VEHICLE EXTRACTION AND SPEED DETECTION FROM DIGITAL AERIAL IMAGES

Fumio Yamazaki[1)], *Wen Liu*[1)], *T. Thuy Vu*[2)]

[1)] Graduate School of Engineering, Chiba University, Japan.
[2)] Geoinformatics, Royal Institute of Technology, Sweden.

ABSTRACT

A new object-based method is developed to extract the moving vehicles and subsequently detect their speeds from two consecutive digital aerial images automatically. Several parameters of gray values and sizes are examined to classify the objects in the image. The vehicles and their associated shadows can be discriminated by removing big objects such as roads. To detect the speed, firstly the vehicles and shadows are extracted from the two images. The corresponding vehicles from these images are linked based on the order, size, and their distance within a threshold. Finally, using the distance between the corresponding vehicles and the time lag between the two images, the moving speed can be detected. Our test shows a promising result of detecting the moving vehicles' speeds. Further development will employ the proposed method for a pair of QuickBird panchromatic and multi-spectral images, which are at a coarser spatial resolution.

Index Terms**—** Digital aerial image, vehicle extraction, speed detection, object-based method

1. INTRODUCTION

As the population in cities continually increases, road traffic becomes more congested than the level which city and infrastructure planning expects. To solve the problem, at first monitoring vehicles must be the important task. Normally, the field-based equipment, like cameras installed at fixed locations or weigh-in motion sensors on the pavements, are used to monitor traffic. Presently remote sensing technique has been emerged as another option to collect the traffic information. Using this way, the wider range of information can be collected over a long time. Thus, vehicle detection by remote sensing can be extensively used to manage traffic, assess fuel demand, estimate automobile emission and also important for planning transport infrastructure.

There have been several researches on vehicle detection using remote sensing data. They can be categorized into two groups: model-based extraction and data-based extraction. Model-based extraction is based on the vehicle models built from a learning process. The models then are used to identify whether the target is a vehicle or not. For example, Gerhardinger et al. [1] tested an automated vehicle extraction approach based on an inductive learning technique, which was implemented using *Features Analyst*, an add-in extension of ArcGIS software. Zhao and Nevatia [5] combined the multiple features of vehicles in a Bayesians network for leaning prior to detecting vehicles.

In data-based extraction, the processing follows a bottom-up procedure to group pixels into objects and the vehicle objects are subsequently discriminated from the others. Hong and Zhang [2] used an object-oriented image processing technique to extract vehicles. A detailed description, which requires a large number of models to cover all types of vehicles, is the key of the former approach. It takes time and cannot be widely applied. The latter is simpler and convenient to be widely used. Also, those recent researches mainly reported on the vehicle position, and few of them go further to speed detection as well as a traffic information database.

In this research, a new method is developed for both vehicle extraction and speed detection. The pixels in an image are firstly formed into objects based on their gray values. The objects are then classified and extracted based on several gray value and size indices. The extracted vehicles and shadows are recorded on the traffic information. Then the vehicle's speed is detected by matching the result of vehicle extraction from two consecutive images. The proposed approach is tested on digital aerial images.

2. DATA AND STUDY FLOW

2.1. Study area and data used

The study area is located in Minato-ku, a central part of Tokyo, Japan. Two pairs of the consecutive aerial images are used in this study. The images were taken by a digital aerial camera UltraCamD on August 4th, 2006, by Geographical Survey Institute of Japan.

The UltraCamD camera offers simultaneously sensing of high-resolution panchromatic channel (pixel size is 9μm) as well as lower-resolution RGB and NIR channels (pixel size is 28μm). It has the ability to capture images with higher overlap, up to 90%, in along track direction.

Fig. 1 Two consecutive pansharpened digital aerial images of Roppongi, Tokyo.
(GPS time: left 437379.7512 sec; right 437376.3633 sec)

A panchromatic image has 7,500×11,500 pixels and a multispectral image 2,400 ×3,680 pixels. A pair of images cover the area near Hamazakibashi Junction and another pair cover the area of Roppongi, Tokyo. Color images with resolution of about 0.12m/pixel, obtained through a pansharpening process, are used in this study. The two consecutive images have an overlap of about 80%, as shown in Fig. 1.

2.2. Flow of the study

A general overview of our approach, consisting of image preprocessing, vehicles extraction and speed detection, is presented in Fig. 2. Firstly, before using the digital images to detect vehicles and their speeds, image preprocessing is needed. Then, the pairs of two consecutive images are obtained. Secondly, an object-based method is used to extract vehicles and record the information on a traffic database. Then vehicles' speeds are detected by matching the same vehicles from the two databases with a time lag. Thirdly, we detect vehicles and their speeds visually as a reference data. Finally, we compare the automated vehicle and speed detection result with the visual one to discuss the accuracy of the proposed automated approach.

3. METHODOLOGY

3.1. Image preprocessing

To detect the speed of vehicles, two images, covering the same area with a time lag, are needed. Firstly, an overlap area from two consecutive images is extracted to obtain two images over the same area. Because of the perspective projection of the camera, geometric distortions between two images exist. Hence registration with the pairs of images using 8 ground control points was conducted. After registration, the two images in a pair have different pixel sizes. Finally, images are arranged in the same pixel sizes by image mosaicing.

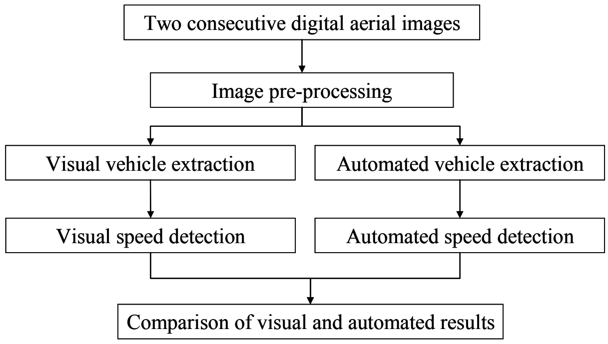

Fig. 2 Flow of the study

3.2. Automated vehicle extraction

Vehicles are extracted from the images and the information of their positions, sizes and brightness (gray value) is stored in a database, which can be used as traffic information. The proposed methodology consists of road extraction, noise removal, object-based vehicle extraction, and verification steps.

3.2.1. Road extraction

Since vehicles are moving on roads, road extraction should be the first processing step. Focusing on the extraction of vehicles and the detection of their speeds, we do not propose a new road extraction approach. There have been numerous researches on road extraction from remote sensing images (Quackenbush [3]). Those can be easily employed to extract the road objects here. Additionally, GIS road data can also be used to extract roads. However, to avoid errors involved in road extraction, which influences the final vehicle extraction results, the roads are extracted manually in this study. Then the areas out of the road areas are masked.

3.2.2. Noise removal (Morphotological filter)

Prior to carrying out vehicle extraction, other irrelevant information such as lines on the road surface should be removed. Concerning the shapes and sizes of the objects, area morphological filtering is chosen (Vu et al. [4]). This filter perfectly removes long and thin road lines and retains the shapes of vehicles. The window size used here is set as 5×5.

3.2.3. Object-based vehicle extraction (initial extraction)

Since vehicles extraction is based on the gray value, color images are converted to black and white images. The object-based vehicle extraction approach is shown in Fig. 3.

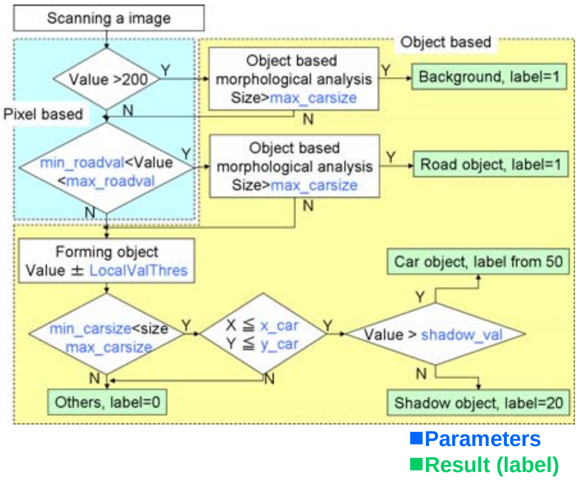

Fig. 3 Flowchart of the automated vehicle extraction

Pixels are scanned and grouped into objects according to the criteria of gray values. In this step, the image represents 4 kinds of objects: background, roads, vehicles (including their shadows) and the others treated as noise. The road extraction step assigns the background as black color. It can be easily discriminated by the lowest range of the gray value. Meanwhile, the road surfaces normally show another specific range of the gray value. Based on those two gray value ranges, the objects are formed. There might be vehicles which show very similar characteristics with the black background. Fortunately, the background and the road are often big objects compared to the others. Then, these two kinds of objects can easily be extracted based on a size threshold.

The remaining pixels are reformed into objects again based on a local threshold of gray value. The fact is that all the pixels belonging to a vehicle should have a similar gray value. Vehicles and their associated shadows generally have a specific range of size. It is the criteria to distinguish them from the others. Consequently, the initial extracted result is obtained, and the information on vehicle position and size is stored in a database.

3.3. Automated speed detection

Speed detection is based on the time lag between two images. Generally, it can be performed using two consecutive aerial images. The proposed vehicle extraction approach can be extended to speed detection. The vehicle and shadow database of each image is developed after the vehicle extraction process. Then, the vehicles in the two databases are linked by the order, moving direction, size and distance (Fig. 4). If a vehicle in the second image is in the range of possible distance from the one in the first image and if they have similar sizes, they are linked as the same vehicle. Subsequently, using their positions stored in the databases, the speed can be computed.

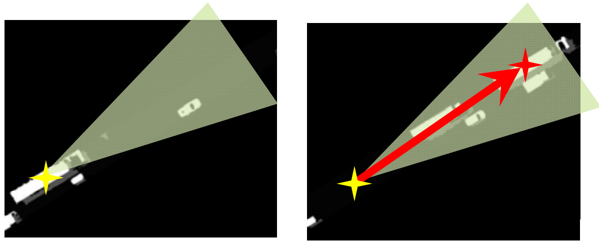

Fig. 4 Condition of matching the same vehicles

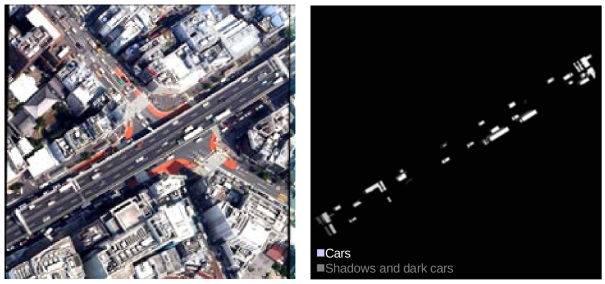

Fig. 5 Original image (left) and the result of vehicle extraction (right)

4. RESULT OF CASE STUDY

4.1. Result of vehicle extraction

The target of extraction is the vehicles on the expressway in the two study areas. As the result, the vehicles with light color are shown in white and the shadows or dark vehicles are extracted as gray (Fig. 5). Additionally, the information of vehicle positions and sizes is stored in a database for speed detection. Then the results are compared with visual extraction results (Table 1).

There were 292 vehicles in the pair images of Hamazakibashi area, and 282 vehicles were extracted correctly by the process of vehicle extraction. Only 10 vehicles were missed, and the noises which are not vehicles but extracted as vehicles were 116. The producer accuracy is 96.5%, and the user accuracy is 71%.

In the image pair of Roppogi, 195 vehicles and 191 vehicles, respectively, were extracted correctly. Four vehicles were missed, and noises were 42. The producer accuracy of these images is 98%, and the user accuracy is 82%.

Overall, almost all the vehicle could be extracted. Not only light-color vehicles but also dark vehicles and some vehicles in shadow were extracted successfully. Because we extracted both vehicles and shadows, even the vehicle’s gray value is similar to that of the road, the vehicle can be extracted by its associated shadow. There still have a few commission errors due to a signboard, its shadow, and some lines on the road. The environment condition around the target area influences the result of vehicle extraction. Accuracy gets higher as an environment becomes simpler.

Table 1 Accuracy of vehicle extraction

Image	Hamazakibashi	Roppogi
Vehicles in image	292	195
Extract result	398	233
Correctly extracted	282	191
Omission	10	4
Commission	116	42
Producer accuracy	97%	98%
User accuracy	71%	82%

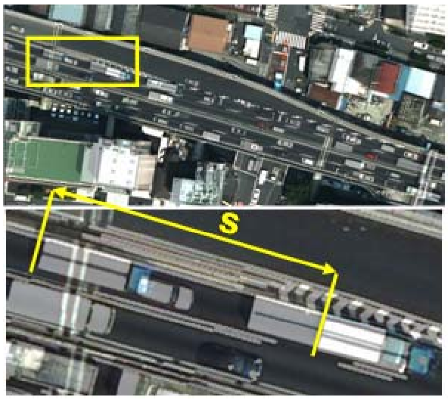

Fig. 6 Method of visual speed detection

4.3. Result of speed detection

Visual speed detection was first carried out to obtain reference data by overlapping the second image to the first one. The speed can be detected by measuring the difference of vehicle's outline as shown in Fig. 6.

Then the speed and moving direction of vehicles were detected automatically by matching the databases for two consecutive images using the parameters of order, direction, size and distance (Fig. 7). About 71% of vehicles' speeds were detected automatically from Roppogi area. The standard deviation for the difference of speed between automated and visual detections is 0.83km/h, and the standard deviation for the difference of direction is 0.38 degrees.

In Hamazakibashi area, only a part of images were used for speed detection since the accuracy of vehicle extraction was low (71%). Vehicle matching depends on the order of vehicles, and matching error occurs when many noises influencing the order of vehicle exist. From the part of images, 64% of vehicles' speed were extracted. The standard deviation for the difference of speed between automated and visual detections is 1.01km/h, and the standard deviation for the difference of direction is 0.59 degrees.

Since the rules for vehicle matching are very severe, not all the vehicles could be matched from the image pairs. The order changes by noises, and the size changes in vehicle extraction are the main reasons for matching error. But the result of speed detection for the matched vehicles showed high accuracy.

Fig. 7 Result of automated speed detection

5. CONCLUSIONS

A new automated approach for vehicle extraction and speed detection was proposed using a pair of digital aerial images with a short time lag. In vehicle extraction, environmental conditions such as road lines, surrounding trees, and road sign boards were found to have significant influence on the accuracy. As the result of the case study for two areas in Tokyo, most vehicles were extracted successfully. The speeds of vehicles were also extracted from two consecutive images with good accuracy. However, the performance of vehicle matching depends on the accuracy of vehicle extraction.

In a further research, we hope to extend the proposed method to extract vehicles from satellite images, likes QuickBird, which has much lower resolution than that of aerial images.

The authors express their gratitude to Geographical Survey Institute of Japan for providing the digital aerial images used in this study.

REFERENCE

[1] Gerhardinger, A., D. Ehrlich, and M. Peseresi, 2005. Vehicles detection from very high resolution satellite imagery. CMRT05. IAPRS, Vol. XXXVI, Part 3/W24.

[2] Hong, G., Zhang, Y. and Lavigne, D.A., 2006. Comparison of car extraction techniques for high resolution airborne images. First Workshop of the EARSeL Special Interest Group on Urban Remote Sensing.

[3] Quackenbush, L. J., 2004. A review of techniques for extracting linear features from imagery. Photogrammetric Engineering & Remote Sensing Vol. 70, No. 12, pp. 1383-1392.

[4] Vu, T.T., M.Matsuoka and F. Yamazaki, 2005. Preliminary results in development of an object-based image analysis method for earthquake damage assessment. Proc. of 3rd International workshop Remote Sensing for Post-Disaster Response, Chiba, Japan.

[5] Zhao, T., and Nevatia, R., 2001. Car detection in low resolution aerial image. In: International Conference on Computer Vision.