Improved gamma corrected layered adaptive background model

Kousuke Sakamoto[1], Hiroki Yoshimura[2], Masashi Nishiyama[2], and Yoshio Iwai[2][0000−0003−2248−4563]

[1] Graduate School of Sustainability Science, Tottori University, Tottori, Japan
s142024@eecs.tottori-u.ac.jp
[2] Cross-Informatics Research Center, Tottori University, Tottori, Japan
{yosimura, nishiyama, iwai}@tottori-u.ac.jp

Abstract. This paper proposes a method for pixel-based background subtraction with improved gamma correction and a layered adaptive background model (IGLABM). The main problems of background subtraction are background oscillation and shadow. To solve these problems, we have proposed the gamma corrected layered adaptive background model (GLABM), however the performance of GLABM is not sufficient for real scenes. We hence improve the gamma estimation and prepossessing step of GLABM in this study using the covariance matrix of each pixel. We demonstrate the performance of the proposed improved method by comparing it with GLABM and other pixel-based background subtraction methods.

Keywords: Blind gamma correction · Background subtraction · Adaptive background model

1 Introduction

In recent years, given the drastic changes in our information-based society, it is necessary to pay close attention to the public safety and security of society. Against this social background, the security camera market has grown, and the need for moving object detection has also grown. Moving object detection is an important task in various practical security systems, such as person identification and traffic monitoring.

Many studies for moving object detection have been performed, and this has substantially improved detection accuracy. Object detection methods use various types of features, inter-frame differences, and background subtraction. In particular, background subtraction is frequently used for its simplicity. However, to deal with dynamic scene changes, background models should be adaptive to changes at pixel, region, or frame level. To deal with dynamic scene changes, we have previously proposed methods that employ adaptive background models (ABMs)[1, 2, 4]. An ABM is a two-color reflection model for dynamic illuminance changes and shadow removal[1]. The layered ABM (LABM) has multiple ABMs for background oscillations such as waving leaves or fluttering flags[2].

The gamma-corrected LABM (GLABM) adds gamma estimation to deal with various images captured by unknown cameras[4]. However, the performance of the gamma estimation implemented in the GLABM is not sufficient for use in real scenes.

In this paper, we propose a method for pixel-based background subtraction called the improved GLABM (IGLABM). The IGLABM method can deal with image sequences containing dynamic scene changes such as illuminance changes, waving leaves, or fluttering flags that were captured by an unknown camera. We demonstrate the performance of the proposed method by comparing it with other pixel-based background subtraction methods.

2 Related Work

As described in the previous section, many related methods have been proposed[5], and background subtraction is often used for detecting moving objects. The simplest model uses the first frame as the background and subtracts all subsequent frames, but this does not remove background oscillation. Lai and Yung uses a weighted moving average over several temporal frames for the background model[6]. For background subtraction, the background model is the key to success. There are many approaches for background modeling such as a statistical model[10], non-parametric model[7], fuzzy model[9], and low-rank sparse decomposition[8].

One of successful approaches to background modeling is statistical modeling. The distribution of pixels is frequently modeled as a univariate Gaussian, even if the input images are in RGB color space[10]. Pixels in the background are represented by six parameters: the sample mean and the standard deviation for each color components. A single Gaussian model, however,is not able to handle the background oscillation problem. Stauffer and Grimson proposed mixture of Gaussian (MoG) models for the background[11] to solve this problem. KaweTraKulPong and Bowden improved the update rules for the MoG model to solve dynamic illumination changes[14].

Our approach to background modeling is also based on statistical modeling, but our base model is the ABM[1], not a single Gaussian model. As described in the previous section, the ABM has two univariate Gaussians to model sunlight and ambient light (sky color), and LABM also has multiple layered ABMs, like a MoG[2]. These properties enable us to handle the background oscillation and remove shadows.

The implicit assumption of image processing is that the color space must be linear RGB space. This is true for images captured by industrial cameras, but not true for images captured by web or commercial cameras because these images are gamma corrected. If the input color space is not a linear RGB color space, the performance of background subtraction decreases. To solve this problem, many methods for estimating gamma have also been proposed[15, 16]. We also tackled this problem by adding a prepossessing stage for blind gamma estimation into LABM[4].

3 IGLABM

Before we explain the detail of IGLABM, we briefly explain the concept of ABM. We assume a static camera capturing images of outdoor scenes.

3.1 ABM, adaptive background model[1]

The adaptive background model is a two-color reflection model of ambient light and sunlight, which is similar to a dichromatic reflection model[3]. The pixel value $\boldsymbol{E}(\boldsymbol{x}$,t) is given by the following equation:

$$\boldsymbol{E}(\boldsymbol{x}, t) = \boldsymbol{L}_a(\boldsymbol{x}, t) + S_d(\boldsymbol{x}, t)\boldsymbol{L}_d(\boldsymbol{x}, t), \tag{1}$$

where $\boldsymbol{L}_a$ and $\boldsymbol{L}_d$ are the reflections of the ambient light and sunlight, respectively. Moreover, $S_d(\boldsymbol{x}, t)$ represents the degree of brightness of the sunlight at point $\boldsymbol{x}$ at time t and ranges from 0 to 1. Additionally, $\boldsymbol{L}_a$ and $\boldsymbol{L}_d$ are assumed to be Gaussian processes.

3.2 GLABM, gamma corrected layered adaptive background model[4]

As described in the previous section, ABM can deal with dynamic illuminance changes and shadow removals because of reflection parameter S_d. However, the ABM has two disadvantages, one is the color space and the other is background oscillation such as waving of leaves or fluttering of flags. We have address these problems using gamma estimation and a layered background model, respectively. Hence, our method, GLABM, has a prepossessing step for gamma estimation and a (multiple) layered ABM (LABM), as shown in Fig. 1.

In this paper, we improve the prepossessing step and gamma estimation step of GLABM from the method in [4] and compare the performance of the proposed method (IGLABM) with those of other existing methods. We explain the details of IGLABM in the next section.

3.3 Proposed Method

Preprocessing for gamma estimation We use ABM for the background model, so the color space of the input image must be linear RGB. We estimate the gamma value of the input images, and then apply inverse gamma correction to them. The true values of gamma are unknown because a camera setting changes in various ways when recording or capturing images.

We evaluate the linearity of background pixels after inverse gamma correction and determine the best gamma value. For better estimation, we eliminate foreground objects from the input images as much as possible. If these objects' pixels are treated as background pixels, the linearity of the background pixels decreases. We therefore first apply LABM to the input images $I(\boldsymbol{x})$ and generate synthesized background images $B(\boldsymbol{x})$ by replacing the foreground pixels with

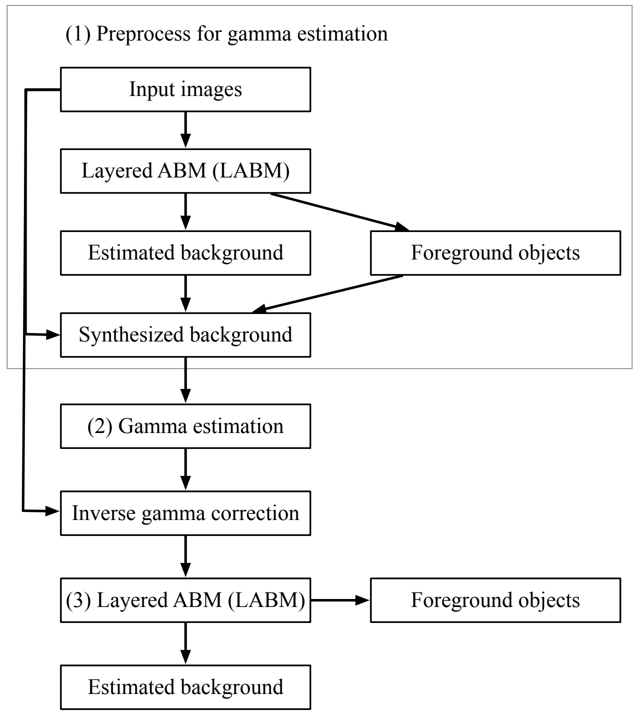

Fig. 1. Process flow of IGLABM

background pixels $S(\boldsymbol{x})$ estimated by LABM. A synthesized background image $B(\boldsymbol{x})$ is expressed by the following equation:

$$B(x) = \begin{cases} I(x) & \text{if } \boldsymbol{x} \text{ is foreground,} \\ S(x) & \text{otherwise.} \end{cases}$$

Gamma estimation We first estimate the gamma values of each pixel. We calculate the covariance matrices of intensity at each pixel for a certain number of frames after inverse gamma correction with various values in $\mathit{\Gamma}$, and then estimate the linearity of each pixel from the eigenvalues of the corresponding covariance matrix.

If the illuminance of each pixel is changed linearly in the direction of sunlight, the most significant eigenvalue of the covariance matrix increases and other eigenvalues decrease or even become zero. From this observation, to determine the best gamma value, we use the index L, which is defined by the following equation:

$$L(\boldsymbol{x}) = \frac{\lambda_2 + \lambda_3}{\lambda_1},$$

where $\lambda_{1,2,3}$ denotes the eigenvalues of the covariance matrix at $\boldsymbol{x}$ in descending order.

The best gamma value $\gamma_{\boldsymbol{x},best}$ at $\boldsymbol{x}$ is determined by the following equation:

$$\gamma_{\boldsymbol{x},best} = \arg\min_{\gamma \in \Gamma} L,$$

where Γ is a set of candidate values of γ. In this paper, we use $\Gamma = \{\gamma | 0.5 \leq \gamma \leq 2.5\}$.

When the gamma decreases, the input images becomes darker and the most significant eigenvalue λ_1 also decreases. Hence, the index L decreases and becomes unreliable when the gamma decreases. Therefore, we remove the outliers of $\gamma_{\boldsymbol{x},best}$ using $\lambda_1 \leq T_1$, where T_1 is the lower threshold of eigenvalue λ_1.

Moreover, if a pixel changes non-linearly, the volume of the covariance matrix is larger and $\lambda_{2,3}$ increases. Hence, the index L increases when a pixel changes non-linearly. We must determine the pixels that change linearly, so the index should be nearly zero. Therefore, we also remove the outliers of $\gamma_{\boldsymbol{x},best}$ using $\lambda_1 > T_2, \text{and} L < T_3$, where T_2 is the upper threshold of the eigenvalue λ_1 and the T_3 is the threshold of the index L.

After removing the outliers, we estimate the gamma value γ_{best} of the all images by voting $\gamma_{\boldsymbol{x},best}$ into Γ parameter space. We use the weighted mean as γ_{best} of the top N-th candidates in the voting Γ parameter as the follows:

$$\gamma_{best} = \frac{\sum_{i=1}^{N} n_i \gamma_i}{\sum_{i}^{\#\Gamma} n_i} \quad (n_1 \geq n_2 \geq \cdots),$$

where n_i is the number of votes, γ_i is the corresponding gamma value in Γ, and $\#\Gamma$ is the number of elements in Γ. In this paper, we use $N = 5, \#\Gamma = 21$.

LABM Now that we have the estimated gamma value γ_{best} for all input images, we can apply an inverse gamma correction to the input images to obtain images in linear RGB color space. Then, we extract foreground regions by using LABM, which has multiple ABM layers. Each layer has a likelihood for the background pixels, and is maintained in descending order of likelihood in the layer list. The idea of a layer structure has been previously proposed[11]; however, our main contribution is that we apply the concept to our ABM instead of a MoG model. The system flow of LABM is depicted in Fig. 2

4 Experimental results

We evaluate the performance of the proposed method and compare the method with existing pixel-based methods. We prepared our datasets “Tree” and “Flag,” which were captured in linear RGB, where $\gamma = 1$, to evaluate the performance of gamma estimation. The Tree dataset contains background oscillation consisting of waving of leaves and object shadows. The Flag dataset also contains background oscillation consisting of a fluttering of a flag and object shadows. Examples of these datasets are shown in Fig. 3.

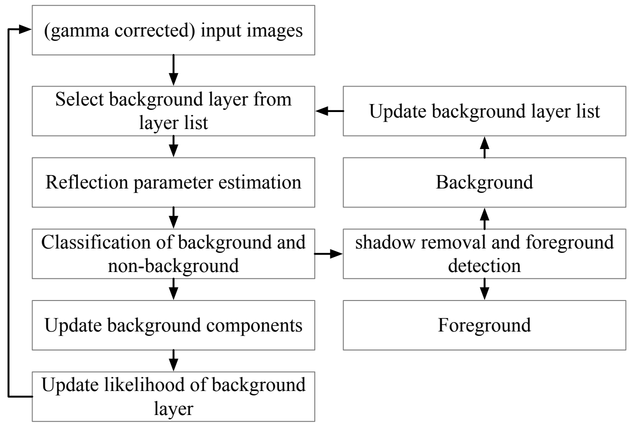

Fig. 2. System flow of LABM

tree flag

Fig. 3. Examples of the Tree and Flag datasets

We also used several image sequences collected from the benchmark datasets: "CDW-2012" [12] and "PETS2001" [13]. We use "BusStation," "Overpass," "Fountain02," "Backdoor," and "Canoe" from CDW-2012 for evaluation, because these image sequences are outdoor scenes. In total, we use six datasets for experiments. We also used "dataset3 testing camera1" of PETS2001, which is referred as "PETS" in this paper.

4.1 Validity of background image synthesis

To evaluate the validity of synthesizing background images, we compare the accuracy of gamma estimation on synthesized background images and input images. We choose 100 images randomly from each dataset, and then applied gamma correction to them from $\gamma = 0.5$ to $\gamma = 2.5$. The results of gamma estimation are shown in Fig. 4. The horizontal axis is threshold value T_3 of the index L, and the vertical axis is the mean absolute error of gamma estimation. The estimation gamma values change when the other thresholds $T_{1,2}$ change. The values of T_1 in this experiment are 5, 10, 15, and 20. The values of T_2 in

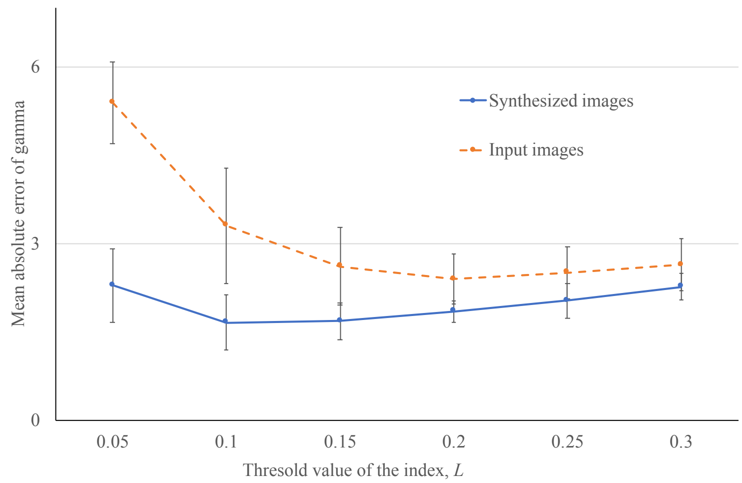

Fig. 4. Accuracy of gamma estimation

this experiment are 100, 200, 300, 400, 500, and 1,000. The standard deviations are also given in the figure.

The figure shows that the accuracy of gamma estimation using the synthesized background is better than that of gamma estimation using the input images. This is because the concept of our gamma estimation is based on the linearity of illuminance changes. If a pixel of interest is background, in other words, from a static object, the linearity is satisfied, but if the pixel is not from a static object, the linearity is violated. This result indicates that our proposed method requires the synthesized background images for better estimation.

Hereafter, we use the best parameter values, $T_1 = 10, T_2 = 200$, and $T_3 = 0.1$, of this experiment for other experiments in this study.

4.2 Comparison of gamma estimation accuracy with existing methods

We compare the accuracy of gamma estimation with other gamma estimation methods: blind inverse gamma correction (BIGC)[15], and MV-Gamma[16], and GLABM[4]. As shown in Table 1, the proposed method (IGLABM) achieves the highest gamma estimation accuracy.

4.3 Performance comparison of background estimation with existing methods

Next, we compare the performance of the proposed method with other pixel-based background subtraction methods: MOG2[17] and MOG. MOG2 is an improved version of MOG, which chooses the number of the mixture adapted for each pixel. We used the OpenCV library for MOG and MOG2.

Table 1. Absolute estimation error of gamma values

Dataset	IGLABM	GLABM	MV-Gamma	BIGC
Tree	**0.00**	0.18	0.30	2.14
Flag	**0.10**	0.41	0.50	0.36
PETS	**0.09**	0.20	0.30	1.69
BusStation	0.30	**0.03**	0.20	1.07
Overpass	**0.10**	0.25	0.20	0.84
Fountain02	**0.38**	0.72	0.50	0.53
Backdoor	**0.07**	0.57	0.40	0.55
Canoe	0.22	0.30	0.40	0.12
Average	**0.16**	0.33	0.35	0.91

Table 2. Comparative analysis of F_1-score

Dataset	IGLABM	GLABM	LABM	MOG	MOG2
Tree	**90.78**	90.64	**90.78**	87.69	84.15
Flag	79.31	**81.27**	79.94	69.75	63.54
PETS	51.14	49.65	49.65	**61.32**	29.34
BusStation	65.89	**66.35**	**66.35**	44.60	58.13
Overpass	**66.38**	64.06	60.24	48.80	51.89
Fountain02	**65.84**	63.17	63.17	57.37	26.84
Backdoor	76.21	71.47	76.48	**78.85**	67.10
Canoe	82.00	79.30	**82.17**	44.90	56.45
Average	**72.19**	70.74	71.10	61.66	54.68

Examples of the estimation results for foreground detection in this experiments are shown in Figs. 5 and 6. The results of IGLABM, GLABM, and LABM are almost the same because the gamma value of Tree is 1.0 and the gamma estimation errors are small. MOG and MOG2 cannot remove shadows from the foreground object, but LABM, GLABM, and IGLABM can remove shadows clearly. MOG2 cannot remove background oscillation such as waving leaves perfectly, but other methods can remove the background oscillation.

We show the F_1-score of each method in Table 2. Here, the proposed method also achieves the highest accuracy of all methods.

5 Conclusion

We proposed the IGLABM method for foreground detection based on pixel-wise background subtraction with improved gamma correction and layered ABMs. Gamma estimation is effective for pixel-wise subtraction but assumes a linear RGB color space. The experimental results show that the performance of foreground detection of IGLABM is superior to those of existing methods from the experimental results. In future work, we will apply our method to various open datasets to improve its performance and extend it to a block-based classification method.

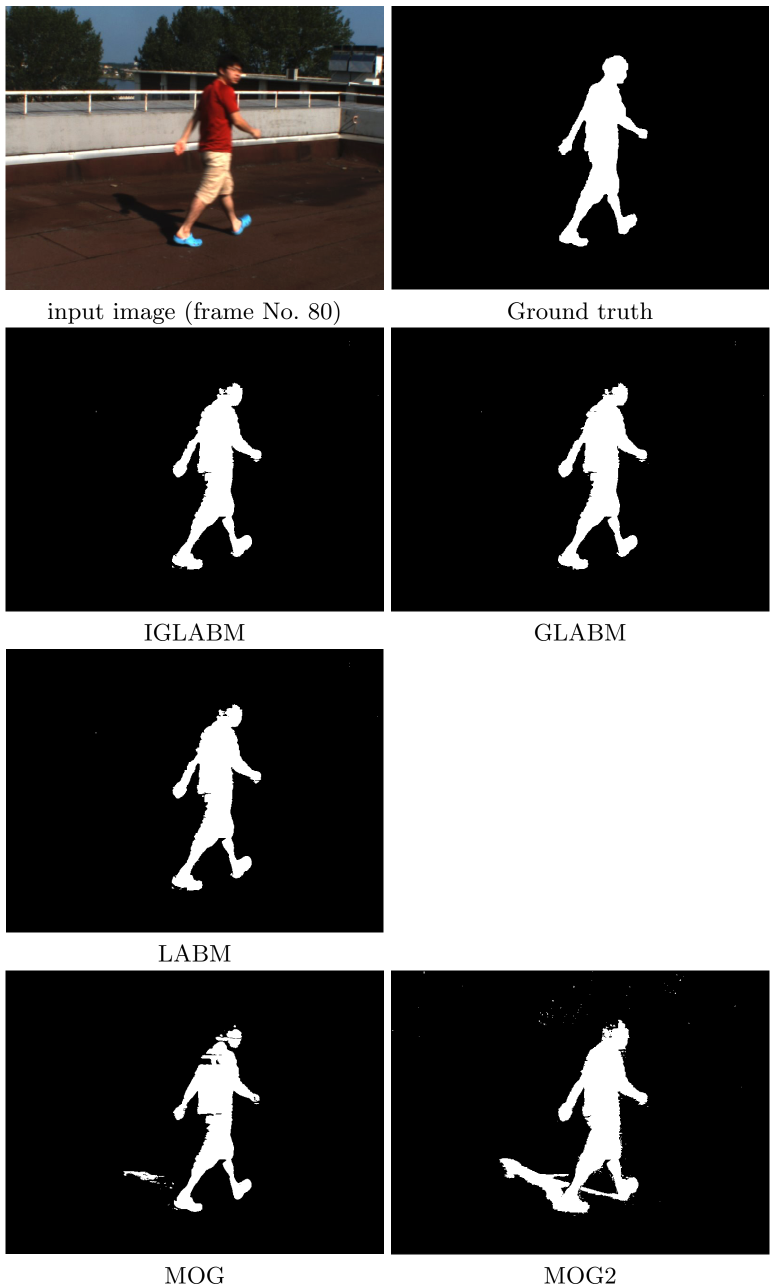

Fig. 5. Examples of foreground detection results for Tree

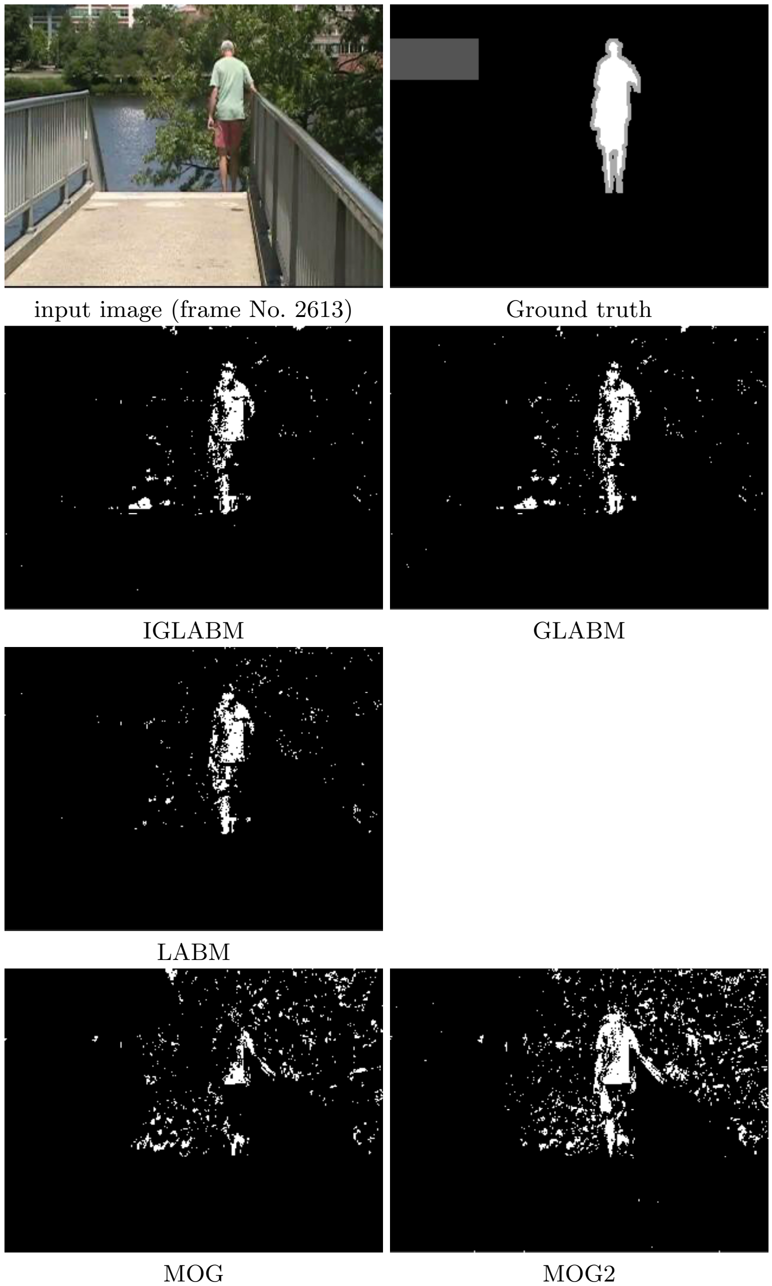

Fig. 6. Examples of foreground detection results for Overpass

Acknowledgment

This work was partially supported by MIC SCOPE Grant Number 172308003, and JSPS KAKENHI Grant Number JP17K00238, JP18H04114.

References

1. H. Yoshimura, Y. Iwai, M. Yachida, “Object Detection with Adaptive Background Model and Margined Sign Cross Correlation,” Proc. Intl. Conf. Pattern Recognit., Vol. III, pp. 19–23, 2005.
2. M. Toyoda, H. Yoshimura, M. Nishiyama, Y. Iwai, “Background subtraction robust for shadow and dynamic background in outdoor scene,” IEICE Tech. report, Vol. 116, No. 366, pp. 71–76, Dec. 2016 (in Japanese).
3. S.A Shafer, “Using color to separate reflection components,” Color: Research and Application, Vol. 10, No. 4, pp. 210–218, December, 1985.
4. K. Sakamoto, H. Yoshimura, M. Nishiyama and Y. Iwai, “GLABM: Gamma Corrected Layered Adaptive Background Model for Outdoor Scenes,” 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE), Nara, pp. 66–67, 2018,
5. I. Setitra, S. Larabi, “Background subtraction algorithms with postprocessing: A review,” in Proc. Int. Conf. Pattern Recognit., pp. 2436–2441, 2014.
6. A. H. S. Lai, N. H. C. Yung, “Afast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence,” in Proc. Int. Symp. Circuits Syst., Vol. 4. pp. 241–2444, May 1998.
7. A. Elgammal, D. Harwood, L. Davis, “Non-parametric model for background subtraction,” in Proc. Eur. Conf. Comput. Vis., pp. 751–767, 2000.
8. J. Wright, A. Ganesh, S. Rao, Y. Peng, Y. Ma, “Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization,” in Proc. Adv. Neural Inf. Process. Syst., pp. 2080–2088, 2009.
9. W. Kim, C. Kim, “Background subtraction for dynamic texture scenes using fuzzy color histograms,” IEEE Signal Process. Lett., Vol. 19, No. 3, pp. 127–130, Mar. 2012.
10. C. R. Wren, A. Azarbayejani, T. Darrell, A. Pentland, “Pfinder: Realtime tracking of the human body,” IEEE Trans. Pattern Anal. Mach. Intell., Vol. 19, No. 7, pp. 780–785, Jul. 1997.
11. C. Stauffer, W. Grimson, “Adaptive background mixture models for real-time tracking,” Proc. Conf. Comput. Vis. Pattern Recognit., pp. 246–252, 1999.
12. N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, “changedetection.net: A new change detection benchmark dataset,” in Proc. IEEE Workshop on Change Detection at CVPR-2012, pp. 16–21, 2012.
13. D. Youngand, J. Ferryman. “PETS metrics: Online performance evaluation service. In Proc. IEEE Int. Workshop on Vis. Surv. and Perf. Eval. of Tracking and Surv., pp. 317–324, 2005.
14. P. KaewTraKulPong and R. Bowden, “An Imporoved adaptive backgorund mixture model for real-time tracking with shadow detection,” in Video-Based Surveillance Systems, pp. 134–144, 2002.
15. H. Farid, “Blind Inverse Gamma Correction,” IEEE Trans. on Image Processing, Vol. 10, No. 10, pp. 1428–1433, Oct. 2001.
16. M. Mahamdioua, B. Mohamed, “New Mean-Variance Gamma Method for Automatic Gamma Correction,” Intl. J. of Image, Graphics and Signal Processing, Vol. 9, No. 3, pp. 41–54, Mar. 2017.

17. Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," Intl. Conf. Pattern Recogniti., Vol. 2, pp. 28–31, 2004.
18. S. K. Choudhury, P. K. Sa, S. Bakshi, B. Majhi, "An Evaluation of Background Subtraction for Object Detection Vis-a-Vis Mitigating Challenging Scenarios," *IEEE Access*, Vol. 4, pp. 6133–6150, 2016.