Detection and Measurement of Displacement and Velocity of Single Moving Object in a Stationary Background

*Amr Reda. R. Almaddah, **Tauseef Ahmad, and ***Abdullah Dubai

***Abstract*- The traditional Harris detector are sensitive to noise and resolution because without the property of scale invariant. In this research, The Harris corner detector algorithm is improved, to work with multi resolution images, the technique has also been working with poor lighting condition by using histogram equalization technique. The work we have done addresses the issue of robustly detection of feature points, detected multiple of local features are characterized by the intensity changes in both horizontal and vertical direction which is called corner features. The goal of this work is to detect the corner of an object through the Harris corner detector with multiple scale of the same image. The scale invariant property applied to the Harris algorithm for improving the corner detection performance in different resolution of the same image with the same interest point. The detected points represented by two independent variables (x, y) in a matrix (x, y) and the dependent variable f are called intensity of interest points. Through these independent variable, we get the displacement and velocity of object by subtracting independent variable f(x,y) at current frame from the previous location f`((x,)`(y,)`) of another frame. For further work, multiple of moving object environment have been taken consideration for developing algorithms.**

***Index Terms*- Conversion of RGB images to gray, Histogram equalization, interest point detection, measurement of displacement and velocity**

I. INTRODUCTION

This is very significant task in the field of image processing that many of the detection method and algorithms has been proposed in different ways over the last decades with the passage of time more advanced method and algorithms developed for detection of objects with different way. In general, to identify and locate objects in images is a deeper understanding of the contents of an image, and where to extract information in the image, the approach to reduce the input data and focus only on that part of the object which contained highly information i.e. To detect the local features of object which can be re-detected in a different captured image [1, 2]. These detected local features are characterized by the intensity changes in both horizontal and vertical direction which is called corner features. Corner is an important local feature, defined as sudden changes of the pixel value, or the intersection point of object edges. An important characteristic of corner detection is stable result, simple extraction processes and strong adaptability to the algorithm. The most widely used corner interest point detectors are Harris corner detector but this algorithm cannot provide a scale invariant and stable corner points [3]. In our proposed work, we improved Harris algorithm to scale invariant property and capable of effective corner detection. Many algorithms developed with the passage of time for detecting interest points with different procedures. The fast detector in [4, 5] compares the brightness of center pixel with its surroundings in each node of tree to quantify salience of each location, cult point detector [6]. David Low in [7] proposed the deference of Gaussian filter, Low, developed an efficient algorithm for object recognition based on local extrema in scale space pyramid with a difference of Gaussian (LOG) [8]. The input images are convolving with a Gaussian kernel for smoothing and sampled. The smoothed images obtained with different scale used in Gaussian kernel. The difference of Gaussian obtained by subtracting two smoothed images filtered with the same scale. The difference of Gaussian algorithm (DOG) is a close approximation of the Laplacian of Gaussian (LOG) algorithm, but a dog can significantly process few images per second. The both algorithms difference of Gaussian and Laplacian of Gaussian representation is a common drawback that is local maxima can also be detected in the neighborhood of contours or straight edges, where the problem is the signal changes in only one direction. These detected maxima are less stable because its detection is more sensitive to noise or small changes in neighboring texture, to solving this problem, we select the scale for which the traces and the determinant of the Hessian matrix simultaneously assume a same local extremum. The Laplacian of Gaussian is equal to the trace of the Hessian matrix, but localizing simultaneously the maxima of the determinant imposes points for which the second derivative of images detects signal changes in only one direction. In Harris detector, the similar idea is

*Department Electrical and Computer Engineering, King Abdul Aziz University, Saudi Arabia.
** Department Electrical and Computer Engineering, King Abdul Aziz University, Saudi Arabia.
*** Department Electrical and Computer Engineering, King Abdul Aziz University, Saudi Arabia

explored using first derivatives, but the response of the second derivative is very small at the point where the signal change is most significant. In our work, we recover the weak response of second derivative by convolving another kernel with the second derivative for localizing strong and stable points and we also explain with detail in our work that how to detect the scale invariant feature by Harris detector. The applied scale invariant property of Harris detects or verifies that it is independent with image resolution. We noticed from several differential operators that the scale adopted Harris detector measures rarely achieve maxima our scale in a scale space. One important challenge in our work is finding the velocity of a detected point of an unknown object from one frame to the next. Usually, a single feature point cannot be tracked, unless it has a very distinctive brightness with respects to all of its neighbors. In fact, the value of a point of an unknown object can heavily change due to noise or occlusion. Detecting and selecting the right feature plays a critical role in tracking. In general, the most desirable property of a visual feature is its uniqueness so that the object can be easily distinguished in the feature space. The interesting point is that area of object which can help us in finding the velocity of an object from one frame to another frame. The cafforio and Rocca in [9] describe a method for segmentation and measurement of displacement of a single moving object in a stationary background. The algorithm we developed here is quite unique from other live applications. Four images have been processed for finding the displacement and velocity of the object. These static imaging approaches create the base for the object tracking in real time video analysis from live camera surveillance. More important technique implemented to overcome the challenges in live tracking system [10]. Where the collected information from video surveillance are used for further data analysis. Johan Sommerfeld also proposed a technique for tracking objects from single camera [11]. For expanding the domain of this work we mapped the images directly to the object properties of image to real world scenarios by redirecting the same interest point over a sequence of images there are many key point descriptors includes for finding descriptor of feature points sift [5], surf [9], brief [10], brisk [11], fern [12], and bin boost [13]. From detected object in the image we relate the object properties, i.e. displacement and velocity of a moving object and the velocity calculation method which can be explored to support other system.

II. CONVERSION OF RGB IMAGES TO GRAY

In this work, we are only interested to detect the interest points of the object in the frame. So, there is no requirement of the full information of the images such as color information or color specification, etc., with these information's it will consume the memory as well as create the time complexity for any computation. At the starting of the process the images are converted into grayscale images. The formula we use here=0.299*R+0.566*G+0.1*B applicable to convert RGB images to grayscale. But the simplest method RGB to grayscale algorithm is Intensity [14]. It is the mean of the RGB channels.

$$G_{\text{intensity}} = \frac{1}{3}(R + G + B) \tag{1}$$

Fig. 1. RGB vs Grey Scale

III. HISTOGRAM EQUALIZATION

In this work, we suggest to apply histogram equalization for the camera with bad quality or poor lighting condition. The histogram equalization a technique using for adjusting the image intensities. It is the statistical representation of pixel intensities.

Before performing histogram equalization, we must know about two important methods used in equalizing histograms. These two concepts are known as Probability Mass Function (PMF) and Cumulative Distribution Function (CDF). First, we have to calculate the probability mass function of all the pixels and then calculate cumulative distributive function for all the pixels of the image.

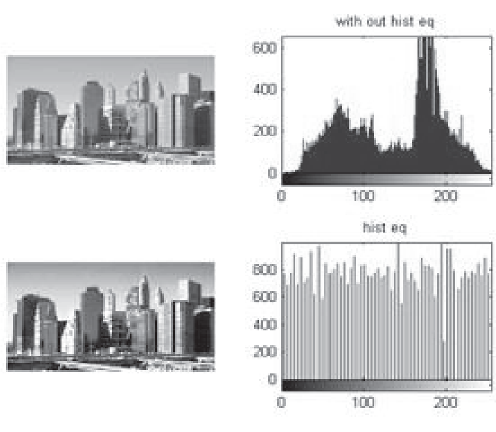

Fig. 2. Result with & without histogram

IV. INTEREST POINT DETECTION

Harris interest point detection

Through Harris detector, we can easily locate the corner point by looking through a shifted window with the shifting window in any direction of an object give a large change in intensity. The Harris detector computes the 1st order derivatives in both X and Y direction for highlighting the directional intensity variation of pixels,

First compute image derivatives (Gaussian)

$$I_x = G^x\sigma^* I I_y = G^y \sigma^* I \quad (2)$$

Compute product of derivatives at each pixel

$$I_{x^2} = I_x \,.\, I_x, \quad I_{y^2} = I_y \,.\, I_y, \; I_{xy} = I_x I_y \quad (3)$$

Compute the sum of the product of derivatives at each pixel

$$S_{x^2} = G\sigma^* l_{x^2}, S_{y^2} = G\sigma^* l_{y^2}, G_{xy} = G\sigma^* l_{xy}, \quad (4)$$

Then find the matrix for each pixel to evaluate each pixel. The idea behind to detect corner points with strong gradients in two directions are easier to distinguish from their surrounding neighborhood and therefore easier to re-detect in another image than homogeneous regions or edges of the object.

$$H(x,y) = w(x,y)\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \quad (5)$$

Where, f_x is the gradient in x direction, f_y is the gradient in y direction, w(x,y) is the window template. Now compute the response of the detector at each pixel

$$R = D(H) - k(trc(H))^2 \quad (6)$$

Weir, R is the response function of the corner point required, set (H) is the Matrix determinant, tr (H) is the Matrix trace, k is the default constant, and generally values are 0.04-0.06. The center value R of an image is calculated and if the value response is maximum in the neighborhood and larger than a given threshold, then the point is corner point.

Scale invariant Harris detector

The Harris detector is based on a matrix which is called auto correlation matrix is often used for feature detection. This matrix must be adopted a property of scale changes to make it independent with image resolution as shown in equ. (15). In this proposed work, first we find the double derivative of the image in the x direction and y direction by convolving image with 3*3 kernel in equ. (7,8).

$$\frac{\partial^2 I(x,y)}{\partial x^2} = \frac{\partial I(x,y)}{\partial x} * \frac{\partial I(x,y)}{\partial x} \quad (7)$$

$$\frac{\partial^2 I(x,y)}{\partial y^2} = \frac{\partial I(x,y)}{\partial y} * \frac{\partial I(x,y)}{\partial y} \quad (8)$$

$$I_x(x,y) = \frac{\partial^2 I(x,y)}{\partial x^2} * h_x \quad (9)$$

$$I_y(x,y) = \frac{\partial^2 I(x,y)}{\partial y^2} * h_y \quad (10)$$

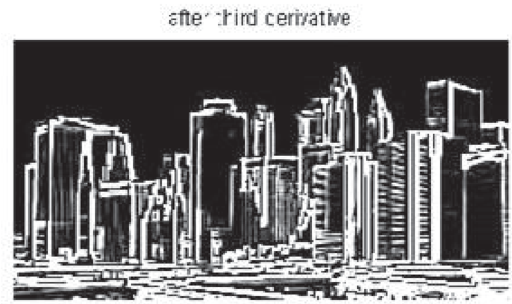

Fig. 3. High resolution invariant

Fig. 4. Medium resolution invariant

Fig. 5. Low resolution invariant

The fig (3,4,5) we get from the mathematical eq. (9,10).

$$G(x,y,\sigma) = e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (11)$$

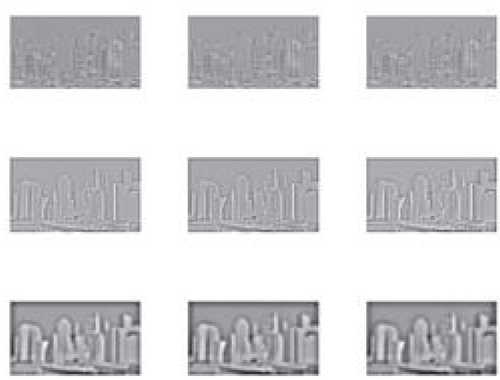

Fig.6. This image explain the scale of the image.

$$I_x(x,y,\sigma_n) = I_x(x,y) * G(x,y,\sigma_n) \quad (12)$$

$$I_y(x,y,\sigma_n) = I_y(x,y) * G(x,y,\sigma_n) \quad (13)$$

$$I_{xy}(x,y,\sigma_n) = I_{xy}(x,y) * G(x,y,\sigma_n) \tag{14}$$

$$M = G(x,y,\sigma_n) * \begin{bmatrix} I_{xx}(x,y,\sigma_n) & I_{xy}(x,y,\sigma_n) \\ I_{xy}(x,y,\sigma_n) & I_{yy}(x,y,\sigma_n) \end{bmatrix} \tag{15}$$

The σ_n is the differentiation scale and (I_{xx}, I_{yy}) is the second derivative in both directions (x,y), the matrix describes the gradient distribution in a local neighborhood of a point. The local derivative is computed with [3*3] matrix kernel. In the neighborhood of points the derivative is averaged by smoothing with a Gaussian window. The eigenvalues of this matrix are represented two principal signal changes in the neighborhood of detecting points. This property enables point extraction for which both curves is important, that is the signal change in both directions is important, i.e. corners points, such point is more stable and better representative of an image. Now here we are combining the trace and determinant of a scale invariant matrix.

$$corner = \frac{\det(M)}{trc(M)} = [(I_{xx})(I_{yy})] - (I_{xy})(I_{xy})/(I_{xx} + I_{yy}) \tag{16}$$

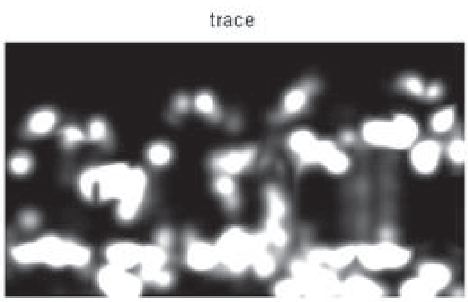

Fig.7. We get maxima of interest points from the eq.(16)

Fig. 8. High Resolution detecting corner points

Fig. 9. Medium Resolution detecting corner points

Fig. 10. Low Resolution detecting corner points

Here the local maxima of above eq (16), determine the location of interest points in fig (h, I, j) with multiple resolution images, we reject the points which get below the threshold value or the second derivative no extremum. In this way, we get more stable points with associated scale. For some point the scale peaks may not correspond to the selected scale of an image. These points are rejected because of the lack of a maximum or the scale is not very accurate.

Fig.11. High Resolution Local Maxima

Fig.12. Medium Resolution Local Maxima

Fig.13. Low Resolution Local Maxima

V. MEASUREMENT OF DISPLACEMENT AND VELOCITY

The method we explained for finding displacement and velocity of an object in a stationary background. we get the detected points from equ. (16) represented by two independent variable (x,y) in a matrix (x,y) and the dependent variable f is called intensity of interest points. The located interest point of object in matrix of (x, y) in the current frame is subtracted from the matrix f` ((x,)` (y,)`) in the previous frame through by Pythagorean equation equ. (17). The covered distance of an object from one frame to another frame are obtained through the located pixel.

The formula for finding distance as:

$$Dx = \sqrt{(x - xi)^2 + (y - yi^2)} \tag{17}$$

As with the movement of an object from one location 'a' to the location 'b' we capture the coverage area of the object. Then we suppose that the movement of an object 72 pixels is equal to the 1 inch of the real-world measurement distance. If we see the object has moved x pixel, then the real-world displacement x of object will be mapped. To obtain the velocity of object convertor map the pixel covered by object in frame with the real world and the time taken by object from one frame to another it depends on frame per second of video

Time of one frame, $dt = \frac{1}{FPS}$ (18)

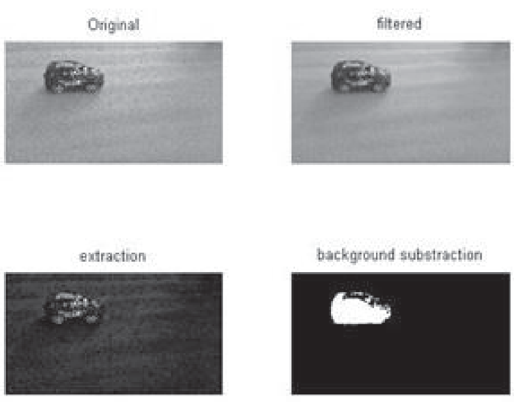

Fig.14. Original vs Filter Image

Fig.15. Combination Result

VI. CONCLUSION

In this paper, we proposed two novel approaches scale invariant point detection and velocity calculation of an object. None of the previous work detects a stable feature with scale invariant property. In our proposed work, we improved detection, stability and robustness of detection. The algorithm we developed working with the grayscale images. The method is used to detect and track the moving object from one frame to another frame from a clutter background and give its total velocity. This work may be incorporated in a robotics, manufacturing or monitoring system for operating upon objects moving along an assembly line.

REFERENCES

[1] Harris, C., & Stephens, M. (1988, August). A combined corner and edge detector. In *Alvey vision conference* (Vol. 15, No. 50, pp. 10-5244).

[2] Shi, J. (1994, June). Good features to track. In Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on (pp. 593-600). IEEE.

[3] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), 91-110.

[4] Se, S., Lowe, D., & Little, J. (2002). Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. The international Journal of robotics Research, 21(8), 735-758.

[5] Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). Computer vision and image understanding, 110(3), 346-359.

[6] Riemenschneider, H., Sternig, S., Donoser, M., Roth, P. M., & Bischof, H. (2012, October). Hough regions for joining instance localization and segmentation. In European Conference on Computer Vision (pp. 258-271). Springer, Berlin, Heidelberg.

[7] Rosten, E., & Drummond, T. (2005, October). Fusing points and lines for high performance tracking. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on(Vol. 2, pp. 1508-1515). IEEE.

[8] Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., & Fua, P. (2012). BRIEF: Computing a local binary descriptor very fast. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(7), 1281-1298.

[9] Cafforio, C., & Rocca, F. (1976). Methods for measuring small displacements of television images. IEEE transactions on Information Theory, 22(5), 573-579

[10] Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In Computer Vision (ICCV), 2011 IEEE International Conference on (pp. 2548-2555). IEEE.

[11] Lepetit, V., & Fua, P. (2006). Keypoint recognition using randomized trees. IEEE transactions on pattern analysis and machine intelligence, 28(9), 1465-1479.

[12] Trzcinski, T., Christoudias, M., Fua, P., & Lepetit, V. (2013, June). Boosting binary keypoint descriptors. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on (pp. 2874-2881). IEEE.

[13] Das, P., Ghoshal, R., Kole, D. K., & Ghosh, R. (2012). Measurement of Displacement and Velocity of a Moving Object from Real Time Video. International Journal of Computer Applications, 49(13).

[14] Kanan, C., & Cottrell, G. W. (2012). Color-to-grayscale: does the method matter in image recognition. PloS one, 7(1), e29740.