(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0071676
(43) 공개일자 2025년05월22일

(51) 국제특허분류(Int. Cl.)
G06V 20/40 (2022.01) *G06N 3/0455* (2023.01)
G06T 5/00 (2024.01) *G06T 7/11* (2017.01)
G06T 7/50 (2017.01) *G06V 10/82* (2022.01)
G06V 20/52 (2022.01)
(52) CPC특허분류
G06V 20/47 (2022.01)
G06N 3/0455 (2023.01)
(21) 출원번호 10-2023-0158470
(22) 출원일자 2023년11월15일
심사청구일자 2023년11월15일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
김영갑
서울특별시 광진구 능동로 209, 대양AI센터 701호
팔라시 잉글
서울특별시 광진구 능동로 209, 대양AI센터 732호
김덕한
서울특별시 광진구 능동로 209, 대양AI센터 732호
(74) 대리인
수안특허법인, 민영준

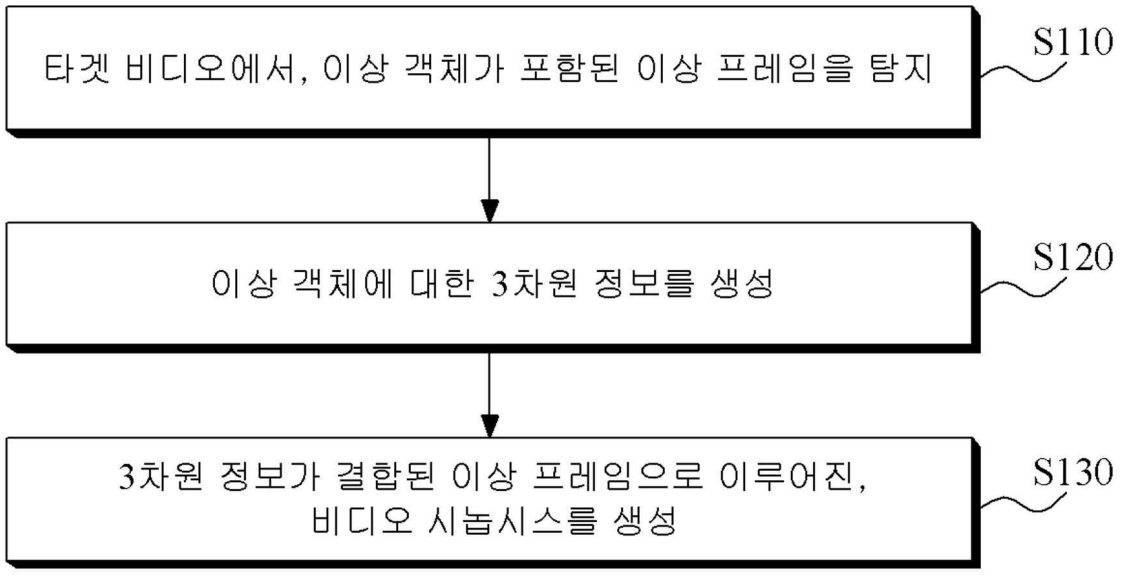
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **이상 객체에 대한 3차원 비디오 시놉시스 생성 방법**

(57) 요 약

이상 객체에 대한 3차원 정보를 제공하는 비디오 시놉시스 생성 방법이 개시된다. 개시된 이상 객체에 대한 비디오 시놉시스 생성 방법은, 타겟 비디오에서, 이상 객체가 포함된 이상 프레임을 탐지하는 단계; 상기 이상 객체에 대한 3차원 정보를 생성하는 단계; 및 상기 3차원 정보가 결합된 이상 프레임으로 이루어진, 비디오 시놉시스를 생성하는 단계를 포함한다.

대 표 도 - 도1



(52) CPC특허분류
G06T 5/00 (2024.01)
G06T 7/11 (2017.01)
G06T 7/50 (2017.01)
G06V 10/82 (2022.01)
G06V 20/52 (2023.08)
G06T 2207/10028 (2013.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193682
과제번호	2019-0-00231-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보보호핵심원천기술개발(R&D)
연구과제명	공공 인프라 안전을 위한 인공지능 기반 영상보안 기술 및 시스템 개발
기 여 율	1/1
과제수행기관명	세종대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

타겟 비디오에서, 이상 객체가 포함된 이상 프레임을 탐지하는 단계;

상기 이상 객체에 대한 3차원 정보를 생성하는 단계; 및

상기 3차원 정보가 결합된 이상 프레임으로 이루어진, 비디오 시놉시스를 생성하는 단계

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 2

제 1항에 있어서,

상기 3차원 정보를 생성하는 단계는

미리 학습된 멀티 태스크 딥러닝 모델을 이용하여, 상기 이상 프레임에서 이상 객체 영역을 분할하고, 상기 이상 프레임에 대한 깊이 맵을 생성하는 단계; 및

상기 깊이 맵을 이용하여, 상기 이상 객체에 대한 3차원 포인트 클라우드를 생성하는 단계

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 3

제 2항에 있어서,

상기 멀티 태스크 딥러닝 모델은

상기 이상 프레임을 인코딩하는 인코더;

상기 인코딩된 데이터를 입력받아, 상기 깊이 맵을 생성하는 제1디코더; 및

상기 인코딩된 데이터를 입력받아, 상기 이상 객체 영역이 분할된 이미지를 생성하는 제2디코더

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 4

제 2항에 있어서,

상기 비디오 시놉시스를 생성하는 단계는

상기 3차원 포인트 클라우드의 포인트 별 z축 좌표값을 이용하여, 포인트별 x축 및 y축 좌표값 중 적어도 하나를 보정하는 단계;

2차원 이미지인 상기 이상 프레임에서, 상기 이상 객체의 위치에 상기 좌표값이 보정된 3차원 포인트 클라우드를 오버랩하는 단계

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 5

제 4항에 있어서,

상기 포인트별 x축 및 y축 좌표값 중 적어도 하나를 보정하는 단계는

상기 z축 좌표값만큼 상기 x축 및 y축 좌표값 중 적어도 하나를 증가 또는 감소시키는

이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 6

제 4항에 있어서,

상기 3차원 포인트 클라우드는

사용자의 요청에 따라 선택적으로 오버랩되는

이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 7

제 1항에 있어서,

상기 비디오 시놉시스를 생성하는 단계는

상기 이상 프레임을 이용하여, 상기 이상 객체에 대한 부가 정보를 생성하는 단계; 및

상기 부가 정보를 상기 이상 프레임에 추가하여, 상기 비디오 시놉시스를 생성하는 단계

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 8

제 7항에 있어서,

상기 부가 정보는

상기 이상 객체의 등장 시간 정보 및 등장 위치 정보

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 9

제 1항에 있어서,

상기 이상 프레임을 탐지하는 단계는

상기 이상 프레임에 대한 이상 점수를 생성하며,

상기 비디오 시놉시스를 생성하는 단계는

동일한 이상 객체가 포함된 이상 프레임이 연속된 경우, 연속된 이상 프레임 중에서, 상기 이상 점수가 최대인 이상 프레임으로 이루어진 비디오 시놉시스를 생성하는

이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 10

타겟 비디오를 입력받는 단계;

상기 타겟 비디오의 프레임 각각에 대해 깊이 맵을 생성하고, 이상 객체 영역 분할을 수행하는 단계; 및

상기 프레임 중, 이상 객체 영역 분할이 수행된 이상 프레임에, 상기 이상 객체에 대한 3차원 정보를 결합하여, 상기 3차원 정보가 결합된 이상 프레임으로 이루어진 비디오 시놉시스를 생성하는 단계

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 11

제 10항에 있어서,

상기 비디오 시놉시스를 생성하는 단계는

상기 깊이 맵을 이용하여, 상기 이상 객체 영역에 대한 3차원 포인트 클라우드를 생성하고, 상기 이상 프레임에 상기 3차원 포인트 클라우드를 결합하는

이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 12

제 10항에 있어서,

상기 멀티 태스크 딥러닝 모델은

상기 타겟 비디오의 프레임을 인코딩하는 인코더;

상기 인코딩된 데이터를 입력받아, 상기 깊이 맵을 생성하는 제1디코더; 및

상기 인코딩된 데이터를 입력받아, 상기 이상 객체 영역이 분할된 이미지를 생성하는 제2디코더

를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법.

청구항 13

타겟 비디오에서, 미리 설정된 클래스의 타겟 객체가 포함된 타겟 프레임을 탐지하는 단계;

상기 타겟 객체에 대한 3차원 정보를 생성하는 단계; 및

상기 3차원 정보가 결합된 타겟 프레임으로 이루어진, 비디오 시놉시스를 생성하는 단계

를 포함하는 비디오 시놉시스 생성 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 이상 객체에 대한 비디오 시놉시스 생성 방법에 관한 것으로서, 더욱 상세하게는 이상 객체에 대한 3차원 정보를 제공하는 비디오 시놉시스 생성 방법에 관한 것이다.

배 경 기 술

[0003] 비디오 감시 카메라의 증가에 따라, 이미지 데이터의 양이 과도하게 증가하였다. 이에 따라, 이미지 데이터의 분석, 추출 및 저장이 어려워지고, 더 많은 데이터 저장 공간과 분석 인력이 필요하게 되었다. 이에 따라 비디오를 짧게 요약하는 비디오 시놉시스 기술에 대한 관심이 높아지고 있다. 비디오 시놉시스 기술은, 이미지 데이터의 양을 감소시키고 몇 시간 분량의 비디오를 단 몇 분 만에 검토할 수 있도록 해주기 때문에, 비디오 시놉시

스 기술을 이용하면, 데이터 저장공간 문제를 해소하고 이미지 분석에 소요되는 비용과 노동력을 크게 감소시킬 수 있다.

[0004] 다만, 비디오 시놉시스는 원본 비디오로부터 일부 이미지, 즉 프레임이 제거되기 때문에, 비디오 시놉시스를 통해서는, 프레임에 포함된 객체 등에 대한 정보가 용이하게 확인되기 어려울 수 있다. 이에 따라, 원본 비디오로부터 프레임 이외 다양한 부가 정보를 추가하여 비디오 시놉시스를 생성하는 방법 등에 대한 연구가 이루어지고 있다.

[0005] 관련 선행문헌으로 특허 문헌인 대한민국 등록특허 제10-2570126호, 대한민국 공개특허 제2023-0112223호, 제2021-0058638호가 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명은, 미리 설정된 클래스의 객체에 대한 비디오 시놉시스 생성 방법을 제공하기 위한 것이다.

[0008] 특히 본 발명은 이상 객체에 대한 비디오 시놉시스 생성 방법을 제공하기 위한 것이다.

[0009] 또한 본 발명은 사용자의 효율적인 분석을 위해 이상 객체에 대한 다양한 부가 정보를 제공하고, 비디오 시놉시스 생성 시간을 줄일 수 있는, 비디오 시놉시스 생성 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0011] 상기한 목적을 달성하기 위한 본 발명의 일 실시예에 따르면, 타겟 비디오에서, 이상 객체가 포함된 이상 프레임을 탐지하는 단계; 상기 이상 객체에 대한 3차원 정보를 생성하는 단계; 및 상기 3차원 정보가 결합된 이상 프레임으로 이루어진, 비디오 시놉시스를 생성하는 단계를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법이 제공된다.

[0012] 또한 상기한 목적을 달성하기 위한 본 발명의 다른 실시예에 따르면, 타겟 비디오를 입력받는 단계; 상기 타겟 비디오의 프레임 각각에 대해 깊이 맵을 생성하고, 이상 객체 영역 분할을 수행하는 단계; 및 상기 프레임 중, 이상 객체 영역 분할이 수행된 이상 프레임에, 상기 이상 객체에 대한 3차원 정보를 결합하여, 상기 3차원 정보가 결합된 이상 프레임으로 이루어진 비디오 시놉시스를 생성하는 단계를 포함하는 이상 객체에 대한 비디오 시놉시스 생성 방법이 제공된다.

[0013] 또한 상기한 목적을 달성하기 위한 본 발명의 또 다른 실시예에 따르면, 타겟 비디오에서, 미리 설정된 클래스의 타겟 객체가 포함된 타겟 프레임을 탐지하는 단계; 상기 타겟 객체에 대한 3차원 정보를 생성하는 단계; 및 상기 3차원 정보가 결합된 타겟 프레임으로 이루어진, 비디오 시놉시스를 생성하는 단계를 포함하는 비디오 시놉시스 생성 방법이 제공된다.

발명의 효과

[0015] 본 발명의 일실시예에 따르면 이상 객체에 대한 3차원 정보가 포함된 이상 객체에 대한 비디오 시놉시스가 생성됨으로써, 사용자는 이상 객체에 대한 보다 효율적인 분석을 수행할 수 있다.

[0016] 또한 본 발명의 일실시예에 따르면, 이상 객체 탐지, 시맨틱 세그멘테이션, 깊이 맵 생성이라는 복수의 태스크를 동시에 수행하는 멀티 태스크 딥러닝 모델을 이용함으로써, 비디오 시놉시스 생성 시간이 단축될 수 있다.

도면의 간단한 설명

[0018] 도 1은 본 발명의 일실시예에 따른 비디오 시놉시스 생성 방법을 설명하기 위한 도면이다.

도 2는 본 발명의 일실시예에 따른 딥러닝 모델을 설명하기 위한 도면이다.

도 3은 본 발명의 일실시예에 따른 3차원 정보 결합 방법을 설명하기 위한 도면이다.

도 4는 본 발명의 다른 실시예에 따른 이상 객체에 대한 비디오 시놉시스 생성 방법을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0019] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0020] 이하에서, 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.

[0022] 도 1은 본 발명의 일실시예에 따른 비디오 시놉시스 생성 방법을 설명하기 위한 도면이며, 도 2는 본 발명의 일실시예에 따른 딥러닝 모델을 설명하기 위한 도면이다.

[0023] 본 발명의 일실시예에 따른 비디오 시놉시스 생성 방법은 메모리 및 프로세서를 포함하는 컴퓨팅 장치에서 수행될 수 있으며, 일예로서, 이상 객체에 대한 비디오 시놉시스를 생성할 수 있다.

[0024] 본 발명의 일실시예에 따른 컴퓨팅 장치는 타겟 비디오에서, 이상 객체가 포함된 이상 프레임을 탐지(S110)한다. 여기서 타겟 비디오는 예컨대 CCTV에서 녹화된 2차원 비디오일 수 있으며, 이상 객체는 총, 칼 등의 무기 등을 소지한 사람일 수 있다. 컴퓨팅 장치는 객체를 탐지하고 탐지된 객체의 클래스를 분류하는 다양한 객체 탐지 모델을 이용하여, 타겟 비디오의 프레임에서 이상 객체를 탐지할 수 있으며, 이상 객체를 포함하는 이상 프레임에 대한 이상 점수를 생성할 수 있다. 여기서, 이상 점수란 객체 탐지 모델에 의해 탐지된 객체가 이상 객체인 확률을 나타내는 점수이다.

[0025] 그리고 컴퓨팅 장치는 이상 객체에 대한 3차원 정보를 생성(S120)하며, 일예로서, 이상 객체에 대한 포인트 클라우드를, 3차원 정보로 생성할 수 있다. 실시예에 따라서 컴퓨팅 장치는 이상 프레임에 포함된 모든 객체에 대한 3차원 정보를 생성할 수 있다.

[0026] 컴퓨팅 장치는, 미리 학습된 멀티 태스크 딥러닝 모델을 이용하여, 이상 프레임에서 이상 객체 영역을 분할하고, 이상 프레임에 대한 깊이 맵을 생성할 수 있다. 그리고 깊이 맵을 이용하여, 이상 객체에 대한 3차원 포인트 클라우드를 생성할 수 있다. 분할된 이상 객체의 화소별로 할당된 깊이 맵의 깊이값으로부터, x, y, z 축 좌표값으로 정의되는 포인트 클라우드가 생성될 수 있다.

[0027] 전술된 멀티 태스크 딥러닝 모델은 입력된 이미지에 대해 깊이 맵을 생성하고 시맨틱 세그멘테이션을 수행하도록 학습된 모델로서, U-NET 기반의 인코더-디코더 모델일 수 있다. 멀티 태스크 딥러닝 모델은 도 2에 도시된 바와 같이, 인코더(210), 제1디코더(220) 및 제2디코더(230)를 포함할 수 있다.

[0028] 인코더(210)는 입력된 이상 프레임(211)을 인코딩하여, 이상 프레임(200)에 대한 잠재 벡터(latent vector)를 생성한다. 그리고 제1 및 제2디코더(210, 220)는 인코딩된 데이터 즉, 잠재 벡터를 입력받아, 각각 깊이 맵(211)과 이상 객체 영역이 분할된 이미지(221)를 생성한다. 본 발명의 일실시예에 따른 멀티 태스크 딥러닝 모델은 복수의 디코더를 이용함으로써, 깊이 맵 생성과 이상 객체 영역 분할을 병렬적으로 동시에 수행할 수 있다.

[0029] 다시 도 1로 돌아가, 컴퓨팅 장치는 이상 프레임에 이상 객체에 대한 3차원 정보를 결합하여, 3차원 정보가 결합된 이상 프레임으로 이루어진, 비디오 시놉시스를 생성(S130)한다. 따라서 사용자는 3차원 정보를 렌더링하는 툴을 통해, 이상 객체를 다양한 시점에서 확인할 수 있으며, 이상 객체를 효율적으로 분석할 수 있다.

[0030] 컴퓨팅 장치는 단계 S130에서, 이상 객체에 대한 부가 정보를 생성할 수 있다. 컴퓨팅 장치는 이상 프레임을 이용하여, 이상 객체에 대한 부가 정보를 생성하고, 부가 정보를 이상 프레임에 추가하여, 비디오 시놉시스를 생성할 수 있다. 부가 정보는 일예로서, 이상 객체의 등장 시간 정보 및 등장 위치 정보 등을 포함할 수 있다.

[0031] 컴퓨팅 장치는 타겟 비디오에서의, 이상 프레임의 재생 시간으로부터, 비디오 영상에서의 이상 객체의 등장 시간 정보를 생성할 수 있으며, 타겟 비디오 가 촬영된 위치 정보로부터 이상 객체의 등장 위치 정보를 생성할 수 있다.

[0032] 또한 컴퓨팅 장치는 더욱 효율적으로 요약된 비디오 시놉시스를 생성하기 위해, 동일한 이상 객체가 포함된 이

상 프레임이 연속된 경우, 연속된 이상 프레임 중 일부 프레임을 선택적으로 비디오 시놉시스에 포함시킬 수 있다. 전술된 바와 같이, 이상 프레임 탐지 과정에서, 이상 점수가 생성되는데, 컴퓨팅 장치는 동일한 이상 객체가 포함된 이상 프레임이 연속된 경우, 연속된 이상 프레임 중에서, 이상 점수가 최대인 이상 프레임으로 이루어진 비디오 시놉시스를 생성할 수 있다.

[0033] 이와 같이, 본 발명의 일실시예에 따르면 이상 객체에 대한 3차원 정보가 포함된 이상 객체에 대한 비디오 시놉시스가 생성됨으로써, 사용자는 이상 객체에 대한 보다 효율적인 분석을 수행할 수 있다.

[0034] 한편, 전술된 실시예에서는 이상 객체에 대한 비디오 시놉시스 생성 방법이 설명되었지만, 본 발명은 이상 객체뿐만 아니라, 미리 설정된 클래스의 객체에 대한 비디오 시놉시스를 생성하기 위해 이용될 수 있다. 이 경우, 본 발명의 일실시예는 타겟 비디오에서, 미리 설정된 클래스의 타겟 객체가 포함된 타겟 프레임을 탐지하고, 타겟 객체에 대한 3차원 정보가 결합된 타겟 프레임으로 이루어진, 비디오 시놉시스를 생성할 수 있다.

[0036] 도 3은 본 발명의 일실시예에 따른 3차원 정보 결합 방법을 설명하기 위한 도면으로서, 도 3(a)는 칼을 소지한 이상 객체를 포함하는 이상 프레임을 나타내며, 도 3(b)는 포인트 클라우드가 오버랩된 이상 프레임의 일예를 나타낸다.

[0037] 전술된 바와 같이 컴퓨팅 장치는 3차원 정보로서 포인트 클라우드를 이상 프레임에 결합할 수 있으며, 2차원 이미지인 이상 프레임에 포인트 클라우드를 오버랩하는 방식으로, 이상 프레임에 3차원 정보를 결합할 수 있다. 2차원 이미지에 3차원 정보인 포인트 클라우드가 오버랩될 경우, z축 좌표값, 즉 깊이 정보가 표현될 수 없으며, 이에 컴퓨팅 장치는 3차원 포인트 클라우드의 포인트 별 z축 좌표값을 이용하여, 포인트별 x축 및 y축 좌표값 중 적어도 하나를 보정한다. 그리고 이상 프레임에서, 분할된 이상 객체의 위치에 x축 및 y축 좌표값이 보정된 3차원 포인트 클라우드를 오버랩함으로써, z축 좌표값을 이상 프레임에 표현한다.

[0038] 컴퓨팅 장치는 일실시예로서, 포인트 클라우드의 z축 좌표값만큼 x축 및 y축 좌표값 중 적어도 하나를 증가시키거나 또는 감소시켜, x축 및 y축 좌표값을 보정한 후, 이상 프레임의 이상 객체에 오버랩함으로써, 이상 객체에 대한 깊이 정보를 이상 프레임에 표현할 수 있다. 깊이값이 클수록 z축 좌표값 역시 커지므로, 포인트 클라우드는 도 3(b)와 같이, 이상 객체를 중심으로 특정 방향으로 분산되는 형태로 이상 프레임에 오버랩될 수 있다.

[0039] 한편, 포인트 클라우드는 실시예에 따라서 사용자의 요청에 따라 선택적으로 오버랩될 수 있으며, 오버랩된 포인트 클라우드는 사용자의 요청에 따라서, 이상 프레임에서 제거될 수 있다.

[0041] 도 4는 본 발명의 다른 실시예에 따른 이상 객체에 대한 비디오 시놉시스 생성 방법을 설명하기 위한 도면이다.

[0042] 전술된 실시예와 달리, 도 4에서는 별도의 이상 프레임 탐지 과정없이 비디오 시놉시스가 생성된다. 이상 객체 영역을 분할하는 과정에서 시맥틱 세그멘테이션에 의해 이상 객체가 탐지되므로, 본 발명의 일실시예에 따른 컴퓨팅 장치는 별도의 객체 탐지 모델을 이용해 이상 프레임을 탐지하지 않고, 전술된 멀티 태스크 딥러닝 모델을 이용하여 이상 프레임 탐지, 깊이 맵 생성, 이상 객체 영역 분할을 모두 수행할 수 있다.

[0043] 도 4를 참조하면, 본 발명의 일실시예에 따른 컴퓨팅 장치는 타겟 비디오를 입력(S410)받고, 입력된 타겟 비디오의 프레임 각각에 대해 깊이 맵을 생성하고, 이상 객체 영역 분할을 수행(S420)한다. 컴퓨팅 장치는 전술된 멀티 태스크 딥러닝 모델을 이용하여, 타겟 비디오의 전체 프레임에 대해 깊이 맵을 생성하고, 이상 객체 영역 분할을 수행할 수 있다. 이 때, 멀티 태스크 딥러닝 모델의 인코더로, 타겟 비디오의 프레임이 입력된다.

[0044] 그리고 컴퓨팅 장치는 타겟 비디오의 프레임 중, 이상 객체 영역 분할이 수행된 이상 프레임에, 이상 객체에 대한 3차원 정보를 결합하여, 3차원 정보가 결합된 이상 프레임으로 이루어진 비디오 시놉시스를 생성(S430)한다. 전술된 바와 같이 컴퓨팅 장치는 깊이 맵을 이용하여, 이상 객체 영역에 대한 3차원 포인트 클라우드를 생성하고, 이상 프레임에 3차원 포인트 클라우드를 결합할 수 있다.

[0045] 실시예에 따라서, 컴퓨팅 장치는 어텐션 매커니즘을 이용하여, 정상 객체와 비정상 객체를 명확하게 식별할 수 있으며, 타겟 비디오의 전체 프레임 중 이상 프레임을 명확하게 탐지할 수 있다.

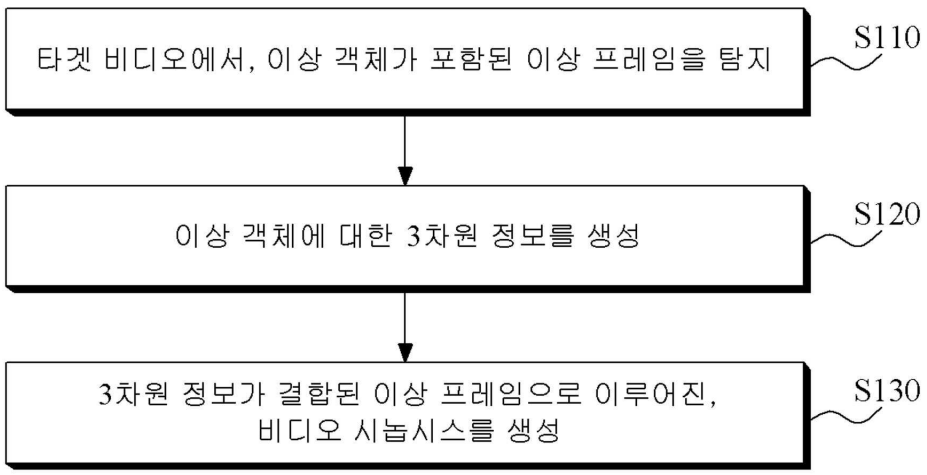
[0046] 이와 같이 본 발명의 일실시예에 따르면, 이상 객체 탐지, 시맨틱 세그멘테이션, 깊이 맵 생성이라는 복수의 태스크를 동시에 수행하는 멀티 태스크 딥러닝 모델을 이용함으로써, 비디오 시놉시스 생성 시간이 단축될 수 있다.

[0048] 앞서 설명한 기술적 내용들은 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 실시예들을 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 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드 뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 하드웨어 장치는 실시예들의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

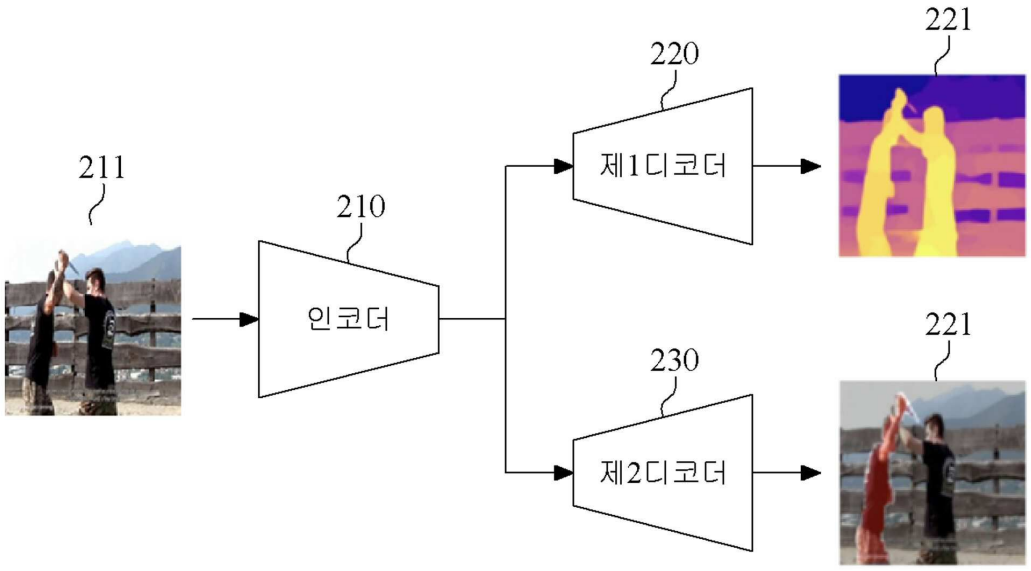
[0050] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

도면

도면1



도면2



도면3

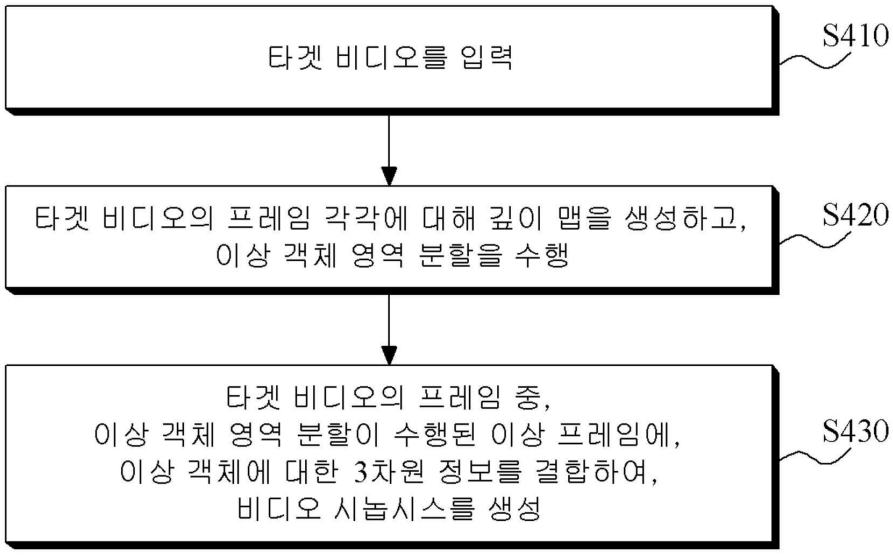


(a)



(b)

도면4


타겟 비디오를 입력
S410
타겟 비디오의 프레임 각각에 대해 깊이 맵을 생성하고,
이상 객체 영역 분할을 수행
S420
타겟 비디오의 프레임 중,
이상 객체 영역 분할이 수행된 이상 프레임에,
이상 객체에 대한 3차원 정보를 결합하여,
비디오 시놉시스를 생성
S430