

Moving Cast Shadow Detection Using Joint Color and Texture Features with Neighboring Information

Bingshu Wang¹, Wenqian Zhu¹, Yong Zhao^{1(✉)}, and Yongjun Zhang²

¹ School of Electronic and Computer Engineering,

Shenzhen Graduate School of Peking University, Shenzhen, China

{wangbingshu,zhuwenqian}@sz.pku.edu.cn, zhaoyong@pku.sz.edu.cn

² College of Computer Science and Technology, Guizhou University, Guiyang, China

zyj6667@126.com

Abstract. Moving cast shadow detection is an important technique that increases the rate of accuracy in object detection. In this paper, we introduce a moving cast shadow detection method that is based on the assumption that the shadow regions are darker than the corresponding background regions and maintain the same chromaticity and texture. The proposed algorithm includes two main stages. The first is to detect candidate shadows by spectral ratio at pixel-level. The other is to extract shadows by jointly using three components of HSV chromaticity, improved local ternary pattern and the gradient for each pixel. Each color or texture component makes use of a pixel's neighboring information and comprises a single result. Three such detected results are combined to determine whether a pixel belongs to a shadow. Experimental results show that the method outperforms some state-of-the-art algorithms on a benchmark dataset of indoor and outdoor scene sequences.

Keywords: Moving cast shadow · HSV chromaticity · Improved local ternary pattern · Gradient · Neighboring information

1 Introduction

Moving object detection is core to a wide variety of visual surveillance applications including perimeter protection, content-based object indices, and behavior analysis. Accurate detection of moving objects is a constant task of image processing. Moving cast shadows—integral to this requirement—are frequently misclassified as moving objects because they share both the similar motion properties and background discrimination with objects. Such erroneous misclassification may affect geometrical properties and trigger distortion of real objects, encumbering effective processing of subsequent vision algorithms such as recognition and tracking. In addition, a shadow cast onto an object may increase the probability of that object being obscured. A moving cast shadow is generated by light source occlusion by an opaque object, making the shadow darker

than its respective covered background while retaining similar color and texture. A shadow is always coupled with the object that cast it and its motion behavior is similar to that object [1].

Diverse shadow detection methods [2–6], such as chromaticity-based methods and texture-based methods, have been proposed in recent years to address multiple moving cast shadow scenarios. Some evaluation metrics, such as shadow detection rate and shadow discrimination rate, have also materialized to evaluate shadow detection methods [1]. In addition, a suite of benchmark test datasets, including indoor and outdoor sequences, has also emerged [7].

The principle contribution of this paper is improvement of shadow detection rate by using a novel approach. Statistical data for shadows demonstrates insignificant change in the color inter-component ratio of shadow pixel point and respect to background pixel point. This results in the detection of all shadows with a modicum of misclassified objects. Following apprehension of candidate shadows, color consistency and texture invariance—including HSV color, improved local ternary pattern and gradient information—are taken into account and undergo rigorous development. Coupled with neighboring information utilization, pixel-level analysis for candidate shadows is performed. Experimental results validate the efficiency of this method and confirm that it outperforms some existing state-of-the-art methods.

In the following section, related works involving several state-of-the-art methods will be discussed. After that, Sect. 3 provides details relevant to the current method’s procedure with joint color and texture features. Section 4 then shows the qualitative and quantitative results. The conclusion will be presented in Sect. 5.

2 Related Works

In the past decades many shadow detection algorithms are proposed and classified as different taxonomies. A two-layer classification is given in [1]. The first layer includes two parts: statistical and deterministic. The statistical class can be subdivided into parametric and non-parametric methods and the deterministic class is further divided into model-based methods and non-model based methods. Deterministic non-model based approach shows the best results for a general-purpose shadow detection by using the HSV color space.

Spectral, spatial and temporal domains are the source of features in methods. Each domain can be subdivided into many algorithms by features which are reckoned as a better taxonomy [3]. Intensity, chromaticity and physical properties belong to spectral domain. Geometry and texture are parts of spatial features. Chromaticity method [8], geometry method [9], physical method [10], small region (SR) texture-based method [11] and large region (LR) texture-based method [12] are common and basic methods. Chromaticity method is based on the assumption that shadow regions are darker but keep the chromaticity, for instance, HSV-based chromaticity method. Geometry method assumes that each blob consists of object and shadow and makes a distinction. Physical method

models the appearance of shadow pixels and some methods try to establish an attenuation model. There are many texture-based methods such as SR and LR. SR considers more about the region-level correlation while LR adopts a method that distinguishes a blob whether it belongs to a shadow. Each method makes use of single property, which may achieve a good result in special scene. The proposed method comprehensively takes chromaticity and texture into account in order to adjust to various scenes.

The main difference between proposed method and the existing methods is that we use two color spaces (RGB and HSV) and combine three shadow detectors: HSV detector, improved local ternary pattern detector and gradient detector simultaneously. The existing methods utilize one or two cues for moving cast shadow detection. Due to the multiple features are utilized, robustness is improved. Experimental results in Sect. 4 show that our method's better performance.

There is a dominant strategy including two steps for moving cast shadow detection. The first step is to extract the candidate regions that contains all the possible shadows, and the second step is to achieve an accurate detection and discrimination. The principle is how to narrow the range of candidate shadows and keep real shadows as much as possible. Our method adopts the strategy, which has been proved efficient for outperforming five methods under the comparison of the standard evaluation dataset.

3 Shadow Detector Method

This paper puts forward to a framework for shadow detection. Two color spaces including RGB and HSV are employed. From the viewpoint of effectiveness, detection result of two color spaces is better than that of one color space. For each pixel, as long as the values of its three channels are lower than those of its background pixel simultaneously, it can be regarded as a candidate shadow pixel. So all real shadow pixels cannot be lost.

Meanwhile, some objects even noise that meet the standard will also be included. Then the detector based on HSV color space can build on the precedent work. Two kinds of texture features are also utilized to extract accurate shadows from candidate shadows and each of which is operated at pixel-level with the consideration of neighboring information. This method principle is shown in Fig. 1 and the objects in the framework are defined as follows:

- Frame, Foreground and Background represent current image, detected foreground image and background image respectively in sequences. Detected foreground image includes the moving objects and moving cast shadows.
- S1 is the candidate shadows detected by spectral ratio of RGB color space in the subsequent Sect. 3.1
- S2, S3 and S4 are the shadow detection results by using HSV detector, improved local ternary pattern (ILTP) detector and gradient detector respectively.
- S5 is the union result of S2, S3 and S4 though a voting process.

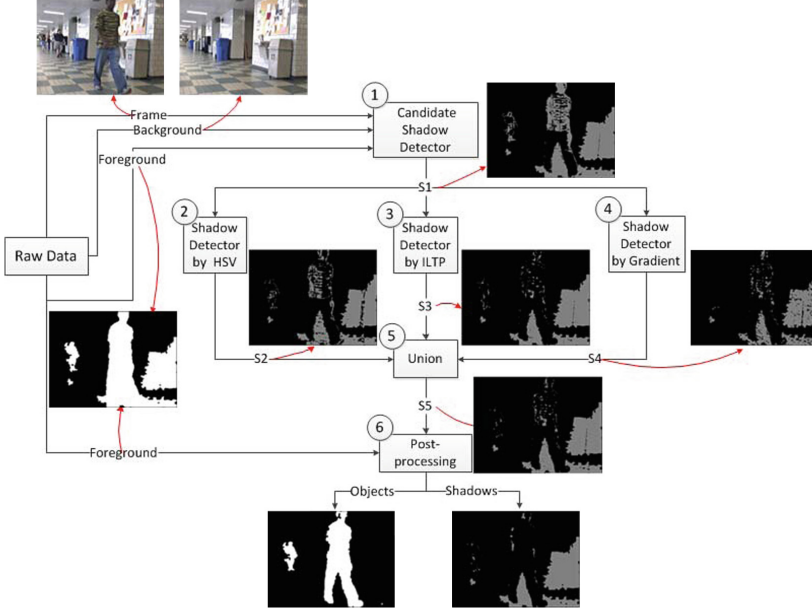


Fig. 1. Flowchart of shadow detection with joint color and texture information.

As shown in Fig. 1, the proposed shadow detection method consists of the following six steps.

1. Obtain the candidate shadows $S1$ from raw data.
2. Use HSV detector to detect shadows and gain $S2$ from $S1$.
3. Use ILTP detector to detect shadows and gain $S3$ from $S1$.
4. Use gradient detector to detect shadows and gain $S4$ from $S1$.
5. Obtain the union shadows $S5$ from $S2$, $S3$ and $S4$.
6. Use Foreground (from raw data) and $S5$ to generate the final Objects image and Shadows image.

These six steps are detailed in the following sections.

3.1 Candidate Shadow Detector by Spectral Ratio

The paper shares the typical assumption that the shadow region is darker than corresponding background, which helps to gain the preliminary candidate shadows. For each pixel, as long as all the channels' values (RGB) are lower than those of corresponding background pixel, it can be treated as a coarse candidate shadow pixel. Inevitably, some objects' or noisy pixels may be included. Experiments on benchmark images validate an assumption that the shadow pixel's ratio between the components of RGB color space does not change significantly

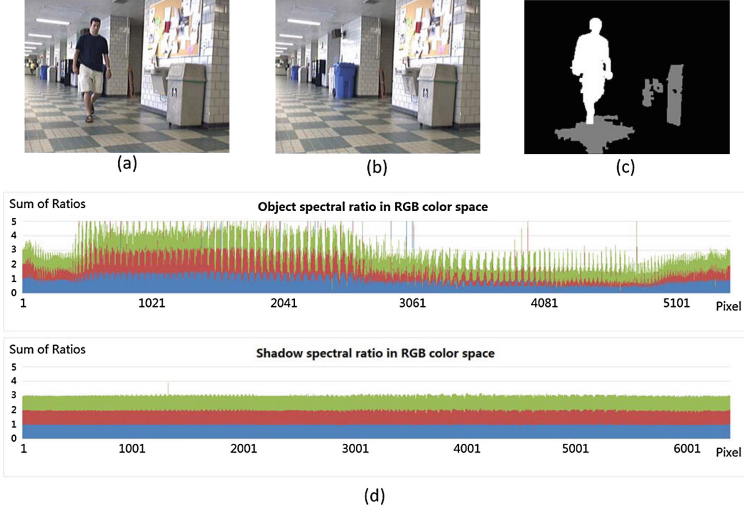


Fig. 2. Object-background spectral ratio and Shadow-background spectral ratio in RGB color space. Image (a) is current frame, (b) is the background image, (c) is the ground truth image, (d) is the statistic spectral ratio for all the object pixels and shadow pixels derived from (c). In (c) white pixels belong to object class while gray pixels belong to shadow class (Color figure online).

when comparing to its background pixel at the same position (see Fig. 2.). This can be applied to detect candidate shadows.

$$\Psi_r = \frac{F_b/F_g}{B_b/B_g}, \Psi_g = \frac{F_b/F_r}{B_b/B_r}, \Psi_b = \frac{F_g/F_r}{B_g/B_r}. \quad (1)$$

$$BGR_Candidate_Shadow = \begin{cases} 1 & \text{if } |\Psi_i - \mu| < \lambda \quad i \in \{b, g, r\} \\ 0 & \text{otherwise} \end{cases}. \quad (2)$$

Here F_i ($i \in \{r, g, b\}$) and B_i represent color component of current frame pixel and background pixel, while Ψ_i is spectral ratio and close to one in shadow regions but not necessarily in object regions. Shadows can be distinguished from objects by using Eq. (2). In shadow regions Ψ_i value changes little while in object regions it has a wide range of values. The parameter μ is a reference value that revolves one around and λ is a smaller value (less than 0.2).

3.2 Shadow Detector by HSV Color Space

Many shadow detection methods are based on HSV color space because the color space enhances the discrimination between objects and shadows [13]. It consists of two parts: color chromaticity and color brightness. We choose color

chromaticity with ignore of the brightness condition due to the use in the pre-selection stage, otherwise it will be repetitive.

$$HSV_Shadow = \begin{cases} 1 & \text{if } \frac{\sum_{i=1}^n |F_i^h - B_i^h|}{n} < \tau_h \wedge \frac{\sum_{i=1}^n (F_i^s - B_i^s)}{n} < \tau_s . \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

The parameters (τ_s, τ_h) represent the empirical thresholds varying from different scenes. Single pixel is isolated and opt to be a bit more arbitrary. When its neighboring information is made use of, the interference caused by uncertain factors such as sudden light changing tends to be reduced. The average difference value of a small region that centered in current pixel (see Eq. (3)) is calculated and used to determine whether current pixel belongs to a shadow class or not.

3.3 Shadow Detector by ILTP

Local binary pattern (LBP) was first introduced as an excellent means to describe local gray-level structure. However, LBP is easily affected by noise especially in the neighborhood-uniform regions. Tan modified LBP and proposed local ternary pattern (LTP) [14]. The difference between LBP and LTP is just from 2-valued to 3-valued codes, which improve the resistance and robustness to noise in many scenes.

$$\nu(x, i_c, t) = \begin{cases} 1, & x \geq i_c + t \\ 0, & |x - i_c| < t \\ -1, & x \leq i_c - t \end{cases} \quad (4)$$

A local neighborhood around each pixel is taken by LTP operator in gray image. The way of LTP code is given in Eq. (4), and x represents the pixel that around centered pixel i_c and t indicates the tolerance to noise.

The LTP operator compares neighboring pixels with centered pixel but ignores the potential information between the neighboring pixels. Actually, the addition of comparison information of neighboring pixels can enhance the integrity of local texture pattern. The ILTP encoding procedure is illustrated in Fig. 3 that considers 3×3 neighborhoods. For one pixel, it is coded with 12-value in single channel. One way to enrich more information of ILTP is to take all the three channels of RGB color space into account.

Here, we denote $C(F_i(x))$ as the code-value for one pixel of current frame with i -th neighboring pixel and $C(B_i(x))$ as code-value of corresponding background pixel. The similarity between them is represented by $x(i)$. It will be assigned to one if $C(F_i(x))$ shares same value with $C(B_i(x))$, otherwise zero. We denote n as the sum of neighboring pixels and δ as the similarity between the current pixel and background pixel at the same position.

$$ILTP_Shadow = \begin{cases} 1 & \text{if } \sum_{i=1}^n x(i)/n > \delta \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

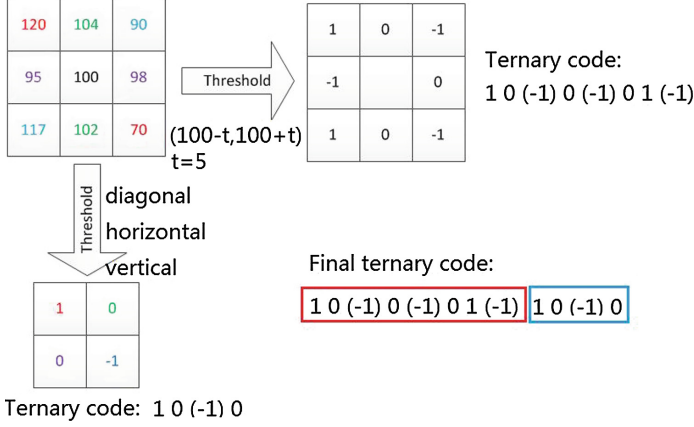


Fig. 3. Improved local ternary pattern for shadow detection.

3.4 Shadow Detector by Gradient

LR method places more importance on gradient information [12]. Like that, ∇_y is denoted as the vertical gradient (difference intensity between pixel in previous row and the pixel in next row) and ∇_x as the horizontal gradient (difference intensity between pixel in previous column and the pixel in next column). Meanwhile, ∇_i and θ_i are defined as gradient magnitude and direction respectively.

$$\nabla_i = \sqrt{\nabla_x^2 + \nabla_y^2}, \theta_i = \arctan\left(\frac{\nabla_y}{\nabla_x}\right). \quad (6)$$

$$Gradient_Shadow = \begin{cases} 1 & \text{if } \frac{\sum_{i=1}^n \sum_{j \in \{b, g, r\}} |F(\nabla_i^j) - B(\nabla_i^j)|}{n} < \phi_m \wedge \frac{\sum_{i=1}^n \sum_{j \in \{b, g, r\}} |F(\theta_i^j) - B(\theta_i^j)|}{n} < \phi_d \\ 0 & \text{otherwise} \end{cases}. \quad (7)$$

Specially, ϕ_m and ϕ_d are defined as deciding thresholds of gradient magnitude difference and direction difference respectively. The neighborhood pixels are exploited to help construct the gradient texture. This is a small region centered to current pixel. The gradient magnitude and direction in each pixel's small region in the foreground are associated with the corresponding pixel's small region in the background. Eq. (7) shows the detailed description of the gradient difference and direction difference simultaneously between the current frame and the corresponding background frame.

3.5 Union and Post-processing

A simple and typically voting mechanism is proposed to decide whether one pixel belongs to shadow class, which is similar with the best-of-three-games.

For each pixel, it will be regarded as a shadow pixel if there are at least more than 2 (including 2) detected shadow results in S2, S3 and S4. So the union shadow S5 will be obtained, and then post-processing is needed for the sake of achieving a preferable object image and shadow image. The post-processing includes open/close morphology processing and filling for small holes. Finally, the shadow will be segmented properly.

4 Experimental Results

4.1 Benchmark Sequences and Evaluation Metrics

The benchmark sequences were introduced in [1, 15] and six typical scenes and sequences [7] were chosen, which consisted of indoor and outdoor scenes with different shadow detection challenges and the ground truth sequences were segmented manually.

$$\eta = \frac{TP_S}{TP_S + FN_S}, \xi = \frac{TP_F}{TP_F + FN_F}. \quad (8)$$

$$F\text{-measure} = \frac{2\eta\xi}{\eta + \xi}. \quad (9)$$

Shadow detection (η) and shadow discrimination rate (ξ) are two metrics (Eq. (8)) widely used to evaluate the performance of shadow detection methods. The comprehensive and cogent evaluate metric is denoted as *F-measure* (Eq. (9)). We denote TP_S and FN_S as the true positive and false negative pixels of detected shadows, TP_F and FN_F as the true positive and false negative pixels of detected objects.

4.2 Qualitative and Quantitative Results

Qualitative results on six sequences are given in Fig. 4 with the straightforward detected shadows and objects. Considering the different scenes, we set two groups of parameters: $\tau_s = 28, \tau_h = 40, \delta = 0.5, \phi_m = 15, \phi_d = 1.2$ and $\tau_s = 20, \tau_h = 40, \delta = 0.55, \phi_m = 15, \phi_d = 1.1$ for indoor scenes and outdoor scenes. The parameters are empirical and a systematic method to generate them needs to be invested in future. Specially, the *Campus* scene uses parameters for indoor scenes due to its weak shadow.

As a whole, the proposed method shows good performance comparing with other five methods: chromaticity method [8], geometry method [9], physical method [10], SR method [11] and LR method [12]. Quantitative results are presented in Table 1 with six kinds of sequences. Outdoor scenes such as *Highway1* and *Highway2* are challenging because of the very less chromaticity and texture information. It is inevitable that some points are easily misclassified as shadows especially at the objects' edges. It is easy to cause the occurrence of false positives by the air disturbance of edges, which results in the thinner detected objects.

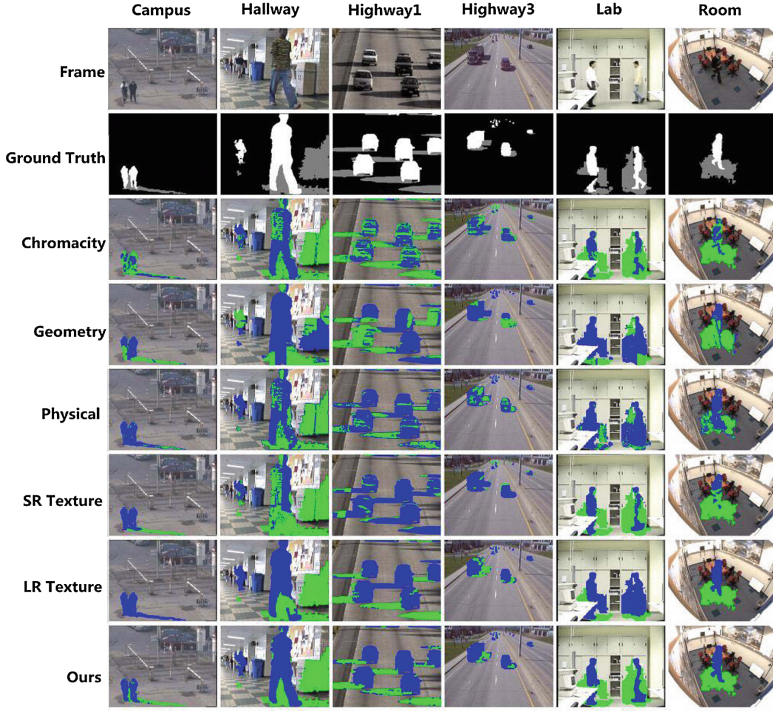


Fig. 4. The qualitative results of shadow detection. The green regions are the shadows and the blue regions are the objects (Color figure online).

Table 1. The η measure evaluation of six methods.

Methods	Campus	Hallway	Highway1	Hallway3	Lab	Room
Chromaticity	0.5386	0.9356	0.7601	0.4508	0.9949	0.9662
Geography	0.6085	0.4866	0.6616	0.4273	0.4533	0.5477
Physical	0.459	0.5608	0.4247	0.3628	0.2428	0.5802
SR	0.5537	0.9609	0.1692	0.0577	0.8151	0.9338
LR	0.5039	0.9508	0.6046	0.3808	0.876	0.818
Ours	0.7359	0.9562	0.6857	0.4192	0.9109	0.9411

It can be found out that Fig. 5 shows the advantages of the proposed method. Results of indoor scenes are better than outdoor scenes due to the abundant chromaticity and texture information provided.

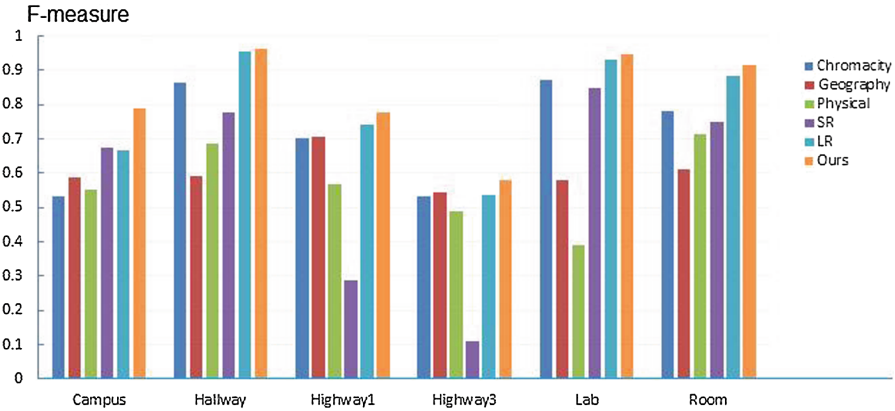


Fig. 5. The comparison of shadow detection results with θ metric by six methods.

5 Conclusion

This paper presents a novel method for moving cast shadow detection by using color and texture information. The innovative contribution is that we combine color information with texture features: two color spaces including RGB and HSV and two kinds of textures including ILTP and gradient. We adopt a two-step strategy for shadow detection: the first step is to detect candidate shadows and the second is to extract real shadows on the basis of first step. The pixel's spectral ratio is fit for efficient candidate shadow detection, and experiments validate it.

For all three detectors of HSV, ILTP and Gradient, neighboring information are utilized to supplement and classify pixel points of candidate shadows. Then a simple voting strategy is adopted to obtain a comprehensive shadow result. Experimental results on both indoor and outdoor scenarios show that our method gains a higher shadow detection rate. For all the tested sequences our method is average 4% larger than LR method at *F-measure* evaluation metric. Specially, the *F-measure* result of campus scene is significant, which is about 12% larger than any compared method. However, the challenges such as hard-shadows exist. Further research under strong light condition needs to be done, especially in outdoor scenes. In addition, an automatic mechanism of obtaining robust and systematical parameters could be invested in the next work.

References

1. Prati, A., Mikic, I., Trivedi, M.M., Cucchiara, R.: Detecting moving shadows: algorithms and evaluation. *J. IEEE Trans. Pattern Analy. Mach. Intell.* **25**(7), 918–923 (2003)
2. Al-Najdawi, N., Bez, H.E., Singhai, J., Edirisinghe, E.A.: A Survey of cast shadow detection algorithms. *J. Pattern Recogn. Lett.* **33**(6), 752–764 (2012)

3. Sanin, A., Sanderson, C., Lovell, B.C.: Shadow detection: a survey and comparative evaluation of recent methods. *J. Pattern Recogn.* **45**(4), 1684–1695 (2012)
4. Jiang, K., Li, A.H., Cui, Z.G., Wang, T., Su, Y.Z.: Adaptive shadow detection using global texture and sampling deduction. *J. IET Comput. Vis.* **7**(2), 115–122 (2013)
5. Zhang, W., Fang, Z.Z., Yang, X.K., Wu, Q.M.J.: Moving cast shadows detection using ratio edge. *J. IEEE Trans. Multimedia* **9**(6), 1202–1214 (2007)
6. Qin, R., Liao, S.C., Lei, Z., Li, S.Z.: Moving cast shadow removal based on local descriptors. In: *IEEE 20th International Conference on Pattern Recognition*, pp. 1377–1380. IEEE Press, New York (2010)
7. The Test Sequences and Ground Truth. <http://arma.sourceforge.net/shadows/>
8. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting moving objects, ghosts, and shadows in video streams. *J. IEEE Trans. Pattern Anal. Mach. Intell.* **25**(10), 1337–1342 (2003)
9. Hsieh, J.W., Hu, W.F., Chang, C.J., Chen, Y.S.: Shadow elimination for effective moving object detection by gaussian shadow modeling. *J. Image Vis. Comput.* **21**(3), 505–516 (2003)
10. Huang, J.B., Chen, C.S.: Moving cast shadow detection using physics-based features. In: *2009 IEEE Conference Computer on Vision and Pattern Recognition*, pp. 2310–2317. IEEE Press, New York (2009)
11. Leone, A., Distanti, C., Buccolieri, F.: Shadow detection for moving objects based on texture analysis. *J. Pattern Recogn.* **40**(4), 1222–1233 (2007)
12. Sanin, A., Sanderson, C., Lovell, B.C.: Improved shadow removal for robust person tracking in surveillance scenarios. In: *IEEE 20th International Conference on Pattern Recognition*, pp. 141–144. IEEE Press, New York (2010)
13. Cucchiara, R., Grana, C., Piccardi, M., Prati, A., Sirotti, S.: Improving shadow suppression in moving object detection with HSV color information. In: *2001 IEEE Conference on Intelligent Transportation Systems*, pp. 334–339. IEEE Press, New York (2001)
14. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. *J. IEEE Trans. Image Process.* **19**(6), 1635–1650 (2010)
15. Martel-Brisson, N., Zaccarin, A.: Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation. In: *2010 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8. IEEE Press, New York (2010)

Image and Video Technology – PSIVT 2015 Workshops
RV 2015, GPID 2013, VG 2015, E04AS 2015, MCBMIA
2015, and VSWS 2015, Auckland, New Zealand,
November 23-27, 2015. Revised Selected Papers
Huang, F.; Sugimoto, A. (Eds.)
2016, XXV, 370 p. 203 illus. in color., Softcover
ISBN: 978-3-319-30284-3