Vol.2022-MBL-104 No.8
Vol.2022-UBI-75 No.8
Vol.2022-CDS-35 No.8
Vol.2022-ASD-24 No.8
2022/9/5

# What Can Data-driven Calibration Do for 6DOF Inertial Odometry?

Huakun Liu[1,a)] Monica Perusquía-Hernández[1,b)] Naoya Isoyama[1,c)]
Hideaki Uchiyama[1,d)] Kiyoshi Kiyokawa[1,e)]

**Abstract:** For low-cost IMU, the uncalibrated bias and noise will quickly propagate 6DOF odometry errors over time. This paper proposes a data-driven accelerometer calibration method based on a dilated convolution network. Then, with a state-of-the-art gyroscope calibration method, we comprehensively analyze the impact of data-driven calibration on 6DOF inertial odometry. The experimental results show that our data-driven accelerometer calibration can reduce the bias by a factor of 5 to 10 and decreases the noise by a factor of 2 to 5. Through our exhaustive evaluations and analysis of data-driven calibration methods, the primary finding is that the data-driven calibration methods can slow down the error growth rate by 40-200 times. However, the effect of accelerometer calibration is only noticeable after calibrating the gyroscope. This fact would be experimental support for the design of future data-driven 6DOF inertial odometry.

**Keywords:** inertial odometry, 6DoF, calibration, neural network

## 1. Introduction

Six degrees of freedom (6DOF) inertial odometry is a process that uses an inertial measurement unit (IMU) to track the 3D space position and 3D orientation of the tracking target relative to its starting pose. An IMU consists of a triaxial accelerometer that measures acceleration and a triaxial gyroscope that measures the angular velocity in the body frame. The 6DOF inertial odometry based on sensor fusion is widely used in intelligent aerial robots [1], underwater drone navigation [2, 3], and a variety of other applications [4–6]. Existing methods typically rely on a high-precision IMU to achieve highly-accurate 6DOF inertial odometry [7]. However, the high cost limits its application scenarios.

Over the last two decades, low-cost, small-size microelectromechanical system (MEMS) IMUs are installed in various smart devices, such as smartphones, AR/VR glasses, and drones. Accordingly, they have enabled an ever-increasing diversity of applications based on inertial sensors [8–10]. In particular, recent research has shown that integrating inertial sensors with machine learning techniques is able to achieve highly accurate pedestrian dead reckoning, a 2D space 3DOF inertial odometry [11, 12]. However, the rapid accumulation of the errors from noise and biases of IMU signals makes 6DOF inertial odometry a challenge.

The MEMS IMU errors include scale factor, axis misalignment, zero bias, and noise. During the integration process, these errors explode exponentially with time. Hence, it is necessary to perform IMU calibration to measure the error parameters as accurately as possible. Classical calibration methods, such as the multi-position method, can accurately identify the unknown error parameters, remove errors, and reduce noise [13]. However, they have strict requirements for expensive experimental equipment or complex setups.

Recent data-driven-based calibration methods, such as Denoising IMU Gyro (DIG) [14] and Temporal Convolutional Network Denoising IMU Gyro (TCN-DIG) [15] use supervised learning to directly output the target correction term, i.e., bias and noise. They demonstrated the effectiveness of data-driven calibration methods. However, only gyroscope calibration is involved. Engelsman *et al.* did an exhaustive literature review of data-driven MEMS-IMU calibration techniques [16]. They showed that data-driven accelerometer calibration is more challenging than data-driven gyroscope calibration. This is because gyroscope data is self-contained, and the offset can be simply optimized to zero. In contrast, an accelerometer measures the specific force that contains both linear acceleration and gravity. In this review, they also proposed four data-driven models based on LSTM, RNN, GRU, and kNN to calibrate accelerometer signals. However, they only verified the proposed methods on static and simulated data. In addition, all the data-driven calibration methods are limited to validating the improvement on the original signal data or the orientation without actually analyzing how the method affects the inertial odom-

[1] Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara, 630–0101, Japan
[a)] liu.huakun.li0@is.naist.jp
[b)] m.perusquia@is.naist.jp
[c)] isoyama@is.naist.jp
[d)] uchiyama@is.naist.jp
[e)] kiyo@is.naist.jp

etry results.

To fill this gap, we propose a data-driven accelerometer calibration method based on a dilated convolution network. We design a new loss function for training the model from the ground truth trajectory. Then, we evaluate the method using a public dataset that captured with a micro aerial vehicles (MAV). Furthermore, we explore the impact of data-driven calibration methods on odometry with a state-of-the-art data-driven gyroscope calibration method. To the best of our knowledge, this is the first detailed evaluation of the data-driven calibration methods on 6DOF inertial odometry. The main contributions of this paper are as follows:

- An accelerometer calibration method based on a dilated convolution network for calibrating real and dynamic motion data.
- A reasonably designed mean-variance loss function based on the sensor error model for training the network.
- A comprehensive analysis on the impact of both data-driven gyroscope and accelerometer calibration methods on the 6DOF inertial odometry. The evaluation from sensor readings comparison to position estimation reveals the secret of data-driven calibration methods.

## 2. Related Work

Calibration should be conducted before or during the estimate process to suppress the negative impact of IMU errors on inertial odometry results. MEMS IMU calibration can be divided into two categories: classical calibration and data-driven calibration.

### 2.1 Classical Calibration

Classical calibration methods use parameter state estimation algorithms, such as least squares [17], maximum likelihood estimation [18] to solve the error coefficient by fusing the IMU readings and external reference information. They are further divided into two categories: non-autonomous and autonomous calibration [19], based on the source of reference information. The reference information of non-autonomous calibration is obtained from high-precision equipment such as high-precision turntables [13]. Autonomous calibrations rely on the external reference excitation, such as the local gravity, magnetic fields, the Earth's rotation rate. For details about traditional calibration, a survey on MEMS inertial sensor calibration is a reference [19].

### 2.2 Data-driven Gyroscope Calibration

Deep learning has provided new possibilities for unimodal position estimation from IMU. For instance, it is used to extract latent features from IMU signals to estimate the velocity [11], orientation [20], and displacement [21–24]. These methods have seen great success in 6DOF inertial odometry owing to avoiding the accumulation of errors caused by the integration process.

Additionally, using data-driven deep learning to calibrate IMU has become popular. In [25], the first LSTM-RNN-based denoising method was proposed to denoise IMU gyroscope signals. Compared with an autoregressive and moving average model, the standard deviation of the denoised signals decreased by at most 42.4% with deep learning. Brassard *et al.* proposed a convolutional neural network to predict the gyroscope correction term, i.e., zero bias and noise, and to find the optimal coefficients of scale factor and axis-misalignment during training from measured accelerometer and gyroscope readings [14]. Then, Huang *et al.* used a temporal convolutional network to improve the performance of gyroscope calibration further [15]. They showed that the orientation estimated from the calibrated gyroscope data could be used to improve the accuracy of position estimation from visual-inertial odometry (VIO). To solve the low generalizability problem of data-driven denoising model, Yao [26] proposed a few-shot domain adaptation gyroscope calibration method that consists of Embedding module, Restructor module and Generator module.

### 2.3 Data-driven Accelerometer Calibration

Compared with data-driven gyroscope calibration, fewer works address accelerometer calibration through the data-driven method. The main reason would be the increased difficulty of predicting biases. MEMS accelerometer senses not only the linear acceleration of itself but also the local gravity. As a result, the presence of non-zero values results in accelerometers with a wider dynamic range and larger noise densities compared to gyroscopes [16].

Chen *et al.* [27] used a convolutional neural network to reduce errors from both accelerometer and gyroscope simultaneously in a laboratory environment. Engelsman and Klein [16] implemented and modified several learning algorithms, including unidirectional bi-layer LSTM, bi-directional one-layer RNN, and bi-directional one-layer GRU, for calibrating the accelerometer. The evaluation of a simulated dataset and static accelerometer data showed clear advantages of the data-driven accelerometer calibration method over classical methods.

## 3. Data-Driven Accelerometer Calibration

In this section, we first present the sensor error model and the kinematic motion model used in inertial odometry, followed by our data-driven accelerometer calibration method.

### 3.1 Preliminaries

The MEMS IMU consists of an accelerometer and a gyroscope measuring the acceleration $\boldsymbol{a}$ and angular velocity $\boldsymbol{\omega}$ of the carrier. However, the measurement contains not only the target $\boldsymbol{a}$ and $\boldsymbol{\omega}$, but also other error terms. These errors are mainly caused by imperfect assembling procedures and the influence of temperature on the MEMS IMU's silicon [19].

**Sensor error model** Given a three-axis strapdown accelerometer and gyroscope, a commonly used error model is established as follows [19]:

$$\tilde{\boldsymbol{u}} = (\boldsymbol{S} + \boldsymbol{N})\boldsymbol{u} + \boldsymbol{b} + \boldsymbol{n} \tag{1}$$

where $\tilde{\boldsymbol{u}}, \boldsymbol{u}$ denote the measurement output and ideal value of acceleration or angular velocity, respectively.

$$\boldsymbol{S} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix}$$

is the scale factor that refers to the ratio between the output quantity and the input quantity. This ratio is caused mainly by sensitivity of the circuit on each axis. It is a $3 \times 3$ identity matrix in the ideal condition.

$$\boldsymbol{N} = \begin{bmatrix} 0 & \gamma_{xy} & \gamma_{xz} \\ \gamma_{yx} & 0 & \gamma_{yz} \\ \gamma_{zx} & \gamma_{zy} & 0 \end{bmatrix}$$

denotes the axis-misalignment error results from the nonorthogonality between each axis, which is caused by the manufacturing technique limitations.

$$\boldsymbol{b} = \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix}$$

is called zero-bias: the output values from the accelerometer or gyroscope when the measured physical quantity equals zero.

$$\boldsymbol{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}$$

is commonly assumed the high-frequency random sensor white noise that follows the zero-mean Gaussian distribution.

**Kinematic motion model** Inertial odometry estimates the trajectory from the acceleration and angular velocity obtained from IMU using a kinematic motion model. The key point is to rotate the acceleration from IMU frame to a fixed global frame, then accumulate it to compute the velocity and moving distance.

Given the angular velocity $\boldsymbol{\omega}$ obtained from gyroscope, the rotation matrix $\boldsymbol{R} \in SO(3)$ that maps from IMU frame to global frame at timestamp $k$ can be expressed as:

$$\boldsymbol{R}_k = \boldsymbol{R}_{k-1} \exp(\boldsymbol{\omega}_{k-1} \Delta t) \tag{2}$$

where $\exp(\cdot)$ is $SO(3)$ exponential map and $\Delta t$ is the time interval of two consecutive frames. The velocity $\boldsymbol{v}$ in global frame is then calculated by rotating the measured acceleration $\boldsymbol{a}$ and removing the local gravity $\boldsymbol{g}$.

$$\boldsymbol{v}_k = \boldsymbol{v}_{k-1} + (\boldsymbol{R}_{k-1}\boldsymbol{a}_{k-1} + \boldsymbol{g})\Delta t \tag{3}$$

Finally, the position in the fixed global frame is calculated.

$$\boldsymbol{p}_k = \boldsymbol{p}_{k-1} + \boldsymbol{v}_{k-1}\Delta t + \frac{1}{2}(\boldsymbol{R}_{k-1}\boldsymbol{a}_{k-1} + \boldsymbol{g})\Delta t^2 \tag{4}$$

### 3.2 Network Structure

According to the sensor error model in (1), the measured acceleration $\tilde{\boldsymbol{a}}_k$ of an IMU accelerometer at timestamp $k$ can be expressed as follows:

$$\tilde{\boldsymbol{a}}_k = (\boldsymbol{S} + \boldsymbol{N})\boldsymbol{a}_k + \boldsymbol{b}_k + \boldsymbol{n}_k$$

Then, the ideal acceleration $\boldsymbol{a}_k$ can be estimated as

$$\boldsymbol{a}_k = (\boldsymbol{S} + \boldsymbol{N})^{-1}(\tilde{\boldsymbol{a}}_k - (\boldsymbol{b}_k + \boldsymbol{n}_k)) \tag{5}$$

To simplify the problem, we denote $(\boldsymbol{S}+\boldsymbol{N})^{-1}$ and $(\boldsymbol{b}_k+\boldsymbol{n}_k)$ in (5) by $\boldsymbol{C} \in \mathbb{R}^{3\times3}$ and $\boldsymbol{\epsilon}_k \in \mathbb{R}^3$, respectively. $\boldsymbol{C}$ contains both scale factor and axis-misalignment, $\boldsymbol{\epsilon}_k$ is the correction term contains zero bias and noise. They are expressed as follows:

$$\boldsymbol{C} = \begin{bmatrix} s_x & \gamma_{xy} & \gamma_{xz} \\ \gamma_{yx} & s_y & \gamma_{yz} \\ \gamma_{zx} & \gamma_{zy} & s_z \end{bmatrix}^{-1} \quad \boldsymbol{\epsilon}_k = \begin{bmatrix} b_x + n_x \\ b_y + n_y \\ b_z + n_z \end{bmatrix} \tag{6}$$

Then, we have

$$\boldsymbol{a}_k = \boldsymbol{C}(\tilde{\boldsymbol{a}}_k - \boldsymbol{\epsilon}_k) \tag{7}$$

We define the neural network structure as

$$\hat{\boldsymbol{\epsilon}}_k = f\Big((\tilde{\boldsymbol{a}}_{k-N}, \tilde{\boldsymbol{\omega}}_{k-N}), \cdots, (\tilde{\boldsymbol{a}}_k, \tilde{\boldsymbol{\omega}}_k)\Big)$$

where $f(\cdot)$ is the function defined by a dilated convolutional neural network, which is a convolutional neural network that achieves a quasi-temporal neural network by using the dilated gap and is widely used for data-driven gyroscope calibration [14, 15, 26, 28].

Compared with recurrent neural network (RNN) which are commonly used to process the time sequence data, the dilated convolutional neural network not only maintains the temporal ordering of the data as RNN, but also achieves higher computational efficiency due to the fewer parameters. In addition, the dilated convolutional neural network achieved higher accuracy in terms of gyroscope calibration with simpler network architecture compared to LSTM.

Leveraging the past $N$ frames information, i.e., 3-axis accelerations and 3-axis angular velocities, we predict the correction term $\hat{\boldsymbol{\epsilon}}_k$ for acceleration $\boldsymbol{a}_k$. Then, as similar to [14], we initialize the $\hat{\boldsymbol{C}}$ in (7) as $3 \times 3$ identity matrix, and optimize it during training. Finally, the corrected acceleration $\hat{\boldsymbol{a}}_k$ is estimated by

$$\hat{\boldsymbol{a}}_k = \hat{\boldsymbol{C}}(\tilde{\boldsymbol{a}}_k - \hat{\boldsymbol{\epsilon}}_k) \tag{8}$$

### 3.3 Loss Function Design

Defining a loss function for calibrating accelerometer is more difficult than gyroscope calibration. Acquiring the ground truth acceleration $\boldsymbol{a}$ and angular velocity $\boldsymbol{\omega}$ at IMU frequency (200Hz or more) is not feasible in practice because the high-precision tracking systems are normally 20-120 Hz. The loss function of data-driven gyroscope calibration normally defined by computing the difference of integrated orientation increments to reduce the frequency of IMU to the same as the tracking systems. However, for accelerometer

calibration, (3) shows that the velocity calculation is not only related to the acceleration, but also to the angular velocity. Thus, computing the integrated increments of velocity requires both the measured acceleration and the ground truth angular velocity $\boldsymbol{\omega}$ at IMU frequency.

In our method, we derive a pseudo ground truth acceleration $\boldsymbol{a}$ in IMU frame from the interpolated ground truth velocity and orientation as

$$\boldsymbol{a}_i = \boldsymbol{R}_i^T \left( \frac{(\boldsymbol{v}_{i+1} - \boldsymbol{v}_i)}{\Delta t} - \boldsymbol{g} \right) \tag{9}$$

According to the sensor error model, the difference between measured acceleration and pseudo ground truth acceleration approximately follows a Gaussian distribution with zero-bias as the mean. The variance of the distribution represents the signal-noise level. Based on this, we define the loss function $\mathcal{L}$ as minimizing the mean and variance of the difference between predicted acceleration and pseudo ground truth acceleration. In addition, zero bias and Gaussian noise change over time as they are affected by temperature, pressure, vibrations, and so forth. Therefore, with window size $t$, we split the sequence into $m$ segments and optimize it based on the assumption that the zero-bias and noise level are constant within each segment. The loss function $\mathcal{L}$ are defined as follows:

$$\mathcal{L}_{1,j} = \frac{1}{t} \sum_{i=1}^{t} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i) \tag{10}$$

$$\mathcal{L}_{2,j} = \frac{1}{t} \sum_{i=1}^{t} (\hat{\boldsymbol{a}}_i - \boldsymbol{a}_i - \mathcal{L}_{1,j})^2 \tag{11}$$

$$\mathcal{L} = \sum_{j=1}^{m} (\lambda_1 \mathcal{L}_{1,j} + \lambda_2 \mathcal{L}_{2,j}) \tag{12}$$

where $\lambda_1$ and $\lambda_2$ are two hyper-parameters to balance the influencing of mean and variance losses.

## 4. Benchmark Setup

To reveal the secrets of data-driven calibration methods and analyze their impact on the 6 DOF inertial odometry, we designed and conducted comprehensive experiments. The experiments are based on EuRoC dataset [29] that is commonly used by data-driven calibration methods with a classic data-driven gyroscope calibration method and our accelerometer calibration proposed in Section 3.

### 4.1 Dataset

The analysis is based on the EuRoC, a visual inertial dataset collected on a micro aerial vehicle. There are two types of data provided in EuRoC. The first batch contains five sequences recorded in a large machine hall. Angular velocity and specific force were measured using an uncalibrated ADIS16448 IMU with 200 Hz. A Leica Nova MS50 laser tracker provided ground truth positions at a rate of 20 Hz. The second batch contains six sequences recorded in a Vicon room equipped with a motion capture system. In addition to the data from the laser tracker and IMU, the 6D pose of the MAV was recorded by the Vicon motion capture system at a rate of 100 Hz.

**Table 1** Train and test sequences of EuRoC.

| Train sequences | Test sequences (No.) |
|---|---|
| MH_01_easy | MH_02_easy (1) |
| MH_03_medium | MH_04_difficult (2) |
| MH_05_difficult | V1_03_difficult (3) |
| V1_02_medium | V1_01_easy (4) |
| V2_01_easy | V2_02_medium (5) |
| V2_03_difficult | |

All data from the laser tracker and Vicon are precisely time-space aligned with the IMU measurements. Then, a classic maximum likelihood state estimator incorporated all ground truth and IMU measurements to estimate the final ground truth orientation, position, velocity, and the biases of the gyroscope and accelerometer.

The data was split into a training set and a test set, as shown in Table 1 [14]. The first 1.5 minutes of each training sequence is used to train the model, and the remaining data is used as the validation dataset.

### 4.2 Evaluation Process and Metrics

We conduct the evaluation following the process of kinematic motion model as introduced in Section 3.1. This enables a detailed evaluation of the impact of data-driven calibration methods on inertial odometry. To simplify the computation, we assume that the acceleration and angular velocity are constant within two consecutive frames, the gravity field is uniform, and the Coriolis forces and the earth's curvature are ignored. The evaluation process is structured into three steps, as illustrated in Table 2.

**Evaluation on the Signal:** First, we measure the calibration effectiveness by comparing the first three of the following signals with the pseudo ground truth signals:

- **raw signals (RAW)**, that is the measurements of uncalibrated accelerometer and gyroscope.
- **calibrated signals with neural network-based methods (NN)**, which is calibrated using data-driven methods. This is the main focus in this study.
- **debiased signals with provided zero-bias (PB)**, which removed the zero-bias near-completely with the biases provided in the dataset.
- **derived pseudo ground truth signals (PGT)**, which is derived from the interpolated ground truth velocity and orientation.

For each test sequence, RAW signals, that is, raw acceleration and angular velocity, are used as input to the data-driven calibration models and output the NN acceleration and angular velocity. For the accelerometer calibration, we use the method we proposed in Section 3. For the gyroscope, we choose DIG [14]. This method outperformed top-ranked calibration algorithms, and the orientation estimation results almost achieved the same level as VIO.

In addition to the NN signals, we use the biases provided in the dataset to perform a debias operation on the RAW signals to obtain PB signals. The provided bias is estimated us-

**Table 2** Overview of the evaluation.

| | Evaluation elements | Description | Metrics |
|---|---|---|---|
| Step 1 | acceleration<br>angular velocity | compare the mean and variance of<br>the differences between RAW, NN, and PB signals and the PGT | mean, variance of<br>({RAW, NN, PB} − PGT) |
| Step 2 | velocity | calculate the velocity from RAW, NN, PB acceleration with<br>GT orientation, then compare with GT velocity | AVE, RVE |
| | orientation | calculate the orientation from RAW, NN, and PB angular velocity,<br>then compare with GT orientation | AOE, ROE |
| Step 3 | position | calculate the position from sixteen combinations of RAW, NN,<br>PB, and PGT acceleration and angular velocity,<br>then compare with GT position | ATE, RTE |

ing a classic maximum likelihood state estimator that fuses all ground truth data and measurements. It shows the best results after a near-complete calibration, thus showing the effect of calibration on inertial odometry.

Since there is no possibility to obtain the real acceleration and angular velocity, we derived pseudo ground truth acceleration and angular velocity, PGT, from the interpolated ground truth velocity and orientation. In this step, we evaluate the RAW, NN, and PB signals by calculating the mean and variance of the differences between them and PGT signals. The closer the mean is to zero, the smaller the variance is, the better.

**Evaluation on the Velocity and Orientation:** The second evaluation is to measure the impact of calibrated signals on the intermediate data, i.e., velocity and orientation, of the kinematic motion model. We calculate the orientation according to (2) from RAW, NN, and PB angular velocity. We also compute the velocity based on (3) from RAW, NN, and PB acceleration with GT orientation. Then, we compare the inferred velocity and orientation with the ground truth (GT) velocity and orientation. The following metrics are used:

- Absolute Velocity Error (AVE): the root mean squared error (RMSE) between the ground truth velocity and estimated velocity for a whole sequence as

$$\text{AVE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} ||v_i - \hat{v}_i||_2^2}$$

  where $i$ is the time step and $n$ is the total frames. $v_i$ denotes the ground truth velocity at time $i$, and $\hat{v}_i$ is the velocity estimated through (3) from RAW, NN, and PB acceleration.
- Time-Normalized Relative Velocity Error (RVE): the RMSE between the ground truth velocity and estimated velocity over window size $t$ as

$$\text{RVE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} ||(v_{i+t} - v_i) - (\hat{v}_{i+t} - \hat{v}_i)||_2^2}$$

  As demonstrated in [14], relative error is more appropriate for comparing odometry results. Because the absolute error is highly related to the length of the sequence, the relative error focuses more on measuring the accuracy within a certain time span. In this analysis, we use $t = 1, 10, 20, 30, 40, 50, 60s$ for all relative error metrics.
- Absolute Orientation Error (AOE): the RMSE between the ground truth and estimated orientation for a whole trajectory sequence as

$$\text{AOE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} ||\log(R_i^T \hat{R}_i)||_2^2}$$

  where $R_i$ and $\hat{R}_i$ denote the ground truth and estimated rotation matrices from RAW, NN, and PB angular velocity at timestamp $i$ through (2), respectively. $\log(\cdot)$ is the $SO(3)$ logarithm map.
- Time-Normalized Relative Orientation Error (ROE): the RMSE between ground truth orientation and estimated orientation over window size $t$ as

$$\text{ROE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} ||\log(\delta R_{i,i+t}^T \delta \hat{R}_{i,i+t})||_2^2}$$

  where $\delta R_{i,i+t}$ and $\delta \hat{R}_{i,i+t}$ are the increment rotation matrices over $t$ timesteps and can be expressed as

$$\delta R_{i,i+t} = R_i^T R_{i+t}$$
$$\delta \hat{R}_{i,i+t} = \prod_{j=i}^{i+t-1} \exp(\hat{\omega}_j)$$

  where $\hat{\omega}$ denotes RAW, NN, and PB angular velocity.

**Evaluation on the Position:** We finally measure the impact of calibrated signals on the position estimation by comparing the inferred and ground truth positions. In order to explore the respective effects of the acceleration and angular velocity on the position estimation, we calculate the position from 16 different combinations of RAW, NN, PB, PGT acceleration, and angular velocity. Similar to the intermediate data evaluation, we use absolute and relative error to evaluate it. They are defined as follows:

- Absolute Translation Error (ATE): the RMSE between the ground truth and estimated position. To compare different sequences, we calculate the ATE for the first 60 seconds rather than the whole sequence.

$$\text{ATE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left|\left|\hat{p_i} - p_i\right|\right|_2^2}$$

- Relative Translation Error (RTE): the RMSE of the position differences over a window of duration $t$.

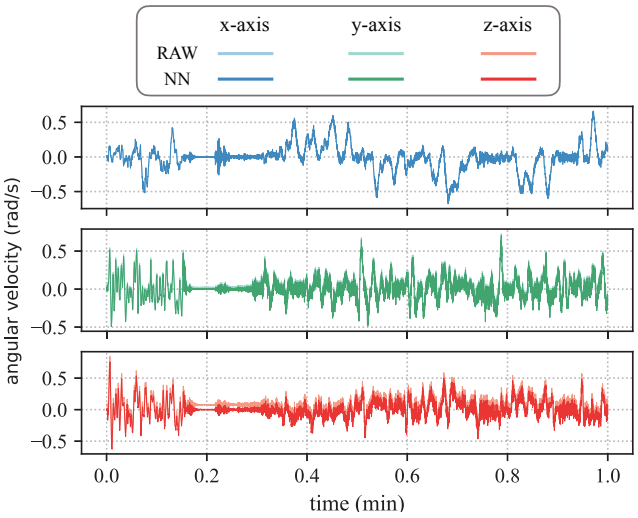


**Fig. 1** RAW and NN angular velocity of the MH_04_difficult.

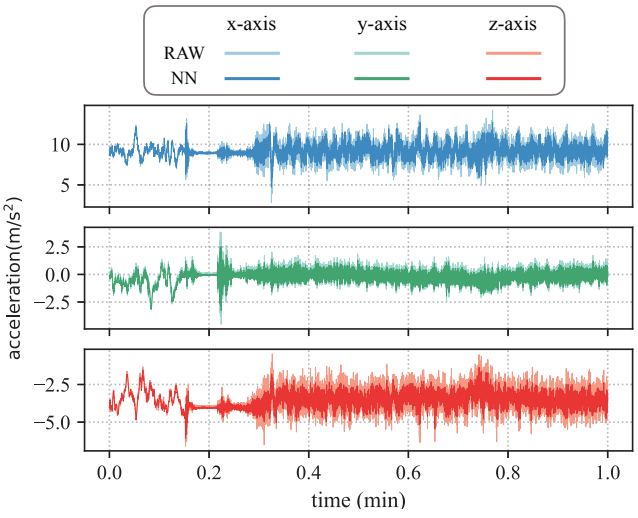


**Fig. 2** RAW and NN acceleration of the MH_04_difficult.

$$\mathrm{RTE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} ||\hat{p}_{i,t} - p_i||_2^2}$$

The RTE defined in this study differs from that used to evaluate the methods that directly output positions. The errors in velocity, orientation, and position are continuously accumulated when using the kinematic motion model for position estimation. Therefore, we correct the estimates with the ground truth every interval $t$ and calculate the position $\hat{p}_{i,t}$ for each frame to accurately measure the accumulated error over a specified time interval $t$.

## 5. Results

This section presents the findings of the detailed experimental evaluation of data-driven calibration methods on the effect of signals, velocity, orientation, and position estimation.

### 5.1 Evaluation on the Signal

The focus of this evaluation is to show how data-driven calibration methods affect the signal quality. As shown in Table 3, the data-driven gyroscope calibration method removes the biases more than 90% along three axes on all test sequences compared with RAW. Particularly, without using any reference data, it achieves almost the accuracy obtained by PB. The variance, on the other hand, are the same as the raw signals. This is due to the fact that the noise of gyroscope in EuRoC is small. Figure 1 illustrates the comparison of raw and data-driven-based calibrated signals for one test sequence. We note that the key to data-driven gyroscope calibration is to shift the signals and reduce the biases, which is regarded as one of the most important inertial errors in IMU.

Our data-driven accelerometer calibration, as shown in Table 4, also reduces biases on most test sequences by a factor 5 to 10. Furthermore, unlike data-driven gyroscope calibration, accelerometer calibration also decreases

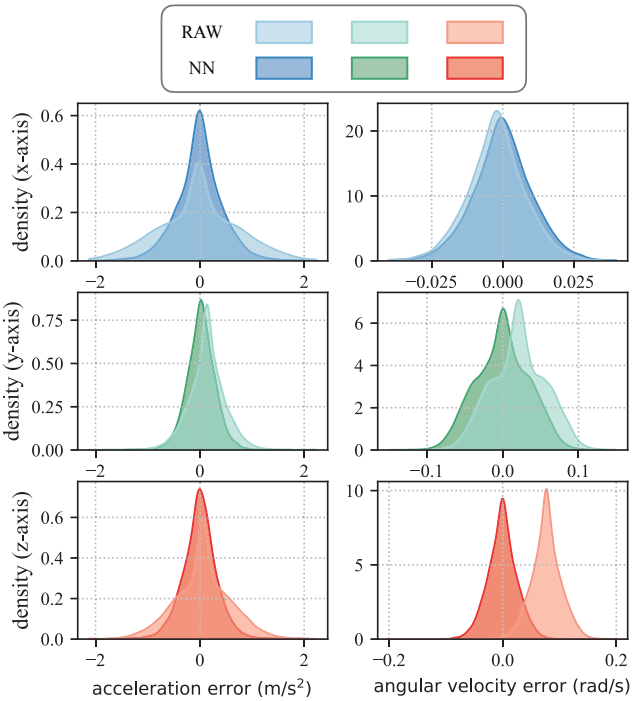


**Fig. 3** Signal error distribution of MH_04_difficult.

the noise by a factor 2 to 5. As shown in Figure 2 and Figure 3, the difference between NN acceleration and the derived acceleration has a higher likelihood of approaching zero compared to the raw signal. This indicates that the variance term in the loss function we designed is able to optimize the noise level.

However, we notice that the biases reduction on the y-axis acceleration of V1_01_easy was not noteworthy. The NN signals on the x-axis and y-axis acceleration of V2_02_medium have larger biases than RAW signals by a factor 2. We analyze this phenomenon by studying the provided biases of train and test sequences. As shown in Figure 4, we calculate the mean bias of each sequence. Then, we discovered that the acceleration bias of V1_01_easy in the y-axis is three times greater than other sequences. Similarly, the acceleration bias of V2_02_medium in the x-axis is larger,

**Table 3** Differences in term of **mean/variance** of the angular velocity ($\times 10^{-3}$ rad/s).

| | x-axis | | | y-axis | | | z-axis | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | RAW | NN | PB | RAW | NN | PB | RAW | NN | PB |
| 1 | 2.55/**0.10** | 0.18/**0.10** | **0.01/0.10** | -21.15/1.09 | **0.03/1.04** | 0.03/1.09 | -77.15/0.43 | 0.60/**0.42** | **0.004**/0.43 |
| 2 | 2.15/**0.14** | -0.08/0.15 | **0.01/0.14** | -21.04/1.43 | 0.05/**1.39** | **0.02**/1.44 | -76.65/0.75 | 0.61/**0.73** | **-0.004**/0.75 |
| 3 | 2.23/**0.15** | 0.12/0.16 | **0.01/0.15** | -20.75/2.05 | **0.02/1.95** | 0.05/2.05 | -76.28/0.48 | 0.19/**0.47** | **0.03**/0.48 |
| 4 | 2.29/**0.16** | **-0.01**/0.17 | -0.07/**0.16** | -21.70/5.82 | -1.39/**5.75** | **0.11**/5.82 | -76.55/**1.00** | -0.08/1.02 | **0.06/1.00** |
| 5 | 1.34/**0.21** | -0.71/0.23 | **-0.05/0.21** | -25.67/**1.54** | -2.61/1.66 | **0.09/1.54** | -78.90/**0.44** | -0.27/0.53 | **-0.03/0.44** |

**Table 4** Differences in terms of **mean/variance** of the acceleration ($\times 10^{-2}/\times 1$ $m/s^2$).

| | x-axis | | | y-axis | | | z-axis | | |
|---|---|---|---|---|---|---|---|---|---|
| No. | RAW | NN | PB | RAW | NN | PB | RAW | NN | PB |
| 1 | 2.43/0.66 | 0.08/**0.17** | **0.01**/0.66 | -14.45/0.10 | -1.49/**0.06** | **-0.003**/0.10 | -6.76/0.31 | 0.22/**0.12** | **0.01**/0.31 |
| 2 | 2.71/0.70 | 0.66/**0.32** | **-0.02**/0.70 | -13.73/0.13 | -2.37/**0.09** | **-0.02**/0.13 | -5.95/0.36 | **0.003/0.20** | 0.004/0.36 |
| 3 | 1.30/1.01 | -0.33/**0.22** | **-0.04**/1.01 | -54.94/0.20 | -43.87/**0.13** | **0.008**/0.20 | -6.82/0.67 | 0.39/**0.17** | **0.02**/0.67 |
| 4 | 2.33/2.79 | 2.29/**0.57** | **-0.04**/2.79 | -17.94/0.23 | -7.66/**0.14** | **0.006**/0.23 | -8.98/0.82 | -0.70/**0.17** | **-0.02**/0.82 |
| 5 | -0.58/0.88 | -1.07/**0.25** | **-0.03**/0.88 | -3.64/0.20 | 6.82/**0.13** | **-0.01**/0.20 | -9.46/0.51 | -3.64/**0.21** | **0.001**/0.51 |

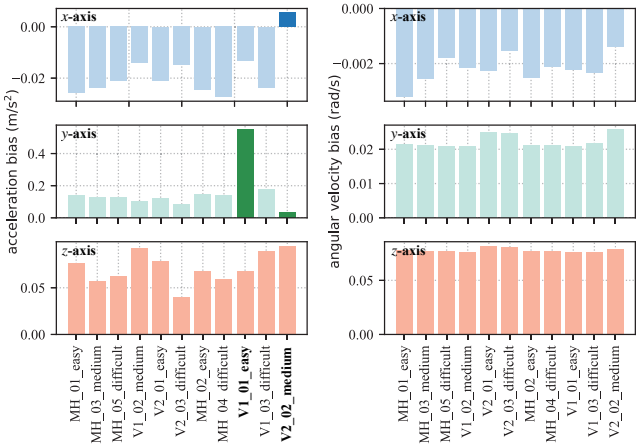


**Fig. 4** Mean biases of acceleration and angular velocity for each sequence, estimated by fusing ground truth measurements.

and the y-axis is a bit smaller than others. These can be considered outliers. It also indicates that the data-driven calibration methods are sensitive to outliers, i.e., low generalizability. Meanwhile, the biases of the angular velocity are more stable than the errors in acceleration among sequences in the EuRoC dataset. We consider that this is one of the critical reasons for the ability of the data-driven gyroscope calibration method to remove errors on the EuRoC dataset significantly.

Figure 3 shows the distribution of differences between RAW, NN signals and PGT signal of one test sequence. It is more intuitive to show the impact of the data-driven calibration on the signals. The density distribution shows not only the mean of the distribution but also the density of each error value. A higher density indicates a higher percentage of the corresponding value. Thus, the ideal distribution should have the highest density at 0. We note that both accelerometer and gyroscope calibration methods shift the mean value of the difference to 0. Additionally, the accelerometer calibration slightly increases the density at 0, making the error more probable near 0.

From the evaluation on the signal, we conclude that:

- Data-driven gyroscope calibration can eliminate the bias significantly, but does not effectively remove noise.
- Data-driven accelerometer calibration can reduce both noise and bias on most test sequences. However, it is sensitive to outliers.
- Data-driven gyroscope calibration is easier than accelerometer calibration because of the smaller norm. Additionally, for EuRoC dataset, it will be much easier because the gyroscope error is more stable.

### 5.2 Evaluation on the Velocity and Orientation

This evaluation focuses on determining how the calibrated signals improve the integration velocity and orientation. Table 5 shows the absolute error of velocities estimated from RAW, NN, and PB acceleration with GT orientation. Our proposed accelerometer calibration method reduces the AVE by a factor 2 to 10 except for the outlier sequence compared with RAW. Furthermore, removing the zero bias almost completely with the provided biases can significantly reduce the AVE to less than 1m/s. As illustrated in Figure 5, the inferred velocity from PB is difficult to distinguish from the ground truth velocities.

Data-driven gyroscope calibration method leads a huge improvement in terms of orientation estimation. As shown

**Table 5** Absolute velocity error (AVE) ($m/s$).

| | RAW | NN | PB |
|---|---|---|---|
| MH_02_easy | 11.66 | 0.84 | **0.13** |
| MH_04_difficult | 6.85 | 1.43 | **0.14** |
| V1_01_easy | 23.72 | 18.67 | **1.03** |
| V1_03_difficult | 4.86 | 1.74 | **1.10** |
| V2_02_medium | 7.20 | **3.16** | 5.35 |

**Table 6** Absolute Orientation Error (AOE) in terms of 3D orientation, in degree.

| | RAW | NN | PB |
|---|---|---|---|
| MH_02_easy | 127.72 | 2.99 | **0.69** |
| MH_04_difficult | 152.05 | 1.69 | **0.47** |
| V1_01_easy | 125.61 | **2.19** | 2.45 |
| V1_03_difficult | 136.08 | **1.52** | 2.04 |
| V2_02_medium | 126.61 | 3.70 | **1.69** |

Vol.2022-MBL-104 No.8
Vol.2022-UBI-75 No.8
Vol.2022-CDS-35 No.8
Vol.2022-ASD-24 No.8
2022/9/5

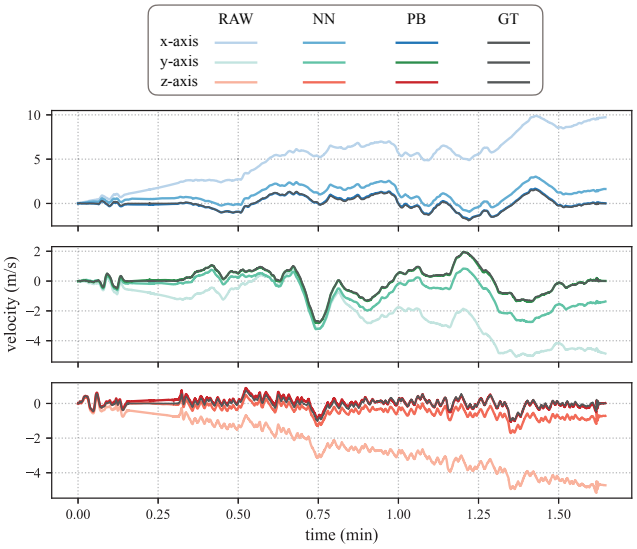


**Fig. 5** Velocity estimates on the sequence MH_04_difficult.

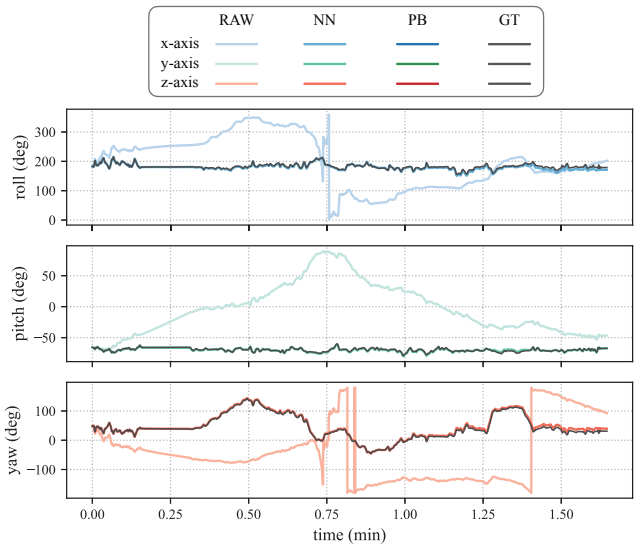


**Fig. 6** Orientation estimates on the sequence MH_04_difficult.

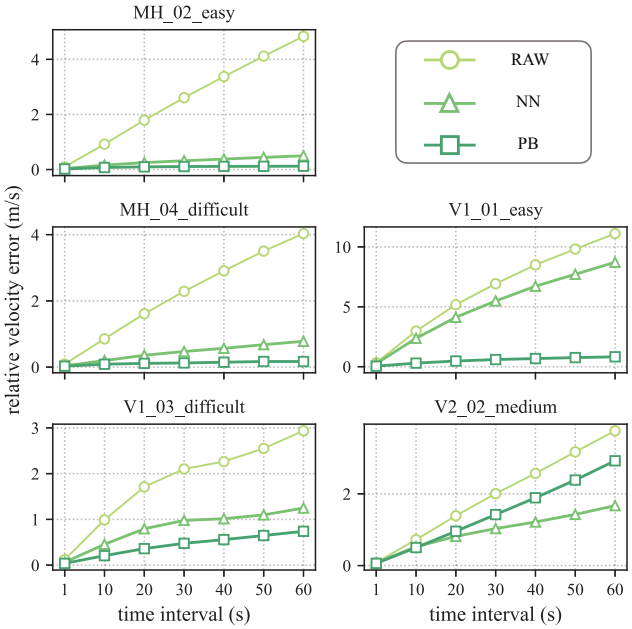


**Fig. 7** Relative velocity error (RVE).

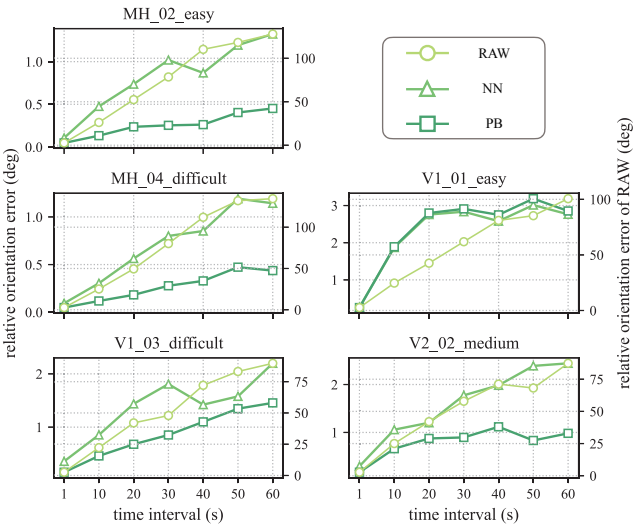


**Fig. 8** Relative orientation error (ROE).

in Table 6 and Figure 6, the orientation estimated from NN angular velocity achieves the same accuracy as from the PB angular velocity. This result is readily derived from the analysis of the impact on the signals, i.e., the signal performance of NN and PB is similar. Furthermore, we note that the orientation estimated from RAW signals completely unreliable. This further reinforces the necessity of calibrating the gyroscope for a low-cost MEMS IMU.

Results of RVE and ROE with different time intervals are given in Figure 7 and Figure 8. The velocity error gradually increases as the running time extends, and the error growth rate of NN and PB is smaller than that of the RAW velocity by a factor 2 to 4. The orientation error has the same changing trend as velocity error does, that is, gets larger as the time interval increases. However, in comparison with the velocity relative error, there are two differences. The first is that the orientation error accumulation rate of the RAW (an average of 50 degrees in 20s) is much faster than that of the velocity error accumulation rate. The second is that the improvement of the orientation estimation after calibration is larger, from an average of 50 degrees in 20s to 1 degree, and it performs well on all test sequences. These two differences support the importance of calibrating the gyroscope and the difficulty of calibrating the accelerometer, especially for the EuRoC dataset. Furthermore, without the correction by other reference information, the velocity and orientation error will accumulate at an approximately linear rate with time.

Based on the analysis we conclude that

- The data-driven calibration can effectively reduce the velocity error and orientation error.
- Data-driven gyroscope calibration provides far more improvement than that provided by accelerometer calibration. This demonstrates the necessity and primacy of calibration for the gyroscope.

Vol.2022-MBL-104 No.8
Vol.2022-UBI-75 No.8
Vol.2022-CDS-35 No.8
Vol.2022-ASD-24 No.8
2022/9/5

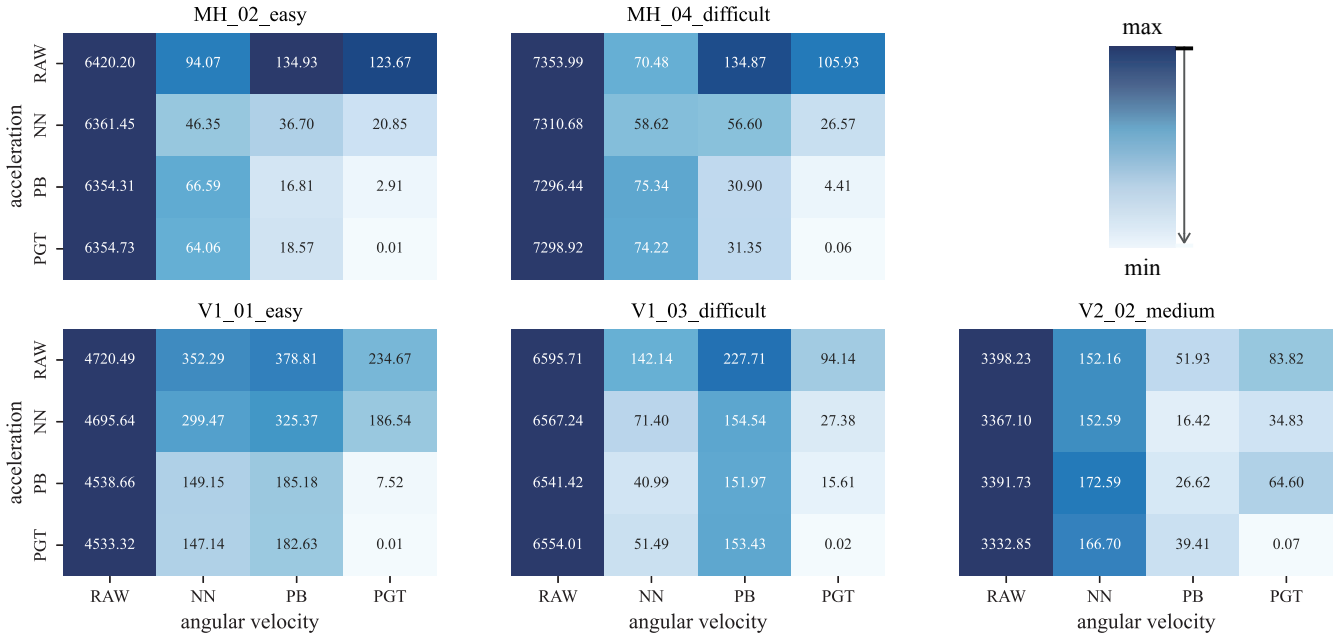


**Fig. 9** Absolute translation error (ATE).

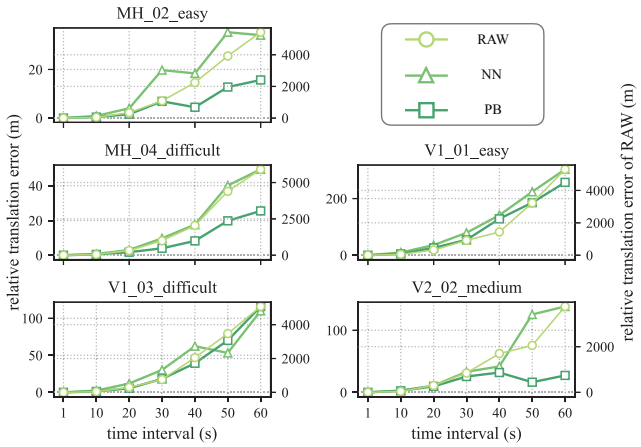


**Fig. 10** Relative translation error (RTE).

### 5.3 Evaluation on the Position

The focus of this evaluation is to measure the impact of data-driven calibrated signals on position estimation. Our findings are as follows.

First, data-driven calibrating accelerometer and gyroscope significantly reduces the position estimation error, from outrageous values, e.g., several kilometers per minute, to less than one hundred meters per minute. Nevertheless, there is still a twofold gap, on average, to the lowest error achieved by the fully compensated signal.

Second, the improvement in position estimation accuracy using data-drive calibrated acceleration increases with the enhancement of angular velocity accuracy. Each column of Figure 9 shows the effect of RAW, NN, and PB accelerations on the position estimation with the same angular velocity. We note that there is no improvement with RAW angular velocity even using PGT acceleration. However, with NN angular velocity, ATE in terms of NN acceleration is reduced by up to double compared with the RAW acceleration. Furthermore, with PGT angular velocity, NN acceleration reduces the ATE by a factor of up to 6. Combining the results in Section 5.2 we conclude that an effective error compensation for the gyroscope is most important. Not only because a slight correction can greatly improve the accuracy of the orientation estimation, but also because the data-driven calibrated acceleration can only improve the accuracy of the position estimation with a relatively accurate orientation.

Third, data-driven calibration methods considerably slow down the accumulation of position estimation errors. As shown in Figure 10, even though they have the same trend of error accumulation, there is a 40-200 times difference in the accumulated position errors with the same time interval. Furthermore, the difference of position estimation from the NN and the PB is no more than 5 meters when the accumulation time is within 30s. This demonstrates that the data-driven calibration methods are as accurate as using the near-complete denoised signal in terms of short-term position estimation.

## 6. Conclusion

In this paper, we propose a data-driven method for calibrating the accelerometer. This method is based on a dilated convolutional network. We design a mean-variance loss function to train the network by fully considering the sensor error model and the characteristics of different frequencies of real data. This loss function allows our method to reduce both bias and noise. Furthermore, exhaustive experiments in terms of data-driven gyroscope and accelerometer calibration methods are performed in this paper. It leads to interesting conclusions: (1) Data-driven gyroscope calibration is much easier than accelerometer calibration, especially for EuRoC dataset; (2) the improvement of odometry

results due to the calibrated accelerometer increases gradually with the increase of orientation accuracy; (3) data-driven calibration considerably slows down the rate of error accumulation; however, it is still far from achieving long-term inertial odometry; (4) data-driven calibration methods are not robust and sensitive to unknown data. According to the analysis of the evaluations, future work will improve the generalizability of the data-driven calibration method, especially for gyroscopes, making the trained model quickly adapt to the error coefficients of new IMUs.

## Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP20K11891.

### References

[1] Paneque, J., Martínez-de Dios, J. and Ollero, A.: Multi-Sensor 6-DoF Localization For Aerial Robots In Complex GNSS-Denied Environments, *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1978–1984 (2019).

[2] Yang, Y. and Huang, G.: Acoustic-Inertial Underwater Navigation, *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4927–4933 (2017).

[3] Bao, J., Li, D., Qiao, X. and Rauschenbach, T.: Integrated Navigation for Autonomous Underwater Vehicles in Aquaculture: A Review, *Information Processing in Agriculture*, Vol. 7, No. 1, pp. 139–151 (2020).

[4] Young, T. S., Teather, R. J. and MacKenzie, I. S.: An Arm-Mounted Inertial Controller for 6DOF Input: Design and Evaluation, *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 26–35 (2017).

[5] Kok, M., Hol, J. D. and Schön, T. B.: Using Inertial Sensors for Position and Orientation Estimation, *Foundations and Trends® in Signal Processing*, Vol. 11, No. 1-2, pp. 1–153 (2017).

[6] Collin, J., Davidson, P., Kirkko-Jaakkola, M. and Leppäkoski, H.: Inertial Sensors and Their Applications, *Handbook of Signal Processing Systems* (Bhattacharyya, S. S., Deprettere, E. F., Leupers, R. and Takala, J., eds.), Springer International Publishing, Cham, pp. 51–85 (2019).

[7] Savage, P. G.: Blazing Gyros: The Evolution of Strapdown Inertial Navigation Technology for Aircraft, *Journal of Guidance, Control, and Dynamics*, Vol. 36, No. 3, pp. 637–655 (2013).

[8] Reyes Leiva, K. M., Jaén-Vargas, M., Codina, B. and Serrano Olmedo, J. J.: Inertial Measurement Unit Sensors in Assistive Technologies for Visually Impaired People, a Review, *Sensors*, Vol. 21, No. 14, p. 4767 (2021).

[9] Ionut-Cristian, S. and Dan-Marius, D.: Using Inertial Sensors to Determine Head Motion—A Review, *Journal of Imaging*, Vol. 7, No. 12, p. 265 (2021).

[10] Zhao, J.: A Review of Wearable IMU (Inertial-Measurement-Unit)-Based Pose Estimation and Drift Reduction Technologies, *Journal of Physics: Conference Series*, Vol. 1087, p. 042003 (2018).

[11] Herath, S., Yan, H. and Furukawa, Y.: RoNIN: Robust Neural Inertial Navigation in the Wild: Benchmark, Evaluations, Amp; New Methods, *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3146–3152 (2020).

[12] Lu, C., Uchiyama, H., Thomas, D., Shimada, A. and Taniguchi, R.-i.: Indoor Positioning System Based on Chest-Mounted IMU, *Sensors*, Vol. 19, No. 2, p. 420 (2019).

[13] Fang, B., Chou, W. and Ding, L.: An Optimal Calibration Method for a MEMS Inertial Measurement Unit, *International Journal of Advanced Robotic Systems*, Vol. 11, No. 2, p. 14 (2014).

[14] Brossard, M., Bonnabel, S. and Barrau, A.: Denoising IMU Gyroscopes With Deep Learning for Open-Loop Attitude Estimation, *IEEE Robotics and Automation Letters*, Vol. 5, No. 3, pp. 4796–4803 (2020).

[15] Huang, F., Wang, Z., Xing, L. and Gao, C.: A MEMS IMU Gyroscope Calibration Method Based on Deep Learning, *IEEE Transactions on Instrumentation and Measurement*, Vol. 71, pp. 1–9 (2022).

[16] Engelsman, D. and Klein, I.: Data-Driven Denoising of Accelerometer Signals (2022).

[17] Frosio, I., Pedersini, F. and Borghese, N. A.: Autocalibration of MEMS Accelerometers, *IEEE Transactions on Instrumentation and Measurement*, Vol. 58, No. 6, pp. 2034–2041 (2009).

[18] Lu, X., Liu, Z. and He, J.: Maximum Likelihood Approach for Low-Cost MEMS Triaxial Accelerometer Calibration, *2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pp. 179–182 (2016).

[19] Ru, X., Gu, N., Shang, H. and Zhang, H.: MEMS Inertial Sensor Calibration Technology: Current Status and Future Trends, *Micromachines*, Vol. 13, No. 6, p. 879 (2022).

[20] Sun, S., Melamed, D. and Kitani, K.: IDOL: Inertial Deep Orientation-Estimation and Localization, *AAAI* (2021).

[21] Silva do Monte Lima, J. P., Uchiyama, H. and Taniguchi, R.-i.: End-to-End Learning Framework for IMU-Based 6-DOF Odometry, *Sensors*, Vol. 19, No. 17, p. 3777 (2019).

[22] Chen, C., Lu, C. X., Wahlström, J., Markham, A. and Trigoni, N.: Deep Neural Network Based Inertial Odometry Using Low-Cost Inertial Measurement Units, *IEEE Transactions on Mobile Computing*, Vol. 20, No. 4, pp. 1351–1364 (2021).

[23] Dugne-Hennequin, Q. A., Uchiyama, H. and Paulo Silva Do Monte Lima, J.: Understanding the Behavior of Data-Driven Inertial Odometry With Kinematics-Mimicking Deep Neural Network, *IEEE Access*, Vol. 9, pp. 36589–36619 (2021).

[24] Liu, W., Caruso, D., Ilg, E., Dong, J., Mourikis, A. I., Daniilidis, K., Kumar, V. and Engel, J.: TLIO: Tight Learned Inertial Odometry, *IEEE Robotics and Automation Letters*, Vol. 5, No. 4, pp. 5653–5660 (2020).

[25] Jiang, C., Chen, S., Chen, Y., Zhang, B., Feng, Z., Zhou, H. and Bo, Y.: A MEMS IMU De-Noising Method Using Long Short Term Memory Recurrent Neural Networks (LSTM-RNN), *Sensors*, Vol. 18, No. 10, p. 3470 (2018).

[26] Yao, F., Wu, Z., Wei, Z. and Wang, D.: Few-Shot Domain Adaptation for IMU Denoising (2022).

[27] Chen, H., Aggarwal, P., Taha, T. M. and Chodavarapu, V. P.: Improving Inertial Sensor by Reducing Errors Using Deep Learning Methodology, *NAECON 2018 - IEEE National Aerospace and Electronics Conference*, pp. 197–202 (2018).

[28] Li, R., Fu, C., Yi, W. and Yi, X.: Calib-Net: Calibrating the Low-Cost IMU via Deep Convolutional Neural Network, *Frontiers in Robotics and AI*, Vol. 8 (2022).

[29] Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W. and Siegwart, R.: The EuRoC Micro Aerial Vehicle Datasets, *The International Journal of Robotics Research*, Vol. 35, No. 10, pp. 1157–1163 (2016).