

Automatic Shadow Removal by Illuminance in HSV Color Space

Wenbo Huang¹, KyoungYeon Kim², Yong Yang¹, Yoo-Sung Kim^{2,*}

¹Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, China

²Knowledge Management System Laboratory, Inha University, South Korea

Copyright © 2015 Horizon Research Publishing All rights reserved.

Abstract In intelligent video surveillance systems, the detected moving objects often contain shadows which may deteriorate the performance of object detections. Therefore, shadow detection and removal is an important step employed after foreground extraction. Since HSV color space gives a better separation of chromaticity and intensity, it has been commonly adopted to detect and remove shadow. However, almost all the HSV color space based methods use static thresholds to separate shadows from foreground. In this paper, a dynamic threshold based method is proposed. In the proposed approach, the threshold prediction model is first established by a statistical analysis tool and then the predicted dynamic thresholds are used for shadow detection. Experiments on a self-built dataset show that the proposed method can get better reliability and robustness than the traditional methods using static thresholds.

Keywords Shadow Detection, Threshold Prediction Model, Dynamic Thresholds

1. Introduction

In recent years, for the sake of public security, the Closed Circuit Television (CCTV) has been installed to many public places, such as school, department store, elevator and parking lot, etc. However, CCTV can't detect pedestrian's action in a short time due to the fact that it only monitors a static surrounding. In order to solve this problem, the intelligent video surveillance system is needed.

In intelligent video surveillance systems, moving pedestrian detection and tracking is the foundation of realizing different intelligent applications. In order to detect pedestrians, some existing foreground segmentation algorithms, such as background subtraction, can be used. However, current moving object detection approaches usually have a typical drawback: moving shadows tend to be classified as part of the foreground. This is because that shadows share the same movement patterns and have a similar magnitude of intensity changes as those of the

foreground objects [1]. Since cast shadow causes the incorrect moving object detection, removing the shadow from foreground is important for robust and reliable intelligent surveillance systems.

In the previous researches, it is observed that the choice of features has great impact on the performance of shadow detection. And three types of features are very popular in shadow detection methods, that is, geometry, texture and chromaticity features. Among the possible features, geometry features are very important. The orientation, size and even shape of the shadows can be used as geometric features [2]. The main advantage of geometry features is that they work directly in the input frame; therefore, they don't need background reference. However, detection methods based on geometric features can be only applied to some specific object types or typical pedestrians. In addition, texture-based methods assume that shadow regions and background share the same texture structures [8]. It does not depend on colors, and would be robust to illumination changes. However, the drawback is that texture-based methods tend to be slow as they often need to compare one or more neighborhoods for each pixel. Furthermore, chromaticity-based methods assume that shadow regions in the given frames are darker compared to the background reference regions. Methods that use chromaticity-based features often choose a proper color space which chromaticity and intensity can be separated effectively than that of the RGB color space. And, the most commonly used color space is HSV [3]. Moreover, most of chromaticity-based methods are easy to implement and with inexpensive computation [7]. In addition, some combinations of the above features have been adopted by some researchers, such as [10]. The combination may improve the performance of shadow detection while the processing time will be increased [11].

Unfortunately, almost all the above methods can't reflect the shadow's change under different situations, such as sun's position or current weather condition which will lead to different illuminance values. And those methods seldom use the illumination information. For most cases, they only use a static experimental threshold for shadow removal. But as

time goes by, the illuminance will change which results in different shadow description, so it's hard to remove shadow correctly by the static thresholds. As shadows are highly related to illuminance, if the surveillance area's illuminance value is known yet, it's possible to remove shadow more correctly by changing the threshold values dynamically. In this paper, a new shadow removal method that can use dynamically the illuminance value is proposed to solve the above problem.

The remainder of this paper is organized as follows. In Section 2, motivation and a brief introduction of the proposed shadow removal method are shown. Then, in Section 3, the proposed dynamic threshold prediction method is described in detail, and the corresponding experiment results are shown in Section 4. Finally, conclusions and discussions are drawn in Section 5.

2. The Proposed Shadow Removal Method

In this paper, a shadow removal method based on dynamic thresholds is proposed. Firstly, Gaussian Mixed Model (GMM) is used to extract the foreground images of video, then, the original RGB color space is converted to HSV color space. For the next step of computing thresholds, α and β are obtained dynamically according to the defined threshold prediction model. After that, a chromaticity-based method is used for shadow detection. At last, mathematical morphology correction is also used in order to get rid of the noise. The whole procedures are shown in Fig. 1.

As mentioned in the previous section, illuminance data

would be used to change the thresholds dynamically for more correctly removing shadows under different conditions. To get the illuminance value for research, we use the illuminance capture device for data collection, and also we developed an Illuminance Input System (IIS) to collect illuminance value.

In order to get various illuminance values, we collect the data from 9:00 AM to 4:00 PM every day. For every hour, we collect the data, so totally 8 times for one day's data collection. As for every collection, the illuminance capture device is laid in the same place, and 10 values are recorded to calculate the average value as the final illuminance value. During the illuminance data collection, the person will stand in the surveillance area for about 5 seconds and then change position for another 5 seconds. This is for the shadow data collection. The collected shadow data is shown in Fig. 2.

For foreground segmentation in video, the background subtraction method is used, so we should get the background reference image first. Some background modeling methods can be used to get that image, such as linear prediction method and median filter method. Among those methods, GMM which was first presented in [4] is the most popular one. This method is used to build the model for the pixel in image sequences whose state has been changed. GMM treats the distribution of a particular pixel as a mixture of Gaussian, which is different from other algorithms modeling the values of many pixels as one particular distribution type. Since, the GMM has achieved good performance on background extraction for different applications in many related literatures, so in this paper, GMM is used to separate the background and foreground which is the first step of shadow removal.

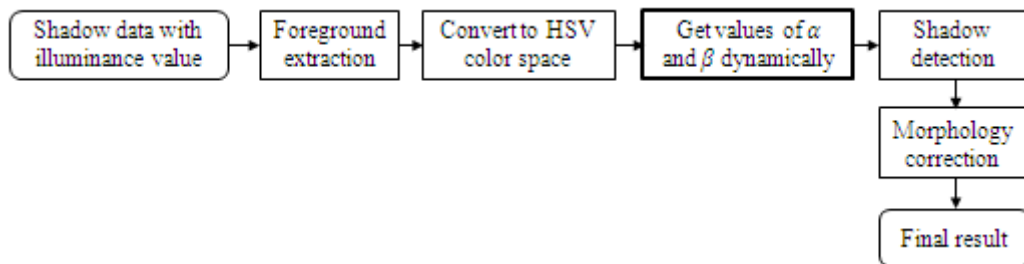


Figure 1. Flowchart of the proposed method.

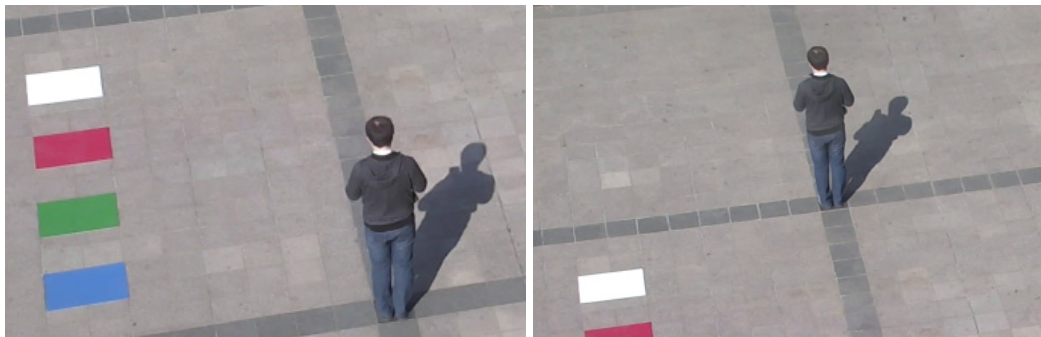


Figure 2. Shadow data collection with different positions.

Among chromaticity-based shadow detection methods, since the HSV color space can separate chromaticity and intensity better than RGB color space, Cucchiara *et al* [3] proposed a shadow detection method based on this color space, and this method has been widely used in surveillance applications [5]. In the HSV color space, the V component is a direct measure of intensity. Pixels belongs to shadow should have a lower value of V than pixels in the background, and also the hue (H) component changes within a certain limited scope. In addition, the saturation (S) component of shadows is often lower. Therefore, a pixel p is considered to be part of a shadow if:

$$\alpha \leq (F_p^V / B_p^V) \leq \beta, \quad (1)$$

$$(F_p^S - B_p^S) \leq \tau_S \quad \text{and} \quad |F_p^H - B_p^H| \leq \tau_H \quad (2)$$

In the above formulas, F_p^C and B_p^C represent the specific component C of HSV space for the pixel position p in the frame (F) and in the background (B) reference image which is obtained from GMM, respectively. α, β, τ_S and τ_H represent thresholds that are set empirically. Among those thresholds, the most important ones are α and β . The lower bound α is used to define a maximum value for the darkening effect of shadows on the background while the upper bound β prevents the system from identifying those points as shadows where they was too dark in the background [3]. In previous research, almost all the researchers use fixed thresholds. However, since the thresholds are sensitive to illuminance, static thresholds can't be applied to real time CCTV effectively.

3. The Proposed Dynamic Threshold Prediction Method

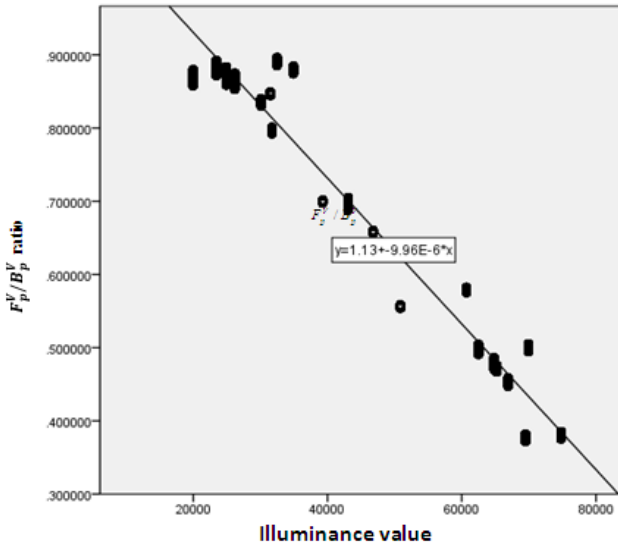


Figure 3. Statistical results of F_p^V/B_p^V ratio.

Based on plenty of collected illuminance and shadow data,

we use SPSS for the threshold prediction. For this task, 24 surveillance videos with different illuminance values are adopted, and for each video, 10 images are extracted. Then, we randomly select 10^4 pixels from the shadow regions of every image for the purpose of statistics. Since α and β are the most important thresholds, the F_p^V/B_p^V ratio is computed for all selected shadow pixels for threshold prediction. The F_p^V/B_p^V ratio is shown in Fig. 3.

According to the statistical results, we find that the F_p^V/B_p^V ratio will decrease along with the illuminance value increased. Furthermore, a linear function which is represented as formula (3) can be used to describe the general trend according to the results in Fig. 3,

$$y = a + b * x \quad (3)$$

where x is the illuminance value and y is a simulated ratio, while a and b are the coefficients of this linear function. According to the experiment results, a can be set to 1.113, and $-9.96 * 10^{-6}$ for b . In order to predict α and β dynamically, according to (1) and (3), it is suitable to design two similar linear formulas for α and β respectively.

$$\alpha = a_1 + b * x, \quad (4)$$

$$\beta = a_2 + b * x \quad (5)$$

From (1), it is obvious that α is smaller than β , so it is easy to draw the conclusion that a_1 is smaller than a_2 , since b is a negative number which we get from the experiment. In order to get the values of a_1 and a_2 , we transform (4) and (5) as follows:

$$a_1 = \alpha - b * x, \quad (6)$$

$$a_2 = \beta - b * x \quad (7)$$

To get the best a_1 and a_2 , 24 surveillance videos with various illuminance values are used for experiments, and the optimum values of α and β are set by manual for every video. Then, the average value of a_1 and a_2 are 1.015112, and 1.251021 respectively. And, the variance of a_1 and a_2 are 0.002987847 and 0.002447514 respectively, which means both a_1 and a_2 are stable. We use the average values as the final coefficients in (4) and (5), then the dynamic α and β can be predicted according to the following formulas:

$$\alpha = 1.015 - 9.96 * 10^{-6} * x, \quad (8)$$

$$\beta = 1.251 - 9.96 * 10^{-6} * x \quad (9)$$

Based on (8) and (9), α and β can be computed dynamically by using the current illuminance value of surveillance video.

4. Experiments

Two metrics proposed by Prati *et al* [6] was adopted for

shadow detection evaluation, which can be described as shadow detection rate (η) and shadow discrimination rate (ξ) and listed as follows:

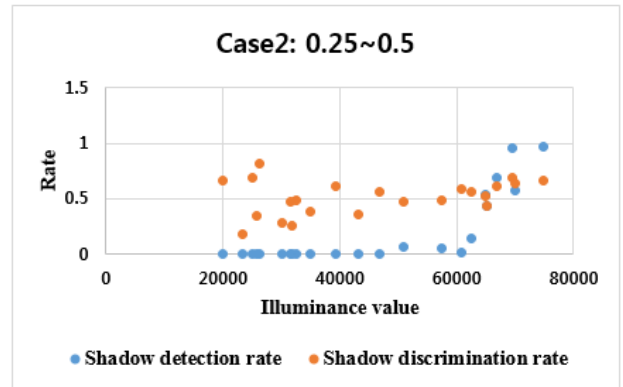
$$\eta = \frac{TP_S}{TP_S + FN_S} \text{ and } \xi = \frac{TP_F}{TP_F + FN_F} \quad (10)$$

Here TP and FN stand for true positive and false negative pixels while subscripts S and F correspond to the foreground and shadow respectively. For example, TP_F is the number of ground-truth pixels of the foreground objects minus the number of pixels detected as shadows, but belong to foreground objects. The shadow detection rate is concerned with labeling the maximum number of cast shadow pixels as shadows while shadow discrimination rate is concerned with maintaining the pixels that belong to the moving objects as foreground [2].

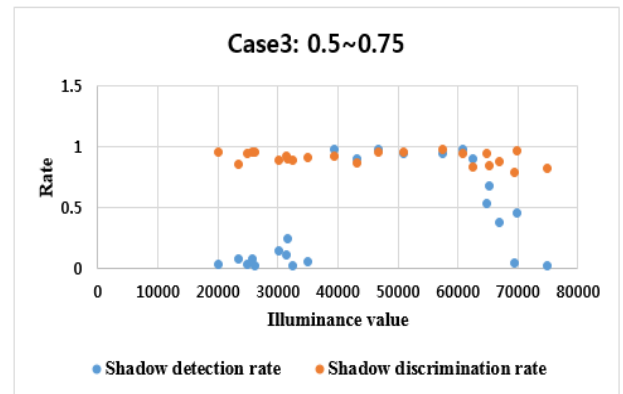
Table 1. 5 cases of static thresholds.

Static Thresholds	Value of α	Value of β
Case 1	0.0	0.25
Case 2	0.25	0.50
Case 3	0.50	0.75
Case 4	0.75	1.0
Case 5	1.0	1.25

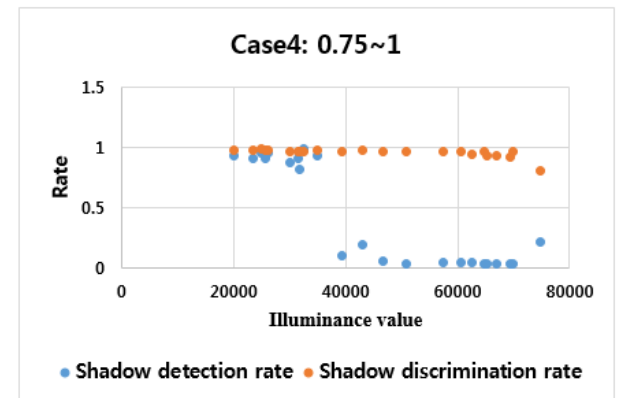
As the available public surveillance datasets do not include illuminance value, we collect the shadow data by ourselves as mentioned in Section 2. 260 images from 26 videos are chosen for experiment, and the average η and ξ are computed for every video. In order to compare our method with the static thresholds based methods, 5 different static α and β are adopted for comparison. The interval of and for static thresholds is 0.25. The detail of the static thresholds is shown in Table 1 while the corresponding final results of the comparison experiments are shown in Fig. 4.



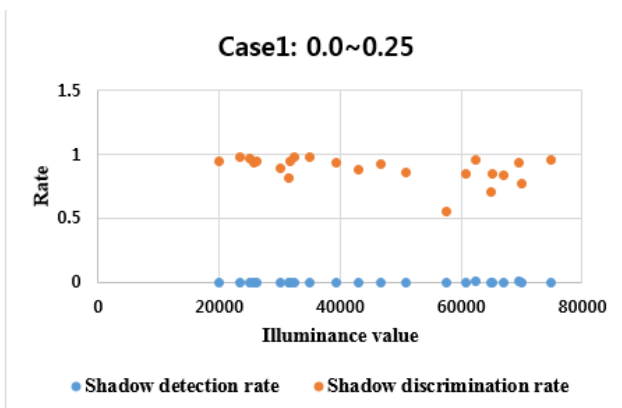
(b)



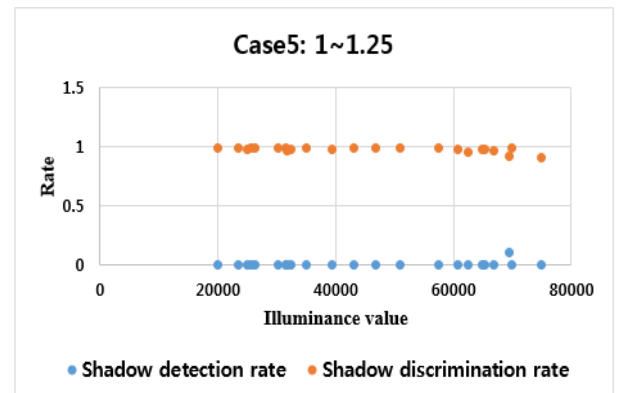
(c)



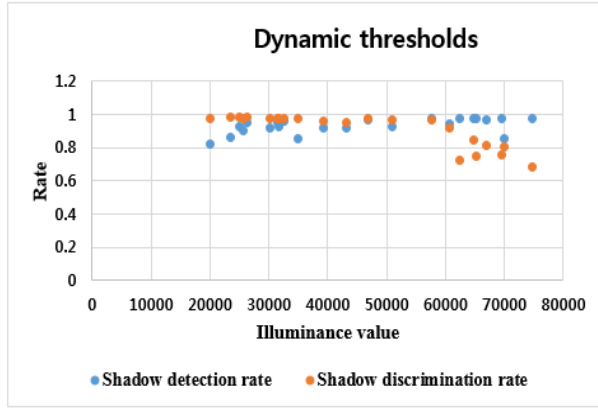
(d)



(a)



(e)



(f)

Figure 4. Results of the comparison experiments**Table 2.** Average η and ξ under different threshold cases.

Threshold Type	Average η (%)	Average ξ (%)
Dynamic thresholds	0.929928814	0.906924948
Case1	0.000429182	0.904265227
Case2	0.201845	0.514467318
Case3	0.394501273	0.909138045
Case4	0.4573625	0.958933318
Case5	0.006405455	0.978898273

We compare the performance of our proposed method with five cases of static thresholds, and the results are shown in **Table 2**. The average shadow detection rate and discrimination rate are computed on all videos for every threshold type. From the results, it's easy to draw the conclusion that only the dynamic thresholds can get both higher shadow detection rate and higher discrimination rate. In contrast, we can see it is very hard for static thresholds to always get both high shadow detection rate and shadow discrimination rate. For some illuminance values, static thresholds may result both high η and ξ , for example, in **Fig. 4(d)**, when the illuminance value is smaller than 3,500, the η and ξ are both high, however, if the illuminance value continually increased, the η begins to decline drastically.

Results of the proposed dynamic thresholds based method are shown in Fig. 4(f) which indicates when the illuminance is greater than 6,000, there also exists some slight decline of shadow discrimination rate. This is mainly because when the illuminance is very high, the color of objects has some impact on the shadow discrimination rate of chromaticity-based methods. For example, when the color of objects is very dark, some pixels belong to objects are likely to be detected as shadows as the shadow pixels are also very dark under high illuminance values. This can be illustrated in Fig. 5.



(a) The color of pedestrian's coat is similar to the shadow's color



(b) The color of pedestrian's coat is different with the shadow's color

Figure 5. Colors of pedestrian's coat are different with same illuminance value(69,500).

5. Conclusion & Discussion

A dynamic thresholds based shadow detection method is proposed in this paper, where a dynamic threshold can be predicted. Experiments on the self-built dataset suggest that proposed method is able to achieve higher shadow detection rate as well as higher shadow discrimination rate. By using the illuminance value, we can remove shadow automatically from surveillance videos. Comparing to static thresholds based methods, the proposed method is robust to illuminance changes which is very import for real applications.

For high illuminance situations, regions belong to objects may be detected as shadows which results in some decline of shadow discrimination rate. To solve this problem, texture features have already considered. the Local Binary Patterns (LBP) [9] was used in the experiments, but the result is not good. This is mainly because the shadow's texture structure is not obvious under high illuminance situations, and it is similar for object region when object's color is also very dark. This fact makes it hard to use texture features in shadow detection in the above situation. For the future work, we will try to find some other effective features to improve the shadow discrimination rate under high illuminance situations.

Acknowledgements

This research is supported by the MSIP(Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2015.

REFERENCES

- [1] Nadimi, S. and Bhanu, B., 2004, Physical models for moving shadow and object detection in video, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26 (8), 1088-1093.
- [2] Sanin, A., Sanderson, C., and Lovell, B. C., 2012, Shadow detection: A survey and comparative evaluation of recent methods, *Pattern Recognition*, 45 (4), 1684-1695.
- [3] Cucchiara, R., Grana, C., Piccardi, M., and Prati, A., 2003, Detecting moving objects, ghosts, and shadows in video streams, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 (10), 1337-1342.
- [4] Stauffer, C. and Grimson, W. E. L., 1999, Adaptive background mixture models for real-time tracking, *Proceedings of the IEEE Conference on Computer Vision and Patter Recognition*, CD-Format, Fort Collins, USA, June 23-25.
- [5] Maddalena, L. and Petrosino, A., 2008, A self-organizing approach to background subtraction for visual surveillance applications, *IEEE Transactions on Image Processing*, 17 (7), 1168-1177.
- [6] Prati, A., Cucchiara, R., Mikic, I., and Trivedi, M., 2001, Analysis and detection of shadows in video streams: a comparative evaluation, *Proceedings of the IEEE Conference on Computer Vision and Patter Recognition*, CD-Format, Kauai, Hawaii, USA, December 8-14.
- [7] Zhang, X., Liu, K., Wang, X., Yu, C., and Zhang, T., 2014, Moving shadow removal algorithm based on HSV color space, *Indonesian Journal of Electrical Engineering*, 12 (4), 2769-2775.
- [8] Leone, A. and Distanto, C., 2007, Shadow detection for moving objects based on texture analysis, *Pattern Recognition*, 40 (4), 1222-1233.
- [9] Zhang, L., Cheng, Y., Xie, Y. and Li, J., 2007, Shadow detection method in video sequences based on LBP, *Systems Engineering and Electronics*, 29 (6), 974-977.
- [10] Huerta, I., Holte, M., Moeslund, T., and Gonzalez, J., 2009, Detection and removal of chromatic moving shadows in surveillance scenarios, *Proceeding of the IEEE 12th International Conference on Computer Vison*, Kyoto, Japan, Sep. 29-Oct. 2.
- [11] Sanin, A., Sanderson, C., and Lovell, B., 2010, Improved shadow removal for robust person tracking in surveillance scenarios, *Proceeding of the 20th International Conference on Pattern Recognition*, Istanbul, Turkey, August 23-26.