



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년05월03일
(11) 등록번호 10-2663519
(24) 등록일자 2024년04월30일

(51) 국제특허분류(Int. Cl.)
G06T 3/00 (2024.01) G06F 18/00 (2023.01)
G06N 3/04 (2023.01) G06N 3/08 (2023.01)
G06T 11/00 (2006.01)
(52) CPC특허분류
G06T 3/04 (2024.01)
G06F 18/2135 (2023.01)
(21) 출원번호 10-2021-7013184
(22) 출원일자(국제) 2019년09월05일
심사청구일자 2022년09월05일
(85) 번역문제출일자 2021년04월30일
(65) 공개번호 10-2021-0083276
(43) 공개일자 2021년07월06일
(86) 국제출원번호 PCT/US2019/049619
(87) 국제공개번호 WO 2020/091891
국제공개일자 2020년05월07일
(30) 우선권주장
201811294026.6 2018년10월31일 중국(CN)
(56) 선행기술조사문헌
US20180307947 A1*
Ziqiang Zheng et al., 'Unpaired
photo-to-caricature translation on faces in
the wild', Neurocomputing, Volume 355, 25
August 2019.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
마이크로소프트 테크놀로지 라이선싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
리아오 징
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 마이크로소프트 테크놀로지
라이선싱, 엘엘씨
유안 루
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 마이크로소프트 테크놀로지
라이선싱, 엘엘씨
카오 카이디
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 마이크로소프트 테크놀로지
라이선싱, 엘엘씨
(74) 대리인
제일특허법인(유)

전체 청구항 수 : 총 11 항

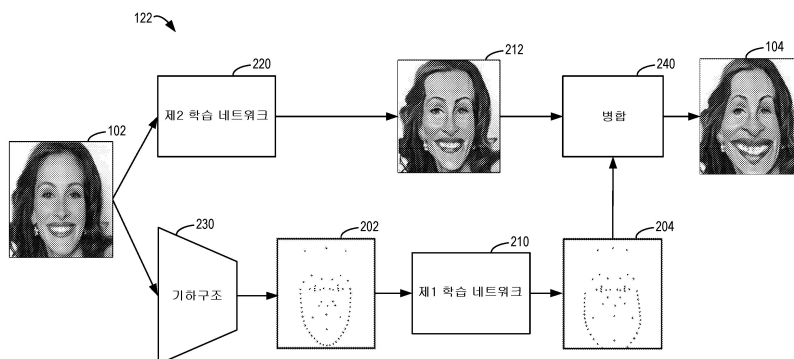
심사관 : 정성윤

(54) 발명의 명칭 교차 도메인 이미지 변환 기법

(57) 요약

본 개시의 구현예는 교차 도메인 이미지 변환에 대한 해결책을 제공한다. 이 해결책에서, 제1 이미지 도메인으로 부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크가, 제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 결정되고, 두 도메인의 이미지들은 상이한 스타일을 가지며, 이들 (뒷면에 계속)

대표도



이미지 내의 객체들은 서로에 대해 기하학적 변형을 갖는다. 제2 이미지에 대해 제2 도메인으로부터 제1 도메인으로의 기하학적 변형이 수행되거나 또는 제1 이미지에 대해 제1 도메인으로부터 제2 도메인으로의 기하학적 변형이 수행되어, 중간 이미지를 생성한다. 제1 이미지와 중간 이미지에 기초하여 또는 제2 이미지와 생성된 중간 이미지에 기초하여 제1 도메인으로부터 제2 도메인으로의 스타일 트랜스퍼(style transfer)를 위한 제2 학습 네트워크가 결정된다. 따라서, 교차 도메인 이미지 변환을 위한 학습 네트워크의 처리 정확도가 개선될 수 있고 복잡도가 낮아진다.

(52) CPC특허분류

G06N 3/044 (2023.01)

G06N 3/045 (2023.01)

G06N 3/084 (2023.01)

G06T 11/00 (2013.01)

G06T 3/18 (2024.01)

명세서

청구범위

청구항 1

컴퓨터로 구현되는 방법으로서,

제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 상기 제1 이미지 도메인으로부터 상기 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크(learning network)를 결정하는 단계 - 상기 제1 이미지 도메인 및 상기 제2 이미지 도메인의 이미지들은 상이한 스타일을 가지며, 상기 이미지들 내의 객체들은 서로에 대해 기하학적 변형을 가지며, 상기 제1 학습 네트워크는 제1 GAN(generative adversarial network)에 기초함 - 와,

상기 제2 이미지에 대해 상기 제2 이미지 도메인으로부터 상기 제1 이미지 도메인으로의 기하학적 변형을 수행하여 중간 이미지를 생성하는 단계 - 상기 중간 이미지는 상기 제2 이미지와 동일한 스타일을 상속함 - 와,

상기 제1 이미지 및 상기 중간 이미지에 기초하여 상기 제1 이미지 도메인으로부터 상기 제2 이미지 도메인으로의 스타일 트랜스퍼(style transfer)를 위한 제2 학습 네트워크를 결정하는 단계를 포함하며,

상기 제1 학습 네트워크를 결정하는 단계는,

상기 제1 GAN의 제1 서브네트워크를 사용하여, 상기 제1 이미지에 대해 상기 제1 이미지 도메인으로부터 상기 제2 이미지 도메인으로의 기하학적 변형을 수행하여, 상기 제1 이미지 내 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 단계와,

상기 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와, 상기 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 단계 - 상기 제1 평균 기하구조는 상기 제1 이미지 도메인의 복수의 이미지 내의 객체들의 기하구조들의 평균이고, 상기 제2 평균 기하구조는 상기 제2 이미지 도메인의 복수의 이미지 내의 객체들의 기하구조들의 평균임 - 와,

상기 제1 차이와 상기 제2 차이 사이의 차가 감소하도록 상기 제1 GAN의 상기 제1 서브네트워크를 업데이트하는 단계 - 업데이트된 상기 제1 서브네트워크는 상기 제1 학습 네트워크로서 결정됨 - 를 포함하는,

방법.

청구항 2

제1항에 있어서,

상기 제1 학습 네트워크를 결정하는 단계는,

상기 제1 이미지 내 객체의 기하구조(geometry)의 제1 랜드마크 포인트 및 상기 제2 이미지 내 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 단계와,

상기 제1 랜드마크 포인트의 제1 PCA(principal component analysis) 표현 및 상기 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 단계와,

상기 제1 PCA 표현 및 제2 PCA 표현에 기초하여 상기 제1 학습 네트워크를 결정하는 단계를 포함하는,

방법.

청구항 3

제1항에 있어서,

상기 제1 GAN의 상기 제1 서브네트워크를 업데이트하는 단계는,

상기 제1 GAN의 제2 서브네트워크를 사용하여, 상기 제2 이미지에 대해 상기 제2 이미지 도메인으로부터 상기 제1 이미지 도메인으로의 기하학적 변형을 수행하여, 상기 제2 이미지 내 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 단계와,

상기 제2 기하구조와 상기 제2 평균 기하구조 사이의 제3 차이와, 상기 제2 변형된 기하구조와 상기 제1 평균 기하구조 사이의 제4 차이를 결정하는 단계와,

상기 제3 차이와 상기 제4 차이 사이의 차 또한 감소하도록 상기 제1 GAN의 상기 제1 서브네트워크 및 상기 제2 서브네트워크를 함께 업데이트하는 단계를 포함하는,

방법.

청구항 4

제3항에 있어서,

상기 제2 이미지 도메인으로부터 상기 제1 이미지 도메인으로의 기하학적 변형을 수행하는 단계는,

상기 업데이트된 제2 서브네트워크를 사용하여 상기 제2 이미지 도메인으로부터 상기 제1 이미지 도메인으로의 기하학적 변형을 수행하는 단계를 포함하는,

방법.

청구항 5

제1항에 있어서,

상기 제2 학습 네트워크는 제2 GAN에 기초하고, 상기 제2 학습 네트워크를 결정하는 단계는,

상기 제2 GAN의 제1 서브네트워크를 사용하여, 상기 제1 이미지에 대해 상기 제1 이미지 도메인으로부터 상기 제2 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제1 스타일의 상기 제1 이미지를 제2 스타일의 제1 예측 이미지로 전이하는(transfer) 단계 - 상기 제2 스타일은 상기 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 와,

상기 제1 예측 이미지와 상기 제1 이미지 사이의 콘텐츠 차이가 감소하도록 상기 제2 GAN의 상기 제1 서브네트워크를 업데이트하는 단계를 포함하되,

업데이트된 상기 제1 서브네트워크는 상기 제2 학습 네트워크로서 결정되는,

방법.

청구항 6

제5항에 있어서,

상기 제2 GAN의 상기 제1 서브네트워크를 업데이트하는 단계는,

상기 제2 GAN의 제2 서브네트워크를 사용하여, 상기 중간 이미지에 대해 상기 제2 이미지 도메인으로부터 상기 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 상기 중간 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 단계 - 상기 제4 스타일은 상기 제1 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 와,

상기 제2 예측 이미지와 상기 중간 이미지 사이의 콘텐츠 차이가 감소하도록 상기 제2 GAN의 상기 제1 서브네트워크 및 상기 제2 서브네트워크를 함께 업데이트하는 단계를 포함하는,

방법.

청구항 7

컴퓨터로 구현되는 방법으로서,

소스 이미지에 대해 기하학적 변형을 수행 - 상기 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 수행하는 것을 포함하고, 상기 제1 학습 네트워크는 제1 GAN(generative adversarial network)에 기초하여 결정됨 - 하여 상기 소스 이미지 내 객체의 제1 기하구조를 제2 기하구조로 변형시키는 단계와,

상기 소스 이미지에 대해 스타일 트랜스퍼를 수행 - 상기 스타일 트랜스퍼를 수행하는 것은 제2 학습 네트워크를 사용하여 수행하는 것을 포함하고, 상기 제2 학습 네트워크는 제2 GAN에 기초하여 결정됨 - 하여 제1 스타일의 상기 소스 이미지를 제2 스타일의 중간 이미지로 전이하는 단계와,

타겟 이미지의 스타일이 상기 중간 이미지의 제2 스타일과 일관성을 유지하면서 상기 타겟 이미지 내 객체가 상기 제2 기하구조를 갖도록 상기 변형된 제2 기하구조에 기초하여 (under the guidance of the deformed second geometry) 상기 중간 이미지에 대해 이미지 워핑을 수행함으로써 상기 소스 이미지의 상기 제2 기하구조 및 상기 중간 이미지의 상기 제2 스타일에 기초하여 상기 타겟 이미지를 생성하는 단계를 포함하는,

방법.

청구항 8

제7항에 있어서,

상기 기하학적 변형을 수행하는 것은,

상기 소스 이미지에서 상기 제1 기하구조를 표현하는 랜드마크 포인트를 결정하는 것과,

상기 랜드마크 포인트의 PCA(principal component analysis) 표현을 생성하는 것과,

상기 PCA 표현에 대해 상기 기하학적 변형을 수행하는 것과,

상기 변형된 PCA 표현에 기초하여 상기 제2 기하구조를 표현하는 변형된 랜드마크 포인트를 결정하는 것을 포함하는,

방법.

청구항 9

제7항에 있어서,

상기 기하학적 변형을 수행하는 것은,

상기 객체의 변형 정도의 표시를 획득하는 것과,

상기 변형 정도에 기초하여 상기 제1 기하구조를 상기 제2 기하구조로 전이하는 것을 포함하는,

방법.

청구항 10

제7항에 있어서,

상기 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은,

상기 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과,

상기 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과,

상기 콘텐츠 관련 특징 표현 및 상기 외형 관련 특징 표현에 기초하여 상기 중간 이미지를 생성하는 것을 포함

하는,
방법.

청구항 11

제10항에 있어서,
상기 외형 관련 특징 표현을 결정하는 것은,
복수의 미리 결정된 외형 관련 특징 표현 중에서 상기 외형 관련 특징 표현을 무작위로 선택하는 것과,
상기 제2 스타일과 관련된 획득된 표시 또는 상기 제2 스타일의 참조 이미지에 기초하여 상기 외형 관련 특징 표현을 결정하는 것을 포함하는,
방법.

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

발명의 설명

기술 분야

배경 기술

[0001] 이미지 변환(image translation)은 소스 이미지를 다른 스타일의 이미지로 변환하는 것을 말한다. 상이한 이미지 유형은 상이한 이미지 도메인(image domain)으로 간주될 수 있다. 예를 들어 사진, 만화, 캐리커처, 유화, 스케치 및 수채화는 상이한 이미지 도메인으로 간주될 수 있다. 상이한 도메인의 이미지로 표현된 스타일 변형 및/또는 기하학적 변형은 일반적으로 상이하다.

[0002] 현재, 교차 도메인(cross-domain) 이미지 변환을 가능하게 하는 많은 이미지 변환 기법이 있다. 그러나 대부분의 이미지 변환 기법은 소스 이미지의 스타일을 다른 원하는 스타일로 바꾸는 데 집중하는 반면에 이미지로 표시된 콘텐츠는 실질적으로 유지된다. 예를 들어, 사진은 스케치 스타일 이미지로 변환될 수 있지만, 이미지 내의 객체의 형상은 거의 동일하게 유지된다. 경우에 따라, 소스 이미지를 다른 기하학적 표현을 사용하여 타겟 이미지로 변환하려는 경우, 예컨대 실제 사람을 보여주는 사진을 과장되게 변형시킨 캐리커처 이미지로 변환하거나 또는 그 반대로 변환하고자 할 경우, 단순히 사진의 스타일을 변환하는 것만으로는 캐리커처의 효과를 표현하기에 충분하지 않을 것이다.

발명의 내용

[0003] 본 명세서에 기술된 발명대상의 구현에 따르면, 교차 도메인 이미지 변환을 위한 해결책이 제공된다. 이 해결책에서, 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크

(learning network)가, 제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 결정되며, 제1 및 제2 이미지 도메인의 이미지들은 상이한 스타일을 가지며 이들 이미지 내 객체들은 서로에 대해 기하학적 변형을 갖는다. 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형이 제2 이미지에서 수행되어 제2 이미지와 동일한 스타일을 상속하는 중간 이미지를 생성하거나, 또는 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형이 제1 이미지에서 수행되어 제1 이미지와 동일한 스타일을 상속하는 중간 이미지를 생성한다. 제1 이미지 도메인으로부터 제2 이미지 도메인으로 스타일을 바꾸기 위한 제2 학습 네트워크가, 제2 이미지로부터 생성된 중간 이미지 및 제1 이미지에 기초하여 또는 제1 이미지로부터 생성된 중간 이미지 및 제2 이미지에 기초하여 결정된다. 이 해결책을 통해, 교차 도메인 이미지 변환을 위한 학습 네트워크의 처리 정확도가 향상될 수 있고 복잡도가 줄어든다.

[0004] 본 요약부는 아래의 상세한 설명에서 추가로 기술되는 개념들의 선택을 단순화된 형식으로 소개하도록 제공된다. 본 요약부는 청구 대상의 중요 특성 또는 기본 특성을 밝히기 위한 것이 아니며, 청구 대상의 범주를 제한하는데 사용하기 위한 것도 아니다.

도면의 간단한 설명

[0005] 도 1은 본 명세서에 기술된 발명대상의 다양한 구현예가 구현될 수 있는 컴퓨팅 환경의 블록도이다.
 도 2a는 본 명세서에 기술된 발명대상의 구현예에 따른 이미지 변환 모듈의 블록도이다.
 도 2b는 본 명세서에 기술된 발명대상의 다른 구현예에 따른 이미지 변환 모듈의 블록도이다.
 도 3은 본 명세서에 기술된 발명대상의 구현예에 따른 제1 학습 네트워크를 트레이닝하기 위한 트레이닝 아키텍처의 블록도이다.
 도 4a는 본 명세서에 기술된 발명대상의 구현예에 따른 제2 학습 네트워크를 트레이닝하기 위한 트레이닝 아키텍처의 블록도이다.
 도 4b는 본 명세서에 기술된 발명대상의 다른 구현예에 따른 제2 학습 네트워크를 트레이닝하기 위한 트레이닝 아키텍처의 블록도이다.
 도 5a는 본 명세서에 기술된 발명대상의 구현예에 따른 교차 도메인 이미지 변환 프로세스의 흐름도이다.
 도 5b는 본 명세서에 기술된 발명대상의 다른 구현예에 따른 교차 도메인 이미지 변환 프로세스의 흐름도이다.
 도 6a는 본 명세서에 기술된 발명대상의 구현예에 따른 학습 네트워크를 트레이닝하는 프로세스의 흐름도이다.
 도 6b는 본 명세서에 기술된 발명대상의 다른 구현예에 따른 학습 네트워크를 트레이닝하는 프로세스의 흐름도이다.
 도면들에 걸쳐, 동일하거나 유사한 참조부호는 동일하거나 또는 유사한 요소들을 지칭한다.

발명을 실시하기 위한 구체적인 내용

[0006] 이하에서는 몇몇 예시적인 구현예를 참조하여 본 명세서에 기술된 발명대상을 설명할 것이다. 이들 구현예는, 발명대상의 범위를 제한하기보다는 당업자가 본 명세서에 기술된 발명대상을 더 잘 이해하고 구현할 수 있도록 하기 위한 것이라는 점을 이해해야 한다.

[0007] 본 명세서에서 사용되는 "포함한다"란 용어와 그 변형어는 "포함하지만 제한되지는 않는다"는 의미의 개방형 용어로 이해해야 한다. "기초하는"이란 용어는 "적어도 부분적으로 기초하는"으로 이해해야 한다. "일 구현예" 및 "일 실시예"란 용어는 "적어도 하나의 구현예"로 이해해야 한다. "다른 실시예"란 용어는 "적어도 하나의 다른 실시예"로 이해해야 한다. "제1", "제2" 등은 상이한 객체 또는 동일한 객체를 지칭할 수도 있다. 명시적이든 또는 암시적이든 다른 정의가 아래에 포함될 수도 있다.

[0008] 전술한 바와 같이, 교차 도메인 이미지 변환은 상이한 기하학적 표현을 갖는 이미지 도메인들 사이의 변환을 포함할 수 있다. 전형적인 예는 사진을 캐리커처 이미지로 또는 그 반대로 변환하는 것이다. 캐리커처는 단순한 또는 과장된 방식으로 객체를 그리는 분야로 정의될 수 있다. 캐리커처는 일반적으로 스케치된 개체의 얼굴, 특히 사람의 얼굴을 스케치하는 데 집중한다. 캐리커처는 객체의 일부를 과장된 형태로 만들어 유머나 풍자 같은 엔터테인먼트 효과를 얻는 것이 특징이다. 따라서, 캐리커처에서 객체의 기하구조(geometry)는 실제 사진과 같은 이미지 속 객체의 기하구조에 비해 변형 효과를 갖는다. 캐리커처는 일반적으로 일상 생활에서 아티스트

들이 그린다. 이들 아티스트는 특정 물체의 독특한 특징을 포착하고 캐리커처를 그릴 때 이러한 특징을 과장할 수 있는 능력이 있다.

[0009] 다른 이미지 도메인의 이미지를 캐리커처로 변환하기 위한 자동 이미지 변환 기법을 개발하려는 시도가 있어왔다. 한 가지 해결책은 사용자가 직접 이미지 변형을 조작할 수 있게 하는 변형(deformation) 시스템을 제공하는 것이다. 이런 종류의 방법은, 사용자에게 전문 지식과 아마도 경험이 풍부한 아티스트의 세부적인 개입을 요구한다. 아티스트가 캐리커처를 그리는 과정에서 적용되는 몇몇 과장 규칙을 결정하는 또 다른 해결책이 있는데, 예를 들면, 얼굴에서 코, 눈 및/또는 입을 강조한 다음, 이들 과장 규칙을 적용하여 소스 이미지를 캐리커처로 바꾸는 것이다. 그러나, 이 해결책은 특정 예술 형식(스케치 또는 일부 만화 변형용)에 국한되며, 변형된 캐리커처의 과장은 단순히 미리 정해진 규칙을 따를 뿐, 상이한 객체들의 특징을 구별하여 보여주지는 못한다.

[0010] 최근에는 이미지 변환에 딥 러닝 기법(머신 러닝이라고도 함)이 성공적으로 사용되고 있다. 딥 러닝 기반 해결책은 트레이닝 이미지를 활용하여 학습 모델을 트레이닝하여 학습 모델이 한 이미지 도메인의 이미지를 다른 이미지 도메인의 이미지로 변환될 수 있도록 한다. 그러나, 캐리커처 도메인을 포함하는 이미지 변환에 딥 러닝 기법을 적용하는 것은 매우 어렵다.

[0011] 첫째, 대부분의 이미지는 캐리커처 버전 쌍이 없다. 예를 들어, 대응하는 캐리커처 이미지는 일반적으로 많은 사진에 사용할 수 없다. 아티스트가 각 사진에 대해 대응하는 캐리커처 버전을 만드는 데에는 시간과 비용이 많이 든다. 따라서, 지도 학습(supervised learning) 방식을 사용하여 학습 모델을 트레이닝하는 것은 어렵다.

[0012] 한편, 캐리커처 이미지를 생성하는 열쇠는 기하학적 과장과 스타일 트랜스퍼(style transfer)에 있다. 그러나 많은 학습 모델은 일반적으로 이미지에 표시된 콘텐츠가 크게 변경되지 않은 상태로 스타일 트랜스퍼에 중점을 둔다. 이는 이미지에 있는 객체의 형상이 크게 변경되지 않음을 의미한다. 원칙적으로 기하학적 변형(geometric deformation)과 스타일 트랜스퍼 모두를 동시에 학습하도록 학습 모델을 트레이닝할 수 있는 몇몇 비지도 교차 도메인 이미지 변환(unsupervised cross-domain image translation) 접근법이 제안되었다. 그러나, 캐리커처 도메인과 다른 이미지 도메인 사이의 형상 및 스타일의 큰 차이는 이들 학습 모델에 큰 어려움이 있었고, 따라서 학습 모델은 대부분 결과가 만족스럽지 못했다.

[0013] 예시적인 환경

[0014] 본 명세서에 기술된 발명대상의 기본 원리 및 몇 가지 예시적인 구현을 도면을 참조하여 아래에서 설명한다. 도 1은 본 명세서에 기술된 발명대상의 다양한 구현예가 구현될 수 있는 컴퓨팅 환경(100)의 블록도이다. 도 1에 도시된 컴퓨팅 장치(100)는 단지 예시를 위한 것이며, 어떤 방식으로든 본 명세서에 기술된 발명대상의 구현예의 기능 및 범위에 대해 어떠한 제한을 하고자 하는 것은 아니라는 것을 이해할 수 있을 것이다. 도 1에 도시된 바와 같이, 컴퓨팅 장치(100)는 범용 컴퓨팅 장치(100)를 포함한다. 컴퓨팅 장치(100)의 구성 요소는 하나 이상의 프로세서 또는 처리 장치(110), 메모리(120), 저장 장치(130), 하나 이상의 통신 유닛(140), 하나 이상의 입력 장치(150) 및 하나 이상의 출력 장치(160)를 포함할 수 있지만, 이에 제한되지는 않는다.

[0015] 일부 구현에서, 컴퓨팅 장치(100)는 컴퓨팅 능력을 갖는 임의의 사용자 단말기 또는 서버 단말로서 구현될 수 있다. 서버 단말은 서비스 제공자가 제공하는 서버, 대규모 컴퓨팅 장치 등일 수 있다. 사용자 단말기는, 예를 들어 이동 전화, 스테이션, 유닛, 장치, 멀티미디어 컴퓨터, 멀티미디어 태블릿, 인터넷 노드, 커뮤니케이터, 데스크톱 컴퓨터, 랩톱 컴퓨터, 노트북 컴퓨터, 넷북 컴퓨터, 태블릿 컴퓨터, PCS(personal communication system) 장치, 개인용 내비게이션 장치, PDA(Personal Digital Assistant), 오디오/비디오 플레이어, 디지털 카메라/비디오 카메라, 포지셔닝 장치, 텔레비전 수신기, 라디오 방송 수신기, 전자책 장치, 게임 장치 또는 이들 장치의 액세서리 및 주변 장치를 포함한 임의의 조합 또는 이들의 조합을 포함한 임의의 유형의 이동 단말기, 고정 단말기 또는 휴대용 단말기일 수 있다. 컴퓨팅 장치(100)는 사용자에게 대한 임의의 유형의 인터페이스(예컨대, "웨어러블" 회로 등)를 지원할 수 있다고 고려될 수 있다.

[0016] 처리 유닛(110)은 물리적 또는 가상 프로세서일 수 있고 메모리(120)에 저장된 프로그램을 기반으로 다양한 프로세스를 구현할 수 있다. 멀티프로세서 시스템에서, 멀티 프로세싱 유닛은 컴퓨팅 장치(100)의 병렬 처리 기능을 개선하기 위해 컴퓨터 실행 가능 명령어를 병렬로 실행한다. 처리 유닛(110)은 또한 중앙 처리 유닛(CPU), 마이크로프로세서, 컨트롤러 또는 마이크로 컨트롤러로 지칭될 수 있다.

[0017] 컴퓨팅 장치(100)는 일반적으로 다양한 컴퓨터 저장 매체를 포함한다. 이런 매체는 휘발성 및 비휘발성 매체, 또는 분리 가능 및 분리 불가능 매체를 포함하지만 이에 제한되지 않는 컴퓨팅 장치(100)에 의해 액세스 가능한

임의의 매체일 수 있다. 메모리(120)는 휘발성 메모리(예를 들어, 레지스터, 캐시, RAM(Random Access Memory)), 비휘발성 메모리(ROM(Read-Only Memory), EEPROM(Electrically Erasable Programmable Read-Only Memory) 또는 플래시 메모리) 또는 이들의 조합일 수 있다. 저장 장치(130)는 임의의 분리 가능 또는 분리 불가능한 매체일 수 있으며, 정보 및/또는 데이터를 저장하는 데 사용될 수 있고 컴퓨팅 장치(100)에서 액세스될 수 있는 메모리, 플래시 메모리 드라이브, 자기 디스크 또는 기타 다른 매체와 같은 기계 판독 가능 매체를 포함할 수 있다.

[0018] 컴퓨팅 장치(100)는 추가의 분리 가능/분리 불가능, 휘발성/비휘발성 메모리 매체를 더 포함할 수 있다. 도 1에 도시되어 있지 않지만, 분리 가능한 비휘발성 자기 디스크로부터 읽고 쓰기 위한 자기 디스크 드라이브와 분리 가능한 비휘발성 광학 디스크로부터 읽기 및/또는 쓰기를 위한 광 디스크 드라이브를 제공하는 것이 가능하다. 이런 경우, 각 드라이브는 하나 이상의 데이터 매체 인터페이스를 통해 버스(도시되어 있지 않음)에 연결될 수 있다.

[0019] 통신 유닛(140)은 통신 매체를 통해 다른 컴퓨팅 장치와 통신한다. 또한, 컴퓨팅 장치(100)의 구성요소들의 기능은 단일 컴퓨팅 클러스터 또는 통신 연결을 통해 통신할 수 있는 다수의 컴퓨팅 머신에 의해 구현될 수 있다. 따라서, 컴퓨팅 장치(100)는 하나 이상의 다른 서버, 네트워크로 연결된 개인용 컴퓨터들(PCs) 또는 추가 일반 네트워크 노드와의 논리적 연결을 사용하여 네트워크 환경에서 동작할 수 있다.

[0020] 입력 장치(150)는 마우스, 키보드, 트랙볼, 음성 입력 장치 등과 같은 다양한 입력 장치 중 하나 이상일 수 있다. 출력 장치(160)는 디스플레이, 라우드 스피커, 프린터 등과 같은 다양한 출력 장치 중 하나 이상일 수 있다. 컴퓨팅 장치(100)는, 통신 유닛(140)에 의해, 저장 장치 및 디스플레이 장치와 같은 하나 이상의 외부 장치(도시되어 있지 않음), 사용자가 컴퓨팅 장치(100)와 상호 작용할 수 있게 하는 하나 이상의 장치, 또는 필요한 경우 컴퓨팅 장치(100)가 하나 이상의 다른 컴퓨팅 장치와 통신할 수 있게 하는 임의의 장치(예컨대, 네트워크 카드, 모뎀 등)와 추가로 통신할 수 있다. 이런 통신은 입력/출력(I/O) 인터페이스(도시되어 있지 않음)를 통해 수행될 수 있다.

[0021] 일부 구현에서, 단일 장치에 통합되는 대안으로서, 컴퓨팅 장치(100)의 일부 또는 모든 구성요소는 또한 클라우드 컴퓨팅 아키텍처로 구성될 수 있다. 클라우드 컴퓨팅 아키텍처에서, 구성요소들은 원격으로 제공될 수 있고 본 명세서에 기술된 발명대상에서 기술된 기능을 구현하기 위해 함께 작동할 수 있다. 일부 구현에서 클라우드 컴퓨팅은 컴퓨팅, 소프트웨어, 데이터 액세스 및 저장 서비스를 제공하며, 이들은 최종 사용자가 이들 서비스를 제공하는 시스템 또는 하드웨어의 물리적 위치 또는 구성을 알 것을 요구하지 않는다. 다양한 구현에서 클라우드 컴퓨팅은 적절한 프로토콜을 사용하여 광역 네트워크(예컨대, 인터넷)를 통해 서비스를 제공한다. 예를 들어, 클라우드 컴퓨팅 제공자는 광역 통신망을 통해 애플리케이션을 제공하며, 이들은 웹 브라우저 또는 기타 컴퓨팅 구성요소를 통해 액세스될 수 있다. 클라우드 컴퓨팅 아키텍처의 소프트웨어 또는 구성요소와 대응하는 데이터는 원격 위치의 서버에 저장될 수 있다. 클라우드 컴퓨팅 환경의 컴퓨팅 리소스는 원격 데이터 센터의 위치에서 병합되거나 분산될 수 있다. 클라우드 컴퓨팅 인프라는, 사용자를 위한 단일 액세스 포인트로 작용하지만, 공유 데이터 센터를 통해 서비스를 제공할 수 있다. 따라서, 클라우드 컴퓨팅 아키텍처는 원격 위치의 서비스 제공자로부터 본 명세서에 기술된 구성요소 및 기능을 제공하는 데 사용될 수 있다. 또는, 이들은 기존 서버에서 제공되거나 클라이언트 장치에 직접 설치되거나 또는 다른 방법으로 설치될 수 있다.

[0022] 컴퓨팅 장치(100)는 본 명세서에 기술된 발명대상의 다양한 구현에서 교차 도메인 이미지 변환을 구현하는 데 사용될 수 있다. 따라서, 이하에서는, 컴퓨팅 장치를 "이미지 처리 장치(100)"라고도 한다. 메모리(120)는 하나 이상의 프로그램 명령어를 갖는 하나 이상의 이미지 변환 모듈(122)을 포함할 수 있다. 이들 모듈은 본 명세서에 기술된 다양한 구현들의 기능을 수행하기 위해 처리 유닛(110)에 의해 액세스 가능하고 실행 가능하다.

[0023] 이미지 변환을 수행할 때, 이미지 처리 장치(100)는 입력 장치(150)를 통해 소스 이미지(102)를 수신할 수 있다. 입력 장치(150)에 의해 획득된 소스 이미지(102)는 이미지 변환 모듈(122)에 입력으로 제공된다. 이미지 변환 모듈(122)은 소스 이미지(102)에 대해 교차 도메인 이미지 변환을 수행한다. 본 개시에서, 이미지의 스타일 및 기하구조 측면 모두에서 변환을 수행하는 것이 바람직하다. 이러한 이미지 변환은 임의의 다른 이미지 도메인으로부터 캐리커처 도메인으로, 또는 캐리커처 도메인에서 다른 이미지 도메인으로 변환하는 데 적합하다. 도 1의 예에서, 이미지 변환 모듈(122)은 사진 도메인의 소스 이미지(102)를 캐리커처 도메인의 타겟 이미지(104)로 변환한다. 출력 장치(160)는 타겟 이미지(104)를 시청자에게 제시하거나 또는 타겟 이미지(104)를 다른 장치로 전송할 수 있다.

[0024] 도 1에 도시된 소스 이미지(102) 및 타겟 이미지(104)는 단지 예시를 위한 것임을 이해해야 한다. 다른 예들에

서, 임의의 소스 이미지가 처리될 수 있고 캐리커처 형태로 다른 타겟 이미지를 생성할 수 있다. 일부 구현에서, 이미지 변환 모듈(122)은 또한 형상 과장 효과 없이 캐리커처 형태의 소스 이미지를 다른 타겟 이미지로 변환하도록 설계될 수도 있다.

[0025] 작동 원리

[0026] 본 명세서에 기술된 발명대상의 구현에 따르면, 교차 도메인 이미지 변환을 위한 해결책이 제공된다. 이 해결책에서, 교차 도메인 이미지 변환은 기하학적 변형 및 스타일 트랜스퍼라는 두 가지 작업으로 명시적으로 나누어진다.

[0027] 제1 구현예에서, 기하학적 변형 및 스타일 트랜스퍼가 변환될 소스 이미지에 대해 각각 수행되고, 두 작업의 결과에 기초하여 타겟 이미지가 획득된다. 구체적으로, 기하학적 변형은 소스 이미지에서 객체의 제1 기하구조가 제2 기하구조로 변형되도록 한다. 스타일 트랜스퍼는 제1 스타일의 소스 이미지를 제2 스타일의 중간 이미지로 전이하도록(transfer) 한다. 그 다음에, 제2 기하구조에 기초하여 중간 이미지가 타겟 이미지로 전이된다. 타겟 이미지는 제2 스타일을 가지며, 그 안에 있는 객체는 변형 후 제2 기하구조를 갖는다. 이 해결책에 따르면, 교차 도메인 이미지 변환의 복잡한 문제가 기하학적 변형(geometric deformation)과 스타일 트랜스퍼라는 두 가지 병렬 작업으로 해결된다. 이것은 처리 정확도를 향상시키고 복잡성을 줄일 수 있다. 따라서, 획득된 타겟 이미지는 기하학적 변형 및 스타일 측면에서 원하는 이미지 도메인의 특성에 매우 부합한다.

[0028] 제1 구현예에서 기하학적 변형 및 스타일 트랜스퍼의 두 가지 작업은 학습 네트워크에 의해 구현될 수 있다. 학습 네트워크에 대한 결정 단계에서, 기하학적 변형을 위한 제1 학습 네트워크를 결정하기 위해 제1 이미지 도메인 및 제2 이미지 도메인의 이미지가 사용된다. 스타일 트랜스퍼를 위한 제2 학습 네트워크를 결정하기 위해, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형이 제2 도메인의 이미지에서 수행되어, 제1 이미지 도메인의 기하구조를 가지며 제2 이미지 도메인의 스타일을 상속하는 중간 이미지를 생성한다. 중간 이미지가 제1 이미지 도메인의 이미지와 함께 사용되어 제2 학습 네트워크를 결정한다. 학습 네트워크를 결정하는 프로세스 동안, 중간 이미지를 도입하여 스타일 트랜스퍼를 위한 학습 네트워크를 결정함으로써, 두 이미지 도메인의 이미지들 사이의 기하학적 차이를 상쇄할 수 있어, 제2 학습 네트워크가 스타일 트랜스퍼 학습에 집중할 수 있다.

[0029] 제2 구현예에서, 기하학적 변형과 스타일 트랜스퍼의 두 작업을 수행할 때, 먼저 변형될 소스 이미지에 대해 기하학적 변형이 수행되고, 그 다음에 기하학적 변형을 기반으로 스타일 트랜스퍼가 수행되어 타겟 이미지를 획득한다. 구체적으로, 기하학적 변형은 소스 이미지 내 객체의 제1 기하구조가 제2 기하구조로 변형되게 한다. 소스 이미지는 제2 기하구조를 기반으로 변환되어 소스 이미지와 동일한 제1 스타일을 상속하는 중간 이미지를 생성한다. 그 다음에, 중간 이미지에 대해 스타일 트랜스퍼가 수행되어 제1 스타일을 갖는 중간 이미지를 제2 스타일을 갖는 타겟 이미지로 변환한다. 이 해결책에 따르면, 복잡한 교차 도메인 이미지 변환이 순차적으로 수행되는 기하학적 변형과 스타일 트랜스퍼라는 두 작업으로 구현된다. 이는 처리 정확도를 향상시키고 복잡성을 줄일 수 있다. 따라서, 획득된 타겟 이미지는 기하학적 변형 및 스타일 측면에서 원하는 이미지 도메인 특성에 매우 부합한다.

[0030] 순차적으로 수행되는 기하학적 변형 및 스타일 트랜스퍼 작업은 또한 학습 네트워크에 의해 구현될 수 있다. 학습 네트워크에 대한 결정 단계에서, 기하학적 변형에 대한 제1 학습 네트워크가 제1 구현예에서와 유사한 방식으로 결정된다. 스타일 트랜스퍼를 위한 제2 학습 네트워크를 결정하기 위해, 제1 이미지 도메인의 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형이 수행되어 제1 이미지 도메인의 스타일을 상속하는 중간 이미지를 생성한다. 중간 이미지가 제2 이미지 도메인의 이미지와 함께 사용되어 제2 학습 네트워크를 결정한다. 학습 네트워크를 결정하는 프로세스 동안, 중간 이미지를 도입하여 스타일 트랜스퍼를 위한 학습 네트워크를 결정함으로써, 두 이미지 도메인의 이미지들 사이의 기하학적 차이를 상쇄할 수 있어, 제2 학습 네트워크가 스타일 트랜스퍼 학습에 집중할 수 있다.

[0031] 이상은 교차 도메인 이미지 변환이 기하학적 변형 및 스타일 트랜스퍼의 두 가지 특정 작업으로 분할되는 상이한 구현예들의 기본 작동 원리를 설명한다. 본 명세서에 기술된 발명대상의 이미지 변환 해결책은 상이한 스타일을 갖고 서로에 대해 기하학적 변형을 갖는 객체를 포함하는 두 이미지 도메인 사이의 이미지 변환을 수행하도록 구성될 수 있으며, 이런 이미지 변환은, 캐리커처 도메인의 객체가 일반적으로 과장된 형상을 사용하여 표현되기 때문에, 다른 이미지 도메인으로부터 캐리커처 도메인으로 또는 캐리커처 도메인으로부터 다른 이미지 도메인으로의 변환에 특히 적합하다. 아래의 특정 구현예에서, 설명을 쉽게 하기 위해, 사진 도메인으로부터 캐리커처 도메인으로의 이미지 변환이 예시의 예로 사용된다. 그러나, 본 명세서에 기술된 발명대상의 해결책

은 또한 다른 이미지 도메인으로부터 캐리커처 도메인으로 또는 캐리커처 도메인으로부터 다른 이미지 도메인의 이미지 변환에 적합하다는 것을 이해해야 한다.

[0032] 본 명세서에서 사용된 바와 같이, "캐리커처"는 과장된 기하구조를 사용하여 표현된 객체를 갖는 이미지로서 광범위하게 이해되어야 한다. 이런 과장은 실제 객체의 기하구조에 대한 변형(예컨대, 강조/확대, 축소, 변위, 회전 및 뒤틀림)이다. 본 명세서에 사용된 바와 같이, 이미지의 스타일은 색상 배열, 명암 전환 및 질감 특징을 포함한 이미지의 색상, 질감 및 밝기와 같은 외형 특징을 의미한다. 이미지의 스타일은 이미지 내 객체의 기하구조를 포함하지 않는다.

[0033] 본 명세서에서 사용되는 바와 같이, 학습 네트워크는 또한 "학습 모델" 또는 "신경망"으로 지칭될 수 있다. 용어 "학습 모델", "학습 네트워크", "신경망", "모델" 및 "네트워크"는 이하에서 같은 의미로 사용된다. 일반적으로 학습 네트워크는 입력을 원하는 출력에 매핑할 수 있는 것으로 이해할 수 있다. 이런 매핑 능력은 트레이닝 데이터로 학습 네트워크를 트레이닝함으로써 유도될 수 있다. 학습 네트워크의 트레이닝 프로세스는 원하는 매핑 출력을 얻기 위해 입력을 조정하기 위한 네트워크 파라미터를 결정하는 것으로 이해할 수 있다.

[0034] 본 명세서에 기술된 발명대상의 특정 구현예를 설명하기 전에, 기하학적 변형과 관련된 교차 도메인 이미지 변환에서의 예상 목표를 먼저 소개한다. 캐리커처 생성의 적절한 결과는 캐리커처로 나타난 객체가 아티스트가 만든 것과 유사하다는 것이다. 캐리커처 이미지의 객체는 과장된 형상으로 보이지만 이 형상의 과장이 왜곡은 아니다. 왜곡은 객체의 진실을 완전히 부인하지만 캐리커처의 과장된 형상은 객체의 상대적인 기하학적 위치(예컨대, 얼굴 특징의 상대적인 기하학적 위치)를 유지하고 주어진 객체의 특징을 다른 객체와 구별하는 것만 강조한다. 따라서, 이미지 변환 후 얻은 캐리커처는 소스 이미지에서 객체의 특징을 유지하고 캐리커처의 스타일을 시각적으로 표현할 수 있을 것으로 기대된다.

[0035] 또한, 이미지 변환의 또 다른 예상 목표는 캐리커처의 생성을 다양화하고 제어 가능하게 만드는 것이다. 즉, 주어진 소스 이미지에 대해 다양한 유형의 캐리커처가 생성될 수 있으며 최종 출력 결과는 사용자에게 의해 또는 다른 방식으로 제어될 수 있다. 또한, 전술한 바와 같이, 비용을 고려했을 때, 사진과 캐리커처 쌍을 획득하여 지도 방식으로 학습 네트워크를 트레이닝하는 것이 어렵다. 따라서, 학습 네트워크를 트레이닝하기 위해, 이미지 변환을 수행하는 학습 네트워크가 쌍을 이루지 않은 사진과 캐리커처를 이용하여 비지도 방식으로 트레이닝될 수 있을 것으로 기대된다.

[0036] 기하구조 변환 및 스타일 트랜스퍼의 병렬 아키텍처

[0037] 도 2a는 기하학적 변형 및 스타일 트랜스퍼가 병렬로 수행되는 제1 구현예에 따른 이미지 변환 모듈의 아키텍처를 도시한 것이다. 이 아키텍처는 학습 네트워크에 기초하여 구현되고, 도 1의 컴퓨터 장치(100)의 이미지 변환 모듈(122)에서 구현될 수 있다. 도시된 바와 같이, 이미지 변환 모듈(122)은 기하학적 변형을 위한 제1 학습 네트워크(210) 및 스타일 트랜스퍼를 위한 제2 학습 네트워크(220)를 포함한다. 제1 학습 네트워크(210) 및 제2 학습 네트워크(220)는 제각기의 작업을 수행하기 위해 독립적으로 동작할 수 있고 독립적으로 트레이닝된다.

[0038] X 및 Y 는 각각 제1 이미지 도메인(예컨대, 사진 도메인) 및 제2 이미지 도메인(예컨대, 캐리커처 도메인)을 나타낸다고 가정한다. 교차 도메인 이미지 변환의 문제는 도메인 X 의 입력을 도메인 Y 의 샘플로, 즉, $y = \Phi(x), y \in Y$ 로 변환할 수 있는 매핑 $\Phi : X \rightarrow Y$ 를 얻는 것으로 설명될 수 있다. 이 문제는, 사진 도메인과 캐리커처 도메인이 기하구조와 스타일이 다르기 때문에 생기는 교차 도메인 이미지 변환의 전형적인 문제이다. 본 명세서에 기술된 발명대상의 구현예에 따르면, 이 문제는 두 가지 작업, 즉 기하학적 변형과 스타일 트랜스퍼로 나누어진다. 따라서, 전체 맵이 기하학적 변형을 위한 맵(Φ_{geo})과 스타일 트랜스퍼를 위한 맵(Φ_{app})의 두 개의 맵으로 명시적으로 분리된다. 도 2a의 예에서, 두 맵 Φ_{geo} 및 Φ_{app} 은 제각기 제1 학습 네트워크(210) 및 제2 학습 네트워크(220)에 의해 학습된다. 즉, 제1 학습 네트워크(210) 및 제2 학습 네트워크(220)의 네트워크 파라미터가 결정되었다고 가정하면, 이들 네트워크는 입력 소스 이미지(102)를 처리하여 타겟 이미지(104)를 획득하는 데 사용될 수 있다.

[0039] 제1 학습 네트워크(210)는 입력 소스 이미지(102)에 대해 기하학적 변형을 수행하여 소스 이미지(102) 내의 객체의 제1 기하구조를 캐리커처 도메인의 제2 기하구조로 변형시키는 능력을 획득하도록 트레이닝된다. 소스 이미지(102)는, 예를 들어, 사용자에게 의해 입력될 수 있으며, 사용자가 변환하고자 하는 이미지이다. 실제 사진

에 비해, 캐리커처 도메인에서의 동일 객체의 기하구조는 과장된 효과를 갖도록 변형될 것이다. 제1 학습 네트워크(210)는 소스 이미지(102)의 어떠한 스타일 트랜스퍼도 고려하지 않고 객체의 기하학적 변형에만 초점을 맞춘다.

[0040] 이미지 변환 모듈(122)은 또한, 제1 학습 네트워크(210)에 의한 처리를 위해 소스 이미지(102)(즉, 제1 기하구조) 내 객체의 기하구조를 검출하기 위한 기하구조 검출기(230)를 포함한다. 소스 이미지(102)는 사람, 사람의 얼굴, 동물, 식물 및/또는 기타 객체와 같은 하나 이상의 객체를 포함할 수 있다. 도 2a에 도시된 소스 이미지(102)에서, 관련 객체는 사람의 얼굴이다. 그러나, 입력 소스 이미지는 특정 기하구조를 갖는 다른 객체를 포함할 수도 있으며, 본 명세서에 기술된 발명대상의 해결책은 이러한 소스 이미지의 객체에 기하학적 변형을 수행할 수도 있다. 기하구조 검출기(230)에 의해 결정된 기하구조는 소스 이미지(102) 내의 객체의 라인(예를 들어, 소스 이미지(102)의 2차원 좌표계 내의 좌표 포인트인, 라인의 각 픽셀을 나타내는 좌표 포인트)에 의해 표현될 수 있다.

[0041] 기하구조 표현의 크기를 줄이기 위해, 소스 이미지(102)에서 제1 기하구조를 나타내는 라인상의 포인트를 샘플링하고 기하구조를 나타내는 샘플링된 포인트의 좌표를 사용하는 것이 가능하다. 이런 샘플링된 포인트는, 얼굴 랜드마크 포인트와 같은 제1 기하구조(202)(도 2a의 기하구조(202)에서 기하구조를 묘사하는 점으로 도시됨)를 나타내는 랜드마크 포인트(키 포인트라고도 함)일 수 있다. 이런 랜드마크 포인트의 수 및 분포는 객체의 각 부분의 형상 및 상대적 위치를 유지하도록 구성될 수 있다. 랜드마크 포인트의 수는, 예컨대 16, 32 또는 63개일 수 있다.

[0042] 랜드마크 포인트는 2차원 좌표로 표현될 수도 있지만, 데이터 크기는 여전히 크다. 이런 입력을 제1 학습 네트워크(210)가 직접 처리하면, 네트워크 구조 복잡성이 높아지고 계산 효율이 감소한다. 크기를 더 줄이기 위해, 제1 기하구조(202)를 나타내는 랜드마크 포인트에 대해 주성분 분석(principal component analysis: PCA)을 수행하여 (예를 들어, 도 2a에 도시되지 않은 다른 모듈에 의해) 랜드마크 포인트의 PCA 표현을 결정할 수 있다. 제1 학습 네트워크(210)는 소스 이미지(102) 내 객체의 제1 기하구조를 나타내는 PCA 표현을 입력으로서 수신하고, PCA 표현에 대해 기하학적 변형을 수행하여 변형된 PCA 표현을 획득한다. 변형된 PCA 표현은(예를 들어, 도 2a에 도시되지 않은 다른 모듈에 의해) 제2 기하구조(204)를 갖는 변형된 랜드마크 포인트를 결정하는데 사용된다.

[0043] PCA는 입력의 주요 특징을 표현하기 위해 주어진 입력에 대해 더 작은 크기의 특징 벡터를 결정하는 것이다. 이런 특징 벡터는 주성분이라고도 한다. 예를 들어, 소스 이미지(102) 내 객체의 63개의 랜드마크 포인트에 대해, 주성분 분석 후에 처음 32개의 주성분이 선택될 수 있다. 이런 주성분은 입력 랜드마크 포인트의 대부분(약 99.03%)을 복구하는 데 사용할 수 있다. 따라서, 객체의 63개의 랜드마크 포인트는 32개의 PCA 계수로 구성된 1차원 벡터로 표현될 수 있다.

[0044] 크기를 줄이는 것 외에도, PCA 표현은 또한 얼굴 구조를 제한하는 것과 같이 제1 학습 네트워크(210)의 처리 동안 객체의 기하구조를 제한하는 데 도움을 주며, 이는 특히 상부 학습 네트워크(210)의 학습(트레이닝)을 용이하게 할 수 있다. 랜드마크 포인트의 PCA 표현은 객체의 기하구조의 임베디드 공간(embedded space)을 재구성하는 데 사용될 수 있으며, 여기서 각각의 주성분은 변형들(예를 들어, 포즈, 형태 또는 크기)의 방향을 나타낸다. 이 임베디드 공간의 임의의 샘플은 객체의 기본 구조를 유지할 수 있다. 2차원 좌표 포인트 위치의 자유도가 특히 트레이닝 동안 크기 때문에, 이런 구조적 제약은 2차원 좌표로 표현되는 랜드마크 점에 의해 거의 보존되지 않고, 이는 객체의 기하구조를 유지하기 어렵게 만든다.

[0045] 일부 구현들에서, 제1 학습 네트워크(210)는 또한 소스 이미지(102) 내의 객체의 제1 기하구조를 제2 기하구조로 변형시키기 위해, 사용자에게 의해 표시된 변형 정도에 기초하여 기하학적 변형을 수행할 수 있다. 변형 정도는 사용자가 지시할 수 있다. 예를 들어, 사용자 조정 가능한 파라미터를 통해 변형 정도를 표시하도록 설정될 수 있다. 제1 학습 네트워크(210)는 변형 정도에 기초하여 제1 기하구조에 대한 제2 기하구조의 변형을 결정할 수 있다. 예를 들어, 제1 학습 네트워크(210)가 제1 기하구조의 일부를 확대 또는 축소할 경우, 변형 정도가 해당 부분이 확대 또는 축소되는 정도를 제어할 수 있다.

[0046] 제2 학습 네트워크(220)는 입력 소스 이미지(102)에 대해 스타일 트랜스퍼를 수행하는 능력을 학습하고, 사진 도메인에서 제1 스타일을 갖는 소스 이미지(102)를 캐리커처 도메인에서 제2 스타일을 갖는 중간 이미지(212)로 변환하도록 트레이닝된다. 상이한 이미지 도메인들의 스타일 유형은 다를 수 있다. 예를 들어, 캐리커처 스타일은 실제 사진과 비교하여 색상과 질감이 손으로 그린 붓놀림을 갖는다. 제2 학습 네트워크(220)는 소스 이미지(102) 내의 객체(들)의 기하구조를 유지하면서 스타일 트랜스퍼만을 수행한다. 따라서, 중간 이미지(212)는

스타일에서 캐리커처와 유사하지만, 기하학적으로는 소스 이미지(102)와 유사하다.

- [0047] 스타일 트랜스퍼를 수행하는 동안, 제2 학습 네트워크(220)는 소스 이미지(102)의 콘텐츠 관련 특징 표현을 추출하고, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하며, 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 중간 이미지(212)를 결정한다. 따라서, 결정된 중간 이미지(212)는 소스 이미지(102)의 콘텐츠를 제2 스타일과 동시에 결합할 것이다.
- [0048] 일부 구현에서, 제2 학습 네트워크(220)는 주어진 소스 이미지를 캐리커처 도메인의 복수의 상이한 스타일로 전이하도록 설계될 수 있다. 이 능력은, 후술하는 바와 같이, 제2 학습 네트워크(220)의 트레이닝 프로세스를 통해 학습될 수 있다. 소스 이미지(102)의 스타일이 전이될 경우, 제2 스타일과 연관되고 제2 학습 네트워크(220)에 의해 결정된 외형 관련 특징 표현이, 미리 정해진 복수의 외형 관련 특징 표현으로부터 무작위로 선택될 수 있다. 미리 결정된 복수의 외형 관련 특징 표현은 제2 학습 네트워크(220)를 트레이닝하기 위한 캐리커처 도메인의 복수의 이미지로부터 결정되며, 이는 제2 학습 네트워크의 트레이닝 동안 상세히 설명될 것이다. 제2 스타일과 연관된 선택된 외형 관련 특징 표현은 캐리커처 도메인의 이미지로부터 추출된 외형 관련 특징 표현일 수도 있고, 복수의 이미지로부터 추출된 복수의 외형 관련 특징 표현의 조합일 수도 있다.
- [0049] 또는, 제2 스타일은 또한 사용자에게 의해 제어될 수도 있다. 구체적으로, 제2 스타일은 사용자에게 의해 표시될 수도 있고(예를 들어, 학습 네트워크(220)를 트레이닝하는 데 사용되는 이미지의 일부 스타일로부터 사용자에게 의해 선택될 수도 있고), 사용자가 참조 이미지에 제2 스타일을 제공할 수도 있다. 제2 학습 네트워크(220)는 표시된 제2 스타일과 동일 또는 유사한 이미지로부터 추출된 외형 관련 특징 표현을 선택하거나, 또는 참조 이미지로부터 외형 관련 특징 표현을 추출하여 스타일 트랜스퍼를 수행할 수 있다.
- [0050] 이미지 변환 모듈(122)은 제2 기하 구조(204)(랜드마크 포인트로 표현됨) 및 중간 이미지(212)에 기초하여 타겟 이미지(104)를 생성하도록 구성된 병합 모듈(240)을 더 포함한다. 병합 모듈(240)은 변형된 제2 기하구조(204)에 기초하여 (under the guidance of the deformed second geometry) 중간 이미지(212)에 대해 이미지 워핑을 수행하여, 획득된 타겟 이미지(104)의 객체가 타겟 이미지(104)의 스타일은 중간 이미지(212)와 일관성을 유지하면서, 제2 기하구조(202)와 동일하거나 유사한 기하구조를 갖도록 한다. 따라서, 포토 도메인의 소스 이미지(102)로부터 캐리커처 도메인의 타겟 이미지(104)로의 이미지 변환이 달성된다.
- [0051] 제1 학습 네트워크(210) 및 제2 학습 네트워크(220)는 이미지 처리에 적합한 딥 러닝 네트워크 구조로 선택될 수 있다. 일부 구현예에서, 제1 학습 네트워크(210) 및 제2 학습 네트워크(220)는 생성적 적대 신경망(generative adversarial network: GAN)에 기초할 수 있다. 이런 네트워크는 특히 서로 다른 이미지 도메인에서 쌍을 이루지 않은 이미지를 기반으로 하는 비지도 학습에 적합하다. GAN은 일련의 변형을 가질 수 있다. 제1 학습 네트워크(210)의 학습 목표는, 각각의 입력 소스 이미지에 대해 최적의 기하학적 변형을 수행하여 캐리커처 도메인의 소스 이미지에 대응하는 변형 결과를 얻을 수 있도록 하는 것으로, 이는 일대일 이미지 변환으로 간주될 수 있다. 이런 일대일 이미지 변환을 구현하도록 설계된 GAN은 제1 학습 네트워크(210)를 결정하는 데 사용될 수 있다. 제2 학습 네트워크(210)의 학습 목표는 각 입력 이미지에 대해 캐리커처 도메인에서 서로 다른 시각적 스타일을 갖는 중간 이미지를 생성할 수 있다는 것으로, 이는 다대다(many-to-many) 이미지 변환으로 간주될 수 있다. 이런 다대다 이미지 변환을 달성하도록 설계된 GAN은 제2 학습 네트워크(220)를 결정하는 데 사용될 수 있다.
- [0052] 여기서 사진 도메인으로부터 캐리커처 도메인으로의 이미지 변환이 예로서 설명되어 있지만, 본 명세서에 기술된 발명대상은 대응하는 이미지 도메인의 이미지가 학습 네트워크를 트레이닝하는 데 사용되는 한, 다른 이미지 도메인으로부터 과장된 형상 변형을 갖는 이미지 도메인으로의 이미지 변환에 유사하게 적용될 수 있음을 이해해야 한다. (캐리커처 도메인과 같은) 형상 과장 변형을 갖는 이미지 도메인으로부터 다른 이미지 도메인으로 변환이 수행되는 경우, 도 2a의 처리 흐름은 반대로 구현될 수 있다. 또한, "이미지"란 용어가 본 명세서에서 사용되지만, 처리될 이미지는 비디오 프레임일 수도 있다는 것을 이해해야 한다. 즉, 본 명세서에 기술된 발명대상의 해결책은 비디오의 각 프레임의 변환에도 적용될 수 있다.
- [0053] 도 2a의 아키텍처에서 제1 학습 네트워크(210) 및 제2 학습 네트워크(220)의 트레이닝은 각각 도 3 및 도 4a를 참조하여 아래에서 설명될 것이다.
- [0054] 순차적 기하학적 변형 및 스타일 트랜스퍼를 위한 아키텍처
- [0055] 도 2b는 기하학적 변형 및 스타일 트랜스퍼가 순차적으로 수행되는 제2 구현예에 따른 이미지 변환 모듈(122)의 아키텍처를 도시한 것이다. 아키텍처의 구성요소 및 그 안에 구현된 특정 기능은 도 2a에 도시된 구현예와 동

일한다. 구체적으로, 이미지 변환 모듈(122)은 기하학적 변형을 위한 제1 학습 네트워크(210)와 스타일 트랜스퍼를 위한 제2 학습 네트워크(220)를 포함하고, 기하학적 검출기(230) 및 병합 모듈(240)을 더 포함할 수 있다.

[0056] 도 2a의 구현예와 달리, 제1 학습 네트워크(210)가 소스 이미지(102)에 대한 기하학적 변형을 수행하여 소스 이미지(102) 내 객체의 제1 기하구조(202)를 제2 기하구조(204)로 변형시킨 후, 제2 기하구조(204) 및 소스 이미지(102)가 병합 모듈(240)에 함께 입력된다. 병합 모듈(240)은 제2 기하구조(204)에 기초하여 소스 이미지(102)를 변환하여 중간 이미지(242)를 생성한다. 병합 모듈(240)은 변형된 제2 기하구조(204)의 안내 하에 소스 이미지(102)에 대해 이미지 워핑을 수행하여, 생성된 중간 이미지(242)의 객체가 제2 기하구조(202)와 동일하거나 유사한 기하구조를 갖도록 한다. 워핑은 소스 이미지(102)에 대해 직접 수행되기 때문에, 중간 이미지(242)는 소스 이미지(102)의 동일한 제1 스타일(예컨대, 실제 사진 스타일)을 유지한다.

[0057] 중간 이미지(242)는, 타겟 이미지(104)를 생성하기 위한 스타일 트랜스퍼를 수행하기 위해, 제2 학습 네트워크(220)에 입력된다. 제2 학습 네트워크(220)에서의 스타일 트랜스퍼의 구체적인 동작은 도 2a를 참조하여 설명한 구현예와 유사하므로, 여기서는 세부사항은 생략한다. 제2 학습 네트워크(220)는 (사용자가 임의로 선택하거나 지정한) 캐리커처 도메인의 특정 스타일을 중간 이미지(242)에 적용할 수 있다. 스타일 트랜스퍼만 수행되므로, 제2 학습 네트워크(220)에 의해 출력된 타겟 이미지(104) 내의 객체의 기하구조는 중간 이미지(242)의 기하구조와 실질적으로 일관되게 유지된다. 따라서, 획득된 타겟 이미지(104)는, 스타일이 캐리커처와 유사하고, 기하구조의 측면에서 캐리커처 도메인에서 기하학적 과장 특성을 나타낸다.

[0058] 전술한 바와 같이, 도 2b의 구성요소들의 기능은 도 2a의 기능과 유사하다. 학습 네트워크의 경우, 입력 처리 능력은 트레이닝을 통해 얻어진다. 도 2a 및 도 2b의 구현예에서, 제1 학습 네트워크(210)는 사진 도메인의 현실 세계 객체의 기하구조를 캐리커처 도메인의 과장된 기하구조로 변형시키는 데 사용된다. 따라서, 제1 학습 네트워크(210)의 트레이닝 프로세스가 동일하며, 이는 도 3을 참조하여 아래에서 상세히 논의할 것이다. 스타일 트랜스퍼를 수행하도록 동작하지만, 제2 학습 네트워크(220)의 입력 유형은 도 2a 및 도 2b에서 상이하다. 도 2a의 구현예에서, 제2 학습 네트워크(220)에 의해 학습되는 능력은 사진 도메인에서 소스 이미지를 처리하는 것인 반면에, 도 2b의 구현예에서, 제2 학습 네트워크(220)에 의해 학습되는 능력은 기하학적 변형 후의 이미지를 처리하는 것이다. 따라서, 도 2b에서 제2 학습 네트워크(220)를 트레이닝하기 위한 트레이닝 데이터는 도 2a에서 제2 학습 네트워크(220)를 트레이닝하기 위한 트레이닝 데이터와 다르다. 도 2b에서 사용된 제2 학습 네트워크(220)의 트레이닝은 도 4b를 참조하여 아래에서 설명할 것이다.

[0059] 기하학적 변형을 위한 학습 네트워크 트레이닝

[0060] 위에서 언급한 바와 같이, 제1 학습 네트워크(210)는 기하학적 변형에 대한 매핑(Φ_{geo})을 획득하여, 제2 기하구조가 제1 기하구조 내의 독특한 특징을 갖는 부분의 과장된 변형을 포함하도록, 입력 소스 이미지 내의 객체의 제1 기하구조를 변형된 제2 기하구조에 매핑하는 것이다. 기하구조를 표현하기 위해 랜드마크 포인트를 사용하는 구현예에서, L_X 및 L_Y 는 사진 도메인 X 및 캐리커처 도메인 Y 에서 랜드마크 포인트의 도메인이라고 가

정한다. 제1 학습 네트워크(210)는 기하학적 변형에 대한 매핑 $\Phi_{geo} : L_X \rightarrow L_Y$ 을 학습하도록 트레이닝되어, 도메인 X 내의 사진(x)의 랜드마크 포인트(l_x)에 대해 도메인 Y 내의 변형된 랜드마크 포인트들 ($l_y \in L_Y$)이 생성된다.

[0061] 제1 학습 네트워크(210)를 트레이닝하기 위한 트레이닝 데이터는 사진 도메인 X 의 제1 이미지 ($\{x_i\}_{i=1, \dots, N}, x_i \in X$ 로 표현되며, N 은 트레이닝할 제1 이미지의 수를 나타냄) 및 캐리커처 도메인 Y 의 제2 이미지 ($\{y_i\}_{i=1, \dots, M}, y_i \in Y$ 로 표현되며, M 은 트레이닝할 제2 이미지의 수를 나타냄)를 포함한다. 도 3 및 도 4a 및 4b에서 사용된 사진 도메인 및 캐리커처 도메인의 사진은 도 2a 및 2b의 트레이닝된 학습 네트워크에 의해 실제로 처리되고 생성된 두 이미지 도메인의 사진과 유사한 것으로 설명되지만, 이것은 단지 설명을 위한 것이다. 본 명세서에 기술된 발명대상의 구현에서, 트레이닝에 사용되는 사진 도메인 및 캐리커처 도메인의 이미지는 임의의 실제 사진 및 생성된 캐리커처일 수 있다.

[0062] 트레이닝시, 트레이닝 이미지가 사진 도메인과 캐리커처 도메인에서 획득되지만, 트레이닝에 사용되는 두 이미지 도메인의 이미지는 쌍을 이루지 않을 수도 있는데, 즉, 주어진 제1 이미지가 매칭된 도메인 Y 의 제2 이미지

를 갖도록 요구되지는 않는다. 사진 및 캐리커처는 기존 데이터베이스에서 임의로 선택할 수 있다. 일부 구현예에서, 여기서의 제1 이미지와 제2 이미지는 동일한 유형의 객체(예컨대, 사람, 초상화, 동물 등)를 가질 수 있지만, 객체의 기하학적 특징과 스타일 특징은 상이하며, 특히 객체의 기하학적 변형의 정도가 상이하다. 일부 구현예에서, 제1 이미지 및 제2 이미지는 동일한 해상도의 동일한 크기로 크기조정될 수 있다.

[0063] 랜드마크 포인트 기반 기하구조 표현에서, 랜드마크 포인트는 트레이닝을 위해 제1 이미지 및 제2 이미지 모두에 표시된다. 따라서, 트레이닝을 위해 이들 이미지에서 랜드마크 포인트를 추출할 수 있다. 가능한 모든 객체의 기하구조를 수집하기 위해, 유사한 변환을 활용하여 트레이닝할 제1 이미지와 제2 이미지를 여러 랜드마크 포인트(예컨대, 두 눈의 중심과 입의 중심을 포함한 사람의 얼굴의 세 개의 랜드마크 포인트)를 통해 객체의 평균 형상으로 정렬할 수 있다).

[0064] 진술한 바와 같이, 제1 학습 네트워크(210)는 GAN 네트워크에 기초할 수 있다. GAN 네트워크는 도메인 X로부터 도메인 Y로의 기하학적 변형을 위한 제1 서브네트워크와 도메인 Y로부터 도메인 X로의 기하학적 변형을 위한 제2 서브네트워크를 포함한다. 제1 서브네트워크는 제1 학습 네트워크(210)에 대응한다. 트레이닝 데이터가 쌍을 이루지 않기 때문에, 변형을 위한 L_X 및 L_Y 의 랜드마크 포인트도 쌍을 이루지 않는다. 따라서, 제1 서브네트워크의 결정된 매핑($\Phi_{geo} : L_X \rightarrow L_Y$)이 적게 제한된다(under-constrained). 따라서, 추가 제한을 위해 제2 서브네트워크의 역 매핑($\Phi_{geo}^{-1} : L_Y \rightarrow L_X$)이 사용될 수 있다. 네트워크의 트레이닝 프로세스는 매핑을 수행하는 네트워크 파라미터(예컨대, 학습 네트워크의 파라미터)를 결정하고 최적화하는 것이다.

[0065] 도 3은 GAN 기반 구현예에서 제1 학습 네트워크(210)를 트레이닝하기 위한 GAN 구조(300)를 도시한다. GAN 구조(300)는 순환 일관성(cycle consistency)에 기초한 GAN 구조이다. GAN 구조(300)는, 이 예에서 2개의 생성기(210, 304) 및 2개의 판별기(320 및 322)인, 복수의 서브네트워크를 포함한다. 생성기 중 하나는 제1 학습 네트워크(210)에 대응하며, 이는 순방향 생성기(210)(G_{LY} 로 표현됨)라고도 하는데, 도메인 X로부터 도메인 Y로의 기하학적 변형(즉, 매핑 Φ_{geo})을 모델링하는 데 사용된다. 다른 생성기는 도메인 Y로부터 도메인 X로의 역 기하학적 변형(즉, 역 매핑 Φ_{geo}^{-1})을 모델링하기 위한 역방향 생성기(304)(G_{LX} 로 표현됨)이다. 여기서, 순방향 생성기(210)를 GAN의 제1 서브네트워크라고도 하고, 역방향 생성기(304)를 GAN의 제2 서브네트워크라고도 한다.

[0066] GAN 구조(300)는 2개의 경로를 포함한다. 제1 경로는 제1 이미지를 기반으로 네트워크를 트레이닝하기 위한 도 3의 상단 부분이고 제2 경로는 제2 이미지를 기반으로 네트워크를 트레이닝하기 위한 도 3의 하단 부분이다. 제1 경로에서, 순방향 생성기(210)는 도메인 X의 제1 이미지의 기하구조(l_x)(311)에 대해 도메인 X로부터 도메인 Y로의 기하학적 변형(즉, 매핑 Φ_{geo})을 수행하여, 변형된 기하구조($\hat{l}_y$)(313)를 생성한다. 변형된 기하구조(313)는 $\hat{l}_y = G_{LY}(l_x)$ 로 표현될 수 있다. 순방향 생성기(210)는 PCA 표현에 기초하여 동작할 수 있다. 보다 잘 볼 수 있도록, 변형 전후의 기하구조에 대응하는 이미지 객체 또한 도 3에 표시되어 있다.

[0067] 구체적으로, 순방향 생성기(210)는 기하구조(311)의 PCA 표현(312)을 입력으로서 취하고 변형 후에 변형된 기하구조(313)의 PCA 표현(314)을 출력한다. 역방향 생성기(304)는 변형된 기하구조(313)에 대해 도메인 Y로부터 도메인 X로의 역 기하학적 변형(즉, 역 매핑 Φ_{geo}^{-1})을 수행하여 L_X 에서 예측된 기하구조($\hat{l}_x$)(315)를 획득한다. 역방향 생성기(304)로의 입력은 PCA 표현(314)일 수 있고, 역방향 생성기(304)는 예측된 기하구조(315)의 PCA 표현(316)을 출력한다. 판별기(320)는 변형된 기하구조($\hat{l}_y$) 및 그라운드 트루스(ground-truth) 기하구조(l_y)가 구별가능한지 여부를 판별하기 위해 사용된다. 그라운드 트루스 기하구조(l_y)는 도메인 Y로부터 샘플링된 이미지의 기하구조이다. 이 경로에서, 예측된 기하구조($\hat{l}_y$)는 순방향 생성기의 결과를 결정하기 위해 판별기(320)에 의해 사용될 뿐만 아니라, 역방향 생성기(304)에 피드백될 수도 있다.

[0068] 도 3의 하단에 있는 제2 경로는 제1 경로와 유사한 동작을 수행하지만 도메인 Y의 제2 이미지를 입력으로 사용

하며, 따라서 생성기들의 순서가 조정된다. 역방향 생성기(304)는 제2 이미지의 기하구조(l_y)에 대한 도메인 Y로부터 도메인 X로의 기하학적 변형(즉, 역 매핑 Φ_{geo}^{-1})을 수행하여, 변형된 기하구조($\hat{l}_x$)(333)를 생성한다. 역방향 생성기(304)는 기하구조(331)의 PCA 표현(332)을 입력으로 취하고 변형 후에 변형된 기하구조(333)의 PCA 표현(334)을 출력할 수 있다. 순방향 생성기(210)는, PCA 표현에 기초하여, 변형된 기하구조(333)에 대해 도메인 X로부터 도메인 Y로의 기하학적 변형(즉, 매핑 Φ_{geo})을 수행하여, L_x 에서 예측된 기하구조($\hat{l}_y$)(315)를 획득한다. 순방향 생성기(210)로의 입력은 PCA 표현(334)일 수 있고, 순방향 생성기(210)는 예측된 기하구조(335)의 PCA 표현(336)을 출력한다. 판별기(322)는 변형된 기하구조($\hat{l}_x$) 및 그라운드 트루스(ground-truth) 기하구조(l_x)가 구별가능한지 여부를 판별하기 위해 사용된다. 그라운드 트루스 기하구조(l_x)는 도메인 X로부터 샘플링된 이미지의 기하구조이다. 제1 및 제2 경로에서, 순방향 생성기(210) 및 역방향 생성기(304)의 파라미터 값이 공유된다.

[0069] 도 3의 GAN 구조(300)에서 순방향 생성기(210)는 도메인 X로부터 도메인 Y로의 기하학적 변형을 모델링하기 때문에, 기하학적 변형을 위한 제1 학습 네트워크(210)로 간주될 수 있다.

[0070] 일부 구현에서, 순방향 생성기(210) 및 역방향 생성기(304)는 PCA 표현에 기초하여 동작을 수행하기 때문에, 이 네트워크 구조는, 복잡한 컨볼루션 레이어를 사용하지 않고, 완전 연결(FC) 계층(예컨대, ReLu와 같은 활성화 기능에 대응하는 노드를 처리함으로써 형성됨)을 포함하도록 구성될 수 있다. 컨볼루션 레이어는 일반적으로 픽셀 단위 이미지 처리에 적합하며 더 복잡한 계산을 요구한다. 그러나, PCA 표현의 낮은 차원으로 인해, 단순한 완전 연결 계층 또한 네트워크 학습을 달성할 수 있다.

[0071] 학습 네트워크의 학습 프로세스는 목적 함수(objective function)를 기반으로 네트워크 파라미터를 지속적으로 반복적으로 업데이트하거나 최적화하는 것이다. 목적 함수는 손실 함수 또는 비용 함수로 결정될 수 있다. 본 명세서에 개시된 발명대상의 구현예에서, 손실 함수가 예로서 취해지지만, 그러한 트레이닝은 또한 비용 함수에 기초하여 수행될 수 있음을 이해해야 한다. 일반 GAN 네트워크의 트레이닝에서, 두 유형의 손실, 즉, 적대적 손실($\mathcal{L}_{adv}^{L_Y}$ 로 표현됨)과 양방향 순환 일관성 손실($\mathcal{L}_{cyc}$ 로 표현됨)이 널리 고려된다.

[0072] 적대적 손실은, 결과의 변형된 기하구조가 도메인 Y로부터 샘플링된 캐리커처 이미지의 기하구조와 유사하도록, 네트워크 파라미터가 지속적으로 업데이트되도록 장려한다. 제1 경로에 대해, 적대적 손실은 다음과 같이 표현될 수 있다.

수학식 1

$$\mathcal{L}_{adv}^{L_Y}(G_{L_Y}, D_{L_Y}) = \mathbb{E}_{l_y \sim L_Y} [\log D_{L_Y}(l_y)] + \mathbb{E}_{l_x \sim L_X} [\log(1 - D_{L_Y}(G_{L_Y}(l_x)))]$$

[0073]

[0074] 적대적 손실에 기초하여, 순방향 생성기(210)는 출력 예측 기하구조가 판별기(320)에 의해 그라운드 트루스 기하구조와 구별될 수 없도록 업데이트될 수 있다. 트레이닝 아키텍처의 제2 경로에서, 적대적 손실($\mathcal{L}_{adv}^{L_X}(G_{L_X}, D_{L_X})$ 로 표현됨)은, 예를 들어 수학식 1로 나타난 방식으로 유사하게 결정될 수 있다.

[0075] GAN 네트워크에서 양방향 순환 일관성 손실은 순방향 매핑(Φ_{geo})과 역방향 매핑(Φ_{geo}^{-1}) 사이의 순환 일관성을 제한하는 데 사용될 수 있다. 그 원리는, 순방향 생성기(210)가 제1 이미지(l_x)에 대해 기하학적 변형을 수행하기 위한 제1 학습 네트워크(210)로 사용되면, 역 기하구조(304)는 변형된 기하구조를 다시 제1 이미지(l_x)의 기하구조로 정확하게 변형할 수 있다는 것이다(즉, $G_{L_X}(G_{L_Y}(l_x)) \approx l_x$). 유사하게, 역 방향에서 일관성은 $G_{L_Y}(G_{L_X}(l_y)) \approx l_y$ 로 정의될 수 있다. 따라서, 양방향 순환 일관성은 다음과 같이 표현할 수 있다.

수학식 2

$$\mathcal{L}_{\text{cyc}}(G_{L_Y}, G_{L_X}) = \mathbb{E}_{l_x \sim L_X} [\|G_{L_X}(G_{L_Y}(l_x)) - l_x\|_1] \\ + \mathbb{E}_{l_y \sim L_Y} [\|G_{L_Y}(G_{L_X}(l_y)) - l_y\|_1].$$

[0076]

[0077]

수학식 2에서, 각 방향에서의 순환 일관성 손실은 두 벡터 사이의 L1 차로 표현될 수 있다. 양방향 순환 일관성 손실은 또한 입력으로부터 출력으로의 매핑을 제한할 수 있다.

[0078]

적대적 손실 및 양방향 순환 일관성 손실은 흔히 GAN 네트워크에서 고려되는 두 가지 손실이다. 그러나, 기하학적 변형 학습 작업에서 두 손실을 기반으로 한 트레이닝은, 변형된 기하구조가 원래 기하구조에서 눈에 띄는 특징을 포착하고 이런 특징을 과장할 수 있는지 확인하기가 여전히 어려울 수 있다. 본 명세서에 기술된 발명 대상의 일부 구현에서, 입력된 제1 이미지 내의 객체의 기하구조와 사진 도메인에서의 이런 유형의 객체의 평균 기하구조 사이의 차를 유지하기 위해 추가 손실이 트레이닝에 도입된다. 이는 특정 객체의 기하구조와 그러한 유형의 객체의 평균 기하구조 사이의 차가 객체의 독특한 눈에 띄는 특징을 나타내므로 기하학적 변형 후에도 유지되거나 또는 강조되어야 하기 때문이다. 예를 들어, 어떤 사람의 얼굴 크기가 평범한 얼굴의 경우보다 크다면, 캐리커처 도메인에서의 변형 후에도 이 독특한 눈에 띄는 특징을 유지하거나 강조해야 한다. 따라서, 트레

닝 동안 새로 도입된 손실을 특성 손실($\mathcal{L}_{\text{cha}}^{L_Y}$ 로 표현됨)이라고도 한다.

[0079]

특성 손실($\mathcal{L}_{\text{cha}}^{L_Y}$)은 GAN 구조(300)의 제1 경로 및 제2 경로 모두에서 고려될 수 있다. 제1 경로에서, 특성 손실($\mathcal{L}_{\text{cha}}^{L_Y}$)은 입력된 제1 이미지의 기하구조(311)에서 대응하는 이미지 도메인에서의 평균 기하구조를 뺀 결과와 순방향 생성기(210)에 의해 생성된 변형된 기하구조(313)에서 대응하는 이미지 도메인에서의 평균 기하구조를 뺀 결과 사이의 차에 페널티를 적용하여 고려될 수 있다. 구체적으로, 트레이닝할 각각의 입력된 제1 이미지에 대해, 현재 순방향 생성기(210)에 의해 변형된 기하구조(313)가 생성된 후에, 제1 이미지에서의 기하구조(311)(l_x 로 표현됨)와 사진 도메인 X에서의 평균 기하구조($\overline{L_X}$ 로 표현됨) 사이의 제1 차이, 및 변형된 기하구조(313)($G_{L_Y}(l_x)$ 로 표현됨)와 캐리커처 도메인 Y에서의 평균 기하구조($\overline{L_Y}$ 로 표현됨) 사이의 제2 차이가 또한 결정된다. 특성 손실($\mathcal{L}_{\text{cha}}^{L_Y}$ 로 표현됨)은 제1 및 제2 차이 사이의 차를 최소화하도록 결정될 수 있다. 이것은, 코사인 차와 같은, 제1 및 제2 차이 사이의 차에 페널티를 적용하여 결정될 수 있다. 여기서, 특성 손실($\mathcal{L}_{\text{cha}}^{L_Y}$ 로 표현됨)은 다음과 같이 표현될 수 있다.

수학식 3

$$\mathcal{L}_{\text{cha}}^{L_Y}(G_{L_Y}) = \mathbb{E}_{l_x \sim L_X} [1 - \cos(l_x - \overline{L_X}, G_{L_Y}(l_x) - \overline{L_Y})]$$

[0080]

[0081]

트레이닝 목표는, 사진 도메인의 객체의 기하구조와 평균 기하구조 사이의 차이와, 캐리커처 도메인의 객체의 기하구조와 평균 기하구조 사이의 차이가 동일하도록 $\mathcal{L}_{\text{cha}}^{L_Y}$ 를 최소화하는 것이다.

[0082]

대칭적으로, GAN 구조(300)의 제2 경로에서, 역방향의 특성 손실($\mathcal{L}_{\text{cha}}^{L_X}(G_{L_X})$ 로 표현됨)이 또한 결정될 수 있다. 구체적으로, 트레이닝할 각각의 입력된 제2 이미지에 대해, 현재 역방향 생성기(305)에 의해 변형된 기하구조(333)가 생성된 후에, 제2 이미지에서의 기하구조(331)(l_y 로 표현됨)와 캐리커처 도메인 Y에서의 평균 기하구조($\overline{L_Y}$ 로 표현됨) 사이의 제3 차이, 및 변형된 기하구조(333)($G_{L_X}(l_y)$ 로 표현됨)와 사진 도메인에서의 평균 기하구조($\overline{L_X}$ 로 표현됨) 사이의 제4 차이가 또한 결정될 수 있다.

조($\bar{L}_X$ 로 표현됨) 사이의 제4 차이가 또한 결정된다. 역 특성 손실($\mathcal{L}_{cha}^{L_X}(G_{L_X})$ 로 표현됨)은 제3 차이 및 제4 차이 사이의 차를 최소화하도록 결정될 수 있다. 이것은, 코사인 차와 같은, 제1 및 제2 차이 사이의 차에 페널티를 적용하여 결정될 수 있다. 역 특성 손실($\mathcal{L}_{cha}^{L_X}(G_{L_X})$ 로 표현됨)은 수학적 식 3과 유사하게 정의될 수 있다.

[0083] 전술한 다양한 손실을 고려한 후에, GAN 구조(300)를 트레이닝하기 위한 목적 함수는 다음 손실 함수를 최적화하도록 결정될 수 있다.

수학적 식 4

$$\mathcal{L}_{geo} = \mathcal{L}_{adv}^{L_X} + \mathcal{L}_{adv}^{L_Y} + \lambda_{cyc} \mathcal{L}_{cyc} + \lambda_{cha} (\mathcal{L}_{cha}^{L_X} + \mathcal{L}_{cha}^{L_Y})$$

[0084]

[0085] 이 손실 함수는, 객체의 기하구조의 독특한 특징을 유지하기 위한 특성 손실을 포함한, 제1 경로 및 제2 경로에 서의 상이한 유형의 손실을 고려한다. 수학적 식 4에서, 파라미터(λ_{cyc} , λ_{cha})는 상이한 유형의 손실 균형을 맞추기 위해 미리 결정된 값이다.

[0086] 트레이닝 프로세스 동안, GAN 구조(300)는 포토 도메인 및 캐리커처 도메인의 복수의 제1 및 제2 이미지를 사용하여 트레이닝될 수 있으며, 여기서 제1 및 제2 이미지 각각은, GAN 구조(300)를 통과한 후 순방향 생성기(210) 및 역방향 생성기(304)의 파라미터 값을 업데이트하는 것을 용이하게 할 수 있다. GAN 구조(300)에서, 순방향 생성기(210) 및 역방향 생성기(304)는 함께 트레이닝된다. 확률적 경사 하강법 알고리즘(stochastic gradient descent algorithm), 순방향 전파, 역방향 전파 등과 같은 머신 러닝을 위한 다양한 트레이닝 기술이 GAN 구조를 트레이닝하는 데 사용될 수 있다. 전체 손실 함수(예컨대, 수학적 식 4)의 집합점(convergence)에 도달한 후, 트레이닝된 순방향 생성기(210) 및 역방향 생성기(304)가 획득된다.

[0087] 트레이닝이 완료된 후, 순방향 생성기(210) 및 역방향 생성기(304)의 파라미터 세트가 결정될 수 있다. 순방향 생성기(210)는, 사진 도메인 X으로부터 캐리커처 도메인 Y으로의 기하학적 변형을 수행할 수 있는, 제1 학습 네트워크(210)로 결정될 수 있다. GAN 구조(300)의 다른 부분은 제1 학습 네트워크(210)를 트레이닝하기 위해 구성된 것으로 간주될 수 있으며, 네트워크의 적용 단계에서 사용되지 않는다.

[0088] 도 3에서 트레이닝을 위한 제1 이미지는 도 2a 및 2b의 적용 단계에서 사용된 것과 유사한 사진을 갖는 것으로 도시되어 있지만, 트레이닝 단계 동안에는, 트레이닝을 수행하기 위해 사진 도메인의 기존의 이용가능한 어떠한 사진도 검색될 수 있다. 트레이닝 목적의 제1 이미지는 특정 유형의 관심 객체를 포함하기만 하면 된다.

[0089] 스타일 트랜스퍼를 위한 학습 네트워크의 제1 예시적인 트레이닝

[0090] 이상에서는 제2 학습 네트워크(220)가 캐리커처 도메인에서 제1 스타일의 이미지를 제2 스타일의 이미지로 전이하기 위해 스타일 트랜스퍼를 위한 매핑(Φ_{app})을 획득하는 것을 언급하였다. 제2 학습 네트워크(220)의 트레이

닝 동안, 먼저 사진 도메인 X 내의 제1 이미지($\{x_i\}_{i=1, \dots, N}, x_i \in X$) 및 캐리커처 도메인 Y 내의 제2 이미지($\{y_i\}_{i=1, \dots, M}, y_i \in Y$)가 트레이닝 이미지로서 직접 사용될 경우, 제2 학습 네트워크(220)에 의해 출력된 이미지 내의 객체들의 기하구조 및 스타일은 서로 혼동될 수 있다.

[0091] 트레이닝 이미지에서 기하학적 차이의 충격을 피하기 위해, 본 명세서에 기술된 발명대상의 구현예에서 중간 도메인을 구성하는 것이 제안된다. 도 2a의 아키텍처에서 사용되는 제2 학습 네트워크(220)가 트레이닝될 경우, 트레이닝을 위한 캐리커처 도메인 Y의 제2 이미지는 기하학적 변형에 의해 중간 도메인(Y'으로 표현됨)의 중간 이미지($\{y'_i\}_{i=1, \dots, M}, y'_i \in Y'$ 로 표현됨)로 변환된다. 중간 도메인 Y'은 사진 도메인 X와 캐리커처 도메인 Y 사이에 있으며, 중간 이미지는 캐리커처 도메인 Y의 제2 이미지와 동일한 스타일을 상속하지만, 사진 도메인 X의 객체와 유사한 기하구조를 갖는다. 중간 도메인 Y'의 중간 이미지는 제2 이미지에 대해 캐리커처 도메인 Y으로부터 포토 도메인 X로 기하학적 변형을 수행하여 생성될 수 있다. 일부 구현예에서, 도메인 Y로부터 도메인 X로의 이런 매핑은 제1 학습 네트워크(210)가 트레이닝될 때 역방향 생성기(304)에 의해 학습될 수 있다. 따라서, 일부 구현예에서, GAN 구조(300)가 먼저 트레이닝될 수 있다. 제1 학습 네트워크(210)가 트레이닝될

경우, 역방향 생성기(304)는 또한 캐리커처 도메인 Y의 제2 이미지를 Y'의 중간 도메인의 중간 이미지, 즉, $G_{LX}(y_i)$ 로 변환하도록 트레이닝된다.

[0092] 제2 학습 네트워크(220)는 포토 도메인 X의 제1 이미지와 중간 도메인 Y'의 중간 이미지에 기초하여 포토 도메인 X로부터 캐리커처 도메인 Y로 스타일을 전이할 수 있는 능력을 획득하도록 결정될 수 있다. 제2 학습 네트워크(220)는, 기하학적 변형은 고려하지 않고 스타일 이미지 변환뿐만 도메인 X로부터 도메인 Y로의 매핑을 획득할 수 있다. 따라서, 제2 학습 네트워크(220)는 스타일 트랜스퍼에 적합한 임의의 학습 네트워크를 활용할 수 있다. 실제로, 스타일화의 관점에서 보면, 도메인 X로부터 도메인 Y'로의 스타일 트랜스퍼는 도메인 X로부터 도메인 Y로의 스타일 트랜스퍼와 같은데, 왜냐하면 도메인 Y'의 스타일은 도메인 Y의 스타일과 일관되기 때문이다.

[0093] 일부 구현예에서, 제1 학습 네트워크(210)와 유사하게, 제2 학습 네트워크(220) 또한 GAN 네트워크에 기초할 수 있다. 일부 예들에서, 제2 학습 네트워크(220)는 (도 3에서와 같이) 제1 학습 네트워크(210)와 유사한 GAN과 같은 순환 일관성을 갖는 GAN 네트워크에 기초할 수 있다. 이러한 GAN 네트워크는 일대일 출력만 생성할 수 있다. 다른 구현예에서, 제2 학습 네트워크(220)는, 다대다 매핑을 구현하며 비지도 학습을 수행할 수 있는 GAN 네트워크에 기초할 수 있다. 현재, 다대다 매핑을 구현하는 일부 학습 네트워크는 주로 특정 레벨에서 순환 일관성을 고려하며, 이는 이미지의 객체의 기하구조를 유지하기 어렵게 한다. 따라서, 제2 학습 네트워크(220)를 트레이닝하는 일부 구현에서, 제2 학습 네트워크(220)의 트레이닝은 이미지 레벨 순환 일관성을 사용하여 제한된다.

[0094] 도 4a는 GAN 기반 구현예에서 제2 학습 네트워크(220)를 트레이닝하기 위한 GAN 구조(400)를 도시한다. 제2 학습 네트워크(220)의 트레이닝은 도 3과 유사하게 반대의 양방향 경로로 이루어진 순환 일관성 GAN 구조에 기초할 수 있다. 그러나, 예시를 쉽게 하기 위해, 도 4a에는 도메인 X로부터 도메인 Y'으로의 제1 경로만 도시되어 있다. 다른 방향의 제2 경로는 도 3의 제2 경로와 유사하게 트레이닝을 위해 대칭으로 구성될 수 있다. 도메인 X로부터 도메인 Y'로의 방향으로의 제1 경로의 처리 및 트레이닝을 아래에서 먼저 설명할 것이다.

[0095] GAN 구조(400)에서, GAN의 생성기는 인코더 부분과 디코더 부분으로 더 세분된다. 입력 이미지의 경우, 두개의 인코더를 사용하여 이미지의 콘텐츠와 스타일을 각각 모델링하고, 입력 이미지의 콘텐츠 관련 특징 표현과 외형 관련 특징 표현을 추출한다. 이러한 콘텐츠와 스타일의 분리는 동일한 콘텐츠에 서로 다른 스타일을 적용하여 서로 다른 결과물을 얻을 수 있도록 한다. 디코더는 인코더의 역 동작을 수행하는 데 사용된다. 이런 구조에서, 스타일 트랜스퍼를 위한 제2 학습 네트워크(220)는 도메인 X의 인코더와 도메인 Y'로 디코딩하는 디코더로 구성된다.

[0096] 도시된 바와 같이, GAN 구조(400)는 도메인 X에 대한 콘텐츠 인코더(410)(E_X^c 로 표현됨), 스타일 인코더(220-1)(E_X^s 로 표현됨) 및 도메인 Y'에 대한 디코더(220-2)로 구성된 순방향 생성기 부분을 포함한다. 스타일 인코더(220-1) 및 디코더(220-2)는 제2 학습 서브네트워크(220)를 형성한다. GAN 네트워크(400)는 도메인 Y'에 대한 콘텐츠 인코더(420)($E_{Y'}^c$ 로 표현됨), 스타일 인코더(430)($E_{Y'}^s$ 로 표현됨), 및 도메인 X에 대한 디코더(440)(R_X)로 이루어진 역방향 생성기 부분을 더 포함한다.

[0097] 각각의 도메인 I (I=X 또는 Y')에 대해, 콘텐츠 관련 특징 표현(C_I , 콘텐츠 코드라고도 함, 예컨대 도 4a에 도시된 도메인 X의 콘텐츠 코드(403) 및 도메인 Y'의 콘텐츠 코드(422))이 콘텐츠 인코더(E_I^c)(예컨대, 제1 경로 내의 도메인 X의 콘텐츠 인코더(410))에 의해 입력 이미지($z_I \in I$ 로 표현되며, 도 4a에 도시된 도메인 X 내의 이미지(402))로부터 추출된다. 복수의 입력 이미지로부터 추출된 콘텐츠 관련 특징 표현은 C_I 로 표현된 콘텐츠 라이브러리를 형성할 수 있다. 각 도메인 I(I=X 또는 Y')에 대해, 스타일 인코더(E_I^s)는 입력 이미지($z_I \in I$ 로 표현되며, 예컨대 도 4a에 도시된 도메인 X 내의 제1 이미지 또는 도시되지 않은 도메인 Y' 내의 중간 이미지)로부터 외형 관련 특징 표현(S_I , 스타일 코드라고도 함)을 추출한다. 각각의 외형 관련 특징은 대응하는 이미지가 외형에 나타나는 스타일을 특징으로 할 수 있다. 복수의 입력 이미지로부터 추출된 외형 관련 특징 표현은 스타일 라이브러리(예컨대, 도메인 X에 대한 스타일 라이브러리(SX)(401) 및 도메인 Y에 대한 스타일 라이브러리(SY)(402))를 형성할 수 있다.

브러리($S_{Y'}$)(405))를 형성한다. 각각의 콘텐츠 관련 특징 표현 및 각각의 외형 관련 특징 표현은 $(c_I, s_I) = (E_I^c(z_I), E_I^s(z_I))$ 로 표현될 수 있다. 각각의 도메인 $I(I=X \text{ 또는 } Y)$ 에 대해, 디코더(R_I)는 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 이미지($z_I = R_I(c_I, s_I)$)를 재구성한다. 트레이닝을 위해 복수의 제1 이미지 및 중간 이미지로부터 얻어진 콘텐츠 라이브러리(C_I) 및 스타일 라이브러리(S_I)는 가우스 분포 $\mathcal{N}(0, 1)$ 를 만족한다고 가정한다.

[0098] 도 4a에 도시된 도메인 X 로부터 도메인 Y' 방향의 제1 경로에서, 주어진 도메인 X 의 제1 이미지($x \in X$)에 대해, 콘텐츠 인코더(220-1)가 콘텐츠 관련 특징 표현(403)($c_x = E_X^c(x)$)을 추출하고, 스타일 라이브러리($S_{Y'}$)(405)로부터 특정 스타일과 연관된 외형 관련 특징 표현($s_{y'}$)을 무작위로 선택한다. 그 후, 디코더(220-2)를 사용하여 캐리커처 도메인의 스타일을 갖는 출력 이미지(y')(이는 $y' = R_{Y'}(c_x, s_{y'})$ 로 표현된다)를 생성한다. 콘텐츠 인코더(420), 스타일 인코더(430), 및 디코더(440)도 유사하게 동작한다.

[0099] 이런 GAN 구조(400)를 트레이닝하기 위해, 손실 함수와 같은 대응하는 목적 함수가 결정될 수 있다. 본 명세서에 기술된 발명대상의 구현예에서, 도 3을 참조하여 기술된 적대적 손실 및 순환 일관성 손실과 같은 GAN 네트워크의 정상적인 트레이닝에서의 전형적인 손실이 고려될 수 있다.

[0100] GAN 구조(400)에서, 적대적 손실은 디코더(220-2)에 의해 생성된 결과 $R_{Y'}(E_X^c(x), s_{y'})$ 를 중간 도메인 Y' 로부터 선택된 그라운드 트루스 중간 이미지와 동일하도록 제한하는데 사용된다. GAN 구조(400)는 또한 $R_{Y'}(E_X^c(x), s_{y'})$ 이 그라운드 트루스 중간 이미지와 동일한지 여부를 판별하기 위한 판별기(432)를 포함할 수 있다. 도 4a에 도시되지 않은 반대 방향의 제2 경로에서, 대칭 적대적 손실($\mathcal{L}_{adv}^X(E_{Y'}, R_X, D_X)$)이 또한 얻어지며, 여기서 D_X 는 도메인 X 에 대한 판별기를 나타낸다.

[0101] 일부 구현예에서, 트레이닝을 위한 또 다른 유형의 손실은 콘텐츠 관련 특징 표현 및 입력 이미지로부터 추출된 외형 관련 특징 표현에 기초하여 결정된 입력 이미지(402)와 예측 이미지(406) 사이의 차이에 페널티를 적용하여, 입력 이미지가 재구성된 이미지와 최대한 유사하도록 하는 재구성 손실일 수 있다. 입력 이미지 및 재구성된 이미지는 방향에 따라 다를 수 있다. 예를 들어, 도 4a에 도시된 바와 같이 도메인 X 로부터 도메인 Y' 로의 방향의 제1 경로에서, 입력 이미지는, 도메인 X 에 있고 사진 스타일을 갖는 이미지(402)이고, 예측 이미지(406)는 입력 이미지(402)의 콘텐츠 관련 특징 표현(403) 및 도메인 X 에 대한 스타일 라이브러리(S_X)(401)로부터 무작위로 선택된 외형 관련 특징 표현에 따라 도메인 X 의 디코더(440)에 의해 재구성된다. 도메인 Y' 로부터 도메인 X 로의 방향의 대칭의 제2 경로에서, 입력 이미지는, 도메인 Y' 에 있고 캐리커처 스타일을 갖는 이미지이다. 재구성된 이미지는 또한 입력 이미지의 콘텐츠 관련 특징 표현 및 도메인 Y' 의 스타일 라이브러리로부터 무작위로 선택된 외형 관련 특징 표현에 따라 유사한 디코더에 의해 재구성된 이미지이다. 재구성 손실은 다음과 같이 나타낼 수 있다.

수학식 5

[0102]
$$\mathcal{L}_{rec}^I(E_I^c, E_I^s, R_I) = \mathbb{E}_{z \sim I} [\|R_I(E_I^c(z), E_I^s(z)) - z\|_1]$$

[0103] 여기서, $I=X$ 또는 Y' 이다. $I=X$ 인 경우, 수학식 5는 도메인 X 로부터 도메인 Y' 으로의 방향의 제1 경로에서의 재구성 손실을 나타내며, $I=Y'$ 인 경우, 수학식 5는 도메인 Y' 으로부터 도메인 X 로의 방향의 제1 경로에서의 재구성 손실을 나타낸다.

[0104] 또 다른 유형의 손실은 순방향 생성부에서의 순방향 매핑 및 역방향 생성부에서의 역방향 매핑을 통과한 후 입력 이미지가 원래 이미지로 다시 변환될 수 있도록 제한하기 위한 순환 일관성 손실일 수 있다. 구체적으로, 도 4a에 예시된 GAN 구조(400)에서, 주어진 입력 이미지(402) $x \in X$ 에 대해, 도메인 X 로부터 도메인 Y' 으로의

매핑이 순방향 매핑에서 콘텐츠 인코더(220-1) 및 디코더(220-2)에 의해 수행되어, 예측 이미지(404)($R_{Y'}(E_X^c(x), s_{y'}), s_{y'} \in S_{Y'}$)로 표현되며, 본 명세서에서 때론 제1 예측 이미지라고도 함)를 획득한다. 전술한 바와 같이, 디코더(220-2)는 전이를 수행하기 위해 도메인 Y'의 스타일(캐리커처 도메인 Y의 스타일과 동일)을 무작위로 선택하며, 따라서 제1 예측 이미지(404)($R_{Y'}(E_X^c(x), s_{y'}), s_{y'} \in S_{Y'}$)는 캐리커처 도메인의 스타일을 갖지만, 객체의 기하구조는 여전히 이미지(402)와 실질적으로 동일하게 유지된다. 따라서, 예측 이미지(404)는 사실 중간 도메인 Y'의 중간 이미지의 예측이다.

[0105] 디코더(220-2)의 결과는, 예측된 중간 이미지의 외형 관련 특징 표현(422)을 추출하도록, 도메인 Y'의 콘텐츠 인코더(420)로 피드백된다. 디코더(440)는 도메인 X에 대한 라이브러리(S_X)(401)로부터 외형 관련 특징 표현을 무작위로 선택하고, 입력 이미지(402)의 콘텐츠 관련 특징 표현(422) 및 선택된 외형 관련 특징 표현에 기초하여 예측 이미지(408)($R_X(E_{Y'}^c(R_{Y'}(E_X^c(x), s_{y'})), s_x)$ 로 표현됨)를 획득한다. 콘텐츠와 스타일 모두가 도메인 X로부터 나오기 때문에, 예측 이미지(408)는 실제로 입력 이미지(402)의 예측이다.

[0106] 도메인 X로부터 도메인 Y'으로의 방향에서의 순환 일관성 손실은 예측 이미지(408)가 일관되게 입력 제1 이미지(402)와 유사하게 되도록 제한하는 데 사용되며, 이는 다음과 같이 표현될 수 있다.

수학식 6

$$\mathcal{L}_{cyc}^X(E_X^c, R_{Y'}, E_{Y'}^c, R_X) = \mathbb{E}_{x \sim X', s_x \sim S_X, s_{y'} \sim S_{Y'}} [\|R_X(E_{Y'}^c(R_{Y'}(E_X^c(x), s_{y'})), s_x) - x\|_2]$$

[0107]

[0108] 수학식 6은 두 이미지의 차이가 L2 차이로 표시됨을 보여준다. 다대다 매핑을 위해 기존의 GAN 네트워크에서 사용되는 순환 일관성 손실과 달리, 본 명세서에 기술된 발명대상의 구현예에서는, 순환 일관성 손실은 이미지 레벨 차이, 즉, 예측 이미지와 그라운드 트루스 입력 이미지 사이의 차이에 기반한다. 트레이닝의 목표는 이러한 차이를 가능한 한 낮추는 것이다. 수학식 6의 예에서 이 차이는 두 이미지 사이의 L2 차이로 표현된다. 물론, 이미지 차이를 결정하기 위한 다른 방법도 사용될 수 있다.

[0109] 일부 구현예에서, 도메인 Y'으로부터 도메인 X로의 방향의 제2 경로에 대한 순환 일관성 손실도 유사하게 계산될 수 있으며, $\mathcal{L}_{cyc}^{Y'}(E_{Y'}^c, R_X, E_X^c, R_{Y'})$ 로 표현된다. 순환 일관성 손실의 결정은 수학식 6과 유사할 수 있지만, 각 변환의 방향은 반대이다(즉, 수학식 6에 포함된 도메인 X로부터 도메인 Y'로의 변환은 도메인 Y'로부터 도메인 X로의 변환으로 돌아간다).

[0110] 전술한 손실 유형은 일반적으로 순환 일관성 및 다대다 매핑을 이용하는 GAN 네트워크에서 고려되는 손실이다. 그러나, 이들 손실만 가지고 GAN 구조(400)를 트레이닝시키는 것은 이미지의 기하학적 특징들로부터 콘텐츠 특성들을 완전히 분리하도록 GAN 구조(400)를 트레이닝시킬 수 없을 수 있으며, 이는 이미지 변환 후에 객체의 기하구조를 유지하는 것을 어렵게 할 수 있다. 본 명세서에 기술된 발명대상의 일부 구현에서, GAN 구조(400)의 트레이닝을 더욱 최적화하기 위해 추가 손실이 또한 고려된다. 이런 손실은 제약 인코더의 입력 이미지를 명시적으로 제한하는 데 사용되는 지각 손실(perceptual loss)이라고 하며, 디코더의 출력 이미지는 콘텐츠에서 변경되지 않고 유지된다.

[0111] 구체적으로, 도 4a에 도시된 GAN 구조(400)에서, 트레이닝을 위해 입력된 제1 이미지(402)에 대해, 콘텐츠 인코더(220-1) 및 디코더(220-2)로 구성된 서브네트워크는 제1 이미지(402)의 스타일 트랜스퍼를 구현하는 데 사용되며, 이것에 의해 제1 예측 이미지(404)($R_{Y'}(E_X^c(x), s_{y'})$ 로 표시됨)를 획득하고, 여기서 제1 예측 이미지(404)의 스타일은 도메인 Y'의 스타일 라이브러리(405)로부터 무작위로 선택된다. 지각 손실은 제1 예측 이미지(404)와 제1 이미지(402) 사이의 콘텐츠의 차이를 의미한다. 이런 차이는 제1 예측 이미지(404)와 제1 이미지(402)로부터 추출된 각 특징 맵들 사이의 차이, 특히 콘텐츠 관련 특징 맵들 사이의 차이를 기반으로 표현될 수 있다. 지각 손실은 다음과 같이 나타낼 수 있다.

수학식 7

$$\mathcal{L}_{\text{per}}^X(E_X^c, R_{Y'}) = \mathbb{E}_{x \sim X, s_{y'} \sim Y'} [\| \text{VGG19}_{5_3}(R_{Y'}(E_X^c(x), s_{y'})) - \text{VGG19}_{5_3}(x) \|_2],$$

[0112]

[0113] 여기서 VGG19_{5_3}(a)는 이미지 a로부터 추출된 콘텐츠 관련 특징 맵을 표현한다. 수학식 7은 특징 맵과 L2 차이 표현 사이의 차이를 나타낸다. 개념 손실($\mathcal{L}_{\text{per}}^X(E_X^c, R_{Y'})$)은 주로 콘텐츠 인코더(220-1)와 디코더(220-2)의 트레이닝을 제한하는 데 사용된다.

[0114] 도 4a에 도시되지 않은 도메인 Y'로부터 도메인 X로의 방향의 제2 경로에서, 지각 손실($\mathcal{L}_{\text{per}}^{Y'}(E_{Y'}^c, R_X)$)로 표시됨)도 유사하게 결정될 수 있다. 이 반대쪽 제2 경로에서, 입력 이미지는 도메인 Y의 제2 이미지에서 변환된 중간 이미지이다. 제2 이미지는 알려진 캐리커처일 수 있으며, 제1 이미지(402)와 매칭되는 캐리커처를 포함할 수도 있고 포함하지 않을 수도 있다. 도메인 Y'로부터 도메인 X로의 스타일 트랜스퍼는 콘텐츠 인코더(420) 및 다른 디코더(도 4a에 포함되지 않음)에 의해 수행된다. 중간 이미지는 도메인 Y'에서의 임의의 스타일(도메인 Y에서의 임의의 스타일과 동일)을 가질 수 있으며(중간 이미지가 제2 이미지로서 도메인 Y에서의 제3 스타일을 갖는다고 가정할 경우), 전이될 스타일은 도메인 X의 스타일 라이브러리(401)로부터 무작위로 선택된 스타일(제4 스타일이라 함)일 수 있다. 지각 손실($\mathcal{L}_{\text{per}}^{Y'}(E_{Y'}^c, R_X)$)은 입력 중간 이미지와 이미지(402)에 대한 제2 예측 이미지 사이의 차이에 기초하여 표현될 수 있다. 양 방향에서의 지각 손실 $\mathcal{L}_{\text{per}}^X(E_X^c, R_{Y'})$ 및 $\mathcal{L}_{\text{per}}^{Y'}(E_{Y'}^c, R_X)$ 은 콘텐츠 인코더(220-1) 및 디코더(220-2)와, 콘텐츠 인코더(420) 및 또 다른 디코더를 함께 트레이닝하는 데 사용될 수 있다.

[0115] 전술한 다양한 손실을 고려함으로써, 도 4a에 도시된 GAN 구조(400) 및 그 대칭 GAN 구조를 트레이닝하기 위한 목적 함수는 다음 손실 함수를 최적화하도록 결정될 수 있다.

수학식 8

$$\mathcal{L}_{\text{app}} = \mathcal{L}_{\text{adv}}^X + \mathcal{L}_{\text{adv}}^{Y'} + \lambda_{\text{rec}}(\mathcal{L}_{\text{rec}}^X + \mathcal{L}_{\text{rec}}^{Y'}) + \lambda_{\text{cyc}}(\mathcal{L}_{\text{cyc}}^X + \mathcal{L}_{\text{cyc}}^{Y'}) + \lambda_{\text{per}}(\mathcal{L}_{\text{per}}^X + \mathcal{L}_{\text{per}}^{Y'})$$

[0116]

[0117] 여기서, 파라미터 λ_{rec} , λ_{cyc} 및 λ_{per} 는 상이한 유형의 손실을 균형있게 조정하는 미리결정된 값이다.

[0118] 트레이닝 프로세스 동안, GAN 구조(400) 및 그 대칭 구조는, 내부 부분의 파라미터 값을 지속적으로 업데이트하기 위해, 사진 도메인 및 중간 도메인의 복수의 이미지를 사용하여 반복적으로 트레이닝될 수 있다. 트레이닝 프로세스 동안, 확률적 경사 하강법, 순방향 전과, 역방향 전과 등과 같은, 머신 러닝을 위한 다양한 트레이닝 기술이 GAN 구조를 트레이닝하는데 사용될 수 있다. 전체 손실 함수(예컨대, 수학식 8)의 집합점에 도달한 후, 트레이닝된 콘텐츠 인코더(220-1) 및 디코더(220-2)가 획득되고, 도메인 X의 주어진 제1 이미지에 대한 스타일 트랜스퍼를 수행하기 위한 제2 학습 네트워크(220)로 결정된다. GAN 구조의 다른 부분은 제2 학습 네트워크(220)를 트레이닝할 목적으로만 구성되고, 네트워크의 적용 단계에서 사용되지 않는다.

[0119] 도 4a의 구조에서 알 수 있는 바와 같이, 제2 학습 네트워크(220)의 디코더(220-2)는 입력 이미지의 동일한 콘텐츠 관련 특징 표현에 대한 전이를 수행하기 위해 다른 스타일과 연관된 외형 관련 특징 표현을 선택할 수 있으며, 이에 따라 상이한 스타일의 출력을 획득한다. 디코더(220-2)의 이런 속성은 유연한 스타일 출력을 가능하게 한다. 전술한 바와 같이, 적용 동안, 트레이닝 이미지에 기초하여 획득한 스타일 라이브러리로부터 전이할 스타일을 무작위로 또는 사용자의 지시를 통해 선택하는 것 외에, 사용자가 특정 스타일의 참조 이미지를 입

력하는 것도 가능하다. 참조 이미지의 외형 관련 특징 표현은 추가적인 서브네트워크에 의해 추출될 수 있으며, 그 다음에 추출된 외형 관련 특징 표현은 스타일 트랜스퍼를 수행하기 위해 디코더(220-2)로 제공된다.

[0120] 스타일 트랜스퍼를 위한 학습 네트워크의 제2 예시적인 트레이닝

[0121] 도 4a는 위의 도 2a에서 사용된 제2 학습 네트워크(220)를 트레이닝하는 방법을 설명하고, 이런 트레이닝을 통해 획득된 제2 학습 네트워크(220)는 원본 사진에 스타일 트랜스퍼를 수행하여 캐리커처 스타일의 사진을 생성하는 방법을 학습할 수 있다. 도 2b의 구현예에서와 같이 본 명세서에 기술된 발명대상의 일부 다른 구현예에서, 제2 학습 네트워크(220)는 변형된 원본 사진에 대해 스타일 트랜스퍼를 수행하여 캐리커처 스타일의 캐리커처를 생성하는 방법을 학습할 수 있는 것이 바람직하다.

[0122] 제2 학습 네트워크(220)는 여전히 다대다 매핑을 구현하며 비지도 학습을 수행할 수 있는 GAN 네트워크에 기초할 수 있다. 트레이닝 데이터를 고려하여, 트레이닝 이미지의 기하학적 차이의 영향을 피하기 위해, 트레이닝할 포토 도메인 X의 제1 이미지는 기하학적 변형에 의해 중간 도메인의 중간 이미지(X'로 표현됨)로 변환된다. 중간 도메인 X'는 사진 도메인 X와 캐리커처 도메인 Y 사이에 있으며, 중간 이미지는 사진 도메인 X의 제1 이미지와 동일한 스타일을 상속하지만, 캐리커처 도메인 Y의 객체와 유사한 기하구조를 갖는다. 중간 도메인 X'의 중간 이미지는 제1 이미지에 대해 포토 도메인 X로부터 캐리커처 도메인 Y로의 기하학적 변형을 수행함으로써 생성될 수 있다. 일부 구현예에서, 도메인 Y로부터 도메인 X로의 이런 매핑은 트레이닝된 제1 학습 네트워크(210)에 의해 수행될 수 있다.

[0123] 제2 학습 네트워크(220)는 중간 도메인 X'의 중간 이미지와 캐리커처 도메인 Y의 제2 이미지에 기초하여 트레이닝되며, 따라서 제2 학습 네트워크(220)는 중간 도메인 X'로부터 캐리커처 도메인 Y로 스타일 트랜스퍼를 수행하는 능력을 학습할 수 있다. 중간 도메인 X'의 스타일은 사진 도메인 X와 동일하므로, 이것은 사진 도메인 X로부터 캐리커처 도메인 Y로 스타일 트랜스퍼를 구현하는 제2 학습 네트워크(220)와 동일하다.

[0124] 도 4b는 GAN 기반 구현예에서 제2 학습 네트워크(220)를 트레이닝하기 위한 GAN 구조(450)를 도시한다. 제2 학습 네트워크(220)의 트레이닝은, 도 3 및 4a와 유사하게, 반대의 양방향 경로로 구성된 순환 일관성 GAN 구조에 기초할 수 있다. 설명을 쉽게 하기 위해, 도메인 X'로부터 도메인 Y로의 방향으로의 제1 경로만 도 4b에 도시되어 있다. 다른 방향의 제2 경로는 도 3의 제2 경로와 유사하게 트레이닝을 위해 대칭으로 구성될 수 있다.

[0125] 특정 트레이닝 동안, GAN 아키텍처(450) 및 그 구성에 포함된 구성요소는 도 4a의 GAN 아키텍처(400)와 유사하다. 그러나, 구성요소들을 트레이닝하기 위한 입력은 다르다. 도시된 바와 같이, GAN 구조(450)는 도메인 X'에 대한 콘텐츠 인코더(460)($E_{X'}^s$ 로 표현됨), 스타일 인코더(220-1)($E_{X'}^s$ 로 표현됨) 및 도메인 Y'에 대한 디코더(220-2)로 구성된 순방향 생성기 부분을 포함한다. 스타일 인코더(220-1) 및 디코더(220-2)는 제2 학습 서브네트워크(220)를 형성한다. GAN 네트워크(450)는 도메인 Y'에 대한 콘텐츠 인코더(470)(E_Y^c 로 표현됨), 스타일 인코더(480)(E_Y^s 로 표현됨), 및 도메인 X'에 대한 디코더(490)($R_{X'}$)로 이루어진 역방향 생성기 부분을 더 포함한다. GAN 네트워크(450)는 판별기(482)를 더 포함할 수 있다. 이들 구성요소는 도 4a에 유사하게 배열된 구성요소들과 유사한 방식으로 동작한다.

[0126] 특정 트레이닝 동안, GAN 아키텍처(450) 및 그 구성에 포함된 구성요소는 도 4a의 GAN 아키텍처(400)와 유사하다. 그러나, 이들 구성요소에 대한 입력들은 상이하다. 예를 들어, 도 4b에 도시된 제1 경로에서, 콘텐츠 인코더(220-1)는 입력 중간 도메인 X'의 중간 이미지(452)로부터 콘텐츠 관련 특징 표현(453)(도메인 X'의 콘텐츠 코드라고도 함)을 추출한다. 복수의 중간 이미지에 의해 추출된 콘텐츠 코드는 콘텐츠 라이브러리를 형성한다. 스타일 인코더(460)는 중간 이미지(452)로부터 외형 관련 특징 표현(스타일 코드라고도 함)을 추출한다. 복수의 중간 이미지로부터 추출된 외형 관련 특징 표현은 스타일 라이브러리(예컨대, 도메인 X'의 경우 스타일 라이브러리($S_{X'}$)(451))를 형성한다. 디코더(490)는 콘텐츠 관련 특징 표현(453) 및 스타일 라이브러리($S_{X'}$)(451)로부터 선택된 외형 관련 특징 표현에 기초하여 이미지 재구성을 수행하여 예측 이미지(456)를 획득한다. 디코더(220-2)는 콘텐츠 관련 특징 표현(453) 및 도메인 Y에 대한 스타일 라이브러리(S_Y)(455)로부터 선택된 외형 관련 특징 표현에 기초하여 이미지 재구성을 수행하여 재구성된 이미지(454)(때로는 제1 예측 이미지라고도 함)를 획득한다.

[0127] 콘텐츠 인코더(470)는 재구성된 이미지(454)로부터 콘텐츠 관련 특징 표현(472)을 추출하고, 디코더(490)는 콘

텐즈 관련 특징 표현(472) 및 스타일 라이브러리(S_X)(451)로부터 에서 선택된 외형 관련 특징 표현에 기초하여 이미지 재구성을 수행하여, 예측 이미지(458)를 획득한다. 스타일 인코더(480)는 재구성된 이미지(454)로부터 외형 관련 특징 표현(스타일 코드라고도 함)을 추출하는 데 사용된다. 복수의 이러한 재구성된 이미지로부터 추출된 스타일 코드는 스타일 라이브러리(즉, 도메인 Y에 대한 스타일 라이브러리(S_Y)(455))를 형성할 수 있다.

[0128] 이들 구성요소는 상이한 트레이닝 이미지로 트레이닝되지만, 손실 함수의 유형과 같이 트레이닝에 사용되는 트레이닝 알고리즘은 도 4a에서 사용된 것과 유사할 수 있다. 트레이닝에서 고려되는 손실 함수는, 위 수학식 5 (여기서, $I=X'$ 또는 Y)와 같은 재구성 손실을 포함한다. 트레이닝 동안 고려되는 손실 함수는 또한 수학식 6에서 도메인 Y로부터 도메인 X'으로의 순환 일관성 손실 또는 도메인 X'으로부터 도메인 Y로의 역방향으로의 순환 일관성 손실을 포함할 수 있다.

[0129] 또한, 손실 함수는 지각 손실을 더 포함할 수 있다. 도 4b의 아키텍처에서, 제1 경로의 지각 손실은 인코더(220-1) 및 디코더(220-2)에 의한 스타일 트랜스퍼 후의 제1 예측 이미지(454)와 입력 중간 이미지(452) 사이의 콘텐츠 차이를 나타낸다. 이런 차이는, 수학식 7에 기술한 바와 같이, 제1 예측 이미지(404)와 제1 이미지(402)로부터 제각기 추출된 특징 맵들 사이의 차이, 특히 콘텐츠 관련 특징 맵들 사이의 차이에 기초하여 표현될 수 있다. 트레이닝 프로세스는 지각 차이를 지속적으로 줄인다. 지각 손실이 반대의 제2 경로에서 고려되는 경우, 지각 손실은 인코더(220-1) 및 디코더(220-2)에 의한 도메인 Y의 제2 이미지에 대한 스타일 트랜스퍼 후의 예측 이미지(이하 제2 예측 이미지라고 함)와 제2 이미지 사이의 콘텐츠 차이의 감소를 나타낸다. 지각 손실의 결정은 또한 도 4a의 아키텍처에서 고려된 지각 손실과 유사하며 여기에서는 자세히 설명하지 않는다.

[0130] 도 4b의 구조에서 알 수 있는 바와 같이, 제2 학습 네트워크(220)를 형성하는 디코더(220-2)는 중간 이미지의 동일한 콘텐츠 관련 특징 표현에 대한 전이를 수행하기 위해 다른 스타일과 연관된 외형 관련 특징 표현을 선택할 수 있으며, 이에 따라 상이한 스타일의 출력을 획득할 수 있다. 디코더(220-2)의 이런 속성은 유연한 스타일 출력을 가능하게 한다. 전술한 바와 같이, 적용 동안, 트레이닝 이미지에 기초하여 획득한 스타일 라이브러리로부터 전이할 스타일을 무작위로 또는 사용자의 지시를 통해 선택하는 것 외에, 사용자가 특정 스타일의 참조 이미지를 입력하는 것도 가능하다. 참조 이미지의 외형 관련 특징 표현은 추가적인 서브네트워크에 의해 추출될 수 있으며, 그 다음에 추출된 외형 관련 특징 표현은 스타일 트랜스퍼를 수행하기 위해 디코더(220-2)로 제공된다.

[0131] 예시적인 프로세스

[0132] 도 5a는 본 명세서에 기술된 발명대상의 일부 구현예에 따른 학습 네트워크를 트레이닝하는 프로세스(500)의 흐름도이다. 프로세스(500)는 컴퓨팅 장치(100) 또는 컴퓨팅 장치(100)와 유사한 다른 장치 또는 장치 클러스터에 의해 구현될 수 있다.

[0133] 블록 510에서, 제1 이미지 도메인의 제1 이미지 및 제2 이미지 도메인의 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크가 결정된다. 제1 및 제2 이미지 도메인의 이미지는 서로 다른 스타일을 가지며, 이들 이미지 내 객체들은 서로에 대해 기하학적으로 변형된다. 블록 520에서, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형이 수행되어, 중간 이미지를 생성하며, 중간 이미지는 제2 이미지와 동일한 스타일을 상속한다. 블록 530에서, 제1 이미지 및 중간 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 스타일 트랜스퍼를 위한 제2 학습 네트워크가 결정된다.

[0134] 일부 구현예에서, 제1 학습 네트워크를 결정하는 것은, 제1 이미지 내의 객체의 기하구조의 제1 랜드마크 포인트 및 제2 이미지 내의 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 것과, 제1 랜드마크 포인트의 제1 주성분 분석(PCA) 표현 및 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 것과, 제1 및 제2 PCA 표현에 기초하여 제1 학습 네트워크를 결정하는 것을 포함한다.

[0135] 일부 구현예에서, 제1 학습 네트워크는 제1 학습 네트워크는 제1 생성적 적대 네트워크(GAN)에 기초하고, 제1 학습 네트워크를 결정하는 것은, 제1 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 제1 이미지 내의 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 것과, 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 것 - 제1 평균 기하구조는 제1 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균이고, 제2 평균 기하구조는 제2 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균임 - 과, 제1 및 제2 차이 사이의 차가 감소하도록 제1 GAN의 제1 서브네트워크를 업데이트하는

것을 포함하되, 업데이트된 제1 서브네트워크는 제1 학습 네트워크로서 결정된다.

- [0136] 일부 구현예에서, 제1 GAN의 제1 서브네트워크를 업데이트하는 것은, 제1 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 도메인으로 기하학적 변형을 수행하여, 제2 이미지 내의 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 것과, 제2 기하구조와 제2 평균 기하구조 사이의 제3 차이 및 제2 변형된 기하구조와 제1 평균 기하구조 사이의 제4 차이를 결정하는 것과, 제3 및 제4 차이 사이의 차가 또한 감소하도록 제1 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.
- [0137] 일부 구현예에서, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하는 것은 업데이트된 제2 서브네트워크를 사용하여 제2 이미지 도메인으로부터 제1 이미지 도메인으로 기하학적 변형을 수행하는 것을 포함한다.
- [0138] 일부 구현예에서, 제2 학습 네트워크는 제2 생성적 적대 네트워크(GAN)에 기초하며, 여기서 제2 학습 네트워크를 결정하는 것은, 제2 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로 스타일 트랜스퍼를 수행하여, 제1 스타일의 제1 이미지를 제2 스타일의 제1 예측 이미지로 전이하는 것 - 제2 스타일은 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제1 예측 이미지와 제1 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하며, 업데이트된 제1 서브네트워크는 제2 학습 네트워크로서 결정된다.
- [0139] 일부 구현예에서, 제2 GAN의 제1 서브네트워크를 업데이트하는 것은, 제2 GAN의 제2 서브네트워크를 사용하여, 중간 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 중간 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 것 - 제4 스타일은 제1 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제2 예측 이미지와 중간 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 포함한다.
- [0140] 도 5b는 본 명세서에 기술된 발명대상의 일부 구현예에 따른 학습 네트워크를 트레이닝하는 프로세스(550)의 흐름도이다. 프로세스(550)는 컴퓨팅 장치(100) 또는 컴퓨팅 장치(100)와 유사한 다른 장치 또는 장치 클러스터에 의해 구현될 수 있다.
- [0141] 블록 560에서, 제1 이미지 도메인의 제1 이미지 및 제2 이미지 도메인의 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크가 결정된다. 제1 및 제2 이미지 도메인의 이미지는 서로 다른 스타일을 가지며, 이미지 내 객체들은 서로에 대해 기하학적으로 변형된다. 블록 570에서, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형이 수행되어 중간 이미지를 생성하며, 중간 이미지는 제1 이미지와 동일한 스타일을 상속한다. 블록 589에서, 중간 이미지 및 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 스타일 트랜스퍼를 위한 제2 학습 네트워크가 결정된다.
- [0142] 일부 구현예에서, 제1 학습 네트워크를 결정하는 것은, 제1 이미지 내 객체의 기하구조의 제1 랜드마크 포인트 및 제2 이미지 내 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 것과, 제1 랜드마크 포인트의 제1 주성분 분석(PCA) 표현 및 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 것과, 제1 및 제2 PCA 표현에 기초하여 제1 학습 네트워크를 결정하는 것을 포함한다.
- [0143] 일부 구현예에서, 제1 학습 네트워크는 생성적 적대 신경망(GAN)에 기초한다. 제1 학습 네트워크를 결정하는 것은, 제1 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 제1 이미지 내의 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 것과, 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 것 - 제1 평균 기하구조는 제1 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균이고, 제2 평균 기하구조는 제2 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균임 - 과, 제1 및 제2 차이 사이의 차가 감소하도록 제1 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하되, 업데이트된 제1 서브네트워크는 제1 학습 네트워크로서 결정된다.
- [0144] 일부 구현예에서, 제1 GAN의 제1 서브네트워크를 업데이트하는 것은, 제1 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 도메인으로 기하학적 변형을 수행하여, 제2 이미지 내의 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 것과, 제2 기하구조와 제2 평균 기하구조 사이의 제3 차이 및 제2 변형된 기하구조와 제1 평균 기하구조 사이의 제4 차이를 결정하는 것과, 제3 및 제4 차이 사이의 차가 감소하도록 제1 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

- [0145] 일부 구현예에서, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하는 것은 업데이트된 제2 서브네트워크를 사용하여 제2 이미지 도메인으로부터 제1 이미지 도메인으로 기하학적 변형을 수행하는 것을 포함한다.
- [0146] 일부 구현예에서, 제2 학습 네트워크는 제2 생성적 적대 신경망(GAN)에 기초한다. 제2 학습 네트워크를 결정하는 것은, 제2 GAN의 제1 서브네트워크를 사용하여, 중간 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로 스타일 트랜스퍼를 수행하여, 제1 스타일의 중간 이미지를 제2 스타일의 제1 예측 이미지로 전이하는 것 - 제2 스타일은 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제1 예측 이미지와 제1 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하며, 업데이트된 제1 서브네트워크는 제2 학습 네트워크로서 결정된다.
- [0147] 일부 구현예에서, 제2 GAN의 제1 서브네트워크를 업데이트하는 것은, 제2 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 제2 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 것 - 제4 스타일은 제1 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제2 예측 이미지와 제2 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.
- [0148] 도 6a는 본 명세서에 기술된 발명대상의 일부 구현예에 따른 교차 도메인 이미지 변환 프로세스(600)의 흐름도이다. 프로세스(600)는 컴퓨팅 장치(100)에 의해 구현될 수 있으며, 예를 들어 컴퓨팅 장치(100)의 메모리(120) 내의 이미지 변환 모듈(122)에서 구현될 수 있다.
- [0149] 블록(610)에서, 기하학적 변형이 소스 이미지에 대해 수행되어 소스 이미지 내의 객체의 제1 기하구조를 제2 기하구조로 변형시킨다. 블록 620에서, 스타일 트랜스퍼가 소스 이미지에 대해 수행되어 제1 스타일의 소스 이미지를 제2 스타일의 중간 이미지로 전이한다. 블록(630)에서, 중간 이미지 내 객체는 제2 기하구조에 기초하여 전이되어 타겟 이미지를 생성하며, 타겟 이미지는 중간 이미지와 동일한 제2 스타일을 상속한다.
- [0150] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 제1 기하구조를 표현하는 소스 이미지 내의 랜드마크 포인트를 결정하는 것과, 랜드마크 포인트의 주성분 분석(PCA) 표현을 생성하는 것과, PCA 표현에 대해 기하학적 변형을 수행하는 것과, 변형된 PCA 표현에 기초하여 제2 기하구조를 나타내는 변형된 랜드마크 포인트를 결정하는 것을 포함한다.
- [0151] 일부 구현예에서, 기하학적 변형을 수행하는 단계는, 객체의 변형 정도의 표시를 획득하는 것과, 변형 정도에 기초하여 제1 기하구조를 제2 기하구조로 전이하는 것을 포함한다.
- [0152] 일부 실시예에서, 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은, 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과, 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 중간 이미지를 생성하는 것을 포함한다.
- [0153] 일부 구현예에서, 외형 관련 특징 표현을 결정하는 것은, 복수의 미리 결정된 외형 관련 특징 표현으로부터 외형 관련 특징 표현을 무작위로 선택하는 것, 또는 제2 스타일과 관련된 획득된 표시 또는 제2 스타일의 참조 이미지에 기초하여 외형 관련 특징 표현을 결정하는 것을 포함한다.
- [0154] 일부 구현예에서, 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 기하학적 변형을 수행하는 것을 포함하고, 제1 학습 네트워크는 생성적 적대 네트워크(GAN)에 기초하여 결정된다. 일부 구현예에서, 스타일 트랜스퍼를 수행하는 것은 제2 학습 네트워크를 사용하여 스타일 트랜스퍼를 수행하는 것을 포함하고, 제2 학습 네트워크는 또 다른 생성적 적대 네트워크(GAN)에 기초하여 결정된다.
- [0155] 도 6b는 본 명세서에 기술된 발명대상의 일부 구현예에 따른 교차 도메인 이미지 변환 프로세스(650)의 흐름도이다. 프로세스(650)는 컴퓨팅 장치(100)에 의해 구현될 수 있으며, 예를 들어 컴퓨팅 장치(100)의 메모리(120) 내의 이미지 변환 모듈(122)에서 구현될 수 있다.
- [0156] 블록(660)에서, 기하학적 변형이 소스 이미지에 대해 수행되어 소스 이미지 내의 객체의 제1 기하구조를 제2 기하구조로 변형시킨다. 블록 670에서, 소스 이미지 내 객체는 제2 기하구조에 기초하여 전이되어 중간 이미지를 생성하며, 중간 이미지는 소스 이미지와 동일한 제1 스타일을 상속한다. 블록 680에서, 스타일 트랜스퍼가 중간 이미지에 대해 수행되어 제1 스타일의 중간 이미지를 제2 스타일의 타겟 이미지로 전이한다.
- [0157] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 제1 기하구조를 표현하는 소스 이미지 내의 랜드마크 포인트를 결정하는 것과, 랜드마크 포인트의 주성분 분석(PCA) 표현을 생성하는 것과, PCA 표현에 대해 기하학적 변형

을 수행하는 것과, 변형된 PCA 표현에 기초하여 제2 기하구조를 표현하는 변형된 랜드마크 포인트를 결정하는 것을 포함한다.

[0158] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 객체의 변형 정도의 표시를 획득하는 것과, 변형 정도에 기초하여 제1 기하구조를 제2 기하구조로 전이하는 것을 포함한다.

[0159] 일부 실시예에서, 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은, 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과, 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 중간 이미지를 생성하는 것을 포함한다.

[0160] 일부 구현예에서, 외형 관련 특징 표현을 결정하는 것은, 복수의 미리 결정된 외형 관련 특징 표현으로부터 외형 관련 특징 표현을 무작위로 선택하는 것, 또는 제2 스타일과 관련된 획득된 표시 또는 제2 스타일의 참조 이미지에 기초하여 외형 관련 특징 표현을 결정하는 것을 포함한다.

[0161] 일부 구현예에서, 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 기하학적 변형을 수행하는 것을 포함하고, 제1 학습 네트워크는 생성적 적대 네트워크(GAN)에 기초하여 결정된다. 일부 구현예에서, 스타일 트랜스퍼를 수행하는 것은 제2 학습 네트워크를 사용하여 스타일 트랜스퍼를 수행하는 것을 포함하고, 제2 학습 네트워크는 또 다른 생성적 적대 네트워크(GAN)에 기초하여 결정된다.

[0162] 구현예들

[0163] 본 명세서에 기술된 발명대상의 몇몇 구현예들을 아래에 나열한다.

[0164] 한 양태에서, 본 명세서에 기술된 발명대상은 컴퓨터로 구현되는 방법으로서, 제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크를 결정하는 것 - 제1 이미지 도메인 및 제2 이미지 도메인의 이미지들은 상이한 스타일을 가지며, 이미지들 내의 객체들은 서로에 대해 기하학적 변형을 가짐 - 과, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하여 중간 이미지를 생성하는 것 - 중간 이미지는 제2 이미지와 동일한 스타일을 상속함 - 과, 제1 이미지 및 중간 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 스타일 트랜스퍼를 위한 제2 학습 네트워크를 결정하는 것을 포함하는, 방법을 제공한다.

[0165] 일부 구현예에서, 제1 학습 네트워크를 결정하는 것은, 제1 이미지 내 객체의 기하구조의 제1 랜드마크 포인트 및 제2 이미지 내 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 것과, 제1 랜드마크 포인트의 제1 주성분 분석(PCA) 표현 및 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 것과, 제1 및 제2 PCA 표현에 기초하여 제1 학습 네트워크를 결정하는 것을 포함한다.

[0166] 일부 구현예에서, 제1 학습 네트워크는 제1 학습 네트워크는 제1 생성적 적대 네트워크(GAN)에 기초하고, 제1 학습 네트워크를 결정하는 것은, 제1 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 제1 이미지 내의 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 것과, 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 것 - 제1 평균 기하구조는 제1 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균이고, 제2 평균 기하구조는 제2 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균임 - 과, 제1 및 제2 차이 사이의 차가 감소하도록 제1 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하되, 업데이트된 제1 서브네트워크는 제1 학습 네트워크로서 결정된다.

[0167] 일부 구현예에서, 제1 GAN의 제1 서브네트워크를 업데이트하는 것은, 제1 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 도메인으로 기하학적 변형을 수행하여, 제2 이미지 내의 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 것과, 제2 기하구조와 제2 평균 기하구조 사이의 제3 차이 및 제2 변형된 기하구조와 제1 평균 기하구조 사이의 제4 차이를 결정하는 것과, 제3 및 제4 차이 사이의 차가 또한 감소하도록 제1 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

[0168] 일부 구현예에서, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하는 것은 업데이트된 제2 서브네트워크를 사용하여 제2 이미지 도메인으로부터 제1 이미지 도메인으로 기하학적 변형을 수행하는 것을 포함한다.

[0169] 일부 구현예에서, 제2 학습 네트워크는 제2 생성적 적대 네트워크(GAN)에 기초하며, 여기서 제2 학습 네트워크를 결정하는 것은, 제2 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2

이미지 도메인으로 스타일 트랜스퍼를 수행하여, 제1 스타일의 제1 이미지를 제2 스타일의 제1 예측 이미지로 전이하는 것 - 제2 스타일은 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제1 예측 이미지와 제1 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하며, 업데이트된 제1 서브네트워크는 제2 학습 네트워크로서 결정된다.

[0170] 일부 구현예에서, 제2 GAN의 제1 서브네트워크를 업데이트하는 것은, 제2 GAN의 제2 서브네트워크를 사용하여, 중간 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 중간 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 것 - 제4 스타일은 제1 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제2 예측 이미지와 중간 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 포함한다.

[0171] 제2 양태에서, 본 명세서에 기술된 발명대상은 전자 장치를 제공한다. 전자 장치는 처리 유닛과, 처리 유닛에 연결되며 명령어가 저장되어 있는 메모리를 포함하되, 이들 명령어는 처리 유닛에 의해 실행될 경우에, 장치로 하여금, 제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크를 결정하는 동작 - 제1 이미지 도메인 및 제2 이미지 도메인의 이미지들은 상이한 스타일을 가지며, 이미지들 내의 객체들은 서로에 대해 기하학적 변형을 가짐 - 과, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하여 중간 이미지를 생성하는 동작 - 중간 이미지는 제2 이미지와 동일한 스타일을 상속함 - 와, 제1 이미지 및 중간 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 스타일 트랜스퍼를 위한 제2 학습 네트워크를 결정하는 동작을 수행하게 한다.

[0172] 일부 구현예에서, 제1 학습 네트워크를 결정하는 것은, 제1 이미지 내 객체의 기하구조의 제1 랜드마크 포인트 및 제2 이미지 내 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 것과, 제1 랜드마크 포인트의 제1 주성분 분석(PCA) 표현 및 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 것과, 제1 및 제2 PCA 표현에 기초하여 제1 학습 네트워크를 결정하는 것을 포함한다.

[0173] 일부 구현예에서, 제1 학습 네트워크는 제1 생성적 적대 네트워크(GAN)에 기초하고, 제1 학습 네트워크를 결정하는 것은, 제1 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 제1 이미지 내의 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 것과, 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 것 - 제1 평균 기하구조는 제1 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균이고, 제2 평균 기하구조는 제2 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균임 - 과, 제1 및 제2 차이 사이의 차가 감소하도록 제1 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하되, 업데이트된 제1 서브네트워크는 제1 학습 네트워크로서 결정된다.

[0174] 일부 구현예에서, 제1 GAN의 제1 서브네트워크를 업데이트하는 것은, 제1 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 도메인으로 기하학적 변형을 수행하여, 제2 이미지 내의 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 것과, 제2 기하구조와 제2 평균 기하구조 사이의 제3 차이 및 제2 변형된 기하구조와 제1 평균 기하구조 사이의 제4 차이를 결정하는 것과, 제3 및 제4 차이 사이의 차가 또한 감소하도록 제1 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

[0175] 일부 구현예에서, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하는 것은 업데이트된 제2 서브네트워크를 사용하여 제2 이미지 도메인으로부터 제1 이미지 도메인으로 기하학적 변형을 수행하는 것을 포함한다.

[0176] 일부 구현예에서, 제2 학습 네트워크는 제2 생성적 적대 네트워크(GAN)에 기초하며, 여기서 제2 학습 네트워크를 결정하는 것은, 제2 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로 스타일 트랜스퍼를 수행하여, 제1 스타일의 제1 이미지를 제2 스타일의 제1 예측 이미지로 전이하는 것 - 제2 스타일은 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제1 예측 이미지와 제1 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하며, 업데이트된 제1 서브네트워크는 제2 학습 네트워크로서 결정된다.

[0177] 일부 구현예에서, 제2 GAN의 제1 서브네트워크를 업데이트하는 것은, 제2 GAN의 제2 서브네트워크를 사용하여, 중간 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 중간 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 것 - 제4 스타일은 제1 이미지 도메인의 복수

의 스타일 중에서 무작위로 선택됨 - 과, 제2 예측 이미지와 중간 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 포함한다.

[0178] 제3 양태에서, 본 명세서에 기술된 발명대상은 컴퓨터로 구현된 방법을 제공한다. 이 방법은, 제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크를 결정하는 것 - 제1 이미지 도메인 및 제2 이미지 도메인의 이미지들은 상이한 스타일을 가지며, 이미지들 내의 객체들은 서로에 대해 기하학적 변형을 가짐 - 과, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 중간 이미지를 생성하는 것 - 중간 이미지는 제1 이미지와 동일한 스타일을 상속함 - 과, 중간 이미지 및 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 스타일 트랜스퍼를 위한 제2 학습 네트워크를 결정하는 것을 포함하는, 방법을 제공한다.

[0179] 일부 구현예에서, 제1 학습 네트워크를 결정하는 것은, 제1 이미지 내 객체의 기하구조의 제1 랜드마크 포인트 및 제2 이미지 내 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 것과, 제1 랜드마크 포인트의 제1 주성분 분석(PCA) 표현 및 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 것과, 제1 및 제2 PCA 표현에 기초하여 제1 학습 네트워크를 결정하는 것을 포함한다.

[0180] 일부 구현에서, 제1 학습 네트워크는 제1 생성적 적대 신경망(GAN)에 기초한다. 제1 학습 네트워크를 결정하는 것은, 제1 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 제1 이미지 내의 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 것과, 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 것 - 제1 평균 기하구조는 제1 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균이고, 제2 평균 기하구조는 제2 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균임 - 과, 제1 및 제2 차이 사이의 차가 감소하도록 제1 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하되, 업데이트된 제1 서브네트워크는 제1 학습 네트워크로서 결정된다.

[0181] 일부 구현예에서, 제1 GAN의 제1 서브네트워크를 업데이트하는 것은, 제1 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 도메인으로 기하학적 변형을 수행하여, 제2 이미지 내의 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 것과, 제2 기하구조와 제2 평균 기하구조 사이의 제3 차이 및 제2 변형된 기하구조와 제1 평균 기하구조 사이의 제4 차이를 결정하는 것과, 제3 및 제4 차이 사이의 차가 감소하도록 제1 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

[0182] 일부 구현예에서, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하는 것은 업데이트된 제2 서브네트워크를 사용하여 제2 이미지 도메인으로부터 제1 이미지 도메인으로 기하학적 변형을 수행하는 것을 포함한다.

[0183] 일부 구현예에서, 제2 학습 네트워크는 제2 생성적 적대 신경망(GAN)에 기초한다. 제2 학습 네트워크를 결정하는 것은, 제2 GAN의 제1 서브네트워크를 사용하여, 중간 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로 스타일 트랜스퍼를 수행하여, 제1 스타일의 중간 이미지를 제2 스타일의 제1 예측 이미지로 전이하는 것 - 제2 스타일은 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제1 예측 이미지와 제1 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하며, 업데이트된 제1 서브네트워크는 제2 학습 네트워크로서 결정된다.

[0184] 일부 구현예에서, 제2 GAN의 제1 서브네트워크를 업데이트하는 것은, 제2 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 제2 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 것 - 제4 스타일은 제1 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제2 예측 이미지와 제2 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

[0185] 제4 양태에서, 본 명세서에 기술된 발명대상은 전자 장치를 제공한다. 전자 장치는 처리 유닛과, 처리 유닛에 연결되며 명령어가 저장되어 있는 메모리를 포함하되, 이들 명령어는 처리 유닛에 의해 실행될 경우에, 장치로 하여금, 제1 이미지 도메인의 제1 이미지와 제2 이미지 도메인의 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 위한 제1 학습 네트워크를 결정하는 동작 - 제1 이미지 도메인 및 제2 이미지 도메인의 이미지들은 상이한 스타일을 가지며, 이미지들 내의 객체들은 서로에 대해 기하학적 변형을 가짐 - 과, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행

하여 중간 이미지를 생성하는 동작 - 중간 이미지는 제1 이미지와 동일한 스타일을 상속함 - 와, 중간 이미지 및 제2 이미지에 기초하여 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 스타일 트랜스퍼를 위한 제2 학습 네트워크를 결정하는 동작을 수행하게 한다.

[0186] 일부 구현예에서, 제1 학습 네트워크를 결정하는 것은, 제1 이미지 내 객체의 기하구조의 제1 랜드마크 포인트 및 제2 이미지 내 객체의 기하구조의 제2 랜드마크 포인트를 추출하는 것과, 제1 랜드마크 포인트의 제1 주성분 분석(PCA) 표현 및 제2 랜드마크 포인트의 제2 PCA 표현을 결정하는 것과, 제1 및 제2 PCA 표현에 기초하여 제1 학습 네트워크를 결정하는 것을 포함한다.

[0187] 일부 구현예에서, 제1 학습 네트워크는 생성적 적대 신경망(GAN)에 기초한다. 제1 학습 네트워크를 결정하는 것은, 제1 GAN의 제1 서브네트워크를 사용하여, 제1 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로의 기하학적 변형을 수행하여 제1 이미지 내의 객체의 제1 기하구조를 제1 변형된 기하구조로 변형시키는 것과, 제1 기하구조와 제1 평균 기하구조 사이의 제1 차이와 제1 변형된 기하구조와 제2 평균 기하구조 사이의 제2 차이를 결정하는 것 - 제1 평균 기하구조는 제1 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균이고, 제2 평균 기하구조는 제2 이미지 도메인의 복수의 이미지 내의 객체의 기하구조의 평균임 - 과, 제1 및 제2 차이 사이의 차가 감소하도록 제1 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하되, 업데이트된 제1 서브네트워크는 제1 학습 네트워크로서 결정된다.

[0188] 일부 구현예에서, 제1 GAN의 제1 서브네트워크를 업데이트하는 것은, 제1 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 도메인으로 기하학적 변형을 수행하여, 제2 이미지 내의 객체의 제2 기하구조를 제2 변형된 기하구조로 변형시키는 것과, 제2 기하구조와 제2 평균 기하구조 사이의 제3 차이 및 제2 변형된 기하구조와 제1 평균 기하구조 사이의 제4 차이를 결정하는 것과, 제3 및 제4 차이 사이의 차가 감소하도록 제1 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

[0189] 일부 구현예에서, 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 기하학적 변형을 수행하는 것은 업데이트된 제2 서브네트워크를 사용하여 제2 이미지 도메인으로부터 제1 이미지 도메인으로 기하학적 변형을 수행하는 것을 포함한다.

[0190] 일부 구현예에서, 제2 학습 네트워크는 제2 생성적 적대 신경망(GAN)에 기초한다. 제2 학습 네트워크를 결정하는 것은, 제2 GAN의 제1 서브네트워크를 사용하여, 중간 이미지에 대해 제1 이미지 도메인으로부터 제2 이미지 도메인으로 스타일 트랜스퍼를 수행하여, 제1 스타일의 중간 이미지를 제2 스타일의 제1 예측 이미지로 전이하는 것 - 제2 스타일은 제2 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제1 예측 이미지와 제1 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 서브네트워크를 업데이트하는 것을 포함하며, 업데이트된 제1 서브네트워크는 제2 학습 네트워크로서 결정된다.

[0191] 일부 구현예에서, 제2 GAN의 제1 서브네트워크를 업데이트하는 것은, 제2 GAN의 제2 서브네트워크를 사용하여, 제2 이미지에 대해 제2 이미지 도메인으로부터 제1 이미지 도메인으로의 스타일 트랜스퍼를 수행하여, 제3 스타일의 제2 이미지를 제4 스타일의 제2 예측 이미지로 전이하는 것 - 제4 스타일은 제1 이미지 도메인의 복수의 스타일 중에서 무작위로 선택됨 - 과, 제2 예측 이미지와 제2 이미지 사이의 콘텐츠 차이가 감소하도록 제2 GAN의 제1 및 제2 서브네트워크를 함께 업데이트하는 것을 더 포함한다.

[0192] 제5 양태에서, 본 명세서에 기술된 발명대상은 컴퓨터로 구현된 방법을 제공한다. 이 방법은, 소스 이미지에 대해 기하학적 변형을 수행하여 소스 이미지 내 객체의 제1 기하구조를 제2 기하구조로 변형시키는 것과, 소스 이미지에 대해 스타일 트랜스퍼를 수행하여 제1 스타일의 소스 이미지를 제2 스타일의 중간 이미지로 전이하는 것과, 제2 기하구조에 기초하여 중간 이미지 내 객체를 전이시켜 타겟 이미지를 생성하는 것을 포함하되, 타겟 이미지는 중간 이미지와 동일한 제2 스타일을 상속한다.

[0193] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 제1 기하구조를 표현하는 소스 이미지 내의 랜드마크 포인트를 결정하는 것과, 랜드마크 포인트의 주성분 분석(PCA) 표현을 생성하는 것과, PCA 표현에 대해 기하학적 변형을 수행하는 것과, 변형된 PCA 표현에 기초하여 제2 기하구조를 나타내는 변형된 랜드마크 포인트를 결정하는 것을 포함한다.

[0194] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 객체의 변형 정도의 표시를 획득하는 것과, 변형 정도에 기초하여 제1 기하구조를 제2 기하구조로 전이하는 것을 포함한다.

[0195] 일부 실시예에서, 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은, 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과, 콘텐츠 관련 특징 표현 및 외형

관련 특징 표현에 기초하여 중간 이미지를 생성하는 것을 포함한다.

- [0196] 일부 구현예에서, 외형 관련 특징 표현을 결정하는 것은, 복수의 미리 결정된 외형 관련 특징 표현으로부터 외형 관련 특징 표현을 무작위로 선택하는 것, 또는 제2 스타일과 관련된 획득된 표시 또는 제2 스타일의 참조 이미지에 기초하여 외형 관련 특징 표현을 결정하는 것을 포함한다.
- [0197] 일부 구현예에서, 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 기하학적 변형을 수행하는 것을 포함하고, 제1 학습 네트워크는 생성적 적대 네트워크(GAN)에 기초하여 결정된다. 일부 구현예에서, 스타일 트랜스퍼를 수행하는 것은 제2 학습 네트워크를 사용하여 스타일 트랜스퍼를 수행하는 것을 포함하고, 제2 학습 네트워크는 또 다른 생성적 적대 네트워크(GAN)에 기초하여 결정된다.
- [0198] 제6 양태에서, 본 명세서에 기술된 발명대상은 전자 장치를 제공한다. 이 전자 장치는, 처리 유닛과, 처리 유닛에 연결되며 명령어가 저장되어 있는 메모리를 포함하되, 이들 명령어는 처리 유닛에 의해 실행될 경우에, 장치로 하여금, 소스 이미지에 대해 기하학적 변형을 수행하여 소스 이미지 내 객체의 제1 기하구조를 제2 기하구조로 변형시키는 것과, 소스 이미지에 대해 스타일 트랜스퍼를 수행하여 제1 스타일의 소스 이미지를 제2 스타일의 중간 이미지로 전이하는 것과, 제2 기하구조에 기초하여 중간 이미지 내 객체를 전이시켜 타겟 이미지를 생성하는 것을 포함하되, 타겟 이미지는 중간 이미지와 동일한 제2 스타일을 상속한다.
- [0199] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 제1 기하구조를 표현하는 소스 이미지 내의 랜드마크 포인트를 결정하는 것과, 랜드마크 포인트의 주성분 분석(PCA) 표현을 생성하는 것과, PCA 표현에 대해 기하학적 변형을 수행하는 것과, 변형된 PCA 표현에 기초하여 제2 기하구조를 나타내는 변형된 랜드마크 포인트를 결정하는 것을 포함한다.
- [0200] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 객체의 변형 정도의 표시를 획득하는 것과, 변형 정도에 기초하여 제1 기하구조를 제2 기하구조로 전이하는 것을 포함한다.
- [0201] 일부 실시예에서, 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은, 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과, 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 중간 이미지를 생성하는 것을 포함한다.
- [0202] 일부 구현예에서, 외형 관련 특징 표현을 결정하는 것은, 복수의 미리 결정된 외형 관련 특징 표현으로부터 외형 관련 특징 표현을 무작위로 선택하는 것, 또는 제2 스타일과 관련된 획득된 표시 또는 제2 스타일의 참조 이미지에 기초하여 외형 관련 특징 표현을 결정하는 것을 포함한다.
- [0203] 일부 구현예에서, 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 기하학적 변형을 수행하는 것을 포함하고, 제1 학습 네트워크는 생성적 적대 네트워크(GAN)에 기초하여 결정된다. 일부 구현예에서, 상기 스타일 트랜스퍼를 수행하는 것은, 제2 학습 네트워크를 사용하여 스타일 트랜스퍼를 수행하는 것을 포함하고, 제2 학습 네트워크는 제2 GAN에 기초하여 결정된다.
- [0204] 제7 양태에서, 본 명세서에 기술된 발명대상은 컴퓨터로 구현된 방법을 제공한다. 이 방법은, 소스 이미지에 대해 기하학적 변형을 수행하여 소스 이미지 내 객체의 제1 기하구조를 제2 기하구조로 변형시키는 것과, 제2 기하구조에 기초하여 소스 이미지 내 객체를 전이시켜 중간 이미지를 생성하는 것 - 중간 이미지는 소스 이미지와 동일한 제1 스타일을 상속함 - 과, 중간 이미지에 대해 스타일 트랜스퍼를 수행하여 제1 스타일의 중간 이미지를 제2 스타일의 타겟 이미지로 전이하는 것을 포함한다.
- [0205] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 제1 기하구조를 표현하는 소스 이미지 내의 랜드마크 포인트를 결정하는 것과, 랜드마크 포인트의 주성분 분석(PCA) 표현을 생성하는 것과, PCA 표현에 대해 기하학적 변형을 수행하는 것과, 변형된 PCA 표현에 기초하여 제2 기하구조를 나타내는 변형된 랜드마크 포인트를 결정하는 것을 포함한다.
- [0206] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 객체의 변형 정도의 표시를 획득하는 것과, 변형 정도에 기초하여 제1 기하구조를 제2 기하구조로 전이하는 것을 포함한다.
- [0207] 일부 실시예에서, 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은, 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과, 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 중간 이미지를 생성하는 것을 포함한다.
- [0208] 일부 구현예에서, 외형 관련 특징 표현을 결정하는 것은, 복수의 미리 결정된 외형 관련 특징 표현으로부터 외

형 관련 특징 표현을 무작위로 선택하는 것, 또는 제2 스타일과 관련된 획득된 표시 또는 제2 스타일의 참조 이미지에 기초하여 외형 관련 특징 표현을 결정하는 것을 포함한다.

- [0209] 일부 구현예에서, 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 기하학적 변형을 수행하는 것을 포함하고, 제1 학습 네트워크는 생성적 적대 네트워크(GAN)에 기초하여 결정된다. 일부 구현예에서, 스타일 트랜스퍼를 수행하는 것은 제2 학습 네트워크를 사용하여 스타일 트랜스퍼를 수행하는 것을 포함하고, 제2 학습 네트워크는 또 다른 생성적 적대 네트워크(GAN)에 기초하여 결정된다.
- [0210] 제8 양태에서, 본 명세서에 기술된 발명대상은 전자 장치를 제공한다. 이 전자 장치는, 처리 유닛과, 처리 유닛에 연결되며 명령어가 저장되어 있는 메모리를 포함하되, 이들 명령어는 처리 유닛에 의해 실행될 경우에, 장치로 하여금, 소스 이미지에 대해 기하학적 변형을 수행하여 소스 이미지 내 객체의 제1 기하구조를 제2 기하구조로 변형시키는 것과, 제2 기하구조에 기초하여 소스 이미지 내 객체를 전이시켜 중간 이미지를 생성하는 것 - 중간 이미지는 소스 이미지와 동일한 제1 스타일을 상속함 - 과, 중간 이미지에 대해 스타일 트랜스퍼를 수행하여 제1 스타일의 중간 이미지를 제2 스타일의 타겟 이미지로 전이하는 것을 포함한다.
- [0211] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 제1 기하구조를 표현하는 소스 이미지 내의 랜드마크 포인트를 결정하는 것과, 랜드마크 포인트의 주성분 분석(PCA) 표현을 생성하는 것과, PCA 표현에 대해 기하학적 변형을 수행하는 것과, 변형된 PCA 표현에 기초하여 제2 기하구조를 나타내는 변형된 랜드마크 포인트를 결정하는 것을 포함한다.
- [0212] 일부 구현예에서, 기하학적 변형을 수행하는 것은, 객체의 변형 정도의 표시를 획득하는 것과, 변형 정도에 기초하여 제1 기하구조를 제2 기하구조로 전이하는 것을 포함한다.
- [0213] 일부 실시예에서, 소스 이미지에 대해 스타일 트랜스퍼를 수행하는 것은, 소스 이미지의 콘텐츠 관련 특징 표현을 추출하는 것과, 제2 스타일과 연관된 외형 관련 특징 표현을 결정하는 것과, 콘텐츠 관련 특징 표현 및 외형 관련 특징 표현에 기초하여 중간 이미지를 생성하는 것을 포함한다.
- [0214] 일부 구현예에서, 외형 관련 특징 표현을 결정하는 것은, 복수의 미리 결정된 외형 관련 특징 표현으로부터 외형 관련 특징 표현을 무작위로 선택하는 것, 또는 제2 스타일과 관련된 획득된 표시 또는 제2 스타일의 참조 이미지에 기초하여 외형 관련 특징 표현을 결정하는 것을 포함한다.
- [0215] 일부 구현예에서, 기하학적 변형을 수행하는 것은 제1 학습 네트워크를 사용하여 기하학적 변형을 수행하는 것을 포함하고, 제1 학습 네트워크는 생성적 적대 네트워크(GAN)에 기초하여 결정된다. 일부 구현예에서, 상기 스타일 트랜스퍼를 수행하는 것은, 제2 학습 네트워크를 사용하여 스타일 트랜스퍼를 수행하는 것을 포함하고, 제2 학습 네트워크는 제2 GAN에 기초하여 결정된다.
- [0216] 제10 양태에서, 본 명세서에 기술된 발명대상은, 비밀식적 컴퓨터 저장 매체에 물리적 방식으로 저장되며 머신 실행가능 명령어를 포함하는 컴퓨터 프로그램 제품을 제공하되, 이들 명령어는 장치에 의해 실행될 경우에, 전술한 양태들 중 어느 한 양태에 따른 방법을 수행하게 한다.
- [0217] 본 명세서에 기술된 기능은 적어도 부분적으로 하나 이상의 하드웨어 로직 컴포넌트에 의해 수행될 수 있다. 예를 들어, 제한적인 것은 아니지만, 사용될 수 있는 예시적인 유형의 하드웨어 로직은 FPGA(Field-programmable Gate Array), ASIC(Program-specific Integrated Circuit), ASSP(Program-specific Standard Product), SOC(System-on-a-chip) system), CPLD(Complex Programmable Logic Device) 등을 포함한다.
- [0218] 본 명세서에 기술된 발명대상의 방법을 수행하기 위한 프로그램 코드는 하나 이상의 프로그래밍 언어의 임의의 조합으로 기록될 수 있다. 프로그램 코드는, 프로세서 또는 컨트롤러에 의해 실행될 경우에 흐름도 및/또는 블록도에서 지칭된 기능/동작이 구현되게 하도록, 범용 컴퓨터, 특수 목적 컴퓨터, 또는 프로그램가능한 데이터 처리 장치의 프로세서 또는 컨트롤러에 제공될 수 있다. 프로그램 코드는 머신 상에서 완전히 또는 부분적으로 실행될 수도 있고, 머신 상에서 독립형 소프트웨어 패키지로서 실행될 수도 있고, 또는 완전히 원격 머신 또는 서버 상에서 실행될 수도 있다.
- [0219] 본 개시의 컨텍스트에서, 머신 판독가능 매체는 명령어 실행 시스템, 장치, 또는 디바이스에 의해 또는 이와 함께 사용할 프로그램을 포함하거나 저장할 수 있는 임의의 물리적인 매체일 수 있다. 머신 판독가능 매체는 머신 판독가능 신호 매체 또는 머신 판독가능 저장 매체일 수 있다. 머신 판독가능 매체는 전자, 자기, 광학, 전자기, 적외선 또는 반도체 시스템, 장치, 또는 디바이스, 또는 이들의 임의의 적절한 조합을 포함할 수 있지만 이들로 제한되지는 않는다. 머신 판독가능 저장 매체의 보다 구체적인 예들은, 하나 이상의 유선을 갖는 전기

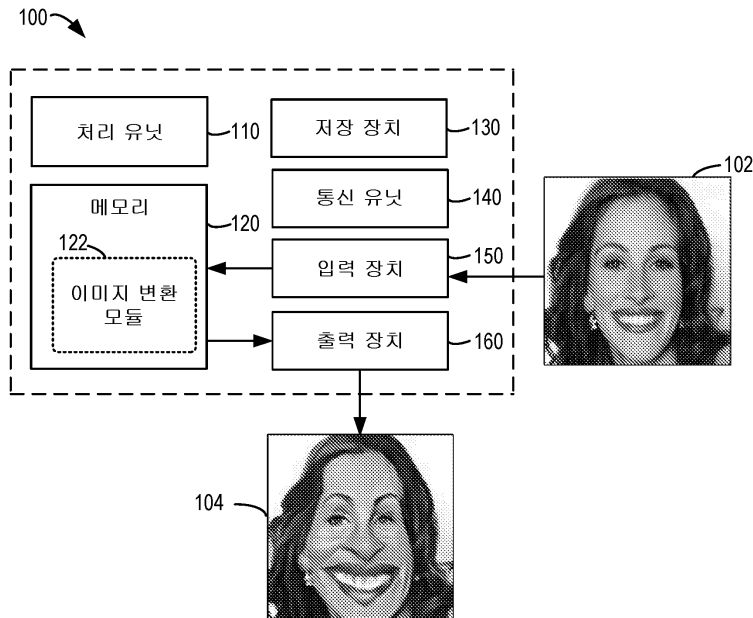
접속, 휴대용 컴퓨터 디스켓, 하드 디스크, RAM(random-access memory), ROM(a read-only memory), EPROM(erasable programmable read-only memory) 또는 플래시 메모리, 광섬유, 휴대용 CD-ROM(compact disc read-only memory), 광학 저장 장치, 자기 저장 장치, 또는 이들의 임의의 적절한 조합을 포함한다.

[0220] 또한, 동작들이 특정 순서로 도시되지만, 이는 그러한 동작들이 도시된 특정 순서 또는 순차적으로 수행된다거나 또는 원하는 결과를 얻기 위해 도시된 모든 동작들이 수행되는 것으로 이해하면 안 된다. 어떤 환경에서는, 멀티태스킹 및 병렬 처리가 유리할 수도 있다. 마찬가지로, 여러 특정 구현 세부사항이 위 설명에 포함되어 있지만, 이들은 본 명세서에 기술된 발명대상의 범위에 대한 제한으로 해석해서는 안 되고, 특정 구현예에 특정될 수 있는 특징들에 대한 설명으로 해석해야 한다. 개별 구현예들의 컨텍스트에서 설명되는 어떤 특징들은 단일 구현예에서 조합으로 구현될 수도 있다. 오히려, 단일 구현예에서 설명되는 다양한 특징들이 다수의 구현예들로 별도로 구현될 수도 있고 아니면 임의의 적절한 하위 조합으로 구현될 수도 있다.

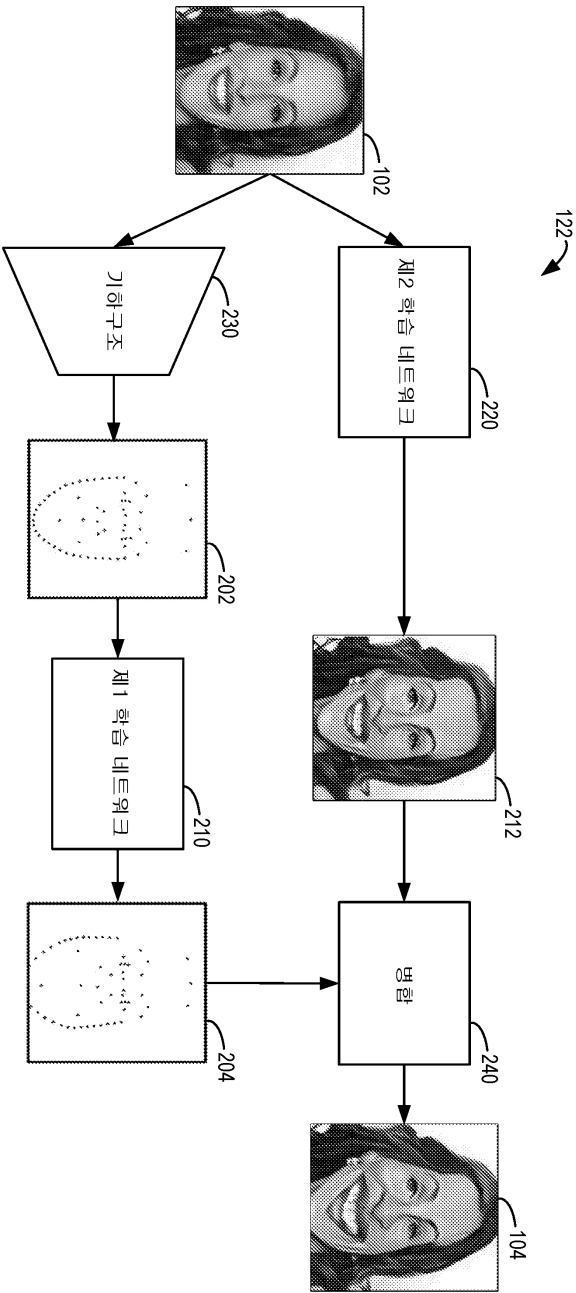
[0221] 발명대상은 구조적인 특징 및/또는 방법의 동작에 특유한 언어로 기술되어 있지만, 첨부된 청구항들에 정의된 청구대상은 반드시 전술한 특정한 특징이나 동작으로 제한되지는 않음을 이해해야 한다. 오히려, 전술한 특정한 특징 및 동작은 청구항들을 구현하는 예시적인 형태로서 개시되어 있다.

도면

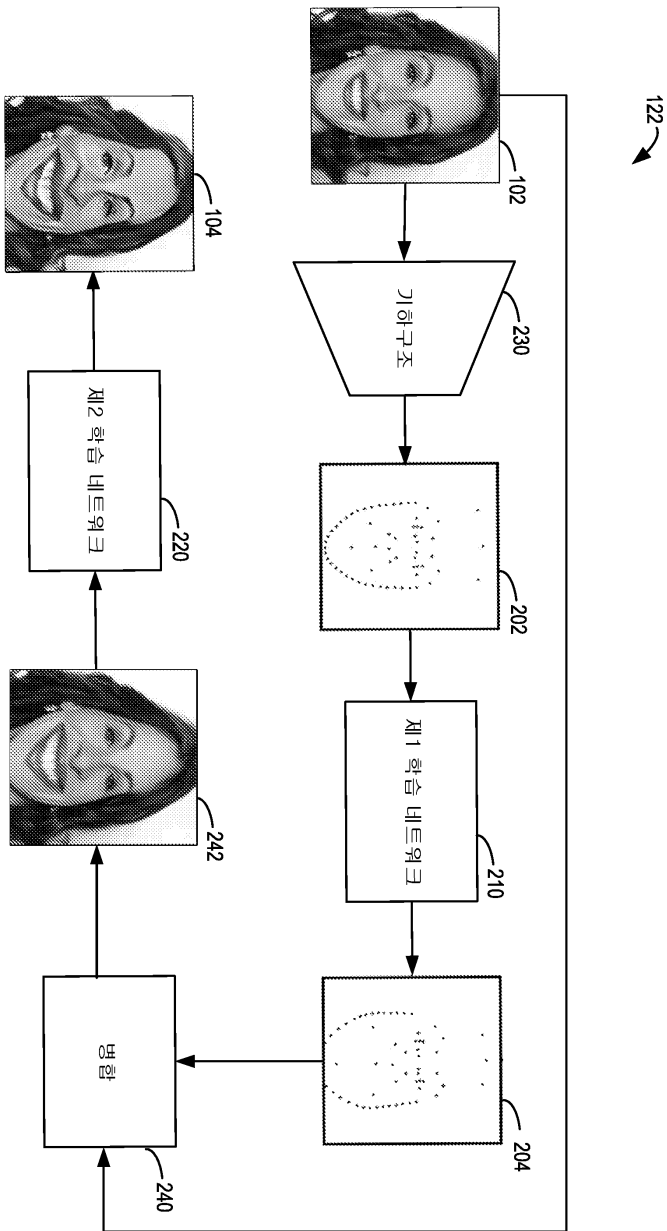
도면1



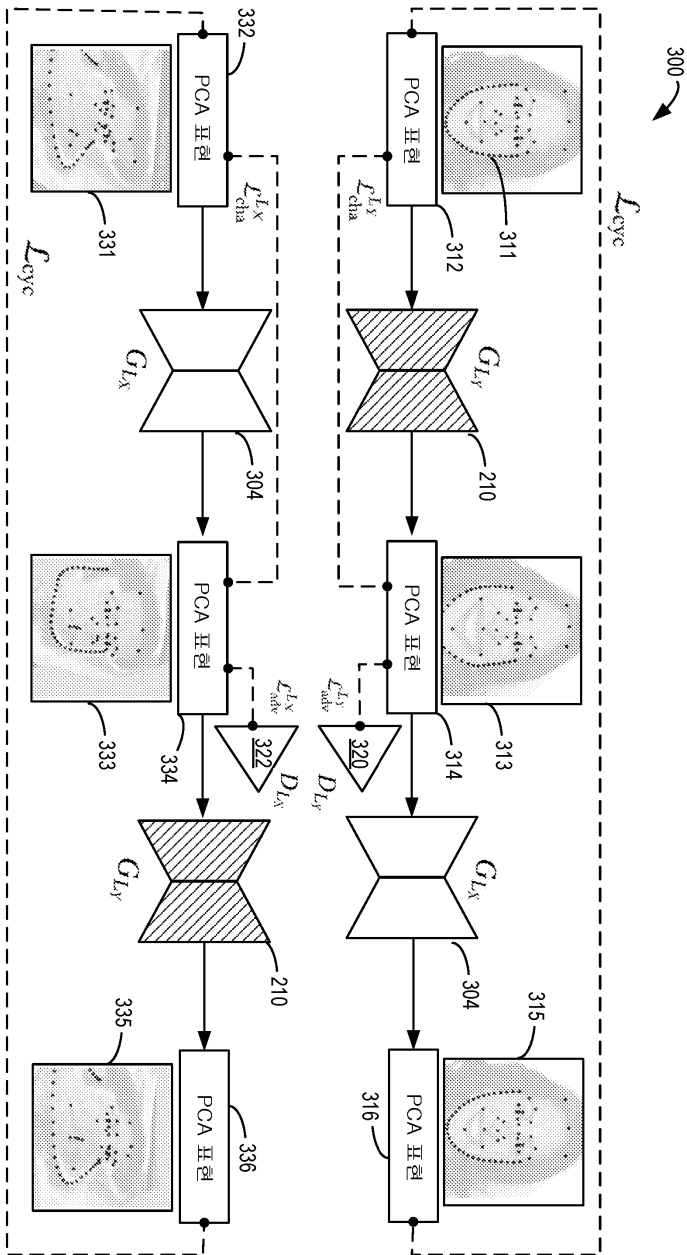
도면2a



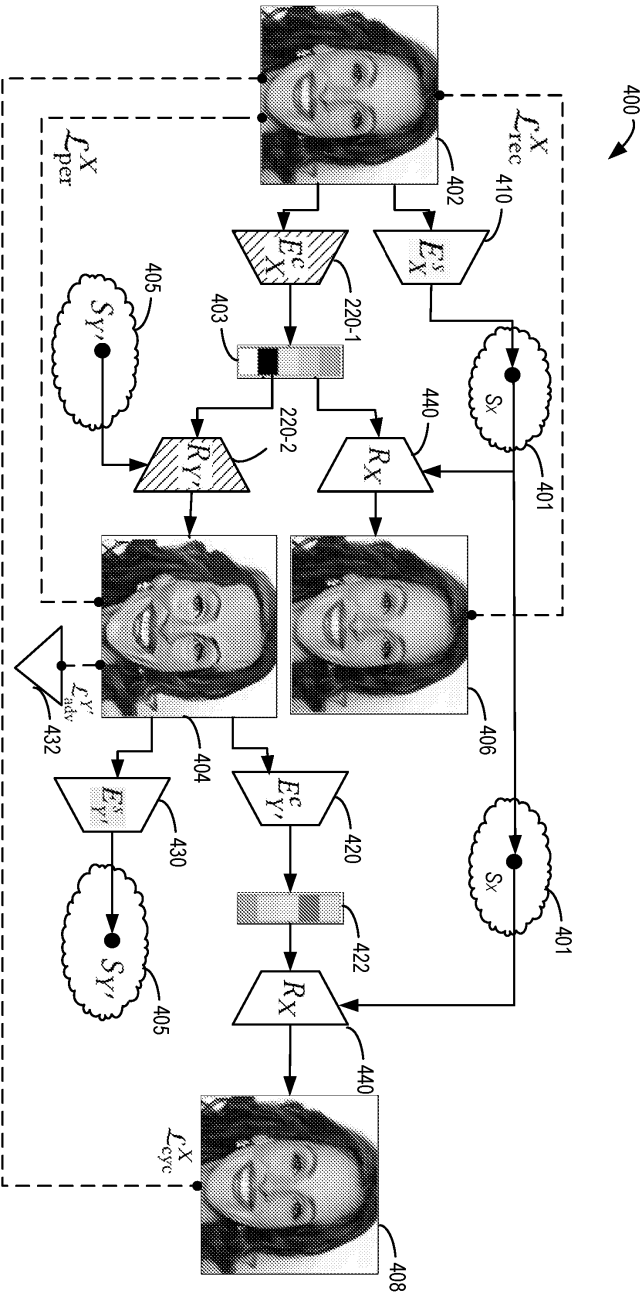
도면2b



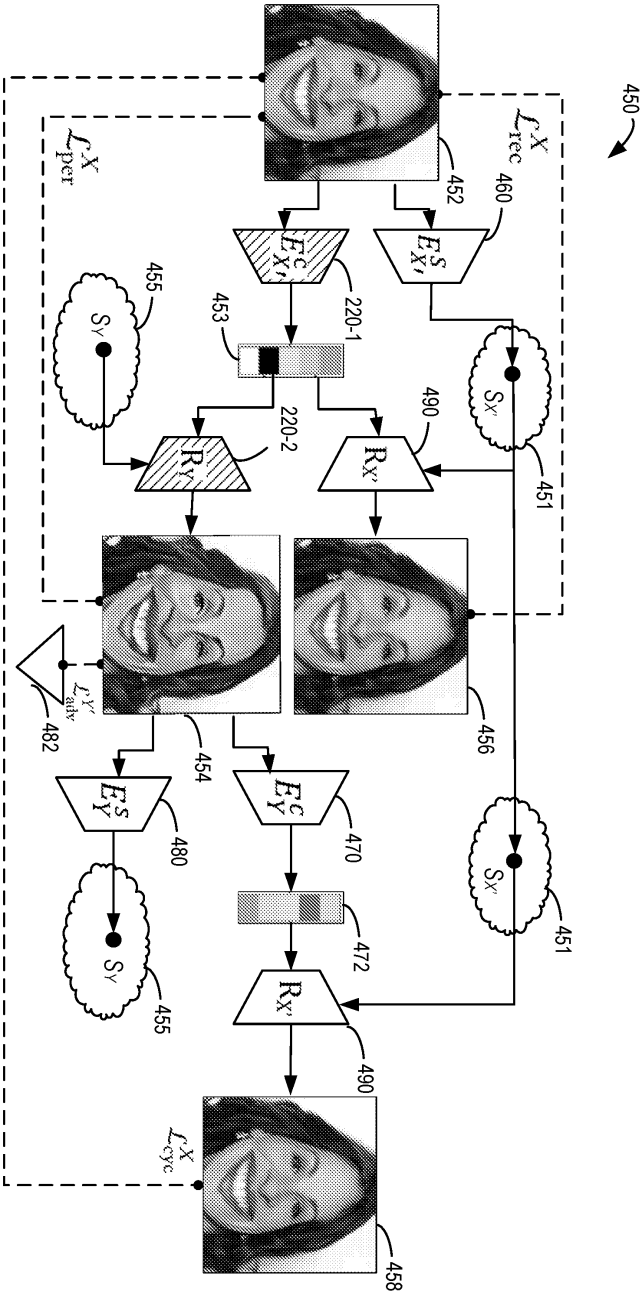
도면3



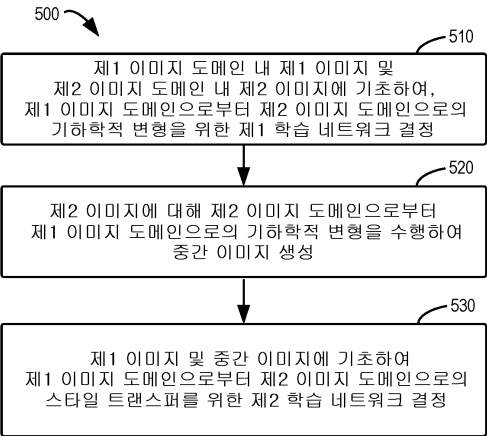
도면4a



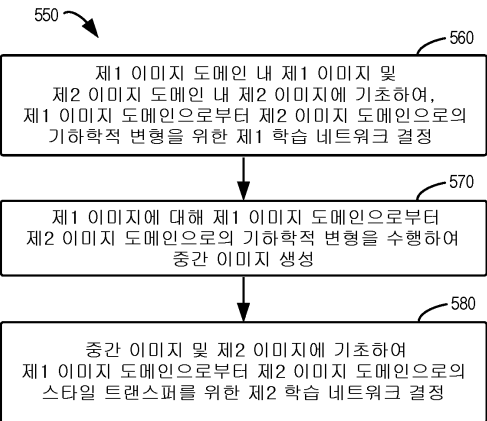
도면4b



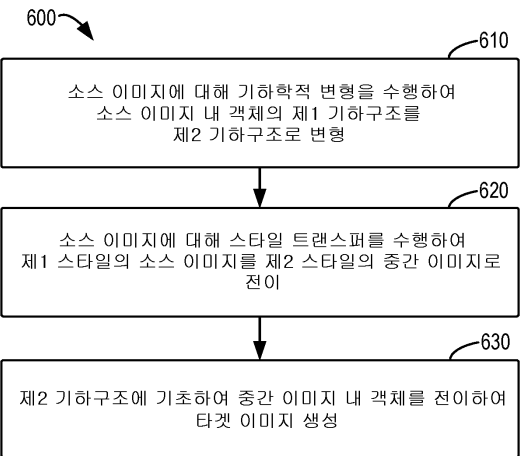
도면5a



도면5b



도면6a



도면6b

