Proceedings of the 2006 IEEE/RSJ
International Conference on Intelligent Robots and Systems
October 9 - 15, 2006, Beijing, China

Robust Background Subtraction with Shadow and Highlight Removal for Indoor Surveillance

Jwu-Sheng Hu[1], Tzung-Min Su[2] and Shr-Chi Jeng[3]
Department of Electrical and Control Engineering
National Chiao-Tung University
Hsinchu, Taiwan, R.O.C.
[1]jshu@cn.nctu.edu.tw; [2]linux.ece89g@nctu.edu.tw; [3]u881403@oz.nthu.edu.tw

***Abstract* - This work describes a new 3D cone-shape illumination model (CSIM) and a robust background subtraction scheme involving shadow and highlight removal for indoor-environmental surveillance. Foreground objects can be precisely extracted for various post-processing procedures such as recognition. Gaussian mixture model (GMM) is applied to construct a color-based probabilistic background model (CBM) that contains the short-term color-based background model (STCBM) and the long-term color-based background model (LTCBM). STCBM and LTCBM are then proposed to build the gradient-based version of the probabilistic background model (GBM) and the CSIM. In the CSIM, a new dynamic cone-shape boundary in the RGB color space is proposed to distinguish pixels among shadow, highlight and foreground. Furthermore, CBM can be used to determine the threshold values of CSIM. A novel scheme combining the CBM, GBM and CSIM is proposed to determine the background. The effectiveness of the proposed method is demonstrated via experiments in a complex indoor environment.**

Index Terms* - *background subtraction, Gaussian mixture model, shadow removal, surveillance

I. Introduction

Background subtraction is an important technology that can be applied in numerous applications, including surveillance, robot vision, objected-based coding, image database, and video teleconferencing. A reference image is generally used to perform background subtraction. The simplest means of obtaining a reference image is by averaging a period of frames [1]. However, it is not suitable to apply time averaging on the applications for the indoor environment. A good background model must also handle the effects of illumination variation, and the variation from background and shadow detection. That is, a single model cannot represent the distribution of pixels with twinkling values.

Two approaches were generally adopted to solve the above problems. The first approach is termed the parametric method, and uses single Gaussian [2] or mixtures of Gaussian [3] to model the background image. The second approach is called the non-parametric method, and uses the kernel function to estimate the background density function [4].

Another important consideration is the shadows and highlights. Numerous recent studies have attempted to detect the shadows and highlights. Stockham [5] proposed that a pixel contains both an intensity value and a reflection factor and a decadent factor should be estimated to remove the shadow. Rosin [6] proposed that shadow is equivalent to a semi-transparent region, and uses two properties for shadow detection. Moreover, Levine [7] tried to convert the RGB color space to the rgb color space (chromaticity coordinate). However, lightness information is lost in the rgb color space. To overcome this problem, a lightness measure is used at each pixel [7]. However, the static thresholds are unsuitable for dynamic environment.

Indoor surveillance applications require solving environmental changes and shadow effects. Despite the existence of abundance of research on individual techniques, as described above, few efforts have been made to investigate the integration of environmental changes and shadow effects. The contribution of this work is the cone-shape illumination model (CSIM) and the scheme to combine the color-based background model (CBM), gradient-based background model (GBM) and CSIM. In CSIM, a new dynamic cone-shape boundary in the RGB color space is proposed for efficiently distinguishing a pixel from the foreground, shadow and highlight. A selection rule combined with the short-term color-based background model (STCBM) and long-term color-based background model (LTCBM) is also proposed to determine the parameters of GBM and CSIM. Fig. 1 illustrates the block diagram of the overall scheme.

The remainder of this paper is organized as follows. Section II describes the statistical learning method used in the probabilistic modeling and defines STCBM and LTCBM. Section III then proposes CSIM using STCBM and LTCBM to classify shadows and highlights efficiently. A hierarchical background subtraction framework that combined with color-based subtraction, gradient-based subtraction and shadow and highlight removal was then described to extract the real foreground of an image. Experimental results are presented to demonstrate the performance of the proposed method in complex indoor environments in section IV. Finally, Section V presents conclusions.

II. Background Modeling

Our previous investigation [8] studied a CBM to record the activity history of a pixel via GMM. However, the foreground regions generally suffer from rapid intensity changes and require a period of time to recover themselves

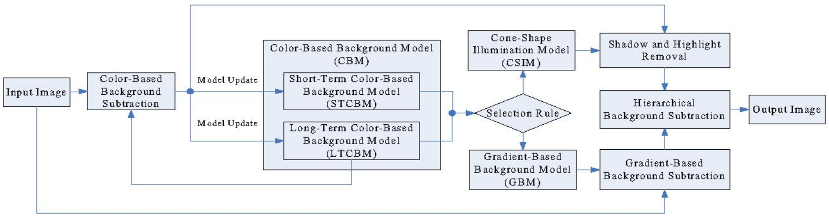

Fig. 1 Block diagram showing the proposed scheme for background subtraction with shadow removal.

when objects leave the background. In this work, STCBM and LTCBM are defined and applied to improve the flexibility of the gradient-based subtraction that proposed by Javed et.al [9]. The features of images used in this work include color and gradient. This study assumes that the density functions of the color and gradient information are both Gaussian distributed.

A. Color-based Background Modeling

First, each pixel x is defined as a 3-dimensional vector (R, G, B) at time t. N Gaussian distributions are used to construct the GMM, which is described as follows:

$$f(x|\lambda)=\sum_{i=1}^{N} w_i \frac{1}{\sqrt{(2\pi)^d |\Sigma_i|}} \exp(-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)) \quad (1)$$

where λ represents the parameters of GMM,

$$\lambda=\{w_i,\mu_i,\Sigma_i\}, i=1,2,...,N \; and \sum_{i=1}^{N} w_i=1 \quad (2)$$

The next step is calculating the GMM parameters λ so that the GMM can match the distribution of training feature vectors with minimal errors. A common method for calculating λ is the expectation maximization (EM) algorithm. Supposing a feature vector set $X=\{x_1,x_2,...,x_m\}$ is gathered from m image frames, then the GMM parameters λ can be obtained by iteratively using the E-step equation and M-step equation in the EM algorithm [8]. Moreover, this study uses the K-means algorithm before the EM algorithm iterations to accelerate the convergence.

B. Model Maintenance of LTCBM and STCBM

According to the above section, an initial color-based probabilistic background model is created using the training data with N Gaussian distributions. However, when the history of background changes is recorded over time, it is not flexible to model the background with exact N Gaussian distributions. Kaew et al [3] proposed a method of sorting the Gaussian distributions based on the fitness value w_i/σ_i ($\Sigma_i=\sigma_i^2 I$), and extracting a representative model with a threshold value B_0. To maintain the representative model and improve the flexibility of the background model simultaneously, LTCBM is defined with extra N new Gaussian distributions (LTCBM contains 2N distributions), an arrangement inspired by [3]. After sorting the first N Gaussian distributions with fitness value, b ($b \le N$) Gaussian distributions are extracted with the following criterion:

$$B=\arg\min_b \sum_{j=1}^{b} w_j > B_0 \quad (3)$$

The first b Gaussian distributions are defined as the ECBM to be the criterion to determine the background. Meanwhile, the remainders (2N-b) of the Gaussian distributions are defined as the CCBM for dealing with the background changes. Finally, LTCBM is defined as the combination of the ECBM and CCBM.

A new pixel value is considered as background when it belongs to any Gaussian distribution in ECBM and has a probability not exceeding 2.5 standard deviations away from the corresponding distribution. If none of the b Gaussian distributions match the new pixel value, a new test is conducted by checking the new pixel value against the Gaussian distributions in CCBM. The parameters of the Gaussian distributions are updated via the following equations:

$$\begin{aligned} w_i^{t+1} &= (1-\alpha)w_i^t + \alpha \hat{p}(w_i^t | X_i^{t+1}) \quad (4)\\ m_i^{t+1} &= (1-\rho)m_i^t + \rho X_i^{t+1} \\ \Sigma_i^{t+1} &= (1-\rho)\Sigma_i^t + \rho(X_i^{t+1}-m_i^{t+1})^T (X_i^{t+1}-m_i^{t+1}) \\ \rho &= \alpha g(X_i^{t+1} | m_i^t, \Sigma_i^t) \end{aligned}$$

ρ and α are termed the learning rates , and determine the update speed of LTCBM. Moreover, $\hat{p}(w_i^t | X_i^{t+1})$ results from background subtraction which is set to 1 if a new pixel value belongs to the i_{th} Gaussian distribution. If a pixel value does not belong to any of the Gaussian distributions in CCBM and the number of Gaussian components in CCBM is below (2N-b), a new Gaussian distribution is added with three parameters: the current pixel value as the mean, a large predefined value as the initial variance, and a low predefined value as the weight. Otherwise, the $(2N\text{-}b)_{th}$ Gaussian distribution in CCBM is replaced by the new one. After updating the parameters of the Gaussian components, all Gaussian distributions in CBM are resorted by recalculating the fitness values.

Unlike LTCBM, STCBM is defined in this study to record the background changes during a short period B_1. First, the corresponding Gaussian distribution of a new pixel is estimated by testing the new pixel with LTCBM. Given a pixel value set $PV=\{p_1,p_2,...,p_k,...,p_{B_1}\}$ collected during a period B_1, the corresponding Gaussian distribution set $CG=\{g_1,g_2,...,g_k,...,g_{B_1}\}$ is calculated by comparing the pixel value set with LTCBM. The histogram of CG is then given using the following equation:

$$H_{CG}(z)=\sum_k \delta(z-g_k)/B_1 \quad (5)$$

The order of the Gaussian distributions in LTCBM may change after re-sorting. A transfer flag set F_{CG} is defined for adjusting the bin order in $H_{CG}(z)$, as follows:

$$F_{CG} = \{F_j, j = 1,...,B, F_j \in \{-2N+1,...,0,1,...2N-1\}\} \quad (6)$$

After adjusting the order of the bins in H_{CG}, the transfer flag set F_{CG} is reset until the next update is performed. When the $(2N)_{th}$ Gaussian distribution is replaced with a new Gaussian distribution, g_j is set to 2N, F_j to 1 and $H_{CG}(2N)$ to zero. A threshold value B_2 is used to determine the short-term tendency of background changes. In this work, B_1 is assigned a value of 300 frames and B_2 is 0.8. If the value of $H_{CG}(k)$ exceeds 0.8, k is used for the representative background component in the STCBM, otherwise, STCBM provides no further information on background selection.

C. *Gradient-Based Background Modeling*

Javed et.al [9] developed a hierarchical approach that combines color and gradient information to solve the problem about rapid intensity changes. The k_{th} Gaussian component of LTCBM is chosen to obtain the gradient information. Javed et.al [9] proposed that the k_{th} Gaussian component is the highest weighted Gaussian distribution, which represents the background color at time t. The choice of k in [9] is similar to selecting k based only on ECBM and it leads to the loss of the short term tendencies of background changes. When a new Gaussian distribution is added into the background model, it is not selected owing to its low weighting value. Consequently, the accuracy of the gradient-based background model is reduced for that the gradient information is not suitable for representing the current gradient information.

To solve this problem, both STCBM and LTCBM are considered in selecting the value of k for developing a more robust gradient-based background model and maintaining the sensitivity to short-term changes. When STCBM can provide a representative background component (says the $k_s th$ bin in STCBM), k is set to k_s rather than the highest weighted Gaussian distribution. The gradient-based background model then can be defined as:

$$F^k(\Delta_m, \Delta_d) = \frac{\Delta_m}{2\prod \sigma_{f_x}^k \sigma_{f_y}^k \sqrt{1-\rho^2}} \exp\left(-\frac{z}{2(1-\rho^2)}\right) \quad (7)$$

where the details are described in [9].

III. Background Subtraction With Shadow Removal

This section describes shadows and highlights removal, and proposes a framework that combines CBM, GBM, and CSIM to improve background subtraction efficiency.

A. *Color-based Background Modeling*

Shadows and highlights are two important phenomenons that should be considered in most cases. The regions influenced by illumination changes are classified as the foreground if shadow and highlight removal is not performed after background subtraction.

Hoprasert [10] proposed a method of detecting highlight and shadow by gathering statistics from collected images. Brightness and chromaticity distortion are used with four threshold values to classify pixels into four classes. The method that used the mean value as the reference image in [10] is not suitable for dynamic background. Furthermore, the threshold values are estimated based on the histogram of brightness distortion and chromaticity distortion with a given detection rate, and are applied to all pixels regardless of the pixel values. Therefore, it is possible to classify the darker pixel value as shadow. Furthermore, it cannot record the background history.

This paper proposes a 3D cone model (Fig.2) and combines LTCBM and STCBM to solve the above problems. In the RGB space, a Gaussian distribution in the LTCBM becomes an ellipsoid whose center is the mean of the Gaussian component, and the length of each principle axis equals 2.5 standard deviations of the Gaussian component. A new pixel $I(R,G,B)$ is considered to belong to background if it is located inside the ellipsoid. The chromaticities of the pixels located outside the ellipsoid but inside the cone (formed by the ellipsoid and the origin) resemble the chromaticity of the background. The brightness difference is then applied to classify the pixel as either highlight or shadow.

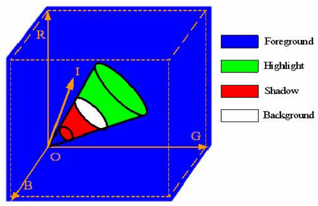

Fig. 2 The proposed 3D cone model in the RGB color space

The threshold values α_{low} and α_{high} are applied to avoid classifying the darker pixel value as shadow or the brighter value as highlight, and can be selected based on the standard deviation of the corresponding Gaussian distribution in CBM. Because the standard deviations of the R, G and B color axes are different, it is difficult to classify the pixel using the angles in the 3D space. Therefore, the 3D cone is projected onto the 2D space to classify a pixel using the slope and the point of tangency. Fig. 3 illustrates the projection onto the RG space.

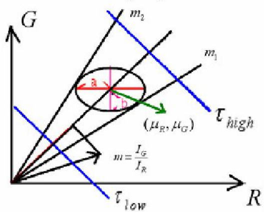

Fig. 3 2D projection of the 3D cone model from RGB space onto RG space

Let $a = 2.5 * \sigma_R$ and $b = 2.5 * \sigma_G$ denote the lengths of major and minor axis. The ellipse center is (μ_R, μ_G) , and the elliptical equation is described as the equation (8):

$$(R-\mu_R)^2/a^2+(G-\mu_G)^2/b^2=1 \quad (8)$$

The line $G=mR$ is assumed to be the tangent line of the ellipse. Equation (8) can then be solved using the line equation $G=mR$ with the following equation:

$$m_{1,2}=\frac{-(2\mu_R\mu_G)\pm\sqrt{(a^2-\mu_R^2)^2-4(2\mu_R\mu_G)(b^2-\mu_G^2)}}{2(a^2-\mu_R^2)} \quad (9)$$

A matching result set is given by $F_b=\{f_{bi},i=1,2,3\}$, where f_{bi} is the matching result of a specific 2D space. A pixel vector $I=[I_R,I_G,I_B]$ is then projected onto the 2D spaces of R-G, G-B, and B-R. The pixel matching result is set to 1 when the slope of the projected pixel vector is between m_1 and m_2. Meanwhile, if the background mean vector is $E=[\mu_R,\mu_G,\mu_B]$, the brightness distortion α_b can be calculated via the following equation

$$\alpha_b=\|I\|\cos(\left|\tan^{-1}(\frac{I_G}{\sqrt{I_R^2+I_B^2}})-\tan^{-1}(\frac{\mu_G}{\sqrt{\mu_R^2+\mu_B^2}})\right|)/\|E\| \quad (10)$$

The image pixel is classified as highlight, shadow or foreground using the matching result set F_b, the brightness distortion α_b and the following criteria:

$$\mathrm{C}(i)=\begin{cases}Shadow & :\sum F_b=3 \text{ and } \tau_{\text{low}}<\alpha_b<1 \quad, else\\ Highlight & :\sum F_b=3 \text{ and } 1<\alpha_b<\tau_{\text{high}} \quad, else\\ Foreground & : otherwise\end{cases} \quad (11)$$

When a pixel is a large standard deviation away from a Gaussian distribution, the Gaussian distribution probability of the pixel approximately equals to zero. It also means the pixel does not belong to the Gaussian distribution. By using the simple concept, τ_{high} and τ_{low} can be chosen using $\mathrm{N_G}$ standard deviation of the corresponding Gaussian distribution in CBM and are described as

$$\tau_{\text{high}}=1+\sqrt{(\mathrm{N_G}\cdot\sigma_R)^2+(\mathrm{N_G}\cdot\sigma_G)^2+(\mathrm{N_G}\cdot\sigma_B)^2}\cdot\cos\theta_\tau\cdot L_\mu \quad (12)$$

$$\tau_{\text{low}}=1-\sqrt{(\mathrm{N_G}\cdot\sigma_R)^2+(\mathrm{N_G}\cdot\sigma_G)^2+(\mathrm{N_G}\cdot\sigma_B)^2}\cdot\cos\theta_\tau\cdot L_\mu$$

where

$$\theta_\tau=|\theta_E-\theta_S|=\left|\tan^{-1}(\frac{\mu_G}{\sqrt{\mu_R^2+\mu_B^2}})-\tan^{-1}(\frac{\sigma_G}{\sqrt{\sigma_R^2+\sigma_B^2}})\right|$$

$$L_\mu=1/\sqrt{(\mu_R)^2+(\mu_G)^2+(\mu_B)^2}$$

B. Color-based Background Modeling

A hierarchical approach inspired by the method proposed by Javed et.al [9] was proposed to extract the foreground pixels. The difference between Javed et.al [9] and the proposed method is that a pixel classifying procedure using CSIM is applied before using the connected component algorithm to group all the foreground pixels in C(I). The robustness of background subtraction is enhanced due to the better accuracy in $|\partial R_a|$.

Given a new image frame I, the 7color-based backgound model is set to LTCBM and STCBM, and gradient-based model is $F^k(\Delta_m,\Delta_d)$. $C(I)$ is defined as the result of color-based background subtraction using CBM. $G(I)$ is defined as the result of gradient-based background subtraction. $C(I)$ and $G(I)$ can be extracted by testing every pixel of frame I using the LTCBM and $F^k(\Delta_m,\Delta_d)$. Moreover, $C(I)$ and $G(I)$ are both defined as a binary image, where 1 represents the foreground pixel and 0 represents the background pixel. The foreground pixels labeled in $C(I)$ are further classified as shadow, highlight and foreground by using the proposed 3D cone model. $C'(I)$ can then be obtained from $C(I)$ after transferring the the foreground pixels which have been labeled as shadow and highlight in $C(I)$ into the background pixel. Moreover, the foreground pixels can be extracted using the following equation

$$\sum_{(i,j)\in\partial R_a}(\nabla I(i,j)G(i,j))/|\partial R_a|\geq P_B \quad (13)$$

where ∇I denotes the edges of image I and ∂R_a represents the number of boundary pixels of region R_a.

IV. Experimental Results

The video data for experiments was obtained using a SONY DVI-D30 PTZ camera in an indoor environment. The same threshold values were used for all experiment $\mathrm{N_G}=15$, $\alpha=0.002$, $P_B=0.1$, $B_0=0.7$, $B_1=300$ and $B_2=0.8$. Meanwhile, the computational speed was around five frames per second on a P4 2.8GHz PC, while the video had a frame size of 320 x 240.

A. Experiments for Local Illumination Changes

The first experiment was performed to test the robustness of the proposed method about the local illumination changes. Local illumination changes resulting from desk lights occur constantly in indoor environments. A video clip containing several changes of desk light is collected to simulate local illumination changes. Fig.4(a) shows 15 representative samples of the video clip. Meanwhile, Fig.4(b) shows the classified result of the foreground pixel and Fig.4(c) displays the result of background subtraction to demonstrate the robustness of the proposed method. The desk light was turned on at the 476_{th} frame and its brightness increased until the 1000_{th} frame. The overall picture becomes the foreground regions of the corresponding frames in Fig.4(b) owing to the lack of such information in CBM. However, the final result of background subtraction of the corresponding frames in Fig.4(c) is still good owing to the proposed scheme combining CBM, CSIM and GBM. The desk light was then turned off at the 1030_{th} frame, and became darker until the 1300_{th} frame. The illumination changes are all modeled into LTCBM when the background model records the background changes. A new representative Gaussian distribution in ECBM is constructed owing to that new background information involved in the new collected frames between the 476_{th} and the 1000_{th} frame became the dominant. Consequently, the area of the red, blue and green regions reduces after the 1300_{th} frame.

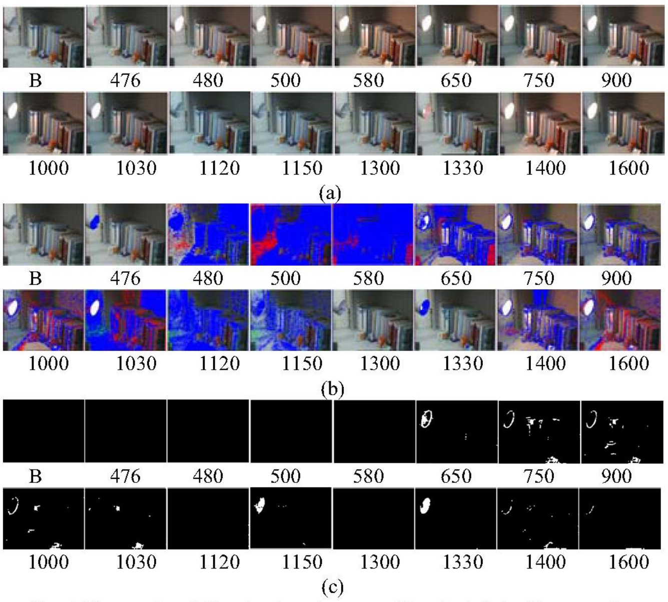

Fig. 4 The results of illumination changes with a desk light, B means the background image and the number below the picture is the index of frame, (a) original images, (b) The results of pixel classification, where red indicates the shadow, green indicates the highlight and blue indicates the foreground, (c) the results of background subtraction with shadow removal using the proposed method, where dark indicates the background and white indicates the foreground.

Table 1 compares the proposed scheme with the method proposed in [10]. Comparison criteria are identified by labeling the foreground regions of a frame manually. CSIM can be constructed based on the appropriate representative Gaussian distribution chosen from LTCBM and STCBM. The ability to handle illumination variation and the accuracy of the background subtraction are improved and the results are shown in Table 1.

B. Experiments for Global Illumination Changes

The second experiment was performed to test the robustness of the proposed method in terms of global illumination changes. The image sequences consist of illumination changes where a fluorescent lamp was turned on at the 381_{th} frame and more lamps were turned on at the 430th frame. The illumination changes are then modeled into LTCBM when the proposed background model recorded the background changes. Notably the area of the red, blue and green regions decreases at the 580_{th} frame. When the third daylight lamp is switched on in the 650_{th} frame, it is clear that less blue regions appear at the 845_{th} frame owing to illumination changes having been modeled in the LTCBM. However, the final results of background subtraction shown in Fig.5(c) are all better than those of pure color-based background subtraction shown in Fig. 5(b). Table 2 shows the comparison results between the proposed scheme and that proposed in [10]. The comparison demonstrates that the proposed scheme is robust to global illumination changes.

C. Experiments for Foreground Detection

In the third experiment, a person goes into the monitoring area, and the foreground region can be effectively extracted regardless of the influence of shadow and highlight in the indoor environment. Owing to the captured video clip having little illumination variation and dynamic background variation, the comparison of the recognition rate of final background subtraction between the proposed method and that of Hoprasert [10] reveals that both methods are about the same, as listed in Table 3.

D. Experiments for STCBM

The final experiment shows the advantage of adding STCBM. Initially, the doll is regarded as foreground at the 360_{th} frame, and at the 560_{th} frame, the foreground region becomes background owing to the LTCBM. Without adding STCBM, when a hand is placed above the doll at the 590_{th} frame, the foreground regions at the 670_{th} frame remain the same as those at the 590_{th} frame, as shown in Fig.7(b). This experiment demonstrates the efficiency of STCBM that a representative Gaussian component of CBM can be selected by giving consideration to long-term tendency and short-term tendency. Besides, the advantage of STCBM helps to reduce the computing time used in GBM and increase the recognition rate of foreground detection.

TABLE 1
THE ROBUSTNESS TEST BETWEEN THE PROPOSED METHOD AND THAT PROPOSED BY HOPRASERT[12] VIA LOCAL ILLUMINATION CHANGES WITH A YELLOW DESK LIGHT

Frame		476		480		500		580		650	
P*	Hoprasert*	100	94.1	99.8	36.4	99.9	22.5	99.9	15.4	83.9	23.4
Frame		750		900		1000		1030		1120	
P*	Hoprasert*	91.5	31.5	93.1	30.9	95.4	34.3	97.8	38.3	99.2	32.9
Frame		1150		1300		1330		1400		1600	
P*	Hoprasert*	93.8	50.7	99.9	99.8	93.3	92.4	96.2	13.0	99.3	34.7

*: The value in the table means the recognition rate that correct background pixels in a frame divide total pixels in a frame(%). P means the method proposed in this work and Hoprasert means the proposed used in [10].

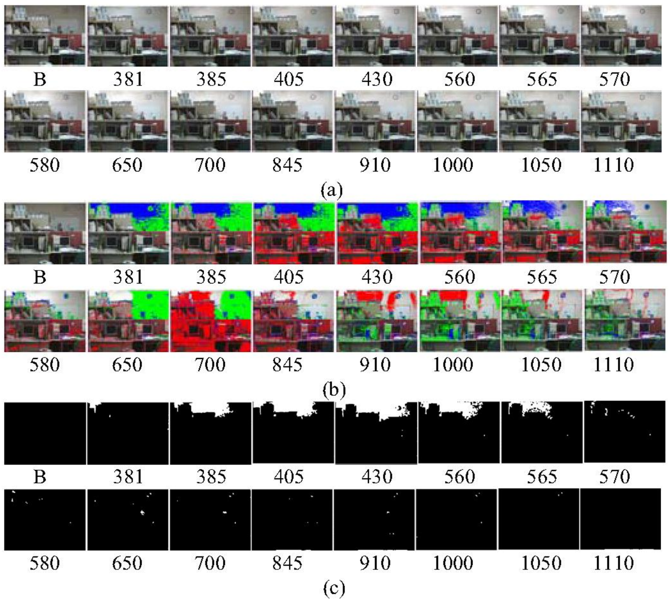

Fig. 5 The results of global illumination changes with fluorescent lamps, (a) original images, (b) The results of pixel classification, (c) The results of background subtraction with shadow removal using our proposed method

TABLE 2
THE COMPARISON BETWEEN THE PROPOSED METHOD AND THAT IN [12] VIA GLOBAL ILLUMINATION CHANGES WITH FLUORESCENT LAMPS

Frame	381 (1**)		385 (1**)		405 (1**)		430 (2**)		560 (2**)	
P* | Hoprasert*	98.2	93.5	88.4	82.1	83.9	78.2	56.5	68.4	66.9	69.8
Frame	565 (2**)		570 (2**)		580 (2**)		650 (3**)		700 (3**)	
P* | Hoprasert*	79.9	69.3	96.9	69.7	99.1	69.6	99.2	45.6	99.5	46.2
Frame	845 (3**)		910 (3**)		1000 (3**)		1050 (3**)		1110 (3**)	
P* | Hoprasert*	99.6	46.2	99.4	53.6	99.9	57.9	99.9	60.8	99.6	60.3

*: the same as * in Table 1

**: The number inside the parentheses indicates the number of fluorescent lamps that have turned on.

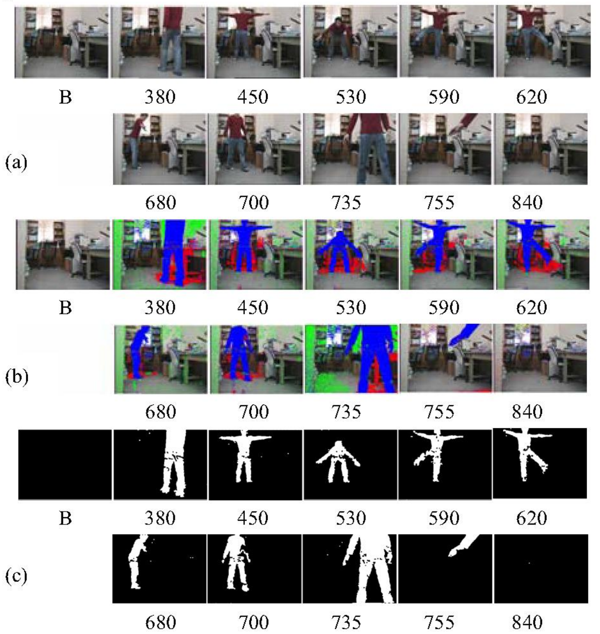

Fig. 6 The results of foreground detection, (a) original images, (b) The results of pixel classification, (c) The results of background subtraction with shadow removal using our proposed method.

TABLE 3
THE COMPARISON BETWEEN THE PROPOSED METHOD AND THAT PROPOSED BY HOPRASERT[12] VIA FOREGROUND DETECTION

Frame		380		450		530		590		620	
P*	Hoprasert*	90.5	89.2	86.5	85.8	89.4	88.9	88.5	87.7	88.7	88.8
Frame		680		700		735		755		840	
P*	Hoprasert*	91.1	90.6	85.6	85.2	82.8	80.7	92.4	92.5	100	99.6

*: the same as * in Table 1

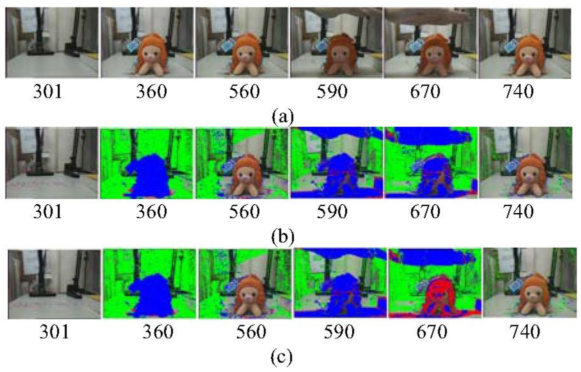

Fig. 7 The results of the advantage of STCBM, (a) original images, (b) The results of background subtraction without STCBM, (c) The results of background subtraction with STCBM.

V. CONCLUSIONS

This work addressed the problem of subtracting the background from an input image using three models, namely the color-based background model (CBM), gradient-based background model (GBM) and cone-shape illumination model (CSIM). In the CBM, the elected color-based background model (ECBM) and candidate color-based model (CCBM) are defined to increase the ability of recording a long period of background changes. The short-term color-based background model (STM) and long-term color-based background model (LTM) are defined to improve the flexibility and robustness of the gradient-based background subtraction. Most important, CSIM is proposed to extract the shadow and highlight in this paper with a 3D cone-shape boundary and combined with CBM in the RGB color space. The threshold values τ_{high} and τ_{low} of CSIM can be calculated automatically using the standard deviation of the Gaussian distribution selected using STCBM and LTCBM. The proposed 3D cone model is compared with the nonparametric model in a complex indoor environment. The experimental results show the effectiveness of the proposed scheme for background subtraction with shadow and highlight removal.

ACKNOWLEDGMENTS

This work was supported by National Science Council of the R.O.C. under grant no. NSC93－2218－E009064.

REFERENCES

[1] N. Friedman and S. Russell, "Image Segmentation in Video Sequences: A Probabilistic Approach," in *Proc. Thirteenth Conf. Uncertainty in Artificial Intelligence*. (UAI 97')

[2] C.R. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, pp. 780–785, July 1997.

[3] P. Kaew, Trakul, Pong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," in *Proc. 2nd Europcan Workshop on Advance Video Based Surveillance Systems*, AVBS01, Sept. 2001.

[4] A. Elgammal, R. Duraiswami, D. Harwood, and L.S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," in *Proc. of the IEEE*, vol. 90, July 2002, pp.1151-1163.

[5] Jr T.G. Stockham, "Image processing in the context of a visual model", in *Proc. of the IEEE*, 60: pp.828-842, July 1972.

[6] P.L. Rosin, and T. Ellis, "Image difference threshold strategies and shadow detection", *in Proc. of the sixth British Machine Vision Conference*, 1994.

[7] A. Elgammal, R. Duraiswami, D. Harwood, and L.S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," in *Proc. of the IEEE*, vol. 90, July 2002, pp.1151-1163.

[8] T. M. Su; J. S. Hu, "Background removal in vision servo system using Gaussian Mixture Model framework," *in Proc. of IEEE Conference on Networking, Sensing and Control*, pages: 3776 -3781, March 2004.

[9] Omar Javed , Khurram Shafique and Mubarak Shah, "A Hierarchical Approach to Robust Background Subtraction using Color and Gradient Information", *IEEE Workshop on Motion and Video Computing*, Orlando, Dec 2002.

[10] T. Hoprasert, D. Harwood, and L.S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection", *in Proc. of IEEE ICCV'99 FRAME-RATE Workshop*, 1999.