Object Recognition Using SIFT

Prachi Pundir
Dept of Information Technology
MAIT, Delhi, India

Abstract- *This research paper presents a novel method for the identification of some image using the distinctive invariant features from images. An object recognition mechanism using the Scale Invariant Feature Transform (SIFT) is proposed in this paper. The SIFT is an algorithm in computer vision that detect and describe local or distinctive invariant features in images and it is a tool for matching of different views of an object.SIFT feature descriptor is invariant to uniform scaling, orientation and illumination changes. Preliminary results show that the proposed SIFT-PCA scheme yields promising performance in terms of detection accuracy.*

Keywords- Scale space, LoG Approximation, Difference of Gaussian, Object, Octaves and Scales, SIFT.

I. INTRODUCTION

Object recogniton is one of the important research fields to realize cognitive ability of the computers, and is expected to be applied to Robot eyes or head mounted display. Recently, manual retrieval and classification of the image become difficult as volume of data becomes huge. Computerized object recognition system becomes prominence in such scenario. The problem in the object recognition is to deal with the rotations of the object, scale changes, and illumination changes. Moreover, there is the problem of occlusion that makes the object recognition difficult. SIFT was proposed by David Lowe as a robust feature for these problems, and the object recognition method. The scale invariant feature transform (SIFT) algorithm that extracts features of an image in a manner that is stable over image translation, rotation, scaling, illumination and camera viewpoint. The SIFT has been selected as it is one of the most widely used algorithms for object recognition, that has been employed in many applications such as face/object recognition, robot localization and mapping, 3D-scene modeling, and action recognition. SIFT accepts an $N \times N$ image as input and produces a set of features. The input bandwidth of N2 pixels can be very high for large values of N. This algorithm is most widely used one for the image feature extraction. SIFT extracts image features that are stable over image translation, rotation and scaling and somewhat invariant to changes in the illumination and camera viewpoint. The development of image matching by using a set of local interest points can be traced back to the work of [1] on stereo matching using a corner detector. The Moravec detector was improved by [2] to make it more repeatable under small image variations and near edges. Harris also showed its value for efficient motion tracking and 3D structure from motion recovery [3], and the Harris corner detector has since been widely used for many other image matching tasks. While these feature detectors are usually called corner detectors. [4] showed that it was possible to match Harris corners over a large image range by using a correlation window around each corner to select likely matches. The ground-breaking work of [5] showed that invariant local feature matching could be extended to general image recognition problems in which a feature was matched against a large database of images. Earlier work by the author [6] extended the local feature approach to achieve scale invariance. Then, there has been an impressive body of work on extending local features to be invariant to full affine transformations [7]. Now, in recent years, wide range of techniques are utilized for object recognition. These are color descriptors [8], genetic [9], unsupervised scale invariant learning [10], appearance information [11]. Some of the other techniques were also used in [12-15].

II. FLOWCHART

The important steps used in this work are given below.

1) Read the reference image and input image.
2) Extract the principal components from images.
3) Apply segmentation
4) SIFT approach
4.1 - create the scale space
4.2 - key point localization
4.3 - orientation and assignment
4.4 – key point descriptor
5) Obtain the key points from reference image and input image.
6) Apply Affine Transformation.
7) Image matching
8) Comparison using transformation function

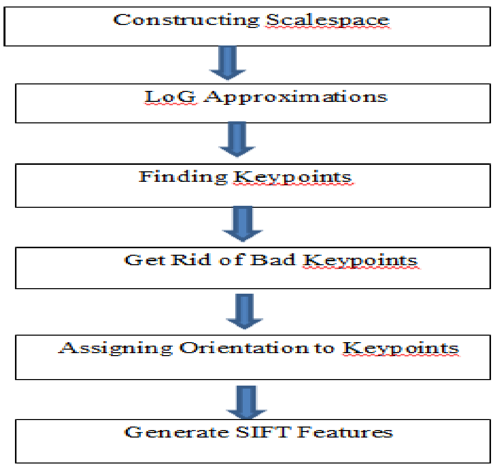

Fig.1.Flow of Proposed Work

III. ALGORITHM AND IMPLEMENTATION

The two images which has to be matched are initially read and then their principal components are calculated, basically PCA and segmentation are done to reduce the data that we are going to process then the process is followed by SIFT, a feature extraction technique and then we match the key points using affine transforms and finally images are matched and detected.

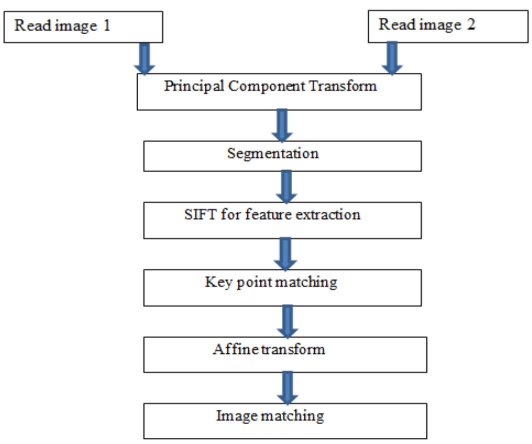

Fig.2.Algorithm and Implementation

IV. PROPOSED METHOD

The proposed work has split into four major phases such as, Extrema detection, keypoint localization, Orient Assignment, key point descriptor generator.

The SIFT algorithm is supposedly able to identify two objects as similar even the object is partly concealed in either one of the images, has changed orientation, or the object is viewed at different angles.

A.Extrema Detection

The first phase examines the image under various scales and octaves to points of the picture that are different from their surroundings. These points are called extrema which is the potential candidates for image features.

B.Keypoint Localization

The KeypointDetection, starts with the extrema and selects some of the points to be key points, that are a whittled down a set of feature candidates. This refinement rejects extrema, which are caused by edges of the picture and by low contrast points.

C.Orientation Assignment

Each keypoint and its neighborhood are converted into a set of vectors by computing a magnitude and a direction for them. It also identifies other keypoints that may have been missed in the first two phases; this is done on the basis of a point having a significant magnitude. The algorithm now has identified a final set of keypoints.

D.Keypoint Descriptor Generation

Keypoint Descriptor Generation, takes a collection of vectors in the neighborhood of each keypoint and consolidates this information into a set of eight vectors called the descriptor. Each descriptor is converted into a feature by computing a normalized sum of these vectors. SIFT provides features characterizing a salient point that remain invariant to changes shown in Fig.3.

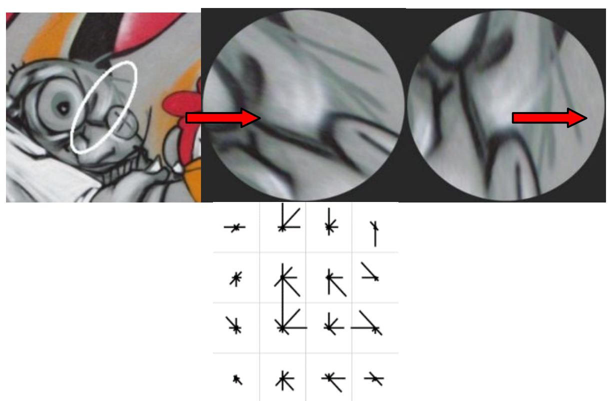

Fig.3. Features characterizing salient point

V. RESULTS & DISCUSSION

The implementation of this algorithm is done in windows platform, with MATLAB– Image processing toolbox. The object recognition flow is executed with the functions like MATCH, SIFT, SHOWKEYS and APPENDIMAGE.

Fig 4a.shows output verified for different image orientation. Fig.4b shows output verified for different size of the input image. Fig.4c. shows output verified for different illumination of the input image, stored image, key points for input image, key points for stored image and also shows the concatenated image so that object can be viewed whether it is present in stored images and also the matches between input and stored image through red line.

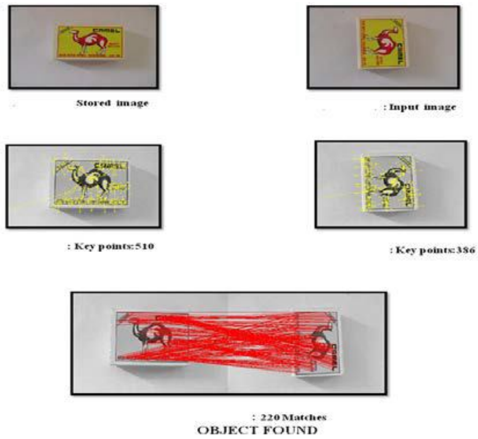

Fig.4aShows output verified for different image orientation

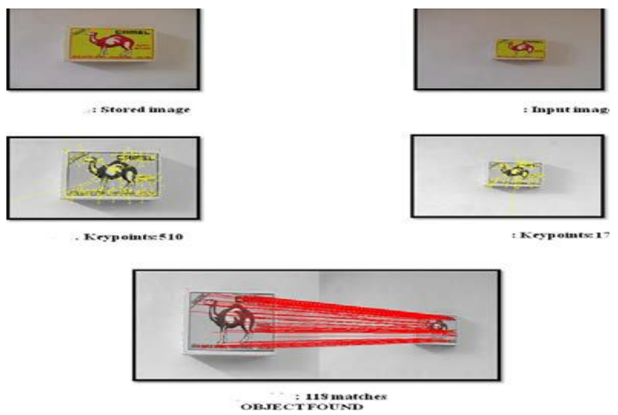

Fig.4b. Shows output verified for different size of the input image.

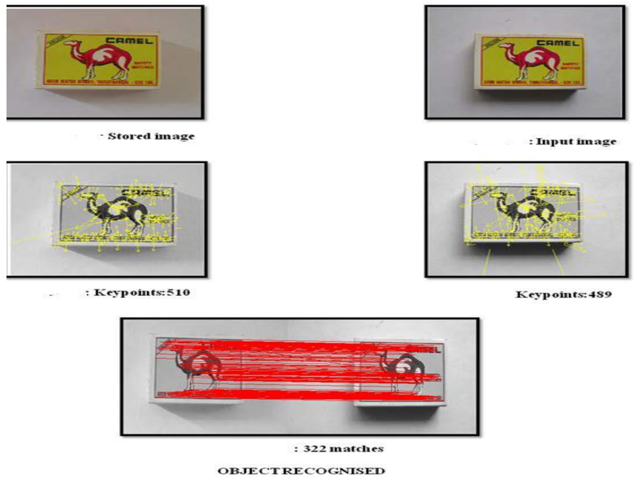

Fig.4c.Shows output verified for different illumination change

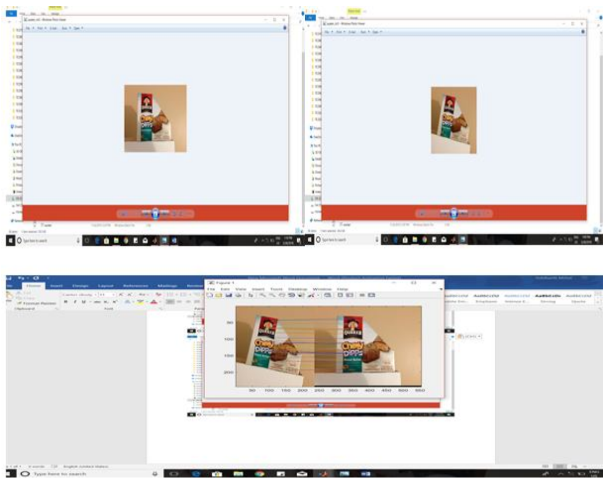

Fig.5.Sample images and output of working project

VI. CONCLUSION

SIFT can be used to detect similar objects in two different images. The SIFT algorithm is supposedly able to identify two objects as similar even the object is partly concealed in either one of the images, has changed orientation, or the object is viewed at different angles. Implementation of such an algorithm could ease the computer vision. As SIFT shows many special features, which are unique in object recognition field, the algorithm could then fulfill the demand of product quality control and object separation in industries. SIFT and SIFT-like GLOH features exhibit the highest matching accuracies (recall rates) for an affine transformation of 50 degrees. After this transformation limit, results start to become unreliable.

Distinctiveness of descriptors is measured by summing the eigenvalues of the descriptors, obtained by the Principal components analysis of the descriptors normalized by their variance. This corresponds to the amount of variance captured by different descriptors, therefore, to their distinctiveness. PCA-SIFT (Principal Components Analysis applied to SIFT descriptors), GLOH and SIFT features give the highest values. SIFT-based descriptors outperform other local descriptors on both textured and structured scenes, with the difference in performance larger on the textured scene.

For scale changes in the range 2-2.5 and image rotations in the range 30 to 45 degrees, SIFT and SIFT-based descriptors again outperform other local descriptors with both textured and structured scene content.

Performance for all local descriptors degraded on images introduced with a significant amount of blur, with the descriptors that are based on edges, like shape context, performing increasingly poorly with increasing amount blur. This is because edges disappear in the case of a strong blur. But GLOH, PCA-SIFT and SIFT still performed better than the others. This is also true for evaluation in the case of illumination changes.

REFERENCES

[1] Moravec, H. 1981. Rover visual obstacle avoidance. In International Joint Conference on Artificial Intelligence, Vancouver, Canada, pp. 785-790.

[2] Harris, C. and Stephens, M. 1988. A combined corner and edge detector. In Fourth Alvey Vision Conference, Manchester, UK, pp. 147-151.

[3] Harris, C. 1992. Geometry from visual motion. In Active Vision, A. Blake and A. Yuille (Eds.), MIT Press, pp. 263-284.

[4] Zhang, Z., Deriche, R., Faugeras, O., and Luong, Q.T. 1995. A robust technique for matching two un-calibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence, 78:87-119.

[5] Schmid, C., and Mohr, R. 1997. Local grayvalue invariants for image retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence, 19(5):530-534.

[6] Lowe, D.G. 1999. Object recognition from local scale-invariant features. In International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157.

[7] Baumberg, A. 2000. Reliable feature matching across widely separated views. In Conference on Computer Vision and Pattern Recognition, Hilton Head, South Carolina, pp. 774-781.

[8] Koen E.A. van deSande, Theo Gevers and Cees G.M. Snoek,“Evaluating Color Descriptors for Object and Scene Recognition,”

[9] George Bebis, Sushil Louis, Yaakov Varol and Angelo Yfantis,“Genetic Object Recognition Using Combinations of Views,” IEEE.Transactions on Evolutionary Computation, Vol. 6, No.,April 2002.

[10] Zachary Pezzementi, ErionPlaku, Caitlin Reyda and Gregory D.Hager, “Tactile-Object Recognition from Appearance Information,”IEEE Transactions on Robotics, Vol. 27, No. 3, pp. 473-487, June 2011.

[11] Bjorn Ommer and Joachim M. Buhmann, “Learning theCompositional Nature of Visual Object Categories forRecognition,” IEEE Transactions on Pattern Analysis and MachineIntelligence, Vol. 32, No. 3, pp. 501-516, March 2010.

[12] Jialue Fan, XiaohuiShen and Ying Wu, “What Are We Tracking: A Unified Approach of Tracking and Recognition,” IEEETransactions on Image Processing, Vol. 22, No. 2, pp. 549-560, February 2013.

[13] Joseph L. Mundy, “Object Recognition in the Geometric Era: ARetrospective,” J. Ponce et al. (Eds.): Toward Category-Level Object Recognition, LNCS 4170, pp. 3–28, 2006

[14] GyuriDorko and CordeliaSchmid, “Object Class Recognition Using Discriminative Local Features,” Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence, October 2004.

[15] K. Matusiak, P. Skulimowski and P. Strumillo, “Object recognitionin a mobile phone application for visually impaired users,” IEEEHSI 2013, Vol. 978-1-4673-5637-4, No. 13, pp. 479-484, June 2013.