

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7559758号  
(P7559758)

(45)発行日 令和6年10月2日(2024.10.2)

(24)登録日 令和6年9月24日(2024.9.24)

(51)国際特許分類	F I
G 0 6 V 10/70 (2022.01)	G 0 6 V 10/70
G 0 6 F 3/04815(2022.01)	G 0 6 F 3/04815
G 0 6 F 3/01 (2006.01)	G 0 6 F 3/01 5 1 0
G 0 6 T 19/00 (2011.01)	G 0 6 T 19/00 6 0 0
G 0 6 V 20/20 (2022.01)	G 0 6 V 20/20

請求項の数 14 (全35頁)

(21)出願番号	特願2021-539168(P2021-539168)	(73)特許権者	000002185 ソニーグループ株式会社 東京都港区港南1丁目7番1号
(86)(22)出願日	令和2年7月7日(2020.7.7)	(74)代理人	100093241 弁理士 宮田 正昭
(86)国際出願番号	PCT/JP2020/026611	(74)代理人	100101801 弁理士 山田 英治
(87)国際公開番号	WO2021/029164	(74)代理人	100086531 弁理士 澤田 俊夫
(87)国際公開日	令和3年2月18日(2021.2.18)	(74)代理人	100095496 弁理士 佐々木 榮二
審査請求日	令和5年5月12日(2023.5.12)	(74)代理人	110000763 弁理士法人大同特許事務所
(31)優先権主張番号	特願2019-147046(P2019-147046)	(72)発明者	後藤 智彦 東京都港区港南1丁目7番1号 ソニー
(32)優先日	令和1年8月9日(2019.8.9)		最終頁に続く
(33)優先権主張国・地域又は機関	日本国(JP)		

(54)【発明の名称】 画像処理装置、および画像処理方法、並びにプログラム

(57)【特許請求の範囲】

【請求項1】

実世界上の実オブジェクトの識別処理を実行するオブジェクト識別部と、  
 実オブジェクトと仮想オブジェクトを重畳表示したAR (Augmented Reality) 画像を生成するコンテンツ表示制御部を有し、  
 前記オブジェクト識別部は、  
 前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行し、  
 前記コンテンツ表示制御部は、  
 前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択する処理を実行し、  
 前記オブジェクト識別部は、  
 オブジェクト種類に応じて予め規定された時間間隔でオブジェクト識別処理を繰り返し実行する画像処理装置。

【請求項2】

前記オブジェクト識別部は、  
 画像認識処理によりオブジェクト識別処理を実行する請求項1に記載の画像処理装置。

【請求項3】

前記オブジェクト識別部は、  
 セマンティック・セグメンテーション処理を適用してオブジェクト識別処理を実行する

請求項 2 に記載の画像処理装置。

【請求項 4】

前記オブジェクト識別部は、

前記仮想オブジェクトの表示領域とするターゲット領域を決定し、決定したターゲット領域の実オブジェクトの識別処理を実行する請求項 1 に記載の画像処理装置。

【請求項 5】

前記オブジェクト識別部は、

前記ターゲット領域をユーザ動作、ユーザ視線、ユーザ操作、ユーザ位置、またはユーザ姿勢の少なくとももいづれかに基づいて決定する請求項 4 に記載の画像処理装置。

【請求項 6】

前記オブジェクト識別部は、

前記ターゲット領域を水平面領域から選択して決定する請求項 4 に記載の画像処理装置。

【請求項 7】

前記コンテンツ表示制御部は、

前記仮想オブジェクトを前記水平面領域に接するように表示する請求項 6 に記載の画像処理装置。

【請求項 8】

前記オブジェクト識別部は、

前記オブジェクト識別処理をリアルタイム処理として実行する請求項 4 に記載の画像処理装置。

【請求項 9】

前記画像処理装置は、

カメラ撮影画像に基づく実世界の 3 次元マップを生成する 3 次元マップ生成部を有し、

前記オブジェクト識別部は、

前記 3 次元マップを利用して前記仮想オブジェクトの表示領域とするターゲット領域を決定する請求項 1 に記載の画像処理装置。

【請求項 10】

前記 3 次元マップ生成部は、

SLAM ( Simultaneous Localization and Mapping ) 処理により、実世界の 3 次元マップを生成する請求項 9 に記載の画像処理装置。

【請求項 11】

前記コンテンツ表示制御部は、

前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するとともに、

前記オブジェクト識別結果に応じて、表示する仮想オブジェクトの表示態様も制御する請求項 1 に記載の画像処理装置。

【請求項 12】

前記画像処理装置は、さらに、

音声出力制御を実行するコンテンツ音声制御部を有し、

前記コンテンツ音声制御部は、

前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて出力する音声を決定して出力する請求項 1 に記載の画像処理装置。

【請求項 13】

画像処理装置において実行する画像処理方法であり、

オブジェクト識別部が、実世界上の実オブジェクトの識別処理を実行するオブジェクト識別処理ステップを実行し、

コンテンツ表示制御部が、実オブジェクトと仮想オブジェクトを重畳表示した AR ( Augmented Reality ) 画像を生成するコンテンツ表示制御ステップを実行し、

前記オブジェクト識別処理ステップは、

10

20

30

40

50

前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行するステップであり、

前記コンテンツ表示制御ステップは、

前記オブジェクト識別処理ステップにおいて識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するステップを実行し、

前記オブジェクト識別処理ステップは、

オブジェクト種類に応じて予め規定された時間間隔でオブジェクト識別処理を繰り返し実行するステップである画像処理方法。

#### 【請求項 14】

画像処理装置において画像処理を実行させるプログラムであり、

オブジェクト識別部に、実世界上の実オブジェクトの識別処理を実行するオブジェクト識別処理ステップを実行させ、

コンテンツ表示制御部に、実オブジェクトと仮想オブジェクトを重畳表示したAR (Augmented Reality) 画像を生成するコンテンツ表示制御ステップを実行させ、

前記オブジェクト識別処理ステップにおいては、

前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行させ、

前記コンテンツ表示制御ステップにおいては、

前記オブジェクト識別処理ステップにおいて識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するステップを実行させ、

前記オブジェクト識別処理ステップにおいては、

オブジェクト種類に応じて予め規定された時間間隔でオブジェクト識別処理を繰り返し実行させるプログラム。

#### 【発明の詳細な説明】

#### 【技術分野】

#### 【0001】

本開示は、画像処理装置、および画像処理方法、並びにプログラムに関する。さらに詳細には、現実に観察可能な実オブジェクトにキャラクタ画像等の仮想コンテンツを重畳して表示する拡張現実 (AR: Augmented Reality) 画像を生成して出力する画像処理装置、および画像処理方法、並びにプログラムに関する。

#### 【背景技術】

#### 【0002】

実空間で観察可能な実オブジェクトや、実オブジェクトと画像に仮想オブジェクトを重畳して表示した画像を拡張現実 (AR: Augmented Reality) 画像と呼ぶ。

#### 【0003】

AR画像を利用したコンテンツやゲームで利用される仮想オブジェクトには様々な種類があるが、例えば人間と同じような動きをする仮想オブジェクト、すなわちキャラクタが利用されることが多い。

#### 【0004】

AR画像は、例えばユーザの目に装着するヘッド・マウント・ディスプレイ (HMD) や、スマホ (スマートフォン) 等の携帯端末等を利用して表示される。

ユーザは、AR画像を視聴することで、例えばAR画像中の表示キャラクタが実世界に存在するような感覚を楽しむことができる。

#### 【0005】

AR画像中にキャラクタを表示させる場合、例えばゲームプログラム等、コンテンツ再生プログラムに従った表示処理がなされる。

具体的には、プログラムに記録されているキャラクタ出力条件を満たす場合にプログラムに規定された手順でキャラクタが表示される。

10

20

30

40

50

しかし、このようなプログラムに規定されたキャラクタ表示を行うと、常に同じような処理が繰り返されることになり、面白みが薄れてしまう。

【0006】

一方、近年、画像内のオブジェクトを識別する技術としてセマンティック・セグメンテーションについての研究や利用が進んでいる。セマンティック・セグメンテーションは、例えばカメラ撮影画像に含まれる様々なオブジェクト、例えば人、車、ビル、道路、木等、オブジェクトの種類を識別する技術である。

【0007】

なお、このセマンティック・セグメンテーションについて開示した文献として、例えば特許文献1（特開2015-207291号公報）がある。

【先行技術文献】

【特許文献】

【0008】

【文献】特開2015-207291号公報

【発明の概要】

【発明が解決しようとする課題】

【0009】

本開示は、実オブジェクトの種類に応じて表示するキャラクタの制御を行う画像処理装置、および画像処理方法、並びにプログラムを提供するものである。

【0010】

本開示の一実施例においては、上述したセマンティック・セグメンテーション等のオブジェクト識別処理によって画像の背景オブジェクトを識別し、識別結果に応じて表示するキャラクタを制御する画像処理装置、および画像処理方法、並びにプログラムを提供するものである。

【課題を解決するための手段】

【0011】

本開示の第1の側面は、  
実世界上の実オブジェクトの識別処理を実行するオブジェクト識別部と、  
実オブジェクトと仮想オブジェクトを重畳表示したAR（Augmented Reality）画像を生成するコンテンツ表示制御部を有し、  
前記オブジェクト識別部は、  
前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行し、  
前記コンテンツ表示制御部は、  
前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択する画像処理装置にある。

【0012】

さらに、本開示の第2の側面は、  
画像処理装置において実行する画像処理方法であり、  
オブジェクト識別部が、実世界上の実オブジェクトの識別処理を実行するオブジェクト識別処理ステップを実行し、  
コンテンツ表示制御部が、実オブジェクトと仮想オブジェクトを重畳表示したAR（Augmented Reality）画像を生成するコンテンツ表示制御ステップを実行し、  
前記オブジェクト識別処理ステップは、  
前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行するステップであり、  
前記コンテンツ表示制御ステップは、  
前記オブジェクト識別処理ステップにおいて識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するステップを実行する画像処理方法にある。

## 【 0 0 1 3 】

さらに、本開示の第 3 の側面は、  
画像処理装置において画像処理を実行させるプログラムであり、  
オブジェクト識別部に、実世界上の実オブジェクトの識別処理を実行するオブジェクト  
識別処理ステップを実行させ、  
コンテンツ表示制御部に、実オブジェクトと仮想オブジェクトを重畳表示した A R ( A  
u g m e n t e d R e a l i t y ) 画像を生成するコンテンツ表示制御ステップを実  
行させ、  
前記オブジェクト識別処理ステップにおいては、  
前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を  
実行させ、  
前記コンテンツ表示制御ステップにおいては、  
前記オブジェクト識別処理ステップにおいて識別されたオブジェクト識別結果に応じて  
表示する仮想オブジェクトを選択するステップを実行させるプログラムにある。

10

## 【 0 0 1 4 】

なお、本開示のプログラムは、例えば、様々なプログラム・コードを実行可能な情報処  
理装置やコンピュータ・システムに対して、コンピュータ可読な形式で提供する記憶媒体  
、通信媒体によって提供可能なプログラムである。このようなプログラムをコンピュータ  
可読な形式で提供することにより、情報処理装置やコンピュータ・システム上でプログラ  
ムに応じた処理が実現される。

20

## 【 0 0 1 5 】

本開示のさらに他の目的、特徴や利点は、後述する本開示の実施例や添付する図面に基  
づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムと  
は、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限ら  
ない。

## 【 0 0 1 6 】

本開示の一実施例の構成によれば、仮想オブジェクトの表示領域とするターゲット領域  
の実オブジェクト種類に応じて表示する仮想オブジェクトの選択や表示態様変更を行う装  
置、方法が実現される。

具体的には、例えば、実世界上の実オブジェクトの識別処理を実行するオブジェクト識  
別部と、実オブジェクトと仮想オブジェクトを重畳表示した A R 画像を生成するコンテ  
ンツ表示制御部を有する。オブジェクト識別部は、仮想オブジェクトの表示領域とするター  
ゲット領域の実オブジェクトを識別し、コンテンツ表示制御部は、オブジェクト識別結果  
に応じて表示する仮想オブジェクトを選択する処理や表示態様の変更処理を行う。

30

本構成により、仮想オブジェクトの表示領域とするターゲット領域の実オブジェクト種  
類に応じて表示する仮想オブジェクトの選択や表示態様変更を行う装置、方法が実現され  
る。

なお、本明細書に記載された効果はあくまで例示であって限定されるものではなく、ま  
た付加的な効果があってもよい。

## 【 図面の簡単な説明 】

40

## 【 0 0 1 7 】

【 図 1 】 本開示の画像処理装置の構成例と実行する処理について説明する図である。

【 図 2 】 本開示の画像処理装置の実行する処理について説明する図である。

【 図 3 】 本開示の画像処理装置の実行する処理について説明する図である。

【 図 4 】 本開示の画像処理装置の実行する処理について説明する図である。

【 図 5 】 本開示の画像処理装置の構成例と実行する処理について説明する図である。

【 図 6 】 本開示の画像処理装置の構成例と実行する処理について説明する図である。

【 図 7 】 本開示の画像処理装置の実行する処理について説明する図である。

【 図 8 】 本開示の画像処理装置の実行する処理について説明する図である。

【 図 9 】 本開示の画像処理装置の一構成例について説明する図である。

50

【図 1 0】空間マップデータのデータ構成例について説明する図である。

【図 1 1】クラス対応更新時間データのデータ構成例について説明する図である。

【図 1 2】クラス対応仮想オブジェクトデータのデータ構成例について説明する図である。

【図 1 3】本開示の画像処理装置の実行する処理のシーケンスについて説明するフローチャートを示す図である。

【図 1 4】本開示の画像処理装置の実行するターゲット領域決定処理のシーケンスについて説明するフローチャートを示す図である。

【図 1 5】本開示の画像処理装置の実行するターゲット領域決定処理のシーケンスについて説明するフローチャートを示す図である。

【図 1 6】本開示の画像処理装置の実行する処理のシーケンスについて説明するフローチャートを示す図である。

10

【図 1 7】本開示の画像処理装置の実行する処理のシーケンスについて説明するフローチャートを示す図である。

【図 1 8】本開示の画像処理装置のハードウェア構成例について説明する図である。

【発明を実施するための形態】

【0018】

以下、図面を参照しながら本開示の画像処理装置、および画像処理方法、並びにプログラムの詳細について説明する。なお、説明は以下の項目に従って行なう。

1. 本開示の画像処理装置の実行する処理の概要について

2. 本開示の画像処理装置の構成例について

20

3. 本開示の画像処理装置の実行する処理のシーケンスについて

3 - (1) 画像処理装置の実行する基本処理シーケンス

3 - (2) ターゲット領域の設定領域をほぼ水平な面に設定する処理のシーケンス

3 - (3) 実オブジェクト識別処理の更新シーケンス

4. 画像処理装置のハードウェア構成例について

5. 本開示の構成のまとめ

【0019】

[ 1. 本開示の画像処理装置の実行する処理の概要について ]

まず、図 1 以下を参照して、本開示の画像処理装置の実行する処理の概要について説明する。

30

【0020】

図 1 には、本開示の画像処理装置の一例として、ヘッド・マウント・ディスプレイ (HMD) 型の光透過型 AR 画像表示デバイス 10 を示している。

ユーザは、ユーザの目を覆うようにヘッド・マウント・ディスプレイ (HMD) 型の光透過型 AR 画像表示デバイス 10 を装着する。

【0021】

光透過型 AR 画像表示デバイス 10 は光透過型の表示部 (ディスプレイ) を有する。光透過型の表示部 (ディスプレイ) をユーザの目の前の位置に設定するようにユーザに装着される。

【0022】

40

ユーザは、光透過型 AR 画像表示デバイス 10 の光透過型の表示部 (ディスプレイ) を介して外部の実オブジェクトをそのまま観察することができる。

さらに、光透過型の表示部 (ディスプレイ) に、仮想オブジェクト、例えばキャラクタ画像等の仮想オブジェクト画像が表示される。

【0023】

ユーザは、光透過型 AR 画像表示デバイス 10 を介して外部の実オブジェクトとキャラクタ等の仮想オブジェクト画像を併せて観察することが可能となり、あたかもキャラクタ等の仮想オブジェクトが実世界状に存在するかのような感覚を体感することができる。

【0024】

図 1 の右側には、ユーザが、光透過型 AR 画像表示デバイス 10 を介して観察可能な画

50

像の例を示している。

【0025】

(a) 観察画像例1は、光透過型AR画像表示デバイス10を介して観察される外部の実オブジェクトからなる透過観察画像21によって構成されている。この画像例1には、仮想オブジェクトは表示されていない。

【0026】

一方、(b) 観察画像例2は、光透過型AR画像表示デバイス10を介して観察される外部の実オブジェクトからなる透過観察画像21に併せて、キャラクタ画像等の仮想オブジェクト画像22を表示した画像例である。この画像例2は、ユーザが実オブジェクトと仮想オブジェクトを併せて観察可能な画像である。

10

【0027】

前述したように、実空間で観察可能な実オブジェクトや、実オブジェクトと画像に仮想オブジェクトを重畳して表示した画像を拡張現実(AR: Augmented Reality)画像と呼ぶ。

本開示の画像処理装置は、このAR画像の表示処理を行う装置である。

【0028】

本開示の画像処理装置、例えば図1に示す光透過型AR画像表示デバイス10は、AR画像表示における仮想オブジェクトの表示制御を実行する。

【0029】

本開示の画像処理装置、例えば図1に示す光透過型AR画像表示デバイス10の実行する具体的な処理は、例えば以下の処理である。

20

(a) 自己位置推定と環境3次元マップ生成を行うSLAM(Simultaneous Localization and Mapping)処理等を適用して、ユーザが光透過型AR画像表示デバイス10を介して観察している実世界の3次元マップを生成する処理。

(b) セマンティック・セグメンテーション等のオブジェクト識別処理によって実世界に含まれるオブジェクトを識別する処理。

(c) 実世界のオブジェクト識別結果に応じて、表示する仮想オブジェクト(キャラクタ等)を選択し、また仮想オブジェクトの表示態様を制御する処理。

本開示の画像処理装置は、例えばこれらの処理を実行する。

30

【0030】

図2以下を参照して、本開示の画像処理装置が実行する仮想オブジェクトの表示制御例について説明する。

【0031】

図2の左に示す図は、光透過型AR画像表示デバイス10を装着したユーザが公園を歩き回り、池の近くで、池の水面を見ながら、池の水面方向に「指さし」を行った状態を示している。

【0032】

光透過型AR画像表示デバイス10は、カメラを有しており、ユーザが指さしを行った画像を撮影し、撮影画像がデバイス内部の画像解析部を介して3次元マップ生成部に入力される。

40

画像解析部は撮影画像から特徴点を抽出し、3次元マップ生成部は、画像から抽出された特徴点を利用して実世界の3次元マップを生成する。

この3次元マップ生成処理は、例えば、SLAM(Simultaneous Localization and Mapping)処理によってリアルタイムで実行される。

【0033】

SLAM(Simultaneous Localization and Mapping)処理は、自己位置推定と環境3次元マップの生成を同時に並列して実行することを可能とした処理である。

【0034】

50

さらに、3次元マップ生成部が生成した3次元マップは、画像処理装置のオブジェクト識別部に入力される。

オブジェクト識別部は、3次元マップ生成部が生成した3次元マップを用いて、ユーザの指さし方向にある実オブジェクト領域をターゲット領域11として決定する。さらに、このターゲット領域11のオブジェクトを識別する。

【0035】

オブジェクト識別部は、例えば、セマンティック・セグメンテーション処理を適用して実世界のオブジェクトの識別処理を行う。

【0036】

セマンティック・セグメンテーションは、画像認識処理の一種であり、ディープラーニングを用いて画像中の物体が画素レベルで何かを認識するタスクを行う手法である。例えば様々な実際のオブジェクトの形状情報やその他の特徴情報を登録したオブジェクト識別用の辞書データ(学習済みデータ)と、例えばカメラ撮影画像内のオブジェクトとの一致度に基づいて、画像の構成画素(ピクセル)各々が、どのオブジェクトカテゴリに属する画素であるかを識別する技術である。

10

このセマンティック・セグメンテーションにより、カメラ撮影画像に含まれる様々なオブジェクト、例えば人、車、ビル、道路、木、池、芝生等、オブジェクトの種類を識別することができる。

【0037】

図2に示す例では、光透過型AR画像表示デバイス10の画像解析部は、カメラ撮影画像に基づいて、ユーザの指さし方向のターゲット領域11のオブジェクトが「池」であると判定する。

20

【0038】

さらに、光透過型AR画像表示デバイス10の仮想オブジェクト表示制御を実行するコンテンツ表示制御部は、ターゲット領域11のオブジェクト識別結果を入力し、オブジェクト識別結果に従って表示すべき仮想オブジェクト(キャラクタ等)の選択処理や表示態様を決定して表示する。

【0039】

図2に示す例では、図2右側の(b)観察画像例2に示すように、仮想オブジェクト画像22として、「水の妖精キャラクタ」の画像を表示する制御を行う。

30

これは、ユーザの指さし方向の実オブジェクトが、「池」であるとの解析結果に基づく表示制御である。

すなわち、コンテンツ表示制御部は、ターゲット領域11のオブジェクト識別結果=「池」に応じた最適な仮想オブジェクトとして、「水の妖精キャラクタ」を選択して表示する処理を行う。

なお、仮想オブジェクト画像22は、例えば3Dコンテンツ画像として表示する。

【0040】

図3は、本開示の画像処理装置が実行する異なる仮想オブジェクトの表示制御例を説明する図である。

図3の左に示す図も、図2と同様、光透過型AR画像表示デバイス10を装着したユーザが公園を歩き回っている図である。ユーザは、公園の芝生を見ながら、芝生方向に指さしを行っている。

40

【0041】

この場合、光透過型AR画像表示デバイス10のオブジェクト識別部は、カメラ撮影画像の解析により、ユーザの指さし方向であるターゲット領域11のオブジェクトが「芝生」であるとの解析結果を出力する。

【0042】

さらに、光透過型AR画像表示デバイス10のコンテンツ表示制御部は、ターゲット領域11のオブジェクト識別結果を入力し、オブジェクト識別結果に従って表示すべき仮想オブジェクト(キャラクタ等)の選択処理や表示態様を決定して表示する。

50

## 【 0 0 4 3 】

図 3 に示す例では、図 3 右側の ( b ) 観察画像例 2 に示すように、仮想オブジェクト画像 2 3 として、「草原の妖精キャラクター」の画像を表示する制御を行う。

これは、ユーザの指さし方向の実オブジェクトが、「芝生」であるとの解析結果に基づく表示制御である。

すなわち、コンテンツ表示制御部は、ターゲット領域 1 1 のオブジェクト識別結果 = 「芝生」に応じた最適な仮想オブジェクトとして、「芝生の妖精キャラクター」を選択して表示する処理を行う。

## 【 0 0 4 4 】

図 4 は、本開示の画像処理装置が実行する異なる仮想オブジェクトの表示制御例を説明する図である。

10

図 4 の左に示す図も、図 2、図 3 と同様、光透過型 A R 画像表示デバイス 1 0 を装着したユーザが公園を歩き回っている図である。ユーザは、公園の木を見ながら、木の方向に指さしを行っている。

## 【 0 0 4 5 】

この場合、光透過型 A R 画像表示デバイス 1 0 のオブジェクト識別部は、カメラ撮影画像の解析により、ユーザの指さし方向であるターゲット領域 1 1 のオブジェクトが「木」であるとの解析結果を出力する。

## 【 0 0 4 6 】

さらに、光透過型 A R 画像表示デバイス 1 0 のコンテンツ表示制御部は、ターゲット領域 1 1 のオブジェクト識別結果を入力し、オブジェクト識別結果に従って表示すべき仮想オブジェクト（キャラクター等）の選択処理や表示態様を決定して表示する。

20

## 【 0 0 4 7 】

図 4 に示す例では、図 4 右側の ( b ) 観察画像例 2 に示すように、仮想オブジェクト画像 2 3 として、「木の妖精キャラクター」の画像を表示する制御を行う。

これは、ユーザの指さし方向の実オブジェクトが、「木」であるとの解析結果に基づく表示制御である。

すなわち、コンテンツ表示制御部は、ターゲット領域 1 1 のオブジェクト識別結果 = 「木」に応じた最適な仮想オブジェクトとして、「木の妖精キャラクター」を選択して表示する処理を行う。

30

## 【 0 0 4 8 】

このように、本開示の画像処理装置は、S L A M 処理等を用いて実世界の 3 次元形状解析を行い、実世界の 3 次元マップを生成し、さらにセマンティック・セグメンテーション等のオブジェクト識別処理によって実世界の 3 次元マップ内のターゲット領域のオブジェクトを識別し、識別結果に応じて表示するキャラクター等の仮想オブジェクトの表示制御を実行する。

## 【 0 0 4 9 】

なお、セマンティック・セグメンテーション等のオブジェクト識別処理によって解析する実オブジェクトの対象は、例えば、ユーザの指によって指定された領域、すなわちターゲット領域のみに限定して処理を行うことが可能となる。このように、解析範囲を限定することにより高速処理が実現される。

40

## 【 0 0 5 0 】

なお、本開示の画像処理装置は、図 1 を参照して説明したヘッド・マウント・ディスプレイ ( H M D ) 型の光透過型 A R 画像表示デバイス 1 0 に限らず、様々な表示部を備えた装置によって構成することができる。

## 【 0 0 5 1 】

例えば、図 5 に示すカメラ撮影画像表示型 A R 画像表示デバイス 3 0 としてもよい。図 5 に示すカメラ撮影画像表示型 A R 画像表示デバイス 3 0 は、非透過型の表示部 ( ディスプレイ ) を有する。光透過型の表示部 ( ディスプレイ ) は、ユーザによって観察可能なユーザの目の前の位置に設定されるようにユーザに装着される。

50

## 【 0 0 5 2 】

カメラ撮影画像表示型 A R 画像表示デバイス 3 0 に一体化されたカメラ 3 1 の撮影画像、すなわち図 5 に示すカメラ撮影画像 3 2 がユーザの目の前の表示部に表示される。すなわちカメラ 3 1 が撮影する実オブジェクトの画像がユーザの目の前の表示部に表示され、ユーザは、カメラ撮影画像 3 2 を見ることで、外部の景色を確認することができる。

## 【 0 0 5 3 】

さらに、表示部（ディスプレイ）には、仮想オブジェクト、例えばキャラクタ画像等の仮想オブジェクト画像 2 2 が表示される。

## 【 0 0 5 4 】

ユーザは、カメラ撮影画像表示型 A R 画像表示デバイス 3 0 の表示部（ディスプレイ）に表示されるカメラ撮影画像 3 2、すなわち実オブジェクト画像と、キャラクタ等の仮想オブジェクト画像 2 2 を併せて観察することが可能となり、あたかもキャラクタ等の仮想オブジェクトが実世界状に存在するかのような感覚を体感することができる。

## 【 0 0 5 5 】

さらに、本開示の画像処理装置は、図 6 に示すようなスマホ 4 0 等の携帯型表示装置とすることも可能である。

## 【 0 0 5 6 】

図 6 に示すスマホ 4 0 は、表示部を有し、カメラ 4 1 を有する。カメラ 4 1 の撮影画像、すなわち図に示すカメラ撮影画像 4 2 が表示部に表示される。すなわちカメラ 4 1 が撮影する実オブジェクトの画像が表示部に表示され、ユーザは、カメラ撮影画像を見ることで、外部の景色を確認することができる。

## 【 0 0 5 7 】

さらに、表示部（ディスプレイ）には、仮想オブジェクト、例えばキャラクタ画像等の仮想オブジェクト画像が表示される。

## 【 0 0 5 8 】

ユーザは、スマホ 4 0 の表示部（ディスプレイ）に表示されるカメラ撮影画像、すなわち実オブジェクト画像と、キャラクタ等の仮想オブジェクト画像を併せて観察することが可能となり、あたかもキャラクタ等の仮想オブジェクトが実世界状に存在するかのような感覚を体感することができる。

## 【 0 0 5 9 】

なお、このスマホ 4 0 の例では、ユーザがスマホ 4 0 の表示部のある位置をタッチした場合に、画像処理装置（スマホ）の画像解析部が、タッチ位置を解析し、さらにタッチ位置の実オブジェクトの種類を判別する。その後、画像処理装置（スマホ）のコンテンツ表示制御部が、この判別結果に応じてキャラクタ等の仮想オブジェクトの表示制御を実行する。

## 【 0 0 6 0 】

上述したように、本開示の画像処理装置は、仮想オブジェクトの表示位置となるターゲット領域にある実オブジェクトが水であるか、草であるか、木であるか等のオブジェクト識別を行い、識別結果に応じて表示するキャラクタ等の仮想オブジェクトを選択して表示する処理等を行う。

## 【 0 0 6 1 】

さらに、本開示の画像処理装置は、ターゲット領域にある実オブジェクトの識別結果に応じた表示仮想オブジェクトの選択処理のみならず、実オブジェクト識別結果に応じてキャラクタ等の仮想オブジェクトの表示態様を変更する処理も実行する。

## 【 0 0 6 2 】

図 7 以下を参照して、実オブジェクト識別結果に応じた仮想オブジェクトの表示態様の変更処理の具体例について説明する。

## 【 0 0 6 3 】

図 7 には、本開示の画像処理装置の表示部に表示される A R 画像の一例として、実オブジェクトである池から仮想オブジェクト（キャラクタ）が飛び出す A R 画像の表示例を示

10

20

30

40

50

している。

【 0 0 6 4 】

本開示の画像処理装置のコンテンツ表示制御部は、ターゲット領域に表示したキャラクタ等の仮想オブジェクトを予め設定されたプログラムに従って移動させる処理や動かす処理を行う。

図 7 に示す例は、ターゲット領域 1 1 に表示したキャラクタを上方向に移動させる表示を行う例を示している。

【 0 0 6 5 】

図 7 ( 1 ) 池から仮想オブジェクト ( キャラクタ ) が飛び出す瞬間 - 1 は、仮想オブジェクト ( キャラクタ ) の上半分のみが水上に表示され、下半分が水中にある状態である。

10

このような状態では、画像処理装置のコンテンツ表示制御部は、図 7 ( 1 ) に示すように、水上の仮想オブジェクト画像 5 0 と、水中の仮想オブジェクト画像 5 1 を異なる表示態様で表示する。

すなわち、水上の仮想オブジェクト画像 5 0 については、通常の輪郭がはっきりした画像として表示するが、水中の仮想オブジェクト画像 5 1 については、水中に存在するように 3 次元的な歪みを持つ画像として表示する。

【 0 0 6 6 】

さらに、画像処理装置のコンテンツ音声制御部は、キャラクタが水上に移動する際の効果音として、水の音 ( ピチャピチャ等 ) を、スピーカを介して出力する。

【 0 0 6 7 】

20

図 7 ( 2 ) 池から仮想オブジェクト ( キャラクタ ) が飛び出す瞬間 - 2 は、図 7 ( 1 ) の後の A R 画像表示例である。この時点では、仮想オブジェクト ( キャラクタ ) の全身が水上に表示されている。この状態では、水上の仮想オブジェクト画像 5 0 がキャラクタ全体の画像となり、画像処理装置のコンテンツ表示制御部は、キャラクタ全体画像を輪郭がはっきりした画像として表示する。

【 0 0 6 8 】

図 8 は、実オブジェクトの識別結果に応じてキャラクタ等の仮想オブジェクト表示態様を変更する処理の別の具体例について説明する図である。

【 0 0 6 9 】

図 8 には、本開示の画像処理装置の表示部に表示される A R 画像の一例として、仮想オブジェクト ( キャラクタ ) の影を表示した A R 画像の表示例を示している。

30

すなわち、コンテンツ表示制御部がターゲット領域 1 1 にキャラクタを表示する際、キャラクタの影の表示も併せて実行する。

【 0 0 7 0 】

図 8 に示す「 ( 1 ) 影の映る面が平面の場合のキャラクタの影の表示例」は、仮想オブジェクト ( キャラクタ ) 画像 5 0 の影の映る面が、室内の床や、外部の歩道等、平面である場合の影の表示例である。

このように仮想オブジェクト ( キャラクタ ) 画像 5 0 の影の映る面が平面である場合、画像処理装置のコンテンツ表示制御部は、ターゲット領域 1 1 に 3 次元キャラクタである仮想オブジェクト画像 5 0 を表示する際、仮想オブジェクト画像 5 0 の影を示す仮想オブジェクト影画像 5 2 を輪郭がはっきりした画像として表示する。

40

【 0 0 7 1 】

一方、図 8 「 ( 2 ) 影の映る面が非平面 ( 砂場等 ) の場合のキャラクタの影の表示例は」、

仮想オブジェクト ( キャラクタ ) 画像 5 0 の影の映る面が、例えば砂場等、平面でない場合の影の表示例である。

このように仮想オブジェクト ( キャラクタ ) 画像 5 0 の影の映る面が砂場等、でこぼこのある面の場合、画像処理装置のコンテンツ表示制御部は、ターゲット領域 1 1 に 3 次元キャラクタである仮想オブジェクト画像 5 0 を表示する際、仮想オブジェクト画像 5 0 の影を示す仮想オブジェクト影画像 5 2 を輪郭がはっきりしない、でこぼこのある画像とし

50

て表示する。

【 0 0 7 2 】

このように、本開示の画像処理装置のコンテンツ表示制御部は、仮想オブジェクトを表示するターゲット領域の実オブジェクトの識別結果に応じて、仮想オブジェクトの表示態様も変更して表示する制御を行い、また、コンテンツ音声制御部はターゲット領域の実オブジェクトの識別結果に応じた効果音の出力制御を行う。

【 0 0 7 3 】

[ 2 . 本開示の画像処理装置の構成例について ]

次に、本開示の画像処理装置の構成例について説明する。

【 0 0 7 4 】

前述したように、本開示の画像処理装置は、図 1 を参照して説明した光透過型 A R 画像表示デバイス 1 0、あるいは図 5 を参照して説明したカメラ撮影画像表示型 A R 画像表示デバイス 3 0、あるいは図 6 を参照して説明したスマホ 4 0 等の携帯型の表示装置等、様々な形態を持つ装置として実現することができる。

【 0 0 7 5 】

図 9 は、これら様々な形態を取り得る本開示の画像処理装置の一構成例を示すブロック図である。

図 9 に示す画像処理装置 1 0 0 の構成について説明する。

【 0 0 7 6 】

図 9 に示すように、画像処理装置 1 0 0 は、データ入力部 1 1 0、データ処理部 1 2 0、データ出力部 1 3 0、通信部 1 4 0 を有する。

【 0 0 7 7 】

データ入力部 1 1 0 は、外部撮影カメラ 1 1 1、内部撮影カメラ 1 1 2、モーションセンサ（ジャイロ、加速度センサ等） 1 1 3、操作部 1 1 4、マイク 1 1 5 を有する。

【 0 0 7 8 】

データ処理部 1 2 0 は、外部撮影画像解析部 1 2 1、3次元マップ生成部 1 2 2、内部撮影画像解析部 1 2 3、デバイス姿勢解析部 1 2 4、音声解析部 1 2 5、オブジェクト識別部 1 2 6、空間マップデータ 1 2 7、クラス対応更新時間データ 1 2 8 を有する。

【 0 0 7 9 】

データ出力部 1 3 0 は、コンテンツ表示制御部 1 3 1、コンテンツ音声制御部 1 3 2、表示部 1 3 3、スピーカ 1 3 4、クラス対応仮想オブジェクトデータ（3Dモデル、音声データ等） 1 3 5 を有する。

【 0 0 8 0 】

データ入力部 1 1 0 の外部撮影カメラ 1 1 1 は、外部の画像を撮影する。例えば、HMD を装着したユーザのいる環境における外の景色等を撮影する。スマホ等の携帯端末の場合はスマホ等に備えられたカメラである。

内部撮影カメラ 1 1 2 は、基本的に HMD 固有の構成であり、ユーザの視線方向を解析するためのユーザの目の領域の画像を撮影する。

【 0 0 8 1 】

モーションセンサ（ジャイロ、加速度センサ等） 1 1 3 は、例えば HMD やスマホ等、画像処理装置 1 0 0 本体の姿勢や動きを検出する。

モーションセンサ 1 1 3 は、例えば、ジャイロ、加速度センサ、方位センサ、一測位センサ、慣性計測装置（IMU）等によって構成される。

【 0 0 8 2 】

操作部 1 1 4 は、ユーザによって操作可能な操作部であり、例えばターゲット領域の入力や、その他の処理指示等の入力に利用される。

マイク 1 1 5 は、ユーザによる音声入力による指示の入力等に用いられる。また、外部の環境音の入力にも利用できる。

【 0 0 8 3 】

次に、データ処理部 1 2 0 の構成部について説明する。

10

20

30

40

50

外部撮影画像解析部 1 2 1 は、外部撮影カメラ 1 1 1 によって撮影された外部の撮影画像を入力し、外部撮影画像から特徴点の抽出を行う。

この特徴点抽出処理は、3次元マップの生成に利用するための特徴点であり、抽出した特徴点情報は、外部撮影カメラ 1 1 1 によって撮影された外部の撮影画像とともに、3次元マップ生成部 1 2 2 に入力する。

【0084】

3次元マップ生成部 1 2 2 は、外部撮影カメラ 1 1 1 によって撮影された外部の撮影画像と、外部撮影画像解析部 1 2 1 が抽出した特徴点に基づいて、外部の実オブジェクトによって構成される3次元マップを生成する。

【0085】

この3次元マップ生成処理は、例えば、SLAM ( Simultaneous Localization and Mapping ) 処理によってリアルタイム処理として実行する。

前述したように、SLAM ( Simultaneous Localization and Mapping ) 処理は、自己位置推定と環境3次元マップの生成を同時に並列して実行することを可能とした処理である。

【0086】

3次元マップ生成部 1 2 2 の生成した外部環境の3次元マップデータは、オブジェクト識別部 1 2 6 に入力される。

【0087】

内部撮影画像解析部 1 2 3 は、内部撮影カメラ 1 1 2 の撮影したユーザの目の領域の画像に基づいてユーザの視線方向を解析する。この内部撮影画像解析部 1 2 3 も、前述した内部撮影カメラ 1 1 2 と同様、基本的にHMD固有の構成である。

内部撮影画像解析部 1 2 3 の解析したユーザ視線情報は、オブジェクト識別部 1 2 6 に入力される。

【0088】

デバイス姿勢解析部 1 2 4 は、モーションセンサ ( ジャイロ、加速度センサ等 ) 1 1 3 によって計測されたセンサ検出情報に基づいて、HMDやスマホ等、画像処理装置 1 0 0 本体の姿勢や動きを解析する。

デバイス姿勢解析部 1 2 4 の解析した画像処理装置 1 0 0 本体の姿勢や動き情報は、オブジェクト識別部 1 2 6 に入力される。

【0089】

音声解析部 1 2 5 は、マイク 1 1 5 から入力するユーザ音声や環境音の解析を行う。解析結果は、オブジェクト識別部 1 2 6 に入力される。

【0090】

オブジェクト識別部 1 2 6 は、3次元マップ生成部 1 2 2 が生成した3次元マップを入力し、仮想オブジェクトの表示領域に設定するためのターゲット領域を決定し、さらに、決定したターゲット領域の実オブジェクトの識別処理を実行する。例えばターゲット領域が池である、木である等のオブジェクト識別処理を実行する。

【0091】

ターゲット領域の特定処理は、様々な方法によって実行可能である。

例えば、3次元マップに含まれるユーザの指の画像を用いて行うことが可能である。

ユーザの指さし方向の延長線と、3次元マップ上の実オブジェクトとの交点を求め、例えば、交点を中心とするあらかじめ規定した半径の円形領域をターゲット領域として決定する。

【0092】

なお、ターゲット領域の指定は、ユーザによる指さし以外の方法で行うことも可能である。オブジェクト識別部 1 2 6 は、ターゲット領域の決定に使用するための情報として、以下のいずれかの情報を利用することができる。

( a ) 内部撮影画像解析部 1 2 3 の解析したユーザ視線情報

10

20

30

40

50

- (b) デバイス姿勢解析部 124 の解析した画像処理装置 100 本体の姿勢や動き情報
- (c) 操作部 114 を介して入力するユーザ操作情報
- (d) 音声解析部 125 が解析したユーザ音声情報

## 【0093】

「(a) 内部撮影画像解析部 123 の解析したユーザ視線情報」を利用する場合、オブジェクト識別部 126 は、ユーザの視線方向の延長線と、3次元マップ上の実オブジェクトとの交点を求め、例えば、交点を中心とする所定半径の円形領域をターゲット領域として決定する。

## 【0094】

「(b) デバイス姿勢解析部 124 の解析した画像処理装置 100 本体の姿勢や動き情報」を利用する場合、オブジェクト識別部 126 は、ユーザによって装着された HMD、あるいはユーザの持つスマホの前方方向の延長線と、3次元マップ上の実オブジェクトとの交点を求め、例えば、交点を中心とする所定半径の円形領域をターゲット領域として決定する。

## 【0095】

「(c) 操作部 114 を介して入力するユーザ操作情報」を利用する場合、オブジェクト識別部 126 は、例えば画像処理装置 100 の入力部を介して入力されたユーザの操作情報に基づいてターゲット領域を決定する。

例えば、先に説明した図 6 に示すスマホを利用した構成において、ユーザの指による画面位置指定情報をユーザ操作情報として入力し、この指定位置をターゲット領域の中心位置とするといった処理が可能である。

## 【0096】

なお、この他、操作部 114 として、画像処理装置 100 と別体の棒状の指示部材を用い、この指示部材の指示方向情報をオブジェクト識別部 126 に入力し、指示方向に基づいてターゲット領域を決定する構成としてもよい。

## 【0097】

「(d) 音声解析部 125 が解析したユーザ音声情報」を利用する場合、オブジェクト識別部 126 は、例えばユーザによる発話を解析してターゲット領域を決定する。

例えば、ユーザ発話が「前の池」といった発話である場合、その色をターゲット領域として決定する。

## 【0098】

さらに、オブジェクト識別部 126 は、これら以外のターゲット領域決定処理を行ってもよい。例えば、外部撮影カメラ 111 によって撮影された画像に基づいて生成される 3次元マップや、モーションセンサ 113 の検出情報に基づいて地面、床面、水面等の水平面の検出処理を実行し、撮影画像の中心領域に最も近い水平面の領域をターゲット領域として決定する処理を行ってもよい。

## 【0099】

また、例えばユーザが仮想的なボールを投げる動作を行い、これを外部撮影カメラ 111 によって撮影し、撮影画像を解析してボールの着地点を解析して、この着地点をターゲット領域の中心位置に設定するといった処理を行ってもよい。

## 【0100】

オブジェクト識別部 126 は、上述した方法のいずれかを用いて、仮想オブジェクト表示領域とするターゲット領域を決定する。さらに、決定したターゲット領域の実オブジェクトの識別処理を実行する。例えばターゲット領域が池である、木である等のオブジェクト識別処理を実行する。

## 【0101】

前述したように、この実オブジェクトのオブジェクト識別処理は、例えば、セマンティック・セグメンテーション処理を適用して実行する。

セマンティック・セグメンテーションは、様々な実際のオブジェクトの形状情報やその他の特徴情報を登録したオブジェクト識別用の辞書データ(学習済みデータ)と、例えば

10

20

30

40

50

カメラ撮影画像内のオブジェクトとの一致度に基づいて、画像の構成画素（ピクセル）各々が、どのオブジェクトカテゴリに属する画素であるかを識別する技術である。

このセマンティック・セグメンテーションにより、カメラ撮影画像に含まれる様々なオブジェクト、例えば人、車、ビル、道路、木、池、芝生等、オブジェクトの種類を識別することができる。

#### 【0102】

なお、オブジェクト識別部126において実行する実オブジェクトの識別処理は、ターゲット領域のみ、あるいはターゲット領域を含む周囲領域の限られた範囲のみを対象として実行する。このような限定範囲の処理を行うことで高速の処理、すなわちリアルタイム処理が可能となる。なお、リアルタイム処理とは、例えばユーザによるターゲット領域の指定後、即座に実オブジェクトの識別処理が実行されることを意味する。これにより、例えばユーザがターゲット領域を観察している間に時間遅れなくオブジェクト識別が完了する。

10

#### 【0103】

オブジェクト識別部126が解析したターゲット領域の実オブジェクトの識別結果は、データ出力部130のコンテンツ表示制御部131と、コンテンツ音声制御部132に入力される。

#### 【0104】

データ出力部130のコンテンツ表示制御部131は、オブジェクト識別部126からターゲット領域のオブジェクト識別結果を入力し、オブジェクト識別結果に従って表示すべき仮想オブジェクト（キャラクタ等）の選択処理や表示態様を決定して表示部133に表示する。

20

具体的には、例えば先に説明した図2～図4、図7、図8に示すような仮想オブジェクト（キャラクタ等）の表示処理を実行する。

#### 【0105】

データ出力部130のコンテンツ音声制御部132は、オブジェクト識別部126からターゲット領域のオブジェクト識別結果を入力し、オブジェクト識別結果に従って出力すべき音声を決定してスピーカ134を介して出力する。

具体的には、例えば先に説明した図7に示すように、実オブジェクトである池から仮想オブジェクトが現れる場合、水の音を出力する処理を実行する。

30

#### 【0106】

なお、データ出力部130のコンテンツ表示制御部131と、コンテンツ音声制御部132は、クラス対応仮想オブジェクトデータ135に記録された仮想オブジェクトの3Dコンテンツや様々な音声データを取得してデータ出力を実行する。

#### 【0107】

クラス対応仮想オブジェクトデータ135には、ターゲット領域の実オブジェクトの識別結果に相当する実オブジェクト種類（クラス）に対応づけられた表示用の仮想オブジェクトの3Dコンテンツや様々な音声データが記録されている。

クラス対応仮想オブジェクトデータ135の具体例については後述する。

#### 【0108】

40

また、コンテンツ表示部131は、画像処理装置100が、例えば図5を参照して説明したカメラ撮影画像表示型AR画像表示デバイス30や、図6を参照して説明したスマホ40のようにカメラ撮影画像を表示する処理を行う構成の場合は、外部撮影カメラ11の撮影画像を入力して、この撮影画像に仮想オブジェクトを重畳させた表示画像を生成して表示部133に表示する。

#### 【0109】

通信部140は、例えば外部サーバと通信し、仮想コンテンツであるキャラクタの3Dコンテンツを取得する。その他、データ処理に必要な様々なデータやパラメータを外部サーバから取得する構成としてもよい。

#### 【0110】

50

なお、オブジェクト識別部 126 は、ターゲット領域の実オブジェクトの識別処理を行う際、識別結果を空間マップデータ 127 として記憶部に格納する。

図 10 に、空間マップデータ 127 のデータ構成例を示す。

【0111】

図 10 に示すように、空間マップデータは以下の各データの対応データが格納される。

- (a) タイムスタンプ (sec)
- (b) 位置情報
- (c) クラス
- (d) 識別処理後経過時間 (sec)

【0112】

(a) タイムスタンプ (sec) は、オブジェクト識別処理を実行した時間情報である。

(b) 位置情報は、オブジェクト識別対象となった実オブジェクトの位置情報である。位置情報の記録方式は、様々な方式が利用可能である。図に示す例は、3次元座標 (x, y, z) のリストによるメッシュとして記述した例である。この他、例えば、ターゲット領域の中心位置の位置情報を記録するといった構成とでもよい。

【0113】

(c) クラスは、オブジェクト識別結果としてのオブジェクト種類情報である。

(d) 識別処理後経過時間 (sec) は、オブジェクト識別処理の完了時からの経過時間である。

【0114】

なお、オブジェクト識別部 126 は、ターゲット領域が決定された直後に、ターゲット領域の実オブジェクトの識別処理を実行するが、その後も、その領域に対するオブジェクト識別処理を繰り返し実行し、図 10 に示す空間マップデータを、逐次、更新する。

【0115】

ただし、更新処理の間隔は、識別された実オブジェクトの種類 (クラス) によって異なる。

この実オブジェクトの種類 (クラス) によって異なる更新時間の規定データが、クラス対応更新時間データ 128 として予め登録されている。

図 11 に、クラス対応更新時間データ 128 のデータ例を示す。

【0116】

図 11 に示すように、クラス対応更新時間データ 128 は、以下の各データを対応付けたデータである。

- (a) ID
- (b) カテゴリ
- (c) クラス
- (d) 更新時間 (sec)

【0117】

(a) ID は、登録データの識別子である。

(b) カテゴリは、実オブジェクトの種類 (クラス) のカテゴリである。

(c) クラスは、実オブジェクトの種類情報である。

(d) 更新時間 (sec) は、実オブジェクト識別処理の更新間隔を示す時間である。

【0118】

例えば、ID001 のクラス (オブジェクト種類) = 芝生の場合、更新時間は 3600 sec (= 1 時間) である。芝生のようなオブジェクトは、時間経過に伴う変化が少なく、更新時間が長く設定される。

一方、例えば ID = 004 のクラス (オブジェクト種類) = 影の場合、更新時間は 2 sec である。影のようなオブジェクトは、時間経過に伴う変化が大きいため、更新時間が短く設定される。

【0119】

オブジェクト識別部 126 は、このクラス対応更新時間データ 128 のデータを参照し

10

20

30

40

50

て、識別したオブジェクトについて規定された時間間隔で、随時、オブジェクト識別処理を繰り返し、実行する。新たな識別処理によって検出された実オブジェクトは、図 10 を参照して説明した空間マップデータ 127 として、順次、登録する。

【0120】

また、前述したように、データ出力部 130 のコンテンツ表示制御部 131 と、コンテンツ音声制御部 132 は、クラス対応仮想オブジェクトデータ 135 に記録された仮想オブジェクトの 3D コンテンツや様々な音声データを取得してデータ出力を実行する。

【0121】

クラス対応仮想オブジェクトデータ 135 には、ターゲット領域の実オブジェクトの識別結果に相当する実オブジェクト種類（クラス）に対応づけられた表示用の仮想オブジェクトの 3D コンテンツや様々な音声データが記録されている。

10

図 12 を参照して、クラス対応仮想オブジェクトデータ 135 の具体例について説明する。

【0122】

図 12 に示すように、クラス対応仮想オブジェクトデータ 135 には、以下の各データが対応付けて記録されている。

- (a) クラス
- (b) 仮想オブジェクト 3D モデル（キャラクタ 3D モデル）
- (c) 出力音声

【0123】

(a) クラスは、実オブジェクトの種類情報である。  
 (b) 仮想オブジェクト 3D モデル（キャラクタ 3D モデル）は、各クラス、すなわちターゲット領域の実オブジェクトの種類に応じて出力（表示）する仮想オブジェクト（キャラクタ）の 3D モデルを登録している。なお、図に示す例では 3D モデルの ID と、3D モデルを併せて記録した構成としているが、例えば ID のみを記録し、ID に基づいて他のデータベースから ID 対応の 3D モデルを取得する構成としてもよい。

20

【0124】

(c) 出力音声は、各クラス、すなわちターゲット領域の実オブジェクトの種類に応じて出力する音声データを登録している。

【0125】

このように、クラス対応仮想オブジェクトデータ 135 には、ターゲット領域の実オブジェクトの識別結果に相当する実オブジェクト種類（クラス）に対応づけられた表示用の仮想オブジェクトの 3D コンテンツや様々な音声データが記録されている。

30

【0126】

なお、クラス対応仮想オブジェクトデータ 135 には、さらに、各仮想オブジェクトの出力態様情報も記録されている。例えば先に図 7 や図 8 を参照して説明したように、ターゲット領域の実オブジェクトが水である場合の表示態様情報や、砂場である場合の表示態様情報等の情報も記録されている。

【0127】

データ出力部 130 のコンテンツ表示制御部 131 と、コンテンツ音声制御部 132 は、この図 12 に示すようなデータを格納したクラス対応仮想オブジェクトデータ 135 に記録された仮想オブジェクトの 3D コンテンツや様々な音声データを取得してデータ出力を実行する。

40

【0128】

[ 3 . 本開示の画像処理装置の実行する処理のシーケンスについて ]

次に、本開示の画像処理装置 100 の実行する処理のシーケンスについて説明する。

【0129】

なお、以下に示す複数の処理シーケンスについて、順次、説明する。

- (1) 画像処理装置の実行する基本処理シーケンス
- (2) ターゲット領域の設定領域をほぼ水平な面に設定する処理のシーケンス

50

### (3) 実オブジェクト識別処理の更新シーケンス

#### 【0130】

##### (3-(1) 画像処理装置の実行する基本処理シーケンス)

まず、図13に示すフローチャートを参照して本開示の画像処理装置100の実行する基本的な処理のシーケンスについて説明する。

#### 【0131】

なお、図13以下に示すフローチャートに従った処理は、主に画像処理装置100のデータ処理部120において実行される処理である。データ処理部120は、プログラム実行機能を持つCPUを備え、記憶部に格納されたプログラムに従ってフローに従った処理を実行する。

10

以下、図13に示すフローの各ステップの処理について説明する。

#### 【0132】

##### (ステップS101)

まず、画像処理装置100のデータ処理部120は、ステップS101において、外部撮影カメラの撮影画像を入力する。

#### 【0133】

##### (ステップS102)

次に、データ処理部120は、ステップS102において、入力した外部撮影カメラの撮影画像から特徴点を抽出する。

20

#### 【0134】

この処理は、図9に示すデータ処理部120の外部撮影画像解析部121が実行する処理である。

外部撮影画像解析部121は、入力した外部撮影カメラの撮影画像から特徴点を抽出する。この特徴点抽出処理は、3次元マップの生成に利用するための特徴点であり、抽出した特徴点情報は、外部撮影カメラによって撮影された外部の撮影画像とともに、3次元マップ生成部122に入力する。

#### 【0135】

##### (ステップS103)

次に、データ処理部は、ステップS103において、外部撮影カメラによって撮影された外部の撮影画像とその特徴点情報を利用して3次元マップを生成する。

30

#### 【0136】

この処理は、図9に示すデータ処理部120の3次元マップ生成部122が実行する処理である。

#### 【0137】

3次元マップ生成部122は、外部撮影カメラ111によって撮影された外部の撮影画像と、外部撮影画像解析部121が抽出した特徴点に基づいて、外部の実オブジェクトによって構成される3次元マップを生成する。

この3次元マップ生成処理は、例えば、SLAM (Simultaneous Localization and Mapping) 処理によってリアルタイム処理として実行する。

40

#### 【0138】

##### (ステップS104)

次に、データ処理部は、ステップS104において、ターゲット領域決定処理を実行する。

#### 【0139】

この処理は、図9に示すデータ処理部120のオブジェクト識別部126が実行する処理である。

オブジェクト識別部126は、仮想オブジェクト表示領域とするターゲット領域を決定する。

#### 【0140】

50

このターゲット領域の決定処理は、前述したように、様々な手法が適用可能である。

例えば、3次元マップに含まれるユーザの指の画像を用いて行うことが可能である。

すなわち、ユーザの指さし方向の延長線と、3次元マップ上の実オブジェクトとの交点を求め、例えば、交点を中心とするあらかじめ規定した半径の円形領域をターゲット領域として決定する。

#### 【0141】

さらに、図9に示すデータ入力部110の各構成部からの入力情報を用いてターゲット領域を決定することも可能である。具体的には、以下の入力情報である。

(a) 内部撮影画像解析部123の解析したユーザ視線情報

(b) デバイス姿勢解析部124の解析した画像処理装置100本体の姿勢や動き情報 10

(c) 操作部114を介して入力するユーザ操作情報

(d) 音声解析部125が解析したユーザ音声情報

例えば、これらの入力情報のいずれかを用いてターゲット領域を決定してもよい。

#### 【0142】

代表的なターゲット領域の決定シーケンスについて、図14、図15を参照して説明する。

#### 【0143】

図14(1)は、ユーザの指さし方向の解析に基づくターゲット領域の決定シーケンスである。ユーザの指さし方向の解析に基づくターゲット領域決定処理は、以下の処理シーケンスで実行される。

まず、ステップS211で、ユーザの指さし方向を解析する。この解析処理は、3次元マップ生成部122の生成した3次元マップを利用して実行する。

#### 【0144】

次に、ステップS212で、ユーザの指さし方向の延長線からなる直線と、実オブジェクトとの交点を検出する。この処理も3次元マップ生成部122の生成した3次元マップを利用して実行する。

#### 【0145】

最後に、ステップS213において、ユーザの指さし方向の延長線からなる直線と、実オブジェクトとの交点を中心とする円形領域をターゲット領域として決定する。

なお、ターゲット領域の形状は任意であり、円形の他、矩形形状としてもよい。またターゲット領域の大きさも任意であり、様々な大きさに設定可能である。

ただし、形状、大きさは、事前に規定し、その規定に従ってターゲット領域を決定することが好ましい。

#### 【0146】

図14(2)は、ユーザの視線方向解析に基づくターゲット領域の決定シーケンスである。ユーザ視線方向の解析に基づくターゲット領域決定処理は、以下の処理シーケンスで実行される。

まず、ステップS221で、ユーザの視線方向を解析する。この解析処理は、内部撮影カメラ112の撮影画像に基づいて、内部撮影画像解析部123が実行する。

#### 【0147】

次に、ステップS222で、ユーザ視線方向の延長線からなる直線と、実オブジェクトとの交点を検出する。この処理は3次元マップ生成部122の生成した3次元マップを利用して実行する。

#### 【0148】

最後に、ステップS223において、ユーザ視線方向の延長線からなる直線と、実オブジェクトとの交点を中心とする円形領域をターゲット領域として決定する。

なお、前述したように、ターゲット領域の形状、大きさはさまざまな設定が可能である。

#### 【0149】

図15(3)は、ユーザの操作情報の解析に基づくターゲット領域の決定シーケンスである。ユーザ操作情報の解析に基づくターゲット領域決定処理は、以下の処理シーケンス

10

20

30

40

50

で実行される。

まず、ステップS 2 3 1で、ユーザの操作情報を解析する。例えば、先に図6を参照して説明したスマホに対するタッチ操作等である。

【0150】

次に、ステップS 2 3 2で、ユーザ操作情報に基づく実オブジェクト指定位置を検出する。この処理は例えばユーザの指の接触位置の検出処理として実行される。

【0151】

最後に、ステップS 2 3 3において、ユーザ操作情報に基づく実オブジェクト指定位置を中心とする円形領域をターゲット領域として決定する。

なお、前述したように、ターゲット領域の形状、大きさはさまざまな設定が可能である。

【0152】

図15(4)は、ユーザの音声情報の解析に基づくターゲット領域の決定シーケンスである。ユーザの音声情報の解析に基づくターゲット領域決定処理は、以下の処理シーケンスで実行される。

まず、ステップS 2 4 1で、ユーザの発話音声を解析する。例えば、「前の池」等のユーザ発話音声を解析する。

【0153】

次に、ステップS 2 4 2で、ユーザの音声情報に基づく実オブジェクト指定位置を検出する。

【0154】

最後に、ステップS 2 4 3において、ユーザ発話音声に基づく実オブジェクト指定位置を中心とする円形領域をターゲット領域として決定する。

なお、前述したように、ターゲット領域の形状、大きさはさまざまな設定が可能である。

【0155】

図14、図15を参照して説明した以外にも、例えば、以下のようなターゲット領域決定処理を行ってもよい。

(a) デバイス姿勢解析部124の解析した画像処理装置100本体の姿勢や動き情報を用いたターゲット領域決定処理。

(b) 外部撮影カメラ111によって撮影された画像に基づいて生成される3次元マップや、モーションセンサ113の検出情報に基づいて地面、床面、水面等の水平面の検出処理を実行し、撮影画像の中心領域に最も近い水平面の領域をターゲット領域として決定する処理。

(c) ユーザが仮想的なボールを投げる動作を行い、これを外部撮影カメラ111によって撮影し、撮影画像を解析してボールの着地点を解析して、この着地点をターゲット領域の中心位置として決定する処理。

(d) その他、ユーザ動作、ユーザ視線、ユーザ操作、ユーザ位置、またはユーザ姿勢の少なくともいずれかを解析して、解析結果に基づいてターゲット領域を決定する処理。

【0156】

図13に示すフローに戻り、説明を続ける。

上述したように、画像処理装置100のデータ処理部120のオブジェクト識別部126は、ステップS 1 0 4において、ターゲット領域決定処理を実行する。

【0157】

(ステップS 1 0 5)

次に、データ処理部は、ステップS 1 0 5において、ターゲット領域の実オブジェクトを識別する。

具体的には、ターゲット領域が池である、木である等のオブジェクト識別処理を実行する。

【0158】

前述したように、この実オブジェクトのオブジェクト識別処理は、例えば、セマンティック・セグメンテーション処理を適用して実行する。

10

20

30

40

50

セマンティック・セグメンテーションは、様々な実際のオブジェクトの形状情報やその他の特徴情報を登録したオブジェクト識別用の辞書データ（学習済みデータ）と、例えばカメラ撮影画像内のオブジェクトとの一致度に基づいて、画像の構成画素（ピクセル）各々が、どのオブジェクトカテゴリに属する画素であるかを識別する技術である。

【0159】

なお、オブジェクト識別部126において実行する実オブジェクトの識別処理は、ターゲット領域のみ、あるいはターゲット領域を含む周囲領域の限られた範囲のみを対象として実行する。このような限定範囲の処理を行うことで高速の処理、すなわちリアルタイム処理が可能となる。

【0160】

オブジェクト識別部126が解析したターゲット領域の実オブジェクトの識別結果は、データ出力部130のコンテンツ表示制御部131と、コンテンツ音声制御部132に入力される。

【0161】

（ステップS106）

次に、ステップS106において、識別されたターゲット領域の実オブジェクト（クラス）に基づいて、ターゲット領域に表示する仮想オブジェクトの種類と出力態様を決定する。

【0162】

この処理は、図9に示す画像処理装置100のデータ出力部130のコンテンツ表示制御部131、コンテンツ音声制御部132が実行する処理である。

データ出力部130のコンテンツ表示制御部131と、コンテンツ音声制御部132は、先に図12を参照して説明したデータを記録したクラス対応仮想オブジェクトデータ135を参照して、ターゲット領域に表示する仮想オブジェクトの種類と出力態様を決定する。

【0163】

すなわち、クラス対応仮想オブジェクトデータ135の各エントリから、ターゲット領域の実オブジェクト種類（クラス）を記録したエントリを選択し、そのエントリに記録された仮想オブジェクトを出力オブジェクトとして決定する処理等を実行する。

【0164】

（ステップS107）

最後に、ステップS107において、ステップS106で決定したターゲット領域に表示する仮想オブジェクトの種類と出力態様に従って仮想オブジェクトをターゲット領域に出力（表示）する。

【0165】

この処理も、図9に示す画像処理装置100のデータ出力部130のコンテンツ表示制御部131、コンテンツ音声制御部132が実行する処理である。

【0166】

コンテンツ表示制御部131は、オブジェクト識別部126からターゲット領域のオブジェクト識別結果を入力し、オブジェクト識別結果に従って表示すべき仮想オブジェクト（キャラクタ等）の選択処理や表示態様を決定して表示部133に表示する。

具体的には、例えば先に説明した図2～図4、図7、図8に示すような仮想オブジェクト（キャラクタ等）の表示処理を実行する。

【0167】

また、コンテンツ音声制御部132は、オブジェクト識別部126からターゲット領域のオブジェクト識別結果を入力し、オブジェクト識別結果に従って出力すべき音声を決定してスピーカ134を介して出力する。

具体的には、例えば先に説明した図7に示すように、実オブジェクトである池から仮想オブジェクトが現れる場合、水の音を出力する処理を実行する。

【0168】

10

20

30

40

50

( 3 - ( 2 ) ターゲット領域の設定領域をほぼ水平な面に設定する処理のシーケンス)

次に、図 1 6 に示すフローチャートを参照して、ターゲット領域の設定領域をほぼ水平な面に設定する処理のシーケンスについて説明する。

【 0 1 6 9 】

キャラクタ等の仮想オブジェクトを実世界の実オブジェクト上に表示する場合、屋外であれば地面の上、室内であれば床の上に表示すれば、より自然なキャラクタ表示が可能となり、ユーザにキャラクタが実際に実世界に存在するような感覚を与えることが可能となる。

【 0 1 7 0 】

このためには、キャラクタである仮想オブジェクトの出力領域となるターゲット領域を地面の上や床の上等、ほぼ水平な面に設定する制御を行うことが有効である。

図 1 6 に示すフローチャートは、このような処理を実行する画像処理装置 1 0 0 の処理シーケンスを説明するフローチャートである。

以下、図 1 6 に示すフローチャートの各ステップの処理について説明する。

【 0 1 7 1 】

なお、図 1 6 に示すフローチャートのステップ S 1 0 1 ~ S 1 0 3 とステップ S 1 0 5 ~ S 1 0 7 の処理は、先に図 1 3 を参照して説明した基本処理フローの各ステップの処理と同様の処理である。

【 0 1 7 2 】

図 1 6 に示すフローのステップ S 3 0 1 ~ S 3 0 3 の処理と、ステップ S 1 0 4 の処理が先に説明した図 1 3 に示すフローと異なる点である。

この各ステップの処理について説明する。

【 0 1 7 3 】

( ステップ S 3 0 1 )

ステップ S 3 0 1 は、図 9 に示す画像処理装置 1 0 0 のデータ入力部 1 1 0 のモーションセンサ 1 1 3 から、データ処理部 1 2 0 のデバイス姿勢解析部 1 2 4 にセンサ検出情報を入力する処理である。

【 0 1 7 4 】

先に図 9 を参照して説明したように、モーションセンサ 1 1 3 は、ジャイロ、加速度センサ等を有し、例えば H M D やスマホ等、画像処理装置 1 0 0 本体の姿勢や動きを検出するセンサである。

このモーションセンサ 1 1 3 から、データ処理部 1 2 0 のデバイス姿勢解析部 1 2 4 にセンサ検出情報が入力される。

【 0 1 7 5 】

( ステップ S 3 0 2 )

次に、ステップ S 3 0 2 において、モーションセンサ検出情報に基づいて重力方向を推定する。

【 0 1 7 6 】

この処理は、図 9 に示すデータ処理部 1 2 0 のデバイス姿勢解析部 1 2 4 が実行する処理である。

データ処理部 1 2 0 のデバイス姿勢解析部 1 2 4 は、モーションセンサ 1 1 3 を構成するジャイロや、加速度センサ等のセンサ検出情報を用いて、重力方向を算出する。

【 0 1 7 7 】

( ステップ S 3 0 3 )

次に、ステップ S 3 0 3 において、水平面領域の検出処理を行う。

この処理は、図 9 に示すオブジェクト識別部 1 2 6 が実行する処理である。

【 0 1 7 8 】

オブジェクト識別部 1 2 6 は、3次元マップ生成部 1 2 2 の生成した3次元マップと、デバイス姿勢解析部 1 2 4 から入力した重力方向情報を用いて、3次元マップ内の水平面領域を検出する。具体的には例えば地面や床面等を検出する。

10

20

30

40

50

なお、検出する水平面領域は、完全な水平面に限定されず、ほぼ水平な領域とすればよい。

【0179】

例えばある程度でこぼこや、ある程度の勾配を持つ坂道等も水平面領域であると判定して検出する。

どの程度でこぼこや傾斜を水平面領域として許容するかについては、事前に設定することが可能である。

【0180】

(ステップS104)

次に、データ処理部は、ステップS104において、ターゲット領域決定処理を実行する。

10

ただし、本処理例では、ターゲット領域は、ステップS303において検出した水平面領域からのみ選択するものとする。

【0181】

この処理は、図9に示すデータ処理部120のオブジェクト識別部126が実行する処理である。

オブジェクト識別部126は、仮想オブジェクト表示領域とするターゲット領域をステップS303において検出した水平面領域内に限定して決定する。

【0182】

このターゲット領域の決定処理は、前述したように、様々な手法が適用可能である。

20

例えば、3次元マップに含まれるユーザの指の画像を用いて行うことが可能である。

すなわち、ユーザの指さし方向の延長線と、3次元マップ上の実オブジェクトであり、地面や床面等の水平面と判定された水平面領域との交点を求め、この水平面との交点を中心とするあらかじめ規定した半径の円形領域をターゲット領域として決定する。

【0183】

なお、ターゲット領域の決定に利用するデータは、先に図13を参照して説明したと同様、様々な情報を利用することが可能である。例えば、以下の入力情報をもちいる構成としてもよい。

(a) 内部撮影画像解析部123の解析したユーザ視線情報

(b) デバイス姿勢解析部124の解析した画像処理装置100本体の姿勢や動き情報

30

(c) 操作部114を介して入力するユーザ操作情報

(d) 音声解析部125が解析したユーザ音声情報

例えば、これらの入力情報のいずれかを用いてターゲット領域を決定してもよい。

【0184】

ステップS101～S103の処理とステップS105以下の処理は、先に説明した図13に示すフローチャートと同様の処理となる。

【0185】

本処理例では、キャラクタである仮想オブジェクトの出力領域となるターゲット領域を地面の上や床の上等、ほぼ水平な面に設定する制御を行うことが可能となる。

これにより、キャラクタ等の仮想オブジェクトを実世界の実オブジェクト上に表示する場合、屋外であれば地面の上、室内であれば床の上等、水平面領域に接するように表示することが可能となり、より自然なキャラクタ表示が可能となり、ユーザにキャラクタが実際に実世界に存在するような感覚を与えることが可能となる。

40

【0186】

(3-(3)実オブジェクト識別処理の更新シーケンス)

次に、オブジェクト識別部が実行する実オブジェクト識別処理の更新シーケンスについて説明する。

【0187】

先に図10、図11等を参照して説明したように、図9に示す画像処理装置100のデータ処理部120内のオブジェクト識別部126は、ターゲット領域が決定された直後に

50

、ターゲット領域の実オブジェクトの識別処理を実行するが、その後も、その領域に対するオブジェクト識別処理を繰り返し実行し、図10に示す空間マップデータを、逐次、更新する。

【0188】

ただし、更新処理の間隔は、識別された実オブジェクトの種類（クラス）によって異なる。

この実オブジェクトの種類（クラス）によって異なる更新時間の規定データは、クラス対応更新時間データ128として予め登録されている。

クラス対応更新時間データ128は、先に図11を参照して説明したように以下の各データを対応付けたデータである。

- (a) ID
- (b) カテゴリ
- (c) クラス
- (d) 更新時間 (sec)

【0189】

- (a) IDは、登録データの識別子である。
- (b) カテゴリは、実オブジェクトの種類（クラス）のカテゴリである。
- (c) クラスは、実オブジェクトの種類情報である。
- (d) 更新時間 (sec) は、実オブジェクト識別処理の更新間隔を示す時間である。

【0190】

例えば、ID001のクラス（オブジェクト種類）＝芝生の場合、更新時間は3600sec（＝1時間）である。芝生のようなオブジェクトは、時間経過に伴う変化が少なく、更新時間が長く設定される。

一方、例えばID＝004のクラス（オブジェクト種類）＝影の場合、更新時間は2secである。影のようなオブジェクトは、時間経過に伴う変化が大きいため、更新時間が短く設定される。

【0191】

オブジェクト識別部126は、このクラス対応更新時間データ128のデータを参照して、識別したオブジェクトについて規定された時間間隔で、随時、オブジェクト識別処理を繰り返し、実行する。新たな識別処理によって検出された実オブジェクトは、図10を参照して説明した空間マップデータ127として、順次、登録する。

【0192】

図17に示すフローチャートは、このオブジェクト識別処理の繰り返し実行シーケンスを含む処理を説明するフローチャートである。

以下、図17に示すフローチャートの各ステップの処理について説明する

【0193】

なお、図17に示すフローチャートのステップS101～S105の処理と、ステップS106～S107の処理は、先に図13を参照して説明した基本処理フローの各ステップの処理と同様の処理である。

【0194】

図17に示すフローのステップS401とステップS402の処理が先に説明した図13に示すフローと異なる点である。

この各ステップの処理について説明する。

【0195】

（ステップS401）

ステップS101～S105において、ターゲット領域の決定と、ターゲット領域の実オブジェクト（クラス）の識別処理が実行された後、ステップS401の処理を実行する。

【0196】

ステップS401では、ステップS105において実行されたターゲット領域のオブジェクト識別結果を空間マップデータに記録する。

10

20

30

40

50

## 【 0 1 9 7 】

空間マップデータは、先に図 1 0 を参照して説明したように、以下の各データの対応データを格納している。

- ( a ) タイムスタンプ ( s e c )
- ( b ) 位置情報
- ( c ) クラス
- ( d ) 識別処理後経過時間 ( s e c )

## 【 0 1 9 8 】

( a ) タイムスタンプ ( s e c ) は、オブジェクト識別処理を実行した時間情報である。  
( b ) 位置情報は、オブジェクト識別対象となった実オブジェクトの位置情報である。  
( c ) クラスは、オブジェクト識別結果としてのオブジェクト種類情報である。  
( d ) 識別処理後経過時間 ( s e c ) は、オブジェクト識別処理の完了時からの経過時間である。

10

## 【 0 1 9 9 】

ステップ S 4 0 1 では、ステップ S 1 0 5 において識別したターゲット領域の実オブジェクトについて、これらの各データを空間マップデータに登録する。

## 【 0 2 0 0 】

( ステップ S 1 0 6 ~ S 1 0 7 )

ステップ S 1 0 6 ~ S 1 0 7 の処理は、先に図 1 3 を参照して説明した処理と同様の処理である。すなわち、以下の処理を実行する。

20

## 【 0 2 0 1 】

ステップ S 1 0 6 において、識別されたターゲット領域の実オブジェクト ( クラス ) に基づいて、ターゲット領域に表示する仮想オブジェクトの種類と出力態様を決定する。

ステップ S 1 0 7 において、ステップ S 1 0 6 で決定したターゲット領域に表示する仮想オブジェクトの種類と出力態様に従って仮想オブジェクトをターゲット領域に出力 ( 表示 ) する。

## 【 0 2 0 2 】

( ステップ S 4 0 2 )

さらに、ステップ S 1 0 7 の処理後、ステップ S 4 0 2 において、ステップ S 1 0 5 において実行したターゲット領域の実オブジェクトの識別処理後の経過時間が、図 1 1 を参照して説明したクラス対応更新時間データに規定された「 ( d ) 更新時間」を超えたか否かを判定する。

30

## 【 0 2 0 3 】

超えたと判定した場合、ステップ S 1 0 1 に戻り、ステップ S 1 0 1 以下の処理を繰り返し、実行する。

すなわち、再度、ターゲット領域の決定、ターゲット領域の実オブジェクト識別処理を実行する。

## 【 0 2 0 4 】

この処理において、ターゲット領域の位置が変更されていなければ、同じ位置のターゲット領域で実オブジェクト識別が再度実行される。

40

一方、ターゲット領域の位置が変更されていれば、新たな位置のターゲット領域で実オブジェクト識別が実行される。

## 【 0 2 0 5 】

これらの処理を行うことで、ターゲット領域の更新処理や、実オブジェクトの識別結果の更新処理を即座に行うことが可能となり、ユーザの動きや指示に応じたタイムリな仮想オブジェクト表示処理を行うことが可能となる。

## 【 0 2 0 6 】

[ 4 . 画像処理装置のハードウェア構成例について ]

次に、上述した実施例に従った処理を実行する画像処理装置のハードウェア構成例について、図 1 8 を参照して説明する。

50

図18に示すハードウェアは、図9を参照して説明した本開示の画像処理装置100のハードウェア構成の一例である。

図18に示すハードウェア構成について説明する。

【0207】

CPU(Central Processing Unit)301は、ROM(Read Only Memory)302、または記憶部308に記憶されているプログラムに従って各種の処理を実行するデータ処理部として機能する。例えば、上述した実施例において説明したシーケンスに従った処理を実行する。RAM(Random Access Memory)303には、CPU301が実行するプログラムやデータなどが記憶される。これらのCPU301、ROM302、およびRAM303は、バス304により相互に接続されている。

10

【0208】

CPU301はバス304を介して入出力インタフェース305に接続され、入出力インタフェース305には、各種センサ、カメラ、スイッチ、キーボード、マウス、マイクロホンなどよりなる入力部306、ディスプレイ、スピーカなどよりなる出力部307が接続されている。

【0209】

入出力インタフェース305に接続されている記憶部308は、例えばハードディスク等からなり、CPU301が実行するプログラムや各種のデータを記憶する。通信部309は、インターネットやローカルエリアネットワークなどのネットワークを介したデータ通信の送受信部、さらに放送波の送受信部として機能し、外部の装置と通信する。

20

【0210】

入出力インタフェース305に接続されているドライブ310は、磁気ディスク、光ディスク、光磁気ディスク、あるいはメモリカード等の半導体メモリなどのリムーバブルメディア311を駆動し、データの記録あるいは読み取りを実行する。

【0211】

[5. 本開示の構成のまとめ]

以上、特定の実施例を参照しながら、本開示の実施例について詳解してきた。しかしながら、本開示の要旨を逸脱しない範囲で当業者が実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本開示の要旨を判断するためには、特許請求の範囲の欄を参酌すべきである。

30

【0212】

なお、本明細書において開示した技術は、以下のような構成をとることができる。

(1) 実世界上の実オブジェクトの識別処理を実行するオブジェクト識別部と、実オブジェクトと仮想オブジェクトを重畳表示したAR(Augmented Reality)画像を生成するコンテンツ表示制御部を有し、前記オブジェクト識別部は、前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行し、前記コンテンツ表示制御部は、前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択する画像処理装置。

40

【0213】

(2) 前記オブジェクト識別部は、画像認識処理によりオブジェクト識別処理を実行する(1)に記載の画像処理装置。

【0214】

(3) 前記オブジェクト識別部は、セマンティック・セグメンテーション処理を適用してオブジェクト識別処理を実行する(2)に記載の画像処理装置。

50

## 【0215】

(4) 前記オブジェクト識別部は、

前記仮想オブジェクトの表示領域とするターゲット領域を決定し、決定したターゲット領域の実オブジェクトの識別処理を実行する(1)~(3)いずれかに記載の画像処理装置。

## 【0216】

(5) 前記オブジェクト識別部は、

前記ターゲット領域をユーザ動作、ユーザ視線、ユーザ操作、ユーザ位置、またはユーザ姿勢の少なくともいずれかに基づいて決定する(4)に記載の画像処理装置。

## 【0217】

(6) 前記オブジェクト識別部は、

前記ターゲット領域を水平面領域から選択して決定する(4)または(5)に記載の画像処理装置。

## 【0218】

(7) 前記コンテンツ表示制御部は、

前記仮想オブジェクトを前記水平面領域に接するように表示する(6)に記載の画像処理装置。

## 【0219】

(8) 前記オブジェクト識別部は、

前記オブジェクト識別処理をリアルタイム処理として実行する(1)~(7)に記載の画像処理装置。

## 【0220】

(9) 前記オブジェクト識別部は、

オブジェクト種類に応じて予め規定された時間間隔でオブジェクト識別処理を繰り返し実行する(1)~(8)いずれかに記載の画像処理装置。

## 【0221】

(10) 前記画像処理装置は、

カメラ撮影画像に基づく実世界の3次元マップを生成する3次元マップ生成部を有し、前記オブジェクト識別部は、前記3次元マップを利用して前記仮想オブジェクトの表示領域とするターゲット領域を決定する(1)~(9)いずれかに記載の画像処理装置。

## 【0222】

(11) 前記3次元マップ生成部は、

SLAM(Simultaneous Localization and Mapping)処理により、実世界の3次元マップを生成する(10)に記載の画像処理装置。

## 【0223】

(12) 前記コンテンツ表示制御部は、

前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するとともに、

前記オブジェクト識別結果に応じて、表示する仮想オブジェクトの表示態様も制御する(1)~(11)いずれかに記載の画像処理装置。

## 【0224】

(13) 前記画像処理装置は、さらに、

音声出力制御を実行するコンテンツ音声制御部を有し、前記コンテンツ音声制御部は、前記オブジェクト識別部において識別されたオブジェクト識別結果に応じて出力する音声を決定して出力する(1)~(12)いずれかに記載の画像処理装置。

## 【0225】

(14) 画像処理装置において実行する画像処理方法であり、

オブジェクト識別部が、実世界上の実オブジェクトの識別処理を実行するオブジェクト

10

20

30

40

50

識別処理ステップを実行し、

コンテンツ表示制御部が、実オブジェクトと仮想オブジェクトを重畳表示したAR (Augmented Reality) 画像を生成するコンテンツ表示制御ステップを実行し、

前記オブジェクト識別処理ステップは、

前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行するステップであり、

前記コンテンツ表示制御ステップは、

前記オブジェクト識別処理ステップにおいて識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するステップを実行する画像処理方法。

10

#### 【0226】

(15) 画像処理装置において画像処理を実行させるプログラムであり、

オブジェクト識別部に、実世界上の実オブジェクトの識別処理を実行するオブジェクト識別処理ステップを実行させ、

コンテンツ表示制御部に、実オブジェクトと仮想オブジェクトを重畳表示したAR (Augmented Reality) 画像を生成するコンテンツ表示制御ステップを実行させ、

前記オブジェクト識別処理ステップにおいては、

前記仮想オブジェクトの表示領域の実オブジェクトを識別するオブジェクト識別処理を実行させ、

20

前記コンテンツ表示制御ステップにおいては、

前記オブジェクト識別処理ステップにおいて識別されたオブジェクト識別結果に応じて表示する仮想オブジェクトを選択するステップを実行させるプログラム。

#### 【0227】

また、明細書中において説明した一連の処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。ソフトウェアによる処理を実行する場合は、処理シーケンスを記録したプログラムを、専用のハードウェアに組み込まれたコンピュータ内のメモリにインストールして実行させるか、あるいは、各種処理が実行可能な汎用コンピュータにプログラムをインストールして実行させることが可能である。例えば、プログラムは記録媒体に予め記録しておくことができる。記録媒体からコンピュータにインストールする他、LAN (Local Area Network)、インターネットといったネットワークを介してプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

30

#### 【0228】

なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。また、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

#### 【産業上の利用可能性】

#### 【0229】

40

以上、説明したように、本開示の一実施例の構成によれば、仮想オブジェクトの表示領域とするターゲット領域の実オブジェクト種類に応じて表示する仮想オブジェクトの選択や表示態様変更を行う装置、方法が実現される。

具体的には、例えば、実世界上の実オブジェクトの識別処理を実行するオブジェクト識別部と、実オブジェクトと仮想オブジェクトを重畳表示したAR画像を生成するコンテンツ表示制御部を有する。オブジェクト識別部は、仮想オブジェクトの表示領域とするターゲット領域の実オブジェクトを識別し、コンテンツ表示制御部は、オブジェクト識別結果に応じて表示する仮想オブジェクトを選択する処理や表示態様の変更処理を行う。

本構成により、仮想オブジェクトの表示領域とするターゲット領域の実オブジェクト種類に応じて表示する仮想オブジェクトの選択や表示態様変更を行う装置、方法が実現され

50

る。

【符号の説明】

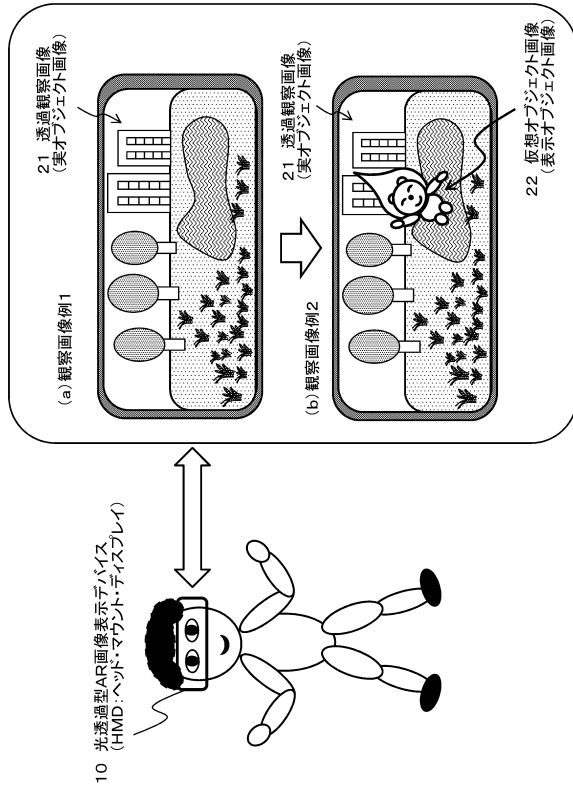
【0230】

10	光透過型AR画像表示デバイス	
11	ターゲット領域	
21	透過観察画像	
22 ~ 24	仮想オブジェクト画像	
30	カメラ撮影画像表示型AR画像表示デバイス	
31	カメラ	
32	カメラ撮影画像	10
40	スマホ	
41	カメラ	
42	カメラ撮影画像	
50	水上の仮想オブジェクト画像	
51	水中の仮想オブジェクト画像	
52	仮想オブジェクト画像	
53 , 54	仮想オブジェクト影画像	
100	画像処理装置	
110	データ入力部	
111	外部撮影カメラ	20
112	内部撮影カメラ	
113	モーションセンサ（ジャイロ、加速度センサ等）	
114	操作部	
115	マイク	
120	データ処理部	
121	外部撮影画像解析部	
122	3次元マップ生成部	
123	内部撮影画像解析部	
124	デバイス姿勢解析部	
125	音声解析部	30
126	オブジェクト識別部	
127	空間マップデータ	
128	クラス対応更新時間データ	
130	データ出力部	
131	コンテンツ表示制御部	
132	コンテンツ音声制御部	
133	表示部	
134	スピーカ	
135	クラス対応仮想オブジェクトデータ（3Dモデル、音声データ等）	
140	通信部	40
301	CPU	
302	ROM	
303	RAM	
304	バス	
305	入出力インタフェース	
306	入力部	
307	出力部	
308	記憶部	
309	通信部	
310	ドライブ	50

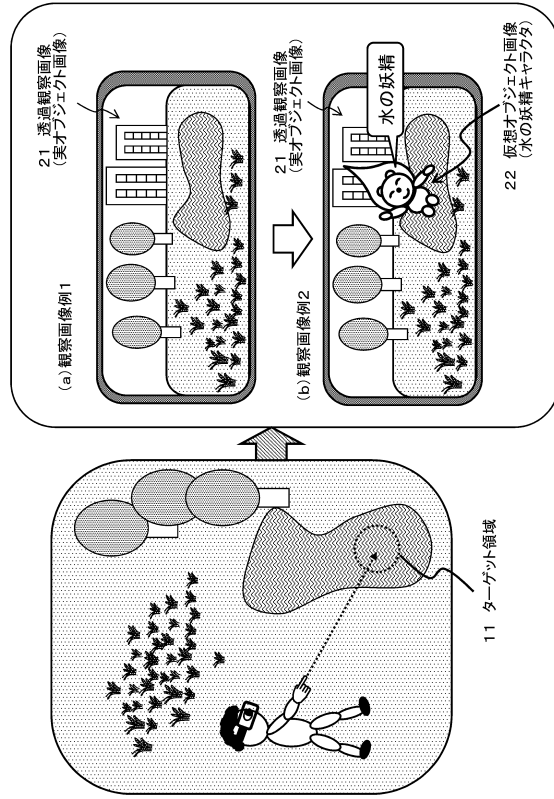
3 1 1 リムーバブルメディア

【 図 面 】

【 図 1 】



【 図 2 】



10

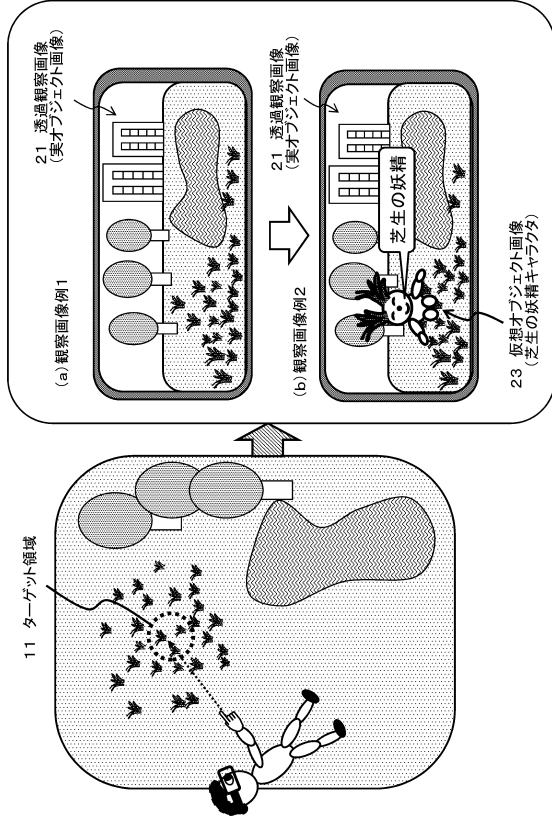
20

30

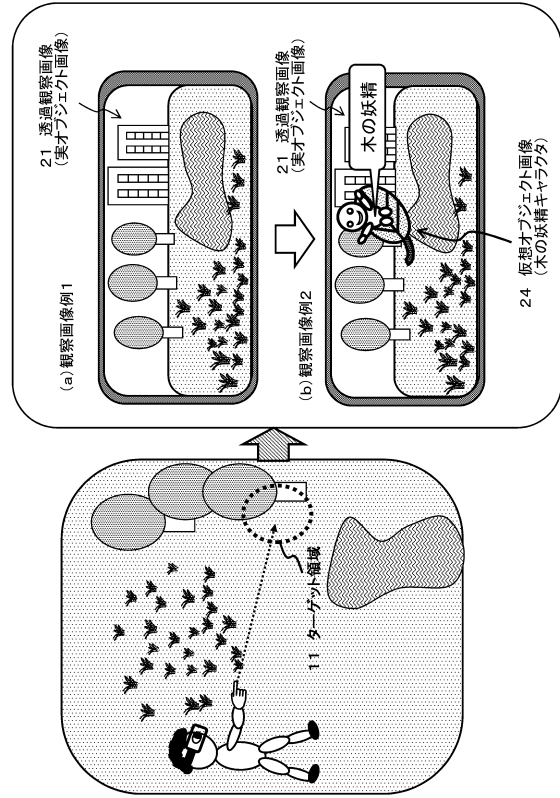
40

50

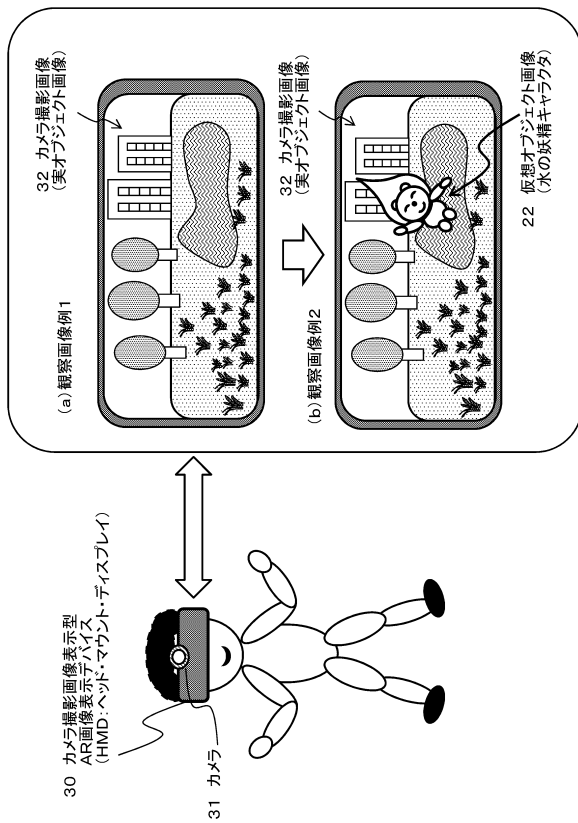
【 図 3 】



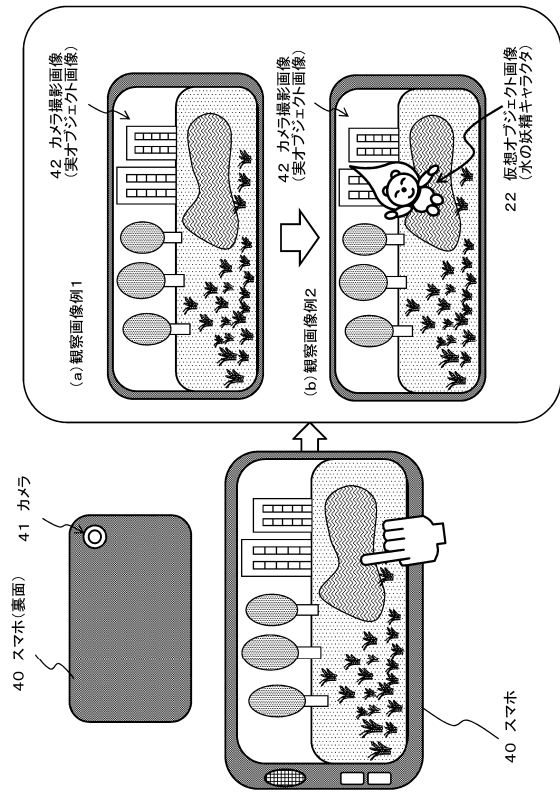
【 図 4 】



【 図 5 】



【 図 6 】



10

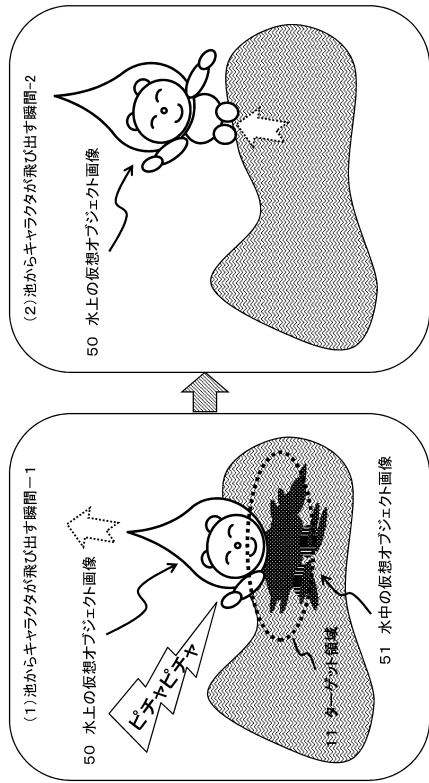
20

30

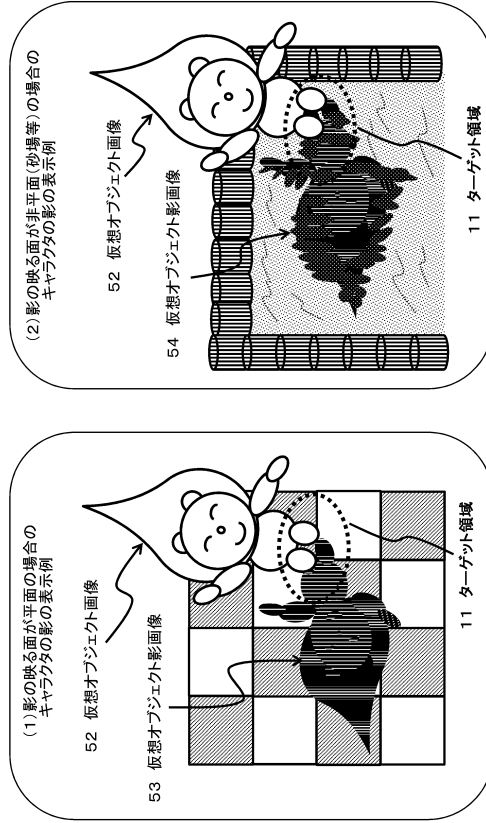
40

50

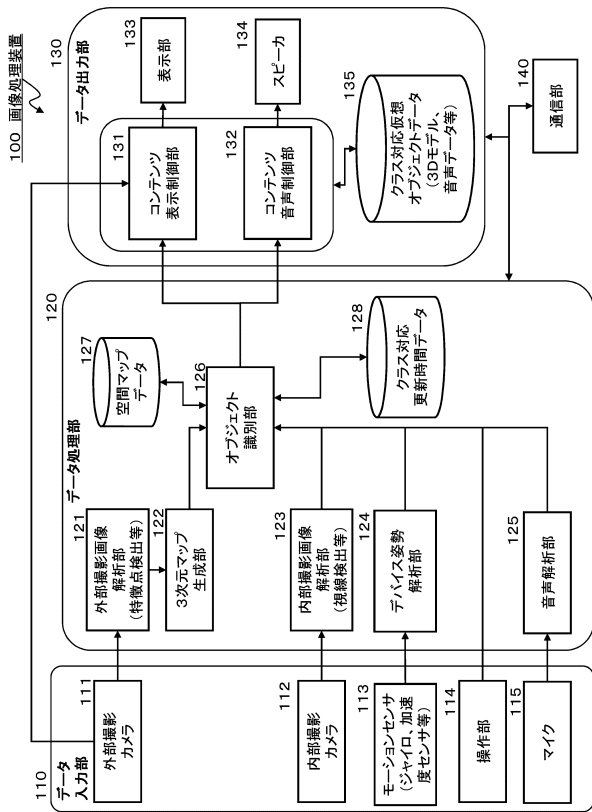
【図 7】



【図 8】



【図 9】






【図 10】

(a) タイムスタンプ(sec)	(b) 位置情報	(c) クラス	(d) 識別処理後経過時間(sec)
10839948	v1355.3356,28 v1232.3460,31 v1344.3577,32 v1234.3456,34	芝生	183
10839950	v890.3351,24 v912.3090,22 v1093.1333,30 v1332.3912,18	アスファルト	3511
10839958	v2239.113,12 v2562.3567,14 v334.134,11 v113.3835,12	水面	452
10839960	v1678.843,31 v879.311,32 v1589.309,30 v865.890,29	影	3
..	..	..	..

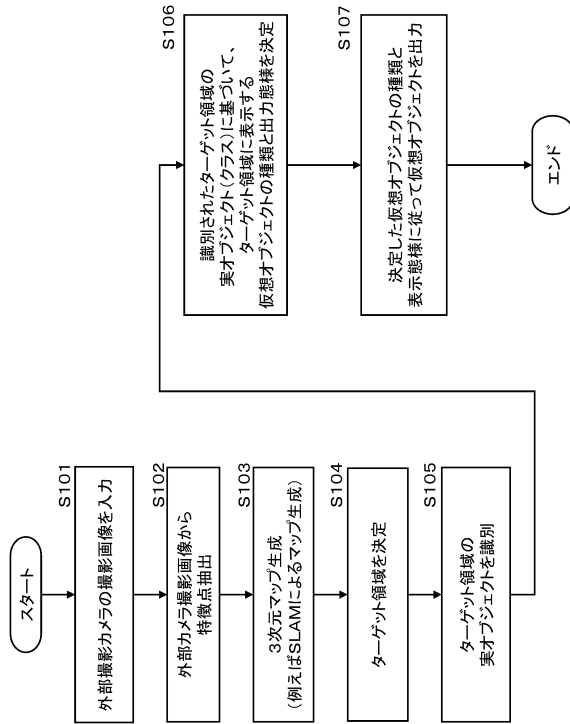
【図 1 1】

(a)ID	(b)カテゴリ	(c)クラス	(d)更新時間(sec)
001	環境	芝生	3600
002	環境	アスファルト	3600
003	環境	水面	600
004	現象	影	2
005	現象	落ち葉に覆われた地面	120
..	..	..	..

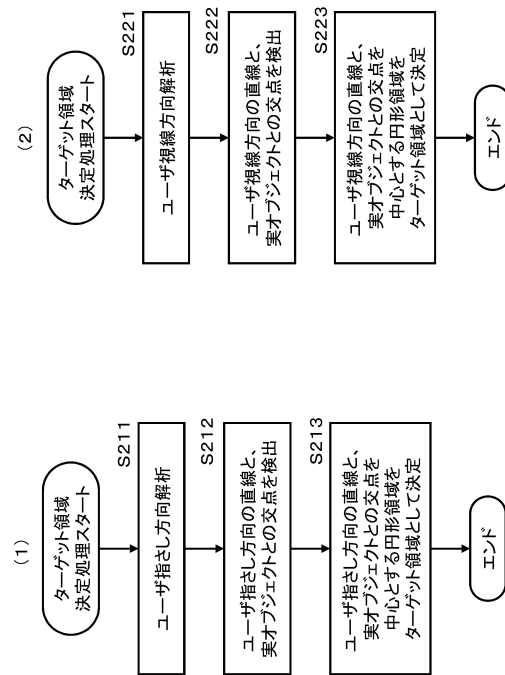
【図 1 2】

(a)クラス	(b)仮想オブジェクト3Dモデル (キャラクター3Dモデル)	(c)出力音声	..
池	0000123(  )	水の音の音声データ	..
芝生	0023459(  )	草の音の音声データ	..
木	0314562(  )	木の葉の音の音声データ	..
..	..	..	..

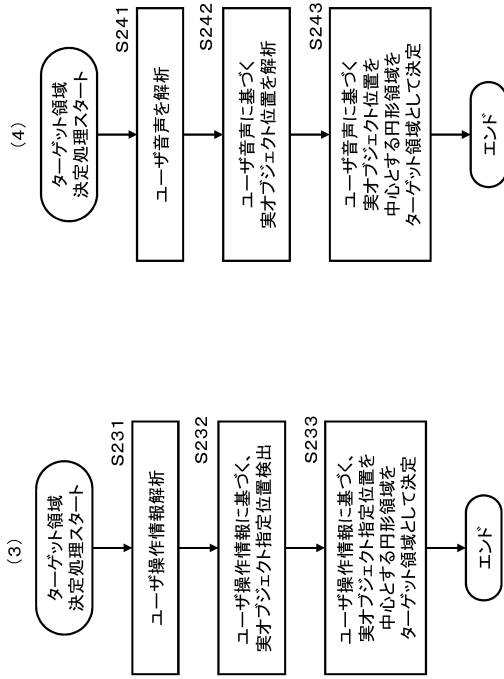
【図 1 3】



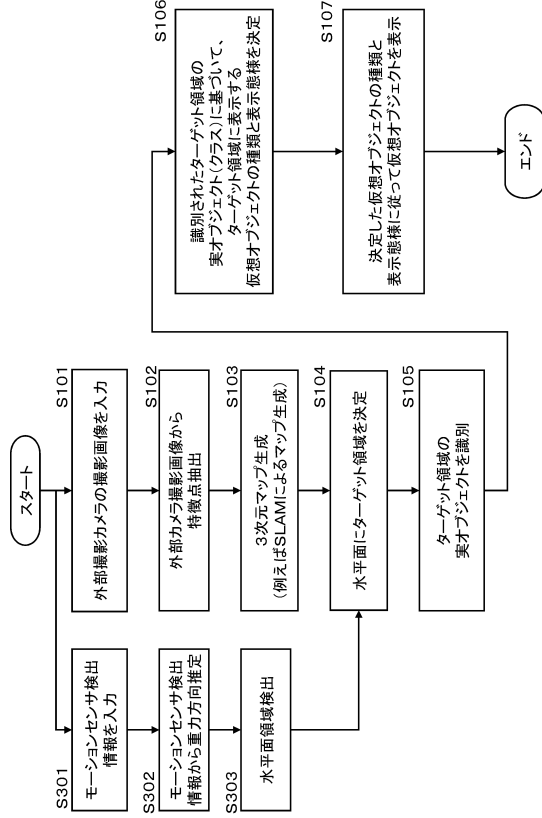
【図 1 4】



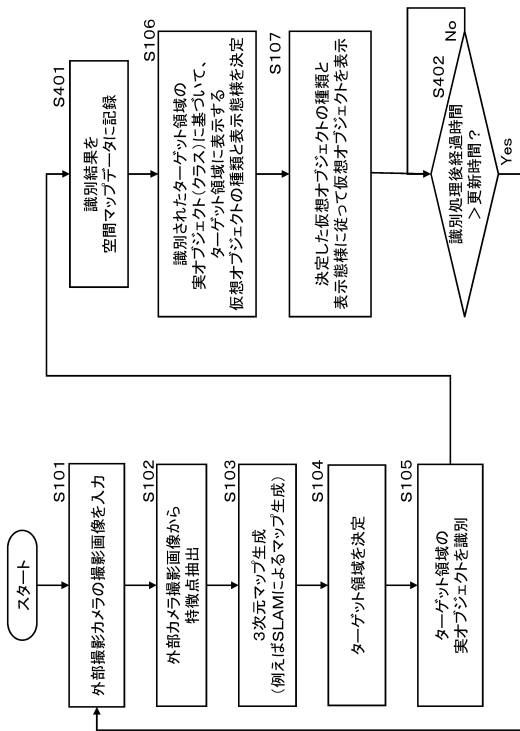
【図 15】



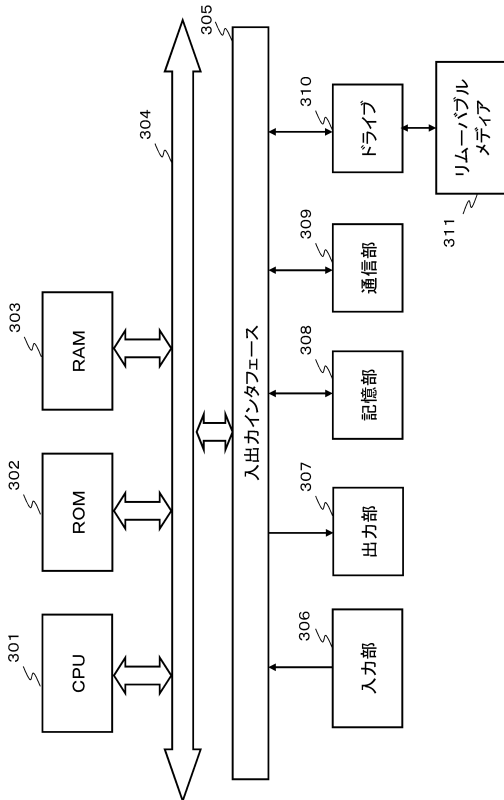
【図 16】



【図 17】



【図 18】



## フロントページの続き

グループ株式会社内

- (72)発明者 青木 秀憲  
東京都港区港南1丁目7番1号 ソニーグループ株式会社内
- (72)発明者 荒井 富士夫  
東京都港区港南1丁目7番1号 ソニーグループ株式会社内
- (72)発明者 永野 京二郎  
東京都港区港南1丁目7番1号 ソニーグループ株式会社内
- (72)発明者 深澤 遼  
東京都港区港南1丁目7番1号 ソニーグループ株式会社内
- (72)発明者 藤澤 春香  
東京都港区港南1丁目7番1号 ソニーグループ株式会社内

審査官 菊池 伸郎

- (56)参考文献 特開2015-060579(JP,A)  
国際公開第2019/016870(WO,A1)  
特開2019-012537(JP,A)  
国際公開第2018/106542(WO,A1)  
特開2017-091433(JP,A)  
特開2017-049658(JP,A)  
特開2014-206904(JP,A)

(58)調査した分野 (Int.Cl., DB名)

G06V 10/00-40/70  
G06T 1/00-7/90  
G06T 19/00  
G06F 3/04815  
G06F 3/01