INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN: 2278 – 0211 (Online)

Moving Object Tracking From A Video Sequence Using Bounding Box Method

Prodip Kumar Sarker
Department Of Computer Science & Engineering
Bangabandhu Sheikh Mujibur Rahman Science And Technology University
Gopalganj, Bangladesh
Md. Jamal Uddin
Department Of Computer Science & Engineering
Gono Bishwabidyalay (University), Dhaka, Bangladesh

Abstract:
The object tracking from a video sequence that contains moving objects is a critical task in real world applications. Object tracking from video sequence is the process of locating moving objects in time using a camera. The purpose of object tracking is to determine the position of the object in images continuously and reliably against dynamic scenes. This paper concerned with the tracking and following of moving object in a sequence of frames from a video sequence. After preprocessing of the original video sequence the moving object is tracked with the proposed moving object tracking algorithm in which the moving object region can be extracted completely. Segmentation is performed to detect the object after reducing the noise from that scene. The bases of the work is the block matching object tracking algorithm for a moving target in a video by plotting a rectangular bounding box around it in each frame, and then process the data within that box to separate the tracked object from the background. The paper also includes experimental results of the tracking using the bounding box algorithm with certain improvements to make it suitable for tracking fast moving object.

Key words*Moving object tracking, feature extraction, segmentation, block matching and bounding box*

1.Introduction

The objective of this paper is to track a moving object from a video sequence. Object tracking is one of the most popular areas of video processing because of its applicability to daily problems and ease of production, e.g. surveillance cameras, adaptive traffic lights, plane detection, vehicle navigation, human-computer interaction, object-based video compression, smart rooms, driver assistance, perceptual user interface, augmented reality etc.[15]. The moving object tracking from a video sequence is an important field of research as security has become a prime concern for every organization and individual nowadays. Surveillance systems have long been in use to monitor security sensitive areas such as banks, department stores, highways, public places and borders. Also in commercial sectors, surveillance systems are used to ensure the safety and security of the employees, visitors, premises and assets. Most of such systems use different techniques for moving object detection[8]. Object tracking is a process to track an object and to take an action on another object with no relationship to the tracked objects, based on changes to the properties of the object being tracked and the moving object tracking is the process of detecting moving objects. Object tracking is important because it enables several important applications such as: Security and surveillance - to recognize people, to provide better sense of security using visual information; Medical therapy - to improve the quality of life for physical therapy patients and disabled people; Retail space instrumentation - to analyze shopping behavior of customers, to enhance building and environment design; Video abstraction - to obtain automatic annotation of videos, to generate object-based summaries; Traffic management - to analyze flow, to detect accidents; Video editing - to eliminate cumbersome human-operator interaction, to design futuristic video effects; Interactive games-to provide natural ways of interaction with intelligent systems such as weightless remote control. The main difficulty in video tracking is to associate target locations in consecutive video frames, especially when the objects are moving fast relative to the frame rate [4]. The proposed method is a low complexity solution to moving object tracking. The authors assume a stationary video camera. As there is no movement of the camera so the background it captures is fixed with respect to the camera frame. The approach to be used depends on the context in which the tracking is performed and the end use for which the tracking information is being sought. The major steps of this method are frame extraction, convert frames to image sequence; reduce noise, boundary tracking, frame differencing, detection of moving object, block matching, bounding box and finally tracked moving object.
The overall process of moving object tracking from a video sequence is summarized below:

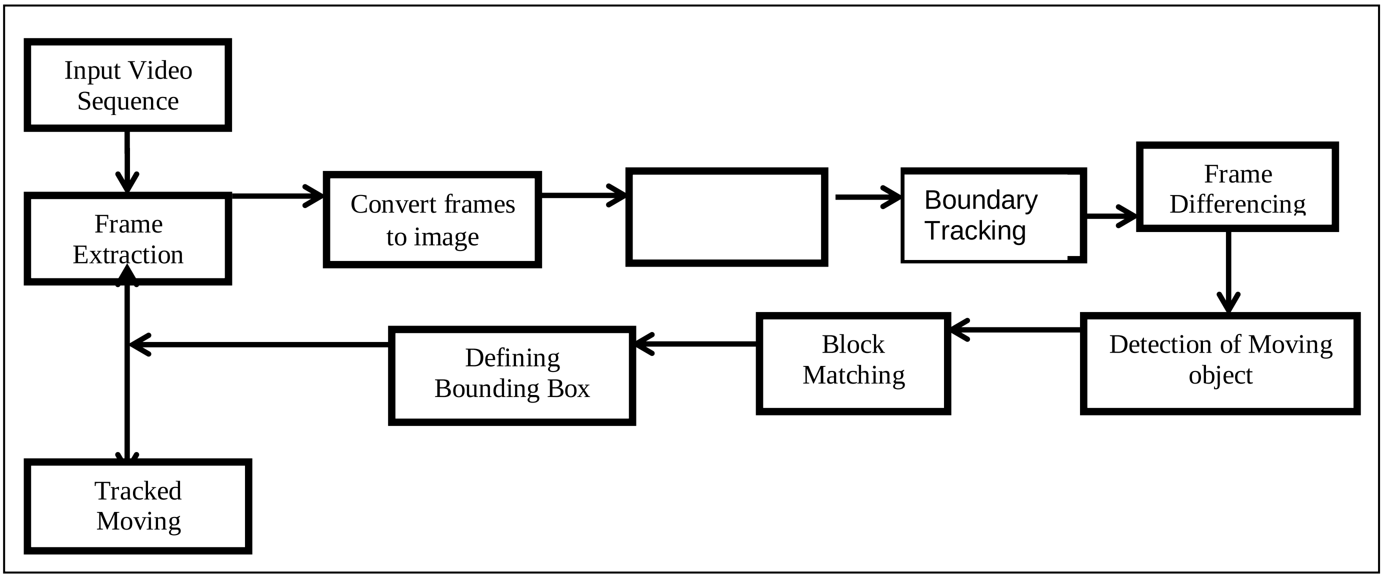

Figure 1: The Overall Process Of Moving Object Tracking From A Video Sequence.

2.Review Of Previous Approaches

There are many approaches for object tracking have been developed for monitoring public areas such as offices, shopping malls or traffic highways.

Tracking based on a moving object region. This method identifies and tracks a blob token or a bounding box, which are calculated for connected components of moving objects in 2D space. The method relies on properties of these blobs such as size, color, shape, velocity, or centroid. A benefit of this method is that it time efficient, and it works well for small numbers of moving objects. Its shortcoming is that problems of occlusion cannot be solved properly in “dense” situations. Grouped regions will form a combined blob and cause tracking errors. For example, [11] presents a method for blob tracking. Kalman filters are used to estimate pedestrian parameters. Region splitting and merging are allowed. Partial overlapping and occlusion is corrected by defining a pedestrian model.

Tracking based on a moving object model. Normally model based tracking refers to a 3D model of a moving object. This method defines a parametric 3D geometry of a moving object. It can solve partially the occlusion problem, but it is (very) time consuming, if it relies on detailed geometric object models. It can only ensure high accuracy for a small number of moving objects. For example, [6] solved the partial occlusion problem by considering 3D models. The definition of parameterized vehicle models makes it possible to exploit the a-priori knowledge about the shape of typical objects in traffic scenes. [2].

Besides these, there are also some other approaches on object tracking. [7] Presents a tracking method based on wavelet analysis. A wavelet-based neural network (NN) is used for recognizing a vehicle in extracted moving regions. The wavelet transform is adopted to decompose an image and a particular frequency band is selected for input into the NN for vehicle recognition. Vehicles are tracked by using position coordinates and wavelet feature differences for identifying correspondences between vehicle regions [7]. Paper [3] employs a second order motion model for each object to estimate its location in subsequent frames, and a “cardboard model” is used for a person’s head and hands. Kalman models and Kalman filters are very important tools and often used for tracking moving objects. Kalman filters are typically used to make predictions for the following frame and to locate the position or to identify related parameters of the moving object. For example, [13] implemented an online method for initializing and maintaining sets of Kalman filters. At each frame, they have an available pool of Kalman models and a new available pool of connected components that they could explain. Paper [12] uses an extended Kalman filter for trajectory prediction. It provides an estimate of each object’s position and velocity. But, as pointed out in [5], Kalman filters are only of limited use, because they are based on unimodal Gaussian densities and hence cannot support simultaneous alternative motion hypotheses. So several methods have also been developed to avoid using Kalman filtering. [5] Presents a new stochastic algorithm for robust tracking which is superior to previous Kalman filter based approaches. Bregler [1] presents a probabilistic decomposition of human dynamics to learn and recognize human beings in video sequences. [9] Presents a much simpler method based on a combination of temporal differencing and image template matching which achieves highly satisfactory tracking performance in the presence of partial occlusions and enables good classification. This avoids probabilistic calculations. Low complexity background subtraction using frame difference method uses the video frame at time t-1 as the background model for the frame at time t. This technique is sensitive to noise and variations in illumination, and does not consider local consistency properties of the change mask. This method also fails to segment the non-background objects if they stop moving. Since it uses only a single previous frame, frame differencing may not be able to identify the interior pixels of a large, uniformly-colored moving object. The main function of Meanshift Tracking algorithm is histogram estimation. Since moving objects can be identified by their color histogram. Mean-shift tracking algorithm is an iterative scheme based on comparing the histogram of the original object in the current image frame and histogram of candidate regions in the next image frame. The aim is to maximize the correlation between two histograms. Motion-based multiple object tracking. The problem of motion-based object tracking can be divided into two parts. First part is detecting moving objects in each frame and second part is associating the detections corresponding to the same object over time.

3.Proposed Procedures Used For Implementation

The procedure first executes a video file to convert the selected avi format file video in to frames. This file generates a series of frames of the video file and creates a folder in which these frames are stored for further processing. These frames are then converted from color scale into Gray Scale files. Then median filtering technique is used to reduce noise from these frames and store them [10]. After reducing noise from these frames the absolute difference between the consecutive frames is calculated. Segmentation is performed using frame difference technique. If the segmentation is performed, the residual image is visualized with rectangular bounding box with the dimensions of the object produced from residual image. For a given image, a scan is performed where the intensity value of the image is more than limit (depends on the assigned value, for accurate assign maximum). In this features is extracted by color and here the intensity value describes the color. The pixel values from the first hit of the intensity values from top, bottom, left and right are stored. By using this dimension values a rectangular bounding box or block is plotted within the limits of the values produced. Blocks are formed in a region without overlapping on the other region. Every block in a frame is compared to the corresponding blocks in the sequence of frames and compares the smallest distance of pixels. By using the position values of object in every frame, we can calculate the position of the moving object.

To perform segmentation operation, frame difference is used because it takes less processing time. Frame difference algorithm performs separation of two sequential frames. Algorithm for the segmentation is explained as follows:

- Read the input images
- for(present position=initial position: final position)
  - Difference between the pixel values at present position of two images is calculated
  - Calculate the absolute value
  - Store the difference in new image at same pixel position that is at present position.

Every object has a specific feature which is used to visualize the object and used for tracking. In this paper to extracted specific feature of an object for tracking, bounding box algorithm is used. Algorithm for bounding box is as follows:

- Take the input video sequence
- Extract frames from video sequence
- Convert frame to image sequence
- Read the image difference
- for(present position=initial value: final value) of Y resolution
- for(present position=initial value: final value) of X resolution
  - calculate the sharp change in intensity of image from top and bottom
  - store the values in an array
- Height of the bounding box is = bottom value – top value
- for(present position=initial value: final value) of X resolution
- for(present position=initial value: final value) of Y resolution
  - calculate the sharp change in intensity of image from left and right
  - store the values in an array
- Width of bounding box = right value – left value
- Using the dimensions, draw boundary to the image
  Initial value: the starting position of the pixel in an image. Final value: the ending the starting position of the pixel in an image
  - $\text{Height} = \frac{\text{bottom.value} - \text{top.value}}{2}$
  - $\text{Width} = \frac{\text{right.value} - \text{left.value}}{2}$
- Add the height value with top value and store it in a variable like mid top
- Add the width value to the left value and store it in a variable like mid left
- Assign the max intensity to the pixel at pixel value at(mid top, mid left)

4.Experimental Results

Our approach is implemented on a PC under Windows operating system using MATLAB 7.0 to execute all the required source code. The technique discussed above is applied on various video files like polygon.avi, inputvideo1.avi, inputvideo2.avi etc. The first code allows the user to select a video file is executed. When the user selects a particular video file then the corresponding frames are generated and stored in a folder. Then the next code is executed that takes the generated frames as input and calculates the absolute difference between each frame and stores it in the same folder that was created in the previous step. As an example, the 1st frame, 2nd frame, 51st Frame and 52nd frame are shown in figure 2(a), 2(b), 2(c) and 2(d) respectively. After calculating the absolute difference between the 1st frame and 2nd frame and 51st frame and 52nd frame are shown in figure 3(a) and 3(b) respectively. After applying the algorithm the moving object is tracked and the result is shown in figure 4. Another example, the 1st and 2nd frame are shown in figure 5(a) and 5(b) respectively. After calculating the absolute difference between the 1st frame and 2nd frame and result are shown in figure 6 and 7 respectively. This process is continuing for every video sequence that contain moving object to tracked moving object.


Figure 2(a): 1st Frame Of "Inputvideo1.Avi"& Figure 2(c): 51st Frame Of "Inputvideo1.Avi"


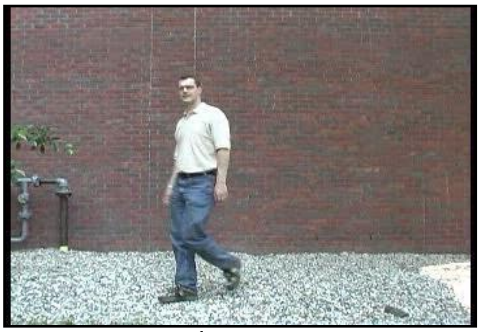

Figure 2(b): 2nd Frame Of "Inputvideo1.Avi" & Figure 2(d): 52nd Frame Of "Inputvideo1.Avi"


Figure 3(a): Absolute Frame Difference Between 1st And 2nd Frames

Figure 3(b): Absolute Frame Difference Between 51st And 52nd Frames


Figure 4: Moving Object Of "Inputvideo1.Avi & Figure 5(a): 1st Frame Of "Inputvideo1.Avi"


Figure 5(b): 2nd Frame Of "Inputvideo1.Avi" & Figure 6: Absolute Frame Difference Between 1st And 2nd Frames

Figure 7: Moving Object Of "Inputvideo2.Avi"

5.Discussion

The main applications of moving object tracking are security, surveillance and vision analysis. In this paper, an attempt has been made to develop a virtual system for moving object tracking from a suitable video sequence. As we know that object detection technique is not completely efficient for all kinds of objects which are available presently all over the world, so this work demonstrated some gateway to overcome those limitations. After all, for test bench for this work, the moving object tracking from a suitable video sequence has been determined at a satisfactory level. Moreover the moving object tracking from a video sequence is a demandable work to find out various types of criminal offence investigation. In this work, the primary works are the video processing as well as image processing for the detection of moving object within the video sequence, but it focuses on the detection of moving object from image in the video sequences detecting the target object based on frame differencing.

6.Conclusion

Moving object tracking is a key task in video monitoring applications. The common problem is occlusion detection. In this case the selection of appropriate features is critical for moving object tracking. We propose a method of block matching to achieve a more accurate but simple moving object tracking system, which can be used in traffic analysis and control applications. In this way we reduced computation time and avoided the common feature grouping problem. This method proved to be easy and efficient, but it only works well on separated regions. So removing noise is an important preprocessing task [14] for the subsequent extraction of moving objects masks, because shadows merge otherwise separated regions.

7.Future Work

In future this method can be modified to differentiate different class objects in real time video also apply 3D analysis, which allows a more detailed classification of cars. The intention is to identify the type of a vehicle. The height value of the car is, for example, easily to extract from the infrared picture.

8.References

1. C. Bregler: Learning and recognizing human dynamics in video sequences. In Proc. IEEE Int. Conf. CVPR'97, pages 568-574, 1997.
2. A. Cavallaro, F. Ziliani, R. Castagno, and T. Ebrahimi: Vehicle extraction based on focus of attention, multi feature segmentation and tracking. In Proc. European signal processing conference EUSIPCO-2000, Tampere, Finland, pages 2161-2164, 2000.
3. I. Haritaoglu, D. Harwood, and L. S. Davis: W4: Who? When? Where? What? A real-time system for detecting and tracking people.In Proc. 3rd Face and Gesture Recognition Conf., pages 222-227, 1998.
4. W. Hu, T. Tan, L. Wang, S. Maybank, "A survey on visual surveillance of object motion and behaviors", IEEE Trans. Syst. Man Cyber.-C 34 (3) (2004) 334–352.
5. M. Isard, and A. Blake: Contour tracking by stochastic propagation of conditional density. In Proc. European Conf. Computer Vision, Cambridge, UK, pages 343-356, 1996.
6. D. Koller, K. Daniilidis, and H. H. Nagel: Model-based object tracking in monocular image sequences of road traffic scenes. Int. Journal Computer Vision, **10**:257–281, 1993.
7. J. B. Kim, C.W. Lee, K. M. Lee, T. S. Yun, and H. H. Kim: Wavelet-based vehicle tracking for automatic traffic surveillance. In proc. IEEE int. Conf. TENCON'01, Aug, Singapore, Vol. 1, pages 313-316, 2001.
8. Soharab Hossain Shaikh and Nabendu Chaki, " A Low Cost Moving Object Detection Method Using Boundary Tracking",Springer- Verlag Berlin Heidelberg 2011, S.S. Al- Majeed et al. (Eds.): WiMoA 2011/ICCSEA 2011, CCIS 154, pp. 127-136, 2011.
9. A. J. Lipton, H. Fujiyoshi, and R. S. Patil: Moving target classification and tracking from real-time video. In Proc.IEEE Workshop Application of Computer Vision, pages 8- 4, 1998.
10. Gonzaleg, R.C. Woods, R.E [1992]. „Digital Image Processing".
11. O. Masoud, and N. P. Papanikolopoulos: "A novel method for tracking and counting pedestrians in real-time using a single camera". IEEE Trans. Vehicular Technology, **50**:1267-1278, 2001.
12. R. Rosales, and S. Sclaroff: "Improved tracking of multiple humans with trajectory prediction and occlusion modeling". In Proc. Workshop on the Interpretation of Visual Motion at CVPR'98, Santa Barbara, CA, pages 228-233, 1998.
13. C. Stauffer and W. E. L. Grimson: "Adaptive background mixture models for real-time tracking". Computer Vision and Pattern Recognition, **2**: 246-252, 1999.
14. W.E.L. Grimson, L. Lee, R. Romano, and C. Stauffer. "Using adaptive tracking to classify and monitor activities in a site". In CVPR98, Pages 22-31, 1998.
15. Madhurima, Madhulika: "Object tracking in a video sequence using Mean-Shift Based Approach: An Implementation using MATLAB7", IJCEM International Journal of Computational Engineering & Management, Vol. 11, January 2011.