Fake Shadow Detection Based on SIFT Features Matching

Liang Zhang Xiaomei He
Tianjin Key Lab for Advanced Signal Processing, Civil Aviation University of China
Tianjin 300300, P.R. China
e-mail: l-zhang@cauc.edu.cn cemn84@126.com

Abstract—**Shadow detection is crucial for robust and reliable visual surveillance system. The shadow detection method based on gray level and color information will fail when object parts have very similar properties with real shadow. In this paper, we present a novel shadow detection scheme which can conveniently verify detected shadows utilizing the extracted SIFT (Scale Invariant Feature Transform) features. After candidate shadow regions are detected by utilizing rgb color model, SIFT algorithm is exploited for local feature detection in two consecutive frames. Then location information of SIFT features in moving object and shadow is analyzed. Finally fake shadow regions are successfully identified. Experiments indicate that the proposed algorithm is fast and exact, and improves the accuracy of the moving object detection. The proposed method can be used for shadow detection in SIFT features tracking and will be highly effective.**

Keywords-moving object detection; fake shadow detection; SIFT; feature matching; location information

I. INTRODUCTION

In Intelligent Visual Surveillance System, moving objects should be first extracted from video sequences. Robust tracking of objects in the scene calls for a reliable and effective moving object segmentation. The methods based on frame difference or background subtraction can't avoid detecting shadows as foreground since shadow points differ significantly from the background points and shadows have the same motion as the objects casting them. Therefore, shadow detection and elimination are the essential steps.

Many methods have been put forward in literature to detect moving shadow. A common way is to use gray level and color information [1-5]. Compared with corresponding background, shadow regions have lower gray level, and shifted color components. Therefore, the two features are mostly utilized to detect shadow regions. Nevertheless, the foreground objects may have very similar features with the real shadow, i.e., the color of the foreground objects and the color of the corresponding background region are similar; in addition, the objects are darker than the corresponding background. So it becomes not reliable to detect moving shadows by using only gray level and color information of the isolated points.

Fake shadow indicates object parts that have very similar features with real cast shadow [6]. If most of the object is detected as shadow, it will cause a large missing of the object, which would lead to a failure of object tracking.

To solve this problem, a verification step can be performed to detect fake shadow after using gray level and color information for shadow detection. For example, Wang et al. [6] made use of gray level and color features to divide the cast shadow region into subregions in traffic images. And then edge refinement algorithm is proposed to find fake shadow regions in previous detection result. Chen et al. [7] employed the shadow features on brightness and chromaticity successively to detect candidate shadow regions in rgb color space. Then region-based geometry information is exploited to exclude fake shadow regions. Later, Zhang et al. [8] introduced a method for shadow detection based on intensity and ratio edge. And then geometric heuristics is used to further improve the detection results. Additionally, Dong et al. [9] proposed a method for the detection of moving object and shadow based on RGB color space and edge ratio. Firstly, extract the likely object and shadow regions according to the chromaticity distortion and brightness distortion of pixels. Then indentify the fake shadow regions according to the area and the edge ratio of each region.

Above verification methods strongly depend on the color and geometric properties of shadow and most of them are used for vehicle fake shadow detection, few referring to human fake shadow detection. It is worth noting that SIFT features is invariant to image scaling and rotation, and partially invariant to change in illumination and 3D camera viewpoint [10]. At present, SIFT method has been widely used for object tracking and tracking is implemented based on SIFT features matching [11-12]. In this paper, we focus on the process in solving fake shadow problem of human objects with SIFT features matching method, combining with traditional gray level and color method.

The rest of this paper is organized as follows. Section II describes rgb shadow model and the detailed procedures of SIFT matching scheme. Experimental results are discussed in Section III. Finally, conclusions and future work are given in Section IV.

II. THE PROPOSED APPROACH

Shadow detection is performed in the course of moving object detection. Firstly, obtain foreground regions in previous frame and current frame through background subtraction. Then detect candidate shadow regions in current frame by exploiting rgb color model proposed in [7]. Instead of using region-based geometry information, SIFT features in foreground regions in previous frame and SIFT features in

978-0-7695-4080-1/10 $26.00 © 2010 IEEE
DOI 10.1109/ICIE.2010.58

the candidate shadow regions in current frame are used for matching by SIFT algorithm. And then location information of matching SIFT features can be used to determine which features belong to fake shadow regions.

A. *Shadow Detection Based on rgb Color Model*

The normalized rgb color model is defined as follows:

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B}. \quad (1)$$

And the luminance of each pixel is defined as follows:

$$Y = (R+G+B)/3. \quad (2)$$

The process of extracting candidate shadow regions is finished by two steps. The first step is Luminance Ratio Test. Let $p(x)$ be the pixel at position x in detected foreground regions, where $I(x)$ and $B(x)$ are the corresponding pixel values in the input image and in the background model, respectively. The Luminance criterion of shadows is as follows:

$$\alpha \le \frac{B(x)}{I(x)} \le \beta \quad (3)$$

where the upper bound β is used to define a maximum value for the darkening effect of shadows on the background. Instead, the lower bound α can avoid identification of the points where the background is slightly darkened by noise as shadows. According to the property of shadows which normally darken the surface upon which they are cast, the pixels satisfying (3) are determined as shadow points.

The second step is color consistency analysis. It is performed in the shadow regions obtained by the first step. Let $p(x)$ be a pixel where $x \in (r,g,b)$ is background pixel and $x \in (r',g',b')$ is foreground pixel, then

$$t_1 = 1 - \frac{|r-r'|}{r+r'}, t_2 = 1 - \frac{|g-g'|}{g+g'}, t_3 = 1 - \frac{|b-b'|}{b+b'}. \quad (4)$$

And the chrominance criterion of shadows is as follows:

$$H(r,g,b) = t_1 * t_2 * t_3, H(r,g,b) \ge H_T \quad (5)$$

where H_T is the threshold to measure color similarity. According to the property of shadow regions which will hardly change the color of corresponding background regions, the pixels satisfying (5) are determined as shadow points.

This process excludes most of the shadow points as a result of the process of luminance ratio test. However, when object parts have very similar properties with real shadow, many points are still preserved because they are highly identical with the true shadow points. Considering there might be some holes inside the shadow regions, morphologic post-processing must be performed so that the shadow regions can be used for later processes of SIFT matching.

Through above steps, the candidate shadow regions have been obtained, including fake shadow regions and real shadow region.

B. *SIFT Matching Scheme*

Scale Invariant Feature Transform [10] has been designed for extracting highly distinctive invariant features from images, which are invariant to image scaling and rotation, and partially invariant to change in illumination and 3D camera viewpoint, and also robust to disruptions due to clutter or noise. So they can be used to perform reliable matching of the same object or scene between different images.

SIFT attempts to extract scale-invariant features by using a staged filtering method. In the first stage, key locations in scale space are identified by looking for locations that are extrema of a difference-of-Gaussian function. Then each keypoint with location, scale and orientation, is used to generate a 128-element feature vector that describes the local image region sampled relative to its scale-space coordinate frame.

For further details the reader is referred to [10].

1) *Matching principle*

Once foreground regions in previous frame and shadow regions in current frame are detected, some SIFT features can be extracted for matching and identifying.

In SIFT algorithm, the sift function extracts SIFT features, including their locations and descriptors. Location information is represented in $(\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}, \boldsymbol{\theta})$. Generally speaking, in the two consecutive frames, the change of real shadow region is very small and locations of SIFT features in real shadow region should be invariable. But locations of SIFT features in object regions are variable when the object is moving. Therefore, SIFT features in foreground regions in previous frame and SIFT features in shadow regions in current frame are used to perform matching. And then location information of matching SIFT features can be used to determine which features belong to fake shadow regions.

In theory, locations of SIFT features in real shadow region are invariable. But in fact, there might be a little change in these locations due to noise or other factors. The change is much less than location variation of SIFT features in object regions. Let T_1 be the threshold for location variation of matching SIFT features, in order to find those features belonging to fake shadow regions. If location variation of SIFT features is greater than T_1, it may be assumed that these SIFT features belong to fake shadow regions.

In real shadow region, some SIFT features whose location variation may be greater than T_1, will be wrongly divided into fake shadow regions. Here, the threshold T_2 is used to define a minimum value of the number of matching features whose location variation is greater than T_1 and divided into each shadow subregion. It can make sure that real shadow region won't be mistaken for fake shadow

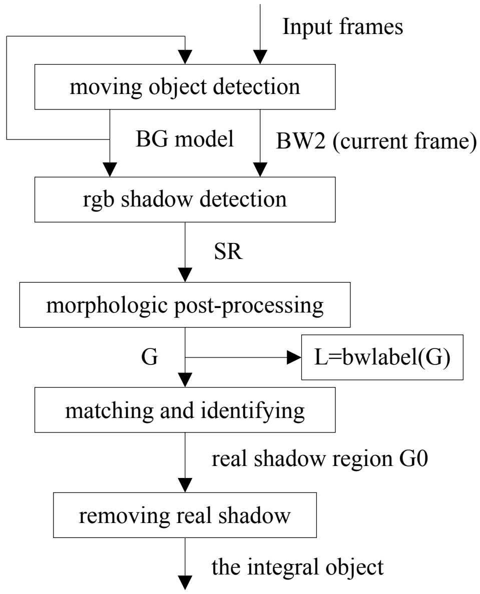


Figure 1. The integral scheme

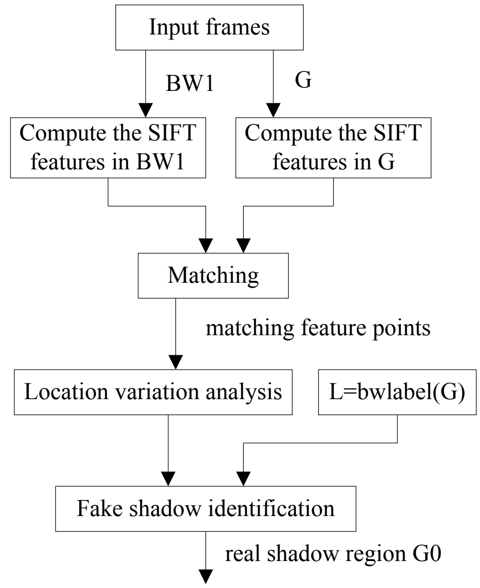


Figure 2. Matching and identifying

regions that will be removed; even if some SIFT features belonging to real shadow region are wrongly determined as SIFT features belonging to fake shadow regions. In our experiment, T_1 is set to 0.2 and T_2 is set to 5.

2) Matching scheme

The integral scheme of our method is depicted in Fig. 1 and the details of matching and identifying are given in Fig. 2. Some notations are as follows:

- BG the background model;
- BW1 the binary mask of foreground regions to be detected in frame n-1;
- BW2 the binary mask of foreground regions to be detected in frame n;
- SR the binary mask of shadow regions to be detected in frame n;
- G the binary mask of SR by morphologic post-processing;
- G0 the binary mask of real shadow region in frame n;
- BW the binary mask of the integral object in frame n.

All steps are given below:

Step 1—Obtain the background model BG and the foreground region BW2 in frame n through background subtraction.

Step 2—Obtain candidate shadow regions SR using (3) and (5) based on rgb color model.

Step 3—Fill some holes and remove isolated pixels inside SR by morphologic post-processing, obtaining regions G, which may consist of several subregions. And then tag these regions with $L = bwlabel(G)$.

Step 4—Matching and identifying

- *Matching:* Extract SIFT features from frame n and frame n-1, and determine which SIFT features belong to G or BW1, respectively. Then these SIFT features are used to perform matching.
- *Location variation analysis:* Let (p_1, p_2) be the matching features, and p_1 is the feature point $(x_1, y_1, \sigma_1, \theta_1)$ in G and p_2 is the feature point $(x_2, y_2, \sigma_2, \theta_2)$ in BW1. Compute location variation, i.e., $\Delta x = |x_1 - x_2|, \Delta y = |y_1 - y_2|$. If any of them is greater than T_1, p_1 will be determined as the feature in fake shadow regions.
- *Fake shadow identification:* Design a counter for each subregion in G and compute the number of SIFT features belonging to each subregion. The subregions where the number is more than T_2 are determined as fake shadow regions. Set the pixels to 0 in fake shadow regions and then real shadow region G0 is obtained.

Step 5—The integral object BW is obtained by BW=BW2-G0.

III. Results and Discussion

Our method is mainly tested with various indoor image sequences which consist of moving people in our experiment. Some results are shown in Fig. 3- Fig. 6.

In Fig. 3 (sequence 1), the person wears a coat with the different color and brightness from the corresponding background. We can see the result of shadow detection is very good only using brightness and chromaticity information. But in Fig. 4 (sequence 2), a large part of object is detected as shadow only using brightness and chromaticity information, as shown in (b) and (c). This is due to the fact that the color of the coat and the color of the corresponding background region are similar. In addition, the coat is slightly darker than the background. This portion of the coat has therefore the same characteristics as a shadow cast on the background. The proposed algorithm based on SIFT features matching is used and SIFT features in candidate shadow region in frame #296 and SIFT features in foreground region in frame #295 are used for matching (the number of matching is 27), then fake shadow regions are identified and real shadow region is obtained, as shown in (g). Good results of shadow detection in other frames are shown in (i) and (j).

When there is more than one subregion in fake shadow regions, our method also achieves good results, as shown in Fig. 5 (sequence 3).

IV. CONCLUSIONS AND FUTURE WORK

This paper proposes a novel shadow detection scheme based on SIFT features matching, combining with traditional gray level and color information. The contributions of our work are mainly focused on dealing with fake shadow problem of human objects. Experimental result verifies that the proposed method is simple, effective and feasible. Even if there is more than one subregion in fake shadow regions, the proposed method also achieves good results. But a limitation of this method is that fake shadow subregion is still removed when real shadow regions and fake shadow subregion are connected, as shown in Fig. 6. Sometimes, shadows are not only cast on the floor but also on walls by moving object, and real shadow regions and fake shadow region may be connected.

A direction for future research will be to find a method to divide connected candidate shadow regions into different subregions. Then SIFT matching method is used to identify fake shadow regions.

Another direction for future research will be that the proposed method is used for shadow detection in SIFT tracking, which not only detects real shadow exactly and effectively but also achieves the real-time performance.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation under Grants 60736009, 60776807, and 60605008.

REFERENCES

[1] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, “Improving shadow suppression in moving object detection with HSV color information,” In Proc. IEEE Int'l Conf. Intelligent Transportation Systems, 2001, pp. 334-339.

[2] O. Schreer, I. Feldmann, U. Goelz, and P. Kauff, “Fast and robust shadow detection in video conference applications,” in: Proc. VIPromCom 2002, 4th EURASIP IEEE International Symposium on Video Processing and Multimedia Communications, 2002, pp. 371–375.

[3] E. Salvador, A. Cavallaro, and T. Ebrahimi, “Cast shadow segmentation using invariant color features,” Computer Vision and Image Understanding, 2004, vol. 95, pp. 238-259.

[4] T. Horprasert, D. Harwood, and L. S. Davis, “A statistical approach for real-time robust background subtraction and shadow detection,” Proc. IEEE Int’l Conf. Computer Vision ’99 FRAME-RATE Workshop, 1999.

[5] F. Bunyak, I. Ersoy, and S. R. Subramanya, “Shadow detection by combined photometric invariants for improved foreground segmentation,” Proceedings of the Seventh IEEE Workshops on Applications of Computer Vision, vol. 1, 2005, pp. 510-515.

[6] Y. Q. Wang, M. Tang, and G. X. Zhu, “An improved cast shadow detection method with edge refinement,” Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications, vol. 2, 2006, pp. 794-799.

[7] B. S. Chen and D. S. Chen, “Shadow detection based on rgb color model,” ICIC, 2006, pp. 1068-1074.

[8] W. Zhang, X. Z. Fang, and X. K. Yang, “Moving cast shadows detection using ratio edge,” IEEE Transactions on Multimedia, vol. 9, no. 6, 2007, pp. 1202-1214.

[9] X. Dong, K. D. Wang, and G. H. Jia, “Moving object and shadow detection based on RGB color space and edge ratio,” 2009. CISP '09. 2nd International Congress on Image and Signal Processing, pp. 1-5.

[10] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, vol. 60, no. 2, 2004, pp. 91-110.

[11] S. Battiato, G. Gallo, and S. Scellato, “SIFT Features tracking for video stabilization,” 14th International Conference on Image Analysis and Processing, 2007.

[12] W. B. Yang, B. Fang, and Y. Y. Tang, “SIFT Features based object tracking with discrete wavelet transform,” Proceedings of the 2009 International Conference on Wavelet Analysis and Pattern Recognition, July 2009, pp. 380-385.

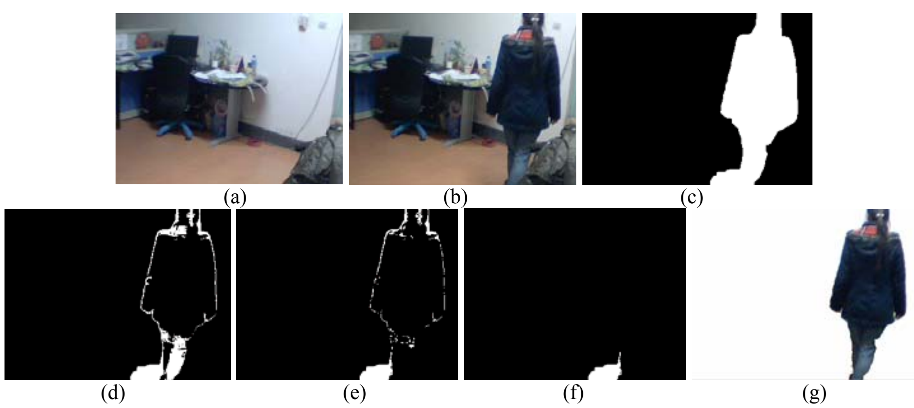
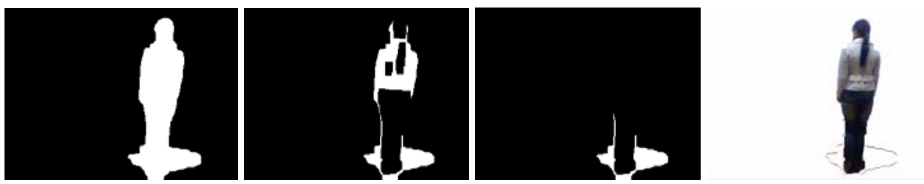


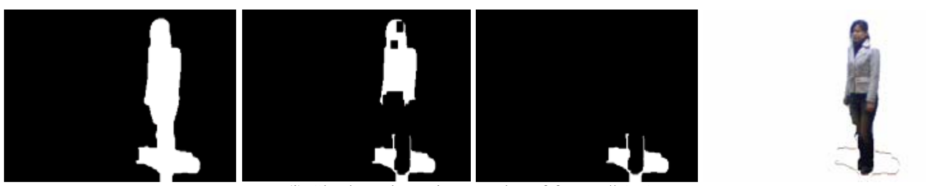
Figure 3. An example shows the result of the algorithm only applying brightness and chromaticity information based on rgb color model. (a) the background image (b) the input image (c) the binary mask of foreground region (d) shadow detection results based on brightness information (e) shadow detection results based on chromaticity information (f) shadow detection results after post-processing (g) the moving object

(a) (b) (c) (d)

(e) (f) (g) (h)



(i) Shadow detection results of frame # 250

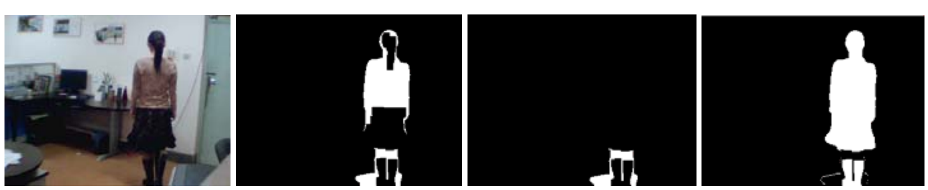


(j) Shadow detection results of frame # 270

Figure 4. Another example shows the result of shadow detection based on SIFT features matching when object parts have very similar features with real cast shadow. (a) the input image (b) shadow detection results based on brightness information (c) shadow detection results based on chromaticity information (d) shadow detection results using morphologic post-processing (e) matching results of SIFT features in candidate shadow region in frame #296 and SIFT features in foreground region in frame #295 (f) fake shadow region (g) real shadow region (h) the integral object (i)(j) shadow detection results of frame #250, #270, respectively

(a) (b) (c)

Figure 5. The third example shows the good results when there is more than one subregion in fake shadow regions. (a) the input image (b) shadow detection results based on brightness and chromaticity information (c) real shadow region after using our method



(a) (b) (c) (d)

Figure 6. The fourth example shows shadow detection results. When real shadow regions and fake shadow subregion are connected, fake shadow subregion is still removed.