

(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  
2023年3月16日(16.03.2023)



(10) 国際公開番号  
**WO 2023/038004 A1**

(51) 国際特許分類:  
A61B 1/045 (2006.01)

(21) 国際出願番号: PCT/JP2022/033260

(22) 国際出願日: 2022年9月5日(05.09.2022)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

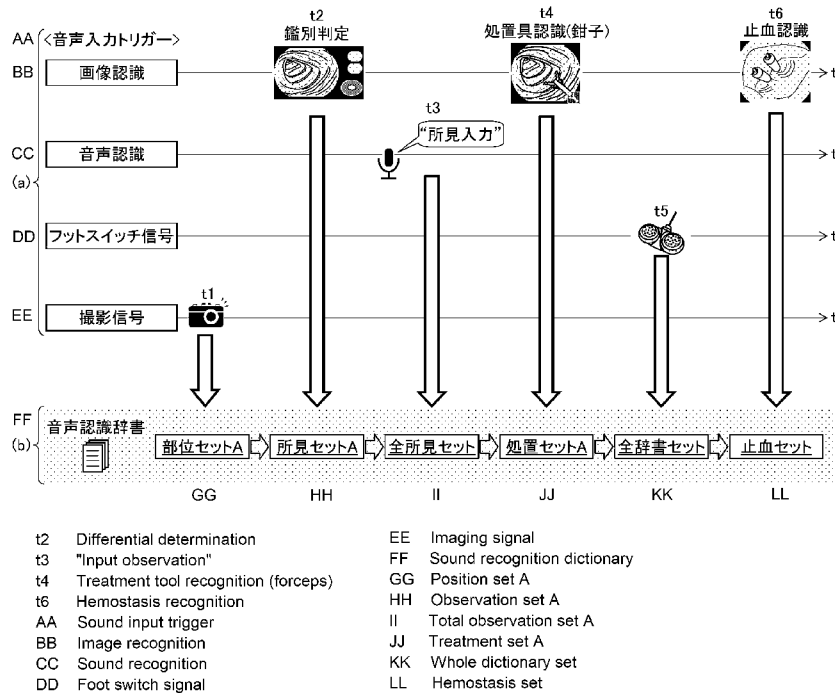
(30) 優先権データ:  
特願 2021-146308 2021年9月8日(08.09.2021) JP

(71) 出願人: 富士フイルム株式会社 (FUJIFILM CORPORATION) [JP/JP]; 〒1068620 東京都港区西麻布2丁目2番30号 Tokyo (JP).

(72) 発明者: 木村 裕哉(KIMURA, Yuya); 〒2588538 神奈川県足柄上郡開成町宮台798番地 富士フイルム株式会社内 Kanagawa (JP). 堀 悠磨(HORI, Yuma); 〒2588538 神奈川県足柄上郡開成町宮台798番地 富士フイルム株式会社内 Kanagawa (JP). 小林 達矢(KOBAYASHI, Tatsuya); 〒2588538 神奈川県足柄上郡開成町宮台798番地 富士フイルム株式会社内 Kanagawa (JP). 坂口 雄一(SAKAGUCHI, Yuichi); 〒2588538 神奈川県足柄上郡開成町宮台798番地 富士フイルム株式会社内 Kanagawa (JP). 原田 憲一(HARADA, Kenichi); 〒2588538 神奈川県足柄上郡開成町宮台798番地 富士フイルム株式会社内 Kanagawa (JP). 杉山 武(SUGIYAMA, Takeru); 〒2588538

(54) Title: ENDOSCOPE SYSTEM, MEDICAL INFORMATION PROCESSING DEVICE, MEDICAL INFORMATION PROCESSING METHOD, MEDICAL INFORMATION PROCESSING PROGRAM, AND STORAGE MEDIUM

(54) 発明の名称: 内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体



(57) Abstract: One embodiment of the invention of the present disclosure provides an endoscope system, medical information processing device, medical information processing method, medical information processing program, and storage medium, capable of improving recognition accuracy of input sound. The endoscope system according to one aspect of the



WO 2023/038004 A1

神奈川県足柄上郡開成町宮台 7 9 8 番地 富士フィルム株式会社内 Kanagawa (JP).

- (74) 代理人: 松浦 憲三 (MATSUURA, Kenzo);  
〒1630223 東京都新宿区西新宿二丁目 6 番 1 号 新宿住友ビル 2 3 階 新都心国際特許事務所 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.
- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

present invention comprises a sound input device, an image sensor for imaging an object, and a processor. The processor causes the image sensor to image an object over time and thereby acquires a plurality of medical images, receives input from a sound input trigger while imaging the plurality of medical images, sets a sound recognition dictionary according to the sound input trigger when the sound input trigger is input, and recognizes, using the set sound recognition dictionary, the sound input into the sound input device after setting the sound recognition dictionary.

(57) 要約: 本開示の技術に係る一つの実施形態は、音声入力の認識精度を向上させることができる内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体を提供する。本発明の一の態様に係る内視鏡システムは、音声入力装置と、被写体を撮影するイメージセンサと、プロセッサと、を備える内視鏡システムであって、プロセッサは、イメージセンサに被写体を時系列に撮影させることで複数の医療画像を取得し、複数の医療画像の撮影中に音声入力トリガの入力を受け付け、音声入力トリガが入力された場合に、音声入力トリガに応じた音声認識辞書を設定し、設定がされた以降に音声入力装置に入力された音声、設定された音声認識辞書を用いて音声認識する。

## 明 細 書

発明の名称：

内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体

### 技術分野

[0001] 本発明は、音声入力及び音声認識を行う内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体に関する。

### 背景技術

[0002] 医療画像を用いた検査や診断支援を行う技術分野では、ユーザが入力した音声を認識し、認識結果に基づく処理を行うことが知られている。例えば、特許文献1には、内視鏡を音声入力で操作することが記載されている。また、特許文献2には、レポート作成用の音声入力を行うことが記載されている。

### 先行技術文献

#### 特許文献

[0003] 特許文献1：特開平8-052105号公報  
特許文献2：特開2004-102509号公報

### 発明の概要

#### 発明が解決しようとする課題

[0004] 医療画像を用いた検査中に音声入力を行う場合、シーンによらずに全ての単語を認識可能とすると単語間の相互誤認識が増えて操作性が低下するおそれがある。しかしながら、上述した特許文献1，2のような従来技術は、このような課題を十分考慮したものではなかった。

[0005] 本発明はこのような事情に鑑みてなされたもので、医療画像に関する音声認識の精度を向上させることができる内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体を提供すること

を目的とする。

### 課題を解決するための手段

- [0006] 上述した目的を達成するため、本発明の第1の態様に係る内視鏡システムは、音声入力装置と、被写体を撮影するイメージセンサと、プロセッサと、を備える内視鏡システムであって、プロセッサは、イメージセンサが被写体を時系列に撮影することで得られた複数の医療画像を取得し、複数の医療画像の撮影中に音声入力トリガの入力を受け付け、音声入力トリガが入力された場合に、音声入力トリガに応じた音声認識辞書を設定し、設定がされた以降に音声入力装置に入力された音声を、設定された音声認識辞書を用いて音声認識する。第1の態様では、音声入力トリガに応じた音声認識辞書を設定し、設定された音声認識辞書を用いて音声認識するので、音声認識のシーンに合わせた音声認識辞書を用い、医療画像に関する音声認識の精度を向上させることができる。
- [0007] 第2の態様に係る内視鏡システムは第1の態様において、プロセッサは、音声認識において、設定された音声認識辞書に登録されている登録単語のみを認識し、登録単語についての音声認識の結果を出力装置に出力させる。第2の態様によれば、設定された音声認識辞書に登録されている登録単語のみを音声認識するので、認識精度を高めることができる。
- [0008] 第3の態様に係る内視鏡システムは第1の態様において、プロセッサは、音声認識において、設定された音声認識辞書に登録されている登録単語及び特定の単語を認識し、認識した単語のうち登録単語についての音声認識の結果を出力装置に出力させる。なお、「特定の単語」の例としては音声入力装置に対するウェイクワードを挙げるができるが、「特定の単語」はこれに限定されるものではない。
- [0009] 第4の態様に係る内視鏡システムは第1から第3の態様のいずれか1つにおいて、プロセッサは、複数の医療画像に特定の被写体が含まれているかを画像認識により判定し、特定の被写体が含まれていることを示す判定結果を音声入力トリガとして受け付ける。

- [0010] 第5の態様に係る内視鏡システムは第1から第4の態様のいずれか1つにおいて、プロセッサは、複数の医療画像に特定の被写体が含まれているかを画像認識により判定し、特定の被写体が含まれていると判定した場合に、特定の被写体を鑑別し、特定の被写体に対する鑑別結果の出力を音声入力トリガとして受け付ける。
- [0011] 第6の態様に係る内視鏡システムは第4または第5の態様において、プロセッサは、複数の医療画像に複数の種類の特定の被写体が含まれているかを、複数の種類の特定の被写体にそれぞれ対応する複数の画像認識により判定し、複数の種類の特定の被写体のうち、複数の画像認識のいずれかにより複数の医療画像に含まれていると判定された特定の被写体の種類に対応する音声認識辞書を設定する。
- [0012] 第7の態様に係る内視鏡システムは第6の態様において、プロセッサは、複数の医療画像に複数の特定の被写体が含まれているかを画像認識により判定し、複数の特定の被写体のうち、複数の医療画像に含まれていると判定された特定の被写体に対応する音声認識辞書を設定する。
- [0013] 第8の態様に係る内視鏡システムは第4から第7の態様のいずれか1つにおいて、プロセッサは、機械学習により構成された画像認識器を用いて画像認識を行う。
- [0014] 第9の態様に係る内視鏡システムは第4から第8の態様のいずれか1つにおいて、プロセッサは、複数の医療画像のうち特定の被写体が写っていると判断された医療画像と、特定の被写体の画像認識による判定の結果と、音声認識の結果と、を関連付けて記録装置に記録させる。
- [0015] 第10の態様に係る内視鏡システムは第4から第9の態様のいずれか1つにおいて、プロセッサは、病変、病変候補領域、ランドマーク、処置後領域、処置具、または止血具のうち少なくとも1つを特定の被写体と判断する。
- [0016] 第11の態様に係る内視鏡システムは第4から第10の態様のいずれか1つにおいて、プロセッサは、設定された音声認識辞書による音声認識を、設定がされた以降のあらかじめ決められた条件を満たす期間において実行する

- 。
- [0017] 第12の態様に係る内視鏡システムは第11の態様において、プロセッサは、画像認識を行った画像認識器ごとに期間を設定する。
- [0018] 第13の態様に係る内視鏡システムは第11または第12の態様において、プロセッサは、音声入力トリガの種類に応じて期間を設定する。
- [0019] 第14の態様に係る内視鏡システムは第11から第13の態様のいずれか1つにおいて、プロセッサは、期間の残り時間を表示装置に画面表示させる。
- 。
- [0020] 第15の態様に係る内視鏡システムは第1から第14の態様のいずれか1つにおいて、プロセッサは、部位情報、所見情報、処置情報、及び止血情報について音声認識を行う。
- [0021] 第16の態様に係る内視鏡システムは第1から第15の態様のいずれか1つにおいて、プロセッサは、複数の医療画像の撮影開始指示、複数の医療画像に対する画像認識の結果の出力、鑑別モードへの切替操作、内視鏡システムに接続された操作デバイスに対する操作、音声入力装置に対するウェイクワードの入力のいずれかがなされた場合に音声入力トリガが入力されたと判断する。
- [0022] 第17の態様に係る内視鏡システムは第1から第16の態様のいずれか1つにおいて、プロセッサは、音声認識の結果を表示装置に表示させる。
- [0023] 上述した目的を達成するため、本発明の第18の態様に係る医療情報処理装置は、プロセッサを備える医療情報処理装置であって、プロセッサは、イメージセンサが被写体を時系列に撮影することで得られた複数の医療画像を取得し、複数の医療画像の入力中に音声入力トリガの入力を受け付け、音声入力トリガが入力された場合に、音声入力トリガに応じた音声認識辞書を設定し、設定がされた以降に音声入力装置に入力された音声を、設定された音声認識辞書を用いて音声認識する。第18の態様によれば、第1の態様と同様に医療画像に関する音声認識の精度を向上させることができる。
- [0024] 上述した目的を達成するため、本発明の第19の態様に係る医療情報処理

方法は、音声入力装置と、被写体を撮影するイメージセンサと、プロセッサと、を備える内視鏡システムにより実行される医療情報処理方法であって、プロセッサが、イメージセンサが被写体を時系列に撮影することで得られた複数の医療画像を取得し、複数の医療画像の撮影中に音声入力トリガの入力を受け付け、音声入力トリガが入力された場合に、音声入力トリガに応じた音声認識辞書を設定し、設定がされた以降に音声入力装置に入力された音声を、設定された音声認識辞書を用いて音声認識する。第19の態様によれば、第1、第18の態様と同様に、医療画像に関する音声入力の認識精度を向上させることができる。第19の態様において、第2から第17の態様と同様の構成を有していてもよい。

[0025] 上述した目的を達成するため、本発明の第20の態様に係る医療情報処理プログラムは、音声入力装置と、被写体を撮影するイメージセンサと、プロセッサと、を備える内視鏡システムに医療情報処理方法を実行させる医療情報処理プログラムであって、医療情報処理方法において、プロセッサは、イメージセンサが被写体を時系列に撮影することで得られた複数の医療画像を取得し、複数の医療画像の撮影中に音声入力トリガの入力を受け付け、音声入力トリガが入力された場合に、音声入力トリガに応じた音声認識辞書を設定し、設定がされた以降に音声入力装置に入力された音声を、設定された音声認識辞書を用いて音声認識する。第20の態様によれば、第1、第18、第19の態様と同様に、医療画像に関する音声認識の精度を向上させることができる。

[0026] 第20の態様に係る医療情報処理プログラムが実行させる医療情報処理方法は、第2～第17の態様と同様の構成を備えていてもよい。

[0027] 上述した目的を達成するため、本発明の第21の態様に係る記録媒体は、非一時的かつ有体の記録媒体であって、第20の態様に係る医療情報処理プログラムのコンピュータ読み取り可能なコードが記録された記録媒体である。第21の態様において、「非一時的 (non-transitory) かつ有体 (tangible) の記録媒体」の例としては、各種の光磁気記録装置及び半導体メモリを挙

げることができる。この「非一時的かつ有体の記録媒体」は、搬送波信号そのもの、及び伝播信号そのもののような非有体の記録媒体を含まない。

[0028] なお、第21の態様において、記録媒体にコードが記録される医療情報処理プログラムは、第2～第17の態様と同様の処理を行う医療情報処理プログラムを内視鏡システムまたは医療情報処理装置に実行させるものでもよい。

## 発明の効果

[0029] 本発明に係る内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体によれば、医療画像に関する音声認識の精度を向上させることができる。

## 図面の簡単な説明

[0030] [図1]図1は、第1の実施形態に係る内視鏡画像診断システムの概略構成を示す図である。

[図2]図2は、内視鏡システムの概略構成を示す図である。

[図3]図3は、内視鏡の概略構成を示す図である。

[図4]図4は、先端部の端面の構成の一例を示す図である。

[図5]図5は、内視鏡画像生成装置の主な機能を示すブロック図である。

[図6]図6は、内視鏡画像処理装置の主な機能を示すブロック図である。

[図7]図7は、画像認識処理部の主な機能を示すブロック図である。

[図8]図8は、検査中の画面表示の一例を示す図である。

[図9]図9は、音声認識の概要を示す図である。

[図10]図10は、音声認識辞書の設定を示す図である。

[図11]図11は、音声認識辞書の設定を示す他の図である。

[図12]図12は、音声認識辞書設定のタイムチャートである。

[図13]図13は、アイコンの画面表示による報知の様子を示す図である。

[図14]図14は、特定の期間において音声入力を実行する様子を示す図である。

[図15]図15は、特定の期間において音声入力を実行する様子を示す他の図

である。

[図16]図16は、音声認識期間の残り表示の画面表示の例を示す図である。

[図17]図17は、音声認識の候補の画面表示例を示す図である。

[図18]図18は、音声認識結果の画面表示例を示す図である。

[図19]図19は、画像認識の品質に応じた処理の様子を示す図である。

## 発明を実施するための形態

[0031] 本発明に係る内視鏡システム、医療情報処理装置、医療情報処理方法、医療情報処理プログラム、及び記録媒体の実施形態について説明する。説明においては、必要に応じて添付図面が参照される。なお、添付図面において、説明の便宜上一部の構成要素の記載を省略する場合がある。

[0032] [第1の実施形態]

[内視鏡画像診断支援システム]

ここでは、本発明を内視鏡画像診断支援システムに適用した場合を例に説明する。内視鏡画像診断支援システムは、内視鏡検査における病変等の検出及び鑑別をサポートするシステムである。以下においては、下部消化管内視鏡検査（大腸検査）における病変等の検出及び鑑別をサポートする内視鏡画像診断支援システムに適用した場合を例に説明する。

[0033] 図1は、内視鏡画像診断支援システムの概略構成を示すブロック図である。

[0034] 図1に示すように、本実施の形態の内視鏡画像診断支援システム1（内視鏡システム）は、内視鏡システム10（内視鏡システム、医療情報処理装置）、内視鏡情報管理システム100及びユーザ端末200を有する。

[0035] [内視鏡システム]

図2は、内視鏡システム10の概略構成を示すブロック図である。

[0036] 本実施形態の内視鏡システム10は、白色光を用いた観察（白色光観察）の他、特殊光を用いた観察（特殊光観察）が可能なシステムとして構成される。特殊光観察には、狭帯域光観察が含まれる。狭帯域光観察には、B L I 観察（Blue laser imaging観察）、N B I 観察（Narrow band imaging観察；

N B I は登録商標)、L C I 観察 (Linked Color Imaging観察) 等が含まれる。なお、特殊光観察自体は、公知の技術であるので、その詳細についての説明は省略する。

[0037] 図2に示すように、本実施の形態の内視鏡システム10は、内視鏡20、光源装置30、内視鏡画像生成装置40、内視鏡画像処理装置60、表示装置70(出力装置、表示装置)、記録装置75(記録装置)、及び入力装置50等を有する。内視鏡画像生成装置40及び内視鏡画像処理装置60は、医療情報処理装置80(医療情報処理装置)を構成する。

[0038] [内視鏡]

図3は、内視鏡20の概略構成を示す図である。

[0039] 本実施形態の内視鏡20は、下部消化器官用の内視鏡である。図3に示すように、内視鏡20は軟性鏡(電子内視鏡)であり、挿入部21、操作部22及び接続部23を有する。

[0040] 挿入部21は、管腔臓器(本実施の形態では大腸)に挿入される部位である。挿入部21は、先端側から順に先端部21A、湾曲部21B、及び軟性部21Cで構成される。

[0041] 図4は、先端部の端面の構成の一例を示す図である。

[0042] 同図に示すように、先端部21Aの端面には、観察窓21a、照明窓21b、送気送水ノズル21c及び鉗子出口21d等が備えられる。観察窓21aは観察用の窓である。観察窓21aを介して管腔臓器内が撮影される。撮影は、先端部21A(観察窓21aの部分)に内蔵されたレンズ等の光学系及びイメージセンサ(不図示)を介して行われる。イメージセンサには、たとえば、CMOSイメージセンサ(Complementary Metal Oxide Semiconductor image sensor)、CCDイメージセンサ(Charge Coupled Device image sensor)等が使用される。照明窓21bは、照明用の窓である。照明窓21bを介して管腔臓器内に照明光が照射される。送気送水ノズル21cは、洗浄用のノズルである。送気送水ノズル21cから観察窓21aに向けて洗浄用の液体及び乾燥用の気体が噴射される。鉗子出口21d、鉗子等の処置具の

出口である。鉗子出口 2 1 d は、体液等を吸引する吸引口としても機能する。

[0043] 湾曲部 2 1 B は、操作部 2 2 に備えられたアングルノブ 2 2 A の操作に応じて湾曲する部位である。湾曲部 2 1 B は、上下左右の 4 方向に湾曲する。

[0044] 軟性部 2 1 C は、湾曲部 2 1 B と操作部 2 2 との間に備えられる長尺な部位である。軟性部 2 1 C は、可撓性を有する。

[0045] 操作部 2 2 は、術者が把持して各種操作を行う部位である。操作部 2 2 には、各種操作部材が備えられる。一例として、操作部 2 2 には、湾曲部 2 1 B を湾曲操作するためのアングルノブ 2 2 A、送気送水の操作を行うための送気送水ボタン 2 2 B、吸引操作を行うための吸引ボタン 2 2 C が備えられる。この他、操作部 2 2 には、静止画像を撮影するための操作部材（シャッターボタン）、観察モードを切り替えるための操作部材、各種支援機能の ON、OFF を切り替えるための操作部材等が備えられる。また、操作部 2 2 には、鉗子等の処置具を挿入するための鉗子挿入口 2 2 D が備えられる。鉗子挿入口 2 2 D から挿入された処置具は、挿入部 2 1 の先端の鉗子出口 2 1 d（図 4 参照）から繰り出される。一例として、処置具には、生検鉗子、スネア等が含まれる。

[0046] 接続部 2 3 は、内視鏡 2 0 を光源装置 3 0 及び内視鏡画像生成装置 4 0 等に接続するための部位である。接続部 2 3 は、操作部 2 2 から延びるコード 2 3 A と、そのコード 2 3 A の先端に備えられるライトガイドコネクタ 2 3 B 及びビデオコネクタ 2 3 C 等とで構成される。ライトガイドコネクタ 2 3 B は、光源装置 3 0 に接続するためのコネクタである。ビデオコネクタ 2 3 C は、内視鏡画像生成装置 4 0 に接続するためのコネクタである。

[0047] [光源装置]

光源装置 3 0 は、照明光を生成する。上記のように、本実施の形態の内視鏡システム 1 0 は、通常の白色光観察の他に特殊光観察が可能なシステムとして構成される。このため、光源装置 3 0 は、通常の白色光の他、特殊光観察に対応した光（たとえば、狭帯域光）を生成可能に構成される。なお、上

記のように、特殊光観察自体は、公知の技術であるので、その光の生成等についての説明は省略する。

[0048] [医療情報処理装置]

[内視鏡画像生成装置]

内視鏡画像生成装置40（プロセッサ）は、内視鏡画像処理装置60（プロセッサ）と共に、内視鏡システム10全体の動作を統括制御する。内視鏡画像生成装置40は、そのハードウェア構成として、プロセッサ、主記憶部（メモリ）、補助記憶部（メモリ）及び通信部等を備える。すなわち、内視鏡画像生成装置40は、そのハードウェア構成として、いわゆるコンピュータの構成を有する。プロセッサは、例えば、CPU（Central Processing Unit）、GPU（Graphics Processing Unit）、FPGA（Field Programmable Gate Array）、PLD（Programmable Logic Device）等で構成される。主記憶部は、たとえば、RAM（Random Access Memory）等で構成される。補助記憶部は、たとえば、フラッシュメモリ等の非一時的かつ有体の記録媒体で構成され、本発明に係る医療情報処理プログラムまたはその一部のコンピュータ読み取り可能なコード、及びその他のデータを記録することができる。また、補助記憶部は、フラッシュメモリに加えて、またはこれに代えて、各種の光磁気記録装置、半導体メモリ等を含んでいてよい。

[0049] 図5は、内視鏡画像生成装置40の主な機能を示すブロック図である。

[0050] 同図に示すように、内視鏡画像生成装置40は、内視鏡制御部41、光源制御部42、画像生成部43、入力制御部44及び出力制御部45等の機能を有する。プロセッサが実行する各種プログラム（本発明に係る医療情報処理プログラムまたはその一部を含んでいてよい）、及び、制御等に必要な各種データ等が上述した補助記憶部に格納され、内視鏡画像生成装置40の各機能は、プロセッサがそれらのプログラムを実行することにより実現される。内視鏡画像生成装置40のプロセッサは、本発明に係る内視鏡システム、医療情報処理装置におけるプロセッサの一例である。

[0051] 内視鏡制御部41は、内視鏡20を制御する。内視鏡20の制御には、イ

メージセンサの駆動制御、送気送水の制御、吸引の制御等が含まれる。

[0052] 光源制御部42は、光源装置30を制御する。光源装置30の制御には、光源の発光制御等が含まれる。

[0053] 画像生成部43は、内視鏡20のイメージセンサから出力される信号に基づいて撮影画像（内視鏡画像）を生成する。画像生成部43は、撮影画像として静止画像及び／または動画像（イメージセンサ25が被写体を時系列に撮影することで得られた複数の医療画像）を生成することができる。画像生成部43は、生成した画像に各種画像処理を施してもよい。

[0054] 入力制御部44は、入力装置50を介した操作の入力及び各種情報の入力を受け付ける。

[0055] 出力制御部45は、内視鏡画像処理装置60への情報の出力を制御する。内視鏡画像処理装置60に出力する情報には、撮影により得られた内視鏡画像の他、入力装置50から入力された各種操作情報等が含まれる。

[0056] [入力装置]

入力装置50は、表示装置70と共に内視鏡システム10におけるユーザインタフェース（user interface）を構成する。入力装置50には、マイク51（音声入力装置）及びフットスイッチ52（操作デバイス）が含まれる。マイク51は後述する音声認識を行うための入力デバイスである。フットスイッチ52は、術者の足元に置かれて、足で操作される操作デバイスであり、ペダルを踏み込むことで、操作信号（例えば、音声入力トリガを示す信号や、音声認識の候補を選択する信号）が出力される。なお、本態様ではマイク51及びフットスイッチ52は内視鏡画像生成装置40の入力制御部44により制御されるが、本発明ではこのような態様に限らず内視鏡画像処理装置60や表示装置70等を介してマイク51及びフットスイッチ52を制御してもよい。また、内視鏡20の操作部22において、フットスイッチ52と同等の機能を有する操作デバイス（ボタン、スイッチ等）を設けてもよい。

[0057] この他、入力装置50には、操作デバイスとしてキーボード、マウス、タ

タッチパネル、視線入力装置等の公知の入力デバイスを含めることができる。

[0058] [内視鏡画像処理装置]

内視鏡画像処理装置60は、そのハードウェア構成として、プロセッサ、主記憶部、補助記憶部、通信部等を備える。すなわち、内視鏡画像処理装置60は、そのハードウェア構成として、いわゆるコンピュータの構成を有する。プロセッサは、たとえば、CPU、GPU (Graphics Processing Unit)、FPGA (Field Programmable Gate Array)、PLD (Programmable Logic Device) 等で構成される。内視鏡画像処理装置60のプロセッサは、本発明に係る内視鏡システム、医療情報処理装置におけるプロセッサの一例である。内視鏡画像生成装置40のプロセッサと内視鏡画像処理装置60のプロセッサとで、本発明に係る内視鏡システムや医療情報処理装置におけるプロセッサの機能を分担してもよい。例えば、内視鏡画像生成装置40は主として内視鏡画像を生成する「内視鏡プロセッサ」の機能を備え、内視鏡画像処理装置60は主として内視鏡画像に画像処理を施す「CADボックス (CAD: Computer Aided Diagnosis)」としての機能を備える態様を採用することができる。しかしながら、本発明では、このような機能の分担と異なる態様を採用してもよい。

[0059] 主記憶部は、たとえば、RAM等のメモリで構成される。補助記憶部は、たとえば、フラッシュメモリ等の非一時的かつ有体の記録媒体 (メモリ) で構成され、プロセッサが実行する各種プログラム (本発明に係る医療情報処理プログラムまたはその一部を含んでよい) のコンピュータ読み取り可能なコード、及び、制御等に必要な各種データ等が格納される。また、補助記憶部は、フラッシュメモリに加えて、またはこれに代えて、各種の光磁気記録装置、半導体メモリ等を含んでよい。通信部は、たとえば、ネットワークに接続可能な通信インターフェースで構成される。内視鏡画像処理装置60は、通信部を介して内視鏡情報管理システム100と通信可能に接続される。

[0060] 図6は、内視鏡画像処理装置60の主な機能を示すブロック図である。

[0061] 同図に示すように、内視鏡画像処理装置60は、主として、内視鏡画像取得部61、入力情報取得部62、画像認識処理部63、音声入力トリガ受付部64、表示制御部65、及び検査情報出力制御部66等の機能を有する。これらの機能は、プロセッサが補助記憶部等に格納されたプログラム（本発明に係る医療情報処理プログラムまたはその一部を含んでよい）を実行することにより実現される。

[0062] [内視鏡画像取得部]

内視鏡画像取得部61は、内視鏡画像生成装置40から内視鏡画像を取得する。画像の取得は、リアルタイムに行うことができる。すなわち、イメージセンサ25（イメージセンサ）が被写体を時系列に撮影することで得られた複数の医療画像をリアルタイムに順次取得（順次入力）することができる。

[0063] [入力情報取得部]

入力情報取得部62（プロセッサ）は、入力装置50及び内視鏡20を介して入力された情報を取得する。入力情報取得部62は、主として音声情報以外の入力情報を取得する情報取得部62Aと、音声情報を取得すると共にマイク51に入力された音声を認識する音声認識部62Bと、音声認識に用いられる音声認識辞書62Cと、を備える。音声認識辞書62Cは、内容が異なる複数の辞書（例えば、部位情報、所見情報、処置情報、及び止血情報に関する辞書）を含んでいてもよい。

[0064] 入力装置50を介して入力情報取得部62に入力される情報には、マイク51、フットスイッチ52、あるいは図示せぬキーボードやマウス等を介して入力される情報（例えば、音声情報、音声入力トリガ、候補選択操作の情報等）が含まれる。また、内視鏡20を介して入力される情報には、内視鏡画像（動画）の撮影開始指示、静止画像の撮影指示等の情報が含まれる。後述するように、本実施形態において、ユーザはマイク51及び／またはフットスイッチ52を介して音声入力トリガの入力や、音声認識候補の選択操作等を行うことができる。入力情報取得部62は、内視鏡画像生成装置40

を介して、フットスイッチ52の操作情報を取得する。

[0065] [画像認識処理部]

画像認識処理部63（プロセッサ）は、内視鏡画像取得部61で取得される内視鏡画像に対し、画像認識を行う。画像認識処理部63は、リアルタイムに画像認識を行うことができる。

[0066] 図7は、画像認識処理部63の主な機能を示すブロック図である。同図に示すように、画像認識処理部63は、病変部検出部63A、鑑別部63B、特定領域検出部63C、処置具検出部63D、止血具検出部63E、及び計測部63F等の機能を有する。これら各部は、「内視鏡画像に特定の被写体が含まれているか」の判断あるいは判定に用いることができる。「特定の被写体」は、以下に説明するように、画像認識処理部63の各部によって違っていてもよい。

[0067] 病変部検出部63Aは、内視鏡画像からポリープ等の病変部（病変；「特定の被写体」の一例）を検出する。病変部を検出する処理には、病変部であることが確定的な部分を検出する処理の他、病変の可能性のある部分（良性の腫瘍または異形成等；病変候補領域）を検出する処理、病変を処置した後の領域（処置後領域）、及び、直接的または間接的に病変に関連する可能性がある特徴を有する部分（発赤等）を認識する処理等が含まれる。

[0068] 鑑別部63Bは、病変部検出部63Aが「内視鏡画像に病変部（特定の被写体）が含まれている」と判定した場合に、病変部検出部63Aで検出された病変部について鑑別処理を行う。本実施形態において、鑑別部63Bは、病変部検出部63Aで検出されたポリープ等の病変部について、腫瘍性（NEOPLASTIC）もしくは非腫瘍性（HYPERPLASTIC）の鑑別処理を行う。なお、鑑別部63Bは、あらかじめ決められた基準を満たす場合に鑑別結果を出力するように構成することができる。「あらかじめ決められた基準」として、例えば、「鑑別結果の信頼度（内視鏡画像の露出、合焦度合い、ぶれ等の条件に依存する）やその統計値（決められた期間内での最大または最小、平均等）がしきい値以上である場合」を採用することができるが、他の基準を用いて

もよい。

- [0069] 特定領域検出部63Cは、内視鏡画像から管腔臓器内の特定領域（ランドマーク）を検出する処理を行う。たとえば、大腸の回盲部を検出する処理等を行う。大腸は管腔臓器の一例であり、回盲部は特定領域の一例である。特定領域検出部63Cは、例えば、肝湾曲部（右結腸部）、脾湾曲部（左結腸部）、直腸S状部等を検出してもよい。また、特定領域検出部63Cは、複数の特定領域を検出してもよい。
- [0070] 処置具検出部63Dは、内視鏡画像から画像内に現れる処置具を検出し、その種類を判別する処理を行う。処置具検出部63Dは、生検鉗子、スネア等、複数の種類の処置具を検出するように構成することができる。同様に、止血具検出部63Eは、止血クリップ等の止血具を検出し、その種類を判別する処理を行う。処置具検出部63Dと止血具検出部63Eを1つの画像認識器で構成してもよい。
- [0071] 計測部63Fは、病変、病変候補領域、特定領域、処置後領域等の計測（形状、寸法等の測定）を行う。
- [0072] 画像認識処理部63の各部（病変部検出部63A、鑑別部63B、特定領域検出部63C、処置具検出部63D、止血具検出部63E、及び計測部63F等）は、機械学習により構成された画像認識器（学習済みモデル）を用いて構成することができる。具体的には、上述の各部は、ニューラルネットワーク（Neural Network：NN）、畳み込みニューラルネットワーク（Convolutional Neural Network：CNN）、アダブースト（AdaBoost）、ランダムフォレスト（Random Forest）等の機械学習アルゴリズムを用いて学習した画像認識器（学習済みモデル）で構成することができる。また、鑑別部63Bについて上述したように、これらの各部は、必要に応じてネットワークの層構成を設定すること等により、最終的な出力（鑑別結果や処置具の種類等）の信頼度を合わせて出力することができる。また、上述の各部は、内視鏡画像の全フレームについて画像認識を行ってもよいし、一部のフレームについて間欠的に画像認識を行ってもよい。

[0073] 後述するように、これらの各部から内視鏡画像の認識結果が出力されることや、あらかじめ決められた基準（信頼度のしきい値等）を満たす認識結果が出力されることを音声入力トリガとしてもよいし、それらの出力がされる期間を、音声認識を実行する期間としてもよい。

[0074] また、画像認識処理部 6 3 を構成する各部の一部または全部を画像認識器（学習済みモデル）で構成する代わりに、内視鏡画像から特徴量を算出し、算出した特徴量を用いて検出等を行う構成とすることもできる。

[0075] [音声入力トリガ受付部]

音声入力トリガ受付部 6 4（プロセッサ）は、内視鏡画像の撮影中（入力中）に音声入力トリガの入力を受け付け、入力された音声入力トリガに応じた音声認識辞書 6 2 C を設定する。本実施形態における音声入力トリガは、例えば内視鏡画像に特定の被写体が含まれていることを示す判定結果（検出結果）であり、この場合、判定結果として病変部検出部 6 3 A の出力を用いることができる。また、音声入力トリガの他の例は特定の被写体に対する鑑別結果の出力であり、この場合、鑑別結果として鑑別部 6 3 B の出力を用いることができる。音声入力トリガのさらに他の例として、複数の医療画像の撮影開始指示、マイク 5 1（音声入力装置）に対するウェイクワードの入力、フットスイッチ 5 2 の操作、内視鏡システムに接続された他の操作デバイス（例えば、大腸内視鏡形状測定装置等）に対する操作等を用いることもできる。これら音声入力トリガに応じた音声認識辞書の設定及び音声認識については、詳細を後述する。

[0076] [表示制御部]

表示制御部 6 5（プロセッサ）は、表示装置 7 0 の表示を制御する。以下、表示制御部 6 5 が行う主な表示制御について説明する。

[0077] 表示制御部 6 5 は、検査中（撮影中）、内視鏡 2 0 で撮影された画像（内視鏡画像）を表示装置 7 0 にリアルタイムに表示させる。図 8 は、検査中の画面表示の一例を示す図である。同図に示すように、画面 7 0 A 内に設定された主表示領域 A 1 に内視鏡画像 1（ライブビュー）が表示される。画面 7

0 Aには、更に副表示領域 A 2 が設定され、検査に関する各種情報が表示される。図 8 に示す例では、患者に関する情報 1 p、及び、検査中に撮影された内視鏡画像の静止画像 1 s を副表示領域 A 2 に表示した場合の例を示している。静止画像 1 s は、たとえば、画面 7 0 A の上から下に向かって撮影された順に表示される。

[0078] また、表示制御部 6 5 は、音声認識の状態を示すアイコン 3 0 0、撮影中の部位を示すアイコン 3 2 0、撮影対象の部位（上行結腸、横行結腸、下行結腸等）及び音声認識の結果をリアルタイムに（時間遅れなしに）文字表示する表示領域 3 4 0 を画面 7 0 A に表示させることができる。表示制御部 6 5 は、内視鏡画像からの画像認識、ユーザによる操作デバイスを介した入力、内視鏡システム 1 0 に接続された外部装置（例えば、内視鏡挿入形状観測装置）等により部位の情報を取得することができる。

[0079] また、後述するように、表示制御部 6 5 は、音声認識の結果を表示装置 7 0（出力装置、表示装置）に表示（出力）させることができる。

[0080] [検査情報出力制御部]

検査情報出力制御部 6 6 は、検査情報を記録装置 7 5 及び／または内視鏡情報管理システム 1 0 0 に出力する。検査情報は、例えば検査中に撮影された内視鏡画像、特定の被写体についての判定の結果、音声認識の結果、検査中に入力された部位の情報、検査中に入力された処置名の情報、検査中に検出された処置具の情報を含む。検査情報は、たとえば、病変ないし検体採取ごとに出力される。この際、各情報が、互いに関連付けられて出力される。たとえば、病変部等を撮影した内視鏡画像に対し、選択中の部位の情報が関連付けられて出力される。また、処置が行われた場合には、選択された処置名の情報及び検出された処置具の情報が、内視鏡画像及び部位の情報に関連付けられて出力される。また、病変部等とは別に撮影された内視鏡画像については、適時、記録装置 7 5 及び／または内視鏡情報管理システム 1 0 0 に出力される。内視鏡画像は、撮影日時情報が付加されて出力される。

[0081] [記録装置]

記録装置 75（記録装置）は、各種の光磁気記録装置や半導体メモリ、及びその制御装置を備え、内視鏡画像（動画像、静止画像）、画像認識の結果、音声認識の結果、検査情報、レポート作成支援情報等を記録することができる。これらの情報は、内視鏡画像生成装置 40 や内視鏡画像処理装置 60 の副記憶部、あるいは内視鏡情報管理システム 100 が備える記録装置に記録してもよい。

[0082] [内視鏡システムにおける音声認識]

上述した構成の内視鏡システム 10 における音声認識について、以下に説明する。

[0083] [音声認識の概要]

図 9 は、音声認識の概要を示す図である。同図に示すように、医療情報処理装置 80（プロセッサ）は、内視鏡画像の撮影中（順次入力中）に音声入力トリガの入力を受け付け、音声入力トリガが入力された場合に、音声入力トリガに応じた音声認識辞書を設定し、音声認識辞書の設定がされた以降にマイク 51（音声入力装置）に入力された音声を、設定された音声認識辞書を用いて音声認識する。上述のように、医療情報処理装置 80 は、病変部検出部 63A による検出結果の出力、鑑別部 63B の鑑別結果の出力、複数の医療画像の撮影開始指示、検出モードから鑑別モードへの切替操作、マイク 51（音声入力装置）に対するウェイクワードの入力、フットスイッチ 52 の操作、内視鏡システムに接続された操作デバイスに対する操作の入力等を「音声入力トリガが入力された」と判断して音声認識を行う。

[0084] なお、音声認識の開始は音声認識辞書の設定に対し遅延していてもよいが、音声認識辞書が設定されたら即座に音声認識を開始する（遅延時間ゼロ）ことが好ましい。

[0085] [音声認識辞書の設定]

図 10 は、音声認識辞書の設定を示す図である。同図の（a）部分～（e）部分において、矢印の左側は音声入力トリガを示し、矢印の右側は音声入力トリガに応じて設定される音声認識辞書及び登録ワードの例を示す。図 1

Oの各部に示すように、音声入力トリガが入力された場合、音声認識部62Bは、音声入力トリガに応じた音声認識辞書62Cを設定する。例えば、鑑別部63Bが鑑別結果を出力した場合、音声認識部62Bは、音声認識辞書として「所見セットA」を設定する。

[0086] 図11は、音声認識辞書の設定を示す他の図である。同図の(a)部分及び(b)部分に示すように、音声認識部62Bは、フットスイッチ52(操作デバイス)への操作を音声入力トリガとして受け付けた場合は「全辞書セット」を設定し、マイク51(音声入力装置)へのウェイクワードの入力を音声入力トリガとして受け付けた場合は、ウェイクワードの内容に応じた音声認識辞書を設定する。なお「ウェイクワード(wake word)」あるいは「ウェイクアップワード(wakeup word)」とは、例えば「音声認識部62Bに音声認識辞書の設定及び音声認識を開始させるための、あらかじめ決められた語句」と規定することができる。

[0087] 上述のウェイクワード(ウェイクアップワード)は、2種類に分けることができる。「レポート入力に関するウェイクワード」と「撮影モード制御に関するウェイクワード」である。「レポート入力に関するウェイクワード」は例えば「所見入力」、「処置入力」であり、このようなウェイクワード認識後に、「所見」、「処置」用の音声認識辞書が設定され、辞書のワードが認識された場合に、音声認識の結果が出力される。音声認識の結果は画像と紐付けたり、レポートで使用したりすることができる。画像との紐付け、レポートでの使用は音声認識の結果の「出力」の一態様であり、表示装置70、記録装置75、医療情報処理装置80の記憶部、あるいは内視鏡情報管理システム100等の記録装置は「出力装置」の一態様である。

[0088] もう一方の「撮影モード制御に関するウェイクワード」は例えば「撮影設定」、「設定」であり、このようなウェイクワードの認識後に、音声で光源のON/OFFあるいは切替え(例えば「ホワイト」、「LCI」、「BLI」等の単語を音声認識することによる)を行ったり、内視鏡A1(人工知能を用いた認識器)による病変検出のON/OFF(例えば「検出オン」、

「検出オフ」等の単語を音声認識することによる)に用いる辞書を設定したりすることができる。なお、「出力」や「出力装置」に関しては「レポート入力に関するウェイクワード」について上述したのと同様である。

[0089] [音声認識辞書設定のタイムチャート]

図12は、音声認識辞書設定のタイムチャートである。なお、図12では具体的に音声入力される語句及びその認識結果は記載していない。図12の(a)部分は、音声入力トリガの種類を示す。同部分に示す例では、音声入力トリガは内視鏡画像の画像認識の結果の出力、マイク51に対するウェイクワードの入力、フットスイッチ52(操作デバイス)の操作による信号、内視鏡画像の撮影開始指示である。また、図12の(b)部分は、音声入力トリガに応じて設定される音声認識辞書を示す。音声認識部62Bは、検査の流れ(撮影開始、病変または病変候補の発見、所見入力、処置具挿入及び処置、止血)に従って、異なる音声認識辞書を設定する。

[0090] 内視鏡システム10では、画像認識処理部63の各部により、判定(認識)の対象となる複数の種類の「特定の被写体」(具体的には、上述した病変、処置具、止血具等)にそれぞれ対応する画像認識(全体として、複数の画像認識)を行うことができ、音声認識部62Bは、これら各部による画像認識のいずれかにより「内視鏡画像に含まれている」と判定された「特定の被写体」の種類に対応する音声認識辞書を設定することができる。

[0091] また、内視鏡システム10では、内視鏡画像に複数の「特定の被写体」が含まれているかをこれら各部により判定し、音声認識部62Bが、複数の「特定の被写体」のうち、「内視鏡画像に含まれている」と判定された特定の被写体に対応する音声認識辞書を設定することもできる。内視鏡画像に複数の「特定の被写体」が含まれているケースとしては、例えば複数の病変部が含まれている場合や、複数の処置具が含まれている場合、複数の止血具が含まれている場合が考えられる。

[0092] なお、上記各部による複数の画像認識のうち一部の画像認識について、「特定の被写体」の種類に応じた音声認識辞書を設定してもよい。

[0093] [音声認識]

音声認識部 6 2 B は、音声認識辞書の設定がされた以降にマイク 5 1（音声入力装置）に入力された音声を、設定された音声認識辞書を用いて音声認識する（図 1 2 では図示を省略する）。表示制御部 6 5 は、音声認識の結果を表示装置 7 0 に表示させることが好ましい。

[0094] 本実施形態では、音声認識部 6 2 B は、部位情報、所見情報、処置情報、及び止血情報について音声認識を行うことができる。なお、病変等が複数存在する場合は、一連の処理（撮影開始～止血のサイクルにおける音声入力トリガの受付、音声認識辞書の設定、及び音声認識）を、病変等ごとに繰り返すことができる。

[0095] [音声認識及び結果表示する単語]

内視鏡システム 1 0 では、音声認識部 6 2 B 及び表示制御部 6 5（プロセッサ）は、音声認識において、設定された音声認識辞書に登録されている登録単語のみを認識し、登録単語についての音声認識の結果を表示装置 7 0（出力装置、表示装置）に表示（出力）させることができる（adaptiveな音声認識）。このような態様によれば、設定された音声認識辞書に登録されている登録単語のみを音声認識するので、認識精度を高めることができる。なお、このようなadaptiveな音声認識では、ウェイクワードを認識しないように音声認識辞書の登録単語を設定してもよいし、ウェイクワードも含めて登録単語を設定してもよい。

[0096] また、内視鏡システム 1 0 では、音声認識部 6 2 B 及び表示制御部 6 5（プロセッサ）は、音声認識において、設定された音声認識辞書に登録されている登録単語及び特定の単語を認識し、認識した単語のうち登録単語についての音声認識の結果を表示装置 7 0（表示装置、出力装置）に表示（出力）させることもできる（adaptiveでない音声認識）。なお、「特定の単語」の例としては音声入力装置に対するウェイクワードを挙げることができるが、「特定の単語」はこれに限定されるものではない。

[0097] なお、内視鏡システム 1 0 において、上述の態様（adaptiveな音声認識、a

daptiveでない音声認識)のいずれにより音声認識及び結果表示を行うかは、入力装置50や操作部22等を介したユーザの指示入力に基づいて設定することができる。

[0098] [アイコンの切替表示によるユーザへの報知]

なお、内視鏡システム10において、表示制御部65(プロセッサ)は、音声認識辞書の設定(設定された事実、及びいずれの辞書が設定されたか)や音声認識が可能である旨を、ユーザに報知することが好ましい。表示制御部65は、図13に示すように、画面表示されるアイコンを切り替えることにより報知を行うことができる。図13に示す例では、表示制御部65が、画像認識処理部63の各部のうち動作している(あるいは認識結果を画面表示させている)画像認識器を示すアイコンを画面70A等に画面表示させ、その画像認識器が特定の被写体(音声入力トリガ)を認識してする音声認識期間になると表示をマイク状のアイコンに切り替えて、ユーザに報知する(図8, 16~18を参照)。

[0099] 具体的には、図13の(a)部分及び(b)部分は処置具検出部63Dが動作している状態であるが、認識対象である特定の被写体が異なる(鉗子、スネア)であるため、表示制御部65は異なるアイコン360, 362を表示させ、実際に鉗子あるいはスネアが認識されるとマイク状のアイコン300に切り替えて、音声認識が可能になった旨をユーザに報知する。同様に、図13の(c)部分、(d)部分に示す状態はそれぞれ止血具検出部63E、鑑別部63Bが動作している状態であり、表示制御部65はアイコン364、アイコン366を表示させているが、止血具、病変が認識されると、マイク状のアイコン300に切り替えて、音声認識が可能になった旨をユーザに報知する。

[0100] このような報知により、ユーザは、特定の画像認識器が動作していることや音声認識が可能な期間であることを容易に把握することができる。なお、表示制御部65は、画像認識処理部63の各部の動作状況だけでなく、マイク51及び/またはフットスイッチ52の動作状況や入力状況に応じてアイ

コンを表示及び切り替えしてもよい。

[0101] [特定の期間における音声認識の実行]

音声認識部62B（プロセッサ）は、設定された音声認識辞書による音声認識を、設定がされた以降の特定の期間（あらかじめ決められた条件を満たす期間）において実行することができる。「あらかじめ決められた条件」は画像認識器から認識結果が出力されることでもよいし、出力の内容に対する条件でもよいし、音声認識の実行時間そのもの（3秒、5秒等）を規定してもよい。実行時間を規定する場合、辞書設定からの経過時間や、音声入力可能である旨をユーザに報知してからの経過時間を規定する事ができる。

[0102] 図14は、特定の期間において音声認識を実行する様子を示す図である。

図14の（a）部分に示す例では、音声認識部62Bは、鑑別モードである期間（鑑別部63Bが動作している期間；時刻t1～時刻t2）のみ音声認識を行う。また、図14の（b）部分に示す例では、鑑別部63Bが鑑別結果（鑑別判定結果）を出力する期間（時刻t2～時刻t3）のみ、音声認識を行う。上述のように、鑑別部63Bは、鑑別結果の信頼度やその統計値がしきい値以上である場合等に出力を行うように構成することができる。また、図14の（c）部分に示す例では、音声認識部62Bは、処置具検出部63Dが処置具を検出している期間（時刻t1～時刻t2）及び止血具検出部63Eが止血具を検出している期間（時刻t3～時刻t4）のみ、音声認識を行う。なお、図14及び以下の図15において、音声入力トリガの受付及び音声認識辞書の設定は図示を省略している。

[0103] このように、特定の期間において音声認識を実行することで、不要な認識や誤認識のおそれを低減し、検査を円滑に行うことができる。

[0104] なお、音声認識部62Bは音声認識の期間を画像認識器ごとに設定してもよいし、音声入力トリガの種類に応じて設定してもよい。また、音声認識部62Bは、「あらかじめ決められた条件」や「音声認識の実行時間」を、入力装置50や操作部22等を介したユーザの指示入力に基づいて設定してもよい。

## [0105] [手動操作後の音声認識]

図15は、特定の期間において音声認識を実行する様子を示す他の図である。図15の(a)部分は、手動操作後の一定時間(同部分では時刻t1～t2、及び時刻t3～t4)に音声認識辞書の設定及び音声認識を実行する例を示す。音声認識部62Bは、入力装置50や操作部22等に対するユーザの操作を「手動操作」として音声認識を行うことができる。具体的には、「手動操作」は上述した各種の操作デバイスに対する操作や、マイク51を介したウェイクワードの入力、フットスイッチ52の操作であってよく、内視鏡画像(動画像、静止画像)の撮影指示、検出モード(病変部検出部63Aが結果を出力する状態)から鑑別モード(鑑別部63Bが結果を出力する状態)への切替操作、内視鏡システム10に接続された操作デバイスに対する操作であってもよい。

[0106] また、図15の(b)部分は、画像認識に基づく音声認識の期間と、上述した「手動操作後の一定時間」とがオーバーラップする場合の処理の例を示す。具体的には、音声認識部62Bは、時刻t1～時刻t3において、鑑別部63Bからの鑑別結果出力に応じた音声認識よりも手動操作に伴う音声認識を優先し、手動操作に基づく音声認識辞書を設定して音声認識を行う。

[0107] このように手動操作に基づく音声認識を優先する場合、画像認識に基づく音声認識の期間が手動操作に伴う音声認識の期間と連続していてもよい。例えば、図15の(b)部分に示す例では、音声認識部62Bは、手動操作による音声認識期間(時刻t1～時刻t2)に続く時刻t3～時刻t4においては、鑑別部63Bの鑑別結果に基づく音声認識辞書を設定し、音声認識を行う。一方、時刻t4～時刻t5においては、手動操作による音声認識期間が終了しているので、音声認識部62Bは音声認識辞書を設定せず、音声認識を行わない。同様に、音声認識部62Bは、時刻t5～時刻t6においては手動操作に基づく音声認識辞書を設定して音声認識を行い、この音声認識期間が終了した時刻t6以降は、音声認識を行わない。

[0108] [残り時間の画面表示]

音声認識部 6 2 B 及び表示制御部 6 5 は、音声認識期間の残り時間を表示装置 7 0 に画面表示してもよい。即ち、音声認識部 6 2 B 及び表示制御部 6 5 は、音声認識辞書が設定された以降の決められた期間において、音声認識を行ってもよい。図 1 6 は、残り時間の画面表示の例を示す図である。図 1 6 の (a) 部分は画面 7 0 A における表示の例であり、残時間メーター 3 5 0 が表示されている。また、同図の (b) 部分は残時間メーター 3 5 0 の拡大図である。残時間メーター 3 5 0 において、斜線で示す領域 3 5 2 が時間の経過につれて伸び、無地で示す領域 3 5 4 が時間の経過につれて縮んでいく。また、これら領域の周辺では黒地領域 3 5 6 A と白地領域 3 5 6 B とから構成される枠 3 5 6 が回転し、ユーザの注意を喚起する。音声認識部 6 2 B 及び表示制御部 6 5 は、音声が入力されていることを検出した場合に枠 3 5 6 を回転表示させてもよい。

[0109] なお、音声認識部 6 2 B 及び表示制御部 6 5 は、音声認識を行う期間として、音声入力トリガや音声認識辞書によって異なる期間を設定してもよい。また、入力装置 5 0 を介したユーザの操作に応じて期間を設定してもよい。

[0110] なお、音声認識部 6 2 B 及び表示制御部 6 5 は、残り時間を数字や音声で出力してもよい。なお、マイク状のアイコン 3 0 0 (図 8, 1 6 ~ 1 8 参照) の画面表示が消えたら残り時間ゼロである。

[0111] [音声認識候補／選択結果の表示]

音声認識部 6 2 B 及び表示制御部 6 5 は、音声認識の候補を画面表示してその候補をユーザに選択させてもよい。また、音声認識結果を表示装置 7 0 に画面表示させてもよい。図 1 7 は音声認識の候補及び音声認識結果の画面表示例を示す図である (同図では、注目領域 R O I 及びフレーム F も表示されている)。図 1 7 は鑑別部 6 3 B が鑑別結果を出力し、鑑別結果の出力に対応する音声認識辞書「所見セット A」 (図 1 0 参照) の内容が画面 7 0 A の領域 3 7 0 に表示された状態を示す。音声認識部 6 2 B は、マイク 5 1 やフットスイッチ 5 2、その他の操作デバイスを介したユーザの選択操作に応じて、変換 (単語の選択) を確定することができる。なお、音声認識部 6 2

B及び表示制御部65は、音声入力トリガの入力、あるいは音声認識辞書の設定を候補表示のトリガとすることができる。

[0112] 図18は音声認識結果の画面表示例を示す図である。表示制御部65は、図18に示すように、ユーザが選択した単語（同図の例では“JNET TYPE 2A”）を画面表示（領域372）することができる。

[0113] [音声認識結果の表示のバリエーション]

本発明において、音声認識結果の表示態様は図18等に例示する態様に限定されるものではない。上述した態様の他に、音声認識部62B及び表示制御部65は、音声認識の結果を表示領域340（図8参照）等にリアルタイムに文字表示させ、その後確定した結果を図18のように領域372に表示してもよい。また、音声認識部62B及び表示制御部65は、選択された、あるいは確定した音声認識の結果を、動画像（例えば、図8、18に示す内視鏡画像1）の表示領域に重畳表示してもよい（図18に示す例では、注目領域ROIやフレームFの近傍に“JNET TYPE 2A”と表示することができる）。

[0114] 音声認識部62B及び表示制御部65は、音声認識結果あるいは認識した被写体の種類等に応じて、音声認識の選択結果や確定結果の表示位置を設定してもよい。音声認識部62B及び表示制御部65は、例えば、「所見」についての音声認識結果を動画像の注目領域（例えば、図18の注目領域ROI）の近傍に重畳表示し、「処置」や「止血」についての音声認識結果を動画像の表示領域の外（例えば、アイコン300やアイコン320、あるいは残時間メーター350の近傍）に表示することができる。

[0115] [画像認識の品質に応じた音声認識辞書の切替]

上述した音声認識において、音声認識部62Bは、図19（画像認識の品質に応じた処理の様子を示す図）を参照して以下に説明するように、画像認識処理部63（図7参照）が実行する画像認識の品質に応じて音声認識辞書62Cを切り替えてもよい。

[0116] 内視鏡画像に病変候補（特定の被写体）が含まれている場合、鑑別部63

Bが鑑別結果を出力する期間が音声認識期間となる（図14の（a）部分と同様）。このような状況で、図19の（a）部分に示すように、時刻t1～時刻t2（検出モード；病変部検出部63Aが結果を出力する）において観察品質（内視鏡画像の画質）が不良であるものとする。観察品質が不良である原因としては、例えば、露出や合焦状態が不適切である、残渣で視野が遮られている、等が考えられる。

[0117] この場合、音声認識部62Bは、図19の（b）部分に示すように、通常ならば（画質が良好ならば）音声認識を行わない時刻t1～時刻t2で音声認識を行い、画質改善操作のコマンドを受け付ける。音声認識部62Bは、例えば「ガス注入、照明オン、センサ感度『高』」等の単語を登録した「画質改善セット」を音声認識辞書62Cとして設定して、音声認識を行うことができる。

[0118] 時刻t3～時刻t4（鑑別モード：鑑別部63Bが結果を出力）では、音声認識部62Bは、通常通り音声認識辞書「所見セット」により音声認識を行う。

[0119] また、時刻t4～時刻t9では検出モードなので、音声認識部62Bは通常ならば音声認識を行わず、時刻t5～時刻t8では、処置具が検出されているので音声認識辞書62Cとして「処置セット」を設定して音声認識を行う。しかしながら、時刻t6～時刻t7で観察品質が不良であるものとする。音声認識部62Bはこの期間（時刻t6～時刻t7）においても、時刻t1～時刻t2と同様に画質改善操作のコマンドを受け付けることができる。

[0120] このように、内視鏡システム10では、観察品質に応じて臨機応変に音声認識辞書を設定し、適切な音声認識を行うことができる。

[0121] [レポート作成支援情報の記録]

音声認識が行われたら、検査情報出力制御部66（プロセッサ）は、内視鏡画像（時系列の医療画像）と音声認識の結果とを関連付けて、記録装置75、医療情報処理装置80の記憶部、内視鏡情報管理システム100等の記録装置に記録させることができる。検査情報出力制御部66は、特定の被写

体が写った内視鏡画像と画像認識による判定の結果（その画像に特定の被写体が写っていること）とを関連付けて記録させてもよい。検査情報出力制御部66は、操作デバイスに対するユーザの操作に応じて記録を行ってもよいし、ユーザの操作によらずに自動的に記録を行ってもよい。内視鏡システム10では、このような記録により、ユーザが検査レポートを作成するのを支援することができる。

[0122] [その他]

上述の実施形態では、本発明を下部消化管用の内視鏡システムに適用した場合について説明したが、本発明は上部消化管内視鏡にも適用することができる。

[0123] 以上で本発明の実施形態について説明してきたが、本発明は上述した態様に限定されず、本発明の精神を逸脱しない範囲で種々の変形が可能である。

## 符号の説明

- [0124] 1 内視鏡画像診断支援システム
- 10 内視鏡システム
  - 20 内視鏡
  - 21 挿入部
    - 21A 先端部
    - 21B 湾曲部
    - 21C 軟性部
      - 21a 観察窓
      - 21b 照明窓
      - 21c 送気送水ノズル
      - 21d 鉗子出口
    - 22 操作部
      - 22A アンクルノブ
      - 22B 送気送水ボタン
      - 22C 吸引ボタン

- 2 2 D 鉗子挿入口
- 2 3 接続部
- 2 3 A コード
- 2 3 B ライトガイドコネクタ
- 2 3 C ビデオコネクタ
- 3 0 光源装置
- 4 0 内視鏡画像生成装置
- 4 1 内視鏡制御部
- 4 2 光源制御部
- 4 3 画像生成部
- 4 4 入力制御部
- 4 5 出力制御部
- 5 0 入力装置
- 5 1 マイク
- 5 2 フットスイッチ
- 6 0 内視鏡画像処理装置
- 6 1 内視鏡画像取得部
- 6 2 入力情報取得部
- 6 2 A 情報取得部
- 6 2 B 音声認識部
- 6 2 C 音声認識辞書
- 6 3 画像認識処理部
- 6 3 A 病変部検出部
- 6 3 B 鑑別部
- 6 3 C 特定領域検出部
- 6 3 D 処置具検出部
- 6 3 E 止血具検出部
- 6 3 F 計測部

- 6 4 音声入力トリガ受付部
- 6 5 表示制御部
- 6 6 検査情報出力制御部
- 7 0 表示装置
- 7 0 A 画面
- 7 5 記録装置
- 8 0 医療情報処理装置
- 1 0 0 内視鏡情報管理システム
- 2 0 0 ユーザ端末
- 3 0 0 アイコン
- 3 2 0 アイコン
- 3 4 0 表示領域
- 3 5 0 残時間メーター
- 3 5 2 領域
- 3 5 4 領域
- 3 5 6 枠
- 3 5 6 A 黒地領域
- 3 5 6 B 白地領域
- 3 6 0 アイコン
- 3 6 2 アイコン
- 3 6 4 アイコン
- 3 6 6 アイコン
- 3 7 0 領域
- 3 7 2 領域
- A 1 主表示領域
- A 2 副表示領域
- F フレーム
- I 内視鏡画像

l p 情報

l s 静止画像

R O I 注目領域

## 請求の範囲

- [請求項1] 音声入力装置と、  
被写体を撮影するイメージセンサと、  
プロセッサと、  
を備える内視鏡システムであって、  
前記プロセッサは、  
前記イメージセンサが前記被写体を時系列に撮影することで得られた複数の医療画像を取得し、  
前記複数の医療画像の撮影中に音声入力トリガの入力を受け付け、  
前記音声入力トリガが入力された場合に、前記音声入力トリガに応じた音声認識辞書を設定し、  
前記設定がされた以降に前記音声入力装置に入力された音声を、前記設定された音声認識辞書を用いて音声認識する、  
内視鏡システム。
- [請求項2] 前記プロセッサは、前記音声認識において、前記設定された前記音声認識辞書に登録されている登録単語のみを認識し、前記登録単語についての前記音声認識の結果を出力装置に出力させる請求項1に記載の内視鏡システム。
- [請求項3] 前記プロセッサは、前記音声認識において、前記設定された前記音声認識辞書に登録されている登録単語及び特定の単語を認識し、認識した単語のうち前記登録単語についての前記音声認識の結果を出力装置に出力させる請求項1に記載の内視鏡システム。
- [請求項4] 前記プロセッサは、  
前記複数の医療画像に特定の被写体が含まれているかを画像認識により判定し、  
前記特定の被写体が含まれていることを示す判定結果を前記音声入力トリガとして受け付ける請求項1から3のいずれか1項に記載の内視鏡システム。

- [請求項5] 前記プロセッサは、  
前記複数の医療画像に特定の被写体が含まれているかを画像認識により判定し、  
前記特定の被写体が含まれていると判定した場合に、前記特定の被写体を鑑別し、  
前記特定の被写体に対する鑑別結果の出力を前記音声入力トリガとして受け付ける請求項1から4のいずれか1項に記載の内視鏡システム。
- [請求項6] 前記プロセッサは、  
前記複数の医療画像に複数の種類の前記特定の被写体が含まれているかを、前記複数の種類の特定の被写体にそれぞれ対応する複数の画像認識により判定し、  
前記複数の種類の前記特定の被写体のうち、前記複数の画像認識のいずれかにより前記複数の医療画像に含まれていると判定された特定の被写体の種類に対応する前記音声認識辞書を設定する請求項4または5に記載の内視鏡システム。
- [請求項7] 前記プロセッサは、  
前記複数の医療画像に複数の特定の被写体が含まれているかを画像認識により判定し、  
前記複数の特定の被写体のうち、前記複数の医療画像に含まれていると判定された前記特定の被写体に対応する前記音声認識辞書を設定する請求項6に記載の内視鏡システム。
- [請求項8] 前記プロセッサは、機械学習により構成された画像認識器を用いて前記画像認識を行う請求項4から7のいずれか1項に記載の内視鏡システム。
- [請求項9] 前記プロセッサは、前記複数の医療画像のうち前記特定の被写体が写っていると判断された医療画像と、前記特定の被写体の前記画像認識による判定の結果と、前記音声認識の結果と、を関連付けて記録装

置に記録させる請求項4から8のいずれか1項に記載の内視鏡システム。

[請求項10] 前記プロセッサは、病変、病変候補領域、ランドマーク、処置後領域、処置具、または止血具のうち少なくとも1つを前記特定の被写体と判断する請求項4から9のいずれか1項に記載の内視鏡システム。

[請求項11] 前記プロセッサは、前記設定された前記音声認識辞書による前記音声認識を、前記設定がされた以降のあらかじめ決められた条件を満たす期間において実行する請求項4から10のいずれか1項に記載の内視鏡システム。

[請求項12] 前記プロセッサは、前記画像認識を行った画像認識器ごとに前記期間を設定する請求項11に記載の内視鏡システム。

[請求項13] 前記プロセッサは、前記音声入力トリガの種類に応じて前記期間を設定する請求項11または12に記載の内視鏡システム。

[請求項14] 前記プロセッサは、前記期間の残り時間を表示装置に画面表示させる請求項11から13のいずれか1項に記載の内視鏡システム。

[請求項15] 前記プロセッサは、部位情報、所見情報、処置情報、及び止血情報について前記音声認識を行う請求項1から14のいずれか1項に記載の内視鏡システム。

[請求項16] 前記プロセッサは、前記複数の医療画像の撮影開始指示、前記複数の医療画像に対する画像認識の結果の出力、鑑別モードへの切替操作、前記内視鏡システムに接続された操作デバイスに対する操作、前記音声入力装置に対するウェイクワードの入力のいずれかがなされた場合に前記音声入力トリガが入力されたと判断する請求項1から15のいずれか1項に記載の内視鏡システム。

[請求項17] 前記プロセッサは、前記音声認識の結果を表示装置に表示させる請求項1から16のいずれか1項に記載の内視鏡システム。

[請求項18] プロセッサを備える医療情報処理装置であって、  
前記プロセッサは、

イメージセンサが被写体を時系列に撮影することで得られた複数の医療画像を取得し、

前記複数の医療画像の入力中に音声入力トリガの入力を受け付け、  
前記音声入力トリガが入力された場合に、前記音声入力トリガに応じた音声認識辞書を設定し、

前記設定がされた以降に音声入力装置に入力された音声を、前記設定された音声認識辞書を用いて音声認識する、  
医療情報処理装置。

[請求項19]

音声入力装置と、被写体を撮影するイメージセンサと、プロセッサと、を備える内視鏡システムにより実行される医療情報処理方法であって、

前記プロセッサが、

前記イメージセンサが前記被写体を時系列に撮影することで得られた複数の医療画像を取得し、

前記複数の医療画像の撮影中に音声入力トリガの入力を受け付け、  
前記音声入力トリガが入力された場合に、前記音声入力トリガに応じた音声認識辞書を設定し、

前記設定がされた以降に前記音声入力装置に入力された音声を、前記設定された音声認識辞書を用いて音声認識する、  
医療情報処理方法。

[請求項20]

音声入力装置と、被写体を撮影するイメージセンサと、プロセッサと、を備える内視鏡システムに医療情報処理方法を実行させる医療情報処理プログラムであって、

前記医療情報処理方法において、前記プロセッサは、

前記イメージセンサが前記被写体を時系列に撮影することで得られた複数の医療画像を取得し、

前記複数の医療画像の撮影中に音声入力トリガの入力を受け付け、

前記音声入力トリガが入力された場合に、前記音声入力トリガに応

じた音声認識辞書を設定し、

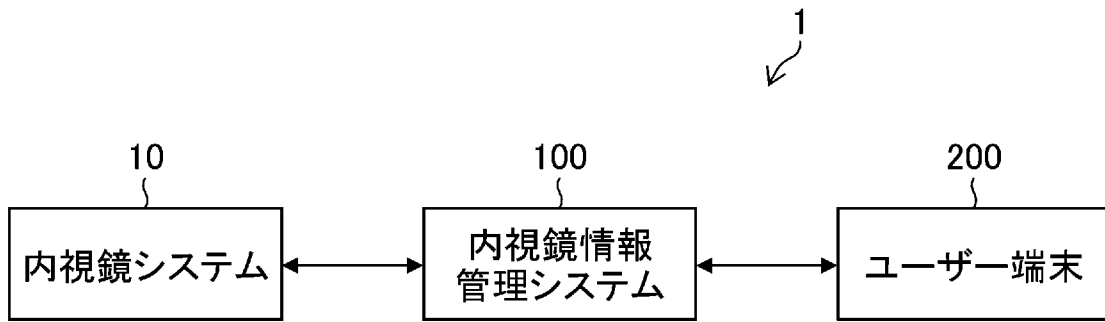
前記設定がされた以降に前記音声入力装置に入力された音声を、前記設定された音声認識辞書を用いて音声認識する、

医療情報処理プログラム。

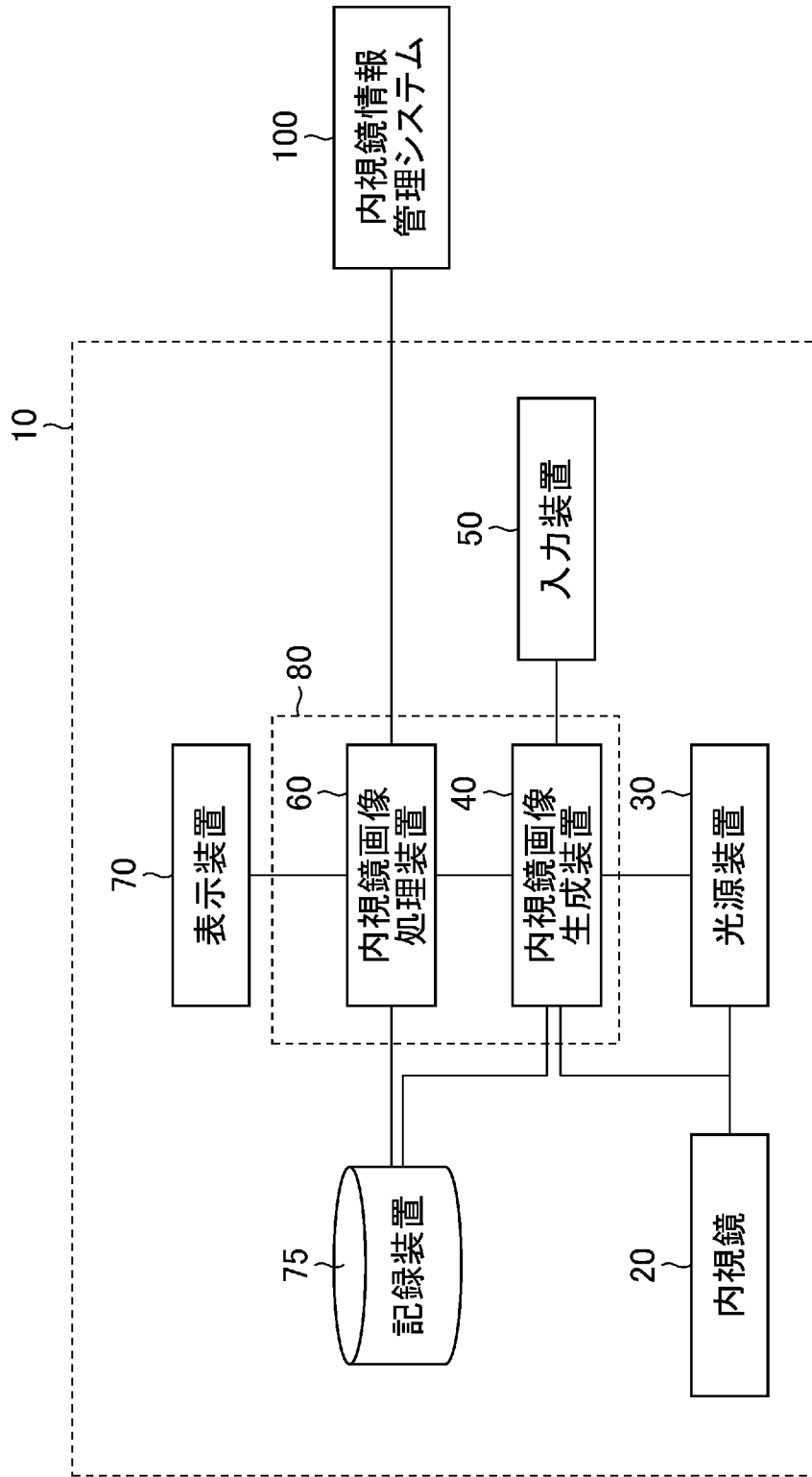
[請求項21]

非一時的かつ有体の記録媒体であって、請求項20に記載の医療情報処理プログラムのコンピュータ読み取り可能なコードが記録された記録媒体。

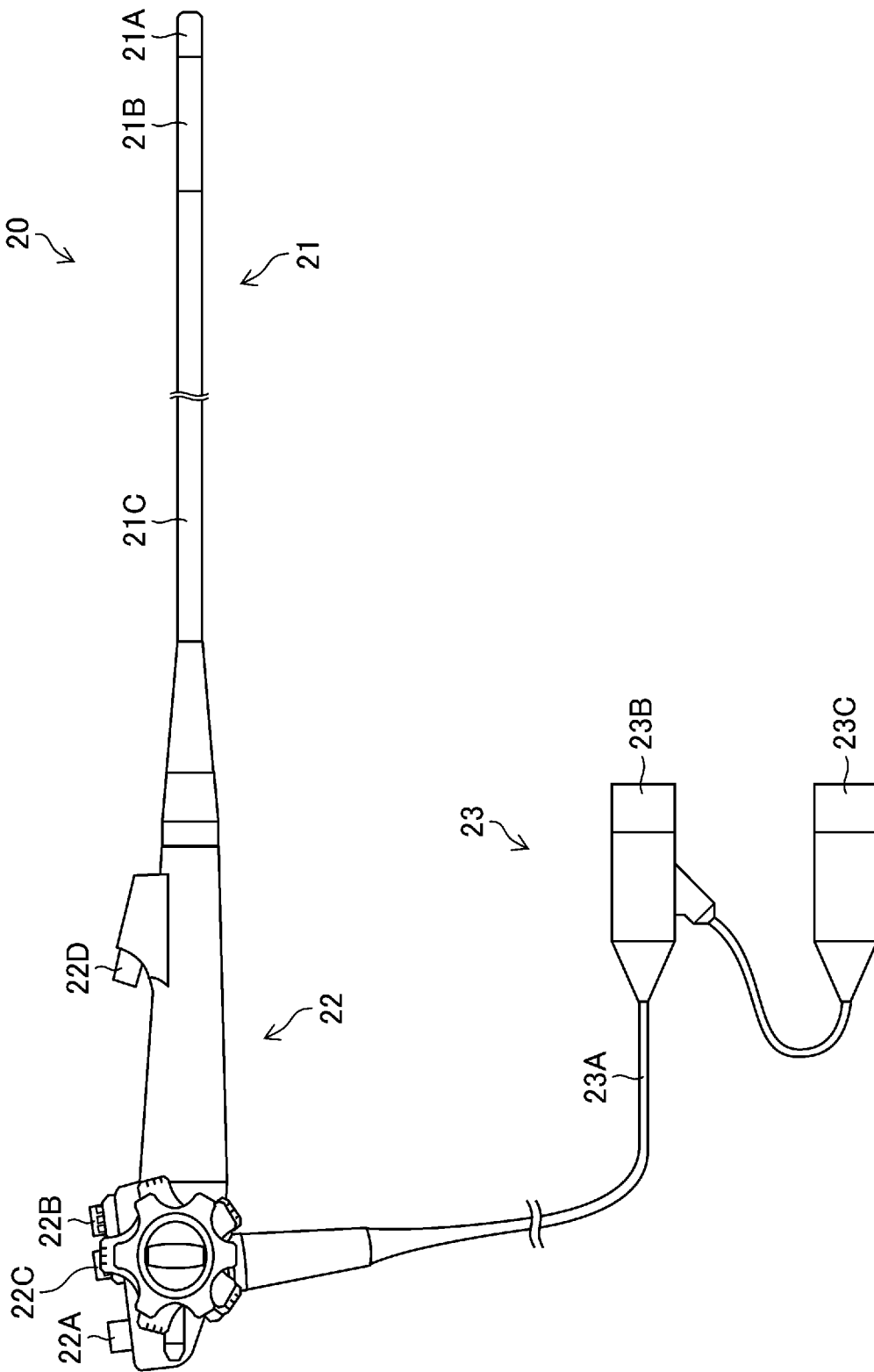
[図1]



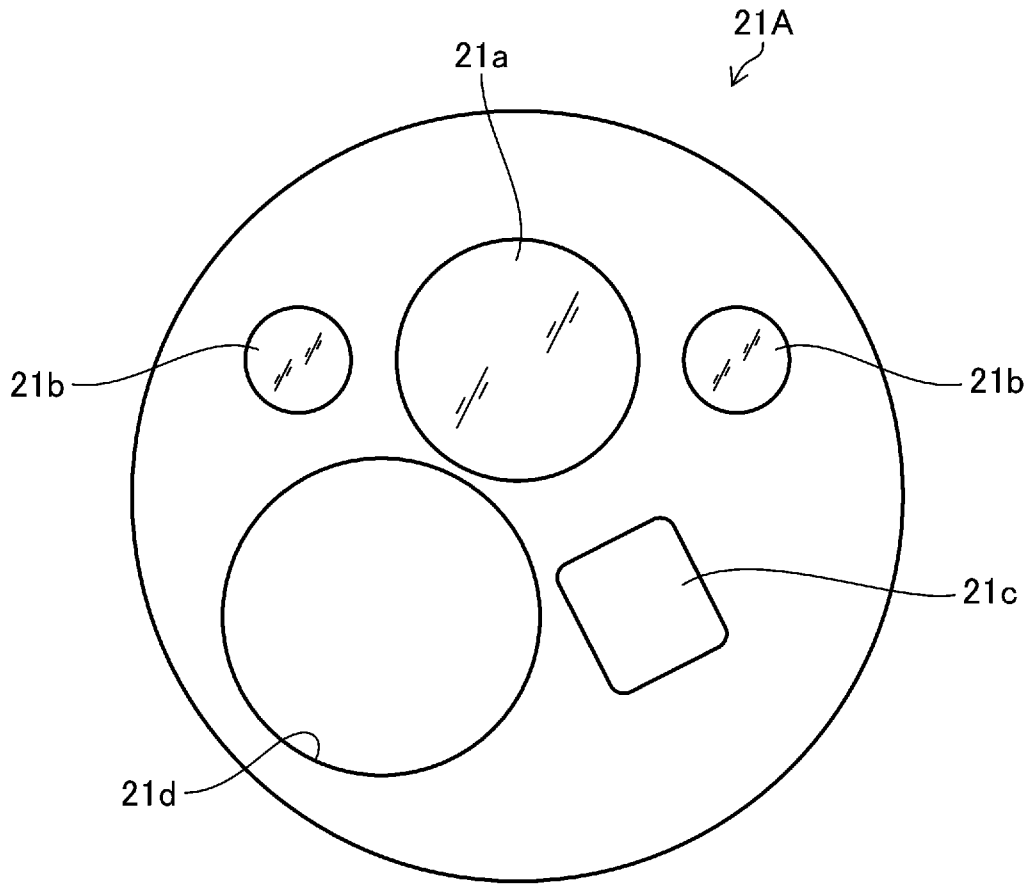
[図2]



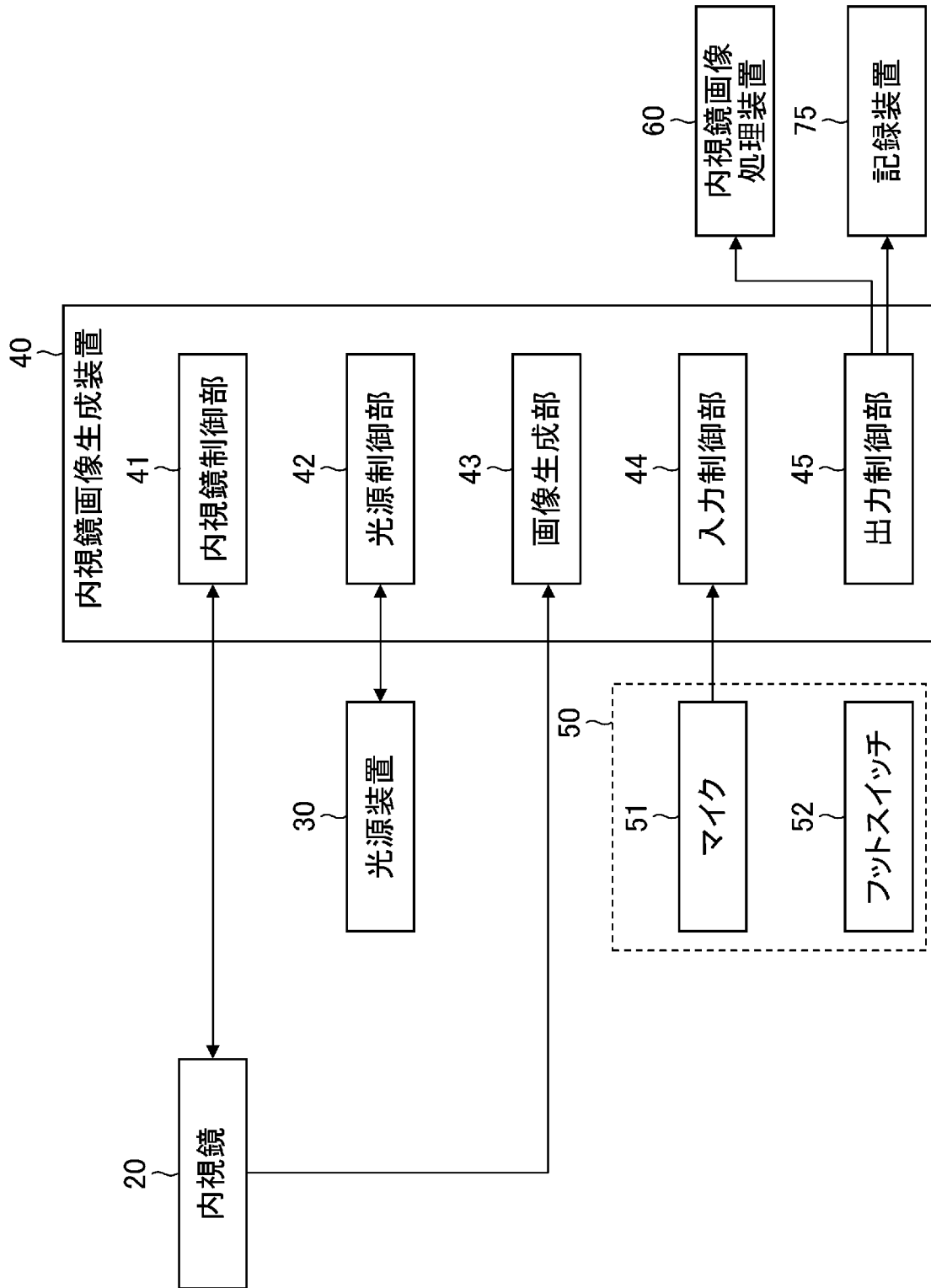
[図3]



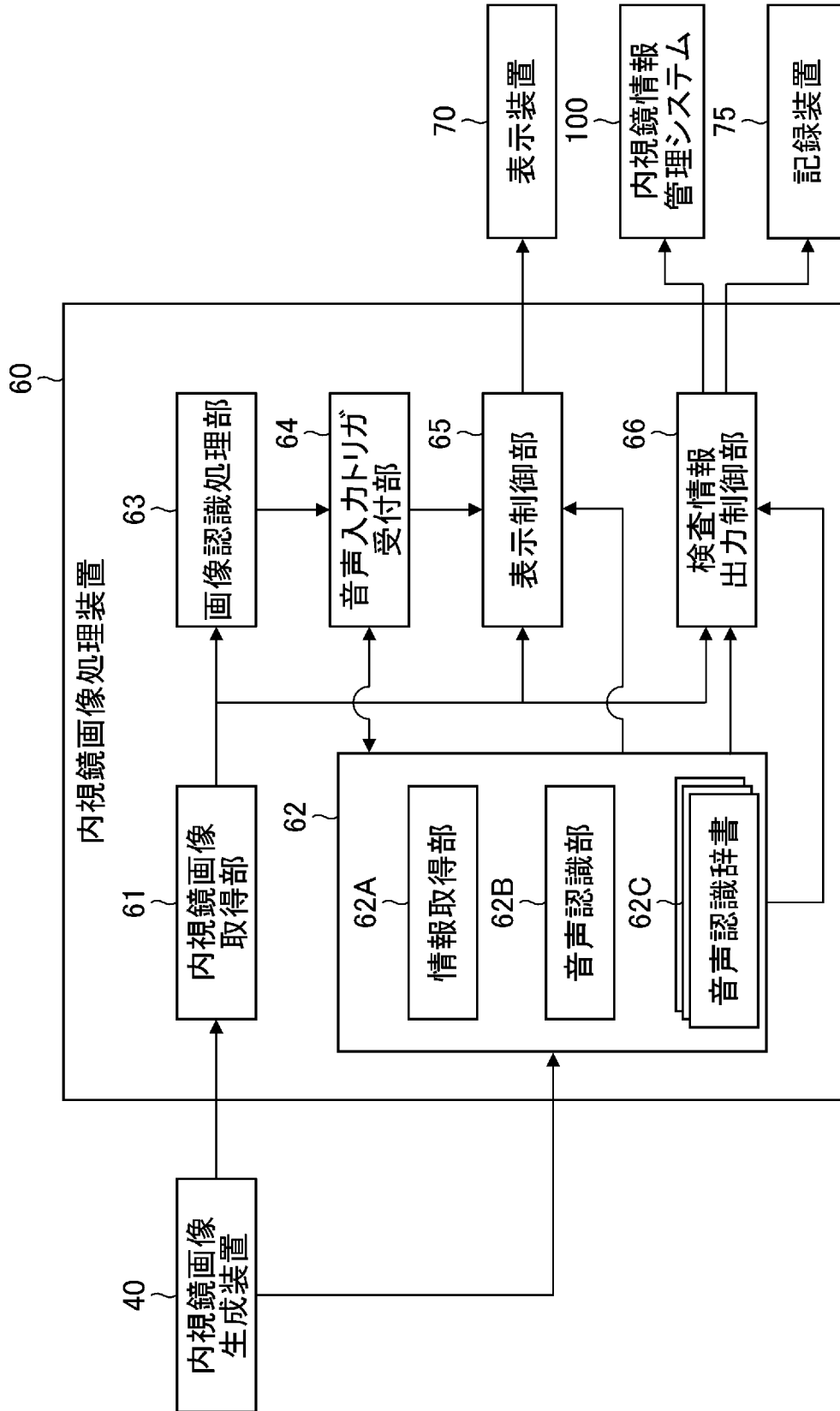
[図4]



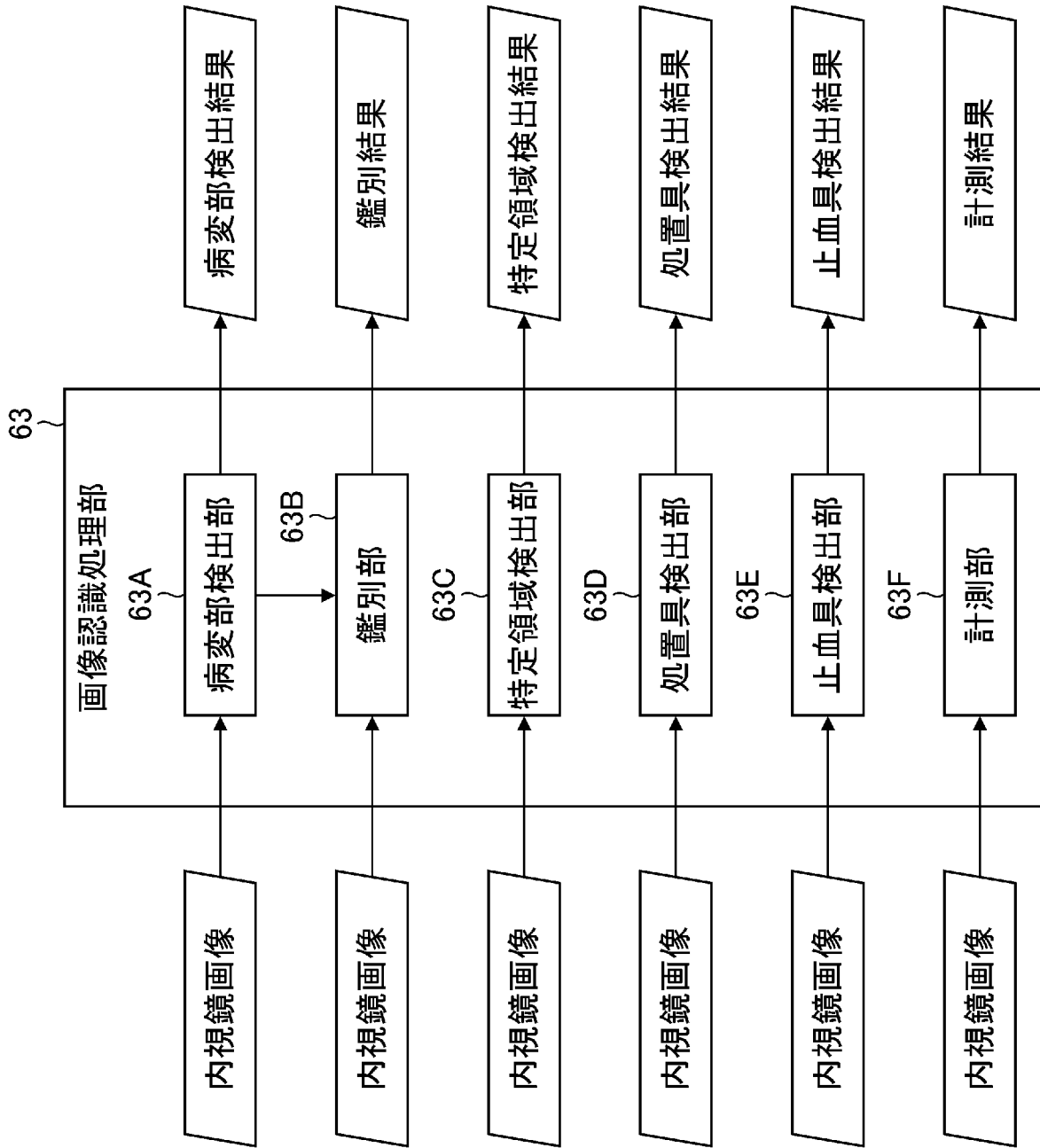
[図5]



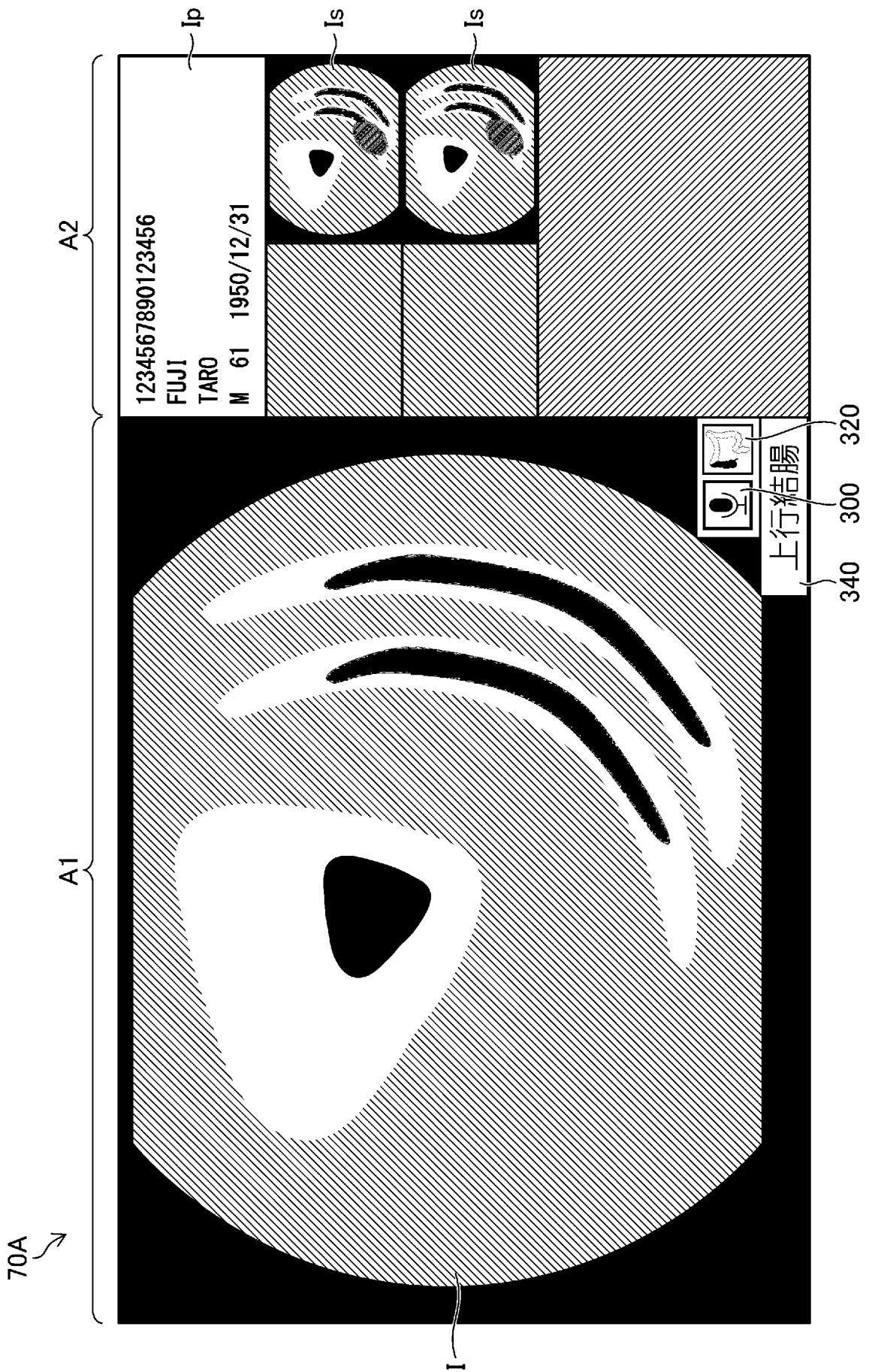
[図6]



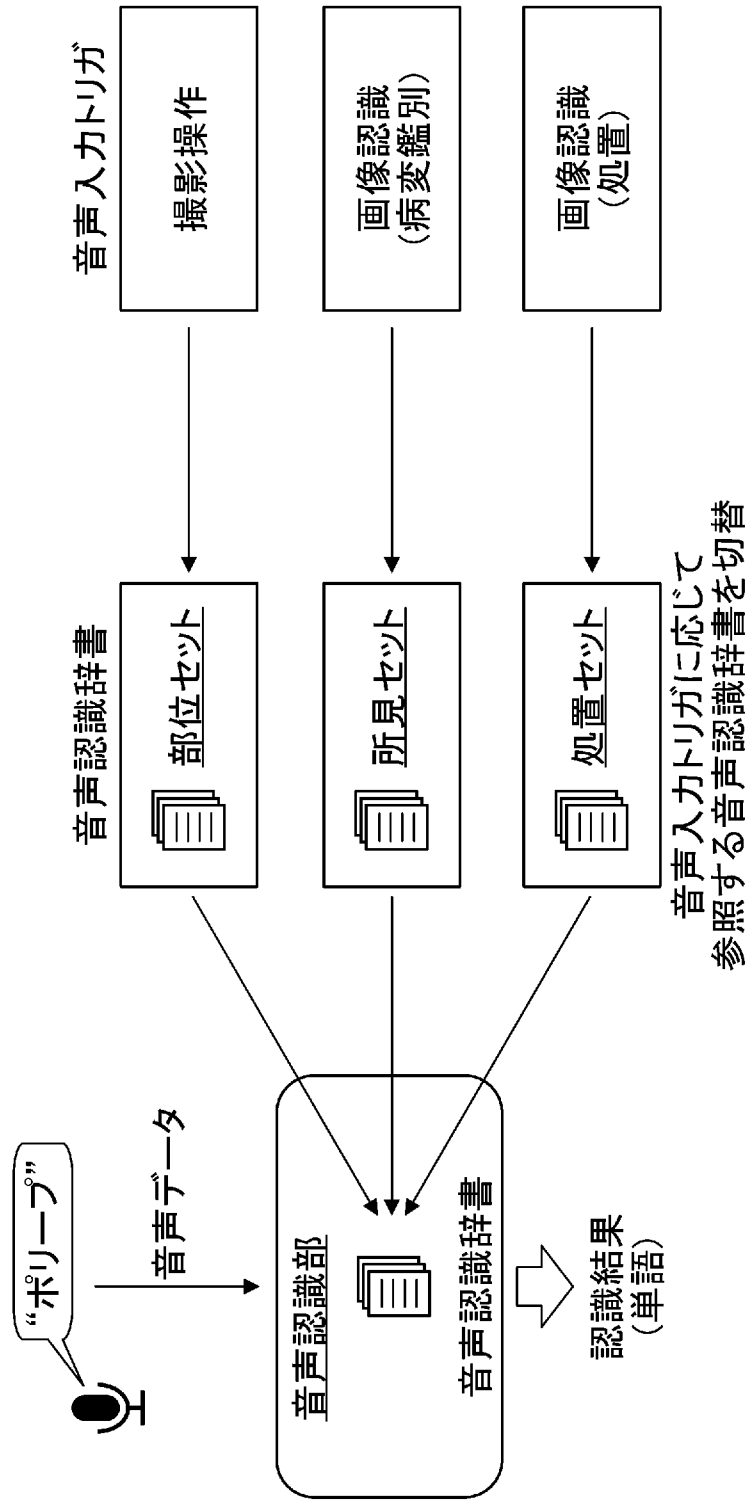
[図7]



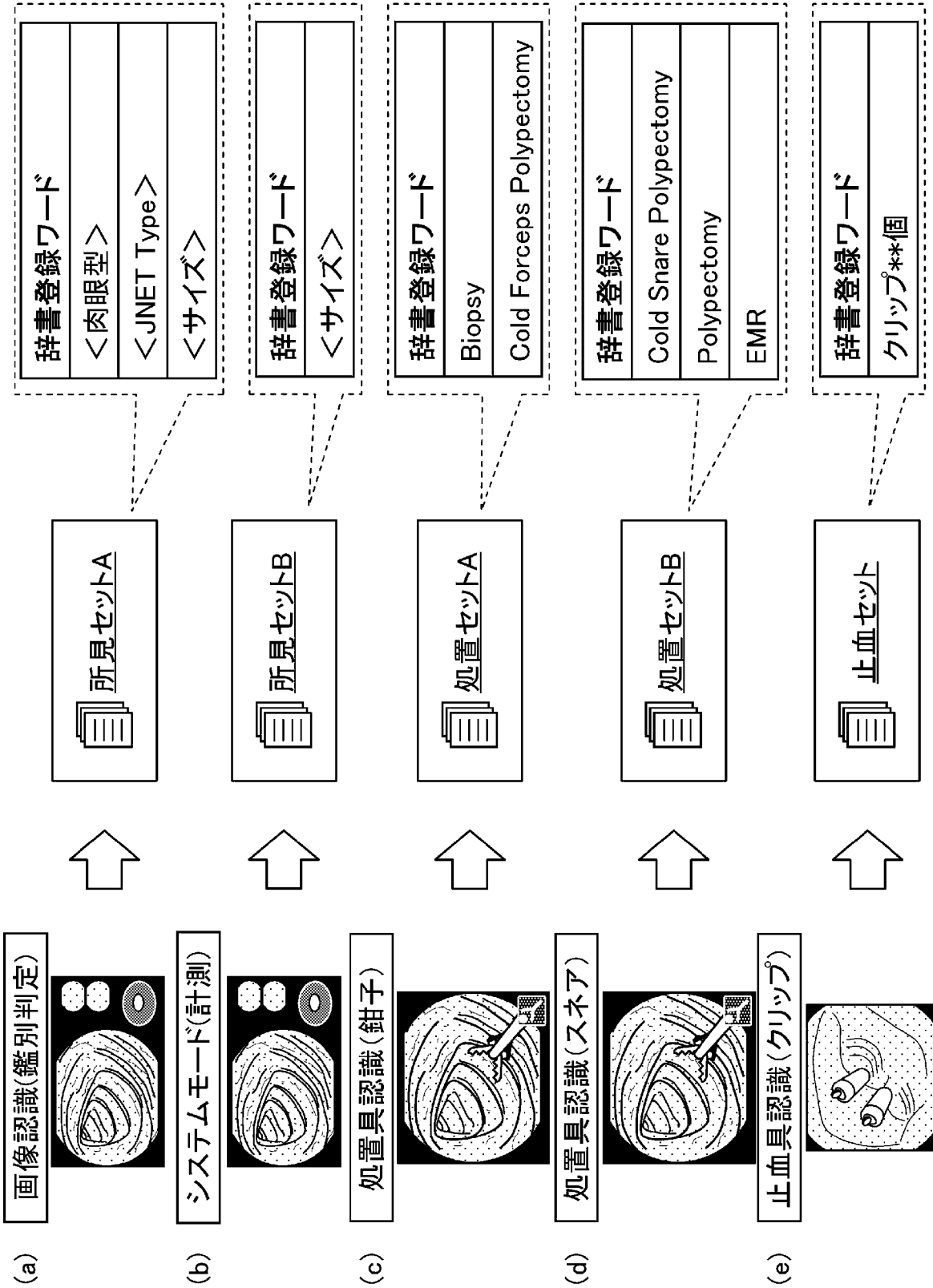
[図8]



[図9]

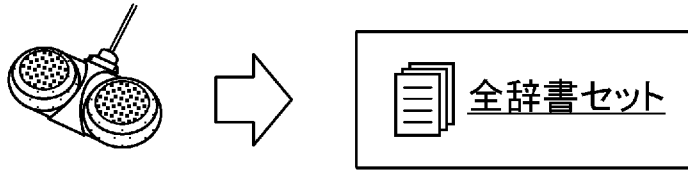


[図10]

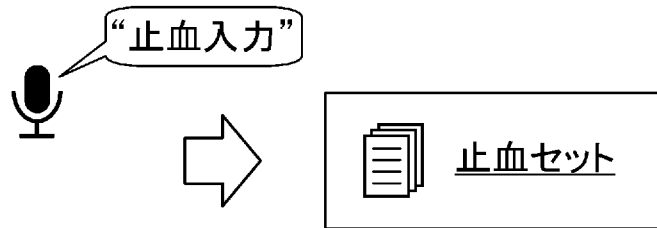
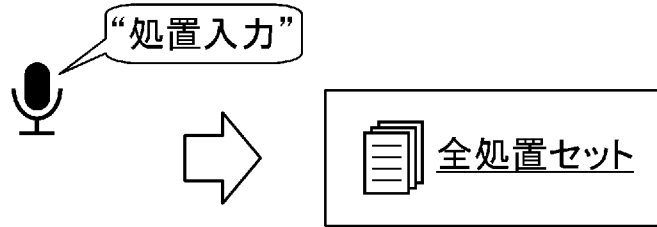
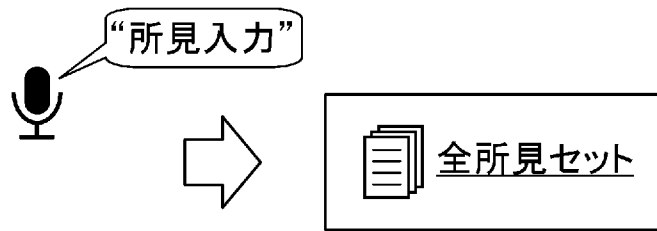


[図11]

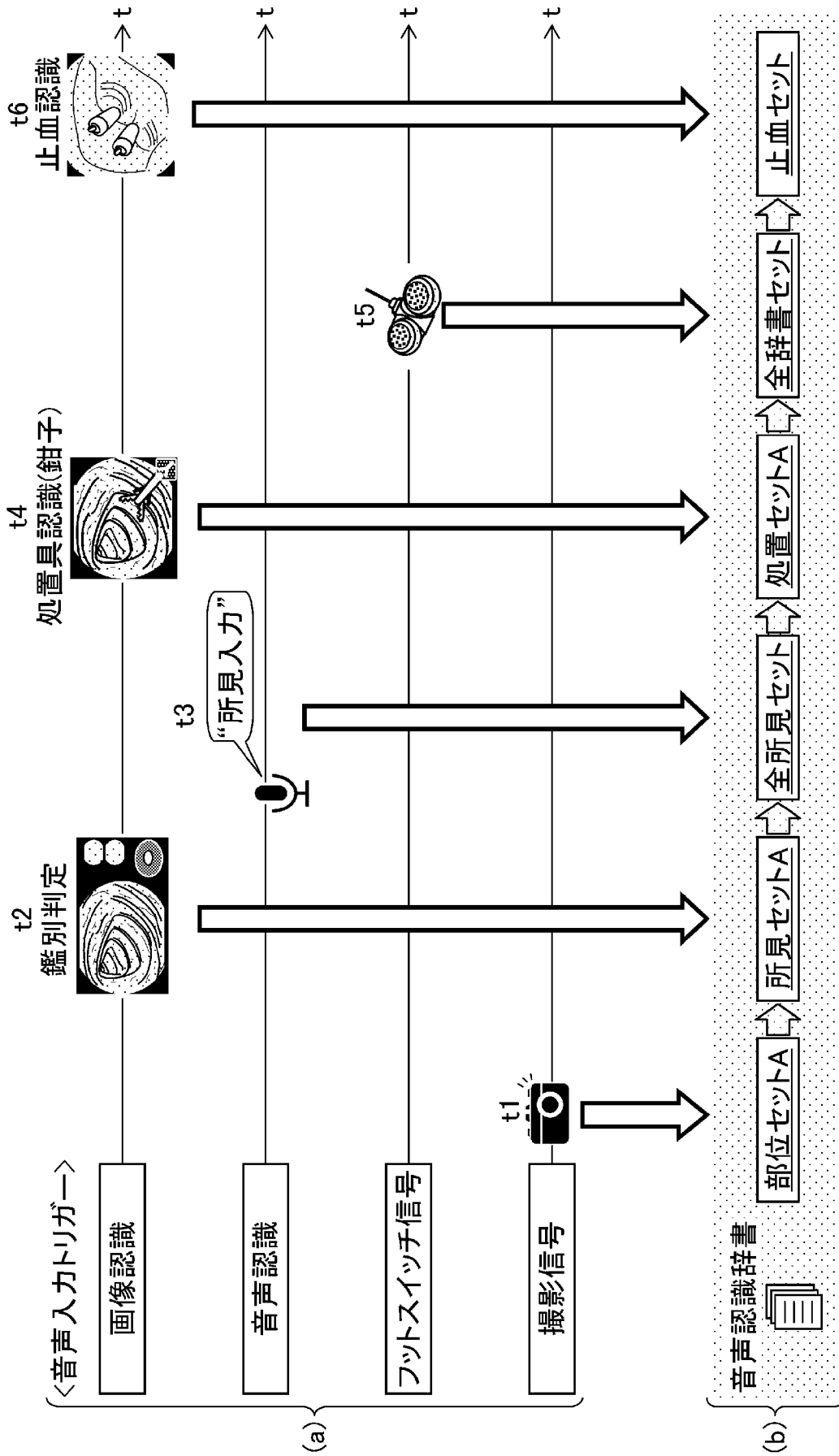
(a) フットスイッチ検出



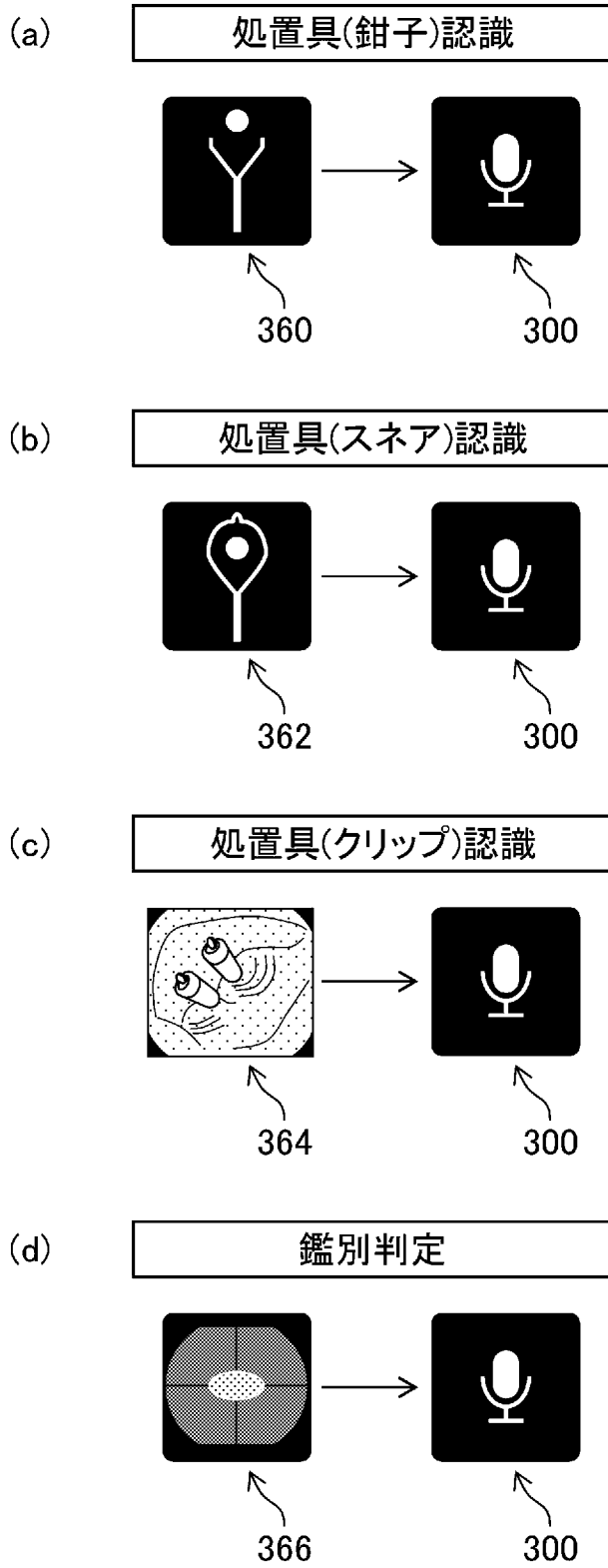
(b) ウェイクワード検出



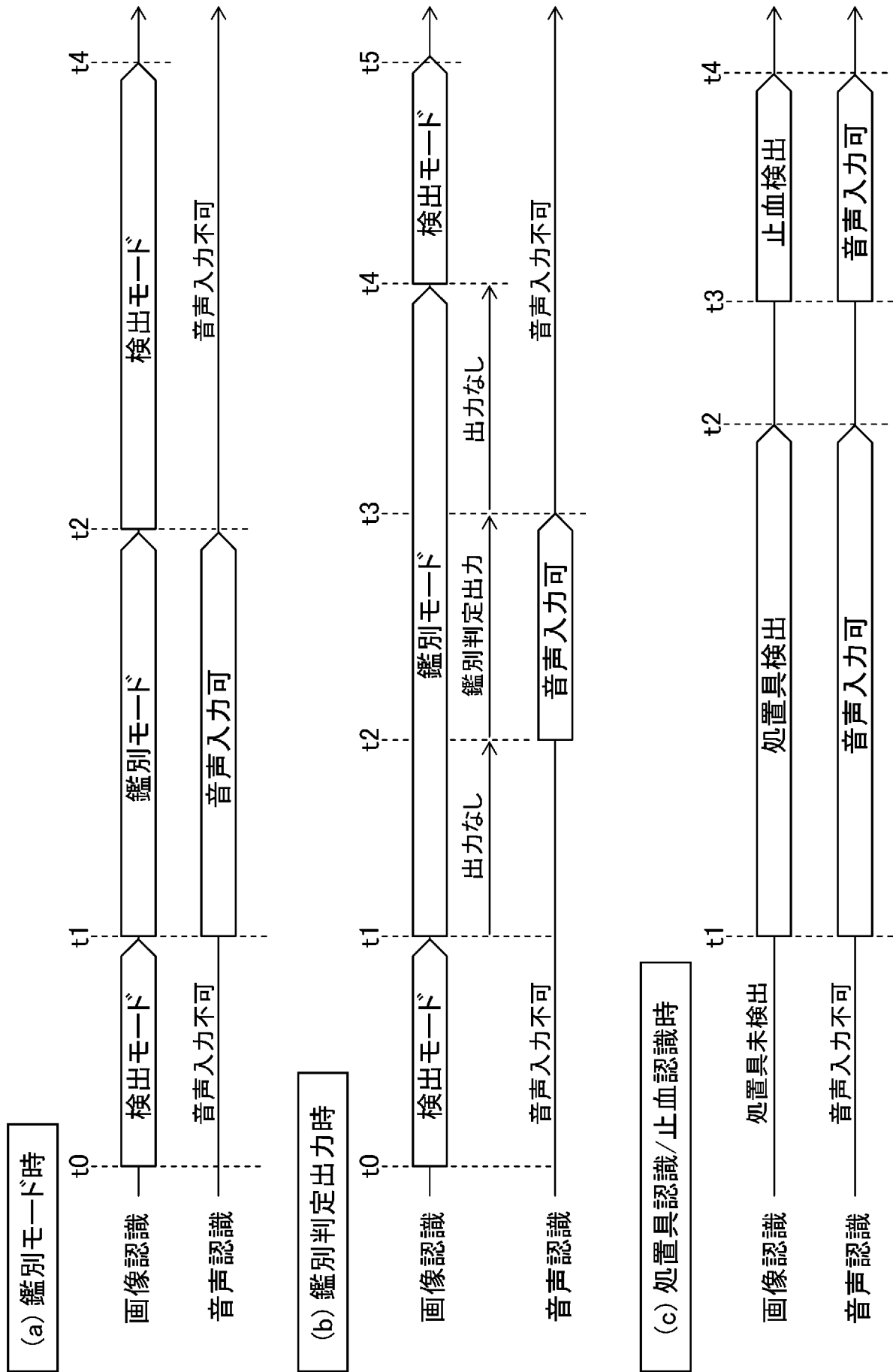
[図12]



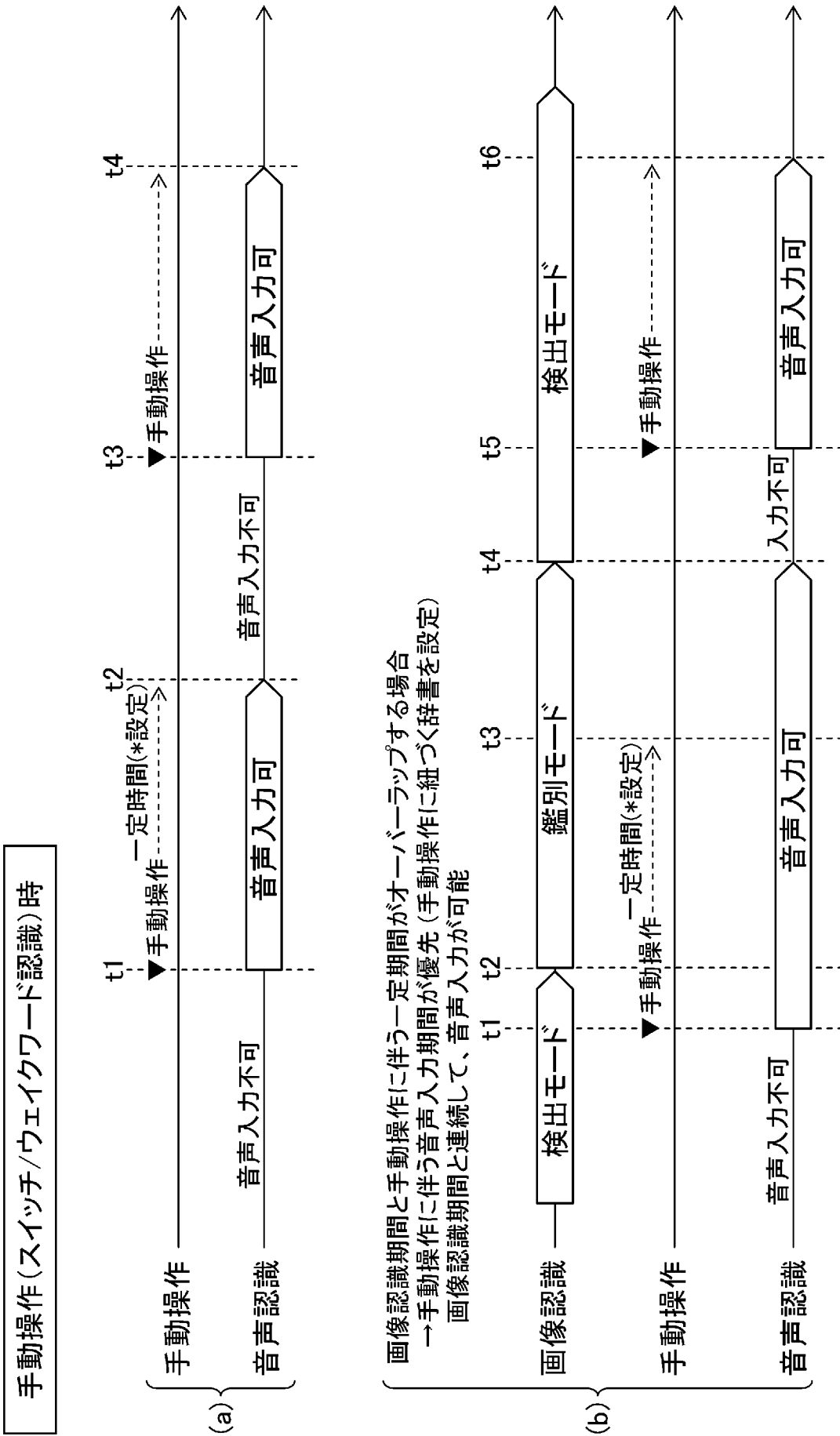
[図13]



[図14]



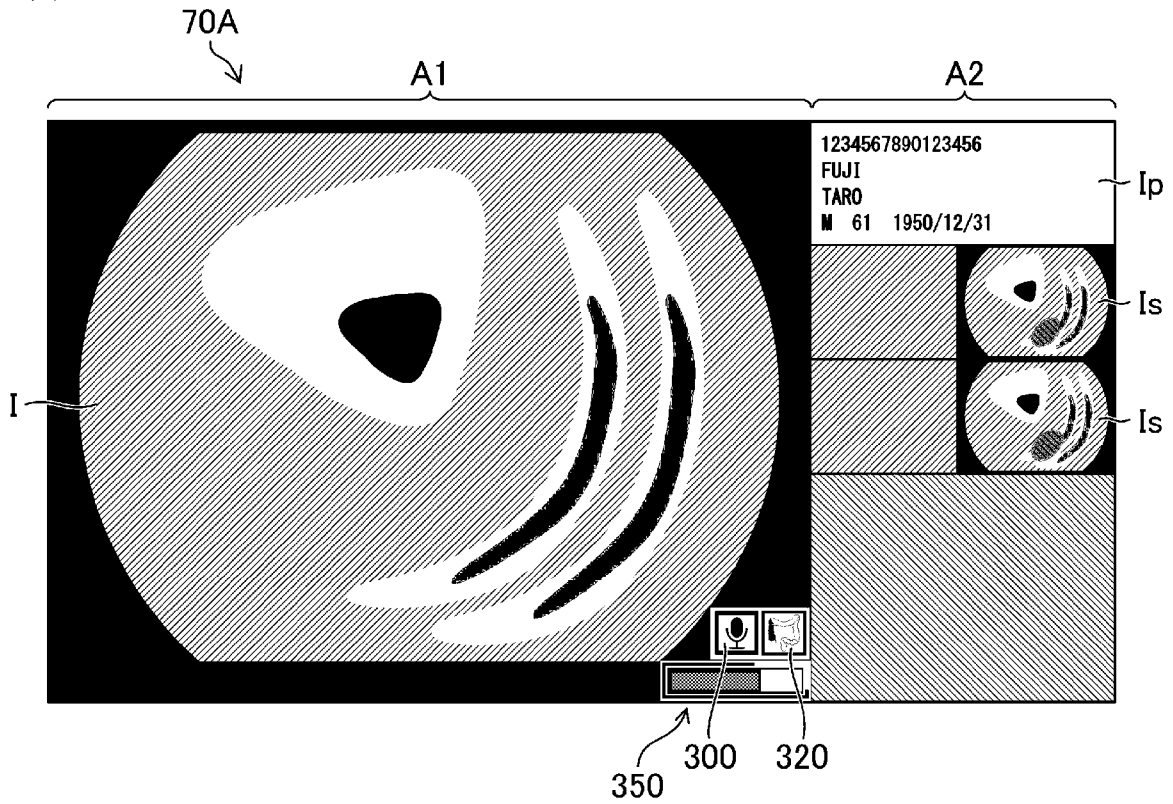
[図15]



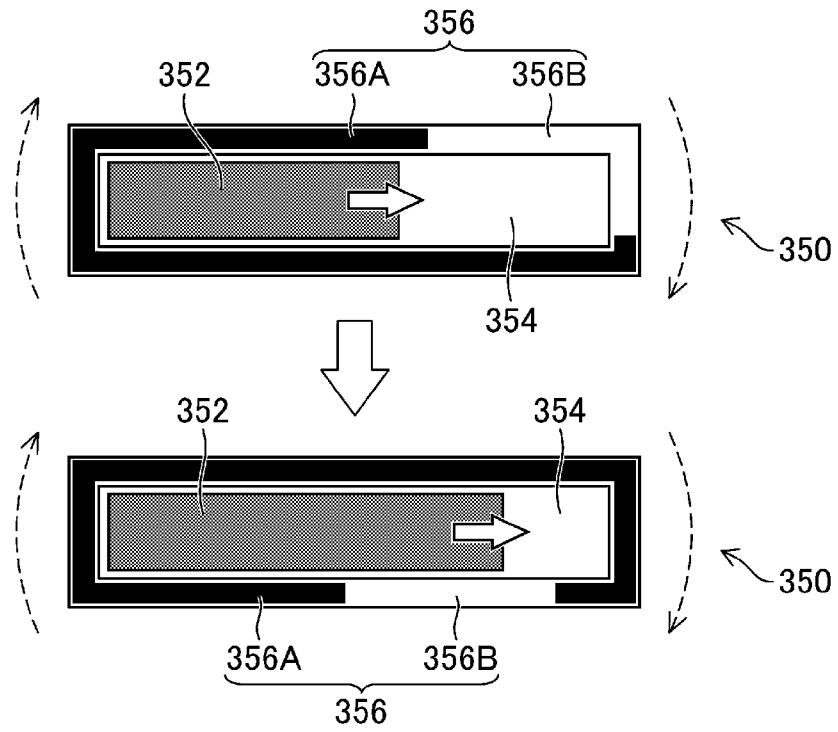
画像認識期間と手動操作に伴う一定期間がオーバーラップする場合  
 →手動操作に伴う音声入力期間が優先(手動操作に紐づく辞書を設定)  
 画像認識期間と連続して、音声入力が可能

[図16]

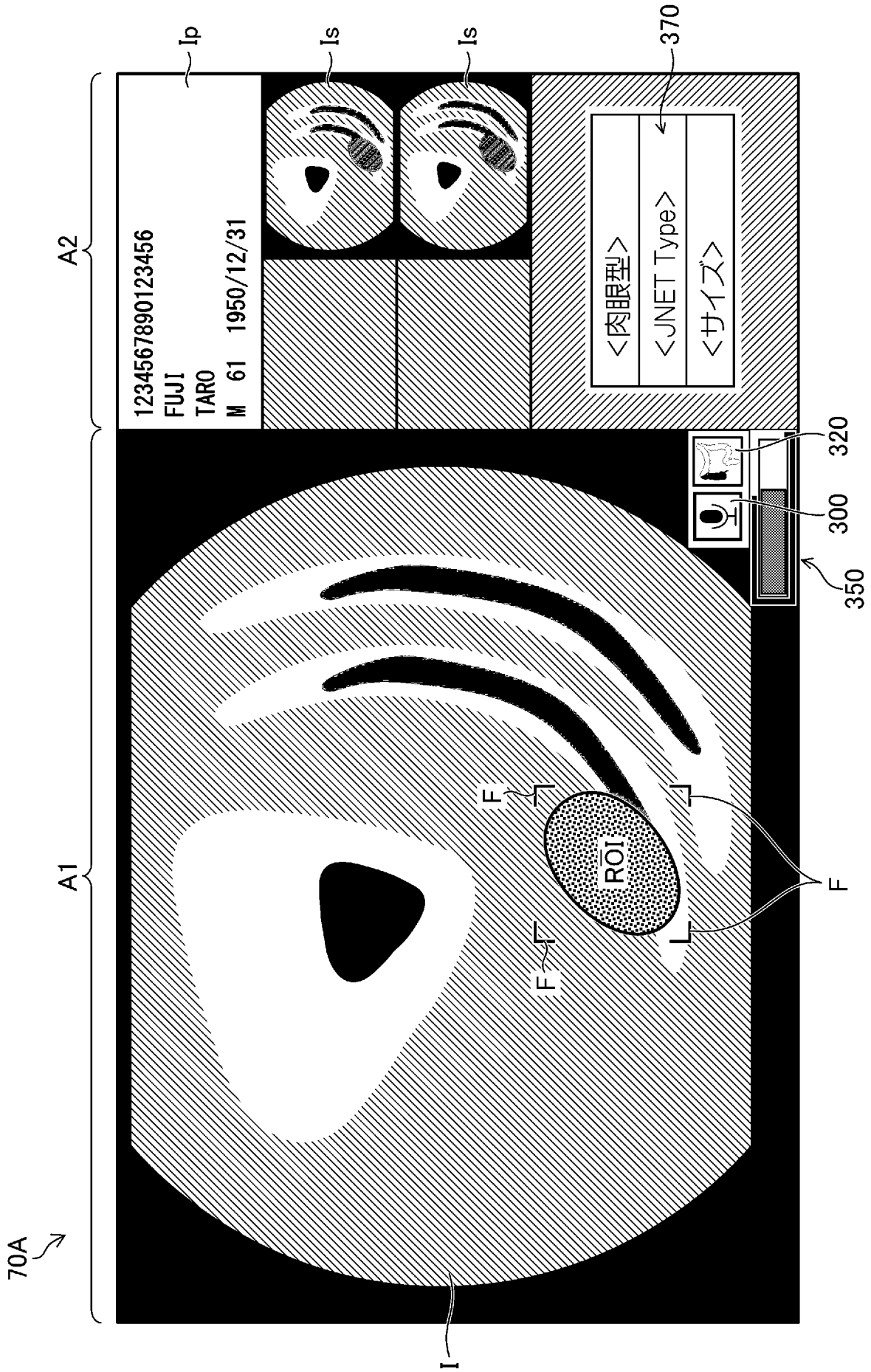
(a)



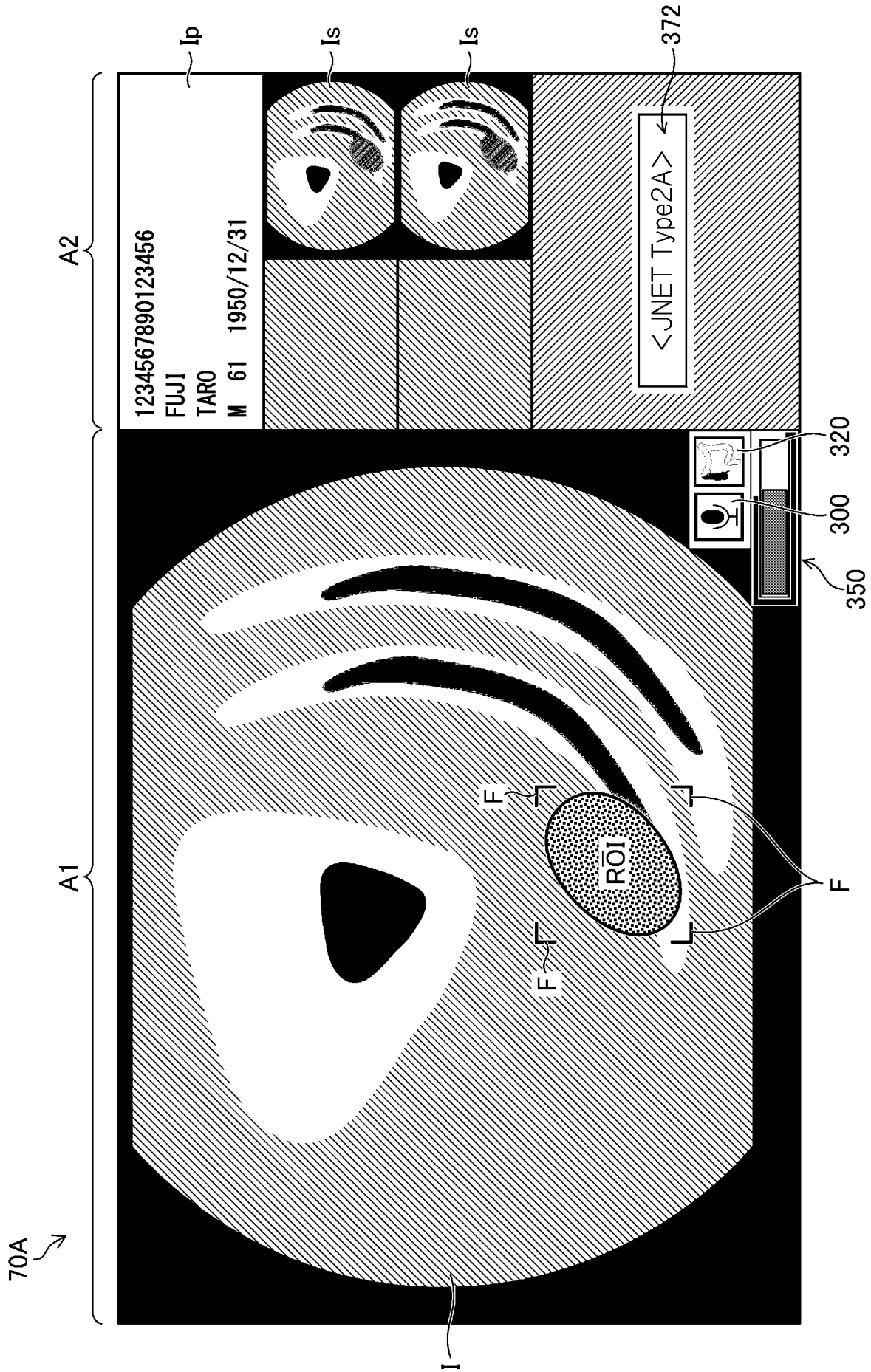
(b)



[図17]



[図18]





## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/033260

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
A61B 1/045(2006.01)j FI: A61B1/045 642; A61B1/045 618		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) A61B1/00-1/32		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Published examined utility model applications of Japan 1922-1996 Published unexamined utility model applications of Japan 1971-2022 Registered utility model specifications of Japan 1996-2022 Published registered utility model applications of Japan 1994-2022		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2017/187676 A1 (SONY CORP) 02 November 2017 (2017-11-02) paragraphs [0014]-[0097], fig. 1-11	1, 15, 17-21
Y		2-14, 16
X	JP 2008-136646 A (TOSHIBA CORP) 19 June 2008 (2008-06-19) paragraphs [0003]-[0004], [0014]-[0039], fig. 1-5	1, 15-21
Y		2-3
A		4-14
Y	中村 健一, 内視鏡検査向け音声認識によるリアルタイム所見入力システム「Voice Capture」, 映像情報インダストリアル, 01 January 2016, vol. 48, no. 1, pp. 67-71 in particular, p. 68, right column, "2.2 Voice recognition", (NAKAMURA, Kenichi. Eizo Joho Industrial.), non-official translation (Real-time observation input system "Voice Capture" with voice recognition for endoscopy)	2-3
Y	WO 2019/078102 A1 (FUJIFILM CORP) 25 April 2019 (2019-04-25) paragraph [0049]	4-14, 16
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 <b>25 October 2022</b>		Date of mailing of the international search report <b>08 November 2022</b>
Name and mailing address of the ISA/JP <b>Japan Patent Office (ISA/JP) 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915 Japan</b>		Authorized officer  Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

**PCT/JP2022/033260**

<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2017-221486 A (SONY CORP) 21 December 2017 (2017-12-21) paragraph [0071]	4-14, 16
Y	JP 2006-136385 A (PENTAX CORP) 01 June 2006 (2006-06-01) paragraph [0006]	11-14

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

International application No.

**PCT/JP2022/033260**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
WO	2017/187676	A1	02 November 2017	US 2019/0125319 A1 paragraphs [0033]-[0116], fig. 1-11 EP 3434219 A1 CN 109069221 A	
JP	2008-136646	A	19 June 2008	US 2008/0133233 A1 paragraphs [0005]-[0006], [0020]-[0054], fig. 1-5	
WO	2019/078102	A1	25 April 2019	US 2020/0242766 A1 paragraph [0086]	
JP	2017-221486	A	21 December 2017	US 2019/0180865 A1 paragraph [0098] WO 2017/217107 A1 EP 3471591 A1 CN 109310279 A	
JP	2006-136385	A	01 June 2006	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） A61B 1/045(2006.01)i FI: A61B1/045 642; A61B1/045 618		
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）） A61B1/00-1/32 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 1971-2022年 日本国実用新案登録公報 1996-2022年 日本国登録実用新案公報 1994-2022年		
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 2017/187676 A1 (ソニー株式会社) 02.11.2017 (2017-11-02) 段落[0014]-[0097], 図1-11	1, 15, 17-21
Y		2-14, 16
X	JP 2008-136646 A (株式会社東芝) 19.06.2008 (2008-06-19) 段落[0003]-[0004], [0014]-[0039], 図1-5	1, 15-21
Y		2-3
A		4-14
Y	中村 健一, 内視鏡検査向け 音声認識によるリアルタイム所見入力システム「Voice Capture」, 映像情報インダストリアル, 2016.01.01, 第48巻, 第1号, p.67-71 特に, p.68右欄「2.2 音声認識」	2-3
Y	WO 2019/078102 A1 (富士フイルム株式会社) 25.04.2019 (2019-04-25) 段落[0049]	4-14, 16
Y	JP 2017-221486 A (ソニー株式会社) 21.12.2017 (2017-12-21) 段落[0071]	4-14, 16
<input checked="" 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 “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 25.10.2022	国際調査報告の発送日 08.11.2022	
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） ▲高▼ 芳徳 2Q 9813 電話番号 03-3581-1101 内線 3292	



国際調査報告  
 パテントファミリーに関する情報

国際出願番号

PCT/JP2022/033260

引用文献	公表日	パテントファミリー文献	公表日
WO 2017/187676 A1	02.11.2017	US 2019/0125319 A1 段落[0033]-[0116], 図1-11 EP 3434219 A1 CN 109069221 A	
JP 2008-136646 A	19.06.2008	US 2008/0133233 A1 段落[0005]-[0006], [0020]- [0054], 図1-5	
WO 2019/078102 A1	25.04.2019	US 2020/0242766 A1 段落[0086]	
JP 2017-221486 A	21.12.2017	US 2019/0180865 A1 段落[0098] WO 2017/217107 A1 EP 3471591 A1 CN 109310279 A	
JP 2006-136385 A	01.06.2006	(ファミリーなし)	