(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0019405
(43) 공개일자 2020년02월24일

(51) 국제특허분류(Int. Cl.)
G06K 9/00 (2006.01) *G06K 9/62* (2006.01)
G06T 7/11 (2017.01)
(52) CPC특허분류
G06K 9/00221 (2013.01)
G06K 9/6276 (2013.01)
(21) 출원번호 10-2018-0094895
(22) 출원일자 2018년08월14일
심사청구일자 2018년08월14일

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
낭종호
서울특별시 강남구 언주로146길 18 (논현2동, 동현아파트) 1동 206호
김현곤
전라북도 전주시 완산구 오두정 1길 27-2
석혜경
부산광역시 동래구 충렬대로 314 (낙민동, 동래한양아파트) 7동 806호
(74) 대리인
이지연

전체 청구항 수 : 총 17 항

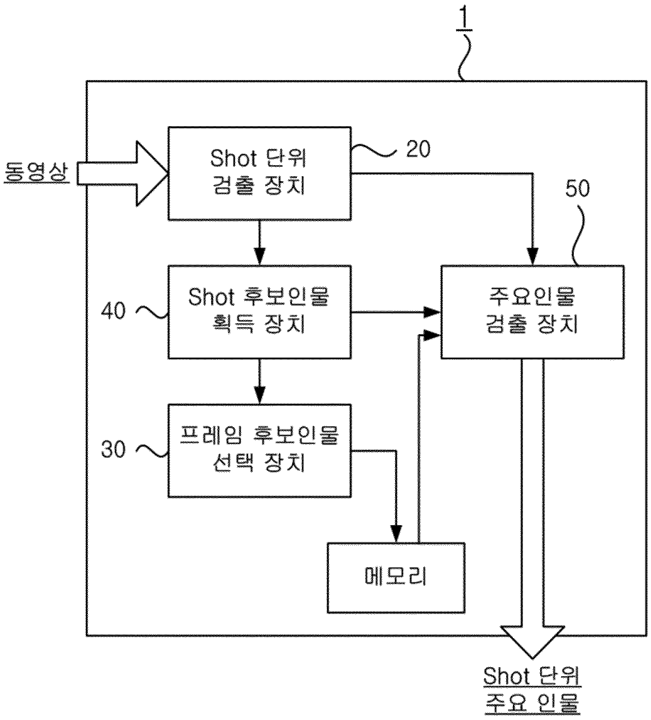
(54) 발명의 명칭 **샷 단위 주요 인물 검출 시스템 및 그 방법**

(57) 요 약

본 발명은 샷 단위 주요 인물 검출 시스템 및 방법에 관한 것이다. 상기 샷 단위 주요 인물 검출 시스템은, 하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 검출된 얼굴 영역의 크기 및 위치에 따른 가중값을 구하고, 상기 가중값을 이용하여 인식 결과의 확률값을 보정하고, 보정된 확률

(뒷면에 계속)

대 표 도 - 도4



값을 이용하여 프레임내의 후보 얼굴 영역들 및 각 후보 얼굴 영역에 대한 후보 인물들을 선택하여 제공하는 프레임 후보 인물 선택 장치; 상기 프레임 후보 인물 선택 장치를 이용하여, 샷을 구성하는 모든 프레임에 대하여 후보 얼굴 영역들 및 각 후보 얼굴 영역들에 대한 후보 인물들로 구성된 리스트를 획득하여 저장하는 샷 후보 인물 획득 장치; 및 샷을 구성하는 모든 프레임에서 획득된 후보 얼굴 영역 정보 및 후보 인물들에 대한 레이블들로 구성된 리스트를 이용하여, 샷의 주요 인물을 검출하는 주요 인물 검출 장치;를 구비하여, 연속되는 복수 개의 프레임으로 구성되는 샷의 주요 인물을 검출하는 것을 특징으로 한다.

(52) CPC특허분류
G06T 7/11 (2017.01)
G06T 2207/30201 (2013.01)

이 발명을 지원한 국가연구개발사업
과제고유번호 2017-0-00271
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 정보통신, 방송 연구개발 사업
연구과제명 아카이브 솔루션 및 콘텐츠 플랫폼 개발
기 여 율 1/1
주관기관 ㈜아이온커뮤니케이션즈
연구기간 2017.04.01 ~ 2018.12.31

명 세 서

청구범위

청구항 1

하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 검출된 얼굴 영역에 대한 정보 및 인식 결과를 이용하여 후보 얼굴 영역들 및 각 후보 얼굴 영역에 대한 후보 인물들을 선택하고, 상기 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 리스트를 제공하는 프레임 후보 인물 선택 장치;

상기 프레임 후보 인물 선택 장치를 이용하여, 샷을 구성하는 모든 프레임에 대하여 후보 얼굴 영역들 및 각 후보 얼굴 영역들에 대한 후보 인물들로 구성된 리스트를 획득하여 저장하는 샷 후보 인물 획득 장치; 및

샷을 구성하는 모든 프레임에서 획득된 후보 얼굴 영역 정보 및 후보 인물들에 대한 레이블들로 구성된 리스트를 이용하여, 샷의 주요 인물을 검출하는 주요 인물 검출 장치;

를 구비하여, 연속되는 복수 개의 프레임으로 구성되는 샷의 주요 인물을 검출하는 것을 특징으로 하는 샷 단위 주요 인물 검출 시스템.

청구항 2

제1항에 있어서, 상기 프레임 후보 인물 선택 장치는,

하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 모듈;

상기 얼굴 영역 검출 모듈에 의해 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 모듈;

주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 모듈;

상기 얼굴 영역 정보 검출 모듈에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 보정 모듈;

사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역으로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 필터 모듈; 및

상기 필터 모듈에 의해 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 출력하는 후보 인물 출력 모듈;

을 구비하여, 상기 프레임 후보 인물 선택 장치가 단일의 프레임에 대하여 후보 인물들을 선택하는 것을 특징으로 하는 샷 단위 주요 인물 검출 시스템.

청구항 3

제2항에 있어서,

상기 얼굴 영역 정보 검출 모듈이 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어진 것을 특징으로 하는 샷 단위 주요 인물 검출 시스템.

청구항 4

제2항에 있어서,

상기 보정 모듈의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고,

상기 제1 가중값은 각 얼굴 영역의 크기를 이용하여 계산된 가중값이며,

상기 제2 가중값은 각 얼굴 영역의 위치를 이용하여 계산된 가중값인 것을 특징으로 하는 샷 단위 주요 인물 검

출 시스템.

청구항 5

제4항에 있어서,

상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며,

상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것을 특징으로 하는 샷 단위 주요 인물 검출 시스템.

청구항 6

제1항에 있어서, 상기 주요 인물 검출 장치는

상기 샷 후보 인물 검출 장치에 의해 모든 프레임에서 선택된 각 후보 얼굴 영역 좌표들에 대한 영역 검출 횟수를 측정하고, 상기 영역 검출 횟수가 사전 설정된 기준 이상인 얼굴 영역들만을 주요 얼굴 영역으로 검출하는 주요 얼굴 영역 검출 모듈; 및

샷을 구성하는 모든 프레임에 대하여 상기 검출된 각 주요 얼굴 영역에 대한 후보 인물 검출 횟수를 측정하고, 상기 후보 인물 검출 횟수가 사전 설정된 기준 이상인 레이블을 샷내의 주요 인물로 선택하는 주요 인물 선택 모듈;

을 구비하는 것을 특징으로 하는 동영상에서의 샷 단위 주요 인물 검출 시스템.

청구항 7

제6항에 있어서, 상기 주요 얼굴 영역 검출 모듈은,

샷내의 연속되는 두 개의 프레임들에서 적어도 일부 영역이 서로 중첩되는 후보 얼굴 영역들이 있는 경우, 상기 두 개의 후보 얼굴 영역의 전체 면적에 대한 중첩 면적의 비율이 사전 설정된 기준값 이상이면 상기 두 개의 후보 얼굴 영역은 동일 얼굴 영역으로 설정하고, 해당 얼굴 영역에 대한 영역 검출 횟수를 증가시키는 것을 특징으로 하는 것을 특징으로 하는 동영상에서의 등장 인물 검출 장치.

청구항 8

하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 모듈;

상기 얼굴 영역 검출 모듈에 의해 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 모듈;

주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 모듈;

상기 얼굴 영역 정보 검출 모듈에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 보정 모듈;

사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역으로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 필터 모듈; 및

상기 필터 모듈에 의해 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 출력하는 후보 인물 출력 모듈;

을 구비하여, 단일의 프레임에 대하여 후보 인물들을 선택하는 것을 특징으로 하는 프레임 후보 인물 선택 장치.

청구항 9

제8항에 있어서,

상기 얼굴 영역 정보 검출 모듈이 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각

얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어진 것을 특징으로 하는 프레임 후보 인물 선택 장치,

청구항 10

제8항에 있어서,

상기 보정 모듈의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고,

상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며,

상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것을 특징으로 하는 프레임 후보 인물 선택 장치.

청구항 11

하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 검출된 얼굴 영역에 대한 정보 및 인식 결과를 이용하여 후보 얼굴 영역들 및 각 후보 얼굴 영역에 대한 후보 인물들을 선택하고, 상기 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 리스트를 제공하는 프레임 후보 인물 선택 단계;

상기 프레임 후보 인물 선택 단계를 이용하여, 샷을 구성하는 모든 프레임에 대한 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 획득하여 저장하는 샷 후보 인물 획득 단계; 및

샷을 구성하는 모든 프레임에서 획득된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 이용하여, 샷의 주요 인물을 검출하는 주요 인물 검출 단계;

를 구비하여, 연속되는 복수 개의 프레임으로 구성되는 샷의 주요 인물을 검출하는 것을 특징으로 하는 샷 단위 주요 인물 검출 방법.

청구항 12

제11항에 있어서, 상기 프레임 후보 인물 선택 단계는,

(a1) 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 단계;

(a2) 상기 얼굴 영역 검출 단계에서 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 단계;

(a3) 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 단계;

(a4) 상기 얼굴 영역 정보 검출 단계에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 단계;

(a5) 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역들로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 단계; 및

(a6) 상기 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 출력하는 단계;

를 구비하여, 상기 프레임 후보 인물 선택 단계에서는 단일의 프레임에 대하여 후보 인물들을 선택하는 것을 특징으로 하는 샷 단위 주요 인물 검출 방법.

청구항 13

제12항에 있어서,

상기 얼굴 영역 정보 검출 단계에서 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어진 것을 특징으로 하는 샷 단위 주요 인물 검출 방법.

청구항 14

제12항에 있어서,

상기 (a4) 단계의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고,

상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며,

상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것을 특징으로 하는 샷 단위 주요 인물 검출 방법.

청구항 15

제11항에 있어서, 상기 주요 인물 검출 단계는

(b1) 상기 샷 후보 인물 검출 단계에서 모든 프레임에서 선택된 각 후보 얼굴 영역들에 대한 영역 검출 횟수를 측정하고, 상기 영역 검출 횟수가 사전 설정된 기준 이상인 후보 얼굴 영역들만을 주요 얼굴 영역으로 검출하는 주요 얼굴 영역 검출 단계; 및

(b2) 샷을 구성하는 모든 프레임에 대하여 상기 검출된 각 주요 얼굴 영역에 대한 후보 인물 검출 횟수를 측정하고, 상기 후보 인물 검출 횟수가 사전 설정된 기준 이상인 레이블을 샷내의 주요 인물로 선택하는 주요 인물 선택 단계;

을 구비하는 것을 특징으로 하는 동영상에서의 샷 단위 주요 인물 검출 방법.

청구항 16

(a1) 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 단계;

(a2) 상기 얼굴 영역 검출 단계에서 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 단계;

(a3) 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 단계;

(a4) 상기 얼굴 영역 정보 검출 단계에서 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 단계;

(a5) 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역들로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 단계; 및

(a6) 상기 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들 을 출력하는 단계;

를 구비하여, 단일의 프레임에 대하여 후보 인물들을 선택하는 것을 특징으로 하는 프레임 후보 인물 선택 방법.

청구항 17

제16항에 있어서,

상기 (a2) 단계에서 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어진 것을 특징으로 하며,

상기 (a4) 단계에서의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고,

상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며,

상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것을 특징으로 하는 프레임 후보 인물 선택 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 샷 단위 주요 인물 검출 시스템 및 그 방법에 관한 것으로서, 더욱 구체적으로는 복수개의 연속되는 프레임들로 구성되는 샷의 주요 인물을 검출하되, 각 프레임에서 검출된 얼굴 영역의 위치와 얼굴 영역의 크기에 대한 정보를 이용하여 가중값을 설정하고, 상기 가중값을 적용하여 샷 내의 주요 인물을 보다 정확하게 검출하는 것을 특징으로 하는 샷 단위 주요 인물 검출 시스템 및 방법에 관한 것이다.

배 경 기 술

[0002] 동영상의 각 프레임에 있어서, 얼굴 영역을 검출하고, 검출된 얼굴 영역의 얼굴을 인식하는 다양한 기술들이 연구되고 있다. 종래의 얼굴 검출 라이브러리로는 OpenCV, Dlib 등이 있으며, CNN(Convolution Neural Network)을 이용한 얼굴 검출 방법도 다양하게 제안되고 있다.

[0003] 도 1은 종래의 기술에 따라, 동영상에서 각 프레임별로 얼굴 영역을 검출한 결과를 예시적으로 도시한 것이다.

[0004] 한편, 검출된 얼굴 영역을 기반으로 하여 CNN을 이용한 얼굴 인식 방법도 꾸준히 연구되고 있으며, 최근 FaceNet을 기반으로 한 모델이 매우 좋은 성능을 보여주고 있는데, 이러한 FaceNet을 기반으로 한 모델을 Fine Tuning하여 사용하고 있다. 이러한 종래의 얼굴 인식 방법은, 인식할 수 있는 사람이 학습에 사용된 사람만으로 이미 정해져 있으며, 학습되지 않은 사람이 들어오는 경우 학습된 사람들 중 가장 비슷하게 생긴 사람으로 인식하게 되는 문제가 발생한다.

[0005] 도 2는 종래의 기술에 따라, 하나의 프레임에서 검출된 얼굴 영역들이 표시된 것이다. 도 2를 참조하면, 해당 프레임에서 10개의 모든 얼굴 영역들이 검출되어 사각형으로 표시되어 있다. 이와 같이 검출된 얼굴 영역들의 각 얼굴들을 인식함에 있어서, 각 얼굴은 학습된 사람들 중 가장 높은 확률값을 갖는 사람으로 인식하게 된다.

[0006] 한편, 드라마 등과 같은 동영상에서 각 Shot 에 대한 정보를 제공하기 위한 방법의 일환으로서, 하나의 Shot 에 나온 주인공들의 얼굴만을 검출하는 방법들이 요구되고 있다. 동영상 콘텐츠에서 샷(Shot)은 일반적으로 카메라가 한 번의 촬영을 시작하여 멈추기 전까지를 의미하므로, 샷은 화면전환의 경계를 나타낼 수 있으며 비디오 콘텐츠를 나누는 기본적인 단위가 되기도 한다.

[0007] 따라서, 하나의 Shot을 구성하는 프레임들에 등장하는 주요 인물들을 검출하여 제공하기 위한 방법들이 제안되고 잇다. 하지만, 도 2에 도시된 바와 같이, 화면안에서 얼굴 영역을 검출할 수는 있으나 검출된 얼굴 영역에 대한 얼굴 인식은 이미 학습된 등장 인물들에 제한되어 인식될 수 있으며, 학습되지 아니한 등장 인물은 이미 학습된 등장 인물들 중 가장 유사한 인물로 인식되어 버리는 문제가 있다. 이와 관련하여, 드라마 등의 동영상의 등장 인물은 주인공과 같은 주요 인물과 단역 배우나 조연 배우등과 같은 엑스트라로 구성될 수 있으며, 얼굴 인식을 위한 사전 학습은 주로 주요 인물들만을 토대로 하여 이루어지게 된다. 따라서, 드라마 등과 같은 동영상에서의 얼굴 인식은 학습된 얼굴들 중 가장 비슷하게 생긴 등장 인물로 인식될 뿐이며, 해당 얼굴이 주인공과 같은 주요 인물인지, 엑스트라인지 판단하기 어려운 문제점이 있다.

[0008] 본 발명에서는 드라마 등과 같은 동영상의 하나의 shot에 포함된 주요 인물들만을 인식하여 제공하는 방안을 제안하고자 한다.

선행기술문헌

특허문헌

[0009] (특허문헌 0001) 한국등록특허공보 제 10-1341808호

(특허문헌 0002) 한국등록특허공보 제 10-1362381호

발명의 내용

해결하려는 과제

[0010] 전술한 문제점을 해결하기 위한 본 발명의 목적은 각 프레임에서 검출된 얼굴 영역의 크기 및 얼굴 영역의 위치 정보를 이용하여, 샷 내의 주요 인물을 보다 정확하게 검출할 수 있는 샷 단위 주요 인물 검출 시스템 및 방법을 제공하는 것이다.

[0011] 본 발명의 다른 목적은 각 프레임에서 검출된 얼굴 영역의 크기 및 얼굴 영역의 위치 정보를 이용하여, 프레임 내의 후보 인물을 정확하게 검출할 수 있는 프레임 후보 인물 선택 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0012] 전술한 기술적 과제를 달성하기 위한 본 발명의 제1 특징에 따른 샷 단위 주요 인물 검출 시스템은, 하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 인식된 얼굴들 중 사전 설정된 기준을 만족하는 후보 인물들을 선택하고, 해당 프레임에 대하여 상기 선택된 후보 인물들에 대한 레이블(Label)들 및 상기 후보 인물들이 위치한 후보 얼굴 영역 정보들을 제공하는 프레임 후보 인물 선택 장치; 상기 프레임 후보 인물 선택 장치를 이용하여, 샷을 구성하는 모든 프레임에 대한 후보 얼굴 영역 및 각 영역의 후보 인물들을 선택하여 저장하는 샷 후보 인물 획득 장치; 및 샷을 구성하는 모든 프레임에서 획득된 후보 얼굴 영역 및 후보 인물들에 대한 레이블들을 이용하여, 샷의 주요 인물을 검출하는 주요 인물 검출 장치;를 구비하여, 연속되는 복수 개의 프레임으로 구성되는 샷의 주요 인물을 검출한다.

[0013] 전술한 제1 특징에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 상기 프레임 후보 인물 선택 장치는, 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 모듈; 상기 얼굴 영역 검출 모듈에 의해 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 모듈; 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 모듈; 상기 얼굴 영역 정보 검출 모듈에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 보정 모듈; 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역으로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 필터 모듈; 및 상기 필터 모듈에 의해 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들의 리스트를 출력하는 후보 인물 출력 모듈;을 구비하여, 상기 프레임 후보 인물 선택 장치가 단일의 프레임에 대한 후보 얼굴 영역들 및 후보 인물들을 선택하는 것이 바람직하다.

[0014] 전술한 제1 특징에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 상기 프레임 후보 인물 선택 장치의 얼굴 영역 정보 검출 모듈이 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어진 것이 바람직하다.

[0015] 전술한 제1 특징에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 상기 프레임 후보 인물 선택 장치의 보정 모듈의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고, 상기 제1 가중값은 각 얼굴 영역의 크기를 이용하여 계산된 가중값이며, 상기 제2 가중값은 각 얼굴 영역의 위치를 이용하여 계산된 가중값인 것이 바람직하다.

[0016] 전술한 제1 특징에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 상기 주요 인물 검출 장치는, 상기 샷 후보 인물 검출 장치에 의해 모든 프레임에서 선택된 각 후보 얼굴 영역들에 대한 영역 검출 횟수를 측정하고, 상기 영역 검출 횟수가 사전 설정된 기준 이상인 후보 얼굴 영역들만을 주요 얼굴 영역으로 검출하는 주요 얼굴 영역 검출 모듈; 및 샷을 구성하는 모든 프레임에 대하여 상기 검출된 각 주요 얼굴 영역에 대한 후보 인물 검출 횟수를 측정하고, 상기 후보 인물 검출 횟수가 사전 설정된 기준 이상인 레이블을 샷내의 주요 인물로 선택하는 주요 인물 선택 모듈;을 구비하는 것이 바람직하다.

[0017] 전술한 제1 특징에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 상기 주요 얼굴 영역 검출 모듈은, 샷내의 연속되는 두 개의 프레임들에서 적어도 일부 영역이 서로 중첩되는 후보 얼굴 영역들이 있는 경우, 상기 두 개의 후보 얼굴 영역의 전체 면적에 대한 중첩 면적의 비율이 사전 설정된 기준값 이상이면 상기 두 개의 후보 얼굴 영역은 동일 얼굴 영역으로 설정하고, 해당 얼굴 영역에 대한 영역 검출 횟수를 증가시키는 것이 바람직하다.

[0018] 본 발명의 제2 특징에 따른 프레임 후보 인물 선택 장치는, 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 모듈; 상기 얼굴 영역 검출 모듈에 의해 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴

영역 정보 검출 모듈; 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 모듈; 상기 얼굴 영역 정보 검출 모듈에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 보정 모듈; 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역들로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 필터 모듈; 및 상기 필터 모듈에 의해 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들의 리스트를 출력하는 후보 인물 출력 모듈;을 구비하여, 단일의 프레임에 대하여 후보 얼굴 영역들 및 후보 인물들을 선택한다.

[0019] 전술한 제2 특징에 따른 프레임 후보 인물 선택 장치에 있어서, 상기 얼굴 영역 정보 검출 모듈이 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어지고,

[0020] 상기 보정 모듈의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고, 상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며, 상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것이 바람직하다.

[0021] 본 발명의 제3 특징에 따른 샷 단위 주요 인물 검출 방법은, 하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 인식된 얼굴들 중 사전 설정된 기준을 만족하는 후보 인물들을 선택하고, 해당 프레임에 대하여 후보 얼굴 영역들 및 후보 인물들에 대한 레이블(Label)들을 제공하는 프레임 후보 인물 선택 단계; 상기 프레임 후보 인물 선택 단계를 이용하여, 샷을 구성하는 모든 프레임에 대한 후보 얼굴 영역들 및 후보 인물들에 대한 정보들을 획득하여 저장하는 샷 후보 인물 획득 단계; 및 샷을 구성하는 모든 프레임에서 획득된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 이용하여, 샷의 주요 인물을 검출하는 주요 인물 검출 단계;를 구비하여, 연속되는 복수 개의 프레임으로 구성되는 샷의 주요 인물을 검출한다.

[0022] 전술한 제3 특징에 따른 샷 단위 주요 인물 검출 방법에 있어서, 상기 프레임 후보 인물 선택 단계는, (a1) 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 단계; (a2) 상기 얼굴 영역 검출 단계에서 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 단계; (a3) 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 단계; (a4) 상기 얼굴 영역 정보 검출 단계에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 단계; (a5) 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역으로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 단계; 및 (a6) 상기 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들의 리스트를 출력하는 단계;를 구비하여, 상기 프레임 후보 인물 선택 단계에서는 단일의 프레임에 대하여 후보 얼굴 영역들 및 후보 인물들을 선택하는 것이 바람직하다.

[0023] 전술한 제3 특징에 따른 샷 단위 주요 인물 검출 방법에 있어서, 상기 얼굴 영역 정보 검출 단계에서 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어진 것이 바람직하고,

[0024] 상기 (a4) 단계의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지고, 상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며, 상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것이 더욱 바람직하다.

[0025] 전술한 제3 특징에 따른 샷 단위 주요 인물 검출 방법에 있어서, 상기 주요 인물 검출 단계는, (b1) 상기 샷 후보 인물 검출 단계에서 모든 프레임에서 선택된 각 후보 얼굴 영역 좌표들에 대한 영역 검출 횟수를 측정하고, 상기 영역 검출 횟수가 사전 설정된 기준 이상인 후보 얼굴 영역들만을 주요 얼굴 영역으로 검출하는 주요 얼굴 영역 검출 단계; 및 (b2) 샷을 구성하는 모든 프레임에 대하여 상기 검출된 각 주요 얼굴 영역에 대한 후보 인

물 검출 횟수를 측정하고, 상기 후보 인물 검출 횟수가 사전 설정된 기준 이상인 레이블을 샷내의 주요 인물로 선택하는 주요 인물 선택 단계;를 구비하는 것이 바람직하다.

[0026] 본 발명의 제4 특징에 따른 프레임 후보 인물 선택 방법은, (a1) 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하는 얼굴 영역 검출 단계; (a2) 상기 얼굴 영역 검출 단계에서 검출된 얼굴 영역들에 대한 정보를 획득하는 얼굴 영역 정보 검출 단계; (a3) 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 얼굴 인식 단계; (a4) 상기 얼굴 영역 정보 검출 단계에서 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 단계; (a5) 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역들로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택하는 단계; 및 (a6) 상기 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들의 리스트를 출력하는 단계;를 구비하여, 단일의 프레임에 대하여 후보 얼굴 영역들 및 후보 인물들을 선택한다.

발명의 효과

[0027] 본 발명에 따른 주요 인물 검출 시스템은, 동영상의 Shot 단위의 등장 인물을 검출함에 있어서, 샷을 구성하는 모든 프레임에서 검출되는 모든 인물를 샷의 등장 인물로 검출하는 것이 아니라 주요 인물만을 샷의 등장 인물로 검출할 수 있게 된다.

[0028] 본 발명에 따른 주요 인물 검출 시스템은 얼굴 영역의 위치와 크기를 반영함으로써, shot내에서 주요 인물만을 검출하고 엑스트라는 제외시킬 수 있게 된다. 도 3은 본 발명에 따른 시스템에 의하여, 주요 인물을 정확하게 검출한 결과를 예시적으로 도시한 것이다. 종래의 얼굴 검출 방법에 따른 도 2의 결과와는 달리, 도 3을 참조하면 본 발명에 따른 시스템은 3명의 주요 인물만을 검출하고 그 외의 얼굴 영역들은 제거할 수 있게 된다.

[0029] 본 발명에 따른 주요 인물 검출 시스템에서 사용하는 프레임 후보 인물 선택 장치는, 주요 인물들의 얼굴만으로 사전 학습시킨 얼굴 인식기를 이용하여 얼굴 인식함에 따라, 모든 얼굴의 확률값을 계산하면 주요 인물의 얼굴은 큰 확률값이 가지게 되며 엑스트라의 얼굴은 낮은 확률값을 가지게 된다. 따라서, 얼굴 인식기의 확률값만으로도 등장 인물을 주요 인물과 엑스트라로 구분할 수 있을 것이다.

[0030] 하지만, 얼굴 인식기의 학습 데이터가 주요 인물들의 정면 얼굴이 대부분이므로, 동영상에서 검출된 얼굴이 측면 얼굴이나 눈을 감고 있는 등의 노이즈가 포함된 얼굴인 경우 확률값이 낮아지므로 분류가 어려워지는 문제가 발생하게 된다.

[0031] 이에 따라 본 발명에 따른 시스템에서는, 학습 데이터의 얼굴들이 주요 인물에 대한 정면 얼굴이 대부분을 차지하는 점을 고려하고, 동영상에서는 주 요인물이 나오는 위치가 화면의 중앙일 확률이 높고 주요 인물의 크기도 엑스트라에 비해 클 확률이 높다는 점을 고려하여, 동영상의 등장 인물을 주요 인물과 엑스트라로 구분하여 분류하게 된다. 따라서, 본 발명에 따른 시스템에서는 얼굴의 위치와 크기에 따른 가중값을 도출하고, 가중값과 얼굴 인식기에 의해 인식된 확률값을 조합한 결과를 Shot 단위별로 누적하여 주요 인물을 필터링함으로써, 프레임과 샷에서 검출되는 인물들 중 Shot 단위의 주요 인물을 분류하게 된다.

도면의 간단한 설명

[0032] 도 1은 종래의 기술에 따라 프레임들로부터 얼굴 검출하는 얼굴 검출 방법에 따른 결과들을 예시적으로 도시한 것이다.

도 2는 종래의 기술에 따라 하나의 프레임에서 다수 개의 얼굴 영역을 검출하고 이로부터 인식된 결과를 예시적으로 도시한 것이다.

도 3은 본 발명에 따른 시스템에 의하여, 주요 인물들만을 정확하게 검출한 결과를 예시적으로 도시한 것이다.

도 4는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템을 전체적으로 도시한 구성도이다.

도 5는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 구성을 개략적으로 도시한 블록도이다.

도 6은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 얼굴 인식 모듈(320)에 의한 결과를 도시한 그래프이다.

도 7은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역의 크기(a) 및 이에 따른 제1 가중값(b)을 구하여 예시적으로 도시한 도표이다.

도 8은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역에 대한 중심 x 좌표(a), 중심 y 좌표(b) 및 프레임의 중심점과 각 영역의 중심과의 거리(c)인 γ를 구하여 예시적으로 도시한 도표이다.

도 9는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역에 대한 α, β, γ를 예시적으로 도시한 도표이며, 도 10은 각 얼굴 영역에 대하여 계산된 보정 확률값들(out)을 예시적으로 도시한 도표이다.

도 11은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역에 보정 확률값들(out)이 사전 설정된 기준값(0.3)보다 큰 얼굴 영역들을 후보 얼굴 영역으로 선택한 상태를 예시적으로 도시한 프레임이다.

도 12는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 주요 인물 검출 장치(50)의 구성을 개략적으로 도시한 블록도이다.

도 13은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 주요 인물 검출 장치(50)의 주요 얼굴 영역 검출 모듈(500)이 영역 검출 횟수를 측정하기 위하여 사용되는 IoU를 설명하기 위하여 도시한 그림이다.

도 14는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 주요 인물이 최종적으로 선택된 출력을 예시적으로 도시한 것이다.

도 15는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 방법을 전체적으로 도시한 흐름도이다.

도 16은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 방법에 있어서, 프레임 후보 인물 선택 단계를 구체적으로 도시한 흐름도이다.

도 17은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 방법에 있어서, 주요 인물 검출 단계를 구체적으로 도시한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0033] 본 발명에 따른 샷 단위 주요 인물 검출 시스템은, 샷를 구성하는 각 프레임들로부터 검출된 얼굴 영역들에 대한 얼굴 영역 위치 및 얼굴 영역 크기에 대한 정보를 이용하여 가중값을 생성하고, 상기 가중값을 이용하여 각 얼굴 영역들에서 인식된 후보 인물들에 대한 확률값을 보정한 후, 보정된 확률값을 이용하여 후보 얼굴 영역들을 선택하고, 선택된 후보 얼굴 영역 및 해당 후보 얼굴 영역들에서 인식된 후보 인물들에 대한 보정 확률값들로 이루어진 리스트를 획득하고, 상기 획득된 얼굴 영역들로부터 주요 얼굴 영역을 다시 검출하고, 주요 얼굴 영역들로부터 주요 인물들을 획득하여 제공하는 것을 특징으로 한다.

[0034] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템 및 방법에 대하여 구체적으로 설명한다.

[0035] < 샷 단위 주요 인물 검출 시스템 >

[0036] 도 4는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템을 전체적으로 도시한 구성도이다. 도 4를 참조하면, 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템(1)은 샷 단위 검출 장치(20), 프레임 후보 인물 선택 장치(30), 샷 후보 인물 획득 장치(40), 주요 인물 검출 장치(50)를 구비하여, 샷 단위로 주요 인물들을 검출하여 제공하는 것을 특징으로 한다.

[0037] 본 명세서에 있어서, '주인공'은 특정 동영상에서 지속적으로 출연하는 인물을 의미하며, '엑스트라'는 주인공의 반의어로서 특정 동영상에서 간헐적으로 출연하는 인물을 의미하며, '후보 인물'은 각 프레임에서 사전 설정된 기준을 만족하는 인물들을 의미하며, '주요 인물'은 샷를 구성하는 모든 프레임에서의 상기 후보 인물들 중 사전 설정된 기준을 만족하는 인물들을 의미한다. 본 발명에 따른 시스템에 의해 샷 단위에서 선택된 주요 인물

이 주인공과 일치될 때, 본 발명에 따른 시스템의 성능이 가장 우수하다고 볼 수 있다.

[0038] 상기 샷 단위 검출 장치(20)는 입력된 동영상에 대하여, 샷 단위를 검출하고, 각 샷 단위를 구성하는 복수 개의 프레임들을 획득한다. 상기 샷 단위 검출 장치는 영상 처리 분야에서 널리 알려진 공지 기술들이므로, 본 명세서에서는 이에 대한 구체적인 설명은 생략한다.

[0039] 상기 프레임 후보 인물 선택 장치(30)는 하나의 프레임에서의 후보 인물들을 검출하는 장치로서, 좀 더 구체적으로는 하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 인식된 확률값을 얼굴 영역의 위치 및 크기를 이용하여 보정하고, 보정된 확률값을 이용하여 후보 얼굴 영역들 및 후보 얼굴 영역에서 인식된 후보 인물들에 대한 레이블들을 제공하게 된다. 상기 샷 후보 인물 획득 장치(40)는, 하나의 샷을 구성하는 모든 프레임들에 대하여, 상기 프레임 후보 인물 선택 장치를 이용하여 각 프레임의 후보 얼굴 영역들 및 후보 인물들을 획득하여 저장 장치에 저장한다. 상기 주요 인물 검출 장치(50)는, 샷을 구성하는 모든 프레임에서 획득된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 이용하여, 샷의 주요 인물을 검출한다.

[0040] 이하, 본 발명에 따른 시스템을 구성하는 각 구성 요소들에 대하여 구체적으로 설명한다.

[0041] < 프레임 후보 인물 선택 장치 >

[0042] 도 5는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 구성을 개략적으로 도시한 블록도이다.

[0043] 도 5를 참조하면, 상기 프레임 후보 인물 선택 장치(30)는, 얼굴 영역 검출 모듈(300), 얼굴 영역 정보 검출 모듈(310), 얼굴 인식 모듈(320), 보정 모듈(330), 필터 모듈(340) 및 후보 인물 출력 모듈(350)을 구비하여, 상기 프레임 후보 인물 선택 장치가 단일의 프레임에 대하여 후보 인물들을 선택하는 것을 특징으로 한다.

[0044] 상기 얼굴 영역 검출 모듈(300)은, 하나의 프레임에 존재하는 모든 얼굴 영역을 검출하고, 검출된 얼굴 영역에 대한 픽셀 좌표값들을 제공한다. 상기 검출된 얼굴 영역에 대한 픽셀 좌표값은 [x, y, x+w, y+h] 로 구성될 수 있다. 얼굴 영역 검출 방법은 이미 널리 알려진 공지의 기술로서, OpenCV, Dlib 와 같은 얼굴 검출 라이브러리들 뿐만 아니라, CNN(Convolution Neural Network)을 이용한 얼굴 검출 방법들이 다양하게 존재하므로, 이에 대하여 본 명세서에서는 구체적으로 설명하지 아니한다.

[0045] 상기 얼굴 영역 정보 검출 모듈(310)은, 상기 얼굴 영역 검출 모듈에 의해 검출된 얼굴 영역들에 대한 정보를 획득하며, 여기서 얼굴 영역들에 대한 정보는 각 검출된 얼굴 영역에 대한 픽셀 좌표값들로부터 획득되는 정보들로서, 각 얼굴 영역의 위치 정보 및 각 얼굴 영역의 크기 정보를 포함한다.

[0046] 상기 얼굴 인식 모듈(320)은, 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값을 제공하는 것이 바람직하다. 도 6은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 얼굴 인식 모듈(320)에 의한 결과를 도시한 그래프이다. 도 6을 참조하면, 상기 얼굴 인식 모듈(320)은, 얼굴 영역 검출 모듈로부터 얻은 각 얼굴 영역들에 대하여 인식된 주요 인물들에 대한 가장 큰 확률값들이 나열되어 있다.

[0047] 상기 보정 모듈(330)은, 상기 얼굴 영역 정보 검출 모듈에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고, 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득하는 것이 바람직하다.

[0048] 상기 얼굴 영역 정보 검출 모듈이 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보 중 하나 또는 둘로 이루어지며, 특히 크기 정보와 위치 정보를 모두 반영하는 것이 더욱 바람직하다. 또한, 상기 보정 모듈의 가중값은 제1 가중값 및 제2 가중값 중 하나로 이루어지거나 제1 및 제2 가중값을 모두 반영하여 이루어지는 바람직하다.

[0049] 상기 제1 가중값은 각 얼굴 영역의 크기를 이용하여 계산된 가중값으로서, 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이다. 상기 제2 가중값은 각 얼굴 영역의 위치를 이용하여 계산된 가중값으로서, 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값

이다.

[0050] 본 발명에 따른 시스템에서의 보정 모듈에 의해, 보정 확률값을 구하는 일 실시 형태를 예시적으로 설명한다. 먼저, 검출된 얼굴 영역들의 좌표값은 [[x11,y11,x12,y12],[x21, y21,x22,y22],...] 이며, 각 얼굴 영역에서의 인식된 확률값은 [α1, α2, ...] 으로 설정한다.

[0051] 이 경우, 각 얼굴 영역의 크기(size)는 각 얼굴 영역들의 좌표값으로부터 계산할 수 있으며, 크기 리스트(size list)에 상기 계산된 얼굴 영역들의 크기를 저장한다. 제1 가중값(β)는 수학식 1에 따라 가장 큰 얼굴 영역의 크기(max(size))와 각 얼굴 영역의 크기(size)의 비율로 결정한다.

수학식 1

[0052] $$\beta = size / max(size)$$

[0053] 도 7은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역의 크기(a) 및 이에 따른 제1 가중값(b)을 구하여 예시적으로 도시한 도표이다.

[0054] 각 얼굴 영역의 위치는 얼굴 영역의 좌표값들의 중간 위치[center_x, center_y] 로 설정하고, 각 얼굴 영역의 중간 위치들을 중간 위치 리스트에 저장한다. 다음, 수학식 2에 따라 각 얼굴 영역의 중간 위치와 프레임의 중간 사이의 거리(γ)를 계산하고, 수학식 3에 따라 프레임 전체의 대각선 길이(C)를 계산한다. 제2 가중값은 1-γ/C 로 설정할 수 있다.

[0055] 도 8은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역에 대한 중심 x 좌표(a), 중심 y 좌표(b) 및 프레임의 중심점과 각 영역의 중심과의 거리(c)인 γ를 구하여 예시적으로 도시한 도표이다.

[0056] 보정 확률값(k)는 수학식 4에 따라 구할 수 있다. 여기서, α, β, (1-γ/C)는 각각 각 얼굴 영역에 대한 확률값, 제1 가중값 및 제2 가중값이다.

수학식 2

[0057] $$\gamma = \sqrt{\left(\left(center_x - \frac{width}{2}\right)^2 + \left(center_y - \frac{height}{2}\right)^2\right)}$$

수학식 3

[0058] $$C = \sqrt{width^2 + height^2}$$

수학식 4

[0059] $$k = \alpha * \beta * (1 - \gamma / C)$$

[0060] 도 9는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역에 대한 α, β, γ를 예시적으로 도시한 도표이며, 도 10은 각 얼굴 영역에 대하여 계산된 보정 확률값들(out)을 예시적으로 도시한 도표이다.

[0061] 상기 필터 모듈(340)은, 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역들로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택한다.

[0062] 도 11은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 프레임 후보 인물 선택 장치(30)의 보정 모듈에서 각 얼굴 영역에 보정 확률값들(out)이 사전 설정된 기준값(0.3)보다 큰 얼굴 영역들을 후보 얼굴 영역들로 선택한 상태를 예시적으로 도시한 프레임이다. 도 11을 참조하면, 사전 설정된 기준값이 0.3 인 경우, 도 10에 도시된 프레임의 모든 얼굴 영역들에서 구한 보정 확률값이 0.3 이상인 영역만을 후보 얼굴 영역들로 선택하고, 이에 따라 3개의 후보 얼굴 영역들이 선택됨을 알 수 있다.

[0063] 상기 후보 인물 출력 모듈(350)은, 상기 필터 모듈에 의해 선택된 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들의 리스트를 출력하는 것이 바람직하다. 도 10 및 도 11을 참조하면, 사전 설정된 기준값을 0.3으로 설정한 경우, 보정 확률값이 0.3보다 큰 후보 얼굴 영역은 [554,104,654,247], [718,144,830,293], [400,305,523,492]이므로, 이들에 대한 얼굴 영역 좌표와 각 얼굴 영역에 대한 후보 인물의 레이블(Label)들의 리스트를 출력한다.

[0064] 전술한 본 발명에 따른 상기 프레임 후보 인물 선택 장치는 전술한 구성들에 의해 단일의 프레임에 대하여 후보 얼굴 영역들 및 후보 인물들을 선택하여 출력하게 된다.

[0065] < 주요 인물 검출 장치 >

[0066] 이하, 본 발명의 시스템에 따른 주요 인물 검출 장치(50)에 대하여 보다 구체적으로 설명한다.

[0067] 상기 샷 후보 인물 획득 장치(40)는, 하나의 샷을 구성하는 모든 프레임들에 대하여, 상기 프레임 후보 인물 선택 장치를 이용하여 샷을 구성하는 모든 프레임의 후보 얼굴 영역들 및 후보 인물들을 획득하여 저장 장치에 저장한다. 그리고, 상기 주요 인물 검출 장치(50)는, 샷 후보 인물 획득 장치 및 프레임 후보 인물 선택 장치를 통해 저장 장치에 저장된 샷를 구성하는 모든 프레임의 후보 얼굴 영역들 및 후보 인물들에 대한 레이블들을 이용하여, 샷의 주요 인물을 검출한다. 도 12는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 주요 인물 검출 장치(50)의 구성을 개략적으로 도시한 블록도이다.

[0068] 도 12를 참조하면, 상기 주요 인물 검출 장치(50)는, 주요 얼굴 영역 검출 모듈(500) 및 주요 인물 선택 모듈(510)을 구비한다.

[0069] 상기 주요 영역 검출 모듈(500)은, 상기 샷 후보 인물 검출 장치에 의해 모든 프레임에서 선택된 후보 얼굴 영역들에 대한 영역 검출 횟수를 측정하고, 상기 영역 검출 횟수가 사전 설정된 기준 이상인 얼굴 영역들만을 주요 얼굴 영역으로 검출한다.

[0070] 상기 주요 영역 검출 모듈의 동작을 보다 구체적으로 설명하면, 먼저 샷 내의 프레임들을 비교하고 IoU가 사전 설정된 값 이상인 영역의 결과는 통합시키게 된다.

[0071] 도 13은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 주요 인물 검출 장치(50)의 주요 얼굴 영역 검출 모듈(500)이 영역 검출 횟수를 측정하기 위하여 사용되는 IoU를 설명하기 위하여 도시한 그림이다. 도 13에 도시된 바와 같이, 샷내의 연속되는 두 개의 프레임들에서 적어도 일부 영역이 서로 중첩되는 후보 얼굴 영역들이 있는 경우, 상기 두 개의 후보 얼굴 영역의 전체 면적(Area of Union)에 대한 중첩 면적(Area of Overlap)의 비율(IoU)이 사전 설정된 기준값 이상이면 상기 두 개의 후보 얼굴 영역은 동일 얼굴 영역으로 설정하고, 해당 얼굴 영역에 대한 영역 검출 횟수를 증가시키게 된다.

[0072] 상기 주요 영역 검출 모듈은 영역 검출 횟수가 사전 설정된 기준보다 작은 후보 얼굴 영역은, 프레임 후보 얼굴 선택 장치에 의한 얼굴 영역의 크기와 얼굴 영역의 위치를 기반으로 한 분석 알고리즘을 통해 필터링되지 않은 소수의 영역에 해당하므로, 이는 샷 내의 노이즈(noise)로 판단하고 이를 필터링하게 된다.

[0073] 상기 주요 인물 선택 모듈(510)은, 샷을 구성하는 모든 프레임에 대하여 상기 검출된 각 주요 얼굴 영역에 대한 후보 인물 검출 횟수를 측정하고, 상기 후보 인물 검출 횟수가 사전 설정된 기준 이상인 레이블을 샷내의 주요 인물로 선택한다. 따라서, 본 발명에 따른 주요 인물 선택 모듈은 같은 주요 얼굴 영역에서 가장 많이 등장한 인물의 label을 선택하여 주요 인물로 출력하게 된다.

[0074] 도 14는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템에 있어서, 최종적으로 선택된 출력을 예시적으로 도시한 것이다. 도 14를 참조하면, 4개의 프레임에 대하여, (a)는 후보 얼굴 영역들 및 이들에

대한 후보 인물 Label을 검출한 리스트 결과이며, (b)는 (a)의 결과로부터 각 얼굴 영역 검출 횟수와 후보 인물의 검출 횟수를 측정한 리스트 결과이며, (c)는 (b)의 결과로부터 최종적으로 주요 인물들을 검출한 결과이다.

[0075] < 샷 단위 주요 인물 검출 방법 >

[0076] 전술한 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 시스템은 프로그램 등에 의해 샷 단위 주요 인물 검출 방법으로도 구현될 수 있다. 이하, 도면을 참조하여 본 발명에 따라 구현될 수 있는 샷 단위 주요 인물 검출 방법에 대하여 구체적으로 설명한다.

[0077] 도 15는 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 방법을 전체적으로 도시한 흐름도이다.

[0078] 도 15를 참조하면, 본 발명에 따른 샷 단위 주요 인물 검출 방법은, 먼저 입력된 동영상으로부터 Shot 단위를 검출한다(단계 600). 다음, shot 내의 첫번째 프레임을 불러와서(단계 610), 프레임 후보 인물 선택 단계를 수행하여 프레임내 후보 얼굴 영역들 및 후보 인물들을 선택한다(단계 620). 상기 프레임 후보 인물 선택 단계는, 하나의 프레임에 존재하는 모든 얼굴 영역들을 검출하고, 각 얼굴 영역의 얼굴들을 인식하고, 상기 인식된 확률값과 얼굴 영역의 위치/크기에 따른 가중값을 이용하여 후보 얼굴 영역들 및 이의 후보 인물들을 선택하고, 해당 프레임에 대하여 상기 선택된 후보 얼굴 영역 및 후보 인물들에 대한 레이블(Label)들을 제공하게 된다.

[0079] 다음, shot 내의 모든 프레임들을 순차적으로 불러와서, 후보 인물 선택 단계를 수행하여 프레임내 후보 얼굴 영역 및 후보 인물을 선택함으로써, shot 내의 모든 프레임에 대한 후보 얼굴 영역들 및 후보 인물들을 획득하게 된다(단계 630). 다음, 주요 인물 검출 단계를 수행하여 shot 내의 모든 프레임에 대한 후보 얼굴 영역들 및 후보 인물 레이블들의 리스트를 이용하여, shot의 주요 인물을 검출한다(단계 640).

[0080] 따라서, 본 발명에 따른 샷 단위 주요 인물 검출 방법은 연속되는 복수 개의 프레임으로 구성되는 샷의 주요 인물을 검출하는 것을 특징으로 한다.

[0081] 이하, 전술한 상기 프레임 후보 인물 선택 단계(단계 620)에 대하여 보다 구체적으로 설명한다.

[0082] 도 16은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 방법에 있어서, 프레임 후보 인물 선택 단계를 구체적으로 도시한 흐름도이다. 도 16을 참조하면, 상기 프레임 후보 인물 선택 단계는, 먼저 하나의 프레임에 존재하는 모든 얼굴 영역을 검출한다(단계 700).

[0083] 다음, 상기 얼굴 영역 검출 단계에서 검출된 얼굴 영역들에 대한 정보를 획득한다(단계 710). 상기 얼굴 영역 정보 검출 단계에서 획득하는 얼굴 영역들에 대한 정보는, 각 얼굴 영역에 대한 크기 정보 및 각 얼굴 영역에 대한 위치 정보들을 포함한다.

[0084] 다음, 주요 인물에 대한 얼굴 사진을 기초로 한 사전 학습에 의해 완성된 얼굴 인식기를 이용하여, 상기 검출된 얼굴 영역들에 대하여 가장 큰 확률값을 갖는 주요 인물로 인식하고, 각 얼굴 영역들에 대하여 인식된 주요 인물에 대한 레이블(Label) 및 그 확률값들로 구성된 리스트를 제공한다(단계 720).

[0085] 다음, 상기 얼굴 영역 정보 검출 단계에 의해 얻은 각 얼굴 영역에 대한 정보를 이용하여 해당 얼굴 영역에 대한 가중값을 획득하고(단계 730) , 각 얼굴 영역에 대하여 상기 얼굴 인식 모듈에 의해 얻은 확률값을 해당 얼굴 영역에 대한 가중값을 이용하여 보정하여, 각 얼굴 영역에서 인식된 주요 인물에 대한 보정 확률값을 획득한다(단계 732). 상기 가중값은 제1 가중값 및 제2 가중값 으로 이루어지고, 상기 제1 가중값은 각 얼굴 영역에 대한 크기와 가장 큰 얼굴 영역의 크기의 비율을 이용하여 계산된 가중값이며, 상기 제2 가중값은 프레임의 중심 좌표와 각 얼굴 영역의 중심 좌표의 이격 거리를 이용하여 계산된 가중값인 것이 바람직하다.

[0086] 다음, 사전 설정된 기준값보다 큰 보정 확률값을 갖는 얼굴 영역들을 후보 얼굴 영역으로 선택하고, 상기 선택된 후보 얼굴 영역들에 대하여 인식된 주요 인물들을 해당 프레임에 대한 후보 인물들로 선택한다(단계 740).

[0087] 다음, 상기 선택된 후보 얼굴 영역 정보 및 각 후보 얼굴 영역의 후보 인물들에 대한 레이블들에 대한 리스트를 출력한다(단계 750).

[0088] 전술한 과정을 통해, 상기 프레임 후보 인물 선택 단계에서는 단일의 프레임에 대하여 선택된 후보 얼굴 영역 정보 및 후보 인물들로 구성된 리스트를 제공하게 된다.

[0089] 이하, 전술한 상기 주요 인물 검출 단계에 대하여 보다 구체적으로 설명한다. 도 17은 본 발명의 바람직한 실시예에 따른 샷 단위 주요 인물 검출 방법에 있어서, 주요 인물 검출 단계를 구체적으로 도시한 흐름도이다.

[0090] 도 17을 참조하면, 상기 주요 인물 검출 단계는, 먼저 상기 샷 후보 인물 검출 단계에서 모든 프레임에서 선택된 각 후보 얼굴 영역들에 대한 영역 검출 횟수를 측정하고(단계 800), 상기 영역 검출 횟수가 사전 설정된 기준 이상인 후보 얼굴 영역들만을 주요 얼굴 영역으로 검출한다(단계 810).

[0091] 다음, 샷을 구성하는 모든 프레임에 대하여 상기 검출된 각 주요 얼굴 영역에 대한 후보 인물 검출 횟수를 측정하고(단계 820), 상기 후보 인물 검출 횟수가 사전 설정된 기준 이상인 레이블을 샷내의 주요 인물로 선택하여 출력하게 된다(단계 830).

[0092] 이상에서 본 발명에 대하여 그 바람직한 실시예를 중심으로 설명하였으나, 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 발명의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 그리고, 이러한 변형과 응용에 관계된 차이점들은 첨부된 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0093] 1 : 샷 단위 주요 인물 검출 시스템

20 : 샷 단위 검출 장치

30 : 프레임 후보 인물 선택 장치

300 : 얼굴 영역 검출 모듈

310 : 얼굴 영역 정보 검출 모듈

320 : 얼굴 인식 모듈

330 : 보정 모듈

340 : 필터 모듈

350 : 후보 인물 출력 모듈

40 : 샷 후보 인물 획득 장치

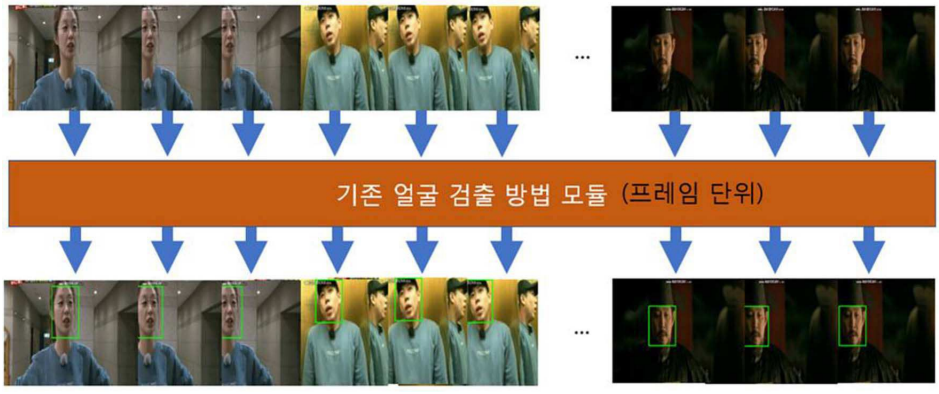
50 : 주요 인물 검출 장치

500 : 주요 얼굴 영역 검출 모듈

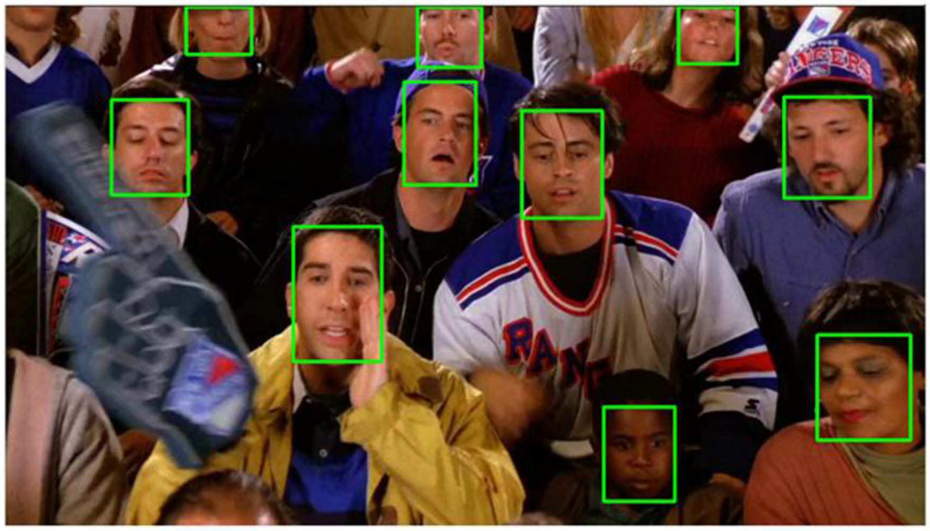
510 : 주요 인물 선택 모듈

도면

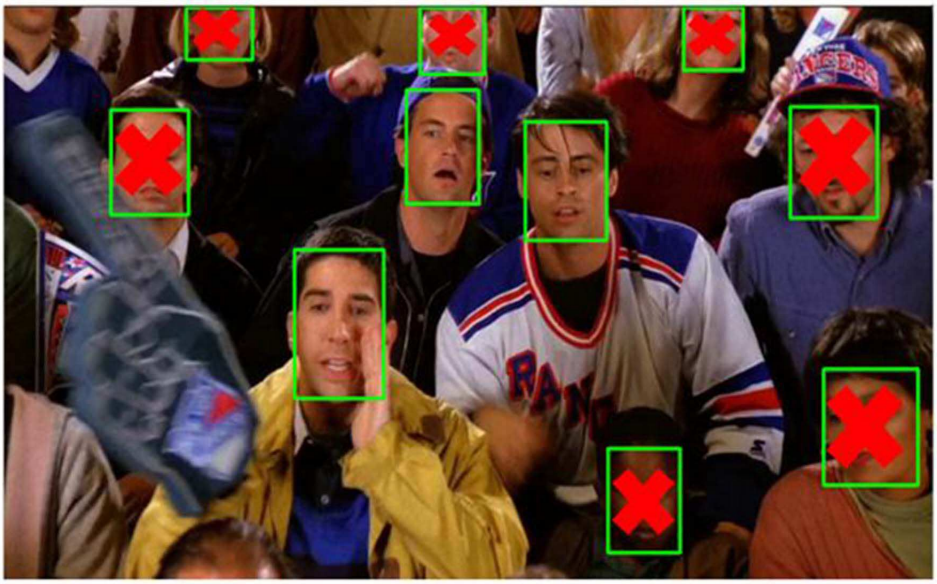
도면1



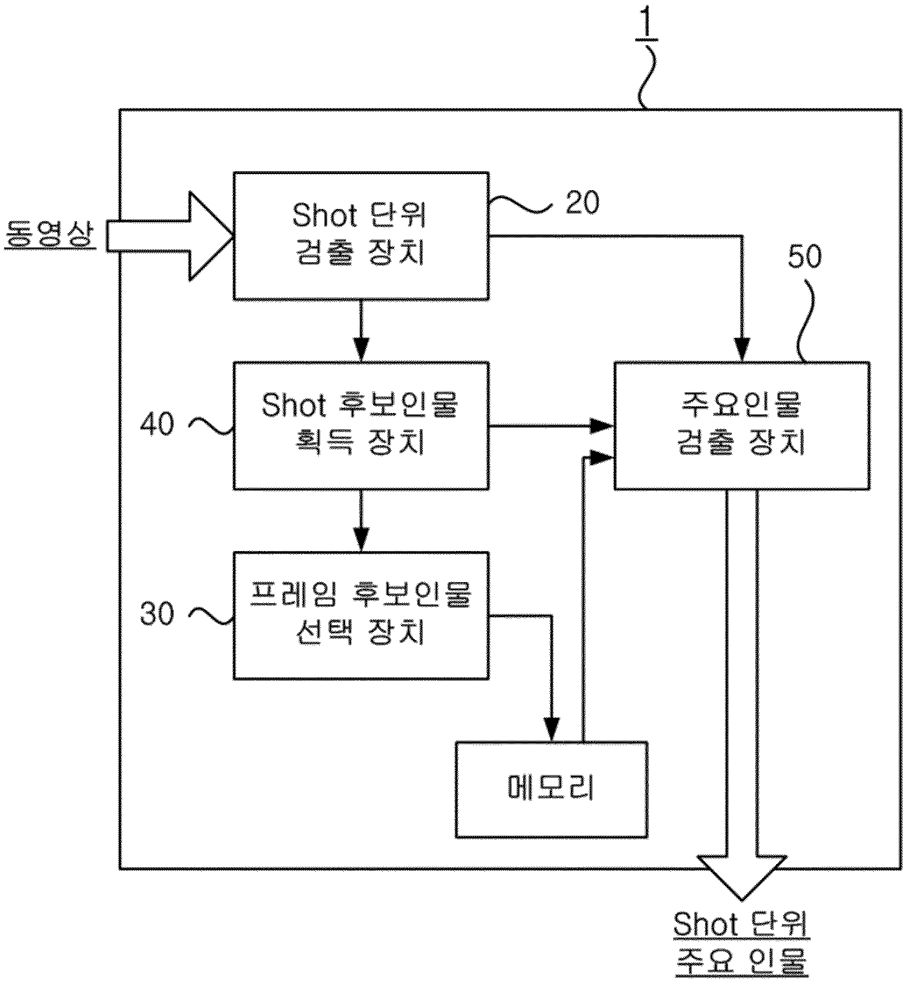
도면2



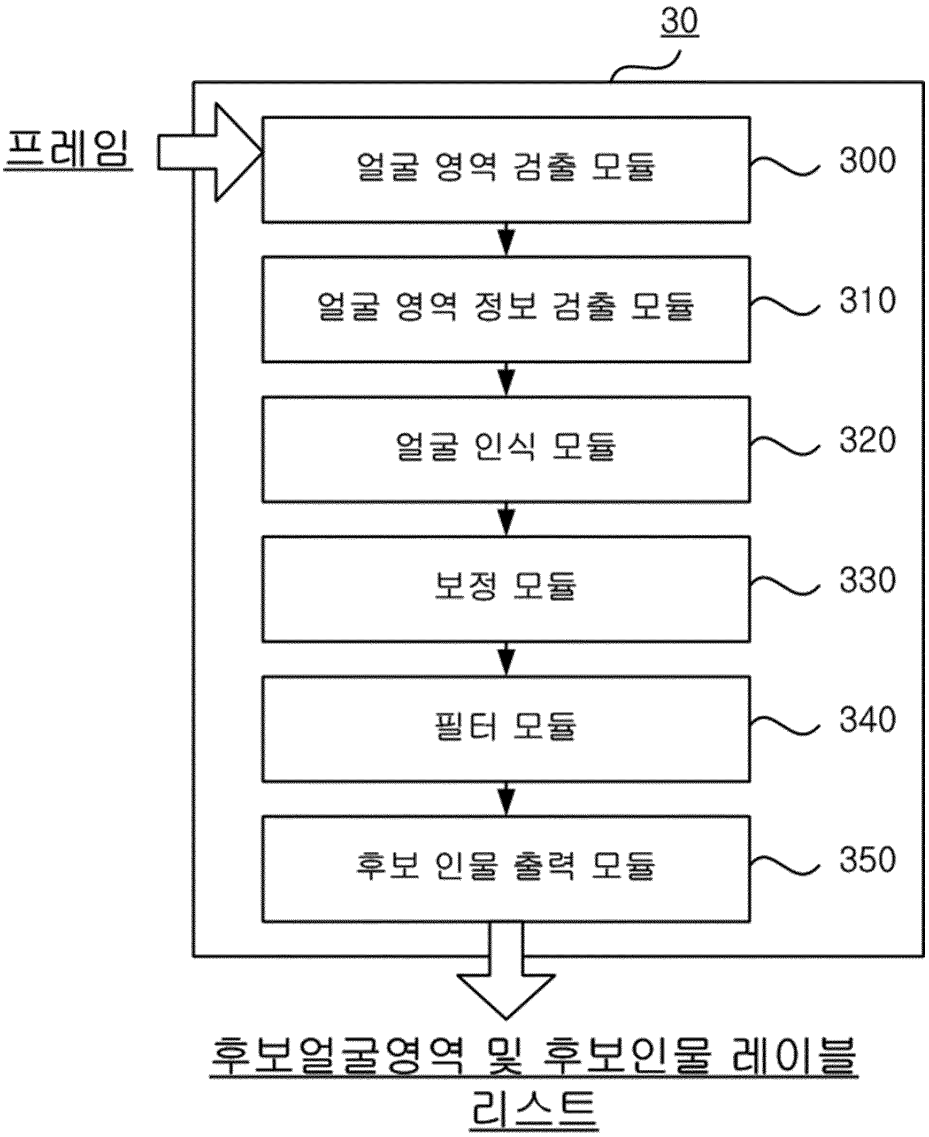
도면3



도면4


1
동영상
Shot 단위
검출 장치
20
50
Shot 후보인물
획득 장치
40
주요인물
검출 장치
프레임 후보인물
선택 장치
30
메모리
Shot 단위
주요 인물

도면5



도면6

얼굴 영역 픽셀 좌표 [x , y, x+w, y+h]	확률값
[147 129 254 261]	0.417893
[255 0 339 65]	0.548898
[400 305 523 492]	0.981599
[554 104 654 247]	0.999777
[573 0 662 84]	0.466968
[718 144 830 293]	0.990940
[832 555 931 686]	0.615315
[936 0 1018 79]	0.378423
[1083 125 1203 265]	0.684197
[1129 455 1259 601]	0.510377

도면7

영역 픽셀 정보	a	b
[147 129 254 261]	14124	0.614060
[554 104 654 247]	14300	0.621712
[1129 455 1259 601]	18980	0.825181
[832 555 931 686]	12969	0.563845
[1083 125 1203 265]	16800	0.730403
[718 144 830 293]	16688	0.725533
[573 0 662 84]	7476	0.325029
[400 305 523 492]	23001	1
[936 0 1018 79]	6478	0.281639
[255 0 339 65]	5460	0.237380

도면8

영역 픽셀 정보	a	b	c
[147 129 254 261]	200.5	195	469.4520742
[554 104 654 247]	604	175.5	187.9793872
[1129 455 1259 601]	1194	528	578.9127741
[832 555 931 686]	881.5	620.5	355.2217617
[1083 125 1203 265]	1143	195	529.3713253
[718 144 830 293]	774	218.5	194.8800913
[573 0 662 84]	617.5	42	318.7949968
[400 305 523 492]	461.5	398.5	182.6047645
[936 0 1018 79]	977	39.5	465.0690809
[255 0 339 65]	297	32.5	474.2417632

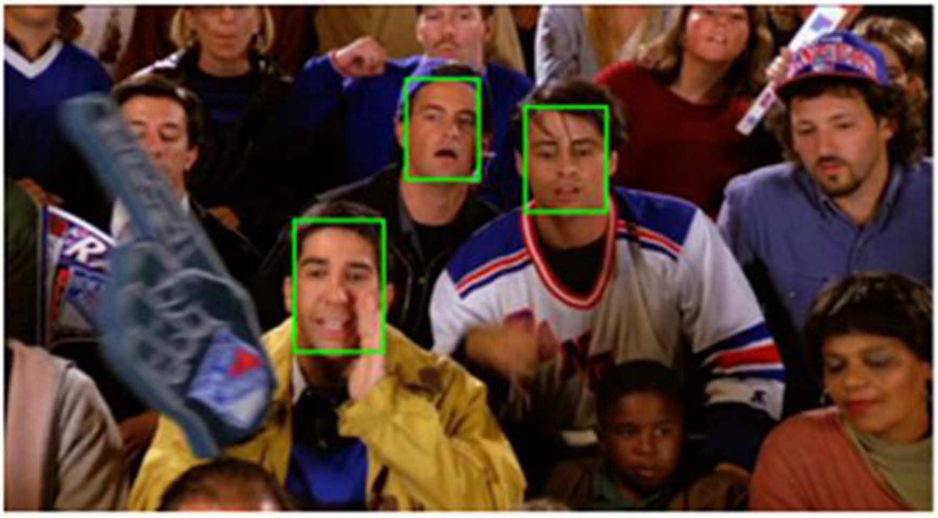
도면9

α	β	γ
0.417893	0.614060	469.4520742
0.999777	0.621712	187.9793872
0.510377	0.825181	578.9127741
0.615315	0.563845	355.2217617
0.684197	0.730403	529.3713253
0.99094	0.725533	194.8800913
0.466968	0.325029	318.7949968
0.981599	1	182.6047645
0.378423	0.281639	465.0690809
0.548898	0.237380	474.2417632

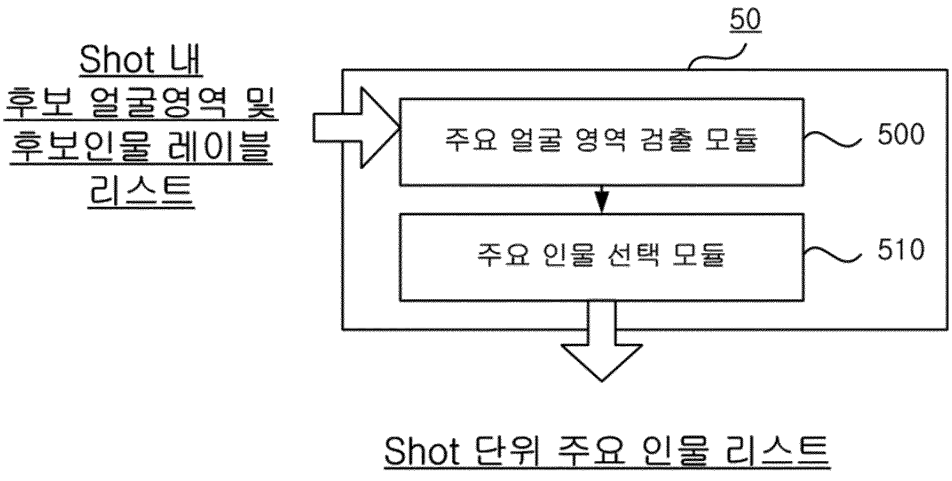
도면10

영역 픽셀 정보	보정확률값
[147 129 254 261]	0.092555
[554 104 654 247]	0.462452
[1129 455 1259 601]	0.089123
[832 555 931 686]	0.179108
[1083 125 1203 265]	0.139469
[718 144 830 293]	0.528152
[573 0 662 84]	0.085884
[400 305 523 492]	0.737497
[936 0 1018 79]	0.039077
[255 0 339 65]	0.046146

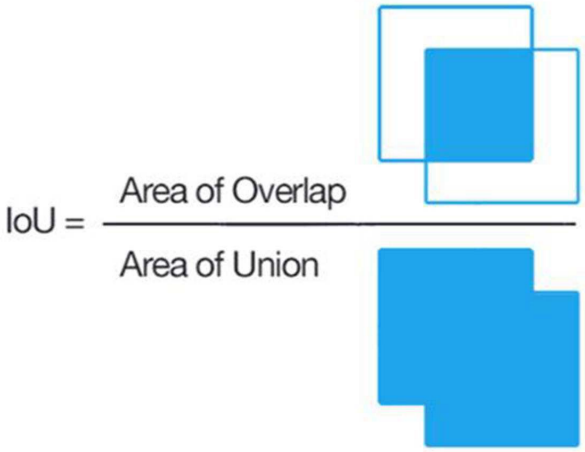
도면11



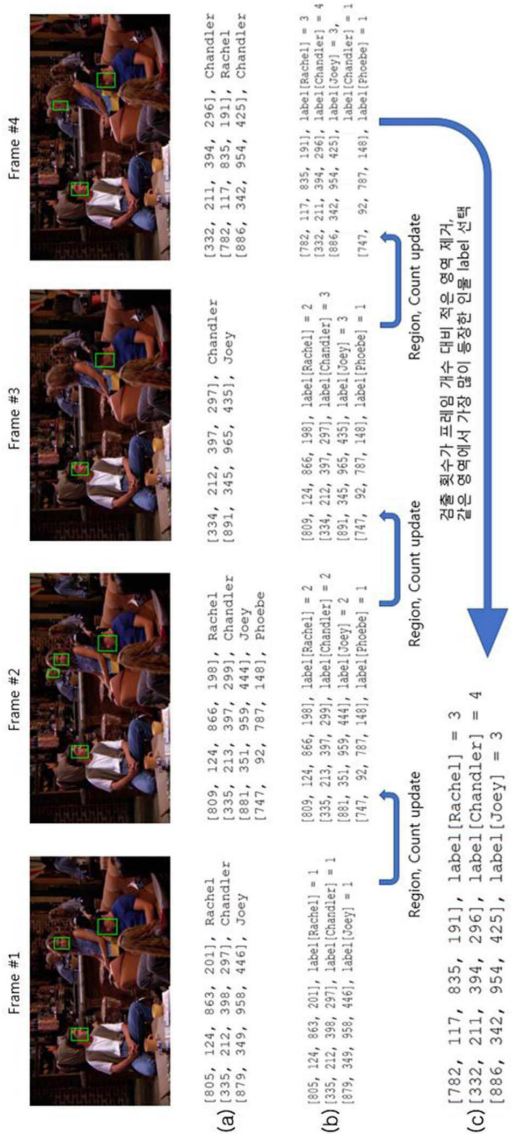
도면12



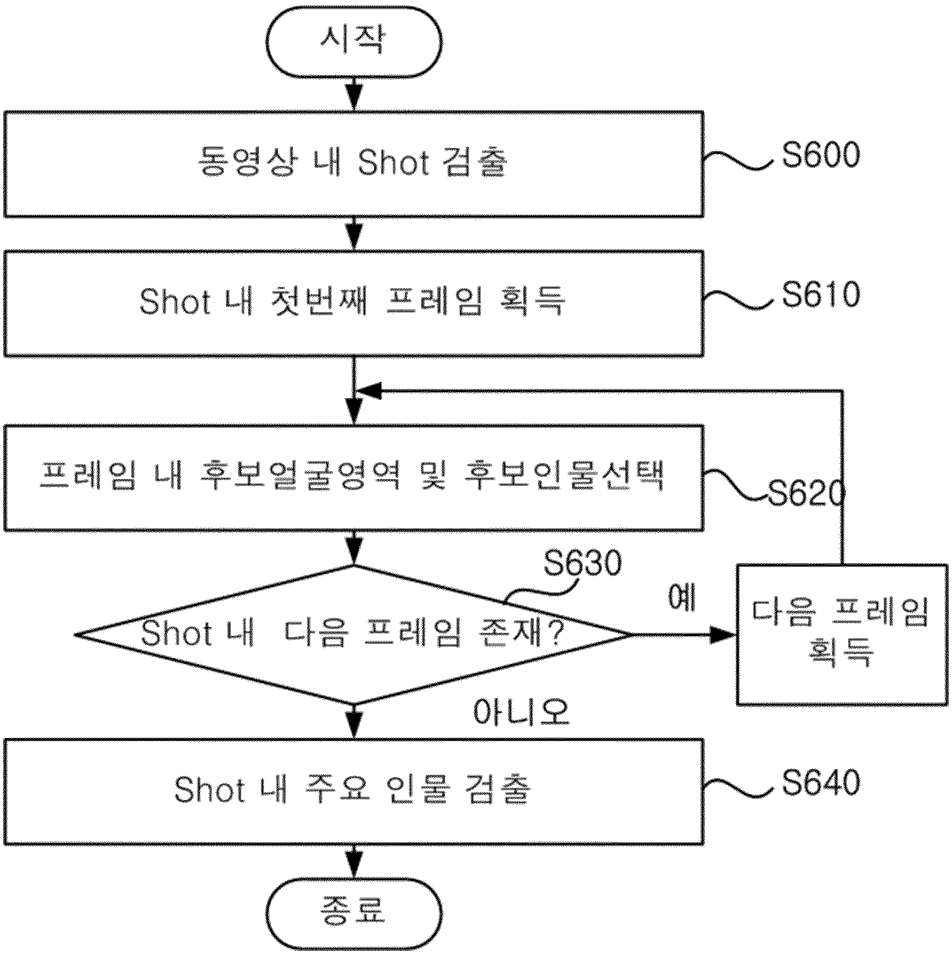
도면13


IoU =
Area of Overlap
Area of Union

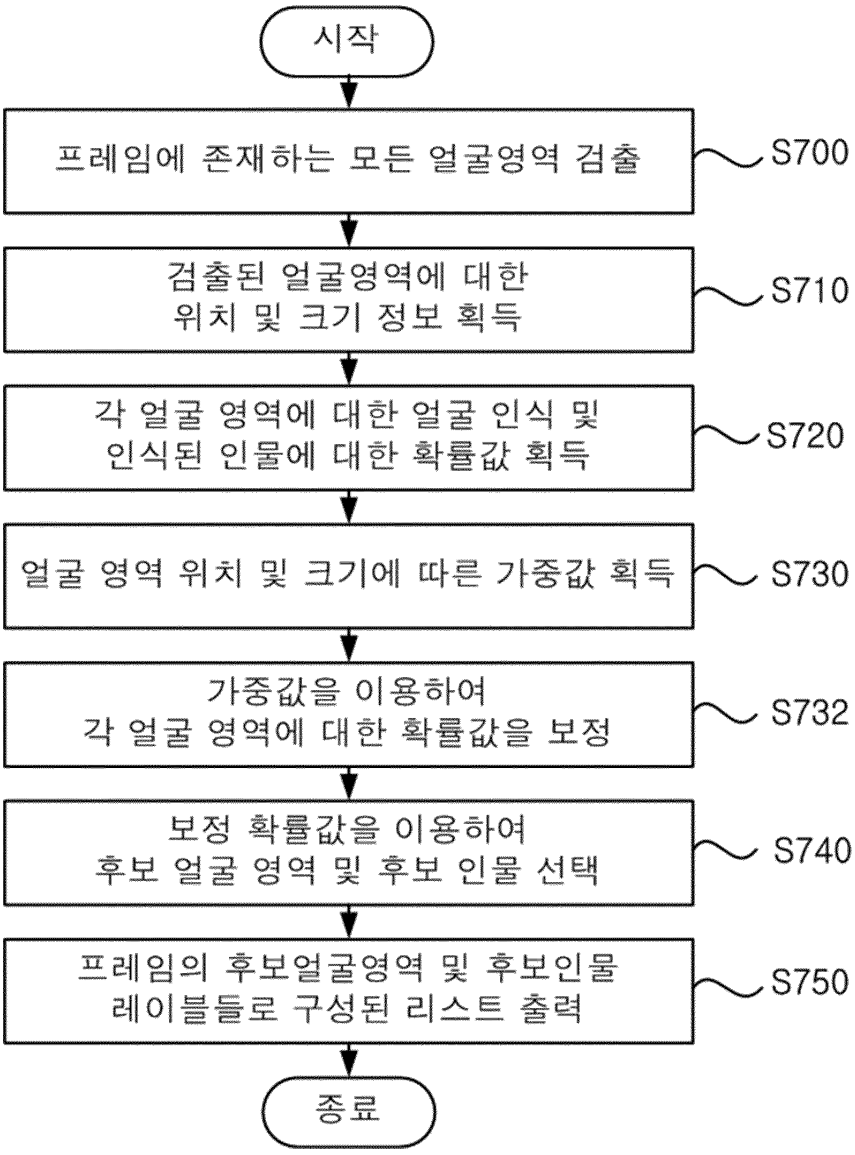
도면14


Frame #1
Frame #2
Frame #3
Frame #4
(a) [805, 124, 863, 201], Rachel
[335, 212, 398, 297], Chandler
[879, 349, 958, 446], Joey
[809, 124, 866, 198], Rachel
[335, 213, 397, 299], Chandler
[881, 351, 959, 444], Joey
[747, 92, 787, 148], Phoebe
[334, 212, 397, 297], Chandler
[891, 345, 965, 435], Joey
[332, 211, 394, 296], Chandler
[782, 117, 835, 191], Rachel
[886, 342, 954, 425], Chandler
(b) [805, 124, 863, 201], label[Rachel] = 1
[335, 212, 398, 297], label[Chandler] = 1
[879, 349, 958, 446], label[Joey] = 1
[809, 124, 866, 198], label[Rachel] = 2
[335, 213, 397, 299], label[Chandler] = 2
[881, 351, 959, 444], label[Joey] = 2
[747, 92, 787, 148], label[Phoebe] = 1
[809, 124, 866, 198], label[Rachel] = 2
[334, 212, 397, 297], label[Chandler] = 3
[891, 345, 965, 435], label[Joey] = 3
[747, 92, 787, 148], label[Phoebe] = 1
[782, 117, 835, 191], label[Rachel] = 3
[332, 211, 394, 296], label[Chandler] = 4
[886, 342, 954, 425], label[Joey] = 3
[747, 92, 787, 148], label[Phoebe] = 1
Region, Count update
Region, Count update
Region, Count update
검출 횟수가 프레임 개수 대비 적은 영역 제거,
같은 영역에서 가장 많이 등장한 인물 label 선택
(c) [782, 117, 835, 191], label[Rachel] = 3
[332, 211, 394, 296], label[Chandler] = 4
[886, 342, 954, 425], label[Joey] = 3

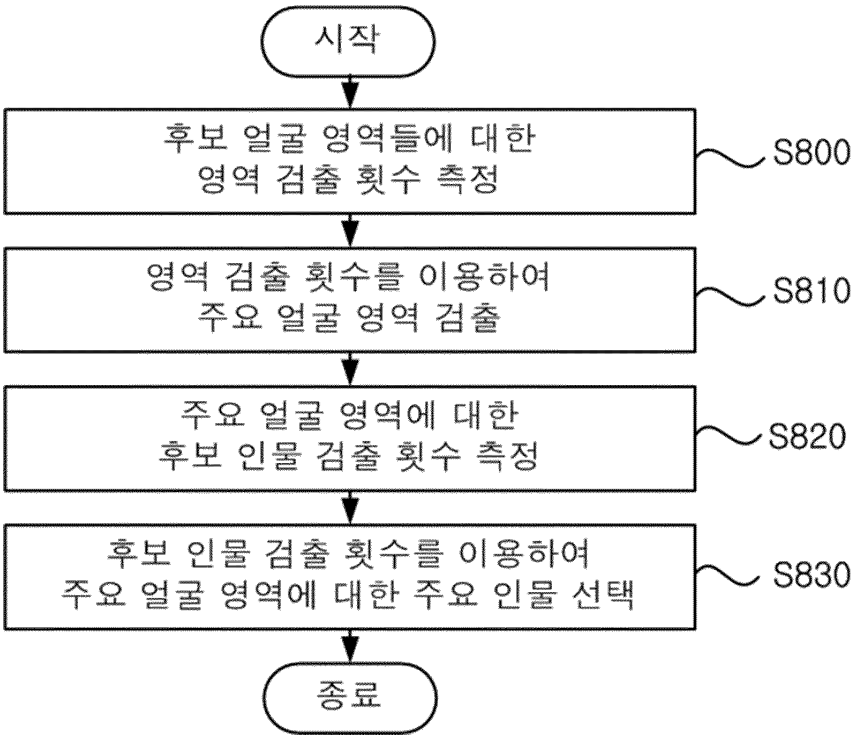
도면15


시작
동영상 내 Shot 검출
S600
Shot 내 첫번째 프레임 획득
S610
프레임 내 후보얼굴영역 및 후보인물선택
S620
S630
Shot 내 다음 프레임 존재?
예
다음 프레임 획득
아니오
Shot 내 주요 인물 검출
S640
종료

도면16


시작
프레임에 존재하는 모든 얼굴영역 검출
S700
검출된 얼굴영역에 대한
위치 및 크기 정보 획득
S710
각 얼굴 영역에 대한 얼굴 인식 및
인식된 인물에 대한 확률값 획득
S720
얼굴 영역 위치 및 크기에 따른 가중값 획득
S730
가중값을 이용하여
각 얼굴 영역에 대한 확률값을 보정
S732
보정 확률값을 이용하여
후보 얼굴 영역 및 후보 인물 선택
S740
프레임의 후보얼굴영역 및 후보인물
레이블들로 구성된 리스트 출력
S750
종료

도면17


시작
후보 얼굴 영역들에 대한
영역 검출 횟수 측정
S800
영역 검출 횟수를 이용하여
주요 얼굴 영역 검출
S810
주요 얼굴 영역에 대한
후보 인물 검출 횟수 측정
S820
후보 인물 검출 횟수를 이용하여
주요 얼굴 영역에 대한 주요 인물 선택
S830
종료