



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0025906
(43) 공개일자 2023년02월23일

- (51) 국제특허분류(Int. Cl.)
G06V 40/16 (2022.01) G06N 3/04 (2023.01)
G06N 3/08 (2023.01) G06Q 30/06 (2023.01)
G06V 10/764 (2022.01) G06V 10/82 (2022.01)
- (52) CPC특허분류
G06V 40/171 (2022.01)
G06N 3/04 (2023.01)
- (21) 출원번호 10-2023-7002610
- (22) 출원일자(국제) 2021년06월29일
심사청구일자 2023년01월20일
- (85) 번역문제출일자 2023년01월20일
- (86) 국제출원번호 PCT/EP2021/067883
- (87) 국제공개번호 WO 2022/002961
국제공개일자 2022년01월06일
- (30) 우선권주장
63/045,303 2020년06월29일 미국(US)
FR2010128 2020년10월05일 프랑스(FR)

- (71) 출원인
로레알
프랑스공화국, 파리 F-75008, 튀 르와이얄 14
- (72) 발명자
아아라비 파르헬
캐나다 엘4비 2씨7 온타리오 리치몬드 힐 스트래
션 애버뉴 7
플라망 프레데릭 앙트와넵 레이몽 세르주
프랑스 92110 끌리쉬 튀 피에르 드레이퓌스 9
(뒷면에 계속)
- (74) 대리인
양영준, 백만기

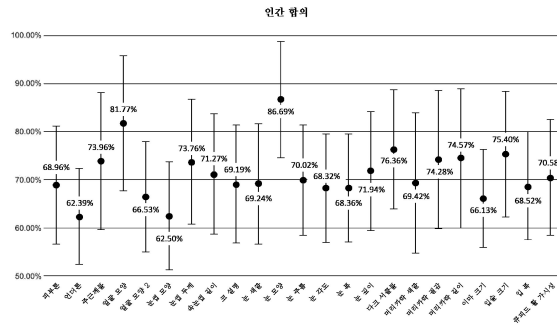
전체 청구항 수 : 총 46 항

(54) 발명의 명칭 **개선된 얼굴 속성 분류 및 그것의 사용을 위한 시스템들 및 방법들**

(57) 요약

얼굴 속성 예측 및 그것의 사용을 위한 방법들 및 시스템들을 포함하는 딥 러닝 지도 회귀 기반 모델이 설명되어 있다. 사용의 예는 이미지로부터 결정된 얼굴 속성 예측들에 응답하여 수정된 이미지를 제공하기 위한 증강 및/또는 가상 현실 인터페이스이다. 얼굴 속성들과 일치하는 얼굴 효과들이 인터페이스에 적용되도록 선택된다.

대표도



(52) CPC특허분류

G06N 3/08 (2023.01)
G06Q 30/0631 (2013.01)
G06Q 30/0643 (2013.01)
G06V 10/764 (2023.01)
G06V 10/82 (2022.01)
G06V 40/172 (2022.01)

(72) 발명자

호턴 제프리

캐나다 엠4씨 4케이3 온타리오 이스트 요크 시더베
일 애버뉴 295

지양 루오웨이

캐나다 엠5에이치 2엠1 온타리오 미씨소거 나스우
드 드라이브 1528

장 위쥬

캐나다 엠1피 3엘4 온타리오 스카버러 레스터우드
크레센트 4

위 즈

캐나다 엠4와이 2엑스5 온타리오 토론토 그린빌 스
트리트 25 유닛 2003

명세서

청구범위

청구항 1

얼굴 속성들(facial attributes)을 결정하기 위한 방법으로서,

상기 얼굴 속성들 각각에 대한 예측을 출력하기 위해 딥 러닝 및 지도 회귀(supervised regression)를 수행하는 네트워크 모델을 사용하여 얼굴의 소스 이미지를 프로세싱하는 단계

를 포함하고, 상기 네트워크 모델은 상기 얼굴 속성들을 예측하기 위해 개별 분류기들에 의한 분류를 위한 공유 피쳐들(shared features)의 피쳐 벡터를 생성하도록 딥 러닝을 수행하는 잔차 블록들을 포함하는 컨볼루션 신경망(CNN) 모델을 포함하는, 방법.

청구항 2

제1항에 있어서, 상기 얼굴 속성들은 상기 얼굴 속성들의 내재적 관계(intrinsic relationship) 및 이질성(heterogeneity)에 따라 복수의 개별 속성 그룹에 연관되고, 상기 소스 이미지는 상기 속성 그룹들 중의 개별 속성 그룹에 대해 구성된 개별 네트워크 모델에 의해 프로세싱되는, 방법.

청구항 3

제2항에 있어서, 상기 개별 속성 그룹들 중 하나는 색상 기반 얼굴 속성들에 대한 색상 속성 그룹이고, 상기 소스 이미지는 색상 기반 네트워크 모델에 의해 프로세싱되는, 방법.

청구항 4

제2항 또는 제3항에 있어서, 상기 개별 속성 그룹들 중 하나는 모양 기반 얼굴 속성들에 대한 모양 속성 그룹이고, 상기 소스 이미지는 모양 기반 네트워크 모델에 의해 프로세싱되는, 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 얼굴 속성들 중의 개별 얼굴 속성에 대해 개별 분류기에 의해 프로세싱될 피쳐 벡터를 복제하는 단계를 포함하는, 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 네트워크 모델은 복수의 개별 분류기를 포함하고, 상기 개별 분류기들 각각은 하나 이상의 완전 연결된 선형 레이어를 포함하고, 상기 개별 분류기들 각각은 상기 얼굴 속성들 중의 개별 얼굴 속성의 예측을 출력으로서 제공하는, 방법.

청구항 7

제6항에 있어서, 상기 복수의 개별 분류기는 상기 얼굴 속성들을 제공하기 위해 병렬로 수행되는, 방법.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 모델은 예측 레이어들이 트리밍된, 적용된 사전 훈련된 ResNet 기반 이미지 프로세싱 네트워크 모델을 포함하는, 방법.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 이미지 검색, 얼굴 인식, 및 제품 및/또는 서비스 추천 중 임의의 것을 수행하기 위해 상기 얼굴 속성들을 제공하는 단계를 포함하는, 방법.

청구항 10

제9항에 있어서, 애플리케이션은 상기 얼굴 속성들에 응답하여 상기 제품 및/또는 서비스 추천을 수행하고, 상

기 애플리케이션은 추천된 제품 또는 서비스에 연관된 효과를 시뮬레이션하도록 상기 소스 이미지를 수정하여 가상 사용 사용자 경험(virtual try on user experience)을 제공하는, 방법.

청구항 11

얼굴 이미지의 얼굴 속성들을 결정하기 위한 방법으로서,

공유 피쳐들의 피쳐 벡터를 생성하기 위해 잔차 블록들을 갖는 컨볼루션 신경망(CNN) 기반 백본 네트워크 모델을 사용하여 얼굴의 소스 이미지를 프로세싱하는 단계; 및

결정될 K개의 개별 얼굴 속성 각각에 대해 하나씩, 복수(K개)의 분류기를 사용하여 상기 피쳐 벡터를 프로세싱하는 단계

를 포함하고, 각각의 개별 분류기는 상기 K개의 개별 얼굴 속성 중 하나의 얼굴 속성의 예측을 출력하기 위해, 하나 이상의 완전 연결된 선형 레이어 및 예측 블록을 포함하는, 방법.

청구항 12

제11항에 있어서, 상기 얼굴의 N개의 소스 이미지에 대해 상기 프로세싱하는 단계들을 수행하고, 각각의 얼굴 속성의 N개의 예측으로부터 상기 K개의 개별 얼굴 속성의 최종 예측을 결정하기 위해 예측 해결 규칙(prediction resolving rule)을 사용하는 단계를 포함하는, 방법.

청구항 13

제11항 또는 제12항에 있어서, 제1 네트워크 모델은 상기 복수의 분류기 중에서 색상 기반 얼굴 속성들을 예측하도록 구성된 개별 분류기들에 의한 프로세싱을 위한 색상 기반 피쳐 벡터를 생성하기 위해 색상 기반 얼굴 속성들에 대해 상기 소스 이미지를 프로세싱하는, 방법.

청구항 14

제11항 내지 제13항 중 어느 한 항에 있어서, 제2 네트워크 모델은 상기 복수의 분류기 중에서 모양 기반 얼굴 속성들을 예측하도록 구성된 개별 분류기들에 의한 프로세싱을 위한 모양 기반 피쳐 벡터를 생성하기 위해 모양 기반 얼굴 속성들에 대해 상기 소스 이미지를 프로세싱하는, 방법.

청구항 15

제11항 내지 제14항 중 어느 한 항에 있어서, 상기 K개의 분류기에 의한 병렬 프로세싱을 위해 상기 피쳐 벡터를 복제하는 단계를 포함하는, 방법.

청구항 16

현실을 시뮬레이션하기 위한 방법으로서,

얼굴의 소스 이미지로부터 복수의 얼굴 속성을 결정하고, 얼굴 속성 분류 네트워크 모델을 사용하여 상기 소스 이미지를 프로세싱하는 단계;

상기 얼굴 속성들 중 적어도 하나에 응답하여 상기 얼굴에 적용하기 위한 적어도 하나의 얼굴 효과를 결정하는 단계; 및

디스플레이를 위해 상기 얼굴에 상기 적어도 하나의 얼굴 효과를 적용하기 위해 상기 소스 이미지를 프로세싱하는 단계

를 포함하는, 방법.

청구항 17

제16항에 있어서, 상기 얼굴 속성들은 제1항 내지 제15항 중 어느 한 항의 방법을 이용하여 결정되는, 방법.

청구항 18

제16항 또는 제17항에 있어서, 상기 적어도 하나의 얼굴 효과는 상기 얼굴에 적용될 적어도 하나의 메이크업 제품 및/또는 기법의 시뮬레이션인, 방법.

청구항 19

제16항 내지 제18항 중 어느 한 항에 있어서, 상기 적어도 하나의 얼굴 효과를 적용하기 위해 상기 소스 이미지를 프로세싱하는 단계는 상기 얼굴 효과를 적용하도록 구성된 딥 러닝 네트워크를 사용하는 단계를 포함하는, 방법.

청구항 20

제18항 또는 제19항에 있어서, 얼굴 속성들과 관련하여 메이크업 제품들 및/또는 기법들을 저장하는 데이터 저장소로부터 하나 이상의 메이크업 제품 및/또는 기법을 선택하기 위해 상기 얼굴 속성들 중 적어도 하나를 사용하는 단계를 포함하는, 방법.

청구항 21

전자상거래 서비스를 위해 제품들을 추천하기 위한 방법으로서,

얼굴의 소스 이미지로부터 결정된 복수의 얼굴 속성을 수신하는 단계 - 상기 소스 이미지는 상기 복수의 얼굴 속성을 생성하기 위해 얼굴 속성 분류 네트워크 모델을 사용하여 프로세싱됨 -;

제품들에 적합한 얼굴 속성들에 관련하여 상기 제품들을 저장하는 데이터 저장소로부터 적어도 하나의 제품을 선택하기 위해 상기 얼굴 속성들 중 적어도 일부를 사용하는 단계; 및

제품들을 구매하기 위해 전자상거래 인터페이스에 프레젠테이션하기 위한 추천으로서 상기 적어도 하나의 제품을 제공하는 단계

를 포함하는, 방법.

청구항 22

제21항에 있어서, 상기 얼굴 속성들은 제1항 내지 제15항 중 어느 한 항의 방법을 이용하여 결정되는, 방법.

청구항 23

제21항 또는 제22항에 있어서, 상기 제품들은 메이크업 제품들을 포함하는, 방법.

청구항 24

제23항에 있어서,

상기 메이크업 제품들은 미리 결정된 메이크업 룩들을 정의하기 위해 상기 데이터 저장소에서 다양하게 연관되고,

상기 방법은 상기 미리 결정된 메이크업 룩들 중 하나의 식별을 수신하는 단계를 포함하고;

상기 얼굴 속성들 중 적어도 일부를 사용하는 단계는 상기 적어도 하나의 제품을 선택할 때 상기 미리 결정된 메이크업 룩들 중 하나에 응답하는, 방법.

청구항 25

제24항에 있어서, 상기 메이크업 제품들 각각은 복수의 메이크업 타입 중 하나에 연관되고, 상기 방법은 상기 추천을 정의하기 위해 메이크업 타입들 각각에 대해, 상기 얼굴 속성들에 응답하여 적어도 하나의 제품을 선택하는 단계를 포함하는, 방법.

청구항 26

제24항에 있어서, 상기 메이크업 타입들은 페이스 제품 타입, 아이 제품 타입, 브로우 제품 타입, 및 립 제품 타입을 포함하는, 방법.

청구항 27

제21항 내지 제26항 중 어느 한 항에 있어서, 추천 제품들을 사용하기 위한 기법을 추천하는 단계를 더 포함하는, 방법.

청구항 28

제21항 내지 제27항 중 어느 한 항에 있어서, 상기 전자상거래 인터페이스는 상기 소스 이미지를 프로세싱함으로써 상기 얼굴에 적용되는 상기 적어도 하나의 제품의 시뮬레이션을 제공하는, 방법.

청구항 29

제21항 내지 제28항 중 어느 한 항에 있어서, 상기 추천 제품들 중 적어도 일부를 구매하기 위해 전자상거래 쇼핑 서비스를 제공하는 단계를 더 포함하는, 방법.

청구항 30

시스템으로서,

제1항 내지 제29항의 방법 청구항들 중 어느 하나를 수행하도록 구성된 컴퓨팅 디바이스를 포함하는, 시스템.

청구항 31

시스템으로서,

얼굴의 소스 이미지로부터 복수의 얼굴 속성을 추출하고, 추출된 복수의 얼굴 속성에 기초하여 하나 이상의 얼굴 효과를 생성하도록 구성된 프로세싱 회로부를 포함하는 얼굴 속성 유닛; 및

상기 소스 이미지에 적어도 하나의 얼굴 효과를 적용하고, 효과 적용된 소스 이미지의 하나 이상의 가상 인스턴스를 전자상거래 인터페이스 상에 생성하도록 구성된 프로세싱 회로부를 포함하는 얼굴 효과 유닛

을 포함하는, 시스템.

청구항 32

제31항에 있어서,

상기 복수의 얼굴 속성 중 하나 이상과 관련하여 제품들을 저장하는 데이터 저장소로부터 적어도 하나의 제품을 결정하고, 제품들을 구매하기 위해 전자상거래 인터페이스 상에서 제품 추천의 하나 이상의 가상 인스턴스를 생성하도록 구성된 프로세싱 회로부를 포함하는 사용자 경험 유닛

을 더 포함하는, 시스템.

청구항 33

제31항 또는 제32항에 있어서, 상기 얼굴 속성 유닛은 얼굴의 소스 이미지로부터 상기 복수의 얼굴 속성 각각에 대한 예측을 출력하기 위해 딥 러닝 및 지도 회귀를 수행하는 네트워크 모델을 실행하도록 구성된 프로세싱 회로부를 포함하는, 시스템.

청구항 34

제31항 내지 제33항 중 어느 한 항에 있어서, 상기 얼굴 효과 유닛은 상기 효과 적용된 소스 이미지의 적어도 일부, 및 메이크업 제품 또는 메이크업 적용 기법 중 적어도 하나를 포함하는 가상 표현을 생성하도록 구성된 프로세싱 회로부를 포함하는, 시스템.

청구항 35

제31항 내지 제33항 중 어느 한 항에 있어서, 상기 얼굴 효과는 상기 얼굴 속성들 중 하나에 응답하는 주석을 포함하는, 시스템.

청구항 36

제35항에 있어서, 상기 주석이 상기 얼굴 속성에 대한 얼굴 속성 값을 포함하는 것; 및 상기 주석이 상기 얼굴 속성이 결정된 상기 소스 이미지의 영역에 위치되는 것 중 적어도 하나인, 시스템.

청구항 37

방법으로서,

추론 시간 이미지(inference time image)로부터 속성들을 예측하는 속성 분류기를 훈련시키기 위한 속성 데이터 세트를 저장하는 단계 - 상기 속성 데이터 세트는 복수의 속성을 나타내는 복수의 이미지를 포함하고, 상기 속성들 각각은 복수의 개별 속성 값을 가짐 -;

상기 복수의 이미지 중의 각각의 이미지에 대해 소프트 레이블들을 상기 데이터 세트에 저장하는 단계 - 상기 소프트 레이블들은 독립적으로 행동하는 복수의 개별 인간 이미지 주석 작성자에 의해 결정되는 속성별 개별 속성 값들을 포함함 -; 및

상기 속성 분류기를 훈련시키기 위해 상기 속성 데이터 세트를 제공하는 단계를 포함하는, 방법.

청구항 38

제37항에 있어서, 상기 속성 데이터 세트를 사용하여 상기 속성 분류기를 훈련시키는 단계를 더 포함하는, 방법.

청구항 39

제38항에 있어서, 훈련 시에, 상기 소프트 레이블들 중에서 진실을 선택하기 위해 "진실" 해결 규칙을 사용하는 단계를 포함하는, 방법.

청구항 40

제37항 내지 제39항 중 어느 한 항에 있어서, 상기 속성들은 상기 속성들의 내재적 관계 및 이질성에 따라 복수의 개별 속성 그룹에 연관되고, 상기 속성 분류기는 상기 속성 그룹들 각각에 대해 하나씩 복수의 서브-모델을 포함하는, 방법.

청구항 41

제40항에 있어서, 상기 개별 속성 그룹들 중 하나는 색상 기반 속성들에 대한 색상 속성 그룹인, 방법.

청구항 42

제40항 또는 제41항에 있어서, 상기 개별 속성 그룹들 중 하나는 모양 기반 속성들에 대한 모양 속성 그룹인, 방법.

청구항 43

제37항 내지 제42항 중 어느 한 항에 있어서, 상기 속성들은 얼굴 속성들인, 방법.

청구항 44

제43항에 있어서, 상기 얼굴 속성들은 표 1의 속성들을 포함하는, 방법.

청구항 45

시스템으로서,

제37항 내지 제43항 중 어느 한 항의 방법을 수행하도록 구성된 컴퓨팅 디바이스를 포함하는, 시스템.

청구항 46

시스템으로서,

제37항 내지 제43항 중 어느 한 항의 방법에 따라 정의된 속성 분류기로 구성된 컴퓨팅 디바이스를 포함하는, 시스템.

발명의 설명

기술 분야

[0001] 상호 참조

[0002] 본 출원은 2020년 6월 29일자로 출원되고 발명의 명칭이 "개선된 얼굴 속성 분류 및 그것의 사용을 위한 시스템 들 및 방법들(Systems and Methods for Improved Facial Attribute Classification and Use thereof)"인 미국 가출원 제63/045,303호, 및 2020년 10월 5일자로 출원되고 발명의 명칭이 "개선된 얼굴 속성 분류 및 그것의 사용을 위한 시스템들 및 방법들(Systems and Methods for Improved Facial Attribute Classification and Use thereof)"인 선행 프랑스 특허 출원 제2010128호에 대한 우선권 또는 이익을 주장하며, 허용가능한 경우, 이들 각각의 전체 내용이 여기에 참조에 의해 포함된다.

[0003] 분야

[0004] 본 출원은 머신 러닝을 이용한 이미지 프로세싱, 및 더 구체적으로는 얼굴 속성 분류 태스크에 관한 것이다. 얼굴 이미지가 주어지면, 얼굴 속성 분류는 원하는 속성들이 얼굴 인물사진(facial portrait)에 존재하는지를 추정하는 것을 목표로 한다. 다운스트림 애플리케이션들은 다양한 이미지 프로세싱 및 다른 애플리케이션들을 포함한다.

배경 기술

[0005] 얼굴 속성 분류(facial attribute classification)(FAC)는 대상의 얼굴 이미지가 주어질 때, 주어진 대상의 얼굴 속성들을 분류하려고 시도하는 태스크이다. FAC는 이미지 검색, 얼굴 인식, 사람 재식별, 및 추천 시스템들을 포함하는 다수의 다운스트림 애플리케이션에 매우 중요하다. 얼굴의 이미지는 입술 크기 및 모양, 눈 색깔 등, 및 머리카락 색깔, 성별 등과 같이 얼굴 자체에 엄격하게 제한되지 않을 수 있는 다른 것들과 같은 다수의 속성(일반적으로 얼굴 속성들로 표시됨)을 예측하기 위해 분석될 수 있다.

[0006] 최근, 컨볼루션 신경망(Convolutional Neural Network)(CNN)의 뛰어난 성능으로 인해, 대부분의 최신 FAC 방법 들은 얼굴 속성들을 분류하기 위해 CNN을 활용한다. 대략적으로 말하자면, 이러한 방법들은 다음과 같이 범주 화될 수 있다: (1) 단일-레이블 학습 기반 FAC 방법들, 및 (2) 다중-레이블 학습 기반 FAC 방법들. 단일-레이 블 학습 기반 FAC 방법들은 일반적으로 얼굴 이미지들의 CNN 피쳐들을 추출한 다음, 서포트 벡터 머신(Support Vector Machine)(SVM) 분류기에 의해 얼굴 속성들을 분류한다. 그러나, 이러한 방법들은 각각의 속성을 개별적 으로 예측하므로, 속성들 간의 상관관계들을 무시한다. 대조적으로, 복수의 속성을 동시에 예측할 수 있는 다 중-레이블 학습 기반 FAC 방법들은 CNN의 하위 레이어들로부터 공유 피쳐들을 추출하고, CNN의 상위 레이어들 상에서 속성 특정 분류기들을 학습한다.

발명의 내용

[0007] 실시예들에 따르면, 얼굴 속성 예측 및 그것의 사용을 위한 방법들 및 시스템들 및/또는 컴퓨터 디바이스들을 포함하는 딥 러닝 지도 회귀 기반 모델(deep learning supervised regression based model)이 설명된다. 사용 의 예는 이미지로부터 결정된 얼굴 속성 예측들에 응답하여 수정된 이미지를 제공하기 위한 (예를 들어, 컴퓨팅 디바이스 또는 방법을 통한) 증강 및/또는 가상 현실 인터페이스이다. 실시예에서, 얼굴 속성들과 일치하는 얼굴 효과들이 인터페이스에 적용되도록 선택된다. 실시예에서, 얼굴 효과들은 메이크업 제품들에 연관된 메이크 업 효과들을 포함한다. 실시예에서, 전자상거래 인터페이스는 증강 및/또는 가상 현실 인터페이스를 제공하고, 얼굴 속성들에 응답하여 메이크업 제품들을 추천하기 위한 추천 컴포넌트를 포함한다.

[0008] 또한, 실시예들에 따라 컴퓨터 프로그램 제품들이 제공되며, 이러한 제품 양태는 명령어들을 저장하는 (예를 들 어, 비-일시적) 저장 디바이스를 포함하고, 이러한 명령어들은 컴퓨팅 디바이스의 프로세서에 의해 실행될 때, 예컨대 본 명세서의 임의의 각각의 실시예에 따른 방법을 수행하도록 컴퓨팅 디바이스를 구성한다.

도면의 간단한 설명

[0009] 도 1 및 도 2는 복수의 (얼굴) 속성에 대한 속성 값 결정을 수행할 때의 인간 합의(human agreement)의 평가, 및 23개의 얼굴 속성에 대한 얼굴 속성 분류 태스크에 대한 인간 성과를 보여주는 그래프들이다.

도 3은 실시예에 따른 딥 러닝 네트워크 아키텍처를 도시하는 블록도이다.

도 4는 실시예에 따른 컴퓨팅 시스템, 예를 들어 네트워크를 도시하는 블록도이다.

도 5, 도 6a, 도 6b, 도 7a, 도 7b, 도 8, 도 9a, 도 9b, 도 10a, 도 10b, 도 10c, 도 10d, 도 11a, 도 11b 및 도 11c는 실시예에 따른 사용자 애플리케이션의 그래픽 사용자 인터페이스에 대한 디스플레이 화면들 또는 그 일부의 도시이다.

도 12, 도 13, 도 14, 도 15 및 도 16은 본 명세서의 실시예들에 따른 개별 동작들의 흐름도이다.

도 17은 본 명세서의 실시예에 따른 시스템의 블록도이다.

본 개념은 본 명세서에서 첨부 도면들을 참조하여 설명되는 특정 실시예들을 통해 가장 잘 설명되며, 여기서 유사한 참조 번호들은 전체적으로 유사한 피쳐들을 나타낸다. 본 명세서에서 사용되는 발명이라는 용어는 단지 실시예들 그 자체가 아니라 아래에 설명되는 실시예들의 기초가 되는 발명적 개념을 의미하는 것으로 의도된다는 점을 이해해야 한다. 일반적인 발명적 개념은 아래에 설명된 예시적인 실시예들에 제한되지 않으며 이하의 설명들은 이러한 견지에서 읽어야 한다는 것을 추가로 이해해야 한다.

발명을 실시하기 위한 구체적인 내용

- [0010] 얼굴 속성 분류(FAC)는 대상의 얼굴 이미지가 주어질 때, 주어진 대상의 얼굴 속성들을 분류하려고 시도하는 태스크이다. FAC는 이미지 검색, 얼굴 인식(예를 들어, 두 장의 사진이 동일한 얼굴을 나타내는지를 결정), 얼굴 로컬라이제이션, 및 추천 시스템들을 포함하는 다수의 다운스트림 애플리케이션에 매우 중요하다.
- [0011] 본 명세서의 실시예들에서, 태스크는 속성들의 미리 결정된 세트에 대한 분류로서 정의된다. 각각의 속성은 각각의 대상이 분류될 수 있는 잠재적 클래스들(예를 들어, 속성 값들)의 세트로 설명된다. 예를 들어, 피부 톤 속성은 아주 밝음(Fair), 밝음(Light), 중간(Medium), 약간 어두움(Tan), 어두움(Deep) 및 아주 어두움(Very Deep)의 총 6개의 클래스를 갖는다. 따라서, 이러한 특정 속성에 대한 태스크는 사람의 이미지가 주어질 때, 6개의 클래스 중 어느 것이 주어진 이미지를 가장 잘 표현하는지를 예측하는 것이다.
- [0012] 속성들의 미리 결정된 세트는 크게 두 개의 상이한 타입: 색상 속성 또는 모양 속성으로 범주화될 수 있다. 색상 속성들의 예들은 머리카락 색상, 피부 톤 및 언더톤을 포함한다. 모양 속성들의 예들은 눈 모양, 눈 주름(eye crease), 및 입술 크기를 포함한다.
- [0013] 최근, 컨볼루션 신경망(CNN)의 뛰어난 성능으로 인해, 대부분의 최신 FAC 방법들은 얼굴 속성들을 분류하기 위해 CNN을 활용한다. 대략적으로 말하자면, 이러한 방법들은 다음과 같이 범주화될 수 있다: (1) 단일-레이블 학습 기반 FAC 방법들, 및 (2) 다중-레이블 학습 기반 FAC 방법들. 단일-레이블 학습 기반 FAC 방법들은 일반적으로 얼굴 이미지들의 CNN 피쳐들을 추출한 다음, 서포트 벡터 머신(SVM) 분류기에 의해 얼굴 속성들을 분류한다. 그러나, 이러한 방법들은 각각의 속성을 개별적으로 예측하므로, 속성들 간의 상관관계들을 무시한다. 대조적으로, 다수의 속성을 동시에 예측할 수 있는 다중-레이블 학습 기반 FAC 방법들은 CNN의 하위 레이어들로부터 공유 피쳐들을 추출하고, CNN의 상위 레이어들 상에서 속성 특정 분류기들을 학습한다.
- [0014] 전형적인 학문적 세팅과 달리, 여기에서 주로 설명되는 시스템 및 그것의 다운스트림 태스크들은 대부분 패션 및 뷰티 산업에 관한 것이다. 일반 학문적 데이터세트 내의 속성들은 현재 태스크에 최소한의 가치를 제공하는 더 넓은 범위의 얼굴 피쳐들을 다루는 경향이 있다. 대조적으로, 새로운 맞춤형 데이터세트는 잘 설계된 속성 타입들에 대한 소프트 레이블들을 갖는 이미지들을 저장함으로써 정의된다. 새로운 데이터세트는 3790명의 대상을 포함하고, 각각은 상이한 환경들에서 촬영한 1개 내지 5개의 이미지를 갖는다. 총 23개의 속성이 있으며, 각각의 속성은 추가로 설명되는 소프트 레이블들을 사용하여, 독립적으로 행동하는 복수(예를 들어, 6개)의 개별 사람 주석 작성자(human annotators)에 의해 레이블 지정된다.
- [0015] 인간 주석 작성자들이 얼굴 속성들의 인스턴스들에 대한 특정 값들에 항상 합의하지는 않는다. 예를 들어, 한 명의 주석 작성자는 값이 평균이라고 결론을 내릴 수 있고, 다른 주석 작성자는 값이 높다고 결론을 내릴 수 있다. 따라서, 얼굴 속성 분류와 같은 분류 태스크들에 대한 인간의 수행은 다양할 수 있다. 즉, 인간은 특정 속성 값들에 관해 상이한 의견들을 갖거나 상이한 결론들에 도달할 수 있다. 실시예에서, 속성 데이터 세트는 그라운드 트루스 속성 값들에 관한 개별 "의견들"을 집계하기 위해 복수의 인간으로부터의 주석들로 컴파일된다.
- [0016] 이러한 접근방식은 하드 또는 단일 그라운드 트루스를 사용하는 전통적인 데이터 세트 주석 접근방식들과 대조된다. 예로서, 하나의 속성은 약한, 중간 또는 심한 주름 값을 갖는 주름들에 관한 것이다. 이 예는 값들 사이의 선형 관계를 보여준다. 대조적으로, 턱 모양은 둥글거나 뾰족하거나 사각형이다. 이들은 수학적으로 관련되어 있지만, 현재 태스크에 대한 새로운 데이터 세트를 생성할 때 수행된 것은 알고리즘이 각각의 이미지에

사용하는 속성들에 대해 고유한 그룹들 또는 버킷들을 정의하는 것이다. 매우 다양한 텍 모양들을 고려하여, 분류기 알고리즘은 각각의 버킷에 대한 신뢰도를 제시하며, 이는 인간이 고객의 뺨 윤곽을 그리는 방법을 결정할 때와 동일한 방식으로, 최종 제품 추천들에서 가장 가능성이 높은 버킷을 사용할 수 있음을 의미한다. 이러한 태스크에 대한 전통적인 접근 방식들은 주석 작성자들 전부가 말한 내용을 평균화함으로써 이미지에 대한 "올바른" 출력이 무엇인지 이해하기 위해 "하드 레이블"을 사용한다. 대신에, "소프트 레이블" 접근 방식을 사용하여, "올바른" 답이 없을 수 있음이 인식된다. 이는 주석 작성자들이 합의하지 않을 수 있지만, 알고리즘이 그것이 가장 높은 신뢰도를 갖는다고 생각하는 것을 결정하더라도, 알고리즘은 모든 출력을 추천할 유연성을 가질 것임을 의미한다.

[0017] 도 1 및 도 2는 23개의 얼굴 속성에 대한 얼굴 속성 분류 태스크에 대한 인간 합의 및 인간 수행을 보여주는 그래프들이다.

[0018] 인간 합의(Human Agreement)는 주석 작성자들이 서로 얼마나 합의하는지를 평가하도록 설정된다. 실시예에서, 6명의 인간 주석 작성자가 있다. 주석을 작성할 때, 모든 경우들에서 주석 작성자들이 서로 잘 합의하지는 않는다. 예를 들어, 머리카락 색깔 주석에 대해, 네 사람은 갈색이라고 말할 수 있지만 두 사람은 짙은 갈색 또는 심지어 검은색이라고 말할 수 있다. 제1 그래프는 일치율을 나타내며, 범위의 중간 지점이 평균 값이다. 범위는 합의의 표준 편차(stddev)를 보여준다. 합의는 대다수 투표에 기초하여 계산되므로, 위의 4/2 예에서, 이러한 특정 속성(예에서 머리카락 색깔)에 대한 이러한 특정 대상에 대한 합의는 66%이다. 3/3은 50%로 카운트될 것이고, 2/2/2는 33%로 카운트될 것이다. 다음으로, 평균 및 stddev는 전체 데이터세트에 걸쳐 각각의 속성에 대해 결정된다.

[0019] 인간 수행(Human Performance)은 인간 자신이 얼마나 잘 수행하는지를 평가하는 것이다. 따라서, 그것은 기본적으로 단지 "얼마나 많은 사람들이 정답을 맞췄는지"이다. "정답"의 정의는 실시예를 위해 정의되는 요소이다. 실시예에서, 그것은 대다수 투표를 따른다. 그러나, 명확한 다수가 없는 인스턴스들(예를 들어, 사진에 대한 주석들의 세트)이 있다. 예를 들어, 사진은 머리카락 색깔에 대해, 2개의 클래스에서 3/3 투표를 받았거나 3개의 클래스에서 2/2/2 투표를 받았는가? 대다수 투표는 없다. 그러한 경우에, 실시예에서, "동일 투표 포함" 및 "동일 투표 포함하지 않음"으로 명명된 두 세트의 메트릭이 정의된다. "동일 투표 포함"은 대다수 투표가 없는 경우들을 포함할 것이고, "동일 투표 포함하지 않음"은 명확한 다수가 없는 경우를 단순히 무시할 것이다. 무시에 의해, 이러한 인스턴스들은 계산의 분자 또는 분모에 포함되지 않는다. 그들은 무시된다. 반면에, "동일 투표 포함"은 모든 최상위 클래스를 올바른 것으로 카운트할 것이다. 따라서, 3/3에 대해, 둘 다 올바르고, 우리는 인간 수행에 대해 100%를 갖는다. 2/2/2에 대해 여전히 100%이다. 2/2/1/1에 대해 66%이다.

[0020] 표 1은 23가지 속성 및 이에 대한 값들을 보여준다.

표 1

[0021]

	속성	속성 값
1	피부톤	아주 밝음, 밝음, 중간, 약간 어두움, 어두움, 아주 어두움
2	언더톤	중성, 적색, 황색
3	주근깨	없음, 약간 있음, 중간, 많음
4	얼굴형	계란형, 둥근형, 하트형, 사각형, 긴형, 다이아몬드형
4a	얼굴 길이	깊, 짧음
4b	턱 모양	둥근형, 사각형, 뾰족형
5	눈썹 모양	아치형, 부드러운 아치형, 직선형, 곡선형
6	눈썹 두께	얇음, 중간, 두꺼움
7	속눈썹 길이	짧음, 중간, 길
8	코 설명	좁음, 중간, 넓은
9	눈 색깔	블루, 그레이, 그린 헤이즐, 브라운, 블랙
10	눈 모양	둥근형, 아몬드형
11	눈 주름	외꺼풀, 처짐, 중성/주름있음
12	눈 각도	하향, 중립, 상향
13	눈 폭	넓은 세트, 중립, 가까운 세트
14	눈 깊이	깊은 세트, 중립, 돌출
15	다크 서클	없음, 있음
16	머리카락 색깔	대머리, 흰색, 회색, 밝은 금발, 금발, 갈색, 짙은 갈색, 검은색, 딸기색, 구리색, 빨간색, 적갈색, 기타
17	머리카락 질감	스트레이트, 웨이브, 컬, 곱슬머리/아프로, 대머리

18	머리카락 길이	매우 짧음, 짧음, 중간, 길, 매우 길, 대머리
19	이마	짧음, 평균, 길
20	입술 크기	얇음, 중간, 두꺼움
21	입 폭	좁음, 평균 폭
22	큐피드의 활 가시성	평평함, 중간, 높음/하트

- [0022] 얼굴 모양은 얼굴 속성들 4a(얼굴 길이) 및 4b(턱 모양)에 대한 예측들로부터의 신호들을 결합함으로써 예측되는 2차 속성이다. 예를 들어, 계란형 = 긴 얼굴 + 둥근 턱, 둥근형 = 짧은 얼굴 + 둥근 턱이다.
- [0023] 훈련 목적을 위해, 데이터는 훈련, 검증 및 테스트 서브세트들로 분할되며, 이들은 그에 따라 2274, 758 및 758 명의 대상을 포함한다.
- [0024] 겹보기 얼굴 아키텍처 예측 문제는 딥 러닝을 사용한 컴퓨터 구현을 위한 지도 회귀 문제로 된다. 속성들은 동일한 개인의 1개 내지 5개의 이미지로부터 평가된다.
- [0025] 얼굴 속성들의 내재적 관계와 이질성을 효과적으로 활용하기 위해, 속성들은 상이한 그룹들로 나누어질 수 있다. 실시예에 따르면, 속성들은 2개의 개별 그룹으로 나누어졌다: 색상 속성들(예를 들어, 머리카락 색깔, 눈 색깔 및 언더톤) 및 모양 속성들(예를 들어, 입술 크기, 눈 주름 및 코 타입). 각각의 속성 그룹에서 유사한 아키텍처를 공유하는 두 개의 서브모델이 훈련되었다. 추론하는 동안, (각각의 개별 속성에 대한) 예측들이 마지막에 집계된다. 편의를 위해, 두 개의 서브모델, 즉 색상에 대한 하나의 서브모델 및 모양에 대한 하나의 서브모델이 집계되어, 얼굴 속성 분류를 위한 수퍼 모델을 형성한다. 그러나, 분류를 위해 개별 속성들만이 요구되는 경우, 서브모델들 중 어느 하나가 유용하다.
- [0026] 두 개의 서브모델(예를 들어, 모양 서브모델 및 색상 서브모델)은 동일한 구조를 공유하지만 상이한 사전 프로세싱된 데이터에 대해 훈련된다(여기서, 사전 프로세싱은 훈련 시작 전의 임의의 데이터 증대를 의미하며, 얼굴 검출 및 얼굴 자르기, 크기 조정 등을 포함하지만 이에 제한되지는 않음). 색상 기호들에 대해, 머리카락 색깔 및 피부톤과 같은 속성들은 또한 색상 추출을 안내하기 위한 배경을 필요로 하기 때문에(예를 들어, 그늘진 환경에서, 배경이 잘리면, 밝은 피부 톤 이미지가 어두운 피부 톤으로 잘못 분류될 수 있음), 데이터세트는 모양 속성들보다 더 많은 배경을 갖도록 사전 프로세싱된다. 모양 모델은 최소한의 배경을 갖도록 잘라낸 얼굴에 대해 훈련되지만, 색상 모델은 더 많은 배경 정보를 포함한다.
- [0027] 실시예에서, 2개의 서브 모델에서의 사전 프로세싱은 다음을 포함한다:
- [0028] 1. 원본 이미지를 RGB 공간으로 변환한다;
- [0029] 2. 이미지에 존재하는 얼굴을 검출한다(임의의 표준 얼굴 검출 알고리즘이 작동해야 한다); 이는 색상에 대한 선택사항이지만 정확도를 높이는 데 도움이 될 것이다.
- [0030] 3. 얼굴 상자 크기 조정 - 색상 모델: 검출된 얼굴 상자를 1.7x 폭 및 높이로 확대, 모양 모델: 검출된 얼굴 상자를 1.2x 폭 및 높이로 확대; 및
- [0031] 4. 단계 3으로부터의 이미지를 모델의 입력 크기로 크기조정.
- [0032] 따라서, 데이터 세트는 관련 기호들: 색상 관련 기호들 및 모양 관련 기호들을 버킷팅함으로써 두 가지 알고리즘을 훈련하기 위해 사용된다. 이것은 관련있는 기호들로부터 학습된 알고리즘들을 의미했다. 예를 들어, 밝은 머리카락을 가진 사람들은 종종 파란 눈을 갖고, 어두운 피부를 가진 사람들은 종종 갈색 눈을 갖는다. 그러나, 파란 눈을 가진 사람들은 상관관계를 거의 갖지 않고서, 가능한 임의의 얼굴 모양을 가질 수 있다. 관련 기호들에 대한 관련 계산들을 재사용하고 관련 기호들을 연결함으로써, 각각의 알고리즘에 대한 결과 코드는 매우 작고 매우 빠르다. 두 가지 알고리즘은 서로를 과도하게 복잡하게 만들지 않으면서, 사용자의 얼굴 특징들의 완벽한 그림을 함께 만들 수 있다.
- [0033] 도 3은 딥 러닝 네트워크 아키텍처를 도시하는 블록도이다. 네트워크(300)는 입력 레이어(304)에서 소스 이미지를 프로세싱하기 위한 컨볼루션 신경망(CNN)(302)을 포함한다. 실시예에서, CNN(302)은 공유 피쳐들을 추출하기 위해, 복수의 잔차 블록(306)을 갖는 잔차 네트워크 기반 백본을 사용하여 구성된다. 예를 들어, 잔차 네트워크 기반 백본은 ResNet을 사용하여 구성된다("Deep Residual Learning for Image Recognition", He, Kaiming et. al, 2015-12-10, URL arxiv.org/abs/1512.03385에서 입수가가능함). 실시예에서, 예를 들어 ImageNet(Stanford University의 Stanford Vision Lab(URL image-net.org)으로부터 입수가가능한 조직화된 이미

지들의 데이터베이스)에서 사전 훈련된 모델들은 적응되어 백본 네트들로서 사용된다. 그러한 모델들의 마지막 몇 개의 예측 레이어는 당면한 목적에 맞게 - 예를 들어, 원하는 대로 공유 피쳐 데이터를 생성하는 인코더로서 - 모델들을 적응시키기 위해 제거된다.

[0034] 편평화된 피쳐 벡터(flattened feature vector)(308)(예를 들어, 평균 풀링 사용)는 백본 네트(302)로부터 획득된다. 예를 들어, 실시예에서, ResNet18 기반 백본 피쳐 네트로부터, 출력은 크기 224×224 의 이미지에 대해 치수 $7 \times 7 \times 512$ 의 벡터를 갖는다. 실시예에서, 평균 풀링은 512개의 위치 각각에서 7×7 항의 평균을 취함으로써 1×512 의 편평화된 벡터를 제공한다.

[0035] 피쳐 벡터(308)는 K개의 얼굴 속성 각각에 대해 복수(K개)의 분류기(310)에 의해 (예를 들어, 병렬로) 프로세싱하기 위해 복제된다. 각각의 개별 분류기(예를 들어, $312_1, 312_2, \dots, 312_k$)는 각각의 속성에 대한 최종 예측을 출력하기 위해, 하나 또는 복수의 완전 연결된 선형 레이어($314_1, 314_2, \dots, 314_k$) 및 예측 블록($316_1, 316_2, \dots, 316_k$)을 포함한다. 모양 모델에서, 각각의 분류기 헤드는 정류된 선형 유닛(rectified linear unit)(ReLU) 활성화들을 갖는 2개의 완전 연결된 선형 레이어로 구성된다. 히든 레이어 크기는 50이다. (이는 피쳐 크기가 N이라고 가정하면, 제1 FC 레이어가 피쳐를 크기 50으로 압축한 다음, 제2 FC 레이어가 크기 50 피쳐에서 작업하고 타겟 출력 크기와 동일한 크기를 갖는 피쳐 벡터를 출력할 것임을 의미한다). 색상 모델에서, 각각의 분류기 헤드는 단 하나의 FC 선형 레이어로만 구성된다. (크기 N인 피쳐를 타겟 출력 크기로 직접 변환함). 전형적으로, 모델이 깊을수록(즉, 더 많은 레이어가 존재할수록) 모델의 용량이 커진다. 모양 속성들은 완전 연결된 레이어들의 수의 차이를 설명하기 위한 학습 난이도 측면에서 색상 속성들보다 상대적으로 더 복잡하다.

[0036] 훈련 및/또는 추론 시간에서의 사용 동안, 실시예에서, 모델은 한 번에 N개의 이미지의 그룹을 프로세싱하기 위해 사용된다. 다음으로, 출력은 (N, K)의 차원을 갖고, 여기서 N은 입력 이미지들의 수이고, K는 속성들의 수이다. 출력 매트릭스 내의 각각의 요소는 정수이고, $O_{i,j}$ 는 i번째 이미지 및 j번째 속성에 대해, 시스템이 이미지가 어느 클래스에 속하는 것으로 예측하는지를 나타낸다. 실시예에서, N은 학습률에 대한 훈련 하이퍼-파라미터이다. N개의 이미지 각각은 동일한 또는 상이한 얼굴들을 포함한다. 모델은 신원에 대한 컨텍스트를 갖지 않으며, 동일한 얼굴의 복수의 이미지에 대해 모순되는 예측 결과들을 해결(resolve)하지도 않는다.

[0037] 개별 속성들에 대한 각각의 분류기들 내에서, 최종 출력 레이어 이전의 레이어는 가능한 속성 값들의 세트에 걸쳐 (점수 값과 같은) 예측들의 세트를 생성한다. 실시예에서, 각각의 예측은 속성 값의 "진실(truth)"에 관한 상대적 가능성이다. 3개의 속성 값의 세트에 대해, 예측들의 세트는 20, 20, 60일 수 있지만, 종종 예측들은 상이한 스케일로 정규화된다. 최종 예측은 예측 값들의 세트의 최대값(예를 들어, 60에 연관된 속성)으로부터 결정된다.

[0038] 언급된 바와 같이, 지도 학습에 사용되는 주석이 달린 데이터세트는 독립적으로 행동하는 6명의 인간 주석 작성자에 의한 주석들을 갖는다. 언급된 바와 같이, 특정 속성에 대한 주석들은 각각의 주석 작성자마다 동일하지 않다. 주석 값들(소프트 레이블들) 중에서 선택하기 위해 "진실" 해결 규칙("truth" resolving rule)이 사용된다. 하나의 규칙은 예측들의 세트로부터 가장 일반적으로 예측되는 속성을 선택하는, 가장 일반적인 예측 규칙(most common prediction rule)이다. 하나의 규칙은 상위 2개의 가장 일반적인 예측 규칙(top two most common prediction rule)이다.

[0039] 색상 모델들에 대해, 실시예에서, "참"은 "가장 일반적인 투표(most common vote)"로 취급된다. 이는 눈 색깔에 대한 이하의 주석들: "블루", "블루", "블루", "그레이", "그레이", "브라운"을 가질 때, 모델이 "블루"를 예측한다면, 가장 일반적인 투표와 일치함을 의미한다. 모델이 "블루" 또는 "그레이"를 예측한다면, 상위 2개의 가장 일반적인 투표와 일치하는 것이다. 주석들이 "블루", "블루", "그레이", "그레이", "브라운", "블랙"인 경우, "블루" 또는 "그레이"가 가장 일반적인 투표(또는 상위 2개의 가장 일반적인 투표)일 것이다.

[0040] 표 2는 훈련된 대로 색상 속성들을 예측하기 위한 모델에 대한 출력들 및 평가들을 보여준다. "*" 주석이 달린 속성 출력들은 모델에 의해 거의 예측되지 않는 출력들을 나타낸다. 실시예에서, 이러한 예측들을 개선하기 위해, 추가적인 훈련 데이터가 이용된다. 실시예에서, 또 다른 접근 방식은 이러한 세분화된 예측들(granular predictions)을 함께 또는 더 일반적인(및 관련된) 예측들과 결합 또는 번들링하여, 예측들의 세분성(granularity) 또는 미세함(fineness)을 감소시키는 것을 포함한다. 예를 들어, 실시예에서, "딸기색", "빨간색" 및 "적갈색"은 함께 번들링된다.

표 2

[0041]

	가장 일반적인 투표와 일치하는 백분율	상위 2개의 가장 일반적인 투표와 일치하는 백분율	출력들
피부톤	0.84180	0.95508	아주 밝음, 밝음, 중간, 약간 어두움*, 어두움, 아주 어두움*
언더톤	0.64355	0.89160	중성, 적색, 황색
머리카락 색깔	0.70752	0.87451	구리색*, 짙은 갈색, 딸기색*, 검은색, 갈색, 미정, 금발, 밝은 금발, 적갈색*, 흰색, 빨간색*, 기타*, 회색, 대머리
눈 색깔	0.69531	0.88281	블랙, 브라운, 그린, 그레이, 헤이즐*, 블루

[0042]

모양 모델에 대해, 가장 일반적인 투표는 또한 그라운드 트루스로 취급되었다. 두 가지 평가 표준(예측 해결 규칙들)이 사용되었다: "대다수 투표에만 일치하는 백분율(percentage matching majority vote only)" 및 "가장 일반적인 투표에 일치하는 백분율(percentage matching most common vote)".

[0043]

"가장 일반적인 투표에 일치하는 백분율"은 모든 최대 득표 클래스가 "그라운드 트루스"로 취급될 때, 테스트 세트 내의 모든 샘플에 대한 정확도를 계산할 것이다.

[0044]

"대다수 투표에만 일치하는 백분율"은 대다수 득표 클래스가 단 하나인 샘플들에 대한 정확도만을 계산할 것이다.

[0045]

표 3은 훈련된 대로 모양 속성들을 예측하기 위한 모델에 대한 출력들 및 평가들을 보여준다.

표 3

[0046]

	대다수 투표에만 일치하는 백분율(즉, 동일 투표 포함하지 않음)	가장 일반적인 투표에 일치하는 백분율	출력들
눈썹 모양	0.6717	0.7124	직선형, 곡선형, 아치형(+ 부드러운 아치형)
눈 모양	0.9333	0.9406	둥근형, 아몬드형
눈 주름	0.7543	0.8021	외꺼풀, 중성/주름있음, 처짐
얼굴 길이 (얼굴형 1)	0.8459	0.8668	더 짧, 더 큼
턱 모양 (얼굴형 2)	0.6604	0.6953	뾰족형, 둥근형, 사각형
입술 크기	0.8346	0.8377	얇음, 중간, 두꺼움

[0047]

도 4는 실시예에서 얼굴 속성 분류에 응답하는 추천 애플리케이션을 제공하는 컴퓨팅 시스템(400)을 도시한다. 애플리케이션은 전자상거래 애플리케이션이다. 또한, 예컨대 딥 러닝 네트워크 모델을 사용하는 것에 의한 이미지 프로세싱 기법들을 통해, 애플리케이션은 상황에 맞는 추천들을 제시하는 얼굴 효과 시뮬레이션을 제공한다. 애플리케이션은 사용자에게 가상 및/또는 증강 현실 경험을 제공한다.

[0048]

시스템(400)에는 사용자 컴퓨팅 디바이스(402)가 도시되어 있다. 사용자 컴퓨팅 디바이스(402)는 스마트폰 형태로 도시되어 있다. 태블릿, 개인용 컴퓨터 등과 같은 다른 폼 팩터들이 유용하다. 사용자 컴퓨팅 디바이스(402)는 애플리케이션(404)과 같은 전자상거래 애플리케이션에 유용한, Apple Inc., Samsung Electronics Co., Ltd. 등으로부터 입수가 가능한 것과 같은 소비자 지향 디바이스이다. 실시예에서, 디바이스(402)는 판매원 등에 의해 조작되는 제품 계산대 키오스크 디바이스이다.

[0049]

실시예에서, 사용자 컴퓨팅 디바이스(402)는 프로세서(들)(예를 들어, CPU, GPU 등), 저장 디바이스(들), 하나 이상의 통신 서브시스템 또는 디바이스, 디스플레이(들), 입력 디바이스(들), 출력 디바이스(들) 등 중 적어도 일부를 포함하여, 통신을 위해 결합된 다양한 컴포넌트들을 갖는 컴퓨팅 디바이스를 포함한다. 실시예에서, 디스플레이는 터치 또는 제스처 가능형이고, 입력 디바이스들은 마우스와 같은 포인팅 디바이스, 마이크로폰, 카메라, 키보드, 버튼(들) 등을 포함하고, 통신 디바이스들은 유선 또는 무선 통신 중 임의의 것을 제공하며, 단

거리 및/또는 장거리 가능형이다. 실시예에서, 통신 디바이스들은 예를 들어 위성 기반 위치 서비스들을 제공하기 위해 위치 디바이스에 결합된다. 실시예에서, 출력 디바이스들은 스피커, 조명, 진동/햅틱 디바이스들 등을 포함하고, 다양한 컴포넌트들은 하나 이상의 통신 버스 또는 다른 구조를 통해 결합된다.

[0050] 실시예에서, 저장 디바이스들은 명령어들 및 데이터를 저장하며, 예를 들어, 이 명령어들은 실행될 때 컴퓨팅 디바이스의 동작을 구성한다. 명령어들은 운영 체제, 애플리케이션들 등을 정의한다. 데이터는 얼굴의 하나 이상의 소스 이미지를 포함한다. 애플리케이션(404)에 더하여, 실시예에서, 이메일들, 텍스트 메시지들, 인스턴트 메시지들, 단문 메시지 서비스(SMS) 메시지들 등과 같은 메시지(데이터)를 통신하기 위한 하나 이상의 애플리케이션이 있다. 실시예에서, 통신은 이미지들 또는 동영상들 등과 같은 첨부물들을 포함한다.

[0051] 실시예에서, 애플리케이션(404)은 사용자 컴퓨팅 디바이스(402)의 일부이거나 그에 결합된 디스플레이 디바이스(예를 들어, 408)에 의한 디스플레이를 위한 하나 이상의 화면(예를 들어, 406)을 포함하는 그래픽 사용자 인터페이스를 제공한다. 실시예에서, 사용자 컴퓨팅 디바이스(402) 및 애플리케이션(404)은 제스처 및/또는 음성 또는 기타 가능형이다. 실시예에서, 애플리케이션(404)은 사용자 디바이스 및 그것의 운영 체제를 위해 특별히 구성되고 다운로드되어 거기에 저장되는 네이티브 애플리케이션이거나, 브라우저 기반 애플리케이션의 형태이고/거나 웹 서비스로서 제공되는 것과 같이, 더 디바이스 중립적이다.

[0052] 실시예에서, 애플리케이션(404)은 예를 들어 인터넷 프로토콜(IP)을 사용하여 네트워크(410)를 통해 (디바이스(402)를 사용하여) 전자상거래 서버(412) 및 이미지 프로세싱 서버(414)와 같은 하나 이상의 서버와 통신한다. 시스템(400)은 다음과 같이 단순화된다: 결제 및 다른 서비스들은 보이지 않고, 단 하나의 사용자 컴퓨팅 디바이스만 보여지는 등.

[0053] 실시예에서, 서버들(412 및 414)은 공지된 바와 같이 적절한 리소스들을 갖는 개별 컴퓨팅 디바이스들에 의해 제공된다(그러나, 하나의 그러한 디바이스가 사용될 수 있음). 적절한 컴퓨팅 디바이스는 프로세서(들)(예를 들어, CPU, GPU 등), 저장 디바이스(들), 하나 이상의 통신 서브시스템 또는 디바이스, 디스플레이(들), 입력 디바이스(들), 출력 디바이스(들) 등 중 적어도 일부를 포함하여, 통신을 위해 결합된 다양한 컴포넌트들을 갖는다. 실시예에서, 디스플레이는 터치 또는 제스처 가능형이고; 입력 디바이스들은 마우스와 같은 포인팅 디바이스, 마이크로폰, 카메라, 키보드, 버튼(들) 등을 포함하고; 통신 디바이스들은 유선 또는 무선 통신 중 임의의 것을 제공하며, 단거리 및/또는 장거리 가능형이다. 실시예에서, 통신 디바이스들은 예를 들어 위성 기반 위치 서비스들을 제공하기 위해 위치 디바이스에 결합되고, 출력 디바이스들은 스피커, 조명, 진동/햅틱 디바이스들 등 중 임의의 것을 포함한다. 다양한 컴포넌트들은 하나 이상의 통신 버스 또는 다른 구조를 통해 결합된다. 실시예에서, 저장 디바이스들은 명령어들 및 데이터를 저장하며, 예를 들어, 이 명령어들은 실행될 때 컴퓨팅 디바이스의 동작을 구성한다.

[0054] 본 실시예에서, 애플리케이션(404)은 makeup 제품들을 위한 전자상거래 애플리케이션이다. 애플리케이션(404)은 얼굴 속성 분석을 얻기 위해 소스 이미지를 제공하고, 소스 이미지로부터 결정된 얼굴 속성들에 응답하여 제품 추천들을 수신하도록 구성된다. 서버(414)는 (서버(412)와의 통신 없이 또는 서버(412)를 통해 디바이스(402)로부터) 소스 이미지를 수신하고 얼굴 피쳐들을 결정하는 방법을 수행하도록 구성된다.

[0055] 서버(414)는 얼굴 속성들 각각에 대한 예측을 출력하기 위해 딥 러닝 및 지도 회귀를 수행하는 네트워크 모델을 사용하여 얼굴의 소스 이미지를 프로세싱한다.

[0056] 실시예에서, 네트워크 모델은 얼굴 속성들을 예측하기 위해 개별 분류기들에 의한 분류를 위한 공유 피쳐들(shared features)의 피쳐 벡터를 생성하도록 딥 러닝을 수행하는 잔차 블록들을 포함하는 컨볼루션 신경망(CNN) 모델을 포함한다. 피쳐 벡터의 복제본들은 얼굴 속성들 중의 개별 얼굴 속성에 대해 개별 분류기에 의해 프로세싱되도록 만들어진다. 실시예에서, 네트워크 모델은 복수의 개별 분류기를 포함하고, 개별 분류기들 각각은 하나 이상의 완전 연결된 선형 레이어를 포함하고, 개별 분류기들 각각은 얼굴 속성들 중의 개별 얼굴 속성의 예측을 출력으로서 제공한다. 실시예에서, 복수의 개별 분류기는 얼굴 속성들을 제공하기 위해 병렬로 수행된다.

[0057] 실시예에서, 얼굴 속성들은 얼굴 속성들(예를 들어, 색상 또는 모양 등)의 내재적 관계(intrinsic relationship) 및 이질성(heterogeneity)에 따라 복수(예를 들어, 2개)의 개별 속성 그룹에 연관된다. 실시예에서, 서버(414)는, 방법이 소스 이미지를 프로세싱하기 위해 속성 그룹들 중의 개별 속성 그룹에 대해 구성된 개별 네트워크 모델을 사용하도록 구성된다.

[0058] 실시예에서, 개별 속성 그룹들 중 하나는 색상 기반 얼굴 속성들에 대한 색상 속성 그룹이고, 소스 이미지는 색

상 기반 네트워크 모델에 의해 프로세싱된다. 실시예에서, 개별 속성 그룹들 중 하나는 모양 기반 얼굴 속성들에 대한 모양 속성 그룹이고, 소스 이미지는 모양 기반 네트워크 모델에 의해 프로세싱된다.

- [0059] 실시예에서, 네트워크 모델은 예측 레이어들이 트리밍된, 적응된 사전 훈련된 ResNet 기반 이미지 프로세싱 네트워크 모델을 포함한다.
- [0060] 서버(414)가 본 예에서는 제품 추천을 제공하기 위해 사용하기 위한 얼굴 속성들을 제공하지만, 다른 실시예들에서, 서버는 이미지 검색, 얼굴 인식, 개인 재식별, 및 제품 및/또는 서비스 추천 등 중 임의의 것을 수행하는 애플리케이션에 의해 사용되도록 얼굴 속성들을 제공한다.
- [0061] 실시예에서, 서버(414)는 공유 피쳐들의 피쳐 벡터를 생성하기 위해 잔차 블록들을 갖는 컨볼루션 신경망(CNN) 기반 백본 네트워크 모델을 사용하여 얼굴의 소스 이미지를 프로세싱하고; 결정될 K개의 개별 얼굴 속성 각각에 대해 하나씩, 복수(K개)의 분류기를 사용하여 피쳐 벡터를 프로세싱하도록 구성되고, 각각의 개별 분류기는 K개의 개별 얼굴 속성 중 하나의 얼굴 속성의 예측을 출력하기 위해, 하나 이상의 완전 연결된 선형 레이어 및 예측 블록을 포함한다. 실시예에서, 제1 네트워크 모델(예를 들어, 서브 모델)은 복수의 분류기 중에서 색상 기반 얼굴 속성들을 예측하도록 구성된 개별 분류기들에 의한 프로세싱을 위한 색상 기반 피쳐 벡터를 생성하기 위해 색상 기반 얼굴 속성들에 대해 소스 이미지를 프로세싱하고, 제2 네트워크 모델은 복수의 분류기 중에서 모양 기반 얼굴 속성들을 예측하도록 구성된 개별 분류기들에 의한 프로세싱을 위한 모양 기반 피쳐 벡터를 생성하기 위해 모양 기반 얼굴 속성들에 대해 소스 이미지를 프로세싱한다.
- [0062] 사용자 컴퓨팅 디바이스가 (예를 들어, 카메라로부터 또는 (예를 들어, 저장 디바이스로부터) 다른 업로드로부터) 소스 이미지를 제공하고 서버(414)가 얼굴 속성 검출 서비스를 수행하는 클라이언트/서버 모델로서 도시되어 있지만, 실시예에서, 사용자 컴퓨팅 디바이스는 충분히 강건하다면 분류를 수행하기 위해 딥 러닝 모델로 구성된다.
- [0063] 실시예에서, 서버(412)는 전자상거래 서비스를 위한 제품 추천들을 제공한다. 서버(412)는 제품 추천들 중 적어도 일부를 구매하기 위해 전자상거래 쇼핑 서비스를 제공한다.
- [0064] 실시예에서, 서버(412)는 예컨대, 얼굴의 소스 이미지로부터 결정된 복수의 얼굴 속성을 수신하는 단계 - 소스 이미지는 복수의 얼굴 속성을 생성하기 위해 얼굴 속성 분류 네트워크 모델을 사용하여 프로세싱됨 -; 제품들에 적합한 얼굴 속성들에 연관하여 제품들을 저장하는 데이터 저장소로부터 적어도 하나의 제품을 선택하기 위해 얼굴 속성들 중 적어도 일부를 사용하는 단계; 및 제품들을 구매하기 위해 전자상거래 인터페이스에 프레젠테이션하기 위한 추천으로서 적어도 하나의 제품을 제공하는 단계에 의해, 전자상거래 서비스를 위해 제품들을 추천하도록 구성된다.
- [0065] 실시예에서, 얼굴 속성은 서버(414), 및/또는 서버(414)에 대해 설명된 방법(들)을 사용하여 결정된다.
- [0066] 실시예에서, 제품들은 메이크업 제품들을 포함한다. 메이크업 제품들은 미리 결정된 메이크업 룩들(make-up looks)을 정의하기 위해 데이터 저장소(예를 들어, 서버(412)에 결합된 데이터베이스(416))에서 다양하게 연관된다. 룩들은 특정 효과들을 갖는 메이크업으로부터 정의되고 및/또는 특정 기법들을 사용하여 적용된다. 각각의 룩은 상이한 타입들의 메이크업으로부터(예를 들어, 그것들을 포함하도록) 정의된다. 타입들은 얼굴의 피부 부위들(영역들)(볼, 코, 턱, 이마, 턱선 등)에, 또는 눈, 입술, 및 눈썹 및 속눈썹과 같은 체모를 포함하는 브로우와 같은 영역들에 관련된다. 얼굴 영역(피부), 눈 영역, 브로우 영역, 입술 영역 등은 설명된 바와 같이 모양, 색상을 포함하는 얼굴 속성들에 연관된다. 개별 영역들에 대한 개별 메이크업 제품들은 예를 들어 속성에 대한 적합성 척도를 사용하는 것에 의한 것과 같이, 개별 속성들에 연관된다. 척도는 "~에 좋음" 또는 "~에 좋지 않음"과 같이 양분법적일 수 있고, 또는 적합성 척도가 더 세분화될 수 있다(예를 들어, 1 내지 10의 스케일로, 또는 다른 스케일로 순위 지정). 실시예에서, 룩은 얼굴의 다수의 영역(예를 들어, 모든 영역)에 대한 개별 제품들을 포함한다. 룩을 위한 제품들은 얼굴 속성들에 연관된 척도들을 갖는다(예를 들어, 모양 및 색상의 입술 속성들에 대한 립 제품들, 브로우 속성들에 대한 브로우 제품들 등). 추천 기능은 사용자 이미지의 이미지 분석으로부터 결정된 얼굴 속성들을 개별 룩들로 그룹화된 제품들에 일치시키기 위해 규칙들 또는 다른 방식들을 사용할 수 있다. 실시예에서, 얼굴 속성들에 대한 일치는 잠재적 룩을 식별한다. 초기 일치에서 하나보다 많은 룩이 식별될 수 있다. 식별된 룩들은 예를 들어 개별 룩에 대한 얼굴 속성 일치들의 카운트, 및/또는 일치의 스케일을 평가하는 얼굴 속성 일치들의 품질에 기초하여 순위가 지정될 수 있다. 순위는 룩들의 프레젠테이션을 정리(order)하고/거나 점수를 제공하기 위해 사용될 수 있다. 실시예에서, 룩이 선택될 수 있고, (예를 들어, 얼굴에 결합될 때) 그 룩을 달성하기 위해 사용가능한 일치 제품들을 결정하기 위해 얼굴 속성 정

보가 사용될 수 있다. 순위는 특을 달성하기 위해 사용될 개별 제품들을 정리하거나 점수를 매기기 위해 사용될 수 있다.

- [0067] 실시예에서, 서버(414)는 미리 결정된 메이크업 룩들 중 하나의 식별을 수신한다. 얼굴 속성들 중 적어도 일부를 사용할 때, 서버(412)는 적어도 하나의 제품을 선택할 때 미리 결정된 메이크업 룩들 중 하나에 응답한다.
- [0068] 실시예에서, 메이크업 제품들 각각은 복수의 메이크업 타입 중 하나에 연관되기 때문에, 서버(412)는 추천을 정의하기 위해 메이크업 타입들 각각에 대해 얼굴 속성들에 응답하여 적어도 하나의 제품을 선택한다. 언급된 바와 같이, 메이크업 타입들은 페이스 제품 타입, 아이 제품 타입, 브로우 제품 타입, 및 립 제품 타입을 포함한다.
- [0069] 실시예에서, 서버(412)는 추천 제품들을 사용하기 위한 기법들을 추천한다.
- [0070] 실시예에서, 서버(412)는 애플리케이션(404)을 위해 (사용자) 컴퓨팅 디바이스(402)에 그것의 출력을 제공한다. 출력은 예컨대 전자상거래 서비스의 그래픽 사용자 인터페이스를 통해, 디스플레이(408)를 통해 프레젠테이션된다.
- [0071] 룩 또는 개별 메이크업 제품에 대한 사용자의 인식을 향상시키기 위해, 실시예에서, 소스 이미지는 예를 들어 얼굴 효과를 프레젠테이션하기 위한 이미지 프로세싱을 통해 수정된다. 소스 이미지는 서버(412)에 의해, 또는 서버(414)에 의해, 또는 본 명세서의 실시예들에서 설명된 바와 같은 사용자 컴퓨팅 디바이스(402)에 의해 수정된다. 딥 러닝 네트워크 모델이 소스 이미지를 수정하거나, 다른 이미지 프로세싱 기법들이 본 명세서의 실시예들에서 설명된 바와 같이 사용된다. 소스 피쳐들을 유지하면서 하나의 도메인으로부터 다른 도메인으로 변환하기 위해 소스 이미지를 수정하기 위한 하나의 그러한 딥 러닝 네트워크 모델은 생성적 대립망(generative adversarial network)(GAN) 기반 모델이다.
- [0072] 실시예에서, 애플리케이션(404)은 사용자 컴퓨팅 디바이스(412)가 현실을 시뮬레이션하기 위한 방법을 수행할 수 있게 한다. 방법은 얼굴의 소스 이미지로부터 복수의 얼굴 속성을 결정하고, 얼굴 속성 분류 네트워크 모델을 사용하여 소스 이미지를 프로세싱하는 단계; 얼굴 속성들 중 적어도 하나에 응답하여 얼굴에 적용하기 위한 적어도 하나의 얼굴 효과를 결정하는 단계; 및 디스플레이를 위해 얼굴에 적어도 하나의 얼굴 효과를 적용하기 위해 소스 이미지를 프로세싱하는 단계를 포함한다. 실시예에서, 애플리케이션(404)은 서버들(412 및 414) 중 하나 또는 둘 다와 통신하여, 방법 단계들이 그를 위해 수행되게 한다. 실시예에서, 사용자 컴퓨팅 디바이스(402)는 소스 이미지(예를 들어, 420)를 생성하기 위해 카메라(418)를 사용한다.
- [0073] 실시예에서, 적어도 하나의 얼굴 효과는 얼굴에 적용될 적어도 하나의 메이크업 제품 및/또는 기법의 시뮬레이션이다. 일 실시예에서, 다른 얼굴 효과들은 보형물의 적용, 성형 수술, 또는 아이웨어 등을 포함한다.
- [0074] 다양한 실시예들에서, 적어도 하나의 얼굴 효과를 적용하기 위해 소스 이미지를 프로세싱하는 것은 얼굴 효과를 적용하도록 구성된 딥 러닝 네트워크(예를 들어, GAN 기반)를 사용하는 것을 포함하거나, 다른 딥 러닝 모델들 또는 다른 이미지 프로세싱 기법들이 사용된다. 제한이 아닌 예로서, (클라이언트 측) 컴퓨팅 디바이스에 구성된 증강 현실(AR) 가상 사용 방법(augmented reality (AR) virtual try on method)은 소스 이미지 상에 추천 제품 또는 서비스에 연관된 효과를 시뮬레이션하여 가상 사용 사용자 경험을 제공하기 위해 사용된다. 요약하면, (예를 들어, 소스 이미지로서의) 카메라 프레임 또는 사진은 물론, 메이크업 제품들을 표현하는 렌더링 값들(예를 들어, 적색, 녹색, 청색(RGB) 색상 및 색상 불투명도)이 방법에 대한 입력으로서 수신된다. 소스 이미지는 소스 이미지에서 사용자 얼굴 주위의 주요 랜드마크들을 결정하기 위해 얼굴 추적을 사용하여 프로세싱된다. 이러한 랜드마크들을 사용하여, 입술 관련 제품에 대한 사용자의 입술과 같이, 메이크업 제품에 대한 얼굴의 부위들이 마스킹된다. 다음으로, 렌더링 값들은 마스킹된 부위에서 사용자 얼굴에 메이크업을 그리기 위해 사용된다. 실시예에서, 실제 제품들이 어떻게 보이는지를 현실적으로 렌더링하기 위해, 소스 이미지로부터의 조명 샘플링, 블렌딩 등과 같은 다른 단계들이 포함된다. 마지막으로, 결과 - 메이크업 또는 얼굴 효과를 갖는 변경된 소스 이미지 - 는 디스플레이될 수 있는 출력으로서 다시 전달된다.
- [0075] 실시예에서, 도 5, 도 6a, 도 6b, 도 7a, 도 7b, 도 8, 도 9a, 도 9b, 도 10a, 도 10b, 도 10c, 도 10d, 도 11a, 도 11b 및 도 11c는 애플리케이션(404)의 그래픽 사용자 인터페이스에 대한 화면들 또는 그 일부를 도시한다.
- [0076] 실시예에서, 도 5는 사용자 컴퓨팅 디바이스(402)의 디스플레이 화면(408) 상에 콘텐츠가 모두 들어맞지 않는 (모두 디스플레이가능하지는 않은) 애플리케이션 열기 화면(500)을 도시한다. 부분들은 애플리케이션 개요를 제공하는 부분(502)과 같이 공지된 대로 스크롤가능하다. 애플리케이션을 진행시키기 위해(예를 들어, 시작)

컨트롤(504)이 제공된다.

- [0077] 실시예에서, 도 6a는 소스 이미지(예를 들어, 420)를 취득하기 위한 초기 화면(600)(예를 들어, "사진 촬영 지시 화면")을 도시한다. 이미지는 카메라(418)를 통한 "셀카"이다. "촬영" 컨트롤은 사진 촬영 인터페이스를 개시하도록 진행된다(도 7 참조). 업로드 컨트롤(604)은 디바이스(402)의 로컬 저장 디바이스로부터, 또는 화면(610)(예를 들어, "업로드/이미지 취득을 위한 옵션들") 및 도 6b를 통해 다르게 이미지를 취득하는 것과 같은 업로드 인터페이스로 진행된다.
- [0078] 실시예에서, 도 7a 및 도 7b는 얼굴(704)의 이미지를 캡처(예를 들어, 카메라로부터 수신)하기 위한 사진 촬영 인터페이스(예를 들어, 각각 화면들(700 및 710))를 도시한다. 이미지 캡처 컨트롤(702)은 소스 이미지(420)를 취득하기 위해 카메라(418)를 호출한다. 후방(반대쪽) 카메라 모드로 전환하는 것들과 같은 다른 컨트롤들은 보여지지 않는다. 실시예에서, 사용자가 얼굴을 적절하게 위치시키고, 메이크업을 최소화하고, 머리카락 또는 다른 물체들이 얼굴을 가리는 것을 최소화하고, 적절한 조명을 사용하는 등을 돕기 위한 지시들이 제시된다. 지시들의 예는 "얼굴을 위치시키세요"(706)이고, 이는 이미지가 촬영되기 전에 얼굴의 프레젠테이션을 오버레이 할 수 있다. 이미지가 캡처되고 나면, 얼굴 위치 및 조명을 확인하고 재촬영을 요청하는 초기 프로세싱이 수행된다(도시되지 않음). 실시예에서, 소스 이미지(420)는 설명된 바와 같이 얼굴 속성들을 결정하기 위한 프로세싱을 위해 예컨대 서버(414)에 전송된다.
- [0079] 실시예에서, 도 8은 소스 이미지(420)로부터 결정된 얼굴 속성들을 프레젠테이션하기 위한 얼굴 분석을 위한 인터페이스(예를 들어, 화면(800))를 도시한다. 현재 인터페이스에서, 소스 이미지는 얼굴 속성들에 관한 정보(802)를 프레젠테이션(예를 들어, 오버레이)하기 위해 수정된 대로 프레젠테이션된다(420A). 얼굴 속성 정보(803)는 속성들에 관련된 얼굴(420)의 영역들에 연관하여 상황에 맞게 프레젠테이션된다. 실시예에서, 얼굴 속성 정보는 소스 이미지(420)에 대해 결정된 바와 같은 표 1로부터의 속성 및 연관된 값을 식별한다. 여기의 화면들에서의 오버레이 스타일들 및 컨트롤 위치는 예들이다.
- [0080] 실시예에서, 영역들은 예를 들어 영역을 찾기 위해 피쳐 검출을 수행하는 이미지 프로세싱에 의한 것과 같이 소스 이미지(420) 상에 위치된다. 위치들은 예를 들어 디바이스(402) 또는 서버(414)에 의해 결정된다. 영역들은 얼굴 전체, 예를 들어 눈썹, 눈 및 입술에 관련된다. 이러한 영역들은 하나 이상의 특정 얼굴 속성에 관련된다. 실시예에서, 각각의 영역은 개별 제품 타입에 관련된다.
- [0081] 실시예에서, 개별 정보(예를 들어, 802A)는 연관된 영역 인터페이스로 진행하기 위한 컨트롤(예를 들어, 탭 컨트롤(802B))에 연관된다. 각각의 연관된 영역 인터페이스는 도시된 바와 같이 영역에 대한 얼굴 속성들 및 관련 제품들 등에 관한 추가 정보를 제공하도록 구성된다. 도 9a(컨트롤(802B)에 연관됨) 및 도 9b(컨트롤(802C)에 연관됨)를 참조한다. 실시예에서, 각각의 영역은 개별 영역 인터페이스를 갖는다.
- [0082] 또한, 얼굴 속성들 및 관련 제품들에 관한 추가 정보를 제공하는 연관된 영역 인터페이스로 진행하기 위해 개별(탭) 컨트롤(804)이 제공된다. 실시예에서, 컨트롤들(804)은 예를 들어 스와이프 제스처(예를 들어, 왼쪽/오른쪽) 및/또는 진행 컨트롤(예를 들어, 804A)의 터치에 의해 진행(호출)된다.
- [0083] 실시예에서, 스크롤은 인터페이스(800)의 부분(800A)에 추가 정보(예를 들어, "내 결과 개요" 부분) 및 컨트롤들을 가져온다. 실시예에서, 저장 컨트롤(806)(예를 들어, 결과 이미지 저장)은 주석이 달린 이미지(420) 및/또는 얼굴 속성 분석의 결과들을 디바이스(402)의 저장 디바이스에 저장하기 위해 제공된다. 컨트롤(808)(예를 들어, "내 메이크업 룩 계속하기")은 추천 인터페이스(예를 들어 도 10a)로 진행된다. 컨트롤들과 같은 다양한 인터페이스 부분들 및 요소들은 인터페이스 탐색 및 사용자 인터페이스 이해를 돕기 위해 텍스트 레이블들을 가질 수 있음이 명백하다.
- [0084] 실시예에서, 도 9a는 얼굴 전체의 얼굴 속성들에 관련된 영역(902)을 시각적으로 강조하고, 얼굴형, 피부톤, 언더톤과 같은 추가 정보(예를 들어 "모양, 피부톤 & 언더톤 정보 메이크업 팁")를 (예를 들어, 부분(904)에) 제공하기 위한 영역 인터페이스(900)를 보여준다. 부분(904) 내의 추가 정보는 영역(902)에 대한 연관 제품들에 관련되고, 그래픽 및/또는 텍스트, 색상 견본 이미지들 등을 포함한다. 실시예에서, 제품 추천 인터페이스(예를 들어, 도 10a)로 진행하기 위해 컨트롤(906)이 제공된다(예를 들어, "내 추천 보기"). 눈썹을 위한 것과 같은 다른 영역 인터페이스(예를 들어, 도 9b)로 진행하기 위해 컨트롤(908)이 제공된다.
- [0085] 실시예에서, 도 9b는 눈썹의 얼굴 속성들에 관련된 영역(912)을 시각적으로 강조하고, 눈썹 모양, 눈썹 색상과 같은 추가 정보(예를 들어, "눈썹 모양 및 색상 정보 메이크업 팁")를 부분(914)에 제공하기 위한 영역 인터페이스(910)를 도시한다. 부분(914) 내의 추가 정보는 영역(912)에 대한 연관 제품들에 관련되고, 그래픽 및/또

는 텍스트, 색상 견본 이미지들 등을 포함한다. 제품 추천 인터페이스(예를 들어, 도 10a)로 진행하기 위해 컨트롤(906)이 제공된다(예를 들어, "내 추천 보기"). 눈을 위한 것과 같은 다른 영역 인터페이스(눈에 대한 영역 인터페이스는 도시되지 않음)로 진행하기 위해 컨트롤(918)이 제공된다. 특정 영역 인터페이스들 또는 "내 결과" 인터페이스(예를 들어, 도 8)로 진행하기 위해 컨트롤들(920)이 제공된다. 특정 영역 인터페이스들로의 진행은 예를 들어 컨트롤들(920) 중 하나를 스와이프하거나 탭핑하여 탐색될 수 있다.

[0086] 영역 인터페이스들(예를 들어, 900 및 910)에 대해, 관심 영역들(902, 912)은 (예를 들어, 오버레이를 통해) 소스 이미지(420) 상에 위치된다. 예를 들어 이미지 프로세싱은 영역들을 찾기 위해 피처 검출을 수행한다. 실시예에서, 개별 마스크는 각각의 수정된 이미지(420B, 420C)를 정의하기 위해 소스 이미지(420)를 오버레이하도록 정의된다. 실시예에서, 마스크는 예를 들어 관심 영역(902, 912) 외부의 영역을 어둡게 하여 개별 영역에 초점을 맞춘다.

[0087] 도시되지는 않았지만, 실시예에서, 입술 및 눈을 위한 영역 인터페이스들이 제공된다. 입술 영역 인터페이스는 입술 크기 정보를 프레젠테이션하고, 입술 모양 정보를 표시 및/또는 설명한다. 눈 영역 인터페이스는 눈 모양, 눈꺼풀 및 눈 색상 정보를 프레젠테이션한다. 도시된 모든 영역 인터페이스에서와 같이, (예를 들어, 영역에 연관된 적용가능한 얼굴 속성들에 관련된) 제품 팁들이 또한 제공된다.

[0088] 실시예에서, 도 10a 내지 도 10d는 추천 인터페이스의 초기 화면들(1000, 1010, 1020, 및 1030)을 도시한 것이다. 화면들(1000, 1010, 1020 및 1030)은 다양한 제품 정보를 프레젠테이션하고, 진행 컨트롤(예를 들어 "다음"(1002))을 포함한다. 화면(1000)은 각각 영역 또는 항목 정보 및 전체 룩 정보로 진행하기 위한 컨트롤들(1004A 및 1004B)을 갖는 "퍼스널 룩 소개 화면"을 제공한다. 도 10b의 룩 선택기 인터페이스(1010)는 "룩 1", "룩 2", "룩 3" 및 "룩 4"와 같은 미리 결정된 "룩"(예를 들어, 전체적인 얼굴 효과를 달성하기 위한 개별 얼굴 효과들을 갖는 미리 결정되거나 선택된 메이크업 제품 그룹들)에 대한 선택들(예를 들어, 선택 컨트롤들(1012A, 1012B, 1012C 및 1012D)에 연관됨)을 프레젠테이션한다. 도시에서는 일반적으로 레이블 지정되어 있지만, 각각의 룩은 설명적이거나 공상적인 레이블(예를 들어, "드라마틱", "우아", "사무적", "지각" 등)을 가질 수 있다. 메이크업 타입들은 각각의 (얼굴) 영역에 연관되며, 인터페이스를 조작하고 추가 정보 프레젠테이션을 요청하기 위해 선택가능(예를 들어, 호출가능)하다. 정보는 예컨대 데이터베이스(416)에 저장된다. 실시예에서, 룩은 얼굴 속성 정보와 키잉되고(예를 들어, 그에 연관하여 저장되고), 예컨대 제품 또는 그것을 적용하기 위한 기법은 하나 이상의 특정 얼굴 속성에 연관(예를 들어, 추천)된다. 예를 들어, 계란형 얼굴, 아몬드형 눈, 두꺼운 눈썹, 갈색 눈썹, 붉은색 언더톤, 금발 머리 등을 갖는 얼굴에 대한 특정한 룩을 달성하기 위해; 얼굴, 눈, 눈썹, 및 입술에 적용가능한 제품들이 결정된다.

[0089] 실시예에서, 피부 타입 선택기 인터페이스(예를 들어, 화면(1020))는 "타입 1", "타입 2", "타입 3", 및 "타입 4"와 같은 피부 타입 선택을 위한 (예를 들어, 선택 컨트롤들(1022A, 1022B, 1022C 및 1022D)에 연관된) 선택들을 프레젠테이션한다. 도시에서는 일반적으로 레이블 지정되어 있지만, 각각의 타입은 설명적이거나 공상적인 레이블을 가질 수 있다. 실시예에서, 피부 타입은 일반, 건성, 지성, 복합성 등의 속성 값들에 관한 것이다. 실시예에서, 피부 타입은 이미지 프로세싱에 의해 자동으로 결정되지 않는, 사용자에게 의해 제공되는 데이터이다.

[0090] 실시예에서, 제품 추천은 피부 타입에 응답한다. 다음 컨트롤(1002)을 선택하면, 예컨대 해당 정보를 요청하는 서버(412)에의 메시지에 의해 제품 추천을 결정하는 데 사용할 룩 및 피부 타입 정보를 제공하기 위한 애플리케이션(402)이 호출된다. 실시예에서, 서비스에 대한 메시지는 소스 이미지(420)에 대해 결정된 사용자의 얼굴 속성 정보를 제공한다. 실시예에서, 소스 이미지(420)는 예컨대 효과를 적용하기 위해 사용되도록 제공된다. 소스 이미지(420)를 보여주는 화면(1030)(도 10d)은 예를 들어 제품 추천들을 수신하기 위해 대기할 때 프레젠테이션된다(예를 들어, "룩 매치"라는 레이블을 가짐). 화면(1030)은 도 11a의 화면(100)에서 시뮬레이션된 제품 추천들과 비교하기 위한 기준선(예를 들어 메이크업 참조 없음)을 제공한다.

[0091] 실시예에서, 도 11a 및 도 11b는 현실이 시뮬레이션되는 제품 추천 인터페이스(예를 들어, 1100A 및 1100B)로서 두 부분으로 도시된 화면(1100))를 도시한다. 얼굴 속성들과 일치하도록 결정된 제품 추천들을 이용한 선택된 룩이 소스 이미지(420) 상에 시뮬레이션되어 이미지(420D)를 프레젠테이션한다.

[0092] 실시예에서, 화면(1100)은 원하는(선택된) 룩(예를 들어, 1102)을 달성하기 위해 제품들을 매칭하는 데 사용되는 얼굴 속성 정보(예를 들어, 1102)를 프레젠테이션한다. 실시예에서, 1002에서의 정보는 룩이 사용자의 개별 속성들 중 2개(예를 들어, 사용자 상세 1 및 사용자 상세 2)에 대해 이상적임을 나타낸다. 실시예에서, 특정 룩에 대한 일반 정보(예를 들어, 1104)가 제공된다. 실시예에서, 일반 정보는 룩의 이름 및 피쳐들을

포함하고, 룩의 제품들 및 튜토리얼 정보를 포함한다. 시뮬레이션된 이미지(420D)를 저장하거나 공유하기 위해 공유/저장 컨트롤(1106)이 제공된다.

[0093] 실시예에서, 화면(1100)은 개별 부분들(1110, 1112, 1114, 및 1116)에서 얼굴의 영역들(예를 들어, 전체 얼굴, 눈, 눈썹 및 입술) 각각에 대한 추천 정보를 프레젠테이션하기 위해 분할된다. 눈에 대한 부분(1112)은 간결함을 위해 부분적으로만 표현되지만, 다른 부분들(1110, 1114 및 1116)과 유사하게 구성된다는 점에 유의해야 한다. 부분들(1110, 1114 및 1116)은 특정 제품 추천들(예를 들어, 1110A, 1114A 및 1116A에서 제품 이미지, 색상, 이름, 가격을 포함함), 및 1110B, 1114B 및 1116B에서의 특정 제품 튜토리얼을 보여준다. 실시예에서, 각각의 개별 제품은 제품을 쇼핑 카트에 항목으로서 추가하기 위해 개별 구매 컨트롤(예를 들어, 1110C, 1114C 및 1116C에서의 "카트에 추가")에 연관된다. 프레젠테이션할 정보의 양에 응답하여, 예로서, 부분들(1110, 1112, 1114 및 1116) 중 임의의 것은 예컨대 영역에 대한 하나보다 많은 제품, 튜토리얼 및 구매 컨트롤을 프레젠테이션하기 위해 (예를 들어, 왼쪽 또는 오른쪽으로의 슬라이프 제스처를 통해) 스크롤가능할 수 있다. 단계 카운트 및/또는 버튼 표시(예를 들어, 1110D, 1114D 및 1116D)는 화면(1100)에 현재 프레젠테이션되는 이러한 정보 내의 위치를 나타내는 데 유용하다.

[0094] 실시예에서, 컨트롤(1118)은 모든 제품을 쇼핑 카트에 추가한다. 예를 들어, 제품 속성들에 응답하여, 개별 제품들은 (예를 들어, 크기 또는 임의의 다른 선택가능한 속성을 선택하기 위해) 선택가능한 피쳐들을 갖는다. 부분들(1110, 1112, 1114 및 1116) 중 임의의 것에서 개별 제품에 대한 연관된 컨트롤(도시되지 않음)을 선택하면, 인터페이스가 제품 선택기 인터페이스(1120)를 프레젠테이션하게 된다. 예를 들어, 부분(1110)에서 제품 추천(1110A)에 대한 연관된 컨트롤을 선택하면, 해당 제품 추천을 위해 1120과 유사한 적용가능한 제품 선택기 인터페이스가 호출되고, 여기서 제품 상세 선택기(1122)는 추천(1110A)을 대체하고, 상세 선택기(1122)를 사용하여 선택을 확인하기 위한 선택 컨트롤(1124)이 카트에 추가 컨트롤(1110C)을 대체한다.

[0095] 컨트롤들(1126)(예를 들어, "내 룩 선택")은 다양한 룩들 사이에서, 또는 하나의 룩 내에서 예컨대 상이한 룩 피쳐들 또는 컨텍스트들에 대해 이동하도록 동작가능하다. 실시예에서, 컨텍스트는 하루 중 시각, 이벤트, 위치 등에 관련된다. 컨트롤들(1126)의 호출은 연관된 얼굴 효과들의 새로운 제품 추천들 및 새로운 현실 시뮬레이션들을 야기한다. 검색 컨트롤(1128)(예를 들어, "메이크업 제품군 검색")은 추가 제품들을 보여준다(인터페이스는 도시되지 않음).

[0096] 실시예에서, 화면(1110)은 예를 들어 룩 변경/내 선택 변경(예를 들어, 룩 인터페이스 탐색에서 한 단계 뒤로 돌아가기), 처음부터 다시 시작(예를 들어, 화면(700)에서 새로운 이미지 캡처를 시작), 결과로 돌아가기(예를 들어, 화면(800)), 종료(닫는 인터페이스는 보이지 않음) 등을 위해 추천 인터페이스에 대한 다양한 내비게이션 컨트롤들(1130A, 1130B)을 프레젠테이션한다.

[0097] 실시예에서, 프라이버시 관련 우려를 위해, 412 또는 414와 같은 서버는 예컨대 소스 이미지를 프로세싱한 후, 소스 이미지 또는 임의의 도출된 이미지를 지속적으로 저장하지 않는다. 실시예에서, 애플리케이션(404)은 얼굴 속성 프로세싱 및 제품 시뮬레이션에 대한 요청과 함께(그에 연관하여) 소스 이미지를 전송한다.

[0098] 결정된 얼굴 속성들은 전자상거래 서비스에서 사용자 경험을 지원하고 적용가능한 사용자 인터페이스의 동작을 정의하는 다수의 목적에 유용하다. 예를 들어 애플리케이션 및 그것의 인터페이스는 다음 중 임의의 하나 이상을 수행하도록 구성가능하다: 1. 사용자가 파운데이션의 잘못된 색조를 보고 있는 경우 피부 검출을 사용하여 버추얼 트라이-온(Virtual Try-on)에 팁을 추가한다; 2. 머리카락 색상 검출을 사용하여 달성가능한 색상들을 추천한다; 3. 머리카락 색상 검출을 사용하여 트렌딩 이미지들로부터 머리카락 색상들을 추출하여 관련 색상들을 추천한다; 4. 더 많은 눈 파라미터를 사용하여 "완벽한 아이 룩(perfect eye look)"을 만든다; 5. 공유된 특성들에 기초하여 사용자에게 맞는 유사한 룩을 찾는다; 6. 얼굴/머리카락 피쳐들을 사용하여, 라이프스타일, 개인 패션 선호도 또는 다른 사용자 특징들을 예측한다; 7. 사용자가 룩을 정한 후, 그들의 얼굴/머리카락 피쳐들에 기초하여 달성해야 할 루틴/액션들을 제공한다; 8. 매력 평가를 계산한다.

[0099] 도 12는 실시예에 따른 동작들(1200)을 보여주는 흐름도이다. 동작들(1200)은 예를 들어 본 명세서에 설명된 실시예에 따른 훈련 및 훈련 데이터 세트에 관련된 방법을 제공한다. 단계(1202)에서, 동작은 추론 시간 이미지(inference time image)로부터 속성들을 예측하는 속성 분류기를 훈련시키기 위한 속성 데이터 세트를 저장하고, 속성 데이터 세트는 복수의 속성을 나타내는 복수의 이미지를 포함하고, 속성들 각각은 복수의 개별 속성 값을 갖는다. 단계(1204)에서, 동작들은 복수의 이미지 중의 각각의 이미지에 대해 소프트 레이블들을 데이터 세트에 저장하고, 소프트 레이블들은 독립적으로 행동하는 복수의 개별 인간 이미지 주석 작성자에 의해 결정되는 속성별 개별 속성 값들을 포함한다. 단계(1206)에서, 동작들은 속성 분류기를 훈련시키기 위해 속성 데이터

세트를 제공한다. 실시예에서, 저장 동작들은 데이터베이스와 같은, 그러나 그에 제한되지는 않는 데이터 저장소에 저장한다.

- [0100] 실시예에서, 동작들(1200)은 (예를 들어, 단계(1208)에서) 속성 데이터 세트를 사용하여 속성 분류기를 훈련시키는 것을 더 포함할 수 있다. 실시예에서, 훈련 시에, 방법은 소프트웨어 레이블들 중에서 진실을 선택하기 위해 "진실" 해결 규칙을 사용하는 것을 포함한다.
- [0101] 실시예에서, 속성들은 속성들의 내재적 관계 및 이질성에 따라 복수의 개별 속성 그룹에 연관되고, 속성 분류기는 속성 그룹들 각각에 대해 하나씩 복수의 서브-모델을 포함한다. 실시예에서, 개별 속성 그룹들 중 하나는 색상 기반 속성들에 대한 색상 속성 그룹이다. 실시예에서, 개별 속성 그룹들 중 하나는 모양 기반 속성들에 대한 모양 속성 그룹이다.
- [0102] 실시예에서, 속성은 얼굴 속성들이다. 실시예에서, 얼굴 속성들은 표 1의 속성들을 포함한다.
- [0103] 동작(1200)의 피쳐들 및 임의의 관련된 실시예는 그에 따라 구성될 때 컴퓨팅 디바이스 양태 및 컴퓨터 프로그램 제품 양태에 따라 제공된다는 것이 이해된다.
- [0104] 도 13은 실시예에 따른 동작(1300)을 보여주는 흐름도이다. 동작들(1300)은 예를 들어 얼굴 속성들을 결정하기 위한 방법을 제공한다. 동작들(1300)은 단계(1302)에서 얼굴 속성들 각각에 대한 예측을 출력하기 위해 딥 러닝 및 지도 회귀를 수행하는 네트워크 모델을 사용하여 얼굴의 소스 이미지를 프로세싱한다. 실시예에서, 네트워크 모델은 얼굴 속성들을 예측하기 위해 개별 분류기들에 의한 분류를 위한 공유 피쳐들의 피쳐 벡터를 생성하도록 딥 러닝을 수행하는 잔차 블록들을 포함하는 컨볼루션 신경망(CNN) 모델을 포함한다.
- [0105] 실시예에서, 얼굴 속성들은 얼굴 속성들의 내재적 관계 및 이질성에 따라 복수의 개별 속성 그룹에 연관되고, 소스 이미지는 속성 그룹들 중의 개별 속성 그룹에 대해 구성된 개별 네트워크 모델에 의해 프로세싱된다. 실시예에서, 개별 속성 그룹들 중 하나는 색상 기반 얼굴 속성들에 대한 색상 속성 그룹이고, 소스 이미지는 색상 기반 네트워크 모델에 의해 프로세싱된다. 실시예에서, 개별 속성 그룹들 중 하나는 모양 기반 얼굴 속성들에 대한 모양 속성 그룹이고, 소스 이미지는 모양 기반 네트워크 모델에 의해 프로세싱된다.
- [0106] 실시예에서, 동작들(1300)은 단계(1304)에서 얼굴 속성들 중의 개별 얼굴 속성에 대해 개별 분류기에 의해 프로세싱될 피쳐 벡터를 복제한다.
- [0107] 실시예에서, 네트워크 모델은 복수의 개별 분류기를 포함하고, 개별 분류기들 각각은 하나 이상의 완전 연결된 선형 레이어를 포함하고, 개별 분류기들 각각은 얼굴 속성들 중의 개별 얼굴 속성의 예측을 출력으로서 제공한다. 실시예에서, 복수의 개별 분류기는 얼굴 속성들을 제공하기 위해 병렬로 수행된다.
- [0108] 실시예에서, 모델은 예측 레이어들이 트리밍된, 적응된 사전 훈련된 ResNet 기반 이미지 프로세싱 네트워크 모델을 포함한다.
- [0109] 실시예에서, 단계(1306)에서, 동작들은 이미지 검색, 얼굴 인식, 및 제품 및/또는 서비스 추천 중 임의의 것을 수행하기 위해 (예를 들어, 애플리케이션에 의해 사용하기 위한) 얼굴 속성들을 제공한다.
- [0110] 실시예에서, 단계(1308)에서, 동작들은 이미지 검색, 얼굴 인식, 및 제품 및/또는 서비스 추천 중 임의의 것을 수행한다.
- [0111] 실시예에서, 애플리케이션은 얼굴 속성들에 응답하여 제품 및/또는 서비스 추천을 수행하고, 애플리케이션은 추천된 제품 또는 서비스에 연관된 효과를 시뮬레이션하도록 소스 이미지를 수정하여 가상 사용 사용자 경험(virtual try on user experience)을 제공한다.
- [0112] 동작(1300)의 피쳐들 및 임의의 관련된 실시예는 그에 따라 구성될 때 컴퓨팅 디바이스 양태 및 컴퓨터 프로그램 제품 양태에 따라 제공된다는 것이 이해된다.
- [0113] 도 14는 실시예에 따른 동작들(1400)을 보여주는 흐름도이다. 동작들(1400)은 얼굴 이미지의 얼굴 속성들을 결정하기 위한 방법을 제공한다. 단계(1402)에서, 동작들은 공유 피쳐들의 피쳐 벡터를 생성하기 위해 잔차 블록들을 갖는 컨볼루션 신경망(CNN) 기반 백본 네트워크 모델을 사용하여 얼굴의 소스 이미지를 프로세싱한다. 단계(1402)에서, 동작들은 결정될 K개의 개별 얼굴 속성 각각에 대해 하나씩, 복수(K개)의 분류기를 사용하여 피쳐 벡터를 프로세싱하고, 각각의 개별 분류기는 K개의 개별 얼굴 속성 중 하나의 얼굴 속성의 예측을 출력하기 위해, 하나 이상의 완전 연결된 선형 레이어 및 예측 블록을 포함한다.

- [0114] 실시예에 따르면, 프로세싱 단계들(1402, 1404)은 얼굴의 N개의 소스 이미지에 대해 수행되고, 각각의 얼굴 속성의 N개의 예측으로부터 K개의 개별 얼굴 속성의 최종 예측을 결정하기 위해 예측 해결 규칙(prediction resolving rule)을 사용한다.
- [0115] 실시예에 따르면, CNN 기반 백본 네트워크 모델은 제1 네트워크 모델 및 제2 네트워크 모델을 포함한다. 실시예에 따르면, 제1 네트워크 모델은 복수의 분류기 중에서 색상 기반 얼굴 속성들을 예측하도록 구성된 개별 분류기들에 의한 프로세싱을 위한 색상 기반 피쳐 벡터를 생성하기 위해 색상 기반 얼굴 속성들에 대해 소스 이미지를 프로세싱한다.
- [0116] 실시예에 따르면, 제2 네트워크 모델은 복수의 분류기 중에서 모양 기반 얼굴 속성들을 예측하도록 구성된 개별 분류기들에 의한 프로세싱을 위한 모양 기반 피쳐 벡터를 생성하기 위해 모양 기반 얼굴 속성들에 대해 소스 이미지를 프로세싱한다.
- [0117] 실시예에 따르면, 피쳐 벡터는 K개의 분류기에 의한 병렬 프로세싱을 위해 복제된다.
- [0118] 실시예에서, 단계(1406)에서, 동작들은 이미지 검색, 얼굴 인식, 및 제품 및/또는 서비스 추천 중 임의의 것을 수행하기 위해 (예를 들어, 애플리케이션에 의해 사용하기 위한) 얼굴 속성들을 제공한다.
- [0119] 실시예에서, 단계(1408)에서, 동작들은 이미지 검색, 얼굴 인식, 및 제품 및/또는 서비스 추천 중 임의의 것을 수행한다.
- [0120] 실시예에서, 애플리케이션은 얼굴 속성들에 응답하여 제품 및/또는 서비스 추천을 수행하고, 애플리케이션은 추천된 제품 또는 서비스에 연관된 효과를 시뮬레이션하도록 소스 이미지를 수정하여 가상 사용 사용자 경험을 제공한다.
- [0121] 동작(1400)의 피쳐들 및 임의의 관련된 실시예는 그에 따라 구성될 때 컴퓨팅 디바이스 양태 및 컴퓨터 프로그램 제품 양태에 따라 제공된다는 것이 이해된다.
- [0122] 도 15는 실시예에 따른 동작들(1500)을 보여주는 흐름도이다. 동작들(1500)은 현실을 시뮬레이션하기 위한 방법을 제공한다. 1502에서의 동작들은 얼굴의 소스 이미지로부터 복수의 얼굴 속성을 결정하고, 얼굴 속성 분류 네트워크 모델을 사용하여 소스 이미지를 프로세싱한다. 1504에서, 동작들은 얼굴 속성들 중 적어도 하나에 응답하여 얼굴에 적용하기 위한 적어도 하나의 얼굴 효과를 결정한다. 동작들(1506)은 디스플레이를 위해 얼굴에 적어도 하나의 얼굴 효과를 적용하기 위해 소스 이미지를 프로세싱한다.
- [0123] 실시예에서, 동작들(1500)의 얼굴 속성들은 개별 동작들(1200 또는 1300)의 임의의 관련 실시예를 포함하여, 그러한 개별 동작들(1200 또는 1300)을 사용하여 결정된다. 따라서, 동작들(1500)은 임의의 관련 실시예를 포함하는 이러한 동작들(1200 또는 1300)을 수행하는 것을 포함할 수 있다.
- [0124] 실시예에서, 적어도 하나의 얼굴 효과는 얼굴에 적용될 적어도 하나의 메이크업 제품 및/또는 기법의 시뮬레이션이다.
- [0125] 실시예에서, 적어도 하나의 얼굴 효과를 적용하기 위해 소스 이미지를 프로세싱하는 것은 얼굴 효과를 적용하도록 구성된 딥 러닝 네트워크를 사용하는 것을 포함한다.
- [0126] 실시예에서, 단계(1508)에서의 동작들(1500)은 얼굴 속성들과 관련하여 메이크업 제품들 및/또는 기법들을 저장하는 데이터 저장소로부터 하나 이상의 메이크업 제품 및/또는 기법을 선택하기 위해 얼굴 속성들 중 적어도 하나를 사용한다.
- [0127] 동작(1500)의 피쳐들 및 임의의 관련된 실시예는 그에 따라 구성될 때 컴퓨팅 디바이스 양태 및 컴퓨터 프로그램 제품 양태에 따라 제공된다는 것이 이해된다.
- [0128] 도 16은 실시예에 따른 동작들(1600)을 보여주는 흐름도이다. 동작들(1600)은 전자상거래 서비스를 위해 제품들을 추천하기 위한 방법을 제공한다. 단계(1602)에서, 동작들은 얼굴의 소스 이미지로부터 결정된 복수의 얼굴 속성을 수신하고, 소스 이미지는 복수의 얼굴 속성을 생성하기 위해 얼굴 속성 분류 네트워크 모델을 사용하여 프로세싱된다. 1604에서, 동작들은 제품들에 적합한 얼굴 속성들에 관련하여 제품들을 저장하는 데이터 저장소로부터 적어도 하나의 제품을 선택하기 위해 얼굴 속성들 중 적어도 일부를 사용한다. 1606에서, 동작들은 제품들을 구매하기 위해 전자상거래 인터페이스에 프레젠테이션하기 위한 추천으로서 적어도 하나의 제품을 제공한다.

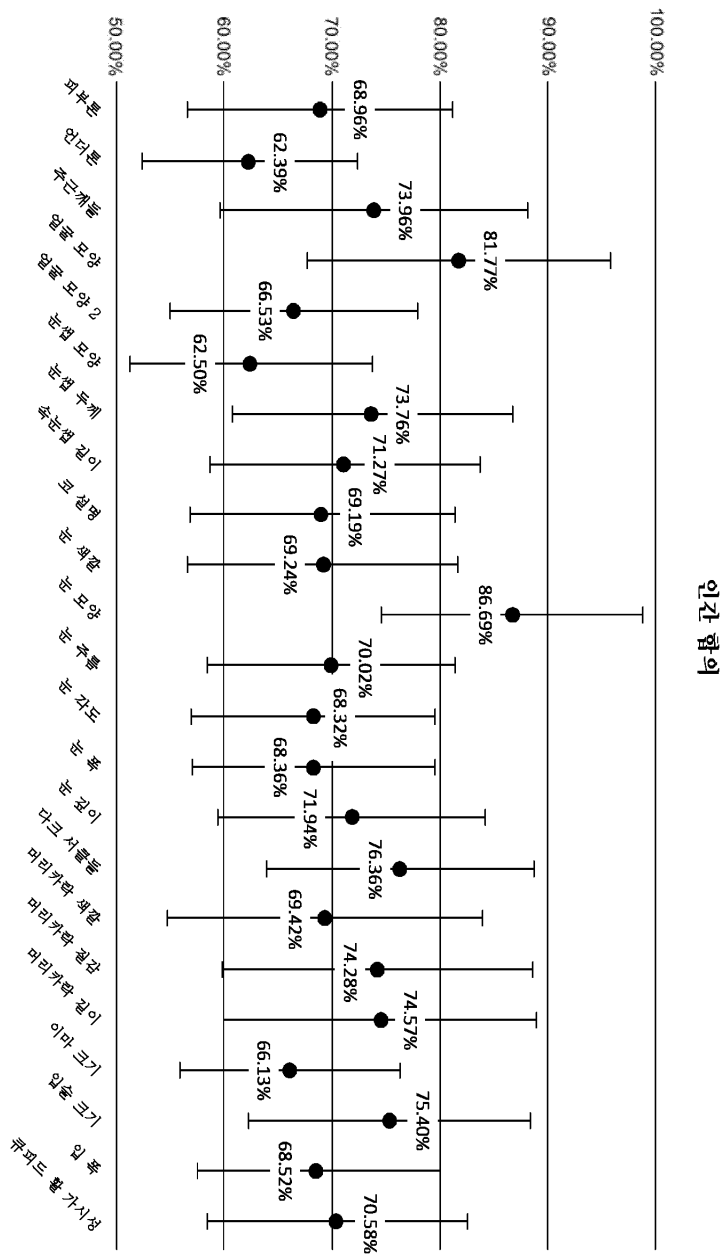
- [0129] 실시예에서, 동작(1600)의 얼굴 속성들은 개별 동작들(1200 또는 1300)의 임의의 관련 실시예를 포함하여, 그러한 개별 동작들(1200 또는 1300)을 사용하여 결정된다. 따라서, 동작들(1600)은 임의의 관련 실시예를 포함하는 이러한 동작들(1200 또는 1300)을 수행하는 것을 포함할 수 있다.
- [0130] 실시예에서, 제품들은 메이크업 제품들을 포함한다.
- [0131] 실시예에서, 메이크업 제품들은 미리 결정된 메이크업 룩들을 정의하기 위해 데이터 저장소에서 다양하게 연관되고, 동작들(도시되지 않음)은 미리 결정된 메이크업 룩들 중 하나의 식별을 수신하고; 얼굴 속성들 중 적어도 일부를 사용하는 단계는 적어도 하나의 제품을 선택할 때 미리 결정된 메이크업 룩들 중 하나에 응답한다. 실시예에서, 메이크업 제품들 각각은 복수의 메이크업 타입 중 하나에 연관되고, 방법은 추천을 정의하기 위해 메이크업 타입들 각각에 대해 얼굴 속성들에 응답하여 적어도 하나의 제품을 선택하는 것을 포함한다. 실시예에서, 메이크업 타입들은 페이스 제품 타입, 아이 제품 타입, 브로우 제품 타입, 및 립 제품 타입을 포함한다.
- [0132] 실시예에서, 동작들(도시되지 않음)은 추천 제품들을 사용하기 위한 기법을 추가로 추천한다.
- [0133] 실시예에서, 예를 들어, 전자상거래 인터페이스를 포함할 수 있는 인터페이스를 통해, 동작들(도시되지 않음)은 소스 이미지를 프로세싱함으로써 얼굴에 적용되는 적어도 하나의 제품의 시뮬레이션을 제공한다.
- [0134] 실시예에서, 동작들(도시되지 않음)은 추천 제품들 중 적어도 일부를 구매하기 위해 전자상거래 쇼핑 서비스를 제공한다.
- [0135] 동작(1600)의 피쳐들 및 임의의 관련된 실시예는 그에 따라 구성될 때 컴퓨팅 디바이스 양태 및 컴퓨터 프로그램 제품 양태에 따라 제공된다는 것이 이해된다.
- [0136] 여기에 도시되고 설명된 컴퓨팅 디바이스(들) 및 인터페이스들은 복수의 상이한 양태를 제공한다. 예를 들어, 실시예에서, 컴퓨팅 디바이스는 시스템의 컴포넌트와 같이 구성가능하며, 시스템은 개별 기능들을 수행하도록 구성된 프로세싱 회로부를 포함하는 하나 이상의 특정 기능 유닛을 포함한다.
- [0137] 하나의 그러한 양태에서, 도 17에 도시된 바와 같이, 시스템(1700)이 제공되며, 이는 얼굴의 소스 이미지로부터 복수의 얼굴 속성을 추출하고, 추출된 복수의 얼굴 속성에 기초하여 하나 이상의 얼굴 효과를 생성하도록 구성된 프로세싱 회로부를 포함하는 얼굴 속성 유닛(1702); 및 소스 이미지에 적어도 하나의 얼굴 효과를 적용하고, 효과 적용된 소스 이미지의 하나 이상의 가상 인스턴스를 전자상거래 인터페이스 상에 생성하도록 구성된 프로세싱 회로부를 포함하는 얼굴 효과 유닛(1704)을 포함한다.
- [0138] 실시예에서, 시스템은 복수의 얼굴 속성 중 하나 이상과 관련하여 제품들을 저장하는 데이터 저장소로부터 적어도 하나의 제품을 결정하고, 제품들을 구매하기 위해 전자상거래 인터페이스 상에서 제품 추천의 하나 이상의 가상 인스턴스를 생성하도록 구성된 프로세싱 회로부를 포함하는 사용자 경험 유닛(1706)을 더 포함한다.
- [0139] 실시예에서, 얼굴 속성 유닛은 얼굴의 소스 이미지로부터 복수의 얼굴 속성 각각에 대한 예측을 출력하기 위해 딥 러닝 및 지도 회귀를 수행하는 네트워크 모델을 실행하도록 구성된 프로세싱 회로부를 포함한다.
- [0140] 실시예에서, 얼굴 효과 유닛은 효과 적용된 소스 이미지의 적어도 일부, 및 메이크업 제품 또는 메이크업 적용 기법 중 적어도 하나를 포함하는 가상 표현을 생성하도록 구성된 프로세싱 회로부를 포함한다.
- [0141] 실시예에서, 얼굴 효과는 얼굴 속성들 중 하나에 응답하는 주석을 포함한다. 실시예에서, 주석이 얼굴 속성에 대한 얼굴 속성 값을 포함하고/거나 주석이 얼굴 속성이 결정된 소스 이미지의 영역에 위치(예를 들어, 소스 이미지 상에(예를 들어, 오버레이로서) 위치)된다.
- [0142] 실시예에서, 얼굴 속성 유닛(1702)에 의해 결정된 얼굴 속성들은 개별 동작들(1200 또는 1300)의 임의의 관련 실시예를 포함하여, 그러한 개별 동작들(1200 또는 1300)을 사용하여 결정된다. 따라서, 얼굴 속성 유닛은 임의의 관련 실시예를 포함하는 이러한 동작들(1200 또는 1300)을 수행하도록 구성될 수 있다.
- [0143] 다른 실시예들에서, 다른 실제 응용들은 본 명세서에 설명된 것과 같은 얼굴 속성 분류기를 사용한다. 다른 실시예들은 이미지 검색, 얼굴 인식 등을 포함한다. 이미지 검색을 위한 주된 태스크는 매우 큰 데이터베이스로부터 이미지를 검색하는 것이다(이러한 큰 크기에서, 순차적으로 검색하는 것은 보통 작동하지 않을 것이다). 실시예에서, 얼굴 속성들은 이미지 데이터베이스에 저장된 이미지들에 대해 계산되고 그에 따라 키인된다. 검색을 위한 후보 이미지는 또한 계산된 속성들을 가지며, 이들은 검색 기준을 좁히거나 신뢰 일치 데이터를 정의하거나 검색 결과들을 정렬하는(예를 들어, 후보 이미지는 하나의 특정 저장된 이미지의 7개 얼굴 속성 중 5개와 일치하고, 두 번째 특정 저장된 이미지의 7개 얼굴 속성 중 6개와 일치하며, 그 결과는 검색 결과들을 정렬

하거나 (부분적으로) 일치하는 저장된 이미지들 중에서 하나를 선택하는 데 사용됨) 등을 위해 사용된다

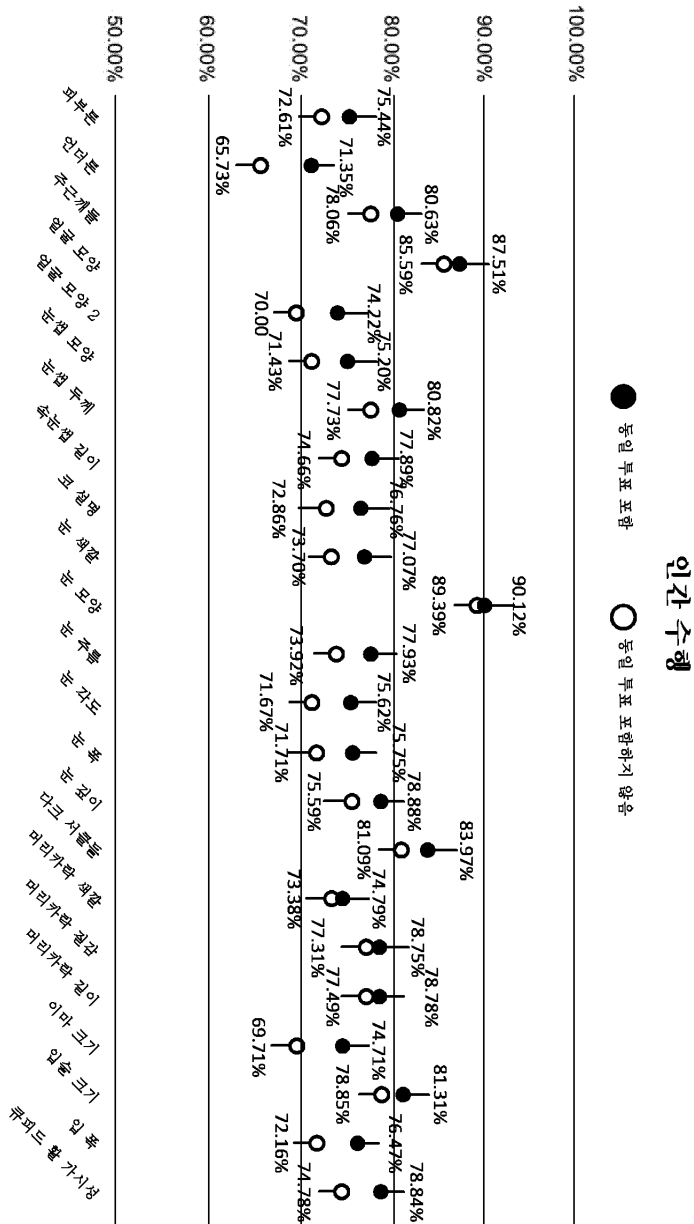
- [0144] 예를 들어, ResNet과 같은 CNN 모델들은 이미지들을 원본 이미지의 압축된 표현인 "피쳐들"로서 표현할 수 있고, 따라서 통상적으로 훨씬 더 작고 검색하기 쉽다.
- [0145] 얼굴 인식에서, 이러한 태스크는 주로 동일한 신원(사람)의 상이한 이미지들을 찾거나 분류하거나 그룹화하는 것에 관련된다. 유사한 기법들이 얼굴 속성들을 사용하여 수행가능하다.
- [0146] 본 개시내용은 어떠한 이전의 작업들 또는 데이터세트들에서도 이전에 다루지 않았던 (얼굴) 속성들의 특별한 세트를 예측하기 위한 방법을 가능하게 하는 포괄적인 데이터세트의 컬렉션을 설명한다. 이전 작업들(예를 들어, "CelebFaces Attributes Dataset"(CelebA) 데이터세트(URL: mmlab.ie.cuhk.edu.hk/projects/CelebA.html에서 입수가능한 홍콩 Chinese University의 Multimedia Laboratory로부터의 얼굴 속성 데이터 세트))은 이전에 컨볼루션 신경망(CNN)을 탐색했지만, 그러한 것은 각각의 속성에 대해 상이한 분기들을 설정하지 않았다. 즉, 각각의 상이한 속성에 대한 별개의 분류기 헤드들이 사용되지 않았다. 색상 및 모양과 같은 내재적 관계들을 갖는 상이한 속성들에 대해, 상이한 서브모델들이 사용되지 않았다.
- [0147] 본 명세서에 도시되고 설명된 데이터세트의 컬렉션의 실시예에서, 훈련 워크플로우는 모델 정확도를 개선하기 위해 소프트 레이블들을 사용했다. 훈련의 결과인 속성 분류기들은 인간의 주석과 동등하다.
- [0148] 실제 구현은 여기에 설명된 피쳐들 중 임의의 것 또는 전부를 포함할 수 있다. 이들 및 다른 양태들, 피쳐들, 및 다양한 조합들은 본 명세서에 설명된 피쳐들을 결합하는 방법들, 장치들, 시스템들, 기능을 수행하기 위한 수단들, 프로그램 제품들로서, 및 다른 방식들로 표현될 수 있다. 다수의 실시예가 설명되었다. 그럼에도 불구하고, 본 명세서에 설명된 프로세스들 및 기법들의 사상 및 범위를 벗어나지 않고서 다양한 수정이 이루어질 수 있음을 이해할 것이다. 추가로, 설명된 프로세스로부터 다른 단계들이 제공될 수 있거나 단계들이 제거될 수 있고, 다른 컴포넌트들이 설명된 시스템들에 추가되거나 그로부터 제거될 수 있다. 따라서, 다른 실시예들은 이하의 청구항들의 범위 내에 있다.
- [0149] 본 명세서의 설명 및 청구항들의 전반에 걸쳐, "포함한다(comprise)" 및 "포함한다(contain)"라는 단어 및 이들의 변형은 "포함하지만 이에 제한되지 않음"을 의미하며, 이들은 다른 컴포넌트들, 정수들 또는 단계들을 배제하도록 의도되지 않는다(그리고 배제하지 않는다). 본 명세서 전체에서, 단수형은 문맥상 달리 요구하지 않는 한 복수형을 포괄한다. 특히, 부정 관사가 사용되는 경우, 본 명세서는 문맥상 달리 요구하지 않는 한 단수형 뿐만 아니라 복수형을 고려하는 것으로 이해되어야 한다.
- [0150] 본 발명의 특정 양태, 실시예 또는 예와 관련하여 설명된 피쳐들, 정수들, 특성들 또는 그룹들은 양립불가능하지 않은 한, 임의의 다른 양태, 실시예 또는 예에 적용가능한 것으로 이해되어야 한다. 여기에 개시된 모든 피쳐(임의의 첨부된 청구항들, 요약 및 도면들을 포함함) 및/또는 그렇게 개시된 임의의 방법 또는 프로세스의 모든 단계는, 그러한 피쳐들 및/또는 단계들의 적어도 일부가 상호배타적인 조합들을 제외하고 임의의 조합으로 조합될 수 있다. 본 발명은 임의의 전술한 예들 또는 실시예들의 세부사항들로 제한되지 않는다. 본 발명은 본 명세서(임의의 첨부된 청구항들, 요약 및 도면들을 포함함)에 개시된 피쳐들 중 임의의 신규한 것 또는 그들의 임의의 신규한 조합, 또는 개시된 임의의 방법 또는 프로세스의 단계들 중 임의의 신규한 것 또는 그들의 임의의 신규한 조합으로 확장된다.
- [0151] 참조들
- [0152] 1. Deep Residual Learning for Image Recognition, He, Kaiming et. al, 2015-12-10, URL arxiv.org/abs/1512.03385에서 입수가능.
- [0153] 2. A Survey of Deep Facial Attribute Analysis, Xin Zheng, Xin et. al, 2018-12-26 제출(v1), 2019-10-27 최종 수정(해당 버전, v3) URL arxiv.org/abs/1812.10265.
- [0154] 3. Deep Learning Face Attributes in the Wild, Liu, Ziwei et. al, 2014-11-28 제출(v1), 2015-9-24 최종 수정(해당 버전, v3), URL: arxiv.org/abs/1411.7766.

도면

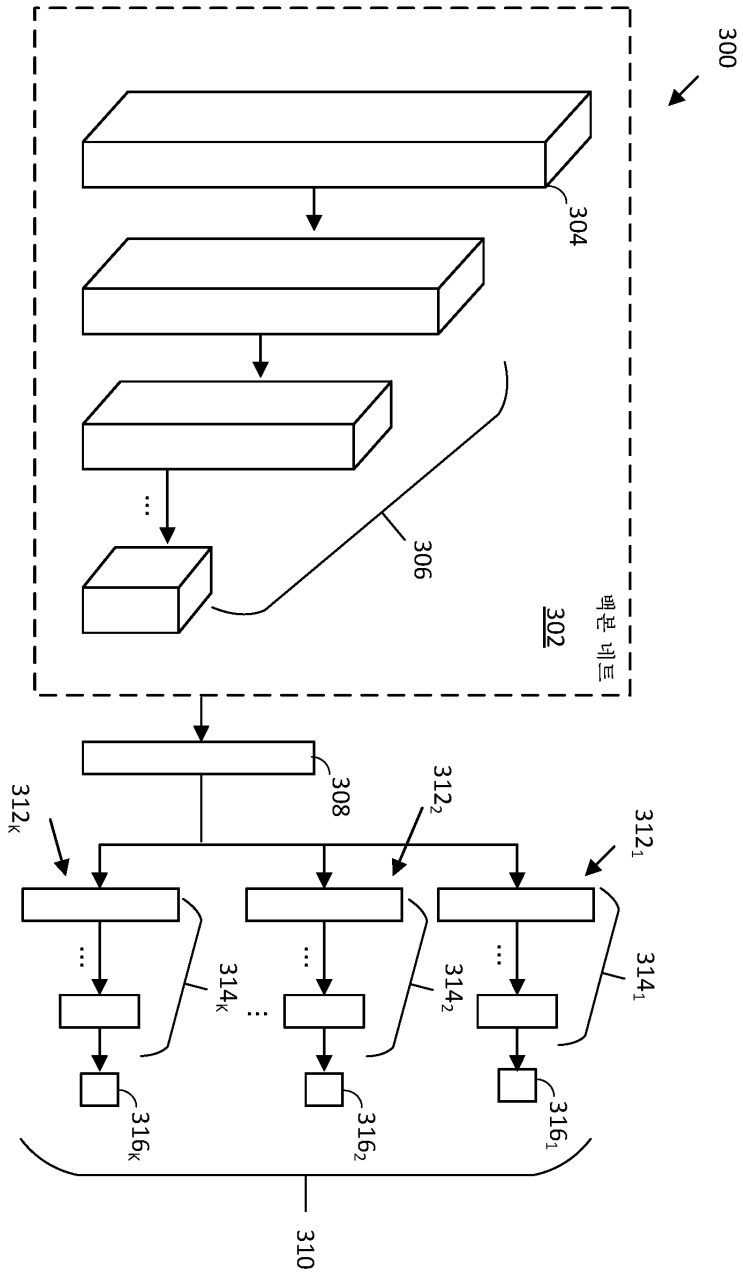
도면1



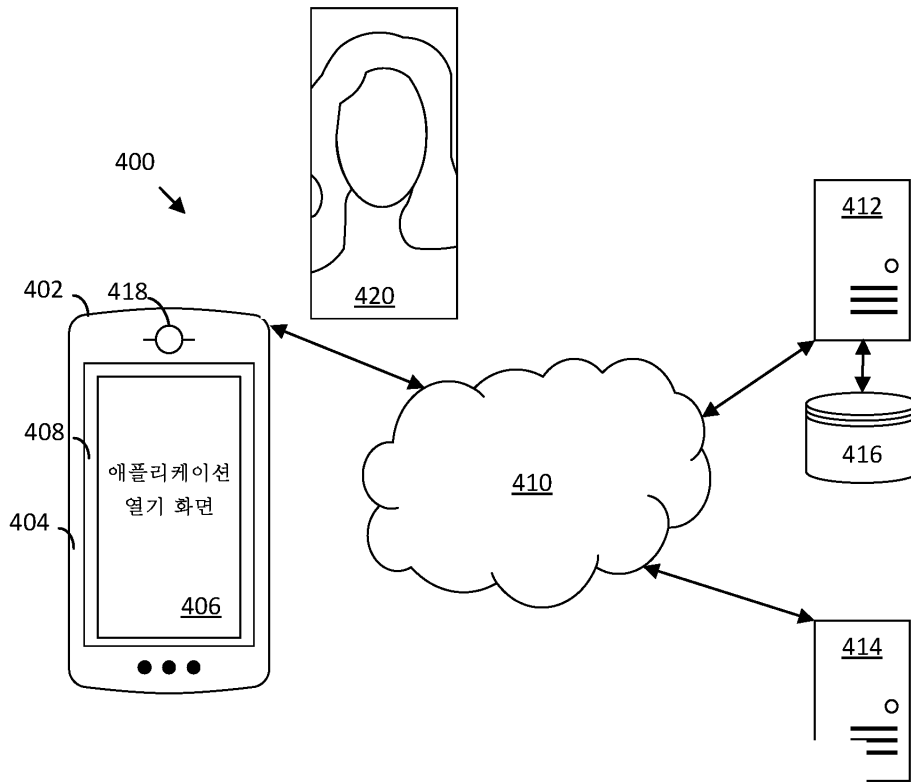
도면2



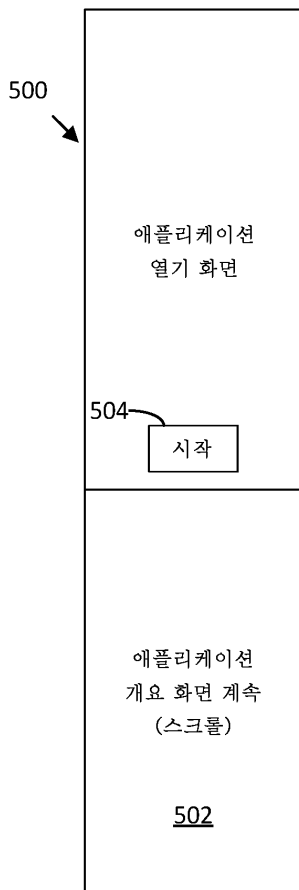
도면3



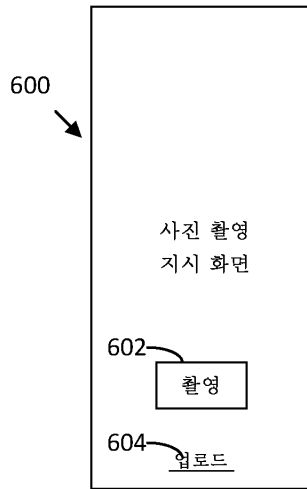
도면4



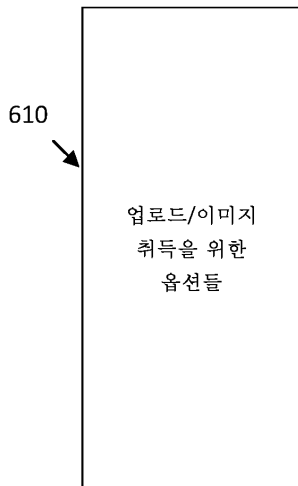
도면5



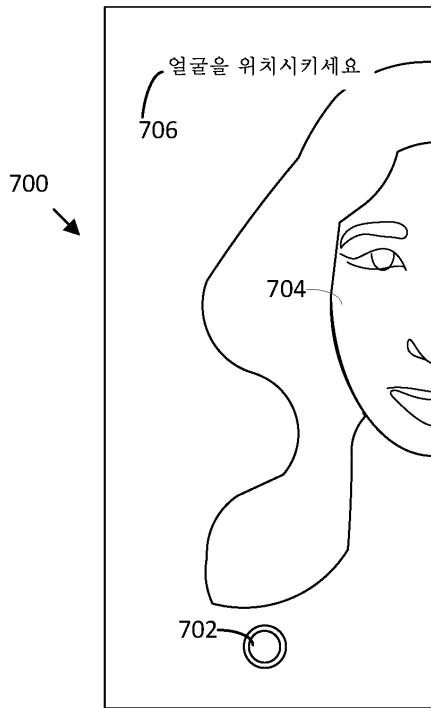
도면6a



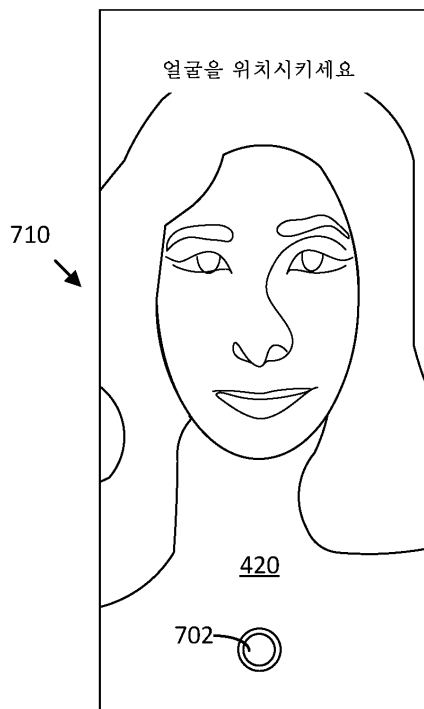
도면6b



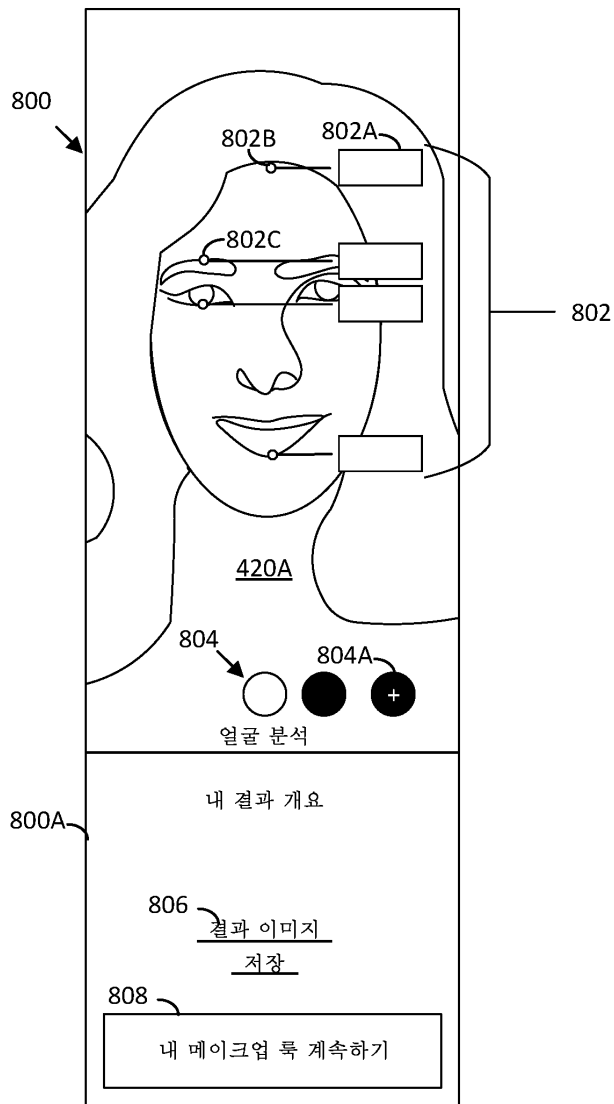
도면7a



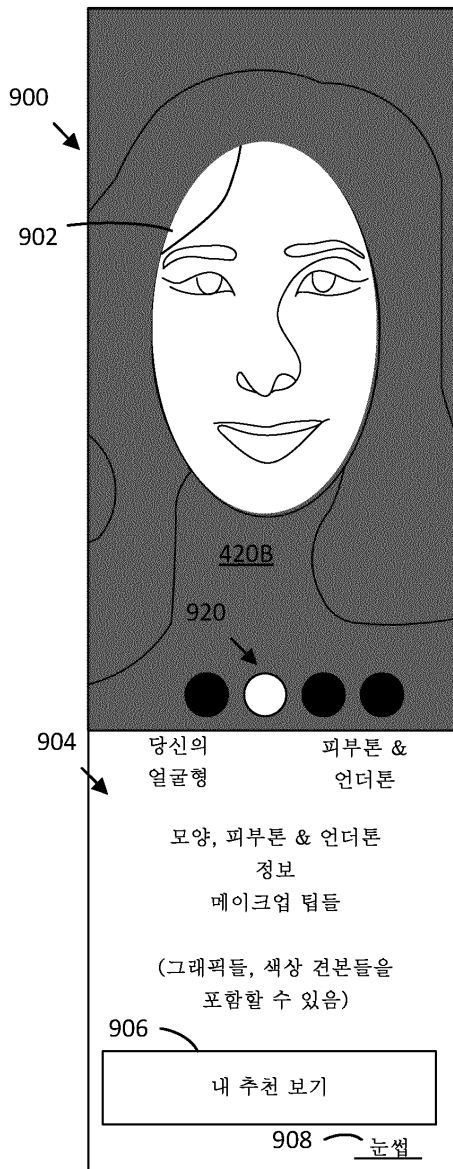
도면7b



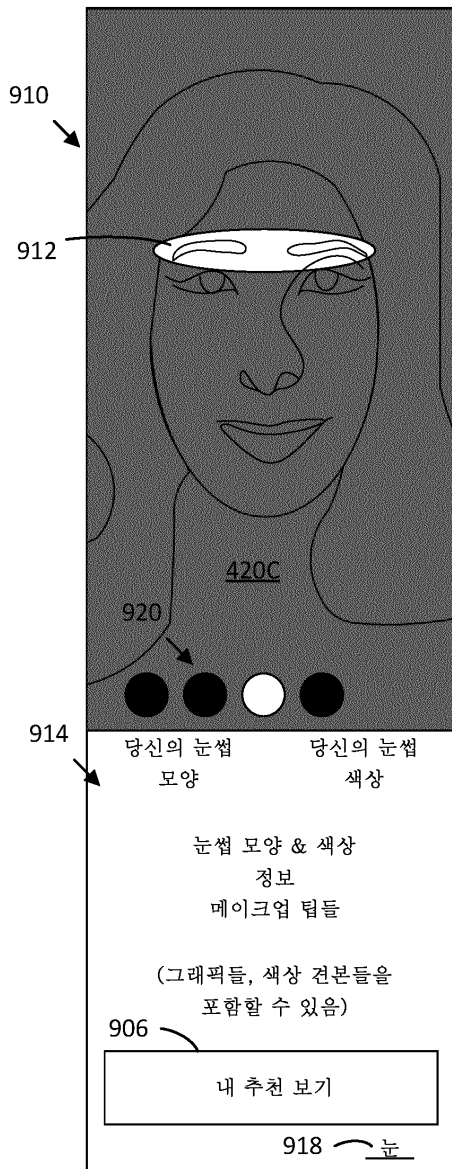
도면8



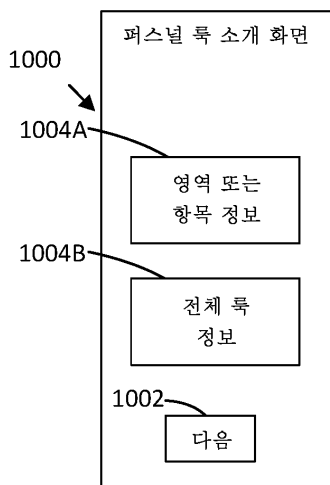
도면9a



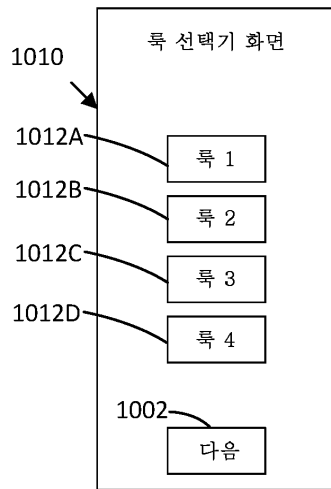
도면9b



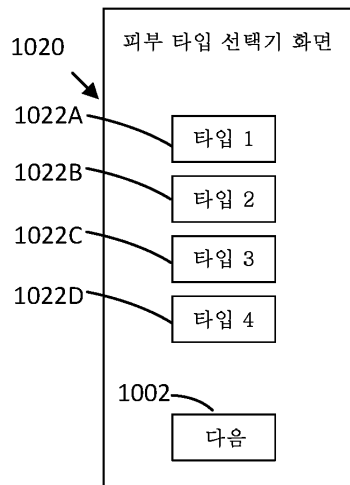
도면10a



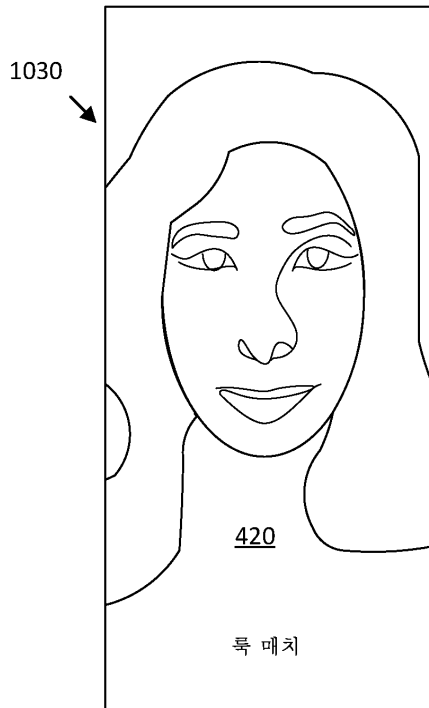
도면10b



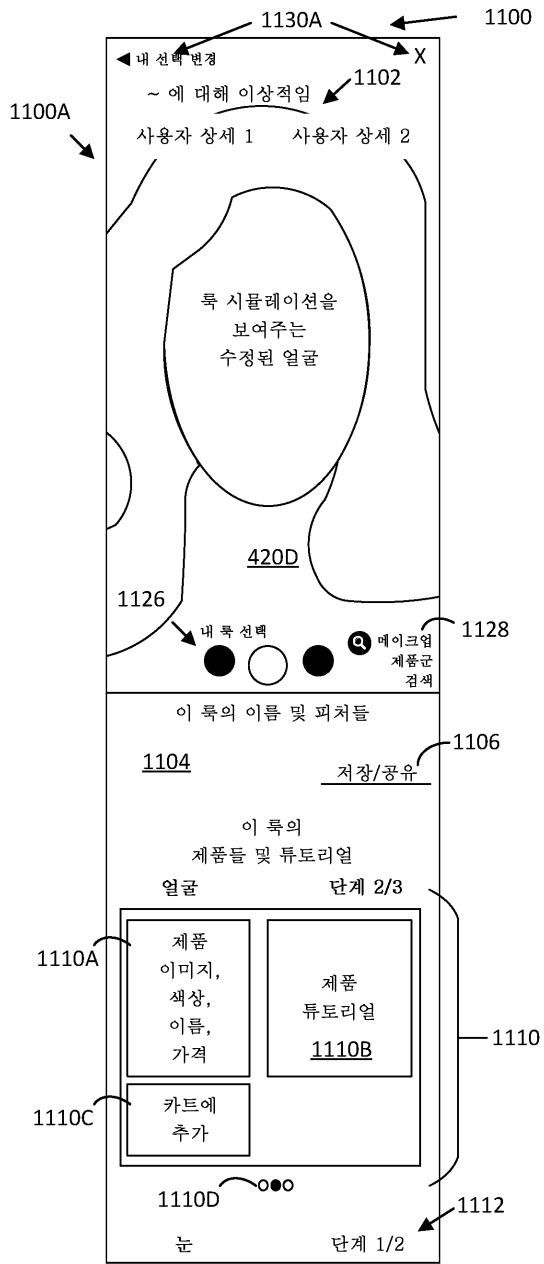
도면10c



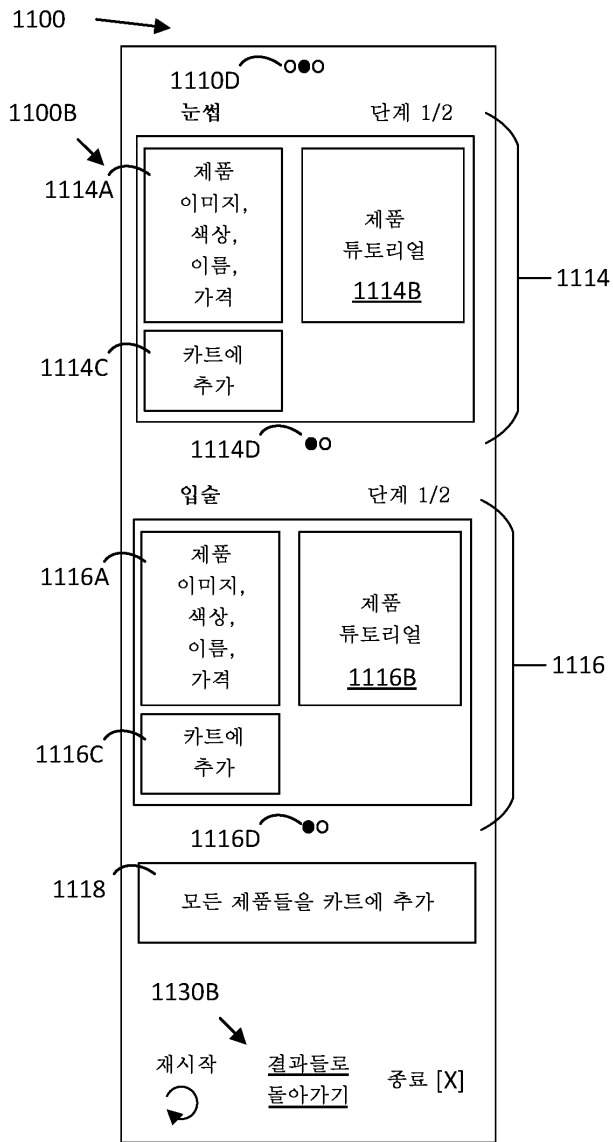
도면10d



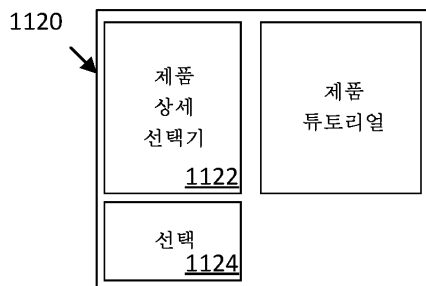
도면11a



도면11b

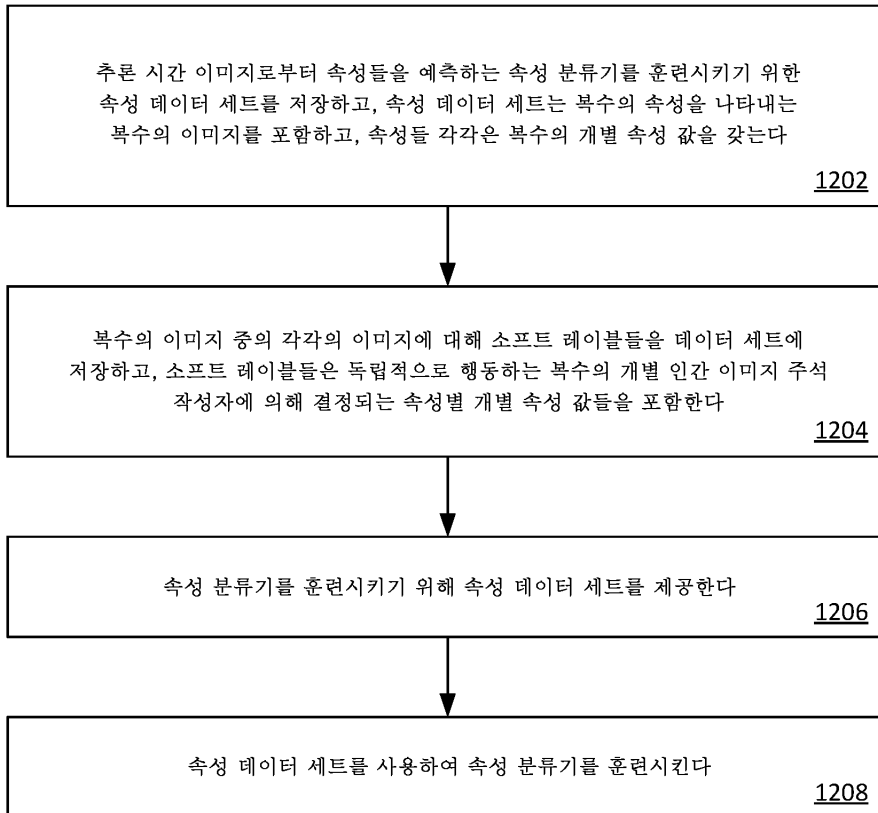


도면11c



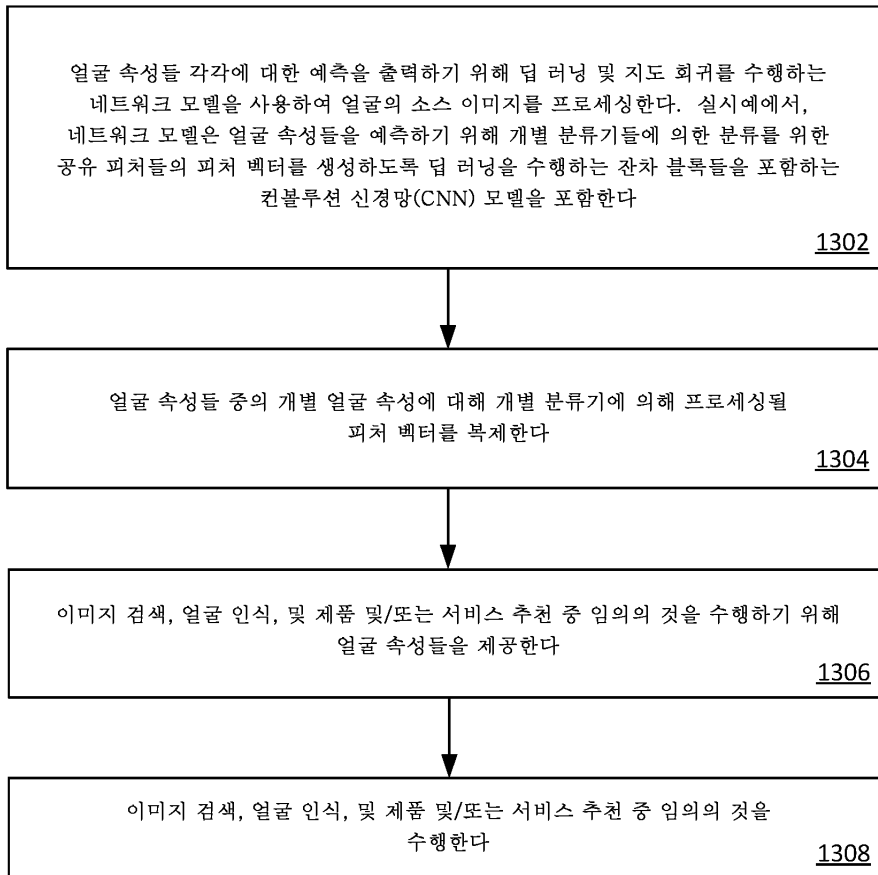
도면12

1200



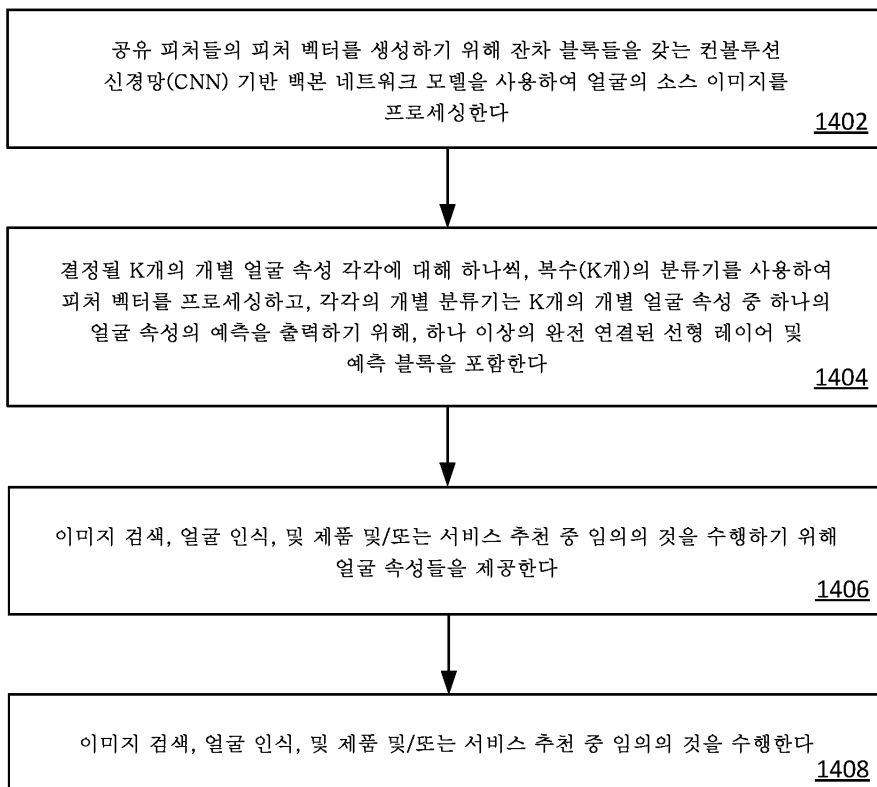
도면13

1300

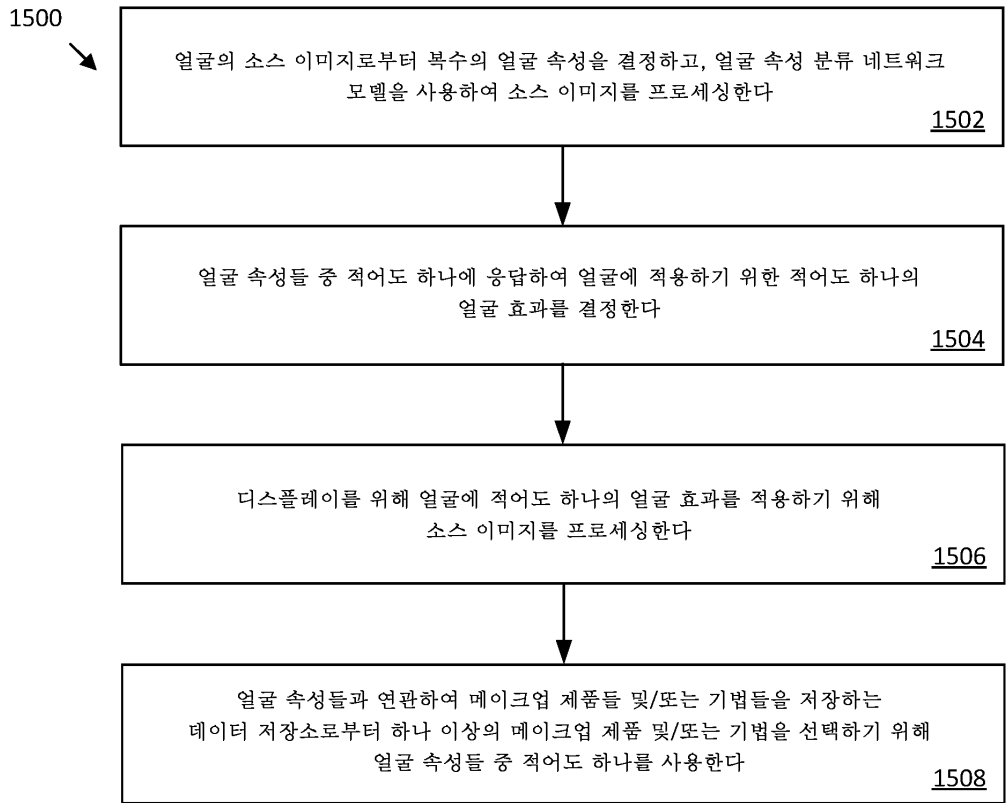


도면14

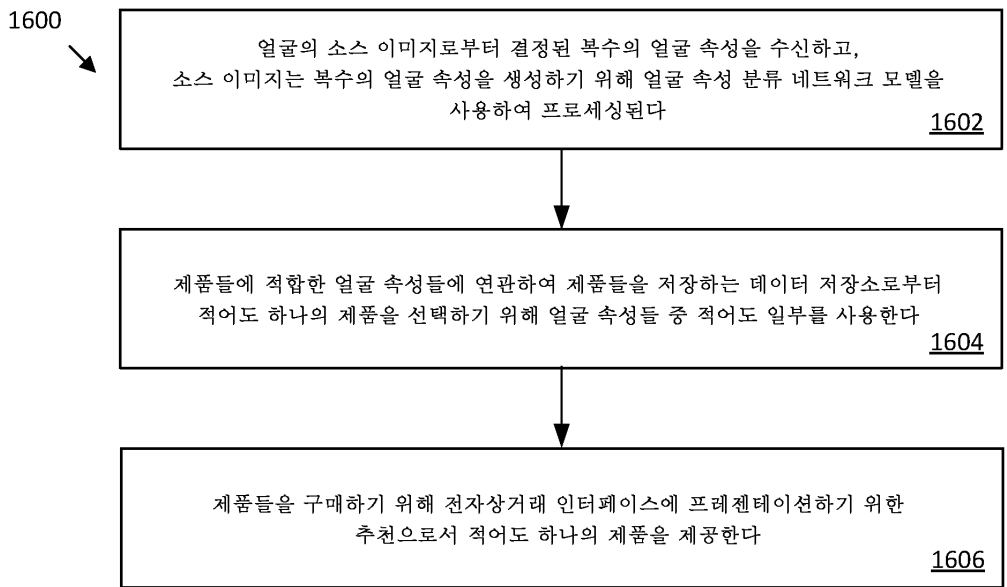
1400



도면15



도면16



도면17

