# Automatic Occlusion Removal System using Optical Flow Method

M.A.Aneesha[1*], K.J.Helen[2]

[1*]Department of Computer Science and Engineering, Government Engineering College Thrissur, APJ Abdul Kalam Technological University, Thiruvananthapuram, India

[2*]Department of Computer Science and Engineering, Government Engineering College Thrissur, APJ Abdul Kalam Technological University, Thiruvananthapuram, India

e-mail :aneeshagec@gmail.com, helenkj28@gmail.com

*[*]Corresponding Author:* aneeshagec@gmail.com, *Tel.: +91 9747630278*

***Abstract—***

We present an automatic occlusion removal methodology for occluded images. The occlusion here considered are the images which contain elements such as grid or fence, the reflection of objects through glass windows and raindrop. The appearance of any object in the space which blocks the complete view of another object or a scene considers as the occlusion. Because of occlusion, lost the aesthetic beauty of the desired scene. To obtain background scene without any discrepancies occlusion removal is essential. Occlusion may happen accidentally, and also there are some situations we can not avoid occlusion. For example, taking photos or videos in a zoo, fence removal is impossible. If the fence obstruction removes from photos, the results became awesome. The Aim is to improve the accuracy of occlusion removal system using optical flow method. The sequence of frames taken as the input to the system. The system automatically detects occlusion. Then decomposition of background component and occlusion components are done using optical flow method. Finally estimates desired background scene while removing the annoying occlusion. We show results of experiments in the various occluded situation while taking photos or videos. .

***Keywords—*** Occlusion, Optical flow, Detection, Decomposition.

## I. Introduction

Now we are living in a world in which people love photography. They want to capture the photos of every occasion as wells as every moment in their life. They want to keep it as a memory in their life.
Nowadays computational photography techniques are available. New processing techniques are emerging to improve the quality of images and videos. But many photographic conditions are far from the most favorable results. For example, visual occlusions forcing us to take photos through the occluded region that we cannot avoid while we are capturing visuals. Occlusion is a problem in the field of photography.

The occlusion appeared in between the camera and desired scene wants to capture. Some real-world situations were forcing us to take photos through the occlusion. Significant occlusion includes some real-world objects, fence, reflection, and also raindrop causes occlusion during the raining season.

Usually taking photos through a fence, the fence grids makes the disturbance in the real view of the image. Actually, during the shooting, the removal of fence grid is an impossible task. Especially in a protected area like the zoo with a camera surveillance system. There the animals are in a cage or a particular protecting infrastructure. No other way shoot through an enclosure.

Visual occlusions are often impossible to remove using the camera techniques. Computational techniques not so much improved to remove an occlusion from the image with ease. A robust image processing computational approach is required to solve the occlusion problem. Through this study, we are focusing the occlusion handling algorithms. Aim to implement more accurate automatic occlusion removal technique through which improve the quality of images and videos.

In this paper, we present an automatic occlusion removal algorithm that allows a user to take photos through an occlusion producing the occlusion free image as the output. Our algorithm requires the user input as a video or a sequence of image frames while moving the camera horizontally from one end to another. Occlusion detection and removal processing are fully automatic.

While considering the various occlusion situations like physical occlusion, transparent occlusion required a common accessible experimental setup. In case of reflection or object occlusion, the captured image contains background scene with the overlying layer of occlusion. As a result image pixel have the different depth than the desired background image, we want to capture. Thus instead of taking a single picture, it is required to catch a sequence of images while moving the camera. Our algorithm considers the pixel intensity difference in image sequences and variations in optical flow motions in layers. Our algorithm produces two components: an image of the background scene, and the occlusion.

The occlusion was handling from images in the past under different methods. Here we review related work in those areas. Various methods were proposed to handle occlusion. GPU accelerated method is the obstruction-free photographic technique [3]. Fence removal algorithms[4] and raindrop removal[5,6] algorithms are another important works.

Rest of the paper is organized as follows, Section I contains the introduction of Automatic occlusion removal system using optical flow method, Section II contains the related work of occlusion handling, Section III contains some measures of optical flow method, Section IV contains the architecture and essential steps of Automatic occlusion removal system, section V explains the optical flow methodology with flow chart, Section VI describes results and discussion of our method, Section VII contains the recommendation of performance evaluation and Section VIII concludes research work with future directions.

## II. Related Work

Here we review related work in those areas.

Occlusion removal methods are related to the image and video processing techniques [1]. Video inpainting[Criminisi et al. 2004; Bertalmio et al. 2000; Bertalmio et al. 2001; Newson et al. 2014]. Used to remove an object from an image or a video. Manual marking methods used here. The user marks some occluded region and fills the regions using the information from the neighboring pixels. Image and video inpainting techniques are focused on the object removal. In this case, the system asks the user to specify the occlusion along with the input. This inpainting method is a manual method. Image In-painting this work uses texture synthesis and image inpainting techniques [2] with additional information from the light field to remove occlusions.

Obstruction-Free Photography While this work focuses primarily on synthesizing unknown parts of the image; occlusion detection is also vital. The obstruction-free photography technique introduced by Xue, et al. [3] can be applied to light field data to detect and remove occlusions.

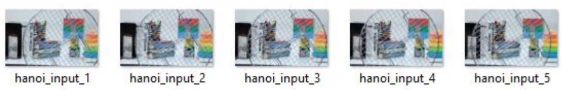
Several other papers proposed to detect and remove partial occlusions from images and also from videos. Several methods proposed to fence occlusion removal [4,5,6]. it involves the regular structure removal. [7] proposed to detect the snow and raindrops in images. It uses frequency space [8] for raindrop removal measures of physical properties is the more significant factor.

The optical flow estimation technique used to determine whether particular image region is occluded or not. [9] used probabilistic formulation methodology to determine occluded regions in the image. Estimated noise model used for this purpose. It generates the histogram of occluded pixel intensities. [10] used level set method along with threshold to determine an occluded region in the image. Above mentioned two methods are trying to minimize the energy function by iteration.

## III. Methodology

The process of our approach is divides into two. First one is the manual approach, and another one is an automatic approach for occlusion removal.

Manual approach: The input given to the system is a sequence of images includes five frames. The data formed by using a horizontal camera moving from one end to the other end. Along with the data input five corresponding occlusion mask given to the system. Occlusion mask is created by the users specifying the object for removal by drawing the thick line over the occlusion region. Figure 1(a) shows an example : Hanoi_input





The five input sequences are hanoi_input_1 to hanoi_input_5. The marked occlusion regions corresponding to the input sequences are hanoi_mask_1 to hanoi_mask_5.


In the resulting image Figure1(b), the occluded region is removed from the input and get an occlusion-free image. The output image sequences stores in a median stack with label median_stack_1 to median_stack_5.These concepts illustrated in Figure1(c).

In our framework, there are three phases involved to achieve the occlusion removal along with the optical flow method: occlusion detection, median filtering, and stack propagation.

The occlusion detection stage computes the motion field for each layer. Along with the motion field, our algorithm uses Edge Flow algorithm to determine the flow between edges in each background and occlusion layer. Canny Edge detector uses to detect edges and corners. From this transformation is calculated to align the results in a focal stack and takes an average. Then calculates the difference in intensity of each pixel with the average value. On the basis of it, the pixels are assigned either into the background layer or occlusion layer.

Median Filtering is the essential step in this occlusion removal method. After loading the input into the focal stack, Compute the median image from the focal stack and set as a reference frame. Lucas Kanade Method is used to approximate the motion between the layers. Based on this focal stack and reference frame generate a transformation to the aligned images. Here in case of a fence or grid occlusion, is the pure translation.

Median Filtering is used to compute edges. It works by moving through the image pixel by pixel, replacing each value with the median value of neighboring pixels. The pattern of neighbors is commonly called window which slides, pixel by pixel over the entire image to the pixel in the image. The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

compute median image

med_filt = nanmedian(stack, number of frames)

There is corresponding Confidence value in between zero and one.

Confidence Ci=Di*Pri, where

Di is the percentage of the known pixel over the total number of pixels. Pri is the similarity of the pixel Pm, Pi each pixel in the set of neighbors N, U and V are pixel along x and y direction and |UV| set of unoccluded pixels.

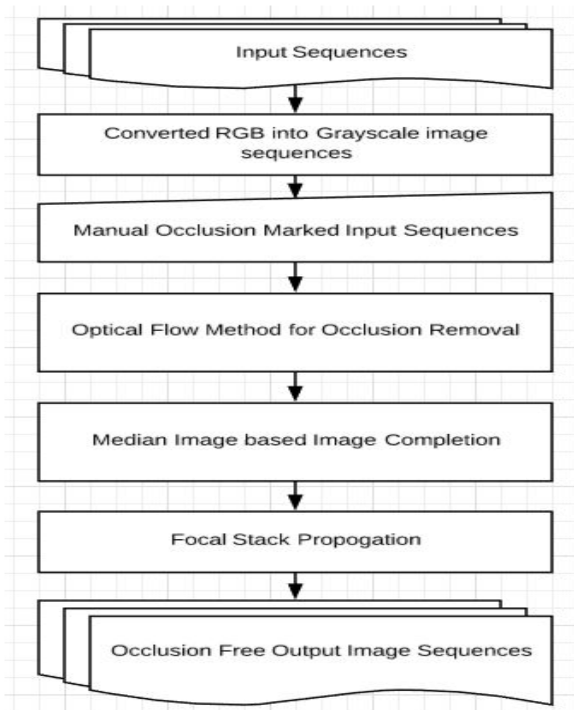
The equation calculates Di,

Di=|UV|/U+V-1 and

$$Pri=\sum_{N} (I(Pi)-I(Pm))\,2 \div |UV|*Imax\,2 \div 2$$

Focal stack propagation performs patch-based synthesis technique. It uses all in focus images to perform the synthesis. Then it checks the depth for each pixel in the image. Finally, the system generates a sequence of an occlusion-free image with more clarity.

## IV. Architectural Design




Optical Flow Method :

Optical flow is the pattern of apparent motion of image objects between two consecutive frames caused by the movement of object or camera. In this case, the camera is moving horizontally from one end to another end then takes a sequence of image frames. The median image set as the reference frame. It is 2D vector field where each vector is a displacement vector showing the movement of points from the first frame to second. In our experiment, we take five consecutive frames as shown in Figure1(a). The arrow indicates its displacement vector. Optical flow works on several assumptions: The pixel intensities of an object do not change between consecutive frames. Neighboring pixels have similar motion.

Consider a pixel I(x,y,t) in first frame. It moves by distance (dx,dy) in next frame taken after dt time. So since those pixels are the same and intensity does not change,

$$I(x,y,t) = I(x+dx, y+dy, t+dt)$$

Then take taylor series approximation of right-hand side, remove common terms and divide by dt to get the following equation:

$$f_x u + f_y v + f_t = 0$$

where:

$$f_x = \frac{\partial f}{\partial x} \ ; \ f_y = \frac{\partial f}{\partial y}$$

$$u = \frac{dx}{dt} \ ; \ v = \frac{dy}{dt}$$

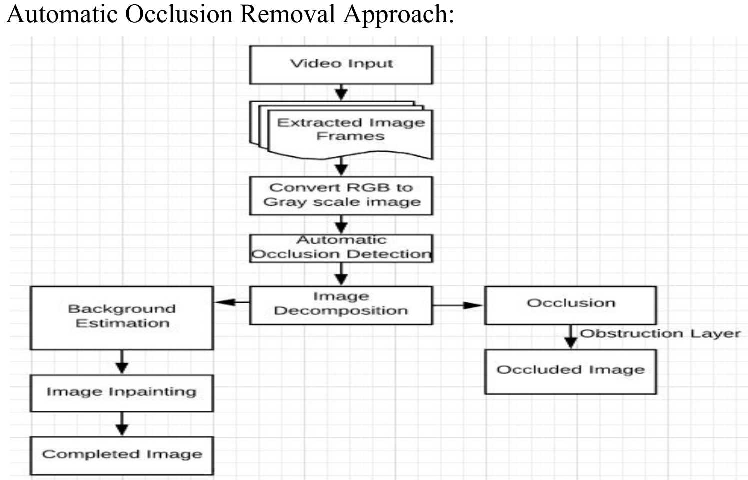
Above equation is called Optical Flow equation. Where fx and fy, are image gradients. Similarly, ft is the gradient along time. But (u,v) is unknown. We cannot solve this one equation with two unknown variables. So Lucas-Kanade method is used to solve this problem.

Lucas-Kanade method: Lucas-Kanade method takes a 3x3 patch around the point. So all the 9 points have the same motion. Then find (fx,fy,ft) for these 9 points. So solving nine equations with two unknown variables which are over-determined. Below is the final solution which is two equation-two unknown problem and solves to get the solution.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i f_{x_i}{}^2 & \sum_i f_{x_i} f_{y_i} \\ \sum_i f_{x_i} f_{y_i} & \sum_i f_{y_i}{}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i f_{x_i} f_{t_i} \\ -\sum_i f_{y_i} f_{t_i} \end{bmatrix}$$

## V. Flow Diagram

Automatic Occlusion Removal Approach:



Consider window reflections as occlusion. When imaging through a window or capture a video of an outdoor scene through a window, the captured image I consist of two layers. The desired background layer IB, and the unwanted obstruction layer IO. And the task is to decompose the input image into these two layers so that we can remove the obstruction from the input. If the system receives a video input, it automatically extracts sequence video frames then turns the color image into grayscale. Extract the reference frame in such a way that a frame in the middle, to have a shorter motion vector is more precise. Then perform the remaining steps as shown in the block diagram. The occlusion detection, Image Decomposition steps are performed by the following algorithm. Image Inpainting and image completion steps are used to fill the unknown regions of the image. Median image-based technique is used to perform inpainting, and patch-based texture synthesis is used to perform image completion.

Algorithm:
Data: Video frames{It}t
Initial guess: IO, IB, A, {VO}t, {VB}t
Result: IO, IB, A, {VO}t, {VB}t

Step1: Set median image as the reference frame.
Step2: Canny edge detection on the reference frame.
Step3: initialize the background,
    I_Background(x,y)=255;
Step4: Extract every frame and apply edge detection.
Step5: Determine Optical flow by Lucas Kanade
    5.1: Find corners
    5.2: Reduce the size of the image.
    5.3: Discard the corners near the margin.
    5.4: Lucas Kanade Method
        5.4.1: Ix_m=conv2(I1_edge,[-1,1;-1,1]);
            Partial on x
        5.4.2: Iy_m=conv2(I1_edge,[-1,-1;1,1]);
            Partial on y
        5.4.3: It_m=conv2(I1_edge,one(2),'valid')+
                conv2(I2_edge,- one(2),'valid');
        5.4.4: u=zeros(length(c),1)
        5.4.5: v=zeros(length(c),1)
    5.5: Calculate Affine transform A=[Ix,Iy]
Step6: Visualize Optical flow vectors, I2_edge
Step7: Find Minimum Motion vector using Euclidean distance,
Sqrt((transition(i,1))2+(transition(i,2))2);
Vmin=[transition(row,1)transition(row,2)];
Vobstruction=[transition(row_max,1)transition(row_max,2)];
Step8: Extract Background

I_background(a,b)=min(I_background(a,b),I2_translate(a,b));
I_obstruction(a,b)=I1(a,b)-I_background(a,b);
if I_obstruction(a,b)< Threshold
    I_obstruction(a,b)=0; Background Layer
else
    I_obstruction(a,b)=1; Occlusion Layer
Step9: End

## VI. Results and Discussion

We tested our algorithm under various situations with occlusion appears. It worked successfully and resulted obtained. The algorithm generates a clean separation of occlusion layer and background layers. Tested our algorithm with the images of occlusions as fence, reflection, raindrops and moving object appeared as occlusion. In all cases, the algorithm produces the proper reconstruction of background scene with the occluded content removed.

Quantitative Evaluation: To evaluate results quantitatively by calculating the normalized cross-correlation (NCC) of recovered decomposition with ground truth decomposition. The following equation calculates the NCC of two images. Let original_image is the ground truth image, output_image is the result of our algorithm. Normalized Cross-correlation, NCC = sum(sum (original_image*output_image)) ÷ sum(sum (original_image*original_image) ).

The NCC of our recovered occlusion free background images with the ground truth backgrounds were Manual approach occlusion removal Hanoi_input, NCC=1.005269, and Automatic occlusion removal using optical flow method for reflection removal from a video input, background scene NCC=0.9975631.

Table 1. NCC Value of Occlusion Removal Methods

| **Occlusion Removal Method** | **NCC Value for Recovered Background** | |
|---|---|---|
| | ***Fence Removal*** | ***Reflection Removal*** |
| [Li and Brown 2013] | 0.9271 | 0.7906 |
| [Guo et al.2014] | 0.9682 | 0.7701 |
| [Xue, Liu and W. T Freeman.2015] | 0.9738 | 0.8985 |
| Our Method | 1.005269 | 0.9975631 |

Table 2. Optical flow method Fence Occlusion NCC Value

| **Fence Occlusion NCC with Ground truth image** | | |
|---|---|---|
| ***Input Sequence*** | ***Algorithm Result*** | ***NCC Value*** |
| Hanoi_input_1.png | Recovered Background_1 | 0.9270291 |
| Hanoi_input_2.png | Recovered Background_2 | 0.9494236 |
| Hanoi_input_3.png | Recovered Background_3 | 0.9526250 |
| Hanoi_input_4.png | Recovered Background_4 | 0.9596487 |
| Hanoi_input_5.png | Recovered Background_5 | 1.005269 |

Table 3. Optical flow method Reflection NCC Value

| **Reflection Removal NCC with Ground truth image** | | | |
|---|---|---|---|
| ***Estimated Background*** | | ***Estimated Occlusion*** | |
| ***Input Video*** | ***Result*** | ***NCC Value*** | ***NCC Value*** |
| Frame1 | Bg_1 | 0.8998939 | 0.3846623 |
| Frame2 | Bg_2 | 0.9432619 | 0.4099005 |
| Frame3 | Bg_3 | 0.9555014 | 0.5081293 |
| Frame4 | Bg_4 | 0.9930570 | 0.6731001 |
| Frame5 | Bg_5 | 0.9975631 | 0.7521180 |

## VII. Conclusion and Future Scope

In this paper, we have demonstrated occlusion removal using optical flow method. Our algorithm increases the quality of photos, and us obtain a background scene with more quality. Our algorithm performs with the sequence of images. Handling occlusion removal with the single photograph is a challenging task. Using image sequences our algorithm produces the better result.

## References


[1] Criminisi, Antonio,Patrik Perez, and Kentaro Toyama, "*Region filling and object removal by exemplar-based image inpainting*", IEEE Transaction on image processing 13.9, pp.**1200-1212**, **2004**

[2] Kao, Shannon, "*Light Field Occlusion Removal*", Sanford University, **2005**

[3] Tianfan Xue, Michael Rubinstein, Ce Liu, William T Freeman , "*PA Computational Approach for Obstruction -Free Photography*", ACM Transactions on Graphics,Vol. 34,No. 4, Article 79, **2015**.

[4] Park, Minwooi et al. "*Image de-fencing*", Asian Conference on Computer Vision. Springer, Berlin, Heidelberg, **2010**

[5] Liuy, Yanxi, et al. "*Image de-fencing*", Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on **IEEE, 2008**

[6] Jonna, Sankaraganesh, et al. "*A multimodal approach for image de-fencing and depth inpainting*", Advances in Pattern Recognition (ICAPR), 2015 Eighth International Conference on. IEEE, **2015**

[7] Yamashita, Atsushi, Akiyoshi Matsui, and Toru Kaneko. "*Fence Removal from multi-focus images.*", Pattern Recognition (ICPR), 2010 20th International Conference on IEEE, **2010**

[8] Garg, Kshitiz, and Shree K. Nayar., "*Detection and removal of rain from videos*", Computer Vision and Pattern Recognition, 2004. CVPR 2004, Proceedings of the 2004, IEEE Computer Society Conference on.Vol.1. IEEE, **2004**

[9] Strecha, Christoph, Rik Fransens, and Luc Van Gool., "*Wide-baseline stereo from multiple views: a probabilistic account.*", Computer Vision and Pattern Recognition, 2004. CVPR 2004, Proceedings of the 2004, IEEE Computer Society Conference on.Vol. 1. IEEE, **2004**

[10] Luan, Xiao, et al., "*Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion.*", Pattern Recognition, 47.2 PP. 495-508. **2004**

Authors Profile

Ms. M A Aneesha pursed Bachelor of Technology from the University of Calicut, 2014 she is currently pursuing M.Tech Computer Science and Engineering, Department of Computer science and Engineering, Government Engineering College, Thrissur, under University of APJ Abdul Kalam Kerala Technological University, Thiruvananthapuram, Kerala 2016-2018. Area of interest is Image Processing and Computer Vision.

Mrs. K J Helen, pursed B. Tech (Computer Engineering), CUSAT, M.Tech (Computer Science), NIT Thiruchirapally. She is currently working as Associate Professor in Department of Computer Science and Engg.
Govt. Engineering College, Thrissur under University of APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala. She has 18 years of teaching experience and area of interest are Image Processing and Computer Networks.