

# Moving Object Detection for Video Surveillance using Different Color Spaces

Adesh Hardas  
Vidyalankar Institute of  
Technology  
Mumbai

Dattatray Bade  
Vidyalankar Institute of  
Technology  
Mumbai

Vibha Wali  
Vidyalankar Institute of  
Technology  
Mumbai

## ABSTRACT

In many vision based application identifying moving objects is important and critical task. For different computer vision application Background subtraction is fast way to detect moving object. Background subtraction separates the foreground from background. However, during background subtraction pixels belonging to shadow are misclassified as foreground object. Moving cast shadow associated with moving object also gets detected making it challenge for video surveillance. Now days many methods are available for background subtraction. The core of background subtraction is background modeling. Gaussian Mixture model is good balance between accuracy and complexity. However, it is difficult to determine the optimal color space in which to remove shadow. In this paper, we study the features of moving object and shadow in different color spaces to solve the problem.

## Keywords

Background subtraction, Gaussian mixture model, Shadow removal, RGB, HSV, and YCbCr color spaces.

## 1. INTRODUCTION

In Video surveillance, motion detection has to deal with several problems [1]. One of the main problems is how to distinguish between a moving object and its moving shadow. Moving cast shadow associated with moving object also gets detected makes it difficult to detect the exact shape of object and to recognize the object. Moving object detection applications are normally car detection, person identification or wild life monitoring. Therefore it is required that background subtraction method should be fast and simple. Background model is also classified as static or dynamic. Dynamic background model consist of moving background mostly in outdoor environment.

Generally moving shadow is detected and eliminated in color space, because many motion detection algorithms are used in color spaces and color video is the main format. In *RGB* color space, there are two properties of shadow, that is, pixel value in moving shadow region is darker than in background scene, and statistic property in shadow region has little variety based on these traits. In *HSV* color space, Cucchiara et al. [5] eliminated outdoor vehicle's moving shadow by using invariance of chrominance. In [3] Kumar studies five color spaces (*RGB*, *HSV*, *XYZ*, *YCrCb* and normalized *rgb*) for motion detection, but the characteristics of moving object their shadow are not mentioned.

There are three basic facts with moving shadow detection and elimination in color spaces.

- shadow have different classes due to various scenes
- In various color spaces shadow is different, therefore different detection results.

- There might be some optimal color space for particular shadow elimination

The shadow points and the object points share two important visual features: motion model and detectability. Since the most common techniques for foreground object detection in dynamic scene are inter-frame difference or background suppression, all the moving points of both objects and shadows are detected at the same time. In this paper we address an approach to select optimal color space for particular shadow elimination among a set of color spaces, such as *RGB*, *HSV* and *YCbCr*.

Background subtraction techniques are mostly used for motion detection in many real-time vision surveillance applications. In these approaches, difference between the coming frame and the background image is performed to detect foreground objects. Background subtraction provides the most complete feature data, but is extremely sensitive to dynamic scene changes due to illumination changes and extraneous events. Most researchers are now devoted to developing robust background models to prevent falseness in motion detection caused by scene changes. For example, the algorithm proposed by Stauffer and Grimson [9] uses a mixture of Gaussian distributions to model a multimodal background image sequence, and an online estimation technique is used to update the background model.

The outline of this paper is as follows. In section 2 Background subtraction method is presented. In section 3 characteristics of shadow in different color space are analyzed and shadow is classified. Section 4 is experimental results and discussions. Section 5 is the conclusions.

## 2. BACKGROUND SUBTRACTION METHOD

In many computer visions application background subtraction is fast way to detect moving object from video by static camera. Background subtraction techniques have to classify each pixel as part of foreground or background. If there is a too much variation in pixel value then it is considered as moving object. According to Cheung and Kamath [12] Background subtraction techniques are categorized as: 1) non-recursive and 2) recursive. Non recursive techniques are adaptive to scene changes depending on the buffer size. Recursive technique do not require buffer. They try to update the background model recursively using either a single or multiple model(s) as each input frame is observed.

Frame differencing is the simplest method in Background subtraction. A background image without any moving objects of interest is taken as the reference image. Pixel value for each co-ordinate (*x*, *y*) for each color channel of the background image is subtracted from the corresponding pixel value of the input image. If the resulting value is greater than a particular threshold value, then that is a foreground pixel otherwise background [6]. If

$$| \text{frame } i - \text{background } i | > \text{threshold} \quad (1)$$

then the pixel  $i$  is foreground.

Although such techniques are very fast, the segmentation performance can be quite poor, especially with fluctuating illumination conditions. In order to manage the change in background several complex background methods have been developed.

## 2.1 Mixture of Gaussian model:

The Mixture of Gaussians technique was first introduced by Stauffer and Grimson in [9]. This method represents each pixel of the scene by using a mixture of normal distributions to handle multimodal background images from video. It finds the difference of the current pixel's intensity value and cumulative average of the previous values. So it keeps a cumulative average ( $\mu_t$ ) of the recent pixel values. If the difference of the current image's pixel values and the cumulative pixel value is greater than the product of a constant value and standard deviation then it is classified as foreground. In this method each pixel is modeled as mixture of  $k$  normal distribution. The probability that a certain pixel has a value  $X_t$  at time can be written as

$$P(X_t) = \sum_{i=1}^k \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2)$$

Where  $k$  is the number of distributions (currently, 3 to 5 is used),  $\omega_{i,t}$  is the weight of the  $k$ th Gaussian in the mixture at time  $t$  and  $\eta(X_t, \mu_{i,t}, \Sigma_{i,t})$  the Gaussian probability density function. For computational reasons, the covariance matrix is assumed to be of the form

$$\Sigma_{k,i} = \sigma_k^2 I \quad (3)$$

Where  $\sigma$  is the standard deviation. This assumes that the red, green, and blue pixel values are independent and have the same variance, allowing us to avoid a costly matrix inversion at the expense of some accuracy. That is, at each  $t$  frame time, the  $I_t$  pixel's value can then be classified as foreground pixel if the inequality:

$$|I_t - \mu_t| > k \sigma \text{ holds;} \quad (4)$$

otherwise, it can be considered as background, where  $k$  is a constant and  $\sigma$  is standard deviation [8].

Here background is updated as the running average:

$$\mu_{t+1} = \mu_t * \alpha + (1 - \alpha) * I_t \quad (5)$$

$$\sigma_{t+1}^2 = \alpha (I_t - \mu_t)^2 + (1 - \alpha) \sigma_t^2 \quad (6)$$

where  $\alpha$ , the learning rate, is typically 0.05,  $I_t$  is the pixels current value and  $\mu_t$  is the previous average. Gaussians are ordered by the value of  $\omega/\sigma$ .

Then the first  $B$  distributions are chosen as the background model, where

$$B = \arg \min_b \left( \sum_{k=1}^b \omega_j > T \right) \quad (7)$$

## 3. SHADOW DETECTION AND REMOVAL

A major drawback of background subtraction techniques is the undesired detection of shadows as foreground objects. Shadows are due to the occlusion of light source by an object

in the scene. In particular, that part of the object not illuminated is called self-shadow, while the area projected on the scene by the object is called cast shadow [11]. This last one is more properly called moving cast shadow if the object is moving.

For Human Vision System moving shadow should be adapted the following conditions.

- Shadow is projection of moving object in background.
- Shadow always is related with moving object. It reflects corresponding motion and behaviours of object.
- The shape of moving shadow could change with motion every time.
- The pixel values of shadow are darker than that of surrounding scene.

As mentioned in [2, 9], shadow can be classified self shadow and cast shadow as follows (Fig. 1).

*Self-shadow* - the part of an object that is not illuminated (Fig. 1 region A)

*Cast shadow* - the area projected on the scene by the object. It can be divided into *Umbra* (dark shadow, Fig. 1, region B) and *Penumbra* (soft shadow, Fig. 1, region C).

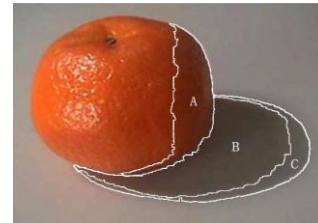


Fig. 1 Shadow Classification

## 3.1 RGB Color Space

RGB is additive in nature. It is sum of three primary colors red, green and blue. RGB is one of most widely used color space for processing and storing the digital image data. RGB space based method proposed by Horprasert in [4] saves the computational cost. The basic idea in [4] is that shadow has similar chromaticity but lower brightness. For a given observed pixel value  $I_i$ , a brightness distortion,  $\alpha_i$ , and a color distortion  $CD_i$ , is calculated by,

$$\alpha_i = \arg \min (I_i - \alpha_i E_i)^2 \quad (8)$$

$$CD_i = \| I_i - \alpha_i E_i \|$$

Where  $E$  is the expected chromaticity line,  $\alpha_i$  equals 1 if the brightness of the given pixel in the current frame is the same as in the background image.  $\alpha_i$  is less than 1 if it is darker and greater than 1 if it becomes brighter than the expected brightness.

In the proposed method to discriminate the shadow pixel and the object pixel, after background subtraction mean value of R, G and B channels is calculated. This mean value is compared with predefined threshold value. If the current pixel value is less than threshold the output is set to 1 as shadow region and if the value is greater than threshold then the output is set to 0 as foreground region. This way shadow mask and foreground mask is calculated.

## 3.2 HSV Color Space

Cucchiara et al. [5] operate brightness, saturation, and hue properties in the HSV color space. HSV (hue, saturation, value) is often more natural to think about a color in terms of

hue and saturation than in terms of additive or subtractive color components. The appearance of a point belonging to a cast shadow can be described as:

$$S_k(x, y) = E_k(x, y) \times \rho_k(x, y) \quad (9)$$

Where  $S_k$  is the luminance of the pixel in location  $(x, y)$ .  $E_k(x, y)$  is the irradiance and it is computed as:

$$E_{(k)}(x, y) = \begin{cases} C_A + C_p \cos(N(x, y), L) & \text{Illuminated} \\ C_A & \text{Shadowed} \end{cases} \quad (10)$$

Where  $C_A$  and  $C_p$  are the intensity of the ambient light and of the light source, respectively,  $L$  the direction of the light source and  $N(x, y)$  is object surface normal.  $\rho_k(x, y)$  is the reflectance of the object surface.

In [14] the difference between the current frame and a reference image is given by

$$D_k(x, y) = S_{k+1}(x, y) - S_k(x, y) \quad (11)$$

According to the assumption in [14] we have

$$\rho_{k+1}(x, y) = \rho_k(x, y) = \rho(x, y) \quad (12)$$

$$D_k(x, y) = \rho(x, y) C_p \cos \angle(N(x, y), L) \quad (13)$$

The approach in [14] exploits the local appearance change due to shadow by computing the ratio  $R_k(x, y)$  between the appearance of the pixel in the actual frame and the appearance in a reference frame:

$$R_k(x, y) = \frac{S_{k+1}(x, y)}{S_k(x, y)} \quad (14)$$

The ratio can be written with respect to irradiance and reflectance by using eq. 9 and eq. 12 as

$$R_k(x, y) = \frac{E_{k+1}(x, y)}{E_k(x, y)} \quad (15)$$

If static background point is covered by a shadow, we have:

$$R_k(x, y) = \frac{C_A}{C_A + C_p \cos \angle(N(x, y), L)} \quad (16)$$

This ratio is less than 1. In fact, the angle between  $N(x, y)$  and  $L$  is in range between  $-\pi/2$  to  $\pi/2$  therefore the Cos function is always positive.

In the proposed method from Hue, Saturation and Value (HSV) Value is compared with threshold as follows:  $0.25 < V < 0.4$ . This simplifies the computational complexity and gives better result.

### 3.3 YCbCr Color Space

In YCbCr color space  $Y'$  is the luma component and Cb and Cr are the blue - difference and red - difference chrominance components. Human eye is more sensitive to light changes compared to color changes. This color space makes use of this property. Here intensity component is stored with higher accuracy than the Cb and Cr components. YCbCr color space is found to be more robust to illumination change and have independent components of luminance and chrominance. For indoor shadow, the pixel values are few smaller than other background scene. Experimental results showed that YCbCr can handle shadows created by moving objects and can also reduce noise. According to shadow model, some channels are shifting invariant.

$$Cb(R+\Delta I, G+\Delta I, B+\Delta I) = Cb(R, G, B)$$

$$Cr(R+\Delta I, G+\Delta I, B+\Delta I) = Cr(R, G, B) \quad (17)$$

So light / indoor shadows could be eliminated easily in these channels. In the proposed method Y channel is compared with threshold ( $105 < th < 65$ ) for shadow detection.

## 4. EXPERIMENTS AND RESULTS

We employ background subtraction model to detect motion regions and shadow removal method to eliminate shadow and compare the results of experiments. The performance of proposed method in different colour spaces is analysed. In [2], the authors proposed two metrics for moving object evaluation: shadow detection rate and shadow discrimination rate. In this paper, we have eliminated background and shadow therefore only moving object detection rate (MODR) is calculated.

$$\tau = \frac{TP}{(TP + FN)} \quad (18)$$

where  $TP$  is the true positives,  $FN$  is the false negatives. Figure 2, 3, 4 and Table 1 are the results of experiments.

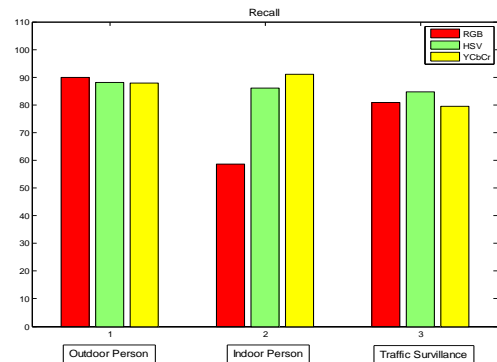
**Table 1. Moving object detection rate ( $\tau$ ) for various color spaces**

	Outdoor person	Indoor person	Traffic_1
RGB	90.40	58.65	80.87
HSV	88.20	86.00	84.79
YCbCr	87.85	91.11	79.55

RGB colour space: Because RGB colour space is sensitive to illumination. It eliminates many foreground objects pixel when used in indoor application (Figure 4). It is suitable to eliminate dark / outdoor Shadow (Figure 2).

HSV colour space: It is also sensitive to illumination and view direction. But because luminance is separated from chrominance, the space is affected less by dark / visible shadows (Figure 2). Although light / indoor shadow can be eliminated, some object's pixels vanished because invisible shadow have similar values with scene. From quantity estimation, HSV space gives better result in all application.

YCbCr colour space: It is not sensitive to illumination. In this colour space luminance and chrominance are separated (same as HSV space). Therefore, YCbCr colour space is fit for light / indoor shadow (Figure 4).



**Fig 5. Graphical representation of  $\tau$  (%)**

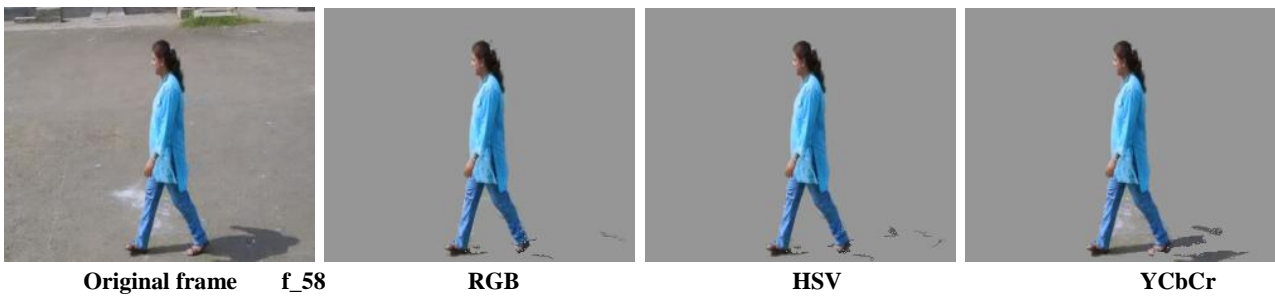


Fig. 2 Detection results for outdoor person



Fig. 3 Detection results for Traffic Surveillance

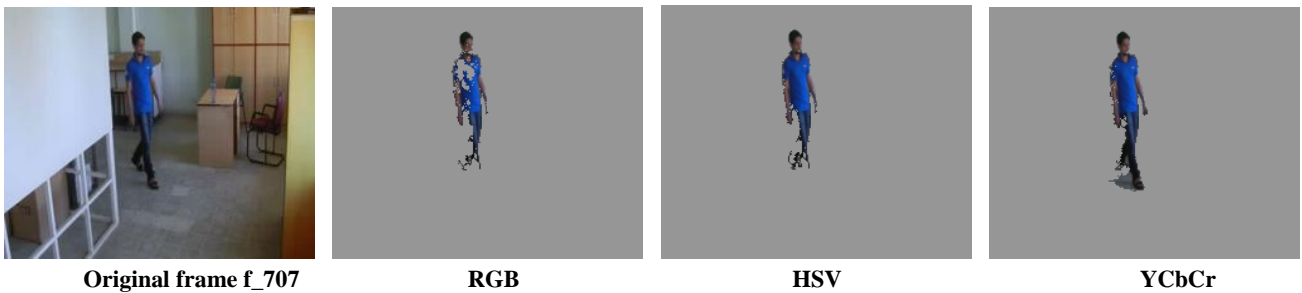


Fig. 4 Detection result for Indoor Person

## 5. CONCLUSION

In the moving object detection process, one of the main challenges is to differentiate moving objects from their cast shadows. Therefore, the accurate detection of a moving object and the acquisition of its exact shape by removing shadows have great influence on the performance of subsequent steps such as tracking, recognition, classification, and activity analysis.

After studying various methods from the literature results indicate that critical tradeoffs are always present between the accuracy and the real time performance of the method. The choice of algorithm for background modeling should be made according to the required application. According to analysis of experiments, every space cannot be suitable for all kinds of shadows. HSV and RGB spaces are suitable for dark / visible shadow, and YCbCr is suitable for light / indoor shadow.

Table shows the comparative analysis of reviewed techniques based on some parameters. This paper will help the newcomers to choose the appropriate technique to their research work. The proposed method is tested on three color spaces; other color spaces can also be compared for better analysis. The current system can be extended to build systems intelligent analysis from video sequences for various applications.

## 6. REFERENCES

[1] W.M.Hu, T.Tan, et al. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on*

*Systems, Man, and Cybernetics - PART C: Applications and Reviews*, 34(3):334–351, 2004.

- [2] A.Prati et al. Detecting moving shadows: algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:918–923, 2003.
- [3] P.Kumar, K.Sengupta, et al. A comparative study of different color spaces for foreground and shadow detection for traffic monitoring system. In *Proceedings of IEEE the 5th International Conference on Intelligent Transportation Systems*, pp.100–105, 2002.
- [4] T. Horprasert, D. Harwood, and L.S. Davis. A robust background subtraction and shadow detection. *Proc. ACCV*, 2000.
- [5] R.Cucchiara. Improving shadow suppression in moving object detection with hsv color information. In *Proceedings of IEEE Conference on Intelligent Transportation Systems*, pp.334–339, 2001.
- [6] Zhan Chaohui Duan Xiaohui, Xu Shuoyu Song Zheng Luo Min, 2007. An improved Moving Object Detection Algorithm Based on Frame Difference and Edge Detection”, Fourth International Conference on Image and Graphics 0-7695-2929-1/07 \$25.00, 2007 IEEE.
- [7] Shih-Chieh Wang, Te-Feng Su and Shang-Hong Lai”Detecting Moving Object From Dynamic Background with Shadow Removal”*ICASSP2011 IEEE*

- [8] A. Elgammal, D. Hardwood, and L. Davis. Non-parametric model for background subtraction. In *ECCV00*, pages 751–767, 2000.
- [9] C. Stauffer, WEL. Grimson, “Adaptive Background Mixture Models for Real- Time Tracking,” *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition CVPR*,; vol.2, pp. 246-252, 1999.
- [10] Jianhua Ye, Tao Gao, Jun Zhang, “Moving object detection with background subtraction and shadow removal” in the proceedings of International Conference on Fuzzy Systems and Knowledge Discovery 978-1-4673-0024-7/10 ©2012 IEEE 2012.
- [11] C. Jiang and M.O. Ward, —Shadow identification, in Proceedings of IEEE Int’l Conference on Computer Vision and Pattern Recognition, , 1992 ,pp. 606–612.
- [12] Sen-Ching S. Cheung, Chandrika Kamath, 2005. Robust Background Subtraction with Foreground Validation for Urban Traffic Video, *EURASIP Journal on Applied Signal Processing*, 14: 2330-2340, Hindawi Publishing Corporation.
- [13] McFarlane and C. Schofield. Segmentation and tracking of piglets in images. *MVA*, 8:187–193, 1995.
- [14] J. Stauder, R. Mech, and J. Ostermann, —Detection of moving cast shadows for object segmentation, in proceedings of IEEE Transactions on Multimedia, vol. 1, no. 1, Mar. 1999, pp. 65-76.