

하이퍼그래프 모델 기반의 장면 이미지 분류 기법

Hypergraph model based Scene Image Classification Method

최선욱 · 이종호[†]

Sun-Wook Choi, and Chong Ho Lee[†]

인하대학교 정보통신공학과

School of Information and Communication Engineering, Inha University

요 약

이미지를 각각의 카테고리로 분류하는 일은 컴퓨터 비전 분야의 중요한 문제 중 하나이다. 그러나 이미지에 존재하는 가변성, 모호성, 스케일 문제 등으로 인해 매우 도전적인 문제라고 할 수 있다. 본 논문에서는 장면 이미지를 구성하는 시멘틱 속성들의 고차원의 상호작용 관계를 고려 가능한 하이퍼그래프 기반의 모델링 기법을 제시하고 이를 장면 이미지 분류에 적용한다. 각 장면 카테고리에 준최적화된 하이퍼그래프를 생성하기 위해 확률 부분공간 기법에 기반을 둔 탐색기법을 제안하고, 이들 부분 공간 내에 속한 시멘틱 속성들의 발현량을 축약하기 위한 우도비 기반의 선형 변환 기법을 제안한다. 제안한 기법의 우수성을 검증하기 위한 실험을 통하여 제시한 기법을 통해 생성된 특징 벡터의 분별력이 기존의 기법들에서 사용된 특징 벡터들의 분별력보다 우수함을 보인다. 또한 제안한 기법을 장면 분류 데이터에 적용한 결과 기존의 기법들과 비교하여 경쟁력 있는 분류 성능을 보인다. 제안한 기법은 이미지 분류에서 일반적으로 사용 되는 기법인 BoW+SPM 모델과 비교하여 3~4%이상의 성능 향상을 보였다.

키워드 : 이미지 분류, 장면 분류, 확률적 부분집합 탐색 기법, 하이퍼그래프, 서포트 벡터 머신

Abstract

Image classification is an important problem in computer vision. However, it is a very challenging problem due to the variability, ambiguity and scale change that exists in images. In this paper, we propose a method of a hypergraph based modeling can consider the higher-order relationships of semantic attributes of a scene image and apply it to a scene image classification. In order to generate the hypergraph optimized for specific scene category, we propose a novel search method based on a probabilistic subspace method and also propose a method to aggregate the expression values of the member semantic attributes that belongs to the searched subsets based on a linear transformation method via likelihood based estimation. To verify the superiority of the proposed method, we showed that the discrimination power of the feature vector generated by the proposed method is better than existing methods through experiments. And also, in a scene classification experiment, the proposed method shows a competitive classification performance compared with the conventional methods.

Key Words : Image Classification, Scene Classification, Probabilistic Subspace Method, Hypergraph, SVM

1. 서 론

물체(object) 및 장면(scene) 등을 분류하는 작업을 포함한 이미지 분류 문제는 컴퓨터 비전 분야의 중요한 문제 중 하나이다[1,13,14]. 그러나 동일 카테고리에 속하는 이미지 들임에도 불구하고 존재하는 가변성, 모호성, 스케일 변화 등의 문제로 인해 매우 어려운 문제라고 할 수 있다[1].

이를 위해 그동안 색상 분포나, 텍스트 기반의 히스토그램 등 저수준의 특징(low-level feature)을 사용하는 기법들이 널리 사용되어 왔다. 특히 SIFT 등의 지역 기술자 등을 바탕으로 시각단어(visual word)들의 집합인 코드북을 생성하고, 이를 기반으로 히스토그램을 생성하여 분류에 사용하고, BoW(Bag-of-Words) 기법을 이용한 연구들이 최근까지 활발하게 진행 중에 있다[2]. 또한 중간 수준의 특징(intermediate-level feature)을 생성하여 분류에 이용하는 기법들도 제안 되었다. 이들은 토픽 모델을 기반으로, 저수준 특징들로부터 공통의 토픽을 추출 하여 이를 중간 수준의 특징으로 활용하고자 하였다[1,3]. 한편 이러한 저수준, 중간 수준의 특징들을 공간적 정보와 결합하여 이미지 분류 성능을 높이는 기법들이 제안되기도 하였다[1,11].

최근 보다 높은 수준의 문맥 정보(contextual information)를 분류에 활용하기 위해 이미지 내에 존재하는 시멘틱 속성(semantic attribute) 등을 이용하는 연구가 활발히 진행 중에 있다[4,5]. 이미지 상에서 나타나는 시멘틱 속성들은 매우 직관적인 것으로 장면을 구성하는 각 사물들

접수일자: 2013년 7월 4일

심사(수정)일자: 2013년 10월 18일

게재확정일자 : 2013년 10월 18일

[†] Corresponding author

본 논문은 인하대학교의 지원에 의하여 연구되었음.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

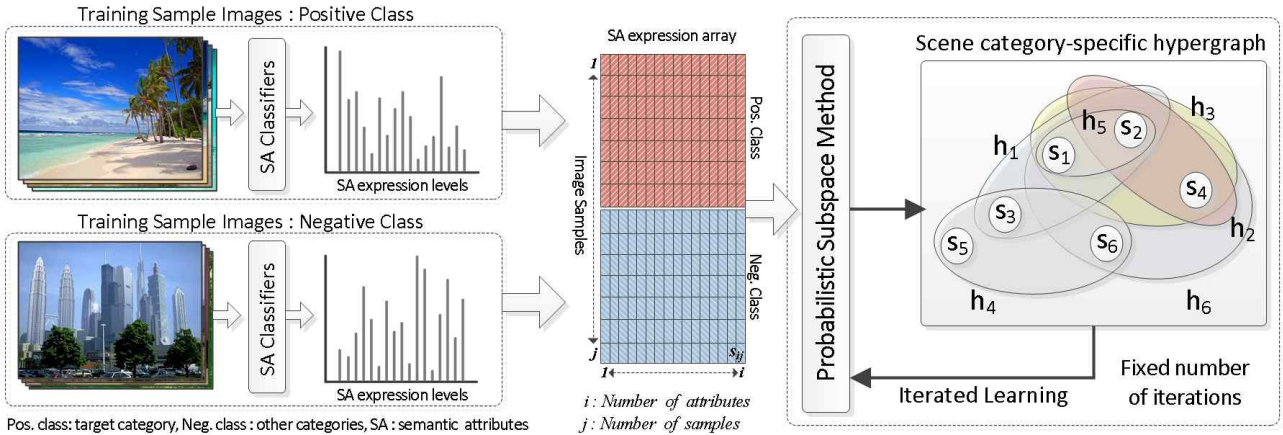


그림 1. 특정 장면 카테고리에 특화된 하이퍼그래프 생성 및 학습 과정
 Fig. 1. A scene category-specific hypergraph generation & learning

을 의미하거나 혹은 하나의 물체를 구성하는 특정한 부분 (구성 물질 혹은 특성)들을 의미하는 것이기도 하다. 이러한 시멘틱 속성들은 일반적으로 이미지 내에서 고차원의 상호작용을 하며 나타나는 특성들이 있는데, 특히 이러한 특성은 장면 이미지들에서 더욱 두드러지게 나타난다. 그러나 기존의 기법들은 이러한 시멘틱 속성들의 고차원의 상호작용을 고려하지 못하고 각 속성들을 개별적으로 다룬다는 한계가 있었다.

시멘틱 속성들 사이의 상호작용을 표현해내기 위해 일반적으로 고려할만한 기법은 그래프 기반의 모델링 기법이라 할 수 있다. 하지만 기존의 그래프 기반의 기법들은 노드와 노드의 단일 상호작용들의 결합으로만 나타낼 수 있다는 한계가 있기 때문에 시멘틱 속성들에서의 고차원의 상호작용 특성을 온전히 표현해내는데 어려움이 있다.

2. 하이퍼그래프 기반의 이미지 모델링

기존의 그래프 기반의 이미지 모델링 기법의 한계를 극복하기 위해 하이퍼그래프(hypergraph)를 이용하여 장면을 모델링하는 것을 고려할 수 있다. 하이퍼그래프는 기존의 그래프를 일반화 한 것으로, 하이퍼그래프를 구성하는 하이퍼에지(hyperedge)들은 임의의 수의 노드들을 연결할 수 있는 특성이 있다[6,15]. 따라서 시멘틱 속성들의 고차원의 상호작용을 모델링하는데 매우 적합한 기법이라 할 수 있다.

특정 장면 카테고리를 하이퍼그래프를 이용하여 모델링하기 위해, 해당 장면 카테고리를 구성하는 학습 샘플 이미지들로부터 얻을 수 있는 시멘틱 속성들의 동시 발현(co-occurrence expression) 양상을 이용하거나, 해당 이미지를 설명하기 위해 추가된 텍스트 기반의 주석(annotation) 정보를 활용할 수 있다.

예를 들어 해안(coast) 카테고리에 대한 하이퍼그래프 $HG=(S,H)$ 를 구성할 때, 해당 카테고리에 속한 학습 샘플 이미지 I_j 들로부터 우리는 하이퍼에지들을 얻을 수 있다. 해당 카테고리에서 발생가능한 시멘틱 속성들의 집합 S 가 $S = \{s_1('sky'), s_2('water'), s_3('sand'), s_4('boat'), s_5('tree'), s_6('rock')\}$ 와 같다고 가정할 경우, 하이퍼에지 h_j 는 다음과 같이 획득할 수 있다. $H = \{I_1 \rightarrow h_1 = \{s_1, s_2, s_3\}, I_2 \rightarrow h_2 = \{s_2, s_4\}, I_3 \rightarrow h_3 = \{s_1, s_2, s_4\}, I_4 \rightarrow h_4 = \{s_3, s_5, s_6\}, I_5 \rightarrow h_5 = \{s_1, s_2\}, I_6 \rightarrow h_6 = \{s_1, s_2, s_3, s_6\}\}$.

여기서 기호 ‘ \rightarrow ’는 이미지 I_j 가 하이퍼에지 h_j 에 대응됨을 의미하고, s 는 시멘틱 속성들의 집합으로써 이는 곧 그래프를 구성하는 노드들의 집합으로 생각할 수 있다. H 는 하이퍼그래프으로써 하이퍼에지들의 집합을 의미하는 것으로, 전체 노드 s 로부터 생성 가능한 부분집합(subset)들로 구성된다. 위의 예에서와 같이 해당 카테고리의 학습 샘플 이미지들로부터 얻어진 하이퍼에지들을 결합하면 특정 카테고리에 대한 하이퍼그래프를 구성할 수 있다.

그림 1에서 확인할 수 있듯이 하이퍼그래프는 하이퍼에지를 이용하여 임의의 수의 속성을 갖는 에지를 표현할 수 있다. 이러한 특성을 기반으로 우리는 각 장면 카테고리를 하이퍼그래프를 통하여 적절하게 모델링할 수 있다.

그러나 학습을 위한 샘플 이미지들로부터 생성된 하이퍼그래프 모델이 원래의 데이터를 잘 표현하고 있음에도, 이를 그대로 장면 분류 작업에 사용하기엔 부족한 면이 있다. 이는 생성된 하이퍼그래프를 구성하는 하이퍼에지를 생성함에 있어서, 대상 카테고리와의 그 밖의 카테고리 사이의 상대적 분포를 고려하지 않았기 때문이다. 따라서 장면 분류에 적합한 하이퍼그래프 모델을 생성하는 것은, 해당 카테고리의 특성을 잘 표현하면서도, 다른 카테고리와의 분별도 고려하는 하이퍼에지들을 탐색하는 작업임을 의미한다. 이는 고차원의 특징 공간으로부터 해당 클래스를 잘 표현하는 특징들의 부분집합을 찾아내는 특징 선택 기법과 유사하다. 그러나 기존의 특징 선택 기법들이 최적의 단일 특징 부분집합을 찾아내는 것을 목표로 한다면, 하이퍼그래프를 구성하는 최적의 하이퍼에지를 찾아내는 작업은 다수의 특징 부분집합을 찾아내는 것을 목표로 한다고 할 수 있다.

일반적으로 고차원의 특징 공간으로부터 최적의 부분집합을 탐색하는 작업은 NP-완비 문제에 해당하는 것으로 현실적으로 모든 공간을 탐색하기에는 문제가 따른다. 이를 위해 본 논문에서는 Kleinberg의 SD(stochastic discrimination) 이론에 따라, 특징 공간으로부터 특징 부분집합을 생성하여 분류에 작업에 적합한 준최적 하이퍼그래프를 생성하고자 한다[7].

SD 이론은 중심극한정리(central limit theorem)에 기반하여 전체 특징 공간으로부터 생성 가능한 부분집합들 중, enrichment, uniformity, projectability의 특성을 만족하는 부분집합 생성을 통해 분류를 위한 모델을 근사할 수 있음을 이론적으로 증명하고, 실험을 통해 입증하였다. 우리는

이를 구현한 기법 중 하나인 RSM(random subspace method)[8]을 개선하여 각 카테고리에 적합한 속성들의 부분집합들을 탐색하고, 이를 분류에 적용하고자 한다.

SD 이론에 따르면, 부분집합들의 생성 수가 많아질수록 본래 데이터의 분포를 잘 근사 할 수 있다. 그러나 이를 위해서는 많은 비용이 따르기 때문에, 본 논문에서는 좀 더 효율적으로 해당 모델을 근사 할 수 있는 특징 부분집합들을 탐색하기 위한 기법을 제안한다. 또한 생성된 하이퍼그래프를 분류에 이용하기 위해 우리는 분류 모델(discriminative model)을 사용한다. 그래프를 기반으로 하는 모델들은 그 특성으로 인해 일반적으로 생성 모델(generative model)을 이용하여 분류에 이용된다. 그러나 분류 모델을 사용 할 경우 보다 빠른 학습이 가능하고, 비교적 많은 분야에서 우수한 분류 능력을 보이는 것으로 알려진 SVM 등 기존의 분류 기법을 활용 할 수 있다는 장점이 있다.

이를 위해서는 하이퍼그래프로부터 새로운 특징 벡터의 생성이 필요한데, 따라서 탐색된 각 하이퍼에지들을 구성하는 각 멤버 속성들의 발현 값을 요약하는 기법을 사용할 필요가 있다. 이를 위해 본 논문에서는 각 특징들의 클래스별 조건부 확률 분포에서의 우도비(likelihood ratio)를 이용하는 요약 기법을 제시한다.

3. 확률적 부분집합 탐색 기법

RSM(random subspace method)은 SD 이론을 기반으로 하여, 랜덤하게 생성된 특징들의 부분 집합을 이용하여 분류에 적용하는 기법이다. 그리고 실험을 통해 생성 되는 부분집합의 수가 늘어날수록 분류 성능이 점진적으로 증가함을 보였다. 기존의 RSM에서 선택 되는 특징들은 모두 균등 분포에 기반한 선택 확률을 갖는다. 그러나 이는 비효율적인 것으로, 실제로 분류를 수행하는데 있어 모든 특징들이 동일한 분별력(discrimination power)을 갖는 것은 아니고 또한 장면 분류와 같은 특정한 문제에서는 전체 특징 공간에서 각 특징 벡터들의 발생 확률이 매우 희박(sparse)하다는 특성이 있다. 따라서 우리는 각 특징들의 중요도를 기반으로 선택 확률을 조정 할 필요가 있다.

본 논문에서는 각 특징들의 중요도에 기반을 두어 효율적으로 부분집합 생성이 가능한 확률적 부분집합 탐색 기법 PSM(probabilistic subset method)을 제안한다. PSM을 이용하여 장면 분류를 위한 시멘틱 속성들의 부분집합을 생성하기 위해서, 먼저 각 속성들의 중요도를 측정 할 필요가 있다. 이는 해당 카테고리와의 연관정도를 이용하여 측정할 수 있다. 이를 위해 상호 정보량(mutual information) 등을 사용할 수도 있지만, 연속형 데이터이면서 샘플의 수가 많지 않을 경우 이를 그대로 적용하기가 용이하지 않기 때문에 two-class 문제에 대한 Student's T -test를 이용 할 수 있다[9].

T -test을 이용하여 positive와 negative 클래스 두 그룹에 속한 특징들의 모집단의 모평균이 같을 것이라고 가정하는 귀무가설(null hypothesis) $H_0: \mu_p = \mu_n$ 를 기각 할 수 있는 확률인 p -value를 계산 할 수 있다. 이때 검정 통계량은 다음과 같이 계산할 수 있다.

표 1. 확률적 부분집합 탐색 및 준최적화 알고리즘
Table 1. Probabilistic subspace search and optimization algorithm of searched hyperedges

<p><Probabilistic subspace search> $p(s) \leftarrow$ Occurrence probability distribution prediction of s_i by T-test $H_0 \leftarrow$ Creation of initial m hyperedges based on $p(s)$ (Each subspace is the same as each hyperedge) 1. Repeat { (a) Fitness calculation of hyperedge h_m by T^2 (b) Sort hyperedges in descending order based on their fitness value (c) Remove the bottom 30% hyperedges, then replace with newly created hyperedges based on $p(s)$ } until(Predetermined number of iteration k) 2. Obtain the top 10% hyperedges</p>
<p><Optimization of searched hyperedges> 3. For $i = 1$ to H (a) Sort member attribute s_i of hyperedge h_m in descending order based on Abs. T-test score (b) Add the 1st ranked s_i to empty set \hat{h}_m, then calculate the discriminative power (c) Repeat { (1) $\hat{h}_m = \hat{h}_m \cup \{s_i\}$: Add the next ranked s_i (2) Calculate discriminative power of \hat{h}_m } until(Greater than prev. discriminative power) 4. End for 5. Remove a duplicate hyperedge</p>
<p>Output : Optimized hyperedge set H</p>

$$T = \frac{(\bar{X}_p - \bar{X}_n)}{\sqrt{S_p^2/n_p + S_n^2/n_n}} \quad (1)$$

여기서 \bar{X}_c 는 각 클래스의 표본 평균, S_c^2 은 각 클래스의 표본 분산, n_c 는 각 클래스에 속한 샘플들의 수를 의미한다.

p -value는 값이 0에 가까울수록 귀무가설을 기각 할 확률이 높은 것으로, 이는 각 클래스에서의 분포 양상이 매우 다름을 의미한다. 즉, 해당 특징의 분별력이 높음을 의미한다. 이를 시멘틱 속성의 선택 확률에 적용하기 위해서 본 논문에서는 $p(s_i) = 1 - p$ -value 값을 이용한다.

이들 시멘틱 속성 선택 확률 $p(s_i)$ 에 기반한 부분집합 탐색 과정은 표 1과 같다. 탐색의 초기에 미리 정해진 수만큼의 부분집합들을 생성하고(본 논문에서는 $m=1000$, $k=20$ 으로 설정), 이들의 적합도(fitness) 값을 구한다. 적합도를 구하기 위한 다양한 기법이 존재 할 수 있지만, 본 논문에서는 빠른 연산이 가능하면서도 견고한 적합도 값을 얻기 위해, 다변량 검정에 사용되는 Hotelling's T^2 -test를 사용한다[10]. Hotelling's T^2 -test는 단변량 검정에 일반적으로 사용 되는 Student's T -test를 일반화 한 것으로, 두 집단 간 최대한의 차이를 창출하도록 2개 이상의 종속변수들로부터 생성된 변량(variate)의 통계적 검정이 가능하다

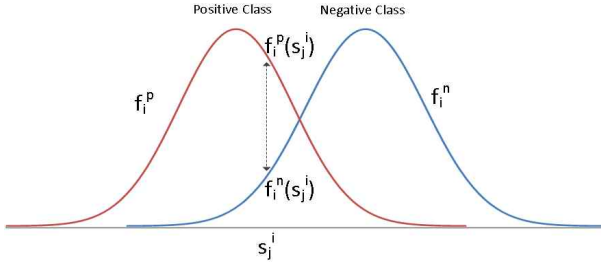


그림 2. 시멘틱 속성에 대한 정규분포 기반 우도비 계산
Fig. 2. Likelihood estimation for each semantic attribute

다. 해당 기법은 집단 간의 차이를 종합적으로 검정하는 기법이기에 때문에 제1종 오류(Type 1 error)의 팽창문제를 완화하는 장점이 있다. 또한 계산 과정에서 구성 요소들 간의 상관관계와 상호의존성을 고려하므로, 부분집합을 구성하는 시멘틱 속성들의 고차원 상호작용 관계를 통계적으로 평가하는데 적합하다. 생성된 부분집합들에 대한 Hotelling's T^2 -test score는 다음과 같이 구할 수 있다.

$$T^2 = \frac{n_p n_n}{n_p + n_n} (\bar{\mathbf{A}}^p - \bar{\mathbf{A}}^n) \mathbf{S}^{-1} (\bar{\mathbf{A}}^p - \bar{\mathbf{A}}^n)^T \quad (2)$$

여기서 n_c 는 각 클래스에 속한 샘플들의 수를 의미하고, $\bar{\mathbf{A}}^c$ 는 각 클래스에 속한 시멘틱 속성 s_j^p 와 s_j^n 들의 평균 발현량을 의미한다. \mathbf{S} 는 시멘틱 속성들에 대한 합동 분산-공분산 행렬(pooled variance-covariance matrix)를 의미한다.

계산된 적합도를 기반으로 각 부분집합들을 내림차순 정렬한다. 정렬된 부분집합들 중 하위 30%를 제거하고, 새롭게 생성한 부분집합들로 대체한다. 이와 같은 학습 과정을 정해진 횟수만큼 반복 수행한 후, 상위 10%의 부분집합들을 하이퍼그래프를 구성하는 부분집합 후보로 획득할 수 있다.(이들 상수 값들은 실험을 통해 경험적으로 설정된 값들이다.) 이후 각 하이퍼에지들은 표 1의 하단의 준최적화 과정과 같이 점진적 학습(incremental learning)을 통해 분별력이 최대화된 멤버 속성들로 구성될 수 있도록 필터링을 수행한다.

4. 우도비 기반 특징 벡터 생성

학습된 하이퍼그래프를 기반으로 장면 분류에 적용하기 위해 분류 모델(discriminative model)을 사용한다. 이를 위해서는 학습 데이터와 테스트 데이터에 대한 각각의 특징 벡터를 생성할 필요가 있는데, 본 논문에서는 우도비 기반의 요약 기법을 사용한 특징 벡터 생성 기법을 제안한다.

분류를 위해 학습된 하이퍼그래프를 구성하는 하이퍼에지들은 멤버 시멘틱 속성들로 구성되어 있다. 이들은 각 장면 이미지에서 각각의 발현 값을 갖는다. 이 발현 값을 그대로 이용하는 대신 해당 시멘틱 속성의 각 클래스에서의 조건부 확률분포(conditional probability distribution)를 예측하고, 각 클래스에서의 우도(likelihood)를 계산하여 발현의 상대적 분포의 차를 아래와 식 (3)과 같이 우도비(likelihood ratio)로 계산하여 나타낸다. (그림 2 참조)

$$\lambda_i(s_j^i) = \frac{f_i^p(s_j^i)}{f_i^n(s_j^i)} \quad (3)$$

여기서 $f_i^c(s)$ 는 클래스 c 에서의 각 시멘틱 속성의 조건부 확률 분포 함수이다. 이는 해당 시멘틱 속성의 발현 값이 어떤 클래스에서 발생했을 가능성이 높은지에 대한 확률을 의미한다. 이를 위해 본 논문에서는 각 클래스에서의 시멘틱 속성의 발현 값이 평균 μ_i^c , 표준편차 σ_i^c 의 값을 갖는 정규분포를 따른다고 가정한다. 이들 파라미터는 학습 데이터로부터 예측한 값을 이용한다.

일반적으로 이들 파라미터는 충분한 양의 학습 데이터로부터 예측된 것이 아니기 때문에 신뢰성이 낮은 편이라고 할 수 있다. 따라서 이들 값은 아주 작은 발현 값의 변화에도 매우 민감하다. 이를 위해 본 논문에서는 평균 0, 표준편차 1이 되도록 표준화(standardization)한 우도비 $\hat{\lambda}_i(s_j^i)$ 를 사용한다. 이들 값을 이용하여 각 하이퍼에지를 구성하는 멤버 속성들의 우도비 값은 다음과 같이 결합되어 하나의 특징 값으로 사용된다.

$$\text{aggregated value of } h_k = \frac{1}{|h_k|} \sum_{i=1}^n \hat{\lambda}_i(s_j^i) \quad (4)$$

이는 Fisher가 제안한 선형 판별함수(linear discriminant function)를 이용한 변환 기법과 유사한 과정으로, 여러 발현 값을 하나의 값으로 요약하는 과정을 통해 발현 편차(variation)에 둔감해지는 효과가 있다. 이로 인해 시멘틱 속성들의 발현 값을 측정하는 과정에 포함된 잡음에 강건한 특징을 생성할 수 있다. 이들 특징 벡터는 3절에서 보인 탐색 과정을 통해 획득되는 부분 집합들로부터 생성되는 값이다. 따라서 분류에 사용되는 특징 벡터의 차원은 이들의 수와 관련이 있다.

특징 벡터를 생성하기 위해 사용되는 하이퍼에지들은 준최적화 과정을 통해 탐색된 분별력이 최대화된 하이퍼에지들이다. 그러나 최고의 분류 성능을 내는 특징 벡터의 차원의 수는 알 수 없으므로 적절한 값을 선택할 필요가 있다. 이를 위해 분류에 사용되는 특징 벡터의 차원의 수는 학습 데이터에 대한 교차 검증 과정을 통하여 최고의 분류 성능을 내는 값으로 설정하였다. 이 과정을 통해 결정된 각 카테고리에 대한 특징 벡터의 차원의 수는 평균 4.69개였다.

5. 실험 및 결과

5.1 실험 데이터셋

실험을 위해 장면 분류와 관련된 연구들에서 일반적으로 사용되는 두 종류의 데이터셋(Scene-13, Scene-15)을 사용하였다. 첫 번째로 Scene-13 데이터셋[3]은 13개의 실내와 실외에서의 장면 카테고리로 다음과 같이 구성되어 있다. Highway (260), Inside city (308), Tall building (356), Streets (292), Suburb residence (241), Forest (328), Coast (360), Mountain (374), Open country (410), Bedroom (174), Kitchen (151), Livingroom (289), Office (216). 여기서 괄호안의 숫자는 구성 샘플 이미지의 수를 의미한다. Scene-13 데이터셋은 Oliva 등이 사용한 8개의 카테고리로 구성된 데이터셋에 5개의 카테고리를 추가하여 Li

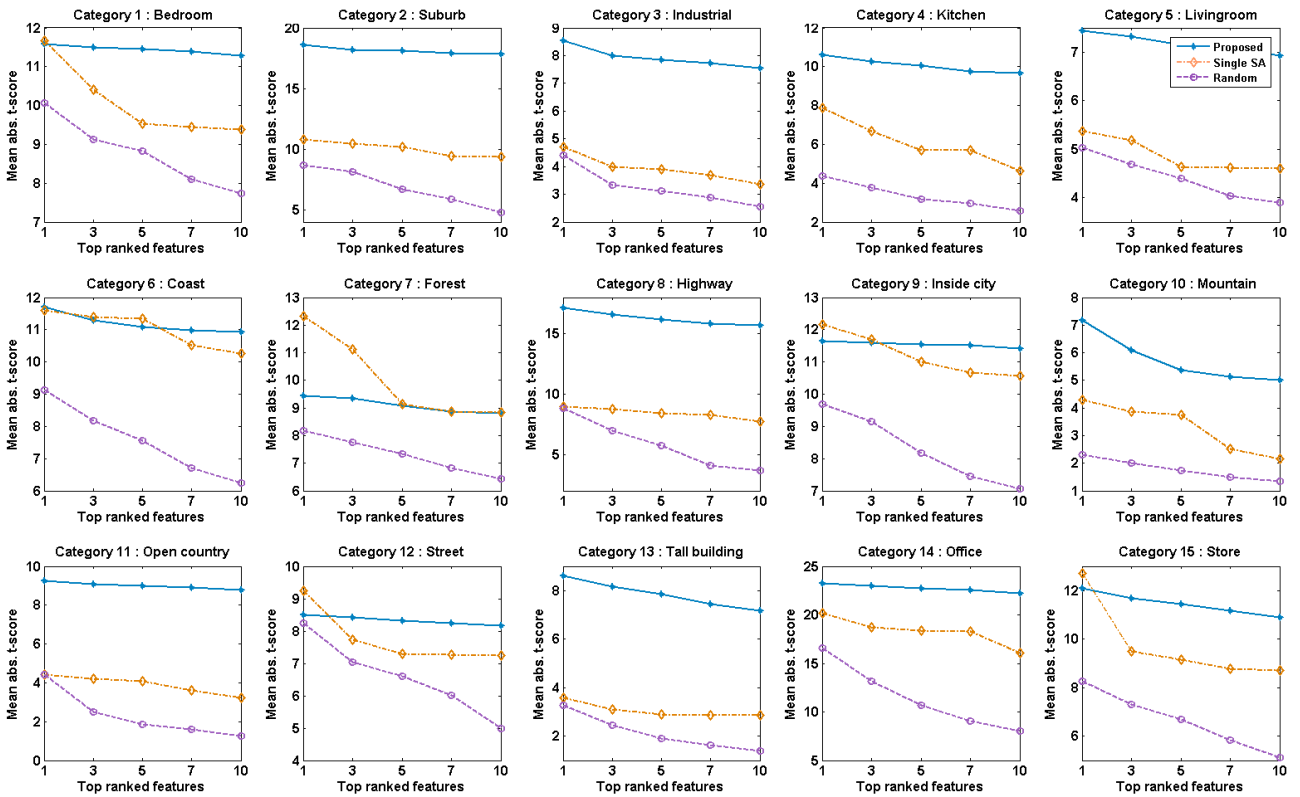


그림 3. 생성된 특징 벡터의 카테고리 별 분별력 평가

Fig. 3. Discriminant power evaluation of the generated feature vector for each category

등이 실험에 사용한 데이터셋이다. 각 이미지들의 평균적인 크기는 250×300 픽셀이고, 그레이스케일 이미지이다. 두 번째 Scene-15 데이터셋은 Scene-13 데이터셋에 2가지의 카테고리 Industrial (311)과 Store (315)가 추가된 것으로 [11]의 실험에서 사용 되었다.

검증을 위해 기존의 연구들에서와 동일한 방식으로 매 반복 실험에서, 각 카테고리에서 100개의 이미지를 랜덤하게 선택하여 훈련 데이터로 사용하고, 나머지는 검증을 위한 테스트 데이터로 사용한다.

5.2 시멘틱 속성 생성 및 장면 분류기

앞서 말했듯이 본 논문에서는 각 이미지를 구성하는 시멘틱 속성들을 탐색하고, 이들의 발현 값을 특징 벡터 생성 시 활용한다. 따라서 각 시멘틱 속성들에 대한 탐색기가 필요한데, 이를 위해 [5]에서 Su 등이 제안한 속성 분류기를 사용한다. 이 속성 분류기는 전역(35) 및 지역(16) 장면 요소, 재료(14), 모양(7), 색상(8), 특정 물체(30) 등으로 구성된 110개의 개별적인 시멘틱 속성들에 대한 발현 값을 구할 수 있다. 본 논문에서는 이들 중 전역 장면 속성 35개와 색상 속성 8개를 제외한 67개의 속성들만을 사용하였다. 이들은 각 속성의 발현 값을 구하기 위해 한 이미지에 대해 100개의 지역 패치를 생성하고 이들을 시멘틱 속성 분류기를 이용하여 분류한 결과를 종합하여 사용한다. 획득 되는 모든 시멘틱 속성 s_j^i 들은 0과 1사이의 발현 값을 갖게 된다.

장면 분류를 수행하기 위해 본 논문에서는 다수의 연구 결과들을 통해 히스토그램 기반의 이미지 분류에 가장 적합한 것으로 알려진 Chi-square 커널 기반의 SVM을 사용하

였다[12]. Chi-square 커널은 아래와 같이 구할 수 있다.

$$k(d_1, d_2) = \exp\left(-\frac{\sigma}{2} \sum_{i=1}^n \frac{(d_1(i) - d_2(i))^2}{d_1(i) + d_2(i)}\right) \quad (5)$$

여기서 d_c 는 각 클래스에서의 특징 벡터를 의미하고, σ 는 정규화 계수이다.

SVM은 공개 라이브러리인 LIBSVM을 이용하여 구현하였다. SVM 분류기 사용 시 필요한 파라미터 C 와 Chi-square 커널의 정규화 계수 σ 그리고 분류에 사용할 특징벡터의 크기는 학습 데이터에 대한 5-fold 교차 검증 과정을 통해 얻어진 최적의 파라미터 값을 사용한다. 또한 다수의 카테고리가 포함된 멀티 클래스 분류 문제에 속하므로 One-versus-all 방식을 사용한 분류를 수행한다. 성능 평가 결과는 전체 카테고리에 대한 평균 분류율(average classification rate)을 사용하여 계산한다. 최종 평가 결과는 동일한 실험을 50회 반복한 결과를 평균하여 얻은 결과를 사용한다.

5.3 실험 결과

제안 한 기법의 성능을 검증하기 위해 먼저 생성된 특징 벡터들의 분별력을 분석하였다. 이는 분별력이 우수한 특징 벡터를 생성 할 수 있어야 높은 분류 성능을 획득할 수 있기 때문이다. 그림 3은 Scene-15 데이터셋의 15개의 각 카테고리에서 탐색된 상위 10개의 특징들의 분별력을 mean absolute t -score를 이용하여 분석한 결과이다. (Scene-13의 데이터는 Scene-15 데이터셋에 포함되어 있으므로 분별

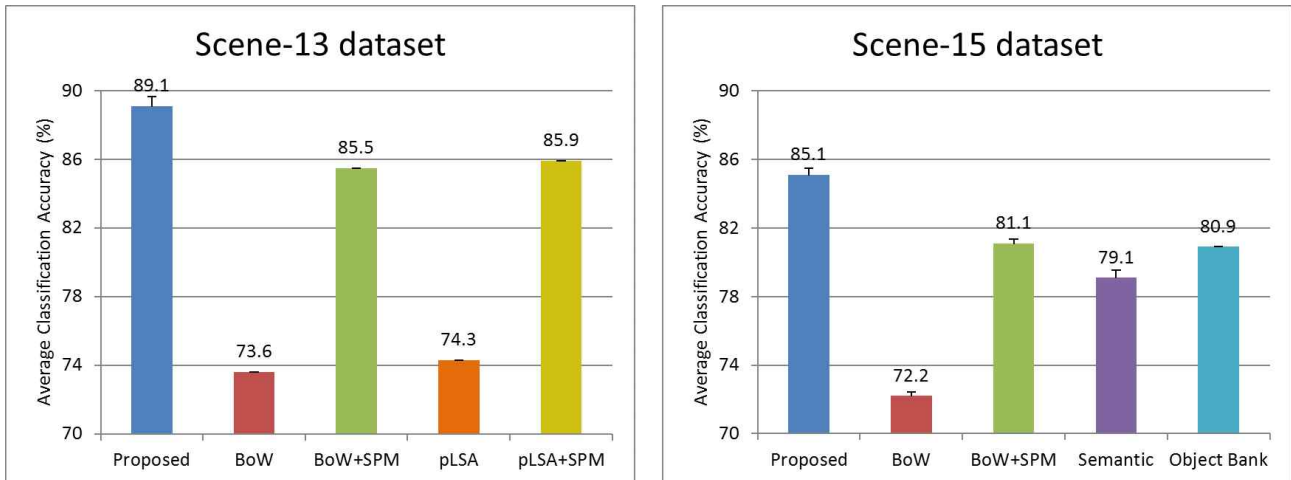


그림 4. 장면 이미지 분류 성능 비교
 Fig. 4. Comparison of scene image classification performance

력 분석 실험에서는 제외) 그림 3의 x축은 특징 벡터들의 순위를 의미하고, y축은 각 특징 벡터들의 mean absolute t-score를 의미한다.

비교를 위해 단일 시멘틱 속성의 분별력과 랜덤하게 선택된 부분집합으로부터 생성한 특징 벡터의 분별력을 함께 도시하였다. (이들은 각각 Single SA와 Random으로 표기) 그림 3에서 알 수 있듯이 본 논문에서 제안된 확률적 부분집합 탐색 기법과 우도비 기반의 요약 기법을 통하여 생성한 특징 벡터들이 대부분의 카테고리에서 우수한 분별력을 보임을 알 수 있다. 이는 곧 분류 성능의 향상과 관련 된다. 특히 다른 기법들과는 달리 특징들의 순위가 낮아지더라도 분별력이 크게 떨어지지 않고, 비교적 높은 분별력을 유지하고 있음을 알 수 있다.

그림 4는 생성된 특징 벡터들을 이용한 분류 성능을 비교한 결과이다. 제안한 기법이 Scene-13 데이터셋과 Scene-15 데이터셋에 적용한 실험에서 기존의 기법들과 비교하여 우수한 분류 성능을 보임을 알 수 있다.

이는 앞서 실험결과를 통해 보였듯이 제안한 기법으로 생성된 특징 벡터들이 기존의 특징벡터들에 비해 높은 분별력을 가지고 있기 때문으로 분석할 수 있다.

Scene-13 데이터셋 실험에서 pLSA(probabilistic latent semantic analysis)로 표기한 실험은 Bag-of-Words 기반의 저수준 특징벡터를 기반으로 확률 모델링 기반의 토픽 모델인 pLSA를 이용하여 중간 수준의 특징벡터를 생성한 후 분류에 적용한 결과이다[1]. 이들 pLSA 기반의 중간 수준의 특징벡터를 사용하는 실험 결과는 기존의 Bag-of-Words 모델 기반의 저수준 특징벡터를 그대로 사용한 경우와 유사한 분류 성능을 보임을 알 수 있다(각각 pLSA, BoW로 표기). 이들을 공간 피라미드 기법과 결합한 경우(각각 BoW+SPM, pLSA+SPM로 표기) 약간의 분류 성능의 향상을 보이지만, 제안한 기법은 공간 피라미드 기법과 결합한 결과들과 비교하여 3.2% 이상 향상된 우수한 분류 성능을 보였다.

Scene-15 데이터셋 실험에서 Semantic[5]과 Object Bank[4]로 표기된 실험들은 단일 시멘틱 속성들을 이용한 기법들으로써 Bag-of-Words 모델만을 사용한 실험에서보다는 우수한 분류 성능을 보였으나, Bag-of-Words 모델을

공간 피라미드 기법과 결합하여 사용한 실험 결과[1] 보다는 낮은 분류 성능을 보임을 알 수 있다(각각 BoW, BoW+SPM로 표기). 제안한 기법은 공간 피라미드 기법과 결합하지 않았음에도 Bag-of-Words 모델에 기반한 다른 두 실험 결과보다 4% 이상 향상된 우수한 분류 성능을 보임을 알 수 있다.

흥미롭게도 두 데이터셋에 대한 실험에서 모두 공간 피라미드 기법(spatial pyramid method)[10]을 결합할 경우 분류 성능 향상이 나타나는 것을 볼 수 있는데, 이는 이미지 상의 공간적 문맥 정보(spatial contextual information)를 결합한 결과라고 할 수 있다. 따라서 향후 제안한 시멘틱 속성 결합 기법과 공간 피라미드 기법을 결합할 경우 더 높은 분류 성능을 얻을 수 있을 것으로 기대할 수 있다.

6. 결론 및 향후 연구

본 논문에서는 장면 이미지 분류 문제에 있어서 시멘틱 속성들을 개별적으로 고려하는 경우 발생할 수 있는 측정 잡음 및 모호성 문제 등을 해결하기 위해, 각 카테고리에 준최적화된 속성 부분 집합(하이퍼에지) 기반의 하이퍼그래프를 생성하고, 이를 장면분류에 활용하는 기법을 제안하였다. 하이퍼그래프 기반의 표현 방법을 통해 특정 장면 카테고리에서 발생하는 시멘틱 속성들의 고차원의 상호작용 관계를 고려한 분석이 가능해진다. 본 논문에서는 학습을 통해 생성된 하이퍼그래프를 장면 분류에 적용하기 위해 분류 모델과 결합하는 기법을 제안하였다. 이를 통해 SVM 등 기존에 여러 분야에서 활용되어 성능이 입증된 분류기를 활용하는 것이 가능하였다.

특히 하이퍼그래프를 구성하는 각 속성들의 부분집합인 하이퍼에지들을 구성하는 멤버 속성들의 발현 값을 하나의 값으로 나타내는 우도비 기반의 선형변환 과정을 통해 단일 속성들에 비해 높은 분별력을 갖는 특징 벡터를 생성할 수 있음을 실험을 통해 확인하였다. 또한 두 가지의 데이터 집합에 대한 실험을 통해 제안한 기법이 경쟁력 있는 분류 성능을 갖고 있음을 확인하였다. 특히 제안한 기법은 일반적으로 이미지 분류 분야에서 높은 분류 성능을 보이는 것으로

로 알려진 Bag-of-Words 모델과 공간 피라미드 기법을 결합한 모델보다 약 3~4% 이상의 성능향상을 보였다.

제안 한 기법은 차원을 효과적으로 감소시키면서도 우수한 분류 성능을 보여주는 모델로서, 향후 속성들의 수가 증가 되었을 때에도 활용 가능성이 매우 크다고 할 수 있다. 하지만 본문에서 언급했던 것과 같이 제안한 기법은 개별 시멘틱 속성 분류기의 성능에 영향을 크게 받을 수밖에 없다. 따라서 보다 우수한 분류 성능을 얻기 위해서는 향후 개별 시멘틱 속성 분류기의 분류 성능 및 재현성을 높이기 위한 연구가 뒷받침 되어야 할 필요가 있다.

References

[1] A. Bosch, A. Zisserman, X. Muñoz, "Scene Classification Using a Hybrid Generative /Discriminative Approach," *IEEE Trans. PAMI*, 30(4), pp.712-727, 2008.

[2] J. Sivic, A. Zisserman, "Video google: A text retrieval approach to object matching in videos," *In: Proc. ICCV*, 2003.

[3] L. Fei-Fei, P. Perona, "A bayesian hierarchical model for learning natural scene categories," *In: Proc. CVPR*, 2005.

[4] L. Li, H. Su, E. Xing, Li Fei-Fei, "Object bank: a high-Level image representation for scene classification and semantic feature sparsification," *In: Proc. NIPS*, 2010.

[5] Yu Su, F. Jurie, "Improving Image Classification Using Semantic Attributes," *International Journal of Computer Vision*, 100(1), pp.59-77, 2012.

[6] Vitaly I. Voloshin. "Introduction to Graph and Hypergraph Theory," *Nova Science Publishers, Inc.*, 2009.

[7] E. M. Kleinberg, "Stochastic Discrimination," *Annals of Mathematics and Artificial Intelligence*, pp.207-239, 1990.

[8] Tin Kam Ho, "The Random Subspace Method for Constructing Decision Forests," *IEEE Trans. PAMI*, 20(8), pp.832-844, 1998.

[9] C. Ding, H. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of Bioinformatics and Computational Biology*, vol.3(2), pp.185-205, 2005.

[10] Y. Lu, P.Y. Liu, P. Xiao, H.W. Deng, "Hotelling's T2 multivariate profiling for detecting differential expression in microarrays," *Bioinformatics*, 21, pp.3105 - 3113, 2005.

[11] S. Lazebnik, C. Schmid, J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," *In: Proc. CVPR*, 2006.

[12] O. Chapelle, P. Haffner, V. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Trans. Neural Networks*, 10(5), pp.1055-1064, 1999.

[13] Jonh-Tae Joo, Dae-Wook Kim, Kwee-Bo Sim, "Pattern Classification Algorithm of DNA Chip Image using ANN," *Journal of The Korean Institute of Intelligent Systems*, No.16, Vol.5, pp.556-561, 2006.

[14] Yoon-Cheol Shin, Young-Hun Park, Hoon Kang, "Image Pattern Classification and Recognition by Using the Associative Memory with Cellular Neural Networks," *Journal of The Korean Institute of Intelligent Systems*, No.13, Vol.2, pp.154-162, 2003.

[15] Hyung Lee-Kwang, Gil Rok Oh, Choong Ho Cho, "Fuzzy Hypergraph," *Journal of The Korean Institute of Intelligent Systems*, No.4 Vol.2, pp.3-8, 1994.

저 자 소 개



최선욱(Sun-Wook Choi)

2007년 : 인하대학교 전자공학
(학사)

2009년 : 인하대학교 정보통신공학
(석사)

2009년~2014년 : 인하대학교 전자공학
(박사)

관심분야 : 로봇 지능, 패턴 인식, 초병렬 연산 컴퓨팅, 컴퓨터 비전, HRI, Visual SLAM
E-mail : swchoi@inhaian.net



이종호(Chong Ho Lee)

1976년 : 서울대학교 전기공학
(학사)

1978년 : 서울대학교 전기공학
(석사)

1986년 : 미국 아이오와 주립대학교 전기 및 컴퓨터공학 (박사)

1986년~1989년 : 미국 노틀담 대학교

조교수

1989년~ 현재 : 인하대학교 정보통신공학과 교수

1997년~1998년 : 인하대 직접회로설계센터 소장

2000년~2010년 : 인하대 슈퍼지능기술연구소 소장

2004년~2005년 : 미국 브라운대학교 두뇌 및 신경회로
망연구소 방문 교수

관심분야 : 지능형시스템, VLSI 설계

E-mail : chlee@inha.ac.kr