

Object Detection using Template and HOG Feature Matching

Marjia Sultana¹

Department of Computer Science and Engineering
Begum Rokeya University
Rangpur, Rangpur-5404
Bangladesh

Tasniya Ahmed²

Institute of Information Technology
Noakhali Science and Technology University
Noakhali – 3814
Bangladesh

Partha Chakraborty³, Mahmuda Khatun⁴

Department of Computer Science & Engineering
Comilla University, Cumilla - 3506, Bangladesh

Md. Rakib Hasan⁵

Department of Information and Communication Technology
Comilla University, Cumilla - 3506, Bangladesh

Mohammad Shorif Uddin⁶

Department of Computer Science and Engineering
Jahangirnagar University, Savar, Dhaka-1342, Bangladesh

Abstract—In the present era, the applications of computer vision is increasing day by day. Computer vision is related to the automatic recognition, exploration and extraction of the necessary information from a particular image or a group of image sets. This paper addresses the method to detect the desired object from an image. Usually, a template of the desired object is used in detection through a matching technique named Template Matching. But it works well when the template image is cropped from the original one, which is not always invariant due to various transformations in the test images. To cope with this difficulty and to develop a generalized approach, we investigate in detail another technique which is known as HOG (Histogram of Oriented Gradient) approach. In HOG, the image is divided into overlapping blocks of template size and then compare each block's normalized HOG with the normalized HOG of the template to find the best match of the object. We perform experiments with a large number of images and have found satisfactory performance.

Keywords—Computer vision; template matching; HOG; feature extraction

I. INTRODUCTION

In order to reduce human sufferings and to speed up a specific task with more precision, we need to train a machine such that it can perform a specific task without any human interaction. The human eye can easily identify an object from an image; however, it is difficult for a machine to recognize the objects from an image automatically. If we want to recognize an object by a machine itself, we need to train the machine with an efficient object detection algorithm [1]. Therefore, the main objective of this research is to find out a better algorithm for machines to recognize objects in a scene. Here we are focusing on object detection using the template as well as HOG (Histogram of Oriented Gradient) feature-based techniques. The most widely used computer vision-based technologies are needed to resolve the problems of object matching and recognition in the field of image processing and

analysis [2]. For any vision-based image processing application, object detection is the most integrated part [3]. An efficient object detection technique enables us to determine the presence of our desired object from a random scene, regardless of object's scaling and rotation, changes in camera orientation, and changes in the types of illumination. Template matching is one of the approaches of great interest in current times which has become a revolution in computer vision. Another widely used approach of object detection is HOG where matching of extracted features is carried out.

Over the past few years, researchers have come up with new and widely used techniques for identifying and tracking objects. Among them, general Template Matching and HOG are widely used techniques. This study focuses on the correct detection of desired objects using template-based methods. This paper uses HOG technique to find the desired object from a testing cluster image using a patch of a template image. We apply HOG based object detection method on images, observe the results and find the advantages and limitations of the technique. We also overview a comparison results of simple Template Matching and HOG based object detection method.

The rest of this paper is embodied as follows. Section II represents the related works of object detection techniques. Section III describes the small description of the methods named HOG feature and Template Matching. This chapter also includes the methods' flow diagrams and the corresponding algorithms for the methods. Experimental results and analysis of the performance based on some criteria are included in Section IV. Section V presents concluding remarks about the research.

II. LITERATURE REVIEW

Al-Mamun and Yousef investigated different types of methods for the segmentation of images and useful feature extraction from the images. They also proposed a model for the identification of flexible desired objects from an image

with asymmetrical shapes [4]. Two feature detection techniques like Scale Invariant Feature Transform (SIFT) and Speeded UP Robust Features (SURF) were used for image registration [5, 6]. SIFT could detect more number of features but its' speed was not remarkable. But SURF showed better results in the case of both speed as well as performance [5]. A variant of a conventional HOG feature named Edge-HOG was proposed by Ren and Li to detect pedestrian and car [7]. Experimental results indicated that Edge-HOG was two times faster in speed compared to the conventional HOG. Chetan in [8] adjoined two external features like the shape and color of the object to detect an object swiftly and with a comparatively accurate detection rate. The results of the paper pointed out that the performance of the proposed method was comparable to other methods. In paper [9], researchers compared four widely used feature detection techniques namely SURF, Harris, FAST (Features from Accelerated Segment) and FREAK (Fast Retina Keypoint) in terms of accuracy and run time. The paper concluded that FREAK algorithm outperformed the other algorithms based on detection accuracy and time complexity. Multiple same instances were detected from a single template image using a new approach which was based on SIFT method but the scale and rotation invariant method named SURF was also used to extract the rotation and scaling of the features [10]. A method was proposed in [11] to detect moving objects from a video. In that paper background subtraction was used for object detection and two methods namely thresholding and edge detection were used for segmentation. Based on Peak Signal to Noise Ratio (PSNR) value, it was experimented that background subtraction showed better performance compared to thresholding. In [12], Nazil Perveen described various types of template matching techniques and emphasized on the applications of those techniques. Haar Cascade Classifier was used to identify human head by considering head image as an object in [13].

Most of the above papers used a feature-based object detection technique to detect a specific type of object like human, car, weapon, etc. This paper offers generalized techniques to identify the desired object from a source image. Again, none of the above researchers compared the prominent approaches, such as simple Template Matching and HOG feature-based method to compare the performance of these techniques in object detection. The paper includes a comparative discussion of HOG and Template Matching to present the drawbacks and advantages of these techniques.

III. METHODOLOGY

A. HOG Feature-based Method

HOG is a feature descriptor based approach that focuses on the extraction of features. In the feature descriptor of an image, only useful information is extracted and unnecessary information is thrown out. In HOG method the whole image is partitioned into small blocks and then the feature descriptor is completed for each block [14]. After extracting the useful features of each block of the image, those blocks are grouped together and then normalized to obtain contrast normalized features.

HOG feature extraction process: First, convert color images to grayscale. This process minimizes the color information. Next, calculate the value of each pixel's luminance gradient. Then, generate a gradient orientation histogram for each block. This process can get the feature quantity that is powerful in changing the form. Finally, the normalize value of the features is obtained for each block.

The stages of HOG feature extraction and overview of HOG method are depicted in Fig. 1 and Fig. 2, respectively.

Gradient image for a HOG descriptor can be represented in several color models like RGB (Red, Green, Blue), LAB (Color representation where L denotes lightness and A, B represent the color-opponent dimensions) and gamma. As most of the cameras would be RGB cameras, we can consider the input image to be RGB image which can be later converted into a gray scale image for processing.

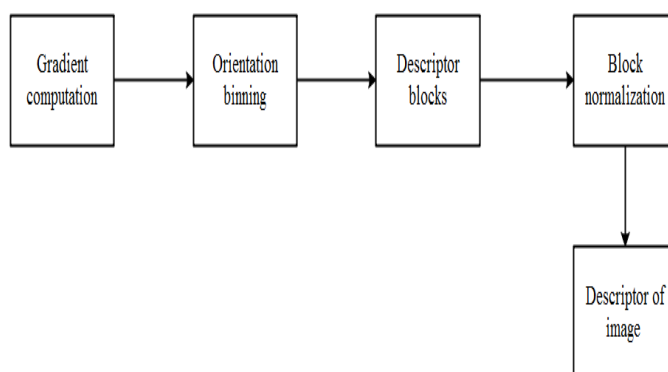


Fig. 1. Implementation of HOG Feature Extraction.

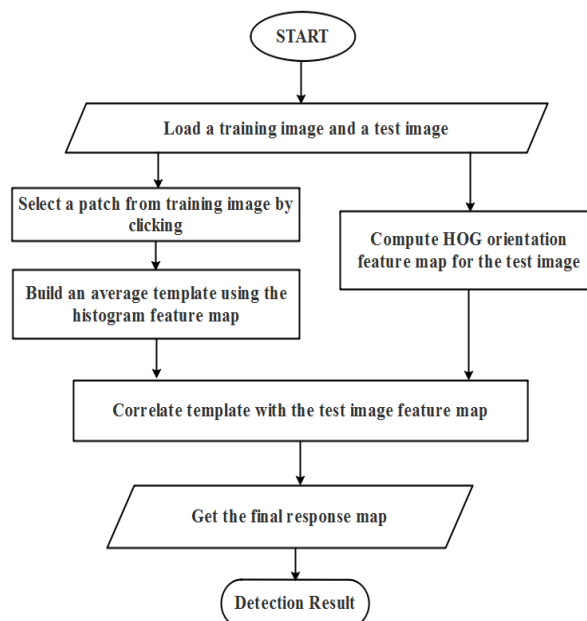


Fig. 2. Overview of HOG based Method.

B. Simple Template Matching Method

Template matching is a popular computer vision based image processing technique used to find out desired object from an input image. This method uses an image patch according to a specific feature of the search image, which we need to identify. The template matching uses a matching criterion to determine the position of an object and calculates a correlation coefficient. Template matching measures the similarity between the template image and the overlapped portion of the original image [15]. This similarity measurement is known as cross correlation. The cross correlation value will be greatest at locations where the input image matches the template patch or mask image [16].

Fig. 3 shows the flow diagram of Template Matching method.

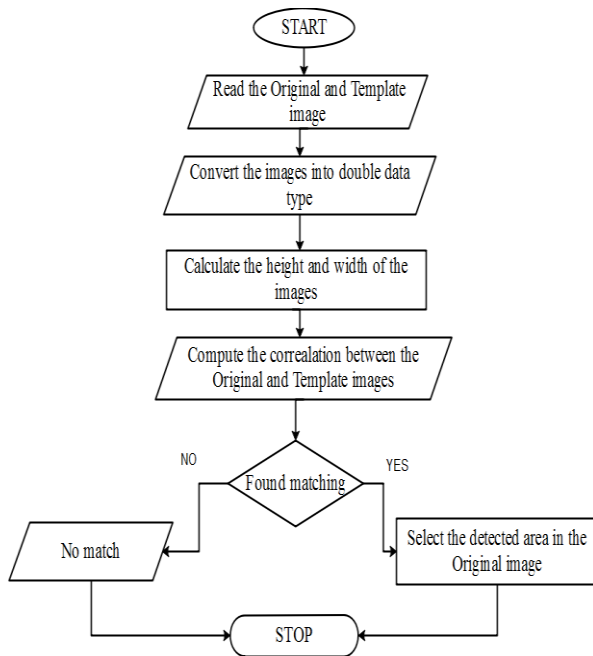


Fig. 3. Flow Diagram of Simple Template Matching.

IV. EXPERIMENTAL RESULTS

A. Experimental Tool

The experiment is performed in MATLAB environment with a number of images.

B. Object Detection using HOG based Method

The experiment is applied to a large number of colored and grayscale images. Here, the paper presents experimental results for some images only. The main motivation of this experiment is to identify the correct location of the template object from the desired test image. For object detection from an image, we have used a template image. For result illustration, we have only used total six template images shown in Fig. 4. A patch (a group of pixels) is taken from a template for detection purpose. The program is run by changing the size of the template patch (128×128, 64×64, 32×32, etc.) and the results for the variants of patch size are observed. We also use the full template image but we do not get the desirable results.

For the detection of objects from a test image, template images have been used. Fig. 4(a), (b), (c), (d), (e) and (f) represent the template images for car, key, elephant, medicine box, staple remover and book, respectively. Fig. 5(a), (b), (c), (d), (e) and (f) show the detection results for the template images of Fig. 4(a), (b), (c), (d), (e) and (f), respectively. It is seen that car, key, elephant and medicine box are detected correctly but staple remover and book objects are detected incorrectly.

The result of detection using HOG features:

True Positive= 4 (Correctly finding the location of the desired objects)

False Positive=2 (Incorrectly detection of template objects)

From Table I, it is observed that 128×128 size of the patch for all the template images is suitable for the correct detection of objects. The size of the template patch 128×128 is divided into 16×16 blocks where each block is of 8 pixels. From this experiment, it can be said that the suitable patch size depends on the original size of the template and the patch size must be smaller than template size to detect the objects accurately.

C. Object Detection using Template Matching Method

In the template matching method, this paper uses the cropped image of objects as a template image. Template matching cannot detect objects when we take the arbitrary size of the template image. Template of arbitrary sizes (e.g. 32×32, 64×64 and 128×128 etc.) are applied but it is observed that the investigated Template Matching system is unable to detect the image. In Fig. 6 we use the original template image and observe that the object is not detected. But in Fig. 7, all the objects are detected correctly.



Fig. 4. Sample of Template Images.



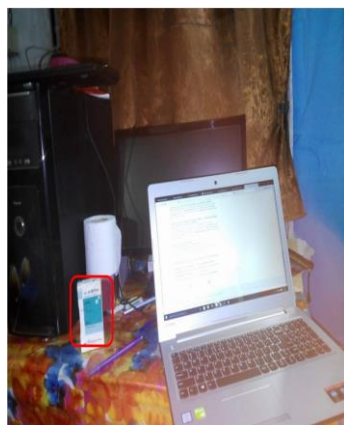
(a) car detected



(b) key detected



(c) elephant detected



(d) medicine box detected



(e) staple remover wrongly detected



(f) book wrongly detected

Fig. 5. Object Detection Results using HOG Method.

TABLE I. THE DETECTION RESULT OF HOG FOR DIFFERENT SIZES OF THE PATCH OF THE TEMPLATE IMAGE

| Patch Size | 256x256 | 128x128 | 64x64 | 32x32 | 16x16 | 8x8 |
|--------------------------|---------|---------|-------|-------|-------|-----|
| Car (225x300) | No | Yes | Yes | Yes | No | No |
| Key (231x149) | No | Yes | No | No | No | No |
| Elephant (389x375) | No | Yes | Yes | No | No | No |
| Medicine box (385x505) | No | Yes | Yes | No | No | No |
| Staple remover (385x504) | No | No | No | No | No | No |
| Book (640x480) | No | No | No | No | No | No |

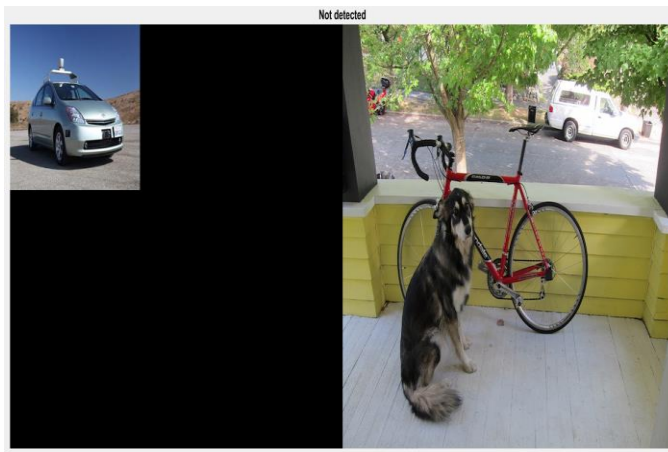


Fig. 6. Object Detection Results using Template Matching Method.

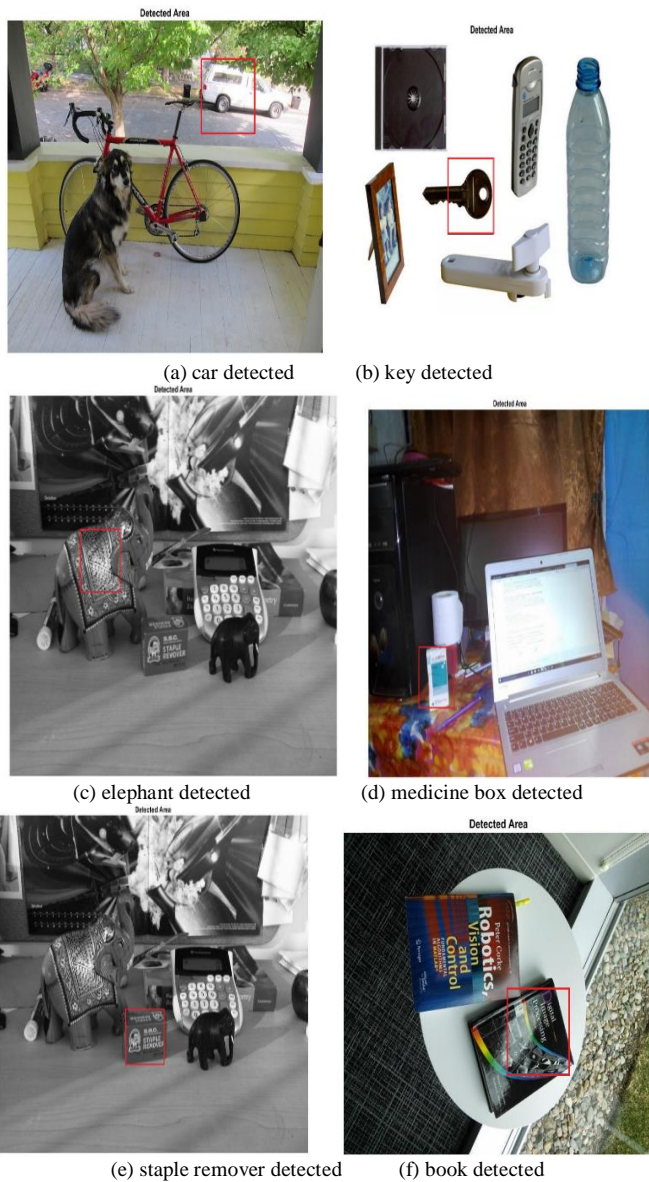


Fig. 7. Object Detection Results of Template Matching using Cropped Template Images.

TABLE II. ELAPSED TIME FOR DETECTION

| Template Image | Elapsed time for HOG (in second) | Elapsed time for Template Matching (in second) |
|----------------|----------------------------------|--|
| Car | 0.731388 | 8.449037 |
| Key | 0.680365 | 5.035367 |
| Elephant | 1.146820 | 8.031569 |
| Medicine box | 1.035807 | 8.09263 |
| Staple Remover | 0.974079 | 8.55565 |
| Book | 0.916373 | 4.727828 |

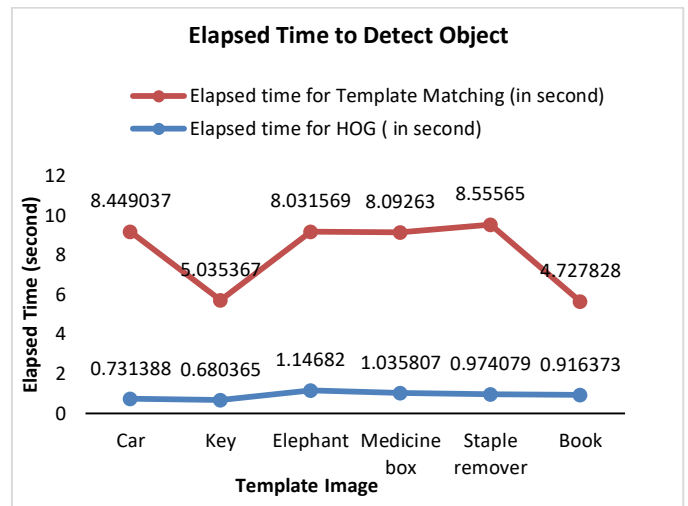


Fig. 8. Detection Time for HOG and Template Matching Methods.

Table II shows the elapsed time in second for the investigated techniques.

From Fig. 8, it is seen graphically that in the detection of the image, template matching elapses more time than the elapsing time of HOG method for each of the object class.

V. CONCLUSION

The paper addresses the issue of the detection of the desired objects from a test image using two methods named HOG feature-based method and template matching method. In this paper, the performance of these methods has been analyzed using several images. The research is tested on a sample of six images from the database of 20 images for different types of objects. Template Matching uses simply pixel-based cross-correlation matches which is easy to implement. It only works well when the template image is cropped from the original image otherwise it shows poor results in object detection. HOG method uses feature descriptors to detect images. In this method, the cross-correlation between the template and feature descriptor is used. It requires less time to run the detection process.

Now, it is concluded that the implementation of HOG based method is comparatively complex but it shows optimum results based on both detection accuracy and elapsed time.

REFERENCES

- [1] Khushboo Khurana, Reetu Awasthi, "Techniques for object recognition in images and multi-object detection," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol 2, Issue 4, April 2013.
- [2] Shleymovich M.P., Medvedev M.V. and Lyasheva S.A., "Object detection in the images in industrial process control systems based on salient points of Wavelet Transform," 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM) of Advanced Research in Computer Science and Software Engineering, 2016.
- [3] Pooja Ghosh, Achala Pandey and Umesh C. Patil, "Comparison of different feature detection techniques for image mosaicing," ACCENTS Transactions on Image Processing and Computer Vision, ISSN(online):2455-4707, vol. 1, Issue 1, November 2015.
- [4] A. Al-Mamun and M. Yousef Ibrahim, "Feature-based real-time object identification in industrial processes," International Conference on Integration Technology, March 20-24, 2007.

- [5] P M Panchal, S R Panchal and S K Shah, "A comparison for SIFT and SURF," *International Journal of Innovative Research in Computer and Communication Engineering (IJRCCE)*, vol. 1, Issue 2, April 2013.
- [6] Vivek Kumar Gupta and Kanchan Cecil, "An analytical study of SIFT and SURF in image registration," *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 3, Issue 9, March 2014.
- [7] Haoyu Ren and Ze-Nian Li, "Object detection using edge histogram of oriented gradient," *ICIP*, 2014.
- [8] Chetan S. Gode and Atish S. Khobragade, "IEEE WiSPNET Conference," 2016.
- [9] Pooja Ghosh, Achala Pandey and Umesh C. Patil, "Comparison of different feature detection techniques for image mosaicing," *ACCENTS Transactions on Image Processing and Computer Vision*, ISSN(online):2455-4707, vol. 1, Issue 1, November 2015.
- [10] Zuzana Haladova and Elena Sikudova, "Multiple instances object detection," *Association for Computing Machinery*, May 2014.
- [11] Anaswara S Mohan and Resmi R, "Video image processing for moving object detection and segmentation using background subtraction," *International Conference on Computational systems and Communications (ICCSC)*, December 2014.
- [12] Nazil Perveen, Darshan Kumar and Ishan Bhardwaj, "An overview on Template Matching methodologies and its applications," *International Journal of Research in Computer and Communication Technology*, Vol 2, Issue 10, October- 2013.
- [13] Chakraborty Partha, Mohammad Abu Yousuf, Md. Zahidur Rahman, and Nuruzzaman Faruqi, "How Can a Robot Calculate the Level of Visual Focus of Human's Attention," In *Proceedings of International Joint Conference on Computational Intelligence*, 329-42. doi:10.1007/978-981-15-3607-6_27.
- [14] Pranav Kumar, S L Happy and Aurobinda Routray, "A real-time robust facial expression recognition system using HOG features," 2016 *International Conference on Computing, Analytics and Security Trends (CAST)* College of Engineering Pune, India. Dec 19-21, 201.
- [15] Nilamani Bhoi and Mihir Narayan Mohanty, "Template Matching based eye detection in facial image," *International Journal of Computer Applications*, Vol.12 (5), pp. 15-18, 2010.
- [16] Kavita Ahuja, Preeti Tuli, "Object recognition by Template Matching using correlations and phase angle method," *International Journal of Advanced Research in Computer and Communication Engineering* ,Vol. 2(3), pp. 8-14, 2013.