

## 과제 3: Scene Recognition with Bag of Words

김의찬

국민대학교 전자공학부

Uichan8@naver.com

### 요 약

전체적인 이미지의 정보를 벡터로 추출하는 BoW, Resize 를 구현하고 벡터들을 분류(클러스터) 해주는 SVC, KNN 을 구현하여 이들의 성능을 비교합니다.

### 1. 서론

이번과제의 sift 기술자는 어느 한 점에서의 기술자를 추출했습니다. 이를 통해 같은 장면을 찍은 다른 두 사진의 같은점을 matching 했습니다.

이번에는 사진 그 자체의 기술자를 뽑아내어 이미지가 어떤 장면은 찍었는지 분류하고자 합니다.

이 문제를 해결하기 위해서는 두 가지가 필요합니다.

1. 이미지에서 벡터를 추출 합니다.(BoW, Resize)
2. 추출된 벡터를 가지고 cluser 을 시행합니다.(SVC, kNN)

이들을 구현하고, 비교하여 각각의 장단점을 확인하고자 합니다.

### 2. 과제 수행 내용

#### 2.1 get tiny image

이미지를 16x16 사이즈로 바꾼 후 1 열로 나열하여 벡터를 생성합니다.

1. 이미지를 가우시안 블러합니다.
2. 이미지를 16x16 으로 줄입니다.
3. 이미지를 1 열로 나열하여 벡터로 만듭니다.
4. 평균을 0으로 만들고 단위벡터화 시켜줍니다.

#### 2.2 bag of words

Train 이미지를 네모 격자 단위로 자른 후 벡터로 만듭니다. 이 벡터들을 k means cluster 하여 대표적인 image voca 를 생성합니다. 생성된 단어들을 이용해 히스토그램을 만들고 이를 벡터로 씁니다.

1. 이미지를 격자 단위로 자릅니다.
2. HOG 기술자로 이를 벡터로 표현합니다.
3. 이 패치들의 많은 벡터를 cluster 하여 대표적인 이미지 패치의 벡터를 선출합니다.
4. Test 이미지를 이미지 격자 단위로 자르고

hog 기술자로 표현합니다.

5. 이 기술자들을 이미지 voca 내의 모든 벡터들과 비교하고, 가장 가까운 항목에 추가하여 이미지 voca 크기의 히스토그램을 만듭니다.
6. 단위벡터화 시켜줍니다.

#### 2.3 k - NN

가장 가까운 거리에 있는 k 개의 라벨 정보로 대상의 라벨을 결정합니다.

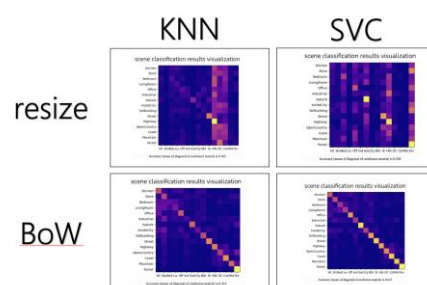
1. 대상 벡터와 train set 의 모든 벡터의 거리를 산출합니다.
2. 가장 가까운 k 개의 점의 라벨을 가져옵니다.
3. 이 중 가장 많은 라벨을 가진 라벨이 대상의 라벨입니다. 그냥 같을 경우 가장 가까운 라벨로 지정합니다. (class 가 none 으로 나오는 것이 오류를 부를 수 있기 때문입니다. K=1 일때는 생각하지 않아도 될 문제입니다.)

#### 2.4 SVM(SVC)

벡터들을 잘 나누는 선을 찾아서 대입 값을 이용해서 분류합니다.

1. 모델 클래스를 생성합니다
2. Train set 으로 학습 시킵니다
3. Test set 을 예측합니다

### 4. 실험 결과 및 분석



	kNN	SVC
resize	18.30%	20.40%
Bow	47.40%	63.70%

kNN 보다 SVC 의 성능이 더 좋고 단순 resize 보다 BoW 의 예측성능이 더 좋았습니다.

## 5. 결론

실험 결과 SVM 과 BoW 의 예측 성능이 가장 좋았습니다. 하지만 BoW 는 resizing 보다 train 특히 k means cluster 과정이 오래 걸렸습니다.

## 참고문헌

[1] 주어진 교재만 참고 함