

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2024-91702

(P2024-91702A)

(43)公開日 令和6年7月5日(2024.7.5)

(51)国際特許分類

F I

G 0 1 C 21/26 (2006.01)

G 0 1 C 21/26 C

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 3 0 0 F

審査請求 有 請求項の数 1 O L (全17頁)

(21)出願番号 特願2024-62673(P2024-62673)
 (22)出願日 令和6年4月9日(2024.4.9)
 (62)分割の表示 特願2020-62289(P2020-62289)の
 分割
 原出願日 令和2年3月31日(2020.3.31)

(71)出願人 000005016
 パイオニア株式会社
 東京都文京区本駒込二丁目2番8号
 (74)代理人 100112656
 弁理士 宮田 英毅
 (72)発明者 大石 智也
 埼玉県川越市山田2-5番地1 パイオニア株式会社 川越事業所内
 (72)発明者 佐藤 祥子
 埼玉県川越市山田2-5番地1 パイオニア株式会社 川越事業所内

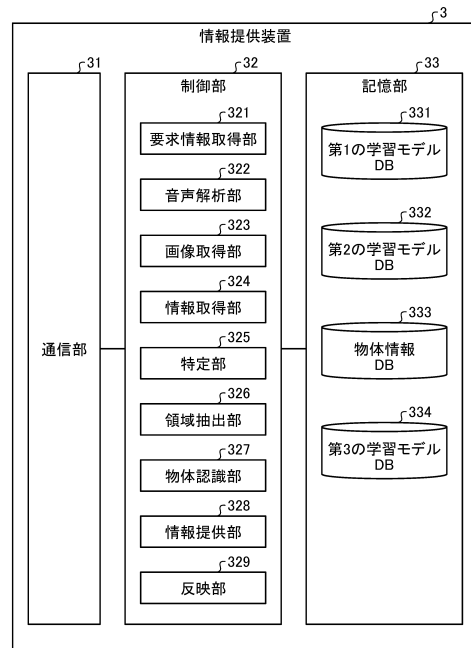
(54)【発明の名称】 情報提供装置

(57)【要約】

【課題】利便性を向上させること。

【解決手段】情報提供装置3は、移動体の周囲を撮影した撮影画像を取得する画像取得部323と、乗員から所定の要求があった場合に、移動体の位置情報を取得する情報取得部324と、位置情報を用いて、統計的に視線が集中する物体を特定する特定部325と、特定部325によって視線が集中する物体を特定できない場合には、撮影画像に基づいて、撮影画像内における視線が集中する注目領域を抽出する領域抽出部326と、撮影画像内における前記注目領域に含まれる物体を認識する物体認識部327と、特定部325によって特定された物体に関する物体情報を提供する情報提供部328とを備える。

【選択図】図3



【特許請求の範囲】**【請求項 1】**

移動体の周囲を撮影した撮影画像を取得する画像取得部と、
乗員から所定の要求があった場合に、前記移動体の位置情報を取得する情報取得部と、
前記位置情報を用いて、統計的に視線が集中する物体を特定する特定部と、
前記特定部によって前記視線が集中する物体を特定できない場合には、前記撮影画像に
基づいて、前記撮影画像内における視線が集中する注目領域を抽出する領域抽出部と、
前記撮影画像内における前記注目領域に含まれる物体を認識する物体認識部と、
前記特定部によって特定された物体に関する物体情報、または、前記注目領域に含まれ
る物体に関する物体情報を提供する情報提供部とを備える
ことを特徴とする情報提供装置。

10

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、情報提供装置、情報提供方法、情報提供プログラム及び記憶媒体に関する。

【背景技術】**【0002】**

従来、車両の周囲に存在する対象物を特定し、当該対象物に関する名称等の情報を音声
にて読み上げる対象物特定装置が知られている（例えば、特許文献1参照）。特許文献1
に記載の対象物特定装置では、車両の乗員が手や指で指し示している指示方向に存在する
地図上の施設等を対象物として特定している。

20

【先行技術文献】**【特許文献】****【0003】**

【特許文献1】特開2007-80060号公報

【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、特許文献1に記載の技術では、対象物に関する情報を得ることを望む車
両の乗員に対して、当該対象物を手や指で指し示すという作業を行わせる必要があり、利
便性を向上させることができない、という問題が一例として挙げられる。

30

【0005】

本発明は、上記に鑑みてなされたものであって、例えば利便性を向上させることができ
る情報提供装置、情報提供方法、情報提供プログラム及び記憶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】**

請求項1に記載の情報提供装置は、移動体の周囲を撮影した撮影画像を取得する画像取
得部と、乗員から所定の要求があった場合に、前記移動体の位置情報を取得する情報取得
部と、前記位置情報を用いて、統計的に視線が集中する物体を特定する特定部と、前記特
定部によって前記視線が集中する物体を特定できない場合には、前記撮影画像に基づいて
、前記撮影画像内における視線が集中する注目領域を抽出する領域抽出部と、前記撮影画
像内における前記注目領域に含まれる物体を認識する物体認識部と、前記特定部によつて
特定された物体に関する物体情報、または、前記注目領域に含まれる物体に関する物体情
報を提供する情報提供部とを備えることを特徴とする。

40

【図面の簡単な説明】**【0007】**

【図1】図1は、実施の形態1に係る情報提供システムの構成を示すブロック図である。

【図2】図2は、車載端末の構成を示すブロック図である。

【図3】図3は、情報提供装置の構成を示すブロック図である。

50

【図 4】図 4 は、情報提供方法を説明する図である。

【図 5】図 5 は、情報提供方法を説明する図である。

【図 6】図 6 は、情報提供方法を示すフローチャートである。

【図 7】図 7 は、情報提供方法を説明する図である。

【図 8】図 8 は、実施の形態 2 に係る情報提供装置の構成を示すブロック図である。

【図 9】図 9 は、情報提供方法を説明する図である。

【図 10】図 10 は、情報提供方法を示すフローチャートである。

【図 11】図 11 は、実施の形態 3 に係る情報提供装置の構成を示すブロック図である。

【図 12】図 12 は、情報提供方法を示すフローチャートである。

【図 13】図 13 は、情報提供装置が記憶するテーブルの一例を示す図である。

10

【図 14】図 14 は、情報提供装置が記憶するテーブルの一例を示す図である。

【発明を実施するための形態】

【0008】

以下に、図面を参照しつつ、本発明を実施するための形態（以下、実施の形態）について説明する。なお、以下に説明する実施の形態によって本発明が限定されるものではない。さらに、図面の記載において、同一の部分には同一の符号を付している。

【0009】

（実施の形態 1）

〔情報提供システムの概略構成〕

図 1 は、実施の形態 1 に係る情報提供システム 1 の構成を示すブロック図である。情報提供システム 1 は、移動体である車両 V E（図 1）の乗員 P A（図 7 参照）に対して、当該車両 V E の周囲に存在する建物等の物体に関する物体情報（例えば当該物体の名称等）を提供するシステムである。この情報提供システム 1 は、図 1 に示すように、車載端末 2 と、情報提供装置 3 とを備える。そして、これら車載端末 2 及び情報提供装置 3 は、無線通信網であるネットワーク N E（図 1）を介して、通信を行う。

20

【0010】

なお、情報提供装置 3 と通信を行う車載端末 2 としては、図 1 では 1 台である場合を示しているが、複数の車両にそれぞれ搭載された複数台であっても構わない。また、1 台の車両に乗車している複数の乗員に対してそれぞれ物体情報を提供するために、1 台の車両に車載端末 2 が複数台、搭載されていても構わない。

30

【0011】

〔車載端末の構成〕

図 2 は、車載端末 2 の構成を示すブロック図である。車載端末 2 は、例えば、車両 V E に設置される据え置き型のナビゲーション装置またはドライブレコーダーである。なお、車載端末 2 としては、ナビゲーション装置またはドライブレコーダーに限らず、車両 V E の乗員 P A が利用するスマートフォン等の携帯型端末を採用しても構わない。この車載端末 2 は、図 2 に示すように、音声入力部 2 1 と、音声出力部 2 2 と、撮像部 2 3 と、入力部 2 4 と、端末本体 2 5 と、センサ部 2 6 と、表示部 2 7 とを備える。

【0012】

音声入力部 2 1 は、音声を入力して電気信号に変換するマイクロフォン 2 1 1（図 7 参照）を含み、当該電気信号に A / D（Analog / Digital）変換等を行うことにより音声情報を生成する。本実施の形態 1 において、音声入力部 2 1 にて生成された音声情報は、デジタル信号である。そして、音声入力部 2 1 は、当該音声情報を端末本体 2 5 に出力する。

40

【0013】

音声出力部 2 2 は、スピーカ 2 2 1（図 7 参照）を含み、端末本体 2 5 から入力したデジタルの音声信号を D / A（Digital / Analog）変換によってアナログの音声信号に変換し、当該スピーカ 2 2 1 から当該アナログの音声信号に応じた音声を出力する。

【0014】

撮像部 2 3 は、端末本体 2 5 による制御の下、車両 V E の周囲を撮影して撮影画像を生

50

成する。そして、撮像部 2 3 は、生成した撮影画像を端末本体 2 5 に出力する。

【 0 0 1 5 】

入力部 2 4 は、タッチパネル、キーボードやマウス等の入力デバイスを含み、乗員 P A の操作に応じて各種データの入力を受け付ける。そして、入力部 2 4 は、入力を受け付けた各種データを端末本体 2 5 に出力する。表示部 2 7 は、液晶または有機 E L (Electro Luminescence) 等を用いた表示ディスプレイで構成され、端末本体 2 5 による制御の下、各種の画像を表示する。

【 0 0 1 6 】

センサ部 2 6 は、GPS (Global Positioning System) センサ、ジャイロセンサ、加速度センサや方位センサ等のセンサ装置を含み、端末本体 2 5 における処理に用いられる情報をセンシングする機能を有する。GPS センサは、GPS 衛星からの GPS 信号を受信することで、対象の緯度、経度及び高度を示す情報を測定する。GPS センサ 1 0 9 により取得される情報は、以下では、「位置情報」とも称される。

【 0 0 1 7 】

端末本体 2 5 は、図 2 に示すように、通信部 2 5 1 と、制御部 2 5 2 と、記憶部 2 5 3 とを備える。通信部 2 5 1 は、制御部 2 5 2 による制御の下、ネットワーク N E を介して情報提供装置 3 との間で情報の送受信を行う。

【 0 0 1 8 】

制御部 2 5 2 は、CPU (Central Processing Unit) や MPU (Micro Processing Unit) 等のコントローラによって、記憶部 2 5 3 に記憶された各種のプログラムが実行されることにより実現され、車載端末 2 全体の動作を制御する。なお、制御部 2 5 2 は、CPU や MPU に限らず、ASIC (Application Specific Integrated Circuit) や FPGA (Field Programmable Gate Array) 等の集積回路によって構成されても構わない。

【 0 0 1 9 】

記憶部 2 5 3 は、制御部 2 5 2 が実行する各種のプログラムや、当該制御部 2 5 2 が処理を行うときに必要なデータ等を記憶する。

【 0 0 2 0 】

〔 情報提供装置の構成 〕

図 3 は、情報提供装置 3 の構成を示すブロック図である。情報提供装置 3 は、例えば、サーバ装置である。この情報提供装置 3 は、図 3 に示すように、通信部 3 1 と、制御部 3 2 と、記憶部 3 3 とを備える。

【 0 0 2 1 】

通信部 3 1 は、制御部 3 2 による制御の下、ネットワーク N E を介して車載端末 2 (通信部 2 5 1) との間で情報の送受信を行う。

【 0 0 2 2 】

制御部 3 2 は、CPU や MPU 等のコントローラによって、記憶部 3 3 に記憶された各種のプログラム (本実施の形態に係る情報提供プログラムを含む) が実行されることにより実現され、情報提供装置 3 全体の動作を制御する。なお、制御部 3 2 は、CPU や MPU に限らず、ASIC や FPGA 等の集積回路によって構成されても構わない。この制御部 3 2 は、図 3 に示すように、要求情報取得部 3 2 1 と、音声解析部 3 2 2 と、画像取得部 3 2 3 と、情報取得部 3 2 4 と、特定部 3 2 5 と、領域抽出部 3 2 6 と、物体認識部 3 2 7 と、情報提供部 3 2 8 と、反映部 3 2 9 とを備える。

【 0 0 2 3 】

要求情報取得部 3 2 1 は、車両 V E の乗員 P A からの物体情報の提供を要求する要求情報を取得する。本実施の形態 1 では、当該要求情報は、車両 V E の乗員 P A が発した言葉 (音声) を音声入力部 2 1 が取り込み、当該音声に基づいて当該音声入力部 2 1 にて生成された音声情報である。すなわち、要求情報取得部 3 2 1 は、通信部 3 1 を介して車載端末 2 から当該要求情報 (音声情報) を取得する。音声解析部 3 2 2 は、要求情報取得部 3

10

20

30

40

50

2 1 が取得した要求情報（音声情報）を解析する。

【0024】

画像取得部 3 2 3 は、撮像部 2 3 にて生成された撮影画像を、通信部 3 1 を介して車載端末 2 から取得する。情報取得部 3 2 4 は、乗員 P A から所定の要求があった場合に、移動体の位置情報を取得する。本実施形態では、情報取得部 3 2 4 は、音声解析部 3 2 2 が要求情報（音声情報）を解析した結果、当該要求情報（音声情報）に特定のキーワードが含まれている場合に、移動体の位置情報を車載端末 2 から取得する。ここで、当該特定のキーワードとしては、車両 V E の乗員 P A が物体情報の提供を要求する言葉であり、「なに」、「なんですか」、「なんだろう」、「おしえて」等の言葉を例示することができる。また、情報取得部 3 2 4 は、移動体の位置情報とともに、乗員の所定の要求内容から、乗員 P A が着目している物体を特定するための特定のキーワードをさらに取得するようにしてもよい。例えば、情報取得部 3 2 4 は、乗員 P A が発した言葉に、「建物」、「お寺」、「お店」などの特定のキーワードが含まれている場合には、該特定のキーワードを取得する。

10

【0025】

特定部 3 2 5 は、位置情報を用いて、統計的に視線が集中する物体を特定する。本実施の形態 1 では、特定部 3 2 5 は、位置情報を入力データとして、統計的に視線が集中する物体を特定する第 3 の学習モデルを用いて、視線が集中する物体を特定する。つまり、特定部 3 2 5 は、第 3 の学習モデルに位置情報を入力し、第 3 の学習モデルからの出力結果として、視線が集中する物体の情報を取得する。なお、ここで統計的に視線が集中する物体とは、乗員 P A の視線が集中する物体と予測される物体、言い換えると、特定部 3 2 5 が乗員 P A の視線が集中する物体であろうと判定した物体のことを意味するものである。また、特定部 3 2 5 は、位置情報およびキーワードを用いて、位置情報およびキーワードに応じて統計的に視線が集中する物体を特定するようにしてもよい。例えば、特定部 3 2 5 は、乗員 P A が「あの建物は何?」、「あそこのお寺について教えて」、「あのお店は何?」などと言葉を発した場合に、言葉が発した際の車両 V E の位置情報の他に、「建物」、「お寺」、「お店」などの乗員 P A が着目している物体を特定するためのキーワードも用いて、統計的に視線が集中する物体を特定するようにしてもよい。この場合には、例えば、特定部 3 2 5 は、位置情報およびキーワードを入力データとして、統計的に視線が集中する物体を特定する第 3 の学習モデルを用いて、視線が集中する物体を特定する。つまり、特定部 3 2 5 は、位置情報とともに、乗員 P A が着目している物体を特定するためのキーワードを第 3 の学習モデルに入力し、第 3 の学習モデルからの出力結果として、視線が集中する物体の情報を取得するようにしてもよい。

20

30

【0026】

第 3 の学習モデルは、例えば、アイトラッカを用いて被験者の視線が集中する領域を判別し、当該領域が予めラベリングされた位置情報とキーワードを教師データとし、当該教師データ像を利用して当該領域を機械学習（例えば深層学習等）することにより得られたモデルである。なお、本実施形態では、第 3 の学習モデルは、後述する反映部 3 2 9 により更新される。

【0027】

領域抽出部 3 2 6 は、特定部 3 2 5 によって視線が集中する物体を特定できない場合には、画像取得部 3 2 3 にて取得された撮影画像に基づいて、当該撮影画像内における視線が集中する（視線が集中し易い）注目領域を抽出（予測）する。本実施の形態 1 では、領域抽出部 3 2 6 は、所謂、視覚的顕著性技術を利用して撮影画像内における注目領域を抽出する。より具体的に、領域抽出部 3 2 6 は、以下に示す第 1 の学習モデルを用いた画像認識（A I（Artificial Intelligence）を用いた画像認識）により、撮影画像内における注目領域を抽出する。

40

【0028】

なお、ここで特定部 3 2 5 によって視線が集中する物体を特定できない場合とは、特定部 3 2 5 が実際に特定できなかった場合のみならず、例えば、第 3 の学習モデルの学習不

50

足により十分な精度で視線が集中する物体を特定できない場合を含むものとする。当該第 1 の学習モデルは、アイトラッカを用いて被験者の視線が集中する領域を判別し、当該領域が予めラベリングされた画像を教師画像とし、当該教師画像を利用して当該領域を機械学習（例えば深層学習等）することにより得られたモデルである。

【 0 0 2 9 】

物体認識部 3 2 7 は、撮影画像内において、領域抽出部 3 2 6 にて抽出された注目領域に含まれる物体を認識する。本実施の形態 1 では、物体認識部 3 2 7 は、以下に示す第 2 の学習モデルを用いた画像認識（AI を用いた画像認識）により、撮影画像内における注目領域に含まれる物体を認識する。

【 0 0 3 0 】

当該第 2 の学習モデルは、動物、山、川、湖、及び施設等の各種の物体が撮影された撮影画像を教師画像とし、当該教師画像に基づいて当該物体の特徴を機械学習（例えば深層学習等）することにより得られたモデルである。

【 0 0 3 1 】

情報提供部 3 2 8 は、特定部 3 2 5 によって特定された物体に関する物体情報、または、物体認識部 3 2 7 にて認識された物体に関する物体情報を提供する。より具体的に、情報提供部 3 2 8 は、記憶部 3 3 における物体情報 DB（Data Base：データベース）3 3 3 から特定部 3 2 5 にて特定された物体に対応する物体情報、または、物体認識部 3 2 7 にて認識された物体に対応する物体情報を読み出す。そして、情報提供部 3 2 8 は、通信部 3 1 を介して車載端末 2 に当該物体情報を送信する。

【 0 0 3 2 】

反映部 3 2 9 は、情報提供部 3 2 8 によって物体情報を提供した結果を学習モデルに反映する。例えば、反映部 3 2 9 は、情報提供部 3 2 8 によって物体情報を提供した結果として、情報提供部 3 2 8 によって提供された物体情報がユーザの希望に沿う情報であったか否かを判定する。この結果、反映部 3 2 9 は、ユーザの希望に沿う情報であった場合には、情報提供部 3 2 8 によって物体情報を提供した結果を正解データとして学習モデルに反映し、ユーザの希望に沿う情報でなかった場合には、情報提供部 3 2 8 によって物体情報を提供した結果を不正解データとして学習モデルに反映する。なお、ユーザの希望に沿う情報であったか否かを判定する手法については、ユーザの手動による入力を受け付けることで判定してもよいし、物体情報を提供した後のユーザの行動等から自動で判定してもよい。なお、反映部 3 2 9 は、位置情報とともに、乗員 P A が着目している物体を特定するためのキーワードを第 3 の学習モデルに入力して物体情報を提供した場合には、キーワードについても学習モデルに反映する。

【 0 0 3 3 】

記憶部 3 3 は、制御部 3 2 が実行する各種のプログラム（本実施の形態に係る情報提供プログラム）の他、制御部 3 2 が処理を行うときに必要なデータ等を記憶する。この記憶部 3 3 は、図 3 に示すように、第 1 の学習モデル DB 3 3 1 と、第 2 の学習モデル DB 3 3 2 と、物体情報 DB 3 3 3 と、第 3 の学習モデル DB 3 3 4 を備える。第 1 の学習モデル DB 3 3 1 は、上述した第 1 の学習モデルを記憶する。第 2 の学習モデル DB 3 3 2 は、上述した第 2 の学習モデルを記憶する。第 3 の学習モデル DB 3 3 4 は、上述した第 3 の学習モデルを記憶する。

【 0 0 3 4 】

物体情報 DB 3 3 3 は、上述した物体情報を記憶する。ここで、物体情報 DB 3 3 3 には、各種の物体に関連付けられた複数の物体情報が記憶されている。当該物体情報としては、物体の名称等の当該物体を説明する情報であって、文字データ、音声データ、あるいは、画像データによって構成されている。

【 0 0 3 5 】

ここで、図 4 および図 5 を用いて、情報提供装置 3 による情報提供方法の一例を説明する。図 4 および図 5 は、情報提供方法を説明する図である。図 4 を用いて、位置情報を考慮した第 3 の学習モデルを用いて統計的に視線が集中する物体を特定できた場合の情報提

10

20

30

40

50

供方法について説明する。

【 0 0 3 6 】

図 4 の例では車両 V E の乗員 P A が建物に注目して「あれは何？」と発話した場合を例に説明する。図 4 に例示するように、車両 V E の乗員 P A から物体情報の提供を要求する言葉として「あれは何？」という言葉が発話された場合には、車両 V E の車載端末 2 は、位置情報を情報提供装置 3 に通知する。そして、情報提供装置 3 は、位置情報を考慮して学習された第 3 の学習モデルを用いて、車両 V E の現在位置で統計的に乗員の視線が集中するであろうと予測される物体を特定する。そして、情報提供装置 3 は、特定した物体に対応する物体情報（建物名称、詳細説明、画像など）を、通信部 3 1 を介して車載端末 2 に送信する。

10

【 0 0 3 7 】

このように、情報提供装置 3 では、位置情報を考慮した第 3 の学習モデルを用いて統計的に視線が集中する物体を特定して、該物体に対応する物体情報を乗員 P A に提供することができるので、視覚顕著性技術を用いる場合と比較して、処理負荷を軽減することが可能である。

【 0 0 3 8 】

ここで、情報提供装置 3 が第 3 の学習モデルを用いて統計的に視線が集中する物体を特定できなかった場合に、視覚顕著性技術と物体認識により視線が集中する物体を特定する情報提供方法について図 5 を用いて説明する。なお、図 5 の例では車両 V E の乗員 P A がライオンに注目して「あれ何？」と発話した場合を例に説明する。

20

【 0 0 3 9 】

情報提供装置 3 は、第 3 の学習モデルを用いて統計的に視線が集中する物体を特定できなかった場合には、第 1 の学習モデルを用いた画像認識により、撮影画像内における注目領域を抽出する。続いて、情報提供装置 3 は、第 2 の学習モデルを用いた画像認識により、撮影画像内における注目領域に含まれるライオンを認識する。そして、情報提供装置 3 は、認識したライオンに対応する物体情報を、通信部 3 1 を介して車載端末 2 に送信する。

【 0 0 4 0 】

このように、情報提供装置 3 は、第 3 の学習モデルを用いて統計的に視線が集中する物体を特定できなかった場合であっても、視覚顕著性技術を用いてリアルタイムに乗員 P A の要求に応じた物体情報を提供することが可能である。

30

【 0 0 4 1 】

〔 情報提供方法 〕

次に、情報提供装置 3（制御部 3 2）が実行する情報提供方法について説明する。図 6 は、情報提供方法を示すフローチャートである。図 7 は、情報提供方法を説明する図である。具体的に、図 7 は、撮像部 2 3 にて生成され、ステップ S 1 0 6 にて取得される撮影画像 I M を示す図である。ここで、図 7 では、車両 V E 内からフロントガラスを介して当該車両 V E の前方が撮影されるように当該車両 V E 内に撮像部 2 3 を設置した場合を例示している。また、図 7 では、撮影画像 I M 内に車両 V E の助手席に座った乗員 P A が被写体として含まれる場合を例示している。さらに、図 7 では、当該乗員 P A が「あれ何？」という言葉を出している場合を例示している。

40

【 0 0 4 2 】

なお、撮像部 2 3 の設置位置としては、上述した設置位置に限らない。例えば、車両 V E 内から当該車両 V E の左側方や右側方、あるいは、後方が撮影されるように当該車両 V E 内に撮像部 2 3 を設置してもよく、車両 V E の周囲が撮影されるように当該車両 V E 外に撮像部 2 3 を設置しても構わない。また、本実施の形態に係る車両の乗員としては、車両 V E の助手席に座った乗員に限らず、運転席や後部座席に座った乗員等を含むものである。また、撮像部 2 3 の数としては、一つに限らず、複数としても構わない。

【 0 0 4 3 】

先ず、要求情報取得部 3 2 1 は、通信部 3 1 を介して車載端末 2 から要求情報（音声情

50

報)を取得する(ステップS101)。ステップS101の後、音声解析部322は、当該ステップS101にて取得された要求情報(音声情報)を解析する(ステップS102)。ステップS102の後、音声解析部322は、当該ステップS102にて要求情報(音声情報)を解析した結果、当該要求情報(音声情報)に特定のキーワードが含まれているか否かを判定する(ステップS103)。ここで、当該特定のキーワードとしては、車両VEの乗員PAが物体情報の提供を要求する言葉であり、「なに」、「なんですか」、「なんだろう」、「おしえて」等の言葉を例示することができる。

【0044】

特定のキーワードが含まれていないと判定された場合(ステップS103:No)には、制御部32は、ステップS101に戻る。一方、特定のキーワードが含まれていると判定された場合(ステップS103:Yes)には、情報取得部324は、乗員PAから所定の要求があった場合に、車両VEの位置情報を取得する(ステップS104)。そして、特定部325は、位置情報を用いて統計的に視線が集中する物体を特定した結果、視線が集中する物体を特定できるかを判定する(ステップS105)。

10

【0045】

この結果、特定部325が、視線が集中する物体を特定できないと判定した場合には(ステップS105:No)、画像取得部323は、通信部31を介して車載端末2から撮像部23にて生成された撮影画像IMを取得する(ステップS106)。

【0046】

なお、図6及び図7では、車両VEの乗員PAが「あれ何?」との言葉を発したタイミング(ステップS103:Yes)で画像取得部323が通信部31を介して車載端末2から撮像部23にて生成された撮影画像IMを取得する構成としているが、これに限らない。例えば、情報提供装置3は、通信部31を介して車載端末2から撮像部23にて生成された撮影画像を順次、取得する。そして、画像取得部323は、当該順次、取得された撮影画像のうち、車両VEの乗員PAが「あれ何?」との言葉を発したタイミング(ステップS103:Yes)で取得された撮影画像をステップS104以降の処理に用いる撮影画像として取得する構成としても構わない。

20

【0047】

ステップS106の後、領域抽出部326は、第1の学習モデルDB331に記憶された第1の学習モデルを用いた画像認識により、撮影画像IM内における視線が集中する注目領域Ar1(図7参照)を抽出する(ステップS107)。

30

【0048】

ステップS107の後、物体認識部327は、第2の学習モデルDB332に記憶された第2の学習モデルを用いた画像認識により、撮影画像IM内において、当該ステップS5にて抽出された注目領域Ar1に含まれる物体OB1を認識する(ステップS108)。

【0049】

ステップS108の後、情報提供部328は、物体情報DB333から、特定部325によって特定された物体OB1、または、物体認識部327によって認識された物体OB1に対応する物体情報を取得し(ステップS109)、通信部31を介して車載端末2に当該物体情報を送信する(ステップS110)。そして、制御部252は、音声出力部22及び表示部27の少なくともいずれかの動作を制御し、情報提供装置3から送信された物体情報を音声、文字、及び画像の少なくともいずれかによって、車両VEの乗員PAに対して報知する。例えば、物体OB1が「ムーン・ルージュ」である場合には、物体情報として「あれはムーン・ルージュです。華やかなダンスショーを夜にやっています。」等の音声(又は文字)が車両VEの乗員PAに対して報知される。また、例えば、物体OB1が建物ではなく動物のバッファローである場合には、物体情報として「あれはバッファローです。バッファローは群れで行動します。」等の音声(又は文字)が車両VEの乗員PAに対して報知される。その後、反映部329は、物体情報を提供した結果を学習モデルに反映する(ステップS111)。

40

50

【 0 0 5 0 】

以上説明した本実施の形態 1 によれば、以下の効果を奏する。本実施の形態 1 に係る情報提供装置 3 は、車両 V E の周囲を撮影した撮影画像を取得し、乗員 P A から所定の要求があった場合に、車両 V E の位置情報を取得する。そして、情報提供装置 3 は、位置情報を用いて、統計的に視線が集中する物体を特定する。また、情報提供装置 3 は、視線が集中する物体を特定できない場合には、撮影画像に基づいて、撮影画像内における視線が集中する注目領域を抽出し、撮影画像内における注目領域に含まれる物体を認識する。その後、情報提供装置 3 は、特定した物体に関する物体情報、または、注目領域に含まれる物体に関する物体情報を提供する。

【 0 0 5 1 】

したがって、物体に関する物体情報を得ることを望む車両 V E の乗員 P A に対して、従来のように当該物体を手や指で指し示すという作業を行わせる必要がなく、利便性を向上させることができる。

【 0 0 5 2 】

また、情報提供装置 3 は、位置情報を考慮した第 3 の学習モデルを用いて統計的に視線が集中する物体を特定して、該物体に対応する物体情報を乗員 P A に提供することができるので、視覚顕著性技術を用いる場合と比較して、処理負荷を軽減することが可能である。

【 0 0 5 3 】

また、情報提供装置 3 は、統計的に視線が集中する物体を特定できなかった場合であっても、視覚的顕著性技術を利用して、撮影画像 I M 内における視線が集中する注目領域 A r 1 を抽出することができる。このため、車両 V E の乗員 P A が物体 O B 1 を手や指で指し示さなくても、当該物体 O B 1 を含む領域を注目領域 A r 1 として精度良く抽出することができる。

【 0 0 5 4 】

また、情報提供装置 3 は、車両 V E の乗員 P A からの物体情報の提供を要求する要求情報に応じて、当該物体情報を提供する。このため、当該要求情報によらず、常時、物体情報を提供する構成と比較して、情報提供装置 3 の処理負荷を軽減することができる。

【 0 0 5 5 】

(実施の形態 2)

次に、本実施の形態 2 について説明する。以下の説明では、上述した実施の形態 1 と同様の構成には同一符号を付し、その詳細な説明は省略または簡略化する。図 8 は、実施の形態 2 に係る情報提供装置 3 A の構成を示すブロック図である。本実施の形態 2 に係る情報提供装置 3 A では、情報取得部 3 2 4 および特定部 3 2 5 の機能が変更されている。以下では、説明の便宜上、本実施の形態 2 に係る情報取得部を情報取得部 3 2 4 A、本実施の形態 2 に係る特定部を特定部 3 2 5 A と記載する(図 8 参照)。さらに、情報提供装置 3 A では、記憶部 3 3 に第 4 の学習モデル D B 3 3 5 (図 8 参照)が追加されている。

【 0 0 5 6 】

第 4 の学習モデルは、例えば、アイトラッカを用いて被験者の視線が集中する領域を判別し、当該領域が予めラベリングされた位置情報とキーワードと属性情報を教師データとし、当該教師データを利用して当該領域を機械学習(例えば深層学習等)することにより得られたモデルである。なお、本実施形態では、第 4 の学習モデルは、反映部 3 2 9 により更新される。そして、第 4 の学習モデル D B 3 3 5 は、当該第 4 の学習モデルを記憶する。

【 0 0 5 7 】

情報取得部 3 2 4 A は、車両 V E の位置情報とともに、乗員 P A の属性情報をさらに取得する。例えば、情報取得部 3 2 4 A は、属性情報として、乗員 P A の年齢、乗員 P A の性別、乗員 P A の国籍、乗員 P A の外観、乗員 P A の嗜好(例えば、コーヒーが好き、旅行が好き等)、および、乗員 P A の言語のうち、いずれか一つまたは複数の情報を車載端末 2 から取得する。なお、車載端末 2 は、どのような手法で属性情報を取得してもよい。

10

20

30

40

50

例えば、車載端末 2 は、入力部 2 4 を介して乗員 P A の手動による属性情報の入力を受け付けてもよし、音声入力部 2 1 や撮像部 2 3 により入力された音声や画像を解析して乗員の属性情報を自動で取得してもよい。また、車載端末 2 は、属性情報を事前に記憶していてもよし、外部の D B から属性情報を取得するようにしてもよい。

【 0 0 5 8 】

特定部 3 2 5 A は、位置情報および属性情報を用いて、位置情報および属性情報に応じて統計的に視線が集中する物体を特定する。例えば、特定部 3 2 5 A は、位置情報および属性情報を入力データとして、統計的に視線が集中する物体を特定する学習モデルを用いて、視線が集中する物体を特定する。

【 0 0 5 9 】

ここで、図 9 を用いて、情報提供装置 3 A による情報提供方法の一例を説明する。図 9 は、情報提供方法を説明する図である。図 9 を用いて、位置情報および属性情報を考慮した第 4 の学習モデルを用いて統計的に視線が集中する物体を特定できた場合の情報提供方法について説明する。

【 0 0 6 0 】

図 9 の例では車両 V E の乗員 P A が建物に注目して「あれは何？」と発話した場合を例に説明する。図 9 に例示するように、車両 V E の乗員 P A から物体情報の提供を要求する言葉として「あれは何？」というが発話された場合には、車両 V E の車載端末 2 は、位置情報および属性情報を情報提供装置 3 A に通知する。そして、情報提供装置 3 A は、位置情報および属性情報を考慮して学習された第 4 の学習モデルを用いて、統計的に視線が集中するであろうと予測される物体を特定する。具体的な例を挙げて図 9 について説明すると、例えば、カフェについて「あれは何？」と発話した乗員 P A の属性情報が、性別「女性」、嗜好「コーヒーが好き」であるものとする。この場合に、情報提供装置 3 A は、言葉を発した際の車両 V E の位置情報とともに属性情報を考慮して学習された第 4 の学習モデルに対して、位置情報とともに、属性情報として、性別「女性」、嗜好「コーヒーが好き」という情報を入力することで、現在位置において同一または類似の属性情報を有する者が統計的に視線を集中させるであろう物体として「カフェ」を特定する。言い換えると、情報提供装置 3 A は、当該場所に置いて、コーヒー好きな女性が統計的に注目しやすい物体として「カフェ」を特定する。そして、図 9 の例では、情報提供装置 3 A は、カフェの情報に対応する物体情報として、カフェのメニューや予約情報を取得して車両 V E の車載端末 2 に送信する。

【 0 0 6 1 】

このように、情報提供装置 3 A では、位置情報および属性情報を考慮した第 4 の学習モデルを用いて統計的に視線が集中する物体を特定して、該物体に対応する物体情報を乗員 P A に提供することができるので、視覚顕著性技術を用いる場合と比較して、処理負荷を軽減することが可能である。なお、情報提供装置 3 A は、第 4 の学習モデルを用いて統計的に視線が集中する物体を特定できなかった場合には、実施の形態 1 と同様に、視覚顕著性技術と物体認識により視線が集中する物体を特定する。

【 0 0 6 2 】

次に、情報提供装置 3 A が実行する情報提供方法について説明する。図 1 0 は、情報提供方法を示すフローチャートである。本実施の形態 2 に係る情報提供方法では、図 1 0 に示すように、上述した実施の形態 1 で説明した情報提供方法（図 6 参照）に対して、ステップ S 2 0 5 の処理が追加されている。図 1 0 のステップ S 2 0 1 ~ 2 0 4 は、図 6 のステップ S 1 0 1 ~ S 1 0 4 と同様の処理であり、ステップ S 2 0 7 ~ S 2 1 1 は、図 6 のステップ S 1 0 6 ~ 1 1 0 で同様の処理である。このため、以下では、ステップ S 2 0 5 ~ S 2 0 6 のみを主に説明する。

【 0 0 6 3 】

ステップ S 2 0 6 は、ステップ S 2 0 5 の後に実行される。ステップ S 2 0 5 において、情報取得部 3 2 4 A は、乗員 P A の属性情報を取得する（ステップ S 2 0 5）。そして、特定部 3 2 5 A は、位置情報および属性情報を用いて、位置情報および属性情報に応じ

10

20

30

40

50

て統計的に視線が集中する物体を特定した結果、視線が集中する物体を特定できるかを判定する（ステップ S 2 0 6）。

【 0 0 6 4 】

この結果、特定部 3 2 5 A が、視線が集中する物体を特定できないと判定した場合には（ステップ S 2 0 6 : N o ）、ステップ S 2 0 7 の処理に進む。また、特定部 3 2 5 A は、視線が集中する物体を特定できると判定した場合には（ステップ S 2 0 6 : Y e s ）、ステップ S 2 1 0 の処理に進む。

【 0 0 6 5 】

なお、情報提供装置 1 0 A は、ステップ S 2 0 4 において、位置情報を取得した後、位置情報のみを用いて、位置情報に応じて統計的に視線が集中する物体を特定し、その結果、視線が集中する物体を特定できるかを判定し、視線が集中する物体を特定できないと判定した場合にのみ、ステップ S 2 0 5 の処理に進むようにしてもよい。

10

【 0 0 6 6 】

以上説明した本実施の形態 2 によれば、上述した実施の形態 1 と同様の効果の他、以下の効果を奏する。本実施の形態 2 に係る情報提供装置 3 A は、位置情報および属性情報に応じて統計的に視線が集中する物体を特定した結果、視線が集中する物体を特定できるかを判定するので、車両 V E の乗員 P A が物体情報を得ることを望む物体を含む領域を注目領域として精度良く特定することができる。したがって、車両 V E の乗員 P A に対して、適切な物体情報を提供することができる。

【 0 0 6 7 】

20

（実施の形態 3）

次に、本実施の形態 3 について説明する。以下の説明では、上述した実施の形態 1、実施の形態 2 と同様の構成には同一符号を付し、その詳細な説明は省略または簡略化する。

【 0 0 6 8 】

図 1 1 は、実施の形態 3 に係る情報提供装置 3 B の構成を示すブロック図である。また、本実施の形態 3 に係る情報提供装置 3 B では、上述した実施の形態 2 で説明した情報提供装置 3 A（図 8 参照）と比べて、画像取得部 3 2 3、領域抽出部 3 2 6 および物体認識部 3 2 7 を有していない点に変更されている。

【 0 0 6 9 】

情報取得部 3 2 4 A は、実施の形態 2 と同様に、車両 V E の位置情報とともに、乗員 P A の属性情報をさらに取得する。例えば、情報取得部 3 2 4 A は、属性情報として、乗員 P A の年齢、乗員 P A の性別、乗員 P A の国籍、乗員 P A の外観、および、乗員 P A の言語のうち、いずれか一つまたは複数の情報を車載端末 2 から取得する。

30

【 0 0 7 0 】

特定部 3 2 5 A は、位置情報および属性情報を用いて、位置情報および属性情報に応じて統計的に視線が集中する物体を特定する。例えば、特定部 3 2 5 A は、位置情報および属性情報を入力データとして、統計的に視線が集中する物体を特定する学習モデルを用いて、視線が集中する物体を特定する。

【 0 0 7 1 】

反映部 3 2 9 は、情報提供部 3 2 8 によって物体情報を提供した結果を学習モデルに反映する。例えば、反映部 3 2 9 は、情報提供部 3 2 8 によって物体情報を提供した結果として、情報提供部 3 2 8 によって提供された物体情報がユーザの希望に沿う情報であったか否かを判定する。この結果、反映部 3 2 9 は、ユーザの希望に沿う情報であった場合には、情報提供部 3 2 8 によって物体情報を提供した結果を正解データとして学習モデルに反映し、ユーザの希望に沿う情報でなかった場合には、情報提供部 3 2 8 によって物体情報を提供した結果を不正解データとして学習モデルに反映する。なお、ユーザの希望に沿う情報であったか否かを判定する手法については、ユーザの手動による入力を受け付けることで判定してもよいし、物体情報を提供した後のユーザの行動等から自動で判定してもよい。

40

【 0 0 7 2 】

50

次に、情報提供装置 3 B が実行する情報提供方法について説明する。図 1 2 は、情報提供方法を示すフローチャートである。まず、要求情報取得部 3 2 1 は、通信部 3 1 を介して車載端末 2 から要求情報（音声情報）を取得する（ステップ S 3 0 1）。ステップ S 3 0 1 の後、音声解析部 3 2 2 は、当該ステップ S 3 0 1 にて取得された要求情報（音声情報）を解析する（ステップ S 3 0 2）。ステップ S 3 0 2 の後、音声解析部 3 2 2 は、当該ステップ S 3 0 2 にて要求情報（音声情報）を解析した結果、当該要求情報（音声情報）に特定のキーワードが含まれているか否かを判定する（ステップ S 3 0 3）。ここで、当該特定のキーワードとしては、車両 V E の乗員 P A が物体情報の提供を要求する言葉であり、「なに」、「なんですか」、「なんだろう」、「おしえて」等の言葉を例示することができる。

10

【 0 0 7 3 】

特定のキーワードが含まれていないと判定された場合（ステップ S 3 0 3 : N o）には、制御部 3 2 は、ステップ S 3 0 1 に戻る。一方、特定のキーワードが含まれていると判定された場合（ステップ S 3 0 3 : Y e s）には、情報取得部 3 2 4 は、乗員 P A から所定の要求があった場合に、車両 V E の位置情報を取得する（ステップ S 3 0 4）。

【 0 0 7 4 】

そして、情報取得部 3 2 4 A は、乗員 P A の属性情報を取得する（ステップ S 3 0 5）。そして、特定部 3 2 5 A は、位置情報および属性情報を用いて、位置情報および属性情報に応じて統計的に視線が集中する物体を特定した結果、視線が集中する物体を特定できるかを判定する（ステップ S 3 0 6）。

20

【 0 0 7 5 】

この結果、特定部 3 2 5 A が、視線が集中する物体を特定できないと判定した場合には（ステップ S 3 0 6 : N o）、そのまま処理を終了する。また、特定部 3 2 5 A は、視線が集中する物体を特定できると判定した場合には（ステップ S 3 0 6 : Y e s）、ステップ S 3 0 7 の処理に進む。その後、情報提供部 3 2 8 は、物体情報 D B 3 3 3 から当該ステップ S 3 0 6 にて特定された物体に対応する物体情報を取得し（ステップ S 3 0 7）、通信部 3 1 を介して車載端末 2 に当該物体情報を送信する（ステップ S 3 0 8）。

【 0 0 7 6 】

そして、制御部 2 5 2 は、音声出力部 2 2 及び表示部 2 7 の少なくともいずれかの動作を制御し、情報提供装置 3 から送信された物体情報を音声、文字、及び画像の少なくともいずれかによって、車両 V E の乗員 P A に対して報知する。例えば、物体 O B 1 が「ムーラン・ルージュ」である場合には、物体情報として「あれはムーラン・ルージュです。華やかなダンスショーを夜にやっています。」等の音声は車両 V E の乗員 P A に対して報知される。また、例えば、物体 O B 1 が建物ではなく動物のバッファローである場合には、物体情報として「あれはバッファローです。バッファローは群れで行動します。」等の音声は車両 V E の乗員 P A に対して報知される。その後、反映部 3 2 9 は、物体情報を提供した結果を学習モデルに反映する（ステップ S 3 0 9）。

30

【 0 0 7 7 】

以上説明した本実施の形態 3 によれば、上述した実施の形態 1、2 と同様の効果の他、以下の効果を奏する。本実施の形態 3 に係る情報提供装置 3 B は、乗員 P A から所定の要求があった場合に、車両 V E の位置情報と乗員 P A の属性を示す属性情報を取得し、位置情報および属性情報を用いて、統計的に視線が集中する物体を特定し、特定した物体に関する物体情報を提供する。このため、上述した実施の形態 1、2 で説明した画像取得部 3 2 3、領域抽出部 3 2 6 および物体認識部 3 2 7 で説明した第 2 の学習モデル D B 3 3 2 を設ける必要がなく、情報提供装置 3 B の構成の簡素化を図ることができる。

40

【 0 0 7 8 】

（その他の実施形態）

ここまで、本発明を実施するための形態を説明してきたが、本発明は上述した実施の形態 1 ~ 3 によってのみ限定されるべきものではない。上述した実施の形態 1 ~ 3 に係る情報提供装置 3、3 A、3 B は、位置情報とキーワードを考慮した第 3 の学習モデル、また

50

は、位置情報、キーワードおよび属性情報を考慮した第4の学習モデルを用いて、統計的に視線が集中する物体を特定した。しかしながら、学習モデルを用いずとも、テーブルを用いて統計的に視線が集中する物体を特定するようにしてもよい。

【0079】

例えば、特定部325Aは、第3の学習モデルを用いずに、図13に例示するテーブルを用いて統計的に視線が集中する物体を特定するようにしてもよい。図13に例示するテーブルでは、「位置情報」と「キーワード」と統計的に視線が集中する「物体」とが対応付けられている。

【0080】

また、例えば、特定部325Aは、第4の学習モデルを用いずに、図14に例示するテーブルを用いて統計的に視線が集中する物体を特定するようにしてもよい。図14に例示するテーブルでは、「位置情報」と「キーワード」と「属性情報」と統計的に視線が集中する「物体」とが対応付けられている。これらのテーブルは、予め記憶部33に記憶されていてもよいし、適宜追加や変更されるようになっていてもよい。

10

【0081】

上述した実施の形態1～3において、情報提供装置3、3A、3Bは、特定のキーワードを含む要求情報（音声情報）を取得したことをトリガとして、各処理を実行していた。しかしながら、本実施の形態に係る情報提供装置としては、特定のキーワードを含む要求情報（音声情報）を取得しなくても、常時、当該各処理を実行する構成としても構わない。また、本実施の形態に係る要求情報としては、音声情報に限らず、車両VEの乗員PAによる車載端末2に設けられたスイッチ等の操作部への操作に応じた操作情報であっても構わない。

20

【0082】

上述した実施の形態1～3において、情報提供装置3、3A、3Bの全ての構成を車載端末2に設けても構わない。この場合には、当該車載端末2は、本実施の形態に係る情報提供装置に相当する。また、情報提供装置3、3A、3Bにおける制御部32の一部の機能、及び記憶部33の一部を車載端末2に設けても構わない。この場合には、情報提供システム1全体が本実施の形態に係る情報提供装置に相当する。

【符号の説明】

【0083】

- 3、3A、3B 情報提供装置
- 321 要求情報取得部
- 322 音声解析部
- 323 画像取得部
- 324、324A 情報取得部
- 325、325A 特定部
- 326 領域抽出部
- 327 物体認識部
- 328 情報提供部
- 329 反映部

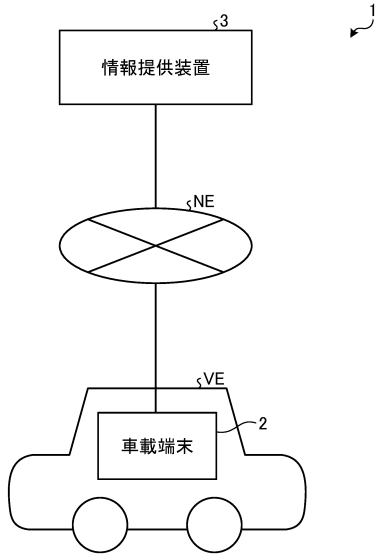
30

40

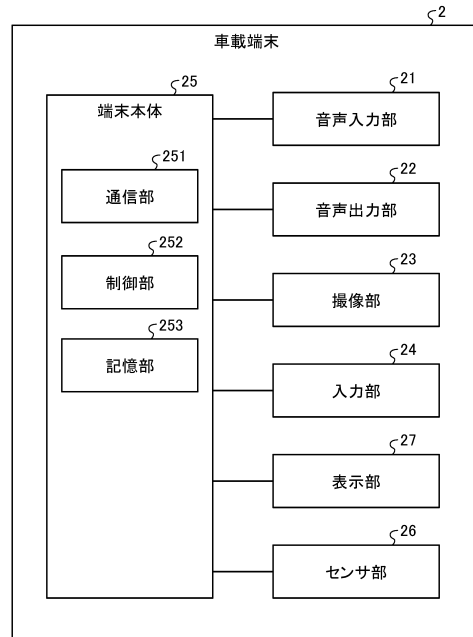
50

【 図 面 】

【 図 1 】



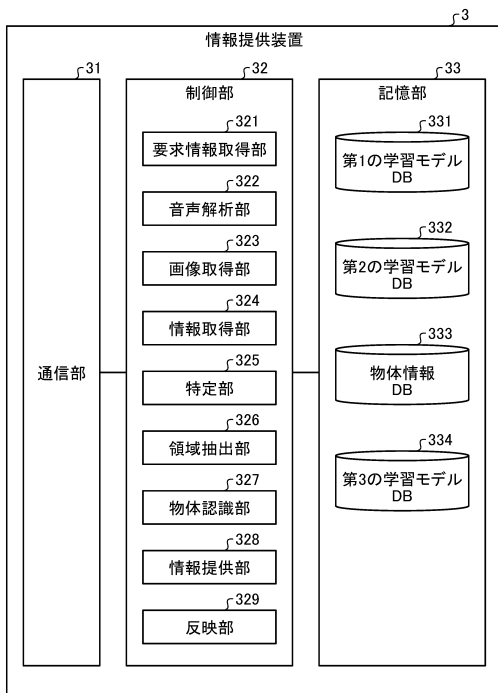
【 図 2 】



10

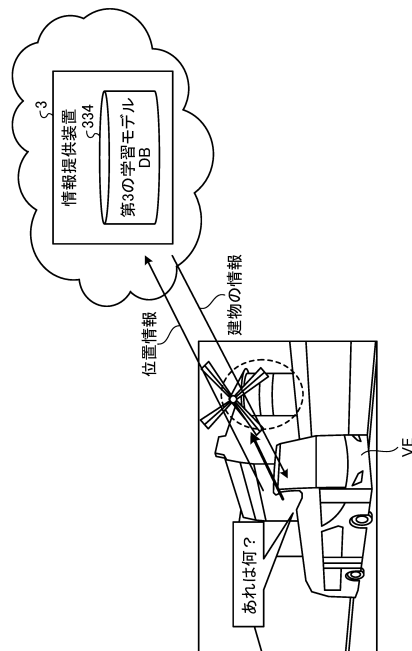
20

【 図 3 】



30

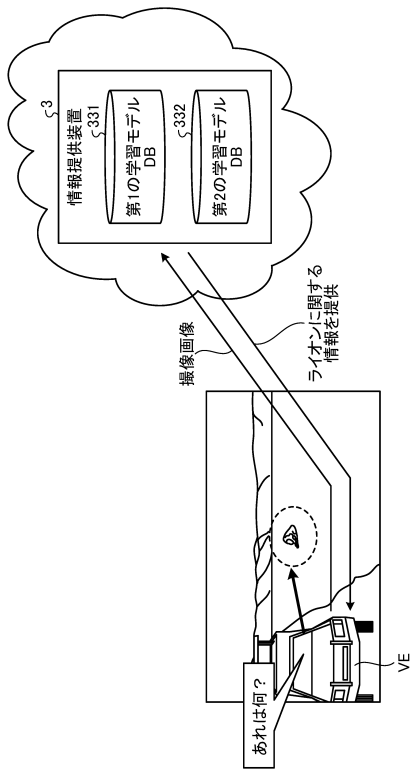
【 図 4 】



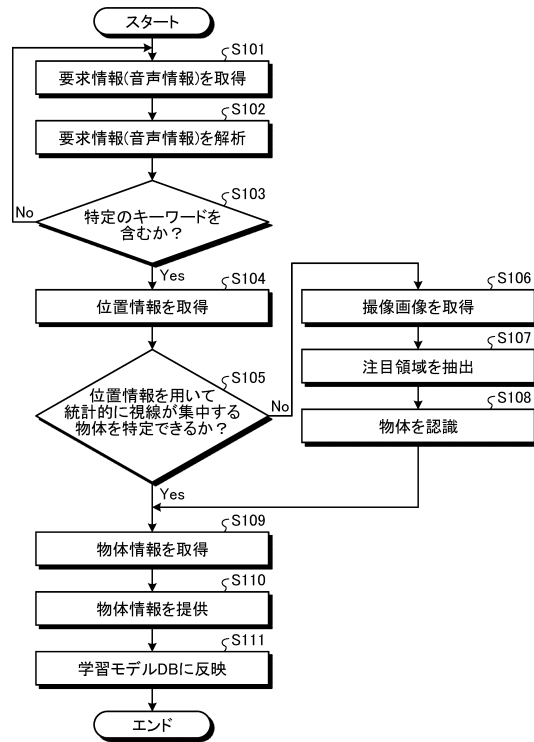
40

50

【 図 5 】



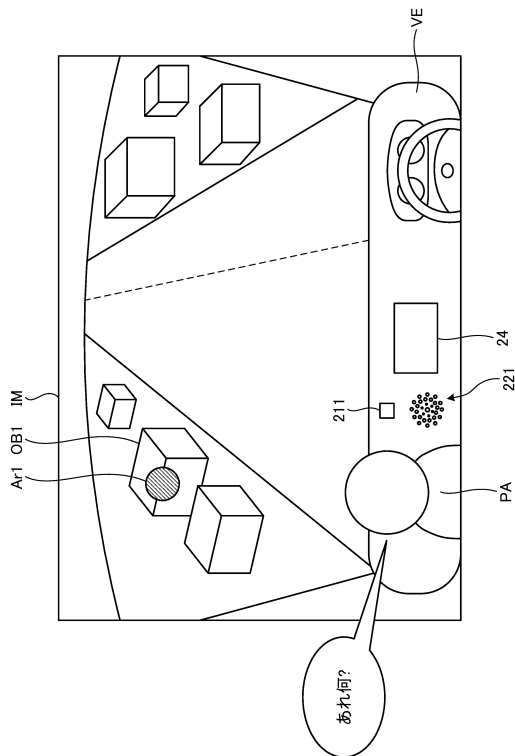
【 図 6 】



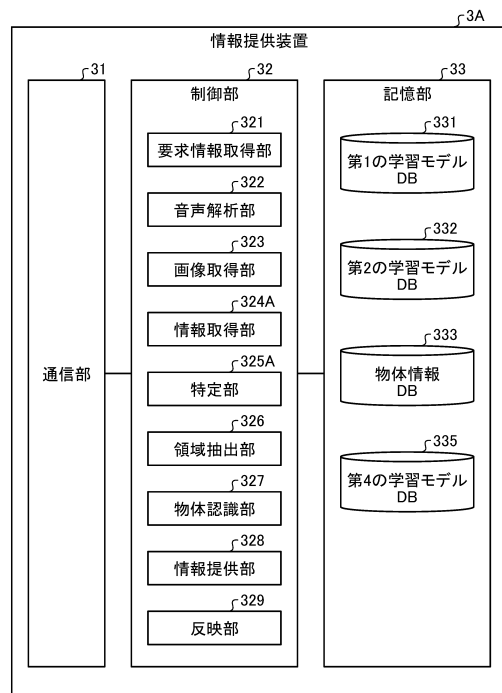
10

20

【 図 7 】



【 図 8 】

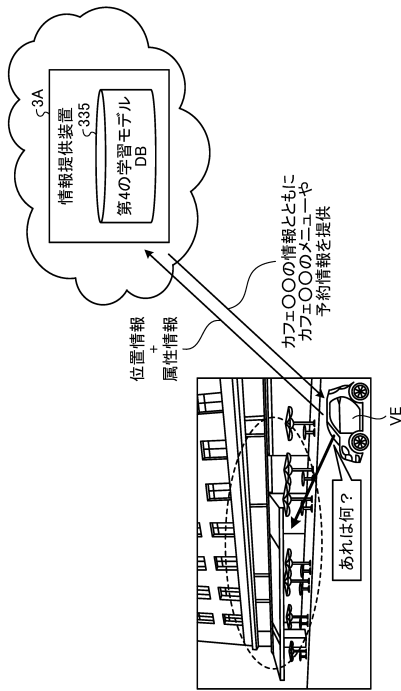


30

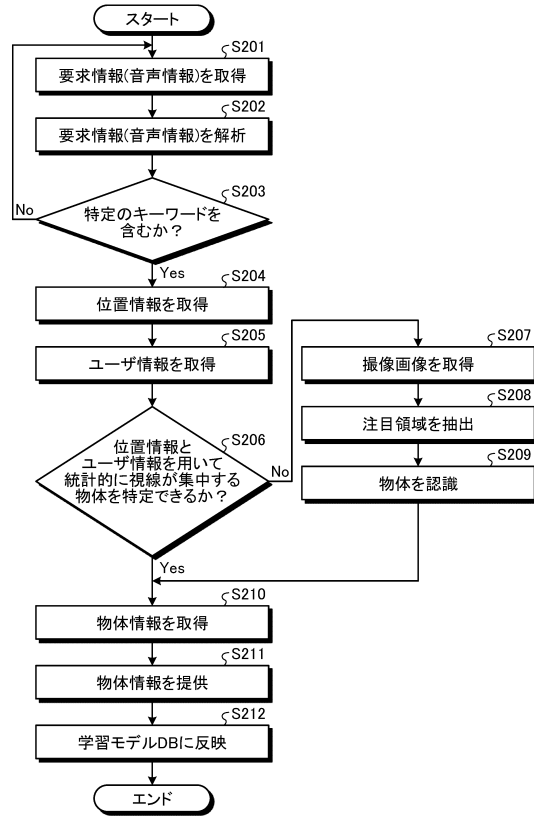
40

50

【図 9】



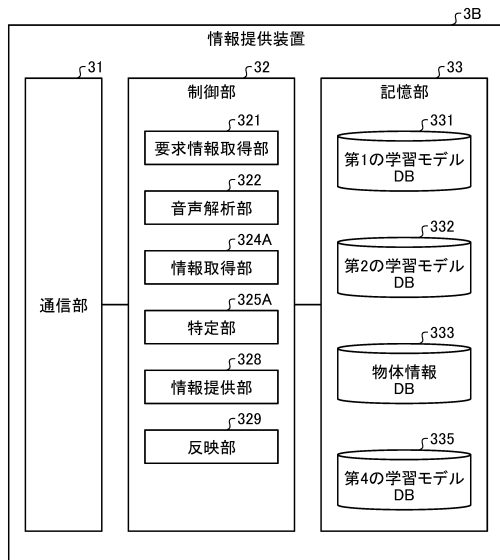
【図 10】



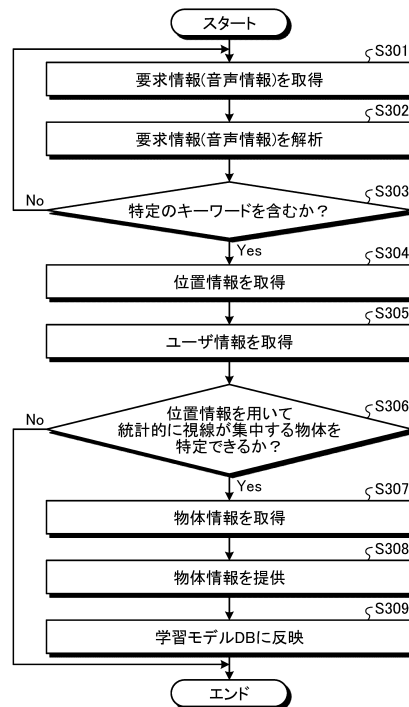
10

20

【図 11】



【図 12】



30

40

50

【 図 1 3 】

位置情報	キーワード	物体
位置X	〇〇〇	物体A
⋮	⋮	⋮

【 図 1 4 】

位置情報	キーワード	属性情報	物体
位置X	〇〇〇	属性Y	物体A
⋮	⋮	⋮	⋮

10

20

30

40

50