An Improved Approach of Iterative Keypoint Selection with Spatial Pyramid Matching for Visual Object Classification

R.M.S. Ranathunga and A. Ramanan
Department of Computer Science, University of Jaffna, Sri Lanka
rmshashis@gmail.com

Introduction

The generic framework of Bag-of-Features (BoF) is depicted in Figure 1. However, one of the problems with this paradigm is the number of keypoints that need to be detected from images to generate the Bag-of-Features is usually very large which causes two problems. First, the computational cost during the feature vector generation step is high and second, some of the detected keypoints are not helpful for recognition. Therefore, this study introduces a framework called Iterative Keypoint Selection (IKS) to select representative keypoints for reducing the computational time to generate the BoF. Also, this work introduces another technique called Spatial Pyramid Matching (SPM) to retrieve more image details in higher resolutions.

Objectives

To make Bag-of-feature representations to be efficient with stable performance by using Iterative Keypoint Selection and Spatial Pyramid Matching techniques.

Methodology

The overall framework is depicted in Figure 2 and the proposed techniques are depicted in Figure 3 and 4.

1. Iterative Keypoint Selection:
   - Resulting in fewer but more representative keypoint descriptors in an image.
2. Spatial Pyramid Matching:
   - Partitioning the image into increasingly fine sub-regions and computing histograms of local features found inside each sub-region. Resulting spatial pyramid is a simple and computationally efficient extension of an orderless BoF image representation.

Experimental Setup

Caltech101: 9,146 images; Xerox7: 17,776 images

- Caltech101: 30 images per class training and testing on the rest.
- Xerox7: 70% training, 30% testing
- Features: Dense SIFT Descriptors
- Vocabulary Construction: K-means algorithm
- Classification: Linear OVA-SVMs
- Distance thresholds in IKS: 0.5, 0.6, 0.7
- L=2 in spatial pyramid matching

Testing Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Dataset</th>
<th>Classification Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Caltech101</td>
<td>36.32%</td>
</tr>
<tr>
<td></td>
<td>Xerox7</td>
<td>84.99%</td>
</tr>
<tr>
<td>IKS</td>
<td>Caltech101</td>
<td>18.15%</td>
</tr>
<tr>
<td></td>
<td>Xerox7</td>
<td>58.72%</td>
</tr>
<tr>
<td>SPM</td>
<td>Caltech101</td>
<td>36.90%</td>
</tr>
<tr>
<td></td>
<td>Xerox7</td>
<td>86.49%</td>
</tr>
<tr>
<td>IKS+SPM</td>
<td>Caltech101</td>
<td>23.12%</td>
</tr>
<tr>
<td></td>
<td>Xerox7</td>
<td>81.61%</td>
</tr>
</tbody>
</table>

Discussion and Conclusion

- IKS extracts spatial-based BoF that can provide greater discriminative power and there is a great reduction in the computational time for generating the BoF and spatial-based BoF.
- SPM improves the performance of BoF approach.
- To improve the performance, a supervised learning based keypoint selection approach can be considered for IKS and Convolution Neural Network (CNN) based features can be used for image classification.