



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0124849  
(43) 공개일자 2022년09월14일

(51) 국제특허분류(Int. Cl.)

G06N 20/00 (2019.01)

(52) CPC특허분류

G06N 20/00 (2021.08)

(21) 출원번호 10-2021-0028456

(22) 출원일자 2021년03월04일

심사청구일자 2021년03월04일

(71) 출원인

계명대학교 산학협력단

대구광역시 달서구 달구벌대로 1095, 계명대학교 산학협력관 201호(신당동)

(72) 발명자

고병철

대구광역시 수성구 범어천로 107 범어 STX KAN, 101동 1002호

김상원

대구광역시 달서구 달구벌대로 1095 계명대학교

(74) 대리인

김건우

전체 청구항 수 : 총 16 항

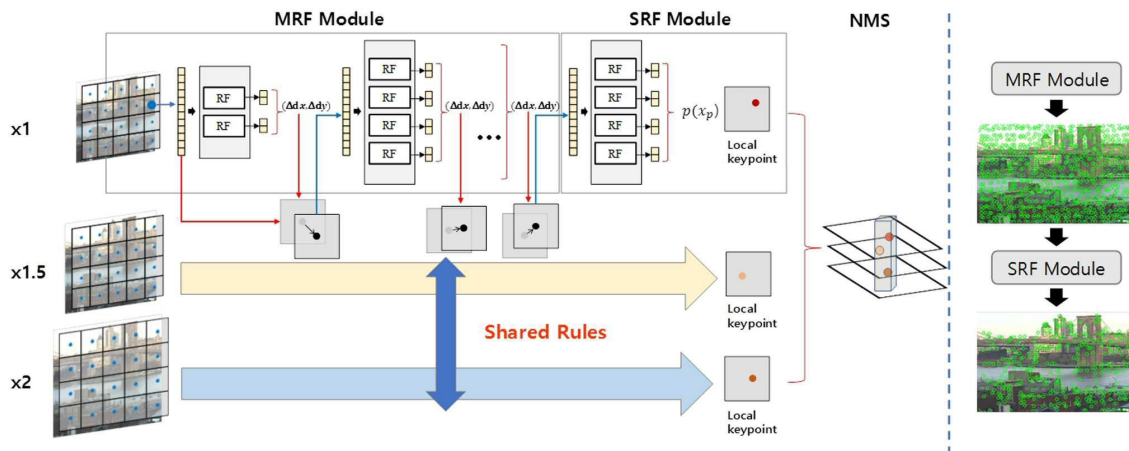
(54) 발명의 명칭 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치

(57) 요약

본 발명은 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법에 관한 것으로서, 보다 구체적으로는 컴퓨터에 의해 각 단계가 수행되는 자기-지도 키포인트 검출 방법으로서, (1) 키포인트 기술자(descriptor) 및 호포그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에

(뒷면에 계속)

대표도 - 도2



의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 단계; (2) 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 상기 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈을 생성하는 단계; (3) 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, 상기 MRF 모듈에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈을 생성하는 단계; 및 (4) 상기 생성한 MRF 모듈과 SRF 모듈을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 상기 MRF 모듈에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

또한, 본 발명은 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치에 관한 것으로서, 보다 구체적으로는 컴퓨터로 구현되는 자기-지도 키포인트 검출 장치로서, 키포인트 기술자(descriptor) 및 호포그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 라벨링 모듈; 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 상기 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈; 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, 상기 MRF 모듈에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈; 및 상기 MRF 모듈과 SRF 모듈을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 상기 MRF 모듈에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 검출 모듈을 포함하는 것을 그 구성상의 특징으로 한다.

본 발명에서 제안하고 있는 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치에 따르면, 다층 랜덤 포레스트 회귀 분석기와 단일 랜덤 포레스트 분류기로 키포인트를 검출하도록 구성함으로써, 다층 랜덤 포레스트 구조를 키포인트의 회귀에 적용하여 깊은 구조를 유지하면서 역전파 없이 반복성 있는 키포인트의 위치를 검출할 수 있고, 원본 이미지와 뒤틀린 이미지 사이의 자기-지도 학습 방법을 이용함으로써 비-라벨 데이터로 학습을 할 수 있으며, 단일 랜덤 포레스트 분류기를 통해 신뢰도가 낮은 키포인트를 제거해 매칭의 정확도를 높일 수 있다.

#### 이 발명을 지원한 국가연구개발사업

|             |                                    |
|-------------|------------------------------------|
| 과제고유번호      | 1345313705                         |
| 과제번호        | 2019R1I1A3A01042506                |
| 부처명         | 교육부                                |
| 과제관리(전문)기관명 | 한국연구재단                             |
| 연구사업명       | 지역대학우수과학자지원사업(1년~5년)               |
| 연구과제명       | 역전파가 필요 없는 심층 랜덤 네트워크 모델 개발과 응용 연구 |
| 기 여 율       | 1/1                                |
| 과제수행기관명     | 계명대학교                              |
| 연구기간        | 2020.03.01 ~ 2021.02.28            |

## 명세서

### 청구범위

#### 청구항 1

컴퓨터에 의해 각 단계가 수행되는 자기-지도 키포인트 검출 방법으로서,

- (1) 키포인트 기술자(descriptor) 및 호포그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 단계;
- (2) 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 상기 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈(200)을 생성하는 단계;
- (3) 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, 상기 MRF 모듈(200)에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈(300)을 생성하는 단계; 및
- (4) 상기 생성한 MRF 모듈(200)과 SRF 모듈(300)을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 상기 MRF 모듈(200)에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 단계를 포함하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

#### 청구항 2

제1항에 있어서, 상기 단계 (1)은,

- (1-1) 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성하는 단계;
- (1-2) 상기 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성하는 단계;
- (1-3) 상기 생성한 키포인트 및 키포인트 기술자를 이용해, 상기 뒤틀린 이미지를 상기 학습 이미지에 매핑하는 단계;
- (1-4) 상기 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키포인트를 찾아, 키포인트-쌍 대응을 생성하는 단계; 및
- (1-5) 상기 학습 이미지에 서로 다른 호모그래피를 적용해 상기 단계 (1-1) 내지 단계 (1-4)를 수행하며, brute force matcher를 통해 자기-키포인트 라벨링을 하여 상기 의사 실측 키포인트를 생성하는 단계를 포함하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

#### 청구항 3

제1항에 있어서, 상기 MRF 회귀 분석기는,

coarse-to-fine 파이프라인 구조로 순차적 회귀를 적용하도록 복수의 레이어로 구성되며, 레이어 당 RF 노드 수는 상기 단계 (2)의 학습 과정에서 자동으로 결정되되, 2개 이상 8개 이하인 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

#### 청구항 4

제1항에 있어서, 상기 단계 (2)에서는,

N-겹 교차 검증을 이용해 상기 MRF 모듈(200)을 생성하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

## 청구항 5

제4항에 있어서, 상기 단계 (2)에서는,

- (2-1) 전체 학습 데이터에서 N-1 폴드를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성하는 단계;
- (2-2) 상기 단계 (2-1)에서 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 상기 단계 (1)에서 자기-키포인트 라벨링에 의해 생성된 의사 실측 키포인트를 상기 학습 이미지에 매핑하는 단계;
- (2-3) 상기 단계 (2-2)의 매핑에 의해 상기 의사 실측 키포인트를 포함하는 격자의 중심점(초기 키포인트)을 기초로 상기 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력하는 단계;
- (2-4) 상기 초기 키포인트를 상기 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 상기 MRF 회귀 분석기의 제2 레이어를 위한 초기 키포인트를 생성하고, 상기 제2 레이어를 학습하여 예측 변위를 출력하는 단계;
- (2-5) 상기 MRF 회귀 분석기의 L개의 레이어에 대해 상기 단계 (2-4)를 진행하는 단계;
- (2-6) 상기 단계 (2-1)에서 선택한 N-1개의 폴드의 학습 후에, 상기 검증 세트의 샘플을 이용해 손실을 계산하는 단계;
- (2-7) 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장하는 단계;
- (2-8) 상기 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 상기 단계 (2-1) 내지 단계 (2-7)을 반복 수행하는 단계; 및
- (2-9) 상기 단계 (2-8)에서 저장된 MRF 회귀 분석기의 구조 중에서 총 손실이 가장 작은 구조를 상기 MRF 모듈 (200)로 결정하는 단계를 포함하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

## 청구항 6

제5항에 있어서, 상기 단계 (2-3)에서는,

상기 단계 (2-2)의 매핑에 의해 키포인트를 포함하는 격자를 미리 정해진 서브 블록으로 분할하고, 상기 분할된 각각의 서브 블록에서 추출한 특징 기술자(descriptors)를 결합해 특징 벡터를 구성하며, 상기 특징 벡터와 상기 초기 키포인트부터 상기 의사 실측 키포인트까지의 변위를 상기 제1 레이어의 입력으로 하여 학습을 수행하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

## 청구항 7

제1항에 있어서, 상기 단계 (3) 이후에는,

상기 MRF 모듈(200) 및 SRF 모듈(300)은 트리로 구성된 규칙을 공유하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

## 청구항 8

제1항에 있어서, 상기 단계 (4)에서는,

상기 입력 이미지를 서로 다른 스케일로 변환한 복수의 스케일 이미지를 상기 MRF 모듈(200) 및 SRF 모듈(300)에 각각 적용하며, NMS를 사용해 대응 창을 이동하면서 신뢰도 점수를 추정하되, 신뢰도 점수가 가장 높으면서 최소 임계값을 초과하는 키포인트를 최종 키포인트로 선정하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법.

## 청구항 9

컴퓨터로 구현되는 자기-지도 키포인트 검출 장치(10)로서,

키포인트 기술자(descriptor) 및 호포그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 라벨링 모듈(100);

다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 상기 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈(200);

단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, 상기 MRF 모듈(200)에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈(300); 및

상기 MRF 모듈(200)과 SRF 모듈(300)을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 상기 MRF 모듈(200)에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 검출 모듈(400)을 포함하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

## 청구항 10

제9항에 있어서, 상기 라벨링 모듈(100)은,

(1-1) 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성하는 단계;

(1-2) 상기 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성하는 단계;

(1-3) 상기 생성한 키포인트 및 키포인트 기술자를 이용해, 상기 뒤틀린 이미지를 상기 학습 이미지에 매핑하는 단계;

(1-4) 상기 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키포인트를 찾아, 키포인트-쌍 대응을 생성하는 단계; 및

(1-5) 상기 학습 이미지에 서로 다른 호모그래피를 적용해 상기 단계 (1-1) 내지 단계 (1-4)를 수행하며, brute force matcher를 통해 자기-키포인트 라벨링을 하여 상기 의사 실측 키포인트를 생성하는 단계를 수행하여,

레이블이 없는 학습 이미지에 의사 실측 키포인트를 생성해 학습 데이터를 구성하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

## 청구항 11

제9항에 있어서, 상기 MRF 회귀 분석기는,

coarse-to-fine 파이프라인 구조로 순차적 회귀를 적용하도록 복수의 레이어로 구성되며, 레이어 당 RF 노드 수는 상기 MRF 회귀 분석기의 학습 과정에서 자동으로 결정되되, 2개 이상 8개 이하인 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

## 청구항 12

제9항에 있어서, 상기 MRF 모듈(200)은,

N-겹 교차 검증을 이용해 생성되는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

### 청구항 13

제12항에 있어서, 상기 MRF 모듈(200)은,

- (2-1) 전체 학습 데이터에서 N-1 폴드를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성하는 단계;
- (2-2) 상기 단계 (2-1)에서 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 상기 라벨링 모듈(100)에서 자기-키포인트 라벨링에 의해 생성된 의사 실측 키포인트를 상기 학습 이미지에 매핑하는 단계;
- (2-3) 상기 단계 (2-2)의 매핑에 의해 상기 의사 실측 키포인트를 포함하는 격자의 중심점(초기 키포인트)을 기초로 상기 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력하는 단계;
- (2-4) 상기 초기 키포인트를 상기 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 상기 MRF 회귀 분석기의 제2 레이어를 위한 초기 키포인트를 생성하고, 상기 제2 레이어를 학습하여 예측 변위를 출력하는 단계;
- (2-5) 상기 MRF 회귀 분석기의 L개의 레이어에 대해 상기 단계 (2-4)를 진행하는 단계;
- (2-6) 상기 단계 (2-1)에서 선택한 N-1개의 폴드의 학습 후에, 상기 검증 세트의 샘플을 이용해 손실을 계산하는 단계;
- (2-7) 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장하는 단계;
- (2-8) 상기 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 상기 단계 (2-1) 내지 단계 (2-7)을 반복 수행하는 단계; 및
- (2-9) 상기 단계 (2-8)에서 저장된 MRF 회귀 분석기의 구조 중에서 총 손실이 가장 작은 구조를 상기 MRF 모듈(200)로 결정하는 단계를 통해 생성되는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

### 청구항 14

제13항에 있어서, 상기 단계 (2-3)에서는,

상기 단계 (2-2)의 매핑에 의해 키포인트를 포함하는 격자를 미리 정해진 서브 블록으로 분할하고, 상기 분할된 각각의 서브 블록에서 추출한 특징 기술자(descriptors)를 결합해 특징 벡터를 구성하며, 상기 특징 벡터와 상기 초기 키포인트부터 상기 의사 실측 키포인트까지의 변위를 상기 제1 레이어의 입력으로 하여 학습을 수행하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

### 청구항 15

제9항에 있어서, 상기 MRF 모듈(200) 및 SRF 모듈(300)은,

트리로 구성된 규칙을 공유하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

### 청구항 16

제9항에 있어서, 상기 검출 모듈(400)은,

상기 입력 이미지를 서로 다른 스케일로 변환한 복수의 스케일 이미지를 상기 MRF 모듈(200) 및 SRF 모듈(300)에 각각 적용하며, NMS를 사용해 대응 창을 이동하면서 신뢰도 점수를 추정하되, 신뢰도 점수가 가장 높으면서 최소 임계값을 초과하는 키포인트를 최종 키포인트로 선정하는 것을 특징으로 하는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10).

### 발명의 설명

## 기술 분야

[0001] 본 발명은 자기-지도 키포인트 검출 방법 및 장치에 관한 것으로서, 보다 구체적으로는 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치에 관한 것이다.

## 배경 기술

[0002] 영상 정합(Image registration)은 여러 영상을 동일한 좌표계로 변환하는 프로세스이다. 정합에 사용되는 영상은 서로 다른 시간, 센서 또는 시점에서 촬영되기 때문에, 단순한 유사성 측정으로는 매칭이 불가능하다. 일반적으로 영상 정합은 키포인트 검출, 특징 기술자(feature descriptor), 특징 매칭(feature matching) 및 이미지 와핑(image warping)의 4가지 주요 단계를 통해 수행될 수 있다. 영상 정합에서 가장 기본적이고 중요한 단계는 영상의 특성을 반영하는 키포인트를 정확하게 추출하는 것이다.

[0004] 키포인트 추출은 SLAM(simultaneous localization and mapping), Sfm(structure-from-motion), 포즈 추정(pose estimation), 객체 검출(object detection) 및 3차원 재구성(3D reconstruction)과 같은 기하학적 컴퓨터 시각화 작업에도 매우 중요한 과정이다.

[0006] 좋은 키포인트는 영상의 인접한 픽셀과 구별되며 고유하게 식별할 수 있어야 한다. 또한, 반복성과 신뢰성을 만족시켜야 한다. 여기서, 반복성(Repeatability)은 기하학적 및 측광 변환에 의해 크게 변경된 이미지에서도 키포인트가 반복적으로 감지되는 속성을 말한다. 신뢰성(Reliability)이란 이미지 매칭을 위해 의미 있는 영역에서 키포인트를 선택해야 함을 의미한다. 예를 들어, 건물 이미지를 매칭해야 하는 경우 나무나 하늘에서 추출한 키포인트는 신뢰할 수 없는 지점이며 이미지 매칭을 위한 중요한 특징으로 간주할 수 없다.

[0008] 키포인트 추출과 관련된 주요 방법은 크게 손으로 만든 특징(handcrafted-feature)과 컨볼루션 신경망(CNN) 기반 접근 방식으로 나눌 수 있다. 특히, 지난 몇 년 동안 딥 러닝이 컴퓨터 비전 분야에서 최첨단 성능을 달성했기 때문에 이미지에 대한 사전 지식 없이 자동으로 키포인트를 검출할 수 있는 딥 네트워크에 대한 연구가 진행되고 있다. 사용 가능한 다양한 딥 네트워크 중에서 CNN은 수작업 특징 방법보다 우수하고 이미지 변환에서 강력한 성능을 달성하는 것으로 나타났다. CNN 기반의 키포인트 추출은 학습 방법에 따라 지도학습(Supervised learning), 자기-지도 학습(Self-supervised learning), 비지도 학습(Unsupervised learning)의 세 가지 유형으로 분류할 수 있다.

[0010] 먼저, 지도 학습 기반 키포인트 추출에는 사람 어노테이터(annotator)가 레이블을 지정한 데이터 세트가 필요하다. 지도 학습을 이용한 키포인트는 비교적 쉽고 정확한 방법으로 널리 연구되고 있으나, 키포인트가 어노테이터에 따라 잘못 정의될 수 있으므로 사람의 주석은 신뢰하기 어려운 한계가 있다. 따라서 최근에는 자기 학습 기반의 의사 실측 레이블을 이용한 자기-지도 학습 또는 준지도 학습, 실측 레이블 자체가 필요하지 않은 비지도 학습 등의 방법이 제안되고 있다.

[0012] 지도 학습과는 달리, 자기-지도 학습(또는 반지도 학습)은 광범위한 사람의 주석을 사용하지 않는다. 따라서 지도 학습에 내재한 실질적인 한계를 완화할 수 있다. 이 방법을 사용하면 기본적으로 원본 이미지와 뒤틀린 이미지(warped image)에 동일한 네트워크를 적용하고 얻은 키포인트를 각 이미지에 매핑하여 키포인트를 얻는다. 그런 다음 두 이미지의 매핑 포인트 사이에 유사한 포인트가 발견되면, K 포인트-쌍 대응(K point-pair correspondences)이 학습을 위해 제공된다.

[0014] 실제로 키포인트 감지 결과의 정확성을 평가하기 위해 실측 데이터(ground-truth data)를 생성하는 것은 여전히 어려운 작업이다. 자기-지도 학습에는 사전 정의된 실측(ground-truth)이 여전히 필요하므로, 이러한 한계를 극복하기 위해 비지도 학습 기반의 키포인트 검출이 제안되었다(Quad-networks, UnsuperPoint 등). 그러나 이

방법은 몇 가지 작업을 학습하기 위해 여전히 자기-지도 학습 프레임워크를 사용한다.

[0016] 자기-지도 또는 비지도 학습에는 특징 맵 추출을 위한 백본 네트워크와 특징 포인트의 위치, 점수 및 기술자(descriptor)를 추출하기 위한 추가 네트워크가 필요하며, 손실에 기반한 복잡한 학습 프로세스를 적용하여 손실 계산을 위해 3개의 출력과 호모그래피를 결합한다. 따라서 이러한 방법은 긴 테스트 및 학습 시간이 필요하고, 키포인트에 대한 반복성은 고려되지만 신뢰성은 고려되지 않는다는 문제가 있다.

[0018] 따라서 반복성과 신뢰성을 모두 고려하고, 적은 학습 데이터로도 학습할 수 있으며, 학습 시간이 짧은 키포인트 추출 방법의 개발이 필요한 실정이다.

[0020] 한편, 본 발명과 관련된 선행기술로서, 등록특허 제10-2188970호(발명의 명칭 경량 다층 랜덤 포레스트 기반의 얼굴 표정 인식 방법 및 장치, 등록일자: 2020년 12월 03일) 등이 개시된 바 있다.

## 발명의 내용

### 해결하려는 과제

[0021] 본 발명은 기존에 제안된 방법들의 상기와 같은 문제점들을 해결하기 위해 제안된 것으로서, 다층 랜덤 포레스트 회귀 분석기와 단일 랜덤 포레스트 분류기로 키포인트를 검출하도록 구성함으로써, 다층 랜덤 포레스트 구조를 키포인트의 회귀에 적용하여 깊은 구조를 유지하면서 역전과 없이 반복성 있는 키포인트의 위치를 검출할 수 있고, 원본 이미지와 뒤틀린 이미지 사이의 자기-지도 학습 방법을 이용함으로써 비-라벨 데이터로 학습을 할 수 있으며, 단일 랜덤 포레스트 분류기를 통해 신뢰도가 낮은 키포인트를 제거해 매칭의 정확도를 높일 수 있는, 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치를 제공하는 것을 그 목적으로 한다.

### 과제의 해결 수단

[0022] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법은,

[0023] 컴퓨터에 의해 각 단계가 수행되는 자기-지도 키포인트 검출 방법으로서,

[0024] (1) 키포인트 기술자(descriptor) 및 호모그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 단계;

[0025] (2) 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 상기 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈을 생성하는 단계;

[0026] (3) 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, 상기 MRF 모듈에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈을 생성하는 단계; 및

[0027] (4) 상기 생성한 MRF 모듈과 SRF 모듈을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 상기 MRF 모듈에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

[0029] 바람직하게는, 상기 단계 (1)은,

[0030] (1-1) 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성하는 단계;

[0031] (1-2) 상기 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성하는 단계;

[0032] (1-3) 상기 생성한 키포인트 및 키포인트 기술자를 이용해, 상기 뒤틀린 이미지를 상기 학습 이미지에 매핑하는 단계;

- [0033] (1-4) 상기 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키폰트를 찾아, 키폰트-쌍 대응을 생성하는 단계; 및
- [0034] (1-5) 상기 학습 이미지에 서로 다른 호모그래피를 적용해 상기 단계 (1-1) 내지 단계 (1-4)를 수행하며, brute force matcher를 통해 자기-키폰트 라벨링을 하여 상기 의사 실측 키폰트를 생성하는 단계를 포함할 수 있다.
- [0036] 바람직하게는, 상기 MRF 회귀 분석기는,
- [0037] coarse-to-fine 파이프라인 구조로 순차적 회귀를 적용하도록 복수의 레이어로 구성되며, 레이어 당 RF 노드는 상기 단계 (2)의 학습 과정에서 자동으로 결정되되, 2개 이상 8개 이하일 수 있다.
- [0039] 바람직하게는, 상기 단계 (2)에서는,
- [0040] N-겹 교차 검증을 이용해 상기 MRF 모듈을 생성할 수 있다.
- [0042] 더욱 바람직하게는, 상기 단계 (2)에서는,
- [0043] (2-1) 전체 학습 데이터에서 N-1 폴드를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성하는 단계;
- [0044] (2-2) 상기 단계 (2-1)에서 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 상기 단계 (1)에서 자기-키폰트 라벨링에 의해 생성된 의사 실측 키폰트를 상기 학습 이미지에 매핑하는 단계;
- [0045] (2-3) 상기 단계 (2-2)의 매핑에 의해 상기 의사 실측 키폰트를 포함하는 격자의 중심점(초기 키폰트)을 기초로 상기 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력하는 단계;
- [0046] (2-4) 상기 초기 키폰트를 상기 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 상기 MRF 회귀 분석기의 제2 레이어를 위한 초기 키폰트를 생성하고, 상기 제2 레이어를 학습하여 예측 변위를 출력하는 단계;
- [0047] (2-5) 상기 MRF 회귀 분석기의 L개의 레이어에 대해 상기 단계 (2-4)를 진행하는 단계;
- [0048] (2-6) 상기 단계 (2-1)에서 선택한 N-1개의 폴드의 학습 후에, 상기 검증 세트의 샘플을 이용해 손실을 계산하는 단계;
- [0049] (2-7) 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장하는 단계;
- [0050] (2-8) 상기 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 상기 단계 (2-1) 내지 단계 (2-7)을 반복 수행하는 단계; 및
- [0051] (2-9) 상기 단계 (2-8)에서 저장된 MRF 회귀 분석기의 구조 중에서 총 손실이 가장 작은 구조를 상기 MRF 모듈로 결정하는 단계를 포함할 수 있다.
- [0053] 더더욱 바람직하게는, 상기 단계 (2-3)에서는,
- [0054] 상기 단계 (2-2)의 매핑에 의해 키폰트를 포함하는 격자를 미리 정해진 서브 블록으로 분할하고, 상기 분할된 각각의 서브 블록에서 추출한 특징 기술자(descriptors)를 결합해 특징 벡터를 구성하며, 상기 특징 벡터와 상기 초기 키폰트부터 상기 의사 실측 키폰트까지의 변위를 상기 제1 레이어의 입력으로 하여 학습을 수행할 수 있다.
- [0056] 바람직하게는, 상기 단계 (3) 이후에는,
- [0057] 상기 MRF 모듈 및 SRF 모듈은 트리로 구성된 규칙을 공유할 수 있다.

- [0059] 바람직하게는, 상기 단계 (4)에서는,
- [0060] 상기 입력 이미지를 서로 다른 스케일로 변환한 복수의 스케일 이미지를 상기 MRF 모듈 및 SRF 모듈에 각각 적용하며, NMS를 사용해 대응 창을 이동하면서 신뢰도 점수를 추정하되, 신뢰도 점수가 가장 높으면서 최소 임계값을 초과하는 키포인트를 최종 키포인트로 선정할 수 있다.
- [0062] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치는,
- [0063] 컴퓨터로 구현되는 자기-지도 키포인트 검출 장치로서,
- [0064] 키포인트 기술자(descriptor) 및 호포그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 라벨링 모듈;
- [0065] 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 상기 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈;
- [0066] 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, 상기 MRF 모듈에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈; 및
- [0067] 상기 MRF 모듈과 SRF 모듈을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 상기 MRF 모듈에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 검출 모듈을 포함하는 것을 그 구성상의 특징으로 한다.
- [0069] 바람직하게는, 상기 라벨링 모듈은,
- [0070] (1-1) 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성하는 단계;
- [0071] (1-2) 상기 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성하는 단계;
- [0072] (1-3) 상기 생성한 키포인트 및 키포인트 기술자를 이용해, 상기 뒤틀린 이미지를 상기 학습 이미지에 매핑하는 단계;
- [0073] (1-4) 상기 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키포인트를 찾아, 키포인트-쌍 대응을 생성하는 단계; 및
- [0074] (1-5) 상기 학습 이미지에 서로 다른 호모그래피를 적용해 상기 단계 (1-1) 내지 단계 (1-4)를 수행하며, brute force matcher를 통해 자기-키포인트 라벨링을 하여 상기 의사 실측 키포인트를 생성하는 단계를 수행하여,
- [0075] 레이블이 없는 학습 이미지에 의사 실측 키포인트를 생성해 학습 데이터를 구성할 수 있다.
- [0077] 바람직하게는, 상기 MRF 회귀 분석기는,
- [0078] coarse-to-fine 파이프라인 구조로 순차적 회귀를 적용하도록 복수의 레이어로 구성되며, 레이어 당 RF 노드 수는 상기 MRF 회귀 분석기의 학습 과정에서 자동으로 결정되되, 2개 이상 8개 이하일 수 있다.
- [0080] 바람직하게는, 상기 MRF 모듈은, N-겹 교차 검증을 이용해 생성될 수 있다.
- [0082] 더욱 바람직하게는, 상기 MRF 모듈은,
- [0083] (2-1) 전체 학습 데이터에서 N-1 폴드를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성하는 단계;
- [0084] (2-2) 상기 단계 (2-1)에서 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 상기 라벨링 모듈에서 자기-키포인트 라벨링에 의해 생성된 의사 실측 키포인트를 상기 학습 이미지에 매핑하는 단계;

- [0085] (2-3) 상기 단계 (2-2)의 매핑에 의해 상기 의사 실측 키폰트를 포함하는 격자의 중심점(초기 키폰트)을 기초로 상기 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력하는 단계;
- [0086] (2-4) 상기 초기 키폰트를 상기 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 상기 MRF 회귀 분석기의 제2 레이어를 위한 초기 키폰트를 생성하고, 상기 제2 레이어를 학습하여 예측 변위를 출력하는 단계;
- [0087] (2-5) 상기 MRF 회귀 분석기의 L개의 레이어에 대해 상기 단계 (2-4)를 진행하는 단계;
- [0088] (2-6) 상기 단계 (2-1)에서 선택한 N-1개의 폴드의 학습 후에, 상기 검증 세트의 샘플을 이용해 손실을 계산하는 단계;
- [0089] (2-7) 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장하는 단계;
- [0090] (2-8) 상기 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 상기 단계 (2-1) 내지 단계 (2-7)을 반복 수행하는 단계; 및
- [0091] (2-9) 상기 단계 (2-8)에서 저장된 MRF 회귀 분석기의 구조 중에서 총 손실이 가장 작은 구조를 상기 MRF 모듈로 결정하는 단계를 통해 생성될 수 있다.
- [0093] 더더욱 바람직하게는, 상기 단계 (2-3)에서는,
- [0094] 상기 단계 (2-2)의 매핑에 의해 키폰트를 포함하는 격자를 미리 정해진 서브 블록으로 분할하고, 상기 분할된 각각의 서브 블록에서 추출한 특징 기술자(descriptors)를 결합해 특징 벡터를 구성하며, 상기 특징 벡터와 상기 초기 키폰트부터 상기 의사 실측 키폰트까지의 변위를 상기 제1 레이어의 입력으로 하여 학습을 수행할 수 있다.
- [0096] 바람직하게는, 상기 MRF 모듈 및 SRF 모듈은, 트리로 구성된 규칙을 공유할 수 있다.
- [0098] 바람직하게는, 상기 검출 모듈은,
- [0099] 상기 입력 이미지를 서로 다른 스케일로 변환한 복수의 스케일 이미지를 상기 MRF 모듈 및 SRF 모듈에 각각 적용하며, NMS를 사용해 대응 창을 이동하면서 신뢰도 점수를 추정하되, 신뢰도 점수가 가장 높으면서 최소 임계값을 초과하는 키폰트를 최종 키폰트로 선정할 수 있다.

### 발명의 효과

- [0100] 본 발명에서 제안하고 있는 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키폰트 검출 방법 및 장치에 따르면, 다층 랜덤 포레스트 회귀 분석기와 단일 랜덤 포레스트 분류기로 키폰트를 검출하도록 구성함으로써, 다층 랜덤 포레스트 구조를 키폰트의 회귀에 적용하여 깊은 구조를 유지하면서 역전과 없이 반복성 있는 키폰트의 위치를 검출할 수 있고, 원본 이미지와 뒤틀린 이미지 사이의 자기-지도 학습 방법을 이용함으로써 비-라벨 데이터로 학습을 할 수 있으며, 단일 랜덤 포레스트 분류기를 통해 신뢰도가 낮은 키폰트를 제거해 매칭의 정확도를 높일 수 있다.

### 도면의 간단한 설명

- [0101] 도 1은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키폰트 검출 장치의 구성을 도시한 도면.
- 도 2는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키폰트 검출 방법의 전체적인 구조를 도시한 도면.
- 도 3은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키폰트 검출 방법의 흐름을 도시한 도면.
- 도 4는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키폰트 검출 방법에서, 단계 S100의 세부적인 흐름을 도시한 도면.

도 5는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법에서, 단계 S200의 MRF 회귀 분석기 학습 과정을 도시한 도면.

도 6은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법에서, 단계 S200의 세부적인 흐름을 도시한 도면.

도 7 내지 도 9는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치를 다른 최첨단 방법과 비교한 결과를 도시한 도면.

도 10은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치와 다른 최첨단 방법의 키포인트 매핑 예제를 도시한 도면.

### 발명을 실시하기 위한 구체적인 내용

- [0102] 이하, 첨부된 도면을 참조하여 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 바람직한 실시예를 상세히 설명한다. 다만, 본 발명의 바람직한 실시예를 상세하게 설명함에 있어, 관련된 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략한다. 또한, 유사한 기능 및 작용을 하는 부분에 대해서는 도면 전체에 걸쳐 동일한 부호를 사용한다.
- [0104] 덧붙여, 명세서 전체에서, 어떤 부분이 다른 부분과 ‘연결’되어 있다고 할 때, 이는 ‘직접적으로 연결’되어 있는 경우뿐만 아니라, 그 중간에 다른 소자를 사이에 두고 ‘간접적으로 연결’되어 있는 경우도 포함한다. 또한, 어떤 구성요소를 ‘포함’한다는 것은, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다.
- [0106] 본 발명은 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치 및 방법에 관한 것으로서, 본 발명의 특징에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치 및 방법은, 메모리 및 프로세서를 포함한 하드웨어에서 기록되는 소프트웨어로 구성될 수 있다. 예를 들어, 본 발명의 본 발명은 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치 및 방법은, 개인용 컴퓨터, 노트북 컴퓨터, 서버 컴퓨터, PDA, 스마트폰, 태블릿 PC 등에 저장 및 구현될 수 있다.
- [0108] 도 1은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10)의 구성을 도시한 도면이다. 도 1에 도시된 바와 같이, 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10)는, 컴퓨터로 구현되는 자기-지도 키포인트 검출 장치(10)로서, 키포인트 기술자(descriptor) 및 호포그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성하는 라벨링 모듈(100); 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈(200); 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, MRF 모듈(200)에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈(300); 및 MRF 모듈(200)과 SRF 모듈(300)을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 MRF 모듈(200)에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 검출 모듈(400)을 포함하여 구성될 수 있다.
- [0110] 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 장치(10)는, 기존의 자기-지도 및 비지도 학습이 직면한 문제를 해결하기 위한 것으로서, CNN 대신에, MRF 모듈(200)과 SRF 모듈(300)을 기반으로 키포인트 회귀자(keypoint regressor, KeyReg)를 학습할 수 있다.
- [0112] RF(랜덤 포레스트, random forest)는 표 형식 데이터에 대해 매우 높은 분류 및 회귀 성능을 보여 주지만, 이미지 및 비디오와 같은 조건이 없는 데이터(unconditioned data)에 대해서는 성능이 좋지 않다. 따라서 본 발명

에서는, 이미지 대신 표 형식으로 특징 기술자를 세분화하고 RF 노드를 기반으로 한 레이어 별 심층 아키텍처를 구축하여, CPU에서 실행할 수 있는 좋은 키포인트 회귀 장치를 구성할 수 있다.

[0114] MRF 모듈(200)과 SRF 모듈(300), 두 개의 RF 기반 모듈로 구성된 KeyReg의 학습을 위해 레이블을 지정할 수 있다. 즉, 라벨링 모듈(100)은, 학습을 위해 복잡한 CNN 백본 네트워크 대신 SIFT 및 DAISY 수작업 특징(handcrafted features)과 호모그래피(homography)를 사용하여 의사 실측(pseudo-ground-truth) 레이블을 생성할 수 있다. 이미지에서 조밀하게 추출된 초기 키포인트는, 점수 및 위치에 대해 병렬 네트워크를 사용하지 않고 MRF 모듈(200)에 의해 coarse-to-fine 방식으로 예측될 수 있다. MRF 모듈(200)은 역전과 없이 N-겹 교차 검증(N-fold cross validation)을 사용해 생성되기 때문에, 적은 수의 데이터로도 학습할 수 있고 학습 시간이 짧다. 또한, 단일 RF 레이어로 구성된 SRF 모듈(300)은 NMS의 필수 요소인 키포인트의 신뢰도를 평가하기 위해 신뢰도 점수를 추정할 수 있다.

[0116] 도 2는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법의 전체적인 구조를 도시한 도면이다. 도 2에 도시된 바와 같이, 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법은, 반복 가능한 키포인트를 추출하기 위한 MRF 모듈(200)과 예측된 키포인트의 신뢰도 점수를 얻기 위한 SRF 모듈(300)로 구성될 수 있다. 여기서, MRF 모듈(200)은 DNN과 유사한 계층 간 구조이며 DNN 계층의 각 뉴런은 RF로 대체된다.

[0118] 보다 구체적으로, 본 발명에서는, 레이어 별(layer-by-layer) MRF 회귀 분석기 기반의 MRF 모듈(200)과 SRF 모듈(300)을 설계하여, 이미지의 뒤틀림과 관계없이 단일 순방향 패스로 키포인트를 검출할 수 있는 KeyReg를 구성하였다. 또한, KeyReg는 GLAMPoints와 유사한 자기-키포인트 라벨링 방법을 통해 학습 데이터를 생성하고 공유 모델을 통해 이미지의 크기 변화에 강한 키포인트를 추출하도록 설계되었다.

[0120] KeyReg의 MRF 모듈(200)의 구조는 분류보다는 회귀를 위해 설계되었으며, 레이어 당 RF 노드 수를 가변적으로 결정하여 coarse-to-fine 구조를 유지할 수 있다. 또한, 연속적인 SRF 모듈(300)을 통해 신뢰도가 낮은 키포인트를 제거하여 매칭의 정확도를 향상시킬 수 있다. KeyReg는 분류가 아닌 키포인트 회귀에 MRF를 적용하는 최초의 방법으로, 각 레이어의 RF 노드 수를 가변적으로 조정하여 회귀 정밀도를 향상시킬 수 있다. 기존 DNN 기반 방식과 비교하여 본 발명의 KeyReg는 GPU가 아닌 CPU에서 실행되도록 설계되었다.

[0122] 도 3은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법의 흐름을 도시한 도면이다. 도 3에 도시된 바와 같이, 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법은, 키포인트 기술자 및 호모그래피를 이용한 자기-키포인트 라벨링에 의해 학습 이미지에 의사 실측 키포인트를 생성하여 학습 데이터를 구성하는 단계(S100), 다층 랜덤 포레스트 회귀 분석기 기반으로 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈(200)을 생성하는 단계(S200), 다층 랜덤 포레스트 분류기 기반으로, MRF 모듈(200)에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈(300)을 생성하는 단계(S300) 및 MRF 모듈(200)과 SRF 모듈(300)을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS를 사용해 MRF 모듈(200)에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정하는 단계(S400)를 포함하여 구현될 수 있다.

[0124] 이하에서는, 도 2 및 도 3을 참조하여 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법의 각 단계에 대해 상세히 설명하도록 한다.

[0126] 단계 S100에서는, 키포인트 기술자(descriptor) 및 호모그래피를 이용한 자기-키포인트 라벨링(self-keypoint labelling)에 의해 학습 이미지에 의사 실측(pseudo-ground-truth) 키포인트를 생성하여 학습 데이터를 구성할 수 있다.

- [0128] 종래 자기-지도 키포인트 추출 방식인 SuperPoints는 라벨이 없는 이미지를 사용하여 특징점과 기술자를 추출할 수 있는 자기-학습 네트워크를 도입했다. GLAMpoints는 이전에 알려진 호모그래피 H를 사용하여 이미지 I 및 뒤틀린 이미지 I'에서 SIFT 특징 기술자를 추출하고, brute force matcher를 통해 양방향으로 일치하는 키포인트를 찾는다. 이 접근 방식을 사용하여 발견된 키포인트는 매칭 네트워크를 학습하기 위한 실측(ground truth)으로 사용된다. 본 발명에서는, SuperPoints, GLAMpoints와 유사한 자기-지도 방식으로 키포인트를 학습하므로, 이미지 학습을 위해 자기-키포인트 라벨링 프로세스가 필요하다.
- [0130] 도 4는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법에서, 단계 S100의 세부적인 흐름을 도시한 도면이다. 도 4에 도시된 바와 같이, 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법의 단계 S100은, 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성하는 단계(S110), 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성하는 단계(S120), 생성한 키포인트 및 키포인트 기술자를 이용해, 뒤틀린 이미지를 학습 이미지에 매핑하는 단계(S130), 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키포인트를 찾아, 키포인트-쌍 대응을 생성하는 단계(S140), 학습 이미지에 서로 다른 호모그래피를 적용해 단계 S110 내지 단계 S140을 수행하며, brute force matcher를 통해 자기-키포인트 라벨링을 하여 의사 실측 키포인트를 생성하는 단계(S150)를 포함하여 구성될 수 있다.
- [0132] 단계 S110에서는, 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성할 수 있다. 뒤틀린 이미지 I'는 레이블이 없는 학습 이미지 I에 i번째 호모그래피 H를 적용하여 생성될 수 있다.
- [0134] 단계 S120에서는, 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성할 수 있다. 단계 S120에서는, 이미지 I와 뒤틀린 이미지 I'에 동일한 SIFT를 적용하여 키포인트를 생성할 수 있다. 여기서, 키포인트 기술자(descriptor)로 DAISY 기술자를 사용할 수 있다. DAISY 기술자는 광도(photometric) 및 기하학적 변환에 강하고 SIFT보다 빠르고 밀도가 높은 기술자를 계산할 수 있다.
- [0136] 단계 S130에서는, 생성한 키포인트 및 키포인트 기술자를 이용해, 뒤틀린 이미지를 학습 이미지에 매핑할 수 있다.
- [0138] 단계 S140에서는, 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키포인트를 찾아, 키포인트-쌍 대응을 생성할 수 있다. 보다 구체적으로, 추출된 키포인트와 해당 키포인트 기술자를 기반으로, 두 이미지의 매핑 포인트 사이에서 유사한 포인트를 찾아 K 키포인트-쌍 대응(keypoint-pair correspondences)을 생성할 수 있다.
- [0140] 단계 S150에서는, 학습 이미지에 서로 다른 호모그래피를 적용해 단계 (1-1) 내지 단계 (1-4)를 수행하며, brute force matcher를 통해 자기-키포인트 라벨링을 하여 의사 실측 키포인트를 생성할 수 있다. 보다 구체적으로, 키포인트는 반복성을 만족해야 하므로 학습 이미지에 서로 다른 호모그래피 H를 적용하여 각 이미지에 대한 키포인트와 기술자를 얻고, 무라벨 이미지(레이블이 없는 학습 이미지)에 대한 키포인트 자기-라벨링은 brute force matcher를 통해 생성될 수 있다. brute force matcher는 한 개의 기술자와 다른 한 개의 기술자 사이의 거리를 계산해 매칭하고, 가까운 것을 반환하는 매칭자이다.
- [0142] 자기-지도 방법을 기반으로 한 이미지 학습을 위한 키포인트 라벨링이 적용된 후, 학습 이미지는 일련의 특징으로 표현될 수 있다. K 학습 세트의 k 번째 이미지는  $I^k = \{x_p^k \in \mathbb{R}^2, f_p^k \in \mathbb{R}^{200}, s_p^k \in \mathbb{R}^1\}$ 로 표시될 수 있다. 여기서,  $x_p^k$ 는 k번째 이미지의 p번째 키포인트의 실측(ground-truth) 위치를 나타낸다. 또한,  $f_p^k$ 는 DAISY 특징

기술자이고,  $s_p^k$ 는 신뢰도 점수이다. brute force distance에 따라 키포인트는 양수와 음수 쌍으로 나뉘고, 양수 쌍은  $s_p^k$ 에 초기 신뢰 점수 1을 할당하고 음수 쌍은  $s_p^k$ 에 0을 할당할 수 있다.

[0144] 단계 S200에서는, 다층 랜덤 포레스트(multi-layer random forest, MRF) 회귀 분석기(regressor) 기반으로 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈(200)을 생성할 수 있다. 이때, MRF 회귀 분석기는, coarse-to-fine 파이프라인 구조로 순차적 회귀를 적용하도록 복수의 레이어로 구성되며, 레이어 당 RF 노드 수는 단계 S200의 학습 과정에서 자동으로 결정되되, 2개 이상 8개 이하일 수 있다. 또한, 단계 S200에서는, N-겹 교차 검증을 이용해 MRF 모듈(200)을 생성할 수 있다.

[0146] 즉, MRF 모듈(200)은, 학습 과정에서 역전과 대신 각 레이어의 최적 규칙 및 RF 매개 변수가 자동으로 결정되는 동시에 N-겹 교차 검증을 사용하여 과적합 위험을 줄일 수 있다. 하나의 글로벌 회귀 분석기만 사용하는 것은 학습 과정에서 과적합되거나 보이지 않는 호모그래피 변환에 대처하지 못하기 때문에, 본 발명의 단계 S200에서는 과적합을 줄이고 이미지 뒤틀림에 강인한 모델을 설계하기 위해 end-to-end 학습에서 영감을 받은 coarse-to-fine 파이프라인 구조로 순차적 회귀를 적용할 수 있는 MRF 회귀 분석기를 구성하였다. 각 레이어의 RF 노드 수는 최소 2개에서 최대 8개 사이에서 순차적으로 증가하도록 설계되었으며, 학습 과정에서 레이어 당 최대 RF 노드 수 내에서 자동으로 결정된다. 첫 번째 레이어에서 초기 키포인트는 거칠게(roughly) 이동하나, 연속되는 레이어로 이동하면서 점차 키포인트의 예측을 구체화하게 된다.

[0148] 도 5는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법에서, 단계 S200의 MRF 회귀 분석기 학습 과정을 도시한 도면이고, 도 6은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법에서, 단계 S200의 세부적인 흐름을 도시한 도면이다. 도 6에 도시된 바와 같이, 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법의 단계 S200은, 전체 학습 데이터에서 N-1 폴드를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성하는 단계(S210), 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 의사 실측 키포인트를 학습 이미지에 매핑하는 단계(S220), 매핑에 의해 의사 실측 키포인트를 포함하는 격자의 중심점(초기 키포인트)을 기초로 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력하는 단계(S230), 초기 키포인트를 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 MRF 회귀 분석기의 제2 레이어를 위한 초기 키포인트를 생성하고, 제2 레이어를 학습하여 예측 변위를 출력하는 단계(S240), MRF 회귀 분석기의 L개의 레이어에 대해 단계 S240을 진행하는 단계(S250), 학습 세트의 N-1개의 폴드의 학습 후에, 검증 세트의 샘플을 이용해 손실을 계산하는 단계(S260), 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장하는 단계(S270), 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 단계 S210 내지 단계 S270을 반복 수행하는 단계(S280) 및 저장된 MRF 회귀 분석기의 구조 중에서 총 손실이 가장 작은 구조를 MRF 모듈(200)로 결정하는 단계(S290)를 포함하여 구현될 수 있다.

[0150] 이하에서는, 도 5 및 도 6을 참조하여 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법의 단계 S200의 흐름에 대해 상세히 설명하도록 한다.

[0152] 단계 S210에서는, 전체 학습 데이터에서 N-1 폴드(fold)를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성할 수 있다.

[0154] 단계 S220에서는, 단계 S210에서 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 단계 S100에서 자기-키포인트 라벨링에 의해 생성된 의사 실측 키포인트를 학습 이미지에 매핑할 수 있다. 보다 구체적으로, 학습 이미지를 9×9 픽셀 해상도로 구성된 격자로 분할하고, 자기-라벨 실측(self-labeled ground-truth) 키포인트가 이미지에 매핑될 수 있다.

[0156] 단계 S230에서는, 단계 S220의 매핑에 의해 의사 실측 키포인트를 포함하는 격자의 중심점(초기 키포인트)을 기초로 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력할 수 있다. 보다 구체적으로, 단계 S230에서는, 단계 S220의 매핑에 의해 의사 실측 키포인트를 포함하는 격자를 미리 정해진 서브 블록으로 분할하고, 분할된 각각의 서브 블록에서 추출한 특징 기술자(descriptors)를 결합해 특징 벡터를 구성하며, 특징 벡터와 초기 키포인트부터 의사 실측 키포인트까지의 변위를 제1 레이어의 입력으로 하여 학습을 수행할 수 있다.

[0158] 즉, 단계 S220의 매핑 후 의사 실측 키포인트를 포함하는 격자만  $3 \times 3$  서브 블록으로 분할하고, 각 서브 블록의 중심점에서 DAISY 기술자  $f_p^k$ 를 추출할 수 있다. 각각의 서브 블록에서 추출한 9개의 기술자를 결합하여 연결된 특징  $\hat{f}_p^k \in \mathbb{R}^{9 \times F}$ 을 생성할 수 있다. 여기서, 위 첨자 F는 특징 기술자의 차원을 나타낸다. 따라서 제1 레이어의 학습을 위한 p번째 학습 키포인트는  $F_p^k(l) = [\hat{f}_p^k, x_p^k, c_p^k, \Delta d_p^k, s_p^k]$ 로 구성될 수 있다. 여기서,  $\Delta d_p^k \in \mathbb{R}^2$ 는  $x_p^k$ 에서  $3 \times 3$  인접 서브 블록의 중심점  $c_p^k$ 까지의 변위 ( $\Delta dx$ ,  $\Delta dy$ )를 나타낸다.

[0160] RF를 구성하는 개별 트리를 학습하기 위해 정보 획득 목적 함수를 최대화하여 j번째 트리 노드의 최적 매개 변수를 계산할 수 있다. 회귀 트리의 최적 목적 함수를 결정하기 위해 다음 수학적 식 1과 같이 최소 제곱 오차 함수(least square error function, E)를 사용하여 오차를 최소화하는 분할 함수를 찾을 수 있다.

### 수학적 식 1

$$E_j = \sum_{s_j} (\Delta d_j - \overline{\Delta d_j})^2 - \sum_{i \in \{L, R\}} \left( \sum_{s_j^i} (\Delta d_j - \overline{\Delta d_j})^2 \right)$$

[0164] 여기서,  $s_j$ 는 노드 j에 도착하는 학습 데이터 세트를 나타낸다. 또한, L과 R은 각각 왼쪽과 오른쪽 분할 노드이며,  $\overline{\Delta d_j}$ 는 j번째 노드에 도달하는 모든 학습 키포인트에 대한  $\Delta d$ 의 평균 오프셋 벡터를 나타낸다. 각 리프 노드는 리프 노드 학습 샘플을 사용하여 2D 오프셋 벡터  $\Delta d_p^k(1)$ 를 예측하고 저장할 수 있다. 오프셋 벡터를 예측하기 위해 리프 노드에서 모든 학습 샘플의 평균이 사용될 수 있다.

[0166] 단계 S240에서는, 초기 키포인트를 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 MRF 회귀 분석기의 제2 레이어를 위한 초기 키포인트를 생성하고, 제2 레이어를 학습하여 예측 변위를 출력할 수 있다. 보다 구체적으로, 제1 레이어의 출력으로 초기 키포인트의 위치가 예측된 변위  $\Delta d_p^k(1)$ 를 이용해  $c_p^k(1+1) = c_p^k(1) + \Delta d_p^k(1)$ 로 이동되고, 이동된 키포인트 주위에  $3 \times 3$  서브 블록이 다시 생성될 수 있다. 같은 방식으로 새로운 특징  $F_p^k(1+1)$ 이 추출되어 제2 레이어에 입력될 수 있다.

[0168] 단계 S250에서는, MRF 회귀 분석기의 L개의 레이어에 대해 단계 S240를 반복 수행할 수 있다.

[0170] 단계 S260에서는, 단계 S210에서 선택한 N-1개의 폴드의 학습 후에, 검증 세트의 샘플을 이용해 손실을 계산할 수 있다. 최종 출력 레이어의 변위로 키포인트의 위치  $\hat{c}_p^k(L) = c(L-1) + \Delta d_p^k(L-1)$ 를 결정할 때, 다음 수학적 식 2의 손실을 계산하여 키포인트 위치를 최적화할 수 있다.

## 수학식 2

[0172]  $\mathcal{L}_{loc} = \sum_{p=1}^P \|\hat{c}_p^k - x_p^k\|$

[0174] 소스 이미지의 각 의사 실측 키포인트  $x_p^k$ 와 연관된 기술자  $f_p^k$ 와 MRF 회귀 모듈의 추정 키포인트 위치  $\hat{c}_p^k$ 에서 추출된 기술자  $\hat{f}_p^k$  간의 손실은 다음 수학식 3과 같을 수 있다.

## 수학식 3

[0176]  $\mathcal{L}_{des} = \sum_{p=1}^P \|\hat{f}_p^k - f_p^k\|$

[0178] 이 프로세스는 n번째 유효성 검증 세트의 모든 이미지에 대해 수행될 수 있다. n번째 검증 세트의 총 손실은 다음 수학식 4로 계산될 수 있다.

## 수학식 4

[0180]  $\mathcal{L}_n = \lambda_1 \mathcal{L}_{loc} + \lambda_2 \mathcal{L}_{des}$

[0182] 여기서,  $\lambda_1$  및  $\lambda_2$ 는 위치 손실과 기술자 손실 간의 균형을 유지하는 데 도움이 되는 가중치이다.

[0184] 단계 S270에서는, 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장할 수 있다. 즉, 학습된 n번째 MRF 구조와 단계 S260에서 계산된 총 손실  $\mathcal{L}_n$ 을 저장할 수 있다.

[0186] 단계 S280에서는, 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 단계 S210 내지 단계 S270을 반복 수행할 수 있다.

[0188] 단계 S290에서는, 다음 수학식 5와 같이, 단계 S280에서 저장된 MRF 회귀 분석기의 구조 중에서 총 손실  $\mathcal{L}_n$ 이 가장 작은 구조를 MRF 모듈(200)로 결정할 수 있다.

## 수학식 5

[0190]  $\text{MRF}^* = \arg \min_{n \in N} \mathcal{L}_n$

[0192] 단계 S300에서는, 단층 랜덤 포레스트(single-layer RF, SRF) 분류기(classifier) 기반으로, MRF 모듈(200)에서 추출한 키포인트의 신뢰도를 추정하는 SRF 모듈(300)을 생성할 수 있다.

[0194] 신뢰도 점수(confidence score)  $s_p^k \in [0, 1]$ 는 키포인트의 신뢰도를 추정할 수 있다. 초기  $s_p^k$  값은 양수 키포인트의 경우 1이고 음수 키포인트의 경우 0이다. 단계 S300에서는, 높은 신뢰도와 낮은 신뢰도를 구별하고 높은 신뢰도를 갖는 특징점에 대한 기술자를 추출하도록 학습시키는 것이 바람직하다. 신뢰도가 낮은 영역은 반복적인 패터를 가지고 있지만 실제 매칭에 중요하지 않은 영역이다. 예를 들어, 하늘, 파도, 땅이 고정되지 않거나 캡처 시간에 따라 달라질 수 있다.

[0196] SRF 모듈(300)은 키포인트에 대한 신뢰도 점수를 얻기 위해 하나의 레이어로 구성된 분류기를 학습한 것이다. SRF 모듈(300)의 목표는 높은 신뢰도와 낮은 신뢰도를 분류하여 높은 신뢰도로 특징점에 대한 기술자를 추출하도록 모델을 학습시키는 것이다. 신뢰도 추정을 위한 SRF 분류기를 학습하기 위한 p번째 키포인트의 특징은  $\mathbf{f}_p(l) = [\mathbf{f}_p, \mathbf{x}_p, s_p]$ 로 구성될 수 있다.

[0198] SRF 분류기는 1개의 레이어와 4개의 RF 노드로 구성될 수 있다. MRF 회귀 분석기와 마찬가지로 N-겹 교차 검증 방법을 사용하여 학습 및 테스트를 반복하고, 손실이 가장 적은 모델을 SRF 모듈(300)로 선택할 수 있다. 최적의 SRF 분류기는 양수 키포인트 사이의 클래스 확률 공간에서 거리를 최소화하고 음수 키포인트 사이의 거리를 최대화하기 위해 다음 수학적 6의 신뢰 손실  $L_{conf}$ 을 학습하려는 시도이다.

## 수학적 6

$$\mathcal{L}_{conf} = - \frac{\sum_{i=1}^N \sum_{j=1}^{class} s_p^{i,j} \log(p(x_p^{i,j}))}{2N_{pos}} - \frac{\sum_{i=1}^N \sum_{j=1}^{class} (1-s_p^{i,j}) \log(p(x_p^{i,j}))}{2N_{neg}}$$

[0200]

[0202] 여기서,  $p(x_p)$ 는 예측된 신뢰 확률이고,  $N_{pos}$  및  $N_{neg}$ 는 각각 양수 및 음수 키포인트의 수이다.

[0204] 동일한 방법을 사용하여 다음 수학적 7과 같이 신뢰 손실  $L_{conf}$ 이 가장 작은 SRF를 최종 SRF로 결정한다.

## 수학적 7

$$SRF^* = \arg \min_{n \in N} \mathcal{L}_{conf}$$

[0206]

[0208] 단계 S300 이후에, MRF 모듈(200) 및 SRF 모듈(300)은 트리로 구성된 규칙을 공유할 수 있다. 즉, 학습 후, DNN과 달리, MRF 모듈(200)과 SRF 모듈(300)은 가중치를 공유하지는 않지만, 도 2와 같이 트리로 구성된 규칙을

공유할 수 있다.

- [0210] 단계 S400에서는, 생성한 MRF 모듈(200)과 SRF 모듈(300)을 이용해 입력 이미지로부터 최종 키포인트를 선정하되, NMS(Non-maximum-suppression)를 사용해 MRF 모듈(200)에서 추출한 키포인트 중에서 신뢰도가 가장 높은 키포인트를 최종 키포인트로 선정할 수 있다. 보다 구체적으로, 단계 S400에서는, 입력 이미지를 서로 다른 스케일로 변환한 복수의 스케일 이미지를 MRF 모듈(200) 및 SRF 모듈(300)에 각각 적용하며, NMS를 사용해 대응창을 이동하면서 신뢰도 점수를 추정하되, 신뢰도 점수가 가장 높으면서 최소 임계값을 초과하는 키포인트를 최종 키포인트로 선정할 수 있다.
- [0212] 예를 들어, 도 2에 도시된 바와 같이, MRF 모듈(200)과 SRF 모듈(300)을 세 가지 스케일( $x_1$ ,  $x_{1.5}$ ,  $x_2$ ) 이미지 각각에 적용할 수 있다. NMS(Non-maximum-suppression)는  $x_{1.5}$  이미지를 기반으로  $x_1$  및  $x_2$  이미지(각 이미지가  $x_1$  크기로 정규화됨)에서 대응 창(corresponding windows)을 이동하면서, 대응 창 내에 두 개 이상의 겹치는 키포인트가 있는 경우 신뢰도 점수를 측정할 수 있다. 그리고 최고 점수를 달성하면서 최소 임계값을 초과하면 ( $p(x_p) > \tau$ ) 최종 키포인트로 선정될 수 있다. 도 2의 우측 말단에는 SRF의 예를 보여 주며, 강의 물결과 하늘에 나타난 키포인트가 사라진 것을 볼 수 있다.
- [0214] 요약하면, 도 2에 도시된 바와 같이, 입력 이미지를  $n \times n$ 의 미리 정해진 격자로 분할하고, 의사 실측 키포인트가 있는 격자에서 특징 벡터를 추출할 수 있다. 추출한 특징 벡터를 이용해 MRF 회귀 분석기가 각 레이어에 대해 초기 키포인트의 변위를 예측할 수 있다. 최종 예측된 변위가 적용된 키포인트로부터 특징 기술자를 추출해 SRF 모듈에 입력하고, SRF 모듈은 해당 키포인트의 신뢰도 점수를 추정할 수 있다. 이 과정은 규칙을 공유하여, 3개의 크기의 이미지에 대해 동일하게 적용될 수 있다. 윈도우 내에 존재하는 키포인트의 신뢰도 점수는 각 이미지 크기의 예측된 키포인트로부터 산출되며, NMS를 사용해 높은 신뢰도 점수를 갖는 키포인트만 남게 된다. SRF 및 NMS로부터 최종적으로 의미있는 키포인트를 얻을 수 있다.
- [0216] 테스트 단계에서는, 입력 테스트 이미지에서  $I^k = \{x_p^k, f_p^k\}$ 가 추출될 수 있다. 입력 이미지는 업 샘플링을 통해 3가지 크기로 변경되며, 공유된 MRF 모듈(200) 및 SRF 모듈(300)을 통해 반복 가능한 키포인트를 추출할 수 있다. 추출된 키포인트는 원본 이미지에 매핑되며 신뢰도 점수가 높은 키포인트만 NMS를 통해 유지될 수 있다.
- [0218] 실험
- [0220] 1. 학습 및 테스트 데이터
- [0222] 학습 레이블이 없는 118k 학습 이미지로 구성된 MS COCO 데이터 세트를 사용하여, 자기-지도 방식으로 KeyReg를 학습하였다. COCO 이미지가 중앙 ( $261 \times 261$  픽셀 해상도) 주변에서 잘린 후, 무작위로 선택된 증강(augmentation) 및 호모그래피 변환을 두 복제된 이미지 각각에 적용하였다. 증강에는 랜덤 가우시안 노이즈(random Gaussian noise)와 밝기 변경이 적용되었으며, 호모그래피 변환에는 원근, 평행 이동, 전단(shearing) 및 회전 배율의 임의 변경이 적용되었다.
- [0224] 테스트 데이터로는 키포인트 매칭 벤치마크 데이터 세트로 널리 사용되는 HPatches를 선정하여 성능을 평가하였다. HPatches에는 696개의 고유한 이미지가 있는 116개의 장면이 있으며, 큰 조명을 사용하는 57개의 시퀀스와 큰 관점의 변화를 가진 59개의 장면이 있다.

- [0226] 2. 평가 지표
- [0228] 제안된 KeyReg 및 기타 최신 방법의 정량적 성능 비교를 위해 반복성, 위치 오류, 호모그래피 정확도 및 일치 점수의 네 가지 실험 방법을 사용하였다. 반복성(repeatability)은 두 카메라 뷰에서 관찰된 키포인트의 품질을 설명하며, 위치 오류(localization error)는 특정 임계 값( $e=3$ )보다 낮은 대응되는 키포인트 간의 평균 유클리드 거리이다. 호모그래피 정확도(homography accuracy)는 전체 호모그래피 수에 대한 올바르게 추정된 호모그래피 수의 비율이다. 일치 점수(matching score)는 두 카메라 뷰에서 관찰된 모든 키포인트에 대한 올바른 일치 비율이다.
- [0230] 3 구현 세부 정보
- [0232] KeyReg의 첫 번째 모듈인 MRF 모듈(200)은 RF 노드의 4개 레이어로 구성되며, 각 레이어는 coarse-to-fine 방식에 따라 점차적으로 증가하는 노드 설정(예: MRF 노드  $\in [2, 4, 6, 8]$ )을 갖는다. 도 2 및 도 5에서 입력 데이터는 레이어가 깊어지고 넓어짐에 따라 지속적으로 변환된다.
- [0234] KeyReg의 두 번째 모듈인 SRF 모듈(300)은 4개의 RF 노드가 있는 단일 레이어 구조로 구성하였다.
- [0236] 평가 절차는 i7-7700K CPU와 32GB RAM을 사용하여 구축된 PC 환경에서 수행되었다. 사용된 소프트웨어 사양은 OpenCV(3.4.2.16) 및 Windows 10을 기반으로 MRF 및 SRF를 구축하기 위한 Scikit-learn(0.22.2.post1)이다.
- [0238] 4. 결과
- [0240] 본 발명에서 제안된 KeyReg의 성능을 측정하기 위해 HPatches 데이터 세트를 사용하여 4개의 평가 지표를 최첨단 방법과 비교하였다. 도 7 내지 도 9는 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치를 다른 최첨단 방법과 비교한 결과를 도시한 도면이다.
- [0242] (1) HPatches 반복성
- [0244] 반복성은 두 카메라 뷰에서 계산된 대응점의 평균값이다. 특정 픽셀 거리 내에 가까운 키포인트가 있는 경우 관찰된 키포인트는 우수한 반복성을 가진 것으로 간주할 수 있다. 도 7은 픽셀 거리( $e$ )가 3일 때 반복성에 대한 기존의 수작업 및 최신 학습 기반 방법을 사용한 평가 결과를 표시한 것이다. 저해상도 및 중해상도 데이터에 대해 비교를 수행하였다. 관찰된 키포인트는 SuperPoint와 UnsuperPoint의 연구와 유사하게 NMS=4를 사용하여 개선하였다. 본 발명(KeyReg)은 두 가지 해상도 모두에서 다른 방법보다 높은 반복성을 보이는 것을 확인할 수 있다.
- [0246] 조도 및 시점의 변화에 따른 반복성 성능 평가에서 KeyReg는 조도 변화에 있어서 다른 방법에 비해 약간 더 나은 성능을 보이지만, 시점의 변화에서 특히 강력한 성능을 보이는 것을 확인할 수 있다.
- [0248] (2) HPatches 위치 오류
- [0250] 도 8은 임계값  $e=3$  및 NMS=4인 위치 오류를 표시한 것이다. 반복성과 유사하게 위치 오류는 두 카메라 뷰에서

대응되는 키포인트 사이의 유클리드 거리의 평균으로 계산된다.

[0252] 도 8에서 확인할 수 있듯이 KeyReg는 대표적인 CNN 기반 방법인 SuperPoint에 비해  $240 \times 320$ 의 이미지 해상도에 대해 0.088,  $480 \times 640$ 의 이미지 해상도에 대해 0.002의 더 낮은 위치 오류를 보인다. 이 결과로부터 KeyReg는 이미지의 크기 변화와 관계없이 잘 맞는 키포인트를 추출하는 것을 알 수 있다.

[0254] (3) HPatches Homography 추정

[0256] HPatches 데이터 세트의 매칭 능력에 대한 KeyReg의 성능을 평가하기 위해, 호모그래피 정확도와 매칭 점수를 사용하여 최신 키포인트 탐지기뿐만 아니라 수작업 방식에 대해 KeyReg를 평가했다. 호모그래피의 정확도는 주어진 임계값( $e=1, 3, 5$ )으로 정량화될 수 있다. 가장 가까운 이웃 아래의 두 지점이 해당 지점 사이에 특정 거리 임계값(예:  $e=3$ )을 가질 때 올바르게 일치한 것으로 판단된다.

[0258] 도 9는 저해상도 및 중해상도에 대한 호모그래피 추정의 평가 결과를 나타낸 것이다. 두 가지 이미지 크기에서, 학습 기반 방법인 SuperPoint, UnsuperPoint는 수작업 방법인 ORB, SURF, AKAZE, BRISK, SIFT보다 더 높은 호모그래피 정확도와 일치 점수를 나타낸다. NMS가 0일 때 KeyReg는  $e=3, 5$ 일 때 이미지 크기와 관계없이 최상의 성능을 보여준다. 또한, 일치 점수는  $240 \times 320$ 의 픽셀 해상도에 대해 0.08,  $480 \times 640$ 의 픽셀 해상도에 대해 0.058의 더 큰 값을 보인다.

[0260] NMS가 4인 경우 SuperPoint, UnsuperPoint, KeyReg 세 가지 방법은  $e$ 가 증가함에 따라 유사한 값을 균등하게 표시한다. KeyReg는 신뢰도 점수를 기반으로 키포인트를 선택하기 때문에  $NMS=0$ 과 같은 좋은 성능을 보이지 않으며, NMS 창 크기 커질수록 주변에서 높은 점수 키포인트를 선택할 가능성이 커진다. KeyReg는 신뢰도가 낮은 키포인트를 제거하기 때문에 NMS의 창 크기가 작더라도 호모그래피 및 일치 점수의 합리적인 정확도를 얻을 수 있다는 것이 실험적으로 밝혀졌다.

[0262] 도 10은 본 발명의 일실시예에 따른 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치와 다른 최첨단 방법의 키포인트 매핑 예제를 도시한 도면이다. 도 10에 도시된 바와 같이, KeyReg는 이미지의 크기, 회전, 조도 변화 측면에서 의미 있는 키포인트를 잘 찾아내어 정확한 매칭이 이루어졌음을 알 수 있다. AKAZE는 조명의 변화에는 강하지만 관점의 변화에는 약했다. SIFT는 크기 변화에 다소 견고했지만, 관점 변화에 부정확한 결과를 제공했고, 특히 조명 변화에 약하다(5번째 행). Superpoint는 최소한의 방식으로 키포인트를 추출하며 키포인트 일치도 정확하게 적용되지만, 네 번째 행과 같이 크기와 회전 변환을 함께 수행하면 일치에 실패하였다.

[0264] 이와 같이, 본 발명에서 제안하고 있는 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치는, 이미지에서 다양한 호모그래피 변환이 발생하더라도 키포인트가 정확하게 감지된다는 것을 실험적으로 보여주었다. KeyReg의 MRF 구조는 키포인트 회귀를 위한 레이어 별 구조 설계에서 처음 시도한 것으로, 레이어 당 노드 수를 가변적으로 결정하여 coarse-to-fine 방식으로 키포인트를 회귀하도록 설계되었다. SRF 분류기는 키포인트의 신뢰도 점수를 출력하여 일치의 정확도를 향상시켰다.

[0266] HPatches 데이터 세트의 경우 반복성, 호모그래피 정확도, 매칭 점수, 위치 오류에 대한 다양한 실험을 통해, 본 발명에서 제안하고 있는 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치는 DNN을 사용하지 않았음에도 불구하고 최신 방법보다 지속적으로 더 나은 결과를 보여주었다.

[0268] 전술한 바와 같이, 본 발명에서 제안하고 있는 다층 랜덤 포레스트 회귀 분석기 기반의 자기-지도 키포인트 검출 방법 및 장치에 따르면, 다층 랜덤 포레스트 회귀 분석기와 단일 랜덤 포레스트 분류기로 키포인트를 검출하도록 구성함으로써, 다층 랜덤 포레스트 구조를 키포인트의 회귀에 적용하여 깊은 구조를 유지하면서 역전과 없이 반복성 있는 키포인트의 위치를 검출할 수 있고, 원본 이미지와 뒤틀린 이미지 사이의 자기-지도 학습 방법을 이용함으로써 비-라벨 데이터로 학습을 할 수 있으며, 단일 랜덤 포레스트 분류기를 통해 신뢰도가 낮은 키포인트를 제거해 매칭의 정확도를 높일 수 있다.

[0270] 한편, 본 발명은 다양한 통신 단말기로 구현되는 동작을 수행하기 위한 프로그램 명령을 포함하는 컴퓨터에서 판독 가능한 매체를 포함할 수 있다. 예를 들어, 컴퓨터에서 판독 가능한 매체는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media) 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치를 포함할 수 있다.

[0272] 이와 같은 컴퓨터에서 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 이때, 컴퓨터에서 판독 가능한 매체에 기록되는 프로그램 명령은 본 발명을 구현하기 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 예를 들어, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함할 수 있다.

[0274] 이상 설명한 본 발명은 본 발명이 속한 기술분야에서 통상의 지식을 가진 자에 의하여 다양한 변형이나 응용이 가능하며, 본 발명에 따른 기술적 사상의 범위는 아래의 특허청구범위에 의하여 정해져야 할 것이다.

## 부호의 설명

[0275] 10: 자기-지도 키포인트 검출 장치  
 100: 라벨링 모듈  
 200: MRF 모듈  
 300: SRF 모듈  
 400: 검출 모듈  
 S100: 키포인트 기술자 및 호모그래피를 이용한 자기-키포인트 라벨링에 의해 학습 이미지에 의사 실측 키포인트를 생성하여 학습 데이터를 구성하는 단계  
 S110: 레이블이 없는 학습 이미지에 호모그래피를 적용해 뒤틀린 이미지를 생성하는 단계  
 S120: 학습 이미지 및 뒤틀린 이미지에서 키포인트 및 키포인트 기술자를 생성하는 단계  
 S130: 생성한 키포인트 및 키포인트 기술자를 이용해, 뒤틀린 이미지를 학습 이미지에 매핑하는 단계  
 S140: 학습 이미지와 뒤틀린 이미지 사이에서 유사한 키포인트를 찾아, 키포인트-쌍 대응을 생성하는 단계  
 S150: 학습 이미지에 서로 다른 호모그래피를 적용해 단계 S110 내지 단계 S140을 수행하며, brute force matcher를 통해 자기-키포인트 라벨링을 하여 의사 실측 키포인트를 생성하는 단계  
 S200: 다층 랜덤 포레스트 회귀 분석기 기반으로 학습 데이터를 학습하여, 이미지로부터 반복 가능한 키포인트를 추출하는 MRF 모듈을 생성하는 단계  
 S210: 전체 학습 데이터에서 N-1 폴드를 학습 세트로 선택하고, 선택에서 제외된 폴드는 검증 세트로 구성하는 단계  
 S220: 선택한 폴드의 학습 이미지를 미리 정해진 격자로 분할하고, 의사 실측 키포인트를 학습 이미지에 매핑하는 단계

S230: 매핑에 의해 의사 실측 키폰트를 포함하는 격자의 중심점(초기 키폰트)을 기초로 MRF 회귀 분석기의 제1 레이어를 학습하여 예측 변위를 출력하는 단계

S240: 초기 키폰트를 제1 레이어에서 출력된 예측 변위를 반영해 이동시켜 MRF 회귀 분석기의 제2 레이어를 위한 초기 키폰트를 생성하고, 제2 레이어를 학습하여 예측 변위를 출력하는 단계

S250: MRF 회귀 분석기의 L개의 레이어에 대해 단계 S240을 진행하는 단계

S260: 학습 세트의 N-1개의 폴드의 학습 후에, 검증 세트의 샘플을 이용해 손실을 계산하는 단계

S270: 학습된 MRF 회귀 분석기의 구조 및 총 손실을 저장하는 단계

S280: 전체 학습 데이터의 모든 폴드가 검증 세트로 사용될 때까지 단계 S210 내지 단계 S270을 반복 수행하는 단계

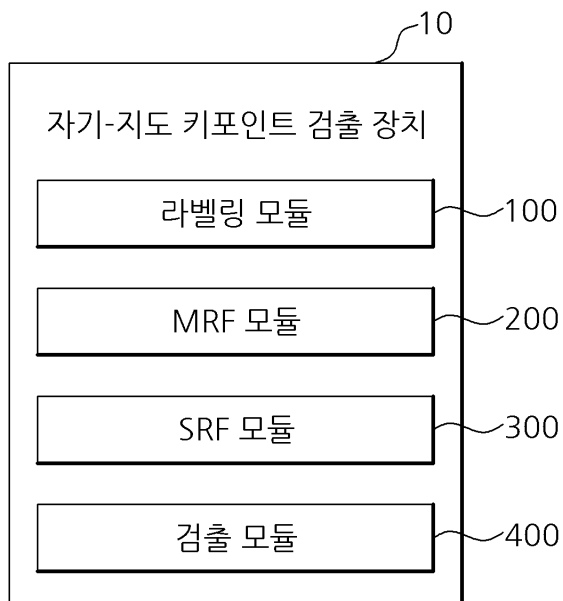
S290: 저장된 MRF 회귀 분석기의 구조 중에서 총 손실이 가장 작은 구조를 MRF 모듈로 결정하는 단계

S300: 단층 랜덤 포레스트 분류기 기반으로, MRF 모듈에서 추출한 키폰트의 신뢰도를 추정하는 SRF 모듈을 생성하는 단계

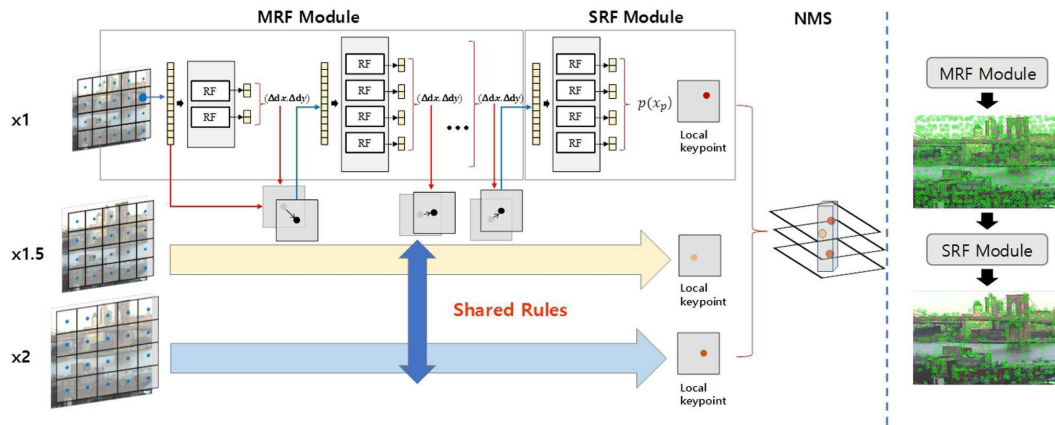
S400: MRF 모듈과 SRF 모듈을 이용해 입력 이미지로부터 최종 키폰트를 선정하되, NMS를 사용해 MRF 모듈에서 추출한 키폰트 중에서 신뢰도가 가장 높은 키폰트를 최종 키폰트로 선정하는 단계

## 도면

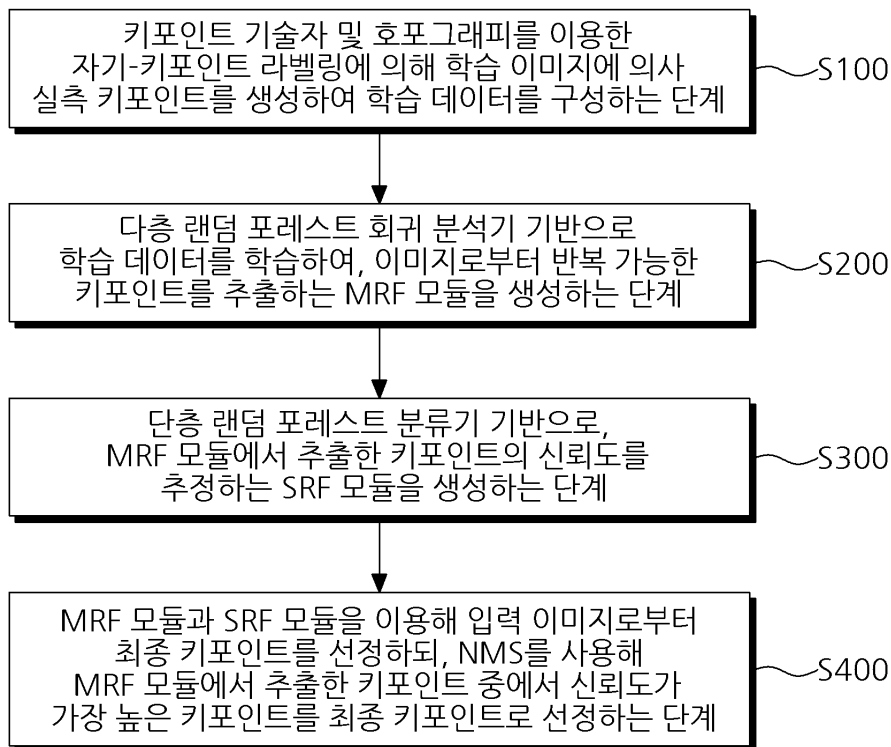
### 도면1



도면2

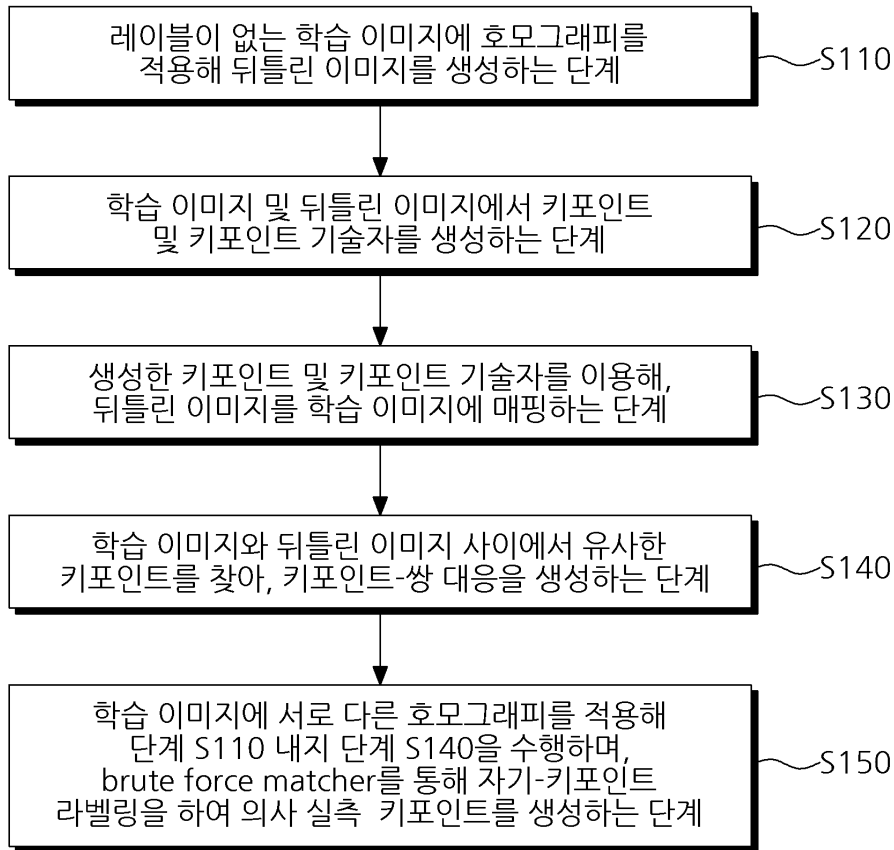


도면3

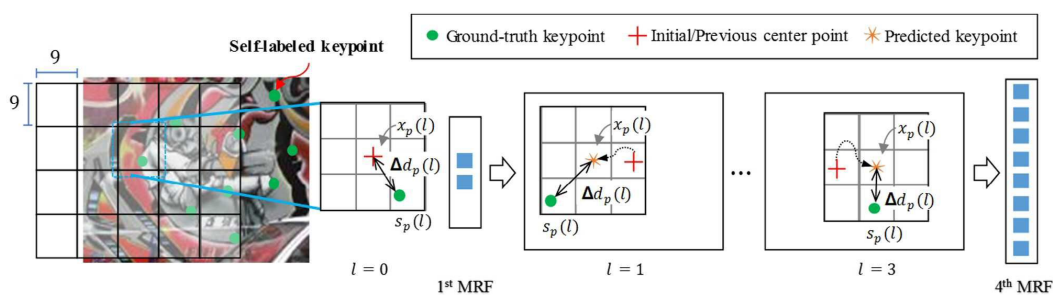


도면4

S100

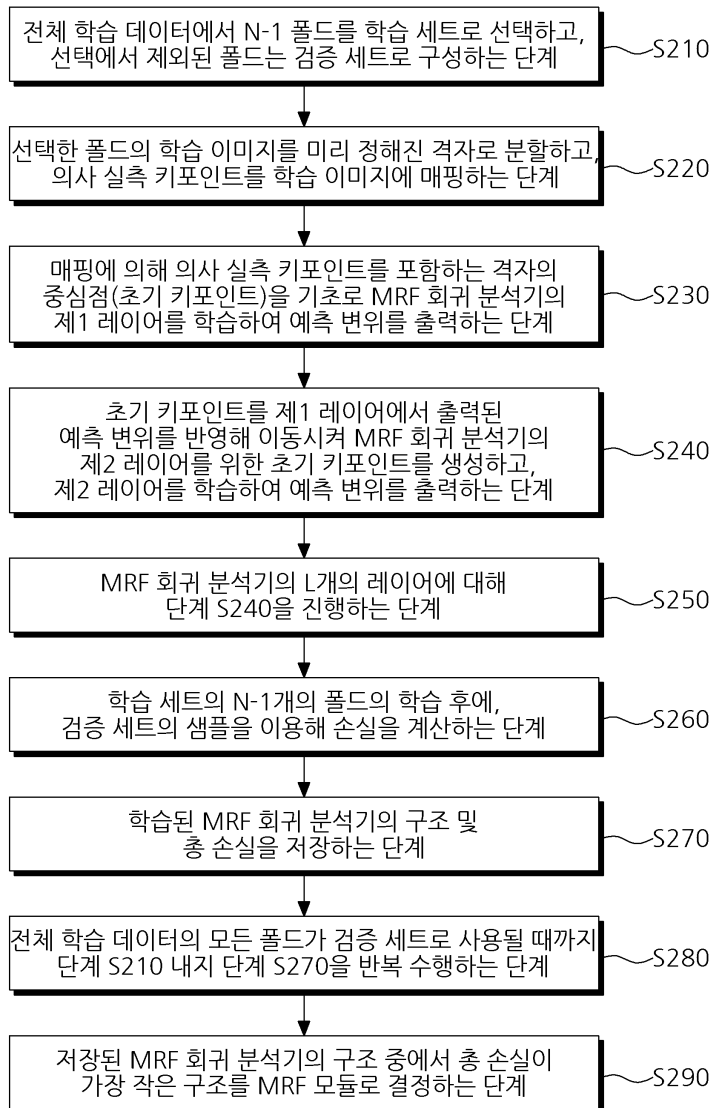


도면5



도면6

S200



도면7

| Detectors         | Repeatability $\uparrow$ |                  |
|-------------------|--------------------------|------------------|
|                   | 240 $\times$ 320         | 480 $\times$ 640 |
| ORB [28]          | 0.532                    | 0.525            |
| SURF [6]          | 0.491                    | 0.468            |
| AKAZE [7]         | 0.555                    | 0.569            |
| BRISK [31]        | 0.566                    | 0.505            |
| SIFT [5]          | 0.490                    | 0.491            |
| SuperPoint [14]   | 0.637                    | 0.592            |
| UnsuperPoint [19] | 0.645                    | 0.612            |
| <b>KeyReg</b>     | <b>0.713</b>             | <b>0.676</b>     |
| Changes           |                          |                  |
|                   | 240 $\times$ 320         | 480 $\times$ 640 |
|                   | Illumi nation            | View point       |
| SIFT [5]          | 0.464                    | 0.516            |
| AKAZE[7]          | 0.558                    | 0.552            |
| SuperPoint [14]   | 0.665                    | 0.609            |
| <b>KeyReg</b>     | <b>0.689</b>             | <b>0.738</b>     |

도면8

| Detectors       | Localization Error $\downarrow$ |                  |
|-----------------|---------------------------------|------------------|
|                 | 240 $\times$ 320                | 480 $\times$ 640 |
| ORB [30]        | 1.429                           | 1.430            |
| SURF [6]        | 1.150                           | 1.244            |
| AKAZE [7]       | 1.007                           | 1.126            |
| BRISK [31]      | 1.077                           | 1.207            |
| SIFT [5]        | <b>0.855</b>                    | <b>1.011</b>     |
| SuperPoint [14] | 1.341                           | 1.212            |
| KeyReg          | 1.253                           | 1.210            |

도면9

| Detector          | 240 × 320           |              |              |                   | 480 × 640           |              |              |                   |
|-------------------|---------------------|--------------|--------------|-------------------|---------------------|--------------|--------------|-------------------|
|                   | Homography Accuracy |              |              | Matching<br>Score | Homography Accuracy |              |              | Matching<br>Score |
|                   | $e = 1$             | $e = 3$      | $e = 5$      |                   | $e = 1$             | $e = 3$      | $e = 5$      |                   |
| $NMS=0$           |                     |              |              |                   |                     |              |              |                   |
| ORB [30]          | 0.131               | 0.422        | 0.540        | 0.218             | 0.286               | 0.607        | 0.710        | 0.204             |
| SURF [6]          | 0.397               | 0.702        | 0.762        | 0.255             | 0.421               | 0.754        | 0.812        | 0.230             |
| AKAZE [7]         | 0.378               | 0.724        | 0.791        | 0.311             | 0.369               | 0.743        | 0.812        | 0.289             |
| BRISK [31]        | 0.414               | 0.767        | 0.826        | 0.258             | 0.300               | 0.653        | 0.746        | 0.211             |
| SIFT [5]          | <b>0.622</b>        | 0.845        | 0.878        | 0.304             | <b>0.602</b>        | 0.833        | 0.876        | 0.265             |
| SuperPoint [14]   | 0.491               | 0.833        | 0.893        | 0.318             | 0.509               | 0.834        | 0.900        | 0.281             |
| UnsuperPoint [19] | 0.579               | 0.855        | 0.903        | 0.424             | 0.493               | 0.843        | 0.905        | 0.383             |
| KeyReg            | 0.612               | <b>0.891</b> | <b>0.930</b> | <b>0.506</b>      | 0.516               | <b>0.872</b> | <b>0.916</b> | <b>0.441</b>      |
| $NMS=4$           |                     |              |              |                   |                     |              |              |                   |
| SuperPoint [14]   | 0.464               | 0.831        | 0.903        | 0.500             | 0.419               | 0.819        | <b>0.912</b> | 0.441             |
| UnsuperPoint [19] | 0.557               | <b>0.871</b> | <b>0.921</b> | <b>0.501</b>      | <b>0.521</b>        | 0.831        | 0.902        | <b>0.452</b>      |
| KeyReg            | <b>0.617</b>        | <b>0.871</b> | 0.910        | 0.494             | 0.486               | <b>0.847</b> | 0.906        | 0.429             |

도면10

