



(19) 대한민국특허청(KR)
 (12) 공개특허공보(A)

(11) 공개번호 10-2015-0091053
 (43) 공개일자 2015년08월07일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01)

(71) 출원인

톰슨 라이센싱

(52) CPC특허분류
G06F 17/30796 (2013.01)
G06F 17/30825 (2013.01)

프랑스 92130 이씨레물리노 잔 다르크 뤼 1-5

(21) 출원번호 10-2015-7013931

(72) 발명자

장, 얀펭

(22) 출원일자(국제) 2012년11월30일

중국 100192 베이징 하이 디안 디스트릭트 슈에
칭 로드 8 테크놀로지 포춘 센터 빌딩 에이 8층
테크니컬러 (중국) 테크니컬러 씨오. 엘디티

심사청구일자 없음

장, 지강

(85) 번역문제출일자 2015년05월27일

중국 100192 베이징 하이 디안 디스트릭트 슈에
칭 로드 8 테크놀로지 포춘 센터 빌딩 에이 8층
테크니컬러 (중국) 테크니컬러 씨오. 엘디티

(86) 국제출원번호 PCT/CN2012/085637

슈, 준
중국 100192 베이징 하이 디안 디스트릭트 슈에
칭 로드 8 테크놀로지 포춘 센터 빌딩 에이 8층
테크니컬러 (중국) 테크니컬러 씨오. 엘디티

(87) 국제공개번호 WO 2014/082288

(74) 대리인

국제공개일자 2014년06월05일

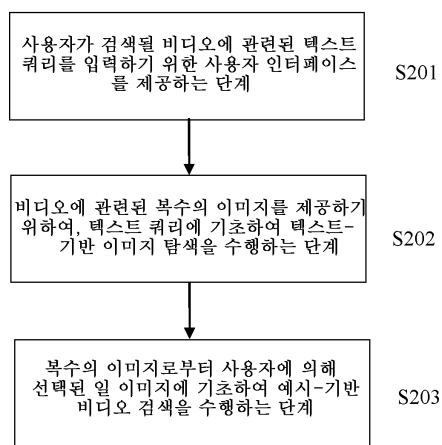
양영준, 전경석, 백만기

전체 청구항 수 : 총 14 항

(54) 발명의 명칭 비디오 검색을 위한 방법 및 장치

(57) 요약

본 발명은 비디오 검색을 위한 방법 및 장치를 제공한다. 본 방법은 사용자가 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위한 사용자 인터페이스를 제공하는 단계; 비디오에 관련된 복수의 이미지를 제공하기 위하여, 텍스트 쿼리에 기초하여 텍스트-기반 이미지 탐색을 수행하는 단계; 및 복수의 이미지로부터 사용자에 의해 선택된 일 이미지에 기초하여 예시-기반 비디오 검색을 수행하는 단계를 포함한다.

대 표 도 - 도2

명세서

청구범위

청구항 1

비디오 검색을 위한 방법으로서,

사용자가 검색될 비디오에 관련된 텍스트 쿼리(text query)를 입력하기 위한 사용자 인터페이스를 제공하는 단계(S201);

상기 비디오에 관련된 복수의 이미지를 제공하기 위하여, 상기 텍스트 쿼리에 기초하여 텍스트-기반 이미지 탐색(text-based image searching)을 수행하는 단계(S202); 및

상기 복수의 이미지로부터 상기 사용자에 의해 선택된 일 이미지에 기초하여 예시-기반 비디오 검색(example-based video retrieval)을 수행하는 단계(S203)

를 포함하는 방법.

청구항 2

제1항에 있어서,

상기 사용자 인터페이스는 비디오 쿼리 다이얼로그인, 방법.

청구항 3

제1항에 있어서,

상기 텍스트-기반 이미지 탐색은 상기 텍스트 쿼리와 이미지의 메타데이터 사이의 텍스트 매칭에 의해 실행되는, 방법.

청구항 4

제3항에 있어서,

상기 메타데이터는 상기 이미지의 텍스트 주석(text annotation), 주변 텍스트(surrounding text) 및 텍스트 태그를 포함하는, 방법.

청구항 5

제1항에 있어서,

상기 예시-기반 비디오 검색은 상기 사용자에 의해 선택된 상기 이미지의 특징과 비디오의 키 프레임의 특징 사이의 이미지 유사성 매칭(image similarity matching)에 의해 실행되는, 방법.

청구항 6

제5항에 있어서,

상기 특징은 상기 키 프레임의 이미지 픽셀들로부터 추출되는 색, 질감 및 모양을 포함하는, 방법.

청구항 7

제1항에 있어서,

상기 예시-기반 비디오 검색의 결과를 상기 결과 내의 비디오의 관련성의 순위에 따라 상기 사용자에게 제시하는 단계

를 더 포함하는 방법.

청구항 8

비디오 검색을 위한 장치(500)로서,

사용자가 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위한 사용자 인터페이스를 제공하기 위한 수단(501);
상기 비디오에 관련된 복수의 이미지를 제공하기 위하여, 상기 사용자에 의해 입력된 상기 텍스트 쿼리에 기초하여 이미지 데이터베이스 내에서 텍스트-기반 이미지 탐색을 수행하기 위한 수단(502); 및

상기 복수의 이미지로부터 상기 사용자에 의해 선택된 일 이미지에 기초하여 비디오 데이터베이스 내에서 예시-기반 비디오 검색을 수행하기 위한 수단(503)

을 포함하는 장치(500).

청구항 9

제8항에 있어서,

상기 사용자 인터페이스는 비디오 쿼리 다이얼로그인, 장치(500).

청구항 10

제8항에 있어서,

상기 이미지 데이터베이스는 외부 데이터베이스이고,

텍스트-기반 이미지 탐색을 수행하기 위한 수단(502)은 상기 이미지 데이터베이스에 대한 애플리케이션 프로그래밍 인터페이스(Application Programming Interface)를 포함하는, 장치(500).

청구항 11

제8항에 있어서,

예시-기반 비디오 검색을 수행하기 위한 수단(503)은 상기 사용자에 의해 선택된 상기 이미지의 특징과 상기 비디오 데이터베이스 내의 비디오의 키 프레임의 특징 사이의 이미지 유사성 매칭을 수행하는, 장치(500).

청구항 12

제11항에 있어서,

상기 예시-기반 비디오 검색은 상기 사용자에 의해 선택된 상기 이미지의 특징과 비디오의 키 프레임의 특징 사이의 이미지 유사성 매칭에 의해 실행되는, 장치(500).

청구항 13

제12항에 있어서,

상기 특징은 상기 키 프레임의 이미지 픽셀들로부터 추출되는 색, 질감 및 모양을 포함하는, 장치(500).

청구항 14

제8항에 있어서,

상기 예시-기반 비디오 검색의 결과를 상기 사용자에게 디스플레이하기 위한 수단을 더 포함하는, 장치(500).

발명의 설명

기술 분야

[0001] 본 발명은 비디오 검색을 위한 방법 및 장치에 관한 것이다.

배경 기술

[0002] 구글 비디오 탐색(Google video searching), 유튜브(Youtube) 등과 같은 종래의 비디오 검색 시스템들은 오로지 사용자들에 의해 입력된 텍스트식 쿼리들(textual queries)에 의존한다. 사용자에 의해 입력된 탐색 텍스트(즉, 키워드)에 기초하여, 종래의 비디오 검색 시스템은 제목, 주석 또는 주변 텍스트에 대한 텍스트 매

칭(text matching)을 실행함으로써 관련 비디오 자료들을 탐색할 것이다. 이러한 텍스트-기반 방법은 2개의 단점을 가진다. 하나는 사용자들이 종종 이러한 텍스트 정보를 입력하는 것, 특히 전체 비디오 문서를 위한 상세한 설명을 입력하는 것을 주저하는 것이다. 다른 단점은 입력된 주석들의 질이 보통 좋지 않고, 이를 대부분은 단지 비디오 문서의 매우 간략한 설명을 제공하는 것이다.

[0003] 카네기 멜론 대학의 인포미디어 디지털 비디오 라이브러리 프로젝트(Informedia Digital Video Library project)(<http://www.informedia.cs.cmu.edu/>)와 같은 낮은-레벨의 콘텐츠-기반 비디오 검색에 대한 많은 연구 활동들이 행해져 왔다. 이 프로젝트는 탐색, 검색, 시각화(visualization) 및 요약(summarization)의 모든 측면들을 포함하는, 비디오 및 필름 매체의 머신 이해(machine understanding)를 성취하려고 노력한다. 발전된 기반 기술은 지능형 탐색 및 이미지 검색을 위해 선형 비디오를 자동적으로 표기(transcribe)하고, 세그먼트화 하며, 인텍싱하기 위하여 음성, 이미지 및 자연 언어 이해(natural language understanding)를 결합한다.

[0004] 예시-기반 탐색 방법들은 낮은-레벨의 콘텐츠-기반 멀티미디어 검색에서 사용자들의 탐색 의도를 설명하기 위해 광범위하게 연구되어 왔다. 예를 들어, 이미지 예시 또는 멜로디의 클립을 이용하여, 유사한 그림들 또는 멜로디를 포함하는 전체 음악이 대응하는 멀티미디어 데이터베이스로부터 검색될 수 있다. 그러나, 낮은-레벨의 콘텐츠-기반 비디오 검색에서, 사용자들이 그들의 비디오 탐색 의도를 설명하고 제시하는 것은 어렵다. 사람들에게 가장 편리한 방식은 그것을 제시하기 위하여 단어들 또는 문장들을 사용하는 것이다. 추가로, 많은 실제의 애플리케이션들에서, 사용자의 정보 필요들을 설명하기 위한 예시를 찾는 것은 힘들다. 이에 따라, 낮은-레벨의 콘텐츠-기반 비디오 검색에서, 사용자들의 의도 설명과 검색 시스템의 능력 사이에는 이해할 큰 의미적 갭(semantic gap)이 존재한다. 콘텐츠-기반 비디오 검색 방법들은 주로 입력된 예시 쿼리에 기초하는 반면, 사용자들은 대부분 그들의 텍스트-스타일 쿼리를 요구하는 것을 선호한다. 사용자들이 비디오 검색을 위한 적합한 쿼리 예시를 만들거나 찾는 것은 어렵다.

[0005] 낮은-레벨의 특징들과 사용자의 탐색 의도 사이의 의미적 갭을 연결하기 위하여, 수동적으로(manually) 입력한 주석에 의한 또는 자동적인 콘텐츠 인식에 의한 텍스트를 사용하여 멀티미디어에 주석을 달기 위한 연구 활동들이 행해져 왔다. 수동적 주석은 텍스트-기반 검색과 동일한 결점들을 가진다. 머신 자동 주석(machine automatic annotation)은 매우 어렵고, 이는 가까운 시일 내에 해결되지 않을 것으로 보인다. 추상적 키워드들은 이미지 콘텐츠와 상호관련시키는 것이 거의 불가능하다.

발명의 내용

[0006] 본 발명의 일 태양에 따르면, 비디오 검색을 위한 방법이 제공된다. 본 방법은 사용자가 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위한 사용자 인터페이스를 제공하는 단계; 비디오에 관련된 복수의 이미지를 제공하기 위하여, 텍스트 쿼리에 기초하여 텍스트-기반 이미지 탐색(text-based image searching)을 수행하는 단계; 및 복수의 이미지로부터 사용자에 의해 선택된 일 이미지에 기초하여 예시-기반 비디오 검색(example-based video retrieval)을 수행하는 단계를 포함한다.

[0007] 본 발명의 일 태양에 따르면, 비디오 검색을 위한 장치가 제공된다. 본 장치는 사용자가 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위한 사용자 인터페이스를 제공하기 위한 수단; 비디오에 관련된 복수의 이미지를 제공하기 위하여, 사용자에 의해 입력된 텍스트 쿼리에 기초하여 이미지 데이터베이스 내에서 텍스트-기반 이미지 탐색을 수행하기 위한 수단; 및 복수의 이미지로부터 사용자에 의해 선택된 일 이미지에 기초하여 비디오 데이터베이스 내에서 예시-기반 비디오 검색을 수행하기 위한 수단을 포함한다.

[0008] 본 발명의 보다 많은 태양들 및 이점들이 본 발명의 후술하는 상세한 설명에서 찾을 수 있음이 이해될 것이다.

도면의 간단한 설명

[0009] 첨부하는 도면들은 실시예들의 원리를 설명하는 역할을 하는 설명과 함께 본 발명의 실시예들의 추가적 이해를 제공하기 위하여 포함된다. 본 발명은 실시예들에 한정되지 않는다.

도면들에서:

도 1은 본 발명의 일 실시예에 따른 비디오 검색을 위한 시스템을 도시하는 예시적 도해;

도 2는 본 발명의 일 실시예에 따른 비디오 검색을 위한 방법의 흐름도;

도 3은 사용자가 텍스트 쿼리를 입력하기 위한 비디오 쿼리 다이얼로그를 도시하는 예시적 도해;

도 4는 텍스트-기반 이미지 탐색을 위해 사용될 수 있는 메타데이터를 가진 플리커(Flickr) 내의 사진의 예시를 도시하는 예시적 도해; 및

도 5는 본 발명의 일 실시예에 따른 비디오 검색을 위한 장치의 블록도.

발명을 실시하기 위한 구체적인 내용

[0010] 본 발명의 일 실시예는 이제부터 도면들과 함께 상세히 설명될 것이다. 후술하는 설명에서, 공지된 기능들 및 구성들의 일부 상세한 설명들은 간명화를 위해 생략될 수 있다.

[0011] 종래 기술들에서의 위의 문제를 고려하여, 본 발명의 일 실시예는 비디오 검색을 위한 방법 및 장치를 제공한다.

[0012] 도 1은 본 발명의 일 실시예에 따른 비디오 검색을 위한 시스템을 도시하는 예시적 도해이다.

[0013] 도 1에 도시된 것과 같이, 본 발명의 일 실시예에 따른 비디오 검색 시스템은 비디오와 관련된 복수의 이미지를 제공하기 위하여 먼저 텍스트-기반 이미지 탐색을 가지도록 제안하고, 일 이미지는 비디오 검색의 출력을 제공하기 위한 예시-기반 비디오 검색을 수행하기 위하여 사용자에 의해 복수의 이미지로부터 선택된다.

[0014] 다음으로, 본 발명의 실시예가 보다 상세히 설명될 것이다.

[0015] 도 2는 본 발명의 일 실시예에 따른 비디오 검색을 위한 방법의 흐름도이다.

[0016] 도 2에 도시된 것과 같이, 본 발명의 일 실시예에 따른 비디오 검색을 위한 방법은 후술하는 단계들을 포함한다:

[0017] S201: 사용자가 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위한 사용자 인터페이스를 제공하는 단계;

[0018] S202: 비디오에 관련된 복수의 이미지를 제공하기 위하여, 텍스트 쿼리에 기초하여 텍스트-기반 이미지 탐색을 수행하는 단계;

[0019] S203: 복수의 이미지로부터 사용자에 의해 선택된 일 이미지에 기초하여 예시-기반 비디오 검색을 수행하는 단계.

[0020] 다음으로, 본 발명의 실시예에 따른 비디오 검색을 위한 방법이 상세히 설명될 것이다.

[0021] 위의 단계(S101)에서, 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위하여, 사용자 인터페이스가 비디오 검색의 사용자에게 제공될 수 있다. 일 예시로서, 사용자 인터페이스는 사용자가 비디오에 관련된 텍스트 쿼리를 입력하기 위한 비디오 쿼리 다이얼로그일 수 있다. 도 3은 사용자가 텍스트 쿼리를 입력하기 위한 비디오 쿼리 다이얼로그를 도시하는 예시적 도해이다. 사용자 인터페이스의 다른 적절한 형태들이 또한 적용될 수 있음이 이해될 것이다. 텍스트 쿼리는 단어들 또는 문장들의 형태의 비디오의 콘텐츠의 설명이다. 텍스트 쿼리를 사용하는 이유는 보통 비디오 검색의 사용자가 자신의 의도를 표현하기 위한 가장 편리한 방식이 이미지 예시들을 준비하거나 목표를 스케치하는 것이 아닌, 텍스트 설명을 사용하는 것이기 때문이다.

[0022] 단계(S102)에서, 텍스트-기반 이미지 탐색은 비디오에 관련된 복수의 이미지를 제공하기 위하여 사용자에 의해 입력된 텍스트 쿼리에 기초하여 수행된다. 텍스트-기반 이미지 탐색은 이미지 공유 소셜 네트워크들(image sharing social networks) 및 이미지 탐색 엔진들과 같은 외부 이미지 데이터베이스 상에서, 또는 사용자의 소유의 이미지 예시 라이브러리와 같은 내부 이미지 데이터베이스 상에서 실행될 수 있다. 외부 이미지 데이터베이스가 사용될 때, 데이터베이스에 의해 요청된 API(Application Programming Interface)가 사용되어야 하는 것이 이해될 것이다. 이에 대하여 임의의 적절한 기술들이 텍스트-기반 이미지 탐색을 위해 사용될 수 있음이 유의 되어야 한다.

[0023] 플리커는 텍스트-기반 이미지 탐색을 위해 사용될 수 있는 이미지 공유 소셜 네트워크들 중 하나이다. 플리커가 단계(S102)에서 사용될 때, 텍스트-기반 이미지 탐색은 예를 들어, 플리커 내에서 사진 제공자들에 의해 부가된 이미지 주석에 관한 텍스트 매칭에 의해 실행될 수 있다. 플리커 내의 사진들은 기술적 상세들부터 더 주관적인 정보에까지 이르는 메타데이터의 상이한 타입들을 포함한다. 낮은 레벨에서, 정보는 카메라, 셔터 속도, 회전 등에 관한 것이다. 높은 레벨에서, 플리커에 사진을 업로드 하였던 사용자는 제목 및 관련 설명을 부가할 수 있고, 이는 사진 내의 이미지를 전체로서 설명하기 위하여 사용될 개연성이 더 크다. 도 4는 텍스트-기반 이미지 탐색을 위해 사용될 수 있는 메타데이터를 가진 플리커 내의 사진의 예시를 도시하는 예시적 도해이다. 백조의 사진은 아마도 이미지 제공자에 의해 부가된 사진의 제목 및 관련 설명과 함께 도 4에 도시된다.

사용자에 의해 입력된 텍스트 쿼리와 사진의 제목 및 관련 설명 사이의 텍스트 매칭은 전화(phone) 내의 이미지가 검색될 비디오와 관련이 있는지 여부를 추정하기 위하여 수행된다.

[0024] 알려진 이미지 탐색 엔진들은 구글 이미지 탐색(Google Image Searching), 야후 이미지(Yahoo Image), 빙 이미지(Bing Image) 등을 포함한다. 단계(S102)에서 구글 이미지 탐색이 사용될 때, 텍스트-기반 이미지 탐색은 예를 들어, 구글 이미지 탐색에 의해 탐색된 주변 텍스트에 의해 실행될 수 있다. 이미지를 포함하는 웹페이지 내의 텍스트는 위에서 언급된 주변 텍스트의 일 예시이다. 구글 이미지 탐색은 이미지들의 주변 텍스트 정보가 사용자에 의해 입력된 키워드 쿼리와 관련성을 가지는 이미지들을 찾는 것을 시도한다.

[0025] 텍스트-기반 이미지 탐색이 내부 이미지 데이터베이스 상에서 실행될 때, 내부 이미지 데이터베이스의 구축자에 의해 부가된 텍스트 주석 및 텍스트 태그들이 사용될 수 있다. 태그들의 사용은 구축자가 단순한 키워드 조합들을 사용하여 그가 생각하는 어떤 것이 이미지에 관련 있는지를 설명하는 것을 허용한다.

[0026] 하나의 관련 이미지는 복수의 이미지를 포함할 수 있는 단계(S102)의 탐색 결과로부터 다음의 비디오 검색을 위한 입력으로서 선택될 수 있다. 이에 대하여, 일부 이미지 공유 소셜 네트워크들 및 이미지 탐색 엔진들은 이미지들의 관련성에 따라 텍스트-기반 이미지 탐색 결과들을 위한 순위 메커니즘(ranking mechanism)을 제공할 수 있기 때문에, 자동적으로 관련 이미지를 선택하는 것이 가능하다. 그러나, 바람직하게, 단계(S102)의 탐색 결과는 사용자가 다음의 비디오 검색을 위한 입력으로서 가장 관련 있는 이미지를 검색하고 선택하기 위하여, 적절한 사용자 인터페이스와 함께 사용자에게 디스플레이된다. 본 발명의 실시예가 사용자에 의한 수동적 선택을 추천하는 이유는 머신(이미지 공유 소셜 네트워크들 및 이미지 탐색 엔진들)이 완전히 쿼리 의도를 이해하고 사용자보다 더 나은 가장 관련 있는 이미지를 선택하는 것이 여전히 매우 어렵기 때문이다.

[0027] 사용자가 단계(S101)의 결과 내의 임의의 이미지들에 만족하지 못하는 경우, 프로세스는 사용자가 텍스트 쿼리를 수정하거나 새로운 텍스트 쿼리를 입력하기 위하여 단계(S101)로 돌아갈 수 있음이 이해될 것이다.

[0028] 다음으로 단계(S103)에서, 예시-기반 비디오 검색이 사용자에 의해 선택된 이미지에 기초하여 수행된다.

[0029] 예를 들어 구어 문서 검색(spoken document retrieval), VOCR(Video Optical Character Recognition) 및 이미지 유사성 매칭(image similarity matching)을 포함하는 일부 종래의 방법들은 예시-기반 비디오 검색의 목적을 위해 개발되어왔다.

[0030] 구어 문서 검색에서, 비디오로부터의 오디오 콘텐츠의 텍스트식 표현은 자동 음성 인식을 통해 입수될 수 있다. 그러나 구어 문서 검색의 사용의 한계는 비디오 자료를 내에서 명확하고 인식 가능한 보이스가 요구되는 점이다.

[0031] VOCR에서, 비디오의 텍스트식 표현은 비디오 이미지 내에서 제시되고 있는 텍스트를 판독함으로써 입수된다. 다음으로 검색은 텍스트(키워드)에 기초하여 수행된다. 그러나 VOCR을 적용하기 위해서는, 비디오 내에 일부 인식 가능한 텍스트 정보가 존재하여야 한다. 그것이 VOCR의 사용을 위한 하나의 한계이다.

[0032] 이미지 유사성 매칭은 비디오 검색 분야로 적용되어온 예시-기반 이미지 검색 방법이다. 이미지 유사성 매칭의 이미지 탐색 엔진은 신중하게 준비된 이미지 예시를 받아들일 수 있고, 다음으로 이미지 데이터베이스로부터 유사한 이미지들을 찾기 위하여 그 예시를 사용할 수 있다. 방법이 비디오 검색에서 사용될 때, 이미지 예시는 비디오로부터 추출되었던 유사한 키 프레임들(key frames)을 찾기 위하여 사용된다. 지금까지, 어떻게 2개의 이미지의 유사성을 평가하는지에 대한 대규모의 표준화된 방법이 없었다. 이에 대하여 대부분의 사용되는 방법들은 이미지 픽셀들로부터 추출된 색, 질감 및 모양과 같은 특징들에 기초한다.

[0033] 위의 방법들은 비디오 검색을 위한 보다 복합적 방법을 형성하기 위하여 결합될 수 있음이 이해될 것이다.

[0034] 본 발명의 실시예에서, 비디오 검색에 대한 입력이 단계(S102)의 탐색 결과로부터 사용자에 의해 선택된 이미지를 포함하기 때문에, 예시-기반 비디오 검색에 이미지 유사성 매칭 방법을 적용하는 것이 바람직하다.

[0035] 다음으로, 상세한 설명은 이미지 유사성 매칭 방법에 의한 예시-기반 비디오 검색에 대하여 주어질 것이다.

[0036] 비디오가 데이터베이스에 저장되기 전에, 세그먼트 및 키 프레임 감지(key frame detection)를 포함하는 비디오 구조 분석(video structure parsing)의 대상이 될 것은 알려져 있다. 세그먼트는 비디오를 개개의 장면들(scenes)로 자르기 위하여 사용된다. 각각의 장면은 연속하는 프레임들의 시리즈들로 구성되고, 동일한 장소에서 촬영되거나 주제 콘텐츠를 공유하는 프레임들은 함께 그룹화될 것이다. 키 프레임 감지는 인테리 이미지로서 개개의 장면으로부터 전형적 이미지를 찾는 것이다. 종래의 비디오 세그먼트 및 키 프레임 추출 알고리즘들은 이에 대하여 사용될 것이다. 예를 들어, 장면 전환 감지 알고리즘(shot boundary detection algorithm)은

비디오 프레임들 내에 포함된 시각적 정보에 의존하여 비디오를 유사한 시각적 콘텐츠들을 가진 프레임들로 분할할 수 있는 그러한 해결책이다. 키 프레임의 추출 후, 메타데이터는 각각의 키 프레임에 부가된다. 메타데이터는 키 프레임이 추출되었던 비디오 및 특정 비디오 내의 키 프레임의 구체적 위치를 제시한다.

[0037] 다음으로, 탐색 쿼리(사용자에 의해 선택된 이미지)의 특징들과 데이터베이스 내에 저장된 비디오의 키 프레임들의 특징들 사이의 유사성의 정도는 매칭 알고리즘을 사용하여 계산될 수 있고, 이는 검색된 비디오의 관련성의 순위를 결정한다. 본 기술분야에서 공지된 종래의 이미지 매칭 알고리즘이 있다. 콘텐츠-기반 이미지 검색을 위한 종래의 방법들은 벡터 모델에 기초한다. 이를 방법들에서, 이미지는 특징들의 집합에 의해 표현되고, 2개의 이미지 사이의 차이는 그들의 특징 벡터들 사이의 거리, 보통 유클리드 거리(Euclidean distance)를 통해 측정된다. 거리는 2개의 이미지의 유사성 정도를 결정하고, 또한 대응하는 비디오의 순위를 결정한다. 대부분의 이미지 검색 시스템들은 이미지 픽셀들로부터 추출된 색, 질감 및 모양과 같은 특징들에 기초한다.

[0038] 유사한 키 프레임들이 발견되고 순위 결정된 후, 비디오 구조 분석 단계에서 부가된 메타데이터는 어떤 비디오들이 검색되어야 하는지, 각각의 비디오의 올바른 첫번째 프레임, 및 사용자의 쿼리와 각각의 비디오 사이의 관련성의 순위들을 결정하기 위하여 사용될 수 있다. 다음으로, 대응하는 순위에 따라 배열될 수 있는 검색된 비디오 문서들의 목록은 사용자에게 제시된다.

[0039] 도 5는 본 발명의 일 실시예에 따른 비디오 검색을 위한 장치의 블록도이다.

[0040] 도 5에 도시된 것과 같이, 비디오 검색을 위한 장치(500)는 사용자가 검색될 비디오에 관련된 텍스트 쿼리를 입력하기 위한 사용자 인터페이스를 제공하기 위한 사용자 인터페이스 제공 유닛(501); 비디오에 관련된 복수의 이미지를 제공하기 위하여, 사용자에 의해 입력된 텍스트 쿼리에 기초하여 이미지 데이터베이스 내에서 텍스트-기반 이미지 탐색을 수행하기 위한 이미지 탐색 유닛(502); 및 복수의 이미지로부터 사용자에 의해 선택된 일 이미지에 기초하여 비디오 데이터베이스 내에서 예시-기반 비디오 검색을 수행하기 위한 비디오 검색 유닛(503)을 포함한다.

[0041] 일 예로서, 사용자 인터페이스 제공 유닛(501)은 사용자가 비디오에 관련된 텍스트 쿼리를 입력하기 위한 비디오 쿼리 다이얼로그를 제공할 수 있다.

[0042] 비디오 검색을 위한 방법에서 설명된 것과 같이, 이미지 데이터베이스는 사용자의 이미지 예시 라이브러리와 같은 내부 이미지 데이터베이스일 수 있다. 이미지 데이터베이스는 또한 이미지 공유 소셜 네트워크들 및 이미지 탐색 엔진들과 같은 외부 이미지 데이터베이스일 수 있다. 이 경우, 이미지 탐색 유닛(502)은 외부 이미지 데이터베이스에 의해 요청된 대응하는 API(Application Programming Interface)를 구비한다.

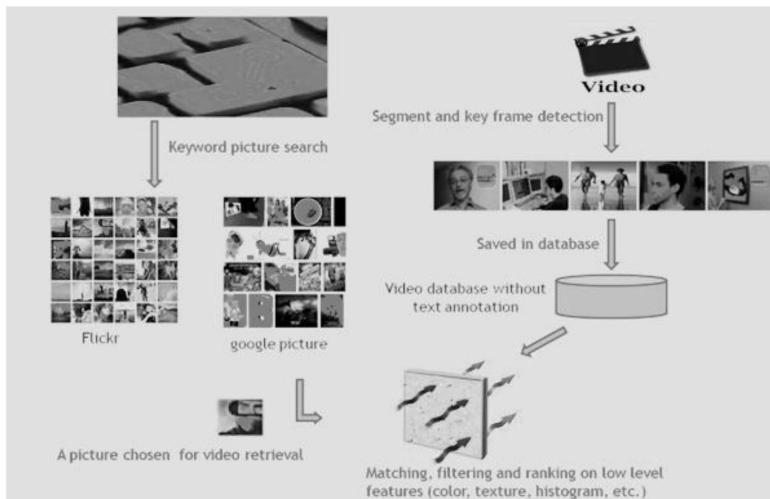
[0043] 비디오 검색 유닛(503)은 이미지 유사성 매칭 알고리즘에 의한 예시-기반 비디오 검색을 수행한다. 이 경우, 비디오 데이터베이스 내의 비디오의 키 프레임들은 키 프레임이 추출되었던 비디오 및 특정 비디오 내의 키 프레임의 구체적 위치를 제시하는 메타데이터를 구비할 필요가 있다. 메타데이터는 데이터베이스에 저장되기 전에, 비디오 데이터에 행해지는 비디오 구조 분석에 의해 입수될 수 있다.

[0044] 비디오 검색을 위한 장치(500)는 또한 예시-기반 비디오 검색 결과를 사용자에게 적절한 형식으로 디스플레이하기 위하여 디스플레이 유닛을 포함할 수 있다. 비디오 검색의 결과는 결과 내에서 비디오의 관련성의 순위에 따라 사용자에게 디스플레이될 수 있다.

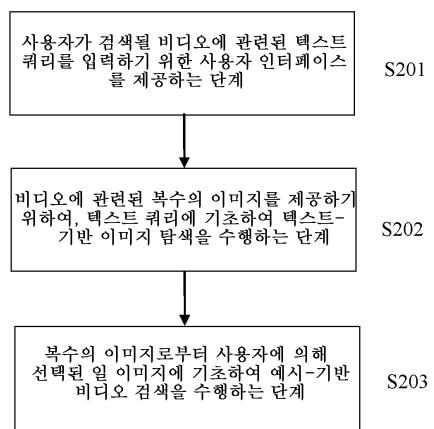
[0045] 본 발명이 하드웨어, 소프트웨어, 펌웨어, 특수 목적 프로세서들, 또는 그들의 조합의 다양한 형태들로 구현될 수 있음이 이해될 것이다.

도면

도면1



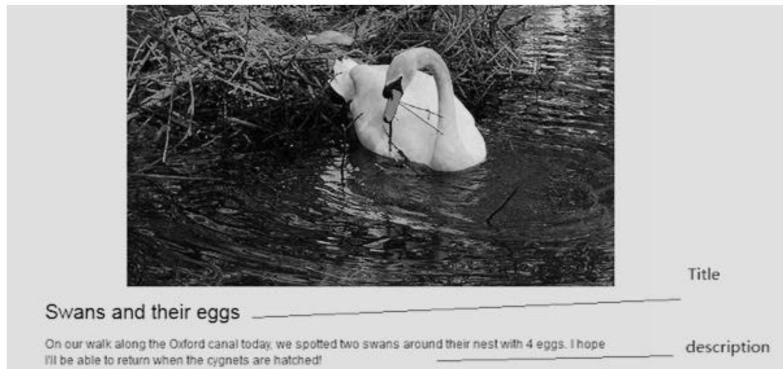
도면2



도면3

탐색하기 위한 텍스트 쿼리를 입력하세요

도면4



도면5

