

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2021-0111590
(43) 공개일자 2021년09월13일

(51) 국제특허분류(Int. Cl.)

G06K 9/00 (2006.01) G06K 9/32 (2006.01)

G06K 9/46 (2006.01) G06K 9/62 (2006.01)

G06N 20/00 (2019.01)

(52) CPC특허분류

G06K 9/00362 (2013.01)

G06K 9/3233 (2013.01)

(21) 출원번호 10-2020-0026714

(22) 출원일자 2020년03월03일

심사청구일자 2020년03월03일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

손광훈

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)

김민수

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)

(74) 대리인

민영준

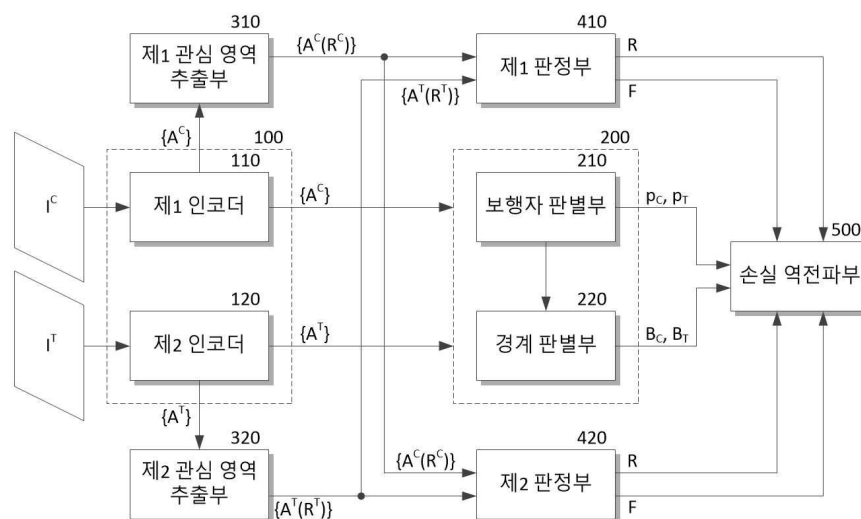
전체 청구항 수 : 총 18 항

(54) 발명의 명칭 적대적 학습을 통한 비쌍 교차 스펙트럴 보행자 검출 장치 및 방법

(57) 요약

본 발명은 서로 매칭되도록 조절되지 않은 컬러 영상과 열 영상인 비쌍 교차 스펙트럴 영상으로부터 적대적 학습 기법을 이용하여 학습이 수행되므로, 컬러 영상과 열 영상을 정렬시키기 위한 별도의 소프트웨어나 하드웨어를 요구하지 않고도 공통 인공 신경망이 용이하게 학습되어 컬러 영상 또는 열 영상으로부터 정확하게 보행자를 검출할 수 있을 뿐만 아니라 컬러 영상 또는 열 영상 중 하나만 인가되는 경우에도 보행자를 검출할 수 있는 보행자 검출 장치를 제공할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06K 9/4652 (2013.01)

G06K 9/627 (2013.01)

G06N 20/00 (2021.08)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711094090

부처명 과학기술정보통신부

과제관리(전문)기관명 한국연구재단

연구사업명 복합인지기술개발사업

연구과제명 (2세부)이종 CCTV 영상에서의 딥러닝 기반 실종자 초동 신원확인 및 추적 시스템
(1단계)(2/2)

기 여 율 1/1

과제수행기관명 연세대학교 산학협력단

연구기간 2019.04.23 ~ 2020.04.22

명세서

청구범위

청구항 1

컬러 영상과 열 영상 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 컬러 특징맵과 열 특징맵을 획득하는 인코더부;

상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어와 보행자 영역을 나타내는 경계 박스를 추출하는 통합 검출부; 및

학습 시에 결합되어 상기 인코더부가 보행자의 위치 및 크기에 대응하여 경계 박스가 미리 설정된 학습 영상인 학습용 컬러 영상과 학습용 열 영상에 대해 추출한 컬러 특징맵과 열 특징맵에서 설정된 경계 박스에 대응하는 관심 영역을 추출하고, 추출된 관심 영역에 대해 적대적 판정을 수행하며, 분류 스코어와 보행자 영역 그리고 적대적 판정 결과를 기초로 분류 손실과 영역 손실 및 적대적 손실을 계산하여 상기 분류 손실과 상기 영역 손실을 상기 통합 검출부로 역전파하여 학습시킴과 동시에 상기 적대적 손실에 따라 적대적 학습이 수행되는 학습부를 포함하는 보행자 검출 장치.

청구항 2

제1 항에 있어서, 상기 통합 검출부는

상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어를 추정하는 보행자 판별부; 및

상기 보행자 판별부와 동일하게 상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식과 상기 보행자 판별부에서 판별된 분류 스코어에 따라 서로 다른 크기 및 형태의 기지정된 개수의 경계 박스 중 보행자의 형상에 따른 보행자 영역을 나타내는 경계 박스를 추출하는 경계 판별부를 포함하는 보행자 검출 장치.

청구항 3

제1 항에 있어서, 상기 학습부는

상기 인코더부가 상기 학습 영상으로 인가된 상기 학습용 컬러 영상 또는 상기 학습용 열 영상 중 적어도 하나에 대해 추출한 컬러 특징맵과 열 특징맵에서 관심 영역을 추출하는 관심 영역 추출부;

상기 적대적 손실에 따라 학습되는 2개의 인공 신경망을 포함하여 상기 컬러 특징맵과 상기 열 특징맵 각각에서 추출된 관심 영역에 대해 리얼 또는 페이크로 판별되도록 서로 상반되도록 판별하는 적대적 판정부; 및

상기 학습 영상에 레이블된 보행자가 존재하는 영역의 개수와 존재하지 않는 영역의 개수 사이의 비와 보행자가 존재하는 영역에 대한 분류 스코어에 따라 상기 분류 손실을 계산하고, 상기 학습용 컬러 영상 또는 상기 학습용 열 영상에 미리 설정된 경계 박스와 상기 통합 검출부에서 추출된 경계 박스 사이의 차이를 상기 영역 손실로 계산하며, 상기 적대적 판정부의 2개의 인공 신경망 각각에 대한 적대적 손실을 개별적으로 계산하여 역전파하는 손실 역전파부를 포함하는 보행자 검출 장치.

청구항 4

제3 항에 있어서, 상기 적대적 판정부는

인공 신경망으로 구현되어 상기 열 특징맵에서 추출된 관심 영역이 인가되면 리얼로 판정하고, 상기 컬러 특징맵에서 추출된 관심 영역이 인가되면 페이크로 판정하도록 학습되는 제1 판정부; 및

인공 신경망으로 구현되어 상기 컬러 특징맵에서 추출된 관심 영역이 인가되면 리얼로 판정하고, 상기 열 특징맵에서 추출된 관심 영역이 인가되면 페이크로 판정하도록 학습되는 제2 판정부를 포함하는 보행자 검출 장치.

청구항 5

제4 항에 있어서, 상기 손실 역전파부는
상기 분류 손실(L_{cls})을 수학식

$$\mathcal{L}_{cls}(p) = \begin{cases} (1/\alpha_t)(1-p^y)\log(p) & \text{if } y = 1 \\ \alpha_t(p)^y\log(1-p) & \text{otherwise} \end{cases}$$

(여기서 α_t 는 학습 영상에 레이블된 진리값에 기반하여 판별되는 보행자가 존재하는 영역($y = 1$)의 개수($\Omega_{y=1}$)와 존재하지 않는 영역($y = -1$)사이의 개수($\Omega_{y=-1}$) 비로서, $\alpha_t = |\Omega_{y=1}|/|\Omega_{y=-1}|$ 로 계산되며, p 는 보행자가 존재하는 영역($y = 1$)에 대한 분류 스코어이며, y 는 미리 지정되는 초점 매개변수(focus parameter)이다.)

에 따라 계산하는 보행자 검출 장치.

청구항 6

제5 항에 있어서, 상기 손실 역전파부는
상기 영역 손실(L_{loc})을 수학식

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

의 평활 $L1$ 함수(smooth_{L1})를 적용하여 수학식

$$\mathcal{L}_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L1}(t_i^u - v_i)$$

(여기서 t^u 는 학습 영상에 레이블된 진리 경계 박스를 나타내고, v 는 통합 검출부에서 획득된 경계 박스를 나타내며, x, y, w, h 는 각각 획득된 경계 박스의 중심 위치와 폭 및 높이를 나타낸다.)

에 따라 계산하는 보행자 검출 장치.

청구항 7

제6 항에 있어서, 상기 손실 역전파부는

상기 제1 판정부에 대한 컬러 적대적 손실(L_{adv}^c)을 수학식

$$\begin{aligned} & \mathcal{L}_{adv}^c(P^C, P^T, D^C) \\ &= \mathbb{E}_{P^C \sim p_{data}(P^C)} \left[\log(D^C(P^C)) + \log(1 - D^C(P^T)) \right] \end{aligned}$$

에 따라 계산하고,

상기 제2 판정부에 대한 열 적대적 손실(L_{adv}^T)을 수학식

$$\begin{aligned} & \mathcal{L}_{adv}^T(P^C, P^T, D^T) \\ &= \mathbb{E}_{P^T \sim p_{data}(P^T)} \left[\log(D^T(P^T)) + \log(1 - D^T(P^C)) \right] \end{aligned}$$

(여기서 P^C 와 P^T 는 각각 학습용 컬러 영상에서 추출된 관심 영역과 학습용 열 영상에서 추출된 관심 영역이고, D^C 와 D^T 는 각각 제1 및 제2 판정부를 판정 함수로 나타낸 것이며, $\mathbb{E}_{P^C \sim p_{data}(P^C)}$ 와 $\mathbb{E}_{P^T \sim p_{data}(P^T)}$ 는 각각 P^C 와 P^T 에서 고정 길이 특징 벡터를 추출하기 위한 ROI 풀링을 나타낸다.)

에 따라 계산하는 보행자 검출 장치.

청구항 8

제7 항에 있어서, 상기 손실 역전파부는

상기 컬러 적대적 손실(L_{adv}^C)과 상기 열 적대적 손실(L_{adv}^T)을 가산하여 적대적 손실(L_{adv})을 계산하고,

상기 분류 손실(L_{cls})과 상기 영역 손실(L_{loc}) 및 상기 적대적 손실(L_{adv})로부터 총 손실(L_{total})을 수학식

$$\mathcal{L}_{total} = \sum_{i \in \Omega^C, \Omega^T} \mathcal{L}_{cls}^i + \sum_{j \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{loc}^j + \lambda \sum_{k \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{adv}^k$$

(여기서 λ 는 손실 밸런스를 조절하기 위한 매개변수이다.)

에 따라 계산하는 보행자 검출 장치.

청구항 9

제8 항에 있어서, 상기 손실 역전파부는

상기 총 손실(L_{total})의 수학식

$$\arg \min_{\{E^C, E^T, U\}} \max_{\{D^C, D^T\}} \mathcal{L}_{total}$$

(여기서 E^C , E^T 는 컬러 영상과 열 영상을 인코딩하는 인코더부의 함수 표현이고, U 는 통합 검출부에 대한 함수 표현이며, D^C , D^T 는 제1 및 제2 판정부에 대한 함수 표현이다.)

에 따라 최적화될 때까지 상기 분류 손실(L_{cls})과 상기 영역 손실(L_{loc})을 상기 통합 검출부로 역전파하고, 상기 컬러 적대적 손실(L_{adv}^C)을 상기 제1 판정부로 역전파하며, 상기 열 적대적 손실(L_{adv}^T)을 상기 제2 판정부로 역전파하여 반복 학습을 수행하는 보행자 검출 장치.

청구항 10

컬러 영상과 열 영상 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 컬러 특징맵과 열 특징맵을 획득하는 인코딩 단계;

상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어와 보행자 영역을 나타내는 경계 박스를 추출하는 보행자 검출 단계; 및

보행자의 위치 및 크기에 대응하여 경계 박스가 미리 설정된 학습 영상인 학습용 컬러 영상과 학습용 열 영상에 대해 추출한 컬러 특징맵과 열 특징맵에서 설정된 경계 박스에 대응하는 관심 영역을 추출하고, 추출된 관심 영

역에 대해 적대적 판정을 수행하며, 분류 스코어와 보행자 영역 그리고 적대적 판정 결과를 기초로 분류 손실과 영역 손실 및 적대적 손실을 계산하여 상기 분류 손실과 상기 영역 손실을 역전파하여 상기 적대적 손실에 따라 적대적 학습이 수행되는 학습 단계를 포함하는 보행자 검출 방법.

청구항 11

제10 항에 있어서, 상기 보행자 검출 단계는

상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어를 추정하는 단계; 및
상기 보행자 판별부와 동일하게 상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식과 상기 보행자 판별부에서 판별된 분류 스코어에 따라 서로 다른 크기 및 형태의 기지정된 개수의 경계 박스 중 보행자의 형상에 따른 보행자 영역을 나타내는 경계 박스를 추출하는 단계를 포함하는 보행자 검출 방법.

청구항 12

제10 항에 있어서, 상기 학습 단계는

상기 학습 영상으로 인가된 상기 학습용 컬러 영상 또는 상기 학습용 열 영상 중 적어도 하나를 인코딩하여 획득한 컬러 특징맵과 열 특징맵에서 관심 영역을 추출하는 단계;

상기 적대적 손실에 따라 학습되는 2개의 인공 신경망을 포함하여 상기 컬러 특징맵과 상기 열 특징맵 각각에서 추출된 관심 영역에 대해 리얼 또는 페이크로 판별하되 서로 상반되도록 판별하는 단계; 및

상기 학습 영상에 레이블된 보행자가 존재하는 영역의 개수와 존재하지 않는 영역의 개수 사이의 비와 보행자가 존재하는 영역에 대한 분류 스코어에 따라 상기 분류 손실을 계산하고, 상기 학습용 컬러 영상 또는 상기 학습용 열 영상에 미리 설정된 경계 박스와 상기 보행자 검출 단계에서 추출된 경계 박스 사이의 차이를 상기 영역 손실로 계산하며, 상기 2개의 인공 신경망 각각에 대한 적대적 손실을 개별적으로 계산하여 역전파하는 단계를 포함하는 보행자 검출 방법.

청구항 13

제12 항에 있어서, 상기 판별하는 단계는

인공 신경망을 이용하여 상기 열 특징맵에서 추출된 관심 영역이 인가되면 리얼로 판정하고, 상기 컬러 특징맵에서 추출된 관심 영역이 인가되면 페이크로 판정하도록 학습되는 단계; 및

인공 신경망으로 이용하여 상기 컬러 특징맵에서 추출된 관심 영역이 인가되면 리얼로 판정하고, 상기 열 특징맵에서 추출된 관심 영역이 인가되면 페이크로 판정하도록 학습되는 단계를 포함하는 보행자 검출 방법.

청구항 14

제13 항에 있어서, 상기 역전파하는 단계는

상기 분류 손실(L_{cls})을 수학식

$$\mathcal{L}_{cls}(p) = \begin{cases} (1/\alpha_t)(1 - p^r)\log(p) & \text{if } y = 1 \\ \alpha_t(p)^r\log(1 - p) & \text{otherwise} \end{cases}$$

(여기서 α_t 는 학습 영상에 레이블된 진리값에 기반하여 판별되는 보행자가 존재하는 영역($y = 1$)의 개수($\Omega_{y=1}$)와 존재하지 않는 영역($y = -1$)사이의 개수($\Omega_{y=-1}$) 비로서, $\alpha_t = |\Omega_{y=1}|/|\Omega_{y=-1}|$ 로 계산되며, p 는 보행자가 존재하는 영역($y = 1$)에 대한 분류 스코어이며, r 는 미리 지정되는 초점 매개변수(focus parameter)이다.)

에 따라 계산하는 보행자 검출 방법.

청구항 15

제14 항에 있어서, 상기 역전파하는 단계는

상기 영역 손실(L_{loc})을 수학식

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

의 평활 L1 함수($smooth_{L1}$)를 적용하여 수학식

$$\mathcal{L}_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L1}(t_i^u - v_i)$$

(여기서 t^u 는 학습 영상에 레이블된 진리 경계 박스를 나타내고, v 는 보행자 검출 단계에서 획득된 경계 박스를 나타내며, x, y, w, h 는 각각 획득된 경계 박스의 중심 위치와 폭 및 높이를 나타낸다.)

에 따라 계산하는 보행자 검출 방법.

청구항 16

제15 항에 있어서, 상기 역전파하는 단계는

상기 2개의 인공 신경망에 대한 적대적 손실을 컬러 적대적 손실(L_{adv}^C)과 열 적대적 손실(L_{adv}^T)로 구분하여 계산하며,

상기 컬러 적대적 손실(L_{adv}^C)을 수학식

$$\begin{aligned} \mathcal{L}_{adv}^C(P^C, P^T, D^C) \\ = \mathbb{E}_{P^C \sim p_{data}(P^C)} \left[\log(D^C(P^C)) + \log(1 - D^C(P^T)) \right] \end{aligned}$$

에 따라 계산하고,

상기 열 적대적 손실(L_{adv}^T)을 수학식

$$\begin{aligned} \mathcal{L}_{adv}^T(P^C, P^T, D^T) \\ = \mathbb{E}_{P^T \sim p_{data}(P^T)} \left[\log(D^T(P^T)) + \log(1 - D^T(P^C)) \right] \end{aligned}$$

(여기서 P^C 와 P^T 는 각각 학습용 컬러 영상에서 추출된 관심 영역과 학습용 열 영상에서 추출된 관심 영역이고, D^C 와 D^T 는 각각 2개의 인공 신경망의 동작을 판정 함수로 나타낸 것이며, $\mathbb{E}_{P^C \sim p_{data}(P^C)}$ 와 $\mathbb{E}_{P^T \sim p_{data}(P^T)}$ 는 각각 P^C 와 P^T 에서 고정 길이 특징 벡터를 추출하기 위한 ROI 풀링을 나타낸다.)

에 따라 계산하는 보행자 검출 방법.

청구항 17

제16 항에 있어서, 상기 역전파하는 단계는

상기 컬러 적대적 손실(L_{adv}^c)과 상기 열 적대적 손실(L_{adv}^T)을 가산하여 적대적 손실(L_{adv})을 계산하고,

상기 분류 손실(L_{cls})과 상기 영역 손실(L_{loc}) 및 상기 적대적 손실(L_{adv})로부터 총 손실(L_{total})을 수학적

$$\mathcal{L}_{total} = \sum_{i \in \Omega^C, \Omega^T} \mathcal{L}_{cls}^i + \sum_{j \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{loc}^j + \lambda \sum_{k \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{adv}^k$$

(여기서 λ 는 손실 밸런스를 조절하기 위한 매개변수이다.)

에 따라 계산하는 보행자 검출 방법.

청구항 18

제17 항에 있어서, 상기 역전파하는 단계는

상기 총 손실(L_{total})이 수학적

$$\arg \min_{\{E^C, E^T, U\}} \max_{\{D^C, D^T\}} \mathcal{L}_{total}$$

(여기서 E^C, E^T 는 컬러 영상과 열 영상을 인코딩하는 인코더부의 함수 표현이고, U 는 통합 검출부에 대한 함수 표현이며, D^C, D^T 는 제1 및 제2 판정부에 대한 함수 표현이다.)

에 따라 최적화될 때까지 상기 분류 손실(L_{cls})과 상기 영역 손실(L_{loc}), 상기 컬러 적대적 손실(L_{adv}^c) 및 상기 열 적대적 손실(L_{adv}^T)을 역전파하여 반복 학습을 수행하는 보행자 검출 방법.

발명의 설명

기술 분야

[0001] 본 발명은 보행자 검출 장치 및 방법에 관한 것으로, 적대적 학습을 통한 비쌍 교차 스펙트럴 보행자 검출 장치 및 방법에 관한 것이다.

배경 기술

[0002] 최근 지능형 교통 시스템(Intelligent Transportation System: ITS) 기술을 활용한 자율 주행(Autonomous Driving) 또는 운전자 보조 시스템(Advanced Driver Assistance System: ADAS)에 대한 관심이 증대되면서 보행자 검출 방법에 대한 연구가 활발하게 이루어지고 있다.

[0003] 보행자 검출은 영상을 획득하고, 획득된 영상에서 보행자의 위치를 판별하는 기술로서 기존에는 일반적인 카메라를 이용하여 획득된 컬러 영상만을 이용하여 보행자를 검출하였다. 그러나 컬러 영상만을 이용하여 보행자를 검출하는 경우, 야간 환경 등과 같이 조명이 충분하지 못한 환경에서 보행자를 정확하게 검출하지 못하는 문제가 있다.

[0004] 이에 일반 카메라에서 촬영되는 컬러 영상뿐만 아니라 주변 조명이 충분하지 않은 환경에서도 사람의 실루엣을 선명하게 포착할 수 있도록 장파장 적외선 카메라를 이용하여 획득되는 열영상을 사용하여 보행자를 검출하는 교차 스펙트럴(Cross-Spectral) 보행자 감지에 대한 연구가 활발하게 진행되고 있다.

[0005] 한편 인공 신경망의 발전으로 인해 현재 보행자 검출에 대한 연구는 주로 인공 신경망을 이용하는 방법으로 이루어지고 있다. 그러나 기존에 인공 신경망을 이용하는 보행자 검출 기법에서는 컬러 영상과 열 영상 각각에 대응하는 별도의 인공 신경망을 이용하여 보행자를 검출하며, 이 경우 검출 결과가 서로 상이하게 나타날 수 있다는 문제가 있다.

[0006] 이에 컬러 영상과 열 영상에 대해 공통 인공 신경망을 이용하는 경우, 공통 인공 신경망이 컬러 영상과 열 영상 모두에서 정확하게 보행자를 검출할 수 있어야 한다. 따라서 공통의 인공 신경망을 학습시키기 위해서는 컬러 영상과 열 영상이 정확하게 매칭되어 쌍을 이루는 학습 데이터가 대량으로 필요하다. 그러나 컬러 영상과 열

영상이 쌍을 이루도록 매칭시키기 위해서는 컬러 영상과 열 영상 사이의 해상도와 화각 등이 차이를 보정하기 위한 별도의 소프트웨어나 하드웨어가 필요하다. 뿐만 아니라 이러한 보정으로도 컬러 영상과 열 영상 사이를 정확하게 정렬시키기 어렵기 때문에 현실적으로 대량의 학습 데이터를 획득하는 것이 매우 어렵다는 한계가 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국 공개 특허 제10-2018-0125278호 (2018.11.23 공개)

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 서로 매칭되도록 조절되지 않은 컬러 영상과 열 영상인 비쌍 교차 스펙트럴 영상으로부터 보행자를 검출하기 위한 공통 인공 신경망을 정확하게 학습시킬 수 있는 보행자 검출 장치 및 방법을 제공하는데 있다.

[0009] 본 발명의 다른 목적은 컬러 영상과 열 영상을 정렬시키기 위한 별도의 소프트웨어나 하드웨어를 요구하지 않고 비쌍 교차 스펙트럴 영상을 기반으로 학습되어 정확하게 보행자를 검출할 수 있는 보행자 검출 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 보행자 검출 장치는 컬러 영상과 열 영상 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 컬러 특징맵과 열 특징맵을 획득하는 인코더부; 상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어와 보행자 영역을 나타내는 경계 박스를 추출하는 통합 검출부; 및 학습 시에 결합되어 상기 인코더부가 보행자의 위치 및 크기에 대응하여 경계 박스가 미리 설정된 학습 영상인 학습용 컬러 영상과 학습용 열 영상에 대해 추출한 컬러 특징맵과 열 특징맵에서 설정된 경계 박스에 대응하는 관심 영역을 추출하고, 추출된 관심 영역에 대해 적대적 판정을 수행하며, 분류 스코어와 보행자 영역 그리고 적대적 판정 결과를 기초로 분류 손실과 영역 손실 및 적대적 손실을 계산하여 상기 분류 손실과 상기 영역 손실을 상기 통합 검출부로 역전파하여 학습시킴과 동시에 상기 적대적 손실에 따라 적대적 학습이 수행되는 학습부를 포함한다.

[0011] 상기 통합 검출부는 상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식에 따라 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어를 추정하는 보행자 판별부; 및 상기 보행자 판별부와 동일하게 상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 미리 학습된 패턴 추정 방식과 상기 보행자 판별부에서 판별된 분류 스코어에 따라 서로 다른 크기 및 형태의 기지정된 개수의 경계 박스 중 보행자의 형상에 따른 보행자 영역을 나타내는 경계 박스를 추출하는 경계 판별부를 포함할 수 있다.

[0012] 상기 학습부는 상기 인코더부가 상기 학습 영상으로 인가된 상기 학습용 컬러 영상 또는 상기 학습용 열 영상 중 적어도 하나에 대해 추출한 컬러 특징맵과 열 특징맵에서 관심 영역을 추출하는 관심 영역 추출부; 상기 적대적 손실에 따라 학습되는 2개의 인공 신경망을 포함하여 상기 컬러 특징맵과 상기 열 특징맵 각각에서 추출된 관심 영역에 대해 리얼 또는 페이크로 판별되도록 서로 상반되도록 판별하는 적대적 판정부; 및 상기 학습 영상에 레이블된 보행자가 존재하는 영역의 개수와 존재하지 않는 영역의 개수 사이의 비와 보행자가 존재하는 영역에 대한 분류 스코어에 따라 상기 분류 손실을 계산하고, 상기 학습용 컬러 영상 또는 상기 학습용 열 영상에 미리 설정된 경계 박스와 상기 통합 검출부에서 추출된 경계 박스 사이의 차이를 상기 영역 손실로 계산하며, 상기 적대적 판정부의 2개의 인공 신경망 각각에 대한 적대적 손실을 개별적으로 계산하여 역전파하는 손실 역전파부를 포함할 수 있다.

[0013] 상기 적대적 판정부는 인공 신경망으로 구현되어 상기 열 특징맵에서 추출된 관심 영역이 인가되면 리얼로 판정하고, 상기 컬러 특징맵에서 추출된 관심 영역이 인가되면 페이크로 판정하도록 학습되는 제1 판정부; 및 인공

신경망으로 구현되어 상기 컬러 특징맵에서 추출된 관심 영역이 인가되면 리얼로 판정하고, 상기 열 특징맵에서 추출된 관심 영역이 인가되면 페이크로 판정하도록 학습되는 제2 판정부를 포함할 수 있다.

[0014] 상기 손실 역전파부는 상기 분류 손실(L_{cls})을 수학식

$$\mathcal{L}_{cls}(p) = \begin{cases} (1/\alpha_t)(1-p^y)\log(p) & \text{if } y = 1 \\ \alpha_t(p)^y\log(1-p) & \text{otherwise} \end{cases}$$

[0015] (여기서 α_t 는 학습 영상에 레이블된 진리값에 기반하여 판별되는 보행자가 존재하는 영역($y = 1$)의 개수($\Omega_{y=1}$)와 존재하지 않는 영역($y = -1$)사이의 개수($\Omega_{y=-1}$) 비로서, $\alpha_t = |\Omega_{y=1}|/|\Omega_{y=-1}|$ 로 계산되며, p 는 보행자가 존재하는 영역($y = 1$)에 대한 분류 스코어이며, y 는 미리 지정되는 초점 매개변수(focus parameter)이다.)에 따라 계산할 수 있다.

[0017] 상기 손실 역전파부는 상기 영역 손실(L_{loc})을 수학식

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

[0019] 의 평활 L1 함수($smooth_{L1}$)를 적용하여 수학식

$$\mathcal{L}_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L1}(t_i^u - v_i)$$

[0021] (여기서 t^u 는 학습 영상에 레이블된 진리 경계 박스를 나타내고, v 는 통합 검출부에서 획득된 경계 박스를 나타내며, x, y, w, h 는 각각 획득된 경계 박스의 중심 위치와 폭 및 높이를 나타낸다.)에 따라 계산할 수 있다.

[0022] 상기 손실 역전파부는 상기 제1 판정부에 대한 컬러 적대적 손실(L_{adv}^C)을 수학식

$$\begin{aligned} \mathcal{L}_{adv}^C(P^C, P^T, D^C) \\ = \mathbb{E}_{P^C \sim p_{data}(P^C)} \left[\log(D^C(P^C)) + \log(1 - D^C(P^T)) \right] \end{aligned}$$

[0024] 에 따라 계산하고, 상기 제2 판정부에 대한 열 적대적 손실(L_{adv}^T)을 수학식

$$\begin{aligned} \mathcal{L}_{adv}^T(P^C, P^T, D^T) \\ = \mathbb{E}_{P^T \sim p_{data}(P^T)} \left[\log(D^T(P^T)) + \log(1 - D^T(P^C)) \right] \end{aligned}$$

[0026] (여기서 P^C 와 P^T 는 각각 학습용 컬러 영상에서 추출된 관심 영역과 학습용 열 영상에서 추출된 관심 영역이고, D^C 와 D^T 는 각각 제1 및 제2 판정부를 판정 함수로 나타낸 것이며, $\mathbb{E}_{P^C \sim p_{data}(P^C)}$ 와 $\mathbb{E}_{P^T \sim p_{data}(P^T)}$ 는 각각 P^C 와 P^T 에서 고정 길이 특징 벡터를 추출하기 위한 ROI 풀링을 나타낸다.)에 따라 계산할 수 있다.

[0027] 상기 손실 역전파부는 상기 컬러 적대적 손실(L_{adv}^C)과 상기 열 적대적 손실(L_{adv}^T)을 가산하여 적대적 손실(L_{adv})을 계산하고, 상기 분류 손실(L_{cls})과 상기 영역 손실(L_{loc}) 및 상기 적대적 손실(L_{adv})로부터 총 손실(L_{total})을 수학식

$$\mathcal{L}_{total} = \sum_{i \in \Omega^C, \Omega^T} \mathcal{L}_{cls}^i + \sum_{j \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{loc}^j + \lambda \sum_{k \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{adv}^k$$

(여기서 λ 는 손실 밸런스를 조절하기 위한 매개변수이다.)에 따라 계산할 수 있다.

상기 손실 역전파부는 상기 총 손실(\mathcal{L}_{total})이 수학식

$$\arg \min_{\{E^C, E^T, U\}} \max_{\{D^C, D^T\}} \mathcal{L}_{total}$$

(여기서 E^C, E^T 는 컬러 영상과 열 영상을 인코딩하는 인코더부의 함수 표현이고, U 는 통합 검출부에 대한 함수 표현이며, D^C, D^T 는 제1 및 제2 판정부에 대한 함수 표현이다.)에 따라 최적화될 때까지 상기 분류 손실(\mathcal{L}_{cls})과 상기 영역 손실(\mathcal{L}_{loc})을 상기 통합 검출부로 역전파하고, 상기 컬러 적대적 손실(\mathcal{L}_{adv}^C)을 상기 제1 판정부로 역전파하며, 상기 열 적대적 손실(\mathcal{L}_{adv}^T)을 상기 제2 판정부로 역전파하여 반복 학습을 수행할 수 있다.

상기 다른 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 보행자 검출 방법은 컬러 영상과 열 영상 중 적어도 하나를 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 컬러 특징맵과 열 특징맵을 획득하는 인코딩 단계; 상기 컬러 특징맵과 상기 열 특징맵 중 적어도 하나를 인가받고, 패턴 추정 방식이 미리 학습된 인공 신경망을 이용하여 상기 컬러 특징맵 또는 상기 열 특징맵 각각에서 위치별 보행자의 존재 확률을 나타내는 분류 스코어와 보행자 영역을 나타내는 경계 박스를 추출하는 보행자 검출 단계; 및 보행자의 위치 및 크기에 대응하여 경계 박스가 미리 설정된 학습 영상인 학습용 컬러 영상과 학습용 열 영상에 대해 추출한 컬러 특징맵과 열 특징맵에서 설정된 경계 박스에 대응하는 관심 영역을 추출하고, 추출된 관심 영역에 대해 적대적 판정을 수행하며, 분류 스코어와 보행자 영역 그리고 적대적 판정 결과를 기초로 분류 손실과 영역 손실 및 적대적 손실을 계산하여 상기 분류 손실과 상기 영역 손실을 역전파하여 상기 적대적 손실에 따라 적대적 학습이 수행되는 학습 단계를 포함한다.

발명의 효과

따라서, 본 발명의 실시예에 따른 보행자 검출 장치 및 방법은 서로 매칭되도록 조절되지 않은 컬러 영상과 열 영상인 비쌍 교차 스펙트럴 영상으로부터 적대적 학습 기법을 이용하여 학습이 수행되므로, 컬러 영상과 열 영상을 정렬시키기 위한 별도의 소프트웨어나 하드웨어를 요구하지 않고도 공통 인공 신경망이 용이하게 학습되어 컬러 영상 또는 열 영상으로부터 정확하게 보행자를 검출할 수 있다. 또한 컬러 영상 또는 열 영상 중 하나만 인가되는 경우에도 보행자를 검출할 수 있다.

도면의 간단한 설명

도 1은 본 발명의 일 실시예에 따른 보행자 검출 장치의 개략적 구조를 나타낸다.

도 2는 도 1의 보행자 검출 장치가 비쌍 교차 스펙트럴 영상으로부터 특징 공간에서 보행자를 검출하는 개념을 설명하기 위한 도면이다.

도 3은 본 발명의 일 실시예에 따른 보행자 검출 방법을 나타낸다.

도 4는 본 실시예에 따른 보행자 검출 방법의 성능을 비교한 결과를 나타낸다.

발명을 실시하기 위한 구체적인 내용

본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

- [0038] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0039] 도 1은 본 발명의 일 실시예에 따른 보행자 검출 장치의 개략적 구조를 나타내고, 도 2는 도 1의 보행자 검출 장치가 비쌍 교차 스펙트럴 영상으로부터 특징 공간에서 보행자를 검출하는 개념을 설명하기 위한 도면이다.
- [0040] 도 1을 참조하면, 본 실시예에 따른 비쌍 교차 스펙트럴 보행자 검출 장치는 인코더부, 통합 검출부, 관심 영역 추출부, 적대적 판정부 및 손실 역전과부를 포함할 수 있다.
- [0041] 인코더부(100)는 교차 스펙트럴 영상인 컬러 영상(I^C)과 열 영상(I^T)을 인가받아 미리 학습된 패턴 추정 방식에 따라 인가된 컬러 영상(I^C)과 열 영상(I^T)을 인코딩하여 컬러 특징맵(A^C)과 열 특징맵(A^T)을 획득한다.
- [0042] 여기서 컬러 영상(I^C)과 열 영상(I^T)은 해상도나 화각 등이 서로 매칭되도록 페어링된 영상일 수도 있으나, 개별적으로 획득되어 페어링되지 않은 비쌍 교차 스펙트럴 영상(Unpaired Cross-Spectral Images)이어도 무방하다.
- [0043] 인코더부(100)는 제1 인코더(110)와 제2 인코더(120)를 포함하며, 제1 인코더는 컬러 영상(I^C)을 인가받아 인코딩하여 컬러 특징맵(A^C)을 획득하고, 제2 인코더(120)는 열 영상(I^T)을 인가받아 인코딩하여 열 특징맵(A^T)을 획득한다.
- [0044] 제1 및 제2 인코더(110, 120)는 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현될 수 있으며, 일 예로 컨볼루션 신경망(Convolutional Neural Network: 이하 CNN) 등으로 구현될 수 있다. 그리고 제1 및 제2 인코더(110, 120)는 동일하게 학습되지 않아도 될 뿐만 아니라 동일한 구조의 인공 신경망이 아니어도 무방하다. 즉 학습된 인공 신경망으로 구현되는 제1 및 제2 인코더(110, 120)의 가중치(W^C , W^T)가 동일하지 않아도 무방하다.
- [0045] 인공 신경망으로 구현되는 제1 및 제2 인코더(110, 120)에 의해 획득되는 컬러 특징맵(A^C)과 열 특징맵(A^T)은 각각 수학적 식 1 및 2로 표현될 수 있다.

수학적 식 1

$$A^C = E^C(I^C; W^C)$$

수학적 식 2

$$A^T = E^T(I^T; W^T)$$

- [0046]
- [0048] 수학적 식 1 및 2에서 E^C 와 E^T 는 각각 제1 인코더(110)와 제2 인코더(120)를 인코딩 함수로 표현한 것이고, W^C 와 W^T 는 각각 미리 학습된 패턴 추정 방식에 의해 설정된 제1 인코더(110) 및 제2 인코더(120)의 가중치를 나타낸다.
- [0049] 여기서 제1 인코더(110)와 제2 인코더(120)에 의해 인코딩되어 획득된 컬러 특징맵(A^C)과 열 특징맵(A^T)은 도 2에 도시된 바와 같이, 비쌍 교차 스펙트럴 영상인 컬러 영상(I^C)과 열 영상(I^T)이 상보 공간인 특징 공간에 공통으로 투영된 것으로 볼 수 있다. 즉 인코더부(100)는 학습된 패턴 추정 방식에 의해 서로 다른 종류의 영상인 컬러 영상(I^C)과 열 영상(I^T)을 공통의 특징 공간에 맵핑할 수 있다.
- [0050] 통합 검출부(200)는 컬러 특징맵(A^C)과 열 특징맵(A^T) 중 적어도 하나를 인가받고, 인가된 컬러 특징맵(A^C) 또는 열 특징맵(A^T)에서 적대적 학습 방식에 기반하여 미리 학습된 패턴 추정 방식에 따라 보행자를 검출한다. 여기

서 통합 검출부(200)는 인가된 컬러 특징맵(A^C) 또는 열 특징맵(A^T) 각각에서 보행자가 존재할 확률을 나타내는 분류 스코어(p_C , p_T)와 보행자의 위치 정보(R_C , R_T)를 획득한다.

[0051] 통합 검출부(200)는 컬러 특징맵(A^C) 또는 열 특징맵(A^T)으로부터 위치별로 보행자의 존재여부를 판별하는 보행자 판별부(210) 및 컬러 특징맵(A^C) 또는 열 특징맵(A^T)에서 판별된 보행자가 포함되는 경계 영역을 추출하여 경계 박스를 획득하는 경계 판별부(220)를 포함할 수 있다.

[0052] 보행자 판별부(210) 또한 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되어, 인가된 컬러 특징맵(A^C) 또는 열 특징맵(A^T) 각각에서 기지정된 크기의 영역별로 보행자의 존재 확률에 따른 분류 스코어(p_C , p_T)를 수학식 3 및 4에 따라 획득할 수 있다.

수학식 3

$$p_C = C(A^C; W^A)$$

수학식 4

$$p_T = C(A^T; W^A)$$

[0055] 여기서 C는 보행자 판별부(210)를 분류 함수로 표현한 것이며, W^A 는 미리 학습된 패턴 추정 방식에 의해 설정된 보행자 판별부(210)의 가중치를 나타낸다.

[0056] 한편 경계 판별부(220)도 패턴 추정 방식이 미리 학습된 인공 신경망으로 구현되고, 컬러 특징맵(A^C) 또는 열 특징맵(A^T) 각각에 대해 보행자 판별부(210)에서 판별된 위치별 분류 스코어(p_C , p_T)를 기반으로 인가된 컬러 특징맵(A^C) 또는 열 특징맵(A^T)으로부터 보행자별 경계 박스를 추정한다.

[0057] 경계 판별부(220)는 각 위치별로 적용 가능한 서로 다른 크기 및 형태의 기지정된 개수(예를 들면 9개)의 경계 박스가 미리 지정되고, 보행자 판별부(210)에서 위치별 분류 스코어(p_C , p_T)가 판별되면, 컬러 특징맵(A^C) 또는 열 특징맵(A^T)의 각 위치에서 보행자의 형상에 대응하는 형상의 경계 박스를 판별할 수 있다. 이때, 경계 판별부(220)는 획득된 분류 스코어(p_C , p_T) 중 기지정된 문턱 스코어 이상인 스코어를 갖는 보행자 판별부(210)에서 획득된 위치별 분류 스코어(p_C , p_T)에 기반하여 보행자가 존재할 것으로 판별되는 위치에 대해서만 경계 박스를 획득할 수도 있다.

[0058] 경계 판별부(220)는 경계 박스(B_C , B_T)를 수학식 5 및 6에 따라 획득할 수 있다.

수학식 5

$$B_C = (x_C, y_C, w_C, h_C) = R(A^C; W^B)$$

수학식 6

$$B_T = (x_T, y_T, w_T, h_T) = R(A^T; W^B)$$

[0060]

[0061]

여기서 B_C 와 B_T 는 각각 컬러 특징맵(A^C)과 열 특징맵(A^T)에 대한 경계 박스로서, (x_C, y_C) , (x_T, y_T) 는 경계 박스의 중심 위치를 나타내고, R 은 경계 판별부(220)를 경계 판별 함수로 표현한 것이며, (w_C, h_C) , (w_T, h_T) 는 경계 박스의 폭과 높이를 나타낸다. 그리고 W^B 는 미리 학습된 패턴 추정 방식에 의해 설정된 경계 판별부(220)의 가중치를 나타낸다.

[0062]

즉 통합 검출부(200)는 컬러 특징맵(A^C) 또는 열 특징맵(A^T) 각각에서 영역별로 보행자가 존재하는 확률에 따른 분류 스코어(p_C , p_T)를 획득하고, 분류 스코어(p_C , p_T)에 기반하여 보행자가 존재하는 위치의 경계 박스를 지정함으로써, 컬러 영상(I^C) 또는 열 영상(I^T)에서 보행자 영역의 위치와 크기를 추출할 수 있다. 따라서 쌍을 이루지 않는 컬러 영상(I^C)과 열 영상(I^T)으로 구성된 비쌍 교차 스펙트럴 영상이 인가되더라도, 비쌍 교차 스펙트럴 영상으로부터 보행자 영역의 위치와 크기를 판별할 수 있다. 또한 컬러 영상(I^C) 또는 열 영상(I^T) 중 하나만 인가되더라도, 인가된 컬러 영상(I^C) 또는 열 영상(I^T)에서 보행자 영역의 위치와 크기를 판별할 수 있다.

[0063]

상기한 통합 검출부(200)는 인코더부(100)에 의해 컬러 특징맵(A^C) 또는 열 특징맵(A^T)의 형태로 맵핑된 공통의 특징 공간에서 각 영역별 보행자 존재 확률과 경계 박스를 판별함으로써, 입력된 영상이 컬러 영상(I^C)인지 열 영상(I^T)인지에 무관하게 보행자 영역의 위치와 크기를 판별할 수 있다.

[0064]

상기에서는 설명의 편의를 위하여 보행자 판별부(210)와 경계 판별부(220)를 별도의 구성으로 도시하였으나, 경우에 따라서 판별부(210)와 경계 판별부(220)를 포함하는 통합 검출부(200)는 컬러 특징맵(A^C)과 열 특징맵(A^T)의 위치별 분류 스코어(p_C , p_T)와 경계 박스(B_C , B_T)를 한번에 추출하는 단일 인공 신경망으로 구현될 수도 있다. 이 경우 통합 검출부(200)는 일 예로 레티나 네트워크(RetinaNet)로 구현될 수 있다.

[0065]

다만 통합 검출부(200)가 공통 특징 공간에 맵핑된 컬러 특징맵(A^C)과 열 특징맵(A^T) 각각으로부터 분류 스코어(p_C , p_T)와 경계 박스(B_C , B_T)를 획득하기 위해서는 학습이 선행되어야만 한다. 즉 상기한 바와 같이, 통합 검출부(200)가 컬러 특징맵(A^C)과 열 특징맵(A^T)이 서로 매칭되지 않은 비쌍 교차 스펙트럴 영상을 기반으로 정확하게 분류 스코어(p_C , p_T)와 경계 박스(B_C , B_T)를 추출할 수 있도록 학습이 수행될 수 있어야 한다.

[0066]

이에 본 실시예의 보행자 검출 장치는 관심 영역 추출부와 적대적 판정부 및 손실 역전파부를 더 포함할 수 있다. 관심 영역 추출부와 적대적 판정부 및 손실 역전파부는 통합 검출부(200)를 학습시키기 위한 학습부로서 기능하므로, 통합 검출부(200)의 학습이 완료된 이후에는 제거될 수 있다. 즉 실제 이용 시, 본 실시예에 따른 보행자 검출 장치는 인코더부(100)와 통합 검출부(200)만으로 구성될 수 있다.

[0067]

관심 영역 추출부는 제1 인코더(110)에서 획득된 컬러 특징맵(A^C)에서 관심 영역을 추출하여 컬러 관심 영역($A^C(R^C)$)을 출력하는 제1 관심 영역 추출부(310)와 제2 인코더(120)에서 획득된 열 특징맵(A^T)에서 관심 영역을 추출하여 열 관심 영역($A^T(R^T)$)을 출력하는 제2 관심 영역 추출부(320)를 포함할 수 있다.

[0068]

여기서 제1 및 제2 관심 영역 추출부(310, 320)는 제1 및 제2 인코더(110, 120)가 보행자의 위치 및 크기에 대응하여 경계 박스가 미리 설정된 학습용 컬러 영상과 학습용 열 영상에 대해 추출한 컬러 특징맵(A^C)과 열 특징맵(A^T)에서 설정된 경계 박스에 대응하는 관심 영역을 추출하여 컬러 관심 영역($A^C(R^C)$)과 열 관심 영역($A^T(R^T)$)을 추출한다.

[0069]

적대적 판정부는 컬러 관심 영역($A^C(R^C)$)과 열 관심 영역($A^T(R^T)$)을 인가받고, 미리 학습된 패턴 분류 방식에 따

라 인가된 컬러 관심 영역($A^C(R^C)$)과 열 관심 영역($A^T(R^T)$)을 리얼(R) 또는 페이크(F)로 판정한다. 적대적 판정부는 컬러 관심 영역($A^C(R^C)$)이 인가되면 페이크(F)로 판정하는 반면, 열 관심 영역($A^T(R^T)$)이 인가되면 리얼(R)로 판정하는 제1 판정부(410)와 컬러 관심 영역($A^C(R^C)$)이 인가되면 리얼(R)로 판정하고, 열 관심 영역($A^T(R^T)$)이 인가되면 페이크(F)로 판정하는 제2 판정부(420)로 포함할 수 있다.

[0070] 그러므로 제1 판정부(410)와 제2 판정부(420)는 컬러 관심 영역($A^C(R^C)$)과 열 관심 영역($A^T(R^T)$)에 대해 서로 상반된 결과를 출력하도록 학습된다. 이는 제1 판정부(410)와 제2 판정부(420)가 서로 적대적 학습 방식으로 학습되어 리얼(R) 또는 페이크(F)를 판정하도록 함으로써, 통합 검출부(200)가 공통 특징 공간에 맵핑된 컬러 특징맵(A^C)과 열 특징맵(A^T) 각각으로부터 분류 스코어(p_c, p_t)와 경계 박스(B_c, B_t)를 획득할 수 있도록 학습시키기 위함이다. 여기서 제1 판정부(410)와 제2 판정부(420)는 통합 검출부(200)와 함께 학습될 수 있다.

[0071] 손실 역전파부(500)는 통합 검출부(200)에서 획득된 분류 스코어(p_c, p_t)와 경계 박스(B_c, B_t) 및 제1 및 제2 판정부(410, 420)에서 판정된 판정 결과로부터 기지정된 방식으로 손실을 계산하여 역전파함으로써, 제1 판정부(410)와 제2 판정부(420) 및 통합 검출부(200)를 학습시킬 수 있다.

[0072] 손실 역전파부(500)는 통합 검출부(200)에 대한 손실로 보행자 검출 성능에 대한 분류 손실(L_{cls})과 경계 박스 추출에 대한 영역 손실(L_{loc}) 그리고 적대적 판정부의 제1 및 제2 판정부(410, 420)의 판정 손실인 적대적 손실(L_{adv})을 각각 계산하고 결합하여 총 손실(L_{total})을 계산하고, 계산된 총 손실(L_{total})이 최적화될 때까지 반복하여, 분류 손실(L_{cls})과 영역 손실(L_{loc})을 통합 검출부(200)로 역전파하고 적대적 판정부로 적대적 손실(L_{adv})을 역전파하여 학습시킨다.

[0073] 손실 역전파부(500)는 우선 분류 손실(L_{cls})을 수학적 식 7에 따라 계산할 수 있다.

수학적 식 7

$$\mathcal{L}_{cls}(p) = \begin{cases} (1/\alpha_t)(1 - p^r)\log(p) & \text{if } y = 1 \\ \alpha_t(p)^r\log(1 - p) & \text{otherwise} \end{cases}$$

[0074]

[0075] 여기서 α_t 는 학습 영상에 레이블된 진리값에 기반하여 판별되는 보행자가 존재하는 영역($y = 1$)의 개수($\Omega_{y=1}$)와 존재하지 않는 영역($y = -1$)의 개수($\Omega_{y=-1}$) 사이의 비로서, $\alpha_t = |\Omega_{y=1}|/|\Omega_{y=-1}|$ 로 계산되며, p 는 보행자가 존재하는 영역($y = 1$)에 대한 통합 검출부(200)의 분류 스코어로서 인가된 학습 영상에 따라 p_c 또는 p_t 이며, r 는 미리 지정되는 초점 매개변수(focus parameter)로서 여기서는 일예로 2($r = 2$)로 설정된 것으로 가정한다.

[0076] 그리고 손실 역전파부(500)는 경계 박스 추출에 대한 영역 손실(L_{loc})을 진리값에 기반하여 판별되는 보행자가 존재하는 영역($y = 1$) 각각에 대해 수학적 식 8의 평활 L1 함수($smooth_{L1}$)를 적용하여 수학적 식 9에 따라 계산하여 획득할 수 있다.

수학적 식 8

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}$$

[0077]

수학식 9

$$\mathcal{L}_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L1}(t_i^u - v_i)$$

[0078]

[0079] 여기서 t^u 는 학습 영상에 레이블된 진리 경계 박스를 나타내고, v 는 통합 검출부(200)에서 획득된 경계 박스를 나타낸다.

[0080] 한편 손실 역전과부(500)는 제1 판정부(410)의 컬러 적대적 손실(L_{adv}^C)을 수학식 10에 따라 계산하고, 제2 판정부(420)의 열 적대적 손실(L_{adv}^T)을 수학식 11에 따라 계산할 수 있다.

수학식 10

$$\begin{aligned} \mathcal{L}_{adv}^C(P^C, P^T, D^C) \\ = \mathbb{E}_{P^C \sim p_{data}(P^C)} \left[\log(D^C(P^C)) + \log(1 - D^C(P^T)) \right] \end{aligned}$$

[0081]

수학식 11

$$\begin{aligned} \mathcal{L}_{adv}^T(P^C, P^T, D^T) \\ = \mathbb{E}_{P^T \sim p_{data}(P^T)} \left[\log(D^T(P^T)) + \log(1 - D^T(P^C)) \right] \end{aligned}$$

[0082]

[0083] 여기서 P^C 와 P^T 는 각각 학습 영상에서 추출된 보행자 특징인 컬러 관심 영역($P^C = A^C(R^C)$)과 열 관심 영역($P^T = A^T(R^T)$)이고, D^C 와 D^T 는 각각 제1 및 제2 판정부(410, 420)를 판정 함수로 나타낸 것이며, $\mathbb{E}_{P^C \sim p_{data}(P^C)}$ 와 $\mathbb{E}_{P^T \sim p_{data}(P^T)}$ 는 각각 다양한 크기의 P^C 와 P^T 에서 고정 길이 특징 벡터를 추출하기 위한 ROI 풀링을 나타낸다.

[0084] 수학식 10 및 11에 따라 손실 역전과부(500)는 통합된 적대적 손실(L_{adv})을 수학식 12과 같이 계산할 수 있다.

수학식 12

$$\begin{aligned} \mathcal{L}_{adv}(P^C, P^T, D^C, D^T) \\ = \mathbb{E}_{P^C \sim p_{data}(P^C)} [\log(D^C(P^C)) + \log(1 - D^T(P^C))] \\ + \mathbb{E}_{P^T \sim p_{data}(P^T)} [\log(D^T(P^T)) + \log(1 - D^C(P^T))] \end{aligned}$$

[0085]

[0086] 그리고 수학식 7, 9 및 12으로부터 손실 역전과부(500)는 총 손실(L_{total})을 수학식 13과 같이 계산할 수 있다.

수학식 13

$$\mathcal{L}_{total} = \sum_{i \in \Omega^C, \Omega^T} \mathcal{L}_{cls}^i + \sum_{j \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{loc}^j + \lambda \sum_{k \in \Omega_{y=1}^C, \Omega_{y=1}^T} \mathcal{L}_{adv}^k$$

여기서 λ 는 통합 검출부(200)와 적대적 판별부 사이의 손실 밸런스를 조절하기 위한 매개변수이다.

그리고 손실 역전파부(500)는 총 손실(\mathcal{L}_{total})을 계산하고, 계산된 총 손실(\mathcal{L}_{total})에 따라 분류 손실(\mathcal{L}_{cls})과 영역 손실(\mathcal{L}_{loc})을 통합 검출부로 역전파하고, 컬러 적대적 손실(\mathcal{L}_{adv}^C), 열 적대적 손실(\mathcal{L}_{adv}^T)을 각각 제1 및 제2 판정부(420)로 역전파한다. 이때 손실 역전파부(500)는 수학식 14의 최소-최대 문제(minmax problem)에 따라 계산된 총 손실(\mathcal{L}_{total})이 최적화될 때까지 반복하여 계산할 수 있다.

수학식 14

$$\arg \min_{\{E^C, E^T, U\}} \max_{\{D^C, D^T\}} \mathcal{L}_{total}$$

여기서 U 는 통합 검출부(200)에 대한 함수 표현으로 보행자 판별부(210)의 분류 함수(C)와 경계 판별부(220)의 경계 판별 함수(R)의 조합으로 볼 수 있다.

즉 수학식 14에 따라 손실 역전파부(500)는 제1 및 제2 판정부(410, 420)의 판정 함수(D^C, D^T)에 따른 총 손실(\mathcal{L}_{total})이 최대가 되되, 제1 및 제2 인코더(110, 120)의 인코딩 함수(E^C, E^T)와 통합 검출부(200)의 통합 함수(U)에 따른 총 손실(\mathcal{L}_{total})이 최소가 되도록 반복하여 총 손실(\mathcal{L}_{total})을 계산하고 분류 손실(\mathcal{L}_{cls})과 영역 손실(\mathcal{L}_{loc}), 컬러 적대적 손실(\mathcal{L}_{adv}^C), 열 적대적 손실(\mathcal{L}_{adv}^T)을 역전파한다.

도 3은 본 발명의 일 실시예에 따른 보행자 검출 방법을 나타낸다.

도 1 및 도 2를 참조하여 도 3의 보행자 검출 방법을 설명하면, 우선 컬러 영상(I^C)과 열 영상(I^T) 중 적어도 하나로부터 보행자를 감지할 수 있도록 보행자 검출 학습을 수행한다(S10).

학습을 수행하기 위해서는 먼저 미리 보행자의 위치 및 크기에 대응하는 경계 박스가 미리 설정된 학습용 영상을 획득한다(S11). 여기서 학습용 영상은 학습용 컬러 영상과 학습용 열 영상을 포함할 수 있으며, 학습용 컬러 영상과 학습용 열 영상은 개별적으로 획득되어 페어링되지 않은 비쌍 교차 스펙트럴 영상일 수 있다. 즉 학습용 영상은 학습용 컬러 영상과 학습용 열 영상이 서로 매치되지 않아도 무방하며, 학습용 컬러 영상과 학습용 열 영상이 교대로 인가되어도 무방하다.

학습용 영상이 획득되면, 획득된 학습용 영상을 이전까지 학습된 패턴 추정 방식에 따라 인코딩하여 컬러 특징맵(A^C)과 열 특징맵(A^T) 중 적어도 하나를 획득한다(S12).

그리고 이전까지 학습된 패턴 추정 방식에 따라 획득된 컬러 특징맵(A^C) 또는 열 특징맵(A^T)의 각 위치에서 보행자의 존재 확률을 추정하여 분류 스코어(p_c, p_t)를 획득한다(S13). 또한 서로 다른 크기 및 형태의 기지정된 개수(예를 들면 9개)의 경계 박스 중 이전까지 학습된 패턴 추정 방식에 따라 컬러 특징맵(A^C) 또는 열 특징맵(A^T)의 각 위치 또는 보행자의 존재하는 것으로 추정된 위치에 대응하는 경계 박스(B_c, B_t)를 추정한다(S14).

이와 함께 획득된 컬러 특징맵(A^C) 또는 열 특징맵(A^T)에서 학습용 영상에 미리 지정된 경계 박스에 대응하는 컬러 관심 영역($A^C(R^C)$) 또는 열 관심 영역($A^T(R^T)$)을 추출한다(S15). 컬러 관심 영역($A^C(R^C)$) 또는 열 관심 영역

($A^T(R^T)$)이 추출되면, 추출된 컬러 관심 영역($A^C(R^C)$) 또는 열 관심 영역($A^T(R^T)$)에 대해 이전까지 학습된 패턴 추정 방식에 따라 적대적 판별을 수행한다(S16). 여기서 적대적 판별은 서로 반대 결과로 판별하도록 학습되는 2개의 인공 신경망을 이용하여 수행된다. 2개의 인공 신경망은 만일 컬러 관심 영역($A^C(R^C)$)이 추출되면, 추출된 컬러 관심 영역($A^C(R^C)$)에 대해 하나의 인공 신경망은 리얼(R)로 판별하는 반면, 나머지 인공 신경망은 페이크(F)를 판별하도록 학습되어야 하며, 반대로 열 관심 영역($A^T(R^T)$)이 추출되면, 추출된 열 관심 영역($A^T(R^T)$)에 대해 하나의 인공 신경망은 페이크(F)로 판별하는 반면, 인공 신경망은 리얼(R)을 판별하도록 학습되어야 한다.

[0099] 분류 스코어(p_c , p_r)와 경계 박스(B_c , B_r) 및 적대적 판별 결과가 획득되면, 획득된 분류 스코어(p_c , p_r)와 경계 박스(B_c , B_r) 및 적대적 판별 결과를 기초로 손실을 계산하고 역전파하여 학습을 수행한다(S17). 여기서 손실은 수학적 7, 9, 10 및 11로부터 분류 손실(L_{cls})과 영역 손실(L_{loc}), 컬러 적대적 손실(L_{adv}^C) 및 열 적대적 손실(L_{adv}^T)로 계산될 수 있다. 그리고 수학적 13과 같이 분류 손실(L_{cls})과 영역 손실(L_{loc}) 및 적대적 손실(L_{adv})을 가산한 총 손실(L_{total})이 계산될 수 있다.

[0100] 총 손실(L_{total})이 계산되면, 계산된 총 손실(L_{total})이 수학적 14에 따른 최적화 조건을 만족하는지 판별한다(S18). 만일 총 손실(L_{total})이 최적화 조건을 만족하는 못하는 것으로 다시 획득된 학습용 영상을 기반으로 특징맵을 획득하여 손실을 반복적으로 계산하고 역전파한다. 즉 총 손실(L_{total})이 최적화 조건을 만족할 때까지 보행자 검출 학습을 반복하여 수행한다.

[0101] 그러나 총 손실(L_{total})이 최적화 조건을 만족하는 것으로 판별되면, 보행자 검출 단계(S20)를 수행한다. 보행자 검출 단계(S20)에서는 보행자가 검출되어야 하는 컬러 영상과 열 영상 중 적어도 하나를 획득한다(S21). 그리고 학습된 패턴 추정 방식에 따라 획득된 컬러 영상과 열 영상 중 적어도 하나의 특징을 추출하여 컬러 특징맵(A^C)과 열 특징맵(A^T) 중 적어도 하나를 획득한다(S22). 그리고 획득된 컬러 특징맵(A^C)과 열 특징맵(A^T)에서 학습된 패턴 추정 방식에 따라 위치별 보행자의 존재 확률을 추정하여 분류 스코어(p_c , p_r)를 획득한다(S23). 즉 컬러 특징맵(A^C)과 열 특징맵(A^T)에서 보행자의 위치를 추정한다.

[0102] 이와 함께 학습된 패턴 추정 방식에 따라 컬러 특징맵(A^C) 또는 열 특징맵(A^T)에서 보행자의 존재하는 것으로 추정된 위치에 대응하는 형태 및 크기의 경계 박스(B_c , B_r)를 추정한다(S24).

[0103] 도 4는 본 실시예에 따른 보행자 검출 방법의 성능을 비교한 결과를 나타낸다.

[0104] 도 4에서 (a)는 보행자의 위치에 경계 박스가 지정된 진리값의 컬러 영상을 나타내고, (b) 및 (c)는 각각 기준의 레티나 네트워크(RetinaNet)가 컬러 영상 및 열 영상에서 보행자를 검출한 결과를 나타내며, (d) 및 (e)는 각각 본 실시예에 따른 적대적 학습으로 학습된 비쌍 교차 스펙트럴 보행자 검출 방법에 의해 컬러 영상 및 열 영상에서 보행자를 검출한 결과를 나타낸다.

[0105] (b)와 (d)를 비교하면, (b)의 경우 보행자를 검출하지 못하거나, 보행자가 아닌 배경을 보행자로 검출하는 경우가 있는 반면, (d)에서는 진리값과 거의 동일하게 경계 박스가 획득된 것을 알 수 있다. 마찬가지로 (c)와 (e)를 비교하면, (e)가 더욱 정확하게 보행자를 검출할 수 있음을 알 수 있다.

[0106] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 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 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0107] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

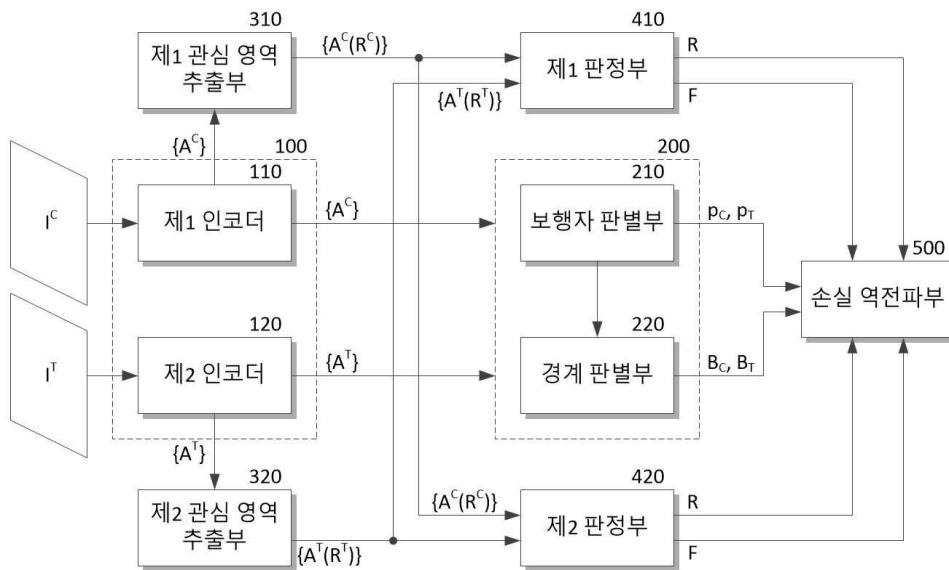
[0108] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

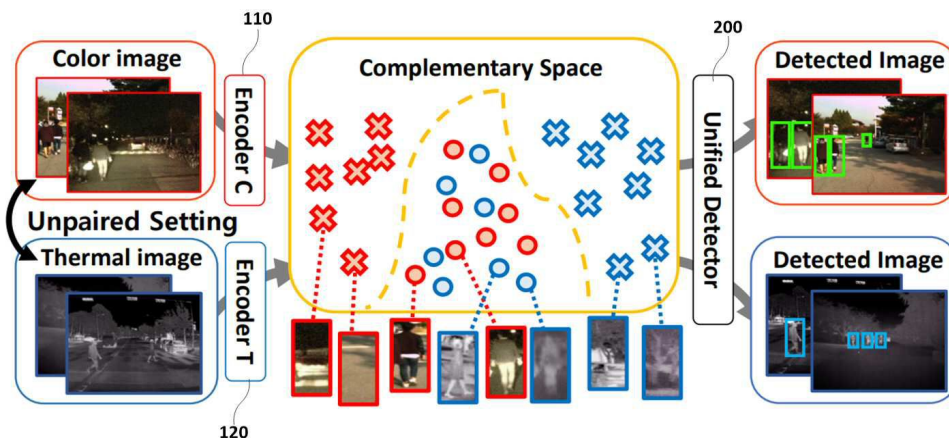
[0109] 100: 인코더부 110, 120: 인코더
200: 통합 검출부 210: 보행자 판별부
220: 경계 판별부 310, 320: 관심 영역 추출부
310, 320: 판정부 500 손실 역전파부

도면

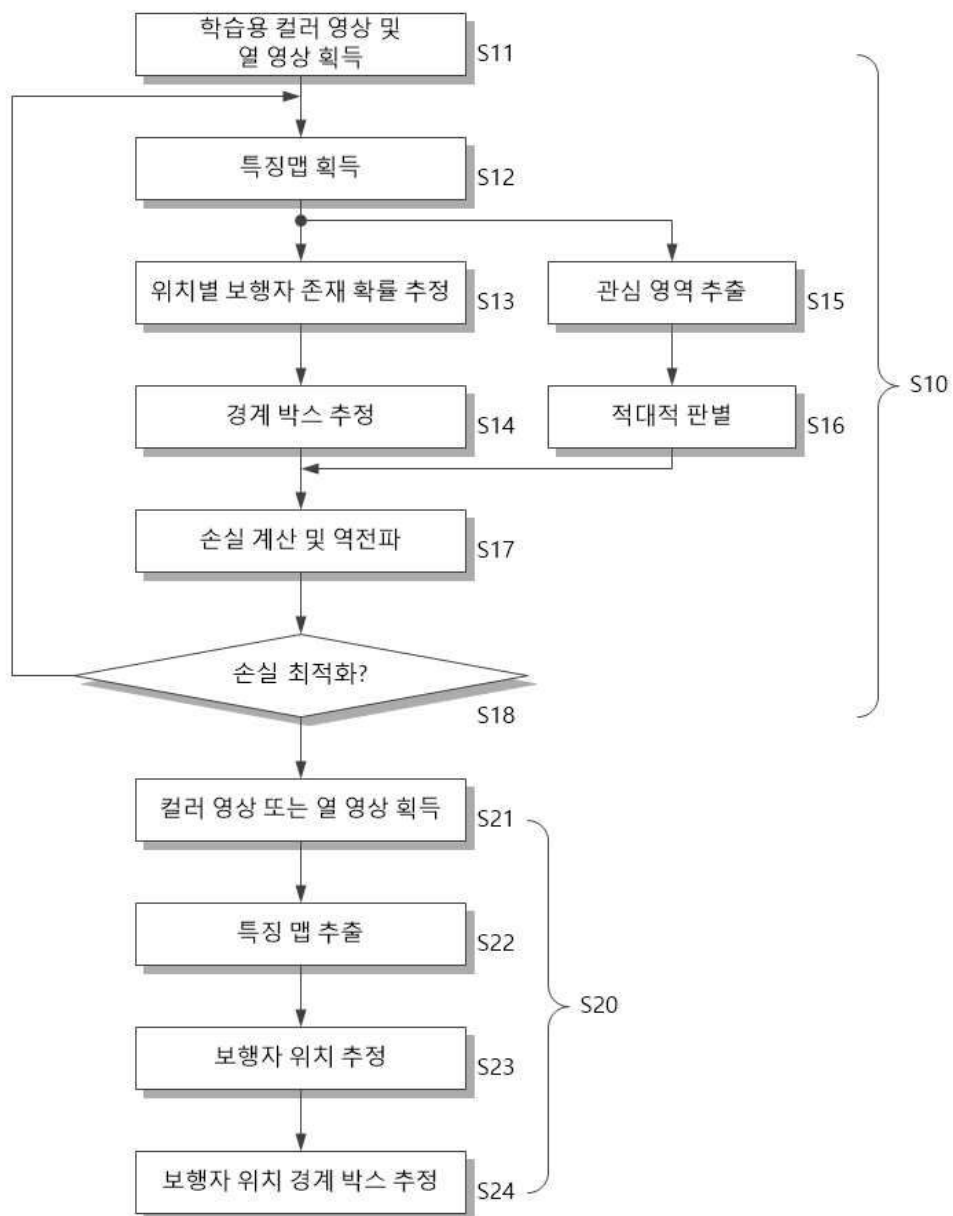
도면1



도면2



도면3



도면4

