Advances in Computer Science and Information Technology (ACSIT)
Print ISSN: 2393-9907; Online ISSN: 2393-9915; Volume 2, Number 10; April-June, 2015 pp. 47-50

http://www.krishisanskriti.org/acsit.html

Image Retrieval using SIFT Descriptor

Shipra Singh[1] and Komal Asrani[2]

[1]Babu Banarasi Das University, Lucknow
[2]Babu Banarasi Das Engineering College, Lucknow
E-mail: [1]shiprasingh86@gmail.com, [2]komalasrani@rediffmail.com

Abstract—*The present paper is based on the retrieval of images from large database. Here we have taken 500 images of leaves in white background. Most of the methods mainly focus on global shape features. In this paper we propose to improve retrieval process by considering both global features as well as local features. Shape context is considered as global feature and SIFT (Scale Invariant Feature Transform) [4] descriptors are considered as local features. Finally K- NN algorithm [14] is used for training purpose. Many different leaves have many similar regions so in this case SIFT local descriptors does not work properly in classification of leaves. In order to reduce false matching rates between leaves SIFT has combined with shape context of boundary points. The proposed approach provides better efficiency in comparison to SIFT algorithm or shape context features alone.*

Keyword*: Image Retrieval, SIFT, Shape Context, Plant Recognition.*

1. INTRODUCTION

Image retrieval from large database is now becoming a hotspot area for researchers or scientists. Due to the digital camera or smart phones nowadays it is very easy to capture digital images which results in creation of large databases. Thus there is a requirement of an efficient system for retrieval of images from databases. Traditionally textual annotations are used for describing and retrieving images but this procedure is very tedious, time consuming and inconsistent. As a result, there is a need of such a mechanism which automatically extracts visual features like color, shape and texture from images and retrieves images based on those features. This type of mechanism is called content based image retrieval (CBIR) [1]. This paper focuses on classification of plants through leaf images based on shape context as global descriptors and SIFT as local descriptors of the keypoints detected on the images of leaves. The reason for considering leaf in plant classification as it is easy to capture the shape of leaf in comparison to flowers and other parts of plants. Moreover leaves of the plants are available throughout the season. Here we focus on single leaf image in white background.

2. BACKGROUND

Image retrieval involves two stages: first is the representation of images in the form of feature vectors and second is matching of feature points of query image and images in the database. Feature points matching involves three steps: First is to determine interest points to be matched, second is to find the descriptors of each keypoints and third is to match each points to those in another images by comparing descriptors. David Lowe used a scale invariant detector that finds extrema in the difference of Guassian scale space.He then used a quadratic equation quadratic equation to the local scale space neighbourhood to improve accuracy.The SIFT algorithm describe the feature points has invariance of the translation, rotation, scale and illumination. But when image has many local areas of similarity, The SIFT algorithm will go wrong with large number of false matching points.Thus to improve the efficiency of retrieval the SIFT algorithm is combined with shape context of images in this paper. Dr Zhiyong Wang et al proposed a paper in which he has classified leaf images using SIFT algorithm with shape context features.Mortensen et al. introduced a new SIFT descriptor with a global context which combined the SIFT descriptor with curvillinear shape information to improve the matching accuracy.

3. FEATURE EXTRACTION

1. First of all convert the colored leaf image into grayscale image and apply the SIFT algorithm on the image of leaf and detect the keypoints and for each keypoint.
2. Then for each keypoint two component vector consisting of a SIFT descriptor as local properties and a global context vector is built.
3. Then combine both vectors to construct the proposed descriptor.

4. SIFT ALGORITHM

Lowe (2004) came up with the idea of invariant technology based feature detection method SIFT [4]. There are following steps involved in the SIFT algorithm:

4.1 Scale space extrema detection

This is the step where keypoints are detected. For this the image $I(x, y)$ is convoled with Guassian filters $G(x, y, \sigma)$ at different scales.

$$L(x, y, \sigma)=G(x, y, \sigma) * I(x, y) \quad \text{-------} \qquad \text{--------} \; (1)$$

where ∗ is the convolution operation in x and y and G(x ,y, σ) is a variable -scale Gaussian and I(x,y) is the input image.

To detect stable key point locations in scale space, it is used as scale space extrema based on difference of Gaussian, D(x, y, σ), which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k:

D(x, y, σ)=L(x, y, kσ)-L(x, y, σ) ----- ---(2)

Fig. 1 shows an efficient approach to construction of D(x, y, σ). To detect the local maxima and minima of D(x, y, σ) each point is compared with its 8 neighbors at the same scale, and its 9 neighbors up and down one scale shown in fig.2. If this value is the minimum or maximum of all these points then this point is an extrema.

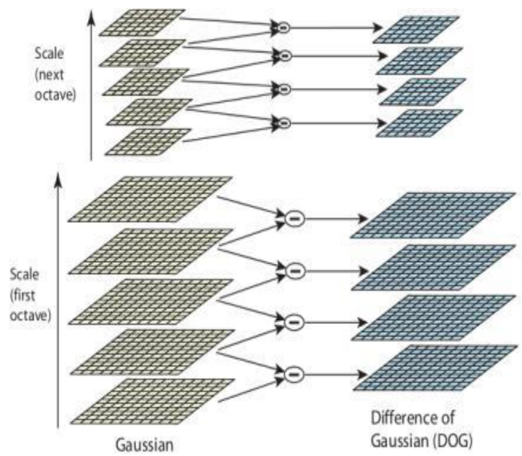

Fig. 1: Scale Space and Difference of Guassian

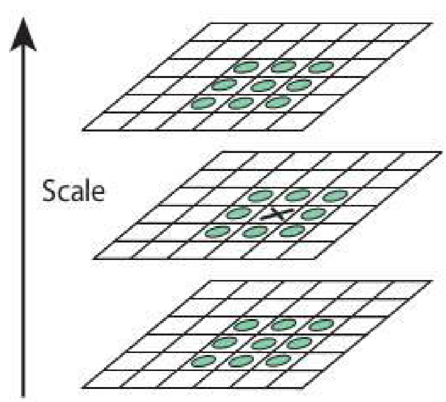

Fig. 2: Neighborhood for extrema detection

Then remove the keypoints with low contrast and edge response point of instability.

4.2 Orientation Assignment

Gradient direction and magnitude of the neighborhood pixels are calculated by following equation:

m (x, y)=((L(x+1, y)−L(x−1, y))2+ (L(x, y+1)−L(x, y−1))2)½ …………… ………..(3)

θ (x, y)=tan−1 ((L(x, y+1)−L(x, y−1))/ (L(x+1, y)−L(x−1, y)))……..…… ……….(4)

where m(x ,y) is the magnitude and θ (x, y) is the direction of keypoint The orientation histogram has 36 bins covering the 360 degree range of orientations.

4.3 Calculation of SIFT Descriptor Vector of given keypoint

The keypoint descriptors typically use a set of 16 histograms,aligned in a 4x4 keypoint descriptor each with 8 orientation bins. This results in a feature vector containing 128 elements**.**

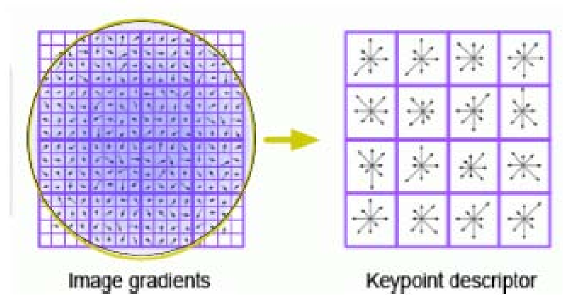

Fig. 3: SIFT Descriptor

4.4 The SIFT Descriptor combined with Global Shape Context Descriptor

For each detected keypoint, it we compute a vector that consists of two parts. One part is the SIFT descriptor of a local feature. Another part is a global texture vector used to distinguish similar local features. Thus a feature vector generated is as follows.

$$F = \begin{bmatrix} \omega S \\ (1-\omega)G \end{bmatrix} \quad \text{----------} \quad (5)$$

where S is a 128-dimensional vector of SIFT, G is a 60-dimensional global vector, and ω is a relative weight factor.

Similar to the local descriptor generation of SIFT, global texture also generates a histogram. The maximum curvature of each pixel is computed, which is defined as

the largest absolute eigen value of a Hessian matrix. For each keypoint log-polar coordinates around it, is established which

divides a circular region, whose diameter is the length of an image diagonal, into a number of regions, and discrete values of angles and radial distances of each

keypoint is computed. Thus, a curvature image can be calculated[16].

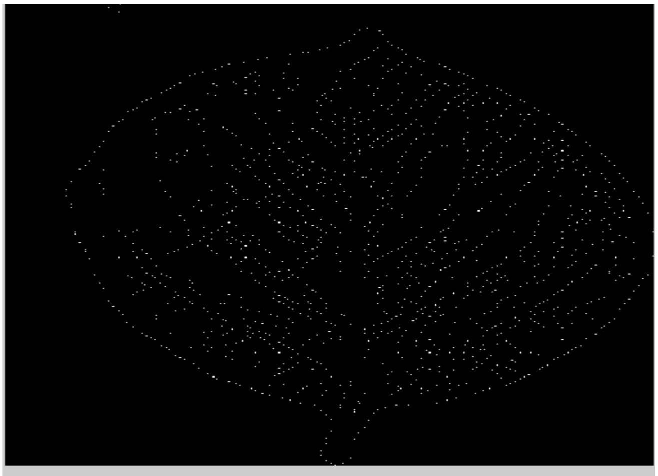

Fig. 4: Leaf Image with SIFT Descriptor

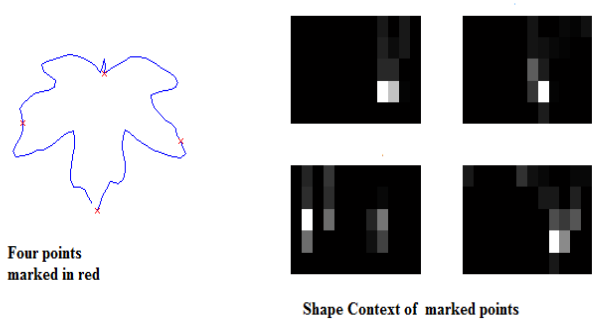

Fig. 5: Shape Context of points marked on leaf image

4.5 Image Matching

Given the definition of our feature descriptor in Eq. (5) and two descriptors, ***Fi*** and ***Fj***, our distance metric is a simple Euclidean distance metric

$$d_L = |L_i - L_j| = \sqrt{\sum_k (L_{i,k} - L_{j,k})^2} \qquad (6)$$

for the SIFT component, ***L***, of the feature vector and a χ2 statistic for the shape context component, ***G***.

$$d_G = \chi^2 = \frac{1}{2}\sum_k \frac{(h_{i,k} - h_{j,k})^2}{h_{i,k} + h_{j,k}} \qquad (7)$$

The final distance measure value is given by

$$d = \omega d_L + (1 - \omega) d_G \quad \ldots\ldots \quad \ldots\ldots\ldots\ldots (8)$$

Experiment

Dataset

The required images for the proposed work are acquired with the help of digital camera. These images are considered in white background and stored in.jpeg format. After storing SIFT descriptor for all leaves, contour lines are extracted of those leaves through any edge detector method like Canny edge detector method. After that shape context method is applied to extract shape context vectors. The subset of database of leaves contains neem, lotus, sadabahar,money plant, grapes, fern etc.The classification is based on KNN method.

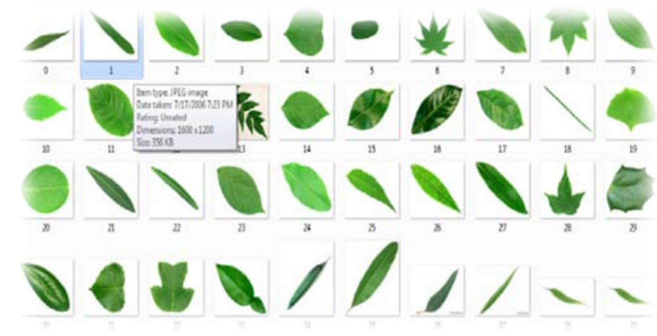

Fig. 6: Subset of database

5. RESULTS

Fig. 7 is showing the retrieval of query image from the database.

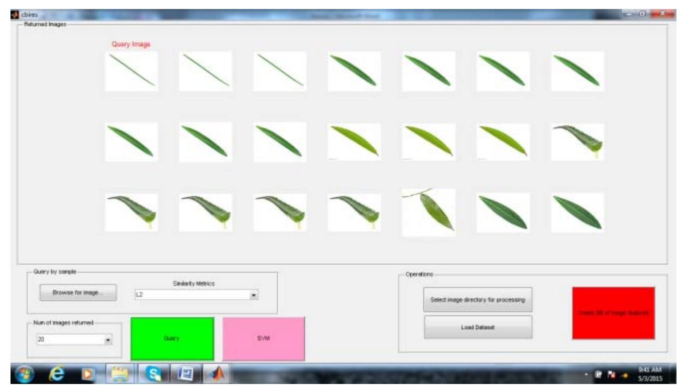

Fig. 7: Retrieval of query image

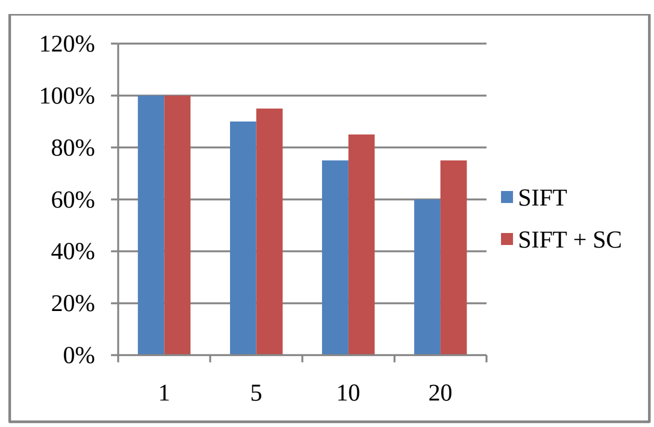

Fig. 8: Graph showing accuracy of system for no. of retrieval

6. EFFECTIVENESS MEASURE

In order to measure the effectiveness of retrieval of the system some measures are suggested below:

Precision–Precision is defined as the ratio of number of relevant retrieved images to total number of retrieved images. [5]

Recall–Recall is defined as the ratio of number of relevant retrieved images to number of all relevant images. [5]

Error Rate- It is defined as number of non relevant images retrieved to total number of images retrieved. [5]

F1 Score–It is the weighted harmonic mean of precision and recall.

7. FUTURE SCOPE

The research work may be improved in the following manner:

Usage of color, texture feature should be considered for future study to improve the efficiency of system.

S. No	Name of leaf	Precision	Recall	Error Rate	F1 Score
1.	Neem	0.475	0.8	0.425	0.5960
2.	Lotus	0.281	0.724	0.719	0.404
3.	Sadabahar	0.530	0.630	0.470	0.5756
4.	Money Plant	0.330	0.68	0.670	0.4443
5.	Banana	0.182	0.437	0.818	0.2569
6.	Grape	0.289	0.525	0.711	0.3727
7.	Aloe Vera	0.241	0.5	0.759	0.3252
8.	Mango	0.219	0.56	0.781	0.3148
9.	Croton	0.425	0.55	0.575	0.4794
10.	Fern	0.616	0.642	0.384	0.627

Fig. 9: Table showing Precision, error rate, recall and F1 score for subset of leaves.

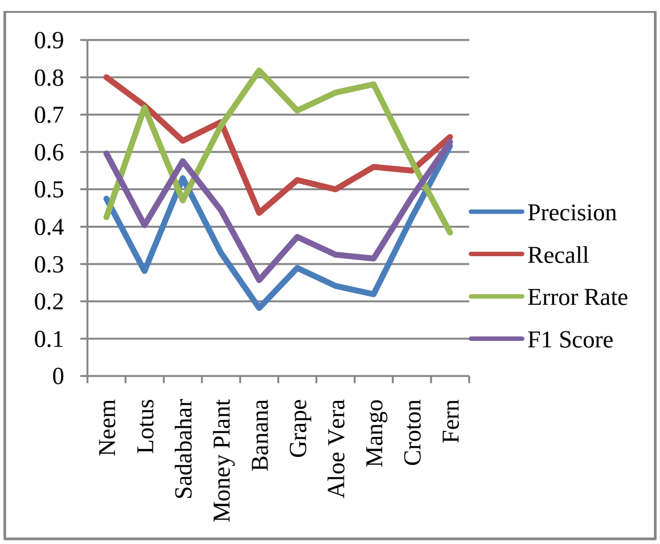

Fig. 10: Graph showing effectiveness measure for subset of leaves

REFERENCES

[1] Michael S. Lew, Nicu Sebe, Chabane Djeraba, Ramesh Jain, "Content-Based Multimedia Information Retrieval: State of the Art and Challenges, ACM Transactions on Multimedia Computing, Communications, and Applications," vol. 2, issue 1, 2006, pp.1-19.

[2] Ying Liu, Dengsheng Zhang, Guojun Lu,Wei-Ying Ma, " A Survey of content based image retrieval with high level semantics", Pattern Recognition 40 (2007), 262-282 available at www.elsevier.com/locate/patcog.

[3] B.Sathya Bama, S.Mohana Valli, S.Raju, V.Abhai Kumar, "Content Based Leaf Image Retrieval (Cblir) Using Shape, Color And Texture Features", ISSN: 0976-5166, Vol. 2 No. 2 Indian Journal of Computer Science and Engineering (IJCSE) 2011.

[4] Lowe, David G., "Object recognition from local scale-invariant features", Proceedings of the International Conference on Computer Vision **2**. pp. 1150–1157. doi:10.1109/ICCV.1999.790410 .

[5] D.H. Ballard and C. M. Brown, 1982, "Computer Vision", Prentice Hall, NewJersey, USA.

[6] Y. Deng, B.S. Manjunath, Unsupervised segmentation of color-texture regions in images and Video, IEEE Trans. Pattern Anal. Mach. Learn. (PAMI) 23(8) (2001) 800-810.

[7] Y.Liu, D.S. Zhang, G. Lu, W-Y. Ma, Region- based image retrieval with perceptual colors, Proceedings of the Pacific-Rim Multimedia Conference (PCM), December 2004, pp.931-938.

[8] F.Jing, M.Li, L.Zhang, H.J. Zhang, B.Zhang ,"Learning in region-based image retrieval, Proceedings of the International Conference on Image and Video Retrieval (CIVR2003), 2003,pp. 206- 215.

[9] J.Z. Wang, J.Li, D. Chan, G. Wiederhold, Semantics-sensitive retrieval for digital picture libraries, Digital Libraries, Digital Library Magazine, vol.5(11), 1999.

[10] E.Chang, S. Tong, SVM_{active}- support vector machine active learning for image retrieval, Proceedings of the ACM International Multimedia Conference, October 2001, pp. 107-118

[11] Jyotismita Chaki, Ranjan Parekh, "Plant Leaf Recognition using Gabor Filter", the ACM International Conference on Multimedia, Ottawa, Canada, 2001, pp. 107-118.

[12] Ricardo Gutierrez-Osuna, Nearest Neighbor, Lecture notes, Texas A&M University**,** Available http://research.cs.tamu.edu/prism/lectures/pr/pr_l8.pdf

[13] J. L. Bentley, "Multidimensional binary search trees used for associative searching." Communications of the ACM (CACM), Vol. 18(9):509–517, 1975

[14] Marius Muja, David G. Lowe, "Fast Approximate Nearest Neighbors With Automatic Algorithm Configuration". In VISAPP International Conference on Computer Vision Theory and Applications 2009, pp 331-340.

[15] A. Kadir, L.E. Nugroho,A.Susanto and P.I. Santosa, "A comparative experiment of Several Shape Methods In Recognizing Plants", International journal of computer science & information technology (IJCSIT), Vol 3, 2011.

[16] Eric N.Mortensen, Hongli Deng, Linda Shapiro, "A SIFT Descriptor with Global Context".

[17] Wang, Zhiyoung, et al. "Leaf Image Classification with shape context and SIFT Descriptiors" Digital image computing techniques and applications,2011 International conference on IEEE,2011.