Moving Targets Tracking on Construction Sites in Complex Scenes

SHEN Qiaonan, AN Xuehui
Tsinghua University, Beijing 100084, China

ABSTRACT**:** This paper mainly discusses the possibility of- and need for-tracking moving targets on construction sites in complex scenes using statically placed video cameras. A reliable target tracking system could work under various conditions, such as changes of lighting and movements of background objects. There are many problems associated with objects detecting and tracking on a construction site, while the major two of them are the changing photometric visual content throughout the course of a day and the presence of moving obstacles. Firstly, to solve the problems for detecting, since the background modeling is one of the most important parts for object detection, this paper presents an effective and adaptive background modeling method for detecting objects on construction sites. The background data is initially obtained from the first several images involving many moving obstacles by improving average value model. Then to deal with the change of lighting, the image is divided into several sub-regions. The background model is updated when the number of different regions between current image and background image is larger than a pre-defined threshold value. Since the scene of construction sites is considered to change slowly, that the updating process does not need to operate at every frame and in every region. As a result, a lot of computational cost is decreasing to ensure real-time object detection in the tracking system. Finally, several video sequences captured from construction sites under different situations are tested and the experimental results compared to the average value model show that the algorithms are efficient and accurate to work under practical construction site conditions.

KEYWORDS: moving targets, tracking, background modeling

1. INTRODUCTION

Traditional management of construction projects often involves human judgment, high costs, and does not provide managers with timely and accurate control data. Intelligent video surveillance systems have recently gained in importance. The ability to monitor people and vehicles automatically is useful for safety, security, and process analysis. To address the need for effective and efficient data collection, we have developed a real-time system called UDTTS (User-Defined Target Tracking System). This system is based on digital image process presented in previous work(Shen, 2008), and substantially solves the problems under the pre-known static background and fixed-camera situations.

Background subtraction and thresholding methods are a preliminary step to moving-object detection and subsequent processing. They are necessary to obtain the masks of moving objects. For detecting moving targets, such as humans and vehicles, the UDTTS subtracts the current frame from the known background. However, as the background is rarely known beforehand at construction sites, the background must be modeled before background-subtraction processes can be used. A robust system should be capable of dealing with lighting changes, the effects of moving elements of the scene(e.g. swaying trees), movement through cluttered areas, objects overlapping in the visual field, and objects being introduced or removed from the scene. In this paper, an adaptive background model is proposed and applied in the UDTTS for targets tracking at construction sites in complex scenes.

1.1 Objectives

Background modeling and subtraction separate the foreground from the static parts of the scene in motion analysis. Often the assumption is made that an initial model can be obtained by using a short training sequence in which no foreground objects are present. However, in some situations, including construction sites, it is difficult or impossible to control the area being monitored. In such cases it might be necessary to train the model by using a sequence that contains foreground objects. Our goal is to create a robust, adaptive tracking system that is flexible enough to handle variations in lighting, moving scene clutter, multiple moving objects and other arbitrary changes to the observed scene.

In this paper, we present a method for background modeling based on previous algorithms that is able to account for dynamic scenes. We provide comparative results of moving-target tracking on construction jobsites by using cost-efficient video cameras. The principal objective is to test and to demonstrate the feasibility of tracking targets from statically placed cameras.

1.2 Outline

The paper is organized as follows: a brief description of the existing methods for background modeling is given in Section 2; Section 3 provides the details on the applied adaptive background modeling method; Section 4 presents the experimental results for a complex scenario and describes the results in terms of tracking stability.

2. EXISTING METHODS

A standard method of adaptive background model is to average the images over time, creating a background approximation which is similar to the current static scene except where motion occurs. While this is effective in situations where objects move continuously and the background is visible a significant portion of the time, it is not robust to scenes with many moving objects, particularly if the objects move slowly. Methods based on differences in adjacent frames are very sensitive to noise and to illumination changes. When the number of frames in the sequence is large and there is little change between consecutive frames, another solution to motion detection is background modeling. This technique is routinely used in surveillance applications by using a fixed camera. Background modeling methods build a probability density function of the intensity at each individual pixel. In a static environment, the statistical distribution of a pixel can be represented by a single Gaussian distribution. The foreground pixels are determined as those for which the intensity value differs from the mean background model. Clustering then allows the detection and the segmentation of foreground objects. A variable number of Gaussian distributions corresponding to each different foreground object can be added (Stauffer, 1999). When the background changes too rapidly, the variance of the Gaussians becomes too large and the computational cost becomes unacceptable.

3. OUR APPROACH

There are many problems associated with object detection and tracking on a construction site, while the major two of them are the changing photometric visual content throughout the course of a day and the presence of any number of moving obstacles. Our approach involves background initialization and background updating based on a gray image which is divided into many regions.

In general, there should be more data supporting the background distributions because they are repeated, whereas pixel values for different objects are often not the same color. Since the average value model is simple and usually easily implemented, an improved method for background initialization is proposed. In our approach, the gray image is divided into several regions. The gray value of all the pixels in each region changes as the illumination changes. When the pixels change slowly and the difference between the current frame and the existing background frame is smaller than a pre-defined threshold, the foreground pixels can be separated from the background pixels, and the background model does not need to be updated. Conversely, when the entire scene changes significantly, it is difficult to detect moving targets and the background model must therefore be updated. The proposed method for adaptive background modeling is applied to the UDTTS and its flow chart is shown in Figure 1. Frames are transformed to gray images as the video sequence acquired by a static camera is input. The first *N* frames are used to generate an initial background. Then, the input current frame is substracted from the background and the remaining input is compared to conditions of background updating, as explained in more detail later. If the pattern matches(“yes”), the background model is updated by using the background pixels in the current frame. If the input does not match(“no”), the existing background model does not need to be updated and the foreground pixels are processed step-by-step for subsequent detection and tracking.

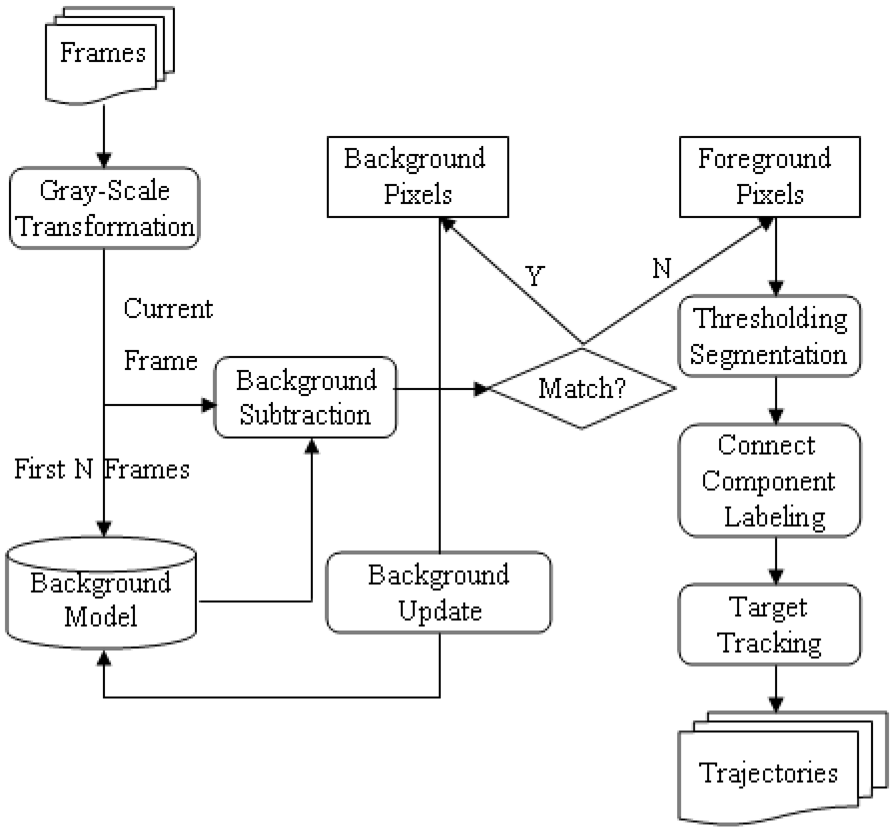

Figure 1: The flow chart of the UDTTS system using background modeling.

3.1 Background Initialization

The background initialization problem is defined as follows. The input is a short video sequence or the first N frames of a video, in which any number of moving objects might be present. The goal is to output a single background model that describes the scene for the subsequent processing.

It is assumed that the first N frames are input in order to generate a background frame. A pixel of background frame at position (x, y) is calculated by

$$\frac{1}{N}\sum_{i=1}^{N} I_i(x,y)\,, \tag{1}$$

where $I_i(x,y)$ denotes the gray value at position (x, y) and i is the frame number. Since the change of pixel values in sequential frames is caused by the movment of targets, the value of N is determined by the scene: when the incoming frames at first are simple with few moving targets, and with targets moving at high speed, N will be small. Otherwise, the pixel values denoted foreground account for a large proportion of the sum, N must be large enough to eliminate the effects of foreground pixels and obtain a perfect background frame.

To avoid this shortcoming, it is noted that pixel values occurred continuously and their frequency at one pixel in N frames. The most-frequent value, rather than the average value is assigned to the pixel in the background image. In cases where there are many moving obstacles, this method can form a background frame quickly and accurately.

3.2 Background Update

In order to adapt the algorithm to illumination changes throughout the day, background updating is required. The image is divided into $M*N$ sub-regions, as shown in Figure 2, and every sub-region background is gotten according to variety of gray grade, then the background image is rebuilt according to the computing results of each sub-region. M, N are determined by the size of the moving targets.

After gray-scale transformation, the image undergoes subtracting of the known background and thresholding segmentation as

$$\Delta I_i(x,y) = I_i(x,y) - BG(x,y) \tag{2}$$

$$T_i(x,y) = \begin{cases} 1, if\, \Delta I_i(x,y) > T \\ 0, else \end{cases}\,, \tag{3}$$

where $BG(x,y)$ denotes the pixel value at (x, y) in the background image, $T_i(x,y)$ denotes the pixel value at (x, y) in the binary image , T is the pre-defined threshold and i is the frame number.

1	2	…	N
N+1	N+2	…	2N
⋮	⋮	⋮	⋮
…	…	…	M*N

Figure 2: The image divided into $M*N$ regions.

The number of "foreground pixels" is calculated in each region by

$$Num_{ij} = \sum_{R(j)} T_i(x,y), (x,y) \in R(j)\,, \tag{4}$$

where $R(j)$ which is the j^{th} sub-region and Num_{ij} is the number of "the foreground pixels" in $R(j)$ at frame i. The match condition is

$$Num_{ij} > MinT\,, \tag{5}$$

where $MinT$ is the minimum number of foreground pixels required for a group of pixels to represent an object. *MinT* is pre-defined according to the scenes. If one region satisfies (5), but it is not along the bolder of the scene, the background region is updates by the “foreground pixels”. If the number of regions matching (5) is more than *MR* which is pre-defined threshold for updating, the entire background is updated with the “foreground pixels” except in the regions involving targets in the previous frame. In these cases, the “foreground pixels” are actually the background pixels.

4. EXPERIMENTAL RESULTS

The proposed adaptive background modeling is applied to the UDTTS. Several video sequences captured from construction sites under different situations are tested and the experimental results of one video sequence are compared to the average value model, as shown in Figure 3. The input video sequence involves many moving targets. The proposed method generates a good background model at frame 80 while the average value model does more than 300 frames.

Main parameters of algorithm implementation: Windows XP, VC++ 6.0, CPU AMD Athlon 2.01 GHz, and memory 1.00 GB of memory. While illumination changes, the average time of background updating using the proposed method on image 360*280 is about 1ms to ensure real-time object detection. The accuracy and rate of the updating process are relied on the pre-defined parameters. So we must choose the appropriate values according to different scenes.

5. CONCLUSIONS AND FUTURE WORKS

An adaptive background model is proposed in this paper based on the average value model and applied in the UDTTS for detecting and tracking moving targets on construction sites in complex scenes. The

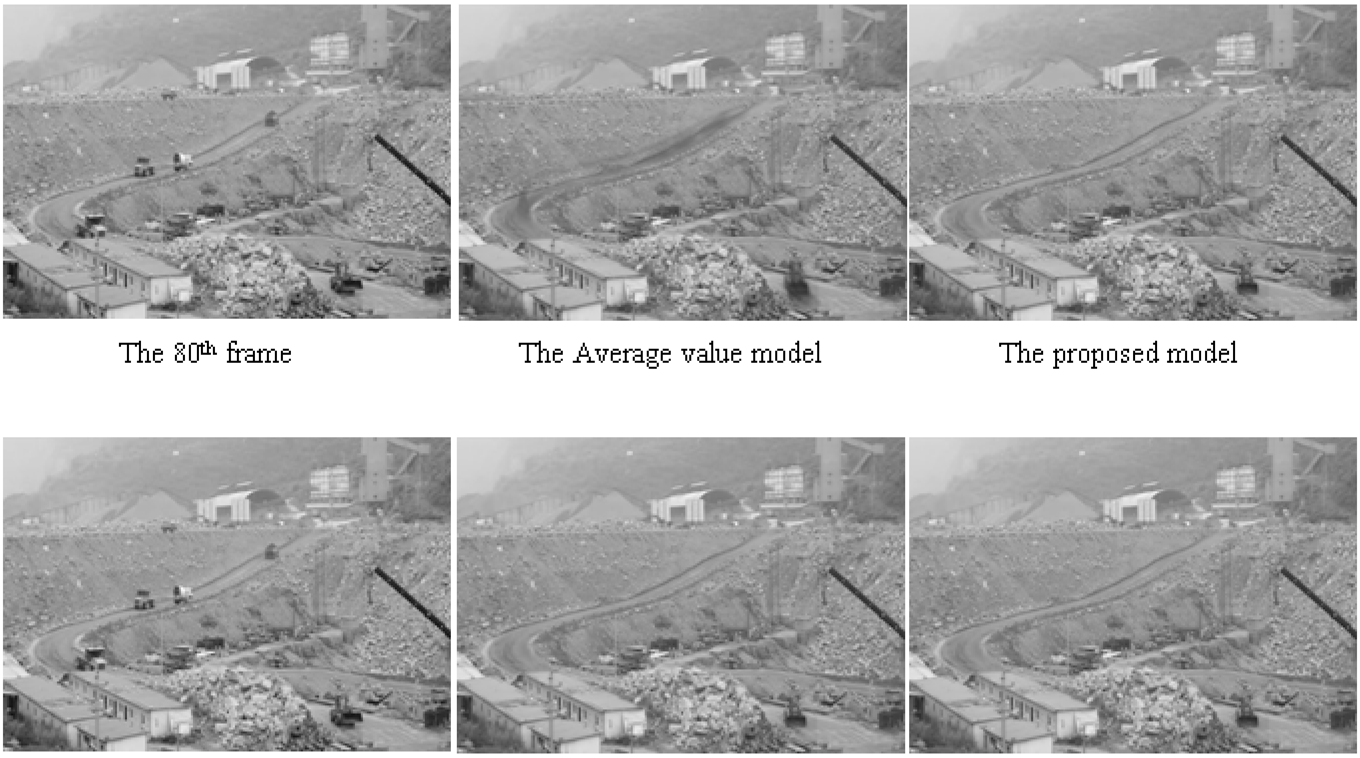

Figure 3: The experimental results of the average value model and the proposed model at frames 80 and 370.

experimental results show that the algorithm gives an accurate initial background in a situation where many moving obstacles are present. When illumination changes throughout a day, the background can be updated if necessary. Since the scene of construction sites is considered to change slowly, the updating process does not need to work at every frame and in every region. As a result, our model reduces a lot of computational cost and ensures real-time object detection.

This method is to be tested by using more complex scenes on construction sites and the efficiency of the algorithm will be improved so that more moving targets can be tracked.

REFERENCES

Aurélie Bugeau, Patrick Pérez, 2009. Detection and segmentation of moving objects in complex scenes, *Computer Vision and Image Understanding*, 113 : 459-476.

Chris Stauffer, W.E.L Grimson, Adaptive background mixture models for real-time tracking, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, USA, pp. V246–252, 1999.

Christopher Richard Wren, Ali Azarbayejani, Trevor Darrell, and Alex Paul Pentland, 1997. Real-time Tracking of the Human Body, *IEEE Transactions on Pattern Anaysis and Machine Intellidence*, 19(7) : 780- 785.

Daniel Gutchess, Miroslav Trajkovic, Eric Cohen-Solal, Damian Lyons, Anil K. Jain, A Background Model Initialization Algorithm for Video Surveillance, *Proceedings of the Eighth IEEE International Conference On Computer Vision*, Canada, pp. V733 – 740, 2001.

J. Teizer, P.A. Vela, 2009. Personnel tracking on construction sites using video cameras, *Advanced Engineering Informatics*, 23 : 452-462.

Qiaonan Shen, Xuehui An, 2008. A Target Tracking System for Applications in Hydraulic Engineering, *Tsinghua Science & Technology*, 13(1) : 343-347.

Y. Tanaka, K. Iwasada and M. Takagi, Broad Application of CALS for Sustainable Construction, *Proceedings of Civil and Environmental Engineering Conference*, Thailand, pp. V137 - 144, 1999 (Proceedings)