



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0119392  
(43) 공개일자 2020년10월20일

(51) 국제특허분류(Int. Cl.)  
G06N 3/08 (2006.01) G06T 7/246 (2017.01)  
(52) CPC특허분류  
G06N 3/084 (2013.01)  
G06T 7/246 (2017.01)  
(21) 출원번호 10-2019-0034833  
(22) 출원일자 2019년03월27일  
심사청구일자 2019년03월27일

(71) 출원인  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
손광훈  
서울특별시 서대문구 연세로 50, 제3공학관 C129호(신촌동, 연세대학교)  
전상률  
서울특별시 서대문구 연세로 50, 제3공학관 C129호(신촌동, 연세대학교)  
(74) 대리인  
민영준

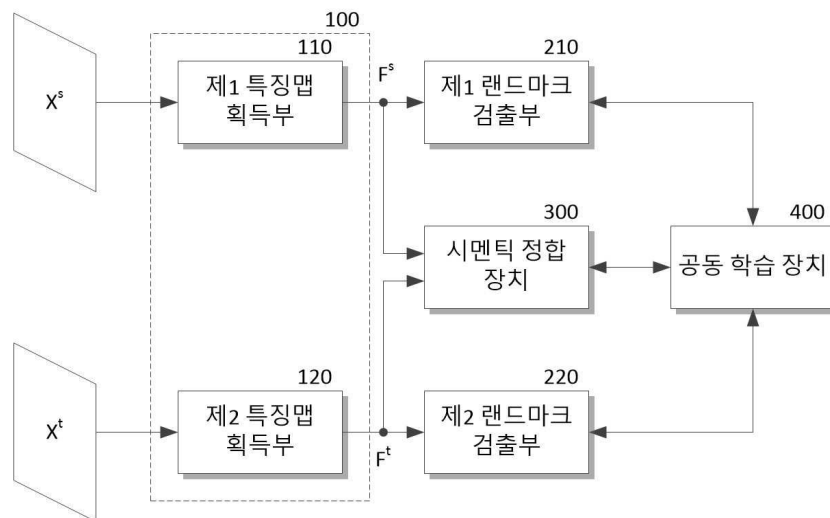
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 장치 및 방법

(57) 요약

본 발명은 비지도 학습 방식으로 객체 랜드마크 검출 장치 및 시멘틱 정합 장치를 학습시킬 수 있어 학습용 데이터를 요구하지 않으며, 학습 시에 시멘틱 정합 장치와 객체 랜드마크 검출 장치가 상호 보완되도록 함으로써, 비지도 학습 방식으로도 우수한 학습 성능을 제공할 수 있는 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 장치 및 방법을 제공할 수 있다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호	2017M3C4A7069370
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	원천기술개발사업
연구과제명	[Ezbaro] (2세부)딥러닝 기반 의미론적 상황 이해 원천기술 연구 (1단계)(2/2)
기 여 율	1/1
과제수행기관명	연세대학교 산학협력단
연구기간	2018.06.01 ~ 2019.05.31

---

## 명세서

### 청구범위

#### 청구항 1

입력 영상들에 대한 특징맵 각각에 자기 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 획득되는 랜드마크 확률 맵으로부터 입력 영상의 다수의 랜드마크의 위치를 추정하기 위한 랜드마크 기대값을 획득하는 객체 랜드마크 검출부;

상기 입력 영상들에 대한 특징맵의 상호 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인턴 인식 방법에 따라 입력 영상들에서 의미론적 대응점의 위치 변화를 나타내는 변환 필드 맵과 상기 변환 필드 맵의 신뢰도를 나타내는 신뢰도 맵을 획득하는 시멘틱 정합부; 및

상기 객체 랜드마크 검출부에 대한 랜드마크 손실과 상기 시멘틱 정합부에 대한 정합 손실 및 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부의 공동 손실을 획득하고 가산하여 총 손실을 획득하며, 상기 총 손실을 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부 각각으로 역전파하여 학습시키는 공동 학습부; 를 포함하는 공동 학습 장치.

#### 청구항 2

제1 항에 있어서, 상기 객체 랜드마크 검출부는

상기 특징맵 각각에 대해 기지정된 크기의 로컬 탐색창을 이용한 자기 상관을 수행하여 자기 상관 스코어로 구성되는 자기 상관 맵을 획득하는 자기 상관부;

상기 특징맵과 대응하는 자기 상관 맵을 결합하고, 미리 학습된 패턴 추정 방식에 따라 결합된 특징맵과 자기 상관 맵에 랜드마크 가중치를 적용하여, 상기 입력 영상들의 기지정된 K개의 랜드마크에 대한 분산 스코어 맵을 추정하는 스코어 맵 획득부;

상기 분산 스코어 맵을 정규화하여 랜드마크 확률 맵을 획득을 획득하는 정규화부; 및

상기 랜드마크 확률 맵에 따라 공간 좌표 각각에 대한 기대값으로 구성되는 기대값 맵을 획득하는 기대값 획득부; 를 포함하는 공동 학습 장치.

#### 청구항 3

제2 항에 있어서, 상기 시멘틱 정합부는

상기 특징맵에 대해 기지정된 크기의 로컬 탐색창을 이용한 상호 상관을 수행하여 상호 상관 스코어로 구성되는 상호 상관 맵을 획득하는 상호 상관부;

미리 학습된 패턴 추정 방식에 따라 상기 상호 상관 맵에 변환 가중치를 적용하여, 특징맵의 각 픽셀에서의 상호 대응점의 위치 변화를 나타내는 변환값으로 구성되는 변환 필드 맵을 획득하는 변환 필드 획득부; 및

미리 학습된 패턴 추정 방식에 따라 상기 상호 상관 맵에 신뢰도 가중치를 적용하여, 각 픽셀에서의 변환 필드 맵의 신뢰도로 구성되는 신뢰도 맵을 획득하는 신뢰도 맵 획득부; 를 포함하는 공동 학습 장치.

#### 청구항 4

제3 항에 있어서, 상기 공동 학습부는

상기 객체 랜드마크 검출부에서 획득되는 기대값과 이에 대응하는 분산 스코어 맵에 따른 집중 손실과 서로 다른 2개의 랜드마크에 대한 기대값 사이의 거리에 따른 이산 손실을 획득하고, 상기 집중 손실과 상기 이산 손실에 각각 집중 손실 가중치와 이산 손실 가중치를 가중하고 합하여 상기 랜드마크 손실을 획득하는 랜드마크 손실 측정부;

변환 필드 맵에 의해 와프된 입력 영상들의 특징에 대한 교차 엔트로피 손실을 기반으로 상기 정합 손실을 획득하는 정합 손실 측정부;

변환 필드 맵에 의해 와프된 입력 영상들의 랜드마크들의 거리를 기반으로 거리 손실을 획득하고, 상기 거리 손실에 거리 손실 가중치를 가중하여 상기 공동 손실을 획득하는 공동 손실 측정부; 및

상기 랜드마크 손실과 상기 정합 손실 및 상기 공동 손실을 가산하여 총 손실을 획득하고, 상기 총 손실을 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부 각각으로 역전파하는 총 손실 역전파부; 를 포함하는 공동 학습 장치.

#### 청구항 5

제4 항에 있어서, 상기 공동 학습부는

학습시에 상기 객체 랜드마크 검출부의 랜드마크 가중치를 고정하고, 상기 입력 영상들에 대한 상기 총 손실을 획득하여, 상기 시멘틱 정합부로 역전파하여 상기 시멘틱 정합부를 학습시키고,

이후, 상기 시멘틱 정합부의 변환 가중치와 신뢰도 가중치를 고정하고, 상기 입력 영상들에 대한 상기 총 손실을 획득하여, 상기 객체 랜드마크 검출부로 역전파하여 상기 객체 랜드마크 검출부를 학습시키며,

상기 총 손실이 기지정된 기준 총 손실 이상이면, 상기 시멘틱 정합부와 상기 객체 랜드마크 검출부를 교대로 반복 학습시키는 공동 학습 장치.

#### 청구항 6

제5 항에 있어서, 상기 공동 학습부는

상기 시멘틱 정합부 학습시에 상기 신뢰도 가중치와 상기 변환 가중치를 교대로 고정하여, 상기 신뢰도 맵 획득부와 상기 변환 필드 획득부를 교대로 학습시키는 공동 학습 장치.

#### 청구항 7

제1 항에 있어서, 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부 각각은

동일 구조를 갖고 패턴 인식 방법이 미리 학습되어 동일한 학습 가중치를 갖는 한 쌍의 삼 신경망을 포함하여, 상기 입력 영상들을 인가받아 상기 특징맵을 획득하는 특징 추출부; 를 더 포함하는 공동 학습 장치.

#### 청구항 8

입력 영상들에 대한 특징맵 각각에 자기 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 획득되는 랜드마크 확률 맵으로부터 입력 영상의 다수의 랜드마크의 위치를 추정하기 위해 획득된 랜드마크 기대값에 대한 오차인 랜드마크 손실을 획득하는 단계;

상기 입력 영상들에 대한 특징맵의 상호 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 획득되는 입력 영상들에서 의미론적 대응점의 위치 변화를 나타내는 변환 필드 맵과 상기 변환 필드 맵의 신뢰도를 나타내는 신뢰도 맵의 오차인 정합 손실을 획득하는 단계;

상기 랜드마크 기대값과 상기 변환 필드 맵 및 상기 신뢰도 맵의 공동 오차인 공동 손실을 획득하는 단계; 및

상기 랜드마크 손실과 상기 정합 손실 및 상기 공동 손실을 가산하여 총 손실을 획득하고, 상기 총 손실을 역전파하는 단계; 를 포함하는 공동 학습 방법.

#### 청구항 9

제8 항에 있어서, 상기 랜드마크 손실을 획득하는 단계는

상기 특징맵 각각에 대해 기지정된 크기의 로컬 탐색창을 이용한 자기 상관을 수행하여 자기 상관 스코어로 구성되는 자기 상관 맵을 획득하는 단계;

상기 특징맵과 대응하는 자기 상관 맵을 결합하고, 미리 학습된 패턴 추정 방식에 따라 결합된 특징맵과 자기 상관 맵에 랜드마크 가중치를 적용하여, 상기 입력 영상들의 기지정된 K개의 랜드마크에 대한 분산 스코어 맵을 추정하는 단계;

상기 분산 스코어 맵을 정규화하여 랜드마크 확률 맵을 획득하는 단계;

상기 랜드마크 확률 맵에 따라 공간 좌표 각각에 대한 기대값으로 구성되는 기대값 맵을 획득하는 단계;

기대값과 이에 대응하는 분산 스코어 맵에 따른 집중 손실을 획득하는 단계;

서로 다른 2개의 랜드마크에 대한 기대값 사이의 거리에 따른 이산 손실을 획득하는 단계; 및

상기 집중 손실과 상기 이산 손실에 각각 집중 손실 가중치와 이산 손실 가중치를 가중하고 합하여 상기 랜드마크 손실을 획득하는 단계; 를 포함하는 공동 학습 방법.

#### 청구항 10

제9 항에 있어서, 상기 정합 손실을 획득하는 단계는

상기 특징맵에 대해 기지정된 크기의 로컬 탐색창을 이용한 상호 상관을 수행하여 상호 상관 스코어로 구성되는 상호 상관 맵을 획득하는 단계;

미리 학습된 패턴 추정 방식에 따라 상기 상호 상관 맵에 변환 가중치를 적용하여, 특징맵의 각 픽셀에서의 상호 대응점의 위치 변화를 나타내는 변환값으로 구성되는 변환 필드 맵을 획득하는 단계;

미리 학습된 패턴 추정 방식에 따라 상기 상호 상관 맵에 신뢰도 가중치를 적용하여, 각 픽셀에서의 변환 필드 맵의 신뢰도로 구성되는 신뢰도 맵을 획득하는 단계; 및

변환 필드 맵에 의해 와프된 입력 영상들의 특징에 대한 교차 엔트로피 손실을 기반으로 상기 정합 손실을 획득하는 단계; 를 포함하는 공동 학습 방법.

#### 청구항 11

제10 항에 있어서, 상기 공동 손실을 획득하는 단계는

상기 변환 필드 맵에 의해 와프된 입력 영상들의 랜드마크들의 거리를 기반으로 거리 손실을 획득하는 단계; 및

상기 거리 손실에 거리 손실 가중치를 가중하여 상기 공동 손실을 획득하는 단계; 를 포함하는 공동 학습 방법.

#### 청구항 12

제11 항에 있어서, 상기 총 손실을 역전파하는 단계는

상기 입력 영상들에 대한 상기 총 손실을 역전파하되, 상기 랜드마크 가중치를 고정하는 단계;

상기 입력 영상들에 대한 상기 총 손실을 역전파하되, 상기 변환 가중치 및 상기 신뢰도 가중치를 고정하는 단계; 및

상기 총 손실이 기지정된 기준 총 손실 이상이면, 상기 랜드마크 가중치를 고정하는 단계와 상기 변환 가중치 및 상기 신뢰도 가중치를 고정하는 단계를 교대로 반복하여 학습시키는 공동 학습 방법.

#### 청구항 13

제12 항에 있어서, 상기 랜드마크 가중치를 고정하는 단계는

상기 신뢰도 가중치와 상기 변환 가중치를 교대로 고정하여 학습시키는 공동 학습 방법.

### 발명의 설명

#### 기술 분야

[0001] 본 발명은 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 장치 및 방법에 관한 것으로, 객체 랜드마크 검출 장치 및 시멘틱 정합 장치를 상호 보완하도록 하여 비지도 학습 방식으로 학습을 수행할 수 있는 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 장치 및 방법에 관한 것이다.

#### 배경 기술

[0002] 객체 랜드마크 검출과 시멘틱 정합(Semantic Alignment)은 각각 이미지 처리(image processing) 및 컴퓨터 비전(computer vision) 분야의 다양한 어플리케이션에서 이용되고 있다.

- [0003] 객체 랜드마크 검출은 이미지에 포함된 객체를 식별할 수 있는 특징의 위치를 나타내는 랜드마크를 검출하는 기법을 의미하고, 시멘틱 정합은 서로 다른 이미지에서 동일한 객체 또는 의미적으로 동일한 범주에 속하는 다양한 형상의 객체를 탐색하여 객체의 형상 변화, 즉 대응점의 위치 변화를 탐색하는 기법을 의미한다.
- [0004] 한편 최근에는 딥 러닝(Deep learning) 기법으로 인공 신경망(artificial neural network)을 학습시켜 이미지 처리에 적용하고자 하는 연구가 활발히 수행되고 있다. 이러한 연구로 인해 객체 랜드마크 검출과 시멘틱 정합에도 인공 신경망이 적용됨으로써 성능이 획기적으로 향상되고 있다. 이때 객체 랜드마크 검출과 시멘틱 정합이 안정적인 성능을 나타내기 위해서는 우선 인공 신경망이 학습되어야 한다. 그리고 인공 신경망을 학습시키기 위해서는 대량의 검증된 학습 데이터(groundtruth training data)가 요구된다.
- [0005] 그러나 일반적으로 학습 데이터는 수작업에 의해 생성되므로, 대량으로 획득하는 것이 매우 어렵다. 또한 기존에는 객체 랜드마크 검출과 시멘틱 정합을 구분하여 개별적으로 연구를 수행하였으며, 이로 인해 객체 랜드마크 검출을 위한 학습 데이터와 시멘틱 정합을 위한 학습 데이터 각각이 별도로 필요하였다. 즉 객체 랜드마크 검출과 시멘틱 정합 각각을 위한 대량의 학습 데이터가 별도로 필요하다는 문제가 있다.

## 선행기술문헌

### 특허문헌

- [0006] (특허문헌 0001) 한국 공개 특허 제10-2018-0033037호 (2018.04.02 공개)

## 발명의 내용

### 해결하려는 과제

- [0007] 본 발명의 목적은 학습용 데이터를 요구하지 않는 비지도 학습 방식으로 객체 랜드마크 검출 장치 및 시멘틱 정합 장치를 함께 학습시킬 수 있는 공동 학습 장치 및 방법을 제공하는데 있다.
- [0008] 본 발명의 다른 목적은 객체 랜드마크 검출 장치와 시멘틱 정합 장치가 학습 시에 상호 보완되도록 하여, 비지도 학습 방식에서도 우수한 성능을 갖도록 하는 공동 학습 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0009] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 객체 랜드마크 검출 장치와 시멘틱 정합 장치의 공동 학습 장치는 입력 영상들에 대한 특징맵 각각에 자기 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 획득되는 랜드마크 확률 맵으로부터 입력 영상의 다수의 랜드마크의 위치를 추정하기 위한 랜드마크 기대값을 획득하는 객체 랜드마크 검출부; 상기 입력 영상들에 대한 특징맵의 상호 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 입력 영상들에서 의미론적 대응점의 위치 변화를 나타내는 변환 필드 맵과 상기 변환 필드 맵의 신뢰도를 나타내는 신뢰도 맵을 획득하는 시멘틱 정합부; 및 상기 객체 랜드마크 검출부에 대한 랜드마크 손실과 상기 시멘틱 정합부에 대한 정합 손실 및 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부의 공동 손실을 획득하고 가산하여 총 손실을 획득하며, 상기 총 손실을 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부 각각으로 역전파하여 학습시키는 공동 학습부;를 포함한다.
- [0010] 상기 객체 랜드마크 검출부는 상기 특징맵 각각에 대해 기지정된 크기의 로컬 탐색창을 이용한 자기 상관을 수행하여 자기 상관 스코어로 구성되는 자기 상관 맵을 획득하는 자기 상관부; 상기 특징맵과 대응하는 자기 상관 맵을 결합하고, 미리 학습된 패턴 추정 방식에 따라 결합된 특징맵과 자기 상관 맵에 랜드마크 가중치를 적용하여, 상기 입력 영상들의 기지정된 K개의 랜드마크에 대한 분산 스코어 맵을 추정하는 스코어 맵 획득부; 상기 분산 스코어 맵을 정규화하여 랜드마크 확률 맵을 획득하는 정규화부; 및 상기 랜드마크 확률 맵에 따라 공간 좌표 각각에 대한 기대값으로 구성되는 기대값 맵을 획득하는 기대값 획득부;를 포함할 수 있다.
- [0011] 상기 시멘틱 정합부는 상기 특징맵에 대해 기지정된 크기의 로컬 탐색창을 이용한 상호 상관을 수행하여 상호 상관 스코어로 구성되는 상호 상관 맵을 획득하는 상호 상관부; 미리 학습된 패턴 추정 방식에 따라 상기 상호 상관 맵에 변환 가중치를 적용하여, 특징맵의 각 픽셀에서의 상호 대응점의 위치 변화를 나타내는 변환값으로 구성되는 변환 필드 맵을 획득하는 변환 필드 획득부; 및 미리 학습된 패턴 추정 방식에 따라 상기 상호 상관 맵에 신뢰도 가중치를 적용하여, 각 픽셀에서의 변환 필드 맵의 신뢰도로 구성되는 신뢰도 맵을 획득하는 신뢰

도 맵 획득부; 를 포함할 수 있다.

[0012] 상기 공동 학습부는 상기 객체 랜드마크 검출부에서 획득되는 기대값과 이에 대응하는 분산 스코어 맵에 따른 집중 손실과 서로 다른 2개의 랜드마크에 대한 기대값 사이의 거리에 따른 이산 손실을 획득하고, 상기 집중 손실과 상기 이산 손실에 각각 집중 손실 가중치와 이산 손실 가중치를 가중하고 합하여 상기 랜드마크 손실을 획득하는 랜드마크 손실 측정부; 변환 필드 맵에 의해 와프된 입력 영상들의 특징에 대한 교차 엔트로피 손실을 기반으로 상기 정합 손실을 획득하는 정합 손실 측정부; 변환 필드 맵에 의해 와프된 입력 영상들의 랜드마크들의 거리를 기반으로 거리 손실을 획득하고, 상기 거리 손실에 거리 손실 가중치를 가중하여 상기 공동 손실을 획득하는 공동 손실 측정부; 및 상기 랜드마크 손실과 상기 정합 손실 및 상기 공동 손실을 가산하여 총 손실을 획득하고, 상기 총 손실을 상기 객체 랜드마크 검출부와 상기 시멘틱 정합부 각각으로 역전파하는 총 손실 역전파부; 를 포함할 수 있다.

[0013] 상기 공동 학습부는 학습시에 상기 객체 랜드마크 검출부의 랜드마크 가중치를 고정하고, 상기 입력 영상들에 대한 상기 총 손실을 획득하여, 상기 시멘틱 정합부로 역전파하여 상기 시멘틱 정합부를 학습시키고, 이후, 상기 시멘틱 정합부의 변환 가중치와 신뢰도 가중치를 고정하고, 상기 입력 영상들에 대한 상기 총 손실을 획득하여, 상기 객체 랜드마크 검출부로 역전파하여 상기 객체 랜드마크 검출부를 학습시키며, 상기 총 손실이 기지정된 기준 총 손실 이상이면, 상기 시멘틱 정합부와 상기 객체 랜드마크 검출부를 교대로 반복 학습시킬 수 있다.

[0014] 상기 공동 학습부는 상기 시멘틱 정합부 학습시에 상기 신뢰도 가중치와 상기 변환 가중치를 교대로 고정하여, 상기 신뢰도 맵 획득부와 상기 변환 필드 획득부를 교대로 학습시킬 수 있다.

[0015] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 객체 랜드마크 검출 장치와 시멘틱 정합 장치의 공동 학습 방법은 입력 영상들에 대한 특징맵 각각에 자기 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 획득되는 랜드마크 확률 맵으로부터 입력 영상의 다수의 랜드마크의 위치를 추정하기 위해 획득된 랜드마크 기대값에 대한 오차인 랜드마크 손실을 획득하는 단계; 상기 입력 영상들에 대한 특징맵의 상호 상관을 수행한 결과를 기반으로, 미리 학습된 패턴 인식 방법에 따라 획득되는 입력 영상들에서 의미론적 대응점의 위치 변화를 나타내는 변환 필드 맵과 상기 변환 필드 맵의 신뢰도를 나타내는 신뢰도 맵의 오차인 정합 손실을 획득하는 단계; 상기 랜드마크 기대값과 상기 변환 필드 맵 및 상기 신뢰도 맵의 공동 오차인 공동 손실을 획득하는 단계; 및 상기 랜드마크 손실과 상기 정합 손실 및 상기 공동 손실을 가산하여 총 손실을 획득하고, 상기 총 손실을 역전파하는 단계; 를 포함한다.

### 발명의 효과

[0016] 따라서, 본 발명의 실시예에 따른 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 장치 및 방법은 비지도 학습 방식으로 객체 랜드마크 검출 장치 및 시멘틱 정합 장치를 학습시킬 수 있어 학습용 데이터를 요구하지 않는다. 또한 학습 시에 시멘틱 정합 장치와 객체 랜드마크 검출 장치가 상호 보완되도록 함으로써, 비지도 학습 방식으로도 우수한 학습 성능을 제공할 수 있다.

### 도면의 간단한 설명

[0017] 도 1은 본 발명의 일 실시예에 따른 객체 랜드마크 검출 장치 및 시멘틱 정합 장치를 학습시키기 위한 공동 학습 장치의 일예를 나타낸다.

도 2는 도 1의 객체 랜드마크 검출 장치의 상세 구성의 일예를 나타낸다.

도 3은 도 2의 자기 상관부에서 획득되는 자기 상관 맵의 특징을 설명하기 위한 도면이다.

도 4는 도 1의 시멘틱 정합 장치의 상세 구성의 일예를 나타낸다.

도 5는 도 1의 공동 학습 장치의 상세 구성의 일예를 나타낸다.

도 6은 본 발명의 일 실시예에 따른 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 방법을 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 부재를 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] 도 1은 본 발명의 일 실시예에 따른 객체 랜드마크 검출 장치 및 시멘틱 정합 장치를 학습시키기 위한 공동 학습 장치의 일예를 나타내고, 도 2는 도 1의 객체 랜드마크 검출 장치의 상세 구성의 일예를 나타내며, 도 3은 도 2의 자기 상관부에서 획득되는 자기 상관 맵의 특징을 설명하기 위한 도면이다. 그리고 도 4는 도 1의 시멘틱 정합 장치의 상세 구성의 일예를 나타내고, 도 5는 도 1의 공동 학습 장치의 상세 구성의 일예를 나타낸다.
- [0022] 도 1을 참조하면, 본 실시예에 따른 공동 학습 장치는 공동 특징맵 획득부(100), 객체 랜드마크 검출부(200), 시멘틱 정합부(300) 및 공동 학습부(400)를 포함한다.
- [0023] 공동 특징맵 획득부(100)는 다수의 입력 영상( $X^s, X^t$ )을 인가받고, 미리 학습된 패턴 추정 방식에 따라 인가된 입력 영상( $X^s, X^t$ )에서 특징을 추출하여 특징맵( $F^s, F^t$ )을 생성한다. 여기서 입력 영상( $X^s, X^t$ )는 시멘틱 정합을 위한 소스 영상( $X^s$ )과 타겟 영상( $X^t$ )일 수 있다.
- [0024] 공동 특징맵 획득부(100)는 도 1에 도시된 바와 같이, 다수의 특징맵 획득부(110, 120)를 포함할 수 있으며, 다수의 특징맵 획득부(110, 120) 각각은 인공 신경망으로 구현되어, 다수의 입력 영상( $X^s, X^t$ ) 중 대응하는 입력 영상에서 특징을 추출하여 각각 특징맵( $F^s, F^t$ )을 생성하여 출력한다. 다수의 특징맵 획득부(110, 120)은 각각 픽셀( $i \in \mathbb{R}^{H \times W}$ )를 갖는  $H \times W$  크기의 입력 영상( $X^s, X^t$ )에 대한  $C$  차원의 특징맵( $F^s, F^t$ )  $\in \mathbb{R}^{H \times W \times C}$ 을 생성할 수 있다.
- [0025] 여기서 다수의 특징맵 획득부(110, 120)는 일예로 컨볼루션 신경망(Convolution neural network: 이하 CNN)으로 구현될 수 있으며, 특히 동일한 구조를 갖고, 동시에 학습되어 동일한 공동 가중치( $W_F$ )가 적용되는 삼(siamese) 신경망으로 구현될 수 있다. 즉 다수의 특징맵 획득부(110, 120)는 다수의 입력 영상( $X^s, X^t$ ) 각각에 대해 동일한 특징을 추출할 수 있도록 삼 신경망으로 구현될 수 있으며, 특징맵( $F^s, F^t$ )은 입력 영상( $X^s, X^t$ ) 각각에 공동 가중치( $W_F$ )가 적용되어  $F^s = F(X^s; W_F)$ ,  $F^t = F(X^t; W_F)$ 로 획득될 수 있다. 여기서  $F()$ 는 인공 신경망의 연산 함수를 나타낸다.
- [0026] 도 1에서는 설명의 편의를 위하여 공동 특징맵 획득부(100)가 객체 랜드마크 검출부(200) 및 시멘틱 정합부(300)와 별개의 구성으로 도시되어 있으나, 실제 공동 특징맵 획득부(100)는 객체 랜드마크 검출부(200) 및 시멘틱 정합부(300)에 각각 포함되어 구성되는 공통 요소이다.
- [0027] 즉 객체 랜드마크 검출부(200) 및 시멘틱 정합부(300)는 각각 동일한 공동 특징맵 획득부(100)를 포함하도록 구성된다.
- [0028] 공동 특징맵 획득부(100)에 포함되는 특징맵 획득부(110, 120)의 개수는 제한되지 않으나, 시멘틱 정합부(300)가 일반적으로 2개의 입력 영상( $X^s, X^t$ )을 상호 비교하여 정합을 수행하고, 객체 랜드마크 검출부(200)는 하나의 입력 영상에서 랜드마크를 검출할 수 있으므로, 도 1에 도시된 바와 같이, 2개의 특징맵 획득부(110, 120)를 포함하는 것이 일반적이다.
- [0029] 객체 랜드마크 검출부(200)는 공동 특징맵 획득부(100)에서 획득된 다수의 특징맵( $F^s, F^t$ )을 인가받고, 인가된 특징맵( $F^s, F^t$ ) 각각에 대한 자기 상관(auto-correlation)을 기반으로 획득되는 랜드마크 확률 맵( $\psi^s, \psi^t$ )으로부터 입력 영상( $X^s, X^t$ )의 픽셀별 랜드마크 기대값( $\phi^s, \phi^t$ )을 획득한다.
- [0030] 도 2를 참조하면, 객체 랜드마크 검출부(200)는 다수(여기서는 2개)의 랜드마크 검출부(210, 220)를 포함할 수



있다. 다만 객체 랜드마크 검출부(200)가 다수의 랜드마크 검출부(210, 220)를 포함하는 것은 비지도 학습을 수행하기 위한 것으로, 객체 랜드마크 검출부(200)가 객체 랜드마크 검출 장치로 이용되는 경우에는 하나의 랜드마크 검출부만을 포함하도록 구성될 수 있다. 즉 다수의 랜드마크 검출부(210, 220) 중 하나가 객체 랜드마크 검출 장치로 이용될 수 있다. 그리고 상기한 바와 같이, 객체 랜드마크 검출부(200)는 학습이 완료된 이후, 객체 랜드마크 검출 장치로 이용되는 경우에, 공동 특징맵 획득부(100)의 다수의 특징맵 획득부(110, 120) 중 적어도 하나를 더 포함할 수 있다.

[0031] 다수의 랜드마크 검출부(210, 220)는 각각 자기 상관부(211, 221), 스코어 맵 획득부(212, 222), 정규화부(213, 223) 및 기대값 획득부(214, 224)를 포함할 수 있다.

[0032] 자기 상관부(211, 221)는 각각 C 차원의 특징맵( $F^s, F^t$ ) 중 대응하는 특징맵을 인가받고, 인가된 특징맵에 대해 수학적 1의 상관 함수를 기반으로 기지정된 크기의 로컬 탐색창( $N_i$ )을 이용한 자기 상관을 수행하여 자기 상관 스코어로 구성되는 자기 상관 맵( $C^{ss}, C^{tt}$ )을 획득한다. 여기서 로컬 탐색창( $N_i$ )은 특징맵( $F^s, F^t$ )의 픽셀(i)을 중심으로 하는 기지정된 크기의 자기 상관 탐색창을 의미한다.

### 수학적 식 1

$$C(F_i^A, F_j^B) = \langle F_i^A, F_j^B \rangle / \sqrt{\sum_l \langle F_i^A, F_l^B \rangle^2}$$

[0033]

[0034] 수학적 식 1에서  $F^A, F^B$ 는 특징 맵으로, 자기 상관부(211, 221)에서는 특징맵( $F^s, F^t$ ) 각각에 대한 자기 상관을 수행하므로,  $F^A = F^B = F^s$  또는  $F^A = F^B = F^t$  이고, j, l는 로컬 탐색창( $N_i$ )에 속하는 특징맵( $F^A, F^B$ )의 픽셀(j, l  $\in N_i$ )을 의미한다. 그리고  $\langle \rangle$  및  $\langle \rangle^2$ 는 각각  $L_1, L_2$ -norm 함수이다.

[0035] 이때, 자기 상관 맵( $C^{ss}, C^{tt}$ )은 특징맵( $F^s, F^t$ ) 각각에서 각 픽셀(i)에 대한 자기 상관 스코어로 구성된다.

[0036] 도 3에서 (a)는 입력 영상( $X^s$ )의 일예를 나타내고, (b)는 입력 영상( $X^s$ )에 대한 자기 상관 맵( $C^{ss}$ )의 분산(variance)을 시각화한 결과를 나타내며, (c)는 입력 영상( $X^s$ )에서 임의의 두 픽셀( $i^A, i^B$ )를 나타낸다. 그리고 (d) 및 (e)는 각각 임의의 픽셀( $i^A, i^B$ )에 대한 자기 상관 맵( $C^{ss}$ )을 나타낸다.

[0037] 도 3에 나타난 바와 같이, 자기 상관 맵( $C^{ss}$ )은 픽셀( $i^B$ )에서와 같이 주변과 유사한 영역에 비해, 픽셀( $i^A$ )에서와 같이 주변에 비해 두드러진 영역에서 큰 분산값을 가지므로, 랜드마크 검출 시에 랜드마크로 검출되어야 할 중요 영역에 대한 정보를 제공할 수 있다.

[0038] 스코어 맵 획득부(212, 222)는 각각 다수의 특징맵( $F^s, F^t$ )과 다수의 자기 상관 맵( $C^{ss}, C^{tt}$ ) 중 대응하는 특징맵과 자기 상관 맵을 인가받고, 인가된 특징맵과 자기 상관 맵을 결합(concatenate)하고, 미리 학습된 패턴 추정 방식에 따라 결합된 특징맵과 자기 상관 맵에 랜드마크 가중치( $W_0$ )를 적용하여, 다수의 랜드마크 스코어 맵( $\hat{\psi} = F(F \parallel C; W_0) \in \mathbb{R}^{H \times W \times (K+1)}$ )을 추정한다. 여기서  $\parallel$ 는 특징맵(F)과 자기 상관 맵(C)을 결합하기 위한 연산자를 나타낸다.

[0039] 이때, 스코어 맵 획득부(212, 222) 또한 동일한 구조에 동일하게 학습되어 동일한 랜드마크 가중치( $W_0$ )를 갖는 인공 신경망인 삼 신경망으로 구현될 수 있다.

[0040] 스코어 맵 획득부(212, 222)는 일예로 각 픽셀에 대한 자기 상관 스코어( $C^{ss} = C(F_i^s, F_j^s), C^{tt} = C(F_i^t, F_j^t)$ )로 구성된 다수의 자기 상관 맵( $C^{ss}, C^{tt}$ )에서 기지정된 K개의 랜드마크에 대한 분산 스코어 맵과 랜드마크가 아닌 배경에 대한 하나의 분산 스코어 맵( $\hat{\psi}^s, \hat{\psi}^t$ )을 추정할 수 있다. 즉 스코어 맵 획득부(212, 222)는 각각 K+1개

의 소스 분산 스코어 맵( $\hat{\psi}^s$ )과 타겟 분산 스코어 맵( $\hat{\psi}^t$ )을 추정할 수 있다.

[0041] 여기서 스코어 맵 획득부(212, 222)는 다수의 자기 상관 맵( $C^{ss}$ ,  $C^{tt}$ ) 중 분산값이 큰 자기 상관 맵( $C^{ss}$ ,  $C^{tt}$ )을 기반으로 K개의 분산 스코어 맵( $\hat{\psi}^s$ ,  $\hat{\psi}^t$ )을 추정할 수 있다.

[0042] 스코어 맵 획득부(212, 222)는 2개의 입력 영상( $X^s$ ,  $X^t$ )의 랜드마크를 공간 좌표로 직접 회귀(regress)시키지 않고, 분산 스코어 맵( $\hat{\psi}$ )의 형태로 추정되도록 하여, 시멘틱 정합에 이용되는 서로 다른 2개의 입력 영상( $X^s$ ,  $X^t$ )에서 대응하는 서로 대응하는 랜드마크 검출을 학습할 수 있도록 한다.

[0043] 정규화부(213, 223)는 스코어 맵 획득부(212, 222)에서 추정된 K+1개의 소스 분산 스코어 맵( $\hat{\psi}^s$ )과 타겟 분산 스코어 맵( $\hat{\psi}^t$ ) 중에서 K개의 랜드마크에 대한 K개의 소스 분산 스코어 맵( $\hat{\psi}^s$ )과 K개의 타겟 분산 스코어 맵( $\hat{\psi}^t$ ) 각각을 수학식 2와 같이 정규화하여 랜드마크 확률 맵( $\psi^s$ ,  $\psi^t$ )을 획득한다.

## 수학식 2

$$\psi_i^k = \exp(\hat{\psi}_i^k) / \sum_{m=0}^K \exp(\hat{\psi}_i^m)$$

[0044]

[0045] 수학식 2에서  $\psi^k$ 는 k번째 랜드마크에 대한 랜드마크 확률 맵을 나타낸다.

[0046] 그리고 기대값 획득부(214, 224)는 k번째 랜드마크의 공간 좌표는 랜드마크 확률 맵( $\psi^k$ )에 따른 공간 좌표(i) 각각에 대해 수학식 3에 따라 계산되는 기대값( $\phi^k$ )으로 구성되는 기대값 맵을 획득할 수 있다.

## 수학식 3

$$\phi^k = \sum_i i \cdot \psi_i^k / \sum_i \psi_i^k$$

[0047]

[0048] 한편 시멘틱 정합부(300)는 공동 특징맵 획득부(100)에서 획득된 다수의 특징맵( $F^s$ ,  $F^t$ )을 인가받고, 인가된 특징맵( $F^s$ ,  $F^t$ )에 대해 상호 상관(cross correlation)을 수행하여 상호 유사성 스코어로 구성되는 상호 상관 맵( $C^{st}$ )을 획득하고, 상호 상관 맵( $C^{st}$ )으로부터 다수의 특징맵( $F^s$ ,  $F^t$ )에서 대응하는 픽셀로의 로컬 변환을 수행할 수 있도록 변환 필드 맵( $\tau$ )을 획득한다. 이와 함께 시멘틱 정합부(300)는 상호 상관 맵( $C^{st}$ )으로부터 변환 필드 맵( $\tau$ )에 의한 로컬 변환의 신뢰도를 검증하기 위한 신뢰도 맵( $\sigma$ )을 생성한다.

[0049] 도 4를 참조하면, 시멘틱 정합부(300)는 상호 상관부(310)와 신뢰도 맵 획득부(320) 및 변환 필드 획득부(330)를 포함한다. 상기한 바와 같이, 시멘틱 정합부(300)는 학습이 완료된 이후, 시멘틱 정합 장치로 이용되는 경우에, 공동 특징맵 획득부(100)를 더 포함한다.

[0050] 상호 상관부(310)는 C 차원의 특징맵( $F^s$ ,  $F^t$ )을 인가받고, 수학식 1의 상관 함수를 기반으로 기지정된 크기의 로컬 탐색창( $N_i$ )을 이용한 상호 상관을 수행하여 상호 상관 스코어로 구성되는 상호 상관 맵( $C^{st}$ )을 획득한다. 상호 상관부(310)는 특징맵( $F^s$ ,  $F^t$ )에 대한 상호 상관을 수행하므로, 수학식 1에서  $F^A = F^s$  이고,  $F^B = F^t$  이다.

[0051] 변환 필드 획득부(330)는 상호 상관 맵( $C^{st}$ )을 인가받고, 미리 학습된 패턴 추정 방식에 따라 상호 상관 맵( $C^{st}$ )에 대해 변환 가중치( $W_A$ )를 적용하여 특징맵( $F^s$ ,  $F^t$ )의 각 픽셀(i)에서의 상호 대응점의 위치 변화를 나타내는

변환값으로 구성되는 변환 필드 맵( $\tau = F(C^{st}; W_A) \in \mathbb{R}^{H \times W \times 2}$ )를 획득한다. 여기서 변환 필드 맵( $\tau$ )은 소스 영상( $X^s$ )에서 타겟 영상( $X^t$ )으로의 변환 및 타겟 영상( $X^t$ )에서 소스 영상( $X^s$ )으로의 변환이 모두 포함되도록 2차원으로 획득될 수 있다.

[0052] 한편, 신뢰도 맵 획득부(320)는 상호 상관 맵( $C^{st}$ )을 인가받고, 확률론적 모델을 기반으로 미리 학습된 패턴 추정 방식에 따라 상호 상관 맵( $C^{st}$ )에 대해 신뢰도 가중치( $W_C$ )를 적용하여 각 픽셀( $i$ )에서의 변환 필드 맵( $\tau$ )의 신뢰도로 구성되는 신뢰도 맵( $\sigma = F(C^{st}; W_C) \in \mathbb{R}^{H \times W \times 1}$ )을 획득한다.

[0053] 본 실시예에서 공동 학습부(400)는 객체 랜드마크 검출부(200)의 랜드마크 손실( $L_{det}$ )과 시멘틱 정합부(300)의 정합 손실( $L_{align}$ ) 및 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)의 공동 학습을 통한 공동 손실( $L_{joint}$ )을 측정하고, 측정된 랜드마크 손실( $L_{det}$ )과 정합 손실( $L_{align}$ ) 및 공동 손실( $L_{joint}$ )의 합을 총 손실( $L_{total}$ )로 획득하여 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)로 역전파하여 가중치( $W_D$ ,  $W_C$ ,  $W_A$ )를 업데이트 함으로써, 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)를 학습시킨다.

[0054] 랜드마크 손실 측정부(410)와 정합 손실 측정부(420), 공동 손실 측정부(430) 및 총 손실 역전파부(440)를 포함할 수 있다.

[0055] 랜드마크 손실 측정부(410)는 랜드마크의 기대값( $\phi^k$ )이 다른 영역과 차별화되는 로컬 영역에 집중되어야 하고, 다수의 랜드마크는 서로 이산되어야 한다는 두 가지 전제에 따라 랜드마크 손실( $L_{det}$ )을 획득한다.

[0056] 랜드마크는 입력 영상( $X$ )에서 특별한 영역을 의미하며, 이는 랜드마크에 기대값( $\phi^k$ )이 다른 영역 기대값보다 매우 집중되어 나타나야 함을 의미한다. 이에 랜드마크 손실 측정부(410)는 기대값( $\phi^k$ )과 이에 대응하는 분산 스코어 맵( $\hat{\psi}^k$ )으로부터 집중 손실(concentration loss)( $L_{con}$ )을 수학적 4에 따라 획득할 수 있다.

#### 수학적 4

$$\mathcal{L}_{con}(\psi) = \sum_k \left( \sum_i (i - \phi^k)^2 \cdot \hat{\psi}_i^k / \sum_i \hat{\psi}_i^k \right)$$

[0057]

[0058] 그리고 다수의 랜드마크가 한 지점에 집중되지 않도록 랜드마크 손실 측정부(410)는 서로 다른 2개의 랜드마크에 대한 기대값( $\phi^k$ ,  $\phi^{k'}$ ) 사이의 거리가 기준 거리값( $c$ )보다 멀어지도록 이산 손실(separation loss)( $L_{sep}$ )을 수학적 5에 따라 획득할 수 있다.

#### 수학적 5

$$\mathcal{L}_{sep}(\psi) = \sum_k \sum_{k' \neq k} \max(0, c - \|\phi^k - \phi^{k'}\|_2^2)$$

[0059]

[0060] 그리고 랜드마크 손실 측정부(410)는 집중 손실( $L_{con}$ )과 이산 손실( $L_{sep}$ ) 각각에 집중 손실 가중치( $\lambda_{con}$ )와 이산 손실 가중치( $\lambda_{sep}$ )를 가중하고 합하여 수학적 6과 같이 랜드마크 손실( $L_{det}$ )을 획득한다.

## 수학식 6

$$\mathcal{L}_{\text{det}}(\psi) = \lambda_{\text{con}} \mathcal{L}_{\text{con}}(\psi) + \lambda_{\text{sep}} \mathcal{L}_{\text{sep}}(\psi)$$

한편, 정합 손실 측정부(420)는 변환 필드 맵( $\tau$ )에 의해 와프(warped)된 소스 영상( $X^s$ )의 특징은 타겟 영상( $X^t$ )의 특징과 동일해야 한다는 전제에 기반하는 분류 작업에 따라 정합 손실( $\mathcal{L}_{\text{align}}$ )을 획득할 수 있다. 다만, 정합 손실 측정부(420)는 신뢰할 수 없는 변환 필드 맵( $\tau$ )으로부터 정합 손실( $\mathcal{L}_{\text{align}}$ )을 예측해야 하므로, 신뢰도 맵( $\sigma$ )을 더 반영한 교차 엔트로피 손실(cross entropy loss)을 수학식 7에 따라 계산하여 정합 손실( $\mathcal{L}_{\text{align}}$ )을 획득한다.

## 수학식 7

$$\mathcal{L}_{\text{align}}(\tau, \sigma) = \sum_i \left( - \sum_{j \in \mathcal{M}_i} \frac{s_j^*}{\sigma_i} \log(s_{i,j}(\tau)) + \log \sigma_i \right)$$

여기서  $s_{i,j}(\tau)$ 는 수학식 7로 정의되는 소프트맥스(softmax) 확률을 나타내고,  $s^*$ 은  $j = i$ 인 경우 1로 정의되고,  $j \neq i$ 이면,  $j \in \mathcal{M}_i$ 에 대해  $\mathcal{M}_i$  내의 중심 픽셀( $i$ )이 양의 샘플이 되고 나머지 픽셀이 음의 픽셀이 되도록 0으로 정의되는 클래스 레이블을 의미한다.

## 수학식 8

$$s_{i,j}(\tau) = \frac{\exp(\mathcal{C}(F_i^s, [\tau \circ F^t]_j))}{\sum_{l \in \mathcal{M}_i} \exp(\mathcal{C}(F_i^s, [\tau \circ F^t]_l))}$$

수학식 7에서는 교차 엔트로피 손실을 예측 불명확도를 나타내는 신뢰도 맵( $\sigma$ )으로 나눔으로써, 실현 불가능한 정합에 대한 정합을 회피하도록 하여 비지도 학습으로 인해 오류가 시멘틱 정합 오류가 발생하는 것을 방지하도록 한다.

공동 손실 측정부(430)는 시멘틱 정합된 소스 영상( $X^s$ )과 타겟 영상( $X^t$ )의 랜드마크는 동일한 위치이어야 한다는 전제에 따라 공동 손실( $\mathcal{L}_{\text{joint}}$ )을 획득한다.

객체 랜드마크 검출부(200)와 시멘틱 정합부(300)가 정상적으로 학습되었다면, 객체 랜드마크 검출부(200)에서 검출되는 타겟 영상( $X^t$ )의 랜드마크를 시멘틱 정합부(300)에서 획득된 변환 필드 맵( $\tau$ )에 따라 와프하면, 와프된 타겟 영상( $X^t$ )의 랜드마크의 위치는 소스 영상( $X^s$ )의 랜드마크의 위치와 동일해야 한다.

따라서 공동 손실 측정부(430)는 타겟 영상( $X^t$ )의 랜드마크의 위치와 소스 영상( $X^s$ )의 랜드마크의 위치 사이의 거리차인 거리 손실( $\mathcal{L}_{\text{dis}}$ )을 수학식 9에 따라 획득한다.

# 수학식 9

$$L_{\text{dis}}(\psi, \tau) = \sum_k \sum_i \|\psi_i^k(\mathbf{x}^s) - \tau \circ \psi_i^k(\mathbf{x}^t)\|_2^2$$

[0070]

[0071]

[0072]

[0073]

[0074]

[0075]

[0076]

[0077]

[0078]

[0079]

[0080]

[0081]

[0082]

그리고 획득된 거리 손실( $L_{\text{dis}}$ )에 거리 손실 가중치( $\lambda_{\text{dis}}$ )를 가중하여 공동 손실( $L_{\text{joint}} = \lambda_{\text{dis}} L_{\text{dis}}$ )을 획득한다.

총 손실 역전과부(440)는 랜드마크 손실( $L_{\text{det}}$ )과 정합 손실( $L_{\text{align}}$ ) 및 공동 손실( $L_{\text{joint}}$ )을 모두 가산하여 총 손실( $L_{\text{total}} = L_{\text{det}} + L_{\text{align}} + L_{\text{joint}}$ )을 획득하고, 획득된 총 손실을 객체 랜드마크 검출부(200) 및 시멘틱 정합부(300)로 역전과하여 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)를 학습시킨다.

여기서 공동 학습부(400)는 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)를 교대로 학습시킴으로써, 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)가 상호 보완하며 학습되도록 할 수 있다.

도 6은 본 발명의 일 실시예에 따른 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 방법을 나타낸다.

본 실시예에서 공동 특징맵 획득부(100)의 다수의 특징맵 획득부(110, 120)가 입력 영상( $X^s, X^t$ )으로부터 특징맵( $F^s, F^t$ )을 획득하기 위한 학습 기법은 공지된 기술이므로, 다수의 특징맵 획득부(110, 120)는 기존의 패턴 추정 방식에 따라 미리 학습되어 다수의 특징맵( $F^s, F^t$ )을 획득할 수 있다. 따라서 여기서는 다수의 특징맵 획득부(110, 120)의 학습 방식은 별도로 설명하지 않는다.

도 1 내지 도 5를 참조하여, 본 실시예에 따른 객체 랜드마크 검출 장치 및 시멘틱 정합 장치의 공동 학습 방법을 설명하면, 공동 학습부(400)는 객체 랜드마크 검출부(200)의 스코어 맵 획득부(212, 222)의 랜드마크 가중치( $W_D$ )가 업데이트 되지 않도록 고정한다(S11). 그리고 공동 학습부(400)는 입력 영상( $X^s, X^t$ )에 대한 객체 랜드마크 검출부(200)의 랜드마크 손실( $L_{\text{det}}$ )과 시멘틱 정합부(300)의 정합 손실( $L_{\text{align}}$ ) 및 객체 랜드마크 검출부(200)와 시멘틱 정합부(300)의 공동 손실( $L_{\text{joint}}$ )을 가산하여 총 손실( $L_{\text{total}} = L_{\text{det}} + L_{\text{align}} + L_{\text{joint}}$ )을 획득하고, 총 손실( $L_{\text{total}}$ )을 시멘틱 정합부(300)로 역전과하여 신뢰도 맵 획득부(320)의 신뢰도 가중치( $W_C$ )와 변환 필드 획득부(330)의 변환 가중치( $W_A$ )는 업데이트함으로써 시멘틱 정합부(300)를 학습시킨다(S12).

즉 공동 학습부(400)는 총 손실( $L_{\text{total}}$ )을 시멘틱 정합부(300)로 역전과하여 시멘틱 정합부(300)의 신뢰도 맵 획득부(320)의 신뢰도 가중치( $W_C$ )와 변환 필드 획득부(330)의 변환 가중치( $W_A$ )는 업데이트 하는 반면, 스코어 맵 획득부(212, 222)의 랜드마크 가중치( $W_D$ )는 고정되도록 객체 랜드마크 검출부(200)로 역전과하지 않는다.

스코어 맵 획득부(212, 222)의 랜드마크 가중치( $W_D$ )가 고정됨에 따라 랜드마크 손실( $L_{\text{det}}$ )은 일정하게 유지되고, 시멘틱 정합부(300)만이 총 손실( $L_{\text{total}}$ )에 기반하여 우선 학습된다.

이때 공동 학습부(400)는 시멘틱 정합부(300)를 학습시키는 경우에도, 신뢰도 맵 획득부(320)의 신뢰도 가중치( $W_C$ )와 변환 필드 획득부(330)의 변환 가중치( $W_A$ )를 교대로 고정하여 신뢰도 맵 획득부(320)와 변환 필드 획득부(330)를 교대로 학습시킨다. 즉 총 손실( $L_{\text{total}}$ )을 신뢰도 맵 획득부(320)로 역전과하여 학습시킨 후, 다시 획득되는 총 손실( $L_{\text{total}}$ )을 변환 필드 획득부(330)로 역전과하여 학습시키는 과정을 반복한다.

공동 학습부(400)는 시멘틱 정합부(300)를 미리 지정된 횟수로 반복 학습시키거나, 총 손실( $L_{\text{total}}$ )의 변화가 기 지정된 기준 손실 변화량 이하이면, 신뢰도 맵 획득부(320)의 신뢰도 가중치( $W_C$ )와 변환 필드 획득부(330)의 변환 가중치( $W_A$ )를 고정한다(S13).

그리고 객체 랜드마크 검출부(200)를 학습시킨다(S14).

신뢰도 맵 획득부(320)의 신뢰도 가중치( $W_C$ )와 변환 필드 획득부(330)의 변환 가중치( $W_A$ )가 고정됨에 따라 정합

손실( $L_{align}$ )은 일정하게 유지되고, 객체 랜드마크 검출부(200)만이 총 손실( $L_{total}$ )에 기반하여 학습된다.

[0083] 공동 학습부(400)는 객체 랜드마크 검출부(200)를 미리 지정된 횟수로 반복 학습시키거나, 총 손실( $L_{total}$ )의 변화가 기지정된 기준 손실 변화량 이하이면, 총 손실( $L_{total}$ )이 기지정된 기준 손실 이상인지 판별한다(S15).

[0084] 만일 총 손실( $L_{total}$ )이 기준 손실 미만이면, 공동 학습부(400)는 객체 랜드마크 검출부(200)와 신뢰도 맵 획득부(320)에 대한 학습을 종료한다(S16). 반면, 총 손실( $L_{total}$ )이 기준 손실 이상이면, 다시 시멘틱 정합부(300)와 객체 랜드마크 검출부(200)를 교대로 학습시킨다(S11 ~ S14).

[0085] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 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.

[0086] 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.

[0087] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

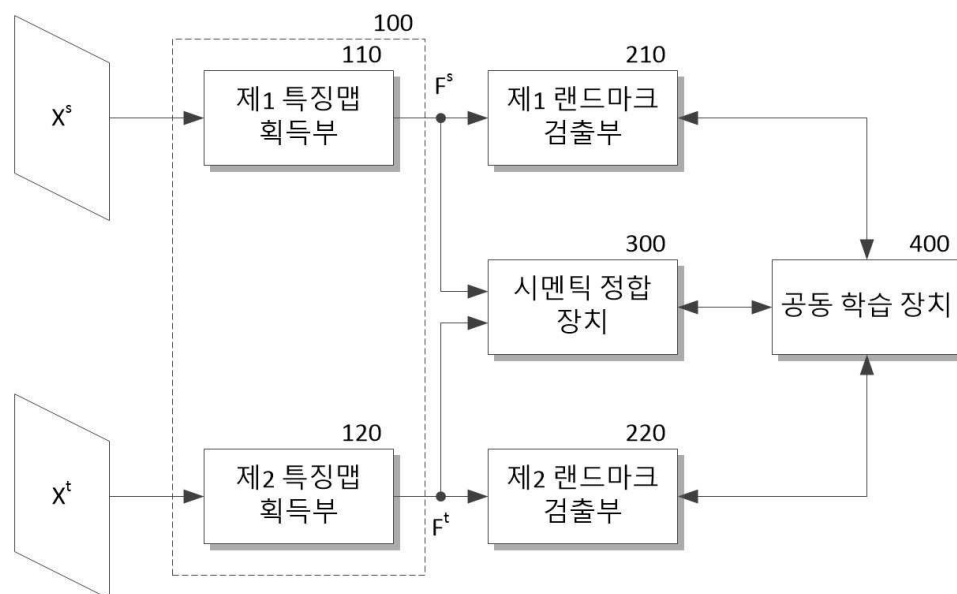
## 부호의 설명

[0088] 100: 공동 특징맵 획득부 200: 객체 랜드마크 검출 장치

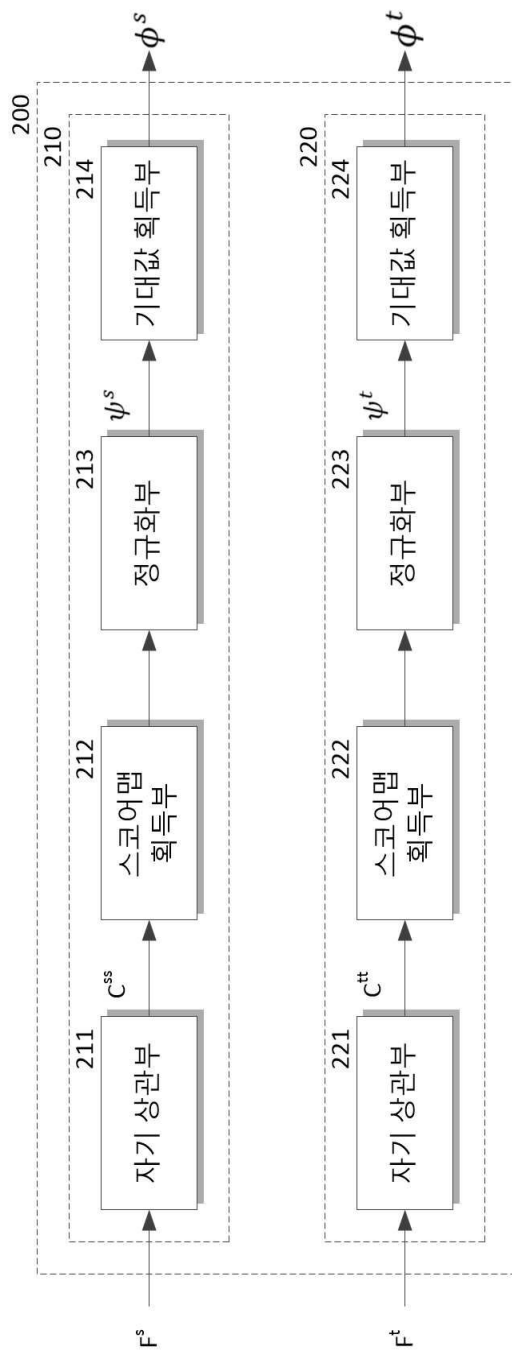
300: 시멘틱 정합 장치 400: 공동 학습 장치

## 도면

### 도면1

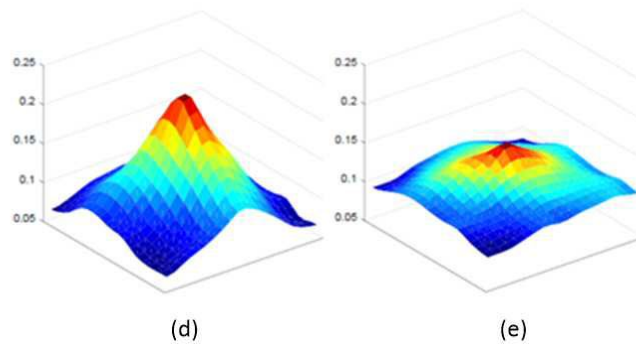
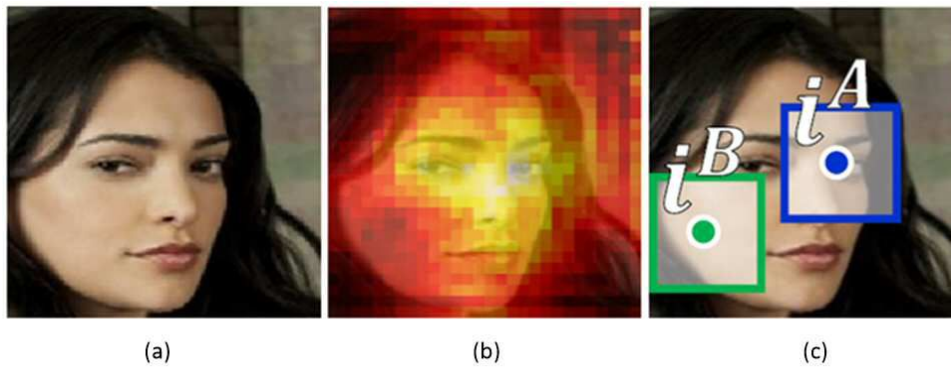


도면2

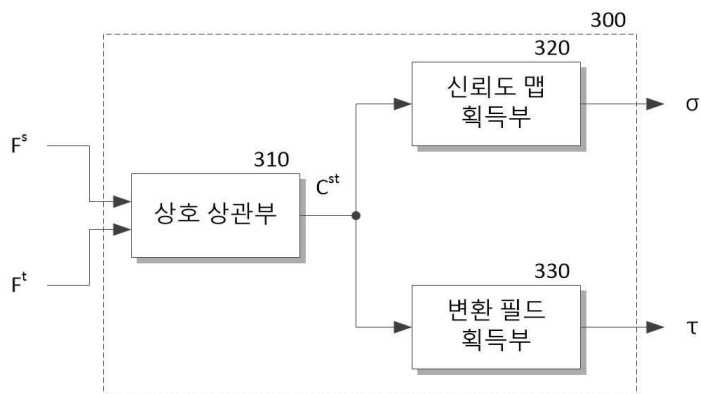




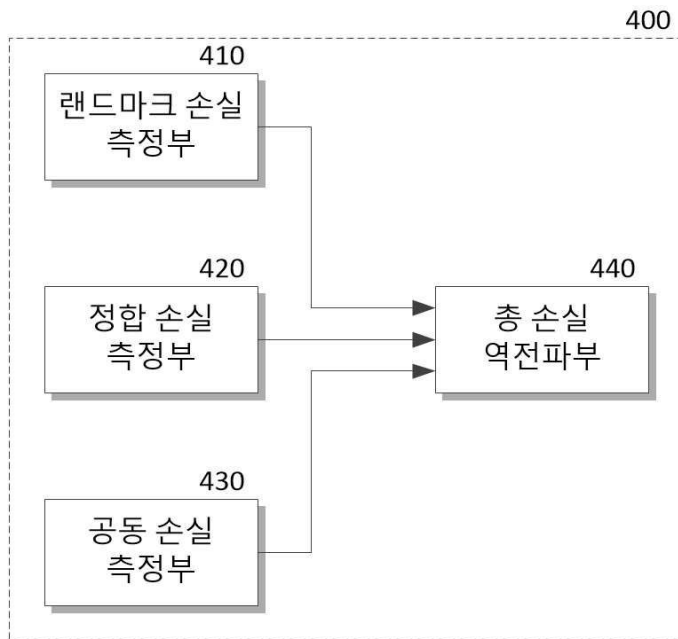
도면3



도면4



도면5



도면6

