Points Descriptor in Pattern Recognition: A New Approach

Dhyaa Shaheed Al-Azzawy

Computer Sciences and Information Technology College / University of Wasitu
dalazzawi@uowasit.edu.iq

Abstract: We presented in the paper a new tactic, the first thing we have done is extracting the points of descriptor, which it is used in pattern recognition, especially in detection of corner algorithm. Scales of samples (images), each image is tuned by a factor (scale), collect the corners, and collect the points of descriptor key in these collected corners, in other words; Hough Transform uses the collected descriptors for classification process, and classify each points of image to its equivalence class. Experimentally, by using MATLAB, we are shown high accuracy of recognition result on the selected samples of objects.

Keywords: SIFT; Pattern Recognition (PR) ; Hough Transform (HT) **;** Corner Detection.

1. Introduction

Several methods as, baseline matching [1], recognition of objects and texture [2][3][4], retrieval of images [5][6], and objects categories recognition [7][8][9][10] use computed local-descriptors of concerned regions. All have assured to be very effective, peculiar, powerful to occlusion, and no segmentation. Harris proposed the outdated style for discovering concerned points [11], which depend on the auto-correlation matrix. Harris-points are effective to scale changing, when undergoing translation and stable rotation. The Harris detector, scale invariant version, was offered by Lind [12]. It is also indicated as the detector of Harris and Laplace. Mikolajcz & Schmd [13] continued improving this manner to provide detector of affine-invariant named Harris-Affine. D. Lowe [3] applied local maxima of Laplacian of Gaussian (LoG) and Difference-of-Gaussian (DoG) to detect interest points, and used the first (DoG)as an approximation to the second (LoG). The detected corners descriptors for scaled and filtered image are calculated by our new approach, which it use a factor to scale each image and Gaussian filter to filter them.

2. Scaled Invariant Features Transforming (SIFT)

A procedure of excavating distinguishing invariant countenance inside images is presented by Lowe [3] [14]. It can be applied to implement efficient matching with several views of an object. The rotation and scale of image are invariant. Powerful matching is provided by showing those features across an intrinsic scope of affine deformity, by adding of noise, illumination changes, and 3D changes. A elementary feature

may be properly identified with high likelihood versus a huge features data base created from numerous images. Subsequently, the features are highly distinctive.[15]

Clusters are identified as relational to a monocular pattern by use algorithm of quick nearest neighbor, pursued by a Hough-Transform, where the recognition take places by identifying monocular features in a database for familiar patterns.

A cascade filtering approach performs features extraction by using four algorithm stages. The central points of calculation that are applied to produce set of features are discussed in the following:

- *Detection of scaled-spaced:*

First scale is calculated by "Eq.(1)".

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \qquad (1)$$
$$G(x, y, \sigma) = 1/(2\pi\sigma^2)e^{-(x^2+y^2)/2\sigma^2} \qquad (2)$$
$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (3)$$

Which * represents convolution factor, $G(x, y, \sigma)$ represents Gaussian-Scale, $I(x, y)$ represents entered picture, and $D(x, y, \sigma)$ is the DoG to find stabilized characteristic. The DoG convolved with the image can be calculated from the variation of 2 closed scales. Next every octet, the Gaussian is increased-observed by 2, and this process repeated. The source image is displayed in figure1, while figure 2 displays the Difference-of-Gaussian DoG.
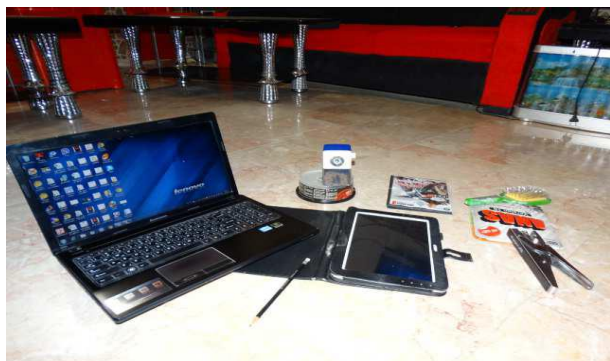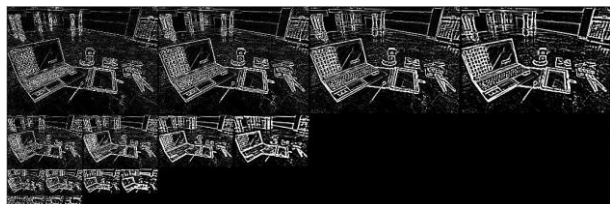
Figure (1): 640x480 image source

Figure (2): Samples of pyramid of *DoG* for filtered and scaled picture.

The regional minima and maxima of *D(x, y, σ)* for every point are detected by comparison to its 8 neighbours of the monocular scale, as well as 8 neighbours of neighbored scales. Insomuch as pixel is either minimum or local maximum, it is picked as elected key-point, as in figure*3*.
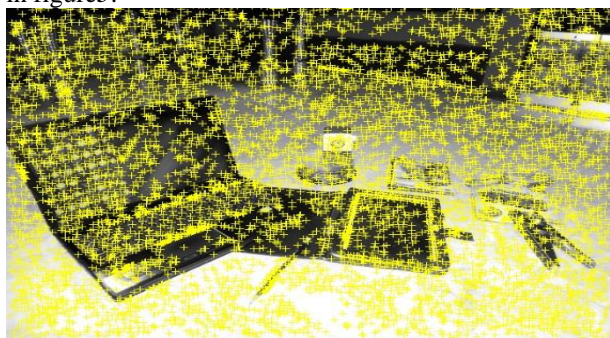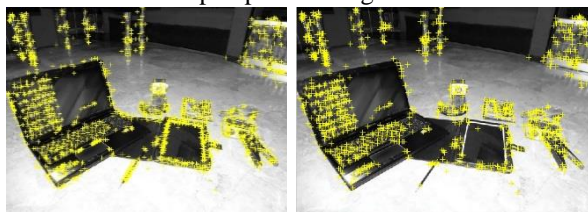

Figure (3): keypoint-locations (1002 features) at minima and maxima of DoG.

- **Localization - filtering of Key points:**
To exclude located keypoints at borders or where the contrast amidst point and its neighbours is soft, our algorithm uses the step depicted in figure 4

(a)   (b)

Figure (4): a. Next implementing a threshold at floor contrast, b. Next implementing a threshold at proportion of principal-bends.

- *Orientation-Alienation:*

Each keypoint has one or further orientation. In this stage each one will be assigned to make it invariant for rotation. *θ(x,y)* may be calculated applying "Eq.(4)".

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

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

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

…(4)

Where L is the picture of the relative-scale for a keypoint, *m(x, y) is* gradient magnitude, and *θ(x,y) is* orientation. A gradient histogram (36 bins) will be created in the next stage where salient within 80% of the maximum salient are applied to produce a keypoint of this orientation- as shown in figure 5.

Figure (5): Next orientation assignment to keypoints, are shown as a vectors scale pointed, location and orientation

- **Keypoint-Descriptor:**
The descriptor is computing step is the last step to make a descriptor invariant to the variations remaining. For this goal, a16x16 slope window is possessed and split into 4x4 sub-regions. After that, the system will create a 4*4 samples histogram in eight bins. The feature vector result contains 128 elements.

Three keys agree the minimum parameters of the pattern, with minimum residual, so there is robust evidence of object's presence. There is dozens of features inside the picture of the object, so it is may have occlusion at material grade in this image to hold over big reliability levels

3. Corner-Detection

Numerous vision missions as localization, tracking, simultaneous localization, image recognition and matching, use corner detection as a first step of their work. This need to develop a big issues of detectors.

Despite the huge growth in calculating capability since corner detectors initiation, when live video-streams are processing at rate of full frame, feature detectors struggle [10]. Eight corner detectors techniques were introduced by Li [16], as shown in the following:

❖ *Moravec Detector*

It appeared one of the going early corner prospector. A corner will be defined by this detector at a point when there is a huge variation of intensity at each direction. The procedure of that detector checks every pixel in this picture to check if corner is extant, by regarding how identical centered pixel is to nigh and broadly overlapping.

❖ *Kitchen-Rosefeld Detector*

A local quadratic fit is used to find corners by Kitchen and Rosenfeld detector. The manner based on the direction change over the edge and the angle amidst generality same neighbors. It also depends on the turning the fitted surface-gradient direction gradient magnitude.

❖ *Harris-Detector*

The Harris detector is constructed using the same ideas as Moravec-detector. That is, the need for a method to identify identical points inside consecutive frames of images, with interest tracking jointly edges and corners amidst frames, see figure 6.

Figure (6): The image next implementing algorithm of Harris-corner detection

The Auto-Correlation function for the signal is the main function, which the Harris corner detector was based on. The main idea of that detector is to find if there is a significant change in all direction of a point are exist or not, and if so, then this point is regarded as a point corner. This function measures the regional changes of windowed signal that was shifted to small values in different directions. [17]

The mathematical calculation of detector is achieved by computing C matrix as Eq(5)

$$C(x,y) = \sum_W [I(x_i, y_i) - I(x_i + \Delta x, y_i + \Delta y)]^2 \dots (5)$$

Where (x,y) are points inside the center of windows W, and $I(x,y)$ represents the image function, the two eigenvalues of matrix $C(x,y)$ is the main two factors that used to decide the case of point and as follow:

1- Insomuch as both of eigenvalues are small, thus the regional correlation function is a line.

2- Insomuch as the first eigenvalue is big and the second small, thus the regional auto-correlation function is an edge.

3- Insomuch as both of eigenvalues are big, thus the regional function is a corner.

❖ *SUSAN Detector*

This detector is implemented by using an orbicular mask to delimit an orbicular-region of picture. Region of the orbicular mask consists of pixels of identical intensity with the central pixel (nucleus) of mask, (called the Univalue Segment Assimilating Nucleus (USAN) area), which reaches to a minimum in a digital image where as the nucleus fall at a corner, as in figure 7
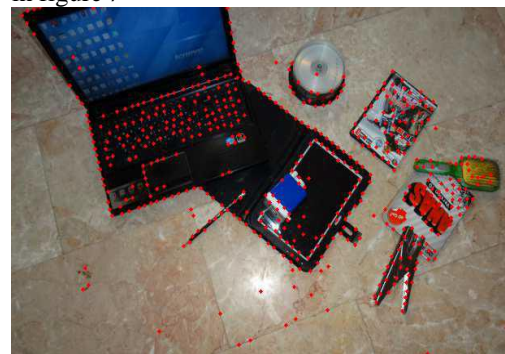

Figure (7): The image after implementing SUSN algorithm

To detect the corner using SUSAN algorithm, Eq(6) represents the mathematical model for this purpose.

$$c(r, r_0) = \begin{cases} 1, & |I(r) - I(r_0)| \le t; \\ 0, & \text{otherwise.} \end{cases} \dots (6)$$

Where r_0 represents nucleus's coordinates point, r represents other points coordinates inside the mask, and $c(r,r_0)$ represents comparison result. If $c(r,r_0)$ value is 1, thus the centered point is a corner; otherwise is not. [18]

❖ *Trajkovic and Hedley Detector*

To obtain comparable repeatability rates and localization performance, Trajkovic and Hedley detector was developed using the most public corner algorithms, which requiring a minimum of computing power.

❖ *CSS Detector*

Corners are detected by seeking local maxima of unlimited

curvature beyond a prethreshold. This technique is performed by the CSS detector [7]. In other words, it identify corners plying the conjectural conception of fixing where the contour produces a heavy turn. The CSS is the lone where first uses extracted brim-contours and calculates curvatures at those brim of big scale value.

❖ IPAN Detector

It marks a corner in a straightforward appealing manner. At, a location, a specified size triangle is inscribed in curve, which is appeared by a series of points in the image. This is performed by using the IPAN algorithm. The arranged points were violently sampled straightforward the curve, in four other algorithms. Non uniform interval amidst them is considered. A coded chain arch may be transacted if transformed to a series of grid.

❖ FAST Detector

A new algorithm is used by the FAST detector to detect fast features by using machine learning. This algorithm is allowed by the FAST detector to be perfect for reiteration, with less loss of qualification, and to implement rigid rapprochement of corner algorithm based within the reiteration norm utilized into three dimensions.

4. THE NEW (PROPOSED) MANNER

We developed our new approach, which extracts descriptors (keypoints) and then uses them, to identify a pattern. That means, electing points (corners) of all scaled-objects of source picture and then computing the tendency (orienting) of the taken away points (corners). It also extracts the keypoints for these oriented-corners, and lastly, uses the taken away the descriptor of the Hough-Transform with regard to the identification task.

Dhyaa Sh. Azzawi [15], describes the major four stages of SIFT, first: detection of Scale-Space extreme, second: Localization of Keypoint, third: Orientation Alienation, and finally Descriptor of the elected Keypoints. In a new approach, we replaced the second stage "KeyPoint Localization" with **corner detection algorithm** for localizing the interested pints.

The processes of eliciting the keypoints from source picture and the identification are described in the following stages:

4.1. Scale-Space stage:

First stage of our approach is the Scale-Space stage, which extracts the descriptor. A picture was scaled by the Gaussian filtrate into 4 octaves. Every octave, the distilled picture dimension is scaled, with specific factors, for five intervals. Thus, 20 scaled images are produced, see Figure8.

Figure (8). Pyramid scaling of the original image scaled into 4 octaves of G filter, scaled into five intervals of dimension in every octave.

4.2. Corners Detection stage:

Hessian and SUSAN algorithms used to elicit all points (corners) through all of the 20 scaled-pictures.

4.3. Orientation assignment stage:

The system assigns one or more orientation for each key point for making it invariants to rotation. As well as orientation, the scale and locations are given to each key point. Those parameters propose an iterated 2D coordinate system that describes the local image region, and thus, provides invariance to those parameters.

4.4. Keypoint descriptor stage:

The system computes a descriptor to the regional picture in this stage, discriminatory yet is as constant as potential to residual diversity, such as modification in 3Dimension viewpoint or illumination. Thus, a 16x16 Gradient-window are taken, that is split into 4x4 sub windows. A histogram of 4x4 patterns in 8 bins is initiated. This results in a features vector which containing 128 elements.

4.5. Recognition stage:

To cluster reliable model hypotheses in an efficient way, we use Hough transform for searching keys that accept an individual pattern pose. Every pattern inside the database catches a record of parameters keys equivalent to the coordinates system of the pattern. Thus, we able to form an entry data in the hash-table presaging the pattern orientation, scale, and location the match supposition. We employ a bin thirteen degrees, argument of two, and 0.25 of the maximum pattern volume for orientation, scale, and location respectively. As well as presence of substantial geometric distortion vast bin, grant to cluster, such as change in 3D.

5. Empirical Results And Analysis

The implementation was undertaken within actual pictures possessed utilizing digital camera. Adding some noise pending the conquest process (zooming, rotation, illumination, and object coordinate). The changes of zooming imply a change intensity of a pixel. Each scale change (sequence level) consists of rotated and scaled images. The horizontal direction variation is (0 to 180) degrees for each viewpoint sequences, and (0 to 60) degrees for vertical direction.

Table (1), depicts the effect of Database quantity (images number) at the performance of proposed manner, the implementation was performed using 25 images for different angle of view point and different numerous of pictures in the DB (1 to 10 pictures), and experimentally we discovered that the best choice is 8 images in the DB. See figure 9, which the chart illustrates analysis of implementation of proposed approach using different number of images in the DB.

The database consists of 8 images (Sharpener object, one Sharpener object but with 8 different degrees of viewpoint) diverse noise. In testing, used 25 images for every sequence appearing variance scenes.

A representative image consists of 1500 or extra features, where may come from various objects in addition to background disorder, and practically, we have discovered that trustworthy identification was possible within as minima as 3 features.

Table 1: Implementation of proposed approach using different number of images in DB

Times of testing processes using different pictures	Proposed approach using	Hit ratio when using different number of images in DB									
		1 picture	2 pictures	3 pictures	4 pictures	5 pictures	6 pictures	7 pictures	8 pictures	9 pictures	10 pictures
25	Harris algorithm	0 %	0.2 %	32.4 %	60.43 %	88.9 %	95 %	99.1 %	100 %	100 %	100 %
	SUSAN algorithm	0 %	0.25 %	33 %	59.8 %	95 %	99.4 %	100 %	100 %	100 %	100 %

The Hough transform recognizes clusters of features consistent version using every characteristic to poll for all poses that are regular with the characteristic. When clusters are matched to poll to the same self-pose of target, the likelihood of the version being right is a lot higher than for any monocular characteristic. Every keypoints determine 4 arguments: 2D location, orientation and scale, every identified keypoint inside the database has record of keypoint parameters proportional to the training image in which it was found.
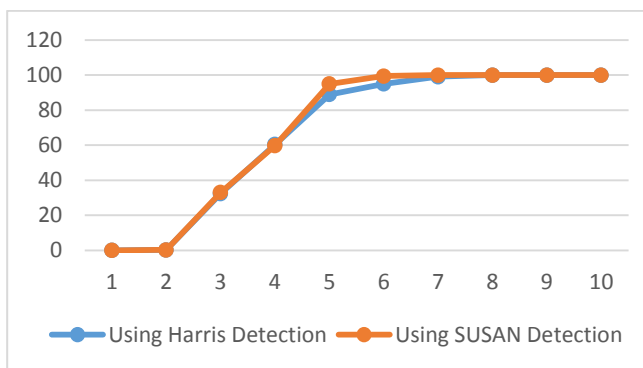

Figure (9): Chart of implementing the proposed approach using different number of images in the DB.

Thus, we initiated entry of Hough transform for prophesy the orientation, location and scale within the identified assumption. Figure10, depicts the applying to both of the new and SIFTS manner, which result identify when the noises are not exist.
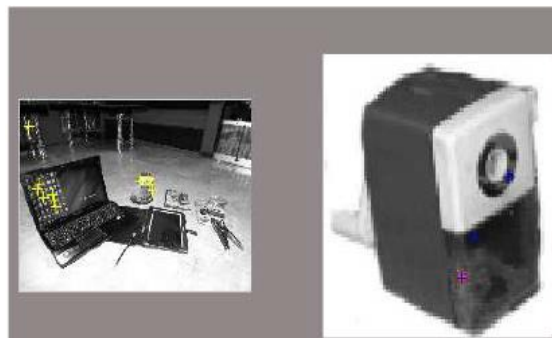

Figure (10): The applying of both the new and SIFT manner. The left displays an image comprises the example sealed with other something. The right one displays the pictures of the wanted example

Figure 11, depicts the enforcement of both of the suggested and SIFT manner together with the presence of noise (3D angle rotation, Illumination, and changes of location). Figure 11 (a), displays the success using our new manner applying the numerous of identified points (4 points) Figure 11 (b), displays the failed plying SIFT.
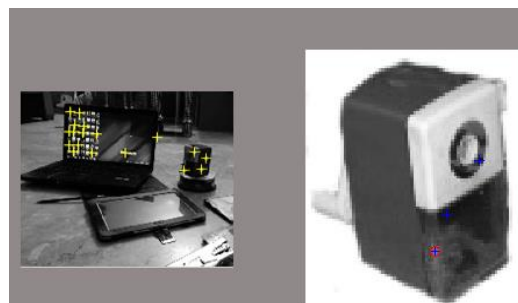
(a)

(b)

Figure (11): (a). Identifying the pattern "in the right side" of object within image of "the left side" when utilizing our new manner. (b) Misidentifying the object "right side is empty" and there are no identification utilizing SIFT manner.

Table 2: Applying of new and SIFT manner

Times of testing processes using different images	No. of images in DB	Approach	Hit ratio %			
			Normal pattern from the environment of testing image	Illumination change	3D rotation	Partial occlusion
25	8	Proposed with Harris algorithm	100%	86.6%	66.6%	83.5%
		Proposed with SUSAN algorithm	100%	82.2%	67.6%	83.5%
		SIFT	100%	16.8%	59.5%	91.8%

Table 2, displays the effect of applying both of the new and SIFT manner for normal and noised medium respectively. It depicts that our new approach gives good effects of identification in a noised medium (3D rotation and illumination variations), but the effectiveness degraded in the fractional overlapping.
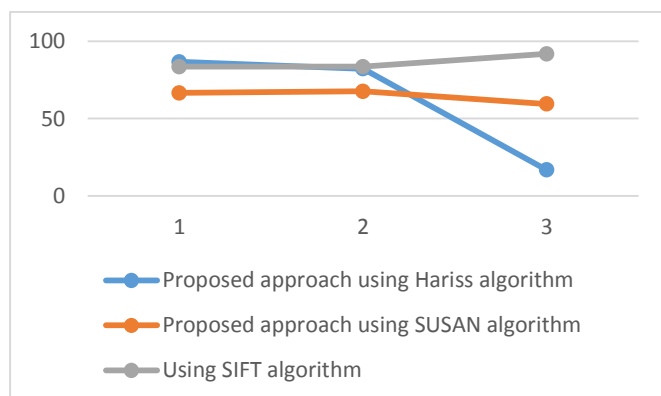

Figure (12): Chart of implementation a proposed approach using Harris and SUSAN algorithm and implementation of conventional SIFT algorithm

Figure 12 illustrates the analysis of implementing a proposed approach using two algorithms of corner detection , Hariss and SUSAN corner detection, and the implementation of conventional algorithm of SIFT. The horizontal axis of chart with numbers 1, 2, and 3 represents illumination change, 3D rotation, and partial occlusion respectively, the chart shows the high performance of proposed approach with two corner algorithms when the illumination change and 3D rotation are occur but it degraded when the partial occlusion of object is applied.

6. Conclusion

We presented a novel tactic for extracting regional (features) descriptors that was used in the Scale Integrated Features Transform method and in this new tactic; we achieved a results characterized high performance. The highlighted performance is the high speed of execution, which happed because of replacing DoG keypoints with calculated corners of detection algorithm (Harris's or SUSAN's algorithm), it extracted rapidly. From the experimental results, we discovered high performance for implementation of a proposed approach even if the high variation of surrounding features (scaling, orientation and rotation in three dimensions, pixels illumination changes). But despite the high succeed in that environment there is some degradation in it, the mismatching happened in the situations of partial occlusion of the object, it was because of the clarity of the selected object was the main factor used in the algorithm of detection depend on it.

References

[1] 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.

[2] 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.

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

[4] 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.

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

[6] 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.

[7] G. Dorko and C. Schmid, "Selection of Scale-Invariant Parts for Object Class Recognition, "Proc. Ninth Int'l Conf. Computer Vision, pp. 634-640, 2013.

[8] 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.

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

[10] 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 2015.

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

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

[13] K. Mikolajczyk, and C. Schmid, "Scale & Affine Invariant Interest Point Detector', Computer Vision, vol. 60, no. 1, pp. 63-86, 2014.[14]D. Lowe, "Object Recognition from Local Scale-Invariant Features", Proc. 7th IEEE International Conference on Computer Vision, vol. 2, pp. 1150-1157, 1999.

[15] Dhyaa Shaheed Sabr Al-Azzawi, "Eigenface and SIFT For Gender Classification", Wasit Journal for Science & Medicine, Vol. 5, No. 1, 2012.

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

[17] Trupti P. Patel, and Sandip R. Panchal, "Corner Detection Techniques: An Introductory Survey", IJEDR, Vol. 2, No. 4, 2014.

[18] Ankita C. Deshmukh, Priya R. Gaikwad, and Madhura P. Patil, "A Comparative Study of Feature Detection Methods", IJEETC, Vol. 4, No. 2, 2015.
.