Journal of Engineering and Applied Sciences 12 (24): 7689-7695, 2017
ISSN: 1816-949X

Adaptive Background Modelling and Shadow Removal for Video Surveillance System

Tawfiq Abdulkhaleq Al Asadi and Fanar Ali Joda
Department of Software, Faculty of Information Technology, University of Babylon, Hillah, Iraq

Abstract: Object detection and moving cast shadow removal are crucial steps in video surveillance system. In this study, we have proposed background modelling algorithm for objects detection has ability to update threshold when illumination is changed which leads to increase the performance of surveillance system instead of other methods that suffer from a general threshold for all cases which fail when there is a change in illumination. In addition, moving cast shadow removal algorithm is proposed it consists of two stages, the first one, candidate shadow regions are obtained by using weak detector. The second one, correlation is calculated between reference and current frames of candidate shadow regions, this stage aims to classify pixels of candidate shadow regions into shadow and foreground object.

Key words: Background modelling, object detection, morphological process, HSV color space, shadow removal, calculated

INTRODUCTION

Tracking of moving object in video has concerned a considerable amount of attention in computer vision, object detection and tracking has main objective is to construct a relationship between object parts or objects in sequential frames and to obtain temporal information related with object such as trajectory, pose, direction and speed. Handling frame by frame in video to track detected objects is an important and difficult task (Hadi and Sabah, 2015).

Detection of foreground object is based on background subtraction by comparing between two frames and will obtain the difference and construct a matrix of distance. Mainly, the difference value will be compared with the threshold value. Therefore, the main situation is that if the difference value satisfies threshold value then is labelled as foreground object if not will label as background.

Background subtraction has several challenges during implementation that appear periodically such as background geometry changes, motion changes and illumination changes. Therefore a simple difference of successive frames is very weak solution for moving object detection accurately. There are several methods to achieve background subtraction such as direct difference of frames, Mixture of Gaussian Model (MOG), Kernel Density Estimation (KDE) and codebook. Different accuracy is obtained by applying different techniques (Desai and Gandhi, 2014). Extracting moving object is an essential step in the video tracking techniques, common problem is that the shadow will be part of the extracted object. At this stage, discrimination the object from the shadow represents a necessary task because the shadow leads to problems in several stages in video tracking such as texture loss of background, object shape analysis and segmentation. Therefore, the process of shadow detection and remove it is a very important task to increase performance of video surveillance (Kar and Deb, 2015).

Shadow possible is larger than real object in case of misclassification of foreground this leads to reduce performance of video tracking. There are several situations related to the performance of detection and tracking the object could be affected by several objects is fused together due to their moving cast shadows and reliability of object appearance model will be reduced due to the existence of pixels shadow which lead to the probability of tracking loss is increased. Figure 1 shows two situations (Sanin *et al*., 2012).

There are many features and ways to detect shadow when the frame contains object and its shadow, most important these features are intensity. It is the simplest feature that can be used to extract cast shadow, it assumes that the areas when become under shadow will be darker due to prevent it from source of lighting. Shadow detection method based intensity set a certain range of illumination to predict regions that fall under the shadow. This feature is used in first step to avoid

Corresponding Author: Fanar Ali Joda, Department of Software, Faculty of Information Technology, University of Babylon, Hillah, Iraq

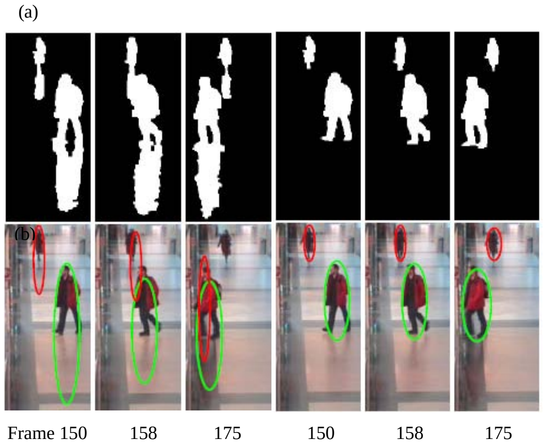

Fig. 1: a) Without shadow removal and b) With shadow removal

non-shadow regions. Note there are no ways mainly dependent on the intensity to discriminate between object and shadow pixels (Huang and Chen, 2009). Many techniques to determine the shadow depend on spectral features associated with color space information, for example when green pixel falls under the shadow effect becomes dark green but preserves same chromacity, the methods that rely on spectral features to detect shadow almost select color space has ability to separate intensity from chromacity such as (HSV, C1C2C3, normalized RGB, YUV) most of these techniques easy to apply and effective in terms of time complexity but it is susceptible to noise due to work on pixel level (Wang *et al*., 2016).

The other direction used to remove the shadow is geometry based technique, this technique use object geometry, camera position and ground surface to detect moving cast shadow (Lakhotiya *et al*., 2015).

Finally, texture based shadow detection technique the idea behind this technique is the texture of foreground object differs from background while texture of shadow preserves same background. The techniques which depend on texture can be classified large region texture and small texture (Hussain and Jebur, 2013).

Literature review: Javed and Shah (2002) proposed technique uses several concepts in a hierarchical style. First, each foreground region in current frame that is darker than reference frame is extracted. Every possible shadow region can include sub regions of cast shadows, self-shadows and parts of foreground object have intensity value less then reference frame. The proposed method achieves color segmentation on every possible shadow region.

The objective is every possible shadow region is divided into sub-regions where each sub-region corresponds only to one of the three (self-shadow, cast shadow and dark object) regions. When the background and object are of different colours then self-shadow and cast shadow regions will have different colours as it is usually the case. A fast color segmentation algorithm is required because the real time requirements are needed by video surveillance systems. Furthermore the segmented regions consist of subset of all foreground regions. As a result the total regions to be segmented are a small ratio of the total frame area.

Tian *et al*. (2005) proposed algorithm based on two types of features the first one is intensity and the second one is texture which are combined to eliminate shadows and to allow to algorithm employing for sudden illumination change. Several assumptions are used based on color information for shadow removal. The intensity information is used rather than color information to preserve the algorithm work with grayscale images. For each pixel of the foreground region the Normalized Cross-Correlation (NCC) of intensities is computed between the reference frame and the current frame.

Huerta *et al*. (2009) present a new technique depended on color and gradient models for separating moving cast shadows which has chromaticity features from detected moving objects. In a first step, both an invariant gradient model and a chromatic invariant colour cone model are constructed to achieve automatic segmentation though detection of possible shadows. In a second step, areas related to possible shadows are gathered by considering “a bluish effect” and an edge partitioning. Finally, spatial similarities between intensity distortions and chrominance angle are analysed for all possible shadow areas in order to finally detecting umbra shadows and temporal similarities between textures. Dissimilar other methods, the proposed system does not make any previous assumptions about camera location, background, surface textures, objects, surface geometries, shapes and types of shadows and objects.

Qin *et al*. (2010) proposed method based on a local texture descriptor called Scale Invariant Local Ternary Pattern (SILTP). A hypothesis turns out that texture features of cast shadow has identical patterns to those of the background but lower than them. The probability of cast shadows is resultant based on information in both texture and color. Adaptively, an online learning pattern is used to update the shadow model. In the final stage, posterior probability of cast shadow areas are constructed by additional combining prior contextual restrictions based on a Markov Random Field (MRF) model.

Kar and Deb (2015) proposed a simple method to detect cast shadow of moving object and then eliminate the shadow area from frames of video. In the proposed system the moving object is extracted by subtraction algorithm depend on thresholding the difference of reference frame and current frame. Intensity in HSV color space, variance and texture features are used to identify the shadow area and shadow removal is based on the information from the reference frame. Background information is primarily kept in reference frame. The next received frames with object are compared with this frame. Color information for both background subtraction and shadow detection to improve object segmentation are guaranteed in proposed system.

Singh and Patnaik (2015) proposed Mixture of Gaussian (MOG) Model with one learning rate only which be used for background modelling and subtraction. Initial classification of foreground pixels into shadow pixels and object pixels are performed using saturation features of HSV color space. In Hue channel difference or intensity ratio based detection of shadow stage, mixture of two Gaussian density functions are used to model the density function calculated on the values of hue difference or brightness ratio. Expectation Maximization (EM) algorithm is used to estimate the Gaussian parameters. Threshold calculations are depended on estimated parameters used to obtain group of shadow pixels. Local region features based on detection of shadow step uses local intensity ratio feature to obtain the group of shadow pixels.

Hu *et al*. (2016) proposed technique based on the hypothesis that shadow region is darker than the corresponding reference but have similar texture and chromacity. The method use both HSV and RGB color spaces to obtain spectral information and merge two texture features to extract moving cast shadows. Firstly, candidate shadow regions are detected by spectral analysis. Secondly, the technique based on proposed extended local ternary pattern, enhanced edge information and HSV color space analysis to achieve three shadow detectors which are used to vote to obtain the final shadows, respectively.

Carmen and Cajote (2016) proposed shadow detection algorithm combines the usage of texture information using interior edges of object, HSV color space information and low-level post processing. Let B_i (x, y) and F_i (x, y) are the current background image and the current foreground image, respectively. The background and foreground frames are converted to grayscale and their individual HSV color space channels equivalent. The candidate shadow pixels are then extracted using an improved form of HSV color space thresholding which is based on the created background image. An additional HSV threshold independent of background image, the foreground hue to value ratio is added to improve the shadow detection rate in cases where the background may be changeable.

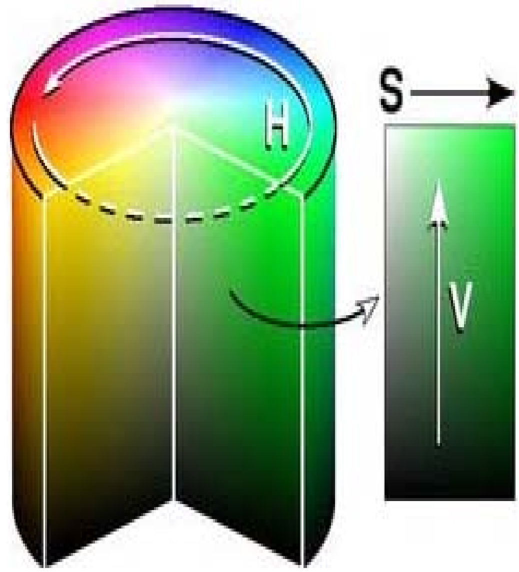

Fig. 2: Block diagram of proposed system

MATERIALS AND METHODS

The proposed system: The general aim of proposed system is to extract moving object from frame sequence which contains object plus shadow. The proposed system focus on object detection using background modelling that based on adaptive threshold then is to detect moving cast shadows and remove it to increase the performance of tracking. Figure 2 shows block diagram of proposed system. As shown in Fig. 2, the proposed system consists of several steps as following.

Convert RGB to HSV color space: Due to the HSV color space has batter ability in term of separation of intensity and chromacity compared with RGB color space. V channel to measure intensity directly and (HS) to measure the chromacity, the HSV channels are normalized to become between 0-255. Figure 3 shows HSV color space model, HSV color space has the following relationships with RGB color space:

$$V = \frac{1}{3}(B+G+R) \tag{1}$$

$$S = 1-\frac{3}{B+G+R}\min(B, G, R) \tag{2}$$

$$H = \begin{cases} \theta \text{ if } G \geq B \\ 360^0 - \theta \text{ if } G < B \end{cases} \tag{3}$$

In which:

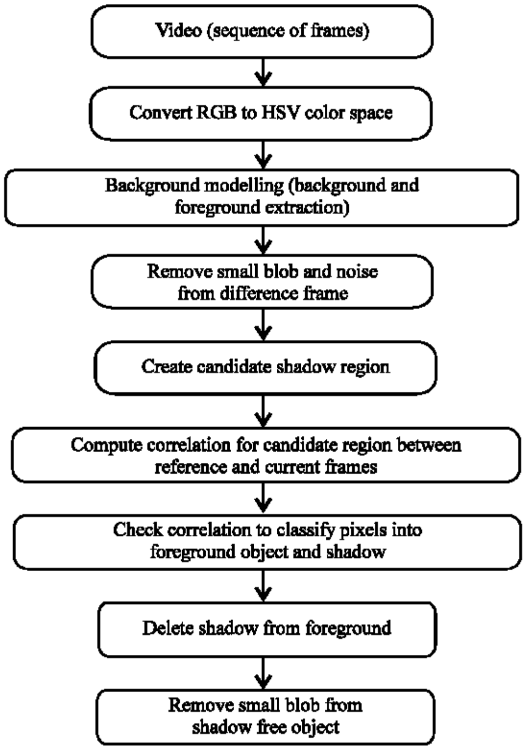

Fig. 3: HSV color space model

$$\theta = \cos^{-1}\left\{\frac{\frac{1}{2}((R\text{-}G)+(R\text{-}B))}{\sqrt{(R\text{-}G)^2+(R\text{-}B)(G\text{-}B)}}\right\} \quad (4)$$

Background modelling (background and foreground extraction): The main stages of background modelling and moving object detection are showed in Fig. 4. Construction of background reference is achieved by investment of N clean frames at beginning of the frames sequence and the frames are consecutive. When the N frames processing is finished the result will be a median value for each pixel.

Foreground detection, the pixel is classified as foreground or background by thresholding the difference between current frame and background reference. Background maintenance in this stage the background reference is updated by using current background reference, current frame and foreground mask. It is incremental online. The background maintenance and foreground detection stages are done frequently during the processing time. The median of n frames is computed which consider as reference frame. Equation 5 of this model is:

$$r(x,y,t) = \text{median}\,(pr(x,y,t\text{-}i)) \quad (5)$$
$$|c(x,y,t)\text{-}r(x,y,t)|>Th$$

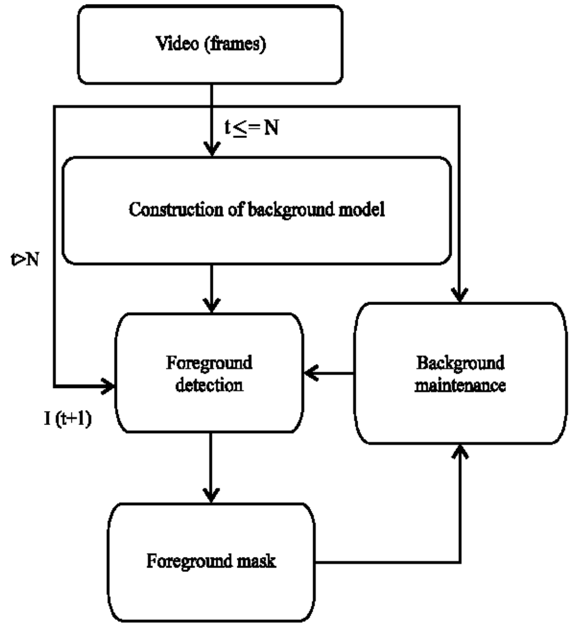

Fig. 4: Background subtraction system

$$i\in\{0, \dots, n\text{-}1\}$$

Where:

r = Reference frame
pr = Previous frames
c = Current frame
Th = Threshold, it depends on the result of Eq. 5 as it is shown in the following paragraph

Median model has several advantages, it is easy to perform and uses, very fast in term of real time requirements and the main advantage for this model is to avoid outsize values such as noise. In most methods which based on direct subtraction there is main problem represented by existence of one fixed threshold, Th for all frames. This lead to major problem is evident when illumination is changed. The proposed system has been solved this problem by using adaptive threshold depend on illumination variance for current frame compared with reference frame:

$$\text{Let } d(x, y) = r(x, y)\text{-}c(x, y) \quad (6)$$

Where:

d(x, y) = Difference frame
r(x, y) = Reference frame
c(x, y) = Current frame

The value is selected in Eq. 7 is Value (V) from HSV color space. The variance of illumination to specific frame with size h×w is:

$$\text{Var. (frame)} = \sum_{x=0}^{h}\sum_{y=0}^{w}\frac{(d(x,y)-\mu)^2}{h\times w} \quad (7)$$

where, μ is the mean of difference frame, the value of Var. (reference) is calculated based on frames have fixed illumination while Var. (current frame) is compared with Var. (reference) to update the value of threshold.

Remove small blob and noise: Difference frame always having small blobs and noise which lead to reduce performance of video tracking because all these blobs need to be addressed, so it is best to remove it. Noise deletes by morphological operations, it can be applied on a frame rely on the wanted results. Removing noise pixels is based on the neighbouring pixels. An example of removing a noise centre pixel is shown:

$$\begin{matrix}1 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 1\end{matrix} \quad \text{Becomes} \quad \begin{matrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{matrix}$$

While in term of small blob will be deleted based on a particular threshold which depends on the area of the blob when the area of blob is smaller than threshold then that blob will be removed.

Create candidate shadow pixels: Cucchiara *et al*. (2003) the intensity is measured directly by Value (V), pixels fall under the shadow have a lower value compared with pixels in the background. Subsequent the chromacity values, Hue (H) does not change in respect of shadow cast on background in term of Saturation (S) usually the shadow pixels have a lower value. Therefore, the pixel p can be classified as candidate shadow pixel if satisfy three conditions in:

$$\beta_1 \leq \left(F_P^V - B_P^V\right) \leq \beta_2 \quad (8)$$

$$\left(F_P^S - B_P^S\right) \leq \tau_s \quad (9)$$

$$|F_P^H - B_P^H| \leq \tau_H \quad (10)$$

In proposed system, the thresholds will be increased to maximize the chance to include all shadow pixels in candidate shadow pixels in addition, the values of β_1 and β_2 will be updated depend on var(current frame) to avoid the fail due to illumination change while τ_s and τ_H will have high values.

Compute correlation for candidate region between reference and current frames: To measure the correlation of candidate regions between reference frame and current frame, initially window with size 3×3 is selected then difference between the centre of window and neighbours are computed for each channel (H, S, V), the correlation is computed using (Eq. 11):

$$\text{corr} = \frac{\sum_{x=0}^{h}\sum_{y=0}^{w}\left(r(x,y)-\mu(r)\right)\left(c(x,y)-\mu(c)\right)}{\sqrt{\sum_{x=0}^{h}\sum_{y=0}^{w}\left(r(x,y)-\mu(r)\right)^2 \sum_{x-0}^{h}\sum_{y-0}^{w}\left(c(x,y)-\mu(c)\right)^2}} \quad (11)$$

Where:

r = Reference frame
c = Current frame
$\mu(r)$ = Mean of difference values of reference frame
$\mu(c)$ = Mean of difference values of current frame

Check correlation to classify pixels into foreground object and shadow: Weighted averaging method is proposed to cope with correlation results (H, S, V) that is weighting channels according to importance. The intensity represented by Value (V) take 0.5 followed by Saturation (S) take 0.3 and finally Hue (H) take 0.2 the equation becomes:

$$\text{corr}_t(x,y) = \text{corrv}(x,y)\,(0.5) + \text{corr}_s(x,y)\,(0.3) + \text{corr}_h(x,y)\times(0.2) \quad (12)$$

If the value of the $\text{corr}_t(x, y) > 0.7$ then the pixel is classified as shadow else is classified as foreground object from object detected regions.

Delete shadow from foreground: In this step the shadow pixels is ignored, the aim of shadow removing is to acquire a shadow free foreground object because remaining the shadow within foreground object leads to several problem in the next steps such as difficulty of shape analysis and feature extraction, etc.

Remove small blob from shadow free object: Sometimes, upon completion of treatment of shadow removal from foreground object, a number of small blobs may be appear with shadow free object these small blobs will be removed by dilation morphological operation using a 3×3 square structuring element.

RESULTS AND DISCUSSION

The benchmark sequences depend on Sanin *et al*. (2012) and six typical scenes which consist of outdoor and indoor scenes and the ground truth sequences were extracted manually. In six scenes, the first step, moving object detection is achieved and as a result of using adaptive threshold, the challenge of illumination change is addressed plus the using of median provides the ability

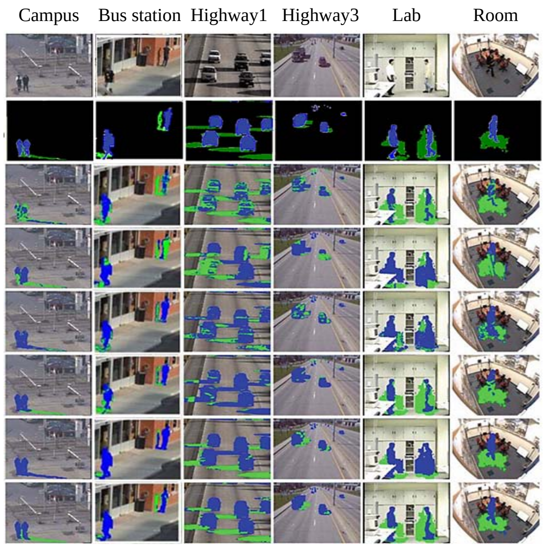

Fig. 5: Qualitative results of six samples, row 1 (frame), row 2 (ground truth), row 3 (chromacity), row 4 (geometry), row 5 (physical), row 6 (texture of small region), row 7 (texture of large region) and row 8 (proposed system)

to ignore noise then the detected object consists of object and shadow (blue and green region). These scenes consist of different shadow detection problems, the metrics of evaluation are as following:

$$\eta = \frac{TP_S}{TP_S+FN_S} \tag{13}$$

$$\xi = \frac{TP_F}{TP_F+FN_F} \tag{14}$$

$$\theta = \frac{2\eta\xi}{\eta+\xi} \tag{15}$$

The True Positive (TP) and False Negative (FN) pixels related to Shadows (S) or Foreground objects (F). The rate of shadow detection (η) and the rate of shadow discrimination (ξ) represent main metrics that commonly used to evaluate the performance of shadow detection techniques. And the comprehensive evaluate metric is represented by (θ) (Table 1).

In view of the different scenes, two values of parameters: τ_S = 48, τ_H = 40 are set which identified to weak detector. Sometimes there small blobs is remained in the difference frame which are filled using the morphological process (dilation). Figure 5 explains the qualitative results of six samples. In addition, Fig. 5 shows the detection rate η to sex benchmark methods.

Table 1: Detection rate (η) of six methods in different status

Methods	Campus	Bus station	High way 1	High way 3	Lab	Room
Chromacity	0.548	0.722	0.740	0.410	0.954	0.916
Geometry	0.628	0.488	0.671	0.427	0.403	0.527
Physical	0.459	0.370	0.454	0.392	0.342	0.530
Texture of SR	0.533	0.625	0.219	0.107	0.715	0.913
Texture of LR	0.043	0.236	0.584	0.390	0.776	0.818
Proposed system	0.745	0.757	0.685	0.439	0.810	0.931

CONCLUSION

In this study, a robust and reliable of adaptive background modelling and moving cast shadow removal approaches have been presented. In background modelling, the threshold is updated based on illumination level of current frame compared with reference frame. The proposed shadow detection and removal technique is a hybrid technique of chromacity that represents weak detector to select candidate shadow region and an enhanced form of the correlation method.

Enhancing of shadow removal results in more accurate, several important effects are obtained in term of video tracking performance such as: decreasing the probability of objects are merged together, decreasing the ratio of false positives produced by segmented shadows.

REFERENCES

Carmen, D.D.J.R. and R.D. Cajote, 2016. Moving shadow detection and removal for video-based traffic monitoring. Proceedings of the Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association (APSIPA), December 13-16, 2016, IEEE, Jeju, South Korea, ISBN: 978-1-5090-2401-8, pp: 1-7.

Cucchiara, R., C. Grana, M. Piccardi and A. Prati, 2003. Detecting moving objects, ghosts and shadows in video streams. IEEE Trans. Pattern Anal. Mach. Intell., 25: 1337-1342.

Desai, H.M. and V. Gandhi, 2014. A survey: Background subtraction techniques. Intl. J. Sci. Eng. Res., 5: 1365-1367.

Hadi, I. and M. Sabah, 2015. Improvement cat swarm optimization for efficient motion estimation. Intl. J. Hybrid Inf. Technol., 8: 279-294.

Hu, Y., B. Wang, Y. Wang and Y. Zhao, 2016. Moving shadows removal using HSV color space and texture analysis. Adv. Sci. Technol. Lett., 122: 240-246.

Huang, J.B. and C.S. Chen, 2009. Moving cast shadow detection using physics-based features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition CVPR, June 20-25, 2009, IEEE, Miami, Florida, ISBN:978-1-4244-3992-8, pp: 2310-2317.

Huerta, I., M. Holte, T. Moeslund and J. Gonzalez, 2009. Detection and removal of chromatic moving shadows in surveillance scenarios. Proceedings of the IEEE 12th International Conference on Computer Vision, September 29-October 2, 2009, IEEE, Kyoto, Japan, ISBN:978-1-4244-4420-5, pp: 1499-1506.

Hussain, Z. and A. Jebur, 2013. Shadow detection and removal in video sequence using color-based method. J. Kerbala Univ., 11: 239-257.

Javed, O. and M. Shah, 2002. Tracking and object classification for automated surveillance. Proceedings of the 7th European Conference on Computer Vision, May 27-June 2, 2002, Copenhagen, Denmark, pp: 343-357.

Kar, A. and K. Deb, 2015. Moving cast shadow detection and removal from Video based on HSV color space. Proceedings of the International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), May 21-23, 2015, IEEE, Dhaka, Bangladesh, ISBN:978-1-4673-6676-2, pp: 1-6.

Lakhotiya, S.A., M.D. Ingole and M.S. Joshi, 2013. Human object tracking using background subtraction and shadow removal technique. Intl. J. Comput. Sci. Appl., 6: 117-121.

Qin, R., S. Liao, Z. Lei and S.Z. Li, 2010. Moving cast shadow removal based on local descriptors. Proceedings of the 20th International Conference on Pattern Recognition (ICPR), August 23-26, 2010, IEEE, Istanbul, Turkey, ISBN:978-1-4244-7542-1, pp: 1377-1380.

Sanin, A., C. Sanderson and B.C. Lovell, 2012. Shadow detection: A survey and comparative evaluation of recent methods. Pattern Recognit., 45: 1684-1695.

Singh, S. and T. Patnaik, 2015. An efficient shadow removal method using HSV color space for video surveillance. Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI), August 10-13, 2015, IEEE, Kochi, India, ISBN:978-1-4799-8790-0, pp: 1454-1460.

Tian, Y.L., M. Lu and A. Hampapur, 2005. Robust and efficient foreground analysis for real-time video surveillance. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR Vol. 1, June 20-25, 2005, IEEE, San Diego, California, ISBN:0-7695-2372-2, pp: 1182-1187.

Wang, B., W. Zhu, Y. Zhao and Y. Zhang, 2015. Moving cast shadow detection using joint color and texture features with neighboring information. Proceedings of the Pacific-Rim Symposium on Image and Video Technology, November 23-27, 2015, Springer, Auckland, New Zealand, ISBN:978-3-319-30284-3, pp: 15-25.