

Speed Determination of a Moving object from a Video Sequence using Bounding Box Approach and Region Based Segmentation

Prodip Kumar Sarker

Lecturer

Department of Computer Science & Engineering

Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj, Bangladesh

Md. Jamal Uddin

Lecturer

Department of Computer Science & Engineering

Gono Bishwabidyalay (University), Dhaka, Bangladesh

Abstract-Traveling speed determination of a moving object from a video sequence is a critical task in real world applications. This paper concerned with the determination of the traveling speed of a moving object from a video sequence based on subsequent object detection techniques. After pre-processing of the original video sequence, which is sampled from the video camera, the target moving object is detected with the improved algorithm in which the moving object region can be extracted completely through several processing of feature extraction and region based. Among the multiple moving objects of the video, the target object has been detected based on particular criteria of region that it occupies. The traveling speed of the target moving object from changes of its coordinate position from the video frames is determined by calculating the distance that the object moved in a sequence of frames with respect to the frame rate that the video is recorded.

Keyword-Object detection, feature extraction, region based segmentation, moving object and traveling speed.

I. INTRODUCTION

The traveling speed determination of a selected moving object from a suitable video sequence is an important aspect in some real world activity. Information about the speed of regions in an image can be also important to the scene interpretation. Object trajectories may be calculated, structural ambiguities may be solved, important events noted, and future situations predicted. At lower levels, information about speed of small structures in an image sequence can be used as an important cue for image segmentation. A discontinuity in the speed of small image areas strongly implies the existence of an object boundary. Visually dissimilar image regions with the same speed may either be part of the same object or semantically linked in the overall scene interpretation. If an observer is moving, perceived speed represents an important depth cue. Because of the value of speed information, increasing attention is being paid to methods analyzing time-varying visual environments.

However, several methods for speed determination of some customized moving object from video sequence have been developed to date. All of the methods required to detect the image object due to the positional shift in each frame in the given video sequence. In this paper our proposed method is quite simple and efficient to determine the traveling speed of the moving object from a suitable video sequence. In this method, firstly, we need to detect the target object that moves from initial frame to the last frame in the given video sequence. An alternate approach is directly calculating the speed information from spatial gradient of the images and local intensity changes over time due to speed. The speed of a moving object is calculated by the distance it traveled with respect to the time. Euclidean distance formula is used to calculate the distance between the sequences of frames. By using the values of distance with respect to frame rate, the speed of the object is defined. The defined is of 2- dimensional.

Each method has limitations. Obviously, procedures which require an accurate segmentation cannot be used to assist in the segmentation process itself. Methods dependent on cross correlation are computationally expensive. While some efficiency are possible, much effort still required to perform the matching at a dense enough sampling of points to be useful for tasks such as segmentation. Techniques using frame-to-frame differences provide efficiency, but require sophisticated higher-level processing to deal effectively with occlusions and situations in which object boundaries are not completely within the image frame. They relate intensity changes over time at appoint to the spatial intensity gradient in the neighborhood of the point. The frame rate is assumed to be sufficiently rapid such that the value of the gradient at a point does not change significantly between frames. Speed along the x and y axes are treated independently.

II. PROPOSED ALGORITHM FOR SPEED DETERMINATION

The procedure first taking a video sequence that contain moving object then process the video sequence to get all the frames and then convert video frames to image format. To reducing noise from image use median filter. Segmentation technique is applied for separation of foreground and background objects. Frame difference algorithm is used for separation of two sequential frames [6]. After performing segmentation operation a rectangular bounding box method is applied for feature extraction. After performing feature extraction operation to detect the moving object from the background scene and calculating the traveling speed between each adjacent pair of frames. Finally calculate the average traveling speed that is the required speed. To calculate traveling speed at first calculates distance between two consecutive frames.

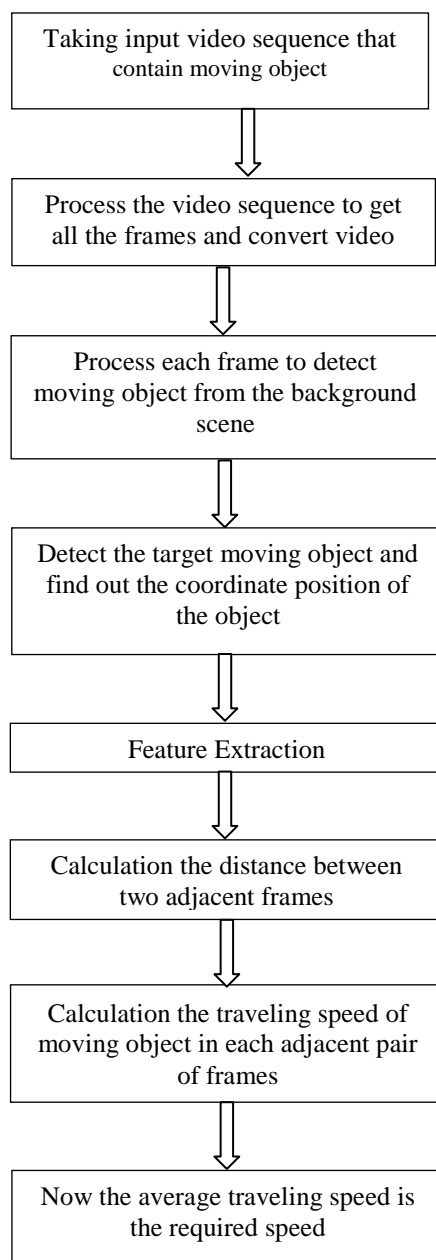


Figure 1: Flowchart of overall process for traveling speed determination from a video sequence

III. PROCESSING VIDEO FOR SPEED DETERMINATIONS

Video processing is a particular case of signal processing, which often employs video filters and where the input and output signals are video files or video streams. A video signal is a sequence of two dimensional (2D) images projected from a dynamic three dimensional (3D) scene on to the image plane of a video camera. The color value at any point in a video frame records the emitted or reflected light at a particular 3D point in the observed scene. Video sequences from the video camera is taken and process it as needed to convert AVI format and get all the frames of that video sequence which are inputted to the next phase of this work.

IV. DETECTION OF ALL MOVING OBJECTS

Detection of all moving objects is composed of the procedure Bounding box extraction and Frame difference Segmentation which is the most important part of this work. Segmentation and feature extraction procedures are given bellow:

Feature Extraction

For detecting object in Speed analysis can be viewed as three different problems [3].

* The first is the case when the camera is moving and the objects in the world are stationary. In this case, the extraction of camera motion is a challenge.

* In the second case, the camera is stationary, and objects in the world are moving.

* It is the combination of the two, where both the camera and some objects in the world are moving.

As, in our work the camera is stationary, so second case is applicable to this point. Different algorithm is usually applied in the second case. In this case, difference algorithm can be divided into two types: one is difference between continuous images; the other is difference between current image and background images. For difference between current image and background image, suppose that the gray value of current image at position (x, y) is $f(x, y)$, the gray value of background image at position (x, y) is $b(x, y)$, the difference between images can be written as: $d(x, y) = f(x, y) - b(x, y) \dots\dots\dots (1)$

For difference between continuous images, suppose that the gray value of image at position (x, y) at time t is $f(x, y, t)$, the gray value of image at position (x, y) at time $t+1$ is $f(x, y, t+1)$, the difference between images can be written as:

$$d(x, y) = f(x, y, t+1) - b(x, y, t) \dots\dots\dots(2)$$

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:

1. Take the input video sequence
2. Extract frames from video sequence
3. Convert frame to image sequence
4. Read the image difference
5. for(present position=initial value: final value) of Y resolution
6. for(present position=initial value: final value) of X resolution
 - (a) calculate the sharp change in intensity of image from top and bottom
 - (b) store the values in an array
7. Height of the bounding box is = bottom value – top value
8. for(present position=initial value: final value) of X resolution
9. for(present position=initial value: final value) of Y resolution
 - (a) calculate the sharp change in intensity of image from left and right
 - (b) store the values in an array
10. Width of bounding box = right value – left value
11. 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

$$\text{Height} = \frac{\text{bottom.value} - \text{top.value}}{2}$$

$$\text{Width} = \frac{\text{right.value} - \text{left.value}}{2}$$
12. Add the height value with top value and store it in a variable like mid top
13. Add the width value to the left value and store it in a variable like mid left
14. Assign the max intensity to the pixel at pixel value at (mid top, mid left).

Object Segmentation

The objective of segmentation is to partition an image into regions. When a moving object is segmented, a region of pixels assigned to the object is available. This region can be tracked using approaches like cross-correlation. The location of the region in the next frame is to be determined. A moving object usually corresponds to one or several tracked regions. Combination of several regions to one object is then performed at a higher level of abstraction [1]. To perform segmentation operation, region based method is used because it takes less processing time. We now turn from segmentation by finding boundaries (pixel differences) to segmentation by finding coherent regions (pixel similarities). This approach has specific advantages over boundary based methods:

- It is guaranteed (by definition) to produce coherent regions. Linking edges, gaps produced by missing edge pixels, etc. are not an issue.
- It works from the inside out, instead of the outside in. The question of which object a pixel belongs to is immediate, not the result of point-in-contour tests.

However, it also has drawbacks:

- Decisions about region membership are often more difficult than applying edge detectors.
- It can't find objects that span multiple disconnected regions. (Whereas edge-based method can be designed to handle "gaps" produced by occlusion—the Hough transform is one example.)

The objectives of region-based approaches can be summarized as follows:

- Produce regions that are as large as possible (i.e., produce as few regions as possible).
- Produce coherent regions, but allow some flexibility for variation within the region. Notice the inherent tradeoffs here. If we require that the pixels in a region be too similar, we get great coherency and probably won't span separate objects, but we over segment the image into regions much smaller than the actual objects. If we allow more flexibility, we can produce larger regions that more likely fill entire objects, but they may cross multiple objects and "leak" across what should otherwise be boundaries. Remember: the real goal is to find regions that correspond to objects as a person sees them—not an easy goal [2].

The basic formulation for Region-Based segmentation is:

- (a) $\bigcup_{i=1}^n R_i = R$.
 - (b) R_i is a connected region, $i=1, 2, \dots, n$
 - (c) $R_i \cap R_j = \emptyset$ for all $i = 1, 2, \dots, n$.
 - (d) $P(R_i) = TRUE$ for $i = 1, 2, \dots, n$.
 - (e) $P(R_i \cup R_j) = FALSE$ for any adjacent $P(R_i)$ is a logical predicate defined over the points in the R_i and \emptyset is the null set.
- (a) Means that the segmentation must be complete; that is, every pixel must be in a region.
- (b) Requires that points in a region must be connected in some predefined sense.
- (c) Indicates that the regions must be disjoint.
- (d) Deals with the properties that must be satisfied by the pixels in a segmented region. For example $P(R_i) = TRUE$ if all pixels in R_i have the same gray level.
- (e) Indicates that region R_i and R_j are different in the sense of predicate P [4].

V. OBJECT DETECTION AND FIND OUT ITS POSITION

To identify a single object as target object with its 2D coordinate position from multiple object in each frame from a video sequence, our algorithm always detect the object that is occupied the maximum region [5]. So, when we will take the video sequence for speed determination of the target object, we will focus on the target object as much as possible that the object will occupy maximum region compared to the other moving object. And of course the camera must static. To identify the position of the target object in each frame of input video sequence the centred point of the total region that is occupied by the object have been considered as reference point.

In the similar way, the reference point of target object in each frame of the video is find out and stores these

positions. Finally from these positions, the movement of target object is measured and the traveling speed is calculated according to the speed calculation procedure.

Procedure for object detection:

- for $i=0$ to (totalFrame-1) do
- Read frame[i],
 - Update frame[i] using Extract bounding box,
 - process frame[i] as follows :
 - Determine the connected components.
 - Run-length encodes the input image.
 - Scan the runs, assigning preliminary labels and recording label equivalences in a local equivalence table.
 - Resolve the equivalence classes. Relabel the runs based on the resolved equivalence classes.
 - Compute the area of each component
 - Remove small objects bellow a threshold.
 - Create morphological structuring element.
 - Close the binary image by the structuring element.
 - Measure image regions
 - Find the maximum region
 - Identify the centered location (x, y) of that region.
 - Return x-coordinate value and y-coordinate value
 - End

VI. Speed Determination of A Selected Moving Object

Several methods for speed determination of some customized moving object from video sequence have developed to date. All of the methods required to detect the image object due to the positional shift in each frame in the given video clip. In our work our proposed method is quite simple and efficient to determinate the traveling speed of the moving objects from video sequence. In this method, firstly, we need to detect the target object that moves from initial frame to the last frame in the given video clip that has already been discussed above. The speed of a moving object is defined as the total amount of distance traveling in unit time.

Algorithm/procedure for speed determination of a selected moving object

- Load the input video file containing moving objects.
- Process the video file to get the required information about the video file.
- Find the number of frames N_F of the video
- Find the frame rate R_F of the video
- Calculate the total duration of the video as $T \leftarrow \frac{N_F}{R_F}$ second and unit time $\Delta t = \frac{T}{N_F - 1}$
- Determinate the displacement D_i of the object between the i -th frame and $(i+1)$ the frame
 - Read the centroid position of each frame.
 - Calculate the distance between two centroid frames.
 - for (present position = initial value : final value) of X resolution.
 - for (present position = initial value : final value) of Y resolution.

- Calculate change in distance by distance

$$D_i = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

Where X_1 = previous pixel position and X_2 = present pixel position in width

Y_1 = previous pixel position and Y_2 = present pixel position in height.

- Store all the distance values in an array.
- Calculate the speed S_i between the frames F_i and F_{i+1} as $S_i = D_i / \Delta t$
- Repeat step 6 to 7 for $i=0$ to $N_F - 2$ to determinate all the speed between the frames.
- Calculate the average value of speed as $\frac{\sum(S_i)}{N_F - 1}$ for $i=0$ to $N_F - 2$
- $S_{final} = \frac{\text{TotalDisanCapteredByCameraMeter (Widely)}}{\text{TotalPixel (widely)}}$
- S_{final} is the real speed (meter/second) of the moving object.

VII. EXPERIMENTAL RESULTS AND DISCUSSION

Our approach is implemented on a PC under Windows operating system using MATLAB to execute all the required source code. The technique discussed above is applied on various video files like as inputvideo1.avi, inputvideo2.avi etc. The first code allows the user to select a video file is executed. The results of the traveling speed of a selected target moving object from a video sequence are given bellow: Several attributes such as file size, number of frames, frame rate are find out from selected input video sequences.

Several frames of inputvideo1.avi are given bellow:



Fig 2a: 1st frame

Fig 2b: 2nd frame



Fig 2c: 4th frame

2d: Last frame

In first input video the file size is 73728 Kb, number of frames 100 and frame rates per second is 15. After getting the object position in each frame of the input video the traveling speed of that object have calculated with the speed calculation procedure. Using the proposed speed determination algorithm the resultant traveling speed of the moving object is 30 meters per second. Several frames of inputvideo1.avi are given bellow:

Fig 3a: 1st frameFig 3b: 2nd frameFig 3c: 5th frame

3d: Last frame

In second input video the file size 23020 Kb, number of frames 140 and frame per second is 25. After getting the object position in each frame of the input video the traveling speed of that object have calculated with the speed calculation procedure. Using the proposed speed determination algorithm the resultant traveling speed of the moving object is 1.52 meter per second.

VIII. CONCLUSION

In this paper, an attempt has been made to develop a virtual system for determination the traveling speed of a selectable moving object of a suitable video clip using subsequent object detection technique based on feature extraction and region based segmentation near to the real time [7]. Feature extraction using bounding box approach and the region based segmentation techniques are relevant to detect multiple moving object from a video clip to determinate the traveling speed of a target moving. 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 has experimentally been shown to be quite accurate and effective in detecting the target moving object from a video sequence and also determined the traveling speed. After all, for the test bench for this work, the traveling speed of a selected moving object of a suitable video clip has been determined at a satisfactory level. In this paper, the primary works are the video processing as well as image processing for the detection of moving object within the video clip, but it focuses on the detection of multiple objects from images in the video sequences and detecting the target object based on region that it occupies to determine the traveling speed of the moving object.

IX. REFERENCES

1. Alper Yilmaz, Omar Javed, and Mubark Shah, "Object tracking: A survey", *ACM Comput. Surv.*, 38(4): 13, 2006.
2. Bryan S. Morse, "Segmentation (Region Based)", Brigham Young University, 1998–2000, Last modified on March 3, 2000.
3. Yong Fan¹, Zhengyu Zhang², "Journal of Communication and Computer, ISSN1548-7709, USA" Jul. 2006, Volume, No.7 (Serial No.20).
4. http://en.wikipedia.org/wiki/Region_growing Region growing From Wikipedia, the free encyclopedia
5. Shuan Wang, Haizhou Ai, Kezhong He, Difference-image-based Multiple Motion Targets Detection and Tracking, *Journal of Image and Graphics*, Vol. 4, No. 6(A), Jun., 1999: pp. 270-273.[8]
6. Gonzalez, R.C. Woods, R.E [1992], "Digital Image Processing".
7. Murat Tekalp, *Digital Video Processing*, Tsinghua University Press and Prentice Hall, Beijing, 1998.[7]