



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년06월17일
(11) 등록번호 10-2266165
(24) 등록일자 2021년06월11일

(51) 국제특허분류(Int. Cl.)
G06T 11/60 (2006.01) G06K 9/00 (2006.01)
G06N 3/08 (2006.01)
(52) CPC특허분류
G06T 11/60 (2013.01)
G06K 9/00268 (2013.01)
(21) 출원번호 10-2021-0039531
(22) 출원일자 2021년03월26일
심사청구일자 2021년03월26일
(56) 선행기술조사문헌
KR101725808 B1*
(뒷면에 계속)

(73) 특허권자
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
박인규
서울특별시 강남구 선릉로 221, 103동 1403호(도곡동, 도곡렉슬아파트)
파르호드
경기도 수원시 영통구 청명로59번길 59-1(영통동)
홍성은
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(74) 대리인
양성보

전체 청구항 수 : 총 6 항

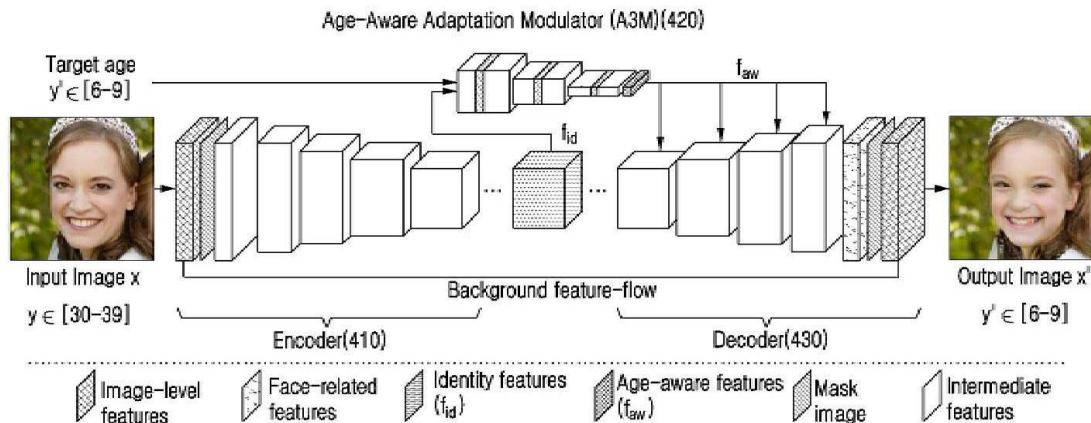
심사관 : 이병우

(54) 발명의 명칭 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법 및 장치

(57) 요약

적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법 및 장치가 제시된다. 본 발명에서 제안하는 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 장치는 얼굴 연령 변환을 위한 원본 이미지 및 목표 연령을 입력 받아 얼굴 연령 변환된 이미지를 생성하는 생성기 및 얼굴 연령 변환된 이미지가 실제 연령 도 (뒷면에 계속)

대표도



메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정하는 판별기를 포함하고, 본 발명의 실시예에 따른 생성기는 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 신원 인코더, 추출된 신원 특징을 입력 받아 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 연령 인식 적응 모듈레이터 및 신원 특징 및 개인별 연령 특성으로 잔차 블록 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 디코더를 포함한다.

(52) CPC특허분류

G06N 3/08 (2013.01)

G06T 2207/20084 (2013.01)

G06T 2207/30201 (2013.01)

(56) 선행기술조사문헌

KR102213494 B1*

KR1020210019182 A*

KR1020200132665 A

KR102206205 B1

*는 심사관에 의하여 인용된 문헌

이 발명을 지원한 국가연구개발사업

과제고유번호	1711110995
과제번호	2019R1A2C1006706
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	중견연구
연구과제명	[Ezbaro] 제한없는 임의 시공간에서 취득한 영상으로부터 완전한 얼굴 복원
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2020.03.01 ~ 2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116910
과제번호	2020-0-01389-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합연구센터지원사업(국고)
연구과제명	[Ezbaro][정부] 인공지능융합연구센터지원(1차년도)
기 여 율	1/2
과제수행기관명	인하대학교 산학협력단
연구기간	2020.04.01 ~ 2020.12.31

공지예외적용 : 있음

명세서

청구범위

청구항 1

얼굴 연령 변환을 위한 원본 이미지를 입력 받아 생성기의 신원 인코더를 통해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 단계;

추출된 신원 특징을 입력 받아 생성기의 연령 인식 적응 모듈레이터를 통해 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 단계;

생성기의 디코더를 통해 신원 특징 및 개인별 연령 특성으로 잔차 블록 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 단계; 및

판별기를 통해 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정하는 단계

를 포함하고,

얼굴 연령 변환을 위한 원본 이미지를 입력 받아 생성기의 신원 인코더를 통해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 단계는,

영상-특징 레벨 컨볼루션 레이어를 포함하고, 잔여 블록을 다운샘플링하는 신원 인코더는 원본 이미지의 얼굴 형태와 동일한 얼굴 형태를 생성하기 위해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 제공하고, 얼굴 영역에서만 신원 특징을 추출하고, 동시에 배경 정보를 유지하기 위해 얼굴 관련 특징뿐만 아니라 배경 관련 특징을 획득하도록 마스킹을 수행하는

적대적 생성 신경망의 얼굴 연령 편집 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,

추출된 신원 특징을 입력 받아 생성기의 연령 인식 적응 모듈레이터를 통해 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 단계는,

디코더 레이어를 변조하기 위해 사용되는 특징 벡터를 생성하는 조건부 배치 정규화 기법과 함께 다운샘플링 레이어의 집합을 포함하는 연령 인식 적응 모듈레이터는 신원 정보를 보존하기 위해, 신원 인코더로부터 신원 특징을 입력 받아 연령 정보를 재구성한 요소 연령 인식 벡터를 생성하여 목표 연령 변화를 학습함으로써 신원 특징 및 연령 특성을 모두 포함하는 개인별 연령 특성을 출력하는

적대적 생성 신경망의 얼굴 연령 편집 방법.

청구항 4

제1항에 있어서,

생성기의 디코더를 통해 신원 특징 및 개인별 연령 특성으로 잔차 블록 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 단계는,

신원 특징 및 개인별 연령 특성을 모두 만족시키기 위해 개인별 연령 특성을 이용하여 형태가 이루어지지 않은 신원 특징에 대한 변조 작업을 통해 디코딩 프로세스의 자체 지침(self-guide)을 수행하고, 신원 특징은 적응형 인스턴스 정규화 레이어에 의해 변조되고, 신원 인코더에서 수행된 마스킹을 통해 획득한 배경 관련 특징을 이용하여 얼굴 관련 특징에 직접 추가하는 대신 학습된 개인별 연령 특성을 포함하는 얼굴 표현에 대해 다시 마스킹 작업을 수행하여 정규화 및 활성화 레이어로 인해 배경에서 발생한 노이즈를 제거한 후 얼굴 관련 특징과 배

경 관련 특징을 결합하고 마지막 컨볼루션 레이어를 사용하여 영상 도메인에 매핑하는 적대적 생성 신경망의 얼굴 연령 편집 방법.

청구항 5

얼굴 연령 변환을 위한 원본 이미지 및 목표 연령을 입력 받아 얼굴 연령 변환된 이미지를 생성하는 생성기; 및
얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정하는 판별기

를 포함하고,

생성기는,

얼굴 연령 변환을 위한 원본 이미지를 입력 받아 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 신원 인코더;

추출된 신원 특징을 입력 받아 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 연령 인식 적응 모듈레이터; 및

신원 특징 및 개인별 연령 특성으로 잔차 블럭 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 디코더

를 포함하고,

신원 인코더는,

영상-특징 레벨 컨볼루션 레이어를 포함하고, 잔여 블록을 다운샘플링하며,

원본 이미지의 얼굴 형태와 동일한 얼굴 형태를 생성하기 위해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 제공하고, 얼굴 영역에서만 신원 특징을 추출하고, 동시에 배경 정보를 유지하기 위해 얼굴 관련 특징뿐만 아니라 배경 관련 특징을 획득하도록 마스킹을 수행하는

적대적 생성 신경망의 얼굴 연령 편집 장치.

청구항 6

삭제

청구항 7

제5항에 있어서,

연령 인식 적응 모듈레이터는,

디코더 레이어를 변조하기 위해 사용되는 특징 벡터를 생성하는 조건부 배치 정규화 기법과 함께 다운샘플링 레이어의 집합을 포함하고,

신원 정보를 보존하기 위해, 신원 인코더로부터 신원 특징을 입력 받아 연령 정보를 재구성한 요소 연령 인식 벡터를 생성하여 목표 연령 변화를 학습함으로써 신원 특징 및 연령 특성을 모두 포함하는 개인별 연령 특성을 출력하는

적대적 생성 신경망의 얼굴 연령 편집 장치.

청구항 8

제5항에 있어서,

디코더는,

신원 특징 및 개인별 연령 특성을 모두 만족시키기 위해 개인별 연령 특성을 이용하여 형태가 이루어지지 않은 신원 특징에 대한 변조 작업을 통해 디코딩 프로세스의 자체 지침(self-guide)을 수행하고, 신원 특징은 적응형 인스턴스 정규화 레이어에 의해 변조되고, 신원 인코더에서 수행된 마스킹을 통해 획득한 배경 관련 특징을 이용하여 얼굴 관련 특징에 직접 추가하는 대신 학습된 개인별 연령 특성을 포함하는 얼굴 표현에 대해 다시 마스

킹 작업을 수행하여 정규화 및 활성화 레이어로 인해 배경에서 발생한 노이즈를 제거한 후 얼굴 관련 특징과 배경 관련 특징을 결합하고 마지막 컨볼루션 레이어를 사용하여 영상 도메인에 매핑하는

적대적 생성 신경망의 얼굴 연령 편집 장치.

발명의 설명

기술 분야

[0001] 본 발명은 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법 및 장치에 관한 것이다.

배경 기술

[0002] 얼굴 연령 편집은 주어진 목표 연령을 고려하여 사람의 외모를 생성하는 것을 목표로 한다. 적대적 생성신경망(Generative Adversarial Networks; GAN)[1]의 발전은 이러한 도전적인 문제를 해결하기 위한 강력한 기반 프레임워크를 제공한다. 연령 편집을 위해서 GAN 을 사용한 여러 시도 [2, 3]가 있었지만, 결과 이미지의 품질은 낮거나 인공적인 결과물(artificial artifact)이 포함된다.

[0003] 얼굴 연령 변환은 주어진 얼굴에 나이 요인을 반영하여 과거 또는 미래의 얼굴 영상을 합성하는 것을 목표로 한다. 이상적으로 이 작업은 정체성을 유지하면서 다양한 연령 그룹에 걸쳐 자연스럽게 보이는 얼굴을 고려해야 한다. 하지만, 대부분의 종래기술에서는 이 중 한 가지에만 초점을 맞췄다. 이러한 문제를 해결하기 위한 기법은 거의 없지만 복잡한 서브 네트워크 때문에 훈련하기 어렵고, 인공적인 결과물(artificial artifact)이 여전히 나타난다는 점에서 한계가 있다.

[0004] 얼굴 연령 변환에 대한 많은 종래기술이 소개되었다[Antipov 외. 2017; He 외. 2019; Or-EI 외. 2020; Wang 외. 2020; Yang 외. 2018; Yao 외. 2020; Zhang 외. 2017]. [Antipov 외. 2017; Zhang 외. 2017]에서 활용된 조건부 적대적 생성망(GAN)[Mirza and Osindero 2014]의 개념은 흐릿한 영상을 생성하고 손상된 신원 정보 [He 외. 2019; Yao 등. 2020]를 생성하게 한다. 차례로, 몇 가지 방법 [Li 외. 2019, 2018; Wang 외. 2018; Yang 외. 2018]은 연령 변환 중 정체성을 보존하기 위해 L2 제약, 사전 훈련된 신원 분류기 등과 같은 메커니즘을 채택하여 개선을 제공했다. 신원 보존 네트워크 외에도 [Le 외. 2019]는 연령 합성을 위한 텍스처 디테일을 캡처하기 위한 파형 변환도 도입했다. 마찬가지로, 파형 정보는 [Liu 외. 2019c,d]에서 사용되지만 판별기에서 다중 스케일 방식으로 사용된다. [Yao 외. 2020]에 따르면, 그들의 결과는 여전히 강력한 아티팩트로부터 어려움을 겪는다. [He 외. 2019]의 연구는 입력 표면의 연령별 표현을 얻고 디코딩하기 위해 인코딩된 개별 연령 기반으로 연령별 변환(계수)을 적용하는 S^2 GAN을 제안했다. 그럴듯한 결과에도 불구하고, S^2 GAN은 [Or-EI 외. 2020]에서 언급한 연령 진행 특성을 묘사하는 능력이 적고 해상도가 비교적 낮은 영상만 처리할 수 있다.

[0005] 고해상도 영상을 생성하기 위해 [Yao 외. 2020]은 고해상도로 작동하는 얼굴 영상 편집 모델(face age editing model operating in high-resolution; HRFAGE)을 제안했다. 이 방법은 완전히 연결된 단일 레이어의 출력을 사용하여 인코딩된 기능의 가중치를 다시 매기는 간단한 접근 방식에 의존한다. 비슷한 방식으로 [Despois 외. 2020]은 고해상도 얼굴 영상을 얻기 위해 디코딩 프로세스에서 SPADE 블록[Park 외. 2019]을 사용하여 노후화된 맵을 통합할 것을 제안했다. 이러한 방법들은 양호한 신원 보존과 나이 정보를 유지하지만, 그것은 대부분 회귀보다 나이 진행을 수행하는 것으로 제한된다. 사실, 앞에서 언급한 대부분의 종래기술은 상위 연령 그룹에 속하는 영상을 생성하도록 제한된다. 이 중요한 문제는 LATS [Or-EI 외. 2020]에서 수명 연령 변환을 합성하는 방법을 제시함으로써 해결되었다. 디코더는 StyleGAN2[Karras 외. 2020]와 비슷한 방식으로 작동하며 매핑 네트워크에서 학습된 목표 연령 잠재 벡터를 주입하는 동안 신원 특징에 대해 변조된 컨볼루션을 수행한다. 인상적인 성능에도 불구하고, LATS는 주로 배경 정보를 사용하지 않고 얼굴 합성에만 초점을 맞추고 있으며, 이는 예상치 못한 아티팩트로 이어진다. 게다가, LATS를 포함한 최근 도입된 접근 방식은 복잡한 서브 네트워크와 모듈로 구성되어 있으며, 이들을 엔드 투 엔드 방식으로 훈련시키기는 어렵다.

[0006] 얼굴 연령별 접근법 외에도 영상 간 변환 [Chen 외. 2019; Choi 외. 2018; Lample 외. 2017; Pumarola 외. 2018; Qian et. Xiao 외. 2018]을 얼굴 연령 변환 작업에 적용할 수 있었다. 다중 도메인 전송 방식으로 얼굴 연령 속성을 조작하려는 시도가 여러 번 있었다. 얼굴 연령 변환을 수행하기 위해(즉, 젊음에서 늙음, 그 반대) StarGAN [Choi 외. 2018] 및 STGAN [Liu 외. 2019a]는 연령 도메인이 다른 얼굴 관련 도메인과 구별된다고 가정한다. 다만 나이가 많아 보이는 얼굴 생김새는 다른 얼굴 속성과 연관성이 높다. 예를 들어 다중 도메인 전송을 사용하여 서로 다른 연령 그룹 간에 변환할 때 성별, 액세서리, 민족성 등 연령과 관련이 없는 스타일도 이

전될 수 있다.

발명의 내용

해결하려는 과제

- [0007] 본 발명이 이루고자 하는 기술적 과제는 생성 모델 내 개인별 연령 특성 학습에 집중하는 GAN 기반 프레임워크를 통해 개인별 특성을 학습함으로써 입력 얼굴이 대상 연령의 자연스런 노화 과정에 상응하도록 얼굴을 편집하며 동시에 신원 정보를 보존하는 방법 및 장치를 제공하는데 있다.

과제의 해결 수단

- [0008] 일 측면에 있어서, 본 발명에서 제안하는 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법은 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 생성기의 신원 인코더를 통해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 단계, 추출된 신원 특징을 입력 받아 생성기의 연령 인식 적응 모듈레이터를 통해 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 단계, 생성기의 디코더를 통해 신원 특징 및 개인별 연령 특성으로 잔차 블럭 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 단계 및 판별기를 통해 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정하는 단계를 포함한다.

- [0009] 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 생성기의 신원 인코더를 통해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 단계는 영상-특징 레벨 컨볼루션 레이어를 포함하고, 잔여 블럭을 다운샘플링하는 신원 인코더는 원본 이미지의 얼굴 형태와 동일한 얼굴 형태를 생성하기 위해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 제공하고, 얼굴 영역에서만 신원 특징을 추출하고, 동시에 배경 정보를 유지하기 위해 얼굴 관련 특징뿐만 아니라 배경 관련 특징을 획득하도록 마스킹을 수행한다.

- [0010] 추출된 신원 특징을 입력 받아 생성기의 연령 인식 적응 모듈레이터를 통해 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 단계는 디코더 레이어를 변조하기 위해 사용되는 특징 벡터를 생성하는 조건부 배치 정규화 기법과 함께 다운샘플링 레이어의 집합을 포함하는 연령 인식 적응 모듈레이터는 신원 정보를 보존하기 위해, 신원 인코더로부터 신원 특징을 입력 받아 연령 정보를 재구성한 요소 연령 인식 벡터를 생성하여 목표 연령 변화를 학습함으로써 신원 특징 및 연령 특성을 모두 포함하는 개인별 연령 특성을 출력한다.

- [0011] 생성기의 디코더를 통해 신원 특징 및 개인별 연령 특성으로 잔차 블럭 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 단계는 신원 특징 및 개인별 연령 특성을 모두 만족시키기 위해 개인별 연령 특성을 이용하여 형태가 이루어지지 않은 신원 특징에 대한 변조 작업을 통해 디코딩 프로세스의 자체 지침(self-guide)을 수행하고, 신원 특징은 적응형 인스턴스 정규화 레이어에 의해 변조되고, 신원 인코더에서 수행된 마스킹을 통해 획득한 배경 관련 특징을 이용하여 얼굴 관련 특징에 직접 추가하는 대신 학습된 개인별 연령 특성을 포함하는 얼굴 표현에 대해 다시 마스킹 작업을 수행하여 정규화 및 활성화 레이어로 인해 배경에서 발생한 노이즈를 제거한 후 얼굴 관련 특징과 배경 관련 특징을 결합하고 마지막 컨볼루션 레이어를 사용하여 영상 도메인에 매핑한다.

- [0012] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 장치는 얼굴 연령 변환을 위한 원본 이미지 및 목표 연령을 입력 받아 얼굴 연령 변환된 이미지를 생성하는 생성기 및 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정하는 판별기를 포함한다.

- [0013] 본 발명의 실시예에 따른 생성기는 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 신원 인코더, 추출된 신원 특징을 입력 받아 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 연령 인식 적응 모듈레이터 및 신원 특징 및 개인별 연령 특성으로 잔차 블럭 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 디코더를 포함한다.

발명의 효과

- [0014] 본 발명의 실시예들에 따르면 연령 변환의 모든 주요 요인을 고려한 새로운 단일 프레임워크인 RAGAN(Re-Aging GAN)을 통해 생성 프로세스의 자체 지침을 수행하도록 입력 신원을 보존하는 개인 얼굴 연령 변환을 달성할 수

있다. 제안하는 RAGAN은 주어진 정체성과 목표 연령 사이의 고차 상호작용을 사용하여 개인 연령 특징을 학습할 수 있다.

도면의 간단한 설명

- [0015] 도 1은 본 발명의 일 실시예에 따른 FFHQ 데이터 셋의 이미지와 StyleGANv2의 합성 이미지에 대한 연속 연령 변환 결과를 나타내는 도면이다.
- 도 2는 본 발명의 일 실시예에 따른 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법을 설명하기 위한 흐름도이다.
- 도 3은 본 발명의 일 실시예에 따른 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 장치의 구성을 나타내는 도면이다.
- 도 4는 본 발명의 일 실시예에 따른 RAGAN 프레임 워크를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 본 발명에서는 생성 모델 내 개인별 연령 특성 학습에 집중하는 GAN 기반 프레임워크를 제안한다. 제안 기법은 개인별 특성을 학습함으로써 입력 얼굴이 목표 연령의 자연스런 노화 과정에 상응하도록 얼굴을 편집하며 동시에 신원 정보를 보존할 수 있다.
- [0017] 본 발명에서는 연령 변환의 모든 주요 요인을 고려한 새로운 단일 프레임워크인 RAGAN(Re-Aging GAN)을 제안한다. 제안하는 프레임워크는 생성 프로세스의 자체 지침을 수행하도록 입력 신원을 보존하는 개인 얼굴 연령 변환을 달성한다. 구체적으로, RAGAN은 주어진 정체성과 목표 연령 사이의 고차 상호작용을 사용하여 개인 연령 특징을 학습할 수 있다. 학습된 개인 연령 특징은 목표 연령에 따라 재조정되는 신원 정보이다. 따라서, 이러한 특징은 입력 신원이 특정 연령에 어떻게 있어야 하는지에 대한 변환 프로세스의 중요한 단서를 제공하는 신원 및 목표 연령 정보를 포함한다. 본 발명의 실시예에 따른 실험에서 가장 낮은 FID 및 KID 점수와 가장 높은 연령 인식 정확도는 이전 방법에 비해 더 나은 정량적 결과를 보여준다. 또한 다양한 연령 그룹에 걸쳐 적은 인공적인 결과물을 나타내고, 신원 보존 및 변환에 대해 제안하는 방법의 시각적 우수성을 보여준다.
- [0018] 이하, 본 발명의 실시 예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0020] 도 1은 본 발명의 일 실시예에 따른 FFHQ 데이터 셋의 이미지와 StyleGANv2의 합성 이미지에 대한 연속 연령 변환 결과를 나타내는 도면이다.
- [0021] 얼굴 연령 변환은 목표 연령에 따라 주어진 얼굴의 연령 진행 또는 회귀를 학습한다. 여기서 목표 연령은 특정 연령의 얼굴 영상을 생성하도록 변환을 안내하는 명시적 조건 요인을 의미한다. 즉, 입력 얼굴 영상의 목표 나이를 설정할 수 있으며 도 1과 같이 목표 연령 특징을 나타내는 출력 얼굴을 가질 수 있다. 이상적으로, 연령 변환 모델은 다음 특성을 만족시켜야 한다. 첫째, 모델은 얼굴 나이를 진행/회귀하는 동안 사람의 독자성을 고려해야 하며, 거의 변경되지 않은 상태, 즉 신원 보존을 유지해야 한다. 둘째, 모델은 다양한 연령 그룹에 걸쳐 목표 연령에 해당하는 자연적으로 보이는 얼굴을 생성할 수 있어야 한다.
- [0022] 얼굴 연령 변환에 있어서, 광범위한 연령 그룹에 걸친 신원 보존 연령 변환과 같은 모든 중요한 요인을 다루는 모델을 제안하기 위해, 본 발명에서는 RAGAN(Re-Aging GAN)이라는 간단하면서도 효과적인 프레임워크를 제안한다.
- [0023] 종래의 얼굴 연령 변환 모델과 달리 제안된 RAGAN은 생성기와 판별기로 구분되고 자체 지침을 활용하여 모델 훈련을 더 쉽고 확장 가능한 단일 프레임워크이다. 생성기 부분은 바닐라 인코더-디코더 네트워크에서 시작한다. 생성기에서는 인코더-디코더의 병목 현상에 연결되는 A3M(Age-Aware Adaptation Modulator)을 도입하는데, 이는 목표 연령 임베딩을 통해 인코더 네트워크의 신원 보존 기능을 변조하는 것을 목표로 한다. 구체적으로는 연령 정보의 통합에 따라 신원 특징을 직접 정렬하기 위해 A3M 모듈에는 주어진 목표 연령에 따라 특징을 변조하는 여러 개의 라벨 주입 장치가 있다. 마지막으로, 디코더 네트워크의 연령 인식 기능에 대한 자체 지침을 통해, 본 발명의 실시예에 따른 생성기는 다양한 연령의 얼굴에서 나이 변환을 안정적으로 수행하기 위해 학습한다. 본 발명의 실시예에 따른 판별기는, 연령 관련 손실로 구성된 기존 접근법[Choi 외. 2020; Liu 외. 2019b; Mescheder 외. 2018]을 따른다.
- [0024] 본 발명의 실시예에 따르면 신원을 유지하면서 다양한 목표 연령 그룹에 걸쳐 입력 얼굴을 변환할 수 있는 개인

화된 자기 지침 방식을 도입한다. 또한, 페어링되지 않은(unpaired) 방식으로 효과적으로 훈련된 단일 생성 및 차별적 모델만을 사용하여 성공적으로 얼굴 연령 변환을 수행한다. 제안하는 방법에 대한 광범위한 실험과 평가를 통해 종래기술에 비해 제안된 RAGAN의 우수성을 정성적 정량적으로 입증한다.

- [0026] 도 2는 본 발명의 일 실시예에 따른 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법을 설명하기 위한 흐름도이다.
- [0027] 제안하는 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 방법은 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 생성기의 신원 인코더를 통해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출하는 단계(210), 추출된 신원 특징을 입력 받아 생성기의 연령 인식 적응 모듈레이터를 통해 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력하는 단계(220), 생성기의 디코더를 통해 신원 특징 및 개인별 연령 특성으로 잔차 블록 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성하는 단계(230) 및 판별기를 통해 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정하는 단계(240)를 포함한다.
- [0028] 단계(210)에서, 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 생성기의 신원 인코더를 통해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출한다. 신원 인코더는 영상-특징 레벨 컨볼루션 레이어를 포함하고, 잔여 블록을 다운샘플링한다. 이때, 원본 이미지의 얼굴 형태와 동일한 얼굴 형태를 생성하기 위해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 제공한다. 신원 인코더는 얼굴 영역에서만 신원 특징을 추출하고, 동시에 배경 정보를 유지하기 위해 얼굴 관련 특징뿐만 아니라 배경 관련 특징을 획득하도록 마스킹을 수행한다.
- [0029] 단계(220)에서, 추출된 신원 특징을 입력 받아 생성기의 연령 인식 적응 모듈레이터를 통해 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력한다. 연령 인식 적응 모듈레이터는 디코더 레이어를 변조하기 위해 사용되는 특징 벡터를 생성하는 조건부 배치 정규화 기법과 함께 다운샘플링 레이어의 집합을 포함한다. 연령 인식 적응 모듈레이터는 신원 정보를 보존하기 위해, 신원 인코더로부터 신원 특징을 입력 받아 연령 정보를 재구성한 요소 연령 인식 벡터를 생성하여 목표 연령 변화를 학습함으로써 신원 특징 및 연령 특성을 모두 포함하는 개인별 연령 특성을 출력한다.
- [0030] 단계(230)에서, 생성기의 디코더를 통해 신원 특징 및 개인별 연령 특성으로 잔차 블록 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성한다.
- [0031] 신원 특징 및 개인별 연령 특성을 모두 만족시키기 위해 개인별 연령 특성을 이용하여 형태가 이루어지지 않은 신원 특징에 대한 변조 작업을 통해 디코딩 프로세스의 자체 지침(self-guide)을 수행하고, 신원 특징은 적응형 인스턴스 정규화 레이어에 의해 변조된다. 디코더는 신원 인코더에서 수행된 마스킹을 통해 획득한 배경 관련 특징을 이용하여 얼굴 관련 특징에 직접 추가하는 대신 학습된 개인별 연령 특성을 포함하는 얼굴 표현에 대해 다시 마스킹 작업을 수행하여 정규화 및 활성화 레이어로 인해 배경에서 발생한 노이즈를 제거한 후 얼굴 관련 특징과 배경 관련 특징을 결합하고 마지막 컨볼루션 레이어를 사용하여 영상 도메인에 매핑한다.
- [0032] 단계(240)에서, 판별기를 통해 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정한다. 판별기는 멀티 태스크(multi-task) 네트워크의 형태로 설계되며 각 분기는 주어진 이미지가 실제 연령 도메인에 속하는지 또는 생성기에 의해 편집되었는지 결정한다. 결과적으로 판별기는 각 연령 그룹의 분류를 위해 다운 샘플링 잔여 블록 세트 및 완전 연결 레이어들(fully connected layers)로 구성된다.
- [0034] 도 3은 본 발명의 일 실시예에 따른 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 장치의 구성을 나타내는 도면이다.
- [0035] 제안하는 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 장치(300)는 생성기(310) 및 판별기(320)를 포함한다.
- [0036] 생성기(310)는 얼굴 연령 변환을 위한 원본 이미지 및 목표 연령을 입력 받아 얼굴 연령 변환된 이미지를 생성한다.
- [0037] 생성기(310)는 신원 인코더(311), 연령 인식 적응 모듈레이터(312) 및 디코더(313)를 포함한다.
- [0038] 신원 인코더(311)는 얼굴 연령 변환을 위한 원본 이미지를 입력 받아 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 포함하는 원본 이미지의 신원 특징을 추출한다. 신원 인코더(311)는 영상-특징 레벨 컨볼루션 레이어

를 포함하고, 잔여 블록을 다운샘플링한다. 이때, 원본 이미지의 얼굴 형태와 동일한 얼굴 형태를 생성하기 위해 얼굴 형태에 관한 정보 및 로컬 레벨의 얼굴 구조를 제공한다. 신원 인코더(311)는 얼굴 영역에서만 신원 특징을 추출하고, 동시에 배경 정보를 유지하기 위해 얼굴 관련 특징뿐만 아니라 배경 관련 특징을 획득하도록 마스킹을 수행한다.

[0039] 연령 인식 적응 모듈레이터(312)는 추출된 신원 특징을 입력 받아 원본 이미지의 목표 연령 변화를 학습하여 개인별 연령 특성을 출력한다. 연령 인식 적응 모듈레이터(312)는 디코더 레이어를 변조하기 위해 사용되는 특징 벡터를 생성하는 조건부 배치 정규화 기법과 함께 다운샘플링 레이어의 집합을 포함한다. 연령 인식 적응 모듈레이터(312)는 신원 정보를 보존하기 위해, 신원 인코더(311)로부터 신원 특징을 입력 받아 연령 정보를 재구성한 요소 연령 인식 벡터를 생성하여 목표 연령 변화를 학습함으로써 신원 특징 및 연령 특성을 모두 포함하는 개인별 연령 특성을 출력한다.

[0040] 디코더(313)는 신원 특징 및 개인별 연령 특성으로 잔차 블록 및 업샘플링 레이어를 사용하여 얼굴 연령 변환된 이미지를 생성한다.

[0041] 디코더(313)는 신원 특징 및 개인별 연령 특성을 모두 만족시키기 위해 개인별 연령 특성을 이용하여 형태가 이루어지지 않은 신원 특징에 대한 변조 작업을 통해 디코딩 프로세스의 자체 지침(self-guide)을 수행하고, 신원 특징은 적응형 인스턴스 정규화 레이어에 의해 변조된다. 디코더(313)는 신원 인코더(311)에서 수행된 마스킹을 통해 획득한 배경 관련 특징을 이용하여 얼굴 관련 특징에 직접 추가하는 대신 학습된 개인별 연령 특성을 포함하는 얼굴 표현에 대해 다시 마스킹 작업을 수행하여 정규화 및 활성화 레이어로 인해 배경에서 발생한 노이즈를 제거한 후 얼굴 관련 특징과 배경 관련 특징을 결합하고 마지막 컨볼루션 레이어를 사용하여 영상 도메인에 매핑한다.

[0042] 판별기(320)는 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지인지 또는 생성기에 의해 변환된 이미지인지 여부를 결정한다. 판별기(320)는 멀티 태스크(multi-task) 네트워크의 형태로 설계되며 각 분기는 주어진 이미지가 실제 연령 도메인에 속하는지 또는 생성기에 의해 편집되었는지 결정한다. 결과적으로 판별기는 각 연령 그룹의 분류를 위해 다운 샘플링 잔여 블록 세트 및 완전 연결 레이어들(fully connected layers)로 구성된다.

[0043] 도 4를 참조하여 본 발명의 일 실시예에 따른 적대적 생성신경망에서의 개인별 변조를 이용한 얼굴 연령 편집 장치(300)의 생성기(310)에 대하여 더욱 상세히 설명한다.

[0045] 도 4는 본 발명의 일 실시예에 따른 RAGAN 프레임 워크를 설명하기 위한 도면이다.

[0046] 각각 X 와 Y 를 이미지 집합과 가능한 연령으로 설정한다. $y' \in Y$ 에서 랜덤으로 추출한 얼굴 영상 $x \in X$ 와 목표 연령을 고려하여, 본 발명의 실시예에 따른 목표는 x 의 신원에 해당하는 특정 연령 y' 의 얼굴 영상 x' 을 생성할 수 있도록 단일 생성기 G 를 훈련시킨다. 또한, G 내에 A3M을 도입하여 목표 연령을 고려하여 신원 특징을 재구성하고 이를 자체 지침 정보(self-guiding information)로 활용한다.

[0047] 종래기술들과 비교하여, 제안하는 방법의 주요 목표는 얼굴 연령을 강력하게 변화시킬 뿐만 아니라 배경, 헤어 스타일, 표현 등과 같이 x' 에서 나이와 무관한 정보를 최대한 유지하는 것이다. 이것은 제안하는 프레임워크가 연령 변환 프로세스 동안 입력 영상에 포함된 연령 관련 정보를 보존해야 한다는 것을 의미한다. 따라서, 본 발명의 실시예에 따른 인코더에 대한 간단한 전략을 고려하고 디코더 네트워크를 통해 중요한 정보를 공유할 수 있다.

[0048] 본 발명의 실시예에 따른 GAN 기반 프레임워크는 신원 인코더(410), 연령 인식 적응 모듈레이터(420) 및 디코더(430)를 포함하는 생성기와 판별기 부분으로 나눌 수 있다. 본 발명의 실시예에 따른 판별기는 종래기술 [Choi 외, 2020; Liu 외, 2019b; Mescheder 외, 2018]을 따르고 본 발명의 주요 초점은 아니기 때문에, 도 4에는 생성기 부분만 도시한다.

[0049] 도 4와 같이, 본 발명의 실시예에 따른 생성기는 영상 생성을 위해 인코더-디코더 아키텍처 개념을 사용하며 신원 인코더(Enc)(410), 연령 인식 적응 모듈레이터(Age-Aware Adaptation Modulator; A3M)(420), 디코더(Dec)(430)를 포함한다. 본 발명의 실시예에 따른 인코더와 디코더 네트워크는 종래기술과 동일한 절차를 수행하지만 몇 가지 개선된 사항이 있다. 본 발명과 종래기술의 주요 차이점은 생성기의 병목 지점(bottleneck)에 추가 서브 네트워크를 통합하는 것이다. 이 네트워크를 통해 특정 사람이 고려 목표 연령에서 어떻게 보여야 하는지에 대한 정보를 제공하는 특징을 얻을 수 있다. 이러한 연령 인식 특징은 주어진 입력 영상 자체를 기반으로 학습되기 때문에, 이후 생성 프로세스에서 자체 지침 정보로 사용될 수 있다. 전반적으로 입력의 연령 특성

을 변환하는 프레임워크는 다음과 같이 수행된다.

$$x' = G(x, y') = \text{Dec}(A3M(\text{Enc}(x), y')) \quad (1)$$

연령 변환을 위한 원본 이미지(x)이 주어지면, 신원 인코더(Enc)(410)는 이미지의 신원 특징(f_{id})을 추출한다. 여기서 $f_{id} = \text{Enc}(x)$ 이다. 특히, 신원 인코더(Enc)(410)는 얼굴 형태에 대한 일반적인 정보와 로컬 레벨의 얼굴 구조를 제공하는 기능을 제공한다. 이러한 기능은 동일한 외관의 얼굴을 생성하기 위해 직관적으로 얼굴 영역에만 집중하게 한다. 따라서 원본 이미지(x)를 특징 도메인으로 전송한 후 마스킹 작업을 수행한다. 본 발명에서는 얼굴 영역에서만 작동하려는 의도를 충족하고 동시에 배경 정보를 유지하기 위해 얼굴 관련 특징뿐만 아니라 배경 관련 특징을 획득할 수 있도록 의도적으로 마스킹을 수행한다. 아키텍처 수준에서, 신원 인코더는 영상-특징 레벨 컨볼루션 레이어를 가지도록 설계되었으며, 이후 잔여 블록을 다운 샘플링하도록 설계되었다[He 등. 2016a].

연령 인식 적응 모듈레이터(Age-Aware Adaptation Modulator; A3M)(420)는 저차원 입력 벡터를 학습하기 위해 널리 사용되는 CNN의 형태로 구성된다. 인코더로부터 신원 특징(f_{id})을 입력 받고 주어진 나이 정보(y') 출력의 재구성된 버전 $f_{aw} = A3M(f_{id}, y')$ 을 고려하며, 여기서 f_{aw} 는 요소 연령 인식 벡터이다. 목표 연령을 3에 포함시키기 위해, 레이블 정보를 네트워크에 통합하는 방법으로 사용되는 조건부 배치 정규화(Conditional Batch Normalization; CBN) 레이어를 추가한다[De Vries 외. 2017; Miyato 외. 2018; Miyato and Koyama 2018; Vincent Dumoulin 2017]. 이를 통해 A3M은 입력 신원을 위한 최적의 연령 인식 특징을 학습하여 신원 특징과 연령 속성을 모두 만족시킬 수 있다. 네트워크가 생성기에 통합됨에 따라, 엔드 투 엔드 방식으로 생성기와 함께 훈련될 수 있다. 본 발명에서는 디코더 레이어를 변조하는 데 사용되는 콤팩트 특징 벡터를 생성하는 조건부 배치 정규화 기법과 함께 다운샘플링 레이어의 집합으로 A3M을 구현한다.

디코더(Dec)(430)는 개인별 연령 특성과 함께 신원 특징을 사용하고 $x' = \text{Dec}(f_{id}, f_{aw})$ 에 의해 연령 변환된 얼굴 이미지를 생성한다. 신원 보존은 연령 특성 변화와 함께 이루어져야 한다. 따라서 이러한 두 가지 특성을 모두 만족시키기 위해 개인별 연령 특성을 사용하여 형태가 이루어지지 않은 신원 특징에 대한 변조 작업을 통해 디코딩 프로세스의 자체 지침(self-guide)을 수행한다. 신원 특징(f_{id})은 적응형 인스턴스 정규화(Adaptive Instance Normalization; AdaIN) 레이어 [Huang and Belongie 2017; Karras 외. 2019]에 의해 변조된다. 위에서 논의한 바와 같이, 신원 인코더(Enc)에서 마스킹 작업을 수행하면 특별한 연결을 통해 배경 관련 특징을 제공한다. 즉, 학습된 특징 표현을 영상 도메인에 매핑하기 전에 이를 사용하여 더 나은 시각적 인식을 가진 영상을 생성할 수 있다. 그러나 이 정보를 얼굴 관련 특징에 직접 추가하는 대신 학습된 얼굴 표현에 대해 다시 마스킹 작업을 수행하여 표준화 및 활성화 레이어로 인해 배경에서 발생한 노이즈를 제거한다. 이후, 얼굴 관련 특징과 배경 관련 특징을 결합하고 마지막 컨볼루션 레이어를 사용하여 영상 도메인에 다시 매핑한다. 본 발명의 실시예에 따르면, 다운샘플링을 업샘플링 레이어로 대체하는 인코더 아키텍처를 미러링하여 이 네트워크를 구성한다.

본 발명의 실시예에 따른 판별기(D)는 [Choi 외. 2020; Liu 외. 2019b; Mescheder 외. 2018]에 제시된 작업을 따르고 다중 작업 분류를 수행한다. 따라서, 마지막으로 완전 연결 레이어는 여러 연령 클래스를 분류하기 위해 많은 출력 분기를 가진다. 이진 분류를 수행함으로써 각 분기는 얼굴 연령 변환된 이미지가 실제 연령 도메인에 속하는지 이미지(x)인지 또는 생성기에 의해 변환된 이미지(x')인지 여부를 결정한다.

본 발명의 실시예에 따른 프레임워크는 입력 신원이 보존되고 목표 연령이 정확하게 표현되는 이미지를 생성하도록 설계되었다. 이러한 이상적인 변환을 보장하기 위해, 적대적(adversarial), 재구성(reconstruction) 및 주기 일관성(cycle-consistency)의 세 가지 유형의 손실이 사용된다. 이 프레임워크는 세 가지 입력 정보, 즉 입력 이미지(x)와 해당 연령 레이블(y)에서 작동하고, 이미지를 변환해야 하는 목표 연령 레이블(y')에서 랜덤하게 샘플링된다. 이후, 이 정보가 제공되면 생성기(G)는 다음과 같은 이미지를 생성한다:

$$x' = G(x, y'), \quad x_{rec} = G(x, y), \quad x_{cycle} = G(x', y) \quad (2)$$

여기서 x' , x_{rec} 및 x_{cycle} 은 각각 연령 변환, 재구성 및 사이클 이미지를 나타내고, 최적화 작업을 수행하는 데 사용된다.

상술된 바와 같이, 판별기의 출력은 특정 연령 영역에 해당한다. 따라서, 적대적 손실이 연령층에 의해 조건화

된다고 생각할 수 있다. 다음과 같은 적대적 손실공식을 사용한다:

$$\mathcal{L}_{adv}(G, D) = \mathbb{E}_{x, y} [\log D_y(x)] + \mathbb{E}_{x, y'} [\log (1 - D_{y'}(x'))] \quad (3)$$

여기서 $D_y(\cdot)$ 는 연령 y 에 해당하는 판별기(D)의 유일한 출력이다.

본 발명의 실시예에 따르면, 생성기(G)를 훈련하면서 입력 이미지(x)의 연령(y)과 목표 연령 y' 이 같은 연령대($y = y'$)에 속하는 경우를 고려해야 한다. 이 경우, 연령 변환된 이미지(x')는 입력 이미지(x)에 최대한 근접해야 한다. 따라서, 생성기(G)가 이러한 경우에 주의를 기울일 수밖에 없는 재구성 손실을 추가한다. 이러한 고려를 위해, 다음 공식을 이용한다.

$$\mathcal{L}_{rec}(G) = \|x - x_{rec}\|_1 \quad (4)$$

본 발명의 실시예에 따르면, 적대적 및 재구성 손실을 최소화하여 목표 연령 측면에서 정확할 뿐만 아니라 현실적인 이미지를 생성하도록 생성기(G)를 훈련할 수 있다. 그러나 이러한 최소화(즉, (3) 및 (4))는 입력 이미지(x)의 연령 특성을 변화시키면서 연령 변환 과정 전반에 걸쳐 신원을 유지하는 것을 생성기(G)가 학습하도록 한다. 이를 해결하기 위해, 주기 일관성[Kim 외. 2017; Zhu 외. 2017] 손실을 다음과 같이 포함한다:

$$\mathcal{L}_{cyc}(G) = \|x - x_{cyc}\|_1 \quad (5)$$

다음과 같이 생성기(G) 및 판별기(D)를 최적화하기 위해 고려된 전체 목표 함수는 다음과 같다:

$$\min_G \max_D \lambda_{adv} \mathcal{L}_{adv}(G, D) + \lambda_{rec} \mathcal{L}_{rec}(G) + \lambda_{cyc} \mathcal{L}_{cyc}(G) \quad (6)$$

여기서 λ_{adv} , λ_{rec} , λ_{cyc} 는 훈련에 미치는 영향의 균형을 맞추기 위한 각각의 손실에 대한 가중치이다.

본 발명의 실시예들에 따르면 연령 변환의 모든 주요 요인을 고려한 새로운 단일 프레임워크인 RAGAN(Re-Aging GAN)을 통해 생성 프로세스의 자체 지침을 수행하도록 입력 신원을 보존하는 개인 얼굴 연령 변환을 달성할 수 있다. 제안하는 RAGAN은 주어진 정체성과 목표 연령 사이의 고차 상호작용을 사용하여 개인 연령 특징을 학습할 수 있다.

이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

실시예에 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.

- [0073] 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.
- [0074] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.
- [0076] <참고자료>
- [0077] Grigory Antipov, Moez Baccouche, and Jean-Luc Dugelay. 2017. Face aging with conditional generative adversarial networks. In ICIP. 2089-2093.
- [0078] Mikolaj Biłkowski, Dougal J. Sutherland, Michael Arbel, and Arthur Gretton. 2018. Demystifying MMD GANs. In ICLR.
- [0079] Bor-Chun Chen, Chu-Song Chen, and Winston H Hsu. 2014. Cross-age reference coding for age-invariant face recognition and retrieval. In ECCV. 768-783.
- [0080] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. 2017. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587 (2017).
- [0081] Ying-Cong Chen, Xiaohui Shen, Zhe Lin, Xin Lu, I Pao, Jiaya Jia, et al. 2019. Semantic component decomposition for face attribute manipulation. In CVPR. 9859-9867.
- [0082] Yunje Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. 2018. StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation. In CVPR. 8789-8797.
- [0083] Yunje Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. 2020. StarGAN v2: Diverse image synthesis for multiple domains. In CVPR. 8188-8197.
- [0084] Harm De Vries, Florian Strub, Jeremie Mary, Hugo Larochelle, Olivier Pietquin, and Aaron C Courville. 2017. Modulating early visual processing by language. In NeurIPS. 6594-6604.
- [0085] Julien Despois, Frederic Flament, and Matthieu Perrot. 2020. AgingMapGAN (AMGAN): High-Resolution Controllable Face Aging with Spatially-Aware Conditional GANs. In ECCV.
- [0086] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016a. Deep residual learning for image recognition. In CVPR. 770-778.
- [0087] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016b. Identity mappings in deep residual networks. In ECCV. 630-645.
- [0088] Zhenliang He, Meina Kan, Shiguang Shan, and Xilin Chen. 2019. S2GAN: Share aging factors across ages and share aging trends among individuals. In ICCV. 9440-9449.
- [0089] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. 2017. Gans trained by a two time-scale update rule converge to a local nash equilibrium. In NeurIPS. 6626-6637.
- [0090] Xun Huang and Serge Belongie. 2017. Arbitrary style transfer in real-time with adaptive instance

normalization. In ICCV. 1501-1510.

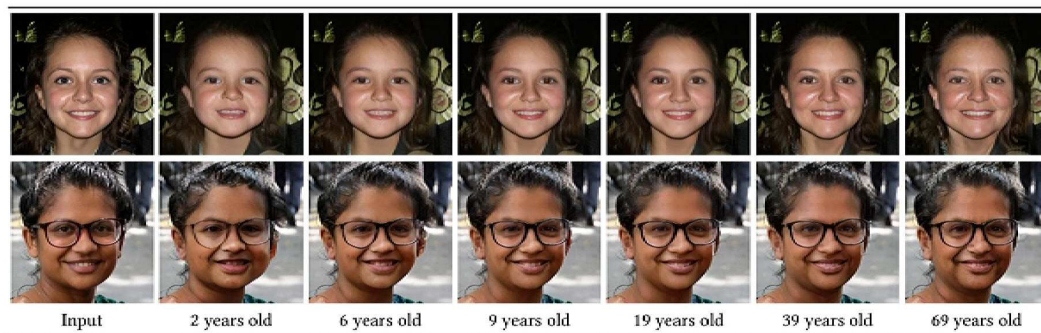
- [0091] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. 2018. Progressive growing of GANs for improved quality, stability, and variation. ICLR (2018).
- [0092] Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In CVPR. 4401-4410.
- [0093] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and improving the image quality of StyleGAN. In CVPR. 8110-8119.
- [0094] Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. 2017. Learning to discover cross-domain relations with generative adversarial networks. ICML (2017).
- [0095] Diederik P Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. ICLR (2015).
- [0096] Guillaume Lample, Neil Zeghidour, Nicolas Usunier, Antoine Bordes, Ludovic Denoyer, and Marc'Aurelio Ranzato. 2017. Fader networks: Manipulating images by sliding attributes. In NeurIPS. 5967-5976.
- [0097] Cheng-Han Lee, Ziwei Liu, Lingyun Wu, and Ping Luo. 2020. MaskGAN: Towards diverse and interactive facial image manipulation. In CVPR. 5549-5558.
- [0098] Peipei Li, Yibo Hu, Ran He, and Zhenan Sun. 2019. Global and local consistent waveletdomain age synthesis. TIFS 14, 11 (2019), 2943-2957.
- [0099] Peipei Li, Yibo Hu, Qi Li, Ran He, and Zhenan Sun. 2018. Global and local consistent age generative adversarial networks. In ICPR. 1073-1078.
- [0100] Ming Liu, Yukang Ding, Min Xia, Xiao Liu, Errui Ding, Wangmeng Zuo, and Shilei Wen. 2019a. STGAN: A unified selective transfer network for arbitrary image attribute editing. In CVPR. 3673-3682.
- [0101] Ming-Yu Liu, Xun Huang, Arun Mallya, Tero Karras, Timo Aila, Jaakko Lehtinen, and Jan Kautz. 2019b. Few-shot unsupervised image-to-image translation. In ICCV. 10551-10560.
- [0102] Yunfan Liu, Qi Li, and Zhenan Sun. 2019c. Attribute-aware face aging with waveletbased generative adversarial networks. In CVPR. 11877-11886.
- [0103] Yunfan Liu, Qi Li, Zhenan Sun, and Tieniu Tan. 2019d. A3GAN: An Attributeaware Attentive Generative Adversarial Network for Face Aging. arXiv preprint arXiv:1911.06531 (2019).
- [0104] Lars Mescheder, Sebastian Nowozin, and Andreas Geiger. 2018. Which Training Methods for GANs do actually Converge?. In ICML.
- [0105] Mehdi Mirza and Simon Osindero. 2014. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784 (2014).
- [0106] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. 2018. Spectral Normalization for Generative Adversarial Networks. In ICLR.
- [0107] Takeru Miyato and Masanori Koyama. 2018. cGANs with Projection Discriminator. In ICLR.
- [0108] Roy Or-El, Soumyadip Sengupta, Ohad Fried, Eli Shechtman, and Ira KemelmacherShlizerman. 2020. Lifespan Age Transformation Synthesis. In ECCV.
- [0109] Taesung Park, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. 2019. Semantic image synthesis with spatially-adaptive normalization. In CVPR. 2337-2346.
- [0110] Albert Pumarola, Antonio Agudo, Aleix M Martinez, Alberto Sanfeliu, and Francesc Moreno-Noguer. 2018. Ganimation: Anatomically-aware facial animation from a single image. In ECCV. 818-833.
- [0111] Shengju Qian, Kwan-Yee Lin, Wayne Wu, Yangxiaokang Liu, Quan Wang, Fumin Shen, Chen Qian, and Ran He. 2019. Make a face: Towards arbitrary high fidelity face manipulation. In ICCV. 10033-10042.
- [0112] Rasmus Rothe, Radu Timofte, and Luc Van Gool. 2018. Deep expectation of real and apparent age from a

single image without facial landmarks. IJCV 126, 2-4 (2018), 144-157.

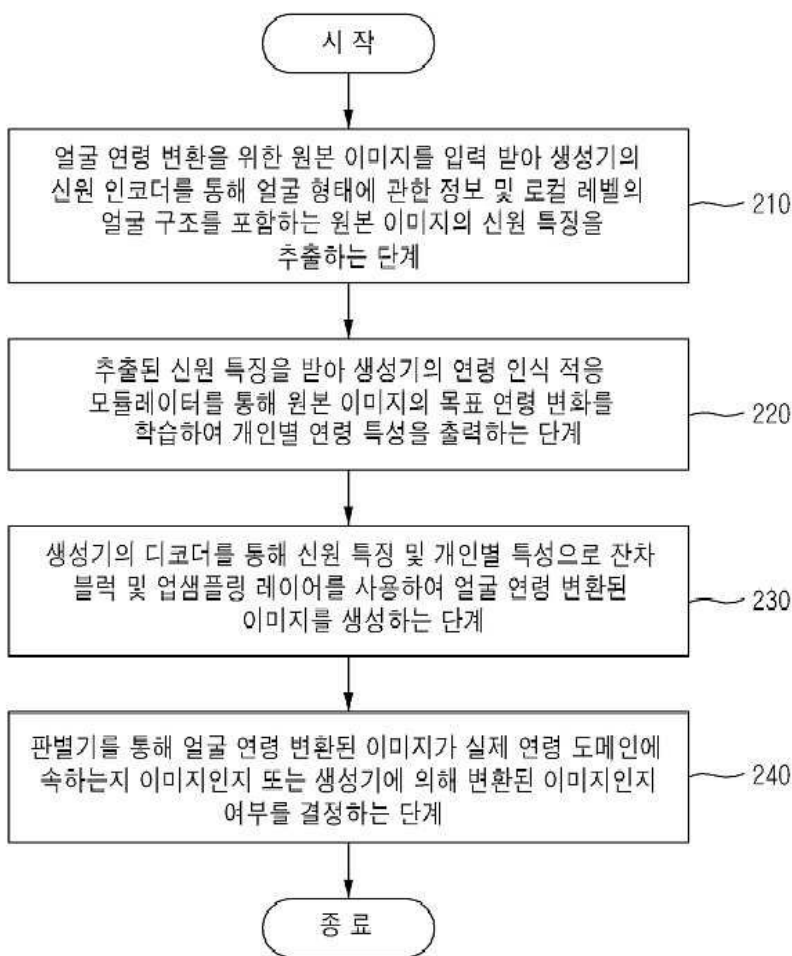
- [0113] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. In ICLR.
- [0114] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. 2016. Rethinking the inception architecture for computer vision. In CVPR. 2818-2826.
- [0115] Manjunath Kudlur Vincent Dumoulin, Jonathon Shlens. 2017. A Learned Representation For Artistic Style. In ICLR.
- [0116] Haoyi Wang, Victor Sanchez, and Chang-Tsun Li. 2020. Age-Oriented Face Synthesis with Conditional Discriminator Pool and Adversarial Triplet Loss. arXiv preprint arXiv:2007.00792 (2020).
- [0117] Zongwei Wang, Xu Tang, Weixin Luo, and Shenghua Gao. 2018. Face aging with identity-preserved conditional generative adversarial networks. In CVPR. 7939-7947.
- [0118] Taihong Xiao, Jiapeng Hong, and Jinwen Ma. 2018. Elegant: Exchanging latent encodings with GAN for transferring multiple face attributes. In ECCV. 168-184.
- [0119] Hongyu Yang, Di Huang, Yunhong Wang, and Anil K Jain. 2018. Learning face age progression: A pyramid architecture of GANs. In CVPR. 31-39.
- [0120] Xu Yao, Gilles Puy, Alasdair Newson, Yann Gousseau, and Pierre Hellier. 2020. High Resolution Face Age Editing. arXiv preprint arXiv:2005.04410 (2020).
- [0121] Zhifei Zhang, Yang Song, and Hairong Qi. 2017. Age progression/regression by conditional adversarial autoencoder. In CVPR. 5810-5818.
- [0122] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In ICCV. 2223-2232.

도면

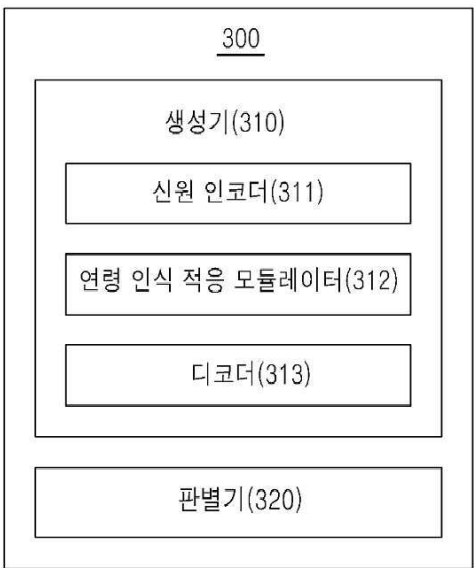
도면1



도면2



도면3



도면4

