

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2014-191237

(P2014-191237A)

(43) 公開日 平成26年10月6日(2014.10.6)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 15/28 (2013.01)	G 1 0 L 15/28 5 0 0	
G 1 0 L 15/00 (2013.01)	G 1 0 L 15/00 2 0 0 G	
G 1 0 L 15/10 (2006.01)	G 1 0 L 15/10 2 0 0 W	
G 0 6 F 3/16 (2006.01)	G 0 6 F 3/16 3 2 0 H	

審査請求 未請求 請求項の数 7 O L (全 16 頁)

(21) 出願番号 特願2013-67631 (P2013-67631)
 (22) 出願日 平成25年3月27日 (2013. 3. 27)

(71) 出願人 000005267
 ブラザー工業株式会社
 愛知県名古屋市瑞穂区苗代町15番1号
 (74) 代理人 100104178
 弁理士 山本 尚
 (72) 発明者 片野 智己
 愛知県名古屋市瑞穂区苗代町15番1号
 ブラザー工業株式会社内
 (72) 発明者 伊藤 邦宏
 愛知県名古屋市瑞穂区苗代町15番1号
 ブラザー工業株式会社内

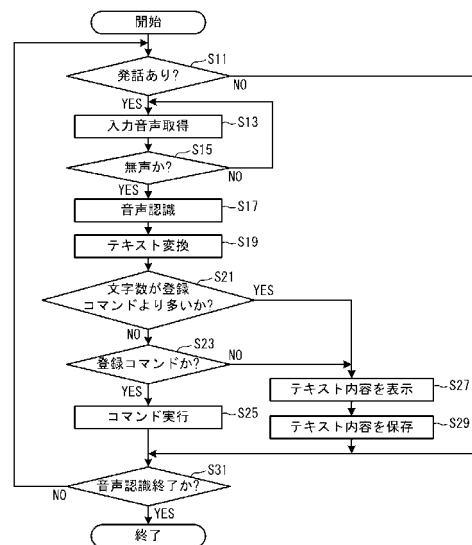
(54) 【発明の名称】 情報記憶装置および情報記憶プログラム

(57) 【要約】

【課題】ユーザの発声内容に基づく文字データで、装置の制御に係る指示データが含まれない文字データを記憶することができる情報記憶装置および情報記憶プログラムを提供する。

【解決手段】CPUは、ユーザの発声した音声データを認識して(S17)、テキスト変換し(S19)、ユーザテキストデータを得る。CPUは、ユーザテキストデータが登録コマンドでなければ(S23:NO)、ユーザテキストデータを保存する(S29)。CPUは、ユーザテキストデータの文字数が登録コマンドの文字数より多ければ(S21:YES)、登録コマンドが否か判断を省略して、ユーザテキストデータを保存する(S29)。CPUは、ユーザテキストデータが登録コマンドである場合(S23:YES)、ユーザテキストデータは保存せず、コマンドを実行する(S25)。

【選択図】 図3



【特許請求の範囲】**【請求項 1】**

ユーザの頭部に装着可能に構成し、前記ユーザの発声内容に基づく文字のデータである文字データを記憶部に記憶可能な情報記憶装置であって、

前記ユーザの発声する音声を取得する取得手段と、

前記取得手段によって取得された音声のデータである音声データに基づいて前記情報記憶装置を制御可能な制御手段と

を備え、

前記制御手段は、

前記音声データを解析し、前記文字データを生成する生成部と、

前記情報記憶装置の制御を指示する指示データと、前記情報記憶装置に実行させる動作の制御信号とが対応付けられた対応データを参照し、前記生成部において生成された前記文字データに前記指示データが含まれるか否か判断する第一判断部と、

前記第一判断部において前記文字データに前記指示データが含まれると判断された場合、前記指示データに対応する制御信号を出力する信号出力部と、

前記第一判断部において前記文字データに前記指示データが含まれないと判断された場合、前記文字データを前記記憶部に記憶する処理を行う一方、前記文字データに前記指示データが含まれる場合には前記文字データを前記記憶部に記憶しない記憶制御部と、

を含むことを特徴とする情報記憶装置。

10

【請求項 2】

20

前記制御手段は、

前記取得手段によって取得された音声に含まれる無声状態を検出する検出部と、

前記検出部において検出された連続する 2 回の無声状態の間に、前記取得手段によって取得され、前記生成部において生成された前記文字データが示す文字の文字数が、所定数より多いか否か判断する第二判断部と

をさらに含み、

前記記憶制御部は、前記第二判断部において前記文字数が所定数よりも多いと判断された場合、前記第一判断部における前記文字データに対する判断処理は行わずに、前記文字データを前記記憶部に記憶する処理を行うことを特徴とする情報記憶装置。

【請求項 3】

30

前記対応データは、複数の前記指示データと前記制御信号との組を含み、

前記指示データは、所定の文字列で示されるデータであり、

前記第二判断部が判断対象とする前記所定数は、前記対応データに含まれる複数の前記指示データのうち、文字数が最多である指示データの文字数以上の数であることを特徴とする請求項 2 に記載の情報記憶装置。

【請求項 4】

前記ユーザが視認可能な画像を表示する表示手段を備え、

前記制御手段は、前記生成部において生成された前記文字データが示す文字を前記表示手段に表示する表示制御部をさらに含むことを特徴とする請求項 1 から 3 のいずれかに記載の情報記憶装置。

40

【請求項 5】

前記表示制御部は、前記第一判断部において前記指示データが含まれると判断された前記文字データが示す文字を前記表示手段に表示せず、前記指示データが含まれない前記文字データが示す文字を前記表示手段に表示する処理を行うことを特徴とする請求項 4 に記載の情報記憶装置。

【請求項 6】

前記ユーザの周囲の様子を撮影可能な撮影手段を備え、

前記記憶制御部は、前記撮影手段によって撮影された撮影画像のデータである撮影データに、前記撮影画像が撮影された後、少なくとも最初に前記生成手段によって生成された前記文字データを対応付けて、前記撮影データと前記文字データとを前記記憶部に記憶す

50

る処理を行うことを特徴とする請求項 1 から 5 のいずれかに記載の情報記憶装置。

【請求項 7】

ユーザの頭部に装着可能に構成し、前記ユーザの発声する音声を取得する取得手段を備え、前記ユーザの発声内容に基づく文字のデータである文字データを記憶部に記憶可能な情報記憶装置のコンピュータが実行可能なプログラムであって、

前記コンピュータに、

前記取得手段によって取得された音声のデータである音声データを解析し、前記文字データを生成する生成ステップと、

前記情報記憶装置の制御を指示する指示データと、前記情報記憶装置に実行させる動作の制御信号とが対応付けられた対応データを参照し、前記生成ステップにおいて生成された前記文字データに前記指示データが含まれるか否か判断する第一判断ステップと、

前記第一判断ステップにおいて前記文字データに前記指示データが含まれると判断された場合、前記指示データに対応する制御信号を出力する信号出力ステップと、

前記第一判断ステップにおいて前記文字データに前記指示データが含まれないと判断された場合、前記文字データを前記記憶部に記憶する処理を行う一方、前記文字データに前記指示データが含まれる場合には前記文字データを前記記憶部に記憶しない記憶制御ステップと、

を実行させることを特徴とする情報記憶プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ユーザが発声した文字データを記憶可能な情報記憶装置および情報記憶プログラムに関する。

【背景技術】

【0002】

ユーザの頭部に装着され、ユーザが視認可能に画像を表示できる画像表示装置において、ユーザが発する音声に基づいて、装置を制御することができる画像表示装置が知られている（例えば特許文献 1 参照）。特許文献 1 の画像表示装置は、ユーザがマイクロフォン（以下、「マイク」と略す。）を介して入力した音声を認識し、認識結果が事前に登録されたキーワードに一致するとき、画像表示部に表示する画像の切り換え等を指示する指示信号を出力する。特許文献 1 の画像表示装置はプレゼンテーションでの使用を意図したものであるが、ユーザの発声内容に基づく文字データを議事録として保存したいという要望がある。

【0003】

ユーザの発声した音声を認識し、文字データに変換して保存することができる情報記憶装置が知られている。ユーザは、特許文献 1 の画像表示装置を使用しながら情報記憶装置を用いて文字データを保存すれば、ユーザの発声内容に基づく文字データを議事録として保存することが可能である。

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開 2002 - 165156 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、情報記憶装置は、ユーザのすべての発声内容に基づく文字データを保存すると、ユーザが画像表示装置を操作するため、例えば表示画像の切り換え等を指示するキーワードとして発声する音声の文字データも含めて議事録に保存してしまうという問題があった。

【0006】

10

20

30

40

50

本発明は、上記の問題点を解決するためになされたものであり、ユーザの発声内容に基づく文字データで、装置の制御に係る指示データが含まれない文字データを記憶することができる情報記憶装置および情報記憶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】

【0007】

本発明の第1態様によれば、ユーザの頭部に装着可能に構成し、前記ユーザの発声内容に基づく文字のデータである文字データを記憶部に記憶可能な情報記憶装置であって、前記ユーザの発声する音声を取得する取得手段と、前記取得手段によって取得された音声のデータである音声データに基づいて前記情報記憶装置を制御可能な制御手段とを備え、前記制御手段は、前記音声データを解析し、前記文字データを生成する生成部と、前記情報記憶装置の制御を指示する指示データと、前記情報記憶装置に実行させる動作の制御信号とが対応付けられた対応データを参照し、前記生成部において生成された前記文字データに前記指示データが含まれるか否か判断する第一判断部と、前記第一判断部において前記文字データに前記指示データが含まれると判断された場合、前記指示データに対応する制御信号を出力する信号出力部と、前記第一判断部において前記文字データに前記指示データが含まれないと判断された場合、前記文字データを前記記憶部に記憶する処理を行う一方、前記文字データに前記指示データが含まれる場合には前記文字データを前記記憶部に記憶しない記憶制御部と、を含むことを特徴とする情報記憶装置が提供される。

10

【0008】

第1態様の情報記憶装置は、情報記憶装置の制御をユーザが発声する音声データに基づいて行いながらも、ユーザが制御のために発声した指示データを含まず、ユーザが制御目的ではなく発声した文字データのみを保存することができる。ユーザの音声に基づく文字データが指示データを含んだ状態で保存された場合、保存した文字データを議事録として用いるには、不要な指示データが含まれているため好ましくないが、第1態様の情報記憶装置を用いれば指示データを除いた文字データを保存できるので、保存した文字データを議事録として利用することができる。

20

【0009】

第1態様において、前記制御手段は、前記取得手段によって取得された音声に含まれる無声状態を検出する検出部と、前記検出部において検出された連続する2回の無声状態の間に、前記取得手段によって取得され、前記生成部において生成された前記文字データが示す文字の文字数が、所定数より多いか否か判断する第二判断部とをさらに含んでもよい。この場合に、前記記憶制御部は、前記第二判断部において前記文字数が所定数より多いと判断された場合、前記第一判断部における前記文字データに対する判断処理は行わずに、前記文字データを前記記憶部に記憶する処理を行ってもよい。

30

【0010】

第1態様の情報記憶装置は、文字データが示す文字の文字数が所定数よりも多いか否か、文字数を比較するという単純な判断処理を行うだけで、多い場合には第一判断部における判断処理を回避できるので、処理の単純化および高速化を図ることができる。所定数としては、例えば、指示データが示す文字の文字数を挙げることができる。この場合、文字データの文字数が指示データの文字数よりも多ければ、文字データに指示データが含まれないとみなすことができるので、第一判断部が文字データと指示データとを比較照合しなくとも済む。

40

【0011】

第1態様において、前記対応データは、複数の前記指示データと前記制御信号との組を含んでもよい。前記指示データは、所定の文字列で示されるデータであってもよい。前記第二判断部が判断対象とする前記所定数は、前記対応データに含まれる複数の前記指示データのうち、文字数が最多である指示データの文字数以上の数であってもよい。所定数が、文字数が最多の指示データの文字数以上であれば、文字データの文字数が所定数よりも多い場合に、文字データに指示データが含まれることはない。

【0012】

50

第1態様の情報記憶装置は、前記ユーザが視認可能な画像を表示する表示手段を備えてもよい。前記制御手段は、前記生成部において生成された前記文字データが示す文字を前記表示手段に表示する表示制御部をさらに含んでもよい。表示手段に文字データが示す文字を表示することで、情報記憶装置は、ユーザの発声内容を正しく認識できたか、ユーザに目視で確認させることができる。

【0013】

第1態様において、前記表示制御部は、前記第一判断部において前記指示データが含まれると判断された前記文字データが示す文字を前記表示手段に表示せず、前記指示データが含まれない前記文字データが示す文字を前記表示手段に表示する処理を行ってもよい。第1態様の情報記憶装置は、指示データが示す文字を表示手段に表示しないことで、ユーザに煩わしさを感じさせない。

10

【0014】

第1態様の情報記憶装置は、前記ユーザの周囲の様子を撮影可能な撮影手段を備えてもよい。この場合に、前記記憶制御部は、前記撮影手段によって撮影された撮影画像のデータである撮影データに、前記撮影画像が撮影された後、少なくとも最初に前記生成手段によって生成された前記文字データを対応付けて、前記撮影データと前記文字データとを前記記憶部に記憶する処理を行ってもよい。撮影画像と、撮影後に生成した文字データとを対応付けて記憶することで、文字データが示す文字を読む者は、撮影画像を参照することができ、内容を把握しやすい。

【0015】

本発明の第2態様によれば、ユーザの頭部に装着可能に構成し、前記ユーザの発声する音声を取得する取得手段を備え、前記ユーザの発声内容に基づく文字のデータである文字データを記憶部に記憶可能な情報記憶装置のコンピュータが実行可能なプログラムであって、前記コンピュータに、前記取得手段によって取得された音声のデータである音声データを解析し、前記文字データを生成する生成ステップと、前記情報記憶装置の制御を指示する指示データと、前記情報記憶装置に実行させる動作の制御信号とが対応付けられた対応データを参照し、前記生成ステップにおいて生成された前記文字データに前記指示データが含まれるか否か判断する第一判断ステップと、前記第一判断ステップにおいて前記文字データに前記指示データが含まれると判断された場合、前記指示データに対応する制御信号を出力する信号出力ステップと、前記第一判断ステップにおいて前記文字データに前記指示データが含まれないと判断された場合、前記文字データを前記記憶部に記憶する処理を行う一方、前記文字データに前記指示データが含まれる場合には前記文字データを前記記憶部に記憶しない記憶制御ステップと、を実行させる情報記憶プログラムが提供される。情報記憶装置のコンピュータが第2態様の情報記憶プログラムを実行することで、第1態様と同様の効果を得ることができる。

20

30

【図面の簡単な説明】

【0016】

【図1】HMD1の外観を示す斜視図である。

【図2】HMD1の電氣的構成を示すブロック図である。

【図3】第一実施形態における情報記憶プログラムを示すフローチャートである。

【図4】第二実施形態における情報記憶プログラムを示すフローチャートである。

【図5】第二実施形態の情報記憶プログラムに従う処理の流れを説明するための図である。

40

【発明を実施するための形態】

【0017】

以下、本発明を具体化した情報記憶装置の第一実施形態について、図面を参照して説明する。なお、参照する図面は、本発明が採用しうる技術的特徴を説明するために用いられるものである。図示された装置の構成等は、その形態のみに限定する趣旨ではなく、単なる説明例である。

【0018】

50

図1に示すように、本発明に係る情報記憶装置（以下、「ヘッドマウントディスプレイ」または「HMD」という。）1は、投影装置（以下、「ヘッドディスプレイ」または「HD」という。）10と制御装置（以下、「コントロールボックス」または「CB」という。）50を備える。HMD1が使用される場合、HD10は、装着対象物（例えばユーザの頭部）に装着される。また、CB50は、装着対象物のHD10とは異なる位置（例えば、ユーザの腰ベルトや腕等）に装着される。以下の説明において、図1の上方、下方、右斜め下方、左斜め上方、右斜め上方および左斜め下方がそれぞれ、HMD1の上方、下方、前方、後方、右方および左方である。

【0019】

HD10は専用の装着具である眼鏡5に装着される。HD10は、ユーザが日常的に使用する眼鏡、ヘルメット、ヘッドホンなど、他の装着具に取り付けてもよい。HD10は、受光対象物（例えばユーザの眼）に画像光を照射する。HD10はハーネス7を介してCB50と着脱可能に接続する。CB50は、HD10を制御する。

10

【0020】

HD10の構成について説明する。HD10は筐体2を備える。筐体2は、ユーザ側から見て右端側（図1における左側）にハーフミラー3を備える。ハーフミラー3は、ユーザがHD10を頭部に装着したとき、ユーザの眼（例えば左眼）の前方に配置される。HD10は、筐体2の内部に、画像表示部14（図2参照）と接眼光学系（図示略）を備える。画像表示部14は、CB50からハーネス7を介して送信される映像信号に基づいて画像を表示する。画像表示部14は、例えば、液晶素子等の空間変調素子および光源である。画像表示部14は、画像信号に応じた強度のレーザ光を2次元走査して画像表示を行う網膜走査型表示部や、有機EL（Organic Electro-luminescence）素子等であってもよい。

20

【0021】

接眼光学系は、画像表示部14に表示された画像を示す画像光を、ユーザの眼に導くために集光する。接眼光学系から射出された画像光は、筐体2の左側に設けられたハーフミラー3によって少なくとも一部（例えば半分）が反射され、ユーザの一方（例えば左）の眼球（図示略）に入射する。ハーフミラー3は外界の実像からの光の少なくとも一部を透過するので、ユーザの視野において実像（外界の風景）に重畳させた画像をユーザに見せることができる。

30

【0022】

眼鏡5は、HD10をユーザの頭部に保持するための構成である。眼鏡5は、フレーム6において、左眼用レンズを支えるリム部の上面右端（ユーザから見て上面左端）に、支持部4を備える。支持部4は、HD10の筐体2を保持し、眼鏡5に取り付ける。支持部4は、筐体2の保持位置を上下方向および左右方向に調整することができる。HD10は、ユーザにハーフミラー3の位置を調整させ、ユーザの眼球の位置に合わせた位置に配置させることができる。

【0023】

眼鏡5は、フレーム6において、左右のリム部を接続するブリッジ部の上面に、カメラ21を備える。カメラ21は、ユーザの視野前方の外界の風景を撮影する。カメラ21の撮影方向および画角は、予め、ユーザの視野に合わせて設定されている。より詳細には、カメラ21の撮影方向および画角は、カメラ21による外界の風景の撮影範囲が、ハーフミラー3を透過する実像に重ねることのできる画像表示部14（図2参照）からの画像光の表示範囲とほぼ一致するように設定されている。なお、カメラ21の撮影範囲はこれに限定されず、画像光の表示範囲よりも大きくてもよく、小さくてもよい。

40

【0024】

眼鏡5は、フレーム6において、右耳に掛けるテンブル部に、マイク17およびイヤホン18を内蔵するヘッドセット16を備える。詳細は後述するが、HMD1は、HMD1の各種操作等の実行命令（コマンド）としてあらかじめ登録された登録コマンドに対応する操作ワードに対応する音声を、ヘッドセット16のマイク17を介して取得することに

50

よって、使用時における各種操作等を行うことが可能である。ヘッドセット16には骨伝導型のものを用いてもよい。なお、眼鏡5自体の形状は通常的眼鏡に類似のため、詳細な説明は省略する。

【0025】

CB50の構成について説明する。CB50は、略直方体状で縁部を丸めた箱型の筐体を有する。CB50は電源ランプ63を内蔵する電源スイッチ62を含む操作部61を備える。電源スイッチ62が操作されることで、HMD1の電源はオンまたはオフされる。HD10における各種設定や、使用時における各種操作等は、操作部61に対して入力される。

【0026】

図2を参照し、HMD1の電氣的構成について説明する。HD10は、HD10全体の制御を司るCPU11を備える。CPU11は、RAM12、プログラムROM13、画像表示部14、インターフェイス15、接続コントローラ19およびカメラ21と電氣的に接続する。CPU11は、インターフェイス15を介してヘッドセット16と電氣的に接続する。RAM12は、各種データを一時的に記憶する。プログラムROM13は、CPU11が実行する各種プログラム等を記憶する。各種プログラムは、HD10の出荷時にプログラムROM13に記憶される。なおCPU11は、後述するCB50のフラッシュROM54に記憶されたプログラムも実行することができる。

【0027】

画像表示部14は前述の通り、映像信号に基づいて画像を表示する。インターフェイス15はヘッドセット16に接続し、信号の入出力を制御する。接続コントローラ19は、ハーネス7を介してCB50の接続コントローラ58に接続し、有線通信を行う。ヘッドセット16はマイク17およびイヤホン18を備える。ヘッドセット16はマイク17が受け取るアナログの音をデジタルの音データに変換し、ハーネス7を介してCB50に送信する。ヘッドセット16はハーネス7を介してCB50から受信するデジタルの音データをアナログの音に変換し、イヤホン18から出力する。なお、HD10は、ヘッドセット16の代わりに、筐体2にマイクとスピーカを内蔵してもよい。カメラ21は前述の通り、映像を撮影する。

【0028】

CB50の電氣的構成について説明する。CB50は、CB50全体の制御を司るCPU51を備える。CPU51は、RAM52、プログラムROM53、フラッシュROM54、インターフェイス55、ビデオRAM56、画像処理部57、接続コントローラ58、およびメディアインターフェイス59と電氣的に接続する。

【0029】

RAM52は、各種フラグ、各種データ等を一時的に記憶する。後述する情報記憶プログラムの実行において、CPU51は、RAM52に、ヘッドセット16から受信する音データを保存する記憶領域を確保する。なお、第一実施形態では、音をデータ化したものを「音データ」といい、ヒトの発声する音声をデータ化したものを「音声データ」という。音声データは音データに含まれる。

【0030】

プログラムROM53は、後述する情報記憶プログラムを含み、CPU51が実行する各種プログラム、各種プログラムが使用するフラグやデータの初期値等を記憶する。情報記憶プログラムを含む各種プログラムは、HMD1の出荷時にプログラムROM53に記憶される。なおCPU51は、フラッシュROM54に記憶されたプログラムも実行可能である。

【0031】

フラッシュROM54は、CB50の出荷後にインストールされた各種プログラム、各種プログラムが使用するフラグやデータの設定値等を記憶する。インターフェイス55は電源スイッチ62および電源ランプ63を含む操作部61に接続し、ユーザによる操作の入力信号やランプの点灯信号等の入出力を行う。画像処理部57は、HD10の画像表示

10

20

30

40

50

部 1 4 に表示する画像を形成する処理を行う。ビデオ R A M 5 6 は、画像処理部 5 7 が形成した画像を画像表示部 1 4 に表示するための映像信号を生成するため、画像を仮想的に表示した仮想画面を記憶領域内に形成する。接続コントローラ 5 8 は、ハーネス 7 を介して H D 1 0 の接続コントローラ 1 9 に接続し、有線通信を行う。メディアインターフェイス 5 9 は、C P U 5 1 がメディアカード 6 5 の記憶領域にアクセスできるように接続するインターフェイスである。メディアカード 6 5 は、例えば、ハードディスク、メモリーカード、コンパクトフラッシュ（登録商標）カード（C F カード）、セキュアデジタル（登録商標）カード（S D カード）、ユニバーサルシリアルバスメモリ（U S B メモリ）等であり、メディアインターフェイス 5 9 に挿抜可能である。

【 0 0 3 2 】

図 3 を参照し、情報記憶プログラムについて説明する。H M D 1 は、上記したように、あらかじめプログラム R O M 5 3 に情報記憶プログラムを記憶した状態で出荷される。情報記憶プログラムは、H M D 1 の C B 5 0 の C P U 5 1 が実行するプログラムである。情報記憶プログラムを実行する C P U 5 1 は、概略、ユーザの発声した音声をテキストデータに変換し、議事録として、メディアカード 6 5 に保存する処理を行う。

【 0 0 3 3 】

ユーザが H M D 1 の C B 5 0 に設けられた電源スイッチ 6 2 を操作すると、C P U 5 1 は起動時における所定の動作をメインプログラム（図示略）の実行に従って行う。C P U 5 1 は、情報記憶プログラムを含む各種プログラムを実行する。

【 0 0 3 4 】

情報記憶プログラムにおいて、C P U 5 1 は、起動時に行う初期設定処理（図示略）を行う。C P U 5 1 は、情報記憶プログラムにおいて使用するフラグやデータの記憶領域を R A M 5 2 に確保する。C P U 5 1 は、確保した記憶領域に、プログラム R O M 5 3 に記憶されているフラグやデータの初期値と、フラッシュ R O M 5 4 に記憶されているフラグやデータの設定値を R A M 5 2 に書き込む。C P U 5 1 は R A M 5 2 に作業用の記憶領域（ワークエリア）と、音データの記憶領域を確保する。C P U 5 1 は、マイク 1 7 が收音してデジタル変換された音データをヘッドセット 1 6 から受信し、F I F O 処理で、R A M 5 2 に保存する処理を開始する。

【 0 0 3 5 】

C P U 5 1 は情報記憶プログラムの初回実行時にフラッシュ R O M 5 4 にコマンド辞書を構築する。「コマンド辞書」は、登録コマンドを記憶するデータベースである。「登録コマンド」は、H M D 1 の各種操作等を実行する制御信号を表す制御コードに対応付けられた実行命令（コマンド）であり、あらかじめ登録されたコマンドである。H M D 1 の各種操作等の例として、画像表示部 1 4 に表示する画像の切り換え（ファイルオープン、ページめくり等）、拡大、縮小、スクロール、回転、輝度変更などが挙げられる。登録コマンドは、例えば上記例にそれぞれ対応付けられたコマンドである。第一実施形態において、登録コマンドはテキスト形式のデータで示される単語として提供される。C P U 5 1 は、情報記憶プログラムの実行中に、情報記憶プログラムの動作をカスタマイズする各種設定を受け付ける設定画面（図示略）を表示することができる。C P U 5 1 は、例えば、ユーザが制御コードに対応付けた任意のコマンドを、コマンド辞書に登録することができる。

【 0 0 3 6 】

C P U 5 1 は上記の初期設定処理が終了すると、C P U 5 1 は、R A M 5 2 に保存された音データの分析を行う。C P U 5 1 は、音データに対し、ヒトの発声する音声の周波数帯でフィルタリングを行う公知のフィルタリング処理を行い、音データが表す音の周波数を検出する（S 1 1）。C P U 5 1 は、検出した周波数に、ヒトの発声する音声の周波数帯が含まれない場合、マイク 1 7 へのユーザの音声入力はないと判断し（S 1 1 : N O）、S 3 1 に処理を進める。C P U 5 1 は、ユーザが電源スイッチ 6 2 を操作していなければ（S 3 1 : N O）、S 1 1 に処理を戻し、R A M 5 2 に保存された新たな音データの分析を行う。

10

20

30

40

50

【 0 0 3 7 】

S 1 1における周波数検出の結果、音データが表す音の周波数に音声の周波数帯が含まれる場合、CPU 5 1は、マイク 1 7へのユーザの音声入力があると判断し(S 1 1 : Y E S)、S 1 3に処理を進める。CPU 5 1は、音データの記憶領域からユーザの発声する音声として認識した音声データを取得し、ワークエリアに保存する(S 1 3)。CPU 5 1は、RAM 5 2に新たに保存された音データに対して上記のフィルタリング処理を行い、音データが表す音の周波数を検出する(S 1 5)。CPU 5 1は、検出した周波数に、ヒトの音声の周波数帯が含まれる場合、ユーザが発声中であって無声状態(発声していない状態)ではないと判断し(S 1 5 : N O)、S 1 3に処理を戻す。CPU 5 1は、RAM 5 2に新たに保存された音声データをワークエリアに保存し、ユーザの音声データを蓄積する。CPU 5 1は、ユーザが発声している間、S 1 3、S 1 5の処理を繰り返して実行し、無声状態になるのを待つ。

10

【 0 0 3 8 】

音データが表す音の周波数にヒトの音声の周波数帯が含まれない場合、CPU 5 1は、ユーザが発声を終了して無声状態になったと判断し(S 1 5 : Y E S)、S 1 7に処理を進める。CPU 5 1は、S 1 3、S 1 5の処理を繰り返して行うたびに分割状態で取得した音声データをワークエリアに蓄積する。よって、ワークエリアには、2回の連続する無音状態の間にユーザが発声した音声に基づく音声データが構築されて保存されている。CPU 5 1は、公知の音声認識処理(図示略)を実行し、ワークエリアに保存されたユーザの音声データに対する音声認識を行う(S 1 7)。CPU 5 1は、音声認識処理を行って認識したユーザの音声データをテキストデータ(以下では便宜上、「ユーザテキストデータ」という。)に変換し(S 1 9)、ワークエリアに保存する。

20

【 0 0 3 9 】

CPU 5 1は、ユーザテキストデータの文字数を、コマンド辞書にテキスト形式のデータとして登録されている登録コマンドのうち、文字数が最多の登録コマンドの文字数と比較する(S 2 1)。ユーザテキストデータの文字数が登録コマンドの文字数よりも多い場合(S 2 1 : Y E S)、CPU 5 1は、ユーザは操作ワードを発声していないと判断し、S 2 7に処理を進める。画像処理部 5 7は、ユーザテキストデータに基づくテキスト画像を形成し、ビデオRAM 5 6に仮想的に表示して、映像信号を生成する。CPU 5 1は映像信号をHD 1 0に送信し、画像表示部 1 4に、ユーザの発声内容をテキスト化したテキスト画像を表示する(S 2 7)。

30

【 0 0 4 0 】

CPU 5 1は、ユーザテキストデータをメディアカード 6 5に転送し、議事録として保存する(S 2 9)。CPU 5 1は、処理をS 3 1に進め、上記同様であれば処理をS 1 1に戻し、RAM 5 2に保存された新たな音データの分析を行う。

【 0 0 4 1 】

CPU 5 1は、新たにユーザが発声した音声データを取得し、ユーザテキストデータに変換した結果、文字数が登録コマンドの文字数以下である場合(S 2 1 : N O)、S 2 3に処理を進める。CPU 5 1は、コマンド辞書を参照し、ユーザテキストデータを登録コマンドと比較して、一致する登録コマンドがあるか否か判断する(S 2 3)。ユーザテキストデータが登録コマンドでない場合(S 2 3 : N O)、CPU 5 1は処理をS 2 7に進め、上記同様、画像表示部 1 4にテキスト画像を表示し(S 2 7)、メディアカード 6 5にユーザテキストデータを保存する(S 2 9)。

40

【 0 0 4 2 】

S 2 3において、ユーザテキストデータが登録コマンドである場合(S 2 3 : Y E S)、CPU 5 1は、登録コマンドを実行する(S 2 5)。CPU 5 1は、ユーザが発声した操作ワードに対応し、HMD 1の操作を実行する制御信号を出力し、ユーザの所望する動作を行う。CPU 5 1は、登録コマンドとして判断したユーザテキストデータについてはメディアカード 6 5に保存せず、処理をS 3 1に進める。ユーザが電源スイッチ 6 2を操作した場合(S 3 1 : Y E S)、CPU 5 1は、情報記憶プログラムの実行を終了する。

50

【 0 0 4 3 】

次に、情報記憶装置の第二実施形態について、図 4、図 5 を参照して説明する。本発明の第二実施形態は、第一実施形態における H M D 1 の C P U 5 1 が実行する情報記憶プログラムが別形態の例である。よって、第二実施形態の情報記憶装置である H M D 1 の構成は、第一実施形態と同様であるので、説明を省略する。また、第二実施形態の情報記憶プログラムに沿って C P U 5 1 が行う処理についても、第一実施形態と同様の部分は、説明を省略または簡略化して行う。

【 0 0 4 4 】

第二実施形態の情報記憶プログラムを実行する C P U 5 1 は、概略、第一実施形態と同様に、ユーザテキストデータを議事録としてメディアカード 6 5 に保存する処理を行う。さらに、C P U 5 1 は、カメラ 2 1 を用いて撮影を行った場合に、撮影後にユーザが発声した発声内容に基づくユーザテキストデータを、撮影画像に対応付ける。C P U 5 1 は、撮影画像とユーザテキストデータとを見ることができる 1 つのファイルを「報告書」として作成する処理を行う。

10

【 0 0 4 5 】

なお、第二実施形態の情報記憶プログラムにおいて、C P U 5 1 は、記録フラグの状態に基づく判断処理を行う。「記録フラグ」は、カメラ 2 1 による撮影がなされた場合に成立し、オンとなるフラグである。記録フラグは、C P U 5 1 が上記の報告書を作成するにあたって、カメラ 2 1 で撮影した撮影画像に対応付けるユーザテキストデータが、撮影後に最初に生成されたテキストデータであるか否かの判断条件として使用される。

20

【 0 0 4 6 】

以下、第二実施形態の情報記憶プログラムの実行に伴い C P U 5 1 が行う処理について説明する。ユーザが H M D 1 の電源スイッチ 6 2 を操作して H M D 1 を起動すると、C P U 5 1 は、情報記憶プログラムを含む各種プログラムを実行する。情報記憶プログラムにおいて、C P U 5 1 は、第一実施形態と同様に、起動時の初期設定処理（図示略）を行った後、R A M 5 2 に保存された音データを分析し、音データが表す音の周波数を検出する（S 4 1）。音データにヒトの音声の周波数帯が含まれなければ（S 4 1 : N O）、C P U 5 1 は S 7 9 に処理を進める。ユーザが電源スイッチ 6 2 を操作していなければ（S 7 9 : N O）、C P U 5 1 は S 4 1 に処理を戻し、R A M 5 2 に保存された新たな音データの分析を行う。

30

【 0 0 4 7 】

S 4 1 において、音データにヒトの音声の周波数帯が含まれる場合（S 4 1 : Y E S）、C P U 5 1 は S 4 3 に処理を進めて音声データを取得し、ワークエリアに保存する（S 4 3）。C P U 5 1 は、R A M 5 2 に新たに保存された音データを分析し、新たな音データにヒトの音声の周波数帯が含まれなければ（S 4 5 : N O）、S 4 3 に処理を戻す。C P U 5 1 は、S 4 3、S 4 5 を繰り返して実行し、ユーザの音声データをワークエリアに蓄積する。

【 0 0 4 8 】

C P U 5 1 は、無声状態を検出したら（S 4 5 : Y E S）、公知の音声認識処理（図示略）を実行して、ユーザの音声データに対する音声認識を行う（S 4 7）。C P U 5 1 は、音声認識処理を行って認識したユーザの音声データをユーザテキストデータに変換し（S 4 9）、ワークエリアに保存する。

40

【 0 0 4 9 】

C P U 5 1 は、コマンド辞書を参照し、ユーザテキストデータを登録コマンドと比較して、一致する登録コマンドがあるか否か判断する（S 5 1）。ユーザテキストデータが登録コマンドでない場合（S 5 1 : N O）、C P U 5 1 は処理を S 7 1 に進め、記録フラグがオンであるか否か判断する（S 7 1）。カメラ 2 1 による撮影が行われておらず、記録フラグがオフである場合（S 7 1 : N O）、C P U 5 1 は処理を S 7 5 に進める。C P U 5 1 は、第一実施形態と同様、画像表示部 1 4 に、ユーザテキストデータに基づくテキスト画像を表示する（S 7 5）。

50

【 0 0 5 0 】

具体的に、例えば図 5 に示す第 1 期に、ユーザが、登録コマンドではない、「このネジが悪かったんだ」という言葉を発声した場合、CPU 5 1 は、ユーザテキストデータに変換後、登録コマンドではないと判断する (S 5 1 : N O) 。 CPU 5 1 は、画像表示部 1 4 に、「このネジが悪かったんだ」というユーザテキストデータに基づくテキスト画像を表示する (S 7 5) 。

【 0 0 5 1 】

図 4 に示すように、CPU 5 1 は、メディアカード 6 5 にユーザテキストデータを議事録として保存する (S 7 7) 。 CPU 5 1 は、処理を S 7 9 に進め、上記同様であれば処理を S 4 1 に戻し、RAM 5 2 に保存された新たな音データの分析を行う。

10

【 0 0 5 2 】

S 5 1 において、ユーザテキストデータが登録コマンドである場合 (S 5 1 : Y E S) 、 CPU 5 1 は、登録コマンドが、カメラ 2 1 による撮影を指示する撮影コマンドであるか否か判断する (S 5 3) 。撮影コマンドでなければ (S 5 3 : N O) 、 CPU 5 1 は、登録コマンドが、報告書作成作業の終了を指示する終了コマンドであるか否か判断する (S 6 1) 。終了コマンドでなければ (S 6 1 : N O) 、 CPU 5 1 は登録コマンドを実行する (S 6 3) 。 CPU 5 1 は、ユーザが発声した操作ワードに対応し、撮影および報告書作成作業終了を除く、HMD 1 の操作を実行する制御信号を出力し、ユーザの所望する動作を行う。 CPU 5 1 は、登録コマンドとして判断したユーザテキストデータについてはメディアカード 6 5 に保存せず、また、画像表示部 1 4 へのテキスト画像の表示も行わない。 CPU 5 1 は、処理を S 7 9 に進め、上記同様であれば処理を S 4 1 に戻し、RAM 5 2 に保存された新たな音データの分析を行う。

20

【 0 0 5 3 】

ユーザテキストデータが撮影コマンドであった場合 (S 5 1 : Y E S 、 S 5 3 : Y E S) 、 CPU 5 1 は、カメラ 2 1 に、撮影を指示する制御信号を送信する (S 5 5) 。カメラ 2 1 はユーザの視野前方の外界の風景を撮影し、ハース 7 を介し、撮影画像のデータ (以下、「撮影データ」という。) を CPU 5 1 に送信する。 CPU 5 1 は、受信した撮影データを RAM 5 2 に保存し、画像表示部 1 4 に撮影画像を表示する (S 5 7) 。

【 0 0 5 4 】

具体的に、例えば図 5 に示す第 2 期に、ユーザが、登録コマンドである、「撮影」という言葉を発声した場合、CPU 5 1 は、ユーザテキストデータに変換後、登録コマンドであり (S 5 1 : Y E S) 、且つ撮影コマンドであると判断する (S 5 3 : Y E S) 。 CPU 5 1 は、カメラ 2 1 が撮影した撮影画像 (例えば自動車のナンバープレートを撮影した画像) を画像表示部 1 4 に表示する (S 5 7) 。

30

【 0 0 5 5 】

図 4 に示すように、CPU 5 1 は、カメラ 2 1 による撮影がなされたことを報せる記録フラグをオンにして (S 5 9) 、処理を S 7 9 に進め、上記同様であれば処理を S 4 1 に戻し、RAM 5 2 に保存された新たな音データの分析を行う。

【 0 0 5 6 】

撮影後、ユーザが新たに発声し、発声内容に基づく今回のユーザテキストデータが登録コマンドでなかった場合 (S 5 1 : N O) 、 CPU 5 1 は、記録フラグがオンであるので S 7 3 に処理を進める (S 7 1 : Y E S) 。 CPU 5 1 は、RAM 5 2 に保存した撮影データと、今回のユーザテキストデータとの対応付けを行う (S 7 3) 。 CPU 5 1 は、画像表示部 1 4 に、今回のユーザテキストデータに基づくテキスト画像を表示する (S 7 5) 。

40

【 0 0 5 7 】

具体的に、例えば図 5 に示す第 3 期に、ユーザが、登録コマンドではない、「この 2 本のネジがゆるんでいた」という言葉を発声した場合、CPU 5 1 は、ユーザテキストデータに変換後、登録コマンドではないと判断する (S 5 1 : N O) 。記録フラグがオンであるので撮影画像に対応付けた後 (S 7 3) 、 CPU 5 1 は、画像表示部 1 4 に、「この 2

50

本のネジがゆるんでいた」というユーザテキストデータに基づくテキスト画像を表示する（S75）。

【0058】

図4に示すように、CPU51は、メディアカード65に、撮影画像と対応付けたユーザテキストデータを議事録として保存する（S77）。CPU51は、処理をS79に進め、上記同様であれば処理をS41に戻し、RAM52に保存された新たな音データの分析を行う。以降同様に、ユーザの発声内容に基づく新たなユーザテキストデータが登録コマンドでない場合、記録フラグはオンのままであるので、CPU51は、ユーザテキストデータを撮影データに対応付けて保存する。

【0059】

ユーザテキストデータが報告書作成作業の終了コマンドであった場合（S51：YES、S53：NO、S61：YES）、CPU51は、撮影画像と、撮影画像に対応付けたユーザテキストデータに基づくテキスト画像とを、画像表示部14に表示する（S65）。すなわち、CPU51は、ユーザの指示に基づき撮影した撮影画像と、その撮影画像に関連してユーザが発声した内容に基づくユーザテキストデータとを組み合わせ、報告書として作成し、報告書の画面を画像表示部14に表示する。

【0060】

具体的に、例えば図5に示す第4期に、ユーザが、登録コマンドである、「作業終了」という言葉が発声した場合、CPU51は、ユーザテキストデータに変換後、登録コマンドであり（S51：YES）、且つ終了コマンドであると判断する（S61：YES）。CPU51は、画像表示部14に、カメラ21が撮影した撮影画像と、第3期において撮影画像に対応付けられた「この2本のネジがゆるんでいた」というユーザテキストデータに基づくテキスト画像を、報告書画面として表示する（S65）。

【0061】

図4に示すように、CPU51は、メディアカード65に、撮影画像と、撮影画像に対応付けたユーザテキストデータとを組み合わせたファイルを、報告書として保存する（S67）。CPU51は、記録フラグをオフにし（S69）、処理をS79に進め、上記同様であれば処理をS41に戻し、RAM52に保存された新たな音データの分析を行う。そして、ユーザが電源スイッチ62を操作した場合（S79：YES）、CPU51は、情報記憶プログラムの実行を終了する。

【0062】

以上説明したように、CPU51は、HMD1の制御をユーザが発声する音声データに基づいて行いながらも、ユーザが制御のために発声した登録コマンドを含まず、ユーザが制御目的ではなく発声したユーザテキストデータのみを保存することができる。ユーザの音声に基づくユーザテキストデータが登録コマンドを含んだ状態で保存された場合、保存したユーザテキストデータを議事録として用いるには、不要な登録コマンドが含まれているため好ましくないが、本実施形態のHMD1であれば、登録コマンドを除いたユーザテキストデータを保存できるので、保存したユーザテキストデータを議事録として利用することができる。

【0063】

第一実施形態のS21において、CPU51は、ユーザテキストデータが示す文字の文字数が所定数よりも多いか否か、文字数を比較するという単純な判断処理を行うだけで、多い場合にはS23においてコマンド辞書を参照して登録コマンドであるか否か判断する処理を回避できるので、情報記憶プログラムの処理の単純化および高速化を図ることができる。所定数としては、例えば、複数種類の登録コマンドがそれぞれ示す文字の文字数のうち、最多の文字数を挙げることができる。この場合、ユーザテキストデータの文字数が登録コマンドの文字数よりも多ければ、ユーザテキストデータに登録コマンドが含まれないとみなすことができるので、コマンド辞書を参照してユーザテキストデータと登録コマンドとを比較照合しなくとも済む。そして、所定数が、文字数が最多の登録コマンドの文字数以上であれば、ユーザテキストデータの文字数が所定数よりも多い場合に、ユーザテ

10

20

30

40

50

キストデータに登録コマンドが含まれることはない。

【0064】

画像表示部14にユーザテキストデータが示すテキスト画像を表示することで、CPU51は、ユーザの発声内容を正しく認識できたか、ユーザに目視で確認させることができる。一方で、CPU51は、登録コマンドが示すテキスト画像を画像表示部14に表示しないことで、ユーザに煩わしさを感じさせない。また第二実施形態において、撮影画像と、撮影後に生成したユーザテキストデータとを対応付けて記憶することで、報告書としてユーザテキストデータが示す文字を読む者は、撮影画像を参照することができ、内容を把握しやすい。

【0065】

なお、本発明は上記実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲内において、種々変更を加え得ることが可能である。CPU51は、ユーザテキストデータをメディアカード65に保存したが、例えばネットワークを介して接続可能なストレージデバイス等に保存してもよい。この場合、CB50は、有線または無線通信部を備え、ネットワークと接続できるようにするとよい。

【0066】

CB50は、HD10を制御する専用コントローラの例として説明したが、HD10に接続する機器は専用コントローラに限るものではない。例えば、スマートフォン、タブレット型端末、ノート型パーソナルコンピュータなど汎用のコンピュータ機器に、情報記憶プログラムに相当する専用のプログラムをインストールしてHD10と接続し、本発明に係る情報記憶装置を構成してもよい。

【0067】

本実施形態では、情報記憶プログラムを含む各種プログラムは、HMD1の出荷時にプログラムROM53に記憶されるものとしたが、必ずしもプログラムROM53に記憶された状態で出荷されなくともよい。例えば、CPU51が、情報記憶プログラムを含む各種プログラムを記憶したメディアカード65から、各種プログラムを読み出してフラッシュROM54にインストールしてもよい。または、CB50が有線通信あるいは無線通信を行う通信部を備え、CPU51は、通信部を介して、ダウンロードサーバのHDなど外部のストレージデバイスに保存された情報記憶プログラムを含む各種プログラムをダウンロードして、フラッシュROM54にインストールしてもよい。なお、情報記憶プログラムも含めて各種プログラムをフラッシュROM54にインストールしてHMD1の駆動を行えば、プログラムROM53はなくともよい。

【0068】

登録コマンドは、テキスト形式のデータとしたが、音声データであってもよい。この場合、ユーザが発声した音声データを、ユーザテキストデータに変換する前に、登録コマンドと比較する処理を行えばよい。また、第一実施形態において、S21でユーザテキストデータの文字数を登録コマンドの文字数と比較する処理は、なくてもよい。また、ユーザテキストデータに基づくテキスト画像を画像表示部14に表示する処理(S27、S75等)はなくともよく、CPU51は、ユーザテキストデータをメディアカード65に保存する処理だけを行ってもよい。また、ヘッドセット16は、公知のノイズキャンセリング機能を搭載したものであってもよい。

【0069】

また、第二実施形態において、撮影後、報告書作成作業の終了コマンドを検出するまで、CPU51は、ユーザテキストデータを撮影画像と対応付けたが、撮影後に最初に取得したユーザテキストデータだけを撮影画像と対応付け、そのときに記録フラグをオフにしてもよい。この場合、CPU51は、ユーザに終了コマンドを発声させなくとも報告書作成作業を終了することができるので、ユーザに手間をかけさせない。

【0070】

なお、本実施の形態においては、HMD1が本発明の「情報記憶装置」に相当する。メディアカード65が、「記憶部」に相当する。ユーザテキストデータが、「文字データ」

10

20

30

40

50

に相当する。登録コマンドが、「指示データ」に相当する。コマンド辞書が、「対応データ」に相当する。CPU 51が、「制御手段」に相当する。S 17またはS 47で、音声データを解析し、S 19またはS 49で、テキストデータに変換するCPU 51が、「生成部」に相当する。S 23またはS 51で、ユーザテキストデータが登録コマンドであるか否か判断するCPU 51が、「第一判断部」に相当する。S 23またはS 51でユーザテキストデータが登録コマンドである場合にS 25またはS 63でコマンドを実行するCPU 51が、「信号出力部」に相当する。S 23またはS 51でユーザテキストデータが登録コマンドである場合にユーザテキストデータの保存は行わず、S 23またはS 51でユーザテキストデータが登録コマンドでない場合にS 29またはS 77でユーザテキストデータの保存を行うCPU 51が、「記憶制御部」に相当する。

10

【0071】

S 15で、無声状態であるか否か判断するCPU 51が、「検出部」に相当する。S 21で、ユーザテキストデータの文字数を登録コマンドの文字数と比較するCPU 51が、「第二判断部」に相当する。画像表示部14が、「表示手段」に相当する。S 27またはS 75で、画像表示部14にユーザテキストデータに基づくテキスト画像を表示するCPU 51が、「表示制御部」に相当する。カメラ21が、「撮影手段」に相当する。

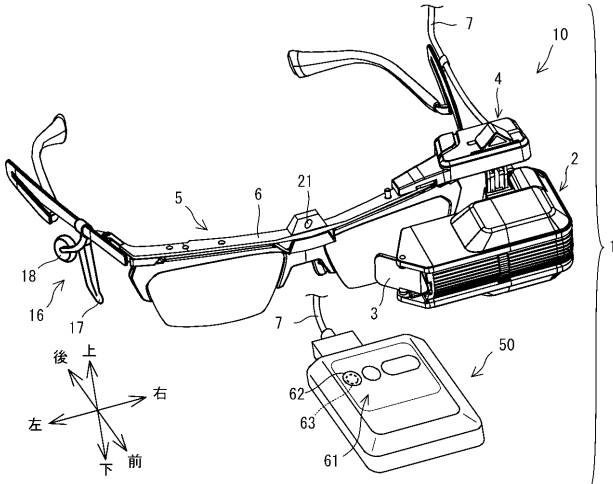
【符号の説明】

【0072】

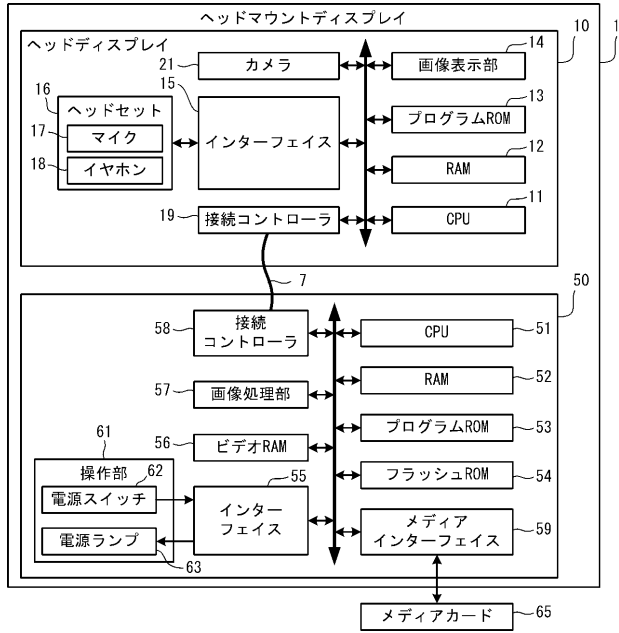
- 1 ヘッドマウントディスプレイ (HMD)
- 10 ヘッドディスプレイ (HD)
- 14 画像表示部
- 17 マイク
- 21 カメラ
- 50 コントロールボックス (CB)
- 51 CPU
- 52 RAM
- 59 メディアインターフェイス
- 65 メディアカード

20

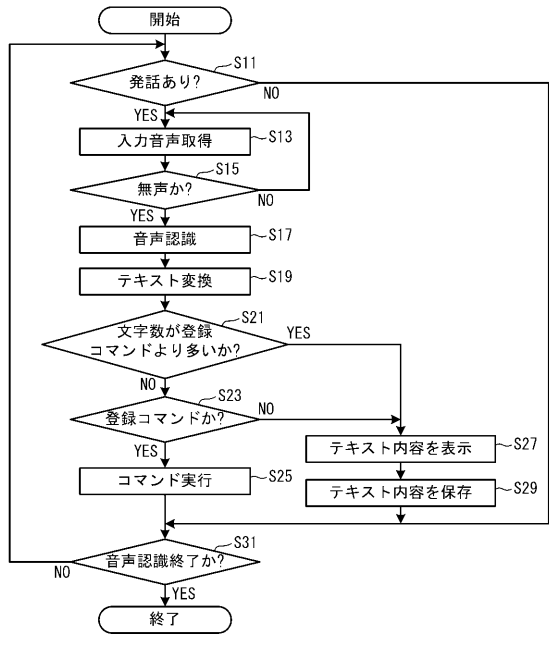
【図1】



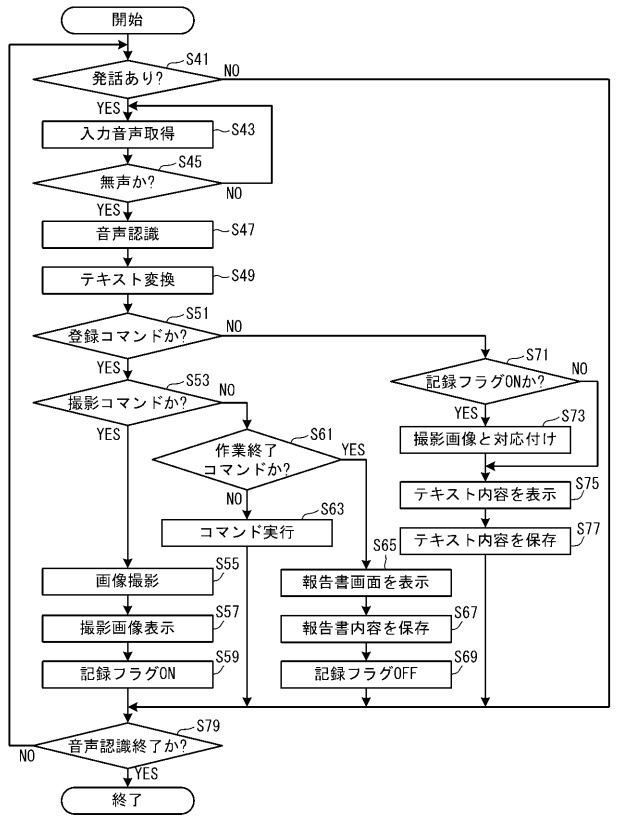
【図2】



【図3】



【図4】



【 図 5 】

