2009 Third International Conference on Genetic and Evolutionary Computing

Video Background Extraction Based on Improved Mode Algorithm

LI Meijin, ZHU Ying
College of Computer and Control
Guilin University of Electronic Technology
Guilin China
{eviloctal, zhuzhu}@guet.edu.cn

HUANG Jiandeng
College of Information and Technology
Guilin University of Electronic Technology
Guilin China
Hjd888@guet.edu.cn

***Abstract*--Background extraction is a fast and efficient moving object segmentation algorithm. This paper presents a new background extraction algorithm based on improved mode algorithm by using frame differential method to obtain the regions of static background. Then, this series of frames containing no moving foreground objects are counted by mode algorithm, thus improving the accuracy of the background extraction. Evaluation of the background extraction is also made by Background Approach Coefficient (BAC) which is presented in this paper. The experimental results show that the proposed method is fast in speed and accurate in background abstract.**

Keywords-background extraction; frame difference method; mode algorithm; background subtraction

Ⅰ. INTRODUCTION

Intelligent Transportation System (ITS) can be define as means of transport and communication trade to realize the modernization of Technology Traffic Management Information, such as information technology, data and communication technology, automatic control technology, Computer Processing, etc, as in [1]. With the fast development of video technology, the way of hidden coil underground have been instead of Video Detecting Technology in Intelligent Transportation System, because it's simple and convenient. Thus, currently, the Video Detecting Technology has become a popular academic branch.

Detecting moving objects from image sequences is a crucial step in Intelligent Transportation System (ITS).The efficiency and accuracy of continual image processing depends on the results of moving objects detecting. Presently, in Intelligent Transportation System (ITS) with fixed camera, the most expeditious and accuracy algorithm of moving objects detecting is background subtraction algorithm. But the expeditious and accuracy of background subtraction depends on the result of background extraction algorithm. Obviously, if there is a scene (frame) without any moving object in the image sequences, it can be use as a background frame—a "pure" background image. However, in the real world, there is hardly to get a "pure" background image. Thus, in such systems as in many other applications a critical issue in this process is extract the pure background image from the videos which include moving objects.

Recently, researchers both at home and abroad have presented some classical algorithms of background extraction, including mean algorithm, median algorithm, determination of stable interval algorithm, the detection of changes algorithm [2] and mode algorithm [3]. In addition, there are some high complexities methods have been use in background extraction algorithm too, such as textural and statistical features as in [4] and genetic algorithm as in [5]. These complexity algorithms are hard to meet the needs of real-time in Intelligent Transportation System.

Based on the research of algorithms above, this paper proposed a new background extraction algorithm: Video Background Extraction based on improved mode algorithm. First, the new algorithm applied frame difference algorithm to obtain the regions of moving objects. Then mark these regions as invalid regions in mode algorithm. At last, calculate the unmarked regions in mode algorithm. The result can be unlimited approach the real value of background pixel. Moreover, the new method can avoid the cause of inaccuracy to join in mode method. Mean algorithm, median algorithm, mode algorithm and the method this paper presented have been compared in numerous experiments. Experimental results show that the new method has better expeditious and accuracy.

Ⅱ. ALGORITHMS

A. *Frame Difference Algorithm*

Frame difference algorithm is proposed in such circumstance: if there is moving object in the video its grey level will changed significantly between two frames as in [5].

The advantage of this detecting method is its insensitive of light changing; its perfect performance in dynamic situation; its simple operation and its locating accurately of moving object. So it adapt to high request of real-time application.

In this paper, we take $I_n(x, y)$ as the value of pixel at (x, y) in frame $t=t_k$. Likewise, $I_{n+1}(x, y)$ is taken as the value of pixel at (x, y) in frame $t=t_{n+1}$.Simple difference image D(x, y) between these two frames is:

$$D(x,y)=|I_{n+1}(x,y)-I_n(x,y)|,\forall(x,y)\in[1,N]\times[1,M] \quad (1)$$

Where N×M is the image frame dimension.

Applying a suitable threshold T on D(x, y) result in a binary image that classifies all pixels into two classifications: Unchanged background and moving objects.

$$BW_n(x,y)=\begin{cases}0\equiv background & if\ D(x,y)<T \\ 1\equiv moving\ object & otherwise\end{cases} \quad (2)$$

After image binaryzation, applying the opening and closing of mathematical morphology on $BW_n(x, y)$ result in devoicing image and still save the results in $BW_n(x, y)$.

$$B\_back_n(x,y)=\begin{cases}1 & if\ BW_n(x,y)=0 \\ 0 & otherwise\end{cases} \quad (3)$$

$B\_back_n(x, y)$ is taken to mark the value of pixel at

978-0-7695-3899-0/09 $29.00 © 2009 IEEE
DOI 10.1109/WGEC.2009.51

(x, y) whether is valid or not. If the value is available ($B\_back_n(x, y) = 1$), this implies that the value of pixel at (x, y) can be joined in the calculation of mode algorithm. Otherwise, $B\_back_n(x, y)$ should be equal to 0(a flag of unavailable value), and it should avoid calculating in Formula (5). A general detection scheme based on this principle is illustrated in Figure 1 for a fixed camera surveillance application.

B. Mode Algorithm

Mode algorithm is proposed under such assumption as in [3]: In a very small space of time(e.g.100 frames), the grey level of background will change among a very small range, but the grey level of the foreground objects is vary with each car. Even though in a car, the grey level will not be the same in different parts. Based on this assumption, a conclusion could be made that the grey of foreground objects will distribute in the whole grey space by chance, and the grey of foreground will distribute around certain value. So we can extract the most frequently certain value of pixel as the background value.

This algorithm is use in 2-D image sequences, for every pixel at (x, y), the corresponding points' values in previous N frames are:

$$B_{t-N}(x,y), B_{t-N+1}(x,y), B_{t-N+2}(x,y) ..., B_{t-2}(x,y), B_{t-1}(x,y)$$

calculate the sequence of values through mode algorithm and take the result as background value of current image, the computing formula of background value is:

$$B(x,y) = \text{mod } e(I_{t-N}(x,y), I_{t-N+1}(x,y), I_{t-N+2}(x,y), ..., I_{t-2}(x,y), I_{t-1}(x,y))$$

Reference [3], where ZENG Yan just make a simple improvement in mode algorithm, and did not solve the problem in mode algorithm from the root. So, based on the research of algorithms above, this paper proposed a new method of using frame difference algorithm to get the value of pixels of unchanging background and calculate these values through mode algorithm. The new method can solve the problem in mode algorithm from root. The new algorithm can eliminate the error result from mode algorithm. And the result from the new method can approach the real background to the limit.

III. ALGORITHMS TO IMPLEMENT

A. Improved Mode Algorithm

To eliminate the deficiency of mode algorithm, this paper made improvement on mode algorithm thought the source of inaccuracy. Using frame differences method classifies the pixels into two classifications: unchanging background and moving objects, and then calculate the unchanged background pixels through mode algorithm. This method can real eliminate the deficiency in mode method. The computing formula of the new method is:

$$\begin{aligned} BG(x,y) = \text{mod} e(&B_{t-N}(x,y)\times\alpha_{t-N}, B_{t-N+1}(x,y)\times\alpha_{t-N+1}, \\ &B_{t-N+2}(x,y)\times\alpha_{t-N+2}, ..., B_{t-2}(x,y)\times\alpha_{t-2}, \\ &B_{t-1}(x,y)\times\alpha_{t-1}) \quad if \end{aligned} \tag{4}$$

$$\alpha_n = \begin{cases} 1 & if\ B\_back_n(x,y) = 1 \\ 0 & otherwise \end{cases}$$

B. Algorithm Steps

According to the discussion above, the train of thought of the new method as follows:

1) Obtain two frames by fixed intervals from the video and save them in $I_n(x, y)$ and $I_{n+1}(x, y)$.For the request of simple calculation and real-time speed, these two frames should be converted into grey images.

2) Using $I_n(x, y)$ and $I_{n+1}(x, y)$ through (1), we can get the frame difference image $D_n(x, y)$. And then, applying the opening and closing operation of mathematical morphology on D(x, y) and save the computed result by (2) in $BW_n(x, y)$.

3) According to the value of pixels in $BW_n(x, y)$, we can classify the pixels into either background or moving objects. According to (3), we can make a flag as the pixel whether it is the moving objects or background. If the flag value is 1, it implies that the pixel belongs to moving object and it's unavailable in mode calculation, otherwise its background is available. These values of flag are saved in $B\_back_n(x, y)$.

4) If n reach at the maximum we set up, the procedure goes on to step 5), or else the procedure should go on to step 1).

5) If we get here, we must have calculated all of $B_n(x, y)$ including in the video and saved as B(x, y, z), namely the values of background of all frame.

Through the steps above, the background image can be extracted accurately. The time interval between two frames is usually 0.25s. Within the limit of the video, the proper times interval can detect the moving objects integrally and do avoid the emergence have not even once of background value.

As a result of the pixels of moving objects have been removed, so even if the pixel of background value just emergence once, it can be extracted accurately by the new method. This is the greatest strength of the new method proposed in this paper.

C. Standard of Algorithm Comparison

For the purpose of performance comparison of various kinds, this paper not only use root-mean-square error (RMS error) θ as the reference standard but also propose a new coefficient-- Background Approach Coefficient(BAC) **C** as the other reference standard to judge a background image is better or not. We should find a pure background frame STBG(x, y) from the video or man-made it for the comparison. The RMS error θ and Background Approach Coefficient **C** are defined as fellow:

$$\theta = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} BG(x,y)^2 - STBG(x,y)^2 \tag{5}$$

The smaller the RMS error θ, the higher efficiency the background image. The smaller θ imply that the pixel of background extraction is distributed closer to the real background scene. Conversely, it is worse.

$$C = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} \partial \tag{6}$$

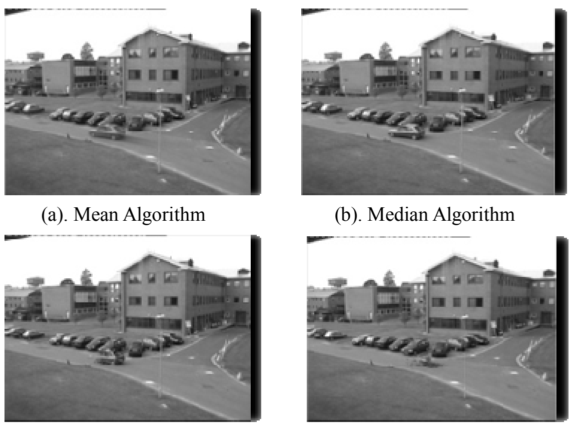
$$if \quad BG(x,y) - STBG(x,y) <= T \quad \partial = 1 \quad ohterwise \quad \partial = 0$$

The threshold T in (6) is usually setting by zero, it means that the pixels of background extraction image are the same as the real scene. Threshold T is able to set other values as you need.

The smaller the BAC **C**, the more pixels of extractive background the same as real scene, and the better the extractive background image. Conversely, the bigger the BAC **C**, the graver extractive pixels deviate from real scene.

Ⅳ. EXPERIMENT RESULT AND ANALYSIS

The algorithms in this paper are tested in PC Common architectures. Test platform is as follow: Intel Pentium Dual T2390 @ 1.86GHz CPU and 0.99GB Memory. The operating system is Windows XP SP3 and the simulation tool is Mat lab R2009a. All the data we test are from the videos we photographed ourselves and the database of currently practical application: http://ftp.pets.rdg.ac.uk/, it is order by PETS (IEEE International Workshop on Performance Evaluation of Tracking and Surveillance).This forum will call a conference yearly since 2000. Simultaneously, it will issue the database of special topic of the year.



(a). Mean Algorithm (b). Median Algorithm

(c). Mode Algorithm (d). The Proposed Method

Figure 2. The Background Extraction Base on Different Methods

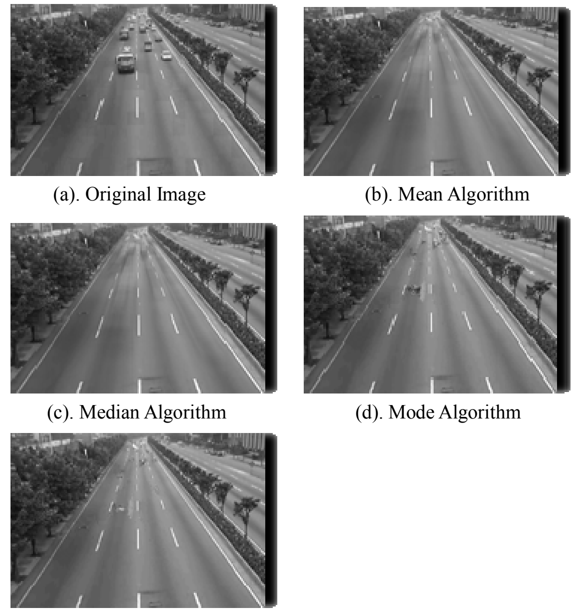
Based on the standard of comparison (5) and (6) in Chapter 3.3, we can calculate the θ and **C** corresponding to Figure 2 - 5.The values are showed in Table Ⅰ.

TABLE Ⅰ. THE VALUES OF θ AND **C** THROUGH DIFFERENT ALGORITHM

	mean	**median**	**mode**	**Proposed**
θ	*0.2599*	*0.2621*	*0.2763*	*0.2571*
C	*0.2826*	*0.2830*	*0.3916*	*0.2796*

The results of experiments under the same conditions are show in Figure 2, and the time of video is 100 frames. The extractive background of figure 2(a) resulted from mean algorithm of the 100 frames. It is obvious that there are ghosts of cars in Figure 2(a).Figure 2(b) resulted from median algorithm and its quality is worse than mean algorithm. The result of mode algorithm show in Figure 2(c) and we can make out the residue pixels of cars from the background image obviously. The result of proposed background extraction showed in Figure 2(d) and the moving car is almost eliminated in this image. The advantage of the proposed method is also proved in Table 1.The values of root-mean-square error (RMS error) θ and Background Approach Coefficient **C** of each algorithm in Table 1 show that the two coefficients of the new method are minimized.

Thus, it is proved that the proposed method in this paper is the best.



(a). Original Image (b). Mean Algorithm

(c). Median Algorithm (d). Mode Algorithm

(e). The Proposed Method in This Paper

Figure 3. Background Extraction Base on Different Methods from another Data

We have the same experiments of another data and the results showed in Figure 3. The video is photographed in one cloudy day. Figure 3(a) is an original frame in the video. The time of the video is 50 frames. The analyses of Figure 3(b)-3(e) have the same reason as above. In addition, it can be tell from that the performance of mean and median are similar, it had wide unexpecting values of background pixels. These experiments also have improved that the method proposed in this paper has better quality than the others and show in Figure 3(e).

Ⅴ. CONCLUSION

Considering the requests of real-time in Intelligent Transportation System and the advantage of background subtraction, this paper have proposed a new algorithm of Video Background Extraction based on improved mode algorithm and have done simulation experiment. The experimental results testify that the method proposed in this paper can extractive background images fast and robust. The quality of this new method will decrease linear if the pixels of real background never appearance. The reasons have been discuss in Chapter ⅢB. Therefore, what should we do in the days ahead is to find a perfect way to eliminate the pixels of foreground objects.

REFERENCES

[1] An Shi and Wang Jian, “Technique and application of city intelligent traffic management,” Beijing, Science Publishing Company, 2005.

[2] Boban Bondzulic, and Vladimir Petrovic, “Multisensor Background Extraction and Updating for Moving Target Detection,” Information Fusion, 2008 11th International Conference on June 30 2008-July 3 2008 Page(s):1 – 8.

[3] Zeng Yan and Yu Lian, "A New Background Subtraction Method for on-road Traffic," China Journal of Image and Graphics, 2008(3), pp. 593-599.

[4] JIANG Yong-lin, QU Zhen-shen, and WANG Chang-hong, "Video background extraction based on textural and statistical features," Optics and Precision Engineering, 2008(1), pp. 172-177.

[5] YANG Yu-zhen, HE Xiao-feng, CHEN Yang-zhou, and BAO Zhan-kuo, "Traffic Parameter Detection Based on Video Images," Control Engineering of China, 2008(5), pp. 349-352.

[6] Seki M,Fujiwara, and H,Sumi K.A, "robust background subtraction method for changing background," Proceeding of IEEE workshop on Applications of Computer vision, 2000, pp.207-213.

[7] WANG Jiang-tao, "Research on Object Detection, Tracking and Behavior Recognition in Video Sequences," Institutes of Technology of Nanjing, 2008, pp.15-17.