



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0101521  
(43) 공개일자 2020년08월28일

(51) 국제특허분류(Int. Cl.)  
G06K 9/62 (2006.01) G06N 3/08 (2006.01)  
(52) CPC특허분류  
G06K 9/6212 (2013.01)  
G06K 9/6265 (2013.01)  
(21) 출원번호 10-2019-0012476  
(22) 출원일자 2019년01월31일  
심사청구일자 2019년01월31일

(71) 출원인  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
함범섭  
서울특별시 강남구 압구정로61길 37, 72동 506호 (압구정동, 한양아파트)  
이중협  
서울특별시 서대문구 연희로8길 26(연희동)  
김도형  
경기도 고양시 덕양구 백양로 8, 1704동 203호(화정동, 옥빛마을17단지아파트)  
(74) 대리인  
민영준

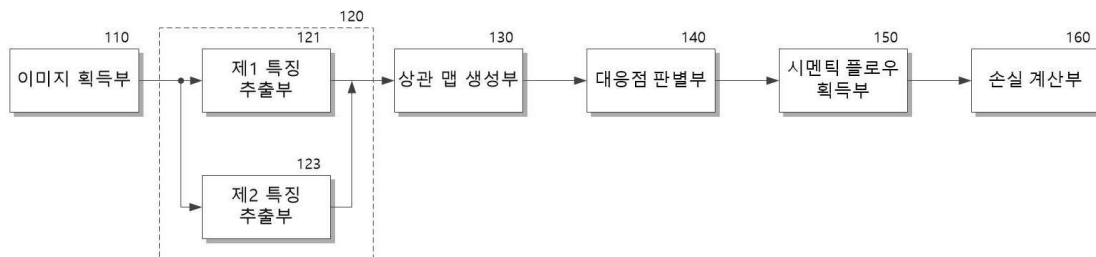
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 시멘틱 매칭 장치 및 방법

(57) 요약

본 발명은 이미지들의 동일 범주의 객체간 외형 차가 크거나 복잡한 질감을 갖는 경우 및 혼잡한 배경이 포함되더라도 정확하게 객체를 매칭할 수 있고, 이미지들에서 추출된 특징맵으로 생성되는 상관 맵에 위치별 가중치가 할당하여 이용함으로써 학습 성능을 향상시킬 수 있으며, 획득이 용이한 전경 마스크를 학습 데이터로 이용하여 학습될 수 있으므로, 학습 데이터 부족으로 인한 성능 저하없이 고성능의 시멘틱 매칭을 수행할 수 있는 시멘틱 매칭 장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

**G06N 3/08** (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2018019545

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 신진연구자지원사업

연구과제명 [Ezbaro] 딥러닝 기반 밀집 시멘틱 영상 정합 기법 연구: 지도 학습부터 비지도 학습까지

(2/3)

기 여 율 1/1

주관기관 연세대학교 산학협력단

연구기간 2018.03.01 ~ 2019.02.28

---

## 명세서

### 청구범위

#### 청구항 1

동일 구조를 갖고 패턴 인식 방법이 미리 학습되어 동일한 학습 가중치를 갖는 한 쌍의 삼 신경망을 포함하여, 인가된 소스 및 타겟 이미지 쌍( $I^s, I^t$ ) 각각에서 특징 맵 쌍( $f^s, f^t$ )을 획득하는 특징 추출부;

상기 특징 맵 쌍( $f^s, f^t$ )의 각 픽셀 사이의 상관 점수를 측정하여 상관 맵을 생성하는 상관 맵 생성부;

상기 상관 맵에서 소스 이미지( $I^s$ )의 픽셀 위치(p)별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하고, 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치를 반영하여 매칭 확률을 기지정된 방식으로 획득하며, 획득된 매칭 확률에 따라 소스 이미지( $I^s$ )의 픽셀 위치(p)에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치(q)인 대응점을 판별하는 대응점 판별부; 및

소스 이미지( $I^s$ )의 픽셀별 위치(p)와 이에 대응하는 타겟 이미지( $I^t$ )의 픽셀별 위치(q)의 관계를 이용하여 픽셀별 위치 변화를 나타내는 플로우 필드로 구성된 시멘틱 플로우 맵을 생성하는 시멘틱 플로우 획득부; 를 포함하는 시멘틱 매칭 장치.

#### 청구항 2

제1 항에 있어서, 상기 대응점 판별부는

상관 맵의 각 픽셀에서 argmax 함수를 이용하여 소스 이미지( $I^s$ )의 픽셀 위치(p)별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하는 상관 픽셀 추출부;

상기 상관 맵의 상기 소스 이미지( $I^s$ )의 픽셀 위치(p)에 대응하는 영역에서 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치로 구성된 가우시안 커널( $k_p$ )을 적용하는 가우시안 커널부;

가우시안 분포에 따른 분포 가중치가 적용된 상기 상관 맵의 픽셀들에 대해 softmax 함수를 기반으로 각 픽셀별 매칭 확률( $m_p(q)$ )을 계산하고, 계산된 매칭 확률( $m_p(q)$ )에 따라 매칭 확률 맵( $m_p$ )을 생성하는 매칭 확률 맵 생성부; 및

매칭 확률 맵( $m_p$ )에 따라 픽셀 위치(q)별 매칭 확률( $m_p(q)$ )을 적용하여 대응점( $\phi(p)$ )을 판별하는 대응점 획득부; 를 포함하는 시멘틱 매칭 장치.

#### 청구항 3

제2 항에 있어서, 상기 매칭 확률 맵 생성부는

수학적식

$$m_p(q) = \frac{\exp(\beta k_p(q) n_p(q))}{\sum_{q' \in n_p} \exp(\beta k_p(q') n_p(q'))}$$

(여기서  $\beta$ 는 온도 파라미터이고,  $n_p$ 는 상관 맵을 정규화한 정규 상관 맵을 나타낸다)

에 따라 매칭 확률( $m_p(q)$ )을 계산하고,

상기 대응점 획득부는

수학식

$$\phi(\mathbf{p}) = \sum_{\mathbf{q}} m_{\mathbf{p}}(\mathbf{q})\mathbf{q}$$

에 따라 대응점( $\phi(\mathbf{p})$ )을 계산하는 시멘틱 매칭 장치.

#### 청구항 4

제1 항에 있어서, 상기 시멘틱 매칭 장치는

학습 과정에서 상기 특징 추출부의 삼 신경망을 학습 시키기 위한 손실 계산부; 를 더 포함하고,

상기 손실 계산부는

상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ ) 각각에서 전경과 배경을 구분하여, 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )를 획득하는 전경 마스크 획득부;

상기 시멘틱 플로우 획득부로부터 소스 이미지( $I^s$ )의 픽셀 위치( $\mathbf{p}$ )로부터 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $\mathbf{q}$ )로의 변화를 나타내는 소스 플로우 필드( $F^s$ )와 타겟 이미지( $I^t$ )의 픽셀 위치( $\mathbf{q}$ )로부터 소스 이미지( $I^s$ )의 픽셀 위치( $\mathbf{p}$ )로의 위치 변화를 나타내는 타겟 플로우 필드( $F^t$ )를 인가받고, 상기 소스 플로우 필드( $F^s$ )와 상기 타겟 플로우 필드( $F^t$ )를 이용하여 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )를 각각 상호 위평함으로써, 추정 소스 마스크( $\hat{M}^s$ )와 추정 타겟 마스크( $\hat{M}^t$ )를 획득하며, 소스 전경 마스크( $M^s$ )와 추정 소스 마스크( $\hat{M}^s$ )와 사이의 차 및 타겟 전경 마스크( $M^t$ )와 추정 타겟 마스크( $\hat{M}^t$ ) 사이의 차를 누적하여 마스크 일관성 손실을 측정하는 마스크 손실 측정부;

상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ )의 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ ) 각각에 대응하는 영역 내의 각 픽셀에 대한 소스 플로우 필드( $F^s(\mathbf{p})$ )와 타겟 플로우 필드( $F^t(\mathbf{q})$ ) 사이의 차이를 누적하여 플로우 일관성 손실을 측정하는 플로우 일관성 손실 측정부;

상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ )의 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ ) 각각에 대응하는 영역 내에서 소스 플로우 필드( $F^s(\mathbf{p})$ )와 타겟 플로우 필드( $F^t(\mathbf{q})$ ) 각각에 대한 평활도 손실을 측정하는 평활화 손실 측정부; 및

상기 마스크 일관성 손실, 플로우 일관성 손실 및 평활화 손실 각각에 기지정된 손실 가중치를 적용하고 합하여 총 손실을 계산하고, 계산된 총 손실을 상기 특징 추출부의 삼 신경망으로 역전파하는 손실 역전파부; 를 포함하는 시멘틱 매칭 장치.

#### 청구항 5

제4 항에 있어서, 상기 마스크 손실 측정부는 마스크 일관성 손실( $L_{\text{mask}}$ )을 수학식

$$\mathcal{L}_{\text{mask}} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} (M^i(\mathbf{p}) - \hat{M}^i(\mathbf{p}))^2$$

(여기서 추정 소스 마스크( $\hat{M}^s$ )는 픽셀 위치( $\mathbf{p}$ )에 대해  $W(M^t; F^s)(\mathbf{p}) = M^t(\mathbf{p} + F^s(\mathbf{p}))$  연산을 수행하는 위평 연산자( $W$ )에 따라  $\hat{M}^s = W(M^t; \mathcal{F}^s)$ 로 획득되고, 추정 타겟 마스크( $\hat{M}^t$ ) 또한 대응하는 방식으로 획득됨)

에 따라 계산하고,

상기 플로우 일관성 손실 측정부는 플로우 일관성 손실( $L_{\text{flow}}$ )을 수학식

$$\mathcal{L}_{\text{flow}} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} \|(\mathcal{F}^i(\mathbf{p}) + \hat{\mathcal{F}}^i(\mathbf{p})) \odot M^i(\mathbf{p})\|_2^2$$

(여기서  $\hat{\mathcal{F}}^s$  는 각각 타겟 플로우 필드( $F^t(q)$ )를 이용하여 워핑 연산자(W)에 따라 소스 플로우 필드( $F^s(p)$ )를 워핑하여 획득되고,  $\hat{\mathcal{F}}^t$  는 소스 플로우 필드( $F^s(p)$ )를 이용하여 워핑 연산자(W)에 따라 타겟 플로우 필드( $F^t(q)$ )를 워핑하여 획득됨)

에 따라 계산하며,

상기 평활화 손실 측정부는 평활화 손실( $L_{\text{smooth}}$ )을 수학식

$$\mathcal{L}_{\text{smooth}} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} \|\nabla \mathcal{F}^i(\mathbf{p}) \odot M^i(\mathbf{p})\|_1$$

(여기서  $\|\cdot\|_1$ 는 L1 norm 함수를 나타내고,  $\nabla$ 는 그래디언트 연산자)

에 따라 계산하는 시멘틱 매칭 장치.

## 청구항 6

동일 구조를 갖고 패턴 인식 방법이 미리 학습되어 동일한 학습 가중치를 갖는 한 쌍의 삼 신경망을 이용하여, 인가된 소스 및 타겟 이미지 쌍( $I^s, I^t$ ) 각각에서 특징 맵 쌍( $f^s, f^t$ )을 획득하는 단계;

상기 특징 맵 쌍( $f^s, f^t$ )의 각 픽셀 사이의 상관 점수를 측정하여 상관 맵을 생성하는 단계;

상기 상관 맵에서 소스 이미지( $I^s$ )의 픽셀 위치(p)별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하고, 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치를 반영하여 매칭 확률을 기지정된 방식으로 획득하며, 획득된 매칭 확률에 따라 소스 이미지( $I^s$ )의 픽셀 위치(p)에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치(q)인 대응점을 판별하는 단계; 및

소스 이미지( $I^s$ )의 픽셀별 위치(p)와 이에 대응하는 타겟 이미지( $I^t$ )의 픽셀별 위치(q)의 관계를 이용하여 픽셀별 위치 변화를 나타내는 플로우 필드로 구성된 시멘틱 플로우 맵을 생성하는 단계; 를 포함하는 시멘틱 매칭 방법.

## 청구항 7

제6 항에 있어서, 상기 대응점을 판별하는 단계는

상관 맵의 각 픽셀에서 argmax 함수를 이용하여 소스 이미지( $I^s$ )의 픽셀 위치(p)별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하는 단계;

상기 상관 맵의 상기 소스 이미지( $I^s$ )의 픽셀 위치(p)에 대응하는 영역에서 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치로 구성된 가우시안 커널( $k_p$ )을 적용하는 단계;

가우시안 분포에 따른 분포 가중치가 적용된 상기 상관 맵의 픽셀들에 대해 softmax 함수를 기반으로 각 픽셀별 매칭 확률( $m_p(q)$ )을 계산하고, 계산된 매칭 확률( $m_p(q)$ )에 따라 매칭 확률 맵( $m_p$ )을 생성하는 단계; 및

매칭 확률 맵( $m_p$ )에 따라 픽셀 위치(q)별 매칭 확률( $m_p(q)$ )을 적용하여 대응점( $\phi(p)$ )을 획득하는 단계; 를 포함하는 시멘틱 매칭 방법.

## 청구항 8

제7 항에 있어서, 상기 매칭 확률 맵을 생성하는 단계는

수학식

$$m_{\mathbf{p}}(\mathbf{q}) = \frac{\exp(\beta k_{\mathbf{p}}(\mathbf{q})n_{\mathbf{p}}(\mathbf{q}))}{\sum_{\mathbf{q}' \in n_{\mathbf{p}}} \exp(\beta k_{\mathbf{p}}(\mathbf{q}')n_{\mathbf{p}}(\mathbf{q}'))}$$

(여기서  $\beta$ 는 온도 파라미터이고,  $n_{\mathbf{p}}$ 는 상관 맵을 정규화한 정규 상관 맵을 나타낸다)

에 따라 매칭 확률( $m_{\mathbf{p}}(\mathbf{q})$ )을 계산하고,

상기 대응점을 획득하는 단계는

수학식

$$\phi(\mathbf{p}) = \sum_{\mathbf{q}} m_{\mathbf{p}}(\mathbf{q})\mathbf{q}$$

에 따라 대응점( $\phi(\mathbf{p})$ )을 계산하는 시멘틱 매칭 방법.

### 청구항 9

제6 항에 있어서, 상기 시멘틱 매칭 방법은

학습 과정에서 상기 삼 신경망을 학습 시키기 위한 손실을 계산하는 단계; 를 더 포함하고,

상기 손실을 계산하는 단계는

상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ ) 각각에서 전경과 배경을 구분하여, 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )를 획득하는 단계;

상기 시멘틱 플로우 맵을 생성하는 단계에서 획득된 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로부터 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로의 변화를 나타내는 소스 플로우 필드( $F^s$ )와 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로부터 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로의 위치 변화를 나타내는 타겟 플로우 필드( $F^t$ )를 이용하여 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )를 각각 상호 위평함으로써, 추정 소스 마스크( $\hat{M}^s$ )와 추정 타겟 마스크( $\hat{M}^t$ )를 획득하며, 소스 전경 마스크( $M^s$ )와 추정 소스 마스크( $\hat{M}^s$ )와 사이의 차 및 타겟 전경 마스크( $M^t$ )와 추정 타겟 마스크( $\hat{M}^t$ ) 사이의 차를 누적하여 마스크 일관성 손실을 측정하는 단계;

상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ )의 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ ) 각각에 대응하는 영역 내의 각 픽셀에 대한 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 사이의 차이를 누적하여 플로우 일관성 손실을 측정하는 단계;

상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ )의 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ ) 각각에 대응하는 영역 내에서 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 각각에 대한 평활도 손실을 측정하는 단계; 및

상기 마스크 일관성 손실, 플로우 일관성 손실 및 평활화 손실 각각에 기지정된 손실 가중치를 적용하고 합하여 총 손실을 계산하고, 계산된 총 손실을 상기 삼 신경망으로 역전파하는 단계; 를 포함하는 시멘틱 매칭 방법.

### 청구항 10

제9 항에 있어서, 상기 마스크 일관성 손실을 측정하는 단계는 마스크 일관성 손실( $L_{\text{mask}}$ )을 수학식

$$\mathcal{L}_{\text{mask}} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} (M^i(\mathbf{p}) - \hat{M}^i(\mathbf{p}))^2$$

(여기서 추정 소스 마스크( $\hat{M}^s$ )는 픽셀 위치( $p$ )에 대해  $W(M^t; F^s)(p) = M^t(p) + F^s(p)$  연산을 수행하는 워핑 연산자( $W$ )에 따라  $\hat{M}^s = W(M^t; F^s)$ 로 획득되고, 추정 타겟 마스크( $\hat{M}^t$ ) 또한 대응하는 방식으로 획득됨)

에 따라 계산하고,

상기 플로우 일관성 손실을 측정하는 단계는 플로우 일관성 손실( $L_{flow}$ )을 수학식

$$\mathcal{L}_{flow} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} \|(\mathcal{F}^i(\mathbf{p}) + \hat{\mathcal{F}}^i(\mathbf{p})) \odot M^i(\mathbf{p})\|_2^2$$

(여기서  $\hat{\mathcal{F}}^s$  는 각각 타겟 플로우 필드( $F^t(q)$ )를 이용하여 워핑 연산자( $W$ )에 따라 소스 플로우 필드( $F^s(p)$ )를 워핑하여 획득되고,  $\hat{\mathcal{F}}^t$  는 소스 플로우 필드( $F^s(p)$ )를 이용하여 워핑 연산자( $W$ )에 따라 타겟 플로우 필드( $F^t(q)$ )를 워핑하여 획득됨)

에 따라 계산하며,

상기 평활도 손실을 측정하는 단계는 평활화 손실( $L_{smooth}$ )을 수학식

$$\mathcal{L}_{smooth} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} \|\nabla \mathcal{F}^i(\mathbf{p}) \odot M^i(\mathbf{p})\|_1$$

(여기서  $\|\cdot\|_1$ 는 L1 norm 함수를 나타내고,  $\nabla$ 는 그래디언트 연산자)

에 따라 계산하는 시멘틱 매칭 방법.

## 발명의 설명

## 기술 분야

[0001] 본 발명은 시멘틱 매칭 장치 및 방법에 관한 것으로, 딥러닝 기반 시멘틱 매칭 장치 및 방법에 관한 것이다.

## 배경 기술

[0002] 이미지 사이의 밀집 대응점(dense correspondence) 분석은 이미지 처리(image processing) 및 컴퓨터 비전(computer vision) 분야의 다양한 어플리케이션에 있어서 기본적인 작업이다. 기존에 이미지 대응점 분석은 동일 시점에서 시차를 갖고 획득된 두 이미지간의 스테레오 매칭(stereo matching) 또는 시간적으로 연속되는 이미지 프레임 사이의 옵티컬플로우(optical flow)를 획득하는데 주로 이용되고 있다.

[0003] 그러나 최근에는 이미지 사이에서 동일한 객체 또는 의미적으로 동일한 범주에 속하는 다양한 형상의 객체를 탐색하는 시멘틱 매칭에 대한 연구가 수행되고 있다. 즉 시멘틱 매칭은 다수의 이미지에서 동일한 의미를 갖지만 다른 형상으로 나타난 객체를 탐색한다.

[0004] 기존의 시멘틱 매칭은 주로 기지정된 수식에 기반하여 이미지들의 특징을 추출하고, 추출된 특징을 이용하여 신뢰성 및 평활화를 함께 고려한 에너지를 정의하여 최적화하는 방식으로 수행되었다. 그러나 수식에 기반하여 이미지의 특징을 추출하기 때문에, 이미지에서 저레벨 특징만을 추출할 수 있다. 따라서 다수 이미지에서 객체간의 외형의 변화가 큰 경우나 복잡한 질감의 객체 또는 배경이 포함된 이미지의 경우에 강인하지 못하다는 한계가 있다.

[0005] 한편, 인공 신경망, 특히 컨볼루션 신경망(Convolutional Neural Network: 이하 CNN)을 이용한 딥러닝 기법에 대한 연구가 발전됨에 따라 딥러닝 기반 시멘틱 매칭 방법이 제안되었다. 기존의 딥러닝 기반 시멘틱 매칭 방법은 기본적으로 지도 학습 방식을 기반으로 학습된 인공 신경망을 이용하여 구현되며, 강한 지도 학습(Strong Supervised learning) 또는 약한 지도 학습(Weakly Supervised learning) 방식으로 인공 신경망을 학습시키고, 학습된 인공 신경망을 이용하여 이미지 사이에 동일한 의미를 갖는 객체를 탐색하는 시멘틱 매칭을 수행한다.

[0006] 강한 지도 학습 방식에서 학습 데이터는 이미지간의 모든 픽셀 또는 일부 픽셀에서의 매칭 정보인 키포인트(keypoint) 데이터가 이용되는 반면, 약한 지도 학습 방식에서 학습 데이터는 키포인트 데이터 이외의 데이터를

이용한다.

- [0007] 이에 강한 지도 학습 방식으로 인공 신경망을 학습시키기 위해서는 다수의 학습 데이터가 요구되지만, 학습 데이터의 생성 과정이 어렵기 때문에 다수의 학습데이터를 획득하기가 용이하지 않다는 문제가 있다. 반면 약한 지도 학습 방식으로 학습된 인공 신경망은 학습 데이터의 획득이 용이한 반면, 시멘틱 매칭 성능이 낮다는 한계가 있다.

## 선행기술문헌

### 특허문헌

- [0008] (특허문헌 0001) 한국 등록 특허 제10-1142163호 (2012.04.25 등록)

## 발명의 내용

### 해결하려는 과제

- [0009] 본 발명의 목적은 이미지들의 동일 범주의 객체간 외형 차가 크거나 복잡한 질감을 갖는 경우 및 혼잡한 배경이 포함되더라도 강인하게 객체를 매칭할 수 있는 시멘틱 매칭 장치 및 방법을 제공하는데 있다.
- [0010] 본 발명의 다른 목적은 용이하게 획득할 수 있는 전경 마스크를 학습 데이터로 이용하여 학습되어 높은 성능으로 이미지 사이의 객체를 인식하여 매칭할 수 있는 시멘틱 매칭 장치 및 방법을 제공하는데 있다.
- [0011] 본 발명의 또다른 목적은 이미지들에서 추출된 특징맵으로 생성되는 상관 맵에 위치별 가중치가 할당하여 이용함으로써 학습 성능을 향상시킬 수 있는 시멘틱 매칭 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0012] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 시멘틱 매칭 장치는 동일 구조를 갖고 패턴 인식 방법이 미리 학습되어 동일한 학습 가중치를 갖는 한 쌍의 삼 신경망을 포함하여, 인가된 소스 및 타겟 이미지 쌍( $I^s$ ,  $I^t$ ) 각각에서 특징 맵 쌍( $f^s$ ,  $f^t$ )을 획득하는 특징 추출부; 상기 특징 맵 쌍( $f^s$ ,  $f^t$ )의 각 픽셀 사이의 상관 점수를 측정하여 상관 맵을 생성하는 상관 맵 생성부; 상기 상관 맵에서 소스 이미지( $I^s$ )의 픽셀 위치(p)별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하고, 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치를 반영하여 매칭 확률을 기지정된 방식으로 획득하며, 획득된 매칭 확률에 따라 소스 이미지( $I^s$ )의 픽셀 위치(p)에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치(q)인 대응점을 판별하는 대응점 판별부; 및 소스 이미지( $I^s$ )의 픽셀별 위치(p)와 이에 대응하는 타겟 이미지( $I^t$ )의 픽셀별 위치(q)의 관계를 이용하여 픽셀별 위치 변화를 나타내는 플로우 필드로 구성된 시멘틱 플로우 맵을 생성하는 시멘틱 플로우 획득부; 를 포함한다.
- [0013] 상기 대응점 판별부는 상관 맵의 각 픽셀에서 argmax 함수를 이용하여 소스 이미지( $I^s$ )의 픽셀 위치(p)별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하는 상관 픽셀 추출부; 상기 상관 맵의 상기 소스 이미지( $I^s$ )의 픽셀 위치(p)에 대응하는 영역에서 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치로 구성된 가우시안 커널( $k_p$ )을 적용하는 가우시안 커널부; 가우시안 분포에 따른 분포 가중치가 적용된 상기 상관 맵의 픽셀들에 대해 softmax 함수를 기반으로 각 픽셀별 매칭 확률( $m_p(q)$ )을 계산하고, 계산된 매칭 확률( $m_p(q)$ )에 따라 매칭 확률 맵( $m_p$ )을 생성하는 매칭 확률 맵 생성부; 및 매칭 확률 맵( $m_p$ )에 따라 픽셀 위치(q)별 매칭 확률( $m_p(q)$ )을 적용하여 대응점( $\phi(p)$ )을 판별하는 대응점 획득부; 를 포함할 수 있다.
- [0014] 상기 시멘틱 매칭 장치는 학습 과정에서 상기 특징 추출부의 삼 신경망을 학습 시키기 위한 손실 계산부; 를 더 포함하고, 상기 손실 계산부는 상기 소스 및 타겟 이미지 쌍( $I^s$ ,  $I^t$ ) 각각에서 전경과 배경을 구분하여, 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )를 획득하는 전경 마스크 획득부; 상기 시멘틱 플로우 획득부로부터 소스 이미지( $I^s$ )의 픽셀 위치(p)로부터 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치(q)로의 변화를 나타내는 소스 플로우



필드( $F^s$ )와 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로부터 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로의 위치 변화를 나타내는 타겟 플로우 필드( $F^t$ )를 인가받고, 상기 소스 플로우 필드( $F^s$ )와 상기 타겟 플로우 필드( $F^t$ )를 이용하여 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )를 각각 상호 위핑함으로써, 추정 소스 마스크( $\hat{M}^s$ )와 추정 타겟 마스크( $\hat{M}^t$ )를 획득하며, 소스 전경 마스크( $M^s$ )와 추정 소스 마스크( $\hat{M}^s$ )와 사이의 차 및 타겟 전경 마스크( $M^t$ )와 추정 타겟 마스크( $\hat{M}^t$ ) 사이의 차를 누적하여 마스크 일관성 손실을 측정하는 마스크 손실 측정부; 상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ )의 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ ) 각각에 대응하는 영역 내의 각 픽셀에 대한 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 사이의 차이를 누적하여 플로우 일관성 손실을 측정하는 플로우 일관성 손실 측정부; 상기 소스 및 타겟 이미지 쌍( $I^s, I^t$ )의 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ ) 각각에 대응하는 영역 내에서 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 각각에 대한 평할도 손실을 측정하는 평할화 손실 측정부; 및 상기 마스크 일관성 손실, 플로우 일관성 손실 및 평할화 손실 각각에 기지정된 손실 가중치를 적용하고 합하여 총 손실을 계산하고, 계산된 총 손실을 상기 특징 추출부의 삼신경망으로 역전파하는 손실 역전파부; 를 포함할 수 있다.

[0015] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 시멘틱 매칭 방법은 동일 구조를 갖고 패턴 인식 방법이 미리 학습되어 동일한 학습 가중치를 갖는 한 쌍의 삼신경망을 이용하여, 인가된 소스 및 타겟 이미지 쌍( $I^s, I^t$ ) 각각에서 특징 맵 쌍( $f^s, f^t$ )을 획득하는 단계; 상기 특징 맵 쌍( $f^s, f^t$ )의 각 픽셀 사이의 상관 점수를 측정하여 상관 맵을 생성하는 단계; 상기 상관 맵에서 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )별 상관 점수가 가장 높은 픽셀 위치인 상관 픽셀 위치를 판별하고, 상기 상관 픽셀 위치를 중심으로 가우시안 분포에 따른 분포 가중치를 반영하여 매칭 확률을 기지정된 방식으로 획득하며, 획득된 매칭 확률에 따라 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )인 대응점을 판별하는 단계; 및 소스 이미지( $I^s$ )의 픽셀별 위치( $p$ )와 이에 대응하는 타겟 이미지( $I^t$ )의 픽셀별 위치( $q$ )의 관계를 이용하여 픽셀별 위치 변화를 나타내는 플로우 필드로 구성된 시멘틱 플로우 맵을 생성하는 단계; 를 포함한다.

### 발명의 효과

[0016] 따라서, 본 발명의 실시예에 따른 시멘틱 매칭 장치 및 방법은 이미지들의 동일 범주의 객체간 외형 차가 크거나 복잡한 질감을 갖는 경우 및 혼잡한 배경이 포함되더라도 정확하게 객체를 매칭할 수 있다. 또한 이미지들에서 추출된 특징맵으로 생성되는 상관 맵에 위치별 가중치가 할당하여 이용함으로써 학습 성능을 향상시킬 수 있다. 뿐만 아니라, 획득이 용이한 전경 마스크를 학습 데이터로 이용하여 학습될 수 있으므로, 학습 데이터 부족으로 인한 성능 저하없이 고성능의 시멘틱 매칭을 수행할 수 있다.

### 도면의 간단한 설명

[0017] 도1 은 본 발명의 일 실시예에 따른 시멘틱 매칭 장치의 개략적 구조를 나타낸다.  
 도2 는 도1 의 손실 계산부의 상세 구성을 나타낸다.  
 도3 은 도1 의 대응점 판별부의 상세 구성을 나타낸다.  
 도4 는 도3 의 대응점 판별부의 동작을 설명하기 위한 도면이다.  
 도5 는 도1 의 손실 계산부의 상세 구성을 나타낸다.  
 도6 은 도5 의 마스크 손실 측정부가 마스크 일관성 손실을 측정하는 개념을 나타낸다.  
 도7 은 플로우 일관성 손실의 개념을 설명하기 위한 도면이다.  
 도8 은 평할화 손실의 개념을 설명하기 위한 도면이다.  
 도9 는 도1 의 시멘틱 매칭 장치를 학습시키기 위한 학습 이미지와 소스 및 타겟 이미지에 대한 시멘틱 플로우의 일예를 나타낸다.

도10 은 본 실시예에 따른 시멘틱 매칭 방법을 나타낸다.

도11 은 소스 이미지와 타겟 이미지 사이의 매칭 결과를 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0018] 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.
- [0019] 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.
- [0020] 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.
- [0021] 도1 은 본 발명의 일 실시예에 따른 시멘틱 매칭 장치의 개략적 구조를 나타낸다.
- [0022] 도1 을 참조하면, 본 실시예에 따른 시멘틱 매칭 장치는 영상 획득부(110), 특징 추출부(120), 상관 맵 생성부(130), 대응점 판별부(140), 시멘틱 플로우 획득부(150) 및 손실 계산부(160)를 포함한다.
- [0023] 우선 영상 획득부(110)는 시멘틱 매칭이 수행되어야 할 다수의 이미지를 획득한다. 여기서는 일례로 영상 획득부(110)가 소스 이미지( $I^s$ )와 타겟 이미지( $I^t$ )의 이미지 쌍을 획득하는 것으로 가정한다.
- [0024] 특징 추출부(120)는 미리 학습된 인공 신경망으로 구현되어 영상 획득부(110)에서 획득된 이미지 쌍( $I^s, I^t$ )에 각각에 대한 2개의 특징맵( $f^s, f^t$ )을 생성한다.
- [0025] 특징 추출부(120)는 이미지 쌍( $I^s, I^t$ )에 대응하는 2개의 특징 추출부(121, 123)를 포함한다. 여기서 2개의 특징 추출부(121, 123) 각각은 패턴 인식 방법이 미리 학습된 인공 신경망, 일례로 Residual Net(이하 ResNet) 또는 컨볼루션 신경망(Convolution neural network: 이하 CNN)으로 구현될 수 있다. 특히 본 실시예에서 특징 추출부(120)의 2개의 특징 추출부(121, 123)는 동일한 구조를 갖고, 동시에 학습되어 동일한 가중치가 적용되는 쌍(siamese) 신경망으로 구현될 수 있다.
- [0026] 제1 특징 추출부(121)는 이미지 쌍( $I^s, I^t$ ) 중 소스 이미지( $I^s$ )를 인가받고, 미리 학습된 패턴 인식 방법에 따라 소스 이미지( $I^s$ )의 특징을 추출하여 소스 특징 맵( $f^s$ )을 생성한다. 그리고 제2 특징 추출부(123)는 타겟 이미지( $I^t$ )를 인가받고, 미리 학습된 패턴 인식 방법에 따라 타겟 이미지( $I^t$ )의 특징을 추출하여 타겟 특징 맵( $f^t$ )을 생성한다.
- [0027] 제1 및 제2 특징 추출부(121, 123)는 시멘틱 매칭 장치의 학습 과정에서 손실 계산부(160)에서 역전파되는 손실에 따라 신경망 가중치를 업데이트함으로써 패턴 인식 방식이 학습되어 이미지 쌍( $I^s, I^t$ )으로부터 특징 맵 쌍( $f^s, f^t$ )을 생성할 수 있다.
- [0028] 상관 맵 생성부(130)는 특징 추출부(120)에서 획득된 특징 맵 쌍( $f^s, f^t$ )의 다수의 픽셀(또는 특징점)에서 서로 매칭되는 픽셀을 추출할 수 있도록 두 특징 맵( $f^s, f^t$ ) 사이의 상관 점수(Correlation score)를 측정하여 상관 맵(correlation map)을 생성한다.
- [0029] 이때 상관 맵 생성부(130)는 수학적 식 1에 따라 상관 점수를 획득할 수 있다.

## 수학식 1

$$c(\mathbf{p}, \mathbf{q}) = f^s(\mathbf{p})^\top f^t(\mathbf{q})$$

[0030]

[0031] 여기서  $f^s(\mathbf{p})$ 와  $f^t(\mathbf{q})$ 는 각각 소스 이미지( $I^s$ )와 타겟 이미지( $I^t$ ) 각각의 픽셀 위치( $\mathbf{p} = (p_x, p_y)$ ,  $\mathbf{q} = (q_x, q_y)$ )에서 특징들을 나타내고,  $T$ 는 전치 행렬(transposed matrix) 기호를 나타내며,  $c(\mathbf{p}, \mathbf{q})$ 는  $\mathbf{p}$ 와  $\mathbf{q}$  사이의 상관 점수를 나타낸다.

[0032] 상관 맵( $c_p$ )은 상관 점수( $c(\mathbf{p}, \mathbf{q})$ )로 구성( $c_p = c(\mathbf{p}, \mathbf{q})$ )되며, 상관 맵 생성부(130)는 획득된 상관 맵( $c_p$ )의 상관 점수( $c(\mathbf{p}, \mathbf{q})$ )를 수학식 2와 같이 L2 정규화 함수(L2 norm)를 이용하여 정규화하여, 정규화된 상관 점수( $n(\mathbf{p}, \mathbf{q})$ )를 갖는 정규 상관 맵( $n_p = n(\mathbf{p}, \mathbf{q})$ )을 획득할 수도 있다.

## 수학식 2

$$n(\mathbf{p}, \mathbf{q}) = \| c(\mathbf{p}, \mathbf{q}) \|_2$$

[0033]

[0034] 이때 본 실시예에서 상관 맵 생성부(130)는 상관 맵을 이용하여 시멘틱 대응점 분석이 용이하도록 적응 레이어를 갖는 상관 맵을 생성할 수 있다.

[0035] 도2는 적응 레이어를 갖는 상관 맵의 개념을 시각화한 예를 나타내며, 도2에서는 상관 맵( $c_p$ )이 정규화된 정규 상관 맵( $n_p$ )을 도시하였다.

[0036] 이미지 쌍( $I^s, I^t$ )에 대응하는 특징 맵 쌍( $f^s, f^t$ )이 각각  $w \times h$ 의 크기를 갖는 경우, 적응 레이어를 갖는 상관 맵( $n_p$ )은  $w \times h \times w \times h$  크기의 4차원 구조를 가질 수 있다. 그러나 도2에 도시된 바와 같이, 상관 맵( $n_p$ )의 깊이( $d$ )를  $w \times h$ 로 구성하는 경우, 상관 맵은  $w \times h \times d$  크기의 3차원 구조로 구성될 수 있다.

[0037] 이는 도2에 도시된 바와 같이, 타겟 이미지( $I^t$ )에서 소스 이미지( $I^s$ )의 픽셀( $\mathbf{p}$ )에 대응하는 픽셀( $\mathbf{q}$ )을 상관 점수를 기반으로 용이하게 탐색할 수 있도록 하기 위함이다.

[0038] 대응점 판별부(140)는 상관 맵 생성부(130)에서 생성된 상관 맵( $n_p$ )으로부터 특징 맵 쌍( $f^s, f^t$ ) 사이의 대응점을 판별한다. 대응점 판별부(140)는 상관 맵을 이용하여 소스 특징 맵( $f^s$ )의 특징( $f^s(\mathbf{p})$ )에 대한 상관 점수( $c_p(\mathbf{q}) = c(\mathbf{p}, \mathbf{q})$ )를 기반으로 타겟 특징 맵( $f^t$ )의 모든 특징 중 대응하는 특징( $f^t(\mathbf{q})$ )의 픽셀 위치( $\mathbf{q}$ )를 대응점으로 판별할 수 있다.

[0039] 도3은 도1의 대응점 판별부의 상세 구성을 나타낸다.

[0040] 도3을 참조하면, 대응점 판별부(140)는 상관 픽셀 추출부(141), 온도 파라미터 적용부(143), 가우시안 커널부(145), 매칭 확률 맵 생성부(147) 및 대응점 획득부(149)를 포함할 수 있다.

[0041] 상관 픽셀 추출부(141)는 argmax 함수를 이용하여,  $w \times h \times d$  크기의 3차원 상관 맵에서 픽셀 위치( $\mathbf{p}$ )로부터 깊이( $d$ ) 방향의  $w \times h$  픽셀 중 상관 점수가 가장 높은 상관 픽셀 위치( $\phi(\mathbf{p})$ )를 수학식 3에 따라 추출할 수 있다.

### 수학식 3

$$\phi(p) = \operatorname{argmax}_q n_p(q)$$

[0042]

[0043] 그러나 수학식 3의  $\operatorname{argmax}$  함수를 이용하는 경우, 단순히 소스 특징 맵( $f^s$ )의 특징( $f^t(p)$ )과 상관 점수가 가장 높은 특징( $f^t(q)$ )을 획득함에 따라 특징( $f^t(p)$ )의 픽셀 위치( $\phi(p)$ )가 이산적으로 획득된다는 문제가 있다.

[0044] 일반적으로 소스 이미지( $I^s$ )에서 픽셀 위치( $p$ )를 중심으로 주변 픽셀(예를 들면,  $p+1$ )들의 특징( $f^t(p+1)$ )은 객체의 경계면을 제외하면 특징( $f^t(p)$ )과 유사하게 나타날 확률이 높다. 그러나  $\operatorname{argmax}$  함수를 적용하는 경우, 특징( $f^t(p+1)$ )에 대응하는 상관 픽셀 위치( $\phi(p)$ )는 매우 이산적으로 나타난다. 즉 타겟 이미지( $I^t$ )에서의 상관 픽셀 위치( $q$ )가 소스 이미지( $I^s$ )에서의 픽셀 위치( $p$ )의 변화와 무관하게 독립적으로 획득될 수 있다. 이는 시멘틱 매칭이 정확하게 수행되지 않을 수 있음을 의미한다. 뿐만 아니라, 특징 추출부(120)의 삼 신경망을 학습시키기 위해 손실을 역전파하는 경우, 역전파되는 손실에는 평활화 손실이 포함되며, 이는 그래디언트( $\operatorname{gradient}$ ) 연산을 전제로 한다. 따라서 대응점 판별로 획득되는 시멘틱 플로우는 모든 위치에서 미분 가능해야 한다.

[0045] 이에 이산적인 결과를 출력하는  $\operatorname{argmax}$  함수 대신, 대응하는 매칭 확률을 갖는 모든 위치의 가중 평균에 대한 출력을 계산하여 미분 가능하도록 하는  $\operatorname{soft argmax}$  함수를 이용할 수도 있으나,  $\operatorname{soft argmax}$  함수는 모든 위치에 동일한 가중치를 반영하도록 하여 부정확한 대응점을 판별할 수 있다는 문제가 있다.

[0046] 그러므로 본 실시예에서 대응점 판별부(140)에서는 상관 픽셀 추출부(141)에서  $\operatorname{argmax}$  함수를 이용하여 상관 점수가 가장 높은 상관 픽셀 위치( $\phi(p)$ )가 획득되면, 가우시안 커널부(145)가 획득된 상관 픽셀 위치( $\phi(p)$ )를 중심으로 인접한 픽셀에 대해 가우시안 분포에 따른 분포 가중치를 반영하는 가우시안 커널( $k_p$ )에 따라 상관 맵( $n_p$ )에 가중치를 적용한다.

[0047] 그리고 매칭 확률 맵 생성부(147)는 가우시안 커널( $k_p$ )에 따른 가중치가 적용된 상관 맵( $n_p$ )에 대해 픽셀 위치( $q$ )별 매칭 확률( $m_p(q)$ )을  $\operatorname{softmax}$  함수를 이용하여 수학식 4와 같이 계산함으로써, 매칭 확률 맵( $m_p$ )을 생성한다.

### 수학식 4

$$m_p(q) = \frac{\exp(\beta k_p(q) n_p(q))}{\sum_{q' \in n_p} \exp(\beta k_p(q') n_p(q'))}$$

[0048]

[0049] 여기서  $\beta$ 는 온도 파라미터이다.

[0050] 대응점 판별부(140)가 매칭 확률 맵 생성부(147)가  $\operatorname{softmax}$  함수를 이용하여 매칭 확률( $m_p(q)$ )을 계산하여 매칭 확률 맵을 생성할 때, 출력 분포를 조절하기 위해, 도3에 도시된 바와 같이 가우시안 커널( $k_p$ )이 상관 맵( $n_p$ )에 가중치를 적용하기 이전에 상관 맵( $n_p$ )에 기지정된 값을 갖는 온도 파라미터( $\beta$ )를 우선 적용하는 온도 파라미터 적용부(143)를 더 구비할 수 있다.

[0051] 수학식 4에 따르면 매칭 확률( $m_p(q)$ )은 상관 맵( $n_p$ )과 가우시안 커널( $k_p$ )을 엘리먼트 단위로 곱셈을 수행하고  $\operatorname{softmax}$  함수를 적용하여 매칭 확률( $m_p(q)$ )을 계산한다. 그리고 계산된 매칭 확률( $m_p(q)$ )로부터 매칭 확률 맵( $m_p$ )을 생성할 수 있다.

[0052] 대응점 획득부(149)는 매칭 확률 맵 생성부(147)에서 생성된 매칭 확률 맵( $m_p$ )에 따라 픽셀 위치( $q$ )별 매칭 확

를( $m_p(q)$ )을 적용하여 수학식 5에 따라 대응점( $\phi(p)$ )을 판별한다.

### 수학식 5

$$\phi(p) = \sum_q m_p(q)q$$

[0053]

[0054] 수학식 5는 매칭 확률( $m_p(p)$ )에 의해 가중치가 반영된 모든 위치( $q = (q_x, q_y)$ )의 평균으로 픽셀 위치( $p$ )에 대응하는 대응점( $\phi(p)$ )을 계산한다.

[0055] 여기서 대응점( $\phi(p)$ )은 상관 맵( $n_p$ )에서 판별된 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )를 나타낸다.

[0056] 도4 은 도3 의 대응점 판별부의 동작을 설명하기 위한 도면이다.

[0057] 도4 에서는 가우시안 분포를 적용하는 가우시안 커널( $k_p$ )의 영향을 확인할 수 있도록 가우시안 커널( $k_p$ )을 이용하지 않는 경우와 본 실시예에 따라 가우시안 커널( $k_p$ )을 이용하는 경우를 함께 나타내었다.

[0058] 가우시안 커널( $k_p$ )을 이용하지 않는 경우는 soft argmax 연산을 수행하여 대응점( $\phi(p)$ )을 판별하는 것으로 볼 수 있으며, 이 경우 도3 의 상단에 도시된 바와 같이 모든 위치( $q$ )에 대해 평균값으로 획득되는 균일한 매칭 확률로 가중치를 적용함에 따라 대응점( $\phi(p)$ )을 오판별 할 수 있다는 문제가 있다.

[0059] 그러나 본 실시예와 같이 argmax 함수를 이용하여 획득되는 상관 픽셀 위치( $q$ )를 중심으로 가우시안 커널( $k_p$ )을 적용하는 경우, 상관 픽셀 위치( $q$ )를 중심으로 인접한 픽셀들에 가우시안 분포에 따른 매칭 확률맵( $m_p$ )으로 가중치를 적용함에 따라 상대적으로 정확하게 대응점( $\phi(p)$ )을 판별 할 수 있다. 즉 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )를 판별할 수 있다.

[0060] 시멘틱 플로우 획득부(150)는 대응점 판별부(140)에서 판별된 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )에 따라 픽셀 위치의 변화를 나타내는 시멘틱 플로우 맵을 생성한다. 즉 입력된 이미지 쌍( $I^s, I^t$ ) 사이에 시멘틱 매칭을 수행한 결과를 출력한다.

[0061] 이때 시멘틱 플로우 획득부(150)는 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로부터 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로의 위치 변화를 나타내는 시멘틱 플로우 맵 뿐만 아니라, 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로부터 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로의 위치 변화를 나타내는 시멘틱 플로우 맵을 함께 생성할 수 있다.

[0062] 여기서 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )에 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )에 따라 픽셀 위치의 변화를 나타내는 시멘틱 플로우 맵은 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로부터 대응하는 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로의 소스 플로우 필드( $F^s$ )로 구성되고, 소스 플로우 필드( $F^s$ )는  $F^s = \phi(p) - p$  로 획득된다.

[0063] 유사하게, 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로부터 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로의 위치 변화를 나타내는 시멘틱 플로우 맵은 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로부터 대응하는 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로의 타겟 플로우 필드( $F^t$ )로 구성되고, 타겟 플로우 필드( $F^t$ )는  $F^t = \phi(q) - q$ 로 획득될 수 있다.

[0064] 손실 계산부(160)는 시멘틱 플로우 획득부(150)에서 생성된 시멘틱 플로우 맵을 분석하여 손실을 계산하고, 계산된 손실을 특징 추출부(120)의 삼 신경망으로 역전파하여 신경망 가중치를 업데이트함으로써, 삼 신경망이 학습되도록 할 수 있다. 본 실시예에서 손실 계산부(160)는 마스크 일관성 손실( $L_{mask}$ ), 플로우 일관성 손실( $L_{flow}$ ) 및 평활화 손실( $L_{smooth}$ )을 각각 계산하고, 계산된 마스크 일관성 손실( $L_{mask}$ ), 플로우 일관성 손실( $L_{flow}$ ) 및 평활

화 손실( $L_{smooth}$ )로부터 총 손실(L)을 획득하여 특징 추출부(120)로 역전파한다.

- [0065] 손실 계산부(160)는 특징 추출부(120)의 삼 신경망을 학습시키기 위한 구성으로, 특징 추출부(120)의 삼 신경망이 학습된 이후에는 생략될 수 있다. 즉 시멘틱 매칭 장치가 실제 이용되는 경우에 손실 계산부(160)는 생략될 수 있다. 또한 손실 계산부(160)는 특징 추출부(120)의 삼 신경망을 학습시키기 위한 별도의 학습 방치로 구성될 수도 있다.
- [0066] 손실 계산부(160)는 미리 지정된 학습 횟수 동안 또는 획득된 총 손실(L)이 기지정된 기준 손실( $L_r$ )이하가 될 때까지, 삼 신경망을 반복학습 시킬 수 있다.
- [0067] 도5 는 도1 의 손실 계산부의 상세 구성을 나타낸다.
- [0068] 도5 를 참조하면, 손실 계산부(160)는 전경 마스크 획득부(161), 마스크 손실 측정부(163), 플로우 일관성 손실 측정부(165), 평활화 손실 측정부(167) 및 손실 역전파부(169)를 포함할 수 있다.
- [0069] 전경 마스크 획득부(161)는 이미지 쌍( $I^s, I^t$ ) 각각에서 이진 전경 마스크(binary foreground mask)를 획득한다. 전경 마스크 획득부(161)는 이진 전경 마스크는 이미지 쌍( $I^s, I^t$ ) 각각에서 전경과 배경을 구분하고, 구분된 전경에 대해서는 1의 비트값을 지정하고, 배경에 대해서는 0의 비트값을 지정하여 전경 마스크( $M^s, M^t$ )를 획득한다. 이미지에서 전경과 배경을 구분하여 이진 전경 마스크를 획득하는 방법은 공지된 기술로서 다양하게 공개되어 있으므로 여기서는 상세한 설명을 생략한다.
- [0070] 마스크 손실 측정부(163)는 전경 마스크 획득부(161)에서 획득된 이진 전경 마스크와 시멘틱 플로우 획득부(150)에서 획득된 시멘틱 플로우 맵을 기반으로 마스크 일관성 손실(mask consistency loss)( $L_{mask}$ )을 측정한다.
- [0071] 마스크 손실 측정부(163)는 시멘틱 플로우 획득부(150)에서 획득된 소스 플로우 필드( $F^s$ )를 이용하여 수학식 6과 같이 타겟 전경 마스크( $M^t$ )를 워핑(warping)하여 추정 소스 마스크( $\hat{M}^s$ )를 획득( $\hat{M}^s = \mathcal{W}(M^t; F^s)$ )한다.

## 수학식 6

[0072] 
$$\hat{M}^s = \mathcal{W}(M^t; F^s)$$

- [0073] 여기서  $\mathcal{W}$ 는 워핑 연산자로서, 픽셀 위치(p)에 대해  $\mathcal{W}(M^t; F^s)(p) = M^t(p + F^s(p))$  연산을 수행하는 연산자이다.
- [0074] 즉 타겟 전경 마스크( $M^t$ )를 소스 플로우 필드( $F^s$ )를 이용한 워핑에 의해 소스 전경 마스크( $M^s$ )에 대응하는 형태로 변환하여 추정 소스 마스크( $\hat{M}^s$ )를 획득한다. 유사하게 타겟 플로우 필드( $F^t$ )를 이용하여 소스 전경 마스크( $M^s$ )를 워핑하여 추정 타겟 마스크( $\hat{M}^t$ )를 획득한다.
- [0075] 그리고 수학식 7과 같이 소스 전경 마스크( $M^s$ )와 추정 소스 마스크( $\hat{M}^s$ ) 사이의 차이 및 타겟 전경 마스크( $M^t$ )와 추정 타겟 마스크( $\hat{M}^t$ ) 사이의 차이의 누적 합을 마스크 일관성 손실( $L_{mask}$ )로 계산한다.

## 수학식 7

[0076] 
$$\mathcal{L}_{mask} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} (M^i(\mathbf{p}) - \hat{M}^i(\mathbf{p}))^2$$

- [0077] 이는 시멘틱 플로우 획득부(150)에서 시멘틱 매칭이 정확하게 수행되었다면, 소스 플로우 필드( $F^s$ ) 및 타겟 플로우 필드( $F^t$ )에 의해 타겟 전경 마스크( $M^t$ )와 소스 전경 마스크( $M^s$ )가 변경된 추정 소스 마스크( $\hat{M}^s$ )와 추정 타



겟 마스크( $\hat{M}^t$ )는 각각 소스 전경 마스크( $M^s$ ) 및 타겟 전경 마스크( $M^t$ )와 유사해야 한다는 마스크 일관성에 따라 측정된 손실이다.

[0078] 도6 은 도5 의 마스크 손실 측정부가 마스크 일관성 손실을 측정하는 개념을 나타낸다.

[0079] 도6 에 도시된 바와 같이, 소스 이미지( $I^s$ )가 타조 이미지이고, 타겟 이미지( $I^t$ )가 독수리 이미지로서 두 이미지가 새라는 의미에서 시멘틱 매칭되는 경우, 소스 이미지( $I^s$ )의 소스 전경 마스크( $M^s$ )가 타겟 플로우 필드( $F^t$ )에 따라 변형되면, 타겟 이미지( $I^t$ )의 독수리 윤곽, 즉 타겟 전경 마스크( $M^t$ )의 형상과 유사하게 나타나야 한다. 또한 타겟 이미지( $I^t$ )의 타겟 전경 마스크( $M^t$ )는 소스 플로우 필드( $F^s$ )에 따라 변형되면, 소스 이미지( $I^s$ )의 타조 윤곽, 즉 소스 전경 마스크( $M^s$ )의 형상과 유사하게 나타나야 한다. 따라서 마스크 손실 측정부(163)는 이들 사이의 오차를 마스크 일관성 손실로 측정할 수 있다.

[0080] 한편, 플로우 일관성 손실 측정부(165)는 대응점에서 소스 플로우 필드( $F^s$ )와 타겟 플로우 필드( $F^t$ )가 일치해야 한다는 플로우 일관성의 오차를 측정한다. 본 실시예에서 시멘틱 매칭 장치는 의미적으로 대응하는 객체 사이의 플로우 필드를 획득하므로, 플로우 일관성 손실 측정부(165) 또한 전경 마스크 획득부(161)에서 획득된 전경 마스크( $M^s$ ,  $M^t$ )의 영역에 기반하여 플로우 일관성 손실을 측정한다.

[0081] 도7 은 플로우 일관성 손실의 개념을 설명하기 위한 도면이다.

[0082] 도7 의 (a)에 도시된 바와 같이, 특정 소스 이미지( $I^s$ )의 픽셀 위치(p)에서 소스 플로우 필드( $F^s(p)$ )에 따른 타겟 이미지( $I^t$ )의 픽셀 위치(q)와, 타겟 이미지( $I^t$ )의 픽셀 위치(q)에서 타겟 플로우 필드( $F^t(q)$ )에 따른 소스 이미지( $I^s$ )의 픽셀 위치(p')가 서로 상이( $p \neq p'$ )하면, 이는 플로우가 일관되지 않는 것으로 시멘틱 매칭에 오류가 발생하였음을 의미한다. 만일 시멘틱 매칭이 정확하게 수행되었다면, (b)에 도시된 바와 같이, 소스 이미지( $I^s$ )의 픽셀 위치(p)에서 소스 플로우 필드( $F^s(p)$ )에 따른 타겟 이미지( $I^t$ )의 픽셀 위치(q)와, 타겟 이미지( $I^t$ )의 픽셀 위치(q)에서 타겟 플로우 필드( $F^t(q)$ )에 따른 소스 이미지( $I^s$ )의 픽셀 위치(p')가 동일( $p = p'$ )하면, 시멘틱 매칭이 정확하게 수행된 것으로 볼 수 있다.

[0083] 이에 플로우 일관성 손실 측정부(165)는 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ )의 차이를 플로우 일관성 손실( $L_{flow}$ )로 측정한다. 이때, 플로우 일관성 손실 측정부(165)는 상기한 바와 같이, 전경 마스크 내의 플로우 필드들( $F^s(p)$ ,  $F^t(q)$ ) 사이의 일관성을 수학식 8에 따라 측정한다.

### 수학식 8

$$\mathcal{L}_{flow} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} \|(\mathcal{F}^i(\mathbf{p}) + \hat{\mathcal{F}}^i(\mathbf{p})) \odot M^i(\mathbf{p})\|_2^2$$

[0084]

[0085] 여기서  $\hat{\mathcal{F}}^s = \mathcal{W}(\mathcal{F}^t; \mathcal{F}^s)$ 로서 소스 플로우 필드( $F^s(p)$ )를 이용하여 타겟 플로우 필드( $F^t(q)$ )를 워핑한 결과를 나타내고,  $\hat{\mathcal{F}}^t$  또한 유사하게 획득될 수 있다. 그리고  $\|\cdot\|_2$ 는 L2 norm 함수를 나타내고,  $\odot$ 는 x 및 y 요소 각각에 대해 개별적 곱셈을 수행하는 element-wise 곱셈을 나타낸다.

[0086] 한편 평활화 손실 측정부(167)는 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 각각에 대해 그래디언트 연산자( $\nabla$ )를 이용하여 평활화 손실( $L_{smooth}$ )을 측정한다.

[0087] 도8 은 평활화 손실의 개념을 설명하기 위한 도면이다.

[0088] 도8 에서 (a)는 평활화 손실이 큰 경우를 나타내고, (b)는 평활화 손실이 적은 경우를 나타낸다. 상기한 바와 같이, 시멘틱 매칭을 수행하는 경우, 의미적으로 동일 범주의 객체간에 대응이 되며, 이는 이미지에서 객체가

포함된 전경이 서로 대응됨을 의미한다. 그리고 시멘틱 매칭에 의해 객체간 형상이 대응하는 경우, 객체 영역, 즉 전경 마스크 영역 내에서 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ )는 (b)에 도시된 바와 같이 매끄럽게 변화하여야 한다. 만일 (a)와 같이, 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ )가 매끄럽게 변화하지 않고 불규칙한 특성을 나타내면, 이는 시멘틱 매칭이 정상적으로 수행되지 않은 것으로 판단할 수 있다. 즉 평활화 손실( $L_{smooth}$ )로서 측정될 수 있다.

[0089] 이에 평활화 손실 측정부(167) 또한 전경 마스크 획득부(161)에서 획득된 전경 마스크( $M^s, M^t$ )의 영역에 기반하여 평활화 손실( $L_{smooth}$ )을 수학적 식 9에 따라 측정한다.

### 수학적 식 9

$$\mathcal{L}_{smooth} = \sum_{i \in \{s, t\}} \sum_{\mathbf{p}} \|\nabla \mathcal{F}^i(\mathbf{p}) \odot M^i(\mathbf{p})\|_1$$

[0090]  $\|\cdot\|_1$ 는 L1 norm 함수를 나타내고,  $\nabla$ 는 그래디언트 연산자를 나타낸다.

[0091] 손실 역전과부(169)는 측정된 마스크 일관성 손실( $L_{mask}$ ), 플로우 일관성 손실( $L_{flow}$ ) 및 평활화 손실( $L_{smooth}$ ) 각각에 기지정된 손실 가중치( $\lambda_{mask}, \lambda_{flow}, \lambda_{smooth}$ )를 적용하고 합하여 총 손실(L)을 계산한다. 그리고 계산된 총 손실(L)을 특징 추출부(120)으로 역전과하여 특징 추출부(120)의 삼 신경망의 가중치를 업데이트 함으로써 학습시킨다.

[0092] 도9 는 도1 의 시멘틱 매칭 장치를 학습시키기 위한 학습 이미지와 소스 및 타겟 이미지에 대한 시멘틱 플로우의 일례를 나타낸다.

[0093] 도9 에서 (a)는 시멘틱 매칭 장치를 학습시키기 위한 학습 이미지의 일례로서, 말과 기수가 전경 마스크로 지정된 경우를 나타낸다. 이러한 학습 이미지는 다수로 획득되어 시멘틱 매칭 장치를 학습시킬 수 있다. 그리고 (b)의 상단은 소스 이미지( $I^s$ )의 일례를 나타내고, 하단은 타겟 이미지( $I^t$ )의 일례를 나타낸다. (a)와 같은 다수의 학습 이미지에 의해 정상적으로 학습된 시멘틱 매칭 장치는 (b)의 이미지 쌍( $I^s, I^t$ )에서 자전거를 시멘틱 객체로 검출할 수 있으며, (c)와 같이 검출된 시멘틱 객체 사이의 픽셀 변화를 소스 플로우 필드( $F^s(p)$ ) 또는 타겟 플로우 필드( $F^t(q)$ )로 획득하여 시멘틱 플로우 맵을 생성할 수 있다.

[0094] 도10 은 본 실시예에 따른 시멘틱 매칭 방법을 나타낸다.

[0095] 도 1 내지 도9 를 참조하여, 도10 의 시멘틱 매칭 방법을 설명하면, 우선 시멘틱 매칭이 수행되어야 할 이미지 쌍( $I^s, I^t$ )을 획득한다(S10). 그리고 동일한 구조를 갖고, 동일하게 패턴 인식 방법이 미리 학습되어 동일한 가중치가 적용되는 삼 신경망 각각에 이미지 쌍( $I^s, I^t$ ) 중 대응하는 이미지를 인가하여 특징 맵 쌍( $f^s, f^t$ )을 획득한다(S20).

[0096] 특징 맵 쌍( $f^s, f^t$ )이 획득되면, 수학적 식 1에 따라 특징 맵 쌍( $f^s, f^t$ )의 각 픽셀 사이의 상관 점수(Correlation score)를 측정하여 상관 맵( $c_p$ )을 생성한다(S30). 여기서 생성된 상관 맵( $c_p$ )의 픽셀 값을 L2 정규화 함수를 이용하여 정규화하여 정규 상관 맵( $n_p$ )을 획득할 수도 있다.

[0097] 상관 맵( $c_p$ )(또는 정규 상관 맵( $n_p$ ))이 획득되면, 획득된 상관 맵의 각 픽셀에서 수학적 식 3과 같이 argmax 함수를 이용하여, 소스 이미지( $I^s$ )에서 픽셀 위치(p)에 대해 상관 점수가 가장 높은 특징( $f^t(q)$ )에 대응하는 상관 픽셀 위치를 확인하고, 확인된 상관 픽셀 위치를 중심으로 인접한 픽셀에 대해 가우시안 분포에 따른 분포 가중치를 반영하는 가우시안 커널( $k_p$ )을 적용한다. 그리고 softmax 함수를 이용하여 매칭 확률( $m_p(q)$ )을 수학적 식 4와 같이 계산한다. 이때, softmax 함수의 출력 분포를 조절하기 위한 온도 파라미터( $\beta$ )를 상관 맵( $c_p$ )(또는 정규 상관 맵( $n_p$ ))에 우선 적용할 수도 있다. 매칭 확률( $m_p(q)$ )이 계산되면, 매칭 확률 맵( $m_p$ )에 따라 픽셀 위치(q)별 매



칭 확률( $m_p(q)$ )을 적용하여 수학식 5에 따라 대응점( $\phi(p)$ )을 판별한다(S40).

- [0099] 대응점이 판별되면, 시멘틱 매칭 장치는 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로부터 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로의 위치 변화를 나타내는 시멘틱 플로우 맵을 획득한다(S50). 이와 함께 타겟 이미지( $I^t$ )의 픽셀 위치( $q$ )로부터 소스 이미지( $I^s$ )의 픽셀 위치( $p$ )로의 위치 변화를 나타내는 시멘틱 플로우 맵을 획득할 수 있다. 즉 시멘틱 매칭을 수행할 수 있다.
- [0100] 한편, 특징 맵 쌍( $f^s, f^t$ )을 획득하는 삼 신경망이 학습되지 않은 경우, 본 실시예에 따른 시멘틱 매칭 방법은 삼 신경망을 학습 시키기 위해 시멘틱 플로우 맵의 손실을 계산하여 역전파하는 단계를 더 포함할 수 있다(S60).
- [0101] 손실 계산 및 역전파 단계(S60)에서는 우선 이미지 쌍( $I^s, I^t$ ) 각각으로부터 이진 전경 마스크( $M^s, M^t$ )를 획득한다.
- [0102] 그리고 시멘틱 플로우 맵 쌍을 구성하는 소스 플로우 필드( $F^s$ )와 타겟 플로우 필드( $F^t$ )를 이용하여 획득된 이진 전경 마스크( $M^s, M^t$ )를 워핑하여, 추정 소스 마스크( $\hat{M}^s$ )와 추정 타겟 마스크( $\hat{M}^t$ )를 획득한다. 추정 소스 마스크( $\hat{M}^s$ )와 추정 타겟 마스크( $\hat{M}^t$ )가 획득되면, 소스 전경 마스크( $M^s$ )와 추정 소스 마스크( $\hat{M}^s$ )와 사이의 차 및 타겟 전경 마스크( $M^t$ )와 추정 타겟 마스크( $\hat{M}^t$ ) 사이의 차를 누적하여 마스크 일관성 손실( $L_{mask}$ )을 계산한다.
- [0103] 한편, 이진 전경 마스크( $M^s, M^t$ )의 대응하는 영역 내의 각 픽셀에 대한 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 사이의 차이를 누적하여 플로우 일관성 손실( $L_{flow}$ )로 측정한다.
- [0104] 또한 이진 전경 마스크( $M^s, M^t$ )의 대응하는 영역 내에서 소스 플로우 필드( $F^s(p)$ )와 타겟 플로우 필드( $F^t(q)$ ) 각각에 대해 그래디언트 연산자( $\nabla$ )를 이용하여 평활화 손실( $L_{smooth}$ )을 측정한다.
- [0105] 측정된 마스크 일관성 손실( $L_{mask}$ ), 플로우 일관성 손실( $L_{flow}$ ) 및 평활화 손실( $L_{smooth}$ ) 각각에 기지정된 손실 가중치( $\lambda_{mask}, \lambda_{flow}, \lambda_{smooth}$ )를 적용하고 합하여 총 손실( $L$ )을 계산하고, 계산된 총 손실( $L$ )을 삼 신경망으로 역전파하여 삼 신경망의 가중치를 업데이트함으로써 학습시킨다.
- [0106] 이하에서는 본 실시예에 따른 시멘틱 매칭 장치 및 방법의 시멘틱 매칭 성능을 시뮬레이션한 결과를 설명한다.
- [0107] 여기서는 삼 신경망이 ImageNet 이미지 데이터 셋의 분류를 위해 학습된 ResNet-101로 구현되는 경우를 가정하였으며,  $320 \times 320$  크기를 갖고 입력되는 이미지 쌍( $I^s, I^t$ )에 대해  $20 \times 20 \times 1024$  및  $20 \times 20 \times 2048$ 의 크기를 갖는 두 개의 특징 맵 쌍( $f^s, f^t$ )을 획득하여,  $20 \times 20 \times 20 \times 20$  크기의 상관 맵( $c_p$ )을 생성하도록 하였다.
- [0108] 그리고 온도 파라미터( $\beta$ )는 50이고, 표준 편차가 5로 설정된 가우시안 커널( $k_p$ )을 적용하였으며, 손실 가중치( $\lambda_{mask}, \lambda_{flow}, \lambda_{smooth}$ )는 각각 1, 10 및 0.5로 지정하였다.
- [0109] 본 실시예에서는 PFWILLOW, PF-PASCAL, Caltech-101 및 TSS 등의 벤치 마크 데이터 셋을 이용하여, 수작업(Hand-crafted)으로 수행된 시멘틱 매칭 결과와 기존의 CNN과 같은 인공 신경망으로만 시멘틱 매칭을 수행한 결과를 본 실시예의 시멘틱 매칭 결과와 비교하였다.
- [0110] 표1 은 PFWILLOW, PF-PASCAL의 벤치 마크 데이터 셋에서의 비교 결과를 나타낸다.

표 1

Type		Methods	PCK ( $\alpha = 0.1$ )	
			WILLOW	PASCAL
Hand-crafted	F	DeepFlow [39]	0.20	0.21
	F	GMK [10]	0.27	0.27
	F	SIFTFlow [30]	0.38	0.33
	F	DSP [26]	0.29	0.30
	F	HOG+PF-LOM [15]	0.56	0.45
CNN-based	A	(T) ResNet-101+CNNGeo [40]	0.68	0.68
	A	(T) ResNet-101+A2Net [42]	0.69	0.67
	A	(T+P) ResNet-101+WS-SA [41]	0.71	0.72
	F	(B+P) FCSS+PF-LOM [27]	0.58	0.46
	F	(M) ResNet-101+Ours	<b>0.74</b>	<b>0.78</b>

[0111]

[0112] 표1 에서는 거리가  $\alpha_{\max}(h, w)$  픽셀 이내 인 키포인트(PCK)의 수에 따라 키포인트(PCK)의 확률을 사용하여 전체적인 매칭의 정확도를 측정하였다.

[0113] 표1 에서 Type 항목의 F와 시멘틱 매칭 결과를 폴로우로 나타냄을 의미하고, A는 정렬로 나타냄을 의미한다. 그리고 Methods 항목에서 () 내부의 문자는 학습을 위해 이용되는 지도 형태를 나타내며, T는 변환 매개 변수, P는 동일 객체 범주의 다른 인스턴스를 나타내는 이미지 쌍을 의미하며, B는 경계 박스를 나타내고, M은 본 실시예에서 이용된 전경 마스크를 의미한다.

[0114] 표1 에 나타난 바와 같이, 본 실시예에 따른 시멘틱 매칭 장치 및 방법은 PFWILLOW, PF-PASCAL의 벤치 마크 데이터 셋에 대해 여러 수작업 기법뿐만 아니라 기존의 인공 신경망 기반 시멘틱 매칭 기법에 비해서도 매우 우수한 성능을 나타냄을 알 수 있다.

[0115] 표2 는 Caltech-101 데이터 세트에서 레이블 전송 정확도(LT-ACC) 및 교차 결합(IoU) 메트릭을 사용하여 매칭 정확도를 평가한 결과를 나타낸다.

표 2

Type		Methods	LT-ACC	IoU
Hand-crafted	F	DeepFlow [39]	0.74	0.40
	F	GMK [10]	0.77	0.42
	F	SIFTFlow [30]	0.75	0.48
	F	DSP [26]	0.77	0.47
	F	HOG+PF-LOM [15]	0.78	0.50
	F	OADSC [46]	0.81	0.55
CNN-based	A	(T) VGG-16+A2Net [42]	0.80	0.57
	A	(T) ResNet-101+CNNGeo [40]	0.83	0.61
	A	(T+P) ResNet-101+WS-SA [41]	0.85	0.63
	F	(C+P) VGG-16+SCNet-AG+ [16]	0.79	0.51
	F	(B+P) FCSS+PF-LOM [27]	0.83	0.52
	F	(M) ResNet-101+Ours	<b>0.88</b>	<b>0.67</b>

[0116]

[0117] 표2 에 나타난 바와 같이, 본 실시예의 시멘틱 매칭 장치 및 방법은 LT-ACC 및 IoU 메트릭의 관점에서도 최상의 성능을 나타냄을 확인할 수 있다.

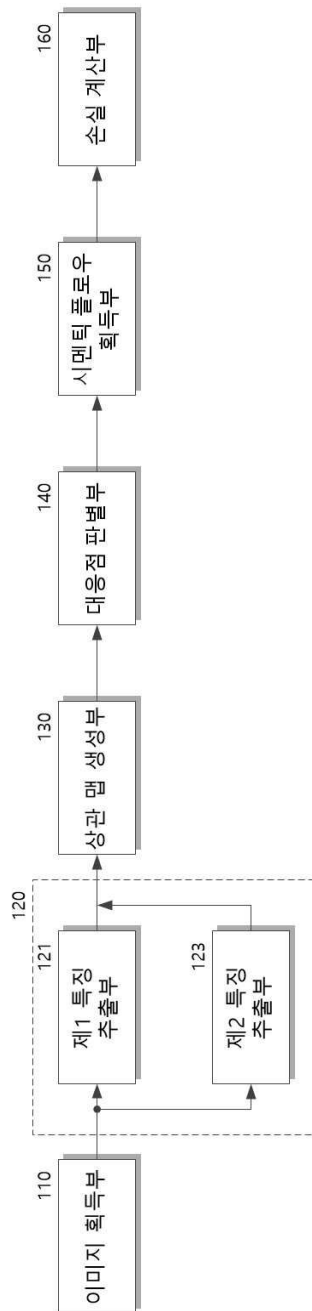
- [0118] 도11 은 소스 이미지와 타겟 이미지 사이의 매칭 결과를 나타낸다.
- [0119] 도11 에서는 PF-PASCAL 데이터 세트의 이미지 쌍에 대한 시멘틱 매칭 결과를 나타낸다. 도11 에서 (a)와 (b)는 각각 소스 이미지( $I^s$ )와 타겟 이미지( $I^t$ )를 나타내고, (c) 내지 (e)는 각각 기존의 시멘틱 매칭 기법인 CNNGeo와 A2Net 및 WS-SA 기법으로 획득된 시멘틱 매칭 결과를 나타내며, (f)는 본 실시예에 따라 획득된 시멘틱 매칭 결과를 나타낸다.
- [0120] 도11 에서 (c) 내지 (f)에서는 시멘틱 매칭 정확도를 시각적으로 표현하기 위해, 소스 이미지( $I^s$ )를 소스 플로우 필드( $F^s$ )에 따라 변형하여 나타내었으며, 그 결과 도11 의 (f)에 도시된 바와 같이, 본 실시예의 시멘틱 매칭 기법이 기존에 비해 우수한 성능을 나타냄을 확인할 수 있다.
- [0121] 본 발명에 따른 방법은 컴퓨터에서 실행 시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 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.
- [0122] 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.
- [0123] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

### 부호의 설명

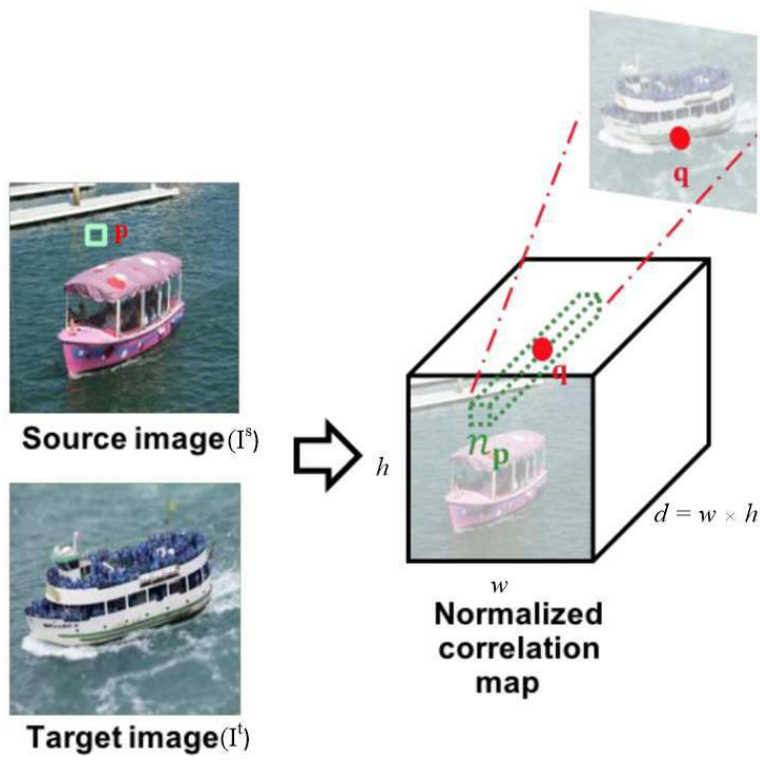
- [0124] 110: 이미지 획득부 120: 특징 추출부
- 130: 상관 맵 생성부 140: 대응점 판별부
- 150: 시멘틱 플로우 획득부 160: 손실 계산부
- 121: 제1 특징 추출부 123: 제2 특징 추출부
- 141: 상관 픽셀 추출부 143: 온도 파라미터 적용부
- 145: 가우시안 커널부 147: 매칭 확률 맵 생성부
- 149: 대응점 획득부 161: 전경 마스크 획득부
- 163: 마스크 손실 측정부 165: 플로우 일관성 손실 측정부
- 167: 평활화 손실 측정부 169: 손실 역전파부

도면

도면1



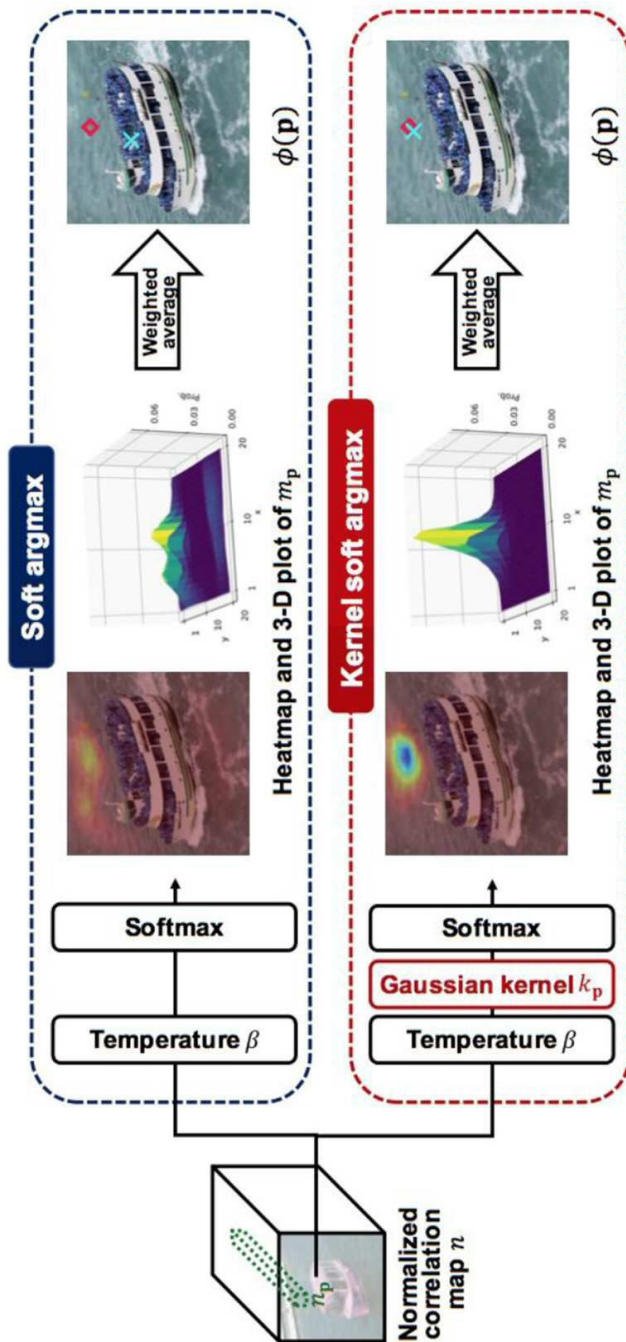
도면2



도면3

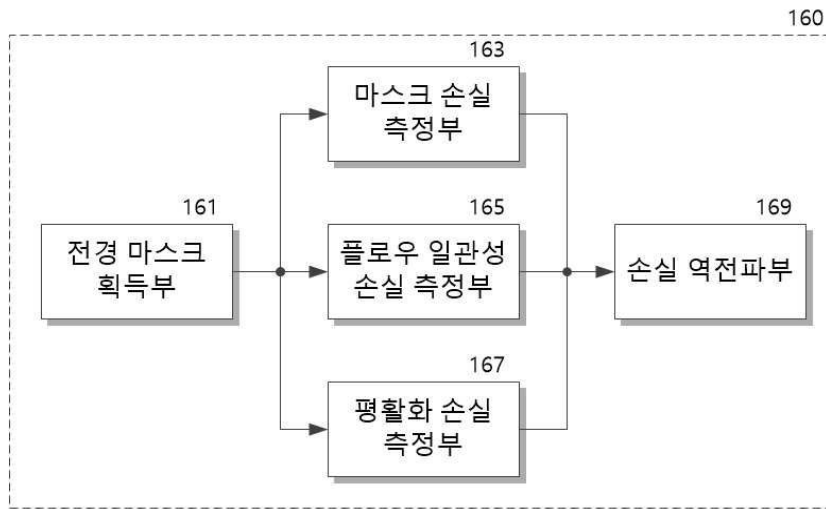


도면4

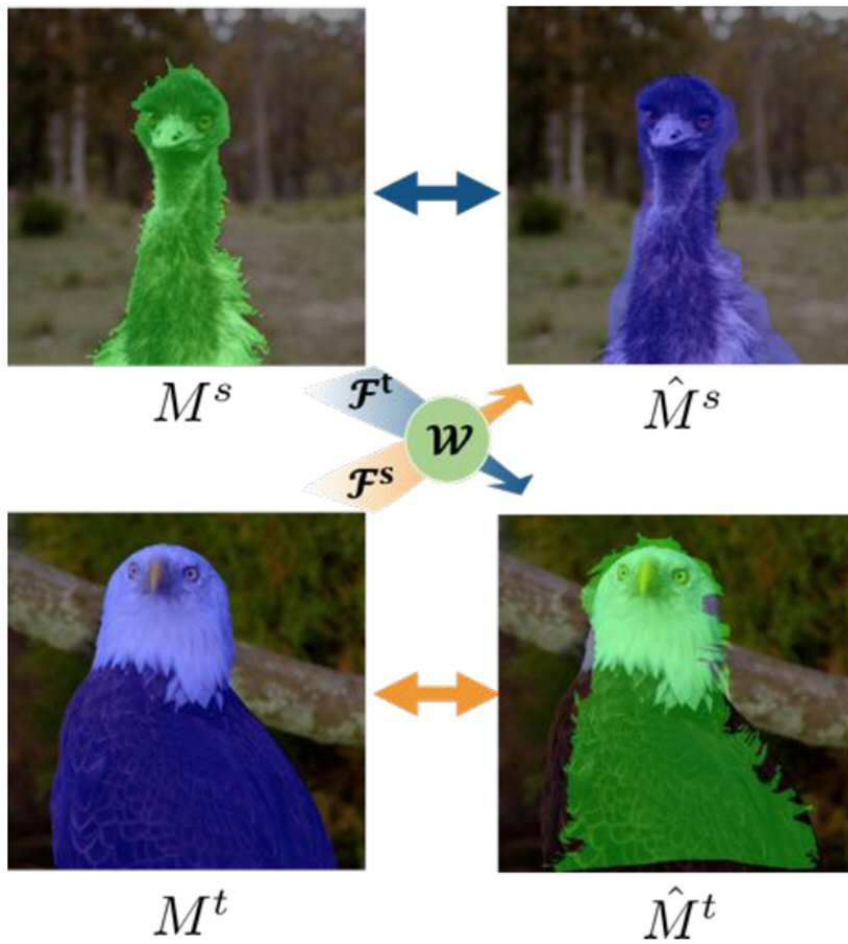




도면5

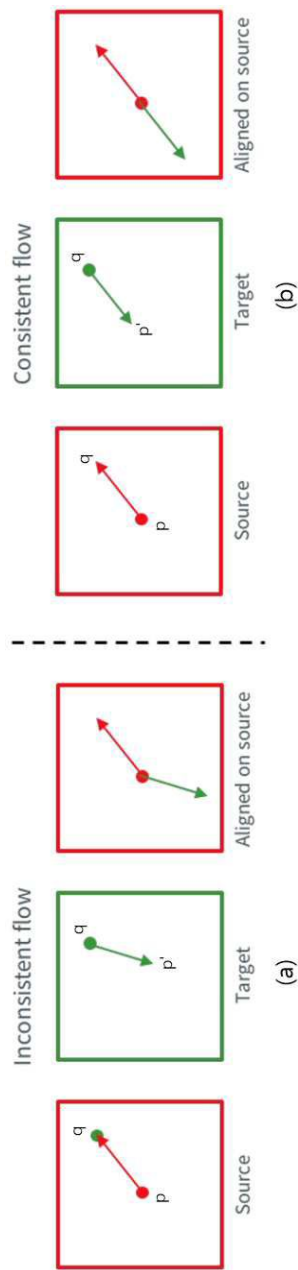


도면6

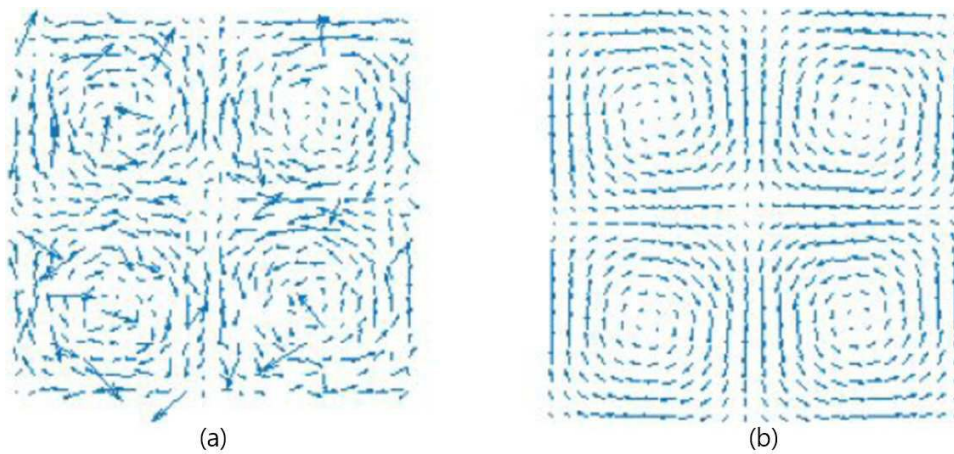




도면7



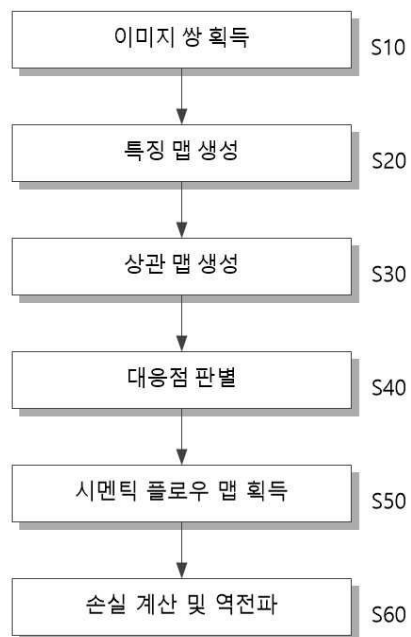
도면8



도면9



도면10



도면11

