IoT-based Vehicle Localization in GNSS-denied Environments

Guoqiang Mao[†], *Fellow*, *IEEE*, Keyin Wang[†*], Zhaozhong Zhang[†], Tianxuan Fu[†], and Xiaozhi Qu[‡]
[†]School of Telecommunications Engineering, Xidian University, Xi'an 710071, China
[‡]Beijing Voyager Technology Co., Ltd, Beijing 100193, China
[*]E-mail: keyinwang@stu.xidian.edu.cn

***Abstract*—Integration of global navigation satellite system (GNSS) and inertial navigation system (INS) presents significant potential for high-precision vehicle localization. However, this approach suffers from cumulative INS errors in GNSS-denied environments. To address this issue, this paper proposes a method to correct the navigation errors due to INS by using Internet of Things (IoT)-based vision positioning. More specifically, this method employs a binocular camera to assist in obtaining the vision positioning of the vehicle through recognition of LED lights installed in a type of ubiquitously deployed IoT devices that are increasingly used in smart transportation systems. The final navigation outcomes are attained by fusing the vision positioning results with INS information using an unscented Kalman filter (UKF). Real-world experiment results validate the effectiveness of the proposed vehicle localization method.**

***Index Terms*—Vehicle localization, Internet of Things (IoT), vision positioning, inertial navigation system (INS), unscented Kalman filter (UKF).**

I. Introduction

An accurate and robust vehicle localization system is important for intelligent driving and intelligent transportation system applications [1], [2]. There are a lot of vehicle localization techniques using diverse sensors, including cameras, LiDAR, RADAR, ultrasonic sensors, and inertial measure units (IMUs) [3], [4]. The most widely used localization method is the global navigation satellite system (GNSS), which can provide a low-cost and readily accessible solution for vehicle localization [5]. However, due to the susceptibility of the GNSS to interference, the navigation frameworks relying on GNSS become unreliable in many scenarios, such as urban canyons, tunnels, and underground parking lots [6]. Hence, it is imperiative to find suitable vehicle localization schemes to achieve reliable localization in GNSS-denied environments.

The odometer-based navigation methods that do not rely on GNSS information are extensively studied [7], [8], [9]. The odometer data can effectively constrain the divergence of the INS navigation errors. However, the position errors of the odometer will still gradually accumulate with time, which is unacceptable for long-term navigation tasks.

The simultaneous localization and mapping (SLAM) based vehicle navigation methods have increasingly become research hotspots, which can improve vehicle localization performance during GNSS outages [10], [11]. SLAM can eliminate the accumulative errors in odometry through closed-loop detection. However, implementing closed-loop detection in open environments is extremely challenging [6].

The localization methods based on high-definition (HD) map can provide high-accuracy vehicle positioning without relying on GNSS [12], [13]. However, the positioning accuracy may diminish due to a lack of timely updates in HD map in dynamic environments. Furthermore, the extensive storage requirements of HD map can impede real-time positioning capabilities [14].

Most recently, some researchers proposed integrated localization methods aided by vision positioning. Qu *et al*. [15] proposed a geo-referenced traffic sign based localization approach. In this method, ground control points (GCPs) are obtained by identifying traffic signs and matching them with a pre-made dataset. The GCPs are integrated in a bundle adjustment process to reduce the drifts during vehicle localization. Wang *et al*. [5] proposed a vehicle localization method to improve the vehicle's positioning accuracy at traffic intersections by fusing the position information of traffic lights and the navigation information of inertial navigation system (INS). In this method, traffic lights are identified by an onboard camera, and their position information is provided by HD map. Hu *et al*. [16] proposed an integrated localization method, which utilizes the vision positioning deduced from kilometer signs to correct the navigation errors of the INS/odometer integration system. The aforementioned localization methods have a notable limitation that they can only correct the vehicle's position in certain scenarios, e.g., traffic intersections, near the kilometer signs, thus failing to meet the demand for vehicle localization throughout the entire area.

In this paper, a method to improve the vehicle localization accuracy in GNSS-denied environments is proposed, in which an unscented Kalman filter (UKF) is used to fuse vision positioning results using a type of ubiquitously deployed Internet of Things (IoT) devices with a prior known position termed smart studs, which are increasingly used in intelligent transportation systems [17], [18], and INS navigation information. The vision positioning results is attained by identifying the LED lights installed inside the smart studs using an onboard binocular camera. Unlike previous studies that utilized an extended Kalman filter (EKF) to fuse vision positioning and INS, a UKF is adopted because its unscented transform (UT) is expected to reduce linearization errors in EKF. Specifically, our main contributions are threefold:

1) An IoT-based vehicle localization method is proposed. In this method, the accumulated positioning errors of INS are

corrected by the vision positioning obtained using IoT devices as landmarks to achieve continuous and reliable vehicle localization.

2) A UKF is used to integrate the vision positioning and the navigation information from INS. Compared to EKF, the UKF achieves higher localization accuracy.

3) Real-world vehicle localization experiments are conducted. The experimental outcomes validate the effectiveness and accuracy of the proposed method.

The paper is organized as follows. Section II gives an overview of the proposed vehicle localization method. Section III introduces the vehicle position fusion based on UKF. The experimental results are given in Section IV. Finally, the conclusion and future work are drawn in Section V.

II. Overview of the Proposed Methodology

A. Framework of the Proposed Vehicle Localization System

As an essential component of smart roads, smart studs are ubiquitously deployed along lane-division lines, road-boundary lines, or near road surfaces. Once deployed, they possess long-term stability, which helps avoid frequent updates to map information. The smart studs can actively emit red/white/yellow light with certain brightness, which can be detected by vehicle onboard cameras. With these prominent features, smart studs can serve as excellent auxiliary positioning devices.

The overall framework of the vehicle localization method is shown in Fig. 1. The global positions of the smart studs are provided by real-time kinematic (RTK) devices, forming the smart stud position database. A vehicle onboard binocular camera is applied to detect smart studs using the hue, saturation, value (HSV) color model [19]. The measured positions of the vehicle can be inferred using smart studs as reference and camera as the sensor through binocular vision and coordinate transformations. These positions of the vehicle are used as observations for UKF, with the INS updates serving as predictions for the filtering algorithm. The state corrections estimated by the UKF are fed back to update the vehicle positions and compensate for the IMU errors.

The following method is adopted to determine the location of the light-emitting smart stud recognized by the binocular camera. An estimated position of the smart stud can be obtained by binocular vision and coordinate transformations. The estimated position of the smart stud is compared to the positions of smart studs in the smart stud position database and the closest match in the database to the estimated value identifies the recognized smart stud and the accurate position of the smart stud is obtained. If there are multiple smart studs in the image captured by the camera, we select the one closest to the vehicle.

B. Coordinate Frames and Notations

Some coordinate systems are defined for clearer explanation. The inertial frame is denoted as *i*-frame and the Earth-Centered, Earth-Fixed (ECEF) frame is denoted as *e*-frame. The navigation coordinate frame is denoted as *n*-frame,

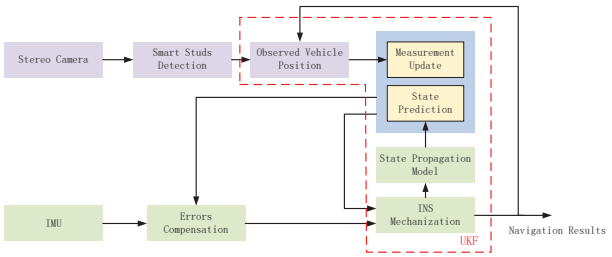

Fig. 1. Framework of the vehicle localization method.

whose coordinate axes point to the geographic north, east, and downward, respectively, with its origin coinciding with the centroid of the vehicle. The IMU body frame is denoted as *b*-frame. Its three axes respectively point to the vehicle's forward, rightward, and downward directions.

The coordinate systems in the model of binocular vision include the camera coordinate system (c-frame), the pixel coordinate system (p-frame), and the image coordinate system (im-frame). Below, the subscripts l and r will respectively represent the left camera and the right camera of the binocular camera. The detailed explanations about these coordinate systems are provided in reference [20].

C. Positional Relationship between Smart Stud and Vehicle

The positional relationship between the smart stud, camera, and IMU is shown in Fig. 2. Parameter $\boldsymbol{l}_c$ refers to the lever arm from the left camera of the binocular camera to the IMU, which can be measured. Let $\boldsymbol{X}_c$ denote the three-dimensional (3-D) coordinate of the smart stud in the c_l-frame, which can be obtained from the binocular vision measurements as follows:

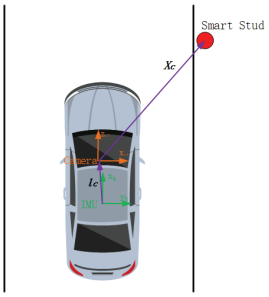

Fig. 2. Relationship between smart stud, camera, and INS.

$$\boldsymbol{X}_c = z_c \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u_l \\ v_l \\ 1 \end{bmatrix} \tag{1}$$

where f_x and f_y are the scale factors in the x and y directions respectively, (u_0, v_0) is the plane center of the image in the p_l-frame, these parameters can be obtained through calibration, (u_l, v_l) is the pixel coordinate of the observed object in the p_l-frame. In this study, (u_l, v_l) represents the pixel coordinates of the geometric center of the smart stud, which can be obtained

from the image, z_c is the depth value of the observed object, which can be described as follows:

$$z_c = f_x \frac{d}{|u_r - u_l|} \tag{2}$$

where d is the distance between the left camera and the right camera of the binocular camera, which can be measured, u_r can be obtained from the pixel coordinate in the p_r-frame.

The positional relationship between the smart stud, camera, and IMU can be represented as follows:

$$\boldsymbol{p}_I = \boldsymbol{p}_S - \boldsymbol{D}_R^{-1}\boldsymbol{C}_b^n\left(\boldsymbol{C}_c^b\boldsymbol{X}_c + \boldsymbol{l}_c\right) \tag{3}$$

where $\boldsymbol{p}_I$ is the position of the IMU, $\boldsymbol{p}_S$ is the position of the smart stud, $\boldsymbol{D}_R^{-1}$ is used to transform the north-east-down meters in the navigation frame into latitude φ, longitude λ, and altitude h, $\boldsymbol{C}_b^n$ is the direction cosine matrix from b-frame to n-frame, and $\boldsymbol{C}_c^b$ is the direction cosine matrix from c_l-frame to b-frame. $\boldsymbol{C}_b^n$ originates from the attitude algorithm of INS [21] and $\boldsymbol{C}_c^b$ can be obtained by calibration. $\boldsymbol{D}_R^{-1}$ can be expressed as follows:

$$\boldsymbol{D}_R^{-1} = diag\left(\begin{bmatrix} \frac{1}{R_M+h} & \frac{1}{(R_N+h)cos\varphi} & -1 \end{bmatrix}^T\right) \tag{4}$$

where R_M and R_N represent the radius of curvature in meridian and prime vertical, respectively.

III. Vehicle Position Fusion Based on Unscented Kalman Filter

A. INS Error Model

The error equations of INS can be established as follows [22]:

$$\dot{\boldsymbol{\phi}} = -\boldsymbol{\omega}_{in}^n \times \boldsymbol{\phi} + \delta\boldsymbol{\omega}_{in}^n - \delta\boldsymbol{\omega}_{ib}^n \tag{5}$$

$$\begin{aligned}\delta\dot{\boldsymbol{v}}^n = {} & \boldsymbol{C}_b^n\delta\boldsymbol{f}^b + \boldsymbol{f}^n \times \boldsymbol{\phi} - (2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \delta\boldsymbol{v}^n \\ & + \boldsymbol{v}^n \times (2\delta\boldsymbol{\omega}_{ie}^n + \delta\boldsymbol{\omega}_{en}^n) + \delta\boldsymbol{g}_p^n\end{aligned} \tag{6}$$

$$\begin{cases} \delta\dot{\varphi} = -\frac{v_N}{(R_M+h)^2}\delta h + \frac{1}{R_M+h}\delta v_N \\ \delta\dot{\lambda} = \frac{v_E tan\varphi}{(R_N+h)cos\varphi}\delta\varphi - \frac{v_E}{(R_N+h)^2 cos\varphi}\delta h \\ \qquad + \frac{1}{(R_N+h)cos\varphi}\delta v_E \\ \delta\dot{h} = -\delta v_D \end{cases} \tag{7}$$

where $\boldsymbol{\phi} = \begin{bmatrix} \phi_N & \phi_E & \phi_D \end{bmatrix}^T$ is the attitude error vector, $\boldsymbol{\omega}_{in}^n = \boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n$, $\delta\boldsymbol{\omega}_{in}^n = \delta\boldsymbol{\omega}_{ie}^n + \delta\boldsymbol{\omega}_{en}^n$, $\boldsymbol{\omega}_{ie}^n$ and $\delta\boldsymbol{\omega}_{ie}^n$ are the angular rate vector of the *e*-frame relative to the *i*-frame under the projection of the *n*-frame due to rotation of the earth and its error, $\boldsymbol{\omega}_{en}^n$ and $\delta\boldsymbol{\omega}_{en}^n$ are the angular rate vector of the *n*-frame relative to the *e*-frame under the projection of the *n*-frame and its error, $\delta\boldsymbol{\omega}_{ib}^n$ is the gyroscope drift in the *n*-frame, $\boldsymbol{v}^n = \begin{bmatrix} v_N & v_E & v_D \end{bmatrix}^T$ is the velocity vector, $\delta\boldsymbol{v}^n = \begin{bmatrix} \delta v_N & \delta v_E & \delta v_D \end{bmatrix}^T$ is the velocity error vector, $\boldsymbol{f}^n$ is the specific force vector, $\delta\boldsymbol{f}^b$ is the measurement error of the accelerometer, $\delta\boldsymbol{g}_p^n$ is the gravity error, $\delta\varphi$, $\delta\lambda$, and δh denote the errors of latitude, longitude, and altitude, respectively.

For UKF design, the state vector $\boldsymbol{X}$ is defined as follows:

$$\boldsymbol{X} = \begin{bmatrix} \delta\boldsymbol{p}^n & \delta\boldsymbol{v}^n & \boldsymbol{\phi} & \boldsymbol{b}_g & \boldsymbol{b}_a \end{bmatrix}^T \tag{8}$$

where $\delta\boldsymbol{p}^n = \begin{bmatrix} \delta\varphi & \delta\lambda & \delta h \end{bmatrix}^T$, $\boldsymbol{b}_g$ and $\boldsymbol{b}_a$ are the gyroscopes' biases and accelerometers' biases, respectively. Both $\boldsymbol{b}_g$ and $\boldsymbol{b}_a$ are modeled as first-order Gaussian Markov processes.

B. Measurement Equation Based on Smart Stud Vision Positioning

The estimated position of the vehicle $\hat{\boldsymbol{p}_I}$ can be derived from INS as follows:

$$\hat{\boldsymbol{p}_I} = \boldsymbol{p}_I + \boldsymbol{D}_R^{-1}\delta\boldsymbol{p}^n \tag{9}$$

The measurement position of the smart stud $\tilde{\boldsymbol{X}}_c$ is

$$\tilde{\boldsymbol{X}}_c = \boldsymbol{X}_c - \boldsymbol{n}_{rc} \tag{10}$$

where $\boldsymbol{n}_{rc}$ represents measurement noise.

The measurement vector $\boldsymbol{Z}$ is defined as the difference between the estimated position and the measurement position of the vehicle

$$\boldsymbol{Z} = \boldsymbol{D}_R(\hat{\boldsymbol{p}}_I - \tilde{\boldsymbol{p}}_I) \tag{11}$$

where $\tilde{\boldsymbol{p}}_I$ is the measurement position, which can be derived from (3) and (10).

Substituting (3), (9), and (10) into (11), the measurement equation can be written as follows:

$$\boldsymbol{Z} = -\boldsymbol{C}_b^n\boldsymbol{C}_c^b\boldsymbol{n}_{rc} + \delta\boldsymbol{p}^n \tag{12}$$

C. Unscented Kalman Filter

UKF employs UT to handle the nonlinearity in the propagation of mean and covariance [23]. Firstly, the UT is introduced. Define a nonlinear transformation $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$, wherein the variable $\boldsymbol{x}$, which is n-dimensional, has a mean of $\bar{\boldsymbol{x}}$ and a covariance matrix of $\boldsymbol{P}$. The sigma points, which are the sampling points of the UT, can be simulated as follows:

$$\begin{cases} \boldsymbol{x}^{(0)} = \bar{\boldsymbol{x}} & i = 0 \\ \boldsymbol{x}^{(i)} = \bar{\boldsymbol{x}} + (\sqrt{(n+\xi)\boldsymbol{P}})_i & i = 1 \sim n \\ \boldsymbol{x}^{(i)} = \bar{\boldsymbol{x}} - (\sqrt{(n+\xi)\boldsymbol{P}})_i & i = n+1 \sim 2n \end{cases} \tag{13}$$

where the superscript i is the index of sample points, $(\sqrt{\boldsymbol{P}})_i$ represents the i-th column of Cholesky decomposition of $\boldsymbol{P}$. The weights associated with these sigma points are computed as follows:

$$\begin{cases} \omega_m^{(0)} = \frac{\xi}{n+\xi} \\ \omega_c^{(0)} = \frac{\xi}{n+\xi} + (1 - a^2 + \beta) \\ \omega_m^{(i)} = \omega_c^{(i)} = \frac{\xi}{2(n+\xi)}, i = 1 \sim 2n \end{cases} \tag{14}$$

where $\omega_m^{(i)}$ and $\omega_c^{(i)}$ are the weights for the mean and covariance of the sigma points, respectively, $\xi = a^2(n+\kappa) - n$, a and κ are candidate parameters [24].

In this study, the nonlinear system can be represented as follows:

$$\begin{cases} \boldsymbol{X}(k+1) = \boldsymbol{f}(\boldsymbol{X}(k), \boldsymbol{\omega}(k)) \\ \boldsymbol{Z}(k) = \boldsymbol{h}(\boldsymbol{X}(k), \boldsymbol{\upsilon}(k)) \end{cases} \tag{15}$$

where $\boldsymbol{X}_{k+1}$ and $\boldsymbol{X}_k$ are the state vectors at time $k+1$ and time k, respectively, $\boldsymbol{Z}_k$ is the measurement vector at time k, $\boldsymbol{\omega}_k$ is the process noise vector, $\boldsymbol{\upsilon}_k$ is the measurement noise vector, $\boldsymbol{f}$ represents the motion equation, and $\boldsymbol{h}$ represents the measurement equation, $\boldsymbol{\omega}_k$ and $\boldsymbol{\upsilon}_k$ are uncorrelated zero-mean white noise sequences, namely:

$$\begin{cases} \boldsymbol{\omega}(k) \sim N(\boldsymbol{0}, \boldsymbol{Q}_k) \\ \boldsymbol{\upsilon}(k) \backsim N(\boldsymbol{0}, \boldsymbol{R}_k) \\ E(\boldsymbol{\omega}_k \boldsymbol{\upsilon}_j^T) = \boldsymbol{0} \end{cases} \tag{16}$$

where $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ are the process noise covariance matrix and the measurement covariance matrix, respectively.

Similar to EKF, the UKF algorithm is also divided into the following two steps.

1) *State Prediction:*

According to (13), sigma points can be obtained as follows:

$$\boldsymbol{X}^{(i)}(k|k) = \begin{bmatrix} \hat{\boldsymbol{X}}(k|k) \\ \hat{\boldsymbol{X}}(k|k) + \sqrt{(n+\xi)\boldsymbol{P}(k|k)} \\ \hat{\boldsymbol{X}}(k|k) - \sqrt{(n+\xi)\boldsymbol{P}(k|k)} \end{bmatrix}^T \tag{17}$$

The sigma points are propagated according to the motion equation:

$$\boldsymbol{X}^{(i)}(k+1|k) = \boldsymbol{f}(\boldsymbol{X}^{(i)}(k|k)) \tag{18}$$

The predicted state $\hat{\boldsymbol{X}}(k+1|k)$ and covariance matrix $\boldsymbol{P}(k+1|k)$ can be obtained as follows:

$$\hat{\boldsymbol{X}}(k+1|k) = \sum_{i=0}^{2n} \omega_m^{(i)} \boldsymbol{X}^{(i)}(k+1|k) \tag{19}$$

$$\begin{aligned} \boldsymbol{P}(k+1|k) = & \sum_{i=0}^{2n} \omega_c^{(i)} \left[\hat{\boldsymbol{X}}(k+1|k) - \boldsymbol{X}^{(i)}(k+1|k)\right] \\ & \left[\hat{\boldsymbol{X}}(k+1|k) - \boldsymbol{X}^{(i)}(k+1|k)\right]^T + \boldsymbol{Q}_k \end{aligned} \tag{20}$$

According to (13), (19), and (20), the new sigma points can be calculated as

$$\boldsymbol{X}^{(i)}(k+1|k) = \begin{bmatrix} \hat{\boldsymbol{X}}(k+1|k) \\ \hat{\boldsymbol{X}}(k+1|k) + \sqrt{(n+\xi)\boldsymbol{P}(k+1|k)} \\ \hat{\boldsymbol{X}}(k+1|k) - \sqrt{(n+\xi)\boldsymbol{P}(k+1|k)} \end{bmatrix}^T \tag{21}$$

The predicted observation is obtained according to the measurement equation:

$$\boldsymbol{Z}^{(i)}(k+1|k) = \boldsymbol{h}\left[\boldsymbol{X}^{(i)}(k+1|k)\right] \tag{22}$$

2) *Measurement Update*:

According to (14) and (22), the observation mean $\bar{\boldsymbol{Z}}(k+1|k)$, covariance matrix $\boldsymbol{P}_{Z_k Z_k}$, cross covariance matrix $\boldsymbol{P}_{X_k Z_k}$ and Kalman gain matrix $\boldsymbol{K}(k+1)$ can be calculated as follows:

$$\bar{\boldsymbol{Z}}(k+1|k) = \sum_{i=0}^{2n} \omega_m^{(i)} \boldsymbol{Z}^{(i)}(k+1|k) \tag{23}$$

$$\begin{aligned} \boldsymbol{P}_{Z_k Z_k} = & \sum_{i=0}^{2n} \omega_c^{(i)} \left[\boldsymbol{Z}^{(i)}(k+1|k) - \bar{\boldsymbol{Z}}(k+1|k)\right] \\ & \left[\boldsymbol{Z}^{(i)}(k+1|k) - \bar{\boldsymbol{Z}}(k+1|k)\right]^T + \boldsymbol{R}_k \end{aligned} \tag{24}$$

$$\begin{aligned} \boldsymbol{P}_{X_k Z_k} = & \sum_{i=0}^{2n} \omega_c^{(i)} \left[\boldsymbol{X}^{(i)}(k+1|k) - \bar{\boldsymbol{Z}}(k+1|k)\right] \\ & \left[\boldsymbol{X}^{(i)}(k+1|k) - \bar{\boldsymbol{Z}}(k+1|k)\right]^T \end{aligned} \tag{25}$$

$$\boldsymbol{K}(k+1) = \boldsymbol{P}_{X_k Z_k} \boldsymbol{P}_{Z_k Z_k}^{-1} \tag{26}$$

Finally, the state $\hat{\boldsymbol{X}}(k+1|k+1)$ and state covariance matrix $\boldsymbol{P}(k+1|k+1)$ are updated as follows:

$$\begin{aligned} \hat{\boldsymbol{X}}(k+1|k+1) = & \hat{\boldsymbol{X}}(k+1|k) + \boldsymbol{K}(k+1) \\ & \left[\boldsymbol{Z}(k+1) - \hat{\boldsymbol{Z}}(k+1|k)\right] \end{aligned} \tag{27}$$

$$\begin{aligned} \boldsymbol{P}(k+1|k+1) = & \boldsymbol{P}(k+1|k) - \boldsymbol{K}(k+1) \\ & \boldsymbol{P}_{Z_k Z_k} \boldsymbol{K}^T(k+1) \end{aligned} \tag{28}$$

IV. Performance Evaluation

To validate the proposed method in this paper, we recorded a driving dataset using a vehicle shown in Fig. 3. It includes a camera (Stereolabs, ZED 2i), a global positioning system (GPS) receiver (Topgnss, GNSS 100G), and an integrated navigation system (Xsens, MTi-680G). The camera mounted at the front of the vehicle, is utilized for smart stud identification. The frame rate of the camera is 30 fps, and it has 1280 $\times$ 720 image resolution with a field of view (FOV) as 90°×60°. The GPS receiver is mounted on the roof of the vehicle. The sampling frequency of the GPS receiver is 1 Hz, and its positioning mode is selected as differential GPS. The integrated navigation system is mounted on the internal platform in the trunk of the vehicle. The system combines IMU data and RTK data to provide reference data. The parameters of the IMU are given in Table I.

A total of 42 smart studs are deployed along both sides of the road in the campus of Xidian University in Xi'an, China. The experimental environment is shown in Fig. 4, with the distance between smart studs set to 15 meters.

Fig. 3. The experiment vehicle with the camera, the GPS receiver, and the integrated navigation module.

TABLE I
PARAMETERS OF IMU

Parameters	Unit	Value
Sample Rate	Hz	100
Gyro Bias Stability	$^\circ/h$	8
Gyro Random Walk	$^\circ/\sqrt{h}$	0.17
Accelerometer Bias Stability	μg	15
Velocity Random Walk	$m/s/\sqrt{h}$	0.02

A. Validation of the Proposed Method

The results of the proposed smart stud-aided INS integration method and GPS-aided INS integration method (GPS/INS) were compared.

Fig. 5 gives the positioning errors of different navigation solutions in different directions. The mean and maximum position errors in the north (N), east (E), and down (D) directions of different methods are reported in Table II. Compared to the GPS/INS, the mean and maximum position errors in different directions of the proposed method, respectively, reduce by 40.7% and 39.3% in the north direction, 33.0% and 14.3% in the east direction, 88.5% and 86.1% in the down direction. The results demonstrate that the localization method assisted by vision positioning achieves higher accuracy than GPS/INS. This is because GNSS positioning accuracy is influenced by various factors such as weather conditions, obstructions, and device processing speed, which may lead to fluctuations or offsets in GNSS positioning information. This will result in significant positioning errors for GNSS-based localization methods. Fig. 6 illustrates a comparison between the trajectories of various navigation solutions and reference trajectories.

B. Performance Comparison between UKF and EKF

To compare the performance of the proposed UKF algorithm and EKF algorithm in fusing vision positioning results and INS information, the same real-world experimental data was also processed using EKF algorithm [16], which is the most widely used fusion localization algorithm. Fig. 7 shows the positioning errors of different algorithms in different directions, and Table III shows the performance of the EKF. From Table II and Table III, we can readily conclude that the localization errors of UKF are smaller than that of EKF.

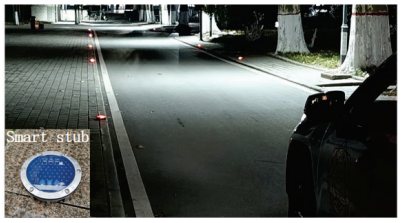

Fig. 4. Experimental scenario.

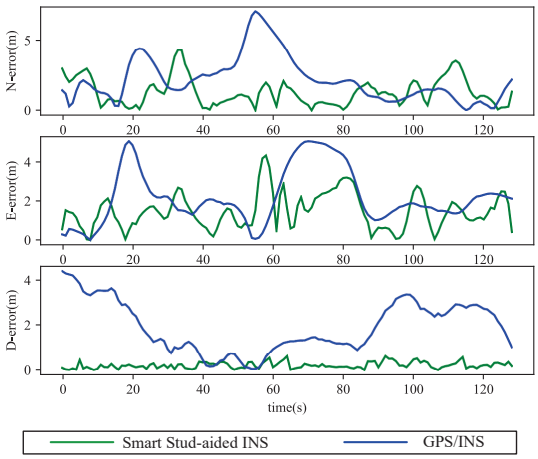

Fig. 5. The 3D (north, east, down) position errors of smart stud-aided INS integration method and GPS/INS.

TABLE II
PERFORMANCE OF SMART STUD-AIDED INS AND GPS/INS

Methods	Orientation	Mean Pos. Error (m)	Max. Pos. Error (m)
Proposed Method	N	1.28	4.28
	E	1.46	4..32
	D	0.22	0.61
GPS/INS	N	2.16	7.05
	E	2.18	5.04
	D	1.91	4.40

TABLE III
PERFORMANCE OF EKF ALGORITHM

Orientation	Mean Pos. Error (m)	Max. Pos. Error (m)
N	1.66	4.76
E	1.72	5.43
D	0.22	0.62

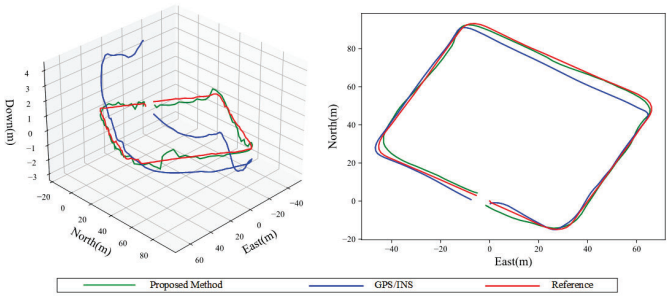

Fig. 6. 3D and 2D trajectories of smart stud-aided INS and GPS/INS.

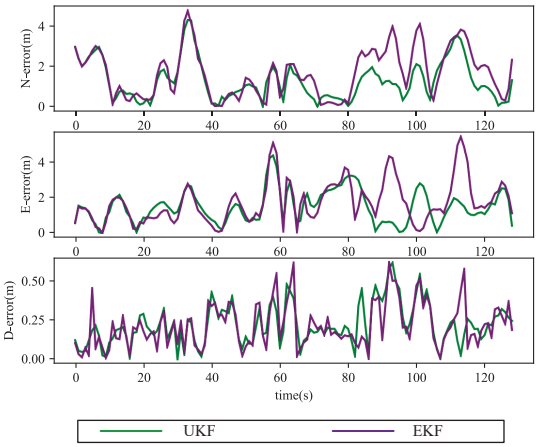

Fig. 7. The 3D position errors of UKF and EKF in fusing vision positioning results and INS information

V. Conclusion and Future Work

This paper proposed an IoT-based vehicle localization method in GNSS-denied environments. In this method, the HSV color model was used to identify the LED lights of ubiquitously deployed IoT devices termed smart studs, thereby obtaining vehicle's positions by vision positioning. Subsequently, the UKF was employed to fuse vision positioning result with INS navigation information. We compared the localization performance of the proposed method with the GPS/INS. Experimental results demonstrated that the proposed method in this paper achieves higher localization accuracy, which shows the potential of the IoT-based localization method to function satisfactorily without relying on GNSS even for a long time. Additionally, the performance of UKF and EKF in fusing vision positioning with INS information were also compared and the results indicated that UKF outperforms EKF in terms of positioning accuracy.

However, the method proposed in this paper also has its limitations. To better utilize the light to identify the IoT devices, the experiments were conducted at night, which may result in larger errors in vision measurements. Our future work will focus on finding superior algorithms to reduce the errors in vision measurements.

Acknowledgment

This research is supported by NSFC grant, Grant number: U21A20446.

References

[1] J. Liu and G. Guo, "Vehicle localization during gps outages with extended kalman filter and deep learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–10, 2021.

[2] S. Lu, Y. Gong, H. Luo, F. Zhao, Z. Li, and J. Jiang, "Heterogeneous multi-task learning for multiple pseudo-measurement estimation to bridge gps outages," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–16, 2021.

[3] H. Tang, T. Zhang, X. Niu, J. Fan, and J. Liu, "Impact of the earth rotation compensation on mems-imu preintegration of factor graph optimization," *IEEE Sensors Journal*, vol. 22, no. 17, pp. 17 194–17 204, 2022.

[4] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 829–846, 2018.

[5] C. Wang, H. Huang, Y. Ji, B. Wang, and M. Yang, "Vehicle localization at an intersection using a traffic light map," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1432–1441, 2019.

[6] J. Dong, X. Ren, S. Han, and S. Luo, "Uav vision aided ins/odometer integration for land vehicle autonomous navigation," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 4825–4840, 2022.

[7] J. Gao, K. Li, and Y. Chen, "Study on integration of fog single-axis rotational ins and odometer for land vehicle," *IEEE Sensors Journal*, vol. 18, no. 2, pp. 752–763, 2018.

[8] W. Ouyang, Y. Wu, and H. Chen, "Ins/odometer land navigation by accurate measurement modeling and multiple-model adaptive estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 1, pp. 245–262, 2021.

[9] N. Vagle, A. Broumandan, and G. Lachapelle, "Multiantenna gnss and inertial sensors/odometer coupling for robust vehicular navigation," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4816–4828, 2018.

[10] C. Campos, R. Elvira, J. J. Gomez Rodriguez, J. M. M. Montiel, and J. D. Tardos, "Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam," *IEEE TRANSACTIONS ON ROBOTICS*, vol. 37, no. 6, pp. 1874–1890, DEC 2021.

[11] J. Li, J. Dai, Z. Su, and C. Zhu, "Rgb-d based visual slam algorithm for indoor crowd environment," *JOURNAL OF INTELLIGENT & ROBOTIC SYSTEMS*, vol. 110, no. 1, MAR 2024.

[12] H. Wang, C. Xue, Y. Zhou, F. Wen, and H. Zhang, "Visual semantic localization based on hd map for autonomous vehicles in urban scenarios," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 11 255–11 261.

[13] A. Chalvatzaras, I. Pratikakis, and A. A. Amanatiadis, "A survey on map-based localization techniques for autonomous vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 2, pp. 1574–1596, 2023.

[14] Y. Lu, H. Ma, E. Smart, and H. Yu, "Real-time performance-focused localization techniques for autonomous vehicle: A review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6082–6100, 2022.

[15] X. Qu, B. Soheilian, and N. Paparoditis, "Vehicle localization using mono-camera and geo-referenced traffic signs," in *2015 IEEE Intelligent Vehicles Symposium (IV)*, 2015, pp. 605–610.

[16] H. Hu, K. Li, W. Liang, Q. Li, and Z. Xie, "Kilometer sign positioning-aided ins/odometer integration for land vehicle autonomous navigation," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 4143–4158, 2023.

[17] G. Mao, Y. Hui, X. Ren, C. Li, and Y. Shao, "The internet of things for smart roads: A road map from present to future road infrastructure," *IEEE Intelligent Transportation Systems Magazine*, vol. 14, no. 6, pp. 66–76, 2022.

[18] Y. Sun, H. Wang, W. Quan, X. Ma, Z. Tao, M. Elhajj, and W. Y. Ochieng, "Smart road stud-empowered vehicle magnetic field distribution and vehicle detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 7, pp. 7357–7362, 2023.

[19] T.-H. S. Li, Y.-H. Wang, C.-C. Chen, and C.-J. Lin, "A fast color information setup using ep-like pso for manipulator grasping color objects," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 645–654, 2014.

[20] S. Lin, W. Li, C. Wang, and Y. Tang, "Distance measurement of underwater target based on stereo vision," in *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2017, pp. 97–102.

[21] X. Niu, Y. Peng, Y. Dai, Q. Chen, C. Guo, and Q. Zhang, "Camera-based lane-aided multi-information integration for land vehicle navigation," *IEEE/ASME Transactions on Mechatronics*, vol. 28, no. 1, pp. 152–163, 2023.

[22] E. Zhang and N. Masoud, "Increasing gps localization accuracy with reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 2615–2626, 2021.

[23] S. J. Julier and J. K. Uhlmann, "A new extension of the kalman filter to nonlinear systems," *Proc.aerosense Int.symp.aerospace/defence Sensing Simulation and Controls*, 1997.

[24] D. Feng, C. Wang, C. He, Y. Zhuang, and X.-G. Xia, "Kalman-filter-based integration of imu and uwb for high-accuracy indoor positioning and navigation," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3133–3146, 2020.