



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0044489  
(43) 공개일자 2022년04월08일

(51) 국제특허분류(Int. Cl.)  
G06T 19/00 (2011.01) G06F 3/01 (2006.01)  
G06F 3/0481 (2022.01)  
(52) CPC특허분류  
G06T 19/00 (2013.01)  
G06F 3/01 (2013.01)  
(21) 출원번호 10-2022-7001726  
(22) 출원일자(국제) 2020년07월07일  
심사청구일자 없음  
(85) 번역문제출일자 2022년01월18일  
(86) 국제출원번호 PCT/JP2020/026611  
(87) 국제공개번호 WO 2021/029164  
국제공개일자 2021년02월18일  
(30) 우선권주장  
JP-P-2019-147046 2019년08월09일 일본(JP)

(71) 출원인  
소니그룹주식회사  
일본국 도쿄도 미나토쿠 코난 1-7-1  
(72) 발명자  
고토 도모히코  
일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니  
주식회사 내  
아오키 히데노리  
일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니  
주식회사 내  
(뒷면에 계속)  
(74) 대리인  
장수길, 이중희

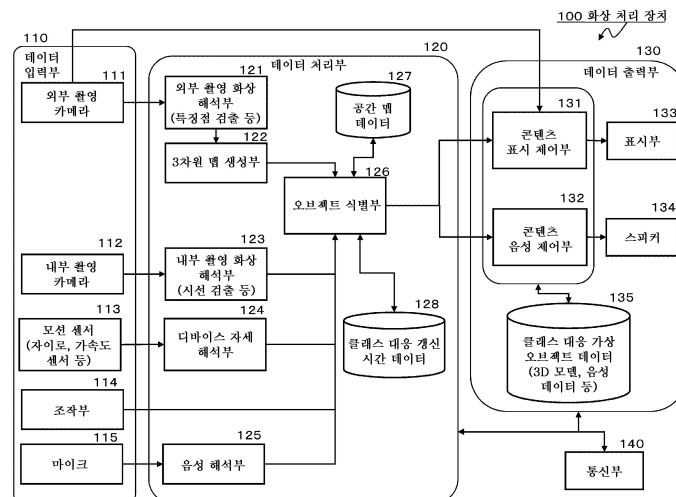
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 화상 처리 장치 및 화상 처리 방법, 그리고 프로그램

(57) 요약

가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트의 종류에 따라 표시할 가상 오브젝트의 선택이나 표시 양태 변경을 행하는 장치, 방법을 제공한다. 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별부와, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR 화상을 생성하는 콘텐츠 표시 제어부를 갖는다. 오브젝트 식별부는, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트를 식별하고, 콘텐츠 표시 제어부는, 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 처리나 표시 양태의 변경 처리를 행한다.

대표도



(52) CPC특허분류

**G06F 3/0481** (2013.01)

(72) 발명자

**아라이 후지오**

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 주식회사 내

**나가노 게이지로**

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 주식회사 내

**후카자와 료**

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 주식회사 내

**후지사와 하루카**

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 주식회사 내

## 명세서

### 청구범위

#### 청구항 1

실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별부와,

실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어부를 갖고,

상기 오브젝트 식별부는,

상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행하고,

상기 콘텐츠 표시 제어부는,

상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 화상 처리 장치.

#### 청구항 2

제1항에 있어서,

상기 오브젝트 식별부는,

화상 인식 처리에 의해 오브젝트 식별 처리를 실행하는 화상 처리 장치.

#### 청구항 3

제2항에 있어서,

상기 오브젝트 식별부는,

시맨틱 · 세그멘테이션 처리를 적용하여 오브젝트 식별 처리를 실행하는 화상 처리 장치.

#### 청구항 4

제1항에 있어서,

상기 오브젝트 식별부는,

상기 가상 오브젝트의 표시 영역으로 할 타깃 영역을 결정하고, 결정된 타깃 영역의 실제 오브젝트의 식별 처리를 실행하는 화상 처리 장치.

#### 청구항 5

제4항에 있어서,

상기 오브젝트 식별부는,

상기 타깃 영역을 유저 동작, 유저 시선, 유저 조작, 유저 위치 또는 유저 자세 중 적어도 어느 것에 기초하여 결정하는 화상 처리 장치.

#### 청구항 6

제4항에 있어서,

상기 오브젝트 식별부는,

상기 타깃 영역을 수평면 영역으로부터 선택하여 결정하는 화상 처리 장치.

#### 청구항 7

제6항에 있어서,  
상기 콘텐츠 표시 제어부는,  
상기 가상 오브젝트를 상기 수평면 영역에 접하도록 표시하는 화상 처리 장치.

#### 청구항 8

제4항에 있어서,  
상기 오브젝트 식별부는,  
상기 오브젝트 식별 처리를 리얼타임 처리로서 실행하는 화상 처리 장치.

#### 청구항 9

제1항에 있어서,  
상기 오브젝트 식별부는,  
오브젝트 종류에 따라 미리 규정된 시간 간격으로 오브젝트 식별 처리를 반복해서 실행하는 화상 처리 장치.

#### 청구항 10

제1항에 있어서,  
상기 화상 처리 장치는,  
카메라 촬영 화상에 기초하는 실세계의 3차원 맵을 생성하는 3차원 맵 생성부를 갖고,  
상기 오브젝트 식별부는,  
상기 3차원 맵을 이용하여 상기 가상 오브젝트의 표시 영역으로 할 타깃 영역을 결정하는 화상 처리 장치.

#### 청구항 11

제10항에 있어서,  
상기 3차원 맵 생성부는,  
SLAM(Simultaneous Localization and Mapping) 처리에 의해, 실세계의 3차원 맵을 생성하는 화상 처리 장치.

#### 청구항 12

제1항에 있어서,  
상기 콘텐츠 표시 제어부는,  
상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택함과 함께,  
상기 오브젝트 식별 결과에 따라서, 표시할 가상 오브젝트의 표시 양태도 제어하는 화상 처리 장치.

#### 청구항 13

제1항에 있어서,  
상기 화상 처리 장치는,  
음성 출력 제어를 실행하는 콘텐츠 음성 제어부를 더 갖고,  
상기 콘텐츠 음성 제어부는,  
상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라서 출력할 음성을 결정하여 출력하는 화상 처리 장치.

#### 청구항 14

화상 처리 장치에 있어서 실행하는 화상 처리 방법이며,  
 오브젝트 식별부가, 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별 처리 스텝을 실행하고,  
 콘텐츠 표시 제어부가, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는  
 콘텐츠 표시 제어 스텝을 실행하고,  
 상기 오브젝트 식별 처리 스텝은,  
 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행하는 스텝이며,  
 상기 콘텐츠 표시 제어 스텝은,  
 상기 오브젝트 식별 처리 스텝에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는  
 스텝을 실행하는 화상 처리 방법.

## 청구항 15

화상 처리 장치에 있어서 화상 처리를 실행시키는 프로그램이며,  
 오브젝트 식별부에, 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별 처리 스텝을 실행시키고,  
 콘텐츠 표시 제어부에, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는  
 콘텐츠 표시 제어 스텝을 실행시키고,  
 상기 오브젝트 식별 처리 스텝에 있어서는,  
 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행키고,  
 상기 콘텐츠 표시 제어 스텝에 있어서는,  
 상기 오브젝트 식별 처리 스텝에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는  
 스텝을 실행시키는 프로그램.

## 발명의 설명

### 기술 분야

[0001] 본 개시는, 화상 처리 장치 및 화상 처리 방법, 그리고 프로그램에 관한 것이다. 더욱 상세하게는, 현실에 관찰 가능한 실제 오브젝트에 캐릭터 화상 등의 가상 콘텐츠를 중첩하여 표시하는 확장 현실(AR: Augmented Reality) 화상을 생성하여 출력하는 화상 처리 장치 및 화상 처리 방법, 그리고 프로그램에 관한 것이다.

### 배경 기술

- [0002] 실공간에서 관찰 가능한 실제 오브젝트나, 실제 오브젝트와 화상에 가상 오브젝트를 중첩해서 표시한 화상을 확장 현실(AR: Augmented Reality) 화상이라 칭한다.
- [0003] AR 화상을 이용한 콘텐츠나 게임에서 이용되는 가상 오브젝트에는 여러 종류가 있지만, 예를 들어 인간과 동일한 움직임을 하는 가상 오브젝트, 즉 캐릭터가 이용되는 경우가 많다.
- [0004] AR 화상은, 예를 들어 유저의 눈에 장착하는 헤드·마운트·디스플레이(HMD)나, 스마트폰(smartphone) 등의 휴대 단말기 등을 이용하여 표시된다.
- [0005] 유저는, AR 화상을 시청함으로써, 예를 들어 AR 화상 중의 표시 캐릭터가 실세계에 존재하는 듯한 감각을 즐길 수 있다.
- [0006] AR 화상 중에 캐릭터를 표시시키는 경우, 예를 들어 게임 프로그램 등, 콘텐츠 재생 프로그램에 따른 표시 처리가 이루어진다.
- [0007] 구체적으로는, 프로그램에 기록되어 있는 캐릭터 출력 조건을 충족하는 경우에 프로그램에 규정된 수순으로 캐릭터가 표시된다.
- [0008] 그러나, 이와 같은 프로그램에 규정된 캐릭터 표시를 행하면, 항상 동일한 처리가 반복되게 되어, 재미가 희석

되어 버린다.

[0009] 한편, 근년, 화상 내의 오브젝트를 식별하는 기술로서 시맨틱·세그멘테이션에 대한 연구나 이용이 진행되고 있다. 시맨틱·세그멘테이션은, 예를 들어 카메라 촬영 화상에 포함되는 다양한 오브젝트, 예를 들어 사람, 차량, 빌딩, 도로, 나무 등, 오브젝트의 종류를 식별하는 기술이다.

[0010] 또한, 이 시맨틱·세그멘테이션에 대하여 개시된 문헌으로서, 예를 들어 특허문헌 1(일본 특허 공개 제2015-207291호 공보)이 있다.

## 선행기술문헌

### 특허문헌

[0011] (특허문헌 0001) 일본 특허 공개 제2015-207291호 공보

## 발명의 내용

### 해결하려는 과제

[0012] 본 개시는, 실제 오브젝트의 종류에 따라 표시할 캐릭터의 제어를 행하는 화상 처리 장치 및 화상 처리 방법, 그리고 프로그램을 제공하는 것이다.

[0013] 본 개시의 일 실시예에 있어서는, 상술한 시맨틱·세그멘테이션 등의 오브젝트 식별 처리에 의해 화상의 배경 오브젝트를 식별하고, 식별 결과에 따라 표시할 캐릭터를 제어하는 화상 처리 장치 및 화상 처리 방법, 그리고 프로그램을 제공하는 것이다.

### 과제의 해결 수단

[0014] 본 개시의 제1 측면은,

[0015] 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별부와,

[0016] 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어부를 갖고,

[0017] 상기 오브젝트 식별부는,

[0018] 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행하고,

[0019] 상기 콘텐츠 표시 제어부는,

[0020] 상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 화상 처리 장치에 있다.

[0021] 본 개시의 제2 측면은,

[0022] 화상 처리 장치에 있어서 실행하는 화상 처리 방법이며,

[0023] 오브젝트 식별부가, 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별 처리 스텝을 더 실행하고,

[0024] 콘텐츠 표시 제어부가, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어 스텝을 실행하고,

[0025] 상기 오브젝트 식별 처리 스텝은,

[0026] 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행하는 스텝이며,

[0027] 상기 콘텐츠 표시 제어 스텝은,

[0028] 상기 오브젝트 식별 처리 스텝에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 스텝을 실행하는 화상 처리 방법에 있다.

- [0029] 또한, 본 개시의 제3 측면은,
- [0030] 화상 처리 장치에 있어서 화상 처리를 실행시키는 프로그램이며,
- [0031] 오브젝트 식별부에, 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별 처리 스텝을 실행시키고,
- [0032] 콘텐츠 표시 제어부에, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어 스텝을 실행시키고,
- [0033] 상기 오브젝트 식별 처리 스텝에 있어서는,
- [0034] 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행시키고,
- [0035] 상기 콘텐츠 표시 제어 스텝에 있어서는,
- [0036] 상기 오브젝트 식별 처리 스텝에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 스텝을 실행시키는 프로그램에 있다.
- [0037] 또한, 본 개시의 프로그램은, 예를 들어 다양한 프로그램·코드를 실행 가능한 정보 처리 장치나 컴퓨터·시스템에 대하여, 컴퓨터 가독 형식으로 제공하는 기억 매체, 통신 매체에 의해 제공 가능한 프로그램이다. 이와 같은 프로그램을 컴퓨터 가독 형식으로 제공함으로써, 정보 처리 장치나 컴퓨터·시스템상에서 프로그램에 따른 처리가 실현된다.
- [0038] 본 개시의 또 다른 목적, 특징이나 이점은, 후술하는 본 개시의 실시예나 첨부하는 도면에 기초하는 보다 상세한 설명에 의해 밝혀질 것이다. 또한, 본 명세서에 있어서 시스템이란, 복수의 장치의 논리적 집합 구성이며, 각 구성의 장치가 동일 하우징 내에 있는 것에 한정되지는 않는다.
- [0039] 본 개시의 일 실시예의 구성에 의하면, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트의 종류에 따라 표시할 가상 오브젝트의 선택이나 표시 양태 변경을 행하는 장치, 방법이 실현된다.
- [0040] 구체적으로는, 예를 들어 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별부와, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR 화상을 생성하는 콘텐츠 표시 제어부를 갖는다. 오브젝트 식별부는, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트를 식별하고, 콘텐츠 표시 제어부는, 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 처리나 표시 양태의 변경 처리를 행한다.
- [0041] 본 구성에 의해, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트의 종류에 따라 표시할 가상 오브젝트의 선택이나 표시 양태 변경을 행하는 장치, 방법이 실현된다.
- [0042] 또한, 본 명세서에 기재된 효과는 어디까지나 예시이지 한정되는 것은 아니며, 또한 부가적인 효과가 있어도 된다.

### 도면의 간단한 설명

- [0043] 도 1은 본 개시의 화상 처리 장치의 구성예와 실행하는 처리에 대하여 설명하는 도면이다.
- 도 2는 본 개시의 화상 처리 장치가 실행하는 처리에 대하여 설명하는 도면이다.
- 도 3은 본 개시의 화상 처리 장치가 실행하는 처리에 대하여 설명하는 도면이다.
- 도 4는 본 개시의 화상 처리 장치가 실행하는 처리에 대하여 설명하는 도면이다.
- 도 5는 본 개시의 화상 처리 장치의 구성예와 실행하는 처리에 대하여 설명하는 도면이다.
- 도 6은 본 개시의 화상 처리 장치의 구성예와 실행하는 처리에 대하여 설명하는 도면이다.
- 도 7은 본 개시의 화상 처리 장치가 실행하는 처리에 대하여 설명하는 도면이다.
- 도 8은 본 개시의 화상 처리 장치가 실행하는 처리에 대하여 설명하는 도면이다.
- 도 9는 본 개시의 화상 처리 장치의 일 구성예에 대하여 설명하는 도면이다.
- 도 10은 공간 맵 데이터의 데이터 구성예에 대하여 설명하는 도면이다.
- 도 11은 클래스 대응 갱신 시간 데이터의 데이터 구성예에 대하여 설명하는 도면이다.

도 12는 클래스 대응 가상 오브젝트 데이터의 데이터 구성예에 대하여 설명하는 도면이다.

도 13은 본 개시의 화상 처리 장치가 실행하는 처리의 시퀀스에 대하여 설명하는 흐름도를 나타내는 도면이다.

도 14는 본 개시의 화상 처리 장치가 실행하는 타깃 영역 결정 처리의 시퀀스에 대하여 설명하는 흐름도를 나타내는 도면이다.

도 15는 본 개시의 화상 처리 장치가 실행하는 타깃 영역 결정 처리의 시퀀스에 대하여 설명하는 흐름도를 나타내는 도면이다.

도 16은 본 개시의 화상 처리 장치가 실행하는 처리의 시퀀스에 대하여 설명하는 흐름도를 나타내는 도면이다.

도 17은 본 개시의 화상 처리 장치가 실행하는 처리의 시퀀스에 대하여 설명하는 흐름도를 나타내는 도면이다.

도 18은 본 개시의 화상 처리 장치의 하드웨어 구성예에 대하여 설명하는 도면이다.

### 발명을 실시하기 위한 구체적인 내용

- [0044] 이하, 도면을 참조하면서 본 개시의 화상 처리 장치 및 화상 처리 방법, 그리고 프로그램의 상세에 대하여 설명한다. 또한, 설명은 이하의 항목에 따라서 행한다.
- [0045] 1. 본 개시의 화상 처리 장치가 실행하는 처리의 개요에 대하여
- [0046] 2. 본 개시의 화상 처리 장치의 구성예에 대하여
- [0047] 3. 본 개시의 화상 처리 장치가 실행하는 처리의 시퀀스에 대하여
- [0048] 3-(1) 화상 처리 장치가 실행하는 기본 처리 시퀀스
- [0049] 3-(2) 타깃 영역의 설정 영역을 거의 수평한 면에 설정하는 처리의 시퀀스
- [0050] 3-(3) 실제 오브젝트 식별 처리의 갱신 시퀀스
- [0051] 4. 화상 처리 장치의 하드웨어 구성예에 대하여
- [0052] 5. 본 개시의 구성의 요약
- [0053] [1. 본 개시의 화상 처리 장치가 실행하는 처리의 개요에 대하여]
- [0054] 우선, 도 1 이하를 참조하여, 본 개시의 화상 처리 장치가 실행하는 처리의 개요에 대하여 설명한다.
- [0055] 도 1에는, 본 개시의 화상 처리 장치의 일례로서, 헤드·마운트·디스플레이(HMD)형의 광투과형 AR 화상 표시 디바이스(10)를 나타내고 있다.
- [0056] 유저는, 유저의 눈을 덮도록 헤드·마운트·디스플레이(HMD)형의 광투과형 AR 화상 표시 디바이스(10)를 장착한다.
- [0057] 광투과형 AR 화상 표시 디바이스(10)는 광투과형의 표시부(디스플레이)를 갖는다. 광투과형의 표시부(디스플레이)를 유저의 눈 앞의 위치에 설정하도록 유저에 장착한다.
- [0058] 유저는, 광투과형 AR 화상 표시 디바이스(10)의 광투과형의 표시부(디스플레이)를 통해 외부의 실제 오브젝트를 그대로 관찰할 수 있다.
- [0059] 또한, 광투과형의 표시부(디스플레이)에, 가상 오브젝트, 예를 들어 캐릭터 화상 등의 가상 오브젝트 화상이 표시된다.
- [0060] 유저는, 광투과형 AR 화상 표시 디바이스(10)를 통해 외부의 실제 오브젝트와 캐릭터 등의 가상 오브젝트 화상을 함께 관찰하는 것이 가능하게 되어, 마치 캐릭터 등의 가상 오브젝트가 실제계상에 존재하는 듯한 감각을 체험할 수 있다.
- [0061] 도 1의 우측에는, 유저가, 광투과형 AR 화상 표시 디바이스(10)를 통해 관찰 가능한 화상의 예를 나타내고 있다.
- [0062] (a) 관찰 화상에 1은, 광투과형 AR 화상 표시 디바이스(10)를 통해 관찰되는 외부의 실제 오브젝트를 포함하는 투과 관찰 화상(21)에 의해 구성되어 있다. 이 화상에 1에는, 가상 오브젝트는 표시되어 있지 않다.

- [0063] 한편, (b) 관찰 화상에 2는, 광투과형 AR 화상 표시 디바이스(10)를 통해 관찰되는 외부의 실제 오브젝트를 포함하는 투과 관찰 화상(21)에 함께, 캐릭터 화상 등의 가상 오브젝트 화상(22)을 표시한 화상예이다. 이 화상에 2는, 사용자가 실제 오브젝트와 가상 오브젝트를 함께 관찰 가능한 화상이다.
- [0064] 진술한 바와 같이, 실공간에서 관찰 가능한 실제 오브젝트나, 실제 오브젝트와 화상에 가상 오브젝트를 중첩해서 표시한 화상을 확장 현실(AR: Augmented Reality) 화상이라 칭한다.
- [0065] 본 개시의 화상 처리 장치는, 이 AR 화상의 표시 처리를 행하는 장치이다.
- [0066] 본 개시의 화상 처리 장치, 예를 들어 도 1에 도시한 광투과형 AR 화상 표시 디바이스(10)는, AR 화상 표시에 있어서의 가상 오브젝트의 표시 제어를 실행한다.
- [0067] 본 개시의 화상 처리 장치, 예를 들어 도 1에 도시한 광투과형 AR 화상 표시 디바이스(10)가 실행하는 구체적인 처리는, 예를 들어 이하의 처리이다.
- [0068] (a) 자기 위치 추정과 환경 3차원 맵 생성을 행하는 SLAM(Simultaneous Localization and Mapping) 처리 등을 적용하여, 사용자가 광투과형 AR 화상 표시 디바이스(10)를 통해 관찰하고 있는 실세계의 3차원 맵을 생성하는 처리.
- [0069] (b) 시맨틱·세그멘테이션 등의 오브젝트 식별 처리에 의해 실세계에 포함되는 오브젝트를 식별하는 처리.
- [0070] (c) 실세계의 오브젝트 식별 결과에 따라서, 표시할 가상 오브젝트(캐릭터 등)를 선택하고, 또한 가상 오브젝트의 표시 양태를 제어하는 처리.
- [0071] 본 개시의 화상 처리 장치는, 예를 들어 이들 처리를 실행한다.
- [0072] 도 2 이하를 참조하여, 본 개시의 화상 처리 장치가 실행하는 가상 오브젝트의 표시 제어에 대하여 설명한다.
- [0073] 도 2의 좌측에 나타난 도면은, 광투과형 AR 화상 표시 디바이스(10)를 장착한 사용자가 공원을 돌아다니고, 연못 근처에서, 연못의 수면을 보면서, 연못의 수면 방향으로 「포인팅」을 행한 상태를 나타내고 있다.
- [0074] 광투과형 AR 화상 표시 디바이스(10)는, 카메라를 갖고 있으며, 사용자가 포인팅을 행한 화상을 촬영하고, 촬영 화상이 디바이스 내부의 화상 해석부를 통해 3차원 맵 생성부에 입력된다.
- [0075] 화상 해석부는 촬영 화상으로부터 특징점을 추출하고, 3차원 맵 생성부는, 화상으로부터 추출된 특징점을 이용하여 실세계의 3차원 맵을 생성한다.
- [0076] 이 3차원 맵 생성 처리는, 예를 들어 SLAM(Simultaneous Localization and Mapping) 처리에 의해 리얼타임으로 실행된다.
- [0077] SLAM(Simultaneous Localization and Mapping) 처리는, 자기 위치 추정과 환경 3차원 맵의 생성을 동시에 병렬하여 실행하는 것을 가능하게 하는 처리이다.
- [0078] 또한, 3차원 맵 생성부가 생성한 3차원 맵은, 화상 처리 장치의 오브젝트 식별부에 입력된다.
- [0079] 오브젝트 식별부는, 3차원 맵 생성부가 생성한 3차원 맵을 사용하여, 유저의 포인팅 방향에 있는 실제 오브젝트 영역을 타깃 영역(11)으로서 결정한다. 또한, 이 타깃 영역(11)의 오브젝트를 식별한다.
- [0080] 오브젝트 식별부는, 예를 들어 시맨틱·세그멘테이션 처리를 적용하여 실세계의 오브젝트의 식별 처리를 행한다.
- [0081] 시맨틱·세그멘테이션은, 화상 인식 처리의 일종이며, 딥 러닝을 사용하여 화상 중의 물체가 화소 레벨에서 무엇인가를 인식하는 태스크를 행하는 방법이다. 예를 들어 다양한 실제의 오브젝트 형상 정보나 그 밖의 특징 정보를 등록한 오브젝트 식별용 사전 데이터(학습 완료 데이터)와, 예를 들어 카메라 촬영 화상 내의 오브젝트의 일치도에 기초하여, 화상의 구성 화소(픽셀) 각각이, 어느 오브젝트 카테고리에 속하는 화소인지를 식별하는 기술이다.
- [0082] 이 시맨틱·세그멘테이션에 의해, 카메라 촬영 화상에 포함되는 다양한 오브젝트, 예를 들어 사람, 차량, 빌딩, 도로, 나무, 연못, 잔디밭 등, 오브젝트의 종류를 식별할 수 있다.
- [0083] 도 2에 도시한 예에서는, 광투과형 AR 화상 표시 디바이스(10)의 화상 해석부는, 카메라 촬영 화상에 기초하여, 유저의 포인팅 방향의 타깃 영역(11)의 오브젝트가 「연못」이라고 판정한다.

- [0084] 또한, 광투과형 AR 화상 표시 디바이스(10)의 가상 오브젝트 표시 제어를 실행하는 콘텐츠 표시 제어부는, 타깃 영역(11)의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라 표시해야 할 가상 오브젝트(캐릭터 등)의 선택 처리나 표시 양태를 결정하여 표시한다.
- [0085] 도 2에 도시한 예에서는, 도 2 우측의 (b) 관찰 화상에 2에 나타난 바와 같이, 가상 오브젝트 화상(22)으로서, 「물의 요정 캐릭터」의 화상을 표시하는 제어를 행한다.
- [0086] 이것은, 유저의 포인팅 방향의 실제 오브젝트가, 「연못」이라는 해석 결과에 기초하는 표시 제어이다.
- [0087] 즉, 콘텐츠 표시 제어부는, 타깃 영역(11)의 오브젝트 식별 결과=「연못」에 따른 최적의 가상 오브젝트로서, 「물 요정 캐릭터」를 선택하여 표시하는 처리를 행한다.
- [0088] 또한, 가상 오브젝트 화상(22)은, 예를 들어 3D 콘텐츠 화상으로서 표시한다.
- [0089] 도 3은, 본 개시의 화상 처리 장치가 실행하는 다른 가상 오브젝트의 표시 제어예를 설명하는 도면이다.
- [0090] 도 3의 좌측에 나타난 도면도, 도 2와 마찬가지로, 광투과형 AR 화상 표시 디바이스(10)를 장착한 유저가 공원을 돌아다니고 있는 도면이다. 유저는, 공원의 잔디밭을 보면서, 잔디밭 방향으로 포인팅을 행하고 있다.
- [0091] 이 경우, 광투과형 AR 화상 표시 디바이스(10)의 오브젝트 식별부는, 카메라 촬영 화상의 해석에 의해, 유저의 포인팅 방향인 타깃 영역(11)의 오브젝트가 「잔디밭」이라는 해석 결과를 출력한다.
- [0092] 또한, 광투과형 AR 화상 표시 디바이스(10)의 콘텐츠 표시 제어부는, 타깃 영역(11)의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라 표시해야 할 가상 오브젝트(캐릭터 등)의 선택 처리나 표시 양태를 결정하여 표시한다.
- [0093] 도 3에 도시한 예에서는, 도 3 우측의 (b) 관찰 화상에 2에 나타난 바와 같이, 가상 오브젝트 화상(23)으로서, 「초원의 요정 캐릭터」의 화상을 표시하는 제어를 행한다.
- [0094] 이것은, 유저의 포인팅 방향의 실제 오브젝트가, 「잔디밭」이라는 해석 결과에 기초하는 표시 제어이다.
- [0095] 즉, 콘텐츠 표시 제어부는, 타깃 영역(11)의 오브젝트 식별 결과=「잔디밭」에 따른 최적의 가상 오브젝트로서, 「잔디밭의 요정 캐릭터」를 선택하여 표시하는 처리를 행한다.
- [0096] 도 4는, 본 개시의 화상 처리 장치가 실행하는 다른 가상 오브젝트의 표시 제어예를 설명하는 도면이다.
- [0097] 도 4의 좌측에 나타난 도면도, 도 2, 도 3과 마찬가지로, 광투과형 AR 화상 표시 디바이스(10)를 장착한 유저가 공원을 돌아다니고 있는 도면이다. 유저는, 공원의 나무를 보면서, 나무 방향으로 포인팅을 행하고 있다.
- [0098] 이 경우, 광투과형 AR 화상 표시 디바이스(10)의 오브젝트 식별부는, 카메라 촬영 화상의 해석에 의해, 유저의 포인팅 방향인 타깃 영역(11)의 오브젝트가 「나무」라는 해석 결과를 출력한다.
- [0099] 또한, 광투과형 AR 화상 표시 디바이스(10)의 콘텐츠 표시 제어부는, 타깃 영역(11)의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라 표시해야 할 가상 오브젝트(캐릭터 등)의 선택 처리나 표시 양태를 결정하여 표시한다.
- [0100] 도 4에 도시한 예에서는, 도 4 우측의 (b) 관찰 화상에 2에 나타난 바와 같이, 가상 오브젝트 화상(23)으로서, 「나무의 요정 캐릭터」의 화상을 표시하는 제어를 행한다.
- [0101] 이것은, 유저의 포인팅 방향의 실제 오브젝트가, 「나무」라는 해석 결과에 기초하는 표시 제어이다.
- [0102] 즉, 콘텐츠 표시 제어부는, 타깃 영역(11)의 오브젝트 식별 결과=「나무」에 따른 최적의 가상 오브젝트로서, 「나무의 요정 캐릭터」를 선택하여 표시하는 처리를 행한다.
- [0103] 이와 같이, 본 개시의 화상 처리 장치는, SLAM 처리 등을 사용하여 실세계의 3차원 형상 해석을 행하고, 실세계의 3차원 맵을 생성하고, 또한 시맨틱·세그멘테이션 등의 오브젝트 식별 처리에 의해 실세계의 3차원 맵 내의 타깃 영역의 오브젝트를 식별하고, 식별 결과에 따라 표시할 캐릭터 등의 가상 오브젝트의 표시 제어를 실행한다.
- [0104] 또한, 시맨틱·세그멘테이션 등의 오브젝트 식별 처리에 의해 해석할 실제 오브젝트의 대상은, 예를 들어 유저의 손가락에 의해 지정된 영역, 즉 타깃 영역만에 한정하여 처리를 행하는 것이 가능해진다. 이와 같이, 해석 범위를 한정함으로써 고속 처리가 실현된다.

- [0105] 또한, 본 개시의 화상 처리 장치는, 도 1을 참조하여 설명한 헤드·마운트·디스플레이(HMD)형의 광투과형 AR 화상 표시 디바이스(10)에 한정되지 않고, 다양한 표시부를 구비한 장치에 의해 구성할 수 있다.
- [0106] 예를 들어, 도 5에 도시한 카메라 촬영 화상 표시형 AR 화상 표시 디바이스(30)로 해도 된다. 도 5에 도시한 카메라 촬영 화상 표시형 AR 화상 표시 디바이스(30)는, 비투과형의 표시부(디스플레이)를 갖는다. 광투과형의 표시부(디스플레이)는, 유저에 의해 관찰 가능한 유저의 눈 앞의 위치에 설정되도록 유저에 장착된다.
- [0107] 카메라 촬영 화상 표시형 AR 화상 표시 디바이스(30)에 일체화된 카메라(31)의 촬영 화상, 즉 도 5에 도시한 카메라 촬영 화상(32)이 유저의 눈앞의 표시부에 표시된다. 즉 카메라(31)가 촬영하는 실제 오브젝트의 화상이 유저의 눈앞의 표시부에 표시되고, 유저는, 카메라 촬영 화상(32)을 봄으로써, 외부 경치를 확인할 수 있다.
- [0108] 또한, 표시부(디스플레이)에는, 가상 오브젝트, 예를 들어 캐릭터 화상 등의 가상 오브젝트 화상(22)이 표시된다.
- [0109] 유저는, 카메라 촬영 화상 표시형 AR 화상 표시 디바이스(30)의 표시부(디스플레이)에 표시되는 카메라 촬영 화상(32), 즉 실제 오브젝트 화상과, 캐릭터 등의 가상 오브젝트 화상(22)을 함께 관찰하는 것이 가능하게 되어, 마치 캐릭터 등의 가상 오브젝트가 실세계상에 존재하는 듯한 감각을 체험할 수 있다.
- [0110] 또한, 본 개시의 화상 처리 장치는, 도 6에 도시한 바와 같은 스마트폰(40) 등의 휴대형 표시 장치로 하는 것도 가능하다.
- [0111] 도 6에 도시한 스마트폰(40)은, 표시부를 갖고, 카메라(41)를 갖는다. 카메라(41)의 촬영 화상, 즉 도면에 도시한 카메라 촬영 화상(42)이 표시부에 표시된다. 즉 카메라(41)가 촬영하는 실제 오브젝트의 화상이 표시부에 표시되고, 유저는, 카메라 촬영 화상을 봄으로써, 외부 경치를 확인할 수 있다.
- [0112] 또한, 표시부(디스플레이)에는, 가상 오브젝트, 예를 들어 캐릭터 화상 등의 가상 오브젝트 화상이 표시된다.
- [0113] 유저는, 스마트폰(40)의 표시부(디스플레이)에 표시되는 카메라 촬영 화상, 즉 실제 오브젝트 화상과, 캐릭터 등의 가상 오브젝트 화상을 함께 관찰하는 것이 가능하게 되어, 마치 캐릭터 등의 가상 오브젝트가 실세계상에 존재하는 듯한 감각을 체험할 수 있다.
- [0114] 또한, 이 스마트폰(40)의 예에서는, 유저가 스마트폰(40)의 표시부가 있는 위치를 터치한 경우에, 화상 처리 장치(스마트폰)의 화상 해석부가, 터치 위치를 해석하고, 또한 터치 위치의 실제 오브젝트의 종류를 판별한다. 그 후, 화상 처리 장치(스마트폰)의 콘텐츠 표시 제어부가, 이 판별 결과에 따라서 캐릭터 등의 가상 오브젝트의 표시 제어를 실행한다.
- [0115] 상술한 바와 같이, 본 개시의 화상 처리 장치는, 가상 오브젝트의 표시 위치가 되는 타깃 영역에 있는 실제 오브젝트가 물인지, 풀인지, 나무인지 등의 오브젝트 식별을 행하고, 식별 결과에 따라 표시할 캐릭터 등의 가상 오브젝트를 선택하여 표시하는 처리 등을 행한다.
- [0116] 또한, 본 개시의 화상 처리 장치는, 타깃 영역에 있는 실제 오브젝트의 식별 결과에 따른 표시 가상 오브젝트의 선택 처리뿐만 아니라, 실제 오브젝트 식별 결과에 따라서 캐릭터 등의 가상 오브젝트의 표시 양태를 변경하는 처리도 실행한다.
- [0117] 도 7 이하를 참조하여, 실제 오브젝트 식별 결과에 따른 가상 오브젝트의 표시 양태의 변경 처리의 구체예에 대하여 설명한다.
- [0118] 도 7에는, 본 개시의 화상 처리 장치의 표시부에 표시되는 AR 화상의 일례로서, 실제 오브젝트인 연못에서 가상 오브젝트(캐릭터)가 튀어나오는 AR 화상의 표시예를 나타내고 있다.
- [0119] 본 개시의 화상 처리 장치의 콘텐츠 표시 제어부는, 타깃 영역에 표시한 캐릭터 등의 가상 오브젝트를 미리 설정된 프로그램에 따라서 이동시키는 처리나 움직이게 하는 처리를 행한다.
- [0120] 도 7에 도시한 예는, 타깃 영역(11)에 표시한 캐릭터를 상측 방향으로 이동시키는 표시를 행하는 예를 나타내고 있다.
- [0121] 도 7의 (1) 연못에서 가상 오브젝트(캐릭터)가 튀어나오는 순간-1은, 가상 오브젝트(캐릭터)의 상반부만이 물 위에 표시되고, 하반부가 물속에 있는 상태이다.
- [0122] 이와 같은 상태에서는, 화상 처리 장치의 콘텐츠 표시 제어부는, 도 7의 (1)에 나타난 바와 같이, 물 위의 가상

오브젝트 화상(50)과, 물속의 가상 오브젝트 화상(51)을 다른 표시 양태로 표시한다.

- [0123] 즉, 물 위의 가상 오브젝트 화상(50)에 대해서는, 통상의 윤곽이 명확한 화상으로서 표시하지만, 물속의 가상 오브젝트 화상(51)에 대해서는, 물속에 존재하도록 3차원적인 변형을 갖는 화상으로서 표시한다.
- [0124] 또한, 화상 처리 장치의 콘텐츠 음성 제어부는, 캐릭터가 물 위로 이동할 때의 효과음으로서, 물소리(철벽철벽 등)를, 스피커를 통해 출력한다.
- [0125] 도 7의 (2) 연못에서 가상 오브젝트(캐릭터)가 튀어나오는 순간-2는, 도 7의 (1)의 후의 AR 화상 표시예이다. 이 시점에서는, 가상 오브젝트(캐릭터)의 전신이 물 위에 표시되어 있다. 이 상태에서는, 물 위의 가상 오브젝트 화상(50)이 캐릭터 전체의 화상이 되고, 화상 처리 장치의 콘텐츠 표시 제어부는, 캐릭터 전체 화상을 윤곽이 명확한 화상으로서 표시한다.
- [0126] 도 8은, 실제 오브젝트의 식별 결과에 따라서 캐릭터 등의 가상 오브젝트 표시 양태를 변경하는 처리의 다른 구체예에 대하여 설명하는 도면이다.
- [0127] 도 8에는, 본 개시의 화상 처리 장치의 표시부에 표시되는 AR 화상의 일례로서, 가상 오브젝트(캐릭터)의 그림자를 표시한 AR 화상의 표시예를 나타내고 있다.
- [0128] 즉, 콘텐츠 표시 제어부가 타깃 영역(11)에 캐릭터를 표시할 때, 캐릭터의 그림자의 표시도 함께 실행한다.
- [0129] 도 8에 도시한 「(1) 그림자가 비치는 면이 평면인 경우의 캐릭터 그림자의 표시예」는,
- [0130] 가상 오브젝트(캐릭터) 화상(50)의 그림자가 비치는 면이, 실내의 바닥이나, 외부의 보도 등, 평면인 경우의 그림자의 표시예이다.
- [0131] 이와 같이 가상 오브젝트(캐릭터) 화상(50)의 그림자가 비치는 면이 평면인 경우, 화상 처리 장치의 콘텐츠 표시 제어부는, 타깃 영역(11)에 3차원 캐릭터인 가상 오브젝트 화상(50)을 표시할 때, 가상 오브젝트 화상(50)의 그림자를 나타내는 가상 오브젝트 그림자 화상(52)을 윤곽이 명확한 화상으로서 표시한다.
- [0132] 한편, 도 8에 도시한 「(2) 그림자가 비치는 면이 비평면(모래밭 등)인 경우의 캐릭터 그림자의 표시예」는,
- [0133] 가상 오브젝트(캐릭터) 화상(50)의 그림자가 비치는 면이, 예를 들어 모래밭 등, 평면이 아닌 경우의 그림자의 표시예이다.
- [0134] 이와 같이 가상 오브젝트(캐릭터) 화상(50)의 그림자가 비치는 면이 모래밭 등, 요철이 있는 면의 경우, 화상 처리 장치의 콘텐츠 표시 제어부는, 타깃 영역(11)에 3차원 캐릭터인 가상 오브젝트 화상(50)을 표시할 때, 가상 오브젝트 화상(50)의 그림자를 나타내는 가상 오브젝트 그림자 화상(52)을 윤곽이 명확치 않은, 요철이 있는 화상으로서 표시한다.
- [0135] 이와 같이, 본 개시의 화상 처리 장치의 콘텐츠 표시 제어부는, 가상 오브젝트를 표시할 타깃 영역의 실제 오브젝트 식별 결과에 따라서, 가상 오브젝트의 표시 양태도 변경하여 표시하는 제어를 행하고, 또한 콘텐츠 음성 제어부는 타깃 영역의 실제 오브젝트의 식별 결과에 따른 효과음의 출력 제어를 행한다.
- [0136] [2. 본 개시의 화상 처리 장치의 구성예에 대하여]
- [0137] 다음으로, 본 개시의 화상 처리 장치의 구성예에 대하여 설명한다.
- [0138] 전술한 바와 같이, 본 개시의 화상 처리 장치는, 도 1을 참조하여 설명한 광투과형 AR 화상 표시 디바이스(10), 혹은 도 5를 참조하여 설명한 카메라 촬영 화상 표시형 AR 화상 표시 디바이스(30), 혹은 도 6을 참조하여 설명한 스마트폰(40) 등의 휴대형 표시 장치 등, 다양한 형태를 갖는 장치로서 실현할 수 있다.
- [0139] 도 9는, 이들 다양한 형태를 취할 수 있는 본 개시의 화상 처리 장치의 일 구성예를 나타내는 블록도이다.
- [0140] 도 9에 도시한 화상 처리 장치(100)의 구성에 대하여 설명한다.
- [0141] 도 9에 도시한 바와 같이, 화상 처리 장치(100)는, 데이터 입력부(110), 데이터 처리부(120), 데이터 출력부(130), 통신부(140)를 갖는다.
- [0142] 데이터 입력부(110)는, 외부 촬영 카메라(111), 내부 촬영 카메라(112), 모션 센서(자이로, 가속도 센서 등)(113), 조작부(114), 마이크(115)를 갖는다.
- [0143] 데이터 처리부(120)는, 외부 촬영 화상 해석부(121), 3차원 맵 생성부(122), 내부 촬영 화상 해석부(123), 디바

이스 자세 해석부(124), 음성 해석부(125), 오브젝트 식별부(126), 공간 맵 데이터(127), 클래스 대응 갱신 시간 데이터(128)를 갖는다.

- [0144] 데이터 출력부(130)는, 콘텐츠 표시 제어부(131), 콘텐츠 음성 제어부(132), 표시부(133), 스피커(134), 클래스 대응 가상 오브젝트 데이터(3D 모델, 음성 데이터 등)(135)를 갖는다.
- [0145] 데이터 입력부(110)의 외부 촬영 카메라(111)는, 외부의 화상을 촬영한다. 예를 들어, HMD를 장착한 유저가 있는 환경에 있어서의 외부 경치 등을 촬영한다. 스마트폰 등의 휴대 단말기의 경우에는 스마트폰 등에 구비된 카메라이다.
- [0146] 내부 촬영 카메라(112)는, 기본적으로 HMD 고유의 구성이며, 유저의 시선 방향을 해석하기 위한 유저의 눈의 영역의 화상을 촬영한다.
- [0147] 모션 센서(자이로, 가속도 센서 등)(113)는, 예를 들어 HMD나 스마트폰 등, 화상 처리 장치(100) 본체의 자세나 움직임을 검출한다.
- [0148] 모션 센서(113)는, 예를 들어 자이로, 가속도 센서, 방위 센서, 단일 측위 센서, 관성 계측 장치(IMU) 등으로 구성된다.
- [0149] 조작부(114)는, 유저에 의해 조작 가능한 조작부이며, 예를 들어 타깃 영역의 입력이나, 그 밖의 처리 지시 등의 입력에 이용된다.
- [0150] 마이크(115)는, 유저에 의한 음성 입력에 의한 지시의 입력 등에 사용된다. 또한, 외부의 환경 소리의 입력에도 이용할 수 있다.
- [0151] 다음으로, 데이터 처리부(120)의 구성부에 대하여 설명한다.
- [0152] 외부 촬영 화상 해석부(121)는, 외부 촬영 카메라(111)에 의해 촬영된 외부의 촬영 화상을 입력하고, 외부 촬영 화상으로부터 특징점의 추출을 행한다.
- [0153] 이 특징점 추출 처리는, 3차원 맵의 생성에 이용하기 위한 특징점이며, 추출된 특징점 정보는, 외부 촬영 카메라(111)에 의해 촬영된 외부의 촬영 화상과 함께, 3차원 맵 생성부(122)에 입력한다.
- [0154] 3차원 맵 생성부(122)는, 외부 촬영 카메라(111)에 의해 촬영된 외부의 촬영 화상과, 외부 촬영 화상 해석부(121)가 추출한 특징점에 기초하여, 외부의 실제 오브젝트에 의해 구성되는 3차원 맵을 생성한다.
- [0155] 이 3차원 맵 생성 처리는, 예를 들어 SLAM(Simultaneous Localization and Mapping) 처리에 의해 리얼타임 처리로서 실행한다.
- [0156] 전술한 바와 같이, SLAM(Simultaneous Localization and Mapping) 처리는, 자기 위치 추정과 환경 3차원 맵의 생성을 동시에 병렬하여 실행하는 것을 가능하게 하는 처리이다.
- [0157] 3차원 맵 생성부(122)가 생성한 외부 환경의 3차원 맵 데이터는, 오브젝트 식별부(126)에 입력된다.
- [0158] 내부 촬영 화상 해석부(123)는, 내부 촬영 카메라(112)에 의해 촬영된 유저의 눈의 영역의 화상에 기초하여 유저의 시선 방향을 해석한다. 이 내부 촬영 화상 해석부(123)도, 전술한 내부 촬영 카메라(112)와 마찬가지로, 기본적으로 HMD 고유의 구성이다.
- [0159] 내부 촬영 화상 해석부(123)가 해석한 유저 시선 정보는, 오브젝트 식별부(126)에 입력된다.
- [0160] 디바이스 자세 해석부(124)는, 모션 센서(자이로, 가속도 센서 등)(113)에 의해 계측된 센서 검출 정보에 기초하여, HMD나 스마트폰 등, 화상 처리 장치(100) 본체의 자세나 움직임을 해석한다.
- [0161] 디바이스 자세 해석부(124)가 해석한 화상 처리 장치(100) 본체의 자세나 움직임 정보는, 오브젝트 식별부(126)에 입력된다.
- [0162] 음성 해석부(125)는, 마이크(115)로부터 입력하는 유저 음성이나 환경 소리의 해석을 행한다. 해석 결과는, 오브젝트 식별부(126)에 입력된다.
- [0163] 오브젝트 식별부(126)는, 3차원 맵 생성부(122)가 생성한 3차원 맵을 입력하고, 가상 오브젝트의 표시 영역에 설정하기 위한 타깃 영역을 결정하고, 또한 결정된 타깃 영역의 실제 오브젝트의 식별 처리를 실행한다. 예를 들어 타깃 영역이 연못, 나무 등인 오브젝트 식별 처리를 실행한다.

- [0164] 타깃 영역의 특정 처리는, 다양한 방법에 의해 실행 가능하다.
- [0165] 예를 들어, 3차원 맵에 포함되는 유저의 손가락 화상을 사용하여 행하는 것이 가능하다.
- [0166] 유저의 포인팅 방향의 연장선과, 3차원 맵상의 실제 오브젝트의 교점을 구하고, 예를 들어 교점을 중심으로 하는 미리 규정된 반경의 원형 영역을 타깃 영역으로서 결정한다.
- [0167] 또한, 타깃 영역의 지정은, 유저에 의한 포인팅 이외의 방법으로 행하는 것도 가능하다. 오브젝트 식별부(126)는, 타깃 영역의 결정에 사용하기 위한 정보로서, 이하 중 어느 정보를 이용할 수 있다.
- [0168] (a) 내부 촬영 화상 해석부(123)가 해석한 유저 시선 정보
- [0169] (b) 디바이스 자세 해석부(124)가 해석한 화상 처리 장치(100) 본체의 자세나 움직임 정보
- [0170] (c) 조작부(114)를 통해 입력하는 유저 조작 정보
- [0171] (d) 음성 해석부(125)가 해석한 유저 음성 정보
- [0172] 「(a) 내부 촬영 화상 해석부(123)가 해석한 유저 시선 정보」를 이용하는 경우, 오브젝트 식별부(126)는, 유저의 시선 방향의 연장선과, 3차원 맵상의 실제 오브젝트의 교점을 구하고, 예를 들어 교점을 중심으로 하는 소정 반경의 원형 영역을 타깃 영역으로서 결정한다.
- [0173] 「(b) 디바이스 자세 해석부(124)가 해석한 화상 처리 장치(100) 본체의 자세나 움직임 정보」를 이용하는 경우, 오브젝트 식별부(126)는, 유저에 의해 장착된 HMD, 혹은 유저가 갖는 스마트폰의 전방 방향의 연장선과, 3차원 맵상의 실제 오브젝트의 교점을 구하고, 예를 들어 교점을 중심으로 하는 소정 반경의 원형 영역을 타깃 영역으로서 결정한다.
- [0174] 「(c) 조작부(114)를 통해 입력하는 유저 조작 정보」를 이용하는 경우, 오브젝트 식별부(126)는, 예를 들어 화상 처리 장치(100)의 입력부를 통해 입력된 유저의 조작 정보에 기초하여 타깃 영역을 결정한다.
- [0175] 예를 들어, 앞에서 설명한 도 6에 도시한 스마트폰을 이용한 구성에 있어서, 유저의 손가락에 의한 화면 위치 지정 정보를 유저 조작 정보로서 입력하고, 이 지정 위치를 타깃 영역의 중심 위치로 한다는 처리가 가능하다.
- [0176] 또한, 이밖에, 조작부(114)로서, 화상 처리 장치(100)와 별체의 막대형의 지시 부재를 사용하여, 이 지시 부재의 지시 방향 정보를 오브젝트 식별부(126)에 입력하여, 지시 방향에 기초하여 타깃 영역을 결정하는 구성으로 해도 된다.
- [0177] 「(d) 음성 해석부(125)가 해석한 유저 음성 정보」를 이용하는 경우, 오브젝트 식별부(126)는, 예를 들어 유저에 의한 발화를 해석하여 타깃 영역을 결정한다.
- [0178] 예를 들어, 유저 발화가 「앞의 연못」이라고 하는 발화인 경우, 그 색을 타깃 영역으로서 결정한다.
- [0179] 또한, 오브젝트 식별부(126)는, 이들 이외의 타깃 영역 결정 처리를 행해도 된다. 예를 들어, 외부 촬영 카메라(111)에 의해 촬영된 화상에 기초하여 생성되는 3차원 맵이나, 모션 센서(113)의 검출 정보에 기초하여 지면, 바닥면, 수면 등의 수평면의 검출 처리를 실행하여, 촬영 화상의 중심 영역에 가장 가까운 수평면의 영역을 타깃 영역으로서 결정하는 처리를 행해도 된다.
- [0180] 또한, 예를 들어 유저가 가상의 불을 던지는 동작을 행하고, 이것을 외부 촬영 카메라(111)에 의해 촬영하고, 촬영 화상을 해석하여 불의 착지점을 해석하고, 이 착지점을 타깃 영역의 중심 위치로 설정한다는 처리를 행해도 된다.
- [0181] 오브젝트 식별부(126)는, 상술한 방법 중 어느 것을 이용하여, 가상 오브젝트 표시 영역으로 할 타깃 영역을 결정한다. 또한, 결정한 타깃 영역의 실제 오브젝트의 식별 처리를 실행한다. 예를 들어 타깃 영역이 연못, 나무 등인 오브젝트 식별 처리를 실행한다.
- [0182] 전술한 바와 같이, 이 실제 오브젝트의 오브젝트 식별 처리는, 예를 들어 시맨틱·세그멘테이션 처리를 적용하여 실행한다.
- [0183] 시맨틱·세그멘테이션은, 다양한 실제의 오브젝트 형상 정보나 그 밖의 특징 정보를 등록된 오브젝트 식별용 사전 데이터(학습 완료 데이터)와, 예를 들어 카메라 촬영 화상 내의 오브젝트의 일치도에 기초하여, 화상의 구성 화소(픽셀) 각각이, 어느 오브젝트 카테고리에 속하는 화소인지를 식별하는 기술이다.

- [0184] 이 시맨틱·세그멘테이션에 의해, 카메라 촬영 화상에 포함되는 다양한 오브젝트, 예를 들어 사람, 차량, 빌딩, 도로, 나무, 연못, 잔디밭 등, 오브젝트의 종류를 식별할 수 있다.
- [0185] 또한, 오브젝트 식별부(126)에 있어서 실행하는 실제 오브젝트의 식별 처리는, 타깃 영역만, 혹은 타깃 영역을 포함하는 주위 영역의 한정된 범위만을 대상으로 하여 실행한다. 이와 같은 한정 범위의 처리를 행함으로써 고속의 처리, 즉 리얼타임 처리가 가능해진다. 또한, 리얼타임 처리란, 예를 들어 유저에 의한 타깃 영역의 지정 후, 바로 실제 오브젝트의 식별 처리가 실행되는 것을 의미한다. 이에 의해, 예를 들어 유저가 타깃 영역을 관찰하고 있는 동안에 시간 지체 없이 오브젝트 식별이 완료된다.
- [0186] 오브젝트 식별부(126)가 해석한 타깃 영역의 실제 오브젝트 식별 결과는, 데이터 출력부(130)의 콘텐츠 표시 제어부(131)와, 콘텐츠 음성 제어부(132)에 입력된다.
- [0187] 데이터 출력부(130)의 콘텐츠 표시 제어부(131)는, 오브젝트 식별부(126)로부터 타깃 영역의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라 표시해야 할 가상 오브젝트(캐릭터 등)의 선택 처리나 표시 양태를 결정하여 표시부(133)에 표시한다.
- [0188] 구체적으로는, 예를 들어 앞에서 설명한 도 2 내지 도 4, 도 7, 도 8에 도시한 바와 같은 가상 오브젝트(캐릭터 등)의 표시 처리를 실행한다.
- [0189] 데이터 출력부(130)의 콘텐츠 음성 제어부(132)는, 오브젝트 식별부(126)로부터 타깃 영역의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라서 출력해야 할 음성을 결정하여 스피커(134)를 통해 출력한다.
- [0190] 구체적으로는, 예를 들어 앞에서 설명한 도 7에 도시한 바와 같이, 실제 오브젝트인 연못에서 가상 오브젝트가 출현하는 경우, 물소리를 출력하는 처리를 실행한다.
- [0191] 또한, 데이터 출력부(130)의 콘텐츠 표시 제어부(131)와, 콘텐츠 음성 제어부(132)는, 클래스 대응 가상 오브젝트 데이터(135)에 기록된 가상 오브젝트의 3D 콘텐츠와 다양한 음성 데이터를 취득하여 데이터 출력을 실행한다.
- [0192] 클래스 대응 가상 오브젝트 데이터(135)에는, 타깃 영역의 실제 오브젝트 식별 결과에 상응하는 실제 오브젝트의 종류(클래스)에 대응지어진 표시용 가상 오브젝트의 3D 콘텐츠와 다양한 음성 데이터가 기록되어 있다.
- [0193] 클래스 대응 가상 오브젝트 데이터(135)의 구체예에 대해서는 후술한다.
- [0194] 또한, 콘텐츠 표시부(131)는, 화상 처리 장치(100)가, 예를 들어 도 5를 참조하여 설명한 카메라 촬영 화상 표시형 AR 화상 표시 디바이스(30)나, 도 6을 참조하여 설명한 스마트폰(40)과 같이 카메라 촬영 화상을 표시하는 처리를 행하는 구성의 경우에는, 외부 촬영 카메라(11)의 촬영 화상을 입력하여, 이 촬영 화상에 가상 오브젝트를 중첩시킨 표시 화상을 생성하여 표시부(133)에 표시한다.
- [0195] 통신부(140)는, 예를 들어 외부 서버와 통신하여, 가상 콘텐츠인 캐릭터의 3D 콘텐츠를 취득한다. 그 밖에, 데이터 처리에 필요한 다양한 데이터나 파라미터를 외부 서버로부터 취득하는 구성으로 해도 된다.
- [0196] 또한, 오브젝트 식별부(126)는, 타깃 영역의 실제 오브젝트의 식별 처리를 행할 때, 식별 결과를 공간 맵 데이터(127)로서 기억부에 저장한다.
- [0197] 도 10에, 공간 맵 데이터(127)의 데이터 구성예를 나타낸다.
- [0198] 도 10에 도시한 바와 같이, 공간 맵 데이터는 이하의 각 데이터의 대응 데이터가 저장된다.
- [0199] (a) 타임 스탬프(sec)
- [0200] (b) 위치 정보
- [0201] (c) 클래스
- [0202] (d) 식별 처리 후 경과 시간(sec)
- [0203] (a) 타임 스탬프(sec)는, 오브젝트 식별 처리를 실행한 시간 정보이다.
- [0204] (b) 위치 정보는, 오브젝트 식별 대상이 된 실제 오브젝트의 위치 정보이다. 위치 정보의 기록 방식은, 다양한 방식이 이용 가능하다. 도면에 도시한 예는, 3차원 좌표(x,y,z)의 리스트에 의한 메쉬로서 기술한 예이다. 이 밖에, 예를 들어 타깃 영역의 중심 위치의 위치 정보를 기록하는 구성이어도 된다.

- [0205] (c) 클래스는, 오브젝트 식별 결과로서의 오브젝트 종류 정보이다.
- [0206] (d) 식별 처리 후 경과 시간(sec)은, 오브젝트 식별 처리의 완료 시부터의 경과 시간이다.
- [0207] 또한, 오브젝트 식별부(126)는, 타깃 영역이 결정된 직후에, 타깃 영역의 실제 오브젝트의 식별 처리를 실행하지만, 그 후에도, 그 영역에 대한 오브젝트 식별 처리를 반복해서 실행하여, 도 10에 도시한 공간 맵 데이터를, 축차 갱신한다.
- [0208] 단, 갱신 처리의 간격은, 식별된 실제 오브젝트의 종류(클래스)에 따라 다르다.
- [0209] 이 실제 오브젝트의 종류(클래스)에 따라 다른 갱신 시간의 규정 데이터가, 클래스 대응 갱신 시간 데이터(128)로서 미리 등록되어 있다.
- [0210] 도 11에, 클래스 대응 갱신 시간 데이터(128)의 데이터예를 나타낸다.
- [0211] 도 11에 도시한 바와 같이, 클래스 대응 갱신 시간 데이터(128)는, 이하의 각 데이터를 대응지은 데이터이다.
- [0212] (a) ID
- [0213] (b) 카테고리
- [0214] (c) 클래스
- [0215] (d) 갱신 시간(sec)
- [0216] (a) ID는, 등록 데이터의 식별자이다.
- [0217] (b) 카테고리는, 실제 오브젝트의 종류(클래스)의 카테고리이다.
- [0218] (c) 클래스는, 실제 오브젝트의 종류 정보이다.
- [0219] (d) 갱신 시간(sec)은, 실제 오브젝트의 식별 처리의 갱신 간격을 나타내는 시간이다.
- [0220] 예를 들어, ID001의 클래스(오브젝트 종류)=잔디밭의 경우, 갱신 시간은 3600sec(=1시간)이다. 잔디밭과 같은 오브젝트는, 시간 경과에 수반되는 변화가 적어, 갱신 시간이 길게 설정된다.
- [0221] 한편, 예를 들어 ID=004의 클래스(오브젝트 종류)=그림자의 경우, 갱신 시간은 2sec이다. 그림자와 같은 오브젝트는, 시간 경과에 수반되는 변화가 크기 때문에, 갱신 시간이 짧게 설정된다.
- [0222] 오브젝트 식별부(126)는, 이 클래스 대응 갱신 시간 데이터(128)의 데이터를 참조하여, 식별된 오브젝트에 대하여 규정된 시간 간격으로, 수시로 오브젝트 식별 처리를 반복하여, 실행한다. 새로운 식별 처리에 의해 검출된 실제 오브젝트는, 도 10을 참조하여 설명한 공간 맵 데이터(127)로서, 순차 등록한다.
- [0223] 또한, 전술한 바와 같이, 데이터 출력부(130)의 콘텐츠 표시 제어부(131)와, 콘텐츠 음성 제어부(132)는, 클래스 대응 가상 오브젝트 데이터(135)에 기록된 가상 오브젝트의 3D 콘텐츠와 다양한 음성 데이터를 취득하여 데이터 출력을 실행한다.
- [0224] 클래스 대응 가상 오브젝트 데이터(135)에는, 타깃 영역의 실제 오브젝트 식별 결과에 상당하는 실제 오브젝트의 종류(클래스)에 대응지어진 표시용 가상 오브젝트의 3D 콘텐츠와 다양한 음성 데이터가 기록되어 있다.
- [0225] 도 12를 참조하여, 클래스 대응 가상 오브젝트 데이터(135)의 구체예에 대하여 설명한다.
- [0226] 도 12에 도시한 바와 같이, 클래스 대응 가상 오브젝트 데이터(135)에는, 이하의 각 데이터가 대응지어 기록되어 있다.
- [0227] (a) 클래스
- [0228] (b) 가상 오브젝트 3D 모델(캐릭터 3D 모델)
- [0229] (c) 출력 음성
- [0230] (a) 클래스는, 실제 오브젝트의 종류 정보이다.
- [0231] (b) 가상 오브젝트 3D 모델(캐릭터 3D 모델)은, 각 클래스, 즉 타깃 영역의 실제 오브젝트의 종류에 따라 출력(표시)할 가상 오브젝트(캐릭터)의 3D 모델을 등록하고 있다. 또한, 도면에 도시한 예에서는 3D 모델 ID와, 3D 모델을 함께 기록한 구성으로 하고 있지만, 예를 들어 ID만을 기록하고, ID에 기초하여 다른 데이터베이스로부터

터 ID 대응의 3D 모델을 취득하는 구성으로 해도 된다.

- [0232] (c) 출력 음성은, 각 클래스, 즉 타깃 영역의 실제 오브젝트의 종류에 따라 출력할 음성 데이터를 등록하고 있다.
- [0233] 이와 같이, 클래스 대응 가상 오브젝트 데이터(135)에는, 타깃 영역의 실제 오브젝트 식별 결과에 상당하는 실제 오브젝트의 종류(클래스)에 대응지어진 표시용 가상 오브젝트의 3D 콘텐츠와 다양한 음성 데이터가 기록되어 있다.
- [0234] 또한, 클래스 대응 가상 오브젝트 데이터(135)에는, 추가로, 각 가상 오브젝트의 출력 양태 정보도 기록되어 있다. 예를 들어 앞에서 도 7이나 도 8을 참조하여 설명한 바와 같이, 타깃 영역의 실제 오브젝트가 물인 경우의 표시 양태 정보나, 모래밭인 경우의 표시 양태 정보 등의 정보도 기록되어 있다.
- [0235] 데이터 출력부(130)의 콘텐츠 표시 제어부(131)와, 콘텐츠 음성 제어부(132)는, 이 도 12에 도시한 바와 같은 데이터를 저장한 클래스 대응 가상 오브젝트 데이터(135)에 기록된 가상 오브젝트의 3D 콘텐츠와 다양한 음성 데이터를 취득하여 데이터 출력을 실행한다.
- [0236] [3. 본 개시의 화상 처리 장치가 실행하는 처리의 시퀀스에 대하여]
- [0237] 다음으로, 본 개시의 화상 처리 장치(100)가 실행하는 처리의 시퀀스에 대하여 설명한다.
- [0238] 또한, 이하에 나타낸 복수의 처리 시퀀스에 대하여, 순차 설명한다.
- [0239] (1) 화상 처리 장치가 실행하는 기본 처리 시퀀스
- [0240] (2) 타깃 영역의 설정 영역을 거의 수평한 면에 설정하는 처리의 시퀀스
- [0241] (3) 실제 오브젝트 식별 처리의 갱신 시퀀스
- [0242] (3-1) 화상 처리 장치가 실행하는 기본 처리 시퀀스)
- [0243] 우선, 도 13에 도시한 흐름도를 참조하여 본 개시의 화상 처리 장치(100)가 실행하는 기본적인 처리의 시퀀스에 대하여 설명한다.
- [0244] 또한, 도 13 이하에 나타낸 흐름도에 따른 처리는, 주로 화상 처리 장치(100)의 데이터 처리부(120)에 있어서 실행되는 처리이다. 데이터 처리부(120)는, 프로그램 실행 기능을 갖는 CPU를 구비하고, 기억부에 저장된 프로그램에 따라서 흐름에 따른 처리를 실행한다.
- [0245] 이하, 도 13에 도시한 흐름의 각 스텝의 처리에 대하여 설명한다.
- [0246] (스텝 S101)
- [0247] 우선, 화상 처리 장치(100)의 데이터 처리부(120)는, 스텝 S101에 있어서, 외부 촬영 카메라의 촬영 화상을 입력한다.
- [0248] (스텝 S102)
- [0249] 다음으로, 데이터 처리부(120)는, 스텝 S102에 있어서, 입력한 외부 촬영 카메라의 촬영 화상으로부터 특징점을 추출한다.
- [0250] 이 처리는, 도 9에 도시한 데이터 처리부(120)의 외부 촬영 화상 해석부(121)가 실행하는 처리이다.
- [0251] 외부 촬영 화상 해석부(121)는, 입력한 외부 촬영 카메라의 촬영 화상으로부터 특징점을 추출한다. 이 특징점 추출 처리는, 3차원 맵의 생성에 이용하기 위한 특징점이며, 추출한 특징점 정보는, 외부 촬영 카메라에 의해 촬영된 외부의 촬영 화상과 함께, 3차원 맵 생성부(122)에 입력한다.
- [0252] (스텝 S103)
- [0253] 다음으로, 데이터 처리부는, 스텝 S103에 있어서, 외부 촬영 카메라에 의해 촬영된 외부의 촬영 화상과 그 특징점 정보를 이용하여 3차원 맵을 생성한다.
- [0254] 이 처리는, 도 9에 도시한 데이터 처리부(120)의 3차원 맵 생성부(122)가 실행하는 처리이다.
- [0255] 3차원 맵 생성부(122)는, 외부 촬영 카메라(111)에 의해 촬영된 외부의 촬영 화상과, 외부 촬영 화상 해석부

(121)가 추출한 특징점에 기초하여, 외부의 실제 오브젝트에 의해 구성되는 3차원 맵을 생성한다.

- [0256] 이 3차원 맵 생성 처리는, 예를 들어 SLAM(Simultaneous Localization and Mapping) 처리에 의해 리얼타임 처리로서 실행한다.
- [0257] (스텝 S104)
- [0258] 다음으로, 데이터 처리부는, 스텝 S104에 있어서, 타깃 영역 결정 처리를 실행한다.
- [0259] 이 처리는, 도 9에 도시한 데이터 처리부(120)의 오브젝트 식별부(126)가 실행하는 처리이다.
- [0260] 오브젝트 식별부(126)는, 가상 오브젝트 표시 영역으로 할 타깃 영역을 결정한다.
- [0261] 이 타깃 영역의 결정 처리는, 전술한 바와 같이, 다양한 방법이 적용 가능하다.
- [0262] 예를 들어, 3차원 맵에 포함되는 유저의 손가락 화상을 사용하여 행하는 것이 가능하다.
- [0263] 즉, 유저의 포인팅 방향의 연장선과, 3차원 맵상의 실제 오브젝트의 교점을 구하고, 예를 들어 교점을 중심으로 하는 미리 규정된 반경의 원형 영역을 타깃 영역으로서 결정한다.
- [0264] 또한, 도 9에 도시한 데이터 입력부(110)의 각 구성부로부터의 입력 정보를 이용하여 타깃 영역을 결정하는 것도 가능하다. 구체적으로는, 이하의 입력 정보이다.
- [0265] (a) 내부 촬영 화상 해석부(123)가 해석한 유저 시선 정보
- [0266] (b) 디바이스 자세 해석부(124)가 해석한 화상 처리 장치(100) 본체의 자세나 움직임 정보
- [0267] (c) 조작부(114)를 통해 입력하는 유저 조작 정보
- [0268] (d) 음성 해석부(125)가 해석한 유저 음성 정보
- [0269] 예를 들어, 이들 입력 정보 중 어느 것을 사용하여 타깃 영역을 결정해도 된다.
- [0270] 대표적인 타깃 영역의 결정 시퀀스에 대하여, 도 14, 도 15를 참조하여 설명한다.
- [0271] 도 14의 (1)은, 유저의 포인팅 방향의 해석에 기초하는 타깃 영역의 결정 시퀀스이다. 유저의 포인팅 방향의 해석에 기초하는 타깃 영역 결정 처리는, 이하의 처리 시퀀스로 실행된다.
- [0272] 우선, 스텝 S211에서, 유저의 포인팅 방향을 해석한다. 이 해석 처리는, 3차원 맵 생성부(122)가 생성한 3차원 맵을 이용하여 실행한다.
- [0273] 다음으로, 스텝 S212에서, 유저의 포인팅 방향의 연장선으로 이루어지는 직선과, 실제 오브젝트의 교점을 검출한다. 이 처리도 3차원 맵 생성부(122)가 생성한 3차원 맵을 이용하여 실행한다.
- [0274] 마지막으로, 스텝 S213에 있어서, 유저의 포인팅 방향의 연장선으로 이루어지는 직선과, 실제 오브젝트의 교점을 중심으로 하는 원형 영역을 타깃 영역으로서 결정한다.
- [0275] 또한, 타깃 영역의 형상은 임의이며, 원형 외에, 직사각형 형상으로 해도 된다. 또한 타깃 영역의 크기도 임의이며, 다양한 크기로 설정 가능하다.
- [0276] 단, 형상, 크기는, 사전에 규정하고, 그 규정에 따라서 타깃 영역을 결정하는 것이 바람직하다.
- [0277] 도 14의 (2)는, 유저의 시선 방향 해석에 기초하는 타깃 영역의 결정 시퀀스이다. 유저 시선 방향의 해석에 기초하는 타깃 영역 결정 처리는, 이하의 처리 시퀀스로 실행된다.
- [0278] 우선, 스텝 S221에서, 유저의 시선 방향을 해석한다. 이 해석 처리는, 내부 촬영 카메라(112)의 촬영 화상에 기초하여, 내부 촬영 화상 해석부(123)가 실행한다.
- [0279] 다음으로, 스텝 S222에서, 유저 시선 방향의 연장선으로 이루어지는 직선과, 실제 오브젝트의 교점을 검출한다. 이 처리는 3차원 맵 생성부(122)가 생성한 3차원 맵을 이용하여 실행한다.
- [0280] 마지막으로, 스텝 S223에 있어서, 유저 시선 방향의 연장선으로 이루어지는 직선과, 실제 오브젝트의 교점을 중심으로 하는 원형 영역을 타깃 영역으로서 결정한다.
- [0281] 또한, 전술한 바와 같이, 타깃 영역의 형상, 크기는 다양한 설정이 가능하다.

- [0282] 도 15의 (3)은, 유저의 조작 정보의 해석에 기초하는 타깃 영역의 결정 시퀀스이다. 유저 조작 정보의 해석에 기초하는 타깃 영역 결정 처리는, 이하의 처리 시퀀스로 실행된다.
- [0283] 우선, 스텝 S231에서, 유저의 조작 정보를 해석한다. 예를 들어, 우선 도 6을 참조하여 설명한 스마트폰에 대한 터치 조작 등이다.
- [0284] 다음으로, 스텝 S232에서, 유저 조작 정보에 기초하는 실제 오브젝트 지정 위치를 검출한다. 이 처리는 예를 들어 유저의 손가락의 접촉 위치의 검출 처리로서 실행된다.
- [0285] 마지막으로, 스텝 S233에 있어서, 유저 조작 정보에 기초하는 실제 오브젝트 지정 위치를 중심으로 하는 원형 영역을 타깃 영역으로서 결정한다.
- [0286] 또한, 전술한 바와 같이, 타깃 영역의 형상, 크기는 다양한 설정이 가능하다.
- [0287] 도 15의 (4)는, 유저의 음성 정보의 해석에 기초하는 타깃 영역의 결정 시퀀스이다. 유저의 음성 정보의 해석에 기초하는 타깃 영역 결정 처리는, 이하의 처리 시퀀스로 실행된다.
- [0288] 우선, 스텝 S241에서, 유저의 발화 음성을 해석한다. 예를 들어, 「앞의 연못」 등의 유저 발화 음성을 해석한다.
- [0289] 다음으로, 스텝 S242에서, 유저의 음성 정보에 기초하는 실제 오브젝트 지정 위치를 검출한다.
- [0290] 마지막으로, 스텝 S243에 있어서, 유저 발화 음성에 기초하는 실제 오브젝트 지정 위치를 중심으로 하는 원형 영역을 타깃 영역으로서 결정한다.
- [0291] 또한, 전술한 바와 같이, 타깃 영역의 형상, 크기는 다양한 설정이 가능하다.
- [0292] 도 14, 도 15를 참조하여 설명한 이외에도, 예를 들어 이하와 같은 타깃 영역 결정 처리를 행해도 된다.
- [0293] (a) 디바이스 자세 해석부(124)가 해석한 화상 처리 장치(100) 본체의 자세나 움직임 정보를 이용한 타깃 영역 결정 처리.
- [0294] (b) 외부 촬영 카메라(111)에 의해 촬영된 화상에 기초하여 생성되는 3차원 맵이나, 모션 센서(113)의 검출 정보에 기초하여 지면, 바닥면, 수면 등의 수평면의 검출 처리를 실행하여, 촬영 화상의 중심 영역에 가장 가까운 수평면의 영역을 타깃 영역으로서 결정하는 처리.
- [0295] (c) 유저가 가상의 볼을 던지는 동작을 행하고, 이것을 외부 촬영 카메라(111)에 의해 촬영하고, 촬영 화상을 해석하여 볼의 착지점을 해석하고, 이 착지점을 타깃 영역의 중심 위치로서 결정하는 처리.
- [0296] (d) 그 밖에, 유저 동작, 유저 시선, 유저 조작, 유저 위치 또는 유저 자세 중 적어도 어느 것을 해석하여, 해석 결과에 기초하여 타깃 영역을 결정하는 처리.
- [0297] 도 13에 도시한 흐름으로 되돌아가서, 설명을 계속한다.
- [0298] 상술한 바와 같이, 화상 처리 장치(100)의 데이터 처리부(120)의 오브젝트 식별부(126)는, 스텝 S104에 있어서, 타깃 영역 결정 처리를 실행한다.
- [0299] (스텝 S105)
- [0300] 다음으로, 데이터 처리부는, 스텝 S105에 있어서, 타깃 영역의 실제 오브젝트를 식별한다.
- [0301] 구체적으로는, 타깃 영역이 연못, 나무 등인 오브젝트 식별 처리를 실행한다.
- [0302] 전술한 바와 같이, 이 실제 오브젝트의 오브젝트 식별 처리는, 예를 들어 시맨틱·세그멘테이션 처리를 적용하여 실행한다.
- [0303] 시맨틱·세그멘테이션은, 다양한 실제의 오브젝트 형상 정보나 그 밖의 특징 정보를 등록된 오브젝트 식별용 사전 데이터(학습 완료 데이터)와, 예를 들어 카메라 촬영 화상 내의 오브젝트의 일치도에 기초하여, 화상의 구성 화소(픽셀) 각각이, 어느 오브젝트 카테고리에 속하는 화소인지를 식별하는 기술이다.
- [0304] 또한, 오브젝트 식별부(126)에 있어서 실행하는 실제 오브젝트의 식별 처리는, 타깃 영역만, 혹은 타깃 영역을 포함하는 주위 영역이 한정된 범위만을 대상으로 하여 실행한다. 이와 같은 한정 범위의 처리를 행함으로써 고속의 처리, 즉 리얼타임 처리가 가능해진다.

- [0305] 오브젝트 식별부(126)가 해석한 타깃 영역의 실제 오브젝트의 식별 결과는, 데이터 출력부(130)의 콘텐츠 표시 제어부(131)와, 콘텐츠 음성 제어부(132)에 입력된다.
- [0306] (스텝 S106)
- [0307] 다음으로, 스텝 S106에 있어서, 식별된 타깃 영역의 실제 오브젝트(클래스)에 기초하여, 타깃 영역에 표시할 가상 오브젝트의 종류와 출력 양태를 결정한다.
- [0308] 이 처리는, 도 9에 도시한 화상 처리 장치(100)의 데이터 출력부(130)의 콘텐츠 표시 제어부(131), 콘텐츠 음성 제어부(132)가 실행하는 처리이다.
- [0309] 데이터 출력부(130)의 콘텐츠 표시 제어부(131)와, 콘텐츠 음성 제어부(132)는, 우선 도 12를 참조하여 설명한 데이터를 기록한 클래스 대응 가상 오브젝트 데이터(135)를 참조하여, 타깃 영역에 표시할 가상 오브젝트의 종류와 출력 양태를 결정한다.
- [0310] 즉, 클래스 대응 가상 오브젝트 데이터(135)의 각 엔트리로부터, 타깃 영역의 실제 오브젝트의 종류(클래스)를 기록한 엔트리를 선택하고, 그 엔트리에 기록된 가상 오브젝트를 출력 오브젝트로서 결정하는 처리 등을 실행한다.
- [0311] (스텝 S107)
- [0312] 마지막으로, 스텝 S107에 있어서, 스텝 S106에서 결정된 타깃 영역에 표시할 가상 오브젝트의 종류와 출력 양태에 따라서 가상 오브젝트를 타깃 영역으로 출력(표시)한다.
- [0313] 이 처리도, 도 9에 도시한 화상 처리 장치(100)의 데이터 출력부(130)의 콘텐츠 표시 제어부(131), 콘텐츠 음성 제어부(132)가 실행하는 처리이다.
- [0314] 콘텐츠 표시 제어부(131)는, 오브젝트 식별부(126)로부터 타깃 영역의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라 표시해야 할 가상 오브젝트(캐릭터 등)의 선택 처리나 표시 양태를 결정하여 표시부(133)에 표시한다.
- [0315] 구체적으로는, 예를 들어 앞에서 설명한 도 2 내지 도 4, 도 7, 도 8에 도시한 바와 같은 가상 오브젝트(캐릭터 등)의 표시 처리를 실행한다.
- [0316] 또한, 콘텐츠 음성 제어부(132)는, 오브젝트 식별부(126)로부터 타깃 영역의 오브젝트 식별 결과를 입력하고, 오브젝트 식별 결과에 따라서 출력해야 할 음성을 결정하여 스피커(134)를 통해 출력한다.
- [0317] 구체적으로는, 예를 들어 앞에서 설명한 도 7에 도시한 바와 같이, 실제 오브젝트인 연못에서 가상 오브젝트가 출현하는 경우, 물소리를 출력하는 처리를 실행한다.
- [0318] (3-(2) 타깃 영역의 설정 영역을 거의 수평한 면에 설정하는 처리의 시퀀스)
- [0319] 다음으로, 도 16에 도시한 흐름도를 참조하여, 타깃 영역의 설정 영역을 거의 수평한 면에 설정하는 처리의 시퀀스에 대하여 설명한다.
- [0320] 캐릭터 등의 가상 오브젝트를 실세계의 실제 오브젝트상에 표시하는 경우, 옥외라면 지면의 위, 실내라면 바닥의 위에 표시하면, 보다 자연스러운 캐릭터 표시가 가능하게 되어, 유저에게 캐릭터가 실제로 실세계에 존재하는 듯한 감각을 부여하는 것이 가능해진다.
- [0321] 이것을 위해서는, 캐릭터인 가상 오브젝트의 출력 영역이 되는 타깃 영역을 지면의 위나 바닥의 위 등, 거의 수평한 면에 설정하는 제어를 행하는 것이 유효하다.
- [0322] 도 16에 도시한 흐름도는, 이와 같은 처리를 실행하는 화상 처리 장치(100)의 처리 시퀀스를 설명하는 흐름도이다.
- [0323] 이하, 도 16에 도시한 흐름도의 각 스텝의 처리에 대하여 설명한다.
- [0324] 또한, 도 16에 도시한 흐름도의 스텝 S101 내지 S103과 스텝 S105 내지 S107의 처리는, 앞에서 도 13을 참조하여 설명한 기본 처리 흐름의 각 스텝의 처리와 마찬가지로 처리이다.
- [0325] 도 16에 도시한 흐름의 스텝 S301 내지 S303의 처리와, 스텝 S104의 처리가 앞에서 설명한 도 13에 도시한 흐름과 다른 점이다.

- [0326] 이 각 스텝의 처리에 대하여 설명한다.
- [0327] (스텝 S301)
- [0328] 스텝 S301은, 도 9에 도시한 화상 처리 장치(100)의 데이터 입력부(110)의 모션 센서(113)로부터, 데이터 처리부(120)의 디바이스 자세 해석부(124)에 센서 검출 정보를 입력하는 처리이다.
- [0329] 앞에서 도 9를 참조하여 설명한 바와 같이, 모션 센서(113)는, 자이로, 가속도 센서 등을 갖고, 예를 들어 HMD나 스마트폰 등, 화상 처리 장치(100) 본체의 자세나 움직임을 검출하는 센서이다.
- [0330] 이 모션 센서(113)로부터, 데이터 처리부(120)의 디바이스 자세 해석부(124)에 센서 검출 정보가 입력된다.
- [0331] (스텝 S302)
- [0332] 다음으로, 스텝 S302에 있어서, 모션 센서 검출 정보에 기초하여 중력 방향을 추정한다.
- [0333] 이 처리는, 도 9에 도시한 데이터 처리부(120)의 디바이스 자세 해석부(124)가 실행하는 처리이다.
- [0334] 데이터 처리부(120)의 디바이스 자세 해석부(124)는, 모션 센서(113)를 구성하는 자이로나, 가속도 센서 등의 센서 검출 정보를 이용하여, 중력 방향을 산출한다.
- [0335] (스텝 S303)
- [0336] 다음으로, 스텝 S303에 있어서, 수평면 영역의 검출 처리를 행한다.
- [0337] 이 처리는, 도 9에 도시한 오브젝트 식별부(126)가 실행하는 처리이다.
- [0338] 오브젝트 식별부(126)는, 3차원 맵 생성부(122)가 생성한 3차원 맵과, 디바이스 자세 해석부(124)로부터 입력한 중력 방향 정보를 이용하여, 3차원 맵 내의 수평면 영역을 검출한다. 구체적으로는 예를 들어 지면이나 바닥면 등을 검출한다.
- [0339] 또한, 검출하는 수평면 영역은, 완전한 수평면에 한정되지 않고, 거의 수평한 영역으로 하면 된다.
- [0340] 예를 들어 어느 정도의 요철이나, 어느 정도의 구배를 갖는 비탈길 등도 수평면 영역이라고 판정하여 검출한다.
- [0341] 어느 정도의 요철이나 경사를 수평면 영역으로서 허용할지에 대해서는, 사전에 설정하는 것이 가능하다.
- [0342] (스텝 S104)
- [0343] 다음으로, 데이터 처리부는, 스텝 S104에 있어서, 타깃 영역 결정 처리를 실행한다.
- [0344] 단, 본 처리예에서는, 타깃 영역은, 스텝 S303에 있어서 검출한 수평면 영역으로부터만 선택하는 것으로 한다.
- [0345] 이 처리는, 도 9에 도시한 데이터 처리부(120)의 오브젝트 식별부(126)가 실행하는 처리이다.
- [0346] 오브젝트 식별부(126)는, 가상 오브젝트 표시 영역으로 할 타깃 영역을 스텝 S303에 있어서 검출한 수평면 영역 내에 한정하여 결정한다.
- [0347] 이 타깃 영역의 결정 처리는, 전술한 바와 같이, 다양한 방법이 적용 가능하다.
- [0348] 예를 들어, 3차원 맵에 포함되는 유저의 손가락 화상을 사용하여 행하는 것이 가능하다.
- [0349] 즉, 유저의 포인팅 방향의 연장선과, 3차원 맵상의 실제 오브젝트이며, 지면이나 바닥면 등의 수평면이라고 판정된 수평면 영역과의 교점을 구하고, 이 수평면과의 교점을 중심으로 하는 미리 규정된 반경의 원형 영역을 타깃 영역으로서 결정한다.
- [0350] 또한, 타깃 영역의 결정에 이용하는 데이터는, 앞에서 도 13을 참조하여 설명한 바와 마찬가지로, 다양한 정보를 이용하는 것이 가능하다. 예를 들어, 이하의 입력 정보를 이용하는 구성으로 해도 된다.
- [0351] (a) 내부 촬영 화상 해석부(123)가 해석한 유저 시선 정보
- [0352] (b) 디바이스 자세 해석부(124)가 해석한 화상 처리 장치(100) 본체의 자세나 움직임 정보
- [0353] (c) 조작부(114)를 통해 입력하는 유저 조작 정보
- [0354] (d) 음성 해석부(125)가 해석한 유저 음성 정보

- [0355] 예를 들어, 이들 입력 정보 중 어느 것을 사용하여 타깃 영역을 결정해도 된다.
- [0356] 스텝 S101 내지 S103의 처리와 스텝 S105 이하의 처리는, 앞에서 설명한 도 13에 도시한 흐름도와 마찬가지로 처리가 된다.
- [0357] 본 처리 예에서는, 캐릭터인 가상 오브젝트의 출력 영역이 되는 타깃 영역을 지면의 위나 바닥의 위 등, 거의 수평한 면에 설정하는 제어를 행하는 것이 가능해진다.
- [0358] 이에 의해, 캐릭터 등의 가상오브젝트를 실세계의 실제 오브젝트상에 표시하는 경우, 옥외라면 지면의 위, 실내라면 바닥의 위 등, 수평면 영역에 접하도록 표시하는 것이 가능하게 되고, 보다 자연스러운 캐릭터 표시가 가능하게 되어, 유저에게 캐릭터가 실제로 실세계에 존재하는 듯한 감각을 부여하는 것이 가능해진다.
- [0359] (3-(3) 실제 오브젝트 식별 처리의 갱신 시퀀스)
- [0360] 다음으로, 오브젝트 식별부가 실행하는 실제 오브젝트 식별 처리의 갱신 시퀀스에 대하여 설명한다.
- [0361] 앞에서 도 10, 도 11 등을 참조하여 설명한 바와 같이, 도 9에 도시한 화상 처리 장치(100)의 데이터 처리부(120) 내의 오브젝트 식별부(126)는, 타깃 영역이 결정된 직후에, 타깃 영역의 실제 오브젝트의 식별 처리를 실행하지만, 그 후에도, 그 영역에 대한 오브젝트 식별 처리를 반복해서 실행하여, 도 10에 도시한 공간 맵 데이터를, 축차 갱신한다.
- [0362] 단, 갱신 처리의 간격은, 식별된 실제 오브젝트의 종류(클래스)에 따라 다르다.
- [0363] 이 실제 오브젝트의 종류(클래스)에 따라 다른 갱신 시간의 규정 데이터는, 클래스 대응 갱신 시간 데이터(128)로서 미리 등록되어 있다.
- [0364] 클래스 대응 갱신 시간 데이터(128)는, 앞에서 도 11을 참조하여 설명한 바와 같이 이하의 각 데이터를 대응지은 데이터이다.
- [0365] (a) ID
- [0366] (b) 카테고리
- [0367] (c) 클래스
- [0368] (d) 갱신 시간(sec)
- [0369] (a) ID는, 등록 데이터의 식별자이다.
- [0370] (b) 카테고리는, 실제 오브젝트의 종류(클래스)의 카테고리이다.
- [0371] (c) 클래스는, 실제 오브젝트의 종류 정보이다.
- [0372] (d) 갱신 시간(sec)은, 실제 오브젝트 식별 처리의 갱신 간격을 나타내는 시간이다.
- [0373] 예를 들어, ID001의 클래스(오브젝트 종류)=잔디밭의 경우, 갱신 시간은 3600sec(=1시간)이다. 잔디밭과 같은 오브젝트는, 시간 경과에 수반되는 변화가 적어, 갱신 시간이 길게 설정된다.
- [0374] 한편, 예를 들어 ID=004의 클래스(오브젝트 종류)=그림자의 경우, 갱신 시간은 2sec이다. 그림자와 같은 오브젝트는, 시간 경과에 수반되는 변화가 크기 때문에, 갱신 시간이 짧게 설정된다.
- [0375] 오브젝트 식별부(126)는, 이 클래스 대응 갱신 시간 데이터(128)의 데이터를 참조하여, 식별한 오브젝트에 대하여 규정된 시간 간격으로, 수시로 오브젝트 식별 처리를 반복하여, 실행한다. 새로운 식별 처리에 의해 검출된 실제 오브젝트는, 도 10을 참조하여 설명한 공간 맵 데이터(127)로서, 순차 등록한다.
- [0376] 도 17에 도시한 흐름도는, 이 오브젝트 식별 처리의 반복 실행 시퀀스를 포함하는 처리를 설명하는 흐름도이다.
- [0377] 이하, 도 17에 도시한 흐름도의 각 스텝의 처리에 대하여 설명한다.
- [0378] 또한, 도 17에 도시한 흐름도의 스텝 S101 내지 S105의 처리와, 스텝 S106 내지 S107의 처리는, 앞에서 도 13을 참조하여 설명한 기본 처리 흐름의 각 스텝의 처리와 마찬가지로 처리이다.
- [0379] 도 17에 도시한 흐름의 스텝 S401과 스텝 S402의 처리가 앞에서 설명한 도 13에 도시한 흐름과 다른 점이다.
- [0380] 이 각 스텝의 처리에 대하여 설명한다.

- [0381] (스텝 S401)
- [0382] 스텝 S101 내지 S105에 있어서, 타깃 영역의 결정과, 타깃 영역의 실제 오브젝트(클래스)의 식별 처리가 실행된 후, 스텝 S401의 처리를 실행한다.
- [0383] 스텝 S401에서는, 스텝 S105에 있어서 실행된 타깃 영역의 오브젝트 식별 결과를 공간 맵 데이터에 기록한다.
- [0384] 공간 맵 데이터는, 앞에서 도 10을 참조하여 설명한 바와 같이, 이하의 각 데이터의 대응 데이터를 저장하고 있다.
- [0385] (a) 타임 스탬프(sec)
- [0386] (b) 위치 정보
- [0387] (c) 클래스
- [0388] (d) 식별 처리 후 경과 시간(sec)
- [0389] (a) 타임 스탬프(sec)는, 오브젝트 식별 처리를 실행한 시간 정보이다.
- [0390] (b) 위치 정보는, 오브젝트 식별 대상이 된 실제 오브젝트의 위치 정보이다.
- [0391] (c) 클래스는, 오브젝트 식별 결과로서의 오브젝트 종류 정보이다.
- [0392] (d) 식별 처리 후 경과 시간(sec)은, 오브젝트 식별 처리의 완료 시부터의 경과 시간이다.
- [0393] 스텝 S401에서는, 스텝 S105에 있어서 식별한 타깃 영역의 실제 오브젝트에 대해서, 이들 각 데이터를 공간 맵 데이터에 등록한다.
- [0394] (스텝 S106 내지 S107)
- [0395] 스텝 S106 내지 S107의 처리는, 앞에서 도 13을 참조하여 설명하는 처리와 마찬가지로의 처리이다. 즉, 이하의 처리를 실행한다.
- [0396] 스텝 S106에 있어서, 식별된 타깃 영역의 실제 오브젝트(클래스)에 기초하여, 타깃 영역에 표시할 가상 오브젝트의 종류와 출력 양태를 결정한다.
- [0397] 스텝 S107에 있어서, 스텝 S106에서 결정한 타깃 영역에 표시할 가상 오브젝트의 종류와 출력 양태에 따라서 가상 오브젝트를 타깃 영역으로 출력(표시)한다.
- [0398] (스텝 S402)
- [0399] 또한, 스텝 S107의 처리 후, 스텝 S402에 있어서, 스텝 S105에 있어서 실행한 타깃 영역의 실제 오브젝트의 식별 처리 후의 경과 시간이, 도 11을 참조하여 설명한 클래스 대응 갱신 시간 데이터에 규정된 「(d) 갱신 시간」을 초과하였는지 여부를 판정한다.
- [0400] 초과하였다고 판정된 경우, 스텝 S101로 되돌아가서, 스텝 S101 이하의 처리를 반복하여, 실행한다.
- [0401] 즉, 다시, 타깃 영역의 결정, 타깃 영역의 실제 오브젝트 식별 처리를 실행한다.
- [0402] 이 처리에 있어서, 타깃 영역의 위치가 변경되어 있지 않으면, 동일한 위치의 타깃 영역에서 실제 오브젝트 식별이 다시 실행된다.
- [0403] 한편, 타깃 영역의 위치가 변경되어 있으면, 새로운 위치의 타깃 영역에서 실제 오브젝트 식별이 실행된다.
- [0404] 이들 처리를 행함으로써, 타깃 영역의 갱신 처리나, 실제 오브젝트의 식별 결과의 갱신 처리를 바로 행하는 것이 가능하게 되어, 유저의 움직임이나 지시에 따른 시의 적절한 가상 오브젝트 표시 처리를 행하는 것이 가능해진다.
- [0405] [4. 화상 처리 장치의 하드웨어 구성예에 대하여]
- [0406] 다음으로, 상술한 실시예를 따른 처리를 실행하는 화상 처리 장치의 하드웨어 구성예에 대하여, 도 18을 참조하여 설명한다.
- [0407] 도 18에 도시한 하드웨어는, 도 9를 참조하여 설명한 본 개시의 화상 처리 장치(100)의 하드웨어 구성의 일례이

다.

- [0408] 도 18에 도시한 하드웨어 구성에 대하여 설명한다.
- [0409] CPU(Central Processing Unit)(301)는, ROM(Read Only Memory)(302) 또는 기억부(308)에 기억되어 있는 프로그램에 따라서 각종 처리를 실행하는 데이터 처리부로서 기능한다. 예를 들어, 상술한 실시예에 있어서 설명한 시퀀스에 따른 처리를 실행한다. RAM(Random Access Memory)(303)에는, CPU(301)가 실행하는 프로그램이나 데이터 등이 기억된다. 이들 CPU(301), ROM(302) 및 RAM(303)은, 버스(304)에 의해 서로 접속되어 있다.
- [0410] CPU(301)는 버스(304)를 통해 입출력 인터페이스(305)에 접속되고, 입출력 인터페이스(305)에는, 각종 센서, 카메라, 스위치, 키보드, 마우스, 마이크론 등을 포함하는 입력부(306), 디스플레이, 스피커 등을 포함하는 출력부(307)가 접속되어 있다.
- [0411] 입출력 인터페이스(305)에 접속되어 있는 기억부(308)는, 예를 들어 하드 디스크 등을 포함하고, CPU(301)가 실행하는 프로그램이나 각종 데이터를 기억한다. 통신부(309)는, 인터넷이나 로컬 에어리어 네트워크 등의 네트워크를 통한 데이터 통신의 송수신부, 또한 방송과의 송수신부로서 기능하여, 외부의 장치와 통신한다.
- [0412] 입출력 인터페이스(305)에 접속되어 있는 드라이브(310)는, 자기 디스크, 광 디스크, 광자기 디스크, 혹은 메모리 카드 등의 반도체 메모리 등의 리무버블 미디어(311)를 구동하여, 데이터의 기록 혹은 판독을 실행한다.
- [0413] [5. 본 개시의 구성의 요약]
- [0414] 이상, 특정 실시예를 참조하면서, 본 개시의 실시예에 대하여 상세히 해석해 왔다. 그러나, 본 개시의 요지를 일탈하지 않는 범위에서 당업자가 실시예의 수정이나 대응을 할 수 있음은 자명하다. 즉, 예시라고 하는 형태로 본 발명을 개시해 온 것이며, 한정적으로 해석되어서는 안된다. 본 개시의 요지를 판단하기 위해서는, 청구 범위의 부분을 참작해야 한다.
- [0415] 또한, 본 명세서에 있어서 개시한 기술은, 이하와 같은 구성을 취할 수 있다.
- [0416] (1) 실제세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별부와,
- [0417] 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어부를 갖고,
- [0418] 상기 오브젝트 식별부는,
- [0419] 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행하고,
- [0420] 상기 콘텐츠 표시 제어부는,
- [0421] 상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 화상 처리 장치.
- [0422] (2) 상기 오브젝트 식별부는,
- [0423] 화상 인식 처리에 의해 오브젝트 식별 처리를 실행하는, 상기 (1)에 기재된 화상 처리 장치.
- [0424] (3) 상기 오브젝트 식별부는,
- [0425] 시맨틱·세그멘테이션 처리를 적용하여 오브젝트 식별 처리를 실행하는, 상기 (2)에 기재된 화상 처리 장치.
- [0426] (4) 상기 오브젝트 식별부는,
- [0427] 상기 가상 오브젝트의 표시 영역으로 할 타깃 영역을 결정하고, 결정된 타깃 영역의 실제 오브젝트의 식별 처리를 실행하는, 상기 (1) 내지 (3) 중 어느 것에 기재된 화상 처리 장치.
- [0428] (5) 상기 오브젝트 식별부는,
- [0429] 상기 타깃 영역을 유저 동작, 유저 시선, 유저 조작, 유저 위치 또는 유저 자세 중 적어도 어느 것에 기초하여 결정하는, 상기 (4)에 기재된 화상 처리 장치.
- [0430] (6) 상기 오브젝트 식별부는,
- [0431] 상기 타깃 영역을 수평면 영역으로부터 선택하여 결정하는, 상기 (4) 또는 (5)에 기재된 화상 처리 장치.

- [0432] (7) 상기 콘텐츠 표시 제어부는,
- [0433] 상기 가상 오브젝트를 상기 수평면 영역에 접하도록 표시하는, 상기 (6)에 기재된 화상 처리 장치.
- [0434] (8) 상기 오브젝트 식별부는,
- [0435] 상기 오브젝트 식별 처리를 리얼타임 처리로서 실행하는, 상기 (1) 내지 (7)에 기재된 화상 처리 장치.
- [0436] (9) 상기 오브젝트 식별부는,
- [0437] 오브젝트 종류에 따라 미리 규정된 시간 간격으로 오브젝트 식별 처리를 반복해서 실행하는, 상기 (1) 내지 (8) 중 어느 것에 기재된 화상 처리 장치.
- [0438] (10) 상기 화상 처리 장치는,
- [0439] 카메라 촬영 화상에 기초하는 실세계의 3차원 맵을 생성하는 3차원 맵 생성부를 갖고,
- [0440] 상기 오브젝트 식별부는,
- [0441] 상기 3차원 맵을 이용하여 상기 가상 오브젝트의 표시 영역으로 할 타깃 영역을 결정하는, 상기 (1) 내지 (9) 중 어느 것에 기재된 화상 처리 장치.
- [0442] (11) 상기 3차원 맵 생성부는,
- [0443] SLAM(Simultaneous Localization and Mapping) 처리에 의해, 실세계의 3차원 맵을 생성하는, 상기 (10)에 기재된 화상 처리 장치.
- [0444] (12) 상기 콘텐츠 표시 제어부는,
- [0445] 상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택함과 함께,
- [0446] 상기 오브젝트 식별 결과에 따라서, 표시할 가상 오브젝트의 표시 양태도 제어하는, 상기 (1) 내지 (11) 중 어느 것에 기재된 화상 처리 장치.
- [0447] (13) 상기 화상 처리 장치는,
- [0448] 음성 출력 제어를 실행하는 콘텐츠 음성 제어부를 더 갖고,
- [0449] 상기 콘텐츠 음성 제어부는,
- [0450] 상기 오브젝트 식별부에 있어서 식별된 오브젝트 식별 결과에 따라서 출력할 음성을 결정하여 출력하는, 상기 (1) 내지 (12) 중 어느 것에 기재된 화상 처리 장치.
- [0451] (14) 화상 처리 장치에 있어서 실행하는 화상 처리 방법이며,
- [0452] 오브젝트 식별부가, 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별 처리 스텝을 실행하고,
- [0453] 콘텐츠 표시 제어부가, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어 스텝을 실행하고,
- [0454] 상기 오브젝트 식별 처리 스텝은,
- [0455] 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행하는 스텝이며,
- [0456] 상기 콘텐츠 표시 제어 스텝은,
- [0457] 상기 오브젝트 식별 처리 스텝에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 스텝을 실행하는 화상 처리 방법.
- [0458] (15) 화상 처리 장치에 있어서 화상 처리를 실행시키는 프로그램이며,
- [0459] 오브젝트 식별부에, 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별 처리 스텝을 실행시키고,
- [0460] 콘텐츠 표시 제어부에, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR(Augmented Reality) 화상을 생성하는 콘텐츠 표시 제어 스텝을 실행시키고,
- [0461] 상기 오브젝트 식별 처리 스텝에 있어서는,

- [0462] 상기 가상 오브젝트의 표시 영역의 실제 오브젝트를 식별하는 오브젝트 식별 처리를 실행시키고,
- [0463] 상기 콘텐츠 표시 제어 스텝에 있어서는,
- [0464] 상기 오브젝트 식별 처리 스텝에 있어서 식별된 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 스텝을 실행시키는 프로그램.
- [0465] 또한, 명세서 중에 있어서 설명한 일련의 처리는 하드웨어 또는 소프트웨어, 혹은 양자의 복합 구성에 의해 실행하는 것이 가능하다. 소프트웨어에 의한 처리를 실행하는 경우에는, 처리 시퀀스를 기록한 프로그램을, 전용의 하드웨어에 내장된 컴퓨터 내의 메모리에 인스톨해서 실행시키거나, 혹은 각종 처리가 실행 가능한 범용 컴퓨터에 프로그램을 인스톨해서 실행시키는 것이 가능하다. 예를 들어, 프로그램은 기록 매체에 미리 기록해 둘 수 있다. 기록 매체로부터 컴퓨터에 인스톨하는 것 외에, LAN(Local Area Network), 인터넷과 같은 네트워크를 통해 프로그램을 수신하여, 내장된 하드 디스크 등의 기록 매체에 인스톨할 수 있다.
- [0466] 또한, 명세서에 기재된 각종 처리는, 기재에 따라서 시계열로 실행될 뿐만 아니라, 처리를 실행하는 장치의 처리 능력 혹은 필요에 따라 병렬적으로 혹은 개별로 실행되어도 된다. 또한, 본 명세서에 있어서 시스템이란, 복수의 장치의 논리적 집합 구성이며, 각 구성의 장치가 동일 하우징 내에 있는 것으로 한정되지는 않는다.

### 산업상 이용가능성

- [0467] 이상, 설명한 바와 같이, 본 개시의 일 실시예의 구성에 의하면, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트의 종류에 따라 표시할 가상 오브젝트의 선택이나 표시 양태 변경을 행하는 장치, 방법이 실현된다.
- [0468] 구체적으로는, 예를 들어 실세계상의 실제 오브젝트의 식별 처리를 실행하는 오브젝트 식별부와, 실제 오브젝트와 가상 오브젝트를 중첩 표시한 AR 화상을 생성하는 콘텐츠 표시 제어부를 갖는다. 오브젝트 식별부는, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트를 식별하고, 콘텐츠 표시 제어부는, 오브젝트 식별 결과에 따라 표시할 가상 오브젝트를 선택하는 처리나 표시 양태의 변경 처리를 행한다.
- [0469] 본 구성에 의해, 가상 오브젝트의 표시 영역으로 할 타깃 영역의 실제 오브젝트의 종류에 따라 표시할 가상 오브젝트의 선택이나 표시 양태 변경을 행하는 장치, 방법이 실현된다.

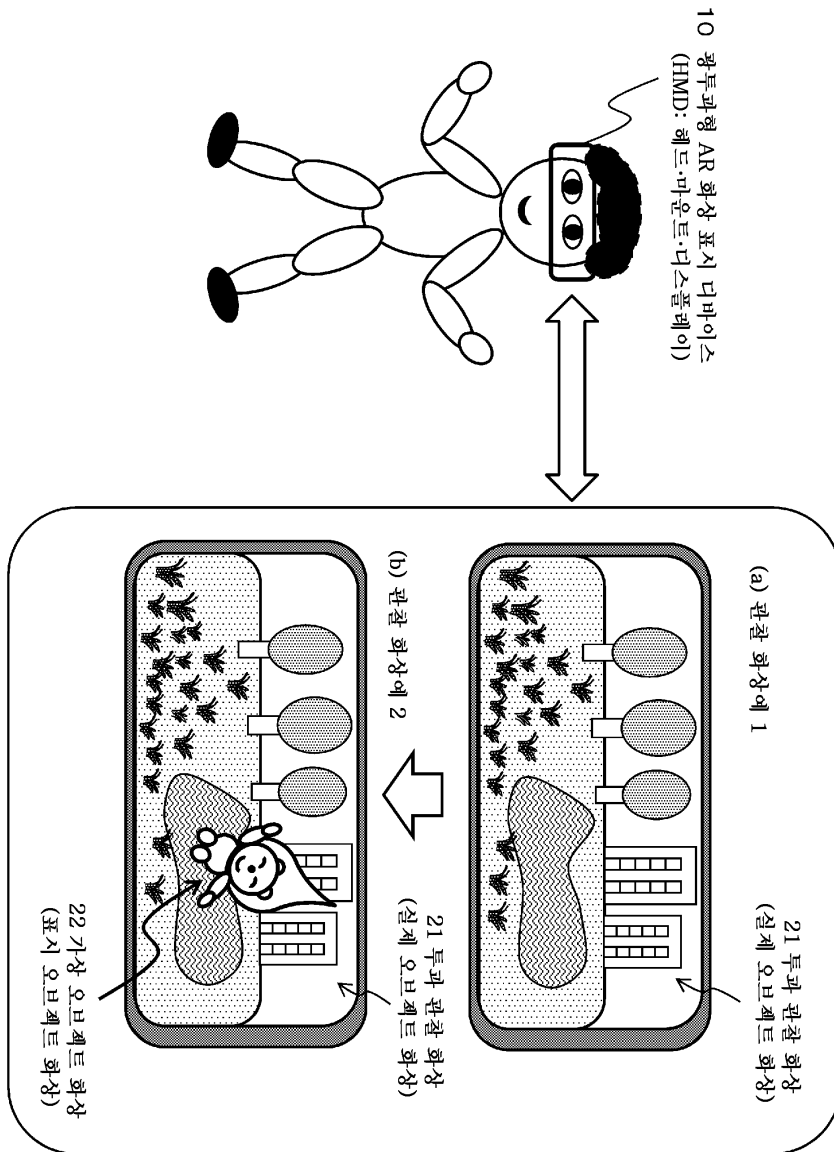
### 부호의 설명

- [0470] 10: 광투과형 AR 화상 표시 디바이스
- 11: 타깃 영역
- 21: 투과 관찰 화상
- 22 내지 24: 가상 오브젝트 화상
- 30: 카메라 촬영 화상 표시형 AR 화상 표시 디바이스
- 31: 카메라
- 32: 카메라 촬영 화상
- 40: 스마트폰
- 41: 카메라
- 42: 카메라 촬영 화상
- 50: 물 위의 가상 오브젝트 화상
- 51: 물속의 가상 오브젝트 화상
- 52: 가상 오브젝트 화상
- 53, 54: 가상 오브젝트 그림자 화상
- 100: 화상 처리 장치

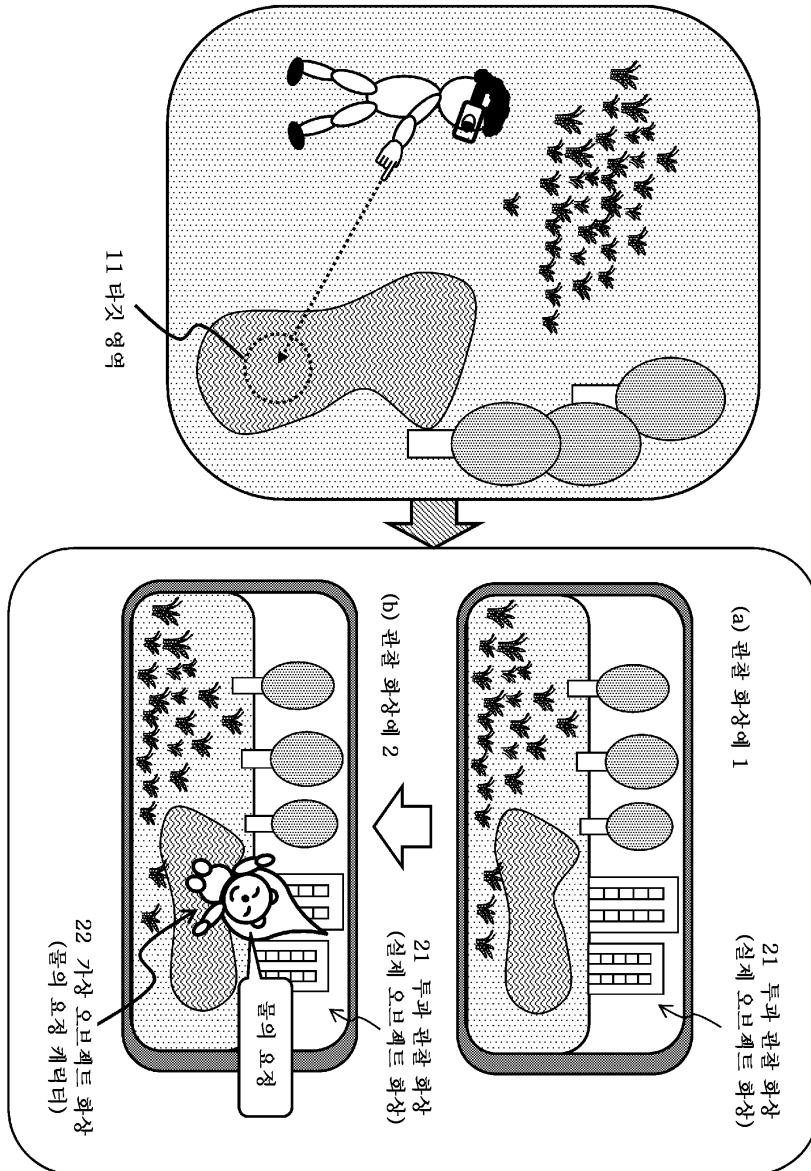
- 110: 데이터 입력부
- 111: 외부 촬영 카메라
- 112: 내부 촬영 카메라
- 113: 모션 센서(자이로, 가속도 센서 등)
- 114: 조작부
- 115: 마이크
- 120: 데이터 처리부
- 121: 외부 촬영 화상 해석부
- 122: 3차원 맵 생성부
- 123: 내부 촬영 화상 해석부
- 124: 디바이스 자세 해석부
- 125: 음성 해석부
- 126: 오브젝트 식별부
- 127: 공간 맵 데이터
- 128: 클래스 대응 갱신 시간 데이터
- 130: 데이터 출력부
- 131: 콘텐츠 표시 제어부
- 132: 콘텐츠 음성 제어부
- 133: 표시부
- 134: 스피커
- 135: 클래스 대응 가상 오브젝트 데이터(3D 모델, 음성 데이터 등)
- 140: 통신부
- 301: CPU
- 302: ROM
- 303: RAM
- 304: 버스
- 305: 입출력 인터페이스
- 306: 입력부
- 307: 출력부
- 308: 기억부
- 309: 통신부
- 310: 드라이브
- 311: 리무버블 미디어

도면

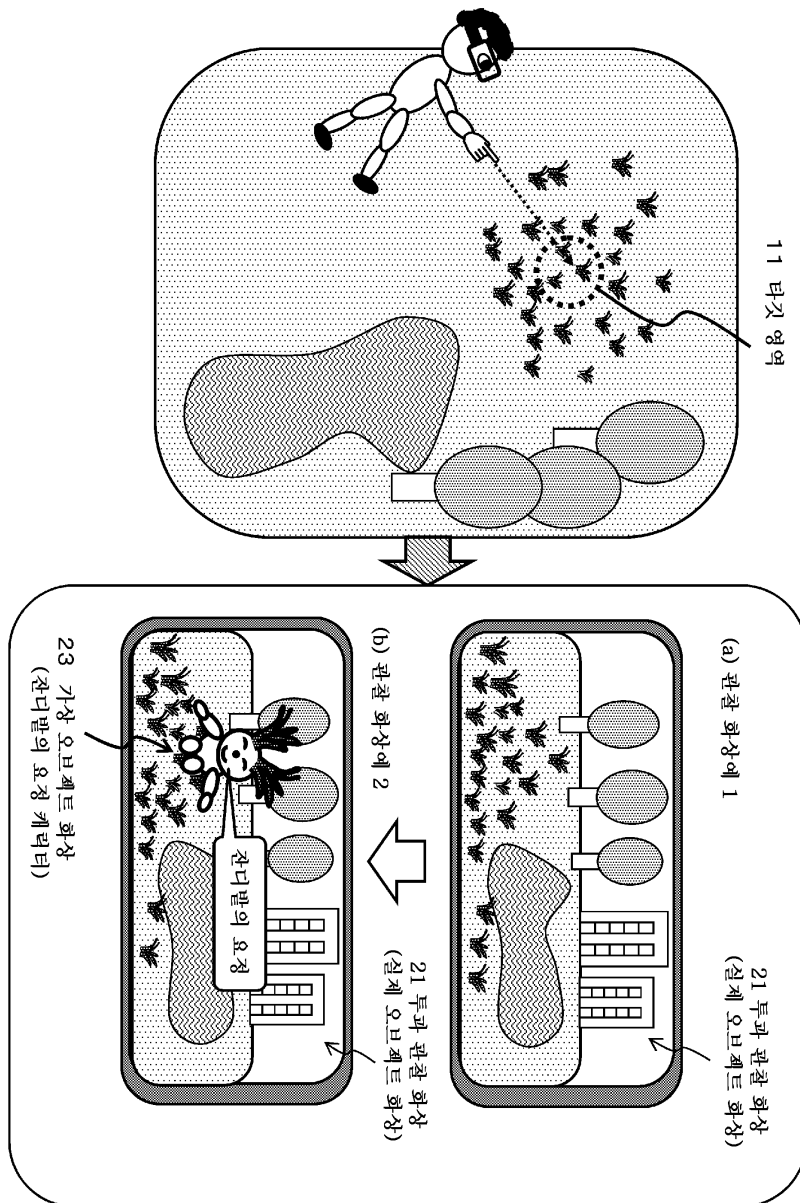
도면1



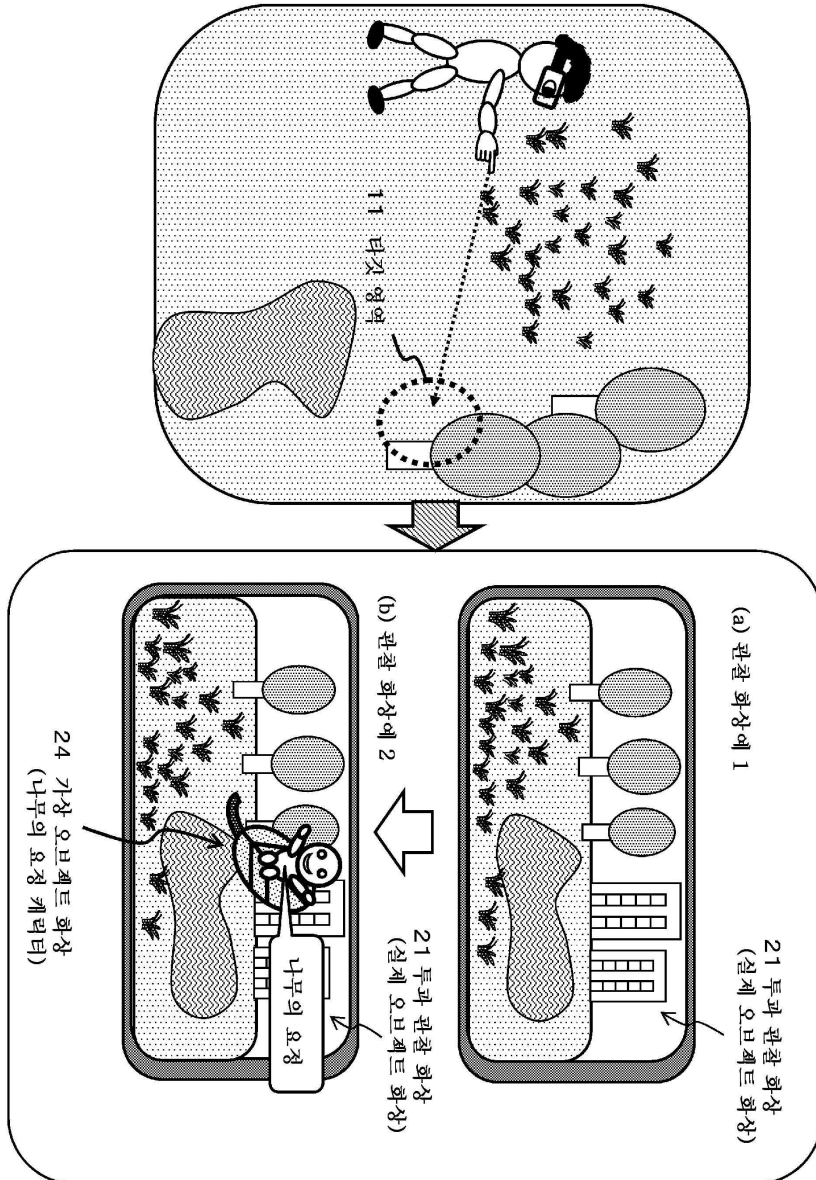
도면2



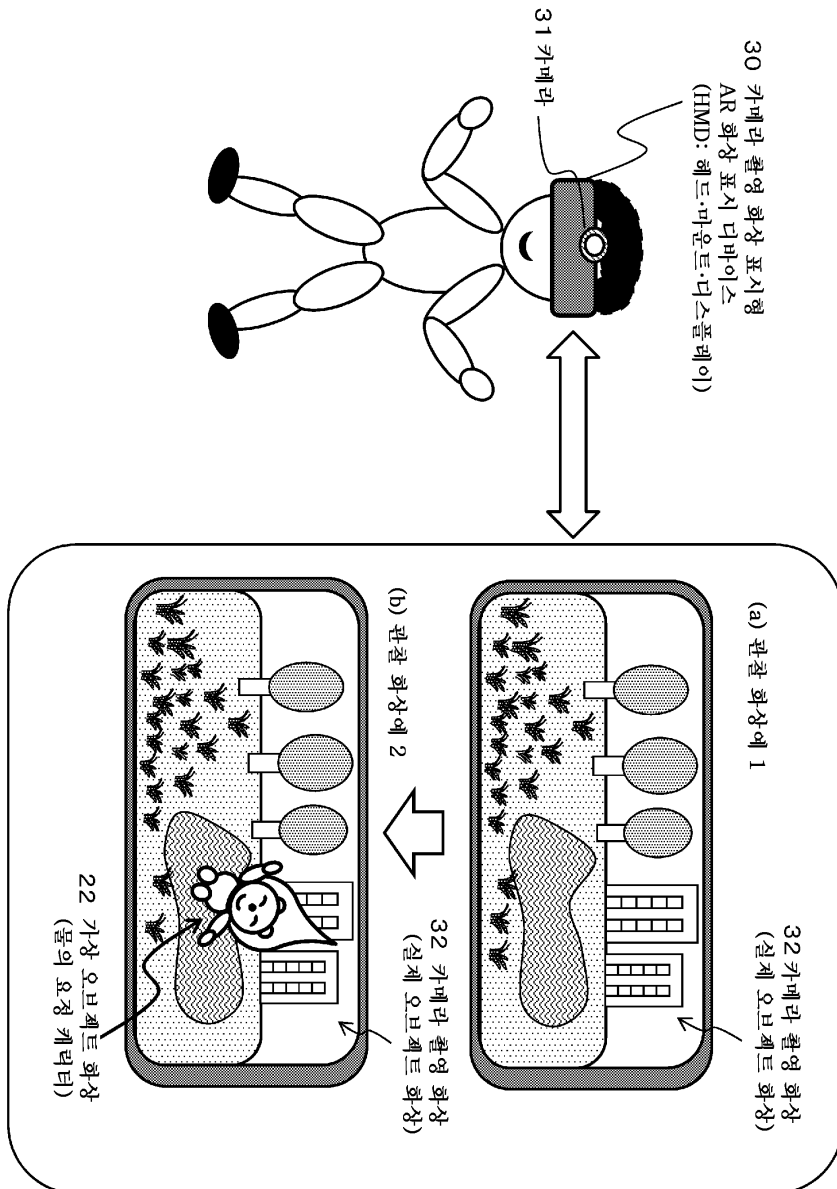
도면3



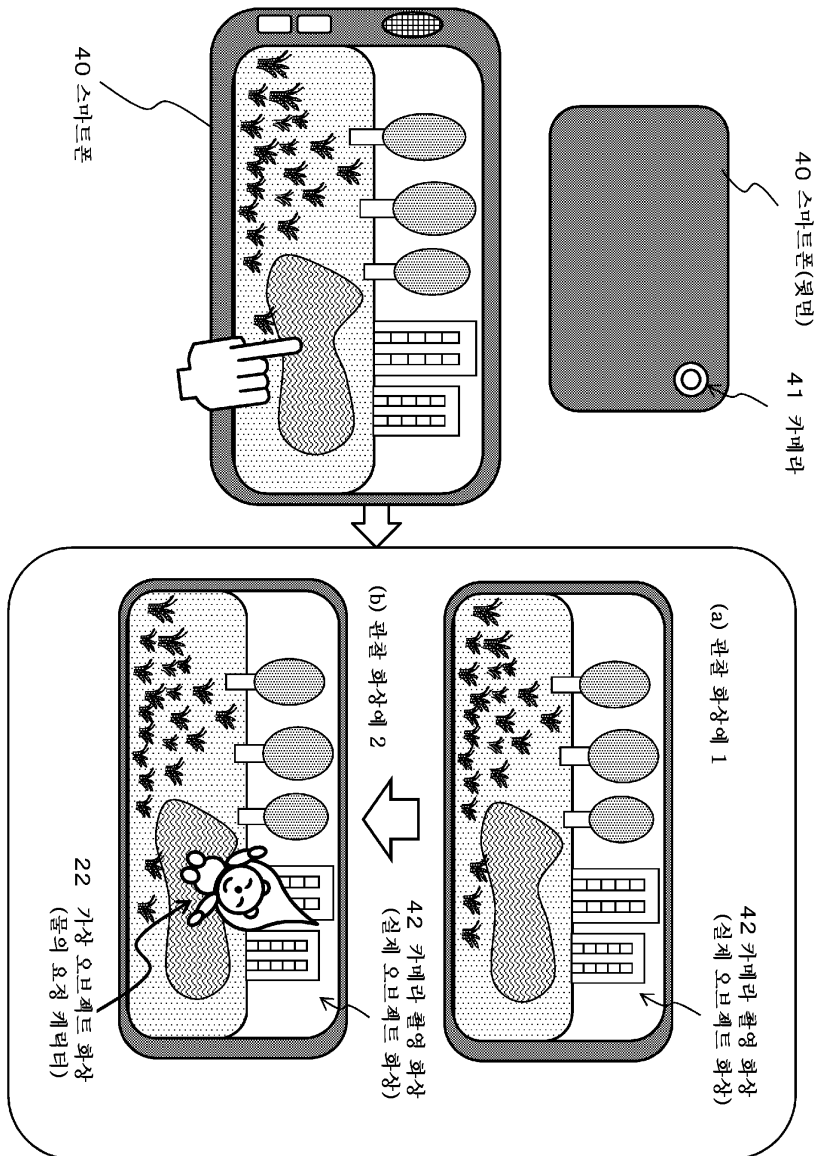
도면4



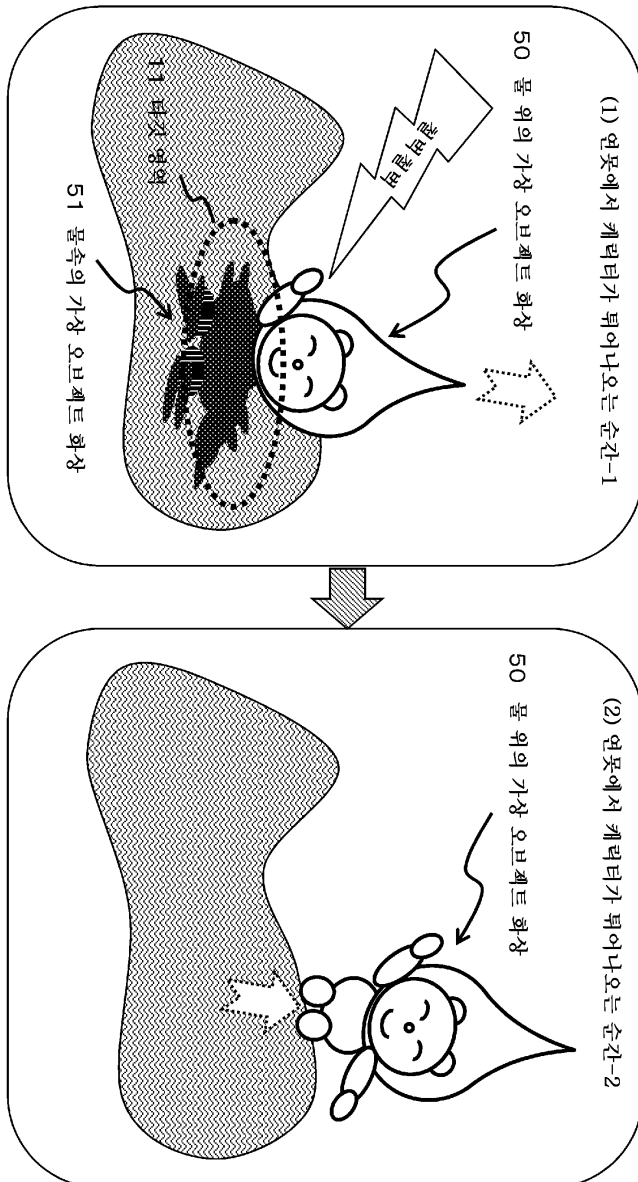
도면5



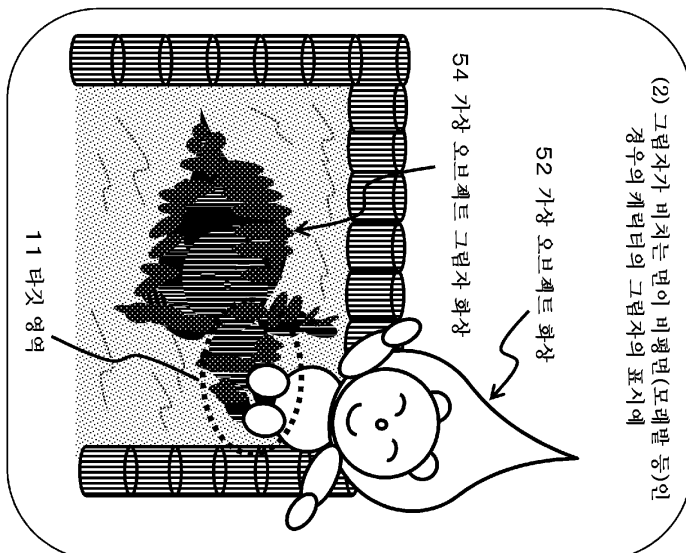
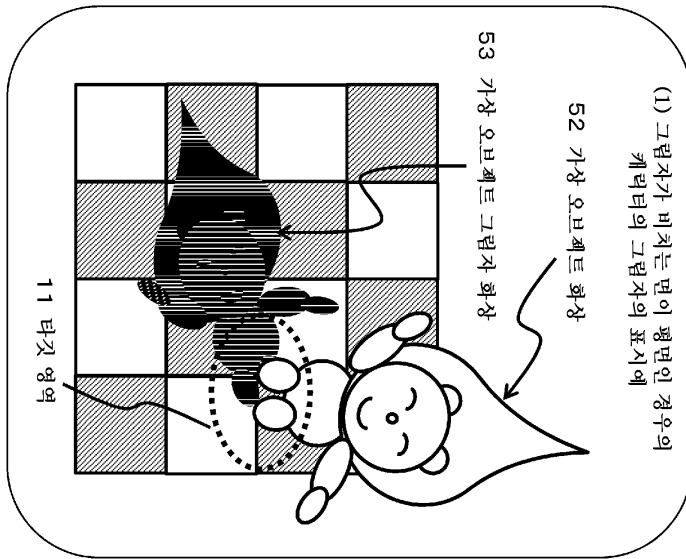
도면6



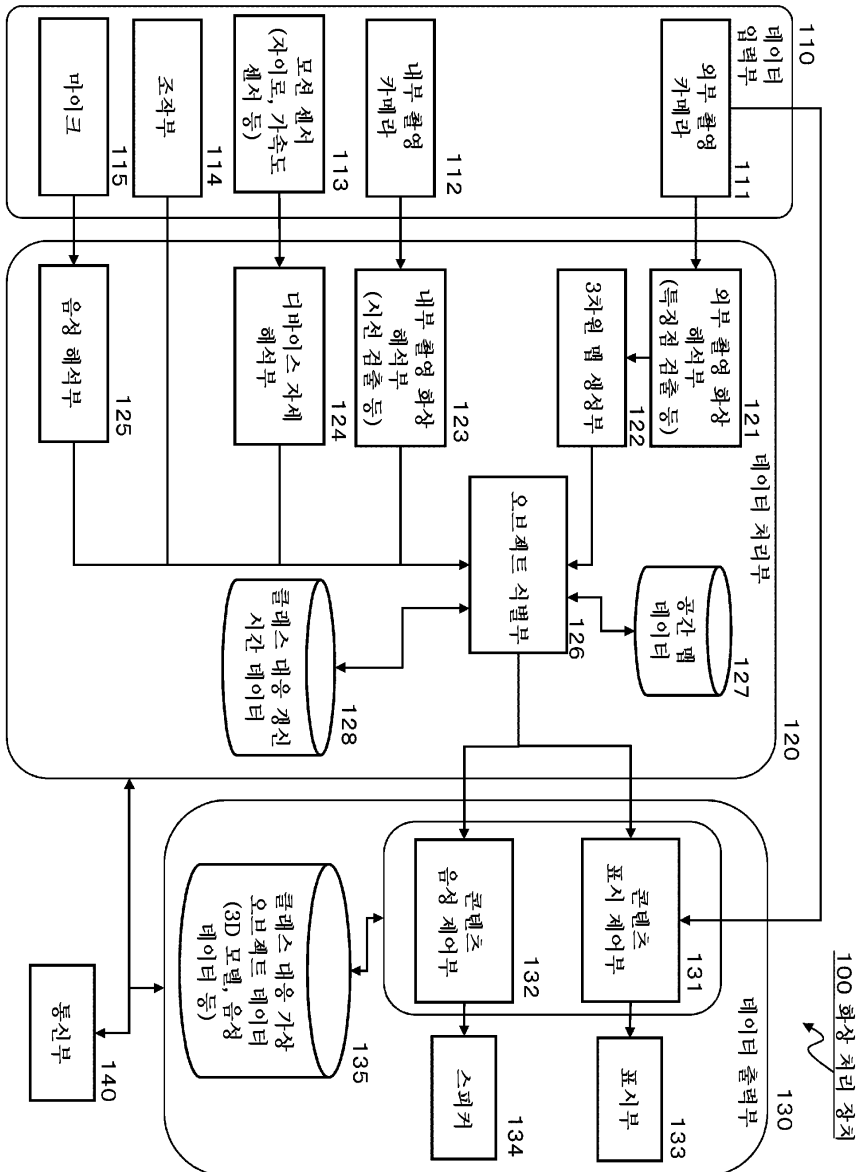
도면7



도면8



도면9



도면10

| (a) 타임 스텝프(sec) | (b) 위치 정보                                                        | (c) 클래스 | (d) 식별 처리 후 경과 시간(sec) |
|-----------------|------------------------------------------------------------------|---------|------------------------|
| 10839948        | v1355,3356,28<br>v1232,3460,31<br>v1344,3577,32<br>v1234,3456,34 | 잔디밭     | 183                    |
| 10839950        | v890,3351,24<br>v912,3090,22<br>v1093,1333,30<br>v1332,3912,18   | 아스팔트    | 3511                   |
| 10839958        | v2239,113,12<br>v2562,3567,14<br>v334,134,11<br>v113,3835,12     | 수면      | 452                    |
| 10839960        | v1678,843,31<br>v879,311,32<br>v1589,309,30<br>v865,890,29       | 그림자     | 3                      |
| ..              | ..                                                               | ..      | ..                     |

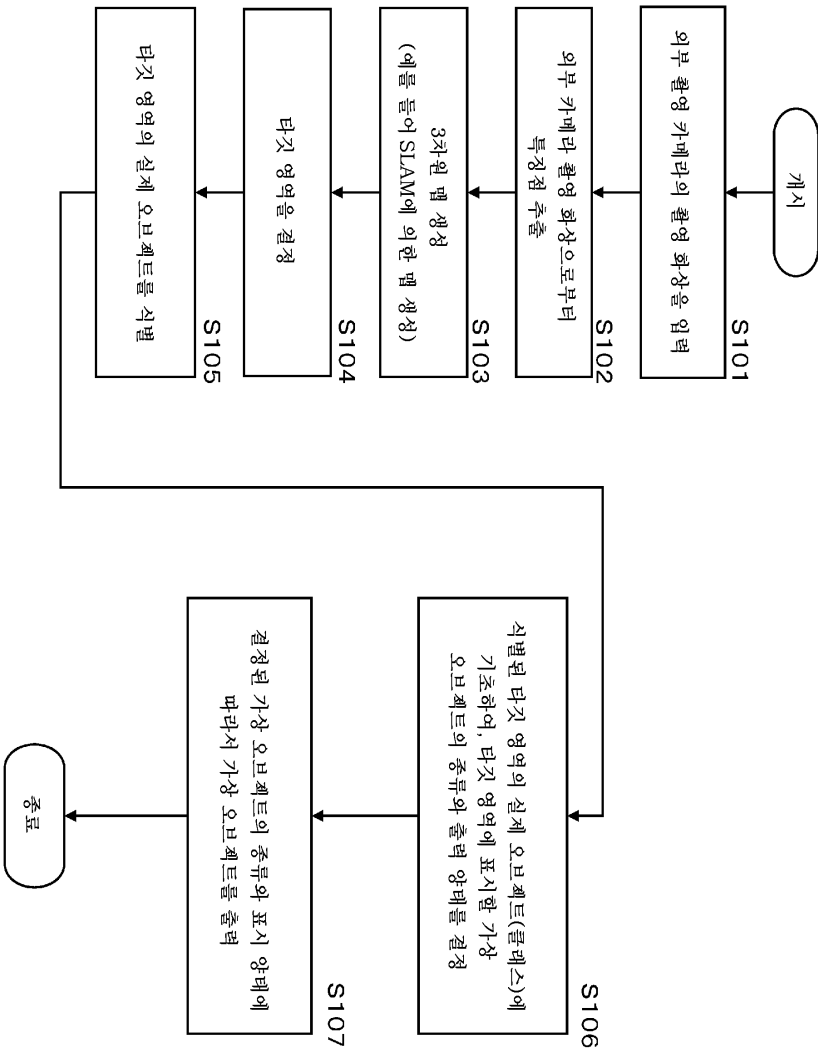
도면11

| (a) ID | (b) 카테고리 | (c) 클래스    | (d) 갱신 시간(sec) |
|--------|----------|------------|----------------|
| 001    | 환경       | 잔디밭        | 3600           |
| 002    | 환경       | 아스팔트       | 3600           |
| 003    | 환경       | 수면         | 600            |
| 004    | 현상       | 그림자        | 2              |
| 005    | 현상       | 낙엽으로 덮인 지면 | 120            |
| ..     | ..       | ..         | ..             |

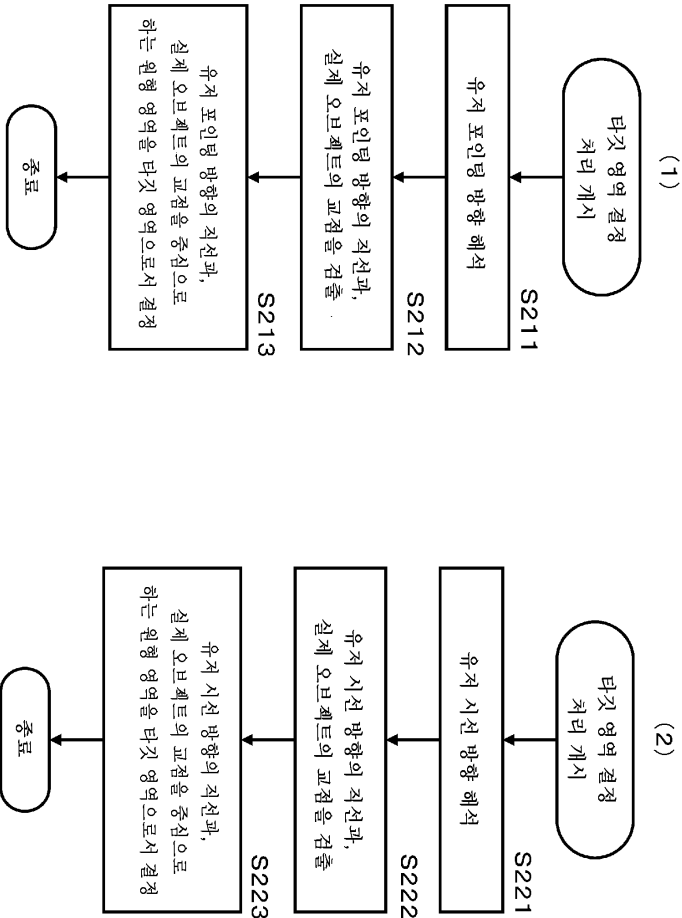
도면12

|         |                                                                                                 |                   |    |
|---------|-------------------------------------------------------------------------------------------------|-------------------|----|
| (a) 클래스 | (b) 가상 오브젝트 3D 모델<br>(캐릭터 3D 모델)                                                                | (c) 출력 음성         | .. |
| 연못      | 0000123(<br> ) | 물소리의<br>음성 데이터    | .. |
| 잔디밭     | 0023459(<br> ) | 풀소리의<br>음성 데이터    | .. |
| 나무      | 0314562(<br> ) | 나뭇잎 소리의<br>음성 데이터 | .. |
| ..      | ..                                                                                              | ..                | .. |

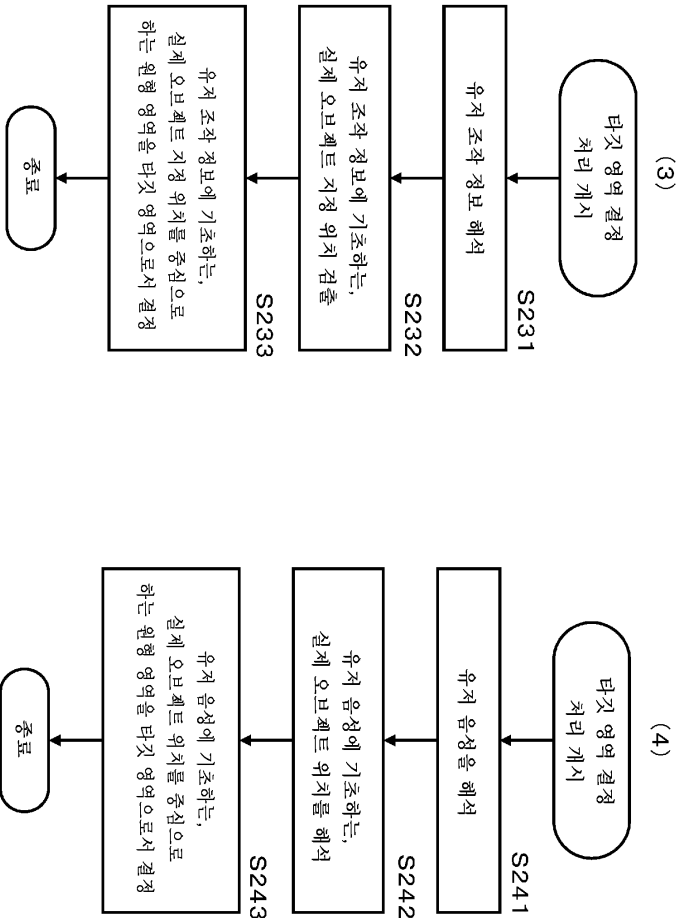
도면13



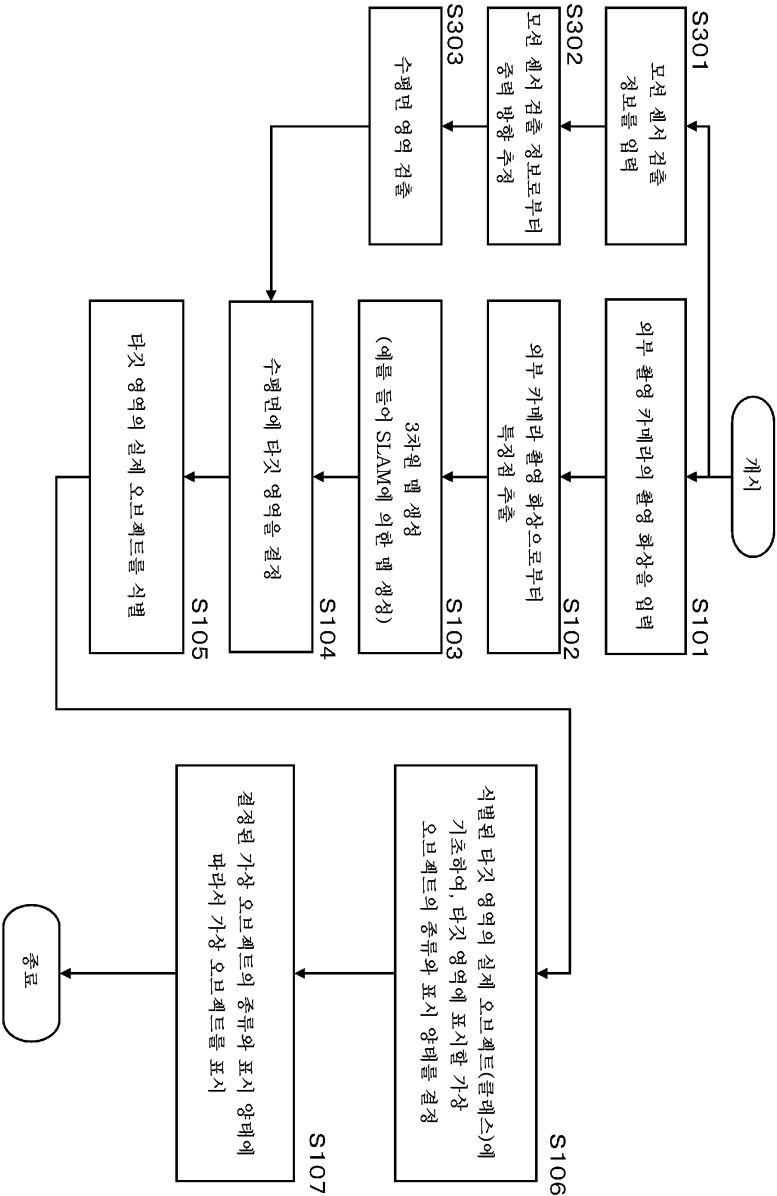
도면14



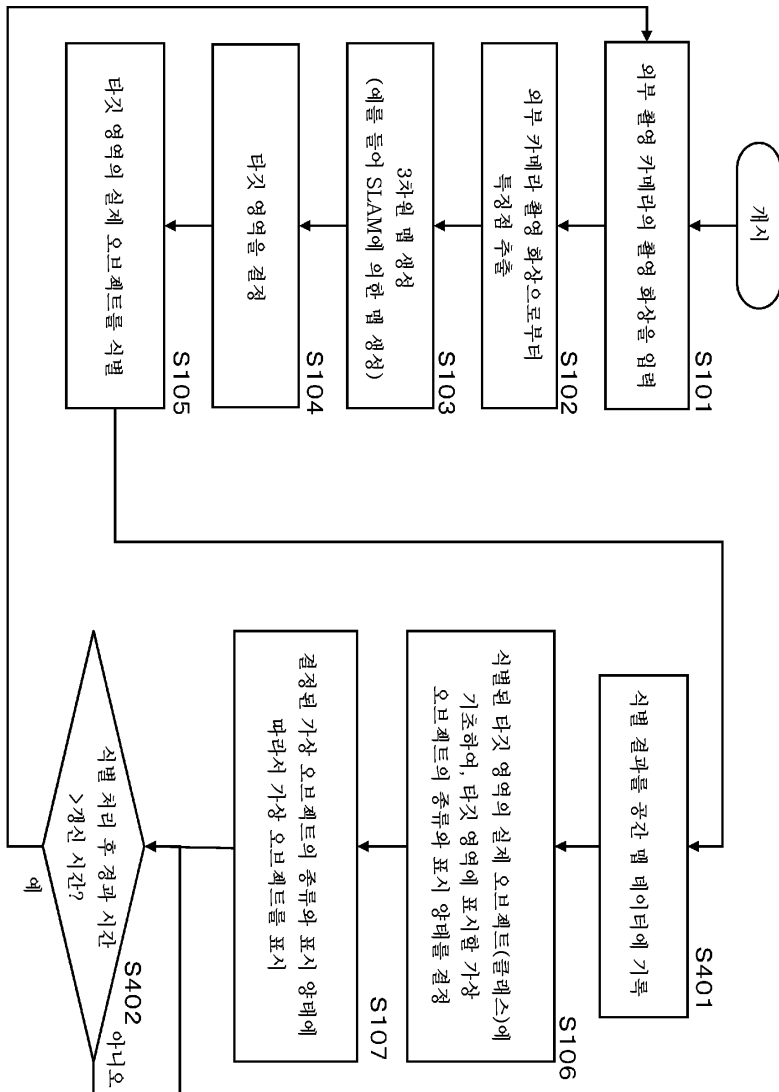
도면15



도면16



도면17



도면18

