



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0112045
(43) 공개일자 2024년07월18일

(51) 국제특허분류(Int. Cl.)
G06V 10/77 (2022.01) G06N 3/08 (2023.01)
G06V 10/14 (2022.01)
(52) CPC특허분류
G06V 10/7715 (2023.08)
G06N 3/08 (2023.01)
(21) 출원번호 10-2023-0004148
(22) 출원일자 2023년01월11일
심사청구일자 2023년01월11일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
이상철
인천광역시 연수구 아트센터대로97번길 76, 1402동 805호(송도동, 송도 더샵 하버뷰 14단지)
정예립
인천광역시 미추홀구 용정공원로83번길 49, 2062호(용현동)
(74) 대리인
양성보

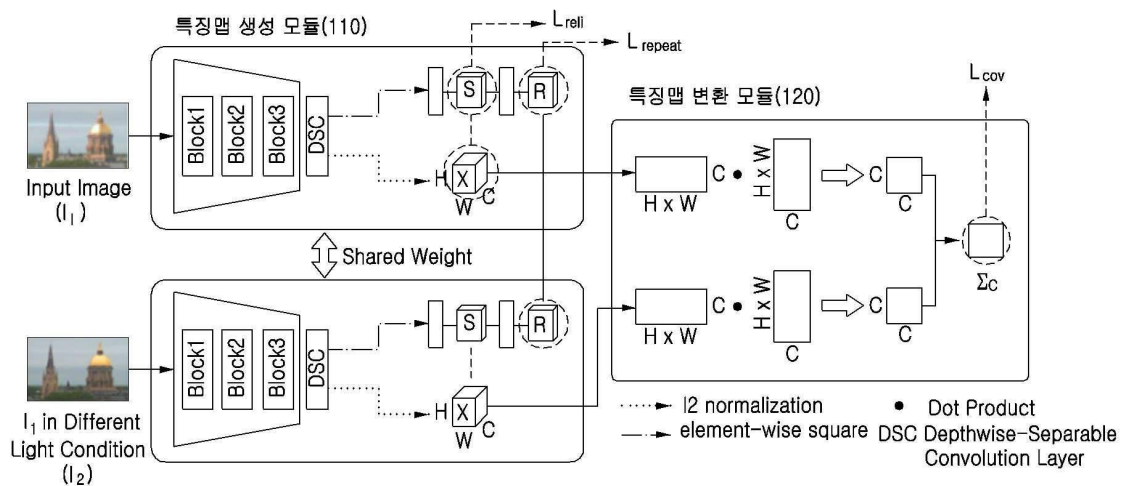
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 딥러닝을 이용한 특징맵 변환을 통해 향상된 특징점을 추출하는 방법 및 시스템

(57) 요약

특징점 추출 시스템에 의해 수행되는 특징점 추출 방법은, 특징맵 생성(Feature Map Generation; FMG) 모듈에서, 이미지 쌍으로부터 추출된 특징 각각에 대해 각각의 특징맵을 획득하는 단계; 및 특징맵 변환(Feature Map Transformation Module; FMT) 모듈에서, 구조적 정보와 스타일 정보를 이용하여 상기 각각의 획득된 특징맵으로부터 영역 예측을 최소화하기 위한 특징점을 추출하는 단계를 포함할 수 있다.

대표도



(52) CPC특허분류

G06V 10/14 (2023.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711162433
과제번호	2021R1A2C2010893
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구
연구과제명	[Ezbaro] 최소 영상에 특화된 영상 분해-합성을 통한 멀티도메인 기반 관심 지표 및
객체 검출 기술 연구	
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2022.03.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711174118
과제번호	00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신인재양성(교육운영수익)
연구과제명	[Ezbaro][정부+민간] 인공지능융합혁신인재양성(1차년도)
기 여 율	1/2
과제수행기관명	인하대학교 산학협력단
연구기간	2022.07.01 ~ 2023.02.28

명세서

청구범위

청구항 1

특징점 추출 시스템에 의해 수행되는 특징점 추출 방법에 있어서,

특징맵 생성(Feature Map Generation; FMG) 모듈에서, 이미지 쌍으로부터 추출된 특징 각각에 대해 각각의 특징맵을 획득하는 단계; 및

특징맵 변환(Feature Map Transformation Module; FMT) 모듈에서, 구조적 정보와 스타일 정보를 이용하여 상기 각각의 획득된 특징맵으로부터 영역 예측을 최소화하기 위한 특징점을 추출하는 단계

를 포함하는 특징점 추출 방법.

청구항 2

제1항에 있어서,

상기 획득하는 단계는,

특징맵 생성 네트워크를 통해 상기 이미지 쌍으로부터 추출된 특징 각각에 대해 설명자와, 신뢰성 맵과 반복성 맵을 포함하는 지점 추출 특징맵을 출력하는 단계

를 포함하는 특징점 추출 방법.

청구항 3

제2항에 있어서,

상기 획득하는 단계는,

상기 출력된 설명자와, 신뢰성 맵과 반복성 맵을 이용하여 반복성 및 신뢰성에 대한 손실 함수를 계산하는 단계를 포함하는 특징점 추출 방법.

청구항 4

제2항에 있어서,

상기 특징맵 생성 네트워크는,

기 이미지 쌍에서 특징 정보를 추출하기 위해 이미지 쌍에 대응되는 각각의 백본 네트워크와, 상기 특징맵 생성 네트워크의 가중치를 줄이기 위해 상기 백본 네트워크의 뒤에 깊이별 분리 가능 컨볼루션로 구성된 것을 특징으로 하는 특징점 추출 방법.

청구항 5

제3항에 있어서,

상기 획득하는 단계는,

제1 이미지의 각 픽셀에서의 로컬 설명자와 제2이미지의 각 픽셀에서의 로컬 설명자를 이용하여 신뢰성 점수를 예측하고, 상기 예측된 신뢰성 점수를 이용하여 미분 가능한 근사치로 최적화된 신뢰성 손실을 계산하는 단계

를 포함하는 특징점 추출 방법.

청구항 6

제3항에 있어서,

상기 획득하는 단계는,

상기 이미지 쌍의 특징 쌍 간의 피크 예측 및 유사성을 이용하여 반복성 손실을 계산하는 단계를 포함하는 특징점 추출 방법.

청구항 7

제1항에 있어서,

상기 추출하는 단계는,

상기 스타일 및 구조적 정보를 이용하여 상기 각각의 획득된 특징맵을 변환하고 조작하기 위해 설명자를 사용하여 변환 손실을 계산하는 단계

를 포함하는 특징점 추출 방법.

청구항 8

제7항에 있어서,

상기 추출하는 단계는,

제1 이미지의 설명자 및 제2 이미지의 설명자로부터 대체된 각각의 표준화된 특징맵에 의해 생성된 제1 공분산 행렬 및 제2 공분산 행렬을 획득하고, 상기 획득된 제1 공분산 행렬 및 상기 획득된 제2 공분산 행렬의 차이를 계산하여 광도 변환에 대한 공분산의 민감도를 의미하는 행렬을 생성하는 단계

를 포함하는 특징점 추출 방법.

청구항 9

제8항에 있어서,

상기 광도 변환에 대한 공분산의 민감도를 의미하는 행렬은,

공분산의 값이 기 설정된 값 이상인 요소에 대해 스타일 정보를 유지시키고, 상기 공분산의 값이 기 설정된 값 이하인 요소에 대해 구조적 정보를 유지시키는 것을 특징으로 하는 특징점 추출 방법.

청구항 10

특징점 추출 시스템에 있어서,

이미지 쌍으로부터 추출된 특징 각각에 대해 각각의 특징맵을 획득하는 특징맵 생성(Feature Map Generation; FMG) 모듈; 및

구조적 정보와 스타일 정보를 이용하여 상기 각각의 획득된 특징맵으로부터 영역 예측을 최소화하기 위한 특징점을 추출하는 특징맵 변환(Feature Map Transformation Module; FMT) 모듈

을 포함하는 특징점 추출 시스템.

발명의 설명

기술 분야

아래의 설명은 특징점 추출 기술에 관한 것이다.

배경 기술

매칭을 위해 로컬 특징을 추출하고 설명하는 것은 특히 이미지 매칭, 검색, 추적 및 3D 재구성을 포함하는 컴퓨터 비전 작업에서 필수적이다. 특징 일치는 특징 감지, 특징 설명, 특징 일치라는 두 개의 유사한 이미지가 지정될 때 세 가지의 주요 단계에 초점을 맞춘다. 특징 일치의 주요 목표는 이전 애플리케이션의 메모리 공간을 최소화하면서 일치 정확도를 최적화하는 것이다. 추출된 특징은 희소하고 반복 가능하며 정밀해야 한다. 각

이미지의 코너와 같은 주요 특징은 처음에 감지 단계에서 관심 지점으로 인식된다. 그런 다음 이러한 관심 지점의 이웃 영역을 기반으로 로컬 설명자를 추출하여 일치 알고리즘에 사용한다. 그러나, 로컬 특징을 추출하는 종래의 기술들은 이미지의 빛과 구조 정보에 대한 고려가 불충분 하였다.

[0004] 참고로, 대한민국 등록특허 제10-2234311호에 이미지로부터 특징점을 추출하고 이미지의 특징 정보를 기술하는 이미지 특징 추출 방법 및 시스템이 개시되어 있다.

발명의 내용

해결하려는 과제

[0006] 빛과 관점에 관계없이 두드러진 로컬 특징을 견고하게 추출하는 방법 및 시스템을 제공할 수 있다.

[0007] 스타일 정보와 구조적 정보를 사용하여 이미지 분산의 불일치 문제를 해결하는 방법 및 시스템을 제공할 수 있다.

과제의 해결 수단

[0009] 특징점 추출 시스템에 의해 수행되는 특징점 추출 방법은, 특징맵 생성(Feature Map Generation; FMG) 모듈에서, 이미지 쌍으로부터 추출된 특징 각각에 대해 각각의 특징맵을 획득하는 단계; 및 특징맵 변환(Feature Map Transformation Module; FMT) 모듈에서, 구조적 정보와 스타일 정보를 이용하여 상기 각각의 획득된 특징맵으로부터 영역 예측을 최소화하기 위한 특징점을 추출하는 단계를 포함할 수 있다.

[0010] 상기 획득하는 단계는, 특징맵 생성 네트워크를 통해 상기 이미지 쌍으로부터 추출된 특징 각각에 대해 설명자와, 신뢰성 맵과 반복성 맵을 포함하는 지점 추출 특징맵을 출력하는 단계를 포함할 수 있다.

[0011] 상기 획득하는 단계는, 상기 출력된 설명자와, 신뢰성 맵과 반복성 맵을 이용하여 반복성 및 신뢰성에 대한 손실 함수를 계산하는 단계를 포함할 수 있다.

[0012] 상기 특징맵 생성 네트워크는, 상기 이미지 쌍에서 특징 정보를 추출하기 위해 이미지 쌍에 대응되는 각각의 백본 네트워크와, 상기 특징맵 생성 네트워크의 가중치를 줄이기 위해 상기 백본 네트워크의 뒤에 깊이별 분리 가능 컨볼루션으로 구성된 것일 수 있다.

[0013] 상기 획득하는 단계는, 제1 이미지의 각 픽셀에서의 로컬 설명자와 제2이미지의 각 픽셀에서의 로컬 설명자를 이용하여 신뢰성 점수를 예측하고, 상기 예측된 신뢰성 점수를 이용하여 미분 가능한 근사치로 최적화된 신뢰성 손실을 계산하는 단계를 포함할 수 있다.

[0014] 상기 획득하는 단계는, 상기 이미지 쌍의 특징 쌍 간의 피크 예측 및 유사성을 이용하여 반복성 손실을 계산하는 단계를 포함할 수 있다.

[0015] 상기 추출하는 단계는, 상기 스타일 및 구조적 정보를 이용하여 상기 각각의 획득된 특징맵을 변환하고 조작하기 위해 설명자를 사용하여 변환 손실을 계산하는 단계를 포함할 수 있다.

[0016] 상기 추출하는 단계는, 제1 이미지의 설명자 및 제2 이미지의 설명자로부터 대체된 각각의 표준화된 특징맵에 의해 생성된 제1 공분산 행렬 및 제2 공분산 행렬을 획득하고, 상기 획득된 제1 공분산 행렬 및 상기 획득된 제2 공분산 행렬의 차이를 계산하여 광도 변환에 대한 공분산의 민감도를 의미하는 행렬을 생성하는 단계를 포함할 수 있다.

[0017] 상기 광도 변환에 대한 공분산의 민감도를 의미하는 행렬은, 공분산의 값이 기 설정된 값 이상인 요소에 대해 스타일 정보를 유지시키고, 상기 공분산의 값이 기 설정된 값 이하인 요소에 대해 구조적 정보를 유지시키는 것일 수 있다.

[0018] 특징점 추출 시스템은, 이미지 쌍으로부터 추출된 특징 각각에 대해 각각의 특징맵을 획득하는 특징맵 생성(Feature Map Generation; FMG) 모듈; 및 구조적 정보와 스타일 정보를 이용하여 상기 각각의 획득된 특징맵으로부터 영역 예측을 최소화하기 위한 특징점을 추출하는 특징맵 변환(Feature Map Transformation Module; FMT) 모듈을 포함할 수 있다.

발명의 효과

- [0020] 빛과 관점에 관계없이 두드러진 로컬 특징을 견고하게 추출할 수 있다.
- [0021] 스타일 정보와 구조적 정보를 사용하여 이미지 분산의 불일치 문제를 해결할 수 있다.

도면의 간단한 설명

- [0023] 도 1은 일 실시예에 있어서, 특징점 추출 시스템의 구성을 설명하기 위한 블록도이다.
- 도 2는 일 실시예에 있어서, 특징점 추출 방법을 설명하기 위한 흐름도이다.
- 도 3은 일 실시예에 있어서, 특징맵 생성 모듈의 네트워크 아키텍처를 설명하기 위한 도면이다.
- 도 4는 일 실시예에 있어서, 특징맵 변환 모듈에서 특징맵을 변환하는 동작을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0024] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다.
- [0026] 도 1은 일 실시예에 있어서, 특징점 추출 시스템의 구성을 설명하기 위한 블록도이고, 도 2는 일 실시예에 있어서, 특징점 추출 방법을 설명하기 위한 흐름도이다.
- [0027] 특징점 추출 시스템(100)의 프로세서는 특징맵 생성 모듈(110) 및 특징맵 변환 모듈(120)을 포함할 수 있다. 이러한 프로세서의 구성요소들은 특징점 추출 시스템에 저장된 프로그램 코드가 제공하는 제어 명령에 따라 프로세서에 의해 수행되는 서로 다른 기능들(different functions)의 표현들일 수 있다. 프로세서 및 프로세서의 구성요소들은 도 2의 특징점 추출 방법이 포함하는 단계들(210 내지 220)을 수행하도록 특징점 추출 시스템을 제어할 수 있다. 이때, 프로세서 및 프로세서의 구성요소들은 메모리가 포함하는 운영체제의 코드와 적어도 하나의 프로그램의 코드에 따른 명령(instruction)을 실행하도록 구현될 수 있다.
- [0028] 프로세서는 특징점 추출 방법을 위한 프로그램의 파일에 저장된 프로그램 코드를 메모리에 로딩할 수 있다. 예를 들면, 특징점 추출 시스템에서 프로그램이 실행되면, 프로세서는 운영체제의 제어에 따라 프로그램의 파일로부터 프로그램 코드를 메모리에 로딩하도록 특징점 추출 시스템을 제어할 수 있다. 이때, 프로세서는 특징맵 생성 모듈(110) 및 특징맵 변환 모듈(120) 각각은 메모리에 로딩된 프로그램 코드 중 대응하는 부분의 명령을 실행하여 이후 단계들(210 내지 220)을 실행하기 위한 프로세서의 서로 다른 기능적 표현들일 수 있다.
- [0029] 단계(210)에서 특징맵 생성 모듈(110)은 이미지 쌍으로부터 추출된 특징 각각에 대해 각각의 특징맵을 획득할 수 있다. 특징맵 생성 모듈(110)은 특징맵 생성 네트워크를 통해 이미지 쌍으로부터 추출된 특징 각각에 대해 설명자와, 신뢰성 맵과 반복성 맵을 포함하는 지점 추출 특징맵을 출력할 수 있다. 특징맵 생성 모듈(110)은 출력된 설명자와, 신뢰성 맵과 반복성 맵을 이용하여 반복성 및 신뢰성에 대한 손실 함수를 계산할 수 있다. 특징맵 생성 모듈(110)은 제1 이미지의 각 픽셀에서의 로컬 설명자와 제2 이미지의 각 픽셀에서의 로컬 설명자를 이용하여 신뢰성 점수를 예측하고, 예측된 신뢰성 점수를 이용하여 미분 가능한 근사치로 최적화된 신뢰성 손실을 계산할 수 있다. 이때, 제2 이미지는 제1 이미지에서 다른 광도 조건을 가진 이미지를 의미할 수 있다. 특징맵 생성 모듈(110)은 이미지 쌍의 특징 쌍 간의 피크 예측 및 유사성을 이용하여 반복성 손실을 계산할 수 있다.
- [0030] 단계(220)에서 특징맵 변환 모듈(120)은 구조적 정보와 스타일 정보를 이용하여 각각의 획득된 특징맵으로부터 영역 예측을 최소화하기 위한 특징점을 추출할 수 있다. 특징맵 변환 모듈(120)은 스타일 및 구조적 정보를 이용하여 각각의 획득된 특징맵을 변환하고 조작하기 위해 설명자를 사용하여 변환 손실을 계산할 수 있다. 특징맵 변환 모듈(120)은 제1 이미지의 설명자 및 제2 이미지의 설명자로부터 대체된 각각의 표준화된 특징맵에 의해 생성된 제1 공분산 행렬 및 제2 공분산 행렬을 획득하고, 획득된 제1 공분산 행렬 및 획득된 제2 공분산 행렬의 차이를 계산하여 광도 변환에 대한 공분산의 민감도를 의미하는 행렬을 생성할 수 있다.
- [0031] 도 3은 일 실시예에 있어서, 특징맵 생성 모듈의 네트워크 아키텍처를 설명하기 위한 도면이다.

- [0032] 특징맵 생성 모듈은 먼저 이미지 쌍 I_1 및 I_2 의 특징을 독립적으로 추출할 수 있으며, 설명자와 지점 추출 특징맵을 포함하는 두 가지의 분기를 출력할 수 있다. 지점 추출 분기는 두 개의 특징 맵으로 구성될 수 있다. 하나는 1×1 컨볼루션 레이어로 다른 하나를 생성한다. 전자는 신뢰성 맵 S이고 후자는 반복성 맵 R이다. 설명자 맵 X에서 파생된 공분산 행렬은 스타일 및 구조적 정보와 함께 중요 영역에 초점을 맞추도록 특징맵을 변환하는 데 사용된다. 그런 다음 특징맵 생성 모듈은 설명자 맵 X, 신뢰성 맵 S, 반복성 맵 R을 사용하여 반복성 및 신뢰성에 대한 손실 함수를 계산할 수 있다.
- [0033] 도 3을 참고하면, 특징맵 생성 모듈의 네트워크 아키텍처에 대하여 보다 상세하게 설명하기로 한다. 특징맵 생성 모듈과 특징맵 변환 모듈로 구성된 네트워크를 나타낸 것이다. 신뢰성 맵 S와 반복성 맵 R은 관심 지점의 영역을 학습하는 반면, 설명자 맵 X에서 파생된 공분산 행렬은 스타일 및 구조적 정보를 사용하여 현재성에 초점을 맞추기 위해 특징맵을 변환하는 데 사용될 수 있다.
- [0034] 깊이별 분리 가능 컨볼루션(depthwise separable convolution; DSC)은 새로운 표현으로 계산과 모델 크기를 크게 줄이는 인수 분해 컨볼루션(factorization convolution) 방법이다. 이에 동기를 부여하여 깊이별 분리 가능 컨볼루션을 활용하여 돌출 영역에 집중할 수 있다. 이미지 쌍에서 특징 정보를 추출하기 위해 백본 네트워크에서 마지막 레이어가 3개의 연속된 레이어 셋으로 대체되는 수정된 L2Net이 채택될 수 있다. 입력 이미지는 이미지 쌍이므로 두 개의 백본 네트워크가 필요하다. 모델 가중치를 줄이기 위해 백본 네트워크 뒤에 추가할 때 깊이별 분리 가능 컨볼루션 레이어에서 가중치 공유를 사용한다. 또한 가중치를 공유함으로써 설명자와 관심 지점 간의 관계를 유지할 수 있다. 그런 다음 백본 네트워크는 세 개의 특징 맵을 생성하는데, ℓ_2 정규화에 의한 설명자 맵 X, 요소별 제곱에 의한 신뢰성 맵 S, 1×1 컨볼루션 레이어를 사용하여 신뢰성 맵 S에서 얻은 반복성 맵 R을 포함하는 세 개의 특징맵을 생성할 수 있다. 대조적으로, 신뢰성 맵 S와 반복성 맵 R은 서로 의존하기 때문에 동일한 분기에서 나온다. 신뢰성 맵 S가 반복성 맵 R에서 높은 반복률에 영향을 미칠 수 있는 매칭 거리를 줄이는 지점을 학습하는 특징맵이라고 가정한다. 따라서, 정보 공유를 통해 모델 가중치가 더 가벼워지고 더 강력한 특징이 개발될 수 있다.
- [0035] 특징맵 생성 모듈은 구별되는 특징점을 획득하기 위해 신뢰성 손실 $\mathcal{L}_{reli}$ 를 계산할 수 있다. X_{ij} 를 이미지 I_1 의 각 픽셀(i, j)의 로컬 설명자로 지정하기로 한다. 그런 다음 X_{ij} 와 X'_{uv} 에서 개별 신뢰성 점수 S_{ij} 가 예측될 수 있다. 여기서 각 픽셀 (i, j)에 해당하는 정확한 좌표(u, v)를 지정하고, 실측(ground truth) 대응 매핑 T를 알고있다. 여기서 $T \in \mathbb{R}^{H \times W \times 2}$ 는 이미지 I_1 과 I_2 사이의 실측 대응이다. X_{ij} 는 X'_{uv} 와 비교되는데, 여기서 X'_{uv} 는 I_2 에서 추출될 수 있다. 그런 다음 평균 정밀도는 개별 신뢰성 점수 S_{ij} 를 사용하여 미분 가능한 근사치로 최적화된 $\mathcal{L}_{reli}$ 를 계산하는 데 사용된다.
- [0036] 또한 특징맵 생성 모듈은 반복 가능한 특징점을 추출하기 위해 반복성 손실인 $\mathcal{L}_{repeated}$ 을 계산할 수 있다. 입력 쌍 이미지의 특징 쌍 간 피크 예측 및 유사성이 사용될 수 있다. 유사성을 위해, R과 R'을 I_1 과 I_2 에 해당하는 반복성 맵이라고 하자. R'은 이미지 쌍 I_1 과 I_2 사이의 실측 호모그래피 관계에 의해 R'이 변환되는 맵으로 설정될 수 있다. 주요 목표는 반복성이 높은 키포인트를 예측하는 것이기 때문에, 반복성 맵 R에서 로컬 최대값의 위치가 시점 구조나 빛 이동과 같은 실제 그림 변환에 공변량이 되도록 네트워크를 훈련시킬 수 있다. 반복성 맵 R의 모든 로컬 최대치가 R'의 로컬 최대치와 일치한다고 가정하여 손실 함수가 정의될 수 있다. 기본 개념은 두 히트맵이 동일하고 최대치가 일치되도록 되도록 R과 R' 사이의 코사인 유사성을 극대화하는 것이다. 손실은 학습 프로세스를 종료할 수 있는 특정 상수에 남아 있을 수 있으므로 피크 예측을 사용하여 이를 방지한다. 최종 반복성 손실은 입력 이미지 쌍의 유사성과 피크성을 모두 고려하여 계산될 수 있다.
- [0037] 도 4는 일 실시예에 있어서, 특징맵 변환 모듈에서 특징맵을 변환하는 동작을 설명하기 위한 도면이다.
- [0038] 특징 매칭의 한 가지 문제는 구조적 관점과 조명(스타일) 차이이다. 스타일은 날씨와 빛과 같은 소음에 해당한다. 이미지의 관점 또는 엡지와 관련된 구조에 집중하면 더 나은 결과가 획득될 수 있다. 특징맵 변환 모듈은 작업 적응을 수행하고 기존의 WT 손실을 광범위한 변환으로 보완하여 스타일 정보를 억제할 수 있다. 특징맵 변환 모듈은 기존의 스타일만 사용하는 한계를 극복하기 위해 구조적 정보를 포함하여 이전 손실을 개선할 수 있다.

[0039] 우선적으로, 스타일/구조 공분산 행렬에 대하여 설명하기로 한다. 스타일 변환의 각 인스턴스에 WT를 적용하면 스타일 정보를 성공적으로 지울 수 있다고 주장되어 왔다. WT는 각 채널의 분산 항을 1로 균등화 하고, 채널 간의 공분산을 0으로 줄이는 선형 변환이다. 중간 특징맵은 $X \in \mathbb{R}^{C \times HW}$ 이며, 여기서 C는 채널 수이고, H와 W는 각각 특징맵의 높이와 너비이다. 채널 쌍 간의 공분산은 다음과 같이 정의될 수 있다.

[0040] 수학적식 1:

$$\Sigma = \frac{1}{HW} (\mathbf{X} - \mu \cdot \mathbf{O}^\top) (\mathbf{X} - \mu \cdot \mathbf{O}^\top)^\top \in \mathbb{R}^{C \times C}$$

[0041]

[0042] 여기서, $\mathbf{O} \in \mathbb{R}^{HW}$ 는 하나의 열 벡터이고, μ 와 Σ 는 각각 평균 벡터와 공분산 행렬이다. 공분산 행렬 Σ 의 요소가 감소하도록 손실 함수를 설계할 때, 입력 이미지에서 특징을 추출할 때 스타일 요소의 영향을 덜 받는다. 특징맵 생성에는 스타일 요소의 정보가 부족하기 때문이다. 이러한 개념을 기반으로 스타일 요소를 사용하여 특징맵을 변환하는 방법이 채택될 수 있다. 또한, 스타일 요소 옆에 있는 그램 행렬의 잔여 요소인 구조에 대한 정보를 사용하여 억제보다는 확장 방향으로 손실 함수가 설계될 수 있다.

[0043] 다음으로, 변환 손실에 대하여 설명하기로 한다. 앞서 언급한 스타일 및 구조 정보를 사용하여 특징맵을 변환하고 조작하기 위해 특징맵 X를 사용하여 $\mathcal{L}_{cov}$ 를 계산하는 특징맵 변환 모듈에 대하여 설명하기로 한다. 설명자가 특징 생성 모듈에서 출력된 세 가지의 특징맵 중에서 가장 많은 정보를 가지고 있기 때문이다. 특징맵 변환 모듈은 각 속성을 다르게 선택적으로 수정하여 특징 표현의 고차 통계로부터 스타일 및 구조 특성을 분리시킬 수 있다. 기존의 특징맵 X를 표준화된 특징맵 X_s 로 대체하면 공분산 행렬의 대각 및 비대각 요소의 최적화 프로세스가 동시에 간소화될 수 있다. Σ_s 는 표준화된 특징맵 X_s 에 의해 생성된 행렬이며, 다음과 같이 Σ_s 가 정의될 수 있다.

[0044] 수학적식 2:

$$\Sigma_s = \frac{1}{HW} \mathbf{X}_s \cdot \mathbf{X}_s^\top \in \mathbb{R}^{C \times C}$$

[0045]

[0046] 특징맵 변환 모듈은 공분산 행렬 $\Sigma_s(I_1)$ 과 $\Sigma_s(I_2)$ 를 얻은 후, 두 행렬 사이의 차이를 계산하여 수학적식 3에 정의된 것과 같이 행렬 Σ_C 를 생성할 수 있다. 행렬 Σ_C 는 광도 변환에 대한 해당 공분산의 민감도를 나타낸다. 분산 값이 높은 요소는 스타일 정보를 유지하는 반면 분산 값이 낮은 요소는 구조 정보를 유지한다. 이때, 수직 바를 써서 절대값 표기법이 사용될 수 있다.

[0047] 수학적식 3:

$$\Sigma_C = |\Sigma_s(I_1) - \Sigma_s(I_2)|$$

[0048]

[0049] 두 구성 요소를 분리하기 위한 임계값을 결정하기 위해 평균 값을 사용하여 스타일 및 구조적 구성 요소를 동일한 양으로 클러스터링할 수 있다. 행렬 Σ_C 요소가 임계값보다 크면 해당 요소는 스타일 요인으로 분류되고 나머지 요소는 구조적 요인으로 분류될 수 있다. 이러한 정의는 행렬 Σ_C 의 주요 요인이 스타일 요인을 의미한다고 가정하기 때문에 확립될 수 있다. 이 경우 스타일 요인은 빛이나 색상의 변화를 의미하며, 구조적 요인은 많은 객체, 엣지 또는 시점으로 복잡성을 의미한다. 제안된 손실 함수 $\mathcal{L}_{cov}$ 는 다음과 같이 공식화될 수 있다.

[0050] 수학적식 4:

$$\mathcal{L}_{cov} = \begin{cases} \mathbb{E} [\|\Sigma_C \odot \mathbf{M}_{sty}\|_1] & \text{if } \Sigma_C > \mu \\ 1 - \mathbb{E} [\|\Sigma_C \odot \mathbf{M}_{str}\|_1] & \text{otherwise} \end{cases}$$

[0051]

[0052] 여기서, $\mathbf{E}$ and $\mathbf{M} \in \mathbb{R}^{C \times C}$ 는 산술 평균과 마스크 행렬이다. $\mathbf{M}_{sty}$ 및 $\mathbf{M}_{str}$ 은 스타일 및 구조 값을 선택하는 마스크이며 $\odot$ 는 요소별 곱이다. 마지막으로, 총 손실 함수는 수학적식 5로 표현될 수 있으며, 각 가중치 λ_i 는 경험적으로 최적 비율인 1:1:2로 조정될 수 있다. 손실 항의 수가 3개이기 때문에 $i \in \{1, 2, 3\}$ 로 정의될 수 있다. 도 2에 도시된 바와 같이 종래의 R2D2의 특징맵과 실시예에서 제안된 특징맵을 비교할 때, 변환 손실을 추가하는 것이 중요 특징만 선택하는 데 우수하다는 것을 확인할 수 있다.

[0053] 수학적식 5:

$$\mathcal{L}_{total} = \lambda_1 \cdot \mathcal{L}_{reli} + \lambda_2 \cdot \mathcal{L}_{repeat} + \lambda_3 \cdot \mathcal{L}_{cov}$$

[0054]

[0055] 집계 손실 함수는 하늘 또는 지면과 같이 정보가 적은 영역의 예측을 최소화하기 위해 중요 지점을 선택하는 데 사용될 수 있다. 특징맵 변환 모듈은 구조와 스타일의 변화에 관계없이 이미지를 촬영하는 위치가 동일한 장소에 있을 경우 제안된 변환 손실 함수가 강건한 특징을 추출할 수 있도록 특징맵 변환을 유도할 수 있다.

[0056] 도 4를 참고하면, 각 네트워크의 특징맵을 비교하면 특징맵을 변경하는 경향을 알 수 있다. 스타일 정보를 제거하면 노이즈가 발생하고 구조적 복잡성이 커진다.

[0057] 일 실시예에 따르면, 로컬 특징의 반복성에 기여하는 특징맵을 변환하는 자체 변환 손실을 사용하여 강력한 네트워크를 제공할 수 있다. 또한, 공분산 벡터를 클러스터링하여 구조와 스타일 특성을 분리하고 각 특징의 특징에 영향을 미칠 수 있다. 또한, 조명 구성 요소를 줄이고 구조적 구성 요소를 선명하게 하여 두 개의 특징맵을 더 가깝게 생성하기 위해 특징맵을 통합시킬 수 있다. 결과적으로 이러한 단계는 지점의 반복성을 높이고 이상값을 줄이며 강력하게 일치하는 설명자를 할당할 수 있다. 이에, 다양한 장면에서 강인한 매칭 지점을 더 효과적으로 추출할 수 있다.

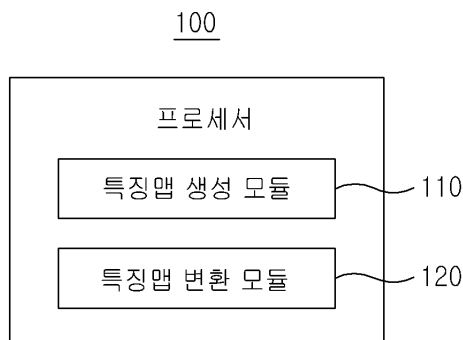
[0058] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0059] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

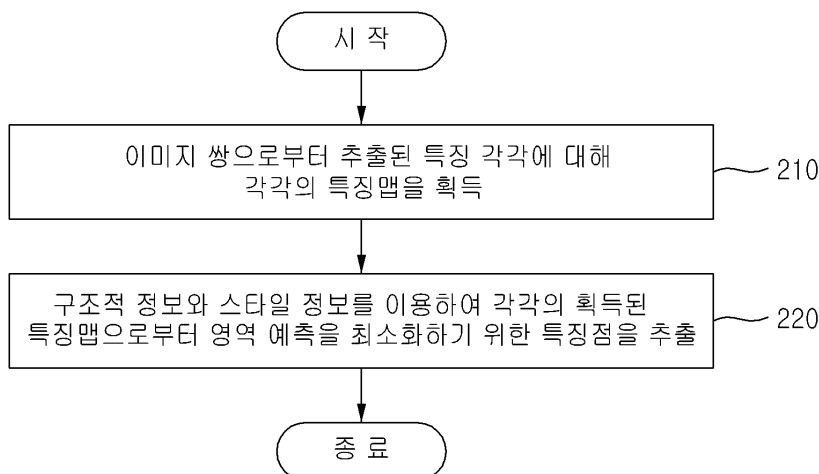
- [0060] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0061] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [0062] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

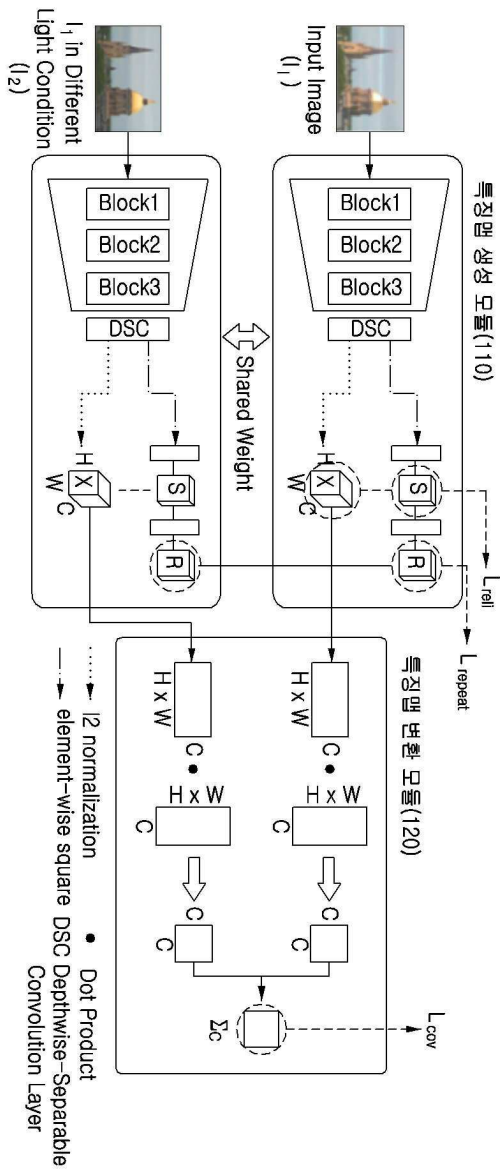
도면1



도면2



도면3



도면4

