Background subtraction in video using recursive mixture models, spatio-temporal filtering and shadow removal

Zezhi Chen[1], Nick Pears[2], Michael Freeman[2] and Jim Austin[1, 2]

[1]Cybula Limited, York, UK
[2]Department of Computer Science, University of York, York, UK

Abstract. We describe our approach to segmenting moving objects from the color video data supplied by a nominally stationary camera. There are two main contributions in our work. The first contribution augments Zivkovic and Heijden's recursively updated Gaussian mixture model approach, with a multidimensional Gaussian kernel spatio-temporal smoothing transform. We show that this improves the segmentation performance of the original approach, particularly in adverse imaging conditions, such as when there is camera vibration. Our second contribution is to present a comprehensive comparative evaluation of shadow and highlight detection appoaches, which is an essential component of background subtraction in unconstrained outdoor scenes. A comparative evelaution of these approaches over different color-spaces is currently lacking in the literature. We show that both segmentation and shadow removal performs best when we use RGB color spaces.

1 Introduction

We consider the case of a nominally static camera observing a scene, such as is the case in many visual surveillance applications, and we aim to generate a background/foreground segmentation, with automatic removal of any shadows cast by the foreground object onto the background. In real applications, cameras are often mounted metal poles, which can oscillate in the wind, thus making the problem more difficult. This problem is also addressed in this paper.

To segment moving objects, a background model is built from the data and objects are segmented if they appear significantly different from this modelled background. Significant problems to be addressed include (i) how to correctly and efficiently model and update the background model, (ii) how to deal with camera vibration and (iii) how to deal with shadows. In this paper our contributions are a spatio-temporal filtering improvement to Zivkovic's recursively updated Gaussian mixture model approach [1], and a comprehensive evaluation of shadow/highlight detection across different color spaces, which is currently lacking in the literature. We also present quantitative results of our complete foreground/background segmentation system with shadow removal in several real-world scenarios. This is valuable to those developing pragmatic visual surveillance solutions that demand a high quality foreground segmentation.

A robust visual segmentation system should not depend on careful placement of the camera, rather it should be robust to whatever is in its visual field, whatever lighting effects occur or whatever the weather conditions. It should be capable of dealing with movement through cluttered areas, objects overlapping in the visual field, shadows, lighting changes, effects of moving elements of the scene (e.g. camera vibration, swaying trees) and slow-moving objects. The simplest form of the background model is a time-averaged background image. However, this method suffers from many problems, for example it requires a large memory and a training period absent of foreground objects. Static foreground objects during the training period would be considered as a part of background. This limits their utility in real time applications.

A Gaussian mixture model (GMM) was proposed by Friedman and Russell [2] and it was refined for real-time tracking by Stauffer and Grimson [3]. The algorithm relies on the assumptions that the background is visible more frequently than any foreground regions and that it has models with relatively narrow variances. The system can deal with real-time outdoor scenes with lighting changes, repetitive motions from clutter, and long-term scene changes. Many adaptive GMM model have been proposed to improve the background subtraction method since that original work. Power and Schoonees [4] presented a GMM model employed with a hysteresis threshold. They introduced a faster and more logical application of the fundamental approximation than that used in the paper [5]. The standard GMM update equations have been extended to improve the speed and adaptation of the model [6][7]. All these GMMs use a fixed number of components. Zivkovic et al. [1] presented an improved GMM model adaptively chooses the number of Gaussian mixture components for each pixel on-line, according to a Bayesian perspective. We call this method the Zivkovic-Heijden Gaussian mixture model (ZHGMM) in the remainder of this paper.

Another main challenge in the application of background subtraction is identifying shadows that objects cast which also move along with them in the scene. Shadows cause serious problems while segmenting and extracting moving objects due to the misclassification of shadow points as foreground. Prati et al. [8] presented a survey of moving shadow detection approaches. Cucchiardi et al. [9] proposed the detection of moving objects, ghosts and shadows in HSV colour space and gave a comparison of different background subtraction methods.

This paper focuses on two issues: 1) How to get a robust GMM, which models the real background as accurately as possible, and can deal with lighting changes in difficult and challenging environments, such as bad weather and camera vibration. 2) How to remove the shadows and highlight reflections, since these can affect many subsequent tasks such as foreground object classification. The contributions of this paper are (i) an improvement to the ZHGMM algorithm, using a multi-dimensional spatio-temporal Gaussian kernel smoothing transform and (ii) a comprehensive survey of moving shadow and highlight reflection detection approaches in various colour spaces for moving object segmentation applications.

The paper is organised as follows: In next section the ZHGMM approach is reviewed. In Section 3, ZHGMM with multi-dimensional Gaussian kernel density transform (MDGKT) is proposed. The training of the MDGKT is given in Section 4. A comprehensive analysis of various shadow removal methods in given in Section 5. Section 6 gives a quantitative evaluation of the background model update and foreground object segmentation. Finally, we present conclusions in Section 7.

2 ZHGMM review

In this section, we provide a brief outline of the recursive mixture model estimation procedure described by Zivkovic et al [1] [10]. First, we choose a reasonable time adaption period of T frames (eg T=100 frames) over which to generate the background model so that, at time t, we have the training set $X_T = \{x^{(t)}, x^{(t-1)}, \cdots, x^{(t-T)}\}$ for each pixel. For each new sample, we update the training data set X_T and re-estimate the density. In general, these samples contain values that belong to both the background (BG) and foreground (FG) object(s). Therefore, we should denote the estimated density as $\hat{p}(x^{(t)} | X_T, BG + FG)$. We use a GMM with *M* components (we set it as 4).

$$\hat{p}(x^{(t)} | X_T, BG + FG) = \sum_{m=1}^{M} w_m N(x^{(t)}; \mu_m, \Sigma_m) \tag{1}$$

where μ_m is the estimate of the mean of *m*th Gaussian and $\sum_m$ is the estimate of the variances that describe the *m*th GMM component. For computational reasons (easily invertible), an assumption is usually made that the dimensions of X_T are independent so that $\sum_m$ is diagonal. A further assumption is that the components (eg red, green and blue pixel values) have the same variances [3] so that the covariance matrix is assumed to be of the form $\sum_m = \sigma_m I$, where *I* is a 3×3 identity matrix. Note that a single σ_m may be a reasonable approximation in a linear colour space, but it may be an excessive simplification in non-linear colour spaces. Thus, in this work, the covariance of a Gaussian component is diagonal, with three separate estimates of variance. The estimated mixing weights (what portion of the data is accounted for by this Gaussian) of *m*th Gaussian in the GMM at time *t*, denoted by w_m, are non-negative and normalized.

Given a new data sample $x^{(t)}$ at time t, the recursive update equations are

$$w_m \leftarrow w_m + \alpha(o_m^{(t)} - w_m) + \alpha c_T \tag{2}$$

$$\mu_m \leftarrow \mu_m + o_m^{(t)} (\alpha / w_m) \delta_m \tag{3}$$

$$\sigma_m^2 \leftarrow \sigma_m^2 + o_m^{(t)} (\alpha / w_m)(\delta_m^T \delta_m - \sigma_m^2) \tag{4}$$

where $x^{(t)} = [x_1, x_2, x_3]^T$, $\mu_m = [\mu_1, \mu_2, \mu_3]^T$, $\delta_m = [\delta_1, \delta_2, \delta_3]^T$, $\sigma_m^2 = [\sigma_1^2, \sigma_2^2, \sigma_3^2]^T$ for a 3 channel colour image, $\delta_m = x^{(t)} - \mu_m$. Instead of the time interval T, mentioned above, here the constant α defines an exponentially decaying envelope that is used to limit the influence of the old data, and we note that $\alpha = 1/T$. c_T is the negative Dirichlet prior evidence weight [1], which means that we will accept that the class exists only if there is enough evidence from the data for its existence. It will suppress the components that are not supported by the data and we discard the components with negative weights. This also ensures that the mixture weights are non-negative.

For a new sample the ownership o_m^T is set to 1 for the "close" component with largest w_m and the others are set to zero. We define that a sample is "close" to a component if the Mahalanobis distance (MD) from the component is, for example, less than three. The squared Mahalanobis distance from the m*th* component is calculated as $D_m^2\left(x^{(t)}\right) = \delta_m^T \Sigma_m^{-1} \delta_m$. If there are no "close" components a new component is generated with $w_{m+1} = \alpha$, $\mu_{m+1} = x^{(t)}$, $\sigma_{m+1} = \sigma_0$, where σ_0 is some appropriate initial variance. If the maximum number of components *M* is reached, we discard the component with smallest w_m. After each weight update, using equation (2), we need to renormalize the weights so that they again sum to unity.

3 ZHGMM with Multi-dimensional Gaussian kernel density transform

An image is typically represented as a two-dimensional matrix of p-dimensional vectors, where p=1 in the gray-level case, p=3 for colour images, and p>3 for multispectral images. The space of the matrix is known as the *spatial* domain, while the gray, colour or multispectral is known as the *spectral* domain [11] [12]. For algorithms that use image sequences, there is also the *temporal* domain.

In order to provide spatio-temporal smoothing for each spectral component, a multivariate kernel is defined as the product of two radially symmetric kernels and the Euclidean metric allows a single bandwidth parameter for each domain.

$$K_{h_t,h_s}(x) = \frac{C}{h_s h_t} k\left(\left\|\frac{x^s}{h_s}\right\|^2\right) k\left(\left\|\frac{x^t}{h_t}\right\|^2\right) \tag{5}$$

where x^s is the spatial part and x^t is the temporal part of the feature vector. $k(x)$ is a common kernel profile (we use Gaussian) used in both spatial and temporal domains, h_s and h_t are the kernel bandwidths, and C is the corresponding normalization constant. In order to improve stability and robustness of the ZHGMM, we have used this *Multi-Dimensional Gaussian Kernel density Transform* (MDGKT) as a pre-process, which only requires a pair of bandwidth parameter (h_s, h_t) to control the size of the kernel, thus determining the resolution and time interval of the ZHGMM.

4 Online training of MDGKT

A sample RGB image is shown in Fig.1 (a). The variation of red and blue values of a pixel stream over 596 frames is shown in Fig.1 (e) and (f). The black curves show the variation of the original red and blue components, and the red curves illustrate the variation of red and blue components in the MDGKT image. A Gaussian kernel with bandwidth $(h_s, h_t) = (5,5)$ and standard deviation (*std*) of 0.5 was chosen as the kernel profile. The *std* of the original image is 1.834 and 1.110, but the *std* of MDGKT output image is only 1.193 and 0.832. Obviously, MDGKT reduces the std figures. Fig.1

(c) and (d) show the scatter plots of the original and MDGKT image (red, blue) values of the same pixel. Fig.1 (d) shows that the distribution of MDGKT image is more localised within two Gaussian components of the mixture model, illustrating the effect of the spatio-temporal filtering in the spectral domain. The mixture of these two Gaussians for the blue colour component of the original pixel and the estimated GMM distribution using MDGKT are shown in Fig.1 (b).

The MDGKT algorithm described above allows us to identify the foreground pixels in each new frame while updating the description of each pixel's background model. This procedure is effective in determining the boundary of moving objects, thus moving regions can be characterized not only by their position, but also size, aspect ratio, moments and other shape and colour information. These characteristics can be used for later processing and classification, for example, using a support vector machine [13]. To analyse the performance of the algorithm, we used a dynamic scene. The results are shown in Fig.2. (a) and (d) are original images. One is outside scene, another is inside scene. (b) and (e) are the results of the ZHGMM algorithm. (c) and (f) are the results of our MDGKT algorithm. Note that the results shown are without the application of any post-processing.

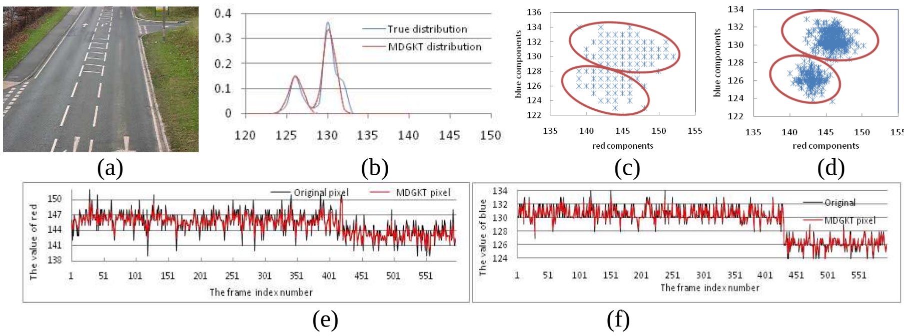


Fig. 1. The effect of spatio-temporal filtering. (a) A sample image. (b) GMM distribution of the blue component value of a sample pixel. (c) and (d) scatter plots of corresponding pixel in original images and MDGKT images respectively. (e) and (f) show the variation of red and blue colour components over time (red trace is spatio-temporally filtered)

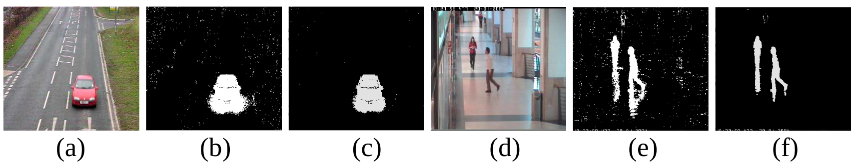


Fig. 2. Comparative results of ZHGMM and MDGKT algorithms

5 Shadow removal

The previous section showed promising initial results for our MDGKT background subtraction algorithm. However, the algorithm is susceptible to both global and local illumination changes such as shadows and highlight reflections (specularities). These often cause subsequent processes, such as tracking and recognition, to fail. Prati et al. [8] present a comprehensive survey of moving shadow detection approaches. It is important to recognize the type of features utilized for shadow detection. Some approaches improve performance by using spatial information working at a region level or at a frame level instead of pixel level [14]. Finlayson et al. [15] proposed a method to remove shadows from a still image using illumination invariance. We give a comparison of several different shadow removal methods, working in different colour spaces, below. For the sake of clarity, we distinguish two different foreground segmentations: segmentation F1, is the foreground segmentation which includes shadows (MDGKT segmentation output), while F2 is the foreground segmentation after we have removed shadows.

5.1 Working with RGB and normalized RGB colour space

(i) RGB colour The observed colour vector is projected onto the expected colour vector, and the *i*th pixel's brightness distortion α_i is a scalar value (less than unity for a shadow) describing the fraction of remaining 'brightness'. This may be obtained by minimizing [16]

$$\phi(\alpha_i) = (I_i - \alpha_i E_i)^2 \tag{6}$$

where $I_i = [I_{Ri}, I_{Gi}, I_{Bi}]$ denotes the *i*th pixel value in RGB space, $E_i = [\mu_{Ri}, \mu_{Gi}, \mu_{Bi}]$ represents the ith pixel's expected (mean) RGB value in MDGKT. The solution to equation (6) is an alpha value equal to the inner product of I_i and E_i, divided by the square of the Euclidean norm of E_i.

Colour distortion is defined as the orthogonal distance between the observed colour and the expected colour vector. Thus, the colour distortion of the *i*th pixel is $CD_i = \|I_i - \alpha_i E_i\|$. If we balance the colour bands by rescaling the colour values by the pixel *std* $s_i = [\sigma_{Ri}, \sigma_{Gi}, \sigma_{Bi}]$, the brightness and chromaticity distortion become

$$\alpha_i = \frac{I_{Ri}\mu_{Ri}/\sigma_{Ri}^2 + I_{Gi}\mu_{Gi}/\sigma_{Gi}^2 + I_{Bi}\mu_{Bi}/\sigma_{Bi}^2}{[\mu_{Ri}/\sigma_{Ri}]^2 + [\mu_{Gi}/\sigma_{Gi}]^2 + [\mu_{Bi}/\sigma_{Bi}]^2} \tag{7}$$

$$CD_i = \sqrt{(I_{Ri} - \alpha_i\mu_{Ri})^2/\sigma_{Ri}^2 + (I_{Gi} - \alpha_i\mu_{Gi})^2/\sigma_{Gi}^2 + (I_{Bi} - \alpha_i\mu_{Bi})^2/\sigma_{Bi}^2} \tag{8}$$

Then a pixel in the foreground segmentation (F1) may be classified as either a shadow or highlight on the true background as follows:

$$\begin{cases} Shadow \quad CD_i < \beta_1 \quad and \quad \alpha_i < 1 \\ Highlight \quad CD_i < \beta_1 \; and \; \alpha_i > \beta_2 \end{cases} \tag{9}$$

β_1 is a selected threshold value, used to determine the similarities of the chromaticity between the MDGKT and the current observed image. If there is a case where a pixel from a moving object in the current image contains a very low RGB value, then this dark pixel will always be misclassified as a shadow, because the value of the dark pixel is close to the origin in RGB space and all chromaticity lines in RGB space meet at the origin. Thus a dark colour point is always considered to be close or similar to any chromaticity line. We introduce a threshold β_2 for the normalized brightness distortion to avoid this problem. This is defined as: $\beta_2 = 1/(1-\varepsilon)$, where ε is a lower band for the normalized brightness distortion. An automatic threshold selection method was provided by Horprasert et al. [16].

(ii) Normalized RGB Given three colour variables, R_i, G_i and B_i, the chromaticity coordinates are $r_i = R_i/(R_i+G_i+B_i)$, $g_i = G_i/(R_i+G_i+B_i)$ and $b_i = B_i/(R_i+G_i+B_i)$, where $r_i + g_i + b_i = 1$ [17]. $s_i = R_i + G_i + B_i$ is a brightness measure. Let a pixel value of the background MDGKT be $< r_i, g_i, s_i >$. Assume that this pixel is covered by a shadow in frame t and let $< r_{ti}, g_{ti}, s_{ti} >$ be the observed value for this pixel at this frame. Then, for a pixel in the foreground segmentation (F1):

$$\begin{cases} Shadow & \beta_1 < s_{ti}/s_i \le \beta_2 \\ Highlight & \beta_3 < s_{ti}/s_i \end{cases} \tag{10}$$

where β_1, β_2 and β_3 are selected threshold values used to determine the similarities of the normalized brightness between the MDGKT and the current observed image. It is expected that, in the shadow area, the observed value s_{ti} will be darker than the normal value s_i, up to a certain limit. On the other hand, in the highlight area, $s_{ti} > s_i$. So that $\beta_1 > 0, \beta_2 \le 1$ and $\beta_3 > 1$. These thresholds may be adapted for different environments (e.g. indoor image, outdoor image or brightness of source light).

5.2 Working with HSV colour space

HSV colour space explicitly separates chromaticity and luminosity and has proven easier than RGB space to set a mathematical formulation for shadow detection [8] [9]. HSV space is more closely related to the human visual system than RGB and it is more sensitive to brightness changes due to shadows. For each pixel in F1, that initially has been segmented as foreground, we check if it is a shadow on the background according to the following consideration. If a shadow is cast on a background, the hue and saturation components change, but within a certain limit. The difference in saturation is an absolute difference, while the difference in hue is an angular difference.

$$\begin{cases} Shadow \;\; \beta_1 < V_{Ii}/V_{Bi} < \beta_2 \; and \left|H_{Ii} - H_{Bi}\right| < \tau_H \; and \left|S_{Ii} - S_{Bi}\right| < \tau_S \\ Highlight \quad V_{Ii}/V_{Bi} > \beta_3 \quad and \left|H_{Ii} - H_{Bi}\right| < \tau_H \; and \left|S_{Ii} - S_{Bi}\right| < \tau_S \end{cases} \tag{11}$$

with $0 < \beta_1, \beta_2, \tau_H, \tau_S < 1$ and $\beta_3 > 1$. Intuitively, this means that a shadow darkens a covered point, and a highlight brightens a covered point, but only within a certain range. Prati et al. [8] state that the shadow often has a lower saturation and, from our experimental results, we see that sometimes the shadow has a higher saturation than that of background sometimes. However, a shadow or highlight cast on a background does not change its hue and saturation as significantly as intensity.

5.3 Working with YCbCr and Lab colour spaces

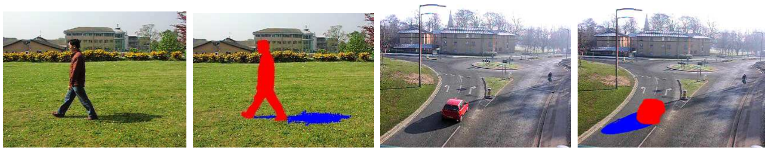
We now consider the luminance and chrominance (YCbCr) colour space to remove shadows from the results of background subtraction. If a shadow is cast on a background, the shadow darkens a point in the MDGKT. The luminance distortion is $\alpha_i = Y_{Ii}/Y_{Bi} < 1$, and chrominance components difference is $CH_i = \left(\left|C_{bi}^I - C_{bi}^B\right| + \left|C_{ri}^I - C_{ri}^B\right|\right)/2 < \beta_1$, where Y_{Ii}, C_{bi}^I, C_{ri}^I and Y_{Bi}, C_{bi}^B, C_{ri}^B are Y, Cb, Cr components in the current image and MDGKT respectively. A pixel in the F1 is classified as follows:

$$\begin{cases} Shadow \quad \alpha_i < 1 \ and \, CH_i < \beta_1 \\ Highlight \, \alpha_i > \beta_2 \, and \, CH_i < \beta_1 \end{cases} \tag{12}$$

where $\beta_1 < 1$ and $\beta_2 > 1$. There is a similar criterion for shadow removal in Lab space.

6 Quantitative evaluation

This section demonstrates the performance of the proposed algorithms above on several videos of both indoor and outdoor scenes, using an image size of 320×240. A quantitative comparison of two GMMs (ZHGMM and MDGKT) with different shadow removal methods is presented. A set of videos to test the algorithms was chosen and, in order to compute the evaluation metrics, the ground truth for each frame is necessary. We obtained this ground truth by segmenting the images with a manual classification of points as foreground, background and shadow regions. We prepared 41 ground truth frames in a 'walking people' sequence, and 26 in a 'moving car' sequence. Sample frames of each sequence and their ground truth mark-up are given in Fig.3. All shadow removal methods in five colour spaces using two GMMs have been fully implemented. Quantitative results for true positive rate (TPR) and specificity (SPC) metrics are reported in Table 1.



(a) Walking (b) Moving car

Fig. 3. Ground truth images: the red manual mark-up shows the foreground segmentation that we are interested in, the blue mark-up shows the shadow cast by the foreground

Table 1. Experimental quantitative results

	ZHGMM		MDGKT	
	TPR	SPC	TPR	SPC
RGB	0.8548	0.9838	0.9552	0.9853
Lab	0.7165	0.9828	0.8499	0.9846
YCbCr	0.6183	0.9811	0.6748	0.9811
Normalized RGB	0.6077	0.9628	0.6356	0.9714
HSV	0.5039	0.9671	0.6327	0.9712

Fig.4 shows sample frames 9 and 17 of the 'walking' video, 3 and 8 of the 'moving car' video. Each two-by-two block of images refers to the same frame in the original video. The top-left image is the original frame. The bottom-left image is the foreground segmentation (F1) results. In this image, all coloured pixels are the foreground segmentation output of the MDGKT algorithm, while the black pixels represent the modelled background. The coloured pixels are categorized as foreground object (coloured yellow), shadow (coloured green) or highlight (coloured red) by our shadow removal algorithm operating in RGB colour space. The shadow and highlight pixels are then removed and this is then followed by a post-processing binary morphology stage of dilatation and erosion to remove sparse noise. This gives the final foreground segmentation, as shown in the bottom right image of each two-by-two block. Finally, the top-right image in each block is a synthetic image, created by using the final foreground segmentation as a mask to extract the foreground object from the original frame, and superimposing this on the background model (mean value of each pixel). Clearly these synthetic images are largely shadow-free. The two videos in fig.4 are scenes with very strong shadows.

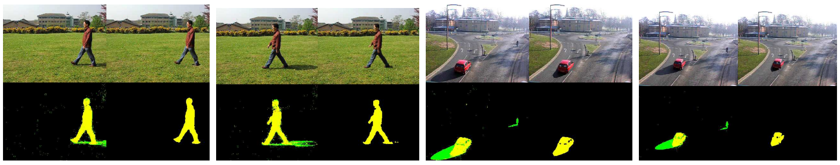


Fig. 4. Segmentation results with heavily shadowed input images.

7 Conclusions

Online learning of adaptive GMMs on nonstationary distributions is an important technique for moving object segmentation. This paper has presented an improvement to an existing adaptive Gaussian mixture model, using a multi-dimensional spatio-temporal Gaussian kernel smoothing transform for background modelling in moving object segmentation applications. The model update process can robustly deal with slow light changes (from clear to cloud or vice versa), blurred images, camera vibration in very strong wind, and difficult environmental conditions, such as rain and snow. The proposed solution has significantly enhanced segmentation results over a commonly used recursive GMM. We have given a comprehensive analysis of performance results in a wide range of environments and using a wide variety of colour space representations. The system has been successfully used to segment objects in both indoor and outdoor scenes, with strong shadows, light shadows, and highlight reflections and we have verified our system with rigorous evaluation. We have found that working in standard RGB colour space provides the best results.

Acknowledgements

The authors would like to acknowledge the provision of data sequence (the indoor scene images) from the Caviar project at the University of Edinburgh. Also funding under the UK Technology Strategy Board project, CLASSAC, and support from Cybula Ltd.

References

1. Zivkovic, Z., Heijden, F. van der, Recursive unsupervised learning of finite mixture models. IEEE Transaction on Pattern Analysis and Machine Intelligence. 26(5), (2004) 651-656
2. Friedman, N., Russell, S., Image segmentation in video sequences: a probabilistic approach. In: Proc 13th Conf on Uncertainty in Artificial Intelligence, (1997) 175-181
3. Stauffer, C., Grimson, W., Adaptive background mixture models for real-time tracking. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, (1999) 246-252
4. Power, P.W., Schoonees, J.A., Understanding background mixture models for foreground segmentation. In: Proceedings of Image and Vision Computing, New Zealand, (2002)
5. Stauffer, C., Grimson, W., Learning patterns of activity using real-time tracking. IEEE Transaction on Pattern Analysis and Machine Intelligence. 22(8), (2000) 747-757
6. KaewTraKulPong, P., Bowden, R., An improved adaptive background mixture model for real-time tracking with shadow detection. In: Proc. of 2nd European workshop on Advanced Video Based Surveillance Systems, chapter 11, (2001) 135-144
7. Lee, D-S., effective Gaussian mixture learning for video background subtraction. IEEE Transaction on Pattern Analysis and Machine Intelligence. 27(5), (2005) 827-832
8. Prati, A., Mikic, I., Trivedi, M.M., Cucchiara, R., Detecting moving shadows: algorithms and evaluation. IEEE Transaction on Pattern Analysis and Machine Intelligence. 25(7), (2003) 918-923
9. Cucchiara, R., Piccardi, M., Prati, A., Detecting moving objects, ghosts and shadows in video streams. IEEE Transaction on Pattern Analysis and Machine Intelligence. 25(10), (2003) 1337-1342
10. Zivkovic, Z., Heijden, F. van der, Efficient adaptive density estimation per image pixel for the task of background subtraction. Pattern Recognition Letters. 27(7), (2006) 773-780.
11. Chen, Z., Husz, Z.L., Wallace, I., Wallace, A.M., Video object tracking based on a Chamfer distance transform. In: IEEE International Conference on Image Processing, San Antonio, Texas, USA, (2007) 357-360
12. Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Transaction on Pattern Analysis and Machine Intelligence. 24(5), 603-619
13. Joachims, T., 2006. Training linear SVMs in linear time. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'06), Philadelphia, Pennsylvania, USA, 217-226
14. Elgammal, A.M., Harwood, A., Davis, L.S., Non-parametric model for background subtraction. Lecture Notes in Computer Science, vol. 1843, In: Proceedings of the 6th European Conference on Computer Vision-Part II, (2000) 751-767
15. Finlayson, G., Hordley, S., Drew, M., Removing shadows from images. In: European Conference on Computer Vision, Lecture Notes in Computer Science Vol. 2353, (2002) 4:823-836
16. Horprasert, T., Harwood, D., Davis, L.S., A statistical approach for real-time robust background subtraction and shadow detection. In: Proceedings of IEEE ICCV'99 Frame rate workshop (1999)
17. Elgammal, A., Duraiswami, R., Harwood, A., Davis, L.S., Background and foreground modelling using nonparametric kernel density estimation for visual surveillance. In: Proceedings of the IEEE, vol. 90, issue 7, (2002) 1151-1163