Object Tracking Using Adaptive Frame Differencing And Dynamic Template Matching Method

Shivam Mittal

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Object Tracking Using Adaptive Frame Differencing and Dynamic Template Matching Method

Thesis submitted in

May 2013

to the department of

Computer Science and Engineering

of

National Institute of Technology Rourkela

in partial fulfillment of the requirements

for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Shivam Mittal

[Roll: 109CS0162]

under the guidance of

Prof. Pankaj Sa

Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, Odisha, India

Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela-769 008, India. www.nitrkl.ac.in

Dr. Pankaj K Sa
Assistant Professor

MAY 2013

Certificate

This is to certify that the work in the project entitled *Object Tracking using Adaptive Frame Differencing and Dynamic Template Matching Method* by *Shivam Mittal* is a record of his work under my supervision in partial fulfillment of the requirements for the award of the degree of *Bachelor of Technology* in *Computer Science and Engineering.*

Pankaj K Sa

Acknowledgment

I take this opportunity to express my profound gratitude and deep regards to my guide Prof Pankaj Sa sir for his exemplary guidance, monitoring and constant encouragement throughout the course of this project. The blessings, help and guidance given by him time to time shall carry me a long way in the journey of life on which I am about to embark.

I am obliged to all the professors of the Department of Computer Science and Engineering, NIT Rourkela for instilling in me the basic knowledge about the field that greatly benefitted me while carrying out the project and achieving the goal.

Also, I am highly indebted to Mr. Rahul Raman Sir, Department of Computer Science and Engineering, National Institute of Technology, Rourkela who invested his full effort in helping me in finalizing this project within the limited time frame and keep motivated me all the time.

Lastly, I am grateful to my friends, for their relentless support in augmenting the value of work; my family, for being considerate and appreciative throughout; and Almighty, for everything.

Shivam Mittal

Abstract

In this project I have used the concept of frame differencing and background subtraction algorithm to propose a modified algorithm which can be used effectively and accurately with comparison to both frame differencing method and background subtraction model used individually for detecting moving objects in a sequence of frames.

In this project we have used the method of frame differencing to propose a new adaptive frame differencing method which shall take account of the velocity of the moving object in order to find the number of frames to be skipped each stage of detection to calculate inter-frame difference in order to get the region of moving object. The above procedure is combined with background subtraction model with a new idea of changing the background dynamically to have a better image of the moving object. The area obtained from adaptive frame differencing is added with the area obtained from adaptive background subtraction model to have a clear view of the pixels associated with the moving object. After getting the detected object the centroid of it is passed to the tracking module in order to track the object in upcoming frames by using the concept of dynamic template matching algorithm which uses a correlation function in order to track the detected object in the region of interest in the upcoming frames. When the tracking fails the algorithm goes back to detection module and the process repeats.

Thus we proposed a effective tracking algorithm which can be use even if the object of interest is far away from the camera independent of the motion of the object. By using the concept of adaptive frame differencing between frames taken into consideration for frame differencing will change accordingly with speed of the moving object and with help of adaptive background subtraction method, background got updated automatically after detection and hence background is changing dynamically. The combination of both will pass through binary thresholding, then through morphological erosion and dilation operation in order to remove insignificant movement due to camera flickering and to remove noise. The obtained image is precisely denotes the moving object region.

Contents

List of Figures

Chapter 1

Introduction

1.1 Object Tracking

Object Detection and tracking is considered as important subject within the area of computer vision. Availability of high definition videos, fast processing computers and exponentially increasing demand for highly reliable automated video analysis has created a new and great deal for modifying object tracking algorithms. Video analysis has three main steps mainly: interesting moving objects detection, the tracking the object detected from frame to frame and visualizing and analyzing to identify the behavior of the object in the entire video.

1.2 Application of Object Tracking

Some of the important applications of object tracking are:

- **Automated video surveillance:** In order to monitor the happenings in a particular area,to detect and recognize moving objects and to detect unlikely events and to report suspicious,criminal activities applications are developed in computer vision system
- **Robot Vision:** for navigation of robot,different obstacles in the path are identified by the steering system in order to avoid collision.In case the obstacles are itself in motion then we need a real time object tracking system.

- **Monitoring of Traffic:**Highway traffic can be continuously monitored using moving cameras.The vehicle breaking any type of law or involved in any illegal activities can be detected and tracked using a object tracking system
- **Animation purpose:** Animation can be supported by using object tracking algorithm
- **Human computer interaction :** it can be used for automatic attendance system in many areas and to record the in and out time of the object.

1.3 Steps of Object Tracking

Object tracking can be performed stage wise stage and the stages are:-

- **O**bject Detection
- **O**bject Tracking
- **D**ynamic Template Matching

1.3.1 Object Detection

The computer technology associated with computer vision and image processing which deals with the task of detecting and identifying instances of semantic objects belonging to a certain class in video and images. Face detection, pedestrian, detection, vehicle tracking etc. are the well-researched domains of object detection. Object detection is widely used in computer vision including areas like image retrieval and surveillance of video.
In a video surveillance system moving detection algorithms can be broadly classified into two categories depending on their relative movement between scene of surveillance and camera.
If the camera is ideal i.e. is fixed then the surveillance scene will remain unchanged leading to method called static background detection method. If camera is not static then camera and scene of surveillance will have some relative movement and the method is called as moving background detection method. In static background

detection method position and size of pixels in the background will remain same in different frames of an image sequence .Thus it can use method of method of difference of the pixels in term of intensity or color value to determine moving region and extract moving object at the same previous position in the different frames. This method of extraction is called as inter-frame difference method

1.3.2 Object Tracking

Tracking is the problem of estimating the trajectory of an object as it moves around a scene. Ultimate aim of object tracking is to associate objects targeted in upcoming video frames. This can be very difficult in case relative motion between moving objects and frame rate is high.
Change of orientation of tracked object with passing of time can increase the complexity of the process. Thus to deal with the above problem we are employ a model for motion of the object to show how the image of the target might get change for every possible motion of the object.
Basically tracker assigns consistent and unique labels to the objects tracked in different frames of a video. Additionally, depending on the domain of tracking, a tracker can also provide centroid information of the object, orientation, shape, or are of an object.

Why Tracking Is Difficult

Tracking objects can become complex due to following reasons:

- projection of the 3D world on a 2D image will cause loss of information
- presence of noises in the image
- irregular object motion,
- non rigid nature of objects,
- occurrence of occlusion(partial or full)
- objects shape is complex,

- change in the illumination or intensity of the scene

1.4 Template Matching

The generated templates from detection module are passed on to the tracking module, which initiate tracking the moving object with a given input reference template. The tracking module makes use of template-matching to search for the input template in the future frames grabbed by the video. . A new template is generated in case the object is lost during tracking due to change in its appearance and used further. Generations of such templates are dynamic which helps to track the object in robust manner. The theory behind template matching is given below

- The matching process moves the template image to all possible positions in a larger source image and computes a numerical index that indicates how well the template matches the image in that position.

- Matching is done on a pixel-by-pixel basis of both input template and masked template.

1.5 Literature Review

1.5.1 Image Segmentation

Jong Bae Kim, Hang Joon Kim presents a paper on efficient region based motion segmentation method of moving with a focus on video monitoring system in a traffic scene [1]. The research is based on using aadaptive thresholding to automatically choose the threshold value instead of determining the threshold value manually.Second,in segematation phase of motion,pixels having similar intensity and motion information are segmented by the help of weighted k means clustering method in order to detect the motion to the binary region of the binary mask.Thus

we have used method of adaptive thresholding in our algorithm to have minimum computational load.
Time difference between current frame and next frame Ta is computed in the kim method(k frame and k+1 frame).Then it calculates the difference between current frame and a static background image to find Tb.After that the difference images Tb and Ta are multiplied to get Tc with Ta and Tb thresholded already.Finally morphological operation and connectivity analyzing is done for Tc to get the moving targets.

1.5.2 Frame Differencing

Karan Gupta and Anjali V. Kulkarni presented a paper on implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching which makes use of interframe difference of frame 2 and two frames are used for frame differencing to obtain moving object region.As in case of slow moving object the frame difference will be nearly zero.Hence this method cannot be applied to slow moving object.But we incorporated the idea of skipping frames from this paper to get the accurate moving object region [2].
Budi Sugandi, Hyoungseop Kim., Joo Kooi Tan and Seiji Ishikawa presented an research work having title A Block Matching Technique for Object Tracking Based on Peripheral Increment Sign Correlation Image.The paper has concept of adding two consective interframe difference in order to determine the moving object contour accurately.But the same problem arises as the adding the diffrence of three frames wont give the moving contour accurately.But this paper gives us the idea of using three frame instead of two frames for frame differencing [3].

1.6 Motivation

The Motivation proposing the new algorithm is:-

- A better method is which can detect and track object simultaneously with maximum efficiency.
- To detect both slow and fast moving object, to makes it independent of speed of moving object and more reliable

1.7 Assumption

In the following algorithm we have imposed certain constraints in the motion of the object and the video to have proper result.The following are the constraints we have putted up : -

- The object motion is smooth with no abrupt changes
- Number of object to be tracked is one and we have prior knowledge about the size ,shape and appearance of the object.
- it should not make sudden change in the direction of motion while moving out of the viewing range of the video.

1.8 Outline of Thesis

The thesis consist of four chapters following this chapter:

Chapter 2: Object Detection

In this chapter we have discussed what are the various existing methods of object detection and what are the shortcoming associated with each detection technique.We have also discussed about binary thresholding operation and morphological operation and its associated role with the image processing.Finally i proposed a new detection method wwhich can be more reliable and accurate.

Chapter 3: Object Tracking

In this chapter i have discussed about the tracking of the moving object with implementing the dynamic template matching algorithm

Chapter 4: Matching algorithm for templates

In this chapter we have discussed the algorithm and the formula used for template matching.

Chapter 5: Results of Implementation

In this chapter we have showed various figures showing the result. **Chapter 5: Advantages of proposed Algorithm**

In this chapter we have stated the advantages of using new proposed algorithm.

Chapter 6:Conclusion

Here we have concluded the main points of the thesis..

Chapter 2

Object Detection

2.1 About the method

Object detection is the method of detecting moving objects from a sequence of farmes.We have deployed method frame differencing for object detection followed by binary thresholding function.Thresholding function output will be again a binary image which is further processed by iterative mathematical morphological dilation and erosion operation.Thus from the resultant image we will find out the centroid of the moving object in order to extract a rectangular template to be used in object tracking module.

2.2 Frame Differencing

Frame differencing, also popularly known as temporal difference is the method of subtracting the video frame at time t-1 with the background model for the frame at time t. Inter frame difference method computes the absolute difference between the previous frame and current frame.

2.2.1 Two Frame Difference Method

In two frame differencing method two frame are taken into consideration and then subtracted in terms of intensities at each pixel resulting in an image giving brief idea about the moving object region.
In this method pixel wise differences current frame and previous frame of the video pixel wise differences are used to get extract the moving object [4].
It is type of background subtraction method in which the last frame become the background for the current frame and difference is calculated.
The disadvantages of two frame differencing method are :-

- All the important pixels are not extracted properly resulting in holes left in the moving entity.
- In case of slow moving object the difference is found to be almost zero thus giving no moving region.
- Unable to detect object in case they stop moving.
- This method is very sensitive to noise and changes in illumination and does not take in account local consistency properties of the change mask.
- It also fails to detect the non-background objects in case they are not in motion anymore.

2.2.2 Three Frame Difference Method

In this method 3 consecutive frames are extracted $frame\_cur, frame\_past, frame\_future$.
And subtraction is performed between frame_cur and frame_past and also frame_cur and frame_future and the resultant images of both the subtraction are multiplied to get moving object region.
The disadvantages of three frame differencing method are :-

- Choice of good frame-rate for three-frame differencing depends on the size and speed of the object.

- In case of slow moving object the difference is found to be almost zero thus giving no moving region.
- Multiplication result in empty space and will get treated as background pixels.
- Objects having intensity values uniformly distributed, the pixels are considered as part of the background.

2.2.3 Backgroun Subtraction Method

In this method current frame is subtracted with background provided in order to create moving foreground object region. This subtraction method can extract the object shape well provided that the background is static and can easily adapt the illumination change [5]. The disadvantages of Background subtraction method are :-

- Background may change lightly with time
- And background subtraction may result in inclusion of noise in resultant image

2.3 Binary Thresholding

Thresholding is used for image segmentation. Binary thresholding operation is performed to separate the pixels corresponding to the moving object to that of the pixels of the background. This operation is used to remove any inaccuracies present due to the camera flickering. This operation result is a binary image where the pixels corresponding to the moving object is set to 1 [6].
In thresholding we choose a parameter called the threshold of brightness (T) and applied to the frame difference image as follows:-
IF f[m,n]>=T
fb[m,n]=object=1
ELSE
fb[m,n]=background=0

2.4 Morphological Operation

Erosion and dilation are two main basic operators in mathematical morphology. The basic function of erosion operator on a binary image is to erode away the boundaries of pixels associated with the foreground (usually the white pixels). Thus areas of foreground pixels shrink in size, and "holes" within those areas become larger in size. Dilation function on binary image is to enlarge the areas of pixels associated with the foreground (i.e. white pixels) at their borders resulting in growth of size of the areas of the foreground and the background pixels with in them shrink.
Above two operations are performed to remove small particles from the binary image. The operation again produces a binary image. Thus this function is used to ensure insignificant small movements in the background are ignored properly in order to ensure better detection of object.

2.5 Algorithm

The proposed algorithm will make use of both three differencing method and background subtraction method The terms used in above algorithm are:-

- **Vel_thresh1**=objects moving with speed less than this value will be considered as slow moving object.
- **Vel_thresh2**=objects moving with speed greater than this value will be considered as fast moving object. Both vel_thresh1 and vel_thresh2 will depend upon the size of the moving object and distance of the moving object from the camera.
- **Frame_diff**=number of frames to be interleaved.
- **Prev_cent**= centroid of the object in previous frame.
- **Cur_cent**= centroid of the object in current frame.

1. Set $vel\_thresh1$ and $vel\_thresh2$ in order to distinguish slow and fast moving object.

2. Set $frame\_diff = 1$ and $prev\_centroid = (0, 0)$

3. Set 2nd frame as $frame\_cur$.

4. $frame\_past = frame\_cur - frame\_diff$ and $frame\_past = frame\_cur + frame\_diff$

5. Perform following operations
   $sub1 = |frame\_cur - frame\_past|$
   $sub2 = |frame\_curfarme\_future|$
   $Resultant\_Image = sub1 + sub2$

6. Perform binary thresholding operation to separate the pixels corresponding to moving object from the background.

7. Perform an Iterative mathematical morphological erosion to remove very small particles from the binary image.

8. Calculate the centroid of the binary image.

9. Calculate frame speed by finding Euclidean distance between $prev\_cent$ and cur_cent.
   $frame\_speed = Euclideandistance/frame\_diff$

10. IF frame_speed<vel_thresh1 AND frame_diff>=num_of_frames THEN Increment $frame\_diff$ and GOTO step 4
    ELSEIF(speed>vel_thresh2 AND frame_diff>=1) THEN Decrement $frame\_diff$ and GOTO step 4
    ELSE GOTO step 11

11. Subtract current frame with background image
    $Resultant\_image2 = |frame\_cur - background\_frame|$

12. Add both the resultant image
    $Resultant\_image = resultant\_image1 + resultant\_image2$

13. Perform binary thresholding and morphological dilation to remove the noises and insignificant movements.

14. Update background by using formula $Background = (1 - a) * background\_frame + a * frame\_cur$ where $a \simeq 0.05$

15. Obtain the $centroid(cog\_X, cog\_Y)$ of the object tracked to tracking module.

2.6 Tracking Algorithm

The tracking algorithm we have employed is based on template matching process which goes on searching the generated template in the region of interest only [7].

1. Get the positional information (cog_x,cog_y) from the object detection module.

2. A image template Ti is generated by extracting a square image from the last frame grabbed by the video. The coordinates of above square will be:-

   - (cog_x-a,cog_y-a)
   - (cog_x-a,cog_y+a)
   - (cog_x+a,cog_y-a)
   - (cog_x+a,cog_y+a)

   Where a is a variable whose value depend upon the size of the object.

3. Now the region of interest in the upcoming frames to track the object will be a square region of height b and width b with coordinates is :-

   - (cog_x-b,cog_y-b)
   - (cog_x-b,cog_y-b)
   - (cog_x-b,cog_y+b)
   - (cog_x+b,cog_y-b)

where b will be equal to twice the height of the template image.

4. We will start implementing the efficient template matching algorithm to the Region of interest with the template Ti

   IF the template matching algorithm is successful
   Co-ordinate positions are saved and the template has been updated with the match and the tracking continues with the next frame having same region of interest and updated template
   ELSE
   go back to the detection module as the object has changed its appearance significantly

Chapter 3

Block Matching Algorithm Psuedocode

Efficient template matching algorithm Psuedocode used is

1. Obtain the height and width of Region of interest;

   r1 c1=size(i3);

   here i3 is the region of interest.

2. Obtain the width and height of the template to be searched

   r2 c2=size(i2);

   Here i2 is the template

3. Build the correlation matrix

   corr=zeros(r1-r2+1,c1-c2+1);

4. for i=1:r1-r2+1

   for j=1:c1-c2+1

   mask=i3(i:i+r2-1,j:j+c2-1);

   Mask will extract a picture of size [r2,c2] to compare with template

5. corr(i,j)=corr_function(mask,i2);

$$cor = \frac{\sum_{i=0}^{N-1}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\sum_{i=0}^{N-1}(x_i - \overline{x})^2 \cdot \sum_{i=0}^{N-1}(y_i - \overline{y})^2}}$$

Figure 3.1: Correlation Formula for Template Matching

3.1 Correlation formula

The terms used in the above correlation formula are:-

- x_i is the template gray level image
- y_i is the source image section
- N is the number of pixels in the section image
  (N= template image size = columns * rows)
- $\bar{x}$ is the average grey level in the template image
- $\bar{y}$ is the average grey level in the source image

The value cor is between 1 and +1, with larger values representing a stronger relationship between the two images.
if corr (i,j)>=Threshold value
THEN template matching is successful;
Else
unsuccessful template matching

Chapter 4

Results

Figure 4.1: background image

Figure 4.2: Current Frame

Figure 4.3: Previous Frame

Figure 4.4: Result of two difference method

Figure 4.5: Result of three frame differencing method

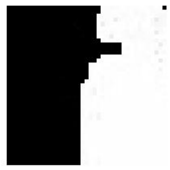

Figure 4.6: Result showing template extracted from a frame

Figure 4.7: Result showing matched template with given Threshold

Figure 4.8: Result showing only Region of Interest

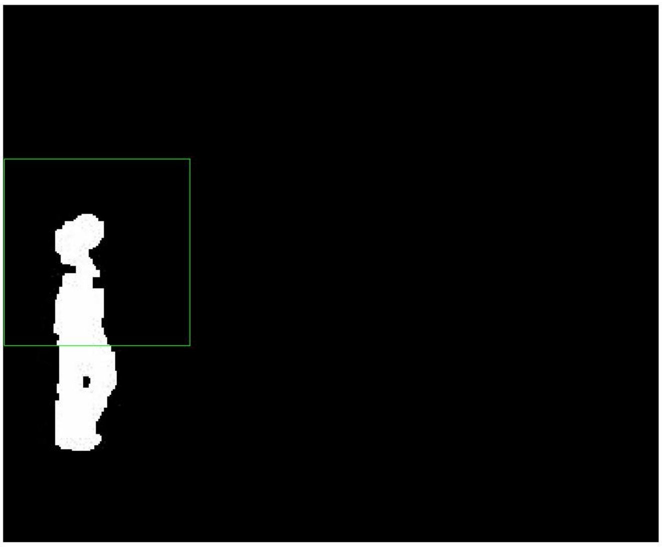

Figure 4.9: Result showing Region of Interest in a Frame

Chapter 5

Advantages Of Using Proposed Algorithm

The following are the advantages of using proposed algorithm deploying both Inter-frame Diffrencing and Dynamic Template Matching Algorithm :-

- The output of Object Tracking can be further utilized by the Pan-Tilt of moving camera to operate the pan or tilt motors according to the behavior of the object.
- The template matching algorithm decreases the computation time as the Region of interest (ROI) is dynamically found out.
- Modest computational load makes the algorithm more effective and cheap in cost.
- Inter-frame difference is adaptively changed depending upon the speed of the moving object.
- The above automated object detection and tracking system can work perfectly under low lighting conditions so it can be used in surveillance and video conferencing.
- Background is based on previous background as well as previous frame and thus it is very highly adaptive and can adopt changing background more faster than any other detection method.

- This method can also be used for small moving objects.
- The template matching technique can successfully tracked the interest moving object in the occlude condition.

Chapter 6

Conclusions

A robust and efficient automated single object tracking system is presented. The system has been implemented using algorithm based on adaptive frame differencing, adaptive background subtraction and dynamic template matching. The algorithm has experimentally been shown to be quite accurate and effective in detecting a single moving object even under bad lighting conditions or occlusions. Results from the experiment shows that the detection effect in case of 2 frame differencing shows only very small part of moving objects with varying intensity, thus cannot show the figure and contour of moving objects correctly.

The experiment producing output doesnt have any noise present and doesnt detect any false object. Effective tracking is performed using dynamic and template matching producing accurate result. Thus the proposed algorithm is better than the existing algorithms.

Bibliography

[1] J.B. Kim and H.J Kim. Pattern recognition letters. In *Efficient region-based motion segmentation for video monitoring system .*, pages 113–128, 2003.

[2] Karan Gupta and Anjali V. Kulkarni. Implementation of an automated single camera object tracking system using frame differencing and dynamic template. `http://home.iitk.ac.in/~anjalik/ObjectTrackingPaper.pdf`.

[3] Joo Kooi Tan Budi Sugandi, Hyoungseop Kim. and Seiji Ishikawa. A block matching technique for object tracking based on peripheral increment sign correlation image. `http://cdn.intechopen.com/pdfs/13884/InTech-A_block_matching_technique_for_object_tracking_based_on_peripheral_increment_sign_correlation_image.pdf`.

[4] Reid A. Shenq Wang Markandey, V. Motion estimation for moving target detection. *IEEE Transactions on Aerospace and Electronic Systems*, 32(3):886–874, 1996.

[5] Liang Yanbing Wang Bin, Pan Jianshou. Moving detection of video images based on matrox card. *Journal of Northwest University*, 2(3):1–7, 2004.

[6] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. on System, Man and Cybernetics*, 9(1):62–66, 1979.

[7] BELL J. W. SCHWEITZER, H. and F. WU. Very fast template matching. *In European Conference on Computer Vision (ECCV)*, -:358–372, 2002.