Pattern Recognition By a Scaled Corners Detection

Asst. Prof. Dr. Dhyaa Shaheed Al-Azzwy
Department of Computer
University of Wasit – College of Computer & Mathematics
Wasit , Iraq
* dhyaaman@yahoo.com

Abstract

In this paper we developed a new approach to extract points descriptor used for pattern recognition with Corner detection approach. We used scales of image, each scale was scaled by a scaling factor, detect the corners in each scale, extract the key points descriptor from these corners, and using these points descriptor as key features of recognition in the Hough Transform to classify the Descriptor to its class. We implemented and analyzed SIFT algorithm, corner detection algorithm, and the proposed approach. The experimental results using MATLAB of a proposed approach gave results of recognition with high accuracy.

Keywords: Pattern Recognition; Corner Detection; SIFT; Hough Transform.

INTRODUCTION

Local descriptors computed for interest regions have proven to be very successful in applications such as wide baseline matching [1], object recognition [2], [3], texture recognition [4], image retrieval [5], [6], and recognition of object categories [7], [8], [9], [10]. They are distinctive, robust to occlusion, and do not require segmentation. One of the oldest and well-known methods for detecting interest-points was proposed by Harris [11] and is based on the auto-correlation matrix. While being translation and rotation invariant, Harris points are not invariant to change in scale. A scale invariant version of the Harris detector was proposed by Lindeberg [12], which is also referred as Harris-Laplace detector. Mikolajczyk and Schmid [13] further improved this method to provide an affine invariant detector called Harris-Affine. Lowe [3] used Difference of Gaussian (DoG) as an approximation to Laplacian of Gaussian (LoG) and used their local maxima to detect interest points.

Our new approach concerned to calculate the descriptors of detected corners for scaled image. Each image scaled by factor and filtered by Gaussian filter. Calculate the descriptors of corners for each scaled and filtered image.

SCALE INVARIANT FEATURES TRANSFORM SIFT

Lowe [3], [14] presents a method for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The features are invariant to image scale and rotation, and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. The features are highly distinctive, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images. The recognition proceeds by matching individual features to a database of features from known objects using a fast nearest-neighbor algorithm, followed by a Hough transform to identify clusters belong to a single object.

Extracting features is performed by a cascade filtering approach using a four stages algorithm; following are the major stages of computation used to generate the set of image features:

Scale-space extraction detection:

First a scale space is defined by "Eq.(1)".

$$L(x ,y, \sigma) = G(x, y, \sigma)*I(x, y) \qquad (1)$$

$$G(x, y, \sigma) = 1/(2\pi\sigma^2)e^{-(x2+y2)/2\sigma 2} \qquad (2)$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \qquad (3)$$

Where * is the convolution operator, $G(x, y, \sigma)$ is a variable-scale Gaussian, $I(x, y)$ is the input image, *Fig 1*, shows original image, $D(x, y, \sigma)$ is the Difference-of-Gaussian DoG, as shown in *Fig 2*, For finding stable features Difference-of-Gaussian function convolved with image, which can be computed with difference of two nearby scales, after each octave, the Gaussian image is down-sampled by a factor of 2, and the process repeated.

To detect the local maxima and minima of *D(x, y, σ)* each point is compared to its eight neighbors at the same scale, plus the nine corresponding neighbors at neighboring scales. If the pixel is a local maxima or minimum, it is selected as a candidate keypoint, see *Fig 3*.

Figure 1. The 640x480 pixel original image.

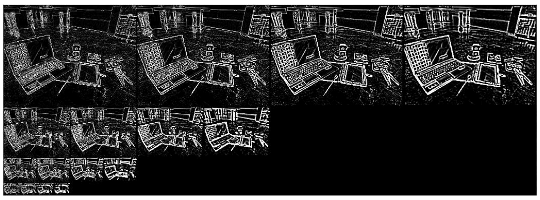

Figure 2. Pyramid of *Difference of Gaussian DoG* for scaled and filtered image.

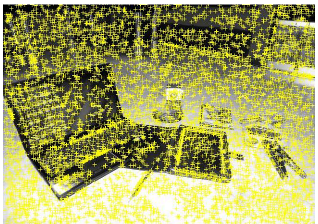

Figure 3. The initial 1002 keypoints locations at maxima and minima of the difference-of-Gaussian function.

Keypoint localization and filtering:

This step attempts to eliminate keypoints which are located on edges or the contrast between point its neighbors and is low, as shown in *Fig 4*.

(a) (b)

Figure 4. (a) After applying a threshold on minimum contrast. (b) After applying a threshold on ratio of principal curvatures.

Orientation assignment:

In this stage, one or more orientation is assigned to each keypoint in order to make them invariant to rotation. Suppose for a keypoint, L is the image with the closest scale, gradient magnitude *m(x, y)* and orientation *θ(x,y)* can be computed using "Eq.(4)".

$$Gradientvector = \begin{bmatrix} L(x+1,y) - L(x-1,y) \\ L(x,y+1) - L(x,y-1) \end{bmatrix}$$

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \qquad (4)$$

$$\theta(x,y) = tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right)$$

In next stage a gradient histogram (36 bins) is created. And peak within 80% of the highest peak is used to create a keypoint with that orientation, see *Fig 5*.

Figure 5. After assigning orientation to the keypoints, keypoints are displayed as vectors indicating scale, orientation, and location.

Keypoint descriptor:

The final step is to compute a descriptor to make it invariant to remaining variations. For this purpose a 16*16 Gradient window is taken that partitioned into 4*4 sub windows. Then, histogram of 4*4 samples in eight bins is created. This result in a feature vector containing 128 elements.

When at least three keys agree on the model parameters with low residual, there is strong evidence for the presence of the object. Since there may be dozens of SIFT keys in the image of a typical object, it is possible to have substantial levels of occlusion in the image and yet retain high levels of reliability.

CORNER DETECTION

Corner detection is used as the first step of many vision tasks such as tracking, localization, simultaneous localization and mapping (SLAM), and image matching and recognition. This need has driven the development of a large number of corner detectors. However, despite the massive increase in computing power since the inception of corner detectors, it is still true that, when processing live video streams at full frame rate, existing feature detectors leave little, if any time, for further processing [10]. Li [15], introduced eight corner detectors techniques as follow:

Moravec Detector

Moravec detector is one of the earliest corner detectors and it defines a corner to be a point where there is a large intensity variation in every direction. The algorithm of it tests each pixel in the image to see if a corner is present, by considering how similar a patch centered on the pixel is to nearby, largely overlapping patches.

Kitchen and Rosenfeld Detector

Kitchen and Rosenfeld detector uses a local quadratic fit to find corners. It includes the methods based on change of direction along the edge, angle between most similar neighbors, turning of the fitted surface and gradient magnitude of gradient direction.

Harris Detector

Harris detector is built on similar ideas to the Moravec detector which needs a method to match corresponding points in consecutive image frames, but is interested in tracking both corners and edges between frames, see *Fig 6*.

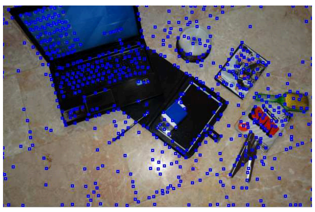

Figure 6. Image after applying Harris corner detection algorithm

SUSAN Detector

The idea of SUSAN corner detector is as follows: Given a circular mask delimiting a circular region of the image, the Univalue Segment Assimilating Nucleus (USAN) area is the area of the circular mask made up of pixels similar in intensity to the intensity of the central pixel (nucleus) of the mask. In a digital image, the USAN area reaches a minimum when the nucleus lies at a corner point, see *Fig 7*.

Trajkovic and Hedley Detector

Trajkovic and Hedley detector was developed with the intent of obtaining comparable repeatability rates and localization performance as the most popular corner detectors, while requiring a minimum of computation.

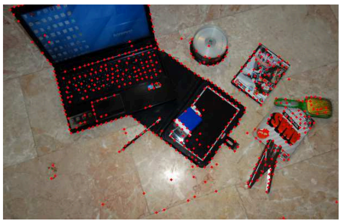

Figure 7. Image after applying SUSN corner detection algorithm

CSS Detector

CSS detector detects corners by directly looking for local maxima of absolute curvature with pre-threshold [7]. That is, it detects corners using the intuitive notion of locating when the contour of an object makes a sharp turn. CSS detector is the only one which uses edge contours extracted firstly and then computes curvatures on the edge contours at a high scale.

IPAN Detector

The algorithm of IPAN detector defines a corner in a simple and intuitively appealing way, as a location where a triangle of specified size and opening angle can be inscribed in a curve. A curve is represented by a sequence of points in the image plane. The ordered points are densely sampled along the curve, but, contrary to the other four algorithms, no regular spacing between them is assumed. A chain-coded curve can also be handled if converted to a sequence of grid points.

FAST Detector

FAST detector is a new algorithm for fast feature detection using machine learning. The detector allows it to be optimized for repeatability, with little loss of efficiency, and carry out a rigorous comparison of corner detectors based on the repeatability criterion applied to 3D scenes.

THE PROPOSED APPROACH

Our proposed approach aims to extract keypoint's descriptor which is used to recognize the pattern, it implies, extracting corners points for all the scaled patterns of the original image, calculating the orientations for the extracted corners points, extracting the descriptor for those oriented points, and finally, use the extracted descriptor in the Hough transform for the recognition process. The following stages describe the process of extracting the descriptor from the original image and the recognition in details:

Scale-Space stage:

This is the first stage of extracting the descriptor, the input image scaled by Gaussian filter for four octaves, in each octaves, the filtered image was scaled in its dimension with particular factor for five intervals, thus, the number of produced images are 20 scaled images, see Fig 8.

Figure 8. The pyramid of scaling of original image, scaled into four octaves of Gaussian filter, and each octave is scaled into five interval of dimension.

Corners Detection stage:

Extract all corners points for all the twenty scaled images. Practically, we used Hessian and SUSAN algorithms only.

Orientation assignment stage:

In this stage, one or more orientation is assigned to each keypoint in order to make them invariant to rotation. In addition to orientation, the location and scale have assigned to each keypoint. These parametrs impose a repeatable local 2D coordinate system in which to describe the local image region, and therefore provide invariant to those parameters.

Keypoint descriptor stage:

*This step is to compute a descriptor for the local image region that is highly distinctive yet is as invariant as possible to remaining variations, such as change in illumination or 3D viewpoint. For this purpose a 16*16 Gradient window is taken that partitioned into 4*4 sub windows. Then, histogram of 4*4 samples in eight bins is created. This result in a feature vector containing 128 elements.*

Recognition stage:

An efficient way to cluster reliable model hypotheses is to use Hough transform to search for keys that agree upon a particular model pose. Each model key in the database contains a record of the key's parameters relative to the model coordinates system. Therefore, we can create an entry in a hash table predicting the model location, orientation, and scale from the match hypothesis. We use a bin size of 30 degrees for orientation, a factor of 2 for scale, and 0.25 times the maximum model dimension for location. These rather broad bin sizes allow for clustering even in the presence of substantial geometric distortion, such as due to a change in 3D viewpoint.

THE EXPERIMENTAL RESULTS

The evaluation is done on real images taken by a digital camera. A significant amount of noise is added during the acquisition process (zoom, viewpoint rotation, light change, and object location change). The zoom changes involve a change in pixel intensity as automatic camera settings are used. Each scale change sequence consists of scaled and rotated images. For viewpoint change sequences the viewpoint varies in the horizontal direction between 0 and 180 degrees, and for vertical direction between 0 and 60 degrees. The database contains 8 of images for the pattern (Rasp object) in different noise. In the testing we used 12 images in each sequence representing different scenes.

A typical image contains 1500 or more features which may come from many different objects as well as background clutter, and experimentally, we have found that reliable recognition is possible with a few as 3 features.

The Hough transform identifies clusters of features with a consistent interpretation by using each feature to vote for all object poses that are consistent with the feature. When clusters of features are found to vote for the same pose of an object, the probability of the interpretation being correct is much higher than for any single feature. Each of our keypoints specifies 4 parameters: 2D location, scale, orientation, and each matched keypoint in the database has a record of the keypoint's parameters relative to the training image in which it was found. Therefore, we create a Hough transform entry for predicting the model location, orientation, and scale from the matched hypothesis. *Fig 9*, illustrates the implementation of both of the proposed approach and SIFTS approach, which gave match of recognition when the noises are not found.

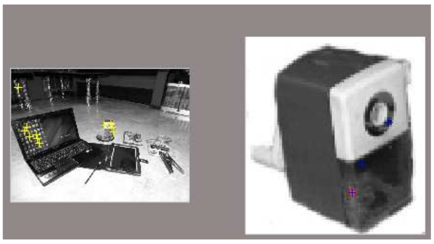

Figure 9. The implementation of both of proposed approach and SIFT approach in the recognition process, the left side shows image that contains the pattern occluded with other objects, right side shows the images of the required object.

Fig 10, illustrates the implementation of both of the proposed approach and SIFT approach with existence of noise (Illumination change, 3D angle change, and location changes), *Fig 10 (a),* shows the recognition success using proposed approach and the number of matched points are four points, *Fig 10 (b),* shows the recognition failed using SIFT approach.

(a) (b)

Figure 10. (a) Matching the object "in the right side" with the object inside the image of "the left side" when using proposed approach, (b) Mismatching the object "right side is empty" and there is no recognition when using SIFT approach.

Table 1: implementation of proposed approach and SIFT approach.

Approach	No. of images in DB	No. of tested images	Hit ratio %			
			Normal (pattern from the environment of testing image	Illumination change	3D rotation	Partial occlusion
Proposed with Harris algorithm	12	8	100%	83.3%	66.6%	83.3%
Proposed with SUSAN algorithm			100%	80.1%	67.6%	83.3%
SIFT			100%	16.6%	58.3%	91.6%

Table 1, shows the result of implementing both of proposed approach and SIFT approach for noised and normal environment, and it illustrates that the proposed approach gets good results of recognition in noised environment (illumination change, and 3D rotation), but the performance decreased in the partial occlusion environment.

CONCLUSION

In this paper, we introduced a new approach to extract local descriptor, which is inspired from earlier one such as SIFT but can be computed much more efficiency for dense matching purposes. Speed increase comes from replacing DoG keypoints by corner points calculated in corner detection algorithms (Harris's or SUSAN's algorithm), which can be computed very quickly. The experiments suggest that although the differencing of environment (scaling, 3D rotation, Illumination change), the performance is very well, but it degraded in partial occlusion. The degradation was because the detection algorithm depends mainly on the appearance of object in the image.

REFERENCES

T. Tuytelaar and L. Van Gool, "Matching Widely Separated Views Based on Affine Invariant Regions". Computer Vision, vol. 1, no.59, pp.61-85, 2004.

V. Ferrari, T. Tuytelaars, and L. Van Gool, "Simultaneous Object Recognition and Segmentation by Image Exploration", Proc. Eighth European Conf. Computer Vision, pp. 40-54, 2004.

D. Lowe, "Distinctive Image Features from Scale-Invariant Key Points". Computer Vission, Vol. 2, no. 60, pp. 91-110, 2004.

S. Lazebnik, C. Schmid, and J. Ponce, "Sparse Texture Representation Using Affine-Invariant Neighborhoods", Proc. Conf. Computer Vision and Pattern Recognition, pp. 319-324, 2003.

K. Mikolajczyk and C. Schmid, "Indexing Based on Scale Invariant Interest Points", Proc. Eighth Int'l Conf. Computer Vision, pp. 525-513, 2001.

C. Schmid and R. Mohr, "Local Grayvalue Invariants for Image Retrieval", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 5, pp. 530-534, may 1997.

G. Dorko and C. Schmid, "Selection of Scale-Invariant Parts for Object Class Recognition, "Proc. Ninth Int'l Conf. Computer Vision, pp. 634-640, 2003.

R. Fergus, P. Perona, and A. Zisserman, "Object Class Recognition by Unsupervisorvised Scale-Invariant Learning" Proc. Conf. Computer Vision and Pattern Recognition, pp. 264-271, 2003.

B. Leibe and B. Schiele, "Interleaved Object Categorizsation and Segmentation", Proc. 14th British Machine Vision Conf., pp. 759-768, 2003.

Edwaed Rosten, Reid Porter, and Tom Drummond, "Faster and Better: A Machine Learning Approach to Corner Detection", IEEE, Pattern Analysis and Machine Intelligence, vol. 32, no. 1, pp. 105-119, January 2010.

C. Harris and M. Stephens, "A Compined Corner Edge Detector", in Proc. Of the 4th Alvey Vision Conference, pp. 147-151, 1988.

T. Lindeberg, "Feature Detection with Automatic Scale Selection", Int'J. Comp. Vis., vol. 30, pp. 79-116, 1998.

K. Mikolajczyk, and C. Schmid, "Scale & Affine Invariant Interest Point Detector', Computer Vision, vol. 60, no. 1, pp. 63-86, 2004.

D. Lowe, "Object Recognition from Local Scale-Invariant Features", Proc. 7th IEEE International Conference on Computer Vision, vol. 2, pp. 1150-1157, 1999.

L. Li, and L. Zhang, "Corner Detection of Hand Gesture", Telkomnika, vol. 10, no. 8, pp. 2088-2094, 2012.