*International Journal of Engineering & Technology*, 7 (4.6) (2018) 78-81

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***



# Tracking and Size Estimation of Objects in Motion based on Median of Localized Thresholding

**Sagar Gujjunoori*,[1], Madhu Oruganti[2], N. Aparna[1], M. Srija[1] , Chaitrali Dangare[3]**

*1 Department of Computer Science and Engineering*
*Vardhaman College of Engineering*
*Telangana, India*
*2 Department of Electronics and Communication Engineering*
*Sreenidhi Institute of Science and Technology*
*Hyderabad, Telangana (501301), India*
*[3]Department of Computer Science and Engineering MLR Institute of Technology, Hyderabad*
**Corresponding author E-mail: sagar.g@vardhaman.org*

**Abstract**

Motion detection and tracking play an important role in Computer vision and Robotics. Optical flow based methods to estimate the motion are widely explored during the last decade. The motion information retrieved from these techniques has enormous applications. Video analysis based on the size, speed, and directions of objects have wider applications in computer vision, robotics and watermarking. Segmentation of moving objects based on the optical flow is very challenging. In this paper, we present a model to estimate the size of a moving object based on the optical flow technique and present localized thresholding technique. Over segmentation is reduced by the proposed local thresholding technique and use of bilateral filtering. We compare our results with Sagar et al. scheme.

***Keywords***: *Filtering; Object segmentation; Machine vision; Motion estimation; Localized thresholding; Bilateral filtering;*

## 1. Introduction

Motion detection, tracking and classification of objects is a challenging task and an emerging area of research that have wider applications from agriculture to space [2], [3]. Feature detection such as size, shape, speed, direction of objects in motion is emerging in computer vision and robotics. The applications of estimating the size of an object in motion are discussed in Sagar et.al scheme [1]. Estimating the size of an object in motion enables a machine or human observer to take certain decisions in different perspectives [1].

Sagar et.al [1] have proposed a model to estimate the size of an object in motion using a thresholding function called MMM thresholding. This function was used to segment the moving objects using optical flow technique. As the MMM thresholding is a global thresholding technique it considered even small magnitudes of optical flow, hence resulted in improper segmentation of the moving objects. The small movements such as leafs of a tree or gray level changes due to light variations have also been detected as moving objects.

To minimize these problems, we propose to use Bilateral filtering technique before applying the optical flow technique for segmenting the size of an object. This is to retain the edge areas of a moving object while smoothing the smaller variations in gray levels. Further, we propose a localized thresholding technique which makes use of the maximum magnitudes of the optical flow patches in deciding the final threshold for segmenting the moving objects. The main idea behind the utilization of maximum magnitude value of every patch in an image sequence is to pick the most significant magnitude values which determine the moving object effectively. We compare our results with Sagar et.al scheme [1] and we use the Farneback dense optical flow technique to estimate optical flow as in [1]. Our experimental results demonstrate that the proposed technique out perform the state of art.

## 2. Proposed Method

We propose a model to efficiently estimate the size of an object in motion. We compare our model with Sagar et.al scheme [1]. The proposed model is shown in Figure1. The proposed model in Figure1 takes two consecutive frames $F_i$ and $F_{i+1}$ from the image sequence and outputs the masked images which contain different objects along with their sizes as in Sagar et.al scheme. The proposed model is broadly divided into four stages namely, Bilateral Filtering, Motion estimation, Segmentation, and Size estimation of objects as in Figure1. We discuss the above stages of the proposed model in the following sections.

### 2.1. Bilateral Filtering

In Sagar et al. scheme [1], even small moving regions such as leafs of a tree and gray level variations due to light changes have been detected as moving objects, hence leading to over segmentation as in Figure 2, which resulted by the use of MMM theshold-ing function in their scheme. To minimize such undesired over segmentation, we propose to employ the Bilateral Filtering on the video sequence before the motion estimation.

**Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.**

The Bilateral Filter [4], [5], [6] is a weighted smoothing filter while preserving the edges and it is defined as in equation (1).

$$B[F_i(x,y)] = \frac{1}{W_{(x,y)}} \sum_{(s,t)\in M\times N} G_{\sigma_{M\times N}}(\| (x,y) - (s,t) \|) G_{\sigma_{[0,255]}}(F_i(x,y) - F_i(s,t)) F_i(s,t) \quad (1)$$

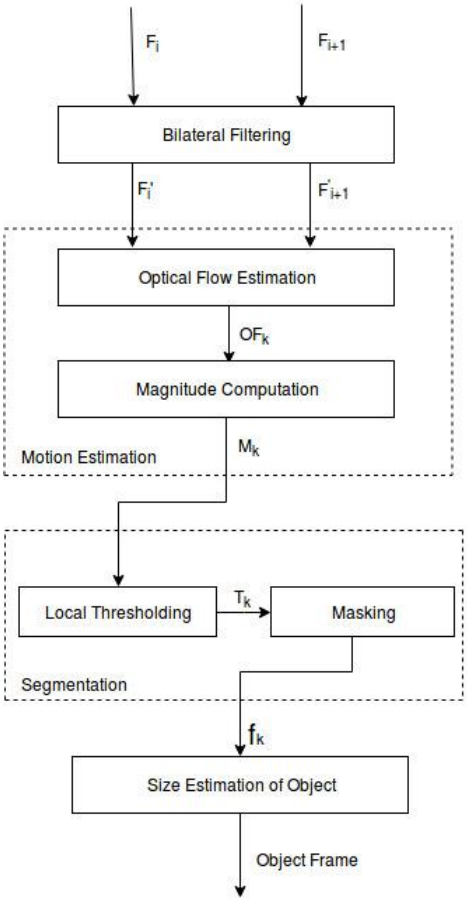
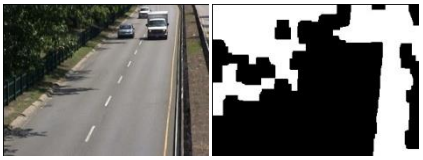



**Fig. 1**: Proposed model



(a) Original frame (b) Segmented mask

**Fig.**2: Over segmentation

$$W_{(x,y)} = \sum_{(s,t)\in M\times N} G_{\sigma_{M\times N}}(\| (x,y) - (s,t) \|) G_{\sigma_{[0,255]}}(F_i(x,y) - F_i(s,t)) \quad (2)$$

And $G_\sigma(\alpha) = \frac{1}{2\Pi\sigma^2}\exp(-\frac{\alpha^2}{2\sigma^2})$ The filtered gray scale image sequences $F_i$ and $F_{i+1}$ are passed to the motion estimation stage as in Figure1.

## 2.2 Motion Estimation

For estimating the motion, Optical flow based techniques are popular. Optical flow based techniques, in simple terms, estimate the current pixel position in the next frame. There are three popular optical flow algorithms to deal with motion, namely, Horn and Schunk algorithm [7], Lukas Kanade algorithm [8] and Gunnar Farneback algorithm [9]. We use Gunner Farneback algorithm, as it is a dense optical flow and computes the flow for all the points in the image sequence when compared to the other optical flow algorithms. Consider the output of the Gunner Farneback optical flow algorithm as $OF_k$ the kth optical flow vector computed between $F_i$ and $F_{i+1}$, where $0 \le k \le L-1$, and L is the total number of frames in the given video sequence. Note that $OF_{k(x,y)}$ is a displacement vector for the pixel position (x; y) in the $F_i$, i.e, the pixel at $(x,y)$ in $F_i$ is displaced by $OF_{k(x,y)}$ in $F_{i+1}$.

Using this displacement vector $OF_k$, we can identify whether the pixels (objects) are moving or not. Usually the pixels in a moving object will have some correlation between the magnitudes, where the magnitudes of the flow vector $OF_k$ are computed using the following equation(3).

$$M_k(OF_k) = \| OF_k(x,y) \| = \sqrt{x^2 + y^2} \quad \textbf{(3)}$$

## 2.3 Segmentation

Once the magnitude are computed by using equation (3), we should be able to automatically distinguish between moving objects and non moving objects which is very challenging. It is performed in segmentation stage in Figure1. Sagar et.al scheme [1] have used the following $MMM$ Thresholding function to find a threshold which is used to distinguish the moving and non moving objects as in (4).

$$T = MMM(M_k, M_{k+1}) = \frac{1}{2\times M\times N}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} M_k(x,y) + M_{k+1}(x,y) \quad (4)$$

Where $M_k$ and $M_{k+1}$ are the two consecutive magnitude matrices and $M\times N$ is the size of a frame and magnitude matrix $M_k$. The magnitude computation in equation (4) is computed globally, i.e, it includes even smaller magnitude values hence resulting in over segmentation. To limit the contribution of smaller magnitudes in deciding the threshold to detect the moving object, we propose a localized thresholding technique.

Let $[p_1, p_2, p_3..., p_m]$ be the $7\times 7$ patches considered from the magnitudes $M_k$ of size $M\times N$. Assume that M and N are multiples of 7 or can be padded with repeated boarder values. The proposed thresholding function is given as below in equation (5).

$$T_k = Median(t_1, t_2, t_3..., t_{\frac{m}{2}}) \quad (5)$$

Where, median(max(P1), max(P2), max(P3),...,max(Pm)) and $\max(p_i)$ returns the maximum magnitude value in the patch

$p_i$. We refer the resulting threshold as localized threshold. Based on the threshold value $T_k$ the masked image $f_k$ is constructed using the following equation in (6)

$$f_k = \begin{cases} 255, M_k(x,y) \geq T_k \\ 0, otherwise \end{cases} \quad (6)$$

### 2.4 Size Estimation

Using the masked frames $f_k$, the size of the object is estimated same as in Sagar et.al scheme [1], by labelling their corresponding sizes based on the count of segment objects binary values. The output of size estimation is called object frame.

## 3. Experiments and Results:

The proposed model is compared with Sagar et. al scheme [1]. As it is stated that the performance of size estimation is directly proportional to the performance measures of moving object segmentation. We make use of the segmentation measures to compare our results. We present the following two segmentation measures as in equation (7) and (8).

$$\text{Precision} = \frac{TruePositive}{TruePositive + FalsePositive} \quad (7)$$

$$Specificity(SP) = \frac{TrueNegative}{TrueNegative + FalsePositive} \quad (8)$$

In the experimentation we used the image sequences available at the Changedetection.net [10]. From the Table I and Figure 3, we can observe that using the proposed model the precision is increased. The percentage of improvement in precision can be in seen in Table II. And from Table III and Figure 4, it can be observed that the specificity measure using our proposed technique is almost nearer to 1. Hence we improved the over segmentation as shown in Figure 5 and Figure 6 caused by Sagar et.al scheme [1].

**Table I:** Comparison of Precision

| Dataset | Proposed scheme | Sagar et.al scheme [1] |
|---|---|---|
| Highway | 0.8716 | 0.8673 |
| Pedestrians | 0.6171 | 0.5847 |
| Office | 0.4718 | 0.4459 |
| PET2006 | 0.7046 | 0.7031 |
| Fountain | 0.1694 | 0.1076 |
| Canoe | 0.538 | 0.5233 |
| Turnpike | 0.3532 | 0.3404 |

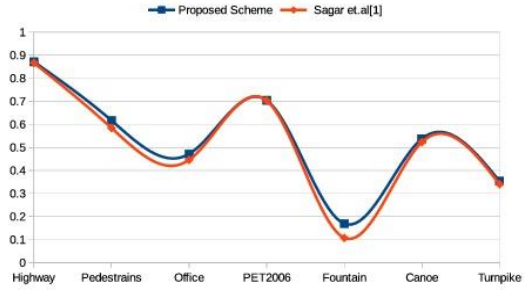



**Figure 3:** Comparison of Precision

**Table II:** Percentage of Improvement in Precision

| Dataset | % of Improvement |
|---|---|
| Highway | 0.0043 |
| Pedestrians | 0.0324 |
| Office | 0.0259 |
| PET2006 | 0.0015 |
| Fountain | 0.0618 |
| Canoe | 0.0147 |
| Turnpike | 0.0128 |

**Table III.** Comparison of Specificity

| Dataset | Spec (Proposed) | Spec(Sagar et.al) |
|---|---|---|
| Highway | 0.9915 | 0.9902 |
| Pedestrians | 0.9969 | 0.9948 |
| Office | 0.9999 | 0.6573 |
| PET2006 | 0.9984 | 0.9965 |
| Fountain | 1.0007 | 0.9908 |
| Canoe | 1.001 | 1 |
| Turnpike | 0.9784 | 0.8964 |

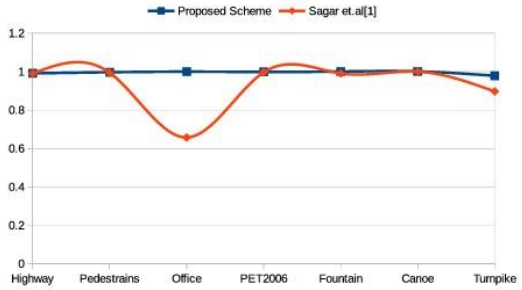



**Figure 4:** Comparison of Specificity



(a) By Sagar et. al [1] (b) Original frame (c) By proposed scheme

**Figure 5:** Comparison of segmentation (Highway 186th frame)



(a) By Sagar et. al [1] (b) Original frame (c) By proposed scheme

**Figure 6:** Comparison of segmentation (Highway 479th frame

## 4. Conclusion

The use of localized thresholding technique and use of bilateral filtering retain the edge areas while smoothing the small gray level variations which result in over segmentation. From the experimental results we conclude that the precision and specificity have been increased.

## Acknowledgement

The work presented in this paper is part of the project Ref: SB/FTP/ETA-0192/2014 and is financially supported by the Department of Science and Technology (DST), Government of India, New Delhi under the Fast Track Young Scientist- Engineering Science Scheme. It is gratefully acknowledged.

## References


[1] Sagar Gujjunoori, S. Sai Satyanarayana Reddy, Gouthaman K V. .Tracking and Size Estimation of Objects in Motion Published in: Machine Vision and Information Technology (CMVIT). International Conference on Date of Conference: 17-19 Feb. 2017 Date Added to IEEE Xplore: 16 March 2017 ISBN Information:INSPEC Accession Number: 16757954 DOI: 10.1109/CMVIT.2017.8 Publisher: IEEE.

[2] D. M. Gavrila,The Visual Analysis of Human Movement: A Survey, Computer Vision and Image Understanding, vol. 73, pages: 82-98, 1999.

[3] AbdullaAl-Kaff, David Martn, FernandoGarca, Arturo de la Escalera, JosMara Armingol, Survey of computer vision algorithms and applications for unmanned aerial vehicles, Expert Systems with Applications Volume 92, February 2018, Pages 447-463.

[4] L. P. Yaroslavsky, Digital Picture Processing. An Introduction. Springer Verlag, 1985.

[5] V. Aurich and J.Weule, Non-linear gaussian filters performing edge preserving diffusion, in Proceedings of the DAGM Symposium, pp. 538545, 1995.

[6] S. M. Smith and J. M. Brady, SUSAN A new approach to low level image processing, International Journal of Computer Vision, vol. 23, no. 1,pp. 4578, May 1997.

[7] B.K.P. Horn and B.G. Schunck, Determining optical flow.Artificial Intelligence, vol. 17, pp. 185-203, 1981.

[8] Bruce D. Lucas and Takeo Kanade, An iterative image registration technique with an application to stereo vision. In Proceedings of the 7th international joint conference on Artificial intelligence - (IJCAI'81), Vol. 2. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 674-679.

[9] Gunnar Farneback, Two-frame motion estimation based on polynomial expansion. In Proceedings of the 13th Scandinavian conference on Image analysis (SCIA'03), Josef Bigun and Tomas Gustavsson (Eds.). Springer-Verlag, Berlin, Heidelberg, pp. 363-370.

[10] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P.Ishwar, changedetection.net: A new change detection benchmark dataset, in Proc. IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence, RI, 16-21 Jun.,2012