



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0077677
(43) 공개일자 2022년06월09일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2022.01) G06N 20/00 (2019.01)
(52) CPC특허분류
G06V 40/161 (2022.01)
G06N 20/00 (2021.08)
(21) 출원번호 10-2020-0166849
(22) 출원일자 2020년12월02일
심사청구일자 2020년12월02일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
아주대학교산학협력단
경기도 수원시 영통구 월드컵로 206 (원천동)
(72) 발명자
홍성은
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
유중빈
경기도 수원시 영통구 월드컵로 206, 팔달관 604호(원천동, 아주대학교)
(74) 대리인
양성보

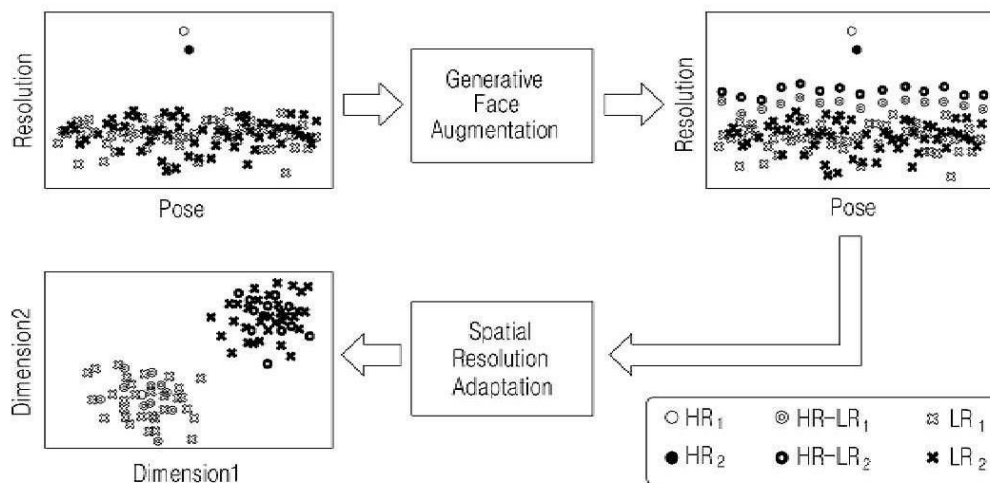
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법 및 장치

(57) 요약

저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법 및 장치가 제시된다. 본 발명에서 제안하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치는 이미지와 특성 레벨에서 모두 정렬된 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습하는 GFA 수행부 및 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블 없이 HR 도메인과 LR 도메인 모두에서 통합 특성 분포를 학습하는 SRA 수행부를 포함한다.

대표도



(52) CPC특허분류

G06V 40/168 (2022.01)

G06V 40/172 (2022.01)

이 발명을 지원한 국가연구개발사업

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

명세서

청구범위

청구항 1

이미지와 특성 레벨에서 모두 정렬된 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습하는 GFA 수행부; 및

SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블 없이 HR 도메인과 LR 도메인 모두에서 통합 특성 분포를 학습하는 SRA 수행부

를 포함하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치.

청구항 2

제1항에 있어서,

GFA 수행부는,

이미지 변환을 수행하는 두 개의 생성기; 및

실제 이미지 변환이 생성기에서 이루어지는지 여부를 분류하는 GFA 감별기

를 포함하고,

두 개의 생성기는,

동일한 대상에 대한 HR 이미지 및 LR 이미지의 페어 데이터를 사용하기 위해 두 개의 생성기로 구성되고, 두 개의 생성기 각각은 어텐션 모듈을 통해 어텐션 유도 도메인 전송을 이용하여 HR 이미지를 LR 이미지에 전송하고 특성 레벨에서 도메인 정렬을 수행하는

저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치.

청구항 3

제2항에 있어서,

GFA 수행부는,

두 개의 생성기 및 어텐션 모듈을 통해 HR 이미지를 입력 받아 도메인 전송 이미지와 어텐션 맵 사이의 전경 영역을 생성하고, 어텐션 맵의 대응으로부터 배경 영역을 획득하여 전경 영역이 생성된 이미지에 추가하는

저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치.

청구항 4

제1항에 있어서,

SRA 수행부는,

특성 추출기, 분류기 및 SRA 감별기를 포함하고,

LR 도메인 레이블 없이, HR과 LR 도메인 모두에서 통합 특성 분포(unified feature distribution)를 학습하여 클래스 레이블과 도메인 레이블을 예측하고,

HR 도메인의 레이블 예측 손실을 최소화하기 위해 특성 추출기 및 클래스 레이블 분류기의 파라미터를 업데이트 하고, SRA 감별기의 파라미터(θ_D)를 검색하는 동시에 도메인 손실을 최대화하는 파라미터(θ_F)를 찾는

저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치.

청구항 5

GFA 수행부를 통해 이미지와 특성 레벨에서 모두 정렬된 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습하는 단계; 및

SRA 수행부를 통해 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블이 없는 HR 도메인과 LR 도메인 모두에서 통합 특성 분포를 학습하는 단계

를 포함하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법.

청구항 6

제5항에 있어서,

GFA 수행부를 통해 이미지와 특성 레벨에서 모두 정렬된 통합 특성 분포를 생성하기 위해 GFA를 수행하여 레이블이 부착된 HR 도메인으로부터 타겟 LR 도메인으로 매핑 기능을 학습하는 단계는,

이미지 변환을 수행하는 두 개의 생성기 및 실제 이미지 변환이 생성기에서 이루어지는지 여부를 분류하는 GFA 감별기를 포함하는 GFA 수행부의 두 개의 생성기를 통해 동일한 대상에 대한 HR 이미지 및 LR 이미지의 페어 데이터를 사용하고, 두 개의 생성기 각각은 어텐션 모듈을 통해 어텐션 유도 도메인 전송을 이용하여 HR 이미지를 LR 이미지에 전송하고 특성 레벨에서 도메인 정렬을 수행하는

저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법.

청구항 7

제6항에 있어서,

두 개의 생성기 및 어텐션 모듈을 통해 HR 이미지를 입력 받아 도메인 전송 이미지와 어텐션 맵 사이의 전경 영역을 생성하고, 어텐션 맵의 대응으로부터 배경 영역을 획득하여 전경 영역이 생성된 이미지에 추가하는

저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법.

청구항 8

제5항에 있어서,

SRA 수행부를 통해 SRA를 수행하여 LR 도메인 레이블이 없는 HR 도메인과 LR 도메인 모두에서 통합 특성 분포를 학습하는 단계는,

특성 추출기, 분류기 및 SRA 감별기를 포함하는 SRA 수행부를 통해 LR 도메인 레이블 없이, HR과 LR 도메인 모두에서 통합 특성 분포(unified feature distribution)를 학습하여 클래스 레이블과 도메인 레이블을 예측하고, HR 도메인의 레이블 예측 손실을 최소화하기 위해 특성 추출기 및 클래스 레이블 분류기의 파라미터를 업데이트 하고, SRA 감별기의 파라미터(θ_D)를 검색하는 동시에 도메인 손실을 최대화하는 파라미터(θ_F)를 찾는

저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법.

발명의 설명

기술 분야

[0001] 본 발명은 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법에 관한 것이다.

배경 기술

[0002] 최근의 얼굴 인식 방법은 유망한 결과를 보여 주었지만 [1]-[3], 갤러리와 프로브 세트 사이의 이질적인 도메인은 여전히 어려운 문제로 남아 있다[4], [5]. 이질성의 일반적인 예는 고품질 갤러리(예를 들어, 식별 사진)와 보안 감시 카메라의 저품질 프로브 이미지 사이의 중요한 차이점이다[6]. 신분증이나 여권에 사용된 사진은 갤러리 이미지로 등록하기 위해 안정된 환경에서 캡처된다. 반면에 프로브 이미지는 [7], [8]에서 언급한 바와 같

이 노이즈, 흐릿함, 임의 포즈 등 불안정한 실제 환경에서 주로 캡처된다. 이로 인해 소스(즉, 갤러리)와 대상(즉, 프로브) 이미지 사이에 심한 도메인 이동이 일어나 인식 성능이 저하된다.

[0003] 지난 수십 년 동안 갤러리와 프로브 세트 사이의 이질적인 조건에 대처하기 위해 상당한 노력[9]-[13]을 기울였다. 이러한 종래기술들 중, 저품질 조건 하에서 프로브 이미지를 촬영하는 저해상도 얼굴 인식(Low-Resolution Face Recognition; LRFR)[14]에 대한 관심이 증가하고 있다[15]-[17]. 기존의 방법은 두 가지 범주로 나눌 수 있다. 환영 접근법은 얼굴 인식을 수행하기 전에 저해상도(Low-Resolution; LR) 프로브 이미지로부터 고해상도(High-Resolution; HR) 이미지를 얻는다[18]-[20]. 반면에 내장형 범어텐션 모델은 HR 및 LR 조건에 대한 공통 특성 공간을 생성한다 [21]-[23]. 결정적으로 두 접근법 모두 HR 및 LR 이미지에 모두 ID 레이블을 사용한다. 즉, 기존의 대부분의 LRFR 방법은 ID 레이블의 감독을 통해 HR 갤러리와 LR 프로브 세트 사이의 관계를 학습한다.

[0004] 그러나 LR 프로브 세트의 레이블링을 요구하지 않는 비감독 접근방식은 감독된 기반 접근방식에 비해 유의적인 이점을 가지고 있다고 주장한다. 우선 LR 이미지를 미리 획득하는 것은 어렵지 않지만, 명시적으로 레이블을 붙인 이미지는 거의 없고 레이블을 붙이는 것은 상당히 노동집약적이고 비실용적이다. 둘째, 비감독 설정은 시간이 지남에 따라 인식 모델의 성능을 향상시킬 수 있는 장점이 있다. 비감독 얼굴 인식 시스템은 LR 프로브 이미지의 레이블 정보가 필요 없어 LR 프로브 이미지에 연속적으로 부여해도 쉽게 진화할 수 있다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 한국 공개특허공보 제10-2019-0008794호(2020.08.10.)

발명의 내용

해결하려는 과제

[0006] 본 발명이 이루고자 하는 기술적 과제는 저해상도 얼굴 인식을 위한 새로운 비감독 얼굴 도메인 전송 방법 및 장치를 제공하는데 있다. 어텐션 메커니즘(attention mechanism)을 활용함으로써 제안하는 생성 얼굴 확대는 이미지 레벨에서 도메인 이동을 감소시키는 반면, 제안하는 공간 해상도 적응은 도메인-불변성 및 차별적 특성 분포를 생성하고자 한다.

과제의 해결 수단

[0007] 일 측면에 있어서, 본 발명에서 제안하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치는 이미지와 특성 레벨에서 모두 정렬된 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습하는 GFA 수행부 및 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블 없이 HR 도메인과 LR 도메인 모두에서 통합 특성 분포를 학습하는 SRA 수행부를 포함한다.

[0008] GFA 수행부는 이미지 변환을 수행하는 두 개의 생성기 및 실제 이미지 변환이 생성기에서 이루어지는지 여부를 분류하는 GFA 감별기를 포함한다.

[0009] 두 개의 생성기는 동일한 대상에 대한 HR 이미지 및 LR 이미지의 페어 데이터를 사용하기 위해 두 개의 생성기로 구성되고, 두 개의 생성기 각각은 어텐션 모듈을 통해 어텐션 유도 도메인 전송을 이용하여 HR 이미지를 LR 이미지에 전송하고 특성 레벨에서 도메인 정렬을 수행한다.

[0010] GFA 수행부는 두 개의 생성기 및 어텐션 모듈을 통해 HR 이미지를 입력 받아 도메인 전송 이미지와 어텐션 맵 사이의 전경 영역을 생성하고, 어텐션 맵의 대응으로부터 배경 영역을 획득하여 전경 영역이 생성된 이미지에 추가한다.

[0011] SRA 수행부는 특성 추출기, 분류기 및 SRA 감별기를 포함하고, LR 도메인 레이블 없이, HR과 LR 도메인 모두에서 통합 특성 분포(unified feature distribution)를 학습하여 클래스 레이블과 도메인 레이블을 예측하고, HR 도메인의 레이블 예측 손실을 최소화하기 위해 특성 추출기 및 클래스 레이블 분류기의 파라미터를 업데이트하

고, SRA 감별기의 파라미터(θ_D)를 검색하는 동시에 도메인 손실을 최대화하는 파라미터(θ_F)를 찾는다.

[0012] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법은 GFA 수행부를 통해 이미지와 특성 레벨에서 모두 정렬된 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습하는 단계 및 SRA 수행부를 통해 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블이 없는 HR 도메인과 LR 도메인 모두에서 통합 특성 분포를 학습하는 단계를 포함한다.

발명의 효과

[0013] 본 발명의 실시예들에 따르면 저해상도 얼굴 인식을 위한 새로운 비감독 얼굴 도메인 전송 방법 및 장치를 통해 어텐션 메커니즘(attention mechanism)을 활용함으로써 제안하는 생성 얼굴 확대는 이미지 레벨에서 도메인 이동을 감소시키는 반면, 제안하는 공간 해상도 적응은 도메인-불변성 및 차별적 특성 분포를 생성할 수 있다.

도면의 간단한 설명

[0014] 도 1은 본 발명의 일 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법을 설명하기 위한 개략도이다.

도 2는 본 발명의 일 실시예에 따른 본 발명의 일 실시예에 따른 HR 갤러리 이미지, 도메인 전송 이미지, LR 프로브 이미지의 샘플 이미지를 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치의 구성을 나타내는 도면이다.

도 4는 본 발명의 일 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0015] 고해상도 갤러리와 저해상도 프로브 세트의 해상도가 다르기 때문에 저해상도 얼굴 인식은 도메인 이동을 겪는다. 기존 방법에서는 동일한 대상에 대해 고해상도 및 저해상도 간의 쌍방향 상관관계를 사용하며, 여기에는 갤러리와 프로브 세트에 대한 레이블 정보가 모두 필요하다. 그러나 저해상도 프로브 이미지에 명시적으로 레이블을 붙인 것은 거의 없으며 레이블을 붙이는 것은 노동 집약적이다. 본 발명에서는 강력한 저해상도 얼굴 인식을 위한 새로운 비감독 얼굴 도메인 전송을 제안한다. 어텐션 메커니즘(attention mechanism)을 활용함으로써 제안된 생성 얼굴 확대는 이미지 레벨에서 도메인 이동을 감소시키는 반면 공간 해상도 적응은 도메인-불변성 및 차별적 특성 분포를 생성한다. 공용 데이터셋에서 이미지 레벨에서의 생성 얼굴 확대와 특성 레벨에서의 공간 해상도 적응 사이의 보완성을 입증한다. 본 발명에서는 저해상도 프로브 세트의 레이블 정보를 전혀 사용하지 않지만, 최첨단 감독 방법을 능가한다.

[0017] 도 1은 본 발명의 일 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법을 설명하기 위한 개략도이다.

[0018] 본 발명에서는 이미지 레벨뿐만 아니라 특성 레벨의 도메인 이동을 줄이기 위한 비감독 LRFR 방식을 제안한다. 이전의 접근법 [18], [19], [21]-[23]과 달리 제안하는 모델은 LR 프로브 세트에 레이블 정보가 없는 이미지만 사용한다. 도 1에서 볼 수 있듯이, 제안하는 방법은 이미지와 특성 레벨에서 모두 정렬된 차별적 분포를 생성하기 위해 생성 얼굴 확대(Generative Face Augmentation; GFA)와 공간 해상도 적응(Spatial Resolution Adaptation; SRA) 네트워크로 구성된다. 제안된 GFA는 레이블이 부착된 HR 이미지를 LR 유사 이미지로 합성하도록 설계되었으며, SRA는 생성된 LR 유사 이미지가 특성 레벨의 실제 LR 이미지와 구별되지 않도록 보장한다.

[0019] 제안된 방법은 이질적인 얼굴 인식을 위한 도메인 적응 네트워크를 고려함에 있어 종래기술[24]과 밀접하게 관련되어 있다. 사전 작업 [24]에서는 포즈 변화에만 초점을 맞춘 3D 얼굴 모델로 합성 이미지를 생성한다. 대조적으로, 본 발명은 도메인 적응 네트워크(즉, SRA)와 함께 학습 기반 생성 네트워크(즉, GFA)의 공동 최적화를 도입하여 도메인 전송 이미지 및 도메인 불변성 특성 분포를 생성한다. 그러나 단순히 HR 도메인에서 LR 도메인으로 이미지를 전송하면 모드 붕괴[25]와 예기치 않은 아티팩트가 자주 발생한다[26]. 따라서, 본 발명의 실시

예에 따른 GFA는 소스 이미지(source images)와 대상 이미지(target images) 사이의 차이를 보상하기 위해 중요한 부분에 초점을 맞추는 어텐션 유도 도메인 전송 모듈(attention-guided domain transfer module)을 포함한다.

- [0021] 도 2는 본 발명의 일 실시예에 따른 HR 갤러리 이미지, 도메인 전송 이미지, LR 프로브 이미지의 샘플 이미지를 나타내는 도면이다.
- [0022] 레이블이 지정된 HR 이미지와 레이블이 지정되지 않은 LR 이미지가 주어지면 GFA는 어텐션 메커니즘을 활용하여 어떠한 레이블 페어링없이 두 도메인을 매핑할 수 있다. 도 2(a)는 HR 갤러리 이미지를 나타내는 것으로, 다양한 포즈의 합성 이미지, 밝은 값이 더 높은 초점을 나타내는 학습된 어텐션 지도, 다양한 포즈의 도메인 전송 HR-LR 이미지(왼쪽에서 오른쪽으로)를 나타내고, 도 2(b)는 LR 프로브 이미지를 나타낸다.
- [0023] 제안된 방법을 공공 데이터셋에서 도전적이고 현실적인 LRFR 프로토콜로 [24], [27], [28]에서 검증한다. 갤러리 세트에는 식별 사진 등 안정적인 환경에서 캡처한 HR 이미지가, 프로브 세트에는 감시 카메라의 레이블이 없는 LR 이미지가 담겨 있다. 제안하는 GFA와 SRA가 보완적으로 성능을 발휘하고 결국 급격한 성능 향상을 보인다는 것을 실험적으로 보여준다. 제안하는 비감독 모델의 정확도는 최첨단 감독 방법[10], [24], [29]-[31]보다 우수하다. 또한 LRFR에 대한 도메인 적응 기반 접근법의 정량적 기준이 될 수 있는 광범위한 연구를 제공한다.
- [0025] 기존의 저해상도 얼굴 인식(Low-Resolution Face Recognition; LRFR) 방식은 해상도 정렬이 적용되는 레벨(예를 들어, 이미지 레벨 또는 특성 레벨)에 따라 두 가지 범주로 분류할 수 있다. 이미지 레벨 방법의 초기 연구에는 사람 특유의 저해상도 얼굴 이미지를 합성하기 위한 단순 값 분해[18] 또는 최소 표현[19]이 포함된다. 입력 LR 이미지에서 HR 이미지에 대한 환영 기반 재구성도 제안되었다. LR 이미지를 HR 이미지에 전송하기 위해 초해상도 기법이나 얼굴 속성 내장 방법이 널리 사용되어 왔다 [11], [20], [32]. 이러한 방법은 프로브 LR 이미지를 HR 이미지에 전송하여 지나치게 부드러워지고 세부 정보가 손실되는 원인이 된다. 반대로, 본 발명에서는 어텐션 유도 도메인 전송을 이용하여 HR 이미지를 LR 이미지에 전송하고 또한 특성 레벨에서 도메인 정렬을 수행한다.
- [0026] 한편, 특성 레벨 범어텐션 목표는 동일 대상에 대해 LR과 HR 이미지의 근접성이 유지되는 통일된 특성 공간을 찾는 것이다. 커널 맵핑 전략은 HR 이미지와 LR 이미지 사이의 투영법을 학습하는 데 널리 사용된다[14], [21]. 사전 기반 접근법(Dictionary-based approaches)도 다른 해상도에서 캡처한 얼굴 이미지와 일치하도록 제안되었다 [22], [23]. 결정적으로, 대부분의 기존 방법론에서는 각 대상에 대해 사용 가능한 LR 버전과 HR 버전이 있다고 가정하며, 이는 실제 적용에 비실용적이다. 이와는 대조적으로, 본 발명에서는 LR 이미지에 어떤 레이블도 요구하지 않는 비감독 방법을 제안한다.
- [0027] 얼굴 관련 작업에 도메인 적응(Domain Adaptation; DA)을 적용하기 위한 여러 시도가 제안되었다. Xie[33]는 프로브 세트에 흐릿한 이미지가 포함되어 있는 동안 갤러리 세트가 선명한 이미지로 구성되는 얼굴 인식 문제를 해결하기 위해 DA와 여러 수작업으로 제작된 디스크립터(descriptor)를 사용한다. Banerjee[34]는 필터 뱅크(예를 들어, Eigenfaces, Fisherfaces, Gaborfaces)와 함께 DA 기반 방법을 제안한다. 이미지에서 수작업으로 제작된 특성을 추출한 후 DA를 적용하는 위의 접근방식과 달리, 통합 딥 아키텍처에서 특성 학습과 분류를 공동으로 실시한다. 또한, 모델에는 1인당 레이블이 부착된 이미지가 1개만 주어지는 얼굴 인식 문제를 해결하는데, 이는 도전적이고 현실적인 프로토콜이다.
- [0029] 도 3은 본 발명의 일 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치의 구성을 나타내는 도면이다.
- [0030] 제안하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치는 GFA 수행부(310) 및 SRA 수행부(320)를 포함한다.
- [0031] GFA 수행부(310)는 이미지와 특성 레벨에서 모두 정렬된 차별적 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습한다.

- [0032] GFA 수행부(310)는 이미지 변환을 수행하는 두 개의 생성기($Gen_{HR \rightarrow LR}(311)$, $Gen_{LR \rightarrow HR}(312)$) 및 실제 이미지 변환이 생성기에서 이루어지는지 여부를 분류하는 GFA 감별기(313)를 포함한다.
- [0033] 두 개의 생성기($Gen_{HR \rightarrow LR}(311)$, $Gen_{LR \rightarrow HR}(312)$)는 동일한 대상에 대한 HR 이미지 및 LR 이미지의 페어 데이터를 사용하기 위해 두 개의 생성기로 구성되고, 두 개의 생성기 각각은 어텐션 모듈을 통해 어텐션 유도 도메인 전송을 이용하여 HR 이미지를 LR 이미지에 전송하고 특성 레벨에서 도메인 정렬을 수행한다.
- [0034] GFA 수행부(310)는 두 개의 생성기 및 어텐션 모듈을 통해 HR 이미지를 입력 받아 도메인 전송 이미지와 어텐션 맵 사이의 전경 영역을 생성하고, 어텐션 맵의 대응으로부터 배경 영역을 획득하여 전경 영역이 생성된 이미지에 추가한다.
- [0035] SRA 수행부(320)는 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블 없이 HR 도메인과 LR 도메인 모두에서 차별적인(discriminative) 통합 특성 분포(unified feature distribution)를 학습한다.
- [0036] SRA 수행부(320)는 특성 추출기(321), 분류기(322) 및 SRA 감별기(323)를 포함한다. SRA 수행부(320)는 LR 도메인 레이블 없이, HR과 LR 도메인 모두에서 통합 특성 분포(unified feature distribution)를 학습하여 클래스 레이블과 도메인 레이블을 예측한다. HR 도메인의 레이블 예측 손실을 최소화하기 위해 특성 추출기 및 클래스 레이블 분류기의 파라미터를 업데이트하고, SRA 감별기의 파라미터(θ_D)를 검색하는 동시에 도메인 손실을 최대화하는 파라미터(θ_F)를 찾는다. 아래에서 제안하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 장치에 대하여 더욱 상세히 설명한다.
- [0037] GFA 수행부(310)의 목표는 소스 HR도메인(source HR domain)으로부터 대상 LR도메인(target LR domain), 즉 $Gen_{HR \rightarrow LR}$ 에 이르는 매핑 기능을 학습하는 것이다. 그 차이를 보완하기 위해 'pix2pix' 프레임워크[35]를 이용한 이미지 레벨 도메인 전송(image-level domain transfers)을 이용하는데, 이 프레임워크는 이미지 변환을 시도하는 '생성기(generator)'($Gen_{HR \rightarrow LR}(311)$, $Gen_{LR \rightarrow HR}(312)$)라고 불리는 자동 인코더 아키텍처와 실제 이미지 변환이 '생성기'에서 이루어지는지 여부를 분류하려는 '감별기(discriminator)'(313) 두 개의 서브 네트워크로 구성된다. 적대적 생성망(Generative Adversarial Networks; GAN)[36]에서와 같이 그들 사이의 이러한 생성적 대립 관계 때문에, 학습이 진행될수록 생성기와 감별기 모두 개선될 수 있다.
- [0038] 동일한 대상에 대한 HR 및 LR 이미지의 페어(pair) 데이터를 사용할 수 없기 때문에, 본 발명에서는 두 개의 생성기($Gen_{HR \rightarrow LR}(311)$, $Gen_{LR \rightarrow HR}(312)$)를 구성하고 주기 일관성 손실을 채택한다[37]. 그러나 이미지 레벨 도메인 전송을 소스 및 대상 도메인에 적용하는 것만으로도 예기치 않은 아티팩트와 성능 저하를 초래할 수 있다. 컴퓨터 비전 분야[38]-[40]에서 최근 어텐션 기반 접근법의 성공에 따라, 도메인 전송 동안 어텐션 메커니즘을 활용한다. 이를 위해 본 발명에서는 두 개의 어텐션 모듈($Att_{HR \rightarrow LR}(311a)$, $Att_{LR \rightarrow HR}(312a)$)을 각각의 생성기에 추가한다. 이미지에서 중요한 부분, 즉 전경(foreground)에만 초점을 맞춰 합성된 이미지의 아티팩트를 줄인다. HR 이미지를 제공하면 먼저 생성기와 어텐션 모듈로 HR 이미지를 공급한다. 본 발명에서는 도메인 전송 이미지와 어텐션 맵 사이에 Hadamard product[41]을 통해 전경 영역을 생성할 수 있다. 그 후 어텐션 맵의 대응으로부터 배경 영역을 획득하여 전경 영역이 생성된 이미지에 추가한다. 그 결과 다음과 같이 도메인 전송 이미지를 얻을 수 있다.
- [0039]
$$hr - lr = Gen_{HR \rightarrow LR}(hr) \odot m_{Att} + hr \odot (1 - m_{Att})$$
- [0040] 여기서 hr 는 소스 HR 도메인에서 가져온 이미지이고 m_{Att} 는 픽셀당 $[0, 1]$ 값을 포함하는 어텐션 모듈 $Att_{HR \rightarrow LR}$ 에서 파생된 어텐션 맵을 가리킨다.
- [0041] 이미지 레벨 도메인 전송 이미지를 생성하지만, 소스 HR 도메인과 특성 레벨의 대상 LR 도메인 사이에는 여전히 차이가 존재한다. 따라서 도 3과 같이 클래스 레이블과 도메인 레이블(즉, HR/LR)을 예측하는 SRA 네트워크를 도입한다. SRA 수행부(320)는 LR 도메인 레이블 없이, HR과 LR 도메인 모두에서 차별적인 통합 특성 분포(unified feature distribution)를 학습하는 것을 목표로 한다. 통합 특성 분포를 위해 F 및 C , θ_F 및 θ_C 의 파라미터를 업데이트하여 HR 도메인의 레이블 예측 손실을 최소화한다. 여기서 F 는 특성 추출기(321)를 나타내며, C 는 클래스 레이블 분류기(322)를 가리킨다. 판별 학습과 동시에 HR 이미지와 LR 이미지의 분포를 일치시켜야 한다. 도메인 불변성 특성을 얻기 위해, SRA는 다음과 같이 도메인 예측 손실을 최소화하기 위한 도메인 감별기

(323) $D(\theta_D)$ 의 파라미터를 검색하는 동시에 도메인 손실을 최대화하는 θ_F 를 찾는다.

$$L = \sum_{i \in HR} L_C^i + \sum_{i \in HR \cup LR} L_D^i \quad \text{when update } \theta_D,$$

$$L = \sum_{i \in HR} L_C^i - \lambda \sum_{i \in HR \cup LR} L_D^i \quad \text{when update } \theta_F \text{ and } \theta_C,$$

여기서 L_C^i 와 L_D^i 는 각각 클래스 및 도메인 예측에서 i 번째 샘플의 손실을 가리킨다. 파라미터 λ 는 λ 의 네거티브 사인이 손실 측면에서 F와 D의 대립적 관계로 이어지는 가장 중요한 요인이다. 그 결과 F의 파라미터는 차별성이 있고 네트워크 손실을 최소화하는 과정을 통해 도메인 불변을 만족시키는 타협점에 수렴한다.

도 4는 본 발명의 일 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법을 설명하기 위한 흐름도이다.

제안하는 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법은 GFA 수행부를 통해 이미지와 특성 레벨에서 모두 정렬된 차별적 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습하는 단계(410) 및 SRA 수행부를 통해 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블이 없는 HR 도메인과 LR 도메인 모두에서 차별적 통합 특성 분포를 학습하는 단계(420)를 포함한다.

단계(410)에서, GFA 수행부를 통해 이미지와 특성 레벨에서 모두 정렬된 차별적 통합 특성 분포를 생성하기 위해 GFA(Generative Face Augmentation)를 수행하여 레이블이 부착된 HR(High-Resolution) 도메인으로부터 타겟 LR(Low-Resolution) 도메인으로 매핑 기능을 학습한다.

GFA 수행부는 이미지 변환을 수행하는 두 개의 생성기 및 실제 이미지 변환이 생성기에서 이루어지는지 여부를 분류하는 GFA 감별기를 포함한다.

두 개의 생성기는 동일한 대상에 대한 HR 이미지 및 LR 이미지의 페어 데이터를 사용하기 위해 두 개의 생성기로 구성되고, 두 개의 생성기 각각은 어텐션 모듈을 통해 어텐션 유도 도메인 전송을 이용하여 HR 이미지를 LR 이미지에 전송하고 특성 레벨에서 도메인 정렬을 수행한다.

GFA 수행부는 두 개의 생성기 및 어텐션 모듈을 통해 HR 이미지를 입력 받아 도메인 전송 이미지와 어텐션 맵 사이의 전경 영역을 생성하고, 어텐션 맵의 대응으로부터 배경 영역을 획득하여 전경 영역이 생성된 이미지에 추가한다.

단계(420)에서, SRA 수행부를 통해 SRA(Spatial Resolution Adaptation)를 수행하여 LR 도메인 레이블이 없는 HR 도메인과 LR 도메인 모두에서 차별적 통합 특성 분포를 학습한다.

SRA 수행부는 특성 추출기, 분류기 및 SRA 감별기를 포함한다. SRA 수행부는 LR 도메인 레이블 없이, HR과 LR 도메인 모두에서 통합 특성 분포(unified feature distribution)를 학습하여 클래스 레이블과 도메인 레이블을 예측한다. HR 도메인의 레이블 예측 손실을 최소화하기 위해 특성 추출기 및 클래스 레이블 분류기의 파라미터를 업데이트하고, SRA 감별기의 파라미터(θ_D)를 검색하는 동시에 도메인 손실을 최대화하는 파라미터(θ_F)를 찾는다. 아래에서 본 발명의 실시예에 따른 저해상도 얼굴 인식을 위한 생성적 도메인 적응 방법의 실험 결과에 대해 더욱 상세히 설명한다.

GFA를 구축하기 위해 '생성기'는 컨볼루션, 잔여블록, 디컨볼루션으로 구성되는 'pix2pix' 프레임워크를 구축하며, '감별기'는 여러 컨볼루션이 있는 2클래스의 분류 네트워크이다. 소스 도메인 이미지의 부족으로 인한 과적합 및 DA 장애 방지를 위해 몇 가지 확대 방식(예를 들어, 플립, 회전, 자르기)[42], [43]을 적용한다. 또한 대상 도메인에서 자세 변동이 심할 경우 선택적으로 3D 이미지 합성을 수행한다[24]. SRA의 경우 사전 훈련된 VGG-Face[1]를 특성 추출기로 채택하고 분류기와 감별기를 위한 얇은 네트워크를 부착한 후 최종적으로 미세 조정한다. 상대적으로 작은 데이터셋에 과도하게 장착되는 것을 방지하기 위해 미리 훈련된 VGG-Face의 Conv1 계층에서 FC6 계층까지의 가중치를 동결하고 FC7 계층을 미세 조정한다.

제안된 방법의 효과를 공개 데이터셋, EK-LFH [24], SCface [27], YouTubeFaces[YTF]에서 입증한다 [28]. 각

데이터셋의 세부사항은 표 1에 제시되어 있다.

<표 1>

Dataset	EK-LFH		SCface		YTF	
Domian	HR	LR	HR	LR	HR	LR
Subjects	30	20/10	130	50/80	1595	798/797
Samples	30	16K	130	2K	3.5K	306K/311K

실제 상황을 고려하여, LRFR[10], [29], [44]의 대상 분리 프로토콜을 따른다. 이 프로토콜은 훈련 단계에서 나타나는 대상의 시험 이미지를 사용하지 못하게 한다.

다음으로 EK-LFH에 대한 평가에 대하여 설명한다.

EKL FH [24]의 대상 분리 프로토콜 [10], [29], [44]에 따라 랜덤으로 선택한 대상 20개를 훈련용으로 사용하고 나머지 10개 대상은 시험용으로 사용한다. 표 1에서 볼 수 있듯이 EK-LFH의 대상 도메인에서 포즈 변화가 심하므로, 이미지 레벨 도메인 전송 중에 3D 이미지 합성[24]을 수행한다. 기존의 비감독 DA [24], [45]의 실험 설정에 의해 동기 부여되어 표 2와 같이 3개의 프로토콜에 대한 실험을 수행한다.

<표 2>

Protocol	Training Set	Accuracy
Source Only	{HR _L }	20.97%
	{HR _L , HR _{pose} }	20.29%
	{HR _L , HR _{pose} , HR-LR}	26.81%
Proposed Adaptation	{HR _L , LR _U }	28.50%
	{HR _L , HR _{pose} , LR _U }	50.12%
	{HR _L , HR _{pose} , HR-LR, LR _U }	52.76%
Labeled Target	{HR _L , LR _L }	63.39%

첫째, 'Source Only' 프로토콜은 소스 HR 도메인에서 레이블이 부착된 샘플만 사용하며, 이는 이론적으로 성능에 대한 하한선을 20.97%로 드러냈다. 또한 기존의 'Source Only' 변형으로 교육 세트에 합성 이미지(예를 들어, HR_{pose} 또는 HR-LR)를 추가한 결과를 보고한다. 둘째, 상위 성능 경계로서 'Labeled Target' 프로토콜의 모델은 레이블이 부착된 소스 HR 이미지와 레이블이 부착된 대상 LR 이미지에 의해 훈련된다.

마지막으로, 본 발명의 주요 접근방식인 '제안된 적응(Proposed Adaptation)'은 레이블이 부착되지 않은 대상 이미지와 함께 레이블이 부착된 소스 이미지를 사용한다. 대상 도메인의 레이블을 사용하지 않더라도 합성 이미지를 사용한 제안 방법은 'Source Only'에 비해 정확도가 향상되는 것을 관찰한다. 5행과 6행은 비감독 DA를 적용할 때 다양한 포즈로 합성된 이미지의 중요성을 검증한다. 또한, 제안된 이미지 레벨 도메인 전송이 성능을 더욱 향상시킬 수 있다는 것을 확인할 수 있다. 전반적인 결과에서 GFA(이미지 레벨)와 SRA(특성 레벨)가 LRFR 과제 해결에 보완적으로 작용하는 것으로 나타났다.

다음으로 SCface 평가에 대해 설명한다.

SCface[9], [10], [29]의 기존 프로토콜에 따라, 훈련 세트와 시험 세트를 50개와 80개 대상으로 분리했다. 표 3에서는 대상 도메인의 레이블 정보가 사용되지 않더라도 'Proposed Adaptation'을 사용하여 상당한 성능 향상을 확인한다.

<표 3>

Protocol	Training Set	Accuracy
Source Only	{HR _L }	23.00%
	{HR _L , HR-LR}	26.08%
Proposed Adaptation	{HR _L , LR _U }	50.50%
	{HR _L , HR-LR, LR _U }	52.25%

SCface에 대한 실험 결과가 EK-LFH 데이터셋의 결과와 일치한다는 것은 명백하다. 마지막으로 표 4에서 볼 수

있는 바와 같이 제안하는 방법을 대상 도메인에서 레이블을 사용하는 최첨단 감독 방법과 비교한다.

<표 4>

Method	Accuracy	Method	Accuracy
SSR [46]	12.78%	LRFRW[9]	24.30%
CSCDN[44]	13.18%	DAlign [29]	41.04%
CCA [47]	15.11%	SKD [10]	48.33%
C-RSDA[30]	18.72%	SSPP-DAN [24]	50.50%
DCA [31]	17.44%	Ours*	52.25%

대상 LR 도메인에서 레이블 정보를 사용하지 않았음에도 불구하고, 제안된 방법은 감독된 LRFR 방법을 능가한다. 마지막으로, 본 발명은 이미지 수준 도메인 전송을 포함한 제안된 방법이 비감독 DA에 근거한 이전 작업[24]보다 우월하다는 것을 알 수 있다.

다음으로 YouTubeFaces에 대한 평가에 대해 설명한다.

제안된 방법의 일반화를 입증하기 위해 대규모 YTF 데이터셋[28]에 대한 평가를 수행한다. YTF 데이터 집합은 총 620k 이미지를 포함하여 1,595개 대상의 3,425개의 비디오로 구성된다. 본 발명에서는 HR 도메인의 경우 나머지 프레임은 16×16 으로 하향 샘플링하면서 HR 도메인의 경우 각 비디오의 첫 번째 프레임의 크기를 조정한다. 다운샘플링된 LR 도메인 샘플 중 절반 세트로 나누어 훈련과 시험 데이터에 각각 사용한다. EK-LFH 및 SCface 데이터셋의 동일한 실험 설정에서 제안된 적응 방법을 'Source Only' 기준과 비교한다. '제안된 적응'이 19.2%를 달성해 8.0%의 정확도로 'Source Only'보다 훨씬 높은 것을 확인한다. 이 결과는 대규모 데이터셋에서 제안된 방법의 일반화 성능을 지원한다.

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, 콘트롤러, 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 콘트롤러를 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(embody)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

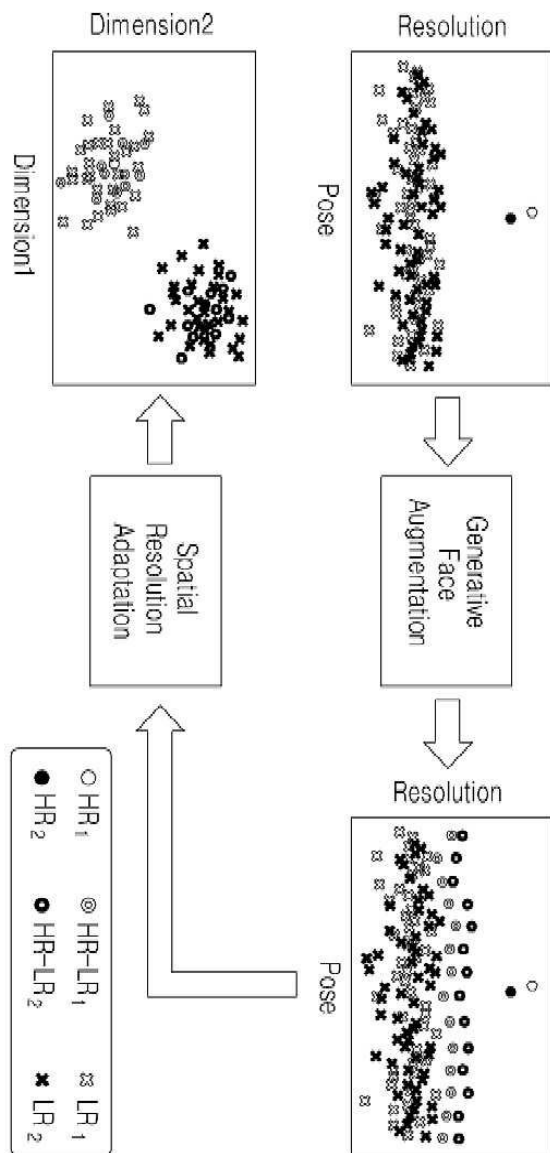
기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0079] 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.

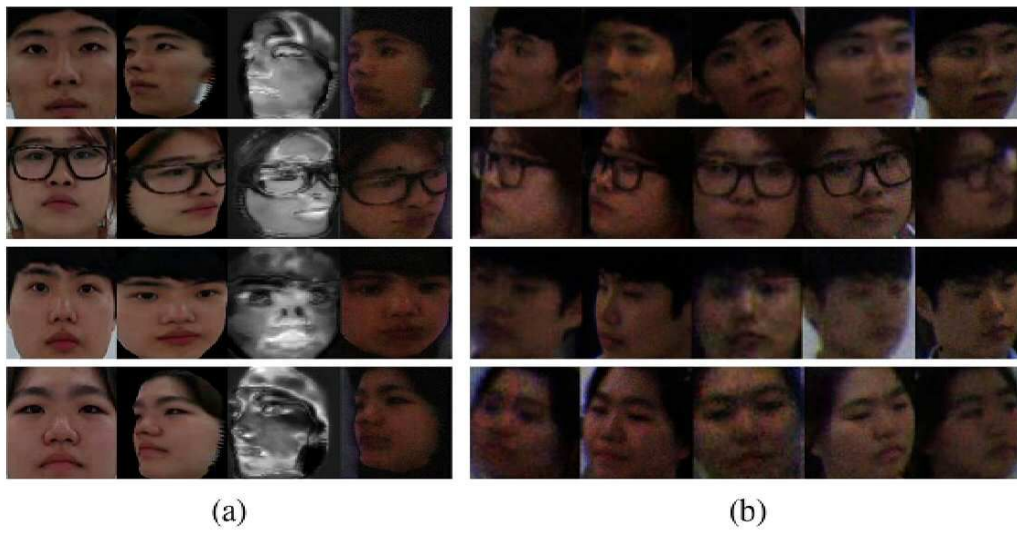
[0080] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

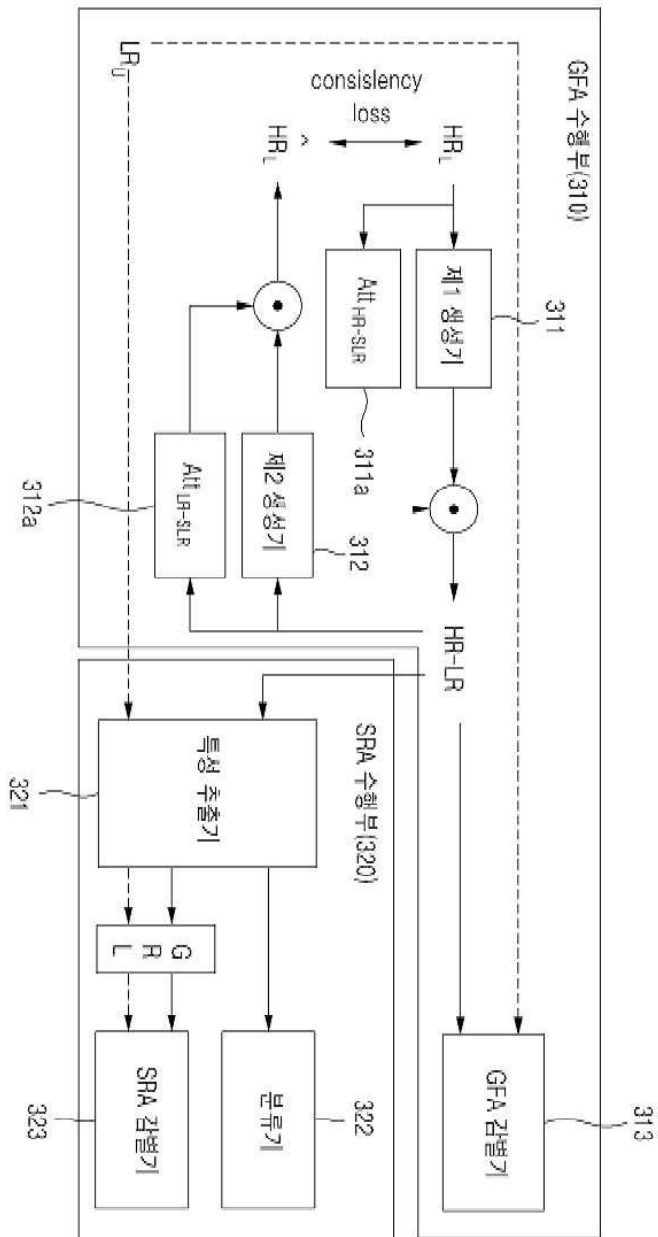
도면1



도면2



도면3



도면4

