Subtractive clustering Takagi-Sugeno position tracking for humans by low-cost inertial sensors and velocity classification

Dariusz Maton[1] [ORCID]**, John Theodoros Economou[1], David Galvão Wall[1], David Ward[2] and Simon Trythall[2]**

Abstract

In this work, open-loop position tracking using low-cost inertial measurement units is aided by Takagi-Sugeno velocity classification using the subtractive clustering algorithm to help generate the fuzzy rule base. Using the grid search approach, a suitable window of classified velocity vectors was obtained and then integrated to generate trajectory segments. Using publicly available experimental data, the reconstruction accuracy of the method is compared against four competitive pedestrian tracking algorithms. The comparison on selected test data, has demonstrated more competitive relative and absolute trajectory error metrics. The proposed method in this paper is also verified on an independent experimental data set. Unlike the methods which use deep learning, the proposed method has shown to be transparent (fuzzy rule base). Lastly, a sensitivity analysis of the velocity classification models to perturbations from the training orientation at test time is investigated, to guide developers of such data-driven algorithms on the granularity required in an ensemble modeling approach. The accuracy and transparency of the approach may positively influence applications requiring low-cost inertial position tracking such as augmented reality headsets for emergency responders.

Introduction

The ad hoc tracking of humans in global navigation satellite system (GNSS)-denied environments is an increasingly urgent requirement given over 55% of the world's population were reported to inhabit urban environments in 2018, places that are prone to GNSS signal fading and multipath effects.[1] In narrowband ranging for instance, the received signal in a severe multipath environment will be the sum of multiple carrier signals having different amplitudes and phases, thus distorting the range measurement.[2] For tracking emergency responders, the development of GNSS-independent systems has the potential to save the lives of casualties through enhanced operational efficiency, as well as the lives of responders themselves through improved situational awareness. A shift in reliance from GNSS is also valuable for maritime and aeronautical navigation due to advances in spoofing and disruption caused by personal privacy devices.[3,4]

While many positional tracking technologies have been proposed to tackle the problem including ultrawideband (UWB), radio frequency identification (RFID), wireless fidelity (Wi-Fi) and vision systems, most are not ad hoc methods, requiring careful deployment and configuration of transmitters and receivers into the tracking space a priori.[5–10] Although breadcrumbing, the opportunistic placement of RFID tags upon arrival at the site of an incident, was proposed by Renaudin et al.,[11] it has been reported sub-optimal due to the potential destruction of tags during operation and difficulties in retrieval afterward.[12] While visual odometry (VO) systems use cameras to track position using natural features and are infrastructure independent, they

[1]Centre for Defence Engineering, Cranfield University, Defence Academy of the United Kingdom, Shrivenham, Swindon, UK
[2]BAE Systems, Rochester, Kent, UK

Corresponding author:
Dariusz Maton, Cranfield University, Centre for Defence Engineering, Defence Academy of the United Kingdom, Shrivenham SN6 8LA, UK.
Email: d.maton@cranfield.ac.uk

suffer from sensitivity to lighting conditions and featureless spaces.[13]

Inertial sensors are self-contained sensors that involve creating an inertial element which remains fixed relative to the inertial reference frame, i, a frame fixed with respect to the distant stars. Inertial measurement units (IMUs) typically consist of three mutually orthogonal gyroscopes measuring the angular velocity of the sensor body frame, b, relative to the i-frame, expressed in b, denoted $\boldsymbol{\omega}_{ib}^{b}$. Similarly, accelerometers measure specific force, $\boldsymbol{f}_{ib}^{b}$. Low-cost, lightweight, small-sized micro-electromechanical systems (MEMS) IMUs are found in many modern electronic devices such as smartphones and virtual reality (VR) headsets. From literature, there are three approaches to low-cost IMU position tracking that trade generalization to different motions for accuracy of trajectory estimation.

Strapdown inertial navigation

In strapdown inertial navigation, position is estimated relative to a known initial condition by integrating measurements, a process of dead reckoning.[14] When navigating small distances on Earth using low-cost IMUs, the local geographic form of the navigation equations suffice assuming[15]:

(i)   The contribution from the Earth's rotation of $\sim 15° \mathrm{h}^{-1}$ is considerably less than the noise and bias stability characteristics of the gyros.
(ii) The transport rate of the navigation frame, n, pointing North, East, and down (NED), over the Earth's surface is negligible due to traveling small distances compared to its radius.
(iii) Coriolis acceleration contributions are negligible compared to the noise and stability characteristics of the accelerometers.

The navigation frame can thus be treated as inertial such that $\boldsymbol{\omega}_{ib}^{b} \approx \boldsymbol{\omega}_{nb}^{b}$ and $\boldsymbol{f}_{ib}^{b} \approx \boldsymbol{f}_{nb}^{b}$ resulting in the simplified acceleration term:

$$\boldsymbol{a}_{nb}^{n} = \mathbf{R}_{b}^{n}\boldsymbol{f}_{nb}^{b} + \boldsymbol{g}_{n} \qquad (1)$$

where $\boldsymbol{a}_{nb}^{n}$ is the acceleration of the sensor relative to the n-frame expressed in the n-frame, $\boldsymbol{f}_{nb}^{b}$ is assumed to be the measured specific force from the accelerometer, $\boldsymbol{g}_{n}$ is the sum of acceleration due to mass attraction and the centrifugal acceleration due to Earth's rotation expressed in the n-frame and $\mathbf{R}_{b}^{n}$ is the rotation matrix whose columns represent the cosines of the angles between the unit basis vectors of the b and n-frames. This matrix evolves with the gyro measurements to first order accuracy according to:

$$\mathbf{R}_{b,t}^{n} = \mathbf{R}_{b,t-1}^{n}\left(\mathbf{I} + \Delta t\, \mathrm{sk}(\boldsymbol{\omega}_{nb,t}^{b})\right) \qquad (2)$$

where Δt is the sample time and $\mathrm{sk}(\boldsymbol{\omega}_{nb,t}^{b})$ is the skew-symmetric matrix of $\boldsymbol{\omega}_{nb}^{b}$:

$$\mathrm{sk}(\boldsymbol{\omega}_{nb,t}^{b}) = \begin{bmatrix} 0 & -\omega_{z} & \omega_{y} \\ \omega_{z} & 0 & -\omega_{x} \\ -\omega_{y} & \omega_{x} & 0 \end{bmatrix} \qquad (3)$$

While rotation matrices have been used to parametrize orientation, quaternions are equally applicable and Euler angles may be preferred for maneuvers not exceeding 10° in attitude.[16]

Due to drift in the gyros and for initialization, measurements of vector fields induced by gravity and the geomagnetic field are used as cues in an Attitude and Heading Reference System (AHRS) to help converge and maintain accurate $\mathbf{R}_{b}^{n}$ estimates.[17,18] As noted by Woodman,[19] accurate $\mathbf{R}_{b}^{n}$ estimates are crucial in positioning as they are required to compensate for the $\boldsymbol{g}_{n}$ term in equation (1).

With $\mathbf{R}_{b}^{n}$ from the AHRS, $\boldsymbol{a}_{nb}^{n}$ may therefore be integrated by Euler discretization as in equations (4) and (5) assuming constant acceleration between epochs.

$$\boldsymbol{v}_{nb,t}^{n} = \boldsymbol{v}_{nb,t-1}^{n} + \Delta t\, \boldsymbol{a}_{nb,t}^{n} \qquad (4)$$

$$\boldsymbol{x}_{nb,t}^{n} = \boldsymbol{x}_{nb,t-1}^{n} + \boldsymbol{v}_{nb,t}^{n}\Delta t + \frac{1}{2}\, \boldsymbol{a}_{nb,t}^{n}\, \Delta t^{2} \qquad (5)$$

Although this approach works well for navigation and strategic grade IMUs, they have noise and bias stability characteristics several orders of magnitude better than commercial IMUs.[20] Thus, despite a well-tuned AHRS algorithm, errors accumulate in $\boldsymbol{a}_{nb}^{n}$ resulting in a positional error that grows cubically with time.[21]

Heuristic approaches

Methods of position tracking that augment the double integration of acceleration with heuristic assumptions are promising alternatives. Two such methods in human tracking include pedestrian dead reckoning (PDR)[22,23] and zero velocity updating (ZUPT).[21,24] Both take advantage of the periodicity of human gait to better estimate displacement between strides while an AHRS handles orientation. Although they improve on strapdown inertial navigation, bounding the positional error growth linearly with the number of steps, PDR suffers from reduced generality to other types of motion such as vehicular and unconventional motions such as side-stepping and jumping.[25] ZUPT restricts placement of the IMU to the foot, risking damage to the sensor.[26] Furthermore, it also mandates a time-consuming calibration of the IMU's gyro and accelerometers beforehand such as the method presented by Glueck et al.[27]

Data-driven approaches

Given the limitations of the previous methods, data-driven approaches have been explored to replace the numerical integration steps in strapdown inertial navigation with deep learning regression models.

Inertial Odometry Network (IONet),[13] and Robust Neural Inertial Navigation in the Wild (RoNIN)[28] used supervised machine learning, performing sequence-to-one regressions of windowed IMU data from smartphones to estimate more accurate trajectories. While RoNIN regressed velocity vectors in the x-y plane of the navigation frame before integration, IONet directly regressed polar displacement vectors from IMU data. In contrast to the sensor placement restrictions imposed by heuristic approaches, these aimed for competitive tracking accuracy regardless of sensor attachment on the human body. Despite presenting improvements in trajectory prediction over conventional strapdown and heuristic methods, Silva do Monte Lima et al.[29] showed that guaranteeing acceptable model performance across disparate data sets remains a problem. Furthermore, the number of parallel nonlinear steps in neural networks learnt by backpropagation makes them not only difficult to interpret but also vulnerable to adversarial examples.[30] They are thus challenging to certify and implement in safety critical applications.

In this work, a classification approach is applied to estimating velocity vectors from IMU data for a given sensor mode of carriage on the human body. The functionality of the approach and sensitivity to orientation errors are investigated using the Oxford Inertial Odometry Data Set (OxIOD)[31] and the data set from RoNIN.[32] The contributions of this paper are:

(i) A subtractive clustering (SC) Takagi-Sugeno (TS) fuzzy classification approach to mapping IMU data to velocity vectors before integration for trajectory estimation. Unlike other data-driven algorithms that use deeply learnt regression models, the proposed SC-TS method produces a model that is transparent as input-output relations can be visualized in surface plots and membership functions (encoded by a mean and standard deviation).

(ii) A novel optimization approach to data pre-processing showing how IMU and ground truth data (experimental data) can be aligned in heading for consistency between sequences during training and testing.

(iii) A sensitivity analysis by Monte Carlo simulation investigating the loss in accuracy of the SC-TS method with increasing disparity between the training and test modes of carriage. To the authors' best knowledge, this has not been reported elsewhere and paves the way to the creation of an ensemble tracking approach offering accurate tracking across a range of sensor-body attachments.
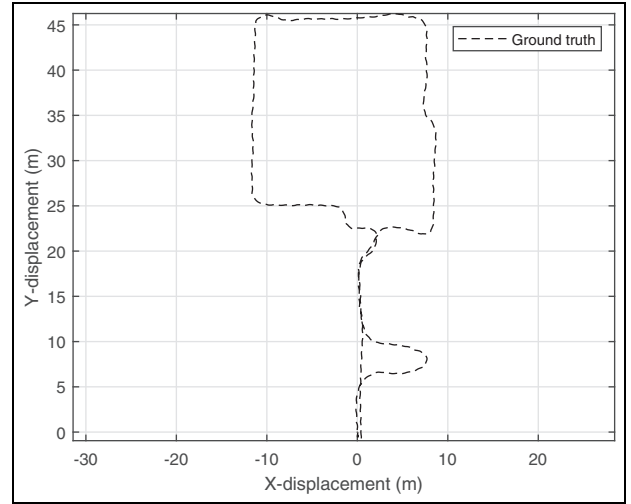

Figure 1. Floor one ground truth trajectory from "large scale" OxIOD data.

The remainder of the paper is structured as follows. Data pre-processing shows how the input and output data were obtained, including the optimization method which ensured their alignment in heading. The SC-TS classification model and its tuning are then outlined. The results show trajectories estimated by the proposed method compared with competitive algorithms, verification on an independent data set and results from the sensitivity analysis. A discussion of the results is presented followed by a conclusion and direction of future work.

Data sets and pre-processing

Data used to train the models and evaluate the approach were obtained from the OxIOD data set, containing 158 sequences of smartphone IMU and ground truth data for various devices and modes of carriage. The subset, "large scale," was used since it featured less repetitive trajectories through a building's corridors, more consistent with natural human motion. Figure 1 shows the ground truth trajectory for a floor one sequence generated by the visual-inertial odometry system.

Training between various sequences was most effective when input IMU data and output ground truth data were expressed in a common reference frame. The locally-level frame, l, was chosen which is like the navigation frame but free in heading. During training, this frame was defined at the start of the sequence by the ground truth system and IMU input data aligned to it such that both inputs and outputs were expressed in l_{true}. This was achieved by first obtaining the IMU's b-frame orientation relative to l_{imu}, $\mathbf{R}_{b_{imu}, k}^{l_{imu}}$, using Madgwick et al.'s Extended Complementary AHRS Filter.[33] AHRS parameters K_{normal}, K_{init} and t_{init} of 0.5, 10.0, and 3.0 s respectively were used for reliable initialization.[34] The magnetometer was not used due to the

potential for hard and soft iron bias interference in the indoor environment causing heading deviations. As a result, the heading estimate from the AHRS was entirely dependent on the initialization conditions, which, relative to the ground truth's reference frame, may be inconsistent between sequences. For each sequence therefore, $\mathbf{R}^{l_{imu}}_{b_{imu},k}$ was composed with a constant heading alignment orientation, $\mathbf{R}^{l_{true}}_{l_{imu}}$, found by solving the nonlinear constrained optimization problem which minimized the mean square velocity error (MSVE) between the ground truth and integrated specific force:

$$\underset{\mathbf{R}^{l_{true}}_{l_{imu}} \in \text{SO}(3)}{\text{argmin}} \frac{1}{n}\sum_{i=1}^{n}\left(v^{l_{true}}_{l_{true}\,b_{true}} - \int\left(\mathbf{R}^{l_{true}}_{l_{imu}}\,\mathbf{R}^{l_{imu}}_{b_{imu}}\right)f^{b_{imu}}_{l_{imu}\,b_{imu}} + \begin{bmatrix}0\\0\\g\end{bmatrix}\mathrm{d}t\right)^2 \quad (6)$$

that is:

$$\underset{\mathbf{R}^{l_{true}}_{l_{imu}} \in \text{SO}(3)}{\text{argmin}} \frac{1}{n}\sum_{i=1}^{n}\left(v^{l_{true}}_{l_{true}\,b_{true}} - v^{l_{true}}_{l_{imu}\,b_{imu}}\right)^2 \quad (7)$$

where $v^{l_{true}}_{l_{true}\,b_{true}}$ was obtained by taking the derivative of the ground truth positional data:

$$v^{l_{true}}_{l_{true}\,b_{true},k} = \frac{x^{l_{true}}_{l_{true}\,b_{true},k} - x^{l_{true}}_{l_{true}\,b_{true},k-1}}{\Delta t} \quad (8)$$

Some portions of $x^{l_{true}}_{l_{true}\,b_{true}}$ were found to contain spurious readings possibly due to the visual inertial odometry failing in occasional featureless spaces. A Hampel filter was therefore applied to clean the data, replacing values greater than 2σ from the local median of a 200-sample window as shown in Figure 2.

Taking the derivative in equation (8) was noisy so a zero-phase Butterworth lowpass filter was applied with a cut-off frequency of 5 Hz. A non-causal filter was used to avoid introducing phase distortion in the training data which may otherwise affect the accuracy of velocity classification at test time. These filtered velocities were the labels for the classification algorithm to learn for the corresponding inputs.

As in RoNIN, input features were windows, length n_{win}, of angular velocity and specific force of the IMU b-frame relative to l_{imu}, coordinatized in l_{true}:

$$\begin{bmatrix}\omega^{l_{true}}_{l_{imu}\,b_{imu}} & f^{l_{true}}_{l_{imu}\,b_{imu}}\end{bmatrix}_{n_{win}\times 6} \quad (9)$$

where:

$$f^{l_{true}}_{l_{imu}\,b_{imu}} = \mathbf{R}^{l_{true}}_{b_{imu}}\,f^{b_{imu}}_{l_{imu}\,b_{imu}} \quad (10)$$

$$\omega^{l_{true}}_{l_{imu}\,b_{imu}} = \mathbf{R}^{l_{true}}_{b_{imu}}\,\omega^{b_{imu}}_{l_{imu}\,b_{imu}} \quad (11)$$
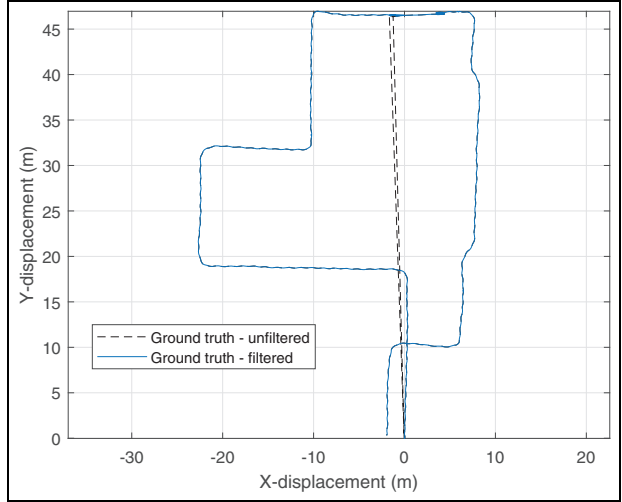

Figure 2. Hampel filter outlier replacement on ground truth trajectory data.

Note that the specific force input in equation (10), is not gravity compensated which represents a departure from the principles of strapdown inertial navigation. It was found that using gravity-compensated specific force reduced the accuracy of trajectory estimated by the algorithm.

At test time, velocity estimates from the classifier constructed trajectories by the first order numerical integration scheme:

$$x^{l_{true}}_{l_{true}\,b_{true},k} = x^{l_{true}}_{l_{true}\,b_{true},k-1} + v^{l_{true}}_{l_{true}\,b_{true},k}\,\Delta t \quad (12)$$

SC-TS velocity classification models

TS velocity classification models were generated based on clusters found by the SC algorithm for a given IMU mode of carriage featured in the data set.

An extension of the grid-based mountain clustering method, SC is a fast method for clustering high dimensional input data.[35] Economou et al.[36] used SC to obtain local models of a skid steer robot's dynamics over its steering envelope and Muhammad et al.[37] used the algorithm for accurate stance detection of human gait. Unlike fuzzy c-means (FCM) and k-means clustering, SC does not require the number of clusters to be defined a priori but rather a radius of influence, r_a. While FCM considers the Euclidean distance of each data point to the cluster center, SC has features that avoid the influence of outliers. The squash factor, f_{squash} controls the extent to which outliers in the feature space are included in a cluster while the r_{accept} and r_{reject} fractions define the potential of the first cluster center above and below which a point may be accepted or rejected as a cluster center respectively. This helps SC avoid returning marginal cluster centers. Figure 3(a) and (b) show how the cluster centers returned by SC are less influenced by outliers than FCM for a small synthetic data set.
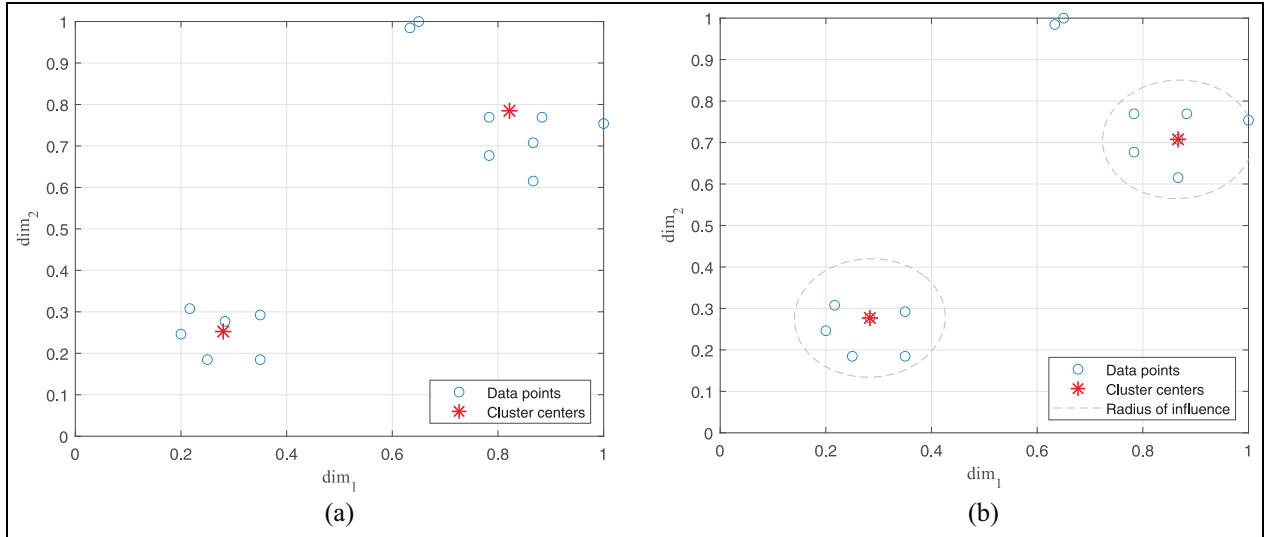
Figure 3. (a) Cluster centers estimated by FCM and (b) cluster centers estimated by SC.

Table 1. Grid search performed to find the optimal parameters of the SC algorithm.

r_a:	0.30	0.45	0.60	0.75	0.90
t_{win} (s):	0.25	0.50	0.75	1.00	1.25
f_{squash}:	0.60	0.80	1.00	1.20	1.40
$r_{,accept}$:	0.30	0.45	0.60	0.75	0.90
r_{reject}:	0.15	0.35	0.50	0.65	0.80

A grid search for the optimal r_a, f_{squash}, r_{accept}, and r_{reject} parameters for the SC algorithm was performed by observing their effect on the MSVE when training and then testing on an unseen sequence. MSVE was calculated according to equation (13).

$$MSVE = \frac{1}{n}\sum_{i=1}^{n}\left(v^{l_{true}}_{l_{true}\ b_{true}} - v^{l_{true}}_{l_{imu}\ b_{imu}}\right)^2 \quad (13)$$

3125 train and test instances were carried out using combinations of the parameters in Table 1. The lowest MSVE of $0.05\,\text{ms}^{-1}$ was achieved for $r_a = 0.3$, $t_{win} = 0.5\,\text{s}$, $f_{squash} = 1.4$, $r_{accept} = 0.45$ and $r_{reject} = 0.35$.

t_{win} was found to have little impact on the test sequence MSVE but an approximately exponential increase in the sequence computation time was observed for $t_{win} < 0.5\,\text{s}$. A value of $0.5\,\text{s}$ was thus chosen as the optimal in terms of frequency of trajectory updating.

Results

Figures 4 and 5 show examples of cluster centers found by the SC algorithm and a surface plot of the TS fuzzy rule base respectively.

The accuracy of the SC-TS velocity classification algorithm was tested against four competitive methods
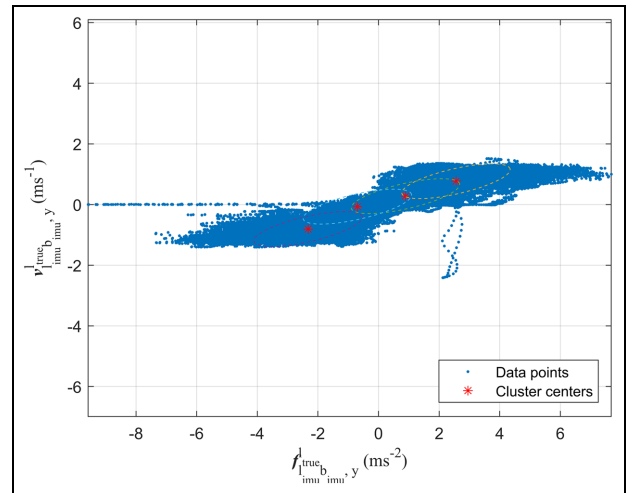

Figure 4. Cluster centers and radii found by the SC method in y-direction specific force.
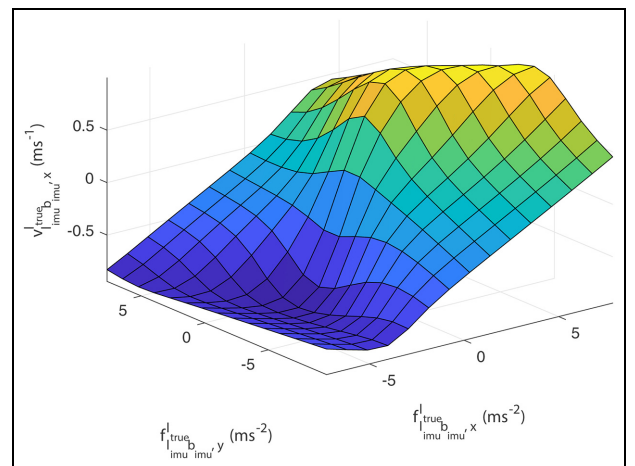

Figure 5. Surface plot mapping x-y acceleration inputs to x velocities.
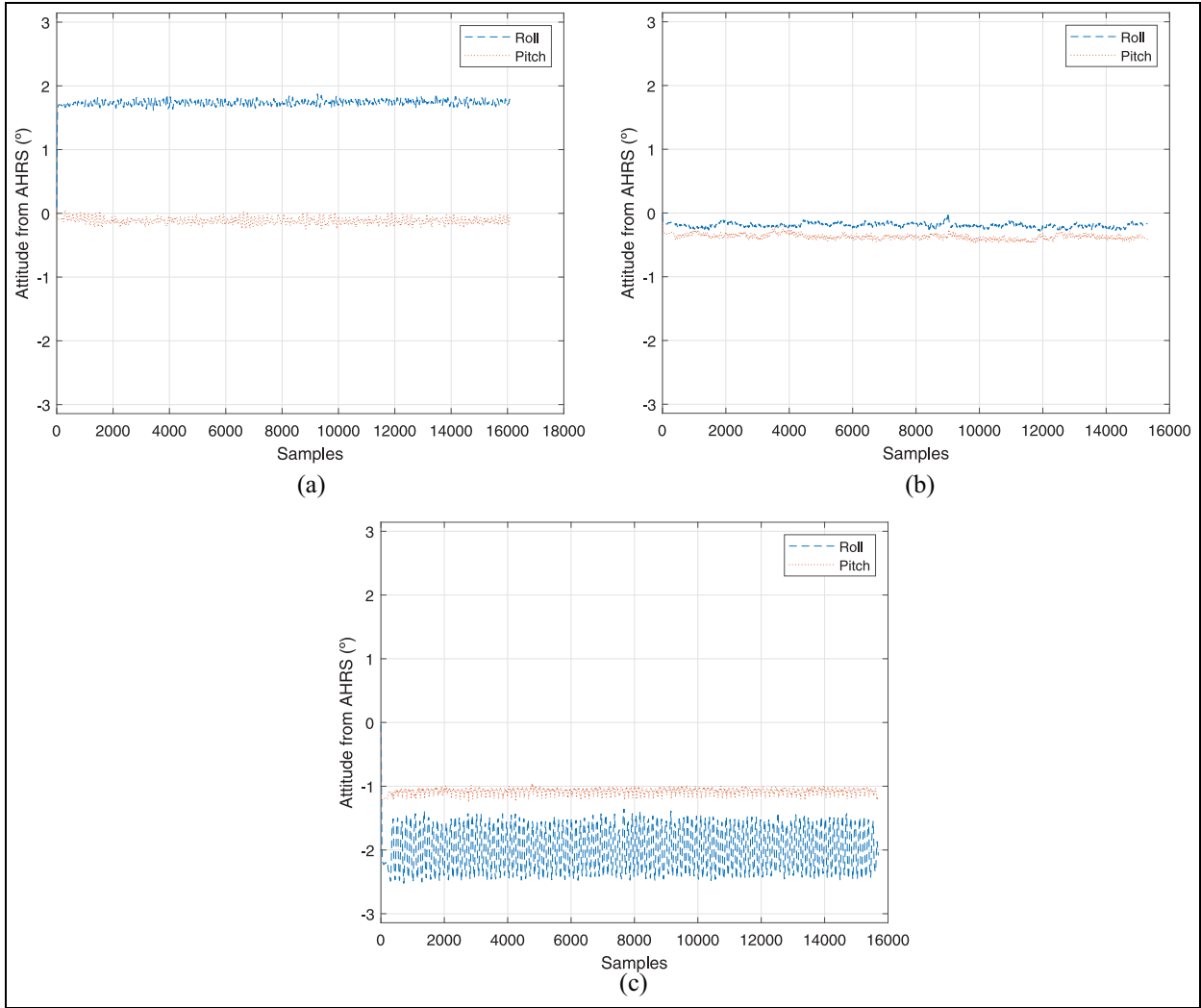
Figure 6. (a) Mode one attitude angles, (b) mode two attitude angles, and (c) mode three attitude angles.

reported in literature: RoNIN, IONet, strapdown inertial navigation and ZUPT.

The RoNIN Python code was obtained from GitHub but as the IONet code was not made public, a local implementation was made in Matlab to the authors' best abilities using details from.[38] The latest gait tracking algorithm from X-IO Technologies was used as the ZUPT baseline with default parameters.[39]

The "large scale" subset of trajectories from the OxIOD data set were used for performance evaluation, having a mean duration of 150 s. Three general IMU modes of carriage were identified by their attitude angles shown in Figure 6(a) to (c). For each sequence and algorithm, the MSVE, relative (RTE) and absolute (ATE) trajectory errors as defined by Strum et al.[40] were used as performance metrics. For the RTE, the same definition implemented in the RoNIN code was used for fairness where it was calculated as the mean root-mean square error (RMSE) for sub-trajectories of length one up to the entire sequence length.

Further tests on the RoNIN data set, were performed to verify the SC-TS method. For this data, the optimization step in equation (6) was not required as

this alignment was provided a priori. Figure 10(a) and (b) depict the trajectories estimated for the "a000" and "a010" set of sequences respectively.

The SC-TS method and other data-driven approaches such as IONet and RoNIN initially trained models to estimate vectors for certain modes of carriage of IMU sensor on the body. It was therefore of interest to investigate how the accuracy of velocity and position estimation changed due to perturbations from the orientation featured in the training data.

Using the sequence of Figure 9 (mode three test sequence), perturbed orientations, $\tilde{\mathbf{R}}_{b_{imu}}^{l_{true}}$, were generated by composing $\mathbf{R}_{b_{imu}}^{l_{true}}$ with a perturbation of randomly generated basis vector, $\boldsymbol{u}$, and angle, ϕ, incremented in steps of 5°:

$$\tilde{\mathbf{R}}_{b_{imu}}^{l_{true}} = \mathbf{R}_{b_{imu}}^{l_{true}} e^{\phi \, \mathrm{sk}(\boldsymbol{u})} \tag{14}$$

Due to the randomness in $\boldsymbol{u}$, the MSVE, RTE and ATE error metrics for each ϕ were determined by a Monte Carlo simulation of 1000 samples for ϕ between
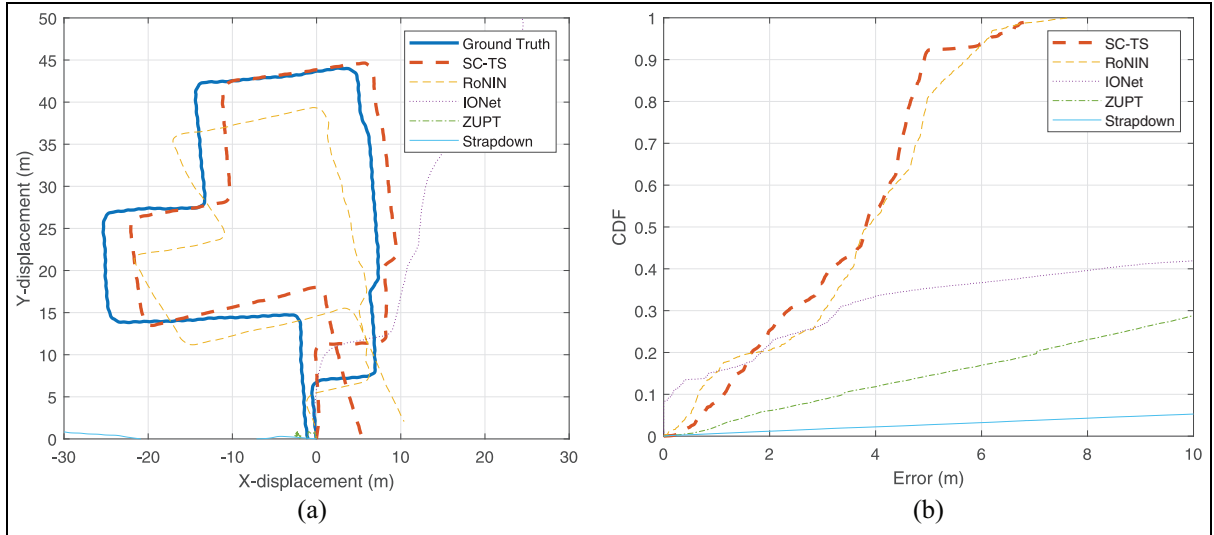
Figure 7. (a) Estimated test trajectories for mode one and (b) CDF of absolute positional error for mode one test trajectory.
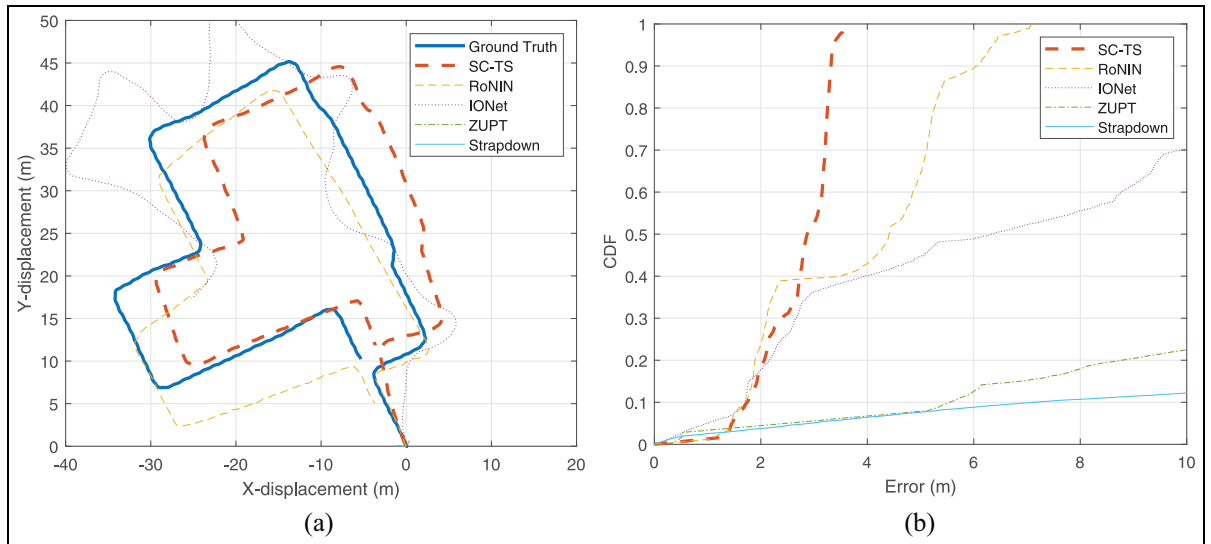

Figure 8. (a) Estimated test trajectories for mode two and (b) CDF of absolute positional error for mode two test trajectory.

$0°$ and $50°$. Mean error metrics and 1σ error bars are shown in Figures 11 to 13.

Discussion

Figures 4 and 5 show examples of clusters found in data and the surface plot of the resulting TS fuzzy rule base respectively. In Figure 5, it can be seen how positive accelerations in the x-direction intuitively map to positive x-velocities in the locally level frame. Observing the gradients in x and y axes, accelerations in the y-direction have less influence on the x-velocity output compared to accelerations in the x-direction. Output velocities in the x-direction are bounded between about $\pm 1 \text{ ms}^{-1}$. The SC-TS model is thus more transparent than other black box models generated by deep learning.

The mean error metrics in Table 2 show how the proposed SC-TS method and RoNIN achieved the same MSVE over the three "large scale" test sequences of the OxIOD data set with the former achieving a more than three-fold reduction in RTE and ATE over the latter. Using the SC method, a mean positional error rate of 4 cms^{-1} of walking motion can be expected. As both were provided with the same alignment orientation described by equation (7), the TS-SC method appears to have reduced the rate of inertially-derived positional error growth at least as well as sequence-to-one regression methods like RoNIN. The cumulative distribution function (CDF) plots of the absolute positional error in Figures 7(b), 8(b), and 9(b) support this finding. While both SC-TS's and RoNIN's input features were specific force and angular velocity projected in the locally-level frame, the superior
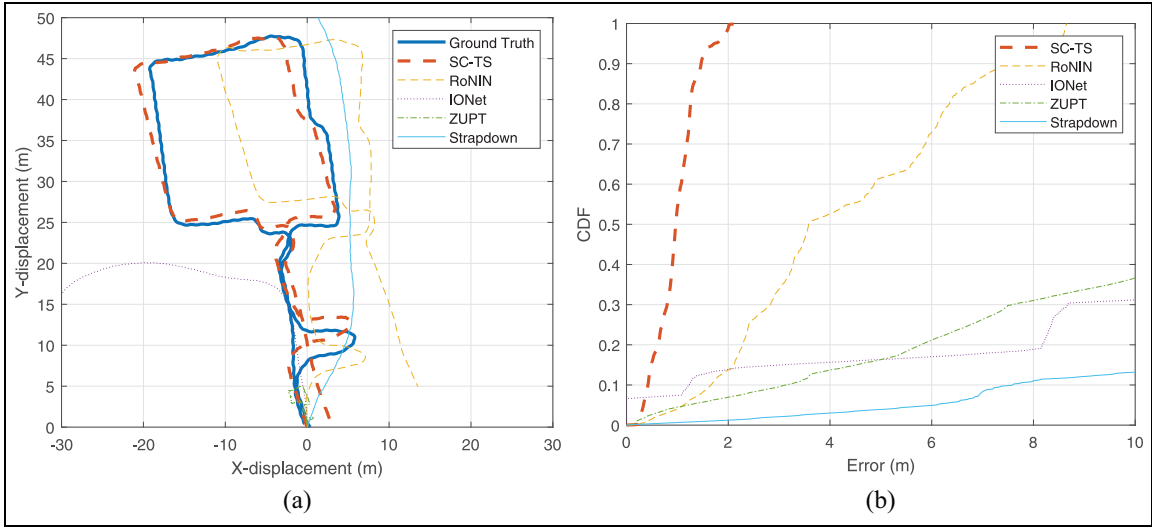
Figure 9. (a) Estimated test trajectories for mode three and (b) CDF of absolute positional error for mode three test trajectory.
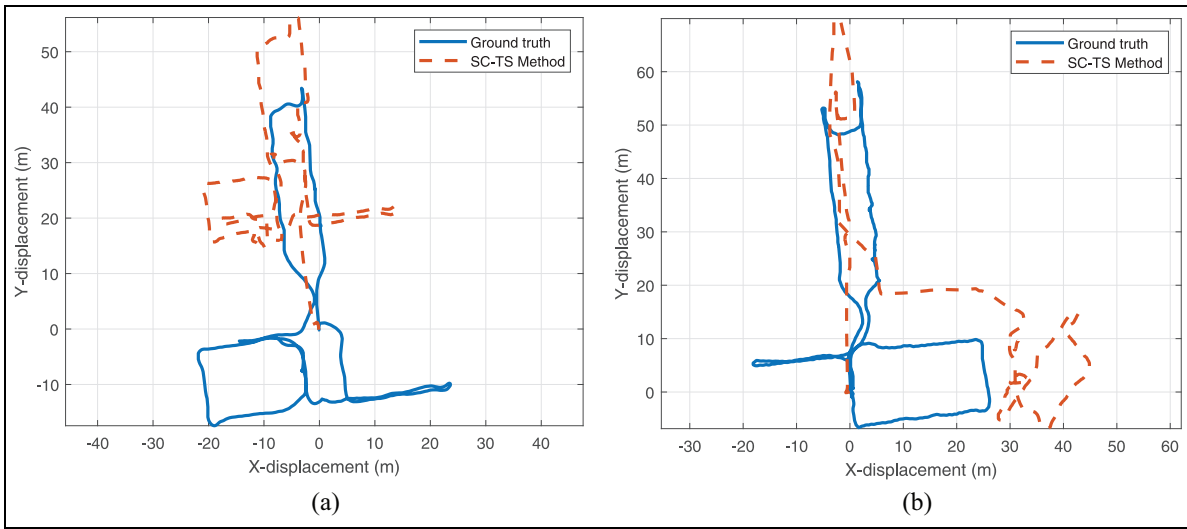

Figure 10. (a) Estimated test trajectory on RoNIN data set sequence a000_3 and (b) estimated test trajectory on RoNIN data set sequence a010_3.
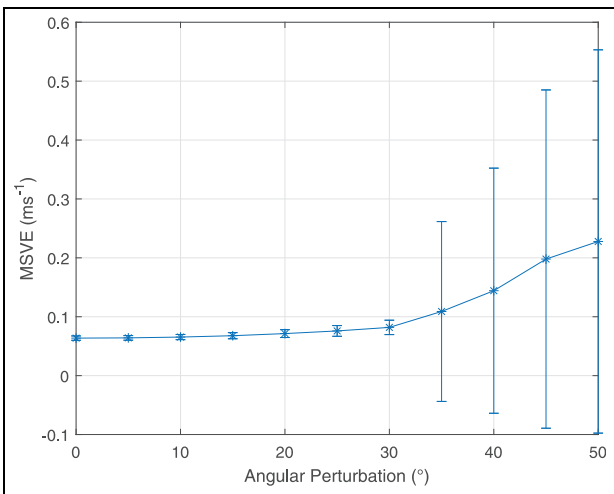

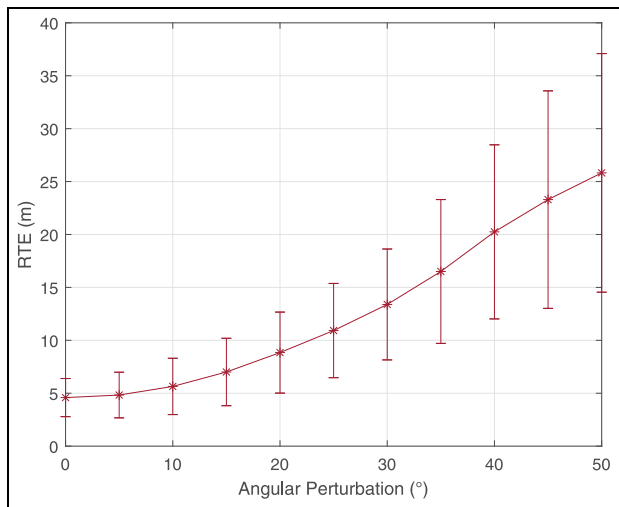**Figure 11.** Plot of MSVE with perturbation angle showing 1σ error bars.

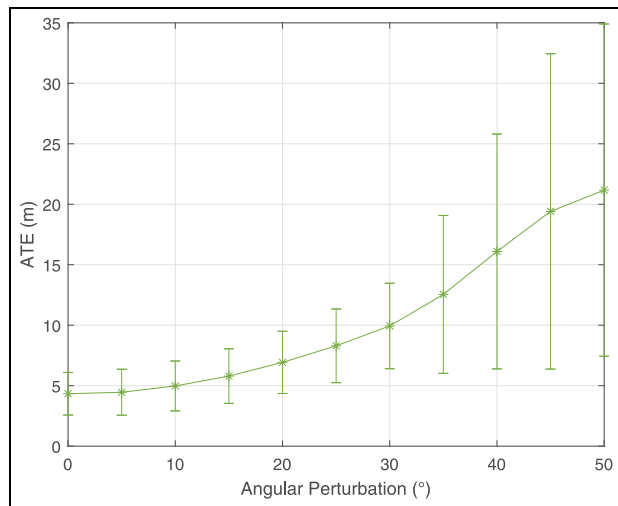**Figure 12.** Plot of RTE with perturbation angle showing 1σ error bars.

Figure 13. Plot of ATE with perturbation angle showing 1σ error bars.

Table 2. Mean test data error metrics for three modes of carriage identified.

	SC-TS:	RoNIN:	IONet:	Strapdown:	ZUPT:
MSVE (ms^{-1}):	0.04	0.04	0.73	1.72	0.31
ATE (m):	3.45	12.03	42.20	121.42	28.03
RTE (m):	2.88	13.44	81.70	173.87	16.18

accuracy in RTE and ATE of the proposed method may have been due to the tuning parameters of the SC algorithm which, found by grid search, helped create a classification rule base that was less affected by outliers compared to a deep learning regression model.

The ZUPT baseline was less accurate than the SC-TS method and RoNIN, with a mean MSVE almost eight times larger than the former methods. This was because ZUPT is intended for foot mounted IMUs whereby frequent zero velocity instances provide opportunities for integral drift estimation and compensation. Such instances were not featured in the OxIOD and RoNIN sequences as the IMUs were sampled from IMU-equipped smartphones carried on the upper body. IONet yielded a MSVE twice that of ZUPT, possibly due to input features being expressed in the b-frame rather than the l-frame, making it more sensitive to disparities in orientation featured in training and test data. Furthermore, since IONet's output features consist of a distance and change in heading angle, the accuracy of its trajectories is highly reliant upon accurate estimation of the change in heading angle by the deep neural network. A single erroneous heading change will thus be penalized heavily by the error metrics, especially the ATE. Finally, strapdown inertial navigation was the worst performing baseline due to the rapid accumulation of errors in the open loop double integration of gravity-compensated acceleration mentioned previously.

The validity of the SC method was demonstrated by applying it to the independent RoNIN data set. Visual inspection of Figure 10(a) and (b) show how the algorithm captured certain features of the ground truth trajectory but with errors visibly larger compared to the results from OxIOD. This may be due to the insufficient capture of stationary periods by the TS models which resulted in the large translational error seen in Figure 10(a) or the momentary changes in the IMU's mode of carriage featured in the RoNIN data set which makes it a more challenging benchmark than OxIOD.

The sensitivity analysis of the SC-TS method presented the increase in the MSVE, RTE and ATE that can be expected with increasing disparity between the orientation featured in the training data and that used at test time. This is useful since to achieve generalization to varying modes of carriage which may occur during natural human motion, RoNIN and IONet both proposed training models for a given orientation and then interpolating between them, akin to an ensemble modeling method. An investigation into the change in tracking accuracy with deviations from the design orientation may therefore help developers decide on the granularity with which the manifold of SO(3) should be covered by their models to meet a required error metric. For instance, referring to Figure 11, if a MSVE no greater than 10 cms^{-1} is required then the algorithm should train and interpolate between models with a granularity no coarser than 30°. The impact of this design choice on the RTE and ATE can also be estimated from Figures 12 and 13.

Conclusion

This work proposed a novel velocity classification method for more accurate positional tracking of humans carrying inertial sensors, contrasting with previous works which used deeply learnt sequence-to-one regression models. The SC-TS algorithm first identified patterns in the input data, projected a priori into a locally-level frame, and encoded the mapping to corresponding velocity vectors in a TS rule base. The integration of gravity compensated acceleration resulting in exponential positional error accumulation seen in strapdown inertial navigation was thus avoided. By solving a nonlinear optimization problem, it was shown for the first time how the misalignment between the locally-level frames defined by ground truth system and IMU can be found and applied to the sequences of the OxIOD data set. While others have used deep learning in their approaches to create black box models, the proposed SC-TS method is more transparent and interpretable. This may be an added benefit when certification of the algorithm is required before implementing in a safety critical application.

In terms of RTE and ATE metrics, the proposed SC-TS method outperformed the other baseline methods achieving mean values of 2.9 and 3.5 m respectively

over sequences of about 150 s in duration from the OxIOD data set. This was likely due to the robustness of the SC algorithm to outliers in data, with the optimal tuning parameters being found by grid search. Visual inspection of results on an independent data set confirmed the validity of the method while also revealing its vulnerability to stationary periods and momentary changes to the mode of carriage on the body which could be addressed in future work.

The sensitivity analysis presented the losses that should be expected due to deviations of the IMU's mode of carriage featured in the training data compared with that used at test time, offering a useful reference for future algorithm developers working toward an ensemble motion tracking model.

Declaration of conflicting interests

Funding

ORCID iD

Dariusz Maton https://orcid.org/0000-0002-1101-6732

Data availability statement

The data that support the findings of this study are openly available to download from Google Drive via http://deepio.cs.ox.ac.uk/ for the OxIOD data set and in the Federated Research Data Repository at https://doi.org/10.20383/102.0543 for the RoNIN data set.

References

1. Ritchie H and Urbanization R. Our world in data, https://ourworldindata.org/urbanization (2018, accessed 15 March 2022).
2. Pahlavan K and Makela JP. Indoor geolocation science and technology. *IEEE Commun Mag* 2002; 40(2): 112–118.
3. C4ADS. Above us only stars: exposing GPS spoofing in Russia and Syria, https://www.c4reports.org/aboveusonlystars (2019, accessed 11 March 2021).
4. Schmidt GT. Navigation sensors and systems in GNSS degraded and denied environments. *Chin J Aeronaut* 2015; 28(1): 1–10.
5. Jimenez Ruiz AR and Seco Granja F. Comparing ubisense, bespoon, and decawave UWB location systems: indoor performance analysis. *IEEE Trans Instrum Meas* 2017; 66(8): 2106–2117.
6. Kok M, Hol JD and Schon TB. Indoor positioning using ultrawideband and inertial measurements. *IEEE Trans Vehicular Technol* 2015; 64(4): 1293–1303.
7. Sczyslo S, Schroeder J, Galler S, et al. Hybrid localization using UWB and inertial sensors. In: *2008 IEEE international conference on ultra-wideband*, Hannover, Germany, 10–12 September 2008, no. 3, pp.89–92. New York: IEEE.
8. Jimenez Ruiz AR, Seco Granja F, Prieto Honorato JC, et al. Accurate pedestrian indoor navigation by tightly coupling foot-mounted IMU and RFID measurements. *IEEE Trans Instrum Meas* 2012; 61(1): 178–189.
9. Saab SS and Nakad ZS. A standalone RFID indoor positioning system using passive tags. *IEEE Trans Ind Electron* 2011; 58(5): 1961–1970.
10. Woodman O and Harle R. RF-based initialisation for inertial pedestrian tracking. In: Tokuda H, Beigl M, Friday A, et al. (ed.) *Pervasive: 7th international conference on pervasive computing 2009*, Nara, Japan, 11–14 May 2009, pp.238–255, Heidelberg: Springer.
11. Renaudin V, Yalak O, Tomé P, et al. Indoor navigation of emergency agents. *Eu. J. Navig* 2007; 5(3): 36–45.
12. ISO/IEC 18305:2016. Test and evaluation of localization and tracking systems.
13. Chen C, Lu X, Markham A, et al. IONet: learning to cure the curse of drift in inertial odometry. In: *32nd AAAI conference on artificial intelligence, AAAI*, New Orleans, LA, USA, 2–7 February 2018, pp.6468–6476. Palo Alto: AAAI Press.
14. Foxlin E. Motion tracking requirements and technologies. In: Stanney K (ed.) *Handbook of virtual environment technology*. Mahwah, NJ: Lawrence Erlbaum Associates, 2002, pp.163–210.
15. Kok M, Hol JD and Schön TB. Using inertial sensors for position and orientation estimation. *Found Trends Signal Process* 2017; 11(1-2): 1–153.
16. Sanz R, Rodenas L, Garcia P, et al. Improving attitude estimation using inertial sensors for quadrotor control systems. In: *2014 international conference on unmanned aircraft systems (ICUAS)*, Orlando, FL, USA, 27–30 May 2014, pp.895–901, New York: IEEE.
17. Madgwick SOH, Harrison AJL and Vaidyanathan R. Estimation of IMU and MARG orientation using a gradient descent algorithm. In: *2011 IEEE international conference on rehabilitation robotics*, Zurich, Switzerland, 29 June–1 July 2011, pp.1–7. New York: IEEE.
18. Pflimlin JM, Hamel T and Souères P. Nonlinear attitude and gyroscope's bias estimation for a VTOL UAV. *Int J Syst Sci* 2007; 38(3): 197–210.
19. Woodman OJ. An introduction to inertial navigation. Report, University of Cambridge, UK, August 2007.
20. El-Sheimy N and Youssef A. Inertial sensors technologies for navigation applications: state of the art and future trends. *Satell Navig* 2020; 1: 2.
21. Foxlin E. Pedestrian tracking with shoe-mounted inertial sensors. *IEEE Comput Graph Appl* 2005; 25(6): 38–46.
22. Cavallo F, Sabatini AM and Genovese V. A step toward GPS/INS personal navigation systems: real-time assessment of gait by foot inertial sensing. In: *2005 IEEE/RSJ international conference on intelligent robots and systems, IROS*, Edmonton, AB, Canada, 2–6 August 2005, pp.1187–1191. New York: IEEE.
23. Fan Q, Zhang H, Pan P, et al. Improved pedestrian dead reckoning based on a robust adaptive Kalman filter for indoor inertial location system. *Sensors* 2019; 19: 294–315.

24. Yun X, Bachmann ER, Moore H, et al. Self-contained position tracking of human movement using small inertial/magnetic sensor modules. In: *Proceedings 2007 IEEE international conference on robotics and automation*, Rome, Italy, 10–14 April 2007, pp.2526–2533. New York: IEEE.

25. Wahlstrom J and Skog I. Fifteen years of progress at zero velocity: a review. *IEEE Sens J* 2021; 21(2): 1139–1151.

26. De Cillis F, Faramondi L, Inderst F, et al. Hybrid indoor positioning system for first responders. *IEEE Trans Syst Man Cybern Syst* 2020; 50(2): 468–479.

27. Glueck M, Oshinubi D, Schopp P, et al. Real-time auto-calibration of MEMS accelerometers. *IEEE Trans Instrum Meas* 2014; 63(1): 96–105.

28. Herath S, Yan H and Furukawa Y. RoNIN: Robust neural inertial navigation in the wild: benchmark, evaluations, new methods. In: *2020 IEEE international conference on robotics and automation (ICRA)*, Paris, France, 31 May–31 August 2020, pp.3146–3152. New York: IEEE.

29. Silva do Monte Lima JP, Uchiyama H and Taniguchi RI. End-to-end learning framework for IMU-based 6-DOF odometry. *Sensors* 2019; 19: 3777.

30. Karim F, Majumdar S and Darabi H. Adversarial attacks on time series. *IEEE Trans Pattern Anal Mach Intell* 2021; 43(10): 3309–3320.

31. Chen C, Zhao P, Lu CX, et al. OxIOD: the dataset for deep inertial odometry, https://doi.org/10.48550/arXiv.1809.07491 (2018, accessed 14 July 2021).

32. Herath S, Furukawa Y and Ronin Y. Robust neural inertial navigation, https://doi.org/10.20383/102.0543 (2022, accessed 12 March 2022).

33. Madgwick SOH, Wilson S, Turk R, et al. An extended complementary filter for full-body MARG orientation estimation. *IEEE/ASME Trans Mechatron* 2020; 25(4): 2054–2064.

34. Madgwick SOH. *AHRS algorithms and calibration solutions to facilitate new applications using low-cost MEMS*. PhD Thesis, University of Bristol, UK, 2014.

35. Chiu S. Method and software for extracting fuzzy classification rules by subtractive clustering. In: Smith MH, Lee MA, Keller J, et al. (ed.) *Proceedings of North American fuzzy information processing*, Berkeley, CA, USA, 19–22 June 1996, pp.461–465. New York: IEEE.

36. Economou JT, Tsourdos A, White BA, et al. Takagi–Sugeno model synthesis of a quasi-linear multiwheeled mobile robot. *Int J Syst Sci* 2003; 34(14-15): 805–818.

37. Muhammad MN, Salcic Z and Wang KIK. Subtractive clustering as ZUPT detector. In: Apduhan BO, Zheng Y, Nakamoto Y, et al. (ed) *Proceedings - 2014 IEEE international conference on ubiquitous intelligence and computing, 2014 IEEE international conference on autonomic and trusted computing, 2014 IEEE international conference on scalable computing and communications and associated SY*, Denpasar, Bali, Indonesia, 9–12 December 2014, pp.349–355. New York: *IEEE*.

38. Chen C. *Learning methods for robust localization*. PhD Thesis, University of Oxford, UK, 2020.

39. X-IO Technologies. Gait tracking with X-IMU, https://x-io.co.uk/gait-tracking-with-x-imu/ (2013, accessed 5 May 2022).

40. Sturm J, Engelhard N, Endres F, et al. A benchmark for the evaluation of RGB-D SLAM systems. In: *2012 IEEE/RSJ international conference on intelligent robots and systems*, Vilamoura-Algarve, Portugal, 7–12 October 2012, pp.573–580. New York: IEEE.