

---

# 특징, 색상 및 텍스처 정보의 가공을 이용한 Bag of Visual Words 이미지 자동 분류

박찬혁\* · 권혁신\* · 강석훈\*

\*인천대학교

## Improved Bag of Visual Words Image Classification Using the Process of Feature , Color and Texture Information

Chan-hyeok Park\* · Hyuk-shin Kwon\* · Seok-hoon Kang\*

\*Incheon University

E-mail : tpzjstkfkd12@gmail.com

### 요 약

이미지를 분류하고 검색하는 기술(Image retrieval)중 하나인 Bag of visual words(BoVW)는 특징점(feature point)을 이용하는 방법으로 데이터베이스의 이미지 특징벡터들의 분포를 통해 쿼리 이미지를 자동으로 분류하고 검색해주는 시스템이다. Words를 구성하는데 특징벡터만을 이용하는 기존의 방법은 이용자가 원하지 않는 이미지를 검색하거나 분류할 수 있다. 이러한 단점을 해결하기 위해 특징벡터뿐만 아니라 이미지의 전체적인 분위기를 표현할 수 있는 색상정보나 반복되는 패턴 정보를 표현할 수 있는 텍스처 정보를 Words를 구성하는데 포함시킴으로서 다양한 검색을 가능하게 한다. 실험 부분에서는 특징정보만을 가진 words를 이용해 이미지를 분류한 결과와 색상정보와 텍스처 정보가 추가된 words를 가지고 이미지를 분류한 결과를 비교하였고 새로운 방법은 80~90%의 정확도를 나타내었다.

### ABSTRACT

Bag of visual words(BoVW) is one of the image classification and retrieval methods, using feature point that automatical sorting and searching system by image feature vector of data base. The existing method using feature point shall search or classify the image that user unwanted. To solve this weakness, when comprise the words, include not only feature point but color information that express overall mood of image or texture information that express repeated pattern. It makes various searching possible. At the test, you could see the result compared between classified image using the words that have only feature point and another image that added color and texture information. New method leads to accuracy of 80-90%.

### Keyword

BoVW, Image Classification ,TBIR(Text based image retrieval) , CBIR(Cotent based image retrieval), GLCM(Gray-Level Co-Occurrence Matrix) , LAB color space

### 1. 서 론

최근 컴퓨터 비전 분야의 발달에 따라 멀티미디어 정보 서비스에 대한 요구가 증가 하고 있으며, 방대한 이미지 데이터베이스 안에서 이미지를 다루는 검색 기술에 대한 연구의 필요성이 증가 하고 있다. 기존의 이미지 검색은 데이터에 직접 색

인을 첨가하여 정보를 검색하는 텍스트 기반 이미지 검색(TBIR:Text Based Image Retrieval)이 검색에 사용되어 왔다. TBIR은 보는 관점에 따라 검색의 효율성이 저하 될 수 있으며 이미지의 데이터를 텍스트만으로 표현하기에는 무리가 있다. 이러한 단점들을 보완하기 위해 이미지안의 내용

데이터를 검색인자로 선택하여 이미지를 검색하는 내용 기반 이미지 검색(CBIR:Content Based Image Retrieval)에 대한 연구[1]가 활발하게 이루어지고 있다. CBIR의 대표적인 방법으로 BoVW (Bag of visual words 혹은 BoF:Bag of Feature)가 사용되며 이는 문서를 자동으로 분류하기 위한 방법 중 하나인 BoW(Bag of Words)를 이미지 분야에 응용한 방식이다. BoW기법은 단어들의 분포를 파악해 이 문서가 어떤 종류의 문서인지를 판단하게 된다. 예를 들어, 어떤 문서에서 '야구', '메시', '스타디움' 등의 단어가 많이 나온다면 스포츠에 관련된 문서로 분류하게 되는 방식이다. 이러한 BoW에서 파생된 BoVW는 단어 대신 이미지의 특징 벡터를 이용해 이미지가 어떠한 종류의 이미지인지 분류할 수 있게 해주는 기법이다. BoVW는 특징 추출 알고리즘[sift, surf, fast 논문]을 이용하여 이미지 내부의 객체들의 엡지 정보를 가지고 이미지를 표현한다. 이러한 엡지 정보들을 visual words라고 부르며 이러한 visaul words들은 이미지를 대표할 수 있는 정보를 압축해 가공 되어 진다[2],[3],[4]. 그러나 이러한 엡지 정보만을 가지고 visual words를 구성 한다면 다른 관점에서 봤을 때 사용자가 원하는 방향으로 검색이나 분류가 이루어지지 않을 수 있다. 예를 들어 이미지안의 객체에 대해 비슷한 사진을 검색하는 것이 아니라 검색하려는 이미지의 전체적인 분위기가 비슷한 이미지를 검색하려 한다면 엡지 정보만으로는 제대로 된 검색을 하지 못할 수 있다. 예를 들어 '비오는 날의 숲'을 가지고 이미지 검색을 한다고 생각했을 때 사용자는 '비오는'이라는 키워드에 대해 초점을 맞출 수 있지만 특징 벡터만 가지고 검색을 한다면 '화창한 날의 숲'이 검색되어 질 수 있다. 본 논문에서는 위와 같은 문제점을 해결하고 좀 더 다양한 검색을 할 수 있게 하기 위해, visual words를 구성 할 때 엡지 정보가 담긴 특징벡터뿐 아니라 칼라정보와 패턴정보를 추가하여 visual words를 가공 하여 실험을 하여 결과를 비교하고 결론을 맺는다.

## II. BoVW 시스템

BoVW시스템[1]은 크게 2가지 훈련 부와 검색 부로 나뉜다.(그림 1.) 훈련 부의 경우 이미지 데이터들에서 비주얼 워드를 뽑아내어 워드들의 모음인 codebook을 생성하는 부분이며 검색 부의 경우 생성된 codebook에서 쿼리 이미지의 visaul words와 비교하여 어떤 클래스에 속하는지 나타내어 주는 부분이다.

### 2.1 FeatureExtraction

BoVW를 위해 첫 번째로는 이미지에서 그 이

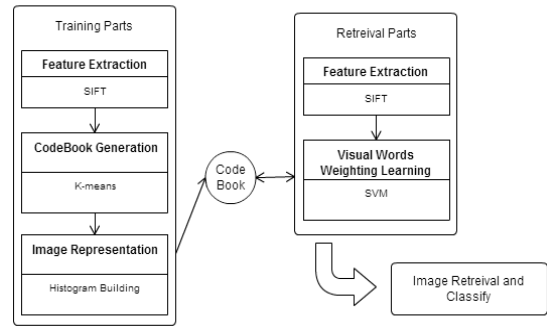


그림 1. BoVW 시스템 전체 흐름도

미지를 대표할 수 있는 정보를 추출하는 과정이다. 본 논문에서는 SIFT(Scale Invariant Feature Transfom) 특징 추출 알고리즘[5]을 이용한 엡지 정보를 사용하였고, 또한 다른 정보를 추가하기 칼라정보와 텍스처 정보는 각각 LAB 칼라 정보와 GLCM(Gray-level co-occurrence Matrix) 텍스처 정보[6]를 이용하여 비주얼 워드를 가공하였다. 기본적으로 엡지 정보를 나타내는 SIFT descriptor의 비중을 50%, 이미지의 전체적인 색감이나 분위기를 나타내어 주는 color descriptor는 30% 이미지의 반복인 패턴을 나타내는 texture descriptor는 20%의 비중을 두어 비주얼 워드를 가공하였다.

### 2.2 CodeBook Generation

CodeBook은 2.1에서 추출한 각각의 descriptor들을 K-means클러스터링을 이용하여 전체 이미지 데이터베이스의 모든 정보를 워드단위로 함축하여 갖고 있는 데이터 묶음이다. K-means클러스터링은 n개의 d-차원 데이터 오브젝트  $(x_1, x_2, \dots, x_n)$  집합이 주어졌을 때, n개의 데이터 오브젝트들을 각 집합 내 오브젝트 간 응집도를 최대화하는  $K(\leq n)$ 개의 집합  $S = S_1, S_2, \dots, S_k$ 으로 분할한다. 예를 들어 모든 이미지들의 특징을 합한 개수가 10000개라고 하고  $K=200$ 일 때 10000개의 데이터는 200개의 집합으로 분할되어 표현되어진다. 이때 K는 코드북에 생성되는 비주얼 워드의 총 개수이다. 그림2.에서와 같이 각각의 이미지에서 특징들을 추출하여 모든 데이터들을 클러스터링하면 하나의 코드북이 생성 된다.

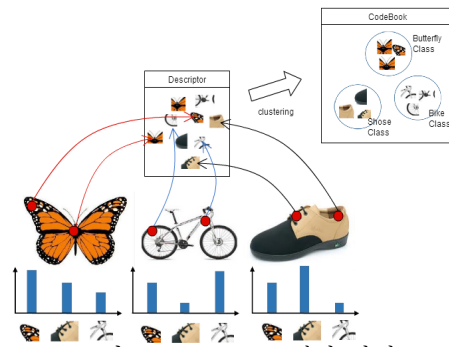


그림 2. CodeBook 생성 과정

### 2.3 Image Representation

이 단계에서는 생성된 코드북을 이용해 각각의 이미지들을 히스토그램으로 나타내는 과정이다. 2.2과정에서 각각의 정보(특징, 색상, 텍스처)에 따라 코드북이 총 3개가 생성되게 된다. 이러한 각각의 코드북을 가지고 훈련 이미지에서 추출한 정보(각 정보들의 descriptor)들과 유클리디안 거리를 이용하여 히스토그램을 이미지당 각각 3개씩 생성할 수 있다. 만들어진 히스토그램을 각 정보의 비율에 맞게 합치면 이미지 하나당 코드북에 대응하는 하나의 히스토그램이 나타내어진다.

### 2.4 Learning and Recognition

위 단계에서 생된 훈련 이미지의 히스토그램 데이터들을 feature vector로 해석하여 SVM(Support Vector Machine)분류기에 넣고 클래스 경계를 학습시켜 훈련 데이터를 완성시킨다. 검색 부의 쿼리 이미지의 경우 훈련부와 마찬가지로 특징들을 추출한 뒤 히스토그램 데이터를 얻을 수 있고 이렇게 만들어진 히스토그램 데이터는 SVM을 이용해 만들어진 학습 데이터에 넣어 해당 이미지가 어떤 클래스에 분류하는 지 알 수 있게 된다.

## III. 실험 결과

실험에 사용된 틀은 매트랩을 사용하였고, 훈련 데이터는 5개의 클래스로 나누어 각 클래스당 100장의 이미지를 훈련 데이터로 만들었다. 각각의 클래스는 '건물', '사막풍경', '사람 얼굴', '밤하늘 풍경', '시계' 사진을 사용 하였다. 각 클래스별로 훈련에 사용되지 않은 이미지 10장, 훈련에 사용된 이미지 5장씩 총 15장의 쿼리 이미지를 분류기에 넣고 원래의 클래스에 분류되는지 정확도를 분석하였다.

표 1. 이미지 검색에 사용된 정보별 정확도

	Feature	Color	Texture	Proposed Method
Building	0.87	0.63	0.80	1
Desert	0.73	0.87	0.43	0.87
Face	0.67	0.80	0.60	0.80
Night Sky	0.93	0.90	0.42	0.94
Watch	0.80	0.60	0.70	0.94

각 이미지마다 비슷한 색상을 많이 가지고 있는 '사막 풍경', '얼굴', '밤하늘 풍경'의 경우 특징 벡터를 이용해 검색한 결과보다 색상을 가지고 검색한 경우 더 좋은 결과를 내었고, 반대로 '건물', '시계'의 경우 다양한 색상을 가지고 있지만 비슷한 패턴이나 비슷한 특징을 가지고 있기 때문에 특징과 텍스처 부분에서 더 정확한 결과를 나타내었다.

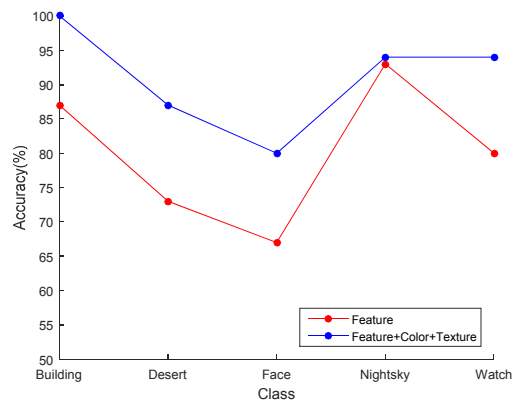


그림 3. 기존의 Feature만 가지고 BoVW 이미지 검색을 한 방법과 새로운 정보 색상정보와 텍스처 정보를 추가해 BoVW 이미지 검색을 한 결과

특징 정보 50%, 색상정보 30%, 텍스처 정보 20%의 비율로 세 가지 정보를 모두 합해 이미지 검색을 한 제안된 메서드가 단일 정보를 가지고 이미지 검색을 했을 때 보다 정확한 결과를 나타내었다.

## IV. 결론

본 논문에서는 기존의 이미지 검색 방법이 옛지 정보만을 가지고 검색하여 사용자에게 제한된 검색 정보를 보여주는 점을 극복하기 위해 새로운 정보를 추가하여 다양한 관점에서 검색이 가능한 방법을 제안한다. 제안된 색상 텍스처 정보가 가미된 BoVW 이미지 검색 시스템은 사용자의 시각적인 선호도를 분석하여 이미지 정보를 다양한 방법으로 나타낼 수 있게 기여 할 수있다. 이러한 사진들을 가공해 의미 있는 정보를 만드는 방법은 앞으로 많은 연구가 진행 될 것이다.

## 참고문헌

- [1] Gudivada, V.N., 'Content based image retrieval systems', IEEE Computer Society, Sep 1995
- [2] SuGil Choi, 'Evaluating weighting schemes for adult image detection using bag of visual words', ICT Convergence (ICTC), 2013 International Conference on, Oct. 2013
- [3] Mukherjee, J, 'A survey on image retrieval performance of different bag of visual words

- indexing techniques', Students' Technology Symposium (TechSym), 2014 IEEE, March 2014
- [4] Cai-Zhi Zhu, 'Bag-of-Words Against Nearest-Neighbor Search for Visual Object Retrieval', Pattern Recognition (ACPR), 2013 2nd IAPR Asian Conference on, Nov 2013
- [5] David G. Lowe, 'Distinctive Image Features from Scale-Invariant Keypoints', 'International Journal of Computer Vision', Nov 2004
- [6] Dr. H.B.Kekre, 'Image Retrieval using Texture Features extracted from GLCM, LBG and KPE', International Journal of Computer Theory and Engineering, Vol. 2, No. 5, Oct 2010