15th IAPR International Conference on Machine Vision Applications (MVA)
Nagoya University, Nagoya, Japan, May 8-12, 2017.

Detection of View-disturbing Noise by using Time-axial Clustering in Spatio-Temporal Image

Taichi ARIMASA, Hiromi YOSHIDA and Yoji IIGUNI
Graduate School of Engineering Science, Osaka University
1 Machikaneyama-cho, Toyonaka-shi, Osaka Japan
arimasa@sip.sys.es.osaka-u.ac.jp
{yoshida, iiguni}@sys.es.osaka-u.ac.jp

Abstract

Lately, surveillance cameras are widely utilized. Nevertheless , raindrops or mud sometimes obstruct the field of view when we use them outside. In this research, we propose a method that can remove noises better than conventional method. In the conventional method, we take photographs while rotating the camera. Then, by considering the relationship between angles of photographing and objects in the images, we can detect noises. But in this method, if the luminance of the objects and noises are so close, we cannot detect them exactly. So we solve the problem by using clustering.

1 Introduction

Recently, surveillance cameras are widely prevalent and used to check traffic situation or prevent crime. When we set up them outside, we typically mount a case to the camera in order to protect it from getting wet. However, raindrops or mud sometimes stick to the glass of camera-protecting case. Then, they appear as noises in images. Thereby, we lose important information in the images. To solve the problem, there has been a lot of ways.

Firstly, there are two methods that can remove noises automatically [1], [2]. But these methods work well only when snow and raindrops are falling.

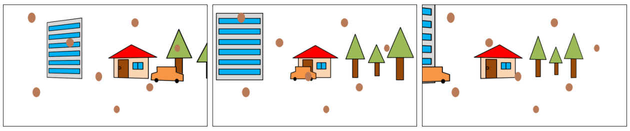
Besides that, there is the Fukuchi's method[3] to detect and remove noises. In this method, they take photographs while rotating the camera and find out noises by considering the relationship between angles of photographing and objects in the images. Yet, there is a problem that we cannot take full advantage of information of R, G and B, because they must convert images to gray. So we propose a method that can remove noises by making use of information of R, G and B. To make use of that information, we use clustering while they use median in conventional method.

In the following, we will explain the camera we use and the conventional method in section 2. In section 3, we will explain our proposed method. Then we will compare our proposed method with conventional one in section 4. Finally, we will conclude in section 5.

2 Conventional Method

2.1 Precondition

As the precondition of noises, we suppose that they move with camera rotation and remain stationary in images while photographing.



(a) t=0 (b) t=(N-1)/2 (c) t=N

Figure 1: Photographed images $I(x, y, t)$

2.2 Photographing

We take photographs while rotating the camera around the center of the lens, then let N images are obtained. Those image examples of frame 0, (N-1)/2 and N are shown in Figure 1. In those figures, brown granular areas are noises and they are in the same point in any image. Besides, a building, trees , a house and a car are in those images. The car moves from right to left and other objects stand stationary.

2.3 Position-matching

Before position-matching, we convert all images to gray and let the images be $G(x, y, t)$. We perform position-matching for all images with reference to the center image (frame (N-1)/2). The figure for explanation of position-matching is shown in Figure 2, and this position-matching is expressed as:

$$x = f\frac{f\tan\theta + \tilde{x}}{f - \tilde{x}\tan\theta} \tag{1}$$

$$y = f\frac{\sqrt{1+\tan^2\theta}}{f - \tilde{x}\tan\theta}\tilde{y} \tag{2}$$

Here, the coordinate of the image before and after conversion is (x, y), $(\tilde{x}, \tilde{y})$ respectively, the focal length is f, and the angular difference of the image is θ. Those image examples of frame 0, (N-1)/4, (N-1)/2, 3(N-1)/4 and N are shown in Figure 3 and the image obtained by superimposing them is shown in Figure 4. In Figure 4, we can see that the car and the noises move but the other objects is at rest.

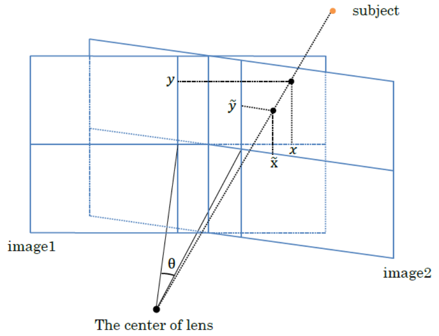


Figure 2: Explanation of position-matching

Figure 3: Position-matching images $G_p(x, y, t)$

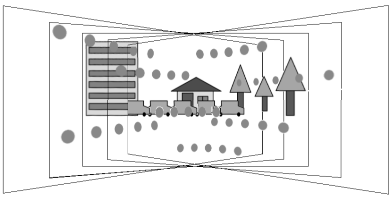


Figure 4: Superimposed image

2.4 Spatio-temporal image

The spatio-temporal image is three-dimensional collection of images which has the horizontal, vertical and time axis. The spatio-temporal image is shown in Figure 5.

2.5 Time-axial median image

The figure for explanation of time-axial median image is shown in Figure 6. For some x, y in spatio-temporal image, we find median of N pixel values(from $t = 0$ to $t = N - 1$), then we change N pixel values to the median. In Figure 6, m is expressed as:

$$m = median\{G_p(x, y, t) \mid t = 0, \cdots, N - 1\} \tag{3}$$

And we do the same processing for all x, y. By doing so, we can create the images from which moving objects

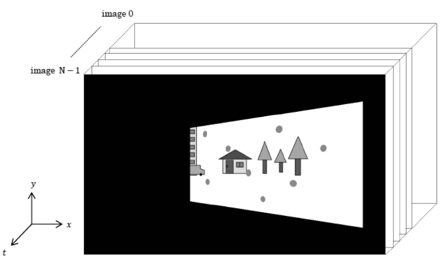


Figure 5: Spatio-temporal image $G_p(x, y, t)$

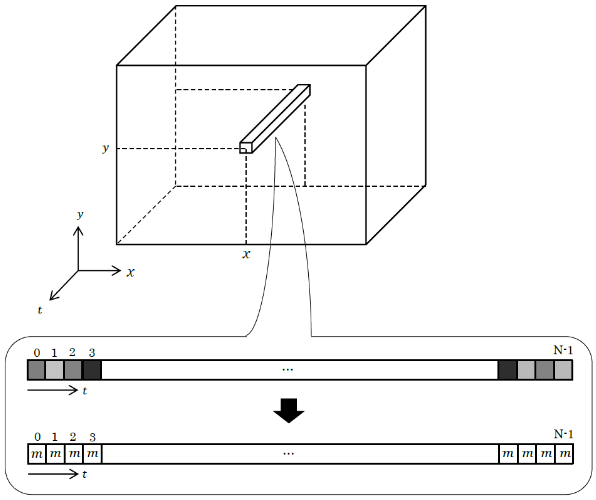


Figure 6: Explanation of time-axial median image

(a) t=0 (b) t=(N-1)/2 (c) t=N

Figure 7: Time-axial median images $M_p(x, y, t)$

are removed. That is because, when we see the pixel values from $t = 0$ to $t = N - 1$ in the same x, y, the stationary objects occupies the majority. So, in finding the median, it will be the pixel value of the stationary objects. Explaining with Figure 5, the car and the noises are cleared away. The example is shown in Figure 7

2.6 Noise-detection

We do the processing opposite to chapter 2.3, then the images such as Figure 8 are created. And we take the difference between the images($M(x, y, t)$) and the photographed images converted to gray($G(x, y, t)$). This image is expressed as:

$$D(x, y, t) = |G(x, y, t) - M(x, y, t)| \tag{4}$$

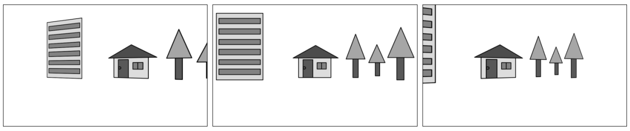
After that, we create binary images of $D(x, y, t)$ with the threshold T_1. The examples of the binary images are shown in Figure 9 and this image is expressed as:

$$D_B(x, y, t) = \begin{cases} 0 & (D(x, y, t) < T_1) \\ 255 & (D(x, y, t) \geq T_1) \end{cases} \tag{5}$$

Here, we set T_1 experimentally. And white area shows moving objects and black area shows stationary objects.

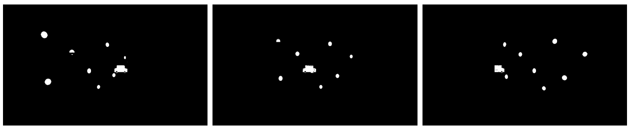
And for some x, y, we count the number of white pixels from $t = 0$ to $t = N - 1$ and we enter the counted number in the $A(x, y)$.(however, if $N \leq 256$) This image is shown in Figure 10.

Finally, by converting $A(x, y)$ to binary images with the threshold T_2, we can detect the noises. This image is shown in Figure 11.



(a) t=0 (b) t=(N-1)/2 (c) t=N

Figure 8: Time-axial median images $M(x, y, t)$



(a) t=0 (b) t=(N-1)/2 (c) t=N

Figure 9: Binary images $D_B(x, y, t)$

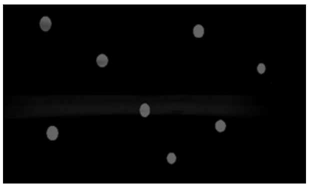


Figure 10: Noise count image $A(x, y)$

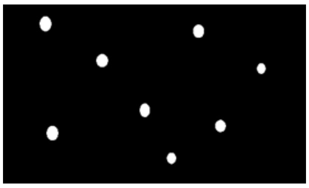


Figure 11: Noise detection image $A_B(x, y)$

2.7 Problem of conventional method

In the conventional method, gray-scaling is necessary in creating the time-axial median image. However, when the noise and the background area have similar luminance values, the difference disappears due to gray-scaling, and it cannot be found out as moving object in binarization.

3 Proposed Method

3.1 Time-axial clustering image

In order to solve the problem discussed in chapter 2.7, it is important to use information of R, G and B. In this research, we attempt to solve by using clustering.

When performing position-matching without gray-scaling, the images as shown in Figure 12 are generated. The explanation of time-axial clustering is shown in Figure 13. For some x, y, we consider the pixel values from $t = 0$ to $t = N - 1$ in the RGB space, and divide them into two clusters by k-means clustering. Then we replace N pixel values(from $t = 0$ to $t = N - 1$) with the pixel value of cluster center with the larger number. In Figure 13, c is the cluster center. By doing this, it is possible to leave stationary objects alone while maintaining color information.

3.2 Comparison with the conventional method

The flowcharts of the conventional and proposed method are shown in Figure 14. In our proposed

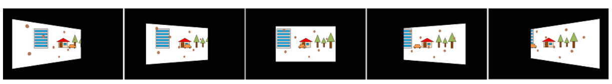


Figure 12: Position-matching images $I_p(x, y, t)$

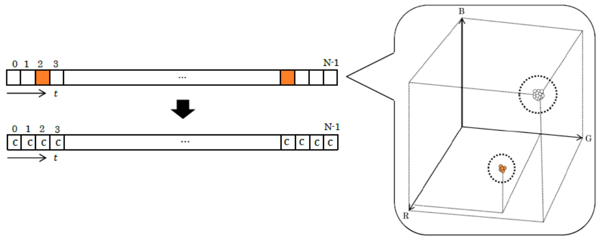


Figure 13: Explanation of proposed method

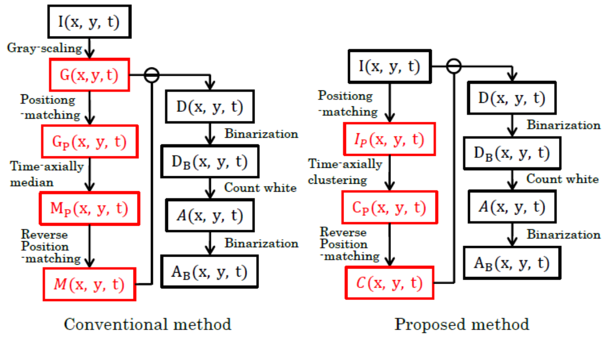


Figure 14: Comparison between the conventional and proposed method

method, it is possible to create color images in which only a stationary objects are left.

Here, the results using the conventional and our proposed method in actual images are shown in Figure 16, 17 and the photographed images are shown in Figure 15. The brown parts in the images are the noises assumed mud. In those images, a person , the person's shadow and noises are white. Though the upper left two noises(red circles) can not be detected enough in the image of conventional method, it can be seen that the noises can be more clearly detected in the image of our proposed method.

4 Experiments

We carried out noise detection experiments of our proposed method and the conventional method with eleven kinds of moving images using brown noise models. The experimental conditions are shown in table 1. For the evaluation scale, we used the recall of noises(%) and the number of false detection. The recall is index of the noise area actually detected divided by the correct noise area. And it shows how much noises are detected. Explaining with Figure 18, the formula is indicated as:

$$Recall(\%) = \frac{A \cap B}{B} \quad (6)$$

And the number of false detection is the number of pixels of the noise incorrectly found out$(A \cap \bar{B})$.

The experimental results are shown in the figure 19, 20. As can be seen from these results, the proposed method shows better noise detection performance in any of the evaluation scales.

(a) color image (b) gray image

Figure 15: Photographed images

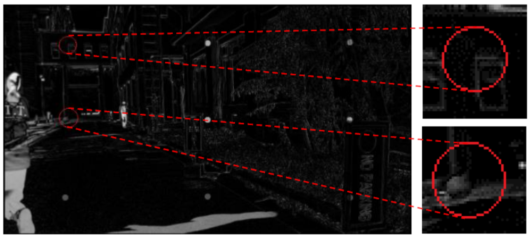


Figure 16: Conventional method

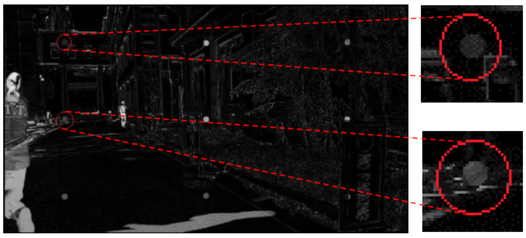


Figure 17: Proposed method

Table 1: Experimental condition

Parameter	number
Number of shots:N	101
Rotation angle per one image:θ	0.288°/image
Image size	640 × 360
Focal length:f	496.6 pixel
Threshold T_1 in conventional method	10
Threshold T_1 in proposed method	20
Threshold T_2	80

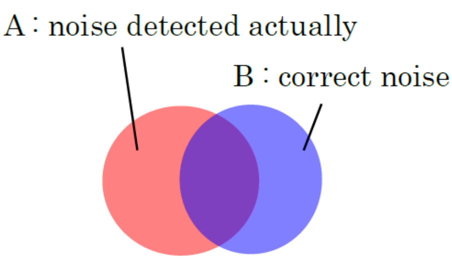


Figure 18: Explanation of evaluation scale

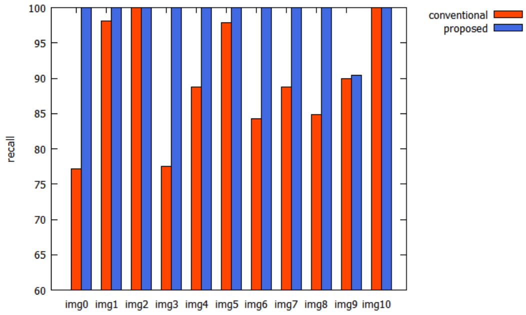


Figure 19: Recall

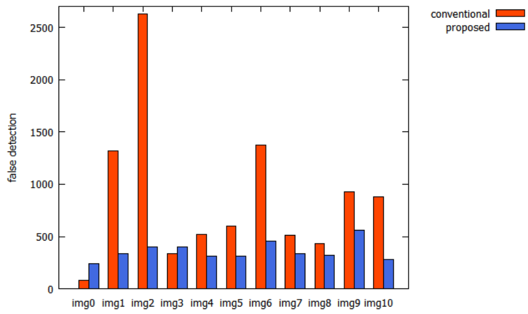


Figure 20: False detection

5 Conclusion

In this paper, we propose a method to create images with only stationary objects using clustering. As a result, it became possible to detect the moving object while maintaining color information. In addition, as a result of the experiments, there were some images in which the proposed method has lower results. Nonetheless, when considering the difference between the average value and the values of the other images, we can say that the recall of noise detection, and the number of false detection could be greatly improved. The reason for those improvements is that the process of gray-scaling becomes unnecessary.

Finally, our future tasks include automatic setting of each threshold value, correspondence when noises moves on glass of camera-protecting case, consideration of noise removal method, and the like.

References

[1] H. Hase, K. Miyake and M. Yoneda: "Real-time Snowfall Noise Elimination" , Proceedings of the 1999 IEEE International Conference on Image Processing (ICIP1999), Vol.2, pp.406409, 1999.

[2] K. Garg, S. K. Nayar: "Detection and Removal of Rain from Videos" , Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2004), Vol.1, pp.528 535, 2004.

[3] I. Fukuchi, A. Yamashita, T. Kaneko and T. Miura: "Removing Adherent Noise from Image Sequences in Rainy Weather by Spatio-Temporal Image Processing", The Journal of The Institute of Image Information and Television Engineers, Vol.62 , No.5 , pp771～777 , 2008