



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0105542
(43) 공개일자 2013년09월25일

- (51) 국제특허분류(Int. Cl.)
G06T 7/00 (2006.01) G06F 17/30 (2006.01)
- (21) 출원번호 10-2013-0028134
- (22) 출원일자 2013년03월15일
심사청구일자 없음
- (30) 우선권주장
12305317.5 2012년03월16일
유럽특허청(EPO)(EP)
12305534.5 2012년05월14일
유럽특허청(EPO)(EP)
- (71) 출원인
툼슨 라이센싱
프랑스 92130 이씨레물리노 잔 다르크 뒤편 1-5
- (72) 발명자
빈터, 마르코
독일 30625 하노버 카를 비에체르트 알레 74 도이
체 톰슨 오에이치지 리서치 앤드 이노베이션
푸츠케-로밍, 볼프람
독일 30625 하노버 카를 비에체르트 알레 74 도이
체 톰슨 오에이치지 리서치 앤드 이노베이션
잭할스키, 요에른
독일 30625 하노버 카를 비에체르트 알레 74 도이
체 톰슨 오에이치지 리서치 앤드 이노베이션
- (74) 대리인
백만기, 양영준, 전경석

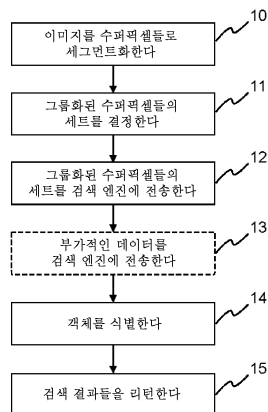
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 이미지들 또는 이미지 시퀀스들에서의 객체 식별

(57) 요약

이미지 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 해결책이 설명된다. 세그멘터(22)가 제1 이미지를 수퍼픽셀들로 분리한다(10). 분석기(23)에 의해 또는 사용자 인터페이스(24)를 통한 사용자 입력에 의해 이들 수퍼픽셀들로부터 그룹화된 수퍼픽셀들의 세트가 결정된다(11). 그룹화된 수퍼픽셀들의 세트는 검색 엔진(30)에 전송되고(12), 검색 엔진(30)은 상기 그룹화된 수퍼픽셀들의 세트에 대해 상기 검색 엔진(30)에 의해 수행된(14) 검색의 결과들을 리턴한다(15).

대표도 - 도9



특허청구의 범위

청구항 1

이미지 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 방법으로서,

- 제1 이미지를 수퍼픽셀들(superpixels)로 세그먼트화하는(segmenting) 단계(10);
- 그룹화된 수퍼픽셀들의 세트를 결정하는 단계(11);
- 상기 그룹화된 수퍼픽셀들의 세트를 검색 엔진(30)에 전송하는 단계(12); 및
- 상기 그룹화된 수퍼픽셀들의 세트에 대해 상기 검색 엔진(30)에 의해 수행된(14) 검색의 결과들을 수신하는 단계(15)

를 포함하는 객체 식별 방법.

청구항 2

제1항에 있어서, 상기 그룹화된 수퍼픽셀들의 세트는 사용자 입력에 의해 또는 상기 수퍼픽셀들의 분석에 의해 결정되는 객체 식별 방법.

청구항 3

제2항에 있어서, 상기 수퍼픽셀들의 분석은 상기 제1 이미지의 수퍼픽셀들과 제2 이미지의 대응하는 수퍼픽셀들 사이의 상대 시프트(relative shift)를 분석하는 것을 포함하는 객체 식별 방법.

청구항 4

제3항에 있어서, 상기 제1 이미지 및 상기 제2 이미지는 이미지들의 시퀀스의 시간적으로 인접한 이미지들 또는 입체 또는 멀티-뷰(multi-view) 이미지들의 세트의 공간적으로 인접한 이미지들인 객체 식별 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 이미지에 관한 부가적인 데이터는 상기 검색 엔진(30)에 전송되는(13) 객체 식별 방법.

청구항 6

이미지 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 장치(20)로서,

- 제1 이미지를 수퍼픽셀들로 세그먼트화(10)하기 위한 세그멘터(segmenter)(22);
- 그룹화된 수퍼픽셀들의 세트를 결정(11)하기 위한 분석기(23) 또는 사용자 인터페이스(24); 및
- 상기 그룹화된 수퍼픽셀들의 세트를 검색 엔진(30)에 전송(12)하고, 상기 그룹화된 수퍼픽셀들의 세트에 대해 상기 검색 엔진(30)에 의해 수행(14)된 검색의 결과들을 수신(15)하기 위한 인터페이스(25)

를 포함하는 객체 식별 장치.

명세서

기술분야

[0001] 본 발명은 이미지에서 또는 이미지들의 시퀀스(a sequence of images)에서 객체를 식별하기 위한 방법 및 장치에 관한 것이다. 더욱 구체적으로, 본 발명은, 수퍼픽셀들(superpixels)을 이용하는, 이미지에서 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 방법 및 장치에 관한 것이다.

배경기술

[0002] 2D 또는 3D 비디오 시퀀스들에서 또는 단일 이미지들에서 객체들을 식별하기 위한 다수의 접근법들이 존재한다.

예를 들어, 얼굴 인식은 콤팩트 카메라들에 대해 최첨단이고, 즉, 그것은 이미 저성능 CPU들에 대한 실시간 애플리케이션으로서 구현되어 있다. 그러나, 모든 이들 알고리즘들은 보통 특정 애플리케이션들에 대해 전용이다. 더욱 일반적인 접근법들은 통상적으로 매우 정교한 알고리즘들 및 증가된 프로세싱 능력을 요구한다.

발명의 내용

해결하려는 과제

[0003] 따라서, 본 발명의 목적은 이미지에서 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 단순하고 요구가 적은 해결책을 제안하는 것이다.

과제의 해결 수단

- [0004] 본 발명에 따르면, 이 목적은 이미지 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 방법에 의해 실현되며, 상기 방법은,
- [0005] - 제1 이미지를 슈퍼픽셀들로 세그먼트화하는(segmenting) 단계;
- [0006] - 그룹화된 슈퍼픽셀들의 세트를 결정하는 단계;
- [0007] - 상기 그룹화된 슈퍼픽셀들의 세트를 검색 엔진에 전송하는 단계; 및
- [0008] - 상기 그룹화된 슈퍼픽셀들의 세트에 대해 상기 검색 엔진에 의해 수행된 검색의 결과들을 수신하는 단계
- [0009] 를 포함한다.
- [0010] 유사하게, 이미지 또는 이미지들의 시퀀스에서 객체를 식별하기 위한 장치는,
- [0011] - 제1 이미지를 슈퍼픽셀들로 세그먼트화하기 위한 세그멘터(segmenter);
- [0012] - 그룹화된 슈퍼픽셀들의 세트를 결정하기 위한 분석기 또는 사용자 인터페이스; 및
- [0013] - 상기 그룹화된 슈퍼픽셀들의 세트를 검색 엔진에 전송하고, 상기 그룹화된 슈퍼픽셀들의 세트에 대해 상기 검색 엔진에 의해 수행된 검색의 결과들을 수신하기 위한 인터페이스
- [0014] 를 포함한다.
- [0015] 본 발명에 따른 해결책은 2D 또는 3D 또는 멀티-뷰(multi-view) 이미지 시퀀스에서 또는 단일 이미지에서 객체를 식별하기 위하여 2개의 상이한 접근법을 결합한다. 먼저, 하나 이상의 이미지들은 슈퍼픽셀들로 세그먼트화된다. 그 다음에, 사용자는 더 조사되어야 하는 슈퍼픽셀들을 표시(mark)하거나, 객체에 속하는 그룹화된 슈퍼픽셀들의 세트가 자동으로 검출된다. 자동 검출은 예를 들어, 이미지들의 시퀀스의 경우에 슈퍼픽셀들의 시간적 분석에, 입체 또는 멀티-뷰 이미지들의 세트의 2개 이상의 이미지들의 공간적 분석에, 또는 다른 이미지 분석 절차들에 기초한다. 임의의 경우, 일단 식별된 슈퍼픽셀들의 세트가 알려지면, 이 슈퍼픽셀들의 세트의 특성들의 리스트가 확립된다. 이들 특성들은 이들 특성들에 대한 가장 적절한 객체를 결정하는 특수 슈퍼픽셀 객체 데이터베이스에 대해 질의로서 전송된다. 그 다음에, 식별된 타입의 객체는 사용자에게 보내진다. 본 발명은 대부분의 객체들이 특징적인 슈퍼픽셀들의 세트를 갖는다는 발견을 활용한다. 따라서, 슈퍼픽셀들에 기초하여 이미지 또는 이미지 시퀀스에서 객체를 식별하는 것이 가능하다. 데이터베이스들은 계속해서 점점 더 커지고 모두를 아우르기 때문에, 검색 엔진 능력의 성장은 셋톱박스들, 스마트폰들, 태블릿들 유사 장치들을 위한 편리한 객체 검색 툴을 가능하게 한다.
- [0016] 유리하게는, 부가적인 정보, 예를 들어, 식별될 객체를 포함하는 이미지 또는 이미지들의 시퀀스에 관한 메타데이터가 슈퍼픽셀 객체 데이터베이스에 전송된다. 이러한 메타데이터의 예는 영화 제목, 영화의 배우들의 리스트 등이다. 이들 부가적인 메타데이터는 특정 정도로 올바르게 분류를 배제할 것이기 때문에 분류를 안정화하는 데 도움을 준다. 또한, 검색 결과들을 향상시키기 위하여 관측된 객체의 시간적 이동이 분석되어 데이터베이스에 전송될 수 있다.
- [0017] 더 나은 이해를 위해, 본 발명은 이제 도면들을 참조하여 다음의 설명에서 더욱 상세히 설명될 것이다. 본 발명은 이 예시적인 실시예로 한정되지 않고, 특정 특징들이 또한 첨부된 청구항들에 정의된 바와 같은 본 발명의 범위에서 벗어나지 않고 편의상 결합 및/또는 수정될 수 있다는 것을 이해한다.

도면의 간단한 설명

- [0018] 도 1은 원래 이미지를 도시한다.
- 도 2는 도 1의 이미지의 인간 표시 세그먼트화(human-marked segmentation)를 도시한다.
- 도 3은 도 1의 이미지로부터 얻어지는 슈퍼픽셀들을 도시한다.
- 도 4는 도 3의 슈퍼픽셀들을 이용하여 도 2의 인간 표시 세그먼트화의 재구성을 도시한다.
- 도 5는 사용자에게 의해 표시된 다수의 슈퍼픽셀들을 갖는 얼룩말의 이미지를 도시한다.
- 도 6은 사용자에게 의해 표시된 슈퍼픽셀들을 갖는 도 5의 확대된 일부를 도시한다.
- 도 7은 슈퍼픽셀들로 세그먼트화된 물고기의 이미지를 도시한다.
- 도 8은 슈퍼픽셀들로 세그먼트화된 빌딩의 이미지를 도시한다.
- 도 9는 객체 식별을 위한 본 발명에 따른 방법을 개략적으로 예시한다.
- 도 10은 객체 식별을 위한 본 발명에 따른 장치를 개략적으로 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0019] 슈퍼픽셀들, 더 정확히 말하면, 슈퍼픽셀 맵들(superpixel maps)의 발생과, 객체 인식을 위한 그것들의 이용은 현재 이미지 프로세싱을 위한 고려사항에 있다. 예를 들어, 기사 X. Ren et al.: "Learning a Classification Model for Segmentation", 9th IEEE International Conference on Computer Vision (ICCV 2003), pp. 10-17은 이미지들의 복잡성을 줄이기 위해서 이미지의 슈퍼픽셀들로의 과세그먼트화(oversegmentation)를 설명한다.
- [0020] X. Ren et al.에 의한 위의 기사로부터 취해진 도 1 내지 4는 슈퍼픽셀들로 세그먼트화된 이미지의 예를 도시하였다. 도 1은 원래 이미지를 도시하고, 도 2는 이 이미지의 인간 표시 세그먼트화를 도시한다. 도 3은 정규화 절단(Normalized Cuts) 알고리즘을 적용함으로써 얻어지는 200개의 슈퍼픽셀들로의 이미지의 세그먼트화를 도시한다. 도 4는 도 3의 슈퍼픽셀들로부터 도 2의 인간 세그먼트화의 재구성이다. 이 목적을 위해, 각각의 슈퍼픽셀들은 최대 중첩 영역(maximum overlapping area)을 갖는 도 2의 세그먼트에 할당되고, 슈퍼픽셀 경계들이 추출된다.
- [0021] 도 5는 사용자에게 의해 표시된 다수의 슈퍼픽셀들을 갖는 얼룩말의 이미지를 도시한다. 표시된 슈퍼픽셀들이 더욱 분명하게 보이는 이 이미지의 확대된 일부가 도 6에 도시된다.
- [0022] 슈퍼픽셀들로 세그먼트화되는 이미지들의 추가 예들이 도 7 및 8에 예시되고, 이것은 물고기의 이미지 및 빌딩의 이미지를 각각 도시한다.
- [0023] 본 발명에 따른 방법은 도 9에 개략적으로 예시된다. 대용량 데이터베이스들이 오늘날 보편적이다. 예를 들어, 인터넷 검색 엔진들은 전세계 서버 팜들에서의 막대한 양의 매우 강력한 데이터베이스들에 기초한다. 이 용가능한 검색 기술들은 셋톱박스들, 스마트폰들, 태블릿 컴퓨터들, 등을 위해서도 강력한 객체 인식을 구현할 수 있게 한다. 이 목적을 위해, 제1 단계에서, 이미지가 슈퍼픽셀들로 분리 또는 세그먼트화된다(10). 그 다음에, 사용자는 도 5 및 6에서 보이는 바와 같이, 객체에 속하는 슈퍼픽셀들의 그룹을 선택하거나(11) 또는 장치 자체가 자동으로 그룹화된 슈퍼픽셀들의 세트를 결정한다. 그 다음에, 장치는 그룹화된 슈퍼픽셀들을 검색 엔진에 전송한다(12). 옵션으로, 장치는 이미지에 관한 부가적인 메타데이터 또는 다른 데이터, 예를 들어, 이미지가 특정 영화의 부분인지 등등을 전송한다(13). 그 다음에, 검색 엔진은 데이터베이스 검색을 이용하여 그룹화된 슈퍼픽셀들에 의해 설명되는 객체를 식별한다(14). 결과가 그룹화되어 장치에 보내진다(15).
- [0024] 본 발명에 따른 장치(20)가 도 10에 도시된다. 장치(20)는 이미지 또는 이미지 시퀀스를 수신하기 위한 입력(21)을 포함한다. 세그멘터(22)가 이미지를 슈퍼픽셀들로 분리한다. 분석기(23)가 그룹화된 슈퍼픽셀들의 세트를 결정한다. 대안적으로, 사용자는 사용자 인터페이스(24)를 통해 객체에 속하는 슈퍼픽셀들의 그룹을 선택할 수 있다. 검색 엔진(30)에 그룹화된 슈퍼픽셀들을 전송하고 검색 엔진(30)에 의해 얻어진 검색의 결과들을 수신하기 위한 인터페이스(25)가 제공된다.
- [0025] 검색 결과는, 객체(예를 들어, 동물)의 대략적인(coarse) 분류; 더욱 특정한 분류(예를 들어, 얼룩말); 대안적

인 분류(예를 들어, 콰가(quagga)); 현재 보는 영화의 다른 곳에서 이 객체가 보이는 곳(즉, 타임 스탬프들); 또는 그러한 객체를 갖는 다른 영화들(예를 들어, 험프리 보가트(Humphrey Bogart)를 갖는 다른 영화들)과 같은, 상이한 타입들의 정보를 커버할 수 있다. 분류와 함께, 분류의 확률치가 또한 제공될 수 있다.

[0026] 배우들의 얼굴들; 동물들의 타입들; 유명한 성들 또는 빌딩들의 이름들; 집의 정면도를 표시함으로써 방송 필름(film) 또는 다큐멘테이션(documentation) 또는 뉴스 등에 나오는 집의 주소; 자동차 타입들; 예를 들어, 영화의 파이널 크레딧들(final credits)의 부분들을 표시함으로써 영화 제목들; 위치를 식별하기 위한 도시들의 특수 표지판, 차량 번호판, 빌딩 간판들 등; 그림 및 다른 미술 객체들, 예를 들어, 조각상; 예를 들어, 제품들에 관한 부가적인 정보를 획득하기 위한, 제품들의 상표들; 나무 타입, 나뭇잎 타입, 열매 타입 등; 예를 들어, 와인의 타입을 식별하기 위한 병 타입과 같은, 매우 다양한 상이한 객체들이 검출될 수 있다.

[0027] 검색을 지원하기 위해 상이한 접근법들이 이용된다. 장치는 바람직하게는 객체들의 시간적 이동들을 주의함으로써 슈퍼픽셀 발생을 최적화한다. 이 목적을 위해, 슈퍼픽셀들의 경계들은 바람직하게는 객체 경계들과 일치한다. 이것은 객체 선택을 단순화한다. 그러나, 객체들의 완벽한 "컷(cut)"을 갖는 것은 필요하지 않고, 즉, 표시된 얼굴이 머리카락으로 둘러싸여 있는지 여부는 관련이 적다. 그룹화된 슈퍼픽셀들의 이동은 객체의 타입을 식별하는 데 도움이 되는 특징적인 거동(characteristic behavior)을 가질 수 있다. 예를 들어, 자동차는 사람과 상이하게 이동하고, 사람은 영양(antelope)과 상이하게 이동하고, 영양은 독수리와 상이하게 이동하는 등등이다. 또한, 이동의 타입은 객체의 상태, 예를 들어, 영양이 달리거나, 서 있거나, 먹고 있거나, 또는 누워 있는지에 대한 힌트를 줄 수 있다. 이 분석은 유리하게는 장치에 의해 수행되어 메타데이터로서 검색 엔진에 전송된다.

부호의 설명

- [0028] 21: 입력
- 22: 세그멘터
- 23: 분석기
- 24: 사용자 인터페이스
- 25: 인터페이스
- 30: 검색 엔진

도면

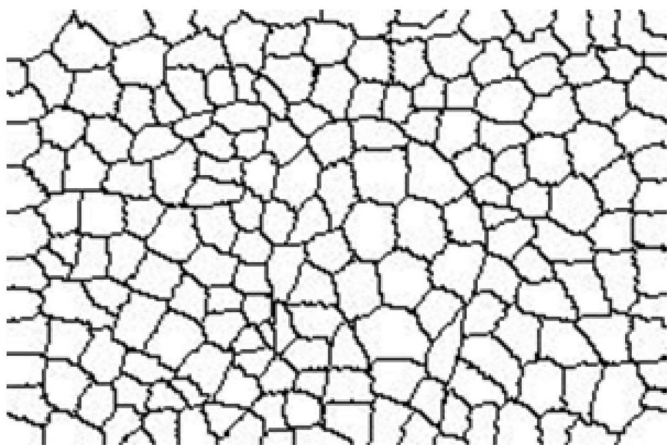
도면1



도면2



도면3



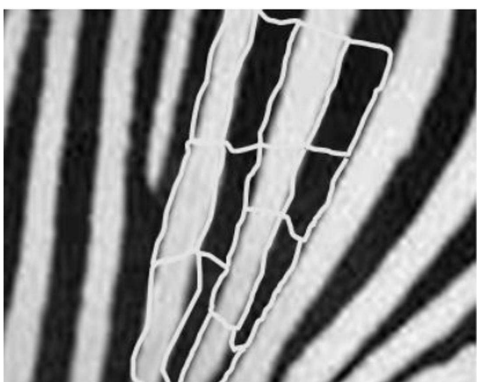
도면4



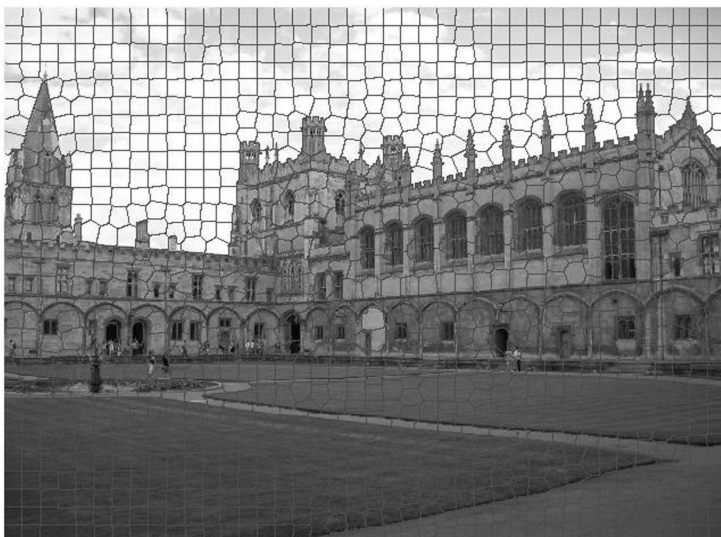
도면5



도면6



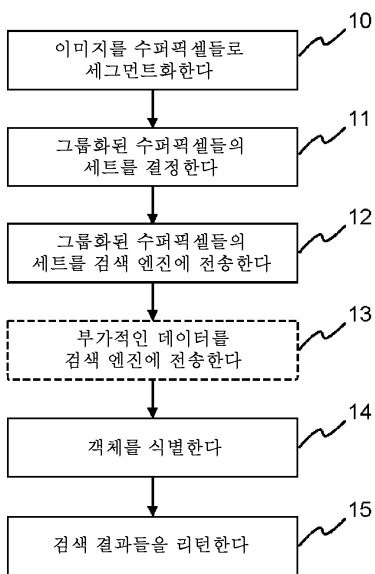
도면7



도면8



도면9



도면10

