(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0081303
(43) 공개일자 2018년07월16일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01)
(52) CPC특허분류
G06F 17/30793 (2013.01)
G06F 17/3082 (2013.01)
(21) 출원번호 10-2017-0002366
(22) 출원일자 2017년01월06일
심사청구일자 2017년01월06일
기술이전 희망 : **기술양도**

(71) 출원인
울산대학교 산학협력단
울산광역시 남구 대학로 93(무거동)
(72) 발명자
조강현
울산광역시 남구 신복로45번길 4, 현대문수팰리스 동 1005호 (무거동, 현대문수팰리스)
이상희
울산광역시 남구 동산로69번길 13, 101동 603호 (신정동, 문수로2차아이파크)
(74) 대리인
김종선, 이형석

전체 청구항 수 : 총 11 항

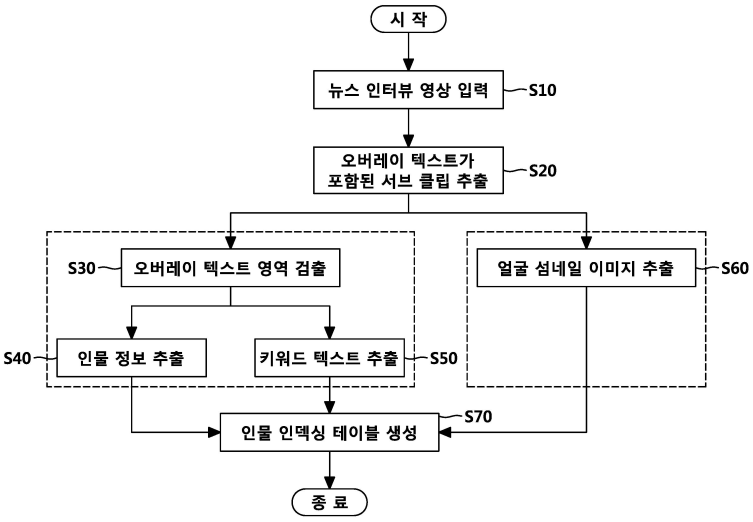
(54) 발명의 명칭 **뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법 및 장치**

(57) 요 약

본 발명은 뉴스 인터뷰 영상을 이용하여 인물 인덱싱 정보를 자동으로 생성 및 수집할 수 있도록 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법 및 장치에 관한 것으로,

상기 방법은 뉴스 인터뷰 영상에서 오버레이 텍스트가 포함된 서브 클립만을 추출하는 단계; 상기 서브 클립의 오버레이 텍스트에서 인터뷰 대상자의 인물 정보 텍스트 및 키워드 텍스트를 추출하는 단계; 상기 서브 클립에서 상기 인터뷰 대상자의 얼굴 섬네일 이미지를 획득하는 단계; 및 상기 인터뷰 대상자의 인물 정보 텍스트, 키워드 텍스트, 및 얼굴 섬네일을 기반으로 인물 인덱싱 테이블을 생성하는 단계를 포함할 수 있다.

대 표 도 - 도2



이 발명을 지원한 국가연구개발사업

과제고유번호	1711054908
부처명	과학기술정보통신부
연구관리전문기관	정보통신기술진흥센터
연구사업명	Grand ICT연구센터지원사업
연구과제명	IoT 및 지능정보 기반 동남권 제조 IT 기술 혁신 및 인재양성
기 여 율	1/1
주관기관	부산대학교 산학협력단
연구기간	2017.01.01 ~ 2017.12.31

명 세 서

청구범위

청구항 1

뉴스 인터뷰 영상에서 오버레이 텍스트가 포함된 서브 클립만을 추출하는 단계;

상기 서브 클립의 오버레이 텍스트에서 인터뷰 대상자의 인물 정보 텍스트 및 키워드 텍스트를 추출하는 단계;

상기 서브 클립에서 상기 인터뷰 대상자의 얼굴 섬네일 이미지를 획득하는 단계; 및

상기 인물 정보 텍스트, 상기 키워드 텍스트, 및 상기 얼굴 섬네일을 기반으로 상기 인터뷰 대상자의 인물 인덱싱 테이블을 생성하는 단계를 포함하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 2

제1 항에 있어서, 상기 서브 클립을 추출하는 단계는

상기 뉴스 영상의 에지 맵 영상으로 변환한 후, 상기 에지 맵 영상을 기반으로 에지 밀도 변화를 추적하는 단계;

상기 에지 밀도 증가량이 제1 설정값 이상이면, 상기 서브 클립의 추출을 시작하는 단계; 및

상기 에지 밀도 감소량 제2 설정값 이상이면, 상기 서브 클립의 추출을 종료하는 단계를 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 3

제1항에 있어서, 상기 인물 정보 텍스트 및 키워드 텍스트를 추출하는 단계는

멀티 에지 맵(multiple edge map) 알고리즘을 이용하여 상기 서브 클립의 오버레이 텍스트 영역을 검출하는 단계;

상기 오버레이 텍스트 영역의 수평 투영 히스토그램을 획득하고, 상기 수평 투영 히스토그램을 기반으로 인물 정보 라인과 키워드 라인을 구분하는 단계; 및

상기 인물 정보 라인 및 상기 키워드 라인으로부터 인물 정보 텍스트와 키워드 텍스트를 각각을 추출하는 단계를 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 4

제3항에 있어서, 상기 서브 클립의 오버레이 텍스트 영역을 검출하는 단계는

상기 서브 클립으로부터 다수의 참고 프레임을 추출하는 단계;

상기 다수의 참고 프레임의 에지 맵 영상을 논리곱하여, 배경 영상 에지가 제거된 멀티 에지 맵 영상을 획득하는 단계; 및

상기 멀티 에지 맵 영상의 흑백 전이 개수를 기반으로 상기 오버레이 텍스트 영역을 검출하는 단계를 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 5

제4항에 있어서, 상기 서브 클립의 오버레이 텍스트 영역을 검출하는 단계는

형태학적 연산을 통해 상기 오버레이 텍스트 영역을 보정하는 단계를 더 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 6

제3항에 있어서, 상기 인물 정보 라인과 키워드 라인을 구분하는 단계는

상기 오버레이 텍스트 영역의 수평 투영 히스토그램을 획득하는 단계; 및

상기 수평 투영 히스토그램이 기 설정값 이상인 수평 라인들 중에서 첫 번째 수평 라인은 인물 정보 라인으로, 나머지 수평라인은 키워드 라인으로 구분하는 단계를 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 7

제6항에 있어서, 상기 수평 투영 히스토그램을 획득하는 단계는

상기 뉴스 영상의 하위 영역에 위치하는 상기 오버레이 텍스트 영역에 대해서만 수평 투영 히스토그램을 획득하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 8

제3항에 있어서, 상기 인물 정보 텍스트와 키워드 텍스트를 각각을 추출하는 단계는

OCR(Optical Character Recognition)을 통해 상기 인물 정보 라인과 상기 키워드 라인 각각의 텍스트를 추출하는 단계; 및

상기 텍스트 추출 결과에 대한 객체명 인식 동작을 수행하여 인터뷰 대상자의 인물 정보와 인터뷰 내용에 대한 키워드를 각각 추출하는 단계를 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 9

제1항에 있어서, 상기 인물 인덱싱 테이블을 생성하는 단계는

상기 인터뷰 대상자의 인물 정보를 기반으로 인터넷 정보 검색을 수행하고, 인터넷 정보 검색 결과를 상기 인물 인덱싱 테이블에 추가 저장하는 단계를 더 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 10

제1항에 있어서, 상기 인물 인덱싱 테이블을 생성하는 단계는

상기 뉴스 인터뷰 영상과 상기 인물 인덱싱 테이블을 데이터베이스에 함께 저장하는 단계를 더 포함하는 것을 특징으로 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법.

청구항 11

뉴스 영상에서 오버레이 텍스트가 포함된 서브 클립만을 추출하는 서브 클립 추출부;

상기 서브 클립의 오버레이 텍스트 영역에서 인터뷰 대상자의 인물 정보 텍스트 및 키워드 텍스트를 추출하는 텍스트 추출부;

상기 서브 클립에서 상기 인터뷰 대상자의 얼굴 섬네일 이미지를 추출하는 얼굴 섬네일 추출부; 및

상기 인터뷰 대상자의 인물 정보 텍스트, 키워드 텍스트, 및 얼굴 섬네일을 기반으로 인물 인덱싱 테이블을 생성하는 인물 인덱싱 테이블 생성부를 포함하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 뉴스 인터뷰 영상을 이용하여 인물 인덱싱 정보를 자동으로 생성 및 수집할 수 있도록 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 장치에 관한 것이다.

배 경 기 술

[0002] 일반적으로 비디오 시퀀스의 텍스트는 오버레이 텍스트와 장면 텍스트로 구분된다. 그래픽 텍스트 또는 자막이라고 하는 오버레이 텍스트는 편집시 그래픽으로 생성되어 비디오 영상 위에 인위적으로 오버레이 되는 특징을

가진다.

[0003] 오버레이 텍스트의 예로는 뉴스 영상의 자막, 스포츠 경기 점수가 있으며, 장면 텍스트의 예로는 거리 표지판, 트럭 번호판, 셔츠 글자 등과 같이 카메라의 영상 촬영 결과물에 기록된 각종 텍스트일 수 있다.

[0004] 오버레이 텍스트는 비디오 시퀀스에 대한 중요 정보 또는 요약 정보를 추가적으로 제공하기 위한 목적으로 주로 사용되므로, 오버레이 텍스트는 장면 이해, 색인 생성, 탐색 및 검색과 같은 자동화된 콘텐츠 분석 시스템에서 필수적인 문제이다.

[0005] 특히, 뉴스 인터뷰 영상의 경우, 도1에서와 같이 인터뷰 대상자에 대한 이름, 나이 등과 같은 인물 정보와 함께 인터뷰 내용이 오버레이 텍스트를 통해 추가적으로 제공하는 특징을 가진다.

[0006] 이에 뉴스 인터뷰 영상을 통해 제공되는 오버레이 텍스트는 인물 인덱싱에 필요한 매우 많은 정보를 제공하고 있으나, 현재까지 이를 활용하도록 하는 기술이 전혀 개발되지 못하고 있다.

[0007] 다만, 국내공개특허 제10-2011-0032347호에서 자막/대본 정보 데이터베이스와 같은 별도의 수단에 접근하여 동영상 등장인물에 대한 부가 정보를 추출하도록 하는 개념만을 언급하고 있을 뿐이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 국내공개특허 제10-2011-0032347호(공개일자: 2011.03.30.)

발명의 내용

해결하려는 과제

[0009] 이에 본 발명에서는 뉴스 인터뷰 영상을 이용하여 인물 인덱싱 정보를 자동으로 생성 및 수집할 수 있도록 하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법 및 장치를 제공하고자 한다.

[0010] 본 발명의 목적은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 본 발명이 속하는 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0011] 상기 과제를 해결하기 위한 수단으로서, 본 발명의 일 실시 형태에 따르면, 뉴스 인터뷰 영상에서 오버레이 텍스트가 포함된 서브 클립만을 추출하는 단계; 상기 서브 클립의 오버레이 텍스트에서 인터뷰 대상자의 인물 정보 텍스트 및 키워드 텍스트를 추출하는 단계; 상기 서브 클립에서 상기 인터뷰 대상자의 얼굴 섬네일(thumbnail) 이미지를 획득하는 단계; 및 상기 인물 정보 텍스트, 상기 키워드 텍스트, 및 상기 얼굴 섬네일을 기반으로 상기 인터뷰 대상자의 인물 인덱싱 테이블을 생성하는 단계를 포함하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법을 제공한다.

[0012] 상기 서브 클립을 추출하는 단계는 상기 뉴스 인터뷰 영상의 에지 맵 영상으로 변환한 후, 상기 에지 맵 영상을 기반으로 에지 밀도 변화를 추적하는 단계; 상기 에지 밀도 증가량이 제1 설정값 이상이면, 상기 서브 클립의 추출을 시작하는 단계; 및 상기 에지 밀도 감소량이 제2 설정값 이상이면, 상기 서브 클립의 추출을 종료하는 단계를 포함할 수 있다.

[0013] 상기 인물 정보 텍스트 및 키워드 텍스트를 추출하는 단계는 멀티 에지 맵(multiple edge map) 알고리즘을 이용하여 상기 서브 클립의 오버레이 텍스트 영역을 검출하는 단계; 상기 오버레이 텍스트 영역의 수평 투영 히스토그램을 획득하고, 상기 수평 투영 히스토그램을 기반으로 인물 정보 라인과 키워드 라인을 구분하는 단계; 및 상기 인물 정보 라인 및 상기 키워드 라인으로부터 인물 정보 텍스트와 키워드 텍스트를 각각을 추출하는 단계를 포함할 수 있다.

[0014] 상기 서브 클립의 오버레이 텍스트 영역을 검출하는 단계는 상기 서브 클립으로부터 다수의 참고 프레임을 추출하는 단계; 상기 다수의 참고 프레임의 에지 맵 영상을 논리곱하여, 배경 영상 에지가 제거된 멀티 에지 맵 영상을 획득하는 단계; 및 상기 멀티 에지 맵 영상의 흑백 전이 개수를 기반으로 상기 오버레이 텍스트 영역을 검

출하는 단계를 포함할 수 있다.

[0015] 상기 서브 클립의 오버레이 텍스트 영역을 검출하는 단계는 형태학적 연산을 통해 상기 오버레이 텍스트 영역을 보정하는 단계를 더 포함할 수 있다.

[0016] 상기 인물 정보 라인과 키워드 라인을 구분하는 단계는 상기 오버레이 텍스트 영역의 수평 투영 히스토그램을 획득하는 단계; 및 상기 수평 투영 히스토그램이 기 설정값 이상인 수평 라인들 중에서 첫 번째 수평 라인은 인물 정보 라인으로, 나머지 수평라인은 키워드 라인으로 구분하는 단계를 포함할 수 있다.

[0017] 상기 수평 투영 히스토그램을 획득하는 단계는 상기 뉴스 영상의 하위 영역에 위치하는 상기 오버레이 텍스트 영역에 대해서만 수평 투영 히스토그램을 획득하는 것을 특징으로 한다.

[0018] 상기 인물 정보 텍스트와 키워드 텍스트를 각각을 추출하는 단계는 OCR(Optical Character Recognition)을 통해 상기 인물 정보 라인과 상기 키워드 라인 각각의 텍스트를 추출하는 단계; 및 상기 텍스트 추출 결과에 대한 객체명 인식 동작을 수행하여 인터뷰 대상자의 인물 정보와 인터뷰 내용에 대한 키워드를 각각 추출하는 단계를 포함할 수 있다.

[0019] 상기 인물 인덱싱 테이블을 생성하는 단계는 상기 인터뷰 대상자의 인물 정보를 기반으로 인터넷 정보 검색을 수행하고, 인터넷 정보 검색 결과를 상기 인물 인덱싱 테이블에 추가 저장하는 단계를 더 포함할 수 있다.

[0020] 상기 인물 인덱싱 테이블을 생성하는 단계는 상기 뉴스 인터뷰 영상과 상기 인물 인덱싱 테이블을 데이터베이스에 함께 저장하는 단계를 더 포함할 수 있다.

[0021] 상기 과제를 해결하기 위한 수단으로서, 본 발명의 다른 실시 형태에 따르면, 뉴스 인터뷰 영상에서 오버레이 텍스트가 포함된 서브 클립만을 추출하는 서브 클립 추출부; 상기 서브 클립의 오버레이 텍스트 영역에서 인터뷰 대상자의 인물 정보 텍스트 및 키워드 텍스트를 추출하는 텍스트 추출부; 상기 서브 클립에서 상기 인터뷰 대상자의 얼굴 섬네일 이미지를 추출하는 얼굴 섬네일 추출부; 및 상기 인터뷰 대상자의 인물 정보 텍스트, 키워드 텍스트, 및 얼굴 섬네일 이미지를 기반으로 인물 인덱싱 테이블을 생성하는 인물 인덱싱 테이블 생성부를 포함하는 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 장치를 제공한다.

발명의 효과

[0022] 본 발명의 뉴스 인터뷰 영상을 이용하여 인물 인덱싱 정보를 자동으로 생성 및 수집할 수 있도록 함으로써, 방대한 양의 인물 인덱싱 정보를 최소한의 시간과 비용으로 구축하고 제공할 수 있도록 해준다.

도면의 간단한 설명

[0023] 도1은 일반적인 뉴스 인터뷰 영상의 일예를 도시한 도면이다.

도2는 본 발명의 일 실시예에 따른 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법을 설명하기 위한 도면이다.

도3은 본 발명의 일 실시예에 따른 서브 클립을 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

도4은 본 발명의 일 실시예에 따른 서브 클립을 추출하는 단계에서 뉴스 인터뷰 영상의 오버레이 텍스트 존재 구간 구분 방법을 설명하기 위한 도면이다.

도5은 본 발명의 일 실시예에 따른 뉴스 인터뷰 영상의 오버레이 텍스트 존재 구간에 따른 에지 밀도 변화를 설명하기 도면이다.

도6은 본 발명의 일 실시예에 따른 오버레이 텍스트 영역을 검출하는 단계를 보다 상세히 설명하기 위한 도면이다.

도7은 본 발명의 일 실시예에 따른 참고 프레임 획득 방법을 설명하기 위한 도면이다.

도8은 본 발명의 일 실시예에 따른 멀티 에지 맵 영상과 오버레이 텍스트 영역 검출 결과를 도시한 도면이다.

도9는 본 발명의 일 실시예에 따른 인물 정보 텍스트 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

도10는 본 발명의 일 실시예에 따른 키워드 텍스트 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

도11은 본 발명의 일 실시예에 따른 얼굴 섬네일 이미지를 추출하는 단계를 보다 상세히 설명하기 위한 도면이

다.

도12는 본 발명의 일 실시예에 따른 인물 인덱싱 테이블을 도시한 도면이다.

도13은 본 발명의 일 실시예에 따른 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 장치를 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0024] 본 발명에서 사용되는 기술적 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아님을 유의해야 한다. 또한, 본 발명에서 사용되는 기술적 용어는 본 발명에서 특별히 다른 의미로 정의되지 않는 한, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 의미로 해석되어야 하며, 과도하게 포괄적인 의미로 해석되거나, 과도하게 축소된 의미로 해석되지 않아야 한다. 또한, 본 발명에서 사용되는 기술적인 용어가 본 발명의 사상을 정확하게 표현하지 못하는 잘못된 기술적 용어일 때에는, 당업자가 올바르게 이해할 수 있는 기술적 용어로 대체되어 이해되어야 할 것이다. 또한, 본 발명에서 사용되는 일반적인 용어는 사전에 정의되어 있는 바에 따라, 또는 전후 문맥상에 따라 해석되어야 하며, 과도하게 축소된 의미로 해석되지 않아야 한다.

[0025] 또한, 본 발명에서 사용되는 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한 복수의 표현을 포함한다. 본 발명에서, "구성된다" 또는 "포함한다" 등의 용어는 발명에 기재된 여러 구성 요소들, 또는 여러 단계를 반드시 모두 포함하는 것으로 해석되지 않아야 하며, 그 중 일부 구성 요소들 또는 일부 단계들은 포함되지 않을 수도 있고, 또는 추가적인 구성 요소 또는 단계들을 더 포함할 수 있는 것으로 해석되어야 한다.

[0026] 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예를 상세히 설명하되, 도면 부호에 관계없이 동일하거나 유사한 구성 요소는 동일한 참조 번호를 부여하고 이에 대한 중복된 설명은 생략하기로 한다.

[0027] 또한, 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다. 또한, 첨부된 도면은 본 발명의 사상을 쉽게 이해할 수 있도록 하기 위한 것일 뿐, 첨부된 도면에 의해 본 발명의 사상이 제한되는 것으로 해석되어서는 아니 됨을 유의해야 한다.

[0028] 도2는 본 발명의 일 실시예에 따른 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 방법을 설명하기 위한 도면이다.

[0029] 도2을 참고하면, 본 발명의 인물 인덱싱 방법은 뉴스 인터뷰 영상을 입력받는 단계(S10), 뉴스 인터뷰 영상에서 오버레이 텍스트 영역이 포함된 서브 클립만을 추출하는 단계(S20), 서브 클립의 오버레이 텍스트를 검출하는 단계(S30), 오버레이 텍스트 영역에서 인물 정보 텍스트를 추출하는 단계(S40), 오버레이 텍스트 영역에서 키워드 텍스트를 추출하는 단계(S50), 서브 클립에서 얼굴 섬네일 이미지를 추출하는 단계(S60), 그리고 인물 정보, 키워드 및 얼굴 섬네일 추출 결과를 기반으로 인물 인덱싱 테이블을 생성하는 단계(S70) 등을 포함할 수 있다.

[0030] 즉, 본 발명에서는 뉴스 인터뷰 영상에서 인터뷰 대상자의 얼굴 이미지 뿐 만 아니라 인터뷰 대상자의 인물 정보 및 키워드까지 추출한 후, 이들 정보를 활용하여 인물 인덱싱 정보를 자동 생성하도록 함을 알 수 있다.

[0031] 도3은 본 발명의 일 실시예에 따른 서브 클립을 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

[0032] 도3을 참고하면, 본 발명의 서브 클립을 추출하는 단계(S20)는 캐니 에지 검출기(Canny edge detector)를 통해 뉴스 영상의 에지 맵 영상을 생성한 후(S21), 에지 맵 영상 기반으로 에지 밀도 변화를 추적하도록 한다(S22). 단계 S22의 에지 밀도 변화량 추적 결과, 에지 밀도 증가량이 제1 값 이상이 되면(S23), 서브 클립을 추출하기 시작하고(S24), 소정의 시간이 경과하여, 에지 밀도 감소량이 기 제2값 이상이 되면(S25), 서브 클립 추출을 종료하도록 한다(S26).

[0033] 뉴스 영상의 오버레이 텍스트는 뉴스 프로그램 제작 규칙을 준수하여 제작 및 제공되며, 예를 들어, 영상 가림 현상 발생을 최소화하기 위해 프레임 하단 또는 상단 좌우 모서리 상에 표시된다. 또한 인터뷰 영상의 경우 첫 번째 라인에는 인터뷰 대상자의 인물 정보가 기재되고, 나머지 라인을 통해 인터뷰 내용이 기재되는 특징이 있다.

[0034] 또한, 뉴스 영상의 경우, 뉴스 영상의 모든 프레임에 오버레이 텍스트가 포함되지 않으며, 서서히 또는 갑자기 표시된 후 일정 시간 후 사라지는 특징이 있다. 이에 뉴스 영상의 제공 구간은 크게 도4에서와 같이 오버레이 텍스트가 표시되지 않는 텍스트 미존재 구간(Non-text period), 텍스트가 오버레이 되기 시작하는 전환 구간(Transition period), 텍스트가 완전히 표시되는 텍스트 존재 구간(Text period)의 세 구간으로 나누어 질 수

있다.

[0035] 이에 본 발명에서는 텍스트 존재 구간(Text period)에 속하는 프레임에 대해서만 텍스트 검출 및 인식 동작을 수행함으로써, 텍스트 검출 및 인식에 소요되는 시간과 비용이 최소화될 수 있도록 한다.

[0036] 텍스트는 많은 에지를 포함하기 때문에 오버레이 텍스트를 포함하는 프레임은 오버레이 텍스트를 포함하지 않는 프레임에 비해 매우 큰 에지 밀도값을 가지게 된다. 예를 들어, 도5에 도시된 바와 같이 전환 구간(Transition period)에 속하는 프레임 13 번에서부터 프레임 20 번에 걸쳐 에지 밀도가 급격히 증가하며, 텍스트 존재 구간(Text period)에 속하는 시작 프레임 21 이후에는 에지 밀도 차이가 거의 발생하지 않음을 알 수 있다.

[0037] 따라서 본 발명에서는 뉴스 인터뷰 영상의 에지 밀도 변화량을 추적함으로써, 시작 프레임을 검출할 수 있으며, 시작 프레임 이후의 프레임을 서브 클립으로 추출하기 시작하도록 한다. 또한, 동일한 원리로 뉴스 인터뷰 영상의 에지 밀도 변화량을 추적함으로써, 종료 프레임의 발생한 후 서브 클립의 추출을 종료함으로써, 텍스트 존재 구간(Text period)에 속하는 프레임만을 서브 클립으로 손쉽게 추출할 수 있게 된다.

[0038] 도6은 본 발명의 일 실시예에 따른 오버레이 텍스트 영역을 검출하는 단계를 보다 상세히 설명하기 위한 도면이다.

[0039] 도6을 참고하면, 본 발명의 오버레이 텍스트 영역을 검출하는 단계(S30)는 서브 클립으로부터 n개의 참조 프레임을 선택하는 단계(S31), n개의 참조 프레임의 에지 맵 영상 모두를 논리 곱 연산하여, 배경 에지 성분이 제거된 멀티 에지 맵 영상을 획득하는 단계(S32), 멀티 에지 맵 영상의 흑백 전이 개수를 기반으로 오버레이 텍스트 영역을 검출하는 단계(S33), 및 오버레이 텍스트 영역에 대한 형태학적 닫힘(closing)과 팽창(dilation) 연산을 수행하여 오버레이 텍스트 영역을 보정하는 단계(S34) 등을 포함한다.

[0040] 서브 클립이 f frames/sec 속도로 재생되는 경우, 오버레이 텍스트는 최소 2f 프레임 이상 동안 고정된 위치에 있게 된다. k를 f 보다 작지 않은 정수로 가정하면, 연속되는 k개의 프레임들은 한 라운드(one round)라고 정의될 수 있다.

[0041] 이에 본 발명에서는 계산 간단화를 위해, 첫 번째 라운드에 속하는 서브 클립을 기 설정된 주기(예를 들어, k/n-1 주기)로 반복 샘플링하여 n개의 참조 프레임을 추출하도록 한다. 즉, 도7에서와 같이 첫 번째 라운드에 속하는 서브 클립을 1, [k/3], 2[k/3], 3 [k/3] 주기로 반복 샘플링하여 4개의 참조 프레임을 추출할 수 있을 것이다.

[0042] 이와 같이 n개의 참조 프레임이 획득되면, n개의 참조 프레임 각각의 에지 맵 영상을 지그재그 스캐닝하면서 문자의 폭(w) 또는 높이(h) 보다 큰 수평 라인 또는 수직 라인을 제거하도록 함으로써, 노이즈 에지 성분을 1차 제거하도록 한다.

[0043] 그리고 나서 n개의 참조 프레임의 에지 맵 영상 모두를 논리 곱하여, 도8의 (a)와 같이 배경 에지 성분이 모두 제거된 멀티 에지 맵 영상을 획득하도록 한다.

[0044] 그리고 멀티 에지 맵 영상을 지그재그 스캐닝하면서 수학식1을 통해 흑백 전이 개수(N_{trans})를 획득한 후, 흑백 전이 개수가 기 설정된 값 이상인 영역을 모두 합하여 오버레이 텍스트 영역을 획득하도록 한다.

[0045] [수학식1]

[0046]
$$N_{trans} = \sum_{i=0}^{h-1}\left(\sum_{j=1}^{w-1}|b(i,j)-b(i,j-1)|\right)+\sum_{j=0}^{w-1}\left(\sum_{i=1}^{h-1}|b(i,j)-b(i-1,j)|\right)$$

[0047] 이때, w 및 h는 스캐닝 윈도우의 폭과 높이이고, b(ㅇ)는 멀티 에지 맵 영상이며, 스캐닝 윈도우의 크기는 문자의 크기에 종속되어 결정된다.

[0048] 그리고 마지막으로, 형태학적 닫힘과 팽창(morphological closing and dilation) 연산을 순차 수행하여, 도8의 (b)와 같이 오버레이 텍스트 영역을 보정하도록 한다. 즉, 형태학적 닫힘과 팽창 연산을 통해 참조 프레임 에지 영상 모두의 논리 곱으로 발생한 문자의 일부 픽셀이 손실되는 문제가 해결되도록 한다.

[0049] 도9는 본 발명의 일 실시예에 따른 인물 정보 텍스트 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

[0050] 도9에 도시된 바와 같이, 본 발명의 인물 정보 텍스트 추출하는 단계(S40)는 오버레이 텍스트 영역에 대한 수평 투영을 수행하여 수평 투영 히스토그램을 구하고, 수평 투영 히스토그램이 기 설정값 이상인 수평라인들 중에서

첫 번째 라인을 인물 정보 라인으로 획득하는 단계(S41), 인물 정보 라인에 대한 OCR(Optical Character Recognition)을 수행하여 인물 정보 라인에 기재된 텍스트를 검출하는 단계(S42), 텍스트 검출 결과에 대한 객체명 인식 동작을 수행하여 인터뷰 대상자의 인물 정보를 추출하는 단계(S43)를 포함할 수 있다.

[0051] 뉴스 영상은 방송국과 같은 전문 기관에 의해 제작되는 영상물로, 뉴스 프로그램 제작 규칙하에 오버레이 텍스트가 생성 및 제공된다. 특히, 인터뷰 영상에 대응되는 오버레이 텍스트는 프레임의 상위 영역, 중간 영역, 및 하위 영역 중 하위 영역에 표시되며, 첫 번째 라인에는 인터뷰 대상자의 인물 정보가 기재되고, 나머지 라인을 통해 인터뷰 내용이 기재되는 특징이 있다.

[0052] 이에 본 발명에서는 뉴스 영상의 전체 영역이 아닌 하위 영역에 위치하는 오버레이 텍스트 영역에 대해서만 수평 투영을 수행함으로써, 텍스트 검출 및 인식에 소요되는 시간과 비용이 더욱 감소될 수 있도록 한다.

[0053] 그리고 본 발명은 상기의 뉴스 프로그램 제작 규칙을 고려하여, 즉 오버레이 텍스트의 첫 번째 라인에는 인터뷰 대상자의 인물 정보가 기재됨을 고려하여 오버레이 텍스트의 첫 번째 라인은 인물 정보 라인으로, 나머지 라인은 키워드 라인으로 구분하도록 한다.

[0054] 또한 본 발명의 객체명 인식 동작은 CRF(Conditional Random Fields) 기반으로 사람 이름, 소속/기관, 직업/직위 등과 같은 개체명을 인식할 수 있으며, 객체명 인식 동작에 필요한 개체명 사전은 기계 학습(machine learning) 기반으로 지속적으로 추가 또는 갱신될 수 있도록 한다.

[0055] 예를 들어, "어네스트 / 백악관 대변인"이라는 텍스트를 입력받아, "사람 이름(Person): 어네스트, 소속/기관(Organization) : 백악관, 직업/직위(Occupation): 대변인, 기타 정보: 없음"이라는 객체명 인식 결과를 획득할 수 있게 된다.

[0056] 도10는 본 발명의 일 실시예에 따른 키워드 텍스트 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

[0057] 도10에 도시된 바와 같이, 본 발명의 키워드 정보 텍스트 추출하는 단계(S50)는 오버레이 텍스트 영역의 수평 투영을 통해 수평 투영 히스토그램을 구하고, 수평 투영 히스토그램이 기 설정값 이상인 수평라인들 중에서 첫 번째 라인 이외의 라인들을 키워드 라인으로 획득하는 단계(S51), 키워드 라인에 대한 OCR을 수행하여 키워드 라인에 기재된 텍스트를 검출하는 단계(S52), 텍스트 검출 결과에 대한 객체명 인식 동작을 수행하여 인터뷰 대상자에 관련된 키워드를 추출하는 단계(S53)를 포함할 수 있다.

[0058] 즉, 본 발명은 오버레이 텍스트의 첫 번째 줄을 인물 정보 라인으로 획득하고, 나머지 라인은 키워드 라인으로 획득한 후, 앞서 설명된 객체명 인식 동작을 동일하게 수행하여, 키워드 라인에서 추출된 텍스트로부터 인터뷰 내용을 규정할 수 있는 적어도 하나의 키워드를 추출할 수 있도록 한다.

[0059] 예를 들어 "제프리 파울씨는 북한을 떠나라는 허락을 받아 가족들과 재회를 위해 집으로 돌아오고 있습니다. 북한의 석방 결정을 환영합니다"라는 텍스트를 입력받아, "제프리 파울, 북한, 허락, 가족, 재회, 집, 석방, 결정, 환영"와 같은 객체명 인식 결과를 획득할 수 있게 된다.

[0060] 도11은 본 발명의 일 실시예에 따른 얼굴 섬네일 이미지를 추출하는 단계를 보다 상세히 설명하기 위한 도면이다.

[0061] 도11에 도시된 바와 같이, 본 발명의 얼굴 영상을 검출하는 단계(S60)는 서브 클립의 시작 프레임 영상을 획득하는 단계(S61), 시작 프레임 영상에 대한 얼굴 검출 동작을 수행하여 인터뷰 대상자의 얼굴을 검출하는 단계(S62), 단계 S62의 얼굴 검출 결과를 섬네일(thumbnail) 이미지로써 저장하는 단계(S63) 등을 포함한다.

[0062] 다만, 시작 프레임 영상을 통해 인터뷰 대상자 얼굴 검출이 실패하는 경우, 다음 프레임 영상을 통해 인터뷰 대상자 얼굴 검출을 수행할 수도 있도록 한다.

[0063] 또한 필요한 경우에는 여러 영상(예를 들어, n 개의 참고 프레임)에 대해 얼굴 검출 동작을 수행한 후, 얼굴 검출 정확도 기준으로 하나의 얼굴 검출 결과를 선택할 수도 있도록 한다. 얼굴 검출 정확도는 인터뷰 대상자의 촬영 시점, 얼굴 가림 정도, 떨림 정도, 및 화면 해상도 등을 고려하여 계산될 수 있을 것이다.

[0064] 도12는 본 발명의 일 실시예에 따른 인물 인덱싱 테이블을 도시한 도면이다.

[0065] 도12을 참고하면, 본 발명의 인물 인덱싱 테이블은 얼굴 섬네일, 인터뷰 대상자의 이름, 소속/기관, 직업/직위, 기타 정보, 및 키워드 등을 포함할 수 있다.

[0066] 얼굴 섬네일에는 인터뷰 대상자의 얼굴 검출 결과가 자동 표시될 것이고, 인터뷰 대상자의 이름, 소속/기관, 직

업/직위에는 인물 정보 검출 결과가 자동으로 입력되고, 키워드에는 키워드 라인의 텍스트 검출 결과가 자동으로 입력되게 된다.

[0067] 본 발명의 인물 인덱싱 테이블은 이름, 소속/기관, 직업/직위를 제외한 인물 정보 검출 결과를 기타 정보에 저장할 수도 있도록 한다.

[0068] 또한 본 발명에서는 인터뷰 대상자의 인물 정보 중 이름을 기반으로 인터넷 인물DB(예를 들어, 네이버 인물정보나 위키피디아 인물정보) 검색을 수행하여, 해당 인터뷰 대상자에 관련된 인물 정보 등이 개시된 인터넷 사이트를 발견하고, 이의 링크 주소를 추가 제공할 수도 있도록 한다. 또한 획득한 인물 인덱싱 테이블 은 입력된 원본 뉴스 인터뷰 영상과 함께 데이터베이스로 저장함으로써, 차후 제3자가 해당 클립을 열람할 수도 있도록 한다.

[0069] 도13은 본 발명의 일 실시예에 따른 뉴스 인터뷰 영상의 오버레이 텍스트 기반 인물 인덱싱 장치를 설명하기 위한 도면이다.

[0070] 도13에 도시된 바와 같이, 본 발명의 인물 인덱싱 장치는 뉴스 인터뷰 영상을 입력받는 뉴스 인터뷰 영상 입력부(10), 뉴스 인터뷰 영상에서 오버레이 텍스트 영역이 포함된 서브 클립을 추출하는 서브 클립 추출부(20), 서브 클립에서 인터뷰 대상자의 얼굴 섬네일 이미지를 추출하는 얼굴 섬네일 추출부(30), 서브 클립의 오버레이 텍스트 영역을 검출한 후, 오버레이 텍스트 영역에서 인터뷰 대상자의 인물정보 및 키워드를 추출해내는 텍스트 추출부(40), 및 얼굴 섬네일 추출부(30)와 텍스트 추출부(40)을 기반으로 인물 인덱싱 테이블을 생성하여, DB(Database)(60)에 원본 뉴스 인터뷰 영상과 함께 저장하는 인물 인덱싱 테이블 생성부(50) 등을 포함할 수 있다.

[0071] 또한 본 발명의 인물 인덱싱 장치는 필요한 경우, 사용자 또는 외부 장치로부터 제공되는 데이터를 수신하거나, 장치 동작 결과를 사용자 또는 외부 장치에 실시간 제공하기 위한 다양한 데이터 입출력 장치, 인물 인덱싱 테이블 생성부(50)가 인터넷 망에 접속하여 인터뷰 대상자 관련 정보를 검색할 수 있도록, 인터넷망에 접속 가능하도록 하는 통신 장치 등을 추가로 구비할 수도 있을 것이다.

[0072] 이와 같이 구성되는 인물 인덱싱 장치는 하나의 하드웨어 장치로써 구현될 수 있으나, 필요한 경우 기존의 컴퓨팅 장치에 결합되는 임베디드 시스템 또는 기존의 컴퓨팅 장치에 다운로드 및 설치되는 소프트웨어 등으로 구현될 수도 있도록 한다.

[0073] 상술한 실시예에 설명된 특징, 구조, 효과 등은 본 발명의 적어도 하나의 실시예에 포함되며, 반드시 하나의 실시예에만 한정되는 것은 아니다. 나아가, 각 실시예에서 예시된 특징, 구조, 효과 등은 실시예들이 속하는 분야의 통상의 지식을 가지는 자에 의하여 다른 실시예들에 대해서도 조합 또는 변형되어 실시 가능하다.

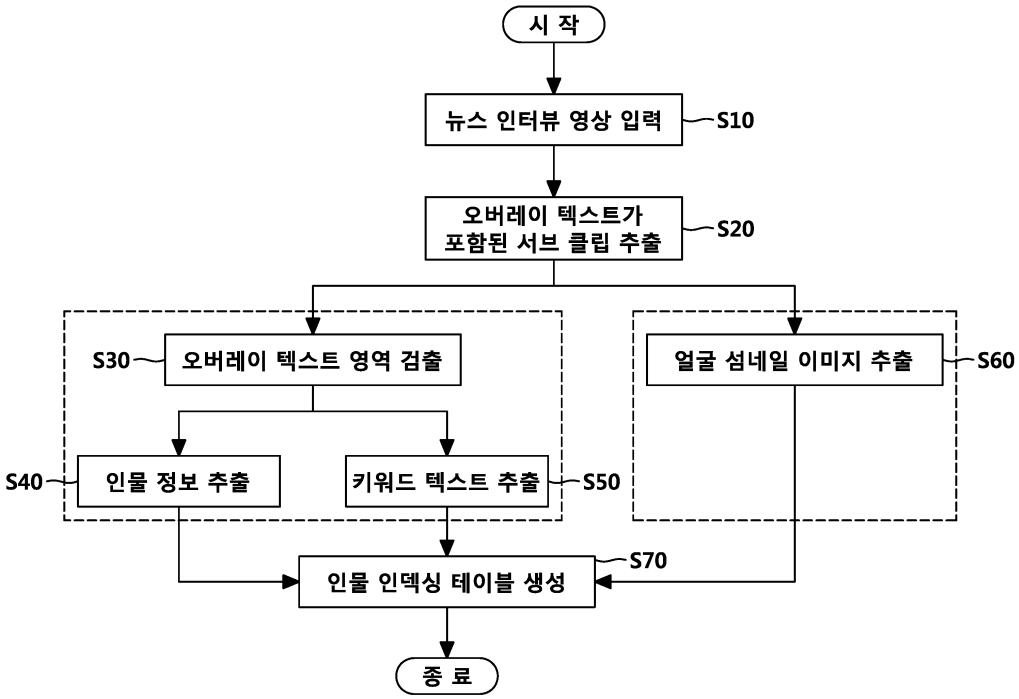
[0074] 따라서 이러한 조합과 변형에 관계된 내용들은 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다. 또한, 이상에서 실시예들을 중심으로 설명하였으나 이는 단지 예시일 뿐 본 발명을 한정하는 것이 아니며, 본 발명이 속하는 분야의 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성을 벗어나지 않는 범위에서 이상에 예시되지 않은 여러 가지의 변형과 응용이 가능함을 알 수 있을 것이다. 예를 들어, 실시예들에 구체적으로 나타난 각 구성 요소는 변형하여 실시할 수 있는 것이다. 그리고 이러한 변형과 응용에 관계된 차이점들은 첨부한 청구 범위에서 규정하는 본 발명의 범위에 포함되는 것으로 해석되어야 할 것이다.

도면

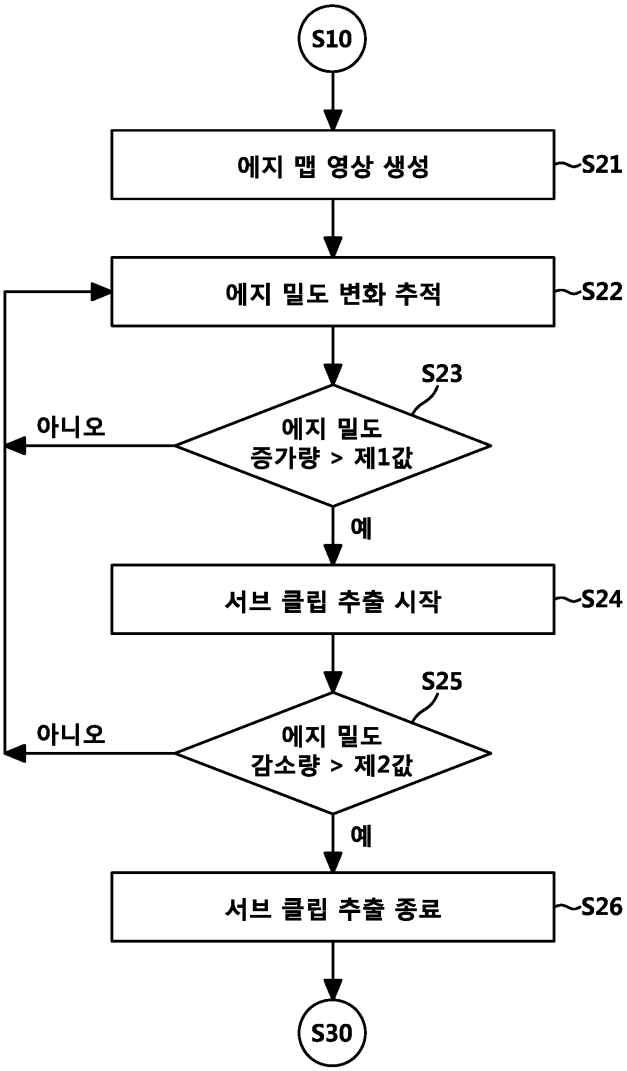
도면1


고경주 열차 탑승 수험생
혹시 못 볼까 봐 긴장되긴 긴장됐는데
시험은 제시간에 들어와서 그냥 푼 것 같아요.

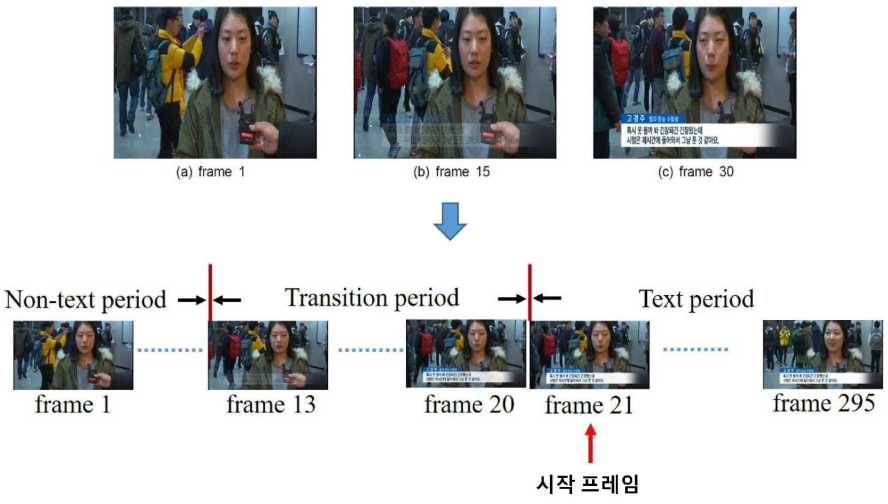
도면2


시 작
뉴스 인터뷰 영상 입력
S10
오버레이 텍스트가
포함된 서브 클립 추출
S20
S30
오버레이 텍스트 영역 검출
얼굴 섬네일 이미지 추출
S60
S40
인물 정보 추출
키워드 텍스트 추출
S50
S70
인물 인덱싱 테이블 생성
종 료

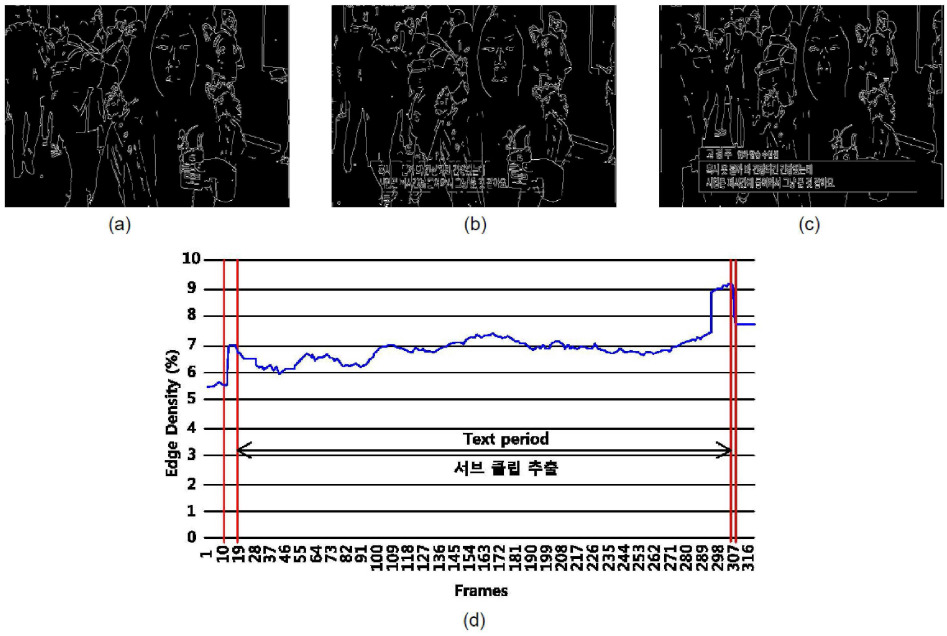
도면3


S10
에지 맵 영상 생성
S21
에지 밀도 변화 추적
S22
S23
에지 밀도
증가량 > 제1값
아니오
예
서브 클립 추출 시작
S24
S25
에지 밀도
감소량 > 제2값
아니오
예
서브 클립 추출 종료
S26
S30

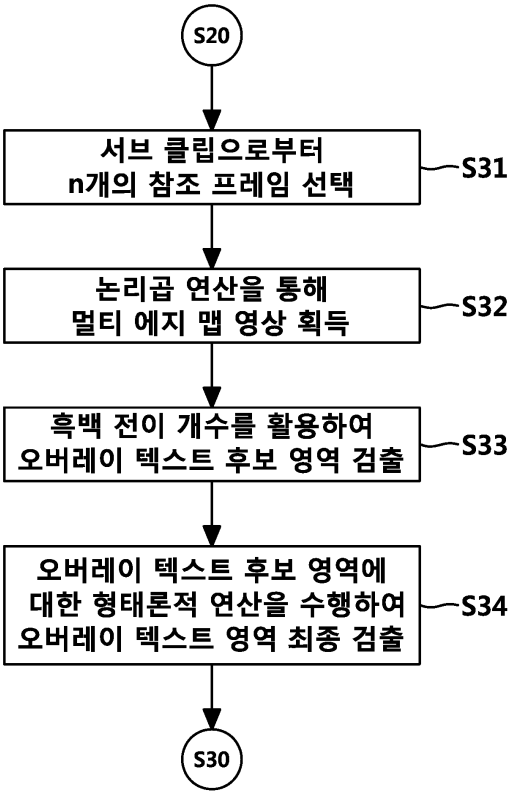
도면4


(a) frame 1
(b) frame 15
(c) frame 30
Non-text period
Transition period
Text period
frame 1
frame 13
frame 20
frame 21
frame 295
시작 프레임

도면5


(a)
(b)
(c)
Edge Density (%)
Text period
서브 클립 추출
Frames
(d)

도면6


S20
서브 클립으로부터
n개의 참조 프레임 선택
S31
논리곱 연산을 통해
멀티 에지 맵 영상 획득
S32
흑백 전이 개수를 활용하여
오버레이 텍스트 후보 영역 검출
S33
오버레이 텍스트 후보 영역에
대한 형태론적 연산을 수행하여
오버레이 텍스트 영역 최종 검출
S34
S30

도면7



(a) 1st reference frame (b) 2nd reference frame (c) 3rd reference frame (d) 4th reference frame

도면8

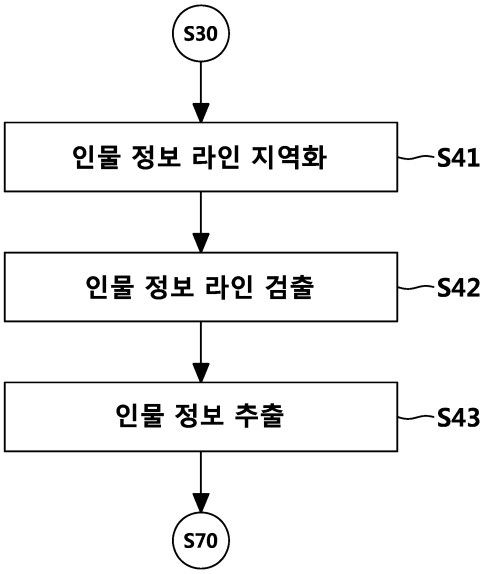


(a)

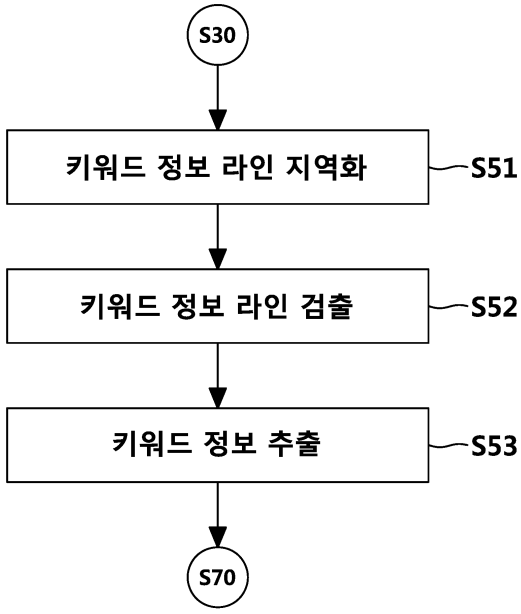


(b)

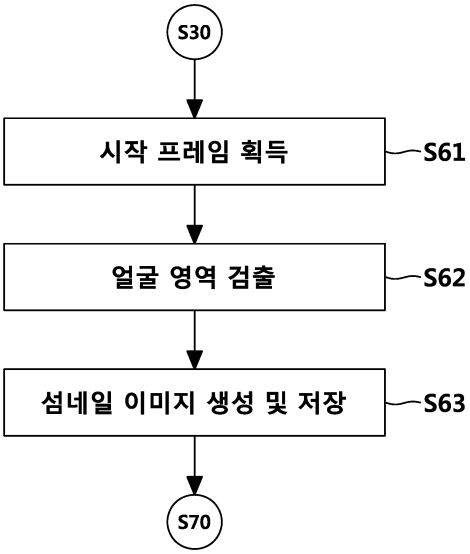
도면9



도면10


S30
키워드 정보 라인 지역화
S51
키워드 정보 라인 검출
S52
키워드 정보 추출
S53
S70

도면11


S30
시작 프레임 획득
S61
얼굴 영역 검출
S62
섬네일 이미지 생성 및 저장
S63
S70

도면12

얼굴 섬네일	이름	인터뷰자 이름
	소속 / 기관	회사명, 기관명 등의 정보
	직업 / 직위	직업, 직위와 관련된 정보
	기타	위 세가지를 제외한 정보 (나이, 성별, 주소 등)
	추가정보	인터넷 인물정보 링크 (네이버 인물정보 등)
키워드	인터뷰 내용에서 주요 단어 추출 (사건, 인명, 장소, 기관, 시간, 날짜 등과 관련된 사항)	

도면13

