



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0138729
(43) 공개일자 2016년12월06일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) G06F 3/01 (2006.01)
(52) CPC특허분류
G06K 9/00335 (2013.01)
G06F 3/017 (2013.01)
(21) 출원번호 10-2015-0072935
(22) 출원일자 2015년05월26일
심사청구일자 2015년05월26일

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 110동 502호(창전동, 서강쌍용에가아파트)
강민주
서울특별시 서대문구 이화여대2다길 15, 봉하숙지하2호 (대현동)
(뒷면에 계속)
(74) 대리인
남정길

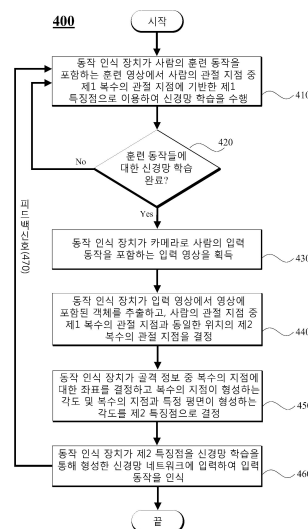
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 영상에서 동작 인식을 위한 특징점 추출 방법 및 골격 정보를 이용한 사용자 동작 인식 방법

(57) 요약

골격 정보를 이용한 사용자 동작 인식 방법은 동작 인식 장치가 사람의 복수의 훈련 동작을 포함하는 훈련 영상에서 사람의 관절 지점 중 제1 복수의 관절 지점에 기반한 제1 특징점으로 이용하여 신경망 학습을 수행하는 단계, 상기 동작 인식 장치가 카메라로 사람의 입력 동작을 포함하는 입력 영상을 획득하는 단계, 상기 동작 인식 장치가 상기 입력 영상에서 상기 사람의 관절 지점 중 상기 제1 복수의 관절 지점과 동일한 위치의 제2 복수의 관절 지점에 기반한 제2 특징점을 추출하는 단계 및 상기 동작 인식 장치가 상기 제2 특징점을 상기 신경망 학습을 통해 형성한 신경망 네트워크에 입력하여 상기 입력 동작에 대응하는 훈련 동작을 결정하는 단계를 포함한다.

대표도 - 도7



(52) CPC특허분류

G06K 9/00362 (2013.01)

(72) 발명자

김나영

서울특별시 성북구 오패산로3길 17, 104동 104
동101호 (하월곡동, 동신아파트)

류수경

서울특별시 송파구 송파대로32길 15, 105동 201호
(가락동, 가락금호아파트)

이지은

경기도 부천시 원미구 소향로 18, 2405동 2102호
(상동, 행복한마을 금호어울림)

이 발명을 지원한 국가연구개발사업

과제고유번호 2014R1A1A2056587

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 기초연구사업(학술진흥)-일반연구자지원사업-기본연구

연구과제명 미래 스마트 미디어 이미지/비디오 서비스를 위한 정보 부호화 기술

기 여 율 1/1

주관기관 이화여자대학교 산학협력단

연구기간 2014.11.01 ~ 2017.04.30

명세서

청구범위

청구항 1

동작 인식 장치가 카메라로 깊이 정보를 갖는 영상을 획득하는 단계;

상기 동작 인식 장치가 상기 영상에 포함된 객체를 추출하고, 상기 객체의 골격 정보를 추출하는 단계;

상기 동작 인식 장치가 상기 골격 정보 중 복수의 지점에 대한 좌표를 결정하는 단계; 및

상기 동작 인식 장치가 상기 복수의 지점이 형성하는 각도 및 상기 복수의 지점과 특정 평면이 형성하는 각도를 특징점으로 결정하는 단계를 포함하는 영상에서 동작 인식을 위한 특징점 추출 방법.

청구항 2

제1항에 있어서,

상기 객체는 사람이고, 상기 골격 정보는 사람의 관절 및 주요뼈에 해당하는 복수의 지점을 포함하는 영상에서 동작 인식을 위한 특징점 추출 방법.

청구항 3

제1항에 있어서,

상기 특징점은 상기 복수의 지점 중 3개의 지점이 형성하는 2개의 직선 사이의 각도 및 상기 3개의 지점 중 2개의 지점이 형성하는 직선과 특정 평면과의 각도 중 적어도 하나인 영상에서 동작 인식을 위한 특징점 추출 방법.

청구항 4

제1항에 있어서,

상기 복수의 지점 중 한쪽 팔의 어깨 지점, 팔꿈치 지점 및 손목 지점에 대해

상기 특징점은 상기 어깨 지점과 상기 팔꿈치 지점이 형성하는 제1 직선과 상기 팔꿈치 지점과 상기 손목 지점이 형성하는 제2 직선 사이의 각도, 상기 제1 직선과 제1 평면이 형성하는 각도 및 상기 제1 직선과 제2 평면이 형성하는 각도 중 적어도 하나인 영상에서 동작 인식을 위한 특징점 추출 방법.

청구항 5

제1항에 있어서,

상기 복수의 지점 중 한쪽 다리의 엉덩이 지점, 무릎 지점 및 발목 지점에 대해

상기 특징점은 상기 엉덩이 지점과 상기 무릎 지점이 형성하는 제1 직선과 상기 무릎 지점과 상기 발목 지점이 형성하는 제2 직선 사이의 각도, 상기 제1 직선과 제1 평면이 형성하는 각도 및 상기 제1 직선과 제2 평면이 형성하는 각도 중 적어도 하나인 영상에서 동작 인식을 위한 특징점 추출 방법.

청구항 6

동작 인식 장치가 사람의 복수의 훈련 동작을 포함하는 훈련 영상에서 사람의 관절 지점 중 제1 복수의 관절 지점에 기반한 제1 특징점으로 이용하여 신경망 학습을 수행하는 단계;

상기 동작 인식 장치가 카메라로 사람의 입력 동작을 포함하는 입력 영상을 획득하는 단계;

상기 동작 인식 장치가 상기 입력 영상에서 상기 사람의 관절 지점 중 상기 제1 복수의 관절 지점과 동일한 위치의 제2 복수의 관절 지점에 기반한 제2 특징점을 추출하는 단계; 및

상기 동작 인식 장치가 상기 제2 특징점을 상기 신경망 학습을 통해 형성한 신경망 네트워크에 입력하여 상기

입력 동작에 대응하는 훈련 동작을 결정하는 단계를 포함하는 골격 정보를 이용한 사용자 동작 인식 방법.

청구항 7

제6항에 있어서,

상기 동작 인식 장치는 상기 훈련 영상 및 상기 입력 영상 각각에서 상기 사람을 추출하고, 추출한 사람 영상의 골격화를 수행하고, 골격화로 생성된 객체를 사람의 골격 정보를 기준으로 사람의 관절 부위 및 주요뼈 부위에 해당하는 복수의 관절 지점을 결정하는 골격 정보를 이용한 사용자 동작 인식 방법.

청구항 8

제6항에 있어서,

상기 수행하는 단계는

상기 동작 인식 장치가 상기 제1 복수의 지점에 대한 좌표를 결정하는 단계;

상기 동작 인식 장치가 상기 제1 복수의 지점이 형성하는 각도 및 상기 복수의 지점과 특정 평면이 형성하는 각도를 상기 제1 특징점으로 결정하는 단계; 및

상기 제1 특징점을 기준으로 신경망 학습을 수행하여 신경망 네트워크를 마련하는 단계를 포함하되,

상기 수행하는 단계는 상기 복수의 훈련 동작 각각에 대해 반복적으로 수행되는 골격 정보를 이용한 사용자 동작 인식 방법.

청구항 9

제6항에 있어서,

상기 추출하는 단계는

상기 동작 인식 장치가 상기 제2 복수의 지점에 대한 좌표를 결정하는 단계; 및

상기 동작 인식 장치가 상기 제2 복수의 지점이 형성하는 각도 및 상기 복수의 지점과 특정 평면이 형성하는 각도를 상기 제2 특징점으로 결정하는 단계를 포함하는 골격 정보를 이용한 사용자 동작 인식 방법.

청구항 10

제6항에 있어서,

상기 제1 복수의 지점은 i)왼쪽 팔의 어깨 지점, 왼쪽 팔의 팔꿈치 지점 및 왼쪽 팔의 손목 지점, ii) 오른쪽 팔의 어깨 지점, 오른쪽 팔의 팔꿈치 지점 및 오른쪽 팔의 손목 지점, iii) 왼쪽 다리의 엉덩이 지점, 왼쪽 다리의 무릎 지점 및 왼쪽 다리의 발목 지점 및 iv) 오른쪽 다리의 엉덩이 지점, 오른쪽 다리의 무릎 지점 및 오른쪽 다리의 발목 지점 중 적어도 하나인 골격 정보를 이용한 사용자 동작 인식 방법.

청구항 11

제6항에 있어서,

상기 특징점은 어깨 지점과 팔꿈치 지점이 형성하는 제1 직선과 상기 팔꿈치 지점과 손목 지점이 형성하는 제2 직선 사이의 각도, 상기 제1 직선과 제1 평면이 형성하는 각도 및 상기 제1 직선과 제2 평면이 형성하는 각도 중 적어도 하나인 골격 정보를 이용한 사용자 동작 인식 방법.

청구항 12

제6항에 있어서,

상기 특징점은 엉덩이 지점과 무릎 지점이 형성하는 제1 직선과 상기 무릎 지점과 발목 지점이 형성하는 제2 직선 사이의 각도, 상기 제1 직선과 제1 평면이 형성하는 각도 및 상기 제1 직선과 제2 평면이 형성하는 각도 중 적어도 하나인 골격 정보를 이용한 사용자 동작 인식 방법.

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 영상에 사용자 동작을 인식하기 위한 기술에 관한 것이다.

배경 기술

[0002] 모션 인식 기술은 사용자의 신체 움직임을 인식하여 컴퓨터와 상호 작용 하는 기술이다. 사람의 움직임을 컴퓨터 장치에 대한 인터페이싱 입력으로 사용할 수 있다.

[0003] 모션 인식을 적용한 대표적인 예는 자이로 센서를 기반으로 한 스마트폰, 컨트롤러 센서를 기반으로 한 게임 콘솔 및 카메라 센서 기반의 스마트폰 등이 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 미국공개특허 US2010-0199228호

(특허문헌 0002) 미국공개특허 US2012-0157198호

발명의 내용

해결하려는 과제

[0005] MS의 Kinect는 영상에 포함된 사람의 골격 정보를 추출하여 동작을 인식한다. 다만 Kinect는 원거리에서 사람의 골격 정보를 추적하므로 정밀한 인식을 제공하지 못하는 한계가 있다.

[0006] 이하 설명하는 기술은 영상에 포함된 사람의 골격 정보 중 동작을 정확하게 인식할 수 있는 특징점을 이용하는 동작 인식 방법을 제공하고자 한다.

과제의 해결 수단

[0007] 영상에서 동작 인식을 위한 특징점 추출 방법은 동작 인식 장치가 카메라로 깊이 정보를 갖는 영상을 획득하는 단계, 동작 인식 장치가 영상에 포함된 객체를 추출하고, 객체의 골격 정보를 추출하는 단계, 동작 인식 장치가 골격 정보 중 복수의 지점에 대한 좌표를 결정하는 단계 및 동작 인식 장치가 복수의 지점이 형성하는 각도 및 복수의 지점과 특정 평면이 형성하는 각도를 특징점으로 결정하는 단계를 포함한다. 객체는 사람이고, 골격 정보는 사람의 관절 및 주요뼈에 해당하는 복수의 지점을 포함할 수 있다.

[0008] 골격 정보를 이용한 사용자 동작 인식 방법은 동작 인식 장치가 사람의 복수의 훈련 동작을 포함하는 훈련 영상에서 사람의 관절 지점 중 제1 복수의 관절 지점에 기반한 제1 특징점으로 이용하여 신경망 학습을 수행하는 단계, 동작 인식 장치가 카메라로 사람의 입력 동작을 포함하는 입력 영상을 획득하는 단계, 동작 인식 장치가 입력 영상에서 사람의 관절 지점 중 제1 복수의 관절 지점과 동일한 위치의 제2 복수의 관절 지점에 기반한 제2 특징점을 추출하는 단계 및 동작 인식 장치가 제2 특징점을 신경망 학습을 통해 형성한 신경망 네트워크에 입력하여 입력 동작에 대응하는 훈련 동작을 결정하는 단계를 포함한다.

[0009] 특징점은 복수의 지점 중 3개의 지점이 형성하는 2개의 직선 사이의 각도 및 3개의 지점 중 2개의 지점이 형성하는 직선과 특정 평면과의 각도 중 적어도 하나일 수 있다.

발명의 효과

[0010] 이하 설명하는 기술은 다른 장비 없이 영상을 통해 사람의 골격 정보를 추적하여 비교적 손쉽게 동작을 인식할 수 있다. 또한 이하 설명하는 기술은 사람의 골격 정보 중 동작을 명확하게 인식할 수 있는 특징점을 사용하고, 신경망 기계학습을 통해 사전에 마련된 기준으로 입력 동작을 명확하게 구분한다.

도면의 간단한 설명

[0011] 도 1은 골격 정보에 기반하여 동작을 인식하는 동작 인식 장치에 대한 구성을 도시한 블록도의 예이다.

도 2는 사람의 골격 정보에 대한 예이다.

도 3은 사람을 포함하는 깊이 영상에서 골격 정보를 도식한 예이다.

도 4는 골격 정보 중 일부를 이용한 특징점에 대한 예이다.

도 5는 골격 정보 중 일부를 이용한 다른 특징점에 대한 예이다.

도 6은 영상에서 동작 인식을 위한 특징점 추출 방법에 대한 순서도의 예이다.

도 7은 골격 정보를 이용한 사용자 동작 인식 방법에 대한 순서도의 예이다.

발명을 실시하기 위한 구체적인 내용

- [0012] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0013] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0014] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 실시된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.
- [0015] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다. 따라서, 본 명세서를 통해 설명되는 각 구성부들의 존재 여부는 기능적으로 해석되어야 할 것이며, 이러한 이유로 이하 설명하는 골격 정보를 이용한 사용자 동작 인식 방법의 구성은 이하 설명하는 기술의 목적을 달성할 수 있는 한도 내에서 대응하는 도면과는 상이해질 수 있음을 명확히 밝혀둔다.
- [0016] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0017] 도 1은 골격 정보에 기반하여 동작을 인식하는 동작 인식 장치(100)에 대한 구성을 도식한 블록도의 예이다. 도 1에서는 사람이 일정한 동작을 취하고 있고, 동작 인식 장치(100)가 제1 카메라(111) 및 제2 카메라(112)를 통해 사용자를 포함한 영상을 획득하고 있다. 동작 인식 장치(100)는 두 개의 카메라로 깊이 정보가 있는 영상을 획득한다. 두 개의 카메라는 RGB 카메라일 수도 있고, 적외선 카메라일 수도 있다. 깊이 정보 영상을 생성하는 과정은 널리 알려진 것이므로 자세한 설명은 생략한다.
- [0018] 도 1에서는 동작 인식 장치(100)의 구성으로 두 개의 카메라(111 및 112)만을 도시하였다. 후술하겠지만 동작 인식 장치(100)는 사람(객체)을 추출하고, 사람의 골격 정보를 기반으로 특징점을 추출한다. 따라서 동작 인식 장치(100)는 획득한 영상을 저장하는 저장 매체, 영상 처리를 위한 중앙 처리 장치 등과 같은 장치를 포함한다. 동작 인식 장치(100)는 일종의 컴퓨터 장치라고 할 수 있다. 동작인식 장치(100)는 게임 콘솔, 스마트 TV, PC, 스마트 도어락, 수화 번역 장치 등과 같은 장치일 수 있다.
- [0019] 동작 인식 장치(100)는 일정한 동작을 취하는 사람이 포함된 영상을 획득하고, 이미지 처리 기법을 통해 획득한 영상에서 사람을 추출한다. 사람 추출은 영상에서 배경을 제외한 특정 객체를 추출하는 기법을 이용한다.

- [0020] 깊이 정보는 인간의 양안을 모델로 한두 개 이상의 카메라를 활용하여 얻은 두 영상/동영상 사이의 변위정보(disparity information)을 3차원 정보로 변환하여 얻거나 적외선 센서를 이용하여 원거리에서 3차원 깊이를 측정할 수 있다. 깊이 정보는 [0,255] 사이의 픽셀 값을 갖는 흑백이미지로 표현한다. 피사체가 카메라의 기준점에서 가까울수록 밝은 영역으로 표현되고, 멀수록 어둡게 표현된다. 깊이 영상에서 사람은 배경보다 앞에 위치하여 배경보다 밝게 표현되는 특성이 있다. 이러한 특성을 반영하여 컬러 정보와 깊이 정보의 혼합을 통해 영상으로부터 사람을 구분할 수 있다.
- [0021] 동작 인식 장치(100)는 영상에서 사람을 추출하고, 추가적인 이미지 처리 기법을 사용하여 사람의 골격(skeleton) 정보를 추출한다. 적외선 카메라로 얻은 깊이 정보 값은 전경 또는 배경 내에서 평탄한 값을 가지며, 가까이 있는 물체 중에서도 사람의 골격을 가진 객체를 사람으로 인식한다. 도 2는 사람의 골격 정보에 대한 예이다.
- [0022] 골격 정보는 사람의 관절 부위 및 몸의 움직임에 관여하는 주요한 뼈의 위치를 포함한다. 팔에서 골격 정보는 손 지점, 손목 지점, 팔꿈치 지점 및 어깨 지점을 포함하고, 다리에서 골격 정보는 발 지점, 발목 지점, 무릎 지점 및 엉덩이 지점을 포함한다. 나머지 골격 정보는 머리 지점, 어깨 중심 지점, 척추 지점, 엉덩이 중심 지점을 포함한다. 도 2는 모두 20개의 위치를 골격 정보로 도시한다.
- [0023] 동작 인식 장치(100)는 일정한 지점을 기준으로 골격 정보에 포함되는 각 부위의 위치를 3차원 좌표로 결정할 수 있다.
- [0024] 도 3은 사람을 포함하는 깊이 영상에서 골격 정보를 도시한 예이다.
- [0025] 도 3(a)는 사람이 나타난 깊이 영상에서 사람에 해당하는 객체를 구분하고, 객체에 대한 골격화를 수행한 예에 해당한다. 도 3(a)에서 사람 내부에 직선으로 표현된 부분이 골격화의 결과물에 해당한다.
- [0026] 골격화는 이진영상에서 물체의 크기와 모양을 요약하는 선과 곡선의 집합을 만드는 과정이다. 골격을 정의하는 방법이 다양하므로 주어진 물체에서 서로 다른 모양의 골격이 있을 수 있다. 다만 동작 인식 장치(100)가 인식하는 영상은 사람에 대한 영상이므로, 사람의 해부학적인 구조는 유사하여 골격화에 따른 결과물로 유사한다. 골격화는 침식연산을 이용해 수행한다. 골격화는 서로 다른 방향성이 여러 개 있는 침식마스크를 이용하여 영상 내의 물체를 서서히 깎아 침식을 반복하여 객체의 하부 구조를 나타낸다. 사람의 구조적인 특징에 대한 정보를 기반으로 골격화 결과물인 직선의 길이, 직선의 위치 및 직선의 휘어짐 등을 기준으로 전술한 골격 정보를 결정할 수 있다. 도 3(a)에서는 골격 정보를 직선에서 작은 점으로 표시하였다. 도 3(a)에서 어깨 지점, 팔꿈치 지점 및 손목 지점은 별도로 표시하였다.
- [0027] 도 3(b)는 깊이 정보 영상에서 골격 정보의 위치만을 십자 모양으로 도시한 예이다. 동작 인식 장치(100)는 획득한 영상에서 사람의 골격 정보를 결정하고, 이후 사람의 동작에 따라 각 골격 정보의 위치를 추적할 수 있다.
- [0028] 동작 인식 장치(100)는 골격 정보의 위치를 기반으로 특징점을 추출한다. 해당 특징점이 사람의 동작을 인식하는 기준이 된다.
- [0029] 동작 인식 장치(100)는 앞서 설명한 20개의 골격 정보를 조합하여 고차원의 특징점을 생성할 수 있다. 예컨대, $R=\{p_1, p_2, \dots, p_n\}$ 이라는 골격 정보가 있을 때 $z_1=p_1+p_2$, $z_2=p_1*p_3$ 과 같은 조합을 통해 특징점 $Z=\{z_1, z_2, \dots, z_m\}$ 를 생성할 수 있다. 결국 복수의 골격 정보를 이용하여 다양한 특징점을 생성할 수 있다.
- [0030] 도 4 및 도 5를 통해 몇 가지 특징점을 추출하는 과정을 설명한다. 도 4는 골격 정보 중 일부를 이용한 특징점에 대한 예이다. 도 4(a)는 일정한 동작을 취한 사람 영상을 도시하고, 도 4(b)는 골격 정보의 위치 중 일부를 이용하여 특징점을 결정하는 예를 도시한다.
- [0031] 먼저 동작 인식 장치(100)는 골격 정보에 대한 3차원 좌표값을 결정한 상태라고 전제한다. 도 4에서 골격 정보는 왼쪽 어깨 지점(P1), 왼쪽 팔꿈치 지점(P2) 및 왼쪽 손목 지점(P3)을 사용한다. 각 골격 정보에 대한 3차원 좌표는 $P1(x_1, y_1, z_1)$, $P2(x_2, y_2, z_2)$ 및 $P3(x_3, y_3, z_3)$ 이다. 아래와 같이 3가지 특징점을 구할 수 있다. 한편 x, y, z 축은 특정한 기준에 따라 설정될 수 있다.
- [0032] (1) 두 개의 직선 사이의 각도(u)
- [0033] P1과 P2는 하나의 직선(제1 직선)을 형성하고, P2와 P3도 하나의 직선(제2 직선)을 형성한다. 따라서 제1 직선과 제2 직선 사이의 각도(u)를 특징점으로 삼을 수 있다. 제1 직선과 제2 직선 사이의 각도(u)는 아래의 수학적 식 1을 통해 연산할 수 있다.

수학식 1

$$u = 180 - \left\{ \text{atan} \left(\frac{y_3 - y_2}{x_3 - x_2} \right) - \text{atan} \left(\frac{y_1 - y_2}{x_1 - x_2} \right) \right\} \times \frac{180}{\pi}$$

(2) 하나의 직선과 제1 평면 사이의 각도(φ)

P1과 P2가 형성하는 제1 직선과 제1 평면 사이의 각도(φ)를 특징점으로 삼을 수 있다. 제1 평면은 다양한 기준의 평면을 사용할 수 있다. 예컨대, 도 4에서는 z축에 평행한 축을 제1 평면으로 사용하였다. 제1 직선과 제1 평면 사이의 각도(φ)는 아래의 수학식 2를 통해 연산할 수 있다.

수학식 2

$$\varphi = \text{atan} \left(\frac{z_2 - z_1}{x_2 - x_1} \right) \times \frac{180}{\pi}$$

또는 P2와 P3가 형성하는 제2 직선과 제1 평면 사이의 각도를 특징점으로 삼을 수도 있을 것이다.

(3) 하나의 직선과 제2 평면 사이의 각도(θ)

P1과 P2가 형성하는 제1 직선과 제2 평면 사이의 각도(θ)를 특징점으로 삼을 수 있다. 제2 평면은 다양한 기준의 평면을 사용할 수 있다. 예컨대, 도 4에서는 y축에 평행한 축을 제2 평면으로 사용하였다. 제1 직선과 제1 평면 사이의 각도(θ)는 아래의 수학식 3을 통해 연산할 수 있다.

수학식 3

$$\theta = \text{atan} \left(\frac{y_2 - y_1}{x_2 - x_1} \right) \times \frac{180}{\pi}$$

도 4는 사람의 왼팔에 대해서 3가지 특징점을 도시한 것이다. 동일한 방법으로 사람의 오른팔에 대해서도 3가지 특징점을 추출할 수 있다.

도 5는 골격 정보 중 일부를 이용한 다른 특징점에 대한 예이다. 도 5는 다리의 골격 정보를 이용한 예를 도시한다. 도 5(a)는 일정한 동작을 취한 사람 영상을 도시하고, 도 5(b)는 골격 정보의 위치 중 일부를 이용하여 특징점을 결정하는 예를 도시한다.

도 5에서 골격 정보는 오른쪽 엉덩이 지점(Q1), 오른쪽 무릎 지점(Q2) 및 오른쪽 발목 지점(Q3)을 사용한다. 각 골격 정보에 대한 3차원 좌표는 Q1(x1, y1, z1), Q2(x2, y2, z2) 및 Q3(x3, y3, z3)이다.

도 4에 대한 설명과 동일한 방법으로 (1) Q1과 Q2가 형성하는 제1 직선과 Q2와 Q3가 형성하는 제2 직선 사이의 각도(u), (2) Q1과 Q2가 형성하는 제1 직선과 제1 평면인 z축에 평행한 평면 사이의 각도(φ) 및 (3) Q1과 Q2가 형성하는 제1 직선과 제2 평면인 y축에 평행한 평면 사이의 각도(θ)를 특징점으로 결정할 수 있다.

또한 왼쪽 발에 대해서도 동일한 방법으로 3개의 특징점을 추출할 수 있다.

도 4 및 도 5에서 설명한 바와 같이 왼쪽 팔, 오른쪽 팔, 왼쪽 다리 및 오른쪽 다리의 일부 골격 정보를 이용하여 각각 3개의 특징점을 추출한다면, 모두 12개의 특징점을 마련할 수 있다.

- [0048] 도 6은 영상에서 동작 인식을 위한 특징점 추출 방법(300)에 대한 순서도의 예이다. 전술한 특징점 추출을 정리하면 동작 인식 장치는 먼저 카메라로 깊이 정보를 갖는 영상을 획득한다(310). 동작 인식 장치는 영상에 포함된 특정 객체(사람)를 추출한다. 동작 인식 장치는 추출한 객체의 골격 정보를 추출한다(320). 동작 인식 장치는 골격 정보 중 특징점의 기준으로 사용하는 복수의 지점에 대한 3차원 위치를 결정한다(330). 마지막으로 동작 인식 장치는 3차원 위치를 결정한 복수의 지점 중 일부 또는 전부가 형성하는 직선 사이의 각도 및 직선과 특정 평면 사이의 각도 중 적어도 하나를 특징점으로 결정한다(340).
- [0049] 전술한 특징점 외에 다른 다양한 특징점이 사용될 수도 있다. 기본적으로 골격 정보 중 다른 지점(머리 지점, 어깨 중심 지점, 척추 지점 등)을 사용할 수도 있다. 어떤 동작을 인터페이스로 사용하느냐에 따라 특징점에 사용하는 골격 정보가 달라질 수 있다. 또한 사람의 동작이 몸 전체의 동작인지, 팔과 같이 일부의 동작인지에 따라 특징점이 달라질 수도 있다.
- [0050] 전술한 특징점(u)는 인접한 골격 정보가 형성하는 직선을 사용하였지만, 반드시 인접한 골격 정보 사이의 직선을 사용하지 않을 수도 있다. 나아가 각도를 결정하기 위한 기준이 되는 직선이나 평면을 다른 것을 사용할 수도 있다. 즉 매우 다양한 조합의 특징점 생성이 가능하다.
- [0051] 도 7은 골격 정보를 이용한 사용자 동작 인식 방법(400)에 대한 순서도의 예이다.
- [0052] 동작 인식 장치가 사용자 동작을 인식하기 위해서는 먼저 사용자 동작에 대한 참조 데이터가 필요하다. 다양한 방법을 사용할 수 있지만, 기본적으로 기계학습 중 신경망 학습을 기준으로 설명한다.
- [0053] 기계학습은 적용 어플리케이션과 데이터의 형태에 따라 감독 학습(supervised learning)과 무감독 학습(unsupervised learning)으로 구분할 수 있다. 무감독 학습의 목적은 학습을 위한 입력과 출력의 관계가 없어 입력 패턴의 유형을 파악하여 입력 간 공통적인 특성을 파악하는 것이다. 반면 감독 학습은 입력과 출력의 쌍이 미리 알려진 사전 정보로부터 입출력을 매핑하는 함수를 학습하는 과정이다. 즉 시스템은 입력 벡터 x 와 이에 대한 올바른 응답 y 를 학습 데이터로 제공받아 함수 f 를 학습하고, 학습하지 않은 새로운 데이터 x' 가 들어왔을 때 출력 y' 를 추론하는 학습을 학습한다. 따라서 사용자 동작 인식을 위해서는 감독 학습 기법을 이용한다.
- [0054] 기계학습의 일종인 신경망 학습은 감독학습 중 하나로 입력 데이터에 X 에 대해 올바른 정답 Y 를 학습하는 방식이다. 동작 인식 장치는 N 개의 동작에 해당하는 Label 1~Label N 모션을 각각 여러 번 입력 받은 후 각 동작에 대한 12개의 특징점(입력 X)들을 입력 후 학습하여 Label 1 ~ Label N 모션을 판별할 수 있는 신경망 네트워크를 생성한다. 신경망 학습의 결과로 노드의 구조 정보와 노드 간 연결 계수를 얻는다.
- [0055] 동작 인식 장치는 사람의 훈련 동작을 포함하는 훈련 영상을 이용하여 특징점을 추출한다. 학습을 위한 동작을 훈련 동작이라고 하고, 훈련 동작이 포함된 영상을 훈련 영상이라고 명명한다. 추출한 특징점은 전술한 바와 같이 골격 정보 중 복수의 관절 지점을 이용하여 생성한다. 학습 과정에 사용되는 복수의 관절 지점을 제1 복수의 관절 지점이라고 명명한다. 학습 과정에서 제1 복수의 관절 지점을 사용하여 추출한 특징점을 제1 특징점이라고 명명한다. 제1 특징점은 전술한 바와 같은 12개의 특징점일 수 있다. 동작 인식 장치는 복수의 훈련 동작을 제1 특징점을 기준으로 신경망 학습을 수행한다(410).
- [0056] 동작 인식 장치는 훈련 동작들에 대한 신경망 학습이 완료되지 않았다면(420의 No) 반복적으로 신경망 학습을 수행하고(410), 훈련 동작들에 대한 신경망 학습이 완료되었다면(420의 Yes) 동작 인식이 필요한 영상을 대기한다.
- [0057] 동작 인식 장치가 두 개의 카메라로 동작 인식이 필요한 입력 영상을 획득한다(430). 두 개의 카메라를 사용하는 이유는 깊이 정보를 갖는 영상을 생성하기 위함이다.
- [0058] 동작 인식 장치는 입력 영상에서 영상에 포함된 객체(사람)를 추출하고, 추출한 사람에서 골격 정보를 추출한다. 사람이라는 객체에서 복수의 관절 지점을 결정한다(440). 입력 영상에서 결정한 복수의 관절 지점을 제2 복수의 관절 지점이라고 명명한다. 제2 복수의 관절 지점은 학습 단계에서 사용한 제1 복수의 관절 지점과 동일한(대응되는) 위치여야 한다.
- [0059] 동작 인식 장치는 제2 복수의 관절 지점 중 복수의 지점을 사용하여 특징점을 결정한다(450). 입력 영상에 대한 특징점을 제2 특징점이라고 명명한다. 제2 특징점은 제1 특징점과 대응된다. 제1 특징점이 도 4 및 도 5에서 설명한 12개의 특징점이라면, 제2 특징점도 동일한 12개의 특징점이어야 한다.
- [0060] 동작 인식 장치는 최종적으로 제2 특징점을 신경망 네트워크에 입력하여 최종적으로 결정되는 동작을 입력 동작으로 인식한다. 예컨대, 학습 과정에서 N 개의 동작에 대한 신경망 네트워크를 생성했다면, 동작 인식 장치는 입

력 동작을 N개의 동작 중 하나로 판단한다.

[0061] 동작 인식 장치는 입력 영상에서 학습과정과 동일한 절차로 특징점 12개를 추출하고, 학습을 이용해 취득한 파라미터 세트(신경망 네트워크의 경우 노드의 구조와 노드 간 연결 계수)를 이용하여 학습한 신경망 네트워크에 입력하면 최종 출력 노드에서 Label1 ~ Label N 모션 중 1개로 판별 한다. 신경망 네트워크를 통해 모션을 판별하는 한편 새로운 동작을 인식하면서 얻어진 신경망 정보는 410으로 피드백 되어 신경망을 업데이트 한다(470).

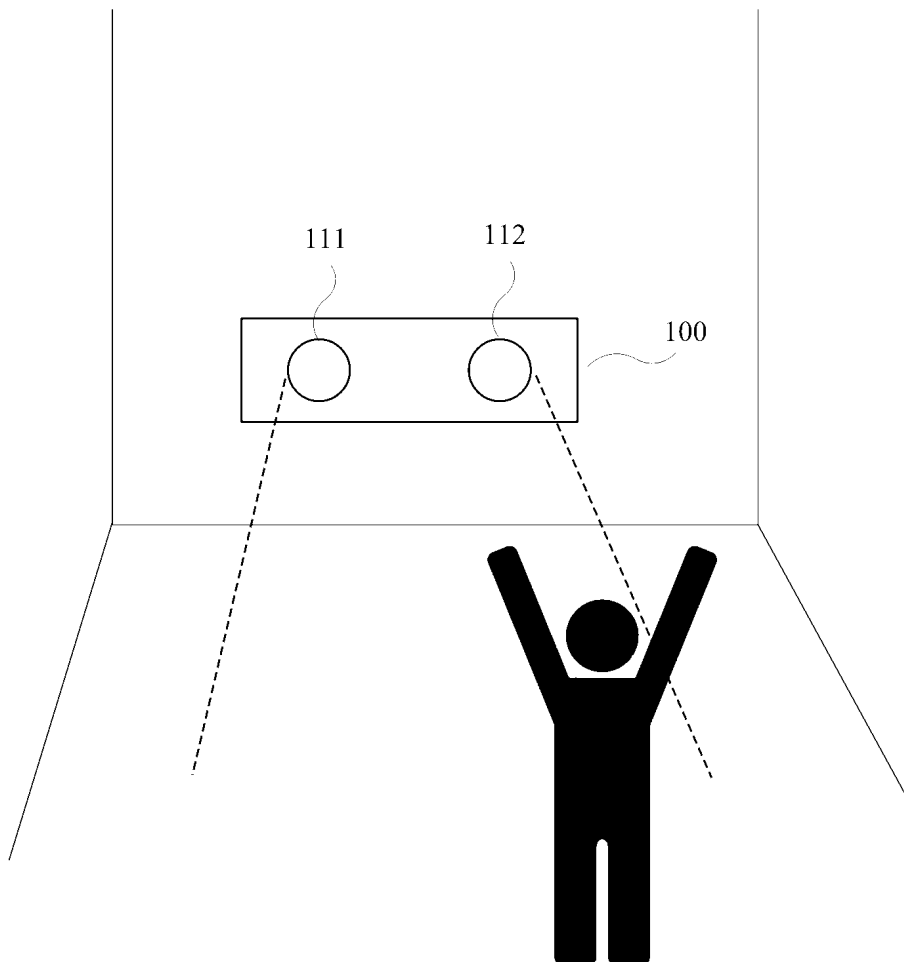
[0062] 본 실시예 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시예는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

부호의 설명

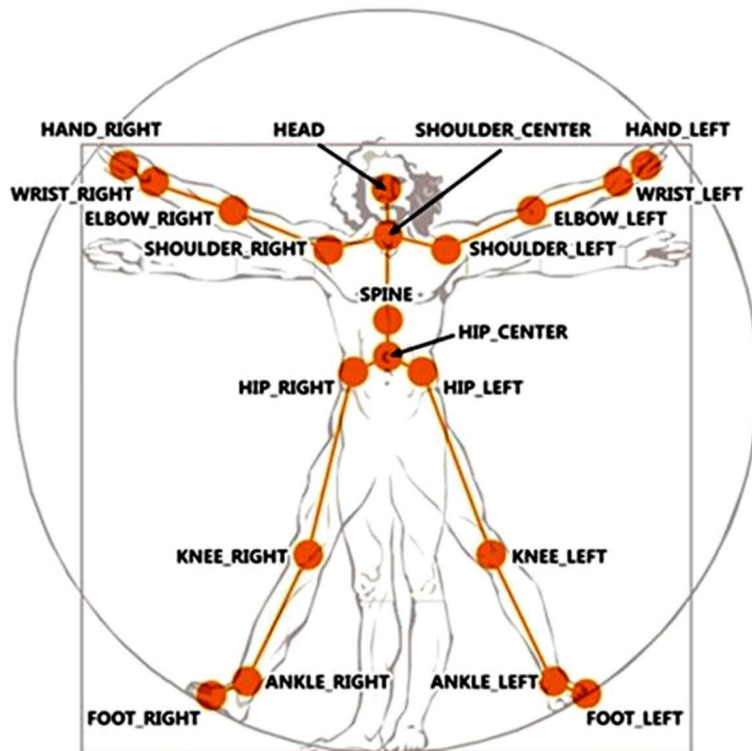
[0063] 100 : 동작 인식 장치 111 : 제1 카메라
112 : 제2 카메라

도면

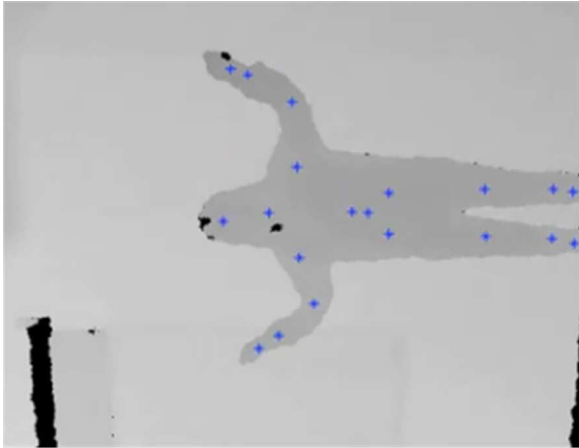
도면1



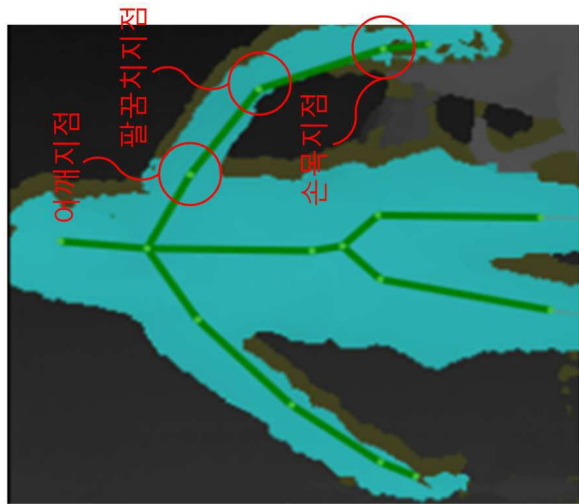
도면2



도면3

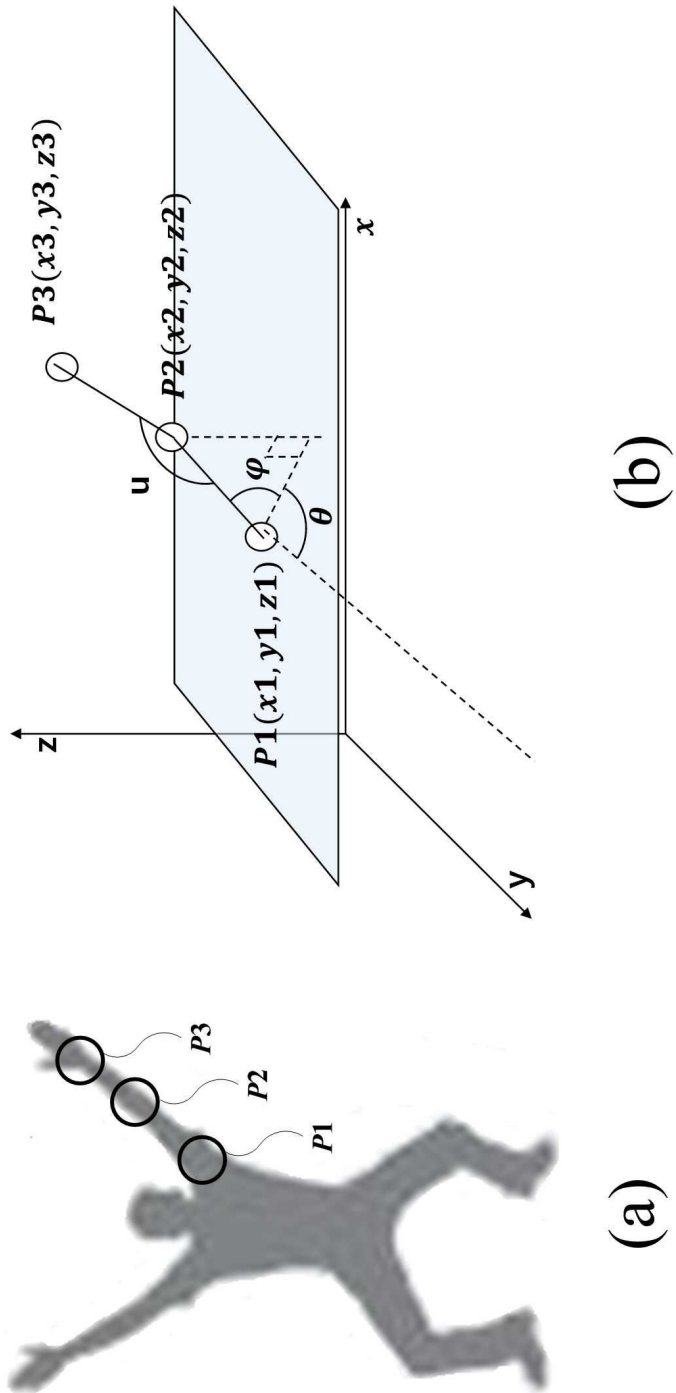


(b)

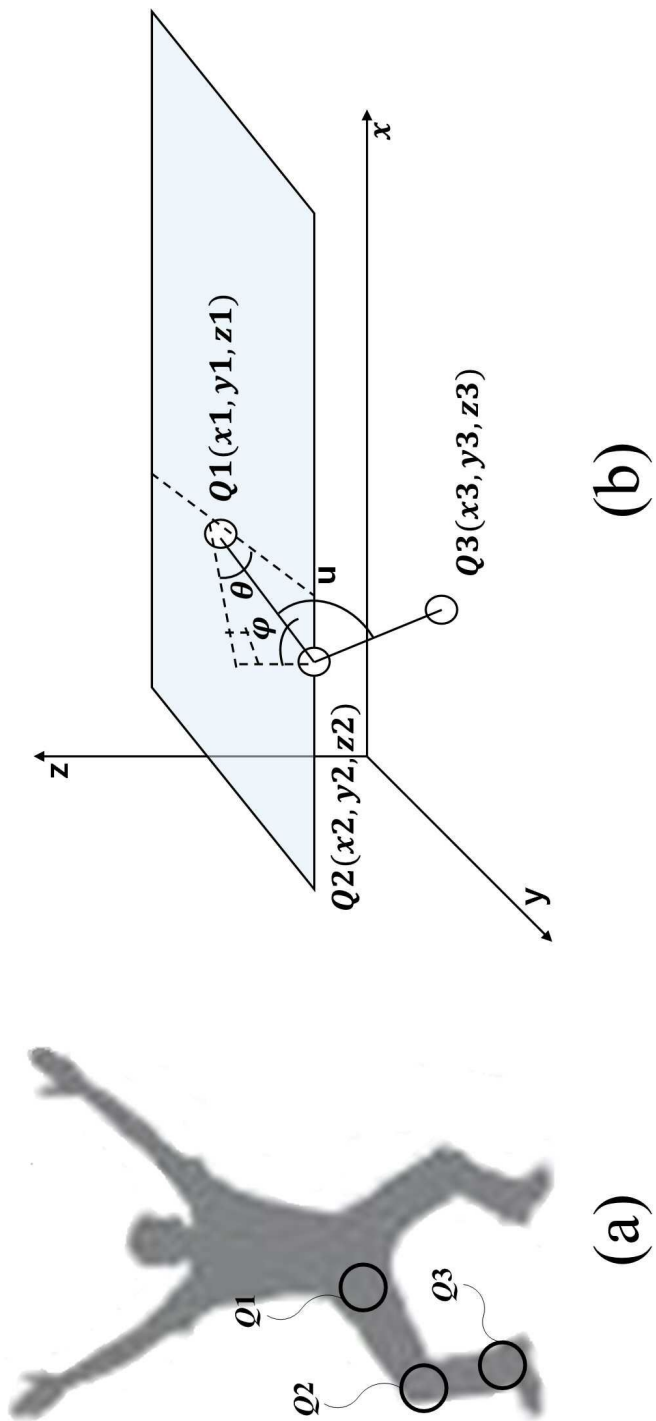


(a)

도면4

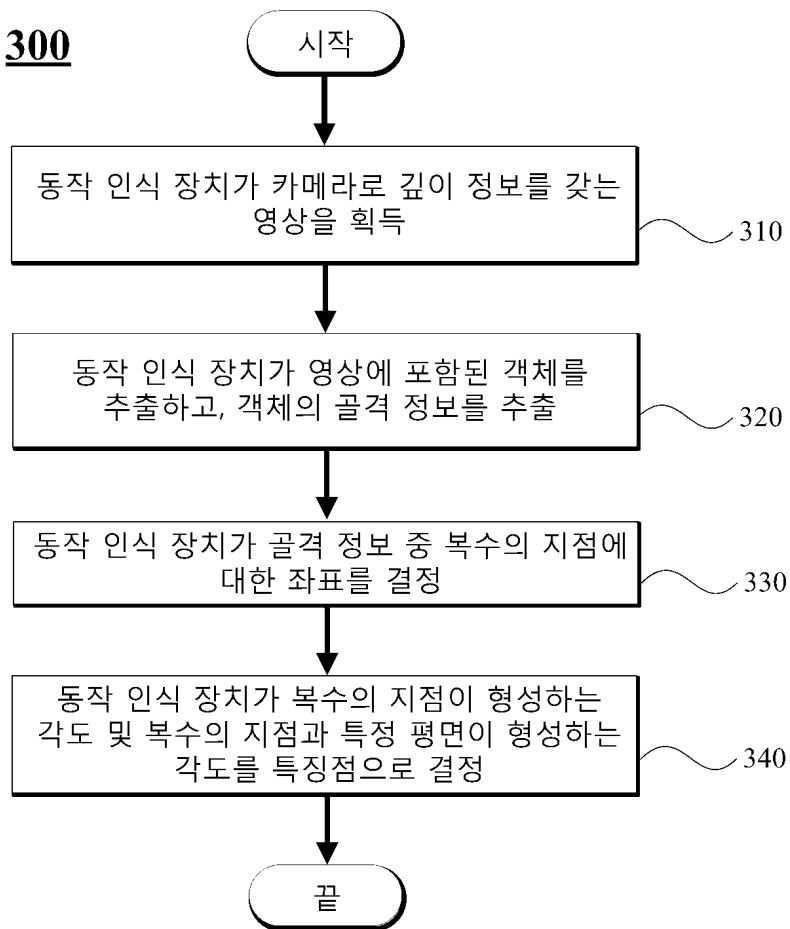


도면5



도면6

300



도면7

