International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE) Vol 3, Issue 5, May 2016

Recognition of Objects in Scenes Containing Multiple Objects Using Matching Technique

[1] Dakshayani N [2] Dr. D.J. Ravi
[1] PG Scholar [2] Professor and Head,
[1][2] Department of ECE, Vidya Vardhaka College of Engineering [VVCE]
[1] dakshu307@gmail.com, [2] ravidj@vvce.ac.in

Abstract: **Recognition of objects in a scene containing multiple objects is a primary challenge that has only in recent times been broadly undertaken by computer vision systems. This paper proposes a new technique how to identify a particular object in jumbled scenes, given a reference image of the object. This paper represents an algorithm for recognizing a particular object based on finding feature points correspondences between the reference and the target image. It can detect objects in spite of a scale change or orientation. It is also robust to small amount of occlusion.**

***Keywords*—object recognition, computer vision system, feature matching, SURF, key points**

I. INTRODUCTION

Vision systems are progressively utilized as a part of the fields of home robotics and industrial automation. Real-time object detection and recognition are essential and difficult tasks in computer vision. Among the application fields that make growth in this area, robotics in particular has a strong requirement for computationally capable approaches, as independent systems constantly have to adjust to a changing and unidentified atmosphere, and to be trained and recognize new objects. Point feature matching is a smart solution for such time-critical applications, as new objects can be easily learned online, compare to statistical-learning techniques that involve many training samples. Our approach is connected to modern and capable matching methods and more particularly to, which uses only images and gradient of that image to detect objects. This method produces best results for objects that demonstrate non-repeating texture patterns, which give rise to distinctive feature matches. The proposed algorithm is intended for detecting a particular object, for example, the cup in the reference image, rather than any cup. In this paper, section ii gives an algorithm implementation; section iii gives simulation outcome and section IV gives conclusion**.**

II. RELEVANT WORK

The object recognition problem because of many factors which are essential to the process of object sensing using imaging techniques is difficult. For example geometric transformation which is a result of changing view point [1]. David G. Low [2] presents a Scale Invariant Feature Transform (SIFT) to detect and describe local features in images. From each reference image in the database we extract a set of SIFT key-points. The extracted descriptors are stored in a database. For a given image we first extract it's SIFT key points, then we compare each SIFT descriptor from the image with all descriptors in the database to find the best match. The best candidate match is determined based on Euclidean distance between SIFT descriptors. SIFT descriptor is invariant to uniform scaling, orientation, and partially robust to affine transformation. Moreover using local descriptors such as SIFT object recognition is feasible even in presence of clutter and under partial occlusion. Bay et al. [3] also present a different descriptor so called Speeded Up Robust Features (SURF). It is inspired by the SIFT descriptor. The standard version of SURF is considerably faster than SIFT and in the seminal paper [3] claimed by its authors to be more robust against different image transformations than SIFT.

III. ALGORITHM IMPLEMENTATION

In this paper, our understanding of an advanced image feature method known as Speeded- Up Robust Features (SURF) is presented. SURF is comprised of a Gaussian second derivative mask based feature detector, and a feature descriptor that depends on local Haar wavelet responses. This structure shares many theoretical similarities with the most commonly used feature detector in the computer vision community, called the Scale-Invariant Feature Transform (SIFT) [2].

The Fig. 1 shows the basic steps involved in the proposed algorithm implementation.

Firstly, the reference image and target image will be read. Reference image is the image containing the object of interest and target image is the image containing a scene of multiple objects. After reading reference and target image, feature detection process will be performed on both images.

Feature detection is the process where we automatically examine an image to extract features, which are unique to the objects in the image, in such a manner that we are able to detect an object based on its features in different images.

The processes of object detection can be divided in to 3 steps.

A. Detection

Automatically identify interesting features, interest points this must be done robustly. The same feature should always be detected regardless of viewpoint.

B. Description

Each interest point should have a unique description that does not depend on the features scale and rotation.

C. Matching

Given input image, determine which objects it contains, and possibly a transformation of the object, based on predetermined interest points. In this paper, we are using SURF algorithm to detect features because of it should provide better results, faster than SIFT algorithm. SURF uses a hessian based blob detector to find interest points. The determinant of a hessian matrix expresses the extent of the response and is an expression of the local change around the area.

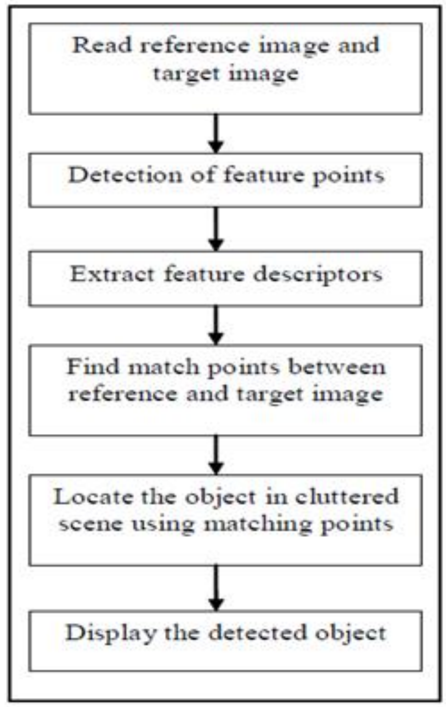

Fig. 1: Algorithm for Recognition of Objects in Jumbled Scene

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \quad (1)$$

Where

$$L_{xx}(X,\sigma) = I(X) * \frac{\partial^2}{\partial x^2}\, g(\sigma) \quad (2)$$

$$L_{xy}(X,\sigma) = I(X) * \frac{\partial^2}{\partial xy}\, g(\sigma) \quad (3)$$

LXX(X, σ) in equation (2) is the convolution of the image with the second derivative of the Gaussian. The heart of the SURF detection is non-maximal-suppression of the determinants of the hessian matrices. The convolution is very costly to calculate and it is approximated and speeded-up with the use of integral images and approximated kernels.

To detect features across scale we have to examine several octaves and levels, where SIFT scales the image down for each octave and use progressively larger Gaussian kernels, the integral images allows the SURF algorithm [1] to calculate the responses with arbitrary large kernels.

The purpose of a descriptor is to provide a unique and robust description of a feature; a descriptor can be generated based on the area surrounding an interest point. The SURF descriptor is based on HAAR wavelet responses and can be calculated efficiently with integral images. SIFT uses another scheme for descriptors based on the Hough transforms. Common to both schemes is the need to determine the orientation. By determining a unique orientation for an interest point, it is possible to achieve rotational invariance. Before the descriptor is calculated the interest area surrounding the interest point are rotated to its direction.

The SURF descriptors are robust to rotations and an upright version, U-SURF, should be robust for rotations ±15 degrees without performing an orientation assignment. I have implemented the upright version, and will not go into further detail on orientation assignment.

Next, match the features including outliers using their descriptors. Finally locate the object in scene using matched points. In order to locate the object in scene estimate geometric transformations i.e. affine transform.

Estimate Geometric Transform calculates the transformation the matched points, while eliminating outliers. This transformation allows us to localize the object in the scene and finally transform the reference image into the coordinate system of the target image. The transformed image indicates the location of the object in the scene.

IV. SIMULATION RESULTS

In the experiments, 5 data sets are used to evaluate the performances of the proposed method for object recognition. The Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8 show the output of each step in implementation for one object. Similarly the experiment was done for 4 other objects.

Fig. 2: Reference Image

Fig. 3: Reference Image

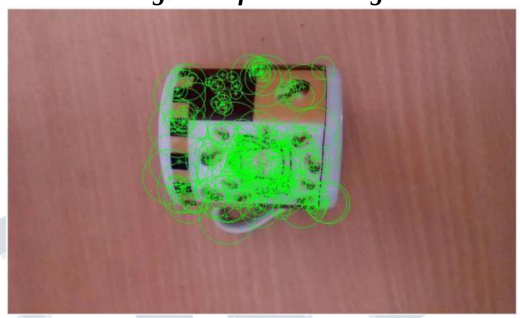

Fig. 4: Feature Points from Reference Image

Fig. 5: feature points from target image

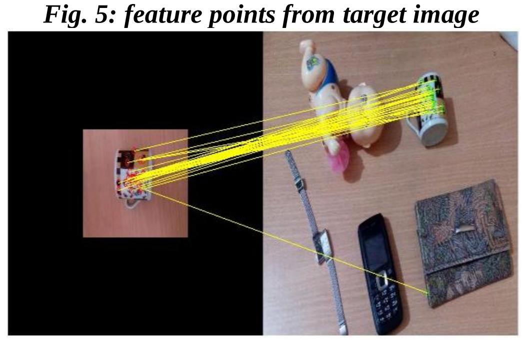

Fig. 6: Matched Points including Outliers

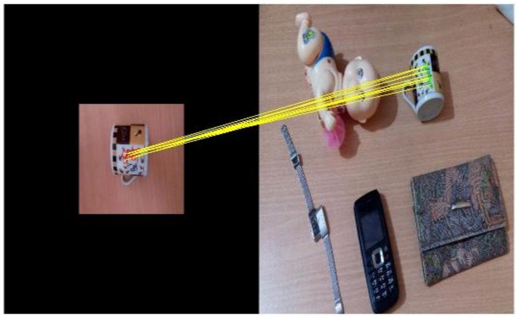

Fig. 7: Matching Points Inliers only

Fig. 8: Recognized Object

V. CONCLUSION

The proposed algorithm is for recognizing a particular object based on finding key point correspondences between the reference and the target image. It can detect objects in spite of changes in scale or rotation. It is also robust to small amount of occlusion. This method of object detection works best for objects that show non-repeating texture patterns, which give rise to unique feature matches

REFERENCES

[1] A. R. Ahmadyfard and J. Kittler, “A multiple classifier system approach to affine invariant object recognition,” in *Computer Vision Systems*, Springer, 2003, pp. 438–447.

[2] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, vol. 2, pp. 1150–1157.

[3] D. L. Waltz, “Generating semantic description from drawings of scenes with shadows,” 1972.

[4] S. U. R. F. (SURF),“Herbet bay, andreas ess, tinne tuytelaars, luc van gool”, Elsevier preprint, 2008.

[5] University of Koblenz-Landau, Tech. Rep., 2009.

[6]Jacob Toft Pedersen,"Study group SURF: Feature detection & description".