(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0137362
(43) 공개일자 2024년09월20일

(51) 국제특허분류(Int. Cl.)
G06N 3/045 (2023.01) *G06N 3/0455* (2023.01)
H04N 1/40 (2006.01)
(52) CPC특허분류
G06N 3/045 (2023.01)
G06N 3/0455 (2023.01)
(21) 출원번호 10-2023-0030667
(22) 출원일자 2023년03월08일
심사청구일자 2023년03월08일

(71) 출원인
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
전해곤
광주광역시 북구 첨단과기로 123 광주과학기술원 AI대학원
신승현
광주광역시 북구 첨단과기로 123 광주과학기술원 AI대학원
(뒷면에 계속)
(74) 대리인
특허법인지원

전체 청구항 수 : 총 12 항

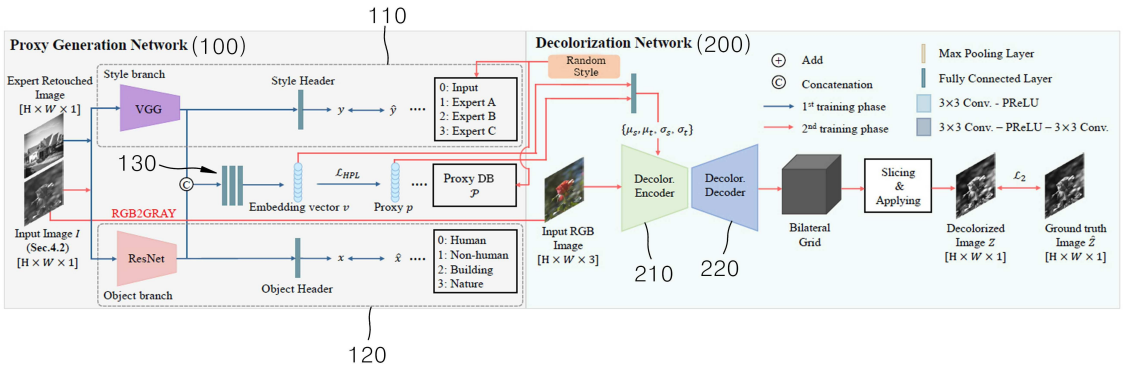
(54) 발명의 명칭 **전문가의 보정 스타일을 모사한 흑백사진 생성 방법**

(57) 요 약

본 발명은 복수 전문가들의 보정 스타일을 모사할 수 있는 신경망 모델을 생성하고, 이를 통해 심미감을 주는 흑백사진을 생성하는 방법에 관한 것이다. 본 발명의 일 실시예에 따른 흑백사진 생성 방법은, 복수의 제1 학습용 흑백사진을 제1 신경망 네트워크에 입력하여 스타일 및 촬영객체에 따른 제1 임베딩 벡터를 생성하는 단계와 제1

(뒷면에 계속)

대 표 도



임베딩 벡터가 형성하는 각 군집을 대표하는 복수의 프록시 벡터를 추출하는 단계, 학습용 컬러사진을 제2 학습용 흑백사진으로 변환하고 제2 학습용 흑백사진을 제1 신경망 네트워크에 입력하여 제2 임베딩 벡터를 생성하는 단계, 복수의 프록시 벡터 중 학습용 컬러사진에 대한 랜덤 스타일과 촬영객체에 대응하는 학습용 프록시 벡터를 식별하는 단계, 학습용 컬러사진, 학습용 프록시 벡터 및 제2 임베딩 벡터를 제2 신경망 네트워크에 입력하여 제2 임베딩 벡터의 분포를 학습용 프록시 벡터의 분포로 변환하는 단계 및 변환된 분포에 대응하는 픽셀별 가중치를 학습용 컬러사진에 적용하여 탈색 이미지를 생성하는 단계를 포함하는 것을 특징으로 한다.

(52) CPC특허분류

H04N 1/40012 (2013.01)

(72) 발명자

신지수

광주광역시 북구 첨단과기로 123 광주과학기술원 AI대학원

배지환

광주광역시 북구 첨단과기로 123 광주과학기술원 전기전자컴퓨터학과

심인욱

대전광역시 유성구 지족북로 33, 102동 1105호(지족동, 노은꿈에그린 1단지)

이 발명을 지원한 국가연구개발사업
과제고유번호 1415177564
과제번호 P0019797(GM15600)
부처명 산업통상자원부
과제관리(전문)기관명 한국산업기술진흥원
연구사업명 국제공동기술개발
연구과제명 뉴럴 휴먼 모델링 기반의 사용자 참여형 메타버스 공연 솔루션 개발
기 여 율 80/99
과제수행기관명 (주)위지윅스튜디오
연구기간 2021.12.01 ~ 2022.11.30
이 발명을 지원한 국가연구개발사업
과제고유번호 1711177232
과제번호 2022-DD-UP-0312-01(NN33370)
부처명 과학기술정보통신부
과제관리(전문)기관명 연구개발특구진흥재단
연구사업명 지역의 미래를 여는 과학기술 프로젝트
연구과제명 인공지능 기반 메타버스 구현을 위한 융·복합 문화 가상 스튜디오
기 여 율 10/99
과제수행기관명 광주과학기술원
연구기간 2022.04.01 ~ 2022.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711139517
과제번호 2021-0-02068-001(GE11480)
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 교육훈련지원
연구과제명 인공지능 혁신 허브 연구 개발
기 여 율 5/99
과제수행기관명 고려대학
연구기간 2021.07.01 ~ 2021.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711166594
과제번호 2020R1C1C1012635(NN31990)
부처명 과학기술정보통신부
과제관리(전문)기관명 한국연구재단
연구사업명 우수신진연구
연구과제명 공간변화에 강인한 인간 시각 인지 기반의 카메라 위치추정
기 여 율 3/99
과제수행기관명 광주과학기술원
연구기간 2022.03.01 ~ 2023.02.28
이 발명을 지원한 국가연구개발사업
과제고유번호 1711097871
과제번호 2019-0-01842-001(NN30900)
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 인공지능대학원지원
연구과제명 인공지능핵심고급인재양성
기 여 율 1/99
과제수행기관명 광주과학기술원
연구기간 2019.09.01 ~ 2023.12.31

명 세 서

청구범위

청구항 1

복수의 제1 학습용 흑백사진을 제1 신경망 네트워크에 입력하여 각 제1 학습용 흑백사진의 스타일 및 촬영객체에 따른 제1 임베딩 벡터를 생성하고, 상기 제1 임베딩 벡터가 형성하는 각 군집을 대표하는 복수의 프록시 벡터를 추출함으로써 상기 제1 신경망 네트워크를 학습시키는 단계;

학습용 컬러사진을 제2 학습용 흑백사진으로 변환하고, 상기 제2 학습용 흑백사진을 상기 제1 신경망 네트워크에 입력하여 제2 임베딩 벡터를 생성하는 단계;

상기 복수의 프록시 벡터 중 상기 학습용 컬러사진에 대한 랜덤 스타일과 상기 학습용 컬러사진의 촬영객체에 대응하는 학습용 프록시 벡터를 식별하는 단계;

상기 학습용 컬러사진, 상기 학습용 프록시 벡터 및 상기 제2 임베딩 벡터를 제2 신경망 네트워크에 입력하여 상기 제2 임베딩 벡터의 분포를 상기 학습용 프록시 벡터의 분포로 변환하고, 상기 변환된 분포에 대응하는 픽셀별 가중치를 상기 학습용 컬러사진에 적용하여 탈색 이미지를 생성함으로써 상기 제2 신경망 네트워크를 학습시키는 단계; 및

학습이 완료된 상기 제1 및 제2 신경망 네트워크를 이용하여 타겟 컬러사진을 타겟 스타일에 대응하는 흑백사진으로 변환하는 단계를 포함하는

흑백사진 생성 방법.

청구항 2

제1항에 있어서,

상기 제1 신경망 네트워크는

상기 제1 학습용 흑백사진의 스타일을 분류하는 스타일 분류 신경망과,

상기 제1 학습용 흑백사진 내 촬영객체를 분류하는 객체 분류 신경망과,

상기 스타일 분류 신경망 및 상기 객체 분류 신경망의 출력을 결합하여 상기 제1 임베딩 벡터를 생성하는 다층 퍼셉트론(Multi-Layer Perceptron; MLP)을 포함하는

흑백사진 생성 방법.

청구항 3

제1항에 있어서,

상기 제1 신경망 네트워크를 학습시키는 단계는

임베딩 공간 상에서 상기 제1 임베딩 벡터가 동일 스타일에 대해서는 가깝게, 다른 촬영객체에 대해서는 멀리 위치하도록 상기 제1 신경망 네트워크를 학습시키는 단계를 포함하는

흑백사진 생성 방법.

청구항 4

제2항에 있어서,

상기 제2 임베딩 벡터를 생성하는 단계는

상기 학습용 컬러사진의 계조값을 그레이스케일(grayscale)로 변환하여 상기 제2 학습용 흑백사진을 생성하는 단계와,

상기 제2 학습용 흑백사진을 상기 스타일 분류 신경망 및 상기 객체 분류 신경망에 각각 입력하고, 상기 다층 퍼셉트론의 출력을 상기 제2 임베딩 벡터로 결정하는 단계를 포함하는

흑백사진 생성 방법.

청구항 5

제1항에 있어서,

상기 학습용 프록시 벡터를 식별하는 단계는

상기 복수의 프록시 벡터 중, 상기 학습용 컬러사진에 대해 임의로 설정되는 랜덤 스타일 및 상기 학습용 컬러사진 내 촬영객체에 대응하는 어느 한 학습용 프록시 벡터를 식별하는 단계를 포함하는

흑백사진 생성 방법.

청구항 6

제5항에 있어서,

상기 학습용 컬러사진 내 촬영객체는 상기 제2 학습용 흑백사진을 상기 제1 신경망 네트워크 내 객체 분류 신경망에 입력하여 식별되는

흑백사진 생성 방법.

청구항 7

제1항에 있어서,

상기 제2 신경망 네트워크는

상기 학습용 컬러사진으로부터 특징을 추출하고, 상기 제2 임베딩 벡터의 분포를 상기 학습용 프록시 벡터의 분포로 변환하는 인코더와,

상기 인코더에서 추출된 특징과 상기 변환된 제2 임베딩 벡터의 분포로부터 상기 픽셀별 가중치를 출력하는 디코더를 포함하는

흑백사진 생성 방법.

청구항 8

제7항에 있어서,

상기 제2 신경망 네트워크는, 상기 제2 임베딩 벡터와 상기 학습용 프록시 벡터로부터 특징맵을 각각 추출하는 완전 연결 레이어(Fully Connected Layer; FCL)를 더 포함하고,

상기 인코더는, 상기 제2 임베딩 벡터에 대한 특징맵의 분포를 상기 학습용 프록시 벡터에 대한 특징맵의 분포로 변환하는

흑백사진 생성 방법.

청구항 9

제8항에 있어서,

상기 인코더는 하기 [수학식 1]에 따라 상기 제2 임베딩 벡터의 특징맵 분포를 상기 학습용 프록시 벡터의 특징맵 분포로 변환하는

[수학식 1]

$$\mathcal{F}' = \sigma_t\left(\frac{\mathcal{F}-\mu_s}{\sigma_s}\right)+\mu_t$$

($\mathcal{F}'$는 변환된 제2 임베딩 벡터에 대한 특징맵, $\mathcal{F}$는 상기 제2 임베딩 벡터에 대한 특징맵, μ_s,μ_t는 각각 상기 제2 임베딩 벡터에 대한 특징맵의 평균 및 분산, μ_t,σ_t는 각각 상기 학습용 프록시 벡터에 대한 특징맵의 평균 및 분산)

흑백사진 생성 방법.

청구항 10

제1항에 있어서,

상기 픽셀별 가중치는 상기 학습용 컬러사진의 픽셀별 계조값 대한 바이레터럴 그리드(bilateral grid) 형태의 가중치인

흑백사진 생성 방법.

청구항 11

제1항에 있어서,

상기 제2 신경망 네트워크를 학습시키는 단계는

상기 학습용 컬러사진의 픽셀별 계조값에 상기 픽셀별 가중치를 곱하여 생성된 상기 탈색 이미지와, 상기 학습용 컬러사진을 상기 랜덤 스타일에 따라 변환한 실측(Ground Truth; GT) 흑백사진 간의 차이가 최소가 되도록 상기 제2 신경망 네트워크를 학습시키는 단계를 포함하는

흑백사진 생성 방법.

청구항 12

제1항에 있어서,

상기 흑백사진으로 변환하는 단계는

상기 타겟 컬러사진을 그레이스케일 이미지로 변환하여 상기 제1 신경망 네트워크에 입력하고, 상기 타겟 스타일에 대응하는 프록시 벡터를 상기 제2 신경망 네트워크에 입력하여 상기 타겟 컬러사진을 흑백사진으로 변환하는 단계를 포함하는

흑백사진 생성 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 복수 전문가들의 보정 스타일을 모사할 수 있는 신경망 모델을 생성하고, 이를 통해 심미감을 주는 흑백사진을 생성하는 방법에 관한 것이다.

배 경 기 술

[0003] 흑백사진은 컬러사진보다 더 높은 다이내믹 레인지(dynamic range)와 선명도를 가짐으로써 더욱 풍부한 텍스처과 대비감을 제공한다. 이를 통해 흑백사진은 컬러사진에서 볼 수 없는 독특한 심미감과 감성을 제공한다.

[0004] 이러한 흑백사진의 심미감은 단순히 스마트폰이나 DSLR 카메라의 부가 기능이 아닌, 흑백 촬영을 위한 고급 카메라나 전문가의 세심한 보정을 통해서만 달성될 수 있다.

[0005] 현재 컴퓨터 비전(computer vision) 분야에서 흑백사진 변환을 위한 다양한 신경망 네트워크들이 제안되고 있으나, 이들은 모두 RGB 계조값을 그레이스케일(grayscale)로 변환함에 있어서 텍스처의 손실을 최소화하는 데 중점을 두고 있을 뿐, 흑백사진 생성에 미학적 요소를 가미하는 것에 대해서는 연구되고 있지 않은 실정이다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 신경망 모델을 통해 전문가들의 보정 스타일을 모사하여 컬러사진을 흑백사진으로 변환하는 것을 목적으로 한다.

[0008] 본 발명의 목적들은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 본 발명의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있고, 본 발명의 실시예에 의해 보다 분명하게 이해될 것이다. 또한, 본 발명의 목적 및 장점들은 특허 청구 범위에 나타낸 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

과제의 해결 수단

[0010] 전술한 목적을 달성하기 위한 본 발명의 일 실시예에 따른 흑백사진 생성 방법은, 복수의 제1 학습용 흑백사진을 제1 신경망 네트워크에 입력하여 각 제1 학습용 흑백사진의 스타일 및 촬영객체에 따른 제1 임베딩 벡터를 생성하고, 상기 제1 임베딩 벡터가 형성하는 각 군집을 대표하는 복수의 프록시 벡터를 추출함으로써 상기 제1 신경망 네트워크를 학습시키는 단계, 학습용 컬러사진을 제2 학습용 흑백사진으로 변환하고, 상기 제2 학습용 흑백사진을 상기 제1 신경망 네트워크에 입력하여 제2 임베딩 벡터를 생성하는 단계, 상기 복수의 프록시 벡터 중 상기 학습용 컬러사진에 대한 랜덤 스타일과 상기 학습용 컬러사진의 촬영객체에 대응하는 학습용 프록시 벡터를 식별하는 단계, 상기 학습용 컬러사진, 상기 학습용 프록시 벡터 및 상기 제2 임베딩 벡터를 제2 신경망 네트워크에 입력하여 상기 제2 임베딩 벡터의 분포를 상기 학습용 프록시 벡터의 분포로 변환하고, 상기 변환된 분포에 대응하는 픽셀별 가중치를 상기 학습용 컬러사진에 적용하여 탈색 이미지를 생성함으로써 상기 제2 신경망 네트워크를 학습시키는 단계 및 학습이 완료된 상기 제1 및 제2 신경망 네트워크를 이용하여 타겟 컬러사진을 타겟 스타일에 대응하는 흑백사진으로 변환하는 단계를 포함하는 것을 특징으로 한다.

[0011] 일 실시예에서, 상기 제1 신경망 네트워크는, 상기 제1 학습용 흑백사진의 스타일을 분류하는 스타일 분류 신경망과, 상기 제1 학습용 흑백사진 내 촬영객체를 분류하는 객체 분류 신경망과, 상기 스타일 분류 신경망 및 상기 객체 분류 신경망의 출력을 결합하여 상기 제1 임베딩 벡터를 생성하는 다층 퍼셉트론(Multi-Layer Perceptron; MLP)을 포함하는 것을 특징으로 한다.

[0012] 일 실시예에서, 상기 제1 신경망 네트워크를 학습시키는 단계는, 임베딩 공간 상에서 상기 제1 임베딩 벡터가 동일 스타일에 대해서는 가깝게, 다른 촬영객체에 대해서는 멀리 위치하도록 상기 제1 신경망 네트워크를 학습시키는 단계를 포함하는 것을 특징으로 한다.

[0013] 일 실시예에서, 상기 제2 임베딩 벡터를 생성하는 단계는, 상기 학습용 컬러사진의 계조값을 그레이스케일(grayscale)로 변환하여 상기 제2 학습용 흑백사진을 생성하는 단계와, 상기 제2 학습용 흑백사진을 상기 스타일 분류 신경망 및 상기 객체 분류 신경망에 각각 입력하고, 상기 다층 퍼셉트론의 출력을 상기 제2 임베딩 벡터로 결정하는 단계를 포함하는 것을 특징으로 한다.

[0014] 일 실시예에서, 상기 학습용 프록시 벡터를 식별하는 단계는, 상기 복수의 프록시 벡터 중, 상기 학습용 컬러사진에 대해 임의로 설정되는 랜덤 스타일 및 상기 학습용 컬러사진 내 촬영객체에 대응하는 어느 한 학습용 프록시 벡터를 식별하는 단계를 포함하는 것을 특징으로 한다.

[0015] 일 실시예에서, 상기 학습용 컬러사진 내 촬영객체는 상기 제2 학습용 흑백사진을 상기 제1 신경망 네트워크 내 객체 분류 신경망에 입력하여 식별되는 것을 특징으로 한다.

[0016] 일 실시예에서, 상기 제2 신경망 네트워크는, 상기 학습용 컬러사진으로부터 특징을 추출하고, 상기 제2 임베딩 벡터의 분포를 상기 학습용 프록시 벡터의 분포로 변환하는 인코더와, 상기 인코더에서 추출된 특징과 상기 변환된 제2 임베딩 벡터의 분포로부터 상기 픽셀별 가중치를 출력하는 디코더를 포함하는 것을 특징으로 한다.

[0017] 일 실시예에서, 상기 제2 신경망 네트워크는, 상기 제2 임베딩 벡터와 상기 학습용 프록시 벡터로부터 특징맵을 각각 추출하는 완전 연결 레이어(Fully Connected Layer; FCL)를 더 포함하고, 상기 인코더는, 상기 제2 임베딩 벡터에 대한 특징맵의 분포를 상기 학습용 프록시 벡터에 대한 특징맵의 분포로 변환하는 것을 특징으로 한다.

[0018] 일 실시예에서, 상기 인코더는 하기 [수학식 1]에 따라 상기 제2 임베딩 벡터의 특징맵 분포를 상기 학습용 프록시 벡터의 특징맵 분포로 변환하는 것을 특징으로 한다.

[0019] [수학식 1]

[0020] $$\mathcal{F}' = \sigma_t\left(\frac{\mathcal{F}-\mu_s}{\sigma_s}\right)+\mu_t$$

[0021] ($\mathcal{F}'$는 변환된 제2 임베딩 벡터에 대한 특징맵, $\mathcal{F}$는 상기 제2 임베딩 벡터에 대한 특징맵, μ_s,μ_t는 각각 상기 제2 임베딩 벡터에 대한 특징맵의 평균 및 분산, μ_t,σ_t는 각각 상기 학습용 프록시 벡터에 대한 특징맵의 평균 및 분산)

[0022] 일 실시예에서, 상기 픽셀별 가중치는 상기 학습용 컬러사진의 픽셀별 계조값 대한 바이레터럴 그리드(bilateral grid) 형태의 가중치인 것을 특징으로 한다.

[0023] 일 실시예에서, 상기 제2 신경망 네트워크를 학습시키는 단계는, 상기 학습용 컬러사진의 픽셀별 계조값에 상기 픽셀별 가중치를 곱하여 생성된 상기 탈색 이미지와, 상기 학습용 컬러사진을 상기 랜덤 스타일에 따라 변환한 실측(Ground Truth; GT) 흑백사진 간의 차이가 최소가 되도록 상기 제2 신경망 네트워크를 학습시키는 단계를 포함하는 것을 특징으로 한다.

[0024] 일 실시예에서, 상기 흑백사진으로 변환하는 단계는, 상기 타겟 컬러사진을 그레이스케일 이미지로 변환하여 상기 제1 신경망 네트워크에 입력하고, 상기 타겟 스타일에 대응하는 프록시 벡터를 상기 제2 신경망 네트워크에 입력하여 상기 타겟 컬러사진을 흑백사진으로 변환하는 단계를 포함하는 것을 특징으로 한다.

발명의 효과

[0026] 본 발명은 전문가들의 흑백사진 변환 스타일을 모사할 수 있도록 신경망 모델을 학습시킴으로써, 딥러닝 모델을 이용한 흑백사진 생성 동작에서 심미적/미학적 향상을 달성할 수 있다.

[0027] 상술한 효과와 더불어 본 발명의 구체적인 효과는 이하 발명을 실시하기 위한 구체적인 사항을 설명하면서 함께 기술한다.

도면의 간단한 설명

[0029] 도 1은 본 발명의 일 실시예에 따른 흑백사진 생성 방법을 도시한 순서도.

도 2a는 각 전문가에 의해 조절된 파라미터들을 예시적으로 도시한 도면.

도 2b는 도 2a에 도시된 파라미터에 따라 보정된 흑백사진들을 예시적으로 도시한 도면.

도 3은 흑백사진에 대한 임베딩 벡터를 생성하는 제1 신경망 네트워크를 예시적으로 도시한 도면.

도 4는 보정 스타일 및 객체별 임베딩 벡터가 임베딩 공간 상에서 군집화된 모습을 도시한 도면.

도 5는 원하는 스타일에 따라 변환된 흑백사진을 생성하는 제2 신경망 네트워크를 예시적으로 도시한 도면.

도 6은 컬러사진에 적용되는 바이레터럴 그리드(bilateral grid) 형태의 가중치를 도시한 도면.

도 7은 제1 및 제2 신경망 네트워크가 결합된 전체 아키텍처를 도시한 도면.

도 8은 본 발명에 따라 생성된 흑백사진들을 예시적으로 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0030] 전술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술되며, 이에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 상세한 설명을 생략한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 도면에서 동일한 참조부호는 동일 또는 유사한 구성요소를 가리키는 것으로 사용된다.

[0031] 본 명세서에서 제1, 제2 등이 다양한 구성요소들을 서술하기 위해서 사용되나, 이들 구성요소들은 이들 용어에 의해 제한되지 않음은 물론이다. 이들 용어들은 단지 하나의 구성요소를 다른 구성요소와 구별하기 위하여 사용하는 것으로, 특별히 반대되는 기재가 없는 한, 제1 구성요소는 제2 구성요소일 수도 있음은 물론이다.

[0032] 또한, 본 명세서에서 "상부 (또는 하부)" 또는 구성요소의 "상 (또는 하)"에 임의의 구성이 배치된다는 것은, 임의의 구성이 상기 구성요소의 상면 (또는 하면)에 접하여 배치되는 것뿐만 아니라, 상기 구성요소와 상기 구성요소 상에 (또는 하에) 배치된 임의의 구성 사이에 다른 구성이 개재될 수 있음을 의미할 수 있다.

[0033] 또한, 본 명세서에서 어떤 구성요소가 다른 구성요소에 "연결", "결합" 또는 "접속"된다고 기재된 경우, 상기 구성요소들은 서로 직접적으로 연결되거나 또는 접속될 수 있지만, 각 구성요소 사이에 다른 구성요소가 "개재"되거나, 각 구성요소가 다른 구성요소를 통해 "연결", "결합" 또는 "접속"될 수도 있는 것으로 이해되어야 할 것이다.

[0034] 또한, 본 명세서에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "구성된다" 또는 "포함한다" 등의 용어는 명세서 상에 기재된 여러 구성 요소들, 또는 여러 단계들을 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다.

[0035] 또한, 본 명세서에서, "A 및/또는 B" 라고 할 때, 이는 특별한 반대되는 기재가 없는 한, A, B 또는 A 및 B를 의미하며, "C 내지 D" 라고 할 때, 이는 특별한 반대되는 기재가 없는 한, C 이상이고 D 이하인 것을 의미한다

[0036] 본 발명은 복수 전문가들의 보정 스타일을 모사할 수 있는 신경망 모델을 생성하고, 이를 통해 심미감을 주는 흑백사진을 생성하는 방법에 관한 것이다. 이하, 도 1 내지 도 8을 참조하여 본 발명의 일 실시예에 따른 흑백사진 생성 방법을 구체적으로 설명하도록 한다.

[0037] 도 1은 본 발명의 일 실시예에 따른 흑백사진 생성 방법을 도시한 순서도이다.

[0038] 도 2a는 각 전문가에 의해 조절된 파라미터들을 예시적으로 도시한 도면이고, 도 2b는 도 2a에 도시된 파라미터에 따라 보정된 흑백사진들을 예시적으로 도시한 도면이다.

[0039] 도 3은 흑백사진에 대한 임베딩 벡터를 생성하는 제1 신경망 네트워크를 예시적으로 도시한 도면이다. 도 4는 보정 스타일 및 객체별 임베딩 벡터가 임베딩 공간 상에서 군집화된 모습을 도시한 도면이다.

[0040] 도 5는 원하는 스타일에 따라 변환된 흑백사진을 생성하는 제2 신경망 네트워크를 예시적으로 도시한 도면이다.

[0041] 도 6은 컬러사진에 적용되는 바이레터럴 그리드(bilateral grid) 형태의 가중치를 도시한 도면이다.

[0042] 도 7은 제1 및 제2 신경망 네트워크가 결합된 전체 아키텍처를 도시한 도면이고, 도 8은 본 발명에 따라 생성된 흑백사진들을 예시적으로 도시한 도면이다.

[0043] 도 1을 참조하면, 본 발명의 일 실시예에 따른 전문가의 보정 스타일을 모사한 흑백사진 생성 방법(이하, 흑백사진 생성 방법)은, 복수의 제1 학습용 흑백사진을 제1 신경망 네트워크에 입력하여 스타일 및 촬영객체에 따른

제1 임베딩 벡터를 생성하는 단계(S10)와 제1 임베딩 벡터가 형성하는 각 군집을 대표하는 복수의 프록시 벡터를 추출하는 단계(S20)를 통해 제1 신경망 네트워크를 학습시키는 단계를 포함할 수 있다.

[0044] 이어서, 흑백사진 생성 방법은, 학습용 컬러사진을 제2 학습용 흑백사진으로 변환하고 제2 학습용 흑백사진을 제1 신경망 네트워크에 입력하여 제2 임베딩 벡터를 생성하는 단계(S30), 복수의 프록시 벡터 중 학습용 컬러사진에 대한 랜덤 스타일과 촬영객체에 대응하는 학습용 프록시 벡터를 식별하는 단계(S40), 학습용 컬러사진, 학습용 프록시 벡터 및 제2 임베딩 벡터를 제2 신경망 네트워크에 입력하여 제2 임베딩 벡터의 분포를 학습용 프록시 벡터의 분포로 변환하는 단계(S50) 및 변환된 분포에 대응하는 픽셀별 가중치를 학습용 컬러사진에 적용하여 탈색 이미지를 생성하는 단계(S60)를 통해 제2 신경망 네트워크를 학습시키는 단계를 포함할 수 있다.

[0045] 이어서, 흑백사진 생성 방법은, 앞서 학습된 제1 및 제2 신경망 네트워크를 이용하여 타겟 컬러사진을 타겟 스타일에 대응하는 흑백사진으로 변환하는 단계(S70)를 포함할 수 있다.

[0046] 다만, 도 1에 도시된 흑백사진 생성 방법은 일 실시예에 따른 것이고, 발명을 이루는 각 단계들이 도 1에 도시된 실시예에 한정되는 것은 아니며, 필요에 따라 일부 단계들이 부가, 변경 또는 삭제될 수 있다.

[0047] 도 1에 도시된 각 단계들은 CPU(central processing unit), GPU(graphics processing unit) 등으로 구현되는 프로세서에 의해 수행될 수 있으며, 프로세서는 후술하는 발명의 동작을 수행하기 위하여 ASICs(application specific integrated circuits), DSPs(digital signal processors), DSPDs(digital signal processing devices), PLDs(programmable logic devices), FPGAs(field programmable gate arrays), 제어기(controller), 마이크로 컨트롤러(micro-controllers) 중 적어도 하나의 물리적인 요소를 더 포함할 수 있다.

[0048] 본 발명에서 흑백사진 생성에 이용되는 신경망 모델은 제1 및 제2 신경망 네트워크(100, 200)를 포함하는데, 먼저 제1 신경망 네트워크(100)를 학습시키는 과정, 즉 단계(S10) 및 단계(S20)을 구체적으로 설명하도록 한다.

[0049] 프로세서는 제1 학습용 흑백사진(11)을 제1 신경망 네트워크(100)에 입력하여 각각의 제1 학습용 흑백사진(11)별 스타일 및 촬영객체에 따른 제1 임베딩 벡터(v)를 생성할 수 있다(S10).

[0050] 제1 학습용 흑백사진(11)은 제1 신경망 네트워크(100)의 학습에 이용되는 흑백사진으로, 전문가의 수작업을 통해 보정된 것일 수 있다. 이에 따라, 제1 학습용 흑백사진(11)은 각 전문가에 따른 스타일을 가질 수 있다.

[0051] 도 2a 및 도 2b를 함께 참조하면, 동일한 컬러사진을 흑백사진으로 변환함에 있어서, 전문가별로 세부적인 보정 방법이 다를 수 있다.

[0052] 구체적으로, 전문가 A, B, C의 보정 방식은 크게 감마 곡선(gamma curve), 노출, 콘트라스트(contrast) 등의 글로벌 파라미터를 조절한 뒤, 사진 내 각 부분별로 비네팅(vignetting), 감마 보정(gamma correction) 등의 필터를 적용하는 점에서 유사할 수 있다.

[0053] 그러나, 전문가 A는 노출값을 높게 조절하고 선명도와 그림자 효과를 이용하여 어두운 영역을 비교적 밝게 조절하는 점에서 다른 전문가들 대비 차별점이 있고, 전문가 B는 텍스처와 선명도 조절을 통해 장면의 디테일을 높이고 비네팅 필터를 통해 사람이나 꽃과 같은 주요 피사체를 강조하는 점에서 다른 전문가들 대비 차별점이 있으며, 전문가 C는 하이라이트 및 그림자 레벨의 제어를 통해 이미지 내 가장 밝은 부분과 가장 어두운 부분을 보정함으로써 흑백사진이 더 높은 다이내믹 레인지(dynamic range)를 갖게 하는 점에서 다른 전문가들 대비 차별점이 있다.

[0054] 또한, 동일한 전문가의 보정이라 하더라도 촬영된 객체(이하, 촬영객체)에 따라 세부적인 보정 방법이 달라질 수 있다.

[0055] 구체적으로, 전문가 A는 건물 사진을 보정할 때 건물의 콘트라스트를 높이면서 풍경에 가장자리 강화 필터(edge enhancement filter)와 높은 노출값을 적용하는 경향이 있다. 이와 달리, 인물 사진을 보정할 때 전문가 A는 강한 그림자 효과를 사용하여 머리카락과 같은 미세한 구조의 디테일을 풍부하게 하는 경향이 있다.

[0056] 마찬가지로 전문가 B, C 역시 촬영객체에 따라 보정 방법을 달리하며, 본 발명에서는 각각의 전문가가 보정한 촬영객체(인간(human), 비인간(non-human), 건물 및 자연)별 흑백사진을 수집하여 제1 학습용 흑백사진(11)으로 이용할 수 있다.

[0057] 프로세서는 각각의 제1 학습용 흑백사진(11)을 제1 신경망 네트워크(100)에 입력하여 제1 임베딩 벡터(v)를 생성할 수 있다(S10). 앞서 설명한 것과 같이 제1 학습용 흑백사진(11)은 스타일, 즉 해당 사진을 보정한 전문가가 누구인지에 따라 경향성을 가질 수 있다. 또한, 제1 학습용 흑백사진(11)은 해당 사진의 촬영객체에 따라 경

향성을 가질 수 있다.

[0058] 제1 임베딩 벡터(v)는 제1 학습용 흑백사진(11)의 특징에 기초하여 생성될 수 있고, 이에 따라 제1 임베딩 벡터(v)는 동일한 스타일 및 동일한 촬영객체에 따라 유사성을 가질 수 있다. 이하, 제1 신경망 네트워크(100)의 예시적인 구조를 참조하여 제1 임베딩 벡터(v) 생성 방법을 설명하도록 한다.

[0059] 도 3을 참조하면, 제1 신경망 네트워크(100)는, 제1 학습용 흑백사진(11)의 스타일을 분류하는 스타일 분류 신경망(110)과, 제1 학습용 흑백사진(11) 내 촬영객체를 분류하는 객체 분류 신경망(120)과, 전술한 두 신경망의 출력을 결합하여 제1 임베딩 벡터(v)를 생성하는 다층 퍼셉트론(Multi-Layer Perceptron; MLP, 130)을 포함할 수 있다.

[0060] 스타일 분류 신경망(110)과 객체 분류 신경망(120)은 이미지를 입력으로 받아 분류 태스크(classification task)를 수행하는 하는 컨볼루션(convolution) 기반의 신경망일 수 있다.

[0061] 분류 태스크를 위해 스타일 분류 신경망(110)은 제1 학습용 흑백사진(11)을 입력 데이터로 하고, 해당 영상의 스타일, 즉 전문가가 누구인지에 따라 결정되는 클래스(0 내지 3)를 출력 데이터로 하는 훈련 데이터셋(training dataset)에 의해 지도 학습(supervised learning)될 수 있다. 또한, 객체 분류 신경망(120)은 제1 학습용 흑백사진(11)을 입력 데이터로 하고 사진 속 촬영객체에 따라 결정되는 클래스(0 내지 3)를 출력 데이터로 하는 훈련 데이터셋에 의해 지도 학습될 수 있다. 여기서 각 클래스들은 각각의 제1 학습용 흑백사진(11)에 대해 미리 라벨링(labeling)될 수 있다.

[0062] 지도 학습에 따라 스타일 분류 신경망(110)은 제1 학습용 흑백사진(11)에서 스타일에 대응하는 특징을 추출할 수 있게 되고, 객체 분류 신경망(120)은 제1 학습용 흑백사진(11)에서 촬영객체에 대응하는 특징을 추출할 수 있게 된다. 본 발명의 일 예시에서 스타일 분류 신경망(110)은 VGG로 구현될 수 있고, 객체 분류 신경망(120)은 ResNet으로 구현될 수 있다.

[0063] 스타일 분류 신경망(110) 및 객체 분류 신경망(120)에서 추출된 각 특징들은 결합되어 다층 퍼셉트론(130)으로 입력될 수 있다. 다층 퍼셉트론(130)은 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성될 수 있고, 앞서 결합된 특징으로부터 제1 임베딩 벡터(v)를 생성할 수 있다.

[0064] 이 때, 다층 퍼셉트론(130)은 임베딩 공간(embedding space)에서 제1 임베딩 벡터(v)가 동일 스타일에 대해서는 가깝게, 다른 촬영객체에 대해서는 멀리 위치하도록 비지도 학습(unsupervised learning)될 수 있다.

[0065] 도 4를 참조하면, 복수의 제1 학습용 흑백사진(11)으로부터 생성된 제1 임베딩 벡터(v)들은 임베딩 공간 상에서 상대적인 거리를 가질 수 있다. 앞서 설명한 것과 같이 제1 임베딩 벡터(v)는 스타일과 촬영객체에 따라 경향성을 가지므로, 동일한 스타일 및 동일한 촬영객체를 갖는 제1 학습용 흑백사진(11)들로부터 생성된 제1 임베딩 벡터(v)들은 군집을 이룰 수 있다.

[0066] 본 발명은 기본적으로 스타일에 따라 서로 다른 흑백사진을 생성하되, 객체별로 나타나는 차이를 부가적으로 고려하는 기술로, 이를 위해 프로세서는 제1 임베딩 벡터(v)가 동일 스타일에 대해서는 가깝게 위치하되 동일 스타일에서도 다른 촬영객체에 대해서는 제1 임베딩 벡터(v)가 멀리 위치하도록, 스타일에 우선한 계층적(hierarchical) 손실함수를 설정할 수 있다.

[0067] 다층 퍼셉트론(130)은 전술한 손실함수가 최소가 되도록 비지도 학습될 수 있고, 이에 따라 도 4에 도시된 바와 같이 같은 스타일에 대한 제1 임베딩 벡터(v)들을 가깝게 위치하면서도 다른 촬영객체에 대한 제1 임베딩 벡터(v)들을 일정 거리를 두고 위치할 수 있다.

[0068] 프로세서는 제1 임베딩 벡터(v)가 형성하는 각 군집을 대표하는 복수의 프록시 벡터(p)를 추출할 수 있다(S20).

[0069] 앞서 설명한 바와 같이 제1 임베딩 벡터(v)들은 스타일과 촬영객체에 따라 군집을 형성할 수 있다. 도 4의 예시에서는 4개의 스타일(RGB2Gray, 전문가 A, B, C)과 4개의 촬영객체(인간, 비인간, 건물 및 자연)에 대해 제1 임베딩 벡터(v)가 생성되었으므로 총 16개의 군집이 형성될 수 있다.

[0070] 프로세서는 앞서 생성된 제1 임베딩 벡터(v)들로부터 각 군집을 대표하는 프록시 벡터(p)를 추출할 수 있고, 추출된 프록시 벡터(p)를 프록시 DB(140)에 저장할 수 있다. 이에 따라 총 16개의 프록시 벡터(p)가 프록시 DB(140)에 저장될 수 있다.

[0071] 이 때, 프록시 벡터(p)는 제1 임베딩 벡터(v)와 같은 차원(dimension)의 벡터로, 각 군집 내 제1 임베딩 벡터(v)들은 임베딩 공간 상에서 프록시 벡터(p) 주위에 위치할 수 있다. 이러한 프록시 벡터(p)는 각 군집 내 제1

임베딩 벡터(v)의 평균값(mean) 또는 중앙값(median)으로 결정될 수도 있고, 이 외에도 대표값을 결정할 수 있는 다양한 방법에 의해 결정될 수도 있다.

[0072] 한편, 전술한 프로세서의 계층적 손실함수 설정 동작은 프록시 벡터(p)에 대해 수행될 수 있다. 다시 말해, 프로세서는 프록시 벡터(p)가 동일 스타일에 대해서는 가깝게 위치하되 동일 스타일에서도 다른 촬영객체에 대해서는 프록시 벡터(p)가 멀리 위치하도록 계층적 손실함수를 설정할 수도 있다.

[0073] 전술한 단계(S10) 및 단계(S20)을 통해 제1 신경망 네트워크(100)는 학습될 수 있다. 다시 도 3을 참조하여 정리하면, 제1 학습용 흑백사진(11)이 제1 신경망 네트워크(100)에 입력되었을 때, 스타일 분류 신경망(110) 및 객체 분류 신경망(120)은 제1 학습용 흑백사진(11)에 대해 미리 라벨링(labeling)된 클래스(스타일 및 촬영객체)를 출력하도록 지도 학습될 수 있고, 다층 퍼셉트론(130)은 앞서 설명한 계층적 손실함수가 최소가 되도록 비지도 학습될 수 있다.

[0074] 다음으로, 제2 신경망 네트워크(200)를 학습시키는 과정, 즉 단계(S30) 내지 단계(S60)을 구체적으로 설명하도록 한다.

[0075] 제1 신경망 네트워크(100)의 학습이 완료되면, 프로세서는 학습용 컬러사진(20)을 제2 학습용 흑백사진(21)으로 변환하고, 제2 학습용 흑백사진(21)을 제1 신경망 네트워크(100)에 입력하여 제2 임베딩 벡터(v)를 생성할 수 있다(S30). 여기서, 학습용 컬러사진(20)은 전문가들의 보정 대상이 된 원본 사진일 수 있다.

[0076] 도 5를 참조하면, 프로세서는 학습용 컬러사진(20)을 제2 학습용 흑백사진(21)으로 변환할 수 있고, 이 때 변환 방식으로는 전문가의 스타일이나 촬영객체를 전혀 고려하지 않는 방식이 이용될 수 있다. 예컨대, 프로세서는 학습용 컬러사진(20)의 계조값을 단순히 그레이스케일(grayscale)로 변환할 수 있고, 이를 위해 RGB2Gray 함수를 이용할 수 있다.

[0077] 이어서, 프로세서는 제2 학습용 흑백사진(21)을 제1 신경망 네트워크(100)에 입력할 수 있다. 단계(S10) 및 단계(S20)을 통해 제1 신경망 네트워크(100)는 흑백사진으로부터 임베딩 벡터(v)를 생성하도록 학습된 바, 프로세서는 제1 신경망 네트워크(100)를 통해 제2 학습용 흑백사진(21)에 대한 제2 임베딩 벡터(v)를 생성할 수 있다.

[0078] 구체적으로 도 3을 다시 참조하면, 프로세서는 제2 학습용 흑백사진(21)을 스타일 분류 신경망(110)과 객체 분류 신경망(120)에 각각 입력할 수 있고, 두 분류 신경망(110, 120)의 출력단에 연결된 다층 퍼셉트론(130)의 출력을 제2 임베딩 벡터(v)로 결정할 수 있다.

[0079] 한편, 프로세서는 복수의 프록시 벡터(p) 중 학습용 컬러사진(20)에 대한 랜덤 스타일과 학습용 컬러사진(20)의 촬영객체에 대응하는 학습용 프록시 벡터(p)를 식별할 수 있다(S40).

[0080] 여기서 랜덤 스타일은 학습용 컬러사진(20)에 적용하고자 하는 스타일로 앞서 예시한 전문가 A, B, C 중 어느 하나의 스타일일 수 있고, 학습 과정에서는 프로세서에 의해 임의로 설정될 수 있다. 또한, 학습용 프록시 벡터(p)는 앞서 제1 신경망 네트워크(100) 학습 과정에서 추출된 프록시 벡터(p)들 중에서 제2 신경망 네트워크(200)의 학습에 이용되는 프록시 벡터(p)를 의미할 수 있다.

[0081] 프로세서는 프록시 DB(140)에 저장된 복수의 프록시 벡터(p) 중에서, 랜덤 스타일과 학습용 컬러사진(20) 내 촬영객체에 대응하는 어느 한 학습용 프록시 벡터(p)를 식별할 수 있다. 이 때, 촬영객체는 제1 신경망 네트워크(100), 구체적으로는 객체 분류 신경망(120)을 통해 식별할 수 있다.

[0082] 예를 들어, 도 3 내지 도 5를 함께 참조하면 랜덤 스타일이 전문가 A의 스타일(클래스 1)이고 객체 분류 신경망(120)을 통해 식별된 촬영객체가 비인간(클래스 1)인 경우, 프로세서는 프록시 DB(140)에 저장된 16개의 프록시 벡터(p) 중에서 클래스 1의 스타일과 클래스 1의 촬영객체에 대응하는 어느 한 학습용 프록시 벡터(p)를 식별할 수 있다.

[0083] 이어서, 프로세서는 학습용 컬러사진(20), 학습용 프록시 벡터(p) 및 제2 임베딩 벡터(v)를 제2 신경망 네트워크(200)에 입력하여 제2 임베딩 벡터(v)의 분포를 학습용 프록시 벡터(p)의 분포로 변환할 수 있다(S50).

[0084] 제2 학습용 흑백사진(21)은 앞서 설명한 바와 같이 스타일이나 촬영객체가 고려되지 않은 단순 그레이스케일로의 변환을 통해 생성된 것이므로, 제2 학습용 흑백사진(21)에 기초하여 생성된 제2 임베딩 벡터(v)는 스타일 또는 촬영객체에 따른 경향성을 갖지 않을 수 있다. 이와 달리, 학습용 프록시 벡터(p)는 랜덤 스타일, 즉 학습용 컬러사진(20)에 적용하고자 하는 스타일과 촬영객체에 대해 높은 경향성을 가질 수 있다.

[0085] 스타일과 촬영객체에 따른 전문가의 보정 방식을 모사하기 위해, 프로세서는 제2 신경망 네트워크(200)를 이용

하여 경향성이 없는 제2 임베딩 벡터(v)의 분포를 학습용 프록시 벡터(p)의 분포로 변환할 수 있다.

[0086] 다시 도 5를 참조하면, 제2 신경망 네트워크(200)는 학습용 컬러사진(20)으로부터 특징을 추출하고, 제2 임베딩 벡터(v)의 분포를 학습용 프록시 벡터(p)의 분포로 변환하는 인코더(210)를 포함할 수 있다.

[0087] 인코더(210)는 제2 임베딩 벡터(v) 및 학습용 프록시 벡터(p)의 평균 벡터와 분산 벡터를 각각 산출하여 분포를 정의하고, 제2 임베딩 벡터(v)의 분포를 학습용 프록시 벡터(p)의 분포로 변환할 수 있다.

[0088] 한편, 제2 임베딩 벡터(v)와 학습용 프록시 벡터(p)는 분포를 정의하기 어려운 데이터 구조를 가질 수 있는데, 이 경우, 제2 신경망 네트워크(200)는 제2 임베딩 벡터(v)와 학습용 프록시 벡터(p)로부터 분포 정의가 가능한 특징맵(feature map)을 각각 추출하는 완전 연결 레이어(Fully Connected Layer; FCL)를 더 포함할 수 있다.

[0089] 이 때, 인코더(210)는 제2 임베딩 벡터(v)에 대한 특징맵의 분포를 학습용 프록시 벡터(p)의 특징맵의 분포로 변환할 수 있다. 구체적으로, 인코더(210)는 하기 [수학식 1]에 따라 분포를 변환할 수 있다.

[0090] [수학식 1]

[0091] $$\mathcal{F}' = \sigma_t\left(\frac{\mathcal{F}-\mu_s}{\sigma_s}\right)+\mu_t$$

[0092] ($\mathcal{F}'$는 변환된 제2 임베딩 벡터(v)에 대한 특징맵, $\mathcal{F}$는 제2 임베딩 벡터(v)에 대한 특징맵, μ_s,μ_t는 각각 제2 임베딩 벡터(v)에 대한 특징맵의 평균 및 분산, μ_t,σ_t는 각각 학습용 프록시 벡터(p)에 대한 특징맵의 평균 및 분산)

[0093] 이어서, 프로세서는 제2 임베딩 벡터(v)의 변환된 분포에 대응하는 픽셀별 가중치를 학습용 컬러사진(20)에 적용하여 탈색 이미지를 생성할 수 있다(S60).

[0094] 프로세서는 앞서 학습용 컬러사진(20)으로부터 추출된 특징과, 변환된 제2 임베딩 벡터(v)의 분포로부터 픽셀별 가중치를 생성할 수 있다. 이를 위해, 프로세서는 디코더(220)를 이용할 수 있다.

[0095] 다시 도 5를 참조하면, 제2 신경망 네트워크(200)는 인코더(210)에서 추출된 특징과 변환된 제2 임베딩 벡터(v)의 분포를 입력받고, 이에 기초하여 픽셀별 가중치를 출력하는 디코더(220)를 포함할 수 있다.

[0096] 디코더(220)에서 출력되는 픽셀별 가중치는 학습용 컬러사진(20)의 탈색을 위해, 학습용 컬러사진(20) 내 각 픽셀의 계조값에 적용되는 가중치의 집합일 수 있다. 이러한 가중치는 각각의 픽셀별로 달리 결정될 수도 있다. 다만, 이 경우 학습에 이용된 컬러사진에 대한 과적합(overfitting)이 발생할 수 있으므로, 가중치는 유사한 계조값을 갖는 픽셀들에 대해서는 동일하게 결정될 수도 있다.

[0097] 일 예에서, 픽셀별 가중치는 학습용 컬러사진(20)의 픽셀별 계조값에 대한 바이레터럴 그리드(bilateral grid) 형태의 가중치를 포함할 수 있다.

[0098] 도 6을 참조하면, 바이레터럴 그리드는 이미지 학습용 컬러사진(20)의 좌표(x축 및 y축 좌표)에 대응하는 계조값의 크기(intensity)로 정의되는 3차원 그리드일 수 있다. 이 때, 하나의 그리드는 하나의 가중치에 매핑될 수 있고, 유사한 계조값을 갖는 픽셀들은 동일한 그리드에 포함될 수 있다. 이에 따라, 유사한 계조값을 갖는 픽셀들에 대해서는 동일한 가중치가 설정될 수 있다.

[0099] 이와 같이, 바이레테럴 그리드 형태의 가중치를 이용함으로써 본 발명은 동일한 계조값을 갖는 인접 픽셀에 다른 가중치가 적용되는 것을 방지할 수 있고, 이를 통해 학습용 컬러사진(20)에 대한 과적합을 억제할 수 있을 뿐만 아니라 탈색 과정에 일관성(consistency)을 보장할 수 있다.

[0100] 다시 도 5를 참조하면, 프로세서는 전술한 방법에 따라 결정된 픽셀별 가중치를 학습용 컬러사진(20)의 픽셀별 계조값에 곱하여 탈색 이미지(decolorized image)를 생성할 수 있다. 이 때, 제2 신경망 네트워크(200)는 앞서 생성된 탈색 이미지와, 학습용 컬러사진(20)을 랜덤 스타일에 따라 변환한 실측(Ground Truth; GT) 흑백사진 간의 차이가 최소가 되도록 학습될 수 있다.

[0101] 여기서 실측 흑백사진(Ground truth image)은 랜덤 스타일에 따라 전문가가 직접 보정한 흑백사진일 수 있다. 앞서 예시한 것과 같이 랜덤 스타일은 전문가 A, B, C 중 어느 하나의 스타일일 수 있고, 각각의 학습용 컬러사진(20)은 전문가 A, B, C에 의해 미리 보정될 수 있다. 이 때, 각 전문가가 보정한 흑백사진이 실측 흑백사진일 수 있다.

[0102] 전술한 학습을 위해 프로세서는 모든 학습용 컬러사진(20)에 대한 탈색 이미지와 실측 흑백사진의 차이값들의 합을 손실함수로 설정할 수 있고, 제2 신경망 네트워크(200), 구체적으로는 인코더(210)와 디코더(220)는 해당 손실함수가 최소가 되도록 학습될 수 있다.

[0103] 전술한 단계(S30) 내지 단계(S60)을 통해 제2 신경망 네트워크(200)는 학습될 수 있다. 한편, 제2 신경망 네트워크(200)의 학습에는 제1 신경망 네트워크(100)에 의해 결정되는 임베딩 벡터(v) 및 프록시 벡터(p)가 이용되므로, 제2 신경망 네트워크(200)의 학습 과정은 제1 신경망 네트워크(100)의 학습이 완료된 이후에 수행될 수 있다.

[0104] 도 7에 도시된 전체 신경망 모델을 참조하여 정리하면, 먼저 프로세서는 제1 학습용 흑백사진(11)을 제1 신경망 네트워크(100)에 입력하여 제1 임베딩 벡터(v)를 생성하고, 군집화된 제1 임베딩 벡터(v)를 대표하는 프록시 벡터(p)를 추출하는 과정을 통해 제1 신경망 네트워크(100)를 학습시킬 수 있다.

[0105] 이 때, 제1 신경망 네트워크(100) 내 스타일 분류 신경망(110) 및 객체 분류 신경망(120)은 제1 학습용 흑백사진(11)과 이에 대응하는 스타일 및 촬영객체에 의해 학습될 수 있고, 제1 신경망 네트워크(100) 내 다층 퍼셉트론(130)은 프록시 벡터(p)가 동일 스타일에 대해서는 가깝게 위치하되 동일 스타일에서도 다른 촬영객체에 대해서는 멀리 위치하도록 설정된 계층적 손실함수(L_{HPL})에 따라 학습될 수 있다.

[0106] 제1 신경망 네트워크(100)의 학습이 완료되면, 프로세서는 제1 신경망 네트워크(100)의 모든 파라미터(가중치(weight) 및 바이어스(bias))를 프리징(freezing)하고, 학습용 컬러사진(20)을 제2 학습용 흑백사진(21)으로 변환하여 제1 신경망 네트워크(100)에 입력함으로써 제2 임베딩 벡터(v)를 생성할 수 있다.

[0107] 이어서, 프로세서는 학습용 컬러사진(20)에 대해 설정된 랜덤 스타일과 촬영객체에 대응하는 학습용 프록시 벡터(p)를 식별하고, 학습용 컬러사진(20), 학습용 프록시 벡터(p) 및 제2 임베딩 벡터(v)를 제2 신경망 네트워크(200)에 입력하여 탈색 이미지를 생성하는 과정을 통해 제2 신경망 네트워크(200)를 학습시킬 수 있다.

[0108] 이 때, 제2 신경망 네트워크(200) 내 인코더(210)와 디코더(220)는 탈색 이미지와, 랜덤 스타일에 따라 보정된 학습용 컬러사진(20), 즉 실측 흑백사진의 차이에 비례하는 손실함수(L_2)에 따라 학습될 수 있다.

[0109] 제1 및 제2 신경망 네트워크(100, 200)의 학습이 모두 완료되면, 프로세서는 제1 및 제2 신경망 네트워크(100, 200)를 이용하여 타겟 컬러사진을 타겟 스타일에 대응하는 흑백사진으로 변환할 수 있다(S70).

[0110] 여기서, 타겟 컬러사진은 학습에 이용되지 않은 사진으로, 사용자가 흑백사진으로의 변환을 위해 제공하는 사진일 수 있다. 한편, 타겟 스타일은 사용자가 타겟 컬러사진에 적용하고자 하는 스타일로, 앞서 예시한 전문가 A, B, C 중 어느 하나의 스타일에 대응할 수 있다.

[0111] 도 7을 참조하여 흑백사진 생성 동작을 설명하면, 프로세서는 타겟 컬러사진을 그레이스케일 이미지로 변환(RGB2Gray)하여 제1 신경망 네트워크(100)에 입력할 수 있다. 이에 따라, 제1 신경망 네트워크(100)는 그레이스케일 이미지에 대한 임베딩 벡터(v)를 생성할 수 있다.

[0112] 이어서, 프로세서는 타겟 스타일(도 7의 랜덤 스타일 대신에 입력됨)에 대응하는 프록시 벡터(p)를 프록시 DB(140)로부터 식별할 수 있다. 구체적으로, 프로세서는 제1 신경망 네트워크(100) 내 객체 분류 신경망(120)을 통해 그레이스케일 이미지 내 촬영객체를 식별하고, 프록시 DB(140)에 저장된 복수의 프록시 벡터(p) 중 타겟 스타일과 촬영객체에 대응하는 어느 한 프록시 벡터(p)를 식별할 수 있다.

[0113] 프로세서는 타겟 컬러사진과 앞서 식별된 어느 한 프록시 벡터(p)를 제2 신경망 네트워크(200)에 입력할 수 있다. 이에 따라, 제2 신경망 네트워크(200)는 제1 신경망 네트워크(100)에서 생성된 임베딩 벡터(v)의 분포를 프록시 벡터(p)의 분포로 변환하고, 변환된 분포에 대응하는 픽셀별 가중치를 타겟 컬러사진에 적용함으로써 타겟 스타일에 대응하는 흑백사진을 생성할 수 있다.

[0114] 도 8을 참조하면, 본 발명의 동작에 따라 신경망 모델은 동일한 컬러사진(input)에 대한 서로 다른 스타일별 흑백사진을 생성할 수 있고, 신경망 모델의 보정 방식은 컬러사진에 포함된 객체(인간, 비인간, 건물, 자연)에 따라 달라지는 것을 확인할 수 있다.

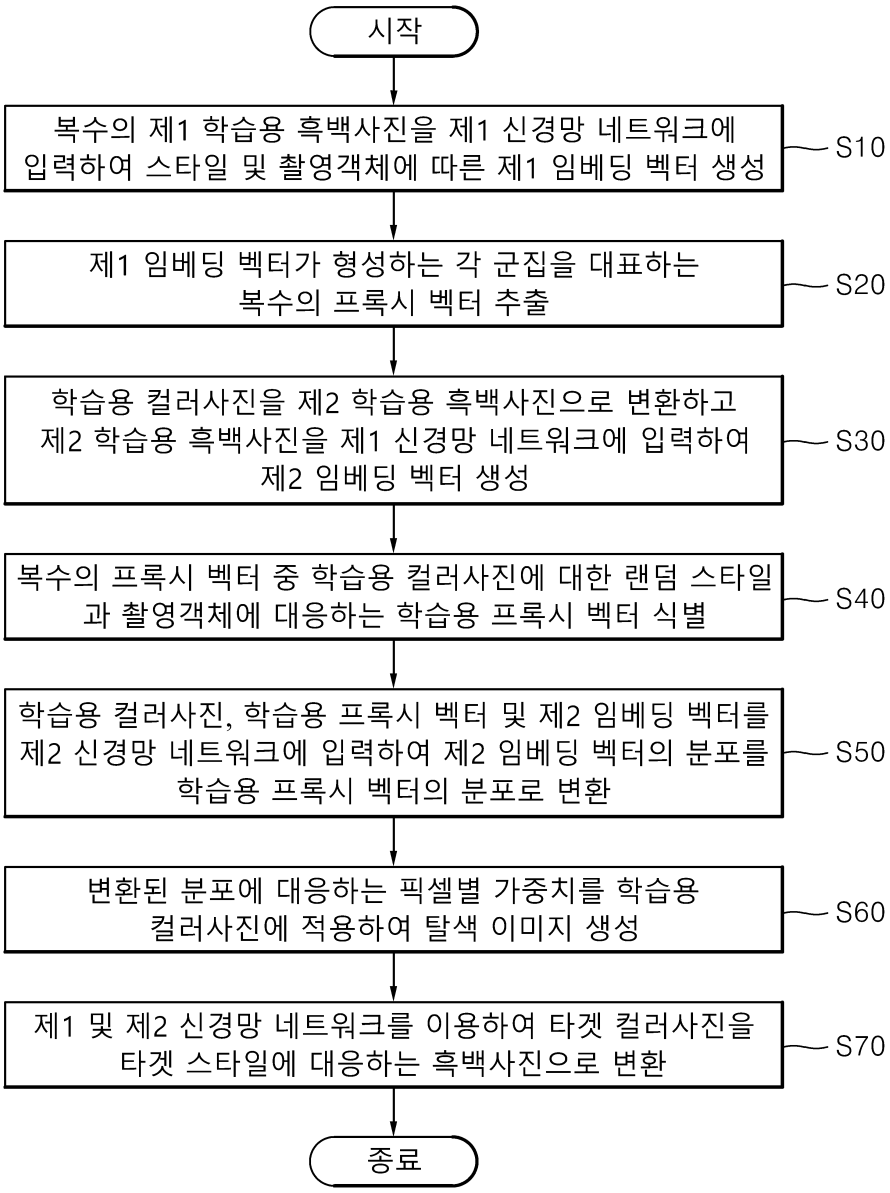
[0115] 전술한 바와 같이, 본 발명은 전문가들의 흑백사진 변환 스타일을 모사할 수 있도록 신경망 모델을 학습시킴으로써, 딥러닝 모델을 이용한 흑백사진 생성 동작에서 심미적/미학적 향상을 달성할 수 있다.

[0116] 이상과 같이 본 발명에 대해서 예시한 도면을 참조로 하여 설명하였으나, 본 명세서에 개시된 실시 예와 도면에

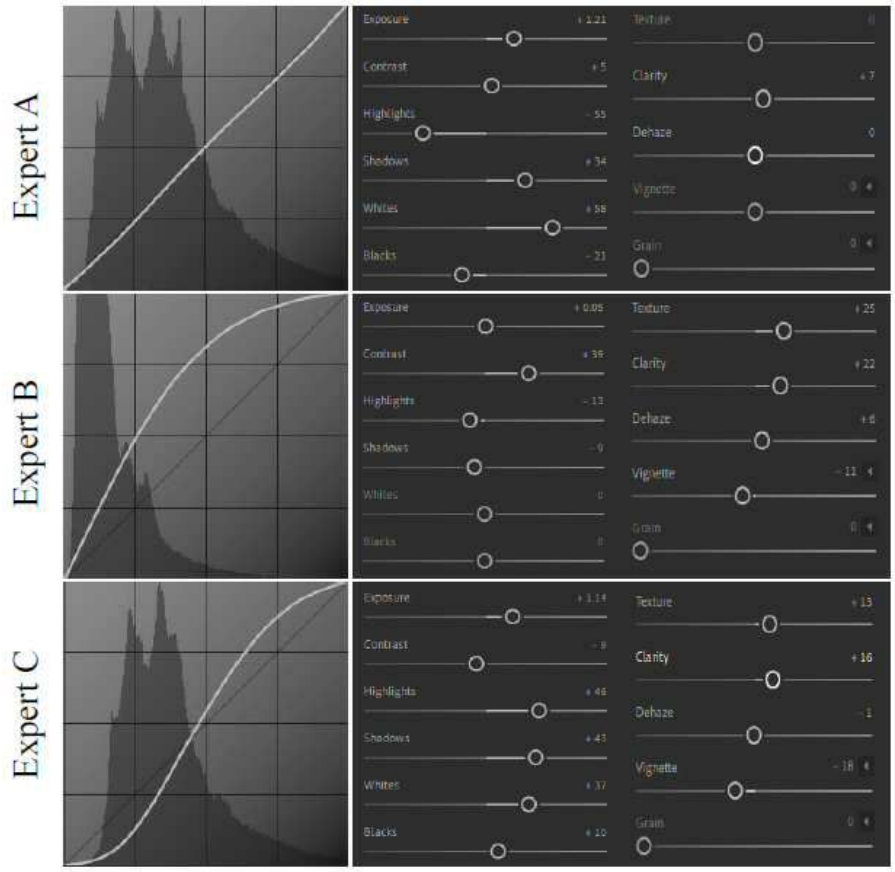
의해 본 발명이 한정되는 것은 아니며, 본 발명의 기술사상의 범위 내에서 통상의 기술자에 의해 다양한 변형이 이루어질 수 있음은 자명하다. 아울러 앞서 본 발명의 실시 예를 설명하면서 본 발명의 구성에 따른 작용 효과를 명시적으로 기재하여 설명하지 않았을 지라도, 해당 구성에 의해 예측 가능한 효과 또한 인정되어야 함은 당연하다.

도면

도면1



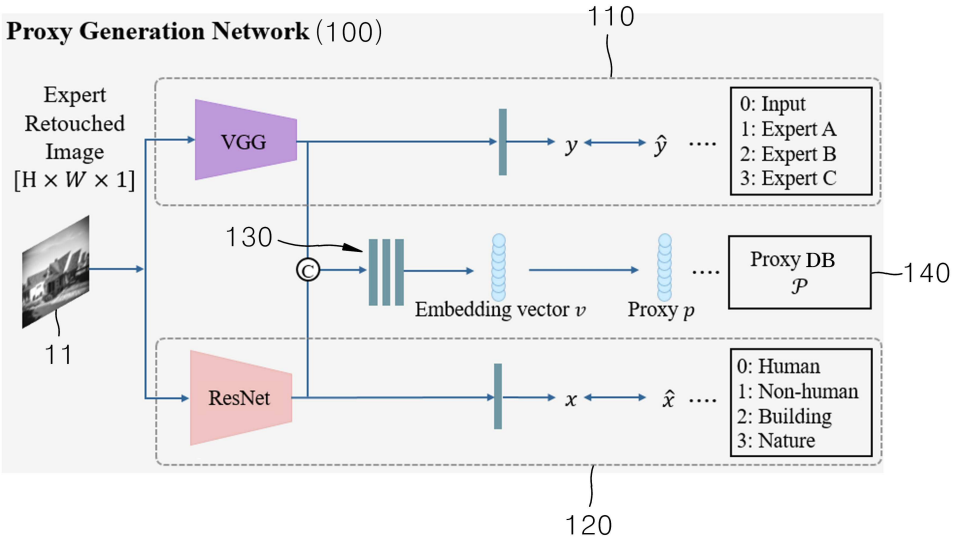
도면2a


Expert A
Exposure
Contrast
Highlights
Shadows
Whites
Blacks
Texture
Clarity
Dehaze
Vignette
Grain
Expert B
Exposure
Contrast
Highlights
Shadows
Whites
Blacks
Texture
Clarity
Dehaze
Vignette
Grain
Expert C
Exposure
Contrast
Highlights
Shadows
Whites
Blacks
Texture
Clarity
Dehaze
Vignette
Grain

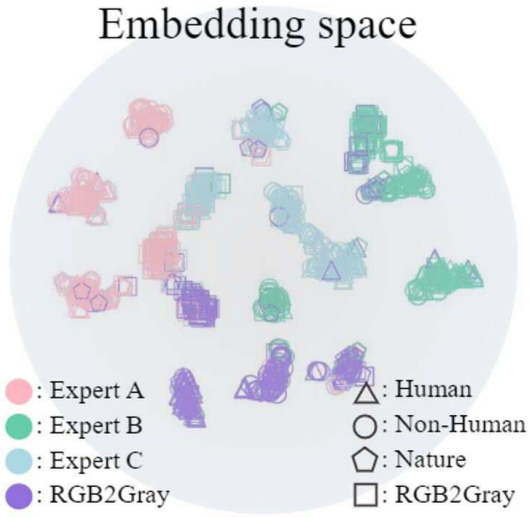
도면2b


Expert A
Expert B
Expert C

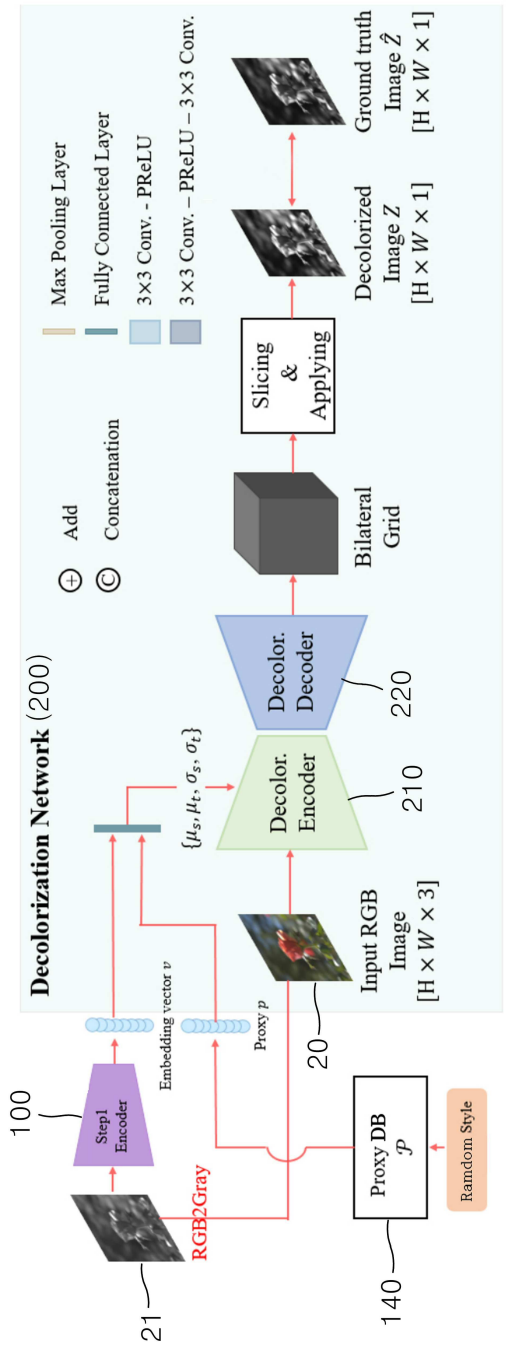
도면3


Proxy Generation Network (100)
110
Expert Retouched Image [H × W × 1]
VGG
y
ŷ ....
0: Input
1: Expert A
2: Expert B
3: Expert C
130
Embedding vector v
Proxy p
....
Proxy DB
P
140
11
ResNet
x
x̂ ....
0: Human
1: Non-human
2: Building
3: Nature
120

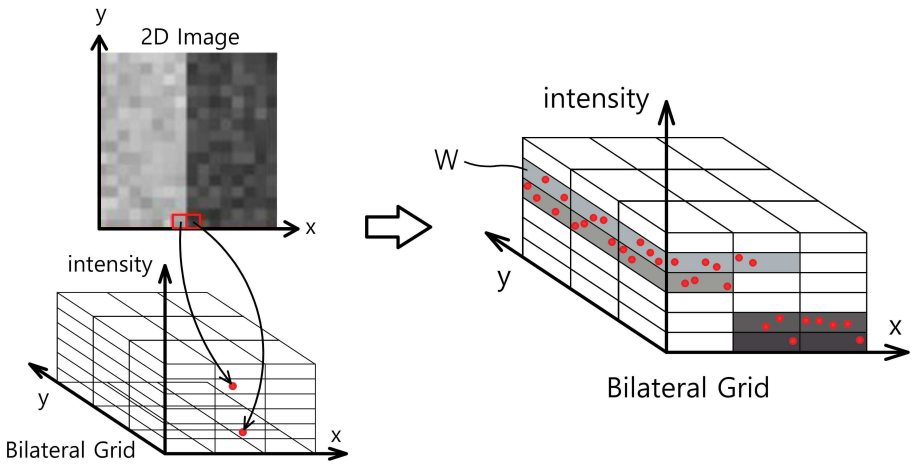
도면4


Embedding space
: Expert A
: Expert B
: Expert C
: RGB2Gray
: Human
: Non-Human
: Nature
: RGB2Gray

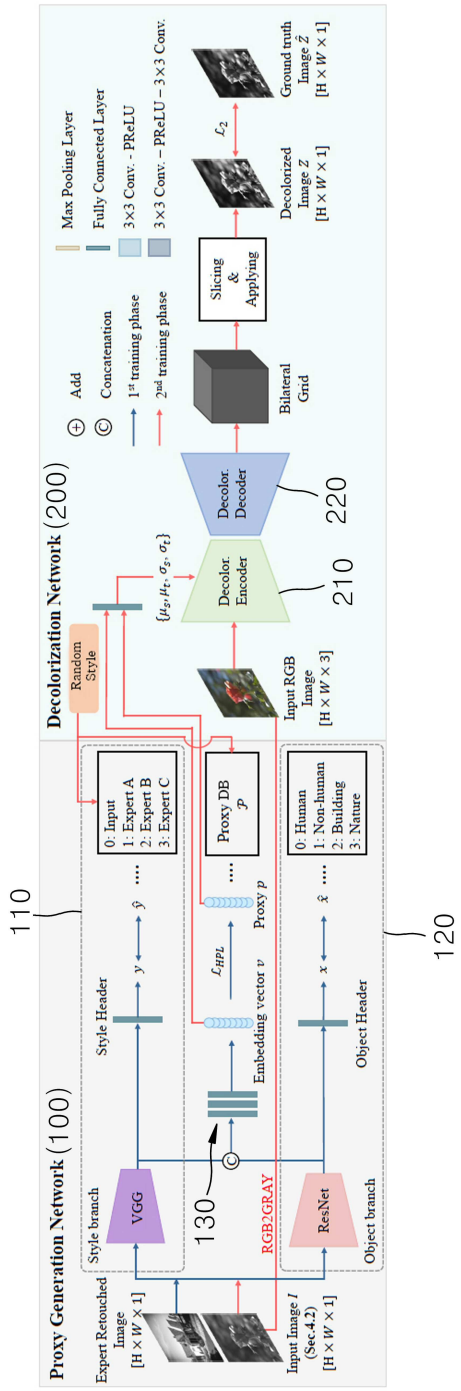
도면5


Decolorization Network (200)
Max Pooling Layer
Fully Connected Layer
3×3 Conv. - PReLU
3×3 Conv. – PReLU – 3×3 Conv.
Add
Concatenation
Input RGB Image [H × W × 3]
Bilateral Grid
Decolorized Image Z [H × W × 1]
Ground truth Image Ẑ [H × W × 1]
Slicing & Applying
Decolor. Encoder
Decolor. Decoder
{μs, μt, σs, σt}
Embedding vector v
Proxy p
Step1 Encoder
RGB2Gray
Proxy DB 𝒫
Ramdom Style
100
210
220
20
21
140

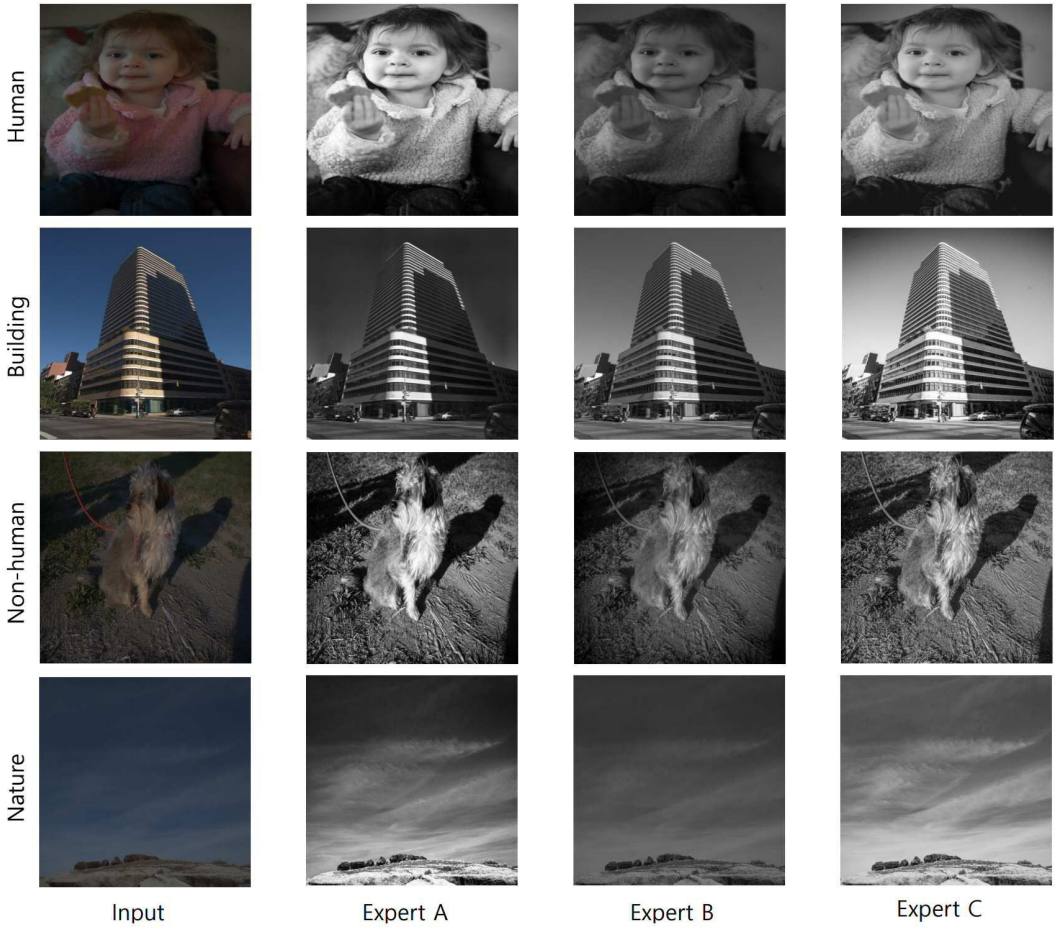
도면6


y
2D Image
x
intensity
y
Bilateral Grid
x
intensity
W
y
x
Bilateral Grid

도면7


Proxy Generation Network (100)
Decolorization Network (200)
Max Pooling Layer
Fully Connected Layer
3×3 Conv. - PReLU
3×3 Conv. – PReLU – 3×3 Conv.
Add
Concatenation
1st training phase
2nd training phase
Expert Retouched Image [H × W × 1]
Style branch
VGG
Style Header
0: Input
1: Expert A
2: Expert B
3: Expert C
Input Image I (Sec.4.2) [H × W × 1]
RGB2GRAY
ResNet
Object branch
Object Header
0: Human
1: Non-human
2: Building
3: Nature
Embedding vector v
Proxy p
Proxy DB P
Input RGB Image [H × W × 3]
Random Style
Decolor. Encoder
Decolor. Decoder
Bilateral Grid
Slicing & Applying
Decolorized Image Z [H × W × 1]
Ground truth Image Ẑ [H × W × 1]
110
120
130
210
220

도면8


Human
Building
Non-human
Nature
Input
Expert A
Expert B
Expert C