# 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

rmshashi5@gmail.com


## Introduction

The generic framework of Bag-of-Features (BoF) is depicted in Figure 1. However, one of the problems with this paradigm raise is the number of keypoint 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 keypoint are not helpful for recognition. Therefore, this study introduces a framework called Iterative Keypoint Selection (IKS) [4] to select representative keypoints for reducing the computational time to generate the Bag-of-Features. Also this work introduces another technique called Spatial Pyramid Matching (SPM) [3] to retrieve more image details in higher resolutions.

## Objectives

To make Bag-of-feature representation 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:**1776 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

## Traditional Bag-of-Features (BoF) Approach

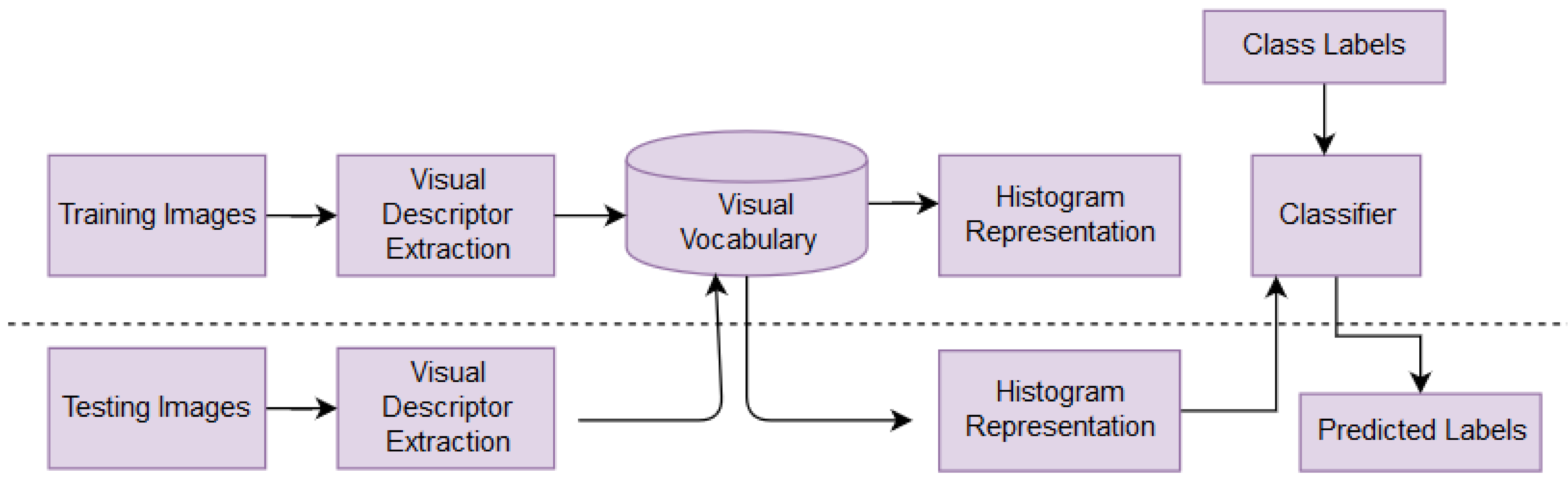


Figure 1. Traditional Bag-of-Features Approach

## Iterative Keypoint Selection (IKS)

Input: training dataset $I_i$ (i.e., the $i$-th image that contains $m$ keypoints)

Output: selected keypoints of $I_i$ (i.e., $SK$)

Reduced set of keypoints (*Reduced_Keypoints*) ← training dataset

Threshold $T$ ← the distance parameter

While any keypoint can be found from *Reduced_Keypoints*

  Get the size of *Reduced_Keypoints*

  Get the random number (*random_number*) between 1 and the size of *Reduced_Keypoints*

  Randomly find a keypoint as the representative keypoint ($RK$) from *Reduced_Keypoints* through *random_number*

  Put $RK$ in $SK$

  For $f$ from 1 to the size of *Reduced_Keypoints*

    If $f$ is not equal to *random_number* then

      Find the distance between $RK$ and the $f$-th keypoint from *Reduced_Keypoints*

      If the distance > $T$ then

        Put the $f$-th keypoint from *Reduced_Keypoints* in a temporary matrix

      End if

    End if

  End for

  *Reduced_Keypoints* ← find all keypoints which are put in the temporary matrix

End while

Return $SK$

Figure 3. Iterative Keypoint Selection Algorithm

## Spatial Pyramid Matching (SPM)

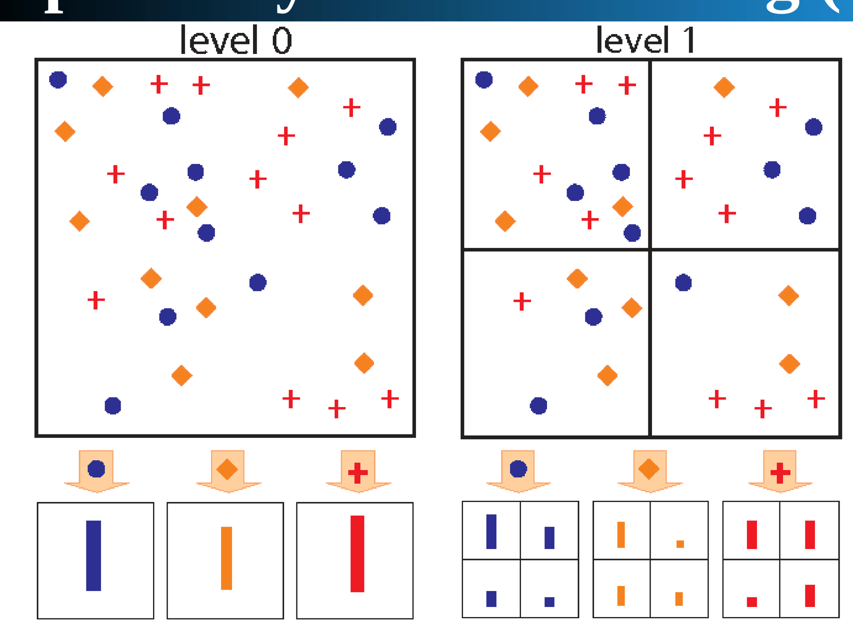


Figure 4. Toy example of constructing a three-level pyramid. The image has three feature types indicated by circles, diamonds and crosses. Subdivide the image at two different levels of resolution. In this study, we choose two levels as not observe any significant in performance beyond two levels.

Construct a sequence of grids at resolutions 0,. . .,L such that the grid at level $\ell$ has $2^{\ell}$ cells along each dimension, where $\ell=0,\ldots,L-1$

The weight associated with each level $\ell$ is given by the equation,

$$\frac{1}{2^{L}-\ell} \qquad (1)$$

For each level of resolution and each channel, count the features that fall in each spatial bin and weight each spatial histogram according to equation 1.

## Proposed methodology

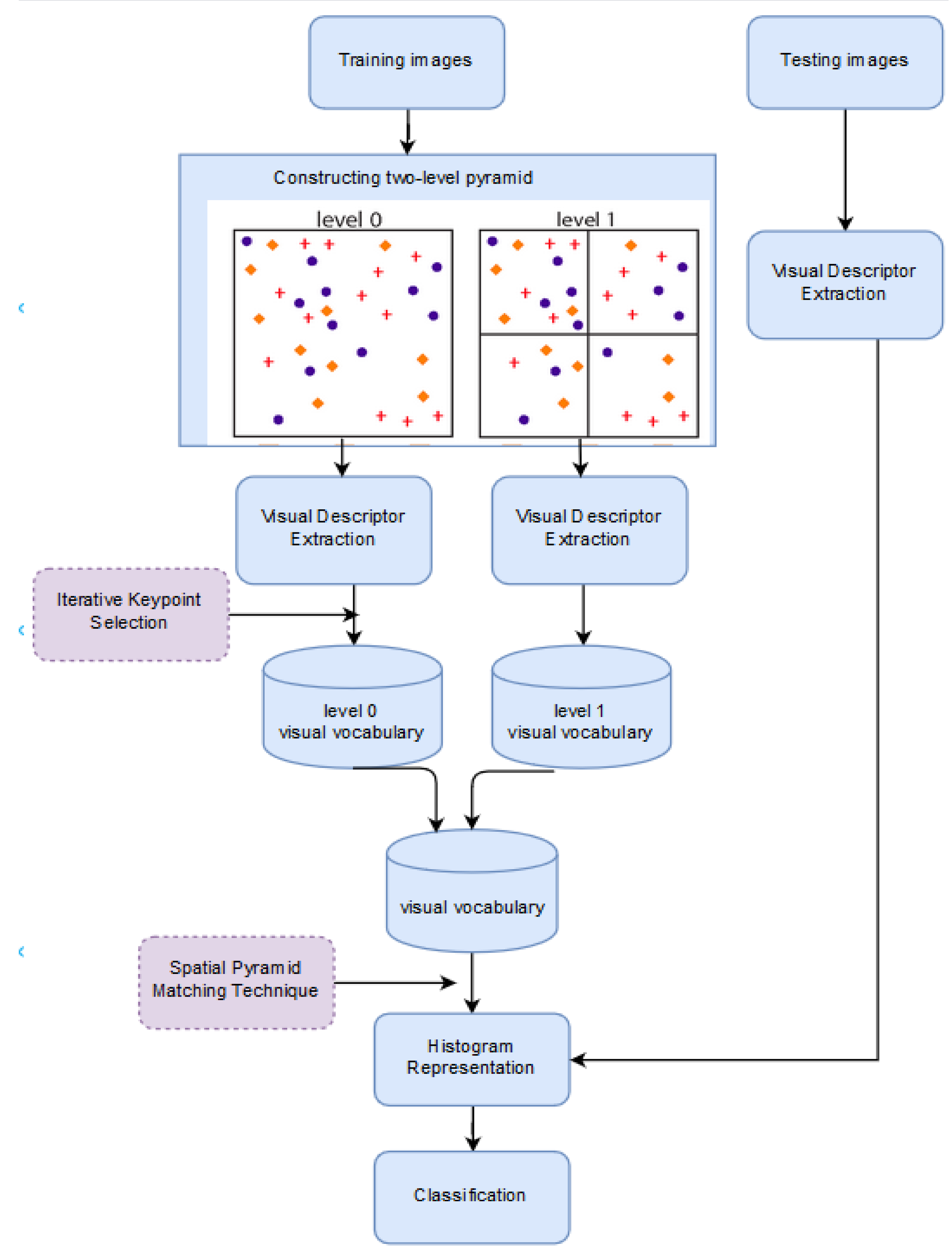


Figure 2. Proposed methodology

## Testing Results

| Method | Dataset | Classification Rate |
|---|---|---|
| Traditional | Caltech101 | 36.32% |
| | Xerox7 | 84.99% |
| IKS | Caltech101 | 18.19% |
| | Xerox7 | 58.72% |
| SPM | Caltech101 | 36.90% |
| | Xerox7 | 86.49% |
| IKS+SPM | Caltech101 | 23.12% |
| | Xerox7 | 81.61% |

Table 1: Comparison of classification rates between Standard BoF approach and proposed techniques; IKS and SPM

## 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.

## References


[1] T. Kirishanthy and A. Ramanan, “Creating compact and discriminative visual vocabularies using visual bits,” in 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2015, pp. 1–6.

[2]V. Vinoharan and A. Ramanan, “Keypoints and codewords selection for efficient bag-of-features representation,” in Future of Information and Communication Conference (FICC). Springer, 2018, pp. 378–390.

[3]S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06),vol. 2. IEEE, 2006, pp. 2169–2178.

[4] W.-C. Lin, C.-F. Tsai, Z.-Y. Chen, and S.-W. Ke, “Keypoint selection for efficient bag-of-words feature generation and effective image classification,” Information Sciences, vol. 329, pp. 33–51, 2016.