(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4月30日(30.04.2020)

WIPO | PCT

(10) 国際公開番号
WO 2020/085070 A1

(51) 国際特許分類:
G10L 15/10 (2006.01) *G10L 15/06* (2013.01)
G06F 17/27 (2006.01) *G10L 15/16* (2006.01)
G06N 3/02 (2006.01) *G10L 25/63* (2013.01)
G06N 20/00 (2019.01)

(21) 国際出願番号： PCT/JP2019/039572

(22) 国際出願日： 2019年10月8日(08.10.2019)

(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 2018-198427 2018年10月22日(22.10.2018) 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: 安藤 厚志(**ANDO, Atsushi**); 〒1808585 東京都武蔵野市緑町三丁目９番１１号 ＮＴＴ 知的財産センタ内 Tokyo (JP). 神山 歩相名(**KAMIYAMA, Hosana**); 〒1808585 東京都武蔵野市緑町三丁目９番１１号 ＮＴＴ 知的財産センタ内 Tokyo (JP). 小橋川 哲(**KOBASHIKAWA, Satoshi**); 〒1808585 東京都武蔵野市緑町三丁目９番１１号 ＮＴＴ 知的財産センタ内 Tokyo (JP).

(74) 代理人: 中尾 直樹, 外 (**NAKAO, Naoki** et al.); 〒1600022 東京都新宿区新宿三丁目１番２２号 新宿ＮＳＯビル６階 Tokyo (JP).

(54) **Title:** PARALANGUAGE INFORMATION ESTIMATION DEVICE, METHOD FOR ESTIMATING PARALANGUAGE INFORMATION, AND PROGRAM

(54) 発明の名称： パラ言語情報推定装置、パラ言語情報推定方法、およびプログラム

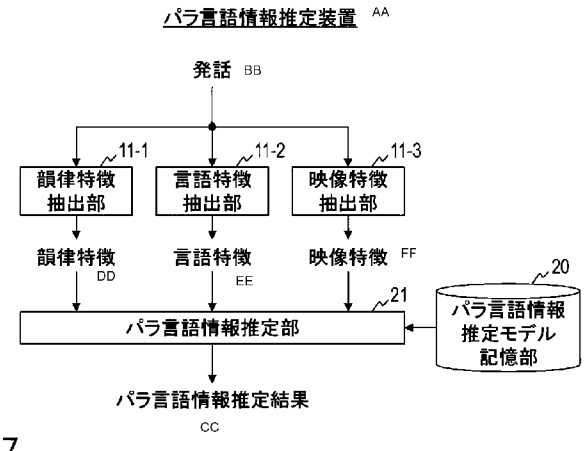

図7

11-1 Prosody feature extraction unit
11-2 Language feature extraction unit
11-3 Video feature extraction unit
20 Paralanguage information estimation model storage unit
21 Paralanguage information estimation unit
AA Paralanguage information estimation device
BB Utterance
CC Paralanguage information estimation results
DD Prosody
EE Language feature
FF Video feature

(57) **Abstract:** The present invention makes it possible to improve the accuracy in estimating paralanguage information. A paralanguage information estimation model storage unit 20 stores a paralanguage information estimation model that uses a plurality of independent feature amounts as input and outputs paralanguage information estimation results. A feature amount extraction unit 11 extracts a feature amount from an input utterance. A paralanguage information estimation unit 20 estimates the paralanguage information of the input utterance from the feature amount extracted from the input utter-

WO 2020/085070 A1

[続葉有]

(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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))

ance using the paralanguage information estimation model. The paralanguage information estimation model comprises: a feature amount sub-model that outputs information used for estimating paralanguage information for each feature amount on the basis of only the feature amount; a feature amount weight calculation unit that calculates a feature amount weight for each feature amount on the basis of the output result from the feature amount sub-model; a feature amount gate that assigns a weight to the output result from the feature amount sub-model for each feature amount using the feature amount weight, and outputs the weighted results; and a result integrating sub-model that estimates paralanguage information on the basis of the output result from all the feature amount gates.

(57) 要約:パラ言語情報推定の精度を向上する。パラ言語情報推定モデル記憶部２０は、複数の独立した特徴量を入力としてパラ言語情報推定結果を出力するパラ言語情報推定モデルを記憶する。特徴量抽出部１１は、入力発話から特徴量を抽出する。パラ言語情報推定部２０は、パラ言語情報推定モデルを用いて入力発話から抽出した特徴量から入力発話のパラ言語情報を推定する。パラ言語情報推定モデルは、特徴量ごとにその特徴量のみに基づいてパラ言語情報の推定に用いる情報を出力する特徴量サブモデルと、特徴量ごとに特徴量サブモデルの出力結果に基づいて特徴量重みを算出する特徴量重み算出部と、特徴量ごとに特徴量サブモデルの出力結果を特徴量重みで重み付けして出力する特徴量ゲートと、すべての特徴量ゲートの出力結果に基づいてパラ言語情報を推定する結果統合サブモデルと、を含む。

明　細　書

発明の名称 ：

パラ言語情報推定装置、パラ言語情報推定方法、およびプログラム

技術分野

[0001]　この発明は、音声からパラ言語情報を推定する技術に関する。

背景技術

[0002]　音声からパラ言語情報（例えば、発話意図が疑問か平叙か、感情が喜び・悲しみ・怒り・平静のいずれか）を推定する技術が求められている。パラ言語情報は、例えば、音声翻訳の高度化（例えば、「明日」という日本語の発話に対して、疑問意図「明日？」と理解して「Is it tomorrow?」と英語に翻訳したり、平叙意図「明日。」と理解して「It is tomorrow.」と英語に翻訳したりと、フランクな発話に対しても発話者の意図を正しく理解した日英翻訳ができる）や、音声対話における話し相手の感情を考慮した対話制御（例えば、相手が怒っていれば話題を変える）などに応用可能である。

[0003]　従来技術として、複数の独立した特徴量を用いたパラ言語情報推定技術が非特許文献１などに示されている。非特許文献１では、音声特徴（音声波形）と映像特徴（複数フレームの画像系列）に基づいて、話者の感情次元値（Valence（感情価）：快－不快、Arousal（覚醒度）：覚醒－睡眠、の二種）を推定する。また、音声の短時間ごとの声の高さなどの韻律特徴の時系列情報と、話した単語などの言語特徴の時系列情報とに基づいて、話者のパラ言語情報を推定する技術も知られている。これらの複数の特徴量を組み合わせる技術は、特徴量単体を利用する技術に比べて高い精度でパラ言語情報を認識できる。

[0004]　図１に、複数の独立した特徴量を用いたパラ言語情報推定モデルの従来技術を例示する。このパラ言語情報推定モデル９００は、各特徴量からパラ言語情報を推定する特徴量サブモデル１０１と、それらの出力を統合して最終的なパラ言語情報推定結果を出力する結果統合サブモデル１０４とで構成さ

れる。この構成は、例えば発話意図推定においては、韻律特徴に疑問や平叙の特性が含まれるか（例えば、語尾が上がっているか否か）、言語特徴に疑問や平叙の特性が表れるか（例えば、疑問詞が含まれるか否か）を推定した後、それらの結果を統合して発話意図が疑問か平叙かを推定する処理に相当する。近年では、各サブモデルを深層学習に基づくモデルで構成し、パラ言語情報推定モデル全体を一体的に学習する、深層学習に基づくパラ言語情報推定モデルが主流となっている。

先行技術文献

非特許文献

[0005] 非特許文献1：Panagiotis Tzirakis, George Trigeorgis, Mihalis A. Nicolaou, Bjorn W. Schuller, Stefanos Zafeiriou, "End-to-End Multimodal Emotion Recognition Using Deep Neural Networks," IEEE Journal of Selected Topics in Signal Processing, vol. 11, No. 8, pp. 1301-1309, 2017.

発明の概要

発明が解決しようとする課題

[0006] パラ言語情報はすべての特徴量にその特性が表れるとは限らず、一部の特徴量だけにパラ言語情報の特性が表れることがある。例えば発話意図では、話し方は語尾上がりだが文章が平叙文である（すなわち、韻律特徴にのみ疑問発話の特性が表れる）発話が存在し、このような発話は疑問発話とみなされる。また、例えば感情では、表情からは平静にみえるが話し方や単語として怒りが強く表れている発話が存在し、このような発話は怒り感情発話とみなされる。

[0007] しかしながら、従来技術では、一部の特徴量だけにパラ言語情報の特性が表れる発話を正しく学習することは困難である。これは、従来技術のパラ言語情報推定モデルでは、すべての特徴量が同じパラ言語情報の特性を示すかのようにモデル学習を行うためである。例えば、疑問発話の学習を行う場合、韻律特徴でも言語特徴でも疑問発話の特性が表れているかのように学習を

行ってしまう。このため、韻律特徴にのみ疑問発話の特性が表れている発話でも、言語特徴にも疑問発話の特性が表れているとみなしてモデル学習をしてしまい、この発話は言語特徴における疑問発話の特性を正しく学習する上でのノイズとなる。その結果、従来技術において、一部の特徴量だけにパラ言語情報の特性が表れる発話が学習データに含まれると、パラ言語情報推定モデルを正しく学習することができず、パラ言語情報推定精度が低下する。

[0008] この発明は、上記のような技術的課題を鑑みて、複数の独立した特徴量を用いたパラ言語情報推定において、一部の特徴量だけにパラ言語情報の特性が表れる発話が学習データに含まれる場合でも、正しくパラ言語情報推定モデルを学習し、正しくパラ言語情報を推定することを目的とする。

課題を解決するための手段

[0009] 上記の課題を解決するために、この発明の一態様のパラ言語情報推定装置は、入力発話からパラ言語情報を推定するパラ言語情報推定装置であって、複数の独立した特徴量を入力としてパラ言語情報推定結果を出力するパラ言語情報推定モデルを記憶するパラ言語情報推定モデル記憶部と、入力発話から複数の独立した特徴量を抽出する特徴量抽出部と、パラ言語情報推定モデルを用いて入力発話から抽出した複数の独立した特徴量から入力発話のパラ言語情報を推定するパラ言語情報推定部と、を含み、パラ言語情報推定モデルは、複数の独立した特徴量ごとにその特徴量のみに基づいてパラ言語情報の推定に用いる情報を出力する特徴量サブモデルと、複数の独立した特徴量ごとに特徴量サブモデルの出力結果に基づいてその特徴量をパラ言語情報の推定に用いるか否かを表す特徴量重みを算出する特徴量重み算出部と、複数の独立した特徴量ごとに特徴量サブモデルの出力結果を特徴量重みで重み付けして出力する特徴量ゲートと、すべての特徴量ゲートの出力結果に基づいてパラ言語情報を推定する結果統合サブモデルと、を含む。

発明の効果

[0010] この発明によれば、複数の独立した特徴量を用いたパラ言語情報推定において、一部の特徴量だけにパラ言語情報の特性が表れる発話に対しても、正

しくパラ言語情報推定モデルを学習し、正しくパラ言語情報を推定することができるようになる。その結果、パラ言語情報推定の精度が向上する。

図面の簡単な説明

[0011] [図1]図１は、従来のパラ言語情報推定モデルを例示する図である。

[図2]図２は、本発明のパラ言語情報推定モデルを例示する図である。

[図3]図３は、パラ言語情報推定モデル学習装置の機能構成を例示する図である。

[図4]図４は、パラ言語情報推定モデル学習方法の処理手順を例示する図である。

[図5]図５は、第一実施形態のパラ言語情報推定モデルを例示する図である。

[図6]図６は、パラ言語情報推定モデル学習部の機能構成を例示する図である。

[図7]図７は、パラ言語情報推定装置の機能構成を例示する図である。

[図8]図８は、パラ言語情報推定方法の処理手順を例示する図である。

[図9]図９は、第二実施形態のパラ言語情報推定モデルを例示する図である。

発明を実施するための形態

[0012] 以下、この発明の実施の形態について詳細に説明する。なお、図面中において同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

[0013] 本発明のポイントは、一部の特徴量だけにパラ言語情報の特性が表れる可能性を考慮し、各特徴量の情報をパラ言語情報推定に利用するかどうかを決定する特徴量ゲートを導入する点にある。一部の特徴量だけにパラ言語情報の特性が表れる発話に対してモデル学習を行うためには、特徴量ごとにパラ言語情報推定に利用するかどうかを選択できる機構を導入すればよいと考えられる。例えば、ある特徴量で特定のパラ言語情報の特性が強く表れている場合はその特徴量を利用してパラ言語情報推定を行うが、別の特徴量で特定のパラ言語情報の特性が表れていない場合は、その特徴量を利用せずにパラ言語情報推定を行う、といった機構を導入する。この選択機構を本発明では特徴量ゲートという形で実現する。

[0014] 図２に、本発明のパラ言語情報推定モデルの例を示す。このパラ言語情報推定モデル１００は、従来と同様の特徴量サブモデル１０１と、特徴量サブモデル１０１の出力をパラ言語情報推定に利用するか否かを決定する特徴量ゲート１０３と、特徴量ゲート１０３の出力に基づいて最終的なパラ言語情報推定結果を出力する結果統合サブモデル１０４とで構成される。

[0015] 特徴量ゲート１０３は、各特徴量サブモデル１０１の出力を結果統合サブモデル１０４に入力するかどうかを決定する役割を持つ。特徴量ゲート１０３は、式（１）に基づいて出力を決定する。

[0016] [数1]

$$y_k = w_k \circ x_k \quad \cdots(1)$$

[0017] ここで、kは特徴量番号（k=1，2，…）、y_kは特徴量ゲート出力ベクトル、x_kは特徴量ゲート入力ベクトル（特徴量サブモデル出力結果）、w_kは特徴量ゲート重みベクトル、

[0018] [数2]

$\circ$

[0019] は、要素積を表す。特徴量ゲート重みベクトルw_kが単位ベクトルのとき、特徴量サブモデル出力結果x_kがそのまま結果統合サブモデル１０４へ出力される。特徴量ゲート重みベクトルw_kがゼロベクトルのとき、特徴量サブモデル出力結果x_kがゼロに変換されて結果統合サブモデル１０４へ出力される。このように、各特徴量に対応する特徴量ゲート重みベクトルw_kを制御することで、ある特徴量を利用するが別の特徴量は利用しないというようにパラ言語情報推定モデルの学習やパラ言語情報の推定を行うことが可能となる。なお、深層学習に基づくパラ言語情報推定モデルの場合、特徴量ゲート重みベクトルw_kもモデルパラメータの一つであるとみなせるため、特徴量ゲート重みベクトルw_kも含めてモデル全体を一体学習することが可能である。

[0020] 具体的には、以下の手順によりパラ言語情報の推定を行う。

[0021] １．複数の独立した特徴量を入力とし、特徴量ごとのサブモデル、特徴量ごとの特徴量ゲート、結果統合サブモデルから構成されるパラ言語情報推定

モデルを用意する。

[0022] ２．パラ言語情報推定モデルの学習を行う。深層学習に基づくパラ言語情報推定モデルの場合、特徴量ゲートの重みベクトルを含めたモデル全体を誤差逆伝搬法により一体学習する。それ以外の場合では特徴量ゲートは学習できないため、特徴量ゲートの重みベクトルは人手によるルールで決定する。例えば、特徴量ごとのサブモデルの出力結果が識別平面からの距離の場合、識別平面からの距離の絶対値が0.5以下なら特徴量ゲートの重みベクトルはゼロベクトル、識別平面からの距離の絶対値が0.5より大きいなら特徴量ゲートの重みベクトルは単位ベクトルとする、というルールを定める。この場合、特徴量ごとのサブモデルを先に学習し、その後結果統合サブモデルを学習するという二段階の学習を行う。

[0023] ３．学習済みのパラ言語情報推定モデルに複数の独立した特徴量を入力し、発話ごとにパラ言語情報推定結果を得る。

[0024] ［第一実施形態］

本実施形態において、入力発話とは、当該発話の音声波形情報および当該発話の話者の表情（顔）の映像情報の両方を指すものとする。本発明でパラ言語情報推定に用いる特徴量は、人間の発話から抽出できる独立した二以上の特徴量であればよいが、本実施形態では、韻律特徴、言語特徴、および映像特徴の互いに独立な三種類の特徴量を用いるものとする。ただし、これら三種類の特徴量のうち、いずれか二種類の特徴量のみを用いてもよい。また、他特徴量と互いに独立であれば、例えば生体信号情報（脈拍、皮膚電位など）などの情報を用いた特徴量を追加で利用してもよい。

[0025] 本実施形態では、特徴量ごとのサブモデルの出力結果として特徴量ごとのパラ言語情報確率を受け取ることもできるが、特徴量ごとのパラ言語情報確率の推定のために必要な中間情報（例えば、ディープニューラルネットワークにおける中間層の出力値）を受け取ることもできる。また、特徴量ゲートの重みベクトルも含めて学習を行う場合、重みベクトルはすべての入力に対して固定値ではなく、入力が変わるたびに動的に重みベクトルを変えること

もできる。具体的には、式（２）または式（３）を用いて、入力から重みベクトルを算出することで、重みベクトルを動的に変化させる。

[0026] [数3]

$$w_k = w_x x_k + b_x \quad \cdots(2)$$

$$w_k = \sigma\left(w_x x_k + b_x\right) \quad \cdots(3)$$

[0027] ここで、kは特徴量番号（k=1，2，…）、x_kは特徴量ゲート入力ベクトル（特徴量サブモデル出力結果）、w_kは特徴量ゲート重みベクトル、w_xは特徴量ゲート重みベクトル算出用行列、b_xは特徴量ゲート重みベクトル算出用バイアス、σは活性化関数（例えば、式（４）のシグモイド関数）を表す。w_x，b_xは予め学習により決定しておく。なお、式（４）においてxがベクトルの場合、ベクトルの各要素に対して式（４）を適用する。

[0028] [数4]

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad \cdots(4)$$

[0029] 上記のように構成することにより、入力発話の話者や発話環境に応じて特徴量ごとのサブモデルの出力結果の利用度合いを変える（例えば、抑揚にパラ言語情報が表れやすい人では韻律特徴を重視してパラ言語情報推定を行う、など）ことができる。そのため、一般的な特徴量ごとのパラ言語情報確率の重み付け和に基づく推定手法に比べて、より多様な入力に対しても高精度にパラ言語情報を推定することが可能となる。すなわち、多様な発話環境に対するパラ言語情報推定精度が向上する。

[0030] ＜パラ言語情報推定モデル学習装置＞

第一実施形態のパラ言語情報推定モデル学習装置は、教師ラベルが付与された発話からパラ言語情報推定モデルを学習する。パラ言語情報推定モデル学習装置は、図３に例示するように、発話記憶部１０－１、教師ラベル記憶部１０－２、韻律特徴抽出部１１－１、言語特徴抽出部１１－２、映像特徴抽出部１１－３、パラ言語情報推定モデル学習部１２、およびパラ言語情報

推定モデル記憶部２０を備える。以下、韻律特徴抽出部１１－１、言語特徴抽出部１１－２、および映像特徴抽出部１１－３を特徴量抽出部１１と総称することもある。特徴量抽出部１１はパラ言語情報推定に用いる特徴量の種類に応じて数や処理内容等の構成を変更する。このパラ言語情報推定モデル学習装置が、図４に例示する各ステップの処理を行うことにより第一実施形態のパラ言語情報推定モデル学習方法が実現される。

[0031] パラ言語情報推定モデル学習装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。パラ言語情報推定モデル学習装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。パラ言語情報推定モデル学習装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて中央演算処理装置へ読み出されて他の処理に利用される。パラ言語情報推定モデル学習装置の各処理部は、少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。パラ言語情報推定モデル学習装置が備える各記憶部は、例えば、RAM（Random Access Memory）などの主記憶装置、ハードディスクや光ディスクもしくはフラッシュメモリ（Flash Memory）のような半導体メモリ素子により構成される補助記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。

[0032] 発話記憶部１０－１には、パラ言語情報推定モデルの学習に用いる発話（以下、「学習発話」ともいう）が記憶されている。本実施形態では、発話は人間の発話音声を収録した音声波形情報と、その発話の話者の表情を収録した映像情報とからなるものとする。発話が具体的にどのような情報から構成されるかはパラ言語情報の推定にどのような特徴量を用いるかに応じて決定される。

[0033] 教師ラベル記憶部１０－２には、発話記憶部１０－１に記憶された各発話

に付与されるパラ言語情報の正解値を表す教師ラベルが記憶されている。発話に対する教師ラベルの付与は、人手で行ってもよいし、周知のラベル分類技術を用いて行ってもよい。具体的にどのような教師ラベルを付与するかはパラ言語情報の推定にどのような特徴量を用いるかに応じて決定する。

[0034] ステップＳ１１－１において、韻律特徴抽出部１１－１は、発話記憶部１０－１に記憶された各発話の音声波形情報から韻律特徴を抽出する。韻律特徴は、基本周波数、短時間パワー、MFCC（Mel-frequency Cepstral Coefficients）、ゼロ交差率、Harmonics-to-Noise-Ratio（HNR）、メルフィルタバンク出力、のいずれか一つ以上の特徴量を含むベクトルである。また、これらの時間ごと（フレームごと）の系列ベクトルであってもよいし、これらの発話全体の統計量（平均、分散、最大値、最小値、勾配など）のベクトルであってもよい。韻律特徴抽出部１１－１は、抽出した韻律特徴をパラ言語情報推定モデル学習部１２へ出力する。

[0035] ステップＳ１１－２において、言語特徴抽出部１１－２は、発話記憶部１０－１に記憶された各発話の音声波形情報から言語特徴を抽出する。言語特徴の抽出には、音声認識技術により取得した単語列または音素認識技術により取得した音素列を利用する。言語特徴はこれらの単語列または音素列を系列ベクトルとして表現したものであってもよいし、発話全体での特定単語の出現数などを表すベクトルとしてもよい。言語特徴抽出部１１－２は、抽出した言語特徴をパラ言語情報推定モデル学習部１２へ出力する。

[0036] ステップＳ１１－３において、映像特徴抽出部１１－３は、発話記憶部１０－１に記憶された各発話の映像情報から映像特徴を抽出する。映像特徴は、各フレームでの顔の特徴点の位置座標、オプティカルフローから算出した小領域ごとの速度成分、局所的な画像勾配のヒストグラム（Histograms of Oriented Gradients: HOG）のいずれか一つ以上を含むベクトルである。また、これらの一定間隔の時間ごと（フレームごと）の系列ベクトルであってもよいし、これらの発話全体の統計量（平均、分散、最大値、最小値、勾配など）のベクトルであってもよい。映像特徴抽出部１１－３は、抽出した映像

特徴をパラ言語情報推定モデル学習部１２へ出力する。

[0037] ステップＳ１２において、パラ言語情報推定モデル学習部１２は、入力された韻律特徴、言語特徴、および映像特徴と、教師ラベル記憶部１０－２に記憶された教師ラベルとを用いて、複数の独立した特徴量を入力としてパラ言語情報推定結果を出力するパラ言語情報推定モデルを学習する。パラ言語情報推定モデル学習部１２は、学習済みのパラ言語情報推定モデルをパラ言語情報推定モデル記憶部２０へ記憶する。

[0038] 図５に、本実施形態で利用するパラ言語情報推定モデルの構成例を示す。このパラ言語情報推定モデルは、韻律特徴サブモデル１０１－１、言語特徴サブモデル１０１－２、映像特徴サブモデル１０１－３、韻律特徴重み算出部１０２－１、言語特徴重み算出部１０２－２、映像特徴重み算出部１０２－３、韻律特徴ゲート１０３－１、言語特徴ゲート１０３－２、映像特徴ゲート１０３－３、および結果統合サブモデル１０４を備える。以下、韻律特徴サブモデル１０１－１、言語特徴サブモデル１０１－２、および映像特徴サブモデル１０１－３を特徴量サブモデル１０１と、韻律特徴重み算出部１０２－１、言語特徴重み算出部１０２－２、および映像特徴重み算出部１０２－３を特徴量重み算出部１０２と、韻律特徴ゲート１０３－１、言語特徴ゲート１０３－２、および映像特徴ゲート１０３－３を特徴量ゲート１０３と総称することもある。特徴量サブモデル１０１は、入力された特徴量のみに基づいてパラ言語情報の推定を行い、パラ言語推定結果もしくはパラ言語推定の際に生成される中間値（以下、「パラ言語情報の推定に用いる情報」ともいう）を出力する。特徴量重み算出部１０２は、特徴量サブモデル１０１の出力結果に基づいてその特徴量をパラ言語情報の推定に用いるか否かを表す特徴量ゲート重みベクトル（以下、「特徴量重み」ともいう）を算出する。特徴量ゲート１０３は、特徴量サブモデル１０１の出力結果を特徴量重み算出部１０２が出力する特徴量ゲート重みベクトルで重み付けして出力する。結果統合サブモデル１０４は、すべての特徴量ゲート１０３の出力結果に基づいてパラ言語情報を推定する。

[0039] パラ言語情報推定モデルは、例えば深層学習に基づくDeep Neural Network（DNN）であってもよいし、Support Vector Machine（SVM）であってもよい。また、時間ごとの系列ベクトルを特徴量に用いる場合、Long Short-Term Memory Recurrent Neural Network（LSTM-RNN）などの系列を考慮できる推定モデルを用いてもよい。なお、パラ言語情報推定モデルがすべてDNNやLSTM-RNNを含む深層学習に基づく手法によって構成される場合、特徴量ゲートの重みベクトルも含めてモデル全体を単一のネットワーク（分類モデル）と見なすことができるため、パラ言語情報推定モデル全体を誤差逆伝搬法により一体学習することが可能である。

[0040] パラ言語情報推定モデルが深層学習に基づく手法以外を含む場合（例えば各特徴量のサブモデルがSVMによって構成される場合）、特徴量ゲートの重みベクトルの数値や重みベクトルの決定規則は人手により与える必要がある。またこの場合、特徴量ごとのサブモデルや結果統合サブモデルは別々に学習する必要がある。このような場合でのパラ言語情報推定モデル学習部１２の構成を図６に示す。この場合のパラ言語情報推定モデル学習部１２は、韻律特徴サブモデル学習部１２１－１、言語特徴サブモデル学習部１２１－２、映像特徴サブモデル学習部１２１－３、韻律特徴重み算出部１２２－１、言語特徴重み算出部１２２－２、映像特徴重み算出部１２２－３、韻律特徴ゲート処理部１２３－１、言語特徴ゲート処理部１２３－２、映像特徴ゲート処理部１２３－３、および結果統合サブモデル学習部１２４を備える。

[0041] 韻律特徴サブモデル学習部１２１－１は、韻律特徴と教師ラベルとの組から、韻律特徴のみに基づいてパラ言語情報を推定する韻律特徴サブモデルを学習する。韻律特徴サブモデルは例えばSVMを用いるが、クラス分類が可能な他の機械学習手法を用いてもよい。また、韻律特徴サブモデルの出力結果とは、例えば韻律特徴サブモデルがSVMであれば識別平面からの距離を指す。

[0042] 言語特徴サブモデル学習部１２１－２および映像特徴サブモデル学習部１２１－３は、韻律特徴サブモデル学習部１２１－１と同様にして、言語特徴サブモデルおよび映像特徴サブモデルを学習する。

[0043] 韻律特徴重み算出部１２２－１は、特徴量ゲートルールを用いて、韻律特徴サブモデルの出力結果から韻律特徴ゲート重みベクトルを算出する。特徴量ゲートルールとは、特徴量ゲートを決定する規則と、特徴量ゲートの重みベクトルとの組を指す。韻律特徴サブモデルがSVMの例であれば、「韻律特徴サブモデルの出力結果において、識別平面からの距離の絶対値が0.5以下なら韻律特徴ゲート重みベクトルはゼロベクトル、識別平面からの距離の絶対値が0.5より大きいなら韻律特徴ゲート重みベクトルは単位ベクトル」といった、人手により与えたルールを指す。これは、SVMの識別平面からの距離が推定結果の尤もらしさであるとみなし、推定結果が尤もらしい（ある特徴量で特定のパラ言語情報の特性が強く表れている可能性が高い）場合は特徴量ゲート重みベクトルを単位ベクトルに、そうでない場合はゼロベクトルに設定する処理に等しい。この人手により与えたルールを韻律特徴サブモデルの出力結果に適用し、出力結果に対する韻律特徴ゲート重みベクトルを算出する。なお、韻律特徴ゲート重みベクトルの次元数は韻律特徴サブモデル出力結果と同じとする（SVMの例であれば１次元のベクトルとする）。

[0044] 言語特徴重み算出部１２２－２および映像特徴重み算出部１２２－３は、韻律特徴重み算出部１２２－１と同様にして、言語特徴重みベクトルおよび映像特徴重みベクトルを算出する。

[0045] 韻律特徴ゲート処理部１２３－１は、韻律特徴サブモデルの出力結果と、韻律特徴ゲート重みベクトルとを用いて、上記式（１）を計算し、韻律特徴ゲート出力ベクトルを求める。

[0046] 言語特徴ゲート処理部１２３－２および映像特徴ゲート処理部１２３－３は、韻律特徴ゲート処理部１２３－１と同様にして、言語特徴ゲート出力ベクトルおよび映像特徴ゲート出力ベクトルを算出する。

[0047] 結果統合サブモデル学習部１２４は、韻律特徴ゲート出力ベクトル、言語特徴ゲート出力ベクトル、映像特徴ゲート出力ベクトル、および教師ラベルの組から、結果統合サブモデルを学習する。結果統合サブモデルは例えばSVMを用いるが、クラス分類が可能な他の機械学習手法を用いてもよい。

[0048] ＜パラ言語情報推定装置＞

第一実施形態のパラ言語情報推定装置は、学習済みのパラ言語情報推定モデルを用いて入力発話からパラ言語情報を推定する。パラ言語情報推定装置は、図７に例示するように、韻律特徴抽出部１１－１、言語特徴抽出部１１－２、映像特徴抽出部１１－３、パラ言語情報推定モデル記憶部２０、およびパラ言語情報推定部２１を備える。このパラ言語情報推定装置が、図８に例示する各ステップの処理を行うことにより第一実施形態のパラ言語情報推定方法が実現される。

[0049] パラ言語情報推定装置は、例えば、中央演算処理装置（CPU: Central Processing Unit）、主記憶装置（RAM: Random Access Memory）などを有する公知又は専用のコンピュータに特別なプログラムが読み込まれて構成された特別な装置である。パラ言語情報推定装置は、例えば、中央演算処理装置の制御のもとで各処理を実行する。パラ言語情報推定装置に入力されたデータや各処理で得られたデータは、例えば、主記憶装置に格納され、主記憶装置に格納されたデータは必要に応じて中央演算処理装置へ読み出されて他の処理に利用される。パラ言語情報推定装置の各処理部は、少なくとも一部が集積回路等のハードウェアによって構成されていてもよい。パラ言語情報推定装置が備える各記憶部は、例えば、RAM（Random Access Memory）などの主記憶装置、ハードディスクや光ディスクもしくはフラッシュメモリ（Flash Memory）のような半導体メモリ素子により構成される補助記憶装置、またはリレーショナルデータベースやキーバリューストアなどのミドルウェアにより構成することができる。

[0050] ステップＳ１１－１において、韻律特徴抽出部１１－１は、入力発話の音声波形情報から韻律特徴を抽出する。韻律特徴の抽出は、パラ言語情報推定モデル学習装置と同様に行えばよい。韻律特徴抽出部１１－１は、抽出した韻律特徴をパラ言語情報推定部２１へ出力する。

[0051] ステップＳ１１－２において、言語特徴抽出部１１－２は、入力発話の音声波形情報から言語特徴を抽出する。言語特徴の抽出は、パラ言語情報推定

モデル学習装置と同様に行えばよい。言語特徴抽出部１１－２は、抽出した言語特徴をパラ言語情報推定部２１へ出力する。

[0052] ステップＳ１１－３において、映像特徴抽出部１１－３は、入力発話の映像情報から映像特徴を抽出する。映像特徴の抽出は、パラ言語情報推定モデル学習装置と同様に行えばよい。映像特徴抽出部１１－３は、抽出した映像特徴をパラ言語情報推定部２１へ出力する。

[0053] ステップＳ２１において、パラ言語情報推定部２１は、入力発話から抽出した韻律特徴、言語特徴、および映像特徴に基づいて、当該発話のパラ言語情報を推定する。推定にはパラ言語情報推定モデル記憶部２０に記憶された学習済みのパラ言語情報推定モデルを用いる。パラ言語情報推定モデルが深層学習に基づくモデルである場合、各特徴量を順伝播することでパラ言語情報推定結果が得られる。深層学習に基づくモデルでない場合、各特徴量をそれぞれ特徴量サブモデルに入力し、各特徴量サブモデルの出力結果に特徴量ゲートルールを適用して特徴量ゲート重みベクトルを求め、上記式（１）に従って特徴量ゲート重みベクトルと特徴量サブモデルの出力結果との要素積を取った結果を結果統合サブモデルに入力することでパラ言語情報推定結果が得られる。

[0054] ［第二実施形態］

第一実施形態では、ある特徴量の特徴量ゲート重みベクトルは、その特徴量の特徴量サブモデルの出力結果から決定している。これは、例えば韻律特徴において特定のパラ言語情報の特性が強く表れていると判断されたとき、韻律特徴をパラ言語情報推定に利用するという構成である。

[0055] 第二実施形態では、ある特徴量の特徴量ゲート重みベクトルは、すべての特徴量の特徴量サブモデルの出力結果から決定する。すべての特徴量の特徴量サブモデルの出力結果を考慮して特徴量ゲート重みベクトルを決定することで、どの特徴量の情報をパラ言語情報推定に利用すべきかを区別しやすくなり、各特徴量にわずかにパラ言語情報の特性が表れる発話に対してもパラ言語情報推定精度が向上する。例えば、韻律特徴でも言語特徴でも特定のパ

ラ言語情報の特性がわずかに表れるような場合、韻律特徴と言語特徴の特性の現れ方を比較し、特性がより強く表れている方の特徴量をパラ言語情報推定に利用できるようになるためである。

[0056] 第二実施形態のパラ言語情報推定モデルは、図９に示すように、各特徴量サブモデル１０１（例えば、韻律特徴サブモデル１０１－１）の出力をすべての特徴量重み算出部１０２（すなわち、韻律特徴重み算出部１０２－１、言語特徴重み算出部１０２－２、および映像特徴重み算出部１０２－３）に入力するように構成する。各特徴量重み算出部１０２（例えば、韻律特徴重み算出部１０２－１）は、すべての特徴量サブモデル１０１（すなわち、韻律特徴サブモデル１０１－１、言語特徴サブモデル１０１－２、および映像特徴サブモデル１０１－３）の出力を比較して、その特徴量の特徴量ゲート重みベクトル（すなわち、韻律特徴ゲート重みベクトル）を決定する。

[0057] 第二実施形態のパラ言語情報推定モデル学習装置およびパラ言語情報推定装置は、図９に示すパラ言語情報推定モデルを用いることで、第一実施形態と同様の手順により、パラ言語情報推定モデルの学習やパラ言語情報の推定が可能である。

[0058] 以上、この発明の実施の形態について説明したが、具体的な構成は、これらの実施の形態に限られるものではなく、この発明の趣旨を逸脱しない範囲で適宜設計の変更等があっても、この発明に含まれることはいうまでもない。実施の形態において説明した各種の処理は、記載の順に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されてもよい。

[0059] ［プログラム、記録媒体］

上記実施形態で説明した各装置における各種の処理機能をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記各装置における各種の処理機能がコンピュータ上で実現される。

[0060] この処理内容を記述したプログラムは、コンピュータで読み取り可能な記

録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等どのようなものでもよい。

[0061] また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

[0062] このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記憶装置に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP（Application Service Provider）型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの（コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等）を含むものとする。

[0063] また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、本装置を構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

符号の説明

[0064] １０－１　発話記憶部

１０－２　教師ラベル記憶部

１１　特徴量抽出部

１２　パラ言語情報推定モデル学習部

２０　パラ言語情報推定モデル記憶部

２１　パラ言語情報推定部

１００，９００　パラ言語情報推定モデル

１０１　特徴量サブモデル

１０２　特徴量重みベクトル

１０３　特徴量ゲート

１０４　結果統合サブモデル

１２１　特徴量サブモデル学習部

１２２　特徴量重み算出部

１２３　特徴量ゲート処理部

１２４　結果統合サブモデル学習部

請求の範囲

[請求項1] 入力発話からパラ言語情報を推定するパラ言語情報推定装置であって、

複数の独立した特徴量を入力としてパラ言語情報推定結果を出力するパラ言語情報推定モデルを記憶するパラ言語情報推定モデル記憶部と、

入力発話から上記複数の独立した特徴量を抽出する特徴量抽出部と、

上記パラ言語情報推定モデルを用いて上記入力発話から抽出した上記複数の独立した特徴量から上記入力発話のパラ言語情報を推定するパラ言語情報推定部と、

を含み、

上記パラ言語情報推定モデルは、

上記複数の独立した特徴量ごとにその特徴量のみに基づいてパラ言語情報の推定に用いる情報を出力する特徴量サブモデルと、

上記複数の独立した特徴量ごとに上記特徴量サブモデルの出力結果に基づいてその特徴量をパラ言語情報の推定に用いるか否かを表す特徴量重みを算出する特徴量重み算出部と、

上記複数の独立した特徴量ごとに上記特徴量サブモデルの出力結果を上記特徴量重みで重み付けして出力する特徴量ゲートと、

すべての上記特徴量ゲートの出力結果に基づいて上記パラ言語情報を推定する結果統合サブモデルと、

を含むパラ言語情報推定装置。

[請求項2] 請求項１に記載のパラ言語情報推定装置であって、

上記特徴量重み算出部は、すべての上記特徴量の上記特徴量サブモデルの出力結果に基づいて上記特徴量重みを算出するものである、

パラ言語情報推定装置。

[請求項3] 請求項１または２に記載のパラ言語情報推定装置であって、

上記特徴量重み算出部は、kを特徴量番号とし、x_kを上記特徴量サブモデルの出力結果とし、w_kを上記特徴量重みとし、w_xをあらかじめ学習した行列とし、b_xをあらかじめ学習したバイアスとし、σを活性化関数とし、

[数5]

$$w_k = w_x x_k + b_x$$

または

[数6]

$$w_k = \sigma\left(w_x x_k + b_x\right)$$

により上記特徴量重みを算出するものである、

パラ言語情報推定装置。

[請求項4] 請求項１から３のいずれかに記載のパラ言語情報推定装置であって、

上記パラ言語情報推定モデルは、ニューラルネットワークに基づくモデルであり、

上記特徴量重みは、固定値または入力に応じた関数であり、

上記特徴量サブモデルと上記特徴量重みと上記結果統合サブモデルとは、複数の学習発話から抽出した上記複数の独立した特徴量と上記学習発話に付与された教師ラベルとを用いて一体で学習したものである、

パラ言語情報推定装置。

[請求項5] 請求項１から３のいずれかに記載のパラ言語情報推定装置であって、

上記特徴量サブモデルは、複数の学習発話から抽出した上記複数の独立した特徴量と上記学習発話に付与された教師ラベルとから学習したものであり、

上記特徴量重みは、上記特徴量ごとにあらかじめ定められたルール

に従って算出されるものであり、

上記結果統合サブモデルは、すべての上記特徴量ゲートの出力結果と上記教師ラベルとから学習したものである、

パラ言語情報推定装置。

[請求項6] 入力発話からパラ言語情報を推定するパラ言語情報推定方法であって、

パラ言語情報推定モデル記憶部に、複数の独立した特徴量を入力としてパラ言語情報推定結果を出力するパラ言語情報推定モデルが記憶されており、

特徴量抽出部が、入力発話から上記複数の独立した特徴量を抽出し、

パラ言語情報推定部が、上記パラ言語情報推定モデルを用いて上記入力発話から抽出した上記複数の独立した特徴量から上記入力発話のパラ言語情報を推定し、

上記パラ言語情報推定モデルは、

上記複数の独立した特徴量ごとにその特徴量のみに基づいてパラ言語情報の推定に用いる情報を出力する特徴量サブモデルと、

上記複数の独立した特徴量ごとに上記特徴量サブモデルの出力結果に基づいてその特徴量をパラ言語情報の推定に用いるか否かを表す特徴量重みを算出する特徴量重み算出部と、

上記複数の独立した特徴量ごとに上記特徴量サブモデルの出力結果を上記特徴量重みで重み付けして出力する特徴量ゲートと、

すべての上記特徴量ゲートの出力結果に基づいて上記パラ言語情報を推定する結果統合サブモデルと、

を含むパラ言語情報推定方法。

[請求項7] 請求項６に記載のパラ言語情報推定方法であって、

上記パラ言語情報推定モデルは、ニューラルネットワークに基づくモデルであり、

上記特徴量重みは、固定値または入力に応じた関数であり、

上記特徴量サブモデルと上記特徴量重みと上記結果統合サブモデルとは、複数の学習発話から抽出した上記複数の独立した特徴量と上記学習発話に付与された教師ラベルとを用いて一体で学習したものである、

パラ言語情報推定方法。

[請求項8] 請求項１から５のいずれかに記載のパラ言語情報推定装置としてコンピュータを機能させるためのプログラム。

[図1]

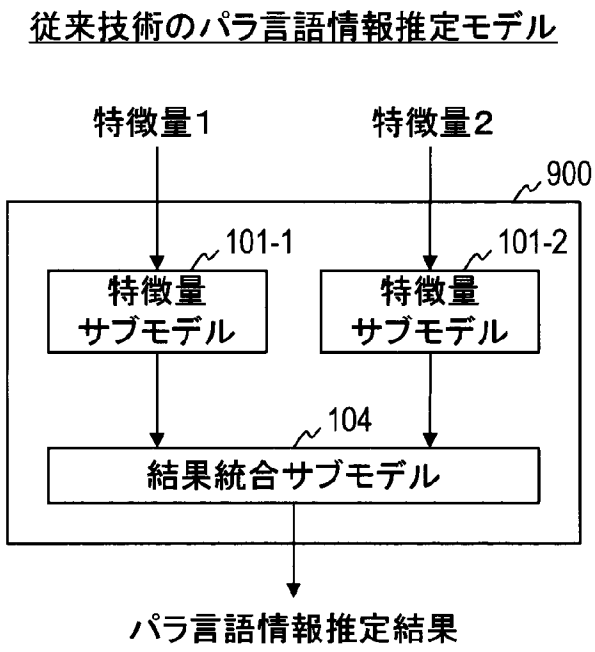
従来技術のパラ言語情報推定モデル
特徴量1
特徴量2
900
101-1
101-2
特徴量
サブモデル
特徴量
サブモデル
104
結果統合サブモデル
パラ言語情報推定結果

図1

[図2]

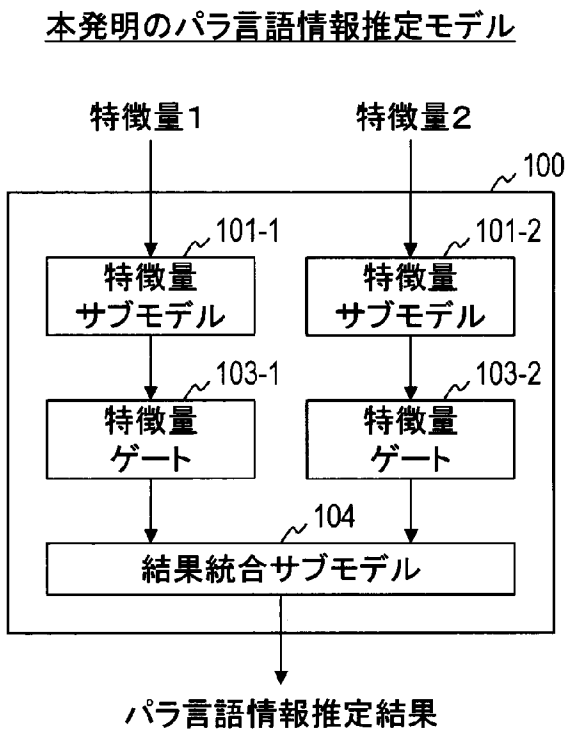
本発明のパラ言語情報推定モデル
特徴量1
特徴量2
100
101-1
101-2
特徴量
サブモデル
特徴量
サブモデル
103-1
103-2
特徴量
ゲート
特徴量
ゲート
104
結果統合サブモデル
パラ言語情報推定結果

図2

[図3]

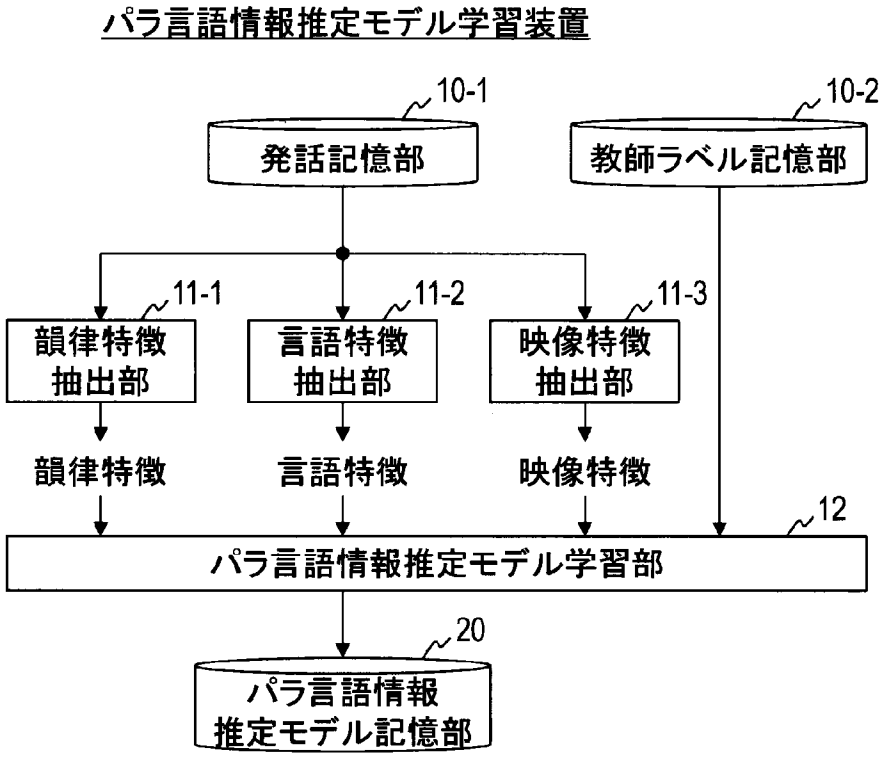
パラ言語情報推定モデル学習装置
10-1
発話記憶部
10-2
教師ラベル記憶部
11-1
韻律特徴
抽出部
11-2
言語特徴
抽出部
11-3
映像特徴
抽出部
韻律特徴
言語特徴
映像特徴
12
パラ言語情報推定モデル学習部
20
パラ言語情報
推定モデル記憶部

図3

[図4]

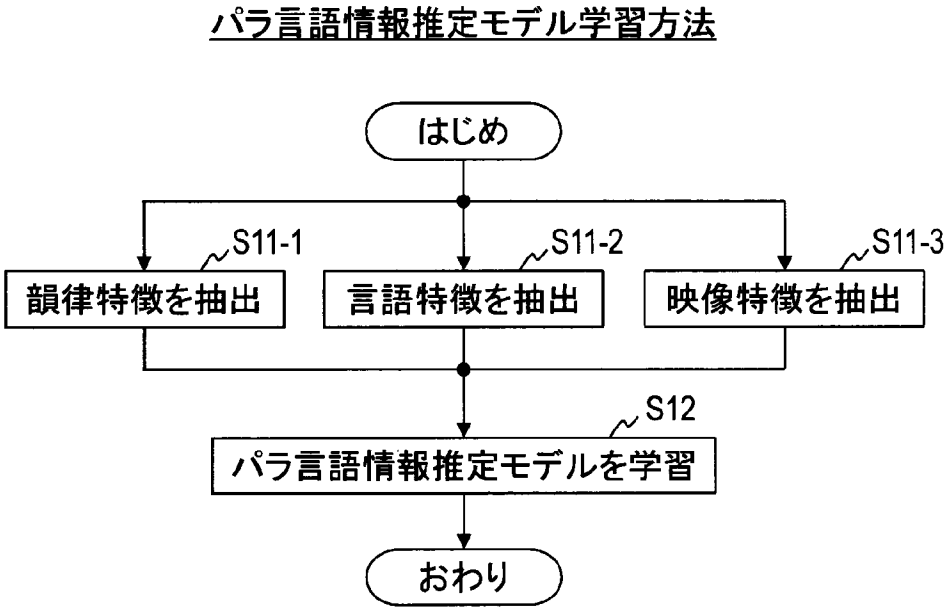
パラ言語情報推定モデル学習方法
はじめ
S11-1
韻律特徴を抽出
S11-2
言語特徴を抽出
S11-3
映像特徴を抽出
S12
パラ言語情報推定モデルを学習
おわり

図4

[図5]

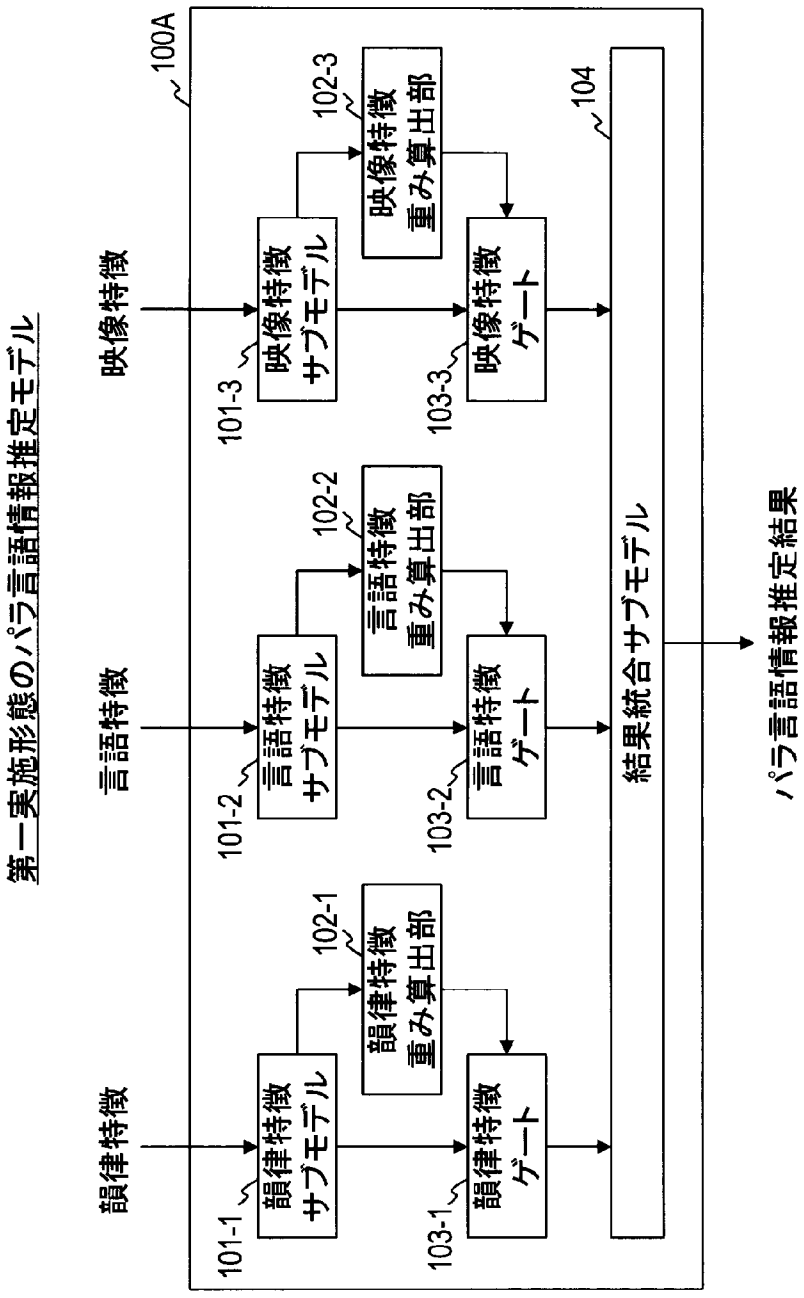
第一実施形態のパラ言語情報推定モデル
100A
韻律特徴
言語特徴
映像特徴
101-1
韻律特徴サブモデル
101-2
言語特徴サブモデル
101-3
映像特徴サブモデル
102-1
韻律特徴重み算出部
102-2
言語特徴重み算出部
102-3
映像特徴重み算出部
103-1
韻律特徴ゲート
103-2
言語特徴ゲート
103-3
映像特徴ゲート
104
結果統合サブモデル
パラ言語情報推定結果

図5

[図6]

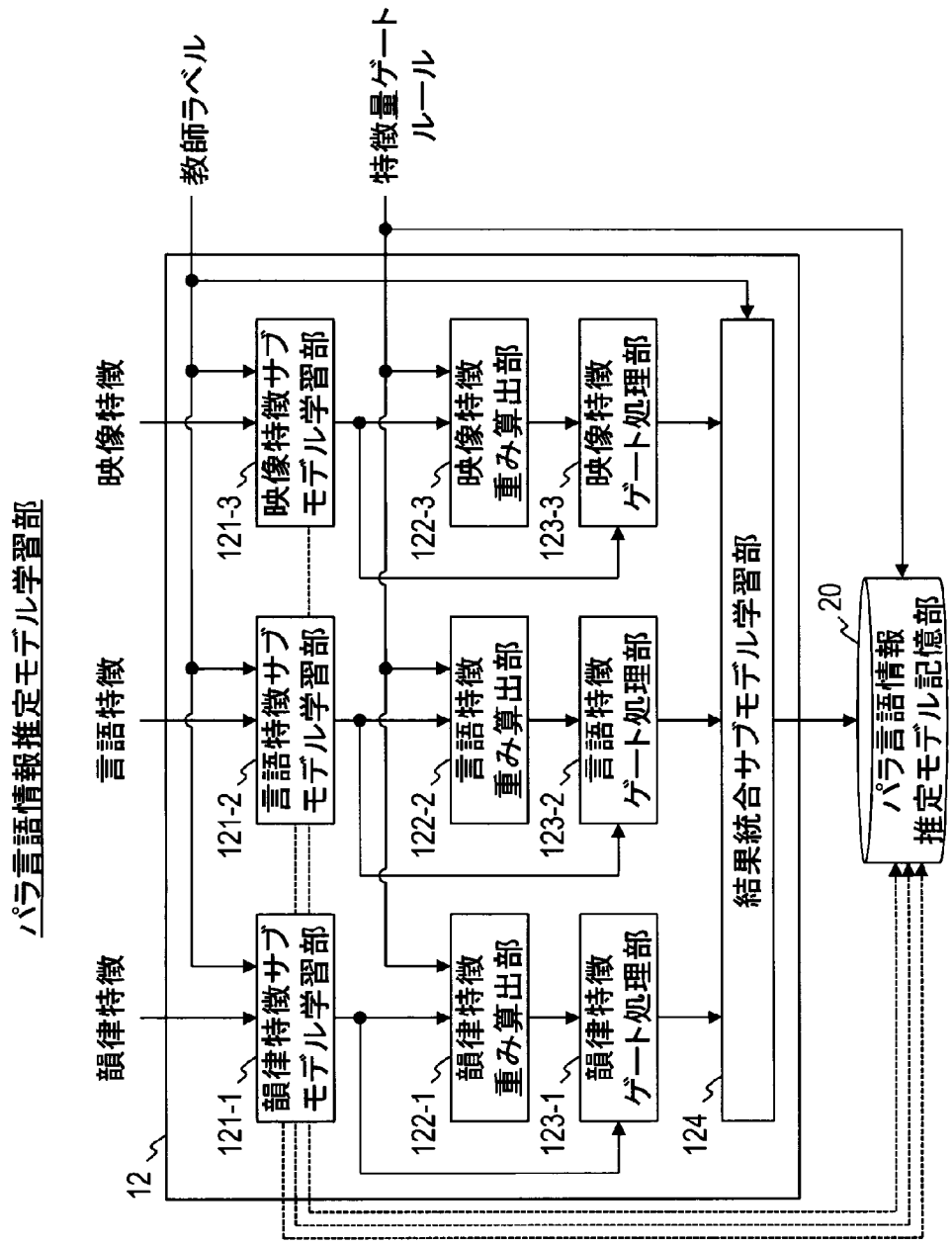
パラ言語情報推定モデル学習部
教師ラベル
特徴量ゲートルール
韻律特徴
言語特徴
映像特徴
12
121-1
韻律特徴サブモデル学習部
121-2
言語特徴サブモデル学習部
121-3
映像特徴サブモデル学習部
122-1
韻律特徴重み算出部
122-2
言語特徴重み算出部
122-3
映像特徴重み算出部
123-1
韻律特徴ゲート処理部
123-2
言語特徴ゲート処理部
123-3
映像特徴ゲート処理部
124
結果統合サブモデル学習部
20
パラ言語情報推定モデル記憶部

図6

[図7]

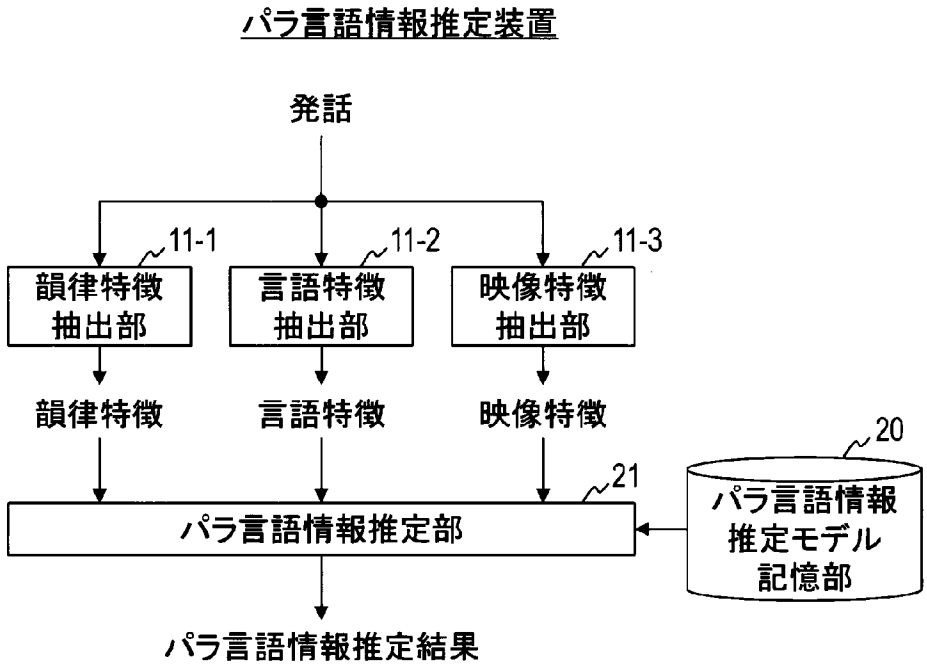
パラ言語情報推定装置
発話
11-1
11-2
11-3
韻律特徴抽出部
言語特徴抽出部
映像特徴抽出部
韻律特徴
言語特徴
映像特徴
20
21
パラ言語情報推定部
パラ言語情報推定モデル記憶部
パラ言語情報推定結果

図7

[図8]

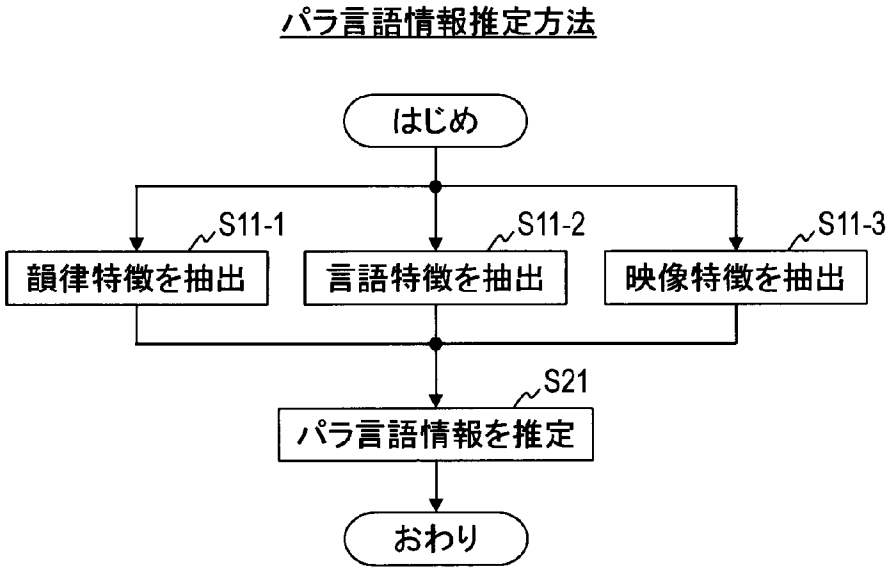
パラ言語情報推定方法
はじめ
S11-1
S11-2
S11-3
韻律特徴を抽出
言語特徴を抽出
映像特徴を抽出
S21
パラ言語情報を推定
おわり

図8

[図9]

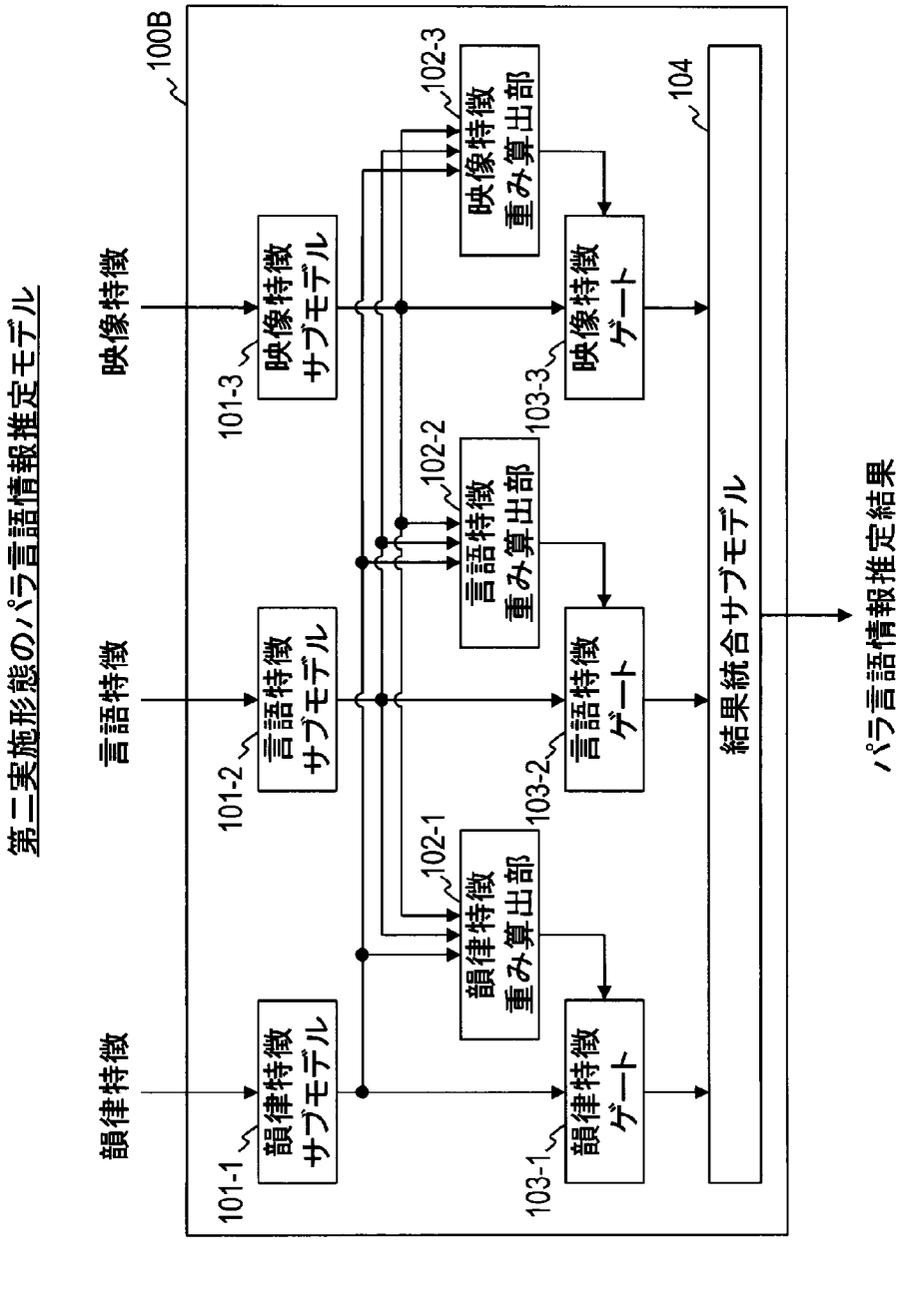
第二実施形態のパラ言語情報推定モデル
100B
韻律特徴
言語特徴
映像特徴
101-1
韻律特徴サブモデル
101-2
言語特徴サブモデル
101-3
映像特徴サブモデル
102-1
韻律特徴重み算出部
102-2
言語特徴重み算出部
102-3
映像特徴重み算出部
103-1
韻律特徴ゲート
103-2
言語特徴ゲート
103-3
映像特徴ゲート
104
結果統合サブモデル
パラ言語情報推定結果

図9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2019/039572

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl. G10L15/10(2006.01)i, G06F17/27(2006.01)i, G06N3/02(2006.01)i, G06N20/00(2019.01)i, G10L15/06(2013.01)i, G10L15/16(2006.01)i, G10L25/63(2013.01)i

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. G10L15/10, G06F17/27, G06N3/02, G06N20/00, G10L15/06, G10L15/16, G10L25/63

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	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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 2018-146898 A (PANASONIC IP MANAGEMENT CO., LTD.) 20 September 2018, entire text, all drawings & US 2018/0257236 A1, entire text, all drawings & EP 3373301 A1 & CN 108573695 A	1-8

☐ 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 06 December 2019 (06.12.2019)	Date of mailing of the international search report 17 December 2019 (17.12.2019)
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１９／０３９５７２

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G10L15/10(2006.01)i, G06F17/27(2006.01)i, G06N3/02(2006.01)i, G06N20/00(2019.01)i, G10L15/06(2013.01)i, G10L15/16(2006.01)i, G10L25/63(2013.01)i

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））

Int.Cl.　G10L15/10, G06F17/27, G06N3/02, G06N20/00, G10L15/06, G10L15/16, G10L25/63

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

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１９年
日本国実用新案登録公報	１９９６－２０１９年
日本国登録実用新案公報	１９９４－２０１９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
Ａ	JP 2018-146898 A (パナソニックＩＰマネジメント株式会社) 2018.09.20, 全文、全図 & US 2018/0257236 A1, 全文、全図 & EP 3373301 A1 & CN 108573695 A	１－８

☐　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矛盾するものではなく、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	国際調査報告の発送日
０６．１２．２０１９	１７．１２．２０１９

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｚ　９１７５
大野　弘
電話番号　０３－３５８１－１１０１　内線　３５９１