



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0076660
(43) 공개일자 2021년06월24일

(51) 국제특허분류(Int. Cl.)

G06N 3/08 (2006.01)

(52) CPC특허분류

G06N 3/08 (2013.01)

H04N 13/00 (2018.05)

(21) 출원번호 10-2019-0168080

(22) 출원일자 2019년12월16일

심사청구일자 2019년12월16일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상훈

서울특별시 강남구 삼성로 151, 3동 504호(대치동, 선경아파트)

이성민

서울특별시 서대문구 신촌로3나길 2, 201호(창천동)

(74) 대리인

특허법인우인

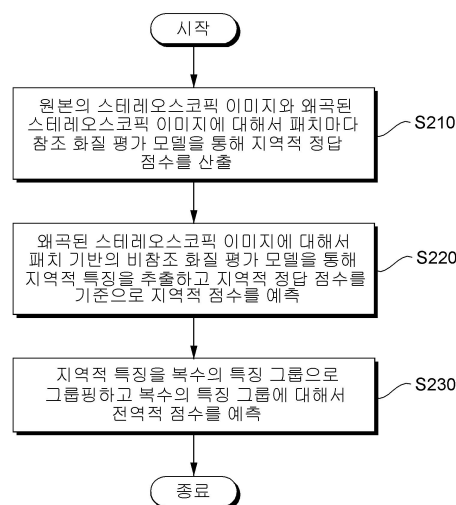
전체 청구항 수 : 총 5 항

(54) 발명의 명칭 **합성곱 신경망 기반의 스테레오스코픽 이미지 화질 평가 방법 및 장치**

(57) 요약

본 실시예들은 스테레오스코픽 이미지에 대해서 패치 기반의 신경망을 사용하여 지역적 특징을 추출하고, 지역적 특징을 복수의 특징 그룹으로 그룹핑하여 전역적 점수를 예측하여, 스테레오스코픽 이미지의 화질을 평가하는 장치 및 방법을 제공한다.

대표도 - 도2



(52) CPC특허분류

H04N 17/00 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711080860
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	첨단융복합콘텐츠기술개발사업-융합형콘텐츠핵심응용기술개발사업
연구과제명	VR 멀미 저감을 위한 휴먼팩터 파라미터 제어기술 개발(표준화연계)
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2017.03.01 ~ 2019.12.31

명세서

청구범위

청구항 1

컴퓨팅 디바이스에 의한 스테레오스코픽 이미지의 화질 평가 방법에 있어서,

원본의 스테레오스코픽 이미지와 왜곡된 스테레오스코픽 이미지에 대해서 패치마다 참조 화질 평가 모델을 통해 지역적 정답 점수를 산출하는 단계;

상기 왜곡된 스테레오스코픽 이미지에 대해서 패치 기반의 비참조 화질 평가 모델을 통해 지역적 특징을 추출하고 상기 지역적 정답 점수를 기준으로 지역적 점수를 예측하는 단계;

상기 지역적 특징을 복수의 특징 그룹으로 그룹핑하고 상기 복수의 특징 그룹에 대해서 전역적 점수를 예측하는 단계

를 포함하는 스테레오스코픽 이미지의 화질 평가 방법.

청구항 2

제1항에 있어서,

상기 지역적 정답 점수를 산출하는 단계는,

스테레오스코픽 이미지(Stereoscopic Image)를 좌우 합성 이미지(Cyclopean Image)로 변환하며, 상기 원본의 스테레오스코픽 이미지를 변환한 기준 좌우 합성 이미지와 상기 왜곡된 스테레오스코픽 이미지를 변환한 왜곡된 좌우 합성 이미지를 패치로 분할하여 상기 참조 화질 평가 모델에 적용하는 것을 특징으로 하는 스테레오스코픽 이미지의 화질 평가 방법.

청구항 3

제1항에 있어서,

상기 지역적 정답 점수를 산출하는 단계는,

이미지의 픽셀마다 상기 참조 화질 평가 모델을 적용하여 산출한 결과 점수를 평균을 내어 상기 패치마다 지역적 정답 점수를 산출하는 것을 특징으로 하는 스테레오스코픽 이미지의 화질 평가 방법.

청구항 4

제1항에 있어서,

상기 전역적 점수를 예측하는 단계는,

(i) 상기 지역적 특징에 대해서 평균을 산출한 제1 특징 그룹, (ii) 상기 지역적 특징에 대해서 분산을 산출한 제2 특징 그룹, (iii) 상기 지역적 특징의 히스토그램에서 상위 비율의 평균을 산출한 제3 특징 그룹, 및 (iv) 상기 지역적 특징의 히스토그램에서 하위 비율의 평균을 산출한 제4 특징 그룹을 전부 결합한 통합 특징 그룹에 대해서 전역적 점수를 예측하는 것을 특징으로 하는 스테레오스코픽 이미지의 화질 평가 방법.

청구항 5

하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 스테레오스코픽 이미지의 화질 평가 장치에 있어서,

상기 프로세서는 원본의 스테레오스코픽 이미지와 왜곡된 스테레오스코픽 이미지에 대해서 패치마다 참조 화질 평가 모델을 통해 지역적 정답 점수를 산출하고,

상기 프로세서는 상기 왜곡된 스테레오스코픽 이미지에 대해서 패치 기반의 비참조 화질 평가 모델을 통해 지역적 특징을 추출하고 상기 지역적 정답 점수를 기준으로 지역적 점수를 예측하고,

상기 프로세서는 상기 지역적 특징을 복수의 특징 그룹으로 그룹핑하고 상기 복수의 특징 그룹에 대해서 전역적 점수를 예측하는 것을 특징으로 하는 스테레오스코픽 이미지의 화질 평가 장치.

발명의 설명

기술 분야

[0001] 본 발명이 속하는 기술 분야는 스테레오스코픽 이미지의 화질 평가 방법 및 장치에 관한 것이다.

배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 스테레오스코픽 이미지는 양안 시차를 활용해서 시청자가 두 눈의 각도 차이를 통해 3차원 이미지를 느낄 수 있도록 하는 입체 이미지이다.

[0004] 기존의 2차원 이미지의 화질 평가 모델을 스테레오스코픽 이미지의 화질 평가에 적용하기에는 크게 3 가지 문제점이 있다.

[0005] 첫 번째로 스테레오스코픽 이미지에 관한 학습 데이터가 부족하고, 두 번째로 이미지의 지역적인 평가 점수가 존재하지 않고, 세 번째로 스테레오스코픽 이미지를 화질 평가할 때 일반적인 이미지 분류 모델을 그대로 사용할 수 없다는 점이다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국등록특허공보 제10-1393621호 (2014.05.02)

발명의 내용

해결하려는 과제

[0007] 본 발명의 실시예들은 스테레오스코픽 이미지에 대해서, 패치 기반의 신경망을 사용하여 지역적 특징을 추출하고, 지역적 특징을 복수의 특징 그룹으로 그룹핑하여 전역적 점수를 예측함으로써, 스테레오스코픽 이미지의 화질을 평가하는 데 주된 목적이 있다.

[0008] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

과제의 해결 수단

[0009] 본 실시예의 일 측면에 의하면, 컴퓨팅 디바이스에 의한 스테레오스코픽 이미지의 화질 평가 방법에 있어서, 원본의 스테레오스코픽 이미지와 왜곡된 스테레오스코픽 이미지에 대해서 패치마다 참조 화질 평가 모델을 통해 지역적 정답 점수를 산출하는 단계, 상기 왜곡된 스테레오스코픽 이미지에 대해서 패치 기반의 비참조 화질 평가 모델을 통해 지역적 특징을 추출하고 상기 지역적 정답 점수를 기준으로 지역적 점수를 예측하는 단계, 상기 지역적 특징을 복수의 특징 그룹으로 그룹핑하고 상기 복수의 특징 그룹에 대해서 전역적 점수를 예측하는 단계를 포함하는 스테레오스코픽 이미지의 화질 평가 방법을 제공한다.

[0010] 상기 지역적 정답 점수를 산출하는 단계는, 스테레오스코픽 이미지(Stereoscopic Image)를 좌우 합성 이미지(Cyclopean Image)로 변환하며, 상기 원본의 스테레오스코픽 이미지를 변환한 기준 좌우 합성 이미지와 상기 왜곡된 스테레오스코픽 이미지를 변환한 왜곡된 좌우 합성 이미지를 패치로 분할하여 상기 참조 화질 평가 모델에 적용할 수 있다.

[0011] 상기 지역적 정답 점수를 산출하는 단계는, 이미지의 픽셀마다 상기 참조 화질 평가 모델을 적용하여 산출한 결과 점수를 평균을 내어 상기 패치마다 지역적 정답 점수를 산출할 수 있다.

[0012] 상기 전역적 점수를 예측하는 단계는, (i) 상기 지역적 특징에 대해서 평균을 산출한 제1 특징 그룹, (ii) 상기

지역적 특징에 대해서 분산을 산출한 제2 특징 그룹, (iii) 상기 지역적 특징의 히스토그램에서 상위 비율의 평균을 산출한 제3 특징 그룹, 및 (iv) 상기 지역적 특징의 히스토그램에서 하위 비율의 평균을 산출한 제4 특징 그룹을 전부 결합한 통합 특징 그룹에 대해서 전역적 점수를 예측할 수 있다.

[0013] 본 실시예의 다른 측면에 의하면, 하나 이상의 프로세서 및 상기 하나 이상의 프로세서에 의해 실행되는 하나 이상의 프로그램을 저장하는 메모리를 포함하는 스테레오스코픽 이미지의 화질 평가 장치에 있어서, 상기 프로세서는 원본의 스테레오스코픽 이미지와 왜곡된 스테레오스코픽 이미지에 대해서 패치마다 참조 화질 평가 모델을 통해 지역적 정답 점수를 산출하고, 상기 프로세서는 상기 왜곡된 스테레오스코픽 이미지에 대해서 패치 기반의 비참조 화질 평가 모델을 통해 지역적 특징을 추출하고 상기 지역적 정답 점수를 기준으로 지역적 점수를 예측하고, 상기 프로세서는 상기 지역적 특징을 복수의 특징 그룹으로 그룹핑하고 상기 복수의 특징 그룹에 대해서 전역적 점수를 예측하는 것을 특징으로 하는 스테레오스코픽 이미지의 화질 평가 장치를 제공할 수 있다.

발명의 효과

[0014] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 스테레오스코픽 이미지에 대해서, 패치 기반의 신경망을 사용하여 지역적 특징을 추출하고, 지역적 특징을 복수의 특징 그룹으로 그룹핑하여 전역적 점수를 예측함으로써, 스테레오스코픽 이미지의 화질을 평가할 수 있는 효과가 있다.

[0015] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 스테레오스코픽 이미지의 화질 평가 장치를 예시한 블록도이다.

도 2는 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법을 예시한 흐름도이다.

도 3은 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법에 적용된 비참조 화질 평가 모델을 예시한 도면이다.

도 4는 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법이 산출한 화질 맵을 예시한 도면이다.

도 5는 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법이 지역적 특징을 추출하고 지역적 점수를 예측하는 것을 예시한 도면이다.

도 6은 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법이 특징 그룹에 대해서 전역적 점수를 예측하는 것을 예시한 도면이다.

도 7은 본 발명의 실시예들에 따라 수행된 모의실험 결과를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0017] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.

[0018] 기존의 컴퓨터 비전 처리 방식과 달리 스테레오스코픽 이미지의 화질을 평가 방식에 딥 러닝을 적용하면 3 가지 문제가 있다.

[0019] 첫 번째로 스테레오스코픽 이미지에 관한 훈련 데이터가 부족하다.

[0020] 이미지 분류에 사용되는 ImageNet 이미지 데이터베이스는 클래스 레이블과 매칭된 1500만 개 이상의 이미지를 포함한다. 기존의 이미지 데이터베이스에 저장된 레이블은 의미적으로 명확하므로 검증자료(Ground Truth)를 획득하는 것이 용이하다.

[0021] 반면에 스테레오스코픽 이미지의 화질 평가를 위한 데이터베이스는 1000 개의 이미지를 초과하지 않는다. LIVE 데이터베이스는 5 가지 왜곡 유형에 대해 365 개의 왜곡된 이미지만 있고 SA 데이터베이스는 9 가지 왜곡 유형에 대해 360 개의 이미지만 보유한 실정이다.

- [0022] 두 번째로 이미지의 지역적인 평가 점수가 존재하지 않는다.
- [0023] 주관적 점수를 획득하기 위해 심리적 테스트를 수행해야 하므로, 검증자료를 획득하는 것이 쉽지 않다. 이미지를 분할하는 패치 방식을 적용하더라도 패치마다 주관적 점수가 필요한데 검증된 데이터베이스가 존재하지 않는다. 전체 이미지에 대한 주관적 점수만 확보하고 지역적 검증자료가 없는 실정이다. 전체 이미지에 대한 주관적 점수의 평균은 각 패치의 점수와 상이하다. 이미지의 각 패치에서의 점수는 균일하지 않다. 예를 들어, 공원 사진에서 중간 부분의 도로 영역이 심하게 왜곡되면, 이 영역의 지역적 화질이 하나의 이미지의 화질 점수보다 낮을 수 있다. 신뢰할만한 지역적 검증자료를 적용하지 않고 패치 기반의 지도 학습을 수행할 수 없다.
- [0024] 세 번째로 스테레오스코픽 이미지를 화질 평가할 때 일반적인 이미지 분류 모델을 그대로 사용할 수 없다.
- [0025] 이미지 인식 모델은 이미지에서 왜곡이 발생하더라도 대상 물체를 감지해야 하므로 왜곡에 강한 특징을 학습해야 한다. 이와 달리 화질 평가 모델은 왜곡 정보를 민감하게 포착해야 하므로, 왜곡에 민감한 특징을 추출해야 한다.
- [0026] 본 실시예에 따른 스테레오스코픽 이미지의 화질 평가 장치는 딥 러닝을 적용하여 비참조 방식으로 스테레오스코픽 이미지의 화질을 평가한다.
- [0027] 스테레오스코픽 이미지의 화질 평가 장치는 훈련 데이터의 부족을 극복하기 위해 패치 기반의 접근 방식을 사용한다. 신뢰할 수 있는 지역적 주관적 점수가 없는 문제를 해결하기 위해, 패치 기반의 학습 및 참조 방식의 평가 모델을 통해 지역적 패치에 대한 슈도 검증자료를 획득한다.
- [0028] 도 1은 본 발명의 일 실시예에 따른 스테레오스코픽 이미지의 화질 평가 장치를 예시한 블록도이다.
- [0029] 스테레오스코픽 이미지의 화질 평가 장치(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.
- [0030] 프로세서(120)는 스테레오스코픽 이미지의 화질 평가 장치(110)로 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 스테레오스코픽 이미지의 화질 평가 장치(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0031] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독한 가능 저장 매체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 스테레오스코픽 이미지의 화질 평가 장치(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0032] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 스테레오스코픽 이미지의 화질 평가 장치(110)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0033] 스테레오스코픽 이미지의 화질 평가 장치(110)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치는 입출력 인터페이스(150)를 통해 스테레오스코픽 이미지의 화질 평가 장치(110)의 다른 컴포넌트들에 연결될 수 있다.
- [0034] 스테레오스코픽 이미지의 화질 평가 장치(110)는 원본의 스테레오스코픽 이미지와 왜곡된 스테레오스코픽 이미지에 대해서 패치마다 참조 화질 평가 모델을 통해 지역적 정답 점수를 산출한다.
- [0035] 스테레오스코픽 이미지의 화질 평가 장치(110)는 왜곡된 스테레오스코픽 이미지에 대해서 패치 기반의 비참조 화질 평가 모델을 통해 지역적 특징을 추출하고 지역적 정답 점수를 기준으로 지역적 점수를 예측한다.
- [0036] 스테레오스코픽 이미지의 화질 평가 장치(110)는 지역적 특징을 복수의 특징 그룹으로 그룹핑하고 복수의 특징 그룹에 대해서 전역적 점수를 예측한다.
- [0037] 도 2는 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법을 예시한 흐름도이다. 스테레오

스코픽 이미지의 화질 평가 방법은 스테레오스코픽 이미지의 화질 평가 장치 또는 컴퓨팅 디바이스 등에 의해 수행될 수 있다.

- [0038] 단계 S210에서 프로세서는 원본의 스테레오스코픽 이미지와 왜곡된 스테레오스코픽 이미지에 대해서 패치마다 참조 화질 평가 모델을 통해 지역적 정답 점수를 산출한다.
- [0039] 단계 S220에서 프로세서는 왜곡된 스테레오스코픽 이미지에 대해서 패치 기반의 비참조 화질 평가 모델을 통해 지역적 특징을 추출하고 지역적 정답 점수를 기준으로 지역적 점수를 예측한다.
- [0040] 단계 S230에서 프로세서는 지역적 특징을 복수의 특징 그룹으로 그룹핑하고 복수의 특징 그룹에 대해서 전역적 점수를 예측한다.
- [0041] 도 3은 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법에 적용된 비참조 화질 평가 모델을 예시한 도면이다.
- [0042] 본 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법에 적용된 화질 평가 모델은 2 가지의 훈련 단계를 수행한다. 첫 번째 단계는 참조 화질 평가 모델을 통해 획득한 슈도 검증자료를 사용하여 지역적 품질 점수를 예측한다. 두 번째 단계는 첫 번째 단계에서 학습한 모델을 사용하여 전역적 주관적 점수를 예측한다.
- [0043] 먼저 스테레오스코픽 이미지의 화질 평가 장치 방법은 참조 화질 평가 모델을 통해 슈도 검증자료를 생성한다.
- [0044] 도 4는 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법이 산출한 화질 맵을 예시한 도면이다.
- [0045] 참조 화질 평가 모델에 기준 이미지와 왜곡된 이미지를 입력한다. 참조 화질 평가 모델은 화질 맵을 생성한다. 화질 맵은 64x36 패치로 나뉘질 수 있으며, 각 패치의 평균을 취하여 지역적 점수를 획득한다. 각 패치에 대해 슈도 검증자료를 획득하고, 지역적 화질 점수 맵을 생성한다. 화질 맵을 생성하는데 구조적 유사도 지수 (Structural Similarity Index, SSIM)를 적용할 수 있다.
- [0046] 프로세서가 지역적 정답 점수를 산출하는 단계는 스테레오스코픽 이미지(Stereoscopic Image)를 좌우 합성 이미지(Cyclopean Image)로 변환하며, 원본의 스테레오스코픽 이미지를 변환한 기준 좌우 합성 이미지와 왜곡된 스테레오스코픽 이미지를 변환한 왜곡된 좌우 합성 이미지를 패치로 분할하여 참조 화질 평가 모델에 적용한다.
- [0047] 좌우 합성 이미지를 획득하는 방식은 수학적 식 1과 같이 표현된다.

수학적 식 1

$$I_c(x, y) = W_l(x, y) \cdot I_l(x, y) + W_r(x + d, y) \cdot I_r(x + d, y)$$

- [0048]
- [0049] d 는 좌우 이미지 I_l 과 I_r 의 픽셀 시차이고, W_l 과 W_r 은 가보 필터 응답에 의한 정규화된 가중치를 나타낸다.
- [0050] 왜곡된 스테레오스코픽 이미지에 대한 좌우 합성 이미지 쌍 I_c 는 매핑 방식으로 합성된다. 왜곡된 스테레오스코픽 이미지 쌍으로부터 신뢰할만한 시차 측정이 없으므로, 기준 이미지 쌍의 시차 맵이 좌우 합성 이미지 쌍에 사용된다.
- [0051] 기준 좌우 합성 이미지 I_c 및 왜곡된 좌우 합성 이미지 I_c' 간의 화질 맵 측정에 구조적 유사도 지수(Structural Similarity Index, SSIM)을 적용할 수 있다. 구조적 유사도 지수(SSIM)는 인간의 시각 체계가 이미지의 구조적 정보에 예민하다는 점을 이용하여 구조적 정보를 이용하여 이미지를 평가하는 대표적인 인간 시각 체계를 만족시키는 평가 기법이다.
- [0052] 왜곡된 스테레오스코픽 이미지에 대한 좌우 합성 이미지에서 획득된 분할된 패치 세트를 P_c 로 칭한다.
- $$P_c = \{p_{c1}, p_{c2}, \dots, p_{cN}\}$$
- N 은 패치의 전체 개수이다.
- [0053] 지역적 정답 점수를 산출하는 단계는 이미지의 픽셀마다 참조 화질 평가 모델을 적용하여 산출한 결과 점수를 평균을 내어 패치마다 지역적 정답 점수를 산출한다. 지역적 정답 점수는 수학적 식 2와 같이 표현된다.

수학식 2

$$\rho_n = \frac{1}{h_p \cdot w_p} \sum_{x,y \in \mathbf{p}_n} m_{SSIM}(x,y)$$

[0054]

[0055] h_p 및 w_p 는 수직 및 수평 패치 사이즈이다. 예컨대, h_p 는 16, w_p 는 18로 설정될 수 있다.

[0056] 프로세서는 비참조 화질 평가 모델에 입력된 데이터를 전처리할 수 있다. 왜곡된 좌우 이미지 I_l 과 I_r 은 정규화

를 거치고, 프로세서는 정규화된 이미지 세트 $\hat{\mathbf{I}}_l = \{\hat{I}_{l1}, \dots, \hat{I}_{lM}\}$ 및 $\hat{\mathbf{I}}_r = \{\hat{I}_{r1}, \dots, \hat{I}_{rM}\}$ 를 생성한다. M 은 왜곡된 스테레오스코픽 이미지 쌍의 전체 개수이다.

[0057] 각 이미지는 중첩되지 않는 패치로 분할된다. 패치 세트 $\mathbf{P}' = \{\{\mathbf{p}'_{l1}, \mathbf{p}'_{r1}\}, \dots, \{\mathbf{p}'_{lN}, \mathbf{p}'_{rN}\}\}$ 가 비참조 화질 평가 모델에 입력된다.

[0058] 도 5는 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법이 지역적 특징을 추출하고 지역적 점수를 예측하는 것을 예시한 도면이다.

[0059] 패치 기반의 비참조 화질 평가 모델이 수행하는 첫 번째 단계는 지역적 주관적 점수의 회귀 단계이다. 비참조 화질 평가 모델은 레이어가 연결된 네트워크이며 가중치 및 바이어스를 학습하는 모델이다. 비참조 화질 평가 모델은 CNN(Convolutional Neural Network) 등의 신경 네트워크로 구현될 수 있다.

[0060] CNN 모델은 각 패치에 대한 지역적 화질 점수를 훈련시키기 위해 적용된다. 2 개의 컨볼루션 레이어와 4 개의 풀리 커넥티드 레이어가 사용될 수 있다. 100 개의 차원을 갖는 추출된 특징 벡터를 사용하여 지역적 화질 점수는 회귀할 수 있다.

[0061] $g_\theta(\cdot)$ 는 파라미터 θ 에 의한 컨볼루션 레이어와 풀리 커넥티드 레이어를 갖는 특징 벡터 추출기이다.

$g_{\phi_1}(\cdot)$ 는 파라미터 ϕ_1 에 의한 회귀 함수이다. 100 개의 차원의 특징 벡터

$\mathbf{f}_n = (f_{n,1}, f_{n,2}, \dots, f_{n,100})$ 는 n 번째 패치 쌍으로부터 $g_\theta(\{\mathbf{p}'_{ln}, \mathbf{p}'_{rn}\})$ 에 의해 추출된다.

[0062] 각 패치 쌍 $\{\mathbf{p}'_{ln}, \mathbf{p}'_{rn}\}$ 은 모델에 독립적으로 입력되고 다른 패치와 상관 관계가 없다.

[0063] 네트워크의 파라미터 Θ_1 를 최적화하기 위한 목적 함수는 슈도 검증자료와 예측된 점수 간의 평균 제곱 오차를 사용할 수 있다.

수학식 3

$$\Theta_1^* = \arg \min_{\Theta_1} \ell_1(\{\mathbf{p}'_{ln}, \mathbf{p}'_{rn}\}, \rho_n; \Theta_1)$$

[0064]

[0065] $\ell_1(\cdot)$ 는 평균 제곱 오차로 정의된 손실 함수이다.

[0066] 도 6은 본 발명의 다른 실시예에 따른 스테레오스코픽 이미지의 화질 평가 방법이 특징 그룹에 대해서 전역적

점수를 예측하는 것을 예시한 도면이다.

[0067] 패치 기반의 비참조 화질 평가 모델이 수행하는 두 번째 단계는 전역적 주관적 점수의 회귀 단계이다.

[0068] 특징 벡터 세트 $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_{N_m})^T$ 는 $\mathbf{f}_n = g_\theta(\{\mathbf{p}'_{ln}, \mathbf{p}'_{rn}\})$ 에 의해 스테레오스코픽 이미지 쌍으로부터 획득된다.

[0069] 전역적 점수를 예측하는 단계는 (i) 지역적 특징에 대해서 평균을 산출한 제1 특징 그룹, (ii) 지역적 특징에 대해서 분산을 산출한 제2 특징 그룹, (iii) 지역적 특징의 히스토그램에서 상위 비율의 평균을 산출한 제3 특징 그룹, 및 (iv) 지역적 특징의 히스토그램에서 하위 비율의 평균을 산출한 제4 특징 그룹을 전부 결합한 통합 특징 그룹에 대해서 전역적 점수를 예측한다.

[0070] 특징 풀링 레이어는 $\mathbf{z} = \{z_1(\cdot), \dots, z_4(\cdot)\}$ 와 같이 4 가지의 풀링 함수를 결합한다. 각 통계적 풀링 함수는 패치 특징을 이미지 특징으로 변환($\mathbb{R}^{N_m} \rightarrow \mathbb{R}^1$)한다.

[0071] 훈련된 모델을 사용하여 각 패치에 대해 100 차원의 특징이 추출된다. 주어진 이미지가 N 개의 패치로 구성되면, 하나의 이미지에서 N x 100 개의 특징 벡터를 생성할 수 있다. 이미지 특징 벡터 $\bar{\mathbf{f}}_n = (\bar{f}_{n,1}, \dots, \bar{f}_{n,100})$ 는 각 풀링 함수에 의해 획득하고, 수학적 4 내지 수학적 7과 같이 표현된다.

[0072] 수학적 4는 평균 풀링, 수학적 5는 분산 풀링, 수학적 6은 높은 백분위 풀링, 수학적 7은 낮은 백분위 풀링을 표현하며, 4 가지 종류의 풀링을 사용하여 통합 특징 벡터를 획득한다.

수학적 4

$$\bar{\mathbf{f}}_1 = z_1(\mathbf{F}) = \frac{1}{N_m} \sum_{n=1}^{N_m} f_{n,l}$$

[0073]

[0074] l은 특징 인덱스이다.

수학적 5

$$\bar{\mathbf{f}}_2 = z_2(\mathbf{F}) = \frac{1}{N_m} \sum_{n=1}^{N_m} (f_{n,l} - \bar{f}_{1,l})^2$$

[0075]

수학적 6

$$\bar{\mathbf{f}}_3 = z_3(\mathbf{F}) = \frac{1}{N_m^p} \sum_{n > n^p+} f_{n,l}^h$$

[0076]

수학식 7

$$\bar{\mathbf{f}}_4 = z_4(\mathbf{F}) = \frac{1}{N_m^p} \sum_{n < n^p-} f_{n,l}^h$$

[0077]

[0078] n^{p+} 와 n^{p-} 는 패치 특징의 히스토그램 $f_{n,l}^h$ 에서 p 번째 비율의 상한과 하한이다. N_m^p 는 p 번째 비율의 품질이며, 예컨대 $N_m^p = N_m \cdot p/100$ 로 설정될 수 있다.

[0079] 이미지 특징 벡터 $\bar{\mathbf{F}} = (\bar{\mathbf{f}}_1, \dots, \bar{\mathbf{f}}_4)$ 는 $z()$ 에 의해 통합된다.

[0080] 폴링된 특징 벡터는 전역적 주관적 점수 $\hat{\rho}_m$ 를 예측하기 위해 회귀한다. 폴링된 특징 벡터의 가중치를 미세 조정하기 위해 역 전파 프로세스에 의해 모델이 학습된다.

[0081] 네트워크의 파라미터 Θ_2 를 최적화하기 위한 목적 함수는 주관적 점수와 예측된 점수 간의 평균 제곱 오차를 사용할 수 있다.

수학식 8

$$\Theta_2^* = \arg \min_{\Theta_2} \ell_2 \left(\left\{ \hat{I}_{lm}', \hat{I}_{rm}' \right\}, \hat{\rho}_m; \Theta_2 \right)$$

[0082]

[0083] $\ell_2(\cdot)$ 는 평균 제곱 오차로 정의된 손실 함수이다.

[0084] 도 7은 본 발명의 실시예들에 따라 수행된 모의실험 결과를 도시한 것이다.

[0085] 도 7에서 첫 번째 행은 왜곡된 이미지이고, 두 번째 행은 참조 방식의 화질 맵이고, 세 번째 행은 예측된 화질 맵을 나타낸다. 예측된 지역 화질 맵은 왜곡된 이미지에 대한 실제 SSIM 맵과 거의 유사하다는 것을 쉽게 확인할 수 있다.

[0086] 스테레오스코픽 이미지의 화질 평가 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.

[0087] 스테레오스코픽 이미지의 화질 평가 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스 또는 서버에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스 또는 서버는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.

[0088] 도 2에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 2에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.

[0089] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨

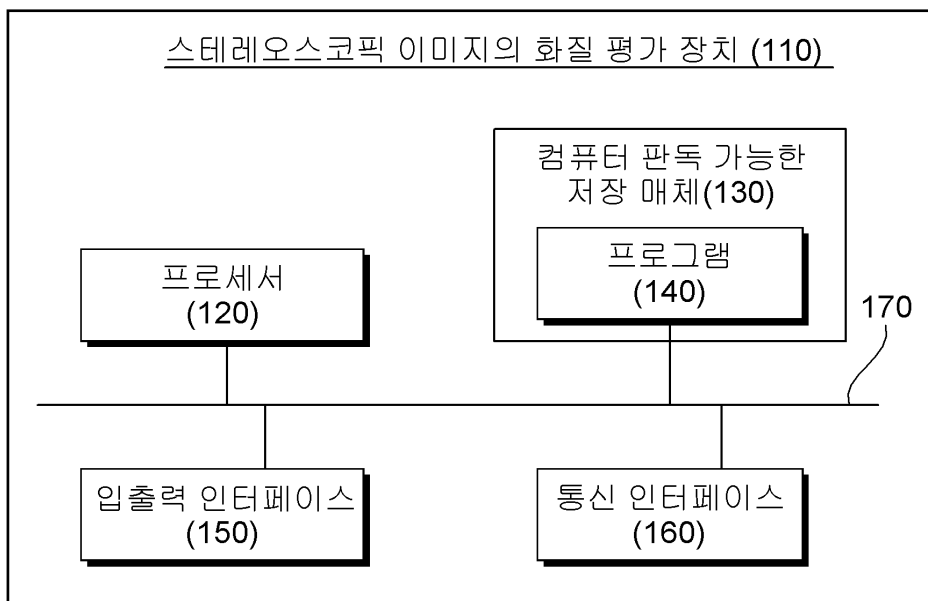
터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.

[0090]

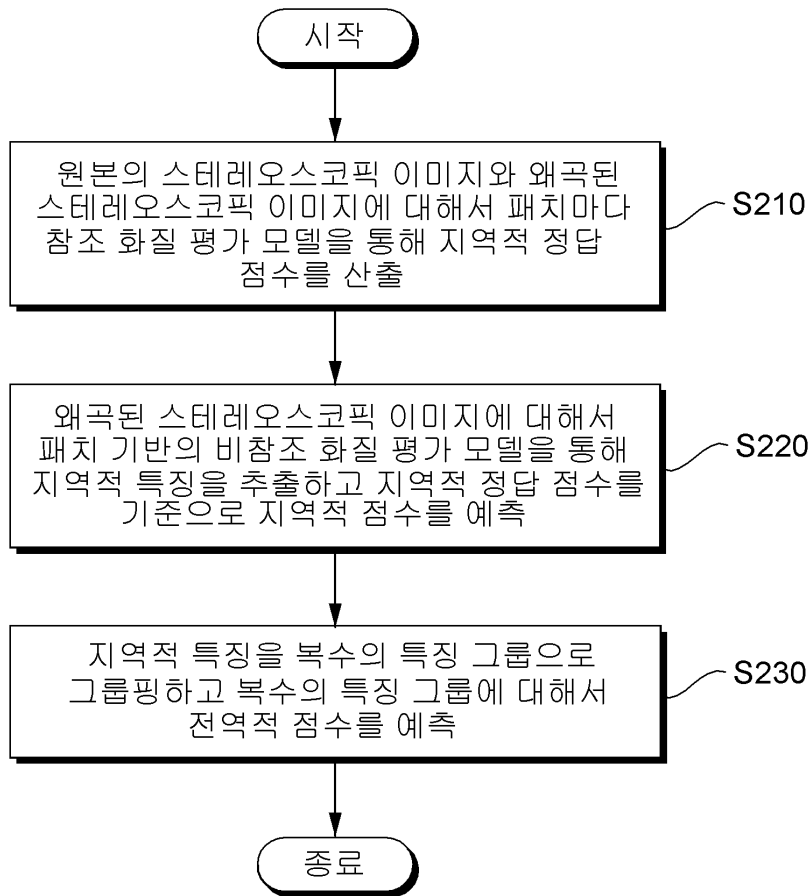
본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

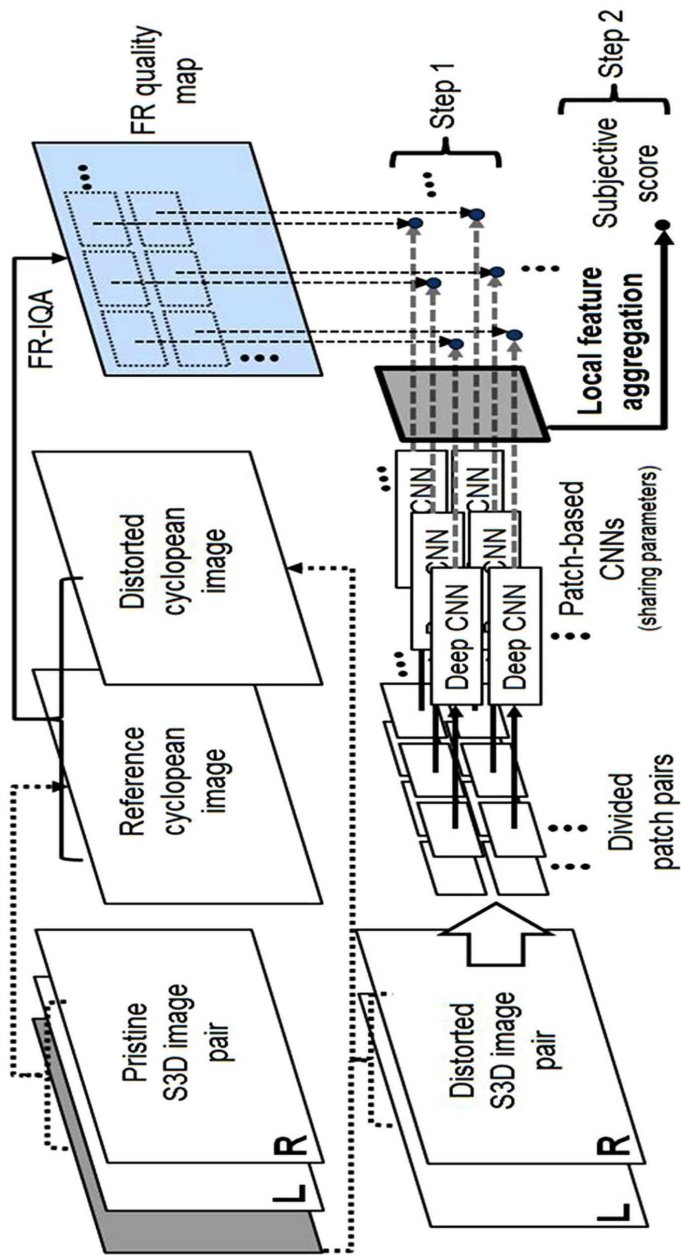
도면1



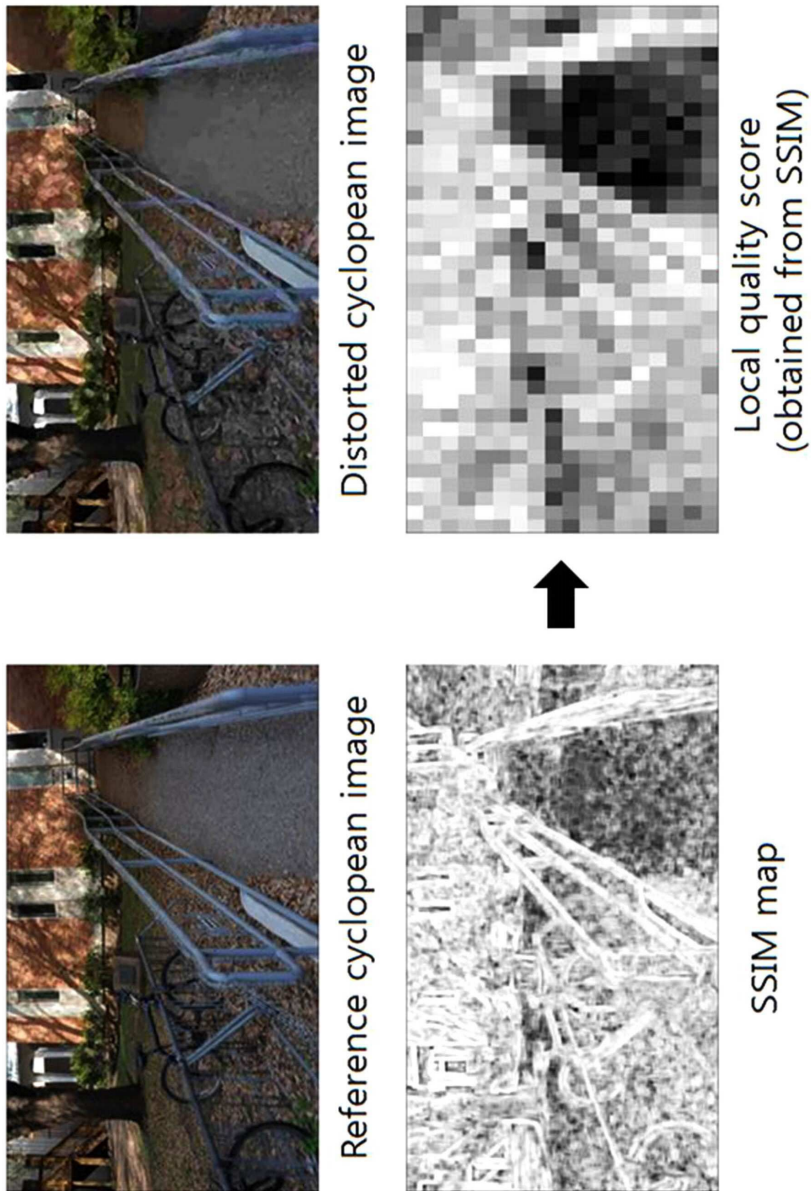
도면2



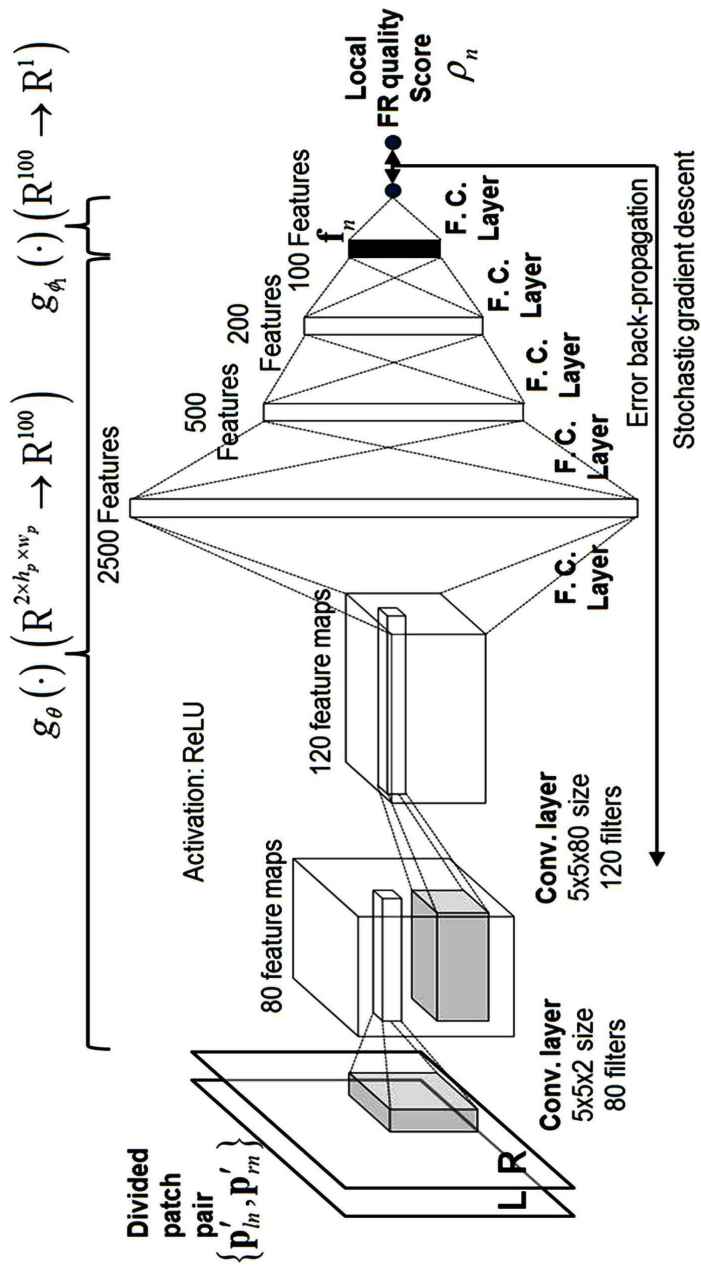
도면3



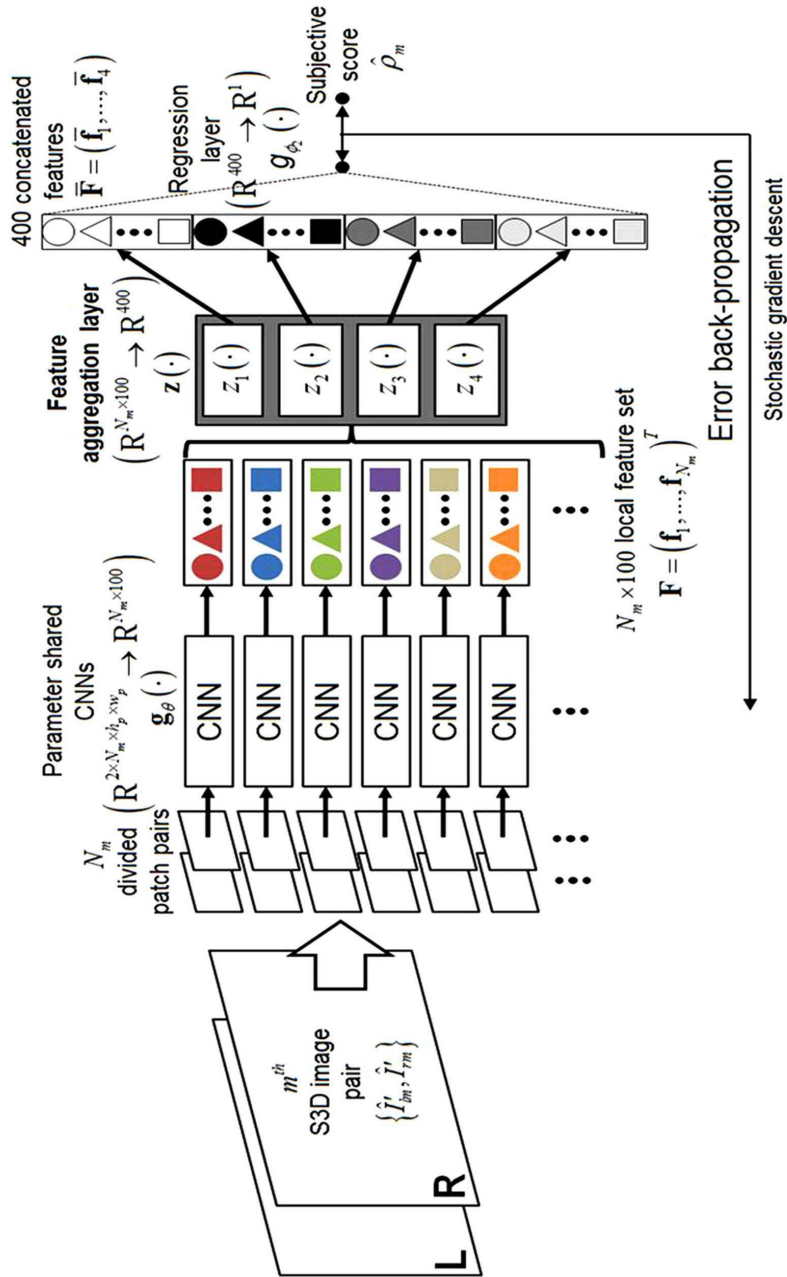
도면4



도면5



도면6



도면7

