



(19) 대한민국특허청(KR)
 (12) 등록특허공보(B1)

(45) 공고일자 2018년09월05일
 (11) 등록번호 10-1895647
 (24) 등록일자 2018년08월30일

- (51) 국제특허분류(Int. Cl.)
H04W 4/02 (2018.01)
 (21) 출원번호 10-2013-7033318
 (22) 출원일자(국제) 2012년06월13일
 심사청구일자 2017년05월16일
 (85) 번역문제출일자 2013년12월16일
 (65) 공개번호 10-2014-0043393
 (43) 공개일자 2014년04월09일
 (86) 국제출원번호 PCT/US2012/042106
 (87) 국제공개번호 WO 2012/174024
 국제공개일자 2012년12월20일
 (30) 우선권주장
 13/162,591 2011년06월17일 미국(US)
 (56) 선행기술조사문헌
 US20110150324 A1
 US20100009713 A1

- (73) 특허권자
마이크로소프트 테크놀로지 라이센싱, 엘엘씨
 미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
 마이크로소프트 웨이
 (72) 발명자
해다우 바르샤
 미국 워싱턴주 98052-6399 레드몬드 원 마이크로
 소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마
 이크로소프트 코포레이션
신하 수답타
 미국 워싱턴주 98052-6399 레드몬드 원 마이크로
 소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마
 이크로소프트 코포레이션
 (뒷면에 계속)
 (74) 대리인
제일특허법인(유)

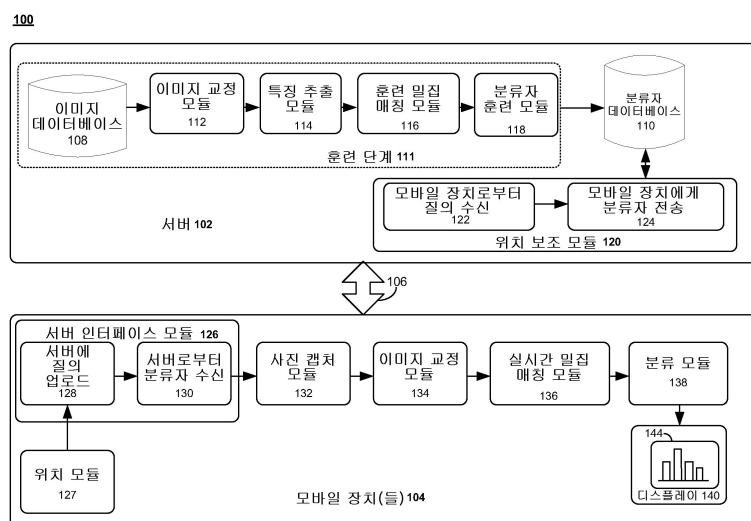
전체 청구항 수 : 총 20 항

심사관 : 조춘근

(54) 발명의 명칭 위치 기반 인식 기법

(57) 요약

서버로부터의 도움으로 실시간 위치 인식을 실행할 수 있는 모바일 장치가 제공된다. 모바일 장치의 대략적인 지구 물리학적 위치가 서버로 업로드된다. 서버는 이에 대한 응답으로서, 모바일 장치의 대략적인 지구 물리학적 위치에 기초하여 모바일 장치에 분류자 및 특징 기술자들의 세트를 포함하는 메시지를 전송한다. 이러한 과정은 이미지가 시각적 질의를 위해 캡처되기 이전에 이루어질 수 있다. 질의에서의 연산을 최소화하기 위해 분류자 및 특징 기술자들은 오프라인 훈련 단계에서 연산된다. 분류자 및 특징 기술자들은 모바일 장치 스스로 분류를 수행하는 것에 의한 실시간 시각적 인식을 실행하기 위해 사용된다.

대 표 도

(72) 발명자

지트닉 찰스 로렌스

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

스제리스키 리차드

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

명세서

청구범위

청구항 1

실행되는 경우 동작들을 수행하는 프로세서 실행가능 명령어가 저장된 컴퓨터 판독가능 저장 매체로서,
상기 동작들은,

하나 이상의 특징 기술자(feature descriptors)를 획득하는 동작과,

캡처된 이미지에 대한 각 특징 기술자의 유사도 스코어를 결정하는 동작- 상기 유사도 스코어는 상기 캡처된 이미지 내의 이미지 패치에 대한 특정 기술자의 가장 근접한 매치를 나타내고, 상기 유사도 스코어는 상기 캡처된 이미지 내의 이미지 패치에 매칭되는 특정 기술자에 대해 계산되고, 상기 매칭되는 특정 기술자는 민 해시 프로세스(min-hash process)에 의해 결정됨 -과,

상기 유사도 스코어의 사용을 통해 상기 캡처된 이미지를 인식하는 동작을 포함하는
컴퓨터 판독가능 저장 매체.

청구항 2

제1항에 있어서,

상기 특징 기술자를 획득하는데 사용되는 지구 물리학적 위치(geophysical location)를 결정하는 동작을 더 포함하되, 상기 특징 기술자는 상기 지구 물리학적 위치와 연관된
컴퓨터 판독가능 저장 매체.

청구항 3

제1항에 있어서,

지구 물리학적 위치(geophysical location)와 연관된 분류자(classifier)를 획득하는 동작- 상기 지구 물리학적 위치는 상기 특징 기술자와 연관됨 -과,

상기 분류자를 상기 특징 기술자 및 상기 유사도 스코어에 적용하여, 상기 캡처된 이미지가 소정의 클래스(a class)에 속하는 적어도 하나의 확률을 결정하는 동작을 더 포함하는
컴퓨터 판독가능 저장 매체.

청구항 4

제1항에 있어서,

상기 유사도 스코어의 결정에 앞서 상기 캡처된 이미지를 교정(rectifying)하는 동작을 더 포함하는
컴퓨터 판독가능 저장 매체.

청구항 5

제1항에 있어서,

상기 유사도 스코어를 해밍 거리로서 표현하는 동작과,

매칭되는 이미지 패치를 갖는 것으로 식별된 특정 기술자에 대해 상기 해밍 거리를 연산하는 동작을 더 포함하는

컴퓨터 판독가능 저장 매체.

청구항 6

실행되는 경우 동작들을 수행하는 프로세서 실행가능 명령어가 저장된 컴퓨터 판독가능 저장 매체로서,

상기 동작들은,

하나 이상의 훈련 이미지로부터 하나 이상의 특정 기술자를 획득하는 동작- 상기 훈련 이미지는 공통의 지구 물리학적 위치(common geophysical location)와 연관됨 -과,

각 훈련 이미지에 대한 각 특정 기술자의 유사도 스코어를 결정하는 동작 - 상기 유사도 스코어는 특정 기술자와 훈련 이미지 내의 이미지 패치에 대응하는 모든 특정 기술자 중 가장 유사한 특정 기술자 사이의 유사도를 나타내고, 매칭되는 특정 기술자는 훈련 이미지에 관해 상기 특정 기술자에 적용된 민 해시 프로세스를 사용하는 것에 의해 가속화되는 해밍 거리 연산을 수행함으로써 결정됨 -과,

상기 유사도 스코어의 사용을 통해 분류자(classifier)를 훈련시키는 동작을 포함하는

컴퓨터 판독가능 저장 매체.

청구항 7

제6항에 있어서,

보다 핵심적인 특징(more salient features)을 갖는 상기 훈련 이미지 내의 이미지 패치로부터 상기 특정 기술자를 선택하는 동작을 더 포함하는

컴퓨터 판독가능 저장 매체.

청구항 8

제6항에 있어서,

상기 분류자로서 하나 이상의 랜덤 결정 트리(random decision tree)를 생성하는 동작을 더 포함하되, 각 랜덤 결정 트리는 상기 특정 기술자 및 상기 유사도 스코어를 사용하여 훈련되는

컴퓨터 판독가능 저장 매체.

청구항 9

제6항에 있어서,

상기 유사도 스코어를 해밍 거리로서 나타내는 동작과,

훈련 이미지 내의 이미지 패치에 매칭되는 특정 기술자에 대한 상기 해밍 거리를 연산하는 동작을 더 포함하는

컴퓨터 판독가능 저장 매체.

청구항 10

제9항에 있어서,

민 해시 값(min-hash values)이 유사한 특정 기술자들을 결정하기 위해 상기 특정 기술자에 민 해시 메소드

(min-hash method)를 적용하는 동작과,

민 해시 스케치(min-hash sketch)가 유사한 특정 기술자들에 대한 상기 해밍 거리를 연산하는 동작을 더 포함하는

컴퓨터 관독가능 저장 매체.

청구항 11

지구 물리학적 위치(geophysical location)와 연관된 제1 세트의 특정 기술자 및 컴팩트한 분류자를 수신하는 단계와,

상기 지구 물리학적 위치와 연관된 이미지를 캡처하는 단계- 상기 이미지는 미지의 클래스를 가짐 -와,

상기 캡처된 이미지로부터 제2 세트의 특정 기술자를 생성하는 단계와,

상기 제2 세트 내의 각 특정 기술자와 함께 상기 제1 세트 내의 각 특정 기술자에 대한 민 해시 값(min-hash value)을 연산하는 단계와,

상기 제1 세트 내의 특정 기술자와 유사한 민 해시 값을 갖는 상기 제2 세트 내의 각 특정 기술자에 대한 유사도 스코어를 생성하는 단계와,

상기 캡처된 이미지의 클래스를 결정하기 위해 상기 유사도 스코어에 기초하여 상기 컴팩트한 분류자를 상기 캡처된 이미지에 적용하는 단계를 포함하는

컴퓨터 구현 방법.

청구항 12

제11항에 있어서,

제1 장치에게 상기 컴팩트한 분류자를 요청하는 단계를 더 포함하되,

상기 요청은 상기 제1 장치의 상기 지구 물리학적 위치를 포함하고,

상기 컴팩트한 분류자는 상기 제1 장치로부터 원격으로 위치한 제2 장치에 의해 수신되는

컴퓨터 구현 방법.

청구항 13

제11항에 있어서,

상기 컴팩트한 분류자는 복수의 랜덤 결정 트리(random decision tree)를 포함하는 랜덤 결정 트리군(random decision tree forest)이고,

상기 컴팩트한 분류자를 상기 캡처된 이미지에 적용하는 단계는 상기 캡처된 이미지가 상기 랜덤 결정 트리 중 특정한 하나의 트리와 연관된 특정 클래스에 매칭되는 가능성(likelihood)을 획득하기 위해 제3 세트의 특정 기술자의 유사도 스코어에 기초하여 각 랜덤 결정 트리를 탐색하는(traversing) 단계를 포함하는

컴퓨터 구현 방법.

청구항 14

제13항에 있어서,

상기 가능성은 확률 분포를 포함하고,

상기 캡처된 이미지와 연관될 가능성이 있는 클래스를 결정하기 위해 상기 랜덤 결정 트리 모두에 걸쳐 각 클래스에 대한 각 확률 분포를 평균화하는 단계를 더 포함하는 컴퓨터 구현 방법.

청구항 15

제13항에 있어서,
랜덤 결정 트리의 각 노드는 임계치와 연관되고,
각 랜덤 결정 트리를 탐색하는 단계는 상기 제3 세트 내의 특정 기술자의 유사도 스코어를 랜덤 결정 트리 내의 각 노드와 연관된 임계치와 비교하는 단계를 포함하는
컴퓨터 구현 방법.

청구항 16

제11항에 있어서,
각 특정 기술자는 BRIEF 특징 기술자인
컴퓨터 구현 방법.

청구항 17

제11항에 있어서,
유사도 스코어에 기초하여 제3 세트의 특정 기술자를 식별하는 단계를 더 포함하며,
상기 유사도 스코어에 기초하여 제3 세트의 특정 기술자를 식별하는 단계는
특정 기술자 쌍에 대한 민 해시 값을 결정하는 단계- 각 쌍은 상기 제1 세트로부터의 하나의 특정 기술자와 상기 제2 세트로부터의 제2 특정 기술자를 포함함 -와,
민 해시 값이 유사한 쌍에 대한 유사도 스코어를 연산하는 단계를 더 포함하는
컴퓨터 구현 방법.

청구항 18

제11항에 있어서,
상기 유사도 스코어는 해밍 거리에 기초하는
컴퓨터 구현 방법.

청구항 19

제11항에 있어서,
상기 컴팩트한 분류자 및 상기 특정 기술자는 오프라인 훈련 단계(offline training phase)를 통해 생성되는
컴퓨터 구현 방법.

청구항 20

제11항에 있어서,
상기 방법은 이동 장치 상에서 구현되는
컴퓨터 구현 방법.

발명의 설명

기술 분야

[0001] 본 발명은 위치 인식을 통해 이미지를 분류하는 기술과 관련된다.

배경 기술

[0002] 카메라를 가진 모바일 장치의 확산은 모바일 장치의 카메라로부터 찍혀진 이미지를 인식할 수 있는 위치 인식 애플리케이션에 대한 필요를 증가시키고 있다. 모바일 장치가 제한적인 저장 용량 및 연산 자원을 가지기 때문에, 시각적 위치 인식의 작업은 종종 원격적으로 실행된다. 모바일 장치는 모바일 장치로부터 찍힌 이미지를 서버에 업로드할 수 있다. 서버는 해당 이미지를 서버 상에 저장된 유사한 이미지와 매칭한다. 저장된 이미지들은 이미지의 위치에 관한 지리적 데이터로 주석이 달려진다. 서버는 모바일 장치에 이미지의 위치를 다운로드한다.

발명의 내용

해결하려는 과제

[0003] 그러나, 모바일 장치는 서버와 통신하기 위해 제한된 대역폭을 가질 수 있다. 모바일 장치 및 서버 사이에서의 이미지 및 데이터의 전송에 관한 지연 시간(latency)은 모바일 장치가 실시간으로 위치 인식을 수행하는데 있어서 장애가 될 수 있다.

과제의 해결 수단

[0004] 본 발명에 따른 일 실시예는 분류자(classifier)를 획득하기 위해 지구 물리학적 위치를 사용하는 단계 - 상기 분류자는 하나 이상의 클래스(class)와 연관되고, 각 클래스는 이미지를 설명하기 위해 사용됨 - ; 상기 지구 물리학적 위치와 연관된 이미지를 캡처하는 단계 - 상기 이미지는 인지되지 않은 클래스를 가짐 - ; 및 상기 이미지에 대한 클래스를 결정하기 위해 상기 이미지에 상기 분류자를 적용하는 단계를 포함하는 컴퓨터 구현된 방법을 제공한다.

발명의 효과

[0005] 본 발명에 따른 실시예들에 따른 방법은 컴팩트한 분류자를 결과로서 만들어내어 분류자가 모바일 장치로 신속하게 다운로드될 수 있으며, 분류 단계를 위한 연산을 모바일 장치 상에서 실시간으로 수행 가능하게 할 수 있어 모바일 장치가 캡처한 이미지를 효과적으로 분류할 수 있도록 한다.

도면의 간단한 설명

[0006] 도 1은 위치 기반 인식을 위한 예시적인 시스템의 블록 다이어그램을 도시한다.
도 2는 훈련 단계에 대한 예시적인 방법을 설명하는 흐름도이다.
도 3은 실시간 질의 단계에 대한 예시적인 방법을 설명하는 흐름도이다.

도 4는 분류자를 훈련하기 위한 예시적인 방법을 설명하는 흐름도이다.

도 5는 특징 추출을 위한 예시적인 방법을 설명하는 블록 다이어그램이다.

도 6은 밀집 매칭(dense matching)을 위한 예시적인 방법을 설명하는 블록 다이어그램이다.

도 7은 밀집 매칭을 위한 다른 예시적인 방법을 설명하는 블록 다이어그램이다.

도 8은 모바일 장치로부터 위치 기반 인식을 수행하기 위한 예시적인 방법을 설명하는 흐름도이다.

도 9는 분류자를 훈련하기 위한 예시적인 방법을 설명하는 블록 다이어그램이다.

도 10은 운영 환경을 설명하는 블록 다이어그램이다.

도 11은 운영 환경 내의 모바일 장치의 예시적인 컴포넌트들을 설명하는 블록 다이어그램이다.

도 12는 운영 환경 내의 서버의 예시적인 컴포넌트들을 설명하는 블록 다이어그램이다.

발명을 실시하기 위한 구체적인 내용

[0007]

본 요약은 이후 상세한 설명에서 더 자세히 설명될 선택적 개념들을 단순화된 형태로 소개하고자 제공된다. 본 요약은 청구 대상 발명의 핵심적이거나 필수적인 특징을 식별하고자 하는 의도를 가지지 않으며, 청구 대상 발명의 범위를 제한하고자 하는 의도를 가지지도 않는다.

[0008]

모바일 장치는 서버로부터의 최소한의 어시스트로 실시간 위치 인식을 수행한다. 모바일 장치의 지구 물리학(geophysical)적 위치는 모바일 장치로부터 서버로 업로드된다. 서버는 모바일 장치의 지구 물리학적 위치에 기초하여 컴팩트한 분류자(compact classifier) 및 특징 기술자(feature descriptors)를 모바일 장치에 제공한다. 이후, 모바일 장치는 이미지를 캡처하고 캡처된 이미지는 컴팩트한 분류자 및 특징 기술자를 사용하여 분류될 수 있다. 이러한 분류의 결과물은 해당 이미지가 특정 클래스(class) 또는 랜드마크(landmark)일 가능성에 대한 추정이다.

[0009]

분류자 및 특징 기술자는 오프라인 훈련 단계를 통해 생성될 수 있으며 연산 프로세싱을 최소화하기 위해 구성된다. 일 실시예에서, 분류자는 동일한 지구 물리학적 위치 및 클래스로부터의 저장된 이미지들을 사용하여 훈련된 랜덤 결정 트리군(random decision forest)이다. 랜덤 결정 트리군을 훈련하기 위한 방법은 이미지를 내의 가장 구별되는 특징들을 자동 선택한다. 이러한 방법은 컴팩트한 분류자를 결과로서 만들어내어 분류자가 모바일 장치로 신속하게 다운로드될 수 있으며, 분류 단계를 위한 연산을 모바일 장치 상에서 실시간으로 수행 가능하게 할 수 있다.

[0010]

전술한 것 및 그 이외의 특징들 및 이점들은 후술되는 상세한 설명을 판독하고 연관된 도면들을 리뷰함에 따라 명확해질 것이다. 후술되는 일반적인 기술 및 상세한 설명은 설명을 위한 것일 뿐이며 청구된 발명을 제한하는 것이 아님이 이해되어야 할 것이다.

[0011]

다양한 실시예들은 실시간 위치 기반 인식을 수행하는 기술에 대한 것이다. 위치 인식은 많은 양의 저장소 및 연산 시간을 필요로 하는 연산적으로 부담스러운 프로세스이다. 실시간 위치 인식은 연산적으로 효율적인 기술이 사용될 때 모바일 장치에 대해 실용적이다. 본 명세서에서 설명하는 기술은 인식의 정확도를 감소시키지 않으면서도 연산적인 부담을 줄이기 위한 압축된 데이터 구조 및 기술의 사용에 기반한다.

[0012]

분류자는 모바일 장치의 카메라에 의해 캡처된 이미지를 특정 지구 물리학적 위치 내의 미리 결정된 클래스들 또는 랜드마크의 세트 중 하나로 분류함으로써 이미지를 인식하기 위해 사용될 수 있다. 일 실시예에서, 분류자는 몇 개의 랜덤 결정 트리들을 가지는 랜덤 결정 트리군이 될 수 있다. 랜덤 결정 트리는 캡처된 이미지의 특징을 동일한 지구 물리학적 위치 및 클래스를 가지는 훈련 이미지와 비교하는데 사용될 수 있다. 클래스는 랜드마크, 특정 빌딩 또는 구조와 같이 이미지를 기술하는 식별자이다. 분류는 일련의 비교를 통해 랜덤 결정 트리를 루트 노드(root node)로부터 잎 노드(leaf node)까지 순회(traverse)하는 과정을 포함한다. 각각의 비교는 특정한 특징이 캡처된 이미지 내에 존재하는지 여부를 평가하는 이진 테스트(binary test)를 나타낸다. 이러한 비교는, 먼저 해밍 거리(Hamming distance)와 같은 측정법을 사용하여 캡처된 이미지 내에서 가장 유사한 특징을 탐색하고, 이러한 특징을 임계치와 비교하는 것에 이루어진다. 임계치는 훈련 이미지들로부터 학습된 매칭 가능성을 계량화한 것이다. 트리에 대한 순회(traversal)는 캡처된 이미지가 특정 클래스 또는 랜드마크일 가능성을 나타내는 잎 노드에서 종료된다. 분류는 랜덤 결정 트리군 내 랜덤 결정 트리 각각에 의해 독립적으로 수행된다. 각각의 랜덤 결정 트리를 순회하는 것에 의해 생성된 모든 가능성의 평균은 이미지에 가장

근접한 클래스를 추정하기 위해 사용될 수 있다.

[0013] 일 실시예에 따르면, 서버는 모바일 장치가 스스로 실시간 위치 인식을 함에 있어서 도움을 주는데 사용될 수 있다. 서버는 오프라인 훈련 단계 동안에 지구 물리학적 위치 및 클래스 각각에 대하여 랜덤 결정 트리군을 생성한다. 랜덤 결정 트리군 내의 랜덤 결정 트리는 동일한 지구 물리학적 위치 및 클래스로부터 저장된 이미지들을 사용하여 훈련된다. 랜덤 결정 트리들은 이미지의 가장 핵심적인 특징을 포함하는 것으로 분석된 특정 기술자들의 세트를 사용하여 생성된다.

[0014] 오프라인 훈련 단계에서, 동일한 지구 물리학적 위치와 연관된 훈련 이미지들은 이미지 내의 가장 핵심적인 특징을 결정하기 위해 분석된다. 훈련 이미지들로부터 특징들을 추출하고 분석하여 구별적인 특성들을 가지는 서브세트를 선택하기 위해 특정 추출 프로세스가 수행된다. 이후, 서브세트 내의 각각의 특징은 특정 기술자로 알려진 보다 컴팩트한 표현으로 변환된다. 밀집 매칭 프로세스가 실행되어 어느 특정 기술자가 이미지 내에서 보다 핵심적인 특징을 나타내는지를 결정한다.

[0015] 핵심적인 특징에 대한 검색은 이미지 각각의 특정 기술자 각각에 대해 해밍 거리와 같이 유사도 스코어를 계산하는 것에 의해 수행된다. 이러한 검색에 관련된 해밍 거리 연산의 수를 최소화하기 위해, 민 해싱 프로세스 (min hashing process)가 사용되어 어느 특정 기술자가 유사한지 결정하고, 이에 해밍 거리의 연산을 필요로 하게 한다. 해밍 거리 연산은 상이한 특정 기술자들에 대해서는 회피되는데, 이는 이들의 유사도 스코어가 낮기 때문이다. 각 훈련 이미지에 대응하는 특정 기술자를 포함하는 테이블이 사용되어 랜덤 결정 트리군을 훈련한다. 특정 기술자의 각 구성 요소는 각 특정 기술자들 사이의 유사도 스코어 및 전체 이미지 내에서 발견된 가장 유사한 기술자를 연산하는 것에 의해 획득된다. 랜덤 결정 트리군을 훈련하기 위한 방법은 다양한 클래스의 이미지를 내에 존재하는 가장 구별되는 특징들을 자동 선택한다. 그 결과, 모바일 장치에 신속히 다운로드될 수 있는 컴팩트한 분류자가 생성되어, 모바일 장치 상에 실시간으로 분류 단계를 위한 연산이 실행될 수 있도록 한다.

[0016] 일 실시예에서, 모바일 장치는 모바일 장치의 지구 물리학적 위치만을 사용하는 질의로부터 랜덤 결정 트리군 및 특정 기술자들을 서버로부터 획득한다. 일 실시예에서, 지구 물리학적 위치는 모바일 장치의 GPS 좌표를 알 수 있다. 이 경우, 서버와 모바일 장치에서 데이터를 송수신하는데 있어서의 지연이 최소화된다. 이후, 모바일 장치는 이후 랜덤 결정 트리군을 사용하여 분류될 이미지를 캡처한다. 분류 결과는 해당 이미지가 특정 클래스에 해당한다는 가능성의 추정이다. 위치 기반 인식에 대한 실시예들에 대해 이하에서 보다 자세히 설명하도록 한다.

[0017] 도 1을 참조하면, 위치 기반 인식을 위한 시스템(100)이 도시된다. 시스템(100)은 하나 이상의 모바일 장치(104)와 통신하는 서버(102)를 포함할 수 있다. 일 실시예에서, 서버(102) 및 모바일 장치(104)는 통신 네트워크(106)를 통해 통신 가능하게 연결될 수 있다. 통신 네트워크(106)는 유선, 무선 통신 매체 및 이들의 조합을 포함할 수 있는 다양한 유형의 통신 매체로 구성될 수 있다. 일례로, 모바일 장치(104)는 무선 통신 매체를 통해, 유선 통신 매체를 통해 서버(102)와 연결되어 있는 무선 액세스 포인트로 통신할 수 있다. 대안적으로, 모바일 장치(104)는 서버(102)로 직접 무선 통신 매체를 통해 통신할 수도 있다. 무선 통신 매체는 임의의 무선 기술, 예를 들어 무선 표준의 IEEE 802 패밀리, WI-FI, WiMax, Bluetooth™, 또는 이들의 조합을 사용하여 동작 가능하며, 이에 대해 한정하지는 않는다.

[0018] 서버(102)는 컴퓨터, 개인용 컴퓨터, 데스크탑 컴퓨터, 랩톱 컴퓨터, 노트북 컴퓨터, 휴대용 컴퓨터, 서버 어레이 또는 서버 팝, 미니 컴퓨터, 메인프레임 컴퓨터, 슈퍼 컴퓨터, 분산 컴퓨팅 시스템, 멀티 프로세서 시스템, 또는 이들의 조합을 포함할 수 있으나 이에 대해 한정하지는 않는다. 모바일 장치(104)는 무선 통신 매체를 통해 통신할 수 있는 능력을 가지는 임의의 휴대용 전자 장치, 예를 들어, PDA(personal digital assistant), 모바일 컴퓨팅 장치, 스마트폰, 태블릿 개인 컴퓨터, 휴대폰, 메시지 장치 또는 이들의 조합이 될 수 있으나, 이에 대해 한정하지는 않는다.

[0019] 서버(102)는 이미지 데이터베이스(108) 및 분류자 데이터베이스(110)를 포함할 수 있다. 이미지 데이터베이스(108) 내에 저장된 이미지들은 지구 물리학적 위치 및 클래스 식별자를 포함하는 지오 태그(geo-tag)를 사용하여 주석이 달릴 수 있다. 일 실시예에서, 지구 물리학적 위치는 지구 상의 위치에 대한 지리적 위도 및 경도를 사용하는 지리학적 좌표 시스템에 기초한다. 지구 물리학적 위치는 두 개의 숫자로 나타날 수 있다. 하나의 숫자는 지리적 위도이고 두 번째 숫자는 지리적 경도이다. 클래스는 이미지 내에서 보이는 랜드마크나 위치를 식별한다. 예를 들어, 클래스는 도로 코너, 빌딩, 기념물과 같은 도시적 랜드마크의 일종 또는 특정 위치일 수 있다. 클래스 식별자는 특정 클래스를 나타내는 정수일 수 있다. 이미지들은 대응하는 이미지 내에서 어떠한

랜드마크 또는 빌딩이 관측되는지를 나타내는 클래스 식별자 및 연관된 지구 물리학적 위치 함께 이미지 데이터베이스 내에 저장될 수 있다.

[0020] 분류자 데이터베이스(110)는 분류자들을 포함한다. 분류자는 동일한 지구 물리학적 위치 및 클래스를 갖는 이미지의 세트를 사용하여 훈련된 학습 함수(learned function)이다. 일 실시예에서, 분류자들은 랜덤 결정 트리군일 수 있다. 랜덤 결정 트리군은 랜덤 결정 트리들의 세트를 포함하고, 각각의 랜덤 결정 트리는 클래스 또는 랜드마크의 세트와 관련된 가능성 분포를 결정하기 위해 사용된다.

[0021] 훈련 단계(111)는 분류자들을 연산하기 위한 몇 개의 단계를 포함할 수 있다. 훈련 단계(111)는 이미지 데이터베이스(108) 내에 저장된 랜드마크에 대응하는 클래스 및 지구 물리학적 위치를 갖는 이미지들의 세트를 사용한다. 각 이미지는 이미지 내에 존재할 수 있는 시각적 왜곡을 제거하기 위해 이미지 교정 모듈(image rectification module)(112)을 통해 교정된다. 이미지 교정은 상이한 시각(perspective)에서 찍혀진 사진들을 시각적 왜곡이 제거되거나 최소화되는 방향으로 변환하기 위한 기하학적 변환이다.

[0022] 반복적이고 상이한 랜드마크들에 대해 고유한 핵심적 특징들은 특징 추출 모듈(114)에 의해 식별된다. 특징 추출 모듈(114)은 이러한 핵심적 특징들의 형상을 인코딩하여 대응하는 이미지 패치들 내의 페어 방식 픽셀 비교(pair wise pixel comparisons)로부터 형성된 이진 스트링(binary string)으로 인코딩한다. 이들이 특징 기술자를 지칭된다. 훈련 밀집 매칭 모듈(training dense matching module)(116)은 특징 기술자의 세트 각각에 대해 가장 유사한 이미지 패치를 탐색한다. 이미지 패치는 디지털화된 이미지 내의 핵심적인 특징의 위치 주변의 사각형 영역 내의 픽셀에 대응된다. 이미지 각각에 대하여, 특정 특징 기술자에 가장 유사한 이미지 패치가 발견되고 유사도 스코어가 연산된다. 모든 특징 기술자들에 대응하는 유사도 스코어가 테이블 내에 저장된다. 테이블 내에는 각 이미지에 대한 하나의 열(column)이 있다. 밀집 매칭은 전체 이미지 내에서 가장 유사한 특징 기술자에 대한 밀집 검색(dense search)을 가속하기 위해 사용되는 민 해시(min-hash) 프로세스를 사용하여 가속된다.

[0023] 서버(102)는 모바일 장치의 대략적인 지구 물리학적 위치에 대응하는 분류자를 모바일 장치(104)에 제공함으로써 모바일 장치(104)의 실시간 위치 인식을 보조한다. 질의 단계 동안, 위치 보조 모듈(120)은 모바일 장치(104)의 지구 물리학적 위치를 포함하는 질의들을 모바일 장치(104)로부터 수신한다(블록 122). 이에 응답하여, 서버는 모바일 장치의 대략적인 지구 물리학적 위치 및 특징 기술자의 세트와 연관된 분류자와 같은 데이터를 모바일 장치(104)로 전송한다(블록 124).

[0024] 모바일 장치(104)는 자신의 지구 물리학적 위치를 결정하기 위한 매커니즘을 포함할 수 있다. 모바일 장치(104)는 모바일 장치의 지구 물리학적 위치를 추적하는 온보드 위성 수신기(onboard satellite receiver)와 상호 작용하는 위치 모듈(127)을 포함할 수 있다. 일 실시예에서, 온보드 위성 수신기는 모바일 장치의 GPS 좌표를 나타내는 GPS 신호들을 수신하도록 구성될 수 있다. 대안적으로, 위치 모듈(127)은 셀룰러 네트워크와 통신하는 무선 트랜시버와 상호 작용할 수 있다. 위치 모듈(127)은 셀룰러 네트워크를 사용하여 모바일 장치의 GPS 좌표를 결정할 수 있다.

[0025] 또한, 모바일 장치(104)는 서버(102)와 상호 작용하는 서버 인터페이스 모듈(126)을 포함할 수 있다. 서버 인터페이스 모듈(126)은 모바일 장치의 GPS 좌표와 같은 지구 물리학적 위치를 포함하는 질의를 서버에 업로드한다(블록 128). 이후, 서버 인터페이스 모듈(126)은 분류자 및 특징 기술자와 같은 데이터를 서버(102)로부터 수신한다(블록 130). 사진 캡처 모듈(132)이 디지털화된 이미지로서 저장된 객체 또는 랜드마크의 이미지를 캡처하기 위해 사용될 수 있다. 이후, 이미지는 시각적 왜곡을 제거하기 위해 이미지 교정 모듈(134)에 의해 교정된다.

[0026] 질의 단계 동안, 실시간 밀집 매칭 모듈(136)은 다운로드된 특징 기술자들 각각에 대응하는 가장 유사한 이미지 패치들에 대해 캡처된 이미지 내에서 밀집 검색(densely searching)을 하는 것에 의해 캡처된 이미지에 대해 고-차원(high-dimensional) 특징 벡터를 구성한다. 특징 벡터의 각 구성 요소는 캡처된 이미지 내에서 발견된 가장 유사한 이미지 패치 또는 특징 기술자와 다운로드된 특징 기술자 사이의 해밍 거리를 연산하는 것에 의해 획득된다. 민 해싱 기술이 이러한 검색을 가속하기 위해 사용된다. 유사한 민 해시 값을 갖는 특징 기술자들의 쌍은 보다 유사할 가능성이 높다. 따라서 특정 특징 기술에 가장 유사한 기술자를 검색하는 경우, 첨조 기술자의 민 해시 값과 상이한 민 해시 값을 갖는 기술자들은 해밍 거리 연산을 실행하지 않고 일찌감치 거부된다. 이러한 과정은 불필요한 해밍 거리 연산을 제거하고 전체 연산의 속도를 향상시킨다.

[0027] 분류 모듈(138)은 각 랜덤 결정 트리를 순회하기 위해 고-차원 특징 벡터를 사용하며 이는 주어진 지구 물리학

적 위치와 연관된 각 클래스에 대한 가능성 추정을 결과물로 한다. 분류 모듈(138)의 결과물은 디스플레이(140) 상에 도시되고 이미지 내의 클래스의 가능성에 의해 분류(sorted)된 랜드마크 또는 위치들의 리스트의 형태를 가질 수 있다. 대안적으로, 결과물은 각 클래스에 대한 가능성 추정을 시각적으로 나타내는 히스토그램의 형태를 가질 수 있다. 또한, 다른 시각적 제시 방법이 분류 결과물을 나타내기 위해 사용될 수 있음이 이해되어야 할 것이다.

[0028] 이하에서는, 다양한 예시적인 방법들을 참조하여 위치 기반 인식 시스템(100)의 실시예들의 동작에 대한 보다 상세한 설명을 하도록 한다. 달리 표현되지 않는 이상, 제시되는 방법들은 제시된 순서나 임의의 특정 순서로 실행될 필요가 없다는 것이 이해되어야 할 것이다. 또한, 방법들과 관련하여 설명된 다양한 동작들은 직렬 또는 병렬 방식으로 실행되거나 직렬 및 병렬 동작의 조합으로서 실행될 수 있다. 방법들은 주어진 세트의 설계 및 성능 제한에서 요구되는 바에 따라 기술된 실시예들 또는 대안적인 실시예들의 하나 이상의 하드웨어 요소 및/또는 소프트웨어 요소를 사용하여 구현될 수 있다. 예를 들어, 방법들은 논리 장치(예: 일반적 목적 또는 특수 목적용 컴퓨터)에 의해 실행되기 위한 논리(예: 컴퓨터 프로그램 명령어)로서 구현될 수 있다.

[0029] 도 2를 참조하면, 서버(102)에 의해 실행되는 훈련이 도시된다(블록 152). 일 실시예에서, 훈련 단계는 질의 프로세싱으로부터 오프라인으로 실행되어 분류자들을 질의 단계에서 용이하게 사용 가능하게 한다. 그러나, 훈련은 의도되는 구현 방법에 의해 요구되는 바에 따라 위치 인식 프로세스의 다른 시점에서 실행될 수도 있다.

[0030] 도 3을 참조하면, 서버(102)는 모바일 장치(104)로부터의 질의가 대기 중인지 여부를 체크할 수 있다(블록 156). 모바일 장치(104)로부터 질의가 수신되면(블록 156-yes), 서버(102)는 분류자 데이터베이스(110)로부터 모바일 장치(104)의 대략적인 지구 물리학적 위치에 대응하는 분류자 및 특징 기술자들을 검색한다(블록 158). 아니면(블록 156-no), 서버(102)는 프로세싱을 종료시킬 것인지 여부를 체크(블록 160-yes)하고, 종료시키지 않는다면(블록 160-no), 서버(102)는 모바일 장치(105)로부터의 질의를 계속적으로 체크한다.

[0031] 도 4는 훈련 단계(152) 동안에 실행되는 동작들을 나타내는 흐름도이다. 먼저, 이미지들이 수집되어 이미지 데이터베이스(108)에 저장된다. 각 이미지들은 전술한 것과 같이 지리적 위치 및 클래스로 주석이 달린다. 분류자들은 이미지 데이터베이스(108)로부터 획득된 특정 지리적 위치 및 클래스에 대응하는 이미지들에서 훈련된다(블록 170).

[0032] 이후, 이미지 교정 모듈(112)은 이미지의 수직 및 수평 라인들이 교정된 이미지 각각에서의 수직 및 수평 라인들로 보이는 것을 보증하기 위해 이미지 내의 시각적 왜곡을 제거하는 기하학적 변환인 이미지 교정을 수행한다(블록 172). 이 경우, 밀집 매칭은 이미지의 모든 위치 및 스케일(scale) 범위에 걸쳐서 수행되면 된다. 캡처된 이미지들에서 직선을 양호하게 감지하고 소실점을 추정하는 것에 기초하여 도시 이미지를 내에서 캡처된 이미지들에 대한 이미지 교정을 수행하기 위한 몇 개의 주지 기술들이 존재하며, 이들 중 임의의 기술이 사용될 수 있다. 또한, 모바일 장치 내에 존재할 수 있는 모션 세서로부터 모바일 장치의 3 차원 로테이션으로부터 추정을 획득하는 다른 기술들도 사용될 수 있다.

[0033] 다음으로, 특징 추출이 수행된다(블록 174). 특징 추출 모듈(114)은 이미지 내에서 핵심적이고 고유한 형상으로 보여지고 반복적인 특징들의 세트, 즉, 동일한 경관에 대한 상이한 이미지들에서 발생하는 특징들을 식별한다(블록 176). 일 실시예에서, 특징들의 후보 세트가 식별된다. 특징들의 후보 세트는 방울 검출(blob detection), 에지 검출, 코너 검출, 능선 검출(ridge detection), 및 스케일 불변 특징 변환을 위한 스케일-불변 관련 포인트 검출(scale invariant feature transformation for scale-invariant interest point detection), 및 이들의 조합과 같은 몇 개의 주지 기술 중 임의의 하나를 사용하여 식별될 수 있다. 이러한 특징들은 각 이미지 내의 특징들의 위치, 스케일 및 지향에 의해 제시된다. 스케일 불변 특징 변환(Scale Invariant Feature Transform, SIFT) 기술자와 같은 특징 기술자는 각 이미지 내의 각 후보 특징에 대해 연산된다. 유사하다고 보여지는 특징들의 쌍은 링크되어 특징 기술자들의 서브세트를 형성한다. 최종적으로, 이러한 서브세트 내의 다른 모든 특징 기술자들에 비해 최소한의 총합 유кли디안 거리(Euclidean distance)를 가지는 특징 기술자를 가지는 특징을 선택하는 것에 의해, 후보 특징은 서브세트로부터 선택된다. 이와 같이 각 클래스에 대응하는 후보 특징들의 세트가 이후의 사용을 위해 선택된다.

[0034] 다음으로, 특징 기술자는 선택된 특징 각각에 대해 생성될 수 있다(블록 178). 각 특징의 형상은 특징 기술자로서 인코딩된다. 일 실시예에서, 각 특징은 BRIEF(Binary Robust Independent Elementary Feature) 기술자에 의해 제시될 수 있다. 이 기술자는 이진 벡터로서, 각 특징에 대응하는 이미지 패치를 32×32 픽셀 패치로 리샘플링(resampling)하는 것에 의해 연산될 수 있다. 다음으로, 이러한 32×32 패치 내로부터 랜덤 픽셀들의 쌍 k개가 샘플된다. 개별적인 픽셀 위치를 랜덤으로 샘플링하기 위해 특징의 중앙에 중앙화된 가우시안 분포가 사

용된다. 일 실시예에서, $k=192$ 로서 이는 특징 기술자를 192-비트의 길이를 가지도록 한다. 특징 기술자의 각 비트는 특징 내에서 랜덤하게 선택된 두 픽셀들 사이의 강도 레벨 사이의 비교를 나타낸다. 만약 첫번째 픽셀이 두번째 픽셀보다 높은 강도를 가지면 비트는 "1"로 설정되고 두번째 픽셀이 첫번째 픽셀보다 높은 강도를 가지면 비트는 "0"으로 설정된다. k 번째 비트는 픽셀의 k 번째 쌍 사이의 비교 결과를 나타낸다.

[0035] 도 5를 참조하면, 이미지(180)가 도시된다. 특징 추출 모듈(114)은 이미지(180) 내의 작은 박스들로 나타나는 몇 개의 이미지 패치 또는 특징 후보들을 식별한다. 특징(182)과 같은 특징 후보들의 서브세트가 선택된다. 특징(182) 내의 픽셀들의 강도 레벨의 쌍 단위 비교는 블록(184)에 도시된다. 다음으로, 특징 추출 모듈(114)은 전술한 블록(184)에 도시된 비교들의 결과를 나타내는 192 비트 BRIEF 특징 기술자(186)을 생성한다. 비록 도 5는 하나의 특징에 대한 특징 추출 모듈(114)의 프로세스를 도시하지만, 이러한 프로세스는 훈련 단계에서 사용된 각 이미지의 각 특징에 대해 실행된다. 따라서, 단일 클래스에 대해 수백 개의 특징 기술자들이 생성될 수 있다.

[0036] 다시 도 4를 참조하면, 훈련 밀집 매칭 모듈(116)이 특징 추출 모듈(114)에 의해 선택된 특징 기술자들에 대해 밀집 매칭을 실행한다(블록 180). 밀집 매칭에서, 이미지의 세트 중 모든 위치 및 스케일에 걸쳐 가장 유사한 특징들을 결정하기 위해 각 특징은 모든 이미지의 각 위치에 걸쳐서 매칭된다. 유사도는 특징 기술자와 이미지 내의 이미지 패치의 특징 기술자 사이의 거리에 기초한다. 몇 실시예들에서, 최소 거리는 해밍 거리가 될 수 있다. 이미지의 차원과 스케일이 크기 때문에, 밀집 매칭은 연산 시간 및 저장 공간 모두에 연산적으로 고비용인 프로세스이다. 신속한 인식을 위해서, 밀집 매칭을 가속하기 위해 민-해시 프로세스(즉, 민 단위 독립 순열(min-wise independent permutations))이 사용될 수 있다. 민 해싱은 이미지 내의 특징 기술자에 대응하는 가장 유사한 이미지 패치를 검색하는 동안에 실행되는 해밍 거리 연산들의 수를 감소시키기 위해 사용된다.

[0037] 민 해시 함수는 두 개의 이진 특징 기술자가 유사한지를 결정하는 세트들 사이의 자카드 유사도 측정(Jaccard similarity measure)에 기초한다. 만약 두 개의 이진 특징 기술자들의 비트의 과반이 매칭되면, 자카드 유사도는 1에 근접하고 특징 기술자들은 유사한 민 해시 값을 가질 가능성이 크다. 민 해시 함수는 이진 특징 기술자에 포함된 비트들의 랜덤 순열(random permutation)로 구현된다. 민 해시 값은 해당 특징 기술자의 랜덤 순열 내의 첫번째 긍정 비트('1')의 인덱스이다. 스케치(sketch)는 독립적인 민 해시 값들의 순서화된 n-튜플(n-tuple)이며, n은 특징 기술자에 대한 민 해시 함수의 애플리케이션의 수이다. 기술자들의 쌍은 다섯 개의 스케치 중 적어도 두 개가 동일하면 유사한 것으로 추정된다.

[0038] 각 특징 기술자에 대해서, 이미지에서 발견된 가장 유사한 기술자에 대한 유사도 스코어는 밀집 매칭 테이블 내에 저장된다(블록 186). 일 실시예에서, 유사도 스코어는 해밍 거리이다. 이미지 y 내의 특징 기술자 x에 대한 유사도 스코어는, 기술자 x와 이에 대한 이미지 y 내의 가장 근접한 매칭된 특징 기술자 z 사이의 해밍 거리를 나타낸다.

[0039] 도 6은 민 해시 프로세스를 도시한다. 도 6을 참조하면, 다양한 특징 기술자들의 세트로부터 두 개의 특징 기술자 188a 및 188b가 도시된다. 각 특징 기술자(188a, 188b)는 다섯 개의 연관된 스케치를 생성하도록 민 해싱되었다. 설명의 목적을 위해 도 6에 각 특징 기술자 190a, 190b에 대해서 세 개의 스케치가 도시되었다. 각 스케치(192a-192e)는 다섯 개의 민 해시 값들을 나타내며 각 민 해시 값은 BRIEF 특징 기술자 내의 비트들의 랜덤 순열에서 1(one)로 설정된 BRIEF 특징 기술자 내의 첫번째 비트의 인덱스를 나타낸다.

[0040] 모든 스케치들을 포함하는 세트(196)가 형성된다. 스케치들(198a-f)의 쌍들 사이의 민 해시 값들은 비교된다. 만약 적어도 두 개의 민 해시 값들이 동일하면, 스케치들은 매칭되는 것으로 추정된다. 도 6에 도시된 것처럼, 스케치들(198b 및 198c)는 함께 원으로 표시된 동일한 민 해시 값들을 갖는다. 스케치들(198e 및 198f)도 함께 원으로 표시된 동일한 민 해시 값들을 갖는다. 스케치들(198b 및 198c)과 연관된 특징 기술자들은 매칭되는 것으로 추정되고 스케치들(198e 및 198f)과 연관된 특징 기술자들도 매칭되는 것으로 추정된다. 해밍 거리는 두 개의 특징 기술자를 다루는 함수이고 상이한 비트들의 수를 결정하는 것에 의해 둘 사이의 거리를 결정한다. 해밍 거리는 적어도 두 개의 스케치들이 동일한 모든 기술자들의 쌍마다 연산된다. 이미지 내의 가장 근접한 특징 기술자는 참조 특징 기술자로부터 가장 낮은 해밍 거리를 갖는 특징 기술자이다.

[0041] 도 7을 참조하면, 스케치들(198a-198f)을 모두 포함하는 세트(196)로부터 스케치들(198b, 198c) 및 스케치들(198e, 198f)을 매칭하는 것이 도시된다. 밀집 매칭 테이블(200)은 m 개의 행 및 n 개의 열을 가지는 $m \times n$ 테이블로서 도시되며, 각각의 행 204A - 204M은 특징 기술자를 나타내고, 각각의 열 202A - 202N은 이미지를 나타낸다. 테이블 내의 각 셀은 특정 행과 연관된 특정 기술자와 특정 열과 연관된 이미지에서 발견된 가장 유사한 특징 기술자 사이의 해밍 거리를 포함한다.

[0042] 도 4를 다시 참조하면, 분류자 훈련 모듈(118)은 분류자를 훈련한다(블록 188). 일 실시예에서, 분류자는 랜덤 결정 트리군으로서 구현될 수 있다. 랜덤 결정 트리군은 개별적으로 훈련된 독립적인 랜덤 결정 트리들의 컬렉션이다. 랜덤 결정 트리는 내부 노드(internal node) 및 잎 노드(leaf node)들을 가지는 이진 트리로서 구성된다. 각 내부 노드는 두 개의 특정 기술자들 사이의 해밍 거리를 임계치와 비교하는 특정 테스트이다. 예를 들어, 일 실시예에서, 특정 테스트는 다음과 같이 구성될 수 있다: $HD(x, s) < T_s$, 여기서 함수 $HD()$ 는 두 개의 특정 기술자들 사이의 해밍 거리이고, x, s 는 특정 기술자들이고, T_s 는 s 에 대한 해밍 거리의 임계치이다. 잎 노드는 클래스 또는 랜드마크들의 세트에 대해 정의된 클래스들의 확률 분포(probability distributions)를 저장한다.

[0043] 랜덤 결정 트리는 루트 노드로부터 시작하여 회귀적으로 빌딩된다. 특정 기술자들은 루트 노드를 왼쪽 및 오른쪽 자식 노드(child node)로 스플릿하기 위해 사용된다. 각 노드에서, 스플릿은 지니 불순도(gini impurity) 기준과 같은 특정 정보 이론적 기준에 기초한다. 각 노드의 특징을 스플릿하는 것은 선택되지 않은 남아 있는 특정 기술자들로부터 랜덤으로 선택된다. 랜덤 결정 트리는 현재 노드에서 스플릿할 특정 기술자가 더 이상 남아 있지 않거나 트리의 깊이가 미리 정해진 한계를 넘어가는 경우에 자라나는 것을 멈춘다. 미리 정해진 수의 랜덤 결정 트리들이 형성될 때까지, 특정 기술자들은 다른 랜덤 결정 트리들을 랜덤하게 훈련시키기 위해 다시 사용된다. 다음으로, 랜덤 결정 트리군은 모든 랜덤 결정 트리들의 총 컬렉션으로서 형성될 수 있다. 이후, 랜덤 결정 트리군은 랜덤 결정 트리군을 훈련하기 위해 사용된 특정 기술자들과 함께 분류자 데이터베이스(110) 내에 저장될 수 있다.

[0044] 일단 랜덤 결정 트리군이 형성되면, 각 랜덤 결정 트리는 독립적인 분류 결정을 획득하기 위해 순회(traverse)될 수 있다. 각 랜덤 결정 트리는 각 내부 노드에서의 고차원 특징 벡터를 사용하여 위에서 아래로 순회되며, 고차원 특징 벡터의 각 구성 요소는 각 내부 노드에서의 특정 임계치와 비교된다. 이러한 순회가 잎 노드에 이르면, 잎 노드에 의해 나타나는 클래스들과 연관된 확률 분포가 저장(recorded)된다. 각 트리로부터의 각 클래스에 대한 확률은 평균화되고 최종 클래스 확률을 결정하기 위해 사용된다.

[0045] 위치 기반 인식에서의 모바일 장치(104)에 의해 실행되는 동작에 대해 다시 주목해보자. 도 8을 다시 참조하면, 모바일 장치(104)는 모바일 장치의 지리학적 위치를 포함하는 질의를 서버(102)에 업로드한다(블록 222). 일 실시예에서, 모바일 장치(104)는 직접 모바일 장치(104)의 지구 물리학적 위치를 획득 및 추적 가능한 위성 또는 무선 트랜시버를 구비할 수 있다. 위성 또는 무선 트랜시버는 GPS, GLONASS, Galileo 및 Beidou 위성을 사용하는 글로벌 네비게이션 위성 시스템(GNSS)과 같은 임의의 위성 위치 탐지 시스템으로부터 신호를 수신하도록 구성될 수 있다. 다른 실시예에서, 모바일 장치(104)는 셀룰러 네트워크로부터 라디오 주파수 신호를 수신 가능한 무선 트랜시버를 구비할 수 있으며, 무선 트랜시버는 셀룰러 네트워크로부터 간접적으로 지구 물리학적 위치를 획득하고 추적하도록 사용될 수 있다. 일 실시예에서, 지구 물리학적 위치는 모바일 장치의 지구 물리학적 위치와 연관된 GPS 경도 및 위도 좌표로서 구성될 수 있다.

[0046] 모바일 장치(104)는 자신의 대략적인 지구 물리학적 위치에 대응하는 분류자 및 특정 기술자를 수신한다(블록 224). 모바일 장치(104)는 사용자가 사진을 캡처할 수 있도록 하는 카메라를 구비할 수 있다(블록 226). 다음으로, 캡처된 이미지는 이미지 교정 모듈(134)에 의해 교정될 수 있다(블록 228). 다른 실시예에서, 모바일 장치가 모션 센서를 포함하는 경우, 모바일 장치의 3D 로테이션의 추정치를 모션 센서로부터 획득하는 기술이 이미지 교정을 수행하기 위해 사용될 수 있다.

[0047] 다운로드된 특정 기술자들 및 이들에 대한 캡처된 이미지 내의 가장 근접한 기술자들에 대한 유사도 스코어를 포함하는 고차원 특징 벡터를 생성하는 것에 의해, 다운로드된 특정 기술자들 각각에 대해 가장 근접하게 매칭되는 캡처된 이미지를 검색하기 위해 실시간 밀집 매칭이 실행된다(블록 230). 실시간 밀집 매칭 모듈(136)은 다운로드된 특정 기술자들 및 캡처된 사진 내의 이미지 패치를 나타내는 특정 기술자들에 대해 민 해싱 기술을 적용하여 다운로드된 특정 기술자들 중 어느 것이 캡처된 이미지 내에 보다 많은 유사 특징들을 가지는지 식별한다(블록 232). 민 해싱 기술은 도 6 및 도 7을 참조하여 전술한 것처럼, 특정 기술자들 각각에 대해 스케치를 생성하고, 유사한 스케치들을 갖는 특정 기술자들을 식별하고, 유사한 특정 기술자들에 대한 유사도 스코어를 생성하되, 최소의 유사도 스코어를 고차원 특징 벡터 내에 저장하는 방식으로 동작한다(블록 234).

[0048] 다음으로, 고차원 특징 벡터는 분류를 실행하기 위해 사용될 수 있다(블록 236). 분류 모듈(138)은 각 랜덤 결정 트리를 다양한 특정 기술자들의 유사도 스코어로서 구성된 고차원 특징 벡터에 기초하여 순회한다. 분류의 결과는 클래스들에 관한 확률이며 이러한 확률은 캡처된 이미지로부터 매칭되는 클래스들을 나타낸다. 이러한 결과는 사용자에게 확률의 내림 차순으로 랭킹된 클래스들의 리스트 형태로 디스플레이될 수 있다(블록 238).

[0049] 도 9를 참조하면, 서버로부터 획득된 특정 기술자들의 세트(242A-242N)가 도시된다. 특정 기술자들(242A-242N) 각각은 해밍 거리와 같은 유사도 스코어의 연산을 통해 캡처된 이미지 내의 이미지 패치들을 나타내는 특정 기술자들과 밀접 매칭될 수 있다. 유사도 스코어는 고차원 특정 벡터(244) 내에 기록될 수 있다. 각 결정 트리는 특정 기술자 및 결정 트리 내의 각 내부 노드에 포함된 임계치에 대응하는 유사도 스코어에 기초하여 순회된다. 랜덤 결정 트리에 대한 순회 결과는 일 노드에 의해 나타나는 다양한 클래스들에 대한 확률 분포이다. 클래스 확률 분포들은 모든 트리들에 대해 평균화되고, 이미지에 대한 가장 확률이 높은 클래스가 결정된다. 이하에서는 다시 예시적인 동작 실시예에 대해 설명하도록 한다.

[0050] 도 10을 다시 참조하면, 예시적인 동작 환경(300)에 대한 대략적인 블록 다이어그램이 도시된다. 동작 환경(300)은 하나 이상의 서버(들)(306)과 통신 프레임워크(304)를 통해 통신하는 하나 이상의 클라이언트(들)(302)을 포함할 수 있다. 일 실시예에서, 클라이언트(302)는 모바일 장치(102)로서 구현되고 서버(306)는 서버(102)로서 구현될 수 있다.

[0051] 클라이언트(302)는 하드웨어 장치, 소프트웨어 모듈 또는 이들의 조합을 포함할 수 있다. 이러한 하드웨어 장치들의 예는 컴퓨터(예: 서버, 개인용 컴퓨터, 랩톱 등), 휴대폰, PDA 또는 임의의 유형의 컴퓨터 장치 등을 포함할 수 있지만 이에 대해 한정하지는 않는다. 클라이언트(302)는 단일의 실행 경로, 복수의 동시적인 실행 경로(예: 스레드, 프로세스 등) 또는 임의의 다른 방식으로 실행되는 명령어들을 갖는 소프트웨어 모듈을 포함할 수 있다.

[0052] 서버(306)는 하드웨어 장치, 소프트웨어 모듈 또는 이들의 조합을 포함할 수 있다. 이러한 하드웨어 장치들의 예는 컴퓨터(예: 서버, 개인용 컴퓨터, 랩톱 등), 휴대폰, PDA 또는 임의의 유형의 컴퓨터 장치 등을 포함할 수 있지만 이에 대해 한정하지는 않는다. 서버(306)는 단일의 실행 경로, 복수의 동시적인 실행 경로(예: 스레드, 프로세스 등) 또는 임의의 다른 방식으로 실행되는 명령어들을 갖는 소프트웨어 모듈을 포함할 수 있다.

[0053] 통신 프레임워크(304)는 클라이언트(302)와 서버(304) 사이의 통신을 가능하게 한다. 일 실시예에서, 통신 프레임워크(304)는 통신 네트워크(106)로서 구현될 수 있다. 통신 프레임워크(304)는 임의의 통신 프로토콜을 사용하는 유선 또는 무선 네트워크들과 같은 임의의 유형의 통신 매체를 포함할 수 있다.

[0054] 도 11을 참조하면, 모바일 장치(102)는 프로세서(312), 메모리(314), 네트워크 인터페이스(316), 무선 트랜시버(318), 및 디스플레이(320)를 가질 수 있다. 프로세서(312)는 임의의 상업적으로 사용 가능한 프로세서일 수 있으며 듀얼 마이크로프로세서들 및 멀티 프로세서 아키텍처를 포함할 수 있다. 네트워크 인터페이스(316)는 모바일 장치(102) 및 통신 네트워크(106) 사이의 유선 또는 무선 통신을 가능하게 하여 모바일 장치(102) 및 서버(102) 사이에 통신 경로를 제공한다. 무선 트랜시버(318)는 위성 기반 라디오 신호, 셀룰러 기반 라디오 신호 및 이들의 조합이 될 수 있는 라디오 주파수 신호를 전송 및 수신할 수 있다. 디스플레이(320)는 모바일 장치의 사용자에게 시각적 표현을 디스플레이할 수 있는 임의의 유형의 스크린이 될 수 있다.

[0055] 메모리(314)는 프로세서 실행 가능한 명령어, 방법, 애플리케이션 및 데이터를 저장할 수 있는 임의의 컴퓨터 판독 가능한 저장 매체 또는 컴퓨터 판독 가능한 매체가 될 수 있다. 컴퓨터 판독 가능한 매체는 캐리어 웨이브를 통해 전송되는 변조된 데이터 신호와 같은 전파 신호와 관련되지 않는다. 이러한 매체는 임의의 유형의 메모리 장치(예: RAM, ROM 등), 자기 저장소, 휘발성 저장소, 비휘발성 저장소, 광학 저장소, DVD, CD, 플로피 드라이브, 디스크 드라이브, 플래쉬 메모리 등이 될 수 있다. 메모리(314)는 하나 이상의 외부 저장 장치 또는 원격에 위치한 저장 장치들을 포함할 수도 있다. 메모리(314)는 다음과 같은 명령어 및 데이터를 포함할 수 있다.

[0056] - 운영 체제(320);

[0057] - 위치 모듈(127);

[0058] - 사진 캡처 모듈(132);

[0059] - 이미지 교정 모듈(134);

[0060] - 실시간 밀집 매칭 모듈(136);

[0061] - 분류 모듈(138);

[0062] - 서버 인터페이스 모듈(126); 및

[0063] - 다양한 다른 애플리케이션 및 데이터(322)

[0064] 도 12를 참조하면, 서버(106)는 프로세서(330), 메모리(332) 및 네트워크 인터페이스(334)를 가질 수 있다. 프로세서(330)는 임의의 상업적으로 사용 가능한 프로세서가 될 수 있으며 듀얼 마이크로프로세서 및 멀티 프로세서 아키텍처를 포함할 수 있다. 네트워크 인터페이스(334)는 서버(106) 및 통신 네트워크(104) 사이에 유선 또는 무선 통신을 가능하게 하여 모바일 장치(102) 및 서버(106) 사이의 통신 경로를 제공한다.

[0065] 메모리(332)는 프로세서 실행 가능한 명령어, 방법, 애플리케이션 및 데이터를 저장할 수 있는 임의의 컴퓨터 판독 가능한 저장 매체 또는 컴퓨터 판독 가능한 매체가 될 수 있다. 컴퓨터 판독 가능한 매체는 캐리어 웨이브를 통해 전송되는 변조된 데이터 신호와 같은 전파 신호와 관련되지 않는다. 이러한 매체는 임의의 유형의 메모리 장치(예: RAM, ROM 등), 자기 저장소, 휘발성 저장소, 비휘발성 저장소, 광학 저장소, DVD, CD, 플로피 드라이브, 디스크 드라이브, 플래시 메모리 등이 될 수 있다. 메모리(332)는 하나 이상의 외부 저장 장치 또는 원격에 위치한 저장 장치들을 포함할 수도 있다. 메모리(332)는 다음과 같은 명령어 및 데이터를 포함할 수 있다.

[0066] - 운영 체제(336);

[0067] - 이미지 데이터베이스(108);

[0068] - 이미지 교정 모듈(112);

[0069] - 특징 추출 모듈(114);

[0070] - 훈련 밀집 매칭 모듈(116);

[0071] - 분류자 훈련 모듈(118);

[0072] - 위치 보조 모듈(120);

[0073] - 분류자 데이터베이스(110); 및

[0074] - 다른 애플리케이션 및 데이터(338)

[0075] 여기서 설명된 기술은 모바일 장치로 하여금 신속한 실시간 위치 인식을 원격 서버로부터의 도움을 받아 수행할 수 있도록 하는 이점을 제공한다. 이와 같은 기술은 모바일 장치 및 서버 사이에 전송되는 데이터의 양을 최소화하여 통신 매체에 기인하는 지연 시간을 향상시킨다. 전체 이미지 대신 지구 물리학적 위치 좌표가 서버에게 전송된다. 랜덤 결정 트리군을 훈련하는 접근 방법은 다양한 랜드마크들 사이에서의 구분을 위해 가장 구별되는 특징들을 자동 선택할 수 있게 한다. 이러한 접근 방법은 랜덤 결정 트리군이 보다 컴팩트한 데이터 구조가 되도록 하여 신속하게 다운로드될 수 있고 낮은 대역폭 통신 애플리케이션에 유리하도록 한다. 이러한 기술들은 모바일 장치가 실시간 위치 인식에 대한 요구를 충족할 수 있도록 한다.

[0076] 비록 본 출원 발명이 구조적 특징 및/또는 방법론적 동작들에 언어 특정적으로 기술되었지만, 첨부된 청구항들에 정의된 발명은 전술한 특정 특징들 또는 동작들에 한정되지 않아야 한다는 것이 이해되어야 한다. 오히려, 전술한 특정 특징들 또는 동작들은 청구항들의 구현의 예시적인 형태로서 개시된 것이다.

[0077] 예를 들어, 시스템(100)의 다양한 실시예들은 하드웨어 구성 요소, 소프트웨어 구성 요소 또는 이들의 조합을 사용하여 구현될 수 있다. 하드웨어 구성 요소의 예시는 장치들, 컴퓨트트들, 프로세서들, 마이크로프로세서들, 회로들, 회로 구성 요소들, 집적 회로들, 애플리케이션 특화된 집적 회로들, 프로그램 가능한 논리 장치들, 디지털 신호 프로세서들, 필드 프로그래밍 가능한 게이트 어레이들, 메모리 유닛들, 논리 게이트들 등을 포함할 수 있다. 소프트웨어 구성 요소의 예시는 소프트웨어 컴퓨트트들, 프로그램들, 애플리케이션들, 컴퓨터 프로그램들, 애플리케이션 프로그램들, 시스템 프로그램들, 기계 프로그램들, 운영 체제 소프트웨어, 미들웨어, 펌웨어, 소프트웨어 모듈, 루틴들, 서브루틴들, 함수들, 방법들, 프로시저들, 소프트웨어 인터페이스들, 애플리케이션 프로그램 인터페이스들, 명령어 세트들, 컴퓨팅 코드, 코드 세그먼트들, 및 이들의 임의의 조합을 포함할 수 있다. 실시예가 하드웨어 구성 요소 및/또는 소프트웨어 구성 요소를 사용하여 구현되는지 여부에 대한 결정은 주어진 구현 예에서 요구되는 요구 연산 속도, 파워 레벨, 대역폭, 연산 시간, 로드 벨런스, 메모리 리소스, 데이터 버스 속도 및 다른 디자인 또는 성능 제한들과 같은 임의의 수의 요소들에 의해 달라질 수 있다.

[0078] 몇 실시예들은 명령어들 또는 논리를 저장하기 위한 저장 매체를 포함할 수 있다. 저장 매체의 예시는 전자 데이터를 저장할 수 있는 컴퓨터 판독 가능한 저장 매체의 하나 이상의 유형, 예를 들어, 휘발성 메모리 또는 비휘발

성 메모리, 제거가능 또는 제거불가능 메모리, 삭제가능 또는 삭제불가능 메모리, 기입가능 또는 재기입가능 메모리 등을 포함할 수 있다. 논리의 예시는 프로그램, 프로시저, 모듈, 애플리케이션, 코드 세그먼트, 프로그램 스택, 미들웨어, 펌웨어, 방법, 루틴 등과 같은 다양한 소프트웨어 컴포넌트들을 포함할 수 있다. 일 실시예에서, 예를 들어, 컴퓨터 판독가능한 저장 매체는 프로세서에 의해 실행되었을 때 프로세서로 하여금 전술한 실시예들에 따른 방법들 및/또는 동작들을 실행하도록 하는 실행가능 컴퓨터 프로그램 명령어들을 저장할 수 있다. 실행가능 컴퓨터 프로그램 명령어들은 컴퓨터에게 특정 기능을 수행하도록 명령하기 위한 미리 정의된 컴퓨터 언어, 매너 또는 신택스에 따라 구현될 수 있다. 명령어들은 임의의 적절한 고-레벨, 저-레벨, 객체 지향적, 비주얼, 컴파일된 및/또는 해석된 프로그래밍 언어를 사용하여 구현 가능하다.

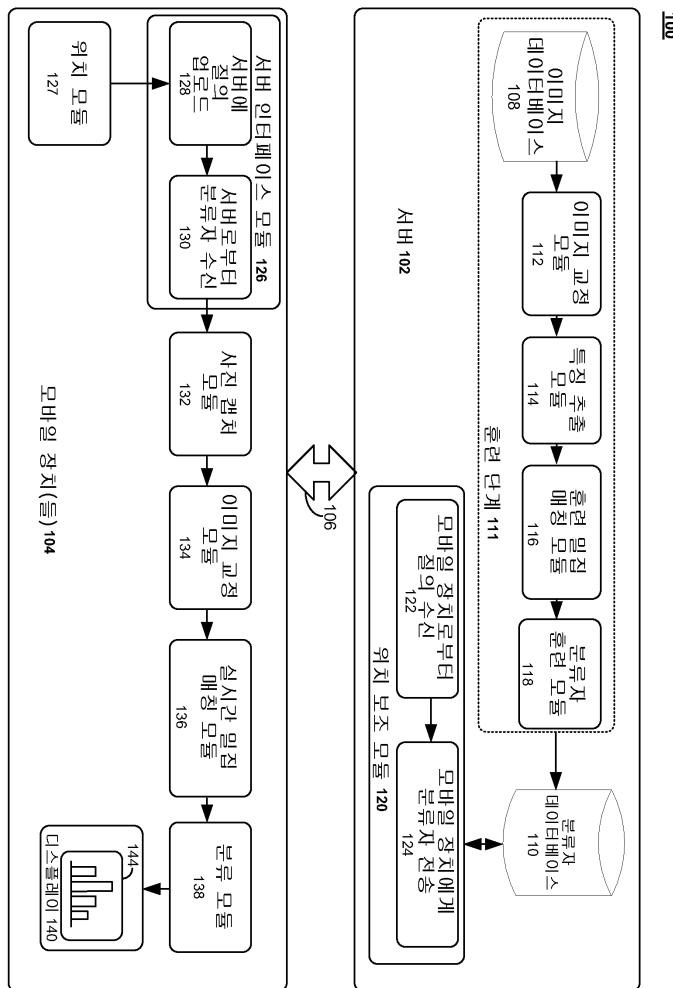
[0079] 다양한 실시예들에서, 여기서 설명된 시스템(100)은 복수의 컴포넌트, 프로그램, 프로시저, 모듈들을 가지는 컴퓨터 구현된 시스템을 포함할 수 있다. 여기서 사용된 용어들은 하드웨어, 소프트웨어, 또는 하드웨어 및 소프트웨어의 조합을 포함하는 컴퓨터 관련 엔티티를 지칭하도록 의도된다. 예를 들어, 컴포넌트는 프로세서 상에 실행되는 프로세스, 하드 디스크 드라이브, (광학 및/또는 자기 저장 매체의) 복수의 저장 드라이브, 객체, 실행의 스레드, 프로그램, 및/또는 컴퓨터로서 구현될 수 있다. 설명을 위해, 서버 상에 실행되는 애플리케이션 및 서버 둘 다가 컴포넌트가 될 수 있다. 주어진 구현 예에서 요구되는 바에 따라, 하나 이상의 컴포넌트들이 프로세스 및/또는 실행 스레드 내에 위치할 수 있으며, 컴포넌트는 하나의 컴퓨터에 로컬화되거나 두 개 이상의 컴퓨터들 사이에 분산될 수 있다. 다만 본 실시예들은 이러한 방식에 의해 한정되지는 않는다.

부호의 설명

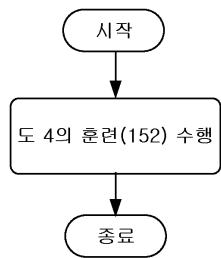
100: 위치 기반 인식 시스템	102: 서버
104: 모바일 장치(들)	108: 이미지 데이터베이스
110: 분류자 데이터베이스	111: 훈련 단계
112: 이미지 교정 모듈	114: 특징 추출 모듈
116: 훈련 밀집 매칭 모듈	118: 분류자 훈련 모듈
120: 위치 보조 모듈	122: 모바일 장치로부터 질의 수신
124: 모바일 장치에게 분류자 전송	126: 서버 인터페이스 모듈
127: 위치 모듈	128: 서버에 질의 업로드
130: 서버로부터 분류자 수신	132: 사진 캡처 모듈
134: 이미지 교정 모듈	136: 실시간 밀집 매칭 모듈
138: 분류 모듈	140: 디스플레이

도면

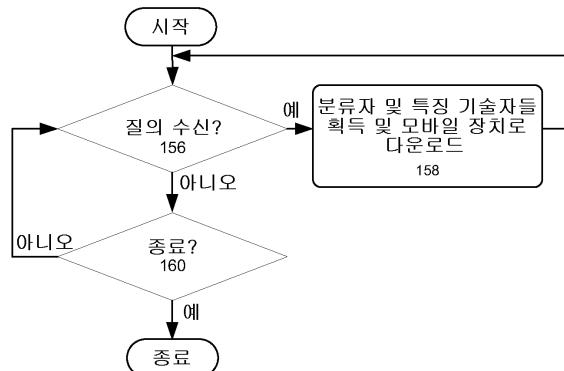
도면1



도면2

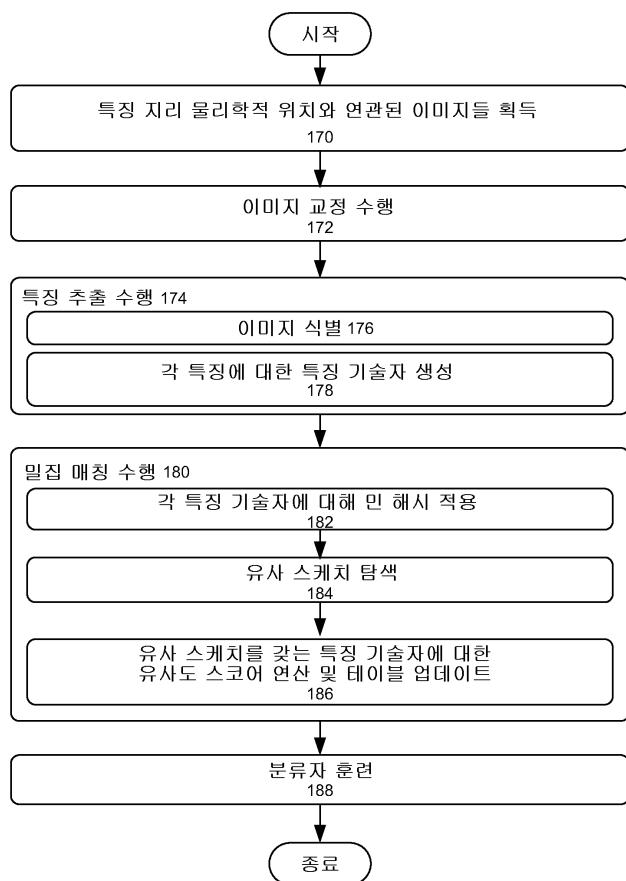


도면3

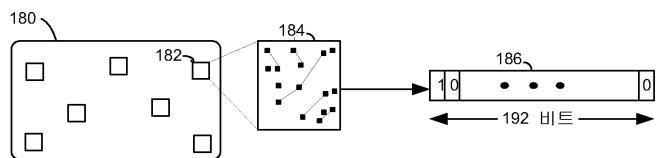


도면4

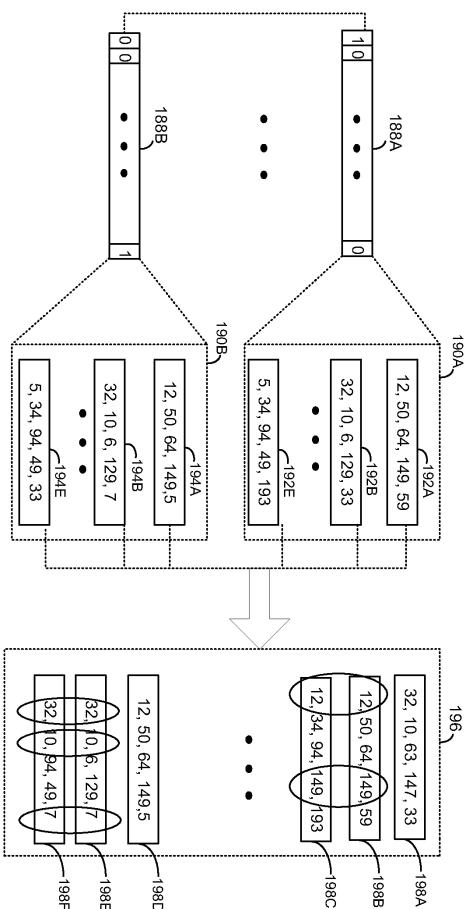
152



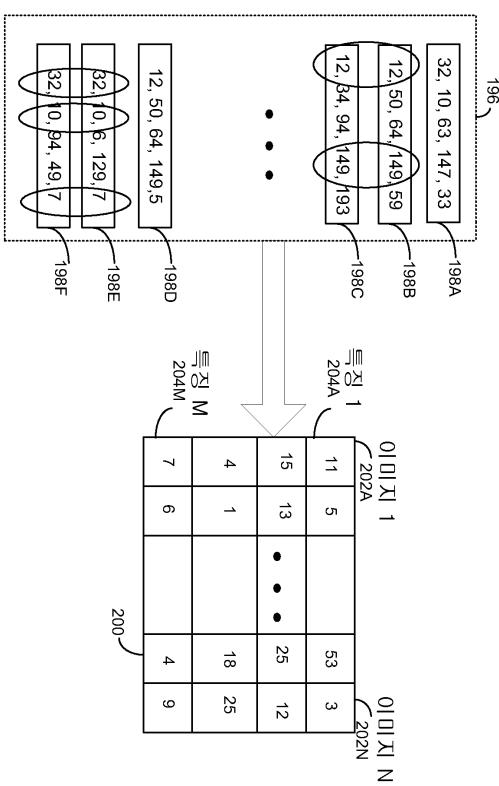
도면5



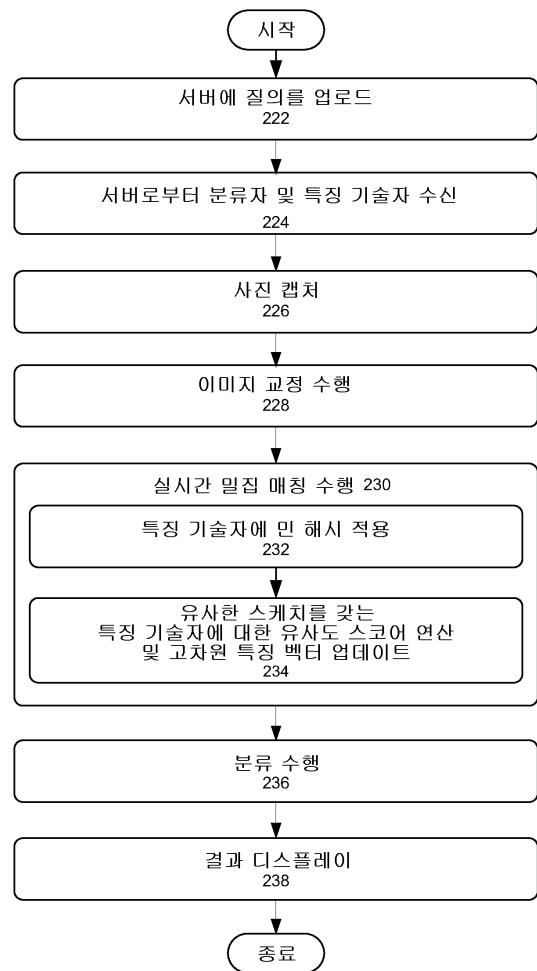
도면6



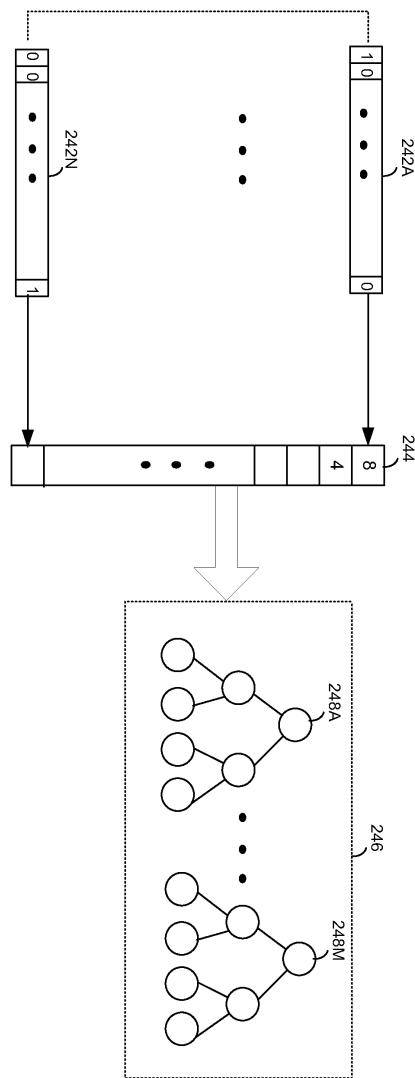
도면7



도면8

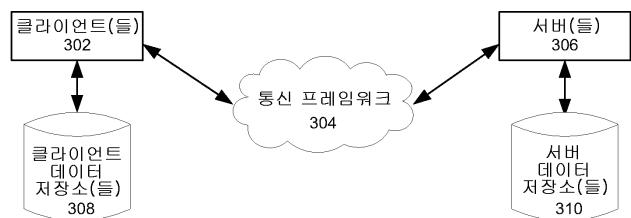


도면9

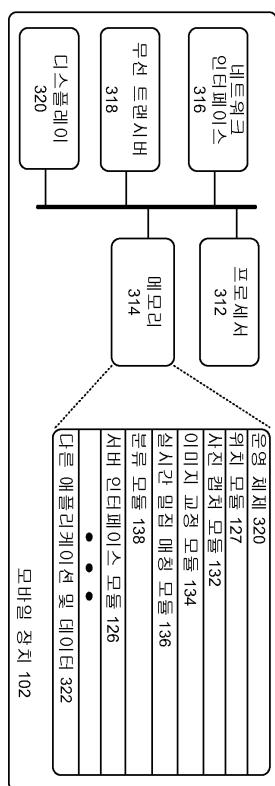


도면10

300



도면11



도면12

