

**International Journal of Innovative Research in Science,  
Engineering and Technology**

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Special Issue 6, May 2015**

# **Moving Object Tracking Using Gaussian Kernel and Template Modelling In High Resolution Remote Sensing Imagery**

C.Abourva Sundari<sup>1</sup>, M.Kamesh<sup>2</sup>

P.G. Student, Department of ECE, GRT Institute of Engineering and Technology, Tiruttani, India<sup>1</sup>

Assistant Professor, Department of ECE, GRT Institute of Engineering and Technology, Tiruttani, India<sup>2</sup>

**ABSTRACT:** Object Tracking is used to associate target objects in consecutive video frames. The association can be especially difficult when the objects are moving fast relative to the frame rate. This paper deals with object tracking from videos based on template matching using gaussian kernel with probability distribution function and mean shift algorithm. Initially, the target will be selected from chosen video sequence to track desired object in consecutive frames. The target will be utilized to determine the probability distribution function for similarity measurement between target and current processing frames. Here, gaussian kernel function and its gradient are used to find the PDF for corresponding templates. Similarity between two different images will be measured by weighted sum of gaussian coefficients and PDFs. Mean shift approach used here to shift the starting coordinates of template to find its similar features in consecutive frames to detect the desired object. The dissimilarity between the target model and target candidates will be expressed by a metric derived from Bhattacharyya coefficient. The project simulated results shows that moving object from video will be tracked accurately at different position and shape with help of templates in a considerable amount of time.

**KEYWORDS:** Bhattacharyya coefficient, Gaussian Coefficient, Mean Shift algorithm, Template Modelling

## **I. INTRODUCTION**

Surveillance systems have emerging needs in the field of commercial, law enforcement and military applications. Though surveillance cameras are already prevalent in banks, stores, and parking lots, video data currently is used only "after the fact" as a forensic tool, thus losing its primary benefit as an active, real-time medium. A continuous 24-hour monitoring of surveillance video to alert security officers is necessary to control a burglary in progress, or to a suspicious individual loitering in the parking lot, while there is still time to prevent the crime. Video surveillance technology also measures traffic flow, detect accidents on highways, monitor pedestrian congestion in public spaces, compile consumer demographics in shopping malls and count endangered species.

The tracking of moving objects in videos is an important and challenging task in the field of computer vision and artificial intelligence. This report presents results obtained by implementing state of the art, kernel-based tracking algorithm using mean shift. In this approach, objects of interest are characterized by the probability density functions of their colour or texture features. By masking the distribution with a monotonically decreasing kernel, spatially-smooth

similarity functions is defined and mean shift iterations can use the gradient of this similarity function as an indicator of the direction of target's movement. The similarity is expressed in terms of Bhattacharyya coefficient, which is argued to be much more suitable than many more commonly employed techniques, such as histogram intersection.

To monitor a scene in high-resolution images it is essential that the processing time per frame be as low as possible. Hence it is important that the techniques which are employed are as simple and as efficient as possible. For that reason the well known technique of background subtraction [1, 2] was selected for this application. Background subtraction allows moving objects to be detected by taking the point-by-point absolute difference of the current image and a

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

background image which must be acquired when there are no moving objects in the scene. The background of the scene may change due to lighting conditions or 'stationary' objects being moved (e.g. a gate being opened and then left open). Algorithms based on direct grey-level comparison are not robust enough against sudden lighting changes [2]. It is possible though to overcome these problems by using a dynamic background together with normalized cross correlation to evaluate any changes. The dynamic background is initialized with the first image acquired and is updated if a point changes and remains changed for a number of frames.

## II. METHODOLOGIES

### 2.1 Background Modelling

Background subtraction includes pre processing, background subtraction, post processing. In pre processing video frames captured from a camera are given as input to the background subtractor. Pre processing stages are used for filtration and to change the raw input video into a processable format.

In background subtraction moving object is detected and its extreme sensitivity to dynamic scene changes due to lighting and extraneous events is also detected. The main objective is to obtain the objects that keep the user attention in accordance with a set of pre-defined features, including gray level, motion and shape features. This approach defines a method for the generation of an active attention to monitor dynamic scenes for the purpose of surveillance. The idea is to build the background model in many background variations, i.e., background motion cycles which is seen as trajectories of pixels in time. Based on the background model through a map of motion and stationary patterns, this algorithm can detect motion and selectively update the background model.

In post processing, foreground detection is done in which the pixels with a significant difference to those in background model are noted as foreground. Data validation is used to examine the found objects of interest and to eliminate any false matches. The foreground detection stage is known as a binary classification problem whereby each pixel in an image is assigned a label to the class of foreground or background. Formally, for every pixel  $p$  in image  $I$ , a label  $p_l$  is assigned where  $I \in \{0, 1\}$  where  $0 = \text{background}$  and  $1 = \text{foreground}$ . After this mask is obtained, background pixels are usually set to white or black to allow focus on the foreground object. This indicates that an object desired has been identified.

### 2.2 Frame Separation

An input video (.avi file) is converted into still images for processing it and to detect the moving objects. These frames are converted into images with help of the command 'frame2im' and a name will be created to each image and this process will be continued for all the video frames. This is explained in the figure 1.

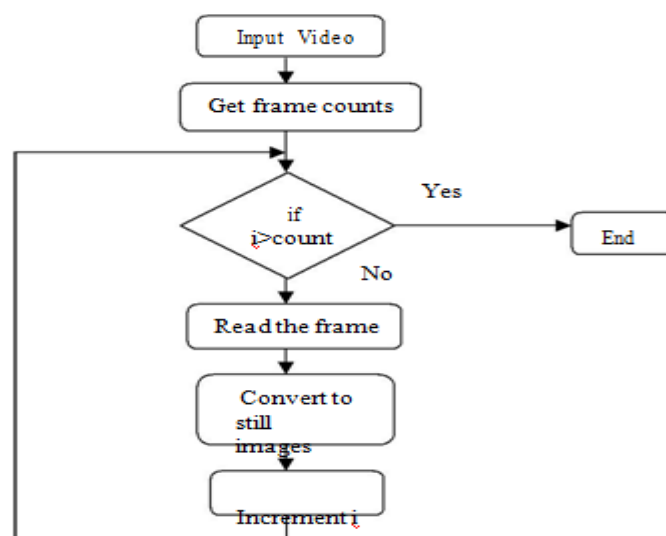


Fig.1 Frame Separation Process

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

## 2.3 Gaussian Kernel and PDF Estimation

Gaussian smoothing is the result of blurring an image by a Gaussian function. It is a widely used in graphics to reduce noise present in an image. Gaussian smoothing is also used as a pre-processing stage in computer vision algorithms in order to enhance image structures at different scales. The idea of Gaussian smoothing is to use this 2-D distribution as a 'point-spread' function, and this is achieved by convolution. The image is stored as a collection of discrete pixels and a discrete approximation to the Gaussian function is produced before the convolution. In theory, the Gaussian distribution is non-zero everywhere, which would require an infinitely large convolution kernel, but in practice it is effectively zero more than about three standard deviations from the mean, hence the kernel can be truncated at this point. The probability density functions with gaussian kernel coefficients [6] are used here to determine the similarity region between target and processing frame. Initially the, the kernel window will be created using gaussian function as follows,

$$\text{Gauss Coefficient} = (1/\sqrt{2*\pi*\text{sig}^2}) (\exp (x^2+y^2/2*\text{sig}^2))$$

Where x and y are input coordinates corresponds to the target and sig is the standard Deviation. Probability function then finds to describe likelihood for this created gaussian random variable corresponds to given input values. This estimated PDF will be normalized by created gaussian kernel to reduce the dynamic changes in input values.

## 2.4 Mean Shift Based Tracking

Tracking of moving objects in video sequences is a very active research area. Objects are tracked based on the texture and edge correlation criteria and the motion model is rigid, affine or homographic. In tracking of people, more complex techniques like covariance-scaled sampling are usually replaced by simpler multi-class statistical color and shape models or adaptive mixtures. In this mean-shift algorithm is used, which is a non-parametric density estimator that optimizes a smooth similarity function to find the direction of the target's motion. The similarity function is obtained by masking the object colour

distributions with Epanechnikov kernel and its smoothness allows the use of a gradient-optimization method. Thus the algorithm focuses on a much smaller neighbourhood and can outperform the exhaustive search. The object's probability density functions are represented with the use of histograms; this is low cost in real-time tracking. The moving object detection in next consecutive frames will be performed by measuring the similarity between template and input data PDF values. The new coordinates belongs to desired target in frames from initial one will be determined based on gradient vector obtained from the gaussian kernel and the coordinates of target dimension will be changed based on this displacement vector, and new PDF will be estimated to similarity between selected targets PDF values.

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

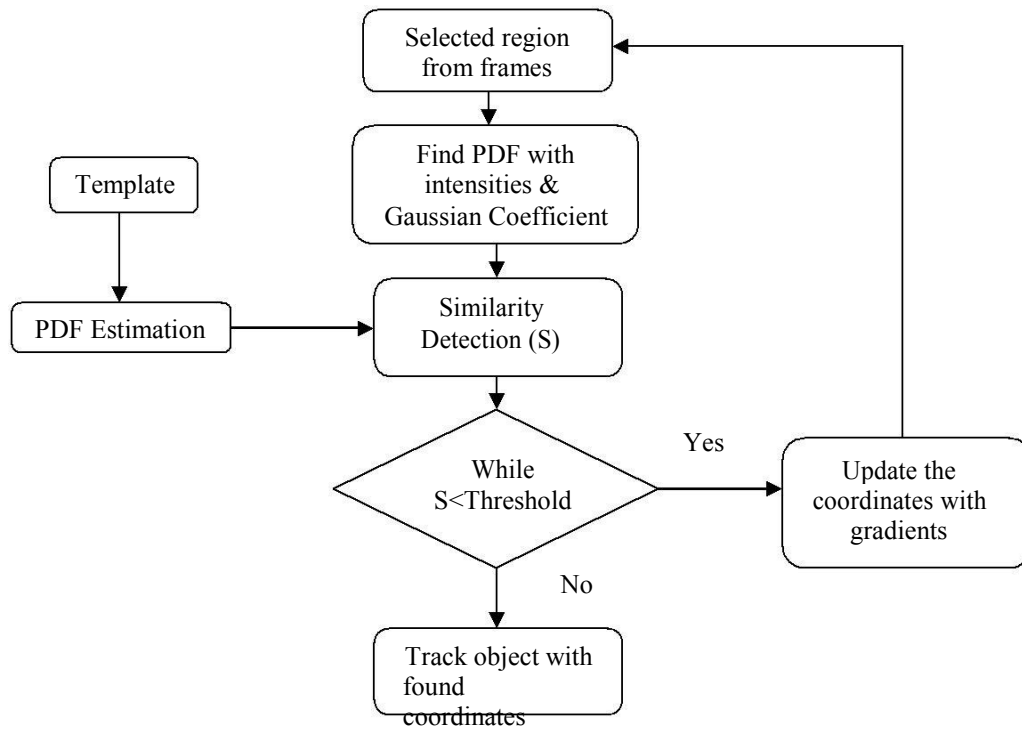


Fig.2 Mean Shift Based Tracking Process

A kernel with Epanechnikov profile was essential for the derivation of the smooth similarity function between the distributions, since its derivative is constant. Thus the kernel masking can lead to a function suitable for gradient optimization, which gives the direction of the target's movement. The search for the matching target in this case is restricted to a smaller area and therefore much faster than the exhaustive search.

Based on the fact that the probability of classification error is directly related to the similarity of the two distributions, the choice of the similarity measure in [1] is such that it is supposed to maximize the Bayes error arising from the comparison of target and candidate PDF. Being a closely related entity to the Bayes error, a Bhattacharyya coefficient was chosen and its maximum searched for to estimate the target localization. Bhattacharyya coefficient of two statistical distributions is defined as

$$\rho [p(y),q] = \int \sqrt{p_z(y)q_z} dz$$

This process is continued to get the desired region or upto loss of target during displacement and the gradient vector. Consider the coordinates  $x$  and  $y$  which represents the origin of selected template and  $k$  gives the gaussian window coefficients. Similarity will be determined by

$$S = (S + w.*k) / (H.*W)$$

Where,  $w$  is  $PDF(T)/PDF(C)$ ,  $H, W$  are Height and Width of template,  $PDF (T),PDF (C)$  are probability distribution of desired template and actual template selected from the frames. The  $X$  coordinates will be updated by the following equation,

$$new\_x = x + (xnew./disvec)$$

The  $Y$  coordinates will be updated by the following equation,

$$new\_y = y + (ynew./disvec)$$

Where  $xnew$  and  $ynew$  determined by,

$$Xnew = Xnew + i*w*Gx$$

$$Ynew = Ynew + j*w*Gy$$

Where  $i$  and  $j$  are row and column indices,  $G$  is Gradient metric. Denominator of displacement vector is given by,  $disvec = disvec + w*norm(Gx, Gy)$

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015

Where, norm is used to find the largest singular values among X and Y derivatives. In order to find the best match of target in the sequential frames, the Bhattacharyya coefficient is maximized.

### III. EXPERIMENTAL RESULTS

The video is given as an input and it is converted into still images using frame separation process. The target will be selected from the first frame of the video as illustrated in figure 3.

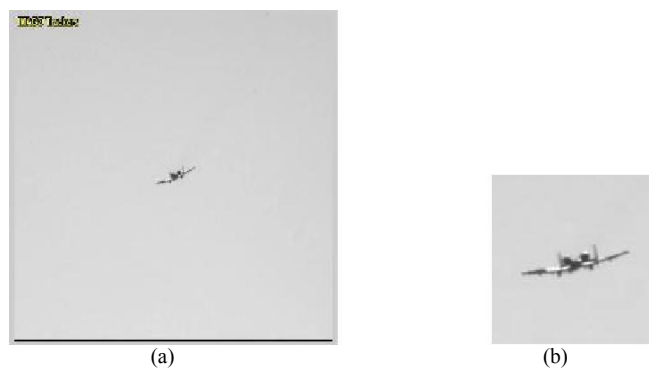


Fig.3 (a) Input Video (b) Selected target image

The input video will be processed into frames. The number of frames of the corresponding input video will be displayed as given in figure 4.



Fig.4 Frames of input video

After the frames being separated, it undergoes gaussian kernel and PDF estimation detection for each and every frame and these are compared with target image values and the object movement in the video is detected. The area of target detection is updated with the help of mean shift algorithm. The video with tracked object is given in figure 5.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Special Issue 6, May 2015



Fig.5 Object Tracking

Any kind of moving object can be tracked with the help of this algorithm and the tracking will be displayed with the help of rectangular box surrounding the desired object.

## IV. CONCLUSION

This paper presents object tracking from videos based on template matching using gaussian kernel with probability distribution function and mean shift algorithm. Gaussian kernel function and its gradient are used here to find the PDF for corresponding templates. Similarity between two different images was measured by weighted sum of gaussian coefficients and PDFs. This was used to update the x and y coordinates of selected template for next frame. Mean shift approach used here to shifting the starting coordinates of template to find its similar features in consecutive frames to detect desired objects. Finally the simulated results shows that used methodologies provided better detection accuracy, suitable for both static and dynamic background videos and less time consumption process.

## REFERENCES

- [1] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," in IEEE Comput Soc. Conf. CVPR, 1999, pp. 246–252
- [2] J. Jacques, C. Jung, and S. Musse, "Background subtraction and shadow detection in grayscale video sequences," in Eighteenth Brazilian Symp, Computer Graphics and Image Processing, Oct. 2005, pp. 189–196.
- [3] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Comput. Surv. vol. 38, Dec. 2006.
- [4] C.Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Realtime tracking of the human body," IEEE Trans. Patt. Anal. Mach. Intell., vol. 19, no. 7, pp. 780–785, Jul. 1997.
- [5] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-time surveillance of people and their activities," IEEE Trans. Patt. Anal. Mach. Intell., vol. 22, no. 8, pp. 809–830, Aug. 2000.
- [6] D. Comaniciu, V. Ramesh, P. Meer, "Kernel-Based Object Tracking", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.25, No. 5, 2003.
- [7] D. Comaniciu, V. Ramesh, P. Meer, "Real-Time Tracking of Non-Rigid Objects using Mean Shift", IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '00), vol.2, 2000.10
- [8] Th. Gevers, "Color in Image Search Engines", Survey on color for image retrieval from Multimedia Search, ed. M. Lew, Springer Verlag, 2001.
- [9] D. Comaniciu and P. Meer, "Mean Shift: A robust approach toward feature space analysis", IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.24, No. 5, 2002
- [10] B. Basclé and R. Deriche, "Region tracking through image sequences", in Proc. of 5th Intl. Conference on Computer Vision, Cambridge, MA, 1995