

# Road Image Update using In-vehicle Camera Images and Aerial Image

Masafumi NODA<sup>1,\*</sup>, Tomokazu TAKAHASHI<sup>1,2</sup>, Daisuke DEGUCHI<sup>1</sup>, Ichiro IDE<sup>1</sup>,  
Hiroshi MURASE<sup>1</sup>, Yoshiko KOJIMA<sup>3</sup> and Takashi NAITO<sup>3</sup>

<sup>1</sup>*Nagoya University, Japan*

<sup>2</sup>*Gifu Shotoku Gakuen University, Japan*

<sup>3</sup>*Toyota Central Research & Development Laboratories, Inc., Japan*

\**mmoda@murase.m.is.nagoya-u.ac.jp*

**Abstract**—Road image is becoming important for several applications such as car navigation systems, traffic environment research, city modeling. Usually, a road image can be obtained from an aerial image but the resolution of the aerial image is often low, or it contains occlusions by obstacles. Therefore, the update of road image is required. In this paper, we propose a road image mosaicing method using in-vehicle camera images and an aerial image. We first perform image registration of road regions between these images, and then, we generate a large road image by performing image mosaicing of road regions in in-vehicle camera images. In an experiment, we achieved resolution improvement and occlusions removal, and also succeeded in update of a large road image.

## I. INTRODUCTION

Road is an important infrastructure of ITS (Intelligent Transportation Systems). Road contains a lot of information for driving (e.g. driving lane and road marking). Therefore, road image is becoming important for several applications such as car navigation systems, traffic environment research, city modeling.

As the one of applications, a road image is used for the ego-localization in previous works [1] [2]. Ego-localization of these methods was performed by matching between road regions in an aerial image (Fig. 1) and in an in-vehicle camera image (Fig. 2). In these methods, however, the road image was obtained from an aerial image, but the aerial image is often in low resolution (Fig. 3(a)), or the road markings are occluded by obstacles (Fig. 3(c)). In such cases, the accuracy of ego-localization degrades. Therefore, it is required to make the road map which quality is high, such as the resolution is high, the occlusion is not included and the accuracy is within 1.0 meter. Also, since the road image is changed over time, the update is required. It is difficult that the aerial image is taken many times because the cost is high. Therefore, we try to update the aerial image using an in-vehicle camera image sequence.

In the paper, we propose the update method of road image. The proposed method achieves followings by updating.

- 1) Resolution improvement: from Fig. 3(a) to Fig. 3(b)
- 2) Occlusion removal: from Fig. 3(c) to Fig. 3(d)

In order to achieve the aim, we perform image mosaicing using in-vehicle camera images and an aerial image. The frontal road



Fig. 1. Aerial image.



Fig. 2. In-vehicle camera image.

region in the in-vehicle camera image is high resolution and does not contain occlusions. We propose the road image mosaicing method by performing the image registration between the in-vehicle camera image and the aerial image.

The rest of this paper is organized as follows: We introduce the related works in Section II, and details of the proposed method are explained in Section III. Then Section IV reports experimental results and discussion of them. Finally the paper is concluded in Section V.

## II. RELATED WORKS

A large image generation method from a partial image sequence is known as *image mosaicing* or *video mosaicing*. Several methods have been proposed for the image mosaicing using camera mounted in a mobile object, such as UGV (Unmanned Ground Vehicle) [3] or UAV (Unmanned Aerial Vehicle) [4] [5] [6]. The methods [3] used only in-vehicle camera. Therefore, image mosaicing in a wide region is

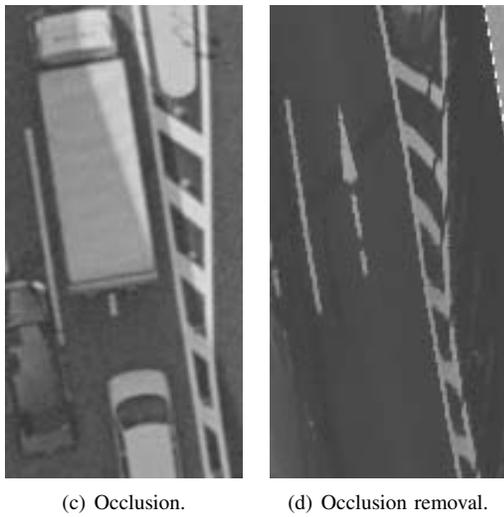
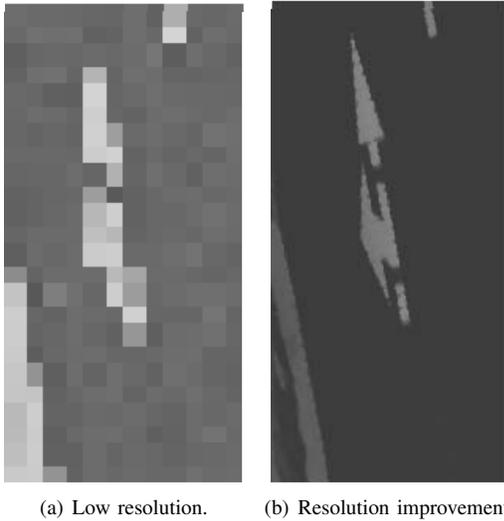


Fig. 3. Our aims are resolution improvement (from (a) to (b)), and occlusion removal (from (c) to (d)).

difficult due to an accumulative error by a vehicle motion. Though it is thought that the vehicle mounted with advanced measurement system is used, the cost is very high. There, we make the road image cheaply without the accumulative error by using an aerial image as a reference for image mosaicing in this paper.

Recently, the aerial image is relatively easily available because it is provided by several web services. The aerial image has high accuracy. Therefore, by using such aerial image as the reference image, the image mosaicing can be performed consistently. The related works [4] [5] [6] performed the image mosaicing of a large image by matching between an aerial image sequence taken from UAV and a satellite image. However, since we have to match between an in-vehicle camera image sequence and a aerial image, there are problems that the image mosaicing is more difficult because we can use road surface only and the vehicle motion is very large, Therefore, we proposed the appropriate image mosaicing

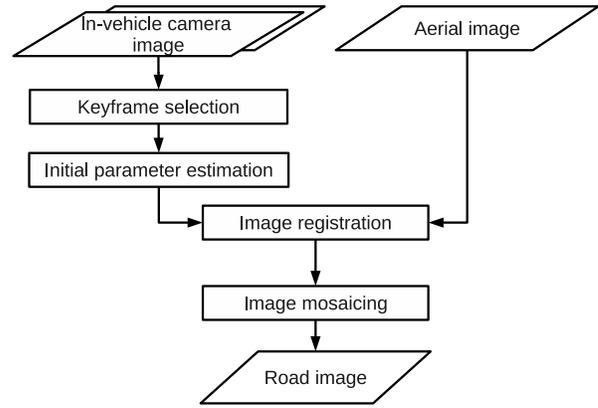


Fig. 4. Processing flow.

method for road image, and solve these problems.

### III. PROPOSED METHOD

#### A. Strategy

The proposed method sequentially computes correspondences between road regions in an in-vehicle camera image and an aerial image by *image registration* [7]. In this research, it is assumed that the aerial image is in low resolution, or it contains occlusions, and the texture of the road region is poor. Therefore, we apply image registration which can consistently match images using entire image.

In order to generate the road image accurately, the proposed method takes the following three strategies.

- 1) Enhancement of road markings to improve the accuracy of the image registration.
- 2) Keyframe selection for avoiding unexpected image mosaicing caused by vehicle motion
- 3) Stable initial parameter estimation for image registration by using the result of the previous frame.

#### B. Processing Flow

The processing flow of the proposed method is shown in Fig. 4. An input is an in-vehicle camera image sequence  $I_t$  ( $t = 0, \dots, n$ ) and an aerial image  $J$  taken in the same place, and an output is a mosaicing image.

First, the proposed method selects keyframes, that can be registered appropriately, from an in-vehicle camera image sequence. Then, initial parameters of image registration are estimated at each keyframe, and the transformation parameter is also estimated by matching the in-vehicle camera image and the aerial image. Finally, a large road image is generated using the in-vehicle camera image sequence.

#### C. Image Registration between an In-vehicle Camera Image and an Aerial Image

The proposed method performs image registration between an in-vehicle camera image  $I_t$  and an aerial image  $J$ , as shown

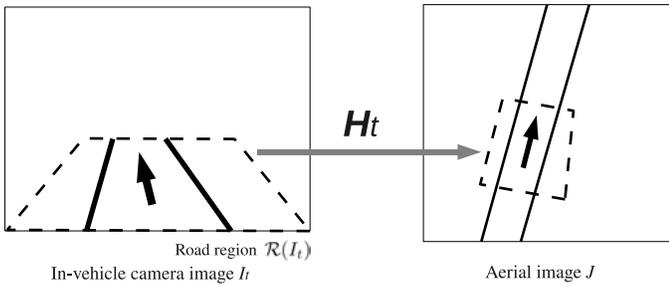


Fig. 5. Registration between an in-vehicle camera image and an aerial image.

in Fig. 5. This image registration is performed by minimizing the following objective function

$$\epsilon(\mathbf{H}_t) = \sum_{\mathbf{x} \in \mathcal{R}(I_t)} w(I_t(\mathbf{x}))(I_t(\mathbf{x}) - J(\mathbf{H}_t \mathbf{x}))^2, \quad (1)$$

where  $\mathcal{R}(I_t)$  shows the road region in  $I_t$ , and details of it is shown in Section III-D. It is assumed that a homography matrix  $\mathbf{H}_t$  gives the correspondence between road regions in  $I_t$  and  $J$ . The relation between  $\mathbf{x} = [x, y, 1]^T$  in  $I_t$  and  $\mathbf{H}_t \mathbf{x}$  satisfies

$$\mathbf{H}_t \mathbf{x} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (2)$$

where  $\mathbf{x}$  and  $\mathbf{H}_t$  are represented by using a homogeneous coordinate system.  $w(\mathbf{x})$  is a weight function defined as

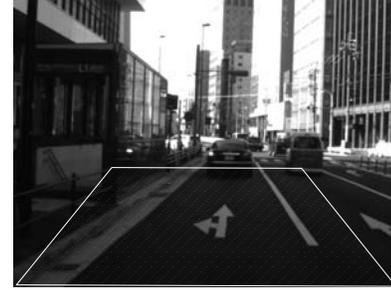
$$w(I_t(\mathbf{x})) = \begin{cases} \omega & I_t(\mathbf{x}) > \theta \\ 1 & \text{otherwise} \end{cases}. \quad (3)$$

Here,  $\omega$  is selected so that its value is larger than or equal to 1. In this paper, in order to improve the accuracy of the image registration, the road markings is enhanced. It is considered that the road marking has a higher intensity than its surroundings. Therefore, this equation weights a pixel whose values is larger than a threshold  $\theta$ .

In this paper,  $\mathbf{H}_t$  is optimized by minimizing Eq. (1) using the steepest descent algorithm [8]. The initial parameter  $\mathbf{H}_t^{(0)}$  is given before the optimization. This is explained in Section III-F

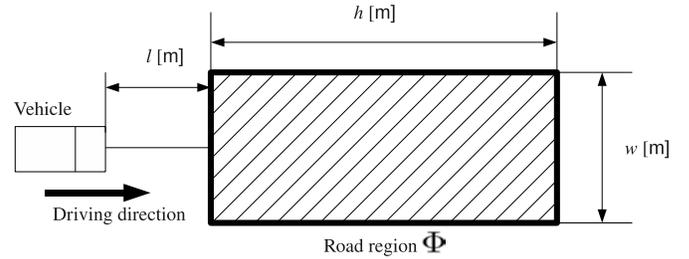
#### D. Road Region

In this paper, the frontal road region  $\mathcal{R}(I_t)$  in the in-vehicle camera image is used for image registration between an in-vehicle camera image and an aerial image, as shown in Fig. 6(a). Here, the road region of both an in-vehicle camera image and an aerial image is assumed to be a planar.  $\mathcal{R}(I_t)$  is transformed from a road region  $\Phi$  in a virtual vehicle/road model, as shown in Fig. 6(b). This transformation is done by using camera parameters, such as mounting height, field angle and so on. These camera parameters are given in advance.



Road region  $\mathcal{R}(I_t)$

(a) In-vehicle camera image.



(b) Virtual vehicle/road model.

Fig. 6. Road region for the image registration.

#### E. Keyframe Selection

The keyframes are selected from the in-vehicle camera image sequence. These keyframes can be registered accurately. In this paper, it is assumed that larger road marking regions are registered easily and accurately. Therefore, the registration accuracy can be approximated by the number of pixels whose value is larger than a threshold  $\theta$  as in Eq. (3).

$$E(t) = \sum_{\mathbf{x} \in \mathcal{R}(I_t)} V(I_t(\mathbf{x})). \quad (4)$$

$$V(I_t(\mathbf{x})) = \begin{cases} 1 & I_t(\mathbf{x}) > \theta \\ 0 & \text{otherwise} \end{cases}. \quad (5)$$

The proposed method selects  $K$  keyframes from the in-vehicle camera image sequence. First, the image sequence is divided into  $K$  subsequences having the same length. Then, for each subsequence, the keyframes are selected as a frame, so that the value of Eq. (4) should be maximized.

After selecting keyframes, a matrix  $\mathbf{H}_t$  ( $t'_1 < t < t'_2$ ) in the subsequence is linearly interpolated using  $\mathbf{H}_{t'_1}$  and  $\mathbf{H}_{t'_2}$  at successive keyframes  $t'_1$  and  $t'_2$ .

#### F. Initial Parameter Estimation

The initial parameter  $\mathbf{H}_t^{(0)}$  for Eq. (1) is an important factor for the image registration. In order to obtain a good initial parameter, the proposed method calculates  $\mathbf{H}_t^{(0)}$  by estimating ego-motion from the previous keyframe. Here,  $\mathbf{H}_t^{(0)}$  is obtained as,

$$\mathbf{H}_t^{(0)} = \mathbf{M}_t \mathbf{M}_{t-1} \cdots \mathbf{H}_{t'}, \quad (6)$$



(a) Dataset 1 (Resolution improvement).



(b) Dataset 2 (Occlusion removal).

Fig. 8. In-vehicle camera image sequences.

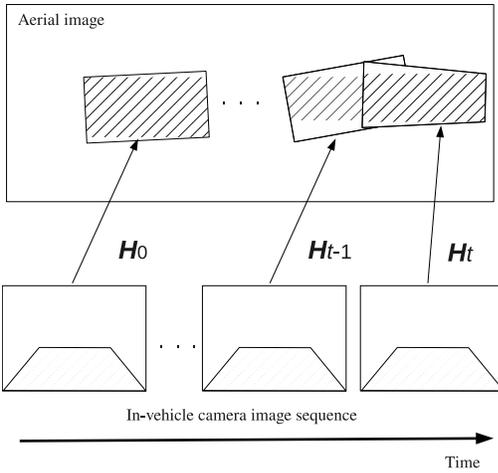


Fig. 7. Image mosaicing.

where  $t' (< t)$  is the time of the previous keyframe.  $\mathbf{M}_t$  is the matrix that transforms  $\mathcal{R}(I_{t-1})$  to  $\mathcal{R}(I_t)$ , and it is obtained by the image registration of these regions. This image registration is performed by minimizing

$$\epsilon'(\mathbf{M}_t) = \sum_{\mathbf{x} \in \mathcal{R}(I_{t-1})} w(I_{t-1}(\mathbf{x}))(I_{t-1}(\mathbf{x}) - I_t(\mathbf{M}_t \mathbf{x}))^2. \quad (7)$$

This minimization is performed by using the steepest descent algorithm, as in Eq. (1). Then,  $\mathbf{M}_{t'+1}, \dots, \mathbf{M}_t$  are calculated, and  $\mathbf{H}_t^{(0)}$  is obtained by Eq. (6). Although  $\mathbf{H}_t^{(0)}$  may contain accumulative error, this can be avoided by shortening keyframe intervals.

### G. Image Mosaicing

Finally, the image mosaicing is performed using the in-vehicle camera image sequence. This mosaicing is performed by sequentially transforming the road region  $\mathcal{R}(I_t)$  into the aerial image coordinate system using  $\mathbf{H}_t$  ( $t = 0, \dots, n$ ), as shown in Fig. 7.

## IV. EXPERIMENTS AND DISCUSSION

We performed road image mosaicing using in-vehicle camera images by the proposed method. We demonstrate the effectiveness of proposed method by using two datasets.

### A. Setup

In this experiment, we used two datasets which consist of pairs of an in-vehicle camera image sequence (Fig. 8) and an aerial image (Fig. 9 and 10). These in-vehicle camera image sequence in each dataset was taken from the vehicle driving in different routes. Dataset 1 was used for the resolution improvement, and Dataset 2 was used for the occlusion removal.

We used an in-vehicle camera whose resolution was  $640 \times 480$  pixels and the frame rate was 30 fps. Both sequences consisted of 500 frames. We used parameters of  $w = 10$ ,  $h = 20$  and  $l = 5$  for obtaining the frontal region of the vehicle, as shown in Fig. 6. The resolution of the aerial image in Fig. 9(a) was about 0.75 meters per pixel, as shown in Fig. 9(a). Fig. 9(a) was created by degrading the resolution of the original image shown in Fig. 10(a). The initial parameter  $\mathbf{H}_0^{(0)}$  was manually specified for both datasets. The other parameters were experimentally given as  $\omega = 10$  and  $\theta = 100$ .

### B. Result and Discussion

Figs. 9(b) and 10(b) show the image mosaicing result. In these figures, the image mosaicing results were overlaid on the original aerial images. From these figures, we can see that the proposed method could mosaic the images in a wide road region without failure. The resolution of these results is quite high, and the results are not occluded by obstacles, therefore it is available for some applications.

For Dataset 1, the resolution was improved from about 0.75 meters per pixel of the original image to about 0.1 meters per pixel. A part of Fig. 9 is zoomed-in in Fig. 11. We can see that the ambiguous shapes of the road marking has become sharp.

For Dataset 2, we can see that the obstacles occluding the road region was removed. A part of Fig. 10 is zoomed-in



(a) Original image.



(b) Image mosaicing result.

Fig. 9. Aerial image: Dataset 1 (Resolution improvement).



(a) Original image.



(b) Image mosaicing result.

Fig. 10. Aerial image: Dataset 2 (Occlusion removal).

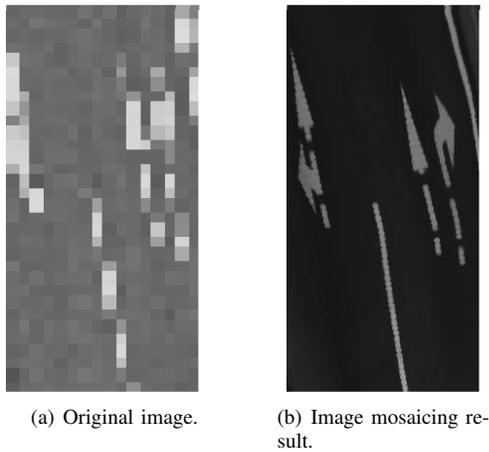


Fig. 11. Resolution improvement.

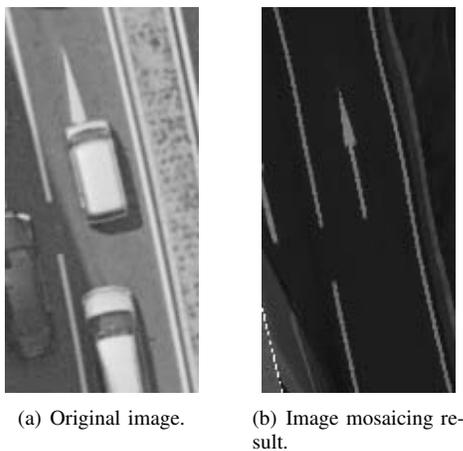


Fig. 12. Occlusion removal.

in Fig. 12. After applying the proposed method, the road markings under the obstacles became visible, but the result was miss-aligned. This is caused by the image registration error due to the obstacle, therefore the obstacles should be detected in advance for improvement.

In this experiment, several problems were revealed. Fig. 13 shows the unsuccessful cases. As shown in Fig. 13(a), the image mosaicing result was distorted. This is because the image registration error was large due to the large ego-motion variation. In another case, as shown in Fig. 13(b), obstacles, such as parking vehicles, still remained. Therefore, we will tackle these problems in the future work. Also, we will evaluate the performance of the proposed method in various conditions, such as night, rain and so on.

## V. CONCLUSION

In the paper, we proposed the update method of road image. The proposed method performed a road image mosaicing using in-vehicle camera images and an aerial image. In the experiment, we could generate a road image in a wide region by the proposed method. In an experiment, we achieved resolution

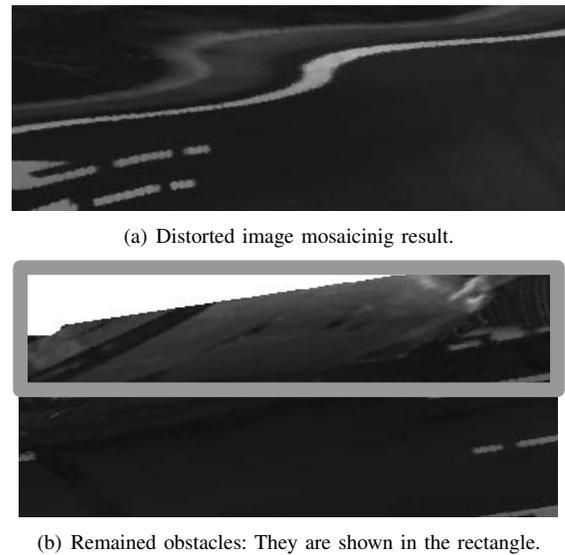


Fig. 13. Unsuccessful cases.

improvement and occlusions removal, and also succeeded in update of a large road image.

In the future work, we will tackle the remaining problems in this paper; large ego-motion, obstacles without movement, and so on.

## ACKNOWLEDGMENTS

Parts of this research were supported by JST CREST and MEXT, Grant-in-Aid for Scientific Research.

## REFERENCES

- [1] M. Noda, T. Takahashi, D. Deguchi, I. Ide, H. Murase, Y. Kojima, and T. Naito, "Vehicle ego-localization by matching in-vehicle camera images to an aerial image," in *Proc. ACCV2010 Workshop on Computer Vision in Vehicle Technology: From Earth to Mars*, pp. 1–10, Nov. 2009.
- [2] O. Pink, F. Moosmann, and A. Bachmann, "Visual features for vehicle localization and ego-motion estimation," in *Proc. Intelligent Vehicles Symp.*, pp. 254–260, Jun. 2009.
- [3] A. Geiger, "Monocular road mosaicing for urban environments," in *Proc. Intelligent Vehicles Symp.*, pp. 140–145, Jun. 2009.
- [4] A. Camargo, R. R. Schultz, and Q. He, "Super-resolution mosaicking of unmanned aircraft system (UAS) surveillance video using Levenberg Marquardt (LM) algorithm," in *Proc. Int. Symp. on Visual Computing*, pp. 698–706, Nov. 2010.
- [5] F. Caballero, L. Merino, J. Ferruz, and A. Ollero, "Unmanned aerial vehicle localization based on monocular vision and online mosaicking," *J. Intelligent Robotics System*, vol. 55, pp. 323–343, 2009.
- [6] Y. Lin and G. Medioni, "Map-enhanced UAV image sequence registration and synchronization of multiple image sequences," in *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 1–7, Jun. 2007.
- [7] B. Zitova, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, 2003.
- [8] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *Int. J. Computer Vision*, vol. 56, no. 1, pp. 221–255, 2004.
- [9] R. Szeliski, "Image mosaicing for tele-reality applications," in *Proc. IEEE Workshop on Applications of Computer Vision*, pp. 230–236, Dec. 1994.
- [10] D. Capel, *Image mosaicing and super-resolution (Cphc/Bcs distinguished dissertations.)*. Springer-Verlag, 2004.