A Real-time Detection for Traffic Surveillance Video Shaking

Yaoyao Niu
College of Information Engineering
Qingdao University
Qingdao, China
e-mail: 15165298304@163.com

Zhenkuan Pan
College of Information Engineering
Qingdao University
Qingdao, China
e-mail: zkpan@qdu.edu.cn

Danfeng Hong
College of Information Engineering
Qingdao University
Qingdao, China
e-mail: hongdanfeng1989@gmail.com

Xin Wu
School of Information and Electronics
Beijing Instituteof Technology
Beijing, China
e-mail: 040251522wuxin@163.com

Abstract—Shaking happens to the traffic surveillance video inevitably, which can lead to the degradation of monitor system performance. A hierarchical gray project the correlation algorithm (HGPC) is proposed to solve the problem of the shaking in the traffic surveillance video. First, global hierarchical gray project correlation algorithm (GHGPC) is used to detect the video roughly. Then, in order to further improve the detection accuracy, we used the local hierarchical gray project correlation algorithm (LHGPC) to deal with the results from the rough matching. The experiment results on the video library HGTVL database show that the accuracy of the proposed algorithm is 91.4%, which is superior to the previous methods. In addition, the detection time is only 47ms, which meets the real-time requirement. Therefore, it is proved to be effective and feasible to dispose the video shaking.

Keywords-video shaking; hierarchical; gray projection; correlation coefficient; real-time

I. INTRODUCTION

In modern society, it is necessary to install traffic video monitoring for reducing the traffic accident. Therefore, the video image with high quality is essential in traffic video monitoring systems. However, it is difficult to obtain the video image with high quality in the natural environment. There are many factors which influences the normal detection, such as image breezing, abnormal brightness [1], video image shaking, etc. Hereinto, the video image shaking is one of the crucial problems because of doing harm to the quality of video image to a large extent, which needs to be solved immediately. From the different perspective to tackle the shaking, it can be mainly classified, for example, frequency domain method [2-4], spatial domain method. Li et al [5] used the image center distance and high-pass filter combination evaluation model to cope with Video image shaking. However, their model is so complex that they cannot meet the real-time requirement. Huang et al [6] proposed a method of the gradient and HSV feature from coke fuzzy image quality evaluation model. They combined color information with gradient to obtain a desirable result, but the problem of real-time still remains to be solved. Mohsen et al [7-8] designed a method of image entropy and gray value of evaluation function to assess the image quality. Their method had got high detection accuracy. Hence, there is a room to improve the performance of method for the shaking in the video.

In normal, continuous images in a video are linked smoothly. When the video shakes, continuous images' correlation changes a lot. Generally speaking, the reason why the traffic surveillance video shakes is that the camera did not be fixed strong enough. The correlation between the pre-image and the next-image is essence to judge whether the traffic surveillance video shakes. If the correlation is large, the contents change between the pre-image and the next-image is relatively small. If the correlation is small, the contents change between the pre-image and the next-image is relatively large and the traffic surveillance video shakes. Through referring to the literature and data, we found that the study of the video image shaking is not many. But the detection of the video image shaking has very important significance for traffic video monitoring system. The detection of the video image shaking can enhance the credibility of the traffic video monitoring system. In the studies of the detection about video image shaking, Peng Yi [9], etc. proposed feature points matching in consecutive images (FM) to detect video image shaking. Wolberg [10] proposed sportive block motion analysis method (SBM) to analyze the motion of pixel block. The two algorithms (FM and SBM) analyze two-dimensional image itself. Its test time is much longer. Its detection accuracy is not high. It can't be better applied to the engineering. As the global video image shakes, the information of the local video image changes largely. When there is moving object in normal video, the local video image changes like the video image shaking. In view of the problem of the video shaking, we propose a hierarchical gray projection correlation coefficient algorithm (HGPC), we use global hierarchical gray project the correlation algorithm to detect the video image firstly.

Then we use local hierarchical gray project the correlation algorithm to do the further detection for getting higher test accuracy. After we did HGPC detection, the algorithm accuracy reached 91.4%, which is higher than that of the traditional test methods. And the test time is only 47 ms, which meet the real-time requirements.

II. HGPC ALGORITHM DESIGN

A. Flow Chart of HGPC Algorithm Design

In view of the video shaking in the traffic surveillance video, we propose a hierarchical gray projection correlation coefficient algorithm (HGPC). Fig .1 shows the flow chart of HGPC briefly.
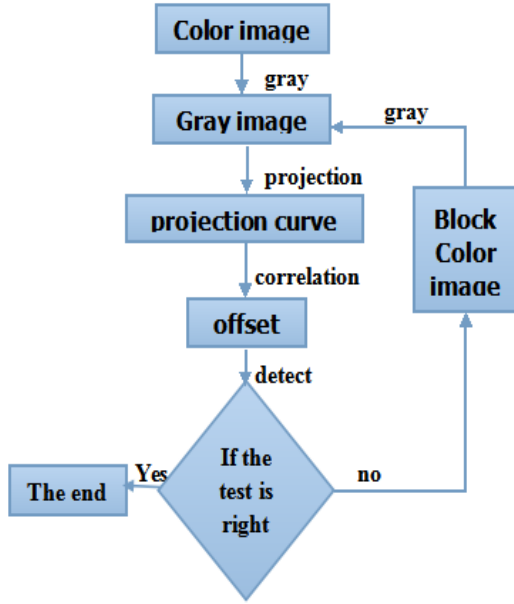

Figure 1. Flow chart of HGPC

B. Basic principle

1) Gray image convertion

In order to speed up the processing speed, it is important and meaningful to convert color image to gray image. In [1], it has been proved that gray image is easy to transplant to the color image. The conversion formula is shown as follows:

$$Gray(i,j) = 0.299 * R(i,j) + 0.578 * G(i,j) + 0.114 * B(i,j) \quad (1)$$

where $R(i,j)$ is the Brightness values of R channel. $G(i,j)$ is the Brightness values of G channel. $B(i,j)$ is the Brightness values of B channel. $Gray(i,j)$ is the gray value.

2) Projection method

Gray image projection includes the row's projection and the column's projection. Projection maps the two-dimension image information to one-dimensional curve information. Projection mapping method can be computed by

$$Row(i) = \sum_{i=1}^{A} f(i,:) \quad (2)$$

$$Col(j) = \sum_{j=1}^{B} f(:,j) \quad (3)$$

where A and B stand for the row and column for the original image, respectively. $Row(i)$ is the sum of gray value of i row for the OI. $Col(j)$ is the gray value of j column of the OI. $f(i,:)$ is the gray value of i-th row, and $f(:,j)$ is j-th column for the OI.

3) The Correlation

After getting the gray image projection curve, we utilize the correlation to further judge whether the image is shaking. The correlation can be described as follows:

$$G(w) = \sum_{j=(B/4)+1}^{3B/4} \left[C(j+w) - Cp(j) \right] * \left[C(j+w) - Cp(j) \right] \quad (4)$$

$$-j+1 \le w \le \frac{1}{2}B - j + 1$$

$$H(s) = \sum_{i=(A/4)+1}^{3A/4} \left[R(i+s) - Rp(i) \right] * \left[R(i+s) - Rp(i) \right] \quad (5)$$

$$-i+1 \le s \le \frac{1}{2}A - i + 1$$

where $C(j)$ and $Cp(j)$ are corresponding gray projection value of pre-image (it is expressed as the image of the first frame) and next-image (it is expressed as the image of the second frame) in column j. $R(i)$ and $Rp(i)$ are corresponding gray projection value of pre-image and next-image in row i. The values of w and s are the translation variation, which are usually equal to 5. Also, $G(w)$ is the Correlation value of column, and $H(s)$ is the Correlation value of row.

C. Algorithm Design forHGPC

1) Global Coarse Detection

We firstly get a set of images from the video library. These images consist of two kinds of images, the pre-image and the next-image. The pre-image and the next-image come from one video, and the time interval is 5ms. Then we turn the two color images to gray images. According to the principle of gray image projection, we can calculate the projection value of gray image, and map relevant projection value curve. Next, we fixed the column projection curve of the pre-image. Then we shift the next-image's column projection curve. At the same time, we computed the relative value of two column projection curve, where we can get the maximum correlation value. Finally, we calculate column projection offset vector at the maximum correlation value. At the time, $G(w)$ takes the minimum, Gmin = g, the next-image shifts g column to the pre-image relatively. $g = 0$ suggests that the video image has no shaking in a horizontal direction. If g > 0, it shows that the next-image shift to right in a horizontal direction, and if g < 0, it can suggest the next-image shift to left in a horizontal direction. We can obtain row projection offset vector by the same principle. The video image is stable when column projection offset vector and row projection offset vector are both 0, vice versa.

2) Local Fine Detection

Global crude shaking anomaly detection is suitable for the video background image which has only simple information. Traffic monitoring videos of Hisense network technology co., LTD. Guiyang library have a large of relative complex background images, such as retrograde traffic flow, etc. The test accuracy of global hierarchical gray project the correlation algorithm is not so desirable

and there is a larger room for improvement. In order to further improve the detection accuracy, we propose local hierarchical gray project the correlation algorithm detection on the basis of global coarse detection. By using different ways of partitioning, we divide the pre-image and the next-image into small pieces of different size. Then we detect the corresponding blocks by gray project the correlation algorithm. At the same time, we analysis offset vector based on block, and select the optimal way of blocking by comparing the results of different partition way. Fig .2 is different ways of image partitioning.
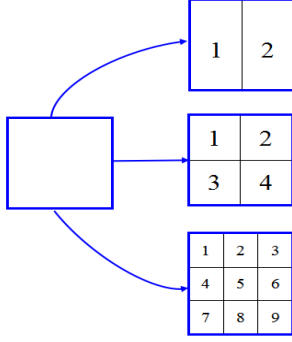

Figure 2.   The different ways of image partitioning

III.   EXPERIMENT RESULTS AND ANALYSIS

A.   Introduction of Video Library

The test videos come from Hisense network technology co., LTD. Guiyang intelligent traffic monitoring video library (HGTVL), which has 570 videos. There are 326 normal videos, 11 shaking videos, 233 other image exception videos. These videos were collected by standard clear camera and high clear camera. We detect the normal videos and the shaking videos by the proposed HGPC algorithm.

B.   The Experiment Result

1)   Global Coarse Detection results

Figure 3.   Normal video 1

TABLE I.   DETECTION RESULT OF NORMAL VIDEO 1

Pre-image	Next-image	A ×B	Row offset	Column offset
Normal 1	Normal 2	1080×1920	0	0

Fig .3 is a set of image, which is selected from normal traffic surveillance video. The value of row offset vector

and column offset vector are both 0, which means the video image is stable. Hence, we can obtain the results from Table 1 that the detection results coincide with the reality.

Figure 4.   Shaking video 1

TABLE II.   DETECTION RESULT OF ABNORMAL VIDEO 1

Pre-image	Next-image	A ×B	Row offset	Column offset
Normal 1	Shaking 2	720×1280	-1	-2

Fig .4 is a set of image, which selected from abnormal traffic surveillance video. The value of row offset vector and the column offset vector is -1 and -2, which means the video shift to the upper left direction. The video shakes. Hence, we can see from Table 2 that the detection result is correct.

Figure 5.   Normal video 2

TABLE III.   DETECTION RESULT OF NORMAL VIDEO 2

Pre-image	Next-image	A ×B	Row offset	Column offset
Normal 3	Normal 4	1080×1920	-127	-20

Fig .5 is a set of image, which is selected from normal traffic surveillance video. The row offset vector and the column offset vector are - 127 and -20, which means the video shift to the upper left direction. But this set of image is selected from normal traffic surveillance video. Hence, we can see from Table 3 that the detection result is not correct.

The global hierarchical gray project the correlation algorithm detects the video library. The detection accuracy

is 63.1%. Although this method based on global has high robustness, but low accuracy. It does not meet the actual demand.

2) Local Fine Detection results

When the video shakes, blocks of the video shake. When some blocks of the video shake, the video does not always shake. In order to further improve the accuracy of the detection, we use local hierarchical gray project the correlation algorithm to detect the wrong detection videos in global coarse detection. We detect Fig .4 and Fig .5 by HGPC. Experiment results are as follows:

a) The result of block with being divided the image into symmetrical two parts(1/2)

TABLE IV.   DETECTION RESULT OF ABNORMAL VIDEO 1 AND NORMAL VIDEO 2 IN 1/2

Figure	Block name	row offset	column offset
4	1	-1	-2
	2	-1	-2
5	1	-53	-87
	2	-62	-165

About Fig. 4, the row offset vector and the column offset vector of 1 block and 2 blocks are -1 and -2.Which means the video shift to the upper left direction. The video shakes. Hence, we can see from Table 4 that the detection result is correct. About Fig .5, the row offset vector and the column offset vector of 1 block and 2 block are not 0. It is shown that the video shift to the upper left direction. However, this video is normal. Hence, we can see from Table 4 that the detection result is not correct.

b) The result of block with being divided the image into four parts of the same size(1/4)

TABLE V.   DETECTION RESULT OF ABNORMAL VIDEO 1 AND NORMAL VIDEO 2 IN 1/4

Figure	Block name	row offset	column offset
4	1	-1	-2
	2	-1	-2
	3	-1	-2
	4	-1	-2
5	1	0	0
	2	+1	-1
	3	0	0
	4	-133	-178

About Fig .4,the row offset vector and the column offset vector of 1, 2 ,3 and 4 block are -1 and -2, which means the video shifts to the upper left direction. The video shakes. Hence, we can see from Table 5 that the detection result is correct. About Fig .5, the row offset vector and the column offset vector of 1 block and 3 block are both 0.The row offset vector and the column offset vector of 2 block and 4 block are not 0, which means the video is moving objects in local, but the whole video is stable, not shaking. Hence, we can see from Table 5 that the detection result is correct.

c) The result of block with being divided the image into nine parts of the same size(1/9)

About Fig .4,the row offset vector and the column offset vector of 1, 2 ,3 ,4 ,5,7and 9 block are -1 and -2.The

row offset vector and the column offset vector of 6 block are 37 and 53, which means that the video shakes. And the video is moving objects in local. Therefore, we can see from Table 6 that the detection result is correct. The row offset vector and the column offset vector of 1 block, 2block, 3 block, 4 block and 7 block are both 0. About Fig .5, the row offset vector and the column offset vector of 5 block ,6 block,8 block and 9 block are not 0, which means the video is moving objects for local region, but the whole video keeps stable. Hence, as shown in Table 6, the detection result is correct.

TABLE VI.   DETECTION RESULT OF ABNORMAL VIDEO 1 AND NORMAL VIDEO 2 IN 1/9

Figure	Block name	row offset	column offset
4	1	-1	-2
	2	-1	-2
	3	-1	-2
	4	-1	-2
	5	-1	-2
	6	37	53
	7	-1	-2
	8	24	47
	9	-1	-2
5	1	0	0
	2	0	0
	3	0	0
	4	0	0
	5	-8	-35
	6	-1	-15
	7	0	0
	8	-43	-127
	9	-57	-162

The video image shaking means the image shifts to the certain direction on the whole. If there is the moving object in the video, the image can be different from the original image. In this case, the detection accuracy can be influenced. Therefore, we used the idea of the block to improve the detection accuracy.

TABLE VII.   THE TEST ACCURACY AND TEST TIME OF DIFFERENT BLOCKED MODES

Partitioning	Test accuracy	Test time
1	63.1%	40ms
1/2	70.1%	44ms
1/4	91.4%	47ms
1/9	82.9%	63ms

However, the video image is divided into too many blocks; the test time will be long. Thereby, it is important to select an optimal block number. Form the result in the Table 7 that the test accuracy of 1/4 partitioning is 91.4% and the test time is 47ms. Taking the test accuracy and test time into comprehensive consideration, the 1/4 partitioning is better than the others.

TABLE VIII.   THE TEST ACCURACY AND TEST TIME OF DIFFERENT METHODS

algorithm name	test accuracy	Test time
SBM	63.2%	480ms
FM	81%	350ms
GPC	63.1%	40ms
HGPC	91.4%	47ms

SBM is proposed by Peng et al [9]. FM is proposed by Wolberg [10]. GPC is used in global coarse detection in

this paper. HGPC is proposed in this paper to detect the video shaking. From Table 8 we can see that the test time of GPC is 40ms, but the test accuracy is 63.1%, which is too low. It can detect the video shaking. But it does not meet the practical needs. The test accuracy of SBM and FM are higher than that of GPC, but they just analyze the two-dimensional images which is easy to neglect the change in the horizontal and vertical orientation. In addition, only using the two-dimension information cannot obtain the superior recognition result. The HGPC firstly turn the two-dimensional images to one-dimensional information. Then, both one-dimensional and two-dimension information are used via the hierarchical idea. So the test time of SBM and FM are much longer than GPC and HGPC. They can't be better applied to the engineering. The test time of HGPC is 47ms, which is much shorter than SBM and FM. The test accuracy is 91.4%, which is higher than GPC, SBM and FM. HGPC not only has short test time and high detection accuracy, but can detect the video which has relative complex background. Therefore, HGPC has also robustness.

IV. CONCLUSION

After about a year's research and test, HGPC detects the whole videos from the video library HGTVL. The detection accuracy is 91.4%. It is higher than traditional algorithms. The test time is only 47ms, which is so fast that it can meet the real-time requirement. HGPC has high robustness. It can be better applied to the engineering. HGPC should be detected by more traffic surveillance videos for further study. We can use HGPC to detect the traffic surveillance video to enhance the reliability of the monitoring system. The detection can provide better service for the related department of monitoring system.

REFERENCES

[1] X. Dong, X.O. Chen, "Color image quality for web browsing automatically adjust strategies," Journal of computer-aided design and image study, 14(6), 2002, pp. 535-541.

[2] D.F. Hong, W.B. Wei, Z.K. Pan, X. Wu, "Image denosing based on Curvelet transform and non-local TV model," Journal of Qingdao University (Natural Science Edition), 27(1), 2014, pp.59-65.

[3] D.F. Hong, Q. Ling, J.J. Peng, "ICA Blind Source Separation Algorithm Based on Wavelet Transform," Journal of Naval Aeronautical and Astronautical University, 28(3), 2013, pp.237-240.

[4] D.F. Hong, Z.K. Pan, X. Wu, "Improved differential box counting with muti-scale and multi-direction: A new palmprint recognition method," Optik-International Journal of Light Electron Optics, 125(15), 2014, pp.4154-4160.

[5] F. Li, Research and realization of the digital image quality detection technology, 2005, Jiangsu University.

[6] L.H. Huang, Z.H. Chen, X.N. Peng, "Based on the gradient and the HVS characteristics from focal fuzzy image quality evaluation," Computer application research, 27(2), 2010, pp.178-183.

[7] M. E. Moghaddam, "A mathematical model to estimate out of focus blur," Proceedings of the 5th International Symposium on image and Signal Processing and Analysis. Proc. 2007, ISPA07:pp.278-281.

[8] W.X. Zhang, B.F. Zhou, "A color image based on gradient domain turned gray image method," Journal of imaging techniques, 3, 2007, pp.20 -22.

[9] Y. Peng , Q.X. Ye, J. Huang, "A complete video content stability algorithm," Chinese journal of image and graphics, 9, 2010, pp.1384-1390.

[10] G. Wolberg, Digital Image Warping, IEEE Computer Society Press, 1990.