Locality-constrained Linear Coding [4] for Image Classification

Qi Zhao

June 4, 2018

1. Introduction

This paper presents a simple but effective coding scheme called Locality-constrained Linear Coding (LLC) in place of the vector quantization (VQ) [1] coding in traditional spatial pyramid matching (SPM) [3]. The recent state-of-the-art image classification systems consist of two major parts: bag-of-features (BoF) [2] and spatial pyramid matching (SPM). The BoF method represents an image as a histogram of its local features. It is especially robust against spatial translations of features, and demonstrates decent performance in whole-image categorization tasks. However, the BoF method disregards the information about the spatial layout of features, hence it is incapable of capturing shapes or locating an object. Motivated by, the SPM method partitions the image into increasingly finer spatial sub-regions and computes histograms of local features from each sub-region. Typically, 2^l x 2^l sub-regions, l = 0, 1, 2 are used. Other partitions such as 3 x 1 has also been attempted to incorporate domain knowledge for images with sky on top and/or ground on bottom. The resulted spatial pyramid is a computationally efficient extension of the orderless BoF representation, and has shown very promising performance on many image classification tasks.

2. Description

A typical flowchart of the SPM approach based on BoF is illustrated on the left of Figure 1. First, feature points are detected or densely located on the input image, and descriptors such as SIFT or color moment are extracted from each feature point (highlighted in blue circle in Figure 1). This obtains the Descriptor layer. Then, a codebook with M entries is applied to quantize each descriptor and generate the Code layer, where each descriptor is converted into an RM code (highlighted in green circle). If hard vector quantization (VQ)is used, each code has only one non-zero element, while for soft-VQ, a small group of elements can be non-zero. Next in the SPM layer, multiple codes from inside each sub-region are pooled together by averaging and normalizing into a histogram. Finally, the histograms from all sub-regions are concatenated together to generate the final
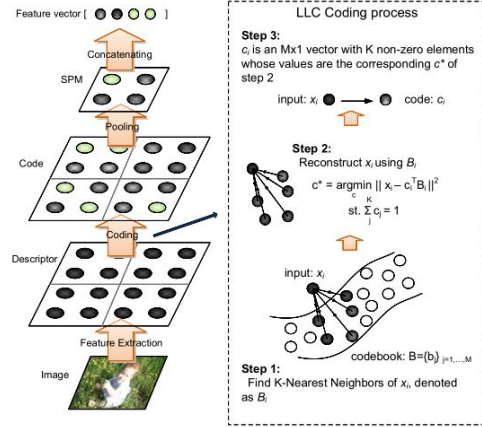

Figure 1. Left: flowchart of the spatial pyramid structure for pooling features for image classification.

representation of the image for classification.

3. Conclusion

The traditional SPM approach based on bag-of-features (BoF) requires nonlinear classifiers to achieve good image classification performance. This paper presents a simple but effective coding scheme called Locality-constrained Linear Coding (LLC) in place of the VQ coding in traditional SPM. LLC is easy to compute and gives superior image classification performance than many existing approaches. LLC applies locality constraint to select similar basis of local image descriptors from a codebook, and learns a linear combination weight of these basis to reconstruct each descriptor.

References

[1] A. Gersho and R. M. Gray. *Vector quantization and signal compression*. Springer Science & Business Media, 2012.

[2] H. Jégou, M. Douze, and C. Schmid. Improving bag-of-features for large scale image search. *International Journal of Computer Vision*, 2010.

[3] S. Lazebnik, C. Schmid, and J. Ponce. Spatial pyramid matching. *Object Categorization: Computer and Human Vision Perspectives*, 2009.

[4] J. Wang, F. Lv, and K. Yu. Locality-constrained linear coding systems and methods for image classification, July 31 2012. US Patent 8,233,711.