The following manuscript was published in

Shadow Elimination for Effective Moving Object Detection with Gaussian Models

Chia-Jung Chang, Wen-Fong Hu, Jun-Wei Hsieh, and Yung-Sheng Chen*
Department of Electrical Engineering, Yuan Ze University, Taiwan, ROC
* eeyschen@ee.yzu.edu.tw
Tel: 886-3-4638800 ext 409

Abstract

This paper presents a coarse-to-fine approach to eliminate unexpected shadows of multiple pedestrians from a static and textured background using Gaussian shadow modeling. At the coarse stage, a moment-based method is proposed to estimate the rough boundaries between shadows and moving objects. Then, at the fine stage, the rough approximation of shadow region provides a key to model shadows. The chosen shadow model is parameterized with several features including the orientation, mean, and center position of a shadow region. With these features, the chosen model can precisely eliminate the unexpected shadows from the scene background and thus improve the quality of further content analysis. Experiments demonstrate approximately 95% ratio of pedestrian-related shadows can be successfully eliminated from the scene background.

1. Introduction

Tracking multiple moving objects is an important problem in many applications such as video surveillance, video retrieval, teleconference, and so on. In the past, there have been researchers investigating methods for tracking moving objects in real time to realize these applications [1]-[3]. For example, Wren et al. [2] proposed a “Pfinder” tracking system for tracking people and observing their behaviors under different conditions. Haritaoglu et al. [3] proposed a real-time visual surveillance system “W4” (Who, When, Where, What) to monitor multiple persons and analyze their activities. All these approaches assume there is no shadow appearing in the acquired video sequences. If the video sequence contains any shadows, the shadows will cause consequent analysis works, e.g., counting the number of objects, estimating the locations of objects, recognizing the objects, etc., to fail.

In order to solve the problems caused by shadows, Tao et al. [1] proposed a pyramid model and a fuzzy neural network approach to eliminate the shadows. Besides, Onoguchi [4] used two cameras to eliminate the shadows of pedestrian-like moving objects based on the information of object heights. Moreover, Ivanov et al. [5] proposed a method for removing occlusion shadows based on a disparity model that is invariant to arbitrarily rapid changes in illumination. However, in order to overcome rapid changes in illumination, at least three cameras are required. Therefore, a simple, reliable and accurate method is urgent to be available for eliminating object shadows in a surveillance system.

In this paper, a novel approach is proposed for detecting moving-object shadows from a static background scene using Gaussian shadow modeling. First, through background subtraction, the regions of moving objects (including shadows) can be extracted. In practice, each detected moving region may contain several pedestrians and shadows. In order to precisely remove the unwanted shadows, this paper presents a histogram projection method to separate each pedestrian from the moving region first. Then, a coarse-to-fine approach is applied for detecting the boundaries between the pedestrian and its shadow. At the coarse stage, a moment-based method is applied to estimate the orientation of the detected pedestrian. According to the orientation and silhouette features of the detected regions, a rough approximation of the exact shadow area can be detected. At the fine stage, the rough approximation of shadow region is further refined through Gaussian shadow modeling. The major difficulty in shadow modeling is the choice of a proper model that can reflect various appearances of shadows at different orientations and lighting. This paper presents several shadow models to demonstrate their performance when modeling shadows for choosing the best shadow model. The chosen shadow model is parameterized by several features including the illumination properties, position, and orientation of a shadow. The representation provides precise information for separating the moving objects from the background. Due to the simplicity of the proposed method, all the shadows can be removed in real-time. Experimental results demonstrated approximately 95% ratio of pedestrian-related shadows can be successfully eliminated from the scene background.

2. Shadow Elimination with Gaussian Models

In this paper, a simple and effective method algorithm is proposed for eliminating moving pedestrian-like shadows using only one camera. In this approach, each moving region is first extracted from the static background using background subtraction. Then, a coarse-to-fine approach is proposed for detecting the correct boundaries between the detected pedestrian and its shad-
ows. In order to illustrate the proposed method more clearly, the paper assumes the extracted moving region has only one pedestrian first and describes a method to tackle this problem in Section 2. Then, more complicated cases when a moving region has several pedestrians and shadows are discussed in Section 3.

2.1. Object Moments and Object Orientation

![Figure 1](image)

**Fig. 1** Object orientation and its contour information.

After image subtraction and thresholding, the moving objects can be detected like Fig. 1 (a). In order to separate the shadow from an object, the orientation of the object should be estimated. In this paper, the orientation of an object is estimated from the properties of object moments. Given a binary region $R$, the central moments of $R$ can be defined as follows:

$$m_{pq} = \sum_{(x,y) \in R} (x - \bar{x})^p (y - \bar{y})^q,$$

where $(\bar{x}, \bar{y})$ is the center of $R$. Then, the orientation $\theta_R$ of $R$ can be estimated as follows:

$$\theta_R = 0.5 \tan^{-1} \left( \frac{2m_{11}}{2m_{02} - m_{20}} \right).$$

2.2. Shadow Detection

After calculating the orientation of the detected object, the following stage is to detect the boundary line for roughly cutting the shadows from the background. Like Fig. 1 (b), the shadow region $R_2$ expects to be cut from the object region $R_1$ by the line $P_{\theta_R}$. The point $P_k$ is the one that has the maximum vertical difference between two adjacent points along the silhouette of the region $R$. The silhouette curve $C_y(x)$ of $R$ can be obtained by tracking the vertical position of the first pixel of the region $R$ when we track all the pixels along the $x$th column from top to bottom. Then, the coordinate $(x_k, y_k)$ of the point $P_k$ can be obtained as follows:

$$x_k = \arg \max_x \left| C_y(x) - C_y(x-1) \right|, \quad y_k = C_y(x_k).$$

Let $\theta_g$ denote the orientation of the object region $R$. The line $P_{\theta_g}$ can be defined as the one that passes through $P_k$ with the orientation $\theta_g$. Let $P_l(x, y)$ denote as the top point of the object $R$. All the pixels $(x, y)$ in the shadow $R_2$ satisfy the following condition:

$$f(x, y) = (y - mx - c)(y - mx - c) < 0 \quad (4)$$

where $m = \tan \theta_g$ and $c = y_k - x_k \tan \theta_g$. Based on Eq. (4), the pixels in the shadow $R_2$ can be identified from the object $R$. However, the line $P_{\theta_R}$ cannot detect all the shadow pixels from the object $R$ due to the irregular shadow contour. The goal of this stage is to find a rough approximation of the shadow region as a base for modeling shadows.

2.3. Gaussian Shadow Modeling

Once the shadow region $R_2$ has been determined, the following is to model the shadow pixels. The first straight-forward model to model shadows is a Gaussian model defined as follows:

$$G_i(x, y) = \exp \left( - \frac{(I(x, y) - \mu_i)^2}{\sigma_i^2} \right), \quad (5)$$

where $\mu_i$ and $\sigma_i$ are the mean and variance of $R_i$, respectively, and $I(x, y)$ is the intensity of a pixel $(x, y)$ in $R_i$. The disadvantage of this model is only the gray intensity is parameterized in this model. Other non-shadow region far from the shadow region but with similar mean intensities will be misclassified as shadows based on this model. In order to overcome this problem, a new model is defined as

$$G_g(x, y) = e^{-\frac{((x-x_i)^2 + (y-y_i)^2 + (I(x, y)-\mu_i)^2)}{\sigma_i^2}},$$

where $\mu_i$, $\sigma_i$, and $\sigma_i$ are the mean and variance of the $x$ and $y$ coordinates of $R_i$, respectively. This model performs better than the model $G_i$ since the shadow positions are included into modeling. However, in this model, small pieces of non-shadow regions near to shadow boundaries still will be misclassified since the orientation of the shadow is not considered into modeling. In order to model shadows more accurately, we map the original coordinates into an elliptic coordinates as $(s, t) = Rotation_{\theta_k} \times (x - \mu_x, y - \mu_y)$, where $\theta_k$ is the major orientation of $R_2$. Then, the final Gaussian object model used in this paper is:

$$G_g(s, t) = e^{-\frac{w_x^2}{\sigma_x^2} \frac{w_y^2}{\sigma_y^2} \frac{(s(s-1)) - \mu_x}{\sigma_x^2} \frac{(t(t-1)) - \mu_y}{\sigma_y^2}}, \quad (6)$$

where $w_x$ is the weight for $s$ coordinate, $w_y$ the weight for $t$ coordinate, $\sigma_x$ the weight for intensity component, $\sigma_y$ the variance of $s$, and $\sigma_z$ the variance of $t$. Here, $w_x$, $w_y$, and $w_z$ are set as 0.2, 0.3, and 0.5, respectively. Once the correct Gaussian model has been chosen, for
each pixel in $R$ if its value in the Gaussian model $G_i$ is larger than a threshold $T_i$, the pixel is classified as a shadow pixel. Here, $T_i$ is the average of $G_i$ for all the pixels in $R$. 

### 3. Multiple-Shadow Elimination

If a moving region contains more than one pedestrian as shown in Fig. 2, their shadows may not be well eliminated by the preceding method. In order to efficiently remove the unwanted shadows from each pedestrian, the following is a histogram-based method for segmenting each pedestrian from the background. Then, the method to eliminate multiple shadows is introduced.

![Shadows caused by multiple pedestrians and their vertical intensity and edge histogram.](image)

Fig. 2 Shadows caused by multiple pedestrians and their vertical intensity and edge histogram.

Let $R_k$ denote the $k$th extracted moving region. In this paper, the vertical intensity and edge projection histograms are used to separate each pedestrian from $R_k$. Let $I_k(x,y)$ denote the intensity of pixel $(x,y)$ in $R_k$. Then, the vertical intensity histogram of $R_k$ along the $y$ direction is defined as

$$SH_k(x) = \frac{1}{2l+1} \sum_{t=-l}^{l} I_k(x+t,y), \quad (7)$$

where $l$ is the smoothing factor and set as 4 in this paper. According to the definition, the vertical intensity histogram of Fig. 2 (a) can be shown in Fig. 2 (b). It is clear the histogram has two major peaks. Each peak corresponds to one head of pedestrians. Therefore, through detecting the peaks of $SH_k(x)$, each of head position can be obtained.

Once the head positions of each pedestrian have been obtained, the following is a method to determine the body boundaries of each pedestrian from the so-called vertical edge histogram. The vertical edge histogram $EH_k(x)$ of $R_k$ is defined as follows:

$$EH_k(x) = \sum_y \sum_{i=-d}^{d} |I_k(x+i,y) - I_k(x-i,y)|, \quad (8)$$

where $d$ is set to be 4 for smoothing consideration. Let $h_k(i)$ denote the head position of the $i$th pedestrian in $R_k$ obtained by searching peaks of $SH_k(x)$ and $N_k$ be the number of heads in $R_k$. Based on $EH_k(x)$, the body boundaries of each person in $R_k$ can be obtained with the following algorithm:

### Pedestrian Boundary Detection (PBD) Algorithm

**Step 1:** Find the maximum value $M_{EH_k}$ from $EH_k(x)$.

**Step 2:** Set the threshold $T_{EH_k}$ as $M_{EH_k}/3$ for determining a pulse function $P_{EH_k}(x)$.

**Step 3:** With $T_{EH_k}$, obtain the pulse function $P_{EH_k}(x)$ as:

$$P_{EH_k}(x) = \begin{cases} 
1, & \text{if } EH_k(x) > EH_k(x-1) \\
0, & \text{otherwise}
\end{cases} \quad \text{and } EH_k(x) \geq T_{EH_k}.$$  

**Step 4:** According to $P_{EH_k}(x)$, if a sequence of continuous pulses forms a segment, detect all the segments if their lengths are less than 3.

**Step 5:** Record the most right position of the $m$th remained segment as $X_{m_{right}}(m)$.

**Step 6:** Let $B_k(i)$ denote the boundaries of each pedestrian body in $R_k$. Then, $B_k(i)$ can be obtained as $B_k(i) = X_{m_{right}}(m)$ and $B_k(i) = X_{m_{left}}(m)$, where $m = \arg \min_{j} h_{k}(j) - X_{m_{left}}(m)$ if $h_{k}(j) - X_{m_{left}}(m)$ is the one of the most right pedestrian. Then, the orientation $O_k$ of shadows can be estimated as follows:

$$O_k = \begin{cases} 
\text{left}, & \text{if } \left| \left| \left| \left| \left( \left( \left( \left( h_{k}(i) - X_{left}(m) \right) \right) \right) \right) \right) \right) \right) \right) \\
\text{right}, & \text{otherwise}
\end{cases} \quad (9)$$

where $left$ and $right$ are the most left and right boundaries of $R_k$, respectively. Based on the above discussions, the whole procedure to eliminate all the unwanted shadows from a moving region $R_k$ can be summarized in details as follows:

### Multiple-Shadow Elimination Algorithm

**Step 1:** According to the PBD algorithm, obtain the boundaries $\{B_k(i)\}_{i=1..N_k}$ of each pedestrian in $R_k$. Then, separate a set of pedestrian regions $\{A_k(i)\}_{i=1..N_k}$ from $R_k$. If the orientation of shadows is “right,” $A_k(i)$ can be obtained as follows:

$$A_k(i) = \{ p(x,y) | p(x,y) \in R_k, B_k(i-1) \leq x < B_k(i+1) \}.$$  

Otherwise, $A_k(i)$ is obtained as follows:

$$A_k(i) = \{ p(x,y) | p(x,y) \in R_k, B_k(i-1) \leq x < B_k(i) \}. $$
Step 2: Determine a region $\overline{A}$, whose shadow is the closest to the boundaries of $R$. That is, if the orientation is “right,” $\overline{A} = A_f(N_f)$. Otherwise, $\overline{A} = A_f(1)$.

Step 3: Calculate the orientation $\theta$ of the region $\overline{A}$ according to Eq. (2);

Step 4: For each region $A_i$, repeat the following steps:

4.1 Let $\phi = \theta$. Find the separating line based on Eq. (3);

4.2 Separate the region $A_i$ into shadow region $S_i$ and non-shadow region $\hat{A}_i$ according to the line $\frac{P_{A_i}Q_{A_i}}{Q_{A_i}}$;

4.3 Build the shadow model $G_{S_i}$ according to Eq. (6) and $S_i$;

4.4 Eliminate each pixel in $R$ if the pixel satisfies $G_{S_i}(x,y) > T_{S_i}(x,y)$, where $T_{S_i}$ is the average of $G_{S_i}$ for all pixels in $S_i$.

4. Experimental Results

In order to analyze the robustness and effectiveness of the proposed method, four experiments under different conditions are demonstrated in this paper. The first experiment is to show the capability of the proposed algorithm to eliminate the shadows from single pedestrian with different orientations. Fig. 3 shows the case when the shadow is cast on the right side of the moving pedestrian. Fig. 3 (b) is the extracted moving object from (a). (c) is the result using the shadow model $G_i$. In this model, though most of shadow pixels have been removed, many non-shadow pixels are also eliminated. (d) is the result using the model $G_2$. Clearly, the result of (d) is better than (c). However, in (d) a few of mistakes are still unsolved. (e) is the result using the suggested model with weights $w_t = 0.2$, $w_s = 0.3$, and $w_f = 0.5$. It is clear the result of (e) is much better than (c) and (d). Fig. 4 shows another case when the shadow is cast on the left side of the moving pedestrian. The suggested model also works better than other models. Based on the results of Fig. 3 and Fig. 4, the proposed method indeed works robustly regardless of changes in shadow orientation. In order to prove the capability of the proposed method to eliminate multiple shadows, more experiments are needed. Fig. 5 shows two pedestrians walk along a road and the left one’s shadow connects to the other pedestrian. (a) is one shot of the video sequence. (b) is the detected moving object through background subtraction and a labelling technique. (c) is the result of shadow elimination with the suggested model. All the shadows are clearly removed with the proposed method. Fig. 6 shows another case when three pedestrians walk along a road. Even though more shadows appear in the video sequence, the proposed method still works well. From these experiments, the superiority of the proposed method can be clearly verified.

![Fig. 3 Results of single shadow elimination.](image)

![Fig. 4 Results of single shadow elimination when the shadow orientation is left.](image)

![Fig. 5 Results of shadow elimination when two moving pedestrians appear in the video sequence.](image)

![Fig. 6 Results of shadow elimination when three moving pedestrians appear in the video sequence.](image)

References


