

(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6日(06.08.2020)



(10) 国際公開番号

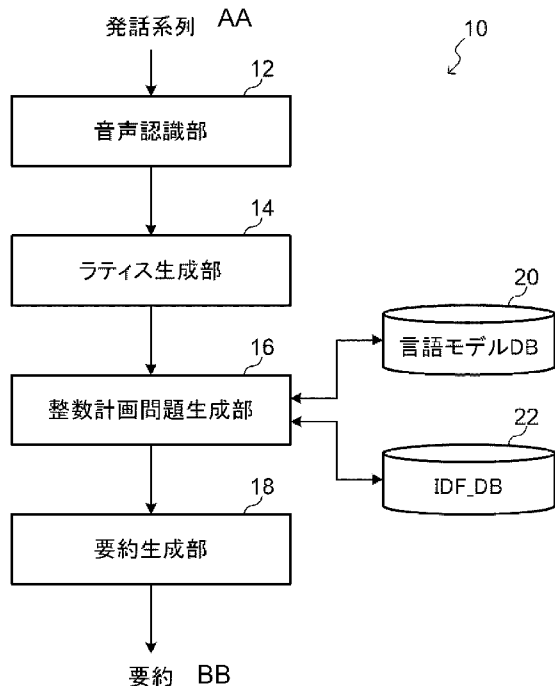
**WO 2020/158409 A1**

(51) 国際特許分類:  
*G10L 15/10* (2006.01) *G10L 15/12* (2006.01)  
(21) 国際出願番号: PCT/JP2020/001250  
(22) 国際出願日: 2020年1月16日(16.01.2020)  
(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 2019-012210 2019年1月28日(28.01.2019) 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: 平尾 努(HIRAO, Tsutomu); 〒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). 小川 厚徳(OGAWA, Atsunori); 〒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). 中谷 智広(NAKATANI, Tomohiro); 〒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). 永田 昌明(NAGATA, Masaaki); 〒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: 特許業務法人太陽国際特許事務所 (TAIYO, NAKAJIMA & KATO); 〒1600022 東京都新宿区新宿4丁目3番17号 Tokyo (JP).

(54) Title: ABSTRACT GENERATION DEVICE, METHOD, PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: 要約生成装置、方法、プログラム、及び記憶媒体



12 Speech recognition unit  
14 Lattice generation unit  
16 Integer programming problem generation unit  
18 Abstract generation unit  
20 Language model DB  
AA Utterance series  
BB Abstract

(57) Abstract: A speech recognition unit (12) converts an utterance series having been input into a confusion network series that is composed of k-best of candidate words as an output of speech recognition; a lattice generation unit (14) designates candidate words as internal nodes and a group of k candidate words for the same speech as an external node from the confusion network series, and generates a lattice series that places edges among internal nodes other than internal nodes contained in the same external node; an integer programming problem generation unit (16) generates an integer

(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))

programming problem for selecting a path to maximize an objective function that includes at least a coverage score for an important word from among paths that follow the internal nodes with the edges placed thereamong in the lattice series; and an abstract generation unit generates a high-quality abstract that has reduced speech recognition errors and lower redundancy by using the candidate words represented by the internal nodes included in the path selected by solving the integer programming problem under constraints associated with the length of the abstract to be generated.

(57) 要約: 音声認識部(12)が、入力された発話系列を、音声認識結果の候補単語の *best* で構成されるコンフュージョンネットワーク系列に変換し、ラティス生成部(14)が、コンフュージョンネットワーク系列から、候補単語を内部ノード、同じ音声に対する *k* 個の候補単語をまとめたものを外部ノードとし、同一の外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成し、整数計画問題生成部(16)が、ラティス系列において、エッジを張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成し、要約生成部が、生成される要約の長さに関する制約の下、整数計画問題を解くことにより選択される経路に含まれる内部ノードが示す候補単語を用いて、音声認識誤りが少なく、かつ冗長性の低い高品質な要約を生成する。

## 明 細 書

### 発明の名称：要約生成装置、方法、プログラム、及び記憶媒体 技術分野

[0001] 開示の技術は、要約生成装置、方法、プログラム、及び記憶媒体に係り、特に、音声認識結果を自動的に要約する要約生成装置、方法、プログラム、及び記憶媒体に関する。

### 背景技術

[0002] 従来の音声認識結果の要約技術は、1 - b e s t の音声認識結果を対象として、文抽出あるいは単語抽出によって要約を生成する。例えば、1 - b e s t の音声認識結果に対して、ある長さの制約の下、単語重要度、音声認識信頼度、言語尤度の和を最大化する部分単語列を要約として生成する手法が提案されている（非特許文献1参照）。

### 先行技術文献

### 非特許文献

[0003] 非特許文献1：C. Hori and S. Furui, "Speech summarization: An approach through word extraction and a method for evaluation," IEICE Trans. on Inf. & Syst., vol. E87-D, no. 1, pp. 1525, Jan. 2004.

### 発明の概要

### 発明が解決しようとする課題

[0004] しかしながら、現状での音声認識技術の性能は高いものの100%ではないため、1 - b e s t の音声認識結果だけを利用して要約を生成すると、要約に認識誤りが含まれる可能性が高くなる。また、従来の音声認識結果の要約技術には、従来のテキスト要約技術によく用いられ、その有効性が実証されている重要単語の被覆という概念が利用されていない。このため、講演のように長く内容に重複があるような発話系列を要約しようとした際、冗長な要約になる可能性が高い。

[0005] 開示の技術は、上記の事情に鑑みてなされたもので、音声認識誤りが少な

く、かつ冗長性の低い高品質な要約を生成することができる要約生成装置、方法、及びプログラム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

[0006] 上記目的を達成するために、開示の技術に係る要約生成装置は、入力された発話系列を、音声認識結果の候補単語の *k* - *b e s t* で構成されるコンフュージョンネットワーク系列に変換する音声認識部と、前記コンフュージョンネットワーク系列から、前記候補単語を内部ノード、同じ音声に対する *k* 個の前記候補単語をまとめたものを外部ノードとし、同一の前記外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成するラティス生成部と、前記ラティス系列において、前記エッジを張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成する整数計画問題生成部と、生成される要約の長さに関する制約の下、前記整数計画問題を解くことにより選択される前記経路に含まれる前記内部ノードが示す前記候補単語を用いて、前記発話系列の要約を生成する要約生成部と、を含んで構成される。

[0007] 開示の技術に係る要約生成装置によれば、音声認識部が、入力された発話系列を、音声認識結果の候補単語の *k* - *b e s t* で構成されるコンフュージョンネットワーク系列に変換し、ラティス生成部が、コンフュージョンネットワーク系列から、候補単語を内部ノード、同じ音声に対する *k* 個の候補単語をまとめたものを外部ノードとし、同一の外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成し、整数計画問題生成部が、ラティス系列において、エッジを張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成し、要約生成部が、生成される要約の長さに関する制約の下、整数計画問題を解くことにより選択される経路に含まれる内部ノードが示す候補単語を用いて、発話系列の要約を生成する。

[0008] このように、1 - b e s t の音声認識結果の候補単語ではなく、k - b e s t の音声認識結果の候補単語を用いることで、音声認識誤りを低減することができる。また、重要単語を多く被覆することで冗長性の低い要約が生成できる。すなわち、開示の技術に係る要約生成装置によれば、音声認識誤りが少なく、かつ冗長性の低い高品質な要約を生成することができる。

[0009] また、前記重要単語の被覆スコアは、前記ラティス系列に含まれる自立語である候補単語のうち、生成される要約に含まれる候補単語の数が多いほど高くなるスコアとすることができる。これにより、目的関数を最大化することで、重要単語としての自立語を多く被覆する要約を生成することができる。

[0010] また、前記目的関数は、さらに、生成される要約に含まれる候補単語の重要度で表される前記内部ノードのスコアと、生成される要約に含まれるエッジの両端の候補単語間のつながりの良さを示す前記エッジのスコアとを含むことができる。また、前記内部ノードのスコアは、前記候補単語の出現頻度及び逆文書頻度と、前記候補単語についての音声認識の信頼度とを含むことができる。また、前記エッジのスコアは、前記エッジの両端の候補単語のバイグラム出現率を含むことができる。これにより、目的関数を最大化することで、重要な単語及び単語間のつながりを含む要約を生成することができる。

[0011] また、開示の技術に係る要約生成方法は、音声認識部と、ラティス生成部と、整数計画問題生成部と、要約生成部とを含む要約生成装置において実行される要約生成方法であって、前記音声認識部が、入力された発話系列を、音声認識結果の候補単語のk - b e s t で構成されるコンフュージョンネットワーク系列に変換し、前記ラティス生成部が、前記コンフュージョンネットワーク系列から、前記候補単語を内部ノード、同じ音声に対するk個の前記候補単語をまとめたものを外部ノードとし、同一の前記外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成し、前記整数計画問題生成部が、前記ラティス系列において、前記エッジを

張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成し、前記要約生成部が、生成される要約の長さに関する制約の下、前記整数計画問題を解くことにより選択される前記経路に含まれる前記内部ノードが示す前記候補単語を用いて、前記発話系列の要約を生成する方法である。

[0012] また、開示の技術に係る要約生成プログラムは、コンピュータを、上記の要約生成装置を構成する各部として機能させるためのプログラムである。

[0013] また、開示の技術に係る記憶媒体は、コンピュータを、上記の要約生成装置を構成する各部として機能させるためのプログラムを記憶した記憶媒体である。

### 発明の効果

[0014] 以上説明したように、開示の技術に係る要約生成装置、方法、プログラム、及び記憶媒体によれば、k - b e s t の音声認識結果の候補単語を用い、重要単語の被覆スコアを含む目的関数を最大化するように要約を生成することで、音声認識誤りが少なく、かつ冗長性の低い高品質な要約を生成することができる。

### 図面の簡単な説明

[0015] [図1]本実施形態に係る要約生成装置の機能ブロック図である

[図2]コンフュージョンネットワーク系列を説明するための図である。

[図3]ラティス系列を説明するための図である。

[図4]内部ノード及びエッジの選択を説明するための図である。

[図5]本実施形態に係る要約生成プログラムによる処理の流れの一例を示すフローチャートである。

### 発明を実施するための形態

[0016] 以下、図面を参照して、開示の技術を実施するための形態の一例について詳細に説明する。

[0017] 本実施形態に係る要約生成装置は、CPU (Central Processing Unit)、RAM (Random Access Memory)、ROM (Read Only Memory)、及びHD

D (Hard Disk Drive) 等を備えたコンピュータとして構成される。ROMには、本実施形態に係る要約生成プログラムが記憶される。なお、要約生成プログラムは、HDDに記憶されてもよい。

[0018] また、要約生成プログラムは、例えば、要約生成装置に予めインストールされていてもよい。この要約生成プログラムは、不揮発性の記憶媒体に記憶して、又は、ネットワークを介して配布して、要約生成装置に適宜インストールすることで実現してもよい。なお、不揮発性の記憶媒体の例としては、CD-ROM (Compact Disc Read Only Memory)、光磁気ディスク、DVD-ROM (Digital Versatile Disc Read Only Memory)、フラッシュメモリ、メモリカード等が挙げられる。

[0019] CPUは、ROMに記憶された要約生成プログラムを読み込んで実行することにより、後述する要約生成装置の各機能部として機能する。

[0020] 図1に示すように、本実施形態に係る要約生成装置10には、例えば、講演などの発話系列であって、U個の発話を含む、すなわち長さUの発話系列（音声データ）が入力される。そして、各機能部により要約生成の処理が実行されて、入力された発話系列の要約（テキストデータ）が出力される。

[0021] 要約生成装置10は、機能的には、図1に示すように、音声認識部12と、ラティス生成部14と、整数計画問題生成部16と、要約生成部18とを含む。また、要約生成装置10の所定の記憶領域、又は、要約生成装置10と通信可能な外部の記憶装置には、言語モデルデータベース(DB)20と、IDF\_\_DB22とが記憶される。言語モデルDB20には、大量のコーパスを用いて予め計算された単語のバイグラム出現率が記憶されている。IDF\_\_DB22には、大量のコーパスを用いて予め計算された単語のidf (Inverse Document Frequency) が記憶されている。

[0022] 音声認識部12は、入力された発話系列をk-e-s-tのコンフュージョンネットワーク系列に変換する。コンフュージョンネットワークとは、複数ある音声認識候補をひとつのネットワークとしてまとめて表現したものである。発話系列に含まれるi番目の発話に対するコンフュージョンネットワー

クの一例を図2に示す。 $w_{i,j,k}$ は、 $i$ 番目のコンフュージョンネットワークの $j$ 番目の単語に対する $k$ 番目の信頼度を持つ認識結果である単語を表す。ノードからノードへ遷移するエッジが複数ある場合、信頼度が高いエッジを選択していくと1-bestの音声認識結果となる。発話系列をコンフュージョンネットワーク系列に変換する手法としては、既存の音声認識システムの手法を用いることができる。

[0023] 音声認識部12は、変換したコンフュージョンネットワーク系列をラティス生成部14へ受け渡す。

[0024] ラティス生成部14は、音声認識部12から受け渡されたコンフュージョンネットワーク系列を、音声認識結果の候補単語を内部ノード、同じ音声に対する $k$ 個の候補単語をまとめたものを外部ノードとするラティス系列に変換する。また、ラティス生成部14は、発話の開始及び終了の各々を表す特別なノードとして、BOU及びEOUというノードを用意する。ラティス生成部14は、任意の内部ノードとそれよりも左側の内部ノード及びBOUの各々との間にエッジを張り、右側の内部ノード及びEOUの各々との間にエッジを張る。同じ外部ノードに属する内部ノード間にはエッジは張らない。ラティス生成部14は、このようにして、BOUからEOUまでの内部ノードを辿る全ての経路を得るためのエッジを張ったラティス系列を生成する。図2に示すコンフュージョンネットワーク系列をラティス系列に変換した例を図3に示す。

[0025] ラティス生成部14は、生成したラティス系列を整数計画問題生成部16へ受け渡す。

[0026] 整数計画問題生成部16は、ラティス生成部14から受け渡されたラティス系列から、生成する要約の文字数の制約の下、目的関数を最大化する内部ノードの経路を選択する整数計画問題を生成する。本実施形態では、内部ノードの重要度の和、エッジの重要度の和、及び重要単語の被覆スコアを最大化する経路を、ラティス系列から選択する整数計画問題を生成する。

[0027] 要約生成の目的関数を、下記(1)式に示す。



[0028] [数1]

$$\begin{aligned} \text{maximize } & \sum_{i=1}^{|U|} \left( \alpha \cdot \sum_{j=1}^{|V_i|} \sum_{k=1}^{|N_{i,j}|} f_{i,j,k} n_{i,j,k} + \beta \cdot \sum_{s=1}^{|V_i|-1} \sum_{p=1}^{|N_{i,s}|} \sum_{t=s+1}^{|V_i|} \sum_{q=1}^{|N_{i,t}|} g_{i,s,p}^{i,t,q} e_{i,s,p}^{i,t,q} \right) \\ & + \sum_{h=1}^{|W|} z_h \end{aligned} \quad (1)$$

$$\text{subject to } \sum_{i=1}^{|U|} \sum_{j=1}^{|V_i|} \sum_{k=1}^{|N_{i,j}|} \ell_{i,j,k} n_{i,j,k} \leq L \quad (2)$$

$$\forall i, j : \sum_{k=1}^{|N_{i,j}|} n_{i,j,k} \leq 1 \quad (3)$$

$$\forall i, j, k, s, p : \sum_{s=1}^{j-1} \sum_{p=1}^{|N_{i,s}|} e_{i,s,p}^{i,j,k} - n_{i,j,k} = 0 \quad (4)$$

$$\forall i, j, k, t, q : \sum_{t=j+1}^{|V_i|} \sum_{q=1}^{|N_{i,t}|} e_{i,j,k}^{i,t,q} - n_{i,j,k} = 0 \quad (5)$$

$$\forall i : \frac{\sum_{j=1}^{|V_i|} \sum_{k=1}^{|N_{i,j}|} n_{i,j,k}}{K} \geq a_{i,1} \quad (6)$$

$$\forall i : 1 - \frac{\sum_{j=1}^{|V_i|} \sum_{k=1}^{|N_{i,j}|} n_{i,j,k}}{|V_i|} \geq a_{i,2} \quad (7)$$

$$\forall i : a_{i,1} + a_{i,2} = 1 \quad (8)$$

$$\forall i, j, k : n_{i,j,k} \in \{0, 1\} \quad (9)$$

$$\forall i, s, t, p, q : e_{i,s,p}^{i,t,q} \in \{0, 1\} \quad (10)$$

$$\forall h : z_h \in \{0, 1\} \quad (11)$$

$$\forall i : a_{i,*} \in \{0, 1\} \quad (12)$$

[0029]  $i$  をラティスのインデックス、 $j$  を  $i$  番目のラティスにおける外部ノードのインデックス、 $k$  を  $i$  番目のラティスの  $j$  番目の外部ノードに含まれる内部ノードのインデックスとする。ラティス集合を  $U$ 、 $i$  番目のラティスにおける外部ノードの集合を  $V_i$ 、 $i$  番目のラティスにおける  $j$  番目の外部ノード

に含まれる内部ノードの集合を  $N_{i,j}$  とする。また、 $W$  を  $U$  に含まれる自立語の集合とする。

[0030] (1) 式の第1項はノードのスコア、第2項はエッジのスコア、第3項は重要単語の被覆スコアを表す。 $n_{i,j,k}$  は  $i$  番目のラティスの  $j$  番目の外部ノードに含まれる  $k$  番目の単語を要約に含めるか否かを表すバイナリ変数、 $f_{i,j,k}$  は  $w_{i,j,k}$  に対する重要度スコアである。 $f_{i,j,k}$  の定義は下記(13)式に示すとおりである。

[0031] [数2]

$$f_{i,j,k} = \text{tfidf}(w_{i,j,k}) + \text{conf}(w_{i,j,k}). \quad (13)$$

[0032]  $\text{tfidf}()$  は単語の  $\text{tfidf}$  スコアであり、 $\text{tf}$  (Term Frequency) は発話系列における単語の出現頻度である。 $\text{idf}$  は  $\text{IDF\_DB22}$  から取得する。 $\text{conf}()$  は単語の認識信頼度スコアであり、音声認識部12が音声認識を行う際に得られる値である。

[0033]  $e_{i,s,p}^{i,t,q}$  は、 $w_{i,s,p}$  と  $w_{i,t,q}$  との間のエッジを要約に含めるか否かのバイナリ変数である。 $g_{i,s,p}^{i,t,q}$  は、 $e_{i,s,p}^{i,t,q}$  の重要度スコアであり、言語モデル  $\text{DB20}$  から取得される単語  $w_{i,s,p}$  と単語  $w_{i,t,q}$  とのバイグラム出現確率とすることができる。 $g_{i,s,p}^{i,t,q}$  の定義は下記(14)式に示すとおりである。なお、 $g_{i,s,p}^{i,t,q}$  は、単語と単語とのつながりの良さをスコア化したものであれば、(14)式に示す例に限定されない。

[0034] [数3]

$$g_{i,s,p}^{i,t,q} = P_{\text{bigram}}(w_{i,t,q} | w_{i,s,p}). \quad (14)$$

[0035]  $\alpha$  及び  $\beta$  は、ノードのスコアの和とエッジのスコアの和とを調整するパラメータであり、検証用データを用いて最適値を決める。 $z_h$  は、 $W$  における  $h$  番目の自立語が要約に含まれれば1、そうでない場合に0となるバイナリ変数であり、このスコアが高いことは多くの重要単語を被覆することを表す。つまり、多くの重要単語を被覆する作用があるため、生成された要約の冗長

性は低くなる。

- [0036] (2) 式は、要約長に関する制約であり、要約の文字数が $L$ 以下であることを保証する。(3) 式は、任意の外部ノードからは高々ひとつの内部ノード(単語)しか選択されないことを表す。(4) 式及び(5) 式は、図4に示すように、任意の内部ノードを要約に含めるためには、その両端のエッジを必ずひとつずつ(入力エッジからひとつ、出力エッジからひとつ)選択しなければならないことを表す。(6) 式及び(7) 式は、ひとつの発話から短い経路を抽出することを避けるための制約であり、 $i$  番目のラティスから単語列を要約として選択するためには、その単語数が $K$ 以上でなければならないことを示す。
- [0037] また、上記以外にも言語知識として要約文に必要な単語があれば、その単語に対応する  $n_{i,j,k}$  を1にしておけばよい。
- [0038] 整数計画問題生成部16は、生成した整数計画問題を要約生成部18へ受け渡す。
- [0039] 要約生成部18は、整数計画問題生成部16から受け渡された整数計画問題を、既存の専用ソルバを用いて解き、 $n_{i,j,k}=1$  となる  $w_{i,j,k}$  を抽出することで要約を生成し、生成した要約を出力する。
- [0040] 次に、図5を参照して、本実施形態に係る要約生成装置10の作用を説明する。なお、図5は、本実施形態に係る要約生成プログラムによる要約生成処理の流れの一例を示すフローチャートである。
- [0041] ステップS12で、音声認識部12が、入力された発話系列を  $k-best$  のコンフュージョンネットワーク系列に変換する。音声認識部12は、変換したコンフュージョンネットワーク系列をラティス生成部14へ受け渡す。
- [0042] 次に、ステップS14で、ラティス生成部14が、音声認識部12から受け渡されたコンフュージョンネットワーク系列を、音声認識結果の候補単語を内部ノード、同じ音声に対する $k$ 個の候補単語をまとめたものを外部ノードとするラティス系列に変換する。また、ラティス生成部14は、発話の開

始及び終了の各々を表すB O U及びE O Uというノードを用意し、同じ外部ノードに属する内部ノード間以外の内部ノード間にエッジを張る。ラティス生成部14は、生成したラティス系列を整数計画問題生成部16へ受け渡す。

[0043] 次に、ステップS16で、整数計画問題生成部16が、ラティス生成部14から受け渡されたラティス系列から、生成する要約の文字数の制約の下、内部ノードのスコア、エッジのスコア、及び重要単語の被覆スコアを含む目的関数を最大化する内部ノードの経路を選択する整数計画問題を生成する。整数計画問題生成部16は、生成した整数計画問題を要約生成部18へ受け渡す。

[0044] 次に、ステップS18で、要約生成部18が、整数計画問題生成部16から受け渡された整数計画問題を、既存の専用ソルバを用いて解き、ラティス系列から選択された経路に含まれる内部ノードが示す候補単語を用いて要約を生成し、生成した要約を出力する。そして、要約生成処理は終了する。

[0045] 以上説明したように、本実施形態に係る要約生成装置によれば、入力された発話系列を、音声認識結果の候補単語のk - b e s tで構成されるコンフュージョンネットワーク系列に変換し、コンフュージョンネットワーク系列から、候補単語を内部ノード、同じ音声に対するk個の候補単語をまとめたものを外部ノードとし、同一の外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成し、ラティス系列において、エッジを張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成し、生成される要約の長さに関する制約の下、整数計画問題を解くことにより選択される経路に含まれる内部ノードが示す候補単語を用いて、発話系列の要約を生成する。これにより、音声認識誤りが少なく、かつ冗長性の低い高品質な要約を生成することができる。

[0046] なお、上記実施形態で説明した要約生成装置の各々の構成及び処理は、一例であり、主旨を逸脱しない範囲内において状況に応じて変更してもよい。

[0047] また、上記実施形態で説明したプログラムの処理の流れも、一例であり、主旨を逸脱しない範囲内において不要なステップを削除したり、新たなステップを追加したり、処理順序を入れ替えたりしてもよい。

[0048] また、上記実施形態では、プログラムを実行することにより、実施形態に係る処理がコンピュータを利用してソフトウェア構成により実現される場合について説明したが、これに限らない。実施形態は、例えば、ハードウェア構成や、ハードウェア構成とソフトウェア構成との組み合わせによって実現してもよい。

### 符号の説明

- [0049] 10 要約生成装置  
12 音声認識部  
14 ラティス生成部  
16 整数計画問題生成部  
18 要約生成部  
20 言語モデルDB  
22 IDF\_\_DB

## 請求の範囲

- [請求項1]           入力された発話系列を、音声認識結果の候補単語の  $k - best$  で構成されるコンフュージョンネットワーク系列に変換する音声認識部と、
- 前記コンフュージョンネットワーク系列から、前記候補単語を内部ノード、同じ音声に対する  $k$  個の前記候補単語をまとめたものを外部ノードとし、同一の前記外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成するラティス生成部と、
- 前記ラティス系列において、前記エッジを張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成する整数計画問題生成部と、
- 生成される要約の長さに関する制約の下、前記整数計画問題を解くことにより選択される前記経路に含まれる前記内部ノードが示す前記候補単語を用いて、前記発話系列の要約を生成する要約生成部と、
- を含む要約生成装置。
- [請求項2]           前記重要単語の被覆スコアは、前記ラティス系列に含まれる自立語である候補単語のうち、生成される要約に含まれる候補単語の数が多いほど高くなるスコアである請求項1に記載の要約生成装置。
- [請求項3]           前記目的関数は、さらに、生成される要約に含まれる候補単語の重要度で表される前記内部ノードのスコアと、生成される要約に含まれるエッジの両端の候補単語間のつながりの良さを示す前記エッジのスコアとを含む請求項1又は請求項2に記載の要約生成装置。
- [請求項4]           前記内部ノードのスコアは、前記候補単語の出現頻度及び逆文書頻度と、前記候補単語についての音声認識の信頼度とを含む請求項3に記載の要約生成装置。
- [請求項5]           前記エッジのスコアは、前記エッジの両端の候補単語のバイグラム

出現率を含む請求項3に記載の要約生成装置。

[請求項6]

音声認識部と、ラティス生成部と、整数計画問題生成部と、要約生成部とを含む要約生成装置において実行される要約生成方法であって、

前記音声認識部が、入力された発話系列を、音声認識結果の候補単語の *k* - *b e s t* で構成されるコンフュージョンネットワーク系列に変換し、

前記ラティス生成部が、前記コンフュージョンネットワーク系列から、前記候補単語を内部ノード、同じ音声に対する *k* 個の前記候補単語をまとめたものを外部ノードとし、同一の前記外部ノードに含まれる内部ノード間以外の内部ノード間にエッジを張ったラティス系列を生成し、

前記整数計画問題生成部が、前記ラティス系列において、前記エッジを張られた内部ノードを辿る経路のうち、少なくとも重要単語の被覆スコアを含む目的関数を最大化する経路を選択するための整数計画問題を生成し、

前記要約生成部が、生成される要約の長さに関する制約の下、前記整数計画問題を解くことにより選択される前記経路に含まれる前記内部ノードが示す前記候補単語を用いて、前記発話系列の要約を生成する

要約生成方法。

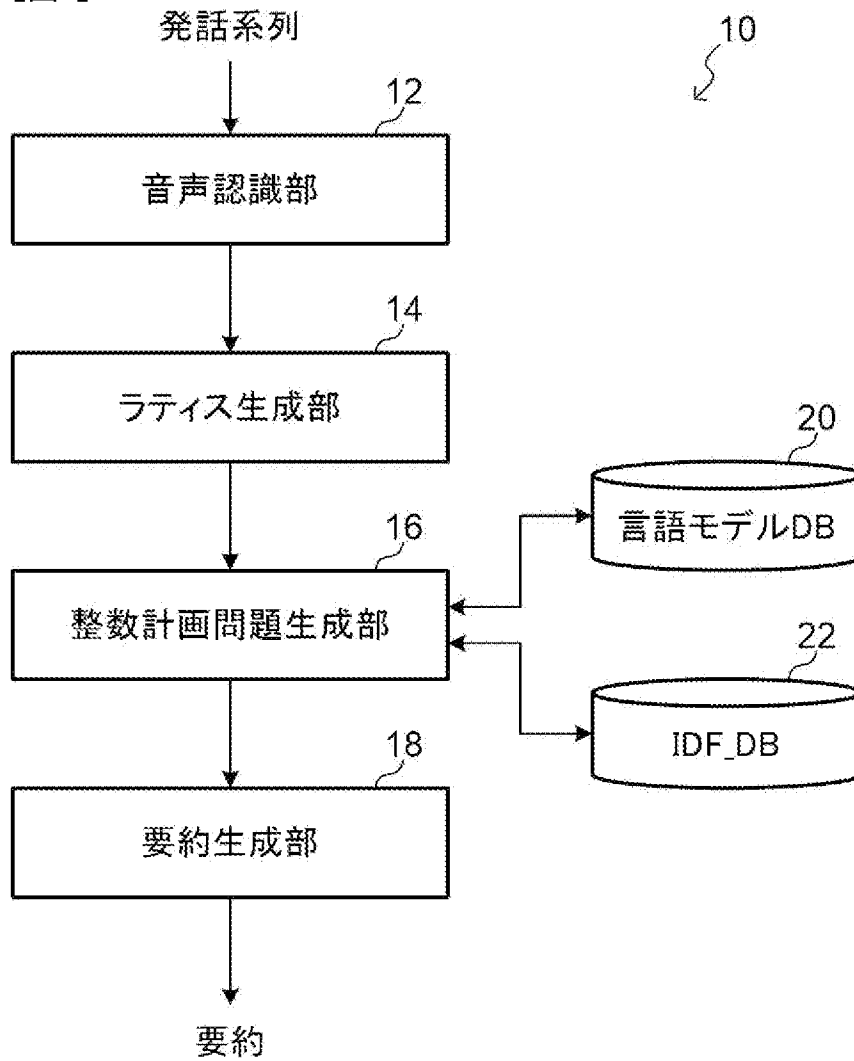
[請求項7]

コンピュータを、請求項1～請求項5のいずれか1項に記載の要約生成装置を構成する各部として機能させるための要約生成プログラム。

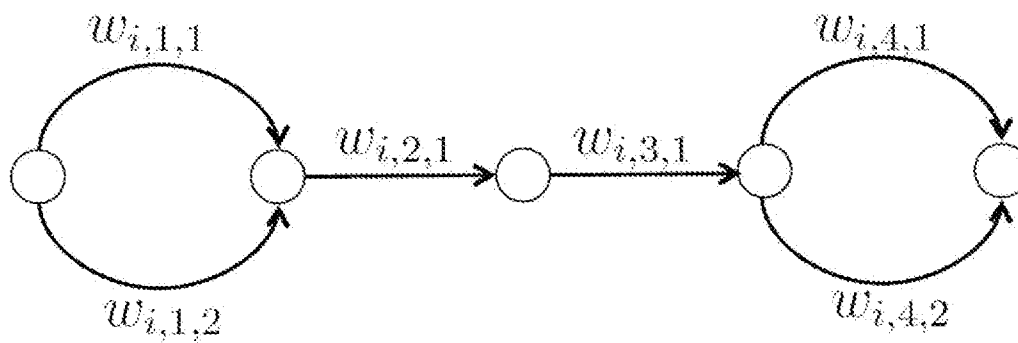
[請求項8]

コンピュータを、請求項1～請求項5のいずれか1項に記載の要約生成装置を構成する各部として機能させるための要約生成プログラムを記憶した記憶媒体。

[図1]

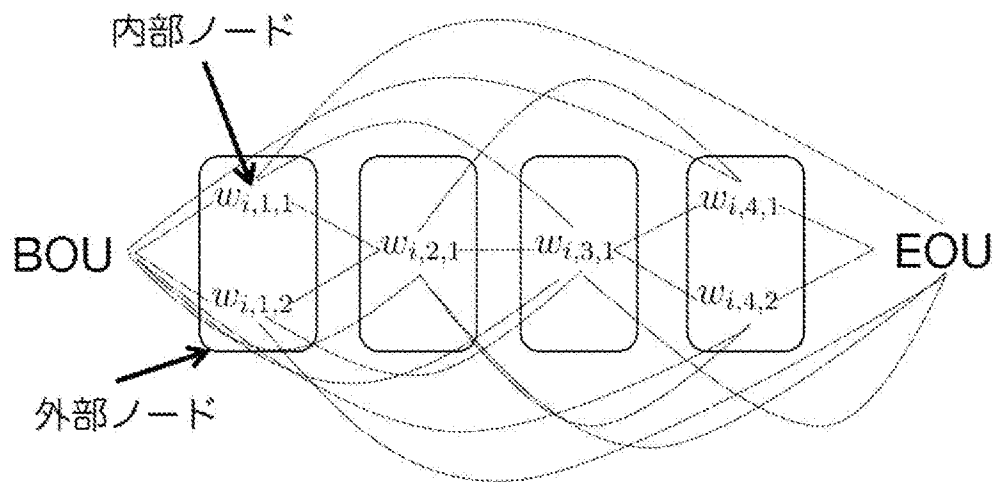


[図2]

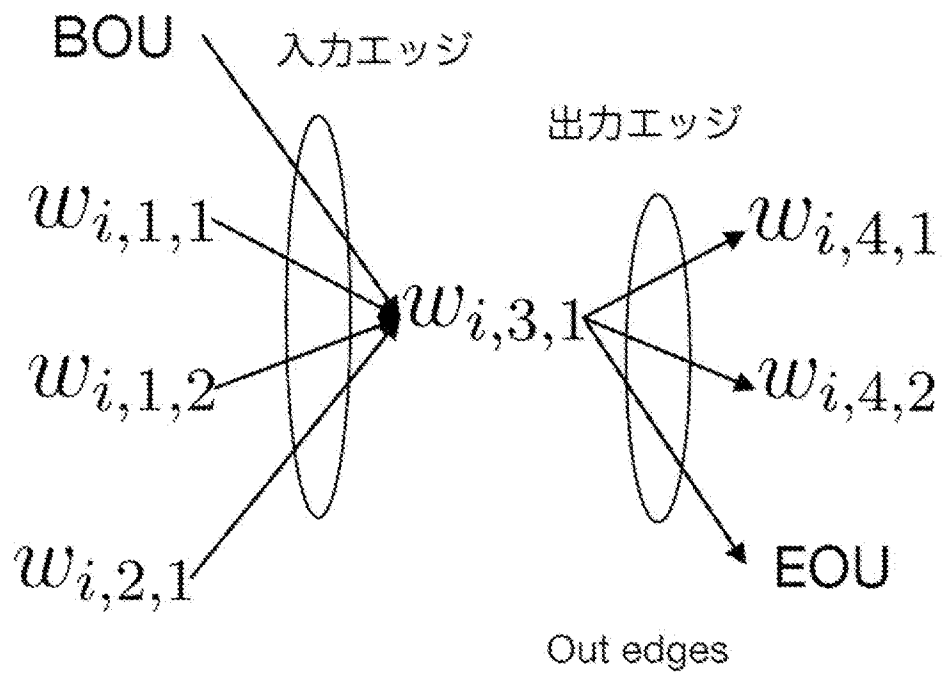




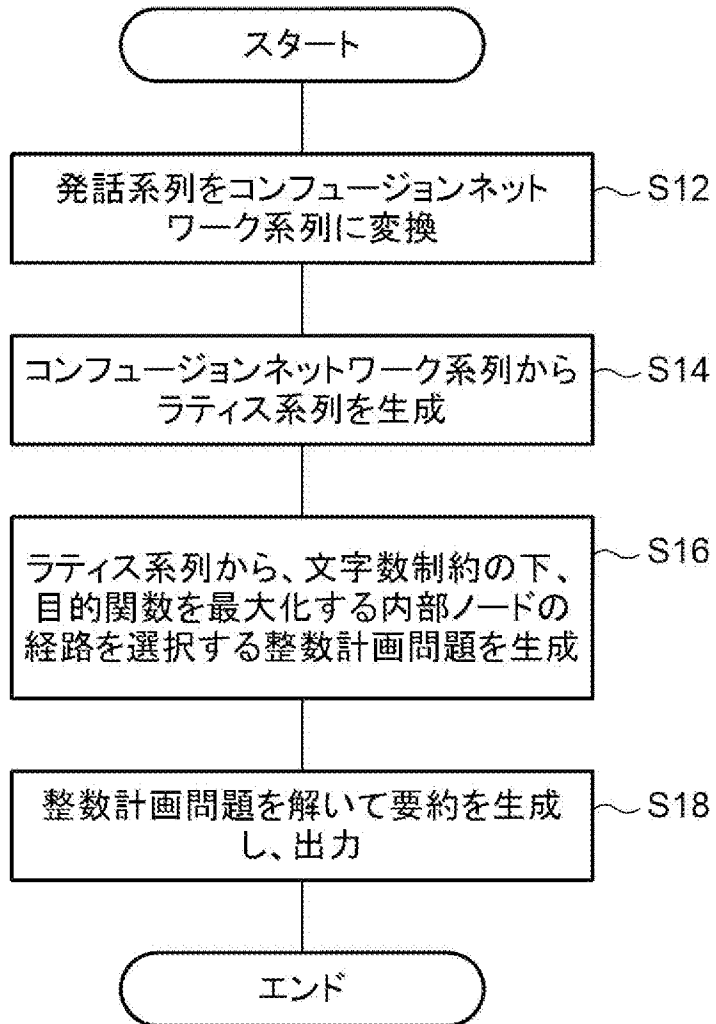
[図3]



[図4]



[図5]



# INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/001250

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G10L15/10 (2006.01) i, G10L15/12 (2006.01) i  
FI: G10L15/10500T, G10L15/12Z

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, G10L15/12

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)

IEEE Xplore

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                    | Relevant to claim No. |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| A         | LIU, F. et al., Towards abstractive speech summarization: Exploring unsupervised and supervised approaches for spoken utterance compression, IEEE Transactions on Audio, Speech, and Language Processing, vol. 21, no. 7, 27 March 2013, <DOI:10.1109/TASL.2013.2255279>, IV Spoken utterance compression                             | 1-8                   |
| A         | LIU, Y. et al., Using N-best recognition output for extractive summarization and keyword extraction in meeting speech, [online], ICASSP 2010, 28 July 2010, etrieved from the Internet: <URL:https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5494972&tag=1>, <DOI:10.1109/ICASSP.2010.5494972>, 4.1 Extractive summarization | 1-8                   |
| A         | JP 2012-37797 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 23.02.2012 (2012-02-23)                                                                                                                                                                                                                                                  | 1-8                   |

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                | "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                                              |
| "A" document defining the general state of the art which is not considered to be of particular relevance                                                                | "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                                                                     |
| "E" earlier application or patent but published on or after the international filing date                                                                               | "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 |
| "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) | "&" document member of the same patent family                                                                                                                                                                                                    |
| "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                                                                  |                                                                                                                                                                                                                                                  |

Date of the actual completion of the international search  
11.03.2020

Date of mailing of the international search report  
24.03.2020

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

Authorized officer  
  
Telephone No.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

PCT/JP2020/001250

JP 2012-37797 A      23.02.2012      (Family: none)

| A. 発明の属する分野の分類（国際特許分類（IPC））<br>G10L 15/10(2006.01)i; G10L 15/12(2006.01)i<br>FI: G10L15/10 500T; G10L15/12 Z                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                                                                      |                                                                       |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| 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））<br>G10L15/10; G10L15/12                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                                                      |                                                                       |
| 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<br>日本国実用新案公報 1922-1996年<br>日本国公開実用新案公報 1971-2020年<br>日本国実用新案登録公報 1996-2020年<br>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）<br>IEEE Xplore                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                      |                                                                       |
| C. 関連すると認められる文献                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                      |                                                                       |
| 引用文献の<br>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<br>請求項の番号                                                        |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                     | FEI, Liu et al., Towards Abstractive Speech Summarization: Exploring Unsupervised and Supervised Approaches for Spoken Utterance Compression, IEEE Transactions on Audio, Speech, and Language Processing, Vol. 21, No. 7, 2013.03.27, <DOI: 10.1109/TASL.2013.2255279>, IV SPOKEN UTTERANCE COMPRESSION<br>IV SPOKEN UTTERANCE COMPRESSION                          | 1-8                                                                   |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                     | YANG, Liu et al., USING N-BEST RECOGNITION OUTPUT FOR EXTRACTIVE SUMMARIZATION AND KEYWORD EXTRACTION IN MEETING SPEECH, [online], ICASSP 2010, 2010.07.28, etrieved from theInternet:<URL:https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5494972&tag=1>, <DOI: 10.1109/ICASSP.2010.5494972>, 4.1 Extractive Summarization<br>4.1 Extractive Summarization | 1-8                                                                   |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                     | JP 2012-37797 A（日本電信電話株式会社）23.02.2012（2012-02-23）                                                                                                                                                                                                                                                                                                                    | 1-8                                                                   |
| <input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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<br>“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<br>“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<br>“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）<br>“O” 口頭による開示、使用、展示等に言及する文献<br>“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<br>“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<br>“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<br>“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<br>“&” 同一パテントファミリー文献 |                                                                                                                                                                                                                                                                                                                                                                      |                                                                       |
| 国際調査を完了した日<br>11.03.2020                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                                                      | 国際調査報告の発送日<br>24.03.2020                                              |
| 名称及びあて先<br>日本国特許庁(ISA/JP)<br>〒100-8915<br>日本国<br>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）<br><br>山下 剛史 5Z 1162<br><br>電話番号 03-3581-1101 内線 3589 |

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/JP2020/001250

| 引用文献            | 公表日        | パテントファミリー文献 | 公表日 |
|-----------------|------------|-------------|-----|
| JP 2012-37797 A | 23.02.2012 | (ファミリーなし)   |     |