



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년03월08일

(11) 등록번호 10-2225024

(24) 등록일자 2021년03월03일

(51) 국제특허분류(Int. Cl.)

G06T 11/40 (2006.01) G06T 5/20 (2006.01)

G06T 9/00 (2019.01)

(52) CPC특허분류

G06T 11/40 (2013.01)

G06T 5/20 (2013.01)

(21) 출원번호 10-2019-0132702

(22) 출원일자 2019년10월24일

심사청구일자 2019년10월24일

(56) 선행기술조사문헌

KR1020170092595 A

(뒷면에 계속)

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

손광훈

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C125호(신촌동)

최성일

서울특별시 서대문구 연세로 50, 연세대학교 제3공학관 C129호(신촌동)

(74) 대리인

민영준

전체 청구항 수 : 총 14 항

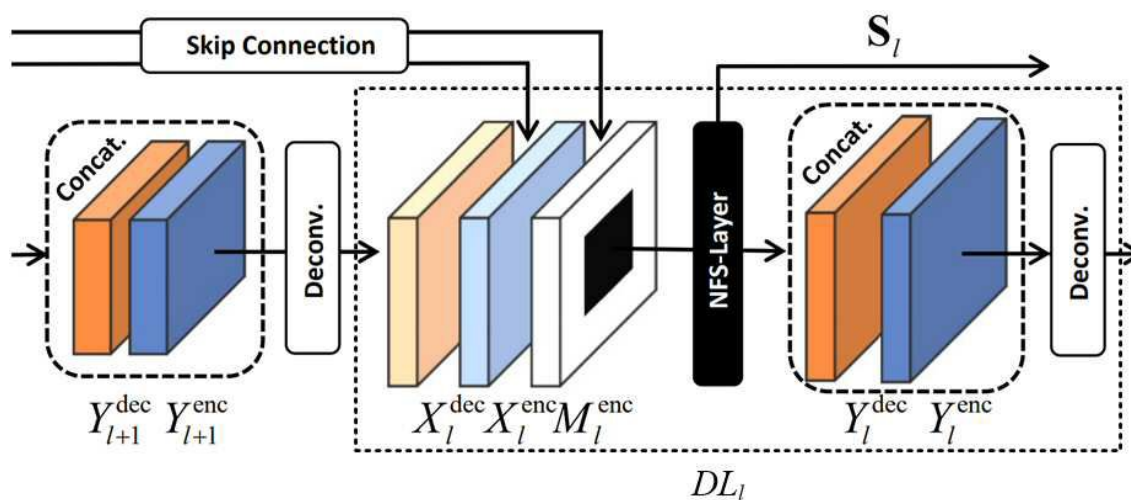
심사관 : 이병준

(54) 발명의 명칭 이미지 채움 장치 및 방법

(57) 요약

본 발명은 스킵 커넥션으로 인코더로부터 디코더로 유효하지 않은 영역의 특징이 전달되더라도, 유효 영역의 특징을 활용하여 유효하지 않은 영역의 특징을 채울 수 있는 비로컬 특징 합성 레이어를 적용하여 비로컬 표현자를 획득하고, 비로컬 표현자를 기반으로 누락된 영역을 재구성함으로써 시각적 오류를 저감하고 의미적으로 일관성 있게 누락된 영역을 채울 수 있는 이미지 채움 장치 및 방법을 제공할 수 있다.

대표도 - 도5



(52) CPC특허분류

G06T 9/00 (2019.01)

G06T 2207/20081 (2013.01)

(56) 선행기술조사문헌

KR1020190039460 A

US20180204111 A1*

US20190295228 A1*

HONGYU LIU et al. "Coherent Semantic Attention for Image Inpainting", Hunan University[online](2019.07.04.)*

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호 1711080916

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원(한국연구재단부설)

연구사업명 정보통신방송연구개발사업

연구과제명 스마트카 다중 센서와 딥러닝을 이용한 초정밀 내추릴 3D 뷰 생성 기술 개발 (창조
씨앗형 2단계)(3/5)

기 여 율 1/1

과제수행기관명 연세대학교 산학협력단

연구기간 2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

마스크에 의해 비유효 영역이 지정된 입력 이미지를 인가받아, 미리 학습된 패턴 추정 방식에 따라 입력 이미지 또는 이전 단계에서 획득된 특징맵과 마스크를 함께 인코딩하여 특징맵을 획득하고, 상기 마스크를 기지정된 방식으로 업데이트 하는 인코더; 및

단계적으로 연결된 다수의 디코딩 레이어를 구비하여, 상기 인코더에서 최종 획득된 특징맵과 단계적으로 획득된 특징맵 및 업데이트된 마스크로부터 복원 특징맵을 획득하는 디코더를 포함하고,

상기 다수의 디코딩 레이어 각각은

상기 인코더에서 최종 획득된 특징맵 또는 이전단에서 획득된 복원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 픽셀별 주의도를 나타내는 특징 주의 행렬을 획득하고, 상기 특징 주의 행렬과 이전 단계에서 획득된 복원 특징맵을 결합하여 디코딩맵을 획득하며, 상기 인코더의 대응하는 단계에서 획득된 마스크와 상기 특징 주의 행렬을 이용하여 특징맵에서 비유효 영역에 채워지기에 적합한 픽셀을 유효 영역에서 탐색하기 위한 홀 채움 유사도 행렬을 획득하고, 홀 채움 유사도 행렬과 대응하는 단계에서 획득된 특징맵을 결합하여 인코딩맵을 획득하는 비로컬 특징 합성 레이어(Non-Local Feature Synthesis-Layer: 이하 NFS 레이어); 및

상기 디코딩맵과 상기 인코딩맵을 결합하고, 미리 학습된 가중치를 기반으로 디컨볼루션하여 복원 특징맵을 획득하는 디컨볼루션 레이어를 포함하며,

상기 NFS 레이어는

상기 인코더에서 최종 획득된 특징맵 또는 이전단에서 획득된 복원 특징맵에 대해 학습에 의해 획득된 서로 다른 3개의 가중치로 컨볼루션하여 3개의 특징을 추출하며, 추출된 3개의 특징 중 두개의 특징 사이의 상관도를 계산하여 상기 특징 주의 행렬을 획득하며, 상기 특징 주의 행렬을 나머지 특징과 곱하고 인가된 특징맵 또는 복원 특징맵을 더하여 상기 디코딩맵을 획득하는 디코딩맵 획득부; 및

상기 인코더의 대응하는 단계에서 획득된 마스크를 기지정된 크기로 슬라이스하고, 슬라이스된 마스크와 슬라이스된 마스크의 픽셀값을 반전하고 전치하여 획득되는 전치 반전 마스크를 곱하여 홀 채움 표시자를 획득하고, 상기 홀 채움 표시자에 상기 특징 주의 행렬을 곱하고 기지정된 방식으로 정규화하여 홀 채움 유사도 행렬을 획득하며, 상기 홀 채움 유사도 행렬과 인가된 특징맵을 곱하고, 그 결과에 다시 인가된 특징맵을 더하여 상기 인코딩맵을 획득하는 인코딩맵 획득부를 포함하는 이미지 채움 장치.

청구항 2

삭제

청구항 3

삭제

청구항 4

제1 항에 있어서, 상기 디코딩맵 획득부는

추출된 3개의 특징 중 하나의 특징을 전치하고, 전치된 특징을 다른 하나의 특징을 곱한 후, 소프트 맥스 함수로 정규화하여 상기 특징 주의 행렬을 획득하는 이미지 채움 장치.

청구항 5

제4 항에 있어서, 상기 디코딩맵 획득부는

상기 특징 주의 행렬과 나머지 특징에 학습에 의해 결정되는 스케일 변수를 함께 곱하는 이미지 채움 장치.

청구항 6

제1 항에 있어서, 상기 인코딩맵 획득부는

상기 홀 채움 표시자에 상기 특징 주의 행렬을 원소별 곱셈하고, 수학적식

$$\sum_j S_l(i, j) = \begin{cases} 1, & \text{if } i \in H, j \in V \\ 0, & \text{otherwise} \end{cases}$$

(여기서 H는 비유효 영역의 픽셀을 나타내고, V는 유효 영역의 픽셀을 나타낸다.)

에 따라 정규화하여, 상기 홀 채움 유사도 행렬(S_l)을 획득하는 이미지 채움 장치.

청구항 7

제1 항에 있어서, 상기 인코더는

상기 다수의 디코딩 레이어와 역순으로 단계적으로 연결되는 다수의 인코딩 레이어를 포함하고,

상기 다수의 인코딩 레이어는

상기 입력 이미지 또는 이전단에서 획득된 특징맵과 대응하는 마스크를 인가받고, 인가된 특징맵과 마스크에 대해 학습에 의해 획득된 가중치를 이용하여 기지정된 연산을 수행하여 특징맵을 획득하고, 상기 마스크를 미리 지정된 방식으로 업데이트하는 이미지 채움 장치.

청구항 8

제7 항에 있어서, 상기 다수의 인코딩 레이어는

인가된 특징맵(X_{l-1}^{enc})과 마스크(M_{l-1}^{enc})와 가중치(W_{l-1}^{enc})를 이용하여 수학적식

$$X_l^{enc}(u, v) = \frac{\sum_{u', v'} W_l^{enc}(u', v') X_{l-1}^{enc}(u + u', v + v') M_{l-1}^{enc}(u + u', v + v')}{\sum_{u', v'} M_{l-1}^{enc}(u + u', v + v')}$$

(여기서 u, v는 특징맵(X_{l-1}^{enc})에서의 좌표를 나타내고, u', v' 는 가중치(W_{l-1}^{enc}) 윈도우(R)에서의 좌표(u', v' ∈ R)를 나타낸다.)

에 따라 특징맵(X_l^{enc})을 획득하고,

상기 마스크(M_{l-1}^{enc})를 수학적식

$$M_l^{enc}(u, v) = \begin{cases} 1, & \text{if } \sum_{u', v'} M_{l-1}^{enc}(u + u', v + v') > \delta \\ 0, & \text{otherwise} \end{cases}$$

(여기서 δ는 유효 영역을 제어하기 위해 미리 설정된 문턱값을 나타낸다.)

에 따라 업데이트 하는 이미지 채움 장치.

청구항 9

제7 항에 있어서, 상기 이미지 채움 장치는

이미지와 상기 이미지에서 비유효 영역을 이진값으로 지정하는 마스크를 원소곱하여 상기 입력 이미지를 획득하는 이미지 획득부를 더 포함하는 이미지 채움 장치.

청구항 10

마스크에 의해 비유효 영역이 지정된 입력 이미지를 인가받아, 미리 학습된 패턴 추정 방식에 따라 단계적으로 입력 이미지와 마스크를 함께 인코딩하여 단계적으로 특징맵을 획득하고, 상기 마스크를 기지정된 방식으로 업데이트 하는 인코딩 단계; 및

상기 인코딩 단계에서 최종 획득된 특징맵과 단계적으로 획득된 특징맵 및 업데이트된 마스크 중 대응하는 단계에서 획득된 특징맵과 마스크를 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 단계적으로 복원 특징맵을 획득하는 디코딩 단계를 포함하고,

상기 디코딩 단계는

상기 인코딩 단계에서 최종 획득된 특징맵 또는 상기 디코딩 단계 내에서 이전 획득된 복원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 픽셀별 주의도를 나타내는 특징 주의 행렬을 획득하는 단계;

상기 특징 주의 행렬과 이전 단계에서 획득된 복원 특징맵을 결합하여 픽셀별 주의도가 가중된 디코딩맵을 획득하는 단계;

상기 인코딩 단계에서 획득된 대응하는 마스크와 상기 특징 주의 행렬을 이용하여 특징맵에서 비유효 영역에 채워지기에 적합한 픽셀을 유효 영역에서 탐색하기 위한 홀 채움 유사도 행렬을 획득하는 단계;

상기 홀 채움 유사도 행렬과 상기 인코딩 단계에서 획득된 대응하는 특징맵을 결합하여 인코딩맵을 획득하는 단계; 및

상기 디코딩맵과 상기 인코딩맵을 결합하고 미리 학습된 특징 복원 방식에 따라 디코딩하여 상기 복원 특징맵을 획득하는 단계를 포함하며,

상기 홀 채움 유사도 행렬을 획득하는 단계는

상기 인코딩 단계에서 획득된 대응하는 마스크를 기지정된 크기로 슬라이스하는 단계;

슬라이스된 마스크와 슬라이스된 마스크의 픽셀값을 반전하고 전치하여 획득되는 전치 반전 마스크를 곱하여 홀 채움 표시자를 획득하는 단계;

상기 홀 채움 표시자에 상기 특징 주의 행렬을 곱하고 기지정된 방식으로 정규화하는 단계를 포함하는 이미지 채움 방법.

청구항 11

삭제

청구항 12

제10 항에 있어서, 상기 특징 주의 행렬을 획득하는 단계는

상기 인코딩 단계에서 최종 획득된 특징맵 또는 이전에서 획득된 복원 특징맵에 대해 학습에 의해 획득된 서로 다른 3개의 가중치로 컨볼루션하여 3개의 특징을 추출하는 단계;

추출된 3개의 특징 중 하나의 특징을 전치하고, 전치된 특징을 다른 하나의 특징을 곱한 후, 소프트 맥스 함수로 정규화하는 단계를 포함하는 이미지 채움 방법.

청구항 13

제12 항에 있어서, 상기 복원 특징맵을 결합하는 단계는

상기 특징 주의 행렬과 나머지 특징 및 학습에 의해 결정되는 스케일 변수를 곱하고 인가된 특징맵 또는 복원 특징맵을 더하여 상기 디코딩맵을 획득하는 이미지 채움 방법.

청구항 14

삭제

청구항 15

삭제

청구항 16

제10 항에 있어서, 상기 정규화하는 단계는

상기 홀 채움 표시자에 상기 특징 주의 행렬을 원소별 곱셈하고, 수학적식

$$\sum_j S_l(i, j) = \begin{cases} 1, & \text{if } i \in H, j \in V \\ 0, & \text{otherwise} \end{cases}$$

(여기서 H는 비유효 영역의 픽셀을 나타내고, V는 유효 영역의 픽셀을 나타낸다.)

에 따라 정규화하여, 상기 홀 채움 유사도 행렬(S_l)을 획득하는 이미지 채움 방법.

청구항 17

제10 항에 있어서, 상기 특징맵을 결합하는 단계는

상기 홀 채움 유사도 행렬과 인가된 특징맵을 곱하고, 그 결과에 다시 인가된 특징맵을 더하여 상기 인코딩맵을 획득하는 단계를 포함하는 이미지 채움 방법.

청구항 18

제10 항에 있어서, 상기 복원 특징맵을 획득하는 단계는

상기 디코딩맵과 상기 인코딩맵을 결합하고, 미리 학습된 가중치를 기반으로 디컨볼루션하여 상기 복원 특징맵을 획득하는 이미지 채움 방법.

청구항 19

제10 항에 있어서, 상기 인코딩 단계는

상기 입력 이미지 또는 이전 획득된 특징맵과 대응하는 마스크를 인가받고, 인가된 특징맵(X_{l-1}^{enc})과 마스크(M_{l-1}^{enc})와 가중치(W_l^{enc})를 이용하여 수학적식

$$X_l^{enc}(u, v) = \frac{\sum_{u', v'} W_l^{enc}(u', v') X_{l-1}^{enc}(u + u', v + v') M_{l-1}^{enc}(u + u', v + v')}{\sum_{u', v'} M_{l-1}^{enc}(u + u', v + v')}$$

(여기서 u, v는 특징맵(X_{l-1}^{enc})에서의 좌표를 나타내고, u', v'는 가중치(W_l^{enc}) 윈도우(R)에서의 좌표(u', v' ∈ R)를 나타낸다.)

에 따라 특징맵(X_l^{enc})을 획득하는 단계; 및

상기 마스크(M_{l-1}^{enc})를 수학적식

$$M_l^{enc}(u, v) = \begin{cases} 1, & \text{if } \sum_{u', v'} M_{l-1}^{enc}(u + u', v + v') > \delta \\ 0, & \text{otherwise} \end{cases}$$

(여기서 δ는 유효 영역을 제어하기 위해 미리 설정된 문턱값을 나타낸다.)

에 따라 업데이트 하는 단계를 포함하는 이미지 채움 방법.

발명의 설명

기술 분야

[0001] 본 발명은 이미지 채움 장치 및 방법에 관한 것으로, 비국부적 특징 동기를 이용한 이미지 채움 장치 및 방법에 관한 것이다.

배경 기술

[0002] 이미지 채움은 이미지에서 누락되지 않은 유효 영역에 대응하여 누락되거나 유효하지 않은 영역에 시각적 그리고 의미적(semantically)으로 대체 가능한 콘텐츠를 합성하는 작업을 의미한다. 이미지 채움은 이미지 인페인팅(Image inpainting), 이미지 홀 필링(image hole filling) 또는 이미지 컴플리션(image completion)으로도 알려져 있으며, 이미지에서 원하지 않는 객체 제거, 변형된 영역 복원 및 불필요 영역 수정 등의 많은 응용 분야에서 유용하게 적용될 수 있다.

[0003] 이미지 채움의 가장 기본적인 방법으로는 예제 기반 채움 기법(exemplar-based inpainting method)이 있다. 예제 기반 채움 기법은 이미지 내에서 누락되지 않은 유효 영역 중 누락된 영역과 가장 유사한 패치를 검색하고, 검색하여 가장 유사한 패치를 복사하여 채우는 방식으로 작업이 수행된다. 이러한 기법은 고주파수 텍스처 디테일을 복원하는 경우에 효과적이지만, 계산 복잡도가 높다. 뿐만 아니라 시멘틱 구조를 보존하기에 적합하지 않아 누락되지 않은 기존 영역에 대한 의미적 일관성을 유지하면서 누락된 영역을 채우기에 적합하지 않아 복잡한 사물과 장면에 대한 누락 영역은 채울 수 없다는 한계가 있다.

[0004] 한편 최근에는 대규모의 학습 데이터를 기반으로 인공 신경망에 이미지의 시멘틱 구조를 학습시킴으로써, 이미지 내에서 누락되지 않은 유효 영역과의 시멘틱 일관성을 유지하면서 누락된 영역을 채울 수 있는 다양한 기법이 제안된 바 있다. 그러나 인공 신경망을 이용하는 기법은 대부분 누락 영역의 픽셀을 기지정된 고정값으로 채운 후, 유효 영역과 누락 영역의 픽셀에 컨볼루션과 같은 동일한 연산을 적용한다. 이로 인해 획득되는 특징의 신뢰도가 낮아 채워진 영역과 기존 영역 사이의 색상 불일치, 흐름, 경계 주변의 아티피컬 에지 응답(artificial edge response)과 같은 시각적 오류가 나타날 수 있다는 문제가 있다.

[0005] 이에 인공 신경망의 각 레이어의 이미지 및 특징맵에서 유효하지 않은 픽셀에 대해 0의 값을, 유효 픽셀에 대해 1의 값을 할당하는 마스크를 적용하여 컨볼루션 연산 시에 유효 및 유효 픽셀을 구별하는 부분 컨볼루션 기법이 제안된 바 있다. 부분 컨볼루션 기법에서는 유효한 특징만을 추출되어 전파되도록 하고 유효하지 않은 픽셀이 전파되지 않도록 함으로써, 안정적인 채움 성능을 획득할 수 있도록 한다.

[0006] 기존의 부분 컨볼루션을 이용하는 기존의 기법에서는 이미지에서 유효한 특징을 단계적으로 추출하는 인코더와 인코더에서 추출된 특징으로부터 누락된 영역을 단계적으로 채워진 이미지로 복원하는 디코더를 포함한다. 이때 디코더의 각 단계는 인코더의 대응하는 단계에서 추출된 특징 및 마스크를 스킵 커넥션을 통해 인가받아 함께 합성하여 누락 영역을 채우며, 이로 인해 스킵 커넥션을 통해 인가된 특징 중 유효하지 않은 픽셀의 무의미한 특징이 전달되어 시각적 오류가 발생된다는 문제가 있다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국 등록 특허 제10-1539013호(2015.07.17 등록)

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은 이미지의 누락되지 않은 유효 영역에서 추출된 특징들 중 누락된 영역과의 거리에 무관하게 가장 의미적으로 유사한 특징을 융합하여 비로컬 표현자를 획득하고, 비로컬 표현자를 기반으로 누락된 영역을 재구성함으로써, 시각적 오류를 저감할 수 있는 이미지 채움 장치 및 방법을 제공하는데 있다.

[0009] 본 발명의 다른 목적은 스킵 커넥션으로 인코더로부터 디코더로 유효하지 않은 영역의 특징이 전달되더라도, 유

효 영역의 특징을 활용하여 유효하지 않은 영역의 특징을 재구성할 수 있는 비로컬 특징 합성 레이어를 적용하여 비로컬 표현자를 획득함으로써, 의미적으로 일관성 있게 누락된 영역을 재구성하여 채울 수 있는 이미지 채움 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

- [0010] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 이미지 채움 장치는 마스크에 의해 비유효 영역이 지정된 입력 이미지를 인가받아, 미리 학습된 패턴 추정 방식에 따라 단계적으로 입력 이미지와 마스크를 함께 인코딩하여 특징맵을 획득하고, 상기 마스크를 기지정된 방식으로 업데이트 하는 인코더; 및 상기 인코더에서 최종 획득된 특징맵과 단계적으로 획득된 특징맵 및 업데이트된 마스크 중 대응하는 단계에서 획득된 특징맵과 마스크를 인가받아 미리 학습된 특징 복원 방식에 따라 단계적으로 디코딩하여 복원 특징맵을 획득하는 디코더를 포함한다.
- [0011] 상기 디코더는 상기 인코더에서 최종 획득된 특징맵 또는 이전 단계에서 획득된 복원 특징맵에 픽셀별 주의도가 가중된 디코딩맵을 획득하고, 상기 인코더의 대응하는 단계에서 획득된 마스크와 상기 특징 주의 행렬을 이용하여 특징맵에서 비유효 영역에 채워지기에 적합한 픽셀을 유효 영역에서 탐색하기 위한 홀 채움 유사도 행렬을 획득하며, 홀 채움 유사도 행렬과 대응하는 단계에서 획득된 특징맵을 결합하여 인코딩맵을 획득하여 상기 디코딩맵과 상기 인코딩맵을 결합하고 미리 학습된 특징 복원 방식에 따라 디코딩하여 상기 복원 특징맵을 획득할 수 있다.
- [0012] 상기 디코더는 단계적으로 연결되어 상기 인코더에서 최종 획득된 특징맵 또는 이전단에서 획득된 복원 특징맵과 상기 인코더의 대응하는 단계에서 획득된 특징맵 및 마스크를 인가받아 복원 특징맵을 획득하는 다수의 디코딩 레이어를 포함하고,
- [0013] 상기 다수의 인코딩 레이어 각각은 상기 인코더에서 최종 획득된 특징맵 또는 이전단에서 획득된 복원 특징맵으로부터 미리 학습된 패턴 추정 방식에 따라 픽셀별 주의도를 나타내는 특징 주의 행렬을 획득하고, 상기 특징 주의 행렬과 이전 단계에서 획득된 복원 특징맵을 결합하여 디코딩맵을 획득하며, 상기 인코더의 대응하는 단계에서 획득된 마스크와 상기 특징 주의 행렬을 이용하여 특징맵에서 비유효 영역에 채워지기에 적합한 픽셀을 유효 영역에서 탐색하기 위한 홀 채움 유사도 행렬을 획득하고, 홀 채움 유사도 행렬과 대응하는 단계에서 획득된 특징맵을 결합하여 인코딩맵을 획득하는 비로컬 특징 합성 레이어(Non-Local Feature Synthesis-Layer: 이하 NFS 레이어); 및 상기 디코딩맵과 상기 인코딩맵을 결합하고, 미리 학습된 가중치를 기반으로 디컨볼루션하여 복원 특징맵을 획득하는 디컨볼루션 레이어를 포함할 수 있다.
- [0014] 상기 NFS 레이어는 상기 인코더에서 최종 획득된 특징맵 또는 이전단에서 획득된 복원 특징맵에 대해 학습에 의해 획득된 서로 다른 3개의 가중치로 컨볼루션하여 3개의 특징을 추출하며, 추출된 3개의 특징 중 두개의 특징 사이의 상관도를 계산하여 상기 특징 주의 행렬을 획득하며, 상기 특징 주의 행렬을 나머지 특징과 곱하고 인가된 특징맵 또는 복원 특징맵을 더하여 상기 디코딩맵을 획득하는 디코딩맵 획득부; 및 상기 인코더의 대응하는 단계에서 획득된 마스크를 기지정된 크기로 슬라이스하고, 슬라이스된 마스크와 슬라이스된 마스크의 픽셀값을 반전하고 전치하여 획득되는 전치 반전 마스크를 곱하여 홀 채움 표시자를 획득하고, 상기 홀 채움 표시자에 상기 특징 주의 행렬을 곱하고 기지정된 방식으로 정규화하여 홀 채움 유사도 행렬을 획득하며, 상기 홀 채움 유사도 행렬과 인가된 특징맵을 곱하고, 그 결과에 다시 인가된 특징맵을 더하여 상기 인코딩맵을 획득하는 인코딩맵 획득부를 포함할 수 있다.
- [0015] 상기 인코더는 상기 다수의 인코딩 레이어와 역순으로 단계적으로 연결되는 다수의 인코딩 레이어를 포함하고, 상기 다수의 인코딩 레이어는 상기 입력 이미지 또는 이전단에서 획득된 특징맵과 대응하는 마스크를 인가받고, 인가된 특징맵과 마스크에 대해 학습에 의해 획득된 가중치를 이용하여 기지정된 연산을 수행하여 특징맵을 획득하고, 상기 마스크를 미리 지정된 방식으로 업데이트할 수 있다.
- [0016] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 이미지 채움 방법은 마스크에 의해 비유효 영역이 지정된 입력 이미지를 인가받아, 미리 학습된 패턴 추정 방식에 따라 단계적으로 입력 이미지와 마스크를 함께 인코딩하여 단계적으로 특징맵을 획득하고, 상기 마스크를 기지정된 방식으로 업데이트 하는 인코딩 단계; 및
- [0017] 상기 인코딩 단계에서 최종 획득된 특징맵과 단계적으로 획득된 특징맵 및 업데이트된 마스크 중 대응하는 단계에서 획득된 특징맵과 마스크를 인가받아 미리 학습된 특징 복원 방식에 따라 디코딩하여 단계적으로 복원 특징맵을 획득하는 디코딩 단계를 포함한다.

[0018] 상기 디코딩 단계는 상기 인코딩 단계에서 최종 획득된 특징맵 또는 상기 디코딩 단계 내에서 이전 획득된 복원 특징맵에 픽셀별 주의도가 가중된 디코딩맵을 획득하는 단계; 상기 인코딩 단계에서 획득된 대응하는 마스크와 상기 특징 주의 행렬을 이용하여 특징맵에서 비유효 영역에 채워지기에 적합한 픽셀을 유효 영역에서 탐색하기 위한 홀 채움 유사도 행렬을 획득하고, 상기 홀 채움 유사도 행렬과 상기 인코딩 단계에서 획득된 대응하는 특징맵을 결합하여 인코딩맵을 획득하는 단계; 및 상기 디코딩맵과 상기 인코딩맵을 결합하고 미리 학습된 특징 복원 방식에 따라 디코딩하여 상기 복원 특징맵을 획득하는 단계를 포함할 수 있다.

발명의 효과

[0019] 따라서, 본 발명의 실시예에 따른 이미지 채움 장치 및 방법은 스킵 커넥션으로 인코더로부터 디코더로 유효하지 않은 영역의 특징이 전달되더라도, 유효 영역의 특징을 활용하여 유효하지 않은 영역의 특징을 채울 수 있는 비로컬 특징 합성 레이어를 적용하여 비로컬 표현자를 획득하고, 비로컬 표현자를 기반으로 누락된 영역을 재구성함으로써 시각적 오류를 저감하고 의미적으로 일관성 있게 누락된 영역을 채울 수 있다.

도면의 간단한 설명

[0020] 도 1은 본 발명의 일 실시예에 따른 이미지 채움 장치 및 방법이 이미지를 채우는 개념을 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시예에 따른 이미지 채움 장치의 개략적 구조를 나타낸다.

도 3은 도 2의 이미지 채움 장치의 개략적 동작을 설명하기 위한 도면이다.

도 4는 도 2의 이미지 채움 장치에서 인코딩 레이어의 상세 구성을 나타낸다.

도 5는 도 2의 이미지 채움 장치에서 디코더부의 상세 구성을 나타낸다.

도 6은 도 5의 디코딩 레이어에서 비로컬 특징 합성 레이어의 상세 구성을 나타낸다.

도 7은 본 발명의 일 실시예에 따른 이미지 채움 방법을 나타낸다.

도 8 및 도 9는 본 실시예에 따른 이미지 채움 방법에 따라 입력 이미지의 비유효 영역을 채운 결과를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0021] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.

[0022] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.

[0023] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0024] 도 1은 본 발명의 일 실시예에 따른 이미지 채움 장치 및 방법이 이미지를 채우는 개념을 설명하기 위한 도면이다.

[0025] 도 1에서 (a)는 채워져야 하는 누락 영역이 포함된 입력 이미지를 나타내고, (b)는 본 실시예의 이미지 채움 장치 및 방법에 의해 누락 영역이 채워진 출력 이미지를 나타내며, (c)는 누락 영역이 발생되지 않은 원본 이미지를 나타낸다.

[0026] 그리고 (d) 내지 (f)는 본 실시예의 이미지 채움 장치 및 방법이 (a)의 누락 영역의 3개의 픽셀 각각에 의미적으로 대응하는 특징을 갖는 유효 영역에서의 픽셀들을 추출하여 이미지를 채우는 개념을 나타낸다.

[0027] (d) 내지 (f)에 도시된 바와 같이, 본 실시예에 따른 이미지 채움 장치 및 방법은 이미지 내에서 채워져야 하는 비유효 영역의 픽셀과 유효 영역의 픽셀들 중 가장 유사한 특징을 갖는 적어도 하나의 픽셀들에 대한 특징을 복

사하여 비유효 영역의 픽셀을 채움으로써, (c)에 도시된 원본 이미지와 거의 동일한 (b)의 출력 이미지를 획득할 수 있도록 한다. 즉 본 실시예에 따른 이미지 채움 장치 및 방법은 비유효 영역 경계 주변의 유효 영역 특징보다 비유효 영역과의 거리에 무관하게 유효 영역에서의 특징 중 의미적으로 가장 유사한 특징을 이용하여 비유효 영역을 채우도록 하여 시각적 오류가 적고 의미적으로 일관성 있는 이미지를 획득할 수 있도록 한다.

[0028] 도 2는 본 발명의 일 실시예에 따른 이미지 채움 장치의 개략적 구조를 나타내고, 도 3은 도 2의 이미지 채움 장치의 개략적 동작을 설명하기 위한 도면이다. 그리고 도 4는 도 2의 이미지 채움 장치에서 인코더부의 상세 구성을 나타내고, 도 5는 도 2의 이미지 채움 장치에서 디코더부의 상세 구성을 나타내며, 도 6은 도 5의 디코딩 레이어에서 비로컬 특징 합성 레이어의 상세 구성을 나타낸다.

[0029] 도 2를 참조하면, 본 실시예에 따른 이미지 채움 장치는 인코딩부(100)와 디코딩부(200)를 포함한다. 인코딩부(100)는 일부 영역이 누락된 입력 이미지를 인가받고, 입력 이미지에서 누락되지 않은 유효 영역의 특징을 추출하고, 디코딩부(200)는 인코딩부(100)에서 추출된 특징을 기반으로 누락된 비유효 영역의 특징을 재구성하여 비유효 영역이 채워진 이미지를 획득한다.

[0030] 우선 인코딩부(100)를 살펴보면, 인코딩부(100)는 입력 이미지 획득부(110)와 다단 구조의 다수의 인코딩 레이어(EL1 ~ EL5)를 포함하는 인코더(120)로 구성된다. 여기서는 일례로 인코더(120)가 5개의 인코딩 레이어(EL1 ~ EL5)를 포함하는 것으로 도시하였으나, 인코딩 레이어의 개수는 다양하게 조절될 수 있다.

[0031] 입력 이미지 획득부(110)는 적어도 일부 영역이 누락된 입력 이미지(IN)를 획득한다. 여기서 입력 이미지 획득부(110)는 이미 일부 영역이 누락된 상태의 입력 이미지(IN)를 획득할 수도 있으나, 도 3에 도시된 바와 같이, 일반적인 이미지와 함께 이미지 내의 일부 영역을 제거하기 위한 마스크(MK)를 함께 인가받아 결합함으로써 입력 이미지를 생성할 수도 있다.

[0032] 상기한 바와 같이, 이미지 처리 분야에서는 객체 제거, 배경 추출, 변형 영역 복원 및 불필요 영역 수정 등의 목적으로 이미지의 일부 영역을 제거하고 제거된 영역을 제거되지 않은 영역과 유사하게 복원해야 하는 경우가 빈번하게 발생한다. 이에 마스크(MK)는 이미지 내에서 불필요한 것으로 결정된 영역을 제거하기 위해 이용되며, 이는 사용자에게 의해 제공될 수 있다. 마스크(MK)는 인가된 이미지에 대응하는 크기를 갖고 제거하고자 하는 영역의 픽셀값은 일례로 0의 값을 갖는 반면 나머지 영역의 픽셀값은 1의 값을 갖도록 하여 제공될 수 있다. 입력 이미지 획득부(110)는 인가된 이미지와 마스크(MK)를 원소간 곱셈 연산을 수행하여 마스크(MK)에 의해 지정된 영역의 픽셀값이 0으로 제거된 입력 이미지(IN)를 획득할 수 있다.

[0033] 도 3에서는 일례로 마스크(MK)가 인가된 이미지에서 제거할 영역이 중앙에 검은색 사각형 형태로 설정된 경우를 나타내며, 이에 입력 이미지(IN)에서 중앙이 사각형 형태로 제거되었음을 알 수 있다.

[0034] 그러나 만일 마스크(MK)가 이미지와 함께 인가되지 않으면, 입력 이미지 획득부(110)는 마스크(MK)를 랜덤하게 생성할 수도 있다.

[0035] 본 실시예에서는 입력 이미지(IN)에서 마스크에 의해 제거된 영역을 비유효 영역이라고 하며, 제거되지 않고 인가된 이미지의 픽셀값을 유지하는 나머지 영역을 유효 영역이라 한다.

[0036] 인코더(120)의 다수의 인코딩 레이어(EL1 ~ EL5) 각각은 입력 이미지 획득부(110)에서 획득된 입력 이미지(IN) 또는 이전 단의 인코딩 레이어(EL1 ~ EL4)에서 출력되는 특징맵을 인가받고, 미리 학습된 패턴 추정 방식에 따라 인가된 입력 이미지(IN) 또는 특징맵을 인코딩하여 특징을 추출함으로써 특징맵을 획득한다. 이때 다수의 인코딩 레이어(EL1 ~ EL5) 각각은 입력 이미지(IN) 또는 특징맵에서 제거된 영역, 즉 비유효 영역에 대응하는 패턴의 마스크를 함께 인가받아 인코딩한다. 다수의 인코딩 레이어(EL1 ~ EL5)는 입력 이미지(IN) 또는 특징맵과 대응하는 마스크를 인가받고, 인가된 특징맵과 마스크를 학습에 의해 획득된 가중치와 부분 컨볼루션(Partial Convolution)을 수행하여 특징맵을 획득한다.

[0037] 여기서 마스크는 각각의 인코딩 레이어(EL1 ~ EL5)에서 유효 영역의 특징은 추출되는 반면 비유효 영역에서는 특징이 추출되지 않도록 하기 위해 적용된다. 마스크가 입력 이미지(IN) 또는 특징맵과 함께 인코딩되므로 인코딩 레이어(EL1 ~ EL5)는 유효 영역의 특징만을 추출할 수 있으며, 결과적으로 인코딩 레이어(EL1 ~ EL5)는 신뢰성 있는 특징을 추출할 수 있다. 여기서 마스크는 입력 이미지 획득부(110)에서 획득된 마스크 또는 이전 인코딩 레이어(EL1 ~ EL5)에서 업데이트 되어 인가된 마스크이다.

[0038] 즉 다수의 인코딩 레이어(EL1 ~ EL5) 각각은 인가된 마스크를 이용하여, 입력 이미지(IN) 또는 특징맵을 필터링하고, 이후, 인가된 마스크를 기지정된 방식에 따라 업데이트하여 다음 인코딩 레이어로 전달한다.

[0039] 도 4에서는 설명의 편의를 위해 인코더(120)의 일부만을 도시하였으며, 인코더(120)의 다수의 인코딩 레이어(EL1 ~ EL5) 중 하나의 제1 인코딩 레이어(EL₁)를 일례로 설명한다.

[0040] 도 4를 참조하면, 제1 인코딩 레이어(EL₁)는 이전 제1-1 인코딩 레이어(EL₁₋₁)에서 획득되어 인가된 특징맵(X_{l-1}^{enc}) (또는 입력 이미지(IN))과 업데이트된 마스크(M_{l-1}^{enc})를 인가받는다. 그리고 인가된 특징맵(X_{l-1}^{enc})과 마스크(M_{l-1}^{enc})에 대해 학습에 의해 획득된 가중치(W_l^{enc})를 이용하여 수학적 식 1과 같이 부분 컨볼루션을 수행함으로써 특징맵(X_l^{enc})을 획득할 수 있다.

수학적 식 1

$$X_l^{enc}(u, v) = \frac{\sum_{u', v'} W_l^{enc}(u', v') X_{l-1}^{enc}(u + u', v + v') M_{l-1}^{enc}(u + u', v + v')}{\sum_{u', v'} M_{l-1}^{enc}(u + u', v + v')}$$

[0041]

[0042] 여기서 u, v는 특징맵(X_l^{enc})에서의 좌표를 나타내고, u', v'는 가중치(W_l^{enc}) 윈도우(R)에서의 좌표(u', v' ∈ R)를 나타낸다.

[0043] 한편 제1 인코딩 레이어(EL₁₋₁)는 인가된 마스크(M_{l-1}^{enc})를 수학적 식 2에 따라 업데이트하여 마스크(M_l^{enc})를 획득할 수 있다.

수학적 식 2

$$M_l^{enc}(u, v) = \begin{cases} 1, & \text{if } \sum_{u', v'} M_{l-1}^{enc}(u + u', v + v') > \delta \\ 0, & \text{otherwise} \end{cases}$$

[0044]

[0045] 여기서 δ는 유효 영역을 제어하기 위해 미리 설정된 문턱값을 나타내며, 문턱값(δ)이 가중치 윈도우(R) 크기의 절반 이상이면, 특징맵(X_l^{enc})의 비유효 영역은 인가된 특징맵(X_{l-1}^{enc})에 비해 줄어들게 된다.

[0046] 그리고 제1 인코딩 레이어(EL₁)는 획득된 특징맵(X_l^{enc})과 마스크(M_l^{enc})를 제1+1 인코딩 레이어(EL₁₊₁)로 전달할 뿐만 아니라 디코딩부(200)의 대응하는 제L 디코딩 레이어(DL₁)로 스킵 커넥션(Skip Connection)전달한다. 여기서 스킵 커넥션은 단순히 인코딩 레이어에서 획득된 특징맵(X_l^{enc})과 마스크(M_l^{enc}) 대응하는 디코딩 레이어로 전달하는 전달 경로를 나타낸다.

[0047] 다시 도 2를 참조하면, 디코딩부(200)는 다단 구조의 다수의 디코딩 레이어(DL1 ~ DL5)를 포함하는 디코더(220)와 이미지 출력부(210)를 포함할 수 있다.

[0048] 디코더(220)의 다수의 디코딩 레이어(DL1 ~ DL5)는 인코더(120)의 다수의 인코딩 레이어(EL1 ~ EL5)의 역순에 대응하는 순서로 구성된다. 다수의 디코딩 레이어(DL1 ~ DL5)는 인코딩부(100)의 마지막 인코딩 레이어(EL5)에서 출력된 특징맵(X^{enc}) 또는 이전 디코딩 레이어에서 출력된 복원 특징맵(X_{l+1}^{dec})과 다수의 인코딩 레이어(EL1 ~ EL5) 중 대응하는 인코딩 레이어에서 전달되는 특징맵(X_l^{enc}) 및 마스크(M_l^{enc})를 인가받아 학습된 패턴 복원 방식에 따라 디코딩하여 복원 특징맵(X_l^{dec})을 출력한다. 이때 본 실시예에 따른 다수의 디코딩 레이어(DL1 ~ DL5) 각각은 비로컬 특징 합성 레이어(Non-Local Feature Synthesis-Layer: 이하 NFS 레이어)를 포함하여, 복원 특

징맵(X_{i+1}^{dec})과 특징맵(X_i^{enc}) 및 마스크(M_i^{enc})를 이용하여 디코딩맵(Y_i^{dec})과 인코딩맵(Y_i^{enc})을 생성하고, 비로컬 특징 합성 레이어에서 생성된 디코딩맵(Y_i^{dec})과 인코딩맵(Y_i^{enc})을 기반으로 기지정된 연산을 수행하여, 복원 특징맵(X_i^{dec})을 획득하도록 구성된다.

[0049] 도 5에서도 설명의 편의를 위해 디코더(220)의 일부만을 도시하였으며, 디코더(220)의 다수의 디코딩 레이어(DL1 ~ DL5) 중 하나의 제1 디코딩 레이어(DL₁)를 일례로 설명한다.

[0050] 도 5를 참조하면, 본 실시예에서 제1 디코딩 레이어(DL₁)는 인코딩 레이어와 역순 배치 구조에 따라 제1+1 디코딩 레이어(DL₁₊₁)로부터 복원 특징맵(X_i^{dec})을 인가받고, 대응하는 인코딩 레이어(EL₁)로부터 특징맵(X_i^{enc}) 및 마스크(M_i^{enc})를 인가받는다. 그리고 제1 디코딩 레이어(DL₁)의 NFS 레이어는 인가된 복원 특징맵(X_i^{dec})과 특징맵(X_i^{enc}) 및 마스크(M_i^{enc})를 이용하여 디코딩맵(Y_i^{dec})과 인코딩맵(Y_i^{enc})을 생성한다. 제1 디코딩 레이어(DL₁)는 NFS 레이어에서 생성된 디코딩맵(Y_i^{dec})과 인코딩맵(Y_i^{enc})을 결합(concatenate)하고, 결합된 맵을 학습에 의해 획득된 가중치로 디콘볼루션 하여 복원 특징맵(X_{i-1}^{dec})을 획득하고, 획득된 복원 특징맵(X_{i-1}^{dec})을 다음 단의 제1-1 디코딩 레이어(DL₁₋₁)로 전달한다.

[0051] 본 실시예에서 NFS 레이어는 다수의 디코딩 레이어(DL1 ~ DL5)가 비유효 영역의 경계 부근의 한정된 영역에서 비유효 영역을 채우기 위한 패턴을 획득하지 않고, 특징맵의 유효 영역 전체에서 의미적으로 가장 유사한 특징을 갖는 패턴 획득하여 비유효 영역을 채울 수 있도록 하기 위해 구비된다.

[0052] 도 6을 참조하면, NFS 레이어는 디코딩맵 획득부(DM)와 인코딩맵 획득부(EM)를 포함할 수 있다.

[0053] 디코딩맵 획득부(DM)는 제1+1 디코딩 레이어(DL₁₊₁)로부터 인가된 $C_1 \times H_1 \times W_1$ 크기의 복원 특징맵(X_i^{dec})을 재구성하고, 재구성된 복원 특징맵(X_i^{dec})을 서로 다른 가중치(W_1^a , W_1^b , W_1^y)로 1 X 1 컨볼루션하여 3개의 특징($\alpha_1(X_i^{dec}; W_1^a)$, $\beta_1(X_i^{dec}; W_1^b)$, $\gamma_1(X_i^{dec}; W_1^y)$)을 추출함으로써, 복원 특징맵(X_i^{dec})을 3개의 서로 다른 특징 공간에 포함시킨다.

[0054] 그리고 복원 특징맵(X_i^{dec})의 모든 픽셀 중 가장 주의해야 할 특징 픽셀을 도출하기 위해 3개의 특징($\alpha_1(X_i^{dec}; W_1^a)$, $\beta_1(X_i^{dec}; W_1^b)$, $\gamma_1(X_i^{dec}; W_1^y)$) 중 특징($\alpha_1(X_i^{dec}; W_1^a)$)과 특징($\beta_1(X_i^{dec}; W_1^b)$) 사이의 상관도를 계산하여 특징 주의 행렬(A_1)을 획득한다.

[0055] 여기서 특징 주의 행렬(A_1)은 특징($\alpha_1(X_i^{dec}; W_1^a)$)을 전치(transpose)한 전치 특징($\alpha_1^T(X_i^{dec}; W_1^a)$)과 특징($\beta_1(X_i^{dec}; W_1^b)$)을 곱하고, 소프트 맥스 함수를 $\sum_{j=1}^{H_1 W_1} A_1(i, j) = 1$ 과 같이 적용하여 정규화함으로써 획득될 수 있다.

[0056] 그리고 수학적 3과 같이 획득된 특징 주의 행렬(A_1)을 특징($\gamma_1(X_i^{dec}; W_1^y)$) 및 스케일 변수(λ_1)와 곱하고 다시 복원 특징맵(X_i^{dec})과 더함으로써 디코딩맵(Y_i^{dec})을 획득할 수 있다. 여기서 스케일 변수(λ_1)는 디코딩 레이어(DL₁)의 특징 업데이트 과정을 제어하기 위한 파라미터로서 초기값은 0으로 설정될 수 있으며, 학습에 의해 조절될 수 있다.

수학식 3

$$Y_l^{\text{dec}} = X_l^{\text{dec}} + \lambda_l \gamma(X_l^{\text{dec}}; \mathbf{W}_\gamma^l) \mathbf{A}_l$$

[0057]

[0058] 즉 디코딩맵 획득부(DM)는 복원 특징맵(X_l^{dec})에서 각 영역별 주의도를 분석하고, 분석된 주의도에 따라 각 영역에 대해 가중치를 가중하여 디코딩맵(Y_l^{dec})을 획득한다.

[0059] 한편, 인코딩맵 획득부(EM)는 대응하는 인코딩 레이어(EL_l)로부터 $C_l \times H_l \times W_l$ 크기의 특징맵(X_l^{enc}) 및 마스크(M_l^{enc})를 인가받고, 인가된 마스크(M_l^{enc})를 $1 \times H_l W_l$ 크기로 슬라이스 한다. 그리고 슬라이스된 마스크(m_l^{enc})를 반전한 반전 마스크($1 - m_l^{\text{enc}}$)를 전치한 전치 반전 마스크와 슬라이스된 마스크(m_l^{enc})를 곱($(1 - m_l^{\text{enc}})^T m_l^{\text{enc}}$)하여 홀 채움 표시자를 획득한다. 이는 인가된 마스크(M_l^{enc})에 의해 지정된 비유효 영역에 대해서만 특징이 합성되어 채워지도록 하기 위해서이다.

[0060] 홀 채움 표시자가 획득되면, 홀 채움 표시자($(1 - m_l^{\text{enc}})^T m_l^{\text{enc}}$)와 디코딩맵 획득부(DM)에서 획득된 특징 주의 행렬(A_l)을 곱($(1 - m_l^{\text{enc}})^T m_l^{\text{enc}} A_l$)하고 정규화하여 홀 채움 유사도 행렬(S_l)을 획득한다. 홀 채움 유사도 행렬(S_l)은 대응하는 인코딩 레이어(EL_l)로부터 인가된 특징맵(X_l^{enc})에서 위치 j 의 특징이 디코딩맵(Y_l^{dec})의 위치 i 의 특징으로 채워지기에 적합한지 여부를 판별하기 위한 행렬이다.

[0061] 홀 채움 유사도 행렬(S_l)은 홀 채움 표시자($(1 - m_l^{\text{enc}})^T m_l^{\text{enc}}$)와 특징 주의 행렬(A_l)의 곱($(1 - m_l^{\text{enc}})^T m_l^{\text{enc}} A_l$)을 수학식 4에 따라 정규화하여 획득될 수 있다.

수학식 4

$$\sum_j S_l(i, j) = \begin{cases} 1, & \text{if } i \in H, j \in V \\ 0, & \text{otherwise} \end{cases}$$

[0062]

[0063] 여기서 H 는 비유효 영역의 픽셀을 나타내고, V 는 유효 영역의 픽셀을 나타낸다.

[0064] 그리고 수학식 5와 같이, 획득된 홀 채움 유사도 행렬(S_l)과 특징맵(X_l^{enc})을 곱하고, 그 결과에 대해 다시 특징맵(X_l^{enc})을 더하여 인코딩맵(Y_l^{dec})을 생성한다.

수학식 5

$$Y_l^{\text{enc}} = X_l^{\text{enc}} + S_l X_l^{\text{enc}}$$

[0065]

[0066] 여기서 홀 채움 유사도 행렬(S_l)과 특징맵(X_l^{enc})을 곱하는 것은 수학식 4에 따른 홀 채움 유사도 행렬(S_l)의 조건식은 NFS 레이어가 비유효 영역(H)에 대해서만 대응하는 인코딩 레이어(EL_l)로부터 인가된 특징맵(X_l^{enc})의 특징을 합성하도록 보장하기 위해서이다.

[0067] 결과적으로 도 6에 도시된 NFS 레이어에서 디코딩맵 획득부(DM)는 복원 특징맵(X_1^{dec})에서 각 영역별 주의도를 분석하고, 분석된 주의도에 따라 각 영역에 대해 가중치를 가중하여 디코딩맵(Y_1^{dec})을 획득하고, 인코딩맵 획득부(EM)는 디코딩맵 획득부(DM)에서 획득된 특징 주의 행렬(A_1)과 대응하는 인코딩 레이어(EL_1)에서 인가된 마스크(M_1^{enc})를 기반으로 비유효 영역(H)에 대한 특징맵(X_1^{enc})의 특징을 도출하여 인코딩맵(Y_1^{dec})을 생성한다.

[0068] 이후, 제1 디코딩 레이어(DL_1)의 디컨볼루션 레이어가 상기한 바와 같이, 디코딩맵(Y_1^{dec})과 인코딩맵(Y_1^{dec})을 결합하고, 가중치로 디컨볼루션을 수행함으로써, 복원 특징맵(X_{1-1}^{dec})을 획득하고, 획득된 복원 특징맵(X_{1-1}^{dec})을 다음 단의 제1-1 디코딩 레이어(DL_{1-1})로 전달한다.

[0069] 본 실시예에서는 다수의 디코딩 레이어($DL1 \sim DL5$) 각각이 NFS 레이어를 포함하여 인가된 복원 특징맵(X_1^{dec})에서 각 영역의 주의도가 가중된 디코딩맵(Y_1^{dec})과 대응하는 인코딩 레이어(EL_1)로부터 인가된 특징맵(X_1^{enc})의 비유효 영역에 대한 특징만이 합성되도록 함으로써, 비유효 영역에 복원 특징맵(X_1^{dec})의 유효 영역에서 가장 높은 주의도를 갖는 픽셀의 특징이 합성되도록 한다. 따라서 비유효 영역의 각 픽셀에 시각적 오류가 적고 의미적으로도 유사한 픽셀이 합성되도록 할 수 있다.

[0070] 이미지 출력부(210)는 디코더(220)의 다수의 디코딩 레이어($DL1 \sim DL5$) 중 최종 단의 디코딩 레이어($DL1$)에서 출력되는 복원 특징맵(X_1^{dec})을 추정 이미지로서 출력한다.

[0071] 도 2에 도시된 본 실시예의 이미지 채움 장치를 학습시키는 경우, 마스크에 의해 비유효 영역이 발생되지 않은 원본 이미지(I_{gt})와 이미지 채움 장치에서 획득된 추정 이미지(I_{pred}) 사이에서 픽셀 단위로 재구성 및 지각 손실을 계산하여 학습시킬 수 있다.

[0072] 여기서 재구성 및 지각 손실(L_{recon})은 수학적 6과 같이 L_1 놈 함수를 이용하여 계산될 수 있다.

수학적 6

$$\mathcal{L}_{recon} = \|I_{pred} - I_{gt}\|_1$$

[0074] 다만 본 실시예에 따른 이미지 채움 장치의 경우, 인코더(120)의 다수의 인코딩 레이어($EL1 \sim EL5$)와 디코더(220)의 다수의 디코딩 레이어($DL1 \sim DL5$) 사이에 스킵 커넥션을 통해 특징맵(X_1^{enc})이 전달됨에 따라 대응하는 레이어들 사이에서의 중간 손실(L_{perc})을 수학적 7과 같이 계산할 수 있다.

수학적 7

$$\mathcal{L}_{perc} = \sum_l \|\Phi_l(I_{pred}) - \Phi_l(I_{gt})\|_1$$

[0076] 여기서 $\Phi_l(I_{pred})$ 와 $\Phi_l(I_{gt})$ 는 원본 이미지(I_{gt})와 추정 이미지(I_{pred})에 대해 제1 인코딩 레이어(EL_1)가 추출한 특징맵(X_1^{enc})을 나타낸다.

[0077] 한편, 이미지 채움 분야에서는 스타일 손실(Style Loss)(L_{scc})이 제안된 바 있으며, 여기서는 스타일 손실(L_{scc})을 수학적 8에 따라 계산한다.

수학식 8

$$\begin{aligned}\mathcal{L}_{\text{SCC}} &= \sum_l \|\mathbf{P}_l^T \mathbf{S}_l \mathbf{P}_l - \mathbf{G}_l^T \mathbf{S}_l \mathbf{G}_l\|_1 \\ &= \sum_l \left\| \sum_i \sum_j (\mathbf{p}_l^T(i) \mathbf{S}_l(i, j) \mathbf{p}_l(j) - \mathbf{g}_l^T(i) \mathbf{S}_l(i, j) \mathbf{g}_l(j)) \right\|_1 \\ &= \sum_l \|\Phi_l^T(I_{\text{pred}}) \mathbf{S}_l \Phi_l(I_{\text{pred}}) - \Phi_l^T(I_{\text{gt}}) \mathbf{S}_l \Phi_l(I_{\text{gt}})\|_1\end{aligned}$$

[0078]

[0079] 여기서 \mathbf{P}_l 및 \mathbf{G}_l 은 각각 $\Phi_l(I_{\text{pred}})$ 와 $\Phi_l(I_{\text{gt}})$ 의 행렬 표현이고, \mathbf{S}_l 은 NFS 레이어에서 획득되는 홀 채움 유사도 행렬이다.

[0080] 또한 홀 채움 유사도 제약 손실($\mathcal{L}_{\text{const-s}}$)과 텍스처 제약 손실($\mathcal{L}_{\text{const-t}}$)이 각각 수학식 9 및 10과 같이 정의될 수 있다.

수학식 9

$$\mathcal{L}_{\text{const-s}} = \sum_l \|\mathbf{S}_l \mathbf{G}_l - \mathbf{G}_l\|_1$$

[0081]

수학식 10

$$\mathcal{L}_{\text{const-t}} = \sum_l \|\mathbf{S}_l \mathbf{P}_l - \mathbf{P}_l\|_1$$

[0082]

[0083] 수학식 8의 스타일 손실(\mathcal{L}_{SCC})은 수학식 4의 홀 채움 유사도 행렬(\mathbf{S}_l) 조건식을 사용하여, 수학식 11로 표현될 수 있다.

수학식 11

$$\sum_l \left\| \sum_{i \in \mathbf{H}} \sum_{j \in \mathbf{V}} (\mathbf{p}_l(i)^T \mathbf{S}_l(i, j) \mathbf{p}_l(j) - \mathbf{g}_l(i)^T \mathbf{S}_l(i, j) \mathbf{g}_l(j)) \right\|_1$$

[0084]

[0085] 또한 수학식 11은 수학식 9 및 10의 홀 채움 유사도 제약 손실($\mathcal{L}_{\text{const-s}}$)과 텍스처 제약 손실($\mathcal{L}_{\text{const-t}}$)을 고려하여 수학식 12로 공식화될 수 있다.

수학식 12

$$\sum_l \left\| \sum_{i \in \mathbf{H}} \sum_{j \in \mathbf{V}} (\mathbf{S}_l^2(i, j) \|\mathbf{p}_l(j)\|^2 - \mathbf{S}_l^2(i, j) \|\mathbf{g}_l(j)\|^2) \right\|_1$$

[0086]

[0087] 그리고 본 실시예에 따른 이미지 채움 장치의 총 손실은 수학식 13으로 계산되며, 계산된 총 손실이 역전파 됨으로써, 이미지 채움 장치는 학습될 수 있다.

수학식 13

$$\mathcal{L}_{\text{total}} = \lambda_{\text{recon}} \mathcal{L}_{\text{recon}} + \lambda_{\text{perc}} \mathcal{L}_{\text{perc}} + \lambda_{\text{SCC}} \mathcal{L}_{\text{SCC}} + \lambda_{\text{const-s}} \mathcal{L}_{\text{const-s}} + \lambda_{\text{const-t}} \mathcal{L}_{\text{const-t}}$$

[0088]

[0089] 여기서 λ_{recon} , λ_{perc} , λ_{SCC} , $\lambda_{\text{const-s}}$ 및 $\lambda_{\text{const-t}}$ 는 손실 가중치이다.

[0090] 도 7은 본 발명의 일 실시예에 따른 이미지 채움 방법을 나타낸다.

[0091] 도 2 내지 도 6을 참조하여, 본 실시예에 따른 이미지 채움 방법을 설명하면, 이미지 채움 방법은 크게 입력 이미지의 유효 영역에서 특징을 추출하는 인코딩 단계(S10)와 인코딩 단계(S10)에서 추출된 유효 영역의 특징을 디코딩하여 유효 영역의 특징 중 의미적 유사도가 높은 특징을 합성하여 비유효 영역을 채움으로써 추정 이미지를 획득하는 디코딩 단계(S20)를 포함한다.

[0092] 인코딩 단계(S10)에서는 우선 입력 이미지를 획득한다(S11). 여기서 입력 이미지(IN)는 원본 이미지와 원본 이미지에서 일부 영역을 제거하기 위한 마스크(MK)를 함께 인가받아, 행렬 곱 연산을 수행하여 입력 이미지(IN)를 획득할 수도 있으나, 경우에 따라서는 랜덤하게 생성된 마스크를 이용하여 입력 이미지(IN)를 획득할 수도 있다.

[0093] 입력 이미지(IN)가 획득되면, 입력 이미지(IN)와 마스크(MK)에 대해 단계적으로 반복 인코딩을 수행한다. 인코딩 시에는 우선 입력 이미지(IN)와 마스크(MK)에 대해 학습에 의해 획득된 가중치를 적용하여 수학식 1에 따라 부분 컨볼루션을 수행하여 특징맵(X^{enc}_1)을 획득한다(S12). 그리고 이용된 마스크(M^{enc}_{1-1})를 수학식 2에 따라 업데이트하여 마스크(M^{enc}_1)를 획득한다(S13).

[0094] 특징맵(X^{enc}_1)과 업데이트된 마스크(M^{enc}_1)가 획득되면, 획득된 특징맵(X^{enc}_1)과 마스크(M^{enc}_1)를 다음 인코딩을 수행하기 위해 전달할 뿐만 아니라, 디코딩 단계(S20)에서 대응하는 디코딩을 수행하기 위해 전달한다(S14).

[0095] 이후 기지정된 횟수로 반복 인코딩이 수행되어 인코딩이 완료되었는지 판별한다(S15). 만일 인코딩이 완료되지 않았으면, 획득된 특징맵(X^{enc}_1)과 마스크(M^{enc}_1)에 대해 다시 부분 컨볼루션을 수행하여 다음 특징맵(X^{enc}_{1+1})을 획득한다(S12).

[0096] 그러나 인코딩이 완료된 것으로 판별되면, 디코딩 단계를 수행한다(S20). 디코딩 단계(S20) 또한 인코딩 단계(S10)와 마찬가지로 인가된 특징맵(X^{enc}_{1+1})에 대해 단계적으로 반복 인코딩을 수행하며, 인코딩 단계(S10)에서 반복되는 인코딩 횟수와 디코딩 단계(S20)에서 반복되는 디코딩 횟수는 동일하게 설정된다.

[0097] 디코딩 단계(S20)에서는 우선 인코딩 단계에서 최종적으로 획득된 특징맵(X^{enc}_{1+1})을 복원 특징맵(X^{dec}_{1+1})으로 인가받을 수 있다. 그리고 복원 특징맵(X^{dec}_{1+1})을 기지정된 크기의 행렬로 재구성하고, 재구성된 복원 특징맵(X^{dec}_{1+1})에 대해 서로 다른 3개의 가중치(W_1^a , W_1^b , W_1^y)로 1 X 1 컨볼루션하여 3개의 특징($\alpha_1(X^{\text{dec}}_1; W_1^a)$, $\beta_1(X^{\text{dec}}_1; W_1^b)$, $\gamma_1(X^{\text{dec}}_1; W_1^y)$)을 추출하고, 3개의 특징 중 2개의 특징($\alpha_1(X^{\text{dec}}_1; W_1^a)$, $\beta_1(X^{\text{dec}}_1; W_1^b)$) 사이의 상관도를 기지정된 방식에 따라 계산하여 특징 주의 행렬(A_1)을 계산한다(S21). 여기서 2개의 특징($\alpha_1(X^{\text{dec}}_1; W_1^a)$, $\beta_1(X^{\text{dec}}_1; W_1^b)$) 사이의 상관도는 2개의 특징 중 하나의 특징을 전치($\alpha_1^T(X^{\text{dec}}_1; W_1^a)$)하고, 전치된 특징($\alpha_1^T(X^{\text{dec}}_1; W_1^a)$)과 특징($\beta_1(X^{\text{dec}}_1; W_1^b)$)을 곱하고 소프트 맥스 함수를 적용하여 정규화 함으로써 획득될 수 있다.

[0098] 특징 주의 행렬(A_1)이 획득되면, 수학식 3과 같이 특징 주의 행렬(A_1)과 나머지 특징($\gamma_1(X^{\text{dec}}_1; W_1^y)$) 및 스케일 변수(λ_1)를 곱하고, 다시 복원 특징맵(X^{dec}_1)과 더하여 디코딩맵(Y^{dec}_1)을 획득한다(S22).

[0099] 한편, 복원 특징맵(X_{i+1}^{dec})과 인코딩 단계(S10)에서 반복 인코딩되어 획득되는 특징맵 중 역순으로 대응하는 특징맵(X_i^{enc})과 마스크(M_i^{enc})를 인가받고, 인가된 마스크(M_i^{enc})를 이용하여 홀 채움 표시자를 획득한다(S23). 홀 채움 표시자는 인가된 마스크(M_i^{enc})를 기지정된 크기로 슬라이스하고, 슬라이스된 마스크(m_i^{enc})를 반전한 반전 마스크($1 - m_i^{enc}$)를 전치한 전치 반전 마스크와 슬라이스된 마스크(m_i^{enc})를 곱($((1 - m_i^{enc})^T m_i^{enc})$)하여 획득될 수 있다.

[0100] 홀 채움 표시자가 획득되면, 홀 채움 표시자 $((1 - m_{enc_1}^{enc})^T m_{enc_1}^{enc})$ 와 획득된 특징 주의 행렬(A_1)을 이용하여 홀 채움 유사도 행렬(S_1)을 획득한다(S24). 홀 채움 표시자 $((1 - m_{enc_1}^{enc})^T m_{enc_1}^{enc})$ 와 특징 주의 행렬(A_1)의 곱 $((1 - m_{enc_1}^{enc})^T m_{enc_1}^{enc} A_1)$ 을 수학적 식 4에 따라 정규화하여 획득될 수 있다.

[0101] 이후 획득된 홀 채움 유사도 행렬(S_1)과 인가된 특징맵(X_1^{enc})을 곱하고, 그 결과에 대해 다시 특징맵(X_1^{enc})을 더 하여 인코딩맵(Y_1^{dec})을 생성한다(S25).

[0102] 디코딩맵(Y_{i-1}^{dec})과 인코딩맵(Y_{i-1}^{enc})이 획득되면, 획득된 디코딩맵(Y_{i-1}^{dec})과 인코딩맵(Y_{i-1}^{enc})을 기지정된 방식으로 결합하고, 결합된 맵에 대해 학습에 의해 획득된 가중치를 이용하여 디컨볼루션 연산을 수행함으로써, 복원 특징맵(X_{i-1}^{dec})을 획득한다(S26).

[0103] 복원 특징맵(X_{i-1}^{dec})이 획득되면, 기지정된 횟수로 반복 디코딩이 수행되어 디코딩이 완료되었는지 판별한다 (S27). 만일 디코딩이 완료되지 않았으면, 획득된 복원 특징맵(X_{i-1}^{dec})에 대해 특징 주의 행렬(A_{i-1})을 획득한다 (S21). 그러나 디코딩이 완료된 것으로 판별되면, 획득된 복원 특징맵(X_{i-1}^{dec})을 추정 이미지로서 출력한다 (S28).

[0104] 도 8 및 도 9는 본 실시예에 따른 이미지 채움 방법에 따라 입력 이미지의 비유효 영역을 채운 결과를 나타낸다.

[0105] 도 8 및 도 9에 도시된 바와 같이, (a)의 입력 이미지에 대해 (e)에 나타난 본 실시예에 따른 이미지 채움 방법은 (b) 내지 (d)에 도시된 기존의 다른 이미지 채움 기법에 비해 (g)에 도시된 원본 이미지와 가장 유사하게 나타남을 알 수 있으며, 이로 인해 시각적 오류가 적고 의미적으로 유사한 이미지를 획득할 수 있음을 알 수 있다.

[0106] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0107] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

[0108] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

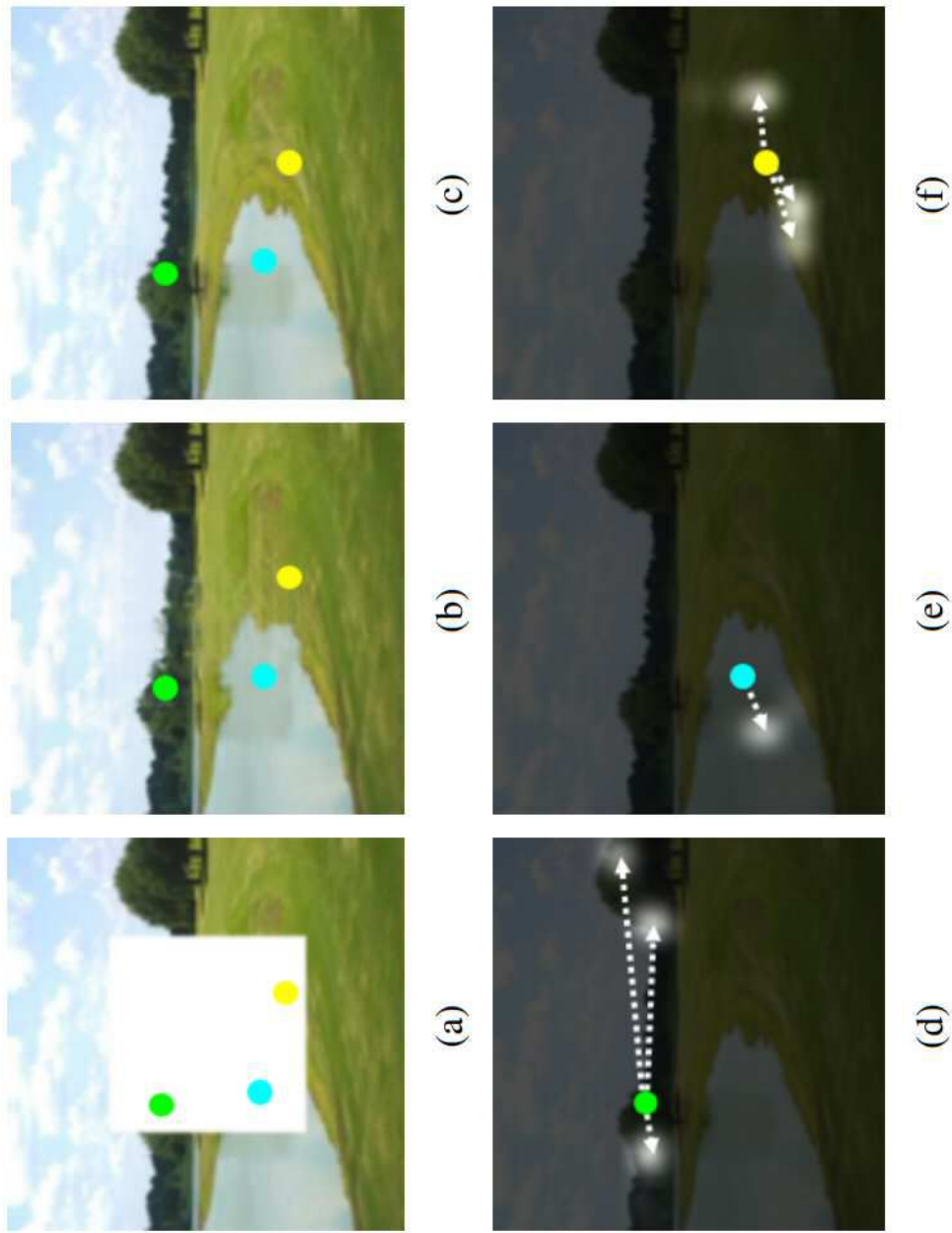
[0109]	100: 인코딩부	110: 입력 이미지 획득부
	120: 인코더	EL1 ~ EL5: 인코딩 레이어
	200: 디코딩부	210: 이미지 출력부

220: 디코더

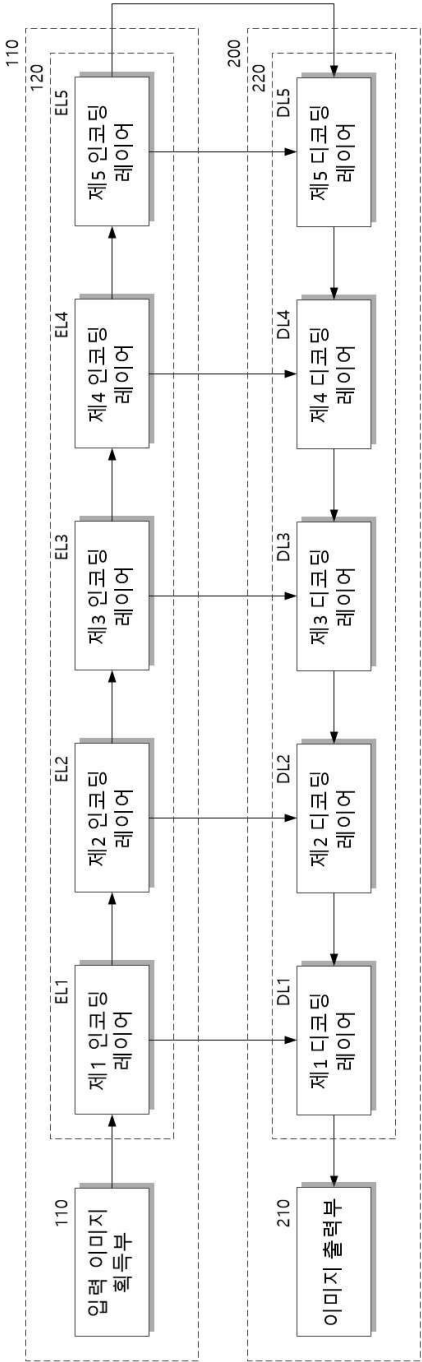
DL1 ~ DL5: 디코딩 레이어

도면

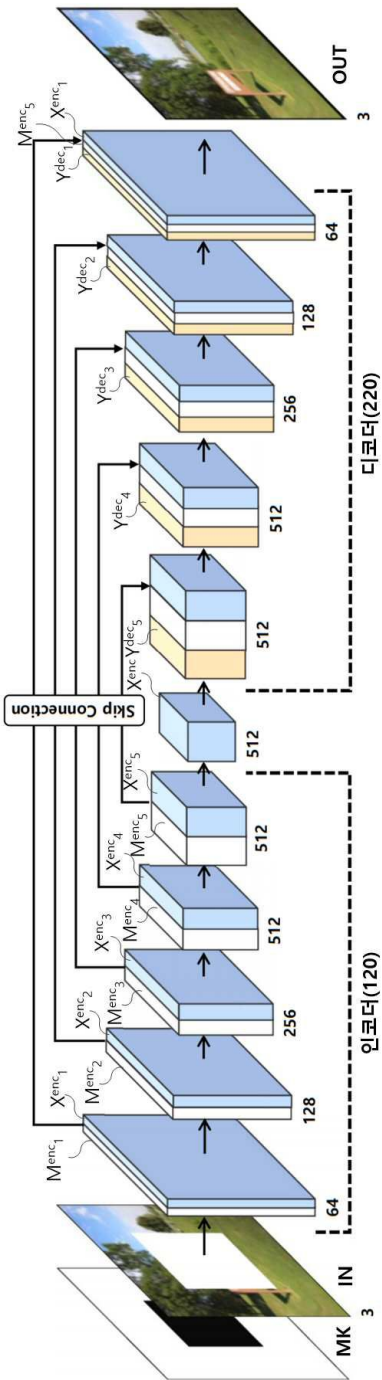
도면1



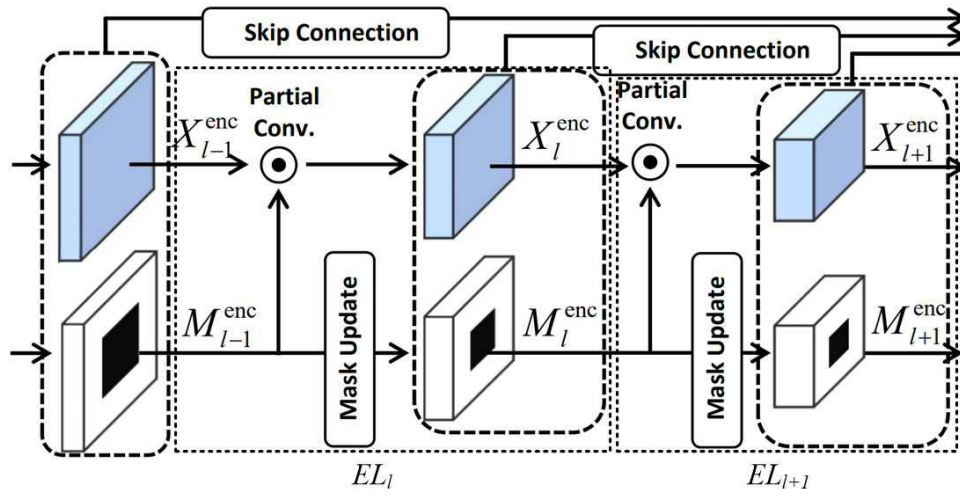
도면2



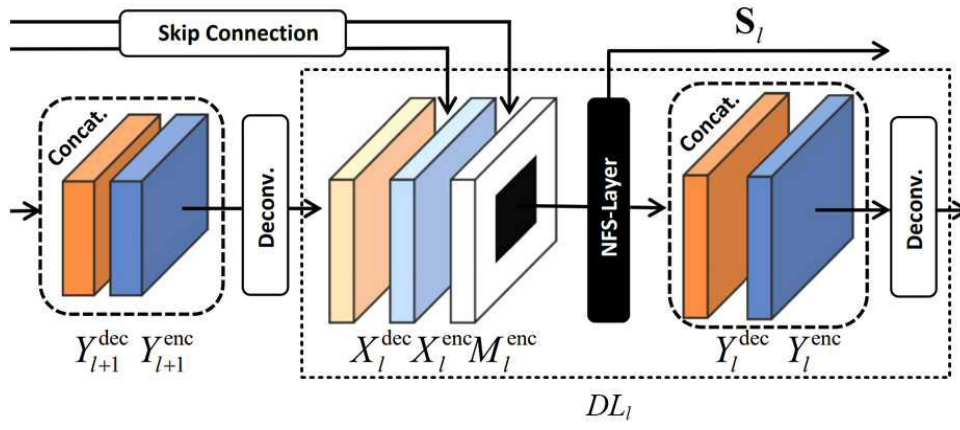
도면3



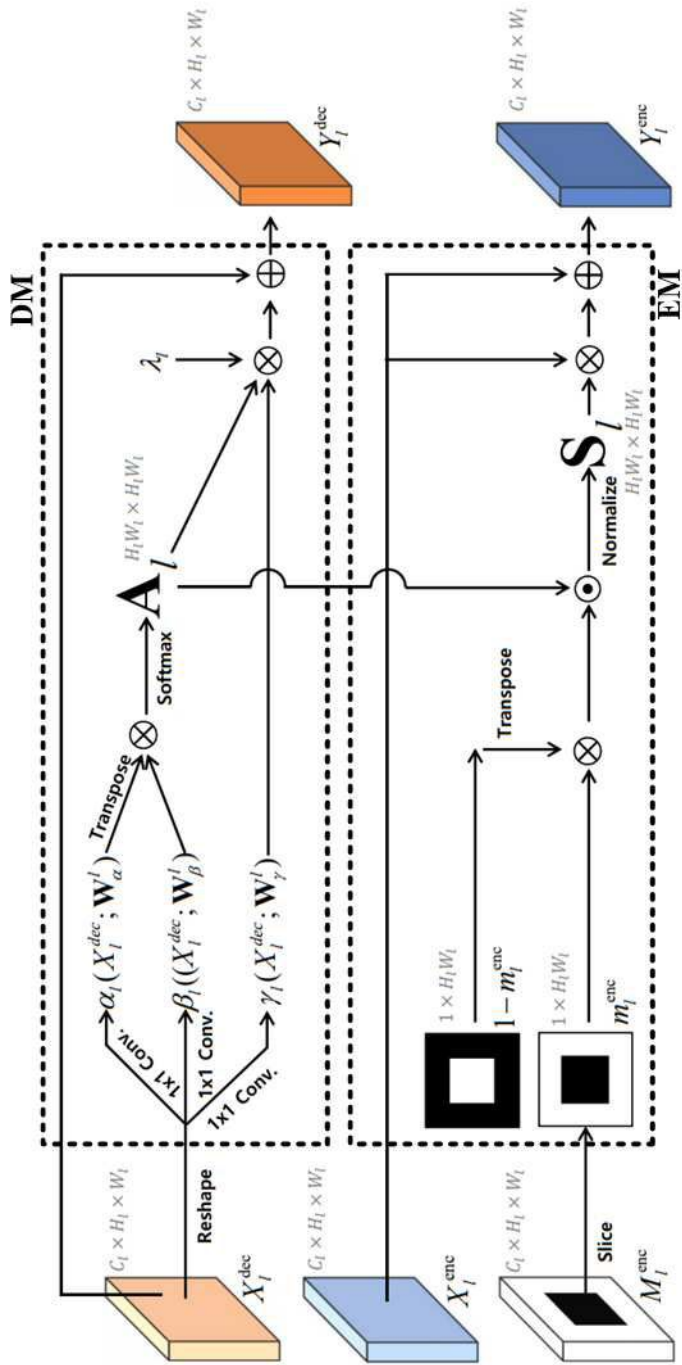
도면4



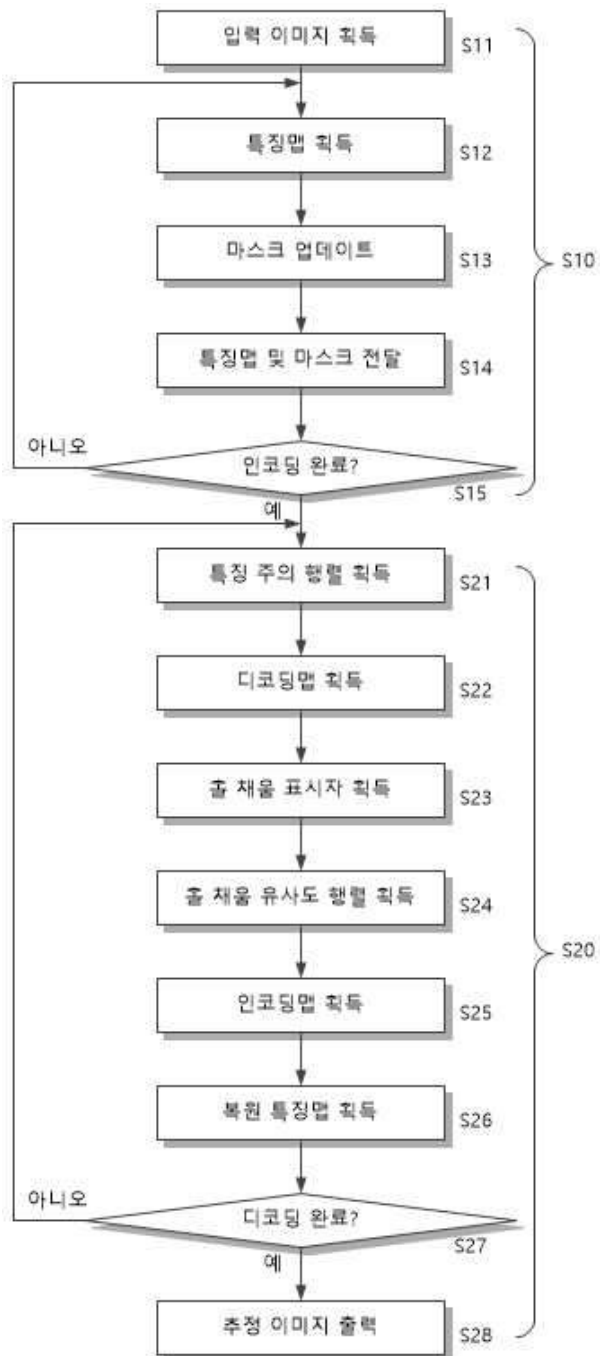
도면5



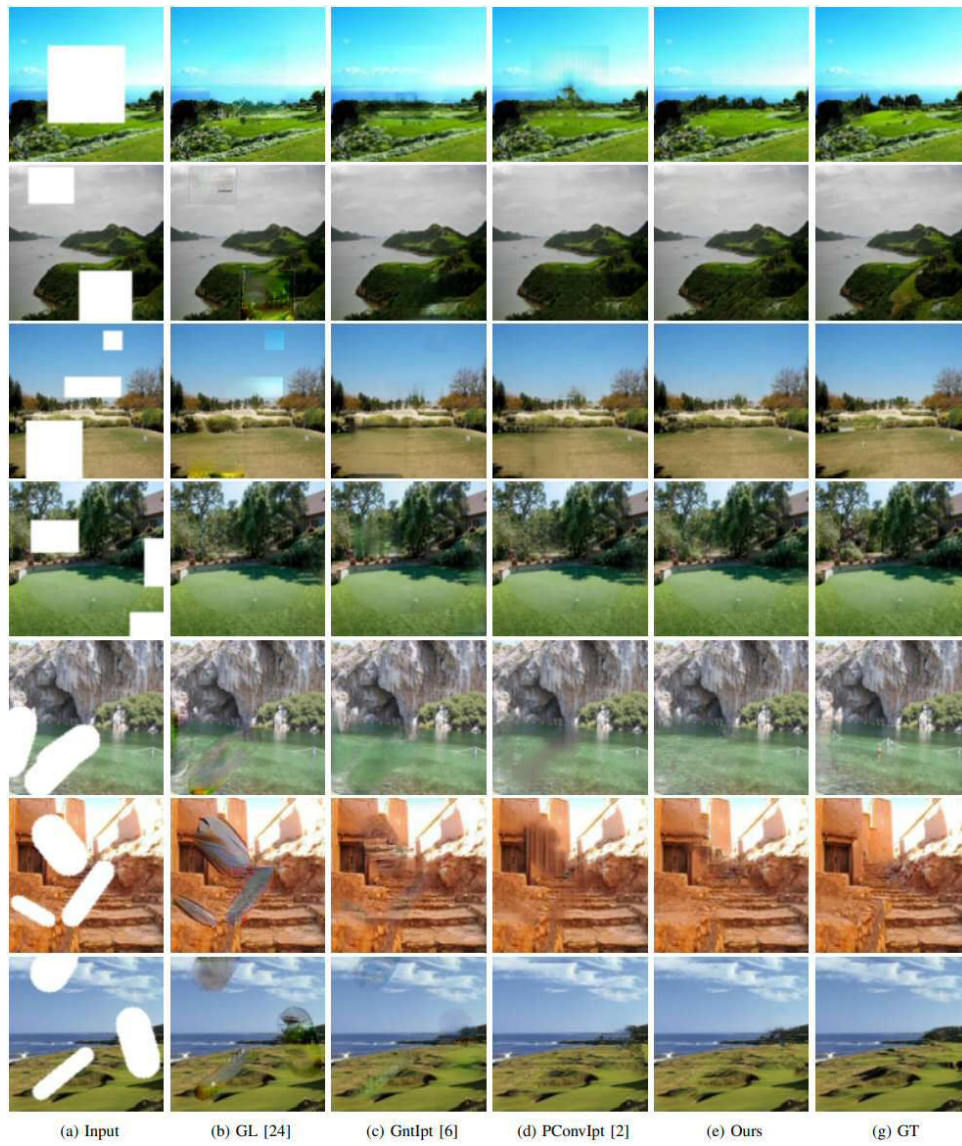
도면6



도면7



도면8



도면9

