

An Adaptive Cast Shadow Detection with Combined Texture and Color Models

Peiwen Liu and Yuesheng Zhu

Abstract—Cast shadows of moving foreground objects can cause miss tracking problem in object detection and tracking, thus shadow detection is an important step used after a moving foreground object is detected. Most of current methods have a significant trade-off between the shadow detection rate and the shadow discrimination rate. In this paper, an effective and adaptive method with combined texture and color models is proposed in order to achieve good shadow detection rate and shadow discrimination rate as well. Firstly, Scale Invariant Local Ternary Pattern (SILTP) is used to select a candidate shadow region. Then HSV color model is employed to detect a new candidate shadow region by using maximum likelihood estimation (MLE) to estimate the thresholds of HSV color model adaptively. Finally the two regions are combined by logical operation and a new shadow region can be obtained. Our experimental results show that the proposed method achieves a better performance in both shadow detection rate and discrimination rate compared to the other current methods. Moreover, the proposed method runs at 100 frames per second and is suitable for the real-time detection and tracking.

Index Terms—Shadow detection, SILTP, HSV color space, adaptive thresholds.

I. INTRODUCTION

Detection of moving foreground objects from video sequences is one of the most important parts in computer vision applications. Generally, the cast shadows would be detected as foreground objects since the shadows share the same movement patterns and have a similar magnitude of intensity change as that of the foreground object [1]. Since the shadow regions are sometimes as big as the object regions, the wrong classification as foreground objects can cause various unwanted consequence such as object shape distortion, false object detection and object merging, which has a bad impact on object detection and tracking. Therefore, shadow detection is a significant step for improving moving foreground objects detection and tracking.

There have been many methods focused on cast shadows detection. A survey and a comparative evaluation of recent techniques for moving cast shadow detection are presented in [2]. The methods of cast shadow detection covered in the survey are divided into four categories according to the features they are based on: chromaticity-based methods, physical-based methods, geometry-based methods and texture-based methods. The chromaticity-based methods are based on the assumption that shadow pixels become darker as

they are blocked from the illumination source but their chromaticity remains the same as the corresponding background pixels, which referred to as color constancy [3] or linear attenuation [4]. These properties have been employed in different colors spaces like HSV [5], YUV [6], normalized RGB [7] and c1c2c3 [8]. Most of these methods are simple to perform and their processing time is the fastest comparatively. However, they can not adapt to variant illumination well. Besides, they have the problem that the thresholds of brightness and chromaticity are set manually. The physical-based methods use physical-based color features to model or learn the specific appearance of shadow pixels [9], [10]. Though these methods tend to have a better performance than the chromaticity-based methods, they still limited to spectral properties. Furthermore, the training samples may be insufficient since the statistical models are learned from background surface variation under cast shadows. The geometry-based methods use the information of shadows to detect them, e.g. the orientation, the size and even the shape, which can be estimated with proper knowledge of the illumination source, object shape and the ground plane [11]-[13]. These methods lose effectiveness when the geometrical relationship changes since they depend on the geometrical relationship of the objects in the scenes. In addition to color properties, texture or gradient information extracted from the spatial domain is used to detect cast shadows. The texture-based methods are based on the fact that shadow regions remain the same texture as the corresponding background regions [14]-[16]. Since textures are greatly differentially and can adapt to variant illumination well, the texture-based methods are potential powerful for shadow detection. However, their processing time tend to be slow since their computation are expensive. What's more, if both the background regions and the moving objects have no texture, the texture-based methods will lose effectiveness.

Moreover, these methods have a big trade-off between the shadow detection rate and discrimination rate except the large region (LR) texture-based method [14]. However, the processing time of LR texture-based method is not fast enough to perform in real-time.

Considering the advantages and disadvantages of the chromaticity-based methods and the texture-based methods respectively, the chromaticity features and texture features can be combined together to detect shadow in order to achieve good performance and fast processing time.

In this paper, an adaptive shadow detection method which achieves good shadow detection rate and shadow discrimination rate is proposed. It is a mixture of the chromaticity-based method and the texture-based method. Firstly, Scale Invariant Local Ternary Pattern (SILTP) [17] is

Manuscript received July 15, 2013; revised September 20, 2013.

The authors are with the Communication & Information Security Lab, Shenzhen Graduate School, Peking University, China (e-mail: zhuys@pku.sz.edu.cn).

used to select a candidate shadow region. Secondly, HSV color model is adopted to detect a new candidate shadow region and maximum likelihood estimation (MLE) is employed to estimate the thresholds of HSV color model. Finally, the two candidate shadow regions are combined and the final shadow region is achieved.

The remainder of this paper is organized as follows. The proposed shadow detection method is introduced in details in Section II. Our experimental results and comparison with the recent methods for moving cast shadow detection is provided in the Section III. The conclusion is drawn in Section IV.

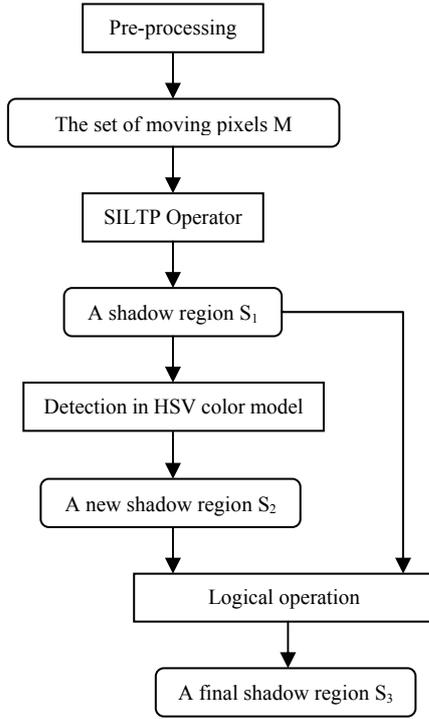


Fig. 1. Process of the proposed method.

II. PROPOSED METHOD

In this section, the local texture descriptor SILTP and the HSV color model are employed to detect shadows. The flowchart of the proposed method is presented in Fig. 1. First, the set of moving pixels is obtained by pre-processing. Second, SILTP is employed to detect the shadow region in the set of moving pixels, by which a shadow region S_1 can be gained. Third, the thresholds of HSV color model are achieved by maximum likelihood estimation (MLE) and a shadow region S_2 is detected by HSV color model. Finally, the above two shadow regions we have detected are synthesized by logical operation and get a final shadow region S_3 .

A. Pre-Processing

The set of moving pixels R is obtained by subtracting the background image from the current image. In the set of moving pixels, both the moving object pixels and shadow pixels are included. The intensity of the shadow pixel is smaller than the background pixel since its illumination source is blocked by the moving foreground object. Therefore, if a pixel from the set of moving pixels is lighter than its corresponding background pixel, it is defined as a

moving object pixel and is extracted from the moving pixels. With this procedure, computation can be reduced. Since the candidate set value (V) in HSV color model is a direct measure of intensity, it is used to estimate. The pixel is determined by

$$\begin{cases} p_F(x, y) \in M, & \text{if } p_F^V < p_B^V, \\ p_F(x, y) \in R - M, & \text{otherwise.} \end{cases} \quad (1)$$

where R is moving pixels obtained in first, M is the set of moving pixels after pre-processing, which is the processing data in the following steps, $R - M$ is the set of object pixels, $p_F(x, y)$ is the pixel p belongs to R in the frame (F), p_F^V and p_B^V are the value (V) in HSV color model for the pixel p in the frame (F) and background reference image (B), respectively.

B. SILTP

Since shadow regions retain most of their textures, SILTP can be used to detect the shadow based on the similarity of texture [15]. Given a pixel location $p(x, y)$, it is encoded by SILTP as

$$SILTP_{N,R}^\tau(x, y) = \bigoplus_{k=0}^{N-1} s_\tau(I_p, I_k) \quad (2)$$

where I_p is the gray intensity value of the center pixel p , I_k are that of its N neighborhood pixels equally spaced on a circle of radius R , \bigoplus denotes concatenation operator of binary strings, τ is a scale factor indicating the comparing range, and s_τ is a piecewise function defined as

$$s_\tau(I_p, I_k) = \begin{cases} 01, & \text{if } I_k > (1 + \tau)I_p, \\ 10, & \text{if } I_k < (1 - \tau)I_p, \\ 00, & \text{otherwise.} \end{cases} \quad (3)$$

SILTP operator has three advantages. Firstly, its computation is efficient since it causes only one more comparison than LBP for each neighbor. Secondly, it is robust to noises since it introduces a tolerative range τ like LTP. Thirdly, it has the property of scale invariance (“scale” here means gray scale pixel value, not spacial scale.) since each comparison can result in one of these three values which can be encoded in two bits (with “11” undefined). Therefore, the SILTP feature is invariant even when the illumination is suddenly changed, so SILTP is robust to illumination changes and can achieve better textures.

The set of moving pixels and the corresponding background pixels are encoded by SILTP and their similarity is estimated by hamming distance. We define that if the Hamming distance of SILTP between the moving pixels of a frame and the corresponding background is less than threshold λ_1 , then the pixel belongs to the set of shadow pixels. It should also consider the situation that heavy shadow regions under strong illumination may lose most of their texture, so we define that the pixel also belongs to the

set of shadow pixels if its SILTP is less than threshold λ_2 :

$$\begin{cases} p_F(x,y) \in S_1, & \text{if } h(\text{SILTP}(p_B(x,y)), \text{SILTP}(p_F(x,y))) < \lambda_1 \\ & \text{or } \text{SILTP}(p_F(x,y)) < \lambda_2, \\ p_F(x,y) \in M - S_1, & \text{otherwise.} \end{cases} \quad (4)$$

where $h(\cdot)$ is the hamming distance of SILTP between the moving pixels of a frame and the corresponding background pixels, λ_1 and λ_2 are the thresholds, S_1 and M is the first of shadow pixels and the set of moving pixels achieved by pre-processing, respectively.

C. HSV Color Model

The HSV color model provides a natural separation between chromaticity and luminosity. The intensity of shadow region is lower than that of object region, and the HSV color model can reflect this problem well since the value (V) is a direct measure of intensity. Following the chromaticity cues, shadow pixels does not change its hue (H) compared with the corresponding background pixels, but their saturation (S) will be lower [5]. Therefore, a pixel p is considered to be a shadow pixel if it satisfies the following three conditions:

$$\begin{cases} \beta_1 \leq p_F^V / p_B^V \leq \beta_2 \\ |p_F^H - p_B^H| \leq \alpha_H \\ (p_F^S - p_B^S) \leq \alpha_S \end{cases} \quad (5)$$

where $p_F^H, p_F^S, p_F^V, p_B^H, p_B^S, p_B^V$ express the component values of HSV for the pixel p in the frame (F) and background reference image (B) respectively. $\beta_1, \beta_2, \alpha_H$, and α_S are thresholds.

The thresholds of the HSV color model are optimized empirically [5]. Therefore, it is not sensitive to strong illumination changes and fails with strong shadows. In the proposed method, MLE is used to estimate the thresholds.

Light ratio R , the hue difference D^H and saturation difference D^S are defined as follows:

$$\begin{cases} R_p = p_F^V / p_B^V \\ D_p^H = p_F^H - p_B^H \\ D_p^S = p_F^S - p_B^S \end{cases} \quad (6)$$

where $p_F(x,y)$ belongs to the set moving pixels M in the frame and $p_B(x,y)$ is its corresponding background pixel.

We assume that the distributions of R_p, D_p^H , and D_p^S for the shadow pixel are Gaussian and each pixel's distribution is independent to each other. To simplify the expression, we use D_p^C ($C = H, S$) to stand for D_p^H and D_p^S .

Now that we have S_1 , the first candidate set of shadow pixels, so we can employ MLE to estimate the mean and standard deviation of R_p and D_p^C of the shadow pixels in

order to get a new set of shadow pixels S_2 . Therefore, we can obtain:

$$\begin{aligned} m_R^V &= (1 / N_{S_1}) \sum_{p(x,y) \in S_1} R_p \\ \sigma_R^V &= (1 / N_{S_1}) \sum_{p(x,y) \in S_1} (R_p - m_R^V)^2 \end{aligned} \quad (7)$$

$$m_D^C = (1 / N_{S_1}) \sum_{p(x,y) \in S_1} D_p^C, C = H, S$$

$$\sigma_D^C = (1 / N_{S_1}) \sum_{p(x,y) \in S_1} (D_p^C - m_D^C)^2, C = H, S$$

where N_{S_1} is the number of pixels in S_1 .

Then we can get the thresholds by:

$$\begin{cases} \beta_1^V = m_R^V + a \times \sigma_R^V \\ \beta_2^V = m_R^V - a \times \sigma_R^V \\ \alpha_1^C = m_D^C + a \times \sigma_D^C, C = H, S \\ \alpha_2^C = m_D^C - a \times \sigma_D^C, C = H, S \end{cases} \quad (8)$$

where $a = 1.645$ with reliability of 90%; $P(-1.645 < Z < 1.645) = 0.9, Z = N(0,1)$.

With the calculated thresholds, we can determine whether or not a pixel belongs to the set of shadow pixels by:

$$\begin{cases} p_F(x,y) \in S_2, & \text{if } \beta_1^V < R_p < \beta_2^V \\ & \text{and } \alpha_1^C < D_p^C < \alpha_2^C, C = H, S, \\ p_F(x,y) \in M - S_2, & \text{otherwise.} \end{cases} \quad (9)$$

where S_2 is a new set of shadow pixels.

D. Logical Operation

Now we have achieved two sets of shadow pixels, and we can get the synthetic result by:

$$\begin{cases} p_F(x,y) \in S_3, & \text{if } p_F(x,y) \in S_1 \text{ and } p_F(x,y) \in S_2, \\ p_F(x,y) \in M - S_3, & \text{otherwise.} \end{cases} \quad (10)$$

where S_3 is the final determined set of shadow pixels and $M - S_3$ is the set of object pixels.

E. Post-Processing

Some unconnected regions may appear after logical operation and this may results in two types of errors: shadow detection failure and object detection failure. Shadow detection failure means that some shadow pixels are misclassified as object pixels and object detection failure means that some shadow pixels are misclassified as object pixels. Hence, morphological operations and connection operation are employed among these unconnected regions to increase the accuracy of shadow detection.

III. EXPERIMENTAL RESULTS

In this section, the proposed method is compared with the recent methods in terms of qualitative observations as well as quantitative performance measures. These methods are

chromaticity-based method [5], geometry-based method [13], physical method [4], small region (SR) texture-based method [15], and large region (LR) texture-based method [16]. We

applied all these methods to five test sequences (Room, Lab, Campus, Hallway, and CAVIAR) which have been commonly used in shadow detection.

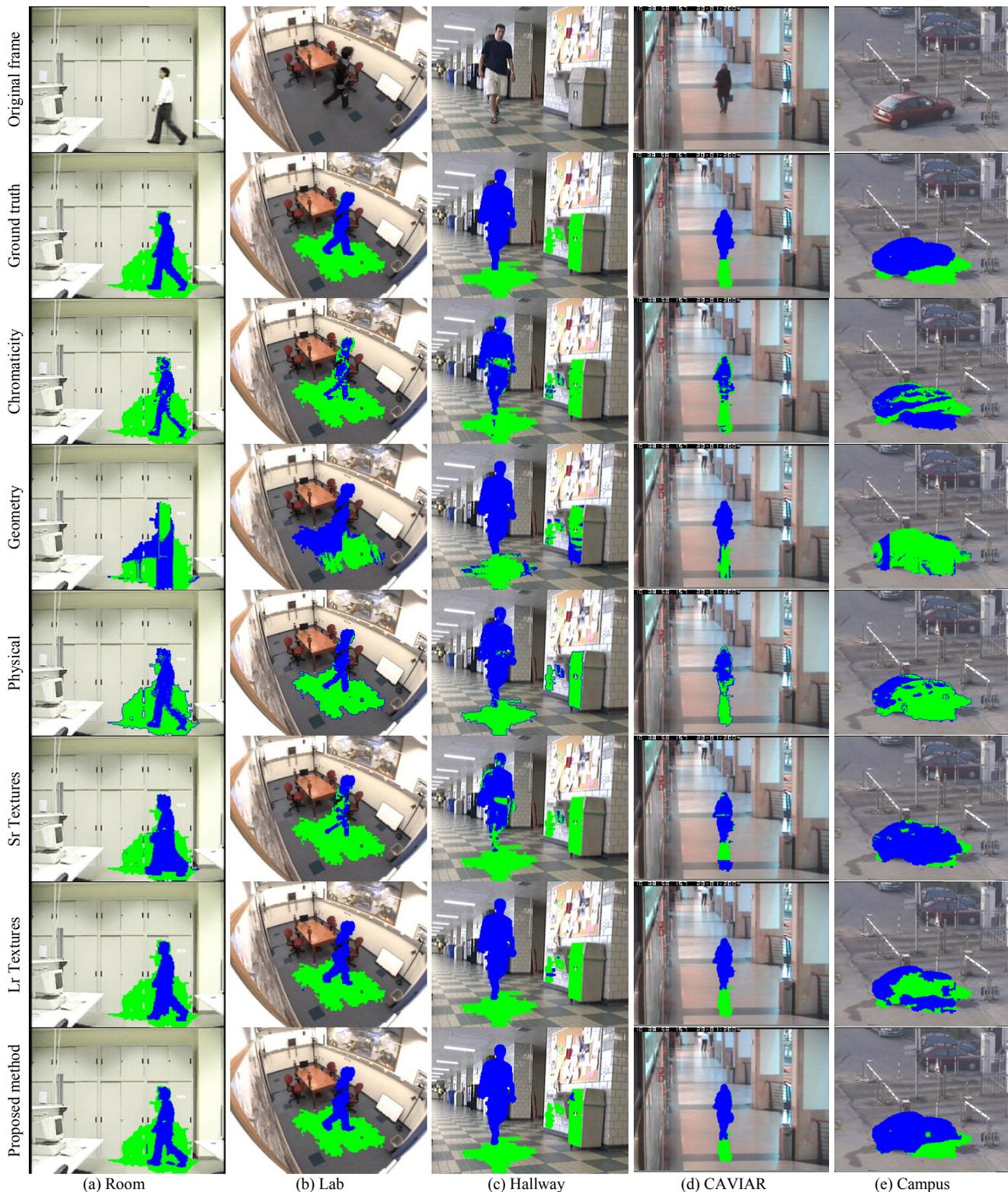


Fig. 2. Detection results of different method.

A. Qualitative Results

The qualitative results are shown in Fig. 2. An example frame for each sequence is shown in the first row. The ground truth images are shown in the second row and the results of different methods are shown in the rest row, where object pixels are marked in blue and shadow pixels are marked in green. From the qualitative results we can see that proposed

method has improvement in both the performance and robustness.

B. Quantitative Results

The quantitative evaluation follows the method proposed by Prati *et al.* [18]. The shadow detection rate η and shadow discrimination rate ξ are defined as follows:

$$\eta = TP_S / (TP_S + FN_S) \quad \xi = TP_F / (TP_F + FN_F) \quad (11)$$

where the subscript S stands for shadow and F for foreground. TP and FN stand for true positive and false negative respectively.

The quantitative results are shown in Fig. 3. It can be noted that the proposed method performs better than the other methods in average, achieving a high value in both shadow detection rate and discrimination rate.

TABLE I: AVERAGE FRAME PROCESSING TIME (IN MILLISECONDS) PER SEQUENCE FOR VARIOUS SHADOW DETECTION METHODS

	Chromaticity	Geometry	Physical	SR textures	LR textures	Proposed method
Room	6.94	8.15	8.26	39.43	15.46	7.65
Lab	8.70	17.08	14.59	65.86	21.71	9.65
Hallway	10.88	8.63	12.25	61.43	20.54	12.11
CAVIAR	10.48	13.26	13.45	67.48	23.60	11.42
Campus	8.36	9.00	9.64	41.03	19.70	9.24
Average	9.07	11.23	11.64	55.04	20.20	10.01

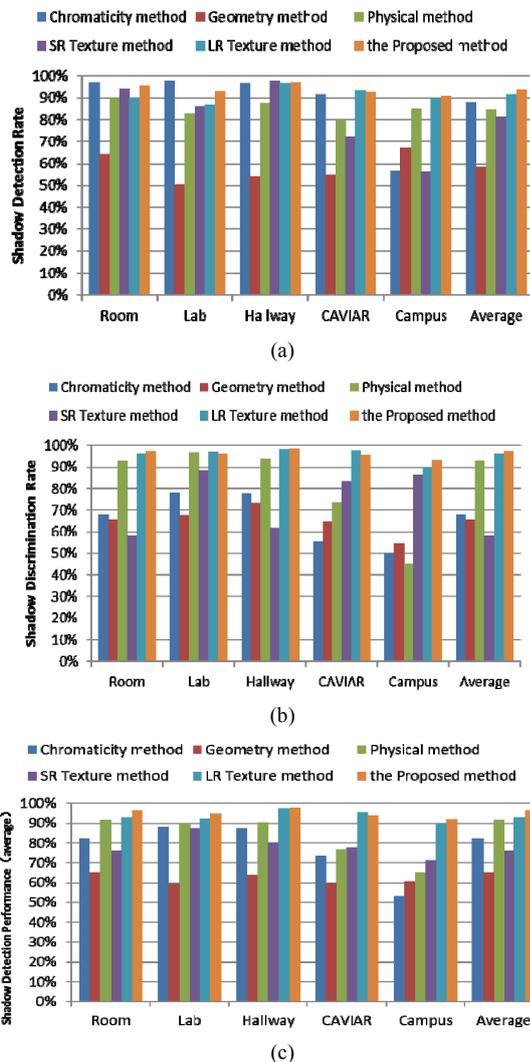


Fig. 3. The quantitative results. The bars represent chromaticity method, geometry method, physical method, SR texture method, LR texture method and the proposed method in sequence. (a) Shadow Detection Rate. (b) Shadow Discrimination Rate. (c) Average of shadow detection rate and shadow discrimination rate.

IV. CONCLUSION

In this paper, an effective and adaptive method of shadow

detection method is shown in Table I. The processing time is achieved by a 32-bit Intel CPU running at 2.8 GHz. Since the code of the proposed method is optimized, it takes only a little bit longer time to process compared with the chromaticity method. It is the second quickest method in these methods and it processes fast enough to be used in surveillance scenarios.

detection method is proposed. The proposed method detects the set of shadow pixels from the moving pixels by using SILTP and HSV color model. The advantages of the method are as follows: Firstly, it achieves higher shadow detection rate and discrimination rate than the existed methods in average. Secondly, it is adaptive and robust to illuminate variation in the environment since the thresholds of HSV color model are estimated automatically without manual setting. Finally, it can be used in surveillance system since it runs at 100 frames per second.

REFERENCES

- [1] S. Nadimi and B. Bhanu, "Physical models for moving shadow and object detection in video," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1079-1087, Aug. 2004.
- [2] A. Sanin, C. Sanderson, and B. C. Lovell, "Shadow detection: A survey and comparative evaluation of recent methods," *Pattern recognition*, vol. 45, no. 4, pp. 1684-1695, Apr. 2012.
- [3] T. Horprasert, D. Harwood, and L. S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *Proc. IEEE International Conference on Computer Vision*, 1999, pp. 1-19.
- [4] J. B. Huang and C. S. Chen, "Moving cast shadow detection using physics-based features," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 2310-2317.
- [5] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1337-1342, Oct. 2003.
- [6] O. Schreer, I. Feldmann, U. Golz, and P. Kauff, "Fast and robust shadow detection in videoconference applications," in *Proc. Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom*, 2002, pp. 371-375.
- [7] A. Cavallaro, E. Salvador, and T. Ebrahimi, "Shadow-Aware object-based video processing," in *Vision, Image and Signal Processing, IEE Proceedings*, 2005, pp. 398-406.
- [8] E. Salvador, A. Cavallaro, and T. Ebrahimi, "Cast shadow segmentation using invariant color features," *Computer vision and image understanding*, vol. 95, no. 2, pp. 238-259, Aug. 2004.
- [9] A. J. Joshi and N. P. Papanikolopoulos, "Learning to detect moving shadows in dynamic environments," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 2055-2063, Nov. 2008.
- [10] N. M. Brisson and A. Zaccarin, "Learning and removing cast shadows through a multidistribution approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1133-1146, July 2007.

- [11] C. C. Chen and J. Aggarwal, "Human shadow removal with unknown light source," in *Proc. 2010 20th International Conference on Pattern Recognition (ICPR)*, 2010, pp. 2407-2410.
- [12] L. Z. Fang, W. Y. Qiong, and Y. Z. Sheng, "A method to segment moving vehicle cast shadow based on wavelet transform," *Pattern Recognition Letters*, vol. 29, no. 16, pp. 2182-2188, Dec. 2008.
- [13] J. W. Hsieh, W. F. Hu, C. J. Chang, and Y. S. Chen, "Shadow elimination for effective moving object detection by Gaussian shadow modeling," *Image and Vision Computing*, vol. 21, no. 6, pp. 505-516, June 2003.
- [14] A. Sanin, C. Sanderson, and B. C. Lovell, "Improved shadow removal for robust person tracking in surveillance scenarios," in *Proc. 2010 20th International Conference on Pattern Recognition (ICPR)*, 2010, pp. 141-144.
- [15] A. Leone and C. Distante, "Shadow detection for moving objects based on texture analysis," *Pattern Recognition*, vol. 40, no. 4, pp. 1222-1233, Apr. 2007.
- [16] Y. L. Tian, M. Lu, and A. Hampapur, "Robust and efficient foreground analysis for real-time video surveillance," in *Proc. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 1182-1187.
- [17] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikainen, and S. Z. Li, "Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes," in *Proc. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 1301-1306.
- [18] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, "Detecting moving shadows: algorithms and evaluation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 918-923, July 2003.



Peiwen Liu received BSc degree in 2011 from Beijing University of posts and telecommunications. She has studied in the communication and information Security Lab, Shenzhen Graduate School, Peking University, Shenzhen, China since 2011 for the MSc degree. Her research interests are computer vision and multimedia technology.



Yuesheng Zhu received his B.Eng. degree in radio engineering, M. Eng. degree in circuits and systems and Ph.D. degree in electronics engineering in 1982, 1989 and 1996, respectively. He is currently working as a professor at the Lab of Communication and Information Security, Shenzhen Graduate School, Peking University. He is a senior member of IEEE, fellow of China Institute of Electronics, and senior member of China Institute of Communications. His interests include digital signal processing, multimedia technology, communication and information security.