Region-based Moving Shadow Detection Using Watershed Algorithm

Jin Gao[1], Jiangyan Dai[2], Peng Zhang[3]

[1,2,3]*School of Computer Engineering, Weifang University, Weifang, China, 261061*

[1]gaojin2008_ny@126.com

[2]daijyan@163.com

[3]sduzhangp@163.com

Abstract— **Moving shadow detection is an important task in computer vision, with applications several fields, such as surveillance, video conference, visual tracking and object recognition. In this paper, we present a Spatio-Temporal based Moving Shadow Detection (STMSD) method, The main idea is: according to the gradient change of current image, we utilize the watershed algorithm to achieve adaptive segmentation foreground sub-regions; Then, the changes in texture of adjacent video frames and current video frame and background are extracted, and each sub-region is classified in terms of texture feature to obtain the final moving shadow and target area. Extensive experiments demonstrate that STMSD is superior to some well-known methods especially pixel-based methods. Especially, our method exhibits much better performance compared with fixed block method, which can maintain the texture consistency in one region adequately.**

***Keywords*— Moving shadow detection, Spatio-Temporal, region-based detection, Pixel-based detection, watershed algorithm**

I. Introduction

Shadows often share the same movement characteristic and have similar intensity change with that of moving objects, which will give rise to detect shadows as parts of moving objects. Therefore, in computer vision, the most important thing at the moving cast shadow detection is to improve accuracy of moving object detection.

In the past decade, most of the research on moving shadow detection methods are based on pixels [1-7]. Pixel-based methods are presented a comprehensive review and are given the advantages and disadvantages about each method by Sanin et al. [1]. Cucchiara et al. [2] pointed a shadow detection method in HSV colour space and detected shadows with colour information. HSV colour space and human visual system can perceive the same colour, and have good ability to distinguish the shadows. In order to improve shadow detection accuracy, two kinds of spatial processing operations were recommended by Song and Tai [3]. Wang et al. [4] used online sub-scene shadow modelling through Gaussian functions and adapted to detect moving shadows. A framework of learning the most relevant features was building at super-pixel level by Khan et al. [5], which detected shadows automatically and in supervised manner along the object boundaries. These methods achieved better performance compared with some state-of-the-art methods.

However, the effect of noise and uncertain factors on pixel-based methods will reduce the shadow detection accuracy easily. In order to solve this problem, many researchers utilized local spatial correlation of pixels and a series of morphological operations to refine shadow detection results further. But this method will increase computational complexity inevitably. Lately, region-based methods have been received widely attention in moving shadow detection areas. In order to distinguish shadows from suspected foreground patches, Bullkich et al. [8] employed Matching by Tone Mapping as non-linear tone mapping existed between shadow patches and that of corresponding background patches.

Inspired by the problem and considering the correlation between adjacent video frames, a region-based moving shadow detection method is proposed in this paper, it has better robustness than pixel-based methods. According to the gradient change of current image, we utilize the watershed algorithm to achieve adaptive segmentation foreground sub-regions; Then, the changes in texture of adjacent video frames and current video frame and background are extracted, and each sub-region is classified in terms of texture feature to obtain the final moving shadow and target area. Extensive experiments demonstrate that STMSD is superior to some well-known methods especially pixel-based methods. Especially, our method exhibits much better performance compared with fixed block method, which can maintain the texture consistency in one region adequately.

The reminding parts of this paper are organized as follows. Section Ⅱ and section Ⅲ describe the proposed method in detail. Experiments and analysis are given in Section Ⅳ, and Section Ⅴ is the conclusions.

II. Based on the Foreground Image Segmentation of Watershed

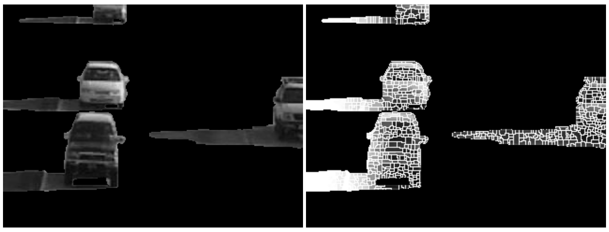
In the concept of geography [9, 10], watershed is referring to the ridge, there are different flows of river systems in the region on both sides of the ridge. Catchment basin means the geographic area where the water enter into the river or reservoir. In digital image processing, watershed is based on the three-dimensional image, and two are the coordinate, another is gray level.

According to the concept of geography, there are three points in the image: (1) Locality minimum point; (2) When a drop of water placed on the position of a point, the water must

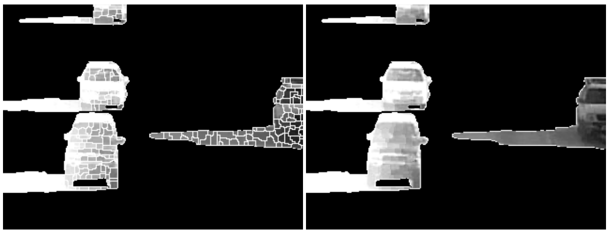
978-1-5090-3071-2/16 $31.00 © 2016 IEEE
DOI 10.1109/IS3C.2016.215

fall on a single minimum point. (3) When the water is at some point on the position, it will flow several such minimum point with an equal probability. Then, for the minimum value of a particular region, the set of points which satisfy the condition (2) will be called the minimum "watershed" or "catchment basin", and the set of points which satisfy the condition (3) and form together will be named the "watershed." In the process of digital image, the purpose of the watershed algorithm [11] is to find the watershed and the watershed line in the image.

Given the foreground image F, defined it as R. Figure 1 shows the segmentation results of he watershed algorithm F foreground image, Figure 1 (a) is a foreground image F. Before using the watershed algorithm, we usually use a gradient magnitude for image pre-processing. STMSD method utilize the Sobel to calculate the foreground image F corresponding gradient image, then break up the algorithm gradient image through the watershed, the result is shown in the Figure 1 (b). Obviously, there is a serious over-segmentation in the segmentation result, which was mainly due to local noise or gradients caused by irregularities. To solve this problem, use the median filtering for smoothing operation process before the watershed algorithm, the result is shown in Figure 1 (c).



(a) Foreground (b) Gradient



(c) Filter_gradie (d) Remove_line

Figure 1 The Watershed Algorithm Results

It can be seen from the Figure 1 (c), using the watershed segmentation algorithm, the foreground region R is divided into several sub-region except watershed line. We have to remove the dividing line so that all pixels in the image are divided into corresponding regions. In this chapter, according to the neighborhood pixel of pixel watershed area, judge which watershed pixels belongs. Watershed line is assumed that the pixel p, its neighbor pixel is shown in Figure 2 (a), after performing a watershed algorithm, we obtain two sub-areas R_1 and R_2, the result is shown in Figure 2 (b). Obviously, the number of pixels belonging to the sub-region R_1 is more than that in the sub-regions R_2. Therefore, the pixel p is divided into sub-areas R_1 finally. The result of the removal of the dividing line is presented in Figure 1 (d), it can be seen that the foreground region R is divided into n sub-region R_i, namely, i = 1, 2, ..., n.

44	45	30
50	*p*	47
42	99	87

(a)

1	1	1
1	*p*	1
1	2	2

(b)

Figure 2 Watershed algorithm

III. MOVING SHADOW DETECTION OF COMBINING THE SPACE AND TIME

After the utilizing the adaptive area division based on watershed algorithm, obtaining several sub-region R_i. Through the analysis of the shadow model and the shadow attribute, the texture in the shaded area is smoother than that in the target region. Hence, we classify each sub-region R_i according to the gradient of the sub-region, changes in texture between adjacent video frames, the current frame and the background respectively, realizing the shadow detection combined with space and time.

A Variation of the Gradient amplitude

Typically, compared with the background, there is a huge difference between the gradient information of the target pixel and the background pixel, the gray level change is more in the target. Therefore, we can detect the shadow through judging the gradient foreground of the sub-region. In this section, we get the gradient image using the foreground image based on the Sobel operator. After that, calculate the gradient amplitude G_i in each sub-region amplitude R_i. According to the following rules, determining the shadow mask image:

$$M_i^G = \begin{cases} 1 & G_i \ge T_G \\ 0 & \text{others} \end{cases} \qquad (1)$$

$$T_G = \varepsilon_G \cdot \text{median}\{G_i\} \qquad (2)$$

G_i represents the gradient magnitude of the sub-region R_i, $M_i^G = 1$ indicates the i-th region, R_i means shadow, otherwise it belongs to the target, T_G is the threshold value according to the gradient magnitude of all the sub- region. When a shadow mask is obtained by the magnitude of gradient of the image, only consider the change of the video information about the region of the current frame neutrons and not utilize the information provided by the video itself. In the following, we will detect the moving shadows through the texture of space-time combination.

B The texture similarity of space-time combination

When we describe the change of texture based on space-time combination. On the one hand, the texture of the similarity between the current video frame airspace and background is considered; On the other hand, the texture change in the time domain adjacent between video frames is considered. This section describes the similarity of the prospect area by the LBP, for the texture change S_i of the i-th sub-region, calculate as follows:

$$S_i = (1-\alpha)\cdot S_i^b + \alpha \cdot S_i^w \qquad (3)$$

$$S_i^b = \rho\left(LBP_t^{R_i^F}, LBP_{t-1}^{R_i^F}\right) \qquad (4)$$

$$S_i^w = \rho\left(LBP_t^{R_i^F}, LBP_t^{R_i^B}\right) \qquad (5)$$

$LBP_t^{R_i^F}$ and $LBP_t^{R_i^B}$ represent the R_i^F of i-th region by the calculation and the background in the corresponding sub-region R_i^B. $LBP_{t-1}^{R_i^F}$ is the first sub-region i R_i calculated in local binary patterns at the time of the t-1. $\rho(\cdot)$ represents the similarity operation calculated from the two local binary patterns, S_i^b indicates the similarity of the adjacent sub-area and S_i^w represents the similarity between the background and foreground sub-region respectively. α means the learning rate, which is related to the movement speed frame rate and motion video sequences of related goals. When α equals 0, S_i is only calculated by the similarity between the sub-region adjacent frames; When α equals 1, S_i is determined by calculating the similarity between the foreground and background of sub-region. In order to achieve a combination of space-time characteristics, $\alpha \in (0,1)$. From the equation (3), the bigger of the S_i value, the more similar of the foreground and the i-th region R_i in the previous frame corresponding to the background texture regions and sub-sub-area.

Hence, according to the change of the texture similarity, the shadow mask image is following:

$$M_i^s = \begin{cases} 1 & S_i \geq T_S \\ 0 & \text{others} \end{cases} \qquad (6)$$

$$T_S = \varepsilon_S \cdot \text{median}\{S_i\} \qquad (7)$$

$M_i^s = 1$, which indicates the i-th region R_i is shadow, otherwise it is the goal. T_S is the threshold value based on the similarity in all sub-regions.

Thus, according to the mask image obtained by combining the gradient magnitude and conjunction of the space-time, obtaining the final shadow mask Mi calculation is following:

$$M_i = M_i^G \wedge M_i^s \qquad (8)$$

$\wedge$ represents a logical "and" operation, $M_i = 1$ means the i-th region R_i is eventually determined to be a shadow, otherwise it is the goal.

IV. Experimental results and analysis

In this paper, we did some experiments about the method of STMSD in the video sequence Highway, Intelligent Room, Hallway and CAVIAR. Figure 3 shows the shadow result at different times and different scenarios. As it can be seen, the method of STMSD can get a good shadow test result at the different time in the same video scene.

25# 45#

65# 185#

Highway

25# 164#

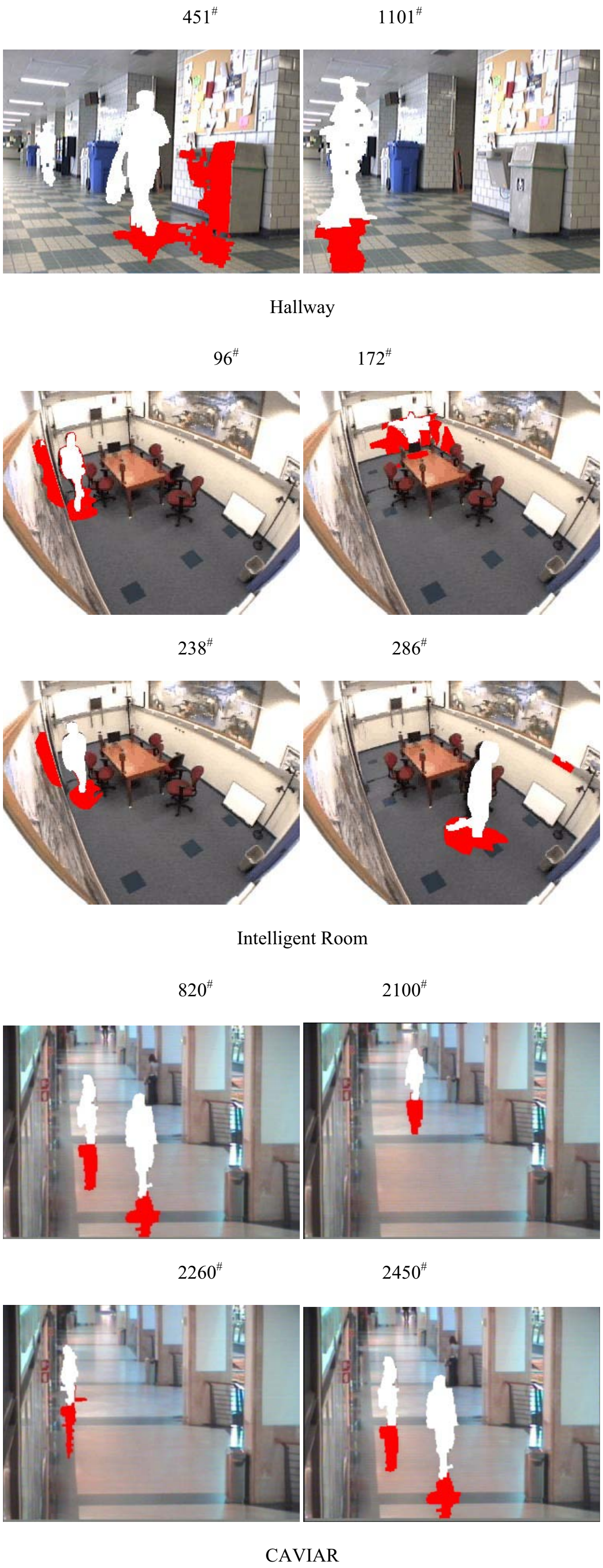


Figure 3 The Qualitative Comparison Results

V. CONCLUSIONS

Compared with pixel-based methods, STMSD detect moving shadows on the basis of adaptive segmentation regions derived from watershed algorithm. The influence of noise or uncertain factors was reduced with STMSD in detection result, and STMSD has good robustness for surveillance videos. Besides, experiment results show that STMSD is superior to fixed block method and some well-known method.

ACKNOWLEDGEMENTS

This research is partially supported by National Natural Science Foundation of China (No. 61471269), the Project of Shandong Province Higher Educational Science and Technology Program (No.J13LN39), the Project of Shandong Province Science and Technology Development Program (No.2011YD01047), the Project of Shandong Province Spark Program (2013XH06031), Doctoral Scientific Research Fund Project of Weifang University (No. 2014BS12) and Weifang Science and Technology Development Plan Project (No. 2014GX021).

REFERENCES

[1] A. Sanin, C. Sanderson and B. C. Lovell, “Pattern Recognition”, vol. 45, no. 4, (2012).

[2] R. Cucchiara, C. Grana, M. Piccardi and A. Prati, “IEEE Transactions on Pattern Analysis and Machine Intelligence”, vol. 25, no. 10, (2003).

[3] K. T. Song and J. C. Tai, “Proceedings of the IEEE”, vol. 95, (2007) February 2.

[4] J. Wang, Y. Wang and M. Jiang, “Journal of Visual Communication and Image Representation”, vol. 25, no. 5, (2014).

[5] S H Khan, M Bennamoun and F. Sohel, “Automatic feature learning for robust shadow detection”, 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (2014), pp. 1939-1946.

[6] E. Salvador, A. Cavallaro and T. Ebrahimi, “Computer Vision and Image Understanding, vol. 95, no. 2, (2004).

[7] R. Qin, S. Liao, Z. Lei and S. Li, “Moving cast shadow removal based on local descriptors”, in: International Conference on Pattern Recognition, (2010), pp. 1377-1380.

[8] E. Bullkich, I. Ilan, Y. Moshe, Y. Hel-Or, and H. Hel-Or. Proc. of 19th Intl. Conf. on Systems, Signals and Image Processing (IWSSIP 2012), (2012) April;Vienna.

[9] Gonzalez R C, Woods R E, Eddins S L．Digital Image Processing, Second Edition [M]．Beijing：Publishing House of Electronics, 2006.

[10] Gonzalez R C, Woods R E, Eddins S L．Digital Image Processing Using MATLAB [M]. Beijing：Publishing House of Electronics, 2005．

[11] Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(6): 583-598.

[12] J. Y. Dai, M. Qi and J. Z. Wang, “Optics & Laser Technology”, vol. 54, (2013).