

(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
2020年8月13日(13.08.2020)



(10) 国際公開番号

WO 2020/162220 A1

(51) 国際特許分類:

H04N 21/442 (2011.01) G10L 15/00 (2013.01)
H04N 17/00 (2006.01) G10L 15/10 (2006.01)
G06F 16/65 (2019.01)

(21) 国際出願番号: PCT/JP2020/002458

(22) 国際出願日: 2020年1月24日(24.01.2020)

(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 2019-020322 2019年2月7日(07.02.2019) JP

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE

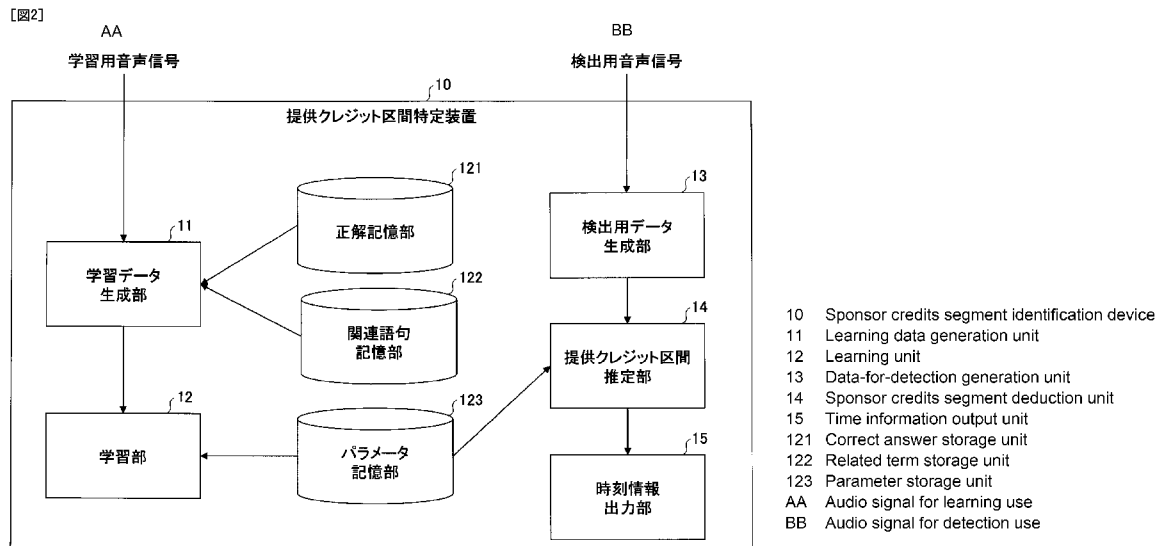
CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 大石 康智(OISHI, Yasunori); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 川西 隆仁 (KAWANISHI, Takahito); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP). 柏野 邦夫(KASHINO, Kunio); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP).

(74) 代理人: 伊東 忠重, 外(ITO, Tadashige et al.); 〒1000005 東京都千代田区丸の内二丁目1番1号 丸の内 M Y P L A Z A (明治安田生命ビル) 16階 Tokyo (JP).

(54) Title: CREDITS SEGMENT IDENTIFICATION DEVICE, CREDITS SEGMENT IDENTIFICATION METHOD, AND PROGRAM

(54) 発明の名称: クレジット区間特定装置、クレジット区間特定方法及びプログラム



(57) Abstract: This credits segment identification device makes identification of a segment including credits more efficient by having: an extraction unit for extracting, from a first audio signal, a plurality of first partial audio signals which each constitute a part of the first audio signal and which are shifted from one another in the time direction; and an identification unit for identifying a segment including credits in the first audio signal by determining, on the basis of an association between respective second partial audio signals extracted from a second audio signal and the presence/absence of credits, whether or not each of the first partial audio signals contains credits.

(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, IR, IS, JO, 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: クレジット区間特定装置は、第1の音声信号から、それぞれが前記第1の音声信号の一部であり、相互に時間方向にずれを有する複数の第1の部分音声信号を抽出する抽出部と、前記各第1の部分音声信号にクレジットが含まれるか否かを、第2の音声信号から抽出される各第2の部分音声信号とクレジットの有無との関連付けに基づいて判定することで、前記第1の音声信号におけるクレジットの区間を特定する特定部と、を有することで、クレジットの区間の特定を効率化する。

明 細 書

発明の名称：

クレジット区間特定装置、クレジット区間特定方法及びプログラム

技術分野

[0001] 本発明は、クレジット区間特定装置、クレジット区間特定方法及びプログラムに関する。

背景技術

[0002] 従来、テレビ放送等について、いずれの企業がいずれの番組のスポンサーであるかを調査することに経済的な価値が認められている。

[0003] このような調査は、テレビ放送等における提供クレジットの表示を目視で見つけ出し、当該提供クレジットから企業名を書き起こすことで行われている。なお、提供クレジットとは、放送番組のスポンサーのロゴの表示やナレーション（例えば、「この番組は、XXXとご覧のスポンサーの提供でお送りします」等）をいう。

先行技術文献

非特許文献

[0004] 非特許文献1：[online]、インターネット<URL：<http://www.jppanet.or.jp/documents/video.html>>

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、提供クレジットが表示される区間は、放送時間の約1%程度に過ぎない。したがって、上記のような調査においては提供クレジットの区間を特定するためのテレビ放送等の視聴作業に多くの時間が費やされてしまう。

[0006] なお、上記では、説明の便宜上、提供クレジットを例として記載したが、例えば、特定のコマーシャル等、提供クレジットだけでなく、他のクレジットの区間を特定したい場合にも同様の課題が生じる。

[0007] 本発明は、上記の点に鑑みてなされたものであって、クレジットの区間の特定を効率化することを目的とする。

課題を解決するための手段

[0008] そこで上記課題を解決するため、クレジット区間特定装置は、第1の音声信号から、それぞれが前記第1の音声信号の一部であり、相互に時間方向にずれを有する複数の第1の部分音声信号を抽出する抽出部と、前記各第1の部分音声信号にクレジットが含まれるか否かを、第2の音声信号から抽出される各第2の部分音声信号とクレジットの有無との関連付けに基づいて判定することで、前記第1の音声信号におけるクレジットの区間を特定する特定部と、を有する。

発明の効果

[0009] クレジットの区間の特定を効率化することができる。

図面の簡単な説明

[0010] [図1]第1の実施の形態における提供クレジット区間特定装置10のハードウェア構成例を示す図である。

[図2]第1の実施の形態における提供クレジット区間特定装置10の機能構成例を示す図である。

[図3]第1の実施の形態における学習処理の処理手順の一例を説明するためのフローチャートである。

[図4]第1の実施の形態における正例の音声セグメントの抽出例を示す図である。

[図5]第1の実施の形態における識別器のモデル例を示す図である。

[図6]第1の実施の形態における提供クレジットの検出処理の処理手順の一例を説明するためのフローチャートである。

[図7]第1の実施の形態における検出用音声信号からの音声セグメントの抽出例を示す図である。

[図8]第2の実施の形態における提供クレジット区間特定装置10の機能構成例を示す図である。

[図9]第2の実施の形態における学習処理の処理手順の一例を説明するためのフローチャートである。

[図10]第2の実施の形態における正例の音声セグメント及び静止画のペアの抽出例を示す図である。

[図11]第2の実施の形態における識別器のモデル例を示す図である。

[図12]第2の実施の形態における提供クレジットの検出処理の処理手順の一例を説明するためのフローチャートである。

[図13]第2の実施の形態における検出用音声信号からの音声セグメントの抽出例を示す図である。

[図14]各実施形態の評価結果の一例を示す図である。

発明を実施するための形態

[0011] 以下、図面に基づいて本発明の実施の形態を説明する。図1は、第1の実施の形態における提供クレジット区間特定装置10のハードウェア構成例を示す図である。図1の提供クレジット区間特定装置10は、それぞれバスBで相互に接続されているドライブ装置100、補助記憶装置102、メモリ装置103、CPU104、及びインタフェース装置105等を有するコンピュータである。

[0012] 提供クレジット区間特定装置10での処理を実現するプログラムは、CD-ROM等の記録媒体101によって提供される。プログラムを記憶した記録媒体101がドライブ装置100にセットされると、プログラムが記録媒体101からドライブ装置100を介して補助記憶装置102にインストールされる。但し、プログラムのインストールは必ずしも記録媒体101より行う必要はなく、ネットワークを介して他のコンピュータよりダウンロードするようにしてもよい。補助記憶装置102は、インストールされたプログラムを格納すると共に、必要なファイルやデータ等を格納する。

[0013] メモリ装置103は、プログラムの起動指示があった場合に、補助記憶装置102からプログラムを読み出して格納する。CPU104は、メモリ装置103に格納されたプログラムに従って提供クレジット区間特定装置10

に係る機能を実行する。インタフェース装置１０５は、ネットワークに接続するためのインタフェースとして用いられる。

[0014] 図２は、第１の実施の形態における提供クレジット区間特定装置１０の機能構成例を示す図である。図２において、提供クレジット区間特定装置１０は、学習データ生成部１１、学習部１２、検出用データ生成部１３、提供クレジット区間推定部１４及び時刻情報出力部１５等を有する。これら各部は、提供クレジット区間特定装置１０にインストールされた１以上のプログラムが、ＣＰＵ１０４に実行させる処理により実現される。提供クレジット区間特定装置１０は、また、正解記憶部１２１、関連語句記憶部１２２及びパラメータ記憶部１２３等を利用する。これら各記憶部は、例えば、補助記憶装置１０２、又は提供クレジット区間特定装置１０にネットワークを介して接続可能な記憶装置等を用いて実現可能である。

[0015] 正解記憶部１２１には、或る期間に放送された学習用のＴＶ放送（以下、「学習用ＴＶ放送」という。）の音声信号（以下、「学習用音声信号」という。）について、提供クレジットの区間（以下、「提供クレジット区間」という。）を示す時刻データ（開始時刻、終了時刻）が記憶されている。なお、提供クレジット区間は、例えば、予めユーザによる目視等によって確認されてもよい。

[0016] 関連語句記憶部１２２には、提供クレジットの表示時のアナウンス（提供クレジット表示の際に流れるアナウンス）に含まれ、提供クレジット表示に関連する関連語句が記憶されている。関連語句の一例として、「ご覧の」、「スポンサー」、「提供」、「お送り（お送りします／お送りしました）」等の語句が挙げられる。また、企業名を示す語句等が関連語句とされてもよい。なお、関連語句は、例えば、予めユーザにより設定される。

[0017] パラメータ記憶部１２３には、音声信号における提供クレジットの有無を識別する識別器のパラメータが記憶される。識別器は、学習用音声信号から抽出される複数の音声信号（後述の「音声セグメント」）と提供クレジット有無との関連付けを学習したモデルである。

[0018] 以下、提供クレジット区間特定装置 10 が実行する処理手順について説明する。図 3 は、第 1 の実施の形態における学習処理の処理手順の一例を説明するためのフローチャートである。

[0019] ステップ S 101 において、学習データ生成部 11 は、学習用音声信号から、正例の音声セグメント（学習用音声信号において提供クレジットを含むと推定される部分（部分音声信号））を抽出する。

[0020] 具体的には、学習データ生成部 11 は、正解記憶部 121 に記憶されている時刻データに基づいて、学習用音声信号における提供クレジット区間を特定する。なお、提供クレジット区間は、複数有ってもよい。学習データ生成部 11 は、学習用音声信号のうち、特定した各提供クレジット区間を対象として音声認識を実行し、提供クレジット区間ごとに音声認識結果（テキストデータ）を生成する。学習データ生成部 11 は、各テキストデータについて、関連語句記憶部 122 に記憶されているいずれかの関連語句を含む部分を特定し、学習用音声信号において当該部分に対応する音声信号を正例の音声セグメントとして抽出する。例えば、関連語句を中心とした前後 N 秒間の部分が正例の音声セグメントとして抽出される。本実施の形態では、 $N = 3$ とする。但し、N は、他の値であってもよい。

[0021] 図 4 は、第 1 の実施の形態における正例の音声セグメントの抽出例を示す図である。図 4 では、学習用音声信号のうち、「ご覧のスポンサーの提供でお送りしました」の部分が提供クレジット区間に対応し、このうち「ご覧」、「スポンサー」、「提供」、「送り」が関連語句である例が示されている。したがって、これらの関連語句を中心とした前後 3 秒間の音声信号が正例の音声セグメントとして抽出されている。

[0022] 続いて、学習データ生成部 11 は、学習用音声信号における提供クレジット区間以外のランダムな部分から、負例の音声セグメントを抽出する（S 102）。負例の音声セグメントの長さは正例の音声セグメントの長さ（6 秒間）と同じである。また、負例の音声セグメントの個数は、正例の音声セグメントの個数と同数であるのが望ましい。

[0023] 続いて、学習部12は、ステップS101において抽出された正例の音声セグメントと、ステップS102において抽出された負例の音声セグメントとを用いて、提供クレジット区間に関する識別器の学習を行う（S103）。

[0024] 具体的には、学習部12は、正例又は負例の各音声セグメントを周波数分析し（例えば、窓長25ms、窓シフト長10ms）、40個のメルフィルタバンク処理を施すことで、600×40のメルスペクトログラムを取得する。学習部12は、音声セグメントごとに、当該音声セグメントに関して取得されたメルスペクトログラムを入力特徴量として、当該音声セグメントに提供クレジットが有るか無いか（当該音声セグメントに提供クレジットが含まれるか否か）を2クラス識別（検出）する識別器を学習する。すなわち、正例の音声セグメントについては、提供クレジットが有ることが学習され、負例の音声セグメントについては、提供クレジットが無いことが学習される。識別器としては、例えば、畳み込みニューラルネットワークが利用されてもよいし、SVM（support vector machine）などの他の識別器が利用されてもよい。

[0025] 図5は、第1の実施の形態における識別器のモデル例を示す図である。図5には、畳み込みニューラルネットワークを利用した例が示されている。

[0026] 続いて、学習部12は、学習された識別器のパラメータをパラメータ記憶部123に記憶する（S104）。

[0027] 図6は、第1の実施の形態における提供クレジットの検出処理の処理手順の一例を説明するためのフローチャートである。図6の処理手順は、図3の処理手順が実行済みであることが前提となる。

[0028] ステップS201において、検出用データ生成部13は、提供クレジットの検出用のTV放送（以下、「検出用TV放送」という。）の音声信号（以下、「検出用音声信号」という。）から、窓長2N秒、窓シフト長1秒で音声セグメントを抽出する。本実施の形態においてN=3であるため、1秒ずつずれた（相互に時間方向にずれを有する）6秒間の複数の音声セグメント

が抽出される。

[0029] 図7は、第1の実施の形態における検出用音声信号からの音声セグメントの抽出例を示す図である。図7では、1秒ずつずれを有する6秒間の音声信号が音声セグメントとして抽出される例が示されている。なお、図7では、便宜上、検出用音声信号の途中までの音声セグメントの抽出例が示されているが、検出用音声信号の全部について、音声セグメントの抽出が行われる。

[0030] 続いて、提供クレジット区間推定部14は、ステップS201において抽出された各音声セグメントを周波数分析し（例えば、窓長25ms、窓シフト長10ms）、40個のメルフィルタバンク処理を施すことで、600×40のメルスペクトログラムを各音声セグメントの特徴量として取得する（S202）。

[0031] 続いて、提供クレジット区間推定部14は、パラメータ記憶部123に記憶されているパラメータを用いて、図3の処理手順によって学習された識別器を復元（生成）する（S203）。

[0032] 続いて、提供クレジット区間推定部14は、ステップS201において抽出された音声セグメントごとに、ステップS202において取得された特徴量を当該識別器に入力して、各音声セグメントにおける提供クレジットの有無（各音声セグメントに提供クレジットが含まれるか否か）を判定する（S204）。例えば、提供クレジット区間推定部14は、識別器の出力値が所定の閾値以上である音声セグメントについては提供クレジット有り「1」と判定し、当該出力値が閾値よりも小さい音声セグメントについては提供クレジット無し「0」と判定する。提供クレジット区間推定部14は、判定結果を音声セグメントの時系列順に配列することで、提供クレジットの有無を時系列的に示すバイナリ時系列信号を生成する。

[0033] 続いて、提供クレジット区間推定部14は、当該バイナリ時系列信号において、提供クレジット表示ありと判定された音声セグメントが所定時間以上連続する区間を、提供クレジットが表示された提供クレジット表示区間として検出（特定）する（S205）。具体的には、提供クレジット区間推定部

14は、ノイズ除去を目的として、バイナリ時系列信号に対して中央値フィルタを適用する。提供クレジット区間推定部14は、中央値フィルタ処理後の時系列信号において、提供クレジット表示有りと判定された音声セグメントが所定時間以上連続する区間（信号「1」が所定時間以上（例えば、音声セグメントの長さ（6秒）×M以上（ $M \geq 2$ ））連続して並ぶ区間）を、提供クレジット表示区間として検出（特定）する。本実施の形態のように、音声セグメントが1秒間隔で（すなわち、1秒のずれを有するように）作成された場合、例えば、300番目から310番目に信号「1」が連続して並んでいれば、提供クレジット区間推定部14は、5分00秒から5分10秒の区間を提供クレジット表示区間として検出（特定）する。

[0034] 続いて、時刻情報出力部15は、検出され提供クレジット表示区間の時刻情報（開始時刻及び終了時刻）を出力する（S206）。

[0035] なお、上記では、TV放送の音声信号を例として説明したが、例えば、ラジオ放送の音声信号における提供クレジットの区間の特定について第1の実施の形態が特定されてもよい。また、特定のコマーシャル（CM）等、提供クレジットだけでなく、他のクレジットの区間の特定について第1の実施の形態が適用されてもよい。この場合、特定のCMに含まれている語句が、関連語句として関連語句記憶部122に記憶されればよい。

[0036] 上述したように、第1の実施の形態によれば、クレジットの区間の特定を効率化することができる。

[0037] 次に、第2の実施の形態について説明する。第2の実施の形態では第1の実施の形態と異なる点について説明する。第2の実施の形態において特に言及されない点については、第1の実施の形態と同様でもよい。

[0038] 図8は、第2の実施の形態における提供クレジット区間特定装置10の機能構成例を示す図である。図8において、図2と同一部分又は対応する部分には同一符号を付し、その説明は適宜省略する。

[0039] 正解記憶部121には、学習用TV放送の映像信号（すなわち、学習用音声信号に対応する（同期した）映像信号。以下、「学習用映像信号」という

）及び音声信号（学習用音声信号）に対して、提供クレジット区間の時刻データ（開始時刻、終了時刻）が記憶されている。

[0040] パラメータ記憶部 123 には、映像信号及び音声信号のペアについて、提供クレジットの有無を識別する識別器のパラメータが記憶される。

[0041] 図 9 は、第 2 の実施の形態における学習処理の処理手順の一例を説明するためのフローチャートである。

[0042] ステップ S101a において、学習データ生成部 11 は、正例の音声セグメント（学習用音声信号において提供クレジットを含む部分）を学習用音声信号から抽出すると共に、当該音声セグメントにおいて関連語句の時刻に対応する静止画を学習用映像信号から抽出する。したがって、正例の音声セグメントと静止画のペアが抽出される。正例の音声セグメントの抽出方法は第 1 の実施の形態と同様でよい。正例の静止画としては、学習用映像信号において、正例の音声セグメントにおける関連語句の時刻のフレーム（静止画）が抽出されればよい。なお、1 つの音声セグメントに対して複数のフレーム（静止画）が抽出されてもよい。

[0043] 図 10 は、第 2 の実施の形態における正例の音声セグメント及び静止画のペアの抽出例を示す図である。図 10 における学習用音声信号は、図 4 における学習用音声信号と同じである。したがって、図 10 では、図 4 と同じ音声セグメントが抽出されている。但し、図 10 では、各音声セグメントにおいて関連語句の出現する時刻における静止画が学習用映像信号から抽出されている。なお、図 10 において、各音声セグメントと静止画との位置関係は、当該音声セグメントに対する当該静止画のタイミングとは無関係である。

[0044] 続いて、学習データ生成部 11 は、学習用音声信号における提供クレジット区間以外の部分から負例の音声セグメントを抽出し、学習用映像信号において当該音声セグメントの中心時刻に対応する静止画を負例の静止画として抽出する（S102a）。したがって、負例の音声セグメントと静止画とのペアが抽出される。なお、負例の音声セグメントの抽出方法は、第 1 の実施の形態と同様でよい。

[0045] 続いて、学習部12は、ステップS101aにおいて抽出された正例の音声セグメント及び静止画のペアと、ステップS102aにおいて抽出された負例の音声セグメント及び静止画のペアとを用いて、提供クレジットに関する識別器（これら各ペアと提供クレジットの有無との関連付け）の学習を行う（S103a）。

[0046] 具体的には、学習部12は、正例又は負例の各音声セグメントを周波数分析し（例えば、窓長25ms、窓シフト長10ms）、40個のメルフィルタバンク処理を施すことで、600×40のメルスペクトログラムを取得する。学習部12は、音声セグメントごとに、当該音声セグメントに関して取得されたメルスペクトログラムと、当該音声セグメントに対応する静止画とのペアを入力特徴量として、当該ペアに提供クレジットが有るか無いか（当該ペアに提供クレジットが含まれているか否か）を2クラス識別（検出）する識別器を学習する。識別器としては、例えば、畳み込みニューラルネットワークが利用されてもよいし、SVMなどの他の識別器が利用されてもよい。

[0047] 図11は、第2の実施の形態における識別器のモデル例を示す図である。図11には、畳み込みニューラルネットワークを利用した例が示されている。

[0048] 続いて、学習部12は、学習された識別器のパラメータをパラメータ記憶部123に記憶する（S104a）。

[0049] 図12は、第2の実施の形態における提供クレジットの検出処理の処理手順の一例を説明するためのフローチャートである。図12中、図6と同一ステップには同一ステップ番号を付し、その説明は適宜省略する。図12の処理手順は、図9の処理手順が実行済みであることが前提となる。

[0050] ステップS201aにおいて、検出用データ生成部13は、窓長2N秒、窓シフト長1秒で音声セグメントを検出用音声信号から抽出すると共に、各音声セグメントの中心時刻（3秒目）の静止画を、検出用TV放送の映像信号（すなわち、検出用音声信号に対応する（同期した）映像信号）から抽出

する。

[0051] 図13は、第2の実施の形態における検出用音声信号からの音声セグメント及び静止画の抽出例を示す図である。図13では、1秒ずつずれを有する6秒間の音声信号が音声セグメントとして抽出され、各音声セグメントの中心時刻における静止画が検出用映像信号から抽出される例が示されている。

[0052] 続いて、第1の実施の形態と同様に、各音声セグメントの特徴量（ 600×40 のメルスペクトログラム）が取得される（S202）。

[0053] 続いて、提供クレジット区間推定部14は、パラメータ記憶部123に記憶されているパラメータを用いて、図9の処理手順によって学習された識別器を復元（生成）する（S203a）。

[0054] 続いて、提供クレジット区間推定部14は、ステップS201aにおいて抽出された音声セグメント及び静止画のペアごとに、当該音声セグメントからステップS202において取得された特徴量と当該静止画とのペアを当該識別器に入力して、各ペアにおける提供クレジットの有無を判定する（S204a）。なお、提供クレジットの有無の判定方法は、第1の実施の形態と同様でよい。その結果、提供クレジットの有無を時系列的に示すバイナリ時系列信号が生成される。

[0055] 以降（S205、S205）は、第1の実施の形態と同様でよい。

[0056] 図14は、各実施形態の評価結果の一例を示す図である。図14には、地上波5局の1週間分の放送について学習し、別の1週間における地上波5局の放送について提供クレジットの区間を特定した際の評価結果（再現率）が示されている。ここで、再現率とは、正解の区間（提供クレジットが実際に表示された区間）に対して、提供クレジット区間特定装置10が、提供クレジットの区間であると判定した区間の割合をいう。例えば、放送の開始から11秒目から20秒目10秒間が正解の区間である場合に、12秒目から20秒目の9秒間が提供クレジットの区間として特定された場合には、再現率は $9 \div 10 = 0.9$ となる。

[0057] また、図14の横軸において「音声」は、音声信号のみを利用した場合、

すなわち、第１の実施の形態に対応し、「画像＋音声」は、音声信号と映像信号を利用した場合、すなわち、第２の実施の形態に対応する。

[0058] 図１４によれば、「音声」及び「画像＋音声」のいずれについても高い再現率が得られている。また、「音声」の場合よりも「画像＋音声」の方が、高い再現率が得られていることが分かる。このことから、第２の実施の形態によれば、第１の実施の形態よりも高精度に提供クレジットの区間を特定できることが分かる。

[0059] なお、上記各実施の形態は、インターネット等において配信される動画におけるクレジットの区間の特定に適用されてもよい。

[0060] なお、上記各実施の形態において、提供クレジット区間特定装置１０は、クレジット区間特定装置の一例である。検出用データ生成部１３は、抽出部の一例である。提供クレジット区間推定部１４は、特定部の一例である。検出用音声信号は、第１の音声信号の一例である。検出用音声信号から抽出される音声セグメントは、第１の部分音声信号の一例である。学習用音声信号は、第２の音声信号の一例である。学習用音声信号から抽出される音声セグメントは、第２の部分音声信号の一例である。検出用映像信号は、第１の映像信号の一例である。検出用映像信号から抽出される静止画は、第１の静止画の一例である。学習用映像信号は、第２の映像信号の一例である。学習用映像信号から抽出される静止画は、第２の静止画の一例である。

[0061] 以上、本発明の実施の形態について詳述したが、本発明は斯かる特定の実施形態に限定されるものではなく、請求の範囲に記載された本発明の要旨の範囲内において、種々の変形・変更が可能である。

符号の説明

[0062]	１０	提供クレジット区間特定装置
	１１	学習データ生成部
	１２	学習部
	１３	検出用データ生成部
	１４	提供クレジット区間推定部

1 5	時刻情報出力部
1 0 0	ドライブ装置
1 0 1	記録媒体
1 0 2	補助記憶装置
1 0 3	メモリ装置
1 0 4	C P U
1 0 5	インタフェース装置
1 2 1	正解記憶部
1 2 2	関連語句記憶部
1 2 3	パラメータ記憶部
B	バス

請求の範囲

- [請求項1] 第1の音声信号から、それぞれが前記第1の音声信号の一部であり、相互に時間方向にずれを有する複数の第1の部分音声信号を抽出する抽出部と、
- 前記各第1の部分音声信号にクレジットが含まれるか否かを、第2の音声信号から抽出される各第2の部分音声信号とクレジットの有無との関連付けに基づいて判定することで、前記第1の音声信号におけるクレジットの区間を特定する特定部と、
- を有することを特徴とするクレジット区間特定装置。
- [請求項2] 前記第2の部分音声信号は、予め設定された語句を含む音声信号であり、
- 前記第2の部分音声信号が前記語句を含むか否かは、当該第2の部分音声信号を対象とした音声認識に基づき判定される、
- ことを特徴とする請求項1記載のクレジット区間特定装置。
- [請求項3] 前記特定部は、前記各第2の部分音声信号とクレジットの有無とを学習した識別器を用いて、前記各第1の部分音声信号にクレジットが含まれるか否かを判定する、
- ことを特徴とする請求項1又は2記載のクレジット区間特定装置。
- [請求項4] 前記抽出部は、前記第1の音声信号に対応する第1の映像信号から、前記各第1の部分音声信号に対応する複数の第1の静止画を抽出し、
- 前記特定部は、前記第1の部分音声信号及び前記第1の静止画の各ペアにクレジットが含まれるか否かを、前記各第2の部分音声信号と、前記第2の音声信号に対応する第2の映像信号から抽出される、前記各第2の部分音声信号に対応する第2の静止画とクレジットの有無との関連付けに基づいて判定することで、前記第1の音声信号及び前記第1の映像信号におけるクレジットの区間を特定する、
- ことを特徴とする請求項1乃至3いずれか一項記載のクレジット区間

特定装置。

[請求項5] 第1の音声信号から、それぞれが前記第1の音声信号の一部であり、相互に時間方向にずれを有する複数の第1の部分音声信号を抽出する抽出手順と、

前記各第1の部分音声信号にクレジットが含まれるか否かを、第2の音声信号から抽出される各第2の部分音声信号とクレジットの有無との関連付けに基づいて判定することで、前記第1の音声信号におけるクレジットの区間を特定する特定手順と、

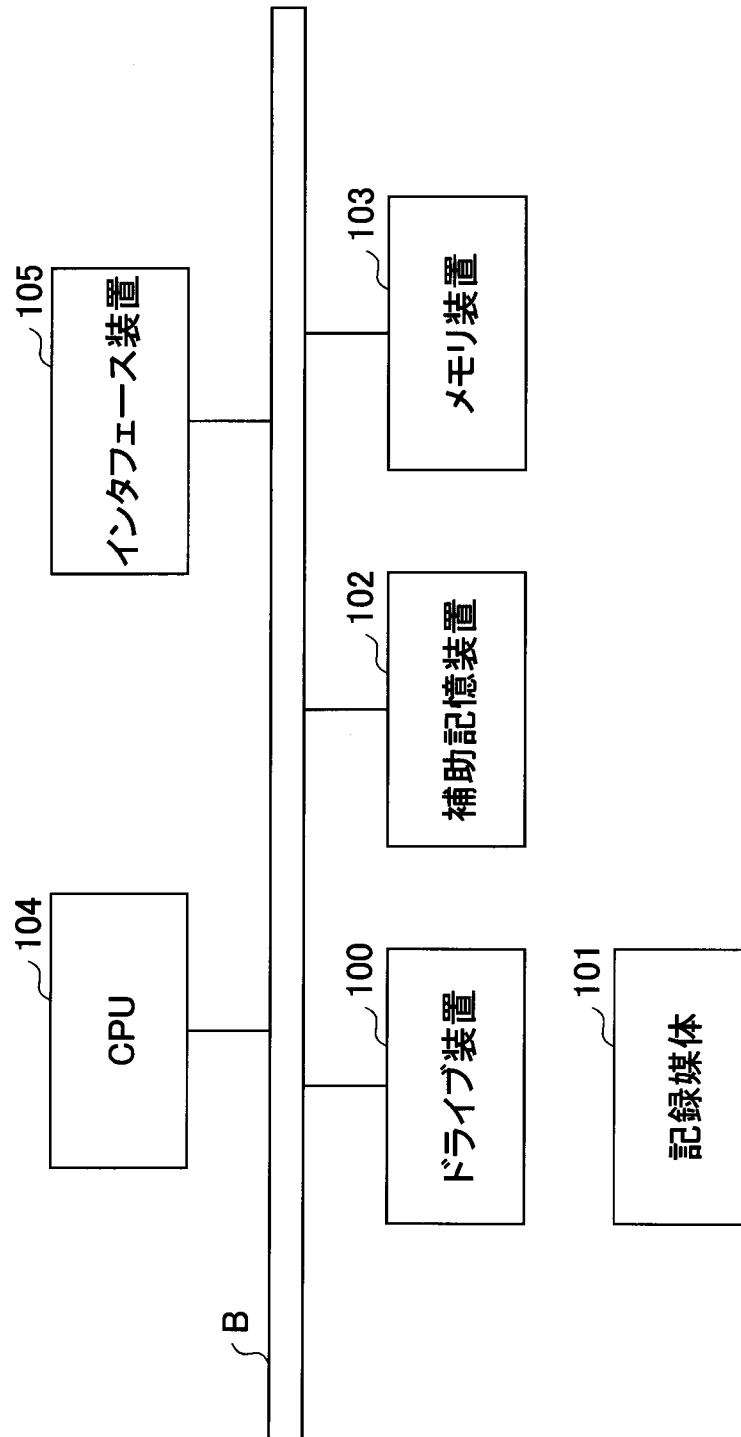
をコンピュータが実行することを特徴とするクレジット区間特定方法

。

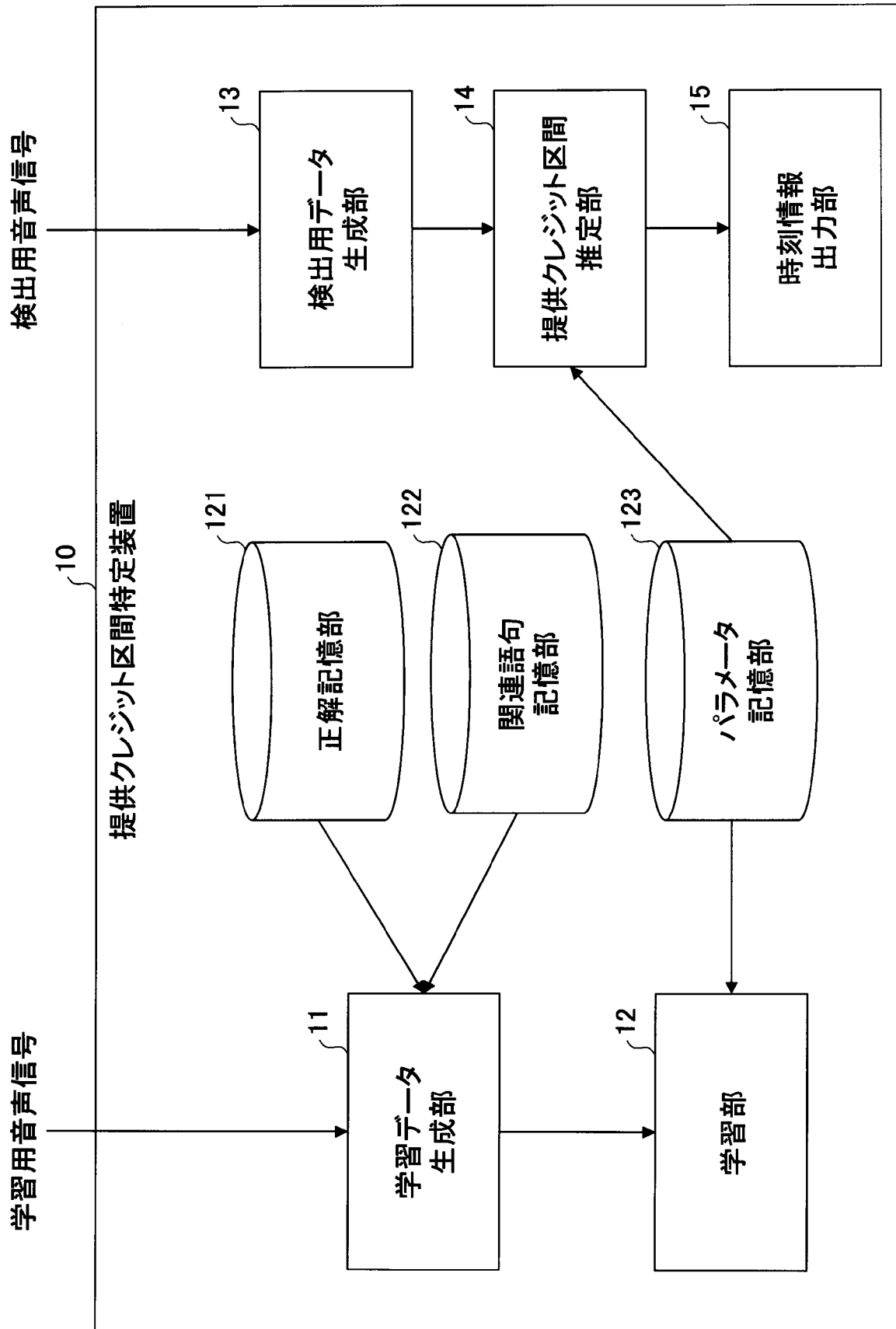
[請求項6] 請求項1乃至4いずれか一項記載のクレジット区間特定装置としてコンピュータを機能させることを特徴とするプログラム。

[図1]

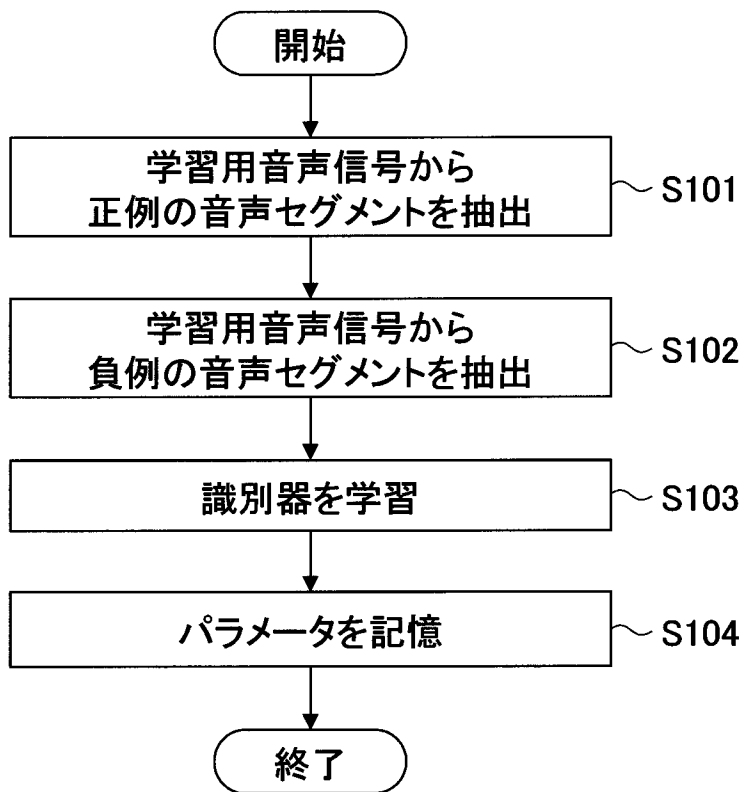
10



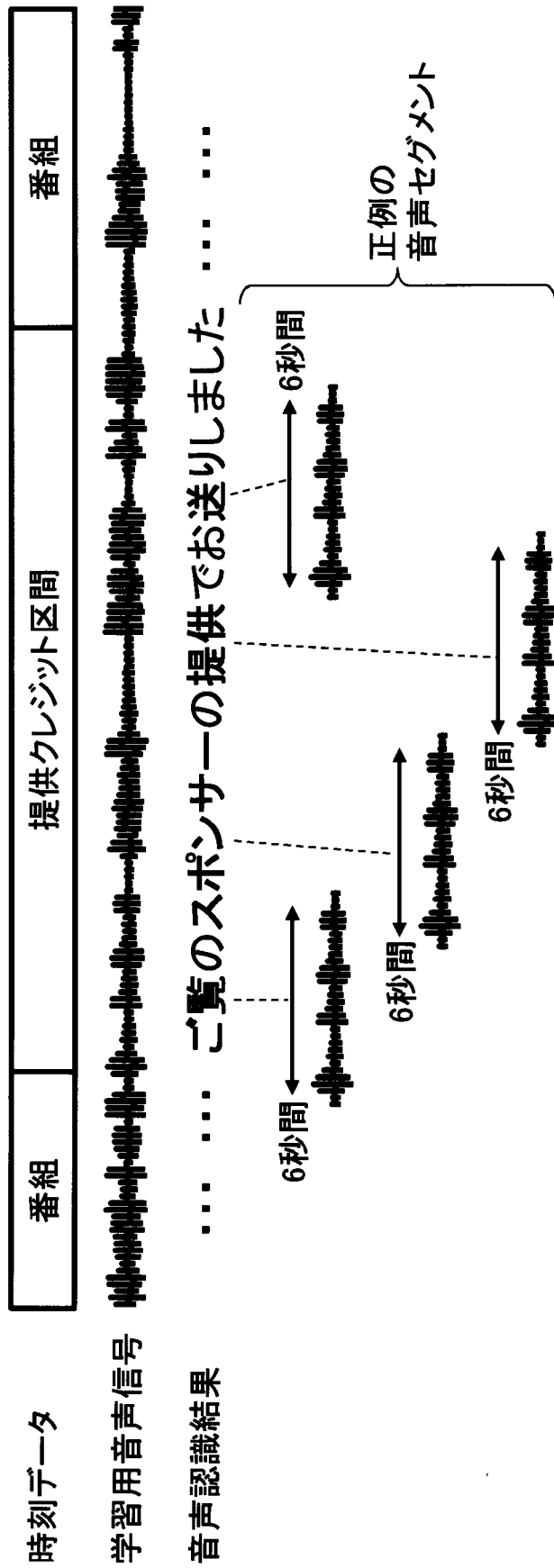
[図2]



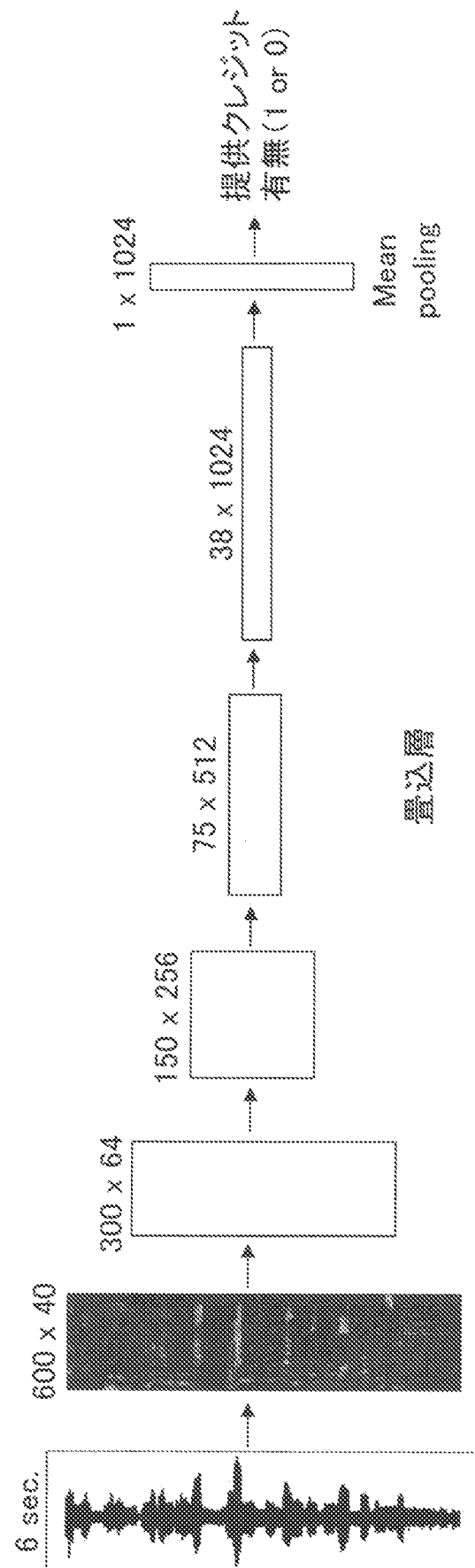
[図3]



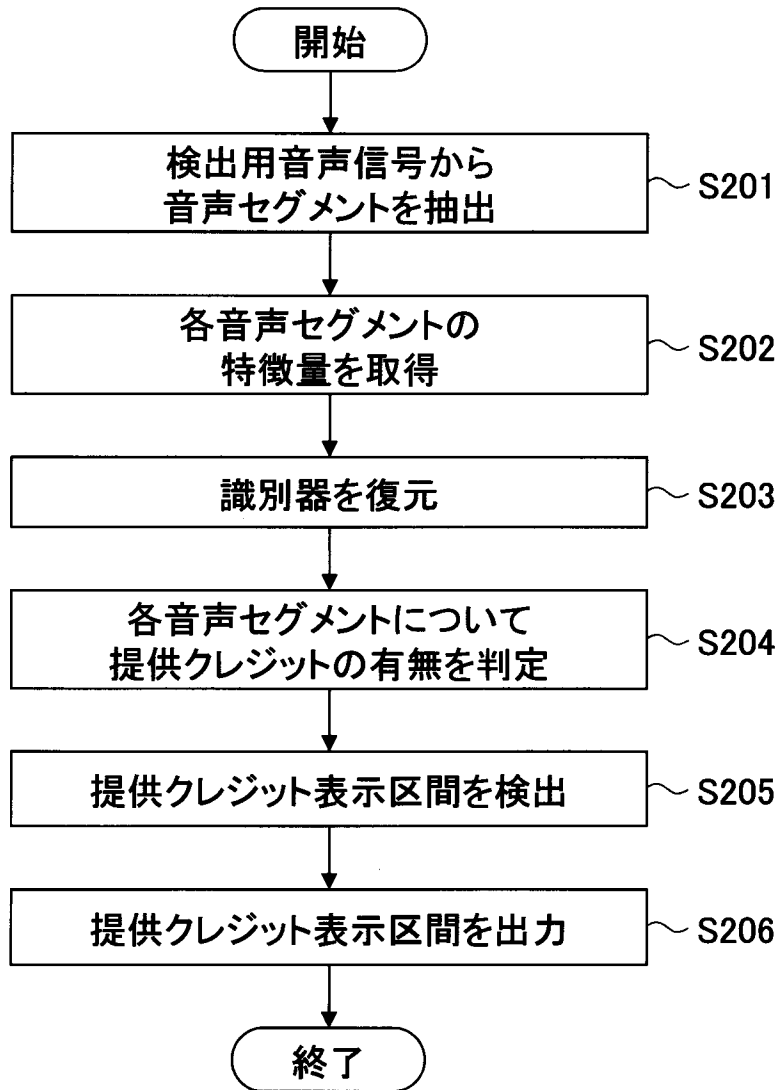
[図4]



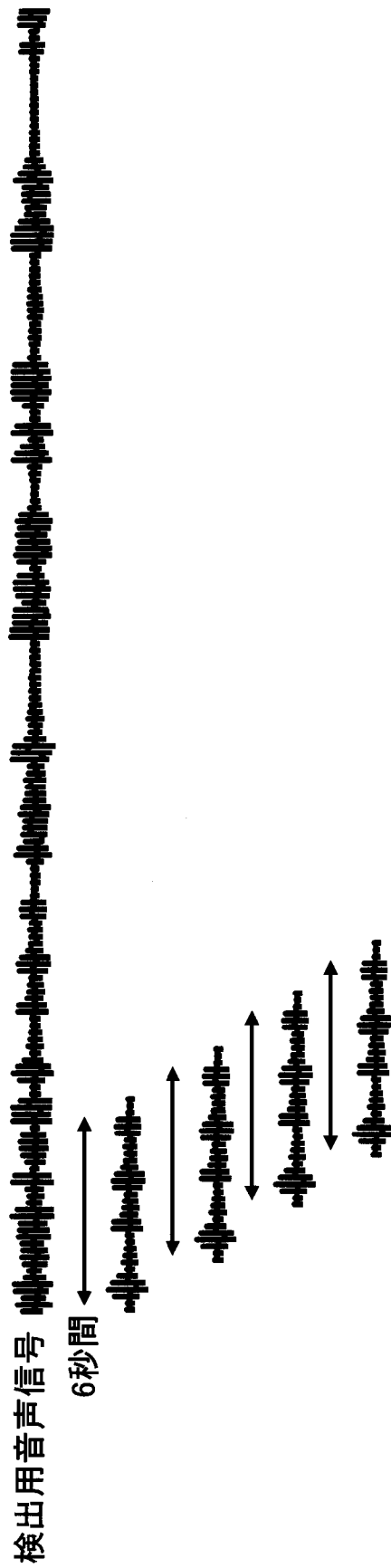
[図5]



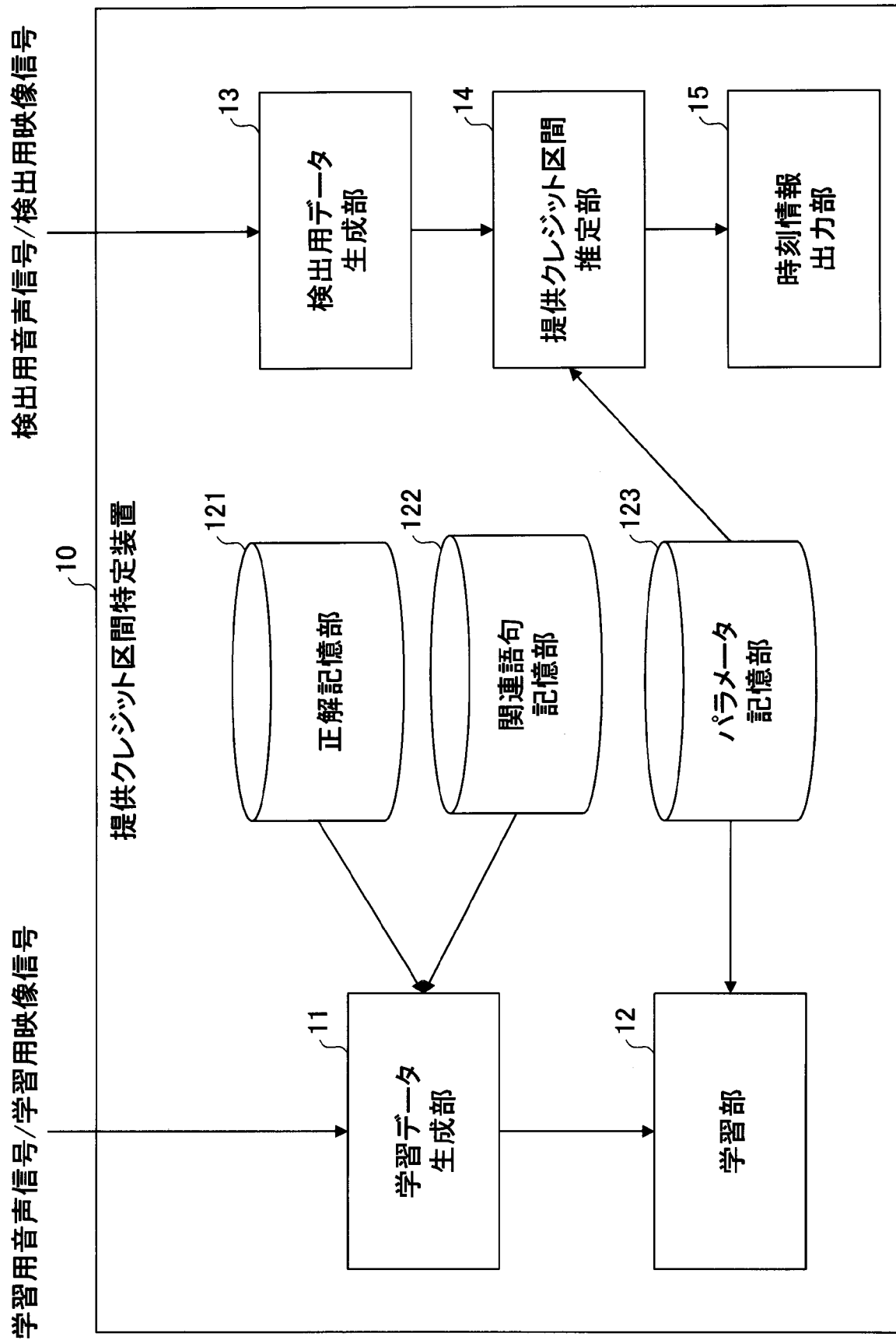
[図6]



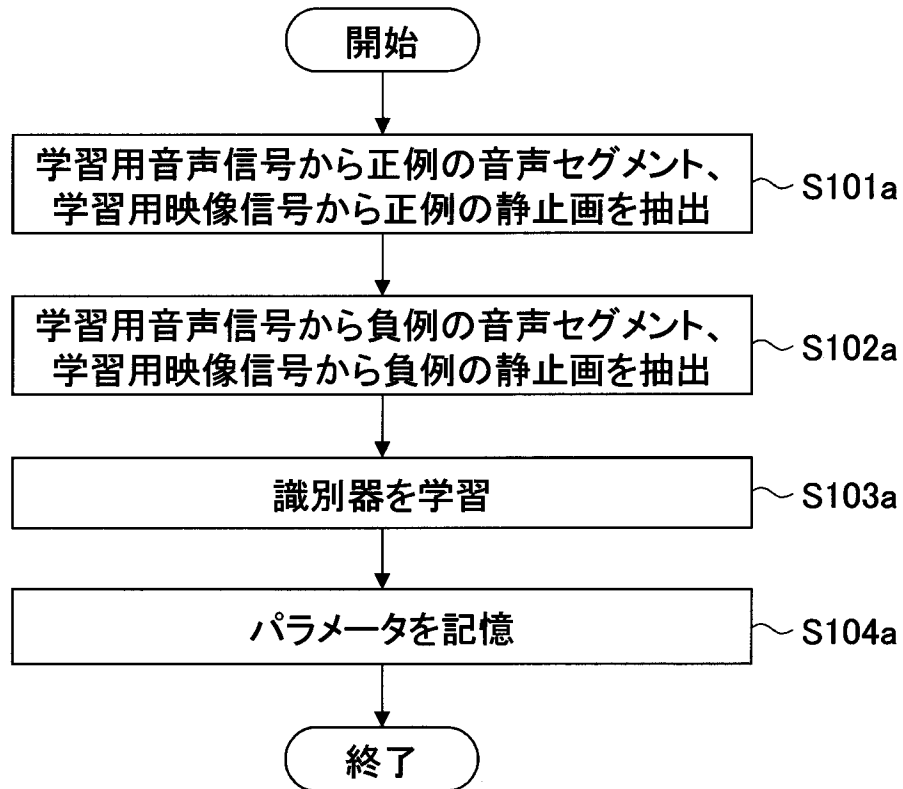
[図7]



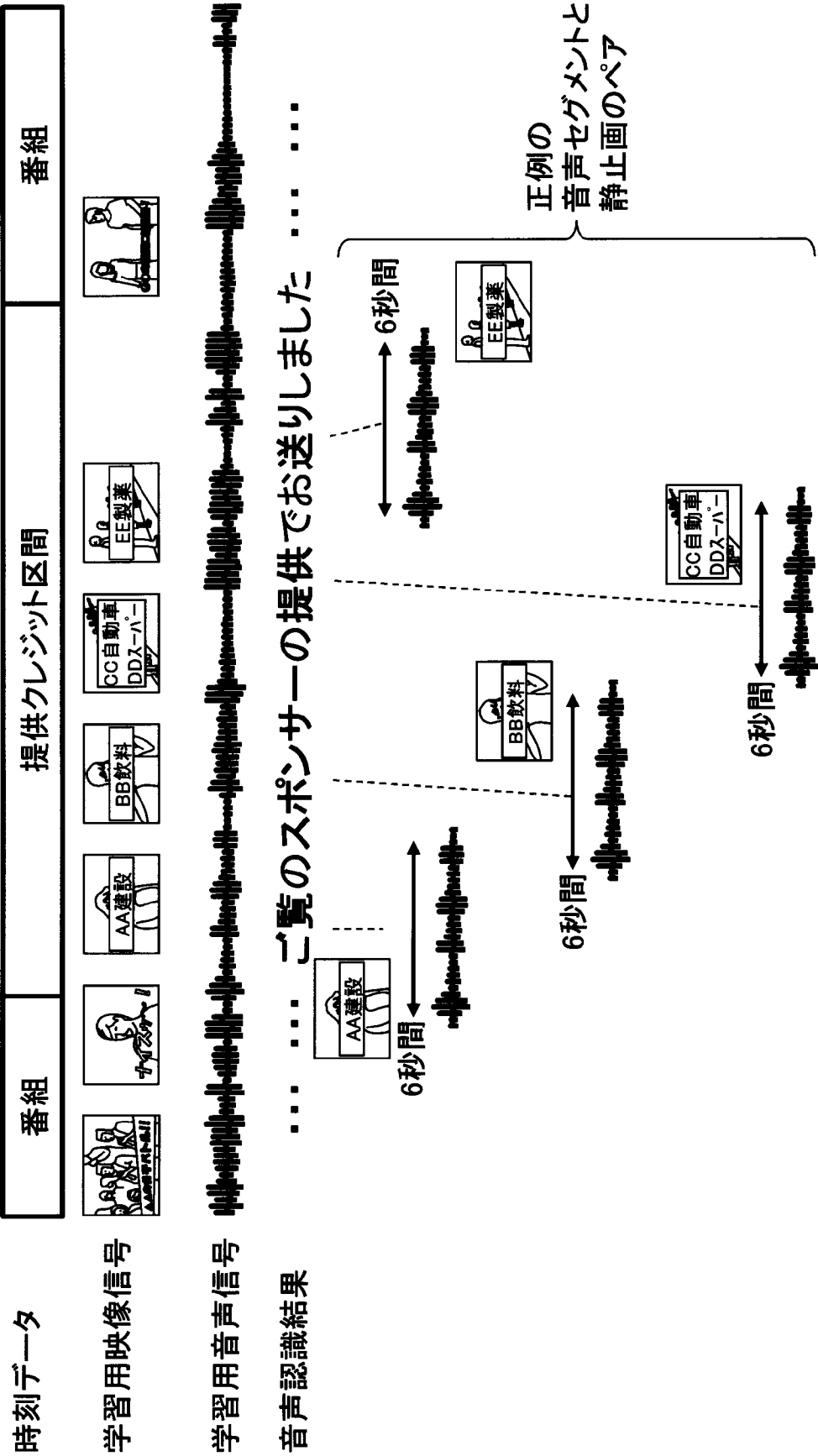
[図8]



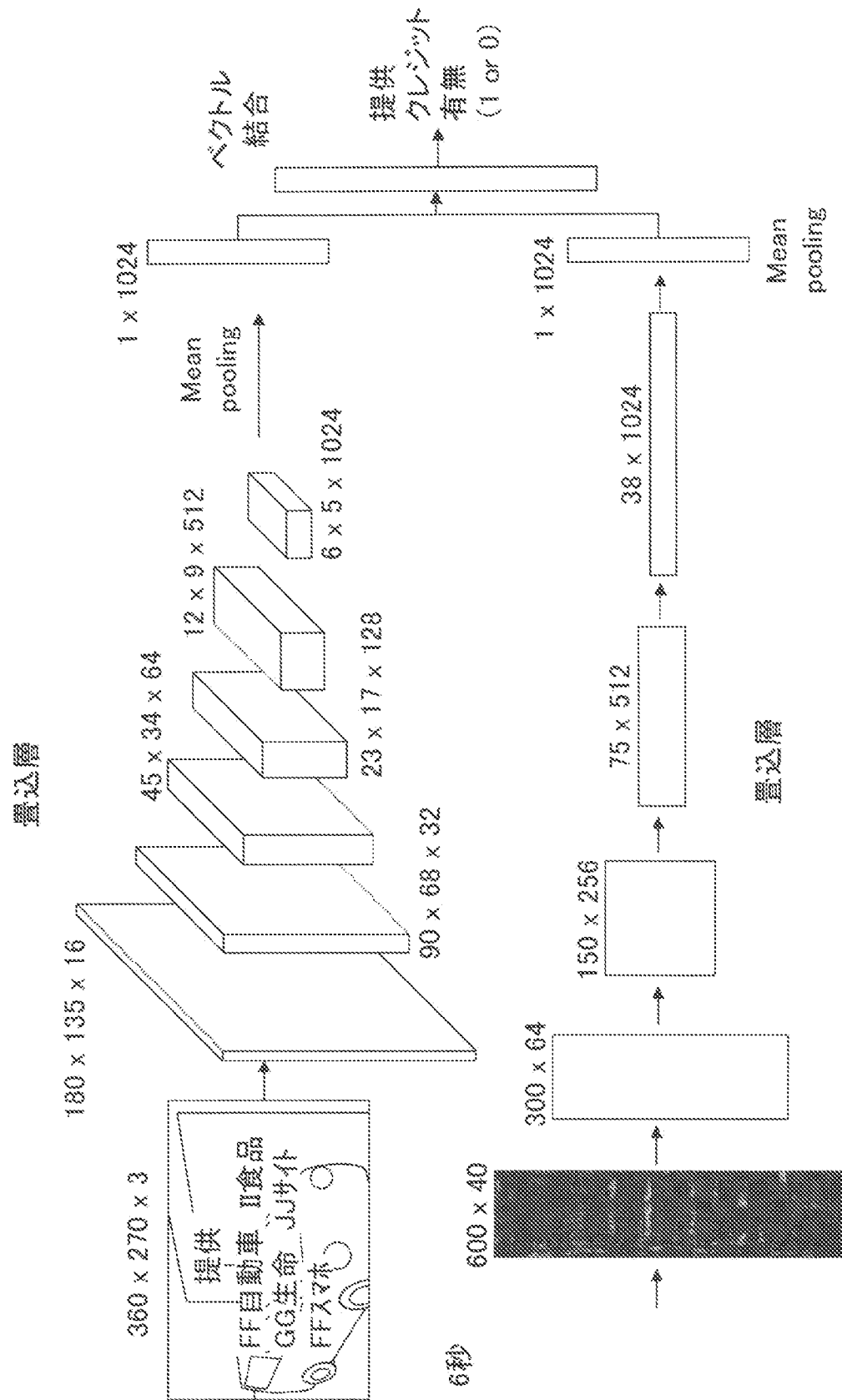
[図9]



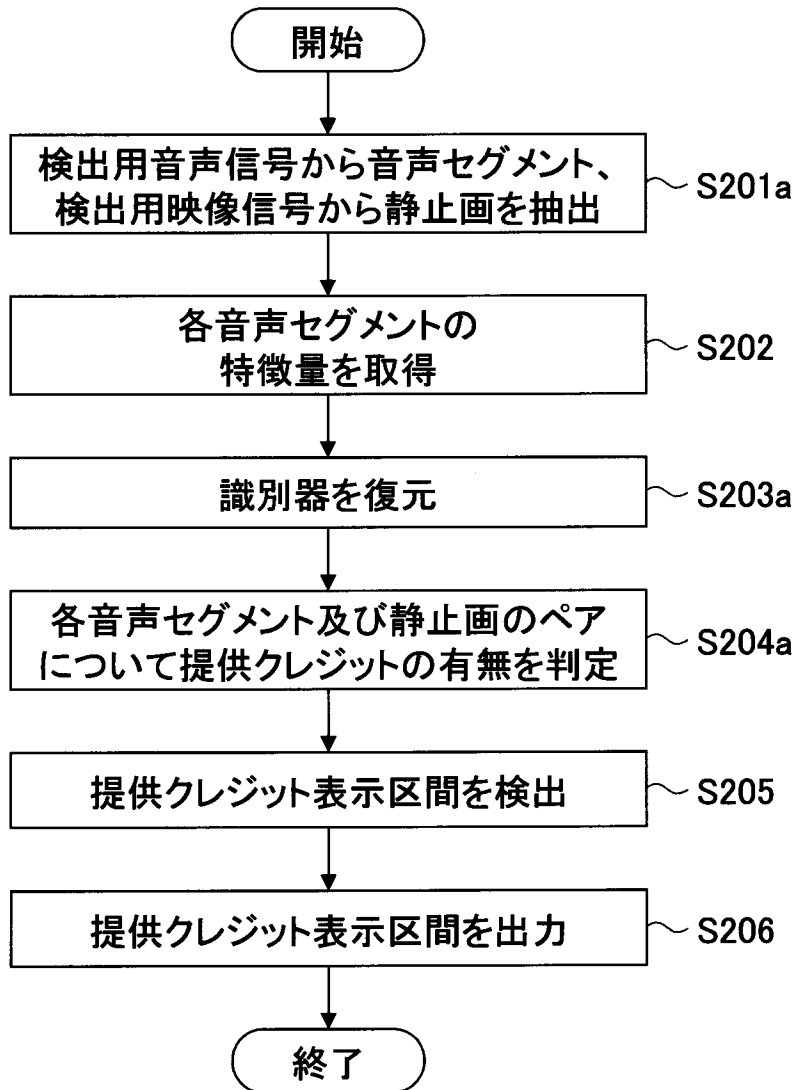
[図10]



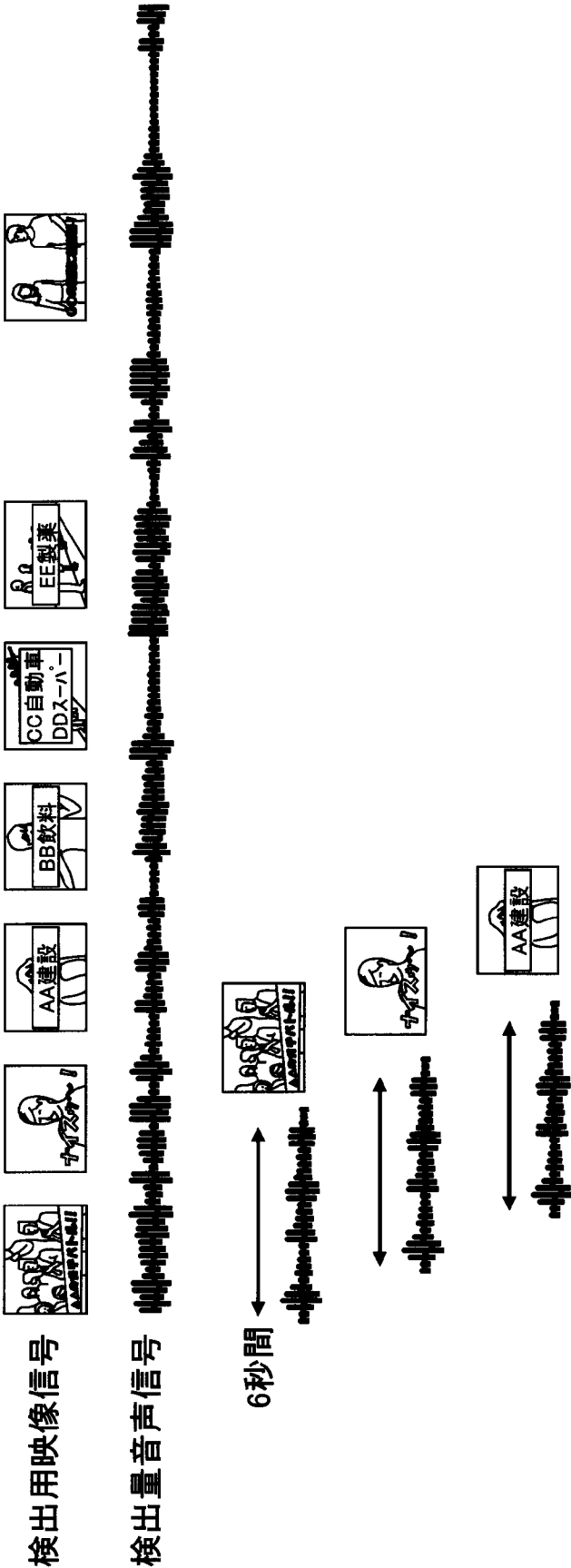
[図11]



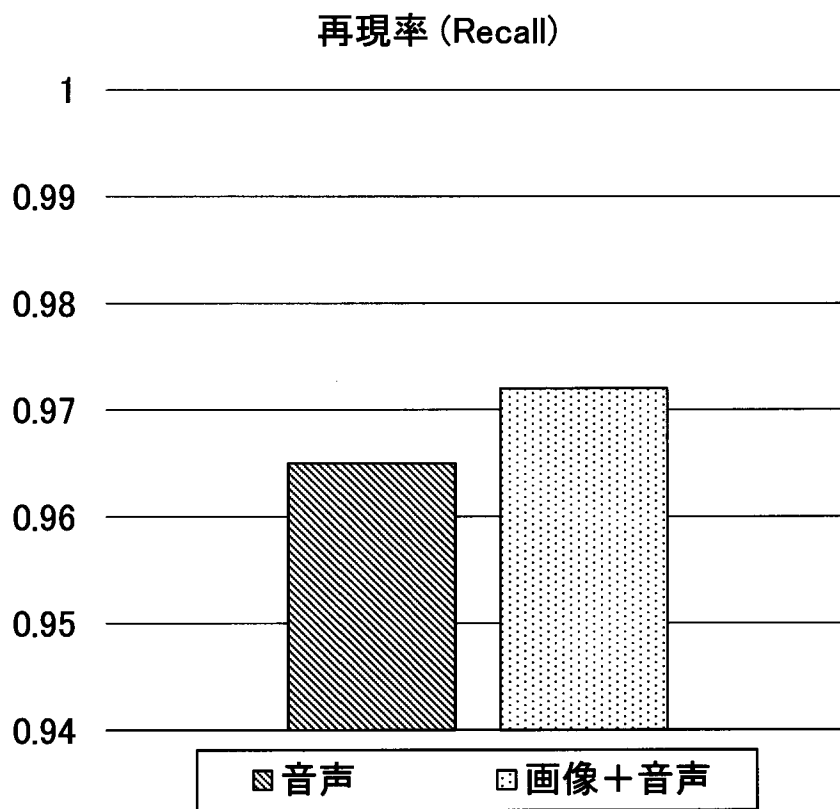
[図12]



[図13]



[図14]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/002458

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. H04N21/442(2011.01)i, H04N17/00(2006.01)i, G06F16/65(2019.01)i, G10L15/00(2013.01)i, G10L15/10(2006.01)i

FI: G06F16/65, H04N21/442, G10L15/10 200W, G10L15/10 200B, G10L15/00 200T, H04N17/00 N

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)

Int. Cl. H04N21/442, H04N17/00, G06F16/65, G10L15/00, G10L15/10

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 1971-2020

Registered utility model specifications of Japan 1996-2020

Published registered utility model applications of Japan 1994-2020

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	WO 2008/050718 A1 (NEC CORP.) 02 May 2008,	1, 3-6
A	paragraphs [0042], [0048], [0055], [0056], [0062], [0072]-[0079] (Family: none)	2
A	JP 2008-108166 A (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) 08 May 2008, paragraphs [0027], [0028] (Family: none)	2

☐ 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
25.02.2020

Date of mailing of the international search report
10.03.2020

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC）） H04N 21/442(2011.01)i; H04N 17/00(2006.01)i; G06F 16/65(2019.01)i; G10L 15/00(2013.01)i; G10L 15/10(2006.01)i FI: G06F16/65; H04N21/442; G10L15/10 200W; G10L15/10 200B; G10L15/00 200T; H04N17/00 N										
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）） H04N21/442; H04N17/00; G06F16/65; G10L15/00; G10L15/10										
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 - 2020年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2020年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2020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 - 2020年	日本国実用新案登録公報	1996 - 2020年	日本国登録実用新案公報	1994 - 2020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 - 2020年									
日本国実用新案登録公報	1996 - 2020年									
日本国登録実用新案公報	1994 - 2020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	WO 2008/050718 A1（日本電気株式会社）02.05.2008（2008 - 05 - 02） [0042], [0048], [0055]-[0056], [0062], [0072]-[0079]	1,3-6								
A	（ファミリーなし）	2								
A	JP 2008-108166 A（松下電器産業株式会社）08.05.2008（2008 - 05 - 08） [0027]-[0028] （ファミリーなし）	2								
☐ 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。										
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的な技術水準を示すもの “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 “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） “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 “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 “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 “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 “&” 同一パテントファミリー文献										
国際調査を完了した日 25.02.2020		国際調査報告の発送日 10.03.2020								
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） 吉田 誠 5N 4807 電話番号 03-3581-1101 内線 3521								