ICIC Express Letters
Volume **18**, Number **9**, September **2024**

ICIC International ©2024 ISSN 1881-803X
pp. 919–925

A MOVING OBJECT DETECTION METHOD BASED ON THE OTSU METHOD

Shanqian Sun and Kohei Inoue

Department of Communication Design Science
Kyushu University
4-9-1 Shiobaru, Minamiku, Fukuoka 815-8540, Japan
sun.shanqian@kyudai.jp; k-inoue@design.kyushu-u.ac.jp

Received November 2023; accepted February 2024

Abstract. *Moving object detection is to recognize the physical movements of the objects in a video or image sequence. It is a preprocessing in our pedestrian re-identification task. We proposed a discrete Fourier transform-based moving object detection method in our former research. However, it depended on an experimental threshold and remained a ghost problem. This paper proposed an Otsu method-based moving object detection method. Specifically, it compared the Otsu method with the entropic method, minimum error method, and moment-preserving method, and the experimental result showed the Ostu method performed the best, and it could determine the threshold automatically and deal with the ghost problem. Moreover, the average nearest neighbor index was used to handle the misclassification problem of no foreground images.*

Keywords: Moving object detection, Pedestrian re-identification, Otsu method, Average nearest neighbor

1. **Introduction.** In computer vision, moving object detection in a video sequence plays a fundamental role in pedestrian detection, pedestrian tracking, person re-identification, etc. Moving object detection aims to divide each video frame into the foreground (moving object) and background (static object). It is always assumed that the video is captured by one stationary camera (and so they do), and the light is stable without flashing [1].

Traditional moving object detection methods can be divided into frame difference method [2, 3], optical flow method [4], and background segment method [5, 6]. The frame difference method calculates the difference between two continuous frames to detect the moving object and would remain holes problem if it moves slowly. The optical flow method only detects the movements of the feature points, which are also not robust with rotation and deformation. The background segment method subtracts the background to get the foreground. Recently, the deep learning-based method has grabbed researchers' attention [7, 8]. However, these methods require a high frame rate and resolution video, which is hard to acquire in an actual situation.

In our former research, [9] proposed a discrete Fourier transform (DFT) based moving object detection method, which was enlightened by signal processing. It transformed the frame sequence, which could be deemed a time-domain signal, into a frequency-domain signal. In the frequency domain, zero frequency was set to 0 to remove the background. And it binarized by an experimental threshold. Figure 1 shows the framework.

This method required an experimental threshold and remained a ghost problem. Figure 2 shows a lousy binarization result by an improper threshold. (a) and (d) are two frames of our dataset. (b) and (e) are the images after DFT. The same person in the red circle in (b) is more apparent than in (e) due to the different contrast to the background. (c) and (f) are the results after binarization. The person can be seen clearly in (c) but almost

DOI: 10.24507/icicel.18.09.919

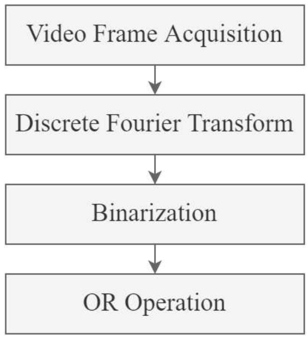


FIGURE 1. Block diagram of the former framework

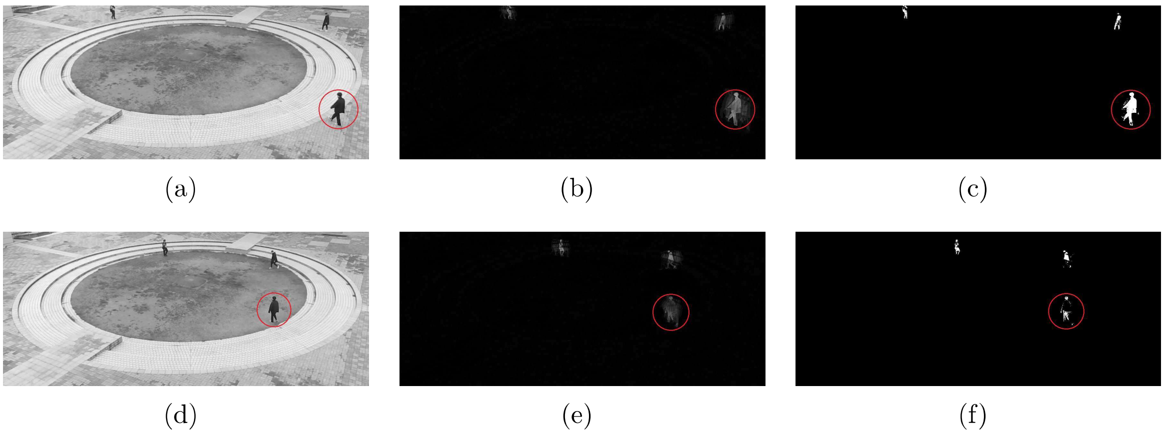


FIGURE 2. Experimental threshold-based binarization

disappears in (f). Also, it should be noted that there is a ghost around the person in (b) and (e).

To solve this problem, this paper proposed an Otsu method-based moving object detection method. Specifically, Section 2 talked about the Otsu, entropic, minimum error, and moment-preserving methods. It also utilized the average nearest neighbor index to correct the misclassification of background-only images. Section 3 shows the experimental results. Our method performs the best and can deal with the ghost problem. Section 4 concludes this paper.

2. **New Method.** This section discusses the improvements to our former moving object detection method. Figure 3 shows the block diagram of the new framework. Compared with the former, our new approach acquires the background through a median filter. Then, it uses the Otsu method for binarization. Finally, the average nearest neighbor index is used to revise the misclassification of background-only images.

2.1. **Dataset preparation.** We made a new dataset to test and verify our new method. Two cameras at Kyushu University took the dataset from different angles. It had 8 video clips, including pedestrian actions like turning around and greeting. The original videos were 1080p at 30 fps and were resized to 1140×480 at 5 fps. Figure 4 shows some frames in the dataset. The first-line and second-line images are captured by two cameras separately.

2.2. **Median filter.** Our method is based on such experience that the median represents the background. With this consideration, we get the background image by a median filter. Then, we subtract the background to get the foreground. As Figure 5 shows, (a) and (d) are the same images in Figure 2, (b) and (e) are the background, and (c) and (f) are the foreground. The median filter deals with the ghost problem that happened in DFT.

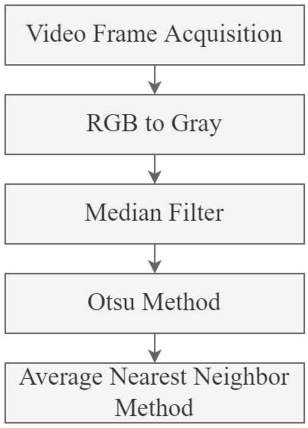
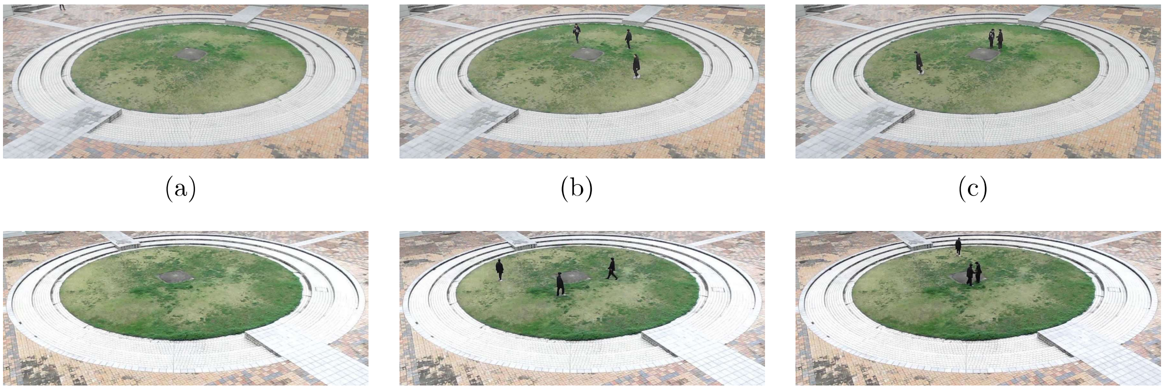


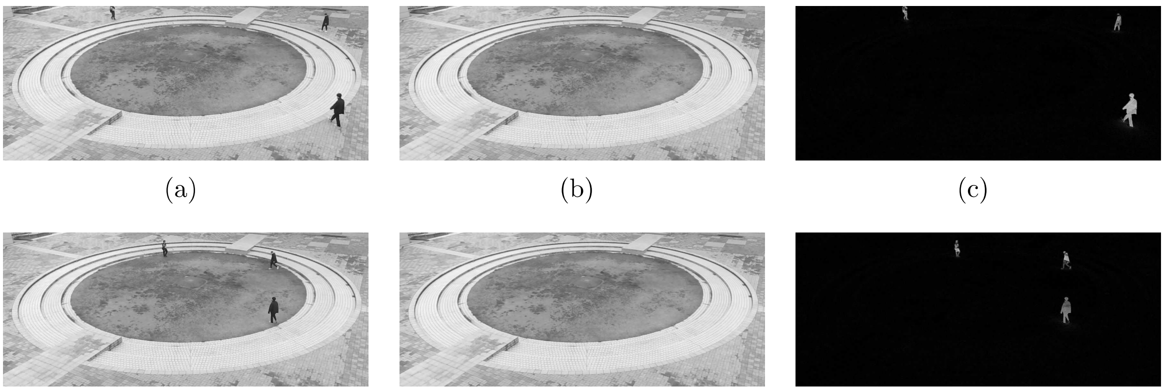
FIGURE 3. Block diagram of the new framework



(a) (b) (c)

(d) (e) (f)

FIGURE 4. Some images from our new dataset



(a) (b) (c)

(d) (e) (f)

FIGURE 5. Median filter-based foreground detection

2.3. **Binarization.** This part discusses the Otsu method with the entropic, minimum error, and moment-preserving methods [10].

The main idea of the Otsu method is to maximize the between-class variance [11]. We assume the proportions of foreground and background are p_f and p_b, and the average gray values are v_f and v_b. We can get Equations (1) and (2), in which v is the average gray value of the foreground and background. The Otsu method determines the threshold by maximizing Equation (3).

$$p_f + p_b = 1 \tag{1}$$
$$v = p_f v_f + p_b v_b \tag{2}$$
$$var = p_f(v_f - v)^2 + p_b(v_b - v)^2 \tag{3}$$

The entropic method obtains the optimal threshold based on information entropy theory [12]. Let t be the threshold, and we can have

$$H_b = -\sum_{i=0}^{t} p_i \log_e p_i \tag{4}$$

$$H_w = -\sum_{i=t+1}^{l-1} p_i \log_e p_i \tag{5}$$

where H_b and H_w can be regarded as a measure of the posteriori information entropy associated with the black and white pixels after the binarization. We can determine the t by maximizing the H as below:

$$H = H_b + H_w \tag{6}$$

In the minimum error method, the gray level histogram is viewed as an estimate of the probability density function of the mixture population comprising the gray level of the object and background pixels [10]. It is usually assumed that each of the two components of the mixture obeys normal distribution. Kittler and Illingworth [13] introduced a criterion function $J(t)$, which is given by

$$J(t) = 1 + 2\{P_1(t)\log_e \sigma_1(t) + P_2(t)\log_e \sigma_2(t)\} - 2\{P_1(t)\log_e P_1(t) + P_2(t)\log_e P_2(t)\} \tag{7}$$

in which

$$P_1(t) = \sum_{g=0}^{t} h(g) \tag{8}$$

$$P_2(t) = \sum_{g=t+1}^{l-1} h(g) \tag{9}$$

$$\sigma_1^2(t) = \frac{\sum_{g=0}^{t}(g - \mu_1(t))^2 h(g)}{P_1(t)} \tag{10}$$

$$\sigma_2^2(t) = \frac{\sum_{g=t+1}^{l-1}(g - \mu_2(t))^2 h(g)}{P_2(t)} \tag{11}$$

$$\mu_1(t) = \frac{\sum_{g=0}^{t} h(g)g}{P_1(t)} \tag{12}$$

$$\mu_2(t) = \frac{\sum_{g=t+1}^{l-1} h(g)g}{P_2(t)} \tag{13}$$

The optimal threshold is obtained by minimizing $J(t)$.

Finally, the moment-preserving method is an improved p-tile method [14], and the optimal threshold p_0 is given by

$$p_0 = \frac{z - m_1}{\sqrt{c_1^2 - 4c_0}} \tag{14}$$

in which

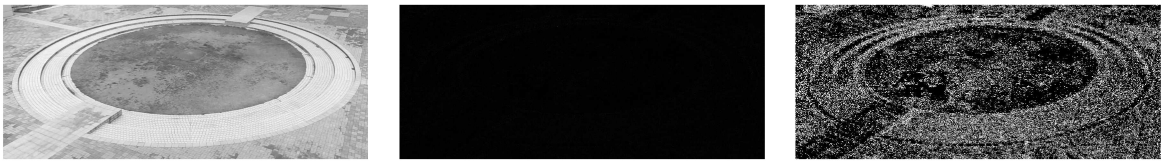
$$z = \frac{1}{2}\left(\sqrt{c_1^2 - 4c_0} - c_1\right) \tag{15}$$

$$c_0 = \frac{m_1 m_3 - m_2^2}{m_2 - m_1^2} \tag{16}$$

$$c_1 = \frac{m_1 m_2 - m_3}{m_2 - m_1^2} \tag{17}$$

$$m_i = \frac{1}{n} \sum_{g=0}^{l-1} g^i h(g), \quad i = 1, 2, 3 \tag{18}$$

2.4. **Average nearest neighbor.** There is a problem that occurs after the binarization. (a) in Figure 6 is the original image, and (b) is the image after the median filter. The original image is almost the same as the background that (b) appears all black. (c) is the image after the Otsu method and mistakenly divides the background into foreground and background.



(a) (b) (c)

Figure 6. Misclassification of background-only images

We use the average nearest neighbor index to solve the misclassification above. This index describes a feature that can be considered as clustered or dispersed [15]. The average-nearest-neighbor ratio is calculated as the observed average distance divided by the expected average distance (with the expected average length being based on a hypothetical random distribution with the same number of features covering the same total area). It can be expressed as

$$ANN = \frac{\overline{D_o}}{\overline{D_e}} \tag{19}$$

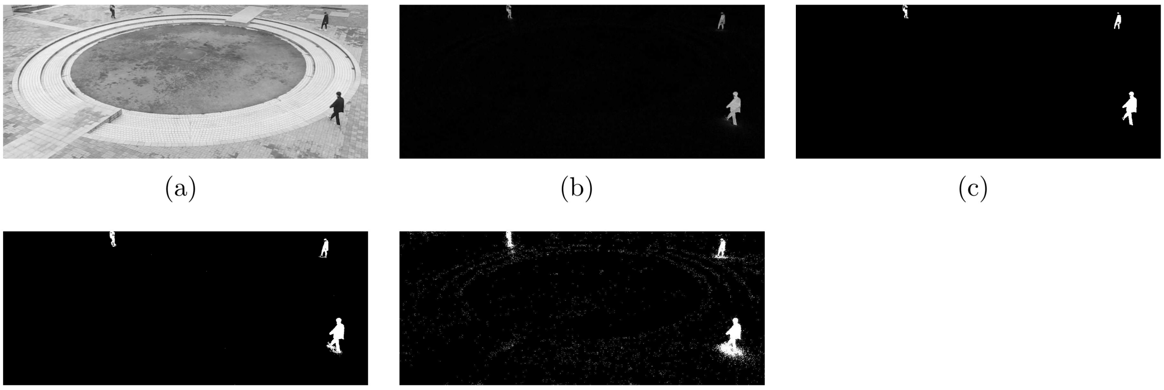
in which $\overline{D_o}$ is the observed mean distance between each feature and its nearest neighbor, and $\overline{D_e}$ is the expected mean distance for the features given in a random pattern, as Equations (20) and (21) show. n corresponds to the total number of features, and A is the area of a minimum enclosing rectangle around all features.

$$\overline{D_o} = \frac{\sum_{i=1}^{n} d_i}{n} \tag{20}$$

$$\overline{D_e} = \frac{0.5}{\sqrt{\frac{n}{A}}} \tag{21}$$

3. **Experimental Result.** Figure 7 shows the binarization result through different methods. (a) is the original image, and (b) is the image after the median filter. (c), (d), (e), and (f) are the results of the Otsu, entropic, minimum error, and moment-preserving methods, respectively. The Otsu method performs better than the other three methods. It should be noticed that the moment-preserving method regards all the pixels as the foreground due to the imbalance histogram that most pixels are background and only a few are foreground.

Then, we want to talk about the result of ANN. In our research, white pixels (foreground) were regarded as features. We separated the images into training data and testing data and used the kd-tree to calculate $\overline{D_o}$ then sorted the ANN, and Tables 1 and 2 show the result. The label we marked manually describes the image, which is all background or



(a) (b) (c)

(d) (e) (f)

FIGURE 7. Binarization result through different methods

TABLE 1. ANN and label in training dataset

ANN	⋯	1.099	1.089	1.083	0.183	0.181	0.174	⋯
Label	⋯	BO	BO	BO	WF	WF	WF	⋯

* BO stands for background-only, and WF stands for with-foreground.

TABLE 2. ANN and label in testing dataset

ANN	⋯	1.194	1.192	1.192	0.156	0.155	0.155	⋯
Label	⋯	BO	BO	BO	WF	WF	WF	⋯

has a foreground. We can discard the image that the ANN is lower than 0.633 (average number of 1.083 and 0.183) as it is without foreground with a high probability.

4. **Conclusions.** This paper proposed an Otsu method-based moving object detection method. Specifically, it used two cameras at different angles to take pictures and made a dataset. Then, a median filter was used to get the background, and the foreground was obtained by subtracting the background. Finally, it used the Otsu method for binarization, which no longer needs an experimental threshold. The experimental result showed it performed well and could solve the ghost problem. Moreover, it utilized the average nearest neighbor index to avoid dividing background-only images into background and foreground. In our future research, we will separate the foreground into pedestrians and pair the pedestrians between two camera images. Finally, we will achieve pedestrian re-identification and improve the accuracy as multi-angle images can provide more information than only one image.

Acknowledgment. This work was supported by JST SPRING, Grant Number JPM JSP2136, and JSPS KAKENHI, Grant Number JP21K11964.

REFERENCES

[1] R. Kalsotra and S. Arora, Background subtraction for moving object detection: Explorations of recent developments and challenges, *The Visual Computer*, vol.38, pp.4151-4178, 2022.

[2] Y. Jusman, L. Hinggis, R. Wiyagi, N. Isa and F. Mujaahid, Comparison of background subtraction and frame differencing methods for indoor moving object detection, *2020 1st International Conference on Information Technology, Advanced Mechanical and Electrical Engineering (ICITAMEE)*, pp.214-219, 2020.

[3] Z. Tian, G. Yang, Y. Zhang, Z. Cui and Z. Bi, A range-gated imaging flash Lidar based on the adjacent frame difference method, *Optics and Lasers in Engineering*, vol.141, 106558, 2021.

[4] H. Xu, J. Zhang, J. Cai, H. Rezatofighi and D. Tao, GMFlow: Learning optical flow via global matching, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.8121-8130, 2022.

[5] X. Wang, S. Wang, Y. Guo, K. Hu and W. Wang, Coal gangue image segmentation method based on edge detection theory of star algorithm, *International Journal of Coal Preparation and Utilization*, vol.43, pp.119-134, 2023.

[6] S. Roy and M. Bhowmik, A comprehensive survey on computer vision based approaches for moving object detection, *2020 IEEE Region 10 Symposium (TENSYMP)*, pp.1531-1534, 2020.

[7] Z. Charouh, A. Ezzouhri, M. Ghogho and Z. Guennoun, A resource-efficient CNN-based method for moving vehicle detection, *Sensors*, vol.22, 1193, 2022.

[8] R. Chandrakar, R. Raja, R. Miri, U. Sinha, A. Kushwaha and H. Raja, Enhanced the moving object detection and object tracking for traffic surveillance using RBF-FDLNN and CBF algorithm, *Expert Systems with Applications*, vol.191, 116306, 2022.

[9] S. Sun and K. Inoue, A moving object detection method based on discrete Fourier transform, *ICIC Express Letters*, vol.16, no.11, pp.1235-1240, 2022.

[10] S. Pare, A. Kumar, G. Singh and V. Bajaj, Image segmentation using multilevel thresholding: A research review, *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol.44, pp.1-29, 2020.

[11] S. Fazilov, O. Yusupov and K. Abdiyeva, Mammography image segmentation in breast cancer identification using the Otsu method, *Web of Scientist: International Scientific Research Journal*, vol.3, pp.196-205, 2022.

[12] A. Caticha, Entropy, information, and the updating of probabilities, *Entropy*, vol.23, 895, 2021.

[13] J. Kittler and J. Illingworth, Minimum error thresholding, *Pattern Recognition*, vol.19, no.1, pp.41-47, 1986.

[14] R. Seelaboyina and R. Vishwakarma, Different thresholding techniques in image processing: A review, *Proc. of the 3rd International Conference on Data Science, Machine Learning and Applications (ICDSMLA 2021)*, pp.23-29, 2023.

[15] D. Yao, P. Hoorn and N. Litvak, Average nearest neighbor degrees in scale-free networks, *arXiv Preprint*, arXiv: 1704.05707, 2017.