2009 Urban Remote Sensing Joint Event

Speed detection of moving vehicles from one scene of QuickBird images

Wen Liu, Fumio Yamazaki
Graduate School of Engineering
Chiba University
Chiba, Japan
nanahachimizu@yahoo.co.jp, yamazaki@tu.chiba-u.ac.jp

***Abstract*— A new method is developed to extract moving vehicles and subsequently detect their speeds from a pair of QuickBird (QB) panchromatic (PAN) and multi-spectral (MS) images automatically. Since PAN and MS sensors of QB have a slight time lag, about 0.2 seconds, the speed of moving vehicles can be detected by the movement between PAN and MS images in the time lag. From a PAN image with 0.6m resolution, vehicles can be extracted by an object-based approach. But it is difficult to extract the accurate position of vehicles from a MS image with 2.4m resolution. Thus an area correlation method is proposed to estimate the location of vehicles from MS images in a sub-pixel level. Using the results of the vehicle extraction, the speed of moving vehicles can be detected. The approach is tested on several parts of the QB image covering the central Tokyo, Japan, and the accuracy of the result is demonstrated.**

I. INTRODUCTION

As the population in cities continually increases, road traffic becomes more congested than the level which city and infrastructure planning has expected. Monitoring vehicles is an important first step to solve the problem. Current technology collects traffic data using field-based equipment, like cameras installed at fixed locations or weigh-in motion sensors on pavement. These equipments provide traffic flow information over time for a point in space. Presently remote sensing technique emerged as another option to collect traffic information. Using this method, a 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 transportation infrastructure planning.

It is known that the panchromatic (PAN) and multi-spectral (MS) sensors of QuickBird have a slight time lag, about 0.2 seconds, depending on the scanning mode of the instrument. Using this time lag between the two sensors of one scene, the speed of moving objects can be detected manually ([1], [2]). To detect speed of moving vehicles automatically, firstly vehicles must be extracted from both PAN and MS images. From a PAN image or an aerial image with high resolution, several researchers have conducted vehicle extraction. Their approaches 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. Then the models are used to identify whether a target is a vehicle or not ([3], [4], [5]). In data-based extraction, the processing follows a bottom-up procedure to group pixels into objects and vehicle objects are subsequently discriminated from the others ([6], [7], [8], [9]). A detailed description, which requires a large number of models to cover all the types of vehicles, is the key for the former approach. It takes time and cannot be widely used. Those recent researches mainly used aerial imagery or high resolution satellite images, e.g. PAN band of QuickBird (QB) or Ikonos, and a few of them extracted vehicles from MS images and their moving speeds.

In this research, a new method is developed to extract moving vehicles and subsequently detect their speeds from a QB's PAN and MS bundle product automatically. From a high resolution (0.6m) PAN image, vehicles can be extracted by an object-based method [8]. Since the resolution of QB's MS image is not high enough (2.4m) to extract vehicles directly, an area correlation approach is introduced to search the best match position with the result of vehicle extraction from a PAN image. The proposed approach is tested on an actual QB and simulated QB images. Using the results of vehicle extraction, the speed of moving vehicles can be obtained.

II. VEHICLE DETECTION FROM PAN IMAGE

There have been several researches on vehicle detection from high resolution satellite images ([3], [5], [9]). In this study, the object-based method [8] is used to extract vehicles from QuickBird's PAN images. The object-based method can extract vehicles from aerial imagery using several local parameters. Since the object-based method follows a bottom-up procedure to extract vehicle objects, it can be used for various images. The approach can not only extract vehicles but also build a database of position and size information. Thus, this object-based approach is tested in this paper to extract vehicles from a QuickBird's PAN image.

A. Study area and data used

The study area is located in the central part of Tokyo, Japan. Three parts of a QB's PAN image are used in this study. The QuickBird image was taken on March 20, 2007 (Fig. 1). The targets are moving vehicles on urban expressways.

978-1-4244-3461-9/09/$25.00

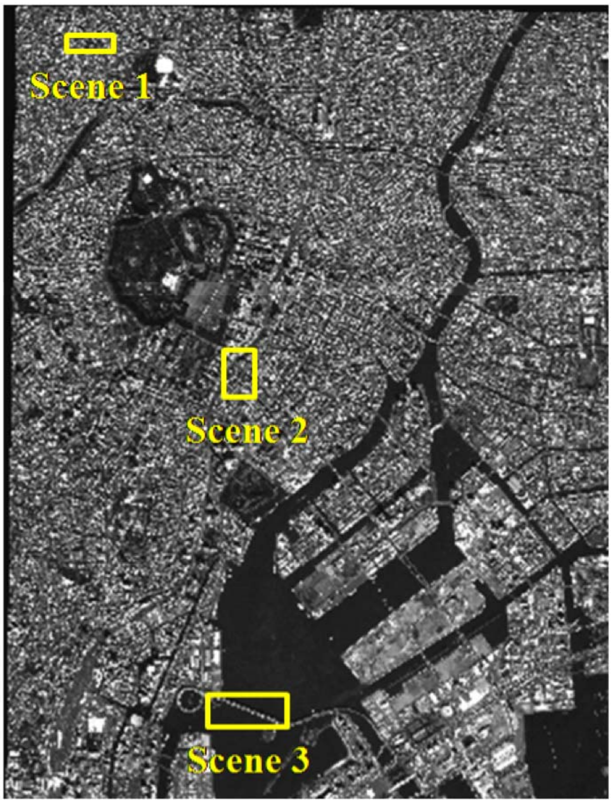

Figure 1. QuickBird' PAN image of the central Tokyo, Japan

B. *Experiment and result*

The vehicles in the three parts of the QuickBird's Pan image with 0.6m resolution, are extracted by the object-based method. First, the PAN image is transformed to 0.24m/pixel by cubic convolution. Then roads are extracted manually to avoid errors involved in road extraction. An area morphological filter [10] is used to remove other irrelevant information such as lines on the road surface. Since the pixel size is 0.24m, a similar resolution with airborne images, the window size of the filter is set as 5×5. Pixels are scanned and grouped into objects according to the criteria of the gray value. The value parameters are defined to estimate roads and background. Several size parameters are defined to find vehicles from objects. As the result, the vehicles with light color are shown in white and shadows or dark vehicles are extracted as gray (Fig. 2). Additionally, the information of vehicles' positions and sizes is stored in a database for vehicle extraction from the corresponding MS image and speed detection.

The results of vehicle extraction are compared with visual inspection results in Table 1. There are 159 vehicles in the PAN image, and 149 vehicles were extracted correctly by the object-based vehicle extraction. Only 10 vehicles were missed, and the noises which are not vehicles but extracted as vehicles were 18. The producer accuracy is obtained as 94%, and the user accuracy is 89%.

Almost all the vehicle could be extracted from the PAN image with 0.6m resolution. The object-based method can extract not only light-color vehicles but also dark ones and some in shadow. A few commission errors still occurred due to a signboard, its shadow, and some lines on the road.

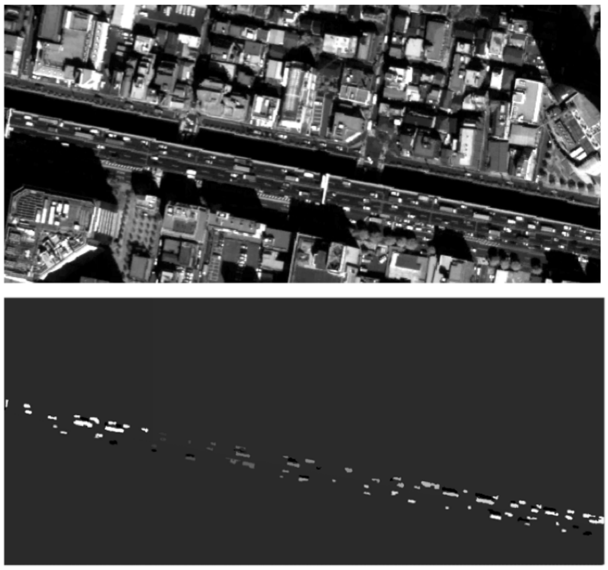

Figure 2. A part of the PAN image (up) and the result of vehicle detection (down) from scene 1 in Fig. 1

TABLE I. ACCURACY OF OBJECT-BASED VEHICLE EXTRACTION

	Vehicles in image	Extraction result			Accuracy	
		Correct	Omission	Commission	Producer	User
Scene 1	96	92	4	2	96%	98%
Scene 2	56	50	6	14	89%	78%
Scene 3	7	7	0	2	100%	78%

III. VEHICLE EXRACTION FROM MS IMAGE

From the PAN image with 0.6m resolution, vehicles can be extracted by the object-based approach. But the resolution of the corresponding MS image is about 2.4m, one vehicle appears in about only 1 or 2 pixels. Most vehicle pixels are mixed with road, and it is difficult to extract the accurate edge and position of vehicles. The object-based approach cannot extract vehicles from the MS image. To detect the speed of vehicles, the shift of the vehicle locations in a pair of PAN and MS images is needed. Thus, an area correlation method is proposed to estimate the location of vehicles from a MS image in a sub-pixel level.

A. *Methodology*

Area correlation is a method for designating Ground Control Points (GCPs) in image-to-image registration [11]. A small window of pixels (template) from the image to be registered is statistically compared within a region of the reference image, which is bigger than the template image. The template of M rows by N columns is selected as shown in Fig. 3. A bigger size window is selected for the reference image. The template is overlaid on the reference image and a

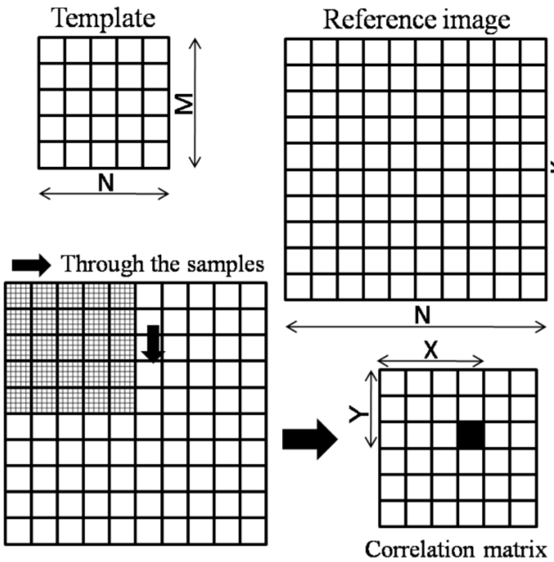

Figure 3. Shifting a template and overlaying a reference image, the correlation matrix is calculated for earch shift.

similarity index is calculated. This procedure is carried for every possible position in the reference image and the similarity matrix is obtained. The content of the similarity matrix is the value of the statistical comparison between the template and the reference image. The position in the similarity matrix where the similarity index is the maximum represents the necessary offset that the template has to be moved horizontally and vertically to match the reference image. This process is shown in Fig. 3. Note that if there is a relative rotation between the template and the reference image, a rotational angle should be introduced for matching.

One of the similarity indexes is the cross-correlation coefficient between the template and the reference image (1). The cross-correlation coefficient is a scalar quantity in the interval [-1.0, 1.0]. The cross-correlation coefficient can be used as a similarity index since it gives a measure of the degree of correspondence between the reference and the template or can be seen as a coefficient that is a direct measure of how well two sample populations vary jointly [12].

$$r_{ij} = \frac{\sum_{m=1}^{N}\sum_{n=1}^{N}(T_{m,n} - \mu_T)(S_{i+m,j+n} - \mu_S)}{K_1 K_2} \tag{1}$$

where

$$K_1 = \left[\sum_{m=1}^{N}\sum_{n=1}^{N}(T_{m,n} - \mu_T)\right]^{1/2}$$

$$K_2 = \left[\sum_{m=1}^{N}\sum_{n=1}^{N}(S_{i+m,j+n} - \mu_S)\right]^{1/2}$$

First, a vehicle is extracted from a PAN image with 0.6m resolution by the object-based approach. A database is obtained after vehicle extraction with the information of vehicles' location. Using this information, a vehicle and the surrounding road is selected as a template. Since the time lag between PAN and MS images is around 0.2s, the maximum moving distance is about 7m (when the maximum moving distance is 120km/h). The reference image is selected with the same center as the template but 7m bigger than that to the moving direction. The cross-correlation coefficient between the two areas is calculated by sliding the template over the reference image, multiplying the two arrays pixel by pixel. The point of maximum correlation indicates the position of the vehicle in the MS image with the highest probability. The template and the reference image are transformed to 0.24m/pixel by cubic convolution. Then the template and the reference image can be matched in a sub-pixel level.

B. Experiment and result

The three parts of the QuickBird's MS image and the results of vehicle extraction from the corresponding PAN image are used to test the area correlation method.

Since the MS image has 4 bands (R, G, B and Near-Infrared), it needs to be transformed to one band image before the area correlation analysis. Principal Component Analysis (PCA) was employed to transform the MS image to a new 4 band image. The first component image with the highest variance (over than 80%) was used to calculate the cross-correlation coefficient with the PAN image. To match the template and the reference image in a sub-pixel level, both the PAN and MS images were transformed to 0.24m/pixel by cubic convolution, 1/10 pixel size of the MS image.

A vehicle in the MS image is mixed with road. Thus, the template selected from the PAN image must include not only a vehicle but also the road around it. After object-based vehicle extraction from the PAN image, the locations of the beginning and ending points in rows and columns are stored in the database. The template is selected 5 pixels bigger than the vehicle object in each direction. Then a reference image is selected around the location of the template from the MS image (Fig. 4). The maximum movement in the time lag between the PAN and MS images is about 7m, which is 29 pixels in the image. Thus the reference image is selected as 29 pixels bigger than the template image to the moving direction. Since the pixels in the MS image are mixed with road and vehicles, to improve the accuracy of the area correlation method, the reference image is selected as 5 pixels bigger than the template in each direction except for the moving direction; e.g. if a vehicle object is 25×17 pixels in an image, and then the template image is selected as 35×27 pixels from the resized PAN image (0.24m/pixel), the reference image is selected as 69×37 pixels from the resized MS image (0.24m/pixel). The cross-correlation coefficient of each shift is calculated as a matrix shown in Fig. 5. The location of the maximum correlation is the upper left point of the template in the reference image.

Comparing the results visually, the vehicle templates were accurately matched with the reference extracted from the MS image. But it is difficult to access the accuracy of sub-pixel level extraction only by visual comparison. Thus, two 0.12m

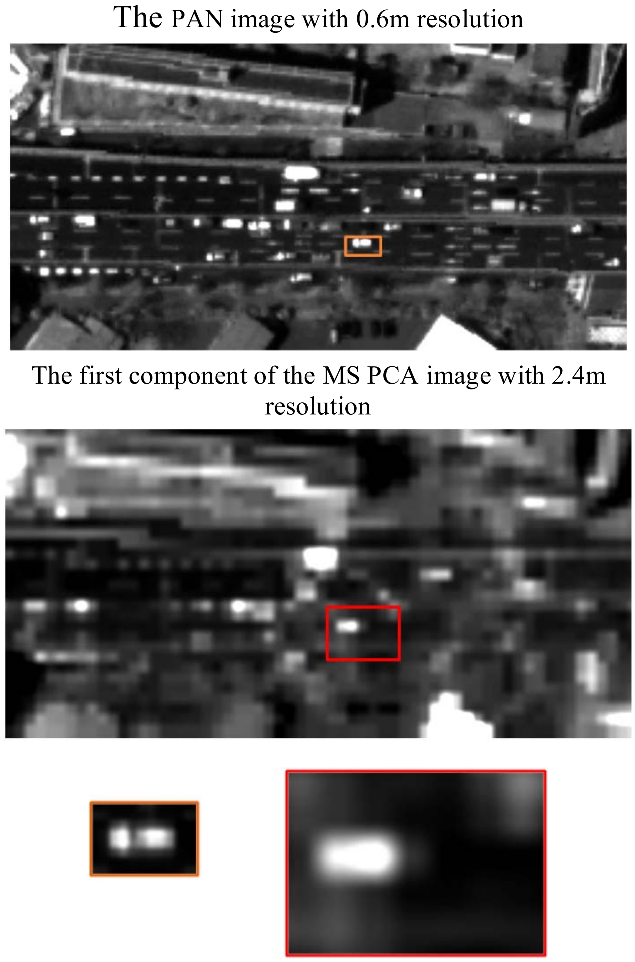

Figure 4. The template selected from the PAN image (orange) and the reference image selected from the MS PCA image (red)

resolution consecutive digital aerial images with both PAN and Reference [8] shows that two consecutive aerial images can be used to detect the speed of moving vehicles. But the time lag between the two consecutive aerial images is about 3 seconds and hence the shift of a vehicle is large. To adjust this difference, the templates of vehicles and the reference image were selected manually. A target vehicle was extracted by the object-based method from the two images. According to the result of extraction, the template was selected from the first image including the target vehicle object and the surrounding road, and the reference was selected from the second image, that is bigger than the template. To compare with the QuickBird study, the pixel size of the original PAN image was also resized from 0.12m to 0.24m. The reference image extracted from the second image was overlaid by the template from the first image and the cross-correlation matrix was obtained. Since the digital aerial images have high spatial resolution, the result is considered to be accurate.

Then the resolution of the PAN image from the first image was converted to 0.6m/pixel, simulating a PAN image of QuickBird. The resolution of the MS image from the second aerial image was also converted to 2.4m/pixel, simulating a MS image of QuickBird. The first component of the MS PCA

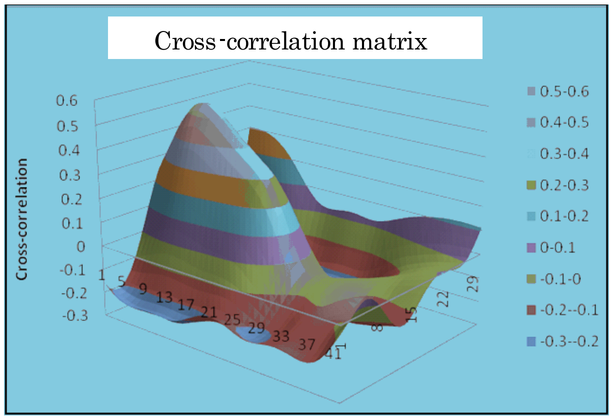

Figure 5. The cross-correlation matrix obtained by shifting the template over the reference image

image was used to calculate the cross-correlation matrix with the simulated PAN image. To register the two images in a sub-pixel level, the pixel sizes of the two images were resized to 0.24m by cubic convolution. The template and the reference image were selected from the simulated PAN and MS images at the same location with the original high-resolution data. The cross-correlation matrix was obtained by shifting the simulated PAN image over the MS image. The maximum correlation point represents the upper left point of the template locates in the reference image.

Comparing the result with the original data, the standard deviation of the difference to the x-axis is about 2 pixels (0.48m), and that to the y-axis is about 3pixels (0.72m). Since the width of vehicles is less than 2.4m, the difference in width is bigger due to a mixed-pixel effect. However, the area correlation method can still extract a vehicle from a MS image with 2.4m resolution in a sub-pixel level.

IV. Speed Detection

Using the time lag between QB's PAN and MS images, the speed of moving vehicles can be detected. Vehicles can be extracted by the object-based method from the PAN image. Using the MS image together, the area correlation method can extract the most probable vehicle's position. The speed of a moving vehicle can be computed by the location change between the PAN and MS images with the time lag about 0.2s.

The moving vehicles in the three parts of Tokyo's QuickBird were extracted from both the PAN and MS image in the former sections. The movement in the time lag can be calculated from the position of a vehicle in the two images. A part of the result is shown in Fig. 6. Since the image was taken at the time of traffic jam, there are many vehicles lining on the expressway and moving slowly. From the 2.4m resolution MS image, vehicles cannot be distinguished clearly. Thus, the result of speed detection cannot be compared with visual inspection. But the result is seen to represent possible speeds, except for several errors.

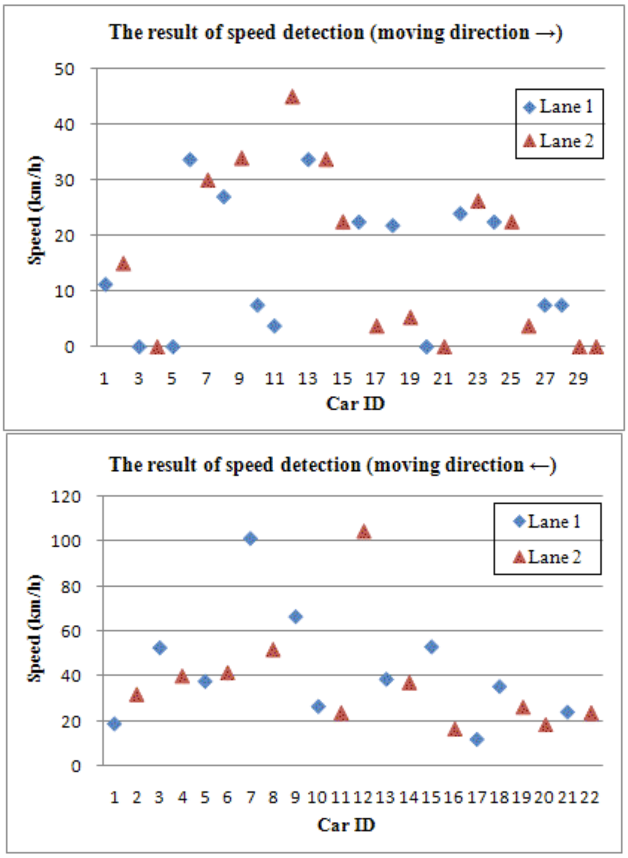

Figure 6. Result of speed detection from scene 1

A simulation was also carried out using the digital aerial images to verify the accuracy of speed detection. First, the target vehicles were extracted and moved manually to make a simulated second image. The movements were recorded as the reference data. Then the original digital image was transformed to be 0.6m resolution, simulating a PAN image, and the second image was transformed to be 2.4m resolution, simulating a MS image. Vehicle extraction and speed detection were applied to the two simulated images. The result of speed detection was compared with the reference data. The speeds of all the target vehicles were detected from the simulated images. Comparing with the reference data, the standard deviation for the difference of speed between the automated result and the reference is about 12km/h. The accuracy of speed detection depends on the accuracy of vehicle extraction. Since the error of vehicle extraction from the MS image is about 0.72m, the accuracy of vehicle movement is in a range of ±0.72m. The accuracy of speed detection is ±12km/h (the speed of vehicle moving 0.72m in 0.2s).

V. CONCLUSIONS

The methods to extract moving vehicles and to measure their speeds from QuickBird images were proposed. First, vehicles were extracted automatically by an object-based approach from a 0.6m resolution QB PAN image. Three parts of a QB image over the central Tokyo were applied to test the approach, and 94% vehicles were successfully extracted.

Then an area correlation method was employed to extract vehicles' location from a 2.4m resolution MS image accurately in a sub-pixel level. A template including a vehicle was selected from a PAN image, and a reference image was selected from a MS image. From the cross-correlation matrix, the position of the maximum correlation could be obtained. The simulation showed that vehicles can be detected with a sub-pixel level accuracy (1/3 pixel of the MS image).

The speed of moving vehicles was detected by the difference of the vehicle position in the PAN and MS image pair. Due to the limitation of vehicle extraction from a MS image, highly accurate speed detection cannot be expected. From the result of the simulation, the speed of moving vehicles can be detected in a range of ±12km/h from a QuickBird bundle product.

More examples are necessary to verify the accuracy of vehicle extraction and speed detection. Accuracy will be improved by introducing a rotation angle between a template and a reference to the area correlation method. The result of this study may be used in assessing traffic conditions of large urban areas without using ground-based traffic sensors.

REFERENCES

[1] M. Etaya, T. Sakata, H. Shimoda, and Y. Matsumae, "An experiment on detecting moving objects using a single scene of QuickBird data," *Journal of the Remote Sensing Society of Japn*, vol. 24, No. 4, pp. 529–551, 2004.

[2] Z. Xiong, and Y. Zhang, "An initial study of moving target detection based on a single set of high spatial resolution satellite imagery," *Proc. ASPRS 2006 Annual Conference*, Reno, Nevada, May 2006.

[3] A. Gerhardinger, D. Ehrlich and M. Peseresi, " Vehicles detection from very high resolution satellite imagery," *CMRT05. IAPRS, Vol. XXXVI, Part3/W24*, 2005.

[4] T. Zhao, and R. Nevatia, "Car detection in low resolution aerial image," *International Coference on Comuter Vision*, 2001.

[5] X. Y. Jin and C. H. Davis, "Vector-Guided vehicle detection from high-resolution satellite imagery," *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International*, September 2004.

[6] G. Sharme, "Vehicle detection and classification in 1-m resolution imagery," *Ohio State University*, Master of Science thesis, 2002.

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

[8] W. Liu, F. Yamazaki, T. T. Vu and Y. Maruyama, "Speed detection of vehicle from aerial photographs," *Proc. 27th Asian Conference on Remote Sensing*, November 2007.

[9] L. Eikvil, L. Aurdal, and H. Koren, "Classification-based vehicle detection in high-resolution satellite images," *International Society for Photogrammetry and Remote Sensing, Inc.*, Elsevier B.V., 2008.

[10] T. T. Vu, M. Matsuoka, and F. Yamazaki, "Preliminary results in development of an object-based image analysis method for earthquake

damage assessment," *Proc. 3th International workshop Remote Sensing for Post-Disaster Response*, Chiba, Japan, 2005.

[11] R. A. Schowengerdt, *Remote Sensing, Models and Methods for Image Processing, Second Edition*, Arizone: Academic Press, 2007.

[12] L. G. Brown, "A survey of image registration techniques," *ACM Computing Survey*,Vol. 24, No. 4, pp.325-376, 1992.

978-1-4244-3461-9/09/$25.00 ©2009 IEEE