ADAPTIVE THRESHOLD BACKGROUND SUBTRACTION FOR DETECTING MOVING OBJECT ON CONVEYOR BELT

D.P.Tripathy*, K.Guru Raghavendra Reddy2
*1Dept. Of Mining Engineering, NIT, Rourkela.
*email: dptripathy@nitrkl.ac.in

Abstract. Moving object detection is an important task in many computer vision classifications applications. The goal of this study is to identify a moving object detection method that provides a reliable and accurate identification of objects on the conveyor belt. In this paper, a study of the moving object detection methods is presented. Firstly, moving object detection pixel by pixel was performed using background subtraction, frame difference method. The threshold value in both background subtraction and frame difference is a fixed value, which determines the accuracy of object identification. The adaptive threshold values were calculated for both the methods to improve the accuracy. The performance of these methods was compared with the ground truth image.

Keyword. Motion segmentation; Moving objects; Background subtraction; Adaptive Threshold; Conveyor.

INTRODUCTION

In the mineral processing industry, minerals extracted were transferred using conveyor belt system. In recent times, research and development efforts were mainly focused on tracking ore properties for monitoring ore quality, size distributions, and ore separation on conveyor belts using various imaging techniques using a static camera. In the image analysis process, the first important task was to identify the ore particles from the moving conveyor from the video sequence recorded using the camera placed above the conveyor belt.

To identify the moving objects, mostly the background subtraction, frame difference and optical flow methods were used [1]. Background subtraction method was widely used approach for detecting moving objects from static cameras [2]. The rationale for the approach was to detect moving objects by the difference between the current frame and a reference frame called background model. This method needs the representation and initialization and the updating of background model [3]. Frame difference method was a technique, which checks the difference between two consecutive video frames. If the pixels value was changed, it apparently indicates some change in the movement in the next frame [4]. This method was commonly used in object detection as it has advantages of small calculations and high performance of real-time situations. Optical flow method makes use of the instantaneous velocity, generated in the continuous movements of the pixels in the moving objects. In the case of unknown scene information, this method can detect the moving objects. Optical flow method has reduced the anti-noise performance, and the calculation is quite complicated, and the hardware requirements are also very high, so it is not suitable for real-time processing [5].

In this paper, in a single static camera condition, background subtraction and frame difference methods were used to identify the moving objects from the conveyor belt. A combined background modeling with dynamic threshold selection method based on the background subtraction and frame difference was used to enhance the effect of moving object detection. These methods were implemented on videos captured from the conveyor belt and performance was compared with ground truth image.

METHODOLOGY

To identify moving objects on the conveyor belt, there are mainly two steps, image segmentation, and feature extraction and identification. Image segmentation is an operation that takes an image as input and returns one or more regions or contours as output.

• Background Subtraction Method

In background subtraction method, moving objects were detected by subtracting the current image from the background image as shown in Figure 1. The first frame of the video sequence was considered as referenced background frame. The current frame was subtracted from the referenced background frame. If the resulted pixel was greater than the predefined threshold value, then the pixel was considered as a background pixel. The choice of the threshold value for the background subtraction was very important to achieve the good motion detection [6].
Figure 1. Background subtraction method

For a video sequence \( f \), \( f_k(x, y) \) represents the current frame, and \( B(x, y) \) represents the background image. \( D_k(x, y) \) is the difference between the current frame and background image as shown in equation (1) & (2).

\[
D_k(x, y) = f_k(x, y) - B(x, y)
\]  
\[
D_k(x, y) = \begin{cases} 
1 & D_k(x, y) \geq T \\
0 & \text{else} 
\end{cases}
\]

Where \( T \) is the predefined threshold value.

**Frame Difference Method**

Frame differencing method was popularly known as temporal difference method, in which the difference between video frame at time \( t \) and the frame at time \( t-1 \) was calculated.

In two-frame differencing method, two consecutive frames were considered and subtracted pixel wise. If the resulted pixel was greater than predefined threshold value, then the pixel was considered as background pixel as shown in Figure 2 [4, 7].

\[ D(x, y) = ||f_t(x, y) - f_{t-1}(x, y)|| \geq T \iff ||f_t(x, y)|| < ||f_{t-1}(x, y)|| < T \]  

**Combining background subtraction and frame difference**

In background subtraction, the outline and shape were conserved quite well. But due to the simplicity, the result may be affected by the noise. In frame difference, the results obtained may be which is affected by the noise little, but contains incomplete moving object information.

Considering all of the above, firstly background subtraction and frame difference were performed and then “AND” operation was performed on both subtracted images as shown in Figure 3. These results of the two methods, not only conserved the most information and outline of moving objects, but also obtained the position and region of objects [8].

---

Figure 2. Frame difference method

For the video sequence \( f \), the frame at \( k \) and frame at \( k+1 \) are considered. Then the frame difference between these frames was calculated. If the resulted pixel was greater than predefined threshold value \( T \), then the pixel was considered as background pixel as expressed in equation (3).

\[ D(x, y) = ||f_k(x, y) - f_{k+1}(x, y)|| \geq T \iff ||f_k(x, y)|| < ||f_{k+1}(x, y)|| < T \]  

---
Figure 3. Combining background subtraction and frame difference

- Adaptive background subtraction
  In background subtraction method, moving objects were detected by subtracting the current image from the background image. The optimal threshold for the subtracted image was calculated using the adaptive threshold method as shown in Figure 4.

  Figure 4. Adaptive background subtraction method

For a video sequence $f$, $f_k(x, y)$ represents the current frame, and $B(x, y)$ represents the background image. $D_k(x, y)$ is the difference between the current frame and background image as shown in equation (4) & (5) [9,10].

$$D_k(x, y) = f_k(x, y) - B(x, y)$$  \hspace{1cm} (4)

$$D_k(x, y) = \begin{cases} 1 & D_k(x, y) \geq T \\ 0 & \text{else} \end{cases}$$  \hspace{1cm} (5)

Where $T$ is the adaptive threshold value calculated using Otsu’s thresholding method.

- Adaptive frame difference method
  In frame differencing method, the difference between video frame at time $t$ and the frame at time $t-1$ were calculated. In two frame differencing method, two frames were considered and subtracted pixel wise. The optimal threshold for the subtracted image was calculated using the adaptive threshold method as shown in Figure 5 [11].

- Adaptive background subtraction frame difference method
  The background subtraction and frame difference methods were performed using adaptive threshold value and then “AND” operation was performed on both subtracted images as shown in Figure 6.

Figure 5. Adaptive frame difference method
RESULTS AND DISCUSSION

The experiment was carried out on videos captured on the conveyor belt with multiple moving objects to check the effectiveness and performance of the moving object detection algorithms. The analysis was performed on 3.40GHz Core i7 processor, and the videos were recorded with frame resolution 640*480 pixels in Avi format.

Moving object detection methods as background subtraction and frame difference were used. The threshold value was selected after analyzing the grey histogram of the background image and frame extracted from the video. Initially the background from the video frame was segmented using different threshold values ranging from 15-20 and it was found that threshold value of 15 gave better results. Hence, the threshold value for both background subtraction and frame difference was considered as 15. A combined background subtraction and frame difference method were performed using AND operator. To improve the accuracy of moving object identification, the adaptive threshold value was calculated for the methods. All these methods were applied to the videos, and the respective ground truth was calculated for the frames as shown in Table 1. The performance of the moving object detection methods was evaluated by Accuracy, F-score, MSE (mean squared error) and RMSE (root mean squared error) using equations (6) (7) (10) & (11).

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)
\]
\[
\text{F-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)
\]
\[
\text{Precision} = \frac{TP}{TP + FP} \quad (8)
\]
\[
\text{Recall} = \frac{TP}{TP + FN} \quad (9)
\]
\[
\text{MSE} = \frac{FP + FN}{M \times N} \quad (10)
\]
\[
\text{RMSE} = \sqrt{\text{MSE}} \quad (11)
\]

F-score was the weighted harmonic mean of Precision and Recall and acts as a single measurement to compare different methods. TP was the number of foreground pixel classified as foreground. FN was the number of foreground pixel classified as background. FP was the number of background pixel classified as foreground and TN was the number of background pixel classified as background that was calculated by comparing the background subtracted frame with the ground truth image of the same frame.
The time taken for identifying moving objects for a single frame was shown in Table 2. The performance of adaptive background subtraction method was better with 99.6% accuracy and MSE value of 0.003 as compared to other methods as shown in Table 3. In adaptive background subtraction frame difference method it was observed that the performance was degraded as the method projects few foreground pixels as background pixels.

<table>
<thead>
<tr>
<th>Method</th>
<th>Frame – 141 (Video-1)</th>
<th>Frame – 91 (Video-2)</th>
<th>Frame – 60 (Video-3)</th>
<th>Frame – 101 (Video-4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Captured frame</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ground Truth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background subtraction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frame difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background subtraction and frame difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive background subtraction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive frame difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adaptive background subtraction frame difference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Comparison of moving object detection methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>Time per frame (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background subtraction</td>
<td>0.252025</td>
</tr>
<tr>
<td>Frame difference</td>
<td>0.324824</td>
</tr>
<tr>
<td>Background subtraction and frame difference</td>
<td>0.314925</td>
</tr>
<tr>
<td>Adaptive background subtraction</td>
<td>0.273534</td>
</tr>
<tr>
<td>Adaptive frame difference</td>
<td>0.294986</td>
</tr>
<tr>
<td>Adaptive background subtraction frame difference</td>
<td>0.581564</td>
</tr>
</tbody>
</table>

Table 2. Comparison of processing time of each method
Table 3. Performance of the moving object detection methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>Accuracy</th>
<th>F-score</th>
<th>MSE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background subtraction</td>
<td>95.9</td>
<td>0.852062</td>
<td>0.040052</td>
<td>0.20013</td>
</tr>
<tr>
<td>Frame difference</td>
<td>89.8</td>
<td>0.640627</td>
<td>0.106862</td>
<td>0.326898</td>
</tr>
<tr>
<td>Background subtraction and frame difference</td>
<td>97.8</td>
<td>0.898121</td>
<td>0.021497</td>
<td>0.14662</td>
</tr>
<tr>
<td>Adaptive background subtraction</td>
<td><strong>99.6</strong></td>
<td><strong>0.986902</strong></td>
<td><strong>0.003005</strong></td>
<td><strong>0.054814</strong></td>
</tr>
<tr>
<td>Adaptive frame difference</td>
<td>89.9</td>
<td>0.570148</td>
<td>0.100505</td>
<td>0.317025</td>
</tr>
<tr>
<td>Adaptive background subtraction frame difference</td>
<td>97.8</td>
<td>0.898121</td>
<td>0.021497</td>
<td>0.14662</td>
</tr>
</tbody>
</table>

CONCLUSION

The primary goal of this study was to provide a reliable and accurate moving object detection method that can identify the moving objects from the conveyor belt. The moving object detection pixel by pixel is done using background subtraction and frame difference perceptively. The threshold value for the both methods was calculated using Otsu thresholding method. These methods were applied to videos captured from the conveyor belt and the accuracy of these methods was compared with ground truth calculated for the frames. Standard background subtraction method was fast and can be implemented in hardware easily. The performance of these methods was increased, when the threshold was calculated dynamically instead of using a fixed value. The comparative results indicated that the adaptive background subtraction method gave better accuracy of 99.6% when compared with the ground truth image.

REFERENCES