(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0124652
(43) 공개일자 2024년08월19일

(51) 국제특허분류(Int. Cl.)
G06T 5/20 (2024.01) *G06N 20/20* (2019.01)
G06T 5/00 (2024.01) *G06T 7/11* (2017.01)
G06V 40/16 (2022.01) *H04N 21/274* (2011.01)
H04N 21/43 (2011.01) *H04N 21/81* (2011.01)
(52) CPC특허분류
G06T 5/20 (2024.01)
G06N 20/20 (2021.08)
(21) 출원번호 10-2023-0017409
(22) 출원일자 2023년02월09일
심사청구일자 2023년02월09일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
하정욱
서울특별시 송파구 잠실로 62, 트리지움 APT 344동 1701호
서현덕
서울특별시 종로구 대학로5길 56, 401호
(뒷면에 계속)
(74) 대리인
특허법인 피씨알

전체 청구항 수 : 총 16 항

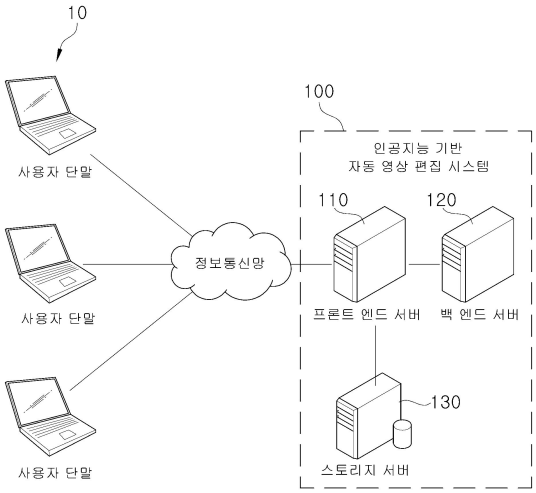
(54) 발명의 명칭 **인공지능 기반 자동 영상 편집 시스템 및 방법**

(57) 요 약

본 발명은 자동 영상 편집 시스템을 개시한다. 보다 상세하게는, 본 발명은 배경에 복수의 객체가 등장하는 영상에서 특정 객체에 대한 마스킹을 자동으로 수행하는 인공지능 기반 자동 영상 편집 시스템 및 방법에 관한 것이다.

본 발명의 실시예에 따르면, 웹 페이지를 통해 사용자가 편집 대상인 영상 및 객체 인식을 위한 이미지를 업로드할 수 있는 웹 인터페이스를 제공하고, 백 엔드 서버에 구축된 학습 데이터 베이스 및 학습된 객체 탐지 모델을 통해 영상 내 이미지에 대응하는 객체를 인식하여 자동으로 마스킹 처리하여 회신함으로써, 일반 사용자들도 손쉽게 영상을 편집할 수 있는 효과가 있다.

대 표 도 - 도1



(52) CPC특허분류
G06T 5/75 (2024.01)
G06T 7/11 (2017.01)
G06V 40/161 (2022.01)
H04N 21/274 (2013.01)
H04N 21/4307 (2022.08)
H04N 21/8173 (2013.01)
G06T 2207/30201 (2013.01)
(72) 발명자
허승희
서울특별시 마포구 독막로29길 38-7, 501호
임정호
인천광역시 서구 청마로 92, 검단힐스테이트4차 410동 1402호

명 세 서

청구범위

청구항 1

사용자 단말로부터 편집 대상인 영상 및 상기 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받고, 편집 프로젝트를 생성하는 웹 페이지를 제공하는 프론트 엔드 서버; 및

탑재된 제1 및 제2 학습 모델을 이용하여, 상기 편집 프로젝트에 포함되는 영상 내 상기 탐지 대상인 객체의 얼굴 영역을 각각 탐지 및 인식하고, 인식된 객체의 얼굴 영역 또는 상기 인식된 객체를 제외한 나머지 타객체의 얼굴 영역에 마스킹 처리를 수행하여 상기 웹 페이지에 회신하는 백 엔드 서버

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 2

제 1 항에 있어서,

상기 백 엔드 서버는,

상기 영상의 각 프레임으로부터 상기 제1 학습 모델을 이용하여 하나 이상의 객체를 탐지하고, 상기 탐지된 객체의 얼굴 영역 중, 상기 추적 이미지에 대응하는 얼굴 영역을 추출하여 데이터 셋을 생성 및 데이터 베이스에 저장하는 데이터 추출부

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 3

제 2 항에 있어서,

상기 백 엔드 서버는,

상기 데이터 셋에 라벨링을 수행하고, 라벨링된 데이터 셋을 이용하여 상기 제2 학습 모델을 학습시키는 모델 학습부

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 4

제 3 항에 있어서,

상기 백 엔드 서버는,

상기 영상의 각 프레임에 대하여, 학습된 상기 제1 및 제2 학습 모델을 이용한 분석 작업을 통해 상기 추적 이미지에 대응하는 객체의 얼굴 영역을 탐지 및 인식하는 객체 인식부

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 5

제 4 항에 있어서,

상기 객체 인식부는,

상기 편집 프로젝트에 포함되는 영상에 대한 분석 작업 진행시 비동기 방식으로 처리하는 비동기 작업 큐 라이브러리

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 6

제 4 항에 있어서,

상기 객체 인식부는,

분석 작업 진행 전, 상기 편집 프로젝트에 포함되는 영상에서 음원을 추출하고, 분석 작업이 완료되면 마스킹된 영상과 추출된 음원을 결합하는 음원 추출 및 영상 변환 라이브러리

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 7

제 4 항에 있어서,

상기 백 엔드 서버는,

탐지된 객체의 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하고 편집 프로젝트를 종료하는 마스킹 처리부

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 8

제 1 항에 있어서,

상기 프론트 엔드 서버는,

정보통신망을 통해, 상기 사용자 단말로부터 영상 및 추적 이미지를 입력받는 UX가 적용된 웹 페이지를 제공하는 웹 서비스 부

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 9

제 8 항에 있어서,

상기 웹 서비스 부는,

상기 웹 페이지를 통해, 상기 사용자 단말로부터 상기 추적 이미지에 대응하는 객체가 탐지 대상인지 또는 상기 탐지 대상을 제외한 나머지 타객체인지 결정하는 옵션정보를 더 입력받는 것인, 인공지능 기반 자동 영상 편집 시스템.

청구항 10

제 8 항에 있어서,

상기 프론트 엔드 서버는,

완료된 편집 프로젝트의 마스킹 영상을 스토리지 서버에 저장하고, 상기 스토리지 서버의 저장된 마스킹된 영상의 인터넷 접근 경로를 전자 메일 형태로 상기 사용자 단말에 전송하는 메일링 서비스 부

를 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 11

제 10 항에 있어서,

상기 프론트 엔드 서버는,

백그라운드에서 실행되고, 상기 스토리지 서버에 저장된 하나 이상의 변환된 영상에 대하여 일정기간이 경과하면 자동 삭제하는 스케쥴러

를 더 포함하는 인공지능 기반 자동 영상 편집 시스템.

청구항 12

제 1 항에 있어서,

상기 제1 및 제2 학습 모델은,

각각 YOLOv5 모델 및 FaceNet 모델인 것인, 인공지능 기반 자동 영상 편집 시스템.

청구항 13

인공지능 기반 자동 영상 편집 시스템에 의한 영상 편집 방법에 있어서,

시스템에 접속한 사용자 단말에 편집 프로젝트를 생성하는 웹 페이지를 제공하는 단계;

상기 사용자 단말로부터 편집 대상인 영상 및 상기 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받아 편집 프로젝트를 생성하는 단계;

상기 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 상기 탐지 대상에 대응하는 객체의 얼굴 영역을 탐지하는 단계;

학습된 제2 학습 모델을 이용한 분석 작업을 통해 탐지된 객체의 얼굴 영역을 인식하는 단계;

인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계; 및

상기 편집 프로젝트를 종료하는 단계

를 포함하는 인공지능 기반 자동 영상 편집 방법.

청구항 14

제 13 항에 있어서,

상기 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 상기 탐지 대상에 대응하는 객체의 얼굴 영역을 탐지하는 단계 이전에,

제1 학습 모델을 이용하여 상기 영상에 포함되는 하나 이상의 객체의 얼굴 영역을 탐지하는 단계;

상기 탐지된 객체의 얼굴 영역을 추출하여 복수의 얼굴 영역을 포함하는 데이터 셋을 생성하는 단계;

상기 데이터 셋을 데이터 베이스에 저장하는 단계; 및

상기 데이터 셋을 이용하여 상기 제2 학습 모델을 학습시키는 단계

를 포함하는 인공지능 기반 자동 영상 편집 방법.

청구항 15

제 13 항에 있어서,

상기 인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계 이후,

완료된 편집 프로젝트의 마스킹 영상을 스토리지 서버에 저장하는 단계; 및

상기 스토리지 서버의 저장된 마스킹 영상의 인터넷 접근 경로를 전자 메일 형태로 상기 사용자 단말에 전송하는 단계

를 포함하는 인공지능 기반 자동 영상 편집 방법.

청구항 16

제 15 항에 있어서,

상기 편집 프로젝트를 종료하는 단계 이후,

백그라운드에서 실행되는 스케쥴러를 통해, 상기 스토리지 서버에 저장된 일정기간이 경과한 하나 이상의 변환된 영상을 자동 삭제하는 단계

를 더 포함하는 인공지능 기반 자동 영상 편집 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 자동 영상 편집 시스템에 관한 것으로, 특히 배경에 복수의 객체가 등장하는 영상에서 특정 객체에 대한 마스킹을 자동으로 수행하는 인공지능 기반 자동 영상 편집 시스템 및 방법에 관한 것이다.

배 경 기 술

[0002] 최근, 정보통신기술의 발전에 따라 기업이나 단체가 아닌 일반 개인도 SNS 플랫폼을 이용하여 라이브 방송을 인터넷 상에 송출할 수 있게 되었으며, 개인 온라인 방송이 증가하면서 실시간 야외 방송 콘텐츠도 쉽게 찾아볼 수 있게 되었다.

[0003] 그러나, 전술한 실시간 온라인 야외 방송 콘텐츠의 경우, 임의의 장소에서 임의의 시간대에 보행자에 대한 통제가 되지 않는 상황에서 방송이 진행될 수 있으며, 이에 방송 콘텐츠에서 의도하지 않은 보행자가 촬영되어 영상에 등장하는 상황이 종종 발생하게 되며, 이러한 영상이 편집없이 송출되는 경우, 지나가는 보행자들에 대한 초상권 문제가 발생할 수 있다.

[0004] 특히, 실시간 방송의 경우에는 다수의 보행자들에 대하여 사전에 양해를 구하는 것이 어려운 경우가 대부분이고, 길거리를 지나는 일반 보행자나, 상점, 식당 등의 특정 촬영 장소에서 방문한 고객 전체를 대상으로 동의를 구하는 것은 현실적으로 불가능하다고 할 수 있다.

[0005] 또한, 이와 비슷하게 공중파 방송에서는 출연 정지 명단의 사람들 혹은 사회적 물의를 일으킨 연예인의 출연을 제한해야 하며, 이에 해당하는 대상인이 등장하는 장면에 대해서는 영상 내 모자이크를 이용하여 식별할 수 없도록 가리는 작업을 거치게 되는 데, 이러한 작업들을 전문가가 일일이 확인하며 작업하는 것은 비효율적이라 할 수 있다.

[0006] 이러한 배경 속에서 개인 방송을 운영하는 크리에이터들은 모자이크 영상 편집 기술을 도입하기 시작하였으나, 영상을 시청하는 데 거슬림이 최소화되도록 정교하게 모자이크 처리를 하기 위해서는 전문 영상 편집 프로그램을 사용해야만 하며, 이는 전문 지식이 없는 크리에이터들이 다루기에는 어려움이 있다.

[0007] 일예로서, 상용화된 프로그램인 "Adobe Premier"와 같은 전문 편집 프로그램은 일반인들이 사용하기에 진입장벽이 높다는 큰 단점이 있다. 반면 영상 내 드래그 조작으로 일정 영역을 설정하고 해상도를 변경하는 방식의 간편한 편집 기능을 제공하는 프로그램들은 편집 결과물의 질이 현저하게 낮다는 한계가 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 등록특허공보 제10-2277929호(공고일자: 2021.07.14.)

발명의 내용

해결하려는 과제

[0009] 본 발명은 전술한 문제점을 해결하기 위해 안출된 것으로, 본 발명은 영상 편집 기술에 관한 전문적인 지식이 없는 일반인들도 웹 상에서 손쉽게 영상을 편집할 수 있는 서비스를 제공하는 데 과제가 있다.

[0010] 또한, 본 발명은 편집 영상에 대하여, 인공지능 기반 학습 모델을 이용하여 마스킹을 위한 특정 객체에 대한 탐지 절차에서 보다 고성능의 객체 탐지 결과를 획득할 수 있는 시스템을 제공하는 데 과제가 있다.

과제의 해결 수단

[0011] 전술한 과제를 해결하기 위해, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템은, 사용자 단말로부터 편집 대상인 영상 및 상기 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받고, 편집 프로젝트를 생성하는 웹 페이지를 제공하는 프론트 엔드 서버 및, 탑재된 제1 및 제2 학습 모델을 이용하여, 상기 편집 프로젝트에 포함되는 영상 내 상기 탐지 대상인 객체의 얼굴 영역을 각각 탐지 및 인식하고, 인식된

객체의 얼굴 영역 또는 상기 인식된 객체를 제외한 나머지 타객체의 얼굴 영역에 마스킹 처리를 수행하여 상기 웹 페이지에 회신하는 백 엔드 서버를 포함할 수 있다.

[0012] 상기 백 엔드 서버는, 상기 영상의 각 프레임으로부터 상기 제1 학습 모델을 이용하여 하나 이상의 객체를 탐지하고, 상기 탐지된 객체의 얼굴 영역 중, 상기 추적 이미지에 대응하는 얼굴 영역을 추출하여 데이터 셋을 생성 및 데이터 베이스에 저장하는 데이터 추출부를 포함할 수 있다.

[0013] 상기 백 엔드 서버는, 상기 데이터 셋에 라벨링을 수행하고, 라벨링된 데이터 셋을 이용하여 상기 제2 학습 모델을 학습시키는 모델 학습부를 포함할 수 있다.

[0014] 상기 백 엔드 서버는, 상기 영상의 각 프레임에 대하여, 학습된 상기 제1 및 제2 학습 모델을 이용한 분석 작업을 통해 상기 추적 이미지에 대응하는 객체의 얼굴 영역을 탐지 및 인식하는 객체 인식부를 포함할 수 있다.

[0015] 상기 객체 인식부는, 상기 편집 프로젝트에 포함되는 영상에 대한 분석 작업 진행시 비동기 방식으로 처리하는 비동기 작업 큐 라이브러리를 포함할 수 있다.

[0016] 상기 객체 인식부는, 분석 작업 진행 전, 상기 편집 프로젝트에 포함되는 영상에서 음원을 추출하고, 분석 작업이 완료되면 마스킹된 영상과 추출된 음원을 결합하는 음원 추출 및 영상 변환 라이브러리를 포함할 수 있다.

[0017] 상기 백 엔드 서버는, 탐지된 객체의 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하고 편집 프로젝트를 종료하는 마스킹 처리부를 포함할 수 있다.

[0018] 상기 프론트 엔드 서버는, 정보통신망을 통해, 상기 사용자 단말로부터 영상 및 추적 이미지를 입력받는 UX가 적용된 웹 페이지를 제공하는 웹 서비스 부를 포함할 수 있다.

[0019] 상기 웹 서비스 부는, 상기 웹 페이지를 통해, 상기 사용자 단말로부터 상기 추적 이미지에 대응하는 객체가 탐지 대상인지 또는 상기 탐지 대상을 제외한 나머지 타객체인지 결정하는 옵션정보를 더 입력 받을 수 있다.

[0020] 상기 프론트 엔드 서버는, 완료된 편집 프로젝트의 마스킹 영상을 스토리지 서버에 저장하고, 상기 스토리지 서버의 저장된 마스킹된 영상의 인터넷 접근 경로를 전자 메일 형태로 상기 사용자 단말에 전송하는 메일링 서비스 부를 포함할 수 있다.

[0021] 상기 프론트 엔드 서버는, 백그라운드에서 실행되고, 상기 스토리지 서버에 저장된 하나 이상의 변환된 영상에 대하여 일정기간이 경과하면 자동 삭제하는 스케쥴러를 더 포함할 수 있다.

[0022] 상기 제1 및 제2 학습 모델은, 각각 YOLOv5 모델 및 FaceNet 모델일 수 있다.

[0023] 또한, 전술한 과제를 해결하기 위해, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템에 의한 영상 편집 방법은, 시스템에 접속한 사용자 단말에 편집 프로젝트를 생성하는 웹 페이지를 제공하는 단계, 상기 사용자 단말로부터 편집 대상인 영상 및 상기 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받아 편집 프로젝트를 생성하는 단계, 상기 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 상기 탐지 대상에 대응하는 객체의 얼굴 영역을 탐지하는 단계, 학습된 제2 학습 모델을 이용한 분석 작업을 통해 탐지된 객체의 얼굴 영역을 인식하는 단계, 인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계; 및, 상기 편집 프로젝트를 종료하는 단계를 포함할 수 있다.

[0024] 상기 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 상기 탐지 대상에 대응하는 객체의 얼굴 영역을 탐지하는 단계 이전에, 제1 학습 모델을 이용하여 상기 영상에 포함되는 하나 이상의 객체의 얼굴 영역을 탐지하는 단계, 상기 탐지된 객체의 얼굴 영역을 추출하여 복수의 얼굴 영역을 포함하는 데이터 셋을 생성하는 단계, 상기 데이터 셋을 데이터 베이스에 저장하는 단계 및, 상기 데이터 셋을 이용하여 상기 제2 학습 모델을 학습시키는 단계를 포함할 수 있다.

[0025] 상기 인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계 이후, 완료된 편집 프로젝트의 마스킹 영상을 스토리지 서버에 저장하는 단계 및, 상기 스토리지 서버의 저장된 마스킹 영상의 인터넷 접근 경로를 전자 메일 형태로 상기 사용자 단말에 전송하는 단계를 포함할 수 있다.

[0026] 상기 편집 프로젝트를 종료하는 단계 이후, 백그라운드에서 실행되는 스케쥴러를 통해, 상기 스토리지 서버에 저장된 일정기간이 경과한 하나 이상의 변환된 영상을 자동 삭제하는 단계를 더 포함할 수 있다.

발명의 효과

[0027] 본 발명의 실시예에 따르면, 웹 페이지를 통해 사용자가 편집 대상인 영상 및 객체 인식을 위한 이미지를 업로드할 수 있는 웹 인터페이스를 제공하고, 백 엔드 서버에 구축된 학습 데이터 베이스 및 학습된 객체 탐지 모델을 통해 영상 내 이미지에 대응하는 객체를 인식하여 자동으로 마스킹 처리하여 회신함으로써, 일반 사용자들도 손쉽게 영상을 편집할 수 있는 효과가 있다.

[0028] 또한, 본 발명의 실시예에 따르면, 로우 데이터에 대하여 객체의 탐지 및 인식을 각각 별도의 알고리즘을 통해 수행하여 라벨링된 다수의 데이터 셋을 획득하고, 이에 기반한 학습을 위한 데이터 베이스를 구축하여 이를 통해 학습된 모델을 통해 영상 내 객체를 탐지함으로써, 보다 정확도가 높은 얼굴 탐지 결과를 제공할 수 있는 효과가 있다.

도면의 간단한 설명

[0029] 도 1은 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 전체 구조를 개략적으로 나타낸 도면이다.

도 2는 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 프론트 엔드 서버의 구조를 나타낸 도면이다.

도 3은 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 백 엔드 서버의 구조를 나타낸 도면이다.

도 4는 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템에 의한 자동 영상 편집 방법을 나타낸 도면이다.

도 5a 내지 도 5c는 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템이 제공하는 웹 페이지의 화면을 예시한 도면이다.

도 6은 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템에 의한 생성된 마스킹 영상을 예시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0030] 설명에 앞서, 명세서 전체에서 어떤 부분이 어떤 구성요소를 "구비" 또는 "포함" 한다고 할 때, 이는 특별히 반대되는 기재가 없는 한, 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부(Unit)", "...서버(Server)", "...시스템(System)" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어, 소프트웨어 또는, 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.

[0031] 또한, 본 명세서에서 "실시예"라는 용어는 예시, 사례 또는 도해의 역할을 하는 것을 의미하나, 발명의 대상은 그러한 예에 의해 제한되지 않는다. 또한, "포함하는", "구비하는", "갖는" 및 다른 유사한 용어가 사용되고 있으나, 청구범위에서 사용되는 경우 임의의 추가적인 또는 다른 구성요소를 배제하지 않는 개방적인 전환어(Transition word)로서, "포함하는(Comprising)"이라는 용어와 유사한 방식으로 포괄적으로 사용된다.

[0032] 본 명세서에 설명된 다양한 기법은 하드웨어 또는 소프트웨어와 함께 구현될 수 있거나, 적합한 경우에 이들 모두의 조합과 함께 구현될 수 있다. 본 명세서에 사용된 바와 같은 "...부(Unit)", "...서버(Server)" 및 "...시스템(System)" 등의 용어는 마찬가지로 컴퓨터 관련 엔티티(Entity), 즉 하드웨어, 하드웨어 및 소프트웨어의 조합, 소프트웨어 또는 실행 시의 소프트웨어와 등가로 취급할 수 있다. 또한, 본 발명의 시스템에서 구현하는 각 기능은 모듈단위의 프로그램으로 구성될 수 있고, 하나의 물리적 메모리에 기록되거나, 둘 이상의 메모리 및 기록매체 사이에 분산되어 기록될 수 있다.

[0033] 또한, 이하의 설명에서 본 발명의 "인공지능 기반 자동 영상 편집 시스템"을 가리키는 용어는, "영상 편집 시스템" 또는 "시스템"으로 병기될 수 있다.

[0034] 이하, 도면을 참조하여 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템 및 방법을 상세히 설명한다.

[0035] 도 1은 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 전체 구조를 개략적으로 나타낸 도면이다.

[0036] 도 1을 참조하면, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템(100)은, 사용자 단말(10)로부터 편집 대상인 영상 및 그 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받고, 편집 프로젝트를 생성하는 웹 페이지를 제공하는 프론트 엔드 서버(110), 탑재된 제1 및 제2 학습 모델을 이용하여, 편집 프로젝트에 포함되는 영상 내 탐지 대상인 객체의 얼굴 영역을 각각 탐지 및 인식하고, 인식된 객체의 얼굴 영역 또는 인식된 객체를 제외한 나머지 타객체의 얼굴 영역에 마스킹 처리를 수행하여 웹 페이지에 회신하는 백 엔드 서버(120) 및, 완료된 편집 프로젝트의 마스킹 영상을 저장하는 스토리지 서버(130)를 포함할 수 있다.

[0037] 프론트 엔드 서버(110)는 온라인을 통해 보유한 영상을 손쉽게 편집하고자 하는 사용자가 PC, 스마트폰 등의 사용자 단말을 이용하여 접속하고 저장된 영상에 대한 편집 프로그램을 생성할 수 있는 웹 페이지를 제공할 수 있다.

[0038] 이를 위해, 사용자 단말(10)에는 본 발명의 실시예에 따른 시스템(100)이 제공하는 웹 사이트에 접속하여 웹 페이지를 표시하고 영상 및 추적 이미지를 업로드할 수 있는 공지의 웹 브라우저(Web browser)가 탑재 및 실행될 수 있다.

[0039] 또한, 프론트 엔드 서버(110)는 웹 페이지를 통해 업로드된 영상 및 이미지를 백 엔드 서버(120)에 전송하여 학습된 복수의 학습 모델을 통해 분석 및 자동 편집을 수행하도록 요청하고, 완료된 영상을 스토리지 서버(130)에 저장함과 아울러 사용자 단말(10)에서 이용할 수 있도록 그 링크 주소를 전자메일 형태로 사용자 단말(10)에 송부할 수 있다.

[0040] 이를 위해, 프론트 엔드 서버(110)는 "React JS"의 "useState hook"을 이용하여 사용자 단말(10)로부터 영상 및 이미지 파일을 입력받고, "React JS"의 "axios" 통신 패키지를 이용하여 파일을 백 엔드 서버(120)에 전송하도록 구현될 수 있다.

[0041] 백 엔드 서버(120)는 복수의 학습 모델을 탑재하고 있고, 구축된 데이터 베이스에 저장된 다수의 학습 데이터를 이용하여 학습 모델을 훈련하고, 그 학습 모델을 이용하여 생성된 편집 프로젝트에 포함되는 영상에 대한 분석을 수행하여 영상내 객체를 탐지하고, 사용자 단말(10)로부터 입력된 인식을 위한 이미지 내 포함되는 객체 얼굴을 인식하여 탐지된 객체의 얼굴 영역과 추적 이미지가 서로 일치하는 지 판단할 수 있다.

[0042] 또한, 백 엔드 서버(120)는 분석된 영상에 대하여, 얼굴이 인식된 영상에 대하여, 해당 객체의 얼굴 영역상에 모자이크와 같은 마스킹 작업을 수행하고, 그 마스킹 영상을 프론트 엔드 서버(110)에 회신할 수 있다. 이에 따라, 백 엔드 서버(110)는 회신받은 편집된 영상을 스토리지 서버(130)에 저장하고, 그 링크 주소를 사용자 단말(10)에 송출하게 된다.

[0043] 스토리지 서버(130)는 고용량의 파일 저장 공간을 탑재할 수 있고, 프론트 엔드 서버(110)와 연계되어 그 요청에 따라 편집 프로젝트가 완료된 마스킹 영상을 저장할 수 있다.

[0044] 이러한 스토리지 서버(130)의 저장 공간내 파일은 URL로 접근할 수 있으며, 프론트 엔드 서버(110)는 스토리지 서버(130)에 마스킹 파일을 저장한 후, 그 마스킹 파일에 접근 가능한 URL을 해당 사용자의 전자메일에만 제한적으로 송부함으로써, 이후 사용자가 자신의 마스킹 파일을 다운로드할 수 있도록 지원할 수 있다.

[0045] 한편, 이러한 구조의 인공지능 기반 자동 영상 편집 시스템(100)은 공지의 클라우드 시스템(AWS 등)상에 구현될 수 있고, "Nginx"를 리버스 프록시로 설정하여 로드 밸런싱(load balancing)을 적용할 수 있다.

[0046] 또한, 시스템은, "AWS EC2" 인스턴스를 생성하여 "Nginx"를 비롯한, "React", "Redis", "Celery", "FastAPI" 인스턴스를 "Docker container"로 관리하는 형태로 구현될 수 있고, 모든 컨테이너가 "docker-compose"로 관리하거나 배포하기 용이하도록 구현될 수 있다.

[0047] 전술한 구조에 따라, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템은, 복수의 로우 데이터에서 탐지된 얼굴 이미지를 통해 학습 데이터에 대한 데이터 베이스를 구축 및 학습 모델을 훈련하고, 자신이 보유한 영상을 편집하고자 하는 사용자가 사용자 단말을 이용하여 온라인을 통해 영상을 업로드하면, 영상 편집에 대한 편집 프로젝트를 생성하여 학습된 학습 모델을 이용하여 해당 객체의 얼굴을 인식하고 옵션에 따라 마스킹 절차를 자동을 수행함으로써, 영상 편집에 대한 전문 지식이 없는 일반 사용자도 손쉽게 초상권 보호를 위한 영상을 제작할 수 있도록 한다.

[0048] 이하, 도면을 참조하여 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 프론트 엔드 서버를 상세히 설명한다.

[0049] 도 2는 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 프론트 엔드 서버의 구조를 나타낸 도면이다.

[0050] 도 2를 참조하면, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 프론트 엔드 서버(110)는, 정보통신망을 통해 사용자 단말로부터 영상 및 이미지를 입력받는 UX가 적용된 웹 페이지를 제공하는 웹 서비스 부(111), 완료된 편집 프로젝트의 마스킹 영상을 스토리지 서버에 저장하고, 스토리지 서버의 저장된 마스킹된 영상의 인터넷 접근 경로를 전자 메일 형태로 사용자 단말에 전송하는 메일링 서비스 부(113) 및, 백그라운드에서 실행되고, 스토리지 서버에 저장된 하나 이상의 변환된 영상에 대하여 일정기간이 경과하면 자동 삭제하는 스케쥴러(115)를 포함할 수 있다.

[0051] 웹 서비스 부(111)는 온라인 상에 본 발명의 실시예에 따른 웹 페이지를 공개하는 것으로, 공지의 웹 서버로 구현될 수 있고, 일반 사용자들의 자신이 소지한 영상 및 이미지를 업로드하여 편집 프로젝트를 생성할 수 있는 UX가 적용된 웹 페이지를 온라인을 통해 사용자 단말에 제공할 수 있다.

[0052] 또한, 웹 서비스 부(111)는, 온라인에 공개된 웹 페이지를 통해, 사용자 단말로부터 영상 내 마스킹 대상이 추적 이미지에 포함되는 탐지 대상인지 또는 탐지 대상을 제외한 나머지 타객체인지 결정하는 옵션정보를 더 입력받을 수 있다.

[0053] 메일링 서비스 부(113)는 편집 프로젝트가 완료된 건에 대하여 해당 마스킹 영상이 저장된 스토리지 서버의 링크 주소를 입력받을 수 있고, 편집 프로젝트의 종료시, 그 편집 프로젝트 생성시 사용자가 입력한 메일주소를 수신인으로 하여 링크 주소를 첨부하여 송부 할 수 있다.

[0054] 스케쥴러(115)는 편집 프로젝트가 완료되어 스토리지 서버(130)에 저장된 마스킹 영상에 대하여, 일정기간 이후 자동으로 삭제하는 기능을 제공할 수 있다. 본 발명의 실시예에 따르면, 편집 완료된 마스킹 영상은 자동으로 스토리지 서버(130)에 저장되며, 영상을 저장하기 위해 할당된 저장공간은 별도로 사용자에 의해 관리되는 것이 아님에 따라, 한정된 공간에 다수의 데이터가 장시간 저장될 경우, 서버의 운용, 관리에 어려움이 발생할 수 있다. 이에, 프론트 엔드 서버(110)는 스토리지 서버(130)에 저장되는 마스킹 영상들을 일정 시간이 경과되면 자동으로 삭제하는 스케쥴러(115)를 더 탑재할 수 있다. 바람직하게는, 마스킹 영상의 저장기간은 24시간 이하로 설정될 수 있다.

[0055] 이하, 도면을 참조하여 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 백 엔드 서버를 상세히 설명한다.

[0056] 도 3은 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 백 엔드 서버의 구조를 나타낸 도면이다.

[0057] 도 3을 참조하면, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템의 백 엔드 서버(120)는, 제1 학습 모델(M1)을 이용하여 편집 대상인 영상 내 하나 이상의 객체를 탐지하고, 탐지된 객체로부터 얼굴 영역을 추출하여 복수의 얼굴 영역을 포함하는 데이터 셋을 생성하는 데이터 추출부(121), 데이터 추출부(121)에 의해 생성된 데이터 셋이 저장되는 데이터 베이스(122), 데이터 셋에 라벨링을 수행하고, 라벨링된 데이터 셋을 이용하여 제2 학습 모델(M2)을 학습시키는 모델 학습부(123), 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 및 제2 학습 모델(M1, M2)을 이용한 분석 작업을 통해 추적 이미지에 등장하는 객체를 탐지 및 얼굴 영역을 인식하는 객체 인식부(124) 및, 탐지된 객체의 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하고 편집 프로젝트를 종료하는 마스킹 처리부(125)를 포함할 수 있다.

[0058] 데이터 추출부(121)는 공개된 로우 데이터를 이용하여 학습된 제1 학습 모델(M1)에 의해 영상 내 탐지되는 하나 이상의 객체로부터, 입력된 추적 이미지에 대응하는 얼굴 영역을 추출하여 제2 학습 모델의 학습을 위한 데이터 셋(data set)을 생성할 수 있다.

[0059] 전술한 데이터 셋은, 학습을 위한 공개된 로우 데이터가 일반적인 인물의 얼굴을 인식시키는 데 최적의 결과를 제공하지 못함에 따라, 얼굴 인식의 정확도를 최대화하기 위해 편집 대상인 영상에 대하여 제1 학습 모델(M1)에 의해 탐지된 추적 이미지에 대응하는 얼굴 영역을 구축한 학습 데이터이다.

[0060] 이를 위한 제1 학습 모델(M1)의 알고리즘으로는 "YOLOv5"가 이용될 수 있다.

[0061] 특히, 본 발명의 실시예에 따르면, 제1 학습 모델을 학습시키기 위한 로우 데이터(data)에는 일반적인 전면이 노출된 얼굴 이미지 뿐만 아니라, 마스크를 착용한 상태의 얼굴 이미지도 추가 이용될 수 있다.

[0062] 또한, 데이터 추출부(121)는 후술하는 제2 학습 모델(M2)의 훈련에 적용할 수 있도록 추출된 데이터 셋에 라벨링(Labeling)을 수행하여 학습 데이터로서 데이터 베이스(122)에 저장할 수 있다.

[0063] 데이터 베이스(122)는 데이터 추출부(121)에 의해 생성된 학습 데이터를 축적, 저장할 수 있다. 또한, 데이터 베이스(122)에는 데이터 셋 이외에도, 본 발명의 백 엔드 서버(120)를 운용, 관리하기 위한 다양한 설정값 등이 저장될 수 있다.

[0064] 모델 학습부(123)는 데이터 베이스(122)에 저장된 라벨링된 데이터 셋을 이용하여 얼굴 인식을 위한 제2 학습 모델(M2)을 학습시킬 수 있다.

[0065] 이러한 제2 학습 모델(M2)의 알고리즘으로는 "FaceNet"가 이용될 수 있다.

[0066] 객체 인식부(124)는 요청된 편집 대상인 영상 내에서, 학습된 제1 및 제2 학습 모델(M1, M2)을 이용하여 편집 프로젝트에 정의된 얼굴 이미지에 대응하는 객체의 얼굴 영역을 인식할 수 있다.

[0067] 상세하게는, 프론트 엔드 서버(110)에서 새로운 편집 프로젝트가 생성되어 편집 대상인 영상 및 이미지가 입력되면 백 엔드 서버(120)에서 제1 학습 모델(M1)을 통해 데이터셋을 도출하여 이를 통해 제2 학습 모델(M2)을 학습시키고, 객체 인식부(124)는 영상의 각 프레임에 등장하는 인물들 중 어느 하나와, 추적 이미지에 포함된 인물과 일치하는 하나의 얼굴 영역을 식별할 수 있다.

[0068] 먼저, 객체 인식부(124)는 제1 학습 모델(M1)을 이용하여 영상 내 하나 이상의 객체 및 추적 이미지에 대응하는 얼굴 영역을 탐지할 수 있고, 이어서 탐지된 하나 이상의 얼굴 영역을 학습 데이터로 이용하여 제2 학습 모델(M2)을 학습시킴으로써, 이후 얼굴 탐지 및 인식 과정에서 영상 내 등장하는 객체의 얼굴 영역의 인물이 편집 프로젝트의 얼굴 이미지와 동일 인물인지를 판단하게 된다.

[0069] 즉, 본 발명의 실시예에 따른 백 엔드 서버(120)는, 다양한 얼굴 이미지를 학습한 제1 학습 모델(M1)과, 제1 학습 모델(M1)을 이용하여 탐지된 얼굴 영역으로 데이터 셋을 구축하고, 그 데이터 셋을 이용하여 훈련된 제2 학습 모델(M2)을 준비하는 것을 특징으로 함과 더불어, 요청된 영상 내 프레임에서 제1 학습 모델(M1)을 이용하여 영상 내 얼굴 영역들을 탐지하고, 제2 학습 모델(M2)을 이용하여 탐지된 얼굴 영역들이 요청된 추적 이미지에 대응하는지 인식하는 것을 특징으로 한다.

[0070] 또한, 본 발명의 실시예에 따른 시스템은, 편집 대상은 정지영상이 아닌, 일정 기간의 재생 시간 동안 다양한 장면을 표시하고, 장면마다 함께 음원이 재생되는 동영상을 대상으로 하는 서비스를 제공하는 것임에 따라, 동기화 문제, 음원손실 문제 등이 발생할 수 있고, 객체 인식부(124)는 편집 중 전술한 문제가 발생하지 않도록 보호하는 하나 이상의 라이브러리를 포함할 수 있다.

[0071] 상세하게는, 객체 인식부(124)는 편집 프로젝트에 포함되는 영상에 대한 분석 작업 진행시 비동기 방식으로 처리하는 비동기 작업 큐 라이브러리(lb1), 분석 작업 진행 전, 편집 프로젝트에 포함되는 영상에서 음원을 추출하고, 분석 작업이 완료되면 마스킹된 영상과 추출된 음원을 결합하는 음원 추출 및 영상 변환 라이브러리(lb2, lb3)를 포함할 수 있다.

[0072] 영상 편집 절차는 영상의 길이에 따라 자동 편집 시간이 상이하며, 예상보다 긴 시간이 소요되는 작업이라 할 수 있다. 일례로서, 현재 학습 모델의 훈련 수준에서 재생시간이 30초 남짓 이라고 할 때, 5분 내지 12분 정도가 소요될 수 있고, 이에 따라 편집 작업을 동기방식으로 처리하는 경우 그 기간동안 사용자는 단지 지루하게 작업의 완료될 때까지 대기해야 하는 치명적인 문제가 발생할 수 있다.

[0073] 이러한 문제를 해결하기 위한 부가적 구성으로서, 비동기 작업 큐 라이브러리(lb1)는 편집 절차에서 동기로 처리되어야 하는 로직과 비동기로 처리되어야 하는 로직을 나누어 처리할 수 있도록 한다.

[0074] 전술한 비동기 작업 큐 라이브러리(lb1)로는, 'Celery' 라이브러리가 이용될 수 있으며, 'Redis'가 제공하는 'Masaage Broker'가 이용될 수 있다.

[0075] 또한, 일반적인 동영상에는 정지영상 뿐만 아니라, 소정의 음원도 포함될 수 있고, 영상 편집 과정에서 음원이 손실되는 문제가 발생할 수 있다. 이러한 문제를 해결하기 위해, 객체 인식부(124)는 영상에서 음원을 추출한 후 처리가 완료된 영상과 추출한 음원을 결합하는 로직을 수행하여 음원을 복원할 수 있다.

[0076] 이를 위한 음원 추출 라이브러리(lb2) 및 영상 변환 라이브러리(lb3)로는, 'Moviepy' 및 'ffmpeg'가 이용될 수 있다.

[0077] 마스킹 처리부(125)는 객체 인식부에 의해 탐지 및 인식된 객체 얼굴 영역 상에 블러 또는 저해상도화 등의 마스킹 작업을 수행하고, 해당 편집 프로젝트를 종료할 수 있다.

[0078] 마스킹 작업은, 영상 내 등장하는 인물의 얼굴에 모자이크 처리를 통해 육안으로 누구인지 식별할 수 없도록 하는 것으로, 블러, 저해상도화 등의 영상 처리 기법이 적용될 수 있다.

[0079] 특히, 본 발명의 실시예에 따르면, 마스킹 처리부(125)는 사용자 단말로부터 입력된 옵션정보에 따라, 영상 내에서 인식된 인물의 얼굴 영역을 대상으로 마스킹 처리를 수행하거나, 또는 영상 내 인식된 얼굴 영역을 제외한 나머지 모든 인물의 얼굴 영역을 대상으로 구분하여 마스킹 처리를 수행할 수 있다.

[0080] 이후, 마스킹 작업이 완료되면 마스킹 처리부(125)는 마스킹 영상을 프론트 엔드 서버(110)에 회신할 수 있고, 해당 편집 프로젝트를 종료 처리할 수 있다.

[0081] 이하, 도면을 참조하여 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템에 의한 자동 영상 편집 방법을 상세히 설명한다.

[0082] 도 4는 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템에 의한 자동 영상 편집 방법을 나타낸 도면이다. 이하의 설명에서, 각 단계별 실행주체는 별도의 기재가 없더라도 본 발명의 시스템 그 구성부가 된다.

[0083] 도 4를 참조하면, 본 발명의 인공지능 기반 자동 영상 편집 방법은, 온라인을 통해 서비스 관련 웹 페이지를 제공하는 인공지능 기반 자동 영상 편집 시스템에 의한 자동 영상 편집 방법으로서, 시스템에 접속한 사용자 단말에 편집 프로젝트를 생성하는 웹 페이지를 제공하는 단계(S100), 사용자 단말로부터 편집 대상인 영상 및 그 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받아 편집 프로젝트를 생성하는 단계(S110), 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 탐지 대상에 대응하는 객체를 탐지하는 단계(S120), 학습된 제2 학습 모델을 이용한 분석 작업을 통해 탐지된 객체의 얼굴 영역을 인식하는 단계(S130), 인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계(S140) 및, 편집 프로젝트를 종료하는 단계(S150)를 포함할 수 있다.

[0084] 전술한 편집 방법에 따른 서비스를 제공하기 앞서, 본 발명에 의한 영상 편집 서비스를 제공하기 위해서는 객체 탐지 및 인식을 위한 학습 모델의 학습 단계가 선행되어야 하며, 특히 객체 인식의 정확도를 높이기 위해 마스크 착용 얼굴을 포함하는 다양한 얼굴 이미지를 획득하여 학습을 진행하게 된다.

[0085] 이를 위한 선행 절차로서, 객체를 탐지하는 단계 이전에, 시스템은 공개된 로우 데이터를 입력받아 YOLOv5 등으로 구현된 제1 학습 모델을 학습시킬 수 있다.

[0086] 이후, 시스템에 접속한 사용자 단말에 편집 프로젝트를 생성하는 웹 페이지를 제공하는 단계(S100)로서, 프론트 엔드 서버가 인터넷을 통해 접속한 사용자 단말에 편집하고자 하는 영상을 업로드할 수 있는 UX가 적용된 웹 페이지를 제공한다.

[0087] 다음으로, 사용자 단말로부터 편집 대상인 영상 및 상기 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받아 편집 프로젝트를 생성하는 단계(S110)에서는, 사용자 단말이 웹 페이지를 통해 영상, 이미지 및 옵션정보를 입력함에 따라, 프론트 엔드 서버가 그 영상에 대한 편집 프로젝트를 생성하고, 입력 데이터를 포함하는 편집 프로젝트를 백 엔드 서버에 제공하게 된다.

[0088] 다음으로, 편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 탐지 대상에 대응하는 객체의 얼굴 영역을 탐지하는 단계(S120)에서는, 백 엔드 서버가 일반 노출 얼굴 이미지 및 마스크를 착용한 얼굴 이미지 등 다양한 얼굴 이미지를 통해 학습된 제1 학습 모델을 이용하여 영상 내 등장하는 추적 이미지에 대응하는 인물에 대한 얼굴 영역을 탐지한다. 이러한 S120 단계를 통해 탐지된 얼굴 영역은 데이터셋으로 이용되며, 후술하는 제2 학습 모델의 훈련에 이용되게 된다.

[0089] 다음으로, 학습된 제2 학습 모델을 이용한 분석 작업을 통해 탐지된 객체의 얼굴 영역을 인식하는 단계(S130)에서는, 백 엔드 서버가 S120 단계에서 탐지된 하나 이상의 얼굴 영역에 기반하여 생성한 데이터셋을 통해 제2 학습 모델을 학습시킴으로써 학습 모델을 준비하고, 영상의 각 프레임에 대하여 제1 학습 모델을 이용하여 얼굴 영역을 탐지하고, 탐지된 얼굴 영역 중에서 제2 학습 모델을 이용하여 편집 프로젝트에 정의된 추적 이미지와 동일 인물의 얼굴을 인식하게 된다.

[0090] 다음으로, 인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계(140)에서는, 백 엔드

서버가 편집 프로젝트에 정의된 사용자에 의해 입력된 옵션 정보에 따라 인식된 얼굴 영역 또는 그 얼굴 영역을 제외한 나머지 모든 얼굴 영역에 대한 모자이크 등의 마스킹 작업을 수행하게 된다.

[0091] 또한, 인식된 얼굴 영역 상에 블러 또는 저해상도화를 통한 마스킹을 수행하는 단계(S140) 이후, 완료된 마스킹 영상을 사용자에게 제공해야 하며, 이를 위해 스토리지 서버를 이용할 수 있다.

[0092] 상세하게는, 마스킹 영상 제공 절차로서, 프론트 엔드 서버는 완료된 편집 프로젝트의 마스킹 영상을 스토리지 서버에 저장하고, 스토리지 서버의 저장된 마스킹 영상의 인터넷 접근 경로를 전자 메일 형태로 사용자 단말에 전송하는 단계를 더 수행할 수 있다.

[0093] 그리고, 편집 프로젝트를 종료하는 단계(S150)로서, 편집 프로젝트를 통해 요청된 편집 작업이 완료됨에 따라 해당 편집 프로젝트는 종료되고, 시스템은 개인정보 보호정책 등에 따라 시스템에 입력된 영상 등은 삭제 처리하게 된다.

[0094] 또한, 전술한 편집 프로젝트를 종료하는 단계(S150) 이후, 스토리지에 축적되는 데이터를 정리하기 위한 절차로서, 시스템은 백그라운드에서 실행되는 스케줄러를 통해 스토리지 서버에 저장된 일정기간이 경과한 하나 이상의 변환된 영상을 자동 삭제하는 단계를 더 수행할 수 있다.

[0095] 이하, 실시예에 따른 인공지능 기반 자동 영상 편집 시스템이 제공하는 웹 페이지 화면 및 편집된 영상의 예시를 통해 본 발명의 기술적 사상을 상세히 설명한다.

[0096] 도 5a 내지 도 5c는 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템이 제공하는 웹 페이지의 화면을 예시한 도면이다.

[0097] 도 5a를 참조하면, 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템은 온라인을 통해 접속하는 사용자 단말에 서비스를 이용할 수 있는 UX가 적용된 웹 페이지를 제공할 수 있다. 사용자 단말의 접속시, 시스템은 초기화면(200)을 통해 사용자가 편리하게 웹 사이트를 이용할 수 있도록 최적화된 인터페이스를 제공할 수 있고, 사용자는 편집하고자 하는 영상을 웹 페이지에 업로드하여 작업을 요청할 수 있다.

[0098] 이를 위한 초기화면(200)은, 도 5a에 도시된 바와 같이, 사용자 단말에 저장된 영상을 첨부파일 형태로 웹 페이지에 업로드하는 영상 업로드 박스(201)와, 영상 내에 등장하는 객체를 탐지하기 위한 인물의 얼굴을 포함하는 추적 이미지를 웹 페이지에 업로드하는 이미지 업로드 박스(202)를 포함할 수 있다.

[0099] 또한, 초기화면(200)에는 사용자가 영상 내 등장하는 어느 하나의 특정 객체를 대상으로 하여 마스킹을 수행할 것인지, 아니면 그 특정 객체를 제외한 타객체들을 대상으로 마스킹을 수행할 것인지를 설정하는 옵션정보 입력란(203)이 포함될 수 있다. 이러한 옵션정보 입력란은 서로 배타적으로 동작하는 체크박스 형태로 구현될 수 있다.

[0100] 또한, 초기화면(200)은 사용자가 편집이 완료된 마스킹 영상을 전자 메일을 통해 제공받을 수 있도록 메일 주소 입력란(204)을 더 포함할 수 있으며, 초기화면(200)의 상단에는 사용자가 언제든 원하는 페이지로 손쉽게 이동할 수 있도록 하나 이상의 메뉴를 포함하는 네비게이션 바(Navigation Bar)가 배치될 수 있다. 이러한 네비게이션 바는 'Home', 'About' 및 'How To' 등의 메뉴로 구성될 수 있다.

[0101] 도 5b를 참조하면, 초기화면(200)의 네비게이션 바에서 'How To' 메뉴를 선택했을 때 표시되는 'How To' 화면(210)으로서, 사용자 단말에 시스템 이용을 위한 각종 사용 방법을 제공할 수 있다.

[0102] 특히, 사용자는 초기화면에서 영상 내 탐지하고자 하는 얼굴 이미지를 표시된 예시화면(211)과 같이 업로드할 수 있음을 확인할 수 있다.

[0103] 또한, 'How To' 화면(210)은 초기화면에서 제공하는 영상 업로드, 얼굴 이미지 업로드 및 옵션 선택 기능 등에 대한 방법에 대한 설명을 제공할 수 있으며, 사용자는 해당 화면을 참조하여 서비스 이용방법을 용이하게 숙지할 수 있다.

[0104] 도 5c를 참조하면, 본 발명의 실시예에 따른 시스템을 통해 영상 편집 절차가 완료됨에 따라 제공되는 영상 다운로드 화면(220)을 예시하고 있으며, 이러한 영상 다운로드 화면(220)에서는, 마스킹 영상이 생성됨에 따라 편집 프로젝트가 종료되고 사용자가 스토리지 서버에 저장된 마스킹 영상을 다운로드할 수 있도록 링크 주소를 제공할 수 있다.

[0105] 특히, 본 발명의 실시예에 따르면 시스템은 마스킹 영상이 생성되고 일정시간, 일례로서 24 시간이 경과되면 스

토리지 서버의 저장공간을 관리하기 위해 저장된 마스킹 영상을 자동으로 삭제 처리 함에 따라, 사용자에게 영상 다운로드 화면을 통해 삭제 처리에 대한 안내 정보를 제공할 수 있다.

[0106] 아울러, 전술한 링크 주소는 프론트 엔드 서버의 메일링 서비스를 통해 편집 프로젝트 생성시 사용자가 입력한 메일 주소를 통해서도 제공될 수 있다.

[0107] 도 6은 본 발명의 실시예에 따른 인공지능 기반 자동 영상 편집 시스템에 의한 생성된 마스킹 영상을 예시한 도면이다.

[0108] 도 6을 참조하면, 본 발명의 실시예에 따른 시스템은 편집 대상인 영상 내 등장하는 인물 중, 특정 인물의 얼굴 영역을 자동으로 탐지 및 인식하여 마스킹 처리하는 것으로, 복수의 객체(obj1, obj2)가 등장하는 영상에 대하여(a), 제1 학습 모델에 의해 복수의 객체(obj1, obj2)의 얼굴 영역에서 추적 이미지에 해당하는 얼굴 영역을 탐지할 수 있고, 그 탐지된 얼굴 영역에 기반한 데이터 셋을 이용하여 제2 학습모델을 학습시킨 후, 영상의 각 프레임에서 제1 및 제2 학습 모델을 통해 얼굴 영역의 탐지를 수행 및 탐지된 얼굴 영역 중, 추적 이미지에 대응하는 해당 객체(obj1)의 얼굴 영역을 인식하게 된다.

[0109] 그리고, 시스템은 옵션정보에 따라, 인식 대상인 객체(obj1)를 제외한 나머지 객체(obj2)의 얼굴 영역을 모자이크 등으로 마스킹할 수 있다.

[0110] 상기한 설명에 많은 사항이 구체적으로 기재되어 있으나 이것은 발명의 범위를 한정하는 것이라기보다 바람직한 실시예의 예시로서 해석되어야 한다. 따라서, 발명은 설명된 실시예에 의하여 정할 것이 아니고 특허청구범위와 특허청구범위에 균등한 것에 의하여 정하여져야 한다.

부호의 설명

[0111] 10 : 사용자 단말 100 : 인공지능 기반 자동 영상 편집 시스템

110 : 프론트 엔드 서버 110 : 웹 서비스 부

113 : 메일링 서비스 부 115 : 스케줄러

120 : 백 엔드 서버 121 : 데이터 추출부

122 : 데이터 베이스 123 : 모델 학습부

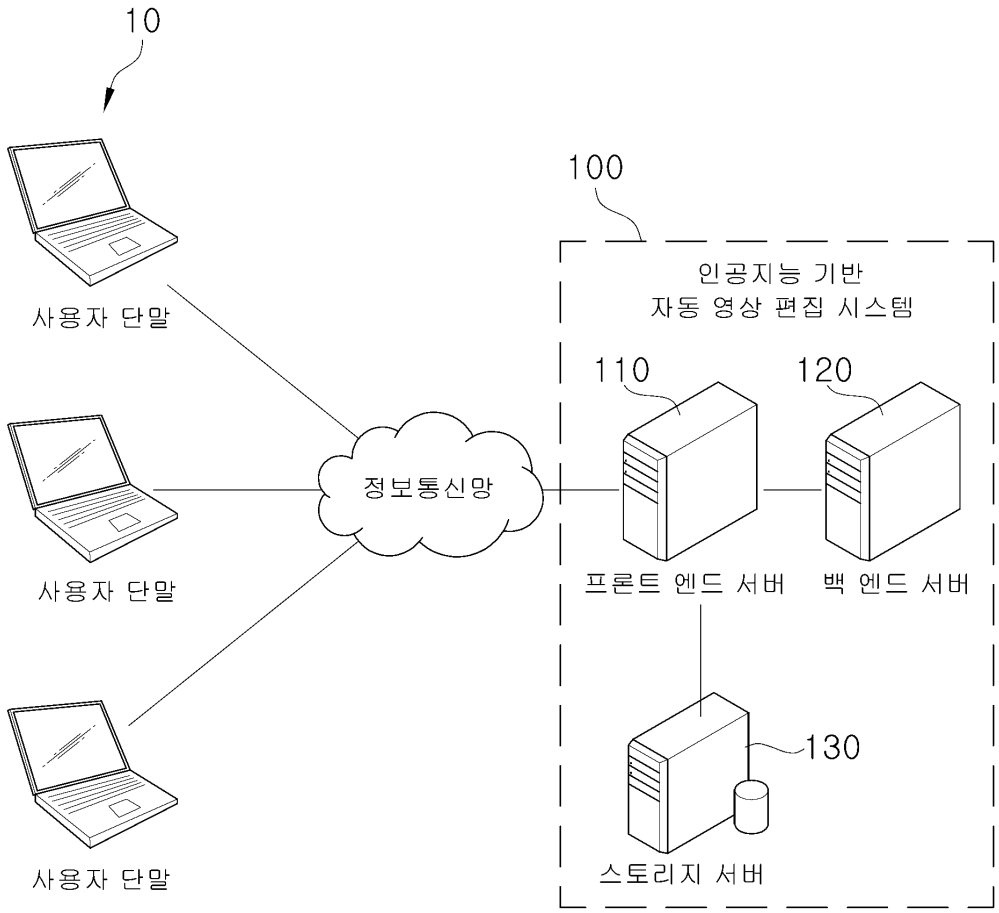
124 : 객체 인식부 125 : 마스킹 처리부

130 : 스토리지 서버 200 ~ 220 : 웹 페이지

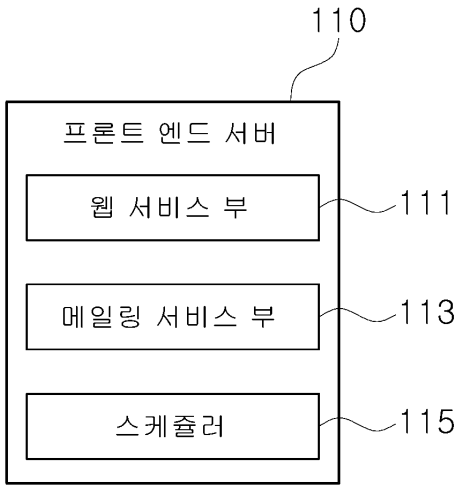
M1, M2 : 학습 모델 lb1 ~ lb3 : 라이브러리

도면

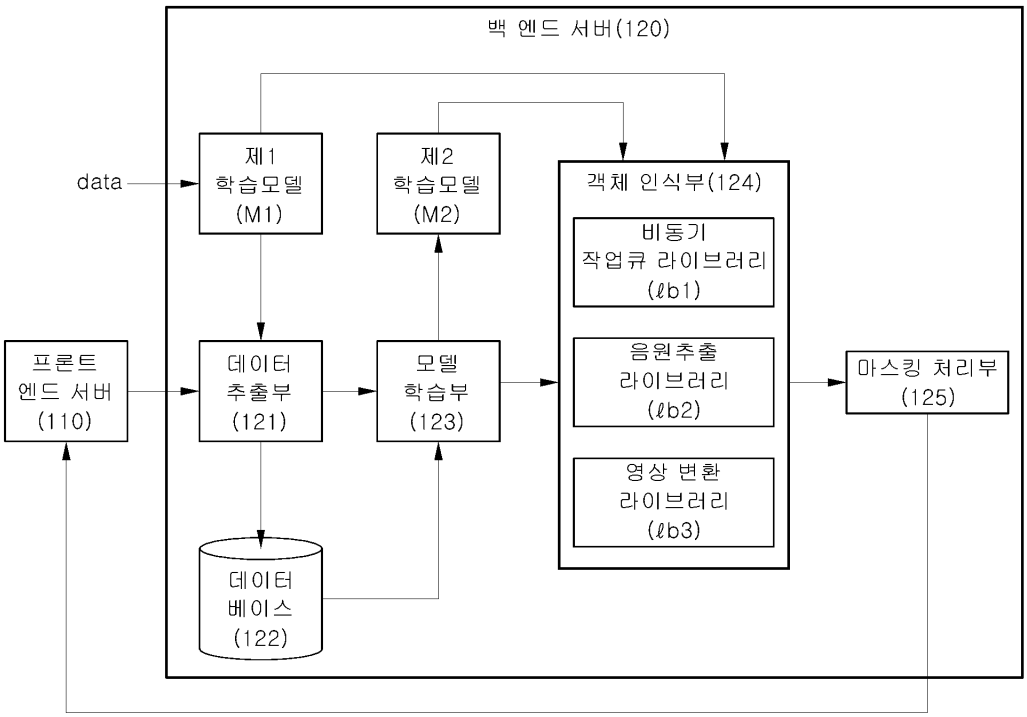
도면1


10
100
인공지능 기반
자동 영상 편집 시스템
110
120
사용자 단말
사용자 단말
사용자 단말
정보통신망
프론트 엔드 서버
백 엔드 서버
130
스토리지 서버

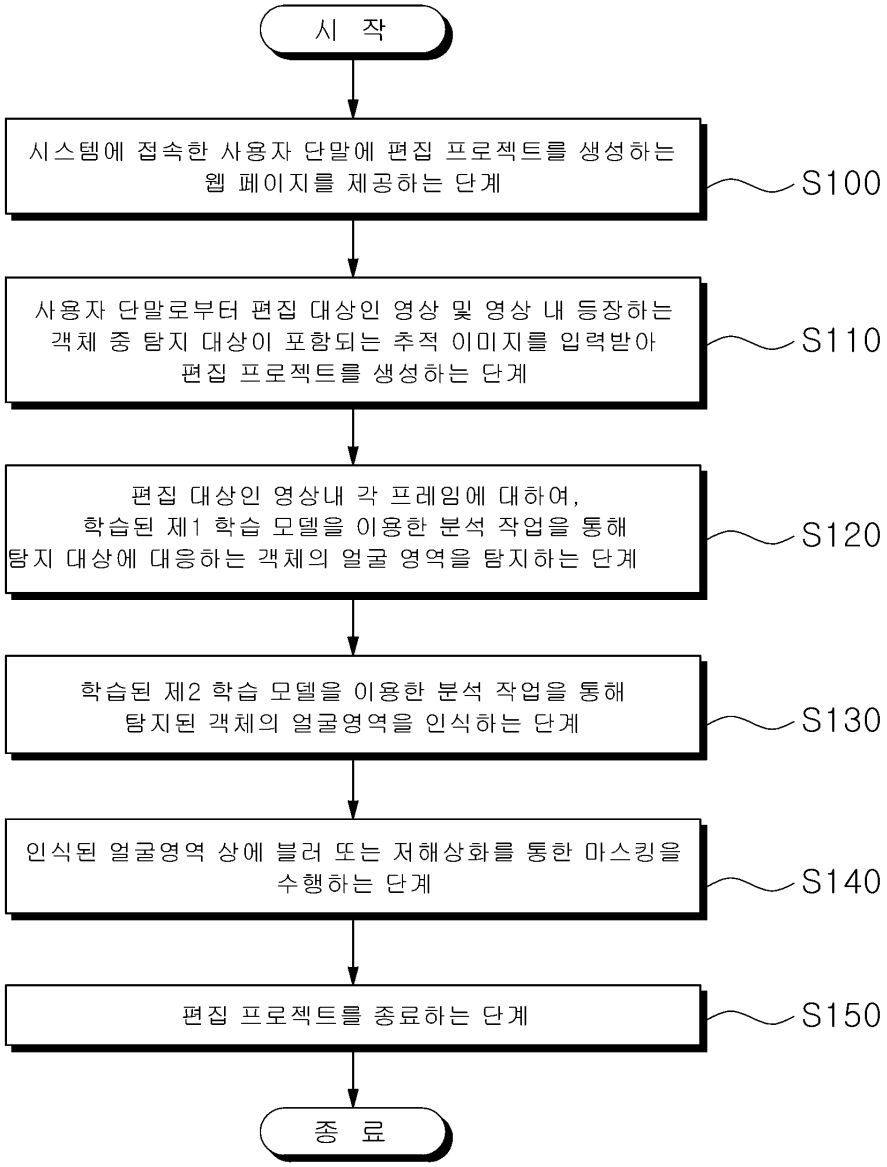
도면2


110
프론트 엔드 서버
웹 서비스 부
111
메일링 서비스 부
113
스케쥴러
115

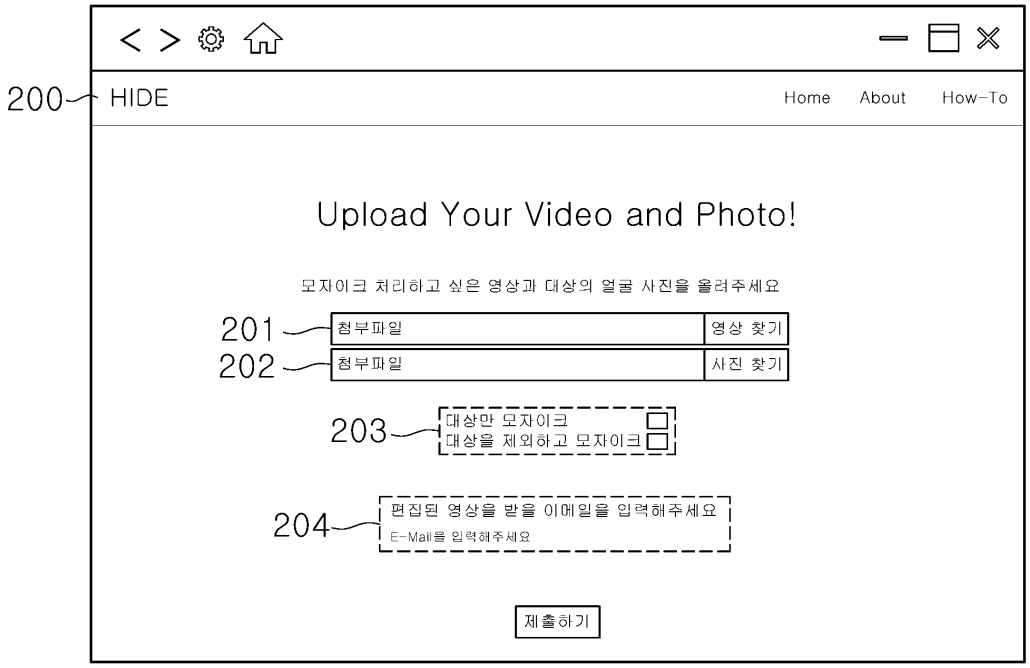
도면3


백 엔드 서버(120)
data
제1
학습모델
(M1)
제2
학습모델
(M2)
객체 인식부(124)
비동기
작업큐 라이브러리
(ℓb1)
음원추출
라이브러리
(ℓb2)
영상 변환
라이브러리
(ℓb3)
프론트
엔드 서버
(110)
데이터
추출부
(121)
모델
학습부
(123)
마스킹 처리부
(125)
데이터
베이스
(122)

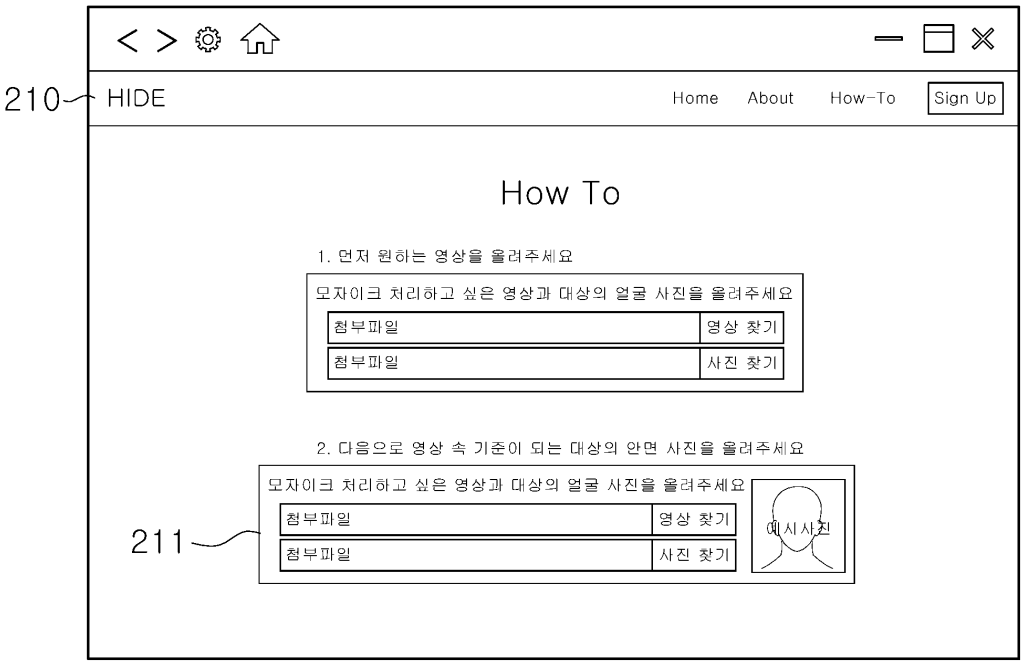
도면4


시 작
시스템에 접속한 사용자 단말에 편집 프로젝트를 생성하는 웹 페이지를 제공하는 단계
S100
사용자 단말로부터 편집 대상인 영상 및 영상 내 등장하는 객체 중 탐지 대상이 포함되는 추적 이미지를 입력받아 편집 프로젝트를 생성하는 단계
S110
편집 대상인 영상내 각 프레임에 대하여, 학습된 제1 학습 모델을 이용한 분석 작업을 통해 탐지 대상에 대응하는 객체의 얼굴 영역을 탐지하는 단계
S120
학습된 제2 학습 모델을 이용한 분석 작업을 통해 탐지된 객체의 얼굴영역을 인식하는 단계
S130
인식된 얼굴영역 상에 블러 또는 저해상화를 통한 마스킹을 수행하는 단계
S140
편집 프로젝트를 종료하는 단계
S150
종 료

도면5a


200
HIDE
Home
About
How-To
Upload Your Video and Photo!
모자이크 처리하고 싶은 영상과 대상의 얼굴 사진을 올려주세요
201
첨부파일
영상 찾기
202
첨부파일
사진 찾기
203
대상만 모자이크
대상을 제외하고 모자이크
204
편집된 영상을 받을 이메일을 입력해주세요
E-Mail을 입력해주세요
제출하기

도면5b


210
HIDE
Home
About
How-To
Sign Up
How To
1. 먼저 원하는 영상을 올려주세요
모자이크 처리하고 싶은 영상과 대상의 얼굴 사진을 올려주세요
첨부파일
영상 찾기
첨부파일
사진 찾기
2. 다음으로 영상 속 기준이 되는 대상의 안면 사진을 올려주세요
211
모자이크 처리하고 싶은 영상과 대상의 얼굴 사진을 올려주세요
첨부파일
영상 찾기
첨부파일
사진 찾기
예시사진

도면5c


220
[HIDE] 영상 변환이 완료되었습니다
COVER UP
영상은 24시간 보관 후 삭제될 예정이니,
시간 내 다운로드 해주시길 바랍니다.
아래 download 버튼을 클릭하시면 변환된 영상을 확인할 수 있습니다.
download

도면6


obj1
obj2
msk
obj1
obj2
(a)
(b)