International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 04 | Apr-2016 www.irjet.net p-ISSN: 2395-0072

Improving Content based Image Retrieval System Using SIFT and SVM

Kimaya S. Meshram[1], Dinkar L. Bhombe[2]

[1]P.G. Scholar, Dept. of E and TC, SSGMCE, Shegaon, Maharashtra, India,

[2] Asso. Professor, Dept. of E and TC, SSGMCE, Shegaon, Maharashtra, India.

---***---

Abstract - *The current improvement in the digital storage media, image capturing devices like scanners, web cameras, digital cameras and rapid development in internet provide a huge collection of images. This leads to the retrieval of these images for visual information efficiently and effectively in different fields of life like medical, medicine, art, architecture, education, crime preventions. To achieve this purpose many image retrieval systems have been developed. Image Retrieval and Content based image retrieval (CBIR) system since it is now a big need of society. CBIR is a method for finding similar images from large image databases. Different approaches are used for Content based image retrieval, out of which Scale invariant feature transform is very popular. In this paper, we discuss the Content based image retrieval by SIFT and SVM. Results obtained in several experiments proposed to evaluate the application of Scale Invariant Feature Transform (SIFT) in content based Image Retrieval (CBIR) systems.*

Key Words: *Content Based Image Retrieval (CBIR), Scale Invariant Feature Transform (SIFT), Support Vector Machine (SVM).*

1. INTRODUCTION

The field of Image Retrieval concerned with the study of searching and browsing digital images from database collection. Different Approaches are used for the image retrieval. Free browsing, keyword based image retrieval or text based image retrieval; these different approaches are previously used in image retrieval. But these methods having some drawbacks. Because of its limitations and scalability problem, the new image retrieval method is emerged named as Content based Image retrieval (CBIR). Though there is a progress in image retrieval, the performance is still not satisfied, since the presence of semantic gaps. Semantic gap is defined as the gap between low level feature and High level feature. Many techniques are used to capture scene semantics in image processing and computer vision [9]. Out of which Scale Invariant feature transform (SIFT) is used in many applications [10]. Prof. David Lowe proposed SIFT Algorithm in ICCV 1999[1] after it is refined in 2004 in IJCV [3], which is cited more than 31,000 times till now. SIFT is used in various applications. Let's have an introduction to SIFT Algorithm.

2. The SIFT APPROACH

Following are the four major stages of computation in SIFT Algorithm used to generate the set of image features:

2.1 Scale-space extrema detection:

All scales and image locations are searched in the first stage of computation and it seeks to identify the potential interest point by difference-of-Gaussian (DoG) function it is implemented efficiently.

The scale space of an image $I(x,y)$ is computed by the convolution of an image with variable-scale Gaussian.

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \tag{1}$$

Where, $*$ is the convolution operation in x and y, $I(x,y)$ is an input image, $G(x,y,\sigma)$ is variable-scale Gaussian and $L(x,y,\sigma)$ is scale space of an image.

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{2}$$

Difference of Gaussian is an algorithm for detecting edges in an image. Which help to identify potential interest points that are invariant to scale and orientation. DoG i.e. Difference of Gaussian function $D(x,y,\sigma)$ can be computed from the difference of two nearby scales separated by a constant multiplicative factor k.

$$D(x,y,\sigma) = \big(G(x,y,k\sigma) - G(x,y,\sigma)\big) * I(x,y) \tag{3}$$

$$D(x,y,k\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \tag{4}$$

2.2 Keypoint localization:

In the second stage, a detailed model is used to determine location and scale it at each candidate location. Based on the measure of keypoints stability they are selected, keypoints resistant to image distortion are stable keypoints.

For detecting local maxima and minima of DoG each point is compared with its eight neighbors on same scale and nine neighbors in the scale above and below as shown in fig 1.

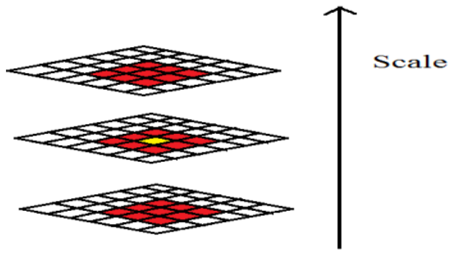

Fig-1 Keypoint localization at different scales.

2.3 Orientation assignment:

In the third stage, SIFT computes the direction of each image gradient. Based on local image gradient directions, one or more orientations are assigned to each keypoint location; its feature vector can be represented to this orientation so that rotation invariance can be achieved.

Keypoint orientation can be calculated by orientation of histogram of local gradients from the closest smoothed image $L(x, y, \sigma)$. The gradient magnitude $m(x, y)$ and orientation $\theta(x, y)$can be computed using pixel differences for each image sample $L(x, y)$.

$$\frac{\partial L}{\partial_x}(x, y) = L(x+1, y) - L(x-1, y) \quad (5)$$

$$\frac{\partial L}{\partial_y}(x, y) = L(x, y+1) - L(x, y-1) \quad (6)$$

$$m(x, y) = \sqrt{\frac{\partial L}{\partial_x}(x, y)^2 + \frac{\partial L}{\partial_y}(x, y)^2} \quad (7)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{\frac{\partial L}{\partial_y}(x,y)}{\frac{\partial L}{\partial_x}(x,y)}\right) \quad (8)$$

2.4 Keypoint descriptor:

In the fourth stage of computation, the local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation which allows for significant levels of local shape distortion and change in illumination. This approach has been named as the Scale Invariant Feature Transform (SIFT) because it transforms image data into scale-invariant coordinates relative to local features. Fig 2 shows the computation of the keypoint descriptors. A 2x2 descriptor array is computed from an 8x8 set of samples is as shown in figure 2, whereas in some experiments 4x4 descriptors are computed from a 16x16 sample array [3,11,14].

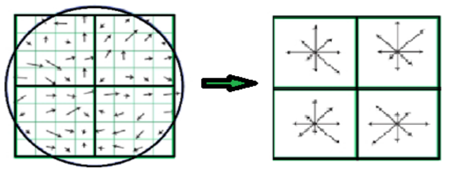

Fig-2 computation of the keypoint descriptors.

By using equation (1),(2),(3),(4),(5),(6),(7) and (8) we have calculated the major stages of SIFT Algorithm.

3. SUPPORT VECTOR MACHINE (SVM)

The Support Vector Machine (SVM) is a state-of-the-art classification method introduced in 1992 by Boser, Guyon, and Vapnik. SVM are an innovative approach to constructing learning machines that minimize generalization error [6]. SVMs belong to the general category of kernel methods. And a kernel method is an algorithm that depends on the data only through dot products. This is the case; the dot product can be replaced by a kernel function which computes a dot product in some possibly high dimensional feature space. It has two advantages: First; the ability to generate non-linear decision boundaries using methods designed for linear classifiers. And second; the use of kernel functions allows the user to apply a classifier to data that have no obvious fixed-dimensional vector space representation. Using SVMs effectively requires an understanding of how they work. When training an SVM the practitioner needs to make a number of decisions: how to pre-process the data, how to the kernel to use or work; and finally; setting the parameters of the SVM and the kernel [5]. Uninformed choices may result in severely reduced performance [7, 8, 12]. Compared with ANN, SVM are faster and can be used with number of genes, are more interpretable and deterministic. SVM offers advantage over other types of multivariate classifiers. They are free of optimization headache of neural network. Because they present the convex programming problem and guarantee of finding global solution [6,13]. Because of these we use SVM classifier.

4. PROJECT FRAMEWORK

When Query image is given to retrieve the same kind of image from the database, SVM is used as classifier which will label the Query image. At the same time the SVM is used for the Image database for labeling the all images in database. Once the Query image and database image is labeled then it will match the label of Query image and Database image to find which class problem was given as Query image. Then the SIFT algorithm is used for finding keypoints of that labelled images. Here dot product matrix, Euclidean distance, mahalanobis distance formula or Earth movers distance formula can be used for matching. But for efficiency in MATLAB, it is cheaper to compute dot products between unit vectors rather than Euclidean distances. So here we used dot products. So that the resultant retrieved images are obtained.

5. FLOWCHART

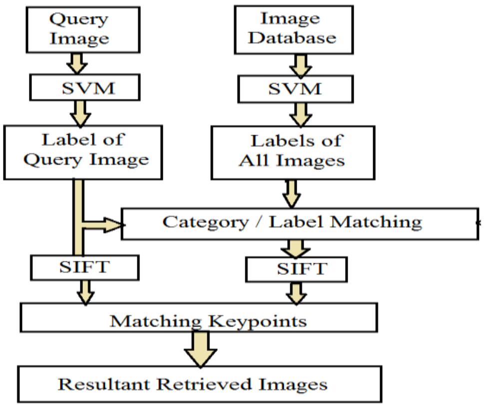

Fig-3 Flowchart of Project work

6. EXPERIMENTAL RESULTS

The experimentation of proposed approach based on SIFT and SVM is carried out over Wang database of 100 images. This image database is prepared of 5 different categories. The 5 different classes we used are owl, bear, butterfly, Flower, sea beach. Each class has 20 images of its own category. Here, by using precision and recall we are finding the output results.

Fig-4 Resultant Retrieved Images

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total no.of irrelevant and relevant images retrieved}}$$

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{All relevant in Database}}$$

By using Precision and recall, we have calculated the final results. Precision is defined as the fraction of retrieved images which is relevant to a query. In contrast, recall measures the fraction of the relevant images which has been retrieved.

Table-1. Precision and recall

Sr. No.	Category	Query Image no.	Precision	Recall
1.	Owl	1	0.8	0.8
2.	Bear	22	0.85	0.85
3.	Butterfly	48	0.9	0.9
4.	Rose flower	77	0.95	0.95
5.	Sea beach	85	1	1

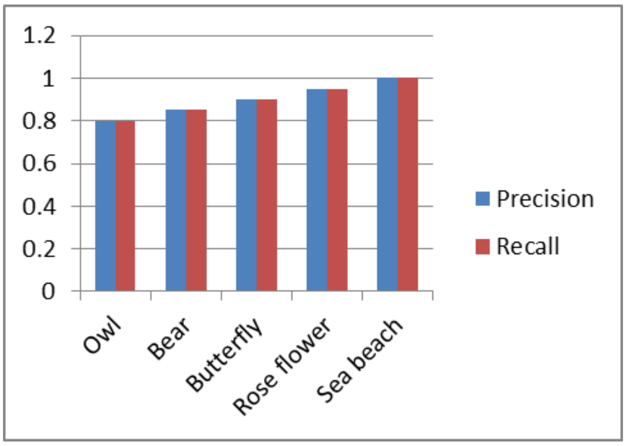

Fig -5: Precision and Recall

7. CONCLUSION

In this paper, we have discussed the results obtained in several experiments conducted to evaluate the application of the SIFT in CBIR tasks along with SVM classifier. In this approach, if no. of matches found in images are Maximum then only the % Match is at its Highest. So that the Resultant Retrieved Images are matched with Query Image with High Accuracy. The Resultant Retrieved Images shown in Results at first are Highly Matched; we can say 100% matched with

Query Image. Precision and Recall is High. SIFT works faster but the training time required is more.

FUTURE WORK

i. In our work, we have used SVM as a Classifier; we may get Better Results if any other Classifier is used along with SVM.

ii. In our work SIFT is used for feature extraction, Along with SIFT we can use other features for Better Results.

iii. In future we would like to Deal with fast processor and Large Memory, so that the retrieving Speed will get increased.

ACKNOWLEDGEMENT

I am thankful to Dr. S. B. Somani, Principal and Dr. G.S. Gawande, HOD (E and TC) of SSGMCE, Shegaon for providing lab facilities. I am highly indebted to Prof. D. L. Bhombe for their valuable Guidance and support during the tenure of the project.

REFERENCES

[1] David G. Lowe, "Object Recognition from Local Scale-Invariant Features", ICCV, 1999.

[2] Arnold W. M. Smeulders, Senior member, IEEE, Mercel Worring, Member, IEEE, Simone Santini, Member,IEEE, Amarnath Gupta,Member, IEEE,and Ramesh Jain, Fellow,IEEE,"Content-Based Image Retrieval at the End of the Early Years", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINEINTELLIGENCE,VOL.22,N0.12,DECEMBER 2000.

[3] David G. Lowe," Distinctive Image Features from Scale-Invariant Keypoints" International Journal of Computer Vision, vol. 6, no. 2,pp. 91-110,2004.

[4] Geoff Gordon," Support Vector Machines and Kernel Methods", ggordon@cs.cmu.edu, June 15, 2004.

[5] B Zitova, J Flusser. Image registration methods: A survey. Image Vis. Comput. 2003; 21(11): 977-1000.

[6] S.N.Sivanandan, S. Sumathi, S.N. Deepa, "Introduction to Neural networks using MATLAB 6.0", McGraw Hill,2006.

[7] Yuan Z, Wu F, Zhuang Y. I-sensor image registration using multiresolution shape analysis. Journal of Zhejiang University Science A. 2006; (4): 549-555.

[8] Tola E, Lepetit V. St local descriptor for dense matching. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society. 2008: 1-8.

[9] Jurandy Almeida, Ricardo da S. Torres and Siome Goldenstein, "SIFT applied to CBIR" Revista de Sistemas de Informacao da FSMA n. 4 (2009) pp. 41-48.

[10] Mamta Kamath, Disha Punjabi, Tejal Sabnis, Divya Upadhyay, Seema Shrawne,"Improving Content Based Image Retrieval using Scale Invariant Feature Transform", International Journal of Engineering and Advanced Technology (IJEAT),ISSN: 2249 – 8958, Volume-1, Issue-5, June 2012.

[11] Dr.K.VELMURUGAN ,"A Survey of Content-Based Image Retrieval Systems using Scale-Invariant Feature Transform (SIFT)", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 1, January 2014.

[12] Sukhmanjeet Kaur , Mr. Prince Verma," Content Based Image Retrieval: Integration of Neural Networks Using Speed-Up Robust Feature and SVM", International Journal of Computer Science and Information Technologies, Vol. 6 (1) , 2015, 243-248.

[13] Pratima Jadhav, Rashmi Phalnikar," Implementing Adaptive Content Based Image Retrieval System Using Neural Network and SIFT", Fourth Post Graduate Conference, iPGCON-2015.

[14] Kimaya S. Meshram, Ajay M. Agarkar," Content Based Image Retrieval Systems using SIFT: A Survey",SSRG International Journal of Electronics and Communication Engineering (SSRG-IJECE) – Volume 2 Issue 10–October 2015.

BIOGRAPHIES

Kimaya S. Meshram is Student of Master of Engineering (Digital Electronics) of Shri Sant Gajanan Maharaj College of Engineering, Shegaon. Received her B.E. degree from Sipna's C.O.E.T,Amravati. Her areas of interests is Image processing, Embedded systems, power Electronics.

Dinkar L.Bhombe received the MS degree in Electronics and Control Engineering from the BITS Pilani, India and doing PhD degree in Electronics engineering from SGB Amravati University Amravati, India. He is currently a Associate Professor at the SSGM College of Engineering Shegaon India.His research interests include Neuro Fuzzy application to Image and Signal Processing.