

(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
2022年9月29日(29.09.2022)



(10) 国際公開番号

WO 2022/202341 A1

(51) 国際特許分類:
G10H 1/02 (2006.01)

(21) 国際出願番号: PCT/JP2022/010400

(22) 国際出願日: 2022年3月9日(09.03.2022)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2021-050384 2021年3月24日(24.03.2021) JP

(71) 出願人: ヤマハ株式会社 (YAMAHA CORPORATION) [JP/JP]; 〒4308650 静岡県浜松市中区中沢町10番1号 Shizuoka (JP).

(72) 発明者: 須見 康平 (SUMI, Kouhei); 〒4308650 静岡県浜松市中区中沢町10番1号 ヤマハ株式会社内 Shizuoka (JP). 浅野 貴裕 (ASANO, Takahiro); 〒4308650 静岡県浜松市中区中沢町10番1号 ヤマハ株式会社内 Shizuoka (JP). 大嵯 郁弥 (OSAKI, Ikumi); 〒4308650 静岡県浜松市中区中沢町10番1号 ヤマハ株式会社内 Shizuoka (JP).

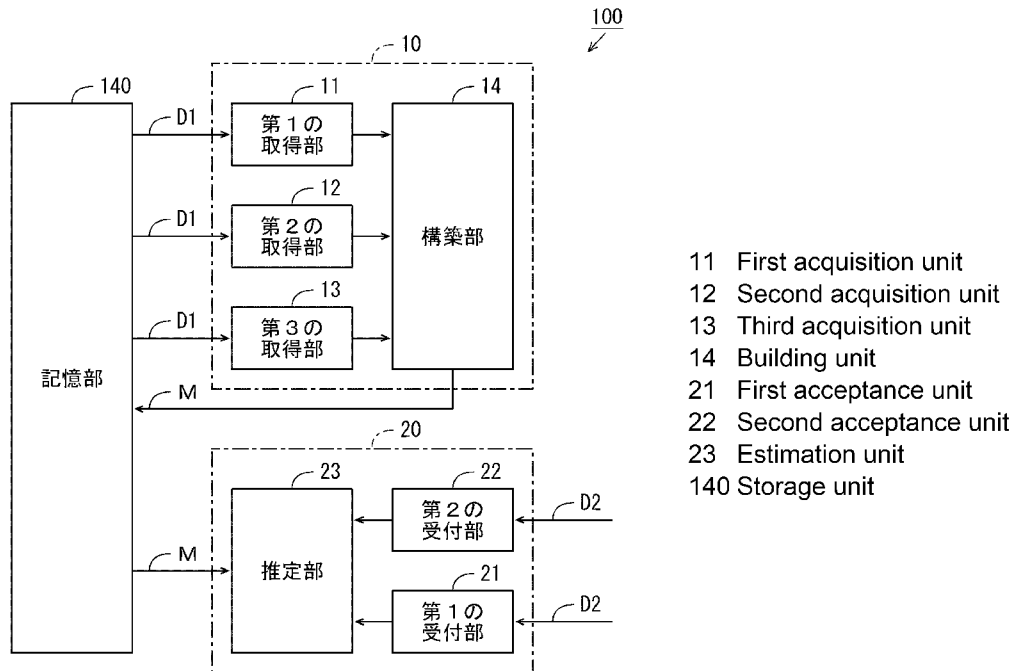
(74) 代理人: 中川 雅博, 外 (NAKAGAWA, Masahiro et al.); 〒5640052 大阪府吹田市広芝町4番1号江坂・ミタカビル3階 Osaka (JP).

(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,

(54) Title: SOUND EDITING DEVICE, SOUND EDITING METHOD, AND SOUND EDITING PROGRAM

(54) 発明の名称: 音編集装置、音編集方法および音編集プログラム

[図2]



(57) Abstract: A first acoustic signal is accepted by a first acceptance unit. A second acoustic signal is accepted by a second acceptance unit. Using a learned model, effect information that reflects an effect to be given to the first acoustic signal is estimated from the first acoustic signal and the second acoustic signal by an estimation unit. The learned model indicates an input-output relation between a first input acoustic signal and a second input acoustic signal, and output effect information that reflects an effect to be given to the first input acoustic signal.

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, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, 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, 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))

(57) 要約: 第1の音響信号が第1の受付部により受け付けられる。第2の音響信号が第2の受付部により受け付けられる。学習済モデルを用いて、第1の音響信号および第2の音響信号から第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報が推定部により推定される。学習済モデルは、第1の入力音響信号および第2の入力音響信号と、第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す。

明 細 書

発明の名称：音編集装置、音編集方法および音編集プログラム
技術分野

[0001] 本発明は、音を編集する音編集装置、音編集方法および音編集プログラムに関する。

背景技術

[0002] 合奏においては、複数の演奏者により同時に楽器が演奏される。したがって、各演奏者は、周囲の演奏者による楽器の音量とのバランスが維持されるように、自己の音量を調整することが好ましい。しかしながら、演奏者は、自己の出音が聴こえにくいため、自己の音量を大きくする傾向がある。この場合、他の演奏者も自己の音量を大きくすることとなるため、音量のバランスを維持することは容易ではない。特に、演奏会場が狭い場合には、音が飽和して会場内を回ることとなるので、音量のバランスを維持することはより困難になる。

[0003] 特許文献1：特開2020-160139号公報

発明の概要

発明が解決しようとする課題

[0004] 音の明瞭度が大きくなるように音響信号にエフェクトを付加することにより、演奏者は、楽器の音量を大きくすることなく自己の出音を認識可能になると考えられる。例えば、特許文献1には、様々なエフェクトを音響信号に付加する効果付加装置が記載されている。しかしながら、各演奏者による音の明瞭度は周囲の演奏者による音に応じて変化するため、音の明瞭度が大きくなるように音響信号にエフェクトを付加することは容易ではない。

[0005] 本発明の目的は、音の明瞭度を容易に大きくすることが可能な音編集装置、音編集方法および音編集プログラムを提供することである。

課題を解決するための手段

[0006] 本発明の一局面に従う音編集装置は、第1の音響信号を受け付ける第1の

受付部と、第2の音響信号を受け付ける第2の受付部と、第1の入力音響信号および第2の入力音響信号と、第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデルを用いて、第1の音響信号および第2の音響信号から第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する推定部とを備える。

[0007] 本発明の他の局面に従う音編集方法は、第1の音響信号を受け付け、第2の音響信号を受け付け、第1の入力音響信号および第2の入力音響信号と、第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデルを用いて、第1の音響信号および第2の音響信号から第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定し、コンピュータにより実行される。

[0008] 本発明のさらに他の局面に従う音編集プログラムは、コンピュータに音編集方法を実行させるプログラムであって、第1の音響信号を受け付ける処理と、第2の音響信号を受け付ける処理と、第1の入力音響信号および第2の入力音響信号と、第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデルを用いて、第1の音響信号および第2の音響信号から第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する処理とを、コンピュータにより実行させる。

発明の効果

[0009] 本発明によれば、音の明瞭度を容易に大きくすることが可能になる。

図面の簡単な説明

[0010] [図1]図1は本発明の第1の実施の形態に係る音編集装置を含む処理システムの構成を示すブロック図である。

[図2]図2は図1の音学習装置および音編集装置の構成を示すブロック図である。

[図3]図3は第1の音響信号および第3の音響信号の一例を示す図である。

[図4]図4は図2の音学習装置による音学習処理の一例を示すフローチャートである。

[図5]図5は図2の音編集装置による音編集処理の一例を示すフローチャートである。

[図6]図6は本発明の第2の実施の形態に係る音編集装置を含む処理システムの構成を示すブロック図である。

[図7]図7は図6の音学習装置および音編集装置の構成を示すブロック図である。

[図8]図8は図7の音学習装置による音学習処理の一例を示すフローチャートである。

[図9]図9は図7の音編集装置による音編集処理の一例を示すフローチャートである。

[図10]図10は他の実施の形態に係る音編集装置の構成を示すブロック図である。

発明を実施するための形態

[0011] [1] 第1の実施の形態

(1) 処理システムの構成

以下、本発明の実施の形態に係る音編集装置、音編集方法および音編集プログラムについて図面を用いて詳細に説明する。図1は、本発明の第1の実施の形態に係る音編集装置を含む処理システムの構成を示すブロック図である。図1に示すように、処理システム100は、RAM（ランダムアクセスメモリ）110、ROM（リードオンリメモリ）120、CPU（中央演算処理装置）130および記憶部140を備える。

[0012] 処理システム100は、例えばエフェクタまたはスピーカに設けられる。また、処理システム100は、パーソナルコンピュータ等の情報処理装置により実現されてもよいし、演奏機能を備えた電子楽器により実現されてもよい。RAM110、ROM120、CPU130および記憶部140は、バス150に接続される。RAM110、ROM120およびCPU130に

より音学習装置１０および音編集装置２０が構成される。本実施の形態では、音学習装置１０と音編集装置２０とは共通の処理システム１００により構成されるが、別個の処理システムにより構成されてもよい。

[0013] RAM１１０は、例えば揮発性メモリからなり、CPU１３０の作業領域として用いられるとともに、各種データを一時的に記憶する。ROM１２０は、例えば不揮発性メモリからなり、音学習プログラムおよび音編集プログラムを記憶する。CPU１３０は、ROM１２０に記憶された音学習プログラムをRAM１１０上で実行することにより音学習処理を行う。また、CPU１３０は、ROM１２０に記憶された音編集プログラムをRAM１１０上で実行することにより音編集処理を行う。音学習処理および音編集処理の詳細については後述する。

[0014] 音学習プログラムまたは音編集プログラムは、ROM１２０ではなく記憶部１４０に記憶されてもよい。あるいは、音学習プログラムまたは音編集プログラムは、コンピュータが読み取り可能な記憶媒体に記憶された形態で提供され、ROM１２０または記憶部１４０にインストールされてもよい。あるいは、処理システム１００がインターネット等のネットワークに接続されている場合には、当該ネットワーク上のサーバ（クラウドサーバを含む。）から配信された音学習プログラムまたは音編集プログラムがROM１２０または記憶部１４０にインストールされてもよい。

[0015] 記憶部１４０は、ハードディスク、光学ディスク、磁気ディスクまたはメモリカード等の記憶媒体を含み、学習済モデルMおよび複数の学習データD１を記憶する。学習済モデルMまたは複数の学習データD１は、記憶部１４０に記憶されず、コンピュータが読み取り可能な記憶媒体に記憶されていてもよい。あるいは、処理システム１００がネットワークに接続されている場合には、学習済モデルMまたは複数の学習データD１は、当該ネットワーク上のサーバに記憶されていてもよい。学習済モデルMは、複数の学習データD１に基づいて構築される。学習済モデルMの詳細については後述する。

[0016] 本実施の形態では、各学習データD１は、第１の入力音響信号、第２の入

力音響信号および出力音響信号をそれぞれ示す複数（マルチトラック）の波形データを含む。第１の入力音響信号は、第１の使用者が使用する楽器と同じ種類の楽器を用いて演奏された音等、第１の使用者により演奏されると仮定された音に対応する。第２の入力音響信号は、第２の使用者が使用する楽器と同じ種類の楽器を用いて演奏された音等、第２の使用者により演奏されると仮定された音に対応する。

[0017] 出力音響信号は、本実施の形態における出力エフェクト情報の例であり、第１の入力音響信号および第２の入力音響信号に基づいて、付与されるべきエフェクトが第１の入力音響信号に付与された音響信号である。第２の入力音響信号が同時に入力される状況において、出力音響信号に対応する音の明瞭度は、第１の入力音響信号に対応する音の明瞭度よりも大きい。出力音響信号を示す波形データは、エフェクトのパラメータが調整されることにより、第１の入力音響信号を示す波形データから生成されてもよい。

[0018] （２）音学習装置および音編集装置

図２は、図１の音学習装置１０および音編集装置２０の構成を示すブロック図である。図２に示すように、音学習装置１０は、機能部として、第１の取得部１１、第２の取得部１２、第３の取得部１３および構築部１４を含む。図１のＣＰＵ１３０が音学習プログラムを実行することにより、音学習装置１０の機能部が実現される。音学習装置１０の機能部の少なくとも一部は、電子回路等のハードウェアにより実現されてもよい。

[0019] 第１の取得部１１は、記憶部１４０等に記憶された各学習データＤ１から第１の入力音響信号を取得する。第２の取得部１２は、各学習データＤ１から第２の入力音響信号を取得する。第３の取得部１３は、各学習データＤ１から出力音響信号を取得する。

[0020] 構築部１４は、各学習データＤ１について、第１の取得部１１および第２の取得部１２によりそれぞれ取得された第１の入力音響信号および第２の入力音響信号に基づいて、第３の取得部１３により取得された出力音響信号を機械学習する。複数の学習データＤ１について機械学習を繰り返すことによ

り、構築部 14 は、第 1 の入力音響信号および第 2 の入力音響信号と、出力音響信号との間の入出力関係を示す学習済モデル M を構築する。

[0021] 本例では、構築部 14 は例えば U-Net を用いて機械学習するが、実施の形態はこれに限定されない。構築部 14 は、CNN (Convolutional Neural Network) または FCN (Fully Convolutional Network) 等の他の方式を用いて機械学習してもよい。構築部 14 により構築された学習済モデル M は、例えば記憶部 140 に記憶される。構築部 14 により構築された学習済モデル M は、ネットワーク上のサーバ等に記憶されてもよい。

[0022] 音編集装置 20 は、機能部として、第 1 の受付部 21、第 2 の受付部 22 および推定部 23 を含む。図 1 の CPU 130 が音編集プログラムを実行することにより、音編集装置 20 の機能部が実現される。音編集装置 20 の機能部の少なくとも一部は、電子回路等のハードウェアにより実現されてもよい。

[0023] 本実施の形態では、第 1 の受付部 21 および第 2 の受付部 22 は、楽曲データ D2 を取得する。楽曲データ D2 は、第 1 の音響信号および第 2 の音響信号をそれぞれ示す複数の波形データを含み、例えば使用者を含む複数の演奏者が合奏することにより生成される。第 1 の音響信号は、使用者により演奏された音に対応する。第 2 の音響信号は、他の演奏者により演奏された音または使用者の周囲で発生した音に対応する。第 1 の受付部 21 は、楽曲データ D2 から第 1 の音響信号を受け付ける。第 2 の受付部 22 は、楽曲データ D2 から第 2 の音響信号を受け付ける。

[0024] 推定部 23 は、記憶部 140 等に記憶された学習済モデル M を用いて、付与されるべきエフェクトが第 1 の音響信号に付与された第 3 の音響信号を楽曲データ D2 に含まれる第 1 の音響信号および第 2 の音響信号から推定する。また、推定部 23 は、推定された第 3 の音響信号を出力する。本実施の形態では、第 3 の音響信号がエフェクト情報の例である。

[0025] 図 3 は、第 1 の音響信号および第 3 の音響信号の一例を示す図である。図 3 の左欄には、楽曲データ D2 に含まれる第 1 の音響信号と、第 1 の音響信

号を周波数解析することにより得られたスペクトルとが示される。図3の右欄には、推定部23により出力された第3の音響信号と、第3の音響信号を周波数解析することにより得られたスペクトルとが示される。

[0026] 図3の例では、一点鎖線で囲まれたA部に示すように、比較的低い周波帯域においては、第3の音響信号の強度が第1の音響信号の強度よりも低減される。一方、二点鎖線で囲まれたB部に示すように、比較的高い周波帯域においては、第3の音響信号の強度が第1の音響信号の強度よりも増強される。これにより、第2の音響信号が同時に発生した状況において、第3の音響信号に対応する音の明瞭度は、第1の音響信号に対応する音の明瞭度よりも大きくなる。

[0027] したがって、使用者は、推定部23により出力された第3の音響信号を用いることにより、楽器の音量を大きくすることなく自己の出音を容易に認識することができる。そのため、使用者は、合奏において、周囲の演奏者による楽器の音量とのバランスが維持されるように、適した音量で自己の楽器を演奏することができる。あるいは、ミキシングエンジニアは、複数の楽器の音量のバランスが維持されるように容易にミキシングを行うことができる。

[0028] (3) 音学習処理および音編集処理

図4は、図2の音学習装置10による音学習処理の一例を示すフローチャートである。図4の音学習処理は、図1のCPU130が音学習プログラムを実行することにより行われる。

[0029] 第1の取得部11は、記憶部140等に記憶されたいずれかの学習データD1から第1の入力音響信号を取得する(ステップS1)。第2の取得部12は、ステップS1の学習データD1から第2の入力音響信号を取得する(ステップS2)。第3の取得部13は、ステップS1の学習データD1から出力音響信号を取得する(ステップS3)。ステップS1～S3は、いずれが先に実行されてもよいし、同時に実行されてもよい。

[0030] 次に、構築部14は、ステップS1で取得された第1の入力音響信号およびステップS2で取得された第2の入力音響信号と、ステップS3で取得さ

れた出力音響信号との間の入出力関係を機械学習する（ステップS 4）。続いて、構築部 1 4 は、所定回数の機械学習が実行されたか否かを判定する（ステップS 5）。所定回数の機械学習が実行されていない場合、構築部 1 4 はステップS 1 に戻る。

[0031] 所定回数の機械学習が実行されるまで、学習データ D 1 または学習パラメータが変更されつつステップS 1 ～S 5 が繰り返される。機械学習の繰り返し回数は、構築される学習済モデルの精度に応じて予め設定される。所定回数の機械学習が実行された場合、構築部 1 4 は、機械学習の結果に基づいて、第 1 の入力音響信号および第 2 の入力音響信号と、出力音響信号との間の入出力関係を示す学習済モデル M を構築し（ステップS 6）、音学習処理を終了する。

[0032] 図 5 は、図 2 の音編集装置 2 0 による音編集処理の一例を示すフローチャートである。図 5 の音編集処理は、図 1 の CPU 1 3 0 が音編集プログラムを実行することにより行われる。

[0033] 第 1 の受付部 2 1 は、楽曲データ D 2 から第 1 の音響信号を受け付ける（ステップS 1 1）。第 2 の受付部 2 2 は、ステップS 1 1 の楽曲データ D 2 から第 2 の音響信号を受け付ける（ステップS 1 2）。ステップS 1 1, S 1 2 は、いずれが先に実行されてもよいし、同時に実行されてもよい。推定部 2 3 は、音学習処理のステップS 6 で構築された学習済モデル M を用いて、ステップS 1 1, S 1 2 でそれぞれ受け付けられた第 1 の音響信号および第 2 の音響信号から第 3 の音響信号を推定し（ステップS 1 3）、音編集処理を終了する。

[0034] （４）実施の形態の効果

以上説明したように、本実施の形態に係る音編集装置 2 0 は、第 1 の音響信号を受け付ける第 1 の受付部 2 1 と、第 2 の音響信号を受け付ける第 2 の受付部 2 2 と、第 1 の入力音響信号および第 2 の入力音響信号と、第 1 の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデル M を用いて、第 1 の音響信号および第 2

の音響信号から第１の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する推定部２３とを備える。

[0035] この構成によれば、第２の音響信号が変化する場合でも、学習済モデルＭを用いて、音の明瞭度が大きくなるように第１の音響信号に付与されるべきエフェクトを反映するエフェクト情報を得ることが可能になる。これにより、音の明瞭度を容易に大きくすることができる。

[0036] エフェクト情報は、付与されるべきエフェクトが付与された第１の音響信号（第３の音響信号）を含んでもよい。この場合、推定された第３の音響信号を用いることにより、明瞭度が大きくされた音を容易に得ることができる。

[0037] 学習済モデルＭは、第１の入力音響信号および第２の入力音響信号に基づいて、付与されるべきエフェクトが付与された第１の入力音響信号（出力音響信号）が出力エフェクト情報として学習されることにより生成されてもよい。この場合、第１の音響信号および第２の音響信号から第３の音響信号を推定するための学習済モデルＭを容易に生成することができる。

[0038] [２] 第２の実施の形態

(１) 処理システムの構成

第２の実施の形態に係る音編集装置２０、音編集方法および音編集プログラムについて、第１の実施の形態に係る音編集装置２０、音編集方法および音編集プログラムと異なる点を説明する。図６は、本発明の第２の実施の形態に係る音編集装置２０を含む処理システム１００の構成を示すブロック図である。図６に示すように、処理システム１００は、エフェクト付与部１６０をさらに備える。エフェクト付与部１６０は、例えばイコライザまたはコンプレッサを含み、バス１５０に接続される。エフェクト付与部１６０は、入力されたパラメータに基づいて、音響信号にエフェクトを付与する。

[0039] 本実施の形態では、記憶部１４０等に記憶される各学習データＤ１は、第１の入力音響信号および第２の入力音響信号をそれぞれ示す複数の波形データを含む。また、各学習データＤ１は、出力音響信号を示す波形データに代

えて、出力音響信号を生成するために第1の入力音響信号に付与されるべきエフェクトを反映するパラメータ（以下、出力パラメータと呼ぶ。）を含む。出力パラメータは、本実施の形態における出力エフェクト情報の例である。

[0040] （2）音学習装置および音編集装置

図7は、図6の音学習装置10および音編集装置20の構成を示すブロック図である。本実施の形態においては、音学習装置10の第3の取得部13は、各学習データD1から出力パラメータを取得する。第1の取得部11および第2の取得部12の動作は、第1の実施の形態における第1の取得部11および第2の取得部12の動作とそれぞれ同様である。

[0041] 構築部14は、各学習データD1について、第1の取得部11および第2の取得部12によりそれぞれ取得された第1の入力音響信号および第2の入力音響信号に基づいて、第3の取得部13により取得された出力パラメータを機械学習する。複数の学習データD1について機械学習を繰り返すことにより、構築部14は、第1の入力音響信号および第2の入力音響信号と、出力パラメータとの間の入出力関係を示す学習済モデルMを構築する。

[0042] 本例では、構築部14は例えばCNNを用いて機械学習するが、実施の形態はこれに限定されない。構築部14は、RNN（Recurrent Neural Network）またはAttention等の他の方式を用いて機械学習してもよい。構築部14により構築された学習済モデルMは、例えば記憶部140に記憶される。構築部14により構築された学習済モデルMは、ネットワーク上のサーバ等に記憶されてもよい。

[0043] 音編集装置20において、第1の受付部21および第2の受付部22は、合奏により生成された第1の音響信号および第2の音響信号をそれぞれリアルタイムで取得する。推定部23は、記憶部140等に記憶された学習済モデルMを用いて、付与されるべきエフェクトが付与された第1の音響信号を生成するためのパラメータを第1の音響信号および第2の音響信号から順次推定する。また、推定部23は、推定されたパラメータを順次出力する。本

実施の形態では、パラメータがエフェクト情報の例である。

[0044] エフェクト付与部 160 は、推定部 23 により出力されたパラメータに基づいて、第 1 の受付部 21 により取得された第 1 の音響信号にエフェクトを付与する。これにより、図 3 の右欄に示された第 3 の音響信号と同様の第 4 の音響信号が生成される。したがって、第 2 の音響信号が同時に発生した状況において、第 4 の音響信号に対応する音の明瞭度は、第 1 の音響信号に対応する音の明瞭度よりも大きくなる。

[0045] (3) 音学習処理および音編集処理

図 8 は、図 7 の音学習装置 10 による音学習処理の一例を示すフローチャートである。図 8 の例では、音学習処理は、ステップ S 21 ~ S 26 を含む。ステップ S 21, S 22 は、図 4 の音学習処理のステップ S 1, S 2 とそれぞれ同様である。ステップ S 23 において、第 3 の取得部 13 は、学習データ D 1 から出力パラメータを取得する（ステップ S 23）。ステップ S 21 ~ S 23 は、いずれが先に実行されてもよいし、同時に実行されてもよい。

[0046] 構築部 14 は、ステップ S 21 で取得された第 1 の入力音響信号およびステップ S 22 で取得された第 2 の入力音響信号と、ステップ S 23 で取得された出力パラメータとの間の入出力関係を機械学習する（ステップ S 24）。ステップ S 25, S 26 は、図 4 の音学習処理のステップ S 5, S 6 とそれぞれ同様である。これにより、ステップ S 26 において、第 1 の入力音響信号および第 2 の入力音響信号と、出力パラメータとの間の入出力関係を示す学習済モデル M が構築される。

[0047] 図 9 は、図 7 の音編集装置 20 による音編集処理の一例を示すフローチャートである。第 1 の受付部 21 は、合奏により生成された第 1 の音響信号を受け付ける（ステップ S 31）。第 2 の受付部 22 は、合奏により生成された第 2 の音響信号を受け付ける（ステップ S 32）。ステップ S 31, S 32 は、略同時に実行される。

[0048] 推定部 23 は、音学習処理のステップ S 26 で構築された学習済モデル M

を用いて、ステップS 3 1, S 3 2でそれぞれ受け付けられた第1の音響信号および第2の音響信号からパラメータを推定する(ステップS 3 3)。その後、推定部2 3は、ステップS 3 3で推定されたパラメータを図7のエフェクト付与部1 6 0に出力し(ステップS 3 4)、ステップS 3 1に戻る。合奏が終了するまでステップS 3 1~S 3 4が繰り返される。

[0049] (4) 実施の形態の効果

本実施の形態においては、第1の実施の形態と同様に、第2の音響信号が変化する場合でも、学習済モデルMを用いて、音の明瞭度が大きくなるように第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を得ることが可能になる。これにより、音の明瞭度を容易に大きくすることができる。

[0050] エフェクト情報は、付与されるべきエフェクトが付与された第1の音響信号を生成するためのパラメータを含んでもよい。この場合、エフェクト情報を高速に得ることができる。また、エフェクト情報に基づいて第1の音響信号にパラメータが付与された第4の音響信号を用いることにより、明瞭度が大きくされた音を容易に得ることができる。

[0051] 学習済モデルMは、第1の入力音響信号および第2の入力音響信号に基づいて、付与されるべきエフェクトが付与された第1の入力音響信号を生成するための出力パラメータが出力エフェクト情報として学習されることにより生成されてもよい。この場合、第1の音響信号および第2の音響信号からパラメータを推定するための学習済モデルMを容易に生成することができる。

[0052] [3] 他の実施の形態

(1) 第1の実施の形態において、第1の入力音響信号および第2の入力音響信号と、出力音響信号との間の入出力関係を示す学習済モデルMが音学習装置1 0により構築されるが、実施の形態はこれに限定されない。第2の実施の形態と同様に、第1の入力音響信号および第2の入力音響信号と、出力パラメータとの間の入出力関係を示す学習済モデルMが音学習装置1 0により構築されてもよい。

- [0053] この場合、構築された学習済モデルMを用いて、付与されるべきエフェクトが付与された第1の音響信号を生成するためのパラメータが第1の音響信号および第2の音響信号から音編集装置20により推定される。この構成においては、音学習装置10または音編集装置20を実現するためのCPU130の処理速度が比較的低くてもよい。また、処理システム100は、エフェクト付与部160を含んでもよい。音編集装置20により推定されたパラメータがエフェクト付与部160に出力されることにより、第4の音響信号が生成される。
- [0054] (2) 第2の実施の形態において、第1の入力音響信号および第2の入力音響信号と、出力パラメータとの間の入出力関係を示す学習済モデルMが音学習装置10により構築されるが、実施の形態はこれに限定されない。第1の実施の形態と同様に、第1の入力音響信号および第2の入力音響信号と、出力音響信号との間の入出力関係を示す学習済モデルMが音学習装置10により構築されてもよい。
- [0055] この場合、構築された学習済モデルMを用いて、付与されるべきエフェクトが第1の音響信号に付与された第3の音響信号が第1の音響信号および第2の音響信号から音編集装置20により推定される。したがって、処理システム100はエフェクト付与部160を含まなくてもよい。この構成においては、音学習装置10または音編集装置20を実現するためのCPU130の処理速度が比較的高いことが好ましい。
- [0056] (3) 上記実施の形態において、学習済モデルMを用いて第1の音響信号および第2の音響信号からエフェクト情報が推定されるが、実施の形態はこれに限定されない。第1の音響信号および第2の音響信号とエフェクト情報との対応関係を示すテーブル等の対応情報が記憶部140等に記憶されている場合には、当該対応情報を用いて第1の音響信号および第2の音響信号からエフェクト情報が推定されてもよい。
- [0057] (4) 図10は、他の実施の形態に係る音編集装置20の構成を示すブロック図である。図10に示すように、他の実施に係る音編集装置20は、機

能部として調整部 24 をさらに含む。調整部 24 は、例えば図示しない表示装置に表示される G U I (Graphical User Interface) であり、使用者により操作される。調整部 24 は、G U I ではなく、物理的なダイヤル、スイッチまたはボタンであってもよい。

[0058] 使用者は、音楽性を低下させてでも音の明瞭度を増加させたい場合等には、エフェクトの程度が大きくなるように調整部 24 を操作することができる。一方、使用者は、リラックスしたい場合等には、エフェクトの程度が小さくなるように調整部 24 を操作することができる。調整部 24 は、使用者による操作に基づいて、第 1 の音響信号に付与されるべきエフェクトの度を調整する。推定部 23 は、学習済モデル M に基づいて、調整部 24 により調整された程度で第 1 の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する。

[0059] この構成においては、学習データ D 1 は、エフェクトの程度に対応して複数用意される。また、音学習装置 10 の構築部 14 は、第 1 の入力音響信号に付与されるべきエフェクトの程度に対応して、学習済モデル M を複数生成する。

請求の範囲

- [請求項1] 第1の音響信号を受け付ける第1の受付部と、
第2の音響信号を受け付ける第2の受付部と、
第1の入力音響信号および第2の入力音響信号と、前記第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデルを用いて、前記第1の音響信号および前記第2の音響信号から前記第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する推定部とを備える、音編集装置。
- [請求項2] 前記エフェクト情報は、付与されるべきエフェクトが付与された前記第1の音響信号を生成するためのパラメータを含む、請求項1記載の音編集装置。
- [請求項3] 前記学習済モデルは、前記第1の入力音響信号および前記第2の入力音響信号に基づいて、付与されるべきエフェクトが付与された前記第1の入力音響信号を生成するための出力パラメータが前記出力エフェクト情報として学習されることにより生成される、請求項2記載の音編集装置。
- [請求項4] 前記エフェクト情報は、付与されるべきエフェクトが付与された前記第1の音響信号を含む、請求項1記載の音編集装置。
- [請求項5] 前記学習済モデルは、前記第1の入力音響信号および前記第2の入力音響信号に基づいて、付与されるべきエフェクトが付与された前記第1の入力音響信号が前記出力エフェクト情報として学習されることにより生成される、請求項4記載の音編集装置。
- [請求項6] 前記第1の音響信号に付与されるべきエフェクトの程度を調整する調整部をさらに備え、
前記推定部は、前記学習済モデルに基づいて、前記調整部により調整された程度で前記第1の音響信号に付与されるべきエフェクトを反映する前記エフェクト情報を推定する、請求項1～5のいずれか一項

に記載の音編集装置。

[請求項7] 前記学習済モデルは、前記第1の入力音響信号に付与されるべきエフェクトの程度に対応して複数生成される、請求項6記載の音編集装置。

[請求項8] 第1の音響信号を受け付け、
第2の音響信号を受け付け、
第1の入力音響信号および第2の入力音響信号と、前記第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデルを用いて、前記第1の音響信号および前記第2の音響信号から前記第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する、
コンピュータにより実行される、音編集方法。

[請求項9] 前記エフェクト情報は、付与されるべきエフェクトが付与された前記第1の音響信号を生成するためのパラメータを含む、請求項8記載の音編集方法。

[請求項10] 前記学習済モデルは、前記第1の入力音響信号および前記第2の入力音響信号に基づいて、付与されるべきエフェクトが付与された前記第1の入力音響信号を生成するための出力パラメータが前記出力エフェクト情報として学習されることにより生成される、請求項9記載の音編集方法。

[請求項11] 前記エフェクト情報は、付与されるべきエフェクトが付与された前記第1の音響信号を含む、請求項8記載の音編集方法。

[請求項12] 前記学習済モデルは、前記第1の入力音響信号および前記第2の入力音響信号に基づいて、付与されるべきエフェクトが付与された前記第1の入力音響信号が前記出力エフェクト情報として学習されることにより生成される、請求項11記載の音編集方法。

[請求項13] 前記コンピュータは、
前記第1の音響信号に付与されるべきエフェクトの程度をさらに調

整し、

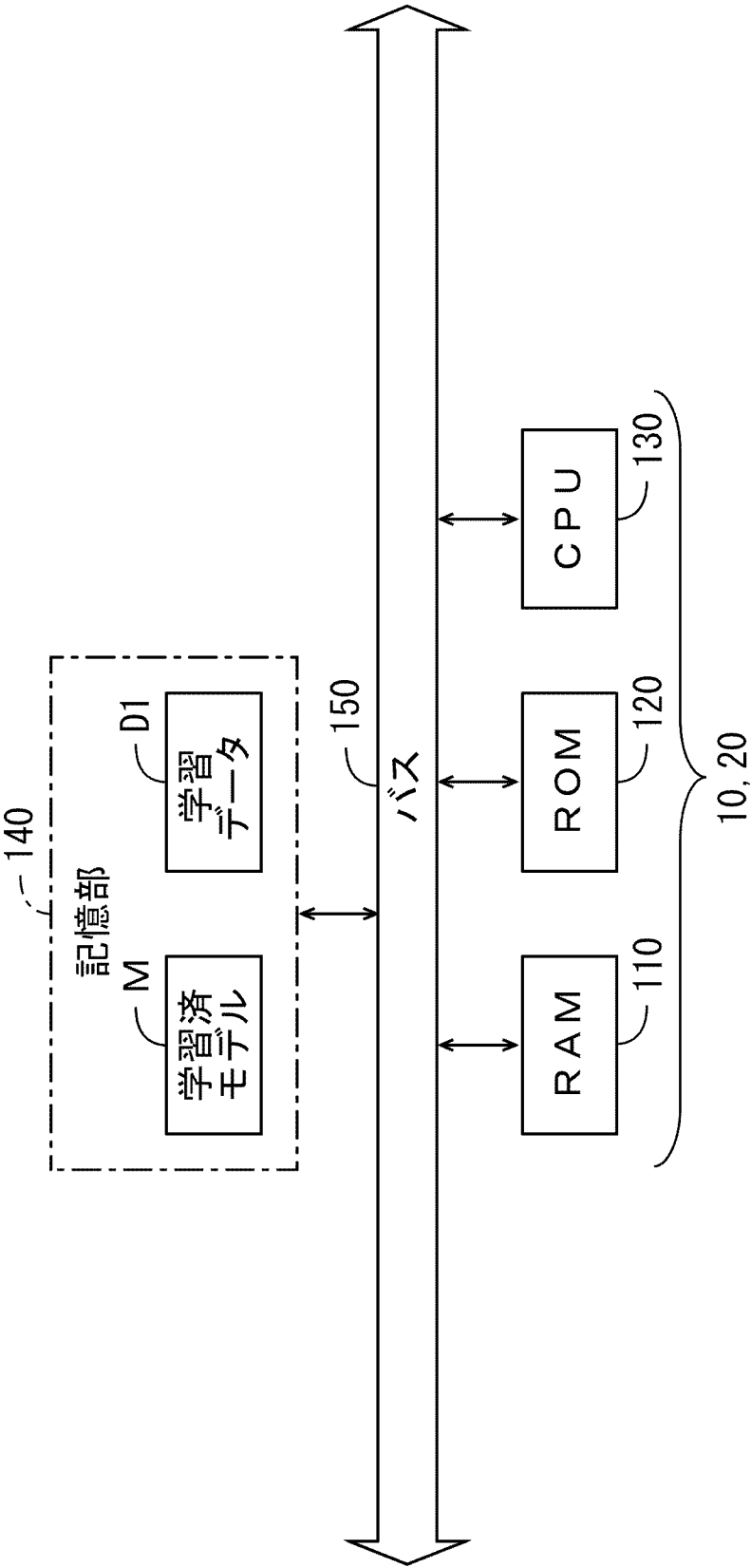
前記エフェクト情報の推定において、前記学習済モデルに基づいて、調整された程度で前記第1の音響信号に付与されるべきエフェクトを反映する前記エフェクト情報を推定する、請求項8～12のいずれか一項に記載の音編集方法。

[請求項14] 前記学習済モデルは、前記第1の入力音響信号に付与されるべきエフェクトの程度に対応して複数生成される、請求項13記載の音編集方法。

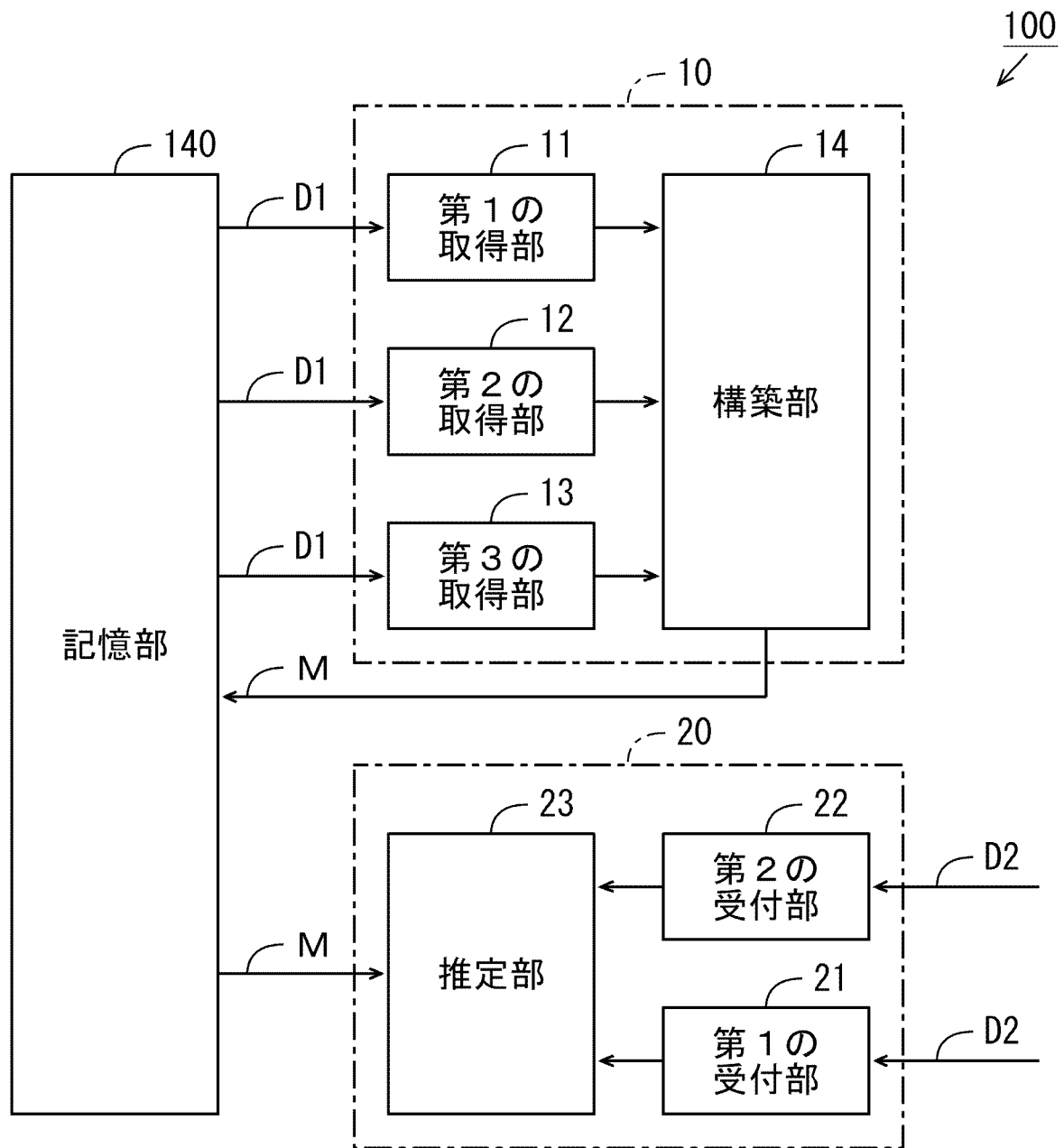
[請求項15] コンピュータに音編集方法を実行させるプログラムであって、
第1の音響信号を受け付ける処理と、
第2の音響信号を受け付ける処理と、
第1の入力音響信号および第2の入力音響信号と、前記第1の入力音響信号に付与されるべきエフェクトを反映する出力エフェクト情報との間の入出力関係を示す学習済モデルを用いて、前記第1の音響信号および前記第2の音響信号から前記第1の音響信号に付与されるべきエフェクトを反映するエフェクト情報を推定する処理とを、
前記コンピュータにより実行させる、音編集プログラム。

[図1]

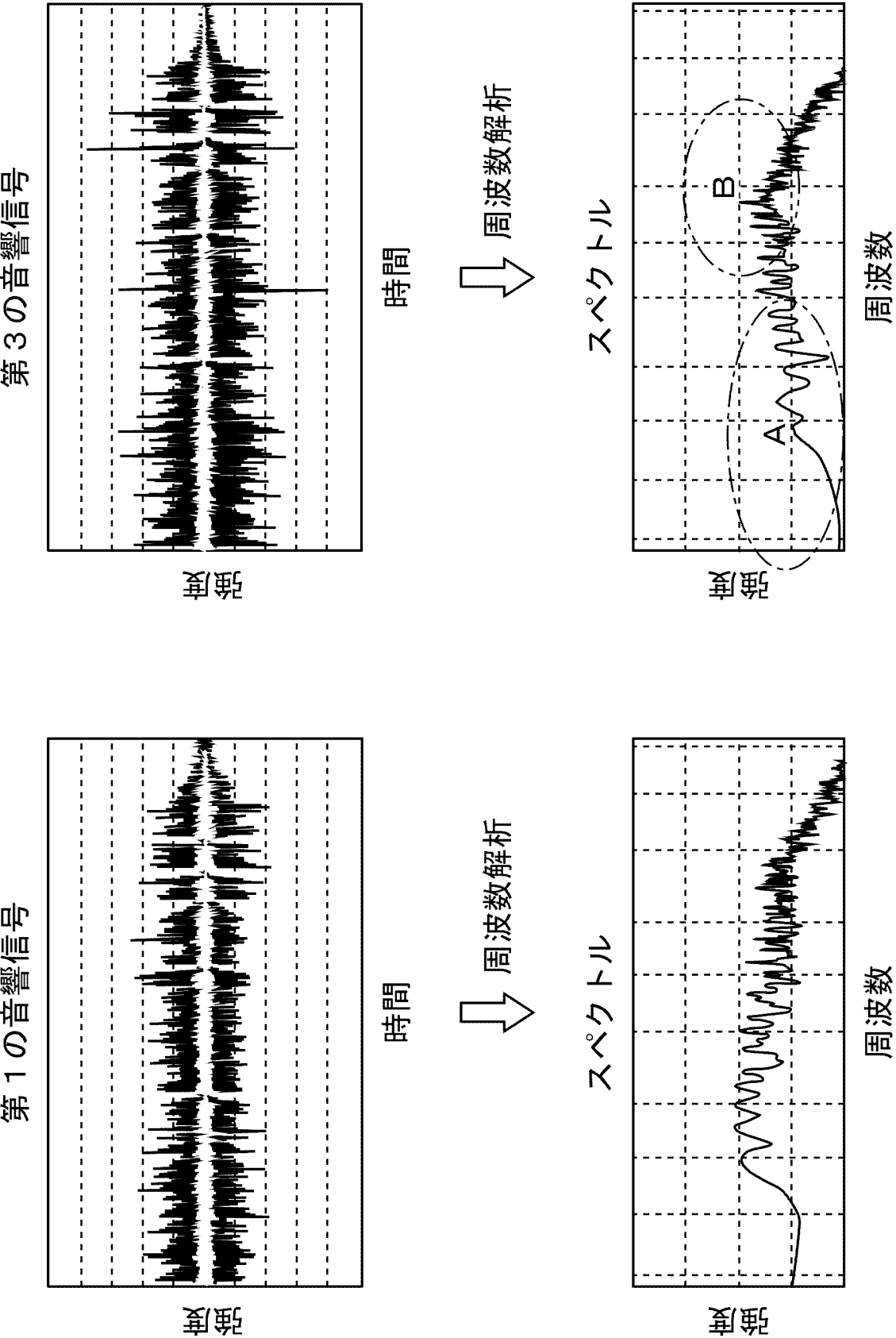
100 ↙



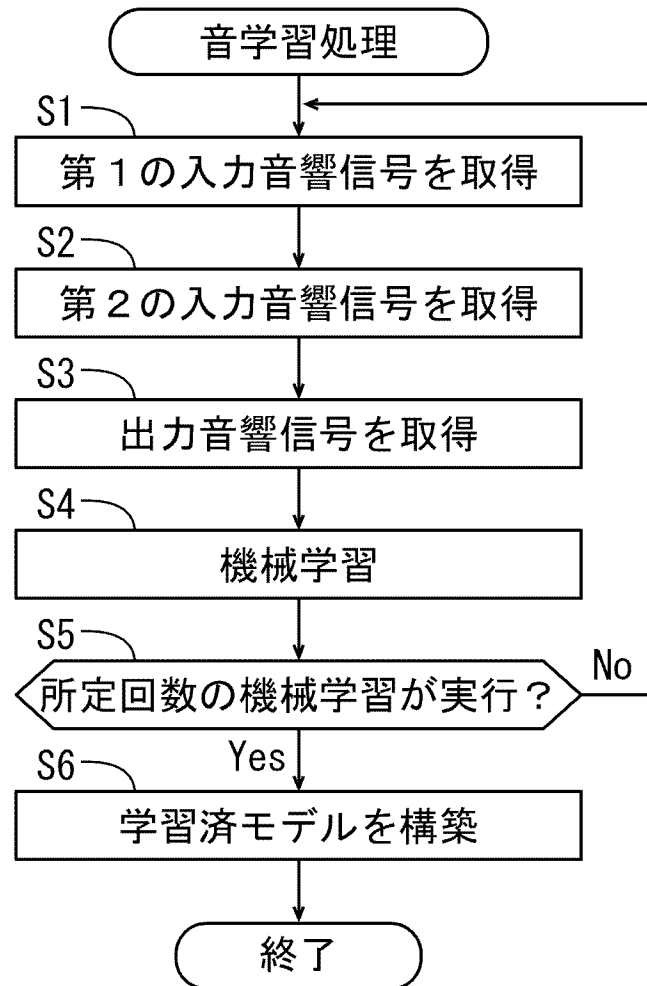
[図2]



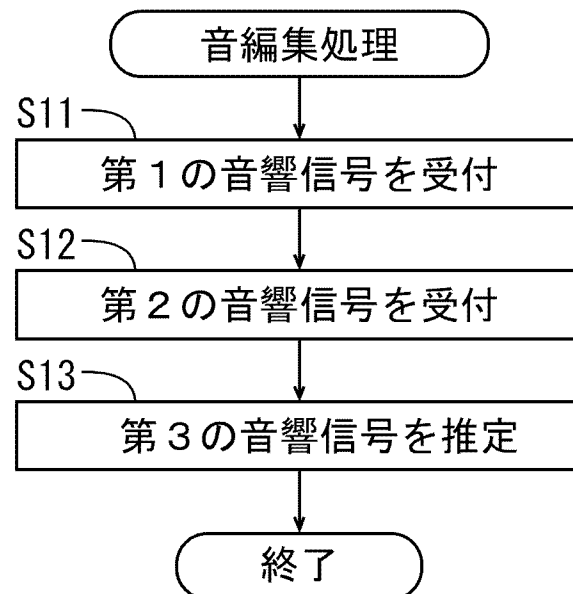
[図3]



[図4]

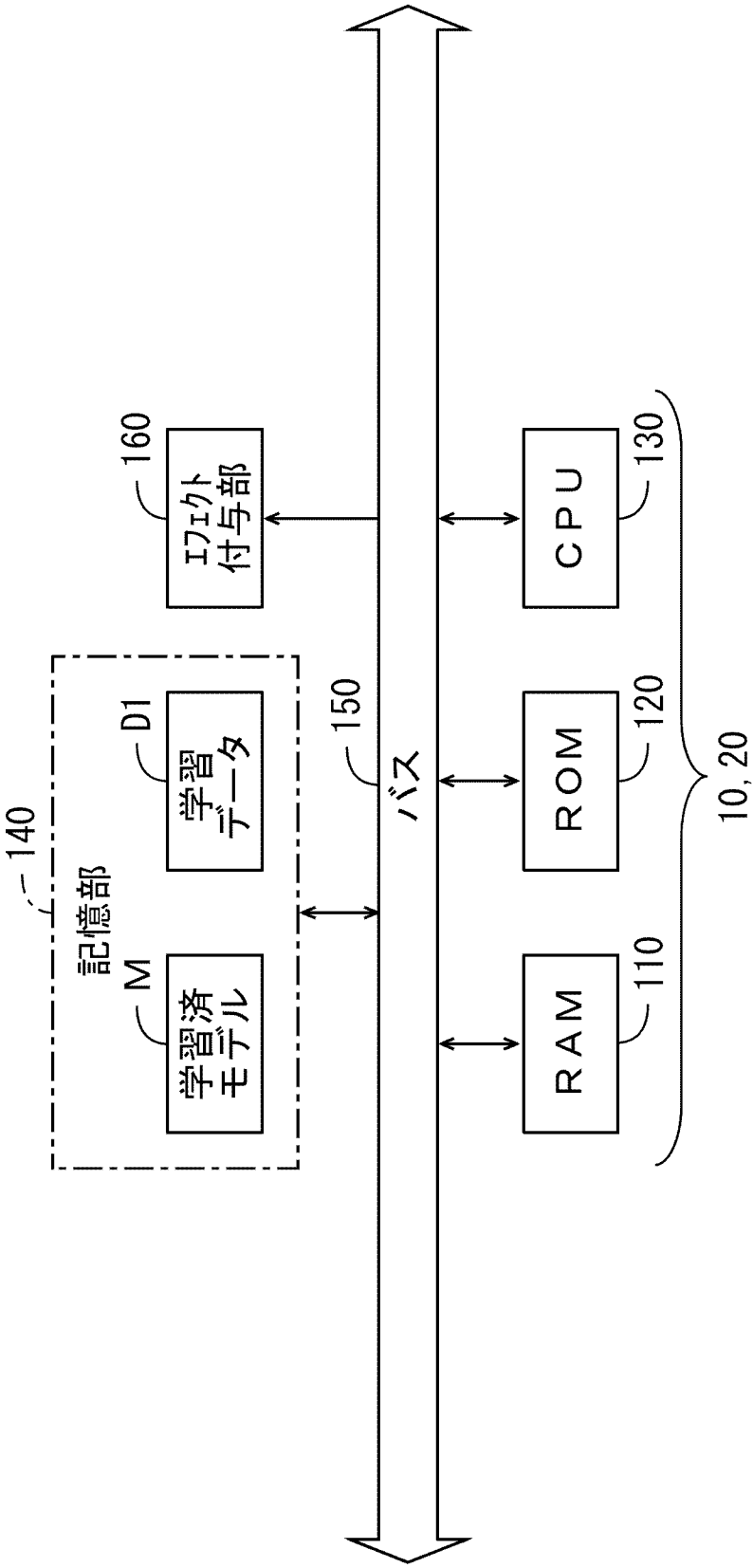


[図5]

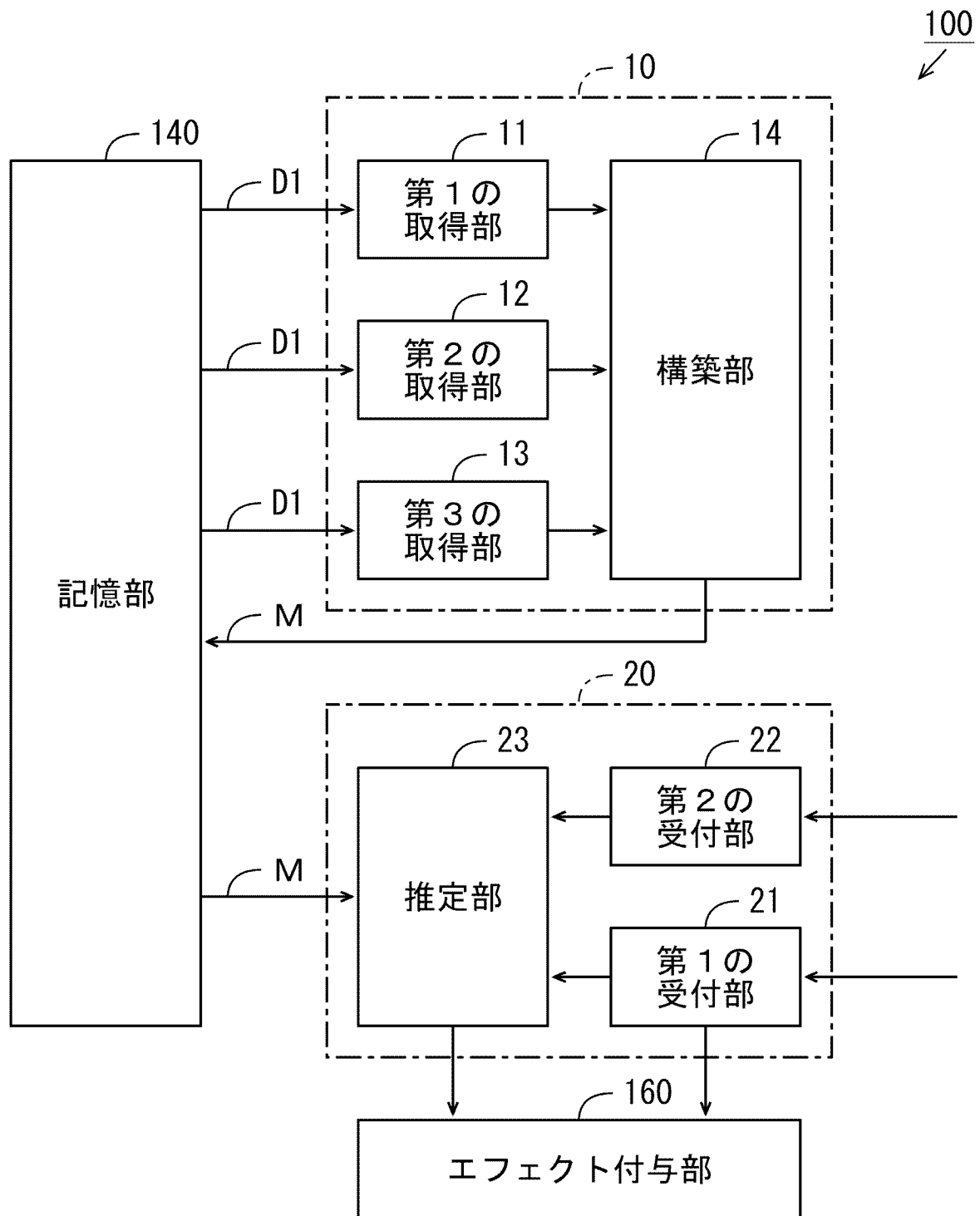


[図6]

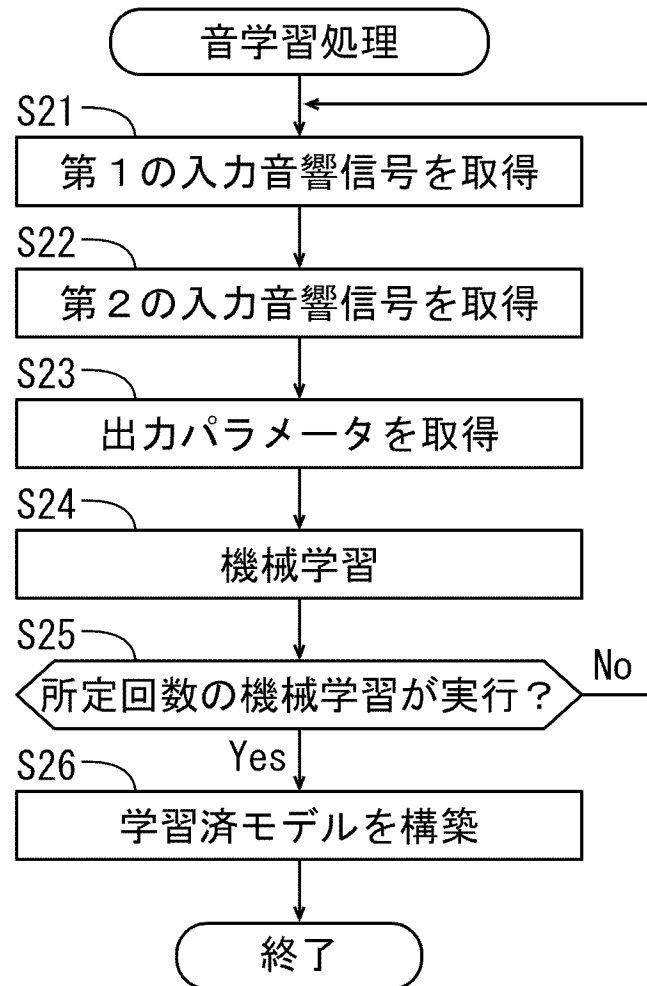
100 ↙



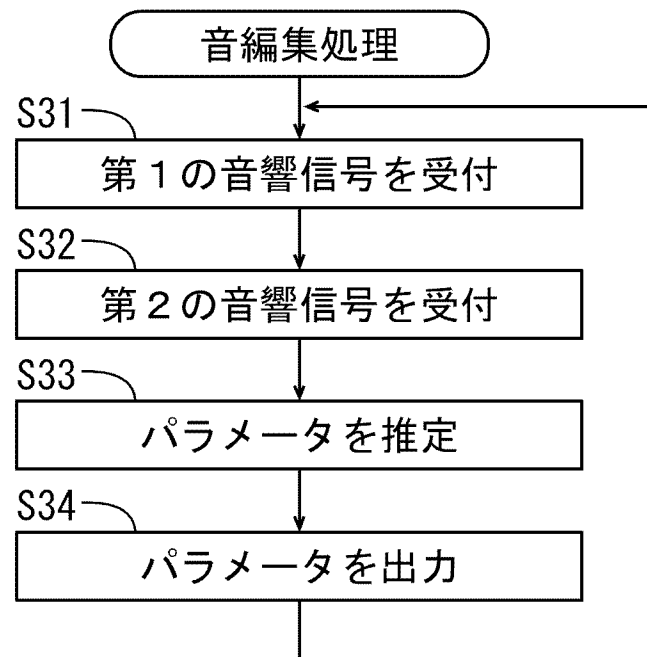
[図7]



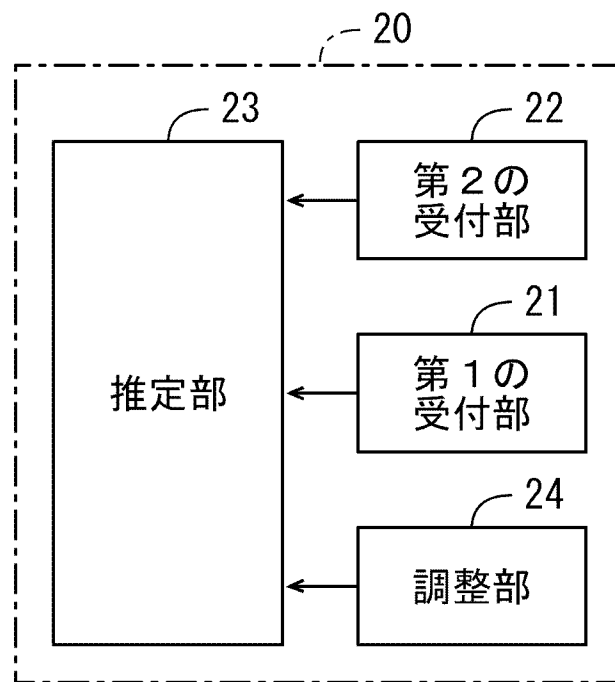
[図8]



[図9]



[図10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/010400**A. CLASSIFICATION OF SUBJECT MATTER****G10H 1/02**(2006.01)i

FI: G10H1/02

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)

G10H1/02

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.
A	JP 2020-13170 A (CASIO COMPUT. CO., LTD.) 23 January 2020 (2020-01-23) entire text, all drawings	1-15

☐ 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

11 May 2022

Date of mailing of the international search report

31 May 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/JP2022/010400

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2020-13170 A	23 January 2020	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G10H 1/02(2006.01)i FI: G10H1/02														
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）） G10H1/02 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> 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）			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													
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												
A	JP 2020-13170 A（カシオ計算機株式会社）23.01.2020（2020 - 01 - 23） 全文、全図	1 - 15												
☐ 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、一般的技術水準を示すもの</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、一般的技術水準を示すもの	“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、一般的技術水準を示すもの	“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 11.05.2022	国際調査報告の発送日 31.05.2022													
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） 大野 弘 5Z 9175 電話番号 03-3581-1101 内線 3591													

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/JP2022/010400

引用文献			公表日	パテントファミリー文献	公表日
JP	2020-13170	A	23.01.2020	(ファミリーなし)	