(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0143228
(43) 공개일자 2020년12월23일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G06K 9/62* (2006.01)
(52) CPC특허분류
G06K 9/00362 (2013.01)
G06K 9/00201 (2013.01)
(21) 출원번호 10-2020-0039564
(22) 출원일자 2020년04월01일
심사청구일자 2020년04월01일
(30) 우선권주장
1020190070829 2019년06월14일 대한민국(KR)

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
도락주
서울특별시 성북구 종암로21길 127 ,104동303호(종암동,현대아이파크1차아파트)
임가현
서울특별시 성북구 인촌로22길 5, 407호
(뒷면에 계속)
(74) 대리인
김연권

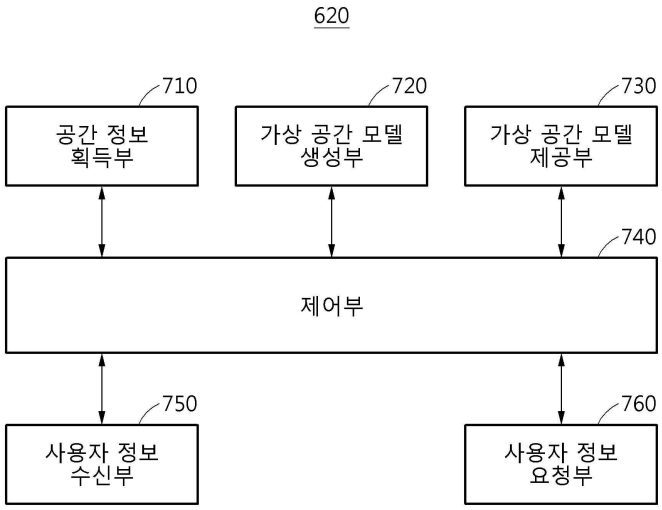
전체 청구항 수 : 총 25 항

(54) 발명의 명칭 **3차원 가상 공간 모델을 이용한 사용자 포즈 추정 방법 및 장치**

(57) 요 약

3차원 가상 공간 모델을 이용한 사용자 포즈 추정 방법 및 장치를 개시한다. 3차원 공간에 대한 사용자의 위치(position) 및 방향(orientation) 정보를 포함하는 사용자 포즈를 추정하는 방법은, 상기 3차원 공간에서 획득된 영상을 포함하는 사용자 정보를 수신하는 단계와, 상기 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모델을 확인하는 단계와, 상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하는 단계와, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하는 단계 및 상기 유사도에 기초하여 사용자 포즈를 추정하는 단계를 포함한다.

대 표 도 - 도7



(52) CPC특허분류
G06K 9/00624 (2013.01)
G06K 9/6201 (2013.01)

(72) 발명자
현장훈
서울특별시 성북구 고려대로17가길 15-6, 301호 (안암동2가)
김동우
서울특별시 동대문구 제기로5길 46, 302호 (제기동)
장범철
서울특별시 동대문구 무학로47길 10, 302호 (용두동)
최형아
서울특별시 동대문구 한빛로 20, 1106호(신설동)

이 발명을 지원한 국가연구개발사업

과제고유번호	10073166
과제번호	10073166
부처명	산업통상자원부
과제관리(전문)기관명	한국산업기술평가관리원
연구사업명	산업핵심기술개발사업
연구과제명	복합 실내 자율주행 솔루션 확보를 위한 3D의미지도-스마트폰-더미로봇 융합기술
기 여 율	1/1
과제수행기관명	고려대학교 산학협력단
연구기간	2016.12.01 ~ 2021.11.30

명 세 서

청구범위

청구항 1

깊이 측정 장치 및 영상 획득 장치를 이용하여 3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득하는 단계;

상기 공간 정보에 기초하여 깊이-영상 연계 정보를 구성하고, 상기 깊이-영상 연계 정보에 기초하여 상기 3차원 공간에 대응하는 3차원 가상 공간 모델을 구축하는 단계;

상기 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신하는 단계;

상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하는 단계;

상기 대응 정보와 상기 사용자 정보의 유사도를 산출하는 단계; 및

상기 유사도에 기초하여 사용자 포즈로 추정하는 단계

를 포함하는 3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 2

제1항에 있어서,

상기 3차원 가상 공간 모델을 구축하는 단계는,

상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 상기 3차원 가상 공간 모델을 구축하는 것을 특징으로하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 3

제1항에 있어서,

상기 대응 정보를 생성하는 단계는,

상기 사용자 정보에 포함된 영상에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하는 단계;

상기 사용자 정보에 포함된 영상의 배경 영역을 이용하여 상기 사용자 정보를 가공하는 단계; 및

상기 3차원 가상 공간 모델 내에서 가공된 사용자 정보에 대응하는 대응 정보를 생성하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 4

제1항에 있어서,

상기 유사도를 산출하는 단계는,

상기 유사도를 높이는 방향으로 상기 대응 정보를 재 생성하는 단계; 및

재 생성된 대응 정보에 기초하여 유사도를 재 산출하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 5

제1항에 있어서,

상기 유사도를 산출하는 단계는,

상기 사용자 정보 및 상기 대응 정보를 비교하기 위한 비교 대상 영역을 추출하는 단계;

상기 사용자 정보에서 추출된 비교 대상 영역과 상기 대응 정보에서 추출된 비교 대상 영역에서 공통 영역을 결정하는 단계; 및

상기 공통 영역에 기초하여 상기 사용자 정보 및 상기 대응 정보를 각각 재 생성하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 6

제1항에 있어서,

상기 유사도를 산출하는 단계는,

상기 사용자 디바이스 주변에 대한 추가적인 사용자 정보를 획득하는 단계; 및

상기 사용자 정보 및 추가적인 사용자 정보에 기초하여 유사도를 산출하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 7

제6항에 있어서,

상기 사용자 포즈를 추정하는 단계는,

상기 사용자 디바이스를 통해 상기 사용자 포즈의 추정에 이용되는 부가적인 정보인 사용자 부가 정보가 획득되는 경우, 상기 사용자 정보 또는 상기 추가적인 사용자 정보와 함께 상기 사용자 부가 정보를 이용하여 상기 사용자 포즈를 추정하는 것을 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 8

제6항에 있어서,

상기 추가적인 사용자 정보를 획득하는 단계는,

상기 3차원 가상 공간 모델에 기초하여 추가적인 사용자 정보 취득을 위한 안내 정보를 상기 사용자 디바이스로 전송하는 것을 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 9

제8항에 있어서,

상기 안내 정보는 상기 3차원 가상 공간 모델 내의 기 설정된 특징점에 대한 사용자 정보 취득 포즈를 포함하고,

상기 추가적인 사용자 정보를 획득하는 단계는 상기 유사도를 높이는 방향으로 반복 수행되는 것을 특징으로 하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 10

3차원 공간에 대한 사용자의 위치(position) 및 방향(orientation) 정보를 포함하는 사용자 포즈를 추정하는 방법에 있어서,

상기 3차원 공간에서 획득된 영상을 포함하는 사용자 정보를 수신하는 단계;

상기 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모

델을 확인하는 단계;

상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하는 단계;

상기 대응 정보와 상기 사용자 정보의 유사도를 산출하는 단계; 및

상기 유사도에 기초하여 사용자 포즈를 추정하는 단계

를 포함하는 3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 11

제10항에 있어서,

상기 3차원 가상 공간 모델은 상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 구축된 것임을 특징으로 하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 12

제10항에 있어서,

상기 대응 정보를 생성하는 단계는,

상기 사용자 정보에 포함된 영상에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하는 단계;

상기 사용자 정보에 포함된 영상의 배경 영역을 이용하여 상기 사용자 정보를 가공하는 단계; 및

상기 3차원 가상 공간 모델 내에서 가공된 사용자 정보에 대응하는 대응 정보를 생성하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 13

제10항에 있어서,

상기 유사도를 산출하는 단계는,

상기 유사도를 높이는 방향으로 상기 대응 정보를 재 생성하는 단계; 및

재 생성된 대응 정보에 기초하여 유사도를 재 산출하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 14

제10항에 있어서,

상기 유사도를 산출하는 단계는,

상기 사용자 정보 및 상기 대응 정보를 비교하기 위한 비교 대상 영역을 추출하는 단계;

상기 사용자 정보에서 추출된 비교 대상 영역과 상기 대응 정보에서 추출된 비교 대상 영역에서 공통 영역을 결정하는 단계; 및

상기 공통 영역에 기초하여 상기 사용자 정보 및 상기 대응 정보를 각각 재 생성하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 15

제10항에 있어서,

상기 유사도를 산출하는 단계는,

사용자 디바이스 주변에 대한 추가적인 사용자 정보를 획득하는 단계; 및

상기 사용자 정보 및 추가적인 사용자 정보에 기초하여 유사도를 산출하는 단계를 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 16

제15항에 있어서,

상기 사용자 포즈를 추정하는 단계는,

상기 사용자 디바이스를 통해 상기 사용자 포즈의 추정에 이용되는 부가적인 정보인 사용자 부가 정보가 획득되는 경우, 상기 사용자 정보 또는 상기 추가적인 사용자 정보와 함께 상기 사용자 부가 정보를 이용하여 상기 사용자 포즈를 추정하는 것을 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 17

제15항에 있어서,

상기 추가적인 사용자 정보를 획득하는 단계는,

상기 3차원 가상 공간 모델에 기초하여 추가적인 사용자 정보 취득을 위한 안내 정보를 상기 사용자 디바이스로 전송하는 것을 포함하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 18

제17항에 있어서,

상기 안내 정보는 상기 3차원 가상 공간 모델 내의 기 설정된 특징점에 대한 사용자 정보 취득 포즈를 포함하고,

상기 추가적인 사용자 정보를 획득하는 단계는 상기 유사도를 높이는 방향으로 반복 수행되는 것을 특징으로 하는

3차원 공간에 대한 사용자 포즈 추정 방법.

청구항 19

3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득하는 공간 정보 획득부;

상기 공간 정보에 기초하여 깊이-영상 연계 정보를 구성하고, 상기 깊이-영상 연계 정보에 기초하여 상기 3차원 공간에 대응하는 3차원 가상 공간 모델을 생성하는 가상 공간 모델 생성부;

상기 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신하는 사용자 정보 수신부; 및

상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하고, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하고, 상기 유사도에 기초하여 상기 사용자 포즈를 추정하도록 구성된 적어도 하나의 프로세서를 포함하는 제어부

를 포함하는 3차원 공간에 대한 사용자 포즈 추정 장치.

청구항 20

제19항에 있어서,

상기 공간 모델 생성부는,

상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 상기 3차원 가상 공간 모델을 구축하

는 것을 특징으로하는

3차원 공간에 대한 사용자 포즈 추정 장치.

청구항 21

제19항에 있어서,

상기 제어부는

상기 사용자 정보에 포함된 영상에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 사용자 정보에 포함된 영상의 배경 영역을 이용하여 상기 사용자 정보를 가공하고, 상기 3차원 가상 공간 모델 내에서 가공된 사용자 정보에 대응하는 대응 정보를 생성하는

3차원 공간에 대한 사용자 포즈 추정 장치.

청구항 22

3차원 공간에 대한 사용자의 위치(position) 및 방향(orientation) 정보를 포함하는 사용자 포즈를 추정하는 장치에 있어서,

상기 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모델을 제공하는 가상 공간 모델 제공부;

상기 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신하는 사용자 정보 수신부; 및

상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하고, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하고, 상기 유사도에 기초하여 상기 사용자 포즈를 추정하도록 구성된 적어도 하나의 프로세서를 포함하는 제어부

를 포함하는 3차원 공간에 대한 사용자 포즈 추정 장치.

청구항 23

제22항에 있어서,

상기 3차원 가상 공간 모델은 상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 구축된 것임을 특징으로 하는

3차원 공간에 대한 사용자 포즈 추정 장치.

청구항 24

3차원 공간에 대한 영상을 포함하는 사용자 정보를 생성하는 사용자 정보 생성부;

상기 사용자 정보를 사용자 포즈 추정 서버로 전송하고, 3차원 가상 공간 모델에서 추정된 사용자 포즈에 대한 정보를 상기 서버로부터 수신하는 통신부; 및

상기 사용자 정보 생성부 및 통신부의 동작을 제어하고, 상기 사용자 포즈에 대한 정보를 현재 구동 중인 어플리케이션 또는 구동 시스템으로 전달하도록 구성된 적어도 하나의 프로세서를 포함하는 제어부

를 포함하는 3차원 공간에 대한 사용자 포즈 추정 장치.

청구항 25

제24항에 있어서,

상기 3차원 가상 공간 모델은 상기 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 생성되고, 상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 구축된 것임을 특징으로 하는

3차원 공간에 대한 사용자 포즈 추정 장치.

발명의 설명

기 술 분 야

[0001] 기술분야는 공간 지도의 생성 및 활용에 관한 것으로서, 보다 구체적으로 3차원 가상 공간 모델을 이용하여 현실 공간 내 사용자의 포즈를 추정하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 공간지도를 활용하여 사용자 포즈(pose)를 추정하는 방법은 지형 값(geometry)을 사용하는 방식, 영상 값을 사용하는 방식, 지형 값과 영상 값을 혼합하여 사용하는 방식이 있다.

[0003] 이때, 라이다(LiDAR) 혹은 유사한 동작원리를 가지는 깊이측정장치를 활용하여 점구름정보를 취득하거나, 카메라 혹은 유사한 동작원리를 가지는 영상측정장치를 활용하여 영상정보를 취득하거나, 키넥트(Kinect) 혹은 유사한 동작원리를 가지는 깊이-영상측정장치를 활용하여 색-점구름정보를 취득하거나 이들의 조합으로 현실 공간을 표현하는 공간지도를 구성할 수 있다.

[0004] 현실 공간에 대한 영상 정보, 깊이 정보 및 깊이-영상 연계 정보는 '공간 정보'라 칭해질 수 있다.

[0005] 현실 공간에서 사용자 디바이스를 통해 획득한 사용자 정보를 공간 지도와 비교함으로써, 사용자 포즈를 추정할 수 있다.

[0006] 여기서, '사용자 정보'는 현실 공간에서 사용자 디바이스를 통해 획득한 영상을 포함하는 정보이다. 그리고, '포즈'는 위치(position)와 방향(orientation)을 모두 포함하는 개념이다. 따라서, '사용자 포즈'는 현실 공간에 대해 영상 데이터를 획득한 위치 정보 및 영상 데이터를 획득한 방향 정보를 포함하는 정보일 수 있다.

[0007] 한편, 종래 기술에 따른 공간 지도를 이용한 사용자 포즈 추정은 다음과 같은 문제점들이 있다.

[0008] 첫째, 공간 지도는 공간 정보를 획득하는 포즈에 민감할 수 있다. 따라서, 공간 지도가 공간 정보를 획득하는 포즈에 민감한 경우 사용자 포즈 추정의 정밀도는 저하될 수 있다. 예를 들어, 이론적으로 가능한 모든 포즈에서 공간 정보를 획득하고 공간 지도를 구성하는 경우, 정밀도가 높은 사용자 포즈 추정이 가능하다.

[0009] 그러나, 현실 공간에 대해 모든 포즈에서 공간 정보를 획득하는 것은 사실상 불가능하다. 공간 지도를 구성하기 위해 현실 공간에서 고른 분포로 다수의 포즈에서 공간 정보를 획득한 경우 공간 정보를 획득하는 포즈의 분포에 의해 사용자 포즈 추정의 민감도를 낮출 수 있다. 그러나, 이 경우 공간 정보의 취득 시간, 공간 정보의 용량 및 처리 속도 등의 시스템 부하 문제가 야기될 수 있다.

[0010] 한편, 시스템 부하 문제의 효율을 고려하여 소수의 포즈에서 공간 정보를 획득하는 경우, 공간 지도가 현실 공간을 충분히 표현하지 못할 수 있다. 또한, 공간 지도를 획득하는 경로가 변경되면 공간 지도의 신뢰성이 저하되어 현실 공간을 강인하게 표현하지 못하게 된다. 현실 공간을 강인하게 표현하지 못하는 공간 지도는 사용자 포즈 추정의 정밀도를 저하시킬 수 있다.

[0011] 둘째, 불연속적인 공간 지도는 사용자 포즈 추정의 정밀도를 저하시킬 수 있다. 도 1은 점구름정보를 활용하여 구성된 불연속적인 공간 지도의 예를 나타내는 도면이다.

[0012] 도 1에 도시된 바와 같이, 점구름정보를 활용하여 공간지도를 구성하는 경우 공간 정보 취득 범위, 경로에 따라 점구름정보가 조밀하게 획득되지 못할 수 있다. 점구름정보가 조밀하게 획득되지 못한 경우 불연속 공간지도가 생성될 수 있고, 이는 사용자 포즈 추정의 정밀도를 저하시킬 수 있다.

[0013] 셋째, 공간지도를 구성하기 위한 공간 정보가 취득된 시점과 사용자 정보를 취득하는 시점의 차이로 인해 사용자 포즈 추정의 정밀도가 저하될 수 있다.

[0014] 도 2 및 도 3은 시간의 흐름에 따른 공간의 변화를 나타내는 예시도들이다.

[0015] 도 2는 빛 또는 조명이 시간의 변화에 따라 변하는 예를 나타낸다.

[0016] 보다 구체적으로, 도 2의 (a), (b) 및 (c)는 동일한 공간이 시간의 흐름에 따라 조명이나 외부로부터 유입되는 빛의 양이 변경된 예시를 나타낸다.

[0017] 또한, 도 3의 (a) 및 (b)는 동일한 공간에서 시간의 흐름에 따라 사물이 달라진 예를 나타낸다.

[0018] 도 3의 (a)에서 탁자(210) 위에는 아무런 사물이 놓여 있지 않지만, 도 3의 (b)는 탁자(220) 위에 여러가지 사물이 놓여 있음을 나타낸다.

[0019] 예를 들어, 도 2에 도시된 공간에 대해 공간 지도를 구성하기 위한 공간 정보는 (a)를 통해 획득되고, 사용자 정보는 (c)에서 획득될 수 있다. 또한, 도 3에 도시된 공간에 대해 공간 지도를 구성하기 위한 공간 정보는 (a)에서 획득되고 사용자 정보는 (b)에서 획득될 수 있다.

[0020] 이와 같이 동일한 공간에 대해 공간 정보가 획득된 시점과 사용자 정보를 획득하는 시점(time)의 차이로 인해 영상 정보는 일치하지 않을 수 있다. 따라서, 동일한 공간에 대해 공간 정보가 획득된 시점과 사용자 정보를 획득하는 시점의 차이로 인해 사용자 포즈 추정의 정밀도는 저하될 수 있다.

[0021] 현실 공간은 시간의 흐름에 따라 빛 또는 조명의 변화, 사람 등 동적인 움직임의 변화, 사물 또는 인테리어 등의 변화가 발생할 수 있다. 이런 변화들을 갱신하지 않은 공간지도를 사용할 경우 사용자 정보와 유사성이 적을 수 있고, 이는 사용자 포즈 추정의 정밀도를 저하시킬 수 있다.

[0022] 따라서, 공간 지도를 바탕으로 사용자 포즈 추정 시 발생하는 기존의 문제점들을 해결할 수 있는 방법이 필요하다.

발명의 내용

해결하려는 과제

[0023] 본 발명은 상술한 문제점을 해결하기 위해 현실 공간에서 획득된 공간 정보를 바탕으로 구성된 3차원 가상 공간 모델과 사용자가 취득한 사용자 정보를 활용하여 사용자 포즈를 추정하는 방법을 제공하고자 한다.

과제의 해결 수단

[0024] 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 방법은, 깊이 측정 장치 및 영상 획득 장치를 이용하여 3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득하는 단계와, 상기 공간 정보에 기초하여 깊이-영상 연계 정보를 구성하고, 상기 깊이-영상 연계 정보에 기초하여 상기 3차원 공간에 대응하는 3차원 가상 공간 모델을 구축하는 단계와, 상기 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신하는 단계와, 상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하는 단계와, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하는 단계 및 상기 유사도에 기초하여 사용자 포즈를 추정하는 단계를 포함한다.

[0025] 상기 3차원 가상 공간 모델을 구축하는 단계는, 상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 상기 3차원 가상 공간 모델을 구축하는 것을 포함할 수 있다.

[0026] 상기 대응 정보를 생성하는 단계는 상기 사용자 정보에 포함된 영상에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하는 단계와, 상기 사용자 정보에 포함된 영상의 배경 영역을 이용하여 상기 사용자 정보를 가공하는 단계 및 상기 3차원 가상 공간 모델 내에서 가공된 사용자 정보에 대응하는 대응 정보를 생성하는 단계를 포함할 수 있다.

[0027] 상기 유사도를 산출하는 단계는, 상기 유사도를 높이는 방향으로 상기 대응 정보를 재 생성하는 단계 및 재 생성된 대응 정보에 기초하여 유사도를 재 산출하는 단계를 포함할 수 있다.

[0028] 상기 유사도를 산출하는 단계는, 상기 사용자 정보 및 상기 대응 정보를 비교하기 위한 비교 대상 영역을 추출하는 단계와, 상기 사용자 정보에서 추출된 비교 대상 영역과 상기 대응 정보에서 추출된 비교 대상 영역에서 공통 영역을 결정하는 단계 및 상기 공통 영역에 기초하여 상기 사용자 정보 및 상기 대응 정보를 각각 재 생성하는 단계를 포함할 수 있다.

[0029] 상기 유사도를 산출하는 단계는, 상기 사용자 디바이스 주변에 대한 추가적인 사용자 정보를 획득하는 단계 및 상기 사용자 정보 및 추가적인 사용자 정보에 기초하여 유사도를 산출하는 단계를 포함할 수 있다.

[0030] 상기 사용자 포즈를 추정하는 단계는, 상기 사용자 디바이스를 통해 상기 사용자 포즈의 추정에 이용되는 부가적인 정보인 사용자 부가 정보가 획득되는 경우, 상기 사용자 정보 또는 상기 추가적인 사용자 정보와 함께 상

기 사용자 부가 정보를 이용하여 상기 사용자 포즈를 추정하는 것을 포함할 수 있다.

[0031] 상기 추가적인 사용자 정보를 획득하는 단계는, 상기 3차원 가상 공간 모델에 기초하여 추가적인 사용자 정보 취득을 위한 안내 정보를 상기 사용자 디바이스로 전송하는 것을 포함할 수 있다.

[0032] 상기 안내 정보는 상기 3차원 가상 공간 모델 내의 기 설정된 특징점에 대한 사용자 정보 취득 포즈를 포함하고, 상기 추가적인 사용자 정보를 획득하는 단계는 상기 유사도를 높이는 방향으로 반복 수행될 수 있다.

[0033] 다른 일 실시예에 따른, 3차원 공간에 대한 사용자의 위치(position) 및 방향(orientation) 정보를 포함하는 사용자 포즈를 추정하는 방법은, 상기 3차원 공간에서 획득된 영상을 포함하는 사용자 정보를 수신하는 단계와, 상기 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모델을 확인하는 단계와, 상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하는 단계와, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하는 단계 및 상기 유사도에 기초하여 사용자 포즈를 추정하는 단계를 포함한다.

[0034] 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 장치는 3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득하는 공간 정보 획득부와, 상기 공간 정보에 기초하여 깊이-영상 연계 정보를 구성하고, 상기 깊이-영상 연계 정보에 기초하여 상기 3차원 공간에 대응하는 3차원 가상 공간 모델을 생성하는 가상 공간 모델 생성부와, 상기 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신하는 사용자 정보 수신부 및 상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하고, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하고, 상기 유사도에 기초하여 상기 사용자 포즈를 추정하도록 구성된 적어도 하나의 프로세서를 포함하는 제어부를 포함한다.

[0035] 상기 공간 모델 생성부는, 상기 3차원 공간에 대한 영상 데이터에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 배경 영역을 이용하여 상기 3차원 가상 공간 모델을 구축할 수 있다.

[0036] 상기 제어부는 상기 사용자 정보에 포함된 영상에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 상기 사용자 정보에 포함된 영상의 배경 영역을 이용하여 상기 사용자 정보를 가공하고, 상기 3차원 가상 공간 모델 내에서 가공된 사용자 정보에 대응하는 대응 정보를 생성할 수 있다.

[0037] 다른 일 실시예에 따른 3차원 공간에 대한 사용자의 위치(position) 및 방향(orientation) 정보를 포함하는 사용자 포즈를 추정하는 장치는, 상기 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모델을 제공하는 가상 공간 모델 제공부와, 상기 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신하는 사용자 정보 수신부 및 상기 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성하고, 상기 대응 정보와 상기 사용자 정보의 유사도를 산출하고, 상기 유사도에 기초하여 상기 사용자 포즈를 추정하도록 구성된 적어도 하나의 프로세서를 포함하는 제어부를 포함한다.

[0038] 또 다른 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 장치는, 3차원 공간에 대한 영상을 포함하는 사용자 정보를 생성하는 사용자 정보 생성부와, 상기 사용자 정보를 사용자 포즈 추정 서버로 전송하고, 3차원 가상 공간 모델에서 추정된 사용자 포즈에 대한 정보를 상기 서버로부터 수신하는 통신부 및 상기 사용자 정보 생성부 및 통신부의 동작을 제어하고, 상기 사용자 포즈에 대한 정보를 현재 구동 중인 어플리케이션 또는 구동 시스템으로 전달하도록 구성된 적어도 하나의 프로세서를 포함하는 제어부를 포함한다.

발명의 효과

[0039] 본 발명의 실시예들은 공간지도로 3차원 가상 공간 모델을 사용함으로써 공간 정보 획득 경로에 강인한 3차원 가상 공간 모델을 구축할 수 있고, 공간 정보 획득 포즈에 따른 사용자 포즈 추정 정밀도의 민감도를 감소시킬 수 있다.

[0040] 또한, 본 발명의 실시예에 따른 3차원 가상 공간 모델은 현실 공간과 유사하게 구성될 수 있음에도 불구하고, 공간 정보 취득시간, 공간 정보 용량, 데이터 처리속도 등이 감소할 수 있다.

[0041] 또한, 시간의 흐름에 따른 현실 공간의 변화에도 강인한 사용자 포즈 추정 방법이 제공될 수 있다.

[0042] 또한, 본 발명의 실시예들은 혼합 현실(Mixed Reality)에서 사용자 포즈를 추정할 때 활용될 수 있다.

[0043] 또한, 정밀한 사용자 포즈 추정으로 인해 현실 공간과 가상 공간 사이의 어색함이 감소할 수 있으며, 혼합 현실에서 사용자의 몰입도를 증가시킬 수 있다. 따라서, 본 발명의 실시예들은 혼합 현실 관련 기술의 상용화 및 발전에 기여할 수 있다.

도면의 간단한 설명

[0044] 도 1은 점구름정보를 활용하여 구성된 불연속적인 공간 지도의 예를 나타내는 도면이다.

도 2는 시간의 흐름에 따른 공간의 변화를 나타내는 예시도이다.

도 3은 시간에 흐름에 따른 공간의 변화를 나타내는 다른 예시도이다.

도 4는 본 발명의 일 실시예에 따른 3차원 가상 공간 모델의 예를 나타내는 도면이다.

도 5는 일 실시예에 따른 3차원 가상 공간 모델 생성의 예를 설명하기 위한 도면이다.

도 6은 일 실시예에 따른 3차원 가상 공간 모델을 이용한 사용자 포즈 추정 시스템을 설명하기 위한 도면이다.

도 7은 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 장치의 구성을 설명하기 위한 도면이다.

도 8은 일 실시예에 따른 사용자 디바이스의 구성을 설명하기 위한 도면이다.

도 9는 일 실시예에 따른 포즈의 개념을 설명하기 위한 예시도이다.

도 10은 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 방법을 설명하기 위한 흐름도이다.

도 11은 다른 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 방법을 설명하기 위한 흐름도이다.

도 12는 일 실시예에 따른 추가 사용자 포즈 획득 방법의 예를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0045] 이하 첨부 도면들 및 첨부 도면들에 기재된 내용들을 참조하여 본 발명의 실시예를 상세하게 설명하지만, 본 발명이 실시예에 의해 제한되거나 한정되는 것은 아니다.

[0046] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소, 단계, 동작 및/또는 소자는 하나 이상의 다른 구성요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0047] 본 명세서에서 사용되는 "실시예", "예", "측면", "예시" 등은 기술된 임의의 양상(aspect) 또는 설계가 다른 양상 또는 설계들보다 양호하다거나, 이점이 있는 것으로 해석되어야 하는 것은 아니다.

[0048] 또한, '또는' 이라는 용어는 배타적 논리합 'exclusive or' 이기보다는 포함적인 논리합 'inclusive or' 를 의미한다. 즉, 달리 언급되지 않는 한 또는 문맥으로부터 명확하지 않는 한, 'x가 a 또는 b를 이용한다' 라는 표현은 포함적인 자연 순열들(natural inclusive permutations) 중 어느 하나를 의미한다.

[0049] 또한, 본 명세서 및 청구항들에서 사용되는 단수 표현("a" 또는 "an")은, 달리 언급하지 않는 한 또는 단수 형태에 관한 것이라고 문맥으로부터 명확하지 않는 한, 일반적으로 "하나 이상"을 의미하는 것으로 해석되어야 한다.

[0050] 또한, 본 명세서 및 청구항들에서 사용되는 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0051] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.

[0052] 한편, 본 발명을 설명함에 있어서, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는, 그 상세한 설명을 생략할 것이다. 그리고, 본 명세서에서 사용되는 용어(terminology)들은 본 발명의 실시예를 적절히 표현하기 위해 사용된 용어들로서, 이는 사용자, 운용자의

의도 또는 본 발명이 속하는 분야의 관례 등에 따라 달라질 수 있다. 따라서, 본 용어들에 대한 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.

[0053] 도 4는 본 발명의 일 실시예에 따른 3차원 가상 공간 모델의 예를 나타내는 도면이다.

[0054] 도 5는 일 실시예에 따른 3차원 가상 공간 모델 생성의 예를 설명하기 위한 도면이다.

[0055] 도 4를 참조하면, 통상적으로 3차원 가상 공간 모델은 obj, x3d 등과 같은 깊이-영상연계정보를 이용하여 구성된 현실 공간에 대한 모델이다. 예를 들어, 3차원 가상 공간 모델은 한국등록특허 제10-1835434호(발명의 명칭: 투영 이미지 생성 방법 및 그 장치, 이미지 픽셀과 깊이 값 간의 매핑 방법)를 통해 obj, x3d 등이 생성된 모델 또는 'TeeVR 모델'을 포함할 수도 있다.

[0056] 이때, 일 실시예에 따른 3차원 가상 공간 모델은 도 5에 도시된 바와 같이 3차원 가상 공간 모델을 구성하기 위해 배경 영역과 비배경 영역을 구분하고, 배경 영역 만을 사용하여 구성될 수 있다.

[0057] 도 5에서 (a)는 공간 정보에 포함된 영상 데이터이고, (b)는 비배경 영역을 제거한 영상을 나타내고, (c)는 배경 영역을 확장하여 영상 데이터를 생성한 예를 나타낸다.

[0058] 예를 들어, 배경 영역은 3차원 공간을 형성하는 건축물 자체의 구조를 의미하거나 또는 방문이나 창문과 같이 그 건축물에 부착된 구조물일 수 있다. 따라서, 영상 데이터에서 배경 영역은 3차원 공간의 구조와 관련된 영역이라 정의될 수 있다.

[0059] 도 5의 (a)에서 실내 공간에 위치한 다양한 물체들(책상, 책장, 의자 등)에 대응되는 영역은 비배경 영역에 해당될 수 있다. 도 5에서 (b)는 실내 공간에 위치한 다양한 물체들(흰색으로 표시된 영역)이 제거된 배역 영역을 나타낸다.

[0060] 3차원 가상 공간 모델은 실내 공간과 실외 공간을 모두 포함하는 개념으로, 독립적인 실내 공간, 독립적인 실외 공간 또는 실내와 실외가 연결된 공간 일수도 있다. 3차원 가상 공간 모델에 obj, x3d 등과 같은 모델(사람, 사물 등)을 추가할 수도 있으며, 3차원 가상 공간 모델은 모델이 추가된 3차원 가상 공간 모델 또한 포함하는 개념일 수 있다. 한편, 공간 지도는 3차원 대신 차원을 낮춰 2차원 가상 공간 모델을 사용할 수도 있다.

[0061] 3차원 공간 모델은 사전에 구축된 obj, x3d와 같은 모델을 사용할 수도 있고, 공간 정보를 새로 획득하여 구축할 수도 있고, 사전에 구축된 모델을 갱신하여 사용할 수도 있다. 3차원 공간 모델은 현실 공간과 유사하다고 판단될 수 있다.

[0062] 도 6은 일 실시예에 따른 3차원 가상 공간 모델을 이용한 사용자 포즈 추정 시스템을 설명하기 위한 도면이다.

[0063] 도 6을 참조하면, 3차원 가상 공간 모델을 이용한 사용자 포즈 추정 시스템은 사용자 디바이스(610) 및 사용자 포즈 추정 장치(620)를 포함한다. 일 실시예에 따른 사용자 포즈 추정 장치(620)는 서버(도시 되지 않음)에 구비되거나 사용자 디바이스(610)에 구비될 수도 있다.

[0064] 사용자 디바이스(610)는 현실 공간(601)에서 사용자 정보(611)를 획득하고 사용자 디바이스(610)에 사용자 정보(611)를 전송할 수 있다.

[0065] 사용자 포즈 추정 장치(620)는 장치 내부 또는 외부의 스토리지 시스템(602)에 저장된 3차원 가상 공간 모델(630) 및 사용자 정보(611)를 이용하여 사용자 포즈를 추정할 수 있다.

[0066] 사용자 포즈 추정 장치(620)는 3차원 가상 공간 모델(630)에서 사용자 포즈에 해당할 확률이 높은 대응 정보들(621)을 사용자 정보(611)와 비교함으로써 정확한 사용자 포즈를 추정할 수 있다.

[0067] 도 7은 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 장치의 구성을 설명하기 위한 도면이다.

[0068] 도 7을 참조하면, 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 장치(620)는 가상 공간 모델 제공부(730), 제어부(740) 및 사용자 정부 수신부(750)를 포함한다. 또한, 사용자 포즈 추정 장치(620)는 공간 정보 획득부(710) 및 가상 공간 모델 생성부(720)를 더 포함할 수 있다. 또한, 사용자 포즈 추정 장치(620)는 사용자 정보 요청부(760)를 더 포함할 수 있다.

[0069] 공간 정보 획득부(710)는 3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득한다. 예를 들어, 공간 정보는 깊이측정장치 및 영상측정장치를 이용하여 획득될 수 있다.

[0070] 3차원 가상 공간 모델을 구성하기 위한 공간 정보를 취득하는 깊이측정장치 또는 영상측정장치 등으로 구성된

측정 장치의 시야각(FoV, Field of View)이 현실 공간을 확보할 수 있는 경로로 공간 정보를 취득하면, 3차원 가상 공간 모델은 현실 공간과 유사하게 구성될 수 있으며, 공간 정보 취득시간, 공간 정보 용량, 데이터 처리 속도 등이 감소할 수 있어 효율적일 수 있다.

[0071] 영상정보는 3차원 공간에 대한 2차원 이미지로서 2자유도의 기저벡터(basis vector)로 표현 가능한 형태를 가질 수 있으며, 카메라와 같이 3차원을 2차원으로 표현하는 형태 또는 카메라에 적외선 필터를 장착하여 3차원 열정보를 2차원으로 표현한 형태일 수 있다.

[0072] 깊이정보는 3자유도의 기저 벡터로 표현 가능한 점 형태를 가지며, 깊이측정장치를 사용하여 취득될 수도 있고, 서로 다른 곳에서 촬영된 두 개 이상의 이미지를 활용하여 추정될 수도 있다. 전자의 예로는 라이다(LiDAR), 소나(SONAR), 적외선(InfraRed), TOF(Time Of Flight) 거리탐지기를 통해 취득된 깊이정보가 있고, 후자의 예로는 스테레오카메라, 다중카메라, 전방향 스테레오카메라 등을 통해 취득된 깊이정보가 있다. 한편, Kinect, JUMP, PrimeSense, Project Beyond 등의 디바이스를 이용하게 되면 깊이정보 및 영상정보를 동시에 취득할 수도 있다.

[0073] 예컨대, 본 발명의 일 실시예 에서는 깊이 측정 장치를 통해 취득한 깊이정보 뿐만 아니라 내삽법(interpolation)을 통해 깊이정보를 새롭게 추정하여 사용할 수도 있다. 보다 구체적으로는, 복수의 취득 깊이정보 중에서 3개 이상의 깊이정보를 선별하여 다각형(삼각형 포함) 메쉬(Mesh)를 구성한 후, 그 다각형 메쉬 내부에 내삽법을 통해 새로운 깊이정보를 추정하여 추가할 수 있다.

[0074] 한편, 본 발명의 일 실시예에 따른 취득 깊이정보 및 영상정보는 통합 센서 시스템을 이용하여 동시에 취득될 수도 있다. 여러 개의 측정장치를 사용하는 경우, 센서 간의 좌표 관계를 구하는 캘리브레이션 과정이 필요할 수 있다.

[0075] 공간 정보를 취득하는 과정에서 관성측정장치(IMU) 등을 추가적으로 사용할 수도 있고, 바퀴형 이동로봇에 센서를 부착하여 측정하는 경우 거리정보(odometry)를 활용할 수도 있다. 현실 공간이 측정장치의 시야각보다 넓은 경우, 센서의 회전, 이동 또는 이를 조합하여 공간 정보를 취득할 수 있다. 이때, 개별 공간 정보가 취득된 3차원 포즈(pose)가 각각 상이할 수 있으며, 개별 공간 정보가 취득된 포즈를 예측하기 위해 슬램(SLAM, Simultaneous Localization And Mapping), 영상-관성 거리정보(VIO, Visual Inertial Odometry), 영상 거리정보(VO, Visual Odometry) 등의 기법을 활용할 수 있다.

[0076] 한편, 공간 정보의 구성은 측정장치의 종류에 따라 상이할 수 있다. 일예로 측정장치가 단일 카메라로만 구성된 경우 사전측정정보는 카메라 영상정보로 구성되며, 해당 영상정보를 활용하여 단일 카메라인 경우 픽셀들 간의 상대적 거리를, 복수의 카메라인 경우에는 픽셀들 간의 절대 거리를 예측할 수 있다. 특히, 특징점을 추출하지 않고도 단일 카메라의 경우 누적된 영상정보를 활용하여 픽셀의 깊이를 예측할 수 있으며, 복수의 카메라의 경우 복수의 카메라 영상이나 이의 누적된 영상정보를 활용하여 픽셀의 깊이를 예측할 수도 있다.

[0077] 나아가, 추가적인 깊이정보, 관성정보 등의 정보가 함께 활용된 경우, 각 측정장치의 고유특성에 맞게 공간 정보 처리가 가능할 수 있다. 일예로 관성측정장치를 사용하여 관성정보가 획득 가능한 경우에 이를 활용하여 슬램의 성능을 향상시키거나, 영상정보 처리 시 영상 취득지점에 대한 추정(prediction) 정보로 사용하여 영상 취득지점에 대한 보정(correction)이 보다 용이하도록 할 수 있다. 또한, 관성정보의 가속도 값 혹은 각속도 값을 활용하여 실제 이동 거리를 예상할 수 있고, 이를 단일 카메라 혹은 복수의 카메라로부터 추출된 깊이정보의 크기(scale)을 보정하는데 활용할 수도 있다.

[0078] 가상 공간 모델 생성부(720)는 공간 정보에 기초하여 깊이-영상 연계 정보를 구성하고, 상기 깊이-영상 연계 정보에 기초하여 상기 3차원 공간에 대응하는 3차원 가상 공간 모델을 생성한다.

[0079] 예를 들어, 실내 공간의 3차원 가상 공간 모델을 생성하는 경우 공간 정보 획득부(710)는 실내 공간에 대한 공간 정보인 실내 공간 이미지를 획득할 수 있다. 이때, 실내 공간 이미지는 실내 공간 내부의 다양한 위치에서 촬영된 이미지일 수 있다.

[0080] 이때, 가상 공간 모델 생성부(720)는 실내 공간의 구조에 대응되는 영역인 배경 영역 및 실내 공간에 위치한 물체 또는 이동 중인 사람에 대응되는 영역인 비배경 영역을 구분할 수 있다.

[0081] 가상 공간 모델 생성부(720)는 실내 공간 이미지를 구성하는 이미지의 픽셀 값에 기초하여 배경 영역과 비배경 영역을 구분할 수 있다.

[0082] 배경영역은 다른 요소들로 인해 일부 가려져 있어 데이터가 온전하지 않으나, 가려지지 않은 부분과 유사성을

가질 수 있을 것으로 추론되어, 가려지지 않은 부분을 홀필링(hole filling)이나 인페이팅(inpainting)기법을 사용하여 재구성해 낼 수 있는 부분에 해당할 수 있다. 또한, 배경영역은 건물 내부의 큰 간판이나 안내데스크 등과 같이 다른 물체를 가릴 수 있으나, 해당 물체의 엣지 부분에서 영상과 지형의 정합도가 모든 데이터내에서 일치하거나 별도의 정합 과정을 통해 일치시킬 수 있는 부분일 수 있다.

[0083] 가상 공간 모델 생성부(720)는 실내 공간 이미지에서 배경영역을 비배경영역으로 확장하여 적어도 하나의 확장 실내 공간 이미지를 생성할 수 있다. 예를 들어, 도 5의 (b)에서 배경 영역을 비배경 영역이 제거된 흰색으로 표현된 부분을 확장할 수 있다.

[0084] 가상 공간 모델 생성부(720)는 배경 영역에 포함된 엣지가 비배경 영역과의 경계선에서 단절된 경우, 엣지의 연장선이 배경 영역과 비배경 영역 간의 경계선을 넘어 비배경 영역으로 이어진다는 추론을 통해 확장된 이미지를 생성할 수 있다.

[0085] 이때, 특정한 실내 공간 이미지 이외에 다른 하나 이상의 실내 공간 이미지를 배경의 보완 이미지로 지정하고, 특정한 실내 공간 이미지의 비배경 영역에 해당하는 영역을 그 배경의 보완 이미지의 정보를 사용하여 줄여 나갈 수 있다.

[0086] 가상 공간 모델 생성부(720)는 적어도 하나의 확장된 실내 공간 이미지와 실내 공간에 대한 깊이 값의 정보를 포함하는 지형 정보에 기초하여 깊이-영상 연계 정보를 생성할 수 있다. 깊이-영상 연계 정보는 적어도 하나의 확장된 실내 공간 이미지의 픽셀들에 대응되는 실내 공간의 깊이 값이 매칭된 정보일 수 있다.

[0087] 가상 공간 모델 생성부(720)는 적어도 하나의 확장된 실내 공간 이미지와 지형정보 외에도, 적어도 하나의 확장된 실내 공간 이미지와 지형정보 각각의 취득 위치 및 취득 각도에 대한 정보를 포함하는 이미지 취득 포즈와 깊이 취득 포즈를 더 이용하여 깊이-이미지연계정보를 생성할 수 있다.

[0088] 가상 공간 모델 생성부(720)는 적어도 하나의 확장된 실내 공간 이미지, 지형정보 및 깊이-영상 연계정보를 이용하여 현실의 3차원 공간에 대한 3차원 가상 공간 모델을 생성한다.

[0089] 가상 공간 모델 제공부(730)는 사용자 포즈 추정이 필요한 경우 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모델을 제공한다.

[0090] 이때, 사용자 포즈 추정은 사용자 디바이스(610) 또는 사용자 포즈 추정 장치(620)에 설치된 어플리케이션이 구동된 후 수행될 수 있다. 가상 공간 모델 제공부(730)는 사용자 디바이스(610) 또는 사용자 포즈 추정 장치(620)에서 구동중인 어플리케이션 또는 해당 장치의 구동 시스템에게 3차원 가상 공간 모델을 제공할 수 있다.

[0091] 제어부(740)는 적어도 하나의 프로세서를 포함할 수 있다. 이때 제어부(740)는 명령어(instructions) 또는 프로그램이 기록된 적어도 하나의 컴퓨터 인식 가능 스토리지(one or more computer - readable storage media)와 연결될 수 있다.

[0092] 따라서, 제어부(740)는 3차원 가상 공간 모델 내에서 사용자 정보에 대응하는 대응 정보를 생성하고, 대응 정보와 사용자 정보의 유사도를 산출하고, 유사도에 기초하여 사용자 포즈를 추정하도록 구성된 적어도 하나의 프로세서를 포함한다.

[0093] 일 실시예에 따른 사용자 포즈 추정은 딥러닝(Deep Learning) 또는 신경망 네트워크를 활용하여 3차원 가상 공간 모델을 학습함으로써 수행될 수 있다.

[0094] 학습 문제의 형태에 따라 강화학습(reinforcement learning), 지도학습(supervised learning), 비지도학습(unsupervised learning)로 나눠질 수 있다. 학습단계에서 필요한 방대한 양의 훈련데이터(training test)가 필요할 수 있으며, 훈련데이터는 영상정보가 포함된 데이터와 그 데이터가 취득된 포즈가 포함된 데이터 등으로 구성될 수 있으며, 학습데이터의 양을 증가시키기 위하여 상기 두 종류의 데이터에 노이즈(noise)를 추가하여 변형시켜 사용할 수도 있다. 합성곱 신경망(CNN) 또는 다양한 신경망의 전체 또는 일부를 사용할 수도 있다. 딥러닝의 성능 또는 속도 향상을 위해 하나 이상의 GPU를 사용할 수도 있으며, 병렬 연산을 수행할 수도 있다. 딥러닝의 결과는 스칼라, 벡터, 확률 등으로 표현될 수 있으며, 이 결과를 이용하여 사용자 정보가 취득된 포즈로 예상되는 사용자 포즈를 추정할 수 있다. 입력으로 사용자 정보의 영상정보를 사용할 수 있으며, 사용자부가정보를 함께 사용할 수 있다. 사용자부가정보를 함께 사용하는 경우, 신경망에 레이어(layer)를 추가하거나 함수를 변화시키거나 파라미터 수를 조절하거나 그 값을 변경할 수도 있다. 신경망 네트워크를 구성하기 위해 파이썬, 씨언어, 매트랩 등 또는 이들의 조합으로 된 컴퓨터언어를 사용할 수도 있다.

[0095] 사용자 정보가 순차적으로 취득되는 경우, 3차원 가상 공간 모델을 기반으로 파티클 필터(Particle Filter), EKF. EIF, UKF 등의 기법을 활용하여 사용자 포즈를 추정할 수 있다. 사용자부가정보로 관성정보 또는 거리정보가 취득되는 경우 추정된 사용자 포즈를 보정할 수 있다. 순차적으로 취득된 사용자 정보에 따라 파티클 필터의 값이 특정 포즈로 수렴할 수 있으며, 이때 수렴된 지점을 사용자 포즈로 추정할 수 있다. 사용자 포즈 추정 시, 가중치(Weight)를 부과할 수 있고 다수의 수렴지점 중에서 사용자 포즈를 결정할 수도 있다.

[0096] 딥러닝으로 추정한 포즈와 파티클 필터 등으로 추정된 포즈를 융합하여 사용자 포즈를 추정할 수 있다. 예컨대, 딥러닝으로 추정한 포즈 주위에서 파티클 필터를 수행하여 사용자 포즈를 추정할 수 있으며, 반대의 방법으로 파티클 필터로 수렴된 포즈 주위에서 딥러닝으로 사용자 포즈를 추정할 수 있다. 사용자 포즈 추정 시, 가중치(Weight)를 부과할 수 있고 다수의 수렴지점 중에서 사용자 포즈를 결정할 수도 있다.

[0097] 유사도는 3차원 가상 공간 모델에서 생성한 대응 정보와 사용자 정보의 유사한 정도를 의미하며, 유사도가 높을수록 상기 대응 정보와 상기 사용자 정보가 비슷하다고 볼 수 있으며, 유사도가 높은 대응 정보가 생성된 3차원 가상 공간 모델의 포즈를 사용자 정보가 취득된 사용자 포즈로 추정할 수 있다. 유사도는 스칼라(scalar), 벡터(vector), 공분산행렬(covariance matrix) 등으로 표현될 수 있고, 유클리디안거리(euclidean distance), 맨해턴거리(manhattan distance), 마하라노비스거리(mahalanobis distance), 구조적유사도(SSIM, structural similarity), 정규정보거리(NID, normalized information distance), 최소평균제곱오차(MMSE, Minimun Mean Square Error), 엔트로피(Entropy) 등으로 산출될 수도 있다.

[0098] 유사도 산출 및 사용자 포즈 추정의 보다 상세한 설명은 도 10을 통해 설명하기로 한다.

[0099] 이때, 3차원 가상 공간 모델은 3차원 공간에 대한 영상 데이터에서 3차원 공간의 구조와 관련된 배경 영역과 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하고, 배경 영역을 이용하여 구축된 것일 수 있다.

[0100] 사용자 정부 수신부(750)는 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신한다.

[0101] 사용자 정보는 영상정보가 포함된 정보로, 영상측정장치를 하나 이상 포함하고, 깊이측정장치 또는 부가장치 등을 이용하여 취득할 수 있다. 측정장치의 시야각이 좁아 충분한 사용자 정보가 취득되지 않는 경우 측정장치 회전, 이동 또는 이를 조합하여 사용자 정보를 취득할 수 있다. 사용자 정보는 단일 또는 복수의 영상센서(카메라)로 취득할 수 있고, 핀홀(pin-hole)모델이나 어안렌즈, 또는 파노라마 형식으로 취득될 수도 있다. 단일 영상정보, 복수의 영상정보 또는 순차적인 영상정보로 취득될 수 있다. 취득한 사용자 정보를 이용하여 영상정보, 깊이정보 또는 깊이-영상연계정보 등을 구성할 수 있다.

[0102] 예컨대, 단일 영상측정장치를 사용하면 영상정보를 취득할 수 있고, 순차적으로 취득된 영상정보를 사용하면 깊이정보를 계산할 수 있고, 이를 통해 깊이-영상연계정보를 구성할 수 있다.

[0103] 예컨대, 복수의 영상측정장치를 사용하면 각 영상측정장치에서 취득된 영상정보와 영상측정장치 간의 관계를 활용하여 깊이정보를 계산할 수 있고, 이를 통해 깊이-영상연계정보를 구성할 수 있다. 영상측정장치 간의 관계는 영상측정장치 간의 캘리브레이션(Calibration)정보 또는 각 영상측정장치에서 취득된 영상정보 간의 변환정보(Homography matrix) 일 수 있다.

[0104] 예컨대, 적어도 하나 이상의 영상측정장치와 적어도 하나 이상의 깊이측정장치를 사용하는 경우 두 장치 간의 캘리브레이션 정보를 이용하여 깊이-영상연계정보를 구성할 수 있다. 딥러닝(Deep Learning)을 활용하여 영상정보에서 깊이정보를 추출할 수도 있다. 신경망 네트워크가 구성될 수 있고, 컨볼루션 네트워크가 사용될 수 있다. 학습 및 시험을 위한 다수의 데이터가 필요할 수 있고, 신경망 네트워크는 선형함수, 비선형함수, 다수의 레이어 등으로 구성될 수 있으며, 딥러닝의 결과는 확률, 스칼라, 벡터 등으로 표현할 수 있다. 반복 학습을 수행할 수도 있으며, 파라미터 튜닝이 필요할 수도 있다. 딥러닝으로 추출된 깊이정보를 활용하여 깊이-영상연계정보를 구성할 수 있다. 영상정보를 가공한 가공된 영상정보를 사용할 수도 있으며, 예컨대, 이미지의 명암, 채도 등을 변화시키거나, 파노라마 이미지를 교정 이미지(Rectified Image)로 변환시키는 작업 등이 수행될 수 있다.

[0105] 사용자 정보 요청부(760)는 추가적인 사용자 정보의 획득이 필요한 경우 사용자 디바이스(610)로 안내 정보를 전송할 수 있다. 안내 정보와 관련된 구체적인 설명은 도 12를 통해 설명하기로 한다.

[0106] 도 8은 일 실시예에 따른 사용자 디바이스의 구성을 설명하기 위한 도면이다.

[0107] 도 8을 참조하면, 사용자 디바이스(610)는 사용자 정보 생성부(810), 통신부(820) 및 제어부(830)를 포함한다.

사용자 디바이스(610)는 사용자와 인터페이스하기 위한 디스플레이, 입력 수단 및 출력 수단을 포함하는 사용자 인터페이스부(840)를 더 포함할 수 있다.

[0108] 사용자 정보 생성부(810)는 3차원 공간에 대한 영상을 포함하는 사용자 정보를 생성한다. 따라서, 사용자 정보 생성부(810)는 영상 측정 장치 및 깊이 측정 장치 중 적어도 하나를 포함할 수 있다.

[0109] 통신부(820)는 사용자 정보를 사용자 포즈 추정 서버로 전송하고, 3차원 가상 공간 모델에서 추정된 사용자 포즈에 대한 정보를 상기 서버로부터 수신한다.

[0110] 이때, 사용자 포즈 추정 서버는 도 7에 도시된 사용자 포즈 추정 장치(620)일 수 있고, 사용자 포즈 추정 서비스를 제공하는 별도의 서버일수 도 있다.

[0111] 제어부(830)는 사용자 정보 생성부(810) 및 통신부(830)의 동작을 제어하고, 사용자 포즈에 대한 정보를 현재 구동 중인 어플리케이션 또는 구동 시스템으로 전달하도록 구성된 적어도 하나의 프로세서를 포함한다.

[0112] 도 9는 일 실시예에 따른 포즈의 개념을 설명하기 위한 예시도이다.

[0113] 3차원 가상 공간 모델을 구축하는데 사용되는 공간 정보는 현실 공간의 일부 포즈에서 취득된 불연속정보로 고려될 수 있다. 여기서 포즈(pose)는 위치(position)과 방향(orientation)을 모두 포함하는 개념이다. 일예로, 2차원에서 포즈는 측정장치의 위치인 x, y와 측정장치의 각도 a로 표현될 수 있다.

[0114] 도 9에 도시된 예는 가로와 세로가 1m인 정사각형의 평면에서 측정장치가 x축, y축에 대해 0~1m의 범위에서 10cm 간격으로 이동하고, 회전각도는 0~360도의 범위에서 10도씩 회전할 수 있는 예를 나타낸다.

[0115] 이때, 가능한 전체 포즈의 경우의 수는 11*11*37으로 4477 가지이다. 마찬가지로, 3차원에서 포즈는 센서의 위치인 x, y, z와 측정장치의 각도 롤(roll), 피치(pitch), 요(yaw)로 표현될 수 있다.

[0116] 가로, 세로. 높이가 각 1m인 정육면체의 공간에서 센서가 x축, y축, z축에 대해 0~1m의 범위에서 10cm 간격으로 움직이고, 회전각도는 0~360도의 범위에서 각 10도씩 회전한다고 가정하면 가능한 전체 포즈의 경우의 수는 11*11*11*37*37*19이다. 즉, 약 3400만 가지의 경우의 수가 존재할 수 있다.

[0117] 측정장치의 이동간격과 회전각도 간격을 감소시키면 불연속정보를 연속정보처럼 보이게 할 순 있으나, 가능한 포즈의 수가 기하급수적으로 커질 것이며, 현실 공간의 부피는 $1m^3$ 보다 매우 크기 때문에 가능한 모든 포즈에서 공간 정보를 취득하는 것은 현실적으로 불가능하다.

[0118] 위와 같은 이유로 공간 정보를 취득하는 단계에서는 현실 공간을 충분히 담을 수 있는 일부 포즈에서 데이터를 취득하고, 상기 취득된 공간 정보를 기초로 하여 깊이-영상연계정보를 구성하여 3차원 가상 공간 모델을 구축함으로써 일부 포즈에서 취득한 공간 정보를 확장시킬 수 있다.

[0119] 3차원 가상 공간 모델은 일부 포즈에서 취득된 공간 정보를 기반으로 구성될 수 있지만, 공간 정보로부터 구성된 깊이-영상연계정보를 활용하기 때문에 3차원 가상 공간 모델 내 어떤 포즈에서도 동일한 현실 공간 내 포즈에서 취득된 공간 정보 또는 구성된 깊이-영상연계정보와 유사한 대응 정보를 생성할 수도 있다.

[0120] 즉, 공간 정보가 취득되지 않은 포즈에서도 3차원 가상 공간 모델을 사용하면 해당 포즈에서의 깊이-영상연계정보 또는 공간 정보와 유사한 대응 정보를 생성할 수 있으며, 상기 생성된 대응 정보는 현실 공간의 같은 포즈에서 취득한 공간 정보로부터 구성된 깊이-영상연계정보와 유사하다고 고려될 수 있다. 3차원 가상 공간 모델은 불연속정보인 공간 정보를 연속정보인 대응 정보로 변환시킬 수 있다.

[0121] 도 10은 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 방법을 설명하기 위한 흐름도이다.

[0122] 도 10에 도시된 방법은 도 7에 도시된 사용자 포즈 추정 장치(620)에 의해 수행될 수 있다.

[0123] S1010 단계에서 장치는, 깊이 측정 장치 및 영상 획득 장치를 이용하여 3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득한다.

[0124] S1020 단계에서 장치는, 공간 정보에 기초하여 깊이-영상 연계 정보를 구성하고, 깊이-영상 연계 정보에 기초하여 3차원 공간에 대응하는 3차원 가상 공간 모델을 구축한다.

[0125] S1030 단계에서 장치는, 3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신한다. 이때, 사용자 정보는 획득된 영상에 대응하는 공간에 대한 깊이 정보를 더 포함할 수 있다.

[0126] S1040 단계에서 장치는, 3차원 가상 공간 모델 내에서 사용자 정보에 대응하는 대응 정보를 생성한다.

[0127] 3차원 가상 공간 모델을 이용하면 공간 정보가 취득되지 않은 포즈일지라도 해당 포즈에서 취득된 공간 정보 또는 깊이-영상 연계정보와 유사할 수 있는 대응 정보를 생성할 수 있다.

[0128] 대응 정보는 깊이정보, 영상정보 또는 깊이-영상연계정보로 표현될 수 있다. 3차원 가상 공간 모델 내에서 3자유도의 기저벡터로 표현되는 포즈에서 대응 정보를 생성할 수 있다.

[0129] 예컨대, 사용자 정보 취득 포즈의 높이가 변하지 않는다면 3차원 가상 공간 모델 내에서 2자유도의 기저벡터로 표현되는 포즈에서 대응 정보를 생성할 수 있다. 대응 정보는 시야각, 영상정보 변환, 깊이정보 변환 등의 과정을 거쳐 생성될 수도 있다.

[0130] 이때, 대응 정보를 생성하는 S1040 단계는 사용자 정보에 포함된 영상에서 상기 3차원 공간의 구조와 관련된 배경 영역과 상기 3차원 공간에 놓인 물체에 해당하는 비배경 영역을 구분하는 단계, 사용자 정보에 포함된 영상의 배경 영역을 이용하여 상기 사용자 정보를 가공하는 단계 및 3차원 가상 공간 모델 내에서 가공된 사용자 정보에 대응하는 대응 정보를 생성하는 단계를 포함할 수 있다.

[0131] 사용자가 3차원 가상 공간 모델의 배경이 된 현실 공간에서 사용자 정보를 획득할 때, 현실 공간은 3차원 가상 공간 모델을 구성하기 위해 공간 정보를 획득한 시점과 동일하지 않을 수 있고, 사람, 사물, 인테리어 등 공간의 모습이 바뀌었을 수도 있다.

[0132] 따라서, 사용자 정보에서 배경부분과 비배경부분을 구분하고 사용자 정보에서 비배경부분을 제거할 수도 있고, 배경부분을 사용하여 사용자 정보를 변환시킬 수 있다. 조명, 빛 등에 의한 효과를 제거하기 위해 사용자 정보를 가공하여 사용할 수 있다. 사용자 정보를 3차원 공간 모델에서 생성한 대응 정보와 비교하는 과정에서 사용자 정보 또는 대응 정보의 형태를 변환하여 비교할 수 있다.

[0133] S1050 단계에서 장치는, 대응 정보와 사용자 정보의 유사도를 산출한다.

[0134] 이때, 유사도를 산출하는 단계는 유사도를 높이는 방향으로 대응 정보를 재 생성하고, 재 생성된 대응 정보에 기초하여 유사도를 재 산출하는 단계를 포함할 수 있다. 이때, 유사도를 높이는 방향은, 사용자 정보를 재 획득하거나 사용자 정보에 대응하는 대응 정보를 재 생성하거나 사용자 정보 이외에 부가적인 정보를 사용하는 것을 포함한다.

[0135] 유사도를 높이기 위하여 유사도를 산출하는 S1050 단계는, 사용자 정보 및 대응 정보를 비교하기 위한 비교 대상 영역을 추출하는 단계, 사용자 정보에서 추출된 비교 대상 영역과 대응 정보에서 추출된 비교 대상 영역에서 공통 영역을 결정하는 단계 및 공통 영역에 기초하여 사용자 정보 및 대응 정보를 각각 재 생성하는 단계를 포함할 수 있다.

[0136] 예를 들어, 대응 정보에서 구조 단순화로 인해 왜곡된 영역 등 소정의 기준에 따른 일정 영역을 제거하고, 사용자 정보의 비배경부분에 대응되는 영역을 제거하여 비교 과정에 사용될 대응 정보를 재 생성할 수 있다. 또한 사용자 정보에서 비배경 부분과 함께 대응 정보의 왜곡된 영역 등에 해당하는 영역을 제거하여 비교과정에 사용될 사용자 정보를 재 생성할 수 있다

[0137] 3차원 가상 공간 모델에서 생성된 대응 정보와 사용자가 획득한 사용자 정보의 유사성을 계산하는 방법은 대응 정보의 영상정보와 사용자 정보의 영상정보를 비교하는 방법, 대응 정보의 깊이정보와 사용자 정보의 깊이정보를 비교하는 방법 또는 깊이-영상연계정보를 비교하는 방법 등이 있을 수 있다.

[0138] 이때, 대응 정보와 사용자 정보의 크기(scale)가 다를 수 있으므로 정규화(Normalization)가 요구되거나 상대적인 비교가 필요할 수도 있다.

[0139] 한편, 영상 정보 간의 비교는, 각 영상 정보의 형식을 비슷하게 하기 위한 영상 정보의 변환이 필요할 수 있다. 예컨대, 파노라마 이미지(Panorama Image)와 보정 이미지(Rectified image) 간의 변환이 있을 수 있으며, 영상정보의 크기를 정규화 할 수 있으며, 시야각의 변환이 있을 수 있다.

[0140] 반대로 정류된 이미지를 파노라마 형식으로 변형시켜 사용할 수도 있다. 두 영상정보에서 RANSAC, SIFT, FAST, SURF 등의 기법 혹은 이들의 조합을 이용하여 영상정보의 특징점(Feature)을 찾고 유사한 특징점 쌍을 연결할 수 있다. 특징점은 에지(Edge), 직선, 선분, 코너(Corner), 원, 타원 등 또는 이들의 조합이 될 수 있고, 크기(scale), 회전(rotation) 등이 다를 수 있다. 특징점 매칭(Feature Matching), SSIM(Structural Similarity), NID(Normalized Information Distance), 호모 그래피 행렬(Homography Matrix) 등의 기법을 통해 영상 정보 간

의 유사도를 계산할 수 있다.

[0141] 특징점 매칭으로 맺어진 다수의 픽셀 좌표를 사용하여 호모그래피 행렬을 계산할 수 있으며, 이를 이용하여 두 이미지정보의 차이(오차, error)를 계산할 수 있다. SSIM는 두 이미지 간의 유사도를 계산하는 방법이고, NID는 확률적인 계산법이다.

[0142] 사용자 정보에서 깊이정보를 추출할 수 있는 경우, 대응 정보의 깊이정보와 유사도를 비교할 수 있다. 깊이 정보는 3차원 점구름정보(PCD, Point Cloud Data), 깊이지도(Depth Map), 메쉬(Mesh) 등으로 표현될 수 있으며, 두 깊이정보의 형식을 통일하는 과정이 필요할 수 있다. 깊이정보는 픽셀 간(포인트 당) 비교가 이루어질 수 있으며, 주변 영역을 고려하여 비교할 수도 있다. 내삽법(interpolation)을 통해 깊이정보를 새롭게 추정하여 비교할 수도 있으며, 가중치를 부과하여 산출할 수도 있다.

[0143] 사용자 정보에서 깊이-영상연계정보가 구성 가능한 경우, 대응 정보의 깊이-영상연계정보와 비교할 수 있다. 깊이정보와 영상정보를 각각 비교하여 각 유사도를 산출하고, 전체 유사도를 산출할 수 있으며, 각 유사도에 가중치를 주어 산출할 수 있다. 또는 깊이-영상연계정보를 복합적으로 비교할 수 있으며, 깊이정보와 영상정보의 유사도를 산출하는 방법들을 조합하여 수행될 수 있다.

[0144] 3차원 가상 공간 모델을 구축하기 위한 공간 정보가 취득된 시점과 사용자 정보가 취득된 시점이 다를 수 있기 때문에, 동일 포즈일지라도 대응 정보와 사용자 정보가 다를 수 있다. 따라서, 대응 정보와 사용자 정보에서 강인한 특징점을 비교할 수 있다. 예컨대, 대응 정보와 사용자 정보에서 배경부분과 비배경부분을 구분하고 배경부분을 사용하여 유사도를 산출할 수 있으며, 배경부분을 이용하여 구축된 3차원 가상 공간 모델에서 대응 정보를 생성하여 사용자 정보의 배경부분과의 유사도를 산출할 수 있다. 대응 정보와 사용자 정보에서 조명 또는 빛에 대한 광원효과를 제거하고 유사도를 산출할 수 있고, 광원효과에 강인한 특징을 비교하여 유사도를 산출할 수 있다.

[0145] 이때, 유사도를 산출하는 S1050 단계는, 사용자 디바이스 주변에 대한 추가적인 사용자 정보를 획득하는 단계 및 사용자 정보 및 추가적인 사용자 정보에 기초하여 유사도를 산출하는 단계를 포함할 수 있다. 추가적인 사용자 정보를 획득하기 위해, 도 12에 도시된 바와 같이 안내 정보를 이용할 수 있다.

[0146] S1060 단계에서 장치는, 유사도가 기 설정된 값 이상인 값을 갖는 후보 대응 정보를 확인하고, 상기 후보 대응 정보에 매칭되는 포즈를 사용자 포즈로 추정한다.

[0147] 유사도가 높을수록 대응 정보가 생성된 3차원 가상 공간 모델의 포즈와 사용자 정보가 취득된 포즈가 동일하다고 고려될 수 있다. 또는, 유사도가 기준값(threshold)보다 높으면 두 데이터가 취득되고 재구성된 포즈는 거의 같다고 볼 수 있으며 기준값은 현실 공간의 환경에 따라 달라질 수 있다. 또는, 다수의 후보 포즈에서 생성된 대응 정보 중 사용자 포즈와 유사도가 제일 높거나 어떤 판단 방식에 의해 선택된 포즈가 사용자 포즈로 고려될 수도 있다.

[0148] 대응 정보 생성, 유사도 계산 단계를 한 번만 수행하여 사용자 포즈를 추정할 수도 있고, 반복수행 할 수도 있다. 반복수행은 추려진 포즈 주위에서 정밀하게 재 추정할 수도 있고, 전체 영역에 대해 무작위(random)로 재 추정할 수도 있으며, 가중치를 부과하여 재 추정할 새로운 포즈를 선택할 수도 있다. 이러한 단계는 정해진 횟수만큼 반복될 수도 있으며, 유사도가 기준값 이상이거나, 반복되어 추정된 포즈가 수렴할 때까지 반복할 수도 있다. 유사도가 높아지도록 최적화기법(Optimization)을 사용할 수도 있다.

[0149] 유사도가 높아지도록 대응 정보를 재생성 할 수도 있으며, 재 생성된 대응 정보는 기존 대응 정보가 생성된 포즈와 유사도의 관계에 따라 사용자 포즈로 예상되는 포즈에서 재 생성될 수 있다. 대응 정보 재 생성 이후에 유사도를 산출하고 필요한 경우 대응 정보 재생성 및 유사도계산과정을 반복할 수 있다.

[0150] 관성정보 및 거리정보 등의 사용자부가정보를 사용하여 유사도가 높아질 것이라 기대되는 포즈에서 대응 정보를 생성 및 재생성 할 수 있다. 이후에 대응 정보와 사용자 정보의 유사도를 산출하고, 필요한 경우 사용자 부가정보를 활용하여 대응 정보를 재생성하고 유사도산출과정을 반복할 수 있다.

[0151] 이때, 사용자 부가 정보란 사용자가 취득한 영상정보 외에 사용자 포즈를 추정하는데 도움이 될 수도 있는 정보로, 관성정보(IMU), 거리정보(odometry) 등으로 구성될 수 있다. 일 예로 관성측정장치를 사용하여 관성정보가 획득 가능한 경우에 영상정보 처리 시 영상 취득 포즈에 대한 추정(prediction)정보로 사용하여 영상 취득 포즈에 대한 보정(correction)이 보다 용이하도록 할 수 있다.

[0152] 따라서, 유사도를 산출하는 S1050 단계 또는 사용자 포즈를 추정하는 S1060단계는 사용자 디바이스를 통해 상기

사용자 포즈의 추정에 이용되는 부가적인 정보인 사용자 부가 정보가 획득되는 경우, 상기 사용자 정보 또는 상기 추가적인 사용자 정보와 함께 상기 사용자 부가 정보를 이용하여 상기 사용자 포즈를 추정하는 것을 포함할 수 있다.

[0153] 이때, 관성 정보의 가속도 값 혹은 각속도 값을 활용하여 실제 이동 거리를 예상할 수 있고, 이를 단일 혹은 복수의 영상측정장치로부터 추출된 깊이정보의 크기(scale)을 보정하는데 활용할 수도 있다.

[0154] 거리정보는 사용자가 취득한 영상정보를 기반으로 구성된 VO(Visual Odometry), VIO(Visual Inertial Odometry)를 활용하여 예측한 거리정보일 수 있으며, 바퀴형 이동 로봇에 측정장치를 부착하여 사용자 정보를 취득하는 경우, 거리정보는 이동 로봇의 거리정보일 수 있다. 이에 관성정보를 활용하면 상기 방법들로 추출된 거리정보를 보정하는데 활용할 수도 있다.

[0155] 사용자 대신 바퀴형 이동 로봇에 센서를 부착하여 사용자 정보를 취득하는 경우, 사용자가 이동 로봇을 조종할 수도 있고, 이동 로봇이 자율주행을 할 수도 있고, 이 둘의 조합으로 사용자 정보를 취득할 수 있다. 이동 로봇 포즈를 사용자 포즈로 고려할 수 있으며, 이동 로봇과 사용자 시야 간의 좌표변환 관계를 알고 있거나, 좌표변환이 가능하면 이동 로봇 포즈를 사용자 포즈로 변환시킬 수 있다.

[0156] 이동 로봇은 영상이 포함된 사용자 정보뿐 아니라 모바일로봇의 거리정보(odometry)를 사용자부가정보로 취득할 수 있다. 거리정보를 활용하여 사용자 포즈를 보정할 수 있다. 순차적으로 취득한 거리정보를 활용하여 이동 로봇의 상대 예상 포즈를 예측할 수 있으며, EKF, EIF, UKF 등의 기법 또는 비슷한 방법을 활용하여 공분산행렬(Covariance matrix) 등의 정보를 계산할 수 있고, 이 정보를 갱신하여 사용자 포즈를 보정할 수 있다.

[0157] 이동 로봇을 이용할 경우, 이동 로봇의 동작, 운전, 조종, 이동, 데이터 취득, 저장 및 처리 등 관련 알고리즘은 로봇운영체제(ROS, robot operating system) 상에서 수행될 수 있다.

[0158] 공간 정보, 깊이-영상연계정보, 3차원 가상 공간 모델, 사용자 정보, 사용자부가정보 등은 외부의 서버(Server)에 저장 및 처리될 수 있다.

[0159] 공간 정보가 취득되는 동시에 깊이-영상연계정보가 구성되고 3차원 가상 공간 모델이 구축될 수 있으며, 사용자 정보가 취득되는 동시에 사용자 포즈를 실시간으로 추정할 수도 있고, 지연(Latency)이 있을 수도 있으며, 사용자 포즈 취득이 끝난 후에 처리될 수도 있다.

[0160] 3차원 가상 공간 모델이 구축되어 있다면, 추가 공간 정보를 취득하지 않을 수 있고, 일부의 공간에 대해 추가 공간 정보가 취득될 수 있다. 추가 공간 정보가 취득되지 않은 경우 구축된 3차원 가상 공간 모델을 사용할 수 있고, 추가공간 정보가 취득된 경우 구축된 3차원 가상 공간 모델을 부분 또는 전체 갱신하여 재구축하여 사용할 수 있다.

[0161] 사용자 정보를 선취득 후 공간 정보를 후 취득하여 3차원 가상 공간 모델을 구축하여 사용자 포즈를 추정할 수도 있고, 3차원 가상 공간 모델을 구축하기 위한 공간 정보를 선취득 후 사용자 정보를 후 취득하여 사용자 포즈를 추정할 수 있다.

[0162] 본 발명은 센서시스템과 컴퓨터가 융합된 시스템에서 수행될 수 있으며, 독립적인 센서시스템과 컴퓨터에서 수행될 수도 있다.

[0163] 사용자 정보 획 득시, 각 측정장치의 포즈와 사용자센서시스템 전체의 포즈가 다를 순 있으나, 각 측정장치 및 센서시스템의 좌표변환 관계를 사용하여 변환이 가능하다. 예컨대, 사용자 센서 시스템의 중심 또는 적절한 위치를 사용자 포즈로 가정하거나, 사용자 센서 시스템을 기준으로 사용자 포즈로 가정할 수 있다. 이 경우 필요한 캘리브레이션 정보 또는 사용자센서시스템으로부터 사용자 포즈까지의 상대 포즈를 알고 있거나 어떤 값으로 가정할 수 있다.

[0164] 도 11은 다른 일 실시예에 따른 3차원 공간에 대한 사용자 포즈 추정 방법을 설명하기 위한 흐름도이다.

[0165] 도 11에 도시된 방법은 도 7에 도시된 사용자 포즈 추정 장치(620)에 의해 수행될 수 있다.

[0166] S1110 단계에서 장치는 3차원 공간에서 획득된 영상을 포함하는 사용자 정보를 수신한다.

[0167] S1120 단계에서 장치는 3차원 공간에 대한 깊이 정보 및 영상 정보를 포함하는 공간 정보에 기초하여 구축된 3차원 가상 공간 모델을 확인한다. 이때, 3차원 가상 공간 모델은 도 7의 가상 공간 모델 제공부(730)에 의해 제공될 수 있다.

[0168] S1130 단계에서 장치는 3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성한다.

[0169] S1140 단계에서 장치는 대응 정보와 상기 사용자 정보의 유사도를 산출한다.

[0170] S1150 단계에서 장치는 유사도에 기초하여 사용자 포즈를 추정한다. 이때, 사용자 포즈는 예를 들어 사용자 정보와 가장 유사도가 높은 대응 정보의 포즈 일 수 있다.

[0171] 도 12는 일 실시예에 따른 추가 사용자 포즈 획득 방법의 예를 설명하기 위한 도면이다.

[0172] 유사도를 향상시키기 위하여 추가사용자 정보를 취득할 수 있으며, 3차원 가상 공간 모델을 활용하여 추가사용자 정보 취득 포즈를 사용자에게 안내할 수 있고, 사용자는 안내받은 포즈에서 추가사용자 정보를 취득할 수 있다.

[0173] 따라서, 도 10의 설명에서 추가적인 사용자 정보를 획득하는 단계는 3차원 가상 공간 모델에 기초하여 추가적인 사용자 정보 취득을 위한 안내 정보를 사용자 디바이스(610)로 전송하는 것을 포함할 수 있다.

[0174] 이때, 안내 정보는 3차원 가상 공간 모델 내의 기 설정된 특징점에 대한 사용자 정보 취득 포즈를 포함하고, 추가적인 사용자 정보를 획득하는 단계는 유사도를 높이는 방향으로 반복 수행될 수 있다.

[0175] 예컨대, 도 12에 도시된 바와 같이, 비슷한 환경이 많은 긴 복도의 경우, 3차원 가상 공간 모델에서 특징점을 고려하여 추가 사용자 정보 취득 포즈를 안내할 수 있다.

[0176] 도 12에서 추가 사용자 정보 취득 포즈는 특징점 1, 2, 3에 대해 순차적으로 영상을 획득하는 포즈 이거나, 특징점 1, 2, 3 중 어느 하나에 대한 포즈일 수 있다.

[0177] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0178]

[0179] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0180]

[0181] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도

록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0182]

[0183] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

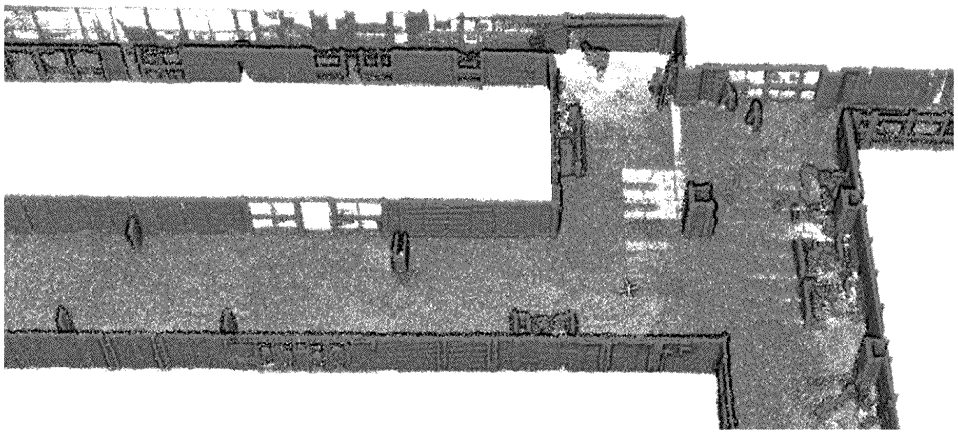
[0184]

[0185] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

[0186]

도면

도면1



도면2



(a)



(b)

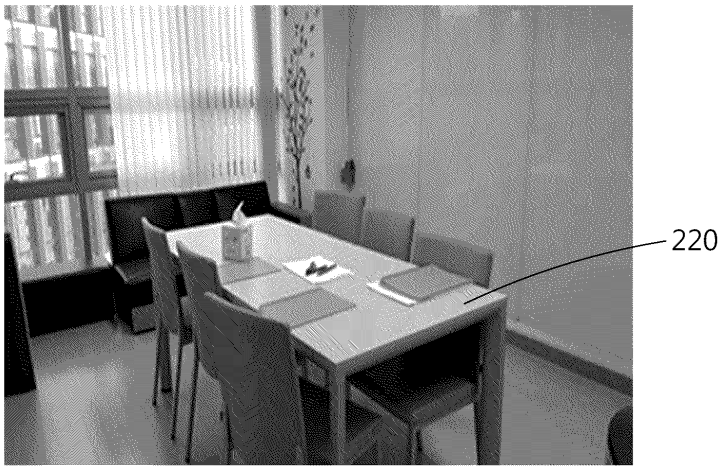


(c)

도면3

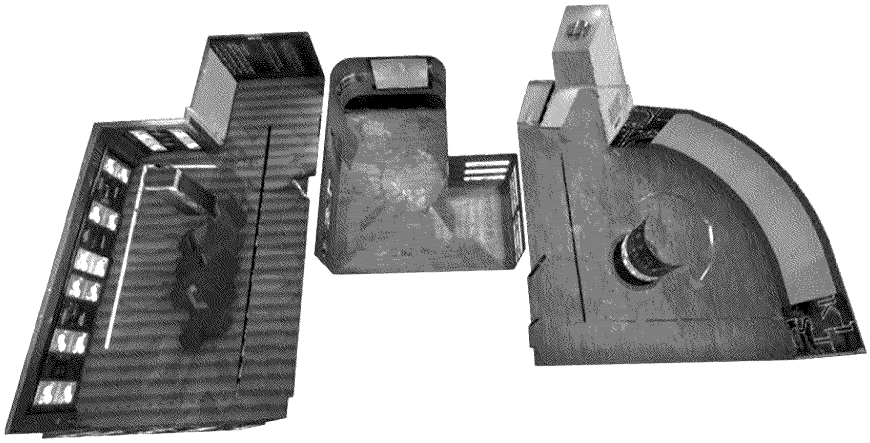


(a)

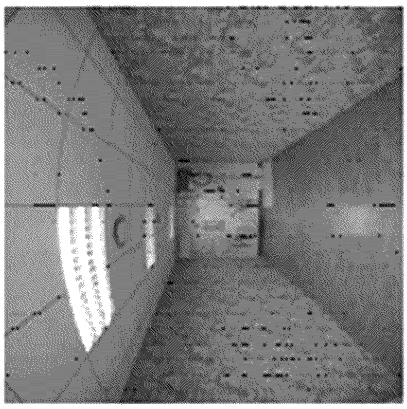


(b)

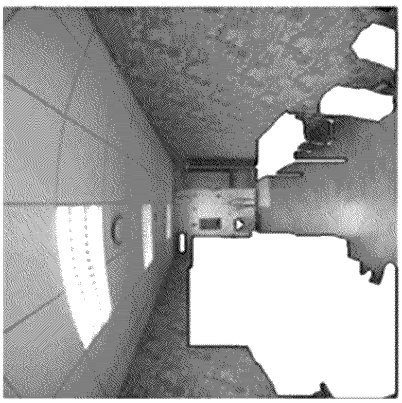
도면4



도면5



(c)

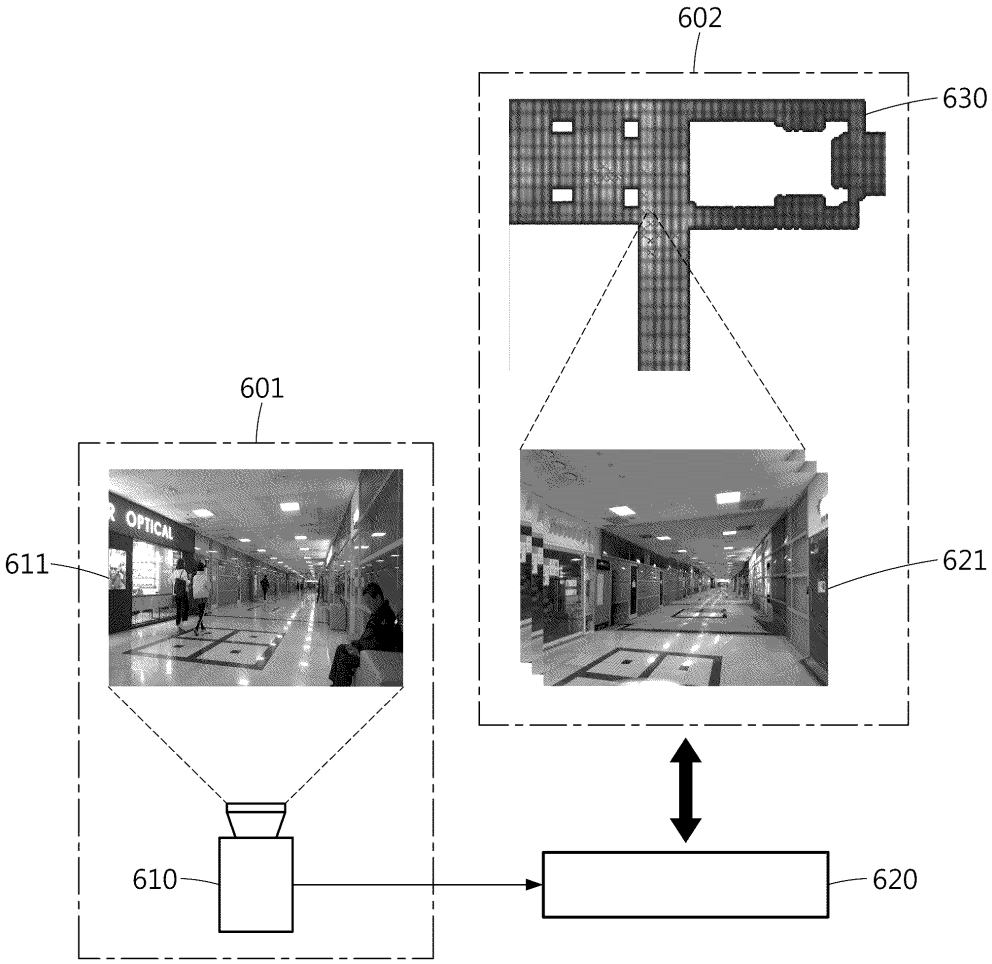


(b)

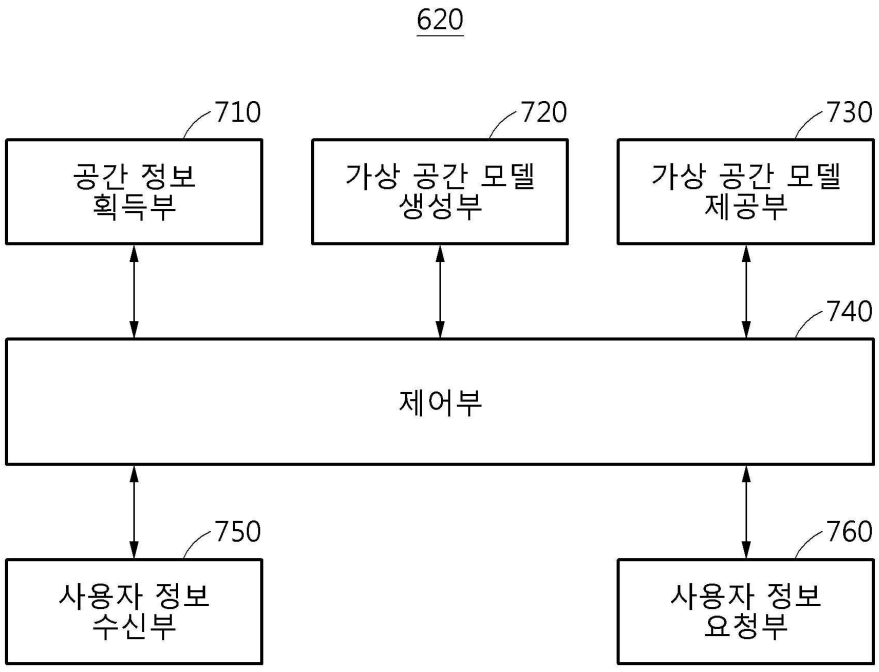


(a)

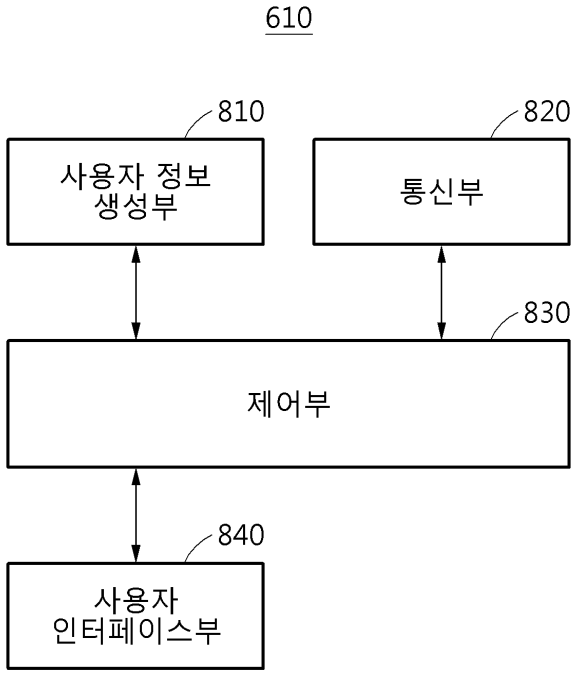
도면6


602
630
601
611
621
610
620

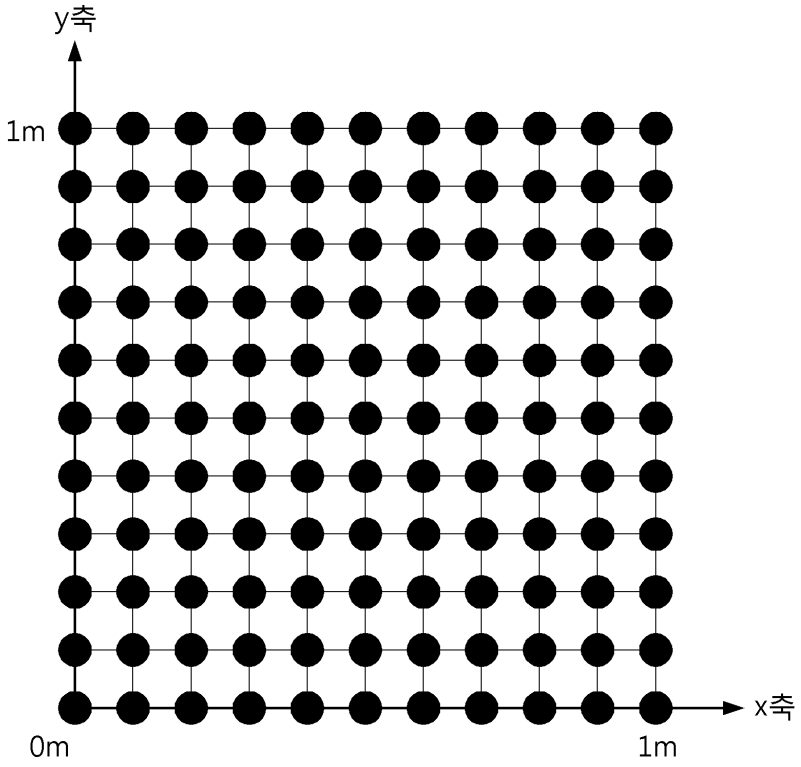
도면7


620
710
공간 정보 획득부
720
가상 공간 모델 생성부
730
가상 공간 모델 제공부
740
제어부
750
사용자 정보 수신부
760
사용자 정보 요청부

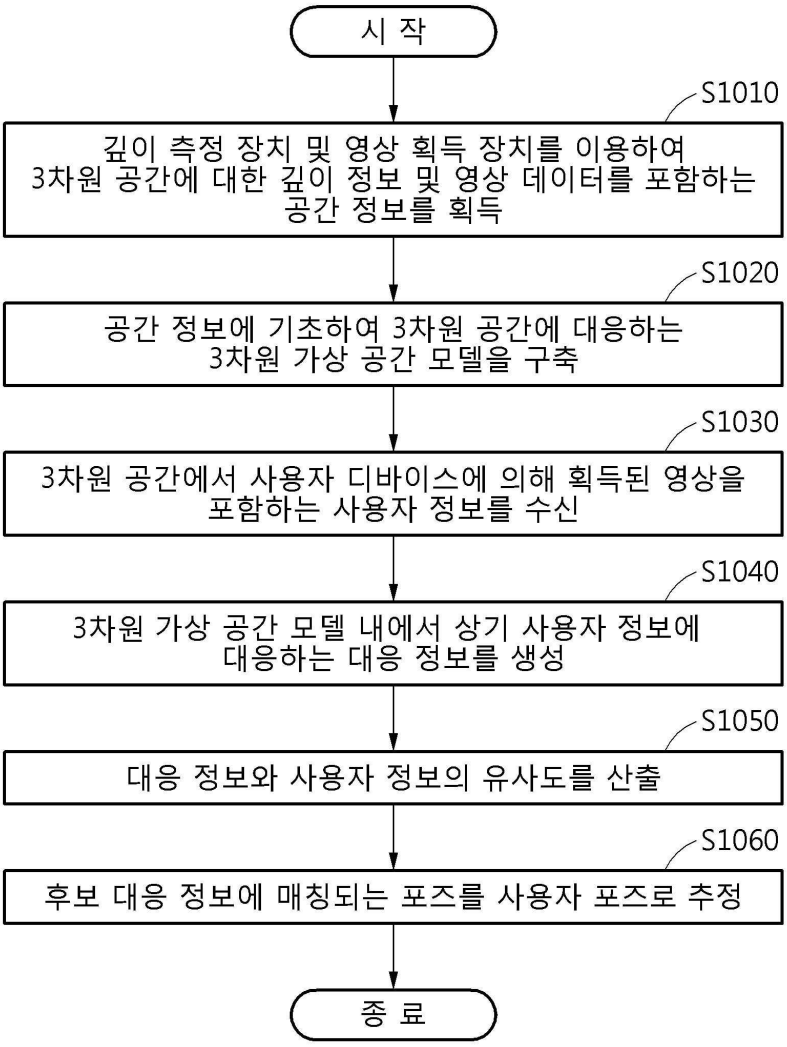
도면8



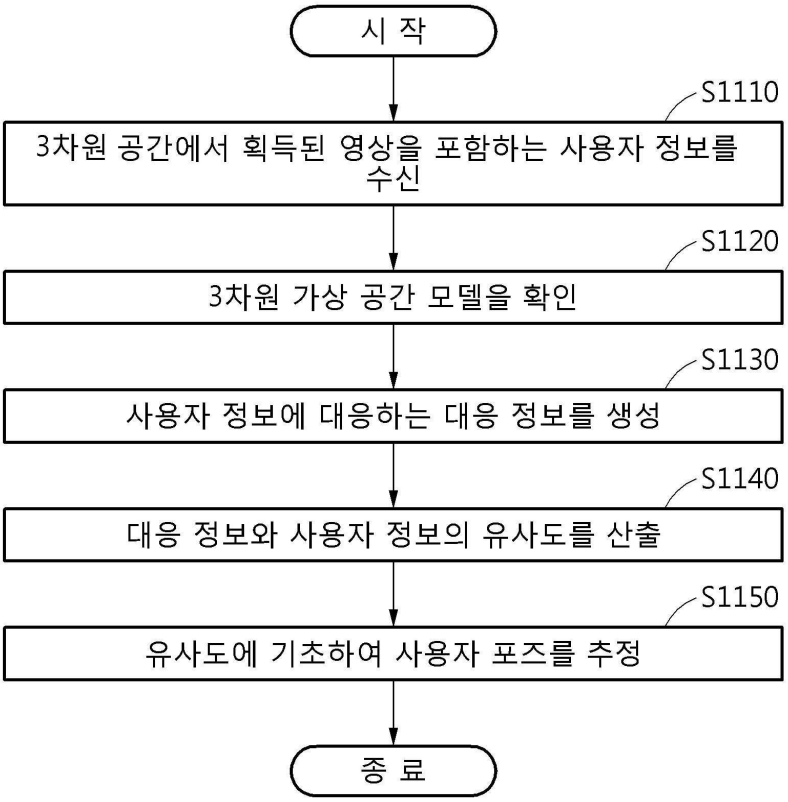
도면9



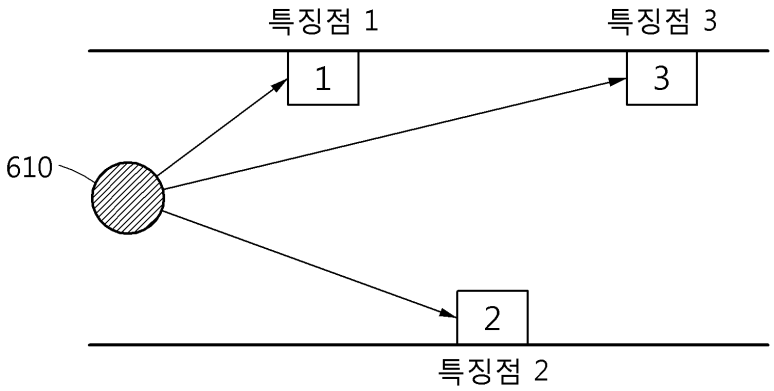
도면10


시 작
S1010
깊이 측정 장치 및 영상 획득 장치를 이용하여 3차원 공간에 대한 깊이 정보 및 영상 데이터를 포함하는 공간 정보를 획득
S1020
공간 정보에 기초하여 3차원 공간에 대응하는 3차원 가상 공간 모델을 구축
S1030
3차원 공간에서 사용자 디바이스에 의해 획득된 영상을 포함하는 사용자 정보를 수신
S1040
3차원 가상 공간 모델 내에서 상기 사용자 정보에 대응하는 대응 정보를 생성
S1050
대응 정보와 사용자 정보의 유사도를 산출
S1060
후보 대응 정보에 매칭되는 포즈를 사용자 포즈로 추정
종 료

도면11


시 작
S1110
3차원 공간에서 획득된 영상을 포함하는 사용자 정보를 수신
S1120
3차원 가상 공간 모델을 확인
S1130
사용자 정보에 대응하는 대응 정보를 생성
S1140
대응 정보와 사용자 정보의 유사도를 산출
S1150
유사도에 기초하여 사용자 포즈를 추정
종 료

도면12


특징점 1
특징점 3
1
3
610
2
특징점 2