Vehicle detection using background subtraction and clustering algorithms

Puguh Budi Prakoso¹, Yuslena Sari*²
¹²Faculty of Engineering, Universitas Lambung Mangkurat, Indonesia
¹²National Center for Sustainable Transportation Technology, Indonesia
*Corresponding author, e-mail: puguh.prakoso@ulm.ac.id, yuzlena@ulm.ac.id

Abstract
Traffic congestion has raised worldwide as a result of growing motorization, urbanization, and population. In fact, congestion reduces the efficiency of transportation infrastructure usage and increases travel time, air pollutions as well as fuel consumption. Then, Intelligent Transportation System (ITS) comes as a solution of this problem by implementing information technology and communications networks. One classical option of Intelligent Transportation Systems is video camera technology. Particularly, the video system has been applied to collect traffic data including vehicle detection and analysis. However, this application still has limitation when it has to deal with a complex traffic and environmental condition. Thus, the research proposes OTSU, FCM and K-means methods and their comparison in video image processing. OTSU is a classical algorithm used in image segmentation, which is able to cluster pixels into foreground and background. However, only FCM (Fuzzy C-Means) and K-means algorithms have been successfully applied to cluster pixels without supervision. Therefore, these methods seem to be more potential to generate the MSE values for defining a clearer threshold for background subtraction on a moving object with varying environmental conditions. Comparison of these methods is assessed from MSE and PSNR values. The best MSE result is demonstrated from K-means and a good PSNR is obtained from FCM. Thus, the application of the clustering algorithms in detection of moving objects in various condition is more promising.

Keywords: FCM, ITS, K-means, OTSU, vehicle detection

1. Introduction
In transportation, vehicle detection systems can be defined as systems capable of detecting vehicles and measuring traffic parameters such as counting, speed, incident, etc [1, 2]. Vehicle detection can also be taken advantages of various transportation applications such as: vehicle navigation system, vehicle security, safety and others. The detection of vehicles with video cameras is one of the most promising non-intrusive technologies for large-scale data collection and adoption of sophisticated traffic control and management. Vehicle detection is also the basis of vehicle tracking. Correct vehicle detection results in better tracking. Vehicle tracking methods are highly influential in tracking results [3-5]. In addition, due to advancement in digital technology, camera is getting cheaper in production, so computer processing of digital image and video is also becoming faster and more cost effective [6-10]. Tracking of recorded videos in varying environmental conditions provides a difficult processing time during image segmentation of video [11, 12]. Various conditions affect the image quality of the resulting video recordings. For instance, the effects of lighting during the night and the rain result in very low quality video recording. The reason is that the lighting at night often undergoes drastic changes as well as rain that can degrade the image quality in video [3, 13].

Most static camera-based tracking systems use background subtraction method [14-16]. The main process in this method is analysing useful video sequences to extract the foreground and background images. This can be achieved by initially modelling the background of the video image. And in separating the foreground and the background, an appropriate threshold value is required [5, 15-17]. OTSU algorithm is the most classical subtraction background method used to segment an image. However, in determining threshold, OTSU algorithm cannot detect image optimally when dealing with other images in grey level. The shortcomings of the OTSU algorithm have been resolved by several researchers. Bhargava [18] noted that in addition to efficient, clustering method is able to provide...
segmentation results in the image better and adaptive to changes in the environment and quick momentary changes. There are various types of clustering: K-means clustering, Fuzzy C-means clustering, mountain clustering method and subtractive clustering method [19-23]. This study is focused on the tracking of moving objects performed under various environmental conditions, which are recorded in a static video camera and a condition where the image quality is low. By proposing adaptive threshold algorithm obtained from FCM and K-means algorithm, it is expected that the accuracy to detect moving objects in the various environment condition is intensified.

2. Related Research

There are numerous studies conducted on the detection of moving objects, which can be seen in the Table 1 [17, 18, 24, 25].

Table 1. Research Related to Moving Object Detection

<table>
<thead>
<tr>
<th>Author and year</th>
<th>Method</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xiangju Sun, Wei Wang, Ping Gao, 2010</td>
<td>Kirsch Operator with Optical Flow (KOF) algorithm and OTSU</td>
<td>Background subtraction algorithm is not optimal for motion detection. 85% accuracy results with background subtraction algorithm and 91% accuracy results with KOF algorithm</td>
</tr>
<tr>
<td>Ruri Suko Basuki, Mochamad Hariadi and R Anggi Pramunendar, 2012</td>
<td>FCM and OTSU</td>
<td>Both algorithms show that the FCM algorithm produces a smaller threshold value so the number of pixel errors is less than the one in the OTSU algorithm. The best results for grouping are achieved using Fuzzy C-Means clustering algorithm. But the best segmentation quality is obtained using Fuzzy C-Means clustering algorithm.</td>
</tr>
<tr>
<td>Bhargava and Sharma, 2016</td>
<td>FCM and K-means</td>
<td>The detection of moving objects has been done using simple background reduction. Single moving object tracking has been done utilizing the modified shifting mean method and Kalman filter.</td>
</tr>
<tr>
<td>Rawat and Raja, 2016</td>
<td>Shift algorithm</td>
<td></td>
</tr>
</tbody>
</table>

3. Technique of Selecting Threshold

Thresholding is the simplest method of image segmentation. Individual pixels in a grayscale image are marked as ‘object’ pixels if their value is greater than certain threshold value (assuming the object is brighter than the background) and as ‘background’ pixels in the other way round [2, 24]. Thresholding techniques, which are used in this research are background subtraction and clustering. Background subtraction is performed by using OTSU while clustering is conducted by employing FCM and K-means.

3.1. OTSU

The OTSU method performs discriminant analysis by determining a variable by distinguishing between two or more groups naturally. The OTSU method starts with normalizing the image histogram as a probability discrete density function as:

\[ p_r(r_q) = \frac{n_q}{n} \text{ where } q = 0,1,2, \ldots, L-1 \] (1)

where \( n \) is the total number of pixels in the image, \( n_q \) is the number of pixels \( r_q \), and \( L \) is the total number of image intensity levels. In determining the value of \( T \) by maximizing between class variances is defined as follows:

\[ \sigma_b^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \] (2)

which is obtained from:

\[ \omega_0 = \sum_{q=0}^{L-1} p_q(r_q) \text{ where } \omega_1 = \sum_{q=0}^{L-1} p_q(r_q) \] (3)
Vehicle detection using background subtraction and clustering algorithms (Yuslena Sari)

\[
\mu_0 = \sum_{q=0}^{k-1} \frac{wp_q(r_q)}{a_o} \quad \text{where} \quad \mu_1 = \sum_{q=k}^{L-1} \frac{wp_q(r_q)}{a_o} \quad (4)
\]

\[
\mu_T = \sum_{q=0}^{L-1} wp_q(r_q) \quad (5)
\]

### 3.2. Fuzzy C-Means (FCM)

FCM generates matrixes that contain ownership of multiple objects in each cluster. In FCM, threshold is defined as the average value of maximum on the object with the smallest and minimum center with middle center.

\[
SSE = \sum_{i=1}^{N} \sum_{j=1}^{c} u_{ij}^m \|x_i - c_j\|^2, \quad 1 \leq m \leq \infty
\]

where for the above formula shows the ownership of each pixel in the cluster:

\[
\sum_{i=1}^{c} u_{ij} = 1, \quad 1 \leq j \leq n \quad (7)
\]

\[
u_{ij} \geq 0,1 \leq i \leq c, 1 \leq j \leq n \quad (8)
\]

\[
\sum_{i=1}^{n} u_{ij} = 1, \quad 1 \leq i \leq c \quad (9)
\]

The first step of FCM algorithm is the searching of data input in the image. After that, it will select the number of the cluster and its value. And then the division of the matrix is calculated using:

\[
u_{ij} = 1 / \sum_{j=1}^{c} \left( \frac{d_{ik}}{d_{jk}} \right)^{2(m-1)} \quad (10)
\]

finally, the cluster in the center is changed with new cluster:

\[
c_j = \frac{\sum_{k=1}^{n} u_{jk}^m x_k}{\sum_{k=1}^{n} u_{jk}^m} \quad (11)
\]

### 3.3. K-Means

The K-means algorithm is the most famous and widely used as clustering method in many fields because it is simple, easy to implement, has the ability to cluster large data, able to handle outlier data and linear time complexity \(O(nKT)\), where \(n\) is the number of documents, \(K\) is the number of clusters and \(T\) is the number of iterations. K-means algorithm is a partitioning clustering method that separates data into different groups. This algorithm works on numeric attribute. With iterative partitioning, the K-means algorithm is able to minimize the average distance of each data to its cluster. K-means algorithm consists of the following steps: (1) determining the value of \(k\) as the number of clusters to be formed, (2) generating the initial centroid (cluster center point) at random, (3) calculating the distance of each data to each cluster center by using the formula of correlation between two objects i.e. Euclidean Distance, (4) grouping any data based on the closest distance between the data and the center, and then (5) determining the position of a new cluster center by calculating the mean value of the data that is in the same cluster center.

### 4. Research Method

The method proposed for this research applies K-means and FCM algorithm to the background subtraction with frame difference in pixel grouping of the image into the foreground or background. Figure 1 shows the proposed method. Video data as input data is image data prepared in the initial data processing, made in the experimental folder. Then a boundary between pixels as foreground and background is given in the process using background subtraction. Finally Morphology is utilized to get a good object in the detection.
5. Results and Analysis

The comparison of the experiment results is summarized in the following Table 2. Table 2 depicts the segmentation result. It can be seen that the K-Means and FCM method is very robust in changes compared to OTSU, so more frequent changes from each frame. Measurement of evaluation utilizes Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) to know the accuracy of each method, which are calculated using the following formulas [26, 27]:

\[
SE(X, Y) = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [X(i,j) - Y(i,j)]
\]

\[
PSNR(X, Y) = 10 \cdot \log \left( \frac{\text{max}^2}{\text{MSE}(X,Y)} \right)
\]

where X is the actual image, Y is the segmentation image of the MxN size and max is the maximum pixel value range of the image. Table 3 shows the MSE and PSNR average result. The histogram images of MSE and PSNR measurements between OTSU, FCM and K-Means can be seen in the following Figures 2 and 3.

Table 2. Segmentation Result

<table>
<thead>
<tr>
<th>Frame</th>
<th>062</th>
<th>103</th>
<th>240</th>
<th>369</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
<tr>
<td>OTSU</td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
</tr>
<tr>
<td>FCM</td>
<td><img src="image9.png" alt="Image" /></td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>K-Means</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
<td><img src="image16.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Table 3. MSE and PSNR Average Result

<table>
<thead>
<tr>
<th></th>
<th>OTSU</th>
<th>FCM</th>
<th>K-Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>25459.8</td>
<td>7872.461</td>
<td>6695.008</td>
</tr>
<tr>
<td>PSNR</td>
<td>8.248</td>
<td>13.438</td>
<td>14.432</td>
</tr>
</tbody>
</table>

Based on the graphs shown in the figures above, there are many values of MSE approaching 0, which can be interpreted that the results of the image and the results of manual tracking do not appear so there is a difference or no moving objects in the frame. Meanwhile, when there are moving objects, the results obtained from OTSU, FCM and K-Means have values higher than 0. It is figured out that the K-Means method is able to obtain detail changes better than FCM and OTSU. As for PSNR tends to be infinite, a very good value is depicted from the one greater than zero.

4. Conclusion

From the experimental research on the detection of moving objects, it can be summarized that K-Means and FCM algorithms are superior to the OTSU algorithm. In addition, many other researches also demonstrated that the cluster algorithm without supervision, which has been successfully applied to some problems in pixels is FCM. It can be said that the mean values of PSNR obtained from the three methods are quite good. Nevertheless, the one generated by FCM produces the best PSNR, which is greater than the other methods. So it can be concluded that clustering algorithms (K-means and FCM) are more prominent to the background subtraction algorithm (OTSU). The MSE values generated from K-Means and FCM algorithms show fewer errors, thus the application of the clustering algorithms in detection of moving objects in various condition is more promising.

Acknowledgement

This paper is supported by USAID through Sustainable Higher Education Research Alliances (SHERA) Program-Centre for Collaborative (CCR) National Center for Sustainable Transportation Technology (NCSTT) with Grant No. IIE00000078-ITB-1.
References


