



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년03월06일
(11) 등록번호 10-2776522
(24) 등록일자 2025년02월27일

(51) 국제특허분류(Int. Cl.)
G06T 5/80 (2024.01) G06N 3/0455 (2023.01)
G06N 3/0475 (2023.01) G06T 7/30 (2017.01)
G06V 40/16 (2022.01)
(52) CPC특허분류
G06T 5/80 (2024.01)
G06N 3/0455 (2023.01)
(21) 출원번호 10-2024-0049597
(22) 출원일자 2024년04월12일
심사청구일자 2024년04월12일
(56) 선행기술조사문헌
KR1020230094951 A*
KR1020150116641 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
국립한밭대학교 산학협력단
대전광역시 유성구 동서대로 125 (덕명동)
(72) 발명자
이승호
대전광역시 유성구 엑스포로123번길 26-30 스마트
시티리버뷰오피스텔 102동 601호
이희열
대전광역시 유성구 월드컵대로290번길 39, 목련아
파트 101동 1601호
(74) 대리인
이은철, 이우영

전체 청구항 수 : 총 3 항

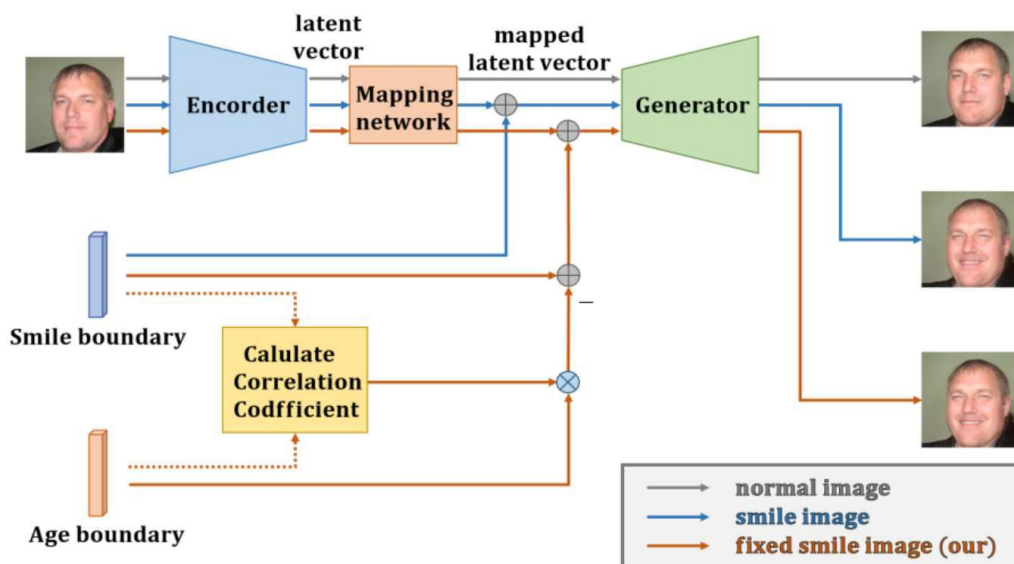
심사관 : 김광식

(54) 발명의 명칭 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템

(57) 요약

본 발명은 표정 이미지 생성 과정에서 발생하는 연령 왜곡 현상을 표정 boundary와 연령 boundary간의 상관계수를 이용하여 연령 왜곡이 개선된 표정 이미지를 생성하는 SyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템에 관한 것으로, 표정 이미지 생성 과정에서 발생하는 연령 왜곡 현상을 표정 boundary와 연령 boundary간의 상관계수를 이용하여 연령 왜곡이 개선된 표정 이미지를 생성하는 효과가 있다. 또한 기존의 표정 이미지 생성에서 발생하는 연령 왜곡 현상을 상관계수를 이용하여 효과적으로 감소시킴으로써, 생성된 이미지의 자연스러움과 신뢰성을 향상시킨다.

대표도 - 도2



(52) CPC특허분류

G06N 3/0475 (2023.01)

G06T 7/30 (2017.01)

G06V 40/174 (2022.01)

G06T 2207/20084 (2013.01)

G06T 2211/441 (2025.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2710013554

과제번호 2022R1F1A1066371

부처명 과학기술정보통신부

과제관리(전문)기관명 한국연구재단

연구사업명 개인기초연구(과기정통부)

연구과제명 메타버스 콘텐츠에서 사용할 수 있는 음성에 따른 자연스러운 감정 표현이 가능한

3D 얼굴 자동 생성 솔루션 개발

과제수행기관명 국립한밭대학교

연구기간 2024.03.01 ~ 2025.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호 1345370812

과제번호 2021RIS-004

부처명 교육부

과제관리(전문)기관명 한국연구재단

연구사업명 지자체대학협력기반지역혁신사업

연구과제명 (대전세종충남지역혁신플랫폼)충남대학교

과제수행기관명 충남대학교

연구기간 2024.03.01 ~ 2025.02.28

공지예외적용 : 있음

명세서

청구범위

청구항 1

실제 얼굴 이미지로부터 잠재 벡터를 추출하는 인코더(100),

상기 인코더로부터 얻은 잠재 벡터를 매핑하는 매핑 네트워크(200),

표정변화와 연령 왜곡 간의 상관계수를 계산하고, 상기 상관계수를 통해 표정변화에 따른 연령 왜곡을 최소화하기 위해 상기 잠재 벡터를 조정하는 왜곡 감소부(300) 및

매핑된 잠재 벡터와 상관계수를 적용한 잠재 벡터 조정을 통해 생성된 데이터를 이용하여 연령 왜곡이 감소된 최종 표정 이미지를 생성하는 생성자(400)를 포함하며,

상기 왜곡 감소부는 StyleGAN ENcoder와 mapping network를 통해 생성된 잠재 벡터에 smile boundary를 적용하여 상관계수에 비례한 중속성을 활용하여 smile boundary와 age boundary의 상관계수를 계산하고, 상기 상관계수를 적용한 잠재 벡터의 age boundary 특성을 편집하는 것을 특징으로 하는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템.

청구항 2

제1항에 있어서,

상기 왜곡 감소부는 표정 변화를 일으키는 잠재 벡터의 방향과 연령 변화를 일으키는 잠재 벡터의 방향 간의 상관관계를 조사하여 상관계수를 추정하여 상기 잠재 벡터를 조정하는 것을 특징으로 하는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템.

청구항 3

삭제

청구항 4

StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템을 이용한 방법에 있어서,

(S1)상기 시스템이 실제 얼굴 이미지를 입력으로 받아, 해당 이미지를 인코더를 통해 잠재 공간의 벡터로 변환하는 StyleGAN Encoder를 사용한 잠재 벡터 추출단계,

(S2)상기 시스템이 인코더로부터 얻은 잠재 벡터를 매핑하는 Mapping Network를 통한 잠재 벡터 맵핑단계,

(S3)상기 시스템이 Smile Boundary와 Age Boundary의 상관계수 추정단계,

(S4)상기 시스템이 Smile Boundary와 Age Boundary의 상관계수를 적용한 잠재 벡터 튜닝 조절단계 및

(S5)상기 시스템이 생성자(Generator)를 통한 연령 왜곡이 감소된 표정 이미지 생성단계를 포함하며,

상기 시스템이 StyleGAN ENcoder와 mapping network를 통해 생성된 잠재 벡터에 smile boundary를 적용하여 상관계수에 비례한 중속성을 활용하여 smile boundary와 age boundary의 상관계수를 계산하고, 상기 상관계수를 적용한 잠재 벡터의 age boundary 특성을 편집하는 것을 특징으로 하는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소방법.

발명의 설명

기술 분야

[0001] 본 발명은 표정 이미지 생성에서의 연령 왜곡 감소 시스템에 관한 것으로, 표정 이미지 생성 과정에서 발생하는 연령 왜곡 현상을 개선하는 SyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템에 관한 것이다.

배경 기술

- [0002] 최근 생성형 딥러닝 모델 기술의 발전으로 이미지 생성 및 특성 편집에 대한 연구가 주목받고 있다.
- [0003] 특히 얼굴 이미지의 표정 및 연령을 수정하는 작업이 가능해지면서 컴퓨터 비전 및 인공지능 분야에서 중요한 주제로 인식되고 있다. 영화산업 및 영상매체 산업에서도 해당 기술은 큰 관심을 받고 있다.
- [0004] 이전까지는 영상 제작이 주로 수작업으로 이루어졌으나, GAN(Generative Adversarial Networks)의 발전과 생성형 딥러닝 모델의 도입으로 인해 영상 제작의 속도와 품질이 현저히 향상되었다.
- [0005] 특히 StyleGAN을 활용한 이미지 생성 기술은 고품질의 이미지를 생성하는 데 큰 기여를 하였다. 그러나 이러한 기술을 통한 표정 편집 시에 연령 왜곡 현상이 발생한다.
- [0006] 이는 표정의 변화에 따라 생기는 얼굴의 주름 등이 영향을 주어 표정에 대한 학습시 연령에 영향을 주는 요소가 같이 학습되기 때문이다.
- [0007] 특히, 웃는 표정에서 생기는 얼굴의 주름이 영향을 크게 주어 연령 왜곡이 심해지는 문제가 발생하고 있다. 이러한 상호작용에 대한 측면을 해결하기 위해 표정 이미지 생성 시 발생하는 연령 왜곡 현상을 개선할 필요가 있다.

선행기술문헌

특허문헌

- [0008] (특허문헌 0001) 공개특허 2015-0116641(2015.10.16)

발명의 내용

해결하려는 과제

- [0009] 본 발명은 상술한 문제를 해결하고자, 표정 이미지 생성 과정에서 발생하는 연령 왜곡 현상을 표정 boundary와 연령 boundary간의 상관계수를 이용하여 연령 왜곡이 개선된 표정 이미지를 생성하는 SyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템을 제안함에 있다.

과제의 해결 수단

- [0010] 본 발명에 따르면, StyleGAN Encoder를 사용하여 실제 얼굴 이미지로부터 잠재 벡터를 추출하는 인코더(100), 인코더로부터 얻은 잠재 벡터를 StyleGAN의 생성자(generator)가 이해할 수 있는 형태로 변환하는 매핑 네트워크(200), 표정변화와 연령 왜곡 간의 상관계수를 계산하고, 상기 상관계수를 통해 표정변화에 따른 연령 왜곡을 최소화하기 위해 상기 잠재 벡터를 조정하는 왜곡 감소부(300) 및 매핑된 잠재 벡터와 상관계수를 적용한 잠재 벡터 조정을 통해 생성된 데이터를 이용하여 연령 왜곡이 감소된 최종 표정 이미지를 생성하는 생성자(400)를 포함한다.
- [0011] 왜곡 감소부는 표정 변화를 일으키는 잠재 벡터의 방향과 연령 변화를 일으키는 잠재 벡터의 방향 간의 상관관계를 조사하여 상관계수를 추정하여 상기 잠재 벡터를 조정한다.
- [0012] 왜곡 감소부는 StyleGAN ENcoder와 mapping network를 통해 생성된 잠재 벡터에 smile boundary를 적용하여 상관계수에 비례한 종속성을 활용하여 두 boundary의 상관계수를 계산하고, 상기 상관계수를 적용한 잠재 벡터의 age boundary 특성을 편집한다.
- [0013] 한편, StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템을 이용한 방법에 있어서, (S1)상기 시스템이 실제 얼굴 이미지를 입력으로 받아, 해당 이미지를 인코더를 통해 잠재 공간의 벡터로 변환하는 StyleGAN Encoder를 사용한 잠재 벡터 추출단계, (S2)상기 시스템이 인코더로부터 얻은 잠재 벡터를 매핑하는 Mapping Network를 통한 잠재 벡터 맵핑단계, (S3)상기 시스템이 Smile Boundary와 Age Boundary의 상관계수 추정단계, (S4)상기 시스템이 Smile Boundary와 Age Boundary의 상관계수를 적용한 잠재 벡터 튜닝 조절단계 및 (S5)상기 시스템이 생성자(Generator)를 통한 연령 왜곡이 감소된 표정 이미지 생성단계를 포함한다.

발명의 효과

- [0014] 본 발명에 의하면, 표정 이미지 생성 과정에서 발생하는 연령 왜곡 현상을 표정 boundary와 연령 boundary간의 상관계수를 이용하여 연령 왜곡이 개선된 표정 이미지를 생성하는 효과가 있다.
- [0015] 기존의 표정 이미지 생성에서 발생하는 연령 왜곡 현상을 상관계수를 이용하여 효과적으로 감소시킴으로써, 생성된 이미지의 자연스러움과 신뢰성을 향상시킨다.
- [0016] 그리고 영화 및 미디어 제작 분야에서 특정 인물의 표정을 자연스럽게 편집하여 수작업의 비율을 줄이고 시간적, 인적 자원을 확보할 수 있다.

도면의 간단한 설명

- [0017] 도 1은 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 구성도이다.
- 도 2는 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 개요도이다.
- 도 3은 본 실시예에 따른 smile boundary를 적용 smile 이미지 생성 구성도이다.
- 도 4는 본 실시예에 따른 기존 smile boundary를 적용하여 추출한 이미지이다.
- 도 5는 본 실시예에 따른 상관계수를 적용한 개선된 smile 이미지 생성 구성도이다.
- 도 6은 본 실시예에 따른 상관계수를 적용한 개선된 smile 이미지이다.
- 도 7은 본 실시예에서 제안된 기법의 실행과정 흐름도이다.
- 도 8은 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 데이터 셋을 나타낸 도면이다.
- 도 9는 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 제안한 방법을 통한 이미지 비교를 나타낸 것이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템은 표정 이미지 생성시 발생하는 연령 왜곡 현상의 관계성을 찾기 위해 표정 이미지 생성 boundary와 연령 이미지 생성 boundary 간의 상관관계를 조사하여 상관계수를 추정한다.
- [0019] 이를 통해 연령 이미지 생성 boundary와 표정 이미지 생성 boundary 간의 관계를 더 잘 이해할 수 있다. 이를 통해 연령 왜곡을 줄이고 더 현실적인 표정 이미지를 생성할 수 있다.
- [0020] 그리고 본 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템은 표정 이미지를 생성시 표정 이미지 생성 boundary와 연령 이미지 생성 boundary 간의 상관계수를 적용하여 연령 왜곡을 감소 및 개선할 수 있다.
- [0021] 이하, 첨부된 도면을 참조하여 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템을 자세히 설명하기로 한다.
- [0022] 도 1은 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 구성도이다.
- [0023] 도 1에 도시된 바와 같이, 본 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템은 인코더(100), 매핑 네트워크(200), 왜곡 감소부(300), 생성자(400)를 포함한다.
- [0024] 인코더(100)는 StyleGAN Encoder를 사용하여 실제 얼굴 이미지로부터 잠재 벡터를 추출한다. 이 과정에서 원본 이미지의 세밀한 특징들이 잠재 공간으로 인코딩되며, 이를 통해 이미지의 고유한 스타일과 특성을 잘 보존하여 생성되는 이미지의 품질을 높인다.
- [0025] 매핑 네트워크(200)는 인코더로부터 얻은 잠재 벡터를 매핑한다. 이 매핑된 벡터는 생성자에 의해 최종적으로

이미지로 변환되기 전에, 얼굴의 다양한 특성들을 조절할 수 있는 정보를 담고 있다. 이러한 특성들은 머리카락 스타일, 미소의 정도, 눈의 크기 등으로 구성된다.

- [0026] 이러한 매핑 네트워크는 인코더로부터 얻은 잠재 벡터를 StyleGAN의 생성자(generator)가 이해할 수 있는 형태로 변환한다.
- [0027] 참고로, StyleGAN의 생성자(generator)가 이해할 수 있는 형태란, StyleGAN 아키텍처 내에서 생성 과정에 사용될 수 있는 잠재 공간의 표현을 의미한다. StyleGAN은 기본적으로 두 가지 잠재 공간을 다룬다.
- [0028] Z 공간 (Z space)은 초기 잠재 벡터가 존재하는 공간으로, 보통 정규 분포로부터 무작위로 샘플링된 벡터들의 집합이다. 이 벡터들은 StyleGAN의 입력으로 처음 사용된다.
- [0029] W 공간 (W space)과 관련하여, 매핑 네트워크는 Z 공간의 벡터를 W 공간으로 변환한다.
- [0030] W 공간의 벡터들은 보다 분산된(disentangled) 특성을 갖기 때문에 각각의 차원이 얼굴의 특정 특성을 보다 명확하게 조절할 수 있다. 예를 들어, 한 차원은 미소의 강도를 조절하고, 다른 차원은 눈의 크기나 모양을 조절하는 등의 역할을 한다.
- [0031] 매핑 네트워크의 주요 목적은 Z 공간에서 추출된 벡터를 W 공간의 더 복잡하고 조절 가능한 표현으로 변환하는 것이다. 이렇게 변환된 W 공간의 벡터는 생성자(generator)가 사용하여 실제와 같은 세밀한 이미지를 생성할 수 있게 한다.
- [0032] W 공간의 특징은 각 차원이 서로 덜 의존적이기 때문에, 조절이 보다 세밀하고 특정 특성에 집중할 수 있도록 한다. StyleGAN에서는 더 나아가 W+ 공간도 도입되었는데, 여기서 각각의 생성자 레이어가 서로 다른 W 공간의 벡터를 받아 더 세밀한 조절이 가능하다. 이는 이미지의 각 부분(예: 헤어 스타일, 안경, 배경 등)을 독립적으로 조절할 수 있는 능력을 제공한다.
- [0033] 매핑 네트워크는 인코더로부터 얻은 초기 잠재 벡터 Z를 가져와, 더 복잡한 W 공간의 잠재 벡터로 변환한다. 이 변환은 벡터를 더 고차원적이고, 제어 가능한 형태로 만든다.
- [0034] 특성분산(Disentanglement)과 관련하여, W 공간의 벡터는 각 차원이 특정 얼굴 특성을 조절하도록 만들어져, 스타일 조절이 더 세밀하게 가능해진다. 이를 통해 생성자가 조절하기 쉽고, 특성 간 상호작용이 줄어들어 예측 가능한 이미지 편집이 가능해진다.
- [0035] 스타일 조절과 관련하여, 매핑 네트워크를 거친 W 공간의 벡터는 이미지의 스타일, 즉 헤어스타일, 피부 질감, 조명 등의 고차원적인 스타일 속성을 조절하는 데 사용된다. 이러한 과정을 통해, 매핑 네트워크는 생성자가 다양한 얼굴 이미지를 자연스럽게 세밀하게 생성할 수 있는 기반이 된다. 이는 StyleGAN의 특징적인 부분으로, 실제와 같은 고품질의 이미지 생성에 중요한 역할을 한다.
- [0036] 왜곡 감소부(300)는 경계 사이의 상관계수를 계산하는 구성으로, 표정변화와 연령 왜곡 간의 상관계수를 계산하고, 상기 상관계수를 통해 표정변화에 따른 연령 왜곡을 최소화하기 위해 상기 잠재 벡터를 조정한다.
- [0037] 이러한 왜곡 감소부는 표정 변화(예: 미소)와 연령 왜곡 간의 상관관계를 조사하고, 이를 상관계수로 계산한다. 이 상관계수를 통해 표정 변화에 따른 연령 왜곡을 최소화하기 위해 잠재 벡터를 조정한다.
- [0038] 특히, smile boundary와 age boundary의 상관계수를 사용하여, 미소 시 나타나는 주름 같은 연령 관련 특징들이 왜곡되지 않도록 잠재 벡터를 세심하게 조절한다.
- [0039] 참고로, 왜곡 감소부(300), 즉 연령 왜곡을 감소시키는 부분의 기능은 StyleGAN을 사용하는 얼굴 이미지 생성 과정에서 표정 변화가 연령 왜곡을 일으키는 문제를 최소화한다.
- [0040] Boundary 상관관계 파악과 관련하여, 표정 변화(예: 미소)를 일으킬 때와 연령 변화가 이미지에 미치는 영향을 분석하여, 두 변화가 서로 얼마나 연관되어 있는지를 파악한다. 이를 위해 표정과 연령을 각각 조절하는 방향을 정의하는 boundary(경계)를 설정한다.
- [0041] 상관계수 계산과 관련하여, 표정 boundary와 연령 boundary 사이의 상관계수를 계산하여, 표정 변경이 연령 특성에 미치는 영향의 정도를 수치적으로 측정한다. 상관계수는 두 변수 사이의 선형 관계를 나타내며, 이 값에 따라 표정 변화 시 연령 왜곡을 조정할 수 있다.
- [0042] 잠재 벡터 편집과 관련하여, 상관계수를 이용하여 잠재 벡터를 조정한다. 이 조정은 표정 변경 시 연령 왜곡을 감소시키기 위해 이루어진다. 예를 들어, 미소를 짓는 표정을 만들 때 나타날 수 있는 주름이나 기타 연령 관련

특징이 과도하게 표현되지 않도록 조정하는 것이다.

- [0043] 본 실시예에 따른 왜곡 감소부는 표정 변화를 일으키는 잠재 벡터의 방향과 연령 변화를 일으키는 잠재 벡터의 방향 간의 상관관계를 조사하여 상관계수를 추정하여 상기 잠재 벡터를 조정한다.
- [0044] 이러한 왜곡 감소부는 StyleGAN ENcoder와 mapping network를 통해 생성된 잠재 벡터에 smile boundary를 적용하여 상관계수에 비례한 중속성을 활용하여 두 boundary의 상관계수를 계산하고, 상기 상관계수를 적용한 잠재 벡터의 age boundary 특성을 편집한다.
- [0046] 생성자(400)는 매핑된 잠재 벡터와 상관계수를 적용한 잠재 벡터 조정을 통해 생성된 데이터를 이용하여 연령 왜곡이 감소된 최종 표정 이미지를 생성한다.
- [0047] 생성자는 여러 층(layer)을 통해 이 잠재 벡터를 실제 이미지로 변환하며, 각 층에서는 다양한 스타일의 조정이 이루어진다. AdaIN(Adaptive Instance Normalization)과 같은 기법을 사용하여 입력된 스타일 특성을 각 이미지에 적응적으로 적용하고, 스타일 변화를 가능하게 한다.
- [0048] 참고로, AdaIN(Adaptive Instance Normalization)과 관련하여, 생성자는 각 레이어에 AdaIN을 적용하여, 잠재 벡터에서 가져온 스타일 정보를 기반으로 피쳐 맵의 정규화를 진행한다. 이 과정은 스타일 정보를 적응적으로 이미지에 적용하여 더 다양하고 세밀한 스타일 변화를 생성하는데 중요하다. 또한 왜곡 감소부를 통해 조정된 잠재 벡터는 표정의 변화가 연령의 특성에 미치는 부정적인 영향을 최소화한다.
- [0049] 이 네 부분은 통합적으로 작동하여, 원본 이미지에 비해 자연스러운 미소를 더하면서도 인물의 연령적 특성을 잘 유지한 새로운 얼굴 이미지를 생성한다.
- [0051] 도 2는 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 개요도이다.
- [0052] 본 발명에서 제안하는 기법의 개요는 도 2와 같다.
- [0053] 본 발명에서 제안하는 기법은 StyleGAN Encoder를 이용하여 표정 이미지를 생성할 때 사용되는 표정 boundary에 연령 boundary와의 상관계수를 이용하여 연령 왜곡을 감소시키는 시스템을 구축한다.
- [0054] 제안된 기법은 크게 4가지 과정으로 구성되며 각 과정은 StyleGAN Encoder를 이용한 잠재 벡터 추출, mapping network를 통한 잠재 벡터 맵핑, 표정 boundary와 연령 boundary의 상관계수를 적용한 잠재 벡터 편집, generator를 통한 연령 왜곡이 감소된 표정 이미지 생성으로 이루어진다.
- [0055] 먼저 첫 번째로, StyleGAN Encoder를 이용하여 잠재 벡터를 추출한다. StyleGAN Encoder는 입력받은 얼굴 이미지로부터 세밀한 특징을 반영한 잠재 벡터를 추출한다. 이 과정에서 이미지의 고유한 스타일과 특성이 잘 보존되어, 생성되는 이미지의 품질을 높일 수 있다.
- [0056] 두 번째로, mapping network를 통하여 잠재 벡터를 맵핑한다. 각 차원이 다양한 얼굴 특징을 담당하게 되어, 머리카락 스타일, 미소의 정도, 눈의 크기 등을 자연스럽게 표현할 수 있다. 이러한 스타일 특성의 맵핑은 Generator의 각 레이어에서 조절되어 최종적으로 고품질의 얼굴 이미지를 생성하는 데 기여한다.
- [0057] 세 번째로, 표정 boundary와 연령 boundary의 상관계수를 활용하여 잠재 벡터를 편집한다. 특히 웃는 표정에서 발생하는 주름 등의 연령 왜곡 현상을 완화하기 위해 두 boundary 간의 상관관계를 고려하여 표정의 특성을 조절한다. 이러한 과정을 통해 생성된 이미지에서 표정과 연령 간의 상호작용이 미세하게 조절된다.
- [0058] 마지막으로 generator를 통하여 연령 왜곡이 감소된 표정 이미지를 생성한다. Generator의 각 레이어에서 스타일이 조절되며, 이러한 과정을 통해 고품질의 얼굴 이미지를 생성된다. StyleGAN의 Generator는 다계층 구조를 가지고 있으며, Mapping Network를 통해 맵핑된 잠재 벡터를 입력으로 받아 이미지를 생성한다.
- [0059] 이때 이미지의 다양성을 높이기 위해 각 레이어에서는 노이즈가 입력되며 이는 스타일의 변화를 증가시키고 디테일한 특성을 부여한다. 높은 차원의 스타일 특성은 각 레이어에서 일정 부분에 전파되어 전체적인 스타일을 형성한다.
- [0060] 또한 AdaIN이라는 정규화 기법이 각 레이어에서 사용된다. 이 기법은 입력된 스타일 특성을 이미지에 적응적으로 적용하여 스타일의 다양한 변화를 가능하게 한다.
- [0062] 상관계수를 적용하여 표정 이미지 생성에서의 연령 왜곡 개선 과정을 설명하면 다음과 같다.

- [0063] 기존의 표정 이미지 생성과정과 관련하여, 표정 이미지를 생성하기 위해서는 StyleGAN Encoder를 통해 추출한 잠재 벡터를 mapping network를 거친 후에 특성 편집이 필요하다.
- [0064] 원하는 특성을 강조하거나 억제하는 등의 편집을 위해 InterFaceGAN을 사용한 추가적인 학습이 필요하다.
- [0065] 먼저, 추출한 잠재 벡터와 특성값을 데이터셋으로 구축한다. 이후, 구축된 데이터셋은 특성값에 따라 Positive 클래스와 Negative 클래스로 나뉘어진다.
- [0066] 여기서 Positive 클래스는 특성값이 있는 데이터로 구성되고, Negative 클래스는 특성값이 없는 데이터로 구성된다.
- [0067] Positive와 Negative로 나누어진 데이터는 SVM(Support Vector Machine)의 이진 분류 학습을 통해 두 클래스로 나누어진다.
- [0068] SVM은 분류 및 회귀 문제에 활용되는 지도 학습 알고리즘으로, 데이터를 고차원 공간에 매핑하여 서로 다른 클래스를 나누는 초평면을 계산한다.
- [0069] 이 초평면은 각 클래스 간의 최대 마진을 갖도록 학습되며, 분류 문제에서 클래스 간의 경계를 형성한다.
- [0070] 분류 학습이 완료되면 학습된 특성 boundary 데이터가 생성된다.
- [0071] 이후, Generator에 입력되는 잠재 벡터에 해당 boundary 데이터를 추가함으로써, 기본 이미지에서 특성값이 증가한 이미지가 출력된다.
- [0072] 이 과정을 통해 이미지의 원하는 특성을 더욱 강조하거나 조절할 수 있다.
- [0073] 도 3은 본 실시예에 따른 smile boundary를 적용 smile 이미지 생성 구성도이다.
- [0075] 표정 이미지 생성과 연령 이미지 생성 간의 상관관계를 조사하여 상관계수 추정과 관련하여 설명하면 다음과 같다.
- [0076] 도 4는 본 실시예에 따른 기존 smile boundary를 적용하여 추출한 이미지이다. (b)는 기본 출력 이미지이고, (a)는 boundary 음의 값을 주어 smile 특성을 줄였고, (c)는 양의 값을 주어 smile 특성을 높였다.
- [0077] 해당 이미지를 관찰하였을 때 상대적으로 (a)는 (b)보다 나이가 적게, (c)는 (b)보다 나이가 많게 출력되었다. 이러한 현상은 웃음 때 발생하는 얼굴의 주름으로 인해 age 특성이 smile 특성의 학습에 영향을 주었다고 판단된다.
- [0078] 따라서, 상관계수를 적용하여 표정 이미지 생성에서의 연령 왜곡을 개선하기 위하여 smile boundary와 age boundary의 상관계수를 이용하여 값을 조정한다.
- [0079] 상관계수는 두 변수 간 상관관계의 정도를 나타내는 수치로, n차원의 두 변수 x , y 가 주어질 때 상관계수 r 은 다음의 식과 같다. 단, $\bar{x}$, $\bar{y}$ 는 각각 x 와 y 의 평균을 나타내고, S_{xx} , S_{yy} 는 각각 x 와 y 의 분산, S_{xy} 는 x 와 y 의 공분산이다.

수학식 1

$$r = \frac{S_{xy}}{\sqrt{S_{xx}} \cdot \sqrt{S_{yy}}}$$

$$\bar{x} = \frac{1}{n} \sum (x_i)$$

$$\bar{y} = \frac{1}{n} \sum (y_i)$$

$$S_{xx} = \sum (x_i - \bar{x})^2$$

$$S_{yy} = \sum (y_i - \bar{y})^2$$

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

[0080]

[0081] 상관계수 r 은 -1과 1 사이의 값이다. -1은 완벽한 음의 상관관계를, 1은 완벽한 양의 상관관계를 나타낸다. 0은 두 변수가 완벽하게 독립(수직)인 것을 의미한다. 높은 상관관계는 두 변수 사이의 종속성이 높아 한 변수의 변화가 다른 변수에 영향을 미친다는 것을 나타낸다.

[0083] 표정과 연령 간의 상관계수를 적용하여 표정 이미지 생성에서의 연령 왜곡 개선에 대해 설명하면 다음과 같다.

수학식 2

$$\overrightarrow{latent'} = \overrightarrow{latent} + \overrightarrow{BDRY_{smile}}$$

[0084]

수학식 3

$$\overrightarrow{latent'} = \overrightarrow{latent} + \overrightarrow{BDRY_{smile}} - r \cdot \overrightarrow{BDRY_{age}}$$

[0085]

[0086] 기존의 표정 이미지 생성을 위해서는 [수식 2]와 같이 잠재 벡터의 특성을 편집한다. StyleGAN ENcoder와 mapping network를 통해 생성된 잠재 벡터에 smile boundary를 적용하여 상관계수 r 에 비례한 종속성을 활용하여 두 boundary의 상관계수를 계산하고, [수식 3]을 사용하여 잠재 벡터의 특성을 편집한다.

[0087] 수식 2는 기존의 잠재 벡터 특성 편집이고, 수식 3은 상관계수를 적용한 잠재 벡터 특성 편집을 나타낸다.

[0088] $\overrightarrow{latent}$ 는 입력 이미지에서 추출된 잠재 벡터이며, $\overrightarrow{BDRY_{smile}}$ 와 $\overrightarrow{BDRY_{age}}$ 는 각각 smile boundary와 age boundary를 나타낸다. 위 수식을 통해 smile boundary로 인한 age 증가 현상을 완화 할 수 있다.

[0089] 도 5는 본 실시예에 따른 상관계수를 적용한 개선된 smile 이미지 생성 구성도이다.

[0090] 도 5에서, 인코더(Encoder)는 사람 얼굴 이미지를 입력으로 받아서 잠재 벡터(latent vector)로 변환한다. 잠재 벡터는 이미지의 핵심 특징을 수치적으로 나타내는 벡터이다.

[0091] 매핑 네트워크(Mapping network)는 인코더로부터의 잠재 벡터를 받아서, 이를 StyleGAN의 생성 네트워크가 이해할 수 있는 형태로 변환하는 매핑된 잠재 벡터(mapped latent vector)로 변환한다.

- [0092] 생성자(Generator)는 매핑된 잠재 벡터를 사용하여 변형된 얼굴 이미지(예를 들어 미소 짓는 표정)를 생성한다.
- [0093] 본 실시예에서 제안하는 표정 이미지 생성 경계(Smile boundary)와 연령 이미지 생성 경계(Age boundary)를 설명하면 다음과 같다.
- [0094] 본 실시예에서는 이 두가지 주요 경계를 사용한다.
- [0095] 표정 이미지 생성 경계(Smile boundary)는 미소와 같은 특정 표정을 생성하는 데 사용되는 방향성을 나타낸다.
- [0096] 연령 이미지 생성 경계(Age boundary)는 연령 변화를 생성하는 데 사용되는 방향성을 나타낸다.
- [0097] 이 두 경계 사이의 상관관계를 조사하여 상관계수(Correlation Coefficient)를 계산한다. 상관계수는 두 변수 사이의 선형 관계의 정도를 수치적으로 나타내며, 이 경우에는 표정 변화가 연령 왜곡에 미치는 영향을 측정한다.
- [0098] 이 시스템은 표정을 변경할 때 발생할 수 있는 연령 왜곡을 감소시키기 위해, 표정 이미지 생성시 발생하는 연령 왜곡 현상의 관계성을 찾는다.
- [0099] 상관계수를 추정함으로써 연령 변화와 표정 변화 사이의 관계를 더 잘 이해하고, 이를 바탕으로 더 현실적인 표정 변화를 달성할 수 있다. 또한, 상관계수를 적용함으로써 연령 왜곡을 줄여 결과 이미지의 품질을 향상시킬 수 있다.
- [0100] 이 도식에서 상단의 이미지는 원본 얼굴 이미지를 나타내고, 하단의 이미지는 미소를 짓는 표정으로 변환된 이미지를 나타낸다. 이러한 변환 과정에서 연령 왜곡을 최소화하기 위해 상관계수를 계산하고 적용하는 것이 핵심이다.
- [0102] 도 6은 본 실시예에 따른 상관계수를 적용한 개선된 smile 이미지이다.
- [0103] 도 6의 (a)는 기존의 smile boundary를 적용하여 생성한 이미지이고, (b)는 제안하는 기법을 적용하여 개선된 이미지이다. (c)는 두 이미지의 차이를 강조하기 위해 차연산을 수행 후 노출을 높인 이미지이다. (a)와 (b)를 자세히 비교하면, (a)가 (b)에 비해 노화에 의한 흰머리와 더 깊은 주름 및 더 거친 피부결 등을 확인할 수 있다. 이 차이는 (c)를 통해 명확하게 확인할 수 있다. 특히, 눈가의 주름이 크게 줄었음을 (c)를 통해 확인할 수 있다.
- [0104] 도 7은 본 실시예에서 제안된 기법의 실행과정 흐름도이다.
- [0105] 실행과정에서 StyleGAN Encoder로 미리 학습된 ReStyle-pSp 모델을 사용하였고 내부의 generator 모델로 stylegan3가 사용한다. 또한 boundary 역시 interfaceGAN의 SVM을 통해 미리 학습된 데이터를 사용한다.
- [0106] 도 7에서 StyleGAN Encoder를 사용한 잠재 벡터 추출(S1) 과정에서, StyleGAN Encoder(예: ReStyle-pSp 모델)는 실제 얼굴 이미지를 입력으로 받아, 해당 이미지를 잠재 공간의 벡터로 변환한다. 이 잠재 벡터는 얼굴의 다양한 특징(예: 표정, 연령 등)을 포함하고 있는 다차원의 데이터 포인트이다.
- [0107] Mapping Network를 통한 잠재 벡터 맵핑(S2)과정에서, 매핑 네트워크는 잠재 벡터를 StyleGAN이 이해할 수 있는 형태로 변환한다. 이 과정은 벡터를 더욱 복잡한 형태로 만들어 생성자(Generator)가 더욱 세밀한 이미지를 생성할 수 있도록 한다.
- [0108] Smile Boundary와 Age Boundary의 상관계수 추정(S3)에서, Smile Boundary는 표정을 조절하는 방향을 나타내며, Age Boundary는 연령을 조절하는 방향을 나타낸다. 이 단계에서는 이 두 경계 사이의 상관계수를 계산하여, 표정 변경 시 연령 변화의 영향을 평가한다.
- [0109] Smile Boundary와 Age Boundary의 상관계수를 적용한 잠재 벡터 튜닝 조절(S4)에서, 계산된 상관계수를 사용하여 잠재 벡터를 조절함으로써, 표정 변화가 연령 왜곡에 미치는 영향을 최소화한다. 이는 SVM(Support Vector Machine)을 사용하는 InterfaceGAN 방법을 응용할 수 있다.
- [0110] Generator를 통한 연령 왜곡이 감소된 표정 이미지 생성(S5)에서, 상관계수로 조절된 잠재 벡터를 사용하여, 내부적인 generator 모델(예: StyleGAN3)이 연령 왜곡 없이 표정이 변한 이미지를 생성한다. 이 단계에서 최종 이미지가 생성되며, 여기서 목표는 원본 이미지의 인물이 자연스러운 미소를 짓는 동시에 연령이 왜곡되지 않도록 하는 것이다.
- [0111] 이 과정은 StyleGAN과 SVM을 이용하여 얼굴 이미지의 표정을 변경하면서도 연령을 왜곡시키지 않도록 한다. 연

령 왜곡을 방지하기 위해 표정과 연령의 상관계수를 산출하고, 이를 통해 잠재 벡터를 조정함으로써, 인공지능이 보다 실제와 가까운 얼굴 이미지를 생성할 수 있게 한다.

- [0113] 도 8은 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 데이터 셋을 나타낸 도면이다.
- [0114] 본 실시예에 따른 데이터 셋과 관련하여, 본 발명에서는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 기법의 성능 검증을 위해 FFHQ(Flickr-Face-HQ) 데이터셋을 사용하였다.
- [0115] FFHQ 데이터셋은 GAN의 벤치마크로 만들어진 고품질 얼굴이미지 데이터셋이다. 해당 데이터셋은 1024x1024 해상도의 PNG 이미지 70000개로 구성되어 있으며 도 8과 같이 연령, 포즈, 표정, 민족, 배경 등의 다양한 환경이 포함되어 있다. 또한 안경, 모자 등의 액세서리를 포함한 이미지도 상당량 분포되어 있다.
- [0117] 성능 평가를 위해 생성된 이미지 간의 FID score와 생성된 이미지들의 연령 추정에 대한 MSE를 측정하여 비교하였다.
- [0118] FID score는 이미지의 분포 및 품질이 원본과 얼마나 유사한지를 나타내는 지표로 생성된 영상의 품질비교에 사용되며 원본 데이터셋과 생성된 데이터셋과의 분포 거리를 나타낸다. 원본 데이터와 유사할수록 거리가 가까워 낮은 점수가 출력된다.
- [0119] 실험 결과에서 FID score는 연령 왜곡 감소가 적용되면서 원본 이미지의 특징을 얼마나 유지하는지를 나타내는 지표로 사용하였다.
- [0120] 또한 연령 추정에 대한 MSE는 연령을 직접적으로 추정하여 왜곡이 얼마나 발생했는지 FID보다 더 명확하게 파악이 가능하다.
- [0121] 도 9는 본 발명의 일 실시예에 따른 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소 시스템의 제안한 방법을 통한 이미지 비교를 나타낸 것이다.
- [0122] 또한 연령 특성에 대해 얼마나 복원을 했는지 확인 할 수 있다. 실험 방법은 도 9와 같이 Ground truth와 InterfaceGAN으로 생성된 smile 이미지, Ground truth와 제안하는 기법으로 생성된 smile 이미지를 비교한다.
- [0123] 또한, StyleGAN Encoder로 생성된 결과와 InterfaceGAN으로 생성된 smile 이미지, StyleGAN Encoder로 생성된 결과와 제안하는 기법으로 생성된 smile 이미지를 비교한다. smile 특성을 음의 값으로 입력하여 생성된 non-smile 이미지에 대해서도 위의 방법으로 실험한다.
- [0124] 실험 결과, [표 1]과 같이 Ground Truth에 대한 smile 이미지들의 FID score는 InterfaceGAN이 약 36.475로 측정되었고, 제안하는 기법이 약 36.015로 측정되었다.
- [0125] 표 1은 smile 이미지에 대한 FID score 비교 결과이다.

표 1

smile	FID score ↓	FID score ↓
	Ground Truth	StyleGAN Encoder Results
InterfaceGAN	36.47488427737483	5.187142119681937
The proposed method	36.01520265680995	4.155880873230103

- [0126]
- [0127] 따라서, 제안하는 기법이 기존 InterfaceGAN의 기법과 비교하여 FID score가 약 0.46 향상되었음을 확인하였다. 또한, StyleGAN Encoder로 생성된 결과에 대한 smile 이미지들의 FID score는 InterfaceGAN이 약 5.187로 측정되었고, 제안하는 기법이 약 4.156로 측정되었다. 따라서, 제안하는 기법이 기존 InterfaceGAN과 비교하여 FID score가 약 1.031 향상되었음을 확인하였다.
- [0129] 표 2는 non-smile 이미지에 대한 FID score 비교 결과이다.

표 2

non-smile	FID score ↓	FID score ↓
	Ground Truth	StyleGAN Encoder Results
InterfaceGAN	36.268957432855984	5.968331511453982
The proposed method	34.018085477022396	4.060174437863338

[0130]

[0131]

[표 2]에서는 Ground Truth에 대한 non-smile 이미지들의 FID score는 InterfaceGAN이 약 36.269로 측정되었고, 제안하는 기법이 약 34.018로 측정되었다. 따라서, 제안하는 기법이 기존 InterfaceGAN과 비교하여 FID score가 약 2.25 향상되었음을 확인하였다. 또한, StyleGAN Encoder로 생성된 결과에 대한 non-smile 이미지들의 FID score는 InterfaceGAN이 약 5.968로 측정되었고, 제안하는 기법이 약 4.06로 측정되었다. 따라서, 제안하는 기법이 기존 InterfaceGAN과 비교하여 FID score가 약 1.908 향상되었음을 확인하였다.

[0133]

표 3은 생성된 이미지들의 연령 추정에 대한 MSE 비교 결과이다.

표 3

age	MSE ↓	
	smile	non-smile
InterfaceGAN	3.52942477678571	3.54871629464285
The proposed method	2.02919665178571	1.91833169642857

[0134]

[0135]

[표 3]는 연령 추정 모델을 사용해 StyleGAN Encoder로 생성된 결과와 각 기법으로 생성된 이미지의 연령 MSE 측정 결과이다. InterfaceGAN으로 생성된 smile 이미지는 약 3.529, InterfaceGAN으로 생성된 non-smile 이미지는 약 3.549, 제안하는 기법으로 생성된 smile 이미지는 약 2.029, 제안하는 기법으로 생성된 non-smile 이미지는 약 1.918의 MSE 측정 결과를 보였다. 따라서, 제안하는 기법이 기존 InterfaceGAN과 비교하여 smile 이미지에서 약 1.5, non-smile 이미지에서 약 1.63의 성능 향상을 확인하였다.

[0137]

본 발명에서는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡을 감소시키는 기법을 제안하였다.

[0138]

표정 이미지 생성 과정은 StyleGAN Encoder를 사용하여 얼굴 이미지를 생성하고, SVM을 이용하여 학습된 boundary를 잠재 벡터에 적용하여 표정을 변화시킨다. 그러나 웃는 표정의 boundary를 학습할 때 표정 변화에 따른 연령 왜곡이 발생하였다. 웃는 표정에 대한 SVM 학습에서 생성된 smile boundary는 표정 변화로 인해 생긴 주름이 학습 요소로 포함되어 있으며 연령에 대한 특성도 함께 학습된 것으로 판단하였다.

[0139]

이를 해결하기 위해, 제안된 기법에서는 smile boundary와 age boundary의 상관계수를 계산하고, 이를 이용하여 smile boundary에서 age boundary를 상관계수에 비례하여 조절하는 방식을 도입하였다.

[0140]

제안된 기법의 효과를 확인하기 위해 공개된 표준 얼굴 데이터셋인 FFHQ 데이터셋을 사용하고 FID score를 측정하여 실험한 결과는 다음과 같다.

[0141]

Smile 이미지에서는 기존 기법에 비하여, Ground Truth와 제안된 기법으로 생성된 smile 이미지의 FID score가 약 0.46 향상되었다.

[0142]

또한, Smile 이미지에서 기존 기법에 비하여, StyleGAN Encoder로 생성된 이미지와 제안된 기법으로 생성된 smile 이미지의 FID score가 약 1.031 향상되었다.

[0143]

Non-smile 이미지에서는 기존 기법에 비하여, Ground Truth와 본 발명에서 제안된 기법으로 생성된 non-smile 이미지의 FID score가 약 2.25 향상되었다.

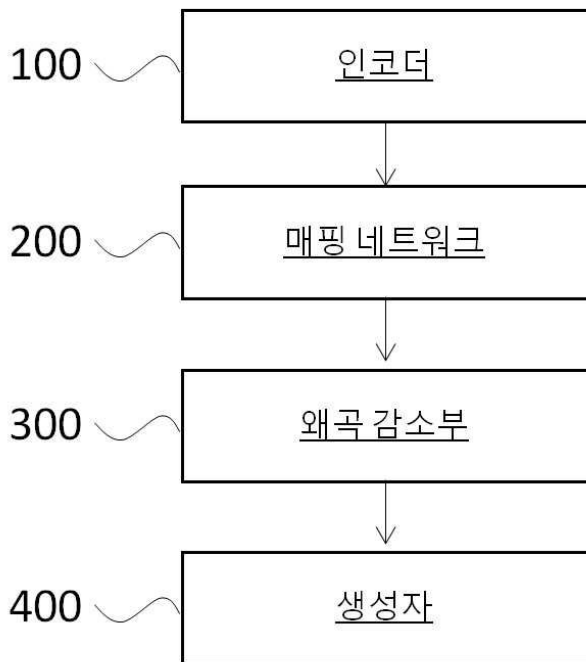
[0144]

또한, Non-smile 이미지에서 기존 기법에 비하여, StyleGAN Encoder로 생성된 이미지와 제안된 기법으로 생성된 non-smile 이미지의 FID score가 약 1.908 향상됨을 확인하였다.

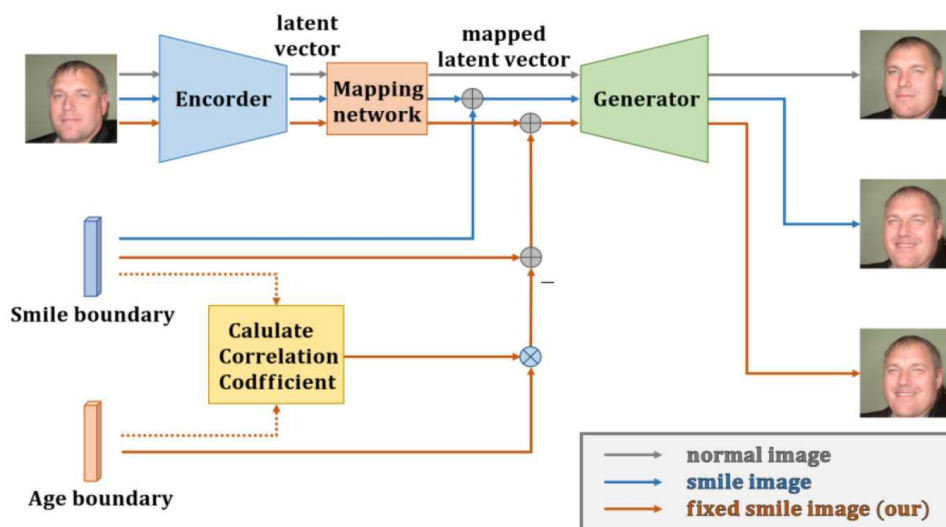
- [0145] 한편, 각 생성된 표정 이미지의 연령을 추정하여 StyleGAN Encoder로 생성된 이미지의 추정된 연령과 MSE를 측정한 결과, 기존 기법 대비 제안하는 기법이 smile 이미지에서 약 1.5, non-smile 이미지에서 약 1.63의 성능 향상되어 제안한 기법에 대한 성능의 효율성이 입증되었다.
- [0146] 본 발명에서 제안하는 기법은 기존의 학습된 데이터를 사용 가능하기 때문에 추가적으로 학습에 소모되는 비용이 없다는 장점이 있다.
- [0147] 그리고, 기존의 프로세스에 추가하여 사용이 가능하기 때문에 시간적, 인력적 비용이 적다는 장점이 있다.
- [0148] 또한 비교적 복잡한 연산을 수행하지 않기 때문에 추가적인 하드웨어 요구가 없다는 장점이 있다.

도면

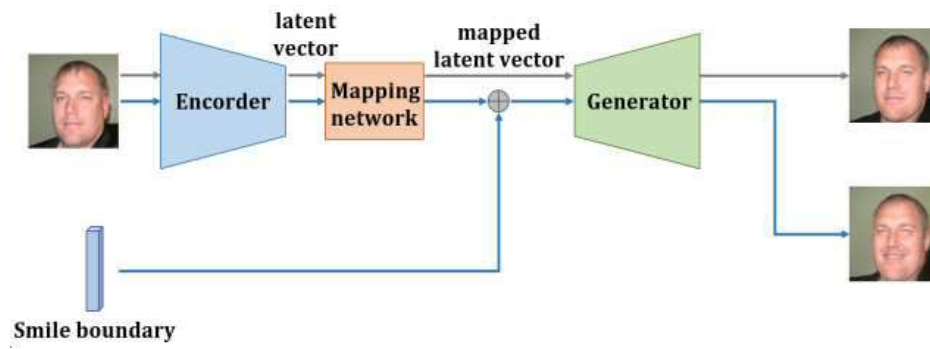
도면1



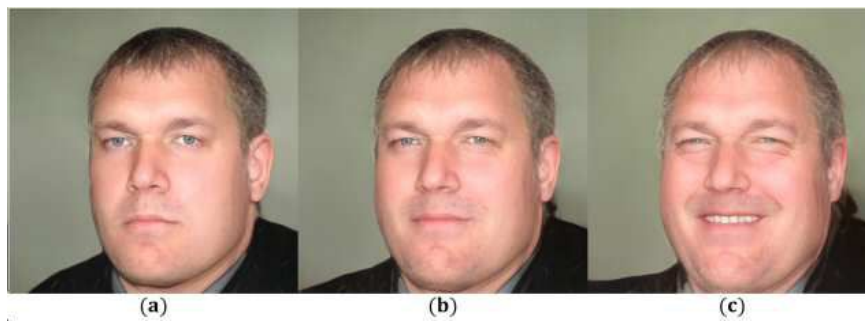
도면2



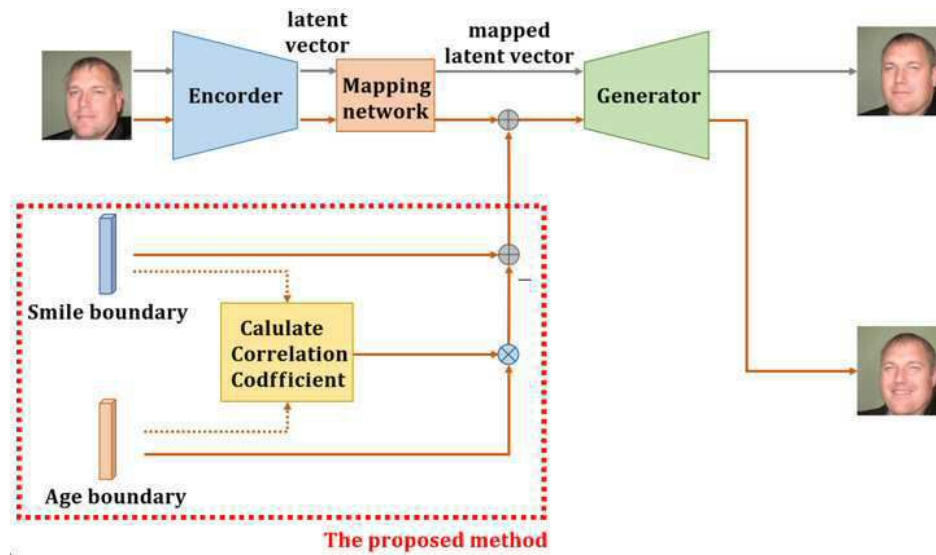
도면3



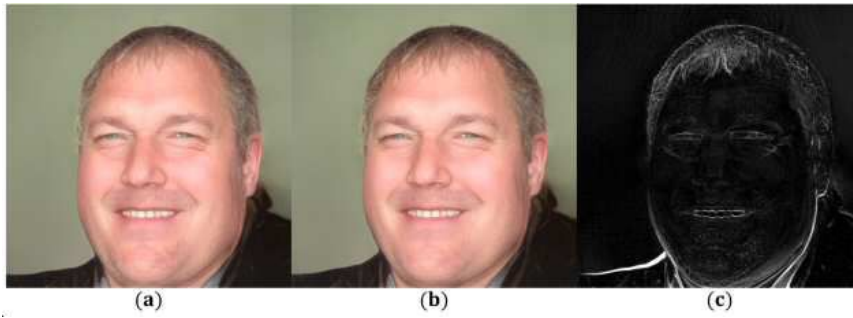
도면4



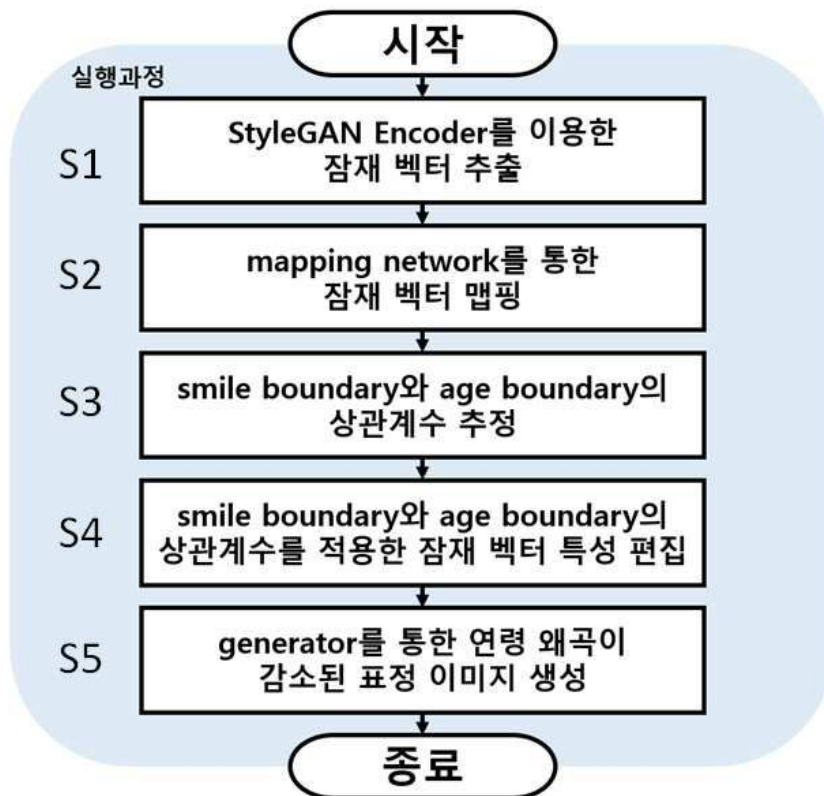
도면5



도면6



도면7



도면8



도면9

