IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

A study of Image classification using BoF model

Charulata Rathi[1], S.V.Jain[2]

Student, Computer Science and Engg., Shri Ramdeobaba College of Engineering and Management, Nagpur, India[1]

Assistant Professor, Computer Science and Engg., Shri Ramdeobaba College of Eng. and Management, Nagpur, India[2]

Abstract: BOFs model is one of the most prominent, successful and used for the purpose of image classification. In spite of having many advantages such as very less complexity, ability to be scaled and generality, it scum to various drawbacks, which includes local descriptors provided limited semantic descriptors ,vulnerable structures depending upon single visual words and spatial weighting is inefficient. Numerous techniques have been proposed to nullify the effects of the above mentioned disadvantages, such as multiple descriptor extraction, (ROI) detection and spatial context modelling .Although these methods has contributed towards the improvement of BOFs model to little extent but still coherent integration scheme of all the modules is lacking to resolve the above mentioned problems, a unique framework with spatial pooling of various features is proposed in this paper. BOFs model is expanded on three aspects by the proposed model. First, SURF (speed up robust feature) descriptor is used which combines texture and edge based local and global feature together. Next, extraction of spatial context depending upon features required for midlevel image representation is done using geometric visual phrases. Finally, combination of effective and useful spatial weighting technique and smoothed edge map is used to capture the required features of image.

Keywords: Image classification, BoF Model, K-medoid clustering, Image matching, SURF.

I. INTRODUCTION

Computer vision community (CVC) which means training of the computer for capturing an image through proper programming is highly dependent on feature extraction and labelling of image. Process of getting relevant information resources to the needed information from collection of information is known as Information Retrieval (IR). Various types of Information retrieval are Image Retrieval, Text Retrieval, Music Retrieval,3D Retrieval ,Speech retrieval and many more. This paper focuses on Image Retrieval, which is a computer system used for purpose of browsing, searching and pull out relevant images from hefty data base of digital images. Most old-style and common methods of Image retrieval make use of some method of addition of metadata such as captioning, keywords or description to the images so that the retrieval can be accomplished over the marginal note words. BoF model, Image Retrieval based on content, Image Meta Search is some of the search methods commonly used in Information Retrieval. Most commonly used algorithm for image classification is BOFs model [1,2,3,4]. Image understanding is one of the major functions of image classification along with its wide range of application. Few of the multimedia applications of BOFs model are information retrieval, image classification and, scene understanding. One of the main ambitions of BOFs model is providing better representation of images through statics bas, keyed model. To achieve the above mentioned aim, local and global descriptors such as SURF are extracted from images, and for further tuning of the gathered information of the images compass operator is used. Depending upon the information haul out from above declared procedure codebook is fabricated using K-mediod algorithm upon all descriptors, removing most possible noises and forming vocabulary for the database. Finally Normalization of the descriptors onto the codebook results into visual words, and these visual words are statistical histograms for image representation [5]. Advantages of above proposed system are, firstly the time complexity for image Retrieval is reduced as a result of SURF descriptor being used. Secondly, as the proposed system takes into account both local and global descriptors of Image resulting into increased accuracy in the obtained result.

II. LITERATURE SURVEY

In the year 2014, LingxiXie, Qi Tian, Senior Member, IEEE, Meng Wang, and Bo Zhang proposed Spatial Pooling of Heterogeneous Features for Image Classification[7]. In this paper texture and edge based local features of input image are extracted using SIFT (Scale Invariant feature Transform) descriptor, then midlevel image representation upon complementary features is build using geometric visual phrase and finally spatial weighting of the image is calculated using the smoothed edge map to capture the image saliency.

Construction of codebook for database vocabulary is done using K-means algorithm. Although BoF model is successful using SIFT, some of the disadvantages of SIFT are it suffers from Synonymy and polysemy[7,8,9], time complexity is more as compared to SURF, also process complexity is increased for calculating local and global features as two different algorithms are used. K-means algorithm also have some drawbacks such as it is not invariant to non-linear transformation which means with different representation of same data we get different results, unable to handle noisy data and outliners.

In the year 2011, Y. Zhang, Z. Jia, and T. Chen, suggested "Image retrieval with geometry-preserving visual phrases", Bag-of-visual-word (BoV) is one of the most popular

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 2, February 2016

methods of large scale image retrieval [10]. RANSAC is the spatial verification used to provide the ranking for the results acquired by adding spatial information as the post processing step. Due to its enormous cost computational techniques r spatial verification cannot be applied to whole hierarchy of images. Approach in this paper uses Geometry-preserving visual phrases (GVP) to encode more spatial information and hence reducing the cost spatial verification. Memory usage or computational time is increased by using GVP. Only disadvantage of the GVP method is that it is translation invariant.

In the year 2010, J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, proposed "Locality constrained linear coding for image classification," in Proc. Comput. Vis. Pattern Recognit. Better results than traditional BoF model can be obtained by using nonlinear classifiers which is the base concept of proposed paper[11]. Very effective yet simple coding scheme, Locality-constrained Linear Coding (LLC) is used replacing the traditional Vector Quotient (VQ). Final representation is generated by integrated max pooling of the projected coordinates, which are generated using locality constraints to project all the local descriptors into their local coordinate system. Constrained least square fitting problems are solved using K-nearest-neighbour search which process a fast approximated LLC. Bearing the advantage that several frames per second can be processed into the system even if the size of the code book is very high. Despite of the above mentioned advantage, disadvantage of the mentioned K-nearest-neighbour search are value of the parameter K prerequisites to be define first, training of which features to be used for different types of distance that means Distance based learning is not clear and working out cost is quite high because we requisite to calculate distance of each request case to all working out models.

In the year 2010, Darui Li yz, LinjunYangz, Xian-Sheng Huaz, Hong-Jiang Zhangx proposed "Large-scale Robust Visual Codebook Construction" in this paper improved approximate K means algorithm is suggested, as image retrieval system used for web-scale requires very large visual codebook and it is difficult to construct it using commonly adopted K means vector quantization. Approximate K-means algorithm has scalability constraint as it needs high correctness estimated bordering neighbour exploration. The proposed algorithm improves the assignment precision same as randomized K-d trees, without any cost surge.[12]

III.PROPOSED WORK

In the proposed architecture improved BOF model will be used for image classification. Local descriptors such as SURF and global descriptors are extracted from input image, information extracted from various descriptors is used to build the codebook and noise depressing techniques are used resulting into formation of the dataset with compacted pictorial vocabulary. At last, codebook is raised by quantization of descriptor, and image representation is done with help of a statistical histogram derived from visual words. Image retrieval and image classification are some of the major application of BOF model. Following are the steps involve in the proposed work:

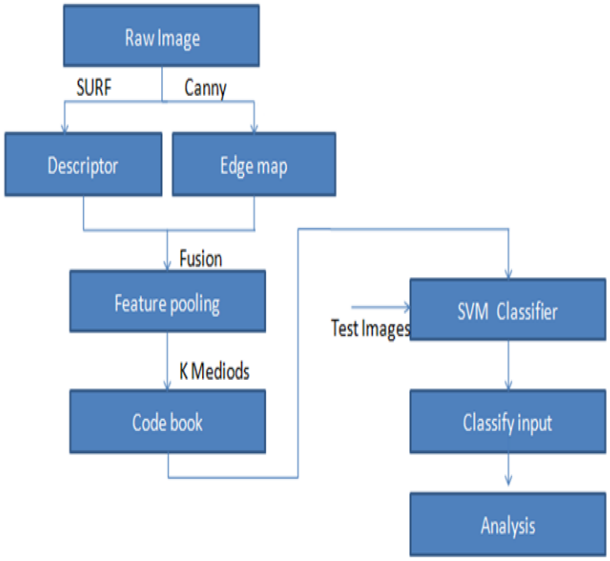

Fig. 1. Flowchart of the proposed work

A. Descriptor extraction

We have an input image I where,

$$I = (B_{ij})_{W \times H} \qquad (1)$$

Where (B_{ij}) is the position of the pixel at coordinate (i,j),w and H are width and height respectively. Local descriptors are extracted from small patches on the image using SURF and global descriptors also known as edge based feature are extracted using canny edge detector. SURF descriptor provide high speed to extract features of the images .

On completion of descriptor extraction process, the input image I can be represented as, set of local and global descriptors, M:

$$M = \{(D_1,I_1),(D_2,I_2),.....,(D_N,I_N)\} \qquad (2)$$

Where D_N denotes the D-dimensional description vector and $I_m$denotes the geometric location of the n-th descriptor.

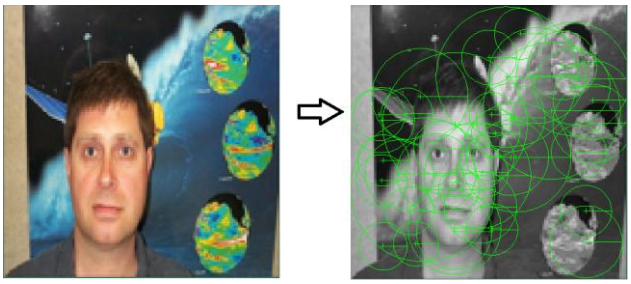

Fig. 2. Example of Descriptor extraction

B. Quantization of Descriptors

Once the descriptors are extracted, then the construction of codebook takes place by compacting of extracted information. K mediod algorithm is used for the purpose of construction of the codebook.

DOI 10.17148/IJARCCE.2016.5254

Then coding is done which refers to the projection of the local descriptors to form a histogram representation. Large quantisation errors takes place because of hard quantization as it uses single codeword for each descriptor. If the descriptor is projected onto the spanned subspace of small codeword group, this technique is known as soft quantization which also results into smaller quantization errors.

C. Feature Pooling

Feature pooling suggests aggregation of all the visual words which are created by quantization of local descriptors and global descriptor so as to form pool of statics for visual words. Most commonly out of the two pooling strategies, Max pooling and Average Pooling any one is used.[7,13].

D. Classification

One of the most crucial step for pre-processing of the data is normalization, in the proposed method the feature vector is normalised using 12 normalization which will result into conversion of all feature vectors into unit length and then they will be used for Image Classification.

Support vector machine (SVM) is one of the most common used classifier as is deals best with very long feature vectors and relatively smaller number of Images. In the proposed method also SVM is used for Image classification. We will use Caltech 101 dataset for classification having 102 classes. We will apply SURF technique on the images from each class and retrieve the key points (local and global descriptors) for each image. Then, by using these descriptors we will generate the dataset of different classes and label them (Training dataset). We will train our SVM classifier by using the above given training dataset. Finally, we give input and check the category in which SVM categorise it.

IV. CONCLUSION

Classification and retrieval of Information from cluster of data still remains an area yet to be explored to its maximum potential .Results obtained using BoF (Bag of Feature) model used for image classification and Information Retrieval yields much faster and stable results when SURF algorithm is used in combination with it. Further improvements in the results achieved can be obtained with more robust BoF model.

ACKNOWLEDGMENT

We sincerely thank the authors mentioned in the references for the valuable information and statistics obtained through their findings.

REFERENCES

[1] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual Categorization with Bags of Key points," in Proc. Workshop Statist. Learn. Comput. Vis., Eur. Con. Comput. Vis., 2004, pp. 1–22.

[2] M. Marszalek and C. Schmid, "Spatial weighting for bag-of-features," in Proc. Comput. Vis. Pattern Recognit., 2006, pp. 2118–2125.

[3] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-offeatures image classification," in Proc. Comput. Vis. Pattern Recognit.,2006, pp. 490–503.

[4] Y. Cao, C. Wang, Z. Li, L. Zhang, and L. Zhang, "Spatial-bag-offeatures," in Proc. Comput. Vis. Pattern Recognit., 2010, pp. 3352–3359

[5] J. Feng, B. Ni, Q. Tian, and S. Yan, "Geometric _p-norm feature pooling for image classification," in Proc. Comput. Vis. Pattern Recognit., 2011, pp. 2609–2704.

[6] Object Classification and Localization Using SURF Descriptors Drew Schmitt, Nicholas McCoy December 13, 2011

[7] Spatial Pooling of Heterogeneous Features for Image Classification LingxiXie, Qi Tian, Senior Member, IEEE, Meng Wang, and Bo Zhang.

[8] D. G. Lowe, "Distinctive image features from scale-invariant key points," Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, 2004.

[9] L. Xie, Q. Tian, and B. Zhang, "Spatial pooling of heterogeneous features for image applications," in Proc. 20th ACM Multimedia, 2012,pp. 539–548.

[10] Image Retrieval with Geometry-Preserving Visual Phrases Yimeng Zhang ZhaoyinJiaTsuhan Chen School of Electrical and Computer Engineering, Cornell University.

[11] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality constrained linear coding for image classification," in Proc. Comput. Vis. Pattern Recognit., 2010, pp. 3360–3367.

[12] D. Li, L. Yang, X. Hua, and H. Zhang, "Large-scale robust visual codebook construction," in Proc. ACM Multimedia, 2010, pp. 1183–1186.

[13] L. Xie, Q. Tian, and B. Zhang, "Spatial pooling of heterogeneous features for image applications," in Proc. 20th ACM Multimedia, 2012, pp. 539–548.

BIOGRAPHIES

Charulata Rathi completed her Bachelor of Engineering in Computer science and Engineering in 2014.She is pursuing her Masters in Technology in Computer Science and Engineering From Shri Ramdeobaba College of Engineering and Management, Nagpur-440013. Her areas of interest include Image Classification, Information Retrieval.

Professor Shweta Jain received the Masters in Technology in Computer Science and Engineering from Nagpur University in 2009 as a first merit holder. She is currently Assistant professor in computer science and engineering department at Shri Ramdeobaba College of Engineering and Management Nagpur. She has a total teaching experience of around 13 Years. Her research interests include Pattern Recognition, Digital Image Processing and Machine Learning.