Shadow Detection by Integrating Multiple Features

Kuo-Hua Lo and Mau-Tsuen Yang
National Dong Hwa University, Taiwan
mtyang@mail.ndhu.edu.tw

Abstract

Cast shadows of moving foreground objects in a scene often result in problems for many applications such as surveillance, object tracking/recognition, video content analysis and intelligent transportation systems. In this paper we presented an algorithm exploiting information of color, shading, texture, neighborhoods and temporal consistency to detect shadows in a scene efficiently and reliably. The experimental results showed that the proposed method can detect umbra as well as penumbra in different kinds of scenarios under various illumination conditions.

1. Introduction

Motion analysis in video sequence is important in many applications such as surveillance, obstacle tracking/recognition, video content analysis and Intelligent Transportation Systems (ITS). However, it is very common that the cast shadows on the background could be classified as foreground objects by mistake. The performance of the successive analysis, recognition or tracking would be seriously degraded due to this problem. A reliable and efficient shadow removal algorithm is required before potential power of these vision-based sensing applications can be realized.

For this purpose, we proposed a shadow detection/removal method considering color, shading, texture, neighborhoods and temporal consistency in the scene. There are two types of cast shadows in a scene: moving and static shadow. We are only interested in the moving shadow since static shadow can be modeled as a part of the background. The remaining parts of this paper are organized as follows: Section 2 outlines and compares the related works. Section 3 shows the assumed luminance model. Section 4 describes the basic ideas of the proposed method. Section 5 compares Toth’s method [2] with the proposed method especially on penumbra regions. Section 6 demonstrates the experimental results. Section 7 concludes this paper.

2. Related Works

Several shadow detection algorithms have been proposed for traffic surveillance [8]. Generally speaking, shadow regions are detected based on information of the luminance, chrominance and gradient density. Large number of false alarms or miss detections can be reduced by the assumption of known geometry information of the foreground vehicles [4] or the lane lines [9]. As a result, these methods are only suitable for the road traffic applications.

It is possible to detect shadow by examining the color information of each pixel. Siala [1] presented a moving shadow detection algorithm by training the shadow samples in RGB color space with Support Vector Domain Description (SVDD). Cucchiara [6] considered the color independence property in the HSV color space to detect shadow. It is observed that if a pixel is covered by shadow, the hue and saturation components of the pixel only change within a certain limit. However, the hue components on pixels with saturated or poor illumination are usually unstable.

Finlayson [10] formed an intrinsic image based on the parallel trends in the logarithm \( \{ G/R, B/R \} \) color space. The method can remove shading in umbra regions, but not work in penumbra region. A few shadow detection algorithms used monocular images as inputs. Stauder [3] relied on brightness, edge and shading information to detect moving cast shadows in a textured background. Dong [7] assumed that shadow often appears around foreground object and tried to detect shadow by extracting moving edges. Without considering color information, problems could occur when both the foreground and background are uniform regions without texture.

Toth [2] proposed a shadow detection algorithm based on color and shading information. It is observed that, for every pixel in a small neighborhood in the umbra of a shadow region, the intensity values with shadow divided by those without shadow should be a constant. This shading property was used to detect umbra region. A successive morphological filter was applied to remove the penumbra region. The edge/texture information was not considered for the discrimination of shadow regions.

Funt [5] presented a method for object recognition under changeful illumination by checking the equality of intensity ratios of neighboring pixels on current image and background image. By extending this color constancy concept, we designed an umbra and penumbra detection method that exploited the information of color, shading, texture, neighborhoods and temporal consistency in a scene.

3. The Luminance Model

Suppose \( I(x,y) \) is the intensity value of the pixel located at \( (x,y) \), \( E(x,y) \) is the irradiance of the 3D point
projecting to \((x,y)\) and \(\rho(x,y)\) is the diffuse reflectance of the same 3D point. A simple luminance model assuming Lambertian reflectance can be defined as follows: 
\[
I(x,y) = E(x,y)\rho(x,y)
\]

Two kinds of shadow can appear in an image: the penumbra and the umbra. The difference between them can be modeled by the following equation:
\[
E(x,y) = \begin{cases} 
C_a + C_p \cdot \cos \theta & \text{no shadow} \\
C_a + k(x,y) \cdot C_p \cdot \cos \theta & \text{penumbra} \\
C_a & \text{umbra}
\end{cases}
\]

where \(C_a\) is the radiance of ambient light, \(C_p\) is the radiance of a distant light source and \(\theta\) is the angle between the direction of the distant light source and the surface normal vector of the 3D point projecting to \((x,y)\). The weighting factor \(k(x,y)\) represents the percentage of the receiving energy when the distant light source is partially occluded (penumbra). The value of \(k(x,y)\) ranges from 0 (umbra) to 1 (no shadow).

4. The Proposed Algorithm

As a preprocessing step, a statistical background subtraction was applied to generate the foreground mask region (FMR). A noise removal algorithm was performed to refine the FMR. The regions outside the FMR were used to dynamically update the background image. Then the proposed shadow detection algorithm was applied to the FMR by considering three factors: the color constancy between pixels, the color constancy within pixel and the temporal consistency between frames. These factors are described in details in the following subsections.

4.1 Color constancy between pixels

Assuming the 3D points projecting to neighboring pixels receive the same irradiance, i.e., \(E(x,y) = E(x+1,y)\), the ratio of the intensity value of a pixel \((x,y)\) to the intensity value of its neighboring pixel \((x+1,y)\) on the current image \(I\) can be calculated as follows:
\[
\frac{I(x,y)}{I(x+1,y)} = \frac{E(x,y)\rho(x,y)}{E(x+1,y)\rho(x+1,y)} = \frac{\rho(x,y)}{\rho(x+1,y)}
\]

Similarly, the ratio of the intensities of the neighboring pixels on the background image \(I'\) can be calculated as follows:
\[
\frac{I'(x,y)}{I'(x+1,y)} = \frac{E'(x,y)\rho'(x,y)}{E'(x+1,y)\rho'(x+1,y)} = \frac{\rho'(x,y)}{\rho'(x+1,y)}
\]

If a pixel locates in a shadow region in the current image, then the pixel in both current and background image are projected by the same 3D point, i.e., \(\rho(x,y) = \rho'(x,y)\) and \(\rho(x+1,y) = \rho'(x+1,y)\). As a result, the ratio of the intensities between neighboring shadow pixels in both current and background image should be the same:
\[
\frac{I(x,y)}{I(x+1,y)} = \frac{I'(x,y)}{I'(x+1,y)}, \text{if } (x,y) \text{ is in shadow region}
\]

We called this shading property as color constancy between pixels. To speedup the examination of this property, two logarithm ratio maps, \(d(x,y)\) for current image and \(d'(x,y)\) for background image, can be computed by convolving the logarithm image with a horizontal first-order derivative mask:
\[
\begin{align*}
    d(x,y) &= \ln \frac{I(x,y)}{I(x+1,y)} = \ln I(x,y) - \ln I(x+1,y) \\
    d'(x,y) &= \ln \frac{I'(x,y)}{I'(x+1,y)} = \ln I'(x,y) - \ln I'(x+1,y)
\end{align*}
\]

Basically, a ratio map keeps the texture information that should not be affected by cast shadow in an image. Based on this idea, a pixel is classified as shadow only if its value in the ratio map (texture information) is similar to that in the background. A simple pixel-wise comparison between \(d(x,y)\) and \(d'(x,y)\) can be used to determine whether a pixel belongs to shadow regions or not. Nevertheless, there could be some outliers due to noise or coincidence. To address this problem, spatial consistency is exploited to remove outliers. It is observed that shadows usually occupy a region instead of a few isolated pixels. The error score for discriminating the pixel \((x,y)\) as shadow can be calculated by summing the difference of \(d\) and \(d'\) over all pixels in a small neighborhood window \(\omega\) centered at \((x,y)\): 
\[
D(x,y) = \sum_{(i,j) \in \omega} |d(i,j) - d'(i,j)|
\]

Furthermore, there are three color channels \(R, G\) and \(B\) if the inputs are color images. The error score \(D_i(x,y)\) of each color channel \(i\) can be processed individually. The score of errors combining three color channels can be obtained as follows:
\[
\Psi(x,y) = \sum_{\omega \in \mathbb{R}, G, B} D_i(x,y)
\]

By considering the neighborhood and combining color channels, the overall error score \(\Psi(x,y)\) for shadow discrimination is much more robust and stable.

4.2 Color constancy within pixel

A shadow casting on a background pixel changes its brightness but not changes much its color. Comparing pixel-wise color information between current image and background image can help to detect cast shadow. When a textureless foreground object is in front of a textureless background, the most important clue for separating cast shadow from foreground object is the color information. The brightness and color information of a pixel can be divided by transferring the color space from RGB to normalized \(r-g\) using the following equations:
\[
\begin{align*}
    C_r(x,y) &= \ln \frac{I_g(x,y)}{I_g(x,y) + I_B(x,y) + I_g(x,y)} \\
    C_g(x,y) &= \ln \frac{I_B(x,y)}{I_g(x,y) + I_B(x,y) + I_g(x,y)}
\end{align*}
\]
Since the value of $C_r$ and $C_g$ remain roughly the same under different illumination condition. The score of error for discriminating the pixel $(x,y)$ as shadow is defined as:
\[
\Theta(x,y) = |C_r(x,y) - C'_r(x,y)| + |C_g(x,y) - C'_g(x,y)|
\] (10)
where $C_r$ and $C_g$ are the color information of the current image. Similarly, $C'_r$ and $C'_g$ are the color information of the background image. A smaller $\Theta(x,y)$ represents that the color of the pixel $(x,y)$ does not change much and it is more likely to be a shadow pixel.

### 4.3 Temporal consistency between frames

Methods considering only the first factor (color constancy between pixels) can not distinguish between a uniform foreground and its shadow on a uniform background with no texture. Methods considering only the second factor (color constancy within pixel) tend to wrongly classify a foreground region with similar color as its background to be a shadow. Assuming the foreground object moves slowly, temporal consistency between frames can provide a clue of potential shadow regions. In the words, a shadow pixel tends to remain in shadow region in next time frame, and vice versa. Exploiting temporal consistency can prevent wrongly classifying the temporally isolated noise as shadow regions.

A reliable shadow detection system should be able to consider all these factors simultaneously. In our system, the error scores corresponding to these factors are fused together using the following recursive linear equation:
\[
T_r(x,y) = a \cdot \Psi_r(x,y) + b \cdot \Theta_r(x,y) + (1 - a - b) \cdot T_r(x,y)
\] (11)
where $a$ and $b$ are weightings that control the importance of each factor and the speed of the recursive update. Their values are determined empirically in our current experiments and should be adjusted dynamically for better adaptability in the future. For example, $a$ should be lowered for images without much texture; $b$ should be lowered for images under saturated or poor illumination; $a + b$ should be raised for images with fast moving objects. $T_r(x,y)$ represents the total score of error for discriminating $(x,y)$ as shadow at time instant $t$. Finally, a thresholding operation is applied on $T_r(x,y)$ to determine whether the pixel $(x,y)$ belongs to foreground object or cast shadow region.

### 5. Comparison of Toth’s method with the proposed method

In this section, the shading property used by Toth’s method is explained for comparison purpose. Suppose $I$ is the current image and $I'$ is the background image. A foreground-to-background ratio $K$ for each pixel $(x,y)$ is defined as:
\[
K(x,y) = \frac{I(x,y) - \bar{I}(x,y)}{\bar{I}(x,y)} = \frac{E(x,y) - \bar{E}(x,y)}{\bar{E}(x,y)}
\] (12)

For every pixel $(x,y)$ in a small neighborhood in the umbra region, it is assumed that $\bar{E}'(x,y) = C_e$ is a constant and $E(x,y) = C_e + C_p \cdot \cos \theta$ is another constant.

As a result, the function $K(x,y)$ should be a constant in a small umbra region. Hence, Toth defined a score of error for discriminating pixel $(x,y)$ as shadow by the following equation:
\[
\Omega(x,y) = \frac{\sum_{i,j} I'(i,j) - I(i,j)}{\hat{K}(x,y)}
\] (13)
where $\hat{K}(x,y)$ represents the overall foreground-to-background ratio of every pixel $(i,j)$ inside a small window $\omega$ centered at $(x,y)$. A thresholding operation is performed on $\Omega(x,y)$ to extract pixels in umbra. Then, a morphological filter is applied to remove the remaining penumbra region.

In umbrella regions, the proposed method shares similar conception with the Toth’s method. The error score $d$ in our method and $K$ in Toth’s method can be related as follows:
\[
d(x,y) - d'(x,y) = \ln \frac{I(x,y)}{I(x+1,y)} = \ln \frac{1}{\hat{K}(x,y)} = \ln \frac{K(x,y)}{K(x+1,y)}
\] (14)

According to the above equation, $d'(x,y) - d(x,y)$ becomes closer to zero when $K(x,y)$ gets closer to $K(x+1,y)$. However, in a penumbra region where the intensity values of pixels changes gradually, the error score $\Omega$ in Toth’s method becomes quite large since the error of $K$ keeps on increasing and accumulating across the neighborhood window $\omega$. As a result, pixels in penumbra region are likely to be wrongly classified as foreground. In figure 1(a), a yellow line is marked across a penumbra shadow region. Figure 1(b) shows the profile of intensity values along the yellow line marked in (a). The red region in figure 1(c) indicates the error score $\Omega$ in Toth’s method by accumulating the difference of $K$. The red region in figure 1(d) represents the error score $\Psi$ in the proposed method by accumulating the difference of $D$. It is clear that the proposed method with smaller error score (area of red region) tends to correctly classify penumbra as shadow, while Toth’s method cannot deal with penumbra properly since their error score is significantly larger especially when the penumbra region becomes broader or $\omega$ becomes bigger. It should be noted that only shading information is considered and no morphological filtering is applied in this comparison.
6. Experimental Results

Several different scenarios are designed to test the performance of the proposed method. Figure 2(a) shows two light sources casting shadows of a car toy on the ground. Figure 2(b) displays the result of Toth’s method in that a broad ring of the penumbra region is not correctly detected as a shadow region. Figure 2(c) demonstrates the results using the proposed method considering all factors. The rim of the shadow region is detected pretty well since the penumbra region is handled properly. In Figure 2(d), the red curve with cross marks shows the performance of the proposed method considering all factors, the green curve with circle marks indicates the performance considering only the proposed shading property, and the blue curve with dot marks represents the results using the shading constraints in Toth’s method. The morphological filters are not applied in both methods since it can not only remove the penumbra ring but also any intrusive parts of the foreground objects.

7. Conclusions

This paper proposed a reliable and efficient moving cast shadow detection algorithm that considers shading, color, texture, neighborhoods and temporal consistency in the scene. The experimental results show that the proposed method can detect umbra as well as penumbra in different kinds of scenarios under various illumination conditions.

Acknowledgements: This research is supported by the National Science Council, Taiwan.

References


