Experiments of Road Vehicle Detection Using Very High-resolution Remote Sensing Images

* Qulin TAN, Hao WANG
Photogrammetric Engineering & Remote Sensing, School of Civil Engineering, Beijing Jiaotong University, Beijing, 100044, China
* Tel.: 086-010-51688207, fax: 086-010-51683764
* E-mail: qltanbjtu@163.com

Received: 13 October 2013 /Accepted: 22 November 2013 /Published: 30 December 2013

Abstract: Road vehicle detection using very high-resolution remote sensing images has a unique advantage of covering a large area at the same time over all ground-based detectors. But the detection of small vehicle-object in remote sensing imagery is still a challenging task. A scheme was proposed to detect road vehicle objects from airborne color digital orthoimagery based on image segmentation and fuzzy logic classification. Firstly, a vector-generated road mask was used to constrain detection of vehicles to road region. Secondly, image segmentation algorithm was performed to form image objects in the preprocessing orthoimagery. Finally, based on a set of fuzzy logic rules defined by membership functions, vehicle objects were detected and separated from other objects. A representative set of road segment images was selected from available images to test the proposed scheme. Experimental results indicate that the detection rates of all test road-segments are high with very few false alarms. Copyright © 2013 IFSA.

Keywords: Remote sensing, Vehicle detection, Segmentation, Fuzzy logic, Classification.

1. Introduction

With the availability of sub-metric resolution remotely sensed imagery, the automatic detection, characterization and monitoring of traffic vehicles using airborne or satellite imagery with its unique wide-area coverage, synoptic view has become an emerging field of research [1-3]. It has great potentiality for supporting traffic-related issues, such as traffic flow monitoring, transportation planning, and estimation of traffic-related pollution over traditional ground-based estimation. Ground-based sensors, like inductive loop detectors, stationary cameras, and vision-based video monitoring systems, can deliver accurate, reliable, timely, yet merely point-wise measurements. High-resolution remote sensing imagery from airborne or satellite based platforms on the other hand can give us synoptic views of complex traffic situations and the associated context, providing road traffic information covered over a large area. Such views are particularly important for urban traffic in big cities, where this type of snapshot information cannot be acquired by other means except space remote sensing data. These data provide a spatially rich dataset to supplement the temporally rich dataset from ground based measurements, and can either be used in research for improving traffic models or as information source for operational monitoring systems.

Currently, more reliable vehicle detection can only be achieved from the very-high-resolution aerial remote sensing images. Existing approaches for vehicle detection can be categorized based on the underlying type of vehicle modeling [4-9]. There are
two basic kinds of vehicle models that have been used: (1) an appearance-based implicit model by clustering appropriate pixel groups into potential vehicles, and (2) an explicit model. The implicit model typically consists of intensity or texture features surrounding each pixel. Detection is performed by checking feature vectors surrounding image pixels. An explicit model usually describes a vehicle as a box or wire-frame representation. Matching the model top-down to the image, or grouping extracted image features bottom-up to create structures similar to the model is performed on the image.

In this paper, we developed a scheme based on image segmentation and fuzzy logic classification to detect and classify road vehicles from airborne color digital orthoimagery at a ground pixel resolution of 20 cm. The basic difference, especially when compared with previously developed pixel-based vehicle detection procedures, is that we don’t process and analyze image pixels to perform vehicle detection and classification, but rather image objects or segments that are extracted from image segmentation.

2. Scheme for Road Vehicle Detection

In comparison to a single pixel, an image object offers substantially more information. Features for classification are computed based on image objects, not on single pixels. Beyond purely gray-level information, image objects contain a lot of additional attributes which can be used for classification: shape, texture and a whole set of relational/contextual information (semantic information). Utilizing not only image object attributes, but also the relationship between networked image objects, results in sophisticated classification incorporating local context. Moreover, the object-oriented method which first extracts homogeneous regions and then classifies them avoids the annoying salt-and-pepper effect of the more or less spatially finely distributed classification results that are typical of pixel-based analysis [10].

Fig. 1 shows the proposed scheme for road vehicle detection. Firstly, a vector-generated road mask was used to constrain detection of vehicles to road region. Secondly, image segmentation algorithm was performed to form image objects in the preprocessing orthoimagery. Finally, based on a set of fuzzy logic rules defined by membership functions, vehicle objects were detected and separated from other objects by fuzzy classification system.

For the preprocessing, road mask was generated from road centerline vector layer by buffer operation. Through road mask, the roadways are isolated from the whole image scene. The generated roadway region image for road vehicle detection is consistent with a large body of previous work using context-supported approaches [2-4]. The following steps were performed just on region of roadway. Fig. 2 shows a representative road segment image. Generally, roads mostly appear as multiple parallel stripes along the road direction. Pavement material changes, electrical poles and shadows occur in road segment of Fig. 2.

3. Segmentation Method

The image segmentation approach used in this study follows the Multiresolution Segmentation algorithm given in Baatz and Schape [10-12]. Multiresolution segmentation is a bottom up region-merging technique starting with one-pixel segments (objects). In numerous subsequent steps, smaller image segments are merged into bigger ones. Segments are located as potential candidates for merger by identifying the neighbouring segment with the smallest difference in heterogeneity for each individual segment. In each step, that pair of adjacent image segments is merged together when increases in heterogeneity formed by the merger are less than a unitless user-defined threshold of spectral and shape
heterogeneity called the scale parameter. If the smallest growth of the defined heterogeneity exceeds the threshold defined by the scale parameter, the process stops. Doing so, multiresolution segmentation is a local optimization procedure. As the scale parameter is increased, the size of segments correspondingly increases. The contiguous and homogeneous image segments act as image object primitives or segments. As such, these segments should be meaningful and ideally outline those objects in the image that are to be extracted.

The heterogeneity $f$ is computed based on the weighting of both spectral heterogeneity $x$ and shape heterogeneity $y$ as follows:

$$f = w_1 \cdot x + (1 - w_1) \cdot y,$$  \hspace{1cm} (1)

where $w_1$ is the user defined weight between 0 and 1 for spectral or color (against shape). The influence of spectral vs. shape heterogeneity is determined by weighting these parameters in ratio to each other; the higher the shape criterion the less that spectral heterogeneity influences the segmentation. If the user wants to place greater emphasis on the spectral characteristics of the desired image objects, then $w_1$ is greater than 0.5, and similarly if the spatial characteristics are considered more important, then $w_1$ is less than 0.5.

The spectral or color heterogeneity $x$ of an image object is calculated as the sum of the standard deviations $\sigma_i$ of spectral values of each multispectral channel weighted with the weights $p_i$ for each layer as follows:

$$x = \sum_{i=1}^{n} p_i \cdot \sigma_i.$$  \hspace{1cm} (2)

The spectral variance of each input channel in the image is used to measure spectral heterogeneity, and the separate channels themselves may be weighted according to the contribution of each channel to the overall variance of the image.

The shape heterogeneity $y$ of an image object is computed based on the weighting of both compactness $u$ and smoothness $v$ as follows:

$$y = w_2 \cdot u + (1 - w_2) \cdot v,$$  \hspace{1cm} (3)

where $w_2$ is the user defined weight between 0 and 1 for compactness. The compactness $u$ is described by the ratio of the de facto border length $E$ and the square root of the number of pixels $N$ forming this image object as follows:

$$u = \frac{E}{\sqrt{N}}.$$  \hspace{1cm} (4)

The smoothness $v$ is the ratio of the de facto border length $E$ and the shortest possible border length $L$ given by the bounding box of an image object parallel to the raster as follows:

$$v = \frac{E}{L}.$$  \hspace{1cm} (5)

In order to determine the outcome of the segmentation algorithm, the user can define several parameters, like the scale parameter, the single layer weights and the mixing of the heterogeneity criterion concerning tone and shape. A comparison of a fusion value with the scale parameter defines the break-up criterion. As mentioned above, the scale parameter is a measure for the maximum change in heterogeneity that may occur when merging two image objects. Internally, this value is squared and serves as the threshold, which terminates the segmentation algorithm. When a possible merge of a pair of image objects is examined, a fusion value between those two objects is calculated and compared to the squared scale parameter. During the segmentation process all generated image objects are linked to each other automatically.

When merging two neighboring image segments (object 1 and object 2), the overall fusion value $f'$ of merged object is computed based on the spectral heterogeneity $x'$ and the shape heterogeneity $y'$ as follows:

$$f' = w_1 \cdot x' + (1 - w_1) \cdot y',$$  \hspace{1cm} (6)

where $w_1$ is the user defined weight between 0 and 1 for spectral or color (against shape). The spectral heterogeneity $x'$ and the shape heterogeneity $y'$ for the merged object are computed as follows:

$$x' = \sum_{i=1}^{n} p_i \cdot (N' \cdot \sigma_i' - (N_1 \cdot \sigma_i + N_2 \cdot \sigma_i^2)),$$

$$y' = w_2 \cdot u' + (1 - w_2) \cdot v',$$  \hspace{1cm} (7)

where $w_2$ is the user defined weight between 0 and 1 for compactness. $N'$, $N_1$, and $N_2$ are the number of pixels forming the bigger merged image object, smaller neighboring objects 1 and 2 respectively. Similarly, $\sigma_i'$, $\sigma_i^1$ and $\sigma_i^2$ represent standard deviations of spectral values of each multispectral channel weighted with the weight $p_i$ for each layer.

The compactness $u'$ and the smoothness $v'$ for the merged object are computed as follows:

$$u' = N' \cdot \left( \frac{E'}{\sqrt{N'}} - \frac{E_1}{\sqrt{N_1}} - \frac{E_2}{\sqrt{N_2}} \right),$$

$$v' = N' \cdot \left( \frac{E'}{L'} - \frac{E_1}{L_1} - \frac{E_2}{L_2} \right).$$  \hspace{1cm} (8)
where $E_1', E_1$, and $E_2$ are the de facto border lengths forming the bigger merged image object, smaller neighboring objects 1 and 2 respectively. Similarly, $L_1', L_1$, and $L_2$ represent the shortest possible border lengths (given by the bounding box of an image object parallel to the raster) forming the bigger merged image object, smaller neighboring objects 1 and 2 respectively.

As the result of segmentation, meaningful object primitives or segments are created. They are unclassified basic image objects. All together, the image objects of a segmentation procedure form an image object level. Two or more image object levels build the image object hierarchy (Fig. 3). Embedded in this hierarchical network of image objects, each object knows its adjacent objects, its sub-objects and its super-object [12, 13]. This hierarchical network of semantic object information can be implemented by two strategies: the bottom-up and top-down approaches. In the first case, small segments are generated using relatively small-scale parameters. Objects generated at higher levels retain the semantic information of the sub-objects. In the top-down approach, large and coarse segments are generated using a relatively large-scale parameter. Objects generated at lower levels are determined by the outlines of the super-objects. Both strategies may be employed in the same segmentation routine. In the image object hierarchy, every image object of a lower level is linked to image objects of its super-level. Two trivial image object levels are the partition of the image into pixels (the pixel level) and the level with only one object covering the entire image [10].

![Fig. 3. Image object hierarchy (Definiens, 2006).](image)

Image object primitives (segments) contain information about their spectral characteristics, their shapes, their positions and textures, as well as information about their neighborhoods. Thus, they are an essential prerequisite for the subsequent image object classification. They serve as both information carrier and building blocks at the starting point of the image analysis. As building blocks, they can be merged, cut, or classified. As the image analysis progresses, image object primitives commonly are merged to build larger image objects. Ideally, they do not need to be split anymore into smaller segments.

Multiresolution segmentation is one basic procedure for object oriented remote sensing image analysis [10, 11]. It is used here to produce image object primitives for a further classification and other processing procedures. For the vehicle detection in the study, we performed the segmentation procedure twice using different parameters. Two image object levels were generated after segmentations. One was road strip level generated from a large-scale segmentation parameter. Another was vehicle detection base level from small-scale segmentation parameter. Fig. 4 shows the road strip level from large-scale segmentation on Fig. 2. Fig. 5 shows the vehicle detection base level from small-scale segmentation on Fig. 2.

![Fig. 4. Road strip level from super-scale segmentation.](image)

![Fig. 5. Vehicle detection base level.](image)

With the generated multi-level image object hierarchy, we can utilize not only image object attributes, but also the relationship between networked image objects in sophisticated classification. Each of the image objects created knows its neighbors as well as its super- and sub-objects along with all their attributes. From previous work in this area and our own experiments, sufficient contrast at least locally is needed for successfully detecting a vehicle from the surrounding pavement. Here, we attempt to exploit the gray-level contrast between the local pavement strip and the embedded possible vehicle objects.
3. Fuzzy Logic Classification

Various types of uncertainty influence information extraction from remote sensing data. Especially important context information is typically only expressed in terms of vague linguistic rules. Fuzzy logic is a mathematical approach to quantifying uncertain statements. The basic idea is to replace the two strictly logical statements “yes” and “no” by the continuous range of [0, 1], where 0 means “exactly no” and 1 means “exactly yes.” All values between 0 and 1 represent a more or less certain state of “yes” and “no.” Thus, fuzzy logic is able to emulate human thinking and take into account even linguistic rules [14-15]. A fuzzy rule base allows the formulation of expert knowledge as well as relations of classes in a very efficient way in class descriptions.

Fuzzy classification systems are well suited to handle most vagueness in remote sensing information extraction. Fuzzy classification was chosen for the analysis of image objects in the study because: 1) by translating feature values into fuzzy values, it standardizes features and allows the combination of features, even of very different range and dimension, 2) it provides a transparent and adaptable feature description, especially compared to neural networks, 3) it enables the formulation of complex feature descriptions by means of logical operations, and allows integration of knowledge close to human thinking. Fuzzy systems consist of three main steps, fuzzification, fuzzy rule base and defuzzification [15].

Fuzzification describes the transition from a crisp system to a fuzzy system. It assigns a membership degree (membership value) between 0 and 1 to each feature value. Therefore, all input values for fuzzy combinations are in the range between 0 and 1, independent of the dynamic of the originally crisp features. The membership value is defined by a so-called membership function. Membership functions are graphic representations of the proximity of an object’s metric for a given feature to a modeled value in a curved or stepped transition between no membership (0) and full membership (1), vice versa, or both. In general, the broader the membership function, the vaguer the underlying concept; the lower the membership values, the more uncertain is the assignment of the set. For successful classification a deliberate choice of membership function is crucial. This allows the introduction of expert knowledge into the system. The better the knowledge about the real system is modeled by the membership functions, the better the final classification result.

A fuzzy rule base is a combination of fuzzy rules, which combine different fuzzy sets. The set of feature values, which allow its members to have different grades of membership (membership values) in the interval [0, 1], can be called a fuzzy set. Parameter and model uncertainties are considered as using fuzzy sets defined by membership functions. The combining of different features (fuzzy sets) within a fuzzy system is always done after the feature is fuzzified. Fuzzy rules are merely a series of if-then statements. To create advanced fuzzy rules, fuzzy sets can be combined by operators, which return a fuzzy value. The basic operators are “AND” and “OR.” “AND” represents the minimum, meaning that the minimum value of all sets defines the return value. “OR” represents the maximum value, meaning that the maximum value of all sets defines the return value. A fuzzy rule base delivers a fuzzy classification, which consists of discrete return values for each of the considered output classes. These values represent the degree of class assignment. The higher the return values for the most possible class, the more reliable the assignment. Technically, a classification with the use of a fuzzy rule base is done by finding out which combination of fuzzy features is suitable to distinguish one class from the others.

The fuzzy classification must be defuzzified to produce a desired crisp classification assigning a single class label to each image object. Defuzzification is performed using the max operator so that each object is classified as the class with the highest membership value.

In the study, once a satisfactory segmentation image for vehicle detection base level is obtained (Fig. 5), it is possible to apply a fuzzy logic rule base to discriminate between Vehicle and other Non-Vehicle objects. The goal of the rule base is to represent the uncertain nature of the properties and relationships that characterize Vehicle and Non-Vehicle object classes in membership functions that permit class likelihood to be described. The degree of likelihood to which an object may be assigned membership in a class is evaluated by comparison of fuzzy values for each image object returned by a rule base that quantifiably describes the contribution of specific object features, or sets of features, to an overall likelihood of the object’s membership in each class. Class membership for given object features are assigned as scores in which the features’ metrics are modeled in membership functions. An increasing function is used when membership is assumed to increase with increasing object feature values. Conversely, when membership is assumed to decrease with increasing object feature values, a decreasing function is used. The transitions in membership functions are governed by assigned control points. Fig. 6 shows the increasing and decreasing Sigmoidal (S-shaped) membership functions, where a, b and c are the control points of the functions. Membership degree µ in class x is calculated from an increasing or decreasing sigmoidal function (S or 1-S), where a is the object feature value for attribute i at which membership is nil or perfect, b is the object feature value at which membership is 0.5, and c is the object feature value at which membership is perfect or nil, inverse to a.

Because object feature values for spectral, spatial, textural and contextual attributes form the independent variables, or x-axes, of both sigmoidal
membership functions and frequency distribution histograms of features, membership in a class itself can be taken to be analogous to frequency for given class attributes.

\[ \mu_s = S(i; a, b, c) \]

\[ \mu_s = 1 - S(i; a, b, c) \]

Fig. 6. Sigmoid (S-shaped) membership functions.

It is therefore possible to assume that distributions of object feature values for independent class attributes that are representative of the class as a whole will be marked by an approximately normal distribution where the accumulation of values within one standard deviation about the mean for each class do not overlap each other. For the Vehicle class in this study, where mean feature values in the modeled class less one standard deviation are greater than mean feature values plus one standard deviation in the other class, the membership function for feature \( i \) is modeled as an increasing sigmoidal curve:

\[ \mu_s = S(i; a, b, c), \quad (9) \]

where \( \mu_s \) is the membership score in the modeled class for feature \( i \) between 0 (no membership) and 1 (full membership), and \( a \) is the mean value minus one standard deviation in the modeled class, \( b \) is the mean value minus one-half standard deviation, and \( c \) is the mean value.

Conversely, where mean feature values in the modeled class plus one standard deviation are less than mean feature values plus one standard deviation in the other class, the membership function for feature \( i \) is modeled as a decreasing sigmoidal curve:

\[ \mu_s = 1 - S(i; a, b, c), \quad (10) \]

where \( a \) is the mean value for the object feature \( i \), \( b \) is the mean value plus one-half standard deviation, and \( c \) is the mean value plus one standard deviation.

Membership scores for the modeled class in each rule \( R \) are combined using an “AND” logical operator:

\[ \mu_R = \min\{\mu_{f_1}, \ldots, \mu_{f_j}\}, \quad (11) \]

for \( f_i, i=1,\ldots,j \) features, or a “OR” logical operator:

\[ \mu_R = \max\{\mu_{f_1}, \ldots, \mu_{f_j}\}, \quad (12) \]

And the membership score for the Non-vehicle class is provided as:

\[ \mu_{NR} = 1 - \mu_R, \quad (13) \]

4. Experiments and Results

Segmentation will result in detection of different types of objects, where the objects corresponding to vehicles need to be identified. For this, we will use a fuzzy logic classification approach, where specific characteristics (features) of the objects are first extracted from each object before the objects are classified based on these features. Descriptions of the spectral, spatial, textural and contextual features as calculated for objects are given in related publications [10-13].

In our experiment, classification is performed in two stages. First a hierarchical, rule-based fuzzy logic classifier is constructed to classify the extracted objects into the possible vehicle objects and the non-vehicle objects on vehicle detection base level. The resulting potential vehicles are then further analyzed through a final step of statistical classification.

For the rule-based fuzzy logic classifier, a preliminary rule composed of membership functions for grey features of the Vehicle class describing properties related to the pixel values of the objects as a whole is used to identify likely Vehicle objects. However, the flat or polymodal distributions of object values for these features in the Non-Vehicle class will result in overestimation of the Vehicle class. For example, the non-vehicle objects may represent road markings, shadows, reflections or parts of houses or other structures along the road. Hence, it is difficult to model this class of clutter objects in a meaningful way. A second rule incorporating independent and representative features of the Non-Vehicle class as a whole using spatial features describing the shape of the objects is then used to identify likely Non-Vehicle objects from the Vehicle class to reduce the overestimation of the Vehicle class. A third rule using relational features between the super-scale object and the present detection level object is then used to further refine the elimination of Non-Vehicle objects from the preliminary Vehicle class. A crisp classification is performed by applying
a final rule that delivers to each object a class label assigned by the highest membership value as calculated from application of the first three rules. These rules with the spectral, spatial, and relational features in order construct the fuzzy logic rule base for vehicle detection.

After the first classification, the possible vehicle objects have been detected from the road segment. However, at this stage, a whole vehicle often is made up of more than one neighboring vehicle objects. Particularly, the big vehicles have divided into two obviously parted vehicle objects where segmentation has failed and resulted in fragmented or connected objects. To form just only one corresponding vehicle object for every whole vehicle, we create a new vehicle detection fusion level by merging the neighboring vehicle objects based on the classified vehicle detection base level. So, a three-level image object hierarchy after multi-scale segmentations and fusion is constructed, that is, road strip level, base level of road vehicle object, and fusion level for road vehicle extraction. For the road strip level, road strips are extracted from scene image which act as super-objects of road vehicle objects in the base level. For the base level of road vehicle object, a fuzzy logic classifier is constructed to divide the extracted object regions into vehicle and non-vehicle classes by using the object-contained and super-object related features. Fusing the same neighboring classes on the base level of road vehicle object form the fusion level for road vehicle extraction. Fig. 7 shows the vehicle detection fusion level based on Fig. 5. On the fusion level, every whole vehicle has just only one corresponding object except for big vehicles.

To detect and count vehicles accurately, we should further classify this fusion level. Features that are used in this stage include different geometric features like length, width, area, elongation, rectangularity, compactness and angle deviation to extract vehicle objects. For big vehicles, some have two parted vehicle objects. One corresponds to the big back part; another corresponds to the small front part. We can use the feature of distance to neighboring object along the main direction to recognize the small vehicle object corresponding to the small front part. Here, we consider the bigger vehicle object corresponding to the big back part as the final detected big vehicle object. Fig. 8 shows the detected and labeled center points of final detected vehicle-objects. We can see that all of the vehicles in the road segment have been detected.

Finally, based on the calculated average length and width of detected vehicle objects on the fusion level, we classify vehicles into three categories, that is, small, medium and big. Fig. 9 shows vehicle classification result. Red objects represent big vehicles, green objects represent medium vehicles, and blue objects represent small vehicles.

Fig. 8. Detected and labeled vehicle points.

Fig. 9. Vehicle classification result.
three vehicle classes are easily derived. The extracted vehicle images are compared with the manually labeled vehicle images. The automatic counts match manual counts very well. It reaches 100% of detection and classification correctness for the road segment illustrated above.

Fig. 10. Frame-labeled vehicles after detection and classification.

The experiments show that the method gives a very good performance even when the vehicles have a low contrast with the road surface or vehicles are close to each other from examination of the detection results. To further validate the vehicle detection approach, relatively complex sample road segments are selected to cover a wide variety of the situations normally encountered. Fig. 11 shows the detection result images of another three road segments. A representative set of 12 road segment images is selected from available images to test the proposed scheme. Experimental results indicate that the detection rates of all test road-segments are high with very few false alarms. Table 1 gives the statistical of detection results for the 12 selected road segments (% of true counts).

Fig. 11. The detection result images of three test road segments.

Table 1. Statistical of detection results for the 12 selected road segments (% of true counts)

<table>
<thead>
<tr>
<th>Road Segments</th>
<th>Detected Vehicles</th>
<th>Omission Errors</th>
<th>Commission Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>97.6</td>
<td>2.4</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>102</td>
<td>0</td>
<td>2.3</td>
</tr>
<tr>
<td>5</td>
<td>93.7</td>
<td>9.4</td>
<td>3.1</td>
</tr>
<tr>
<td>6</td>
<td>106.4</td>
<td>0</td>
<td>6.4</td>
</tr>
<tr>
<td>7</td>
<td>97.6</td>
<td>2.4</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>3.1</td>
<td>3.1</td>
</tr>
<tr>
<td>9</td>
<td>97.6</td>
<td>2.4</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>93.7</td>
<td>6.2</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>105.6</td>
<td>0</td>
<td>5.5</td>
</tr>
<tr>
<td>12</td>
<td>102.8</td>
<td>1.4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

5. Conclusions and Discussion

Experimental results indicate that the proposed scheme has a good performance under varying conditions of road geometry, vehicle contrast, variability of pavement characteristics, and vehicle density. The detection rates of all test road-segments are high with very few false alarms. Comparing with pixel-based vehicle detection approaches, the proposed method directly detects vehicles in single image and background image is not needed. In our experiments, morphological or post-processing operations are not necessary, which are usually performed on the detected image to sieve out spurious responses and cluster appropriate pixel groups into potential vehicles and finally derives the vehicle-only image. The scheme may find utility in estimation for vehicle counts. And further investigation and improvement could include improving automation by incorporating automated road extraction techniques before applying vehicle detection algorithms.

Acknowledgements

The work was supported by the National Natural Science Foundation of China (Grant No. 40801121, 51078020) and the Fundamental Research Funds for the Central Universities (2012JBM078).

References

[3]. X. Jin, C. H. Davis, Vehicle detection from high-resolution satellite imagery using morphological


