL TECHNOLOGY CO LLC) 25 April 2013 (2013-04-25) entire text, all drawings	1-5



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

05 July 2022

Date of mailing of the international search report

19 July 2022

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)

3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915

Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2022/019792

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2012-103840	A	31 May 2012	US 2012/0112995 A1 paragraphs [0065]-[0204] EP 2450879 A1 CN 102591448 A RU 2011144585 A	
JP	2007-121576	A	17 May 2007	(Family: none)	
JP	2013-75005	A	25 April 2013	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 3/16(2006.01)i; G06F 3/01(2006.01)i FI: G06F3/16 610; G06F3/01 570														
B. 調査を行った分野														
調査を行った最小限資料（国際特許分類（IPC）） G06F3/01; G06F3/16														
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2022年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2022年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2022年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2022年	日本国実用新案登録公報	1996 - 2022年	日本国登録実用新案公報	1994 - 2022年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2022年													
日本国実用新案登録公報	1996 - 2022年													
日本国登録実用新案公報	1994 - 2022年													
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）														
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
X	JP 2012-103840 A（ソニー株式会社）31.05.2012（2012 - 05 - 31） 段落[0030]-[0169]	1-4												
X	JP 2007-121576 A（松下電工株式会社）17.05.2007（2007 - 05 - 17） 段落[0017]-[0132]	1, 4-5												
A	JP 2013-75005 A（ジーイー・メディカル・システムズ・グローバル・テクノロジー・カンパニー・エルエルシー）25.04.2013（2013 - 04 - 25） 全文、全図	1-5												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的技术水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技术水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技术水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 05.07.2022	国際調査報告の発送日 19.07.2022													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 田川 泰宏 5E 4236 電話番号 03-3581-1101 内線 3521													

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2022/019792

引用文献			公表日	パテントファミリー文献	公表日
JP	2012-103840	A	31.05.2012	US 2012/0112995 A1 段落[0065]-[0204] EP 2450879 A1 CN 102591448 A RU 2011144585 A	
JP	2007-121576	A	17.05.2007	(ファミリーなし)	
JP	2013-75005	A	25.04.2013	(ファミリーなし)	