Human Detection from Surveillance Cameras with Considering the Influence of Shadows

Ryo Yonekawa† Masao Izumi†
†Graduate School of Engineering, Osaka Prefecture University

Abstract
Nowadays, there are huge numbers of surveillance cameras installed in the world everywhere. And many researches cover surveillance cameras installed outdoors. Surveillance cameras installed outdoors receive considerable environmental influence. Particularly, shadows raise the incidence of occlusions of the object regions. In this paper, we propose the novel method of human detection with considering the influence of shadows. At first, our proposal method performs background subtraction using normalized distance. Next, it updates background using mixed Gaussian distribution dynamically. Afterwards, it changes the size of the extracted object regions into four phases and performs human detection by using HOG (Histograms of Oriented Gradients) features to classify detection windows into human or non-human category with SVM (Support Vector Machine). Experimental results show that the proposed method is effective to extract correct object regions with surveillance cameras installed outdoors.

Keywords: background subtraction, normalized distance, mixed Gaussian distribution, human detection, Histograms of Oriented Gradients, SVM

1 Introduction
Surveillance cameras have been installed in the world everywhere. They have been installed in such areas as public accommodations, stations, buildings, public loads, etc, where a large number of people come into or go out. Main purpose of using surveillance cameras is crime prevention. However, it is difficult to keep watching the surveillance camera images with human eyes because the abnormal situation rarely happens. Thus we hope the system can automatically detect abnormalities and inform a watchman. To that end, a new image processing technology to detect target objects only from images is necessary. In this study, we perform the experiment for the purpose of detecting only humans precisely from images of fixed cameras installed outdoors.

2 Outline of the proposed method
This study covers surveillance cameras installed outdoors. Therefore, the camera is affected by the dim change of the background and sunshine condition. Because the shadow in particular makes it easy to generate the occlusion of the object regions, it is an obstacle in securing correct object regions. Many conventional methods can adapt to the dim change of the background, but were not able to remove the shadow completely. Accordingly, our proposal method performs background subtraction using normalized distance [1] that is effective for the shadow removal. Next, it updates the background with mixed Gaussian distribution [2][3] to be adapted to the dim change of the background. And it removes the noise of the extracted object regions, performs labeling in each domain and performs human detection using HOG features to classify detection windows into human or non-human category with SVM. We show a flow of the proposed method in figure 1.

![Figure 1. Flow of the proposed method](image-url)
3 Object extraction in consideration of the influence of the shadows

In this section, we suggest the method that combines mixed Gaussian distribution with background subtraction using normalized distance.

3.1 Background subtraction using normalized distance

As the background subtraction that can secure the object regions except the shadows, there is a method using normalized distance [1]. In local area, this method calculates normalized distance from a background image and an input image. When each luminosity value of the background image and that of the input image are correlative to each other, the normalized distance is close to 0. The normalized distance \( D_N \) is calculated in the next formula.

\[
D_N(i, b) = \frac{|i - b|}{|i|} \quad \quad (1)
\]

Where, “\( i \)” is a vector of local area’s luminosity values of the input image. “\( b \)” is a vector of local area’s luminosity values of the background image. \( |i|, |b| \) is calculated with the square root of sum of squares of local area’s luminosity value respectively. Shadows tend to darken the shadowed area with the certain range of proportions. In other words, the ratio of luminosity between shadowed and non-shadowed in each local area keeps constant. For that reason, lower value of normalized distance tends to be shadowed area, and higher value of normalized distance tends to be non-shadowed area. (Figure 2.)

![Figure 2. Background subtraction using normalized distance](image)

3.2 Background update using the mixed Gaussian distribution

Background subtraction using normalized distance needs to fix the background image which we acquired from camera images beforehand. For that reason, this method is very weak to camera shake and the dim change of the background.

Thus we suggest the method that combines mixed Gaussian distribution [2][3] with background subtraction using normalized distance. (Figure 3.) The mixed Gaussian distribution is used to update dynamically the background image. The following is a brief flow of the handling of background update using the mixed Gaussian distribution. Refer to documents [2] for the details of the method.

**Step1**
This method makes mixed Gaussian distribution from a pixel value of the constant section. And, it divides the Gaussian distribution into plural Gaussian distribution by EM algorithm. It calculates a collation degree of each Gaussian distribution on this occasion.

**Step2**
If a distance between a pixel value of the input and the mean of Gaussian distribution is less than 2.5 times of standard deviation of Gaussian distribution, it considers the pixel matches the Gaussian distribution.

**Step3**
If a distance between a pixel value of the input and the mean of Gaussian distribution is less than 2.5 times of standard deviation of Gaussian distribution, it considers the pixel matches the Gaussian distribution. If Gaussian distribution is found, it estimates the background of the pixel than the collation degree of the Gaussian distribution.

**Step4**
It updates the collation degree of each divided Gaussian distribution.

![Figure 3. Proposal method](image)

3.3 Corrective processing (noise reduction, labeling)

After the processing mentioned above, the left noises can be eliminated using expansion reduction. Besides, we perform labeling in order to
human detection using HOG

4.1 Summary of human detection using HOG and SVM

The human detection using HOG [4] is divided into a learning process and a detection process. In the learning process, we calculate HOG from a large quantity of samples and SVM learns a shape characteristic of the target for the detection. In the detection process, a detection window to detect HOG performs raster scan on an image. The detection window changes size and performs multiple raster scan. SVM judges whether HOG which is calculated by a detection window is human or not.

4.2 Calculation of HOG

HOG[4] is calculated by the following flows. In addition, we show the relations among an input image, blocks, cells and pixels be used for calculation of HOG in following figure 4.

Step1 Calculation of gradient intensity and gradient orientation

\[(x, y)\] is the coordinate of images. Luminosity gradient \( f_x, f_y \) is calculated by the next formula.

\[
\begin{cases}
    f_x(x, y) = L(x + 1, y) - L(x - 1, y) \\
    f_y(x, y) = L(x, y + 1) - L(x, y - 1)
\end{cases}
\]  

In addition, gradient intensity and gradient orientation are calculated by the next formula.

Gradient Intensity : \( m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \)  
Gradient Orientation : \( \theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \)

Step2 Making of the gradient orientation histogram of the luminosity

The calculated gradient orientation becomes 0°~360°. However, because it is not necessary to consider the direction of the gradient orientation, we convert it into 0°~180° and divide it into 9 directions by 20° steps. Moreover, add gradient intensity of the luminosity in each calculated directions and make gradient orientation histogram.

Step3 Normalize each block

We normalize each block by using the following formula.

\[ v(n) = \frac{v(n)}{\left( \sum_{k=1}^{N} v(n)^2 \right)^{\frac{1}{2}}} \]  
\( \epsilon = 1 \)

\( N \) : Gradient Orientation Number  
\( \epsilon \) is a coefficient to prevent a denominator from becoming 0.

5 Experiment and consideration

We show an inspection result based on the procedure of Section 3 and Section 4.

5.1 Experimental environment

There were three camera images of each 50 seconds which varied in the photography date, a sunshine condition, and a congestion condition. Besides, we used HOG data which SVM learned for human detection. This data was prepared for by OpenCV2.3.1 [5].

The followings show the condition of this experiment.

- photography scale : 30 meters - 35 meters in width
- camera images① : influence of shadows and trees, camera images② : influence of the shadow, camera images③ : no influence of shadows and trees
- scene number : 3 (50 second) )
- resolution : 640x480
- frame rate : 6 frame/second
- threshold 2.7 of normalized distance
- resized size of object regions for human detection:80x140,90x140,100x140,110x140
- HOG calculation parameter : (cell size 8x8 , block size 16x16 , block movement size 8x8 , the number of edge directions 9 )
5.2 Experimental results

As a preliminary experiment, we test detection precision of HOG data which was prepared for by OpenCV2.3.1 [5]. We found the rate of detection for 1,033 pieces of positive samples (Figure 6) in the preliminary experiment. Following table 1 shows the result.

<table>
<thead>
<tr>
<th>Preliminary experiment</th>
<th>Positive sample</th>
<th>Rate of detection</th>
<th>Detection error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1033</td>
<td>91.7%</td>
<td>8.3%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Example of positive sample

Next, following table 2 shows the result which compared among the mixed Gaussian distribution, normalized distance and proposed method.

<table>
<thead>
<tr>
<th></th>
<th>Mixed Gaussian</th>
<th>Normalized distance</th>
<th>Proposal method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of region</td>
<td>3356</td>
<td>2098</td>
<td>1442</td>
</tr>
<tr>
<td>Rate of human detection</td>
<td>36.5%</td>
<td>55.7%</td>
<td>70.1%</td>
</tr>
<tr>
<td>Number of region</td>
<td>2637</td>
<td>1112</td>
<td>1102</td>
</tr>
<tr>
<td>Rate of human detection</td>
<td>37.5%</td>
<td>81.7%</td>
<td>76.8%</td>
</tr>
<tr>
<td>Number of region</td>
<td>2551</td>
<td>1984</td>
<td>1828</td>
</tr>
<tr>
<td>Rate of human detection</td>
<td>72.3%</td>
<td>80.2%</td>
<td>70.9%</td>
</tr>
</tbody>
</table>

5.3 Consideration

In case of camera images①, the object regions extracted by mixed Gaussian distribution have the low rate of human detection. It is because an object domain received the considerable influence of shadows. Further, background subtraction using normalized distance is necessary to fix a background image. Hence, shaking of trees caused the false detection. On the other hand, the proposal method was able to get the highest rate of human detection because it could cope with the influence of shadows and trees. In the case of camera images② and camera images③, the proposal method gets high rate of human detection. The proposal method is effective in surveillance cameras installed outdoors which cannot avoid environmental changes.

6 Conclusion

We performed background subtraction using normalized distance that is effective for the shadow removal in this study. This method reduced the occlusion of object regions which shadows caused and made it possible to secure a more correct object domain. Surveillance cameras installed outdoors is affected by the dim change of the background and sunshine condition. Therefore, we have proposed the method that combined mixed Gaussian distribution (which was one of methods to update a background dynamically) with background subtraction using normalized distance. As a result of this experiment, we found that the proposal method is effective in surveillance cameras installed outdoors which cannot avoid environmental changes.

References