HARD SHADOWS REMOVAL USING AN APPROXIMATE ILLUMINATION INVARIANT

Bingshu Wang[1], C.L. Philip Chen[1,], Yuyuan Li[2], Yong Zhao[2]*

[1]Faculty of Science and Technology, University of Macau, Macau, China
[2]Shenzhen Graduate School, Peking University, Shenzhen, China
{yb77408, philiphen}@umac.mo, 1501213870@sz.pku.edu.cn, zhaoyong@pkusz.edu.cn

ABSTRACT

Hard shadows detection and removal from foreground masks is a challenging step in change detection. This paper gives a simple and effective method to address hard shadows. There are inside portion and boundary portion in hard shadows. Pixel-wise neighborhood ratio is calculated to remove the most of inside shadow points. For the boundaries of shadow regions, we take advantage of color constancy to eliminate the edges of hard shadows and obtain relative accurate objects contours. Then, morphology processing is explored to enhance the integrity of objects. The main contribution of this paper is to design an approximate estimation strategy for illumination invariant based on Lambertian reflectance model without prior knowledge. The proposed method is unsupervised and experimental results on six challenging sequences show the effectiveness and robustness of our approach.

Index Terms— Hard shadows, Illumination invariant, Lambertian reflectance, Color constancy

1. INTRODUCTION

In surveillance application scenes, moving cast shadows are generated when illumination source is partially or totally occluded by moving objects in the field of change detection [1]. According to the extent of darkness, shadows can be divided into two categories: weak (or soft) shadows and hard (or strong) shadows [2]. Due to the lack of illumination source and low values of incoming energy, hard shadow regions often have the lowest intensity among image pixels [3]. Moving shadows usually share same motion property with moving objects and have obvious discrimination from corresponding background. This easily results in occurrence of shadows misclassification as objects [4,5] and distorted object contours, multiple objects merging. etc. Most of background subtraction methods have difficulty in handling hard shadows although some can suppress weak shadows. Therefore, algorithms or approaches aimed at removing hard shadows need to be developed independently.

Many shadow detection methods were proposed and surveyed in [5-8]. These algorithms are based on one or several certain assumptions about shadow properties but not limit to: (i) shadow regions are darker than the corresponding background but the color and texture information does not change significantly (also called color constancy and texture consistency); (ii) the shadow is adjacent to the moving objects; (iii) both the direction and strength of the illumination source are known; (iv) the shadow falls on ground plane. For shadows occurring in indoor conditions, it seems not hard to detect them because the assumptions are always easy to be satisfied. Various color-based [5,9,10,11,12], edge-based [13,14], texture-based features [15-19], multiple features fusion [20-23] and other methods or techniques [24-30] were given and they have achieved good results. But these methods face challenge on handling hard shadows occurring in outdoor conditions since it is difficult to meet some certain assumptions for the complex scenes, for example, the color or texture information is less and shadows even dark to black. Recent techniques using convolutional deep neural networks [29] or generative adversarial networks [30] were explored to remove shadows from single image. They achieved good performance with complex network design, which might be time-consuming and not suitable for the fast removal of moving shadows.

Our motivation is to design an approach to remove hard shadows from foreground masks. Two problems should be issued mainly: (a) shadow camouflage happens when shadow-foreground discrimination is little, (b) object and background have high similarity. To address these issues, we present a framework depending on illumination invariant [31] and color constancy [5]. Hard shadows can be divided into two parts: the inside and the boundary. The inside shadows are much darker than corresponding background while the shadow boundaries are changing areas from background to shadow whose textures change significantly but the chromaticity remains unchanged. According to the characteristics above, our strategy is to deal with the inside parts using illumination invariant and remove the boundary parts by color constancy, following by a morphological methodology that only rely on objects' contours. The experiments are performed on six typical outdoor scenes of hard shadows. Six methods, as the representative of the-state-of-the-art, is compared with the proposed method. The comparison results demonstrate our method's effectiveness.

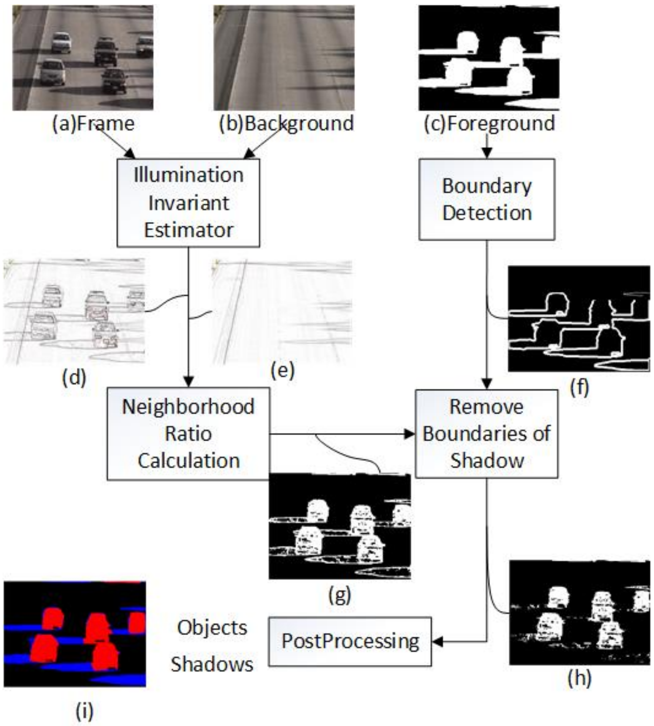

Fig.1. Flowchart of the proposed method.

2. SYSTEM OVERVIEW

The proposed method requires three inputs for each frame in sequences (see Fig.1. (a), (b) and (c)): 1) Frame as the current image in sequences; 2) Foreground as the mask of detected changes; 3) Background that is generated from background subtraction or other methods. In this paper, supposing that three inputs are given in advance.

Firstly, the approximate illumination invariants of input frame and background are estimated by the proposed illumination invariant estimator (see Fig.1. (d) and (e)). The neighborhood ratio between the illumination invariants of frame and background is calculated to discriminate object from shadow (see Fig.1. (g)). It can be seen from Fig.1. (g) that the edges of hard shadows are misclassified as objects, which will be processed individually.

Then the input foreground is used to detect the outlines of connected regions (see Fig.1. (f)). By color constancy method based on HSV color space [5] we discriminate shadows' boundaries from objects' boundaries (see Fig.1. (h)). Finally, post-processing with morphological operation is operated to fit the objects' contours and produces the desired mask (the red) and shadows (the blue). In the process, our method is unsupervised, which is more capable of adapting to complex scenes.

3. THE PROPOSED METHOD

In this section, we give hard shadows removal algorithm. Hard shadows are those points that have low intensity, much darker than corresponding background and lack of color and texture information. Mostly in outdoor scenes, the strong illumination source creates a high brightness environment for scenes. But the brightness of umbra belonging to hard shadows is mostly affected by ambient light other than illumination source. Generally, the ambient light is much weaker than illumination source. Due to the weakness, shadows tend to be darker, while the reflectance of shadow regions changes little. In this paper, it is assumed that the reflectance is scattered reflection according to Lambertian reflectance model [32]. Therefore, the surface reflectance is a good property as illumination invariant feature.

3.1. Illumination invariant estimator

According to Lambertian reflectance, any pixel of an image obtained from fixed and static scene can be described by a simple luminance model [4,11]:

$$\mathbf{L}(x) = \mathbf{I}(x)\,\mathbf{R}(x) \tag{1}$$

where $\mathbf{L}(x) = [L^B(x), L^G(x), L^R(x)]^{\mathrm{T}}$ is the illumination vector of RGB color space at pixel x. $\mathbf{I}(x) = [I^B(x), I^G(x), I^R(x)]^{\mathrm{T}}$ is the irradiance vector of the illumination source and $\mathbf{R}(x) = [R^B(x), R^G(x), R^R(x)]^{\mathrm{T}}$ is the reflectance vector of surface at the pixel x. Assuming the irradiance component given, the intrinsic reflectance property of objects and background can be expressed by:

$$\mathbf{R}(x) = \mathbf{L}(x) / \mathbf{I}(x) \tag{2}$$

$\mathbf{L}(x)$ can be acquired by standard RGB camera while $\mathbf{I}(x)$ cannot be gained directly and should be estimated approximately. The shadow irradiance of input signal for one illumination source can be described by [33]:

$$\mathbf{I}(x) = \mathbf{C}_a + \mathbf{C}_b \cos(\theta(x))\varsigma(x) \tag{3}$$

where $\mathbf{C}_a$, $\mathbf{C}_b$ and $\theta(x)$ represent the intensity of ambient light, the intensity of illumination source and the angle between the direction of illumination source and the normal vector of surface, respectively. The term $0 \le \varsigma(x) \le 1$ describes the transition inside the penumbra or umbra, when $\varsigma(x)$ equals to zero, the shadow is only affected by ambient light and belongs to umbra.

For hard shadows, the points of inside region mostly belong to umbra while the points of boundary region tend to belong to penumbra. The division line or points between penumbra and umbra is not fixed and rigid under various and complex scenes, which is difficult to define the accurate positions. This is because that neighboring pixels in shadow regions tend to have spatial consistency and influence each other a lot. Therefore, we process penumbra and umbra together with the same strategy. For pixel x in shadow region, the intensity of illumination source should be greater than or equal to $\mathbf{L}(x)$. For this reason, we carry out a selection policy from its neighboring pixels to calculate an approximation value of the irradiance $\mathbf{I}(x)$.

$$\mathbf{I}_{max}(x) = \max\,(L(t)),\ \mathbf{I}_{min}(x) = \min\,(L(t)),\ t \in \mathbf{W} \tag{4}$$

$$\alpha = (\mathbf{I}_{max}(x) - \mathbf{I}_{min}(x)) / \mathbf{I}_{max}(x) \quad (5)$$

$$\mathbf{I}(x) \approx \mathbf{I}_{max}(x) \cdot (1-\alpha) + \mathbf{I}_{min}(x) \cdot \alpha + \epsilon \quad (6)$$

In Equation (4), $\mathbf{W}$ represents the neighboring window of pixel x in RGB color space. The $\mathbf{I}_{max}(x)$ and $\mathbf{I}_{min}(x)$ represent the maximum and minimum values in $\mathbf{W}$, respectively. In Equation (5), α is defined as the fusion factor of illumination correction that is used to overcome the underestimate of single illumination source for quick changing areas. In Equation (6), ϵ is the adjust factor to ensure that the estimation value $\mathbf{I}(x)$ is greater than or equals to $\mathbf{L}(x)$. An example of visual results of illumination invariant is shown in Fig.1. (d) and (e).

3.2. Neighborhood ratio calculation

This step is based on the results of illumination invariant estimation from section3.1. Compared to the calculation of single pixel, a small neighborhood region (eg, 3×3 patch) is more robust to light changes. The formulation is defined as:

$$\Omega(x) = \begin{cases} 1, & if\ |I_c(x) - I_b(x)| < \lambda \\ 0, & otherwise \end{cases} \quad (7)$$

where $I_c(x)$ represents the illumination invariant of current frame at pixel x position, $I_b(x)$ is the illumination invariant of corresponding background point and λ is the range threshold parameter. The shadow point is determined by:

$$shadow = \begin{cases} 1, & if\ (\sum_{j=1}^{n} \sum_{k\in\{b,g,r\}} \Omega_k(j)) / 3n > \tau \\ 0, & otherwise \end{cases} \quad (8)$$

where τ is the determination parameter of neighborhood ratio. This step makes a rough determination to discriminate shadow from object (see Fig.1. (f)).

3.3. Boundary detection

It can be seen from Fig.1. (g) that the boundaries of hard shadows are easily misclassified as objects by neighborhood ratio calculation. Due to the penumbra points are distributed in the light changing areas from background to shadow, they are lighter than the inside parts of hard shadows. To correct the misclassification, we take advantage of color constancy techniques to detect boundaries according to the size of connected region. The larger the size of connected region is, the wider the boundaries are. This is because that in terms of small objects, its boundaries account for a high percentage of the whole connected region. In practice, the upper limit of boundary width with three pixels may provide satisfactory results. If the contour size of a connected region (CSCR) is more than one-third of the sum of frame's height and width, the boundary width is set to three pixels. If CSCR is less than one-fourth of the sum of frame's height and width, the boundary width is set to one pixel. In other situations, the boundary width is set to two pixels.

3.4. Removal boundaries of shadows

For every connected region, the boundaries consist of object points and shadow points. The shadow points mostly belong to penumbra and reserve the chromaticity information, which is to coincide with the assumption of color constancy, while the object points are not. Therefore, we choose to utilize HSV color space to detect the edges of shadows. It is because that HSV color space with fast processing speed [5] and has revealed the accuracy in distinguish shadows from objects. Then most boundaries of shadows can be removed from object regions (Fig.1. (g)).

In addition, the post-processing process is also of importance to maintain the structure and contours of real objects. In this paper, it is performed by polygonal fitting operation and convex hull operation.

4. EXPERIMENTAL RESULTS

To validate the proposed method's effectiveness, six typical and challenging sequences from [1,7,26] are used to test shadow detection methods. They are all outdoor scenes. It should be noted that *Seam, Senoon* and *Sepm* are three sequences from the same camera with different time. These are typical real application with hard shadows. For fairness and respect to the original authors, the parameters of compared methods are kept in default given by their papers. In our method, only λ and τ need to be considered globally and set since multiple scenarios are included. Selecting a higher value than λ=23 or lower value than τ=0.6 may provide a better shadow detection, but it will lead to misclassification that object points are classified as shadow class. Visual results and quantitative comparisons are both presented in Fig.2., Fig.3., Table 1 and Table 2.

Figure 2 shows visual results of the proposed method. It can be seen from the figure that hard shadows with diverse direction and strength changes can be detected accurately. The contours and insides of objects are retained in a relative complete way. The quantitative results are reported in Table 1 and Table 2 and evaluated by shadow detection (η) and shadow discrimination (ξ) [5,23] as follows:

$$\eta = \frac{TP_S}{TP_S + FN_S}, \xi = \frac{TP_F}{TP_F + FN_F}, F-measure = \frac{2\eta\xi}{\eta+\xi} \quad (9)$$

where TP_S and FN_S are the true positives and false negatives for shadows, TP_F and FN_F are the true positives and false negatives for objects. The proposed approach achieves much higher shadow detection rate than other methods in the five sequences except *HighwayI* scene in Table 1 and the shadow discrimination rate of our method is also high (Table 2). For *Bungalows* scene, the shadow detection rates of compared methods are lower than 70% while ours is more than 90%. The compared methods show the instability in handling hard shadows, for instance, LR method [7] could detect shadows in three sequences with only over 40%. Multiple features

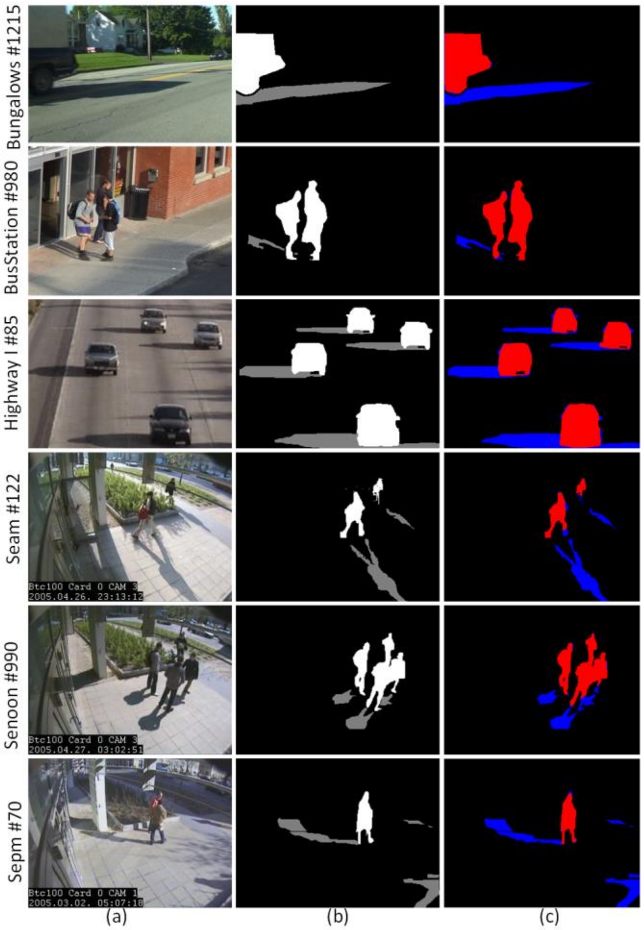

Fig.2. The visual results of our method. (a) input frame. (b) ground truth manually. (c) the blue represents objects and the red represents hard shadows, individually.

Table 1. Shadow detection rates on six sequences

η %	Chr [5]	Geo [24]	Phy [10]	SR [15]	LR [7]	MFF [23]	Ours
Bungalows	1.1	59.8	13.3	2.2	62.6	3.2	90.9
BusStation	56.1	25.1	40.1	58.4	49.2	75.0	93.8
HighwayI	90.5	71.0	45.1	25.1	74.7	69.2	88.2
Seam	19.2	57.4	56.0	69.1	22.7	73.4	99.6
Senoon	18.4	53.4	64.9	3.2	4.8	48.8	95.6
Sepm	46.7	60.3	35.5	35.5	18.8	58.4	95.8

Table 2. Shadow discrimination rates on six sequences

ξ %	Chr [5]	Geo [24]	Phy [10]	SR [15]	LR [7]	MFF [23]	Ours
Bungalows	78.9	55.4	92.5	81.8	69.1	95.6	92.1
BusStation	83.7	70.8	94.7	85.7	94.0	91.2	88.9
HighwayI	54.2	74.7	83.9	88.0	82.9	85.6	90.1
Seam	59.9	64.6	87.1	67.7	79.1	77.7	72.8
Senoon	59.2	58.4	81.7	86.3	98.2	60.4	85.3
Sepm	66.9	63.1	86.6	79.7	98.1	85.6	78.9

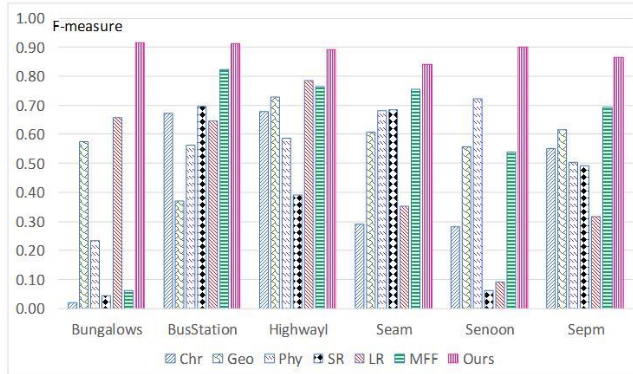

Fig.3. The comprehensive quantitative results of seven methods on six sequences.

fusion [23] obtains relative high shadow detection rate in five scenes but not in *Bungalows* scene. What needs to be pointed out is that the high rates of the (η%) and (ξ%) simultaneously verify the effectiveness of shadow detection methods convincingly. It can be seen clearly from Fig. 3. that our method outperforms the-state-of-the-art methods by comprehensive metric (F-measure) in all the sequences.

In addition, our algorithm has been implemented in C++ and processes 50~60 frames a second for a 360×240 resolution frame. The algorithm has been tested on a 3.5GHz Xeon machine with memory size of 8G. After optimization, the time consumption can be reduced further.

5. CONCLUSION

In this paper, a simple, effective approach of removing hard shadows is proposed. It is unsupervised and without prior knowledge. Visual and quantitative results validate the proposed method's effectiveness and robustness. The most important innovation of this paper is the introduction of illumination invariant and estimation by spatial information. Based on estimating illumination with pixel neighboring, the inside part of shadows is detected by the illumination invariant that is expressed in an approximate way. The boundary part of hard shadows is removed by the method of HSV color space-based. The proposed method is applied to multiple outdoor sequences and has promising results in accuracy and processing speed.

6. ACKNOWLEDGEMENT

This research is supported by FDCT under grant: 019/2015/A1, 079/2017/A2, 024/2015/AMJ, partly MYRG: 2016-00159-FST, 2016-00169-FST and 2017-00137-FST, and partly NSFC: 61572540. This work is also supported partly by National Sci-Tech Support Plan: 2015BAK01B04 and Shenzhen Technology Program: JCYJ20160506172651253.

7. REFERENCES

[1] Y. Wang, P. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, P. Ishwar, "CDnet 2014: An expanded change detection benchmark dataset," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 387-394, 2014.

[2] Q. Wu, W. Zhang, B. V. Kumar, "Strong shadow removal via patch-based shadow edge detection," *IEEE International Conference on In Robotics and Automation*, pp. 2177-2182, 2012.

[3] E. A. Mosabbeb, M. Sadeghi, M. Fathy, "A new approach for vehicle detection in congested traffic scenes based on strong shadow segmentation," *In International Symposium on Visual Computing.* Springer, Berlin, Heidelberg, pp. 427-436, 2007.

[4] J. Stander, R. Mech, J. Ostermann, "Detection of moving cast shadows for object segmentation," *IEEE Transactions on Multimedia*, Vol. 1, No. 1, pp.65-76, 1999.

[5] A. Prati, I. Mikic, M. M. Trivedi, R. Cucchiara, "Detecting moving shadows: Algorithms and evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 7, pp. 918-923, 2003.

[6] N. Al-Najdawi, H. E. Bez, J. Singhai, E. A. Edirisinghe, "A survey of cast shadow detection algorithms," *Pattern Recognition Letters*, Vol. 33, No. 6, pp. 752-764, 2012.

[7] A. Sanin, C. Sanderson, B. C. Lovell, "Shadow detection: A survey and comparative evaluation of recent methods," *Pattern Recognition*, Vol. 45, No. 4, pp. 1684-1695, 2012.

[8] M. Russell, J. J. Zou, G. Fang, "An evaluation of moving shadow detection techniques," *Computational Visual Media*, Vol. 2, No. 3, pp. 195-217, 2016.

[9] Z. Liu, K. Huang, T. Tan, L. Wang, "Cast shadow removal combining local and global features," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2007.

[10] J. Huang, C. Chen, "A physical approach to moving cast shadow detection," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 769-772, 2009.

[11] A. Amato, M. G. Mozerov, A. D. Bagdanov, J. Gonzalez, "Accurate moving cast shadow suppression based on local color constancy detection," *IEEE Transactions on Image Processing*, Vol. 20, No. 10, pp. 2954-2966, 2011.

[12] I. Huerta, M. B. Holte, T. B. Moeslund, J. Gonzlez, "Chromatic shadow detection and tracking for moving foreground segmentaion segmentation*," Image and Vision Computing*, Vol. 41, pp. 42-53, 2015.

[13] W. Zhang, X. Z. Fang, X. K. Yang, Q. J. Wu, "Moving cast shadows detection using ratio edge," *IEEE Transactions on Multimedia*, Vol. 9, No. 6, pp. 1202-1214, 2007.

[14] J. Wang, Y. Wang, M. Jiang, X. Yan, M. Song, "Moving cast shadow detection using online sub-scene shadow modeling and object inner-edges analysis," *Journal of Visual Communication and Image Representation*, Vol. 25, No. 5, pp. 978-993, 2014.

[15] A. Leone, C. Distante, "Shadow detection for moving objects based on texture analysis," *Pattern Recognition*, Vol. 40, No. 4, pp. 1222-1233, 2007.

[16] M. Shoaib, R. Dragon, J. Ostermann, "Shadow detection for moving humans using gradient-based background subtraction," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 773-776, 2009.

[17] K. Jiang, A. Li, Z. Cui, T. Wang, Y. Su, "Adaptive shadow detection using global texture and sampling deduction," *IET Computer Vision*, Vol. 7, No. 2, pp.115-122, 2013.

[18] B. N. Subudhi, S. Ghosh, S. Cho, A. Ghosh, "Integration of fuzzy markov random field and local information for separation of moving objects and shadows," *Information Sciences*, Vol. 331, pp. 15-31, 2016.

[19] Z. Zhu, X. You, C. P. Chen, D. Tao, W. Ou, X. Jiang, J. Zou, "An adaptive hybrid pattern for noise-robust texture analysis," *Pattern Recognition*, Vol. 48, No.8, pp. 2592-2608, 2015.

[20] J. Dai, M. Qi, J. Wang, J. Dai, J. Kong, "Robust and accurate moving shadow detection based on multiple features fusion," *Optics & Laser Technology*, Vol. 54, pp. 232-241, 2013.

[21] V. Gomes, P. Barcellos, J. Scharcanski, "Stochastic shadow detection using a hypergraph partitioning approach," *Pattern Recognition*, No. 63, pp. 30-44, 2017.

[22] M. Russell, J. Zou, G. Fang, W. Cai, "Feature-based image patch classification for moving shadow detection," *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 14, No. 8, 2017.

[23] B. Wang, Y. Yuan, Y. Zhao, W. Zou, "Adaptive moving shadows detection using local neighboring information," *In Asian Conference on Computer Vision*, pp. 521-535, 2016.

[24] J. Hsieh, W. Hu, C. Chang, Y Chen, "Shadow elimination for effective moving object detection by gaussian shadow modeling," *Image and Vision Computing*, Vol. 21, No. 6, pp. 505-516, 2003.

[25] Y. Zhou, J. Peng, C. P. Chen, "Extreme learning machine with composite kernels for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 8, No. 6, pp. 2351-2360, 2015.

[26] C. Benedek, T. Szirnyi, "Bayesian foreground and shadow detection in uncertain frame rate surveillance videos*," IEEE Transactions on Image Processing*, Vol. 17, No. 4, 608-621, 2008.

[27] J. T. Lee, K. Lim, Y. Chung, "Moving shadow detection from background image and deep learning," *Pacific-Rim Symposium on Image and Video Technology*, pp. 299-306, 2015.

[28] C. P. Chen, Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29 no.1, pp. 10-24, 2018.

[29] S. H. Khan, M. Bennamoun, F. Sohel, Togneri, R, "Automatic shadow detection and removal from a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 38, No. 3, pp. 431-446, 2016.

[30] V. Nguyen, T. F. Y. Vicente, M. Zhao, M. Hoai, D. Samaras, "Shadow detection with conditional generative adversarial networks," *IEEE International Conference on Computer Vision,* pp. 4520-4528, 2017.

[31] Y. Cheng, Y. Hou, C. Zhao, Z. Li, Y. Hu, C. Wang, "Robust face recognition based on illumination invariant in nonsubsampled contourlet transform domain," *Neurocomputing*, Vol. 73, No. 10, pp. 2217-2224, 2010.

[32] M. Oren, S. K. Nayar, "Generalization of lambert's reflectance model," *21st Annual Conference on Computer Graphics and Interactive Techniques*, pp. 239-246, 1994.

[33] B. T. Phong, "Illumination for computer generated pictures," *Communications of the ACM*, Vol, 18, No. 6, pp. 311-317, 1975.