



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년06월02일
(11) 등록번호 10-2540290
(24) 등록일자 2023년05월31일

(51) 국제특허분류(Int. Cl.)
G06F 18/00 (2023.01) G06N 20/00 (2019.01)
G06T 3/00 (2019.01) G06V 10/40 (2022.01)
(52) CPC특허분류
G06V 40/10 (2022.01)
G06F 18/22 (2023.01)
(21) 출원번호 10-2021-0006683
(22) 출원일자 2021년01월18일
심사청구일자 2021년01월18일
(65) 공개번호 10-2022-0104426
(43) 공개일자 2022년07월26일
(56) 선행기술조사문헌
JP2012098170 A*
KR1020200101521 A*
US20200242333 A1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
함범섭
서울특별시 강남구 압구정로61길 37, 72동 506호
(압구정동, 한양아파트)
박현중
서울특별시 송파구 양재대로 1218, 204동 504호(방이동, 올림픽선수기자촌아파트)
이상훈
서울특별시 서대문구 연희로26가길 14, 301호(연희동)
(74) 대리인
민영준

전체 청구항 수 : 총 18 항

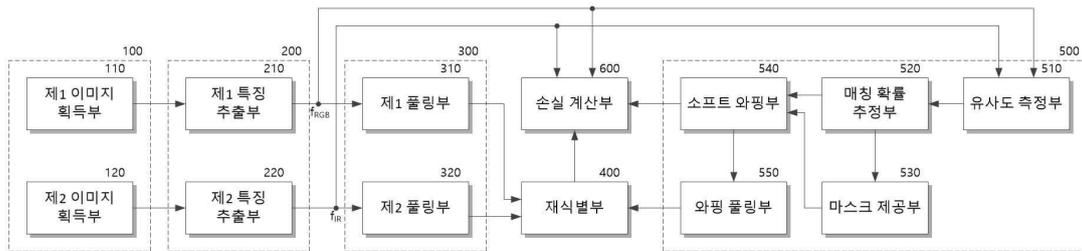
심사관 : 노용완

(54) 발명의 명칭 이중 센서 카메라 기반 사람 재식별 장치 및 방법

(57) 요약

본 발명은 이중 센서 카메라에서 획득된 이미지들에서 추출된 특징맵과 함께 특징맵에서 서로 대응하는 픽셀을 와핑한 와핑 특징맵을 기반으로 학습을 수행하며, 특히 특징맵과 와핑 특징맵에 대한 밀집 삼중 손실을 계산하여 학습에 반영함으로써, 정렬되지 않은 이미지들에서도 세분화된 로컬 정보를 추출할 수 있어 다양한 카메라에서 매우 상이한 조건에서 획득된 이미지로부터도 사람을 정확하게 재식별할 수 있는 사람 재식별 장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

G06N 20/00 (2021.08)

G06T 3/0093 (2013.01)

G06V 10/40 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711119679
과제번호	2018M3E3A1057289
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	복합인지기술개발사업
연구과제명	이중 CCTV 영상에서의 딥러닝 기반 실종자 초동 신원확인 및 추적 시스템 (2
단계)(2/3)	
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

서로 다른 이종의 센서를 이용하여 획득된 제1 및 제2 이미지를 인가받아 학습되는 방식에 따라 특징을 추출하여 제1 및 제2 특징맵을 획득하고, 기지정된 방식으로 풀링하여 제1 및 제2 표현자를 획득하는 특징 풀링부;

학습 시에 구비되어, 상기 제1 및 제2 특징맵을 인가받아 서로 대응하는 픽셀을 추정하고, 추정된 대응 픽셀의 위치를 기반으로 상호 와핑을 수행하여, 상기 제2 특징맵이 상기 제1 특징맵에 대응하도록 와핑된 제1 와핑 특징맵과 상기 제1 특징맵이 상기 제2 특징맵에 대응하도록 와핑된 제2 와핑 특징맵 획득하며, 상기 제1 및 제2 와핑 특징맵을 풀링하여 제1 및 제2 와핑 표현자를 획득하는 교차 모달 정렬부;

학습되는 방식에 따라 상기 제1 및 제2 표현자와 상기 제1 및 제2 와핑 표현자 각각이 기지정된 다수의 클래스 각각에 대응하는 확률을 나타내는 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵을 획득하고, 획득된 제1 및 제2 스코어맵에서 가장 높은 확률의 클래스의 식별자를 획득하는 재식별부; 및

학습 시에 구비되어, 상기 제1 특징맵과 상기 제1 와핑 특징맵 사이의 차이와 상기 제2 특징맵과 상기 제2 와핑 특징맵 사이의 차이를 동일한 식별자인 경우의 포지티브와 서로 다른 식별자인 경우의 네거티브로 구분하여 밀집 삼중 손실을 계산하는 손실 계산부를 포함하되,

상기 교차 모달 정렬부는

기지정된 방식에 따라 상기 제1 및 제2 특징맵 사이의 픽셀간 유사도를 계산하여 유사도맵을 획득하는 유사도 측정부;

상기 유사도 맵을 기반으로 상기 제1 및 제2 특징맵의 각 픽셀별 매칭 확률을 계산하여 매칭 확률맵을 획득하는 매칭 확률 추정부;

상기 매칭 확률맵을 이용하여 상기 제1 및 제2 특징맵 각각의 픽셀들이 상대 특징맵의 각 픽셀에 매칭될 확률을 가중 집계하여 상대 특징맵으로 와핑함으로써 상기 제1 및 제2 와핑 특징맵을 획득하는 소프트 와핑부; 및

상기 제1 및 제2 와핑 특징맵을 기지정된 방식으로 풀링하여 제1 및 제2 와핑 표현자를 획득하는 와핑 풀링부를 포함하는 사람 재식별 장치.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 교차 모달 정렬부는

상기 제1 및 제2 특징맵의 각 위치별 픽셀의 크기를 정규화하여 제1 및 제2 마스크를 획득하는 마스크 제공부를 더 포함하는 사람 재식별 장치.

청구항 4

제3항에 있어서, 상기 소프트 와핑부는

상기 제1 및 제2 마스크를 인가받아 상기 제1 및 제2 특징맵 각각의 픽셀들이 상대 특징맵의 각 픽셀에 매칭될 확률을 가중 집계하고, 가중 집계된 와핑 픽셀값에 마스크의 픽셀값을 추가로 가중하여 와핑하는 사람 재식별 장치.

청구항 5

제4항에 있어서, 상기 마스크 제공부는

상기 제1 특징맵(f_{RGB})으로부터 상기 제1 마스크(M_{RGB})를 수학식

$$M_{RGB}(\mathbf{p}) = f(\|f_{RGB}(\mathbf{p})\|_2)$$

(여기서 $\|\cdot\|_2$ 는 L_2 놈 함수이고, f 는 최대-최소 정규화 함수이다.)

에 따라 획득하고,

상기 소프트 와핑부는

상기 제1 특징맵(f_{RGB})과 상기 제2 특징맵(f_{IR}), 상기 제1 마스크(M_{RGB})로부터 상기 제1 와핑 특징맵(\hat{f}_{RGB})을 수학식

$$\hat{f}_{RGB}(\mathbf{p}) = M_{RGB}(\mathbf{p})\mathcal{W}(f_{IR}(\mathbf{p})) + (1 - M_{RGB}(\mathbf{p}))f_{RGB}(\mathbf{p})$$

(여기서 \mathcal{W} 는 매칭 확률맵($P(p, q)$)에 따라 $\mathcal{W}(f_{IR}(\mathbf{p})) = \sum_q P(\mathbf{p}, \mathbf{q})f_{IR}(\mathbf{q})$ 로 나타나는 소프트 와핑 함수이다.)

에 따라 획득하는 사람 재식별 장치.

청구항 6

제4항에 있어서, 상기 손실 계산부는

상기 제1 및 제2 이미지 모두에서 모두 확인 가능한 사람 영역을 강조하는 제1 및 제2 공통 주의맵을 획득하고, 상기 제1 공통 주의맵에서 앵커인 상기 제1 특징맵의 식별자와 동일한 식별자로 추정되는 제1 와핑 특징맵인 포지티브 및 서로 다른 식별자로 추정되는 제1 와핑 특징맵인 네거티브 각각의 사이의 차이와 상기 제2 공통 주의맵에서 앵커인 상기 제2 특징맵의 식별자와 동일한 식별자로 추정되는 포지티브 및 서로 다른 식별자로 추정되는 제2 와핑 특징맵인 네거티브 각각의 사이의 차이를 누적하여 상기 밀집 삼중 손실을 계산하는 사람 재식별 장치.

청구항 7

제6항에 있어서, 상기 손실 계산부는

상기 제1 및 제2 공통 주의맵(A_i , $i \in \{RGB, IR\}$)을 기반으로 상기 밀집 삼중 손실(\mathcal{L}_{DT})을 수학식

$$\mathcal{L}_{DT} = \sum_{i \in \{RGB, IR\}} \sum_{\mathbf{p}} A_i(\mathbf{p}) [d_i^+(\mathbf{p}) - d_i^-(\mathbf{p}) + \alpha]_+$$

(여기서 α 는 미리 정의된 마진을 나타내고, $d_i^+(\mathbf{p})$ 와 $d_i^-(\mathbf{p})$ 는 각각

$d_i^+(\mathbf{p}) = \|f_i^a(\mathbf{p}) - \hat{f}_i^p(\mathbf{p})\|_2$, $d_i^-(\mathbf{p}) = \|f_i^a(\mathbf{p}) - \hat{f}_i^n(\mathbf{p})\|_2$ 로 계산되는 앵커(f_i^a)와 포지티브(f_i^p) 사이의 오차 및 앵커(f_i^a)와 네거티브(f_i^n) 각각 사이의 오차를 나타낸다.)

에 따라 계산하는 사람 재식별 장치.

청구항 8

제7항에 있어서, 상기 손실 계산부는

상기 제1 및 제2 스코어맵 각각과 학습 시에 상기 제1 및 제2 이미지에 레이블되어 인가되는 식별자 사이의 차이에 따른 식별 손실과 상기 제1 및 제2 스코어맵과 상기 제1 및 제2 와핑 스코어맵 사이의 차이에 따른 일관성 손실을 더 계산하고, 상기 식별 손실과 상기 일관성 손실 및 상기 밀집 삼중 손실을 가중합하여 계산되는 총 손실을 역전파하여 학습을 수행하는 사람 재식별 장치.

청구항 9

제8항에 있어서, 상기 손실 계산부는

상기 제1 및 제2 스코어맵과 상기 제1 및 제2 이미지에 레이블되어 인가되는 식별자 사이의 음의 크로스 엔트로피를 계산하여 상기 식별 손실을 획득하고, 상기 제1 및 제2 스코어맵 각각과 상기 제1 및 제2 와핑 스코어맵 중 대응하는 와핑 스코어맵 사이의 음의 크로스 엔트로피를 계산하여 상기 일관성 손실을 계산하는 사람 재식별 장치.

청구항 10

제1항에 있어서, 상기 제1 이미지는 가시 카메라에서 획득된 RGB 이미지이고 상기 제2 이미지는 IR 카메라에서 획득된 IR 이미지인 사람 재식별 장치.

청구항 11

학습 단계; 및

인공 신경망을 이용하여, 서로 다른 이종의 센서를 이용하여 획득된 이미지들을 인가받아 각 이미지에 포함된 사람에 대한 식별자를 획득하는 사람 재식별 단계를 포함하고,

상기 학습 단계는

학습 데이터로서 서로 다른 이종의 센서를 이용하여 획득되고 식별자가 미리 레이블된 제1 및 제2 이미지를 인가받아 학습되는 방식에 따라 특징을 추출하여 제1 및 제2 특징맵을 획득하고, 기지정된 방식으로 풀링하여 제1 및 제2 표현자를 획득하는 단계;

상기 제1 및 제2 특징맵을 인가받아 서로 대응하는 픽셀을 추정하고, 추정된 대응 픽셀의 위치를 기반으로 상호 와핑을 수행하여, 상기 제2 특징맵이 상기 제1 특징맵에 대응하도록 와핑된 제1 와핑 특징맵과 상기 제1 특징맵이 상기 제2 특징맵에 대응하도록 와핑된 제2 와핑 특징맵 획득하며, 상기 제1 및 제2 와핑 특징맵을 풀링하여 제1 및 제2 와핑 표현자를 획득하는 단계;

상기 제1 및 제2 표현자와 상기 제1 및 제2 와핑 표현자 각각이 기지정된 다수의 클래스 각각에 대응하는 확률을 나타내는 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵을 획득하고, 획득된 제1 및 제2 스코어맵에서 가장 높은 확률의 클래스의 식별자를 획득하는 단계; 및

상기 제1 특징맵과 상기 제1 와핑 특징맵 사이의 차이와 상기 제2 특징맵과 상기 제2 와핑 특징맵 사이의 차이를 동일한 식별자인 경우의 포지티브와 서로 다른 식별자인 경우의 네거티브로 구분하여 손실을 계산하는 단계를 포함하되,

상기 제1 및 제2 와핑 표현자를 획득하는 단계는

기지정된 방식에 따라 상기 제1 및 제2 특징맵 사이의 픽셀간 유사도를 계산하여 유사도맵을 획득하는 단계;

상기 유사도 맵을 기반으로 상기 제1 및 제2 특징맵의 각 픽셀별 매칭 확률을 계산하여 매칭 확률맵을 획득하는 단계;

상기 매칭 확률맵을 이용하여 상기 제1 및 제2 특징맵 각각의 픽셀들이 상대 특징맵의 각 픽셀에 매칭될 확률을 가중 집계하여 상대 특징맵으로 와핑함으로써 상기 제1 및 제2 와핑 특징맵을 획득하는 단계; 및

제1 및 제2 와핑 표현자를 획득하기 위해 기지정된 방식으로 상기 제1 및 제2 와핑 특징맵을 풀링하는 단계를 포함하는 사람 재식별 방법.

청구항 12

삭제

청구항 13

제11항에 있어서, 상기 제1 및 제2 와핑 표현자를 획득하는 단계는

상기 제1 및 제2 특징맵의 각 위치별 픽셀의 크기를 정규화하여 제1 및 제2 마스크를 획득하는 단계를 더 포함하는 사람 재식별 방법.

청구항 14

제13항에 있어서, 상기 제1 및 제2 와핑 특징맵을 획득하는 단계는

상기 제1 및 제2 마스크를 인가받아 상기 제1 및 제2 특징맵 각각의 픽셀들이 상대 특징맵의 각 픽셀에 매칭될 확률을 가중 집계하는 단계; 및

가중 집계된 와핑 픽셀값에 마스크의 픽셀값을 추가로 가중하여 와핑하는 단계를 포함하는 사람 재식별 방법.

청구항 15

제14항에 있어서, 상기 마스크를 획득하는 단계는

상기 제1 특징맵(f_{RGB})으로부터 상기 제1 마스크(M_{RGB})를 수학식

$$M_{RGB}(\mathbf{p}) = f(\|\mathbf{f}_{RGB}(\mathbf{p})\|_2)$$

(여기서 $\|\cdot\|_2$ 는 L_2 놈 함수이고, f 는 최대-최소 정규화 함수이다.)

에 따라 획득하고,

상기 추가로 가중하여 와핑하는 단계는

상기 제1 특징맵(f_{RGB})과 상기 제2 특징맵(f_{IR}), 상기 제1 마스크(M_{RGB})로부터 상기 제1 와핑 특징맵(\hat{f}_{RGB})을 수학식

$$\hat{f}_{RGB}(\mathbf{p}) = M_{RGB}(\mathbf{p})\mathcal{W}(f_{IR}(\mathbf{p})) + (1 - M_{RGB}(\mathbf{p}))f_{RGB}(\mathbf{p})$$

(여기서 \mathcal{W} 는 매칭 확률맵($P(p,q)$)에 따라 $\mathcal{W}(f_{IR}(\mathbf{p})) = \sum_q P(\mathbf{p}, \mathbf{q})f_{IR}(\mathbf{q})$ 로 나타나는 소프트 와핑 함수이다.)

에 따라 획득하는 사람 재식별 방법.

청구항 16

제14항에 있어서, 상기 손실을 계산하는 단계는

상기 제1 및 제2 이미지 모두에서 모두 확인 가능한 사람 영역을 강조하는 제1 및 제2 공통 주의맵을 획득하는 단계; 및

상기 제1 공통 주의맵에서 앵커인 상기 제1 특징맵의 식별자와 동일한 식별자로 추정되는 제1 와핑 특징맵인 포지티브 및 서로 다른 식별자로 추정되는 제1 와핑 특징맵인 네거티브 각각의 사이의 차이와 상기 제2 공통 주의맵에서 앵커인 상기 제2 특징맵의 식별자와 동일한 식별자로 추정되는 포지티브 및 서로 다른 식별자로 추정되는 제2 와핑 특징맵인 네거티브 각각의 사이의 차이를 누적하여 계산되는 밀집 삼중 손실을 획득하는 단계를 포함하는 사람 재식별 방법.

청구항 17

제16항에 있어서, 상기 밀집 삼중 손실을 획득하는 단계는

상기 제1 및 제2 공통 주의맵(A_i , $i \in \{RGB, IR\}$)을 기반으로 상기 밀집 삼중 손실(\mathcal{L}_{DT})을 수학식

$$\mathcal{L}_{DT} = \sum_{i \in \{RGB, IR\}} \sum_{\mathbf{p}} A_i(\mathbf{p}) [d_i^+(\mathbf{p}) - d_i^-(\mathbf{p}) + \alpha]_+$$

(여기서 α 는 미리 정의된 마진을 나타내고, $d_i^+(\mathbf{p})$ 와 $d_i^-(\mathbf{p})$ 는 각각 $d_i^+(\mathbf{p}) = \|\mathbf{f}_i^a(\mathbf{p}) - \hat{\mathbf{f}}_i^p(\mathbf{p})\|_2$, $d_i^-(\mathbf{p}) = \|\mathbf{f}_i^a(\mathbf{p}) - \hat{\mathbf{f}}_i^n(\mathbf{p})\|_2$ 로 계산되는 앵커(f_i^a)와 포지티브(f_i^p) 사이의 오차 및 앵커(f_i^a)와 네거티브(f_i^n) 각각 사이의 오차를 나타낸다.)

에 따라 계산하는 사람 재식별 방법.

청구항 18

제17항에 있어서, 상기 손실을 계산하는 단계는

상기 제1 및 제2 스코어맵 각각과 학습 시에 상기 제1 및 제2 이미지에 레이블되어 인가되는 식별자 사이의 차이에 따른 식별 손실을 계산하는 단계;

상기 제1 및 제2 스코어맵과 상기 제1 및 제2 와핑 스코어맵 사이의 차이에 따른 일관성 손실을 계산하는 단계; 및

상기 식별 손실과 상기 일관성 손실 및 상기 밀집 삼중 손실을 가중합하여 계산되는 총 손실을 역전파하는 단계를 더 포함하는 사람 재식별 방법.

청구항 19

제18항에 있어서, 상기 식별 손실을 계산하는 단계는

상기 제1 및 제2 스코어맵과 상기 제1 및 제2 이미지에 레이블되어 인가되는 식별자 사이의 음의 크로스 엔트로피를 계산하여 상기 식별 손실을 획득하고,

상기 일관성 손실을 계산하는 단계는

상기 제1 및 제2 스코어맵 각각과 상기 제1 및 제2 와핑 스코어맵 중 대응하는 와핑 스코어맵 사이의 음의 크로스 엔트로피를 계산하여 상기 일관성 손실을 계산하는 사람 재식별 방법.

청구항 20

제11항에 있어서, 상기 사람 재식별 단계는

상기 학습 단계에서 미리 학습된 방식에 따라 인가되는 이종 이미지 각각의 특징을 추출하여 특징맵을 획득하는 단계;

이종 이미지 각각으로부터 획득된 특징맵을 기지정된 방식으로 풀링하여 대응하는 표현자를 획득하는 단계; 및

미리 학습된 방식에 따라 획득된 표현자 각각에 대응하는 스코어맵을 획득하고, 획득된 스코어맵을 기반으로 인가된 이종 이미지 각각에 포함된 사람에 대한 식별자를 획득하는 단계를 포함하는 사람 재식별 방법.

발명의 설명

기술 분야

[0001] 본 발명은 사람 재식별 장치 및 방법에 관한 것으로, 이종 센서 카메라 기반 사람 재식별 장치 및 방법에 관한 것이다.

배경 기술

[0002] 서로 다른 환경에서 촬영된 동일한 사람을 탐색할 수 있는 사람 재식별(Person Re-identification: reID 라고 함) 기술에 대한 연구가 활발하게 이루어지고 있다.

[0003] 도 1은 사람 재식별의 개념을 나타내고, 도 2는 사람 재식별을 이용하는 분야 중 하나로 사람 추적 기술의 일 예를 나타낸다.

[0004] 도 1에 도시된 바와 같이, 사람 재식별은 서로 다른 조건에서 촬영된 동일한 사람을 식별하여 검출하는 기법으로서, 자세의 변화, 배경의 변화, 조명의 변화, 촬영 거리 및 각도 변화와 같이 여러 환경 조건이 변화하여도

동일한 사람이 포함된 이미지를 정확하게 검출하는 것을 목적으로 한다.

- [0005] 이러한 사람 재식별 기술은 도 2에 도시된 바와 같이, 실종자 탐색 또는 범죄자 추적 등 다수의 이미지에서 특정인을 탐색 및 추적하는 다양한 분야에 이용될 수 있다.
- [0006] 이러한 사람 재식별 기술은 이미지로부터 주로 인공 신경망을 이용하여 개인을 식별하기 위해 사람의 특정 부분이 아닌 정체성을 나타내는 특징을 추출하고, 추출된 특징을 기반으로 동일한 사람을 식별할 수 있도록 한다. 즉 사람의 얼굴 형태나 특정한 자세, 배경, 촬영 각도, 조명 등에 한정되지 않고 개개인 각각의 정체성을 나타내는 특징을 추출하여 사람을 재식별함으로써, 서로 다른 다수의 이미지에서 동일한 사람을 검출할 수 있도록 한다. 다만 기존의 사람 재식별은 주로 동일한 종류의 센서 카메라에서 획득된 이미지에서 동일한 사람을 재식별하기 위해 집중되어 있었다.
- [0007] 도 3은 이중 센서 카메라에서 획득된 이미지의 일 예를 나타내고, 도 4는 이중 센서 카메라 기반 사람 재식별 개념을 나타낸다.
- [0008] 도 3과 상단 3개의 이미지 중 왼쪽 2개의 이미지 각각 일반적인 가시 카메라(RGB camera 또는 Visible camera)에서 주간 및 야간에 획득된 이미지이고, 오른쪽 이미지는 적외선(Infrared: 이하 IR) 카메라에서 야간에 획득된 이미지를 나타낸다.
- [0009] 가운데 좌측 이미지는 주간에 가시 카메라에서 획득된 이미지로부터 추출된 사람 영역을 나타내고, 우측 다수의 이미지는 재식별을 위해 비교 대상이 되는 이미지를 나타낸다. 그리고 하단 우측 이미지는 야간에 IR 카메라에서 획득된 이미지로부터 추출된 사람 영역을 나타내고, 좌측 다수의 이미지는 재식별을 위해 비교 대상이 되는 이미지를 나타낸다.
- [0010] 도 3에 도시된 바와 같이, 가시 카메라는 주간에 이미지를 획득하는 경우, 사람 식별이 가능한 수준의 이미지를 획득할 수 있으나, 야간에는 주변 환경과 구분되지 않아 사람 식별이 불가능하다. 그에 반해, 가시 카메라와 다른 IR 센서를 이용하는 IR 카메라를 이용하는 경우, 야간에 사람을 정확하게 식별할 수 있다.
- [0011] 따라서 도 4에 도시된 바와 같이, 환경 변화에 무관하게 사람을 식별하고 추적할 수 있도록 하기 위해서는 서로 다른 이중 센서 카메라에서 획득된 이미지에서 사람을 재식별할 수 있어야 한다.
- [0012] 도 5은 사람 재식별에서의 오정렬 문제를 설명하기 위한 도면이다.
- [0013] 도 5에서 상단과 하단에서 동일한 행에 위치한 이미지는 동일한 객체, 즉 동일한 사람에 대한 이미지이다. 도 5에 도시된 바와 같이, 동일한 가시 카메라로 획득된 이미지에서도 사람은 여러 조건에 따라 매우 상이한 형태로 나타날 수 있으며, 이로 인해 각 이미지에서 동일한 사람이 식별하기 어렵게 될 수 있다. 다만 동일한 가시 카메라에서 획득된 이미지의 경우, 컬러 분포 등에 따른 특징을 기반으로도 사람을 재식별할 수도 있다.
- [0014] 그러나 가시 카메라와 달리 IR 카메라의 경우, 색상 정보가 포함되지 않으므로 가시 카메라에서 획득된 컬러 이미지의 특징과 IR 카메라에서 획득된 IR 이미지의 특징 사이의 비교가 용이하지 않다. 즉 재식별을 위해 서로 비교할 수 특징이 제한된다. 이와 같은 모달리티 차이를 극복하기 위해 기존에는 RGB 이미지와 IR 이미지를 모두 가상의 특징 공간(common feature space)에 임베딩(embedding)하고, 특징 공간에 임베딩된 특징을 이미지 또는 파트 수준으로 추출하여 비교하는 방식으로 학습을 수행한다. 그러나 이와 같은 공통 특징 공간에 임베딩하는 방식에서는 특징점들에 대한 밀집 정합 오류로 인한 객체 오정렬(misalignment)에 의한 문제가 발생할 수 있다.
- [0015] 기존에는 서로 다른 이중 센서 카메라에서 획득된 이미지에서 사람을 재식별하는 경우, 각 이미지에 포함된 사람이 서로 정렬되어 있다고 가정하여 이미지 또는 파트 수준에서의 특징을 학습하는데 집중하였다. 즉 두 이미지가 개략적일지라도 서로 일정 수준 이상 정렬된 것으로 가정하여 학습을 수행하였다. 일 예로 이미지를 기지 정된 횡방향 그리드로 구분하고 두 이미지에서 각 그리드에 포함되는 신체 부위(예를 들면, 어깨, 무릎 등)가 서로 대응하도록 정렬된 것으로 가정하여 학습된다.
- [0016] 그러나 도 5와 같이 비록 동일 객체에 대한 이미지일지라도 실제로 이미지 내에서 객체의 위치, 크기, 포즈 등은 매우 상이한 경우가 많으며, 이와 같이 두 이미지가 정렬되지 않은 상태에서 학습이 수행되면 오정렬로 인해 학습이 정상적으로 수행되지 않아 사람 재식별 오차가 크게 발생하는 문제가 발생할 수 있다. 상기한 바와 같이, 이러한 오정렬 문제는 특히 특성이 다른 두 모달리티 간 영상의 객체 재식별에서 문제가 더 크게 나타날 수 있다.

[0017] 또한 공통 특징 공간에 임베딩하는 방식에서는 객체 재식별에 크게 도움이 될 수 있는 세분화된 로컬 정보가 간과되는 경향이 있어 재식별 성능에 한계가 있다.

선행기술문헌

특허문헌

[0018] (특허문헌 0001) 한국 공개 특허 제10-2019-0078270호 (2019.07.04 공개)

발명의 내용

해결하려는 과제

[0019] 본 발명의 목적은 이중 센서 카메라에서 획득된 이미지로부터 사람을 정확하게 재식별할 수 있는 사람 재식별 장치 및 방법을 제공하는데 있다.

[0020] 본 발명의 다른 목적은 정렬되지 않은 이미지에서도 교차 모달리티 객체 이미지간 밀집 정합을 이용하여 정확하게 사람을 재식별할 수 있는 사람 재식별 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0021] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 사람 재식별 장치는 서로 다른 이중의 센서를 이용하여 획득된 제1 및 제2 이미지를 인가받아 학습되는 방식에 따라 특징을 추출하여 제1 및 제2 특징맵을 획득하고, 기지정된 방식으로 풀링하여 제1 및 제2 표현자를 획득하는 특징 풀링부; 학습 시에 구비되어, 상기 제1 및 제2 특징맵을 인가받아 서로 대응하는 픽셀을 추정하고, 추정된 대응 픽셀의 위치를 기반으로 상호 와핑을 수행하여, 상기 제2 특징맵이 상기 제1 특징맵에 대응하도록 와핑된 제1 와핑 특징맵과 상기 제1 특징맵이 상기 제2 특징맵에 대응하도록 와핑된 제2 와핑 특징맵 획득하며, 상기 제1 및 제2 와핑 특징맵을 풀링하여 제1 및 제2 와핑 표현자를 획득하는 교차 모달 정렬부; 학습되는 방식에 따라 상기 제1 및 제2 표현자와 상기 제1 및 제2 와핑 표현자 각각이 기지정된 다수의 클래스 각각에 대응하는 확률을 나타내는 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵을 획득하고, 획득된 제1 및 제2 스코어맵에서 가장 높은 확률의 클래스의 식별자를 획득하는 재식별부; 및 학습 시에 구비되어, 상기 제1 특징맵과 상기 제1 와핑 특징맵 사이의 차이와 상기 제2 특징맵과 상기 제2 와핑 특징맵 사이의 차이를 동일한 식별자인 경우의 포지티브와 서로 다른 식별자인 경우의 네거티브로 구분하여 밀집 삼중 손실을 계산하는 손실 계산부를 포함한다.

[0022] 상기 교차 모달 정렬부는 기지정된 방식에 따라 상기 제1 및 제2 특징맵 사이의 픽셀간 유사도를 계산하여 유사도맵을 획득하는 유사도 측정부; 상기 유사도 맵을 기반으로 상기 제1 및 제2 특징맵의 각 픽셀별 매칭 확률을 계산하여 매칭 확률맵을 획득하는 매칭 확률 추정부; 상기 매칭 확률맵을 이용하여 상기 제1 및 제2 특징맵 각각의 픽셀들이 상대 특징맵의 각 픽셀에 매칭될 확률을 가중 집계하여 상대 특징맵으로 와핑함으로써 상기 제1 및 제2 와핑 특징맵을 획득하는 소프트 와핑부; 및 상기 제1 및 제2 와핑 특징맵을 기지정된 방식으로 풀링하여 제1 및 제2 와핑 표현자를 획득하는 와핑 풀링부를 포함할 수 있다.

[0023] 상기 교차 모달 정렬부는 상기 제1 및 제2 특징맵의 각 위치별 픽셀의 크기를 정규화하여 제1 및 제2 마스크를 획득하는 마스크 제공부를 더 포함할 수 있다.

[0024] 상기 소프트 와핑부는 상기 제1 및 제2 마스크를 인가받아 상기 제1 및 제2 특징맵 각각의 픽셀들이 상대 특징맵의 각 픽셀에 매칭될 확률을 가중 집계하고, 가중 집계된 와핑 픽셀값에 마스크의 픽셀값을 추가로 가중하여 와핑할 수 있다.

[0025] 상기 손실 계산부는 상기 제1 및 제2 이미지 모두에서 모두 확인 가능한 사람 영역을 강조하는 제1 및 제2 공통 주의맵을 획득하고, 상기 제1 공통 주의맵에서 앵커인 상기 제1 특징맵의 식별자와 동일한 식별자로 추정되는 제1 와핑 특징맵인 포지티브 및 서로 다른 식별자로 추정되는 제1 와핑 특징맵인 네거티브 각각의 사이의 차이와 상기 제2 공통 주의맵에서 앵커인 상기 제2 특징맵의 식별자와 동일한 식별자로 추정되는 포지티브 및 서로 다른 식별자로 추정되는 제2 와핑 특징맵인 네거티브 각각의 사이의 차이를 누적하여 상기 밀집 삼중 손실을 계산할 수 있다.

[0026] 상기 손실 계산부는 상기 제1 및 제2 스코어맵 각각과 학습 시에 상기 제1 및 제2 이미지에 레이블되어 인가되

는 식별자 사이의 차이에 따른 식별 손실과 상기 제1 및 제2 스코어맵과 상기 제1 및 제2 와핑 스코어맵 사이의 차이에 따른 일관성 손실을 더 계산하고, 상기 식별 손실과 상기 일관성 손실 및 상기 밀집 삼중 손실을 가중합하여 계산되는 총 손실을 역전과하여 학습을 수행할 수 있다.

[0027] 상기 손실 계산부는 상기 제1 및 제2 스코어맵과 상기 제1 및 제2 이미지에 레이블되어 인가되는 식별자 사이의 음의 크로스 엔트로피를 계산하여 상기 식별 손실을 획득하고, 상기 제1 및 제2 스코어맵 각각과 상기 제1 및 제2 와핑 스코어맵 중 대응하는 와핑 스코어맵 사이의 음의 크로스 엔트로피를 계산하여 상기 일관성 손실을 계산할 수 있다.

[0028] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 사람 재식별 방법은 학습 단계; 및 인공 신경망을 이용하여, 서로 다른 이종의 센서를 이용하여 획득된 이미지들을 인가받아 각 이미지에 포함된 사람에 대한 식별자를 획득하는 사람 재식별 단계를 포함하고, 상기 학습 단계는 학습 데이터로서 서로 다른 이종의 센서를 이용하여 획득되고 식별자가 미리 레이블된 제1 및 제2 이미지를 인가받아 학습되는 방식에 따라 특징을 추출하여 제1 및 제2 특징맵을 획득하고, 기지정된 방식으로 풀링하여 제1 및 제2 표현자를 획득하는 단계; 상기 제1 및 제2 특징맵을 인가받아 서로 대응하는 픽셀을 추정하고, 추정된 대응 픽셀의 위치를 기반으로 상호 와핑을 수행하여, 상기 제2 특징맵이 상기 제1 특징맵에 대응하도록 와핑된 제1 와핑 특징맵과 상기 제1 특징맵이 상기 제2 특징맵에 대응하도록 와핑된 제2 와핑 특징맵 획득하며, 상기 제1 및 제2 와핑 특징맵을 풀링하여 제1 및 제2 와핑 표현자를 획득하는 단계; 상기 제1 및 제2 표현자와 상기 제1 및 제2 와핑 표현자 각각이 기지정된 다수의 클래스 각각에 대응하는 확률을 나타내는 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵을 획득하고, 획득된 제1 및 제2 스코어맵에서 가장 높은 확률의 클래스의 식별자를 획득하는 단계; 및 상기 제1 특징맵과 상기 제1 와핑 특징맵 사이의 차이와 상기 제2 특징맵과 상기 제2 와핑 특징맵 사이의 차이를 동일한 식별자인 경우의 포지티브와 서로 다른 식별자인 경우의 네거티브로 구분하여 손실을 계산하는 단계를 포함한다.

발명의 효과

[0029] 따라서, 본 발명의 실시예에 따른 사람 재식별 장치 및 방법은 이종 센서 카메라에서 획득된 이미지들에서 추출된 특징맵과 함께 특징맵에서 서로 대응하는 픽셀을 와핑한 와핑 특징맵을 기반으로 학습을 수행하며, 특히 특징맵과 와핑 특징맵에 대한 밀집 삼중 손실을 계산하여 학습에 반영함으로써, 정렬되지 않은 이미지들에서도 세분화된 로컬 정보를 추출할 수 있어 다양한 카메라에서 매우 상이한 조건에서 획득된 이미지로부터도 사람을 정확하게 재식별할 수 있다.

도면의 간단한 설명

- [0030] 도 1은 사람 재식별의 개념을 나타낸다.
- 도 2는 사람 재식별을 이용하는 분야 중 하나로 사람 추적 기술의 일 예를 나타낸다.
- 도 3은 이종 센서 카메라에서 획득된 이미지의 일 예를 나타낸다.
- 도 4는 이종 센서 카메라 기반 사람 재식별 개념을 나타낸다.
- 도 5는 사람 재식별에서의 오정렬 문제를 설명하기 위한 도면이다.
- 도 6은 본 발명의 일 실시예에 따른 이종 센서 카메라 기반 사람 재식별 장치의 개략적 구조를 나타낸다.
- 도 7은 도 6의 사람 재식별 장치의 동작을 설명하기 위한 도면이다.
- 도 8은 공통 주의맵을 설명하기 위한 도면이다.
- 도 9는 사람 재식별 장치의 손실 설정에 따른 성능 변화를 나타낸다.
- 도 10은 본 실시예에 따른 사람 재식별 장치의 성능을 시뮬레이션한 결과를 나타낸다.
- 도 11은 본 발명의 일 실시예에 따른 이종 센서 카메라 기반 사람 재식별 방법을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0031] 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.

[0032] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러

나, 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.

- [0033] 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.
- [0034] 도 6은 본 발명의 일 실시예에 따른 이중 센서 카메라 기반 사람 재식별 장치의 개략적 구조를 나타내고, 도 7은 도 6의 사람 재식별 장치의 동작을 설명하기 위한 도면이며, 도 8은 공통 주의맵을 설명하기 위한 도면이다.
- [0035] 도 6 및 도 7을 참조하면, 본 실시예의 이중 센서 카메라 기반 사람 재식별 장치는 이중 이미지 획득부(100), 특징 추출부(200), 풀링부(300), 재식별부(400), 교차 모달 정렬부(500) 및 손실 계산부(600)를 포함할 수 있다.
- [0036] 이중 이미지 획득부(100)는 서로 다른 센서 카메라에서 획득된 이중 이미지를 획득한다. 이중 이미지 획득부(100)는 제1 이미지와 제2 이미지를 획득할 수 있으며, 여기서는 일 예로 제1 이미지는 가시 카메라를 이용하여 획득된 RGB 이미지이고, 제2 이미지는 IR 카메라를 이용하여 획득된 IR 이미지인 것으로 가정하여 설명한다. 이중 이미지 획득부(100)는 제1 이미지를 획득하는 제1 이미지 획득부(110)와 제2 이미지를 획득하는 제2 이미지 획득부(120)를 포함할 수 있다. 다만 본 실시예의 이중 센서 카메라 기반 사람 재식별 장치는 가시 카메라나 IR 카메라 이외에 다른 종류의 센서 카메라를 이용하여 이미지를 획득할 수도 있다. 이때, 제1 이미지와 제2 이미지는 각각 단일 프레임의 이미지일 수도 있으나, 다수 프레임으로 구성된 이미지일 수 있으며, 동일한 사람이 포함되거나 다른 사람이 포함된 이미지일 수 있다.
- [0037] 제1 이미지 획득부(110)와 제2 이미지 획득부(120)는 각각 서로 다른 센서를 포함하는 이중 센서 카메라로 구현되거나, 이중 센서 카메라를 이용하여 획득된 이미지를 저장하는 저장 장치 또는 외부에서 이중 센서 카메라를 이용하여 획득된 이미지를 인가받는 통신 모듈 등으로 구현될 수도 있다.
- [0038] 특징 추출부(200)는 이중 이미지 획득부(100)에서 획득된 이중 이미지 각각으로부터 특징을 추출하여 특징맵을 획득한다. 특징 추출부(200)는 각각 CNN(Convolutional neural network)과 같은 인공 신경망으로 구현되는 제1 특징 추출부(210)와 제2 특징 추출부(220)를 포함할 수 있다. 제1 특징 추출부(210)는 제1 이미지 획득부(110)에서 획득된 제1 이미지를 인가받아 미리 학습된 방식에 따라 특징을 추출하여 제1 특징맵(f_{RGB})을 획득하고, 제2 특징 추출부(220)는 제2 이미지 획득부(120)에서 획득된 제2 이미지를 인가받아 미리 학습된 방식에 따라 특징을 추출하여 제2 특징맵(f_{IR})을 획득한다.
- [0039] 이때 제1 특징맵($f_{RGB} \in \mathbb{R}^{h \times w \times d}$)과 제2 특징맵($f_{IR} \in \mathbb{R}^{h \times w \times d}$)은 동일한 높이(h)와 폭(w) 및 채널(d)을 갖는 $h \times w \times d$ 크기로 추출될 수 있다.
- [0040] 풀링부(300)는 특징 추출부(200)에서 획득된 제1 및 제2 특징맵(f_{RGB} , f_{IR})을 인가받아 각각 미리 지정된 방식으로 풀링하여, 제1 및 제2 표현자를 획득한다. 풀링부(300)는 제1 풀링부(310)와 제2 풀링부(320)를 포함할 수 있다. 제1 풀링부(310)는 제1 특징맵(f_{RGB})을 인가받아 풀링하여 제1 표현자($\phi(f_{RGB})$)를 획득하고, 제2 풀링부(320)는 제2 특징맵(f_{IR})을 인가받아 풀링하여 제2 표현자($\phi(f_{IR})$)를 획득할 수 있다. 여기서 제1 및 제2 표현자($\phi(f_{RGB})$, $\phi(f_{IR})$) 각각은 채널(d) 길이에 대응하는 1차원 벡터 형태를 갖고, 이미지에 포함된 사람을 식별할 수 있도록 하는 특징 표현자로 볼 수 있다.
- [0041] 여기서 특징 추출부(200)와 풀링부(300)는 특징 풀링부로 통합될 수 있다.
- [0042] 재식별부(400)는 이미지에 포함된 사람의 특징 표현자인 제1 및 제2 표현자($\phi(f_{RGB})$, $\phi(f_{IR})$)를 인가받아, 제1 및 제2 이미지에 포함된 사람을 재식별한다. 재식별부(400)는 일 예로 제1 및 제2 표현자($\phi(f_{RGB})$, $\phi(f_{IR})$) 사이의 유사도를 측정하여 제1 및 제2 이미지에 포함된 사람이 동일한 사람인지 여부를 판별할 수 있다. 그러나 재식별부(400)는 미리 지정된 다수의 사람에 대응하도록 설정된 다수의 클래스 중 제1 및 제2 표현자($\phi(f_{RGB})$, $\phi(f_{IR})$)에 대응하는 클래스를 확인하여 식별자를 획득하도록 구성될 수도 있다. 즉 재식별부(400)는 이중 이미

지 각각에 포함된 사람의 식별자를 획득하여 동일한 사람인지 여부를 확인할 수 있다.

[0043] 여기서 재식별부(400)는 일 예로 소프트 맥스와 같은 활성화 함수와 완전 연결 레이어(Fully Connected Layer)를 포함하는 구성되는 인공 신경망으로 구현되어, 제1 및 제2 표현자($\phi(f_{RGB})$, $\phi(f_{IR})$) 각각이 기지정된 다수의 클래스 각각에 대응하는 확률을 나타내는 스코어맵을 획득하고, 획득된 스코어맵에서 가장 높은 확률의 클래스를 확인하여 식별자를 선택할 수 있다.

[0044] 도 6에서는 설명의 편의를 위하여 특징 추출부(200)와 풀링부(300) 및 재식별부(400)를 별도의 구성으로 도시하였으나, 특징 추출부(200)와 풀링부(300) 및 재식별부(400)는 사람을 재식별하기 위한 재식별 인공 신경망에서 서로 다른 기능을 수행하는 레이어들로 구성될 수 있다.

[0045] 사람 재식별 장치는 실제 적용 시에 이중 이미지 획득부(100), 특징 추출부(200), 풀링부(300) 및 재식별부(400)만으로 구성될 수 있다. 다만, 사람 재식별 장치는 실제 적용되기 이전에 인공 신경망으로 구성되는 특징 추출부(200), 풀링부(300) 및 재식별부(400) 등이 학습되어야만 한다. 이에 본 실시예의 사람 재식별 장치는 학습을 수행하는 동안 교차 모달 정렬부(500) 및 손실 계산부(600)를 더 포함할 수 있으며, 학습이 완료되면 제외할 수 있다.

[0046] 교차 모달 정렬부(500)는 서로 다른 이중 이미지에서 추출된 제1 및 제2 특징맵(f_{RGB} , f_{IR})을 인가받고, 인가된 제1 특징맵(f_{RGB})과 제2 특징맵(f_{IR})의 픽셀간 유사도에 기반하여 서로 대응하는 픽셀을 추정하고, 추정된 대응 픽셀의 위치를 기반으로 상호 와핑을 수행하여, 제1 및 제2 와핑 특징맵(\hat{f}_{RGB} , \hat{f}_{IR})을 획득한다. 여기서 제1 와핑 특징맵(\hat{f}_{RGB})은 제2 이미지에서 추출된 제2 특징맵(f_{IR})을 와핑하여 획득되는 특징맵이고, 제2 와핑 특징맵(\hat{f}_{IR})은 제1 이미지에서 추출된 제1 특징맵(f_{RGB})을 와핑하여 획득되는 특징맵이다. 즉 제1 및 제2 와핑 특징맵(\hat{f}_{RGB} , \hat{f}_{IR})은 서로 상대의 이중 이미지로부터 추출된 특징맵을 와핑하여 모달리티를 교차하여 획득된다.

[0047] 구체적으로 살펴보면, 교차 모달 정렬부(500)는 유사도 측정부(510), 매칭 확률 추정부(520), 마스크 제공부(530), 소프트 와핑부(540) 및 와핑 풀링부(550)를 포함할 수 있다.

[0048] 유사도 측정부(510)는 제1 및 제2 특징맵(f_{RGB} , f_{IR})을 인가받고, 인가된 제1 및 제2 특징맵(f_{RGB} , f_{IR}) 사이의 픽셀간 유사도를 계산하여 유사도맵($C(p,q)$)을 획득한다. 유사도 측정부(510)는 제1 특징맵(f_{RGB})의 $h \times w$ 평면 상에서 각 위치(p)별 채널 방향 길이(d)를 갖는 1차원 벡터인 픽셀 벡터와 제2 특징맵(f_{IR})의 $h \times w$ 평면 상에서 각 위치(q)별 채널 방향 길이(d)를 갖는 1차원 벡터인 픽셀 벡터 사이의 유사도를 수학적 식 1과 같이 코사인 유사도를 기반으로 계산하여 유사도맵($C(p,q)$)을 획득할 수 있다.

수학적 식 1

[0049]
$$C(p, q) = \frac{f_{RGB}(p)^T f_{IR}(q)}{\|f_{RGB}(p)\|_2 \|f_{IR}(q)\|_2}$$

[0050] 여기서 $f_{RGB}(p)$ 는 제1 특징맵(f_{RGB})의 위치(p)별 픽셀 벡터를 나타내고, $f_{IR}(q)$ 는 제2 특징맵(f_{IR})의 위치(q)별 픽셀 벡터를 나타내며, $\|\cdot\|_2$ 는 L2 놈 함수를 나타내며, T는 전치 행렬을 나타낸다.

[0051] 매칭 확률 추정부(520)는 유사도 측정부(510)에서 획득된 유사도맵($C(p,q)$)을 인가받아 각 픽셀별 일치 확률을 계산한다. 매칭 확률 추정부(520)는 제1 특징맵(f_{RGB})과 제2 특징맵(f_{IR})의 각 픽셀 벡터 사이의 유사성을 기반으로 제1 특징맵(f_{RGB})과 제2 특징맵(f_{IR})의 픽셀 위치별 매칭 확률을 나타내는 $h \times w \times h \times w$ 의 크기의 매칭 확률맵($P(p,q)$)을 수학적 식 2에 따라 계산하여 획득한다.

수학식 2

$$P(\mathbf{p}, \mathbf{q}) = \frac{\exp(\beta C(\mathbf{p}, \mathbf{q}))}{\sum_{\mathbf{q}'} \exp(\beta C(\mathbf{p}, \mathbf{q}'))}$$

[0052]

[0053] 여기서 β 는 온도 파라미터이다.

[0054] 마스크 제공부(530)는 제1 및 제2 특징맵(f_{RGB} , f_{IR})을 인가받아, 제1 특징맵(f_{RGB})에 대응하는 제1 마스크(M_{RGB})와 제2 특징맵(f_{IR})에 대응하는 제2 마스크(M_{IR})를 획득하여 소프트 와핑부(540)로 출력한다.

[0055] 마스크 제공부(530)는 학습이 수행됨에 따라 제1 및 제2 이미지에서 사람에 해당하는 영역, 즉 전경이 나머지 영역에 비해 특징이 두드러진다는 가정 하에, 실제 사람 영역에 대한 레이블이 없는 제1 이미지에서 획득된 제1 특징맵(f_{RGB})으로부터 사람 영역이 강조된 제1 마스크(M_{RGB})를 수학식 3에 따라 획득할 수 있다.

수학식 3

$$M_{RGB}(\mathbf{p}) = f(\|\mathbf{f}_{RGB}(\mathbf{p})\|_2)$$

[0056]

[0057] 여기서 f 는 수학식 4로 정의되는 최대-최소 정규화 함수이다.

수학식 4

$$f(x) = \frac{x - \min(x)}{\max(x) - \min(x)}$$

[0058]

[0059] 이와 유사하게 마스크 제공부(530)는 제2 이미지에서 획득된 제2 특징맵(f_{IR})으로부터 사람 영역에 대응하는 제2 마스크(M_{IR}) 또한 획득할 수 있다.

[0060] 즉 마스크 제공부(530)는 제1 및 제2 특징맵(f_{RGB} , f_{IR})의 각 위치별 픽셀 벡터의 크기를 정규화하여 마스크를 획득한다. 만일 마스크 제공부(530)가 이진 마스크를 이용하는 경우, 전경과 배경이 명확하게 구분될 수 있는 경우에는 유효하지만 명확하게 구분되지 않는 경우에는 오히려 큰 오차를 유발할 수 있게 된다. 이러한 오차가 발생하는 것을 방지하기 위해, 본 실시예의 마스크 제공부(530)는 제1 및 제2 특징맵(f_{RGB} , f_{IR})의 각 위치별 픽셀 벡터의 크기를 정규화하여 전경 영역이 배경 영역에 비해 강조되는 제1 및 제2 마스크(M_{RGB} , M_{IR})를 획득한다.

[0061] 한편 소프트 와핑부(540)는 수학식 2에 따라 획득된 픽셀간 매칭 확률맵($P(p,q)$)을 기반으로 제1 및 제2 특징맵(f_{RGB} , f_{IR})의 각 픽셀을 서로 대응하는 픽셀의 위치로 소프트 와핑함으로써 제1 및 제2 와핑 특징맵(\hat{f}_{RGB} , \hat{f}_{IR})을 획득한다.

[0062] 이때 소프트 와핑부(540)는 마스크 제공부(530)로부터 제공되는 제1 및 제2 마스크(M_{RGB} , M_{IR})를 이용하여, 제1 및 제2 특징맵(f_{RGB} , f_{IR})에서 사람에 해당하는 전경 영역에 집중적으로 와핑을 수행할 수 있다.

[0063] 소프트 와핑부(540)는 제2 특징맵(f_{IR})의 각 픽셀이 제1 특징맵(f_{RGB})의 사람에 해당하는 전경 영역이 강조되어 와핑되도록 수학식 5에 따라 제1 마스크(M_{RGB})를 이용하여 소프트 와핑을 수행함으로써 제1 와핑 특징맵(\hat{f}_{RGB})을 획득할 수 있다.

수학식 5

$$\hat{\mathbf{f}}_{\text{RGB}}(\mathbf{p}) = M_{\text{RGB}}(\mathbf{p})\mathcal{W}(\mathbf{f}_{\text{IR}}(\mathbf{p})) + (1 - M_{\text{RGB}}(\mathbf{p}))\mathbf{f}_{\text{RGB}}(\mathbf{p})$$

여기서 \mathcal{W} 는 수학식 6으로 정의되는 소프트 와핑 함수이다.

수학식 6

$$\mathcal{W}(\mathbf{f}_{\text{IR}}(\mathbf{p})) = \sum_{\mathbf{q}} P(\mathbf{p}, \mathbf{q})\mathbf{f}_{\text{IR}}(\mathbf{q})$$

수학식 5 및 6에 따르면, 소프트 와핑부(540)는 소프트 와핑 함수(\mathcal{W})를 이용하여 제2 특징맵(\mathbf{f}_{IR})의 각 위치(\mathbf{q})별 픽셀 벡터를 매칭 확률에 따라 가중 집계하여 제1 특징맵(\mathbf{f}_{RGB})의 대응하는 픽셀 위치(\mathbf{p})로 와핑하되, 제1 마스크(M_{RGB})에 의해 지정되는 사람 영역에 제2 특징맵(\mathbf{f}_{IR})의 매칭 확률 가중 집계된 픽셀 벡터가 강조되어 와핑되도록 하고, 나머지 영역은 제1 특징맵(\mathbf{f}_{RGB})의 픽셀 벡터가 가급적 유지되도록 한다. 즉 제1 특징맵(\mathbf{f}_{RGB})에서 사람에 해당하는 전경 영역은 제2 특징맵(\mathbf{f}_{IR})에서 매칭 확률에 따라 가중 집계되어 와핑된 픽셀 벡터가 강조되는 반면, 나머지 배경 영역은 제1 특징맵(\mathbf{f}_{RGB})의 픽셀 벡터가 강조되도록 하여 제1 와핑 특징맵($\hat{\mathbf{f}}_{\text{RGB}}$)을 획득한다.

또한 소프트 와핑부(540)는 수학식 5 및 6과 유사하게 소프트 와핑 함수와 제2 마스크(M_{IR})를 이용하여, 제2 특징맵(\mathbf{f}_{IR})에서 사람에 해당하는 전경 영역은 제1 특징맵(\mathbf{f}_{RGB})에서 와핑된 픽셀 벡터가 강조되고, 나머지 배경 영역은 제2 특징맵(\mathbf{f}_{IR})의 픽셀 벡터가 강조되도록 하여 제2 와핑 특징맵($\hat{\mathbf{f}}_{\text{IR}}$)을 획득한다.

본 실시예에서 교차 모달 정렬부(500)가 소프트 와핑 함수(\mathcal{W})를 이용하는 것은 서로 다른 이미지에서 픽셀간 매칭을 위해 일반적으로 사용되는 argmax 함수를 이용하는 경우, 매우 하드하게 픽셀간 매칭이 수행되기 때문에 명시적인 대응 픽셀을 설정할 수 있는 반면, 서로 다른 이미지의 질감이나 폐색 등으로 인해 대응점이 조밀하지 않고 분산되거나 대응점이 존재하지 않는 경우가 빈번하게 발생할 수 있기 때문이다. 특히 본 실시예에서와 같이 이중의 이미지에서 추출된 제1 및 제2 특징맵($\mathbf{f}_{\text{RGB}}, \mathbf{f}_{\text{IR}}$) 사이의 대응 픽셀을 검출하고자 하는 경우, 이미지의 색상이나 질감과 같은 디테일한 특징보다 이미지의 개략적 형상 변화에 더욱 치중되어야 한다. 이에 본 실시예에서는 argmax 함수가 아닌 소프트 와핑 함수(\mathcal{W})를 이용하여 이중 이미지에서 서로 대응하는 픽셀의 위치로 와핑을 수행한다.

와핑 풀링부(550)는 제1 및 제2 풀링부(310, 320)와 동일한 방식으로 소프트 와핑부(540)에서 획득된 제1 와핑 특징맵($\hat{\mathbf{f}}_{\text{RGB}}$)과 제2 와핑 특징맵($\hat{\mathbf{f}}_{\text{IR}}$) 각각을 풀링하여 제1 와핑 표현자($\emptyset(\hat{\mathbf{f}}_{\text{RGB}})$)과 제2 와핑 표현자($\emptyset(\hat{\mathbf{f}}_{\text{IR}})$)를 획득한다.

그리고 재식별부(400)는 제1 및 제2 표현자($\emptyset(\mathbf{f}_{\text{RGB}}), \emptyset(\mathbf{f}_{\text{IR}})$)과 마찬가지로, 제1 및 제2 와핑 표현자($\emptyset(\hat{\mathbf{f}}_{\text{RGB}}), \emptyset(\hat{\mathbf{f}}_{\text{IR}})$)을 인가받아 제1 및 제2 표현자($\emptyset(\mathbf{f}_{\text{RGB}}), \emptyset(\mathbf{f}_{\text{IR}})$)에 대한 식별자를 추출할 수 있다.

손실 계산부(600)는 학습 과정에서 발생하는 손실을 기지정된 방식으로 계산하고, 계산된 손실을 역전파함으로써, 특징 추출부(200), 풀링부(300), 재식별부(400) 및 교차 모달 정렬부(500)를 학습시킨다.

본 실시예에서 손실 계산부(600)는 식별 손실(Identity loss)(L_{ID})과 일관성 손실(Identity Consistency loss)(L_{IC}) 및 밀집 삼중 손실(Dense Triplet loss)(L_{DT})를 각각 계산하고, 계산된 식별 손실(L_{ID})과 일관성 손실

(L_{IC}) 및 밀집 삼중 손실(L_{DT})을 결합하여 계산되는 총 손실(L)을 역전파함으로써 학습을 수행할 수 있다.

[0074] 손실 계산부(600)는 우선 제1 및 제2 표현자($\phi(f_{RGB}), \phi(f_{IR})$)을 기반으로 식별 손실(L_{ID})을 획득할 수 있다. 여기서 식별 손실(L_{ID})은 제1 및 제2 표현자($\phi(f_{RGB}), \phi(f_{IR})$)으로부터 획득되는 제1 및 제2 스코어맵과 학습시에 입력되는 제1 및 제2 이미지에 미리 레이블된 사람에 대한 식별자 사이의 차이를 기반으로 계산되어 획득될 수 있다. 즉 본 실시예의 사람 재식별 장치는 학습 시에 사람에 대한 식별자가 미리 레이블된 이중 이미지를 학습 데이터로 인가받아 지도 학습(Supervised Learning) 방식으로 학습이 수행될 수 있다.

[0075] 손실 계산부(600)는 제1 및 제2 스코어맵과 학습 데이터로서 인가된 식별자 사이의 음의 크로스 엔트로피(negative cross-entropy)를 계산하여 식별 손실(L_{ID})을 계산할 수 있다. 여기서 음의 크로스 엔트로피를 기반으로 식별 손실(L_{ID})을 계산하는 방법은 공지된 기술이므로 여기서는 상세하게 설명하지 않는다.

[0076] 한편, 손실 계산부(600)는 교차 모달 정렬부(500)에 의해 이중 이미지에서 획득된 제1 및 제2 특징맵(f_{RGB}, f_{IR})의 전경 영역이 정상적으로 와핑되었다면, 제1 및 제2 스코어맵로부터 추출되는 식별자와 제1 및 제2 와핑 스코어맵에서 추출되는 식별자가 동일해야 한다는 일관성 손실(L_{IC})을 계산한다.

[0077] 이때 손실 계산부(600)는 학습에 의해 최종적으로는 제1 및 제2 스코어맵에서 추출되는 식별자가 학습 데이터에 레이블된 식별자와 동일해지므로, 제1 및 제2 스코어맵으로부터 식별 손실(L_{ID})을 계산하는 방식과 동일하게 제1 및 제2 스코어맵과 학습 데이터에 레이블된 식별자를 기반으로 일관성 손실(L_{IC})을 계산할 수 있다.

[0078] 그러나 일관성 손실(L_{IC})은 실제 이미지에 대한 식별자보다는 제1 및 제2 스코어맵로부터 추출되는 식별자와 제1 및 제2 와핑 스코어맵에서 추출되는 식별자의 차이를 나타내므로, 손실 계산부(600)는 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵 사이의 음의 크로스 엔트로피로 일관성 손실(L_{IC})을 계산할 수 있다.

[0079] 한편 특징 추출부(200)에서 획득된 제1 및 제2 특징맵(f_{RGB}, f_{IR})과 소프트 와핑부(540)에서 획득된 제1 및 제2 와핑 특징맵($\hat{f}_{RGB}, \hat{f}_{IR}$)을 인가받고, 이중 서로 대응하는 특징맵과 와핑 특징맵을 기반으로 제1 및 제2 공통 주의맵(co-attention map)을 획득하고, 획득된 제1 및 제2 공통 주의맵을 이용하여 밀집 삼중 손실(L_{DT})을 계산한다.

[0080] 제1 및 제2 특징맵(f_{RGB}, f_{IR})과 소프트 와핑되어 재구성된 제1 및 제2 와핑 특징맵($\hat{f}_{RGB}, \hat{f}_{IR}$) 사이의 오차를 계산하는 가장 직접적인 방식은 L_2 놈 함수를 적용하여 대응하는 픽셀간 L_2 거리를 계산하는 것이다. 그러나, 이는 폐색 영역 등을 고려하지 않는다는 한계가 있다. 이에 본 실시예에서 손실 계산부(600)는 이중의 이미지 양측에서 모두 확인 가능한 사람 영역을 강조하는 공통 주의맵을 획득하고, 공통 주의맵 상에서 특징맵과 와핑 특징맵 사이의 오차를 계산하여 밀집 삼중 손실(L_{DT})을 획득한다.

[0081] 우선 제1 이미지에 대한 제1 공통 주의맵(A_{RGB})은 수학적 식 7에 따라 획득될 수 있다.

수학적 식 7

[0082]
$$A_{RGB}(\mathbf{p}) = f(M_{RGB}(\mathbf{p})\mathcal{W}(M_{IR}(\mathbf{p})))$$

[0083] 수학적 식 7을 살펴보면, 제1 공통 주의맵($A_{RGB}(\mathbf{p})$)은 제1 마스크(M_{RGB})와 와핑된 제2 마스크(M_{IR}) 사이의 교차점으로 구성되는 맵을 나타낸다.

[0084] 마찬가지로 제2 이미지에 대한 제2 공통 주의맵(A_{IR}) 또한 동일한 방식으로 획득될 수 있다.

[0085] 도 8에서 왼쪽 그림은 제2 마스크(M_{IR})를 나타내고, 오른쪽 그림은 제1 마스크(M_{RGB})를 나타내며, 가운데 그림은 제2 공통 주의맵(A_{IR})을 나타낸다. 도 8에 도시된 바와 같이, 제2 마스크(M_{IR})와 와핑된 제1 마스크(M_{RGB}) 사이의 교차점을 나타내는 제2 공통 주의맵(A_{IR})은 비록 제1 이미지에 포함된 사람의 일부 영역이 가려져 있음에도

제1 및 제2 이미지에서 공통적으로 표시되는 영역만이 강조되고, 적어도 하나의 이미지에서 표시되지 않은 영역(예를 들면 무릎)은 억제되도록 할 수 있다.

[0086] 제1 및 제2 공통 주의맵(A_{RGB}, A_{IR})이 획득되면, 손실 계산부(600)는 앵커를 기준으로 동일한 식별자에 대한 특징맵을 나타내는 포지티브와 다른 식별자에 대한 특징맵을 나타내는 네거티브를 비교하는 삼중(triplet) 학습 방식을 이용하여 손실을 계산한다. 여기서 포지티브 및 네거티브는 앵커와 다른 이종의 이미지(예를 들면, 앵커가 제1 이미지일 때, 포지티브 및 네거티브 쌍은 이종의 제2 이미지)에서 획득된다.

[0087] 일 예로 제1 특징맵(f_{RGB})이 앵커(f_{RGB}^a)인 경우, 포지티브(\hat{f}_{RGB}^p)는 앵커(f_{RGB}^a)와 동일한 식별자를 갖는 포지티브 제2 특징맵(f_{IR}^p)을 와핑하여 재구성된 제1 와핑 특징맵(\hat{f}_{RGB})을 나타내고, 네거티브(\hat{f}_{RGB}^n)는 앵커(f_{RGB}^a)와 다른 식별자를 갖는 네거티브 제2 특징맵(f_{IR}^n)을 와핑하여 재구성된 제1 와핑 특징맵(\hat{f}_{RGB})을 나타낸다.

[0088] 유사하게 제2 특징맵(f_{IR})이 앵커(f_{IR}^a)인 경우, 포지티브(\hat{f}_{IR}^p)는 앵커(f_{IR}^a)와 동일한 식별자를 갖는 포지티브 제1 특징맵(f_{RGB}^p)을 와핑하여 재구성된 제2 와핑 특징맵(\hat{f}_{IR})을 나타내고, 네거티브(\hat{f}_{IR}^n)는 앵커(f_{IR}^a)와 다른 식별자를 갖는 네거티브 제1 특징맵(f_{RGB}^n)을 와핑하여 재구성된 제2 와핑 특징맵(\hat{f}_{IR})을 나타낸다.

[0089] 이에 손실 계산부(600)는 제1 및 제2 공통 주의맵($A_i, i \in \{RGB, IR\}$) 각각을 기반으로 밀집 삼중 손실(L_{DT})을 수학적 식 8에 따라 계산할 수 있다.

수학적 식 8

[0090]
$$\mathcal{L}_{DT} = \sum_{i \in \{RGB, IR\}} \sum_{\mathbf{p}} A_i(\mathbf{p}) [d_i^+(\mathbf{p}) - d_i^-(\mathbf{p}) + \alpha]_+$$

[0091] 여기서 α 는 미리 정의된 마진을 나타내고, $d_i^+(\mathbf{p})$ 와 $d_i^-(\mathbf{p})$ 는 각각 수학적 식 9에 따라 계산되는 앵커와 포지티브 사이의 오차 및 앵커와 네거티브 각각 사이의 오차를 나타낸다.

수학적 식 9

[0092]
$$d_i^+(\mathbf{p}) = \|\mathbf{f}_i^a(\mathbf{p}) - \hat{\mathbf{f}}_i^p(\mathbf{p})\|_2, d_i^-(\mathbf{p}) = \|\mathbf{f}_i^a(\mathbf{p}) - \hat{\mathbf{f}}_i^n(\mathbf{p})\|_2$$

[0093] 식별 손실(L_{IC})과 일관성 손실(L_{IC}) 및 밀집 삼중 손실(L_{DT})이 계산되면, 손실 계산부(600)는 수학적 식 10에 따라 총 손실(\mathcal{L})을 계산하고 역전파함으로써, 학습을 수행할 수 있다.

수학적 식 10

[0094]
$$\mathcal{L} = \mathcal{L}_{ID} + \lambda_{IC} \mathcal{L}_{IC} + \lambda_{DT} \mathcal{L}_{DT}$$

[0095] 여기서 $\lambda_{IC}, \lambda_{DT}$ 는 각각 일관성 손실 가중치 및 밀집 삼중 손실 가중치이다.

[0096] 도 9는 사람 재식별 장치의 손실 설정에 따른 성능 변화를 나타낸다.

[0097] 도 9에서 (a)는 식별 손실(L_{IC})만을 적용하여 학습을 수행한 경우, 이종 이미지에서 상호 매칭되는 픽셀을 나타내고, (b)는 식별 손실(L_{IC})과 일관성 손실(L_{IC})을 적용하여 학습을 수행한 경우, 이종 이미지에서 상호 매칭되

는 픽셀을 나타낸다. 그리고 (c)는 식별 손실(L_{IC})과 일관성 손실(L_{IC}) 및 밀집 삼중 손실(L_{DT})을 적용하여 학습을 수행한 경우, 상호 매칭되는 픽셀을 나타낸다.

[0098] 도 9의 (c)에 도시된 바와 같이, 식별 손실(L_{IC})과 일관성 손실(L_{IC}) 및 밀집 삼중 손실(L_{DT})을 적용하여 학습을 수행하는 경우, 이중 이미지에서 사람 영역의 각 픽셀이 서로 대응하는 위치에 정확하게 매칭됨을 알 수 있다. 그리고 이는 본 실시예의 사람 재식별 장치가 매우 높은 정확도로 사람을 재식별할 수 있음을 나타낸다.

[0099] 도 10은 본 실시예에 따른 사람 재식별 장치의 성능을 시뮬레이션한 결과를 나타낸다.

[0100] 상기한 바와 같이, 이중 이미지에서 사람 영역의 각 픽셀이 서로 대응하는 위치에 정확하게 매칭시킬 수 있으며, 도 10의 (a)에 도시된 바와 같이 서로 다른 스케일이나 (b)의 시점 변화 및 (c)와 같이 여러 사람이 포함된 이미지에서도 정확하게 대응하는 픽셀을 매칭시킬 수 있어 사람을 재식별 할 수 있도록 한다.

[0101] 도 11은 본 발명의 일 실시예에 따른 이중 센서 카메라 기반 사람 재식별 방법을 나타낸다.

[0102] 도 11을 참조하면, 본 실시예에 따른 이중 센서 카메라 기반 사람 재식별 방법은 크게 학습 단계(S10)와 이중 이미지 사람 재식별 단계(S20)로 구분될 수 있다. 도 6 및 도 7을 참조하여, 우선 학습 단계(S10)를 설명하면, 이중 센서 카메라에서 획득되고, 각각 포함된 사람의 식별자가 레이블된 이중 이미지인 제1 및 제2 이미지를 학습 데이터로서 획득한다(S11). 이때, 제1 이미지와 제2 이미지는 각각 단일 프레임의 이미지일 수도 있으며, 다수 프레임으로 구성된 이미지일 수 있다. 또한 동일한 사람이 포함된 이미지이거나 서로 다른 사람이 포함된 이미지일 수도 있다.

[0103] 그리고 학습되는 방식에 따라 제1 및 제2 이미지 각각의 특징을 추출하여 제1 및 제2 학습 특징맵을 획득한다(S12). 이후 획득된 제1 및 제2 학습 특징맵 사이의 기지정된 방식으로 픽셀간 유사도를 계산하여 유사도맵($C(p,q)$)을 획득한다(S14). 여기서 유사도맵($C(p,q)$)은 일 예로 코사인 유사도를 기반으로 수학적 식 1에 따라 계산할 수 있다.

[0104] 유사도맵($C(p,q)$)이 획득되면, 획득된 유사도맵($C(p,q)$)을 기반으로 제1 및 제2 학습 특징맵에서 픽셀 위치별 매칭 확률을 수학적 식 2에 따라 계산하여 매칭 확률맵($P(p,q)$)을 획득한다(S15). 그리고 매칭 확률맵($P(p,q)$)을 이용하여 제1 및 제2 학습 특징맵에서 서로 대응하는 픽셀을 소프트 와핑하여 와핑 특징맵을 획득한다(S15). 이때 소프트 와핑은 제1 및 제2 학습 특징맵 각각에서 미리 사람 영역이 강조된 제1 및 제2 마스크를 수학적 식 3 및 4에 따라 획득하고, 수학적 식 5 및 6으로 나타난 바와 같이, 획득된 제1 및 제2 마스크 중 대응하는 마스크를 이용하여 픽셀이 와핑되는 특징맵의 사람 영역에 집중되어 와핑되도록 하고 나머지 영역은 기존 특징맵의 픽셀이 유지되도록 수행된다. 즉 제1 학습 특징맵의 픽셀은 제2 마스크를 이용하여 제2 특징맵의 사람 영역에 해당하는 위치로 와핑되고, 나머지 영역은 제2 학습 특징맵의 픽셀이 유지되어 제2 와핑 특징맵이 획득된다. 그리고 제2 학습 특징맵의 픽셀은 제1 마스크를 이용하여 제1 특징맵의 사람 영역에 해당하는 위치로 와핑되고, 나머지 영역은 제2 학습 특징맵의 픽셀이 유지되어 제1 와핑 특징맵이 획득된다. 제1 및 제2 와핑 특징맵이 획득되면, 제1 및 제2 학습 특징맵과 제1 및 제2 와핑 특징맵 각각을 풀링하여, 제1 및 제2 표현자와 제1 및 제2 와핑 표현자를 획득한다(S16).

[0105] 이후 획득된 제1 및 제2 표현자와 제1 및 제2 와핑 표현자 각각을 기반으로 기지정된 클래스별 대응확률을 나타내는 클래스를 분류를 수행하여 각각에 대응하는 식별자를 획득한다(S17). 일 예로 식별자는 제1 및 제2 표현자와 제1 및 제2 와핑 표현자 각각이 기지정된 다수의 클래스 각각에 대응할 확률을 나타내는 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵을 획득하고, 획득된 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵 각각을 기반으로 가장 높은 확률을 나타내는 클래스의 식별자를 추출하여 획득할 수 있다.

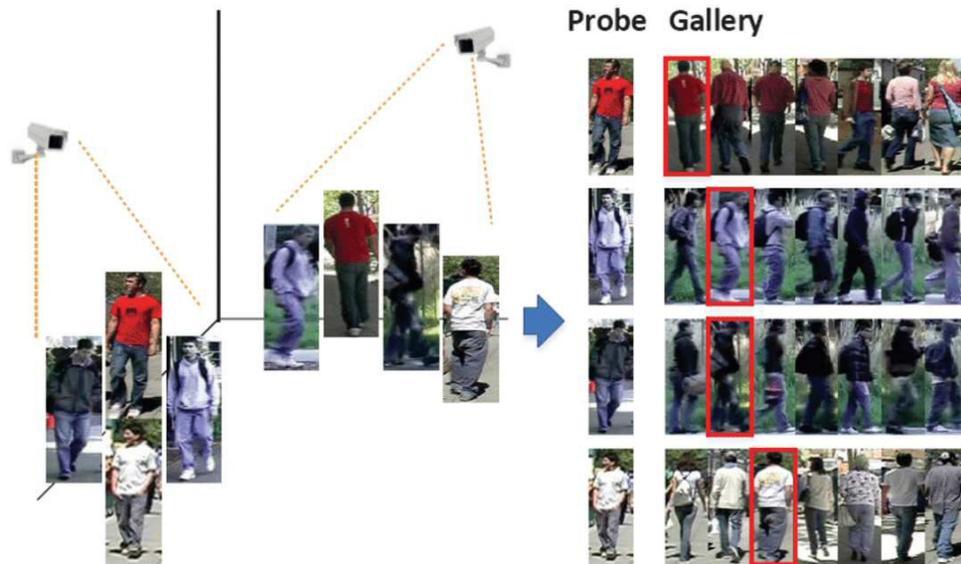
[0106] 그리고 식별자가 획득되면, 획득된 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵 그리고 제1 및 제2 학습 이미지에 레이블된 식별자를 이용하여 식별 손실(L_{ID})과 일관성 손실(L_{IC})을 계산하고, 제1 및 제2 특징맵과 제1 및 제2 와핑 특징맵을 이용하여 밀집 삼중 손실(L_{DT})을 계산하여 총 손실을 계산하고 역전파한다(S18).

[0107] 여기서 식별 손실(L_{ID})은 제1 및 제2 표현자에서 획득된 제1 및 제2 스코어맵과 제1 및 제2 학습 이미지에 레이블된 식별자 사이의 차이로 계산되며, 일 예로 제1 및 제2 스코어맵과 레이블된 식별자들 사이의 음의 크로스 엔트로피로 계산될 수 있다.

[0108] 그리고 일관성 손실(L_{IC})은 제1 및 제2 스코어맵과 제1 및 제2 와핑 스코어맵 사이의 음의 크로스 엔트로피를 계

도면

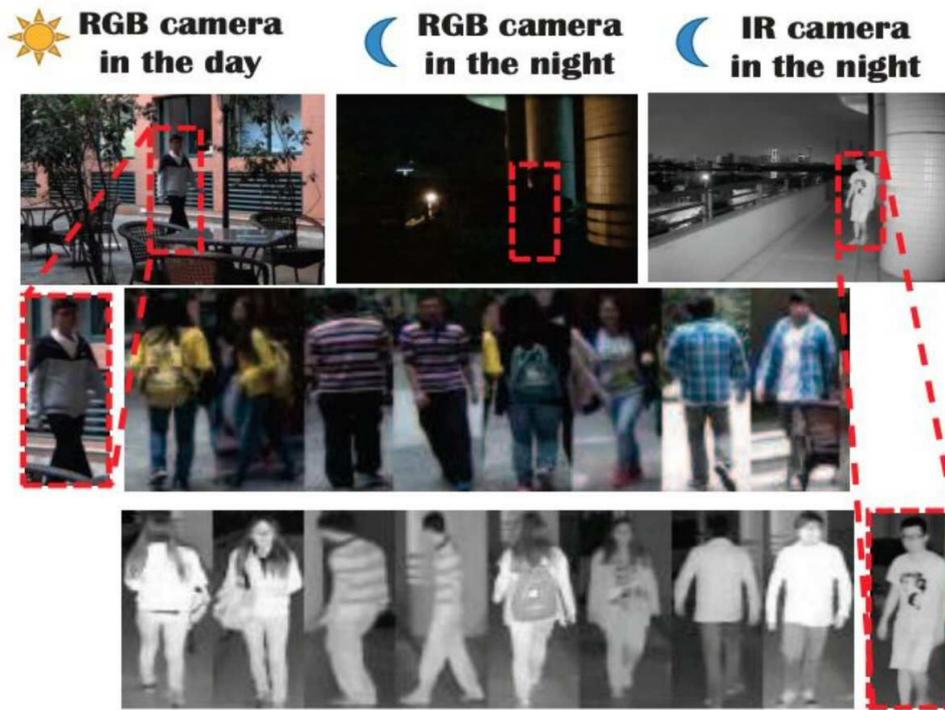
도면1



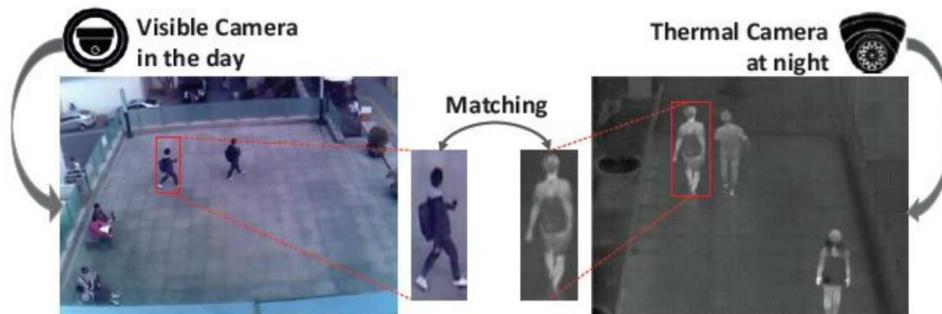
도면2



도면3



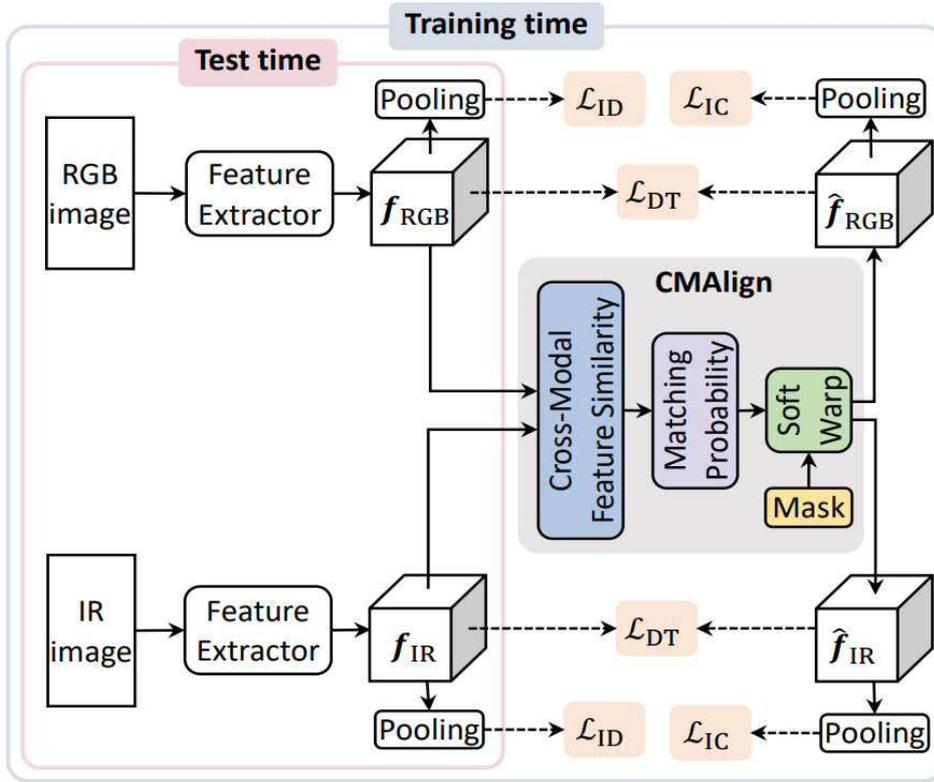
도면4



도면5



도면7



도면8



도면9

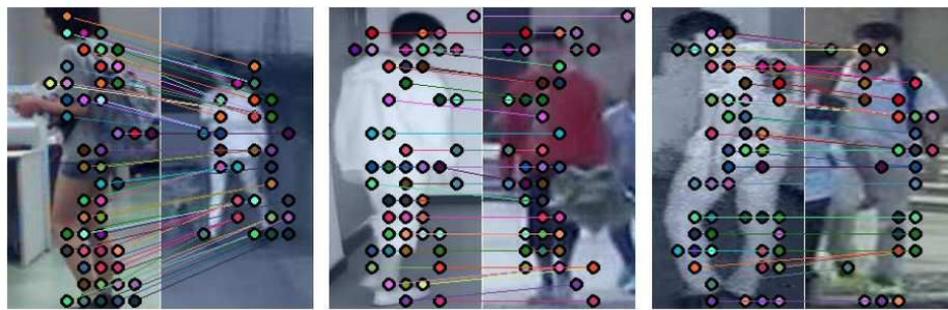


(a) \mathcal{L}_{ID}

(b) $\mathcal{L}_{ID} + \mathcal{L}_{IC}$

(c) $\mathcal{L}_{ID} + \mathcal{L}_{IC} + \mathcal{L}_{DT}$

도면10



(a) Scale variation

(b) Viewpoint variation

(c) Multiple persons

도면11

