Localized Background Subtraction Feature-Based Approach for Vehicle Counting

Najeeb Ur Rehman Malik*

* Department of Electronics and Communications, Arab Academy for Science, Technology and Maritime Transport, Egypt
* School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia
* Iskandar Region Development Authority (IRDA)
† Mehran University of Engineering & Technology Jamshoro, Sindh, Pakistan
{m.atef.elkhoreby@aast.edu; syed@fke.utm.my; musa@fke.utm.my; syaril@irda.com.my; rehman@graduate.utm.my}

Abstract—This paper presents vehicle counting using local features derived from background subtraction method combined with the traditional SVM classifier. The main advantage of the proposed method is that it manages to overcome counting problem in different challenging situations such as during rainy day condition, which is yet a difficult problem in the standard background subtraction method. In this method, the background subtraction is performed within a small predefined region of a lane or road, thus increasing the computational speed. Next, two key features i.e. area covered by the motion indexed graph and number of edge pixel computed in the designated areas in the image are computed to form a 2D feature vector. Finally, we use SVM for classifying this vector into either vehicle or noise. Comparisons between the proposed method and other method show the potential of our approach.

Keywords—background subtraction; feature extraction; vehicle counting

I. INTRODUCTION

Estimating the number of vehicles in intelligent transportation systems (ITS) based on traffic video sequences is an important task, as it can provide reliable information for traffic management and control [1]. It can be used for understanding the traffic status, such as road-traffic intensity, lane occupancy, and congestion level, which helps the drivers to avoid traffic congestion and spend less time in traffic. Subsequently, development authority can use such information to design a better solution for road traffic.

Vehicle detection is a primary step in the traffic monitoring system. Some of the traditional methods for vehicle detection are inductive loops, passive magnetic sensors, and microwave detectors. These methods, in general, have limitations such as unable to detect stationary vehicles in addition to being highly complex in the hardware design which translates to high cost [2]. Therefore, the vision-based traffic monitoring system has attracted many researchers due to its two-fold advantages: less costly and easier to deploy [3]. Moreover, the advancement in the development of computational technologies has made vision-based vehicle counting an extremely attractive option for ITS. Nevertheless, such a system does face some challenges such as changing lighting conditions, occlusion, and unfavorable weather conditions. Needless to say, such problems have opened up a new horizon of research opportunities.

Vehicle detection can be performed in a number of ways such as frame differencing, background subtraction and optical flow. Frame differencing is the simplest method. It can detect only part of the moving object by comparing two or three consecutive frames using a simple threshold [8]. It has low computational complexity and easy to implement, however it has low accuracy and not robust against noise. Background subtraction, on the other hand, is the most common method used for vehicle detection. It is achieved by subtracting each frame from the background model. The main challenge of this method is in designing an optimum background model [9]. Lastly, the optical flow approach is to compute the two-dimensional velocity vector, which is formed when relative motion occurs between the camera and the target. The main drawback of the optical flow method is that it is computationally intensive [10].

In this paper, a vehicle counting algorithm that can also cater to rainy weather condition is proposed. The algorithm consists of three main stages (background subtraction, feature extraction, and support vector machine (SVM) classifier). Prior to the background subtraction stage, a few specific detection regions will be identified. These are the regions where background subtraction will be performed. Motion index which measures the occupancy of these regions will be computed. In the feature extraction stage, two features will be extracted i.e. the area and the edge. Finally, SVM classifier is applied to discriminate between the moving objects and noise. The technique has shown better performance when compared to the one proposed by Juan et al. in [11] for counting vehicles under rainy and normal weather condition.

The rest of this paper is organized as follows: Section II gives a survey of the existing techniques in vehicle counting. Section III describes the proposed method. Section IV presents the results and discussion, and finally, the conclusion is discussed in Section V.

II. RELATED WORK

In one of a recent works, Ershadi et al. [4] proposed a vehicle counting under different weather conditions with a vibrating camera based on background subtraction and some
analysis of the headlight size, location, and area. In addition, they also used particle filter for the vehicle tracking. Their results showed that there was a big difference in the detection rate between detection during normal day and detection during poor weather condition. They found out that during poor weather conditions such as (snowy and dusty) the detection rate went as low as around 82% while during normal weather condition the detection rate was around 98.9%.

In another approach, a multiscale edge fusion method was proposed for vehicle detection under different weather conditions [5]. Once again, the obtained results showed a big gap between the precision value during abnormal and normal weather conditions. The precision value for abnormal weather condition was around 85% while for normal weather condition was around 95%.

In [6], El-khoreby and Abu-Bakar proposed a vehicle counting method based on approximate median filter background modeling combined with triangle threshold method. The main problem in this method is in setting the appropriate number of frames to be used to model the background. They realized that this number is influenced by the frame rate and the amount of the movement.

In [7], Tian et al. proposed a vehicle counting method under complex weather conditions based on analyzing and extracting two features i.e. gradient and range feature, on detection lines. Their results showed that the worst case happens in rainy weather conditions.

The basic idea in the proposed method is inspired by the work proposed by Juan et al. in [11]. Particularly, we adopt both the background subtraction within predefined regions and motion index (MI) values, but differ significantly in the post processing steps. Instead of using adaptive threshold over the MI graph as in [11] method, we extract two features, one from the MI graph and the other from the background subtraction result, and form them into a 2D feature vector. Finally, support vector machine (SVM) classifier is applied to discriminate the 2D feature vectors between vehicles and noise. As shall be illustrated in this paper, by introducing these two features, we manage to improve the performance of the vehicle counting both during normal and abnormal weather conditions.

III. PROPOSED METHOD

As stated previously, our proposed method consists of three stages. The first stage is applying background subtraction over the predefined regions and calculates the motion index (MI) value for each frame. The second stage is extracting two features, *area feature* which is computed from motion index graph and *edge feature* which calculated from the result of the background subtraction process. These two features are then concatenated to form a 2D feature vector. Based on these 2D feature vectors, SVM classifier is applied to differentiate between the vehicles and noise. Fig.1 illustrates the process flow of the proposed method.

A. Background Subtraction

At the beginning of this stage the detection regions will be identified with few constraints:

• The number of the detection regions is equal to the number of lanes and each one covers only one lane.
• Each detection region should approximately have the same width and length of a typical compact car. Any vehicle larger than this will be detected without a problem.
• The video should be recorded from a camera which is placed above the road to avoid occlusion from vehicles coming from adjacent lanes. Fig. 2 shows an example of the detection regions.

\[
D(x, y) = |L(x, y) - B(x, y)| \quad \forall (x, y) \in \text{detection region.} \quad (1)
\]
where $L(x, y)$ and $B(x, y)$ represent the intensity level of pixels in the current frame and the reference image inside the detection region respectively. The average intensity value of background subtraction result is calculated to find the motion index (MI) value for each frame as shown in Eq. (2).

$$MI = \frac{1}{255^2} \sum_{i=1}^{N} D(x, y)$$ (2)

where $N$ is the total number of pixels in the detection region. The motion index values are calculated for each detection region as shown in Fig. 3.

It is clear from Fig. 3 that when a vehicle goes inside the detection region the value of the MI value increases till it reaches the peak when the vehicle is totally inside, and the MI value decreases as the vehicle goes outside the detection region. The method proposed by Juan et al. applied an adaptive threshold based on these MI values. On the other hand, in our method we propose an extension to idea by forming a 2D feature vector and then employ SVM classifier to discriminate between vehicles and noise as described in the next subsection.

B. Feature Extraction

In this stage, two features namely MI area and edge pixels are extracted. Based on these two features SVM classifier will be used to differentiate between vehicles and noise.

1) Area: The area here refers to the area under the MI curve at the vicinity of the peak. To compute the area, the initial step is to find all local maximum points $z^k$ in the motion index graph (these maximum must be at least 30 frames apart so as to avoid double counting) and all local minimum points. Next, for every local maximum point, the area is estimated from two local minimum points with one point to the immediate left and another to the immediate right of the maximum point. The computation is given as in Eq. 3.

$$A(z^k) = \int_{m^k}^{n^k} MI \, df$$ (3)

where $A(z)$ is the area value for the local maximum point $z^k$. $m^k$ and $n^k$ are the two adjacent local minimum points, to right and left of $z^k$, respectively. This computation is repeated for all the local maximum points. Higher area value indicates moving vehicle while low value indicates noise. Fig. 4 illustrates the area values for every local maximum point based on motion index graph.

2) Edge pixels: The edge is obtained by applying Sobel edge detection on the background subtraction result $D(x, y)$. Following this, the edge value can be calculated by summing up all the non-zero detected edge pixels. Eq. 4 is employed for this computing this value.

$$E(x, y) = \sum(M(D(x, y)) > 0) \quad \forall (x, y) \in detection \, region$$ (4)

where $M(D(x, y))$ is the magnitude of the Sobel gradients performed on $D(x, y)$ at all $(x, y)$ locations in the detection region.
The performance of the proposed method is evaluated by calculating the precision (P), recall (R), and F-measure as given in Eqs. (5–7).

\[
\text{Precision} (P) = \frac{TP}{ND} = \frac{TP}{TP+FP} \tag{5}
\]

\[
\text{Recall} (R) = \frac{TP}{NR} = \frac{TP}{TP+FN} \tag{6}
\]

\[
F - \text{measure} (F) = 2 \frac{P \times R}{P + R} \tag{7}
\]

where ND is the number of detected vehicles, NR is the actual number of vehicles or the ground truth, true positive (TP) is the number of detected vehicles that are correct, false negative (FN) is the number of the vehicles that are not detected and false positive (FP) is the number of detected vehicles that are incorrect.

As for comparison, we also utilized K-means as another potential candidate for a classifier. We then benchmarked our method to that of Juan’s method. The results are given below.

**TABLE II. COMPARISON BETWEEN THE PROPOSED METHOD AND METHOD BY JUAN ET AL.**

<table>
<thead>
<tr>
<th>K-means</th>
<th>Juan</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R</td>
</tr>
<tr>
<td>Video1</td>
<td>88.8</td>
<td>88.8</td>
</tr>
<tr>
<td>Video2</td>
<td>100</td>
<td>92.9</td>
</tr>
<tr>
<td>Video3</td>
<td>100</td>
<td>78.6</td>
</tr>
<tr>
<td>Video4</td>
<td>84.2</td>
<td>88.9</td>
</tr>
<tr>
<td>Video5</td>
<td>100</td>
<td>90.9</td>
</tr>
</tbody>
</table>

Table II shows the counting accuracy between the proposed method based on (SVM and K-means) and method by Juan et al. for five different videos in rainy and normal weather conditions. It is evident from these results that the proposed method based on SVM performs better in terms of precision, recall and f-measure values.

**TABLE III. COMPARISON BETWEEN THE PROPOSED METHOD AND METHOD BY JUAN ET AL.**

<table>
<thead>
<tr>
<th></th>
<th>K-means</th>
<th>Juan</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>94.6</td>
<td>84.04</td>
<td>95.02</td>
</tr>
<tr>
<td>Recall</td>
<td>88.02</td>
<td>97.46</td>
<td>95.72</td>
</tr>
<tr>
<td>F-measure</td>
<td>90.96</td>
<td>90</td>
<td>95.04</td>
</tr>
</tbody>
</table>

Table III shows the average precision, recall and f-measure values for the five videos for the proposed method based on (SVM and K-means) and method by Juan et al. The average f-measure for the proposed method based on SVM is 95% while it is around 90% in Juan et al. method.

**V. CONCLUSION**

In this paper, we have proposed a vehicle-counting approach based on background subtraction, feature extraction, and SVM classifier. The proposed method works well even under rainy weather conditions. As for the performance, it has
been demonstrated that the proposed method gives better results compared to the method by Juan et al. in terms of precision and F-measure values. For future work, we would explore further on how to reduce the false negative and positive rates.

ACKNOWLEDGEMENT

This work was supported by Universiti Teknologi Malaysia (UTM) and Ministry of Education (MoE) under Research University Grant Q.J130000.2523.19H86.

REFERENCES


