2018 IEEE Intelligent Vehicles Symposium (IV)
Changshu, Suzhou, China, June 26-30, 2018

# Bayesian Framework for Vehicle Localization Using Crowdsourced Data*

Sergey Verentsov[1], Emil Magerramov[1], Vlad Vinogradov[1], Ramil Gizatullin[1], Andrey Alekseenko[2], Yaroslav Kholodov[1] and Evgeniy Nikolskiy[3]

*Abstract*— **Recently, the vehicle localization has become a major problem in the field of intelligent transportation systems (ITS). Modern vehicles support multiple systems for localization, such as Global Positioning System (GPS), Inertial Measurement Unit (IMU), lidar, and video feed. However, the problem of merging data from different sources for reducing errors remains challenging. This problem becomes increasingly hard if we use data with high uncertainty — for example, crowdsourced environmental features. Such data can provide invaluable information, especially when other data sources, such as GPS, are not available. Unfortunately, using it for vehicle localization requires great care, as crowdsourced data can be very noisy and imprecise. In this paper we present a Bayesian approach for sensor fusion, and use it to improve vehicle localization using crowdsourced data of traffic sign positions. We show that this method offers noticeable improvements (error reduction by 8.2%) compared to using GNSS and IMU only, and is comparable to the use of precise traffic signs positions — which are superior to crowdsourced positions, but very hard to obtain at scale.**

## I. INTRODUCTION

Precise vehicle localization has become a subject of great interest in recent years in the field of Intelligent Transportation Systems. Vehicle localization is needed for both autonomous and non-autonomous vehicles. One of the areas where precise localization is essential are Advanced Driver-Assistance Systems (ADAS), which contain components like lane detector and road rules checker. Another application is the problem with dangerous cargo delivery, where it is crucial to know exactly where the car is at a given point of time and to check whether the driving behavior is safe.

One way to localize a vehicle is to use Global Navigation Satellite Systems (GNSS), such as Global Positioning System (GPS) [1], GLONASS, and BeiDou Navigation Satellite System. However, most consumer devices provide low accuracy: in dynamic outdoor tests, GNSS positioning error could rise up to 20 meters with an average of about 3 meters [2]. In urban areas, high buildings and narrow streets further aggravate GNSS localization errors. For GNSS the measurement errors are independent (no time-correlation), and the error does not accumulate over time. The downside is that GNSS measurements occur with a frequency of approximately 1 Hz, which leads to a lag in acceleration measurements and reduces the efficiency of many denoising techniques that rely on averaging consequent measurements.

The relative position of the vehicle can be measured using Inertial Measurement Unit (IMU) [3], present inside most modern cell phones. IMU, which typically consists of an accelerometer and a gyroscope, has the update frequency around 400 Hz. On the downside, it can only measure position relative to the previous point in time and, consequently, velocity and position errors accumulate over time [4]. This makes IMU not suitable as the only method for position estimation on longer timescales.

The advantages of both methods can be leveraged using sensor fusion approach: by combining the measurements from GNSS and IMU, both systems will compensate each other's errors, resulting in more precise measurements [5].

Another approach for vehicle localization is Simultaneous Localization and Mapping (SLAM) [6], which is a popular method to obtain the information about the surrounding environment using a video of the front view of a vehicle. One more approach uses computer vision techniques to detect traffic lights and achieve lane-level navigation [7]. However, traffic lights are not as numerous as other visual landmarks.

The work of Qu *et al.* [8] incorporates the data about geo-referenced traffic signs as Ground Control Points into the Local Bound Adjustment method to correct the accumulated error in the vehicle position measurements. The downside of this approach is that it requires a database of very precise traffic signs positions. The authors address this issue by processing a video of the area where the algorithm was tested and extracting the signs' positions beforehand. Only then, the localization algorithm could be applied.

We propose an algorithm that uses a crowdsourced database to correct vehicle's position obtained from GNSS and IMU. It applies the probabilistic Bayesian framework to estimation of the current vehicle's coordinates using prior information about the positions where a traffic sign has been previously spotted. The new estimated position is then added to the set of the prior observations. This way, a crowdsourcing is used to continuously improves localization accuracy.

In the proposed algorithm we only use data from GPS, IMU, and the camera feed to detect the traffic signs (using custom algorithm). We do not use the SLAM approach to recreating the trajectory of a vehicle since it is much more

*This work was supported by the Russian Science Foundation (grant ID 14-11-00877)

[1]Sergey Verentsov, Emil Magerramov, Vlad Vinogradov, Ramil Gizatullin, and Yaroslav Kholodov are with the Intelligent Transportation Systems Lab, Innopolis University, Innopolis, Russia {s.verentsov, e.magerramov, v.vinogradov, r.gizatullin, ya.kholodov}@innopolis.ru

[2]Andrey Alekseenko is with the Institute of Computer Aided Design, Moscow, Russia aland@phystech.edu

[3]Evgeniy Nikolskiy is with the RoadAR, Inc., Innopolis, Russia eugeniy@gmail.com

978-1-5386-4451-5/18/$31.00 ©2018 IEEE

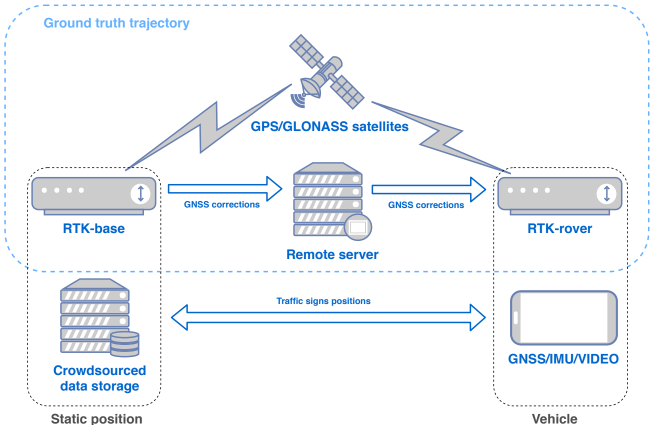


Fig. 1. Overall system setup. The upper part show the setup of Real Time Kinematic GNSS system, used for Ground truth trajectory measurement. The lower part show the proposed crowdsourcing localization system, which uses smartphone installed in the vehicle to obtain GNSS, IMU, and traffic signs sightings data, and exchanges this data with crowdsourced data storage server to improve localization.


Fig. 2. Detected traffic signs in the area of test run.

computationally expensive than the traffic sign detection, and thus might be unsuitable for use in embedded or mobile devices.

We show that probabilistic trajectory error correction based on the traffic sign detection provides better precision than using only GNSS and IMU, and is a viable tool for lane-level vehicle positioning.

## II. METHODOLOGY

We compared the following approaches built trajectories using different approaches: 1) GNSS — using raw data from smartphone GNSS system; 2) GNSS+IMU (KF) — combining GNSS and IMU data from smartphone using Kalman filter; 3) GNSS+IMU (LTO) — combining GNSS and IMU data from smartphone using Local Transformation Optimization; 4) GNSS+IMU+PSP — refining trajectory from GNSS+IMU (LTO) method using proposed Bayesian approach with precisely measured traffic signs positions; 5) GNSS+IMU+CSP — refining trajectory from GNSS+IMU (LTO) method using proposed Bayesian approach with crowdsourced traffic signs positions. We use high-precision trajectory measured with Real Time Kinematic Global Navigation Satellite System (RTK GNSS) system as a ground truth.

### *A. Data Collection Setup*

The setup for the data collection was the following: a car with a smartphone installed in the vehicles interior above the windshield and an RTK-rover GNSS receiver for precise GNSS measurements. Fig. 1 shows the system setup for data collection.

The OnePlus 5T smartphone was used to collect data from the built-in GNSS and IMU sensors and to record the video feed from the built-in monocular camera.

Real Time Kinematic Global Navigation Satellite System (RTK GNSS) [9] was used to collect ground truth GPS/GLONASS coordinates during the ride. It consists of two modules: an RTK-rover receiver, which is installed in the vehicle, and an RTK-base station, which is a stationary device. The base station transmits GNSS corrections to the RTK-rover, which allows precise localization with precision up to 2 centimeters. We used NV08C-RTK[1] devices from NVS Technologies.

We also used this device to collect precise traffic signs coordinates for GNSS+IMU+PSP method.

### *B. Experimental Setup*

We recorded a 28-minute long trajectory in an urban environment in Innopolis, Republic of Tatarstan, Russia.

The video has 1920x1080 resolution with 30 frames per second. The route includes low- and middle-height buildings, traffic signs and irregular road crossing pedestrians.

The imprecise GNSS coordinates of the vehicle with a sampling rate of 1 Hz and IMU velocities with a sampling rate of 400 Hz were recorded from the built-in OnePlus 5T smartphone sensors.

The high-precision GNSS coordinates from the RTK GNSS system with a sampling rate of 5 Hz were recorded to be used as ground truth.

Traffic signs sightings (shown in Fig. 2) were crowdsourced using the recorded video and RoadAR[2] dash camera mobile application. The data consists of GNSS coordinates of a car in the time point of the last video frame when a concrete traffic sign was visible. We collected the data for the experiment during a ride in Innopolis city. However, the proposed system can work in any other environments provided proper crowdsourced data collecting tool.

To test the Bayesian coordinate approximation, we generated a realistic test set of previous observations for the traffic signs. Fig. 2 shows traffic signs once detected on a camera video stream during an extra ride. For each observed sign, there are 100 simulated traffic signs which are sampled

[1] http://nvs-gnss.com/products/receivers/item/39-nv08c-rtk.html
[2] http://www.roadar.ru/

from a Gaussian distribution with the mean of the sign's GPS coordinate and the covariance matrix built according to the car direction at the detection moment.

### *C. GNSS+IMU (KF) — Fusing GNSS and IMU Data with Kalman Filter*

One of the classical approaches to trajectory correction is Kalman Filter [10]. It is often used for vehicle localization in two dimensions [11]. Due to the nature of the algorithm, it is most useful when applied to systems with linear dynamics. However, most systems are non-linear and, in these cases, the Extended Kalman Filter (eKF) is used [12].

The algorithm has two steps: prediction and update steps. In the prediction step, the algorithm predicts the next state of the variables which has to be estimated. Then the outcome of the next measurement becomes known and the estimates are updated based on the new data. This allows us to use Kalman filter as both smoothing and sensor fusion technique to obtain the combined GNSS + IMU trajectory.

### *D. GNSS+IMU (LTO) — Fusing GNSS and IMU Data with Local Transformation Matrix Optimization*

Another approach for combining global (GNSS) and local (IMU) data is through use of local transformation matrices.

We obtain accelerometer and gyroscope data in local coordinate space, which needs to be converted to global space. We use the following linear transformation:

$$a_{\text{global}}(x_i) = R(x_i) \cdot a_{\text{local}}(x_i),$$

where $a_{\text{global}}(x_i)$ is the acceleration vector at point $x_i$ in the global space, $R(x_i)$ is the rotation matrix at point $x_i$, $a_{\text{local}}(x_i)$ is the acceleration vector in the point $x_i$ in the local space (raw accelerometer data). We model error in the accelerometer measurements as two constant biases — for local ($b_1$) and global ($b_2$) accelerations.

$$\hat{a}_{\text{global}}(x_i) = R(x_i)(a_{\text{local}}(x_i) + b_1) + b_2.$$

$R(x_i)$ can be calculated as follows:

$$R(x_i) = R(x_{i-1})(I + W_i \Delta t_i) = \prod j = 1^i \left(I + W_j \Delta t_j\right),$$

where $\Delta t_i$ is the time span between the $i$-th and the $(i+1)$-st measurements, and $W_i$ is the derivative of $R(x_i)$ obtained directly from the gyroscope sensor data. $I$ is the $3\times3$ identity matrix.

For a known $\hat{a}_{\text{global}}(x_i)$ we can find the speed at point $x_i$ given the initial speed $v_0$ as

$$v_i^{\text{IMU}} = v_0 + \sum_{k=1}^{i} \hat{a}_{\text{global}}(x_k) \Delta t_k. \tag{1}$$

The values of $b_1$, $b_2$, and $v_0$ should be tuned to fit the data obtained from the GPS. In other words, the position estimated using IMU should be close to the position estimated using GPS. We achieve this by solving an optimization problem:

$$\min_{v_0, b_0, b_1} \sum_i \left( \left\| \sum_{j \in I_i} v_j^{\text{IMU}} \Delta t_j \right\| - v_i^{\text{GPS}} \Delta T_i \right)^2,$$

where $I_i$ is a set of indices of IMU measurements between $i$-th and $(i+1)$-st GPS measurements, $\Delta T_i$ is the time between $i$-th and $(i+1)$-st GPS measurement, $v_j^{\text{IMU}}$ is the speed approximation made using IMU for $j$-th IMU measurement as defined in (1), $v_i^{\text{GPS}}$ is the speed approximation made directly from GPS measurements.

The problem above is a quadratic optimization problem and can be efficiently solved. We use L-BFGS algorithm [13] due to its low computational cost.

However, constant bias model shows poor results for long rides. Therefore, we perform parameter fitting regularly using 40 seconds long sliding windows.

### *E. GNSS+IMU+PSP — Using Precisely Measured Traffic Signs Positions*

We also measured precise traffic signs positions along the testing trajectory. We build the vehicles trajectory by correcting the vehicles position towards the closest traffic sign location.

In order to shift a trajectory towards the known positions of the traffic signs the first derivative over the trajectory is calculated. Then we calculate the correction value for both axes - the cumulative sum of the obtained differences between the positions of the nearest traffic signs.

At each timestep the correction is added to the current vehicle location until the new traffic sign is detected. At this point the correction is set to 0 and the process is repeated.

### *F. GNSS+IMU+CSP — Bayesian Approach with Crowd-sourced Traffic Signs Positions*

One of the ways to improve the estimation of the vehicle position based on arbitrary prior information (especially if the data has varying uncertainty) is to use Bayesian Statistics, specifically the maximum *a posteriori* probability (MAP) estimate. It is an estimate of an unknown quantity that is equal to the mode of the posterior distribution. Assuming prior distribution over $g(\theta)$ is known, we can use Bayes' Theorem to calculate $f(\theta|x)$ — the posterior distribution of given observations.

We can calculate an estimate $\hat{\theta}_{\text{MAP}}(x)$ of the unknown parameter as a mode of the posterior distribution. The MAP estimate coincides with the Maximum Likelihood Estimation method when the prior $g(\theta)$ is uniform.

In this paper we estimate the corrected position of the vehicle $x_{\text{corrected}}$ given $n$ previous observations $\vec{x}_{\text{old}} = [x_1, \ldots, x_n]$ of the traffic sign $s$ and the current vehicle position $x_{\text{current}}$ obtained from the GNSS+IMU (LTO) sensor fusion. To do this we calculate the mode of the posterior distribution in an $\varepsilon$-neighborhood of the $x_{\text{current}}$ and set the $x_{\text{corrected}}$ to the coordinates of this mode. Then the optimization problem is:

$$x_{\text{corrected}} = \underset{x_i \in B_{x_{\text{current}}}[\varepsilon]}{\arg\max} \; p(s|x_i; \vec{x}_{\text{old}}) \cdot p(x_i|x_{\text{current}}), \tag{2}$$

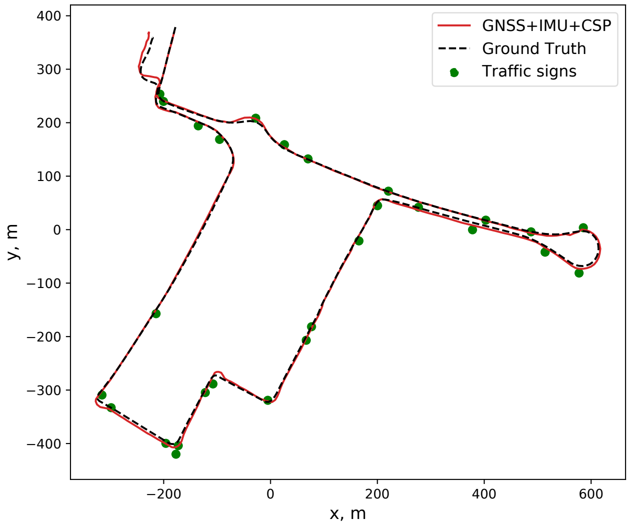


Fig. 3. Trajectory obtained by the proposed GNSS+IMU+CSP method compared to ground truth. Static positions of traffic signs depicted by green dots.

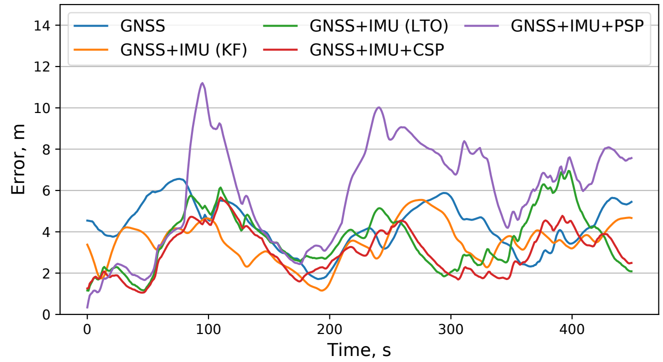


Fig. 4. Comparison of localization methods. The translation errors for all methods are shown for the first 450 seconds of trajectory.

where $B_{x_{\text{current}}}[\varepsilon] = x \in \mathbb{R}^2 | \|x - x_{\text{current}}\|^2 \leq \varepsilon$ with $\varepsilon$ tuned empirically.

We model the probability to observe the sign $s$ given current coordinates and previous observations $\vec{x}_{\text{old}}$ as a bivariate Gaussian distribution with the covariance matrix $\Sigma$ calculated from the set of the previous observations. Also, we assume our prior distribution $p(x_i; x_{\text{current}})$ to be Gaussian with $\mu = x_{\text{current}}$, and $\sigma$ tuned empirically during the experiment phase.

The $\varepsilon$-neighborhood around the current vehicle's position is quantized and constrained to a small radius around the vehicle. We calculate these points the probabilities of the posterior distribution are calculated so that the $\arg\max$ in (2) is taken only for these points.

Furthermore, in the proposed framework the Bayesian Updating method can improve the estimates of the vehicle position given new data. The calculated posterior position $x_{\text{corrected}}$ where the sign has been observed is added to the set of observations $\vec{x}_{\text{old}}$ for a specific sign. These data can later be used to recalculate the distribution of the probability to observe this sign given some other uncorrected position. The new posterior probability distribution can be computed to get a better estimate in the future. This way the observations needed for the algorithm to make the corrections are crowd-sourced instead of being collected beforehand and require only the approximate locations of the previous sightings of for each traffic sign to model the distribution.

After the corrected vehicle locations have been obtained, the final trajectory is calculated using the same method as for GNSS+IMU+PSP (see subsection II-E).

## III. RESULTS

We used a 28-minutes car ride in an urban environment to collect the following data: imprecise GNSS and IMU (accelerometer, gyroscope) measurements from a smartphone; a video from a smartphone on a windshield to detect traffic signs; precise GNSS measurements from RTK GNSS as the ground truth; precise GNSS coordinates of the detected traffic signs.

We compare several sensor fusion approaches to vehicle localization: 1) GNSS — using raw data from smartphone GNSS system; 2) GNSS+IMU (KF) — combining GNSS and IMU data from smartphone using Kalman filter; 3) GNSS+IMU (LTO) — combining GNSS and IMU data from smartphone using local transformation matrix optimization; 4) GNSS+IMU+PSP — refining trajectory from GNSS+IMU (LTO) method using proposed Bayesian approach with precisely measured traffic signs positions; 5) GNSS+IMU+CSP — refining trajectory from GNSS+IMU (LTO) method using proposed Bayesian approach with crowdsourced traffic signs positions.

The ground truth and one of the fused trajectories are shown in Fig. 3. We see that the biggest discrepancies (for all methods — data not shown) are at the turns.

We use mean translation error as an evaluation metric, which is calculated as follows:

$$\mathrm{Err}(\hat{X}, X) = \frac{1}{n} \sum_{i=1}^{n} \|\hat{X}_i - X_i\|,$$

where $n$ is number of measurements, $X_i$ and $\hat{X}_i$ are $i$-th points of compared paths. The translation error as a function of time is shown in Fig. 4, and mean translation errors are summarized in Table I — all errors are calculated in comparison with RTK-GNSS rover trajectories.

TABLE I
COMPARISON OF THE ERRORS FOR DIFFERENT LOCALIZATION METHODOLOGIES.

| Method | Error, m |
|---|---|
| GNSS | 4.40 |
| GNSS+IMU (KF) | 3.69 |
| GNSS+IMU (LTO) | 3.55 |
| GNSS+IMU+PSP | 4.86 |
| GNSS+IMU+CSP | **3.27** |

The first approach (GNSS) uses only data from smartphone's built-in GPS/GLONASS, and has the error a little over 4 meters, consistent with previous research [2]. The GNSS+IMU (KF) approach uses Kalman filter to combine measurements from smartphone's built-it GNSS and IMU sensors. The error per measurement is 3.69 meters. The GNSS+IMU (LTO) uses optimization of the local transformation matrix to fix accumulated IMU error using GNSS data. The mean error per measurement for this trajectory is 3.55 meters. The GNSS+IMU+PSP and GNSS+IMU+CSP approaches use GNSS+IMU (LTO) trajectory as a baseline and correct it using precise traffic signs positions and Bayesian approach with crowdsourced traffic signs positions, respectively. The mean error per measurement for GNSS+IMU+PSP trajectory equals 4.86 meters, and 3.27 meters for GNSS+IMU+CSP.

Therefore, adding information about traffic signs data results in the 8.2% mean error reduction compared to GNSS+IMU (LTO), even if the prior data of traffic signs positions are not very accurate.

## IV. CONCLUSIONS

We propose a Bayesian accumulated error correction algorithm that uses crowdsourced traffic signs sightings to improve vehicle localization accuracy. The proposed framework corrects the input trajectory based on the vehicles coordinates where the traffic signs have been detected in the past, and these data can be collected to improve localization in the future.

The input trajectory is obtained using a traditional sensor fusion frameworks which combines the trajectories obtained from the GNSS and IMU sensors. Examples of such algorithms are Kalman filter and GNSS/IMU local transformation matrix solving optimization, both of which were tested in this work.

We tested the algorithms on the data collected from the ride in an urban environment, and compared the calculated trajectories to the ground truth data from the RTK GNSS. The addition of the traffic signs data resulted in the mean error per measurement reduction by 8.2%.

Besides improvement in localization accuracy, the algorithm can also be used to correct the accumulated IMU error in situations when GNSS is not available. It can be further extended to incorporate other visual landmarks such as traffic lights, road markings and buildings, or use the trajectories generated with SLAM.

## ACKNOWLEDGMENT

The authors thank RoadAR Inc. for providing the traffic signs sightings data and the software library for road sign recognition.

## REFERENCES

[1] W. Rahiman and Z. Zainal, "An overview of development GPS navigation for autonomous car," in *2013 IEEE 8th Conf. Ind. Electron. Appl.* IEEE, jun 2013, pp. 1112–1118.

[2] P. A. Zandbergen and S. J. Barbeau, "Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones," *J. Navig.*, vol. 64, no. 03, pp. 381–399, jul 2011.

[3] M. M. Morrison, "Inertial measurement unit," Granted Patent US 4 711 125 A, 1987.

[4] J. Knuth and P. Barooah, "Error growth in position estimation from noisy relative pose measurements," *Rob. Auton. Syst.*, vol. 61, no. 3, pp. 229–244, mar 2013.

[5] D. A. Tazartes and J. G. Mark, "Integration of GPS Receivers into Existing Inertial Navigation Systems," *Navigation*, vol. 35, no. 1, pp. 105–119, mar 1988.

[6] M. Dissanayake, P. Newman, S. Clark, H. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Trans. Robot. Autom.*, vol. 17, no. 3, pp. 229–241, jun 2001.

[7] A. Vu, A. Ramanandan, A. Chen, J. A. Farrell, and M. Barth, "Real-Time Computer Vision/DGPS-Aided Inertial Navigation System for Lane-Level Vehicle Navigation," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 899–913, jun 2012.

[8] X. Qu, B. Soheilian, and N. Paparoditis, "Vehicle localization using mono-camera and geo-referenced traffic signs," in *2015 IEEE Intell. Veh. Symp.* IEEE, jun 2015, pp. 605–610.

[9] L. Wanninger, "Introduction to Network RTK," 2008. [Online]. Available: http://www.wasoft.de/e/iagwg451/intro/introduction.html

[10] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *J. Basic Eng.*, vol. 82, no. 1, p. 35, 1960.

[11] J. Woods and C. Radewan, "Kalman filtering in two dimensions," *IEEE Trans. Inf. Theory*, vol. 23, no. 4, pp. 473–482, jul 1977.

[12] S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proc. 1995 Am. Control Conf.*, vol. 3. American Autom Control Council, 1995, pp. 1628–1632.

[13] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Math. Program.*, vol. 45, no. 1-3, pp. 503–528, aug 1989.