TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 09/15 pp692–705
DOI: 10.26599/TST.2021.9010017
Volume 26, Number 5, October 2021

Inertial Motion Tracking on Mobile and Wearable Devices: Recent Advancements and Challenges

Zhipeng Song, Zhichao Cao, Zhenjiang Li, Jiliang Wang*, and Yunhao Liu

Abstract: Motion tracking via Inertial Measurement Units (IMUs) on mobile and wearable devices has attracted significant interest in recent years. High-accuracy IMU-tracking can be applied in various applications, such as indoor navigation, gesture recognition, text input, etc. Many efforts have been devoted to improving IMU-based motion tracking in the last two decades, from early calibration techniques on ships or airplanes, to recent arm motion models used on wearable smart devices. In this paper, we present a comprehensive survey on IMU-tracking techniques on mobile and wearable devices. We also reveal the key challenges in IMU-based motion tracking on mobile and wearable devices and possible directions to address these challenges.

Key words: Inertial Measurement Units (IMUs); motion tracking; accelerometer; gyroscope; magnetometer; sensor fusion

1 Introduction

Inertial motion tracking, which uses Inertial Measurement Units (IMUs) or inertial sensors for motion tracking, has been a hot research topic for the last two decades. Typical inertial sensors, such as accelerometers, gyroscopes, and magnetometers, are equipped in most mobile and wearable devices and can provide different types of sensing information. The goal of inertial motion tracking is to use the sensing information provided by these sensors to derive three-dimensional (3D) motion, which can be further used in different applications, such as indoor navigation, gesture recognition, and text input. For example, a health center can utilize a smart wristband to study patients' physical motion during their rehabilitation[1]. Inertial tracking can also be used for gesture recognition to control appliances in smart home scenarios[2]. The tracking-based text-input method can offer a more convenient way to perform text input on many small-screen or even no-screen devices, such as smartwatches and wristbands[3]. With the prevalence of wearable smart devices, including but not limited to smartphones, smartwatches, and wristbands, inertial motion tracking will play a more critical role in wearable and mobile applications in the future. IMU tracking has its advantages. Unlike tracking based on wireless signals, such as Wi-Fi[4] and UWB[5], IMU tracking does not rely on extra infrastructure (e.g., base station and router). Moreover, recent work on the embedded terminals[6, 7] can both secure sensitive data transmitted from IMU sensors and save battery energy of wearable devices. The information provided by IMU sensors will facilitate more applications in different scenarios, such as healthcare, smart home, and human-computer interaction.

In this survey, we focus on 3D IMU tracking where

• Zhipeng Song and Jiliang Wang are with the School of Software, Tsinghua University, Beijing 100084, China. E-mail: songzp18@mails.tsinghua.edu.cn; jiliangwang@tsinghua.edu.cn.
• Zhichao Cao is with the Department of Computer Science and Engineering, Michigan State University, Michigan, MI 48824, USA. E-mail: caozc@msu.edu.
• Zhenjiang Li is with the Department of Computer Science, City University of Hong Kong, Hong Kong 999077, China. E-mail: zhenjiang.li@cityu.edu.hk.
• Yunhao Liu is with the Department of Automation and the Global Innovation eXchange Institute (GIX), Tsinghua University, Beijing 100084, China. E-mail: yunhao@tsinghua.edu.cn.
* To whom correspondence should be addressed.
Manuscript received: 2021-02-07; accepted: 2021-02-26

a user holds or wears smart devices with different types of IMUs on her/his hand or wrist. We survey works that focus on tracking fine-grained movements performed by fingers, hands, wrists, and arms, and conduct a comprehensive study of inertial motion tracking approaches.

We consider inertial motion tracking with an accelerometer, gyroscope, and magnetometer. These sensors measure the device's acceleration, angular rate, and amplitude of the nearby magnetic field.

Initially, a simple Two-Dimensional (2D) motion tracking can be achieved based on the accelerometer's acceleration on a device. We integrate acceleration over its time to derive the velocity, and further integrate velocity to derive the displacement. Consequently, the displacement can be used to track the device's position.

This intuitive double-integration method can only work well under several constraints: (1) The device's orientation keeps unchanged, that is, the device has no rotations. (2) Due to the error in direction, most acceleration-based tracking methods only work well for 2D planes. (3) The tracking duration should be short[8]; otherwise, the accumulated error can be very high. If a device rotates during the tracking period, then the acceleration cannot directly represent the movement of the device anymore. It then requires more information from the gyroscope and magnetometer, which will be further elaborated in Section 2.1. If the movement is along Z-axis, then the corresponding acceleration can be impacted by gravity, which can be a non-negligible noise.

Therefore, typical IMU-tracking methods can combine information from the gyroscope, magnetometer, and other available sensors to address the limitations mentioned above. The angular rate from the gyroscope can be used to derive the rotation angle and rotation matrix, which rotates sensor readings from the Local Reference Frame (LRF) (see Fig. 1) to the Global Reference Frame (GRF). The magnetometer can provide the magnetic north in the LRF, served as an "anchor" to calibrate orientation.

Fig. 1 LRF of IMU sensors in a smartphone and a smartwatch.

For mobile and wearable devices, IMUs on these devices are typically made lighter, smaller, and inexpensive, and have lower power consumption, enabling them to be embedded into smartphones, smartwatches, and virtual reality headsets. These low-cost sensors may have lower precision compared to those in airplanes or robotics. However, the acceleration significantly varies in practice. The relatively low sampling rate (100 Hz in most smartwatches nowadays) is insufficient to capture instant movements, such as sharp turns and frequent flips.

Numerous research works have been conducted to make IMU tracking feasible on modern wearable devices. Based on different requirements and scenarios, we categorize them into three main groups, i.e., physical model based methods, arm motion model based methods, and activity model based methods.

The physical model-based methods leverage sensor fusion techniques and sensor error modeling to calibrate errors in sensor readings. The arm motion model based methods use kinetic knowledge to model the arm movements of humans, including hand or wrist movements in a particular region. The activity model based methods leverage different features to recognize different activities, such as gestures, text, and other movements, including sign language and smoking.

Based on IMU tracking, many applications have been proposed, which can be further categorized into body tracking and wrist tracking. The former normally focuses on tracking the human body's position, whereas the latter focuses on tracking wrist movements. The two categories have several main differences:

- **Granularity:** Considering the indoor localization scenario, body tracking will allow meter-level errors, whereas wrist tracking has finer granularity and generally requires decimeter- or centimeter-level errors.
- **Periodicity:** Most body-tracking works are based on the human's walking or running pattern, which tends to be periodical. This periodicity can be helpful in refining sensor data. Conversely, wrist-tracking movements can be more complicated.
- **Model:** Body tracking and wrist tracking focus on different parts of humans and use different models. The stride length model is used in body tracking; whereas, the arm motion model is used in wrist tracking.

Many research works on IMU body tracking are

summarized in Ref. [9]. Body tracking includes pedestrian localization, heading estimation, step counting, and activity recognition[10]. This paper will focus on wrist tracking.

This survey is organized as follows: Section 2 introduces IMU-tracking frameworks based on the three models mentioned above. Section 3 analyzes noises in IMU sensor readings. Sections 4, 5, and 6 elaborate on the three models (i.e., physical model, arm motion model, and activity model, respectively). Section 7 summarizes the challenges and shows possible solutions and research directions for IMU-tracking works. Finally, Section 8 concludes our survey.

2 IMU-Tracking Framework

Figure 2 presents the general IMU-tracking framework. First, we collect readings from three IMU sensors, namely, acceleration, magnetic north, and angular rate. These data are then preprocessed to obtain fine-grained motion information, such as linear acceleration, initial orientation, and rotation matrix. Three different models are applied with different techniques to process motion data and derive the tracking results. Tracking results can be 3D locations, orientations, or certain recognized activities (e.g., gestures and text). We will briefly introduce each part in this section and present more details in the next sections.

2.1 IMU sensors in motion tracking

Typical IMUs in modern mobile devices consist of three components: accelerometer, gyroscope, and magnetometer. The accelerometer measures the device's acceleration along three axes caused by all external forces applied to it. The device's rotation angles around three axes are captured by the gyroscope. The magnetometer measures the distribution of the magnetic field.

Ideally, three axes (noted as X-, Y-, and Z-axes) of the three sensors are considered to be perfectly aligned. When we describe sensor readings, such as "acceleration along the X-axis", "angular velocity around the X-axis", and "magnetic amplitude along the X-axis", all three X-axes refer to the same direction. However, in a real device, these sensors may have some misalignment due to manufacturing imperfection[8]. Different approaches have been proposed to remove this misalignment (e.g., in the physical and arm motion models) or compensate for the misalignment in the tracking results (e.g., in the activity models). Therefore, we usually assume that the three axes are properly aligned.

For motion tracking, we need to identify the reference frame clearly. In practice, there are many reference frames for tracking via IMU. Typically, there are four kinds of frames[8]: body frame, navigation frame, inertial frame, and earth frame. The body frame is the frame defined by the device, e.g., sensor readings. The navigation frame is a global and absolute frame. In this study, we are interested in tracking traces in the navigation frame. The inertial and earth frames are introduced considering the centrifugal acceleration and Coriolis acceleration due to the earth's self-rotation. In a wearable sensing scenario, the earth's rotation can be

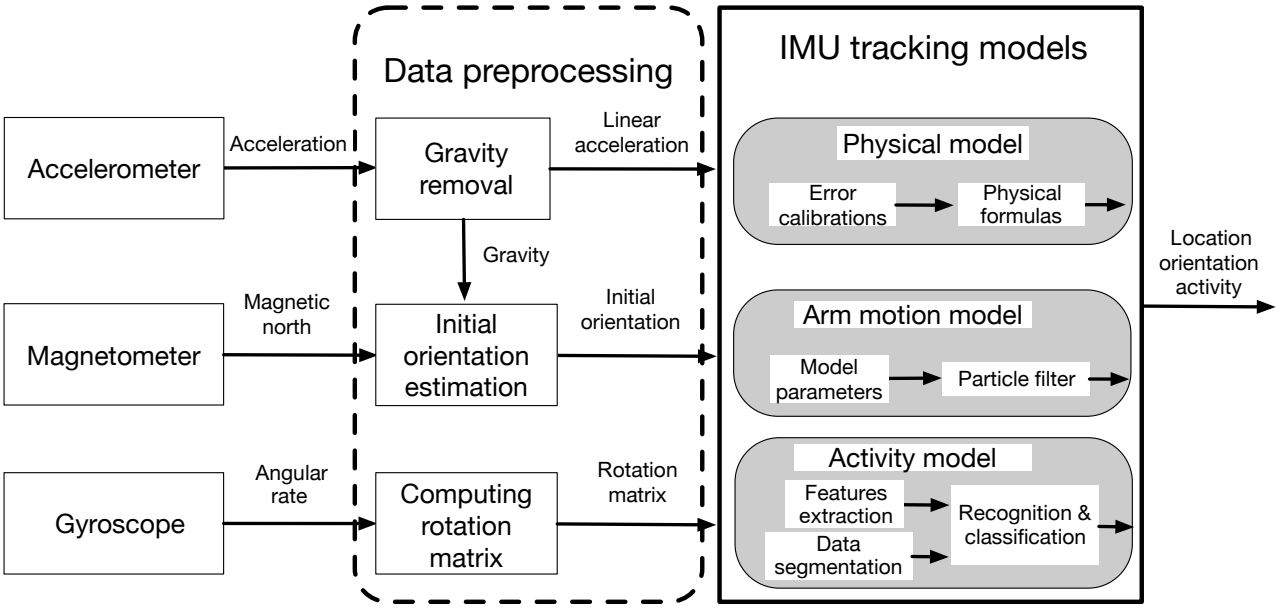


Fig. 2 General framework of IMU-tracking.

negligible, as the device's moving range and velocity are limited compared with those of the earth. Hence, the earth's rotation should be considered for navigating a plane or ship that travels either in high speed or long range. For inertial motion tracking on mobile and wearable devices, we typically consider the GRF and LRF[11], which correspond to the navigation and body frame mentioned above, respectively. The GRF is defined as ⟨ *North, East, Up* ⟩.

All the three inertial sensors work in their LRFs. The same readings of IMU sensors can be different in GRF given different orientations or attitudes of a device. Therefore, we need to distinguish the LRF and GRF in tracking, as shown in Fig. 3. Sensor readings can be transformed from the LRF into the GRF through a rotation matrix.

The goal of IMU tracking is to derive the tracking traces in GRF, and the output of IMUs is in the LRF. For example, if an accelerometer has readings on the X-axis, then this acceleration is along the device's X-axis, not the X-axis in the global framework. Therefore, we need to transfer three sensors' readings from the LRF to the GRF to obtain correct tracking results.

The LRF of a smartphone and a smartwatch is illustrated in Fig. 1, where the Z-axis is perpendicular to the surface of the device and points from back to front.

2.2 Data preprocessing

After we obtain readings from the three sensors, we can further preprocess the data. The sampling rates of different IMUs are not necessarily the same. For example, for the Huawei P30 smartphone, the highest sampling rates of the accelerometer and gyroscope are 500 Hz, and the sampling rate of the magnetometer is 100 Hz. The difference among the sampling rates implies that samplings from different sensors may not be synchronized. To manually synchronize the three sensors, a usually adopted method is to apply interpolation to the data[8, 12] with available timestamps of each sensor reading. To make a better measurement of the acceleration and capture high-frequency features, ViBand[13] developed a custom smartwatch whose sampling rate increases to 4 kHz by modifying the kernel of the operating system. The accelerometer's sampling rate is 100 Hz in current off-the-shelf smartwatches (e.g., Samsung LG Sport and Huawei Watch 2), and smartphones can reach as high as 500 Hz (e.g., Huawei P30 Pro). However, these rates are much lower than 4 kHz. ViBand[13] can leverage high-fidelity bio-acoustic features to make a wide range of applications, providing a new direction for IMU tracking.

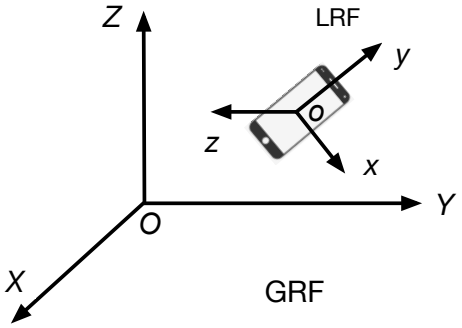


Fig. 3 GRF and LRF.

2.2.1 Gravity removal

Accelerometer readings are polluted by gravitational acceleration by default, so we need to apply gravity removal to separate gravity from raw acceleration data. One method is to apply a low-pass filter on acceleration data[14]. For instance, Android provides an Application Programming Interface (API) for removing gravity[15].

2.2.2 Initial orientation estimation

The orientation here describes the device's facing direction concerning the GRF, which determines the rotation matrix from the LRF to the GRF. Initial orientation requires "global anchors" to obtain, and a global anchor means that its measurement keeps unchanged in the GRF, regardless of how the LRF changes. The magnetic north direction provided by the magnetometer and the gravity from the accelerometer can serve as global anchors. These two directions can be used to compute initial orientation, e.g., the methods in Refs. [11, 16] can estimate the initial orientation.

2.2.3 Computing rotation matrix

When the device rotates, its orientation changes over time. Thus, we need to determine the initial orientation and orientation's change between every two samples. Angular rates from the gyroscope can be used to compute the rotation matrix, and thus the orientation changes. The works in Refs. [8, 12] present the details of the representation of 3D rotations and how to compute the rotation matrix.

2.3 Tracking models

After processing the data, we can apply the data to different models for motion tracking. We summarize three tracking models: physical model, arm motion model, and activity model. Section 4 presents the physical model, which computes the orientation and

position according to the physical motion relationship. Section 5 presents the arm tracking model, which leverages the body constraints for tracking. Section 6 presents the activity model, which focuses on recognizing certain activities, such as gestures and text inputs.

3 Noises in IMU Data

Similar to other sensors, IMU sensors suffer from many kinds of noises. The use of three different sensors further makes noise removal more complicated. This section will introduce typical IMU sensor noises and discuss why IMU tracking is nontrivial.

3.1 Inherent noises

Inherent noises or hardware noises exist in all kinds of sensors. The inherent noise on IMU sensors has been extensively studied in the past decade[8, 12, 17]. For the completeness of this survey, we follow definitions in Refs. [8, 12, 17] and present a brief introduction of different sources of inherent noises.

From the aspect of effect on IMU sensors, there are three types of noises: bias, scale factor, and nonorthogonality misalignment.

3.1.1 Bias

When no external input is fed to a sensor, it is expected to output zero; otherwise, this nonzero output is called a bias. The bias consists of deterministic and random parts. Reference [8] found that when the sensor is static, the expected output of the accelerometer and gyroscope is nonzero. Furthermore, the output is not constant but varies along with time.

3.1.2 Scale factor

The scale factor is defined as the ratio of changes of the output to that of the input. Because MicroElectroMechanical System (MEMS) sensors usually transfer voltage into the corresponding measurement, they need to scale the measured voltage to derive acceleration (angular rate or magnetic amplitude), which may introduce this kind of noise.

3.1.3 Nonorthogonality misalignment

Due to the manufacturing imperfection, the three axes of a sensor may not perfectly align those of the device. Therefore, a misalignment may exist between sensors' and device's coordinate frames.

Apart from the noises mentioned above, IMU sensors also suffer from saturated measurements, i.e., when the measured physical quantity is too large for the sensors, the measured voltage may be clipped or scaled to fit in the range of measurement, thus causing an error. The gyroscope can be sensitive to accelerations along the three axes as if the accelerometer and gyroscope are coupled, which is called a g-dependent bias[18]. Moreover, the magnetometer is prone to hard- and soft-iron effects[19].

3.2 Environmental noise

Besides inherent noises, IMU sensors can also be affected by environmental factors, such as gravity, self-rotation of earth, magnetic interference, and temperature. These environmental factors also play important roles in the noises of IMU sensors.

3.2.1 Gravity

When an accelerometer measures acceleration, it takes all kinds of external forces applied on the device into consideration, including gravity if the sensor is near the earth, which is the typical case. Therefore, gravitational acceleration is a natural noise for the accelerometer. The amplitude of gravitational acceleration (9.81 m/s^2) can be at the same magnitude with that from the device's movement, e.g., while writing or performing some gestures (normally does not exceed 15 m/s^2[20]). If the gravitational acceleration is not thoroughly removed, then it can significantly introduce noise to tracking. Many works have performed gravity removal by first calculating the IMU orientation, projecting acceleration in the LRF into the GRF, and directly subtracting gravity in the GRF, because the amplitude and direction of gravity in the GRF are kept constant regardless of how the device moves. The performance of this method heavily relies on the estimation precision of the device's orientation. In the GRF, gravitational acceleration is generally 9.81 m/s^2 and points to the negative direction of the Z-axis. The state-of-the-art research work on orientation estimation[11] has a median error of 10° within a 5-min measurement. Thus, part of the gravitational acceleration is projected into the X-Y plane due to the orientation estimation error, and this part can be as large as 1.7 m/s^2. Compared with the inherent noise of the accelerometer, such as constant bias (in the magnitude of 10^{-2} m/s^2 [21]), the error from gravity is considerably large.

3.2.2 Self-rotation of the earth

The self-rotation rate of the earth is approximately 7.29×10^{-5} rad/s, which can be captured by a gyroscope. The amplitude of the earth's rotation rate is very small,

which typically impacts the performance of lengthy and time-consuming tracking[22, 23].

3.2.3 Magnetic interference

In motion tracking, the magnetometer is often used as a compass. That is, it determines the magnetic north direction in the LRF, assuming there are no ferromagnetic materials nearby. In many indoor environments, such as offices, classrooms, laboratories, and living rooms, where computers and cables are everywhere, the magnetometer is exposed to interferences, and its readings can be irregular and unstable. Some research works point out that the magnetometer has a high accuracy for outdoor areas, but its performance degrades in the indoor environment. The experiments show that the compass output can cause an error of 50° for the measurement of half a minute[16]. Reference [11] tested its attitude-estimating system in different environments, such as outdoor and indoor open places, crowded engineering buildings, and laboratories. The results show that the orientation error can be very high, e.g., when the fluctuation (standard deviation) of the magnetic field density reaches 13 μT (which is common in houses with computers and cables), the orientation error can be as high as 100°.

3.2.4 Temperature

All the three IMU sensors are sensitive to temperatures, and the factors of inherent noises can vary in different temperatures[24].

3.3 Summary

In this section, we summarize some properties of IMU sensor noises:

(1) There are many types of IMU sensor noises

On the top level, IMU sensor noises can be divided into two categories: inherent noise and environmental noise. Inherent noises exist in various sensors due to hardware imperfection, including bias, scaling factors, and nonorthogonality misalignment. Environmental noise comes from the earth and surroundings. For example, gravity impacts the accelerometer, the self-rotation of the earth impacts the gyroscope, and nearby ferromagnetic materials impact the magnetometer. Some noises can be further categorized as constant or random according to their statistical properties.

(2) The same type of noise may have different effects on tracking

Take constant bias as an example. For the accelerometer, this acceleration bias will be integrated twice to calculate displacement. For the gyroscope, the angular velocity needs to be integrated once to get the angular shift, whereas the magnetic amplitude does not need integration. Therefore, the accelerometer's constant bias is considered to have more impact on the tracking, because the double-integration process makes the position error grow quadratically with time. Reference [8] conducted an experiment on a smartphone and found that even when the phone is kept stationary, the computed position can drift several meters within 10 s due to noise. Constant bias only makes the angular error grow linearly, and the magnetometer's error will not accumulate with time.

(3) Noises can be time-varying

The noises presented above can change over time. Even the constant bias can only be considered "constant" in a short period, and it may drift in an extended range of time. Reference [8] collected a dataset for 55 min of a static gyroscope and found that this sensor's bias in the first minute is different from that of the last minute. Furthermore, the bias of an axis could even change its sign. The instability or time-varying property of these sensor noises makes offline calibration techniques ineffective and exacerbates the complexity of modeling these noises.

4 Physical Model Approach

The physical model investigates the physical relationship among sensor data and computes the device's orientation and location. Generally, a physical model consists of a series of physical formulas, which take IMU sensor readings as the input and output of the device's orientation and location. References [8, 12] presented an introduction of inertial navigation and classic physical models.

However, it is impracticable to directly use raw data and perform computations, because **raw sensor data are full of noises and need fine-grained calibrations**. Thus, all physical model approaches need calibration methods to improve the data quality.

This section will introduce a series of calibration-based IMU-tracking research works, which can be classified into two categories: equipment-aided and filter-based calibration.

4.1 Equipment-aided calibration

Equipment-aided calibration first uses special equipment to measure the IMU sensor noise parameters and then applies the measured noise parameters into the physical

model to calibrate IMU tracking.

The special equipment mentioned above can be classified into two kinds: electric slide rail and turntable. We can manually set the moving direction and speed of a slide rail or change the 3D orientation of a turntable. Therefore, when we mount the device with IMU sensors on the special equipment, we can determine the tracking ground truth of this device. We denote the ground-truth 3D position of this device as (Y_x, Y_y, Y_z). Meanwhile, we derive IMU sensor readings from the device, and through the physical model, we can also compute the device's position (X_x, X_y, X_z). Basically, there is a difference between (Y_x, Y_y, Y_z) and (X_x, X_y, X_z) because the physical model contains IMU sensor noises.

To calibrate the physical model, we introduce a typical method to measure IMU sensor noises. This method combines three kinds of noise parameters (e.g., bias, scale factor, and nonorthogonality misalignment) and describes their relationships in the following equation:

$$\begin{bmatrix} Y_x \\ Y_y \\ Y_z \end{bmatrix} = \begin{bmatrix} S_x & M_{xy} & M_{xz} \\ M_{yx} & S_y & M_{yz} \\ M_{zx} & M_{zy} & S_z \end{bmatrix} \begin{bmatrix} X_x \\ X_y \\ X_z \end{bmatrix} + \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} + \begin{bmatrix} \eta_x \\ \eta_y \\ \eta_z \end{bmatrix} \tag{1}$$

where B_x, B_y, and B_z are bias noises, S_x, S_y, and S_z are scale factors, M_{xy}, M_{xz}, M_{yx}, M_{yz}, M_{zx}, and M_{zy} are misalignment coefficients, and η_x, η_y, and η_z are random noises.

Theoretically, we can place IMU sensors in many different positions or movement states to solve unknown noise parameters, as presented in Eq. (1). This scheme is called the multi-position calibration technique[23]. We feed Eq. (1) with (Y_x, Y_y, Y_z) and (X_x, X_y, X_z), and then the noise parameters can be estimated. During the IMU tracking, we can use these estimated noise parameters to calibrate our physical model.

However, equipment-aided calibration has two drawbacks: (1) The calibration is complicated and requires special high-precision equipment. Thus, it is not easy to use in practice for wearable devices. (2) As mentioned in Section 3.1, some inherent noises can be time-varying; thus, calibrating IMU sensors for once is not sufficient.

Therefore, equipment-aided calibration is not practical in practice for smart devices that require plug-and-play usage.

4.2 Filter-based calibration

Filter-based calibration is a popular method in spacecraft attitude determination[25] and has been used in MEMS IMU sensors on mobile devices. It uses different filters, including the Kalman Filter (KF), Extended Kalman Filter (EKF), Unscented Kalman Filter (UKF), and Complimentary Filter (CF), for calibration.

Besides the classic filters mentioned above, some research works have developed their own filters. Reference [16] proposed an opportunistic replacement filter, which determines whether a gyroscope, accelerometer, or magnetometer is more reliable when estimating the orientation and replaces unreliable data with reliable data. The Android platform[15] leverages gravity and magnetic north to filter out noises of the gyroscope, and this filter is provided as an API *getRotationMatrix*.

We summarize filter-based calibration works in Table 1 and list their respective assumptions and limitations.

As stated in Table 1, filter-based calibration has its limitations. Many works are feasible because they assume the random noise to be Gaussian. Further optimizations, such as maximum likelihood, are based on the Gaussian noise assumption. In practice, when the device is stationary, IMU sensors' noises can be considered Gaussian noise[8]. However, no observation or experiment has proven that the assumption still holds when the device moves very fast or makes complicated movements.

5 Arm Motion Model Approach

Arm motion model or kinematic model based IMU tracking has gained its popularity in recent years[11, 49–52]. It leverages the property of humans' biological characteristics and models the relationship among the shoulder, arm, and wrist during their movements for wearable sensing. This model can also be applied to control robots to imitate human arms' motions[53].

After building an arm motion model for a specific user, we can track the device worn by this user with the help of this model. Specifically, we make some assumptions on the initial positions and orientations of the device. Then, as the device moves, we compute the movements based on the IMU sensor readings. Our arm motion model puts a constraint on the possible movements, and it can rule out incorrect assumptions to determine the correct one. Therefore, the device can be correctly tracked.

Table 1 Related works of filter-based calibration on IMU tracking.

Related work	Filter type	Assumption / limitation
[15]	Gravity and magnetic north	Simple algorithm to reduce computational complexity at the cost of low accuracy
[16]	Opportunistic replacement	Require nearby magnetic field to be stable; require frequent pause to reset the system with gravity
[21]	KF	Require velocity readings as input
[26, 27]	KF	Rely on stable magnetic field or other sensors to get inclination angles
[28]	KF, gradient descent	Require stable nearby magnetic field; assume initial orientation is known
[29]	KF	Need redundant IMU sensors
[30, 31]	KF	Require stable nearby magnetic field; devices are stationary or in low acceleration consitions
[32]	KF	Rotational motion of devices is slow
[33]	KF	Estimate heading direction rather than 3D orientation
[11, 34]	KF[34], CF[11]	Require stable nearby magnetic field
[35]	EKF	Assume gyroscope's bias is constant; require stable nearby magnetic field
[36, 37]	EKF	Assume the device is stationary
[38]	EKF	Need stationary condition to do initialization; require stable nearby magnetic field
[39]	UKF	Require devices to be stationary or move slowly
[40]	UKF	Require GPS as reference
[41–46]	EKF[41–43, 46], UKF[44], CF[45, 46]	Require magnetic measurement to be accurate and the device to move slowly so that the gravity can be extracted
[47, 48]	CF[47], gradient descent[48]	Require devices to move slowly, so that it can leverage gravity to compensate for magnetic distortion

To leverage arm motion model constraints and rule out incorrect assumptions, many research works have used particle filters[11, 49–52]. Particle filter[54] has been proposed for more than two decades, and it can be used in many nonlinear estimation problems.

The particle filter algorithm starts from a group of particles. Each particle represents a legal state in the state space with a weight representing its possibility. Using the transition equations and new input data, these particles will be transferred to their next states. During the process, a portion of the particles will become illegal and will be immediately dropped, and a portion of the particles will change their weight. Finally, the algorithm converges to some individual particles representing the real transition path.

We make assumptions of the device's initial positions and orientations by defining the state space of the particle filter. During the transition of the particles, the computation relies on the input of IMU sensors. As we have mentioned above, the existence of inherent noises will make particles move in the wrong way. The arm motion model can be used here to detect illegal particles and filter them out, which means that our wrong assumptions are ruled out.

Arm motion model approaches can achieve higher tracking precisions compared with physical model approaches. However, the arm motion model also has limitations: First, the arm motion model assumes the user's torso keeps stationary, and only the arms, wrists, or hands move during the IMU tracking. Thus, the application scenarios are limited. Second, to achieve a high tracking precision, the particle filter should either enlarge the number of its initial particles or resample more particles after each transition. This process will result in a higher computational complexity. For example, MUSE[11] cannot be afforded on modern off-the-shelf smartphones due to its heavy computation.

In summary, the particle filter can effectively address the error accumulation of IMU sensor data and further introduces a trade-off between high accuracy and high computational complexity.

6 Activity Model Approach

In this section, we introduce the activity model based IMU tracking approaches, whose main idea is to derive the corresponding gestures instead of the accurate trajectory.

Theoretically, if IMU tracking's precision can be sufficiently high, then it would not be difficult to distinguish or recognize activities according to their traces. Due to the difficulty in high-precision tracking, activity model based approaches recognize related gestures instead of accurate positions.

The main steps for activity model based approaches

are as follows: First, segment the IMU sensor readings to determine which segments of the readings contain gestures. Second, extract features from the IMU sensor readings to represent certain gestures. Third, build a classifier to recognize gestures from our extracted features. These steps are further elaborated below.

6.1 Data segmentation

We need to perform data segmentation on the IMU sensor readings to distinguish the segments for different gestures, which are usually of two types: A segment containing gestures is denoted as active segment, and a segment without gestures is denoted as passive segment. Data segmentation needs to achieve two goals: (1) Distinguish active segments from passive segments. (2) If an active segment consists of many gestures, then we need to segment it, such that after segmentation, each part contains only one gesture.

Furthermore, it is important to achieve an accurate data segment, as this is the basis for gesture recognition.

Intuitively, threshold-based methods are proposed for data segmentation. Some works[20, 55–57] proposed acceleration- or energy-based segmentation. For example, Ref. [55] used the energy threshold based method for data segmentation. Here energy is defined as the summation of the square of a three-axis acceleration. Reference [55] set a window over data and determined whether the value of energy exceeds a predefined threshold.

Threshold-based methods may not work well sometimes, because it is not necessary that the user should stay static during passive segments. For example, when the user is performing gestures for a text input, there is also connecting movements between two letters. This kind of movement does not contain any direct information to recognize the letters and thus should be regarded as passive segments. However, the connecting movement can cause the energy to exceed the threshold, thus leading to an incorrect segment.

Different approaches have also been proposed to address the problem of movements between two gestures. SignSpeaker[58] requires users to move their hands back on their knees between two signs. By defining this special movement, SignSpeaker can recognize continuous gestures and make correct segmentation. RisQ[59] found that hand-to-mouth gesture is typical for a smoking person. The special information in between gestures can help data segmentation.

Furthermore, some research works have leveraged machine learning algorithms and combined gesture recognition and segmentation into one solution, such as Support Vector Machines (SVM)[60].

6.2 Feature extraction

We need to extract meaningful features from IMU sensor readings. The better these features can represent a gesture's trace, the better the recognition accuracy that the activity model will achieve. The categories of features can vary depending on the activities.

Raw sensor readings can be used as features[61]. Moreover, statistical values, such as mean, variance, correlation, or frequency-domain entropy of sensor readings, can be used as features[56, 60, 62, 63].

The use of activity models can bypass the difficulties met in actual 3D IMU tracking. By extracting useful features, activity recognition can be achieved without knowing the real traces of devices.

6.3 Gesture recognition

To recognize gestures from the features we extract, we need to build a classifier, which is usually achieved through machine learning approaches.

Machine learning based approaches are popular in gesture recognition, and different methods have been proposed, such as random forest[20, 64, 65], SVM[66], k-Nearest Neighbor (kNN)[55], and Hidden Markov Model (HMM)[60, 61]. Recent deep learning methods have also been leveraged, such as the Convolutional Neural Network (CNN)[20] and Long Short-Term Memory network (LSTM)[58].

6.4 Summary

Based on application scenarios, activity model based approaches can be classified into three groups: daily activities (e.g., smoking, bicycling, and sign language), text input (e.g., writing alphabet letters and Arabic numbers or typing on keyboards), and customized gestures. We summarize the related works into three tables: Table 2 for daily activities, Table 3 for text input, and Table 4 for customized gestures, where A, G, and M denote accelerometer, gyroscope, and magnetometer, respectively.

7 Challenges and Future Work

7.1 Measurement of the initial orientation

The initial orientation is very important for IMU tracking. When estimating the initial orientation, a popular method is to find anchors, including the gravity and magnetic

Table 2 Recent works on gesture-recognition applications via IMU tracking (daily activity).

Related work	Activity type	Sensor type	Recognition technique	Performance
[58]	American sign language, including 1900 word signs, 26 alphabet signs, and 9 digit signs	A+G	LSTM, Connectionist Temporal Classification (CTC)	99.2% detection ratio, 99.5% reliability of sign recognition, and 1.04% word error rate in continuous sentences
[59]	Smoking gestures in the wild	A+G+M	Random forest, conditional random field	95.7% accuracy of smoking gestures, smoking duration time detection error less than 1 minute
[63]	Daily activites including walking, running, bicycling, etc.	A	SVM	94.8% F-score of 7 activities on condition that pocket position is known
[67]	Table-tennis-related hand movements	A+G+M	HMM	90.5% precision and 95.3% recall of 6 strokes, median accuracy of 6.2 cm

Table 3 Recent works on gesture-recognition applications via IMU tracking (text input).

Related work	Activity type	Sensor type	Recognition technique	Performance
[20]	26 lowercase alphabet letters	A+G+M	Principle Component Analysis (PCA), CNN	91.6% online accuracy and 94.3% offline accuracy of 26 letters
[55]	Movement on Point-Of-Scale (POS) terminal and QWERTY keyboard	A, microphone	kNN, HMM	65% Personal Identification Numbers (PINs) accuracy in top 3 candidates, 54.8% English words accuracy in top 5 candidates
[60]	8000 English words	A+G	HMM, SVM, language model	11% word error rate in user-independent case
[61]	10 Arabic numbers	A	HMM, Dynamic Time Warping (DTW)	90.8% accuracy in writer-independent case
[64]	Handwriting lowercase letters	A+G	Random forest, dictionary filter	42.8% word recognition accuracy in top 5 items
[65]	Press and movement on QWERTY keyboard	A+G	Random forest	73.85% accuracy with an error bound of ± 2 neighboring keys
[68]	Press and movement on QWERTY keyboards	A+G	Could fitting, Bayesian inference	With one hand wearing smartwatch, it is possible to shortlist a median of 24 words which contains the words that user has typed
[69]	6 basic strokes and 26 English letters	A	Correlation, spelling correction	91.9% accuracy if users conform to a few constraints
[70]	Movement over Swype and Cirrin keyboard	A+G+M	Directly tracking trajectory	Residual error rate of 1.6%, 10 words per minute
[71]	26 capital alphabet letters	A+G, microphone	DTW	Recognition rate of 94% in the first guess, 99% using 3 guesses
[72]	Hand-writing strokes	A+G	Directly tracking trajectory	Character error rate of 18%, $15-20$ words per minute

Table 4 Recent works on gesture-recognition applications via IMU tracking (customized gesture).

Related work	Activity type	Sensor type	Recognition technique	Performance
[56]	Customized gestures of arm, hand, and finger	A+G	Naive Bayes, logistic regression, decision tree	98% accuracy of 37 gestures
[62]	4 direction gestures, 3 shape gestures, and 5 alphabet letters	A	Frame-based descriptor and multi-class SVM	89.29% accuracy of 12 gestures in user-independent case
[66]	Tapping and sliding gestures on smartwatch	A+G	SVM	$88.7\,\% - 99.4\,\%$ accuracy of 3 gestures among different users

north direction[11, 16]. In the GRF, the gravity's direction and magnetic north direction are always orthogonal to each other. Thus, in the LRF, they are also orthogonal. Through a cross product of the two directions, we can obtain the third axis. The three axes altogether form a rotation matrix, which determines the orientation.

Gravity can be measured via an accelerometer when there is only gravity applied on this sensor and there are no other external forces, e.g., when the device is stationary. Thus, we assume that the device is static before the tracking begins. However, during tracking, it is challenging to determine whether the accelerometer only contains gravitational acceleration. Reference [16] pointed out that in the long run of orientation estimation (within 10 min), the opportunity that the device keeps static is rare in practice.

Moreover, a magnetometer is often used to determine the magnetic north direction in motion tracking. However, when there is only a geomagnetic field applied to this sensor, the magnetic north is measured rather than the north direction in the GRF. The magnetic north may vary in different locations on the earth. Particularly, it points to the sky in the southern hemisphere and points to the ground in the northern hemisphere. Thus, the magnetometer readings are usually combined with gravity readings from the accelerometer to calculate the north direction in the GRF. Due to magnetic interference, the magnetometer's performance is impacted by different factors, such as ferromagnetic materials.

7.2 Unbounded measurement

In Section 2, we showed that the error of tracking could be unbounded. We lack anchors for calibration during the tracking period, and the accumulated error can be very large.

All the three models presented in Fig. 2 encounter this challenge. In the physical model, the location error accumulates with time after integrating acceleration twice, and it can be very large after a time period. The arm motion model places constraints on the wrist's movement and can bound tracking locations in a particular area. However, it cannot tackle the problem of accumulated orientation error. Although the activity models do not track a device movement's real trace, the features extracted from IMU sensor readings can be severely affected by these unbounded errors. If these features cannot represent the gesture's trace well, then the gesture recognition part may suffer from degradation.

7.3 Data segmentation

For gesture recognition, there exists a challenge that the user's movement is unpredictable during passive segments (defined in Section 6.1). For example, when a user is writing in the air, the user's hands need to move from the ending point of a letter to the beginning point of the next one, leading to passive segments in data segmentation. The passive segments significantly degrade the recognition performance.

Some machine learning based methods can take the whole data sequences as input without segmentation. However, without knowing how many segments (e.g., letters) are in the data, the training process can be very wardriving for these methods.

8 Conclusion

Motion tracking via IMUs on mobile and wearable devices has attracted significant interest in recent years from both industry and academia. It can be used in different applications. In this work, we study IMU-tracking techniques on mobile and wearable devices. We show the main categories for the existing IMU-tracking techniques, and introduce representative works in each category. Finally, we reveal the main challenges and future directions for IMU-tracking on mobile and wearable devices.

Acknowledgment

This work was in part supported by the National Key R&D Program of China (No. 2018YFB1004800) and the National Natural Science Foundation of China (No. 61932013).

References

[1] A. Altenbuchner, S. Haug, R. Kretschmer, and K. Weber, How to measure physical motion and the impact of individualized feedback in the field of rehabilitation of geriatric trauma patients, *Stud. Health Technol. Inform.*, vol. 248, pp. 226–232, 2018.

[2] F. Alemuda and F. J. Lin, Gesture-based control in a smart home environment, in *Proc. IEEE Int. Conf. Internet of Things* (*iThings*) *and IEEE Green Computing and Communications* (*GreenCom*) *and IEEE Cyber, Physical and Social Computing* (*CPSCom*) *and IEEE Smart Data* (*SmartData*), Exeter, UK, 2017, pp. 784–791.

[3] Y. Shen, B. Du, W. Xu, C. Luo, B. Wei, L. Cui, and H. Wen, Securing cyber-physical social interactions on wrist-worn devices, *ACM Transactions on Sensor Networks*, vol. 16, no. 2, pp. 1–22, 2020.

[4] Z. M. Zhou, C. S. Wu, Z. Yang, and Y. H. Liu, Sensorless sensing with WiFi, *Tsinghua Science and Technology*, vol.

20, no. 1, pp. 1–6, 2015.

[5] Q. Shi, S. H. Zhao, X. W. Cui, M. Q. Lu, and M. D. Jia, Anchor self-localization algorithm based on UWB ranging and inertial measurements, *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 728–737, 2019.

[6] S. Jiang, Z. Li, and P. Zhou. Memento: An emotion-driven lifelogging system with wearables, *ACM Transactions on Sensor Networks*, vol. 15, no. 1, pp. 1–23, 2019.

[7] Z. Zhang, X. Cong, W. Feng, H. P. Zhang, G. D. Fu, and J. Y. Chen, WAEAS: An optimization scheme of EAS scheduler for wearable applications, *Tsinghua Science and Technology*, vol. 26, no. 1, pp. 72–84, 2021.

[8] M. Kok, J. D. Hol, and T. B. Schön, Using inertial sensors for position and orientation estimation, arXiv preprint arXiv: 1704.06053, 2017.

[9] Z. Yang, C. S. Wu, Z. M. Zhou, X. L. Zhang, X. Wang, and Y. H. Liu, Mobility increases localizability: A survey on wireless indoor localization using inertial sensors, *ACM Comput. Surv.*, vol. 47, no. 3, p. 54, 2015.

[10] H. Liu, R. Li, S. C. Liu, S. B. Tian, and J. Z. Du, Smartcare: Energy-efficient long-term physical activity tracking using smartphones, *Tsinghua Science and Technology*, vol. 20, no. 4, pp. 348–363, 2015.

[11] S. Shen, M. Gowda, and R. R. Choudhury, Closing the gaps in inertial motion tracking, in *Proc. 24th Annu. Int. Conf. Mobile Computing and Networking* (*MobiCom*), New Delhi, India, 2018, pp. 429–444.

[12] O. J. Woodman, An introduction to inertial navigation, Technical Report UCAM-CL-TR-696, Computer Laboratory, University of Cambridge, https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-696.html，2007.

[13] G. Laput, R. Xiao, and C. Harrison, ViBand: High-fidelity bio-acoustic sensing using commodity smartwatch accelerometers, in *Proc. 29th Annu. Symp. User Interface Software and Technology* (*UIST*), New York, NY, USA, 2016, pp. 321–333.

[14] V. T. Van Hees, L. Gorzelniak, E. C. D. León, M. Eder, M. Pias, S. Taherian, U. Ekelund, F. Renström, P. W. Franks, A. Horsch, et al., Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity, *PLoS One*, vol. 8, no. 4, p. e61691, 2013.

[15] Android developer, Motion sensors, https://developer.android.com/guide/topics/sensors/sensors˙motion, 2021.

[16] P. F. Zhou, M. Li, and G. B. Shen, Use it free: Instantly knowing your phone attitude, in *Proc. 20th Annu. Int. Conf. Mobile Computing and Networking* (*MobiCom*), Maui, HI, USA, 2014, pp. 605–616.

[17] S. Poddar, V. Kumar, and A. Kumar, A comprehensive overview of inertial sensor calibration techniques, *J. Dyn. Syst. Meas. Control*, vol. 139, no. 1, p. 011006, 2016.

[18] P. D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*. 2nd ed. London, UK: Artech House, 2013.

[19] M. J. Caruso, Applications of magnetic sensors for low cost compass systems, in *Proc. IEEE 2000 Position Location and Navigation Symp.*, San Diego, CA, USA, 2000, pp. 177–184.

[20] Y. F. Yin, L. Xie, T. Gu, Y. J. Lu, and S. L. Lu, Aircontour: Building contour-based model for in-air writing gesture recognition, *ACM Trans. Sensor Net.*, vol. 15, no. 4, p. 44, 2019.

[21] P. Batista, C. Silvestre, P. Oliveira, and B. Cardeira, Accelerometer calibration and dynamic bias and gravity estimation: Analysis, design, and experimental evaluation, *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 5, pp. 1128–1137, 2011.

[22] E. H. Shin and E. S. Naser, Accuracy improvement of low cost INS/GPS for land applications, in *Proc. 2002 National Technical Meeting of the Institute of Navigation*, San Diego, CA, USA, 2002, pp. 146–157.

[23] E. H. Shin and N. El-Sheimy, A new calibration method for strapdown inertial navigation systems, *Zfv*, vol. 127, pp. 41–50, 2002.

[24] X. J. Niu, Y. Li, H. P. Zhang, Q. J. Wang, and Y. L. Ban, Fast thermal calibration of low-grade inertial sensors and inertial measurement units, *Sensors*, vol. 13, no. 9, pp. 12192–12217, 2013.

[25] J. L. Crassidis, F. L. Markley, and Y. Cheng, Survey of nonlinear attitude estimation methods, *J. Guid. Control Dyn.*, vol. 30, no. 1, pp. 12–28, 2007.

[26] N. H. Q. Phuong, H. J. Kang, Y. S. Suh, and Y. S. Ro, A DCM based orientation estimation algorithm with an inertial measurement unit and a magnetic compass, *J. Univ. Comp. Sci.*, vol. 15, no. 4, pp. 859–876, 2009.

[27] E. Edwan, J. Y. Zhang, J. C. Zhou, and O. Loffeld, Reduced DCM based attitude estimation using low-cost IMU and magnetometer triad, in *Proc. 2011 8th Workshop on Positioning, Navigation and Communication*, Dresden, Germany, 2011, pp. 1–6.

[28] S. Yean, B. S. Lee, C. K. Yeo, and C. H. Vun, Algorithm for 3D orientation estimation based on Kalman filter and gradient descent, in *Proc. 2016 IEEE 7th Annu. Information Technology, Electronics and Mobile Communication Conf.* (*IEMCON*), Vancouver, Canada, 2016, pp. 1–6.

[29] L. Lou, X. Xu, J. Cao, Z. L. Chen, and Y. Xu, Sensor fusion-based attitude estimation using low-cost MEMS-IMU for mobile robot navigation, in *Proc. 2011 6th IEEE Joint Int. Information Technology and Artificial Intelligence Conf.*, Chongqing, China, 2011, pp. 465–468.

[30] W. Li and J. L. Wang, Effective adaptive Kalman filter for MEMS-IMU/magnetometers integrated attitude and heading reference systems, *J. Navigat.*, vol. 66, no. 1, pp. 99–113, 2013.

[31] R. Zhu, D. Sun, Z. Y. Zhou, and D. Q. Wang, A linear fusion algorithm for attitude determination using low cost mems-based sensors, *Measurement*, vol. 40, no. 3, pp. 322–328, 2007.

[32] D. Jurman, M. Jankovec, R. Kamnik, and M. Topic, Calibration and data fusion solution for the miniature attitude and heading reference system, *Sens. Actuators A: Phys.*, vol. 138, no. 2, pp. 411–420, 2007.

[33] Z. Ercan, V. Sezer, H. Heceoglu, C. Dikilitas, M. Gokasan, A. Mugan, and S. Bogosyan, Multi-sensor data fusion of DCM based orientation estimation for land vehicles, in *Proc. 2011 IEEE Int. Conf. Mechatronics*, Istanbul, Turkey, 2011, pp. 672–677.

[34] N. Shantha and T. Jann, Estimation of attitudes from a low-cost miniaturized inertial platform using kalman filter-based sensor fusion algorithm, *Sadhana*, vol. 29, no. 2, pp. 217–235, 2004.

[35] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi, Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration, *Automatica*, vol. 51, pp. 158–166, 2015.

[36] F. Olsson, M. Kok, K. Halvorsen, and T. B. Schön, Accelerometer calibration using sensor fusion with a gyroscope, in *Proc. 2016 IEEE Statistical Signal Proc. Workshop* (*SSP*), Palma de Mallorca, Spain, 2016, pp. 1–5.

[37] E. Foxlin, Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter, in *Proc. IEEE 1996 Virtual Reality Annu. Int. Symp.*, Santa Clara, CA, USA, 1996, pp. 185–194.

[38] R. Munguia and A. Grau, Attitude and heading system based on EKF total state configuration, in *Proc. 2011 IEEE Int. Symp. Industrial Electronics*, Gdansk, Poland, 2011, pp. 2147–2152.

[39] M. Romanovas, L. Klingbeil, M. Trachtler, and Y. Manoli, Efficient orientation estimation algorithm for low cost inertial and magnetic sensor systems, in *Proc. 2009 IEEE/SP 15th Workshop on Statistical Signal Proc.*, Cardiff, UK, 2009, pp. 586–589.

[40] H. G. de Marina, F. J. Pereda, J. M. Giron-Sierra, and F. Espinosa, UAV attitude estimation using unscented Kalman filter and TRIAD, *IEEE Trans. Indust. Electron.*, vol. 59, no. 11, pp. 4465–4474, 2012.

[41] J. L. Marins, X. P. Yun, E. R. Bachmann, R. B. McGhee, and M. J. Zyda, An extended Kalman filter for quaternion-based orientation estimation using MARG sensors, in *Proc. 2001 IEEE/RSJ Int. Conf. Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the Next Millennium* (*Cat. No. 01CH37180*), Maui, HI, USA, 2002, pp. 2003–2011.

[42] A. M. Sabatini, Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing, *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1346–1356, 2006.

[43] J. Goslinski, M. Nowicki, and P. Skrzypczynski, Performance comparison of EKF-based algorithms for orientation estimation on android platform, *IEEE Sens. J.*, vol. 15, no. 7, pp. 3781–3792, 2015.

[44] B. Huyghe, J. Doutreloigne, and J. Vanfleteren, 3D orientation tracking based on unscented Kalman filtering of accelerometer and magnetometer data, in *Proc. 2009 IEEE Sensors Applications Symp*, New Orleans, LA, USA, 2009, pp. 148–152.

[45] R. Mahony, T. Hamel, and J. M. Pflimlin, Nonlinear complementary filters on the special orthogonal group, *IEEE Trans. Automat. Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[46] M. Nowicki, J. Wietrzykowski, and P. Skrzypczynski, Simplicity or flexibility? complementary filter vs. EKF for orientation estimation on mobile devices, in *Proc. 2015 IEEE 2nd Int. Conf. Cybernetics* (*CYBCONF*), Gdynia, Poland, 2015, pp. 166–171.

[47] S. O. H. Madgwick, *An Efficient Orientation Filter for Inertial and Inertial/Magnetic Sensor Arrays*. UK: University of Bristol, 2010.

[48] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, Estimation of IMU and MARG orientation using a gradient descent algorithm, in *Proc. 2011 IEEE Int. Conf. Rehabilitation Robotics*, Zurich, Switzerland, 2011, pp. 1–7.

[49] M. El-Gohary, S. Pearson, and J. McNames, Joint angle tracking with inertial sensors, *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 208, pp. 1068–1071, 2008.

[50] M. El-Gohary and J. McNames, Shoulder and elbow joint angle tracking with inertial sensors, *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2635–2641, 2012.

[51] S. Shen, H. Wang, and R. R. Choudhury, I am a smartwatch and I can track my user's arm, in *Proc. 14th Annu. Int. Conf. Mobile Systems, Applications, and Services, MobiSys'16*, Singapore, 2016, pp. 85–96.

[52] Y. Liu, C. D. Lin, Z. J. Li, Z. D. Liu, and K. S. Wu, When wearable sensing meets arm tracking (poster), in *Proc. 17th Annu. Int. Conf. Mobile Systems, Applications, and Services, MobiSys'19*, Seoul, Republic of Korea, 2019, pp. 518–519.

[53] B. Fang, X. Wei, F. C. Sun, H. M. Huang, Y. L. Yu, and H. P. Liu, Skill learning for human-robot interaction using wearable device, *Tsinghua Sci. Technol.*, vol. 24, no. 6, pp. 654–662, 2019.

[54] P. del Moral, Nonlinear filtering: Interacting particle solution, *Markov Process. Relat. Fields*, vol. 2, no. 4, pp. 555–580, 1996.

[55] X. Y. Liu, Z. Zhou, W. R. Diao, Z. Li, and K. H. Zhang, When good becomes evil: Keystroke inference with smartwatch, in *Proc. 22nd ACM SIGSAC Conf. Computer and Communications Security*, Denver, CO, USA, 2015, pp. 1273–1285.

[56] C. Xu, P. H. Pathak, and P. Mohapatra, Finger-writing with smartwatch: A case for finger and hand gesture recognition using smartwatch, in *Proc.16th Int. Workshop on Mobile Computing Systems and Applications*, Santa Fe, NM, USA, 2015, pp. 9–14.

[57] Z. J. Ba, T. H. Zheng, X. Y. Zhang, Z. Qin, B. C. Li, X. Liu, and K. Ren, Learning-based practical smartphone eavesdropping with built-in accelerometer, in *Proc. Network and Distributed Systems Security* (*NDSS*) *Symp.*, San Diego, CA, USA, 2020, pp. 23–26.

[58] J. H. Hou, X. Y. Li, P. D. Zhu, Z. F. Wang, Y. Wang, J. W. Qian, and P. L. Yang, SignSpeaker: A real-time, high-precision smartwatch-based sign language translator, in *Proc. 25th Annu. Int. Conf. Mobile Computing and Networking, MobiCom'19*, Los Cabos, Mexico, 2019, pp. 24:1–24:15.

[59] A. Parate, M. C. Chiu, C. Chadowitz, D. Ganesan, and E. Kalogerakis, RisQ: Recognizing smoking gestures with inertial sensors on a wristband, in *Proc. 12th Annu. Int. Conf. Mobile Systems, Applications, and Services, MobiSys'14*, Bretton Woods, NH, USA, 2014, pp. 149–161.

[60] C. Amma, M. Georgi, and T. Schultz, Airwriting: A wearable handwriting recognition system, *Pers. Ubiquit. Comput.*, vol. 18, no. 1, pp. 191–203, 2014.

[61] S. D. Choi, A. S. Lee, and S. Y. Lee, On-line handwritten

character recognition with 3D accelerometer, in *Proc. 2006 IEEE Int. Conf. Information Acquisition*, Veihai, China, 2006, pp. 845–850.

[62] J. H. Wu, G. Pan, D. Q. Zhang, G. D. Qi, and S. J. Li, Gesture recognition with a 3-D accelerometer, in *Proc. 6^{th} Int. Conf. Ubiquitous Intelligence and Computing* (*UIC*) *2009*, Heidelberg, Germany, 2009, pp. 25–38.

[63] L. Sun, D. Q. Zhang, B. Li, B. Guo, and S. J. Li, Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations, in *Proc. 7^{th} Int. Conf. Ubiquitous Intelligence and Computing* (*UIC*), Berlin, Germany, 2010, pp. 548–562.

[64] Q. X. Xia, F. Hong, Y. Feng, and Z. W. Guo, Motionhacker: Motion sensor-based eavesdropping on handwriting via smartwatch, in *Proc. IEEE INFOCOM 2018-IEEE Conf. Computer Communications Workshops* (*INFOCOM WKSHPS*), Honolulu, HI, USA, 2018, pp. 468–473.

[65] S. Sen, K. Grover, V. Subbaraju, and A. Misra, Inferring smartphone keypress via smartwatch inertial sensing, in *Proc. 2017 IEEE Int. Conf. Pervasive Computing and Communications Workshops* (*PerCom Workshops*), Kona, HI, USA, 2017, pp. 685–690.

[66] C. Zhang, J. R.Yang, C. Southern, T. E. Starner, and G. D. Abowd, Watchout: Extending interactions on a smartwatch with inertial sensing, in *Proc. ACM Int. Symp. Wearable Computers*, Kobe, Japan, 2016, pp. 136–143.

[67] T. H. Vu, A. Misra, Q. Roy, K. C. T. Wei, and Y. Lee, Smartwatch-based early gesture detection 8 trajectory tracking for interactive gesture-driven applications, *Proc. ACM Interact. Mobile Wearab. Ubiquit. Technol.*, vol. 2, no. 1, p. 39, 2018.

[68] H. Wang, T. T. T. Lai, and R. R. Choudhury, MoLe: Motion leaks through smartwatch sensors, in *Proc. 21^{st} Annu. Int. Conf. Mobile Computing and Networking*, *MobiCom'15*, Paris, France, 2015, pp. 155–166.

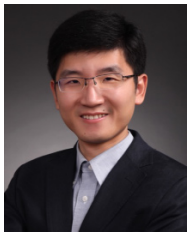
[69] S. Agrawal, I. Constandache, S. Gaonkar, R. R. Choudhury, K. Caves, and F. DeRuyter, Using mobile phones to write in air, in *Proc. 9^{th} Int. Conf. Mobile Systems*, *Applications, and Services*, *MobiSys'11*, Bethesda, MD, USA, 2011, pp. 15–28.

[70] K. Katsuragawa, J. R. Wallace, and E. Lank, Gestural text input using a smartwatch, in *Proc. Working Conf. Advanced Visual Interfaces*, Bari, Italy, 2016, pp. 220–223.

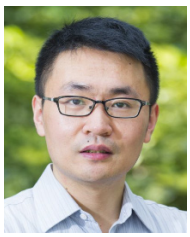
[71] L. Ardüser, P. Bissig, P. Brandes, and R. Wattenhofer, Recognizing text using motion data from a smartwatch, in *Proc. 2016 IEEE Int. Conf. Pervasive Computing and Communication Workshops* (*PerCom Workshops*), Sydney, Australia, 2016, pp. 1–6.

[72] T. Deselaers, D. Keysers, J. Hosang, and H. A. Rowley, Gyropen: Gyroscopes for pen-input with mobile phones, *IEEE Trans. Human Mach. Syst.*, vol. 45, no. 2, pp. 263–271, 2015.

Zhipeng Song received the BEng degree from Tsinghua University, China in 2018, where he is currently a PhD candidate. His research interests include low power wireless communication, wireless sensing, and IoTs applications.

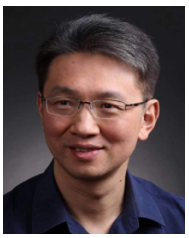


Zhichao Cao received the BEng degree from Tsinghua University, and the PhD degree from Hong Kong University of Science and Technology, China in 2013. He is currently an assistant professor at the Department of Computer Science and Engineering, Michigan State University, USA. His research interests lie broadly in IoT systems, edge computing, and mobile computing.



Zhenjiang Li received the BEng degree from Xi'an Jiaotong University, China in 2001, and the M. Phil and PhD degrees from Hong Kong University of Science and Technology, China in 2009 and 2012, respectively. He is an assistant professor at the Department of Computer Science, City University of Hong Kong. His research interests include distributed and mobile sensing computing and learning system.

Jiliang Wang received the BEng from University of Science and Technology of China in 2007, and the PhD degree from Hong Kong University of Science and Technology, China in 2011. He is currently an associate professor at School of Software, Tsinghua University. His research interests include IoTs, mobile networks, and wireless and sensor networks.



Yunhao Liu received the BEng degree from Tsinghua University, China in 1995, and the MA degree from Beijing Foreign Studies University, China in 1997. He received the MEng and PhD degrees in computer science and engineering from Michigan State University, USA in 2003 and 2004, respectively. He is now a professor at the Department of Automation and the dean of the GIX, Tsinghua University, China. He is a fellow of IEEE and ACM. His research interests include IoTs and wireless sensor network; indoor localization and network diagnosis; RFID; supply chain and industrial internet; and distributed systems and cloud computing.