(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년06월05일
(11) 등록번호 10-2119136
(24) 등록일자 2020년05월29일

(51) 국제특허분류(Int. Cl.)
G06K 9/62 (2006.01) *G06K 9/46* (2006.01)
G06T 5/40 (2006.01) *G06T 7/11* (2017.01)
(52) CPC특허분류
G06K 9/6267 (2013.01)
G06K 9/4647 (2013.01)
(21) 출원번호 10-2018-0169026
(22) 출원일자 2018년12월26일
심사청구일자 2018년12월26일
(56) 선행기술조사문헌
JP2014215852 A
JP2018013927 A
KR1020150092546 A
KR1020110103409 A

(73) 특허권자
인천대학교 산학협력단
인천광역시 연수구 아카데미로 119 (송도동)
(72) 발명자
전광길
인천광역시 연수구 경원대로119번길 21, 114동 904호(동춘동, 연수2차풍림아파트)
(74) 대리인
특허법인충정

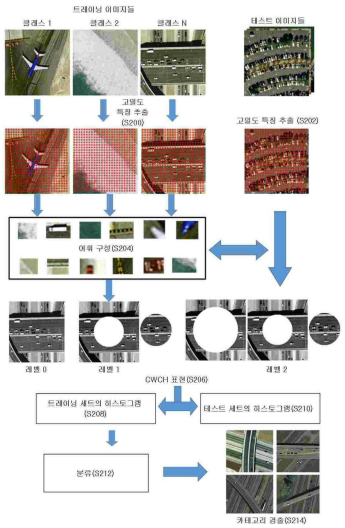
전체 청구항 수 : 총 5 항 심사관 : 강현일

(54) 발명의 명칭 **지능형 이미지 분류 방법**

(57) 요 약

본 발명의 일 실시예에 의한 지능형 이미지 분류 방법은, (A) 트레이닝 이미지 표현 획득부가, 각각의 클래스의 복수의 트레이닝 이미지에 대한 시각적 어휘를 구성하고 각각의 클래스의 트레이닝 이미지들에 대한 시각적 단어의 히스토그램을 생성하여 각각의 클래스의 이미지에 대한 최종 이미지 표현을 획득하는 단계; (B) 분류 모델 생성부가, 상기 각각의 클래스의 이미지에 대한 최종 이미지 표현에 기반하여 분류 모델을 생성하는 단계; (C) 테스트 이미지 표현 획득부가, 테스트 이미지에 대한 시각적 단어의 히스토그램을 생성하여 테스트 이미지에 대한 최종 이미지 표현을 획득하는 단계; 및 (D) 분류부가, 상기 테스트 이미지에 대한 최종 이미지 표현을 상기 분류 모델에 기반하여 분류하는 단계를 포함한다.

대 표 도 - 도2



(52) CPC특허분류
G06K 9/6201 (2013.01)
G06T 5/40 (2013.01)
G06T 7/11 (2017.01)

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

(A) 트레이닝 이미지 표현 획득부가, 각각의 클래스의 복수의 트레이닝 이미지에 대한 시각적 어휘를 구성하고 각각의 클래스의 트레이닝 이미지들에 대한 시각적 단어의 히스토그램을 생성하여 각각의 클래스의 이미지에 대한 최종 이미지 표현을 획득하는 단계;

(B) 분류 모델 생성부가, 상기 각각의 클래스의 이미지에 대한 최종 이미지 표현에 기반하여 분류 모델을 생성하는 단계;

(C) 테스트 이미지 표현 획득부가, 테스트 이미지에 대한 시각적 단어의 히스토그램을 생성하여 테스트 이미지에 대한 최종 이미지 표현을 획득하는 단계; 및

(D) 분류부가, 상기 테스트 이미지에 대한 최종 이미지 표현을 상기 분류 모델에 기반하여 분류하는 단계를 포함하고,

상기 단계 (A)는,

(A-1) 상기 트레이닝 이미지들을 그레일스케일 이미지로 변환하는 단계;

(A-2) 상기 단계 (A-1)에서 변환된 그레이스케일 이미지에서 고밀도 특징들을 추출하는 단계;

(A-3) 상기 단계 (A-2)에서 추출된 특징들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득하여 시각적 어휘를 구성하는 단계;

(A-4) 상기 트레이닝 이미지들을 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계;

(A-5) 각 레벨에 상이한 가중치를 적용하여 상기 트레이닝 이미지들의 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성하는 단계; 및

(A-6) 상기 각 클래스의 트레이닝 이미지들과 관련하여 생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 상기 각 클래스의 트레이닝 이미지들에 대한 최종 이미지 표현을 획득하는 단계를 포함하며,

상기 단계 (C)는,

(C-1) 상기 테스트 이미지를 그레일스케일 이미지로 변환하는 단계;

(C-2) 상기 단계 (C-1)에서 변환된 그레이스케일 이미지에서 고밀도 특징들을 추출하는 단계;

(C-3) 상기 단계 (C-2)에서 추출된 특징들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득하는 단계;

(C-4) 상기 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계;

(C-5) 각 레벨에 상이한 가중치를 적용하여 상기 테스트 이미지의 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성하는 단계; 및

(C-6) 상기 테스트 이미지와 관련하여 생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 상기 테스트 이미지에 대한 최종 이미지 표현을 획득하는 단계를 포함하는, 지능형 이미지 분류 방법.

청구항 2

청구항 1에 있어서,

상기 단계 (A-4)에서 상기 트레이닝 이미지들을 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계, 및 상기 단계 (C-4) 상기 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계에서,

상기 트레이닝 이미지들과 상기 테스트 이미지는, k번째 레벨이 (k+1)개의 세그먼트를 갖도록 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할되고,

상기 동심원의 k번째 레벨의 반경(r_k)은,

[수학식 5]

$$r_k = \frac{k}{L} min\{c_x, c_y\}$$

에 따라 계산되며,

c_x, c_y는,

[수학식 6]

$$c_x = \frac{1}{|I|}\sum_{i=1}^{|I|} x_i, \quad c_y = \frac{1}{|I|}\sum_{i=1}^{|I|} y_i$$

에 따라 계산되고,

상기 I는 R×C의 크기를 갖는 이미지를 나타내며,

$$I = \{(x_i, y_i) \mid 1 \le x_i \le C, 1 \le y_i \le R\}$$

이고,

c = (c_x, c_y)는 이미지 I의 중심을 나타내며,

L은 레벨의 수를 나타내고, 가장 작은 원의 중심은 r_1이며, $|I|$는 이미지 I 내의 요소들의 수이고, k는 0 이상인 정수이며, R과 C는 0 이상인 정수인, 지능형 이미지 분류 방법.

청구항 3

청구항 2에 있어서,

상기 단계 (A-5) 및 상기 단계 (C-5)에서 각 레벨에 상이한 가중치를 적용하여 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성할 때, 레벨이 높을수록 더 큰 가중치를 할당하는, 지능형 이미지 분류 방법.

청구항 4

청구항 2에 있어서,

상기 단계 (D)는,

상기 분류부가, 상기 각 클래스의 트레이닝 이미지의 히스토그램과 상기 테스트 이미지의 히스토그램의 인터섹션(intersection)에 기반하여 각 클래스의 트레이닝 이미지와 테스트 이미지 간의 유사도를 계산하고, 계산된 유사도에 기반하여 상기 테스트 이미지를 분류하는 단계를 포함하는, 지능형 이미지 분류 방법.

청구항 5

청구항 4에 있어서,

상기 유사도는,

[수학식 7]

$$S = \sum_{k=0}^{L} w_k N_k$$

에 따라 계산되고,

$$w_k = 1/2^{L-k}$$

으로서 각 레벨에 할당되는 가중치이며,

N_k는 두 연속적인 레벨들의 히스토그램 인터섹션의 차를 나타내고,

상기 두 이미지의 히스토그램 인터섹션은,

[수학식 8]

$$HI(P,Q) = \sum_{i=1}^{c} min(P_i, Q_i)$$

에 따라 계산되며,

P와 Q는 각각 c 빈(bin)을 갖는 상기 각 클래스의 트레이닝 이미지와 상기 테스트 이미지의 히스토그램을 나타내고,

P_i는 P의 i번째 빈의 카운트(count)를 나타내며, Q_i는 Q의 i번째 빈의 카운트(count)를 나타내는, 지능형 이미지 분류 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 지능형 이미지 분류 방법에 관한 것이다.

배 경 기 술

[0002] 최근 몇 년 동안 이미지 콘텐츠의 양이 증가함에 따라, 이미지 분류는 산업계와 학계에서 상당한 주목을 받았다[20]. 분류 기반 접근법의 주요 목표는 사용 가능한 범주[39]에 따라 이미지를 분류하는 것이다. 규모, 공간 레이아웃 및 조명의 변화는 이미지 분류를 어려운 과제로 만든다[1]. 과거에는 전역 특징의 구현이 간단하기 때문에 전역 색상, 텍스처 및 모양 특징이 이미지 분류에 사용되었다. 그러나 그들의 성능은 낮다.

[0003] BoVW(Bag of Visual Words) 모델은 이미지 분류, 이미지 검색 및 객체 인식 문제[24]에 사용된다. BoVW 모델에서 시각 정보는 국부적인 특징을 이용하여 영상으로부터 추출되며, 시각적 단어의 순서가 없는 히스토그램이 영상 표현을 위해 사용된다[29]. 접근 방식에서 이미지는 단어로 구성된 문서처럼 취급된다. 일군의 트레이닝 이미지는 코드북 또는 시각적 어휘로 알려진 이미지 패치를 나타내는 데 사용된다. BoVW 모델에서 사용된 단계는 (i) 키-포인트 검출기를 이용한 특징 추출, (ii) 코드북 구성, (iii) 시각적 단어의 순서 없는 히스토그램과 같은 이미지 표현[29, 23]이다. BoVW 모델에서 시각적 단어(visual words)의 히스토그램은 2-D 이미지 공간의 공간 레이아웃을 고려하지 않고 계산된다. BoVW의 성능은 이미지의 공간적 세부 사항에 대한 정보 부족으로 인해 어려움을 겪는다. 그러나 연구에 따르면 이 공간 정보는 이미지 분류 기반 문제에 유용하다[1, 18].

[0004] 이 문제를 해결하기 위해 두 가지 접근법이 제시된다: (i) 시각적 단어 동시 발생 간의 기하학적 관계와, (ii) 시각적 단어의 히스토그램 구성을 위한 다른 영역으로의 이미지 분할[1, 4, 16]. 시각적 단어 사이의 복잡한 관계로 인해, 시각적 단어 동시 발생에 기반한 접근법은 계산상으로 비싸다고 보고되고 있다[1, 4]. 이미지를 다른 영역으로 분할하는 기술 중에서 공간 피라미드 매칭(SPM: Spatial Pyramid Matching)이 가장 많이 사용되는 것으로 간주된다. 그것은 이미지를 사각형으로 나누고 각 서브 영역으로부터 시각적 단어의 히스토그램을 계산한다.

[0005] SPM은 순서 없는 BoVW의 연장으로 간주되며 공간 정보 부족의 단점을 해결한다[5, 37, 6]. 그러나 SPM은 회전 불변적이 아니며 포즈와 뷰 포인트를 변경할 수 없고 매우 높은 차원 공간의 이미지를 나타내기 때문에, 이미지의 물체를 뒤집거나 돌리면 SPM의 식별력은 감소한다[37, 15].

[0006] Zhao et al.[37]는 이미지의 회전 불변 공간 레이아웃 정보를 토지 이용 장면 분류를 위한 BoVW 모델에 통합하기 위해 CCM-BOVW (Concentric Circle-structured Multiscale BoVW) 모델을 제안한다. 첫 번째 단계로 CCM-BOVW는 여러 해상도 이미지를 만들고 모든 해상도 이미지에서 여러 국부적 특징을 추출한다. 그 다음에 시각적 어휘 구성 단계가 수행되고, 그 후에 각 이미지는 동심원 집합에 의해 서브 영역으로 분할된다. 각 서브 영역은 시각적 단어의 히스토그램으로 개별적으로 표시되며, 이미지의 각 서브 영역으로부터의 히스토그램들은 연결된

다. 이미지의 최종 히스토그램 표현을 만들려면 다양한 해상도 이미지의 히스토그램을 결합해야 한다.

[0007] Karmakar et al.[15]에 의하면, 동심 사각형 위에 히스토그램을 구축하면 이미지 회전과 시점 변화를 처리할 수 있다. 본 발명에서는 동심 가중 원의 히스토그램에 기반한 새로운 이미지 표현을 제안함으로써 이 문제를 해결하고자 한다. 도 1에 도시된 두 가지 경우의 예를 고려해 보자.

[0008] 도 1의 위쪽 영역은 동심 직사각형 영역[15] 위에 구축된 히스토그램의 개념을 나타내며, 아래 영역은 가중 동심원을 기반으로 제안된 접근법을 사용하여 구성된 히스토그램을 나타낸다. 도 1에서, (a) 이미지가 회전되지 않으며, (b) 이미지가 110도 회전하고, (c) 이미지가 300도 회전한다.

[0009] 직사각형 좌표를 사용하는 시각적 단어는 중심에서 등거리가 아니므로 회전 불변성은 달성되지 않지만, 원에서 모든 점은 중심에서 등거리에 있으므로 동심원을 사용하는 시각적 단어는 회전 불변성을 얻는다([15]와 비교).

[0010] BoVW 모델의 주요 한계는 공간 정보를 무시한다는 사실이다[30,16]. 이 사실에도 불구하고, BoVW는 높은 판별력을 나타내며 이미지 분류에서 우수한 결과를 보여주었다[13, 12, 36]. 그러나 수많은 연구 결과에 따르면 누락된 공간 정보를 통합하여 성능을 향상시킬 수 있다. 일반적으로 관련 연구는 두 그룹으로 분류할 수 있다[3, 17]. 첫 번째 그룹은 시각적 단어의 관계 또는 동시 발생을 인코딩하는 방법으로 구성된다[16, 26]. Khan et al. [16]은 이 도메인에서 주목할만한 공헌을 했고 BoVW 모델에 전역 공간 정보를 통합했다. 이는 동일한 단어 쌍(Pair of Identical Words, PIW) 간의 전역 기하학적 관계를 고려한 것이다. 동일한 시각적 단어 사이의 각도를 기반으로 하는 정규화된 히스토그램이 만들어지는데, 이것은 동일한 시각적 단어 각도 히스토그램의 쌍(PIWAH: Pairs of Identical Visual Words Angle Histogram)이라고 불린다. 이러한. PIWAH 표현은 기하학적 변환, 즉 스케일링 및 평행 이동에는 불변하지만 회전에 민감하다[16, 3]. Anwar et al.[3]은 회전 불변성을 획득하기 위해 연구를 확장하고 동일한 시각적 단어의 트리플릿(Triplets of Identical Visual Words: TIW) 사이의 관계를 계산하여 동일한 시각적 단어 각도 히스토그램의 트리플릿(TIWAH: Triplets of Identical Visual Words Angle Histogram)을 제안한다. 그러나 시각적 단어들의 동시 발생 사이의 관계를 계산하는 것에 기반한 접근법은 계산상으로 비싸다는 것이 알려져 있다.

[0011] 두 번째 그룹은 이미지를 다른 모양의 서브 영역으로 나누는 방법을 포함한다. 시각적 단어에 관한 정보는 선택된 영역들 각각으로부터 계산된다. 이 영역에서 가장 주목할만한 연구는 Lazebnik et al.[18]의 연구인데, 이 연구자는 원래 SPM을 제안한 사람이다. SPM에서 이미지는 직사각형의 서브 영역으로 나누어지며 각 영역의 시각적 단어 통계가 집계된다. 최종 히스토그램은 각 영역에서 추출한 히스토그램의 연결이다. Anwar et al.[2]는 고밀도 샘플링 기반 BoVW 접근법을 제시하고 문화 유산 분류 기반 문제에 대한 세 가지 유형의 방식, 즉, 직사각형 타일링, 로그-극좌표 타일링 및 원형 타일링을 평가한다.

[0012] 제안된 접근법의 효과를 증명하기 위해 본 발명의 접근 방법과 동시에 [18, 37, 15] 방법을 살펴보고, 또한 분류 성능을 향상시키기 위한 특징 추출[39, 8, 9] 및 중간 특징 표현 [21, 22, 7] 기법과 같은 다른 접근법에 초점을 맞춘 최근의 최신 발명이 선택되었다.

[0013] Mekhalfi et al.[21]은 압축 센싱과 다중-특징 프레임워크를 사용하여 이미지를 콤팩트하게 표현하는 새로운 방안을 제안한다. 그들의 방법은 토지 이용 이미지 데이터 세트에 대한 최첨단 방법과 비교하여 상당한 성능 향상을 가져온다. 더 나은 분류 성능을 얻기 위해, Zou et al. [39]는 국부적 및 전역 이미지 특징의 융합의 생성을 제안한다. 이를 위해 먼저 BoVW 및 SPM을 사용하여 국부적 특징을 추출한다. 전역 특징을 추출하기 위해 멀티 스케일 CLBP(MS-CLBP)를 사용한다. 특징 표현을 위해 KCRC(Kernel Collaborative Representation-Based Classification)를 사용한다. 표현 후에는 잔상을 두 가지 유형의 특징들로부터 획득하고, 라벨은 가중된 잔상의 합에 기반하여 할당된다.

[0014] Zhu et al.[38]은 고 공간 해상도(HSR: High Spatial Resolution) 장면 이미지 분류를 위한 국부적 및 전역 특징의 융합 전략을 제안한다. 논문[9]에서 Bian et al.은 분류 성능을 높이기 위해 국부적 및 전역 서술자의 융합을 제안한다. 그것들은 이미지 구조의 전역 구조와 정밀한 국부적 디테일을 결합하여 특징 표현을 풍부하게 한다. 전역 회전 불변 특징을 추출하려면, 전역 돌극성(saliency) 기반 다중스케일, 다중해상도 및 다중구조 LBP가 사용되고, 국부적 코드북 없는 모델(CLM: CodeBookless Model)이 국부적 식별 특징을 나타내기 위해 사용된다. Bian et al.은 최첨단 분류 방법보다 더 좋거나 유사한 성능을 나타낸다. Chen et al.[10]은 토지 이용 장면 분류를 위한 텍스쳐 기반 패러다임, 즉 MS-CLBP(Multi-Scale Completed local binary pattern) 서술자를 제안한다. CLBP는 회전 불변의 텍스처 분류에 효과적이라고 알려져 있다. CLBP를 향상시켜 여러 해상도에서 지배적인 텍스처 특징을 특성화하고 최첨단 기술에 비해 성능을 지속적으로 향상시킨다.

[0015] 오브젝트 뱅크(OB: Object Bank)는 공간 정보와 의미 정보를 인코딩하는 하이-레벨 이미지 표현이다[19]. 그러나 OB 접근법은 고차원의 단점을 가지고 있으며, OB의 크기를 줄이고 성능을 향상시키기 위해 다양한 접근법이 문헌에서 제안되었다[19,34]. OB 표현의 성능을 향상시키기 위해 Zang et al. [34]는 임계 값 필터 방법을 제안했다. 그들은 Matthew 효과 정규화 방법을 사용하여 OB 표현을 단순화하고 보다 컴팩트한 서술자를 작성했으며 세 가지 실제 데이터 세트에서 이미지 서술자의 실질적인 차원 감소와 함께 향상된 성능을 보였다.

[0016] Cheng 등[11]은 심층 컨볼루션(convolutional) 특징을 사용하여 원격 센싱 이미지 분류를 위한 시각적 어휘를 생성하는 컨볼루션 특징 모음(BoCF: Bag of Convolutional Features)을 제안한다. Lu et al.[14]는 원격 센싱 장면 분류를 위한 감독되지 않은 표현 학습 방법을 제안한다. Scott et al.[27]은 미세 조정 및 증강과 함께 전송 학습(TL: Transfer Learning)을 사용할 것을 제안한다. 그들은 고해상도 영상 분류에 대한 제안된 접근법의 효과를 평가하고 가장 효과적인 방법보다 훨씬 더 높은 정확도를 달성한다. Scott et al.[28]은 투표, 가중 평균 및 퍼지 적분을 포함하는 다양한 융합 기법을 조사하여 CaffeNet, GoogLeNet 및 ResNet50과 같은 다중 DCNN 토지 표지 분류자를 단일 집계 분류기에 통합한다. 심층 학습과 신경망 기반 접근법은 트레이닝을 위해 엄청난 양의 데이터(수백만)와 시간이 필요하다는 것을 여기서 언급할 가치가 있다. BoVW 모델은 사전 초기화 또는 트레이닝 없이 채택할 수 있는 플러그 앤 플레이 방식이다.

선행기술문헌

비특허문헌

[0017] (비특허문헌 0001) [1] Ali, N., Bajwa, K.B., Sablatnig, R., Mehmood, Z.: Image retrieval by addition of spatial information based on histograms of triangular regions. Computers & Electrical Engineering 54, 539-550 (2016)

(비특허문헌 0002) [2] Anwar, H., Zambanini, S., Kampel, M.: Supporting ancient coin classification by image-based reverse side symbol recognition. In: International Conference on Computer Analysis of Images and Patterns. pp. 17-25. Springer (2013)

(비특허문헌 0003) [3] Anwar, H., Zambanini, S., Kampel, M.: Encoding spatial arrangements of visual words for rotation-invariant image classification. In: German Conference on Pattern Recognition. pp. 443-452. Springer (2014)

(비특허문헌 0004) [4] Anwar, H., Zambanini, S., Kampel, M.: Efficient scale-and rotation-invariant encoding of visual words for image classification. IEEE Signal Processing Letters 22(10), 1762-1765 (2015)

(비특허문헌 0005) [5] Ashraf, R., Ahmed, M., Ahmad, U., Habib, M.A., Jabbar, S., Naseer, K.: Mdcbir-mf: multimedia data for content-based image retrieval by using multiple features. Multimedia Tools and Applications pp. 1-27 (2018)

(비특허문헌 0006) [6] Ashraf, R., Ahmed, M., Jabbar, S., Khalid, S., Ahmad, A., Din, S., Jeon, G.: Content based image retrieval by using color descriptor and discrete wavelet transform. Journal of medical systems 42(3), 44 (2018)

(비특허문헌 0007) [7] Ashraf, R., Bajwa, K.B., Mahmood, T.: Content-based image retrieval by exploring bandletized regions through support vector machines. J. Inf. Sci. Eng. 32(2), 245-269 (2016)

(비특허문헌 0008) [8] Ashraf, R., Bashir, K., Irtaza, A., Mahmood, M.T.: Content based image retrieval using embedded neural networks with bandletized regions. Entropy 17(6), 3552-3580 (2015)

(비특허문헌 0009) [9] Bian, X., Chen, C., Tian, L., Du, Q.: Fusing local and global features for high-resolution scene classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2017)

(비특허문헌 0010) [10] Chen, C., Zhang, B., Su, H., Li, W., Wang, L.: Land-use scene classification using multi-scale completed local binary patterns. Signal, image and video processing 10(4), 745-752 (2016)

(비특허문헌 0011) [11] Cheng, G., Li, Z., Yao, X., Guo, L., Wei, Z.: Remote sensing image scene classification using bag of convolutional features. IEEE Geoscience and Remote Sensing Letters 14(10), 1735-1739 (2017)

(비특허문헌 0012) [12] Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Workshop on statistical learning in computer vision, ECCV. pp. 1-2. Prague (2004)

(비특허문헌 0013) [13] Fei-Fei, L., Perona, P.: A bayesian hierarchical model for learning natural scene categories. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. vol. 2, pp. 524-531. IEEE (2005)

(비특허문헌 0014) [14] Fu, M., Yuan, Y., Lu, X.: Remote sensing scene classification by unsupervised representation learning. IEEE Trans. Geoscience and Remote Sensing 55(9), 5148-5157 (2017)

(비특허문헌 0015) [15] Karmakar, P., Teng, S.W., Lu, G., Zhang, D.: Rotation invariant spatial pyramid matching for image classification. In: Digital Image Computing: Techniques and Applications (DICTA), 2015 International Conference on. pp. 1-8. IEEE (2015)

(비특허문헌 0016) [16] Khan, R., Barat, C., Muselet, D., Ducottet, C.: Spatial orientations of visual word pairs to improve bag-of-visual-words model. In: Proceedings of the British Machine Vision Conference. pp. 89-1. BMVA Press (2012)

(비특허문헌 0017) [17] Khan, R., Barat, C., Muselet, D., Ducottet, C.: Spatial histograms of soft pairwise similar patches to improve the bag-of-visual-words model. Computer Vision and Image Understanding 132, 102-112 (2015)

(비특허문헌 0018) [18] Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: Computer vision and pattern recognition, 2006 IEEE computer society conference on. vol. 2, pp. 2169-2178. IEEE (2006)

(비특허문헌 0019) [19] Li, L.J., Su, H., Lim, Y., Fei-Fei, L.: Object bank: An object-level image representation for high-level visual recognition. International journal of computer vision 107(1), 20-39 (2014)

(비특허문헌 0020) [20] Ma, L., Li, M., Ma, X., Cheng, L., Du, P., Liu, Y.: A review of supervised object-based landcover image classification. ISPRS Journal of Photogrammetry and Remote Sensing 130, 277-293 (2017)

(비특허문헌 0021) [21] Mekhalfi, M.L., Melgani, F., Bazi, Y., Alajlan, N.: Land-use classification with compressive sensing multifeature fusion. IEEE Geoscience and Remote Sensing Letters 12(10), 2155-2159 (2015)

(비특허문헌 0022) [22] Nazir, A., Ashraf, R., Hamdani, T., Ali, N.: Content based image retrieval system by using hsv color histogram, discrete wavelet transform and edge histogram descriptor. In: Computing, Mathematics and Engineering Technologies (iCoMET), 2018 International Conference on. pp. 1-6. IEEE (2018)

(비특허문헌 0023) [23] Nowak, E., Jurie, F., Triggs, B.: Sampling strategies for bag-of-features image classification. Computer Vision-ECCV 2006 pp. 490-503 (2006)

(비특허문헌 0024) [24] O'Hara, S., Draper, B.A.: Introduction to the bag of features paradigm for image classification and retrieval. arXiv preprint arXiv:1101.3354 (2011)

(비특허문헌 0025) [25] Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. International journal of computer vision 42(3), 145-175 (2001)

(비특허문헌 0026) [26] Savarese, S., Winn, J., Criminisi, A.: Discriminative object class models of appearance and shape by correlatons. In: Computer Vision and Pattern Recognition, 2006 IEEE Computer

Society Conference on. vol. 2, pp. 2033-2040. IEEE (2006)

(비특허문헌 0027) [27] Scott, G.J., England, M.R., Starms, W.A., Marcum, R.A., Davis, C.H.: Training deep convolutional neural networks for land-cover classification of high-resolution imagery. IEEE Geoscience and Remote Sensing Letters 14(4), 549-553 (2017)

(비특허문헌 0028) [28] Scott, G.J., Marcum, R.A., Davis, C.H., Nivin, T.W.: Fusion of deep convolutional neural networks for land cover classification of high-resolution imagery. IEEE Geoscience and Remote Sensing Letters 14(9), 1638-1642 (2017)

(비특허문헌 0029) [29] Sivic, J., Zisserman, A.: Video google: A text retrieval approach to object matching in videos. In: null. p. 1470. IEEE (2003)

(비특허문헌 0030) [30] Song, Y., McLoughlin, I.V., Dai, L.R.: Local coding based matching kernel method for image classification. PloS one 9(8), e103575 (2014)

(비특허문헌 0031) [31] Tuytelaars, T.: Dense interest points. In: Computer Vision and Pattern Recognition. IEEE (2010)

(비특허문헌 0032) [32] Vedaldi, A., Zisserman, A.: Sparse kernel approximations for efficient classification and detection. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. pp. 2320-2327. IEEE (2012)

(비특허문헌 0033) [33] Yang, Y., Newsam, S.: Bag-of-visual-words and spatial extensions for land-use classification. In: Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. pp. 270-279. ACM (2010)

(비특허문헌 0034) [34] Zang, M., Wen, D., Liu, T., Zou, H., Liu, C.: A pooled object bank descriptor for image scene classification. Expert Systems with Applications 94, 250-264 (2018)

(비특허문헌 0035) [35] Zhang, D., Islam, M.M., Lu, G.: A review on automatic image annotation techniques. Pattern Recognition 45(1), 346-362 (2012)

(비특허문헌 0036) [36] Zhang, J., Marszałek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: A comprehensive study. International journal of computer vision 73(2), 213-238 (2007)

(비특허문헌 0037) [37] Zhao, L.J., Tang, P., Huo, L.Z.: Land-use scene classification using a concentric circlestructured multiscale bag-of-visual-words model. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7(12), 4620-4631 (2014)

(비특허문헌 0038) [38] Zhu, Q., Zhong, Y., Zhao, B., Xia, G.S., Zhang, L.: Bag-of-visual-words scene classifier with local and global features for high spatial resolution remote sensing imagery. IEEE Geoscience and Remote Sensing Letters 13(6), 747-751 (2016)

(비특허문헌 0039) [39] Zou, J., Li, W., Chen, C., Du, Q.: Scene classification using local and global features with collaborative representation fusion. Information Sciences 348, 209-226 (2016)

발명의 내용

해결하려는 과제

[0018] 본 발명이 해결하고자 하는 과제는 이미지의 히스토그램 표현에 공간 정보를 추가함으로써 이미지 표현에 대해 회전 불변성을 획득하여 이미지를 더 정확하게 식별하여 분류할 수 있는 지능형 이미지 분류 방법을 제공하는 것이다.

과제의 해결 수단

[0019] 상기 과제를 해결하기 위한 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법은,

[0020] (A) 트레이닝 이미지 표현 획득부가, 각각의 클래스의 복수의 트레이닝 이미지에 대한 시각적 어휘를 구성하고 각각의 클래스의 트레이닝 이미지들에 대한 시각적 단어의 히스토그램을 생성하여 각각의 클래스의 이미지에 대한 최종 이미지 표현을 획득하는 단계;

[0021] (B) 분류 모델 생성부가, 상기 각각의 클래스의 이미지에 대한 최종 이미지 표현에 기반하여 분류 모델을 생성하는 단계;

[0022] (C) 테스트 이미지 표현 획득부가, 테스트 이미지에 대한 시각적 단어의 히스토그램을 생성하여 테스트 이미지에 대한 최종 이미지 표현을 획득하는 단계; 및

[0023] (D) 분류부가, 상기 테스트 이미지에 대한 최종 이미지 표현을 상기 분류 모델에 기반하여 분류하는 단계를 포함하고,

[0024] 상기 단계 (A)는,

[0025] (A-1) 상기 트레이닝 이미지들을 그레일스케일 이미지로 변환하는 단계;

[0026] (A-2) 상기 단계 (A-1)에서 변환된 그레이스케일 이미지에서 고밀도 특징들을 추출하는 단계;

[0027] (A-3) 상기 단계 (A-2)에서 추출된 특징들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득하여 시각적 어휘를 구성하는 단계;

[0028] (A-4) 상기 트레이닝 이미지들을 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계;

[0029] (A-5) 각 레벨에 상이한 가중치를 적용하여 상기 트레이닝 이미지들의 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성하는 단계; 및

[0030] (A-6) 상기 각 클래스의 트레이닝 이미지들과 관련하여 생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 상기 각 클래스의 트레이닝 이미지들에 대한 최종 이미지 표현을 획득하는 단계를 포함하며,

[0031] 상기 단계 (C)는,

[0032] (C-1) 상기 테스트 이미지를 그레일스케일 이미지로 변환하는 단계;

[0033] (C-2) 상기 단계 (C-1)에서 변환된 그레이스케일 이미지에서 고밀도 특징들을 추출하는 단계;

[0034] (C-3) 상기 단계 (C-2)에서 추출된 특징들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득하는 단계;

[0035] (C-4) 상기 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계;

[0036] (C-5) 각 레벨에 상이한 가중치를 적용하여 상기 테스트 이미지의 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성하는 단계; 및

[0037] (C-6) 상기 테스트 이미지와 관련하여 생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 상기 테스트 이미지에 대한 최종 이미지 표현을 획득하는 단계를 포함한다.

[0038] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 있어서, 상기 단계 (A-4)에서 상기 트레이닝 이미지들을 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계, 및 상기 단계 (C-4) 상기 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계에서,

[0039] 상기 트레이닝 이미지들과 상기 테스트 이미지는, k번째 레벨이 (k+1)개의 세그먼트를 갖도록 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할되고,

[0040] 상기 동심원의 k번째 레벨의 반경(r_k)은,

[0041] [수학식 5]

[0042] $$r_k = \frac{k}{L} min\{c_x, c_y\}$$ 에 따라 계산되며,

[0043] c_x, c_y는,

[0044] [수학식 6]

$$c_x = \frac{1}{|I|}\sum_{i=1}^{|I|} x_i, \quad c_y = \frac{1}{|I|}\sum_{i=1}^{|I|} y_i$$

[0045] 에 따라 계산되고,

[0046] 상기에서 I는 R×C의 크기를 갖는 이미지를 나타내며,

[0047] $I = \{(x_i, y_i) \mid 1 \le x_i \le C, 1 \le y_i \le R\}$ 이고,

[0048] c = (c_x, c_y)는 이미지 I의 중심을 나타내며,

[0049] L은 레벨의 수를 나타내고, 가장 작은 원의 중심은 r_1이며, $|I|$는 이미지 I 내의 요소들의 수이고, k는 0 이상인 정수이며, R과 C는 0 이상인 정수일 수 있다.

[0050] 또한, 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 있어서, 상기 단계 (A-5) 및 상기 단계 (C-5)에서 각 레벨에 상이한 가중치를 적용하여 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성할 때, 레벨이 높을수록 더 큰 가중치를 할당할 수 있다.

[0051] 또한, 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 있어서, 상기 단계 (D)는,

[0052] 상기 분류부가, 상기 각 클래스의 트레이닝 이미지의 히스토그램과 상기 테스트 이미지의 히스토그램의 인터섹션(intersection)에 기반하여 각 클래스의 트레이닝 이미지와 테스트 이미지 간의 유사도를 계산하고, 계산된 유사도에 기반하여 상기 테스트 이미지를 분류하는 단계를 포함할 수 있다.

[0053] 또한, 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 있어서, 상기 유사도는,

[0054] [수학식 7]

$$S = \sum_{k=0}^{L} w_k N_k$$

[0055] 에 따라 계산되고,

[0056] $w_k = 1/2^{L-k}$ 으로서 각 레벨에 할당되는 가중치이며,

[0057] N_k는 두 연속적인 레벨들의 히스토그램 인터섹션의 차를 나타내고,

[0058] 상기 두 이미지의 히스토그램 인터섹션은,

[0059] [수학식 8]

$$HI(P,Q) = \sum_{i=1}^{c} min(P_i, Q_i)$$

[0060] 에 따라 계산되며,

[0061] P와 Q는 각각 c 빈(bin)을 갖는 상기 각 클래스의 트레이닝 이미지와 상기 테스트 이미지의 히스토그램을 나타내고,

[0062] P_i는 P의 i번째 빈의 카운트(count)를 나타내며, Q_i는 Q의 i번째 빈의 카운트(count)를 나타낼 수 있다.

발명의 효과

[0063] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 의하면, 이미지의 히스토그램 표현에 공간 정보를 추가

함으로써 이미지 표현에 대해 회전 불변성을 획득하여 이미지를 더 정확하게 식별하여 분류할 수 있다.

도면의 간단한 설명

[0064] 도 1은 동심원에 기반한 본 발명의 일 실예에 의한 이미지 분류 방법이 동심 사각형보다 더 식별력이 있다는 것을 나타내는 도면.

도 2는 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법의 개념도.

도 3은 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법의 흐름도.

도 4는 도 3에 도시된 각각의 클래스의 트레이닝 이미지들에 대한 최종 이미지 표현 획득 단계의 상세 흐름도.

도 5는 도 3에 도시된 테스트 이미지에 대한 최종 이미지 표현 획득 단계의 상세 흐름도.

도 6은 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법이 적용될 수 있는 예시적인 장치의 블록도.

도 7은 15-장면 이미지 데이터 세트[18]의 각 카테고리의 예시적인 이미지를 도시한 도면.

도 8은 15-장면 이미지 데이터 세트를 이용한 어휘 크기의 함수로서의 평균 정확도를 도시한 그래프.

도 9는 15-장면 이미지 데이터 세트에 대한 LGF[39]와 CWCH의 클래스별 비교를 도시한 그래프.

도 10은 15-장면 이미지 데이터 세트에 대한 혼동 행렬을 도시한 도면.

도 11은 UCM 이미지 데이터 세트[33]의 각 카테고리에서의 예시적인 이미지를 도시한 도면.

도 12는 UCM 이미지 데이터 세트를 사용한 어휘 크기의 함수로서의 평균 정확도를 도시한 도면.

도 13은 UCM 이미지 데이터 세트에 대한 CCM-BOVW[37]와 CWCH의 클래스별 비교를 도시한 그래프.

도 14는 UCM 이미지 데이터 세트에 대한 혼동 행렬을 도시한 도면.

발명을 실시하기 위한 구체적인 내용

[0065] 본 발명의 목적, 특정한 장점들 및 신규한 특징들은 첨부된 도면들과 연관되어지는 이하의 상세한 설명과 바람직한 실시예들로부터 더욱 명백해질 것이다.

[0066] 이에 앞서 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이고 사전적인 의미로 해석되어서는 아니되며, 발명자가 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있는 원칙에 입각하여 본 발명의 기술적 사상에 부합되는 의미와 개념으로 해석되어야 한다.

[0067] 본 명세서에서 각 도면의 구성요소들에 참조번호를 부가함에 있어서, 동일한 구성 요소들에 한해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 번호를 가지도록 하고 있음에 유의하여야 한다.

[0068] 또한, "제1", "제2", "일면", "타면" 등의 용어는, 하나의 구성요소를 다른 구성요소로부터 구별하기 위해 사용되는 것으로, 구성요소가 상기 용어들에 의해 제한되는 것은 아니다.

[0069] 이하, 본 발명을 설명함에 있어, 본 발명의 요지를 불필요하게 흐릴 수 있는 관련된 공지 기술에 대한 상세한 설명은 생략한다.

[0070] 이하, 첨부된 도면을 참조하여 본 발명의 바람직한 실시형태를 상세히 설명하기로 한다.

[0071] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에서는 공간 콘텐츠를 포함하는 접근법에 비해 이미지의 공간 콘텐츠가 추가되고 회전이 효율적으로 처리된다. 원의 히스토그램을 구성하여 공간 정보를 추가하고 동심원을 사용하여 회전을 처리한다. 또한, 가중치를 적용하여 시각적 단어의 히스토그램 형태로 이미지를 표현한다. 벤치마크 데이터 세트의 광범위한 평가와 최근 분류 모델과의 비교는 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법의 효과를 입증한다. 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법은 분류 정확도 측면에서 최첨단 방법보다 우수하다.

[0072] 우선, 동심 가중 원의 히스토그램을 기반으로 한 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 대한 설명 및 기본 표기법과 함께 BoVW 모델의 개요를 설명하기로 한다.

[0074] BoVW 모델

[0075] BoVW 모델은 문서를 비교하기 위해 광범위하게 사용되는 텍스트 기반 검색 시스템을 기반으로 한다. BoVW는 시각적 단어의 히스토그램을 사용하여 이미지를 나타낸다. 시각적 단어는 문서의 단어와 유사하며 작은 픽처 요소(이미지의 국부적인 영역)를 나타낸다. BoVW 이미지 히스토그램 표현은 순서가 없으며 2D 이미지의 시각적 단어의 위치를 무시하여 만들어진다.

[0076] BoVW에서, 제1 단계로서, 각각의 이미지 I는 수학식 1에서와 같은 이미지 서술자 세트에 의해 표현된다.

수학식 1

[0077] $$Im = \{d_1, d_2, d_3, ..., d_I\}$$

[0078] 여기서 di는 이미지 패치의 색상, 모양 또는 텍스처 서술이고, I는 총 이미지 서술자 수를 나타낸다.

[0079] 시각적 어휘(v)를 생성하기 위해, 수학식 2와 같이, 감독되지 않은 클러스터링 기법인 K-평균(K-means)이 추출된 서술자)에 적용된다.

수학식 2

[0080] $$v = \{w_1, w_2, w_3, ..., w_K\}$$

[0081] 여기서 K는 클러스터 또는 시각적 단어의 사전 정의된 수이다.

[0082] 시각적 단어와 가장 가까운 서술자의 매핑은 수학식 3에 따라 수행된다.

수학식 3

[0083] $$w(d_j) = \underset{w \in v}{argmin} Dist(w, d_j)$$

[0084] 여기서 w(dj)는 j번째 서술자에 할당된 시각적 단어를 나타내고, Dist(w, dj)는 서술자 dj와 시각적 단어 w 사이의 거리를 나타낸다.

[0085] 클러스터링은 고차원 특징 공간을 줄인다. 그러므로 클러스터링의 결과로 각 서술자는 시각적 단어에 매핑되고, 이미지의 최종 표현은 시각적 단어의 히스토그램이 된다. 히스토그램에서의 빈(bin)의 카운트(count)는 어휘 내의 시각적 단어의 수(즉, K)와 동일하다. 각 빈이 어휘에서 시각적 단어 wi를 나타내는 경우, 수학식 4와 같이 된다.

수학식 4

[0086] $$bin_i = card(D_i) \ \ where \ \ D_i = \{d_j, j \in 1, ..., n \mid w(d_j) = w_i\}$$

[0087] Di는 이미지에서 특정 시각적 단어 wi에 속하는 모든 서술자의 집합이다. 집합 Di의 기수(카디널리티)는 집합의 요소 수를 제공하는 Card(Di)에 의해 제공된다. 이것은 최종 표현을 얻기 위해 이미지의 모든 시각적 단어에 대해 반복된다. 하지만, 이렇게 생성된 히스토그램은 관심 지점의 공간 정보를 보유하지 않는다.

[0089] **동심 가중 원 히스토그램(CWCH: Concentric Weighted Circles Histogram)**

[0090] 이미지의 공간 콘텐츠를 추가하고 회전 불변성을 얻기 위하여 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에서는 동심원의 히스토그램을 생성한다. 시각적 단어의 히스토그램 형태로 이미지를 나타내기 위해 가중치 방식이 적용된다. 이미지는 k번째 레벨이 (k+1)개의 세그먼트를 갖는 동심원 방식으로 서로 다른 레벨의 세그먼

트로 분할된다. 추출된 각 세그먼트는 시각적 단어의 히스토그램으로 표시된다.

[0091] 한편, 도 2는 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법의 개념도이다.

[0092] 도 2를 참조하면, 각 클래스의 트레이닝 이미지들에서 고밀도 특징들이 추출되고(S200), 상기 추출된 특징들에 기반하여 시각적 단어들이 생성되어 시각적 어휘가 구성된다(S504). 또한, 동심 가중 원 히스토그램 표현(S206)에 기반하여 각 클래스의 트레이닝 이미지들의 시각적 단어들에 대한 히스토그램이 생성된 후 각 레벨의 히스토그램이 연결된다(S208).

[0093] 테스트 이미지에서 고밀도 특징들이 추출되고(S202), 상기 추출된 특징들에 기반하여 시각적 단어들이 생성되며, 동심 가중 원 히스토그램 표현(S206)에 기반하여 테스트 이미지의 시각적 단어에 대한 히스토그램이 생성된 후 각 레벨의 히스토그램이 연결된다(S210).

[0094] 테스트 이미지에 대한 히스토그램은 각 클래스의 트레이닝 이미지들에 대한 히스토그램과 비교되어, 분류되고(S212), 테스트 이미지의 카테고리가 검출된다(S214).

[0095] 도 3 내지 도 6을 참조하여, 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에 대해 더 상세히 설명하기로 한다.

[0096] 도 6은 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법이 적용되는 예시적인 장치의 블록도이다.

[0097] 도 6에 도시된 장치는, 각 클래스의 트레이닝 이미지들에서 각 클래스의 트레이닝 이미지들에 대한 최종 이미지 표현을 획득하기 위한 트레이닝 이미지 표현 획득부(602), 상기 트레이닝 이미지들에 대한 최종 이미지 표현에 기반하여 분류 모델을 생성하기 위한 분류 모델 생성부(606), 테스트 이미지에 대한 최종 이미지 표현을 획득하기 위한 테스트 이미지 표현 획득부(600), 및 테스트 이미지에 대한 최종 이미지 표현을 분류 모델에 기반하여 분류하기 위한 분류부(604)를 포함한다.

[0098] 도 3을 참조하면, 단계 S300에서, 트레이닝 이미지 표현 획득부(602)가, 각각의 클래스의 복수의 트레이닝 이미지에 대한 시각적 어휘를 구성하고 각각의 클래스의 트레이닝 이미지들에 대한 시각적 단어의 히스토그램을 생성하여 각각의 클래스의 이미지에 대한 최종 이미지 표현을 획득한다.

[0099] 단계 S302에서, 분류 모델 생성부(606)가, 상기 각각의 클래스의 이미지에 대한 최종 이미지 표현에 기반하여 분류 모델을 생성한다.

[0100] 단계 S304에서, 테스트 이미지 표현 획득부(600)가, 테스트 이미지에 대한 시각적 단어의 히스토그램을 생성하여 테스트 이미지에 대한 최종 이미지 표현을 획득한다.

[0101] 단계 S306에서, 분류부(604)가, 상기 테스트 이미지에 대한 최종 이미지 표현을 상기 분류 모델에 기반하여 분류한다.

[0102] 한편, 상기 단계 S300은, 트레이닝 이미지들을 그레일스케일 이미지로 변환하는 단계(단계 S400), 그레이스케일 이미지에서 고밀도 특징들을 추출하고 특징 서술자들을 생성하는 단계(단계 S402), 특징 서술자들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득하여 시각적 어휘를 구성하는 단계(단계 S404), k번째 레벨이 (k+1)개의 세그먼트들을 갖도록 트레이닝 이미지들을 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계(단계 S406), 각 레벨에 상이한 가중치를 적용하여 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성하는 단계(단계 S408), 및 생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 각 클래스의 트레이닝 이미지들에 대한 최종 이미지 표현을 획득하는 단계(단계 S410)를 포함한다.

[0103] 또한, 상기 단계 S304는, 테스트 이미지를 그레일스케일 이미지로 변환하는 단계(단계 S500), 그레이스케일 이미지에서 고밀도 특징들을 추출하고 특징 서술자들을 생성하는 단계(단계 S502), 특징 서술자들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득하는 단계(단계 S504), k번째 레벨이 (k+1)개의 세그먼트들을 갖도록 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계(단계 S506), 각 레벨에 상이한 가중치를 적용하여 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성하는 단계(단계 S508), 및 생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 각 클래스의 테스트 이미지에 대한 최종 이미지 표현을 획득하는 단계(단계 S510)를 포함한다.

[0104] 상기 단계 S406에서 상기 트레이닝 이미지들을 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계, 및 상기 단계 S506에서 상기 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할하는 단계에서, 상기 트레이닝 이미지들과 상기 테스트 이미지는, k번째 레벨이 (k+1)개의 세그먼트를 갖도록 동심원

방식으로 서로 다른 레벨의 세그먼트로 분할된다.

[0105] 상기 동심원의 k번째 레벨의 반경(r_k)은,

수학식 5

$$r_k = \frac{k}{L} min\{c_x, c_y\}$$

[0106]

[0107] 상기에서 c_x, c_y는,

수학식 6

$$c_x = \frac{1}{|I|}\sum_{i=1}^{|I|} x_i, \quad c_y = \frac{1}{|I|}\sum_{i=1}^{|I|} y_i$$

[0108]

[0109] 상기에서 I는 R×C의 크기를 갖는 이미지를 나타내며, $I = \{(x_i, y_i) \mid 1 \le x_i \le C, 1 \le y_i \le R\}$ 이고, $c = (c_x, c_y)$는 이미지 I의 중심을 나타내며, L은 레벨의 수를 나타내고, 가장 작은 원의 중심은 r_1이며, $|I|$는 이미지 I 내의 요소들의 수이고, k는 0 이상인 정수이며, R과 C는 0 이상인 정수이다.

[0110] 또한, 상기 단계 S408 및 상기 단계 S508에서 각 레벨에 상이한 가중치를 적용하여 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성할 때, 레벨이 높을수록 더 큰 가중치를 할당하여 히스토그램을 생성한다.

[0111] 또한, 상기 단계 S306은, 상기 분류부(604)가, 상기 각 클래스의 트레이닝 이미지의 히스토그램과 상기 테스트 이미지의 히스토그램의 인터섹션(intersection)에 기반하여 각 클래스의 트레이닝 이미지와 테스트 이미지 간의 유사도를 계산하고, 계산된 유사도에 기반하여 상기 테스트 이미지를 분류하는 단계를 포함한다.

[0112] L개의 레벨을 갖는 2개의 이미지의 유사도는,

수학식 7

$$S = \sum_{k=0}^{L} w_k N_k$$

[0113]

[0114] 상기에서 $w_k = 1/2^{L-k}$ 으로서 각 레벨에 할당되는 가중치이며, N_k는 두 연속적인 레벨들의 히스토그램 인터섹션의 차를 나타낸다.

[0115] 두 이미지의 히스토그램 인터섹션은,

수학식 8

[0116]
$$HI(P,Q) = \sum_{i=1}^{c} min(P_i, Q_i)$$

[0117] 상기에서 P와 Q는 각각 c 빈(bin)을 갖는 상기 각 클래스의 트레이닝 이미지와 상기 테스트 이미지의 히스토그램을 나타내고, P_i는 P의 i번째 빈의 카운트(count)를 나타내며, Q_i는 Q의 i번째 빈의 카운트(count)를 나타낸다.

[0118] 일반적으로 사용되는 세 가지 유형의 특징 추출 기법, 즉 고밀도 샘플링, 고밀도 관심 지점 및 관심 지점이 있다. 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에서는 고밀도 관심 지점을 선택했는데, 그것은 하이브리드 접근 방식이며 관심 지점과 고밀도 샘플링의 두 가지 장점을 결합한다. 특징 추출을 위해, 모든 이미지는 그레이스케일로 변환되고, 멀티 스케일 및 스텝 사이즈 8의 고밀도 SIFT가 특징 서술자를 생성하는데 사용된다. 이 서술자들은 시각적 어휘를 생성하기 위해 k-평균(means) 클러스터링 기법을 사용하여 비슷한 그룹으로 클러스터링된다. k-평균(means) 클러스터링의 감독되지 않은 특성으로 인해 실험을 무작위로 선택한 트레이닝 및 테스트 이미지의 하위 집합을 사용하여 10번 반복한다. 평균값은 테이블과 그래프로 보고된다. 시각적 어휘의 크기는 시스템 성능에 중요한 영향을 미치는 주요 매개 변수이다. 이 성능은 어휘 크기에 직접적으로 비례하며 큰 어휘 크기는 과다 적합 경향이 있다. 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에서는 이미지 표현의 최적 성능을 결정하기 위해 다양한 크기의 시각적 어휘로 실험을 수행한다.

[0119] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에서 이미지 공간은 동심원으로 분할되고, 적절한 가중치가 각 레벨에 할당되며, 각 레벨의 히스토그램이 연결되어 최종 표현을 얻는다.

[0120] 트레이닝 세트로부터 얻어진 이미지 표현은 분류부(604)를 트레이닝시키는데 사용된다; 평가는 테스트 세트로 수행된다. 분류를 위해, 감독 학습 방법으로 특징 지어지는 서포트 벡터 머신(SVM: Support Vector Machines)이 사용된다. SVM은 비선형 결정 경계를 생성하는 특징을 획득하고 커널 방법을 사용하여 고차원 특징 공간에서 내적(dot product)을 계산한다. 포지티브 및 네거티브 트레이닝 이미지가 주어지면, 테스트 이미지를 각각의 객체 클래스로 분류하는 것이 목적이다. 제안된 프레임 워크를 사용하여 구성된 히스토그램을 정규화하고 SVM Hellinger Kernel[32]을 정규화된 히스토그램에 적용한다. SVM Hellinger 커널은 계산 비용이 적기 때문에 선택된다. 정규화 매개 변수 C에 대한 최적 값을 결정하기 위해 10-겹(fold) 교차 검증이 트레이닝 데이터 세트에 적용된다.

[0122] **실험 및 결과**

[0123] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법의 유효성을 평가하기 위해 15-장면 이미지 벤치마크를 사용하여 자연 장면 분류에 적용되었다. 이는 지금까지 문헌에서 가장 어렵고 광범위하게 사용되는 데이터 세트이다. 잘 알려진 UC Merced 토지 사용 데이터 세트를 사용하여 원격 센싱 이미지 장면 분류를 위한 실험도 수행된다. 모든 데이터 세트에 대해 제안된 이미지 표현을 생성하기 위해 동일한 실험 설정을 따랐다. 회전 불변성에 대한 실험을 수행하기 위해, 관련 연구와 일치하는 각 데이터 세트에서 회전 데이터 세트가 생성된다.

[0124] 여기에서는 실험 결과 및 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법을 평가하는 데 사용된 이미지 벤치마크에 대한 개요를 제공한다.

[0125] 실험에 사용된 첫 번째 데이터 세트는 실내 및 실외 장면 카테고리로 구성된 잘 알려진 15-장면 이미지 데이터 세트이다. 도 7은 카테고리별 이미지의 이름과 개수를 나타내는 15-장면 데이터 세트의 예시적인 이미지를 보여준다. 처음에 Oliva와 Torralba[25]에 의해 8개의 카테고리들이 제공되었고, Li와 perona[13]이 5개의 클래스를 추가하였으며, 나머지는 [18]에 의해 수집되었다. 이미지의 주된 소스는 인터넷, 개인 사진 및 COREL 컬렉션이다. 이는 까다로운 데이터 세트이며 평균 이미지 크기가 300×250 픽셀인 총 4485개의 이미지를 가지고 있다. 그것은 지금까지 가장 널리 사용되는 데이터 세트이다.

[0126] 공정한 비교를 보장하기 위해 테스트 샘플 및 트레이닝 샘플은 최신 방법[18, 15]에 따라 선택된다. 트레이닝 세트는 무작위로 선택된 100개의 이미지로 구성되며 나머지 이미지는 테스트에 사용된다. 따라서 총 트레이닝

이미지 수는 1500개이며 2985개의 테스트 이미지가 있다. 본 실험에서 클래스 번호는 'Bedroom=1', 'Calsubrub=2', 'Industrial=3', 'Kitchen=4', 'Living room=5', 'MITcoast=6', 'MITforest=7, 'MIThighway=8', 'MITinsidecity=9', 'MITmountain=10', 'MITopen country=11', 'MITstreet=12', 'MITtallbuilding=13', 'PARoffice=14' 및 'Store=15'이다.

[0127] 정확한 특징 표현을 위한 최적의 성능을 얻기 위하여, 실험은 15-장면 및 회전된 데이터 세트에서 다양한 크기의 시각적 어휘를 가지고 수행된다. 도 8은 두 개의 데이터 세트에서 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법의 견고함을 보여주고 다양한 크기의 어휘에 대한 성능을 그래프로 나타낸 것이다. 두 데이터 세트의 경우, 크기가 400인 어휘에 대해 CWCH에 대한 최적 성능을 얻는다. 결과 특징 벡터의 차원(dimensions)은 2400이다.

[0128] 표 1은 제안된 CWCH와 BoVW 모델에 공간적 상황을 통합한 최근의 방법을 비교한 것이다. 실험 결과는 최신 방법에 비해 본 발명의 접근 방법의 견고함을 입증한다.

표 1

[0129]

Algorithms	Feature Dimensionality	Accuracy
PIWAH [16]	1800	76%
SPM [18]	4200	81.1%
Zang *et al.* [34]	3717	81.5%
SPS_{ad}+ [17]	13200	83.7%
Karmakar *et al.* [15]	4200	84.20%
LGF [39]	X	85.8%
CWCH	2400	**88.04%**

[0130] 표 1에서 본 발명의 방법이 최신 방법에 비해 가장 높은 정확도를 제공한다는 것이 분명하다. PIWAH[16]의 차원은 CWCH에 비해 낮지만 본 발명의 방법은 상대적으로 간단하고 기존 방법에 비해 12.04% 더 높은 정확도를 달성한다. CWCH의 차원은 정확도를 크게 떨어뜨리지 않으면서 일부 차원 축소 기술을 적용하여 더 줄일 수 있다. 본 발명의 방법은 공간 정보를 BoVW 모델에 통합하는 선구적 연구인 SPM[18]에 비해 6.94% 더 높은 정확도를 달성한다.

[0131] 최근의 다른 논문[34]에서, Zang et al.은 오브젝트 뱅크(OB: Object Bank)의 크기를 줄여 81.5%의 정확도를 달성했고, CWCH는 6.54% 더 높은 정확도로 낮은 차원의 히스토그램 표현을 생성했다. 직사각형 링의 접근 방법[15]에 비해 본 발명의 방법은 그들의 연구에서 보고되는 더 적은 차원으로 3.84%의 더 높은 정확도를 달성한다. 본 발명의 방법은 88.04%의 정확도로 정확도와 차원의 측면에서 최신 방법을 능가한다.

[0132] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법에서의 이미지 표현의 지속 가능한 성능을 입증하기 위해, 도 9에 최신 방법 [39]과의 클래스별 비교를 수행한 결과를 도시한다. Zou et al.[39]은 SPM[18]을 그들의 구현에 통합하여 공간 내용을 고려한다.

[0133] 15-장면 이미지 데이터 세트에 대한 평균 혼동 행렬이 도 10에 도시된다. 대각선 값은 각 클래스의 정밀도 정규화 백분율을 보여준다.

[0135] **UC 머세드(Merced) 토지-이용(UCM) 이미지 데이터 세트**

[0136] UCM 이미지 데이터 세트는 Yang과 Newsam[33]에 의해 만들어졌으며 21개의 토지 이용 이미지 장면 카테고리로 구성된다. 이 데이터 세트는 큰 지형을 가지고 있으며 미국 지질 조사(USGS) 내셔널 지도에서 이미지를 다운로드했다. 각 클래스는 평균 크기가 256×256 픽셀인 100개의 이미지로 구성된다. 도 11은 클래스 레이블과 각 범주의 총 이미지 수와 함께 각 클래스의 예시적인 이미지를 보여준다.

[0137] 실험에 사용된 클래스 번호는 '농경지=1', '비행기=2', '야구장=3', '해변=4', '건물=5', '수풀 지역=6', '밀도가 높은 주거지=7', '숲=8', '고속도로=9', '골프 코스=10', '항구=11', '교차로=12', '중형 주거지=13', '이동 주택 공원=14, '고가 도로=15', '주차장=16', '강=17', '활주로=18', '희박한 주거지=19', '저장 탱크=20', '테니스 코트=21'이다. 트레이닝과 테스트 세트는 비교를 위해 관련 연구[39, 9, 33]에 따라 선택된다. 트레이

닝을 위해 80개의 이미지가 각 클래스에서 무작위로 선택되고 나머지는 테스트에 사용된다.

[0138] UCM을 위한 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법과 다양한 크기의 어휘에 대한 회전된 데이터 세트의 성능 비교가 도 12에 도시되어 있다. 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법은 이미지 회전에 견고하다는 것을 알 수 있다. 양쪽 데이터 세트에 제안된 접근법을 사용하여 크기가 200인 어휘에 대해 최적의 성능을 얻었고, 1200차원의 특징 벡터가 생성되었다.

[0139] 표 2에서 본 발명의 제안된 CWCH와 최신 연구 결과를 비교한다. CWCH는 본 발명의 접근법과 일치하는 방법뿐만 아니라 공간적 상황과의 융합에서 특징 표현을 향상시키는 다른 연구를 능가하고, 또한 중간 표현을 향상시키는 연구를 훨씬 능가한다는 것이 분명하다.

표 2

[0140]

Algorithms	Accuracy
SPM [38]	82.3%
BoCF [11]	84.32%
CCM-BOVW [37]	86.64%
$MSCLBP_1$ [10]	90.6%
SOS [21]	94.33%
LGF [39]	95.48%
Lu *et al.* [14]	95.71%
$salM^3LBP$-CLM [9]	95.75%
LGFBOVW [38]	96.88%
ResNet50 [27]	98.5%
Evolved Sugeno [28]	99.33%
CWCH	99.4%

[0141] [37]에서, Zhao et al.은 동심원 구조의 다중 스케일 BoVW 모델을 사용하여 공간 콘텍스트를 통합했다. 그들은 SIFT, 컬러 모멘트 및 LBP와 같은 여러 특징을 사용하여 표현력을 향상시켰지만 본 발명의 방법은 그들의 접근 방식에 비해 12.76% 더 높은 정확도를 제공한다. Scott et al.[28]은 UCM 데이터 세트에 가장 적합한 분류를 제공한다. CWCH 표현은 특징 추출[39], 공간 콘텍스트[37], 중간 특징 표현의 적용[21] 및 심층 학습 기술 [28]을 기반으로 최신 기술에 경쟁력 있는 성능을 제공한다.

[0142] CWCH와 CCM-BOVW[37]의 클래스별 비교는 도 13에 나와있다. CCM-BOVW[37]에서 공간 정보는 동심원 기반 접근법을 사용하여 통합되었고, 그들의 접근법은 비행기, 야구장, 골프 코스 및 저장 탱크와 같은 방향에 민감한, 클래스들에 대해서만 좋은 것으로 나타난다[37]. 반면에 본 발명의 CWCH 모델은 단순한 패턴을 가지고 있는 카테고리에 큰 영향을 받지 않았고, 숲, 강, 농경지 및 수풀 지역과 같은 방향으로 고통받지 않는다. 본 발명의 제안된 접근 방식은 모든 클래스에서 상당한 개선을 보여준다.

[0143] UCM 영상 데이터 세트에 대한 혼동 행렬을 도 14에 나타내었다. 대각선 값은 각 클래스의 정밀도 정규화 백분율을 보여준다.

[0145] **복잡성 성능**

[0146] 알고리즘은 인텔(®) 코어 i7(6 세대) 2.59GHz CPU, 8GB RAM 및 Windows-10 운영 체제를 사용하여 계산된다. 본 발명의 제안된 알고리즘은 MATLAB에서 구현되고 시각적 어휘(코드북)는 트레이닝 데이터 세트를 사용하여 오프라인으로 구성되며 테스트 데이터 세트를 사용하여 런타임에 테스트된다. 테스트 데이터 세트에 대한 CWCH 및 BoVW를 위한 특징 추출에서 이미지 분류까지 필요한 평균 CPU 시간(초)은 표 3에 나와 있다.

표 3

Dataset	BoVW		CWCH	
	Accuracy	Time	Accuracy	Time
15-scene	84.11%	74.40s	88.04%	118.67s
UCM	97.62%	9.93s	99.4%	21.09s

[0147]

[0148] 본 발명의 일 실시예에 의한 지능형 이미지 분류 방법은 공간 정보가 없는 BoVW에 비해 두 데이터 세트 모두에서 더 나은 분류율을 제공한다. 그러나 BoVW는 공간 정보 추출 단계를 포함하지 않으므로 계산상 효율적인 것으로 보인다. 이에 비해 본 발명의 접근 방식은 15-장면 이미지 데이터 세트에서 3.93% 더 높은 정확도를 제공하고 UCM 이미지 데이터 세트에서 1.78% 더 높은 정확도를 제공한다. CCM-BOVW[37]의 계산 시간은 보고되지 않았지만 여러 가지 특징 추출 단계가 포함되며 특징 추출의 일부로 여러 개의 다중 해상도 이미지를 생성한다. 본 발명의 접근법은 단일 특징 추출 단계로 구성되어 단순하고, 더 나은 분류 정확도를 제공한다.

[0150] **결론**

[0151] 최근 기술의 발전으로 멀티미디어 콘텐츠의 복잡성과 볼륨이 크게 증가하고 있다. 이로 인해 매우 효과적인 이미지 분류 시스템에 대한 요구가 높아졌다. 본 발명에서는 BoVW 모델의 분류 기반 프레임워크와 관련된 두 가지 문제를 다루었다. 표준 BoVW 모델에는 공간 정보가 없고, 시각적 단어의 히스토그램을 만들기 위해 이미지를 셀로 나누는 방식을 기반으로 한 접근법은 회전 및 시점 변경을 허용하지 않는다. 제안된 접근법은 동심원을 사용하여 시점의 변화, 회전 및 공간 정보의 계산을 처리함으로써 시각적 단어의 히스토그램을 구성한다. 가중치 방식이 이미지 표현에 적용되고 성능이 동심원의 두 레벨을 사용하여 평가된다. 가중 방식의 적용은 과다 적합(over-fitting) 문제를 해결하는 저 차원 공간에서 이미지를 표현하는 것을 가능하게 한다.

[0152] 제안된 히스토그램 표현은 간단하고 견고하며 회전 및 시점의 변화를 처리할 수 있다. 두 개의 표준 이미지 벤치마크가 제안된 접근법의 성능 평가에 사용된다. 스케일링(scaling)과 변환(translation)과 같은 다른 변형에 대해서도 관대한 통일된 접근법을 만들어 이 연구는 확장될 수 있다. 본 발명에서 제안된 프레임워크는 방대한 멀티미디어 콘텐츠의 이미지 분류에 적용될 수 있다. 심층 회선 신경망을 사용하여 ImageNet과 Flicker와 같은 더 큰 데이터 세트에 대해 제안된 방법이 평가될 수 있다. 이 연구를 직접 확장하면 색상 및 모양과 같은 다른 공간적인 큐를 통합하거나 이를 분류 성능을 향상시키는 보완적인 상대 공간 형상 추출 방법과 결합할 수 있다.

[0153] 이상 본 발명을 구체적인 실시예를 통하여 상세하게 설명하였으나, 이는 본 발명을 구체적으로 설명하기 위한 것으로, 본 발명은 이에 한정되지 않으며, 본 발명의 기술적 사상 내에서 당 분야의 통상의 지식을 가진 자에 의해 그 변형이나 개량이 가능함은 명백하다고 할 것이다.

[0154] 본 발명의 단순한 변형 내지 변경은 모두 본 발명의 영역에 속하는 것으로, 본 발명의 구체적인 보호 범위는 첨부된 청구범위에 의하여 명확해질 것이다.

부호의 설명

[0155] 600 : 테스트 이미지 표현 획득부

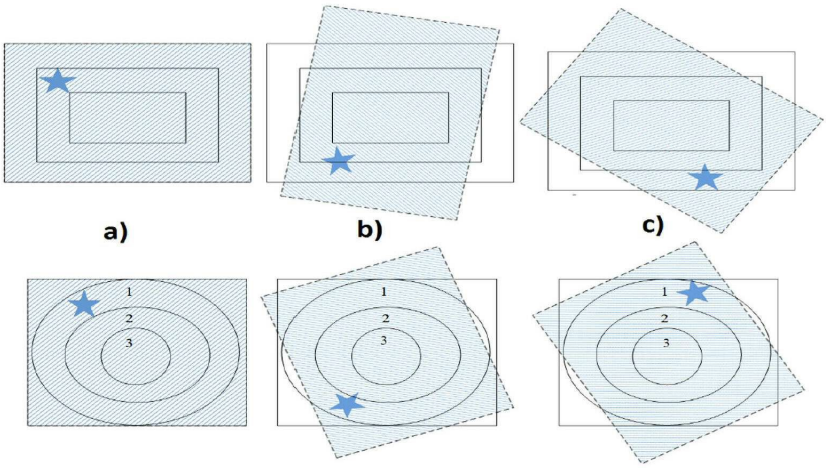
602 : 트레이닝 이미지 표현 획득부

604 : 분류부

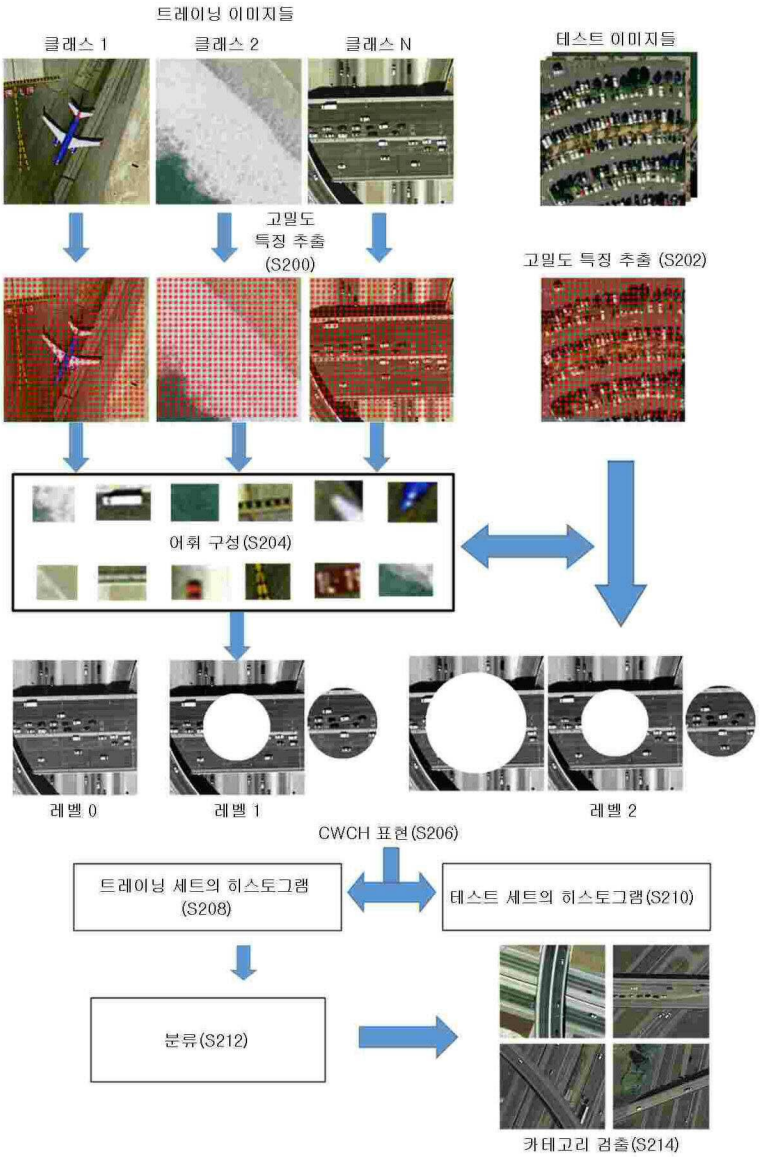
606 : 분류 모델 생성부

도면

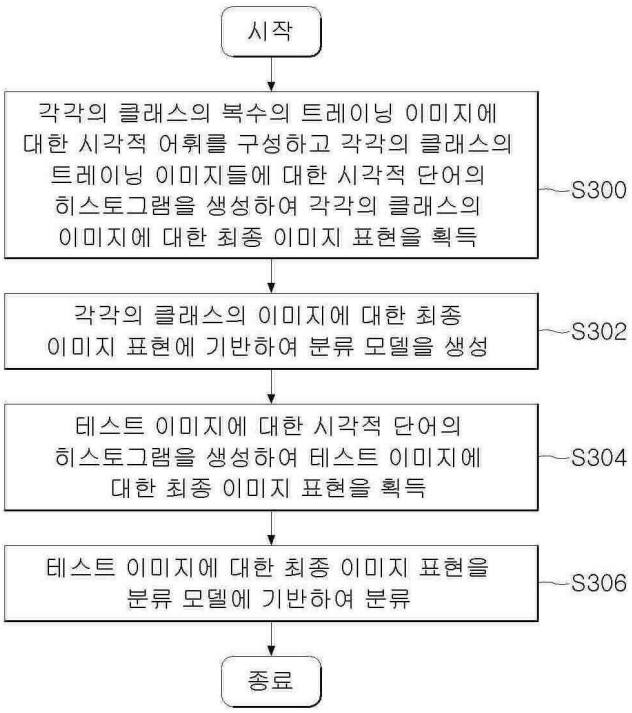
도면1


a)
b)
c)
1
2
3

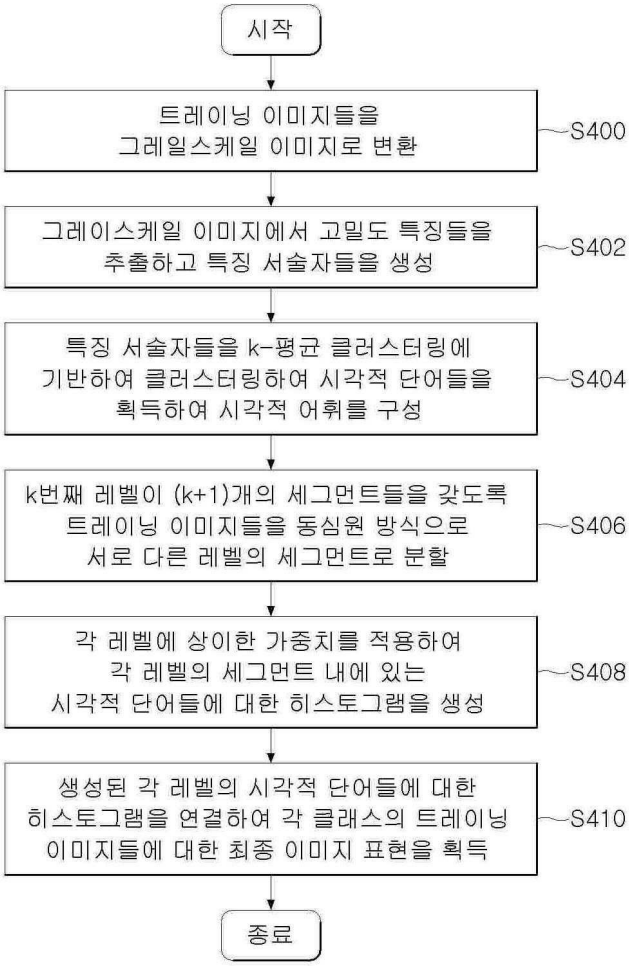
도면2


트레이닝 이미지들
클래스 1
클래스 2
클래스 N
테스트 이미지들
고밀도
특징 추출
(S200)
고밀도 특징 추출 (S202)
어휘 구성 (S204)
레벨 0
레벨 1
레벨 2
CWCH 표현 (S206)
트레이닝 세트의 히스토그램
(S208)
테스트 세트의 히스토그램 (S210)
분류 (S212)
카테고리 검출 (S214)

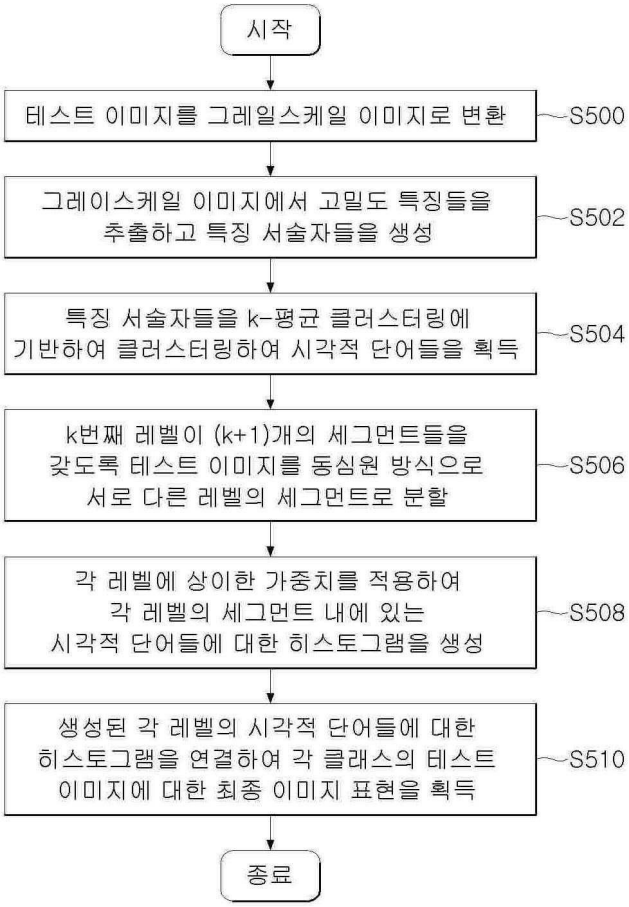
도면3


시작
각각의 클래스의 복수의 트레이닝 이미지에 대한 시각적 어휘를 구성하고 각각의 클래스의 트레이닝 이미지들에 대한 시각적 단어의 히스토그램을 생성하여 각각의 클래스의 이미지에 대한 최종 이미지 표현을 획득
S300
각각의 클래스의 이미지에 대한 최종 이미지 표현에 기반하여 분류 모델을 생성
S302
테스트 이미지에 대한 시각적 단어의 히스토그램을 생성하여 테스트 이미지에 대한 최종 이미지 표현을 획득
S304
테스트 이미지에 대한 최종 이미지 표현을 분류 모델에 기반하여 분류
S306
종료

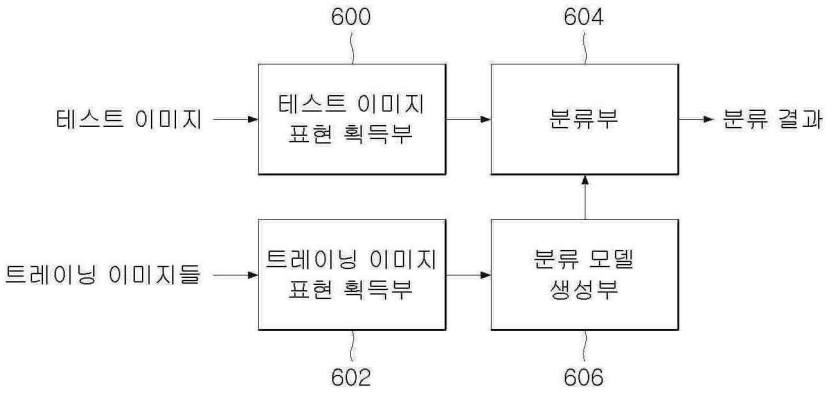
도면4


시작
트레이닝 이미지들을
그레일스케일 이미지로 변환
S400
그레이스케일 이미지에서 고밀도 특징들을
추출하고 특징 서술자들을 생성
S402
특징 서술자들을 k-평균 클러스터링에
기반하여 클러스터링하여 시각적 단어들을
획득하여 시각적 어휘를 구성
S404
k번째 레벨이 (k+1)개의 세그먼트들을 갖도록
트레이닝 이미지들을 동심원 방식으로
서로 다른 레벨의 세그먼트로 분할
S406
각 레벨에 상이한 가중치를 적용하여
각 레벨의 세그먼트 내에 있는
시각적 단어들에 대한 히스토그램을 생성
S408
생성된 각 레벨의 시각적 단어들에 대한
히스토그램을 연결하여 각 클래스의 트레이닝
이미지들에 대한 최종 이미지 표현을 획득
S410
종료

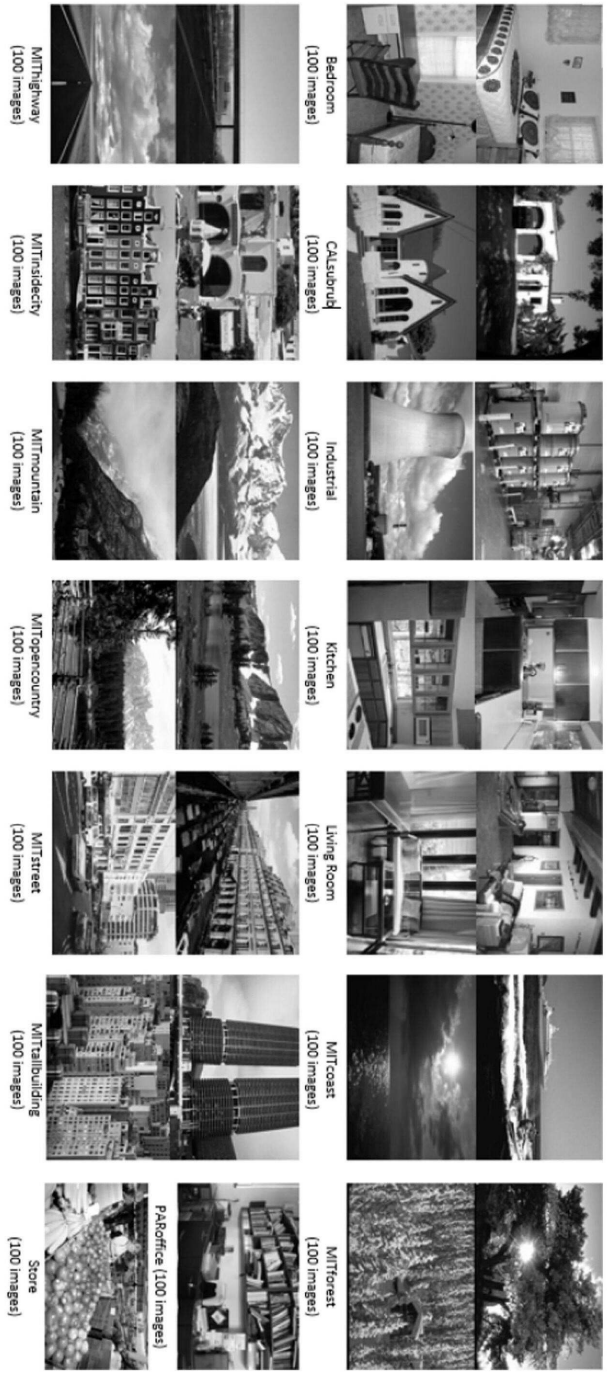
도면5


시작
테스트 이미지를 그레일스케일 이미지로 변환
S500
그레이스케일 이미지에서 고밀도 특징들을 추출하고 특징 서술자들을 생성
S502
특징 서술자들을 k-평균 클러스터링에 기반하여 클러스터링하여 시각적 단어들을 획득
S504
k번째 레벨이 (k+1)개의 세그먼트들을 갖도록 테스트 이미지를 동심원 방식으로 서로 다른 레벨의 세그먼트로 분할
S506
각 레벨에 상이한 가중치를 적용하여 각 레벨의 세그먼트 내에 있는 시각적 단어들에 대한 히스토그램을 생성
S508
생성된 각 레벨의 시각적 단어들에 대한 히스토그램을 연결하여 각 클래스의 테스트 이미지에 대한 최종 이미지 표현을 획득
S510
종료

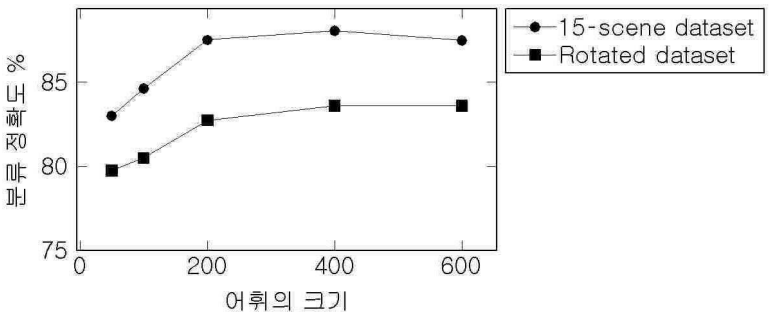
도면6


600
604
테스트 이미지
테스트 이미지 표현 획득부
분류부
분류 결과
트레이닝 이미지들
트레이닝 이미지 표현 획득부
분류 모델 생성부
602
606

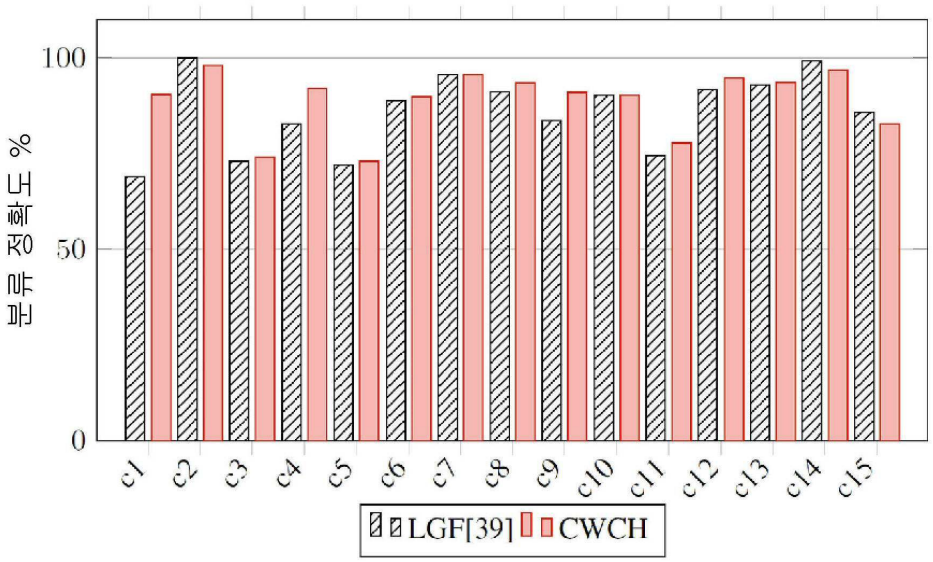
도면7


MITkighway (100 images)
Bedroom (100 images)
MITinsidecity (100 images)
CALsubrub (100 images)
MITmountain (100 images)
Industrial (100 images)
MITopencountry (100 images)
Kitchen (100 images)
MITstreet (100 images)
Living Room (100 images)
MITtallbuilding (100 images)
MITcoast (100 images)
Store (100 images)
PARoffice (100 images)
MITforest (100 images)

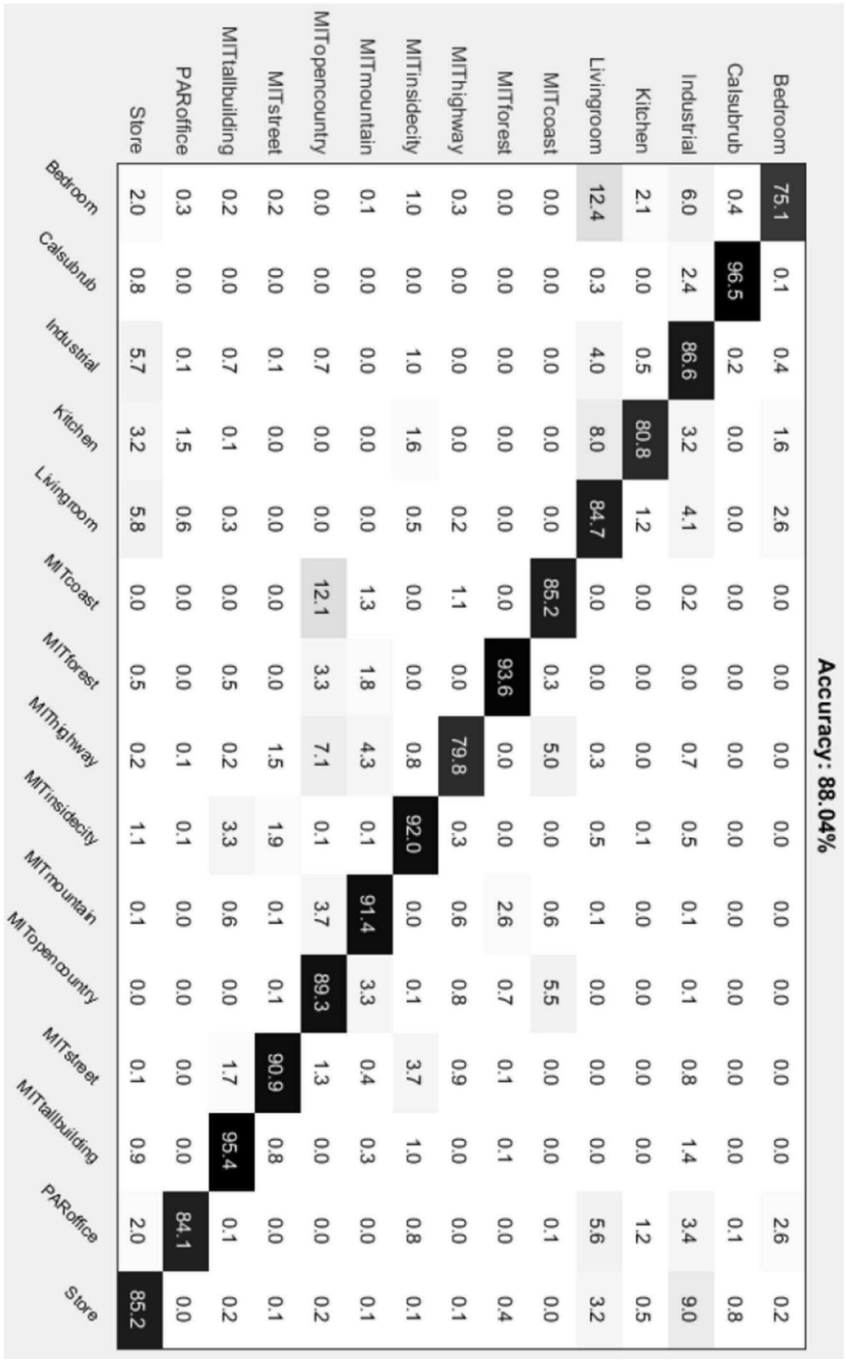
도면8


15-scene dataset
Rotated dataset
분류 정확도 %
85
80
75
0
200
400
600
어휘의 크기

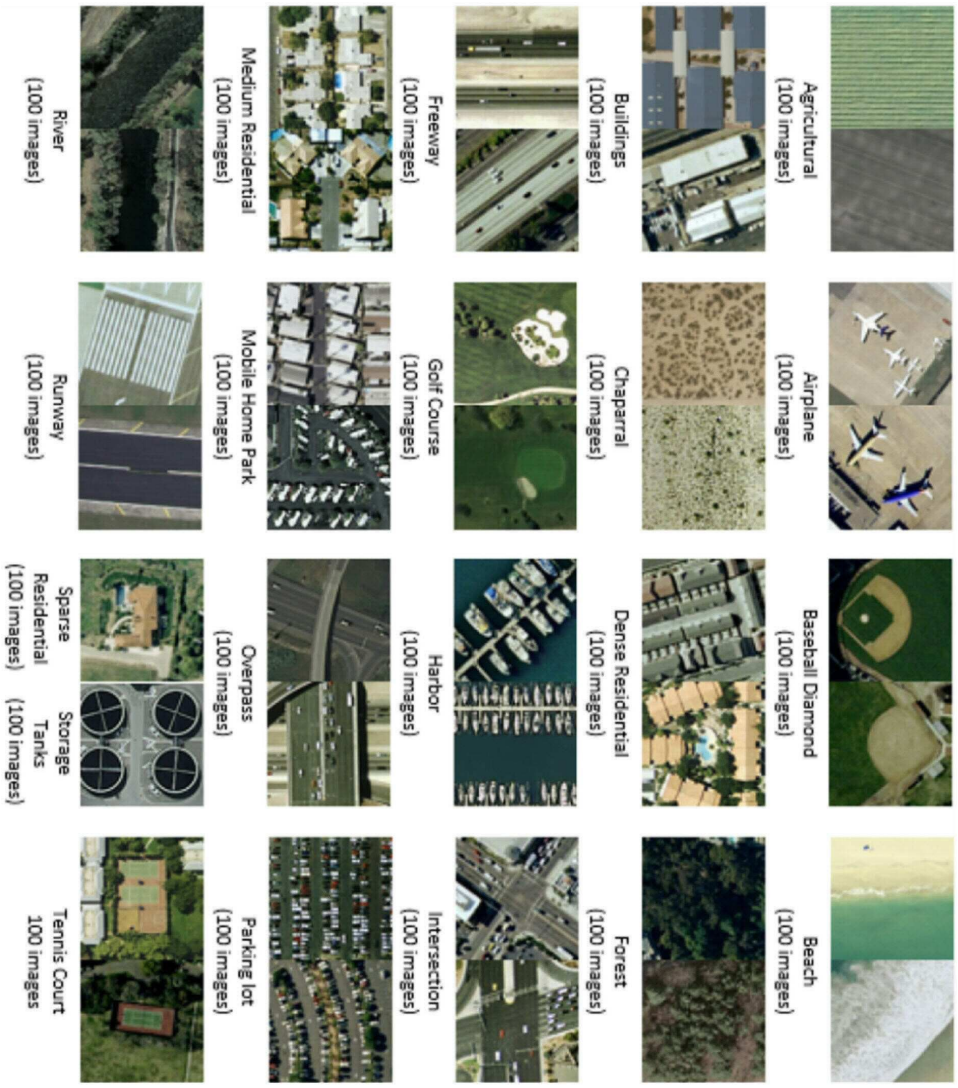
도면9


100
50
0
분류 정확도 %
c1
c2
c3
c4
c5
c6
c7
c8
c9
c10
c11
c12
c13
c14
c15
LGF[39]
CWCH

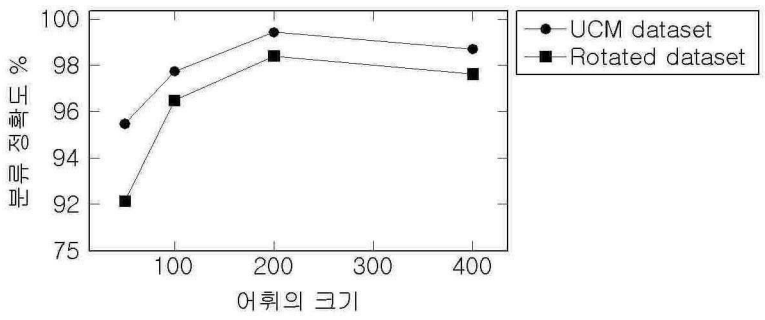
도면10


Accuracy: 88.04%
Bedroom
Calsubrub
Industrial
Kitchen
Livingroom
MITcoast
MITforest
MIThighway
MITinsidecity
MITmountain
MITopencountry
MITstreet
MITtallbuilding
PARoffice
Store

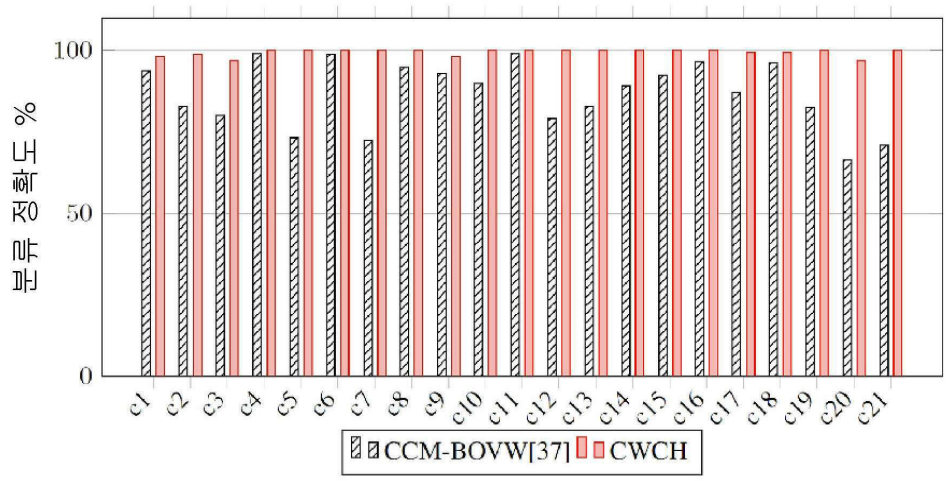
도면11


Agricultural (100 images)
Buildings (100 images)
Freeway (100 images)
Medium Residential (100 images)
River (100 images)
Airplane (100 images)
Chaparral (100 images)
Golf Course (100 images)
Mobile Home Park (100 images)
Runway (100 images)
Baseball Diamond (100 images)
Dense Residential (100 images)
Harbor (100 images)
Overpass (100 images)
Storage Tanks (100 images)
Sparse Residential (100 images)
Beach (100 images)
Forest (100 images)
Intersection (100 images)
Parking lot (100 images)
Tennis Court 100 images

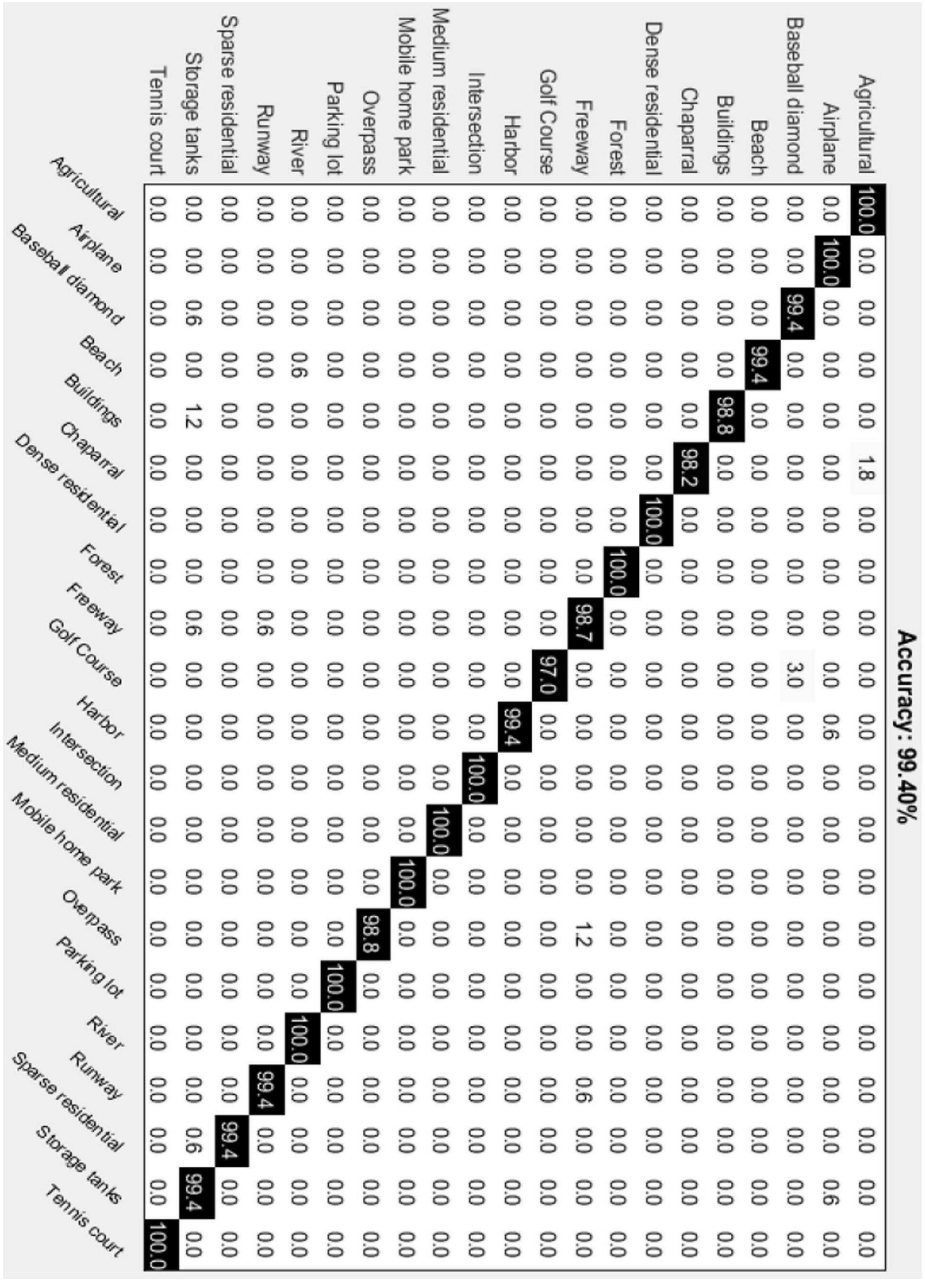
도면12


UCM dataset
Rotated dataset
분류 정확도 %
100
98
96
94
92
75
100
200
300
400
어휘의 크기

도면13


분류 정확도 %
100
50
0
c1
c2
c3
c4
c5
c6
c7
c8
c9
c10
c11
c12
c13
c14
c15
c16
c17
c18
c19
c20
c21
CCM-BOVW[37]
CWCH

도면14


Accuracy: 99.40%
Agricultural
Airplane
Baseball diamond
Beach
Buildings
Chaparral
Dense residential
Forest
Freeway
Golf Course
Harbor
Intersection
Medium residential
Mobile home park
Overpass
Parking lot
River
Runway
Sparse residential
Storage tanks
Tennis court