# Novel Approach of Small and Large Displacement in Optical Flow for Object Detection

## Bipinkumar Rajput[1],Dr. Anand Mankodia[2]

*[1]PG student [2]Asst. Professor, EC U.V. Patel College of Engineering, Ganpat University, Mehsana, India*

***Abstract***: *Optical flow estimation is oldest but still most active research domain in image processing. This method is used in motion estimation process. Here, we have discussed the Optical flow method to detect the moving object in video. We have also discussed object detection technique which is used in object detection for moving object. Thereafter, we have chosen optical flow method of detection for further improvement in object detection task. We have discussed different types of most popular techniques of optical flow and also recent survey paper and identified problems (research gaps) in the existing techniques. We found the most recent benchmark of optical flow MPI Sintel. According to the benchmark small displacement, large displacements, occlusions problem are there in object detection using optical flow method. In this report, we have discussed a novel approach for small and large displacement in optical flow for object detection. Finally, we compared the existing work with our approach using simulation results obtained in MATLAB & Visual Studio platform.*
***Keywords***: *Object detection, Optical Flow, SURF Feature, Patch Matching, Near Neighbourhood Field (NNF), Sub-pixel Refinement.*

## I. Introduction

In computer vision system motion analysis is one of important task. Motion characterizations can be used so many application like we can be track objects, quantifying deformation, retrieving dominant motion, detection abnormal behaviours. Now recent days video tracking is very research topic. In video tracking we have to find out the object in video and track that object through all sequence of frame. So for tracking first of all we have to identify the object in video. And for identifying first of all object detection process has done than after it we classify the object with different feature based techniques like shaped based, motion based, colour based and texture. Than using that feature we track the object with different tracking algorithms.So recent years the lots of research is going on to find the best approach for object detection of moving target and optical flow is the one of those technique. The most low-level characterization is the estimation of each pixel which is known as optical flow. Most high-level motion analysis task employ flow as a fundamental basis upon which more interpretation is build. So after that there is number of methodological innovation have progressively changed the field and improved the performance. After studying the benchmark research papers of Middlebury and MPI Sintel we found some difficulties in terms of motion discontinuities, computational coast, illumination changes, small displacements, large displacements and occlusion. To handle this type of problem in a unique method still remains an open problem for object detection. Object detection techniques are classified in to main three categories as below,

1. Background Subtraction
2. Frame Differencing
3. Optical Flow

### 1.1 Background Subtraction

According to the researcher the background subtraction method is most beneficial as compare to the other method because it is very fast and it doesn't require the large memory. In this method what we can do is we take a difference between current frame and the background model which is taken as assumption. In this differential the method is take a difference of the pixel of assume frame and the current frame. When difference between color is greater than threshold, so the pixel is taken as foreground pixel.Now because of some problem like dynamic background, illumination changes or shadow the performance of background method get affected. Massimo Poccardi et al. [1] have reviewed different type of method on background Subtraction like concurrence of image variation, Eigen background, running Gaussian average, Kernel density estimation, Mixture of Gaussians and temporal median filter.

### 1.2 Frame Differencing

We identify the moving object with this method by take a difference between two consecutive frames. In this technique we take an operator which is known as frame subtraction operator to find output image and it will be dine to subtract the current frame from the previous frame. Problem with this method is to obtain the

complete contour of the object as a result of which morphology operations are general used to obtain better result[2].

### 1.3 Optical Flow

This approach is based on optical flow field. To detect the moving target we have to calculate optical flow field of image or videos. This method is give the full information of moving object and this method is very help full to find the moving objects from the background of image. This method has problem like we have to do large calculation to obtain optical flow information and it is very sensitive to noise. There are several algorithms have been developed so far but still some problem like motion discontinuities, computational coast, illumination changes large displacements and occlusion are there.

## II. Evolution Of Optical Flow

The first benchmark has come at year of 1995 by Barron et al. [3]. They simple do parametric transformation on real images like. Their result is qualified by absence of discontinuities and small displacement. Than after Middlebury benchmarked has write for more challenging sequences and they get smooth information to the sequence explain by Barron et al. [4]. But in this benchmark they added motion discontinuity.Now recent year MPI Sintel benchmark [5] is solved all most raised issues in Middlebury benchmark. They have experiment their result on synthetic movie and they find the new issues like occlusion, illumination changes, large displacement and effects like blur or defocus. In recent time most of researches are try to solve small and large displacement problem and all method are based on feature matching technique. In recent paper like[6] & [7] are work on this problem both paper are use feature based method but in first [6] they use spare matching technique and NNF for matching nearest pixel of image. This paper applies on pyramidal approach. In paper [7] they use edge preserving technique in patch matching and find improved result in optical flow.Jian Yao et al.[8]proposed a technique of coarse-to-fine optimization to tackle the problem of unsupervised segmentation in the form of super pixels and their main aim is on speed and accuracy. In Moritz Menze et al. [9]here the author founded that the sub pixel accuracy is accurate for estimating the optical flow. They found some problem in that so they use the Sub-pixel refinement strategies solved the problem like large displacement in estimating the optical flow. Marius leordeanu et al.[10]have propose the new method which is sparse-t0-dense method. In this they use the technique of sparse matching method. This method is very use full for field estimation and also for the occlusion problem which founds at the time of extracting the image. Here this proposes method is different from the coarse-to-fine method. In this method they start at the higher level and apply the sparse matching algorithm and then they move toward the dense flow. After that they apply the interpolation procedure.

## III. Feature Based Methods in Optical Flow

After reading the survey paper[11], feature based method are best for small and large displacement problem in optical flow. So in feature based method we simply we ignored vectors and focus on some part of images. The method matching of feature between two images are correlation methods. In next we take one feature to find the correspondence in images. We can use two direction features like corners. So in this proposed work SURF feature method is find out the feature of the image to match with the other.

### 3.1 SURF feature extraction

Since the SURF features have scale and rotational invariance, it additionally keeps up some soundness on the enlightenment also, relative change, and contrasted and the SIFT features, the figuring rate of the SURF components is quick.In SURF feature Hessian detection is utilized to acquire the interesting point of the image frame.To enhance the execution, the case channel is utilized to estimated Gauss filter in the SURF calculation, The integral picture is utilized to quicken the convolution calculation speed. The SURF algorithm is not the same as the conventional strategy in light of the sub sampling and repeated Gauss filtering for the structure of the scale pyramid.

$$H\big(f(x,y)\big) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial xy} \\ \frac{\partial^2 f}{\partial xy} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \qquad (3.1)$$

$$H(\bar{x}, \sigma) = \begin{bmatrix} L_{xx}(\bar{x}, \sigma) & L_{xy}(\bar{x}, \sigma) \\ L_{xy}(\bar{x}, \sigma) & L_{yy}(\bar{x}, \sigma) \end{bmatrix} \quad (3.2)$$

Over the location of the interest points is executed in the single scale, and after that 26 neighbourhood estimations of the three dimensional neighbourhood in multi scale space are thought about with the interest point in the SURF algorithm.

### 3.2 Hierarchical Patch matching

In patch matching process, keeping in mind the end goal to abstain from catching into nearby minima, a few arbitrary theories are also strove for every pixel amid the spread. The coordinating expense between two patches is initially characterized as the patch Euclidean separation. In particular, assume two patches with span r are focused at area a (xa,ya) in picture A what's more, area b (xb,yb) in picture B, separately. The coordinating taken a toll between the two patches is

$$d(a,b) = \sum_{\Delta(\Delta x,\Delta y):|\Delta x|\leqslant r,|\Delta y|\leqslant r} ||I^A(a+\Delta) - I^B(b+\Delta)||^2 \quad (3.3)$$

Where $I^A$ and $I^B$ denote the CIElab color appearance of image A and B, respectively.
Than this process is followed to every layered of image frame.The points of interest in info picture can be greatly improved safeguarded in NNF when utilizing our edge-saving rendition.

### 3.3 Occlusion Handling

In the wake of processing bidirectional NNFs (at every determination) between two pictures, we expressly perform forward-backward consistency check between the two NNFs to distinguish blocked districts. Conflicting mapping pixel is then settled by adjacent pixels as per their bilateral weights. Indeed, even thus, there will in any case be some wrong mapping pixels that can't be identified by the consistency check, which we regard as outliers. A weighted median filtering is subsequently performed on the stream fields to evacuate the outliers (filtering is performed on all pixels). A second pass consistency check and settling is at that point performed to ensure the sifting does not present irregularity. Note that the consistency check and settling is typically quick, the computational overhead in this progression is basically the weighted median filtering performed on all pixels.

### 3.4 Sub-pixel Refinement

Assume the discrete correspondence for every pixel an in picture.An is $NN_{A \to B}(a) = b$, and the patch focused at pixel an is indicated by Ωa. We then figure the coordinating costs between patch Ωa and m distinctive patches around patch b, separately, which is signified as D = {d1, d2,....,dm}. Note that when processing the coordinating expense, the quick calculation in past area still applies. Accept the taken a toll takes after a paraboloid surface on the 2D picture network:

$$d = f(x,y) = \theta * [x^2, y^2, xy, x, y, 1]^T \quad (3.4)$$

where θ = [θ1,θ2,...,θ6] are the unknown, value of m is known and we can solve the linear system and found out the unknowns. Than $b^* = (x^*, y^*)$ can be found by the following

$$x^* = \frac{2\theta_2\theta_4 - \theta_3\theta_5}{\theta_3{}^2 - 2\theta_1\theta_2} \text{and} y^* = \frac{2\theta_1\theta_5 - \theta_3\theta_4}{\theta_3{}^2 - 2\theta_1\theta_2} \quad (3.5)$$

Which are the location of a's related with sub-pixel accuracy.to increase the sub—pixel accuracy we calculate matching cost of the m points foe not sample image or frame in the place of original ones. Now this is not increase the calculation time cost here we only need to calculate the matching cost for every pixel in main image or frame. The main difference is that the m points around the pixels of b now with sub-pixel offset to b. Finally, an edge-preserving filtering with small parameters is performed on the flow fields to smooth out small outliers that might be introduced in this step.
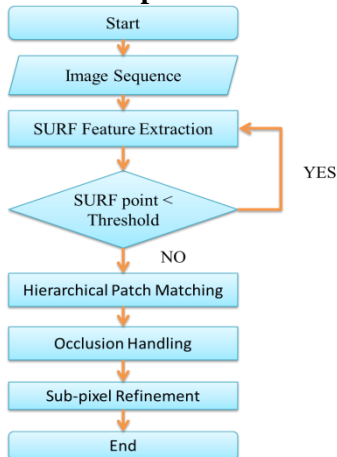
## IV. Proposed Method



**Figure 1.** Flow chart of proposed method

In our proposed work first of all the video is converted into image sequence and take two sequential images I1 and I2 for feature matching process but before that the pyramidal level is made for both the images. First find the SURF feature from the image 1$^{st}$ pyramid level and if the first feature point is less than threshold value than add some more feature point or location. According to the patch idea apply SURF feature extraction method to extract the feature from a square image neighbourhoodcentred on the corner pixel or location from the image 1$^{st}$ pyramid level of I1 and this patch of pixel is fixed in size. This patch is match with the help of Euclidean distance as explain in section 3.2This process is repeated for every level of pyramid. This process is called hierarchical process. As we know that this sparse matching is good for small displacement problem in optical flow. So this SURF feature patch matching process is solved the problem of small displacement in our proposed work.

Than after the next step of our proposed work is to match the patches for every pyramid level of image. This process is useful because in simple feature matching process we missed the pixel that displaced larger distance. Here to find the matches of every patches in 1$^{st}$ level if I1 against I2. The patches are constructed for every pyramid level, and define the patches for every level from the top level to the lowest level of pyramid. Than search the patches on the NNF bases. We define the r as search radius on the top level. This search radius is change according to the level. We have the largest search radius at the lowest level or fine level. In our proposed work we found the 8×8 patches size is good, and this size is used in SURF feature extraction. Furthermore, in hierarchical process forward-backward consistency check have detected the occlusion and remove the outliers. As explained in section 3.3 in this step the pixel are eliminated witch are not moving so result clearly saw two separated object. Than after sub-pixel refinement is use to preserve the edges of object boundary and performed on the flow fields to smooth out small outliers that might be introduced in this step.

## V. Result & Simulation

A. Result on Middlebury after Sub-pixel refinement



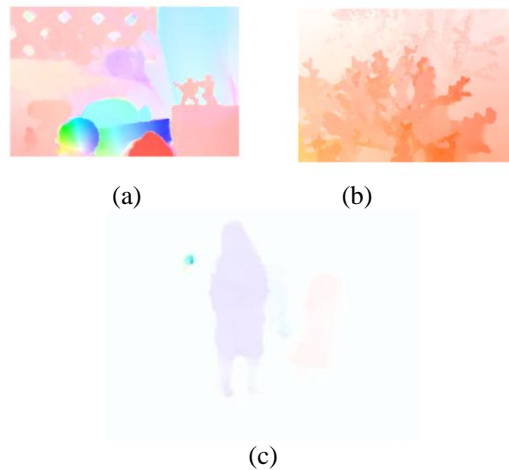(a)                              (b)

(c)

**Figure 2:** (a) Army (b) Grove (c) Backyard

Figure 2 is the result of Middlebury database by using proposed method. These result shows the proposed method is good for small displacement problem. The object in the frame has in good shape which is nearer to the original frame. In other method objects shape is not so clear but in our proposed method the boundary of object is clearly define that give the good knowledge about object shape.

B. Result on MPI-Sintel after occlusion handling



**Figure 3:** Occlusion region after occlusion handling

Above figures shows result of after occlusion handling. In this figure is saw matching process is very good. And we can also see that occlusion problem is solved figure shows two different object separately in figure 3. After patch matching the occlusion handling process eliminate the pixel which are hide behind the first object. In figure the dragon and girls occlusion region is black. So that occlusion handling is good to separate the object from the other object and compare the performance of proposed method with other methods.
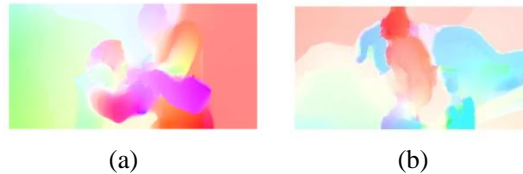
C. Result on MPI-Sintel after Sub-pixel refinement



| (a) | (b) |

**Figure 4:** (a) Non Occlusion region (b) Occlusion region

Figure4 is the result of MTI-Sintel database by using our proposed method. This result got after the process of sub pixel refinement after the occlusion handling process the sub-pixel refinement is refine the pixel around the main pixel and smooth the portion around the pixel. This process fills the occlusion region and gives the fine shape idea about the objects. So now in figure 4 gives the good idea about the both the object for occlusion and non–occlusion region. In figure the object in good shape and separated from other object and background and compare the performance of proposed method with other methods with the help of AE parameter.

D. Parameter comparison

**Table 5:** Angular error (AE)

| Method | AE |
| --- | --- |
| Propose work | 2.95 |
| S2D-Matching | 3.36 |
| MDP-Flow | 3.48 |
| EPPM | 4.25 |
| LDOF | 4.60 |
| DEEP Flow | 4.49 |

**Table 6**: Endpoint error (EPE) for occlusion region&non-occlusion region

| Method | EPE (Occ) | EPE (N.Occ) |
| --- | --- | --- |
| Propose work | 23.189 | 1.358 |
| Flow Field | 25.700 | 1.056 |
| Deep Flow | 34.751 | 1.771 |
| PCA Layer | 32.468 | 2.445 |

Table 5 and 6 shows bar chart comparison. Table5 shows that our method is very good for small displacement by comparing the AE parameter. We got 2.95 AE for proposed method which is less than other method. Table 7 are shows that this proposed method is also good for non-occlusion and occlusion region. In occlusion we reduce the error better than any other method. We can see that flow field give better result but other method give more error for non-occlusion region so this method is better than other method except flow field.After analysis from the bar chart we can say that the proposed method give good performance for occlusion region and non-occlusion region for large displacement and small displacement.

## VI. Conclusion

After study different research paper large displacement, small displacement and occlusion problem occur in optical flow. In propose work we work on these problems and we get the better result from other technique. In our work we use two technique spears flow and patch matching using pyramidal approach. As can be seen from the result obtained, the proposed method reduces the AE and EPE errorwe can say that the proposed method give good performance for occlusion region and non-occlusion region for large displacement and small displacement. But Flow Field which is based on our logic gives less EPE for non-occlusion. So in future we can reduce the EPE for non-occlusion region.In this propose work we solve the problem of both technique like small displacement and large displacement. We compare our results using two parameters for average endpoint error and angular error. After comparison from other method we can say that our proposed work is much better than other method. We used SURF feature in our work and it give the better result from SIFT feature. In future work we can be use different feature extraction technique and enhance the performance of our work or can be using other regularization method to enhance the performance of optical flow.

## Reference

[1]. Piccardi, Massimo. "Background subtraction techniques: a review." In*Systems, man and cybernetics, 2004 IEEE international conference on*, vol. 4, pp. 3099-3104.IEEE, 2004.

[2]. Kulchandani, Jaya S., and Kruti J. Dangarwala. "Moving object detection: Review of recent research trends." In *Pervasive Computing (ICPC), 2015 International Conference on*, pp. 1-5. IEEE, 2015.

[3]. Beauchemin, Steven S., and John L. Barron. "The computation of optical flow." *ACM computing surveys (CSUR)* 27, no. 3, pp. 433-466, 1995.

[4]. Brox, Thomas, and Jitendra Malik. "Large displacement optical flow: descriptor matching in variational motion estimation." *IEEE transactions on pattern analysis and machine intelligence* 33, no. 3, pp. 500-513,2011.

[5]. Butler, Daniel J., Jonas Wulff, Garrett B. Stanley, and Michael J. Black. "A naturalistic open source movie for optical flow evaluation."In *European Conference on Computer Vision*, pp. 611-625.Springer Berlin Heidelberg, 2012.

[6]. Timofte, Radu, and Luc Van Gool. "Sparse flow: Sparse matching for small to large displacement optical flow." In *2015 IEEE Winter Conference on Applications of Computer Vision*, pp. 1100-1106. IEEE, 2015.

[7]. Bao, Linchao, Qingxiong Yang, and Hailin Jin. "Fast edge-preserving patchmatch for large displacement optical flow." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3534-3541. 2014.

[8]. Yao, Jian, Marko Boben, Sanja Fidler, and Raquel Urtasun. "Real-time coarse-to-fine topologically preserving segmentation." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2947-2955. 2015.

[9]. Menze, Moritz, Christian Heipke, and Andreas Geiger. "Discrete optimization for optical flow." In *German Conference on Pattern Recognition*, pp. 16-28. Springer International Publishing, 2015.

[10]. Leordeanu, Marius, Andrei Zanfir, and Cristian Sminchisescu. "Locally affine sparse-to-dense matching for motion and occlusion estimation." In*Proceedings of the IEEE International Conference on Computer Vision*, pp. 1721-1728. 2013.

[11]. Fortun, Denis, Patrick Bouthemy, and Charles Kervrann. "Optical flow modeling and computation: a survey." *Computer Vision and Image Understanding* 134,pp. 1-21, 2015.