

# Algorithm and Data Structure Coursework: K-Means Feature for Image Retrieval

Qiwei Feng  
2011011250, IIS-10  
Tsinghua University  
gdfqw93@163.com

Pufan He  
2011011307, IIS-10  
Tsinghua University  
hpfdf@126.com

## ABSTRACT

This project implements a similar image search algorithm (image retrieval) based on multiclass classification and K-Means feature. Our training phase includes image resizing, image patch extraction, patch sampling, PCA whitening, K-Means for patches, feature extraction and multiclass SVM. We use 218-dimension K-Means and RGB, HSV color moment. The training phase takes no greater than one hour in time, 8GB in memory. Finally we obtained 69.49% accuracy on test data classification.

We have made our work open, and the full project codes can be found at <https://github.com/caiwaifung/lastcourse>.

## Keywords

Image Retrieval, Image Classification, SVM, Whitening, K-Means

## 1. INTRODUCTION

## 2. IMPLEMENTATION

We built a system to extracting features from images as well as training model and answering queries of finding related images. The whole system is written in MATLAB.

The system contains feature extracting part and SVM training and testing part. In the feature extracting part, we use both features from K-Means method [3], and the RGB and HSV color moments. We use 200-dimension K-Means features and 18-dimension color moment features (9 for RGB and 9 for HSV). The K-Means method contains patch extracting, patch whitening, patch sampling and K-Means clustering, and feature extrating. The color moment method makes use of 3 common color moments for each channel of RGB and HSV colors. At last, we put the features into SVM and make the classification possible. The related image search process is done by simply finding the closest feature in the same catelogy.

The following subsections includes the details of our algorithms.

## 2.1 Patch Extracting and Sampling

The key idea of K-Means featuring is to find the most common patches in all images, and build features based on that common patches (called centroids). Let  $w$  be the size of the centroids (set to 6 in our implementation). For an  $l$ -by- $m$  image, we will consider all its sub-images (patches) of size  $w$ -by- $w$ ; there are totally  $(l - w + 1)(m - w + 1)$  such patches. Each patch can be represented into a vector of length  $3w^2 = p$ , so one image can be represented into a  $n$ -by- $p$  matrix, where  $n = (l - w + 1)(m - w + 1)$ . We represent all images into this matrix form. This process is called **patch extracting**.

We need a large amount of sampling patches for the K-Means clustering process. We simply random selected  $t \approx 1000000$  patches from all the images' matrices. The sampled patches is a matrix  $P$  of size  $t$ -by- $p$ .

## 2.2 Whitening

## 2.3 K-Means Clustering

We want  $k = 200$  centroids of all sampled patches. We do the K-Means clustering for those patches, and take the final  $k$  "means". The final centroids can be represented by a matrix  $C$  of size  $k$ -by- $p$ .

## 2.4 Feature Extracting

Now consider an  $l$ -by- $m$  image. We know that it has  $n$  patches, so it's an matrix  $G$  of size  $n$ -by- $p$ . We first calculate the distance between each patch and each centroids, resulting in a distance matrix  $D \in \mathbb{R}^{n \times k}$  where  $D_{ij}$  is the distance from the  $i$ -th patch and the  $j$ -th centroid. Then, we normalize each row of  $D$  by subtracting its mean and dividing its standard variance. We negate each elements of the normalized  $D$ , elimtate all negative elements by setting them to zeros, and get the new matrix  $D^*$ . Here,  $D_{ij}^*$  being large means that the  $i$ -th

## 2.5 Color Moment

## 2.6 Multiclass SVM

## 2.7 Related Image Search

## 3. EXPERIMENTS

### 3.1 Data Set

### 3.2 Without Whitening

### 3.3 With Whitening

### 3.4 Final Test

Please see `a.html` under `result.zip` for a more detailed demo.

## 4. CONCLUSION

## 5. REFERENCES

- [1] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [2] A. Coates and A. Y. Ng. Learning feature representations with k-means. In *Neural Networks: Tricks of the Trade*, pages 561–580. Springer, 2012.
- [3] A. Coates, A. Y. Ng, and H. Lee. An analysis of single-layer networks in unsupervised feature learning. In *International Conference on Artificial Intelligence and Statistics*, pages 215–223, 2011.
- [4] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2002.
- [5] D. Lu and Q. Weng. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, 28(5):823–870, 2007.
- [6] J. A. Suykens and J. Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.
- [7] A. Vailaya, M. A. Figueiredo, A. K. Jain, and H.-J. Zhang. Image classification for content-based indexing. *Image Processing, IEEE Transactions on*, 10(1):117–130, 2001.