Constructing Discriminative and Compact Codebook Using Statistical Measures

Veerapathirapillai Vinoharan and Amirthalingam Ramanan
Department of Computer Science, Faculty of Science, University of Jaffna
{vinoharan,a.ramanan}@jfn.ac.lk

SLSWCS Student Workshop on Computer Science

Introduction

In this work, we propose a two-stage approach to create a discriminative and compact BoF representation for visual object recognition.

- Eliminate ambiguous patch-based descriptors using an entropy-based filtering approach to reduce the features causing false positives in object classification.
- Select the informative subset of codewords based on categorical confidence measures to enhance the discriminative power of the codebook and make it more compact.

Methodology

- A compact visual codebook has advantages in terms of computing efficiency and storage requirement [3, 4].
- Keypoints are selected using an entropy-based filtering method.
- A visual codebook is constructed using resource allocating codebook (RAC) algorithm.
  - Entropy of SIFT descriptor F is computed as,

$$E(F) = -\Sigma_{i=0}^{255} p_i(F) log_2 p_i(F)$$

where $p_i(F) = \frac{|\{k|f_k=i\}|}{128}$; $k = 0, 1, 2, \cdots, 255$.
$F = [f_1, f_2, \cdots, f_{128}]$.

- In-distinctive codewords are eliminated based on categorical confidence measures.
  - The *inter-category confidence* of the i^{th} codeword is represented as follows:

$$C_{inter,i} = \Sigma_{j=1}^{N} max\left(\frac{f_{ij}}{n_i} - \frac{1}{m_i}, 0\right)$$

where K - is the size of the codebook.
N - number of object categories in classification.
f_{ij} - number of j^{th} category training keypoints in the i^{th} codeword, $i = 1, 2, ..., K., j = 1, 2, ..., N$.
n_i - total number of keypoints in i^{th} codeword
m_i - number of object categories in i^{th} codeword

Methodology ...

- The *intra-category confidence* of the i^{th} codeword is represented as follows:

$$C_{intra,i} = \frac{1}{\sum_{j=1}^{N} var(h_{ij})}$$

where K - is the size of the codebook.
N - number of object categories in classification.
h_{ij} - i^{th} codeword value of each image belonging to the j^{th} category in the BoF histogram domain, $i = 1, 2, .., K., j = 1, 2, .., N$.

- *Combined confidence* of the i^{th} codeword is shown as follows:

$$C_{com,i} = \alpha C_{inter,i} + \beta C_{intra,i}$$

where α and β are constant values.

- The overall framework of the proposed method in this work is illustrated in Figure 1.

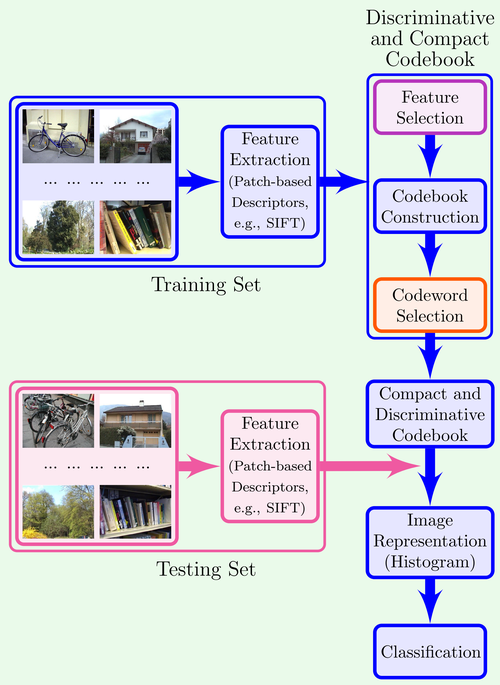

Fig 1: The overall framework of the proposed technique.

Experimental Setup

- We tested our approach on Xerox7, UIUCTex, and Caltech101 image datasets.
- SIFT features were clustered using RAC with r = 0.85, 0.825, 0.86 for Xerox7, UIUCTex, Caltech101 dataset, respectively. We used the linear OVA-SVMs in classification.
- Training keypoints were selected from the initially extracted descriptor with an entropy value E(F)>4.1, 4.4, and 3.8 for Xerox7, UIUCTex and Caltech101 datasets, respectively.
- In the inter-category confidence, the selection criteria is:
  - $\widehat{C}_{inter} = 0$ having a single category, or
  - $\widehat{C}_{inter} > 20^{th} Percentile_{1 \le i \le K}(C_{inter}, i)$
- In the intra-category confidence, the selection criteria is:
  - $\widehat{C}_{intra} > 25^{th} Percentile_{1 \le i \le K}(C_{intra}, i)$
- In the combined confidence, the selection criteria is:
  - $C_{com} > \alpha \widehat{C}_{inter} + \beta \widehat{C}_{intra}$, $0 \le \alpha, \beta \le 1$.

Results

- The performance comparison of BoF approach with entropy-based filtering, the technique with both filtering and codeword selection is shown in Table 1.
- The entropy-based filtering technique eliminates around 45% of the descriptors that outperforms traditional BoF approach.
- On average 80% of the codewords were selected using inter-category, intra-category confidence, and combined confidence from the initially constructed codebook.
- Our proposed technique yields on average 45% of reduction in the initially constructed codebook while maintaining comparable performance with the traditional approach.

Results ...

Table 1: Mean Average Precision (mAP) rate with codebook size using categorical confidences obtained by the proposed method

Approach	Dataset	Before FS		After FS	
		CB	mAP	CB	mAP
Standard	Xerox7	987	67.64	659	69.07
	UIUCTex	1032	93.40	617	95.04
	Caltech101	958	74.71	753	**78.36**
Inter-category Confidence	Xerox7	803	65.77	546	70.11
	UIUCTex	835	93.70	496	**95.84**
	Caltech101	742	75.34	603	76.23
Intra-category Confidence	Xerox7	740	67.63	**494**	69.03
	UIUCTex	774	93.78	**463**	93.95
	Caltech101	694	75.53	**565**	77.32
Combined confidence	Xerox7	833	65.71	598	**70.32**
	UIUCTex	842	93.73	501	95.60
	Caltech101	850	75.16	564	76.41

Discussion and Conclusion

- The central idea of the proposed algorithm is to select representative keypoints and informative codewords, so that the cluster structure of the image dataset can be best respected.
- The proposed method provides an effective way to reduce the BoF representation to low-dimension while maintaining the BoF model to be efficient with stable performance.

References

[1] C. Csurka, R. Dance, L. Fan, J. Willamowski, C. Bray, *Visual Categorization with Bags of Keypoints*, In Workshop on Statistical Learning in Computer Vision, Pages 1–22, 2004.
[2] A. Ramanan, M. Niranjan, *A One-pass Resource-Allocating Codebook for Patch-based Visual Object Recognition*, In proceedings of the IEEE International Workshop on Machine Learning for Signal Processing, Pages 35–40, 2010.
[3] J. Cui, M. Cui, B. Xiao, G. Li, *Compact and Discriminative Representation of Bag-of-features*, In Neurocomputing, Vol. 169, Pages 55–67, 2015.
[4] W. C. Lin, C. F. Tsai, Z. Y. Chen, S. W. Ke, *Keypoint Selection for Efficient Bag-of-words Feature Generation and Effective Image Classification*, In Information Sciences, Vol. 329, Pages 33–51, 2016.
[5] A. Nasirahmadi, S. H. M. Ashtiani, *Bag-of-feature Model for Sweet and Bitter Almond Classification*, In Biosystems Engineering, Vol. 156, Pages 51–60, 2017.