Object detection and recognition by using enhanced Speeded Up Robust Feature

Tawfiq A. Al-asadi† and Ahmed J. Obaid 2††,

CS Dept. College of IT, Babylon University, IRAQ CS Dept., Faculty of Education KUFA University , IRAQ

Abstract

In image processing field there is an attention for detection objects, regions and points then made decision in case found it in a single or collection of images may called test or image data set, for this task we have used an algorithm that used in many computer vision application and also considered very fast by compared to others this algorithm can detect and describe local features for any interest object and extract features or descriptor points from it and compare these features/ descriptor by the features that extracted from origin image, matching process has been done among features and decision made based on similar features found, this algorithm called Speeded Up Robust Features (SURF) algorithm. In this paper we used enhanced Speeded Up Robust Features "SURF" algorithm, our model counting the features in either object and origin image in data set, then matching percentage calculated using a metric of counting the size of inlier matching features towards outlier features, Radom Sample Consensus (RANSAC) algorithm has been combined with SURF for eliminated error matching that happen in features, then decision has been given based on that metric if the object is present or not. In case object found Speeded UP Robust Features "SURF" algorithm can detect the position of the interest object in origin image by using geometric transform. In this paper we have used our metric and enhanced model to made decision and write result finally for each compared process and also write some information that used for matching procedure finally we can distinguish each calculating percentage and valid strength features matched that used for finding the interest objects under different circumstances.

Key words:

Object detection, object recognition, feature matching, SURF.

1. Introduction

The main task for object detection and recognition systems is to detect and recognize if any query object from origin image was known prior, many computer vision application recently consider the problems arises with this concept because there are problems of different images acquisition methods and the corruption of background image and noise affects in order to deal with this problems one of the robust algorithm are consider in many computer vision application called Speeded Up Robust Features "SURF" algorithm. Speeded Up Robust Features "SURF" algorithm is a local feature and descriptor algorithm that can be used in many application such as object recognition , SURF use much larger number of features descriptor from origin image which can reduce contribution of the errors caused by local variation in the average of all feature matching . SURF can robustly recognize and identify objects in origin images even in case of clutter and partial occlusion because SURF has feature descriptor which is invariant to scale, partial variant in illumination changes and orientation.

The process of Speeded UP Robust Features "SURF" algorithm can be divided into three main steps. First step is "Detection step", in this step interest points are selected at distinctive locations in the origin image, such as corners, blobs and T-junctions and this process must be robustly. The most valuable property of an interest points it's a repeatability. Repeatability express the reliability of the detector for finding the same physical interest points under different scene conditions. Second step is "Description step", in this step interest points should have unique identifiers does not depend on features scale and rotations which are called descriptor, the information of interest points represented by descriptor which are vectors that contain information about the points itself and the surroundings. Third step is "Matching step", in this step descriptor vectors are compared between the object image and the new input or origin image, the matching score is calculated based on the distance between vectors e.g. Euclidian distance and the sign of laplacian. Then if the object is found then give a message for that and view the percentage of matching score and store the result in predefined file, otherwise will give an underline message that object was not found. In this paper we are used enhanced Speeded up Robust Features "SURF" algorithm in order to detect and recognize our interest objects and proposed a metric for counting matching score to give better result.

SURF algorithm was first presented by [1] in 2006, Use an integer approximation of determinant of Hessian blob detector, which can be computed with three integer operations using an integral image. SURF features descriptor are calculated by the sum of Haar wavelet response around interest points. And these can be computed by the concept of integral image. This algorithm can be used in many application such as recognize and locate of objects, track objects, face recognition, make 3D

Manuscript received April 5, 2016
Manuscript revised April 20, 2016

scenes. In this paper we used enhanced SURF algorithm for detect and locate the objects if it's found in origin image.

2. Related Work

SURF algorithm used in various computer vision and real time application, SURF algorithm one of the fast and robust method for object detection and recognition which has been proposed by [1]. SURF features used for detection the traffic sign in Field Programmable Gate Array (FPGA) which is hardware to process video streams in real time has been used in [4]. There are several research published in order to improve the performance of SURF algorithm for varies application, an image matching algorithm combined with SURF and DAISY descriptor is proposed by [2] for increase the matching , running time, capability of SURF in rotation situation. The Robust of SURF feature rather than other algorithm for fast matching lead to use in some specific applications such as face recognition, object detection, object recognition and image retrieval using SURF and interest points' detection and description concepts as it used by [7, 8, and 9]. [10] Has been proposed a way to increase matching performance which can be obtain in relation of underline detector of interest points in both algorithms Scale Invariant Feature Transform "SIFT" and SURF. Many research work and computer vision application use SURF algorithm for image retrieval and content based image retrieval (CBIR) by indexing the features vectors and calculate the features for input or query image then find the matching images based on similarity measures after that retrieving matching images result such used in [11]. Our proposed methods that used enhanced SURF technique in real time for detect interest objects from tested images by finding the strongest features then RANSAC algorithm combined in our work for eliminated the error matching features in matching process, RANSAC is insensitive to outlier then our formula for counting the percentage of inlier and outlier features employed here to give a decision in report file and visualize mode.

3. Surf Algorithm

Speeded UP robust Features "SURF" algorithm is considered a robust local feature detector and extractor algorithm and can be used in many computer vision application like object recognition, 3D reconstruction and its one the best approaches suitable for real-time application [1]. The interest point detection which is represented by a vector call descriptor in SURF algorithm is based on scale space theory , SURF algorithm use an Integer approximations as the determinant of Hessian blob detector which can be computed fast with an integral image [2]. The integral image is an image where each point in this image X which equal to (x, y) T Stores the sum of all pixels in the input image (I) within a rectangular region which is formed by the origin and X, see Eq. 1:

$$I(X) = \sum_{i=0}^{i\le x} \sum_{j=0}^{j\le y} I(i,j) \quad (1).$$

The integral images are used in Hessian matrix approximation which reduce the time of computation effectively. Since Hessian matrix has good performance and also has good accuracy, in image I, X=(x, y) is a given point in an origin image, the Hessian matrix H(X, σ) in X at scale σ is defined in Eq. 2:

$$H(X,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{yx}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \quad (2)$$

Where $L_{xx}(x, \sigma)$ is the convolution result of the second order derivation of Gaussian Filter [(∂^2)/ (∂x^2)] g (σ) with the image in point X, and similarity for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$. To reduce the computation time, set of 9×9 box filter is used for approximations for Gaussian second order derivatives with σ =1.2 and this value represent the lowest scale (i.e. highest spatial resolution) for computing the blob response maps. We will denote them by D_{xx}, D_{yy} and D_{xy}. The weights applied to the rectangular regions are kept simple for computational efficiency, see Eq. 3:

$$\det(H_{approx}) = D_{xx} D_{yy} - (\omega D_{xy})^2 \quad (3)$$

Where ω is the relative weight of filter response and given by this formula for box filter 9×9 and σ=1.2 [3], for ω see Eq. 4:

$$\omega = \frac{|L_{xy}(1.2)|_F \, |D_{yy}(9)|_F}{|L_{yy}(1.2)|_F \, |D_{xy}(9)|_F} \quad (4)$$

The relative weight of filter ω is used in this formula to keep the balance for the Hessian determination process and the weight can be changes depending on the scale by default this weight did not have a significant impact on the result [4] because we are using a metric for counting the percentage of inliers points founded as well as outliers. SURF applies different sizes from box filters to search and compare interest points, so box filters has different size can be construct the scale space and which can be divided into octaves. Scale space representation is defined as the convolution of a given image with Gaussian kernel. Usually scale space are implemented as an image pyramid, Scale space is a continuous function which can be used to find the maximum values across all possible scales [5].

Scale space in SURF algorithm is analyzed by up-scaling the filter size instead of iteratively reducing the size of an image as shown in the figure 1.

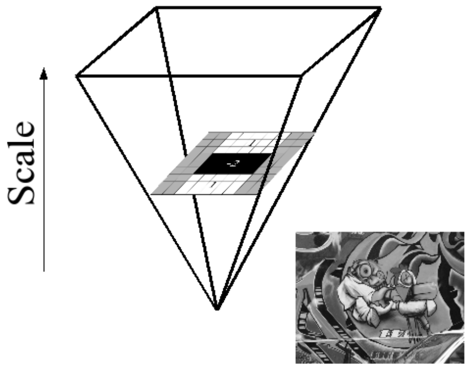

Figure 1 SURF using different filter size while Origin image unchanged

The Scale space then divided into a number of octaves, octave refers to a series of response maps result from convolving the same input image with a filter has been sized increased. So each octave is subdivided into constant number of scale level, to ensure the size of an image in odd and the central pixel is present box filter increased by fixed values, the box filter starts off with a 9×9 size filter as an initial scale layer and scale value is an s=1.2(the approximated Gaussian derivative with σ=1.2), Table 1 represent the box filter edge size for three Octaves as we are used in our proposed method.

Table 1. Box filter edge size of three octaves along with scale values

Octaves	**1**				**2**				**3**			
Box filter size l	9	15	21	27	15	27	39	51	27	51	75	99
Scale value s=1.2* (l/9)	1.2	2.0	2.8	3.6	2.0	3.6	5.2	6.8	3.6	6.8	10.0	13.2
Increased value	6				12				24			

After the approximation of hessian matrix determinant is obtained in each layer, non-maxima suppression is applied in 3×3×3 neighborhood to localize the interest points over the scales of image. Non Maximum Suppression (NMS) can be defined is a process which can find the candidate interest points within certain neighborhood around the pixel. That means each pixel in the scale-space it compared to its 26 neighbors in the above and below scales, at this stage we get set of interest points that has minimum strength determined by threshold value and also local maxima and minima in the scale-space. After the pixel is selected as a maxima if it's greater than of it surrounding pixels and above / below intervals, then the interpolated in scale image space based on fats Hessian matrix detector method is applied [6]. The interpolated location of interest point can be determined by finding the blob response of 3D neighborhood, blob response is based on the first order Haar wavelet response in both x and y direction, using Haar wavelet increase robustness and can minimize the computation time [3].

Once the localization of interest point has been completed, each of these interest points must has unique descriptor in order to find the correspondences between two images and also can evaluated both. The main goal from descriptor is to provide a unique and strong unique description of an image features which can describe the distribution of intensities of pixels within neighborhood of interest points [13].SURF descriptor is based on two steps, the first step is use the information arise in circular region around the interest point which can lead to reproducible orientation information, second step is a square region aligned to the selected orientation has been construct to extract SURF descriptor from it this window contains the pixels that will form the descriptor vectors that used for matching process. In case of fast indexing within matching stage sign of Laplacian is used [1]. Sign of Laplacian can distinguish bright blob on dark background from reverse one, to increase the matching process the same type of contrast features are compared "minimal information can increase matching process without reducing descriptor performance" [12].

4. Proposed Methodology

In Proposed method as shown in figure 2, at first step, the interest point for object and origin image has been detected this step also called SURF point's detection, these points contain information about features also called blob features. We select strongest features by attend a specific threshold and selection criteria then all features are return represent the strongest features for our interest object and origin image. Then in second step, feature descriptor also called features vector are extracted from pixels that surrounding an interest point. Pixels represent match features specified by single point location and this specify the center location of neighborhood pixels at the end of this step we get strongest feature point (interest point) descriptors also called object representative points because its carry information that can distinguish and recognize it [15,16]. Third step is a matching step that match features from first set of (object image) to second features set (origin image). Matching step return indices of the matching features for tow features set. Second Threshold

Is present here to evaluate the matching features, some of these features in object image might match features in origin image which are not belong to the object.

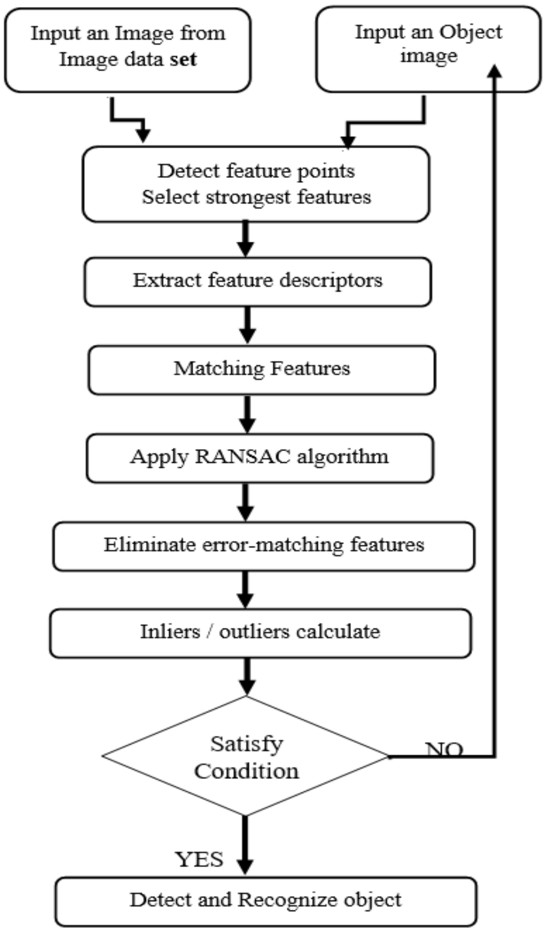

Figure 2 Block diagram of Proposed Methodology

At this step are use our comparison schema to find the total number of feature descriptors in the index pairs and then used this result in the next step. Step 4 we eliminate the error matched features using RANSAC algorithm [12].Result may include outliers matched features, to remove outliers in matched feature sets, RANSAC (Radom Sample Consensus) algorithm will be used here by: Establish an initial inliers combination and calculate fundamental matrix by fetch 8 matched pairs randomly, compute Sampson error and compare matched pairs based on our selected threshold if larger discard and check next pair otherwise update fundamental matrix , this process repeated till no more inliers are included. Here we are using metrics to calculate the matched pairs points and distinguish the inlier and outliers points based on this we have built our decision if the object is exist or not , and measure matching percent based on it [13]. Figure 3 show the proposed algorithm steps for our model,

Object Detect and locate Algorithm:

Step1 Read an integral image
Read an image I from archiving set
Read an image J (Object Image).

Step2 Detect feature points (interest points)
Detect feature point for image I1
Detect feature point for image J1
Image I1 (feature1, feature2 ….. feature $_m$)
ImageJ1 (feature1, feature2 ….. feature $_n$)
For I = 1 to M do
For J = 1 to N do
{Select strongest feature points} >= Th1 (threshold)

Step3 Extract feature descriptors
Image I1 {f1, f2, f3 ……. f $_m$}
Image J1 {f1, f2, f3 ……. f $_n$}

Step4 Matching features
Find matched features indices = IP (features pairs)
If IP >= Th2 (threshold) %check if there is sufficient number of features pairs.
Go to step 5
Else Object is not found,

Step5 Apply RANSAC algorithm to eliminate error matched features
1-Inliers combination and built a fundamental matrix
2-Compute Sampson error and compare matched feature with Threshold.
3-Repeat till no more inlier included.
Count inlier and outlier features C_{in} and C_{out}.

If C_{in} >= C_{out} object found,
Else object is not found,
Find the matching percentage MP.

The result of this step we are compute geometric transform between two images to align it.
END.

5. Experimental Results

Object detection and recognition simulation done by using Matlab 2015a, Main GUI interface for input the selected image from image data set and entered object image are shown in the figure 4.

Figure 4: Main GUI interface for Object and Tested image

Then find interest points for both images and from these points select strongest points based on threshold, figure 5 show strongest features for both origin and interest object images.

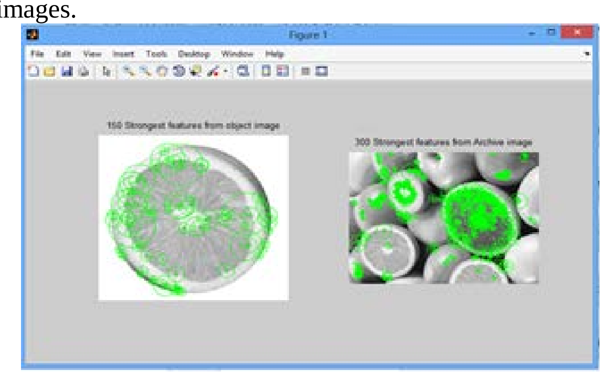

Figure 5: Strongest points for both Images

Feature descriptors for both images has been found, these descriptors are represented in vectors also called representative points which can recognize object in the scene , then matched features from origin image to the object image. Matching may contain some error-matched features (outliers) as shown in the figure 6.

In order of eliminate error-matched features, we apply RANSAC algorithm, the result can produce matched features, then we calculate the percentage of matching there , if the matching score does not have enough matching points then our system will show an message that object is not found, the implementation of this step can be shown in the figure 7.

In our scenario different images are used from image data set and tested against our interest objects, object can be found in these images in different scene, formula for counting the percentage of inliers / outliers to the total number of features has been calculated then based on this formula result is saved in file that contain tested images along with interest objects were selected.

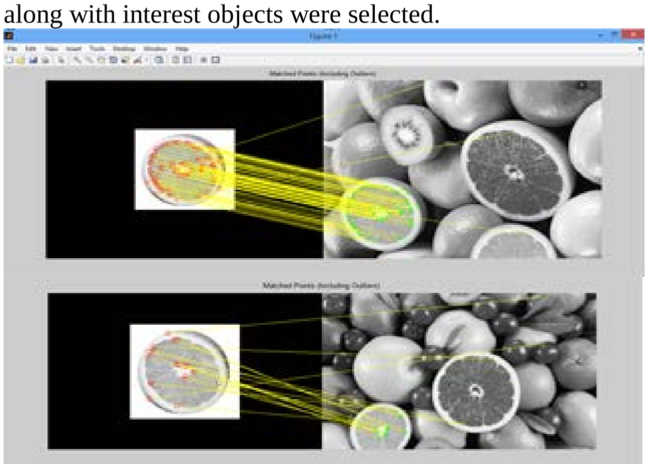

Figure 6: Matched process and applying RANSAC step

Figure 7: Object detection in the Origin image

5. Conclusion

Proposed method for detect and recognize object in the scene is based on SURF algorithm, we enhanced the performance of object detection by selecting the strongest features descriptor, our proposed method it's successfully detect one or more objects in data set of images and calculate matching score for object in the scene by applying three types of thresholds and accuracy measures are objects recognition under variable conditions of rotation, partial occlusion, orientation and illumination changes by enhanced illumination of image inputs. Many real time application that use SURF algorithm can detect objects by visualize mode, our model calculate many information that used through object detection scenario therefore our proposed model its user-friendly where we can select and changed many parameters such as selected threshold and octaves that used for detection and recognition process. Matching score can represent interest value for how much accuracy our selected parameters were affected and also support the accuracy for detection interest objects in variant origin images from data set.

References

[1] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346-359, June 2008.

[2] Li Li, "Image matching algorithm based on feature point and DAISY descriptor", Journal of multimedia, vol.9, No.6, pp.830-831, June 2014.

[3] Bassem Sheta, Mohamed Elhabiby, Nase El-Sheimy, " Assessments of different Speeded UP Robust Features SURF algorithm resolution for Pos estimation of UAV ", " IJCSES, vol.3, No.5, pp.21-22, October 2012.

[4] Jin Zhao , Sichao Zhu , Xinming Huang , " Real-Time traffic sign detection using SURF features on FPGA " , HPEC , 10.1109 , IEEE pp.2-3, Sep 2013.

[5] Christopher Evans, "Notes on the Open SURF Library", pp.4-5, 2009.

[6] Matthew Brown, David Lowe, "Invariant features from interest points group ", BMVC, 2002.

[7] Geng Du , Sichao ZhFei SU , Anni Cai , " Face recognition using SURF features " , SPIE , vol.7496 , Dec. 2009.
[8] Bhosale swapnali, Kayastha Vijay, Harpale Varsha , " Feature extraction using SURF algorithm for object recognition " , " IJTRA , vol.2, Issue 4 ,pp.197-199 July-August 2014.
[9] Miriam Lopez ,Takayuki Nagai, Mohammed Attamimi, Mariko Nakano, Hector Perez , " Object detection using SURF and super pixels " , " JSEA , SciRes , 6, 511-518 , Sep 2013.
[10] Tony Lindeberg, "Image matching using generalized scale-Space interest points", Springer, SSVM, vol.7893, pp355-367, 2013.
[11] K.Velmurugan, Santhosh Baboo, "Content-Based image retrieval using SURF and colour moment", GJCSA, vol.11, Issue 10, May, 2011.
[12] Peng-le, C., "Study over high-precision vision inspection based on RANSAC algorithm", Journal of Convergence Information Technology, vol.7, no.20, pp.33-38, 2012.
[13] Peng-le, C., "Study over high-precision vision inspection based on RANSAC algorithm", Journal of Convergence Information Technology, vol.7, no.20, pp.33-38, 2012.
[14] Kong Jun, Jitang Min, Alimujiang Yiming, "Research on Speeded UP Robust Features with RANSAC", IJACT, vol.5, no.8 April 2013.