



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0023492
(43) 공개일자 2015년03월05일

(51) 국제특허분류(Int. Cl.)
H04N 21/8549 (2011.01) H04N 21/462 (2011.01)
H04N 21/234 (2014.01) G06F 17/30 (2006.01)
(21) 출원번호 10-2014-7036413
(22) 출원일자(국제) 2013년06월18일
심사청구일자 없음
(85) 번역문제출일자 2014년12월24일
(86) 국제출원번호 PCT/EP2013/062568
(87) 국제공개번호 WO 2014/001137
국제공개일자 2014년01월03일
(30) 우선권주장
12305733.3 2012년06월25일
유럽특허청(EPO)(EP)

(71) 출원인
툼슨 라이센싱
프랑스 92130 이씨레물리노 잔 다르크 뒤편 1-5
(72) 발명자
와젤, 리오넬
프랑스 35576 쉐송 쉐비네 쉐에스 17616 아브뉴
데 상 블랑 975 떼끄니폴로르 에르 에 데 프랑스
제페다, 살바디에라, 조아킨
프랑스 35576 쉐송 쉐비네 쉐에스 17616 아브뉴
데 상 블랑 975 떼끄니폴로르 에르 에 데 프랑스
(뒷면에 계속)
(74) 대리인
양영준, 전경석, 백만기

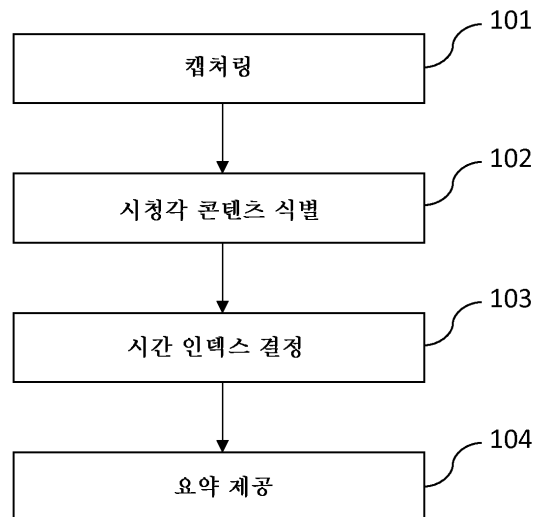
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 동기화된 영화 요약

(57) 요약

본 발명은 시청각 오브젝트의 요약을 제공하는(104) 방법에 관한 것이다. 본 방법은, 시청각 오브젝트로부터 정보를 캡처링하는 단계(101), 시청각 오브젝트를 식별하는 단계(102), 시청각 오브젝트에 대한 캡처링된 정보의 타임 인덱스를 결정하는 단계(103), 식별된 시청각 오브젝트의 부분의 요약을 제공하는 단계(104)를 포함하고, 상기 부분은 식별된 시청각 오브젝트의 시작부와 결정된 타임 인덱스 사이에 포함된다.

대표도 - 도1



(72) 발명자

슈발리에, 루이

프랑스 35576 쉐송 셰비네 셰에스 17616 아브뉴 데
상 블랑 975 페르니폴로르 에르 에 데 프랑스

페레즈, 파트릭

프랑스 35576 쉐송 셰비네 셰에스 17616 아브뉴 데
상 블랑 975 페르니폴로르 에르 에 데 프랑스

엘리에, 피에르

프랑스 35576 쉐송 셰비네 셰에스 17616 아브뉴 데
상 블랑 975 페르니폴로르 에르 에 데 프랑스

특허청구의 범위

청구항 1

시청각 오브젝트의 요약을 제공하기(104) 위한 방법으로서,

(i) 상기 시청각 오브젝트를 식별하고 상기 시청각 오브젝트에 대한 타임 인덱스를 결정할 수 있게 하는 정보를 시청각 오브젝트로부터 캡처링하는 단계(101);

(ii) 상기 시청각 오브젝트를 식별하는 단계(102);

(iii) 상기 시청각 오브젝트에 대한 캡처링된 정보의 타임 인덱스를 결정하는 단계(103); 및

(iv) 식별된 시청각 오브젝트의 부분의 요약을 제공하는 단계(104) - 상기 부분은 상기 식별된 시청각 오브젝트의 시작부와 결정된 타임 인덱스 사이에 포함됨 -

를 포함하는 요약 제공 방법.

청구항 2

제1항에 있어서,

상기 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터를 포함하는 데이터베이스(204)가 제공되고,

상기 캡처링된 정보는 상기 캡처링 시간에서의 시청각 오브젝트의 이미지의 데이터이며,

상기 타임 인덱스는, 상기 캡처링 시간에서의 시청각 오브젝트의 이미지의 데이터와 상기 데이터베이스(204)에서의 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터 간의 유사도 매칭에 따라 결정되는, 요약 제공 방법.

청구항 3

제2항에 있어서,

상기 시청각 오브젝트의 이미지의 데이터의 방식과 상기 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터의 방식은 서명 방식인, 요약 제공 방법.

청구항 4

제1항에 있어서,

상기 식별된 시청각 오브젝트의 타임-인덱싱된 오디오 신호들의 데이터를 포함하는 데이터베이스(204)가 제공되고,

상기 캡처링된 정보는 상기 캡처링 시간에서의 시청각 오브젝트의 오디오 신호의 데이터이며,

상기 타임 인덱스는, 상기 캡처링 시간에서의 시청각 오브젝트의 오디오 신호의 데이터와 상기 데이터베이스(204)에서의 식별된 시청각 오브젝트의 타임-인덱싱된 오디오 신호들의 데이터 간의 유사도 매칭에 따라 결정되는, 요약 제공 방법.

청구항 5

제2항에 있어서,

상기 시청각 오브젝트의 오디오 신호의 데이터의 방식과 상기 식별된 시청각 오브젝트의 타임-인덱싱된 오디오 신호들의 데이터의 방식은 서명 방식(signature nature)인, 요약 제공 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 캡처링 단계(101)는 모바일 디바이스(202)에 의해 수행되는, 요약 제공 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 식별 단계(102), 결정 단계(103) 및 제공 단계(104)는 전용 서버(205) 상에서 수행되는, 요약 제공 방법.

명세서

기술 분야

[0001] 본 발명은 시청각 오브젝트의 요약을 제공하기 위한 방법에 관한 것이다.

배경 기술

[0002] 시청자는 재생되고 있는 시청각 오브젝트의 초반부를 놓칠 수 있다. 그러한 문제에 직면하게 되면, 시청자는 무엇을 놓쳤는지를 알고 싶을 것이다. 미국 특허 출원 제11/568,122호는, 프로그램을 신규 세그먼트 공간에 맵핑하고 콘텐츠 부분이 그 콘텐츠 스트림의 시작 부분인지, 중간 부분인지 또는 종료 부분인지에 따른 요약 기능을 이용하여 그 프로그램에 대한 콘텐츠 스트림의 부분을 자동으로 요약함으로써 이 문제를 해결한다.

발명의 내용

과제의 해결 수단

[0003] 최종 사용자에게 그가 실제로 놓쳤던 콘텐츠에 더 바람직하게 맞춰진 요약을 제공하는 것이 본 발명의 한 목적이다.

[0004] 이를 위해, 본 발명은 시청각 오브젝트의 요약을 제공하기 위한 방법을 제안하고, 본 방법은,

[0005] (i) 시청각 오브젝트를 식별하고 시청각 오브젝트에 대한 타임 인덱스를 결정할 수 있게 하는 시청각 오브젝트로부터의 정보를 캡처링하는 단계,

[0006] (ii) 시청각 오브젝트를 식별하는 단계,

[0007] (iii) 시청각 오브젝트에 대한 캡처링된 정보의 타임 인덱스를 결정하는 단계, 및

[0008] (iv) 식별된 시청각 오브젝트의 부분의 요약을 제공하는 단계 - 상기 부분은 식별된 시청각 오브젝트의 시작부와 결정된 타임 인덱스 사이에 포함됨 - 를 포함한다.

[0009] 타임 인덱스의 결정은 사용자가 놓쳤던 시청각 오브젝트의 부분을 정확하게 평가하여, 놓쳤던 부분에 맞춰진 요약을 생성하고 제공할 수 있게 한다. 결과적으로, 사용자가 무엇을 놓쳤는지에 대한 정보를 포함하고 결정된 타임 인덱스에 한정된 요약을 사용자에게 제공하게 된다. 예를 들어, 시청각 오브젝트의 스포일러는 제공된 요약에서는 제공되지 않는다.

[0010] 본 발명은 또한 방법에 관한 것으로,

[0011] 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터를 포함하는 데이터베이스가 제공되고,

[0012] 캡처링된 정보는 캡처링 시간에서의 시청각 오브젝트의 이미지의 데이터이며,

[0013] 타임 인덱스는, 캡처링 시간에서의 시청각 오브젝트의 이미지의 데이터와 상기 데이터베이스에서의 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터 간의 유사도 매칭에 따라 결정된다.

[0014] 바람직하게, 시청각 오브젝트의 이미지의 데이터의 방식과 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터의 방식은 서명 방식(signature nature)이다.

[0015] 서명들을 이용하는 장점은, 특히 데이터가 원본 데이터보다 더 가볍게 되고, 그에 따라 더 신속하게 매칭할 수 있을 뿐만 아니라 더 신속하게 식별할 수 있게 한다는 것을 포함한다.

[0016] 대안적으로, 본 발명은 방법에 관한 것으로, 여기서,

[0017] 식별된 시청각 오브젝트의 타임-인덱싱된 오디오 신호들의 데이터를 포함하는 데이터베이스가 제공되고,

[0018] 캡처링된 정보는 캡처링 시간에서의 시청각 오브젝트의 오디오 신호의 데이터이며,

- [0019] 타임 인덱스는, 캡처링 시간에서의 시청각 오브젝트의 오디오 신호의 데이터와 상기 데이터베이스에서의 식별된 시청각 오브젝트의 타임-인덱싱된 오디오 신호들의 데이터 간의 유사도 매칭에 따라 결정된다.
- [0020] 바람직하게, 시청각 오브젝트의 오디오 신호의 데이터의 방식과 식별된 시청각 오브젝트의 타임-인덱싱된 오디오 신호들의 데이터의 방식은 서명 방식이다.
- [0021] 바람직하게, 캡처링하는 단계는 모바일 디바이스에 의해 수행된다.
- [0022] 바람직하게, 식별하는 단계, 결정하는 단계 및 제공하는 단계는 전용 서버 상에서 수행된다.
- [0023] 이러한 방식으로, 캡처링 면에서 더 적은 프로세싱 전력이 요구되고, 요약물 제공하는 프로세스가 가속화된다.
- [0024] 더 나은 이해를 위해, 본 발명은 이제 도면들을 참조하여 다음 설명에서 더 상세하게 설명될 것이다. 본 발명은 설명된 실시예들에 제한되지 않으며, 또한 첨부된 청구항들에 정의된 바와 같은 본 발명의 범위를 벗어나지 않고 특정 특징들이 편의상 결합 및/또는 수정될 수 있다.

도면의 간단한 설명

- [0025] 도 1은 본 발명에 따른 방법의 예시적인 흐름도를 도시한다.
- 도 2는 본 발명에 따른 방법의 구현을 가능하게 하는 장치의 예를 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0026] 도 2를 참조하면, 본 발명의 방법을 구현하도록 구성되는 예시적인 장치가 도시된다. 본 장치는 렌더링 디바이스(201), 캡처링 디바이스(202) 및 데이터베이스(204)를 포함하고, 선택적으로 전용 서버(205)를 포함한다. 도 1에서의 흐름도 및 도 2에서의 장치를 참조하여 본 발명의 방법에 대한 제1 바람직한 실시예가 더 상세하게 설명될 것이다.
- [0027] 렌더링 디바이스(201)는 시청각 오브젝트를 렌더링하는 데 이용된다. 예를 들어, 시청각 오브젝트는 영화이고, 렌더링 디바이스(201)는 디스플레이이다. 캡처링 수단을 갖춘 캡처링 디바이스(202)에 의해, 렌더링되는 시청각 오브젝트의 정보, 예를 들어, 디스플레이되는 영화의 이미지의 데이터가 캡처링된다(101). 그러한 디바이스(202)는 예를 들어 디지털 카메라가 장착된 모바일 폰이다. 캡처링된 정보는 시청각 오브젝트를 식별하고(102) 시청각 오브젝트에 대한 타임 인덱스를 결정하는데(103) 이용된다. 후속하여, 식별된 시청각 오브젝트의 부분의 요약이 제공되는데(104), 오브젝트의 부분은 식별된 시청각 오브젝트의 시작부와 결정된 타임 인덱스 사이에 포함된다.
- [0028] 구체적으로, 캡처링된 정보, 즉, 영화의 이미지의 데이터는 예를 들어 네트워크(203)를 통해 데이터베이스(204)로 송신된다. 이 바람직한 실시예에서, 데이터베이스(204)는 영화들의 세트와 같은 식별된 시청각 오브젝트들의 타임-인덱싱된 이미지들의 데이터를 포함한다. 바람직하게, 시청각 오브젝트의 이미지의 데이터와 데이터베이스에서의 식별된 시청각 오브젝트의 타임-인덱싱된 이미지들의 데이터는 이미지들의 서명들이다. 예를 들어, 그러한 서명은 키 포인트 서술자(key point descriptor), 예를 들어, SIFT 서술자를 이용하여 추출될 수 있다. 그 후, 시청각 오브젝트를 식별하는 단계(102) 및 캡처링된 정보의 타임 인덱스를 결정하는 단계(103)는, 캡처링 시간에서의 시청각 오브젝트의 이미지의 데이터와 데이터베이스(204)에서의 타임-인덱싱된 이미지들의 이미지의 데이터 간의, 즉 이미지들의 서명들 간의 유사도 매칭에 따라 수행된다. 캡처링 시간에서의 시청각 오브젝트의 이미지에 대해 데이터베이스(204) 내의 가장 유사한 타임-인덱싱된 이미지가 식별되면, 시청각 오브젝트를 식별하고 시청각 오브젝트에 대한 캡처링된 정보의 타임 인덱스를 결정할 수 있게 된다. 그 후, 식별된 시청각 오브젝트의 시작부와 결정된 타임 인덱스 사이에 포함되는 식별된 시청각 오브젝트의 부분의 요약이 얻어져 사용자에게 제공된다(104).
- [0029] 캡처링 수단을 갖춘 캡처링 디바이스(202)에 의해 직접, 또는 대안으로 전용 서버(205) 상에서 시청각 오브젝트의 이미지의 데이터, 예를 들어, 이미지 서명이 캡처링될 수 있다. 유사하게, 시청각 오브젝트를 식별하는 단계(102), 캡처링된 정보의 타임 인덱스를 결정하는 단계(103) 및 요약물 제공하는 단계(104)는 대안적으로 전용 서버(205) 상에서 수행될 수 있다.
- [0030] 디바이스(202) 상에서 직접적으로 이미지 서명 캡처를 수행하는 것의 장점은 전용 서버(205)로 송신되는 데이터의 무게가 메모리 면에서 보다 더 가볍다는 것이다.
- [0031] 전용 서버(205) 상에서 서명 캡처를 수행하는 것의 장점은 서명의 방식이 서버 측에서 제어될 수 있다는

것이다. 그에 따라, 시청각 오브젝트의 이미지의 서명의 방식과 데이터베이스(204)에서의 타임-인덱싱된 이미지들의 서명들의 방식이 동일하게 되고, 직접적으로 비교될 수 있게 된다.

[0032] 데이터베이스(204)는 전용 서버(205)에 위치될 수 있다. 물론 전용 서버(205) 외에 또한 위치될 수 있다.

[0033] 상기 바람직한 실시예에서, 캡처링된 정보는 이미지의 데이터이다. 더 일반화된 방식으로, 캡처링된 데이터가 시청각 오브젝트를 식별하고(102) 그 시청각 오브젝트에 대한 캡처링된 정보의 타임 인덱스를 결정하는(103) 것을 가능하게 한다면, 정보는 적용된 캡처링 수단을 프로세싱하는 캡처링 디바이스(202)에 의해 캡처링될 수 있는 임의의 데이터일 수 있다.

[0034] 본 발명의 방법에 대한 제2 바람직한 실시예에서, 캡처링된 정보는 캡처링 시간에서의 시청각 오브젝트의 오디오 신호의 데이터이다. 마이크로폰 또는 라우즈스피커가 장착된 모바일 디바이스에 의해 정보가 캡처링될 수 있다. 시청각 오브젝트의 오디오 신호의 데이터는 오디오 신호의 서명일 수 있는데, 이것은 데이터베이스(204)에 포함된 오디오 서명들의 모음(collection) 중에서 가장 유사한 오디오 서명에 매칭된다. 시청각 오브젝트를 식별하고(102) 시청각 오브젝트에 대한 캡처링된 정보의 타임 인덱스를 결정하기(103) 위해 유사도 매칭이 이용된다. 식별된 시청각 오브젝트의 부분의 요약이 후속하여 제공되는데(104), 여기서, 오브젝트의 부분은 식별된 시청각 오브젝트의 시작부와 결정된 타임 인덱스 사이에 포함된다.

[0035] 데이터베이스(204) 및 식별된 시청각 오브젝트의 부분의 요약에 대한 예가 이제 설명될 것이다. 기존의 및/또는 공공의 데이터베이스를 이용하여, 데이터베이스(204)를 생성하기 위한 오프라인 프로세스가 수행된다. 이제 영화들의 세트의 모음에 대한 예시적인 데이터베이스가 설명될 것이지만, 본 발명은 이하 설명에 제한되지 않는다.

[0036] 데이터베이스(204)의 요약 데이터베이스에 대해, 전체 영화에 시간적으로 동기화된 요약이 생성된다. 이것은, 예를 들어, 인터넷 영화 데이터베이스(IMDB) 상에서 이용가능한 것들과 같은 기존의 시놉시스(synopsis)에 의존한다. 그러한 시놉시스는 영화의 이름으로부터 직접적으로 검색될 수 있다. 예를 들어, 주어진 영화의 오디오 트랙의 트랜스크립션(transcription)을 이용하여, 주어진 영화의 텍스트 설명을 주어진 영화의 시청각 오브젝트와 동기화시킴으로써 동기화가 수행될 수 있다. 그 다음, 트랜스크립션과 텍스트 설명 모두로부터 추출된 단어들과 컨셉들의 매칭이 수행되고, 그 결과로 그 영화에 대한 동기화된 시놉시스를 얻게 된다. 동기화된 시놉시스는 물론 수동적으로 얻을 수 있다.

[0037] 선택적으로, 추가 정보가 또한 추출될 수 있다. 얼굴 검출 및 클러스터링(clustering) 프로세스가 전체 영화에 적용되고, 그에 따라 영화에서 알아볼 수 있는 얼굴들의 클러스터들을 제공하게 된다. 클러스터들 각각은 동일한 등장 인물에 대응하는 얼굴들을 포함한다. 이러한 클러스터링 프로세스는 M. Everingham, J. Sivic, and A. Zisserman "Hello! My name is ... Buffy" - Automatic naming of characters in TV video" Proceedings of the 17th British Machine Vision Conference (BMVC 2006)에 자세히 설명된 기술들을 이용하여 수행될 수 있다. 그 후, 특정 등장 인물의 존재에 연관된 영화 시간 코드들의 리스트와 연관된 등장 인물들의 리스트가 얻어진다. 더 양호한 클러스터링 결과를 위해, 주어진 영화의 IMDB 등장 인물 리스트에 비교하여 얻어진 클러스터들이 매칭될 수 있다. 이 매칭 프로세스는 수동적인 단계들을 포함할 수 있다.

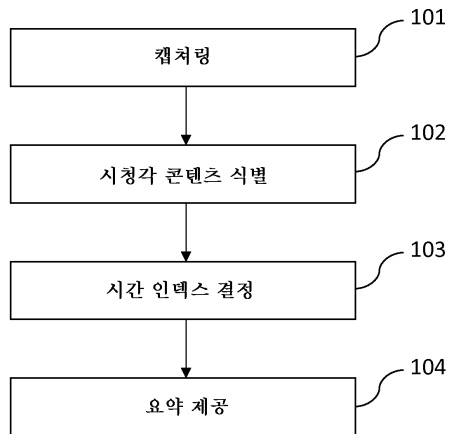
[0038] 얻어진 동기화된 시놉시스 요약 및 클러스터 리스트들은 데이터베이스(204)에 저장된다. 데이터베이스(204) 내의 영화들은 복수의 프레임들로 분할되고, 각각의 프레임들이 추출된다. 영화에 대한 캡처링된 정보의 타임 인덱스를 결정하는 단계(103)와 같은 후-동기화(post-synchronization) 프로세스들을 용이하게 하도록 영화의 프레임들이 인덱싱된다. 대안적으로, 영화의 각각의 프레임을 추출하는 대신에, 프로세싱될 총 데이터를 감소시키기 위해 충분한 서브-샘플링에 의한 일부 프레임들만이 추출된다. 각각의 추출된 프레임에 대해, 이미지 서명, 예를 들어, 키 포인트 설명에 기초한 핑거프린트가 생성된다. H. Jegou, M. Douze, and C. Schmid - Hamming embedding and weak geometric consistency for large scale image search - ECCV, October 2008에 설명된 기술들을 이용하여 실시될 수 있는 효율적인 방식으로, 그러한 키 포인트들 및 그와 연관된 서술자들이 인덱싱된다. 그 다음, 이미지 서명들과 연관된 영화들의 프레임들은 데이터베이스(204)에 저장된다.

[0039] 식별된 시청각 오브젝트(즉, 영화)의 부분의 요약을 얻기 위해, 시청각 오브젝트의 정보, 예를 들어, 그의 이미지의 데이터가 캡처링 디바이스(202)에 의해 캡처링된다. 그 정보는 데이터베이스(204)에 송신되고 시청각 오브젝트를 식별하기 위해 데이터베이스(204)와 비교하게 된다. 예를 들어, 영화의 프레임은 캡처링된 정보에 대응하는 데이터베이스(204)에서 식별된다. 식별된 프레임은 캡처링된 정보와 데이터베이스(204) 내의 동기화된 시놉시스 요약 간의 매칭을 용이하게 하고, 그에 따라 영화에 대한 캡처링된 정보의 타임 인덱스를 결정하게 된

다. 영화의 부분의 동기화된 요약이 사용자에게 제공되는데, 여기서, 영화의 부분은 식별된 영화의 시작부와 결정된 타임 인덱스 사이에 포함된다. 예를 들어, 요약은 모바일 디바이스(202)에 디스플레이됨으로써 사용자가 읽을 수 있게 제공될 수 있다. 선택적으로, 요약은 영화의 부분에 나타나는 등장 인물들의 클러스터 리스트들을 포함할 수 있다.

도면

도면1



도면2

