

(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2015-207181
(P2015-207181A)

(43) 公開日 平成27年11月19日(2015. 11. 19)

(51) Int. Cl.	F I	テーマコード (参考)
G06T 7/00 (2006.01)	G06T 7/00 U	5B091
G06F 17/27 (2006.01)	G06F 17/27 Z	5E555
H04M 1/00 (2006.01)	H04M 1/00 U	5K127
G10L 15/00 (2013.01)	G10L 15/00 200G	5L096
G06F 3/048 (2013.01)	G06F 3/048 654A	

審査請求 未請求 請求項の数 17 OL (全 34 頁) 最終頁に続く

(21) 出願番号 特願2014-87883 (P2014-87883)
(22) 出願日 平成26年4月22日 (2014. 4. 22)

(特許庁注：以下のものは登録商標)

1. Blu-ray

(71) 出願人 000002185
ソニー株式会社
東京都港区港南1丁目7番1号

(74) 代理人 100095957
弁理士 亀谷 美明

(74) 代理人 100096389
弁理士 金本 哲男

(74) 代理人 100101557
弁理士 萩原 康司

(74) 代理人 100128587
弁理士 松本 一騎

(72) 発明者 小林 賢一郎
東京都港区港南1丁目7番1号 ソニー株式会社内

Fターム(参考) 5B091 AA15 BA19 CB32
最終頁に続く

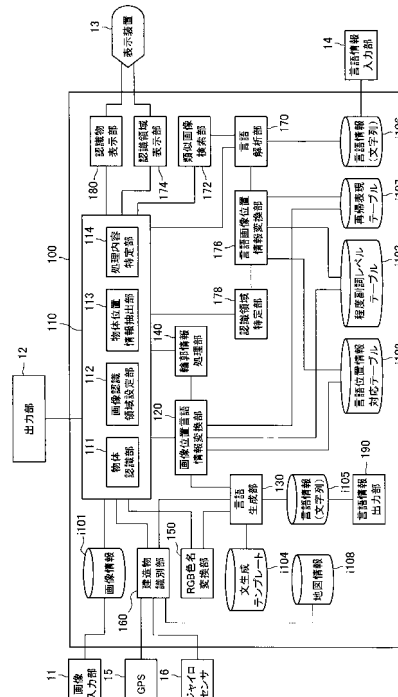
(54) 【発明の名称】 情報処理装置、情報処理方法及びコンピュータプログラム

(57) 【要約】

【課題】画像情報に対する処理内容をユーザから入力される言語情報の内容に応じて決定することが可能な情報処理装置を提供する。

【解決手段】入力された言語情報に基づいて画像の中の領域を特定する画像領域特定部と、前記入力された言語情報に基づいて、前記画像領域特定部が特定した画像の中の領域に対する、前記画像を使用した処理の内容を特定する処理内容特定部と、を備える、情報処理装置が提供される。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

入力された言語情報に基づいて画像の中の領域を特定する画像領域特定部と、前記入力された言語情報に基づいて、前記画像領域特定部が特定した画像の中の領域に対する、前記画像を使用した処理の内容を特定する処理内容特定部と、を備える、情報処理装置。

【請求項 2】

前記処理内容特定部は、前記入力された言語情報に基づいて、前記画像領域特定部が特定した前記画像の中の領域に対する物体の認識処理を実行することを特定する、請求項 1 に記載の情報処理装置。

10

【請求項 3】

前記画像領域特定部は、特定した画像の中の領域で認識された物体を基準として、さらに入力された言語情報に基づいて画像の中の領域を特定する、請求項 2 に記載の情報処理装置。

【請求項 4】

前記処理内容特定部は、前記入力された言語情報に基づいて、前記画像領域特定部が特定した画像の中の領域に含まれる物体の情報を取得する処理を実行することを特定する、請求項 1 に記載の情報処理装置。

【請求項 5】

前記処理内容特定部は、前記物体の情報として該物体の名称を取得する処理を実行することを特定する、請求項 4 に記載の情報処理装置。

20

【請求項 6】

前記画像領域特定部は、特定した画像の中の領域で情報が取得された物体を基準として、さらに入力された言語情報に基づいて画像の中の領域を特定する、請求項 4 に記載の情報処理装置。

【請求項 7】

前記画像領域特定部が前記画像の中の領域を特定した後に、該特定の取り消しを意味する言語情報が入力されると、前記画像領域特定部は、前記画像に対する直前の領域の特定を解除する、請求項 1 に記載の情報処理装置。

【請求項 8】

前記言語情報として前記画像に対する時間の限定を意味する言語情報が入力されていれば、前記処理内容特定部は、前記画像を使用した処理の対象となる前記画像の時間を限定する、請求項 1 に記載の情報処理装置。

30

【請求項 9】

前記処理内容特定部が特定した処理に基づいて、前記画像の物体を認識する物体認識部をさらに備える、請求項 1 に記載の情報処理装置。

【請求項 10】

前記物体認識部での認識の結果を表す言語情報を生成する言語生成部をさらに備える、請求項 9 に記載の情報処理装置。

【請求項 11】

前記処理内容特定部が特定した処理に基づいて、前記画像の物体を認識する処理を前記物体認識部が実行した結果、前記入力された言語情報から物体を一意に特定できない場合は、前記言語生成部は、認識された物体の特徴を付加して言語情報を生成する、請求項 10 に記載の情報処理装置。

40

【請求項 12】

前記処理内容特定部が特定した処理に基づいて、前記画像の物体を認識する処理を前記物体認識部が実行した結果、前記言語生成部は、前記物体認識部による前記入力された言語情報に基づいた物体認識結果の度合いに応じた言語情報を生成する、請求項 10 に記載の情報処理装置。

【請求項 13】

50

前記画像領域特定部が特定した領域を前記画像に重畳して表示させる認識領域表示部をさらに備える、請求項 1 に記載の情報処理装置。

【請求項 1 4】

前記画像領域特定部は、入力された言語情報に基づいて前記画像を所定数に分割し、該分割後の各領域を用いて前記画像の中の領域を特定する、請求項 1 に記載の情報処理装置。

【請求項 1 5】

前記画像領域特定部は、前記入力された言語情報に程度を表す副詞が含まれていれば、前記副詞を参照して前記画像の中の領域を特定する、請求項 1 に記載の情報処理装置。

【請求項 1 6】

入力された言語情報に基づいて画像の中の領域を特定することと、
前記入力された言語情報に基づいて、前記特定された画像の中の領域に対する、前記画像を使用した処理の内容を特定することと、
を含む、情報処理方法。

【請求項 1 7】

コンピュータに、
入力された言語情報に基づいて画像の中の領域を特定することと、
前記入力された言語情報に基づいて、前記特定された画像の中の領域に対する、前記画像を使用した処理の内容を特定することと、
を実行させる、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、情報処理装置、情報処理方法及びコンピュータプログラムに関する。

【背景技術】

【0002】

画像情報を基に物体を認識する技術が広く用いられている（例えば特許文献 1 等参照）。例えば、画像情報から顔や所定の物体を認識する技術は、デジタルカメラ、スマートフォン、パーソナルコンピュータで動作する画像処理ソフトウェア、ロボット装置等で広く用いられている。

【先行技術文献】

【特許文献】

【0003】

【特許文献 1】特開 2013 - 175956 号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

画像情報の認識結果は、画面上の X - Y の 2 軸、または奥行きを考慮した 3 軸上の位置関係として捉えることが出来る。また画像認識の対象が顔や物体など、特定のものであれば、画像中のどこにあるかという位置情報は取得できる。ここで、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定することが出来ると、より効果的に画像情報に対する処理が出来ると考えられる。

【0005】

そこで本開示では、画像情報に対する処理内容をユーザから入力される言語情報の内容に応じて決定することが可能な、新規かつ改良された情報処理装置、情報処理方法及びコンピュータプログラムを提案する。

【課題を解決するための手段】

【0006】

本開示によれば、入力された言語情報に基づいて画像の中の領域を特定する画像領域特定部と、前記入力された言語情報に基づいて、前記画像領域特定部が特定した画像の中の領域に対する、前記画像を使用した処理の内容を特定する処理内容特定部と、を備える、情報処理装置が提供される。

【0007】

また本開示によれば、入力された言語情報に基づいて画像の中の領域を特定することと、前記入力された言語情報に基づいて、前記特定された画像の中の領域に対する、前記画像を使用した処理の内容を特定することと、を含む、情報処理方法が提供される。

【0008】

また本開示によれば、コンピュータに、入力された言語情報に基づいて画像の中の領域を特定することと、前記入力された言語情報に基づいて、前記特定された画像の中の領域に対する、前記画像を使用した処理の内容を特定することと、を実行させる、コンピュータプログラムが提供される。

10

【発明の効果】

【0009】

以上説明したように本開示によれば、画像情報に対する処理内容をユーザから入力される言語情報の内容に応じて決定することが可能な、新規かつ改良された情報処理装置、情報処理方法及びコンピュータプログラムを提供することが出来る。

【0010】

なお、上記の効果は必ずしも限定的なものではなく、上記の効果とともに、または上記の効果に代えて、本明細書に示されたいずれかの効果、または本明細書から把握され得る他の効果が奏されてもよい。

20

【図面の簡単な説明】

【0011】

【図1】本開示の一実施形態に係る情報処理システム1の構成例を示す説明図である。

【図2】本開示の一実施形態に係る情報処理システム1の機能構成例を示す説明図である。

【図3】画像位置言語情報変換部120による画像情報i101の画像領域の分割例を示す説明図である。

【図4】言語位置情報対応テーブルi102に格納される、領域名と言語としての名称との対応例である。

30

【図5】程度副詞レベルテーブルi103に格納される、レベルと副詞との対応例である。

【図6】再帰表現テーブルi107に格納される情報の例を示す説明図である。

【図7】各領域を再分割する際の、分割領域と程度副詞のレベルと組み合わせを示す説明図である。

【図8】画像情報i101の画像領域の再分割例を示す説明図である。

【図9】情報処理装置100が使用する文生成テンプレートi104の一例を示す説明図である。

【図10A】本開示の一実施形態に係る情報処理装置100の動作例を示す流れ図である。

40

【図10B】本開示の一実施形態に係る情報処理装置100の動作例を示す流れ図である。

【図11】画像入力部11から情報処理装置100に入力される画像情報i101の一例を示す説明図である。

【図12】認識物体の重なりの状態の表現について示す説明図である。

【図13】画像入力部11から情報処理装置100に入力される画像情報i101の一例を示す説明図である。

【図14】画像入力部11から情報処理装置100に入力される画像情報i101の一例を示す説明図である。

50

【図15】情報処理装置100による物体認識処理の結果の出力例を示す説明図である。
【図16】情報処理装置100による物体認識処理の結果の出力例を示す説明図である。
【図17】情報処理装置100が認識した物体を強調して表示する例を示す説明図である。

【図18】情報処理装置100による物体認識処理の結果の出力例を示す説明図である。
【図19】情報処理装置100が画像情報i101に対して物体認識処理を実行した際の認識結果を表示する例を示す説明図である。

【図20】情報処理装置100が画像情報i101に対して物体認識処理を実行した際の認識結果を表示する例を示す説明図である。

【図21】画像情報i101とユーザの方向との位置関係の例を示す説明図である。

10

【図22】画像入力部11から情報処理装置100に入力される画像情報i101の一例を示す説明図である。

【図23】情報処理装置100による物体認識処理の結果の出力例を示す説明図である。

【図24】本開示の一実施形態に係る情報処理システム2を示す説明図である。

【図25】本開示の実施形態にかかる情報処理装置100のハードウェア構成例を示すブロック図である。

【発明を実施するための形態】

【0012】

以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

20

【0013】

なお、説明は以下の順序で行うものとする。

1. 本開示の一実施形態
 - 1.1. 概要
 - 1.2. システム構成例
 - 1.3. 機能構成例
 - 1.4. 動作例
2. まとめ

【0014】

< 1. 本開示の一実施形態 >

[1.1. 概要]

まず、本開示の好適な実施の形態について詳細に説明するにあたり、本実施形態の概要について説明する。上述したように、画像情報を基に物体を認識する技術が広く用いられ、画像情報から顔や所定の物体を認識する技術は、デジタルカメラ、スマートフォン、パーソナルコンピュータで動作する画像処理ソフトウェア、ロボット装置等で広く用いられている。

30

【0015】

画像情報の認識結果は、画面上のX-Yの2軸、または奥行きを考慮した3軸上の位置関係として捉えることが出来る。また画像認識の対象が顔や物体など、特定のものがあれば、画像中のどこにあるかという位置情報は取得できる。ここで、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定することが出来ると、より効果的に画像情報に対する処理が出来ると考えられる。

40

【0016】

そこで、本件開示者は、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定するための技術について検討を行なった。そして本件開示者らは、以下で説明するように、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定する処理を実行することで、画像情報に対する効果的な処理が可能となる技術を想到するに至った。

【0017】

50

[1 . 2 . システム構成例]

まず、図面を参照しながら本開示の一実施形態に係る情報処理システムの構成例について説明する。図 1 は、本開示の一実施形態に係る情報処理システム 1 の構成例を示す説明図である。図 1 に示した情報処理システム 1 は、ユーザから入力される言語情報の内容に応じて処理内容を決定することを可能としたシステムである。以下、図 1 を用いて本開示の一実施形態に係る情報処理システム 1 の構成例について説明する。

【 0 0 1 8 】

図 1 に示したように、本開示の一実施形態に係る情報処理システム 1 は、ヘッドマウントディスプレイ (HMD) 10 と、ヘッドマウントディスプレイ 10 から情報を取得し、ヘッドマウントディスプレイ 10 へ情報を出力する情報処理装置 100 と、を含んで構成される。ヘッドマウントディスプレイ 10 は、ユーザの頭部に装着されることを目的とした装置である。ヘッドマウントディスプレイ 10 は、ユーザの頭部に装着されることで、ユーザの眼前に情報を提示することが出来るように構成される。またヘッドマウントディスプレイ 10 は、ユーザの前方または周囲を撮像するカメラや、ユーザが発話した言葉等を收音するマイクを備える。

10

【 0 0 1 9 】

情報処理装置 100 は、ヘッドマウントディスプレイ 10 から取得した情報に対する処理を実行して、その処理結果として得られる情報をヘッドマウントディスプレイ 10 へ出力する。本実施形態では、情報処理装置 100 は、ヘッドマウントディスプレイ 10 から送られる、ヘッドマウントディスプレイ 10 に備えられたカメラで撮像された画像に対する処理を実行する。情報処理装置 100 が実行する処理は、例えば、カメラで撮像された画像に対する画像認識処理である。情報処理装置 100 は、カメラで撮像された画像に対する画像認識処理を実行してその画像認識処理の結果をヘッドマウントディスプレイ 10 へ出力する。

20

【 0 0 2 0 】

そして情報処理装置 100 は、ヘッドマウントディスプレイ 10 に備えられたカメラで撮像された画像に対する処理を実行する際に、ヘッドマウントディスプレイ 10 で入力された言語情報、例えば、ヘッドマウントディスプレイ 10 に備えられたマイクで收音されたユーザの言葉に基づいて処理内容を決定する。

【 0 0 2 1 】

本開示の一実施形態に係る情報処理装置 100 は、ヘッドマウントディスプレイ 10 に備えられたカメラで撮像された画像に対する処理内容を、ヘッドマウントディスプレイ 10 を装着したユーザから入力される言語情報の内容に応じて決定することが可能となる。

30

【 0 0 2 2 】

以上、図 1 を用いて本開示の一実施形態に係る情報処理システム 1 の構成例について説明した。続いて、本開示の一実施形態に係る情報処理システム 1 の機能構成例について説明する。

【 0 0 2 3 】

[1 . 3 . 機能構成例]

図 2 は、本開示の一実施形態に係る情報処理システム 1 の機能構成例を示す説明図である。図 2 には、主に本開示の一実施形態に係る情報処理装置 100 の機能構成例が示されている。また図 2 には、情報処理装置 100 の構成と併せ、ヘッドマウントディスプレイ 10 の構成も示されている。以下、図 2 を用いて本開示の一実施形態に係る情報処理システム 1 の機能構成例について説明する。

40

【 0 0 2 4 】

図 2 に示した情報処理装置 100 は、画像入力部 11 から送られる画像情報 i 101 に対する画像処理を実行し、画像情報 i 101 に含まれる物体を認識することを目的とした装置である。また図 2 に示した情報処理装置 100 は、言語情報入力部 14 から入力された言語情報に基づいて、画像入力部 11 から送られる画像情報 i 101 に対する処理を実行することを目的とした装置である。

50

【0025】

図2に示したように、本開示の一実施形態に係る情報処理装置100は、画像処理部110と、画像位置言語情報変換部120と、言語生成部130と、輪郭情報処理部140と、RGB色名変換部150と、建造物識別部160と、言語解析部170と、類似画像検索部172と、認識領域表示部174と、言語画像位置情報変換部176と、認識領域特定部178と、認識物表示部180と、言語情報出力部190と、を含んで構成される。

【0026】

また図2に示したように、本開示の一実施形態に係るヘッドマウントディスプレイ10は、画像入力部11と、出力部12と、表示装置13と、言語情報入力部14と、GPS15と、ジャイロセンサ16と、を含んで構成される。

10

【0027】

画像処理部110は、画像入力部11が提供する画像情報i101に対する画像処理を実行し、画像情報i101に含まれる物体を認識する。図2に示したように、画像処理部110は、物体認識部111と、画像認識領域設定部112と、物体位置情報抽出部113と、処理内容特定部114と、を含んで構成される。

【0028】

物体認識部111は、画像情報i101に含まれている所定の物体の認識を実行する。物体認識部111が認識の対象とする物体は特定の物体に限定されるものではなく、様々な物体が認識の対象となり得る。物体認識部111は、画像情報i101に含まれている物体の認識の実行により、画像情報i101にどのような物体がいくつ存在しているか、を認識することが出来る。物体認識部111は、画像情報i101に対する認識処理の結果として得られる認識物体に関する情報、例えば認識した物体の名称や、認識した物体の数の情報等を画像位置言語情報変換部120に送る。

20

【0029】

画像認識領域設定部112は、画像情報i101に対して物体認識部111が物体の認識を実行する範囲を設定する。物体認識部111は、画像認識領域設定部112が設定した画像情報i101の範囲に対して、画像情報i101に含まれている所定の物体の認識を実行する。画像認識領域設定部112は、最初に画像情報i101に対して物体認識部111が物体の認識を実行する際には画像情報i101の全範囲について処理を実行するように領域を設定し、例えばユーザなどから、言語情報入力部14を通じて、画像情報i101の範囲を限定しての認識の要求や、検出した物体の近傍について再認識の要求があれば、その物体の近傍に限定して処理を実行するように領域を設定してもよい。例えば、認識領域特定部178が画像情報i101における認識領域を特定すれば、画像認識領域設定部112は認識領域特定部178が特定した認識領域に基づいて、画像情報i101に対して物体認識部111が物体の認識を実行する範囲を設定する。

30

【0030】

物体位置情報抽出部113は、物体認識部111が認識した、画像情報i101に含まれている所定の物体の、画像情報i101における位置を位置情報として抽出する。物体位置情報抽出部113は、画像情報i101に含まれている所定の物体の重心の位置を、位置情報として抽出しても良い。物体位置情報抽出部113は、画像情報i101に含まれている所定の物体の重心の位置を、位置情報として抽出する場合は、その重心の位置を積分により求めても良い。物体位置情報抽出部113は、画像情報i101に含まれている所定の物体の位置情報を画像位置言語情報変換部120に送る。

40

【0031】

処理内容特定部114は、画像処理部110による処理内容を、言語情報入力部14を通じて送られてくる言語情報i106の内容に基づいて特定する。詳細については後述するが、処理内容特定部114は、例えば言語情報i106に、物体名を問い合わせる内容が含まれている場合は、画像処理部110による処理内容として、画像情報i101に含まれている物体の名称を取得するための処理を実行することを特定する。処理内容特定部

50

114は、また例えば言語情報i106に、物体の検索を指示する内容が含まれている場合は、画像処理部110による処理内容として、画像情報i101に含まれている物体を検索するための処理を実行することを特定する。

【0032】

画像位置言語情報変換部120は、画像処理部110から送られてくる情報、具体的には、物体認識部111から送られる認識物体に関する情報、及び物体位置情報抽出部113が抽出した位置情報を、言語情報に変換する処理を実行する。画像位置言語情報変換部120は、言語情報として、画像情報i101のどこにどのような物体が存在しているか、についての情報を生成する。画像位置言語情報変換部120による具体的な変換処理については、後に詳述するが、変換に際しては、画像位置言語情報変換部120は、言語位置情報対応テーブルi102と、程度副詞レベルテーブルi103と、再帰表現テーブルi107と、を用いる。図3は、画像位置言語情報変換部120による画像情報i101の画像領域の分割例を示す説明図である。また図4は、言語位置情報対応テーブルi102に格納される、領域名と言語としての名称との対応例である。また図5は、程度副詞レベルテーブルi103に格納される、レベルと副詞との対応例である。そして図6は、再帰表現テーブルi107に格納される情報の例を示す説明図である。画像位置言語情報変換部120は、画像処理部110から送られてくる情報を言語情報に変換すると、その言語情報を言語生成部130に送る。

10

【0033】

図3は、画像位置言語情報変換部120による画像情報i101の画像領域の分割例を示す説明図である。図3に示した例では、画像位置言語情報変換部120は、画像全体のX-Y座標の最大値を基に、X軸、Y軸それぞれに対して3等分するように領域を分割した上で、各領域にAからIまでのアルファベットで領域名を付与している。

20

【0034】

また画像位置言語情報変換部120は、分割した各領域に対して、言語としての名称を割り当てる。この割り当てに際しては、画像位置言語情報変換部120は言語位置情報対応テーブルi102を参照する。図4は、言語位置情報対応テーブルi102に格納される、領域名と言語としての名称との対応例である。例えば領域Aであれば「左上」や「左の上」という名称に対応する。

【0035】

副詞の付与に際しては、画像位置言語情報変換部120は程度副詞レベルテーブルi103を参照する。図5は、程度副詞レベルテーブルi103に格納される、レベルと副詞との対応例である。図5に示した例では、レベル「大」に対応する副詞として「とても」、「かなり」、「すごく」、「たいへん」等の副詞が割り当てられている。

30

【0036】

画像位置言語情報変換部120は、「もっと」や「あと少し」などの再帰表現を伴う言語情報への変換に、図6に示した再帰表現テーブルi107を用いる。図6に示した再帰表現テーブルi107には、レベル「大」に対応する表現として「もっと」、「さらに」、「それより」等の単語が割り当てられ、レベル「小」に対応する表現として「あと少し」、「もうちょっと」、「ほんのちょっと」等の句が割り当てられている。

40

【0037】

言語生成部130は、画像位置言語情報変換部120から送られてくる言語情報を用いて、画像情報i101の状態を表す言語情報i105、すなわち画像情報i101の状態を表す文章を生成する。本実施形態では、言語生成部130は、画像情報i101の状態を表す文章の生成に際し、文生成テンプレートi104に格納されているテンプレートに、画像位置言語情報変換部120から送られてくる言語情報を適用することで、画像情報i101の状態を表す文章を生成する。言語生成部130による、画像情報i101の状態を表す文章の生成処理の詳細については後述する。

【0038】

輪郭情報処理部140は、画像処理部110から送られてくる情報を用いて、物体認識

50

部 1 1 1 が認識した物体の輪郭の座標情報を取得する。輪郭情報処理部 1 4 0 は、物体の輪郭の座標情報の取得に際し、例えば公知のエッジ検出処理を用いることが出来る。輪郭情報処理部 1 4 0 は、物体の輪郭の座標情報を画像位置言語情報変換部 1 2 0 に送る。画像位置言語情報変換部 1 2 0 は、物体の輪郭の座標情報を取得することで、画像情報 i 1 0 1 に物体が複数存在する場合に、その物体が重なっているかどうかを判断することが出来るるとともに、その重なり状態に応じた言語情報を生成することが出来る。

【 0 0 3 9 】

R G B 色名変換部 1 5 0 は、画像処理部 1 1 0 から送られてくる情報、具体的には、物体認識部 1 1 1 から送られる認識物体に関する情報を用いて、その認識物体の色を取得する。R G B 色名変換部 1 5 0 は、認識物体の色の情報を言語生成部 1 3 0 に送る。言語生成部 1 3 0 は、R G B 色名変換部 1 5 0 から認識物体の色の情報を得ることで、画像情報 i 1 0 1 の状態を表す文章に、物体認識部 1 1 1 が認識した物体の色の情報を加えることが出来る。

10

【 0 0 4 0 】

建造物識別部 1 6 0 は、画像情報 i 1 0 1 に写っている建造物を識別する処理を実行する。具体的には、建造物識別部 1 6 0 は、画像処理部 1 1 0 から送られてくる情報を用いて、物体認識部 1 1 1 が認識した認識物体がどのような建造物であるかどうかを識別する。建造物識別部 1 6 0 は、画像情報 i 1 0 1 に写っている建造物に関する情報を言語生成部 1 3 0 に送る。言語生成部 1 3 0 は、建造物識別部 1 6 0 から画像情報 i 1 0 1 に写っている建造物に関する情報を得ることで、画像情報 i 1 0 1 の状態を表す文章に、画像情報 i 1 0 1 に写っている建造物の情報を加える事ができる。建造物識別部 1 6 0 は、画像情報 i 1 0 1 に写っている建造物に関する情報を取得する際に、地図情報 i 1 0 8 を参照する。また建造物識別部 1 6 0 は、画像情報 i 1 0 1 に写っている建造物に関する情報を取得する際に、GPS 1 5 から送られる現在位置の情報や、ジャイロセンサ 1 6 から送られるヘッドマウントディスプレイ 1 0 の向きの情報を参照してもよい。

20

【 0 0 4 1 】

言語解析部 1 7 0 は、言語情報入力部 1 4 から入力された言語情報（文字列）i 1 0 6 に対する解析処理を実行する。言語情報入力部 1 4 から入力された言語情報 i 1 0 6 は、例えば音声データであってもよく、テキストデータであってもよい。言語情報 i 1 0 6 が音声データである場合は、言語解析部 1 7 0 は、まず音声データをテキストデータへ変換し、変換後のテキストデータに対して解析処理を実行する。言語情報 i 1 0 6 がテキストデータである場合は、そのテキストデータに対して直接解析処理を実行する。

30

【 0 0 4 2 】

言語解析部 1 7 0 は、言語情報 i 1 0 6 に対する解析処理の結果、画像情報 i 1 0 1 に対する問い合わせであることが解れば、画像処理部 1 1 0 に対し、画像情報 i 1 0 1 に含まれている物体に対する物体認識処理の開始を指示する。例えば言語情報 i 1 0 6 が「あれは何？」というように問いかける文章であった場合は、言語解析部 1 7 0 は画像処理部 1 1 0 に対し、画像情報 i 1 0 1 に含まれている物体に対する物体認識処理の開始を指示する。

40

【 0 0 4 3 】

また言語解析部 1 7 0 は、言語情報 i 1 0 6 に対する解析処理の結果、画像情報 i 1 0 1 に対する問い合わせであることが解り、かつ、既知の物体に類似する物体に対する問い合わせであることが解れば、画像処理部 1 1 0 に対し、画像情報 i 1 0 1 に含まれている物体に対する物体認識処理の開始を指示するとともに、類似画像検索部 1 7 2 に対する類似画像の検索を指示する。例えば言語情報 i 1 0 6 が「あのロケットに似ている物体は何？」というように問いかける文章であった場合は、言語解析部 1 7 0 は画像処理部 1 1 0 に対し、画像情報 i 1 0 1 に含まれている物体に対する物体認識処理の開始を指示するとともに、類似画像検索部 1 7 2 に対し、画像情報 i 1 0 1 に含まれているロケットに似ている物体を検索する類似画像の検索処理を指示する。

50

【 0 0 4 4 】

また言語解析部 170 は、言語情報 i 106 に対する解析処理の結果、画像情報 i 101 に対する問い合わせであることが解り、かつ、領域を特定した問い合わせであることが解れば、画像処理部 110 に対し、画像情報 i 101 に含まれている物体に対する物体認識処理の開始を指示するとともに、言語画像位置情報変換部 176 に対し、言語情報 i 106 から画像情報 i 101 における位置の情報に変換する処理を指示する。例えば言語情報 i 106 が「あの右下にある物体は何？」というように問いかける文章であった場合は、画像情報 i 101 に含まれている物体に対する物体認識処理の開始を指示するとともに、言語画像位置情報変換部 176 に対し、「右下」という領域を特定した言語情報を画像情報 i 101 における位置の情報に変換する処理を指示する。

【0045】

言語解析部 170 は、形態素解析により単語単位に分解し、それぞれの単語の品詞、係り受けを認識することで、画像中に領域をどのように設定するのか、その画像を使ってどのような処理を行うのかについて判定する。例えば、入力された言語情報 i 106 が「富士山はどこ？」というように問いかける文章であった場合は、言語解析部 170 は、形態素解析により言語情報 i 106 を「富士山」「は」「どこ」「？」と分割する。そして言語解析部 170 は、分割された言語情報 i 106 における「どこ」というキーワードによって場所検索を行っている判断し、また分割された言語情報 i 106 における「富士山」という名詞が、その場所の対象であると判断する。従って言語解析部 170 は、言語情報 i 106 が「富士山はどこ？」というように問いかける文章であった場合、形態素解析により、その言語情報 i 106 が富士山の場所の検索を要求するための言語情報であると判定することができる。

【0046】

なお、言語解析部 170 は、例えば特開 2011-216071 号公報や、特開 2012-238062 号公報などで開示されている言語解析に関する技術を用いて、上述したような形態素解析による言語情報 i 106 の解析処理を実行しても良い。もちろん、言語解析部 170 は、上記文献で開示されているもの以外の技術を用いて形態素解析による言語情報 i 106 の解析処理を実行しても良いことは言うまでもない。

【0047】

類似画像検索部 172 は、類似画像の検索処理を実行する。本実施形態では、類似画像検索部 172 は、言語情報 i 106 に含まれている物体に似ている、画像情報 i 101 に含まれている物体を検索する処理を実行する。そして本実施形態では、類似画像検索部 172 は、言語情報 i 106 に含まれている物体に、画像情報 i 101 に含まれている物体がどの程度似ているかを表す指標である類似度を計算する。類似画像検索部 172 は、類似度を計算すると、計算した類似度の値を画像処理部 110 に渡す。

【0048】

なお類似度の算出は特定の方法に限定されるものではないが、例えば、特開 2010-3021 号公報等に画像の類似度を計算するための技術が開示されており、類似画像検索部 172 は、上記公報に開示されている技術を用いて画像の類似度を計算しても良い。

【0049】

言語画像位置情報変換部 176 は、言語解析部 170 での言語情報 i 106 に対する解析処理の結果を用いて、言語情報 i 106 を画像情報 i 101 上の位置情報に変換する処理を実行する。言語画像位置情報変換部 176 は、言語情報 i 106 を画像情報 i 101 上の位置情報に変換する際に、画像位置言語情報変換部 120 と同様に、言語位置情報対応テーブル i 102 と、程度副詞レベルテーブル i 103 と、再帰表現テーブル i 107 と、を用いる。言語画像位置情報変換部 176 は、言語情報 i 106 を画像情報 i 101 上の位置情報に変換した結果を認識領域特定部 178 に送る。

【0050】

言語画像位置情報変換部 176 は、例えば、画像情報 i 101 を X 軸において「左、中、右」と3分割して、また Y 軸において「上、中、下」と3分割して、また X 軸と Y 軸との両方で3分割、合計9分割して、言語情報 i 106 と対応させる。

10

20

30

40

50

【 0 0 5 1 】

認識領域特定部 178 は、本開示の画像領域特定部の一例であり、言語画像位置情報変換部 176 による変換結果を用いて、言語解析部 170 での言語情報 i 106 に対する解析処理の結果から画像情報 i 101 上での認識領域を特定する。認識領域特定部 178 は、画像情報 i 101 上での認識領域を特定すると、その特定した領域に対して物体の認識を実行するように画像処理部 110 に通知する。

【 0 0 5 2 】

例えば、言語解析部 170 での言語情報 i 106 に対する解析処理の結果、「右下」という言葉が言語情報 i 106 に含まれていることが解れば、言語画像位置情報変換部 176 は、画像情報 i 101 における「右下」に対応する領域が指定されたと判断し、その判断結果を認識領域特定部 178 に送る。認識領域特定部 178 は、画像情報 i 101 における「右下」に対応する領域に対して物体の認識を実行するように画像処理部 110 に通知する。

10

【 0 0 5 3 】

認識領域表示部 174 は、画像情報 i 101 のどの領域が画像処理部 110 での認識対象になっているかを、画像情報 i 101 に重ね合わせて表示する。認識領域表示部 174 による表示については後に詳述するが、例えば画像情報 i 101 の中央部分が画像処理部 110 での認識対象となっていれば、認識領域表示部 174 は、その中央部分を枠で囲ったり、中央部分以外の部分を暗くしたりすることで、画像情報 i 101 のどの領域が画像処理部 110 での認識対象になっているかを示す。認識領域表示部 174 による認識対象の領域の表示例は後に詳述する。

20

【 0 0 5 4 】

認識物表示部 180 は、画像情報 i 101 に対する画像処理部 110 での画像認識処理によって認識された物体の情報（例えば名称）や輪郭を、画像情報 i 101 に重ねて表示する。認識物表示部 180 による表示については後に詳述するが、例えば画像情報 i 101 に対する画像処理部 110 での画像認識処理によって富士山が認識された場合、認識物表示部 180 は、「富士山です。」のような文章を画像情報 i 101 に重ねたり、富士山の輪郭を強調したりすることで、画像情報 i 101 から富士山が認識されたことを示す。

【 0 0 5 5 】

言語情報出力部 190 は、言語生成部 130 が生成した言語情報 i 105 を出力する。言語情報出力部 190 は、言語情報 i 105 を出力する際に文字情報として出力してもよく、文字情報から生成される音声情報として出力してもよい。

30

【 0 0 5 6 】

以上、図 2 を用いて本開示の一実施形態に係る情報処理システム 1 の機能構成例について説明した。図 2 を用いて本開示の一実施形態に係る情報処理システム 1 を構成する情報処理装置 100 は、図 2 に示した様な構成を有することで、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定する処理を実行することが出来る。画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定する処理を実行することで、情報処理装置 100 は、画像情報に対する効果的な処理が可能となる。

40

【 0 0 5 7 】

[1 . 3 . 動作例]

続いて、本開示の一実施形態に係る情報処理装置 100 の動作例について説明する。図 10A 及び図 10B は、本開示の一実施形態に係る情報処理装置 100 の動作例を示す流れ図である。図 10A 及び図 10B に示したのは、言語情報入力部 14 から言語情報 i 106 の入力を受けて、画像情報 i 101 の認識結果を表す言語情報 i 105 を生成する際の、本開示の一実施形態に係る情報処理装置 100 の動作例である。以下、図 10A 及び図 10B を用いて本開示の一実施形態に係る情報処理装置 100 の動作例について説明する。

【 0 0 5 8 】

50

ユーザがヘッドマウントディスプレイ 10 の言語情報入力部 14 に対して発話すると (ステップ S 101)、情報処理装置 100 は、ユーザが発話した内容に対する音声認識を行なって、言語情報 i 106 を言語情報入力部 14 から取得する (ステップ S 102)。言語情報 i 106 を言語情報入力部 14 から取得した情報処理装置 100 は、言語情報 i 106 に対する言語解析処理を実行する (ステップ S 103)。ステップ S 103 の言語解析処理は、言語解析部 170 が実行する。

【0059】

続いて情報処理装置 100 は、上記ステップ S 103 による言語解析処理の結果、ユーザから入力された言語情報 i 106 に、物体認識を開始する内容が含まれているかどうかを判断する (ステップ S 104)。ステップ S 104 の処理は処理内容特定部 114 が実行し得る。例えば、ユーザから入力された言語情報 i 106 が「あれは何?」というように物体名を問い合わせる内容であった場合は、情報処理装置 100 は、物体認識を開始すると判断する。一方、ユーザから入力された言語情報 i 106 が「富士山はどこ?」というように物体の位置を問い合わせる内容であった場合は、物体認識を開始するものではないと判断する。

10

【0060】

上記ステップ S 104 の判断の結果、ユーザから入力された言語情報 i 106 に、物体認識を開始する内容が含まれていることが解れば (ステップ S 104、Yes)、続いて情報処理装置 100 は、上記ステップ S 103 による言語解析処理の結果、ユーザから入力された言語情報 i 106 に、領域を限定する内容が含まれているかどうかを判断する (ステップ S 105)。ステップ S 105 の処理は処理内容特定部 114 が実行し得る。例えば、ユーザから入力された言語情報 i 106 が「右下にあるものは何?」というように領域を限定した上で物体名を問い合わせる内容であった場合は、情報処理装置 100 は、ユーザから入力された言語情報 i 106 に領域を限定する内容が含まれていると判断する。

20

【0061】

上記ステップ S 105 の判断の結果、ユーザから入力された言語情報 i 106 に、領域を限定する内容が含まれていることが解れば (ステップ S 105、Yes)、続いて情報処理装置 100 は、画像情報 i 101 における物体を認識する領域を限定する処理を実行する (ステップ S 106)。ステップ S 106 の領域を限定する処理は、例えば言語画像位置情報変換部 176 及び認識領域特定部 178 が実行し得る。一方、上記ステップ S 105 の判断の結果、ユーザから入力された言語情報 i 106 に、領域を限定する内容が含まれていないことが解れば (ステップ S 105、No)、情報処理装置 100 はステップ S 106 の処理をスキップする。

30

【0062】

画像情報 i 101 における物体を認識する領域を限定する処理の詳細について説明する。例えばユーザから入力された言語情報 i 106 が「右下にあるものは何?」であることが解析処理で解れば、言語解析部 170 は、その「右下」という言語情報を言語画像位置情報変換部 176 に送る。言語画像位置情報変換部 176 は、言語位置情報対応テーブル i 102 を参照することで、「右下」という言語情報は図 3 における「I」の領域であることを判断する。言語画像位置情報変換部 176 は、この図 3 における「I」の領域の情報を認識領域特定部 178 に送る。そして認識領域特定部 178 は、「I」の領域について物体認識処理を実行するよう画像処理部 110 に指示する。画像処理部 110 は、「I」の領域について物体認識処理を実行する。

40

【0063】

また例えばユーザから入力された言語情報 i 106 が「すごく右下にあるものは何?」や「ちょっと左にあるものは何?」というように、程度を示す副詞が付与されていることが解析処理で解れば、言語画像位置情報変換部 176 は、分割される認識対象領域をさらに分割する。分割される認識対象領域をさらに分割する際には、言語画像位置情報変換部 176 は、程度副詞レベルテーブル i 103 を用いる。図 7 は、各領域を再分割する際の

50

、分割領域と程度副詞のレベルと組み合わせを示す説明図である。そして図 8 は、画像情報 i 1 0 1 の画像領域の再分割例を示す説明図である。図 8 に示したのは、領域「I」を再分割する際の各領域をアルファベットで示す説明図である。

【0064】

例えばユーザから入力された言語情報 i 1 0 6 が「すごく右下」という表現を含んでいる場合は、言語画像位置情報変換部 1 7 6 は、言語位置情報対応テーブル i 1 0 2 から「右下」という表現は「I」の領域であり、「すごく」という程度副詞は程度が「大」なので、図 8 から、領域「I」を再分割した「I - F」、「I - H」、「I - I」の 3 つの領域が「すごく右下」の領域として決定する。そして認識領域特定部 1 7 8 は、言語画像位置情報変換部 1 7 6 が決定した領域を、認識対象の領域として特定する。

10

【0065】

続いて情報処理装置 1 0 0 は、画像情報 i 1 0 1 に対する物体認識処理を実行する（ステップ S 1 0 7）。このステップ S 1 0 7 の物体認識処理は、例えば画像処理部 1 1 0 が、具体的には物体認識部 1 1 1 が実行する。情報処理装置 1 0 0 は、ステップ S 1 0 7 の物体認識処理により、物体の名称、物体の色、画像情報 i 1 0 1 における物体の大きさや位置を取得する。情報処理装置 1 0 0 は、ステップ S 1 0 7 の認識処理の際に、例えば、公知の物体認識処理を実行することが出来る。入力画像から、人物や自動車等の物体を認識するとともに、その認識した物体が何であるかを認識する技術は広く知られており、この物体認識処理は、例えばカメラの自動シャッターや監視カメラに適用されるなどして広く実用化されているものである。具体的には、この物体認識処理は、入力画像の解析により

20

【0066】

画像情報 i 1 0 1 に含まれる物体の認識をステップ S 1 0 7 で行なうと、続いて情報処理装置 1 0 0 は、認識した物体が画像中のどの場所にあるかを抽出するために、認識した物体の重心を算出する（ステップ S 1 0 8）。このステップ S 1 0 8 の処理は、例えば画像処理部 1 1 0 が、具体的には物体位置情報抽出部 1 1 3 が実行する。ステップ S 1 0 8 では、例えば認識した物体に対する積分により重心が算出されても良い。

30

【0067】

図 1 1 は、画像入力部 1 1 から情報処理装置 1 0 0 に入力される画像情報 i 1 0 1 の一例を示す説明図である。図 1 1 には、画像情報 i 1 0 1 に 4 つの物体 3 0 0、3 1 0、3 2 0、3 3 0 が含まれており、情報処理装置 1 0 0 は、上記ステップ S 1 0 7 の物体認識処理により、この 4 つの物体 3 0 0、3 1 0、3 2 0、3 3 0 を認識する。そして情報処理装置 1 0 0 は、上記ステップ S 1 0 8 の重心算出処理により、それぞれの物体 3 0 0、3 1 0、3 2 0、3 3 0 の重心位置 3 0 1、3 1 1、3 2 1、3 3 1 を求める。

【0068】

なお、上記ステップ S 1 0 7 の物体認識処理で認識した物体が複数存在し、言語情報 i 1 0 6 に領域を限定する内容が含まれていなければ、情報処理装置 1 0 0 は、所定の規則に基づいて 1 つの物体を選択する。情報処理装置 1 0 0 は、所定の規則として、例えば画像情報 i 1 0 1 の中央付近の物体について物体認識処理を実行してもよく、一番大きな物体について物体認識処理を実行してもよい。

40

【0069】

図 3 は、画像情報 i 1 0 1 の画像領域の分割例を示す説明図である。図 3 に示した例では、画像情報 i 1 0 1 は、画像全体の X - Y 座標の最大値を基に、X 軸、Y 軸それぞれに対して 3 等分するように領域が分割された上で、各領域に A から I までのアルファベットで領域名が付与されている。画像情報 i 1 0 1 の中央付近の物体について物体認識処理を実行する場合は、情報処理装置 1 0 0 は、画像情報 i 1 0 1 の「E」の部分に重心があるものを認識物体として選択する。図 1 1 のように画像情報 i 1 0 1 に 4 つの物体 3 0 0、3 1 0、3 2 0、3 3 0 が含まれている場合は、物体 3 1 0 の重心位置 3 1 1 が画像情報

50

i 1 0 1の「E」の部分に存在しているので、情報処理装置100は、物体310を認識物体として選択する。

【0070】

上記ステップS107の物体認識処理で認識した物体が複数存在し、言語情報i106に領域を限定する内容が含まれていない場合に、所定の規則として一番大きな物体について物体認識処理を実行するときは、情報処理装置100は、図11に示した画像情報i101については、最も面積が大きい物体300を認識物体として選択する。

【0071】

また上記ステップS107の物体認識処理で認識した物体が複数存在し、言語情報i106に領域を限定する内容が含まれている場合も、所定の規則として一番大きな物体について物体認識処理を実行するときは、情報処理装置100は、図11に示した画像情報i101については、その限定された領域の中で最も面積が大きい物体300を認識物体として選択する。例えばユーザから入力された言語情報i106が「右下にあるものは何？」であれば、情報処理装置100は、図11に示した画像情報i101について、画像情報i101の右下部分(図3の「I」の領域)に存在する最も面積が大きい物体320を認識物体として選択する。

10

【0072】

上記ステップS108で物体の重心を算出すると、続いて情報処理装置100は、上記ステップS107の物体認識処理により認識した物体が1つかどうか判断する(ステップS109)。このステップS109の判断は、例えば画像位置言語情報変換部120が実行し得る。

20

【0073】

上記ステップS109の判断の結果、上記ステップS107の物体認識処理で認識した物体が複数あった場合は(ステップS109、No)、続いて情報処理装置100は、ユーザから入力された言語情報i106の解析によって、物体の類似検索を行なう内容が言語情報i106に含まれているかどうか判断する(ステップS110)。ステップS110の処理は処理内容特定部114が実行し得る。例えば、ユーザから入力された言語情報i106が「あのロケットに似たものは何？」というように物体名を問い合わせる内容であり、かつ物体の類似検索を行なう内容であった場合は、情報処理装置100は、ユーザから入力された言語情報i106に物体の類似検索を行なう内容が含まれていると判断する。

30

【0074】

上記ステップS110の判断の結果、物体の類似検索を行なう内容が言語情報i106に含まれていることが解れば(ステップS110、Yes)、情報処理装置100は、物体の類似検索を行なう(ステップS111)。ステップS111の類似検索処理は、例えば類似画像検索部172が実行する。上述した例のように、言語情報i106が「あのロケットに似たものは何？」というものであれば、情報処理装置100はステップS111でロケットの画像と物体認識処理で認識した物体の画像との類似検索を実行する。

【0075】

情報処理装置100は、物体の類似検索を行なう際に、言語情報i106に含まれている物体に、画像情報i101に含まれている物体がどの程度似ているかを表す指標である類似度を計算する。情報処理装置100は、例えば上述したように特開2010-3021号公報等で開示されている技術を用いて類似度を算出する。

40

【0076】

上記ステップS111で物体の類似検索を行なうと、続いて情報処理装置100は、検索候補を選択する(ステップS112)。具体的には、情報処理装置100は、上記ステップS111での物体の類似検索の結果を取得し、複数の認識物体の中から類似度が最も高い物体を選択する。ステップS111の選択処理は、例えば物体認識部111が実行し得る。

【0077】

50

上記ステップ S 1 0 7 の物体認識処理で認識した物体が 1 つだけであった場合（ステップ S 1 0 9、Yes）、または、上記ステップ S 1 1 2 で検索候補を選択すると、続いて情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 を参照する（ステップ S 1 1 3）。このステップ S 1 1 3 の参照処理は、例えば言語生成部 1 3 0 が実行する。

【0078】

図 9 は、情報処理装置 1 0 0 が使用する文生成テンプレート i 1 0 4 の一例を示す説明図である。図 9 に示した文生成テンプレート i 1 0 4 には、物体認識処理の応答用のテンプレート、物体検索処理の応答用のテンプレート、方向指示処理の応答用のテンプレートが含まれている。

【0079】

ステップ S 1 1 3 では、情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 に含まれている物体認識処理の応答用のテンプレートを参照する。図 9 に示した文生成テンプレート i 1 0 4 の < target - n > (n = 1 , 2 , . . .) タグには、物体認識部 1 1 1 の認識処理によって得られた物体名が当てはめられる。< size - n > (n = 1 , 2 , . . .) タグには、物体認識部 1 1 1 の認識処理によって得られた物体の大きさが当てはめられる。< position - n > (n = 1 , 2 , . . .) タグには、画像位置言語情報変換部 1 2 0 の変換処理より得られるその画像の位置の言語情報が当てはめられる。< color - n > (n = 1 , 2 , . . .) タグには、RGB 色名変換部 1 5 0 より得られる物体の色情報が当てはめられる。

【0080】

上記ステップ S 1 1 3 で文生成テンプレート i 1 0 4 を参照すると、続いて情報処理装置 1 0 0 は、上記ステップ S 1 0 7 の物体認識処理で認識した物体の名称を文生成テンプレート i 1 0 4 のテンプレートに補完して、文を作成する（ステップ S 1 1 4）。このステップ S 1 1 4 の作成処理は、例えば言語生成部 1 3 0 が実行する。ステップ S 1 1 4 の処理により、情報処理装置 1 0 0 は、物体認識処理の応答としての言語情報 i 1 0 5 を生成することが出来る。

【0081】

例えば物体認識部 1 1 1 の認識処理によって、画像情報 i 1 0 1 に含まれている物体が 1 つだけであり、それが富士山であることが認識出来た場合は、情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 を参照して、「富士山です。」という言語情報 i 1 0 5 を生成する。

【0082】

ユーザが「もっと右にあるのは何？」という発話をさらに行うと、言語解析部 1 7 0 は、「もっと」という再帰的な分割を行なう指示とであると判断し、言語画像位置情報変換部 1 7 6 は、「右」という方向を表す位置情報から領域を再度絞り込む。言語画像位置情報変換部 1 7 6 は、絞り込む領域の決定に際して言語位置情報対応テーブル i 1 0 2 を用いる。この場合、言語画像位置情報変換部 1 7 6 は、元の右下の領域「I」をさらに絞り込み、領域「I - F」を認識対象領域として決定する。

【0083】

なお、言語画像位置情報変換部 1 7 6 が認識対象領域を再決定した後に、ユーザから言語情報 i 1 0 6 として取り消しを意味する表現、例えば「やっぱり」、「前の」、「元の」等の言葉を含んだ表現が入力された場合、言語画像位置情報変換部 1 7 6 は、1 つ前に決定した認識対象領域に戻る。

【0084】

図 1 1 のような画像情報 i 1 0 1 の例では、情報処理装置 1 0 0 が最初に右下の物体として物体 3 2 0 を認識した場合、ユーザの「もっと右にあるのは何？」という発話に対して、情報処理装置 1 0 0 は認識領域を図 8 の領域「I - C」、「I - F」、「I - I」に絞り込み、それらの領域にある物体 3 3 0 を認識することが出来る。そして情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 を参照して、例えば「人間です。」という言語情報 i 1 0 5 を生成することが出来る。この再帰的な分割処理は何度も繰り返してもよく、

10

20

30

40

50

情報処理装置 100 は、ユーザの発話に応じて段階的に認識領域を絞り込むことができる。

【0085】

情報処理装置 100 は、認識物体の輪郭から認識物体の面積を求め、画面全体における割合に応じて、認識物体の大きさを言語により表現してもよい。例えば、画面全体の面積に対して認識物体の面積が 30% 以上ある場合は「大きい」と定義し、3% 以下の場合は「小さい」と定義する。図 11 に示した画像情報 i101 の例では、情報処理装置 100 は、例えば「小さい東京タワーです。」や、「大きな富士山です。」という言語情報 i105 を生成することができる。

【0086】

言語生成部 130 は、物体認識部 111 による、言語情報 i106 に基づいた物体認識結果の度合いに応じた言語情報 i105 を生成してもよい。例えば、物体の類似検索を行なう内容が言語情報 i106 に含まれている場合は、情報処理装置 100 は、算出した類似度に応じて生成する言語情報 i105 の中身を変化させても良い。例えば言語情報 i106 が「あのロケットに似ている物体は何？」というように問いかける文章であった場合は、物体の類似検索を行なって算出した類似度がある閾値以上であれば、情報処理装置 100 は、「東京タワーです。」と断言するような言語情報 i105 を生成し、類似度がその閾値より低い場合は、情報処理装置 100 は、「タワーだと思います。」や、「タワーかもしれませんが、」のような確信度を下げた表現で言語情報 i105 を生成してもよい。そして物体の類似検索を行なって算出した類似度が所定の閾値より低ければ、情報処理装置 100 は、「わかりません。」のような全く確信の無い表現で言語情報 i105 を生成してもよい。

【0087】

情報処理装置 100 は、認識物体の色を得ることができる。認識物体の RGB 情報は RGB 色名変換部 150 に送られ、RGB 色名変換部 150 は認識物体の代表的な色を得ることができる。例えば、代表的な色は、認識物体の中で一番多くの面積を占める色とすることができる。

【0088】

RGB 色名変換部 150 で得られた色情報は言語生成部 130 に渡される。言語生成部 130 は、上述したように文生成テンプレート i104 を参照して文の生成を行うが、文の生成の際に、文生成テンプレート i104 上の <color-n> タグを、RGB 色名変換部 150 で得られた色情報に置き換える形で文生成を行う。

【0089】

上述したように、輪郭情報処理部 140 は、画像処理部 110 から送られてくる情報を用いて、物体認識部 111 が認識した物体の輪郭の座標情報を取得する。そして輪郭情報処理部 140 は、物体の輪郭の座標情報を画像位置言語情報変換部 120 に送る。画像位置言語情報変換部 120 は、物体の輪郭の座標情報を取得することで、画像情報 i101 に物体が複数存在する場合に、その物体が重なっているかどうかを判断することができる。とともに、その重なり状態に応じた言語情報を生成することができる。

【0090】

本実施形態では、「含んでいる / 含まれる」「重なっている」「接している」「離れている」という 4 つのパターンによって認識物体の重なり状態を表現する。図 12 は、それぞれの認識物体の重なり状態の表現について示す説明図である。図 12 には、認識物体である三角形の物体と丸い物体とが、どのような重なり状態にあるかが示されている。情報処理装置 100 は、上記ステップ S107 での物体認識処理において、認識物体の重なりを判別し、上記 4 つのパターンのいずれかを用いて、認識物体の重なり関係を変換してもよい。例えば、認識物体の状態が図 12 の左上の状態であれば、情報処理装置 100 は、「三角形：接している：丸」というような言語情報に変換することができる。

【0091】

10

20

30

40

50

情報処理装置 100 は、2つの認識物体が重なっている場合は、その重なり面積の比率の程度によりその重なり程度を表現しても良い。情報処理装置 100 は、例えば2つの認識物体である物体 1 と物体 2 とが重なっている場合、物体 1 と物体 2 とが重なっている部分の面積が物体 1 の面積の 25% 以下の場合には重なり程度を「小」、75% 以上の場合には「大」、それ以外の場合を「中」としてもよい。この程度の決定は輪郭情報処理部 140 が実行してもよい。また2つの認識物体が離れている場合は、情報処理装置 100 は、それぞれの認識物体の重心の距離を求め、認識物体の重心間の距離に応じて程度情報を決定してもよい。情報処理装置 100 は、認識物体の重心間の距離が画面全体の対角線の長さに対して、例えば25% 以内である場合は離れている距離の程度を「小」、75% 以上の場合には「大」、それ以外の場合を「中」としてもよい。

10

【0092】

図 13 は、画像入力部 11 から情報処理装置 100 へ入力される画像情報 i101 の別の例を示す説明図である。図 13 には、画像情報 i101 として2個の赤いリングが机の上に置いてある様子が示されている。従って、図 13 に示した画像情報 i101 を情報処理装置 100 で処理すると、情報処理装置 100 は上述の一連の処理により、言語生成部 130 において「机の上に赤いリングと赤いリングがあります」ではなく、「机の上に赤いリングが2個あります」または「机の上に2つの赤いリングがあります」のような文を生成すること出来る。

【0093】

図 14 は、画像入力部 11 から情報処理装置 100 へ入力される画像情報 i101 の別の例を示す説明図である。図 14 には、机の上に赤いリングが2個ある画像情報 i101 が示されている。従って、図 14 に示した画像情報 i101 を情報処理装置 100 で処理すると、情報処理装置 100 は上述の一連の処理により、言語生成部 130 において「机の上と下に赤いリングがあります」のような文を生成すること出来る。

20

【0094】

上記ステップ S114 で、物体の名称をテンプレートに補完して文を作成すると、続いて情報処理装置 100 は、例えばテキストを音声合成することで生成した文を出力する（ステップ S115）。このステップ S115 の出力処理は、例えば言語情報出力部 190 が実行し得る。もちろん、情報処理装置 100 は、音声合成でなく単に文字列として言語情報 i105 を出力しても良いことは言うまでもない。文字列として言語情報 i105 を出力する場合は、情報処理装置 100 は例えばヘッドマウントディスプレイ 10 の表示装置 13 に言語情報 i105 を出力する。

30

【0095】

一方、上記ステップ S104 の判断の結果、ユーザから入力された言語情報 i106 に、物体認識を開始する内容が含まれていないことが解れば（ステップ S104、No）、続いて情報処理装置 100 は、上記ステップ S103 による言語解析処理の結果、ユーザから入力された言語情報 i106 に、物体検索を行なって、物体の情報を取得する処理を開始する内容が含まれているかどうかを判断する（ステップ S116）。ステップ S116 の処理は、例えば処理内容特定部 114 が実行し得る。

【0096】

上記ステップ S116 の判断の結果、ユーザから入力された言語情報 i106 に、物体検索を行なって、物体の情報を取得する処理を開始する内容が含まれていないことが解れば（ステップ S116、No）、情報処理装置 100 はユーザから入力された言語情報 i106 に対する処理は実行せず、処理を終了する。一方上記ステップ S116 の判断の結果、ユーザから入力された言語情報 i106 に、物体検索を行なって、物体の情報を取得する処理を開始する内容が含まれていることが解れば（ステップ S116、Yes）、続いて情報処理装置 100 は、画像情報 i101 に対する物体認識処理を実行する（ステップ S117）。このステップ S117 の物体認識処理は、例えば画像処理部 110 が、具体的には物体認識部 111 が実行する。情報処理装置 100 は、ステップ S117 の物体認識処理により、物体の情報、具体的には、物体の名称、物体の色、画像情報 i101 に

40

50

おける物体の大きさや位置を取得する。

【0097】

画像情報 i 1 0 1 に含まれる物体の認識をステップ S 1 1 7 で行なうと、続いて情報処理装置 1 0 0 は、認識した物体が画像中のどの場所にあるかを抽出するために、認識した物体の重心を算出する（ステップ S 1 1 8）。このステップ S 1 1 8 の処理は、例えば画像処理部 1 1 0 が、具体的には物体位置情報抽出部 1 1 3 が実行する。ステップ S 1 1 8 では、例えば認識した物体に対する積分により重心が算出されても良い。

【0098】

上記ステップ S 1 1 8 で物体の重心を算出すると、続いて情報処理装置 1 0 0 は、ユーザから入力された言語情報 i 1 0 6 に含まれる物体名と、上記ステップ S 1 1 7 で認識した物体の名前とが一致しているかどうか判断する（ステップ S 1 1 9）。ステップ S 1 1 9 の判断処理は、例えば物体認識部 1 1 1 が実行する。例えばユーザから入力された言語情報 i 1 0 6 が「富士山はどこ？」というものであり、形態素解析により言語情報 i 1 0 6 が富士山の場所を探すことを意味するものであると判断され、上記ステップ S 1 1 7 で認識した物体が富士山であれば、情報処理装置 1 0 0 は、ユーザから入力された言語情報 i 1 0 6 に含まれる物体名と、上記ステップ S 1 1 7 で認識した物体の名前とが一致していると判断することができる。

【0099】

上記ステップ S 1 1 9 の判断の結果、ユーザから入力された言語情報 i 1 0 6 に含まれる物体名と、上記ステップ S 1 1 7 で認識した物体の名前とが一致していないことが解れば（ステップ S 1 1 9、No）、続いて情報処理装置 1 0 0 は、現在のユーザの位置から目標物（すなわち、言語情報 i 1 0 6 に含まれる物体名）がどの方向にどのくらいの距離にあるかを判断する。判断に際して本実施形態では、情報処理装置 1 0 0 は、地図情報 i 1 0 8 を用いた地図検索（ステップ S 1 2 0）、GPS 1 5 が取得した現在位置の情報を用いた GPS 検索（ステップ S 1 2 1）、ジャイロセンサ 1 6 が取得したヘッドマウントディスプレイ 1 0 の向きの情報を用いた方向検出（ステップ S 1 2 2）を実行する。

【0100】

一方上記ステップ S 1 1 9 の判断の結果、ユーザから入力された言語情報 i 1 0 6 に含まれる物体名と、上記ステップ S 1 1 7 で認識した物体の名前とが一致していることが解れば（ステップ S 1 1 9、Yes）、上記ステップ S 1 2 0 ~ 1 2 2 の処理をスキップする。

【0101】

続いて情報処理装置 1 0 0 は、検索結果の強調処理を実行する（ステップ S 1 2 3）。ステップ S 1 2 3 の検索結果の強調処理は、例えば認識物表示部 1 8 0 が実行し得る。ユーザから入力された言語情報 i 1 0 6 に含まれる物体名と、上記ステップ S 1 1 7 で認識した物体の名前とが一致している場合は、情報処理装置 1 0 0 は、上記ステップ S 1 1 7 で認識した物体を強調する処理を実行する。認識した物体を強調する処理としては、例えば、画像情報 i 1 0 1 や、ヘッドマウントディスプレイ 1 0 の表示装置 1 3 を通して見える認識物体の位置を枠で囲う、認識物体を枠で囲った上で枠を点滅させる、認識物体の輪郭を強調する、認識物体を矢印で示す、認識物体または認識物体を囲う枠以外の領域を網掛けにする、等の様々な処理がある。

【0102】

続いて情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 を参照する（ステップ S 1 2 4）。このステップ S 1 2 4 の参照処理は、例えば言語生成部 1 3 0 が実行する。ステップ S 1 2 4 では、情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 に含まれている物体検索処理の応答用のテンプレートを参照する。

【0103】

上記ステップ S 1 2 4 で文生成テンプレート i 1 0 4 を参照すると、続いて情報処理装置 1 0 0 は、上記ステップ S 1 1 7 の物体認識処理で認識した物体の名称を文生成テンプレート i 1 0 4 のテンプレートに補完して、文を作成する（ステップ S 1 2 5）。このス

10

20

30

40

50

ステップ S 1 2 5 の作成処理は、例えば言語生成部 1 3 0 が実行する。ステップ S 1 2 5 の処理により、情報処理装置 1 0 0 は、物体認識処理の応答としての言語情報 i 1 0 5 を生成することが出来る。

【 0 1 0 4 】

例えばユーザが「富士山はどこ？」という言語情報 i 1 0 6 を入力し、言語解析部 1 7 0 での形態素解析により言語情報 i 1 0 6 が富士山の場所を探すことを意味するものであると判断され、物体認識部 1 1 1 の認識処理によって、画像情報 i 1 0 1 から富士山を認識出来た場合は、情報処理装置 1 0 0 は、文生成テンプレート i 1 0 4 を参照して、「富士山はこちらです。」という言語情報 i 1 0 5 を生成する。

【 0 1 0 5 】

上記ステップ S 1 2 5 で、物体の名称をテンプレートに補完して文を作成すると、続いて情報処理装置 1 0 0 は、例えばテキストを音声合成することで生成した文を出力する（ステップ S 1 2 6）。このステップ S 1 2 6 の出力処理は、例えば言語情報出力部 1 9 0 が実行し得る。

【 0 1 0 6 】

情報処理装置 1 0 0 は、物体認識の対象領域を分かりやすくユーザに提示しても良い。図 1 5、図 1 6 及び図 1 8 は、情報処理装置 1 0 0 による物体認識処理の結果の出力例を示す説明図であり、情報処理装置 1 0 0 が画像情報 i 1 0 1 に対して物体認識処理を実行した際の認識領域を表示する例を示す説明図である。図 1 5 に示したのは、画像情報 i 1 0 1 における認識物体が中央部分に位置する物体であり、その認識物体を枠 3 7 1 で囲った上で、ステップ S 1 2 5 で生成した言語情報 i 1 0 5 を表示する領域 3 7 2 を画像情報 i 1 0 1 に重畳させている例である。図 1 6 に示したのは、画像情報 i 1 0 1 における認識物体が右下部分に位置する物体であり、その認識物体を枠 3 7 1 で囲っている例である。

【 0 1 0 7 】

図 1 8 に示したのは、例えばユーザが「富士山はどこ？」と発話した際に表示される認識結果の例であり、画像情報 i 1 0 1 における認識物体が左上部分に位置する物体であり、その認識物体を枠 3 7 1 で囲っている例である。情報処理装置 1 0 0 は、GPS 1 5 が取得した情報、ジャイロセンサ 1 6 が取得した情報、地図情報 i 1 0 8 を用いることにより、ヘッドマウントディスプレイ 1 0 を装着しているユーザの現在位置から、目標の物体がどの方向にどのくらいの距離にあるかを判断できる。

【 0 1 0 8 】

また図 1 7 は、情報処理装置 1 0 0 が認識した物体を強調して表示する例を示す説明図である。図 1 7 に示したのは、画像情報 i 1 0 1 における認識物体が中央部分に位置する物体であり、その認識物体を枠 3 7 1 で囲った上で、枠 3 7 1 で囲われていない領域を網掛けにしている例である。

【 0 1 0 9 】

図 1 9 は、情報処理装置 1 0 0 が画像情報 i 1 0 1 に対して物体認識処理を実行した際の認識結果を表示する例を示す説明図である。図 1 9 に示したのは、例えばユーザが「富士山はどこ？」と発話した際に表示される認識結果の例である。情報処理装置 1 0 0 は、画像情報 i 1 0 1 に対して物体認識処理を実行した結果、画像情報 i 1 0 1 から富士山を認識出来た場合は、富士山を指し示す矢印 3 8 2 をヘッドマウントディスプレイ 1 0 に表示させたり、目標物（富士山）を発見したことを示すマーク 3 8 3 をヘッドマウントディスプレイ 1 0 に表示させたりしてもよい。また情報処理装置 1 0 0 は、画像情報 i 1 0 1 から富士山を認識出来た場合は、富士山を矢印 3 8 2 で示すことに加え、「富士山はこちらです」のような文章 3 8 4 をヘッドマウントディスプレイ 1 0 に表示させてもよい。

【 0 1 1 0 】

状況によっては、画像情報 i 1 0 1 から物体を認識できない場合がある。例えば天候が悪く、画像情報 i 1 0 1 から富士山を認識できない場合が考えられる。そのような場合、情報処理装置 1 0 0 は、GPS 1 5 が取得した情報、ジャイロセンサ 1 6 が取得した情報

10

20

30

40

50

、地図情報 i 1 0 8 を用いることにより、富士山が本来存在するであろう場所を画像情報 i 1 0 1 の中から特定することができる。

【 0 1 1 1 】

図 2 0 は、情報処理装置 1 0 0 が画像情報 i 1 0 1 に対して物体認識処理を実行した際の認識結果を表示する例を示す説明図である。図 2 0 に示したのは、富士山が雲で隠れている状態が写っている画像情報 i 1 0 1 に対して、例えばユーザが「富士山はどこ？」と発話した際に表示される認識結果の例である。このように富士山は雲で隠れて見えないが、情報処理装置 1 0 0 は、GPS 1 5 が取得した情報、ジャイロセンサ 1 6 が取得した情報、地図情報 i 1 0 8 を用いることにより、富士山が本来存在する場所を特定し、枠 3 7 1 で囲ってヘッドマウントディスプレイ 1 0 に表示させることができる。

10

【 0 1 1 2 】

以上、図 1 0 A、1 0 B を用いて本開示の一実施形態にかかる情報処理装置 1 0 0 の動作例について説明した。

【 0 1 1 3 】

上述の説明では、ユーザから入力された言語情報 i 1 0 6 が「あの右下にある物体は何？」というように位置を特定するものである例を示したが、ユーザから入力される言語情報 i 1 0 6 は、例えば物体の持つ色、大きさ、形などの特徴を特定するものであってもよい。例えばユーザから入力された言語情報 i 1 0 6 が「赤くて大きなものは何？」というものであれば、情報処理装置 1 0 0 は、画像情報 i 1 0 1 の中から赤くて大きな物体を認識し、その物体について生成した言語情報 i 1 0 5 を生成してもよい。もちろん、ユーザから入力された言語情報 i 1 0 6 に位置と特徴の両方が特定されていてもいいことは言うまでもない。すなわち、ユーザから入力された言語情報 i 1 0 6 が「あの右下にある赤くて大きなものは何？」というものであれば、情報処理装置 1 0 0 は、画像情報 i 1 0 1 の右下部分の赤くて大きな物体を認識し、その物体について生成した言語情報 i 1 0 5 を生成してもよい。

20

【 0 1 1 4 】

情報処理装置 1 0 0 は、認識対象となる物体の画像情報 i 1 0 1 での他の物体との位置関係が指定された場合に、複数の認識物体から任意の認識物体の名称を取得することもできる。例えば図 1 1 に示した画像情報 i 1 0 1 の場合、画面の中央部分に位置する物体 3 1 0 (タワー) が認識された時点で、ユーザが「タワーの右にあるのは何？」という発話をしたとする。言語解析部 1 7 0 は、「タワー」が予め認識されており、さらに発話されたのが物体の名前なので、その「タワー」が認識の起点となる物体の名前であると判断する。そして言語画像位置情報変換部 1 7 6 は、「タワー」と「右」という単語から認識領域を決定し、認識領域特定部 1 7 8 は、言語画像位置情報変換部 1 7 6 の決定により、X 座標方向で物体 3 1 0 の重心位置 3 1 1 のより正の方向を認識領域として特定する。

30

【 0 1 1 5 】

この場合、情報処理装置 1 0 0 はユーザの発話等により認識領域を設定し直して物体認識を行なっても良い。また情報処理装置 1 0 0、X 座標方向で正の方向に取得済みの認識物体がある場合は、その認識物体の結果を利用してもよい。

【 0 1 1 6 】

例えば図 1 1 に示した画像情報 i 1 0 1 の場合、ユーザが「タワーの右にあるのは何？」という発話をする、情報処理装置 1 0 0 は、物体 3 2 0、3 3 0 の 2 つの物体を認識することが出来る。しかし、物体 3 2 0の方が物体 3 1 0に近いので、情報処理装置 1 0 0 は、言語情報 i 1 0 5 として「自動車です。」という文章を生成することが出来る。もちろん情報処理装置 1 0 0 は、起点となる物体に最も近いものを認識物体として決定してもよく、再認識領域の中で一番面積が大きなものを認識物体として決定してもよい。

40

【 0 1 1 7 】

またユーザが「タワーのちょっと右にあるものは何？」や「タワーよりかなり右にあるものは何？」などのように「ちょっと」や「かなり」などの程度を示す副詞を含んだ内容を発話した場合は、情報処理装置 1 0 0 は、その副詞に対応するレベルの大小により選択

50

範囲を変更してもよい。また位置関係だけでなく、ユーザが「タワーよりもっと赤いのは何？」や「もっと大きいのは何？」などの属性の違いを指定する内容を発話した場合は、情報処理装置100は、その属性による認識物体の絞り込みを行なっても良い。

【0118】

認識物体が画像情報i101の中にある場合は、情報処理装置100は、例えばジャイロセンサ16が取得した情報を用いて物体が存在する方向を検出し、物体が存在する方向をユーザに通知してもよい。図21は、画像情報i101とユーザの方向との位置関係の例を示す説明図である。情報処理装置100は、ユーザの方向に対する対象物の方向の言語情報の対応に従って、認識対象物体がユーザから見てどの方向にあるかをユーザに指示する。

10

【0119】

例えばジャイロセンサ16が取得した情報を用いることで富士山がユーザの後方にあることが解った場合は、情報処理装置100は、文生成テンプレートi104の方向指示用のテンプレートを用いて文を生成することが出来る。具体的には、情報処理装置100は、文生成テンプレートi104の方向指示用のテンプレートの<direction-n>(n=1,2,...)タグに方向の情報を当てはめて、例えば「富士山は後ろです。」という文を生成することが出来る。そして情報処理装置100は、その生成した文に基づいてユーザに指示を与えることができる。また例えば、目標物体が画像情報i101の中に含まれるためには(すなわち、目標物体がユーザの視界に入るためには)ユーザが左右方向で画角30度以内での移動で済む場合、言語生成部130は文生成テンプレートi104の方向指示用のテンプレートの<direction-n>タグに「右」や「左」という情報を当てはめて、「もう少し右です。」や「もう少し左です。」などの文を生成することができる。

20

【0120】

なお上述の例では、物体認識の開始を促す指示がユーザの発話によって行われていたが、本開示に係る例に限定されるものではない。例えば、情報処理装置100に対して送出した所定の信号、例えばヘッドマウントディスプレイ10に備えられているボタン、キーボード、スイッチ等に対するユーザ操作に応じて生成される信号が、情報処理装置100での物体認識開始のトリガとして用いられても良い。

【0121】

また情報処理装置100は、物体認識処理に際してヘッドマウントディスプレイ10を使用しているユーザの手と指を認識し、ユーザの指が指している方向の物体を認識してもよい。ユーザの指が指している方向の物体を認識する場合は、情報処理装置100は、画像情報i101で手と指を認識し、手及び指が認識された領域の重心からY軸上の正の方向を対象物体の方向とする。

30

【0122】

図22は、画像入力部11から情報処理装置100に入力される画像情報i101の一例を示す説明図である。図22には、ヘッドマウントディスプレイ10を使用しているユーザの手が画像情報i101に含まれている状態が示されている。符号401は、手及び指が認識された領域の重心位置を示し、符号402は重心位置401からユーザの指先へ向かうベクトルを示している。情報処理装置100は、画像情報i101に対する物体認識処理によりベクトル402を求めると、そのベクトル402の方向にある物体を対象物体として認識する。図22に示した例では、情報処理装置100は富士山を認識し、「富士山です。」のような文章を生成することができる。このようにユーザの手と指を認識し、ユーザの指が指している方向の物体を認識することで、ユーザが発話等をしなくても情報処理装置100はユーザが意図する物体を認識し、その物体の認識結果を言語情報で出力することができる。

40

【0123】

なお、情報処理装置100が物体認識処理を行なった結果、「わかりません。」のような全く確信の無い表現で言語情報i105を生成した場合に、ユーザがその物体を指で指

50

し示したときは、情報処理装置 100 は、物体認識の閾値を下げて、その物体についての確からしさが低い状態で物体認識を行うようにしてもよい。

【0124】

情報処理装置 100 は、AR（拡張現実）の技術を適用して認識結果を出力しても良い。例えば地図情報 i108 として星座情報を用いれば、ユーザが「みずがめ座はどこ？」と夜空を見ながら発話すると、情報処理装置 100 は、星座を見つけた時点で「みずがめ座はこちらです。」という言語情報 i105 をヘッドマウントディスプレイ 10 に出力することが出来るとともに、星座の画像を重ねあわせてヘッドマウントディスプレイ 10 に出力することができる。図 23 は、情報処理装置 100 による物体認識処理の結果の出力例を示す説明図である。図 23 には、物体認識処理の結果として、情報処理装置 100 が星座の画像を重ねあわせてヘッドマウントディスプレイ 10 に出力している様子を示したものである。

10

【0125】

情報処理装置 100 は、ユーザから入力される言語情報を解析し、その解析結果に基づいて物体認識処理を実行し、ユーザから言語情報によって指定された物体が画像情報 i101 の中央に位置した場合は、ヘッドマウントディスプレイ 10 に対してユーザから言語情報によって指定された物体が画像情報 i101 の中央に位置したことを示す通知、例えば振動や音を出力する指示を送出しても良い。

【0126】

情報処理装置 100 は、ユーザから入力される言語情報 i106 に基づく画像処理の実行により、認識物体が複数検出して一意に特定できない場合は、言語生成部 130 で言語情報 i105 を生成する際に物体の特徴を付加した形で生成する。特徴には、例えば色、大きさ、位置、明るさ、距離などがある。例えばユーザが入力した言語情報 i106 が、「あの赤いものは何？」という内容であった場合に、画像情報 i101 に例えばリンゴとトマトのように異なる赤い物体が複数存在したとする。このとき、情報処理装置 100 は、例えば「赤いものが複数ありますが、一番大きなものはリンゴです。」など物体の特徴を付加した言語情報 i105 を生成する。

20

【0127】

例えばユーザが入力した言語情報 i106 が、「あの赤いものは何？」という内容であった場合に、画像情報 i101 に例えばリンゴとトマトのように異なる赤い物体が複数存在したとする。このとき、情報処理装置 100 は、「赤いものが複数あります。どれですか？」や、「赤いものが複数あります。指で示してください。」などのさらなる情報の要求をユーザに求める言語情報 i105 を生成してもよい。

30

【0128】

情報処理装置 100 は、ユーザから入力される言語情報 i106 に基づく画像処理の実行によって、物体認識できない場合は、例えば所定のメッセージを含んだ言語情報 i105 を生成する。例えば情報処理装置 100 は、「物体が認識できません」という言語情報 i105 を生成してもよく、「赤いものは見つかりませんがオレンジがあります。」と認識された物体の特徴の違いを説明して見つかった物体名のみを返す言語情報 i105 を生成してもよい。また例えば情報処理装置 100 は、物体認識できない場合は、「指でさしてください。」と方向を特定させる言語情報 i105 を生成してもよい、方向を特定させる言語情報 i105 を生成することで、情報処理装置 100 は、物体認識の閾値を下げて、確からしさが低い状態で物体認識を行うことができる。

40

【0129】

情報処理装置 100 は、静止画だけでなく、動画像に対してもユーザが入力した言語情報 i106 に対する言語解析処理及び物体認識処理を実行することが出来る。

【0130】

上述の例では、ヘッドマウントディスプレイ 10 と、ヘッドマウントディスプレイ 10 から情報を取得し、ヘッドマウントディスプレイ 10 へ情報を出力する情報処理装置 100 と、を有する情報処理システム 1 を例に挙げて説明したが、本開示に係る例に限定され

50

るものではない。

【0131】

図24は、本開示の一実施形態に係る情報処理システム2を示す説明図である。図24に示した情報処理システム2は、出力装置20と、情報処理装置100と、入力装置30と、を含んで構成される。出力装置20は、動画像データ等のデジタルコンテンツを出力する装置であり、情報処理装置100によってそのデジタルコンテンツの出力が制御される装置である。入力装置30は、情報処理装置100に対して言語情報を入力する装置であり、例えばマイクやキーボード等の入力デバイスで構成される。

【0132】

図24に示したように構成された情報処理システム2において、ユーザが入力装置30を用いて情報処理装置100に言語情報を入力すると、情報処理装置100は、その入力させた言語情報を解析し、出力装置20へのデジタルコンテンツの出力を制御することが出来る。

【0133】

例えばユーザが入力装置30を用いて入力した言語情報i106が、「今のは何？」や、「さっき赤い服を着ていたのは誰？」というような時間を指定した内容であれば、情報処理装置100は、動画像である画像情報i101に対する物体認識処理を、例えば物体認識部111で実行する。その際に情報処理装置100は、画像情報i101の中でユーザが発話した時点では含まれない物体の中から、ある基準に合致する物体を認識物体として選択する。基準としては、例えば最も面積が大きいもの、最も大きな動きをしたもの、最も輝度が大きいもの、最も大きな音を出していたもの、フレーム間の差分を検出して特徴的な動きや状態の変化をしたもの、などがあり得る。

【0134】

そして情報処理装置100は、認識物体を選択すると、その認識物体が含まれている時点の画像（フレーム画像）を抽出し、抽出した画像に対して上述してきた画像処理を実行する。情報処理装置100は、抽出した画像に対して上述してきた画像処理を実行することで、例えばユーザからの「今のは何？」という問いかけに対して「自動車です。」という言語情報i105を生成することが可能になる。

【0135】

動画像に対する処理によって、情報処理装置100は、ユーザからの質問ではなく、ユーザからの言語情報の内容に対して応答することが可能になる。例えばユーザが入力した言語情報i106が、「さっき赤い服を着た人が出ていたシーンまで戻して」というような巻き戻し再生を指示した内容であれば、情報処理装置100は、動画像である画像情報i101に対して赤い服を着た人を探す物体認識処理を、例えば物体認識部111で実行する。巻き戻し再生が指示された場合、情報処理装置100は動画像を静止画に変換して、上述の認識処理を逐次繰り返して行うことで、ユーザが言語情報を入力した時点から、ユーザに指定された物体が最初に見つかる時間まで出力装置20に巻き戻させることで、ユーザに指定された物体（赤い服を着た人）が登場する場面まで動画像を巻き戻すことができる。そして情報処理装置100は、ユーザに指定された物体が最初に見つかる時間から、ユーザに指定された物体が画像から消える時点までさらに巻き戻すことにより、その物体が直近に出現したシーンの先頭まで出力装置20に巻き戻させることができる。

【0136】

また例えばユーザが入力した言語情報i106が、「さっき赤い服を着た人が最初に出たシーンまで戻して」というような巻き戻し再生を指示した内容であれば、情報処理装置100は、動画像である画像情報i101に対して赤い服を着た人が最初に登場する画像を探す物体認識処理を、例えば物体認識部111で実行することができる。

【0137】

情報処理装置100は、動画像である画像情報i101に対する認識処理の結果得られる文章によって、画像情報i101がどのようなシーンであるかを示すシーン情報を表現出来る。そして動画像である画像情報i101からシーン情報が生成される場合は、情報

10

20

30

40

50

処理装置 100 は、そのシーン情報を用いた物体認識処理を、例えば物体認識部 111 で実行することができる。例えば動画像である画像情報 i101 に対する認識処理の結果、シーン情報として、画像情報 i101 には A さんと B さんの共演シーンが存在するという情報が得られた場合に、ユーザが「A さんと B さんの共演シーンを見せて」という言語情報 i106 を情報処理装置 100 に入力したとする。情報処理装置 100 は、その「A さんと B さんの共演シーンを見せて」という言語情報 i106 に基づいて、シーン情報から得られる A さんと B さんの共演シーンから画像情報 i101 を出力するよう出力装置 20 に指示することが可能になる。

【0138】

今までの説明で 2 次元の画像に関して説明したが、本開示に係る例に限定されるものではない。すなわち情報処理装置 100 は、さらに Z 軸が設けられる画像情報 i101 について、奥行方向に関して同様の処理を行うことも可能である。また、上述した各種テーブルや閾値はユーザによって任意の値に変更することが可能であることは言うまでもない。

10

【0139】

< 2 . ハードウェア構成例 >

次に、図 25 を参照して、本開示の一実施形態にかかる情報処理装置 100 のハードウェア構成について説明する。図 25 は、本開示の実施形態にかかる情報処理装置 100 のハードウェア構成例を示すブロック図である。上記の各アルゴリズムは、例えば、図 25 に示す情報処理装置のハードウェア構成を用いて実行することが可能である。つまり、当該各アルゴリズムの処理は、コンピュータプログラムを用いて図 25 に示すハードウェアを制御することにより実現される。なお、このハードウェアの形態は任意であり、例えば、パーソナルコンピュータ、携帯電話、PHS、PDA 等の携帯情報端末、ゲーム機、接触式又は非接触式の IC チップ、接触式又は非接触式の IC カード、又は種々の情報家電がこれに含まれる。但し、上記の PHS は、Personal Handy - phone System の略である。また、上記の PDA は、Personal Digital Assistant の略である。

20

【0140】

図 25 に示すように、このハードウェアは、主に、CPU 902 と、ROM 904 と、RAM 906 と、ホストバス 908 と、ブリッジ 910 と、を有する。さらに、このハードウェアは、外部バス 912 と、インターフェース 914 と、入力部 916 と、出力部 918 と、記憶部 920 と、ドライブ 922 と、接続ポート 924 と、通信部 926 と、を有する。但し、上記の CPU は、Central Processing Unit の略である。また、上記の ROM は、Read Only Memory の略である。そして、上記の RAM は、Random Access Memory の略である。

30

【0141】

CPU 902 は、例えば、演算処理装置又は制御装置として機能し、ROM 904、RAM 906、記憶部 920、又はリムーバブル記録媒体 928 に記録された各種プログラムに基づいて各構成要素の動作全般又はその一部を制御する。ROM 904 は、CPU 902 に読み込まれるプログラムや演算に用いるデータ等を格納する手段である。RAM 906 には、例えば、CPU 902 に読み込まれるプログラムや、そのプログラムを実行する際に適宜変化する各種パラメータ等が一時的又は永続的に格納される。

40

【0142】

これらの構成要素は、例えば、高速なデータ伝送が可能なホストバス 908 を介して相互に接続される。一方、ホストバス 908 は、例えば、ブリッジ 910 を介して比較的データ伝送速度が低速な外部バス 912 に接続される。また、入力部 916 としては、例えば、マウス、キーボード、タッチパネル、ボタン、スイッチ、及びレバー等が用いられる。さらに、入力部 916 としては、赤外線やその他の電波を利用して制御信号を送信することが可能なリモートコントローラ（以下、リモコン）が用いられることもある。

【0143】

出力部 918 としては、例えば、CRT、LCD、PDP、又は ELD 等のディスプレ

50

イ装置、スピーカ、ヘッドホン等のオーディオ出力装置、プリンタ、携帯電話、又はファクシミリ等、取得した情報を利用者に対して視覚的又は聴覚的に通知することが可能な装置である。但し、上記のCRTは、Cathode Ray Tubeの略である。また、上記のLCDは、Liquid Crystal Displayの略である。そして、上記のPDPは、Plasma Display Panelの略である。さらに、上記のELDは、Electro-Luminescence Displayの略である。

【0144】

記憶部920は、各種のデータを格納するための装置である。記憶部920としては、例えば、ハードディスクドライブ(HDD)等の磁気記憶デバイス、半導体記憶デバイス、光記憶デバイス、又は光磁気記憶デバイス等が用いられる。但し、上記のHDDは、Hard Disk Driveの略である。

10

【0145】

ドライブ922は、例えば、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリ等のリムーバブル記録媒体928に記録された情報を読み出し、又はリムーバブル記録媒体928に情報を書き込む装置である。リムーバブル記録媒体928は、例えば、DVDメディア、Blu-rayメディア、HDDVDメディア、各種の半導体記憶メディア等である。もちろん、リムーバブル記録媒体928は、例えば、非接触型ICチップを搭載したICカード、又は電子機器等であってもよい。但し、上記のICは、Integrated Circuitの略である。

【0146】

20

接続ポート924は、例えば、USBポート、IEEE1394ポート、SCSI、RS-232Cポート、又は光オーディオ端子等のような外部接続機器930を接続するためのポートである。外部接続機器930は、例えば、プリンタ、携帯音楽プレーヤ、デジタルカメラ、デジタルビデオカメラ、又はICレコーダ等である。但し、上記のUSBは、Universal Serial Busの略である。また、上記のSCSIは、Small Computer System Interfaceの略である。

【0147】

通信部926は、ネットワーク932に接続するための通信デバイスであり、例えば、有線又は無線LAN、Bluetooth(登録商標)、又はWUSB用の通信カード、光通信用のルータ、ADSL用のルータ、又は接触又は非接触通信用のデバイス等である。また、通信部926に接続されるネットワーク932は、有線又は無線により接続されたネットワークにより構成され、例えば、インターネット、家庭内LAN、赤外線通信、可視光通信、放送、又は衛星通信等である。但し、上記のLANは、Local Area Networkの略である。また、上記のWUSBは、Wireless USBの略である。そして、上記のADSLは、Asymmetric Digital Subscriber Lineの略である。

30

【0148】

以上、情報処理装置100のハードウェア構成の一例を示した。上記の各構成要素は、汎用的な部材を用いて構成されていてもよいし、各構成要素の機能に特化したハードウェアにより構成されていてもよい。かかる構成は、実施する時々の技術レベルに応じて適宜変更されうる。

40

【0149】

<3.まとめ>

以上説明したように本開示の一実施形態によれば、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定する処理を実行する情報処理装置100が提供される。本開示の一実施形態に係る情報処理装置100は、画像情報に対する処理内容を、ユーザから入力される言語情報の内容に応じて決定する処理を実行することで、画像情報に対する効果的な処理が可能となる。

【0150】

本開示の一実施形態に係る情報処理装置100は、「画面の右下」や「画面の中央やや

50

右上」などの言語的な表現に位置を表現したり、「赤い自動車のすぐ右隣」や「赤い自動車の手前」、「赤い自動車からかなり右側」「テーブルの上」など対象物と目標物の位置関係を言語で表現したりすることが可能になる。

【0151】

本開示の一実施形態に係る情報処理装置100は、相対的な位置関係が入力された場合に、画像認識を行うために複数の物体から任意の物体を認識するためにカメラやレンズを認識対象物の方向に向ける必要がなくなる。本開示の一実施形態に係る情報処理装置100は、ユーザから「テーブルの上に乗っているものは何？」や「太郎さんの右隣にいるのは誰？」などの言語情報が入力された場合に、その内容を適切に解釈し、認識領域を設定することが出来る。

10

【0152】

本開示の一実施形態に係る情報処理装置100は、「リモコンはどこにある？」という問い合わせがユーザから入力された場合、ヘッドマウントディスプレイ10のような眼鏡型のウェアラブルコンピュータなどに対して、画面上で物体のある領域を強調して提示したり、「右にある机の上にあります」というようなテキストや音声で答えたり、認識物体が画面の中央に来た際に音や振動で通知するなどの処理が可能になる。

【0153】

本開示の一実施形態に係る情報処理装置100は、GPS情報や地図情報を用いることにより、ユーザからの「右に見えている大きな建物は何？」や「タワーの左に見える山の名前は？」などの問い合わせに対する回答の生成や、ユーザからの「富士山はどこに見えるの？」などの問い合わせに対する実際の建造物や自然地形の検索の実行が可能になる。従って本開示の一実施形態に係る情報処理装置100は、GPS情報や地図情報を用いることで、ユーザの問い合わせに対する適切な処理を実行することが可能になる。

20

【0154】

本明細書の各装置が実行する処理における各ステップは、必ずしもシーケンス図またはフローチャートとして記載された順序に沿って時系列に処理する必要はない。例えば、各装置が実行する処理における各ステップは、フローチャートとして記載した順序と異なる順序で処理されても、並列的に処理されてもよい。

【0155】

また、各装置に内蔵されるCPU、ROMおよびRAMなどのハードウェアを、上述した各装置の構成と同等の機能を発揮させるためのコンピュータプログラムも作成可能である。また、該コンピュータプログラムを記憶させた記憶媒体も提供されることが可能である。また、機能ブロック図で示したそれぞれの機能ブロックをハードウェアで構成することで、一連の処理をハードウェアで実現することもできる。

30

【0156】

また本開示の一実施形態に係る情報処理装置100は、情報処理装置100の処理の結果として表示される画像を表示するディスプレイを備える装置とは別の装置（たとえばインターネット等のネットワークを介してディスプレイを備える装置と接続されるサーバ装置）として実施されてもよいし、サーバ装置から情報を受ける端末装置で実施されてもよい。また本開示の一実施形態に係る情報処理装置100の構成は、単独の装置で実現されてもよく、複数の装置が連携するシステムで実現されても良い。複数の装置が連携するシステムには、例えば複数のサーバ装置の組み合わせ、サーバ装置と端末装置との組み合わせ等が含まれ得る。

40

【0157】

なお、上記実施形態で示したユーザインタフェースやアプリケーションを実現するソフトウェアは、インターネット等のネットワークを介して使用するウェブアプリケーションとして実現されてもよい。ウェブアプリケーションは、例えば、HTML(HyperText Markup Language)、SGML(Standard Generalized Markup Language)、XML(Extensible Markup Language)などのマークアップ言語により実現されてもよい。

50

【 0 1 5 8 】

以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、特許請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

【 0 1 5 9 】

また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

10

【 0 1 6 0 】

なお、以下のような構成も本開示の技術的範囲に属する。

(1)

入力された言語情報に基づいて画像の中の領域を特定する画像領域特定部と、

前記入力された言語情報に基づいて、前記画像領域特定部が特定した画像の中の領域に対する、前記画像を使用した処理の内容を特定する処理内容特定部と、
を備える、情報処理装置。

(2)

前記処理内容特定部は、前記入力された言語情報に基づいて、前記画像領域特定部が特定した前記画像の中の領域に対する物体の認識処理を実行することを特定する、前記 (1) に記載の情報処理装置。

20

(3)

前記画像領域特定部は、特定した画像の中の領域で認識された物体を基準として、さらに入力された言語情報に基づいて画像の中の領域を特定する、前記 (2) に記載の情報処理装置。

(4)

前記処理内容特定部は、前記入力された言語情報に基づいて、前記画像領域特定部が特定した画像の中の領域に含まれる物体の情報を取得する処理を実行することを特定する、前記 (1) ~ (3) のいずれかに記載の情報処理装置。

(5)

前記処理内容特定部は、前記物体の情報として該物体の名称を取得する処理を実行することを特定する、前記 (4) に記載の情報処理装置。

30

(6)

前記画像領域特定部は、特定した画像の中の領域で情報が取得された物体を基準として、さらに入力された言語情報に基づいて画像の中の領域を特定する、前記 (4) または (5) に記載の情報処理装置。

(7)

前記画像領域特定部が前記画像の中の領域を特定した後に、該特定の取り消しを意味する言語情報が入力されると、前記画像領域特定部は、前記画像に対する直前の領域の特定を解除する、前記 (1) ~ (6) のいずれかに記載の情報処理装置。

40

(8)

前記言語情報として前記画像に対する時間の限定を意味する言語情報が入力されていれば、前記処理内容特定部は、前記画像を使用した処理の対象となる前記画像の時間を限定する、前記 (1) ~ (7) のいずれかに記載の情報処理装置。

(9)

前記処理内容特定部が特定した処理に基づいて、前記画像の物体を認識する物体認識部をさらに備える、前記 (1) ~ (8) のいずれかに記載の情報処理装置。

(1 0)

前記画像処理部が実行した処理の結果を表す言語情報を生成する言語生成部をさらに備える、前記 (9) に記載の情報処理装置。

50

(1 1)

前記処理内容特定部が特定した処理に基づいて、前記画像の物体を認識する処理を前記物体認識部が実行した結果、前記入力された言語情報から物体を一意に特定できない場合は、前記言語生成部は、認識された物体の特徴を付加して言語情報を生成する、前記(10)に記載の情報処理装置。

(1 2)

前記処理内容特定部が特定した処理に基づいて、前記画像の物体を認識する処理を前記物体認識部が実行した結果、前記言語生成部は、前記物体認識部による前記入力された言語情報に基づいた物体認識結果の度合いに応じた言語情報を生成する、前記(10)に記載の情報処理装置。

10

(1 3)

前記画像領域特定部が特定した領域を前記画像に重畳して表示させる認識領域表示部をさらに備える、前記(1)~(12)のいずれかに記載の情報処理装置。

(1 4)

前記画像領域特定部は、入力された言語情報に基づいて前記画像を所定数に分割し、該分割後の各領域を用いて前記画像の中の領域を特定する、前記(1)~(13)のいずれかに記載の情報処理装置。

(1 5)

前記画像領域特定部は、前記入力された言語情報に程度を表す副詞が含まれていれば、前記副詞を参照して前記画像の中の領域を特定する、前記(1)~(14)のいずれかに記載の情報処理装置。

20

(1 6)

入力された言語情報に基づいて画像の中の領域を特定することと、
前記入力された言語情報に基づいて、前記特定された画像の中の領域に対する、前記画像を使用した処理の内容を特定することと、
を含む、情報処理方法。

(1 7)

コンピュータに、
入力された言語情報に基づいて画像の中の領域を特定することと、
前記入力された言語情報に基づいて、前記特定された画像の中の領域に対する、前記画像を使用した処理の内容を特定することと、
を実行させる、コンピュータプログラム。

30

【符号の説明】

【0161】

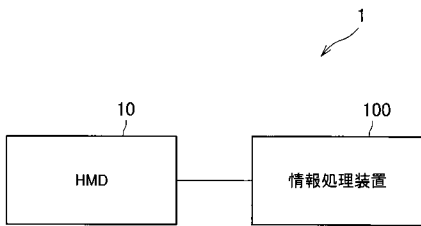
- 10 ヘッドマウントディスプレイ(HMD)
- 100 情報処理装置
- 110 画像処理部
- 111 物体認識部
- 112 画像認識領域設定部
- 113 物体位置情報抽出部
- 114 処理内容特定部
- 120 画像位置言語情報変換部
- 130 言語生成部
- 140 輪郭情報処理部
- 150 RGB色名変換部
- 160 建造物識別部
- 170 言語解析部
- 172 類似画像検索部
- 174 認識領域表示部
- 176 言語画像位置情報変換部

40

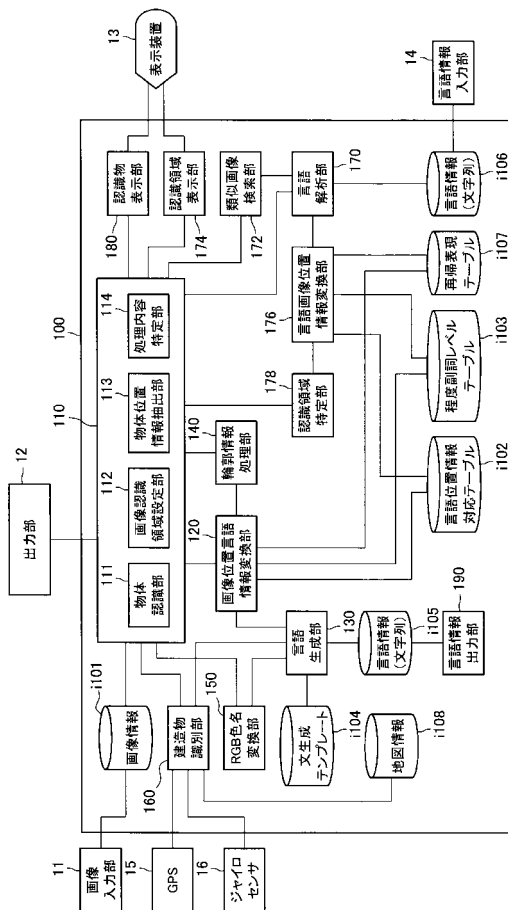
50

- 1 7 8 認識領域特定部
- 1 8 0 認識物表示部
- 1 9 0 言語情報出力部

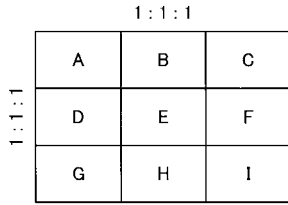
【 図 1 】



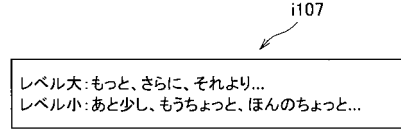
【 図 2 】



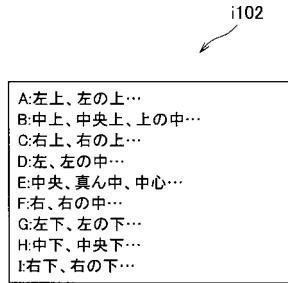
【 図 3 】



【 図 6 】



【 図 4 】



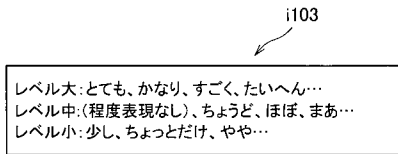
【 図 7 】

	A	B	C	D	E	F	G	H	I
大	ABD	ABC	B,C,F	A,D,G	E	C,F,I	D,G,H	G,H,I	F,H,I
中	C,E,G	D,E,F	A,E,I	B,E,H	全て	B,E,H	A,E,I	D,E,F	C,E,G
小	F,H,I	G,H,I	D,G,H	C,F,I	全て	A,D,G	B,C,F	A,B,C	A,B,D

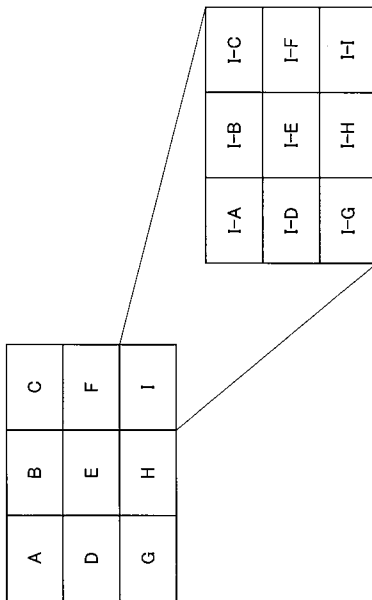
分割領域

程度レベル

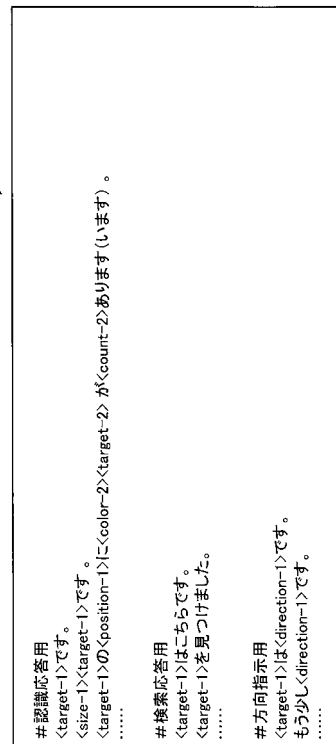
【 図 5 】



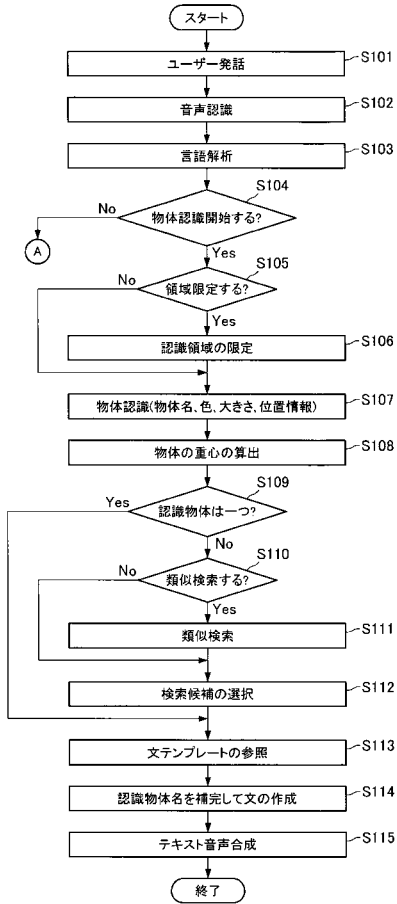
【 図 8 】



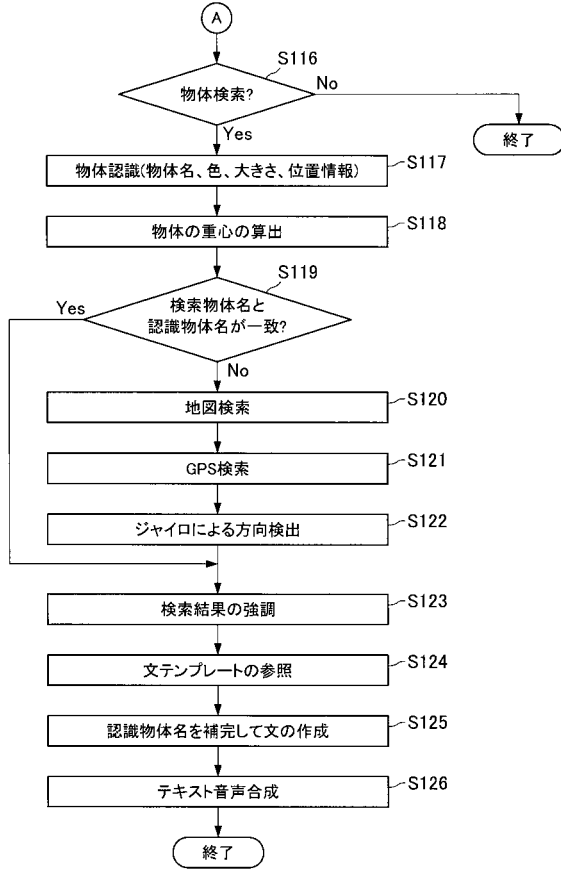
【 図 9 】



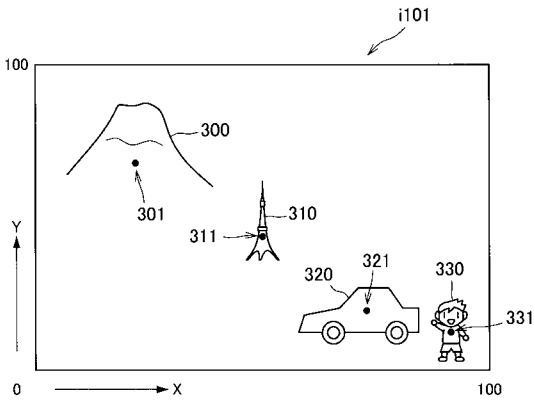
【図10A】



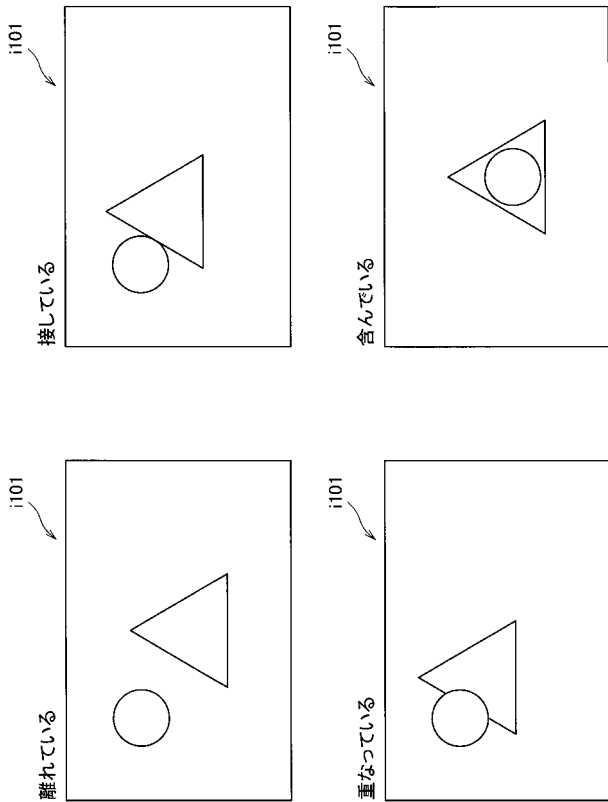
【図10B】



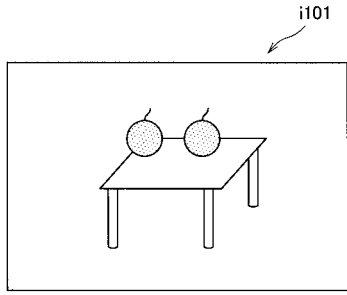
【図11】



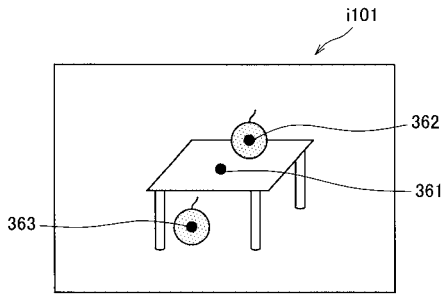
【図12】



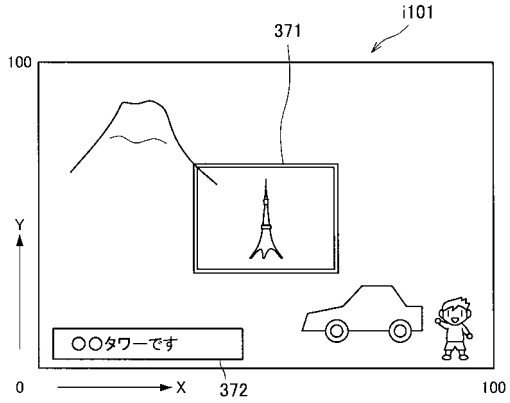
【図 13】



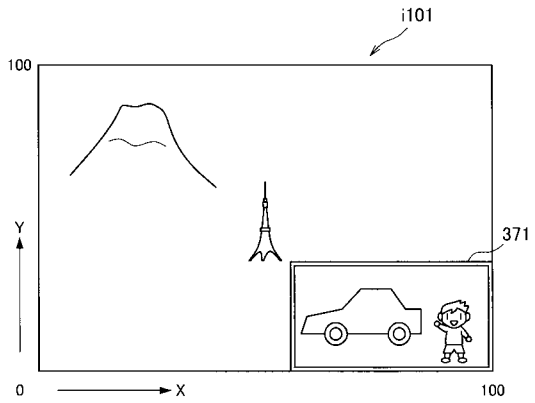
【図 14】



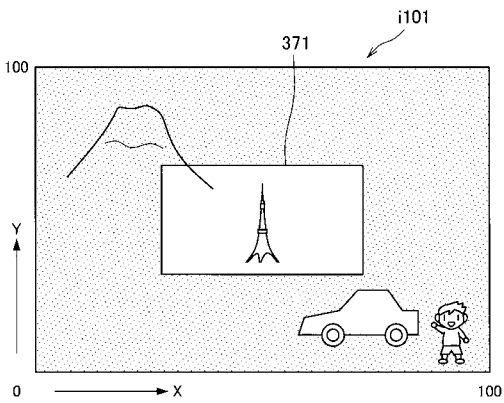
【図 15】



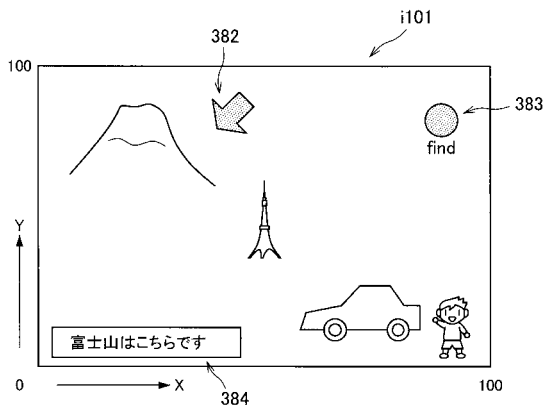
【図 16】



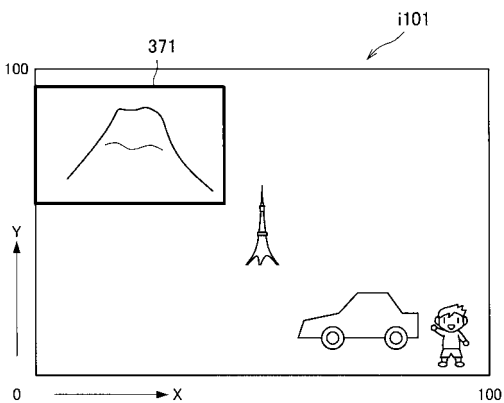
【図 17】



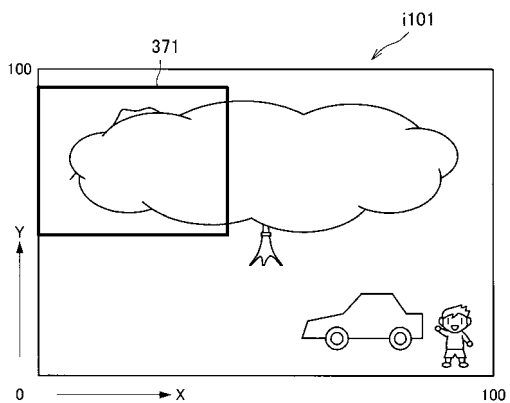
【図 19】



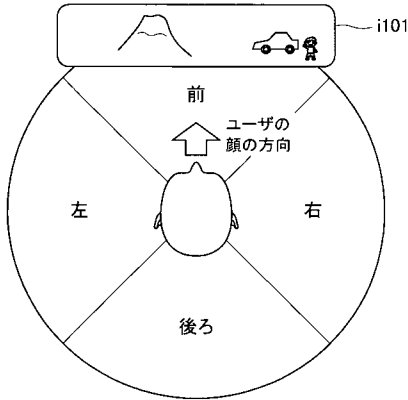
【図 18】



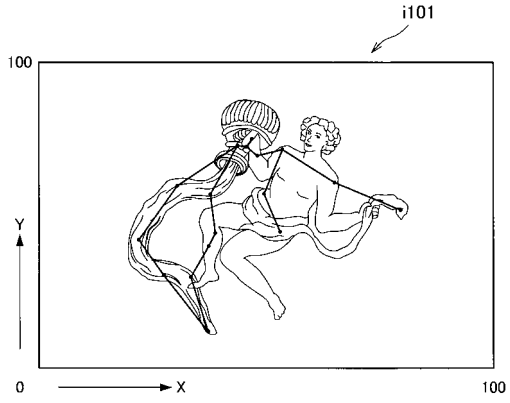
【図 20】



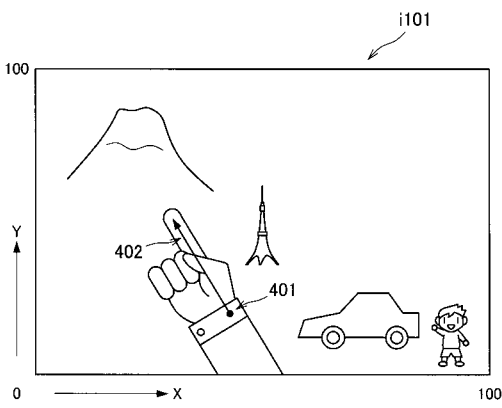
【図 2 1】



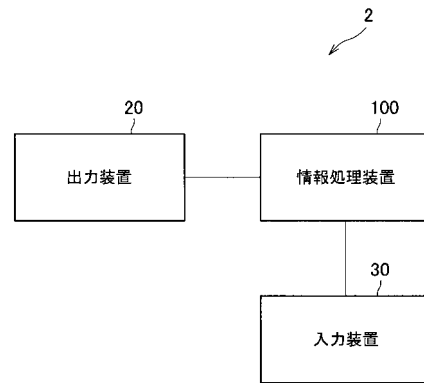
【図 2 3】



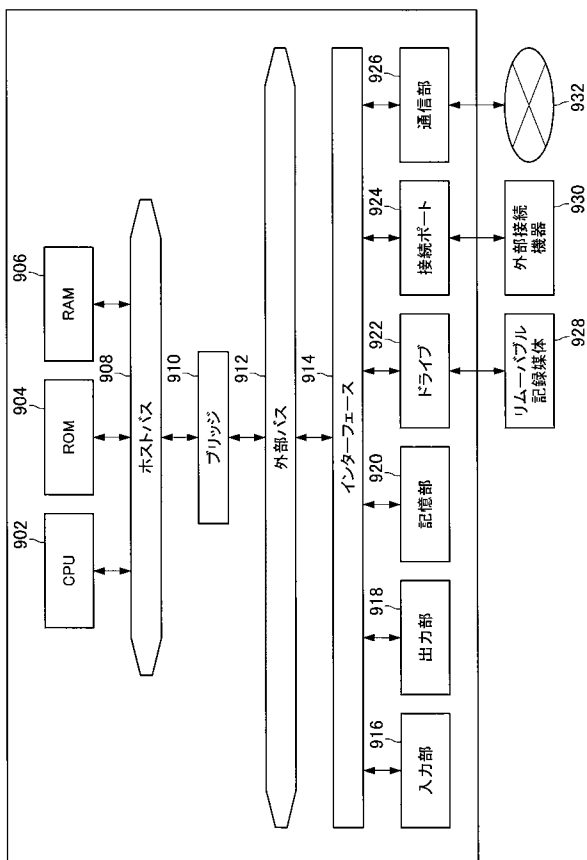
【図 2 2】



【図 2 4】



【図 2 5】



フロントページの続き

(51)Int.Cl. F I テーマコード(参考)
G 0 6 F 3/16 (2006.01) G 0 6 F 3/16 3 2 0 B

Fターム(参考) 5E555 AA11 BA02 BB02 BC04 BC17 BE17 CA42 CA45 CA47 CB64
CB74 CB82 CC01 CC05 DA08 DB41 DB53 DB56 DB57 DC09
DC13 DD07 EA23 FA30
5K127 BA03 CA27 CA34 CB02 CB13 KA04 KA05
5L096 CA24 DA02 FA69 GA19 JA11