객체검출에서의 개선된 투영 그림자 제거
An Improved Cast Shadow Removal in Object Detection

빈흐타한*, 정선태*, 김유성**, 김재민**
숭실대학교*, 홍익대학교**

Thanh Binh Nguyen*, Sun-Tae Chung*,
Yu-Sung Kim**, Jae-Min Kim**
*School of Electronics Engineering, Soongsil University,
**School of Electronics and Electrical Engineering, Hongik Univ.

Abstract

Accompanied by the rapid development of Computer Vision, Visual surveillance has achieved great evolution with more and more complicated processing. However there are still many problems to be resolved for robust and reliable visual surveillance, and the cast shadow occurring in motion detection process is one of them. Shadow pixels are often misclassified as object pixels so that they cause errors in localization, segmentation, tracking and classification of objects. This paper proposes a novel cast shadow removal method. As opposed to previous conventional methods, which considers pixel properties like intensity properties, color distortion, HSV color system, and etc., the proposed method utilizes observations about edge patterns in the shadow region in the current frame and the corresponding region in the background scene, and applies Laplacian edge detector to the blob regions in the current frame and the background scene. Then, the product of the outcomes of application determines whether the blob pixels in the foreground mask comes from object blob regions or shadow regions. The proposed method is simple but turns out practically very effective for Gaussian Mixture Model, which is verified through experiments.

I. Introduction

Shadow removal is a challenge for not only object detection but also object tracking and object classification in a visual surveillance system. Especially in the case of moving object, the cast shadow is a critical problem(refer to Fig.1) since cast shadows always cause problems such as object merging, shape distortion and even object losses (due to the shadow cast over another object).

Fig. 1. Cast shadow

It is usually detected as foreground by most common background/foreground modeling, since they differ significantly from the background. Furthermore cast shadow has the same motion as the object casting it.

Many ideas were put forward by using some features such as color to remove shadow but they are not still robust under various environments. Cucchiara et. al [1] first defined an approach of shadow detection based on HSV color space. They assume that the V component is smaller than its priori value, when shadowed. It needs to tune up four parameters every time the context changes. The authors in [2] use YUV information to reduce time consumed by the transformation to HSV color space. Martel-Brisson et al. [3] assume that, for a given pixel, the shadow cast by different moving foreground objects is relatively similar and model every pixel by Gaussian Mixture Shadow Model (GMSM). The comparative and evaluative study can be found in [4]. Cast shadow on the background is generated by an object moving between a light source and the background [5].

In this paper, we propose a simple but effective cast shadow removal method for blobs in Gaussian Mixture Model (GMM) [6] in the context of gray scale video. By distinguishing edge detection filtering with selected pixels and suitable neighbors on scene frame data and background, we can extract blobs without cast shadow.

Our paper is organized as follows. In Section 2, an overview of GMM and shadow problem are introduced. After our proposed cast shadow removal method is described in Section 3, the experimental results about the proposed method are explained in Section 4, and finally conclusion is given in Section 5.

Ⅱ. Moving Foreground Detection and Shadow Problem

1. Moving Foreground Detection with Gaussian Mixture Model (GMM)

The GMM can adapt to slow illumination changes, periodical motions from clutter background, and etc. However, it cannot deal with shadows and fast illumination changes [8]. In this paper, we propose a cast shadow removing method which can correct the blob region result of GMM.

In [6], Stauffer and Grimson model the RGB value history by a mixture of K Gaussian distributions, while, in our paper, we use the K Gaussian distributions to Model for just the Y element in YUV format. Our implementation relies on the improvement idea for GMM in [7]. Each pixel in the scene is modeled by a mixture of three Gaussian distributions. The probability that a certain pixel has a value of X_N at time N can be written as (1).

$$P(X_N) = \sum_{j=1}^{k} W_j \eta(X_N; \theta_j) \qquad (1)$$

where W_k is the weight parameter of the K^{th} Gaussian component. $\eta(X_N; \theta_k)$ is the normal distribution of K^{th} component represented by (2).

$$\eta(X;\theta_k) = \eta(X;\mu_k, \Sigma_k) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(X-\mu_k)^T \Sigma_k^{-1} (X-\mu_k)} \qquad (2)$$

In formula (2) μ_k is the mean and $\Sigma_k = \sigma_k^2 I$ is the covariance of the K^{th} component. The K distributions are ordered based on the fitness value W_k/σ_k and the first B distributions are used as a model of the background of the scene where B is estimated like (3).

$$B = \arg\min_b \left(\sum_{j=1}^{b} W_j > T \right) \qquad (3)$$

The threshold T = 0.6 is the minimum fraction of the background model. Background subtraction is performed by determining any pixel that is more than 2.5 standard deviations away from any of the B distributions as a foreground pixel. The online EM algorithms by expected sufficient statistics are applied in (4).

$$\begin{aligned}
\widehat{W}_k^{N+1} &= \widehat{W}_k^N + \frac{1}{L}\left(\hat{P}(\omega_k | X_{N+1}) - \widehat{W}_k^N\right) \\
\widehat{\mu}_k^{N+1} &= \widehat{\mu}_k^N + \frac{1}{L}\left(\frac{\hat{P}(\omega_k | X_{N+1}) X_{N+1}}{\widehat{W}_k^{N+1}} - \widehat{\mu}_k^N\right) \\
\widehat{\Sigma}_k^{N+1} &= \widehat{\Sigma}_k^N + \frac{1}{L}\left(\frac{\hat{P}(\omega_k | X_{N+1})(X_{N+1} - \widehat{\mu}_k^N)(X_{N+1} - \widehat{\mu}_k^N)^T}{\widehat{W}_k^{N+1}} - \widehat{\Sigma}_k^N\right)
\end{aligned} \qquad (4)$$

And in context of L-recent window, in (5) the Gaussian component that matches with current pixel is updated.

$$
\begin{aligned}
\widehat{W}_k^{N+1} &= \widehat{W}_k^N + \frac{1}{N+1}\left(\widehat{P}(\omega_k|X_{N+1}) - \widehat{W}_k^N\right) \\
\widehat{\mu}_k^{N+1} &= \widehat{\mu}_k^N + \frac{\widehat{P}(\omega_k|X_{N+1})}{\sum_{i=1}^{N+1}\widehat{P}(\omega_k|X_i)}(X_{N+1} - \widehat{\mu}_k^N) \\
\widehat{\Sigma}_k^{N+1} &= \widehat{\Sigma}_k^N + \frac{\widehat{P}(\omega_k|X_{N+1})}{\sum_{i=1}^{N+1}\widehat{P}(\omega_k|X_i)}\left((X_{N+1} - \widehat{\mu}_k^N)(X_{N+1} - \widehat{\mu}_k^N)^T - \widehat{\Sigma}_k^N\right)
\end{aligned}
\tag{5}
$$

2. Shadow Problem

GMM often mistakenly indicates cast shadow of a object as foreground. Fig.2 illustrate this situation. This can happen when one applies GMM for scenes with the complex environments with a lot of noises. Many approaches have been proposed to improve these weaknesses. These approaches are usually classified into two main groups. One is to improve the Gaussian Model by using information such as color distortion [7] or applying GMM for reflectance component with the preprocessing [9]. Another is to use an independent processing as a sub-process of GMM to remove shadow. The solutions in the former group is commonly used to remove both attached and cast shadow, while the latter group mainly focuses on the cast shadow to standardize the results of GMM and a list of blobs for better handling higher analysis.

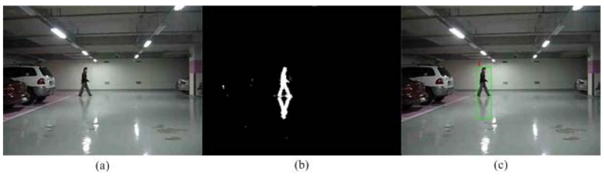

Fig. 2. (a) Original scene, (b) The result of GMM, (c) Redundant rectangle region in object tracking

We have implemented the GMM [7] but for the purpose of optimizing speed that performs background modeling for the Y element of YUV exported from the camera. Therefore the solution of the color is not applicable. Then We focus on the ideas of group two and propose a solution in the next part of this paper that based on the blob Rectangle We already calculated. The work flow of our approach is shown in Fig.3.

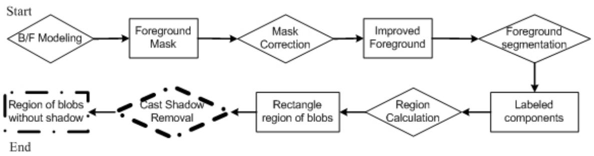

Fig. 3. Work flow of applying shadow removal in detection

Ⅲ. The Proposed Cast Shadow Removal in Motion Detection

1. Cast Shadow

Shadows are created by the effects of light so the shadows itself do not have edges, but surfaces overlaid with cast shadows can have. Cast shadows are divided into two types: shallow cast shadows and deep cast shadows. A deep cast shadow means the shadow when the region overlaide with it has much lower gray-level intensities than what the original region has. We call a region overlaid with a shadow as a shadow region. A shallow cast shadow region will usually have similar edge pattern to those of the original region without the shallow cast shadow has because the overlaid shallow cast does not bring big changes in the texture inside the original region. The deep cast shadow region will have a persistent texture and not show edges inside the region.

2. Cast Shadow Removal by Laplacian Edge Detection

Fig. 4 shows related images for arguments in this section.

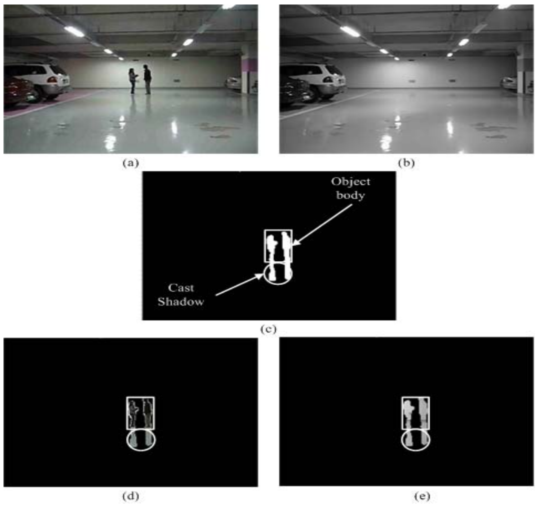

Fig. 4. (a) Current frame, (b) Current background scene, (c) Foreground mask, (d) blob region on the current frame, (e) corresonding region in the background scene

Fig. 4(c) is a foreground mask image extracted from comparison the current frame, Fig. 4(a) and the background scene, Fig. 4(b). Foreground mask image, Fig. 4(c) has a blob (white pixel region). Fig. 4(d) and 4(e) show the gray-level blob region in the current frame of Fig. 4(a) and the corresponding region int the background scene of, Fig. 4(b).

Among many types of edge detection methods such as Canny [10] and Sobel, we choose the Laplacian edge detector with kernel 3x3 of Fig.5 to reduce computation costs and also ensure a good result in common conditions.

-1	-1	-1
-1	8	-1
-1	-1	-1

Fig. 5. Laplacian Kernel

We consider the outcome results from applying the Laplacian edge detector of Fig. 5 to shadow regions of current frame and the original region of the background scene.

case 1: When cast shadow is shallow.

Since shallow cast shadow regions have similar edge patterns to those of the original regions, the product of the outcome result for the current frame and the outcome result for the background scene is likely to be greater than zero.

case 2: When cast shadow is deep.

In this case, deep shadow regions in the current frame have persistent textures and the outcome result of applying the Laplacian edge detector to the deep shadow regions is likely to have zero value.

From these considerations, we propose the following cast shadow removal rule using Laplacian edge detector.

We let $L_{CF}(x,y)$ and $L_{BG}(x,y)$ denote the outcome results of convolution of the Laplacian kernel of Fig.5 with a pixel at (x,y) in the blob regions of the current frame, and with a pixel at (x,y) in the corresponding region of the background, respectively. During convolution of the Lalacian kernel, if the neighbor pixel of the pixel at (x,y) doesn' t belong to blob regions in the foreground mask, we ignore it (that is, we take the gray-level intensity of the neighbor pixel as 0).

Now, we define $Diff(x,y)$ as

$$Diff(x,y) \equiv L_{BG}(x,y) * L_{CF}(x,y).$$

<u>Cast shadow removal Rule</u>

If for a pixel at (x,y) in the foreground mask,

$$Diff(x,y) \geqq 0 \qquad (6)$$

then the pixel at (x,y) is considered as a shadow region pixel.

However, there exist situations where even if the formula (6) is satisfied, the pixel at (x,y) in the foreground mask may not be a shadow region pixel but a object blob pixel. This can often happen in the situation with a complex background. But practically there are some reasons why we believe this rule is still acceptable in general situations. First, for many visual surveillance applications, we just need the rectangle region enclosing an blob from object, not a exact blob shape of object. Thus, even if some shadow pixels in the foreground mask are misclassified as object pixels, the outcome rectangle region is not likely to be much different from that from a exact object blob. Second, in this paper the above rule is considered for GMM background model where background is continuously updating, so that the failing situations of the proposed rule will not be persistent.

Ⅳ. Experimental Results

In order to see the effectiveness of the proposed cast shadow removal method in this paper, we conduct experiments which compare between status before and after applying the proposed method in three different types of scenes such as highway, indoor car parking lot and playground. Fig. 6(c) shows the proposed method effectively removes the cast shadow even for moving objects with deep cast shadows.

Again, the results in Fig.7 and Fig.8, fortifies robustness of our proposed method in removing cast shadow in various environments.

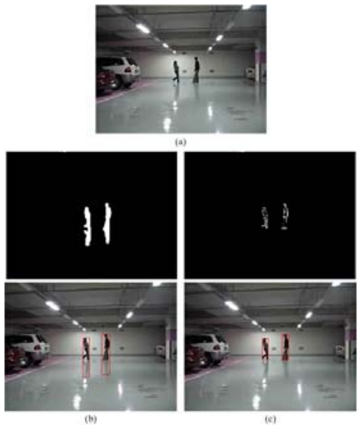

Fig. 6. (a) Origin frame, (b) The result without shadow removal, (c) The result after the proposed method applies, (parking lot)

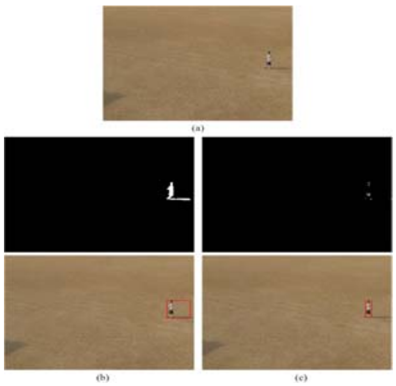

Fig. 7. (a) Origin frame, (b) The result without shadow removal, (c) The result after shadow removal, (Playground)

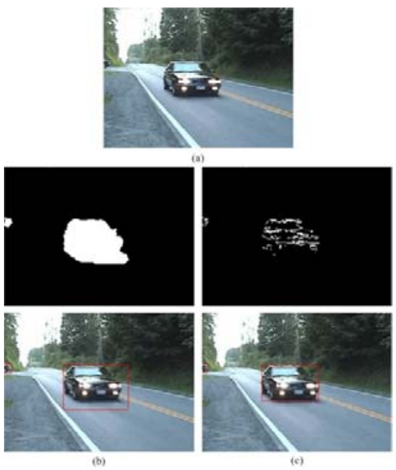

Fig. 8. (a) Origin frame, (b) The result without shadow removal, (c) The result after shadow removal, (Highway)

V. Conclusions

It has been a challenge to remove cast shadows in motion detection. In this paper, we proposed a simple but effective cast shadow removal method. The proposed method applies the Laplacian edge detector to each pixel in blob regions in the in the current frame and the background scene, and from the product of the outcomes of application, determine whether the blob pixels in the foreground mask comes from object blob regions or shadow regions. Then, product of the outcomes of application determines whether the pixels in the foreground mask comes from object blob regions or shadow regions. Experimental results show that our method worked well, especially in the case where cast shadow is deep and the background is not really similar to object texture. However, when object has low contrast, it shows weak results. Our future work will focus on how to estimate the contrast relation, texture and decide a suitable manner for detail contexts.

ACKNOWLEDGMENT

The authors would like to acknowledge the following for their support in the accomplishment of this work: Soongsil University and BK21 of Korea.

■ REFERENCES ■

[1] R.Cucchiara, C.Grana, M. Piccardi et al., "Improving Shadow Suppression in Moving Object Detection with HSV Color Information" Proc. IEEE Int' l Conf. Intelligent Transportation Systems, pp.334-339, Aug. 2001.

[2] O.Schreer, I. Feldmann, U. Golz et al., "Fast and Robust Shadow Detection in Videoconference Applications" , 4th IEEE Intern. Symposium on Video Proces. And Multimedia Comm., pp. 371-375, 2002.

[3] Martel-Brisson, and André Zaccarin, "Moving Cast Shadow Detection from a Gaussian Mixture Shadow Model" , Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, Vol. 2, pp. 643-648, June 2005.

[4] A. Prati, I. Mikic , M.M. Trivedi, and R. Cucchiara, "Detecting Moving Shadows: Algorithms and Evaluation" , IEEE Trans. PAMI, vol. 25, pp. 918-923, 2003.

[5] J. Stauder, R. Mech, J. Ostermann, "Detection of Moving Cast Shadows for Object Segmentation" , IEEE Transactions on Multimedia, vol.1, no. 1, Mar. 1999, pp. 65-76.

[6] C. Stauffer C, W.E.L Grimson, "Adaptive background mixture models for real-time tracking," 1999 IEEE CVPR.

[7] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," ICPR 2004, pp. 28-31.

[8] Ying-Li Tian, M. Lu, A. Hampapur, "Robust and efficient foreground analysis for real-time video surveillance" , Proc. IEEE Int' l Conf. Computer Vision and Pattern Recognition, 2005, pp. 1182 - 1187.

[9] Zhou Liu, Kaiqi Huang, Tieniu Tan, Liangsheng Wang, "Cast Shadow Removal with GMM for Surface Reflectance Component", Proc. IEEE Int' l Conf. Computer Vision and Pattern Recognition - Volume 01, 2006, pp. 727 - 730.

[10] Canny, J. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, 1986, pp. 679-698.