Special Feature: Spatial Information Technology towards Intelligent Vehicle Systems

Research Report

Automatic Lane-level Map Generation Using Low-cost Sensors in Urban Traffic Scenarios

Chunzhao Guo, Kiyosumi Kidono and Junichi Meguro

Report received on Dec. 28, 2017

ABSTRACT A detailed and up-to-date digital map is crucial to automated vehicles. In this work, we proposed an approach that can generate the lane-level map of urban streets in an automatic manner using low-cost onboard sensors. First, it divides the global map into fixed local segments. Subsequently, it integrates the road surface images acquired from a rear-view camera into large synthetic orthographic images. Finally, it constructs the lane-level map from the integrated images together with a large amount of vehicle trajectories without manual processing. Experimental results substantiated the effectiveness of the proposed system, which can provide strong priors about the road environment, thereby enhancing the performance of automated driving.

KEYWORDS Digital Map, Mapping, Lane Graph, ADAS, Automated Driving

1. Introduction

In order to achieve zero traffic fatalities, it is crucial to promote technologies of automated driving and deploy them in the next generation cars. Among them, the detailed and up-to-date digital map is one of the most important components, since it can not only provide necessary route information for contemporary navigation systems, but also simplify the driving tasks with strong priors about the road environment. Generally, the map construction methods can be divided into two major categories. One is to use the aerial images of road to construct the digital map.[1-3] These methods can generate the map automatically, whereas, the cost of the aerial images is quite expensive. Moreover, it cannot deal with the roads with multiple layers, such as overpasses and tunnels. The other category of methods depends on specialized probe cars with expensive sensors to collect data of a predefined course, and use manual efforts for data processing.[4-10] Such methods can deal with the multi-layer roads and the generated digital map usually has high accuracy. However, the coverage is limited and update rate is low due to the high cost. These limitations prevent their benefits to the normal cars.

The lane-level map generation approach proposed in this work belongs to the latter category, nevertheless, it is able to construct the map in low cost. We have developed the following two ways to reduce the cost. One is to use low-cost sensors. In our system, we only make use of conventional sensors available in contemporary cars, such as GPS (Global Positioning System), INS (Inertial Navigation System) and a rear-view camera, which makes a large fleet of standard vehicles possible to sample the road environment during their daily driving. The other is to remove the manual data processing. In our system, the spatial information of the environment is extracted in an automatic manner and the map is constructed without human operation or intervention. The flow chart of the proposed system is shown in **Fig. 1**. First, the global map is divided into fixed local segments according to the Open Street Map (OSM).[11] Meanwhile, the GPS/INS-based vehicle positioning[12] is used to initialize an optimization-based alignment of the sequential road images acquired by the rear-view camera. Subsequently, the aligned road images are integrated into large synthetical orthographic images within and between the local map segments. Finally, the spatial information is extracted from the orthographic road image, which is then used to construct the lane graph in an automatic manner.

The rest of the paper is organized as follows: Sec. 2 describes the details of the proposed approach. In

Sec. 3, the experimental results are provided, which proved the effectiveness of the system. Finally, Sec. 4 draws the conclusions and discusses the future work.

2. Proposed Approach

2. 1 Map Segmentation

The global map is firstly divided into fixed local segments according to the road network topology such that the information can be processed locally. As shown in **Fig. 2**, for any spot in the global world, a region map will be defined in the East-North-Up (ENU) coordinate. The information of the region map, such as intersections/road ends (nodes) and waypoints of the roadways (links) as well as their connectivity, is imported automatically from the OSM. For each node and link in the region map, a fixed local segment will be generated accordingly. More specifically, the local segment for a node is a square with a pre-defined fixed range, as indicated by the magenta boxes in Fig. 2. The local segment for a link is a rectangle that contains both of the nodes it connects and all of its waypoints with a pre-defined fixed margin, as indicated by the white boxes in Fig. 2. It should be noted that the overlapping between the adjacent node and link segments can be ensured due to the preserved margin.

2. 2 Optimization-based Image Alignment

The sequential road images acquired from a rear-view camera during vehicles' daily driving are aligned to generate the local bird's eye view images of the road surface, from which, the map information will be extracted. As shown in **Fig. 3**, in the projective geometry, the mapping from the point p to p' can be decomposed into three individual projections, i.e., the inter-frame vehicle's motion, the projection from ground to camera, and the projection from camera to image described by the intrinsic matrix. Generally, the intrinsic parameters and the camera height and roll angle barely change during vehicle driving, thereby, we further simplify the mapping by the vector $\boldsymbol{\beta} = (\Delta x, \Delta y, \Delta\alpha, \Delta\theta)^T$, and optimize it by minimizing the compatible cost $E(\beta)$ in Eq. (1) with the LMA (Levenberg-Marquardt algorithm)[13] between the consecutive frames.

$$E(\beta) = \sum_{p\in\bar{P}} \sum_k \lambda_k [\Phi_k^t(p) - \Phi_k^{t-1}(p')]^2 , \quad (1)$$

where Δx, Δy are the vehicle translation, $\Delta\alpha$, $\Delta\theta$ are the yaw and pitch rates, respectively. λ_k is the k^{th} feature's weight. $\bar{P}$ is a computational region, as indicated by the green box in Fig. 3, which is a predefined ROI (region of interest) in the free road space. p is a road pixel in the region of the current frame and p' is the pixel projected from the same road point in the adjacent previous frame. $\boldsymbol{\Phi}(\cdot)$ is a feature vector which includes the intensity value and a 2D gradient vector in the horizontal and vertical axes. The GPS/INS-based vehicle positioning[12] results are used to initialize $\boldsymbol{\beta}$ for each frame. Thanks to the tightly integration of the high-resolution GPS Doppler and INS measurements together with the bundle adjustment over a number of preserved GPS pseudorange data, the accurate and robust initial values of $\boldsymbol{\beta}$ are provided, which is very important for optimization particularly in the cases of low-textured roads and occlusion by nearby cars. With the optimized $\boldsymbol{\beta}$, the image patch, as indicated by the red box in Fig. 3, is projected to the ground plane π and the aligned into a local bird's eye view image of the road surface for each local map segment.

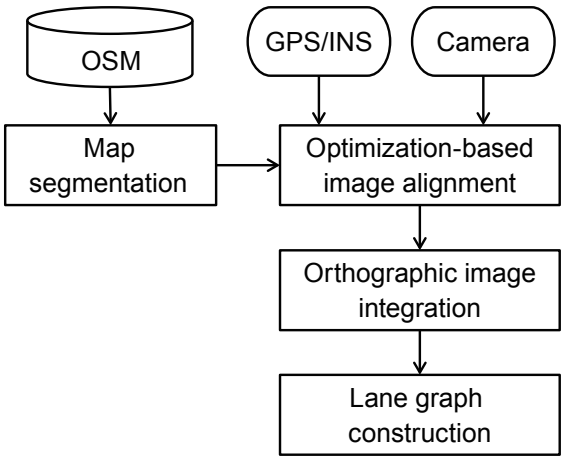


Fig. 1 Flow chart of the proposed system.

2. 3 Orthographic Image Integration

The absolute position and heading angle of the generated local road images are assumed to have zero mean Gaussian errors. Therefore, for each local map segment, all of the aligned local road images from multiple replicated runs within it are integrated

Fig. 2 Illustration of the global map segmentation with a region map, indicated by the cyan box, in the world map. The map images are from OSM.[11]

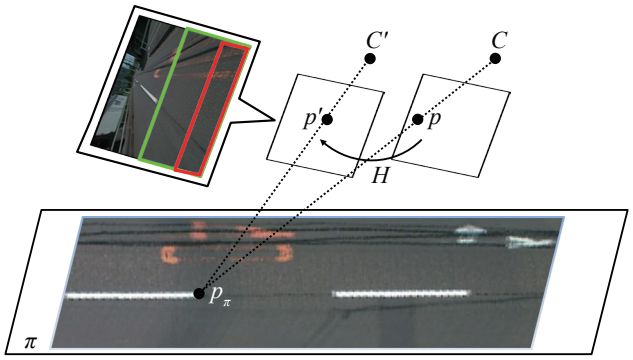


Fig. 3 Mapping of a road point between consecutive frames during vehicle's driving from C' to C.

to enhance the spatial accuracy and the image quality. Specifically, for each pair of overlapped local images, a vector $\boldsymbol{\beta}' = (\Delta x', \Delta y', \Delta\alpha')^T$, which measures the transformation between the image pair, is optimized by Eq. (1) with the overlapped image portion as the computation region. Such a local road image can provide the rich spatial information of the corresponding driving lane for the lane graph construction.

Once the above processing for all of the map segments is finished, the enhanced local road images between adjacent map segments will be further integrated into large synthetical orthographic images in the same manner. Such an integration procedure can improve the global position accuracy of the local maps, as well as their extracted spatial information, based on the optimization over the image evidence. An example of the orthographic image integration is given in **Fig. 4**, which shows the image integrations inside a local map segment and between the adjacent ones.

2. 4 Lane Graph Construction

In the proposed approach, the lane graph is defined to represent the driving lane information. It consists of the link graph for a roadway and the node graph for intersections.

2. 4. 1 Link Graph Construction

The link graph is the centerlines of the driving lanes in a link segment, which is defined by a clothoid spline model[14] as

$$y(x) = y_0 + \tan(\varphi_0)(x - x_0) + \frac{C_0(x - x_0)^2}{2} + \frac{C_1(x - x_0)^3}{6}, \quad (2)$$

where (x_0, y_0) is the origin of the map segment. φ_0 is the orientation of the base of the curve. C_0 and C_1 are the curvature and the rate of curvature of the lane, respectively.

As shown in **Fig. 5**, for a link segment, the lane markings are detected in the integrated local road image at first, which are then refined to find the lane boundaries by using an elongated filter[15] with the spatial prior provided by the vehicle trajectories. Subsequently, the centerline points inside the lane boundaries will be extracted as follows. For each trajectory point $o = (x, y)$, a circle D with the radius of r is defined. For a point $q = (x + u, y + v)$ on the circle, its cost $I(q) = 0$ if the line $\overline{oq}$ does not touch the lane boundary, otherwise $I(q) = 1$. To localize the point o to the center of the lane is equivalent to maximizing the circle D while without touching the lane boundary. We accomplish this by minimizing the following energy function:[16]

$$\in(x, y, r) = \iint_D \left(r - \sqrt{u^2 + v^2}\right) I(x + u, y + v)\, dudv - r^\rho / \rho, \quad (3)$$

where the first term will keep the circle inside the lane while the second term will try to maximize the circle. ρ is a constant for weight adjustment.

For the circles inside the driving lane, their sizes will be almost the same, which is half of the lane width. For the circles which enter into the intersection, their sizes will increase since there are no lane boundaries at intersections. By analyzing the mean and variance of the circles' sizes, as shown in Fig. 5, the entry point and the exit point of the driving lane could be determined automatically. Subsequently, all of the centerline points, i.e., the centers of the circles, between the entry and exit points will be used to fit the clothoid spline to obtain the lane graph of the driving lane. More formally, the lane graph in the link segment will be stored as (N_1, N_2, O, W, H, ppm, n_{lane}, en_k, ex_k, l_k, w_k), which are endpoints (nodes) of the link, ENU coordinates of the upper left corner and ranges of the link, resolution, number of lanes, entry point, exit point, centerline equation and lane width of k$^{\text{th}}$ lane, respectively.

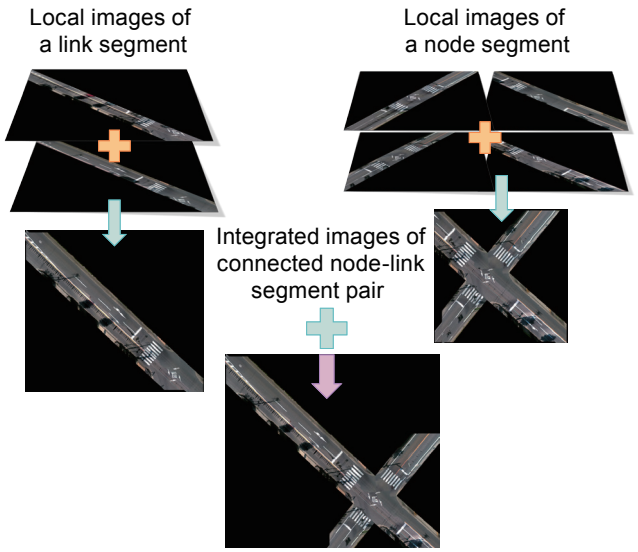


Fig. 4 Example results of the orthographic image integration.

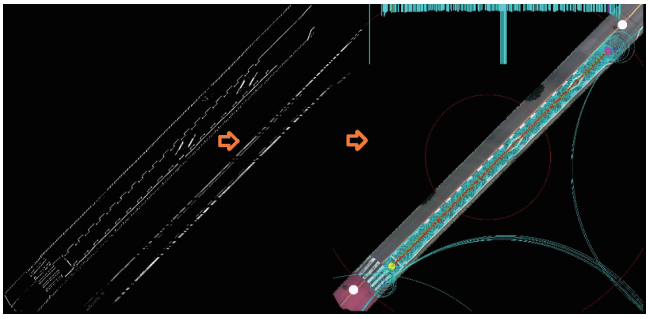


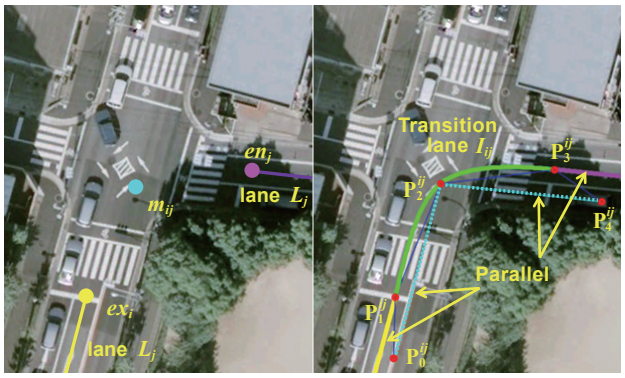
Fig. 5 Example of the link graph construction.

2. 4. 2 Node Graph Construction

After the lane graphs of all the link segments are obtained, the lane graphs in the node segments will be constructed automatically. As shown in **Fig. 6**, it is a smooth and safe path (green line) to pass through the intersection from lane L_i (yellow line) to lane L_j (purple line) while maintaining the continuity at the joint points. At first, the link graphs connected to the node segment will be input. Subsequently, the center of the node will be extracted by averaging the locations of all the entry and exit points. From a large amount of past vehicle trajectories, the connectivity and category of each exit-entry-point pair are obtained

automatically. The node graphs will be generated for each pair, i.e. from exit point ex_i (yellow dots) to entry point en_j (purple dots) via an inter point m_{ij} (cyan dots). The cubic spline with five control points are used to generate the transition lanes. Since the control points are actually in the spline, P_1 and P_3 are firstly assigned with ex_i and en_j. The inter point P_2 is determined according to the category of the transition lane. For driving forward, P_2 (green dot) is set as the middle point of ex_i and en_j. For left or right turn, P_2 (blue and cyan dots) is set as a point in the bisector of angle ex_iOen_j with a given distance D_l or D_r to ensure a safety margin. O (red dot) is the center of the intersection. P_0 and P_4 will be set as points that ensure the continuity between the link graph and the transition lane graph at ex_i and en_j.

(a)



(b)

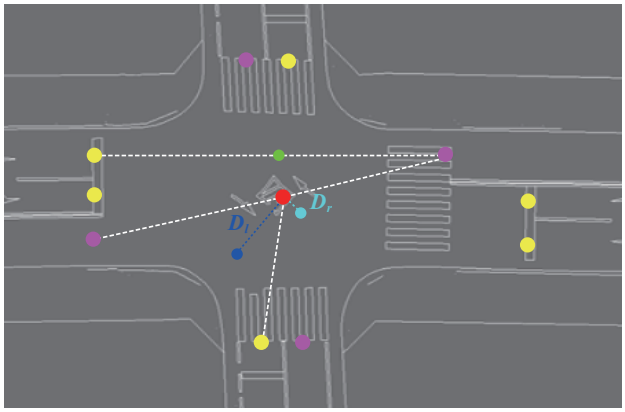


Fig. 6 Example of the node graph construction. (a) Lane model of a node lane graph with five control points. (b) Determinations of the inter point for driving forward and left/right turning.

3. Experimental Results

3. 1 Data Collection

In the experiments, three different-sized cars have been used to collect the driving data with standard low-cost GPS/INS and camera sensors in urban environment with daily traffic. **Figure 7** gives example data collected from the same road but at different dates by the probe cars. The red lines in Figs. 7 (a)-(c) show the trajectories during the data collection, which are very smooth thanks to the use of GPS Doppler. Moreover, the heading angles of the trajectories in the same road segment are compared. As shown in Figs. 7 (d)-(e), the differences are less than ± 0.5°. From these experimental results, we can see that conventional contemporary cars can be used to sample the road environment with the proposed system, thereby resulting in a huge amount of probe vehicles to ensure the larger coverage and faster update for map construction.

3. 2 Synthetic Orthographic Image Integration

Here, we evaluate the accuracy of image alignment with a continuous driving loop. An example result of image alignment is shown in **Fig. 8**, from which, we can see that the aligned image is locally accurate. Furthermore, the aligned image of the loop also coincides with the global world map very well. The positioning error between the same point in a 1636 m driving loop is only 1.3 m.

Furthermore, we evaluate the accuracy of image integration by comparing the generated synthetical orthographic image against the high precision map, which was constructed by using a specialized probe car with manual data processing. As shown in **Fig. 9**, the generated orthographic image and the ground truth map are in good agreement. The maximum error $Error_{\max}$ of a 260 m long roadway is only 0.4 m. Such limited errors will be further reduced by integrating all of the local images of the same segment from multiple runs.

3. 3 Lane Graph Construction

An example result of the proposed lane graph construction system is shown in **Fig. 10**. From the results, we can see that the proposed approach can

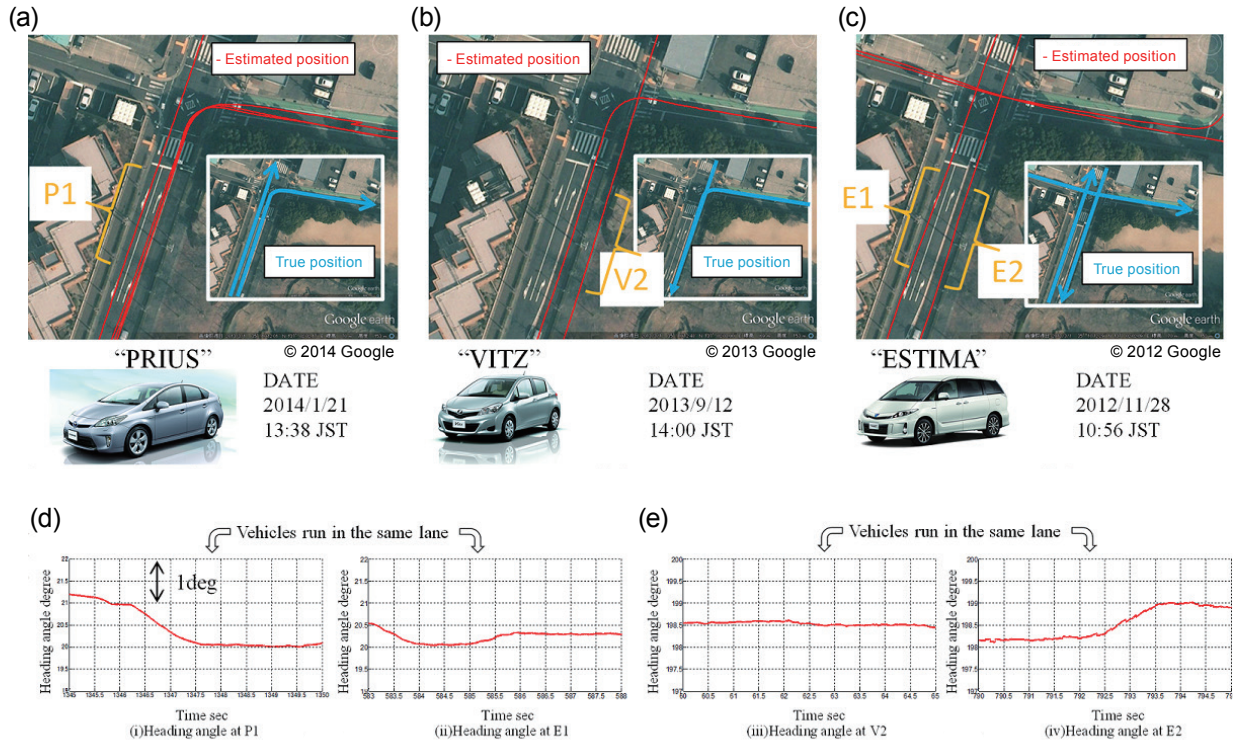


Fig. 7 Data collection with middle-sized (a), small-sized (b) and large-sized (c) probe cars, and comparisons of heading angles between the trajectories of different vehicles driving in the same lanes, i.e., P1 and E1 (d), as well as V2 and E2 (e).

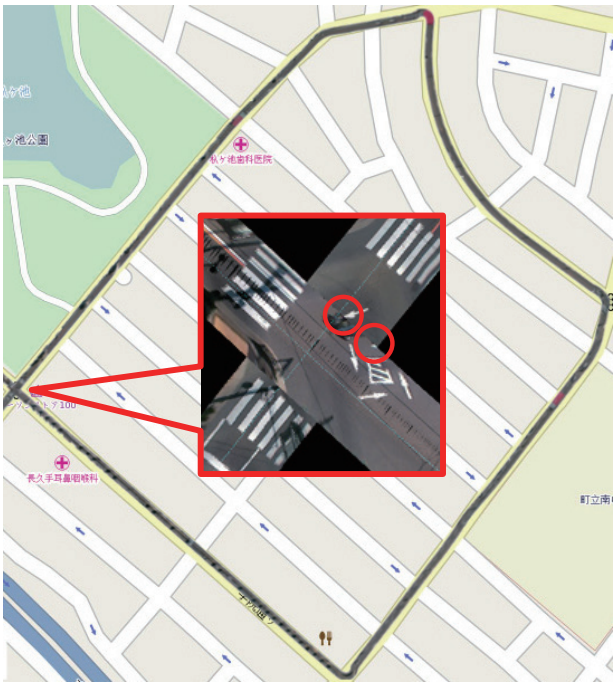


Fig. 8 Example result of image alignment in a 1636 m driving loop.

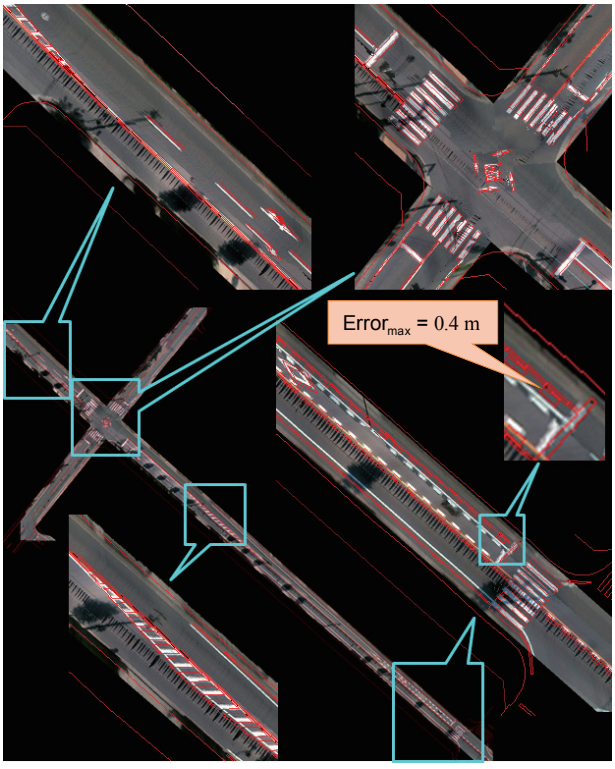


Fig. 9 Comparison between the generated synthetical orthographic image and the high-precision map.

extract the fundamental properties of the driving lanes efficiently and generate safe and reasonable path for driving through both roadways and intersections. Moreover, **Table 1** gives the quantitative evaluation. The success rate shows the percentage of the number of local segments with correct construction result in the total local segments. Although the success rate of the centerline point extraction is a little bit low due to dashed lane markings and worn or complex lane markings, the rates of lane graph construction are high thanks to the lane fitting procedure which can remove most of the outliers of the extracted centerline points.

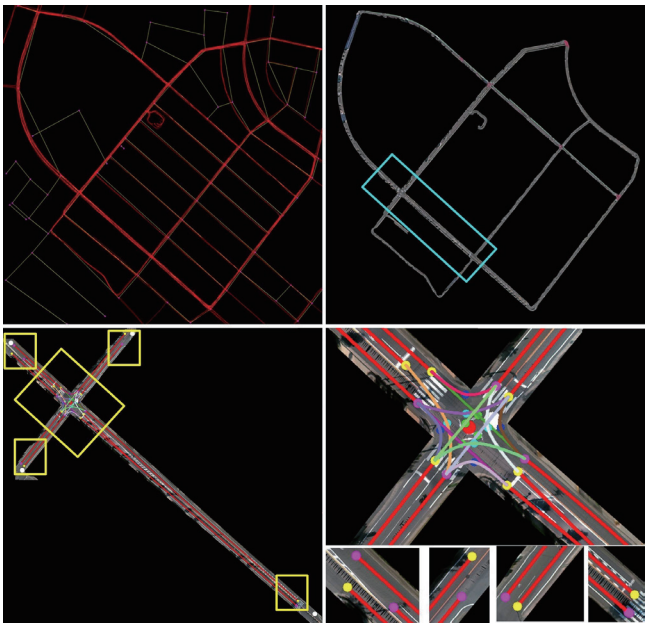


Fig. 10 Example result of the proposed lane graph construction system. Top-left: Trajectories of the vehicles' daily driving. Top-right: Aligned road surface images of the major streets. Bottom-left: Enlarged result of the lane graph generation in the cyan box. Bottom-right: Enlarged results of lane graph in the yellow boxes.

Table 1 Quantitative evaluation of the proposed lane graph construction system.

Step	Success rate
Centerline point extraction	83%
Link lane graph construction	95%
Node lane graph construction	90%

4. Conclusions

In this paper, we propose an approach for automatic generation of lane-level map. The unique features and contributions of this work include the use of conventional low-cost sensors and the automatic manner of lane graph construction for both roadways and intersections. These features result in the proposed low-cost digital map construction system, which can ensure a larger and fast coverage of the accessible area and higher update rate for the up-to-date information. The proposed approach may benefit the deployment of automated cars and thereby enhancing the road safety and comfort for the whole society.

Future work will be mainly focused on the improvement of the exit/entry points determination and the automatic 3D map construction with more environmental prior information.

References

(1) Pink, O. and Stiller, C., "Automated Map Generation from Aerial Images for Precise Vehicle Localization", *Proc. 13th IEEE Conf. Intell. Transp. Syst.* (2010), pp. 1517-1522.

(2) Kummerle, R., Steder, B., Dornhege, C., Kleiner, A., Grisetti, G. and Burgard, W., "Large Scale Graph-based SLAM Using Aerial Images as Prior Information", *Proc. Robotics: Science and Syst.* (2009).

(3) Joshi, A. and James, M., "Generation of Accurate Lane-level Maps from Coarse Prior Maps and Lidar", *Intell. Transp. Syst. Mag.*, Vol. 7, No. 1 (2015), pp. 19-29.

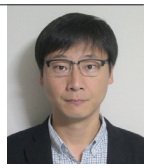
(4) Levinson, J., Askeland, J., Becker, J., Dolson, J., Held, D., Kammel, S., Kolter, J., Langer, D., Pink, O., Pratt, V., Sokolsky, M., Stanek, G., Stavens, D., Teichman, A., Werling, M. and Thrun, S., "Towards Fully Autonomous Driving: Systems and Algorithms", *Proc. Intell. Vehicles Symp.* (IV) (2011), pp. 163-168.

(5) Betaille, D. and Toledo-Moreo, R., "Creating Enhanced Maps for Lanelevel Vehicle Navigation", *IEEE Trans. Intell. Transp. Syst.*, Vol. 11, No. 4 (2010), pp. 786-798.

(6) Zott, C., Yuen, S. Y., Brown, C. L., Bertels, C., Papp, Z. and Netten, B., "Safespot Local Dynamic Maps: Context-dependent View Generation of a Platformfs State and Environment", *Proc. Intell. Transp. Syst. World Congr.* (2008), No. 20260.

(7) Wevers, K. and Dreher, S., "Digital Maps for Lane Level Positioning", *15th Intell. Transp. Syst. World Congr.*, New York, Nov. (2008), No. 20487.
(8) "Mobile Mapping System (MMS)", <http://www.aisantec.co.jp/english/mms.html>, (accessed 2017-12-01).
(9) Jo, K. and Sunwoo, M., "Generation of a Precise Roadway Map for Autonomous Cars", *IEEE Trans. Intell. Transp. Syst.*, Vol. 15, No. 3 (2014), pp. 925-937.
(10) Schindler, A., Maier, G. and Janda, F., "Generation of High Precision Digital Maps Using Circular Arc Splines", *Proc. IEEE Intell. Vehicles Symp.* (2012), pp. 246-251.
(11) *OpenStreetMap Foundation*, <http://www.openstreetmap.org/>, (accessed 2017-12-01).
(12) Kojima, Y., Suzuki, N., Hattori, Y. and Teramoto, E., "Proposal for a New Localisation Method Using Tightly Coupled Integration Based on a Precise Estimation of Trajectory from GPS Doppler", *Vehicle Syst. Dyn.*, Vol. 50, No. 6 (2012), pp. 987-1000.
(13) Levenberg, K., "A Method for the Solution of Certain Non-linear Problems in Least Squares", *Q. Appl. Math.*, Vol. 2, No. 2 (1944), pp. 164-168.
(14) Baass, K., "The Use of Clothoid Templates in Highway Design", *Transp. Forum,* Vol. 1 (1984), pp. 47-52.
(15) Guo, C., Mita, S. and McAllester, D., "Lane Detection and Tracking in Challenging Environments Based on a Weighted Graph and Integrated Cues", *Proc. 2010 IEEE/RSJ Intel. Conf. Intell. Robots and Syst.* (2010), pp. 5543-5550.
(16) Zhu, S. and Yuiller, A., "FORMS: a Flexible Object Recognition and Modeling System", *Proc. IEEE Int. Conf. Comput. Vision* (1995), pp. 465-472.

Figs. 2-10
Reprinted from Trans. IEEE Intell. Transport. Syst., Vol. 17, No. 8 (2016), pp. 2355-2366, Guo, C., Kidono, K., Meguro, J., Kojima, Y., Ogawa, M. and Naito, T., A Low-cost Solution for Automatic Lane-level Map Generation Using Conventional In-car Sensors,



Chunzhao Guo
Research Fields:
- Machine Learning
- Computer Vision
- Intelligent Vehicles
- Cooperative Driving
- Robotics

Academic Degree: Dr.Eng.
Award:
- IAPR Best Application Paper Award, Pacific Rim Symposium on Image and Video Technology, 2015

Kiyosumi Kidono
Research Fields:
- Computer Vision
- Intelligent Robot
- ITS

Academic Degree: Dr.Eng.
Academic Societies:
- The Institute of Electronics, Information and Communication Engineers
- The Robotics Society of Japan

Awards:
- SSII Audience Award, Symposium on Sensing via Image Information, 2012
- SSII Best Paper Award, Symposium on Sensing via Image Information, 2013
- IAPR Best Application Paper Award, Pacific Rim Symposium on Image and Video Technology, 2015

Junichi Meguro
Research Fields:
- Land Vehicle Positioning and Navigation
- Road Environment Reconstruction

Academic Degree: Dr.Eng.
Academic Societies:
- The Institute of Navigation
- The Robotics Society of Japan
- Information Processing Society of Japan

Awards:
- SICE Paper Award (Tomoda Award), 2006
- IPSJ Yamashita SIG Research Award, 2011

Present Affilication: Meijo University