



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년07월01일
(11) 등록번호 10-2416227
(24) 등록일자 2022년06월29일

- (51) 국제특허분류(Int. Cl.)
G06Q 50/08 (2012.01) B64C 39/02 (2006.01)
B64D 47/00 (2006.01) G06N 20/00 (2019.01)
G06N 3/08 (2006.01) G06Q 10/06 (2012.01)
H04N 5/14 (2006.01)
- (52) CPC특허분류
G06Q 50/08 (2013.01)
B64C 39/024 (2013.01)
- (21) 출원번호 10-2019-0173757
- (22) 출원일자 2019년12월24일
심사청구일자 2019년12월24일
- (65) 공개번호 10-2021-0081618
- (43) 공개일자 2021년07월02일
- (56) 선행기술조사문헌
KR101548639 B1*
KR101656642 B1*
KR1020180045745 A*
KR1020190072077 A*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
- (72) 발명자
김형관
서울특별시 강남구 도산대로1길 40, 대영빌딩 5층(신사동)
방성덕
서울시 강북구 오패산로 227 태성빌리지 C동 202호
- (74) 대리인
김인철

전체 청구항 수 : 총 6 항

심사관 : 육성원

(54) 발명의 명칭 건설 객체 실시간 모니터링 장치 및 모니터링 방법 및 이를 실행하기 위한 컴퓨터프로그램

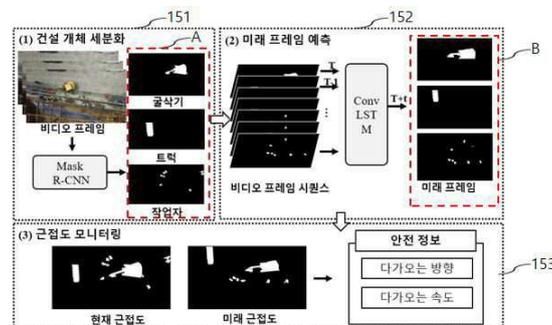
(57) 요약

본 발명은 드론의 센싱 능력을 활용하여 안전 정보의 양을 극대화하고, 실질적인 안전 사고 저감에 기여할 수 있는 건설 객체 실시간 모니터링 장치, 방법 및 이를 실행하기 위한 컴퓨터프로그램에 관한 것으로,

본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치는,

건설 현장을 촬영한 영상 데이터를 수신하고, 외부 디바이스로 안전정보를 전송하는 통신부; 상기 건설 현장의 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비하며, 상기 통신부를 통해 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하여 이를 이진 이미지로 변환하고, 변환된 이진 이미지를 상기 예측 모델에 입력하여 미래 프레임을 생성하고, 생성된 미래 프레임과 상기 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하여 상기 안전정보를 생성하는 모니터링부;를 포함한다.

대표도



(52) CPC특허분류

- B64D 47/00 (2013.01)
- G06N 20/00 (2021.08)
- G06N 3/08 (2013.01)
- G06Q 10/063 (2013.01)
- H04N 5/145 (2013.01)
- B64C 2201/127 (2013.01)

이 발명을 지원한 국가연구개발사업

- 과제고유번호 1711084302
- 부처명 과학기술정보통신부
- 과제관리(전문)기관명 한국연구재단
- 연구사업명 개인기초연구(과기정통부)(R&D)
- 연구과제명 증강현실 및 심층학습 기반 위험요인 선제적 경보 인터페이스 개발을 위한 건설현장 위험요인 모니터링 및 시각화 기술
- 기여율 1/2
- 과제수행기관명 연세대학교
- 연구기간 2019.03.01 ~ 2020.02.29

이 발명을 지원한 국가연구개발사업

- 과제고유번호 1345295664
 - 부처명 교육부
 - 과제관리(전문)기관명 한국연구재단
 - 연구사업명 이공학학술연구기반구축(R&D)
 - 연구과제명 기후변화 적응형 사회기반시설 연구센터
 - 기여율 1/2
 - 과제수행기관명 연세대학교
 - 연구기간 2019.03.01 ~ 2020.02.29
-

명세서

청구범위

청구항 1

건설 현장을 촬영한 영상 데이터를 수신하고, 외부 디바이스로 안전정보를 전송하는 통신부;

상기 건설 현장의 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비하며, 상기 통신부를 통해 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하여 이를 이진 이미지로 변환하고, 변환된 이진 이미지를 상기 예측 모델에 입력하여 미래 프레임을 생성하고, 생성된 미래 프레임과 상기 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하여 상기 안전정보를 생성하는 모니터링부;를 포함하고,

상기 모니터링부는,

상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하고, 검출된 복수개의 건설 객체의 수만큼 복수개의 이진 이미지를 생성하되, 각각의 이진 이미지는 단일 건설 객체를 포함하는 객체 세분화 모듈;

상기 객체 세분화 모듈에 의해 생성된 복수개의 이진 이미지 중에서 동일한 건설 객체를 포함하는 복수개의 이진 이미지를 시간 순으로 정렬한 이진 이미지 시퀀스를 상기 예측 모델의 입력값으로 하여 수초 내지 수십초 후의 건설 객체의 위치 및 자세가 포함된 이진 이미지의 미래 프레임을 생성하는 미래 프레임 예측 모듈;

상기 미래 프레임 예측 모듈에 의해 생성된 미래 프레임과, 상기 시간 순으로 정렬한 이진 이미지 시퀀스 중에서 선택된 비교 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하여 상기 안전정보를 생성하는 근접도 모니터링 모듈;을 포함하며,

상기 근접도 모니터링 모듈은,

상기 미래 프레임과 상기 비교 이진 이미지에 포함된 건설 장비에 해당하는 건설 객체의 다각형 꼭지점을 검출하는 코너 검출 유닛과,

건설 장비에 해당하는 건설 객체의 모든 꼭지점 및 점으로 처리된 건설 현장의 인부로부터 다른 건설 객체의 모든 꼭지점까지의 각각의 직선 거리를 계산하여 그 중 가장 작은 최소 거리를 두 건설 객체 간의 근접도로 결정하여, 픽셀 단위의 근접도를 추출하고, 상기 픽셀 단위의 근접도를 실제 크기로 스케일링하여 실제 크기 단위의 근접도를 산출하는 것을 특징으로 하는 건설 객체 실시간 모니터링 장치.

청구항 2

삭제

청구항 3

청구항 1에 있어서, 상기 객체 세분화 모듈은,

상기 건설 현장의 건설 객체들을 학습 데이터로 하여 미리 학습된 기계학습 모델인 Mask R-CNN을 포함하고, 상기 Mask R-CNN을 이용하여 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하는 건설 객체 실시간 모니터링 장치.

청구항 4

청구항 1에 있어서, 상기 미래 프레임 예측 모듈은,

상기 건설 현장의 건설 객체의 이진 이미지 시퀀스를 이용하여 미리 학습된 예측 모델인 convolutional LSTM Network를 포함하는 건설 객체 실시간 모니터링 장치.

청구항 5

삭제

청구항 6

건설 현장의 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비하는 모니터 링부가,

통신부를 통해 수신된 건설 현장을 촬영한 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하 여 이를 이진 이미지로 변환하는 제1 단계;

상기 변환된 이진 이미지를 상기 예측 모델에 입력하여 미래 프레임을 생성하는 제2 단계;

상기 생성된 미래 프레임과 상기 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하는 제3 단계;

상기 도출된 근접도를 이용하여 안전정보를 생성하는 제4 단계를 포함하고,

상기 제1 단계는 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하고, 검출된 복수개의 건설 객체의 수만큼 복수개의 이진 이미지를 생성하되, 각각의 이진 이미지는 단일 건설 객체를 포함 하도록 하고,

상기 제2 단계는 생성된 복수개의 이진 이미지 중에서 동일한 건설 객체를 포함하는 복수개의 이진 이미지를 시 간 순으로 정렬한 이진 이미지 시퀀스를 상기 예측 모델의 입력값으로 하여 수초 내지 수십초 후의 건설 객체의 위치 및 자세가 포함된 이진 이미지의 미래 프레임을 생성하며,

상기 제3 단계는 상기 미래 프레임과, 상기 시간 순으로 정렬한 이진 이미지 시퀀스 중에서 선택된 비교 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하며,

상기 제4 단계는 현재 시점의 근접도와 상기 제3 단계에 의해 도출된 근접도를 비교하여 수행하되, 건설 장비에 해당하는 건설 객체의 모든 꼭지점 및 점으로 처리된 건설 현장의 인부로부터 다른 건설 객체의 모든 꼭지점까 지의 각각의 직선 거리를 계산하여 그 중 가장 작은 최소 거리를 두 건설 객체 간의 근접도로 결정하여, 픽셀 단위의 근접도를 추출하고, 상기 픽셀 단위의 근접도를 실제 크기로 스케일링하여 실제 크기 단위의 근접도를 산출하는 것을 특징으로 하는 건설 객체 실시간 모니터링 방법.

청구항 7

삭제

청구항 8

청구항 6에 있어서,

상기 제1 단계는 Mask R-CNN을 이용하여 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객 체를 검출하고,

상기 제2 단계의 예측 모델은 convolutional LSTM Network를 포함하며,

상기 제3 단계는 Harris 코너 검출 알고리즘을 이용하는 건설 객체 실시간 모니터링 방법.

청구항 9

하드웨어와 결합되어, 청구항 6 또는 청구항 8에 따른 건설 객체 실시간 모니터링 방법을 실행시키기 위하여, 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 드론의 센싱 능력을 활용하여 안전 정보의 양을 극대화하고, 실질적인 안전 사고 저감에 기여할 수 있는 건설 객체 실시간 모니터링 장치, 방법 및 이를 실행하기 위한 컴퓨터프로그램에 관한 것이다.

배경 기술

[0003] 최근 드론 등의 다양한 무인항공기(UAV; Unmanned Aerial Vehicle)의 발전으로 넓은 지역의 이미지 정보를 효율적으로 수집할 수 있게 되었다. 카메라가 장착 된 무인항공기는 폐쇄회로 텔레비전(CCTV) 카메라와 달리 효율적으로 이미지를 수집 할 수 있기 때문에 모니터링에 매우 유용한 도구이다.

[0004] 이러한 드론을 건설 현장 관리에 활용하고자 하는 연구가 활발히 진행되고 있다. 현재 드론의 촬영 이미지로부터 건설 객체 간의 근접도를 모니터링하여 충돌 사고를 예방하려는 연구는 많으나, 건설 객체의 미래 움직임을 예측하여 선제적인 안전 사고 예방 정보를 생성하려는 연구는 많지 않다. 또한, 건설 객체의 크기와 모양을 고려하기 위해 객체 세분화를 활용하여 근접도를 측정하는 연구 역시 수행되지 않았다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) (문헌 1) 한국공개특허공보 제10-2016-0034013호 (2016.03.29)
 (특허문헌 0002) (문헌 2) 한국등록특허공보 제10-1996371호 (2019.07.03)

발명의 내용

해결하려는 과제

[0007] 본 발명은 드론의 센싱 능력을 활용하여 안전 정보의 양을 극대화하고, 실질적인 안전 사고 저감에 기여할 수 있는 건설 객체 실시간 모니터링 장치, 방법 및 이를 실행하기 위한 컴퓨터프로그램을 제공하는 것을 목적으로 한다.

과제의 해결 수단

- [0009] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치는,
- [0010] 건설 현장을 촬영한 영상 데이터를 수신하고, 외부 디바이스로 안전정보를 전송하는 통신부; 상기 건설 현장의 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비하며, 상기 통신부를 통해 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하여 이를 이진 이미지로 변환하고, 변환된 이진 이미지를 상기 예측 모델에 입력하여 미래 프레임을 생성하고, 생성된 미래 프레임과 상기 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하여 상기 안전정보를 생성하는 모니터링부;를 포함한다.
- [0011] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치에 있어서, 상기 모니터링부는, 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하고, 검출된 복수개의 건설 객체의 수만큼 복수개의 이진 이미지를 생성하되, 각각의 이진 이미지는 단일 건설 객체를 포함하는 객체 세분화 모듈; 상기 객체 세분화 모듈에 의해 생성된 복수개의 이진 이미지 중에서 동일한 건설 객체를 포함하는 복수개의 이진 이미지를 시간 순으로 정렬한 이진 이미지 시퀀스를 상기 예측 모델의 입력값으로 하여 상기 미래 프레임을 생성하는 미래 프레임 예측 모듈; 상기 미래 프레임 예측 모듈에 의해 생성된 미래 프레임과, 상기 시간 순으로 정렬한 이진 이미지 시퀀스 중 적어도 어느 하나의 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하여 상기 안전정보를 생성하는 근접도 모니터링 모듈;을 포함할 수 있다.
- [0012] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치에 있어서, 상기 객체 세분화 모듈은, 상기 건설 현장의 건설 객체들을 학습 데이터로 하여 미리 학습된 기계학습 모델인 Mask R-CNN을 포함하고, 상기 Mask R-CNN을 이용하여 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출할 수 있다.
- [0013] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치에 있어서, 상기 미래 프레임 예측 모듈은, 상기 건설 현장의 건설 객체의 이진 이미지 시퀀스를 이용하여 미리 학습된 예측 모델인 convolutional LSTM Network를 포

함할 수 있다.

- [0014] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치에 있어서, 상기 근접도 모니터링 모듈은, 상기 이진 이미지에 포함된 건설 객체의 코너를 검출하는 코너 검출 유닛과, 상기 검출된 코너를 기반으로 도출된 건설 객체 간의 최소 거리를 이용하여 근접도 정보를 산출하는 근접도 측정 유닛을 포함할 수 있다.
- [0016] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 방법은,
- [0017] 건설 현장의 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비하는 모니터링부가, 통신부를 통해 수신된 건설 현장을 촬영한 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하여 이를 이진 이미지로 변환하는 제1 단계; 상기 변환된 이진 이미지를 상기 예측 모델에 입력하여 미래 프레임을 생성하는 제2 단계; 상기 생성된 미래 프레임과 상기 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하는 제3 단계; 상기 도출된 근접도를 이용하여 안전정보를 생성하는 제4 단계를 포함한다.
- [0018] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 방법에 있어서, 상기 제1 단계는 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하고, 검출된 복수개의 건설 객체의 수만큼 복수개의 이진 이미지를 생성하되, 각각의 이진 이미지는 단일 건설 객체를 포함하도록 하고, 상기 제2 단계는 생성된 복수개의 이진 이미지 중에서 동일한 건설 객체를 포함하는 복수개의 이진 이미지를 시간 순으로 정렬한 이진 이미지 시퀀스를 상기 예측 모델의 입력값으로 하여 상기 미래 프레임을 생성하며, 상기 제3 단계는 상기 미래 프레임과, 상기 시간 순으로 정렬한 이진 이미지 시퀀스 중 적어도 어느 하나의 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하며, 상기 제4 단계는 현재 시점의 근접도와 상기 제3 단계에 의해 도출된 근접도를 비교하여 수행할 수 있다.
- [0019] 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 방법에 있어서, 상기 제1 단계는 Mask R-CNN을 이용하여 상기 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하고, 상기 제2 단계의 예측 모델은 convolutional LSTM Network를 포함하며, 상기 제3 단계는 상기 이진 이미지에 포함된 건설 객체의 코너를 검출하는 Harris 코너 검출 알고리즘을 이용하고, 어느 한 건설 객체의 모든 코너점으로부터 다른 건설 객체의 모든 코너점까지의 각각의 직선 거리를 계산하여, 그 중 가장 작은 최소 거리를 두 건설 객체 간의 근접도로 결정할 수 있다.
- [0021] 본 발명은 하드웨어와 결합되어, 본 발명에 따른 건설 객체 실시간 모니터링 방법을 실행시키기 위하여, 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.
- [0023] 기타 본 발명의 다양한 측면에 따른 구현예들의 구체적인 사항은 이하의 상세한 설명에 포함되어 있다.

발명의 효과

- [0025] 본 발명의 일 실시예에 따르면, 기존의 이미지 기반 안전 관리 방법과 상이한 새로운 방법을 구현하는 것으로, 드론의 센싱 능력을 활용하여 안전 정보의 양을 극대화하고, 실질적인 안전 사고 저감에 기여할 수 있다.

도면의 간단한 설명

- [0027] 도 1은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치를 포함하는 건설 객체 실시간 모니터링 시스템이 도시된 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치의 구성이 도시된 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치의 모니터링부가 도시된 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 예측 모델의 학습을 위한 학습 데이터 생성방법을 설명하기 위한 도면이다.
- 도 5는 본 발명의 일 실시예에 따른 근접도 모니터링 모듈에서 근접도를 산출하여 안전정보를 생성하는 과정을 설명하기 위한 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치에 의해 생성된 안전정보의 일 예를 보여주는 도면이다.
- 도 7은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 방법이 도시된 순서도이다.
- 도 8은 본 발명의 실시예에 따른 컴퓨팅 장치를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0028] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예를 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0030] 본 발명에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 발명에서, '포함하다' 또는 '가지다' 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다. 이하, 도면을 참조하여 본 발명의 실시예에 따른 건설 객체 실시간 모니터링 장치 및 모니터링 방법을 설명한다.
- [0032] 도 1은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치를 포함하는 건설 객체 실시간 모니터링 시스템이 도시된 도면이다.
- [0033] 도 1에 도시된 바와 같이, 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 시스템은, 드론(10)으로 건설 현장(20)을 실시간으로 촬영하고, 드론(10)에 의해 촬영된 영상 데이터를 실시간으로 건설 객체 실시간 모니터링 장치(100, 이하 “모니터링 장치” 라고도 함)로 전송하면, 모니터링 장치(100)는 영상 데이터에서 건설 객체를 검출하고, 검출된 건설 객체 관련 이미지를 미리 학습된 예측 모델에 입력하여 수초 ~ 수초십 이후의 미래 예측값(미래 프레임)을 생성하고, 이를 예측 모델 입력 당시의 건설 객체 이미지와 비교하여 건설 객체 간의 근접도를 도출하고, 근접도에 따라 위험 여부를 알리는 안전정보를 작업자 단말기(30)로 전송한다. 작업자 단말기(30)는 건설 현장의 인부 또는 건설 현장에서 사용되는 설비(굴삭기, 트럭 등)의 운전자가 소지하는 이동 단말기일 수 있다.
- [0034] 이와 같은 과정을 통해 건설 현장에서의 건설 객체(굴삭기, 트럭 등)에 의해 다른 건설 객체(작업자 등)에 가해지는 충돌 위험을 미리 작업자(인부 또는 설비 운전자)에게 알려서 건설 현장의 충돌 사고를 예방할 수 있게 된다.
- [0036] 도 2를 참조하여 본 발명의 일 실시예에 따른 모니터링 장치(100)를 설명한다. 도 2는 본 발명의 일 실시예에 따른 모니터링 장치의 구성이 도시된 도면이다.
- [0037] 도 2에 도시된 바와 같이, 본 발명의 일 실시예에 따른 모니터링 장치(100)는, 통신부(110), 입력부(120), 출력부(130), 저장부(140), 모니터링부(150), 제어부(160)를 포함할 수 있다.
- [0039] 통신부(110)는 통신망을 통해 데이터를 송수신하기 위한 기능을 수행한다. 통신부(110)는 드론(10)에서 송출되는 영상 데이터를 수신한다. 또한, 통신부(110)는 위험 여부를 알리는 안전정보를 작업자 단말기(30)로 전송한다.
- [0040] 입력부(120)는 숫자 및 문자 정보 등의 다양한 정보를 입력 받고, 각종 기능을 설정 및 모니터링 장치(100)의 기능 제어와 관련하여 입력되는 신호를 제어부(160)로 전달한다.
- [0041] 출력부(130)는 사용자 입력에 대응하는 사용자 인터페이스를 출력할 수 있다. 출력부(130)는 모니터링 장치(100)의 메뉴 및 사용자가 입력한 사용자 데이터 등을 표시할 수 있다. 출력부(130)가 터치스크린(Touch screen) 형태로 구성된 경우, 출력부(130)는 입력부(120)의 기능 중 일부 또는 전부를 수행할 수 있다.
- [0042] 저장부(140)는 데이터를 저장하기 위한 장치로, 주 기억 장치 및 보조 기억 장치를 포함하고, 모니터링 장치(100)의 기능 동작에 필요한 응용 프로그램을 저장할 수 있다. 저장부(140)는 각종 기계학습 모델, 기능 수행에 필요한 알고리즘을 저장할 수 있다.
- [0043] 모니터링부(150)는 건설 현장의 건설 객체의 이진 이미지 시퀀스를 이용하여 미리 학습된 예측 모델을 구비하며, 통신부(110)를 통해 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하여 이를 이진 이미지로 변환하고, 변환된 이진 이미지를 예측 모델에 입력하여 미래 프레임을 생성하고, 생성된 미래 프레임과 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하여 안전정보를 생성할 수 있다. 모니터링부(150)에 대해 도 3 내지 도 6을 참조하여 후술하도록 한다.
- [0044] 제어부(160)는 모니터링 장치(100)의 제반 동작을 처리하도록 제어한다. 즉, 제어부(160)은 입력부(120), 출력부(130), 통신부(110), 저장부(140), 모니터링부(150)의 동작을 각각 제어한다.

- [0046] 도 3 내지 도 6을 참조하여 모니터링부(150)에 대해 설명한다. 도 3은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치(100)의 모니터링부(150)가 도시된 도면이고, 도 4는 본 발명의 일 실시예에 따른 예측 모델의 학습을 위한 학습 데이터 생성방법을 설명하기 위한 도면이며, 도 5는 본 발명의 일 실시예에 따른 근접도 모니터링 모듈에서 근접도를 산출하여 안전정보를 생성하는 과정을 설명하기 위한 도면이고, 도 6은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치에 의해 생성된 안전정보의 일 예를 보여주는 도면이다.
- [0047] 도 3에 도시된 바와 같이, 모니터링부(150)는 객체 세분화 모듈(151), 미래 프레임 예측 모듈(152), 근접도 모니터링 모듈(153)을 포함한다.
- [0049] 객체 세분화 모듈(151)은 통신부(110)를 통해 수신된 건설 현장을 촬영한 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출한다. 여기서, 건설 객체는 건설 현장에 투입되는 모든 건설 자원으로, 예를 들어 굴삭기, 트럭, 불도저, 크레인, 지게차 등의 건설 장비와, 건설 현장의 인부를 포함한다.
- [0050] 객체 세분화 모듈(151)은 건설 현장의 건설 객체들을 학습 데이터로 하여 미리 학습된 기계학습 모델을 이용하여 영상 프레임에 포함된 건설 객체 각각을 픽셀 단위로 검출하고, 검출된 복수개의 건설 객체의 수만큼 복수개의 이진 이미지를 생성한다. 예를 들어, 영상 프레임에 포함된 건설 객체가 굴삭기, 트럭, 작업자(인부)인 경우, 객체 세분화 모듈(151)은 3개의 건설 객체를 분리하여 각각 별도의 영상 프레임을 생성하고, 각각의 건설 객체가 포함된 영상 프레임에 대해 건설 객체는 “1”, 건설 현장의 배경은 “0”으로 처리하여 3개의 이진 이미지로 변환한다. (도 3의 A 참조)
- [0051] 객체 세분화 모듈(151)에서 사용되는 미리 학습된 기계학습 모델은 Mask R-CNN인 것이 바람직하다. Mask R-CNN(Convolutional Neural Network)은 개념적으로 단순하고 유연하며 일반적인 객체 분류에 대한 프레임워크(framework)이다. Mask R-CNN은 각 객체에 대한 고성능의 분류 마스크를 생성하는 동안 이미지에서 객체를 효율적으로 검출한다. Mask R-CNN은 Faster R-CNN에 대한 확장이며 박스를 통한 인식에 대한 기존의 branch와 더불어 병렬적으로 객체 마스크를 예측하는 것에 대한 branch를 제공한다.
- [0052] Mask R-CNN은 연속적인 영상 프레임들을 입력 받아서 각 프레임 내 건설 객체들을 픽셀 단위로 인식한다. 기존 비전 기반 안전 평가 연구들은 바운딩 박스에 기반한 객체 검출을 활용하여 건설 객체 간 거리를 측정하였다. 그들은 건설 객체의 위치를 점으로 간주하고 객체를 검출한 경계 상자의 중심점 사이의 거리를 객체 간 최단 거리로 계산하였다. 하지만, 충돌 사고는 객체의 중심부가 아닌 끝부분의 충돌로 인해 발생한다. 예를 들어, 굴삭기의 팔의 스윙은 충돌 사고를 직접적으로 유발할 수 있으나 중심부로 근접도를 측정하는 기존의 방식에서 고려되지 않는다. 보다 정확한 객체 간 근접도 측정을 위해 본 발명에서는 객체 세분화에 기반하여 건설 객체를 인식한다. 따라서, 본 발명은 건설 객체의 크기 및 모양을 고려한 건설 객체 간 최단 거리를 측정하는 것을 가능하게 한다.
- [0054] 미래 프레임 예측 모듈(152)은 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비한다. 미래 프레임 예측 모듈(152)은 객체 세분화 모듈(151)에 의해 생성된 복수개의 이진 이미지 중에서 동일한 건설 객체를 포함하는 복수개의 이진 이미지를 시간 순으로 정렬한 이진 이미지 시퀀스를 예측 모델의 입력값으로 하여 수초 내지 수십초 후의 건설 객체의 위치 및 자세(posture)가 포함된 이진 이미지의 미래 프레임을 생성한다. 여기서, 예측 모델은 딥러닝의 일종인 ConvLSTM(Convolutional Long Short-Term Memory)을 사용할 수 있다.
- [0055] 도 4를 참조하여, 미리 학습된 예측 모델을 구축하는 방법을 설명한다. 도 4는 ConvLSTM 학습을 위한 학습 데이터 생성방법을 설명하기 위한 도면으로, 드론으로 획득한 건설 현장의 영상 데이터(비디오)로부터 영상 프레임(비디오 프레임)으로 구성된 시퀀스들의 집합인 ConvLSTM의 학습 데이터를 생성하는 과정을 보여준다.
- [0056] 하나의 시퀀스는 네트워크가 건설 객체의 변화 패턴을 이해하는 단위를 의미하며, 연속적인 n개의 이진 이미지로 구성된다. 이진 이미지는 객체 세분화 모듈(151)에서 영상 프레임으로부터 작업자, 굴삭기 및 덤프 트럭을 각각 인식한 결과이다. 네트워크가 t 시점 후의 프레임을 예측하도록 학습시키기 위해, n개의 연속적인 이진 이미지로 구성된 시퀀스의 레이블로 각 이미지의 t 시점 후의 이미지들로 구성된 새로운 시퀀스가 활용된다. 이러한 방식으로 3 종류의 건설 객체를 인식한 이진 이미지들을 활용하여 학습 데이터와 레이블이 생성된다. 따라서, 전체 비디오 프레임 개수가 N개일 때 총 시퀀스의 개수는 $3*(N-n-t+1)$ 개가 된다. 여기서, 3은 건설 객체의 수를 의미하는 것으로, 이는 예시일 뿐, 건설 현장의 건설 객체 수를 m이라 할 때, 총 시퀀스의 개수는 $m*(N-n-t+1)$ 개가 된다. 학습 데이터와 그 레이블은 시퀀스의 개수, 한 시퀀스 당 프레임의 개수, 이미지의 폭, 이미지의 높이 및 이미지의 채널 등의 차원을 가진 5차원 텐서(tensor)이다. ConvLSTM은 같은 비디오 프레임으

로부터 학습 시퀀스와 그 레이블을 생성하기 때문에 별도의 레이블링 작업이 필요없는 비지도 학습에 의존한다.

[0057] ConvLSTM 네트워크 설계 시 주의할 점은 커널 사이즈가 입력 이미지에서 하나의 타임 유닛 당 특정 객체가 이동하는 픽셀보다 커야 한다는 것이다. 이는 컨볼루션 연산의 특성 상 특정 오브젝트의 변화량이 커널 사이즈보다 클 때, 이전 셀에서 연관성이 있는 정보를 얻어올 수 없기 때문이다. 예를 들어, 비디오에서 t-1번째 프레임의 건설 장비가 t번째 프레임에서 커널 사이즈보다 더 큰 픽셀값으로 이동했다면 t번째 프레임의 같은 건설 장비에 대한 정보와 t-1번째 프레임의 정보를 연관시킬 수 있는 근거가 없다. 여러 연구들은 이러한 커널 사이즈 설정의 한계를 해결하기 위해 ConvLSTM layer의 stacking을 시도하였다. 이러한 맥락에서 목표 수행에 최적화된 레이어 수와 커널 사이즈를 가진 새로운 네트워크를 개발하는 것이 필요하다. 하나의 타임 유닛 당 객체가 이동하는 이미지 픽셀의 수가 건설 현장의 상황 및 객체의 특성에 따라 다를 수 있기 때문에, 이를 대표할 수 있는 네트워크가 설계되어야 한다. [표 1]은 4가지의 컨볼루션 필터 크기와 4가지의 ConvLSTM layer에 따른 16개의 네트워크 후보를 나타낸다. 건설 비디오에 대한 미래 프레임 예측을 위한 후보 아키텍처로 레이어 수와 필터 크기를 달리한 16개의 모델이 선택되었다.

표 1

Number of ConvLSTM layers	Network structure
1	N(3), N(5), N(7), N(9)
2	N(3)-N(3), N(5)-N(5), N(7)-N(7), N(9)-N(9)
3	N(3)-N(3)-N(3), N(5)-N(5)-N(5), N(7)-N(7)-N(7), N(9)-N(9)-N(9)
4	N(3)-N(3)-N(3)-N(3), N(5)-N(5)-N(5)-N(5), N(7)-N(7)-N(7)-N(7), N(9)-N(9)-N(9)-N(9)

N(n) = n×n의 filter size를 갖는 ConvLSTM

[0059]

[0061] 도 3의 B는, 객체 세분화 모듈(151)에 의해 생성된 굴삭기 이미지 시퀀스, 트럭 이미지 시퀀스, 작업자 이미지 시퀀스를 미리 학습된 ConvLSTM에 입력하여 생성된 t초 후의 굴삭기의 미래 프레임, 트럭의 미래 프레임, 작업자의 미래 프레임을 예시하고 있다. 생성된 미래 프레임은 이진 이미지 형태이다.

[0063] 근접도 모니터링 모듈(153)은 미래 프레임 예측 모듈(152)에 의해 생성된 미래 프레임과, 시간 순으로 정렬한 이진 이미지 시퀀스(즉, 예측 모델에 입력된 이미지 시퀀스) 중 적어도 어느 하나의 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출한다. 여기서, 근접도 도출을 위해 미래 프레임과 비교되는 이진 이미지는 이진 이미지 시퀀스를 구성하는 이진 이미지 중에서 가장 최근의 이진 이미지인 것이 바람직하다. 또한, 근접도 모니터링 모듈(153)은 도출된 근접도를 이용하여 건설 객체간의 충돌 위험을 알리는 안전정보를 생성한다.

[0064] 도 5를 참조하여 근접도를 산출하고 안전정보를 생성하는 과정을 설명한다. 도 5는 본 발명의 일 실시예에 따른 근접도 모니터링 모듈(153)에서 근접도를 산출하여 안전정보를 생성하는 과정을 설명하기 위한 도면이다.

[0065] 도 5에 도시된 바와 같이, 근접도 모니터링 모듈(153)은 코너 검출유닛(153a)과 근접도 측정유닛(153b)과 안전정보 생성유닛(153c)을 포함한다.

[0067] 객체 세분화 모듈(151)에 의해 생성된 건설 객체의 이진 이미지에서 건설 장비에 해당하는 건설 객체는 다각형 형태로 인식될 수 있다. 다각형 형태의 여러 건설 객체 간의 최단 거리를 측정하기 위해, 코너 검출유닛(153a)은 미래 프레임과, 미래 프레임과 비교될 이진 이미지(이하, “비교 이진 이미지”라 함)에 포함된 건설 객체의 코너를 검출한다. 코너 검출유닛(153a)은 Harris 코너 검출 알고리즘을 포함하며, Harris 코너 검출 알고리즘을 기반으로 다각형의 꼭지점을 검출한다. 건설 객체의 하나인 건설 현장의 인부는 건설 장비에 비해 상대적인 크기가 작으므로 계산의 편의를 위해 코너를 검출하지 않고 점으로 간주한다.

[0068] 한편, 미래 프레임 예측 모듈(152)에서 생성된 미래 프레임은 건설 객체와 무관한 픽셀들을 포함할 수 있으므로, 코너 검출유닛(153a)은 가우스 필터를 적용하여 입력 이미지(미래 프레임과 비교 이진 이미지)의 노이즈를 제거한다. 건설 객체의 상대적인 크기에 따라 적용될 가우스 필터의 크기는 가변할 수 있다. 예를 들어, 작업자의 크기는 건설 장비에 비해 상대적으로 작기 때문에 작업자를 인식한 이진 이미지에는 크기가 3,3인 가

우스 필터를 적용하고, 건설 장비인 트럭과 굴삭기 이미지는 7,7 크기의 가우시안 필터를 적용할 수 있다.

[0069] 노이즈를 제거한 다음, 윈도우를 모든 방향으로 이동시키며 픽셀 강도의 변화를 측정한다. Harris 코너 검출 알고리즘은 모든 방향의 픽셀 강도 변화가 큰 지점을 코너로 판단한다. 본 발명에서 Harris 코너 검출 알고리즘에 입력되는 입력 이미지는 0과 1의 값으로 구성된 이진 이미지를 처리하기 때문에, 픽셀 강도의 변화값과 방향에 관계없이 변화가 발생하는 모든 지점을 다각형의 코너로 간주할 수 있다.

[0071] 3차원의 실제 건설 현장을 드론에 장착된 카메라가 2차원의 영상 프레임(비디오 프레임)으로 매핑하는 과정에서 투사 왜곡이 발생하므로, 영상 데이터 획득 과정에서 정사 투영을 수행하여 투사 왜곡을 최소화하는 것이 바람직하다. 정사 투영에 의해 건설 객체의 높이 정보를 손실한다는 단점은 있으나, z축 변위를 제외한 x축 및 y축 방향의 변위를 측정할 수 있다. 본 발명에서는 2차원 영상 데이터로부터 측정되는 높이를 제외한 두 건설 객체 간의 x 및 y축 변위를 이용하여 근접도를 산출한다.

[0072] 근접도 측정유닛(153b)은 코너 검출유닛(153a)의 한 쌍의 결과 이미지에서 건설 객체의 미래 위치 및 자세를 기반으로 건설 객체 간의 근접도를 산출한다. 상기 결과 이미지는 미래 프레임에서 검출한 코너와 비교 이진 이미지에서 검출한 코너가 포함된 이미지이다. 근접도 측정유닛(153b)은 어느 한 건설 객체의 모든 코너점으로부터 다른 건설 객체의 모든 코너점까지의 각각의 직선 거리를 계산하여, 그 중 가장 작은 최소 거리가 두 건설 객체 간의 근접도로 결정된다. 이때, 계산의 복잡성을 줄이기 위해 상대적인 크기가 작은 건설 현장의 인부는 점으로 간주한다. 즉, 인부는 코너 검출을 적용하는 대신 중심점을 계산하여 근접도 결정에 활용한다. 실제 근접도를 측정하기 위해 근접도 측정유닛(153b)은 크기가 이미 알려진 참조 건설 객체를 활용한다.

[0073] 이와 같이, 한 쌍의 결과 이미지로 건설 객체 간의 근접도를 측정하여 픽셀 단위의 근접도를 추출하고, 픽셀 단위의 근접도를 하기 식 (1)을 이용하여 실제 크기로 스케일링하여 실제 크기 단위의 근접도를 산출한다.

[0075] 식 (1) : $근접도_{실제} = scale * 근접도_{이미지}$ ($scale = 근접도_{실제} / 근접도_{이미지}$)

[0077] 안전정보 생성유닛(153c)은 미래의 예측되는 근접도를 작업자에게 효과적으로 전달하기 위한 직관적인 안전정보를 생성한다. 안전정보 생성유닛(153c)은 현재 시점의 근접도 정보와 미래 프레임 예측 모듈(152)에 의해 예측된 미래 프레임을 이용하여 근접도 측정유닛(153b)에 의해 산출된 미래 시점의 근접도 정보를 비교하여, 작업자가 직관적으로 파악할 수 있는 다른 건설 객체가 작업자에게 다가오는 방향 및 속도를 계산한다.

[0078] 현재 시점의 근접도 정보는 드론(10)으로부터 송출되는 현재의 영상 데이터에 포함된 영상 프레임으로부터 객체 세분화 모듈(151)을 통해 파악되는 건설 객체의 위치 및 자세를 활용하여 코너 검출 및 근접도 측정을 수행한 결과로써 얻어진다. 즉, 현재 시점의 근접도 정보는 객체 세분화 모듈(151)을 거친 이진 이미지를 코너 검출유닛(153a) 및 근접도 측정유닛(153b)에 순차적으로 입력하여 산출되고, 미래 시점의 근접도 정보는 객체 세분화 모듈(151) 및 미래 프레임 예측 모듈(152)을 거친 이진 이미지를 코너 검출유닛(153a) 및 근접도 측정유닛(153b)에 순차적으로 입력하여 산출된다.

[0079] 미래 프레임 예측 모듈(152)을 거치지 않은 현재 프레임과 미래 프레임 예측 모듈(152)을 거친 미래 프레임 간의 시차와 근접도의 차이, 최소 거리에 해당하는 두 지점 사이의 각도를 활용하여 다른 건설 객체가 작업자에게 다가오는 방향 및 속도를 식 (2) 및 식 (3)과 같이 계산할 수 있다.

$$Speed = \frac{Proximity_{future} - Proximity_{current}}{t}$$

[0081] 식 (2) :

$$Heading\ direction = \arctan\left(\frac{Proximity_{future}(y)}{Proximity_{future}(x)}\right)$$

[0082] 식 (3) :

[0084] 도 6은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치에 의해 생성된 안전정보의 일 예를 보여주는 도면으로, 도 6의 좌측은 현재 근접도 정보이고, 우측은 미래 근접도 정보이며, 도면 아래 표시된 표는 각각의 작업자와 굴삭기 간의 현재 근접도, 미래 근접도, 접근 속도, 각도 등을 보여 준다.

[0085] 도 6의 표에 의하면, 작업자 1은 굴삭기가 멀어지므로 안전한 상태로 추정할 수 있고, 작업자 3과 작업자 4는 굴삭기가 접근하는 속도가 크므로 위험한 상태로 추정할 수 있다. 작업자 2는 굴삭기와의 거리는 가까우나 작업자 3과 작업자 4에 비해 굴삭기의 상대적인 접근 속도는 작으므로 준위험 상태로 추정할 수 있다. 위험(준위험)

또는 안전 여부는 작업자와 건설 장비인 건설 객체 간의 최소 거리(미래 근접도), 접근 속도를 고려하여 복합적으로 판단할 수 있으며, 이는 설계시에 해당 건설 현장의 특수성을 고려하여 설정될 수 있다.

- [0087] 다음으로, 도 7을 참조하여 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 방법을 설명한다. 도 7은 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 방법이 도시된 순서도이다.
- [0088] 도 7에 도시된 바와 같이, 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 방법은, 건설 현장의 건설 객체의 이진 이미지 시퀀스를 학습 데이터로 하여 미리 학습된 예측 모델을 구비하는 모니터링부(150)에 의해 수행되며, 통신부(110)를 통해 수신된 건설 현장을 촬영한 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하여 이를 이진 이미지로 변환하는 제1 단계(S100)와, 상기 변환된 이진 이미지를 상기 예측 모델에 입력하여 미래 프레임을 생성하는 제2 단계(S200)와, 상기 생성된 미래 프레임과 상기 변환된 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출하는 제3 단계(S300)와, 상기 도출된 근접도를 이용하여 안전정보를 생성하는 제4 단계(S400)를 포함한다.
- [0089] 제1 단계(S100)는 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출하고, 검출된 복수개의 건설 객체의 수만큼 복수개의 이진 이미지를 생성하되, 각각의 이진 이미지는 단일 건설 객체를 포함하도록 한다. 제1 단계(S100)는 Mask R-CNN을 이용하여 수신된 영상 데이터에 포함된 영상 프레임에서 복수개의 건설 객체를 검출한다.
- [0090] 제2 단계(S200)는 생성된 복수개의 이진 이미지 중에서 동일한 건설 객체를 포함하는 복수개의 이진 이미지를 시간 순으로 정렬한 이진 이미지 시퀀스를 예측 모델의 입력값으로 하여 미래 프레임을 생성한다. 예측 모델은 convolutional LSTM Network를 포함한다.
- [0091] 제3 단계(S300)는 제2 단계(S200)에서 생성된 미래 프레임과, 시간 순으로 정렬한 이진 이미지 시퀀스 중 적어도 어느 하나의 이진 이미지를 비교하여 건설 객체 간의 근접도를 도출한다. 제3 단계(S300)는 이진 이미지에 포함된 건설 객체의 코너를 검출하는 Harris 코너 검출 알고리즘을 이용하고, 어느 한 건설 객체의 모든 코너점으로부터 다른 건설 객체의 모든 코너점까지의 각각의 직선 거리를 계산하여, 그 중 가장 작은 최소 거리를 두 건설 객체간의 근접도로 결정한다.
- [0092] 제4 단계(S400)는 현재 시점의 근접도와 제3 단계(S300)에 의해 도출된 근접도를 비교하여 수행한다. 현재 시점의 근접도 정보는 드론(10)으로부터 송출되는 현재의 영상 데이터에 포함된 영상 프레임으로부터 파악되는 건설 객체의 위치 및 자세를 활용하여 코너 검출 및 근접도 측정을 수행한 결과로써 얻어진다. 즉, 현재 시점의 근접도 정보는 제1 단계(S100)를 거친 이진 이미지를 제3 단계(S300)에 적용하여 산출되고, 미래 시점의 근접도 정보는 제1 단계(S100) 및 제2 단계(S200)를 거친 이진 이미지를 제3 단계(S300)에 적용하여 산출된다.
- [0094] 도 8은 본 발명의 실시예에 따른 컴퓨팅 장치를 나타내는 도면이다. 도 8의 컴퓨팅 장치(TN100)는 본 명세서에서 기술된 건설 객체 실시간 모니터링 장치일 수 있다.
- [0095] 도 8의 실시예에서, 컴퓨팅 장치(TN100)는 적어도 하나의 프로세서(TN110), 송수신 장치(TN120), 및 메모리(TN130)를 포함할 수 있다. 또한, 컴퓨팅 장치(TN100)는 저장 장치(TN140), 입력 인터페이스 장치(TN150), 출력 인터페이스 장치(TN160) 등을 더 포함할 수 있다. 컴퓨팅 장치(TN100)에 포함된 구성 요소들은 버스(bus)(TN170)에 의해 연결되어 서로 통신을 수행할 수 있다.
- [0096] 프로세서(TN110)는 메모리(TN130) 및 저장 장치(TN140) 중에서 적어도 하나에 저장된 프로그램 명령(program command)을 실행할 수 있다. 프로세서(TN110)는 중앙 처리 장치(CPU: central processing unit), 그래픽 처리 장치(GPU: graphics processing unit), 또는 본 발명의 실시예에 따른 방법들이 수행되는 전용의 프로세서를 의미할 수 있다. 프로세서(TN110)는 본 발명의 실시예와 관련하여 기술된 절차, 기능, 및 방법 등을 구현하도록 구성될 수 있다. 프로세서(TN110)는 컴퓨팅 장치(TN100)의 각 구성 요소를 제어할 수 있다.
- [0097] 메모리(TN130) 및 저장 장치(TN140) 각각은 프로세서(TN110)의 동작과 관련된 다양한 정보를 저장할 수 있다. 메모리(TN130) 및 저장 장치(TN140) 각각은 휘발성 저장 매체 및 비휘발성 저장 매체 중에서 적어도 하나로 구성될 수 있다. 예를 들어, 메모리(TN130)는 읽기 전용 메모리(ROM: read only memory) 및 랜덤 액세스 메모리(RAM: random access memory) 중에서 적어도 하나로 구성될 수 있다.
- [0098] 송수신 장치(TN120)는 유선 신호 또는 무선 신호를 송신 또는 수신할 수 있다. 송수신 장치(TN120)는 네트워크에 연결되어 통신을 수행할 수 있다.
- [0100] 한편, 본 발명은 컴퓨터프로그램으로 구현될 수도 있다. 본 발명은 하드웨어와 결합되어, 본 발명에 따른 건설

객체 실시간 모니터링 방법을 실행시키기 위하여, 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.

[0101] 본 발명의 실시예에 따른 방법들은 다양한 컴퓨터수단을 통하여 판독 가능한 프로그램 형태로 구현되어 컴퓨터로 판독 가능한 기록매체에 기록될 수 있다. 여기서, 기록매체는 프로그램 명령, 데이터 파일, 데이터구조 등을 단독으로 또는 조합하여 포함할 수 있다.

[0102] 기록매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다.

[0103] 예컨대 기록매체는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CDROM, DVD와 같은 광 기록 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 장치를 포함한다.

[0104] 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어를 포함할 수 있다.

[0105] 이러한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0107] 다음으로, 본 발명의 일 실시예에 따른 건설 객체 실시간 모니터링 장치의 모니터링부(150)를 구성하는 객체 세분화 모듈(151), 미래 프레임 예측 모듈(152), 근접도 모니터링 모듈(153)에 대한 성능 검증 결과를 설명한다.

[0109] 1. 객체 세분화 모듈

[0110] 객체 세분화 모듈의 성능 검증을 위해, 건설 현장을 촬영한 드론 영상 데이터로부터 1180개의 프레임을 추출하였다. VGG Image Annotator를 사용하여 건설 현장의 좌표에 픽셀 수준으로 수동으로 주석을 달았다.

[0111] Mask R-CNN의 성능을 평가하기 위해, 1180장의 이미지 중 860장의 이미지를 학습 데이터 세트로 사용하고, 나머지 320장의 이미지를 validation 데이터 세트로 사용하였다.

[0112] [표 2]는 객체 세분화의 클래스 별 AP(average precision), mAP(mean average precision) 및 평균 IOU를 보여준다. IoU는 기준 마스크와 예측 마스크 사이의 퍼센트 오버랩을 정량화한 메트릭으로 픽셀 단위의 객체 인식 정확도를 나타낸다.

표 2

Average precision			mAP	IoU
Excavator	Dump truck	Worker		
99.77%	95.05%	89.88%	94.32%	85.79%

[0114]

[0116] 평균 IOU는 테스트 데이터셋의 모든 검출 결과에서 획득되는 IoU 값을 평균한 것이다. 평균 정밀도는 IoU값에 따른 인식 성능의 변화를 하나의 값으로 표현한다. 본 명세서에서는 COCO (Common Objects in Context) 챌린지의 객체 세분화 분야에서 정밀도를 평가하는 방식과 동일하게 0.5 ~ 0.95의 IoU 임계 값 범위에서 정밀도를 계산하였다. mAP는 물체 클래스가 여러 개일 때 각 클래스당 AP의 평균을 구하여 네트워크의 성능을 하나의 메트릭으로 표현한다. 실험 결과는 Mask R-CNN이 충분한 정확도로 건설 객체를 인식한다는 것을 증명한다.

[0118] 2. 미래 프레임 예측 모듈

[0119] 미래 프레임 예측 모듈의 성능 검증을 위해, 총 9개의 영상 데이터로부터 760장의 연속적인 영상 프레임을 1초 단위로 추출하였다. 760장의 영상 프레임은 전술한 1180장의 이미지와 완전히 독립적이다. 학습된 Mask R-CNN을 통해 760장의 영상 프레임은 각 영상 프레임으로부터 3가지 객체를 각각 인식한 2,280장의 이진 이미지를 생성하였다. 이 중 7개의 영상 데이터에서 추출한 694개의 영상 프레임(2082장의 이진 이미지)은 학습 데이터로 활용되었고, 2개의 영상 데이터에서 추출한 66개의 영상 프레임(198장의 이진 이미지) test 데이터로 활용되었다.

[0120] 본 명세서에서 데이터 한 개는 하나의 영상 데이터로부터 추출된 영상 프레임들로 생성된 시퀀스들의 집합으로 정의될 수 있다. 이 때, 1개의 시퀀스는 5개의 연속적인 이진 이미지로 구성될 수 있다. 하지만, 네트워크는 학습 데이터 개수가 7개 뿐이므로 다양한 패턴의 건설 객체의 맥락을 포함한다고 하기는 어렵다. 상대적으로 적은 데이터로 ConvLSTM의 최적화된 stack 수와 filter를 찾기 위해 본 명세서에서는 Leave-One-Out 교차 검증 방법(LOOCV)을 활용하였다. 이 방법은 샘플의 수 만큼 모델을 만들고 각 모델을 학습할 때 하나의 샘플만 제외하면서 제외된 하나의 샘플로 성능을 측정하여 이들을 평균내는 방법이다. Leave-one-out 교차 검증 방법은 모든 샘플에 대해서 한번씩 test하기 때문에 randomness을 완벽하게 배제할 수 있을 뿐만 아니라 특정 데이터셋에 편중되는 모델의 overfitting을 극복할 수 있다.

[0121] [표 1]의 16개 모델 후보 중에서 가장 높은 평균 퍼포먼스를 갖는 ConvLSTM의 stack 수와 filter가 선택된다. 각 모델의 성능은 recall과 precision의 조화평균인 F-measure를 활용하여 평가되었다. recall은 올바르게 예측된 predicted mask의 픽셀 수에서 reference mask의 픽셀 수로 나눈 indicator이며, precision은 올바르게 예측된 predicted mask의 픽셀 수에서 predicted mask의 픽셀 수로 나눈 것이다.

[0122] [표 3]은 16개의 네트워크 후보에 대해 LOOCV를 수행한 결과를 with F-measure을 보여준다. 이는 9×9 필터를 가진 1개의 ConvLSTM이 본 데이터셋에 최적화된 모델임을 논증한다. 선택된 모델을 기반으로 2개의 테스트 데이터를 활용하여 본 방법론의 성능이 실험되었다. convLSTM 네트워크는 하나의 미래 프레임에 대해 recall, precision은 각각 81.24%, 79.95%의 픽셀 단위 정확도로 예측하였다.

표 3

		Number of stack			
		1	2	3	4
	3×3	80.06%	79.80%	80.25%	80.20%
Filter	5×5	80.72%	81.36%	81.29%	80.94%
size	7×7	81.69%	81.53%	80.96%	81.10%
	9×9	82.00%	81.71%	80.98%	81.21%

[0124]

[0126] 3. 근접도 모니터링 모듈

[0127] 근접도 모니터링 모듈을 검증하기 위해, 실제 객체 간 거리와 예측된 객체 간 거리 간의 오차를 측정하였다. 계산된 픽셀 단위의 proximity는 식 (1)을 기반으로 실제 metric 단위의 proximity로 환산되었다. 식 (1)의 참조 객체의 실제 길이는 5.205m, 이미지 내 차지하는 픽셀의 크기는 92 pixel로 scene scale는 약 0.05658m/pixel로 계산되었다.

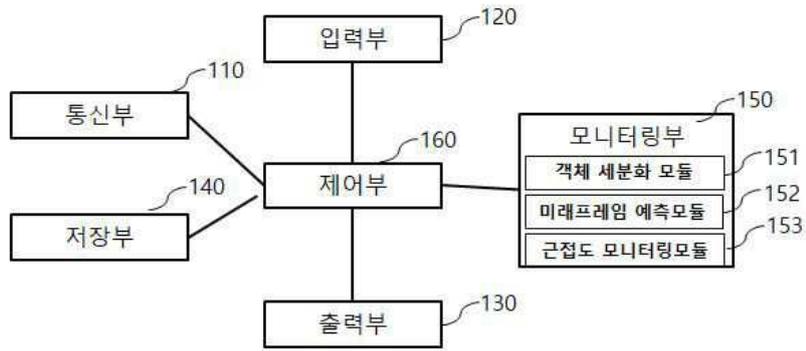
[0128] 이를 기반으로, 66개의 테스트 프레임을 대상으로 평균 proximity error를 측정하였다. [표 4]는 미래 프레임 예측 모듈의 time-step에 따른 proximity의 평균 error를 보여준다. 식 (2)를 활용하여 현재 프레임과 미래 프레임의 근접도(proximity) 정보를 비교하여 선제적인 안전 정보인 접근 속도(heading speed)와 접근 방향(heading direction)을 제공할 수 있다.

표 4

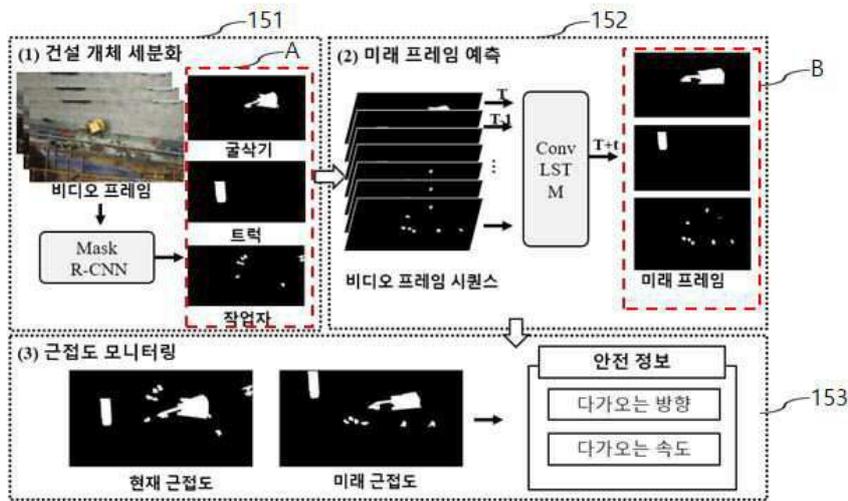
		Pixel error (pixel)	Actual error (m)
Time-step of frame prediction (s)	1	7.54	0.43
	2	16.11	0.91
	3	21.61	1.22

[0130]

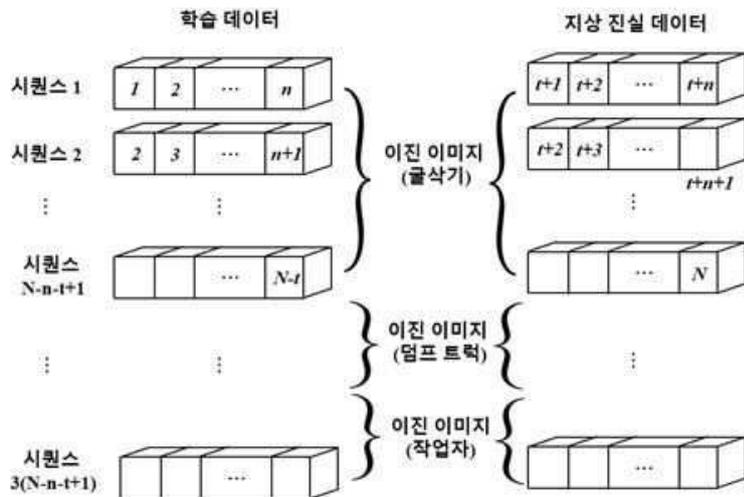
도면2



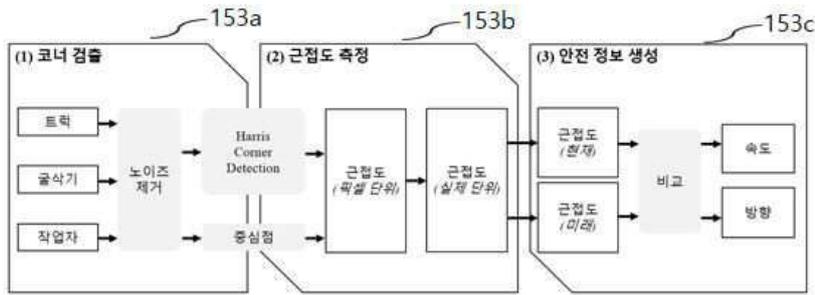
도면3



도면4



도면5

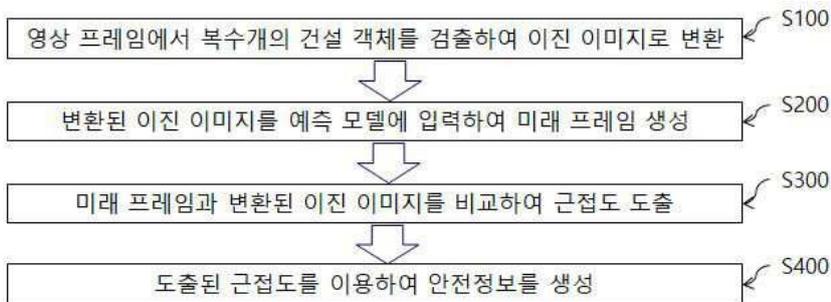


도면6



작업자	현재 근접도		미래 근접도	
	작업자와 굴삭기의 현재 근접도 (m)	작업자와 굴삭기의 미래 근접도 (m)	굴삭기와 작업자의 미래 각도 (°)	굴삭기의 작업자로 접근하는 속도 (m/s)
1	2.68	2.73	343.1	-0.05
2	4.18	2.23	225.00	1.95
3	5.14	2.98	120.96	2.16
4	3.86	1.76	236.82	2.10

도면7



도면8

