

Image Classification using Frequent Itemset Mining

Vyoma Patel
Computer Engineering Dept.
Sardar Vallabhbhai Patel
Institute of Technology,
Vasad, Gujarat, India

G. J. Sahani
Assistant Professor
Computer Engineering Dept.
Sardar Vallabhbhai Patel
Institute of Technology,
Vasad, Gujarat, India

ABSTRACT

Image classification is one of the most useful and essential research field in computer vision domain and challenging task in the image management and retrieval system. The growing demands for image classification in computer vision having application such as video surveillance, image and video retrieval, web content analysis, biometrics etc. have pushed application developers to search and classify images more efficiently. The main goal of image classification is to classifying image into different classes according to their visual characteristics. In this paper, we propose an approach for image classification by applying frequent itemset mining. Frequent Itemset Mining is used to finding the frequent patterns which is referring as Frequent Local Histograms or FLHs. All Local Histogram information must be maintained for obtaining these FLHs patterns and demonstrate the Bag-of-FLHs based image representation.

The main aim of this research is to use PCA-SIFT (Principal Component Analysis- Scale Invariant Feature Transform) local descriptor for feature extraction. It is more distinctive, robust to image deformations and more compact than the standard SIFT descriptor. The proposed work reduces the overall time of image classification task and maintaining its overall accuracy.

General Terms

Image Classification, Frequent Itemset Mining (FIM)

Keywords

SIFT, PCA-SIFT, BOW

1. INTRODUCTION

One of the challenging tasks is classifying images based on their similarity. Image classification remains to be a major challenge in field of computer vision. With the development of computer, image classification have been widely used in many military areas and civilian areas such as safety monitoring, human-computer interaction, medical diagnosis, vehicle navigation, industrial visual inspection, robot sensing and remote sensing. Despite some research progress, the ability of programs to automatically interpret the content of images is still limited in terms of accuracy and scalability. For example, searching keyword by Google image thousands of images are easily retrieved but typically they contain negative images that do not match with what the user wanted. Image classification consists of database that contains predefined patterns and compares those patterns to classify objects in to proper category.

Frequent itemset mining (FIM) [1] is used to tackle computer vision problem like image classification, action recognition, scene understanding, object recognition and object-part recognition. It is an important data mining technique to extract informative patterns. It allows the construction of compound

features which can capture more discriminative information from images. The main aim is to optimize the process of finding patterns which should be efficient, scalable and can detect the important patterns which can be used in various ways. In image classification, frequent pattern-based classification can achieve high accuracy and good scalability in classifying large datasets [2].

The main objective of image classification is to identify the feature occurring in an image. Local features are of the most successful approaches compare to global feature to object class detection and image classification. Global features like histograms of color, texture and edge information were used to express and classify images but drawback of this approach was that it captured noisy background with interesting target. There exists a variety of local or global image features, but one of the most frequently used local features are SIFT [7] features which is quite robust and invariant to translation, scaling and rotation. Principal Component Analysis (PCA) is a standard technique for dimensionality reduction and has been applied to a variety of computer vision problems, including face recognition [10] feature selection, object recognition. So, by applying PCA to SIFT descriptor another local descriptor called PCA-SIFT [8, 9] which is more accurate and compact than the standard SIFT representation.

In most image classification methods, Bag-Of-Words (BOW) [11] is used for image representation and it can be computed either Locally (LBOW) [12] or globally (BOW) in a neighborhood around an interest point. Local bag-of-words are usually preferred for FIM-based image classification since they result in sparse representation and increased robustness to image clutter. Applying FIM on the bag-of-words representation the new method called bag-of-FLHs [3, 4, 5] . For applying FIM technique to image classification these bags (histogram) need to be converted into sets of items known as transactions using Bag-to-set (B2S) [2] method. This is done by considering each visual word as an item [13].

FIM generates a large number of patterns, of which only a subset is relevant for classification. So, Relevant Pattern mining method selects most discriminative, representative and non-redundant FLHs pattern. By counting the occurrences of these FLHs, each image is represented by bag-of-FLHs [3, 4].

2. RELATED WORK

Frequent pattern mining techniques have been used in computer vision problems, including image classification [2, 13, 14], object recognition and object-part recognition [12]. These methods used different image representation, the way they convert image representation into transactional description which is suitable for pattern mining techniques and selects relevant and discriminative patterns. FIM ignore the issues such that redundancy and repeatability across image of the mined patterns in computer vision.

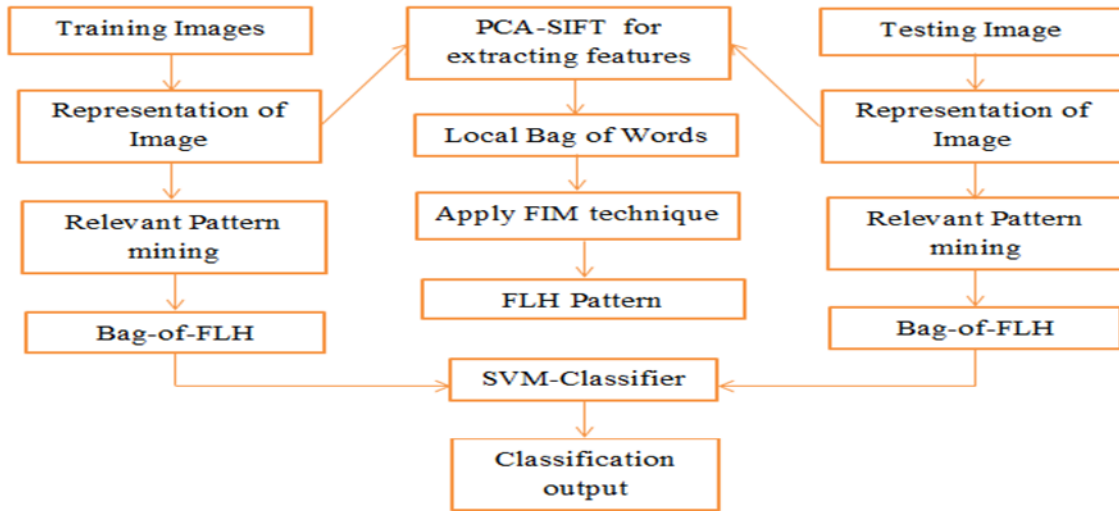


Fig 1: Proposed Work

Many Image Representation techniques were studied in a survey done in [5]. Bag-of-visual words [11] are a standard method for image representation. The BOW can be computed either globally or locally (LBOW) in a neighborhood around an interest point. Local Bag of Words usually preferred than BOWs since they result in sparse representation and an increased robustness to image clutter [3]. Spatial Pyramid matching (SPM) [17] is an advanced method based on orderless bag-of-features. This method works by partitioning the image into increasingly fine sub-regions and computing histograms of local features found inside each sub-region but it is sensitive to image clutter and geometric deformation. The ‘grouplet’ [16] is an another representation technique which captures the structured information of an image by encoding a number of discriminative visual features and their spatial configurations. Bag-of-FLHs [3,4] is a method which uses frequent itemset mining algorithm for finding frequent closed patterns which is compact representation of frequent patterns. It is robust to the occlusion and image clutter. In this paper, we look this Bag-of-FLHs method and use PCA-SIFT descriptor for feature extraction process and obtain promising result for image classification.

3. PROPOSED WORK

In this paper, a new algorithm has been proposed for image classification. PCA-SIFT descriptor is used for feature extraction which reduces overall time of image classification task. The overall workflow can be seen in Fig 1.

3.1 Feature Extraction:

The feature is defined as a function which specifies some quantifiable property of an object, and computed some significant characteristics of the object. One common problem of SIFT descriptor is its relatively high dimensionality compare to PCA-SIFT descriptor. Stages of PCA-SIFT algorithm includes: (i) scale-space peak selection, (ii) keypoint localization and (iii) orientation assignment (iv) gradient map vector generation. The first three stages of the standard SIFT algorithm [7] are applied to detecting the scales and locations of the keypoints. Therefore, the scale invariant and rotation invariant properties of SIFT descriptor is preserve by the PCA-SIFT descriptor. PCA-SIFT can be summarized as[8] : (1) To express the gradient images of local patches pre-compute an eigenspace (2) For given patch, compute its local image gradient (3) Project the gradient image vector using the eigenspace to derive a compact feature vector. PCA-SIFT method calculates the gradient magnitudes at each position of the local region in horizontal and vertical orientations. The last stage of gradient magnitudes represents the feature of the local region.

The 32×32 patch centered on the keypoint is extracted for each keypoint from the image with the nearest scale. The patch is rotated relatively to the dominant orientation of the keypoint. Then, the horizontal gradient and vertical gradient at position (x, y) of the patch are calculated. Then, these gradients are concatenated as a gradient map vector.

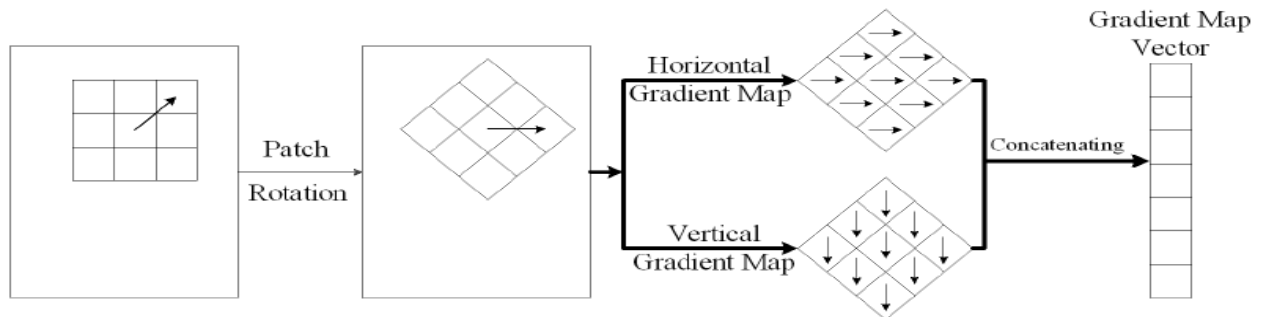


Fig 2: The gradient map vector generation [10]

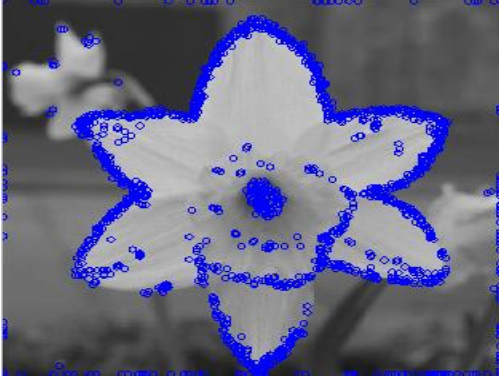


Fig 3: PCA-SIFT keypoint

3.2 FLH-Based Image Representation and Classification

Each image I is described by set of keypoint which is computed by PCA-SIFT descriptor. Here all the descriptor has been clustered using K-means algorithm to obtain a set of visual words. Then for each keypoint k_i compute local bag-of-words (LBOW) using the K nearest neighbours. For each LBOW generates the transactions x which include item as a pair of (l, p) where l is the keypoint label known as the visual word index and p being the frequency of the visual word. The transaction x contains all the items (l_i, p_i) such that the bin corresponding to l_i in transaction x has the nonzero value p_i .

Frequent Local Histogram (FLH) mining: Linear time Closed itemset Miner (LCM) [15] algorithm find the set of FLHs. During the mining process it allows counting the support of an itemset efficiently. LCM supports database reduction. It also generates closed itemsets with no duplication. So, LCM is preferred over the well-known APRIORI algorithm [1]. The input to the algorithm is the set of local histograms and building the transaction from the local histogram as described above. The set of frequent closed patterns is a compact representation of the frequent patterns. This frequent and closed local histogram pattern is referred as a Frequent Local Histogram or FLH.

3.3 Relevant pattern mining

Relevant pattern mining approach is same as described in [3,4]. The FLH set is a new set of mid-level features to represent an image. We need to select best FLH patterns because generated FLH pattern is huge and not all FLH patterns are useful for the image classification. So, relevant pattern mining method selects those patterns which are discriminative, representative and not-redundant.

Relevance Criterion $S(t)$ use two criteria for pattern relevance: a discriminativity score $D(t)$ and a new representativity score $O(t)$.

$$S(t) = O(t) \times D(t) \quad (1)$$

Discriminativity Score finds the discriminative patterns and follows the entropy-based approach [6]. $D(t)$ ($0 \leq D(t) \leq 1$) for pattern t is defined as :

$$D(t) = 1 - \frac{\sum_c p(c|t) \cdot \log p(c|t)}{\log C} \quad (2)$$

$p(c|t)$ is the probability of class c given the pattern t , and it computed as follows :

$$p(c|t) = \frac{\sum_{j=1}^N F(t|I_j) \cdot p(c|I_j)}{\sum_{j=1}^N F(t|I_j)} \quad (3)$$

where I_j is the j^{th} image and N is the total number of images in dataset. $p(c|I)$ is 1 if the class label of I_j is c otherwise it is 0.

Representativity score find patterns with distributions close to the optimal one and defined as:

$$O(t) = \max_c (\exp - [D_{KL}(P(I|t_c^*) || P(I|t))]) \quad (4)$$

Where $D_{KL}(\cdot || \cdot)$ is the Kullback-Liebler divergence between two distributions, $P(I|t_c^*)$ is the optimal distribution for the class c and $P(I|t)$ is the distribution for pattern t . $P(I|t)$ is computed from the frequencies $F(t|I_j)$ of the pattern t :

$$P(I|t) = \left(\frac{F(t|I)}{\sum_j F(t|I_j)} \right) \quad (5)$$

Redundant pattern is used to remove redundant patterns in order to obtain a compact representative set of FLHs. It is define the redundancy $R(s, t)$ between two patterns s, t as follows :

$$R(s, t) = \exp - [p(t) \cdot D_{KL}(P(I|t) || P(I|\{t, s\})) + p(s) \cdot D_{KL}(P(I|s) || P(I|\{t, s\}))] \quad (6)$$

Two pattern t and s are redundant if $P(I|t) \approx P(I|s) \approx P(I|\{t, s\})$ where $P(I|\{t, s\})$ gives the document distribution given both patterns $\{t, s\}$.

After finding the k most relevant and non-redundant FLHs pattern, each image is represented as bag-of-FLHs by counting the occurrences of such FLHs in the image. Let L be such a bag-of-FLHs for the image I_L and M be the bag-of-FLHs for the image I_M . The kernel function $K(L, M) = \sum_i \min(\sqrt{L(i)}, \sqrt{M(i)})$ [3,4] is used to find the similarities between the bag-of-FLHs of L and M . For frequent pattern-based image representation, this kernel function provides good classification accuracies.

LIBSVM[19] is used to train an SVM (Support Vector Machine). The goal of SVM Classification is to produce a model, based on the training data, which can be able to predict class labels of the test data accurately.

4. EXPERIMENTAL RESULTS

In this section we introduce dataset and compare our Bag-of-FLH based method using PCA-SIFT and SIFT algorithm.

4.1 Dataset and Performance Evaluation

To derive the experimental results, we use Oxford-Flowers 17 dataset [18]. Dataset containing 1360 images of 17 classes. Each class contains 80 images. Some sample images of this dataset is shown in fig. 4.

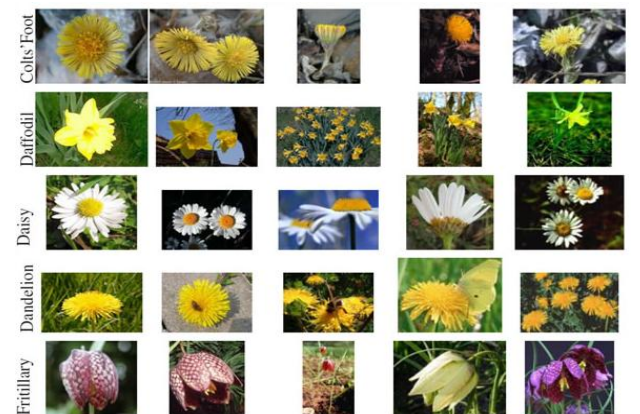


Fig 4. Oxford Flowers 17 dataset Sample images

The experiments performed here calculates visual word dictionary by using K-means algorithm. We use dictionary of 100 words and a local neighbourhood size of 5. PCA-SIFT is used for feature extraction with patch size 32×32 and grid spacing of 8 pixels. Then for each keypoint LBOWs are calculated.

By using LCM [15] algorithm we find Frequent Patterns of images and then calculates Relevant patterns described in section 3.2. Classification result on test image which is shown in Fig 5 of Oxford Flower Dataset is shown in Fig. 6.



Fig 5: Test Image



Fig 6: Classification Images

Here, time consuming computation of feature extraction has been reduced by 15%. As a result, the overall computational time of classification reduces.

Analyses on Feature Extraction time using SIFT and PCA-SIFT descriptor is given in Table 1.

Table 1. Analysis on Feature Extraction time

No. Of Images	SIFT Descriptor(Seconds)	PCA-SIFT Descriptor(Seconds)
50	50.95	31.22
100	85.50	67.44
150	115.89	97.66
200	148.90	130.67

Table 2. Classification accuracy and time on Oxford Flower Dataset

Method	Time (seconds)	Accuracy
Bag-of-FLH using SIFT descriptor	481.269	88.5294%
Bag-of-FLH using PCA-SIFT descriptor	455.567	89.6784%

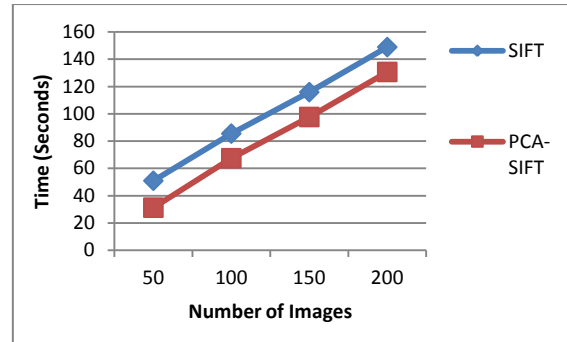


Fig7: Comparison of SIFT and PCA-SIFT Descriptor based on feature extraction time

Comparison of the classification methods based on a time and accuracy are shown in table 2.

5. CONCLUSION

In this paper, we propose a frequent pattern-based image classification and represent each image as Bag-of-FLH model. Bag-of-FLH using SIFT descriptor take more time for feature extraction so using PCA-SIFT descriptor reduces the time of feature extraction process and thus reduce the overall time of image classification task and maintaining its accuracy. In future work use more visual categories to achieve high efficiency for image classification.

6. REFERENCES

- [1] Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases. In: VLDB, pp. 487–499 (1994).
- [2] Kim, S., Jin, X., Han, J.: Disiclass: discriminative frequent pattern-based image classification. In: Tenth Int. Workshop on Multimedia Data Mining (2010).
- [3] Fernando, B., Fromont, É., and Tuytelaars, T. “ Effective use of frequent itemset mining for image classification.” Computer Science-Springer, Volume.7572, 214–227,2012.
- [4] Fernando, B., Fromont, É., and Tuytelaars, T. “ Mining Mid-level Features for Image Classification.” International Journal of Computer Vision-Springer, Volume. 108, 186-203, 2014.
- [5] Vyoma Patel and G. J. Sahani. “A Survey on Image Classification using Data Mining Techniques.” International Journal for Scientific Research & Development, Volume. 2, Issue 10, 2014.
- [6] Cheng, H., Yan, X., Han, J., Hsu, C.W.: Discriminative frequent pattern analysis for effective classification. In: ICDE, pp. 716–725 (2007).
- [7] David, G. L. “Distinctive Image Features from Scale-Invariant Keypoints.” International Journal of Computer Vision-Springer, Volume.60, Issue 2, 91–110,2004.
- [8] Ke, Y., Sukthanka, R. “PCA-SIFT: A More Distinctive Representation for Local Image Descriptors.” Computer Vision and Pattern Recognition-IEEE, Volume.2, 506–513, 2004.
- [9] Juan, L.,Gwun, O. “A Comparison of SIFT, PCA-SIFT and SURF.” International Journal of Image Processing, Volume.3, Issue 4, 187-245, 2009.

- [10] Lin S. D., Lin J., and Chiang C. "Using gradient features from scale-invariant keypoints on face recognition." *International Journal of Innovative Computing, Information and Control*, Volume.7, 2011.
- [11] Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C. "Visual categorization with bags of keypoints." *International Workshop on Statistical Learning in Computer Vision, ECCV*, 1–22, 2004.
- [12] Quack, T., Ferrari, V., Leibe, B., Van Gool, L. "Efficient mining of frequent and distinctive feature configurations." *IEEE 11th International Conference on Computer Vision*, 1 – 8, 2007.
- [13] Yuan, J., Wu, Y., Yang, M.: "Discovery of collocation patterns: from visual words to visual phrases." In: *CVPR* (2007).
- [14] Nowozin, S., Tsuda, K., Uno, T., Kudo, T., Bakir, G. "Weighted substructure mining for image analysis." *Computer Vision and Pattern Recognition-IEEE*, 1-8, 2007.
- [15] Uno, T., Asai, T., Uchida, Y., Arimura, H.: Lcm: An efficient algorithm for enumerating frequent closed item sets. In: *FIMI* (2003).
- [16] Yao, B., Fei-Fei, L.: Grouplet: A structured image representation for recognizing human and object interactions. In: *CVPR* (2010).
- [17] Svetlana Lazebnik, C.S., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: *CVPR*, pp. 2169–2178 (2006).
- [18] Image Dataset of Oxford Flowers 17 <http://www.robots.ox.ac.uk/~vgg/data/flowers/17/index.html>
- [19] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM -- A Library for Support Vector Machines", <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.