Sequential Input Space Carving for visual codebook design

B. Mayurathan[1], U.A.J. Pinidiyaarachchi[2] and M. Niranjan[3]

[1]Department of Computer Science, University of Jaffna, Sri Lanka
[2]Department of Statistics and Computer Science, University of Peradeniya, Sri Lanka
[3]School of Electronics and Computer Science, University of Southampton, Southampton
barathy@jfn.ac.lk, ajp@pdn.ac.lk, mn@ecs.soton.ac.uk

Abstract

In well-known framework for visual scene recognition literature, clustering to construct a visual codebook is an important step, and is usually achieved by k-means clustering. This is known to be computational and performance bottleneck. On current benchmark tasks like the PASCAL VOC challenge, with only 20 classes and a few hundred images per class, the computational problem is one of clustering millions of 128 dimensional vectors into codebooks of a few thousand clusters. K-means being computationally hard on problems of such a scale implies that scaling up to even larger tasks such as the ImageNet challenge, with thousands of classes, becomes impossible. Additionally, there is an inherent compromise between constructing a large codebook, which can potentially retain noise in the data as cluster centres and have the undesirable effect of posing the subsequent classification problem in high dimensions, and a small codebook which loses resolution of the distribution of image features. We present a novel approach to the design of codebooks in patch-based, bag-of-feature visual scene recognition problems. The Sequential Input Space Carving (SISC) approach that we present achieves compact codebooks in a fraction of the computation time needed by the k-means clustering method usually employed in this setting. We demonstrate the performance of the SISC using several recognition tasks including the visual object recognition tasks: PASCAL VOC challenge, MPEG-7 Part-B silhouette image, Caltech-101 and Caltech-256 datasets, human action classification task: KTH and WEIZMANN datasets and texture classification tasks: UIUC and CUReT datasets. Hence, we compare and contrast the recognition performance of SISC evaluated with two different clustering techniques: k-means and Resource-Allocating Codebook (RAC).In all these, the SISC approach achieves classification performances comparable to those reported by other authors, and sometimes outperforms them, in a fraction of the computing time and at significantly smaller codebook sizes.

Author Keywords

Visual codebook, Sequential Input Space Carving, K-means, Mean-shift, Resource Allocating Codebook