

GROUND VEHICLE LOCALIZATION WITH PARTICLE FILTER BASED ON SIMULATED ROAD MARKING IMAGE

Oleg Shipitko and Anton Grigoryev
Institute for Information Transmission Problems – IITP RAS,
Bol'shoy Karetnyy Pereulok 19,
Moscow, Russia, 127051
E-mail: shipitko@visillect.com

KEYWORDS

Model-based localization, vehicle localization, particle filter, lane markings, autonomous vehicle, pose estimation.

ABSTRACT

Precise localization is a prerequisite and a cornerstone for successful operation of any autonomous vehicle. In this paper, consideration is given to a lane feature-based approach to a self-driving vehicle localization. Proposed map-relative localization method is built upon a combination of vision-based lane markings detection and odometry data. Detected lane markings are aligned with a reference map in order to derive global pose estimate while odometry provides path consistency. To combine heterogeneous sensory data use is made of particle filter method. It allows for non-Gaussian noise common for vision-based detectors as well as a further extension of data sources. The approach described in this work was tested on a real vehicle in urban environment and proved itself to be precise and reliable enough for real-world applications. It was able to provide lateral and longitudinal map-relative localization with a precision of 0.2 m.

INTRODUCTION

Recent advances in autonomous vehicle development have drawn attention of a wide audience including scientists, engineers, and the general public. Although nowadays there are already ongoing experiments on autonomous driving in the natural environment, the increasing safety requirements lead to the high demand for improvement of all vehicle subsystems. Precise and reliable localization is a prerequisite for any complex task such as path planning, maneuvering, and navigation in general.

It has been long understood that lane markings in particular and all types of road markings in general contain essential information used by human drivers for decision making and road situation analysis. The idea to provide autonomous vehicles with the same capabilities is not new as well. Lane detection is widely used in both partially and fully autonomous vehicle projects for various purposes (Krokhina et al. 2015; Lu et al. 2014). Hillel et al. (Hillel et al. 2014) provides an extensive overview of lane detection techniques, their limitations, and application. One of the applications where lane detection has proved itself to be an efficient foundation is localization (Ziegler et al. 2014; Jo et al.

2015; Du and Tan 2016). Many systems presented in literature exploit LIDAR as a major source of information for road understanding and lane detection (Huang et al. 2009; Hernández and Marcotegui 2009). One of the subtasks of lane marking detection where LIDAR is superior to other sensors is curb detection. There are many studies of localization systems build upon detector of this type exclusively. Thus in (Hata et al. 2014), authors propose a system using a method named least trimmed squares to fit a curve model to detected curb points and validate the proposed detector in a localization task. However, curbs on their own are prone to occlusion and cannot guarantee high precision in many real-world scenarios, while lane features (which may as well include detection of curbs as shown in (Ziegler et al. 2014)) provide more solid localization foundation. In general, the high (though constantly decreasing) cost of LIDAR systems prevents such kind of systems from becoming wide-spread. On the other hand, vision-based approaches require only relatively inexpensive camera systems and are able to provide a comparable level of detection precision.

The majority of known autonomous ground vehicles use a predefined map as a reference and derive their global pose by associating incoming sensory data with mapped data corresponding to a particular reference pose (Chausse et al. 2005; Tao et al. 2013; Volkov et al. 2017). As shown in (Li et al. 2017), size is an important property of a digital map. For instance, maps built from multidimensional features as (e.g. 3D Point Cloud maps used in (Sheehan et al. 2013)) may require significant storage capacity and relatively high amount of computational resources. At the same time curve-based vector maps require much less storage and are specifically efficient for representing 2D structures like lane markings.

A crucial issue to solve in any localization pipeline is combining together more than one sensor or several concurrent algorithms of vehicle pose estimation. There are two major approaches to data fusion in localization applications, namely Kalman filter, and particle filter (also known as Sequential Monte Carlo localization technique). Both approaches have been used widely in robotics as emphasized in (Thrun 2002) and have their own advantages and drawbacks. Kalman filters provide exact, optimal solution and have fast computational speed, but can handle only linear (or linearized) Gaussian systems (Kalman 1960). Particle filters, on the other hand, provide only approximate so-

lution and in general, require much more computational resources (since probability evaluation is performed for each particle), but allow to accommodate arbitrary noise models. The latter is crucial for a vision-based systems, where detection noise model often cannot be straightforwardly derived from the architecture. Another attractive feature of particle filters is that they are time-independent and thus are applicable in cases with strongly varying sensor sampling time (Thrun 2002).

In this work, we propose a map-relative vision-based localization approach which matches detected lane markings to a reference vector map. A particle filter algorithm is exploited in order to combine sensory data from heterogeneous sources such as lane marking detectors (including a pedestrian crossings detectors), an a priori occupancy map and odometry.

The rest of the paper is structured as follows: section II explains the details of digital map preparation, section III describes generic particle filter algorithm and emphasizes the implementation details of our particular method, and section IV presents and discusses obtained results.

MAP PREPARATION

Predefined map is a crucial component of the proposed localization system. It allows storing information such as lane marking features and occupancy data in a compact format of a vector image. Such format allows for a natural representation of 2D geometric primitives such as lane markings and makes it easy to access and edit an existing map markup.

Although automatic mapping technique can be exploited, current implementation of map generation process requires manual specification of road marking in the form of polylines as shown in Fig. 1. There are no additional properties associated with line segments which significantly simplifies and speeds up the map preparation process. Apart from lanes detected pedestrian crossings are also used as an input to the localization. They are distinguished by the separate detector and approximated by rectangles in the road plane. As well as lines pedestrian crossings are manually marked on a map preparation step. Another important component of the map is the occupancy data marking places with zero probability of the vehicle presence. Accommodation of this information leads to lower requirements to computational speed by reducing the number of considered pose hypotheses.

Once prepared in vector graphic form the map is transformed to a raster image which allows to access any point on the map in $O(1)$ time. Another important property of the raster representation is an ability to explicitly accommodate the measurement noise model as will be shown below. Although a simple geometric representation of mapped lane markings is used in this work, other properties (e.g. line orientation) may be computed offline and encoded in a multi-channel raster image.

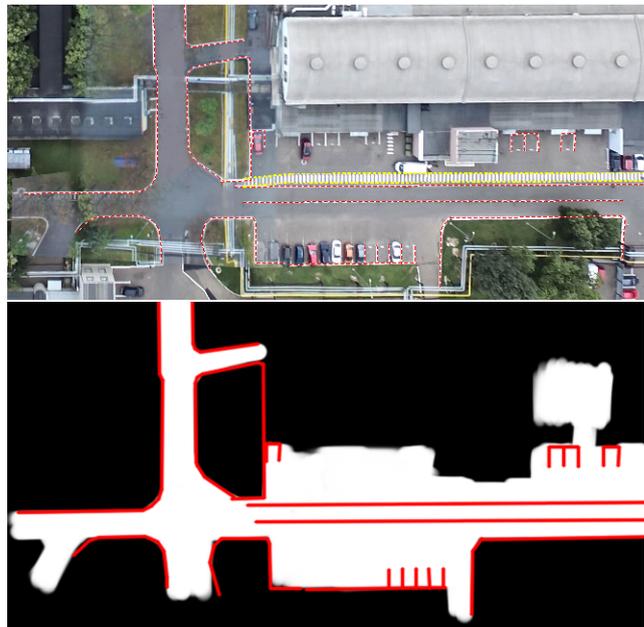


Figure 1. A fragment of a manually prepared map (on the top) and the resulting raster map used by the vehicle (on the bottom). Preprocessed lane markings are shown as red dashed line. Pedestrian crossings are marked as yellow rectangle. The resulting raster markup is shown as solid red lines. Black areas denote occupied space.

PARTICLE FILTERS

In the context of localization, particle filters are also known as Monte Carlo Localization (MCL) technique. They allow to derive the posterior distribution of state variables in a partially observable Markov chains. Let's denote a Markov chain state at time t as \mathbf{x}_t . The state \mathbf{x}_t depends on the previous state \mathbf{x}_{t-1} according to the probabilistic law $p(\mathbf{x}_t|\mathbf{u}_t, \mathbf{x}_{t-1})$, where \mathbf{u}_t is the vector of control signals applied at time $t-1$. In the field of robotics, $p(\mathbf{x}_t|\mathbf{u}_t, \mathbf{x}_{t-1})$ is called an actuation model. The state of partially observable Markov chain can be described by measurement vector \mathbf{z}_t , generated via probabilistic law $p(\mathbf{z}_t|\mathbf{x}_t)$, called a measurement model. The problem being solved by particle filters can be formulated as follows: given all consecutive sensors measurements $\mathbf{z}^t = \mathbf{z}_0, \dots, \mathbf{z}_t$, and control vectors $\mathbf{u}^t = \mathbf{u}_0, \dots, \mathbf{u}_t$ recover posterior distribution of the state \mathbf{x}_t at any given time t (Thrun 2002). The idea behind MCL is to approximate unknown posterior by the set of particles $\{\mathbf{x}_t^n\}$ with associated weight, where $n = 1, \dots, N$ - number of particles used. In our application each particle represents a hypothesis of the whole state vector of the vehicle pose. As discussed in (Thrun 2002), the original particle filter algorithm adopted to field of robotics has the form presented below.

Subsequent sections describe in details implementation specifics of each step of presented general particle filter algorithm.

Algorithm 1 Generic particle filter algorithm

```
1: procedure PARTICLE FILTER( $\mathbf{x}_{t-1}, \mathbf{u}_t, \mathbf{z}_t$ )
2:    $\{\mathbf{x}_t^n\} = \{\widetilde{\mathbf{x}}_t^n\} = \emptyset$ 
3:   for  $n = 1$  to  $N$  do
4:     sample  $x_t^n \sim p(\mathbf{x}_t | \mathbf{u}_t, \mathbf{x}_{t-1}^n)$ 
5:      $w_t^n = p(\mathbf{z}_t | \mathbf{x}_t^n)$ 
6:      $\{\mathbf{x}_t^n\} = \{\mathbf{x}_t^n\} + (x_t^n, w_t^n)$ 
7:   end for
8:   for  $n = 1$  to  $N$  do
9:     draw  $i$  with probability  $\propto w_t^i$ 
10:     $\{\mathbf{x}_t^n\} = \{\mathbf{x}_t^n\} + (\widetilde{\mathbf{x}}_t^i, \widetilde{\mathbf{w}}_t^i)$ 
11:   end for
12:   return  $\{\mathbf{x}_t^n\}$ 
13: end procedure
```

Actuation model

The actuation model used in this work relies upon odometry data. Thus, on each iteration of particle filter each particle is evolved according to the relative odometry measurements. Gaussian noise with zero mean is applied to the motion of particles in order to accommodate measurement errors. The vehicle state is described by the state vector $\mathbf{x}_t = (x_t, y_t, \theta_t)^T$, where x_t and y_t - vehicle 2D map-relative coordinates at the time t and θ_t is a corresponding heading angle. Given the difference between two consecutive odometry measurements:

$$\begin{cases} \Delta x_t = x_t^{odom} - x_{t-1}^{odom}, \\ \Delta y_t = y_t^{odom} - y_{t-1}^{odom}, \\ \Delta \theta_t = \theta_t^{odom} - \theta_{t-1}^{odom}, \end{cases} \quad (1)$$

the actuation applied to particles is expressed as:

$$\begin{cases} x_t^n = x_{t-1}^n + d_t^n \sin(\theta_{t-1}^n + \Delta\theta^n + \delta_t^n), \\ y_t^n = y_{t-1}^n + d_t^n \cos(\theta_{t-1}^n + \Delta\theta^n + \delta_t^n), \\ \theta_t^n = \theta_{t-1}^n + \Delta\theta^n + \delta_t^n, \end{cases} \quad (2)$$

where $d_t^n = \sqrt{(\Delta x_t)^2 + (\Delta y_t)^2} + \eta_t^n$, δ_t^n and η_t^n represent Gaussian white noise.

The set of sensors used in this work to support pose estimation includes wheel speed sensor and yaw-rate sensor. It is important to note that the presented set does not provide acceptable pose estimation on its own, and prone to significant drift over time. Due to this reason, no absolute pose estimation derived from odometry data is used for localization and only relative measurements $(\Delta x_t, \Delta y_t, \Delta \theta_t)$ are fed to the actuation model. The trajectory derived from odometry is shown in Fig. 2.

Measurement model

In order to compute the weight of each sample, the measurement model expressed as a likelihood function is applied as a next filtering step. The function indicates how well each particle corresponds to sensor

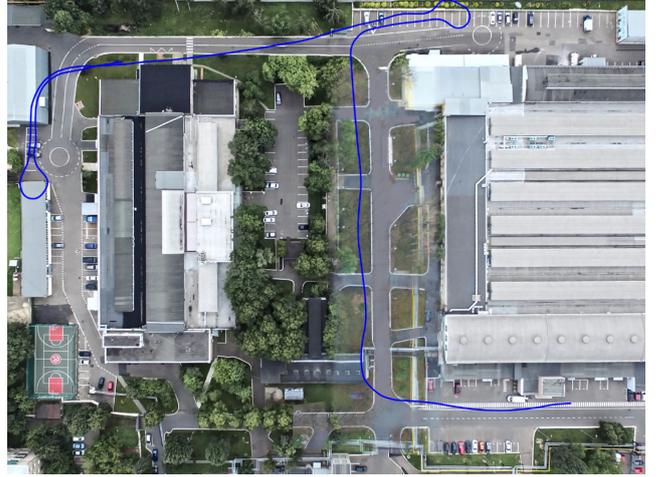


Figure 2. Vehicle trajectory estimated from odometry data.

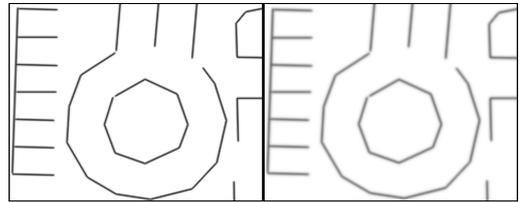


Figure 3. A fragment of the map image before (left) and after (right) Gaussian smoothing.

measurements obtained at the current moment of time. This section describes how the likelihood function is derived for each type of measurements used in this work.

The main data source for the proposed localization method is the lane marking detector, which works on images from a single monocular camera. Lane markings detected on each incoming image are approximated by polylines. Then each polyline is transformed from the reference frame associated with camera origin to the map reference frame. Particle weight based on lane markings may be expressed via log-likelihood function as:

$$w_{mt}^n = \sum_{i=1}^I \sum_{j=1}^J \ln f(x_{i,j}, y_{i,j} | \mathbf{x}_t^n), \quad (3)$$

where I - number of line segments detected on image, J - number of points corresponding to i th line segment and $f(x_{i,j}, y_{i,j} | \mathbf{x}_t^n)$ is a function indicating the probability of point $(x_{i,j}, y_{i,j})$ to be detected given the state \mathbf{x}_t^n .

As it was mentioned previously measurement error model associated with lane markings detection is embedded into the map on offline map preparation stage. We assume that the error follows a normal distribution with zero mean. In order to account for it, Gaussian smoothing is applied to the raster map image as shown in Fig. 3.

The resulting likelihood value is computed as the sum values of the map pixels matched to the detected set of polylines. We propose an efficient approximation of equation (3) presented below:

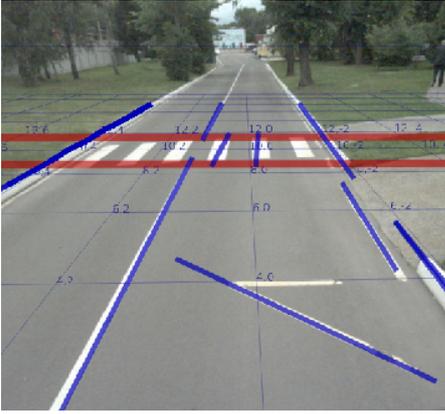


Figure 4. Example of lane markings (blue) and pedestrian crossing (red) recognition.

$$w_{m_t}^n = \sum_{i=1}^I \sum_{j=1}^J \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{I(x_{i,j}^n, y_{i,j}^n)}{\sigma}\right)^2}, \quad (4)$$

where $x_{i,j}^n, y_{i,j}^n$ are coordinates of j th point of i th detected line segment in the map reference frame, $I(x_{i,j}^n, y_{i,j}^n)$ is a function which accepts coordinates in a map reference frame and returns intensity value of the pixel corresponding to this coordinates on a digital map image, σ is a standard deviation of lane markings detection error.

In addition to lane markings, other types of features and prior information can be used to improve the localization precision. As an example, a detector of pedestrian crossings similar to one proposed in (Povolotskiy et al. 2017) is used in this work to reduce longitudinal error. It becomes possible due to the higher specificity of the detector in comparison to the more generic lane marking detector. Fig. 4 shows the example of pedestrian crossing as well as lane marking detection. As it was discussed previously, pedestrian crossings are represented by rectangles on the map. The likelihood function for detected crossings is represented by a shifted sigmoid function as indicated below:

$$w_{c_t}^n = \begin{cases} \sum_{i=1}^I \frac{1}{1+e^{(-C_n+S_i)}}, & \text{if crossing was detected,} \\ 1, & \text{otherwise.} \end{cases} \quad (5)$$

In equation (5) C_n is an intersection area of a crossing marked on the reference map with the rectangle approximating the detected crossing transformed to the map reference frame according to a position of n th particle; S_i corresponds to a bounding rectangle area of i th crosswalk on the map.

Another type of information used to increase accuracy and computational speed is the occupancy map. It constrains the area where particle presence is possible (utilizing so-called "vehicle on road" assumption) and therefore discards unlikely particles as early as possible and improves the posterior distribution accuracy. The likelihood function for occupancy map is expressed as:

$$w_{om_t}^n = \begin{cases} 1, & \text{if pose is not occupied,} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

The resulting measurement model combining likelihood functions (4), (5), and (6) is presented below:

$$p(z_t | x_t^n) \propto w_t^n = \frac{w_{m_t}^n w_{c_t}^n w_{om_t}^n}{\sum_{n=1}^N w_t^n} \quad (7)$$

EVALUATION AND DISCUSSION

The proposed method was experimentally evaluated on a real ground vehicle of a small bus kind.

The analysis and comparison of localization methods are complicated by the fact that there is no straightforward way to obtain the reference trajectory from the sensors. In the scope of this work, in order to generate reference trajectories, the described approach described below was utilized. During the test run of a localization algorithm, on each iteration of particle filter the posterior particle distribution is saved with the corresponding indices relating each particle to one from the prior distribution from which it was derived. Together these measurements form the data structure referred to within the scope of this work as a particle transition graph. To obtain the reference trajectory the graph is traversed backward in time and the reference pose is computed on each step by averaging poses of particles which are known to survive on the subsequent steps.

While many systems proposed in the literature had been tested in relatively simple scenarios, we have deliberately chosen a challenging test bed with two 180 degree turns on the roundabouts to evaluate the proposed localization technique. The comparison of trajectory recovered from online localization versus reference trajectory is presented in the Fig. 5 along with quantitative localization results presented in table 1. The number of particles used in this experiment was set to 1000, which can be easily computed on any modern low-end computer. The x and y coordinates of starting position are assumed to be known in all the presented experiments unless otherwise indicated. Meanwhile the initial heading direction is unknown and has a uniform distribution.

Table 1: Summary of localization precision in experimental run.

	x, m	y, m	θ , degrees	Euclidean distance, m
Max abs. error	1.310	1.207	0.246	1.346
Mean abs. error	0.159	0.127	0.029	0.235
Standard deviation	0.246	0.237	0.044	0.251

The obtained results demonstrate that the proposed method is able to perform precise localization in com-

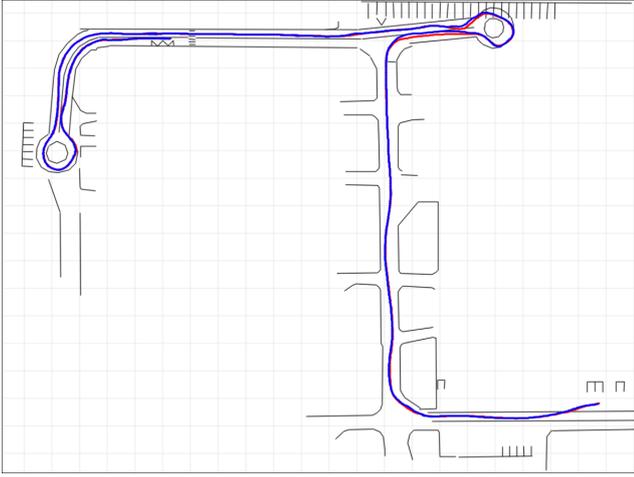


Figure 5. Reference trajectory (red) vs. trajectory recovered by particle filter algorithm (blue). Cell size is 10 m along both directions.

Table 2: Summary of localization precision in experimental run without lane marking detection.

	x, m	y, m	θ , degrees	Euclidean distance, m
Max abs. error	3.958	6.724	1.002	6.725
Mean abs. error	0.793	1.395	0.122	1.862
Standard deviation	1.107	1.826	0.221	1.159

plex real-world scenarios. Even though the maximum absolute error is quite high the average precision and standard deviation suggest that most of the time the localization precision stays within 0.5 m tolerance. It is important to note that most of the previous research is focused on regular cars where only limited use of lane markings can be made for localization purposes due to the relatively low camera position (and consequently narrow field of view), while for autonomous buses it is possible to install the camera at around 2 m height above the road surface and therefore to obtain a more informative picture of road marking.

In order to analyze the precision gain introduced by using lane feature-based approach trajectory obtained by using all the available data (i.e. odometry, occupancy map, crosswalk detection) except lane markings was compared with the reference one. The experimental results are presented in Fig. 6 and in table 2.

The results show that the use of lane markings yields about fivefold increase in localization precision.

It is worth mentioning that the developed approach is capable of solving the global localization problem, i.e. localization under global uncertainty. Fig. 7 shows the convergence of particles distribution over time in the scenario when all starting positions on the map have equal probability. The number of particles used in this

experiment was increased to 30000 in order to populate the whole map uniformly.

CONCLUSION

A vision-based localization method for autonomous ground vehicles was proposed in this work. It is based on a combination of visual detectors (i.e. lane marking detector and pedestrian crossing detector) and odometry data. Particle filter technique was applied in order to fuse heterogeneous data sources. The proposed method has a straightforward yet computationally efficient implementation which allows it to be used even in applications with limited computational resources. Another advantage of this method is an ease of map preparation and use. The required vector map contains lane markings and pedestrian crossings marked as simple geometric primitives. Once transformed into a raster image the map can also embed extra pre-computed properties as it was demonstrated with lane detection error represented by Gaussian blur. Such map representation allows to reduce the number of operations in the online localization algorithm and guarantees constant data access time.

As it was demonstrated, another possible application of the proposed method is to be used as a preliminary reference trajectory generation approach for further map refinement and enrichment with other data, especially in cases where precise external localization is impossible, and implementing a full-featured simultaneous localization and mapping (SLAM) approach is impractical due to the algorithm complexity.

Tests on a real vehicle have demonstrated the applicability of the proposed method to the real-world scenarios. In considering the applicability of this method one should take into account that the tests were conducted on a bus-like platform which has a camera system placed at a considerable height, which may be unachievable for regular cars. Therefore a degradation of localization precision may be expected when used on conventional platforms.

Further improvement may include the use of additional lane marking features in the likelihood estimation such as line orientation as well as an extension of the set of visual detectors, e.g. stable local image features or stereo vision-based detection of buildings and curbs.

ACKNOWLEDGMENT

This work was supported by the Russian Science Foundation, project no 14-50-00150.

REFERENCES

- Chausse, Frederic; Jean Laneurit; and Roland Chapuis. 2005. "Vehicle localization on a digital map using particles filtering." In *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, 243–248.
- Du, Xinxin and Kok Kiong Tan. 2016. "Vision-based approach towards lane line detection and vehicle localization." *Machine Vision and Applications*, 27(2):175–191.
- Hata, Alberto Y; Fernando S Osorio; and Denis F Wolf. 2014. "Robust curb detection and vehicle localization in

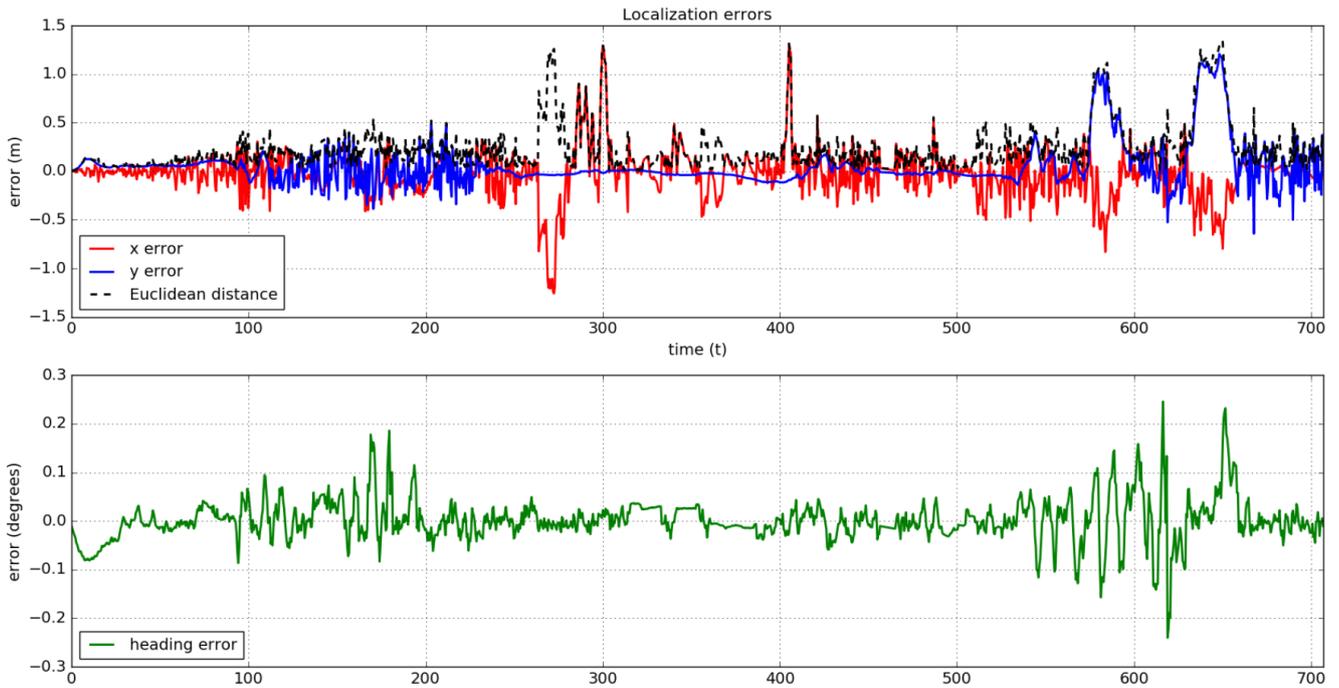


Figure 6. Localization errors in a testing run.

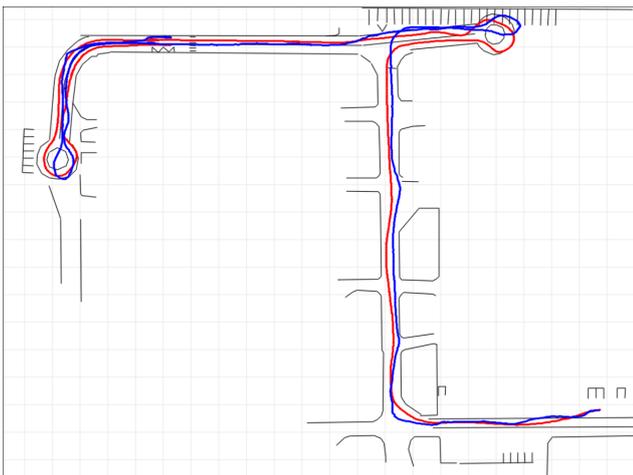


Figure 7. Reference trajectory (red) vs. trajectory recovered by particle filter algorithm in the absence of lane markings detector (blue). Cell size is 10 m along both directions.

urban environments.” In *Intelligent vehicles symposium proceedings, 2014 IEEE*, 1257–1262.

- Hernández, Jorge and Beatriz Marcotegui. 2009. “Filtering of artifacts and pavement segmentation from mobile lidar data.” In *ISPRS Workshop Laserscanning 2009*.
- Hillel, Aharon Bar; Ronen Lerner; Dan Levi; and Guy Raz. 2014. “Recent progress in road and lane detection: a survey.” *Machine vision and applications*, 25(3):727–745.
- Huang, Albert S; David Moore; Matthew Antone; Edwin Olson; and Seth Teller. 2009. “Finding multiple lanes in urban road networks with vision and lidar.” *Autonomous Robots*, 26(2-3):103–122.
- Jo, Kichun; Yongwoo Jo; Jae Kyu Suhr; Ho Gi Jung; and Myoungcho Sunwoo. 2015. “Precise localization of an autonomous car based on probabilistic noise models of

road surface marker features using multiple cameras.” *IEEE Transactions on Intelligent Transportation Systems*, 16(6):3377–3392.

- Kalman, Rudolph Emil. 1960. “A new approach to linear filtering and prediction problems.” *Journal of basic Engineering*, 82(1):35–45.
- Krokhina, D.; V. Blinov; S. Gladilin; I. Tarxanov; and V. Postnikov. 2015. “Fast roadway detection using car cabin video camera.” In *ICMV 2015*, volume 9875, 98751F–1–98751F–5 (SPIE, 2015).
- Li, Liang; Ming Yang; Bing Wang; and Chunxiang Wang. 2017. “An overview on sensor map based localization for automated driving.” In *Urban Remote Sensing Event (JURSE), 2017 Joint*, 1–4.
- Lu, Wenjie; Emmanuel Seignez; F Sergio A Rodriguez; and Roger Reynaud. 2014. “Lane marking based vehicle localization using particle filter and multi-kernel estimation.” In *Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on*, 601–606.
- Povolotskiy, Mikhail A; Elena G Kuznetsova; and Timur Khanipov. 2017. “Russian license plate segmentation based on dynamic time warping.”
- Sheehan, Mark; Alastair Harrison; and Paul Newman. 2013. “Continuous vehicle localisation using sparse 3d sensing, kernelised rényi distance and fast gauss transforms.” In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, 398–405.
- Tao, Zui; Ph Bonifait; Vincent Fremont; and Javier Ibanez-Guzman. 2013. “Lane marking aided vehicle localization.” In *Intelligent Transportation Systems (ITSC), 2013 16th International IEEE Conference on*, 1509–1515.
- Thrun, Sebastian. 2002. “Particle filters in robotics.” In *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence*, 511–518.
- Volkov, Alexey; Egor Ershov; Sergey Gladilin; and Dmitry Nikolaev. 2017. “Stereo-based visual localization without triangulation for unmanned robotics platform.” In *2016 International Conference on Robotics and Machine Vision*, volume 10253, page 102530D.
- Ziegler, Julius; Henning Lategahn; Markus Schreiber;

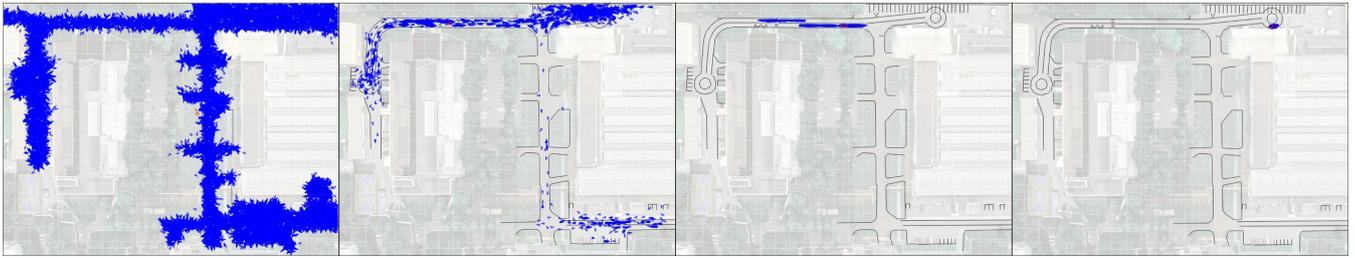


Figure 8. Particles evolution under global uncertainty. Blue markers denote particles positions. Red marker shows the estimated vehicle pose.

Christoph G Keller; Carsten Knoppel; Jochen Hipp; Martin Haueis; and Christoph Stiller. 2014. “Video based localization for berthä.” In *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, 1231–1238.

AUTHOR BIOGRAPHIES

OLEG SHIPITKO



was born in Yeisk, Russia. He obtained his bachelor degree in Electrical Engineering from Bauman Moscow State Technical University in 2015 and masters degree in Computer Science from Skolkovo Institute of Science and Technology in 2017. Since then he has works in the Laboratory of Vision Systems at the Institute for Information Transmission Problems. His research is focused on vision-based localization algorithms for robotics. Send mail to oleg.shipitko@iitp.ru.

ANTON GRIGORYEV



was born in Petropavlovsk-Kamchatskiy, Russia. Having graduated from Moscow Institute of Physics and Technology, he has been developing industrial computer vision systems with the Laboratory of Vision Systems at the Institute for Information Transmission Problems since 2010. Now he is heading the Autonomous Machine Vision Systems Research Group developing visual navigation solutions for robotic vehicles. Send mail to me@ansgri.com.