

# Efficient Motion Estimation and Detection, Background Subtraction, Shadow Removal and Occlusion Detection.

**Prof. Pravin R. Lakhe**

Department of EXTC Engg.  
Prof. Ram Meghe Institute of tech.  
and research, Badnera, Amravati

**Prof. Dr. V. U. Kale**

Department of EXTC Engg.  
Prof. Ram Meghe Institute of tech.  
and research, Badnera, Amravati

## Abstract

*The main objective of this project is to develop multiple human object tracking approach based on motion estimation and detection, background subtraction, shadow removal and occlusion detection. A reference frame is initially used and considered as background information. While a new object enters into the frame, the foreground information and background information are identified using the reference frame as background model. Most of the times, the shadow of the background information is merged with the foreground object and makes the tracking process a complex one. The algorithm involves modeling of the desired background as a reference model for later used in background subtraction to produce foreground pixel which is deviation of the current frame from the reference frame. In the approach, morphological operations will be used for identifying and removed the shadow. The occlusion is one of the most common events in object tracking and object centroid of each object is used for detecting the occlusion and identifying each object separately. Video sequences will be captured and will be detected with the proposed algorithm.*

**Keywords:** Background modeling and subtraction, human motion detection, object tracking, shadow removal, occlusion.

## 1. Introduction

Object tracking can be defined as the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc. in order to extract useful information. Object tracking in video processing follows the segmentation step and is more or less equivalent to the 'recognition' step in the image processing. Detection of moving objects in video streams is the

first relevant step of information extraction in many computer vision applications, including traffic monitoring, automated remote video surveillance, and people tracking.

The capability of extracting moving objects from a video sequence is a fundamental and crucial problem of many vision systems that include video surveillance [1,2], traffic monitoring [3], human detection and tracking for video conferencing or human-machine interface [4, 5, 6], video editing, among other applications.

In applications using fixed cameras with respect to the static background (e.g. stationary surveillance cameras), a very common approach is to use background subtraction to obtain an initial estimate of moving objects. Basically, background subtraction consists of comparing each new frame with a representation of the scene background: significant differences usually correspond to foreground objects. Ideally, background subtraction should detect real moving objects with high accuracy, limiting false negatives (objects pixels that are not detected) as much as possible; at the same time, it should extract pixels of moving objects with the maximum responsiveness possible, avoiding detection of transient spurious objects, such as cast shadows, static objects, or noise.

In this paper, we present a shadow removal technique which effectively eliminates a human shadow cast from an unknown direction of light source. A multi-cue shadow descriptor is proposed to characterize the distinctive properties of shadows. We employ a 3-stage process to detect then remove shadows. Our algorithm improves the shadow detection accuracy by imposing the spatial constraint between the foreground subregions of human and shadow.

The existence of human shadows is a general problem in tracking and recognizing human activities. Shadows not only distort the color properties of the area being shaded but also complicate the edge structure of the figure as a

whole. There are several factors that together determine the appearance of a shadow, for example, the view point of camera, the angle of incidence, the light intensity, and the number of light sources, etc. Further, under the sun, the dominant orientation of a human shadow changes as a function of time. Therefore, a human tracker becomes more prone to miss the target, and the motion pattern of a single action varies considerably. For simplification, by human shadow we mean a human cast shadow in contrast with a human self shadow

## 2. Desired implication

### 2.1 Object Tracking

Basic steps in object tracking can be listed as:

1. Segmentation
2. Foreground / background extraction
3. Camera modeling
4. Feature extraction and tracking

#### 2.1.1. Segmentation

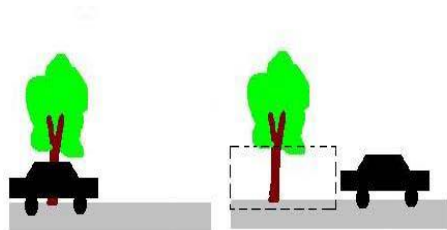
Segmentation is the process of identifying components of the image. Segmentation involves operations such as boundary detection, connected component labeling, thresholding etc. Boundary detection finds out edges in the image. Any differential operator can be used for boundary detection [7,8]. Thresholding is the process of reducing the grey levels in the image. Many algorithms exist for thresholding [7, 8]. Refer [8] for connected component labeling algorithms.

#### 2.1.2 Foreground extraction

As the name suggests this is the process of separating the foreground and background of the image. Here it is assumed that foreground contains the objects of interest.

#### 2.1.3. Background extraction

Once foreground is extracted a simple subtraction operation can be used to extract the background [1]. Following figure illustrates this operation:



Another method that can be used in object tracking is *Background learning*. This approach can be used when fixed cameras are used for video capturing. In

this method, an initial training step is carried out before deploying the system. In the training step the system constantly records the background in order to 'learn' it. Once the training is complete the system has complete (or almost complete) information about the background. Though this step is slightly lengthy, it has a very important advantage. Once we know the background, extracting the foreground is matter of simple image subtraction!

#### 2.1.3. Camera modeling

Camera model is an important aspect of any object-tracking algorithm. All the existing objects tracking systems use a preset camera model. In words camera model is directly derived from the domain knowledge are required to adjust all the inputs. This what is done in [10]. For a moving camera, we need some heuristic about camera motion. If exact information about the camera movement is available then it can be included in the form of transformations. Having multiple moving cameras is very complicated situation (but can be faced with in many real world applications). It needs the algorithm to model motion of all the cameras as well as to integrate results from all the cameras.

#### 2.1.4. Feature

This is an area of image processing that uses algorithms to detect and isolate various desired portions of a digitized image. A feature is a significant piece of information extracted from an image which provides more detailed understanding of the image. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

## 3. Shadow removal technique

### 3.1. Algorithm

The flowchart in figure 1 shows the main algorithm of the project that has been proposed and the resultant image for every process is shown with an example in figure 4. It has been assumed that the input (object's blob and background's blob) is obtained from some background subtraction. All process in figure 1 is explained in the following sections.

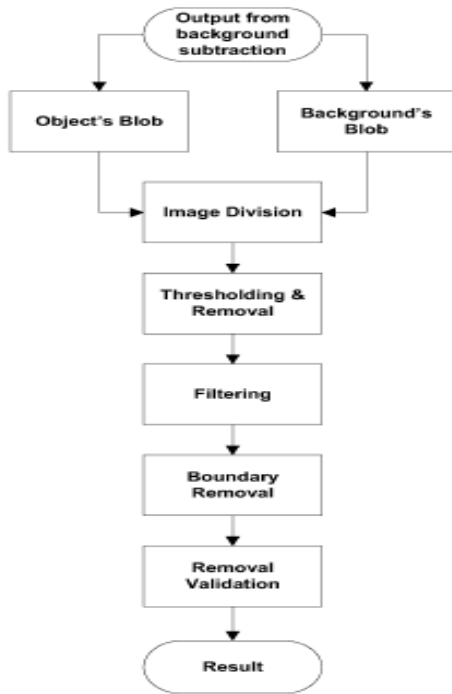


Fig1: Overall algorithm of proposed shadow removal technique

### 3.1.1. Image Division

In this process, the object's blob,  $ob(x, y)$ ,  $\{x, y \in Z^2\}$  is divided with the background's blob,  $bk(x, y)$ ,  $\{x, y \in Z^2\}$ . It has been said before that the purpose of image division is to highlight the homogeneity property of shadows. Resultant image after the division process is multiplied with a constant for the purpose of increasing the signal of the resultant image. In this case, the constant value is 100 (Eq.1). The result of this process is define as  $Img\_Div(x, y)$ .

$$Img\_Div(x, y) = \frac{ob(x, y)}{bk(x, y)} \times 100, \forall x \in X, \forall y \in Y \quad (1)$$

### 3.1.2. Thresholding

The purpose of thresholding is to decide the shadow's blob in the resultant image after the image division process ( $Img\_Div$ ). In this proposed technique, the range has been set according to the scene (Eq.2) and this is done by studying the histogram of the division image over a few samples.

$$Img\_Th = \begin{cases} 1, & t\_min \leq Img\_Div \leq t\_max \\ 0, & otherwise \end{cases}, \forall(x, y) \quad (2)$$

### 3.1.3 Filtering

The purpose of filtering is to enhance the resultant image after the thresholding process ( $Img\_Th$ ) and to

find the biggest blob which is predicted as the shadow's blob or shadow region. Filtering process include filling, erosion and dilation to enhance the image and labeling to predict the shadow. It is assumed that the biggest blob after the labeling process or connected component process as a shadow region.

### 3.1.4 Boundary Removal

The purpose is to remove the penumbra region of shadow, or in other words to remove the shadow's boundary. The first step in this process is to get the coordinates of the boundary (object's blob). This is also called as boundary tracing process. After that, each boundary pixel and its neighbor is checked whether it is a shadow pixel or not. In this case, neighbor pixels that are only located in the horizontal, vertical, and diagonal (45 and -45 degree) of the boundary pixel with certain offset (range between neighbor pixels and boundary pixel) are checked.

### 3.1.5 Removal Validation

Removal validation process consists of two sub processes which are the percentage checking process and the Vertical Scan process. The purpose of the Percentage Checking process is to check whether the removal process was correct or not. This is done by checking the percentage of area that has been removed in the removal process (Eq.3), where BR represent the percentage of area that has been removed over the area of whole object's blob. Based on the study and analysis of sample images, the shadow removal is correct if the percentage value is within a range that is dependent on a scene (Eq.4), where RV, percent\_min and percent\_max represent removal validation result, minimum percentage and maximum percentage (the range). This percentage range will be explained later in section IV. If the percentage value does not fall in that range, it is assumed that the removal did not work correctly.

$$BR = \frac{\text{area that has been removed}}{\text{object's blob area}} \times 100 \quad (3)$$

$$RV = \begin{cases} \text{true}, & \text{percent\_min} \leq BR \leq \text{percent\_max} \\ \text{false}, & \text{otherwise} \end{cases} \quad (4)$$

The second sub process is the Vertical Scan process which will check which part of the object's blob is predicted as a shadow region. In the Filtering process, it is assumed that the biggest blob after that labeling process is the shadow's region. However based on the study of input samples, sometimes, the

second biggest blob is the correct shadow region and the biggest blob is not a shadow region.

Figure 3 shows an example where the Filtering process has done a wrong prediction and based on the analysis this biggest blob (wrong predicted shadow's region) is always located at the center of the object's blob.

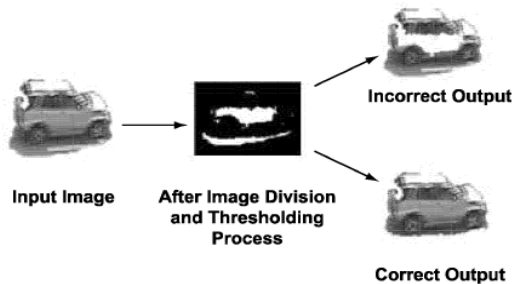


Figure 3. Vertical scan to determine correct removal

so in the vertical scan process, a vertical scan is performed through the centroid of the object's blob just to make sure that the predicted shadow's region is not located at the center of object's blob. However, the Vertical Scan process can only be applied on certain scenes. Some scenes are not suitable because it will only cause a poorer result.

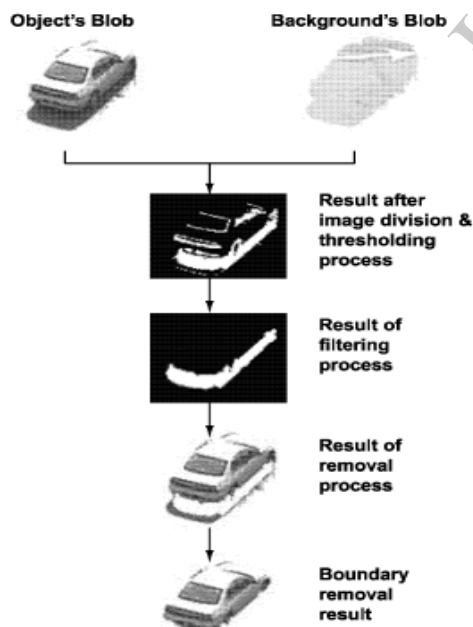


Fig 4. Shadow removal process

In each scene, the vehicle is monitored and analyzed for a period of time and the overall success rate is

calculated. This is calculated by the percentage of result from every video sample,  $PV_i$   $\{i= 1..N\}$  (by getting the number of frames that have the correct result over the number of frames) and then, to get the average percentage of correct result from video samples in the same scene (PSS). Eq. 5 and Eq. 6 are the formulas that are applied in this analysis where the PV, and PSS represent the percentage of correct result from a video samples, number of video samples in a scene and percentage of correct removal in a scene.

$$PV = \frac{\text{no. of frames with correct result}}{\text{no. of frames in video sample}} \quad (5)$$

$$PSS = \frac{\sum_{i=1}^N PV_i}{N} \quad (6)$$

#### 4. Proposed work and objectives

The main objective of this project is to develop an algorithm that can detect human motion at certain distance for object tracking applications. Various tasks are carried out such as motion detection, background modeling and subtraction, foreground detection, shadow detection and removal, morphological operations and identifying occlusion.

#### 5. Conclusion

In this paper, an approach capable of detecting motion and extracting object information which involves human as object will be described. The algorithm involves modeling of the desired background as a reference model for later used in background subtraction to produce foreground pixels which is the deviation of the current frame from the reference frame. The deviation which represents the moving object within the analyzed frame is further processed to localize and extracts the information. We present an effective technique to remove human shadows. Our method has led to accurate recognition of activities.

#### 6. References

- [1] I. Haritaoglu, D. Harwood, and L.S.Davis. W4: Who? when? where? what? a real-time system for detecting and tracking people. In *Proc. the third IEEE Int'l Conf. Automatic Face and Gesture Recognition (Nara, Japan)*, pages 222–227. IEEE Computer Society Press, Los Alamitos, Calif., 1998.
- [2] P.L. Rosin. Thresholding for change detection. In *Proc. IEEE Int'l Conf. on Computer Vision*, 1998.
- [3] N. Friedman and S. Russell. Image segmentation in video sequences: A probabilistic approach. In *Proc. 13th Conf. Uncertainty in Artificial Intelligence*. Morgan Kaufmann, 1997.

- [4] C.R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):780–785, July 1998.
- [5] J. Ohya and et al. Virtual metamorphosis. *IEEE Multimedia*, 6(2):29–39, 1999.
- [6] J. Davis and A. Bobick. The representation and recognition of action using temporal templates. In *Proc. The Computer Vision and Pattern Recognition*, 1997.
- [7] R. Gonzalez and R. Woods – Digital image processing – second edition – Prentice Hall.
- [8] L. G. Shapiro and R. M. Haralick – Computer and robot vision – volume 1 and 2 – 1<sup>st</sup> edition – Prentice Hall.
- [9] C. Ridder, O. Munkelt, and H. Kirchner Adaptive Background Estimation and Foreground Detection using Kalman-Filtering - *Proceedings of International Conference on recent Advances in Mechatronics (ICRAM)*, pp. 193-199, 1995.
- [10] A. Turolla, L. Marchesotti and C.S.Regazzoni – Multicamera object tracking in video surveillance applications.
- [11] Mohamed Ibrahim M, Anupama R, "Scene Adaptive Shadow Detection Algorithm." *Transactions on Engineering, Computing and Technology*, V2, December, pp 88-91, 2004.
- [12] Alessandro Bevilacqua, "Effective Shadow Detection in Traffic Monitoring Applications." *Journal of WSCG*, VII, pp 57-64, 2003.
- [13] A. Bevilacqua, M. Roffilli, "Robust denoising and moving shadows detection in traffic scenes." *Proc. Of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai Marriott, Hawaii*, pp 1-4, 2001. [14] Falah E. Alsaqre, Yuan Baozong, "Moving Shadows Detection In Video Sequences." *Proc. of 7th ICSP 04*, V2, pp 1306-1309, 2004.
- [15] R. Cucchiara, C. Grana, M. Piccardi, A. Prati: "Detecting moving objects, ghosts, and shadows in video streams" *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 25(10), 1337 - 1342, 2003. [27]
- [16] Cran H.D. and Steele, C.M.(1968), "Translation-tolerant Mask Matching using Noncoherent Optics" *Pattern Recognition*, Vol. 1, No. 2, pp. 129-136.
- [17] Comaniciu, D., Ramesh, V. and Meer, P. (2000) Real-time Tracking of Non-rigid Objects using Mean shift", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00)*, Vol.2, pp.142-149.
- [18] Heisele, B. (2000) "Motion based Object Detection and Tracking in Color Image Sequence" 4<sup>th</sup> Asian Conference on Computer Vision.