Information Fusion 46 (2019) 51–62

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Multi-sensor fusion methodology for enhanced land vehicle positioning

Li Xu[a,*], Chen Wei[b], Chan Chingyao[c], Li Bin[d], Song Xianghui[d]

[a] School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China
[b] College of Electrical Engineering and Control Science, Nanjing Tech University, Nanjing 211816, China
[c] Institute of Transportation Studies, University of California at Berkeley, Berkeley, CA 94720, USA
[d] Key Laboratory of Technology on Intelligent Transportation Systems, Ministry of Transport, Research Institute of Highway Ministry of Transport, Beijing 100088, China

ARTICLE INFO

Keywords:
Vehicle positioning
Empirical mode decomposition
NARX
INS/GPS
MEMS

ABSTRACT

It is a main challenge for land vehicles to achieve reliable and low-cost navigation solution in various situations, especially when Global Positioning System (GPS) is not available. To address this challenge, we propose an enhanced multi-sensor fusion methodology to fuse the information from low-cost GPS, MEMS Inertial Measurement Unit (IMU), and digital compass in this paper. First, a key data preprocessing algorithm based on Empirical Mode Decomposition (EMD) interval threshold filter is developed to remove the noises in inertial sensors so as to offer more accurate information for subsequent modeling. Then, a Least-Squares Support Vector Machine (LSSVM)-based nonlinear autoregressive with exogenous input (NARX) model (LSSVM-NARX) is designed and augmented with Kalman filter (KF) to construct a novel LSSVM-NARX/KF hybrid strategy. In case of GPS outages, the recently updated LSSVM-NARX is adopted to predict and compensate for the INS position errors. Finally, the performance of proposed methodology was evaluated with real-world data collected in urban settings including typical driving maneuvers. The results indicate that the proposed methodology can achieve remarkable enhancement in positioning accuracy in GPS-denied environments.

1. Introduction

In recent years, reliable and accurate vehicle position information has increasingly emphasized its importance in many applications of intelligent transportation system (ITS), including advanced driver assistance, route guidance, freight management, automatic emergency calls, and electronic fee collection [1–6]. For example, in advanced driver assistance systems (ADASs), continuous and precise knowledge of the vehicle's position is required to assess safety threats and trigger warning or intervention of these systems [4].

Presently, Global Positioning System (GPS) is the most accessible vehicle navigation technique and has been considered as a standard solution in a wide range of transportation applications [2,7]. However, when satellite signal is intermittent or obstructed in GPS-denied environments such as urban canyons and tunnels, the standalone GPS cannot provide continuous and reliable positioning information [8,9].

To address such problem, GPS is commonly integrated with Inertial Navigation System (INS) to create an integrated INS /GPS positioning system, which is more robust to environmental factors such as GPS signal blockage [10]. INS is an autonomous system which contains an Inertial Measurement Unit (IMU) and an onboard computer [11]. The IMU measures the linear accelerations and angular velocities of a land vehicle applying three accelerometers and three gyroscopes [12]. When the initial conditions of navigation parameters are given, the onboard computer can process IMU measurements to provide current position, velocity, and attitude. Thus, INS can bridge GPS outages and maintain the continuity of navigation solution [13].

For INS/GPS integrated system, the navigation solution is traditionally achieved by using Kalman filter (KF) [14,15]. However, the major inadequacy related to KF is the requirement of accurate stochastic modeling of sensor errors [14,16]. For high-grade INS sensors whose error statistical properties can be accurately acquired, KF can provide a proper positioning performance even when it works in prediction mode for a long time owing to GPS unavailability. However, the high cost of high-grade INS inhibits it from being integrated in commercially affordable land vehicle navigation systems [14]. Recently, the low-cost INS based on microelectromechanical system (MEMS) technology has emerged as a particularly attractive solution for manufacturers due to its outstanding advantages such as low cost, low power consumption, and small size [2,14,17]. Nevertheless, it is difficult to accurately determine the error statistical properties of MEMS INS sensors because they are characterized by much large noise, bias, and drift errors in their measurements when compared with higher-grade INS sensors [17–19]. Even for a short-duration GPS outage, the accuracy of

* Corresponding author.
E-mail addresses: lixu.mail@163.com (X. Li), henuchenwei.20@163.com (W. Chen), cychan@berkeley.edu (C. Chan), libin@itsc.cn (B. Li), sxh@itsc.cn (X. Song).

https://doi.org/10.1016/j.inffus.2018.04.006
Received 15 November 2016; Received in revised form 23 March 2017; Accepted 25 April 2018
Available online 27 April 2018
1566-2535/ © 2018 Elsevier B.V. All rights reserved.

MEMS INS/GPS integrated positioning system will degrade sharply when KF cannot be updated by GPS measurements.

To overcome the drawbacks of KF, Artificial Neural Network (ANN) approaches have been proposed as alternative methods to improve traditional INS/GPS integration. The existing ANN-based approaches include Multilayer Perceptron Neural Networks (MLPNN) [20–23], Adaptive Neuro-fuzzy Inference System (ANFIS) [14,24–26], Radial Basis Function Neural Network (RBFNN) [27], Random Forest Regression (RFR) [11], and Support Vector Machine (SVM) [28], etc. Basically, the ANN-based methods mentioned above are based on relating INS error at a certain time instant to the corresponding INS output at the same instant. These methods didn't consider the error dependency on past outputs of INS. The INS position errors commonly accumulate over time. Therefore, during relatively long GPS outages, these methods may not be able to provide a reliable positioning solution.

To address the weakness related to the aforementioned ANN-based methods, Noureldin et al. [29] suggested the utilization of Input-Delayed Neural Network (IDNN) to model both the INS velocity and position errors based on current and some past samples of INS velocity and position, and demonstrated better performance during GPS outages in comparison with conventional ANN-based methods. According to Noureldin et al. [29], Saadeddin et al. [30] developed a dynamic ANFIS to forecast the INS position errors in the case of GPS signal blockage based on the current and previous INS output. However, the methods adopted in [29] and [30] attempted to replace KF with neural network, which would show a very limited effectiveness when applied to a MEMS INS/GPS integration owing to the relatively high instrument bias, noise and random error of MEMS inertial sensors. Moreover, the influence of vehicle running states on positioning performance was not considered.

In addition, for the ANN-based methods mentioned above, one common drawback is that they all lack a suitable and effective pretreatment to mitigate the high-level noises in MEMS inertial sensor raw measurement before the data integration occurs. For the land vehicle navigation application, the high-level noises in MEMS IMU measurement usually comprise uncertain colored noises including the vibratory noises caused by the engine or road conditions and slowly varying bias drifts, in addition to white noises. The de-noising of inertial sensor is traditionally achieved using wavelet filter [14]. However, this method is based on white noise model and is incapable of removing the uncertain colored noises. Compared with white noises, the uncertain colored noises make the non-linear input/output function relationship, which is to be modeled using ANN, more complex. This will limit the generalization ability of ANN and distort its prediction accuracy during GPS outages.

To achieve reliable and accurate positioning information for land vehicles during GPS unavailability, an enhanced fusion methodology is proposed to bridge the GPS outages. In this methodology, a key data preprocessing algorithm based on Empirical Mode Decomposition (EMD) interval threshold filter for inertial data de-noising and a novel LSSVM-NARX/KF hybrid strategy for INS error compensation are developed. These two novel aspects can ensure to obtain more reliable positioning solution even in GPS-denied environments, which is further explained as follows:

1) A key data preprocessing algorithm based on Empirical Mode Decomposition (EMD) interval threshold filter, abbreviated as EITF, is first proposed to de-noise the MEMS inertial sensor. The proposed preprocessing algorithm is capable of suppressing the vibratory noises, partial bias drift noises and high-frequency white noises in MEMS inertial sensor measurements prior to INS strap-down calculation mechanization and INS/GPS integration. In comparison with raw inertial sensor measurements, the preprocessed inertial data is more accurate, which can considerably enhance the performance of subsequent information fusion.
2) A LSSVM-NARX/KF hybrid strategy is then designed to compensate for the INS position errors during GPS outages. Specifically, a nonlinear autoregressive with exogenous input (NARX) model based on Least-Squares Support Vector Machine (LSSVM) is designed and augmented with KF, termed as LSSVM-NARX/KF hybrid strategy, to predict and compensate for the INS position errors when GPS is unavailable. The developed NARX model takes not only the INS error trend but also the vehicle running states into account, which will provide a better prediction of MEMS INS position error. The LSSVM is utilized to identify the proposed NARX model because of its improved ability to model the highly non-linear input-output functional relationship.

The organization of the paper is as follows: Section 2 gives an outline of the enhanced multi-sensor fusion methodology proposed in this paper. Section 3 presents the INS error models and KF. The key data preprocessing algorithm is designed in Section 4. Section 5 explains the detailed implementation of the novel LSSVM-NARX model. Section 6 provides the road experimental results while Section 7 makes the concluding remarks.

2. Overview of proposed multi-sensor fusion methodology

The whole mechanism of the proposed multi-sensor fusion methodology is shown in Fig. 1.

First, the key data preprocessing algorithm based on EITF is developed to preprocess the MEMS inertial sensor data. The EITF is developed by integrating the advantages of EMD and fractional Gaussion noise (fGn). EMD is a powerful adaptive and data-driven decomposition tool that is suitable for nonlinear and non-stationary signals, and fGn is the generalization of white noise and capable of reflecting the correlation characteristic of noises in inertial sensor output [31,32]. The proposed preprocessing algorithm can effectively and efficiently suppress the vibratory noises, partial bias drift noises and high-frequency white noises in IMU raw measurements. Therefore, more accurate information is provided for subsequent modeling. Moreover, the real-time implementation of the proposed preprocessing algorithm is realized by using a sliding window, where data window moves forward when the current preprocessing is completed.

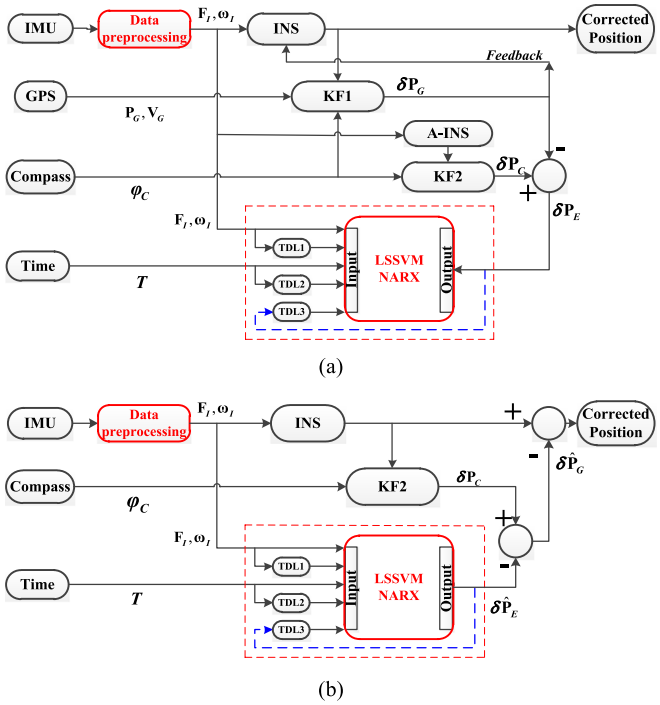

Fig. 1. Proposed system structures. (a) Update mode; (b) prediction mode. TDL: Tapped Delay Line.

To compensate for the MEMS INS position errors during GPS outages, the LSSVM-NARX/KF hybrid strategy is designed. The hybrid strategy comprises update mode and prediction mode, which are illustrated in Fig. 1(a) and (b), respectively. When the GPS signal is available, the strategy operates in the update mode, as shown in Fig. 1(a). A closed-loop KF1 is designed and updated by the measurements of GPS and digital compass, which can produce accurate fully updated positioning errors (FUPEs) due to good observability. The error states estimated by the KF1 including the position, velocity and attitude error states, and the inertial sensor errors are fed back to the INS for the error correction. Besides, an open-loop KF2 and an auxiliary-INS (A-INS) are constructed. With the assistance of A-INS, the KF2 is updated by the measurement of digital compass and provides partly updated positioning errors (PUPEs), but does not fed back to the INS. In actual implementation of KF2 and A-INS, a periodic moving window with a certain size (e.g., 45 s) is applied in the update mode, which is used to simulate the error growth trend under the condition of inertial navigation aided only by the compass (assumed that GPS completely fails for a period, e.g., 45 s). At the beginning of every moving window, the initial values of KF2 and A-INS are set equal to those of KF1 and INS, respectively. Thereafter, the KF2 and A-INS do not receive the information from KF1 and INS, and they are operating alone in open-loop mode until the end of every moving window. During this period, the error between FUPE and PUPE at each time is transferred to the LSSVM-NARX model as the desired output. The inputs of LSSVM-NARX model include exogenous inputs and inner inputs. The exogenous inputs contain current and some past samples of IMU measurements and the time elapsing since GPS is assumed to fail, while the inner inputs comprise some past samples of the error between FUPE and PUPE. The parameters of the LSSVM-NARX model are continuously updated until a GPS outage occurs. Obviously, similar to the KF2, the LSSVM-NARX model works with the periodic moving window mechanism in the update mode. As a time delay unit, the Tapped Delay Line (TDL) gives the LSSVM-NARX model the latest short-time memories. In addition, a sequence of IMU measurements provide some information about the vehicle dynamic variation, which is important for the prediction accuracy of the error between FUPE and PUPE.

When satellite signal blockage occurs in GPS-denied environments, the proposed strategy switches to the prediction mode shown in Fig. 1(b). In this situation, there are no KF1 and A-INS. However, the open-loop KF2 can still be updated by the heading angle information and products PUPEs all the time. Current and some past samples of IMU measurements and the time elapsing since the GPS outage happens, together with some past samples of the error between FUPE and PUPE are used as inputs to the recently updated LSSVM-NARX model. Based on these inputs, the LSSVM-NARX model provides the estimation of current error between FUPE and PUPE. Then, utilizing the available PUPE at current time, the predicted current FUPE during GPS outages can be acquired. Finally, the predicted current FUPE can be removed from the INS corresponding position component to achieve open-loop correction to obtain reliable and accurate positioning solution.

3. INS error models and Kalman filter

3.1. INS error models

The coordinate systems and frames used in this paper are given as follows:

(1) Earth-Centered Inertial Frame (i-frame): Its origin is at the center of mass of the Earth. The z-axis is along axis of the Earth's rotation through the conventional terrestrial pole, and x-axis is in the equatorial plane pointing towards the vernal equinox. The y-axis completes a right-handed system.
(2) Earth-Centered Earth-Fixed Frame (e-frame): It shares the same origin and z-axis as the i-frame, while the x-axis passes through the intersection of the equatorial plane and the reference meridian, and the y-axis completes the right-hand coordinate system in the equatorial plane.
(3) Local geographic frame (n-frame): It is chosen as navigation frame in this paper. The origin is at a point near the vehicle on the surface of earth. The x-axis, y-axis and z-axis point to east, north and up (ENU), respectively.
(4) Body-frame (b-frame): Its origin is at the center of gravity of the vehicle. The x-axis, y-axis and z-axis are along the longitudinal direction forward, the transversal direction left and vertical direction upward, respectively.

According to INS mechanization, the INS attitude error equation can be described as follow [27,33]

$$\dot{\boldsymbol{\psi}}^n = \delta\boldsymbol{\omega}_{ie}^n + \delta\boldsymbol{\omega}_{en}^n - (\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \boldsymbol{\psi}^n - \boldsymbol{\varepsilon}^n \tag{1}$$

where $\boldsymbol{\psi}^n = [\psi_e \ \ \psi_n \ \ \psi_u]^T$ denotes attitude error vector, the superscript "T" denotes the matrix transposition. $\boldsymbol{\omega}_{ie}^n = [0 \ \ \omega_e cosL \ \ \omega_e sinL]^T$ represents the angular rate of rotation of the Earth to the n-frame. $\boldsymbol{\omega}_{en}^n = \left[\frac{-V_n}{R_{n+h}} \ \ \frac{V_e}{R_{e+h}} \ \ \frac{V_e tanL}{R_e+h}\right]^T$ represents the angular rate of the rotation of n-frame relative to Earth. $\boldsymbol{\varepsilon}^n = [\varepsilon_e \ \ \varepsilon_n \ \ \varepsilon_u]^T$ is the random drifts of gyroscopes in the n-frame, and each element of it is modeled as a first-order Gauss Markov process in this paper. L and h denote latitude and height of the land vehicle location, respectively. V_e and V_n represent the velocities of the land vehicle in east and north direction, respectively.

The velocity error equation of INS is given by

$$\delta\dot{\mathbf{V}}^n = \mathbf{f}^n \times \boldsymbol{\psi}^n - (2\delta\boldsymbol{\omega}_{ie}^n + \delta\boldsymbol{\omega}_{en}^n) \times \mathbf{V}^n - (2\boldsymbol{\omega}_{ie}^n + \boldsymbol{\omega}_{en}^n) \times \delta\mathbf{V}^n + \nabla^n \tag{2}$$

where $\delta\mathbf{V}^n = [\delta V_e \ \ \delta V_n \ \ \delta V_u]^T$ denotes the velocity errors. $\mathbf{f}^n = [f_e \ \ f_n \ \ f_u]^T$ is the specific force vector. $\mathbf{V}^n = [V_e \ \ V_n \ \ V_u]^T$ represents the velocity. $\nabla^n = [\nabla_e \ \ \nabla_n \ \ \nabla_u]^T$ stands for the biases of accelerometer in the n-frame, and each element of it is modeled as a first-order Gauss Markov process in this paper.

According to $\dot{L} = \frac{V_n}{R_n+h}$, $\dot{\lambda} = \frac{V_e \mathrm{sec}L}{R_e+h}$ and $\dot{h} = V_u$, the position error equation of INS can be presented as

$$\begin{cases} \delta\dot{L} = \frac{\delta V_n}{R_n+h} - \frac{V_n}{(R_n+h)^2}\delta h \\ \delta\dot{\lambda} = \frac{\mathrm{sec}L}{R_e+h}\delta V_e + \frac{V_e}{R_e+h}\mathrm{sec}L\cdot\tan L\cdot\delta L - \frac{V_e \mathrm{sec}L}{(R_{ne}+h)^2}\delta L \\ \delta\dot{h} = \delta V_u \end{cases} \tag{3}$$

where λ represents the longitude. δL, $\delta\lambda$, and δh denote the errors of latitude, longitude and height, respectively. More information about the error models of INS can be found in References [33].

3.2. Kalman filter (KF)

In this paper, two fifteen-state Kalman filters are constructed based on the perturbation error model of INS. Derived from the error models in (1)–(3), the state equation of KF can be expressed as [33]

$$\dot{\mathbf{X}} = \mathbf{F}\mathbf{X} + \mathbf{G}\boldsymbol{\mu} \tag{4}$$

where
$\mathbf{X} = [\delta L \ \ \delta\lambda \ \ \delta h \ \ \delta V_e \ \ \delta V_n \ \ \delta V_u \ \ \psi_e \ \ \psi_n \ \ \psi_u \ \ \nabla_{bx} \ \ \nabla_{by} \ \ \nabla_{bz} \ \ \varepsilon_{bx} \ \ \varepsilon_{by} \ \ \varepsilon_{bz}]^T$; δV_e, δV_n, and δV_u represent three velocity errors, respectively. ψ_e, ψ_n and ψ_u denote pitch error, roll error, and yaw error, respectively. ∇_{bx}, ∇_{by} and ∇_{bz} represent biases of three accelerometers in the b-frame, respectively. ε_{bx}, ε_{by} and ε_{bz} represent random drifts of three gyroscopes in the b-frame, respectively. $\mathbf{F}$ and $\mathbf{G}$ represent the state transition matrix and noise input matrix, respectively. $\boldsymbol{\mu}$ is the system noise and is assumed to be uncorrelated and zero-mean white noise.

It should be noted that three accelerometer drifts and three gyro drifts above are all modeled as first-order Gauss Markov processes, which have the similar formulation as follows:

$$\varepsilon(k) = (1 - \beta\Delta t)\varepsilon(k-1) + \sqrt{2\beta\sigma_\varepsilon^2}w(k)\Delta t \tag{5}$$

where β is the reciprocal of correlation time, Δt is the sampling interval, and σ_ε^2 is the variance of sensor white noise $w(k)$. The specific model parameters can be determined through autocorrelation analysis and Allan analysis. For the KF1 and KF2 in Fig. 1, they have the similar state equation form as shown in (4). Besides, the FUPEs $\boldsymbol{\delta P_G}$ and the PUPEs $\boldsymbol{\delta P_C}$ have the similar form, i.e., $[\delta L \ \ \delta\lambda \ \ \delta h]^T$. However, $\boldsymbol{\delta P_G}$ is estimated by KF1 while $\boldsymbol{\delta P_C}$ is estimated by KF2, as shown in Fig. 1(a). The difference between them is denoted as $\boldsymbol{\delta P_E}$. The measurement equation is given by

$$\mathbf{Z} = \mathbf{H}\cdot\mathbf{X} + \mathbf{v} \tag{6}$$

where $\mathbf{Z}$ is the observation vector. $\mathbf{H}$ is the observation matrix. $\mathbf{v}$ represents the observation noise and is assumed to be uncorrelated and zero-mean white noise.

KF1 and KF2 have different measurement vectors. For KF1, $\mathbf{Z} = [\delta L \ \ \delta\lambda \ \ \delta h \ \ \delta V_e \ \ \delta V_n \ \ \delta V_u \ \ \delta\varphi]^T$ because it is updated based on the measurements of GPS and digital compass. For KF2, $\mathbf{Z} = [\delta\varphi]$ since it is updated only using the measurements of digital compass. In this paper, the measurement information are given as

$$\begin{cases} \delta L = L_I - L_G \\ \delta\lambda = \lambda_I - \lambda_G \\ \delta h = h_I - h_G \\ \delta V_e = V_{e,I} - V_{e,G} \\ \delta V_n = V_{n,I} - V_{n,G} \\ \delta V_u = V_{u,I} - V_{u,G} \\ \delta\varphi = \varphi_I - \varphi_C \end{cases} \tag{7}$$

where L_I, λ_I, and h_I denote INS predicted latitude, longitude, and height, respectively. L_G, λ_G, and h_G denote GPS-derived latitude, longitude, and height, respectively. $V_{e,\,I}$, $V_{n,\,I}$, and $V_{u,\,I}$ represent the INS predicted velocities in east, north and upward directions, respectively. $V_{e,\,G}$, $V_{n,\,G}$, and $V_{u,\,G}$ represent the GPS-derived velocities in east, north and upward directions, respectively. $\delta\varphi$ is the heading angle error. φ_I is the heading angle predicted by INS, and φ_C is the heading angle measured by compass.

4. Key data preprocessing algorithm

For the low-cost MEMS inertial sensors installed on the vehicle, their outputs are usually contaminated by many uncertain noises including the vibratory noises caused by the engine or road conditions, slowly varying bias drifts and high-frequency white noises [17]. The vibratory noises are caused by external disturbances, while the high-frequency white noises and slowly varying bias drift noises can be considered as sensor internal noises because they are mainly caused by MEMS inertial sensors themselves. To remove these noises in inertial sensors, the commonly-used method is wavelet filter [14]. However, the complex noise characteristics of MEMS inertial sensors will distort the performance of classic wavelet filter because it is based on the assumption that the actual signal is contaminated by pure white noise. To overcome the limitation of white noise assumption, this paper proposes an improved Empirical Mode Decomposition (EMD) interval threshold filter (EITF) by combining the advantages of EMD and fractional Gaussion noise (fGn) model. The proposed EITF algorithm can effectively suppress the colored noises such as the vibratory noises caused by the engine or road conditions and partial slowly varying bias drifts, as well as the white noises in IMU raw measurements, and thus can offer more accurate information for subsequent modeling. To achieve a real-time solution, a W-length sliding window is adopted. After processing the data of current window, the data window moves forward for the next processing.

4.1. Brief review of EMD

Empirical mode decomposition (EMD) is a powerful time-frequency analysis techniques proposed by Huang et al. [31,34]. The goal of EMD is to break a signal down into a number of intrinsic mode function (IMF) by their local characteristic time scales [34].

To perform EMD on a signal, the first step is to identify the local maxima and minima of it. Then, the upper and lower envelopes are constructed from these maxima and minima by using cubic-spline interpolations. After that, the mean envelope is obtained by averaging the two envelopes. The difference between the signal and mean envelope, named as an intermediate signal, is considered as an IMF if it satisfies the following conditions: (a) the number of extrema and the number of zero crossings must be the same or differ at most by one; (b) at any point, the local average is zero. Otherwise, the above process is repeatedly cycled until the intermediate signal is identified as an IMF and subtracted from the signal. This iteration process is carried out to extract additional IMFs until the intermediate signal becomes a constant or a monotonic function.

The major advantage of the EMD is that the basic functions are derived from the signal itself, rather than the fixed and pre-defined basis functions that utilized by the traditional methods, such as wavelet decomposition [31]. Thus, EMD is a completely self-adaptive signal processing tool that is especially suited for non-stationary and nonlinear signals, such as inertial sensor signals utilized in this paper.

4.2. FGn model for inertial sensor noises

A fractional Gaussion noise is a zero-mean stationary process whose statistical properties are determined merely by a key parameter: Hurst parameter H $(0 < H < 1)$ [32]. It is a versatile model for a broadband noise that has not any dominant frequency band. The autocorrelation sequence of fGn is given as following [35]

$$C_H(\tau) = \frac{\sigma^2}{2}(|\tau - 1|^{2H} - 2|\tau|^{2H} + |\tau + 1|^{2H}) \tag{8}$$

where τ is the lag length and σ^2 is the variance. Particularly, when $H = 0.5$ the fGn is equivalent to white noise, i.e. $C_{0.5}(\tau) = \sigma^2$ for $\tau = 0$, and $C_{0.5}(\tau) = 0$ for all $\tau \neq 0$. Therefore, fGn is the generalization of white noise.

To model inertial sensor errors utilizing fGn, the main task is to estimate the Hurst parameter H to reveal the correlation degree of errors. Several useful methods are developed to estimate H, which include R/S method, autocorrelation analysis method, Fourier analysis, maximum likelihood method, and Aggregated Variance method [36], etc. In this paper, the Aggregated Variance method is adopted because it requires smaller datasets and exhibits less bias, leading to a fast convergence and a better accuracy. More details about Aggregated Variance method can be found in [36].

4.3. Design of data preprocessing algorithm based on EMD and FGn

In the proposed data preprocessing algorithm, fGn model serves as the fundament in order to alleviate the impact of colored noises. As the core of fGn, the Hurst parameter H is estimated utilizing Aggregated Variance method in this paper. For the sake of simplicity, the yaw rate is selected as an example to show the estimation of H because other inertial data have the similar process. For a given yaw rate $\omega_z(i)$, $i = 1,2,\ldots N$, we can estimate H according to the following five steps.

Step 1: Set the block size to $s = 1$.

Step 2: Calculate the standard deviation σ_s of the N data points and record the point $(s, s \cdot \sigma_s)$. The standard deviation is calculated as

$$\sigma_s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\omega_z(i) - \frac{1}{N}\sum_{i=1}^{N}\omega_z(i)\right)^2} \tag{9}$$

Step 3: Calculate the mean of neighboring data points and refresh the original dataset as

$$\omega_z(j) \leftarrow \frac{1}{2}(\omega_z(2j-1) + \omega_z(2j)) \tag{10}$$

At the same time, N and s are changed as

$$N \leftarrow \frac{N}{2},\ \mathrm{s} \leftarrow 2\mathrm{s} \tag{11}$$

where $j = 1, 2,\ldots, N/2$.

Step 4: If N is less than 64, jump to *step 5*. Otherwise, jump to *step 2*.

Step 5: Perform a linear regression of the record points on the log-log graph

$$\log_{10}(s\cdot\sigma_s) = H\log_{10}(s) + C \tag{12}$$

and the calculated slope is the estimation of H. The $\log_{10}(\cdot)$ refers to the base-10 logarithm function and C is a constant.

The yaw rate during a real road test and the estimation of Hurst parameter H are illustrated in Fig. 2. The yaw rate data is obtained from a MEMS IMU mounted inside a test vehicle. The estimation result of H is 0.933, i.e., the slope of fitted line, which indicates that the noise in gyro is distinct from white noise. This confirms that the yaw rate is contaminated by the colored noises such as the vibratory noises caused by the engine or road conditions and the slowly varying bias drifts.

Because the distribution of colored noise varies with the EMD decomposition level, it is not suitable if the threshold de-noising utilizes a fixed threshold. To eliminate the vibratory noises and part of slowly varying bias drifts in inertial signal, the paper adopts the level-dependent threshold based on the variance relation among the IMFs of fGn.

After implementing the EMD on inertial signal, the original signal is decomposed into a number, l, of IMFs, named as $d_M(i)$ ($M = 1, 2,\ldots, l$, $i = 1, 2,\ldots N$), and residue r_l. According to the theory of EMD, several

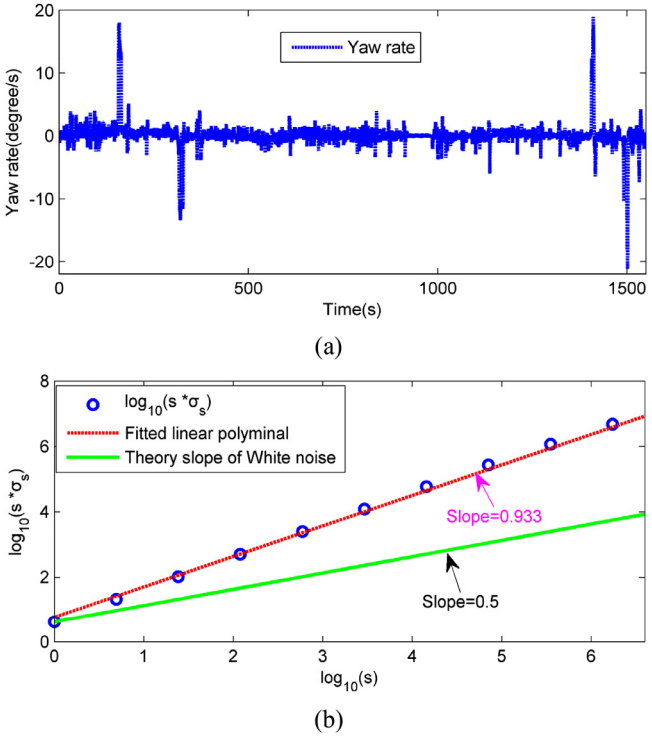

Fig. 2. Yaw rate and estimation of H. (a) Yaw rate; (b) estimation of H.

low-order IMFs capture mostly noises. Therefore, the variances of IMF1 and IMF2 can be estimated by employing a robust estimator

$$V_M = \left[\frac{median(|d_M(i)|)}{0.6745}\right]^2,\ M = 1, 2 \tag{13}$$

where V_M is the variance of Mth IMF.

As the IMF order increases, there is more useful information retained in the IMF. Thus, it is not valid to adopt (13) to estimate the noise variances for high order IMFs. To solve the problem, an alternative way is introduced based on the variance relation among the IMFs of fGn.

After obtaining the decomposition of fGn, the variance relation among all the IMFs can be expressed as [35]

$$V_{M'} = \rho_H^{(2H-2)(M'-M)}V_M,\ \ M' > M \geq 2 \tag{14}$$

where $V_{M'}$ is the variance of M'th IMF, $\rho_H \approx 2.01 + 0.2(H-0.5) + 0.12(H-0.5)^2$. Because the errors of inertial sensors are modeled using fGn, (14) is also valid for high order $d_M(i)$ obtained by decomposing inertial signal. Thus, it can be derived from (14) that

$$V_M = \rho_H^{(2H-2)(M-2)}V_2,\ \ M > 2 \tag{15}$$

Then, the level-dependent thresholds Th_M of proposed de-noising method are determined by

$$Th_M = \sqrt{V_M\cdot 2\ln N} = \begin{cases}\sqrt{V_M\cdot 2\ln N}, & M = 1, 2\\ \rho_H^{(H-1)(M-2)}\sqrt{V_2\cdot 2\ln N}, & M > 2\end{cases} \tag{16}$$

After obtaining the thresholds, we can adopt shrinkage scheme to perform thresholding on the IMFs to remove the noise components. Soft-thresholding is the most widely used shrinkage scheme for threshold filtering because hard-thresholding may cause discontinuity in the de-noised signal. However, soft-thresholding scheme produces biased outputs, which may introduce additional errors to the de-noised signal in certain conditions. To overcome abovementioned disadvantages, a new thresholding technique is proposed as

$$\tilde{d}_M(i) = \begin{cases}(d_M(i) - Th_M) - \left[\frac{2}{1+\exp\left(\frac{d_M(i)-Th_M}{Th_M}\right)} - 1\right]Th_M, & d_M(i) \geq Th_M\\ 0, & |d_M(i)| < Th_M\\ (d_M(i) + Th_M) - \left[\frac{2}{1+\exp\left(\frac{d_M(i)+Th_M}{Th_M}\right)} - 1\right]Th_M, & d_M(i) \leq -Th_M\end{cases} \tag{17}$$

where $\tilde{d}_M(i)$ represents the thresholded IMF, $\exp(\cdot)$ is the natural exponential function. The proposed thresholding scheme is a more continuous approach which preserves the high amplitude points and has a smooth transition from noise point to important feature point. Thus, it can help keep the continuity and reduce the impact of a biased signal.

Although the proposed thresholding scheme possesses better performances, the direct application of it to the IMFs will trigger devastating consequences for the continuity of the de-noised inertial signal. This is due to the fact that in any interval of zero-crossings $\theta_M(j) = [\theta_M(j)\ \ \theta_M(j+1)]$ in the Mth IMF, the absolute amplitude of $d_M(i)$ ($i = 1,2,\ldots N$) will be smaller than any non-zero threshold in the proximity of the zero-crossings $\theta_M(j)$ and $\theta_M(j+1)$. To deal with this problem, a novel interval thresholding scheme is designed based on the consideration that the signal between two neighboring zero-crossings are treated as the basic unit of analysis. For the two adjacent zero-crossings $\theta_M(j)$ and $\theta_M(j+1)$ within $d_M(i)$, the developed interval thresholding is defined in (18). In (18), $d_M((E_M(j)))$ represents the signal extremum of the corresponding zero-crossing interval, $d_M(\boldsymbol{\theta}_M(j))$ denotes all data points from zero-crossing $\theta_M(j)$ to $\theta_M(j+1)$.

In order to provide a real-time solution, the sliding window

technique is adopted. The proposed EITF algorithm is applied to process the data window. After de-noising the data of current window, the data window moves forward for the next processing.

$$\tilde{d}_M(\theta_M(j)) = \begin{cases} (d_M(\theta_M(j)) - Th_M) - \left[\frac{2}{1+\exp\left(\frac{d_M(\theta_M(j)) - Th_M}{Th_M}\right)} - 1 \right] Th_M, & d_M((E_M(j))) \geq Th_M \\ 0, & |d_M((E_M(j)))| < Th_M \\ (d_M(\theta_M(j)) + Th_M) - \left[\frac{2}{1+\exp\left(\frac{d_M(\theta_M(j)) + Th_M}{Th_M}\right)} - 1 \right] Th_M, & d_M((E_M(j))) \leq -Th_M \end{cases} \tag{18}$$

Generally, the improved EMD interval threshold filter comprises the following six steps

Step 1: Estimate Hurst parameter H of the inertial data of the current window utilizing Aggregated Variance method.
Step 2: Perform EMD for inertial data of current data window to obtain IMFs and residue.
Step 3: Calculate level-dependent thresholds using (16).
Step 4: Implement interval thresholding on IMFs according to (18).
Step 5: Reconstruct the de-noised inertial signal of current data window by the combination of thresholded IMFs and residue.
Step 6: The data window moves forward and the process jumps to *Step 1*. The algorithm terminates when the data window moves to the end of raw data.

Compared with the traditional and widely-used wavelet filter, the improved EMD interval threshold filter above can achieve better performance due to the following two advantages: (a) EMD is complete self-adaptive signal processing tool, whereas the wavelet decomposition is non-adaptive in essence; (2) The fGn model adopted in this method is suitable to not only white noises but also colored noises in inertial sensors, while the wavelet filter is based on white noise assumption and ignores the color of noises. Besides, due to that the basic functions adopted in EMD are derived from the original signal itself, the proposed EITF only introduces slight phase delay to the filtered signal, unlike the classical low-pass filter which brings about a notable phase delay.

5. Proposed LSSVM-NARX model

5.1. Brief review of NARX model

Due to its strength in accommodating the complex, dynamic, and nonlinear nature of real-time series applications, the NARX model is one of the most popular tools for time series modeling and predicting, and has been extensively adopted in practical applications [37,38]. The NARX model can be described as

$$y(t) = f\left[y(t-1), y(t-2), ..., y(t-q), u(t), u(t-1), ..., u(t-p)\right] + e(t) \tag{19}$$

where $u(t)$, $u(t\text{-}1)$,…, $u(t\text{-}p)$ represent the exogenous inputs of the model, $y(t)$ denotes the output of the model, $y(t\text{-}1)$, $y(t\text{-}2)$,…, $y(t\text{-}q)$ are the feedbacks of model output, p and q are the input-memory and output-memory orders, $e(t)$ is a zero-mean uncorrelated observation noise, $f(\cdot)$ is the nonlinear function which should be estimated. The time delay structure of the NARX model can be achieved by using the Tapped Delay Line (TDL), which is shown in Fig. 3. For an input variable $u(t)$, the p values $u(t\text{-}1)$, $u(t\text{-}2)$, …, $u(t\text{-}p)$ are simultaneously presented at the input of the model after processing by a p-order TDL.

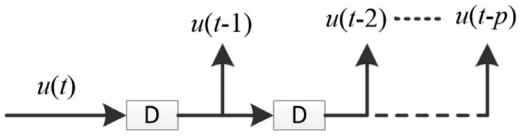

Fig. 3. The structure of TDL (order = p). D: time delay.

For the NARX model, the estimation of the nonlinear function $f(\cdot)$ is the main consideration. To identify the NARX model, the widely-used estimation method is based on neural network, as reported in [39]. However, this method may suffer from local minima problem, which will limit the generalization ability of the identified model. Due to its improved ability to avoid local minima, the LSSVM is suggested to identify the NARX model which is involved in highly non-linear black-box modeling applications. LSSVM is an enhancement of standard SVM, which is based on the structural risk minimization principle and presents superior performance when compared with neural network [40]. In this paper, it is adopted to identify the NARX model.

5.2. Design of the LSSVM-NARX model

As illustrated in Fig. 1, $\delta \boldsymbol{P}_E$ is the difference between positioning error $\delta \boldsymbol{P}_G$ provided by KF1, and positioning error $\delta \boldsymbol{P}_C$ offered by KF2.Therefore, the $\delta \boldsymbol{P}_G$ can be acquired when $\delta \boldsymbol{P}_E$ and $\delta \boldsymbol{P}_C$ are available. For the $\delta \boldsymbol{P}_C$, it is available all the time. So, if the positioning error $\delta \widehat{\boldsymbol{P}}_E$ can be predicted during GPS outages, the estimation of $\delta \widehat{\boldsymbol{P}}_G$ can be acquired subsequently. By correcting the INS position output employing the predicted $\delta \widehat{\boldsymbol{P}}_G$, a more accurate position solution can be obtained when GPS fails.

Aiming to accurately estimate the $\delta \widehat{\boldsymbol{P}}_E$ when GPS is completely unavailable, a NARX model is proposed to accommodate the potential factors affecting the INS position errors. The proposed NARX model for $\delta \boldsymbol{P}_E$ is expressed as

$$\delta \mathbf{P}_E(t) = f\left[\delta \mathbf{P}_E(t-1), \delta \mathbf{P}_E(t-2)..., \delta \mathbf{P}_E(t-q), \mathbf{u}(t), \mathbf{u}(t-1), \mathbf{u}(t-2)..., \mathbf{u}(t-p)\right] + \mathbf{e}(t) \tag{20}$$

where the exogenous input $\mathbf{u}(t)$ consists of three accelerations $\mathbf{F}_I(t)$ at time t, three angular rates $\boldsymbol{\omega}_I(t)$ at time t, and the time elapsing since GPS fails. The accelerations and angular rates are adopted to determine the vehicle running states, which poses a certain influence on INS position error [41]. For convenience, the feedbacks of model output, i.e. $\delta \mathbf{P}_E(t-1)$, $\delta \mathbf{P}_E(t-2)$…, and $\delta \mathbf{P}_E(t-q)$ are defined as inner inputs. Moreover, the exogenous inputs $\mathbf{u}(t)$, $\mathbf{u}(t\text{-}1)$, and $\mathbf{u}(t\text{-}p)$ are combined with the inner inputs to form the real inputs $\mathbf{g}(t)$ of the designed NARX model. The real inputs $\mathbf{g}(t)$ is given in (21). In this paper, three separate NARX models are designed simultaneously to model the $\delta \boldsymbol{P}_E$ in the east, north, and up components.

$$\mathbf{g}(t) = \left[\delta \mathbf{P}_E(t-1), \delta \mathbf{P}_E(t-2)..., \delta \mathbf{P}_E(t-q), \mathbf{u}(t), \mathbf{u}(t-1), ..., \mathbf{u}(t-p)\right] \tag{21}$$

A proper order of input/output memories can be determined through trials. As the order increases, more accurate model may be achieved, but it can also increase the computation complexity. Through successive tests, we can determine proper p and q to achieve a balance between the efficiency and accuracy. According to this idea, via sufficient trials, it is appropriate when p and q both take the value of 5 in this paper. The short-memory structure is realized by applying TDL with the order of 5, as explained in Fig. 3. After selecting the memory orders, the LSSVM is employed to identify the NARX model for $\delta \boldsymbol{P}_E$. First, the proposed NARX model in (20) is rewritten as follow

$$\delta \mathbf{P}_E(t) = \boldsymbol{\eta}^T \varphi\left[\mathbf{g}(t)\right] + \mathbf{b} + \mathbf{e}(t) \tag{22}$$

where $\boldsymbol{\eta}$ is an unknown parameter vector, $\varphi(\cdot)$ is a nonlinear feature map, $\mathbf{b}$ is a bias term.

Then, let $\{\mathbf{g}(i), \delta\mathbf{P}_E(i)\}(i = 1, 2,...,N)$ be a set of input/ouput data for the designed NARX model when the GPS works well. According to the LSSVM theory, the regularized cost function is proposed as

$$\begin{aligned}&\mathbf{min}L(\boldsymbol{\eta}, \mathbf{e}(\mathrm{i})) = \tfrac{1}{2}\boldsymbol{\eta}^T\boldsymbol{\eta} + \tfrac{r}{2}\textstyle\sum_{i=1}^{N}[\mathbf{e}(i)]^2\\ &\text{subject to:}\delta\mathbf{P}_E(i) = \boldsymbol{\eta}^T\varphi[\mathbf{g}(i)] + \mathbf{b} + \mathbf{e}(i),\ \ i = 1, 2,...,N\end{aligned} \tag{23}$$

where $r > 0$ is the regularization constant. For the purpose of solving the constrained optimization problem in (23), a Lagrangian is constructed as

$$\begin{aligned}Q(\boldsymbol{\eta}, \mathbf{b}, \mathbf{e}(i); \boldsymbol{\alpha}) &= L(\boldsymbol{\eta}, \mathbf{e}(i)) - \textstyle\sum_{i=1}^{N}\boldsymbol{\alpha}_i\{\boldsymbol{\eta}^T\varphi[\mathbf{g}(i)] + \mathbf{b} + \mathbf{e}(i) - \delta\mathbf{P}_E(i)\}\\ &=\tfrac{1}{2}\boldsymbol{\eta}^T\boldsymbol{\eta} + \tfrac{r}{2}\textstyle\sum_{i=1}^{N}[\mathbf{e}(i)]^2\\ &\quad - \textstyle\sum_{i=1}^{N}\boldsymbol{\alpha}_i\{\boldsymbol{\eta}^T\varphi[\mathbf{g}(i)] + \mathbf{b} + \mathbf{e}(i) - \delta\mathbf{P}_E(i)\}\end{aligned} \tag{24}$$

where $\boldsymbol{\alpha} = [\boldsymbol{\alpha}(1)\ \ \boldsymbol{\alpha}(2)\ \cdots\ \boldsymbol{\alpha}(N)]^T$ refers to Lagrangian multiplier vector. On the basis of the first order conditions of optimal theory, we can obtain the following equations

$$\frac{\partial Q}{\partial\boldsymbol{\omega}} = 0 \rightarrow \boldsymbol{\eta} = \sum_{i=1}^{N}\boldsymbol{\alpha}(i)\varphi[\mathbf{g}(i)] \tag{25}$$

$$\frac{\partial Q}{\partial\mathbf{b}} = 0 \rightarrow \sum_{i=1}^{N}\boldsymbol{\alpha}(i) = 0 \tag{26}$$

$$\frac{\partial Q}{\partial\mathbf{e}(i)} = 0 \rightarrow \boldsymbol{\alpha}(i) = r\mathbf{e}(i),\ \ i = 1, 2, ...,N \tag{27}$$

$$\frac{\partial Q}{\partial\boldsymbol{\alpha}(i)} = 0 \rightarrow \delta\mathbf{P}_E(i) = \boldsymbol{\eta}^T\varphi[\mathbf{g}(i)] + \mathbf{b} + \mathbf{e}(i),\ \ i = 1, 2, ...,N \tag{28}$$

Substituting (25)–(27) into (28) and eliminating $\boldsymbol{\eta}$ and $\mathbf{e}(i)$ gives

$$\begin{bmatrix}0 & \boldsymbol{\Delta}^T\\ \boldsymbol{\Delta} & \boldsymbol{\omega} + r^{-1}I\end{bmatrix}\begin{bmatrix}\mathbf{b}\\ \boldsymbol{\alpha}\end{bmatrix} = \begin{bmatrix}0\\ \delta\mathbf{P}_E\end{bmatrix} \tag{29}$$

where $\delta\mathbf{P}_E = [\delta\mathbf{P}_E(1)\ \ \delta\mathbf{P}_E(2)\ \cdots\ \delta\mathbf{P}_E(N)]^T$, $\boldsymbol{\Delta}$ is a N-dimension vector $[1\ 1\ \cdots\ 1]^T$, $\boldsymbol{\omega}$ is a $N \times N$ kernel matrix whose element $\omega_{m,n} = \varphi^T[\mathbf{g}(m)]\varphi[\mathbf{g}(n)] = K_{er}[\mathbf{g}(m), \mathbf{g}(n)]$ $(m, n = 1,2,..., N)$, I is the identity matrix with order of N. In this paper, the kernel matrix $\boldsymbol{\omega}$ is constructed by employing the Radial Basis Function (RBF) kernel function $K_{er}[\mathbf{g}(m), \mathbf{g}(n)] = \exp\left(-\frac{\|\mathbf{g}(m)-\mathbf{g}(n)\|^2}{\varsigma^2}\right)$, which is the most popular kernel function and has shown better performance than other kernels such as Linear and Polynomial [42]. The ς denotes the bandwidth of the kernel. By solving the linear (29) for $\boldsymbol{\alpha}$ and $\mathbf{b}$, we can obtain the estimation of nonlinear function $f(\cdot)$ in (20) as

$$f[\mathbf{g}(t)] = \sum_{i=1}^{N}\boldsymbol{\alpha}(i)K_{er}[\mathbf{g}(t), \mathbf{g}(i)] + \mathbf{b} \tag{30}$$

Combining (20) and (30), the final LSSVM-NARX model for $\delta\boldsymbol{P}_E$ is derived as

$$\delta\mathbf{P}_E(t) = \sum_{i=1}^{N}\boldsymbol{\alpha}(i)K_{er}[\mathbf{g}(t), \mathbf{g}(i)] + \mathbf{b} + \mathbf{e}(t) \tag{31}$$

It should be noted that the choice of proper regularization constant r and RBF kernel bandwidth ς is crucial for the performance of the proposed LSSVM-NARX model. In this paper, the 5-fold cross-validation is utilized for selecting these parameters.

6. Experimental results

6.1. Test setup

The positioning performance of the proposed methodology is verified with several real road tests on a Buick Sail SRV vehicle. The test vehicle was equipped with a low-cost NovAtel Superstar II receiver, MEMSIC MEMS-based IMU-440 inertial sensors as well as a digital compass (KVH, Middletown, CT, USA). All sensor data were recorded during the experiments, and then the logged data are utilized to evaluate the positioning solutions. Considering the computational burden of the proposed method, to simulate and achieve real-time execution, the rate of INS strap-down calculation is set as 20 Hz whereas the LSSVM-NARX/KF module in Fig. 1 is 1 Hz. In indirect integration mode as shown in Fig. 1, the overall system provides navigation solution at a rate of 20 Hz, which is the same as that of INS. The measurement noise accuracies (1σ) for the GPS velocity and position are set as 0.05 m/s and 3 m, respectively.

A series of typical road tests were carried out according to the above-described setup. Test 1 was carried out in an open suburban area of Nanjing City and the number of accessible satellites was above 5. The length of trajectory is about 23 km and it took about 25 min to finish one-round test. Some representative driving maneuvers, such as stops, turns, and sudden accelerations/decelerations were carried out. Eight simulated 45-s GPS outages were artificially introduced within the trajectories in order to evaluate the performance of the proposed method. Several typical driving conditions, including straight roads, curved roads and intersections, were considered to select the GPS outages, as illustrated in Fig. 4. An accurate differential GPS (DGPS) NovAtel L1L2/RT2 with centimeter-level accuracy was adopted as a reference for quantitative comparison.

Test 2 was carried out on the Fourth Ring Road in Beijing, which was a typical urban scenario with real GPS-denied environments in some parts. The whole test trajectory is about 65 km and it took about 1 h to complete one-round test. A series of typical driving maneuvers such as acceleration or deceleration, straight-driving, curved-driving, lane change and turn etc., were intentionally executed during the test. Eight representative 45 s GPS outages were chosen to execute the evaluation, as shown in Fig. 5. When the real outage time was less than 45 s, it was extended to 45 s. An accurate and reliable NovAtel SPAN-CPT system was used as the reference.

Test 3 was executed on the Fifth Ring Road in Beijing, which also

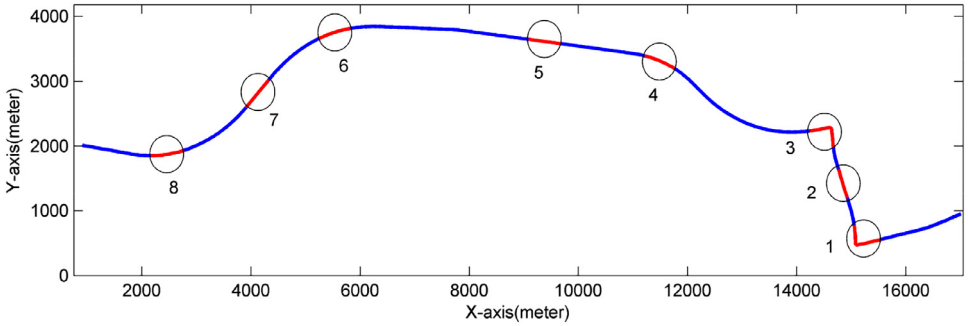

Fig. 4. Test trajectory 1(in blue) and the simulated GPS outages (in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

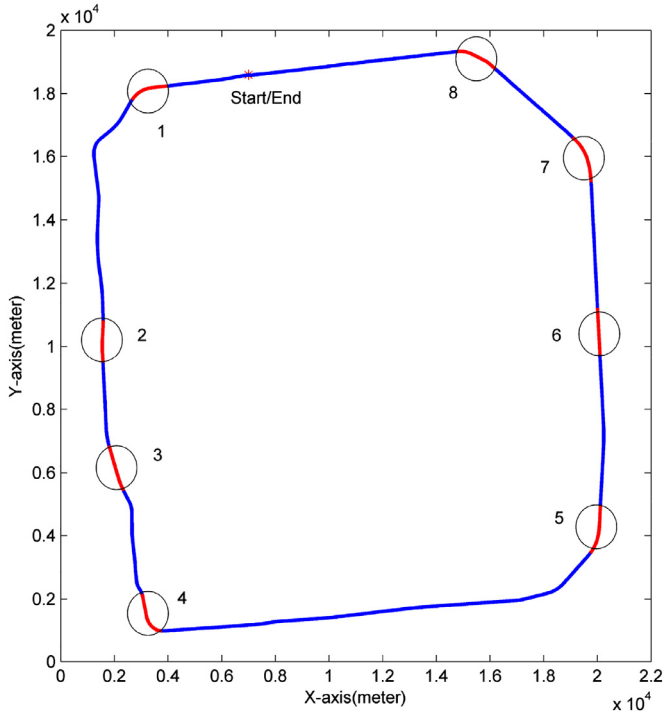

Fig. 5. Test trajectory 2(in blue) and the simulated GPS outages (in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

has the features in urban scenario. During the test, in some regions there are high buildings on both sides and in some other regions there are tunnels, which will cause GPS outages. Moreover, the vehicle was running at a low speed (less than 50 km/h or even stop) in some regions during the test. For these regions, six representative 45 s GPS outages were chosen to execute the evaluation, as shown in Fig. 6. When the real outage time was less than 45 s, it was extended to 45 s.

In Figs. 4–6, the solid blue line denotes the field test trajectory and black circles represent the simulated GPS outages.

6.2. Performance evaluation of data preprocessing algorithm

To verify the effectiveness of the proposed data preprocessing algorithm, other representative filtering algorithms including the wavelet filter, the EMD de-noising algorithm based on partial reconstruction (EMD-PR) and the EMD de-noising algorithm based on soft thresholding (EMD-ST) are also investigated for comparison. The wavelet filter is widely used in the accuracy enhancement of inertial sensors, while EMD-PR and EMD-ST are two typical de-noising filters based on EMD. In this paper, the yaw rate shown in Fig. 2(a) is chosen as an example to demonstrate the comparison since the results of other inertial data are consistent with those of yaw rate. Fig. 7 displays the performance of filtering based on wavelet filter, EMD-PR, EMD-ST and the proposed EITF. Fig. 7(b)–(d) are the zoomed-in figures of the sections marked by green squares in Fig. 7(a). To implement the wavelet filter in this paper, wavelet function “db4” with soft thresholding scheme based on Stein's Unbiased Risk Estimate (SURE) is adopted, because these are typical parameters employed to pre-process inertial sensors [17]. The EMD-ST employs the same threshold as the wavelet filter does.

From Fig. 7(b)–(d), it can be found that EMD-PR provides the worst results. The reason is that EMD-PR discards the first several irrelevant IMFs which contain some vehicle motion information. Compared with EMD-PR, the wavelet filter yields better performance. However, it produces biased result and distorts the high-amplitude section of the original yaw rate, which is due to both the non-adaptive feature of wavelet decomposition in nature and the soft thresholding schemes it adopts. In contrast with the wavelet filter, EMD-ST can suppress the high-frequency white noises effectively and does not apparently distort the original, which offers more preferable de-noising performance than the wavelet filter. This improvement can be attributed to that the self-adaptive EMD is more effective than the wavelet decomposition to process the nonlinear and non-stationary inertial signal. Nevertheless, it has limited capacity to tackle the colored noises such as the vibratory noises caused by the engine or road conditions and slowly varying bias drift noises, due to the fact that the threshold is selected based on white noise assumption, which ignores the color of noises. Among four filtering algorithms, the proposed EITF algorithm achieves the best performance. It can effectively suppress the vibratory noises, partial bias drift noises and high-frequency white noises. This can be attributed to that the EITF combines the advantages of EMD and colored noise model based on fGn.

6.3. Evaluation of the positioning performance

To evaluate the positioning performance, the following four representative methods are investigated for comparison.

1) *KF method* [24]. In this method, the information of INS and compass are fused to provide position solution during GPS outages. There is not any compensation for INS position errors. Besides, it does not adopt the proposed data preprocessing algorithm to de-noise the inertial data. The KF method is the most popular method for land vehicle positioning.
2) *LSSVM-NARX/KF method*. In this method, the LSSVM-NARX/KF is

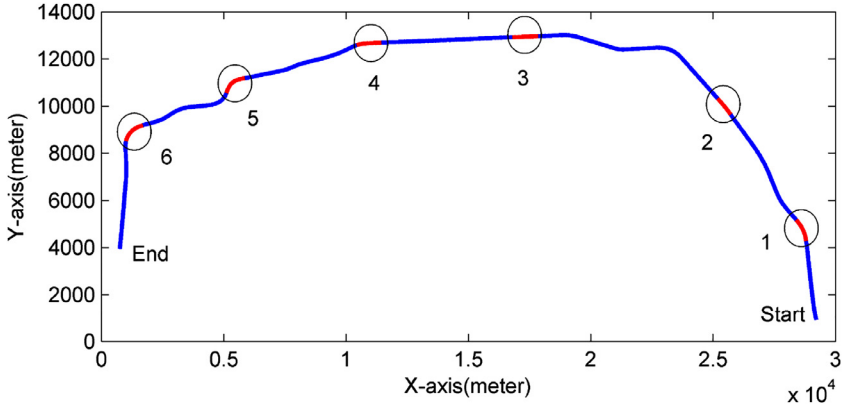

Fig. 6. Test trajectory 3(in blue) and the GPS outages (in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

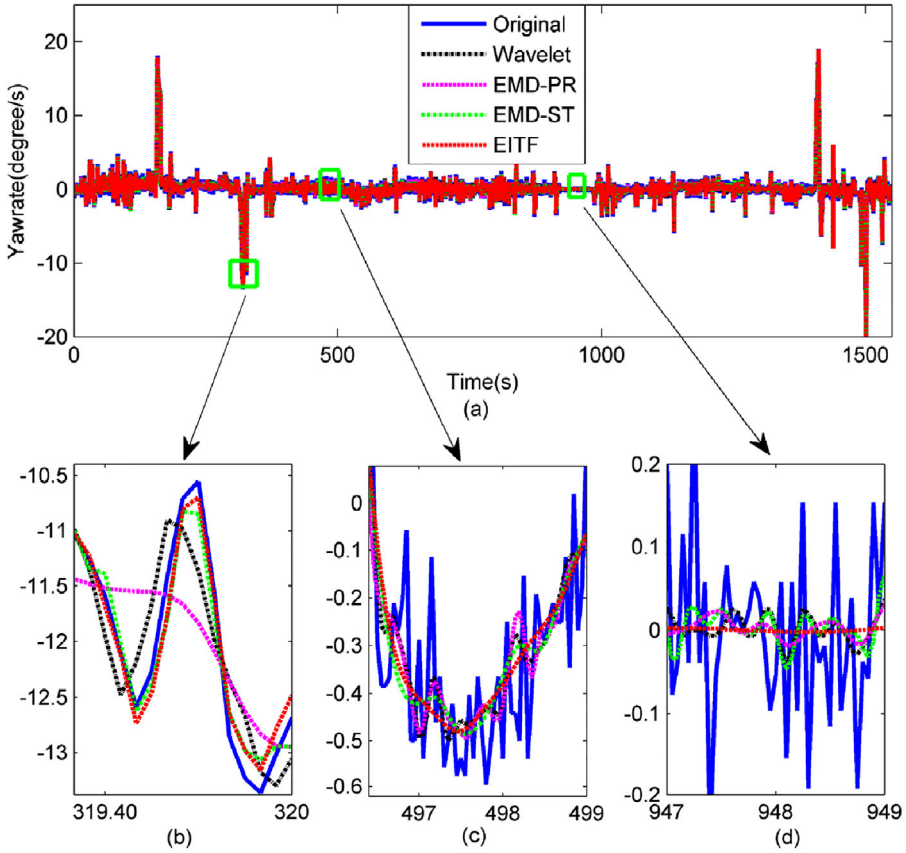

Fig. 7. Comparison of EITF and other filtering algorithms. (b), (c), and (d) are zoom-in parts of (a). (For interpretation of the references to color in the text, the reader is referred to the web version of this article.).

used to predict and compensate for the INS position errors when GPS signal is blocked. The proposed data preprocessing algorithm is not adopted to de-noise the raw inertial data.

3) *Wavelet-LSSVM-NARX/KF method*. Compared with LSSVM-NARX/KF method, the improvement in Wavelet-LSSVM-NARX/KF method is that the wavelet filter is used to de-noise the measurements of MEMS IMU prior to the data integration.
4) *EITF-LSSVM-NARX/KF method*. This is the method proposed in this paper. It utilizes the EITF to eliminate the noises in MEMS IMU before the data integration occurs, and employs LSSVM-NARX/KF to predict and compensate for the INS position errors When GPS signal is unavailable.

(1) Results of Test Trajectory 1

For four positioning methods, i.e. KF, LSSVM-NARX/KF, Wavelet-LSSVM-NARX/KF, and EITF-LSSVM-NARX/KF, Table 1 gives their statistics of the horizontal position errors during the 8 simulated GPS outages for test trajectory 1. From Table 1, it can be found that positioning performances of the four methods are significantly different during GPS outages.

Clearly, the KF method provides the worst results among the four methods during GPS outages. In all regions, the maximum (Max) value as well as root-mean-square (RMS) value of the position errors of KF method are the largest, and it cannot satisfy the requirements of land vehicle positioning. This is due to the fact that the accuracy of raw measurements of low-cost MEMS IMU is poor, and the KF is only updated based on heading angle, without considering the compensation for INS position errors when GPS signal is unavailable.

In comparison with KF method, the LSSVM-NARX/KF method accomplishes remarkable improvement in positioning performance. Over

Table 1
Statistics of the horizontal position errors for test trajectory 1 during GPS outages (Unit: m).

Outage Region	KF		LSSVM-NARX/KF		Wavelet-LSSVM-NARX/KF		EITF-LSSVM-NARX/KF	
	Max value	RMS value	Max value	RMS value	Max value	RMS value	Max value	RMS value
Out.1	71.16	36.17	31.53	18.23	21.32	11.76	17.25	9.48
Out. 2	97.31	46.10	54.52	30.75	27.28	18.01	18.26	8.34
Out. 3	99.43	43.21	45.14	24.58	20.36	10.17	13.73	7.08
Out. 4	132.22	59.04	41.25	22.02	32.13	18.74	25.36	15.21
Out. 5	37.56	18.61	21.24	9.81	16.36	9.18	13.71	6.18
Out. 6	73.82	35.14	35.73	21.01	22.85	15.70	17.59	10.89
Out. 7	53.60	23.66	23.28	10.20	15.71	6.75	13.38	6.25
Out. 8	77.02	33.78	28.51	9.83	17.81	6.92	7.11	4.70

the eight GPS outages, LSSVM-NARX/KF method can maintain the Max errors in the range of 21.24–54.52 m (with an RMS of 9.81–30.75 m), as opposed to the range of 37.56–132.22 m (with an RMS of 18.61–59.04 m) obtained by KF method. The reason is that the INS position errors are effectively compensated by the designed LSSVM-NARX, and thus positioning accuracy is significantly enhanced during GPS outages. Nevertheless, since the raw measurements of low-cost IMU are still directly used in this method, the positioning accuracy can be further improved to some extent.

Compared with LSSVM-NARX/KF method, the Wavelet-LSSVM-NARX/KF method can achieve better positioning performance in all GPS outages. For instance, in GPS outage 3, the maximum value of horizontal position errors is decreased to 20.36 m, i.e. about 55% accuracy improvement over LSSVM-NARX/KF method. The reason is that

the wavelet filter can eliminate the white noises in inertial signals to provide relatively higher quality information for LSSVM-NARX modeling.

Of the four methods, the EITF-LSSVM-NARX/KF method obtains the best positioning accuracy. As demonstrated in Table 1, this method is able to achieve an average enhancement in Max error by 27% and in RMS error by 30% over Wavelet-LSSVM-NARX/KF method. This can be attributed to that the proposed EITF can efficiently remove the vibratory noises and part of slowly varying bias drifts, in addition to the white noises in inertial signals. Therefore, more accurate LSSVM-NARX model is obtained due to the additional accuracy improvement in inertial signals.

To further illustrate the enhancements of the proposed method over the other three methods, three typical GPS outages, i.e. Outage 5, Outage 8 and Outage 1, are selected to show the trajectories output by four solutions, as demonstrate in Figs. 8 to 10.

Fig. 8 shows the positioning results for outage 5, during which the test vehicle moves along a straight line. Compared with KF (in cyan), the other three methods all obtain better position solution. For the LSSVM-NARX/KF method (in green), it achieves a 43% improvement in Max error against KF method, which validates the usefulness of adopting proposed LSSVM-NARX to compensate for the unbounded position errors of INS during the outages of the satellite signal. In addition, after eliminating the white noises in MEMS IMU measurements employing wavelet filtering algorithm, the Wavelet-LSSVM-NARX/KF method (in blue) can realize a 23% enhancement in Max error over the LSSVM-NARX/KF method. Furthermore, the EITF-LSSVM-NARX/KF method(in red) is able to achieve an additional accuracy improvement of about 16% over the Wavelet-LSSVM-NARX/KF method, which verifies the usefulness of the proposed EITF to suppress the colored noises such as the vibratory noises and partial bias drifts in MEMS IMU measurements.

Fig. 9 descripts the GPS outage 8, during which the test vehicle moves along a curve line. For the RMS error, the EITF-LSSVM-NARX/KF method accomplishes accuracy improvements of about 32%, 52% and 86% when compared with those of the Wavelet-LSSVM-NARX/KF method, LSSVM-NARX/KF, and KF method, respectively.

Fig. 10 shows the positioning results for GPS outage 1, which is a more challenging scenario because it consists of a sharp turn. The maximum horizontal position error for EITF-LSSVM-NARX/KF method is 17.25 m, whereas it is 21.32 m for EITF-LSSVM-NARX/KF method, 31.53 m for LSSVM-NARX/KF method, and 71.16 m for KF method. The advantage of LSSVM-NARX/KF method over KF method is again obvious from these results, as well as the advantage of EITF over the wavelet filter.

According to the results in Table 1 and Figs. 8–10, whether the trajectory of the vehicle is a straight line, a curve line or a sharp turn, the proposed multi-sensor fusion methodology is able to achieve continuous and reliable positioning solution.

(2) Results of Test Trajectory 2

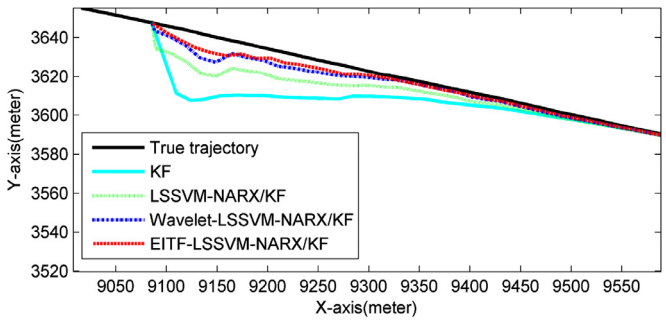

Fig. 8. Positioning results during outage 5 in trajectory 1. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.).

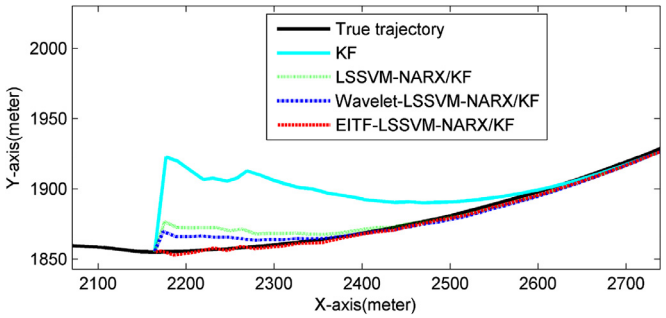

Fig. 9. Positioning results during outage 8 in trajectory 1.

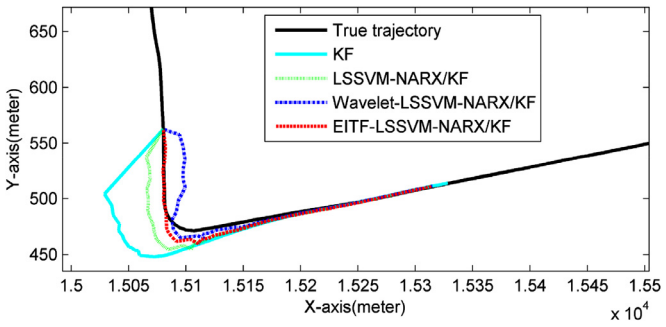

Fig. 10. Positioning results during GPS outage 1 in trajectory 1.

To further examine the positioning performance of the proposed methodology, test trajectory 2 with eight GPS outages in urban road was utilized for this purpose. Moreover, to compare with traditional methods, the radial basis function network (RBF) as utilized in much literature [11,27,28] is also designed. Here, the RBF/KF method denotes that the RBF network with similar inputs/outputs is developed to replace the LSSVM-NARX module in Fig. 1, but the EITF-based preprocessing algorithm is not utilized.

Table 2 shows the Max value and RMS value of the horizontal position errors during eight GPS outages in trajectory 2 for five methods, i.e., KF, LSSVM-NARX/KF, Wavelet-LSSVM-NARX/KF, EITF-LSSVM-NARX/KF and newly-added RBF/KF. Moreover, two representative outages, i.e., Outage 4 (curved) and Outage 6 (straight), were selected to show the positioning results output by different solutions, as demonstrated in Figs. 11 and 12 respectively. Note that for brevity and clarity, the results of only four methods, i.e., KF, RBF/KF, LSSVM-NARX/KF and EITF-LSSVM-NARX/KF, are shown in Figs. 11 and 12.

From Table 2 and Figs. 11–12, the proposed EITF-LSSVM-NARX/KF method demonstrates repeatable superior performance, as we have found in test trajectory 1. As far as LSSVM-NARX/KF and RBF/KF are concerned, the LSSVM-NARX/KF method exhibits better performance whether for the maximum error or for the RMS error, which verifies that the proposed LSSVM-NARX module has stronger nonlinear predicting ability and better generalization. For the proposed EITF-LSSVM-NARX/KF method, it achieves the best performance. As for the RMS position error, it obtains accuracy improvements of about 75%, 64%, 59% and 39% on average in comparison with those of KF, RBF/KF, LSSVM-NARX/KF and Wavelet-LSSVM-NARX/KF, respectively. The similar results can also be obtained from the maximum error.

(3) Results of Test Trajectory 3

In this section, the test trajectory 3 was utilized to further evaluate the positioning performance of the proposed methodology in GPS-denied environments. In addition, the Unscented Kalman Filter (UKF) method is also investigated for comparison. Here, the UKF method is similar to the KF method except that the KF algorithm is replaced with

Table 2
Statistics of the horizontal position errors for test trajectory 2 during GPS outages (Unit: m).

Outage Region	KF		RBF/KF		LSSVM-NARX/KF		Wavelet-LSSVM-NARX/KF		EITF-LSSVM-NARX/KF	
	Max value	RMS value	Max value	RMS value	Max value	RMS value	Max value	RMS value	Max value	RMS value
Out. 1	22.88	12.38	22.02	8.53	20.13	7.15	12.28	4.15	7.28	2.81
Out. 2	90.59	42.53	71.31	31.75	63.39	28.84	49.38	20.12	21.49	10.52
Out. 3	48.33	24.20	32.64	18.39	24.13	17.02	19.44	13.78	14.89	10.48
Out. 4	84.89	41.65	69.30	23.57	62.11	20.19	45.52	13.87	30.91	9.34
Out. 5	86.30	34.30	64.40	28.59	52.07	25.19	22.65	9.01	17.41	5.86
Out. 6	72.45	29.43	47.92	21.60	36.68	18.76	31.83	13.29	15.27	7.81
Out. 7	58.27	25.72	52.80	22.05	51.59	20.01	36.67	14.13	22.41	10.52
Out. 8	91.07	40.99	45.30	20.39	39.18	18.91	33.84	16.41	14.82	6.43

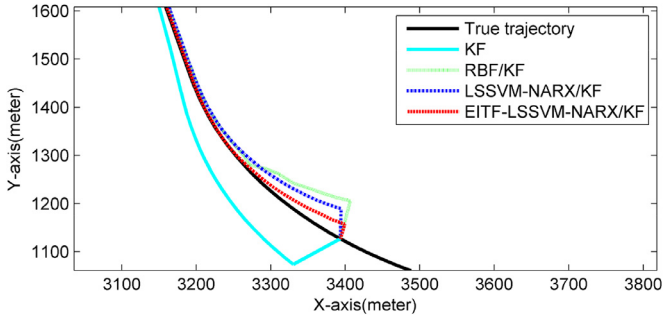

Fig. 11. Positioning results during GPS outage 4 in trajectory 2.

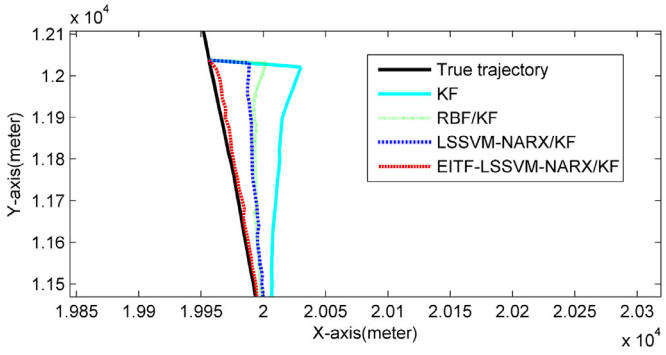

Fig. 12. Positioning results during GPS outage 6 in trajectory 2.

the UKF algorithm.

For brevity, Table 3 only shows the RMS value of the horizontal position errors during six GPS outages in trajectory 3 for five methods, i.e., KF, UKF, RBF/KF, LSSVM-NARX/KF and EITF-LSSVM-NARX/KF. For the Max values, we can obtain the similar conclusion. For GPS outages 1 and 4, the vehicle ran at low speed, and they were utilized to simulate the low speed scenarios. Besides, as a representative one, Outage 6 (curved, and large speed change during it) was selected to show the positioning results, as shown in Fig. 13.

As can be seen from Table 3 and Fig. 13, the proposed EITF-LSSVM-NARX/KF method remains achieves the best performance, as demonstrated in test trajectories 1 and 2. As far as KF and UKF are concerned, the KF method generally has better performance than the UKF method in most GPS outages. For instance, in GPS outage 3, the RMS error for KF method is 24.72 m, while it is 26.49 m for UKF method.

Moreover, from the results above, the performance of EITF-LSSVM-NARX/KF method in Outage 5 is slightly better than that of KF method, unlike the remarkable improvements in other GPS outages. It reveals that the improvement for the proposed method may be limited in the situation that the vehicle dynamics in the prediction mode due to the GPS failure is significantly different from that in the update mode, as Outage 5 was in.

Table 3
Statistics of the horizontal position errors for test trajectory 3 during GPS outages (Unit: m).

Outage Region	KF	UKF	RBF/KF	LSSVM-NARX/KF	EITF-LSSVM-NARX/KF
	RMS value	RMS value	RMS value	RMS value	RMS value
Out. 1	45.66	50.89	25.93	23.40	10.81
Out. 2	27.15	28.01	17.16	16.11	5.12
Out. 3	24.72	26.49	14.21	13.12	8.48
Out. 4	31.23	33.52	18.39	16.92	13.58
Out. 5	52.49	57.87	55.16	54.23	51.07
Out. 6	71.15	70.02	32.76	25.83	11.96

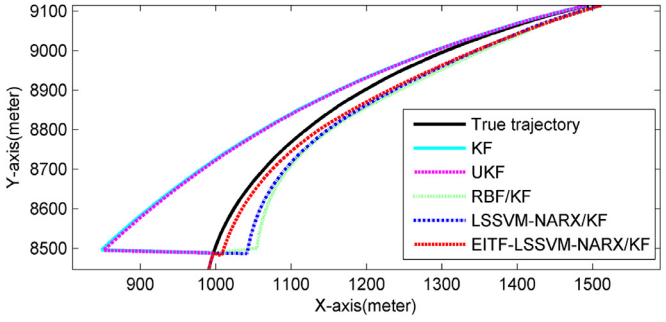

Fig. 13. Positioning results during GPS outage 6 in trajectory 3.

7. Conclusions

To provide a reliable and accurate solution for land vehicle positioning even under GPS complete unavailability, this paper proposes an enhanced multi-sensor fusion methodology.

In the proposed methodology, both the key data preprocessing algorithm and the novel hybrid strategy have been designed to improve the overall system performance. First, the data preprocessing algorithm based on EITF is developed to effectively suppress the vibratory noises caused by the engine or road conditions, partial slowly-varying bias drifts, as well as the high-frequency white noises in MEMS IMU raw measurements before the data fusion occurs, which can enhance the model accuracy of subsequent LSSVM-NARX model. Then, the LSSVM-NARX/KF hybrid strategy is developed to predict and compensate for the INS position errors in case of GPS outages.

Typical road experiments have been performed to assess the positioning performance of the proposed multi-sensor fusion methodology. The experimental results validate the effectiveness of the proposed data preprocessing and the designed LSSVM-NARX/KF hybrid strategy.

However, for the proposed methodology, there is still room for improvement in real-time implementation, better generalization capability and accuracy. Future work can focus on the following directions:

(1) Improving real-time capability. More efforts will be dedicated to optimize the operation of the proposed EITF to further improve its real-time capability.
(2) Enhancing generalization capability. More nonlinear factors, such as vehicle attitude and speed etc., will be considered in INS error modeling in order to further increase the generalization capability of the proposed method.
(3) Introducing more auxiliary sensors. The proposed methodology is an open structure, in which more information source (e.g., digital map) can be introduced to further enhance the accuracy of the positioning system.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 61273236), the Jiangsu Provincial Basic Research Program (Natural Science Foundation, Grant No. BK2010239) and the Doctoral Fund for Youth Teachers of Ministry of Education of China (No. 200802861061).

References

[1] N.R. Velaga, M.A. Quddus, A.L. Bristow, Y. Zheng, Map-aided integrity monitoring of a land vehicle navigation system, IEEE Trans. Intell. Transp. Syst. 13 (2) (2012) 848–858.
[2] I. Skog, P. Handel, In-car positioning and navigation technologies—a survey, IEEE Trans. Intell. Transp. Syst. 10 (1) (2009) 4–21.
[3] N. Salameh, G. Challita, S. Mousset, A. Bensrhair, S. Ramaswamy, Collaborative positioning and embedded multi-sensors fusion cooperation in advanced driver assistance system, Transp. Res. C: Emerg. Technol. 29 (2013) 197–213.
[4] J. Piao, M. Beecroft, M. McDonald, Vehicle positioning for improving road safety, Transp. Rev. 30 (6) (2010) 701–715.
[5] J. Zhang, F.Y. Wang, K. Wang, W.H. Lin, X. Xu, C. Chen, Data-driven intelligent transportation systems: a survey, IEEE Trans. Intell. Transp. Syst. 12 (4) (2011) 1624–1639.
[6] L. Heng, D.B. Work, G.X. Gao, GPS signal authentication from cooperative peers, IEEE Trans. Intell. Transp. Syst. 16 (4) (2015) 1794–1805.
[7] M.M. Atia, S. Liu, H. Nematallah, T.B. Karamat, A. Noureldin, Integrated indoor navigation system for ground vehicles with automatic 3-D alignment and position initialization, IEEE Trans. Veh. Technol. 65 (4) (2015) 1279–1292.
[8] G. Panahandeh, M. Jansson, Vision-aided inertial navigation based on ground plane feature detection, IEEE/ASME Trans. Mechatron. 19 (4) (2014) 1206–1215.
[9] K. Wang, Y.H. Liu, L. Li, A simple and parallel algorithm for real-time robot localization by fusing monocular vision and odometry/AHRS sensors, IEEE/ASME Trans. Mechatron. 19 (4) (2014) 1447–1457.
[10] J. Huang, H.S. Tan, A low-order DGPS-based vehicle positioning system under urban environment, IEEE/ASME Trans. Mechatron. 11 (5) (2006) 567–575.
[11] S. Adusumilli, D. Bhatt, H. Wang, P. Bhattacharya, V. Devabhaktuni, A low-cost INS/GPS integration methodology based on random forest regression, Expert Syst. Appl. 40 (11) (2013) 4653–4659.
[12] H. Hur, H.S. Ahn, Unknown input H observer-based localization of a mobile robot with sensor failure, IEEE/ASME Trans. Mechatron. 19 (6) (2014) 1830–1838.
[13] N.E. El Faouzi, H. Leung, A. Kurian, Data fusion in intelligent transportation systems: progress and challenges–a survey, Inf. Fusion 12 (1) (2011) 4–10.
[14] A. Noureldin, T.B. Karamat, M.D. Eberts, A. El-Shafie, Performance enhancement of MEMS-Based INS/GPS integration for low-cost navigation applications, IEEE Trans. Veh. Technol. 58 (3) (2009) 1077–1096.
[15] J. Li, N. Song, G. Yang, et al., Improving positioning accuracy of vehicular navigation system during GPS outages utilizing ensemble learning algorithm, Inf. Fusion 35 (2017) 1–10.
[16] S. Hong, C. Lee, F. Borrelli, J.K. Hedrick, A novel approach for vehicle inertial parameter identification using a dual Kalman filter, IEEE Trans. Intell. Transp. Syst. 16 (1) (2015) 51–161.
[17] A.G. Quinchia, G. Falco, E. Falletti, F. Dovis, C. Ferrer, A comparison between different error modeling of MEMS applied to GPS/INS integrated systems, Sensors 13 (8) (2013) 9549–9588.
[18] H. Sahli, N. El-Sheimy, A novel method to enhance pipeline trajectory determination using pipeline junctions, Sensors 16 (4) (2016) 567.
[19] T. Brunner, J.P. Lauffenburger, S. Changey, et al., Magnetometer-augmented IMU simulator: in-depth elaboration, Sensors 15 (3) (2015) 5293–5310.
[20] N. El-Sheimy, K.W. Chiang, A. Noureldin, The utilization of artificial neural networks for multisensor system integration in navigation and positioning instruments, IEEE Trans. Instrum. Meas. 55 (5) (2006) 1606–1615.
[21] B.H. Kaygısız, A.M. Erkmen, İ. Erkmen, Enhancing positioning accuracy of GPS/INS system during GPS outages utilizing artificial neural network, Neural Process. lett. 25 (3) (2007) 171–186.
[22] D. Bhatt, P. Aggarwal, V. Devabhaktuni, P. Bhattacharya, A new source difference artificial neural network for enhanced positioning accuracy, Meas. Sci. Technol. 23 (10) (2012).
[23] P. Aggarwal, D. Bhatt, V. Devabhaktuni, P. Bhattacharya, Dempster Shafer neural network algorithm for land vehicle navigation application, Inf. Sci. 253 (2013) 26–33.
[24] W. Abdel-Hamid, A. Noureldin, N. El-Sheimy, Adaptive fuzzy prediction of low-cost inertial-based positioning errors, IEEE Trans. Fuzzy. Syst. 15 (3) (2007) 519–529.
[25] A.M. Hasan, K. Samsudin, A.R. Ramli, Optimizing of ANFIS for estimating INS error during GPS outages, J. Chin. Inst. Eng. 34 (7) (2011) 967–982.
[26] A.M. Hasan, K. Samsudin, A.R. Ramli, R.S. Azmir, Automatic estimation of inertial navigation system errors for global positioning system outage recovery, Proc. IMechE. G: J. Aerosp. Eng. 225 (1) (2011) 86–96.
[27] X. Lei, J. Li, An adaptive navigation method for a small unmanned aerial rotorcraft under complex environment, Measurement 46 (10) (2013) 4166–4171.
[28] D. Bhatt, P. Aggarwal, V. Devabhaktuni, P. Bhattacharya, A novel hybrid fusion algorithm to bridge the period of GPS outages using low-cost INS, Expert Syst. Appl. 4 (5) (2014) 2166–2173.
[29] A. Noureldin, A. El-Shafie, M. Bayoumi, GPS/INS integration utilizing dynamic neural networks for vehicular navigation, Inf. Fusion 12 (1) (2011) 48–57.
[30] K. Saadeddin, M.F. Abdel-Hafez, M.A. Jaradat, M.A. Jarrah, Optimization of intelligent approach for low-cost INS/GPS navigation system, J. Intell. Robot. Syst. 73 (1) (2014) 325–348.
[31] Y. Lei, J. Lin, Z. He, M.J. Zuo, A review on empirical mode decomposition in fault diagnosis of rotating machinery, Mech. Syst. Signal. Process. 35 (1) (2013) 108–126.
[32] V. Maxim, L. Şendur, J. Fadili, J. Suckling, R. Gould, R. Howard, E. Bullmore, Fractional Gaussian noise, functional MRI and Alzheimer's disease, Neuroimage 25 (1) (2005) 141–158.
[33] A. Noureldin, T. Karamat, J. Georgy, Fundamentals of Inertial Navigation, Satellite-Based Positioning and Their Integration, Springer, New York, 2014.
[34] M.S. Tsai, C.L. Yen, H.T. Yau, Integration of an empirical mode decomposition algorithm with iterative learning control for high-precision machining, IEEE/ASME Trans. Mechatron. 18 (3) (2013) 878–886.
[35] P. Flandrin, G. Rilling, P. Goncalves, Empirical mode decomposition as a filter bank, IEEE Signal Process. Lett. 11 (2) (2004) 112–114.
[36] H.J. Block, On the Nature of the Stock Market: Simulation and Experiments, Phd thesis University of British Columbia, 2000.
[37] H.T. Pham, B.S. Yang, A hybrid of nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting, Expert Syst. Appl. 37 (4) (2010) 3310–3317.
[38] S. De Vito, P. Delli Veneri, E. Esposito, M. Salvato, V. Bright, R.L. Jones, O. Popoola, Dynamic multivariate regression for on-field calibration of high speed air quality chemical multi-sensor systems, 2015 AISEM Annual Conference, Trento, Italy, February, 2015, pp. 1–3.
[39] S.M. Ahmad, M.H. Shaheed, A.J. Chipperfield, M.O. Tokhi, Non-linear modelling of a one-degree-of-freedom twin-rotor multi-input multi-output system using radial basis function networks, Proc. IMechE. G: J. Aerosp. Eng. 216 (4) (2002) 197–208.
[40] M.M. Adankon, M. Cheriet, A. Biem, Semisupervised learning using Bayesian interpretation: application to LS-SVM, IEEE Trans. Neural Netw. 22 (4) (2011) 513–524.
[41] J. Wang, H.K. Lee, S. Hewitson, H.K. Lee, Influence of dynamic and trajectory on integrated GPS/INS navigation performance, Positioning 1 (5) (2002) 109–116.
[42] M. Aydogdu, M. Firat, Estimation of failure rate in water distribution network using fuzzy clustering and LS-SVM methods, Water Resour. Manag. 29 (5) (2015) 1575–1590.