

Motion Object Detection and Tracking in long Distance Imaging through Turbulent Medium

Ashwini Thakare¹, Swati Nandusekar²

¹Department of EXTC, Pillai's College, Mumbai University, Mumbai, India

²Professor, Department of EXTC, Pillai's College, Mumbai University, Mumbai, India

Abstract: *This paper represents, automatic detection and tracking of moving objects such as people and vehicles through relatively long distances are important computer applications, but still are very challenging tasks, because of the effects of the atmospheric path, which induces turbulence caused movements in the scene and also blurs the video sequence. Such degraded videos may increase the miss detection (false negative) and false detection (false positive) rates. There are numerous methods for detecting and tracking moving objects in both static and non static environments but only a few methods having specific features with background movements caused by atmospheric turbulence which may have unique characteristics straight forward background subtraction method is not directly used in turbulence medium so we proposed method for detecting and tracking moving objects significantly degraded for atmospheric effects.*

Keywords: Atmospheric turbulence, Moving object detection, tracking, adaptive threshold, masking

1. Introduction

Videos are actually sequences of images, each of which called a frame, displayed in fast enough frequency so that human eyes can percept the continuity of its content. It is obvious that all image processing techniques can be applied to individual frames. Besides, the contents of two consecutive frames are usually closely related. Object detection in videos involves verifying the presence of an object in image sequences and possibly locating it precisely for recognition. Object tracking is to monitor an objects spatial and temporal changes during a video sequence, including its presence, position, size, shape, etc. This is done by solving the temporal correspondence problem, the problem of matching the target region in successive frames of a sequence of images taken at closely-spaced time intervals. These two processes are closely related because tracking usually starts with detecting objects, while detecting an object repeatedly in subsequent image sequence is often necessary to help and verify tracking.

Automatic detection and tracking of moving objects such as people and vehicles through relatively long distances (about two kilometers and above) are important computer vision applications, but yet are very challenging tasks. This is due to the effects of the atmospheric path, which induces turbulence caused movements in the scene (temporal clutter) and also blurs the video sequence. Such degradation sources may increase the miss detection (false negative) and false detection (false positive) rates. Because of the long-distance imaging conditions, the objects in the video frames are affected by turbulence-based movements, and are usually small and blurred, and sometimes their real movement is slow. There are numerous methods for detecting and tracking moving objects in both static and nonstatic environments but only a few methods having specific features with background movements caused by atmospheric turbulence which may have unique characteristics straightforward background subtraction method is not well perform in turbulence medium so we proposed method for detecting and tracking moving objects significantly degraded

for atmospheric effects. This method we are applied to the various images of video sequences (Rainy, and windy days)

2. Literature Review

Authors Baldani and D. Cozzi [1] proposed a simple and robust method for tracking moving targets in an outdoor scene which may contain background movements (as snow, hail, and swaying leaves). The motion detection is based on a probabilistic method for background subtraction, while the tracking is based on matching the data found by the motion detection algorithm with those found by block matching. Only after having tracked a moving object for several frames (at least three), is it not considered a false target. The algorithm has been applied to various image sequences in different weather conditions (sunny, rainy, snowy, and windy days)

Authors O. Barnich and M Van [2] proposed a universal background subtraction algorithm, which takes into account the temporal behavior of each pixel, in order to determine whether that pixel belongs to the background, and adapts the model by choosing randomly which values to substitute from the background model (i.e., instead of replacing the oldest values first).

Authors B. Fishbain and L.P. Yaroslavsky proposed a real-time algorithm that compensates for spatio-temporal image movements due to atmospheric turbulence in video sequences, while keeping the real moving objects in the video unharmed. The method is based on (a) generating a background image using a temporal median filter, (b) detecting moving objects using background subtraction with a single (global) threshold, (c) estimating the motion of each detected object (relative to a reference frame) using optical flow estimation (the estimated motion vectors of an object are considered to be turbulence-induced if their magnitudes are low and their direction variations are high), and (d) stabilizing the video by applying a longer temporal median filter to moving objects with higher probability of being turbulence-induced.

3. Proposed System

This paper represents the proposed system for detecting and tracking moving object in degraded videos. In this paper, used three types of video sequences that are rain, fog, and

wind. So this paper proposed adaptive threshold per pixel method for detecting moving object and for tracking used background activity mask method and then reduced the miss detection and false detection rate.

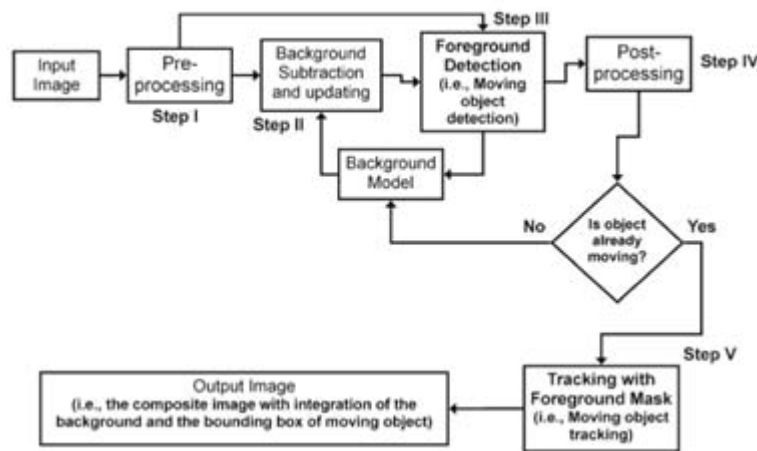


Figure 1: Operation of proposed method

4. Foreground Detection Using a Per-Pixel Adaptive Threshold

Background subtraction of the detection of moving (foreground) regions is based on subtracting each frame of the video from a background model followed by thresholding. For initial background estimation, we have used the common temporal median filter and calculated for each pixel over the first two seconds of the input video. (N=60 frames). A two second time period required, which limits the size and velocity of the moving objects. This period was defined to be close to minimum time required for the case of the regular human walking and common automobile driving. A normal human walking speed is about 5 km/h and the pixel field of view is 0.012 mrad pixel.

Then, starting from frame N + 1, the background image (model) is updated by a weighted average between the current frame and background model,

$$B_k(\vec{x}) = \alpha \cdot B_{k-1}(\vec{x}) + (1 - \alpha) \cdot I_k(\vec{x}),$$

Where \vec{x} is the image coordinator vector of (x,y), $B_k(\vec{x})$ is the kth background image, $I_k(\vec{x})$ is the kth frame and α is the weight (in the range of 0–1). We used a conventional value of 0.97 for α . Now we have starting from detection and tracking procedure using smaller number of frames are possible but it may lead to detect less accurate real moving object in background modelling, because of this increase miss and false detection rate.

The updating is done only for background pixels in the previous frame. To avoid cases of discontinuity in the scene revise the background is performed by reemploying a temporal median for an initial background model estimation. A scene change is detected if a large change in the number of foreground pixels occurs (relative to that number in the previous frame).

Foreground pixel $F_k(x)$ are detected using background subtraction with a per-pixel adaptive threshold

$$F_k(\vec{x}) = \begin{cases} 1 & \text{if } \text{Diff}_k(\vec{x}) > T(\vec{x}), \\ 0 & \text{otherwise} \end{cases}$$

Where $\text{Diff}_k(x)$ is the absolute background subtraction

$$\text{Diff}_k(\vec{x}) = |I_k(\vec{x}) - B_k(\vec{x})|,$$

And $T(x)$ is an adaptive threshold

$$T(\vec{x}) = \text{Gain} \cdot \text{median}_k[\text{Diff}_k(\vec{x})] + \text{Offset}.$$

Offset represents the common global motion detection threshold. Its value depends on the level of the noisiness in the images, which causes differences between images of the same scene. The main aim of equation of $T(x)$ is to increase the threshold values at pixels belonging to dynamic (turbulent) background regions, thus reducing the detection probability there (lower false alarm rate). We are taking 5% value of maximum image intensity for the offset and a multiplier equal to three for the Gain. Normally in background subtraction method use a single (global) threshold for detecting the moving objects in each frame.

The threshold may be adapted to each frame of the video but using a single threshold is still inferior in the case of turbulence-degraded videos since the turbulence random motion properties may change from one pixel to the other. Now after the moving pixels are identified, we use morphological operations of opening and closing with a small (3 × 3 pixels) structure element. Opening generally smoothes the contour of an object and eliminates small objects (smaller than the structure element) considered as noise. Closing tends to eliminate small holes and fills small gaps. Finally, we use connected component labelling (grouping all the neighbouring pixels that are connected to each other) for forming the initial set of detected moving objects (blobs).

5. Tracking Moving Objects

Tracking foreground-detected objects frame by frame in video sequences extracts cohesive temporal information about the objects, and thus can contribute to higher-level processing such as classification or behaviour analysis. Out of various techniques for object tracking, two strategies can be distinguished ; one uses correspondence to match objects between successive frames, and the other performs explicit tracking using a position prediction strategy (such as Kalman filtering) or motion estimation. Thus in this paper used basic tracking process which is based on matching the data found by the foreground detection algorithm with those found by a block-matching motion estimation

The method is based on the following stages:

- 1) For each foreground object, its bounding box (“motion window”) is used.
- 2) For each frame k , starting from $N + 1$ (i.e., after the learning time): for each motion window, block matching within a predefined search radius is performed to find its best match (“matching window”) in the next frame ($k + 1$).
- 3) For each frame (starting from $k + 1$): tracking windows (that define the tracked objects) are created by fusing their motion windows and matching windows according to the overlap between them, giving higher priority to the motion windows.

There are some conditions,

- 1) If the bounding box center of the motion window is contained within the bounding box of one or more matching windows, then the tracking window is equal to the motion window.
- 2) If there is only a motion window without an overlapped matching window, the tracking window is set to equal the motion window.
- 3) If there is only a matching window without an overlapped motion window, then the tracking of the object ceases.
- 4) If one or more motion window centers are included within a matching window, then the tracking window is the bounding box of the motion windows.

The atmospheric blur may reduce the contrast of the moving objects, thus increasing the probability of temporal discontinuities in the detection of the moving object, mainly at nonuniform background conditions. In order to deal with this problems and cases such as temporary object concealments, we added an additional stage. In this stage, when a tracked object disappears, a “grace window” (in time and space) is opened for 10 frames, at which the object (motion window) may appear again and be considered as the same one, i.e., continue the above tracking process disregarding its disappearance, given that its motion characteristics are kept within a small region of possible change.

6. Locating Suspected Turbulence-based Moving Objects Using a Background Activity Mask

The tracked objects from the previous stage may contain false detections caused by the turbulence. Such false objects are usually turbulence-induced moving edge regions in the background. To locate such turbulence-suspected objects, we first define a background activity mask by applying an edge detection operation to the background frame. Then, for each tracked object, we check whether its overlap with the background activity mask is above a third of its bounding box area. the background activity mask is calculated from the background image, which contains local shifts relative to each frame, we calculate the overlap several times, each time with a different shift of the mask (performed within a square search area of 9 pixels around the center of the object’s bounding box), and take the maximum value of the overlap. Such turbulence-suspected objects should satisfy stiffer conditions in the next stage of the algorithm

7. Reduction of Foreground False Alarms Caused by Turbulence Effects

Background movements of static regions caused by the turbulence may somewhat resemble movements of real targets when these objects are relatively small, which is a common situation in long-range imaging. This paper defined the specific characteristics to distinguish false region movements that characterize turbulence motion. A detected moving object is considered false if it has at least one of the following characteristics:

- 1) It cannot be tracked for at least a minimal number of frames. We use two values for this parameter, one for “regular” objects N_{min} and the second, higher (a stiffer condition), $3N_{min}$, for the “suspected as turbulence” regions obtained in the previous min stage. We have found that a value of $N = 10$ was robust for all the video sequences examined.
- 2) It has an average movement extent below a pixel and a half.
- 3) It makes rapid changes in its motion direction. This is done by calculating the following term in the time domain on the larger motion component V of the object (X or Y):

Video sequence	FP	FN	TP	Precision	Recall	F1 Measure
Rain+wind	1.58	5.14	9.67	0.58993	0.9237	0.80731
Fog	1.46	3.02	9.24	0.48793	0.7107	0.70731
Fog	1.30	2.57	8.59	0.44293	0.6319	0.60031
Fog+wind	1.34	4.51	9.39	0.55023	0.8639	0.64231
Rain+wind	1.27	4.43	8.78	0.56513	0.8208	0.67831
Rain+wind	1.32	4.55	9.46	0.48723	0.8199	0.70431

$$\left| \sum_{t=k-N_{min}}^k \text{sign}(V_t) \right|$$

It makes significant rapid changes in its size properties, measured by the standard deviation(STD) of the differences in the width or height dimensions along several frames. Such

type of false alarm occurs usually at large edges of the background, and rapid changes are considered when the STD is three times the average object size.

FP (False positive): detected regions that do not correspond to a moving object,
 FN (False negative): moving objects not detected

8. Results and Implementation

The Detection and Tracking algorithm is implemented on real atmospheric degraded video sequences

The proposed method reduces false alarm rate by applying various measures that discriminate real moving objects from turbulence induced motion

TP (True positive): detected regions that correspond to moving objects

Table 1: Statistical Performance Parameters:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_1 = 2 \frac{Precision \cdot Recall}{Precision + Recall}$$

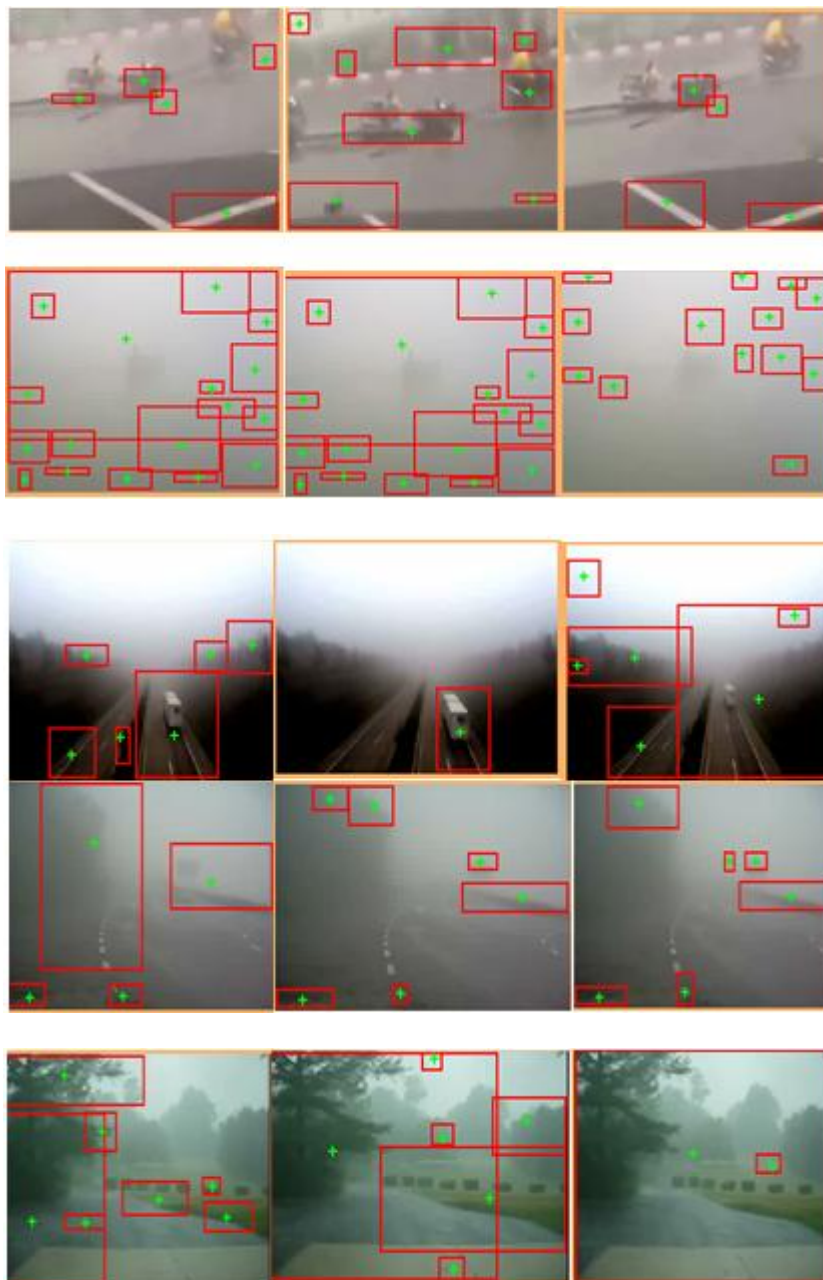


Figure 2: Represents 3 samples of 6 videos with turbulence effect showing detecting and tracking results obtained by proposed method. Each row represents different samples of different videos. To detected objects in each frame are mark by red bounding boxes

References

- [1] G. Baldini, P. Campadelli, D. Cozzi, and R. Lanzarotti, "A simple and robust method for moving target tracking," in Proceedings of IASTED International Conference on Signal Processing, Pattern Recognition and Applications, (ACTA, 2012), pp. 108–112.
- [2] O. Barnich and M. Van Droogenbroeck, "ViBe: a universal background subtraction algorithm for video sequences," IEEE Trans. Image Process. 20, 1709–1724 (2011).
- [3] B. Fishbain, L. P. Yaroslavsky, and I. A. Ideses, "Real time stabilization of long range observation system turbulent video," J. Real-Time Image Proc. 2, 11–22 (2007).
- [4] O. Haik and Y. Yitzhaky, "Effects of image restoration on automatic acquisition of moving objects in thermal video sequences degraded by the atmosphere," Appl. Opt. 46, 8562–8572 (2007).
- [5] O. Oreifej, L. Xin, and M. Shah, "Simultaneous video stabilization and moving object detection in turbulence," IEEE Trans. Pattern Anal. Mach. Intell. 35, 450–462 (2013).
- [6] A. Elkabetz and Y. Yitzhaky, "Background modeling for moving object detection in long-distance imaging through turbulent medium," Appl. Opt. (2014), to be published.
- [7] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in Proceedings of 6th European Conference on Computer Vision, Dublin, Ireland, Vol. 2 (Springer, 2000), pp. 751–767.
- [8] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (IEEE, 1999).
- [9] L. Zhang and Y. Liang, "Motion human detection based on background subtraction," in Second International Workshop on Education Technology and Computer Science (IEEE, 2010).
- [10] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed. (Prentice-Hall, 2008), Chap. 9.

Author Profile

Ashwini Thakare received her Bachelors Degree in Electronics and Telecommunications from Sipna College of Engineering in 2005, Maharashtra, India and currently pursuing Masters from Pillai College of Engineering, Maharashtra, India.