SURF and RANSAC: A Conglomerative Approach to Object Recognition

Silica Kole
Department of Computer Science and Engineering, Bharati Vidyapeth's College of Engineering
New Delhi, India

Charvi Agarwal
Department of Computer Science and Engineering, Bharati Vidyapeth's College of Engineering
New Delhi, India

Tripti Gupta
Department of Computer Science and Engineering, Bharati Vidyapeth's College of Engineering
New Delhi, India

Sanya Singh
Department of Computer Science and Engineering, Bharati Vidyapeth's College of Engineering
New Delhi, India

ABSTRACT
In this paper, an object recognition system [7] has been developed that uses SURF(Speeded-Up Robust Features) and RANSAC(Random Sample Consensus) algorithms to identify a series of real-life objects in a given scene using their 2-D images. SURF algorithm has been used for feature detection, extraction and matching. The features are invariant to image scaling, translation, and rotation, and partially invariant to illumination changes and affine or 3D projection[2]. RANSAC algorithm has been used to filter out the results obtained by the SURF algorithm and remove the outliers. Ten different objects have been successfully recognized using this system.

Keywords
SURF, RANSAC, object recognition, computer vision

1. INTRODUCTION
Object recognition needs image features that can be uniquely identified, are distinct and repeatable in the 2-D image of the object. The features must be partially invariant to 3D projective transforms, illumination, and common object variations.[3] At the same time, the features must be distinctive to an extent that they can be used to identify specific objects under various circumstances.

This paper presents the use of a robust feature detection algorithm, SURF (Speeded-Up Robust Features) algorithm for feature detection, extraction and matching. It is a fast and rotation-invariant algorithm used for interest point detection and description [1]. It is also partially invariant to occlusion and illumination changes. SURF algorithm was chosen over SIFT algorithm as it outperforms SIFT in various aspects such as speed, efficiency and space complexity [5].The RANSAC (Random Sample Consensus) algorithm is used to filter out the results obtained by the SURF algorithm.

In the program, the input image of the object that has to be recognized is provided, along with the image of the scene in which the object is to be recognized. First, the SURF algorithm detects various "interest points" in both the images, which are distinct and repeatable so that they can be accurately detected in various orientations and scales of the same object[6]. For each interest point, a feature vector is defined which stores information about the region surrounding the point. After this, the strongest features are selected from both the object and the scene and are matched against each other. This provides a mapping between the features in one image that have also been detected in the other one. There may be some erroneous matches in the scene which do not correspond to the object that is to be detected. To remove these outliers, the RANSAC algorithm is used to filter out the erroneous matches and retain only the interest points that accurately recognize the object. After that, a polygon is drawn around the final interest points that give the boundary of the object in the scene which marks the completion of the task. Ten different objects have been successfully recognized using this system.

2. METHODOLOGY
The algorithms used in this program are SURF and RANSAC algorithms. The entire approach can be explained in the following steps:

1. Converting to gray scale image: The input images of the object as well as the scene are first converted into a gray scale image since the SURF algorithm takes gray scale images as inputs. This is done to improve accuracy so that the object can be recognized, even if it is represented in a different color combination.

2. Feature Detection: The SURF algorithm now identifies local features and interest points which are unique, distinct and repeatable in the object image. The information about the regions surrounding these interest points is stored in a feature vector which stores 64 elements.

3. Feature Matching: The similar kinds of features are identified in the scene. These features are then matched to the corresponding features of the object and a mapping of these features between these two images is stored in a vector.

4. Filtering of Matched points: In feature matching, several erroneous matches might occur. The RANSAC algorithm is used to overcome this problem and refine the matched points. The algorithm takes all the matched points as input, formulates a mathematical model that incorporates the majority of the points, and filters out the remaining points which are considered as outliers [3]. At the end of this step, a set of matched points is found which accurately detect the object in the given scene.

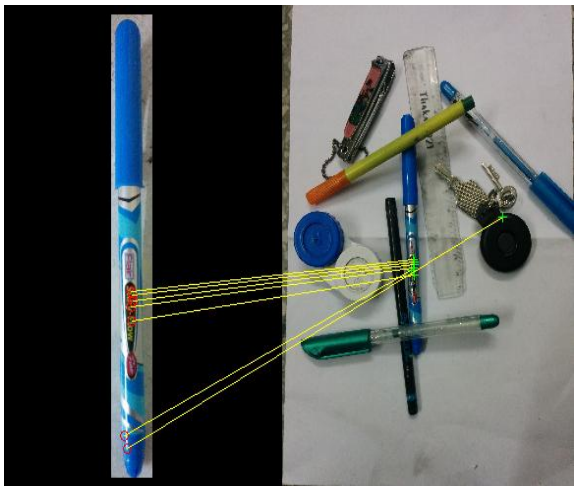
3. SURF vs SIFT

Table 1: Comparision between SIFT and SURF algorithm [5]

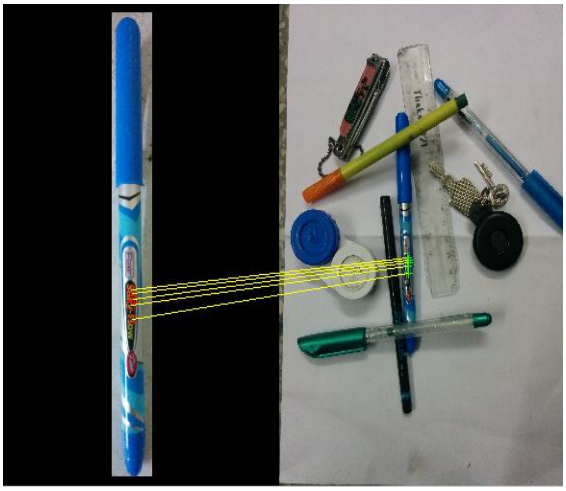
	SIFT	SURF
Average Interest Points	222	431
Runtime	2887.13 ms	959.87 ms
Power Consumption	0.04%	0.02%
Memory Used	102.30 KB	385.02 KB
Effectiveness	0.33	0.40

The main difference between the SIFT and SURF algorithms is the speed of implementation. SURF outperforms SIFT in various aspects. As can be deduced from the above table, the runtime of the SURF algorithm is one third of that of the SIFT algorithm. The power consumption is less and the overall efficiency is high. Owing to these advantages, SURF algorithm was chosen over SIFT algorithm.

4. RESULTS



(a)



(b)

(c)

Fig. 1: (a) Output after applying SURF algorithm (b) Output after applying RANSAC algorithm (c) Final object recognition

As can be seen from Fig.1, the object features were identified and matched with the corresponding features in the scene, using the SURF algorithm. The refining of the matched points was done thereafter using the RANSAC algorithm. Finally, the object was successfully recognized in the given scene. Ten such objects have been accurately recognized using this method so far, under varying circumstances such as variations in rotation, scaling as well as partial occlusion of the identified object. Some of the experimental results have been listed in Table 2, specifying the amount of rotation, scaling and occlusion of the objects in the given scene images.

Table 2: Objects recognized under varying conditions

Object	Rotation	Scaling	Occlusion
Book	180 degrees	25%	5%
Pen	NIL	50%	20%
Toy	30 degrees	70%	5%
Chocolate	90 degrees	25%	10%
Identity Card	310 degrees	20%	NIL
Transparent Bottle	NIL	80%	30%
Watch	20 degrees	50%	5%
Pencil Box	10 degrees	75%	15%
Phone	30 degrees	30%	35%
Wallet	45 degrees	90%	5%

5. CONCLUSION

It can be concluded that a combination of SURF and RANSAC makes a highly efficient object recognition system. SURF algorithm is fast and robust, while RANSAC ensures high accuracy. This system can be used in applications such as navigation and visual surveillance. For instance, landmarks can be recognized using this system which could facilitate the user to navigate through an unknown region. The dataset can expanded to incorporate a large number of identifiable objects as well. Further, it can be used to identify objects in everyday life, such as finding items at a grocery store, or books in a library. Depending on the size of the data set, a sophisticated system for such applications can be developed.

6. REFERENCES

[1] Bay, Herbert, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded up robust features." Computer Vision–ECCV 2006. Springer Berlin Heidelberg, 2006. 404-417.

[2] Lowe, David G. "Object recognition from local scale-invariant features."Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Vol. 2. Ieee, 1999.

[3] Fischler, Martin A., and Robert C. Bolles. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." Communications of the ACM 24.6 (1981): 381-395.

[4] Lowe, David G. "Three-dimensional object recognition from single two-dimensional images." Artificial intelligence 31.3 (1987): 355-395.

[5] Panchal, P. M., S. R. Panchal, and S. K. Shah. "A Comparison of SIFT and SURF." International Journal of Innovative Research in Computer and Communication Engineering 1.2320 (2013): 9798.

[6] Bay, Herbert, et al. "Speeded-up robust features (SURF)." Computer vision and image understanding 110.3 (2008): 346-359.

[7] Alexander Andreopoulos, John K. Tsotsos "50 Years of Object Recognition: Directions Forward", ∗Department of Computer Science and Engineering,Centre for Vision Research, New York University, Toronto