

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7286894号
(P7286894)

(45)発行日 令和5年6月6日(2023.6.6)

(24)登録日 令和5年5月29日(2023.5.29)

(51)国際特許分類

F I

G 1 0 L 25/51 (2013.01)

G 1 0 L 25/51

G 1 0 L 25/30 (2013.01)

G 1 0 L 25/30

請求項の数 11 (全18頁)

(21)出願番号	特願2020-571028(P2020-571028)	(73)特許権者	304023994
(86)(22)出願日	令和1年12月17日(2019.12.17)		国立大学法人山梨大学
(86)国際出願番号	PCT/JP2019/049337		山梨県甲府市武田四丁目4番37号
(87)国際公開番号	WO2020/162048	(74)代理人	110000660
(87)国際公開日	令和2年8月13日(2020.8.13)		Knowledge Partners
審査請求日	令和4年3月18日(2022.3.18)		弁理士法人
(31)優先権主張番号	特願2019-20267(P2019-20267)	(72)発明者	西 崎 博光
(32)優先日	平成31年2月7日(2019.2.7)		山梨県甲府市武田四丁目4番37号 国
(33)優先権主張国・地域又は機関	日本国(JP)	(72)発明者	澤田 直輝
			山梨県甲府市武田四丁目4番37号 国
			立大学法人山梨大学内
		審査官	山下 剛史

最終頁に続く

(54)【発明の名称】 信号変換システム、機械学習システムおよび信号変換プログラム

(57)【特許請求の範囲】

【請求項1】

時間変化する信号であって、時間軸上の複数の位置における値のそれぞれが1個の成分で表現された前記信号を取得する信号取得部と、

複数の前記位置毎の値のそれぞれを、複数の成分の値で表現した多次元量に変換する変換部と、

前記多次元量を、連続する複数の前記位置における前記多次元量の同一の成分の値を少なくとも含む、L個（Lは1以上、前記多次元量の成分の数以下の整数）の時系列データとして出力する出力部と、

L個の前記時系列データを入力して前記信号の特徴に関する情報を出力するように機械学習された機械学習モデルに対して、L個の前記時系列データを入力し、前記信号の特徴に関する情報を出力する特徴出力部と、
を備える信号変換システム。

【請求項2】

前記機械学習モデルは、
畳み込みニューラルネットワークを含む、
請求項1に記載の信号変換システム。

【請求項3】

前記変換部は、
複数の前記位置毎の値のそれぞれを、n進数（nは2以上の整数）の値に変換し、変換

された値の各桁を前記多次元量の成分とする、
請求項 1 または請求項 2 に記載の信号変換システム。

【請求項 4】

n は 2 である、
請求項 3 に記載の信号変換システム。

【請求項 5】

前記信号は音を示す、
請求項 1 ~ 請求項 4 のいずれかに記載の信号変換システム。

【請求項 6】

前記時系列データは、
前記多次元量から、同一の成分の値の時間変化を示す 1 次元データが L 個（L は 1 以上、前記多次元量の成分の数以下の整数）抽出されたデータである、
請求項 1 ~ 請求項 5 のいずれかに記載の信号変換システム。

10

【請求項 7】

前記時系列データは、
前記多次元量から、
時間軸上の同一の前記位置における異なる成分の値が第 1 軸方向に並び、
時間軸上の異なる前記位置における同一の成分の値が前記第 1 軸方向と異なる第 2 軸方向に並ぶ 2 次元データが抽出されたデータである、
請求項 1 ~ 請求項 5 のいずれかに記載の信号変換システム。

20

【請求項 8】

前記 2 次元データを入力する畳み込みニューラルネットワークを含む機械学習済モデルに基づいて、前記信号の特徴に関する情報を出力する特徴出力部、をさらに備える、
請求項 7 に記載の信号変換システム。

【請求項 9】

機械学習システムに利用されるニューラルネットワークへの入力信号の信号変換システムであって、
時間変化する信号であって、時間軸上の複数の位置における値のそれぞれが 1 個の成分で表現された前記信号を取得する信号取得部と、
複数の前記位置毎の値のそれぞれを、2 進数の値に変換し、変換された値の各桁が成分となる多次元量に変換する変換部と、
前記多次元量から、連続する複数の前記位置における前記多次元量の同一の成分の値の時間変化を示す 1 次元データを L 個（L は 1 以上、前記多次元量の成分の数以下の整数）抽出して出力する出力部と、
L 個の前記 1 次元データを入力して前記信号の特徴に関する情報を出力するように機械学習された機械学習モデルに対して、L 個の前記 1 次元データを入力し、前記信号の特徴に関する情報を出力する特徴出力部と、
を備える、信号変換システム。

30

【請求項 10】

請求項 1 ~ 請求項 7 のいずれかに記載の信号変換システムと、
前記時系列データを入力し、前記信号の特徴に関する情報を出力する機械学習済モデルを機械学習する機械学習部と、
を備える機械学習システム。

40

【請求項 11】

コンピュータを、
時間変化する信号であって、時間軸上の複数の位置における値のそれぞれが 1 個の成分で表現された前記信号を取得する信号取得部、
複数の前記位置毎の値のそれぞれを、複数の成分の値で表現した多次元量に変換する変換部、
前記多次元量を、連続する複数の前記位置における前記多次元量の同一の成分の値を少

50

なくとも含む、 L 個 (L は 1 以上、前記多次元量の成分の数以下の整数) の時系列データとして出力する出力部、

L 個の前記時系列データを入力して前記信号の特徴に関する情報を出力するように機械学習された機械学習モデルに対して、 L 個の前記時系列データを入力し、前記信号の特徴に関する情報を出力する特徴出力部、
として機能させる信号変換プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、信号変換システム、機械学習システムおよび信号変換プログラムに関する。

10

【背景技術】

【0002】

従来、各種の分類や認識等に機械学習が利用されている。例えば、特許文献 1 には、入力情報を示す数値を 2 進数のビット系列に変換する技術が開示されている。具体的には、単語を示す ID の数値を 2 進数のビット系列に変換する構成が開示されている。

【先行技術文献】

【特許文献】

【0003】

【文献】特許第 6 2 5 9 9 4 6 号公報

【発明の概要】

20

【発明が解決しようとする課題】

【0004】

時間変化する信号に関して分類や認識、あるいは未来の信号変化の予測を精度良く実行するためには、当該信号の時間変化の特徴を詳細に捉えることが可能なフォーマットで信号を表現する必要があるが、従来、このような表現は知られていなかった。例えば、時間変化する波形信号の振幅をサンプリングして 10 進数の値で表現した場合、微細な振幅の変化や大局的な振幅の時間変化を直接的に捉えることは困難である。フーリエ変換等によって周波数空間の情報を得ようとしても、時間空間で短期間の信号を利用すると周波数空間の情報の精度が低下する。時間空間で長期間の信号を利用すると周波数空間の情報の精度が向上するが、短期間の信号の分類や認識に不適切な信号となってしまう。

30

本発明は、前記課題にかんがみてなされたもので、信号の時間変化の特徴を捉えやすいデータを生成する技術を提供することを目的とする。

【課題を解決するための手段】

【0005】

上述の目的を達成するため、信号変換システムは、時間変化する信号であって、時間軸上の複数の位置における値のそれぞれが 1 個の成分で表現された信号を取得する信号取得部と、複数の位置毎の値のそれぞれを、複数の成分の値で表現した多次元量に変換する変換部と、多次元量を、連続する複数の位置における多次元量の同一の成分の値を少なくとも含む、 L 個 (L は 1 以上、多次元量の成分の数以下の整数) の時系列データとして出力する出力部と、 L 個の時系列データを入力する機械学習モデルに基づいて、信号の特徴に関する情報を出力する特徴出力部と、を備える。

40

【0006】

すなわち、信号変換システムにおいては、時間軸上の 1 箇所の位置の値が 1 個の成分で表現された信号を、変換部が、複数の成分の値で表現した多次元量に変換する。この結果、元の信号よりも多様な特徴を捉えることが可能なフォーマットで信号を表現することができる。そして、出力部においては、多次元量を、時間軸上で連続する複数の位置における多次元量の同一の成分の値を少なくとも含む L 個の時系列データとして出力する。さらに、当該 L 個の時系列データを機械学習モデルに適用することで信号の特徴に関する情報を出力する。すなわち、時系列データにおいては、多次元量の同一の成分の値の時間変化を示しており、成分毎に信号の時間変化の特徴を捉えることが可能である。このような構

50

成によれば、信号の時間変化の特徴を捉えやすいデータを生成することが可能である。

【図面の簡単な説明】

【 0 0 0 7 】

【図 1】一実施形態のシステム構成を示す図である。

【図 2】データの変換および出力を説明する図である。

【図 3】学習対象のモデルを示す図である。

【図 4】機械学習処理のフローチャートである。

【図 5】特徴出力処理のフローチャートである。

【図 6】時系列データの例を示す図である。

【図 7】他の実施形態における学習対象のモデルを示す図である。

10

【発明を実施するための形態】

【 0 0 0 8 】

ここでは、下記の順序に従って本発明の実施の形態について説明する。

(1) システムの構成：

(1 - 1) 機械学習処理：

(1 - 2) 特徴出力処理：

(2) 他の実施形態：

【 0 0 0 9 】

(1) システムの構成：

図 1 は、本発明の一実施形態である信号変換システム、特徴出力システム、機械学習システムとして機能するコンピュータ 10 の概略構成を示す図である。コンピュータ 10 は、マイクロホン 10 a、A/D 変換部 10 b、制御部 20、記憶媒体 30、表示部 40 を備えている。制御部 20 は、図示しない CPU、RAM、ROM を備えており、記憶媒体 30 等に記憶された各種プログラムを実行することができる。マイクロホン 10 a、A/D 変換部 10 b、制御部 20、記憶媒体 30、表示部 40 は、一体的なコンピュータで構成されていても良いし、少なくとも一部が別の装置であり、USB ケーブル等によって接続される構成であっても良い。コンピュータ 10 の態様は種々の態様であって良く、例えば、据置型の汎用コンピュータであっても良いし、スマートフォンやタブレット端末などの可搬型のコンピュータであっても良い。

20

【 0 0 1 0 】

マイクロホン 10 a は、周囲の音を集音し、音圧の時間変化を示すアナログ音波信号を出力する。A/D 変換部 10 b は、アナログ信号をデジタル信号に変換する装置であり、制御部 20 に指示されたサンプリング周波数でアナログ音波信号をサンプリングしてデジタル音波信号として出力する。本実施形態において、デジタル音波信号は、時間軸上の複数の位置毎の音圧の振幅値で構成され、当該振幅値は 10 進数で表現されている。従って、当該デジタル音波信号は、時間軸上の複数の位置における値のそれぞれが 1 個の成分で表現された信号である。

30

【 0 0 1 1 】

記憶媒体 30 は、各種の情報を記憶可能な媒体であり、本実施形態においては、機械学習のための教師データ 30 a が記憶媒体 30 に記憶される。また、機械学習によって機械学習済モデルが生成されると、当該モデルを示す情報が機械学習済モデル 30 b として記憶媒体 30 に記憶される。表示部 40 は、各種の情報を表示するディスプレイである。

40

【 0 0 1 2 】

本実施形態において、制御部 20 は、信号変換プログラムを実行することができる。信号変換プログラムは、時間変化する信号を変換して出力する機能を制御部 20 に実行させるプログラムである。また、本実施形態において信号変換プログラムは、変換された信号に基づいて機械学習を行う機能と、変換された信号に基づいて音の特徴を示す出力を行う機能とを有している。

【 0 0 1 3 】

信号変換プログラムが実行されると、制御部 20 は、信号取得部 20 a、変換部 20 b

50

、出力部 20c、機械学習部 20d、特徴出力部 20e として機能する。本実施形態においては、音源の種類が音源の特徴として出力される。すなわち、制御部 20 は、音源の種類を分類することができる。分類の数は任意であって良いが、ここでは、音源が音楽であるか否かを分類する例を説明する。

【0014】

本実施形態においては、機械学習を利用して音源の種類を分類する。すなわち、分類を行う前に予め機械学習済モデル 30b が生成され、分類対象の音を示すデジタル信号を機械学習済モデル 30b に入力することによって分類結果を出力する。このような分類を正確に行うためには、音の詳細な特徴を捉えられるデジタル信号を機械学習済モデル 30b への入力値とすることが好ましい。

10

【0015】

このため、本実施形態において制御部 20 は、音を示すデジタル音波信号を多次元量に変換し、音の詳細な特徴を捉えられる時系列データとして出力する機能を有している。時系列データの例としては、図 2 の点線で囲われたデータ列が挙げられる。詳細については後述する。本実施形態においては、この時系列データを用いて機械学習し、この時系列データを用いて分類を行うことによって、高精度に音の分類を実行するように構成されている。すなわち、本実施形態においては、機械学習対象のモデルを学習する際の入力データとして時系列データが利用され、また、機械学習済モデル 30b に基づいて分類を行うための入力データとしても時系列データが利用される。

【0016】

20

本実施形態において、当該時系列データの生成は、制御部 20 が、信号取得部 20a、変換部 20b、出力部 20c の機能を実行することによって実現される。また、時系列データを利用した機械学習は、制御部 20 が、機械学習部 20d の機能を実行することによって実現される。さらに、時系列データを利用した音源の種類の分類は、制御部 20 が、特徴出力部 20e の機能を実行することによって実現される。以下においては、機械学習と、音源の種類の分類（特徴出力）を順に説明する。

【0017】

（1-1）機械学習処理：

本実施形態においては、機械学習を行うために予め教師データ 30a が用意される。教師データ 30a は、音波信号と当該音波信号が示す音の音源の種類とを対応づけた情報である。例えば、音楽を示す音波信号に音楽であることが対応づけられた 1 組のデータや、音楽以外の音（人の発話等）を示す音波信号に音楽ではないことが対応づけられた 1 組のデータが教師データ 30a となり得る。むろん、教師データ 30a としては、機械学習を行うために十分な量のデータが予め用意される。

30

【0018】

教師データ 30a は、種々の手法で用意されて良く、例えば、複数のクライアントで取得されたデータがサーバ等で収集されるなどして用意されて良い。本実施形態においては、コンピュータ 10 で教師データ 30a を生成することも可能である。すなわち、コンピュータ 10 が備えるマイクロホン 10a で音が集音され、A/D 変換部 10b で変換されたデジタル音波信号に対して音源の種類を示す情報が対応づけられて記憶媒体 30 に教師データ 30a として記憶されても良い。いずれにしても、教師データ 30a においては、時間軸上の複数の位置の音圧を示す値が 10 進数で表現されている。

40

【0019】

図 2 は、データの変換を説明するための図である。図 2 においては、音波信号 Ss の一部がグラフ上に示されている。マイクロホン 10a によって音波信号 Ss が集音されると、A/D 変換部 10b は、制御部 20 に指示されたサンプリング周波数で音波信号をデジタル音波信号に変換する。図 2 において音波信号 Ss に重ねられた黒丸はサンプリングされた振幅を示しており、その下部にサンプリング値が示されている。例えば、図 2 においては、最初のサンプリング値は 10 進数で 0、次のサンプリング値は 10 進数で 12 である。教師データ 30a においては、このようにして得られた時間軸上の複数の位置におけ

50

る値のセットに対して、音源の種類を示す情報（ラベル）が対応づけられている。

【 0 0 2 0 】

本実施形態において、機械学習処理は、ニューラルネットワークを形成する訓練モデルを最適化する処理であり、本実施形態においては、図 4 に示すフローチャートに従って実行される。ここで、モデルとは、分類対象のデータと分類結果のデータとの対応関係を導出する式を示す情報であり、本実施形態において分類結果は、音楽であるか否かである。分類対象は音波信号から変換、出力された時系列データである。すなわち、本実施形態においては、機械学習のモデルに対して、A/D 変換部 1 0 b されたデジタル音波信号（1 0 進数表現）を直接入力するのではなく、デジタル音波信号を変換し、出力した時系列データを入力する。

10

【 0 0 2 1 】

このように時系列データを利用した処理を行うために、制御部 2 0 は、信号取得部 2 0 a、変換部 2 0 b、出力部 2 0 c を備えている。信号取得部 2 0 a は、時間変化する信号であって、時間軸上の複数の位置における値のそれぞれが 1 個の成分で表現された信号を取得する機能を制御部 2 0 に実行させる。本実施形態において当該時間変化する信号は、音波信号を既定のサンプリング周波数でサンプリングして 1 0 進数の値で表現した信号であり、機械学習の際には教師データ 3 0 a が当該信号に相当する。すなわち、機械学習処理が開始されると、制御部 2 0 は、記憶媒体 3 0 を参照し、各教師データ 3 0 a から音波信号を取得する（ステップ S 1 0 0）。この結果、図 2 に示すように、音波信号の複数の時間軸上の位置における振幅を 1 0 進数で表現したデータセットが複数個得られる。

20

【 0 0 2 2 】

変換部 2 0 b は、複数の位置毎の値のそれぞれを、複数の成分の値で表現した多次元量に変換する機能を制御部 2 0 に実行させるプログラムモジュールである。すなわち、1 0 進数による振幅の表現は、時間軸上のある位置における振幅の大きさを 1 個の成分で示しているため、制御部 2 0 は、より多数の成分によって音波信号を表現するため、音波信号を多次元量に変換する（ステップ S 1 0 5）。

【 0 0 2 3 】

多次元量への変換では、1 個の成分での表現を複数個の成分での表現に変換することができればよく、本実施形態において制御部 2 0 は、時間軸上の各位置における振幅の大きさを 2 進数の値に変換する。すなわち、変換後の 2 進数の値の各桁が多次元量の成分となる。図 2 においては、音圧の振幅の大きさを示す値の下方に変換後の 2 進数の値（8 ビット）を示している。この例では、上方から下方に向けて下位の桁から上位の桁が並ぶように表示してある。例えば、最初のサンプリング値である 1 0 進数の 0 は 2 進数で 0 0 0 0 0 0 0 0、次のサンプリング値である 1 0 進数の 1 2 は 2 進数で 0 0 0 0 1 1 0 0 である。なお、実際のデジタル音波信号における時間長は、図 2 に示された時間長より長く、例えば、1 0 k H z でサンプリングされた 1 秒分のデータ（1 0 0 0 0 個のデータ）等でデジタル音波信号が構成され、当該デジタル音波信号が変換対象となる。

30

【 0 0 2 4 】

出力部 2 0 c は、多次元量を、連続する複数の位置における多次元量の同一の成分の値を少なくとも含む時系列データとして出力する機能を制御部 2 0 に実行させるプログラムモジュールである。本実施形態においては、ステップ S 1 0 5 で変換された 2 進数の値の桁を成分と見なしており、制御部 2 0 は、出力部 2 0 c の機能により、桁毎の値を抽出することで時系列データを出力する（ステップ S 1 1 0）。

40

【 0 0 2 5 】

例えば、図 2 に示す例においては、2 進数で表現された多次元量の最下位桁を一点鎖線で囲んで示している。制御部 2 0 は、当該最下位桁において時間軸上で異なる位置に存在する各値を、時間軸上の順序に従って並べたデータを、最下位桁成分の時系列データ D_1 として生成する。このように、本実施形態において時系列データは、同一の成分の値の時間変化を示す 1 次元データである。1 次元データは、多次元量の各成分について実施されてよい。すなわち、2 進数の上位桁について時系列データが抽出された場合、当該時系列デ

50

ータはデジタル音波信号の振幅の大きい変化を示している。一方、2進数の下位桁について時系列データが抽出された場合、当該時系列データはデジタル音波信号の振幅の細かい変化を示している。

【0026】

従って、機械学習の目的や分類対象に応じて、時系列データとして抽出すべき成分の数は、1以上、多次元量の成分の数以下の整数の中から選択されてよい。本実施形態においては、8桁の2進数の全成分（全桁）について時系列データが出力される。このため、図2においては、時系列データ D_1 以外にも、時系列データ $D_2 \sim D_8$ が出力されることが示されている。

【0027】

本実施形態においては、以上のようにして出力された時系列データに基づいて音の分類を行う。すなわち、時系列データを入力するモデルを機械学習によって最適化する機械学習処理を行う。機械学習は、種々の手法で行われて良いが、本実施形態においては、畳み込みニューラルネットワーク（CNN：Convolutional Neural Network）と再帰ニューラルネットワーク（RNN：Recurrent Neural Network）とを含むニューラルネットワークによって機械学習が行われる例を説明する。

【0028】

図3は、図2に示す例に適用可能なモデルの一例を示す図である。図3においては、下部から上方に向けて演算が進行するようにモデルが記載されている。図3に示す例において、初期の3層はCNNによって構成されている。また、図3に示す例において、CNNによる出力結果はRNNに入力され、RNNによる出力は全結合を経て出力層に至る。

【0029】

図3の最下部には入力データとなる8チャンネル分の時系列データ $D_1 \sim D_8$ が例示されている。ここでは、時間軸上の位置が10000個、すなわち、10kHzで10000回サンプリングされることによって得られたデジタル音波信号が変換され、時系列データとして出力された例を示している。本実施形態においては、時系列データ $D_1 \sim D_8$ の時間方向の情報を畳み込む演算が行われる。

【0030】

すなわち、時系列データ $D_1 \sim D_8$ のそれぞれは1次元の時系列データであり、合計で8チャンネルである。そこで、本実施形態においては、2次元（時間軸方向とチャンネル方向）のフィルタが用意され、畳み込み演算が行われる。図3に示す例において、フィルタは2次元であり、 30×8 個の値（時間軸上で連続する30個の位置の値が8チャンネル分存在する）を畳み込むフィルタである。図3においては、当該フィルタの大きさが $1 \times 30 \times 8$ として示されている。むしろ、フィルタの大きさは30に限定されず、例えば50、40、20または10でもよいし、これら例示した数値の何れか2つの間の範囲であってもよい。フィルタの大きさが短い方が信号を取得後の処理を早く始めることができる。また図3では入力データを8チャンネルとしているがこれも8に限定されない。フィルタの数も限定されないが、図3に示す例においてCNNの最初の層でのフィルタ数は128個である。図3においては、これらのフィルタが $F_{1-1} \sim F_{1-128}$ として示されている。本実施形態においては、このように時間軸方向の大きさが短いフィルタを用いることができるので、従来技術に比べて時系列データの微細な変化を捉えることができる。従来技術では、時間変化する波形信号の振幅をサンプリングして10進数の値で表現した場合、時間軸方向の大きさが短いフィルタでは微細な振幅の変化や大局的な振幅の時間変化を直接的に捉えることは困難である。フーリエ変換等では、時間空間で短期間の信号を利用すると周波数空間の情報の精度が低下してしまう。すなわち信号を取得してから特徴を出力するまでにより時間が必要となる。

【0031】

各フィルタは、既定のパディングやストライドで時系列データ $D_1 \sim D_8$ のそれぞれに適用される。例えば、ストライド10で適宜パディングを行えば10000個のデータは1000個になる。これらを足し合わせることによって1個のフィルタによる出力結果が得

10

20

30

40

50

られる。従って、フィルタ数が128個の場合、1000個のデータを有する1次元のデータが128個生成される。図3においては、当該 $1 \times 1000 \times 128$ のデータが長方形によって示されている。むろん、ストライドやパディング等のハイパーパラメータは一例であり、出力されるデータの数(1000個等)等も一例である(他の層も同様)。

【0032】

いずれにしてもCNNの最初の層による演算によって $1 \times 1000 \times 128$ のデータが出力されると、このデータは次のCNNの層への入力データとなる。図3においては、第2層で $1 \times 30 \times 128$ の大きさのフィルタが256個用意され、さらに第3層で $1 \times 30 \times 256$ の大きさのフィルタが512個用意される例が示されている。ストライドやパディングは最初の層と同様である。この例であれば、以上の3層の演算により、 $1 \times 10 \times 512$ 個のデータ D_{CNN} が出力される。

10

【0033】

当該データ D_{CNN} は、RNNへの入力になる。図3に示すRNNを構成する要素はLSTM(Long Short-Term Memory)であるが、要素はGRU(Gated Recurrent Unit)であってもよいし、双方向のモデルが採用されてもよく、種々の構成を採用可能である。むろん、ハイパーパラメータは種々のパラメータとして良い。

【0034】

いずれにしても、図3に示す例においては、RNNに入力されるデータ D_{CNN} が $1 \times 10 \times 512$ 個のデータであるため、この例ではLSTMに対して512個のデータを10回入力する演算が行われる。図3においては、最初に入力される512個の入力値を $X_{1-1} \sim X_{1-512}$ 、次に入力される512個の入力値を $X_{2-1} \sim X_{2-512}$ などのようにして示している。

20

【0035】

さらに、図3に示すLSTMではLSTMからの出力が1024個である例が想定されており、最後のLSTMから出力されるデータ $Y_1 \sim Y_{1024}$ が次の層に入力される。すなわち、データ $Y_1 \sim Y_{1024}$ が全結合層に入力され、全結合層の次の層で音楽であるか否かの分類結果を出力する。

【0036】

本実施形態において制御部20は、機械学習部20dの機能により、以上のようなモデルに基づいて機械学習処理を実行する。すなわち、制御部20は、機械学習部20dの機能により、時系列データを入力し、信号の音源の特徴に関する情報を出力する機械学習済モデルを機械学習する。具体的には、制御部20は、機械学習部20dの機能により、訓練モデルを取得する(ステップS115)。すなわち、図3に示すようなニューラルネットワークの構造に対応した各種の情報(モデルを示すフィルタや活性化関数等の情報)が予め定義されており、制御部20は、当該情報を取得することで訓練モデルを取得する。

30

【0037】

次に、制御部20は、機械学習部20dの機能により、教師データ30aのラベルを取得する(ステップS120)。すなわち、制御部20は、ステップS100で取得した各音波信号の音源の種類を示すラベルを取得する。例えば、図3に示す例であれば、ラベルは出力層のノードの出力値が1, 0のいずれであるかを示しており、ステップS100で取得された音波信号の音源が音楽である場合に1, 音楽以外の音である場合に0が取得される。

40

【0038】

次に、制御部20は、機械学習部20dの機能により、テストデータを取得する(ステップS125)。本実施形態において制御部20は、ステップS110で取得されたデータの一部を抽出し、ステップS120で取得されたラベルを対応づけてテストデータとする。テストデータは、学習の汎化が行われたか否かを確認するためのデータであり、機械学習には使用されない。

【0039】

次に、制御部20は、機械学習部20dの機能により、初期値を決定する(ステップS

50

130)。すなわち、制御部20は、ステップS115で取得した訓練モデルのうち、学習対象となる可変のパラメータ（フィルタの重みやバイアス等）に対して初期値を与える。初期値は、種々の手法で決定されて良い。むしろ、学習の過程でパラメータが最適化されるように初期値が調整されても良いし、各種のデータベース等から学習済のパラメータが取得されて利用されても良い。

【0040】

次に、制御部20は、機械学習部20dの機能により、学習を行う（ステップS135）。すなわち、制御部20は、ステップS115で取得した訓練モデルにステップS110で出力された時系列データを入力し、分類結果を示す情報を出力する。分類結果を示す情報が出力されると、制御部20は、当該出力と、ステップS120で取得されたラベルとの誤差を示す損失関数によって誤差を特定する。損失関数が得られたら、制御部20は、既定の最適化アルゴリズム、例えば、確率的勾配降下法等によってパラメータを更新する。すなわち、制御部20は、損失関数のパラメータによる微分に基づいてパラメータを更新する処理を既定回数繰り返す。

10

【0041】

以上のようにして、既定回数のパラメータの更新が行われると、制御部20は、訓練モデルの汎化が完了したか否かを判定する（ステップS140）。すなわち、制御部20は、ステップS125で取得したテストデータを訓練モデルに入力して音波信号の分類結果を出力する。そして、制御部20は、出力された分類結果と、テストデータに対応づけられた分類結果とが一致している数（分類結果とラベルとの誤差が既定値以下である数）を取得し、サンプル数で除することで分類精度を取得する。本実施形態において、制御部20は、分類精度が閾値以上である場合に汎化が完了したと判定する。

20

【0042】

なお、汎化性能の評価に加え、ハイパーパラメータの妥当性の検証が行われてもよい。すなわち、学習対象となる可変のパラメータ以外の可変量であるハイパーパラメータ、例えば、フィルタサイズやノードの数等がチューニングされる構成において、制御部20は、検証データに基づいてハイパーパラメータの妥当性を検証しても良い。検証データは、ステップS125と同様の処理により、検証データを予め抽出し、訓練に用いないデータとして確保しておくことで取得すれば良い。

【0043】

ステップS140において、訓練モデルの汎化が完了したと判定されない場合、制御部20は、ステップS135を繰り返す。すなわち、さらに学習対象となる可変のパラメータを更新する処理を行う。一方、ステップS140において、訓練モデルの汎化が完了したと判定された場合、制御部20は、機械学習済モデルを記録する（ステップS145）。すなわち、制御部20は、訓練モデルを機械学習済モデル30bとして記憶媒体30に記録する。

30

【0044】

（1-2）特徴出力処理：

次に、音源の種類を分類する特徴出力処理を図5に示すフローチャートに基づいて説明する。コンピュータ10の周辺で分類対象の音源から音が出力されている状態で、コンピュータ10の利用者は、キーボードやマウスなどの図示しない入力部を操作して特徴出力処理の実行開始指示を行う。利用者が当該実行開始指示を行うと、制御部20は、特徴出力処理の実行を開始する。

40

【0045】

特徴出力処理の実行が開始されると、制御部20は、音波信号を取得する（ステップS200）。すなわち、音源から音波が出力されると、マイクロホン10aは当該音波を取得してアナログ音波信号を出力する。制御部20は機械学習の際に使用されたサンプリング周波数と同一のサンプリング周波数を予めA/D変換部10bに指示している。A/D変換部10bは当該サンプリング周波数でアナログ音波信号をサンプリングし、デジタル音波信号に変換する。そして、A/D変換部10bからデジタル音波信号が出力されると

50

、制御部 20 は、信号取得部 20 a の機能により、当該デジタル音波信号を取得する。例えば、図 2 に示す音波信号 S_s であれば、A/D 変換部 10 b によって 10 進数のデジタル音波信号 (0, 12, 6, , , ,) が出力されるため、制御部 20 は、当該デジタル音波信号を取得する。

【0046】

次に、制御部 20 は、A/D 変換部 10 b の機能により、音波信号を多次元量に変換する (ステップ S205)。すなわち、制御部 20 は、ステップ S200 で取得されたデジタル音波信号を取得し、時間軸上の複数の位置のそれぞれの振幅を示す 10 進数の値を 2 進数に変換する。例えば、ステップ S200 において、図 2 に示すデジタル音波信号 (0, 12, 6, , , ,) が取得された場合、図 2 に示す 2 進数の多次元量 (00000000, 00001100, , , ,) が取得される。

10

【0047】

次に、制御部 20 は、出力部 20 c の機能により、時系列データを出力する (ステップ S210)。すなわち、制御部 20 は、多次元量の桁毎の値を抽出し、各桁の時系列データを生成する。例えば、図 2 に示す例であれば、制御部 20 は、時系列データ $D_1 \sim D_8$ を取得する。

【0048】

次に、制御部 20 は、特徴出力部 20 e の機能により、音の特徴を判定する (ステップ S215)。すなわち、制御部 20 は、機械学習済モデル 30 b を取得し、ステップ S210 で出力された時系列データ $D_1 \sim D_8$ を当該機械学習済モデル 30 b に対して入力する。制御部 20 は、当該機械学習済モデル 30 b が示すパラメータを利用して図 3 に示す CNN, RNN 等の演算を行う。この結果、出力層の値が決定される。なお、図 2, 図 3 に示す例においては、8 チャンネルの 1 次元データである時系列データ $D_1 \sim D_8$ を機械学習済モデル 30 b に入力しているが、チャンネル数は 8 に限定されず、1 以上、多次元量の成分の数以下の整数であって良い。

20

【0049】

制御部 20 は、出力層の値と閾値とを比較し、閾値以上であれば音源の種類が音楽であると判定し、閾値より小さければ音源の種類が音楽ではないと判定する。例えば、出力層の値が 0 ~ 1 の範囲で変化し、閾値が 0.9 である場合、出力層の値が 0.9 以上であれば、制御部 20 は、音源の種類が音楽であると見なす。

30

【0050】

次に、制御部 20 は、特徴出力部 20 e の機能により、判定結果を出力する (ステップ S220)。すなわち、制御部 20 は、表示部 40 を制御して、音源の種類を示す情報を表示させる。以上の構成においては、10 進数の数値で表現された振幅をより多数の成分の値で表現した多次元量に変換して機械学習した結果に基づいて音の特徴を判定することができる。従って、10 進数で表現された元の信号よりも多様な特徴を捉えた機械学習を行って音の特徴を判定することができ、高精度に音の特徴を判定することができる。

【0051】

さらに、本実施形態においては、多次元量を、同一成分 (同一桁毎) 毎に分解し、同一成分の値の時間変化を示す時系列データを用いて機械学習および特徴の取得を行う。この結果、大局的な時間変化と微細な時間変化の双方を含んだ音波信号の時間変化を詳細に捉えることができる。従って、10 進数で表現された元の信号よりも多様な時間変化を捉えた機械学習を行って音の特徴を判定することができ、高精度に音の特徴を判定することができる。なお、微細な時間変化がノイズであり、微細な時間変化を除いた部分に信号の特徴が現れる場合には、微細な時間変化が分類に与える重みが小さくなるように学習される。従って、ノイズが含まれていても高精度に音の特徴を判定することができる。

40

【0052】

ここで、本実施形態を用いて環境音イベント分類を行った場合と、公知の手法を用いて環境音イベント分類を行った場合とにおける分類精度を比較する。環境音は、参考文献の Table 2 に記載された 28 種類のイベントを含むデータベースを利用した。このデータベ

50

ースに基づいて、参考文献に記載されたようにパワースペクトル（256次元）とその変化量、の3チャンネルを分類モデルに入力し、環境音イベントを分類する例を対比例とする。なお、この場合の分類モデルは、6層のCNNを通した結果を3層の全結合層に通すモデルである。当該分類モデルで機械学習を行った場合の分類精度は、80.3%であった。

【0053】

一方、このデータベースに基づいて、本実施形態の図3に基づいて機械学習を行った場合、分類精度は88.3%であった。このように、本実施形態によれば、公知の手法と比較して分類精度が大幅に改善している。（参考文献）N. Takahashi, et al., "Deep Convolutional Neural Networks and Data Augmentation for Acoustic Event Detection," Proc. of INTERSPEECH2016, pp.2982-2986, 2016

10

【0054】

（2）他の実施形態：

以上の実施形態は本発明を実施するための一例であり、時間変化する信号の時間軸上の各位置における値を多次元量で表現し、多次元量の成分毎の時間変化を示す時系列データを出力する限りにおいて、他にも種々の実施形態を採用可能である。例えば、1台のコンピュータ10によって、信号変換システム、機械学習システム、特徴出力システムが構成されていたが、それぞれが別のシステムであっても良いし、2個のシステムが1台のシステムとして構成されてもよい。

【0055】

20

むろん、別個のシステムの連携法は種々の手法であって良く、クライアントとサーバとして構成されていても良い。さらに、信号取得部20a、変換部20b、出力部20c、機械学習部20d、特徴出力部20eの少なくとも一部が複数の装置に分かれて存在してもよい。例えば、信号取得部20aによって教師データ30aが取得される処理と、分類対象のデジタル音波信号が取得される処理とが異なる装置で実施される構成等であっても良い。むろん、上述の実施形態の一部の構成が省略されてもよいし、処理の順序が変動または省略されてもよい。

【0056】

例えば、上述の実施形態に示す機械学習のモデルによる処理過程で得られる情報が出力されるシステムが構成されても良い。具体的には、図3に示すモデルの過程でCNNのデータ D_{CNN} が得られるが、当該データ D_{CNN} が音の特徴を示す情報であるとして出力されても良い。当該情報は、音の特徴を示しているため、例えば、特定の音の種類の場合におけるデータ D_{CNN} のパターンを予め特定しておけば、音の種類の分類を行うことが可能である。むろん、機械学習を行うモデルは図3に示す例に限定されず、例えば、RNNが省略され、CNNのデータ D_{CNN} が全結合層に入力され、その後の層で分類結果を示す情報が出力される構成等であっても良い。

30

【0057】

機械学習の態様は限定されず、例えばニューラルネットワークによる機械学習が行われる場合、モデルを構成する層の数やノードの数、活性化関数の種類、損失関数の種類、勾配降下法の種類、勾配降下法の最適化アルゴリズムの種類、ミニバッチ学習の有無やバッチの数、学習率、初期値、過学習抑制手法の種類や有無、畳み込み層の有無、畳み込み演算におけるフィルタのサイズ、フィルタの種類、パディングやストライドの種類、プーリング層の種類や有無、全結合層の有無、再帰的な構造の有無など、種々の要素を適宜選択して機械学習が行われればよい。むろん、他の機械学習、例えば、深層学習（ディープラーニング）、サポートベクターマシンやクラスタリング、強化学習等によって学習が行われてもよい。さらに、モデルの構造（例えば、層の数や層毎のノードの数等）が自動的に最適化される機械学習が行われてもよい。

40

【0058】

分類態様は音楽と非音楽に限定されず、音楽と音声の分類であってもよいし、より多くの種類の音源の分類であっても良い。むろん、信号変換システムや機械学習システム、特

50

徴出力システムの用途は、音源の種類の分類に限定されない。例えば、打音検査などの異常音検出に利用されても良いし、音楽の詳細分類（ジャンル分類や楽曲名分類等）に利用されても良いし、音声認識に利用されても良いし、音が関連するイベントの分類（環境音の種類の分類）に利用されても良い。利用場面も限定されず、例えば、放送音声に対する分類等が行われることによって放送内容の統計管理に利用されても良い。また、本発明は音源の分離に対しても適用できる。例えば音楽に対し音声为重畳された音源をニューラルネットワークに入力し音楽のみまたは音声のみを分離する技術が知られているが、本発明の時系列データをニューラルネットワークへ入力することにより、信号の時間変化の特徴を捉えやすくなるので、高精度・高効率な分離を実現できる。

【 0 0 5 9 】

10

信号取得部は、時間変化する信号であって、時間軸上の複数の位置における値のそれぞれが1個の成分で表現された信号を取得することができればよい。すなわち、信号取得部は、時間軸上の各位置における値を単一成分で表現した信号を取得する。成分は、信号の特徴を示しており、成分が1個であれば信号を1個の特徴で表現することになる。従って、上述の実施形態のように、音波信号の振幅を一定期間毎にサンプリングし、時間軸上の1箇所について値が1個（スカラー量）であるような信号は、1個の成分で表現された信号である。また、1個の成分で表現された信号としては他にも種々の態様が採用され得る。

【 0 0 6 0 】

また、時間変化する信号は、時間軸上の異なる位置における値が異なり得る信号であれば良く、音波に限定されない。例えば、超音波や脳波、心電図、筋電図などの生体信号や、温度や湿度、気圧といった環境信号などの任意の波は、時間変化する信号になり得る。さらに、株価や為替チャートなどのように、着目している対象が時間変化する場合、対象の時間変化を示す情報は時間変化する信号に該当し得る。さらに、動画等の画像信号などにも適応が可能である。動画の画像信号においては、複数枚の画像の時系列の変化によって動画が表現される。そして、各画像においては、カラー画像の場合に赤・緑・青等の3チャンネルで1画素の色が表現され、グレー画像の場合に1チャンネルで1画素の輝度が表現され、各チャンネルのそれぞれが階調値を持つ（通常は0～255の値）。したがって、これらの画素の階調値を多次元量に変換し、変換後の成分毎の値を示す時系列データを出力する構成を採用することもできる。むしろ、本発明の一実施形態で出力された特徴は、分類や認識以外にも種々の目的で利用されて良く、例えば、未来の信号変化の予測等に利用されても良い。以上のように、時間変化する信号の特徴が音源の種類である構成に限定されず、信号における種々の特徴が特徴出力部によって出力されて良い。

20

30

【 0 0 6 1 】

時間軸上の位置は、離散的な位置であれば良く、その間隔は任意の間隔であって良い。例えば、波形信号をサンプリングして信号を取得する場合、そのサンプリング周期は任意であるし、サンプリング周期は固定であっても良いし時間変化しても良い。

【 0 0 6 2 】

変換部は、複数の位置毎の値のそれぞれを、複数の成分の値で表現した多次元量に変換することができればよい。すなわち、変換部は、時間軸上の各位置で1個の値を、時間上の各位置で複数の値となるように変換する。多次元量への変換は、種々の手法でおこなわれてもよい。すなわち、時間軸上の各位置における1個の値を2進数に変換し、変換された値の各桁が多次元量の各成分となる構成以外にも種々の構成が採用されてよい。

40

【 0 0 6 3 】

例えば、信号波形を示すグラフを2次元の画像と捉え、当該画像を2値化すると、信号波が存在する位置が1、信号波が存在しない位置が0のデータが得られる。図6は、図2に示す音波信号 S_s と同様の信号を時間軸および振幅軸からなるグラフで表現した例を示している。また、図6においては、このグラフを画像と捉え、時間軸方向の複数の位置において振幅値が存在する部分を1、存在しない部分を0として2値化した例を示している。このような変換であっても、時間変化する信号の時間軸上の各位置における値が多次元量で表現されており、この例を含めて、多次元量への変換としては種々の変換が採用され

50

てよい。

【 0 0 6 4 】

むろん、多次元量の各成分の値を表現するための桁数は1に限定されず、例えば、2進数の複数桁分の値が1個の成分の値であっても良い。さらに、変換前の1個の成分の解釈法も種々の手法で行われてよい。例えば、変換前の1個の成分が10進数の1個の値であり複数桁に渡る場合、この値を桁毎の値に分解した後に桁毎にn進数（nは2以上の整数）の値に変換し、変換された値を多次元量の各成分の値としても良いし、変換された値の各桁を多次元量の各成分の値としても良い。

【 0 0 6 5 】

より具体的な例としては、例えば、変換前の1個の成分の値が10進数の"806"である場合に、"8","0","6"のそれぞれを2進数の"1000","0000","0110"とし、"1000","0000","0110"のそれぞれを多次元量の3個の成分の各値としても良い。また、"1000000000110"の各桁を多次元量の12個の成分の各値としても良い。

【 0 0 6 6 】

多次元量は、複数の成分で表現された値であれば良く、種々の手法で決められて良い。また、成分毎の関係も種々の関係であっても良い。例えば、上述の実施形態のように、2進数表現の各桁が成分であっても良いし、各成分が線形独立となるように表現された量であっても良い。

【 0 0 6 7 】

出力部は、多次元量を、連続する複数の位置における多次元量の同一の成分の値を少なくとも含む時系列データとして出力することができればよい。すなわち、時系列データは、多次元量の同一の成分の値が時間変化する様子を示していればよい。時系列データで表現された時間軸上の位置の数（時系列データが示す時間長）は、複数であれば任意の数であって良く、時系列データの用途等によって時間軸上の位置の数が選択されればよい。

【 0 0 6 8 】

また、時系列データは、多次元量の同一の成分の値が時間変化する様子を示していれば良いので、成分の数は1以上、多次元量の成分の数以下の整数の範囲で任意である。ここでも、時系列データの用途等によって成分の数が選択されればよい。

【 0 0 6 9 】

時系列データとされる多次元量の成分の数が複数である場合、時系列データは2次元データとなる。この場合、例えば、多次元量から、時間軸上の同一の位置における異なる成分の値が第1軸方向に並び、時間軸上の異なる位置における同一の成分の値が第1軸方向と異なる第2軸方向に並ぶ2次元データが抽出される構成であっても良い。

【 0 0 7 0 】

図2に示す例であれば、左下に示された8桁の2進数の全桁を含む2次元データが抽出されても良い。すなわち、図2に示す多次元量は、図の上下方向に見ると時間軸上の同一の位置における異なる成分（桁）の値が並んでいるため上下方向を第1軸と見なすことができる。さらに、図の左右方向に見ると時間軸上の異なる位置における同一の成分の値が並んでいるため左右方向（時間方向）を第2軸と見なすことができる。

【 0 0 7 1 】

このような2次元データであっても、各成分に着目すると、連続する複数の位置における多次元量の同一の成分の値を含んでいるため、このようなデータを出力部20cによる出力とする実施形態が構成されても良い。この場合、当該2次元データを入力する畳み込みニューラルネットワークを含む機械学習済モデル30bが機械学習される。また、当該2次元データを入力する畳み込みニューラルネットワークを含む機械学習済モデル30bに基づいて、信号の音源の特徴に関する情報が出力される。

【 0 0 7 2 】

このような構成は、図1に示す例において、機械学習モデルの構造、機械学習部20dおよび特徴出力部20eの処理を変更することによって実現される。図7は、2次元デー

10

20

30

40

50

タを入力する機械学習のモデル例である。図7においては、下部から上方に向けて演算が進行するようにモデルが記載されている。図7に示す例において、初期の3層はCNNによって構成されている。また、図7に示す例において、CNNによる出力結果はRNNに入力され、RNNによる出力は全結合を経て出力層に至る。

【0073】

図7の最下部には入力データとなる2次元データ D_{2D} が例示されている。ここでは、時間軸上の位置の数は任意である。例えば、10kHzで10000回サンプリングされることによって得られたデジタル音波信号が変換され10000個の時間軸上の位置についてのデータが得られた場合、当該データを2次元データ D_{2D} とすることができる。

【0074】

本例においては、2次元データ D_{2D} の時間方向および成分方向（桁方向）の情報を畳み込む演算が行われる。すなわち、2次元のフィルタが用意され、2次元データ D_{2D} に対して適用されて畳み込み演算が行われる。図7に示す例においては、フィルタの大きさが時間方向の大きさ x 、成分方向の大きさ y 、チャンネル方向の大きさ ch の立方体で示されている。すなわち、第1層のCNNにおけるフィルタの大きさは時間方向の大きさ x_1 、成分方向の大きさ y_1 、チャンネル方向の大きさ1、第2層のCNNにおけるフィルタの大きさは時間方向の大きさ x_2 、成分方向の大きさ y_2 、チャンネル方向の大きさ128、第3層のCNNにおけるフィルタの大きさは時間方向の大きさ x_3 、成分方向の大きさ y_3 、チャンネル方向の大きさ256である。

【0075】

また、図7に示す例において、第1層のCNNにおけるフィルタの数は128個、第2層のCNNにおけるフィルタの数は256個、第3層のCNNにおけるフィルタの数は512個である。むろん、フィルタの大きさやフィルタの数は適宜調整可能である。いずれにしても、各フィルタの演算において、パディングやストライドがハイパーパラメータとして予め決められ、畳み込み演算が行われると、順次、ニューラルネットワークの演算が行われる。例えば、第1層のCNNの畳み込み演算が行われると、出力が $i_1 \times j_1 \times 128$ 個のデータとなり、第2層のCNNの入力となる。

【0076】

このようなCNNの演算を第2層、第3層と繰り返すと、第2層の出力が $i_2 \times j_2 \times 256$ 個のデータとなり、第3層の出力が $i_3 \times j_3 \times 512$ 個のデータ D_{CNN} となる。図7に示すモデルにおいても当該データ D_{CNN} は、RNNへの入力になる。図7に示すRNNを構成する要素はBiGRU（Bidirectional Gated Recurrent Unit）であるが、要素はLSTM等であってもよく、種々の構成を採用可能である。むろん、ハイパーパラメータは種々のパラメータとして良い。いずれにしても、図7に示す例においては、RNNに入力されるデータ D_{CNN} が全結合層に入力され、全結合層の次の層で特徴の出力、例えば、音楽であるか否かの分類結果の出力が行われればよい。

【0077】

機械学習のモデルが予め決められた状態において、図1に示す構成によって図4に示す機械学習処理が実行されると、2次元データを入力として音の特徴を出力する機械学習済モデル30bが得られる。ただし、図4に示すステップS110において時系列データとして出力されるデータは、2次元データである。この状態において、制御部20は、ステップS115において図7に示すようなモデルを示す情報を訓練モデルとして取得する。そして、ステップS120～S130を経て制御部20がステップS135の学習を行えば、2次元データを入力として音の特徴を出力する機械学習済モデル30bが得られる。

【0078】

図5に示す特徴出力処理も同様であり、ステップS210で取得される時系列データが2次元データとなる。そして、ステップS215において、当該2次元データを制御部20が機械学習済モデル30bに入力すれば、音の特徴が判定される。以上のような2次元データであっても、時間軸上の位置毎の値が複数の成分に分解された状態で各成分の時間変化を含んでいる。従って、当該2次元データに基づいて機械学習および特徴の判定が行

10

20

30

40

50

われることにより、多様な時間変化を捉えた機械学習を行って音の特徴を判定することができ、高精度に音の特徴を判定することができる。

【0079】

なお、出力部による出力は、同一装置に対して行われてもよいし、異なる装置に対して行われてもよい。すなわち、上述の実施形態のように、出力部が、多次元量を時系列データに変換した結果を出力してさらに演算する構成であってもよいし、出力部が他の装置（例えば、外部に接続されたサーバ等）に出力する構成であってもよい。

【0080】

さらに、本発明のように、時間変化する信号の時間軸上の各位置における値を多次元量で表現し、多次元量の成分毎の時間変化を示す時系列データを出力する手法は、プログラムや方法としても適用可能である。また、以上のようなシステム、プログラム、方法は、単独の装置として実現される場合や、複数の装置によって実現される場合が想定可能であり、各種の態様を含むものである。また、一部がソフトウェアであり一部がハードウェアであったりするなど、適宜、変更可能である。さらに、システムを制御するプログラムの記録媒体としても発明は成立する。むろん、そのソフトウェアの記録媒体は、磁気記録媒体であってもよいし半導体メモリであってもよいし、今後開発されるいかなる記録媒体においても全く同様に考えることができる。

【0081】

以上説明したように、上記した例では信号取得部20aと、変換部20bと、出力部20cを少なくとも備えるシステムとして信号変換システムを構成している。また、この信号変換システムに、機械学習済モデル30bに基づいて、入力信号の特徴に関する情報を出力する特徴出力部20eを備えるようにして特徴出力システムを、同様にこの信号変換システムに、入力信号の特徴に関する情報を出力する機械学習済モデル30bを機械学習する機械学習部20dをさらに備えるようにして機械学習システムを構成している。なお、本実施形態での信号変換システムは上述のように信号変換プログラムとして実行可能なようにしてもよい。

【符号の説明】

【0082】

10...コンピュータ、10a...マイクロホン、10b...A/D変換部、20...制御部、20a...信号取得部、20b...変換部、20c...出力部、20d...機械学習部、20e...特徴出力部、30...記憶媒体、30a...教師データ、30b...機械学習済モデル、40...表示部

10

20

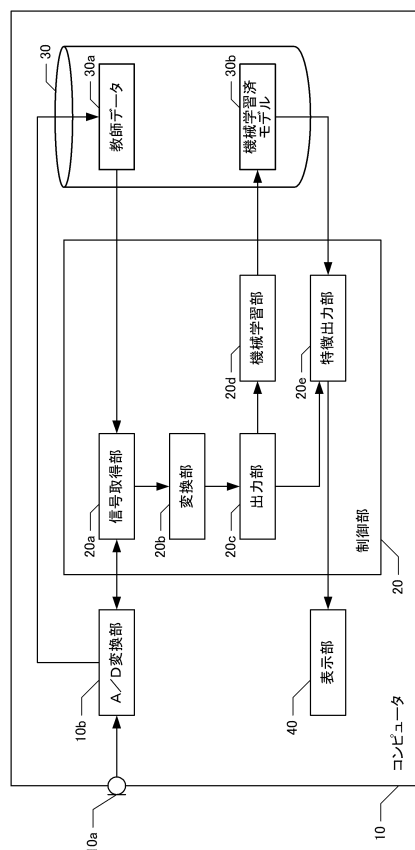
30

40

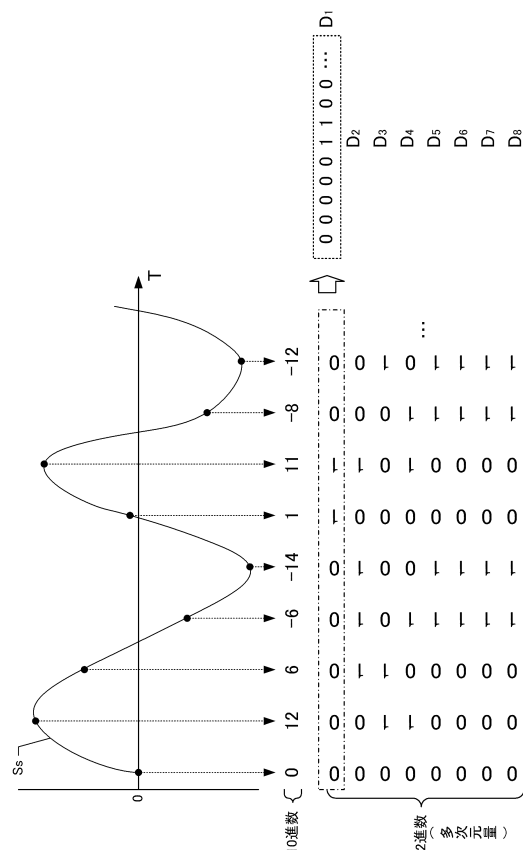
50

【図面】

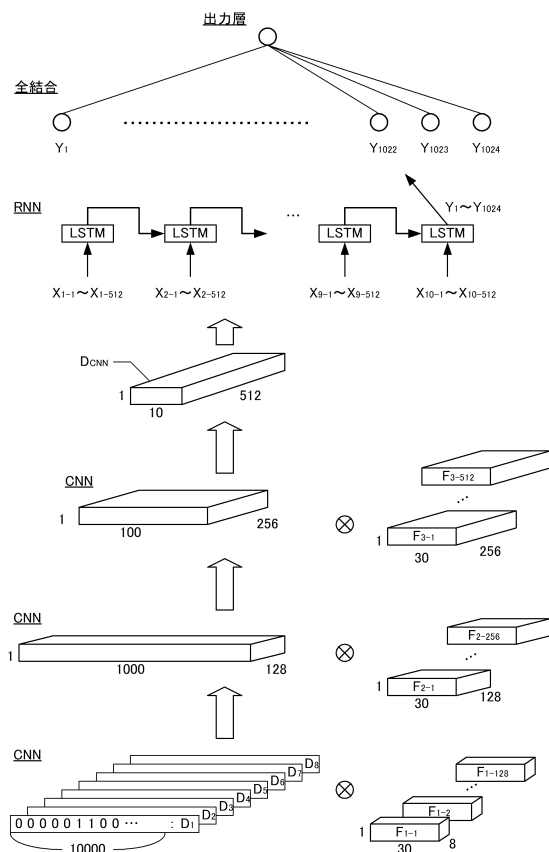
【 図 1 】



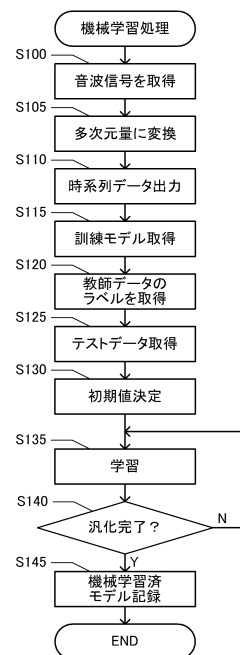
【圖 2】



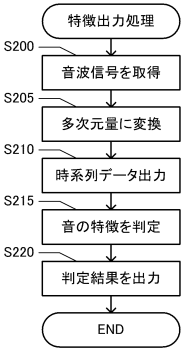
【圖 3】



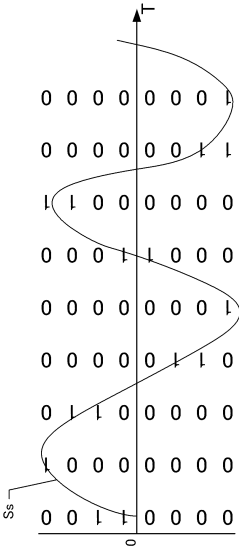
【圖 4】



【図 5】

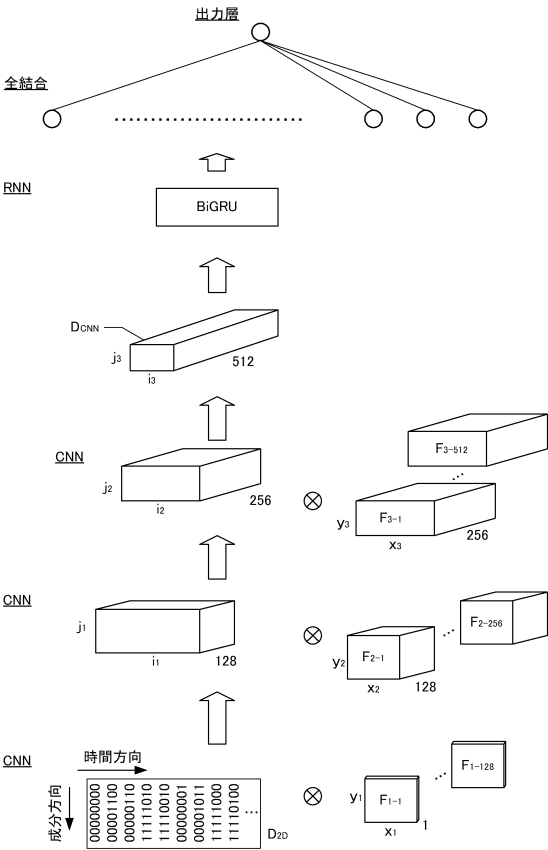


【図 6】



10

【図 7】



20

30

40

50

フロントページの続き

- (56)参考文献 特表 2 0 1 8 - 5 1 7 9 2 8 (J P , A)
特開 2 0 0 3 - 3 3 2 9 1 4 (J P , A)
特許第 6 2 5 9 9 4 6 (J P , B 1)
特開 2 0 1 5 - 9 5 2 1 5 (J P , A)
- (58)調査した分野 (Int.Cl. , D B 名)
G 1 0 L 1 3 / 0 0 - 9 9 / 0 0