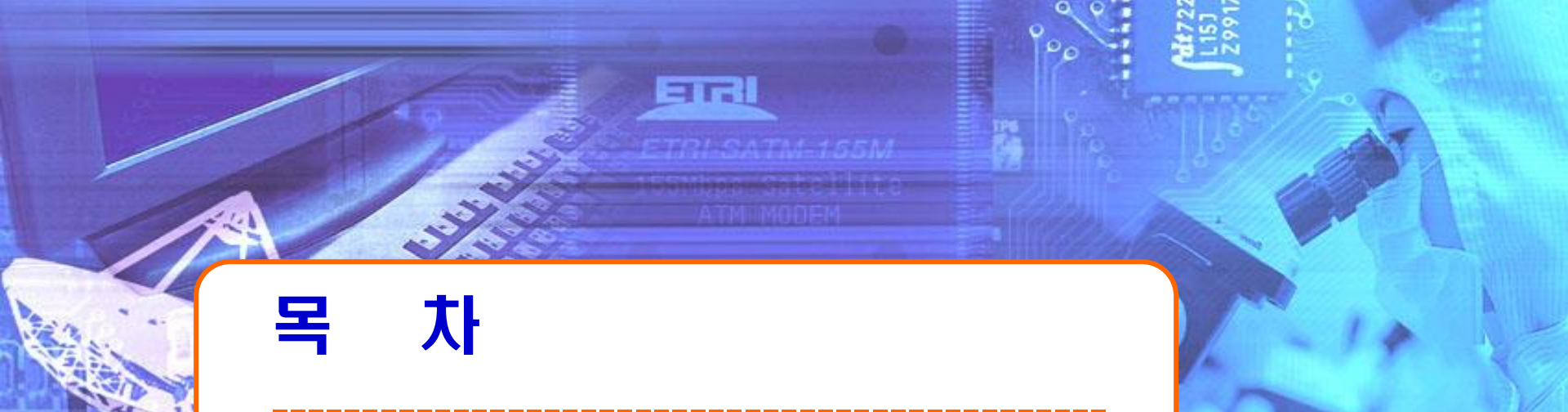


태블릿 환경에서 구동이 가능한 GAN 기반의 고화질 얼굴 편집 기술





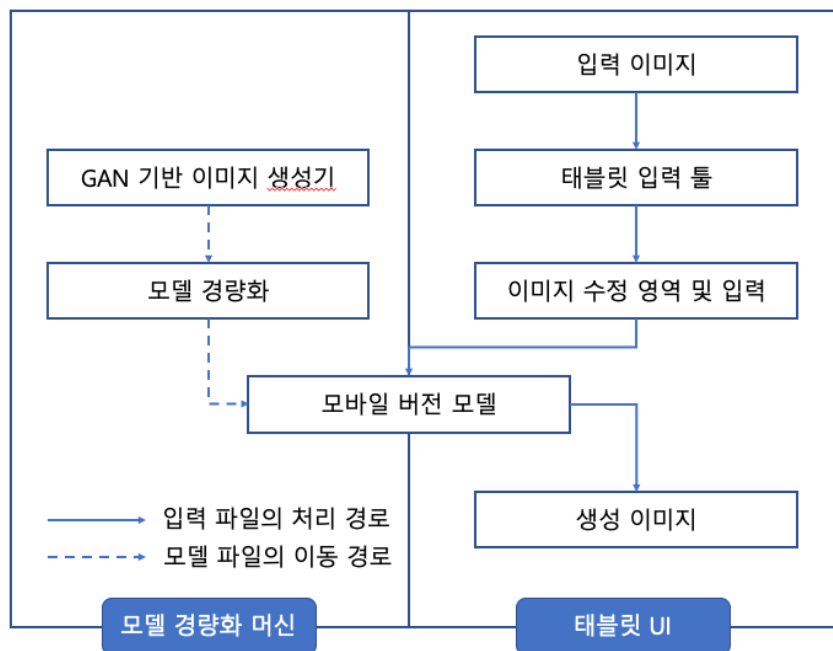
목 차

1. 기술의 개요
2. 기술이전 내용 및 범위
3. 경쟁기술과 비교
4. 기술의 사업성
 - 활용분야 및 기대효과
5. 국내외 시장 동향

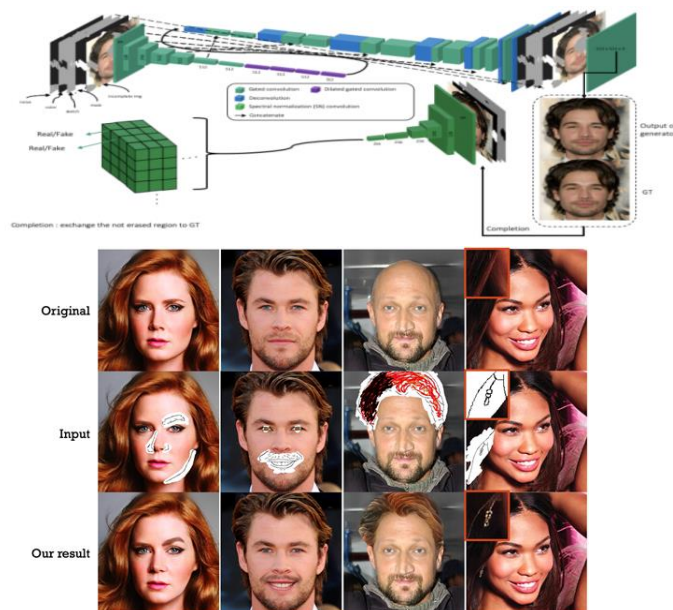
1. 기술의 개요

■ 태블릿 환경에서 사용자의 스케치 입력에 따라 얼굴을 편집하고 고화질의 얼굴 이미지를 생성 복원하는 기술

본 기술이전은 얼굴 편집 및 복원 기술로서, 고화질의 얼굴 이미지 편집과 복원을 위한 GAN 모델 생성 기술, 태블릿 환경에서 동작이 가능하도록 모델을 변환하는 기술, 이미지 편집 및 결과 출력을 위한 UI를 포함하여 구성됨



< 본 기술의 전체적인 구조도 >



< 본 기술에 활용된 딥러닝 네트워크 구조 및 실행 결과 >

※ 본 기술 우수성은 SC-FEGAN 논문으로 ICCV 2019에서 발표

2. 기술이전 내용 및 범위

□ 기술이전 내용

- ❖ 태블릿 환경에서 구동 가능한 GAN 기반의 얼굴 이미지 편집 엔진
 - ❖ 고화질 얼굴 이미지 편집 및 복원을 위한 GAN 모델 생성 기술
 - ❖ 태블릿 환경에서 구동이 가능하도록 모델을 변환해주는 기술
 - ❖ 이미지 수정 및 결과 출력을 위한 초기 태블릿 UI, 데스크톱 UI 모듈 (GUI 기반의 시각화된 출력 인터페이스 제공)

□ 기술이전 범위

- ❖ 기술문서
- ❖ 관련 소스 코드, 샘플 데이터 및 구동 프로그램
- ❖ 공개 데이터에 학습된 모델
- ❖ 시험 절차서 및 결과서

2. 기술이전 내용 및 범위

■ 기술 개발 현황

❖ 기술성숙도(TRL : Technology Readiness Level) 단계 : (6)단계

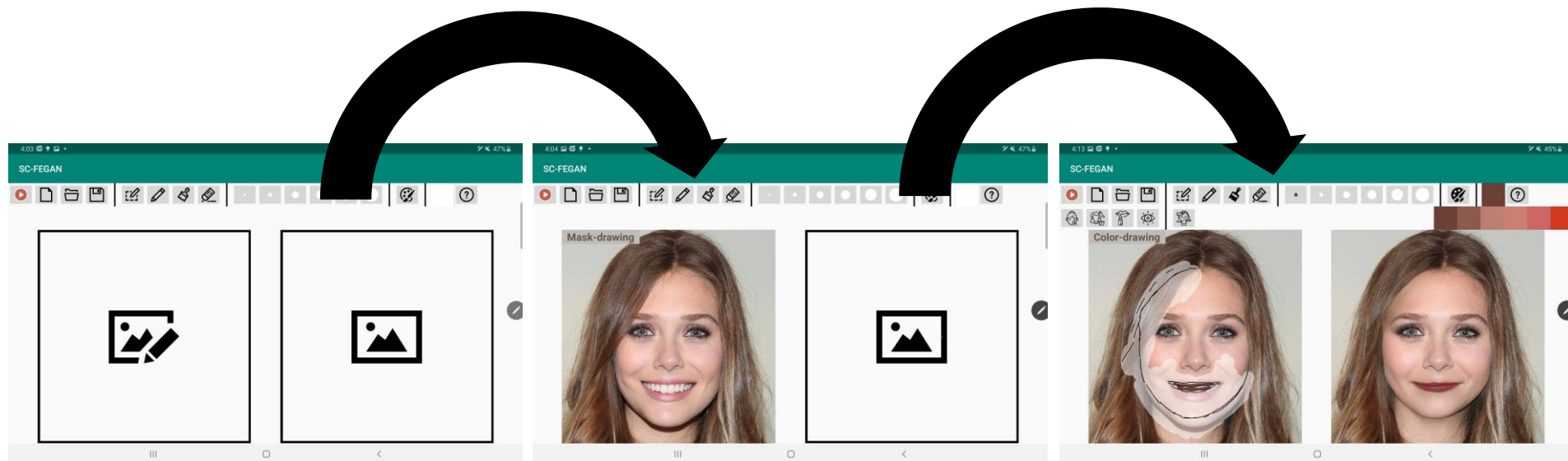
구 분	단계	정 의	세 부 설 명
기초 연구 단계	1	기초 이론/실험	기초이론 정립 단계
	2	실용 목적의 아이디어 특허 등 개념정립	기술개발 개념 정립 및 아이디어에 대한 특허 출원 단계
실험 단계	3	실험실 규모의 기본성능 검증	실험실 환경에서 실험 또는 전산 시뮬레이션을 통해 기본성능이 검증될 수 있는 단계 개발하려는 부품/시스템의 기본 설계도면을 확보하는 단계
	4	실험실 규모의 소재/부품/시스템 핵심성능 평가	시험생품을 제작하여 핵심성능에 대한 평가가 완료된 단계 3단계에서 도출된 다양한 결과 중에서 최적의 결과를 선택하려는 단계 컴퓨터 모사가 가능한 경우 최적화를 완료하는 단계
시작품 단계	5	확정된 소재/부품/ 시스템 시작품 제작 및 성능 평가	확정된 소재/부품/시스템의 실험실 시작품 제작 및 성능 평가가 완료된 단계 개발 대상의 생산을 고려하여 설계하나 실제 제작한 시작품 샘플은 1~수개 미만인 단계 경제성을 고려하지 않고 기술의 핵심성능으로만 볼 때, 실제로 판매가 될 수 있는 정도로 목표 성능을 달성한 단계
	6	파일럿 규모 시작품 제작 및 성능 평가	파일럿 규모(복수 개~양산규모의 1/10정도)의 시작품 제작 및 평가가 완료된 단계 파일럿 규모 생산품에 대해 생산량, 생산용량, 불량률 등 제시 파일럿 생산을 위한 대규모 투자가 동반되는 단계 생산기업이 수요기업 적용환경에 유사하게 자체 현장테스트를 실시하여 목표 성능을 만족시킨 단계 성능 평가 결과에 대해 가능하면 공인인증 기관의 성적서 확보
실용화 단계	7	신뢰성평가 및 수요기업 평가	실제 환경에서 성능 검증이 이루어지는 단계 부품 및 소재개발의 경우 수요업체에서 직접 파일럿 시작품을 현장 평가(성능 및 신뢰성 평가) 가능하면 인증기관의 신뢰성 평가 결과 제출
	8	시제품 인증 및 표준화	표준화 및 인허가 취득 단계
사업화	9	사업화	본격적인 양산 및 사업화 단계 6-시그마 등 품질관리가 중요한 단계

3. 경쟁기술과 비교

■ 기술의 주요 특징

❖ 태블릿 환경에서 구동이 가능한 GAN 기반의 고화질 얼굴 편집 기술

- 태블릿 환경에서의 고화질의 얼굴 이미지 편집 및 복원 지원
- 사용자의 자유로운 입력을 반영하여 고화질의 얼굴 이미지를 편집하고 특별한 툴 없이 GAN 기술을 활용하여 고품질의 얼굴 이미지 생성 복원이 가능



4. 기술의 사업성

■ 활용 분야

예상 제품 / 서비스	예상 수요자
얼굴 이미지 편집 시스템	- 성형외과 업체 - 모바일 어플리케이션 공급 업체
얼굴 이미지 복원 시스템	- 영상 분석 사업자 - CCTV 분석 사업자

■ 기대 효과

❖ 본 기술은 사람의 얼굴 이미지를 PC 및 모바일 환경에서 특정한 틀 없이 자유롭게 수정이 가능하고 이를 위한 기초 UI를 제공. 본 기술은 기학습된 얼굴에 대한 이미지와 비슷한 이미지를 생성할 수 있으므로, 사업화에 필요한 추가적인 기술 개발이 필요함

❖ 기대 활용처

- 1. 모바일 어플리케이션 : 모바일 환경에서도 간단한 조작으로 높은 성능으로 얼굴 이미지를 수정하여 보여주는 분야. 성형외과 등에서 성형 전후 얼굴 비교를 통한 모바일 홍보 어플리케이션으로 활용 가능
- 2. CCTV 몽타주 분석: CCTV에서 범인과 같은 특정 사람의 얼굴을 다시 복원하거나, 몽타주를 스케치하여 자연스러운 얼굴을 생성하는 방법으로 활용 가능

5. 국내외 시장 동향

■ 시장전망

- ❖ 이미지/영상 분석 SW 관련 국내 시장은 2017년 605억원에서 2021년 1,444억원으로 연 평균 15.59% 정도의 성장세, 세계 시장은 2016년 86.7억 달러에서 2021년 170.9억달러로 연평균 약 11.97% 정도의 성장세를 전망

출처: Intelligent Video Analytics-Global intelligent video analytics market(2017-2021)

(단위 : 억달러, 억원)

관련 제품 /서비스	시장	1차년도 (2017)	2차년도 (2018)	3차년도 (2019)	4차년도 (2020)	5차년도 (2021)	합계
지능형 영상 분석	해외	97.7	110.8	127.1	146.7	170.9	653.2
	국내	698	815	966	1,167	1,443	5,091

(출처 :Intelligent Video Analytics-Global intelligent video analytics market(2017-2021))

* 국내 시장은 아시아 지역 시장의 합에서 20% 정도의 시장 규모로 산정, (환율 1\$ = 1,100원)

감사합니다.



www.etri.re.kr



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0107742
(43) 공개일자 2020년09월16일

(51) 국제특허분류(Int. Cl.)
G06T 11/60 (2006.01) G06T 11/00 (2006.01)
G06T 11/20 (2006.01)
(52) CPC특허분류
G06T 11/60 (2013.01)
G06T 11/001 (2013.01)
(21) 출원번호 10-2019-0130282
(22) 출원일자 2019년10월18일
심사청구일자 없음
(30) 우선권주장
1020190026006 2019년03월06일 대한민국(KR)

(71) 출원인
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
조영주
경기도 화성시 향남읍 행정축전로2길 42-9 201호
박중열
대전광역시 중구 서문로 96, 203동 1503호 (문화동, 센트럴파크2단지아파트)
배유석
대전광역시 유성구 관평1로 12, 704동 1401호 (관평동, 대덕테크노밸리7단지아파트)
(74) 대리인
특허법인지명

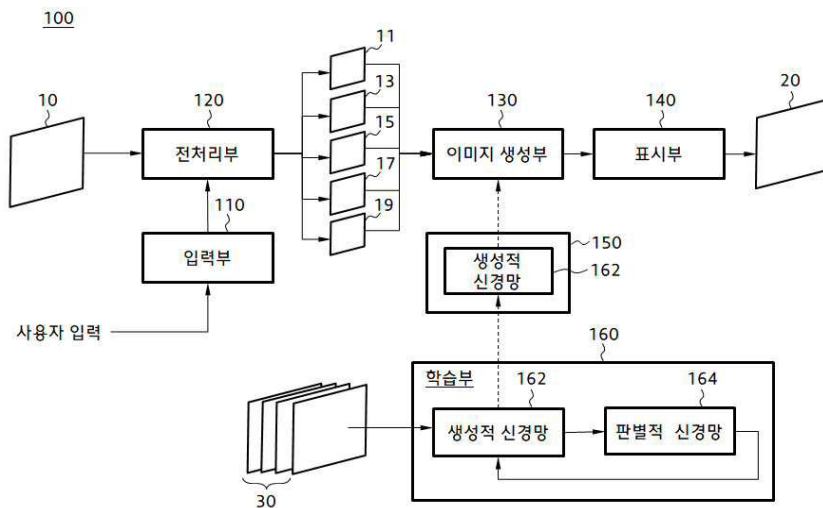
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 이미지 수정 시스템 및 이의 이미지 수정 방법

(57) 요약

본 발명의 이미지 수정 방법은 원본 이미지에 대해 전처리 과정을 수행하여 상기 원본 이미지 내에서 지워진 영역만을 포함하는 마스크(mask) 이미지를 생성하는 단계; 생성적 적대 신경망(Generative Adversarial Networks)을 이용하여 상기 마스크 이미지 내에서 상기 지워진 영역에 합성될 이미지를 예측하는 단계; 및 상기 예측된 이미지를 상기 원본 이미지 내에서 상기 지워진 영역에 합성하여 새로운 이미지를 생성하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06T 11/20 (2013.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

G06T 2207/30201 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	2014-3-00123
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	2019년 RnD 재발견프로젝트
연구과제명	(딥뷰-1세부) 실시간 대규모 영상 데이터 이해·예측을 위한 고성능 비주얼 디스커
버리 플랫폼 개발	
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

원본 이미지에 대해 전처리 과정을 수행하여 상기 원본 이미지 내에서 지워진 영역만을 포함하는 마스크(mask) 이미지를 생성하는 단계;

생성적 적대 신경망(Generative Adversarial Networks)을 이용하여 상기 마스크 이미지 내에서 상기 지워진 영역에 합성될 이미지를 예측하는 단계; 및

상기 예측된 이미지를 상기 상기 원본 이미지 내에서 상기 지워진 영역에 합성하여 새로운 이미지를 생성하는 단계

를 포함하는 이미지 수정 방법.

청구항 2

제1항에서,

상기 생성적 적대 신경망은 서로 적대적인(Adversarial) 관계를 유지하도록 학습된 생성적 신경망(Generative neural network)과 판별적 신경망(Discriminative neural network)을 포함하고,

상기 예측하는 단계는,

상기 생성적 신경망을 이용하여 상기 합성될 이미지를 예측하는 단계인 것인 이미지 수정 방법.

청구항 3

제1항에서,

상기 마스크 이미지를 생성하는 단계 이전에, 대량의 훈련 데이터를 이용하여 상기 생성적 적대 신경망을 학습시키는 단계를 더 포함하고,

상기 대량의 훈련 데이터는,

훈련용 원본 이미지들, 상기 지워진 영역만을 포함하는 훈련용 마스크 이미지들, 상기 지워진 영역에 다양한 모양이 스케치된 훈련용 스케치 이미지들, 상기 지워진 영역에 다양한 색깔이 채워진 훈련용 색깔 이미지들을 포함하는 것인 이미지 수정 방법.

청구항 4

제1항에서,

상기 마스크 이미지를 생성하는 단계 이전에, 대량의 훈련 데이터를 이용하여 상기 생성적 적대 신경망을 학습시키는 단계를 더 포함하고,

상기 생성적 적대 신경망을 학습시키는 단계는,

상기 대량의 훈련 데이터를 이용하여 생성적 신경망(Generative neural network)을 학습시키는 단계; 및

상기 생성적 신경망과 적대적인(Adversarial) 관계를 유지하도록 판별적 신경망(Discriminative neural network)을 학습시키는 단계를 포함하고,

상기 예측하는 단계는,

학습이 완료된 상기 생성적 신경망을 이용하여 상기 합성될 이미지를 예측하는 단계인 것인 이미지 수정 방법.

청구항 5

제4항에서,

상기 생성적 신경망을 학습시키는 단계는,

상기 판별적 신경망으로부터 출력되는 손실(L_G)을 이용하여 상기 생성적 신경망을 학습시키는 단계이고,

상기 손실(L_G)은,

상기 대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 포함하는 것인 이미지 수정 방법.

청구항 6

제4항에서,

상기 생성적 신경망을 학습시키는 단계는,

상기 대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 이용하여, 상기 생성적 신경망에 포함된 게이트 컨볼루션 레이어(gated convolution layer)를 학습시키는 단계인 것인 이미지 수정 방법.

청구항 7

원본 이미지에 대해 전처리 과정을 수행하여 상기 원본 이미지 내에서 사용자 입력에 의해 지워진 영역만을 포함하는 마스크(mask) 이미지, 상기 사용자 입력에 의해 상기 지워진 영역에 스케치된 모양만을 포함하는 스케치 이미지 및 상기 사용자 입력에 의해 상기 지워진 영역에 칠해진 색깔만을 포함하는 색깔 이미지를 생성하는 전처리부;

생성적 적대 신경망(Generative Adversarial Networks)을 이용하여, 상기 마스크 이미지, 상기 스케치 이미지 및 상기 색깔 이미지로부터 상기 지워진 영역에 합성될 이미지를 예측하고, 상기 예측된 이미지를 상기 지워진 영역에 합성하여 상기 원본 이미지로부터 새로운 이미지를 생성하는 이미지 생성부; 및

상기 새로운 이미지를 표시하는 표시부

를 포함하는 이미지 수정 시스템.

청구항 8

제7항에서,

상기 생성적 적대 신경망은 서로 적대적인(Adversarial) 관계를 유지하도록 학습된 생성적 신경망(Generative neural network)과 판별적 신경망(Discriminative neural network)을 포함하고,

상기 이미지 생성부는,

상기 생성적 신경망을 이용하여 상기 합성될 이미지를 예측하는 것인 이미지 수정 시스템.

청구항 9

제7항에서,

대량의 훈련 데이터를 이용하여 상기 생성적 적대 신경망을 학습시키는 학습부; 및

상기 학습부에 의해 학습이 완료된 상기 생성적 적대 신경망을 저장하는 저장부를 더 포함하고,

상기 대량의 훈련 데이터는,

훈련용 원본 이미지들, 상기 지워진 영역만을 포함하는 훈련용 마스크 이미지들, 상기 지워진 영역에 다양한 모양이 스케치된 훈련용 스케치 이미지들, 상기 지워진 영역에 다양한 색깔이 채워진 훈련용 색깔 이미지들을 포함하는 것인 이미지 수정 시스템.

청구항 10

제9항에서,

상기 학습부는,

서로 적대적인(Adversarial) 관계를 유지하도록 생성적 신경망(Generative neural network)과 판별적 신경망(Discriminative neural network)을 포함하는 상기 생성적 적대 신경망을 학습시키고,

상기 저장부는,

학습이 완료된 상기 생성적 신경망만이 저장되는 것인 이미지 수정 시스템.

청구항 11

제10항에서,

상기 학습부는,

상기 판별적 신경망으로부터 출력되는 손실(L_d)을 감소시키는 방향으로 상기 생성적 신경망을 학습시키고,

상기 손실(L_d)은,

상기 대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 포함하는 것인 이미지 수정 시스템.

청구항 12

제10항에서,

상기 학습부는,

상기 대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 이용하여, 상기 생성적 신경망에 포함된 게이트 컨볼루션 레이어(gated convolution layer)를 학습시키는 것인 이미지 수정 시스템.

청구항 13

서로 적대적인 관계에 있는 생성적 신경망과 판별적 신경망으로 구성된 생성적 적대 신경망(Generative Adversarial Networks)을 학습시키는 단계;

학습이 완료된 상기 생성적 신경망을 저장부에 저장하는 단계;

전치치 과정을 통해 원본 이미지로부터 상기 원본 이미지 내에서 사용자 입력에 의해 지워진 영역만을 포함하는 마스크(mask) 이미지, 상기 사용자 입력에 의해 상기 지워진 영역에 스케치된 모양만을 포함하는 스케치 이미지 및 상기 사용자 입력에 의해 상기 지워진 영역에 칠해진 색깔만을 포함하는 색깔 이미지를 생성하는 단계;

상기 저장부에 저장된 생성적 신경망을 이용하여, 상기 마스크 이미지, 상기 스케치 이미지 및 상기 색깔 이미지로부터 상기 지워진 영역에 합성될 이미지를 예측하는 단계; 및

상기 예측된 이미지를 상기 지워진 영역에 합성하여 상기 원본 이미지로부터 새로운 이미지를 생성하는 단계를 포함하는 이미지 수정 방법.

청구항 14

제13항에서,

상기 학습시키는 단계는,

서로 적대적인(Adversarial) 관계를 유지하도록 생성적 신경망(Generative neural network)과 판별적 신경망(Discriminative neural network)을 학습시키는 단계이고,

상기 생성적 신경망을 학습시키는 단계는,

대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 이용하여, 상기 생성적 신경망에 포함된 게이트 컨볼루션 레이

어(gated convolution layer)를 학습시키는 단계인 것인 이미지 수정 방법.

청구항 15

제13항에서,

상기 지워진 영역에 합성될 이미지를 예측하는 단계에서,

상기 합성될 이미지는,

상기 지워진 영역에 스케치된 모양에 따라 예측된 사실적인 이미지인 것인 이미지 수정 방법.

청구항 16

제13항에서,

상기 원본 이미지는 얼굴 이미지인 것인 이미지 수정 방법.

발명의 설명

기술 분야

[0001] 본 발명은 이미지를 수정하기 위한 기술에 관한 것으로, 더욱 상세하게는 얼굴 이미지를 수정하기 위한 기술에 관한 것이다.

배경 기술

[0003] 최근 SNS 를 통해 이미지와 같은 다양한 정보를 공유하는 사람들이 증가하는 추세이고, 이러한 추세에 따라, 이미지 편집 프로그램이나 어플리케이션에 대한 관심이 높아지고 있다.

[0004] 종래의 이미지 편집 프로그램(또는 이미지 편집 툴)이나 이와 관련된 어플리케이션은 주로 이미지의 픽셀 값을 조절하는 방식으로 이미지를 수정한다. 사실적으로 수정된 이미지는 이미지 편집 프로그램을 다루는 사용자의 숙련도에 따라 결정된다.

[0005] 따라서, 이미지 편집 프로그램에 대해 전문적인 지식이나 경험이 적은 일반 사용자가 이미지 편집 프로그램을 이용하여 이미지를 수정할 경우, 그 결과물은 사실적이지 못하고, 어색한 이미지일 가능성이 높다.

[0006] 이미지 편집과 관련해, 이미지에 대해 3D 모델링을 수행하여 획득한 3D 모델을 엔진(이하, 3D 모델 엔진)을 이용하여 편집하는 방식도 있다. 그러나, 이러한 방식은 3D 모델을 구현하는 3D 모델 엔진이 필요하며 이러한 3D 모델 엔진은 다양하고, 사용하는 3D 모델 엔진마다 결과물의 완성도나 완성에 필요한 지식이 천차만별로 다르다.

[0007] 또한 각 3D 모델 엔진에 대한 사용법을 사용자가 충분히 숙지하고 있어야 하는 한계가 존재한다. 즉, 기존의 방법들은 단순한 이미지 입력만으로 이미지를 사실적으로 수정하는 것이 매우 어렵다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 얼굴 이미지와 사용자 입력을 이용하여 사전에 학습시킨 신경망을 이용하여, 기존 방식에 비해 쉽고 빠르게 사실적인 합성 이미지를 제공할 수 있는 얼굴 이미지 수정 시스템 및 그 방법을 제공하는 데 목적이 있다.

[0010] 즉, 본 발명에서는 사용자가 쉽고 직관적으로 얼굴 이미지를 수정하고 사실적인 결과물을 얻을 수 있도록 하는 시스템과 방법을 제시한다. 사실적인 결과물을 얻기 위하여 전문적인 지식이나 경험을 필요로 하는 것이 아닌 누구나 쉽게 얼굴 이미지를 수정하도록 하는 것이 본 발명의 목적이다.

[0011] 본 발명의 기술한 목적들 및 그 이외의 목적과 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부된 도면과

함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다.

과제의 해결 수단

- [0013] 상술한 목적을 달성하기 위한 본 발명의 일측면에 따른 이미지 수정 방법은 원본 이미지에 대해 전처리 과정을 수행하여 상기 원본 이미지 내에서 지워진 영역만을 포함하는 마스크(mask) 이미지를 생성하는 단계; 생성적 적대 신경망(Generative Adversarial Networks)을 이용하여 상기 마스크 이미지 내에서 상기 지워진 영역에 합성될 이미지를 예측하는 단계; 및 상기 예측된 이미지를 상기 원본 이미지 내에서 상기 지워진 영역에 합성하여 새로운 이미지를 생성하는 단계를 포함한다.
- [0014] 본 발명의 다른 측면에 따른 이미지 수정 시스템은, 원본 이미지에 대해 전처리 과정을 수행하여 상기 원본 이미지 내에서 사용자 입력에 의해 지워진 영역만을 포함하는 마스크(mask) 이미지, 상기 사용자 입력에 의해 상기 지워진 영역에 스케치된 모양만을 포함하는 스케치 이미지 및 상기 사용자 입력에 의해 상기 지워진 영역에 칠해진 색깔만을 포함하는 색깔 이미지를 생성하는 전처리부;
- [0015] 생성적 적대 신경망(Generative Adversarial Networks)을 이용하여, 상기 마스크 이미지, 상기 스케치 이미지 및 상기 색깔 이미지로부터 상기 지워진 영역에 합성될 이미지를 예측하고, 상기 예측된 이미지를 상기 지워진 영역에 합성하여 상기 원본 이미지로부터 새로운 이미지를 생성하는 이미지 생성부; 및 상기 새로운 이미지를 표시하는 표시부를 포함한다.
- [0016] 본 발명의 또 다른 측면에 따른 이미지 수정 방법은, 서로 적대적인 관계에 있는 생성적 신경망과 판별적 신경망으로 구성된 생성적 적대 신경망(Generative Adversarial Networks)을 학습시키는 단계; 학습이 완료된 상기 생성적 신경망을 저장부에 저장하는 단계; 전처리 과정을 통해 원본 이미지로부터 상기 원본 이미지 내에서 사용자 입력에 의해 지워진 영역만을 포함하는 마스크(mask) 이미지, 상기 사용자 입력에 의해 상기 지워진 영역에 스케치된 모양만을 포함하는 스케치 이미지 및 상기 사용자 입력에 의해 상기 지워진 영역에 칠해진 색깔만을 포함하는 색깔 이미지를 생성하는 단계; 상기 저장부에 저장된 생성적 신경망을 이용하여, 상기 마스크 이미지, 상기 스케치 이미지 및 상기 색깔 이미지로부터 상기 지워진 영역에 합성될 이미지를 예측하는 단계; 및 상기 예측된 이미지를 상기 지워진 영역에 합성하여 상기 원본 이미지로부터 새로운 이미지를 생성하는 단계를 포함한다.

발명의 효과

- [0018] 본 발명에 의하면, 적대적 신경망을 이용하여 얼굴 이미지를 수정함으로써, 별도의 이미지 툴에 대한 전문적인 지식이나 경험 없이 사용자가 원하는 방식으로 얼굴 이미지를 빠르고 간편하게 할 수 있다.
- [0019] 기존의 이미지 수정 프로그램은 얼굴을 가늠게 만들거나 눈을 키우거나 하기 위하여 각각에 해당하는 툴이 따로 존재하고 이를 잘 활용하기 위해서는 사용자의 많은 경험이 필요했다.
- [0020] 하지만, 본 발명에서 제공하는 시스템은 별도의 이미지 툴이 필요하지 않고, 마스크, 스케치 또는 색깔을 입력 정보로 사용하여 사용자가 원하는 방향으로 얼굴 이미지를 수정할 수 있다.

도면의 간단한 설명

- [0022] 도 1은 본 발명의 실시 예에 따른 얼굴 이미지 수정 시스템의 블록도.
- 도 2는 도 1에 도시한 전처리부로부터 이미지 생성부로 입력되는 다수의 입력 이미지를 설명하기 위한 도면.
- 도 3 내지 도 7은 본 발명의 실시 예에 따른 사용자 인터페이스의 화면 구성을 나타내는 도면들.
- 도 8은 본 발명에 적용되는 생성적 적대 신경망(GANs)의 전체 네트워크 구조를 나타내는 도면.
- 도 9는 본 발명의 실시 예에 따른 이미지 수정 방법을 보여주는 흐름도.
- 도 10 내지 12는 본 발명의 이미지 수정 방법에 따라 원본 이미지를 수정한 결과 이미지들의 예시한 도면들.

발명을 실시하기 위한 구체적인 내용

- [0023] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시 예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0024] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 변경, 균등물, 또는 대체물을 포함한다.
- [0025] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만, 예를 들어 본 발명의 개념에 따른 권리 범위로부터 이탈되지 않은 채, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0026] 본 명세서에서 사용한 용어는 단지 특정한 실시예들을 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 실시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0028] 이하, 실시 예들을 첨부된 도면을 참조하여 상세하게 설명한다. 그러나, 특허출원의 범위가 이러한 실시 예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0030] 도 1은 본 발명의 실시 예에 따른 이미지 수정 시스템의 블록도이다.
- [0031] 도 1을 참조하면, 본 발명의 실시 예에 따른 이미지 수정 시스템(100)은 이미지 수정 프로그램과 같은 틀에 대한 전문적인 지식이나 충분한 경험이 없는 사용자가 이미지를 원하는 방식으로 빠르고 간편하게 수정 할 수 있다.
- [0032] 본 실시 예에 따른 이미지 수정 시스템이 얼굴 이미지를 수정하는 것으로 한정한다. 따라서, 이하에서는 이미지 수정 시스템을 '얼굴 이미지 수정 시스템'으로 지칭한다. 그러나 본 발명이 얼굴 이미지 수정에 제한적으로 적용되는 것은 아니며, 차량 디자인, 건축 디자인, 가전 제품 디자인 등과 관련된 모든 종류의 이미지 수정에 적용될 수 있다.
- [0033] 사용자가 간편하고 빠르게 얼굴 이미지를 수정하는 시스템을 설계하기 위해, 얼굴 이미지 수정 시스템(100)은 입력부(110), 전처리부(120), 이미지 생성부(130), 표시부(140), 저장부(150) 및 학습부(160)를 포함한다.
- [0035] **입력부(110)**
- [0036] 입력부(110)는 사용자 입력을 레이어 생성부(120)로 전달하는 구성으로, 키보드, 마우스, 터치 패드, 터치 패널과 같은 하드웨어 수단과 이러한 하드웨어 수단과 연동하도록 프로그래밍된 소프트웨어 수단(이하, '사용자 인터페이스'라 함)을 포함한다.
- [0037] 사용자는 입력부(110)를 통해 다음과 같은 작업을 수행할 수 있다.
- [0038] - 원본 얼굴 이미지(Original face image, 10)에서 사용자가 수정하고자 하는 영역을 지우는 작업.
- [0039] - 원본 이미지(10)에서 지워진 영역(erased area) 내에 사용자가 수정하고자 하는 모양(또는 형태)으로 스케치하는 작업. 여기서, 스케치는 그 모양을 알아볼 수 있는 수준을 의미하며, 전문가 수준의 스케치는 불필요하다.
- [0040] - 스케치한 모양을 사용자가 수정하고자 하는 컬러로 색칠하는 작업
- [0041] 실시 예에 따르면, 사용자가 수정하고자 하는 영역은 원본 얼굴 이미지에서 눈 영역, 귀 영역, 입 영역, 코 영

역, 헤어 영역 등을 포함한다.

[0042] 실시 예에 따르면, 사용자가 수정하고자 하는 모양(또는 형태)은 눈 영역에 착용하는 안경, 귀 영역에 착용하는 귀걸이, 입술 모양, 헤어 스타일 등을 포함한다.

[0043] 실시 예에 따르면, 사용자가 수정하고자 하는 컬러는 눈동자의 색깔, 귀걸이의 색깔, 입술의 색깔, 헤어 색깔 등을 포함한다.

[0044] 입력부(110)는, 상기의 작업들에 대응하는 사용자 입력에 따라, 원본 이미지에서 지워진 영역을 나타내는 제1 입력 값, 상기 지워진 영역에 스케치된 모양(또는 형태)를 나타내는 제2 입력 값 및 사용자가 수정하고자 하는 색깔을 나타내는 제3 입력 값을 생성하여 이를 전처리부(120)로 입력한다.

[0046] 전처리부(120)

[0047] 전처리부(120)는 상기 입력부(110)로부터 입력된 제1 내지 제3 입력 값에 따라 원본 얼굴 이미지를 전처리하여, 전처리 된 다수의 입력 이미지를 생성한다. 생성된 다수의 입력 이미지는 이미지 생성부(130)로 입력되는 정보로 사용된다.

[0048] 도 2는 본 발명의 실시 예에 따른 전처리부로부터 이미지 생성부로 입력되는 다수의 입력 이미지를 설명하기 위한 도면이다.

[0049] 도 2를 참조하면, 다수의 입력 이미지는 상기 제1 입력 값에 따라 사용자에게 의해 지워진 영역을 포함하는 원본 얼굴 이미지(11), 상기 제1 입력 값에 따라 상기 지워진 영역만을 포함하는 마스크 이미지(mask image)(13), 상기 제2 입력 값에 따라 사용자에게 의해 상기 지워진 영역에 스케치된 모양만을 나타내는 스케치 이미지(15), 상기 제3 입력 값에 따라 상기 지워진 영역에 나타나는 색깔만을 포함하는 색깔 이미지(17) 및 상기 지워진 영역의 가우시안 노이즈만을 나타내는 노이즈 이미지(19)를 포함한다.

[0051] 이미지 생성부(130)

[0052] 다시 도 1을 참조하면, 이미지 생성부(130)는 사전에 학습된 신경망을 이용하여 상기 전처리부(120)로부터 입력된 상기 다수의 입력 이미지를 분석하고, 그 분석 결과에 따라 상기 원본 얼굴 이미지(10) 내에서 상기 지워진 영역에 사용자가 합성하고자 하는 사실적인 이미지를 예측하여, 상기 예측된 사실적인 이미지를 상기 원본 얼굴 이미지(10)의 지워진 영역에 자동으로 합성한다.

[0053] 실시 예에 따르면, 상기 사전에 학습된 신경망은 생성적 적대 신경망(Generative Adversarial Networks, GANs)일 수 있다. 이에 따라, 이미지 생성부(130)는 생성적 적대 신경망(GANs)을 이용하여 원본 얼굴 이미지를 수정한다.

[0054] 생성적 적대 신경망(GANs)은 2014년 몬트리올 대학의 Ian Goodfellow와 Yoshua Bengio를 포함하는 연구자들이 작성한 논문에서 소개된 심층 신경망 기술이다.

[0055] 생성적 적대 신경망(GANs)에 대한 상세한 설명은 상기 논문으로 대신하고, 본 명세서에서는 생성적 적대 신경망(GANs)의 개략적인 동작 원리와 본 발명에 적용되도록 수정 및 변경된 부분에 대해서만 설명하기로 한다.

[0056] 생성적 적대 신경망(GANs)은 생성기(Generator)로 불릴 수 있는 생성적 신경망(162, Generative neural network)과 판별기(Discriminator)로 불릴 수 있는 판별적 신경망(164, Discriminative neural network)으로 구성된 심층 신경망 구조를 갖는다.

[0057] 생성적 신경망(162)은 판별적 신경망(164)으로 전달하는 새로운 이미지를 생성하고, 판별적 신경망(164)은 생성적 신경망(162)로부터 입력된 새로운 이미지의 진위 여부를 판단한다.

[0058] 생성적 신경망(162)은 새로운 이미지를 생성하고, 판별적 신경망(164)은 생성적 신경망(162)으로부터 입력되는 새로운 이미지의 진위를 판별하고, 그 판별 결과를 생성적 신경망(162)의 입력으로 피드백한다.

[0059] 이때, 생성적 신경망(162)은 자신이 생성한 새로운 이미지를 판별적 신경망(164)이 진짜 이미지(real image)로 판별하도록 이미지 생성 과정을 학습하고, 반대로 판별적 신경망(164)은 생성적 신경망(162)으로부터 입력된 새로운 이미지를 가짜 이미지로 판별하도록 이미지 판별 과정을 학습한다.

- [0060] 이처럼 생성적 신경망(162)과 판별적 신경망(164)은 제로섬 게임처럼 서로 반대되는 목적 함수 또는 손실 함수를 통해 학습된다. 즉, 생성적 신경망(162)과 판별적 신경망(164)은 서로 적대적인(Adversarial) 관계를 유지하도록 학습되고 진화한다.
- [0061] 이미지 생성부(130)는 생성적 신경망(162)과 판별적 신경망(164) 중에서 학습이 완료된 생성적 신경망(162)만을 이용하여 원본 얼굴 이미지에서 지워진 영역에 합성되는 사실적인 이미지를 생성한다.
- [0063] **저장부(150)**
- [0064] 저장부(150)에는 이미지 생성부(130)가 사용하는 생성적 신경망(162)이 저장된다. 생성적 신경망(162)은 생성 알고리즘(Generative Algorithms)이라는 용어로 대체될 수 있다. 유사하게, 판별적 신경망(164)은 판별 알고리즘(Discriminative Algorithms)이라는 용어로 대체될 수 있다. 저장부(150)는 휘발성 메모리 및 비휘발성 메모리로 구현될 수 있다.
- [0066] **학습부(160)**
- [0067] 학습부(160)는 사전에 수집된 대용량의 훈련 데이터를 이용하여 생성적 적대 신경망(GANs), 즉, 생성적 신경망(162)과 판별적 신경망(164)을 학습시킨다.
- [0068] 여기서, 대용량의 훈련 데이터는 대용량의 훈련용 얼굴 이미지, 서로 다른 위치에서 지워진 영역을 포함하도록 구성된 대용량의 훈련용 얼굴 이미지, 상기 지워진 영역만을 나타내는 대용량의 훈련용 마스크 이미지, 상기 지워진 영역에 다양한 모양이 스케치된 대용량의 훈련용 스케치 이미지, 상기 지워진 영역에 서로 다른 색깔이 채워진 대용량의 훈련용 색깔 이미지 및 대용량의 노이즈 이미지를 포함한다.
- [0069] 이미지 생성부는 학습이 완료된 생성적 신경망(162)과 판별적 신경망(164) 중에서 생성적 신경망(162)만을 이용하므로, 학습부(160)는 학습이 완료된 생성적 신경망(162)과 판별적 신경망(164) 중에서 생성적 신경망(162)을 저장부(150)에 저장함으로써, 이미지 생성부(130)는 저장부(150)에 저장된 생성적 신경망(162)을 이용할 수 있게 된다.
- [0071] **표시부(140)**
- [0072] 표시부(140)는 이미지 생성부(130)에서 생성한 이미지를 표시하는 구성으로, 상기 원본 얼굴 이미지(10) 내에서 사용자 입력에 따라 지워진 영역에 사용자가 합성하고자 하는 사실적인 이미지가 자동으로 합성된 합성 이미지를 표시한다. 이러한 표시부(140)는 LCD, LED, OLE 등일 수 있다.
- [0073] 또한 표시부(140)는 사용자 인터페이스의 화면 구성을 표시한다.
- [0075] 도 3 내지 도 7은 본 발명의 실시 예에 따른 사용자 인터페이스의 화면 구성을 나타내는 도면들이다.
- [0076] 사용자 인터페이스는 사용자에게 원본 얼굴이미지의 수정에 필요한 입력, 즉, 이미지 생성부로 입력되는 입력 정보를 생성하기 위한 환경을 제공한다.
- [0077] 이를 위해, 사용자 인터페이스의 화면 구성은 원본 얼굴 이미지 및 원본 얼굴 이미지에 대한 진행 과정이 표시되는 영역(31)과 수정된 결과(합성 이미지)가 표시되는 영역(32)을 포함한다.
- [0078] 또한, 사용자 인터페이스의 화면 구성은 입력 정보 생성과 관련된 다수의 아이콘 형태의 버튼(33, 34, 35, 36 및 37)을 포함한다.
- [0079] 먼저, 도 3을 참조하면, 버튼(33)은 원본 이미지의 불러오기 기능을 제공한다. 사용자가 버튼(33)을 터치 또는 클릭하면, 원본 이미지가 영역(31)에 표시된다.
- [0080] 도 4를 참조하면, 버튼(34)은 사용자가 원본 얼굴 이미지(10)에서 수정할 영역을 지우는 기능을 제공한다. 예를 들면, 사용자가 버튼(34)을 터치 또는 클릭하면, 지우개 형태의 아이콘 도구가 생성되고, 지우개 형태의 아이콘 도구를 이용하여 원본 이미지(10)에서 수정할 부분을 지운다. 도 4는 사용자가 원본 얼굴 이미지(10)에서 수정

하고자 하는 영역이 눈 영역, 코 영역, 입술 영역인 경우를 도시한 것이다.

- [0081] 도 5를 참조하면, 버튼(35)은 사용자가 원본 얼굴 이미지(10)에서 지워진 영역에 수정하고자 하는 모양을 스케치하는 기능을 제공한다. 예를 들면, 사용자가 버튼(35)을 클릭 또는 터치하면, 펜 형태의 아이콘 도구가 생성되고, 펜 형태의 아이콘 도구를 이용하여 원본 얼굴 이미지(10)에서 지워진 영역에 수정하고자 하는 모양을 스케치한다. 이때, 스케치는 그 모양을 알아볼 수 있는 수준을 의미하며, 전문가 수준의 스케치는 불필요하다.
- [0082] 도 6을 참조하면, 버튼(36)은 스케치한 모양을 사용자가 수정하고자 하는 컬러로 색칠하는 기능을 제공한다. 예를 들면, 사용자가 버튼(36)을 클릭 또는 터치하면, 붓 형태의 아이콘 도구가 생성되고, 붓 형태의 아이콘 도구를 이용하여 스케치한 모양 또는 그 주변에 사용자가 결정한 색깔을 칠한다. 도 6에서는 눈동자의 색깔을 사용자가 정한 색깔로 칠한 경우를 도시한 것이다.
- [0083] 도 7을 참조하면, 버튼(37)은 수정된 결과를 보여주는 기능을 제공한다. 사용자가 버튼(37)을 클릭 또는 터치하면, 영역(32)에 상기 원본 얼굴 이미지(10) 내에서 사용자 입력에 따라 지워진 영역에 사용자가 합성하고자 하는 사실적인 이미지가 자동으로 합성된 합성 이미지가 표시된다. 도 7에서는 생성적 신경망(162)의 이미지 생성 과정에서 수정된 코 형상 이미지, 입술 형상 이미지 및 눈동자 색상 이미지가 원본 얼굴 이미지에 합성된 예를 도시한 것이다.
- [0084] 도 7에서는 생성적 신경망(162)이 마스크 작업(원본 얼굴 이미지에서 사용자가 수정하고자 하는 영역을 지우는 작업), 스케치 작업 및 색깔 작업에 따라 생성된 모든 입력 정보를 이용하여 합성된 이미지를 생성한 예를 도시하고 있으나, 합성된 이미지를 생성하기 위해 모든 입력 정보가 필요한 것은 아니다.
- [0085] 예를 들면, 생성적 신경망(162)이 마스크 작업에 따라 생성된 오직 하나의 입력 정보, 즉, 마스크 이미지만을 이용하여 합성된 이미지를 생성할 수도 있다. 이는 생성적 신경망(162)이 대용량의 훈련 데이터, 즉, 대용량의 마스크 이미지, 대용량의 스케치 이미지 및 대용량의 색깔 이미지를 모두 이용하여 학습된 것이기 때문이다.
- [0086] 이것은 원본 얼굴 이미지(10)에서 사용자가 수정하고자 하는 영역의 위치를 나타내는 마스크 이미지는 반드시 생성적 신경망(162)의 입력으로 사용되어야 함을 의미하기도 한다. 즉, 마스크 이미지는 사용자 수정하고자 하는 의도를 생성적 신경망(162)에게 알리는 최소한 정보로 해석할 수 있다.
- [0087] 물론, 생성적 신경망(162)으로 입력되는 정보의 양이 많을수록 사용자가 수정하고자 하는 의도에 가장 부합하는 수정된 얼굴 이미지가 생성될 확률은 당연히 가장 높을 것이다.
- [0089] 도 8은 본 발명에 적용되는 생성적 적대 신경망(GANs)의 전체 네트워크 구조를 나타내는 도면이다.
- [0090] 전술한 바와 같이, 생성적 신경망(162)을 학습시키기 위해서는 생성적 신경망(162)과 판별적 신경망(164)으로 구성된 전체 네트워크(GANs)를 같이 학습시켜야 한다.
- [0091] 학습이 완료되면, 얼굴 이미지의 수정에는 생성적 신경망(162)만이 사용되므로, 생성적 신경망(162)의 속도 향상과 수정된 얼굴 이미지의 사실성을 높이기 위해, 생성적 적대 신경망(GANs)은, 도 8에 도시된 바와 같은 구조로 구성될 수 있다.
- [0092] 도 8에 도시된 바와 같이, 생성적 신경망(162)은 U-net 구조로 이루어져 있으며 게이트 컨볼루션 레이어(Gated convolution layer)들, 확장된 게이트 컨볼루션 레이어(Dilated gated convolution layer)들, 디컨볼루션 레이어(Deconvolution layer)들을 포함한다. 각 컨볼루션 레이어는 채널 수에 따라 서로 다른 사이즈의 육면체 형상으로 도시된다.
- [0093] 생성적 신경망(162)은 일반적인 컨볼루션 레이어(convolution layer)가 아니라 게이트 컨볼루션 레이어(gated convolution layer)를 사용함에 특징이 있다. 일반적인 컨볼루션 레이어는 이전단의 컨볼루션 레이어로부터 입력된 특징값에 대해 다른 특징값을 출력한다. 이에 반해, 게이트 컨볼루션 레이어는 이전단의 게이트 컨볼루션 레이어로부터 입력된 특징에 대해 다른 특징값과 마스크 이미지에 대한 특징값을 출력한다. 즉, 일반적인 컨볼루션 레이어는 하나의 데이터를 출력하고, 게이트 컨볼루션 레이어는 두 개의 데이터를 출력하는 점에서 차이가 있다. 게이트 컨볼루션 레이어에서 출력되는 2개의 데이터 중 하나의 데이터는 마스크 이미지의 특징값이고, 이 마스크 이미지의 특징값은 인접하지 않은 다른 게이트 컨볼루션 레이어로 입력된다. 도 8에서 점선의 화살표는 현재의 게이트 컨볼루션 레이어가 출력하는 마스크 이미지의 특징값이 인접하지 않은 다른 게이트 컨볼루션 레이어로의 입력을 나타낸 것이다.

[0094] 판별적 신경망(164)은 스펙트럼 정규화(Spectral normalization)가 적용된 게이트 컨볼루션 레이어(gated convolution layer)로 이루어진 patchGAN 형태이다. 판별적 신경망(164)을 학습시키는 과정은 데이터셋(dataset)에 대하여 일반적인 생성적 적대 신경망(GANs)을 학습시키는 방법과 동일하다.

[0095] 다만, 판별적 신경망(164)의 학습에 사용되는 손실 변수(Loss)는 일반적인 손실(Loss)와는 차이가 있다. 여기서, 손실(Loss)는 원본 얼굴 이미지와 생성적 신경망(162)이 생성한 새로운 얼굴 이미지와의 차이를 의미한다.

[0096] 본 발명의 실시 예에 따른 생성적 신경망(162)을 학습시키는데 사용되는 파라미터(L_G)는 아래의 수학적 식 1과 같고, 판별적 신경망(164)을 학습시키는데 사용되는 파라미터(L_D)는 아래의 수학적 식 2와 같다.

수학적 식 1

$$L_G = L_{per-pixel} + \sigma L_{percept} + \beta L_{G_SN} + \gamma(L_{style}(I_{gen}) + L_{style}(I_{comp})) + \nu L_{tv} + \epsilon \mathbb{E} [D(I_{gt})^2]$$

[0098]

수학적 식 2

$$L_D = \mathbb{E} [1 - D(I_{gt})] + \mathbb{E} [1 + D(I_{comp})] + \theta L_{GP}$$

[0099]

[0101] 파라미터 L_G 는 생성적 신경망(162)의 레이어(layer)를 학습시키는 손실(loss)이다. 여기서, 생성적 신경망(162)의 레이어(layer)는 도 8에 도시된 바와 같이, 다수의 게이트 컨볼루션 레이어(Gated convolution layer), 다수의 확장된 게이트 컨볼루션 레이어(Dilated gated convolution layer) 및 다수의 디컨볼루션 레이어(Deconvolution layer)를 포함한다.

[0102] 파라미터 L_D 는 판별적 신경망(164)의 레이어(layer)를 학습시키는 손실(loss)이다. 여기서, 판별적 신경망(164)의 레이어(layer)는 다수의 스펙트럼 정규화(Spectral Normalization: SN) 컨볼루션 레이어를 포함한다.

[0104] 수학적 식 1에서 $L_{per-pixel}$ 은 아래의 수학적 식 3과 같다.

수학적 식 3

$$L_{per-pixel} = \frac{1}{N_{I_{gt}}} \|M \odot (I_{gen} - I_{gt})\|_1 + \alpha \frac{1}{N_{I_{gt}}} \|(1 - M) \odot (I_{gen} - I_{gt})\|_1$$

[0105]

[0106] 여기서, M 은 마스크 이미지에 포함된 지워진 영역을 나타내는 1채널의 픽셀값으로서, 예를 들어, '1'일 수 있다. 이때, 마스크 이미지에서 지워진 영역을 제외한 나머지 영역의 각 픽셀값은 '0'이다. I_{gen} 은 생성적 신경망(162)이 지워진 영역을 갖는 원본 얼굴 이미지, 상기 지워진 영역만을 갖는 마스크 이미지, 스케치 이미지,

색깔 이미지 및 노이즈 이미지를 입력받아서 생성한 새로운 얼굴 이미지, 즉, 원본 얼굴 이미지에서 지워진 영역이 다른 이미지로 채워진 이미지를 의미하고, I_{gen} 는 생성적 신경망(162)이 생성한 새로운 얼굴 이미지를 3차원 벡터 공간에서 표현한 3차원 벡터값일 수 있다. I_{gt} 는 지워진 영역이 없는 원본 얼굴 이미지를 의미하고, I_{gt} 는 원본 얼굴 이미지를 3차원 벡터 공간에서 표현한 3차원 벡터값일 수 있다. $L_{per-pixel}$ 은 생성적 신경망에서 생성한 새로운 얼굴 이미지와 원본 얼굴 이미지 간의 거리 L1를 계산하는 파라미터 이다. 거리 L은 새로운 얼굴 이미지의 픽셀과 상기 픽셀에 대응하는 원본 얼굴 이미지의 픽셀 간의 거리이다. α 는 원본 얼굴 이미지 내에서 지워진 영역에 대한 손실(loss)를 강화시키기 위한(줄이기 위한) 가중치(weight)이다. $\odot$ 는 요소 별 곱셈 연산(Element-wise multiplication)을 나타내는 기호이고, $N_{I_{gt}}$ 는 이미지 크기에 따른 총 픽셀 수로서, 예컨대, 786432(= 512 x 512 x 3)일 수 있다.

[0108] 수학식 1에서 $L_{percept}$ 는 아래의 수학식 4와 같다.

수학식 4

$$L_{percept} = \sum_q \frac{\|\Theta_q(I_{gen}) - \Theta_q(I_{gt})\|_1}{N_{\Theta_q(I_{gt})}} + \sum_q \frac{\|\Theta_q(I_{comp}) - \Theta_q(I_{gt})\|_1}{N_{\Theta_q(I_{gt})}}$$

[0109]

[0110] $L_{percept}$ 는 스타일 손실(style-loss)로 활용되는 손실(loss) 중 하나이다. Θ_q 는 대량의 훈련 데이터 셋에 대한 이미지 분류(classification)을 위하여 학습된 생성적 신경망의 q번째 레이어(layer)의 특징을 의미한다. 즉, $L_{percept}$ 는 기존의 학습된 네트워크를 활용하여 픽셀 단위로 계산된 특징에 대한 손실(loss)이다. $N_{\Theta_q(I_{gt})}$ 에서 N은 픽셀 총 개수이고, Θ_q 는 생성적 신경망과는 다른 기존의 이미지 분류를 위해 대량의 이미지를 학습한 인공 신경망의 q번째 특징맵이다. 즉, $N_{\Theta_q(I_{gt})}$ 는 원본 얼굴 이미지를 이미지 분류 인공신경망에 통과시켜서 얻은 q번째 특징맵의 총 픽셀 수이다. $\|\sim\|_1$ 는 L1 거리에 대한 손실(loss)로서, 모든 픽셀 차이의 절대값의 합이다. I_{comp} 는 지워진 영역을 갖는 원본 얼굴 이미지와 생성적 신경망(162)에 의해 생성된 새로운 얼굴 이미지(I_{gen})에서 상기 지워진 영역에 대응하는 부분만을 추출한 이미지를 합성한 합성 이미지이다. 즉, I_{gen} 는 생성적 신경망(162)의 예측 및/또는 추론 과정을 통해 생성된 이미지를 의미하는 것인 반면, I_{comp} 는 예측 및/또는 추론 과정이 아니라 지워진 영역을 갖는 원본 이미지와 생성적 신경망(162)에 의해 생성된 이미지(I_{gen})로부터 추출된 이미지(상기 지워진 영역에 대응하는 이미지)를 합성하는 합성 과정을 통해 생성된 이미지를 의미한다. I_{comp} 는 3차원 벡터 공간에서 표현되는 3차원 벡터 값일 수 있다.

[0111] 수학식 1에서 $L_{style}(I_{gen})$ 은 아래의 수학식 5와 같다.

수학식 5

$$L_{style}(I) = \sum_q \frac{1}{C_q C_q} \left\| \frac{(G_q(I) - G_q(I_{gt}))}{N_q} \right\|_1$$

[0112]

[0113] L_{style} 은 대량의 얼굴 데이터 셋에 대하여 이미지 분류(classification)를 위하여 학습된 네트워크의 q번째 layer

특징에 대한 손실(loss)이다. 이러한 L_{style} 을 계산하기 위해 Gram matrix가 활용될 수 있다. C_q 는 q 번째 레이어의 채널 수이고, N_q 는 q 번째 레이어의 전체 픽셀 수이고, G_q 는 q 번째 레이어의 Gram matrix 값을 나타낸다.

추가로, 생성적 신경망(162)을 학습시키는데 분산 손실(variance loss)이 이용된다. 분산 손실(variance loss)은 판별적 신경망(164)이 생성적 신경망(162)로부터 입력된 새로운 얼굴 이미지 내의 모드 픽셀을 1 pixel만큼 강제로 이동시켜서 획득한 새로운 얼굴 이미지와 원본 얼굴 이미지 간의 차이를 나타내는 손실(loss)이다. 이러한 분산 손실(variance loss)은 아래의 수학적 식 6과 같다.

수학적 식 6

$$L_{tv-col} = \sum_{(i,j) \in R} \frac{\|I_{comp}^{i,j+1} - I_{comp}^{i,j}\|_1}{N_{comp}}$$

$$L_{tv-row} = \sum_{(i,j) \in R} \frac{\|I_{comp}^{i+1,j} - I_{comp}^{i,j}\|_1}{N_{comp}}$$

분산 손실(variance loss)은 L_{tv-col} , L_{tv-row} 를 포함하고, 위의 수학적 식 6에서 아래 첨자 comp는 원본 얼굴 이미지에서 지워진 영역이 사용자 입력에 따른 이미지로 채워진 이미지를 의미한다. N_{comp} 는 생성적 신경망(162)이 생성한 새로운 얼굴 이미지의 전체 픽셀 수이다.

생성적 신경망(162)은 위와 같은 분산 손실(variance loss)을 이용하여 블러링에 대하여 강인하도록 학습된다.

마지막으로 판별적 신경망(164)이 훈련 데이터에 수렴하지 못하도록 gradient penalty term(L_{GP})이 아래와 같이 추가된다. 여기서, 판별적 신경망(164)이 훈련 데이터에 수렴하지 못하게 한다는 의미는 판별적 신경망(164)이 생성적 신경망(162)보다 더 높은 레벨로 학습되어서는 안됨을 의미하는 것이다. 즉, 생성적 신경망(162)과 판별적 신경망(164)은 서로 적대적인 관계를 유지하도록 동일한 레벨로 학습되어야 함을 의미한다.

수학적 식 7

$$L_{GP} = \mathbb{E} \left[(\|\nabla_U D(U) \odot M\|_2 - 1)^2 \right]$$

이상 설명한 바와 같이, 파라미터 L_G 와 파라미터 L_D 를 감소시키는 방향으로 생성적 신경망(162)과 판별적 신경망(164)을 각각 학습시키는 과정이 완료되면, 이후, 이미지 생성부(130)는 생성적 신경망(162)만을 이용하여 새로운 얼굴 이미지를 생성한다.

생성적 신경망(162)은 크기가 가볍기 때문에, 새로운 얼굴 이미지를 생성하는데 걸리는 시간은 일반적인 CPU를 기준으로 2초 이내이다.

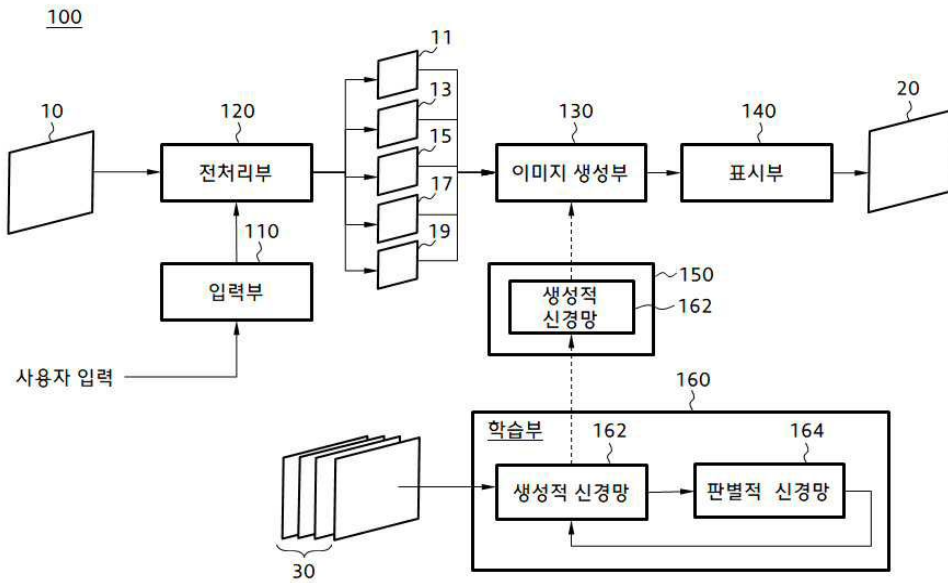
도 9는 본 발명의 실시 예에 따른 이미지 수정 방법을 보여주는 흐름도이다.

- [0128] 도 9를 참조하면, S910에서, 생성적 적대 신경망(160)을 학습시키는 학습 과정이 수행된다. 이 학습 과정은 서로 적대적인(Adversarial) 관계를 유지하도록 생성적 신경망과 판별적 신경망을 학습시키는 과정이다. 아래에서 설명하겠지만, 생성적 신경망(162)과 판별적 신경망(164)의 학습이 완료되면, 생성적 신경망만을 이용하여 원본 이미지에 합성될 이미지가 예측되고, 예측된 이미지와 원본 이미지가 합성된다. 생성적 신경망(162)을 학습시키기 위해, 사전에 수집된 대량의 훈련 데이터가 이용된다. 대량의 훈련 데이터는 훈련용 원본 이미지들, 상기 지워진 영역만을 포함하는 훈련용 마스크 이미지들, 상기 지워진 영역에 다양한 모양이 스케치된 훈련용 스케치 이미지들, 상기 지워진 영역에 다양한 색깔이 채워진 훈련용 색깔 이미지들을 포함한다. 추가로, 상기 지워진 영역의 노이즈를 나타내는 노이즈 이미지가 더 포함될 수 있다. 판별적 신경망(164)과 적대적 관계를 유지하도록, 생성적 신경망(162)을 학습시키는데, 상기 판별적 신경망(164)으로부터 출력되는 손실(L_G)이 이용될 수 있다. 생성적 신경망(162)을 학습시키는 과정은 상기 손실(L_G)을 줄이는 방향으로 학습시키는 과정이다. 이때, 상기 손실(L_G)은, 상기 대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 포함한다. 이에 대한 설명은 전술한 수학적 식 1 및 3에 대한 설명으로 대신한다. 또한 생성적 신경망(162)을 학습시키는 과정은 상기 대량의 훈련 데이터에 포함된 훈련용 마스크 이미지 내에서 지워진 영역의 픽셀값(M)과 상기 원본 이미지와 상기 새로운 이미지 사이의 픽셀 차이값($I_{gen}-I_{gt}$)을 이용하여, 상기 생성적 신경망에 포함된 게이트 컨볼루션 레이어(gated convolution layer)를 학습시키는 과정일 수 있다.
- [0129] 이어, S920에서, 생성적 적대 신경망(160)에 대한 학습 과정이 완료되면, 즉, 생성적 신경망(162)과 판별적 신경망(164)의 학습이 완료되면, 학습이 완료된 생성적 신경망(162)을 저장부(150)에 저장하는 과정이 수행된다. 이렇게 함으로써, 이미지 생성부(130)는 생성적 신경망(162)의 접근이 가능하다.
- [0130] 이어, S930에서, 원본 이미지(10)에 대한 전처리 과정이 수행되고, 이러한 전처리 과정을 통해 원본 이미지(10)로부터 상기 학습이 완료된 생성적 신경망(162)으로 입력되는 정보가 생성된다. 생성적 신경망(162)으로 입력되는 정보는 상기 원본 이미지 내에서 사용자 입력에 의해 지워진 영역만을 포함하는 마스크(mask) 이미지, 상기 사용자 입력에 의해 상기 지워진 영역에 스케치된 모양만을 포함하는 스케치 이미지 및 상기 사용자 입력에 의해 상기 지워진 영역에 칠해진 색깔만을 포함하는 색깔 이미지를 포함한다. 이때, 생성적 신경망(162)으로 입력되는 정보는 상기 원본 이미지 내에서 사용자 입력에 의해 지워진 영역만을 포함하는 마스크(mask) 이미지로만 구성될 수도 있다. 이처럼 본 발명에서는 사용자에 의해 지워진 영역의 위치 정보만을 포함하고 있는 마스크 이미지만을 생성적 신경망(162)에 입력하는 것만으로도 원본 이미지의 일부를 사용자 의도에 맞게 수정할 수 있다.
- [0131] 이어, S940에서, 이미지 생성부(130)가 학습이 완료된 상기 생성적 적대 신경망, 즉, 학습이 완료된 생성적 신경망(162)을 이용하여 상기 마스크(mask) 이미지, 스케치 이미지 및 색깔 이미지를 분석하여, 그 분석 결과에 따라 상기 원본 이미지 내에서 지워진 영역에 합성될 이미지를 예측하는 과정이 수행된다.
- [0132] 이어, S950에서, 이미지 생성부(130)가 상기 예측된 이미지를 상기 원본 이미지 내에서 상기 지워진 영역에 합성하여 새로운 이미지를 생성하고, 표시부(140)가 상기 생성된 새로운 이미지를 표시하여, 사용자에게 제공된다.
- [0134] 도 10 내지 12는 본 발명의 이미지 수정 방법에 따라 생성된 합성 이미지들의 예시한 도면들이다.
- [0135] 도 10의 (A) 및 (B)에 도시한 바와 같이, 본 발명에 따르면, 원본 얼굴 이미지에서 입 모양, 눈 형상을 수정하여 새로운 이미지를 자동으로 생성할 수 있다. 또한, 도 10의 (C)에 도시한 바와 같이, 원본 얼굴 이미지에서 안경을 지워서 새로운 이미지를 자동으로 생성할 수 있다. 여기서 중요한 점은 생성적 신경망의 입력 정보(Free-form input)를 생성하기 위해 지워진 영역에 스케치되는 모양은 전문가 수준의 스케치를 필요로 하지 않는 점이다. 즉, 본 발명은 대량의 훈련 데이터를 이용하여 사전에 학습시킨 생성적 신경망을 이용하기 때문에, 마스크 이미지로부터 지워진 영역의 위치 정보와 일반인 수준으로 스케치된 모양만으로 원본 얼굴 이미지에 합성될 이미지의 추론과 예측이 가능하다. 이것은 마스크 이미지, 스케치 이미지 또는 색깔 이미지와 같은 단순한 이미지 입력만을 이용하여 원본 얼굴 이미지를 사실적으로 수정할 수 있음을 의미한다.
- [0136] 또한, 본 발명에 따르면, 도 11의 (A)에 도시된 바와 같이, 헤어 스타일의 일부를 수정하거나 (B)에 도시된 바와 같이, 헤어 스타일의 전부를 수정할 수 있다.

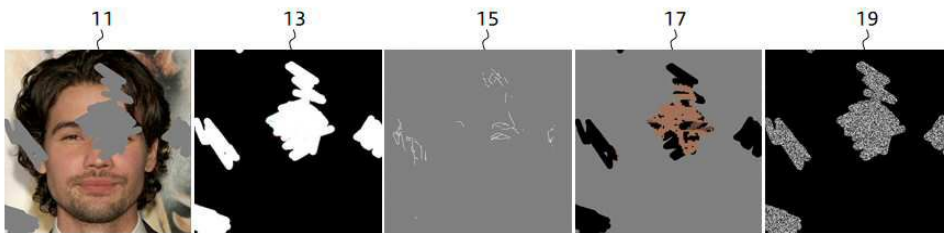
- [0137] 또한, 본 발명에 따르면, 도 12의 (A)에 도시된 바와 같이, 원본 얼굴 이미지에서 귀 영역의 일부 영역을 지우고, 그 지운 영역에 사용자가 간단하게 스케치한 귀걸이 모양을 생성적 신경망의 입력 정보로 구성하면, 원본 얼굴 이미지에 귀걸이 이미지가 사실적으로 합성된 이미지를 자동으로 생성할 수 있다.
- [0138] 또한, 도 12의 (B)에 도시된 바와 같이, 원본 얼굴 이미지에 포함된 귀걸이를 다른 형상의 귀걸이로 수정하는 것도 가능할 것이다.
- [0140] 이상에서 설명된 이미지 수정 시스템은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다.
- [0141] 예를 들어, 실시 예들에서 설명된 전처리부(120), 이미지 생성부(130), 학습부(160)와 같은 구성요소들은, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 그래픽 프로세서(GPU), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서로 구현될 수 있다.
- [0142] 또한, 이미지 수정 시스템은 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 이미지 수정 시스템은 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다.
- [0143] 실시 예에 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. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0144] 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.
- [0145] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

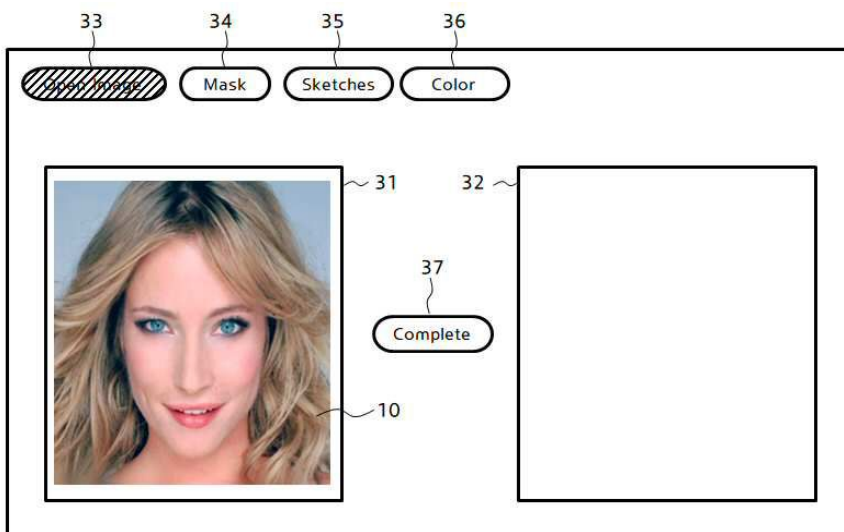
도면1



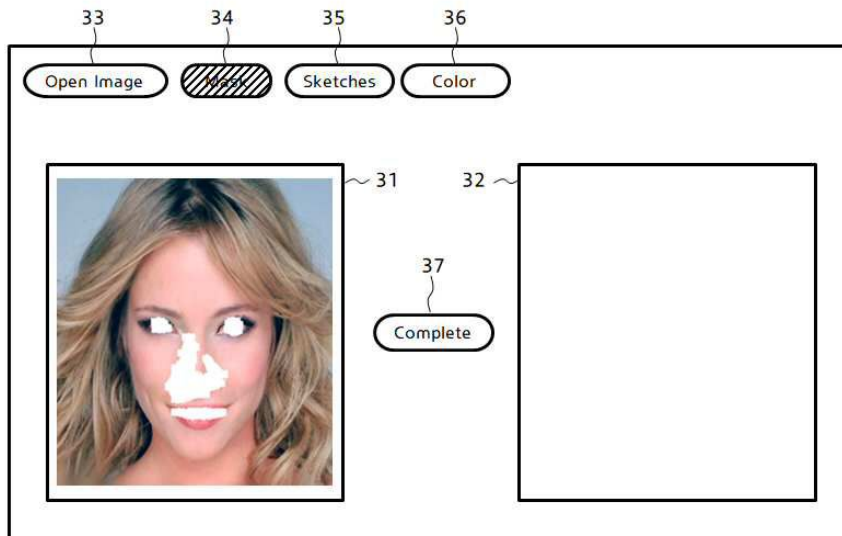
도면2



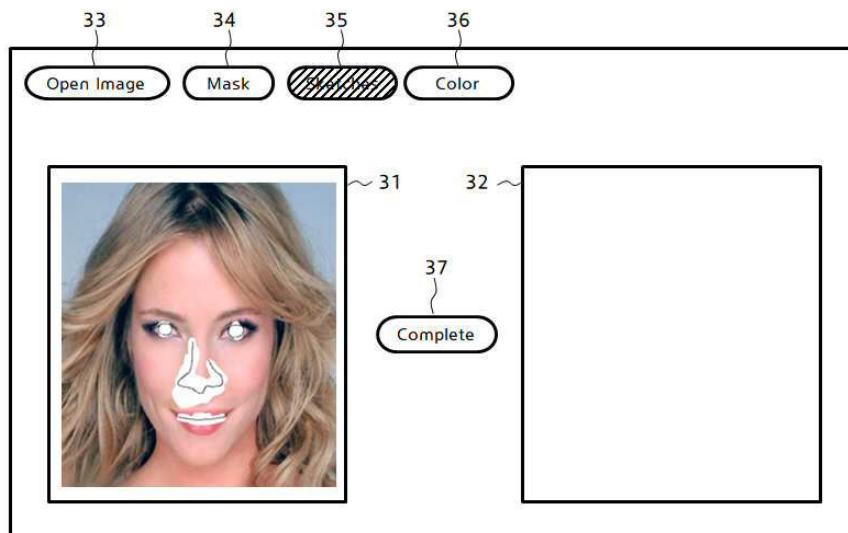
도면3



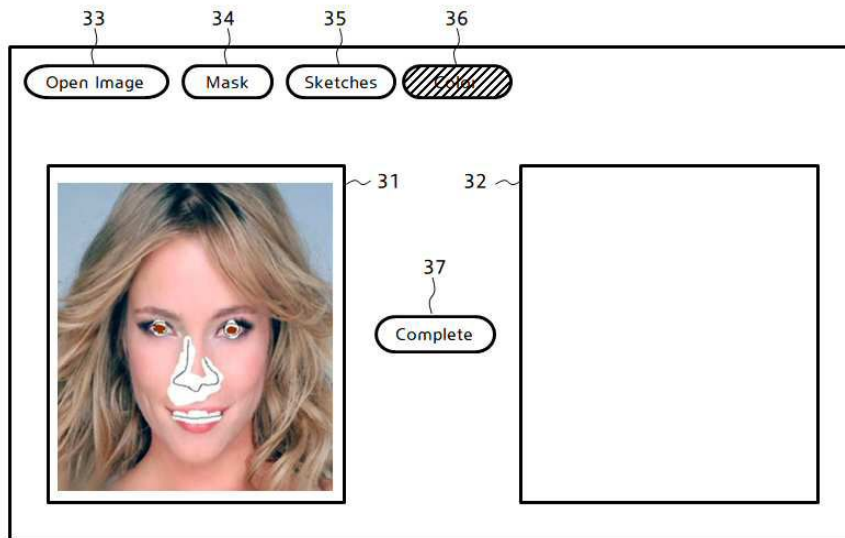
도면4



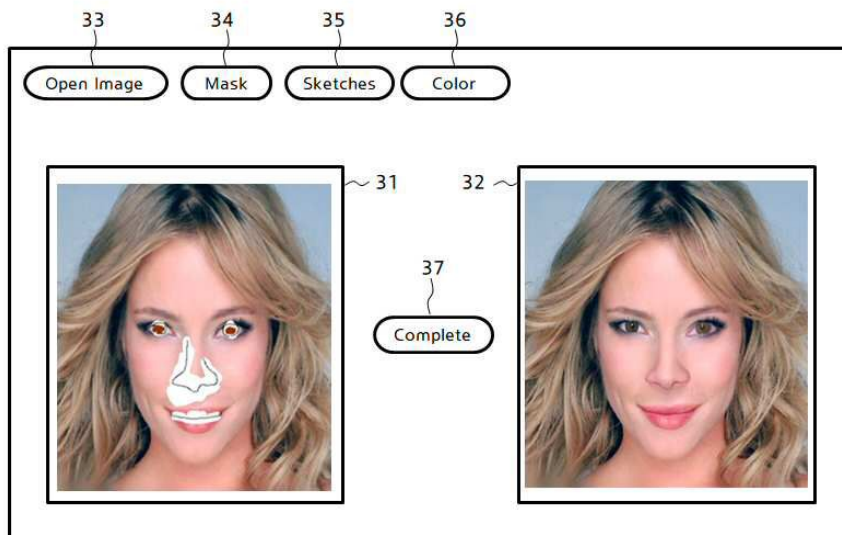
도면5



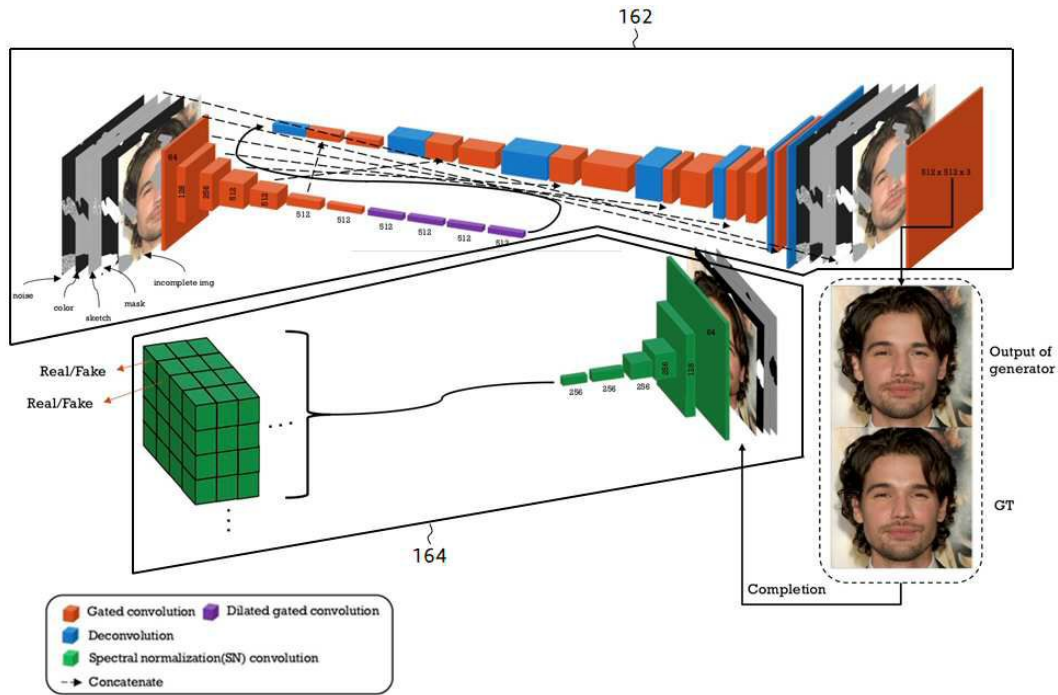
도면6



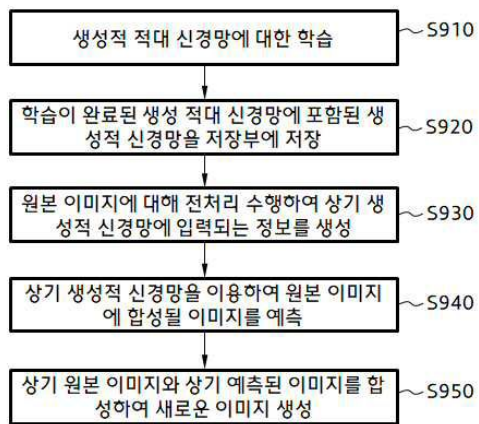
도면7



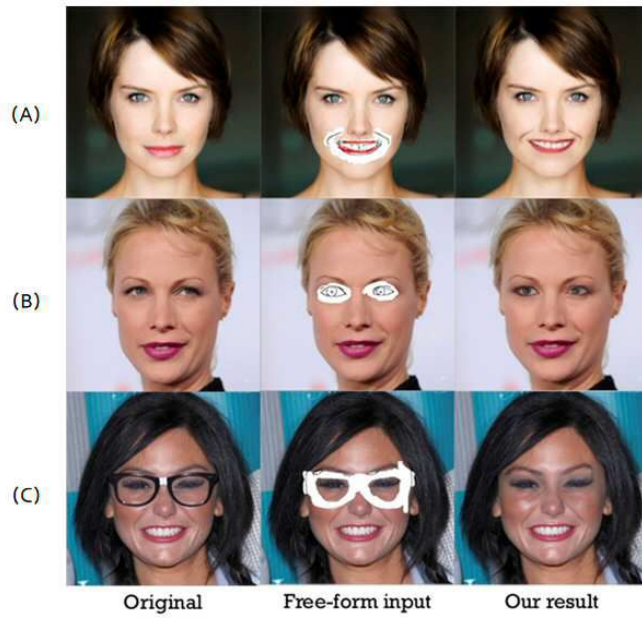
도면8



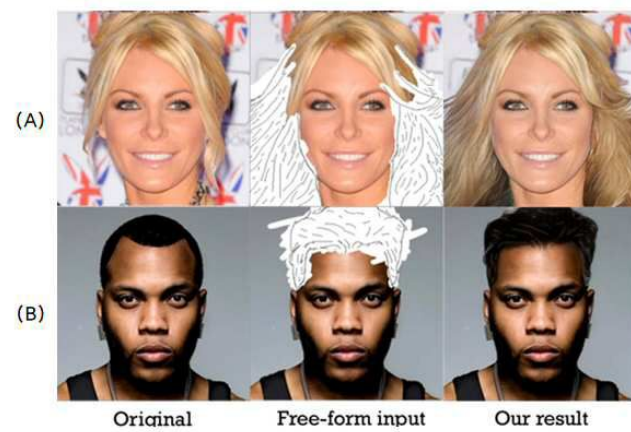
도면9



도면10



도면11



도면12

