Object Motion estimation using edge detection and background subtraction with block matching algorithm

Shardul Thapliyal[1], Yogendra Pratap Pundir[2], Arun Shekhar Bahuguna[3], Sunil Semwal[4]
[1, 2, 3,4] Department of Electronics and Communication Engineering, SOET,
HNB Garhwal University, Srinagar Garhwal

Abstract:—**Detection of an object motion is the growing research field of image processing which revealed the several applications. Several techniques (including the proposed one) are discussed so far in literatures. In this paper the edge detection and frame differencing also known as background subtraction technique with block matching algorithm has been implemented to detect the object motion. The object taken for experimentation is arbitrary having no fixed shape and size. The MATLAB output result showing the practicability of the both algorithms.**

Index Terms—**edge detection, block matching, filtering, Thresholding**

I INTRODUCTION

Following a object is characterized as procedure of finding the element of interest for a video stream and monitoring its movement, introduction and so forth outline by casing in whole video stream. Following item in edges require arrangement of division ventures to detach the question from its environment or foundation. Such comparable strides of edge recognition and edge differencing with layout coordinating were utilized as a part of the calculation, which are portrayed in detail later in this paper. Recognition of question in the video stream is the initial move towards extraction of significant data in numerous PC vision application, including activity observing, video observation and following individuals. Object following is the essential progressing research range in the field of PC vision and mixed media. Object following manages assortment of clients, for example, human PC communication, reconnaissance and security, pressure, restorative imaging and so on. For the most part object following is tedious process because of the measure of information required to be prepared. Henceforth, time of handling is kept low by utilizing calculation. Promote, many-sided quality is added because of the need of acknowledgment of object in every edge of video.

II METHODOLOGY

MATLAB provides us various toolboxes to make our work easier. Two such toolboxes are image/video processing toolbox and image/video acquisition toolbox [1]. Since the completion of this research work requires both the toolboxes. Object detection is a process of detecting object in a video or live feed. Challenges are met in object detection due to many variations in visual appearance. For example, vehicles vary in shape, size color and in small details like tiers, grills, etc. over a wide range. Appearance of an object also depends on environment. Light source may vary in intensity or may be changing position with respect to object and in addition varying shadows may affect the intensity of object in different frames. The appearance also depends on stance of an object like man could be standing or sitting; front view of face is different than side view, etc. Object detection must occur even if there is change in stance in different frames of video. Hence to get rid of the anomaly two part strategy is used for object detection.

There are various methods available for object detection like; Statistical Method for Object Detection, and Robust Real-time Object Detection

Background Subtraction: According to the name of this method it recommends that, it is the way toward isolating the forefront and foundation of the picture. Here it is expected that closer view contains the objects of intrigue. Everything with the exception of a question is said to be a foundation. For instance leaves, trees, structures and so on are the foundation. In a given picture (generally prone to be a video outline), we need to recognize the frontal area questions in that picture. By and large, items are of intrigue, not the scene i.e. the articles in a video other than the objects of intrigue are of no utilized. Foundation subtraction is a generally utilized approach for identifying moving articles in recordings from static cameras. The reason is that of identifying the moving articles from the distinction between the present casing and a reference outline, called foundation display. The foundation picture must be a representation of the scene with no moving items and must be kept consistently overhauled in order to adjust to the changing illuminating presences conditions and geometry settings [2]. The more straightforward calculation processes the distinction between the present edge and the foundation, relegating a name: either 0 if the question has a place with the foundation or 1 if the object is in forefront. The outcomes would be tasteful if the articles shading, being on the scene, was not the same as foundation. In the event that another object enters in the zone

marked as closer view, it won't identify and name accurately. Another negative part of this is the accompanying: if a foundation object moves, both the genuine question and its apparition are identified. In addition, Foundation Subtraction is extremely touchy to enlightenment and foundation changes and it doesn't deal with camera developments despite the fact that we have attempted to limit the imperatives. Technique for frontal area extraction is contrast picture strategy.

Image differencing: In this method the difference of images is used in order to find object which is moving and the object which is not moving. The result of the difference is store as another grey image called the difference image [3]. Three types of difference images are defined.

• Complete accumulative subtraction or difference image is given by

$$f(x, y) = f(x, y) + 1 \ldots\ldots\ldots\text{if } |g(x, y, t_{i+1}) - g(x, y, t_i)| > T$$

• Positive accumulative subtraction or difference image is given by

$$f(x, y) = f(x, y) + 1 \ldots\ldots\ldots\text{if } g(x, y, t_{i+1}) - g(x, y, t_i) > T$$

• Negative accumulative subtraction or difference image is given by

$$f(x, y) = f(x, y) + 1 \ldots\ldots\ldots \text{ if } g(x, y, t_i) - g(x, y, t_{i+1}) > T$$

A consistent and strong background subtraction methodology algorithm should take care of the abrupt or regular illumination changes, High frequency, recurring motion in the background i.e. tree leaves, fog, polluted surrounding, body heat and Long-term outlook changes like a vehicle is parked for a month.

III Algorithms

A. Edge Detection

Edge ditection is the most widely recognized way to deal with get the most significant discontinuities in intermittence values. Such discontinuities are identified utilizing first and second request subsidiaries [4]. The fig. 1 is demonstrating the yield consequence of edge discovery.

The main request subordinate of decision in picture preparing is inclination. We rehash the germane conditions here for comfort .The inclination of 2D capacity f(x, y) is characterized as a vector.

Fig.1: output of edge detection algorithm

$$\Delta f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

The magnitude of vector is

$$\nabla f = mag(\nabla f) = [G_x^2 + G_y^2]^{1/2}$$
$$= [\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial x}\right)^2]^{1/2}\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

To simplify computation this quantity is approximated sometimes by omitting the square root operation:

$$\nabla f \cong G_x^2 + G_y^2\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

These estimation still behaves as derivatives i.e. they are zero in areas of continuous intensity and their values are proportional to degree of intensity change in areas whose pixel values are variable[5]. It is common practice to refer to magnitude of gradient or its approximation simply as 'gradient'. The fundamental property of gradient vector is that it points to direction of maximum rate of change of coordinate f(x, y).

2nd Order derivative of function f(x, y):

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}\ldots\ldots\ldots\ldots\ldots(4)$$

They are computed using laplacian. In MATLAB this is done by function 'edge'.

[g t]=edge (f, 'method', 'parameter')

Where f is image and method is one of the procedures

A. Frame Differencing

The frame differencing is a strategy where the PC checks the contrast between two video outlines. On the off chance that the pixels have changed there obviously was something changing in the picture (moving for instance). Most strategies work with some obscure and limit, to particular genuine development from clamor. Casing could vary when light conditions in a room change (and camera auto center, shine adjustment and so forth). On the off chance that the foundation is stationary, the face area can be secluded from the foundation by first catching the foundation picture and after that subtracting it from each consequent picture. Districts of the distinction picture with high plentifulness are thought to be face areas [6].

Firstly the image is investigations by the system, which is taken by the camera, for identification of any moving article. The Edge Differencing calculation is utilized for this reason, which gives as yield the position of the moving article in the picture. This data is then used to separate a square picture layout (of settled size) from that locale of the picture. The formats are

created as and when the presence of the object changes essentially. The Fig.2 is the yield of the edge differencing calculation.

B. *Thresholding*

Thresholding is a non-straight operation that changes over a dark scale picture into a parallel picture where the two levels are appointed to pixels that are beneath or over the predetermined edge esteem. It is however significantly more productive to utilize the Picture [7]. Fig. 3 demonstrates the yield of thesholding. Thresholding operation, which likewise give a strategy to finding the "ideal" limit an incentive for a given picture. One technique that is moderately straightforward, does not require much particular information of the picture, and is hearty against picture clamor, is the accompanying iterative strategy
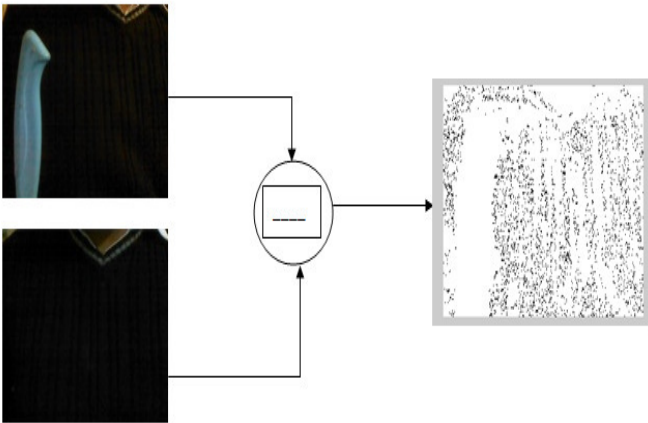

Fig 2: Result after Frame Differencing

1. Initially threshold (T) is chosen; this can be done arbitrarily or done according to any other desired method.
2. Image is divided into the object and the background pixels as described above, creating two sets:
   1. $G_1 = \{f(m, n): f(m, n) > T\}$ (object pixels)
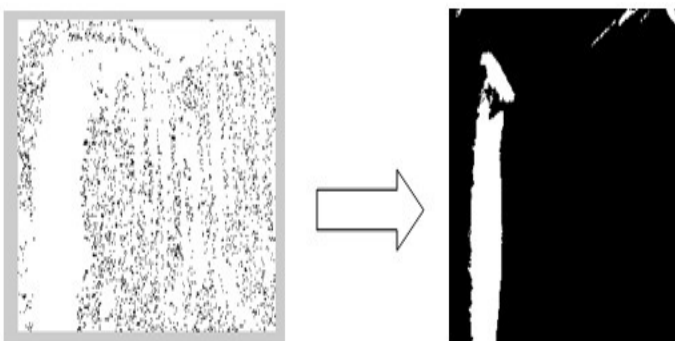   2. $G_2 = \{f(m,n): f(m,n) < T\}$ (background pixels)

Fig 3 Image after Thresholding

C. *Median Filtering and morphological erosion*

Among the various type of filtering techniques available for improving the quality of our image the most the median filtering is used to avoid any flicker noise. Further, morphological erosion operation is performed on the thresholed image to remove any variation in background which might have slipped into the thresholded image [8]. Fig.5 showing the result of morphological erosion it is clear from the visual inspection that the variation due to external element is removed by this algorithm.
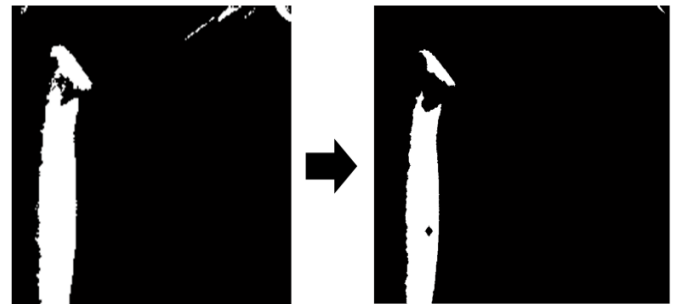
Sample image   After filtering and erosion

Fig 4: image after morphological erosion

D. *Generation of Centroid*

The vast majority of these techniques give therefore the in all probability focal pixel of each recognized question, either specifically by method for a channel reaction or in a roundabout way by method for the centroid of all focuses on the limit shape. At the point when the places of the items should be known with sub-pixel exactness, a precise and strong gauge can be gotten by figuring its focal point of gravity [8]. In the event that the acknowledgment calculation brings about a paired yield (pixel having a place with the question or not), the focal point of gravity of every object is decreased to registering the normal arrange along the (x, y and potentially z) tomahawks among the pixels or vexes that are connected with the perceived question. At the point when, then again, the recognition brings about an arrangement of yield qualities circulated in an area around the question focus (e.g., a channel reaction) and the yield esteem is directly identified with the separation to the object focus, the weighted focal point of gravity may bring about a more correct question area. In the wake of distinguishing the question from the video we have to keep a track on the object we are occupied with.

E. *Calculation of Centroid*

Estimation the center of gravity (COG) or centroid of a binary image needed a set of two integers C(cog_x ,cog_y) which determines the position of the moving object in the given scene.

The COG is calculated by:

$$cog\_x = cog\_x + x \ ......(1)$$

$$cog\_y = cog\_y + y \ ......(2)$$

$$Total = Total + 1 \ ....(3)$$

For each pixel where x, y is the current location of picture, the resulting COG is then divided by the Total value:

$$cog\_x = cog\_x/Total \ ....(4)$$

$$cog\_y = cog\_y/Total \ ...... (5)$$

Fig. 5 Centroid identification

F.   Template matching

The layout picture has solid components, an element based approach might be viewed as; the approach may demonstrate facilitate helpful if the match in the hunt picture may be changed in some mold [9]. Since this approach does not consider the sum of the format picture, it can be all the more computationally proficient when working with source pictures of bigger determination, as the option approach, layout based, may require looking possibly a lot of focuses keeping in mind the end goal to decide the best coordinating location.
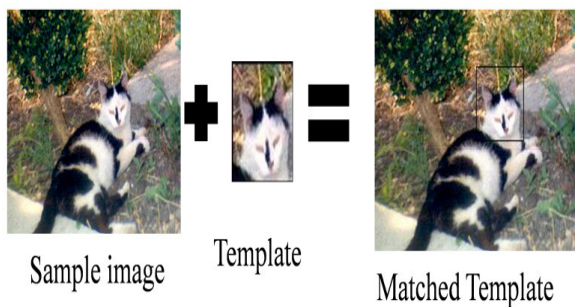

Fig.6. Template matching

V Conclusion

A strong and effective computerized single object following framework is introduced. The framework has been executed utilizing a calculation in view of casing differencing and element format coordinating. The calculation has tentatively been appeared to be very precise and viable in recognizing a solitary moving article even under awful lighting conditions or impediments. Such a mechanized question following framework can be utilized as a part of utilizations where exact following is required yet great lighting conditions can't be given. The framework is additionally especially appropriate to zones like reconnaissance and video conferencing. Future work concentrates on following numerous items in the meantime and also on enhancing tracker exactness amid camera movement

References

[1]. Yilmaz, A., Javed, O., And Shah, M. 2006. Object tracking: Asurvey. *ACM Comput. Surv. 38*, 4, Article 13 (Dec. 2006)

[2]. Jain, R. And Nagel, H. 1979. On the analysis of accumulative difference pictures from image sequences of real world scenes. *IEEE Trans. Patt. Analy. Mach. Intell. 1*, 2, 206–214.

[3]. Comaniciu, D., Ramesh, V., And Meer, P. 2003. Kernel-based object tracking. *IEEE Trans. Patt. Analy. Mach. Intell. 25*, 564–575.

[4] Bar-Shalom, Y. And Foreman, T. 1988. *Tracking and Data Association*.Academic Press Inc

[5] L. Wang, W. Hu, And T. Tan. Face tracking using motion guided dynamic template matching. In *ACCV*, 2002.

[6] D. Murray And A. Basu, "Motion tracking with an active camera," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 449–459, May 1994.

[7] Rafael C. Gonzalez, Richard E. Woods. (2002): Digital Image processing, Second Edition. Prentice Hall International

[8] Edward R. Dougherty (1993) Mathematical Morphology in Image Processing. CRC Press.

[9] A. K Jain, digital image processing, second edition,Tata mcgraw hill.