2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Marker-Based Mapping and Localization for Autonomous Valet Parking

Zheng Fang*, Yongnan Chen, Ming Zhou and Chao Lu

***Abstract*— Autonomous valet parking (AVP) is one of the most important research topics of autonomous driving in low-speed scenes, with accurate mapping and localization being its key technologies. The traditional visual-based method, due to the change of illumination and appearance of the scene, easily causes localization failure in long-term applications. In order to solve this problem, we introduce visual fiducial markers as artificial landmarks for robust mapping and localization in parking lots. Firstly, the absolute scale information is acquired from fiducial markers, and a robust and accurate monocular mapping method is proposed by fusing wheel odometry. Secondly, on the basis of the map of fiducial markers that are sparsely placed in the parking lot, we propose a robust and efficient filtering-based localization method, which realizes accurate real-time localization of vehicles in parking lot. Compared with the traditional visual localization methods, we adopt artificial landmarks, which have strong stability and robustness to illumination and viewpoint changes. Meanwhile, because the fiducial markers can be selectively placed on the columns and walls of the parking lot, it is not easy to be occluded compared to the ground information, ensuring the reliability of the system. We have verified the effectiveness of our methods in real scenes. The experiment results show that the average localization error is about 0.3 m in a typical autonomous parking operation at a speed of 10km/h.**

I. INTRODUCTION

Autonomous valet parking is one of the most important research topics of autonomous driving in low-speed scenes. With the increasing density of vehicles in the city, parking space is tight and accidents are frequent during parking operations [1]. Autonomous valet parking technology can help realize high density parking, make full use of limited parking space, reduce accidents caused by human errors during parking, and also bring great convenience to drivers. After the vehicle is switched to AVP mode, it will automatically enter the parking lot to look for free parking spaces and park into any parking space available. However, the technology is quite far from mature yet and there are still many problems to be solved. One of the key problems is the lack of robust and accurate localization information in the absence of GNSS signals [2]. The schemes of traditional indoor localization technology , such as localization based on UWB, fixed laser scanners and other sensors [3], require a large number of high-cost modifications to the environment. Among SLAM-based self-localization methods, visual methods are preferred for its low cost compared to laser methods.

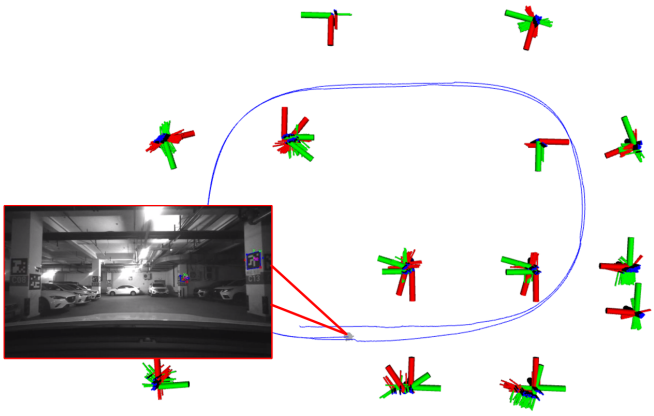

Fig. 1: The smaller figure in the bottom left corner shows a common scene in underground parking lot with complex illumination condition. Fiducial markers are applied to pillars and walls in this scene. As the bigger figure shows, our system realizes robust and accurate real-time localization in parking lot by fusing marker detections(as visualized in smaller figure) and odometry, with the help of a previously-built map of markers.

Visual SLAM can accurately estimate the current camera pose and establish the corresponding environmental map. ORB-SLAM2 [4], [5] and other feature-based methods have good results in the scene with rich texture. However, these methods suffer from environment appearances changes and complex illumination conditions. Thus, these methods could only provide visual maps that need to be established within a short period of time of localization usage, lacking long-term stability and practicability.

Fiducial marker [6] is a commonly used landmark, which is often used to estimate the pose of robots [7]. Compared with traditional geometric features, the fiducial marker has strong adaptability to illumination changes [8] and has larger identifiable angle range. In this paper, we propose a mapping and localization system based on fiducial markers, and utilize fiducial markers that are sparsely placed in real scenes and low-cost processors to realize accurate mapping and localization. Because the fiducial markers can be selectively placed on the columns and walls of the parking lot, it is not easy to be obscured compared to the ground information, which can ensure the reliability of the system. Also, due to the adoption of visual markers, this system only needs a low-performance ARM processor to realize robust localization,

This work is supported by the Major Science and Technology Projects Fundation of Liaoning Province (2019JH1/10100026), the Fundamental Research Funds for the Central Universities(No.N182608003, No.N172608005), Natural Science Foundation of Liaoning Province(20180520006)

Zheng Fang, Yongnan Chen, Ming Zhou and Chao Lu are with the Faculty of Robot Science and Engineering, Northeastern University, Shenyang, China; Corresponding author: Zheng Fang, e-mail: fangzheng@mail.neu.edu.cn

which lays a foundation for the practicality of the system. The method proposed in this paper can establish a long-term stable and reusable parking lot map and provide accurate localization information for vehicles. We test the system on actual vehicles to verify the effectiveness and accuracy of our method. The experimental results show that the average localization error of our proposed methods is about 0.3m in the low-speed parking process with a vehicle speed of 10km/h. In summary, our main contributions are:

- Propose a robust and accurate marker-based mapping method by fusing scale information extracted from fiducial markers with odometry and feature points;
- Based on sparse fiducial marker map, propose a robust localization method with low computational resource consumption, by fusing marker detection and wheel odometry with a particle filter;
- Experiments in real scenes are carried out to verify the validity of our methods.

The rest of the paper is organized as follows. Section II describes the related works. Section III details the proposed mapping and localization methods. We validate our method in Section IV. Finally, Section V concludes the paper.

II. Related Works

In the past ten years, there have been many visual mapping and localization works in the field of AVP. According to the different emphases of these methods, we classify them into mapping and localization methods. Due to parking lots usually being private area, there are generally no maps established in advance, and vehicles need to establish their own maps. Visual SLAM-based mapping method is one of the commonly used methods. The V-Charge project [9] uses SFM framework to build a three-dimensional map of the environment through images collected by the multi-camera system configured by the vehicle. Chirca et al. [10] created a three-dimensional map of the environment through EKF-based visual SLAM. However, the above methods all utilize traditional geometric features, such as sparse points and straight lines in the environment. These traditional feature-based methods will be affected by changes in illumination, viewpoint and appearance when used in long term. In order to overcome these influences, high-dimensional environmental features are used for mapping. Huang et al [11] extract the ID information of the parking space through the fisheye camera, and established the semantic map of the parking lot environment by combining the monocular camera, wheel odometry and IMU. However, the parking space information in the parking lot is easily blocked by vehicles. Huang et al. additionally introduced visual tags to assist in localization. Similarly, Zong et al. [12] also introduce visual tags, combine with vehicle kinematics model, to improve the performance of ORB-SLAM in underground parking lots. In addition, road-based semantic features [13], [14], such as lane lines, speed bumps, turn signs and other features, are also applied to the mapping system. However, most of these ground semantic features may suffer from occlusion or be worn out in usage, which can lead to system failure. The computational consumption is also relatively high compared with traditional methods.

The vision-based localization methods [15]–[17] use the established map to obtain the pose of the camera relative to the map through descriptor matching. However, they are subject to localization failures in indoor parking lots and other low illumination environments [18]. Jeevan et al. [19] proposed a localization method, which fuse fiducial markers placed on the ground and wheel odometry. Compared with the feature-based method, it is more robust, but the map is generated by georeferencing each marker with GPS, thus the mapping can only be applied to outdoor scenes. For indoor parking lots, Qin et al. [20] utilize a variety of road semantic features and combined with wheel odometry to achieve centimeter-level parking accuracy. However, this method puts forward higher requirements for onboard hardware (high-performance processors, high-resolution cameras, etc.)

III. APPROACH

In this paper, we use monocular camera to get the image information. The monocular camera is installed to the center of the vehicle, behind the windshield to capture front-view scenes. Vehicle odometry information formed by steering wheel angle and vehicle speed is also used in our system. Intrinsic and extrinsic parameters of all sensors were calibrated offline in advance.

The framework consists of two parts, as shown in Fig.2. The first part is mapping, in which we use the front-view monocular camera to detect fiducial markers, extract scale and pose information and then fuse with odometry data to build a global fiducial marker map. This marker map is saved for localization. Then the vehicle is localized by matching fiducial markers extracted from monocular image to the marker map. In the end, a particle filter fuses visual localization results with odometry, which guarantees the system survives in the marker-less region and has a smooth output.

A. Mapping with Fiducial Markers and Vehicle Odometry

The proposed mapping method contains three main modules: tracking, local mapping and loop closing [21]. In the initialization part of the tracking module, we use fiducial markers to recover the scale of monocular camera. In the local mapping module, the map is extended by adding newly observed markers and new map points. In addition, the poses of local keyframes and local map points are optimized jointly in this module, called local bundle adjustment(BA). Accumulated drift will be eliminated by loop closing.

1) Scale Recovery From Visual Fiducial Markers: There are many different kinds of visual fiducial markers. We choose ArUco marker in our system due to its robustness and high-efficiency and it is included in OpenCV [22].

Adding the ArUco marker to monocular SLAM [23], [24] solves the problem of scale ambiguity. At initialization, we can recover the scale factor s of the monocular camera trajectory by obtaining the same ArUco Marker observed

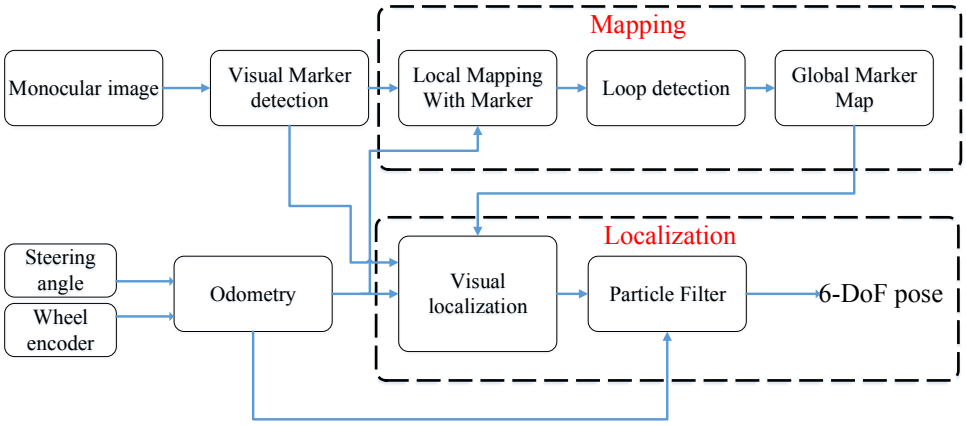

Fig. 2: Block diagram illustrating the full pipeline of the proposed system. In the mapping procedure, it builds a map of the large-scale indoor parking lot with fiducial markers. Based on this prior map, the localization procedure can provide precise 6-DoF pose through a particle filter fusing fiducial markers with odometry.

by both keyframes as follows.

$$R_{WC}^{k+1}R_{CM_i}^{k+1}t_{M_iC}^{k+1} - R_{WC}^{k}R_{CM_i}^{k}t_{M_iC}^{k} = s(t_{WC}^{k+1} - t_{WC}^{k}) \quad (1)$$

where $[R_{CM_i}^{k}\ t_{CM_i}^{k}]$ is the pose from ith marker to camera at frame k and $[R_{WC}^{k}\ t_{WC}^{k}]$ is the pose from camera to world at frame k.

2) Pose Optimization with Vehicle Odometry Constraints: Most cars are equipped with wheel encoders [25]. In most cases, wheel encoders provide a reliable measurement of the distance traveled by the wheel. In our case, we can directly read the wheel speed v of the rear wheels and the corresponding steering wheel angle δ through the vehicle's CAN bus. Then we can get the pose R_{WV}^{k+1} and t_{WV}^{k+1} of the vehicle for frame $k+1$ according to the vehicle odometry as Equation 2.

$$\begin{cases} x_{k+1} = x_k + \Delta x cos(\theta_k) - \Delta y sin(\theta_k) \\ y_{k+1} = y_k + \Delta x sin(\theta_k) + \Delta y cos(\theta_k) \\ \theta_{k+1} = \theta_k + \Delta\theta \end{cases} \quad (2)$$

Therefore, in local map optimization, we additionally introduce vehicle odometry error term on top of the reprojection error term by Equation 3

$$\begin{aligned} &\gamma = \{R_{CW}^{j}, t_{CW}^{j}\} \\ &\gamma^{*} = argmin(\textstyle\sum_k E_{proj}(k,j) + E_{vehicle}(i,j)) \end{aligned} \quad (3)$$

where E_{proj} is the reprojection error of current frame j for given match k . And the vehicle odometry error term $E_{vehicle}$ between keyframe i and j is denoted by Equation 4.

$$\begin{aligned} &E_{vehicle}(i,j) = \rho([e_R^T e_t^T]\textstyle\sum_I [e_R^T e_t^T]^T) \\ &e_R = Log((Exp(w_v \Delta t(i,j))^T R_{VW}^{i} R_{WV}^{j}) \\ &e_p = R_{VW}^{i}(t_{WV}^{j} - t_{WV}^{i}) - v_{WV}^{i}\Delta t \end{aligned} \quad (4)$$

where $\rho(\cdot)$ is the robust Huber cost function, $\sum_I$ is the information matrix of vehicle odometry error term.

B. Marker Map-based Real-time Localization

To better suit the need of performing real-time localization on automotive-grade embedded processors, we propose a particle filter-based method to fuse visual and odometry information for localization in indoor parking zone. The marker map we use is created in previous part.

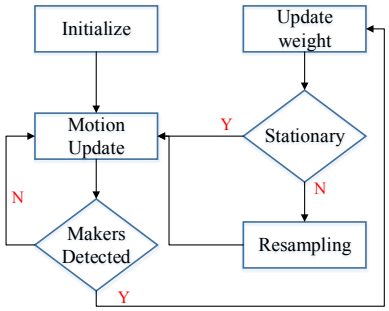

Fig. 3: Localization Algorithm Structure

1) Initialization: Our system requires at least one marker detection to initialize. After localization system is started, markers in surrounding area are detected by vision and their IDs and relative poses to vehicle are used for initialization. With known ID, a marker's absolute pose can be acquired from marker map created earlier by matching its ID to the marker with same ID in the map. Then the vehicle's initial pose in map coordinate can be calculated from marker's absolute pose and marker's relative pose to vehicle. Being vehicle pose in 2D space (x, y, θ), for kth marker detected during initialization with a relative pose to vehicle as (x', y', θ'), its absolute pose in map coordinate being $m_k = [x_k\ y_k\ \theta_k]^T$, then

$$\begin{bmatrix} x_0 \\ y_0 \\ \theta_0 \end{bmatrix} = \begin{bmatrix} x_k - x'cos\theta' + y'sin\theta' \\ y_k - x'sin\theta' + y'cos\theta' \\ \theta_k - \theta' \end{bmatrix} \quad (5)$$

where $X_0 = [x_0\ y_0\ \theta_0]^T$ is vehicle's initial pose under map coordinate.

To improve initialization accuracy, we do this calculation multiple times and use the results' average as vehicle initial pose, and then generate our set of particles with different poses (x_n, y_n, θ_n) around this pose according to normal distribution, where $(std_x, std_y, std_\theta)$ are preset initialization parameters.

2) Motion Update: By using wheel odometry we can obtain vehicle's relative movements sequentially and use them to perform a prediction update to our particles. Because in our system the frequency of marker observations is significantly lower than that of odometry feedbacks, so the motion and observation could be considered separately while updating state probability.

If at moment t-1 the nth particle's pose state is $X_{t-1}^n = [x_{t-1}^n \ y_{t-1}^n \ \theta_{t-1}^n]^T$, then after a motion update its state is

$$\begin{gathered}\begin{bmatrix} x_t^n \\ y_t^n \\ \theta_t^n \end{bmatrix} = \begin{bmatrix} \mathrm{N}\left(0, x_\sigma^2\right) \\ \mathrm{N}\left(0, y_\sigma^2\right) \\ \mathrm{N}\left(0, \theta_\sigma^2\right) \end{bmatrix} + \\ \begin{bmatrix} x_{t-1}^n + \cos\left(\theta_{t-1}^n + \Delta\theta\right)\Delta x - \sin\left(\theta_{t-1}^n + \Delta\theta\right)\Delta y \\ y_{t-1}^n + \sin\left(\theta_{t-1}^n + \Delta\theta\right)\Delta x \\ \theta_{t-1}^n + \Delta\theta \end{bmatrix}\end{gathered} \tag{6}$$

where $u_t = [\Delta x \ \Delta y \ \Delta\theta]^T$ are translation and rotation increments measured by wheel odometry, and $(x_\sigma, y_\sigma, \theta_\sigma)$ are preset motion noises. Motion noises is estimated by experiments and in our case set to $(0.005m, 0.005m, 0.001rad)$.

3) Observation Update: While doing motion updates the marker detector is detecting markers in monocular images at the same time. Due to the pose ambiguity problem, marker detection's accuracy is slightly lower on longer distance, so we only take detections within a distance threshold($10m$). Once one or more such detections are returned, observation update is carried out using these detections.

To evaluate the error of each particle's pose to vehicle's actual pose, based on relative pose of marker to vehicle measured by marker detector and each particle's pose state, we calculate the marker poses observed from each particle and compare them to this marker's actual pose in the map. If the relative pose of a marker to vehicle is measured as $Z_t = [x_{ob} \ y_{ob} \ \theta_{ob}]^T$, then the nth particle in particle set, with a pose state of $X_n = [x_n \ y_n \ \theta_n]^T$, the marker pose observed from this particle under map coordinate is:

$$\begin{bmatrix} x_{ob\_\mathrm{map}} \\ y_{ob\_\mathrm{map}} \\ \theta_{ob\_\mathrm{map}} \end{bmatrix} = \begin{bmatrix} x_{ob}\cos\theta_n - y_{ob}\sin\theta_n + x_n \\ x_{ob}\sin\theta_n + y_{ob}\cos\theta_n + y_n \\ \theta_{ob} + \theta_n \end{bmatrix} \tag{7}$$

The particle's weight can be calculated using error between this pose and the marker's real pose, thus the weight w_n of nth particle is:

$$w_n = \frac{e^{-\left(\left(\left(x_{ob\_map} - x_{ob}\right)^2/\sigma_x^2\right) + \left(\left(y_{ob\_map} - y_{ob}\right)^2/\sigma_y^2\right)\right)/2}}{2\pi\sigma_x\sigma_y} \tag{8}$$

where σ_x and σ_y are observation noises, set to 0.3m and 0.3m in our case respectively. They are set slightly bigger than translational errors of marker detections intendedly. Considered that markers in parking garage is relatively sparse, this can help the system correct odometry accumulation errors more gradually and avoid local sharp changes in pose output, which may have negative effects on motion control and path planning. Then we complete weights update by normalize weights of the whole particle set:

$$w_{\mathrm{norm}} = \frac{w_{1:n}}{\sum_{n=1}^{n=num} w_n} \tag{9}$$

Where num is particles' number and w_{norm} is the array of normalized particle weights.

4) Resampling: With particle weights updated, we firstly check vehicle's moving state through latest odometry readout. If odometry shows that vehicle is stationary, we keep the particle weights update without resampling because vehicle's pose is not supposed to change at this moment. If the vehicle is moving, we resample the particles by their weights and reinitialize weights of the new particle set as:

$$w_{1:num} = \frac{1}{num} \tag{10}$$

The X_t probability distribution is approximated by the new particle set.

Finally, we output vehicle pose every time system state is updated. To smooth estimated trajectory and reduce pose jumps, we choose average pose of all particles $(x_{avg}, y_{avg}, \theta_{avg})$ as the final pose output.

IV. Experiments and Analysis

In order to verify the validity of the system proposed in this paper, we conducted separate experiments for mapping and real-time localization. The experiment environment is an underground parking garage with uneven lighting and an area of about 500 m^2. Twenty markers of 0.552m*0.552m size are placed on the walls and pillars of parking lots where they are easy to be observed and not easily to be obscured by the vehicles in the parking lot. The average interval between the markers is about 8m, excluding the case where there are multiple markers on different sides of the same pillar. The vehicle is equipped with two wheel encoders, an Intel RealSense D435i camera , VLP16 LiDAR and an embedded platform with Ubuntu 16.04. The vehicle travels at a constant speed of 10km/h. The video link for a demonstration of the proposed system is: `https://youtu.be/11r3eRAjFVA`

A. Mapping Metric Evaluation

For mapping metric evaluation, considering the high accuracy and robustness and maturity of the 3D laser SLAM algorithm in indoor scenarios, we collect laser point cloud data during the experiment. We use the Lego-Loam [26] algorithm to process the acquired data and treat the resulting laser trajectory as ground truth. The total trajectory length is 143 m. We recorded the camera trajectory as well as the laser trajectory. Due to the uniqueness of our sensor configuration, it is hard to directly compare against other existing algorithms. We compared our method with ground truth in terms of mapping accuracy.

The mapping result and estimated trajectories are shown in Fig. 4. The RMSE of absolute trajectory error is 0.438m and

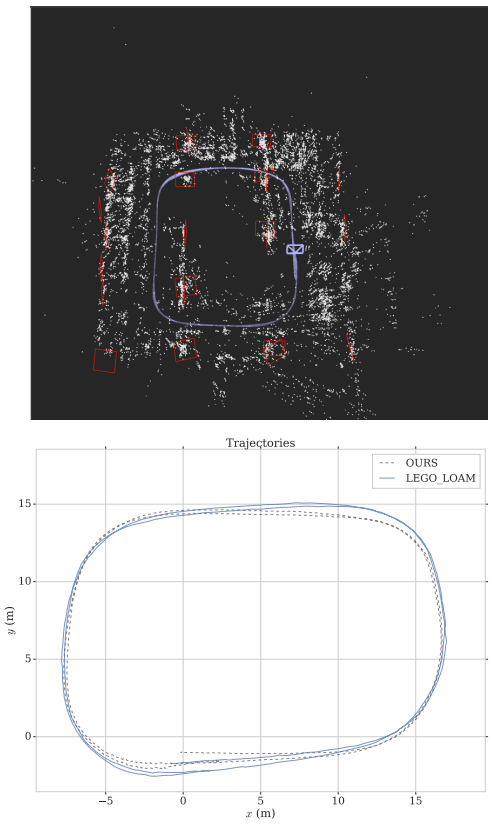

Fig. 4: (top) Map of the parking lot. The red squares are visual fiducial markers. (bottom) Estimated trajectories by our methods and the ground truth.

the normalized estimation error squared(NESS) is 0.306%. It can be seen that our algorithm performs well by fusing feature points, visual fiducial markers and vehicle odometry.

B. Localization Accuracy Evaluation

After the marker map is created, we performed real-time localization experiments in underground parking garage mentioned above, as shown in Fig. 1, where the blue line is estimated trajectory, the bigger axes show the absolute poses of markers in the map and smaller axes show the poses of markers observed from different particles. We also use trajectories estimated by laser SLAM as ground truth to evaluate accuracy of our localization method. The results of two independent experiments are shown in Fig. 5, Fig. 6 and TABLE I.

TABLE I: Errors in two experiments

Error	Mean[m]	Max[m]	Min[m]	RMSE[m]
Experiment 1	0.301	0.775	0.0153	0.347
Experiment 2	0.264	0.687	0.0248	0.307

As shown above, our method is low on most of the errors, with an average error around 0.3m. The localization performance is also stable throughout the whole trajectory. Due to the fact that marker landmarks are sparsely distributed in the parking garage(as stated above, the average distance between markers is 8m), localization error at some places with few or no markers detectable will be slightly bigger, especially during turns(as shown at the upper right and bottom left corners of trajectories in Fig. 5), but these errors are still acceptable and could be corrected quickly as the trajectories show.

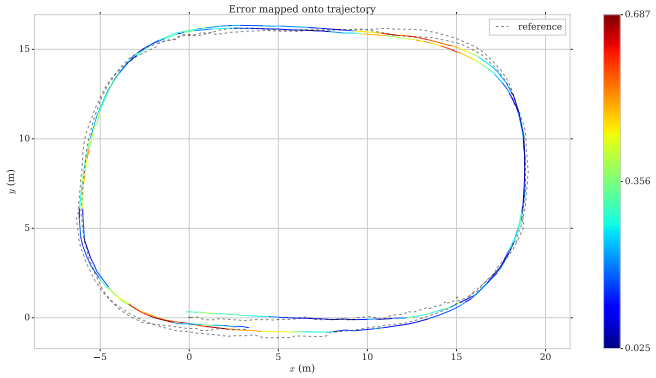

Fig. 5: Comparison of marker-based localization and ground truth(grey line)

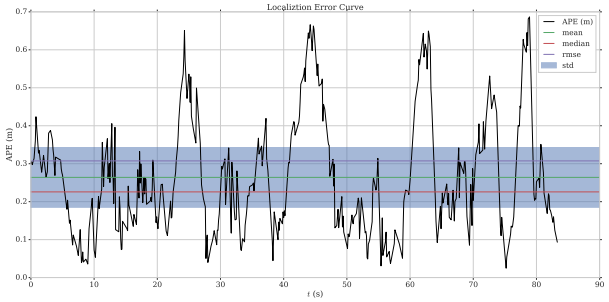

Fig. 6: Localization error graph of two experiments respectively

TABLE II: Running performance of our method on different hardware

	CPU occupation	Memory used[MB]	Frequency[Hz]
i7 laptop	14%	500	100
A53 embedded	25%	510	100

C. Computational Resource Demand

As mentioned above, our localization method is developed for online usage on intelligent vehicle's onboard embedded processor, so the algorithm's computational resource demand needs to be as low as possible. We tested our algorithm on 8-core i7-7700HQ equipped laptop and 4-core A53 equipped embedded platform respectively. As shown in TABLE II, while performing localization successfully, the computational resource consumption of our method is also suitable for real-time application on intelligent vehicles.

D. System Robustness

To verify our method's environment robustness over feature point-based methods, we tested these methods in long-term localization. Experiments showed that after significant changes occurred in operation environment, appearance changes in certain locations will lead to false feature point matches to map(as shown in Fig. 7), causing localization failures. On contrary, because marker detections are not affected by appearance changes of surrounding area(as shown in Fig. 8), localization based on marker map is still robust and effective.

Fig. 7: The false feature point matches of the same place at different time due to appearance changes

Fig. 8: Marker detections are not affected by changes in surrounding area

V. CONCLUSIONS

In order to realize the robust localization for autonomous parking in underground parking lots, we introduce visual fiducial marker as a stable artificial landmark to establish a robust and long-term usable map. On this basis, an efficient localization algorithm based on particle filter is proposed to perform robust and accurate localization. However, The method we proposed still requires manually placing markers in the parking lot. In the future, we plan to replace fiducial markers with the existing text landmarks in the parking lot to further improve the practicability of our system.

REFERENCES

[1] H. Banzhaf, D. Nienhuser, S. Knoop, and J. Marius Zollner, "The future of parking: A survey on automated valet parking with an outlook on high density parking," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 1827–1834, 2017.

[2] B. Li, L. Yang, J. Xiao, R. Valde, M. Wrenn, and J. Leflar, "Collaborative Mapping and Autonomous Parking for Multi-Story Parking Garage," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1629–1639, 2018.

[3] H. M. Hussien, Y. N. Shiferaw, and N. B. Teshale, "Survey on indoor positioning techniques and systems," in *Information and Communication Technology for Development for Africa*, pp. 46–55, Springer, 2018.

[4] R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[5] J. B. M. L. Felix Nobis, Odysseas Papanikolaou, "Persistent map saving for visual localization for autonomous vehicles: An orb-slam extension," in *International Conference on Ecological Vehicles and Renewable Energies (EVER)*, 2020.

[6] F. J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, pp. 38–47, 2018.

[7] H. Lim and Y. S. Lee, "Real-time single camera slam using fiducial markers," in *2009 ICCAS-SICE*, pp. 177–182, IEEE, 2009.

[8] D. Hu, D. Detone, and T. Malisiewicz, "Deep charuco: Dark charuco marker pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8436–8444, 2019.

[9] U. Schwesinger, M. Burki, J. Timpner, S. Rottmann, L. Wolf, L. M. Paz, H. Grimmett, I. Posner, P. Newman, C. Hane, L. Heng, G. H. Lee, T. Sattler, M. Pollefeys, M. Allodi, F. Valenti, K. Mimura, B. Goebelsmann, W. Derendarz, P. Muhlfellner, S. Wonneberger, R. Waldmann, S. Grysczyk, C. Last, S. Bruning, S. Horstmann, M. Bartholomaus, C. Brummer, M. Stellmacher, F. Pucks, M. Nicklas, and R. Siegwart, "Automated valet parking and charging for e-mobility," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 157–164, 2016.

[10] M. Chirca, R. Chapuis, and R. Lenain, "Autonomous Valet Parking System Architecture," in *Proceedings of IEEE Conference on Intelligent Transportation Systems, ITSC*, pp. 2619–2624, 2015.

[11] Y. Huang, J. Zhao, X. He, S. Zhang, and T. Feng, "Vision-based Semantic Mapping and Localization for Autonomous Indoor Parking," in *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 636–641, 2018.

[12] W. Zong, L. Chen, C. Zhang, Z. Wang, and Q. Cheny, "Vehicle model based visual-tag monocular ORB-SLAM," in *2017 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2017*, pp. 1441–1446, 2017.

[13] A. Ranganathan, D. Ilstrup, and T. Wu, "Light-weight localization for vehicles using road markings," in *IEEE International Conference on Intelligent Robots and Systems*, pp. 921–927, 2013.

[14] Y. Lu, J. Huang, Y. T. Chen, and B. Heisele, "Monocular localization in urban environments using road markings," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 468–474, 2017.

[15] P. Mühlfellner, M. Bürki, M. Bosse, W. Derendarz, R. Philippsen, and P. Furgale, "Summary Maps for Lifelong Visual Localization," *Journal of Field Robotics*, vol. 33, no. 5, pp. 561–590, 2016.

[16] H. Lategahn, M. Schreiber, J. Ziegler, and C. Stiller, "Urban localization with camera and inertial measurement unit," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 719–724, 2013.

[17] J. Ziegler, H. Lategahn, M. Schreiber, C. G. Keller, C. Knoppel, J. Hipp, M. Haueis, and C. Stiller, "Video based localization for Bertha," in *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 1231–1238, 2014.

[18] P. Nelson, W. Churchill, I. Posner, and P. Newman, "From dusk till dawn: Localisation at night using artificial light sources," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 5245–5252, 2015.

[19] P. Jeevan, F. Harchut, B. Mueller-Bessler, and B. Huhnke, "Realizing autonomous valet parking with automotive grade sensors," in *The IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3824–3829, 2010.

[20] T. Qin, T. Chen, Y. Chen, and Q. Su, "AVP-SLAM: Semantic Visual Mapping and Localization for Autonomous Vehicles in the Parking Lot," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, accepted.

[21] S. Sumikura, M. Shibuya, and K. Sakurada, "OpenVSLAM: A versatile visual SLAM framework," in *Proceedings of the 27th ACM International Conference on Multimedia*, pp. 2292–2295, 2019.

[22] G. Bradski, "The OpenCV Library," *Dr Dobbs Journal of Software Tools*, vol. 25, 2000.

[23] R. Munozsalinas, M. J. Marinjimenez, E. Yeguasbolivar, and R. Medinacarnicer, "Mapping and localization from planar markers," *Pattern Recognition*, vol. 73, pp. 158–171, 2018.

[24] R. Muñoz-Salinas and R. Medina-Carnicer, "Ucoslam: Simultaneous localization and mapping by fusion of keypoints and squared planar markers," *Pattern Recognition*, vol. 101, p. 107193, 2020.

[25] H. Banzhaf, D. Nienhuser, S. Knoop, and J. Marius Zollner, "The future of parking: A survey on automated valet parking with an outlook on high density parking," pp. 1827–1834, 2017.

[26] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4758–4765, IEEE, 2018.