



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0076398
(43) 공개일자 2022년06월08일

(51) 국제특허분류(Int. Cl.)
G06T 19/00 (2011.01) G06K 9/62 (2022.01)
G06N 3/08 (2006.01) G06N 5/04 (2006.01)
G06T 7/00 (2017.01) G06T 7/11 (2017.01)
(52) CPC특허분류
G06T 19/006 (2013.01)
G06K 9/6217 (2022.01)
(21) 출원번호 10-2021-0168994
(22) 출원일자 2021년11월30일
심사청구일자 2021년11월30일
(30) 우선권주장
1020200164494 2020년11월30일 대한민국(KR)

(71) 출원인
서울대학교산학협력단
서울특별시 관악구 관악로 1 (신림동)
(72) 발명자
이경한
서울특별시 관악구 관악로 1, 301동 1006호 (신림동, 서울대학교)
함성민
울산광역시 울주군 언양읍 반천강변길 51, 106동 1402호 (언양읍, 현대아파트)
(74) 대리인
특허법인엠에이피에스

전체 청구항 수 : 총 19 항

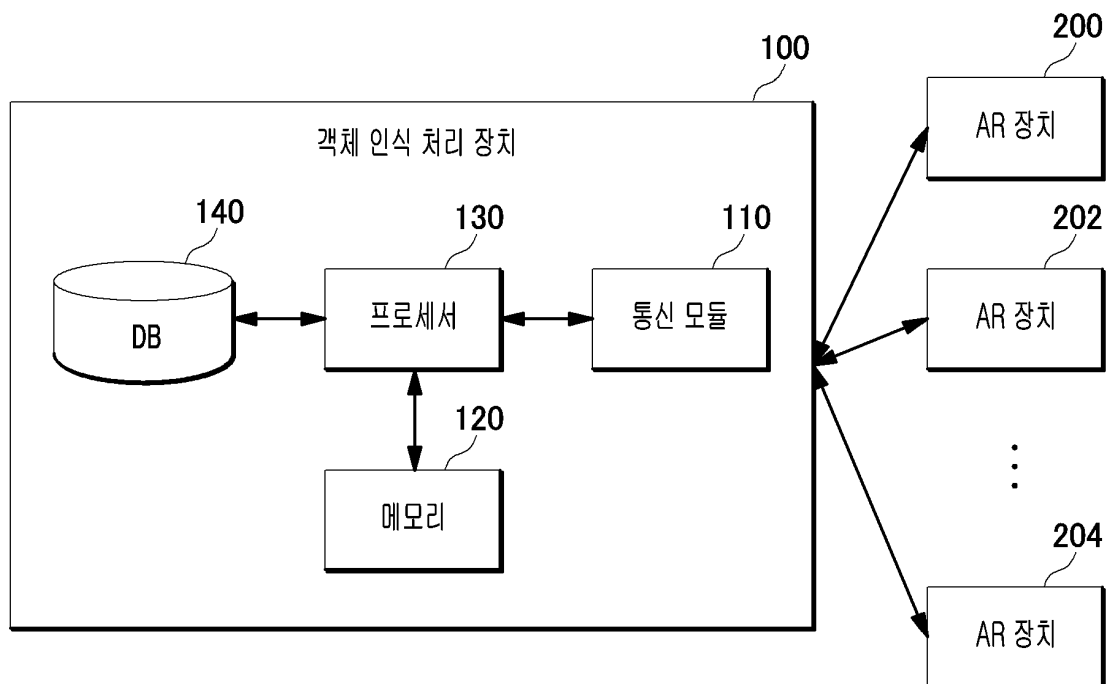
(54) 발명의 명칭 AR장치를 위한 객체 인식 처리 장치 및 방법

(57) 요약

본 발명의 일 실시예에 따른 AR(Augmented Reality) 장치를 위한 객체 인식 처리 장치는 통신모듈; 객체 인식 프로그램이 저장된 메모리; 및 상기 객체 인식 프로그램을 실행하는 프로세서를 포함한다. 이때, 객체 인식 프로그램은 적어도 하나 이상의 AR 장치로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해

(뒷면에 계속)

대표도 - 도1



당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 AR 정보 수집 모듈, 상기 AR 영상에서 각각의 객체를 구분하는 객체 탐지 모듈, 상기 객체 탐지 모듈을 통해 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출하는 객체 추적 및 데이터셋 수집 모듈, 상기 객체 추적 및 데이터셋 수집 모듈로부터 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하는 객체 학습 모듈 및 입력된 AR 영상으로부터 상기 객체 탐지 모듈을 통해 구분된 객체를 상기 객체 학습 모듈에 입력하여 객체의 식별 정보를 출력하는 객체 인지 모듈을 포함한다.

(52) CPC특허분류

G06K 9/6267 (2022.01)

G06N 3/08 (2013.01)

G06N 5/04 (2013.01)

G06T 7/001 (2013.01)

G06T 7/11 (2017.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103020
과제번호	2017-0-00667-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	디지털콘텐츠원천기술개발사업
연구과제명	이동 객체에 대한 증강현실 서비스 실현을 위한 정보-영상 정합기술 연구
기 여 율	1/1
과제수행기관명	울산과학기술원
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

AR(Augmented Reality) 장치를 위한 객체 인식 처리 장치에 있어서,

통신모듈;

객체 인식 프로그램이 저장된 메모리; 및

상기 객체 인식 프로그램을 실행하는 프로세서를 포함하되,

상기 객체 인식 프로그램은 적어도 하나 이상의 AR 장치로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 AR 정보 수집 모듈, 상기 AR 영상에서 각각의 객체를 구분하는 객체 탐지 모듈, 상기 객체 탐지 모듈을 통해 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출하는 객체 추적 및 데이터셋 수집 모듈, 상기 객체 추적 및 데이터셋 수집 모듈로부터 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하는 객체 학습 모듈 및 입력된 AR 영상으로부터 상기 객체 탐지 모듈을 통해 구분된 객체를 상기 객체 학습 모듈에 입력하여 객체의 식별 정보를 출력하는 객체 인지 모듈을 포함하는 것인, 객체 인식 처리 장치.

청구항 2

제1항에 있어서,

상기 객체 탐지 모듈은 상기 AR 영상에서 각 객체별로 바운딩 박스를 생성하고, 각 객체별 바운딩 박스에 상기 객체 카테고리 정보를 부여하여 상기 객체 추적 및 데이터셋 수집 모듈에 전달하는 것인, 객체 인식 처리 장치.

청구항 3

제1항에 있어서,

상기 객체 탐지 모듈은 상기 AR 영상에서 각 객체별로 바운딩 박스를 생성하고, 각 객체별 바운딩 박스에 상기 객체 카테고리 정보를 부여하여 상기 객체 인지 모듈에 전달하고,

상기 객체 인지 모듈은 상기 객체 카테고리 정보를 추가로 활용하여 상기 객체의 식별 정보를 출력하는 것인, 객체 인식 처리 장치.

청구항 4

제1항에 있어서,

상기 객체 탐지 모듈은 상기 AR 영상에서 기탐지되어 추적중인 객체가 존재하는 경우 해당 객체의 영역을 제외한 나머지 영역에 대해 객체 탐지를 수행하는 것인, 객체 인식 처리 장치.

청구항 5

제1항에 있어서,

상기 객체 추적 및 데이터셋 수집 모듈은 상기 AR 영상으로부터 상기 객체 탐지 모듈을 통해 구분된 각 프레임별 객체에 삼 네트워크(Siamese Network)를 적용하여, 각 프레임별로 유사도가 가장 큰 객체를 추출하는 방식으로 객체를 추적하는 것인, 객체 인식 처리 장치.

청구항 6

제5항에 있어서,

상기 객체 추적 및 데이터셋 수집 모듈은 상기 객체 추적에 사용된 전체 프레임들 중에서의 각 객체간의 유사도

의 합이 최소가 되는 복수의 프레임 또는 각 프레임들간의 유사도가 임계값 이하인 소정 개수의 프레임을 선택하여 학습을 위한 데이터 셋으로 선택하는 것인, 객체 인식 처리 장치.

청구항 7

제1항에 있어서,

상기 객체 추적 및 데이터셋 수집 모듈은 상기 객체의 탐지 또는 객체의 추적 과정에서 식별된 객체의 특징 정보로부터 추출되는 각 객체의 부가 식별 정보를 상기 학습을 위한 데이터 셋에 추가하고,

상기 객체 학습 모듈은 상기 부가 식별 정보가 추가된 데이터 셋을 기초로 각 객체를 학습하는 것인, 객체 인식 처리 장치.

청구항 8

제1항에 있어서,

상기 객체 학습 모듈은 상기 수신된 데이터 셋의 유사도에 기초하여, 각 객체를 구분하는 상기 객체 식별 정보를 부여하되, AR 영상의 촬영 시간 및 AR 영상이 촬영된 장소의 위치 정보를 추가적으로 고려하여 데이터 셋을 갱신하는 방식으로 학습을 수행하는 것인, 객체 인식 처리 장치.

청구항 9

제1항에 있어서,

상기 객체 인지 모듈은

상기 객체 탐지 모듈을 통해 구분된 객체와 상기 객체 학습 모듈의 각 객체별 데이터 셋과의 유사도에 기초하여, 기존 객체 식별 정보로 분류하거나, 신규 객체 식별 정보를 부여하거나, 미확정으로 분류하는 것인, 객체 인식 처리 장치.

청구항 10

AR(Augmented Reality) 장치를 위한 객체 인식 처리 장치를 이용한 객체 인식 처리 방법에 있어서,

(a) 적어도 하나 이상의 AR 장치로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 단계;

(b) 상기 AR 영상에서 각각의 객체를 구분하고, 구분된 객체 카테고리 정보를 생성하는 단계;

(c) 상기 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출하는 단계;

(d) 상기 추출된 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하여 객체 학습 모듈을 구축하는 단계; 및

(e) 입력된 AR 영상으로부터 상기 구분된 객체를 상기 객체 학습 모듈에 입력하여 객체의 식별 정보를 출력하는 객체 인지 단계를 포함하는, 객체 인식 처리 방법.

청구항 11

제10항에 있어서,

상기 (b) 단계는 상기 AR 영상에서 각 객체별로 바운딩 박스를 생성하고, 각 객체별 바운딩 박스에 상기 객체 카테고리 정보를 부여하는 것인, 객체 인식 처리 방법.

청구항 12

제10항에 있어서,

상기 (b) 단계는 상기 AR 영상에서 각 객체별로 바운딩 박스를 생성하고, 각 객체별 바운딩 박스에 상기 객체 카테고리 정보를 부여하여 상기 (e) 단계로 전달하고,

상기 (e) 단계는 상기 객체 카테고리 정보를 추가로 활용하여 상기 객체의 식별 정보를 출력하는 것인, 객체 인

식 처리 방법.

청구항 13

제10항에 있어서,

상기 (b) 단계는 상기 AR 영상에서 기탐지되어 추적중인 객체가 존재하는 경우 해당 객체의 영역을 제외한 나머지 영역에 대해 객체 탐지를 수행하는 것인, 객체 인식 처리 방법.

청구항 14

제10항에 있어서,

상기 (c) 단계는 상기 AR 영상으로부터 각 프레임별 객체에 삼 네트워크(Siamese Network)를 적용하여, 각 프레임별로 유사도가 가장 큰 객체를 추출하는 방식으로 객체를 추적하는 것인, 객체 인식 처리 방법.

청구항 15

제14항에 있어서,

상기 (c) 단계는 상기 객체 추적에 사용된 전체 프레임들 중에서의 각 객체간의 유사도의 합이 최소가 되는 복수의 프레임 또는 각 프레임들간의 유사도가 임계값 이하인 소정 개수의 프레임을 선택하여 학습을 위한 데이터 셋으로 선택하는 것인, 객체 인식 처리 방법.

청구항 16

제10항에 있어서,

상기 (c) 단계는 상기 객체의 탐지 또는 객체의 추적 과정에서 식별된 객체의 특징 정보로부터 추출되는 각 객체의 부가 식별 정보를 상기 학습을 위한 데이터 셋에 추가하고,

상기 (d) 단계는 상기 부가 식별 정보가 추가된 데이터 셋을 기초로 각 객체를 학습하는 것인, 객체 인식 처리 방법.

청구항 17

제10항에 있어서,

상기 (d) 단계는 상기 수신된 데이터 셋의 유사도에 기초하여, 각 객체를 구분하는 상기 객체 식별 정보를 부여하되, AR 영상의 촬영 시간 및 AR 영상이 촬영된 장소의 위치 정보를 추가적으로 고려하여 데이터 셋을 갱신하는 방식으로 학습을 수행하는 것인, 객체 인식 처리 방법.

청구항 18

제10항에 있어서,

상기 (e) 단계는 상기 (b) 단계를 통해 구분된 객체와 상기 객체 학습 모듈의 각 객체별 데이터 셋과의 유사도에 기초하여, 기존 객체 식별 정보로 분류하거나, 신규 객체 식별 정보를 부여하거나, 미확정으로 분류하는 것인, 객체 인식 처리 방법.

청구항 19

제10항 내지 제 18 항 중 어느 한 항에 따른 방법을 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 매체.

발명의 설명

기술 분야

[0001] 본 발명은 AR 장치를 위한 객체 인식 처리 장치 및 방법에 관한 것이다.

배경 기술

- [0002] 증강현실(AR) 시스템은 특별한 이벤트(박물관, 기계 조립 등)에서만 스마트폰 화면을 통해 장면을 인지하는 이벤트형 사용방식에서 출발하여, 최근에는 스마트 글라스와 같은 웨어러블 장치를 착용하는 형태로 변화하고 있다.
- [0003] 기존의 이벤트형 AR 시스템을 위해서는 특정한 조건을 인식하기 위한 식별 수단(QR코드, 번호판, 위치 좌표)을 제공하거나 소수의 구성요소(전시물, 기계 부품)에 대해 학습하여 증강할 정보를 얻을 수 있었다. 하지만 기존 학습 및 추론 방식을 새로운 사용방식에 적용할 경우, 기존 이벤트형 AR시스템과 같이 소수의 구성요소로 전체 객체의 범위를 한정하지 않는다면, 학습 난이도의 문제로 객체들의 카테고리(사람, 자동차, 의자 등)를 분류(classify)하는 정도의 수준만을 얻을 수 있게 된다.
- [0004] 이를 극복하기 위해서는 단순히 학습, 모델을 늘리는 것이 아닌 AR의 특성을 고려한 효율적인 대안이 필요하다. 그 이유는, 첫번째로 모든 객체에 대해서 충분히 학습할 수 있는 데이터 셋(dataset)을 만드는 것이 매우 어렵고, 두번째로 시간과 공간을 한정하여 범위 내의 모든 객체에 대한 데이터 셋을 확보한다고 하더라도 한정된 범위를 넘어가는 상황에서 매번 새로운 객체에 대한 모델을 단시간 내에 수정하는 것은 한계가 있기 때문이다.
- [0005] 또한, 통제된 상황에서의 기존 딥러닝 학습과는 달리 객체 이미지의 훼손(블러, 가려짐 등)이 빈번한 AR 환경 또한 고려되어야 한다.
- [0006] 새로운 AR 시스템용 딥러닝 알고리즘의 요구사항은 다음과 같이 크게 2가지이다.
- [0007] 첫번째는 객체의 카테고리를 넘어서 객체별 구분을 가능한 성능을 확보하는 것이며, 두번째로 객체 이미지가 훼손된 경우에도 이를 극복할 수 있어야 한다.
- [0008] 객체 분류(object detection/classification)와 관련하여 일반적으로 널리 알려진 이미지 딥러닝 기법은 전체 이미지 안에서 특정 객체가 존재하는 영역(Bounding box)과 해당 객체의 카테고리를 동시에 추론하는 기법으로 Mask RCNN, YOLO v3, SSD 등의 다양한 연구가 진행되고 있다. 또한 입력 데이터가 동영상인 경우에는 매 프레임마다 이미지 딥러닝 기법을 사용하는 것은 비효율적이므로, 객체의 카테고리나 무관(class-agnostic)하게 특정 객체의 경계 상자(bounding box)가 주어지면 연속된 프레임에서 해당 객체의 이동을 파악하여 이동된 경계 상자를 제공하는 객체 추적(tracking) 딥러닝 기법이 사용되며, MDNet, SiamFC, SiamRPN++ 등의 다양한 연구가 진행되고 있다.
- [0009] 이미지 기반의 객체 검출(object detection)과 대비하여 객체 추적(object tracking)은 같은 객체가 시간에 따라 각도, 조명 등의 변화에 의해 다른 모습을 보이게 되는 객체 변형(object deformation)에 더 쉬운 대처가 가능하다는 특징이 있다.
- [0010] 따라서 이러한 특징을 활용하면 AR 시스템에서 객체 추적을 활용하여 학습 데이터 셋을 구축하는 동시에, 이 데이터 셋을 통해 학습된 분류 모델(classification model)과 삼 유사성 모델(Siamese similarity model)의 조합을 통해서 카테고리보다 더 구체적인 객체 ID별 구분이 가능하며 객체 이미지 훼손(deformation) 또한 극복할 수 있을 것으로 생각된다.

발명의 내용

해결하려는 과제

- [0011] 본 발명의 일 과제는, 증강현실 시스템의 객체 인식 알고리즘의 성능을 강화시킬 수 있는 새로운 학습 및 추론 방법이 적용된 객체 인식 처리 장치를 제공하는 것을 목적으로 한다.
- [0012] 본 발명의 과제는 이상에서 언급된 과제들로 제한되지 않으며, 언급되지 않은 다른 과제들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0013] 상기한 기술적 과제를 해결하기 위한 기술적 수단으로서, 본 발명의 일 실시예에 따른 AR(Augmented Reality) 장치를 위한 객체 인식 처리 장치는, 통신모듈; 객체 인식 프로그램이 저장된 메모리; 및 상기 객체 인식 프로그램을 실행하는 프로세서를 포함하되, 상기 객체 인식 프로그램은 적어도 하나 이상의 AR 장치로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 AR 정보 수집 모듈, 상기 AR 영상에서 각각의 객체를 구분하는 객체 탐지 모듈, 상기 객체 탐지 모듈을 통해 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위

한 데이터 셋을 추출하는 객체 추적 및 데이터셋 수집 모듈, 상기 객체 추적 및 데이터셋 수집 모듈로부터 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하는 객체 학습 모듈 및 입력된 AR 영상으로부터 상기 객체 탐지 모듈을 통해 구분된 객체를 상기 객체 학습 모듈에 입력하여 객체의 식별 정보를 출력하는 객체 인지 모듈을 포함한다.

[0014] 본 발명의 다른 실시예에 따른, AR(Augmented Reality) 장치를 위한 객체 인식 처리 장치를 이용한 객체 인식 처리 방법은, 적어도 하나 이상의 AR 장치로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 단계; 상기 AR 영상에서 각각의 객체를 구분하고, 구분된 객체 카테고리 정보를 생성하는 단계; 상기 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출하는 단계; 상기 추출된 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하여 객체 학습 모듈을 구축하는 단계; 및 입력된 AR 영상으로부터 상기 구분된 객체를 상기 객체 학습 모듈에 입력하여 객체의 식별 정보를 출력하는 객체 인지 단계를 포함한다.

발명의 효과

[0015] 전술한 본원의 과제의 해결 수단에 의하면, AR 시스템에서 객체 추적을 활용하여 학습 데이터 셋을 구축하는 동시에, 이 데이터 셋을 통해 학습된 분류 모델(classification model)과 삼 유사성 모델(Siamese similarity model)의 조합을 통해서 카테고리보다 더 구체적인 객체 ID별 구분이 가능하며 객체 이미지 왜손(deformation) 또한 극복할 수 있다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 AR 장치를 위한 객체 인식 처리 장치의 구성을 도시한 블록도이다.
 도 2는 본 발명의 일 실시예에 따른 객체 인식 프로그램의 구성을 도시한 블록도이다.
 도 3은 본 발명의 일 실시예에 따른 객체 탐지 모듈의 동작을 설명하기 위한 도면이다.
 도 4와 도 5는 본 발명의 일 실시예에 따른 객체 추적 및 데이터셋 수집 모듈의 동작을 설명하기 위한 도면이다.
 도 6은 본 발명의 일 실시예에 따른 객체 학습 모듈의 동작을 설명하기 위한 도면이다.
 도 7은 본 발명의 일 실시예에 따른 객체 인지 모듈의 동작을 설명하기 위한 도면이다.
 도 8은 본 발명의 일 실시예에 따른 객체 인식 방법을 도시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 아래에서는 첨부한 도면을 참조하여 본원이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본원의 실시예를 상세히 설명한다. 그러나 본원은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본원을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0018] 본원 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다.

[0019] 본원 명세서 전체에서, 어떤 부재가 다른 부재 "상에" 위치하고 있다고 할 때, 이는 어떤 부재가 다른 부재에 접해 있는 경우뿐 아니라 두 부재 사이에 또 다른 부재가 존재하는 경우도 포함한다.

[0020] 본원 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함" 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것을 의미한다. 본원 명세서 전체에서 사용되는 정도의 용어 "약", "실질적으로" 등은 언급된 의미에 고유한 제조 및 물질 허용오차가 제시될 때 그 수치에서 또는 그 수치에 근접한 의미로 사용되고, 본원의 이해를 돕기 위해 정확하거나 절대적인 수치가 언급된 개시 내용을 비양심적인 침해자가 부당하게 이용하는 것을 방지하기 위해 사용된다. 본원 명세서 전체에서 사용되는 정도의 용어 "~(하는) 단계" 또는 "~의 단계"는 "~를 위한 단계"를 의미하지 않는다.

[0021] 이하, 첨부한 도면들 및 후술되어 있는 내용을 참조하여 본 발명의 바람직한 실시예들을 상세히 설명한다. 그러나, 본 발명은 여기서 설명되어지는 실시예들에 한정되지 않고 다른 형태로 구체화될 수도 있다. 명세서 전체에

걸쳐서 동일한 참조번호들은 동일한 구성요소들을 나타낸다.

- [0022] 아래에서는 첨부한 도면을 참조하여 본원이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본원의 실시예를 상세히 설명한다. 그러나 본원은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본원을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0023] 본원 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다.
- [0024] 본원 명세서 전체에서, 어떤 부재가 다른 부재 "상에" 위치하고 있다고 할 때, 이는 어떤 부재가 다른 부재에 접해 있는 경우뿐 아니라 두 부재 사이에 또 다른 부재가 존재하는 경우도 포함한다.
- [0025] 본원 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함" 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성 요소를 더 포함할 수 있는 것을 의미한다. 본원 명세서 전체에서 사용되는 정도의 용어 "약", "실질적으로" 등은 언급된 의미에 고유한 제조 및 물질 허용오차가 제시될 때 그 수치에서 또는 그 수치에 근접한 의미로 사용되고, 본원의 이해를 돕기 위해 정확하거나 절대적인 수치가 언급된 개시 내용을 비양심적인 침해자가 부당하게 이용하는 것을 방지하기 위해 사용된다. 본원 명세서 전체에서 사용되는 정도의 용어 "~(하는) 단계" 또는 "~의 단계"는 "~를 위한 단계"를 의미하지 않는다.
- [0026] 이하, 첨부한 도면들 및 후술되어 있는 내용을 참조하여 본 발명의 바람직한 실시예들을 상세히 설명한다. 그러나, 본 발명은 여기서 설명되어지는 실시예들에 한정되지 않고 다른 형태로 구체화될 수도 있다. 명세서 전체에 걸쳐서 동일한 참조번호들은 동일한 구성요소들을 나타낸다.
- [0027] 이하, 본 발명의 일 실시예에 따른 산업재해 사건 추출 시스템에 대하여 설명한다.
- [0028] 도 1은 본 발명의 일 실시예에 따른 AR 장치를 위한 객체 인식 처리 장치의 구성을 도시한 블록도이다.
- [0029] 도 1을 참조하여 설명하면, 객체 인식 처리 장치(100)는 복수의 AR 장치(200~204)로부터 수집되는 AR 영상을 이용하여, 객체 인식을 처리하는 것으로서, 마치 서버와 같은 동작을 수행할 수 있다. 이를 위해, 객체 인식 처리 장치(100)는 통신모듈(110), 메모리(120), 프로세서(130), 데이터베이스(140)를 포함한다.
- [0030] 다음으로, 통신 모듈(110)은 근거리 통신망(Local Area Network; LAN), 광역 통신망(Wide Area Network; WAN) 또는 부가가치 통신망(Value Added Network; VAN) 등과 같은 유선 네트워크나 이동 통신망(mobile radio communication network) 또는 위성 통신망 등과 같은 모든 종류의 무선 네트워크를 사용하는 통신 모듈을 포함할 수 있다. 특히, 통신 모듈(110)은 외부 통신망을 통해 접속하는 각 AR 장치(200~204)와의 통신 인터페이스를 제공한다.
- [0031] 메모리(120)는 객체 인식 프로그램이 저장된 것이다. 객체 인식 프로그램은 적어도 각각의 AR 장치(200~204)로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 AR 정보 수집 모듈(310), AR 영상에서 각각의 객체를 구분하는 객체 탐지 모듈(320), 객체 탐지 모듈(320)을 통해 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출하는 객체 추적 및 데이터셋 수집 모듈(330), 객체 추적 및 데이터셋 수집 모듈(330)로부터 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하는 객체 학습 모듈(340) 및 입력된 AR 영상으로부터 객체 탐지 모듈(320)을 통해 구분된 객체를 객체 학습 모듈(340)에 입력하여 객체의 식별 정보를 출력하는 객체 인지 모듈(350)을 포함한다.
- [0032] 메모리(120)는 전원이 공급되지 않아도 저장된 정보를 계속 유지하는 비휘발성 저장장치 및 저장된 정보를 유지하기 위하여 전력을 필요로 하는 휘발성 저장장치를 통칭하는 것으로 해석되어야 한다. 또한, 메모리(120)는 프로세서(130)가 처리하는 데이터를 일시적 또는 영구적으로 저장하는 기능을 수행할 수 있다. 메모리(120)는 저장된 정보를 유지하기 위하여 전력이 필요한 휘발성 저장장치 외에 자기 저장 매체(magnetic storage media) 또는 플래시 저장 매체(flash storage media)를 포함할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.
- [0033] 프로세서(130)는 메모리(120)에 저장된 객체 인식 프로그램을 실행한다. 프로세서(130)는 데이터를 제어 및 처리하는 다양한 종류의 장치들을 포함할 수 있다. 프로세서(130)는 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 일 예에서, 프로세서(200)는 마이크로프로세서(microprocessor), 중앙처리장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific

integrated circuit), FPGA(field programmable gate array) 등의 형태로 구현될 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다.

- [0034] 또한, 데이터베이스(140)는 각 AR 장치로부터 수집되는 각종 AR 영상과 AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보 등을 관리할 수 있다.
- [0035] 한편, 객체 인식 처리 장치(100)는 각 AR 장치(200~204)로부터 AR 영상을 수신하고, 객체 인식 프로그램을 통해 식별된 객체에 대한 정보를 제공하는 서버로서 동작할 수 있다. 이때, 객체 인식 처리 장치(100)는 SaaS (Software as a Service), PaaS (Platform as a Service) 또는 IaaS (Infrastructure as a Service)와 같은 클라우드 컴퓨팅 서비스 모델에서 동작할 수 있다. 또한, 객체 인식 처리 장치(100)는 사설(private) 클라우드, 공용(public) 클라우드 또는 하이브리드(hybrid) 클라우드와 같은 형태로 구축될 수 있다.
- [0036] 도 2는 본 발명의 일 실시예에 따른 객체 인식 프로그램의 구성을 도시한 블록도이다.
- [0037] 객체 인식 프로그램은 적어도 각각의 AR 장치(200~204)로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신하는 AR 정보 수집 모듈 (310), AR 영상에서 각각의 객체를 구분하는 객체 탐지 모듈(320), 객체 탐지 모듈(320)을 통해 구분된 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출하는 객체 추적 및 데이터셋 수집 모듈(330), 객체 추적 및 데이터셋 수집 모듈(330)로부터 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하는 객체 학습 모듈(340) 및 입력된 AR 영상으로부터 객체 탐지 모듈(320)을 통해 구분된 객체를 객체 학습 모듈(340)에 입력하여 객체의 식별 정보를 출력하는 객체 인지 모듈(350)을 포함한다.
- [0038] AR 정보 수집 모듈 (310)은 앞서 정의한 바와 같이, 각 AR 장치(200~204)로부터 AR 영상뿐만 아니라, 해당 AR 영상이 촬영되는 시간, 해당 AR 영상이 촬영된 장소의 위치 정보(예를 들면 GPS 정보 등), 해당 영상을 촬영한 AR 장치에 대한 정보(예를 들면, 단말기 고유 번호)를 각각 수집한다. 특히, AR 장치에 대한 정보는 향후, 각 AR 장치를 사용하는 사용자를 구분하는데 사용될 수 있다.
- [0039] 도 3은 본 발명의 일 실시예에 따른 객체 탐지 모듈의 동작을 설명하기 위한 도면이다.
- [0040] 객체 탐지 모듈(320)은 AR 영상에서 각 객체별로 바운딩 박스를 생성하고, 각 객체별 바운딩 박스에 객체 카테고리 정보를 부여하여 객체 추적 및 데이터셋 수집 모듈(330)에 전달한다.
- [0041] 도 3에 도시된 바와 같이, AR 영상에서 각 객체별로 바운딩 박스를 생성하는데, 바운딩 박스 생성을 위한 알고리즘으로는 YOLO 등 종래의 기술을 사용한다. 그리고, 각 객체별 바운딩 박스에 객체 카테고리 정보를 부여하여 객체 추적 및 데이터셋 수집 모듈(330)에 전달한다. 객체 추적 및 데이터셋 수집 모듈(330)은 객체 탐지 모듈(320)에서 수신한 객체 카테고리 정보를 기초로, 객체 추적을 수행하며, 객체 추적 수행 결과를 다시 객체 탐지 모듈(320)에 전달하여, 객체 탐지에 사용되도록 한다.
- [0042] 한편, 객체의 탐지 단계에서 객체 카테고리 분류 결과가 도출되면, 객체 카테고리 정보를 객체 인지 모듈(350)에 전달하고, 객체 인지 모듈(350)은 객체 카테고리 분류 결과를 추가로 활용하여 객체의 식별 정보를 출력한다. 객체 탐지 모듈(320)은 객체의 카테고리 정보로서, 예를 들면, 사람, 자동차, 동물 등의 분류 정보를 제공할 수 있다.
- [0043] 또한, 도 3에 도시된 바와 같이, AR 영상에서 기탐지되어 추적중인 객체와 기존에 추적하지 않던 객체가 함께 존재하는 경우 추적 중인 객체의 영역을 제외한 나머지 영역에 대해, 즉 새롭게 등장한 객체에 대하여 탐지를 수행한다. 이와 같이, 새롭게 탐지된 객체에 대해서는 새로운 객체로서 인식하여 객체 추적 및 데이터셋 수집 모듈(330)과 객체 인지 모듈(350)에 각각 전달할 수 있다.
- [0044] 도 4와 도 5는 본 발명의 일 실시예에 따른 객체 추적 및 데이터셋 수집 모듈의 동작을 설명하기 위한 도면이다.
- [0045] 객체 추적 및 데이터셋 수집 모듈(330)은 객체 탐지 모듈(320)을 통해 구분된 객체 카테고리 정보를 수신하고, 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출한다.
- [0046] 도 4에 도시된 바와 같이, AR 영상으로부터 복수의 구분된 프레임의 이미지를 수신하는데, 각 프레임별로 객체 탐지 모듈(320)을 통해 구분된 각 프레임별 객체에 대한 정보를 수신한다. 그리고, 객체 추적 및 데이터셋 수집 모듈(330)은 삼 네트워크(Siamese Network)를 적용하여, 각 프레임별로 유사도가 가장 큰 객체를 추출하는

방식으로 객체를 추적할 수 있다.

[0047] 특히, 아래의 수식식 1과 같이, 객체 추적에 사용된 전체 프레임들 중에서의 각 객체간의 유사도의 합이 최소가 되는 복수의 프레임 또는 각 프레임들간의 유사도가 임계값 이하인 소정 개수의 프레임을 선택하여 학습을 위한 데이터 셋으로 선택한다.

[0048] [수식식 1]

$$\operatorname{argmin}_{K \subset U} \sum_{i \in K} \sum_{j \in K} S_{ij}, n(K) = 4$$

[0049]

[0050] i, j는 프레임 번호를 나타냄, S_{ij} 는 두 프레임 i, j간 유사도를 나타냄

[0051] K는 전체 프레임의 집합인 U의 부분 집합이며, $n(K)$ 는 집합 K에 속하는 프레임의 개수임

[0052] 이와 같이, 특정 객체를 추적한 영상 내에서 유사도의 차이가 큰 프레임을 대상으로 각 객체에 대한 데이터 셋을 수집하므로, 각 객체의 다양한 변형 형태에 대한 학습 모델 구축이 가능해진다.

[0053] 이때, 유사도 판단을 위한 삼 네트워크의 구성은 도 5에 도시된 바와 같으며, 각 프레임에 포함된 객체별로 유사도를 산출하는 알고리즘으로 알려져 있다.

[0054] 그리고, 객체 추적 및 데이터셋 수집 모듈(330)은 객체의 탐지 또는 객체의 추적 과정에서 식별된 객체의 특징 정보로부터 추출되는 각 객체의 부가 식별 정보를 상기 학습을 위한 데이터 셋에 추가하고, 객체 학습 모듈(340)은 부가 식별 정보가 추가된 데이터 셋을 기초로 각 객체를 학습한다.

[0055] 예를 들면, 객체의 특징 정보로는 사람의 경우 객체의 얼굴의 특징 정보를 이용하거나, 자동차의 경우 자동차 번호판을 통해 식별되는 특징 정보, 또는 각 물품에 부착된 바코드 등의 정보를 특징 정보로서 이용할 수 있다.

[0056] 이와 같은 정보는 객체를 대표하는 특징 정보인 만큼, 이를 이용하면, 객체를 보다 정확하게 식별할 수 있고, 이를 기반으로 각 객체의 식별 정보를 보다 용이하게 설정할 수 있게 된다.

[0057] 도 6은 본 발명의 일 실시예에 따른 객체 학습 모듈의 동작을 설명하기 위한 도면이다.

[0058] 객체 학습 모듈(340)은 객체 추적 및 데이터셋 수집 모듈(330)로부터 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행한다.

[0059] 도 6에 도시된 바와 같이, 객체 학습 모듈(340)은 수신된 데이터 셋의 유사도에 기초하여, 각 객체를 구분하는 객체 식별 정보를 부여하되, AR 영상의 촬영 시간 및 AR 영상이 촬영된 장소의 위치 정보를 추가적으로 고려하여 데이터 셋을 갱신하는 방식으로 학습을 수행한다.

[0060] 예를 들면, 각 AR 장치로부터 수집된 AR 영상이 촬영된 장소의 위치 정보를 기준으로, 서로 구분된 위치(위치 A, 위치 B) 별로 각 객체 식별 정보가 부여된 데이터 셋이 수집된다. 이때, 데이터 셋은 각 객체 식별 정보 별로 유사한 프레임 이미지, 객체 탐지 결과(객체의 카테고리에 대한 정보)등을 포함하고 있는 것이다.

[0061] 그리고, 앞선 단계를 순차적으로 진행하여, 새로운 데이터 셋이 입력되면, 기존의 데이터 셋과의 유사도 비교를 통해, 유사도가 큰 데이터 셋을 중심으로 데이터 셋을 갱신한다. 이 과정에서 서로 다른 AR 장치에서 수집된 데이터 셋을 포함하도록, 데이터 셋을 갱신하여, 다양한 특징이 포함하면서 고른 분포를 가진 데이터 셋이 포함되도록 한다.

[0062] 한편, 데이터 셋 간의 유사도를 판별하거나, 각 지역에 적합한 형태로 데이터 셋을 갱신하는 과정에 앞서 설명한 삼 네트워크가 사용될 수 있다.

[0063] 도 7은 본 발명의 일 실시예에 따른 객체 인지 모듈의 동작을 설명하기 위한 도면이다.

[0064] 객체 인지 모듈(350)은 입력된 AR 영상으로부터 객체 탐지 모듈(320)을 통해 구분된 객체를 객체 학습 모듈(340)에 입력하여 객체의 식별 정보를 출력한다.

[0065] 이때, 객체 인지 모듈(350)은 객체 탐지 모듈(320)을 통해 구분된 객체와 객체 학습 모듈(340)의 각 객체별 데이터 셋과의 유사도에 기초하여, 기존 객체 식별 정보로 분류하거나, 신규 객체 식별 정보를 부여하거나, 미확

정으로 분류할 수 있다. 이때, 입력 AR 영상과 객체 학습 모듈(340)의 데이터 셋 간의 유사도를 판별하기위해 앞서 설명한 삼 네트워크가 사용될 수 있다.

- [0066] 입력 AR 영상과 특정 데이터 셋간의 유사도가 임계값 이상인 경우에는 기존에 객체 학습 모듈(340)에서 해당 데이터 셋을 대표하는 객체 식별 정보(ID)에 해당하는 것으로 분류한다.
- [0067] 입력 AR 영상과 전체 데이터 셋간의 유사도가 임계값에 도달하지 못한 경우에는 객체 학습 모듈(340)에서 해당 데이터 셋을 확보하지 못한 것으로 보고, 새로운 객체 식별 정보(ID)를 부여한 후, 해당 객체 식별 정보(ID)를 기초로 데이터 셋을 구축하도록 한다.
- [0068] 입력 AR 영상과 복수의 데이터 셋간의 유사도가 임계값 이상인 경우에는 경우에는 특정 객체 식별 정보(ID)를 부여하지 않고, 미확정으로 분류한다.
- [0069] 도 8은 본 발명의 일 실시예에 따른 객체 인식 방법을 도시한 순서도이다.
- [0070] 먼저, 적어도 하나 이상의 AR 장치로부터 AR 영상, AR 영상의 촬영 시간, AR 영상이 촬영된 장소의 위치 정보, 해당 영상을 촬영한 AR 장치에 대한 정보를 수신한다(S810). 객체 인식 처리 장치(100)의 AR 정보 수집 모듈(310)에서 수행하는 동작으로서, 이와 같이 수신된 AR 데이터는 학습 모듈 구축에 사용되고, 이후 새롭게 입력된 AR 영상에 대하여 추론 과정을 수행하는데 사용된다.
- [0071] 다음으로, AR 영상에서 각각의 객체를 구분하고, 구분된 객체 카테고리 정보를 생성한다(S820). 객체 인식 처리 장치(100)의 객체 탐지 모듈(320)에서 수행하는 동작으로서, 영상에서 객체별로 바운딩 박스를 생성하고, 해당 객체에 대한 카테고리 분류를 수행한다. 이와 같이 탐지된 객체 정보는 이후 객체 추적 및 데이터 셋 수집 단계(S830)와 객체 인지 단계(S850)로 각각 전달된다.
- [0072] 다음으로, 객체 카테고리 정보를 수신하고 각 객체의 유사도에 기초하여 객체를 추적하고, 해당 객체의 학습을 위한 데이터 셋을 추출한다(S830). 객체 인식 처리 장치(100)의 객체 추적 및 데이터 셋 수집 모듈(330)에서 에서 수행하는 동작으로서, 삼 네트워크를 이용하여 유사도를 산출하고, 앞서 설명한 수학적 식 1을 통해 최적의 프레임 데이터 셋으로서 추출하되, 예를 들면, 각 객체간의 유사도의 합이 최소가 되는 복수의 프레임 또는 각 프레임들간의 유사도가 임계값 이하인 소정 개수의 프레임을 선택하여 학습을 위한 데이터 셋으로 선택한다. 또한, 객체의 탐지 또는 객체의 추적 과정에서 식별된 객체의 특징 정보로부터 추출되는 각 객체의 부가 식별 정보를 학습을 위한 데이터 셋에 추가하여, 각 ID 별로 데이터 셋의 정확성을 향상시킬 수 있다.
- [0073] 다음으로, 추출된 데이터 셋을 수신하고, 수신된 데이터 셋의 유사도에 기초하여 각 객체를 구분하기 위한 학습을 수행하여 객체 학습 모듈을 구축한다(S840). 수신된 데이터 셋의 유사도에 기초하여, 각 객체를 구분하는 상 기 객체 식별 정보를 부여하되, AR 영상의 촬영 시간 및 AR 영상이 촬영된 장소의 위치 정보를 추가적으로 고려하여 데이터 셋을 갱신하는 방식으로 학습을 수행한다.
- [0074] 다음으로, 입력된 AR 영상으로부터 구분된 객체를 객체 학습 모듈에 입력하여 객체의 식별 정보를 출력하는 객체 인지 단계를 수행한다(S850). 객체와 객체 학습 모듈의 각 객체별 데이터 셋과의 유사도에 기초하여, 기존 객체 식별 정보로 분류하거나, 신규 객체 식별 정보를 부여하거나, 미확정으로 분류할 수 있다.
- [0075] 본 발명의 일 실시예에 따른 객체 인식 방법은 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독가능 매체는 컴퓨터 저장 매체를 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다.
- [0076] 본 발명의 방법 및 시스템은 특정 실시예와 관련하여 설명되었지만, 그것들의 구성 요소 또는 동작의 일부 또는 전부는 범용 하드웨어 아키텍처를 갖는 컴퓨터 시스템을 사용하여 구현될 수 있다.
- [0077] 진술한 본원의 설명은 예시를 위한 것이며, 본원이 속하는 기술분야의 통상의 지식을 가진 자는 본원의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

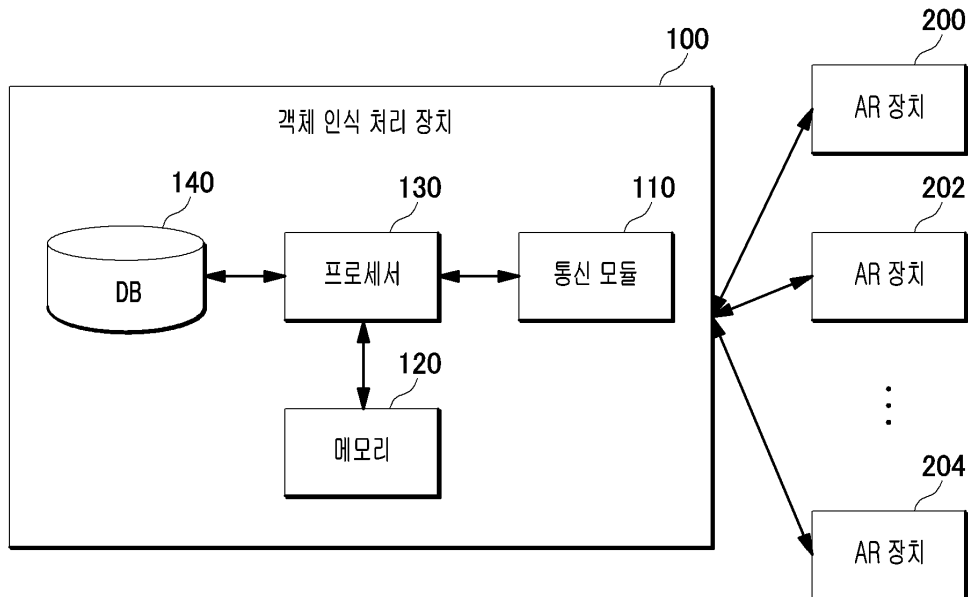
[0078] 본원의 범위는 상기 상세한 설명보다는 후술하는 청구범위에 의하여 나타내어지며, 청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본원의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

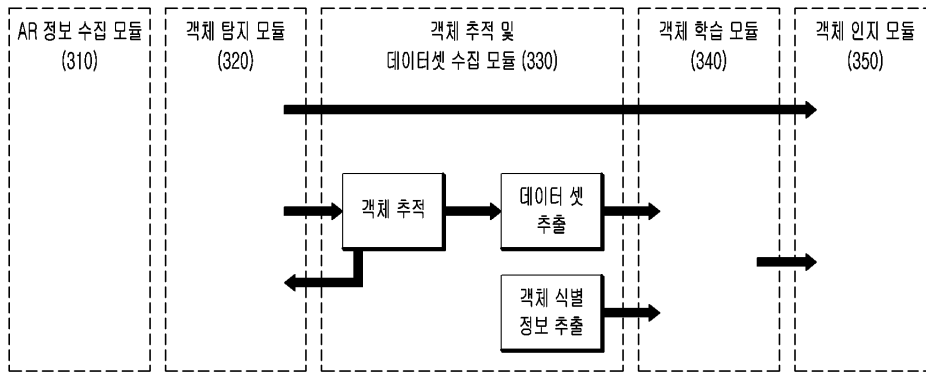
[0079] 100: 객체 인식 처리 장치
 110: 통신 모듈
 120: 메모리
 130: 프로세서
 140: 데이터베이스
 300: 객체 인식 프로그램
 310: AR 정보 수집 모듈
 320: 객체 탐지 모듈
 330: 객체 추적 및 데이터 셋 수집 모듈
 340: 객체 학습 모듈
 350: 객체 인지 모듈

도면

도면1



도면2

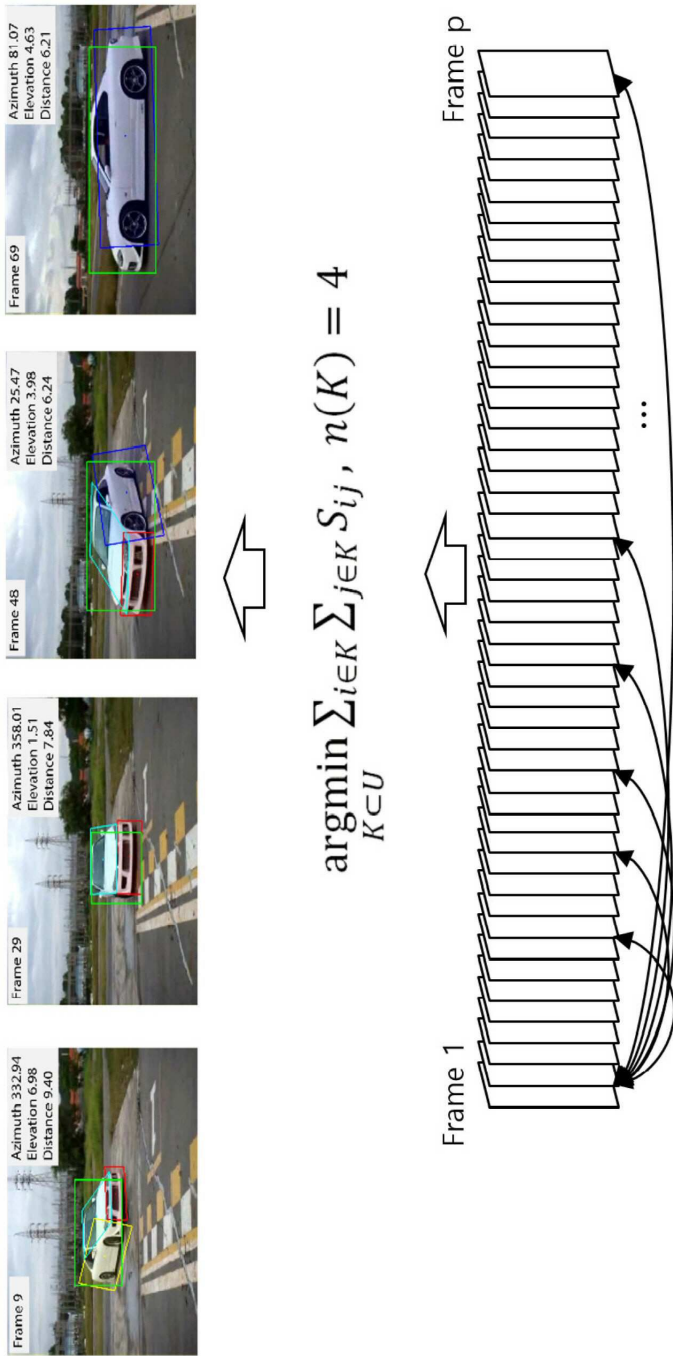


300

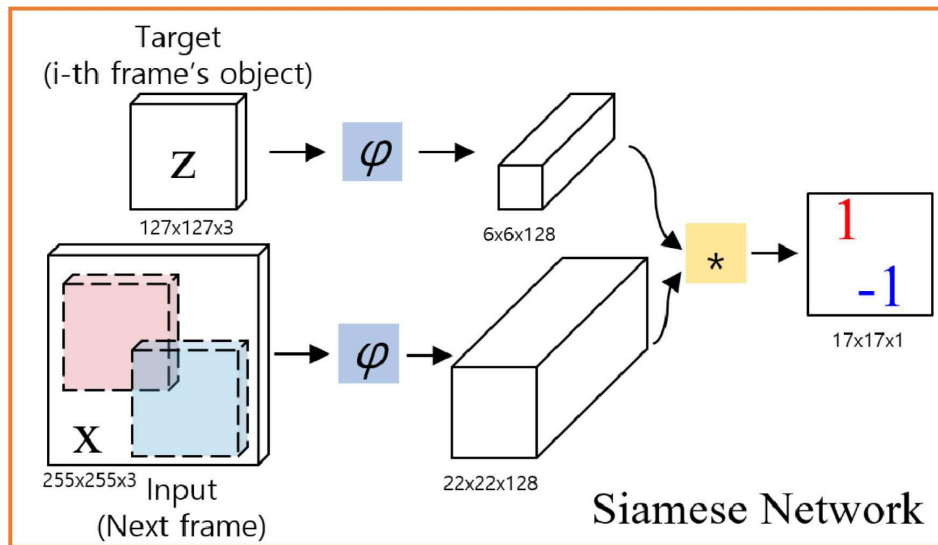
도면3



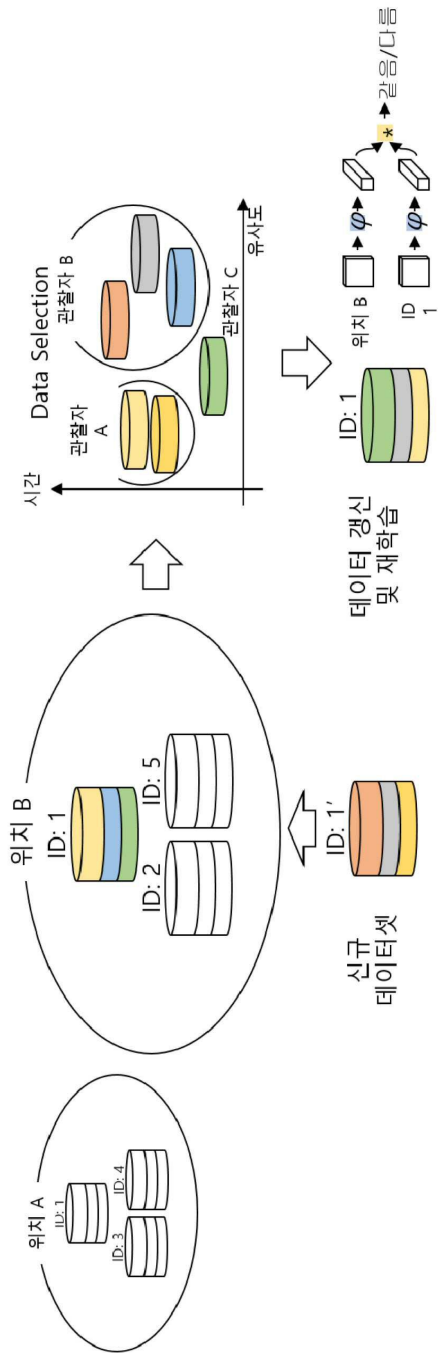
도면4



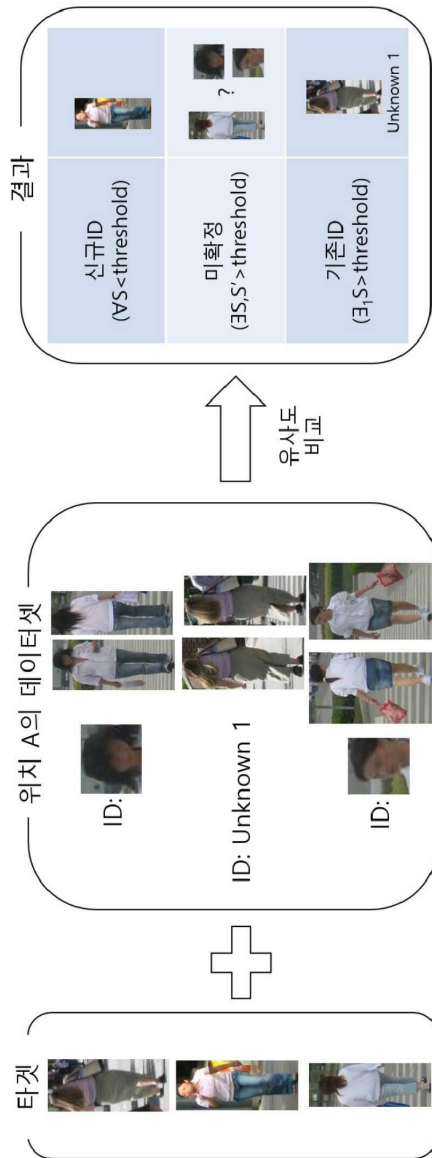
도면5



도면6



도면7



도면8

