Recognition and Tracing Scheme Study of Moving Objects by Video Monitoring System

Peilong XU[1]

[1] The Growing Base for State Key Laboratory, Qingdao University, No. 308, Ningxia Road, Qingdao 266071, China.

Abstract

Objective: In this paper a recognition and tracing scheme for moving objects by video monitoring system was studied. Methods: During moving objects recognition, Multi frame sampling method was used to establish the initial background. Edges of the moving objects were drawn according to the changes of the images, and the influence factors for edges drawing were eliminated. In the end, tracing for moving objects could be realized by recognition of morphological characteristics. Results: The experiments indicate that, for field environment, the number of collected frames between 120 to 180 could get better image background. The shadows of the moving objects are the main factor which influence detection of object edges, and they could be eliminated by Shadow edge detection operator. For vehicles tracing, adjacent frame matching method could be used to reflect the time-space transformation of the vehicles. Conclusion: This scheme could realize the recognition and tracing of moving objects by video monitoring system effectively.

Keywords: *Image Recognition, Image Tracing, Video Monitoring System.*

1. Introduction

Video monitoring systems are widely used in fields like industry, transport, security and military, and are playing more and more important roles. With time progressing, the video monitoring systems have even reached scales of hundreds and thousands ways. It is impossible to rely entirely on human to monitor so many systems. In this case, all kinds of intelligent monitoring system emerge as the times require. Nowadays, the monitoring systems could realize auto alarming of the invasion objects in certain distances through all kinds of sensors and image recognition software[1]. But this is far from enough. How to realize recognition and tracing of the moving objects by video monitoring system has become an important problem to solve.

For recent years, the objects recognition and tracing techniques have made great progress. In many brands of digital cameras, face recognition function has been developed, and the face could be positioned and be focused automatically. But these techniques could just realize recognition and comparison of some preset specific shapes, and then track them. They techniques still could not realize recognition and tracing of the multi type and multi angle objects.

This study developed a intelligent monitoring system scheme according to society video monitor needs, which could recognize and real timely trace the targets. This scheme realized classification statistics and tracing of the moving objects with high recognition rate.

2. Method for Realizing the System

Traditional target segmentation algorithms are mainly based on Iteration threshold segmentation algorithm. The general progress is to detect edges and acquire difference images, then two value the data through threshold segmentation so as to highlight the parts in the images we are interested in. But this method usually needs large amounts of computation, and demands highly time complexity[2]. So it is not suitable for real time analysis for images of the moving objects. In order to ensure the real time of the system reaction, this study used background subtraction algorithm to realize quick and effective segmentation of moving objects. The key of the algorithm is that how to establish the background models.

2.1 Background subtraction algorithm

Usually, there are two kinds of background development scheme for using background subtraction algorithm. First, select two frames of adjacent images, and put the former image as background of the latter one. Detect changes of the two images using differential operation. Second, preset an unified background, and detect image changes through comparing all acquired real time images with the preset background image by differential operation. For the first scheme, it is no need to preset background, so that it is more suitable for moving monitoring facilities. But the targets volume it could monitor is related to the targets movement speeds. If the target stopped moving after coming into monitoring views, then it could not be detected by the monitoring system. This is called "Hole" phenomenon in video monitoring. For this reason, the first

scheme is not suitable for video monitoring systems with fixed camera. The second scheme could customer the problem of "Hole" phenomenon caused by the algorithm in the first scheme and is better for tracing moving objects, but this scheme need monitoring systems to establish and real time update backgrounds by themselves.
Through analysis of the above two schemes, our study try to establish a background subtraction algorithm with an uniformed background, and based on this algorithm to set up a monitoring system. Since the background quality has great influence on the targets recognition and tracing, this study used a background estimation method by multi frame difference image to set up the initial background. And through a background changing estimate strategy, the background updates only on certain conditions. To ensure the system could work normally in different light conditions and monitoring scene changes accidentally, this study used a filtering algorithm based on differential two value image processing, and this increased robustness of the algorithm effectively.

2.2 Method for background development and update

This scheme used improved background differential method to separate moving objects with background. The scheme includes two parts which are development and update of background.

2.2.1 Method for acquire background images
The background images are developed by a multi-frame subtraction images dynamic evaluation method, and the images are collected in accordance with certain time intervals. It is supposed that 3 continues images named a, b, c, and $\{B_{i,j}^t\}$, $\{O_{i,j}^t\}$ are background and moving object of the t frame. The processing procedure should be:
1) Separate the 3 frame continues images into 2 groups. The first group include frame a and b, and the second group include frame b and c. Gray scale difference subtractions between frames are operated to each pixel of the two groups, then the absolute values were preserved in $\{N_{i,j}^{-1}\}$ and $\{N_{i,j}^{+1}\}$. See in formula (1):

$$N_{i,j}^{-1} = \left| I_{i,j}^{a} - I_{i,j}^{b} \right| \quad , \quad N_{i,j}^{+1} = \left| I_{i,j}^{b} - I_{i,j}^{c} \right| \qquad (1)$$

2) Because $\{N_{i,j}^{-1}\}$ and $\{N_{i,j}^{+1}\}$ are difference values of adjacent two frames, they are highly similar and the histogram shows double peaks. The threshold T_0 of $\{N_{i,j}^{-1}\}$ calculated through OTSU method could be used as the best separation threshold of foreground and background in $\{N_{i,j}^{-1}\}$ and $\{N_{i,j}^{+1}\}$. Compare values in $\{N_{i,j}^{-1}\}$ and $\{N_{i,j}^{+1}\}$ with T_0 separately. If for any point X（i, j）, the corresponding values in $\{N_{i,j}^{-1}\}$ and $\{N_{i,j}^{+1}\}$ are all bigger than T0, then it could be judged that this point are moving in all the 3 frames, so that the point could be included into moving objects $\{O_{i,j}^t\}$ in foreground. See in formula (2):

$$O_{i,j}^{t} = \begin{cases} 255 & \text{if } D_{i,j}^{-1} > T_0 \text{ AND } D_{i,j}^{+1} > T_0 \\ 0 & \text{else} \end{cases} \qquad (2)$$

3) According to moving images $\{O_{i,j}^t\}$, all pixels valued 255 are eliminated from input b frames, and the rests are background images selected from b frames$\{B_{i,j}^t\}$. See in formula (3):

$$B_{i,j}^{t} = \begin{cases} I_{i,j}^{t} & \text{if } O_{i,j}^{t} = 0 \\ 0 & \text{else } O_{i,j}^{t} = 255 \end{cases} \qquad (3)$$

4) Deal all collected frames with method above, and supplement lost parts in background, and then a entire initial background image could be acquired.

2.2.2 Method for background image update.
After the monitoring system is used for a certain periods, the background would have some changes inevitably, for example influence of weather, light and displacement because of other factors. If the background is not updated throughout working, it must influence accuracy of recognition and tracing for moving objects[3].
This study could realize background update through method of establishing statistical models for each pixels, and calculating probabilities to judge background changes. When a certain pixel values were changing constantly during a certain times, then the grey value of the previous background would be replaced by the value of this point. Or else, the background would not be changed. In actual operation, the background judging model is established by calculating two parameters of each images—the mean value μ and variance σ. For the newly collected sample value S for a certain point （x, y）, if formula (4) were satisfied, the point would be regarded to be the new background pixel.

$$f(s) = \frac{1}{\sqrt{2\pi\sigma}} Exp(-\frac{(s-\mu)^2}{2\sigma^2}) \geq T \qquad (4)$$

T is the probability threshold value in this formula, and could be set dynamically.

2.3 Recognition and tracing of moving objects

After background was established, the background difference method was used to collect shapes of moving objects in the background. The method is to use the grey values of each pixels of the current image to subtract with those of the background. See in formula (5):

$$N_{i,j}^t = \left| I_{i,j}^t - B_{i,j}^t \right| \quad (5)$$

Because this system need to recognize moving objects, but through experiments we got to know that the moving objects produced shadows are important factors to influence system to recognize targets accurately. So that the first step to recognize objects is to eliminate influences of the shadows. Through studies of moving objects, two characters were found: one is that the difference values of the moving object shadows to the background are smaller than the those of objects themselves to the background. And the other is that the grey values of the shadows are usually smaller than those of the surrounding areas. Then the shadows could be eliminated by threshold setting. But only use threshold to eliminate moving targets would cause lost of details of moving objects. So edge detecting becomes a key problem for moving objects recognition.

2.3.1 Edge detection of the moving objects and their shadows

In detection of the objects edges, the Prewitt operator and the Sobel gradient operator are usually used. The Prewitt operator finish the calculation by counting grey differences of each pixels with their adjacent points, and then by neighbor convolution of the model and images using the horizontal and vertical vectors. This method has better effect on smoothing noises, but also has shortcoming of lower positioning precision. The Sobel gradient operator also detects image edges using the horizontal and vertical vectors, which is similar with the Prewitt operator. But the Sobel gradient operator did weighted processing on position influences of the pixels, so that it is more accurate in detection. This study used the Sobel gradient operator to detect image edges, and the calculation method is shown in formula (6).

$$E_{i,j}^{'} = \max\left(\frac{1}{4} \sum_{y=j-1}^{j+1} \gamma_y \left| I_{i-1}^t - I_{i+1}^t \right|, \frac{1}{4} \sum_{x=i-1}^{i+1} \gamma_x \left| I_{x,j+1}^t - I_{x,j-1}^t \right| \right) \quad (6)$$

Through the above calculation, and after image binaryzation, we could get the entire binary image of the whole moving object area. After this, the edges of the shadows could be get through method below: A window of 5×5 moves on the binary image, and if the central pixel grey value reached 255 in the window, then count convolution of the window subgraph with 4 sensitive one dimensional Laplasse operator. The maximum value of the 4 convolution absolute value is regarded as the basis of judging whether the central pixel is the moving object edge pixel. See in formula (7):

$$EC_{i,j}^t = \max\left\{ \left| BE_{i,j}^t \times K_p \right| : p = 1,2,3,4 \right\} \quad (7)$$

K_p is the pth convolution operator in the formula.

Through the above process, shadow areas were subtracted from the motion edge depicted binary image, influences of shadows on moving objects recognition could be eliminated, and the texture features of the moving objects are entirely conserved. This is the basic of realizing moving objects recognition and tracing.

2.3.2 Recognition method for moving objects

The study testing targets for this study is the monitoring cameras set on a school gate area. In this circumstance, most popular moving objects are vehicles and pedestrians[5]. Usually, the target recognition work through recognizing shapes, colors and textures features of the targets. But in this complicated circumstance, high misjudging rate will happen if this recognition method were used. So our scheme used targets recognition method based on morphology.

The first step is to set up a three dimensional feature library for vehicles and pedestrians. In feature depicting for vehicles and pedestrians, their length, width and shape features could be shown by morphological parameters. Because morphological features could reflect quantitative difference of shapes of vehicles and pedestrians, it is effective for recognition of the targets. According to study needs, our monitoring system selected morphological features include the following:

- Size ratio: Deal with the targets as a rectangular and get their ratio of length, width and height.
- Rectangle filling degree: Each object has its own shape filling degree. We could recognize the shape of the targets according to this.
- Projection rate: It is ratio of A size to convex polygon area size, and this feature could depict the irregular of the object edges.
- Eccentricity: This is a parameter for detecting whether the target is concentrated or not, and through calculating the eccentricity, compact degrees of the targets could be depicted.

To solve problems of highly recognition time complexity caused by vast number of data in feature database, fast linear classifier for self adaption feature selection is used to realize targets fast recognition in this study. The processing procedure of the fast linear classifier is that, select features with big differences to start initial classification for the targets, for example, large cars, small cars, pedestrians and other, and then self adaption select features fit for further classification and finish targets recognition (See in Fig 1). In occasion of too many moving targets causes high presser for the monitoring system, the system could lower its classification level to insure system recognition speed.

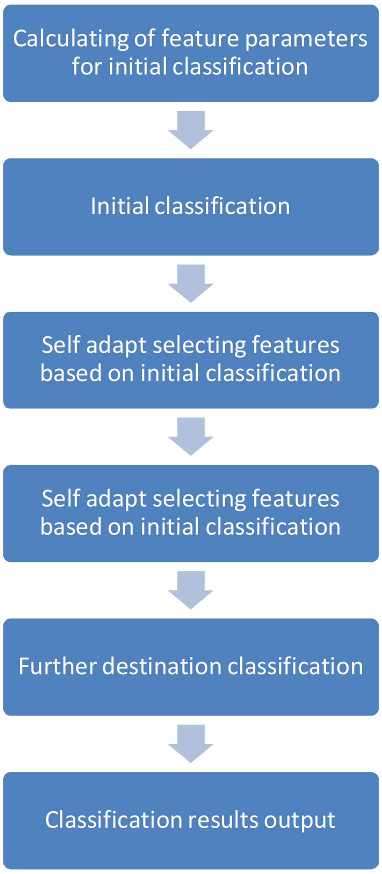

Fig 1. Target recognition chart

2.3.3 Tracing for the moving objects

Method for tracing moving objects relies on calculating matching degrees of moving targets between former and latter frames. In this study, moving objects in adjacent two frames are matched according to object position, size, average color grey value etc., to realize consistency labeling for the same targets. And the action path for the objects are recorded according to the matched position. The method is shown in formula (8):

$$\begin{aligned} C_{a,b} &= \alpha DIS_{a,b} + \beta AVE_{a,b} + \gamma AREA_{a,b} \\ DIS_{a,b} &= \sqrt{\left(\frac{X_{c,a}-X_{c,b}}{R_x}\right)^2 + \left(\frac{Y_{c,a}-Y_{c,b}}{R_y}\right)^2} \\ AVE_{a,b} &= \begin{bmatrix} |M_{R,a}-M_{R,b}| + |M_{G,a}-M_{G,b}| \\ +|M_{B,a}-M_{B,b}| \end{bmatrix} / 256 \\ AREA_{a,b} &= |(S_a - S_b)/S_a| \end{aligned} \tag{8}$$

In the formula, (R_x, R_y) symbols image resolution, α, β, γ show target position (X_c, Y_c), average color value (MR, MG,MB), and weight parameters for relative changing of target size S in function calculation. And $\alpha+\beta+\gamma=1$. The system used feature relative variation to increase adaption of target matching.

3. Result and analysis

In the experiment, IK-HD1 type 3CCD camera produced by Toshiba company is used as image collection facility. This camera owns following features: 1) Output pixel: 1920×1080; 2) Output port: digital HD-SDI（SMPTE 292M）, DVI output; 3) Manual / automatic mode white balance settings are available. Digital collection facility is MV9300HD video capture card from WOSHI company. Its mainly parameters are : Output quality: 10 bit; Compress mode: H.264; 8 video collection channels and 4 voice collection channels. is Precision T7500 type graphics workstation from Dell company is used to be data collection platform. This workstation has 4 channels memory system and NVIDIA Quadro FX3800 display chip, which has relatively high speed for graphics processing.

In software development, Microsoft Visual C++（VC++）is used and Matlab numerical calculation software is used in programming of formulas in this study, these programs are eventually compiled into VC++ procedure for call. There are 3 compiling methods used in this study: 1) Use keil compiler MCC of Matlab; 2) Use Matcom compiler; 3) Use COM Build tool of Matlab. Among these methods, method 1 is the simplest, but could not call powerful Matlab image toolbox. Method 2 is high efficient, but imperfect in supporting graphics and image functions. Method 3 is fast in program working, could be used out from the Matlab environment, and supports almost all Matlab functions. It is also perfect in supporting graphics functions, and is a MathWorks company recommended Matlab mixed programming method. Therefore, The 3rd method was used in this study to develop image processing program.

3.1 Establishment of background image

The calculation results from the multi-frame subtraction method indicate that, under the same frame sampling frequency, background quality is positively related to frame sampling time. Let *t* be sampling time, and n be sampling frames, the background collection effect is shown in Fig 2. When t equals to 2 sec, the background image C came from image A and B showed large blanks; and when t equals to 6 sec, the background image F came from D and E could basically fulfill background establishment requirements.

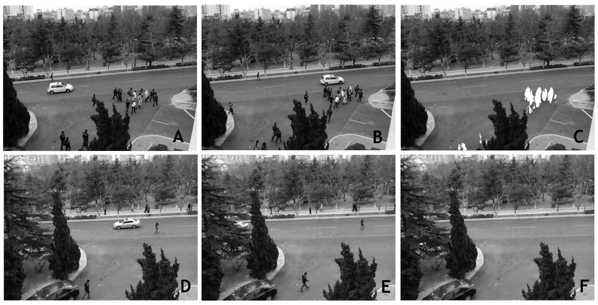

Fig 2. Initial background established under different t and n values (In the figure, C is the background established from A and B, t= 2; F is the background established from D and E, t=6)

In general transportation and traffic condition, when sampling rate is 5 per second, the bigger t value is, the more clear the background will be. When t equals to12, the background is imperfect, and when t equals to 36, the background is clear, and when t equals to 72, the background definition shows almost no difference comparing with t equals to 36. After many experiments, the results indicates that, in general transportation and traffic condition, a background clear enough for monitoring system could be established when t value reach 36. But in conditions of too many pedestrians or traffic jam, the t value needs to become bigger to establish a better background for the monitoring system to recognize and track targets effectively.

3.2 Targets recognition and tracing

In target edges detection aspects, in order to decrease pressers for calculation system, firstly the collected RGB images are converted into grayscale images, and then Sobel gradient operator is used. The target edges are detected through Matlab programming. Then the images are processed with edge function, and the input parameter of the function is the two dimensional matrix, indicator string and some numerical parameters with restricted method after imread. The edge detection procedures used Matlab are as follows:

```
f=imread('1.jpg');
f=rgb2gray(f);% convert to Gray scale map
f=im2double(f);% The funtion im2double, the value is
normalized to 0 ～ 1
% Use vertical Sobcl operator, and select threshold
grates automatically
[VSFAT Threshold]=edge(f, 'sobel','vertical'); % Edge
detection
figure,imshow(f),title(' initial image， ');% Show the
initial image
figure,imshow(VSFAT),title( ' Vertical image edge
detection ');
% Show the edge detection image
% Using the horizontal and vertical Sobel operator to
select threshold automatically.
SFST=edge(f,'sobel',Threshold);
figure,imshow(SFST),title(' Horizontal and vertical
image edge detection ');
% Show the edge detection image
% Use specified angles of 45 degrees Sobel operator
filter to specified the threshold
s45=[-2 -1 0;-1 0 1;0 1 2];
SFST45=imfilter(f,s45,'replicate');% Function ： to filter
arrays of arbitrary types or multidimensional image
SFST45=SFST45>=Threshold;
figure,imshow(SFST45),title(' angles of 45 degrees edge
detection') ;
% Show the edge detection image
```

The background elimination effect for moving object after edge detection is shown in fig 3.

Fig 3. The background elimination effect for moving object after edge detection

During target recognition and tracing, the system frame rate is 15/ second, and image resolution is 1920×1080 pixels. In the actual experiment, videos stored with local hard disk were used to test the function of recognition, and the experiment results were calculated artificial counting and statistical classification results comparison. The results got from testing 1200 second images recognition is shown in table 1.

Table 1: Results of recognition rate experiments

Name of moving object	*Recognition rate （%）*	*Average recognition time （ms）*
Pedestrians	91.6	27.11
Small cars	87.1	29.52
Big buses	93.9	25.70
Small buses	92.8	25.87
Trucks	89.8	26.64
Bicycles	90.5	27.46

The experiment shows that this system could finish targets recognition under frame rate of 15 / second, and recognition rates are higher than 87%. It also could finish target tracing according to front and rear frames matching degree, but the recognition of small cars are relatively lower, because that the rectangular degree of the cars are usually difficult to control, and the system often mistake them with small buses. In future studies, this system could be further developed by improving the target morphological database.

4. Conclusion

This study designed a set of moving objects recognition and tracing methods according to unique requests of monitoring system. And researchers also developed an effective scheme in background establishment and moving objects shadows elimination regarding to the actual conditions. This study proved through experiments that this scheme could fulfill video monitoring system needs for moving objects recognition and tracing. And the recognition rate and tracing speed both reached the design standards.

References

[1] Weihua Liu. An image restoration algorithm based on image fusion, International Review on Computers and Software, 2012,Vol.7, n.3,pp. 1245-1249.
[2] Abo-Eleneen Z.A., Abdel-Azim Gamil. An improved image segmentation algorithm based on MET method, International Journal of Computer Science Issues, Vol.9 n.5-3, 2012, pp.346-351.
[3] Jun Sun, Yan Wang, Xiaohong Wu.A New Image Segmentation Algorithm and Its Application in Lettuce Object Segmentation, TELKOMNIKA, Vol.10, n.3,2012,pp. 227-563.
[4] Kondapalli Varaprasad S., Chiranjeevi Manike, Mishra Kundan Kumar, Tanuja Kdbs. Image processing and analysis for DTMRI, International Journal of Computer Science Issues, Vol.9 No. 1-1,2012, pp.266-272.
[5] Ren Mingwu, Yang Jingyu, Sun Han. Tracing boundary contours in a binary image, Image and Vision Computing, Vol.20 No.2,2002,pp:125-131.
[6] Kang Lie, Zhong Sheng, Wang Fang. A new contour tracing method in a binary image, 2011 International Conference on Multimedia Technology, ICMT2011,July 26, 2011 - July 28, 2011,Hangzhou,China,2011,pp.6183-6186.

Peilong XU, born in 1977. The author has achieved Master degree from Tongji University of Shanghai in 2007. He is currently an engineer of computer science in the State Key Laboratory, Qingdao University. The author's research interests include software engineering, image processing, and spectral analysis. Recently, he have published paper named Design and Implementation of Landscape System for East and West Huashi Street in Beijing Based on Virtual Reality Technology on journal Applied Mechanics and Materials（Trans Tech Publications Inc, Switzerland）etc.