



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2014년04월24일

(11) 등록번호 10-1388638

(24) 등록일자 2014년04월17일

(51) 국제특허분류(Int. Cl.)

G06F 17/30 (2006.01) G06F 17/40 (2006.01)

(21) 출원번호 10-2011-7000140

(22) 출원일자(국제) 2009년04월17일

심사청구일자 2013년05월13일

(85) 번역문제출일자 2011년01월04일

(65) 공개번호 10-2011-0033179

(43) 공개일자 2011년03월30일

(86) 국제출원번호 PCT/US2009/040975

(87) 국제공개번호 WO 2009/154861

국제공개일자 2009년12월23일

(30) 우선권주장

61/059,702 2008년06월06일 미국(US)

(56) 선행기술조사문헌

WO2007-136861 A2

KR1020070115194 A

US2007/0271226 A1

US20070296993 A1

전체 청구항 수 : 총 34 항

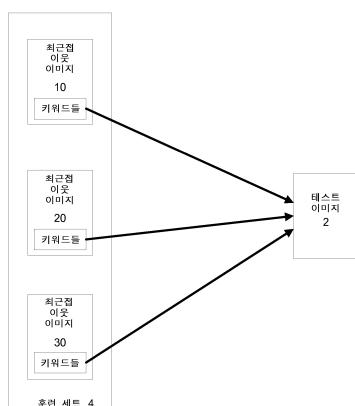
심사관 : 윤혜숙

(54) 발명의 명칭 이미지에 주석 달기

(57) 요약

자동적으로 이미지에 주석을 달기 위한 데이터를 생성하는 컴퓨터 프로그램 제품을 포함하는 방법, 시스템, 및 장치들에 관한 것이다. 일 양태에서, 방법은 상기 데이터 프로세싱 장치에서 입력 이미지를 수신하는 단계; 상기 입력 이미지에 대한 하나 이상의 최근접 이웃 이미지-여기서, 하나 이상의 최근접 이미지 각각은 하나 이상의 이미지 라벨에 각각 연관됨-를, 컴퓨터 판독가능 매체 상에 저장된 디지털 이미지들의 집합 중에서 상기 데이터 프로세싱 장치의 동작에 의해 식별하는 단계; 복수의 이미지 라벨-여기서, 복수의 이미지 라벨은 상기 데이터 프로세싱 장치에 의해 상기 하나 이상의 최근접 이웃 이미지와 연관된 상기 이미지 라벨들로부터 선택됨-을 상기 입력 이미지에 할당하는 단계; 및 상기 할당된 복수의 이미지 라벨을 갖는 상기 입력 이미지를 디지털 데이터 저장소에 저장하는 단계를 포함한다. 다른 양태에서, 방법은 단일 이미지 라벨을 상기 입력 이미지에 할당하는 단계를 포함하고, 상기 단일 이미지 라벨은 복수의 순위가 부여된 최근접 이웃 이미지와 연관된 라벨들로부터 선택된다.

대표 도 - 도1



특허청구의 범위

청구항 1

데이터 처리 장치에 의해 수행되는 이미지에 주석을 달는 방법으로서,

상기 데이터 처리 장치에서 입력 이미지를 수신하는 단계;

상기 입력 이미지에 대한 복수의 최근접 이웃 이미지들(nearest neighbor image) — 상기 복수의 최근접 이웃 이미지들 각각은 하나 이상의 이미지 라벨에 각각 연관됨 — 을, 컴퓨터 판독가능 매체 상에 저장된 디지털 이미지들의 집합 중에서 상기 데이터 처리 장치의 동작에 의해 식별하는 단계;

복수의 이미지 라벨을 상기 입력 이미지에 할당하는 단계 — 상기 복수의 이미지 라벨은 상기 데이터 처리 장치에 의해 상기 복수의 최근접 이웃 이미지들과 연관된 상기 이미지 라벨들로부터 선택됨 —;

상기 디지털 이미지들의 집합에서의 각각의 발생 빈도(frequency of occurrence)에 따라 하나 이상의 제1 이미지 라벨에 순위를 부여(ranking)하는 단계 — 상기 하나 이상의 제1 이미지 라벨 각각은, 상기 복수의 최근접 이웃 이미지를 중 제1 최근접 이웃 이미지와 연관됨 —;

상기 복수의 최근접 이웃 이미지를 중 하나 이상의 잔여 최근접 이웃 이미지와 각각 연관되는, 하나 이상의 제2 이미지 라벨에 순위를 부여하는 단계; 및

상기 할당된 복수의 이미지 라벨을 갖는 상기 입력 이미지를 디지털 데이터 저장소에 저장하는 단계를 포함하고,

상기 하나 이상의 제2 이미지 라벨에 순위를 부여하는 단계는, 상기 디지털 이미지들의 집합에서 상기 제2 이미지 라벨 각각과 상기 제1 이미지 라벨 각각이 중복적으로 발생(co-occurrence)하는 정도에 따라 상기 하나 이상의 제2 이미지 라벨을 분류하는 단계를 포함하고,

상기 복수의 이미지 라벨을 할당하는 단계는

(1) 상기 하나 이상의 제1 이미지 라벨의 순위에 기초하여, 적어도 하나의 상기 제1 이미지 라벨을 상기 입력 이미지에 할당하는 단계; 및

(2) 상기 하나 이상의 제2 이미지 라벨의 순위에 기초하여, 적어도 하나의 상기 제2 이미지 라벨을 상기 입력 이미지에 할당하는 단계를 포함하는 방법.

청구항 2

청구항 1에 있어서,

상기 입력 이미지는 복수의 이미지 라벨을 메타데이터로서 포함하는 이미지 파일로 상기 디지털 데이터 저장소에 저장되는 방법.

청구항 3

청구항 1에 있어서,

상기 디지털 이미지들의 집합은 복수의 참조 이미지(reference image)를 포함하고,

상기 복수의 최근접 이웃 이미지를 식별하는 단계는

각 참조 이미지에 대해, 상기 입력 이미지와 상기 참조 이미지 간의 차이 정도(degree of difference)를 나타내는 상응하는 전체 이미지 거리(whole-image distance)를 결정하는 단계와, 상기 전체 이미지 거리에 의해 측정된 것에 따라, 상기 입력 이미지에 가장 가까운(closest) 참조 이미지들을 상기 복수의 최근접 이웃으로 식별하는 단계를 포함하는 방법.

청구항 4

청구항 1에 있어서,

상기 하나 이상의 제2 이미지 라벨에 순위를 부여하는 단계는, 상기 하나 이상의 잔여 최근접 이웃 이미지에서의 상기 제2 이미지 라벨 각각의 국소적 빈도(local frequency)에 따라 상기 하나 이상의 제2 이미지 라벨을 분류하는 단계-상기 제2 이미지 라벨 각각의 국소적 빈도는, 상기 제2 이미지 라벨 각각이 상기 하나 이상의 잔여 최근접 이웃 이미지들과 연관된 이미지 라벨로서 나타나는 정도임-를 포함하는 방법.

청구항 5

청구항 3에 있어서,

상기 전체 이미지 거리는 특징 거리들(feature distances)의 조합을 포함하고, 각 특징 거리는 상기 입력 이미지와 연관된 이미지 특징과, 상응하는 참조 이미지와 연관된 상응하는 이미지 특징 각각 간에 차이 정도를 나타내는 방법.

청구항 6

청구항 5에 있어서,

적어도 하나의 이미지 특징은 상기 입력 이미지 및 상기 참조 이미지 모두에서 추출된 전체적 이미지 특징(global image feature)인 방법.

청구항 7

청구항 5에 있어서,

적어도 하나의 이미지 특징은 상기 입력 이미지 및 상기 참조 이미지 모두에서 추출된 국부적(局部的) 이미지 특징(local image feature)인 방법.

청구항 8

청구항 5에 있어서,

각 특징 거리는 상기 전체 이미지 거리에서 동등하게 가중(weight)되는 방법.

청구항 9

청구항 5에 있어서,

둘 이상의 특징 거리는 상기 전체 이미지 거리에서 서로 다르게 가중되는 방법.

청구항 10

청구항 5에 있어서,

상기 디지털 이미지들의 집합에 기초하여, 상기 특징 거리 각각에 대한 가중치를 산출하는 단계를 더 포함하고, 상기 디지털 이미지들의 집합은 유사한 이미지와 유사하지 않은 이미지의 쌍들을 포함하는 훈련 이미지(training image)들의 그룹인 방법.

청구항 11

청구항 5에 있어서,

상기 입력 이미지의 텍스쳐 특징(texture feature)과 상응하는 참고 이미지의 상응하는 텍스쳐 특징 간의 차이로서, 상기 특징 거리를 중 적어도 하나를 산출하는 단계를 더 포함하는 방법.

청구항 12

청구항 5에 있어서,

상기 입력 이미지의 색상 특징(color feature)과 상응하는 참고 이미지의 상응하는 색상 특징 간의 차이로서, 상기 특징 거리를 중 적어도 하나를 산출하는 단계를 더 포함하는 것을 특징으로 하는 방법.

청구항 13

하나 이상의 컴퓨터에 구현된 서버를 포함하고, 상기 서버는

상기 서버에서 입력 이미지를 수신하는 동작;

상기 입력 이미지에 대한 복수의 최근접 이웃 이미지들—상기 복수의 최근접 이웃 이미지들 각각은 하나 이상의 이미지 라벨에 각각 연관됨—을, 컴퓨터 판독가능 매체 상에 저장된 디지털 이미지들의 집합 중에서 상기 서버의 동작에 의해 식별하는 동작;

복수의 이미지 라벨을 상기 입력 이미지에 할당하는 동작—상기 복수의 이미지 라벨은 상기 서버에 의해 상기 복수의 최근접 이웃 이미지들과 연관된 상기 이미지 라벨들로부터 선택됨—;

상기 디지털 이미지들의 집합에서의 각각의 발생 빈도에 따라 하나 이상의 제1 이미지 라벨에 순위를 부여 (ranking)하는 동작—상기 하나 이상의 제1 이미지 라벨 각각은 상기 복수의 최근접 이웃 이미지들 중 제1 최근접 이웃 이미지와 연관됨—;

상기 복수의 최근접 이웃 이미지들 중 하나 이상의 잔여 최근접 이웃 이미지와 각각 연관되는, 하나 이상의 제2 이미지 라벨에 순위를 부여하는 동작; 및

상기 할당된 복수의 이미지 라벨을 갖는 상기 입력 이미지를 디지털 데이터 저장소에 저장하는 동작을 포함하고,

상기 하나 이상의 제2 이미지 라벨에 순위를 부여하는 동작은, 상기 디지털 이미지들의 집합에서 상기 제2 이미지 라벨 각각과 상기 제1 이미지 라벨 각각이 중복적으로 발생하는 정도에 따라 상기 하나 이상의 제2 이미지 라벨을 분류하는 동작을 포함하고,

상기 복수의 이미지 라벨을 할당하는 동작은

(1) 상기 하나 이상의 제1 이미지 라벨의 순위에 기초하여, 적어도 하나의 상기 제1 이미지 라벨을 상기 입력 이미지에 할당하는 동작; 및

(2) 상기 하나 이상의 제2 이미지 라벨의 순위에 기초하여, 적어도 하나의 상기 제2 이미지 라벨을 상기 입력 이미지에 할당하는 동작을 수행할 수 있는 시스템.

청구항 14

청구항 13에 있어서,

상기 디지털 이미지들의 집합은 복수의 참조 이미지를 포함하고,

상기 복수의 최근접 이웃 이미지를 식별하는 동작은

각 참조 이미지에 대해, 상기 입력 이미지와 상기 참조 이미지 간의 차이 정도를 나타내는 상응하는 전체 이미지 거리를 결정하는 동작과, 상기 전체 이미지 거리에 의해 측정된 것에 따라, 상기 입력 이미지에 가장 가까운 참조 이미지들을 상기 복수의 최근접 이웃으로 식별하는 동작을 포함하는 시스템.

청구항 15

청구항 13에 있어서,

상기 하나 이상의 제2 이미지 라벨에 순위를 부여하는 동작은, 상기 하나 이상의 잔여 최근접 이웃 이미지에서의 상기 제2 이미지 라벨 각각의 국소적 빈도에 따라 상기 하나 이상의 제2 이미지 라벨을 분류하는 동작—상기 제2 이미지 라벨 각각의 국소적 빈도는, 상기 제2 이미지 라벨 각각이 상기 하나 이상의 잔여 최근접 이웃 이미지와 연관된 이미지 라벨로서 나타나는 정도임—을 포함하는 시스템.

청구항 16

청구항 14에 있어서,

상기 전체 이미지 거리는 특정 거리들의 조합을 포함하고, 각 특정 거리는 상기 입력 이미지와 연관된 이미지 특징과, 상응하는 참조 이미지와 연관된 상응하는 이미지 특징 각각 간에 차이 정도를 나타내는 시스템.

청구항 17

청구항 16에 있어서,

적어도 하나의 이미지 특징은 상기 입력 이미지 및 상기 참조 이미지 모두에서 추출된 전체적 이미지 특징인 시스템.

청구항 18

청구항 16에 있어서,

적어도 하나의 이미지 특징은 상기 입력 이미지 및 상기 참조 이미지 모두에서 추출된 국부적 이미지 특징인 시스템.

청구항 19

청구항 16에 있어서,

각 특징 거리는 상기 전체 이미지 거리에서 동등하게 가중되는 시스템.

청구항 20

청구항 16에 있어서,

둘 이상의 특징 거리는 상기 전체 이미지 거리에서 서로 다르게 가중되는 시스템.

청구항 21

청구항 16에 있어서, 상기 서버는

상기 디지털 이미지들의 집합에 기초하여, 상기 특징 거리 각각에 대한 가중치를 산출하는 동작을 더 수행할 수 있고,

상기 디지털 이미지들의 집합은 유사한 이미지와 유사하지 않은 이미지의 쌍들을 포함하는 훈련 이미지들의 그룹인 시스템.

청구항 22

청구항 16에 있어서, 상기 서버는

상기 입력 이미지의 텍스처 특징과 상응하는 참고 이미지의 상응하는 텍스처 특징 간의 차이로서, 상기 특징 거리들 중 적어도 하나를 산출하는 동작을 더 수행할 수 있는 시스템.

청구항 23

청구항 16에 있어서, 상기 서버는

상기 입력 이미지의 색상 특징과 상응하는 참고 이미지의 상응하는 색상 특징 간의 차이로서, 상기 특징 거리들 중 적어도 하나를 산출하는 동작을 더 수행할 수 있는 시스템.

청구항 24

명령어를 포함하는 컴퓨터 프로그램이 인코드된 컴퓨터 저장 매체로서, 데이터 처리 장치에 의해 실행되었을 때, 상기 데이터 처리 장치로 하여금

상기 데이터 처리 장치에서 입력 이미지를 수신하는 동작;

상기 입력 이미지에 대한 복수의 최근접 이웃 이미지들—상기 복수의 최근접 이웃 이미지들 각각은 하나 이상의 이미지 라벨에 각각 연관됨—을, 컴퓨터 판독가능 매체 상에 저장된 디지털 이미지들의 집합 중에서 상기 데이터 처리 장치의 동작에 의해 식별하는 동작;

복수의 이미지 라벨을 상기 입력 이미지에 할당하는 동작—상기 복수의 이미지 라벨은 상기 데이터 처리 장치에 의해 상기 복수의 최근접 이웃 이미지들과 연관된 상기 이미지 라벨들로부터 선택됨—;

상기 디지털 이미지들의 집합에서의 각각의 발생 빈도에 따라 하나 이상의 제1 이미지 라벨에 순위를 부여 (ranking)하는 동작—상기 하나 이상의 제1 이미지 라벨 각각은 상기 복수의 최근접 이웃 이미지들 중 제1 최근

접 이웃 이미지와 연관됨—;

상기 복수의 최근접 이웃 이미지를 중 하나 이상의 잔여 최근접 이웃 이미지와 각각 연관되는, 하나 이상의 제2 이미지 라벨에 순위를 부여하는 동작; 및

상기 할당된 복수의 이미지 라벨을 갖는 상기 입력 이미지를 디지털 데이터 저장소에 저장하는 동작을 포함하고,

상기 하나 이상의 제2 이미지 라벨에 순위를 부여하는 동작은, 상기 디지털 이미지들의 집합에서 상기 제2 이미지 라벨 각각과 상기 제1 이미지 라벨 각각이 중복적으로 발생하는 정도에 따라 상기 하나 이상의 제2 이미지 라벨을 분류하는 동작을 포함하고,

상기 복수의 이미지 라벨을 할당하는 동작은

(1) 상기 하나 이상의 제1 이미지 라벨의 순위에 기초하여, 적어도 하나의 상기 제1 이미지 라벨을 상기 입력 이미지에 할당하는 동작; 및

(2) 상기 하나 이상의 제2 이미지 라벨의 순위에 기초하여, 적어도 하나의 상기 제2 이미지 라벨을 상기 입력 이미지에 할당하는 동작을 수행하도록 하는 컴퓨터 기록 매체.

청구항 25

청구항 24에 있어서,

상기 디지털 이미지들의 집합은 복수의 참조 이미지를 포함하고,

상기 복수의 최근접 이웃 이미지를 식별하는 동작은

각 참조 이미지에 대해, 상기 입력 이미지와 상기 참조 이미지 간의 차이 정도를 나타내는 상응하는 전체 이미지 거리를 결정하는 동작과, 상기 전체 이미지 거리에 의해 측정된 것에 따라, 상기 입력 이미지에 가장 가까운 하나 이상의 참조 이미지들을 상기 복수의 최근접 이웃으로 식별하는 동작을 포함하는 컴퓨터 기록 매체.

청구항 26

청구항 24에 있어서,

상기 하나 이상의 제2 이미지 라벨에 순위를 부여하는 동작은 상기 하나 이상의 잔여 최근접 이웃 이미지에서의 상기 제2 이미지 라벨 각각의 국소적 빈도에 따라 상기 하나 이상의 제2 이미지 라벨을 분류하는 동작-상기 제2 이미지 라벨 각각의 국소적 빈도는, 상기 제2 이미지 라벨 각각이 상기 하나 이상의 잔여 최근접 이웃 이미지와 연관된 이미지 라벨로서 나타나는 정도임-을 포함하는 컴퓨터 기록 매체.

청구항 27

청구항 25에 있어서,

상기 전체 이미지 거리는 특정 거리들의 조합을 포함하고, 각 특정 거리는 상기 입력 이미지와 연관된 이미지 특징과, 상응하는 참조 이미지와 연관된 상응하는 이미지 특징 각각 간에 차이 정도를 나타내는 컴퓨터 기록 매체.

청구항 28

청구항 27에 있어서,

적어도 하나의 이미지 특징은 상기 입력 이미지 및 상기 참조 이미지 모두에서 추출된 전체적 이미지 특징인 컴퓨터 기록 매체.

청구항 29

청구항 27에 있어서,

적어도 하나의 이미지 특징은 상기 입력 이미지 및 상기 참조 이미지 모두에서 추출된 국부적 이미지 특징인 컴퓨터 기록 매체.

청구항 30

청구항 27에 있어서,

각 특징 거리는 상기 전체 이미지 거리에서 동등하게 가중되는 컴퓨터 기록 매체.

청구항 31

청구항 27에 있어서,

둘 이상의 특징 거리는 상기 전체 이미지 거리에서 서로 다르게 가중되는 컴퓨터 기록 매체.

청구항 32

청구항 27에 있어서, 상기 데이터 처리 장치로 하여금

상기 디지털 이미지들의 집합에 기초하여, 상기 특징 거리 각각에 대한 가중치를 산출하는 동작을 더 수행하도록 하고,

상기 디지털 이미지들의 집합은 유사한 이미지와 유사하지 않은 이미지의 쌍들을 포함하는 훈련 이미지들의 그룹인 컴퓨터 기록 매체.

청구항 33

청구항 27에 있어서, 상기 데이터 처리 장치로 하여금

상기 입력 이미지의 텍스쳐 특징과 상응하는 참고 이미지의 상응하는 텍스쳐 특징 간의 차이로서, 상기 특징 거리들 중 적어도 하나를 산출하는 동작을 더 수행하도록 하는 컴퓨터 기록 매체.

청구항 34

청구항 27에 있어서, 상기 데이터 처리 장치로 하여금

상기 입력 이미지의 색상 특징과 상응하는 참고 이미지의 상응하는 색상 특징 간의 차이로서, 상기 특징 거리들 중 적어도 하나를 산출하는 동작을 더 수행하도록 하는 컴퓨터 기록 매체.

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

명세서

기술 분야

[0001]

본 출원은 2008년 6월 6일 출원되고 발명의 명칭이 "이미지에 주석 달기"인 미국 특허출원 제61/059,702호에 대한 우선권을 미국 특허법(35USC) 119(e)항 하에 주장하며, 이 미국 출원의 모든 내용은 참조로서 본 명세서에 포함된다.

[0002]

본 명세서는 이미지 주석 달기에 관한 것이다.

배경기술

[0003]

텍스트 기반 이미지 주석 달기는 컴퓨터 시각(computer vision)과 정보 검색 커뮤니티에서 근본적이고 상당히 실제적인 과제로 지속되어 왔다. 실제적인 전망으로 볼 때, 현재 이미지 탐색 솔루션은 이미지 탐색을 위하여 이미지 컨텐츠를 효율적으로 이용하지 못한다. 이것은 때때로 적용성이 제한된 탐색 결과를 가져온다.

[0004]

입력 이미지를 고려했을 때, 자동 이미지 주석 달기의 목표는 이미지의 비주얼 컨텐츠를 반영하는 몇 개의 관련 텍스트 키워드(relevant text keyword; 또는 라벨로 일컬어짐)들을 이미지에 할당하는 것이다. 키워드들은 다양한 방법 예를 들어, 이미지를 포함하는 디지털 파일 내에 메타 데이터(meta data)로서 키워드들을 저장하거나, 키워드들을 이미지에 연결시키는 링크 또는 참조물이 저장된 데이터 베이스 내에 메타 데이터로서 키워드들을 저장하거나, 또는 키워드와 이미지를 링크하는 데이터가 포함된 XML 파일 내에 메타 데이터로서 키워드들을 저장함으로써 이미지에 할당(또는 연관) 될 수 있다.

[0005]

온 및 오프 웹 상의 이미지 데이터 집합(collection of image data)들이 급속하게 증가함에 따라, 탄탄한 이미지 탐색 및 검색이 빠르게 중요 요구 사항이 되고 있다. 현재 인터넷 이미지 탐색 엔진은 이미지 컨텐츠는 무시하면서, 관련 이미지를 검색하기 위하여 일반적으로 텍스트 기반 탐색을 이용하고 있다. 더 관련있는 키워드 세트를 더 많이 할당하기 위해 이미지 컨텐츠를 이용함으로써, 개선된 이미지 탐색을 위하여 탐색 엔진들의 고속 색인화와 검색 구조를 더욱 이용할 수 있다. 이는 방대한 실제적 흥미에 대한 관련 텍스트 키워드들로 이미지에 주석을 달아야 한다는 문제를 야기한다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 이미지에 주석을 다는 기술을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0007] 본 명세서는 이미지에 주석을 자동적으로 다는 것과 관련된 기술을 설명한다. 전체적으로, 본 명세서에서 설명된 요지에 대한 한 양태는, 데이터 프로세싱 장치에 의해 실행되는 이미지에 주석다는 방법으로 구현될 수 있고, 그 방법은 데이터 프로세싱 장치에 의해 수행되는 이미지에 주석을 다는 방법으로서, 데이터 프로세싱 장치에서 입력 이미지를 수신하는 단계; 입력 이미지에 대한 하나 이상의 최근접 이웃 이미지-여기서, 하나 이상의 최근접 이미지 각각은 하나 이상의 이미지 라벨에 각각 연관됨-를, 컴퓨터 판독가능 매체 상에 저장된 디지털 이미지들의 집합 중에서 데이터 프로세싱 장치의 동작에 의해 식별하는 단계; 복수의 이미지 라벨-여기서, 복수의 이미지 라벨은 데이터 프로세싱 장치에 의해 하나 이상의 최근접 이웃 이미지와 연관된 이미지 라벨들로부터 선택됨-을 입력 이미지에 할당하는 단계; 및 할당된 복수의 이미지 라벨을 갖는 입력 이미지의 군집(association)과 입력 이미지를 저장하는 단계를 포함한다.

[0008] 다른 양태에서, 컴퓨터 판독 가능 방법은 입력 이미지를 수신하는 단계; 이미지 집합 중에서 입력 이미지에 대한 복수의 순위가 부여된 최근접 이웃 이미지- 순위가 부여된 최근접 이웃 이미지 각각은 하나 이상의 이미지 라벨과 각각 연관됨-를 식별하는 단계; 단일 이미지 라벨- 단일 이미지 라벨은 복수의 순위가 부여된 최근접 이웃 이미지와 연관된 이미지 라벨들로부터 선택됨-을 입력 이미지에 할당하는 단계; 할당된 단일 이미지 라벨을 갖는 입력 이미지를 데이터 저장소에 저장하는 단계를 포함한다. 상술한 양태들에 대한 다른 실시예들은 상응하는 시스템들, 장치, 및 컴퓨터 프로그램 제품들을 포함한다.

[0009] 다른 양태에서, 시스템은 입력 이미지의 최근접 이웃 이미지인 하나 이상의 참조 이미지(reference image)- 하나 이상의 참조 이미지는 다중 이미지 라벨과 집합적으로 연관됨-를 결정하는 수단; 및 둘 이상의 이미지 라벨-들 이상의 이미지 라벨은 하나 이상의 참조 이미지에 연관된 다중 이미지 레벨 중에서 선택됨- 을 입력 이미지에 할당하는 수단을 포함한다. 하나 이상의 이미지 각각은 복수의 이미지 라벨과 연관될 수 있다. 이러한 실시예들 또는 다른 실시예들은 하나 이상의 후술하는 특징들을 선택적으로 포함할 수 있다. 복수의 이미지는 단일 최근접 이웃 이미지로부터 선택될 수 있다.

[0010] 복수의 이미지를 할당하는 단계는 다중 최근접 이웃 이미지들을, 최근접 이웃 이미지 각각과 연관되는 전체 이미지 거리(whole-image distance) 각각에 따라 분류하는 단계를 포함할 수 있다. 전체 이미지 거리 각각은 입력 이미지 및 상응하는 최근접 이웃 이미지 간의 차이 정도(degree of difference)를 나타낼 수 있다.

[0011] 복수의 이미지 라벨을 할당하는 단계는 디지털 이미지들의 집합에서 하나 이상의 제1 이미지 라벨-하나 이상의 제1 이미지 라벨 각각은 제1 최근접 이웃과 연관됨-에 각 이미지 라벨의 발생 빈도(frequency of occurrence)에 따라 순위를 부여하는 단계; 및 하나 이상의 제1 이미지 라벨들에 대한 순위에 기초하여 적어도 하나의 제1 이미지 라벨을 입력 이미지에 할당하는 단계를 더 포함할 수 있다.

[0012] 또한, 복수의 이미지 라벨을 할당하는 단계는 복수의 이미지 라벨을 입력 이미지에 할당하는 단계는, 하나 이상의 제2 이미지 라벨-제2 이미지 라벨 각각은 하나 이상의 잔여 최근접 이웃 이미지와 연관됨-에 순위를 부여하는 단계; 및 하나 이상의 제2 이미지 라벨들의 순위에 기초하여 적어도 하나의 제2 이미지 라벨을 입력 이미지에 할당하는 단계를 포함할 수 있다.

[0013] 하나 이상의 제2 이미지 라벨에 순위를 부여하는 단계는, 디지털 이미지들의 집합에서 제2 이미지 라벨 각각과 각 제1 이미지의 중복 발생(co-occurrence)에 따라 하나 이상의 제2 이미지 라벨을 분류하는 단계를 포함할 수 있다. 대안적으로, 또는 부가적으로, 하나 이상의 제2 이미지 라벨에 순위를 부여하는 단계는 하나 이상의 잔여 최근접 이웃 이미지에 있는 제2 이미지 라벨 각각의 국소적 빈도(local frequency)에 따라 하나 이상의 제2 이미지 라벨을 분류하는 단계를 포함할 수 있다.

[0014] 전체 이미지 거리는 특징 거리들(feature distances)의 조합을 포함할 수 있고, 각 특징 거리는 입력 이미지와 연관된 이미지 특징과, 참조 이미지와 연관된 이미지 특징 각각 간에 차이 정도를 나타낸다. 입력 이미지에 연관된 이미지 특징 및 참조 이미지에 연관된 각 이미지 특징은 전체적 이미지 특징(global image feature)을 각각 포함할 수 있다. 대안적으로, 또는 부가적으로, 입력 이미지에 연관된 이미지 특징 및 참조 이미지에 연관된 각 이미지 특징은 국소적 이미지 특징(local image feature)을 각각 포함할 수 있다.

- [0015] 전체 이미지 거리는 특정 거리의 평균으로서 유도될 수 있다. 그 평균은 실질적으로 특정 거리 각각으로부터의 동일한 기여에 기초할 수 있다. 그 평균은 특정 거리 각각으로부터의 가중된 기여에 기초할 수 있다. 특정 거리 각각에 대한 가중치는 디지털 이미지들의 집합에 기초하여 산출될 수 있고, 디지털 이미지들의 집합은 유사한 이미지와 유사하지 않은 이미지 쌍들을 포함하는 훈련 이미지(training image)들의 그룹이다.
- [0016] 특정 거리들 중 적어도 하나는 입력 이미지의 텍스처 특징(texture feature)과 참조 이미지의 상응하는 텍스처 특징 간의 차이로서 산출될 수 있다. 특정 거리들 중 적어도 하나는 입력 이미지의 색상 특징(color feature)과 참조 이미지의 상응하는 색상 특징 간의 차이로서 산출될 수 있다.
- [0017] 본 명세서에서 설명된 요지에 대한 특정 실시예는 하나 이상의 후술되는 효과를 실현하도록 구현될 수 있다. 일부 실시예에서, 이미지 주석 달기 기술은 최소 훈련 요구에 의해 특징화된다. 특정 구현예에서, 검색 문제로서 이미지에 주석 달기를 훈련하는 것은 주석 달기 프로세스를 단순화시킨다.
- [0018] 본 발명의 하나 이상의 실시예는 첨부 도면 및 아래의 설명에 상세하게 설명된다. 본 발명의 다른 특징, 양태, 및 장점은 상세한 설명, 도면, 및 청구항으로부터 명백하게 된다.

발명의 효과

- [0019] 본 발명에 따르면, 입력 이미지에 최근접 이웃 이미지들로부터의 복수의 이미지 라벨을 할당하여 웹 상에서 탐색될 수 있는 이미지에 주석을 다는 기술을 제공할 수 있다.
- 도면의 간단한 설명**
- [0020] 도 1은 키워드로 테스트 이미지에 주석을 다는 예시를 나타낸다.
- 도 2a는 입력 이미지의 최근접 이미지를 결정하는 단계에 대한 개관을 제공하는 순서도이다.
- 도 2b는 합성 거리(composite distance)를 유도하는 단계에 대한 개관을 제공하는 순서도이다.
- 도 3은 입력 이미지의 최근접 이웃들로부터의 키워드들을 입력 이미지에 전송하는 단계의 개관을 제공하는 순서도이다.
- 도 4는 적어도 4 개 키워드를 공통으로 갖는 이미지 쌍들의 예시를 나타낸다.
- 도 5는 공통인 키워드가 없는 이미지 쌍들의 예시를 나타낸다.
- 도 6과 도 7은 별개의 이미지 데이터 세트들로부터의 예시적 이미지들을 나타낸다.
- 도 8은 주석이 달린 이미지들에 대한 예시를 나타낸다.
- 도 9는 주석이 달린 이미지들에 대한 예시를 나타낸다.
- 도 10, 도 11, 및 도 12는 복수의 다른 키워드에 대하여 세 개의 다른 이미지 데이터 세트 내에서 각각 검색된 처음 몇 개의 이미지에 대한 예시를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0021] 이미지에 대해 키워드들을 자동적으로 할당함으로써, 대용량 이미지 데이터 집합을 검색하고, 색인하고, 조직화하고, 및 이해할 수 있다. 본 명세서는 이미지에 주석을 다는 기술을 설명하며, 이 기술은 검색 문제로서 주석 달기를 다룬다. 이 기술은 주어진 이미지에 대한 최근접 이웃들을 찾기 위하여, 저-레벨 이미지 특징(low-level image feature)을 이용하고, 또한 기본 거리 계측들의 단순 결합을 이용한다. 이어서, 키워드가 원하는 라벨 전송 메카니즘(greedy label transfer mechanism)을 사용하여 할당된다.
- [0022] 이미지 주석 달기는 후술되는 2가지 중요한 이유 때문에 어려운 작업이다. 첫째, 픽셀 대 속성 예측(pixel-to-predicate) 또는 의미 차이 과제(semantic gap problem)가 있는데, 단지 저 레벨 이미지 특징(예를 들어, 색상이나 텍스처)을 사용하여 의미적으로 유의한 실체(entity)를 추출하는 것은 어렵다. 수 많은 객체 또는 클래스를 확실하고 신뢰성있게 인식하는 것은 아직 해결하지 못한 문제이다. 두 번째 어려움은 훈련 데이터 내의 이미지 영역들과 키워드들 사이의 상응성(correspondence) 결여로 인해 발생한다. 각 이미지에 대하여, 그 이미지 전체에 할당된 키워드들에 접근해야 하지만, 이미지의 어느 영역이 그 키워드들에 상응하는지 알려지지 않는다. 이 때문에 클래스 분류자(classifier)들을 직접 학습할 수 없어서, 각 키워드는 개별 클래스인 것으로 간주된다.

- [0023] 본 명세서는 최소 훈련 요구에 의해 특징화되는 기술들을 설명한다. 이 기술들은 대용량 웹 데이터 세트뿐만 아니라 몇몇의 표준 데이터 세트에 대하여 복잡한 최신의 이미지 주석 달기 방법보다 우월한 성능을 갖는다.
- [0024] 도 1은 키워드로 테스트 이미지에 주석을 달는 예시를 나타낸다. 테스트 이미지(2)를 고려했을 때, 이미지의 훈련 세트(4)로부터 테스트 미지(2)에 대한 최근접 이웃(예를 들어, 미리-명시된 거리 단위로 임의 특정 영역에 정의된 제1 최근접 이웃, 제2 최근접 이웃(20), 제3 최근접 이웃(30))을 찾을 수 있고, 최근접 이웃 이미지들과 연관된 일부 또는 모든 키워드를 입력 테스트 이미지(2)에 할당할 수 있다. 일부 경우에는, 전체적 이미지 특징으로 정의된 단순 거리 단위들을 사용하는 것이 다른 주석 달기 기술보다 더 바람직하게 실행할 수 있다. 일부 구현예에서, K-최근접 이웃들은 정확하게 최근접한 이미지를 대신하여 키워드들을 할당하는데 사용된다. 다중 이웃의 경우, 적절한 키워드들이 주석 달기 성능을 보다 향상시키는 원하는 접근법을 사용하여 입력 이미지에 할당될 수도 있다.
- [0025] K-최근접 이웃 접근법은 특정 특정 영역을 통해 정의될 수 있는 다중 거리 단위를 포함하도록 확대될 수 있다. 다른 거리 또는 커널들의 결합은 객체 인식 작업에서 양호한 성능을 가져올 수 있다. 주석 달기 방법을 생성하기 위하여 다른 거리들을 결합하는 두 가지 다른 방법을 설명한다. 그 첫 번째 방법은 각 거리를 적절하게 스케일링한 후 다른 거리들의 평균을 산출한다. 두 번째 방법은 라소(Lasso)로 알려진 스파스 로지티스 회귀 방법(sparse logistic regression method)을 사용하여 관련 거리를 선택하는 것에 기초한다. 회귀 방법을 위해서, 유사한 이미지와 유사하지 않은 이미지들을 포함하는 훈련 세트가 사용된다. 주석 달기 작업을 위해 제공되는 전형적인 훈련 세트는 이러한 정보를 직접적으로 포함하지 않는다. 일부 구현예에서, 라소는 주석 달기 훈련 데이터로부터 라벨이 부여된 세트를 생성함으로써 훈련된다. 이처럼 미약하게 훈련된 라소이더라도 좋은 성능을 제공한다. 일부 경우에는, 평균된 거리 기술이 잘 알려진 라소 기술과 같거나 그 이상으로 수행된다.
- [0026] 이제, 모양이 비슷한 이미지들이 키워드를 공유할 수 있다는 전체 하에 구축된 방법인 이미지 주석 달기 방법을 설명한다. 이를 위하여, 이미지 주석 달기는 최근접 이웃으로부터의 키워드들을 전송하는 프로세스를 포함한다. 지역 구조(neighborhood structure)는 이미지 특징들을 사용하여 구축되어, 입력 이미지와 상응하는 참조 이미지에 있는 개별적 특징들 간의 거리 개념(notion of distance)에 의존하는 초기 모델(rudimentary model)이 된다.
- [0027] 도 2a는 입력 이미지의 최근접 이웃들을 결정하는 단계에 대한 전체적인 개관을 제공하는 순서도이다. 하나 이상의 컴퓨터에 구현된 서버는 디지털 입력 이미지를 수신(200)하기 위해 사용될 수 있다. 또한 이 서버는 선택된 하나 이상의 참조 이미지에 대한 디지털 이미지들의 집합을 수신할 수 있다. 디지털 이미지들은 서버의 데이터 저장소 또는 다른 컴퓨터-판독 가능 매체에 저장될 수 있다. 이어 서버는 입력 이미지와 디지털 이미지들의 집합으로부터 선택된 참조 이미지 간의 전체 이미지 거리를 유도한다(202). 전체 이미지 거리는 전체로서의 입력 이미지와 복수의 이미지 특징을 참조하는 전체로서의 참조 이미지 간의 차이 정도를 나타낸다. 이어 전체 이미지 거리는 서버의 디지털 데이터 저장소에 저장된다(204).
- [0028] 이미지 특징들은 전체적 특징(전체 이미지로부터 생성됨), 또는 국부적(局部的) 특징(예를 들어, 전체 이미지의 어느 한 부분에 한정되는, 흥미있는 부분으로부터 생성됨) 중 하나일 수 있다. 전체적 이미지 특징의 예시는 색상과 텍스처를 포함한다. 색상과 텍스처는 이미지를 표시하기 위한 2가지 저-레벨 시각적 지시(low-level visual cue)이다. 공통 색상 서술자(common color descriptor)는 정제되지 않은 픽셀 색상 값의 히스토그램에 기초한다. 이 색상 특징은 본질적인 그들의 효율성과 산출의 간단성 때문에, 이미지 매칭과 스키마 색인에 이용될 수 있다. 텍스처는 이미지 표시 요소가 될 수 있는 다른 저-레벨 시각적 특징이다. 이미지 텍스처는 웨이블릿 특징(Wavelet features)으로 취득될 수 있다. 특히, 가버(Gobor)와 하버(Haar) 웨이블릿은 아직 부족한 특유한 이미지 특징들을 생성하는데 상당히 효과적이다. 개별적 특징들에 대한 영향과 편향을 제한하고, 추출된 정보의 양을 최대화하기 위하여, 간단하고 쉽게 산출하는 복수의 색상 및 텍스처 특징들이 이용된다.
- [0029] 3가지 다른 색상 영역에 있는 이미지들로부터의 특징들이 생성된다. 특징들은 적색-녹색-파랑(RGB), 색도-채도-값(HSV), 및 CIE 1976 L, a*, b* (LAB) 색상 영역을 포함한다. RGB는 이미지 취득과 표시를 위한 기준 색상 영역인 반면, HSV와 LSB는 모두 RGB에 의해 취득되지 않는 중요한 외형 특성(appearance characteristics)을 분리시킨다. 예를 들어, HSV 색상 영역은 값 채널(Value channel)에서 색상을 비추는 빛의 양을 인코드하고, LAB의 휘도 채널(Luminance channel)은 밝기에 대한 사람의 인식을 반영하기 위한 것이다. RGB 특징은 RGB 픽셀 값의 표준화된 3D 히스토그램으로서 산출된다. 유사하게, HSV(및 LAB) 특징은 HSV(및 LAB) 색상 영역에서의 16-빈-퍼-채널(16-bin-per-channel histogram) 히스토그램이다. 각 칼라 영역을 위해 이용되는 거리 단위들을 결정하기 위하여, 히스토그램과 분포(KL-분기, L1-거리, L2-거리)를 위해 사용되는 3가지 거리 단위가 코렐5K 데이

터 세트(Core15K dataset)로부터 인간에 의해 레벨화된 훈련 데이터에 대해 평가된다. KL-분기는 2가지 잠재적 분포 들간의 차이에 대한 비호환적(commutative) 단위이다. 이산 랜덤 변수(discrete random variable)에 대한 2 가지 분포가 P1과 P2이면, KL-분기는 $\text{sum_i} (P1[i] * \log(P1[i] / P2[i]))$ 로서 산출된다. KL-분기는 LAB 거리에 적합하도록 찾았다는 반면, L1은 RGB와 HSV에 최적으로 수행된다. 이후, 이 명세서의 전체에서, RGB와 HSV 거리는 L1(맨하탄; Manhattan) 단위를 의미하고, LAB 거리는 KL-분기를 의미한다. 다른 거리 단위도 사용될 수 있다. 예를 들어, 일부 경우에는, 코사인 거리 단위 또는 EMD(earth mover's distance)가 사용될 수 있다.

[0030] 이미지에 대한 텍스쳐는 가바 웨이블렛과 하르 웨이블렛으로 표현될 수 있다. 본 구현예에서, 각 이미지는 3개의 스케일과 4 방향에서 가바 웨이블렛으로 필터링된다. 12개의 응답 이미지는 비-중첩 지역으로 나뉘지고, 각 지역으로부터의 평균 응답 크기는 특징 벡터로 연결된다(이 명세서 전체에서 이 특징은 "가버"라고 일컬어짐). 제2 특징은 양자화된 가버 위상을 취득한다. 응답 픽셀 각각에서의 위상각은 12개의 가버 응답 이미지의 각각에 있는 16×16 블록들에 대해 평균화된 것이다. 이 평균 위상각은 3비트(8개의 값)로 양자화되고, 특징 벡터("가버Q"로 명세서 전체에서 참조됨)로 연결된다. L1 거리는 가버 특징과 가버Q 특징을 위해 사용된다.

[0031] 하르 필터는 2×2 에지 필터(edge filter)이다. 하르 웨이블렛 응답은 3개의 다른 방위(수평, 대각선, 및 수직)에서 하르 필터가 이미지를 블록-컨볼루션(block-convolution)함으로써 생성된다. 스케일 차이에 대한 응답은 적절하게 서브 샘플링된 이미지로 컨볼루션을 수행함으로써 얻어진다. 이미지를 64×64 픽셀 크기로 리스케일링한 후, 하르 특징은 하르 응답 크기(이 특징은 단지 "하르"로 일컬어짐)를 연결함으로써 생성된다. 가버 특징과 마찬가지로, 하르 응답의 표시가 3개의 값(응답이 0, 양수, 또는 음수이면, 각각 "0", "1" 또는 "-1" 중 하나 임)으로 양자화되는 양자화된 버전이 또한 고려된다. 이후 본 명세서 전체에서, 이 양자화된 특징은 "하르Q"로 일컬어진다. L1-거리는 가버 특징과 마찬가지로 하르 특징 및 하르Q 특징을 위해 사용된다.

[0032] 전체적 이미지 특징의 다른 예시는 1) 아주 작은 크기(예를 들어, 엄지 손톱 크기)로 이미지 축소되고, 픽셀 대 픽셀로 비교되는 "아주 작은 이미지(Tiny Image)"; 2) 조정 가능한 필터에 의해 변환되고 취득되는 웨이블렛과 유사한 요지 변환(Gist transforms); 3) 라인 또는 다른 윤곽의 통계와 같은 기하학적 특징의 분포; 및 4) 모든 이미지에 대한 경사 방향 히스토그램을 포함한다. 다른 전체적 이미지 특징도 사용될 수 있다. 상술한 전체적 이미지 특징을 위한 거리 단위(distance measure)는 예를 들어, L1, L2, KL 분기, 코사인, 및 EMD를 포함할 수 있다.

[0033] 국부적 특징에 관련하여, 이미지로부터 특징들을 얻기 위한 2가지 요소가 있다. 첫째, 식별된 이미지들간 매칭 또는 비교에 유용할 수 있는 이미지 내에 지점(point)이나 지역(region)의 위치인 "홍미 지점 검출(interest point detection)"이다. 예를 들어, 코너는 공통 홍미 지점이다. 홍미 지점 검출 기술의 예시는 에지 검출(edge detection), 얼룩 검출(blob detection), 융기 검출(ridge detection), 및 아핀-불변 검출(affine-invariant detection)을 포함하지만, 이에 한정되지는 않는다. 두 번째 단계는 홍미 지점으로부터 생성되는 기술적 특징 벡터(descriptive feature vectors)인 "특징 추출"이다. 예를 들어, 특징 벡터는 코너 지역내 색상 분포를 설명하거나, 또는 특징 벡터는 코너각을 설명할 수 있다. 다른 국소적 기술적 특징(local descriptive features)의 예시는 스케일-불변(예를 들어, SIFT 서술자), 회전-불변, 기울기 크기, 기울기 방향 및 SURF(Speeded Up Robust Features; 예를 들어 SURF 서술자)를 포함하지만, 이에 한정되지 않는다.

[0034] 상술된 것처럼, 본 구현예에서 각 이미지는 7개의 특징(예를 들어, 3개의 색상 히스토그램과, 4개의 텍스쳐 특징)으로 표현되었다. 다른 이미지에 있는 상응하는 이미지 특징들간의 거리는 "기본 거리"이다. "합성 거리"는 7개의 특징의 일부 또는 전부를 포함하는 이미지들간의 거리 단위이다. 일부 구현예에서, 합성 거리는 추가 특징을 포함할 수 있다.

[0035] 도 2b는 합성 거리를 유도하는 단계에 대한 개관을 제공하는 순서도이다. 이전과 마찬가지로, 하나 이상의 컴퓨터 상에 구현된 서버는 입력 이미지와 참조 이미지를 수신한다(210). 이어서, 입력 이미지와 참조 이미지로부터 7개의 전체-이미지 특징이 추출된다(212). 추출된 특징은 3가지 색상 특징과 4개의 텍스쳐 특징을 포함한다. 색상 특징은 RGB 색상 영역에 있는 이미지 색상들의 히스토그램, HSV 색상 영역에 있는 이미지 색상들의 히스토그램, 및 LAB 색상 영역에 있는 이미지 색상들의 히스토그램을 포함한다. 텍스쳐 특징은 가버 응답의 크기에 대한 벡터, 가버 응답의 양자화된 위상에 대한 벡터, 하르 응답의 크기에 대한 벡터, 및 하르 응답의 표시에 대한 벡터를 포함한다.

[0036] 이어서, 입력 이미지와 참조 이미지로부터 추출된 특징들에 기초하여 기본 거리가 유도된다(214). RGB 색상 영역의 경우, 기본 거리는 각각의 RGB 특징 벡터들간의 L1 거리이다. L1 거리는 때때로 맨하탄 또는 시티 블록 거

리(city block distance)로 일컬어진다. 유사하게, HSV 색상 영역에 대한 기본 거리는 HSV 특징 벡터들간의 L1 거리이다. LAB 색상 영역에 대한 기본 거리는 LAB 특징 벡터들간의 KL-분기이다.

[0037] 입력 이미지와 참조 이미지에 대한 가버 특징, 가버Q 특징, 하르 특징, 및 하르Q 특징들 간의 기본 거리는 L1 거리 단위를 사용하여 결정된다. 각 특징을 위해 사용되는 거리 단위(L1, KL-분기)는 몇 개의 다른 거리 단위를 위해 설정된 간단한 훈련에서 각 특징들의 성능을 평가하고, 각 특징을 위한 최선을 선택함으로써 결정될 수 있다. L1과 KL-분기 이외의 거리 단위가 기본 거리를 산출하기 위해 사용될 수도 있다. 예를 들어, 임의 지역에 걸쳐진 2개의 분포들간의 차이에 대한 수학적 단위인 임의의 L_p 거리가 히스토그램 교점(histogram intersection) 또는 EMD에 사용될 수도 있다.

[0038] 전체 이미지 특징 각각에 대한 기본 거리를 얻은 후에, 그 거리가 스케일된다(216). 7개 특징 유형의 각각을 위하여, 스케일링 기간은 "0"과 "1"로 제한된 기본 거리 즉, 표준화된 기본 거리가 보장되는 훈련 데이터로부터 결정된다.

[0039] 이어서, 스케일링된 기본 거리는 합성 거리에 결합된다(218). 단순 기선 방법(simple baseline)은, 예를 들어 합성 거리 단위를 산출하기 위하여 기본 거리에 대한 선형 결합을 포함한다. 즉, 입력 이미지와 참조 이미지간의 합성 거리는 7개의 기본 거리의 평균화된 합이다. 7개의 특징이 사용되고 있지만, 그 알고리즘은 임의 개수의 특징(한개도 포함함)으로 용이하게 동작할 수 있다.

[0040] 일 실시예에서, 선형 결합은 기본 거리 각각이 총 결합 거리와 동일하게 기여하도록 함으로서 얻어진다. 이 방법은 JEC(Joint Equal Contribution)로 불리운다. 다른 실시예에서, 기본 거리는 이미지 유사성을 취득하기 위하여 더 관련되는 특징들을 우위함으로써 즉, 기본 거리에 가중치를 부여함으로써, 평등하지 않게 결합된다. 기본 거리를 결합하기 위한 가중치는 스페스 로지티스 회귀 기술(using the sparse logistic regression technique) 즉 라소(Lasso)를 사용하여 얻어질 수 있다. 추가 방법들이 이용될 수도 있다. 예를 들어, 일부 실시예에서, 선형 결합은 컴퓨터 비전 2007의 국제 회의에서 프로메(Frome) 등에 의해 발표된 " Learning Globally-Consistent Local Distance Functions for Shape-Based Image Retrieval and Classification"에 설명된 것과 같은, 최대 마진 접근법(max-margin approach)에 기초한다.

[0041] 라벨이 부여된 훈련 데이터를 이용할 수 없거나, 라벨이 극도로 잡음이 많다면, 다른 특징들로부터 합성 거리를 만드는 간단한 방법은 JEC 방법을 사용하는 것이며, JEC 방법은 개별적 기본 거리 각각이 총 결합 비용 또는 거리를 동일하게 기여하는 것이다. I_i 이 I 번째 이미지라 하고, 추출된 N 특징들 $F_i = f_i^1, \dots, f_i^N$ (본 명세서의 경우 $N=7$ 임)이 추출되어졌다고 가정한다. 기본 거리, $d_{(i,j)}^k$ 는 두 개의 이미지 I_i 와 I_j 에 있는 상응하는 특징 f_i^k

과 f_j^k 간에 산출된다. N 개별적 기본 거리 $d_{(i,j)}^k$, $k=1, \dots, N$, 는 이미지 I_i 와 I_j 간의 포괄적 거리 (comprehensive distance)를 제공하기 위해 결합된다. 기본 거리 각각이 0과 1 간에 놓여지도록 스케일되는 JEC에서, 각 스케일링된 기본 거리는 동일하게 기여한다. 스케일링 기간(scaling terms)은 훈련 데이터로부터 경험적으로 결정될 수 있다. $\tilde{d}_{(i,j)}$ 가 적절하게 스케일링된 거리를 나타내면, 이미지 I_i 와 I_j 간의 포괄적 이미지 거리는 $\frac{1}{N} \sum_{k=1}^N \tilde{d}_{(i,j)}^k$ 로 정의될 수 있다. 이 거리는 JEC(Joint Equal Contribution) 즉 간략 JEC이다.

[0042] 특징 거리를 결합하는 다른 접근법은 이미지 유사도(image similarity)를 취득하기 위하여 더 관련된 특징들을 식별하는 것일 수 있다. 다른 색상(및 텍스쳐) 특징이 완전하게 독립적이지 않기 때문에, 어떤 색상(또는 텍스쳐) 특징이 과잉되는지를 결정하는 것이 바람직하다. L1 패널티를 갖는 로지티스 회귀(즉, 라소)는 다른 특징들에 대한 관련성을 결정하는 간단한 방법을 제공할 수 있다.

[0043] 특징 선택에 로지티스 회귀를 적용하기 위하여, 이미지 주석 달기 시나리오는 라소 훈련을 위해 사용될 수 있는 것으로 변형되어야 한다. 이를 위하여, 새로운 세트 X 가 정의되며, 각 데이터 포인트 $x_i \in X$ 는 이미지(I_i , I_j)의 쌍이다. 훈련 세트는 $X = \{x_i = (I_i, I_j) | I_i, I_j \in S, i \neq j\}$ 로 주어지는데, 여기서 S 는 모든 훈련 이미지의

입력 세트이다. $y_l \in \{+1, -1\}$ 가 각 훈련 포인트(x_l)에 첨부되는 라벨이라고 한다. 쌍(I_i, I_j)이 "유사한" 이미지들을 포함하면, x_l 은 라벨 $y_l=1$ 이 할당되고, 이와 달리 쌍(I_i, I_j)이 "유사한" 이미지들을 포함하지 않으면, x_l 은 라벨 $y_l=-1$ 이 할당된다. 라소에서, 최선의 가중치 $\hat{\omega}$ 는 후술되는 패널티 부여를 최소화함으로써 얻어지고, 음 로그 우도(negative log-likelihood)은 [수학식 1]과 같다.

수학식 1

$$\hat{\omega} = \underset{\omega}{\operatorname{arg\,min}} \sum_{l=1}^L \log(1 + \exp(-\omega^T \mathbf{d}_{x_l, y_l})) + \lambda \|\omega\|_1$$

[0044] 여기서, L 은 훈련에 사용되는 이미지 쌍의 개수이고, $\|\cdot\|_1$ 은 L1 평균(norm)이고, \mathbf{d}_{x_l} 은 이미지 쌍 x_l 에 대한 개별적 기본 거리를 포함하는 벡터이고, λ 는 교차-검증(cross-validation)을 통해 조절된 양 가중치 파라미터(positive weighting parameter)이다. 훈련 데이터 $\{(x_l, y_l)\}$ 를 고려했을 때, 방정식 (a)는 그 훈련 데이터를 조건부 최적화 과제(constrained optimization problem)로 변환시킴으로써 해결될 수 있다. 상용하는 거리를 산출하기 위해 그 결과가 상쇄되도록 [수학식 1]에서 산출된 가중치를 이용한 기본 거리들의 선형 결합은 이미지 유사도 단위를 제공한다.

[0045] 상술한 이미지 주석 달기 스키마를 적용하는 시도는 유사한 이미지와 유사하지 않은 이미지를 포함하는 훈련 세트를 생성하는 것에 달려있다. 전형적인 이미지 주석 달기 데이터 세트는 각 이미지가 소수의 텍스트 키워드만을 포함하기 때문에 유사한 이미지와 유사하지 않은 이미지에 대한 정보들이 있지 않고, 이미지들간의 유사성(또는 비유사성)에 대한 개념이 없다. 이 설정에서, 충분한 키워드를 공유하는 이미지들의 임의 쌍은 양 훈련 예시(positive training example)이고, 공통적인 키워드가 없는 임의의 쌍은 부 훈련 예시(negative training example)이다. 이러한 훈련 세트의 품질(quality)은 이미지 쌍이 "유사하다"고 불리기 전에 매칭하기 위해 요구되는 키워드의 개수에 따라 달라진다. 높은 임계값은 더 순수한 훈련 세트를 보장하지만, 양 쌍(positive pair)의 개수를 감소시킨다. 반대로, 낮은 임계값은 이 쌍들의 품질을 희생하여 훈련을 위해 충분한 양 쌍들을 생성한다. 이 과정에서, 훈련 샘플은 지시된 코렐RK 기준점 훈련 세트(training set of the Corel15K benchmark)로부터 얻어진다. 적어도 4개 키워드를 공통으로 갖는 이미지들이 훈련용 양 샘플들로서 다뤄진다. 도 4는 적어도 4개의 키워드를 공통으로 갖는 이미지 쌍의 예시를 나타내고, 도 5는 0 키워드를 공통으로 갖는 이미지 쌍의 예시를 나타낸다. 키워드가 많이 중복된다고 해서 항상 더 좋은 이미지 유사도로 해석되지 않으며, 훈련 세트가 본질적으로 잡음이 많다는 것을 의미한다는 것에 주의한다.

[0046] JEC 또는 라소를 사용하여 기본 거리들을 결합하는 것은 이미지들간에 거리를 산출하는 단순한 방법을 제공한다. 이러한 합성 거리를 사용함으로써, 훈련 세트 내에 있는 테스트 세트로부터 이미지에 대한 K개의 최근접한 이웃을 찾아낼 수 있다. 이어서, 키워드들은 최근접 이웃 이미지들로부터의 테스트 이미지에 할당된다.

[0047] 도 3은 입력 이미지에 대한 최근접 이웃으로부터의 키워드를 입력 이미지에 전송하는 단계에 대한 전체적인 개관을 제공하는 순서도이다. 하나 이상의 컴퓨터 상에 구현된 서버는 입력 이미지를 수신하고(300), 일부 구현 예에서는, 참조 이미지 집합을 수신한다. 이어, 서버는 이미지 집합 중에서 입력 이미지에 대한 하나 이상식별하는데(302), 하나 이상의 최근접 이웃 이미지 각각은 하나 이상의 이미지 라벨 각각과 연관된다. 이어서, 서버는 하나 이상의 최근접 이웃 이미지와 연관된 이미지 라벨들로부터 선택된 복수의 이미지 라벨을 입력 이미지에 할당한다(304). 이어서, 할당된 복수의 이미지 라벨을 갖는 입력 이미지가 서버의 디지털 데이터 저장소에 저장된다(306). 라벨들을 포함하는 메타 데이터는 입력 데이터와 함께 저장소 또는 다른 곳에 저장될 수 있다.

[0048] 훈련 세트에 있는 쿼리의 K개 최근접 이웃들로부터의 이미지 I 를 쿼리하는 n개의 키워드들을 전송하기 위한 간단한 방법이 아래와 같이 설명된다. $I_i, i \in \{1, \dots, K\}$, 가 훈련 세트에 있는 이미지 I 에 대한 K개의 최근접 이웃들이라 하고, 거리가 증가함에 따라 순서가 정해진다고 가정한다(즉, I_1 는 가장 유사한 이미지임). I_i 와 연관

되는 키워드들의 개수는 $|I_i|$ 로 나타내진다. 원하는 라벨 전송 알고리즘의 단계는 아래 단계를 포함한다.

[0050] 1. 훈련 세트에 있는 키워드들의 빈도에 따라 최근접 이웃(I_1)의 키워드 각각에 점수를 부여함.

[0051] 2. I_1 의 $|I_1|$ 키워드들에 대해, n 개의 최고 점수 키워드들을 쿼리 \tilde{I} 에 전송함. 만약 $|I_1| < n$ 이면, 더 많은 키워드들을 전송해야 하므로, 3 단계로 진행함.

[0052] 3. 아래 두개의 인자에 따라서 I_2 내지 I_k 이웃들의 키워드 각각에 대한 순위를 부여함: 1) 훈련 세트 내에서 단계 2에서 전송된 키워드들과 함께, 키워드들이 중복 발생(co-occurrence), 및 2) 그 키워드들의 국소적 빈도(I_2 내지 I_k 이웃들의 키워드들로서 그들이 나타나는 정도). 이들 두 인자의 곱은, 정규화 후에, 이 키워드들의 순위를

부여하기 위해 필요한 점수를 제공한다. 이 키워드 순위에 기초하여, 쿼리 \tilde{I} 에 전송할 최선의 $n - |I_1|$ 키워드들을 선택함.

[0053] 반드시, 라벨 전송 스키마는 제1 최근접 이웃의 모든 키워드를 적용한다. 더 많은 키워드가 요구되면, N 을 통해 이웃들 2로부터 선택된다[2개의 변수(중복 발생과 빈도)에 기초함].

[0054] 요약하면, 설명된 이미지 주석 달기의 구현에는 아래의 단계들을 포함한다. 첫째, 합성 이미지 거리(JEC 또는 라소로 산출됨)는 입력 이미지에 대한 최근접 이웃들을 식별하는데 사용된다. 다음으로, 소망하는 개수의 키워드들이 최근접 이웃들로부터 입력 이미지에 전송된다.

[0055] 이미지 주석 달기 방법의 성능은 다른 이미지 데이터 세트들로 평가되었다. 도 6과 도 7은 2개의 분리된 이미지 데이터 세트(코렐5K 세트와 ESP 세트)로부터의 예시적 이미지를 나타낸다. 도 6에는 이미지 주석 달기 커뮤니티에서 사설 표준 평가 기준점(de facto evaluation benchmark)이 되는 코렐5K 데이터 세트로부터의 이미지가 도시되어 있다. 왼쪽에는 데이터 세트로부터 25개의 랜덤하게 선택된 이미지가 있다. 오른쪽에는 2개의 샘플 이미지와 그 이미지와 연관된 주석들이 있다. 이 세트는 대량 코렐 CD 세트로부터 수집된 5000개의 이미지를 포함한다. 이 세트는 374개의 키워드 사전으로부터 주석이 달려지고, 이미지 각각은 1개와 5개 사이, 및 평균 3.5개의 키워드로 주석이 달려져 있다. 374개의 키워드 중에서, 260 개만이 테스트 세트에 나타난다.

[0056] 도 7에는 ESP 이미지 데이터 세트로부터의 이미지가 도시되어 있다. 왼쪽에는 25개의 데이터 세트로부터 랜덤하게 선택된 이미지가 있다. 오른쪽에는 2개의 이미지와 그 이미지와 연관된 주석들이 있다. ESP 세트는 ESP 협력적 이미지 라벨링 게임으로부터 수집된 21844 이미지를 포함한다. ESP 게임은 2인용 게임이고, 서로간에 통신 할 수 있는 2명의 게임 참가자는 동일한 이미지에 대해 라벨을 할당하도록 요구받는다. 그들이 한 레벨을 공통적으로 갖는 즉시, 이미지에 성공적으로 라벨링한 것에 대한 신용(credit)을 제공받고, 다음 이미지를 보게 된다. 따라서, 대개 한 라벨은 한 쌍의 게임 참가자에게 이미지가 보여질 때마다 얻어진다. 이미지 각각이 더 많은 게임 참가자에게 보여질 때, 금기어(taboo words) 목록이 생성된다. 이어서 게임의 참가자들은 동일한 이미지가 보여졌을 때 금기어를 할당하는 것이 허용되지 않는다. 이 규칙들은 각 이미지에 많은 다른 게임 참가자에 의해 많은 다른 라벨이 할당되는 것을 보장한다. 사용된 세트는 자연 풍경, 인공 풍경, 및 269 키워드에 의해 주석이 달려진 객체들의 매우 다양한 이미지들을 포함한다. 각 이미지에는 적어도 1개의 키워드, 최대 15개의 키워드, 및 평균 4.6개의 키워드가 달려진다.

[0057] 5개의 키워드가 라벨 전송을 사용하여 각 이미지에 할당된다. 일 실시예에서, JEC 스키마는 5개의 키워드가 코렐 5K 데이터 세트에 있는 테스트 이미지 각각에 할당하기 위하여 라벨 전송 알고리즘에 사용된다. 도 8은 주석이 달린 이미지에 대한 예시 즉, 복수의 샘플 이미지를 위한 지상 조사 정보(ground-truth; 예를 들어, 사람이 할당한 것) 키워드에 대응하는 예측된 키워드들의 비교를 나타낸다. 사람에 의해 달린 주석(human-annotations)은 주로 5개 미만의 키워드를 갖기 때문에, 일부 경우 JEC는 지상 조사 정보 세트에 있지 않지만, 그럼에도 불구하고 정확하게 이미지 컨텐츠를 설명하는 키워드를 예측한다. 예를 들어, 도면에 있는 첫 번째 이미지는 키워드 대형(formation)을 갖는다고 예측되었다. 분명히, 사람에 의해 할당한 키워드 중 하나는 아니더라도, 이미지에 있는 비행기들에 대한 정확한 설명이다.

[0058] 도 9는 주석이 달려진 이미지의 예시를 나타낸다. 이미지들은 ESP 이미지 데이터 세트로 JEC 스키마를 사용하여 주석이 달려졌다. JEC 주석 달기 방법을 사용한 예측된 키워드들이더라도, 사람에 의해 달린 주석과 완전하게 중복되지 않고, 많은 경우 "사실과는 다른" 예측된 키워드들이 정확하게 이미지를 설명한다. 예를 들어, 사진

액자가 가득 걸린 벽 앞에 있는 소파에 앉아있는 남자를 나타내는 네 번째 이미지에서, 분명히 JEC로 할당된 키워드는 ESP 게임을 통해 생성된 키워드들만큼(또는 더욱) 정확하거나 더 정확하게 이미지를 설명한다.

[0059] 이미지에 주석 달기 작업에 대한 문제는 얼마나 많은 키워드가 이미지에 대한 컨텐츠를 설명하기 위해 필요한지를 아는 것이다. 라벨 전송 단계 동안 5개의 키워드 만을 할당하는 것은 많은 이미지를 위해 정확하게 회수될 수 있는 키워드의 개수를 인위적으로 제한한다. 이미지에 대해 할당된 키워드 개수가 증가하는 것은 회수할 키워드를 증가시킬 수 있지만(예를 들어, 극단적인 경우에, 모든 키워드가 이미지 데이터 세트에 있는 각 이미지에 할당되면, 100% 회수가 모든 키워드에 대해 보장될 수 있음), 이는 정확도에 감소를 가져온다. 주석 달기 방법을 사용하여 이미지에 대해 5개 이상의 키워드를 할당하기 위하여, 라벨 전송 단계 동안 사용된 최근접 이웃의 개수는 고유한 키워드들을 충분히 확인하기 위해 필요한 최소로 설정된다. 그러나, 이는 회수되는 키워드의 증가로 인하여 정확도에 감소를 가져올 수 있다. 이는 적은 수의 이웃에 적합한 최근접 이웃 구조가 라벨 전송을 위해 사용된다는 사실에 기인한다. 이는 이웃의 수가 증가함에 따라 더 많은 에러를 유발하지만, 복수의 이웃은 많은 키워드를 할당하기 위해 필요하다.

[0060] 이미지에 대해 기술적인 키워드들을 할당하는 것은 사용자가 텍스트 기반 쿼리를 만을 사용하여 이미지를 검색 할 수 있도록 한다. 이미지 검색 엔진에 대한 성능을 평가하는 것은 검색에 대한 것이기 때문에 주석 달기 엔진에 대한 성능 평가와는 다르고, 주어진 키워드에 연관되는 처음 몇 개 이미지의 품질에만 관심이 있다. 도 10, 도 11, 및 도 12는 3개의 다른 이미지 데이터시트 내에 있는 복수의 다른 키워드에 대해 각각 검색된 처음 몇 개 이미지에 대한 예시를 나타낸다.

[0061] 특정 도전 키워드(예를 들어, 자전거 타는 사람, 두개골, 다이아그램, 및 타이)에 대해서, 많은 상위 검색 이미지가 정확하다. 또한, 많은 키워드는 다중 의미를 갖는데, 일반적으로 "단어 의미(word sense)"로 일컬어진다. 이러한 실시예의 일부에서, 검색된 이미지가 단어의 다양한 의미에 따라 놓여진다. 예를 들어, 도 12에 있는 키워드 "링"에 대해 검색된 이미지들은 단어 "링"의 다른 의미들을 나타낸다.

[0062] 제안된 이미지 주석 달기 방법들은 기본 거리 계측들을 아주 단순한 전체적 색상 특징과 텍스처 특징에 결합한다. 이 결합된 거리를 사용하여 산출된 K개의 최근접 이웃은 원하는 레벨 전송 알고리즘의 기준을 형성한다.

[0063] 본 명세서에 기재된 요지와 기능적 동작들의 실시예들은 디지털 전자 회로로 구현되거나, 또는 상세한 설명에 기재된 구조 및 그들의 구조적 등가물을 포함하는 컴퓨터 소프트웨어, 펌웨어, 또는 하드웨어로 구현되거나, 또는 이들 중 하나 이상의 조합으로 구현될 수 있다. 본 명세서에 기재된 요지의 실시예들은 하나 이상의 컴퓨터 프로그램 제품, 즉, 데이터 프로세싱 장치에 의해 실행되거나 또는 그 장치의 동작을 제어하도록, 컴퓨터 저장 매체에 인코딩된 컴퓨터 프로그램 명령의 하나 이상의 모듈로서 구현될 수 있다. 대안적 또는 부가적으로, 프로그램 명령들은 인공적으로 생성된 전파되는(propagated) 신호(예를 들어, 기계-생성 전기, 광 또는 전자기 신호)로서, 데이터 처리 장치에 의한 실행을 위하여 적절한 수신 장치로 전송하기 위한 정보를 인코딩하기 위해 생성된다. 컴퓨터 저장 매체는 컴퓨터-판독가능 저장 디바이스, 컴퓨터 판독가능 저장 기판(substrate), 랜덤 또는 시리얼 액세스 메모리 어레이 또는 디바이스, 이들 중 하나 이상의 조합일 수 있다.

[0064] 본 명세서에서 설명된 동작들은 하나 이상의 컴퓨터 판독가능 매체 디바이스에 저장되거나 다른 소스로부터 수신된 데이터에 대해 데이터 처리 장치에 의해 동작들이 실행됨으로써 구현될 수 있다.

[0065] "데이터 프로세싱 장치"라는 용어는 데이터를 처리하기 위한 모든 장치, 디바이스 및 기계를 포함하며, 예를 들어, 프로그래머블 프로세서, 컴퓨터, 칩상의 시스템, 또는 이들의 조합을 포함한다. 장치는 FPGA(field programmable gate array) 또는 ASIC(application specific integrated circuit)과 같은 전용 논리 회로가 프로세스를 포함할 수 있다. 장치는 또한 하드웨어 외에도, 당해 컴퓨터 프로그램에 대한 실행 환경을 생성하는 코드를 포함하고, 코드는 예를 들어, 프로세서 펌웨어, 프로토콜 스택, 데이터베이스 관리 시스템, 운영 시스템, 가상 기계와 같은 교차-플랫폼 런타임 환경, 또는 이들 중 하나 이상의 조합을 구성한다. 장치와 실행 환경은 예를 들어 웹 서비스, 분산된 컴퓨팅 구조, 및 그리드 컴퓨팅 구조와 같이, 다양한 다른 컴퓨팅 모델 구조로 구현될 수 있다.

[0066] 컴퓨터 프로그램(프로그램, 소프트웨어, 소프트웨어 애플리케이션, 스크립트 또는 코드로도 알려짐)은 컴파일 또는 인터프리터 언어나 선언적 또는 절차적 언어를 포함하는 모든 형태의 프로그래밍 언어로 작성될 수 있으며, 독립형 프로그램이나 모듈, 컴포넌트, 서브루틴 또는 컴퓨터 환경에서 사용하기에 적합한 그 밖의 유닛을 포함하는 임의의 형태로도 배치될 수 있다. 컴퓨터 프로그램은 파일 시스템의 파일에 반드시 상응해야 하는 것은 아니다. 프로그램은 다른 프로그램 또는 데이터를 보유하는 파일의 일부에 저장되거나(예를 들어, 마크업

언어 문서 내에 저장되는 하나 이상의 스크립트), 당해 프로그램 전용의 단일 파일에 저장되거나, 또는 다수의 조화된(coordinated) 파일들(예를 들어, 하나 이상의 모듈, 서브프로그램, 코드의 부분을 저장하는 파일)에 저장될 수 있다. 컴퓨터 프로그램은 하나의 컴퓨터에서, 또는 한 위치에 배치되거나 또는 다수의 위치에 걸쳐서 분산되고 통신 네트워크에 의해 접속된 다수의 컴퓨터에서 실행되도록 배치될 수 있다.

[0067] 본 명세서에 설명된 프로세스와 논리 흐름은 하나 이상의 프로그래머블 프로세서에 의해 수행될 수 있고, 이 프로그래머블 프로세서는 입력 데이터에 작용하여 출력을 생성함으로써 기능을 수행하는 하나 이상의 컴퓨터 프로그램들을 실행한다. 예를 들어, FPGA(field programmable gate array) 또는 ASIC(application specific integrated circuit)과 같은 전용 논리 회로가 프로세스와 논리 흐름을 수행하거나, 장치를 구현할 수 있다.

[0068] 컴퓨터 프로그램의 실행에 적합한 프로세서에는, 예를 들어, 범용 및 전용 마이크로프로세서, 및 임의 종류의 디지털 컴퓨터 중 하나 이상의 프로세서가 있다. 일반적으로, 프로세서는 판독 전용 메모리(ROM), 또는 랜덤 액세스 메모리(RAM), 또는 양자로부터 명령과 데이터를 수신한다. 컴퓨터의 필수 구성요소는 명령을 실행하는 프로세서, 및 명령과 데이터를 저장하는 하나 이상의 메모리 디바이스이다. 일반적으로, 컴퓨터는 데이터를 저장하기 위한 하나 이상의 대용량 저장 디바이스(예를 들어, 자기 디스크, 광자기 디스크, 또는 광디스크)를 포함하거나, 또는 이 디바이스와 데이터를 송수신하기 위하여 동작적으로(operatively) 결합될 수 있다. 하지만 컴퓨터는 이러한 디바이스를 반드시 구비할 필요는 없다. 더욱이, 컴퓨터는 예를 들어, 모바일 전화기, 개인 정보 단말(PDA), 모바일 오디오 또는 비디오 재생기, 게임 콘솔, GPS(global positioning system) 수신기 등과 같은 다른 디바이스에 내장될 수 있다. 컴퓨터 프로그램 명령어와 데이터를 저장하기 적합한 컴퓨터 판독가능 매체에는, 예를 들어, 반도체 메모리 디바이스(예를 들어, EPROM, EEPROM, 플래시 메모리 디바이스); 자기 디스크(예를 들어, 내부 하드디스크, 착탈식 디스크); 광자기 디스크; 및 CD ROM과 DVD-ROM 디스크를 포함하는 모든 형태의 비휘발성 메모리, 매체 및 메모리 디바이스가 포함된다. 프로세서와 메모리는 전용 논리 회로에 의해 보완되거나 또는 전용 논리 회로에 통합될 수 있다.

[0069] 사용자와의 상호작용을 제공하기 위하여, 본 명세서에 기술된 요지의 실시예는, 정보를 사용자에게 디스플레이하기 위한 디스플레이 디바이스(예를 들어, CRT(cathode ray tube) 또는 LCD(liquid crystal display) 모니터), 키보드 및 포인팅 디바이스(예를 들어, 마우스 또는 트랙볼)를 구비한 컴퓨터에 구현될 수 있다. 사용자는 키보드와 포인팅 디바이스를 이용하여 컴퓨터에 입력을 제공할 수 있다. 사용자와의 상호작용을 제공하기 위하여 다른 종류의 디바이스가 또한 사용될 수 있다. 예를 들어, 사용자에게 제공되는 피드백(feedback)은 예를 들어, 시각 피드백, 청각 피드백 또는 촉각 피드백인 임의 형태의 감각 피드백일 수 있고, 사용자로부터의 입력은 음향, 음성 또는 촉각 입력을 포함하는 임의의 형태로 수신될 수 있다.

[0070] 본 명세서에 기술된 요지의 실시예는, 예를 들어, 데이터 서버와 같은 백엔드(back-end) 구성요소를 구비하는 컴퓨팅 시스템; 또는 예를 들어, 애플리케이션 서버와 같은 미들웨어 구성요소를 구비하는 컴퓨팅 시스템; 또는 예를 들어, 사용자가 본 명세서에 기술된 요지의 구현에와 상호작용할 수 있는 그래픽 사용자 인터페이스 또는 웹브라우저를 구비한 클라이언트 컴퓨터와 같은 프론트엔드(front-end) 구성요소를 구비하는 컴퓨터 시스템; 또는 이러한 백엔드, 미들웨어 또는 프론트엔드 구성요소들의 임의 조합을 구비하는 컴퓨팅 시스템으로 구현될 수 있다. 시스템의 구성요소는 디지털 데이터 통신의 임의 형태 또는 매체(예를 들어, 통신 네트워크)에 의해 상호 접속될 수 있다. 통신 네트워크의 예에는 근거리 네트워크(LAN)와 인터넷과 같은 광역 네트워크(WAN)가 포함된다.

[0071] 컴퓨팅 시스템은 클라이언트와 서버를 포함할 수 있다. 클라이언트와 서버는 보통 서로 떨어져 있으며, 일반적으로는 통신 네트워크를 통하여 상호작용한다. 클라이언트와 서버의 관계는 각각의 컴퓨터상에서 실행되고 상호 클라이언트-서버 관계를 갖는 컴퓨터 프로그램에 의하여 발생한다.

[0072] 본 명세서가 다수의 특정한 구현 세부사항을 포함하고 있지만, 이는 발명의 범위나 청구할 사항의 범위에 대한 어떠한 제한으로서도 이해되어서는 안 되며, 본 발명의 특정한 실시예에 고유할 수 있는 특징의 설명으로서 이해되어야 한다. 별개의 실시예의 문맥으로 본 명세서에서 설명된 소정 특징은 조합되어 단일 실시예로 구현될 수 있다. 반대로, 단일 실시예의 문맥에서 설명한 다양한 특징은 복수의 실시예에서 별개로 구현되거나 어떤 적당한 하위 조합으로서도 구현 가능하다. 또한, 앞에서 특징이 소정 조합에서 동작하는 것으로서 설명되고 그와 같이 청구되었지만, 청구된 조합으로부터의 하나 이상의 특징은 일부 경우에 해당 조합으로부터 삭제될 수 있으며, 청구된 조합은 하위 조합이나 하위 조합의 변형으로 될 수 있다.

[0073] 마찬가지로, 도면에서 특정한 순서로 동작을 묘사하고 있지만, 그러한 동작이 바람직한 결과를 얻기 위해, 도시한 특정 순서나 순차적인 순서로 수행되어야 한다거나, 설명한 모든 동작이 수행되어야 한다는 것을 의미하는

것은 아니다. 소정 환경에서, 멀티태스킹 및 병렬 프로세싱이 바람직할 수 있다. 또한, 상술한 실시예에 있어서 다양한 시스템 구성요소의 분리는 모든 실시예에서 그러한 분리를 요구하는 것으로 이해되어서는 안 되며, 설명한 프로그램 구성요소와 시스템은 단일 소프트웨어 제품으로 통합되거나 또는 복수의 소프트웨어 제품으로 패키지될 수 있다는 점을 이해되어야 한다.

[0074]

본 명세서에서 설명한 요지의 특정 실시예가 기술되었다. 그 밖의 실시예는 후술하는 청구범위 내에 속한다. 예를 들어, 청구항에 인용된 동작들은 상이한 순서로 수행될 수 있지만, 여전히 바람직한 결과를 달성한다.

부호의 설명

[0075]

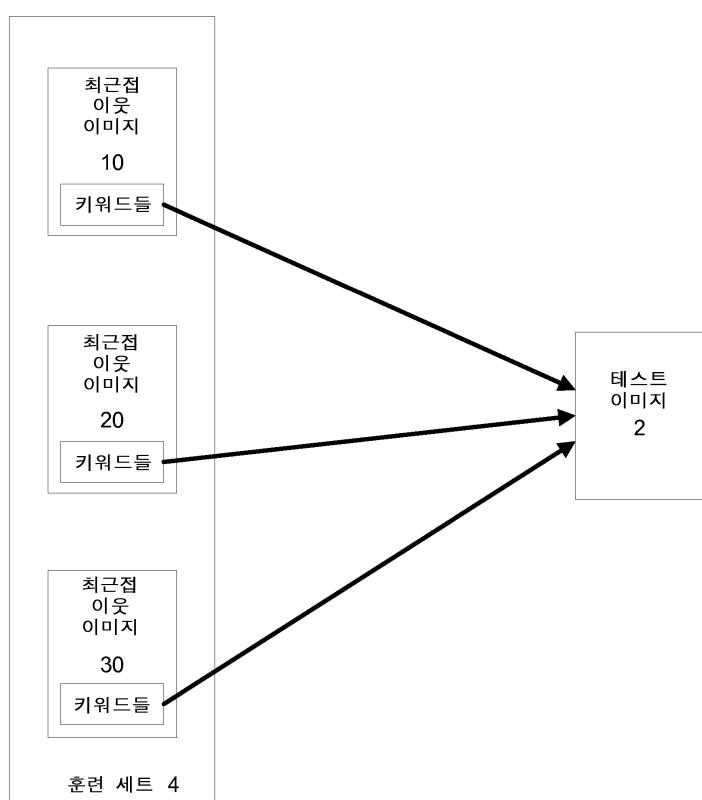
2: 테스트 이미지

3: 훈련 세트

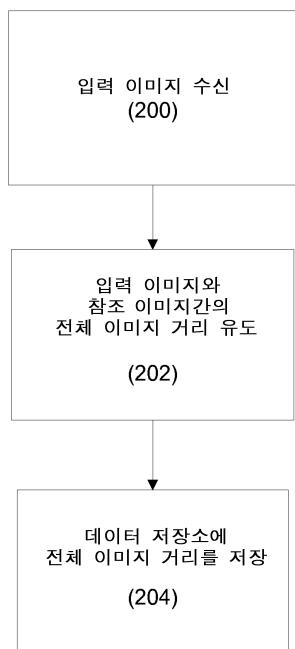
10, 20, 40: 최근접 이웃 이미지

도면

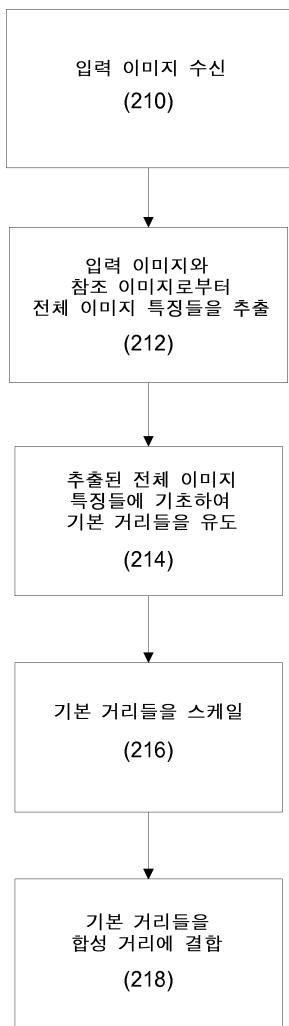
도면1



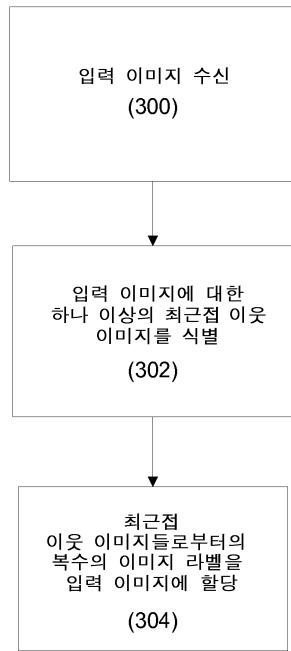
도면2a



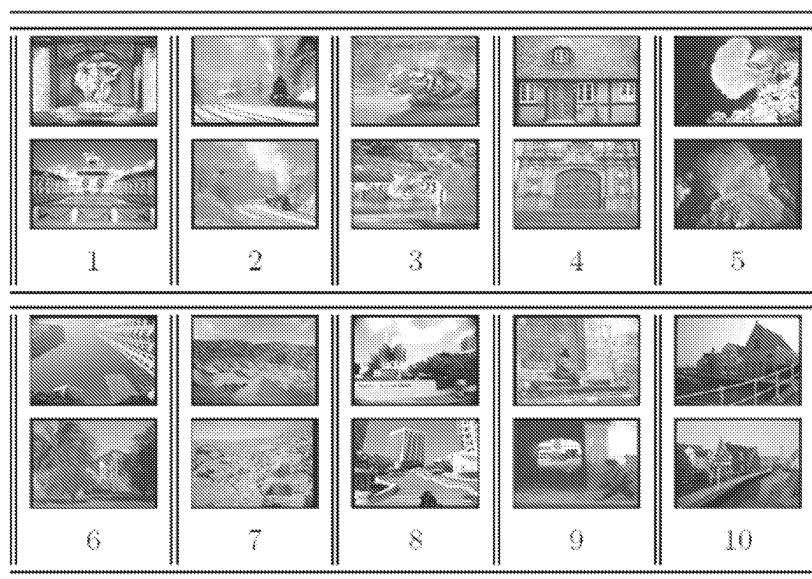
도면2b



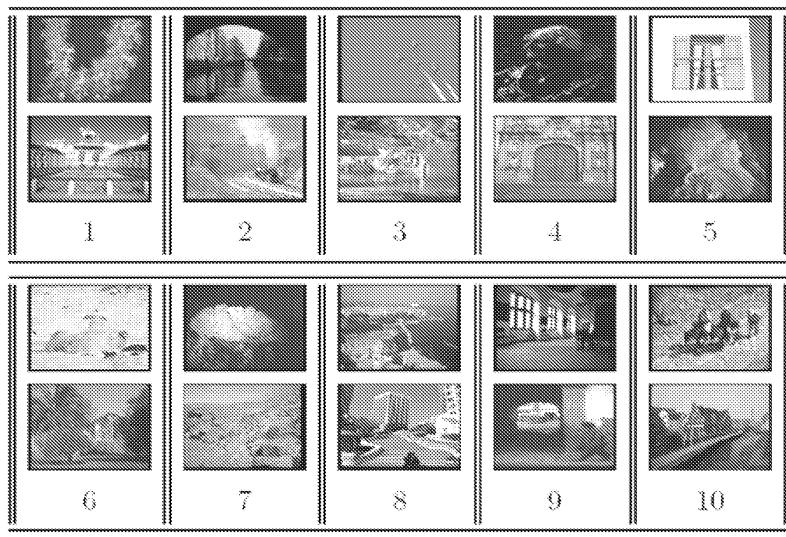
도면3



도면4



도면5



도면6



도면7



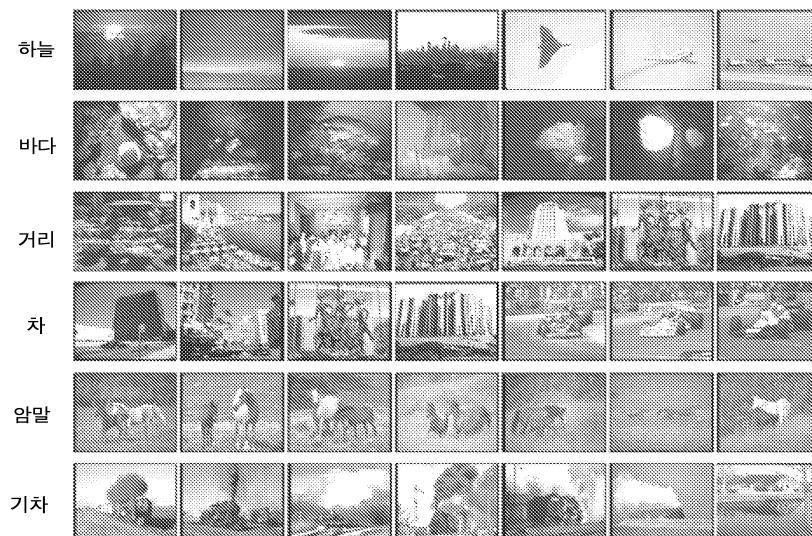
도면8

					
예측된 키워드	하늘, 제트 비행기, 연기, 대형	풀, 바위들, 모래, 계곡, 협곡	태양, 물, 바다, 파도, 새들	물, 나무, 풀, 사슴, 흰꼬리	곰, 눈, 시슴, 흰꼬리
사람에 의해 달린 주석	하늘, 제트 비행기, 연기	바위들, 모래, 계곡, 협곡	태양, 물, 구름들, 새들	나무, 숲, 사슴, 흰꼬리	나무, 눈, 목재, 여우

도면9

					
예측된 키워드	비키니, 소녀, 풀, 머리카락, 여자	곰, 검정, 갈색, 코, 흰색	밴드, 조명, 남자, 음악, 연주	남자, 노인, 사진, 빨강, 벽	구름, 풀, 녹색, 언덕, 빨강
사람에 의해 달린 주석	침대, 소녀, 여자	동물, 곰, 검정, 갈색, 머리, 코	밴드, 조명, 남자, 음악, 빨강, 바퀴	검정, 남자, 노인, 빨강, 안다	구름, 회색, 녹색, 산, 사진, 바위, 하늘, 돌

도면10



도면11



도면12

